UNIVERSIDAD NACIONAL SANTIAGO ANTUNEZ DE MAYOLO

FACULTAD DE CIENCIAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

"DATA MART COMO PARTE DE UNA SOLUCION DE INTELIGENCIA DE NEGOCIOS, PARA EL SOPORTE DE LA TOMA DE DECISIONES DE LA GESTIÓN ACADÉMICA DE LA FACULTAD DE CIENCIAS EN LA UNASAM"

TESIS

PARA OPTAR EL TÍTULO DE

INGENIERO DE SISTEMAS E INFORMÁTICA

AUTOR

BACH. DURAN COLONIA EVELYN ANABEL

ASESOR

ING. FLORES CHACÓN ERICK GIOVANNY

HUARAZ – PERU

2017

Nº Registro. T042

A la memoria Estanislao Durán quien fue la fuente de inspiración, ejemplo de sabiduría, humildad y perseverancia.

A mis padres Angel y Aurora que me dedicaron un tiempo especial.

Evelyn.

AGRADECIMIENTOS

A Dios por ser fuente de vida, y ser el soporte de todas las acciones realizadas durante mi formación profesional.

Así mismo a la Facultad de Ciencias de la Universidad Nacional Santiago Antúnez de Mayolo, en la persona de los Jefes de Departamento y Directores de Escuela por el apoyo brindado para la realización de la presente investigación.

A la Escuela Profesional de Ingeniería de Sistemas e Informática de la Facultad de Ciencia de la Universidad Nacional Santiago Antúnez de Mayolo, en la persona de los docentes por su profesionalismo y su dedicación en la formación profesional.

Al asesor Ing. Erick Flores Chacón, por sus acertadas orientaciones, el soporte profesional oportuno y el tiempo dedicado para el desarrollo de la presente Investigación.

PRESENTACIÓN

Señores Miembros del Jurado Calificador:

En cumplimiento con el reglamento de grados y títulos de la Escuela Académico- Profesional de Ingeniería de sistemas e Informática, Facultad de Ciencias, de la Universidad Nacional Santiago Antúnez de Mayolo, me permito presentar la tesis titulada "Data Mart como parte de una solución de inteligencia de negocios, para el soporte de la toma de decisiones de la gestión académica de la facultad de ciencias en la UNASAM", que es el producto de la inquietud por determinar como la aplicación de un Data Mart, utilizando inteligencia de Negocios, influye en la toma de decisiones de la gestión académica de la Facultad de Ciencias de la Universidad Nacional Santiago Antúnez de Mayolo.

HOJA DE VISTO BUENO

Ing. NARRO CACHAY CESAR AUGUSTO

PRESIDENTE

Reg. CIP Nº 169491

Ing. MEDINA VILLACORTA ALBERTO MARTIN

SECRETARIO

Reg. CIP Nº 143211

Ing. FLORES CHACON ERICK GIOVANNY

VOCAL

Reg. CIP Nº 89540

RESUMEN

La presente investigación pretende implementar un Data Mart, como parte de una solución de inteligencia de negocio, para mejorar el proceso de Toma de Decisiones en la Gestión Académica de Facultad de Ciencias en la UNASAM.

El estudio responde a una metodología experimental, diseño cuasi-experimental, y el método general empleado fue la experimental, la muestra de estudio estuvo constituida por los tomadores de decisiones de la Facultad de Ciencias, siendo el grupo experimental la Escuela de Ingeniería de Sistemas e Informática y el grupo control conformada por la Escuela de Estadística e Informática y la escuela de Matemática. para la contratación de la hipótesis se utilizó la prueba de hipótesis de t-student.

Cuya conclusión arribada fue, la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora significativamente el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM. Según los resultados el 66.7% de los tomadores de decisiones se encuentra en el nivel muy bueno y el 33.3% de los tomadores de decisiones en el nivel bueno. El p=0,000 siendo altamente significativo; por lo que queda demostrado que el Data Mart, como parte de una solución de inteligencia de negocio, se mejora significativamente el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM (ver Cuadro Nº 55, Gráfico Nº 1 y gráfico Nº 2).

Palabras Claves: Data Mart – Inteligencia de Negocios – Toma de decisiones, Metodología DWEP, Gestión Académica.

ABSTRACT

The present research intends to implement a Data Mart, as part of a business intelligence solution, to improve the process of Decision Making in the Academic Management of Faculty of Sciences in "Santiago Antunez de Mayolo" University.

This research has an experimental methodology, quasi-experimental design, and the general method employed was experimental method, the study sample was constituted by the decision makers of the Faculty of Sciences, the experimental group being the School of Systems Engineering and Computer Science And the control group formed by the School of Statistics and Informatics and the school of Mathematics. The t-student hypothesis test is used to contract the hypothesis.

Whose conclusion was reached, the implementation of a Data Mart, as part of a business intelligence solution, is the process of growth of Decision Making in Academic Management of the Faculty of Sciences of UNASAM. According to the results, 66.7% of the decision-makers are at the very good level and 33.3% of the decision-makers at the good level. P = 0.000 being highly significant; Therefore, it is demonstrated that the Data Mart, as part of a business intelligence solution, is the process of growth of Decision Making in the Academic Management of the Faculty of Sciences of UNASAM (see Table Nº 55, figure Nº 1 and Graphic Nº 2).

Key words: Data Mart – Business Intelligence – Decision Making, DWEP methodology, Academic management.

SUMARIO

AGRADECIMIENTOS ii
PRESENTACIÓN iii
HOJA DE VISTO BUENO iv
RESUMEN v
ABSTRACT vi
SUMARIO vii
SUMARIO DE CUADROS xii
SUMARIO DE FIGURAS xv
SUMARIO DE GRÁFICOS xvii
CAPÍTULO I 18
GENERALIDADES 18
1.1. REALIDAD PROBLEMÁTICA 18
1.2. ENUNCIADO DEL PROBLEMA 21
1.2.1. GENERAL 21
1.2.2. ESPECÍFICOS 21
1.3. HIPÓTESIS 21
1.3.1. GENERAL 21
1.3.2. ESPECÍFICO 22
1.4. OBJETIVOS 23
1.4.1. GENERAL 23
1.4.2. ESPECÍFICOS 23
1.5. JUSTIFICACIÓN 23
1.6. LIMITACIONES 25
1.7. DESCRIPCIÓN Y SUSTENTACIÓN DE LA SOLUCIÓN 26
1.7.1. DESCRIPCIÓN DE LA SOLUCIÓN 26
1.7.1.1. FUNCIONALIDAD 26
1.7.1.2. MODELO MULTIDIMENSIONAL 26
1.7.1.3. PROCESO ETL 27
1.7.1.4. HERRAMIENTA DE EXPLOTACIÓN 28

1.7.2. SUSTENTACIÓN DE LA SOLUCIÓN 28
1.7.2.1. SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS 29
1.7.2.2. CREACIÓN DE DATA MART 29
CAPÍTULO II 31
MARCO TEÓRICO 31
2.1. ANTECEDENTES 31
2.2. INTELIGENCIA DE NEGOCIOS 34
2.2.1. ARQUITECTURAS DE INTELIGENCIAS DE NEGOCIOS 35
2.2.2. COMPONENTES DE UN SISTEMA DE INTELIGENCIA DE NEGOCIO 35
2.3. DATA WAREHOUSE 39
2.4. DATA MART 44
2.5. ARQUITECTURA DATA MART 45
2.6. MODELO DIMENSIONAL 45
2.6.1. MODELO MULTIDIMENSIONAL CONCEPTUAL 46
2.6.2. MODELO MULTIDIMENSIONAL LÓGICO 47
2.7. HERRAMIENTAS ETL 50
2.8. INGENIERÍA DE SOFTWARE 51
2.8.1. PROCESO DE INGENIERÍA DE DATA WAREHOUSE 52
2.8.1.1. FASES 53
2.8.1.2. DISCIPLINAS 54
2.8.1.3. ENTREGABLES 56
2.8.1.4. ARQUITECTURA DE UN DATA WAREHOUSE 59
2.8.1.5. ESQUEMAS 60
2.9. TOMA DE DECISIONES 62
2.9.1. EL PROCESO DE TOMA DE DECISIONES 64
2.9.2. SISTEMAS DE SOPORTE A LA TOMA DE DECISIONES 66
2.9.3. HERRAMIENTAS DE TOMA DE DECISIONES 68
2.9.4. GESTIÓN ACADÉMICA 68
2.10. DEFINICIÓN DE TÉRMINOS 70
CAPÍTULO III 74

MATERIALES Y MÉTODOS 74
3.1. TIPO DE ESTUDIO 74
3.2. DISEÑO DE ESTUDIO 75
3.3. POBLACIÓN Y MUESTRA 76
3.3.1. POBLACIÓN 76
3.3.2. MUESTRA 76
3.3.3. MUESTREO 76
3.4. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS 77
3.5. TÉCNICAS DE PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN 78
3.6. MATERIALES 78
3.7. MÉTODOS 79
3.7.1. MÉTODO GENERAL 79
3.7.2. MÉTODOS ESPECÍFICOS: 79
3.8. OPERACIONALIZACIÓN DE VARIABLES 80
3.9. PROCEDIMIENTO 81
3.9.1. IDENTIFICACIÓN DE REQUERIMIENTOS 81
3.9.2. DISEÑO DEL DATA MART 82
3.9.3. COMPROMETER A LOS STAKEHOLDER 82
CAPÍTULO IV 83
ANÁLISIS 83
4.1. ANÁLISIS DE LA SITUACIÓN ACTUAL 83
4.1.1. ANÁLISIS DEL ORGANIGRAMA FUNCIONAL ESTRATÉGICO 83
4.1.2. EVALUACIÓN DE LA CAPACIDAD INSTALADA 85
4.1.3. ANÁLISIS DE LAS FORTALEZAS, OPORTUNIDADES, DEBILIDADES Y AMENAZAS 87
4.1.4. ANÁLISIS DE LA SITUACIÓN ACTUAL 89
4.2. IDENTIFICACIÓN Y DESCRIPCIÓN DE REQUERIMIENTOS 91
4.2.1. REQUERIMIENTOS FUNCIONALES 91
4.2.2. REQUERIMIENTOS NO FUNCIONALES 94
4.3. ANÁLISIS DE REQUERIMIENTOS 95

4.3.1. DIAGRAMA DE CONTEXTO 95
4.4. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL 96
CAPÍTULO V 98
DISEÑO DE LA SOLUCIÓN 98
5.1. ARQUITECTURA TECNOLÓGICA DE LA SOLUCIÓN 98
5.1.1. FUENTES DE INFORMACIÓN 99
5.1.2. ETL 100
5.1.3. DATA MART 102
5.1.4. OLAP 103
5.1.5. INTERFAZ DE USUARIO 103
5.2. ANÁLISIS DIMENSIONAL 104
5.2.1. DISEÑO CONCEPTUAL DEL DATA MART 104
5.2.1.1. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 1 104
5.2.1.2. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 2 105
5.2.1.3. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 3 107
5.3. DISEÑO DE EXTRACCIÓN 111
5.4. DISEÑO DE EXPLOTACIÓN 113
CAPÍTULO VI 116
CONSTRUCCIÓN DE LA SOLUCIÓN 116
6.1. CONSTRUCCIÓN 116
6.1.1. HERRAMIENTAS Y SOLUCIONES 117
6.1.1.1. BASE DE DATOS 117
6.1.1.2. FRAMEWORK 117
6.1.1.3. LENGUAJE DE PROGRAMACIÓN 118
6.1.2. ESPECIFICACIÓN DE CONSTRUCCIÓN 118
6.1.3. PROCEDIMIENTOS DE OPERACIÓN Y ADMINISTRACIÓN 119
6.1.4. PROCEDIMIENTOS DE SEGURIDAD Y CONTROL DE ACCESO 120
6.1.5. PROCEDIMIENTOS DE OPERACIÓN Y MANUALES DEL USUARIO. 122

6.1.6. CONSTRUCCIÓN DE LOS COMPONENTES Y PROCEDIMIENTOS DE MIGRACIÓN Y CARGA INICIAL DE DATOS 122
6.2. PRUEBAS 123
6.2.1. CASO DE PRUEBA PARA ETL ESTUDIANTE 125
6.2.2. CASO DE PRUEBA PARA ETL CURRICULA 128
6.2.3. CASO DE PRUEBA PARA ETL DOCENTE 131
6.2.4. CASO DE PRUEBA PARA ETL MODALIDAD_APROBACION . 135
6.2.5. CASO DE PRUEBA PARA ETL SEMESTRE 137
6.2.6. CASO DE PRUEBA PARA EL HECHO COMPORTAMIENTO ACADÉMICO ESTUDIANTE 139
6.2.7. CASO DE PRUEBA PARA EL HECHO COMPORTAMIENTO ACADÉMICO DOCENTE 141
CAPÍTULO VII 144
IMPLEMENTACIÓN 144
7.1. MONITOREO Y EVALUACIÓN DE LA SOLUCIÓN. 144
7.1.1. ELEMENTOS DEL MONITOREO Y EVALUACIÓN. 145
7.1.2. POLÍTICAS Y REGLAS DE PROCEDIMIENTO 145
7.1.3. PLAN DE MONITOREO Y EVALUACIÓN. 146
7.2. BITÁCORA Y PUESTA A PUNTO. 147
7.2.1. MIGRACIÓN Y CARGA INICIAL DE DATOS. 147
7.2.2. APROBACIÓN DE LA SOLUCIÓN TECNOLÓGICA 148
CAPÍTULO VIII 149
RESULTADOS 149
CAPÍTULO IX 156
DISCUSIÓN DE RESULTADOS 156
CONCLUSIONES 160
RECOMENDACIONES 162
REFERENCIAS BIBLIOGRÁFICAS 163

SUMARIO DE CUADROS

CUADRO Nº 01. DIFERENCIAS ENTRE OLTP Y OLAP 42
CUADRO Nº 02.ENTREGABLES POR DISCIPLINAS DEL DWEP 57
CUADRO Nº 03. DISEÑO DE UN DATA WAREHOUSE.....60
CUADRO Nº 04. POBLACIÓN JERÁRQUICA.....76
CUADRO Nº 05. RECURSOS MATERIALES 78
CUADRO Nº 06. RECURSOS EN EQUIPOS 79
CUADRO Nº 07. OPERACIONALIZACIÓN DE VARIABLES 81
CUADRO Nº 08. PERSONAL DE LA FACULTAD DE CIENCIAS.....86
CUADRO Nº 09. MATRIZ DE FORTALEZAS, OPORTUNIDADES, DEBILIDADES Y AMENAZAS 87
CUADRO Nº 10. ESTRATEGIAS DEL ANÁLISIS FODA 88
CUADRO Nº 11. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL 97
CUADRO Nº 12. FACTS VS DIMENSIONES.....111
CUADRO Nº 13. MAPEO DE DATOS DEL DATA MART DE LA TABLA ESTUDIANTE.....111
CUADRO Nº 14. MAPEO DE DATOS DEL DATA MART DE LA TABLA DOCENTE 112
CUADRO Nº 15. MAPEO DE DATOS DEL DATA MART DE LA TABLA SEMESTRE 112
CUADRO Nº 16. MAPEO DE DATOS DEL DATA MART DE LA TABLA MODALIDADAPROBACION 112
CUADRO Nº 17. MAPEO DE DATOS DEL DATA MART DE LA TABLA COMACADESTUDIANTE.....112
CUADRO Nº 18. MAPEO DE DATOS DEL DATA MART DE LA TABLA COMPDOC 113
CUADRO Nº 19. MAPEO DE DATOS DEL DATA MART DE LA TABLA CURRICULA 113
CUADRO Nº 20. PROCEDIMIENTO DE OPERACIÓN Y ADMINISTRACIÓN.....119
CUADRO Nº 21. PROCEDIMIENTO DE OPERACIÓN Y ADMINISTRACIÓN.....120
CUADRO Nº 22. GENERACIÓN DEL CÓDIGO DE LOS COMPONENTES Y PROCEDIMIENTOS.....121

CUADRO Nº 23. PROCEDIMIENTO DE OPERACIÓN Y MANUALES DE USUARIO 122
CUADRO Nº 24. CONSTRUCCIÓN DE LOS COMPONENTES Y PROCEDIMIENTOS DE MIGRACIÓN Y CARGA INICIAL DE DATOS 123
CUADRO Nº 25. ELEMENTOS DE PRUEBA Y CRITERIOS DE ACEPTACIÓN 124
CUADRO Nº 26. CASO DE PRUEBA 1 125
CUADRO Nº 27. REGISTROS DE LA TABLA ALUMNO 125
CUADRO Nº 28. REGISTROS DE LA TABLA SEDE 126
CUADRO Nº 29. REGISTROS DE LA TABLA FACULTAD 126
CUADRO Nº 30. REGISTROS DE LA TABLA ESCUELA 126
CUADRO Nº 31. CASO DE PRUEBA Nº2 128
CUADRO Nº 32. REGISTROS DE LA TABLA CURRICULA 128
CUADRO Nº 33. REGISTROS DE LA TABLA SEDE 129
CUADRO Nº 34. REGISTROS DE LA TABLA CURSOS 129
CUADRO Nº 35. REGISTROS DE LA TABLA ESCUELA 129
CUADRO Nº 36. REGISTROS DE LA TABLA FACULTAD 130
CUADRO Nº 37. CASO DE PRUEBA Nº 03 131
CUADRO Nº 38. REGISTROS DE LA TABLA TRABAJADOR 132
CUADRO Nº 39. REGISTROS DE LA TABLA DEDICACION 132
CUADRO Nº 40. REGISTROS DE LA TABLA CATEGORIA 132
CUADRO Nº 41. REGISTROS DE LA TABLA CONDICION 133
CUADRO Nº 42. REGISTROS DE LA TABLA GRADO 133
CUADRO Nº 43. REGISTROS DE LA TABLA DEPARTAMENTOACADEMICO 133
CUADRO Nº 44. CASO DE PRUEBA Nº 04 135
CUADRO Nº 45. REGISTROS DE LA TIPO_APROBACION 135
CUADRO Nº 46. CASO DE PRUEBA Nº 05 137
CUADRO Nº 47. REGISTROS DE LA TABLA SEMESTRE 137
CUADRO Nº 48. CASO DE PRUEBA Nº 06 139
CUADRO Nº 49. REGISTROS DE LA TABLA RENDIMIENTO 139
CUADRO Nº 50. CASO DE PRUEBA Nº 07 141
CUADRO Nº 51. REGISTROS DE LA TABLA ENCUESTA 141
CUADRO Nº 52. REGISTROS DE LA TABLA PREGUNTA 142
CUADRO Nº 53. REGISTROS DE LA TABLA RESPUESTA 142
CUADRO Nº 54. PLAN DE MONITOREO Y EVALUACIÓN 147

CUADRO Nº 55. DISTRIBUCIÓN DE FRECUENCIAS POR NIVEL Y MEDIADAS DE TENDENCIA CENTRAL DEL POST-TEST DEL GRUPO EXPERIMENTAL Y GRUPO CONTROL.....150
CUADRO Nº 56. DISTRIBUCIÓN DE FRECUENCIAS POR NIVEL Y MEDIADAS DE TENDENCIA CENTRAL DEL POST-TEST DEL GRUPO EXPERIMENTAL Y GRUPO CONTROL DE LA DIMENSIÓN DISEÑO.....152
CUADRO Nº 57. DISTRIBUCIÓN DE FRECUENCIAS POR NIVEL Y MEDIADAS DE TENDENCIA CENTRAL DEL POST-TEST DEL GRUPO EXPERIMENTAL Y GRUPO CONTROL DIMENSIÓN ELECCIÓN. 154

SUMARIO DE FIGURAS

FIGURA Nº 01. UNA ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS TÍPICO 35
FIGURA Nº 02. LOS PRINCIPALES COMPONENTES DE LOS SISTEMAS DE NEGOCIO INTELIGENTES....36
FIGURA Nº 03 ARQUITECTURA Y FUNCIONES DE UNA DATA WAREHOUSE....45
FIGURA Nº 04. ESQUEMA ESTRELLA....48
FIGURA Nº 05. ESQUEMA COPO DE NIEVE....49
FIGURA Nº 06. ESQUEMA CONSTELACIÓN DE HECHOS....49
FIGURA Nº 07. PROCESO ETL....50
FIGURA Nº 08. INGENIERÍA DE SOFTWARE....52
FIGURA Nº 09. LAS SIETE DISCIPLINAS PROPUESTAS SE RELACIONAN CON LAS FASES DE DWEP....56
FIGURA Nº 10. LOS DIFERENTES ESQUEMAS DEL DATA MART QUE PROPONE LA METODOLOGÍA DWEP: ESQUEMA CONCEPTUAL, LÓGICO Y FÍSICO....61
FIGURA Nº 11. PROCESO DE TOMA DE DECISIONES....65
FIGURA Nº 12. ESTRUCTURA DE LOS SISTEMAS DE SOPORTE A LA TOMA DE DECISIONES....66
FIGURA Nº 13. ORGANIGRAMA ESTRUCTURAL DE LA UNIVERSIDAD NACIONAL SANTIAGO ANTÚNEZ DE MAYOLO....84
FIGURA Nº 14. DIAGRAMA DE ACTIVIDADES DEL NEGOCIO....90
FIGURA Nº 15. DIAGRAMA DE CONTEXTO....96
FIGURA Nº 16. TOMADORES DE DECISIONES....96
FIGURA Nº 17. DIAGRAMA DE LA ARQUITECTURA TECNOLÓGICA DE LA SOLUCIÓN....99
FIGURA Nº 18. DIAGRAMA DE CLASES DE LA BASE DE DATOS FUENTE....100
FIGURA Nº 19. DIAGRAMA DE MAPEO DE ESTUDIANTE....101
FIGURA Nº 20. PROCESO ETL DE LA DIMENSIÓN ESTUDIANTE....102
FIGURA Nº 21. PROCESO ETL DE LA DIMENSIÓN ESTUDIANTE....103
FIGURA Nº 22. ESQUEMA CONCEPTUAL DEL DATA MART – NIVEL 1....105
FIGURA Nº 23. ESQUEMA CONCEPTUAL DEL COMPORTAMIENTO ACADÉMICO DEL ESTUDIANTE....106
FIGURA Nº 24. ESQUEMA CONCEPTUAL DEL COMPORTAMIENTO DOCENTE..107
FIGURA Nº 25. ESQUEMA CONCEPTUAL DE LA DIMENSIÓN ESTUDIANTE....108

FIGURA Nº 26. ESQUEMA CONCEPTUAL DE LA DIMENSIÓN SEMESTRE 108
FIGURA Nº 27. ESQUEMA CONCEPTUAL DE LA DIMENSIÓN CURRICULA 109
FIGURA Nº 28. ESQUEMA CONCEPTUAL DE LA DIMENSIÓN MODALIDAD DE APROBACIÓN 109
FIGURA Nº 29. ESQUEMA CONCEPTUAL DE LA DIMENSIÓN MODALIDAD DE APROBACIÓN 110
FIGURA Nº 30. PATRÓN DE DISEÑO DEL APLICATIVO WEB 114
FIGURA Nº 31. FLUJO DE ADMINISTRACIÓN DE SEGURIDAD Y CONTROL DE ACCESO.....121
FIGURA Nº 32. CONSULTA A BASE DE ITUNASAM, TABLA ALUMNO 127
FIGURA Nº 33. CONSULTA A BASE DE DATA MARTUNASAM, TABLA ESTUDIANTE.....127
FIGURA Nº 34. CONSULTA A BASE DE ITUNASAM, TABLA CURRICULA 130
FIGURA Nº 35. CONSULTA A BASE DE DATA MARTUNASAM, TABLA CURRICULA 131
FIGURA Nº 36. CONSULTA A BASE DE ITUNASAM, UTILIZANDO EL SCRIPT DE DOCENTE.....134
FIGURA Nº 37. CONSULTA A BASE DE DATA MARTUNASAM, TABLA DOCENTE 134
FIGURA Nº 38. CONSULTA A BASE DE ITUNASAM, TABLA MODALIDAD_APROBACION 136
FIGURA Nº 39. CONSULTA A BASE DE DATA MARTUNASAM, TABLA MODALIDAD_APROBACION 136
FIGURA Nº 40. CONSULTA A BASE DE ITUNASAM, TABLA SEMESTRE.....138
FIGURA Nº 41. CONSULTA A BASE DE DATA MARTUNASAM, TABLA SEMESTRE 138
FIGURA Nº 42. CONSULTA A BASE DE ITUNASAM, TABLA CURRICULA 140
FIGURA Nº 43. CONSULTA A BASE DE DATA MARTUNASAM, TABLA COMACADALUMNO.....140
FIGURA Nº 44. CONSULTA A BASE DE ITUNASAM, AL SCRIPT DE COMPORTAMIENTO DOCENTE.....143
FIGURA Nº 45. CONSULTA A BASE DE DATA MARTUNASAM, TABLA COMDOC 143

SUMARIO DE GRÁFICOS

GRÁFICO Nº 01. GRÁFICO DE BARRAS DE LA FRECUENCIA PORCENTUAL DEL POST-TEST DEL GRUPO EXPERIMENTAL Y CONTROL. 150
GRÁFICO Nº 02. PRUEBA DE HIPÓTESIS DEL GRUPO EXPERIMENTAL.....151
GRÁFICO Nº 03. GRÁFICO DE BARRAS DE LA FRECUENCIA PORCENTUAL DEL POST-TEST DEL GRUPO EXPERIMENTAL Y CONTROL. 152
GRÁFICO Nº 04. PRUEBA DE HIPÓTESIS DEL GRUPO EXPERIMENTAL, DIMENSIÓN DISEÑO.....153
GRÁFICO Nº 05. GRÁFICO DE BARRAS DE LA FRECUENCIA PORCENTUAL DEL POST-TEST DEL GRUPO EXPERIMENTAL Y CONTROL DIMENSIÓN ELECCIÓN. 154
GRÁFICO Nº 06. PRUEBA DE HIPÓTESIS DEL GRUPO EXPERIMENTAL, DIMENSIÓN ELECCIÓN.....155

CAPÍTULO I

GENERALIDADES

1.1. REALIDAD PROBLEMÁTICA

En la actualidad, la información se ha convertido en un recurso y elemento esencial para la empresa, "las empresas buscan emplear dicha información para generar conocimiento útil dirigido a la mejora de sus procesos empresariales" (Zambrano Alarcón 2011, 3). Desde este punto de vista, la ventaja competitiva de las organizaciones radica en la forma de interpretar la información y convertirla en un elemento relevante para las empresas. Además, coincide con Tansey (2003), que menciona que la "Information is a –perhaps the- major resource of modern businesses" (Tansey 2003, 5).

La Universidad, es una entidad orgánica o sistema de unidades operativas, y al igual que la empresa maneja una gran cantidad de información día a día. Realizando así un alto número de transacciones, la cual genera un gran repositorio o almacén de datos. Sin embargo, muchas de las universidades no saben cómo administrar dicha información adecuadamente, debido a que sus

sistemas de información implantados para el control académico, y que además soportan las transacciones académicas, no soportan el manejo adecuado de grandes volúmenes de información. Así, las universidades tienen problemas al momento de utilizar su información contenida en sus repositorios o almacenes de datos para emplearla en la toma de decisiones, ya que esta información necesita ser procesada y sistematizada, y para lograrlo requiere un tiempo, personal calificado, generando un alto costo para esta entidad. Siendo considerado a este proceso como ineficiente, ya que retrasa la toma de decisiones oportunas.

Esto es uno de los problemas por la cual la Universidad Nacional Santiago Antúnez de Mayolo (UNASAM) está atravesando, específicamente la Facultad de Ciencias y sus escuelas. Para realizar el control y seguimiento de los procesos académicos de las escuelas profesionales de la UNASAM, las cuales apoyan su operatividad en el SIGA, cuyas siglas significan “Sistema Integral de Gestión Académica” , el cual como parte de su estructura tiene a la base de datos OLTP (Procesamiento de Transacciones en Línea) llamada “ITUnasam”, en la cual se encuentran la mayoría de los datos relevantes para la institución, tales como los datos de los alumnos, docentes, mallas curriculares, cursos, entre otras.

Siendo el SIGA administrada directamente por la oficina de “Administración de la Base de Datos” (DBA) de la oficina “General de Estudios”, quienes haciendo uso de herramientas y técnicas especializadas de base de datos y DBMS SQL Server, preparan y organizan la información requerida para ser entregada posteriormente al jefe de la oficina general de estudios o a los decanos o jefes de departamento para que efectúen sus análisis respectivos posteriores y finalmente tomen las decisiones que correspondan, de esta forma la consulta de información relevante se encuentra centralizada, generando colas y retrasos a la toma de decisiones oportunas.

Por otro lado, la Facultad de Ciencias de la Universidad Nacional Santiago Antúnez de Mayolo, cuenta con tres escuelas profesionales como son la escuela de Estadística e Informática, la escuela de Ingeniería de Sistemas e Informática y la escuela de Matemática. El decano, quien es la máxima autoridad de la Facultad, los jefes de departamento, directores de escuela y docentes que conforman la Facultad, son los encargados de la toma de decisiones y la planificación diaria, a mediano plazo y/o largo plazo, siendo la calidad, la disponibilidad y la presentación de la información un papel categórico para la toma de decisiones oportunas. Siendo estos los principales usuarios que requieren disponer de información tanto consolidada como detallada de cómo marchan las actividades ya cumplidas, para así predecir tendencias y comportamientos para la toma de decisiones proactivas.

El control y procesamiento de todos los datos generados durante los diversos procesos en la que está inmersa la Universidad nacional Santiago Antúnez de Mayolo son almacenados en la base de datos, para luego ser administradas directamente por la Oficina General de Estudios, lo que genera para las Facultades la dificultad de acceder a la información relevante para la toma de decisiones por parte de cada uno de los usuarios.

Además de ello, dentro del proceso de acreditación por cual la Universidad nacional Santiago Antúnez de Mayolo y sobre todo la Facultad de Ciencias está atravesando, es necesario la disponibilidad de información procesada como se menciona en la dimensión formación profesional factor Enseñanza – Aprendizaje, pero actualmente la Facultad de Ciencias no dispone de información, datos procesados, que permita responder a cuestiones del negocio que serán utilizados para la toma de decisiones; esto dificulta y retrasa este proceso.

Por todo lo sustentado en la realidad problemática, se planteó el siguiente problema:

1.2. ENUNCIADO DEL PROBLEMA

Los enunciados del problema planteados para el presente proyecto de investigación, fueron las siguientes:

1.2.1. GENERAL

¿En qué medida la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, mejora el proceso de toma de decisiones de la gestión académica de la Facultad de Ciencias de la UNASAM?

1.2.2. ESPECÍFICOS

- ¿De qué manera la implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye en el diseño de la toma de decisiones de la gestión académica?

- ¿De qué manera la implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye en la elección de las decisiones de la gestión académica?

1.3. HIPÓTESIS

Las hipótesis planteadas para el presente proyecto de investigación, son los siguientes:

1.3.1. GENERAL

H_A : Con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora el proceso de

Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM.

H_0 : Con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, no mejora el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM.

1.3.2. ESPECÍFICO

H_{1A} : La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en el diseño de la toma de decisiones de la gestión académica.

H_{10} : La implementación de un Data Mart, como parte de una solución de inteligencia de negocio no influye significativamente en el diseño de la toma de decisiones de la gestión académica.

H_{2A} : La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en la elección de la toma de decisiones de la gestión académica.

H_{20} : La implementación de un Data Mart, como parte de una solución de inteligencia de negocio no influye significativamente en la elección de las decisiones de la gestión académica.

1.4. OBJETIVOS

1.4.1. GENERAL

Implementar un Data Mart, como parte de una solución de inteligencia de negocio, para mejorar el proceso de Toma de Decisiones en La Gestión Académica de Facultad de Ciencias en la UNASAM.

1.4.2. ESPECÍFICOS

- Identificar y describir los requerimientos de análisis de información para el comportamiento académico de la Facultad de Ciencias de la UNASAM.

- Diseñar el modelo dimensional que permita el análisis y explotación de la información identificada.

- Desarrollar la estructura y funcionalidad del Data Mart, como parte de una solución de Inteligencia de Negocio.

- Implementar el Data Mart, como parte de una solución de Inteligencia de Negocio.

- Facilitar la toma de decisiones, en base a la Información presentada por el Data Mart.

1.5. JUSTIFICACIÓN

Esta investigación propuso el desarrollo de un Data Mart para la mejora de la Toma de Decisiones en la Facultad de Ciencias, aportando con una herramienta de gran utilidad para el nivel estratégico de la Facultad de Ciencias de la UNASAM.

El presente proyecto se justifica porque actualmente se requiere saber acerca del comportamiento Académico, a fin de identificar indicadores tales como la cantidad de alumnos que aprobaron un determinado curso en un determinado semestre, entre otros. Todos estos indicadores ayudarán a la toma de decisiones en el ámbito académico.

A. JUSTIFICACIÓN ACADÉMICA

Desde el punto de vista académico es justificable, debido que se contribuye con nuevos conocimientos a los alumnos de la carrera profesional de ingeniería de Sistemas e Informática de la UNASAM, fortaleciendo su formación profesional y sirviendo de ayuda para posteriores trabajos.

B. JUSTIFICACIÓN OPERATIVA

Desde el punto de vista operativo es justificable, debido que el proceso de obtención de información relevante es más rápido, confiable, eficiente y eficaz; además este proceso se realiza en un tiempo menor que en las etapas anteriores.

C. JUSTIFICACIÓN ECONÓMICA

Desde el punto de vista económico es justificable, ya que se reduce los costos de obtención, procesamiento y resultados de información, tales como las horas/hombre y la reducción de actividades que no agregan valor a la institución.

D. JUSTIFICACIÓN TÉCNOLOGICA

Desde el punto de vista técnico es justificable, debido a que para el desarrollo del presente proyecto se contó con los conocimientos técnicos y herramientas necesarias para la realización del presente.

E. JUSTIFICACIÓN LEGAL

Según la ley universitaria, ley Nº30220, la investigación constituye una función esencial y obligatoria de la universidad, y apoya a la producción de conocimiento y al desarrollo de tecnología, la que es mencionada en el artículo 48 de la presente ley. Además, según la presente la investigación es requisito indispensable para la obtención de título profesional.

1.6. LIMITACIONES

Dentro de las limitaciones del presente trabajo de investigación tenemos:

Las limitaciones de carácter económico, se evidenciaron en cuanto a la adquisición y acceso a material especializado, libros, artículos entre otros, las mismas que fueron solucionados utilizando los foros de las redes sociales, tales como Facebook, Twitter entre otros en las cuales algunos usuarios de todo el mundo comparten materiales educativos de forma gratuita.

Las limitaciones de tiempo, cuando se hace investigación de desarrollo tecnológico, el tiempo es un factor determinante, en cuanto al tiempo de desarrollo de la herramienta y al tiempo de experimentación para la obtención de resultados; sin embargo, se tuvo que someterse a limitaciones que se presentaron por parte de la universidad.

Teniendo en cuenta las limitaciones económicas y de tiempo, la investigación superó estos problemas adecuando el presente trabajo de acuerdo al avance de la ciencia y tecnología.

1.7. DESCRIPCIÓN Y SUSTENTACIÓN DE LA SOLUCIÓN

1.7.1. DESCRIPCIÓN DE LA SOLUCIÓN

La solución desarrollada en la presente investigación consistió en la creación de un Data Mart, como soporte para la toma de decisiones de la Facultad de Ciencias de la Universidad Nacional Santiago Antúnez de Mayolo, con el fin que sea una herramienta diseñada bajo la concepción de la Inteligencia de Negocios para que sea el soporte a la toma de decisiones de todos los usuarios.

Teniendo en cuenta que la presente solución no solo consistió en el diseño del modelo multidimensional del Data Mart. Esta solución abarco varios puntos que se detallan a continuación. La integración de todos estos puntos forma parte de la solución integral al problema.

1.7.1.1. FUNCIONALIDAD

El Data Mart debe permitir cumplir con las necesidades de información requeridas, es decir, debe contener la funcionalidad adecuada. Sin ella, el Data Mart no cumple con su objetivo principal, por tanto, el proyecto no logra solucionar el problema planteado y finalmente no comprueba la hipótesis. Por ello, el Data Mart debe estar orientado a satisfacer las necesidades de la gestión Académica de la Facultad de Ciencias.

1.7.1.2. MODELO MULTIDIMENSIONAL

Constituye el punto fundamental de la solución debido a que el modelo multidimensional es la base del Data Mart. El modelo multidimensional a desarrollar es de tipo estrella y tiene como dimensiones a: Alumnos, Curricula, Docentes y Modalidad de Aprobación y Semestre; y siendo los hechos:

Comportamiento Alumno y Comportamiento Docente. Logrando así abarcar los posibles escenarios y supuestos de la gestión académica de la Facultad de Ciencias.

Para llevar a cabo esta labor se ha tenido que levantar información y haber definido claramente los requerimientos de los usuarios en cuanto a la gestión académica de la Facultad de Ciencias.

Para luego pasar a la construcción del Data Mart. En la construcción del Data Mart, primero se elaboró el modelo lógico en la cual nos muestra las relaciones e interacciones entre las tablas. Luego de ello se elaboró el modelo físico, en el cual se definió los tipos de datos y tamaños más adecuados para cada uno de los campos de cada una de las tablas.

1.7.1.3. PROCESO ETL

Un Data Mart es un almacén de datos especializado, y como tal requiere de datos. Para lograr ello fue necesario algún procedimiento que permita cargar de datos válidos al Data Mart.

La solución planteada incluye un proceso ETL el cual extrae los datos de la base de datos del SIGA o ITUNASAM que viene hacer la principal fuente de datos, para luego transformar la estructura de datos a fin de que coincida con la estructura del modelo del Data Mart y para luego realizar la carga de datos a éste.

1.7.1.4. HERRAMIENTA DE EXPLOTACIÓN

Si bien el modelo multidimensional constituía en gran parte la solución, es la herramienta de explotación la que finalmente con la que el usuario puede interactuar y ver el Data Mart.

La herramienta de explotación es un sistema que recibe una base de datos de modelo multidimensional y permite visualizarla de una manera fácil e intuitiva.

La solución incluye la implementación de una herramienta de explotación de datos, la cual fue diseñada y construida en base a las necesidades del usuario final, adicionalmente esta herramienta fue construida utilizando las herramientas y procedimientos de la Ingeniería de Software, el ultimo que incluye los métodos y procesos que facilitan el desarrollo de un software (Pressman 2006).

1.7.2. SUSTENTACIÓN DE LA SOLUCIÓN

La solución planteada es la más adecuada debido a dos razones principales: que es una solución de Inteligencia de Negocios y está basado en la creación de un Data Mart.

A continuación, se explicará por qué estas dos razones hacen que la solución propuesta sea considerada como la más adecuada y ventajosa.

1.7.2.1. SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS

Al ser una solución de Inteligencia de Negocios, el Data Mart propuesto permite generar conocimiento útil a partir de una gran cantidad de información organizacional.

Las soluciones de este tipo permiten administrar adecuadamente la información para utilizar sólo la información que los usuarios requieran al momento deseado. Los tiempos de respuesta a la información deseada son considerablemente superiores a los sistemas transaccionales.

Esta información presentada puede ser resumida o detallada según la necesidad del usuario.

Además, ofrece la capacidad de crecimiento de la información a medida que la organización realizar sus operaciones del día a día y permite actualizar la información a través de los procesos ETL. De esta manera se asegura que siempre se cuente con la información actualizada y real.

En conclusión, el Data Mart, con el uso correcto, podrá ser una herramienta de soporte en la toma de decisiones.

1.7.2.2. CREACIÓN DE DATA MART

Se eligió crear un Data Mart frente a un Data Warehouse debido a que el primero ofrece una serie de ventajas. Un Data Mart es más fácil de implementar e instalar que un Data Warehouse corporativo. De esto se concluye que el costo de construcción de un Data Mart es considerablemente menor al del Data Warehouse. Así muchas empresas pueden tener acceso

al Data Mart debido a que está dentro de sus posibilidades económicas.

Por otro lado, los Data Marts, al ser más pequeños, están enfocados a satisfacer las necesidades de un grupo de usuarios en concreto y permite que el tiempo de respuesta a sus consultas sea más rápido.

CAPÍTULO II

MARCO TEÓRICO

2.1. ANTECEDENTES

Ante la creciente demanda de información procesada dentro del campo de la inteligencia de negocio se han desarrollado innumerables investigaciones dentro de este campo, y luego de haber realizado las indagaciones respectivas, se encontró los siguientes trabajos de Investigación que se asemejan al presente:

Reyes Marroquín y Rosales Tejada (2007) en su estudio titulado "Desarrollo de un Data Mart de información académica de estudiantes de la escuela de ciencias y sistemas de la Facultad de Ingeniería de la USAC", presentado para la obtención del título de Ingeniero en Ciencias y Sistemas, estudio de tipo aplicado, concluye en lo siguiente: En la escuela de ciencias y sistemas se obtiene grandes beneficios al utilizarse el Data Mart académico, puesto que se podrá analizar el comportamiento de los estudiantes de la escuela

y por ende se podrá tomar mejores decisiones en cuanto al uso de los recursos y el enfoque de la carrera de sistemas (Reyes Marroquín y Rosales Tejada 2007). Además “La creación del presente Data Mart permite que las diferentes áreas o unidades de la Facultad cuenten con información académica de una forma autónoma, sin que exista la dependencia de centro de cálculo, siempre guardando los debidos controles de seguridad y acceso a la información” (Reyes Marroquín y Rosales Tejada 2007) y finalmente “Contar con una herramienta de inteligencia de negocio en la Escuela de Ciencias y Sistemas, facilitará la información que muchas empresas requieren sobre referencias de los estudiantes que solicitan empleos” (Reyes Marroquín y Rosales Tejada 2007). Los resultados obtenidos en este estudio demuestran que la implementación de un Data Mart apoya a la toma de decisiones a la universidad, permitiendo a los usuarios finales que puedan conectarse desde cualquier parte de la Facultad sin tener que instalar ningún software-cliente para poder hacer análisis de la información. Al mismo tiempo desde sus conclusiones se infiere que es una herramienta de negocio es útil y beneficia a la organización

Zambrano Alarcón en su tesis titulada “Análisis, diseño e implementación de un Data Mart para el área de mantenimiento y logística de una empresa de transporte público de pasajeros”, presentada a la Pontificia Universidad Católica del Perú en el año 2011 para obtener el título de Ingeniero Informático, cuyo estudio fue de tipo aplicado, orientada para el apoyo al proceso de toma de decisiones del área de mantenimiento y logística de una empresa de transporte público de pasajeros. Cuya investigación arribo a las siguientes conclusiones, primero “La metodología DWEP logró guiar exitosamente toda la construcción de la solución. Su integración con la metodología RUP y los diagramas UML permitió documentar los requerimientos, análisis, diseño e implementación de la solución de manera clara y sin ambigüedades. Si bien UML está orientado a la construcción de sistemas orientados a objetos, se ha demostrado en el presente proyecto que, empleando estereotipos basados en UML, puede modelarse la construcción de una solución

basada en inteligencia de negocios" (Zambrano Alarcón 2011), segundo "Las necesidades de información del área de mantenimiento y logística de una empresa de transporte público fueron identificadas satisfactoriamente debido a que se consideró los posibles escenarios, actores y supuestos en toda empresa del mismo rubro. Esto contribuyó a identificar requerimientos claros y precisos que fueron documentados y utilizados para la construcción del modelo multidimensional" (Zambrano Alarcón 2011), tercero "El modelo multidimensional de la solución logró abarcar las necesidades de información identificadas y fue representado utilizando diagramas de fácil comprensión que permitieron una correcta validación del mismo" (Zambrano Alarcón 2011), cuarto "Los procesos de extracción, transformación y carga de los datos lograron contar con un Data Mart con datos correctos y coherentes provenientes de la base de datos fuente. Los procesos fueron implementados empleando estándares de programación y luego se verificó su funcionamiento a través de un plan de casos de prueba" (Zambrano Alarcón 2011) finalmente "Los reportes y tableros de control elaborados permitieron mostrar la importancia de la explotación de la información y de los indicadores que generan una ventaja competitiva en las empresas de transporte público" (Zambrano Alarcón 2011). Ante ello el responsable del estudio, hace hincapié sobre la utilización de la metodología DWEP, que es una metodología dedicada originalmente al desarrollo de un Data Warehouse y que propone una gran cantidad de entregables. Al mismo tiempo menciona la importancia de cada proceso del Data Mart.

Flores Chacón at al (2008) en su trabajo de investigación titulado "Diseño, implantación y explotación de Data Mart" presentado a la Universidad Nacional Santiago Antúnez de Mayolo, de tipo de investigación aplicada, arribó a las siguientes conclusiones: El usuario del sistema puede efectuar cualquier consulta de acuerdo a las variables y criterios de consulta del Data Warehouse así como hacer uso de múltiples variables; - La aplicación de la Ingeniería Empresarial a la solución de problemas de la gestión académica de la universidad, como un medio de sinergia de la gestión de los negocios y de las

tecnologías de la información y la comunicación en un contexto de competitividad, calidad y globalización; - El proyecto ha permitido aplicar conceptos de las tecnologías de la información y la comunicación, aplicados a la gestión de la universidad (de los negocios) como son: Data Warehouse y Business Intelligence.

2.2. INTELIGENCIA DE NEGOCIOS

La Inteligencia de Negocios (en inglés Business Intelligence) nació como un concepto que se asociaba totalmente con los niveles de los altos directivos ante la necesidad de contar con información para dirigir el rumbo de la empresa. En la actualidad, la información se ha convertido en un bien muy preciado para las empresas.

Las empresas que son capaces de transformar los datos en información y el conocimiento pueden utilizarlos para tomar decisiones más rápidas y más eficaces y por lo tanto para lograr una ventaja competitiva.

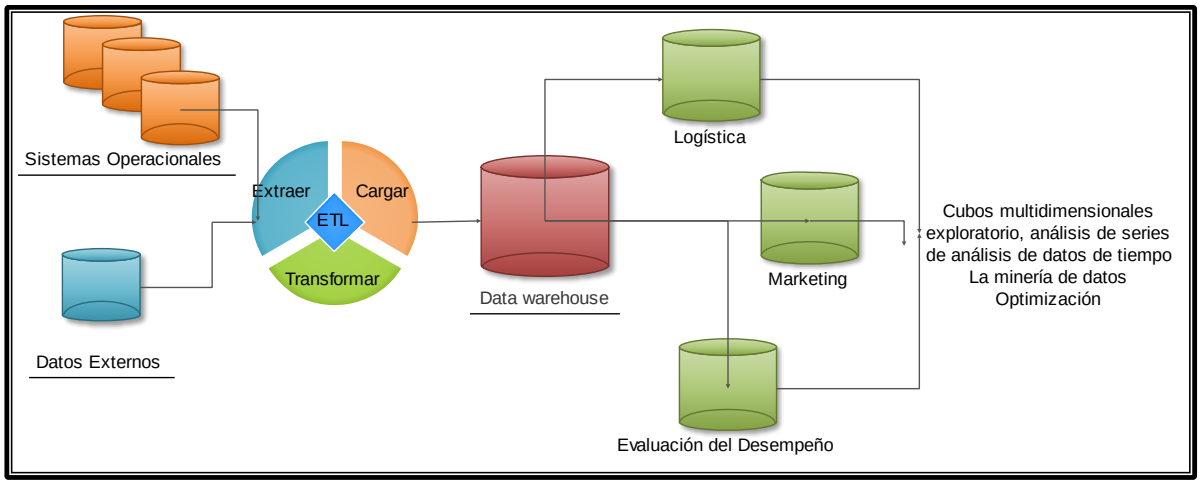
Carlo Vercellis, en su libro "Business Intelligence, Data Mining and Optimization for Decision Making" define a la inteligencia de negocios como un conjunto de modelos matemáticos y metodologías de análisis que explotan sistemáticamente los datos disponibles para recuperar la información y conocimiento útil en el apoyo a los procesos de toma de decisiones complejas. (Vercellis 2009)

Luis Méndez en su libro "Más allá del Business Intelligence" lo definió como un conjunto de herramientas y aplicaciones para la ayuda a la toma de decisiones que posibilitan acceso interactivo, análisis y manipulación de información corporativa de misión crítica (Río 2006). Estas aplicaciones aportan un conocimiento valioso sobre la información operativa identificando problemas y oportunidades de negocio.

2.2.1. ARQUITECTURAS DE INTELIGENCIAS DE NEGOCIOS

La arquitectura de un sistema de inteligencia de negocios, puede incluir dos componentes principales de los sistemas de Inteligencia de Negocio (Las Fuentes de Datos, Almacenes de Datos), como es representado en la figura siguiente.

Figura Nº 01. Una arquitectura de inteligencia de negocios típico

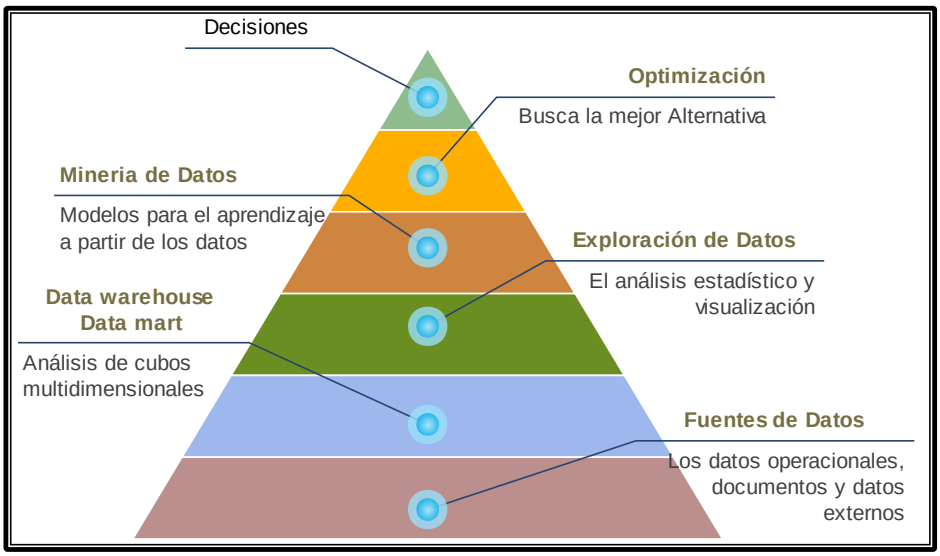


***Fuente**: Business Intelligence Data Mining and Optimization for Decision Making (Vercellis, Carlo)*

2.2.2. COMPONENTES DE UN SISTEMA DE INTELIGENCIA DE NEGOCIO

Está conformada por seis componentes, que es Ilustrada en la figura Nº 02 - Los Principales componentes de los sistemas de Negocio Inteligentes.

Figura Nº 02. Los Principales componentes de los sistemas de Negocio Inteligentes



***Fuente**: Business Intelligence Data Mining and Optimization for Decision Making (Vercellis, Carlo)*

A. LAS FUENTES DE DATOS.

En una primera etapa, es necesario reunir e integrar los datos almacenados en las diversas fuentes primarias y secundarias, que son heterogéneos en origen y tipo.

Las fuentes consisten en su mayor parte de los datos pertenecientes a los sistemas de operacionales, pero también pueden incluir documentos no estructurados, tales como mensajes de correo electrónico y datos recibidos de los proveedores externos. (Vercellis 2009)

B. DATA WAREHOUSE Y DATA MART.

El uso de herramientas de extracción y transformación conocido como extracción, transformación, carga (ETL), los datos procedentes de las diferentes fuentes se almacenan en bases de datos destinadas a apoyar los análisis de inteligencia de negocios.

C. METODOLOGÍAS DE INTELIGENCIA DE NEGOCIOS

Los datos se extraen y finalmente son utilizados para alimentar los modelos matemáticos y metodologías de análisis destinados a apoyar a los tomadores de decisiones.

D. LA EXPLORACIÓN DE DATOS

En el tercer nivel de la pirámide se encuentran las herramientas para formación de un análisis de inteligencia empresarial pasiva, que consisten en sistemas de consulta y presentación de informes, así como los métodos estadísticos.

E. LA MINERÍA DE DATOS

El cuarto nivel de inteligencia de negocios incluye metodologías activas, cuyo objetivo es la extracción de la información y el conocimiento a partir de datos.

Estos incluyen modelos matemáticos para el reconocimiento de patrones, aprendizaje automático y técnicas de minería de datos. A diferencia de las herramientas descritas en el nivel anterior de la pirámide, los modelos de un tipo activo no requieren los tomadores de decisiones para formular ninguna hipótesis previa que ser verificado más tarde.

F. OPTIMIZACIÓN

Al mover un nivel hacia arriba en la pirámide encontramos modelos de optimización que nos permiten determinar la mejor solución de un conjunto de acciones alternativos, que suele ser bastante extensa y, a veces incluso infinito.

G. DECISIONES

Por último, la parte superior de la pirámide corresponde a la elección y la adopción de una decisión específica, y de alguna manera representa la conclusión natural del proceso de toma de decisiones. Incluso cuando las metodologías de inteligencia de negocios están disponibles y adoptadas con éxito, que también pueden tomar ventajas de la información informal y no estructurada disponible para adaptar y modificar las recomendaciones y las conclusiones alcanzadas a través del uso de modelos matemáticos.

A medida que avanzamos desde el fondo hasta la parte superior de la pirámide, los sistemas de negocio de la inteligencia ofrecen herramientas de apoyo cada vez más avanzadas de un tipo de activo. Incluso las funciones y competencias cambian.

En la parte inferior, se proporcionan las competencias requeridas en su mayor parte por los especialistas en sistemas de información dentro de la organización, por lo general se hace referencia como administradores de bases de datos. Analistas y expertos en modelos matemáticos y estadísticos son responsables de las fases intermedias. Por último, las actividades de los encargados de tomar decisiones responsables para el dominio de aplicación aparecen dominantes en la parte superior.

Como se ha descrito anteriormente, los sistemas de inteligencia de negocios frente a las necesidades de los diferentes tipos de organizaciones complejas, incluidos los organismos de administración y las asociaciones públicas.

2.3. DATA WAREHOUSE

Los Data Warehouse nacen debido a la necesidad de contar con información útil de apoyo a la toma de decisiones, dado que los datos operacionales no cumplen con este objetivo (Kimball y Ross 2013). Un Data Warehouse es una colección de datos integrados, temáticos, no volátiles y variantes en el tiempo, organizados para soportar necesidades empresariales orientadas a la toma de decisiones (Inmon 2005)

Un Data Warehouse es un almacén de datos disponibles más importante para el desarrollo de arquitecturas de inteligencia de negocios y sistemas de soporte a la toma de decisiones (Vercellis 2009).

El almacenamiento de datos a largo plazo indica que el conjunto de actividades interrelacionadas que intervienen en el diseño, la implementación y el uso de un almacén de datos (Kimball y Ross 2013).

Es posible identificar tres categorías principales de datos que alimentan a un almacén de datos: los datos internos, datos externos y datos personales.

A. DATOS INTERNOS

Los datos internos se almacenan en su mayor parte en las bases de datos, que se refiere a los sistemas como transaccionales o sistemas operativos, que son la columna vertebral de un sistema de información de la empresa.

B. DATOS EXTERNOS

Hay varias fuentes de datos externos que puede utilizarse para ampliar la cantidad de información almacenada en las bases de datos internas.

C. DATOS PERSONALES

En la mayoría de los casos, los tomadores de decisiones que realizan un análisis de inteligencia de negocios también se basan en la información y apreciaciones personales almacenados en el interior de las hojas de cálculo o bases de datos locales situados en sus computadoras. La recuperación de dicha información y su integración con datos estructurados procedentes de fuentes internas y externas es uno de los objetivos de los sistemas de gestión del conocimiento.

Las aplicaciones de software operativas y/o transaccionales son tipo OLTP (on-line transaction processing). Por otro lado, el conjunto de herramientas destinadas a la realización de los análisis de inteligencia empresarial y el apoyo a los procesos de toma de decisiones son del Tipo OLAP (on-line analytical processing) (Kimball y Ross 2013).

Hay varias razones para implementar un almacén de datos por separado de las bases de datos que apoyan las aplicaciones OLTP en una empresa. Entre ellos, recordamos aquí las más relevantes.

1. INTEGRACIÓN.

En muchos casos, los sistemas de apoyo de decisiones deben tener acceso de la información procedente de varias fuentes de datos, distribuidos en diferentes partes de la organización o que se deriven de fuentes externas. Entonces se requiere un almacén de datos que integren múltiples ya menudo heterogéneas fuentes de promover y facilitar el acceso a la información.

2. CALIDAD.

Los datos transferidos de los sistemas operativos en el almacén de datos se examinan y corrigen a fin de obtener información fiable y libre de errores, tanto como sea posible.

3. EFICIENCIA

Consultas encaminadas a la extracción de información para un análisis de inteligencia de negocios pueden llegar a ser una carga en términos de recursos informáticos y tiempo de procesamiento.

4. POSIBILIDAD DE AMPLIACIÓN

Los datos almacenados en los sistemas transaccionales se extienden en un lapso de tiempo limitado en el pasado. Se puede definir un almacén de datos como una recolección de datos que apoyen los procesos de toma de decisiones y los sistemas de inteligencia de negocios que tienen las siguientes características.

5. ENTIDAD ORIENTADA

Los datos contenidos en un almacén de datos se refieren principalmente a las principales entidades de interés para el análisis.

6. INTEGRADO

Los datos procedentes de las distintas fuentes se integran y se homogeneizaron a medida que se cargan en un almacén de datos.

7. VARIANTE EN EL TIEMPO

Todos los datos introducidos en un almacén de datos se etiquetan con el período de tiempo al que se refieren. De este modo, las aplicaciones

de apoyo a las decisiones pueden desarrollar análisis de tendencias históricas.

8. PERSISTENTES

Una vez que se han cargado en un almacén de datos, los datos son por lo general sin modificar aún más y se llevan a cabo de forma permanente. Esta característica hace que sea más fácil organizar el acceso de sólo lectura por los usuarios y simplifica el proceso de actualización, evitando la concurrencia que es de importancia crítica para los sistemas operativos.

9. CONSOLIDADO

Por lo general, algunos datos almacenados en un almacén de datos se obtienen como resúmenes parciales de los datos primarios que pertenecen a los sistemas operativos de las que proceden.

Cuadro Nº 01. Diferencias entre OLTP y OLAP

Características	OLTP	OLAP
Volatilidad	Datos dinámicos	Datos estáticos
Oportunidad	Datos actuales sólo	Los datos actuales e históricos
Dimensión temporal	Implícita y corriente	Explícitos y la variante
Granularidad	Datos detallados	Datos agregados y consolidados
Actualización	Continuas e irregulares	Periódicas y regulares
Actividades	Repetitivo	Impredecible
Flexibilidad	Baja	Alta
Rendimiento	Alto, pocos segundos por consulta	Puede ser baja para consultas complejas
Usuarios	Empleados	Trabajadores del conocimiento
Funciones	Operacionales	Analíticas
Prioridad	Alto rendimiento	Alta flexibilidad
Métricas	Tasa de transacciones	Respuesta eficaz
Tamaño	Megabytes a Gigabytes	Gigabytes a Terabytes

***Fuente**: Elaboración Propia*

Se puede concluir, que un Data Warehouse, es el proceso de extraer datos de las operaciones diarias de la empresa, procedentes de diversos subsistemas,

para integrarlos, sumarlos y almacenarlos en un depósito de datos, para poder acceder a ellos cada vez que el usuario lo requiera.

Basándose en la definición de Inmon, un Data Warehouse se caracteriza por ser:

1. **INTEGRADO**

Su característica más importante, pues posee la información integrada. Los datos almacenados en un Data Warehouse deben integrarse en una estructura consistente. Esta estructura permite tener la información en distintos niveles de detalle para adecuarse a las necesidades del usuario.

2. TEMÁTICO

Los datos se organizan por temas para facilitar su acceso y su entendimiento por parte de los usuarios finales.

3. NO VOLÁTIL

La perspectiva estratégica que permite el análisis y la toma de decisiones requiere una base de datos estable, no cambiante.

El almacén de información de un datawarehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del datawarehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

4. VARIANTE EN EL TIEMPO

La información contenida en el Data Warehouse sirve para predecir tendencias. Por ello, esta se carga con los distintos valores que tiene una variable en el tiempo permitiendo comparaciones.

2.4. DATA MART

El acceso a toda la información por parte de los usuarios de la empresa no es conveniente, ya que muchas veces sólo necesitan un subconjunto de esta información. En estos casos utilizan los Data Mart.

Según Inmon, el concepto de Data Mart es una especialización de un Data Warehouse enfocado a un departamento (Inmon 2005).

Los Data Marts son sistemas que recopilan todos los datos requeridos por un departamento específico de la empresa, tales como el marketing o la logística, con el fin de llevar a cabo la inteligencia empresarial y análisis de la ejecución de aplicaciones de apoyo a las decisiones específicas de la función misma. Por lo tanto, un Data Mart se puede considerar como un almacén de datos funcional o departamental de un tamaño más pequeño y un tipo más específico que el almacén de datos de la empresa en general (Kimball y Ross 2013).

Por tanto, un Data Mart contiene un subconjunto de los datos almacenados en el Data Warehouse de la empresa, que por lo general están integradas con otros datos que el departamento compañía responsable del mercado de datos posee y considere de interés.

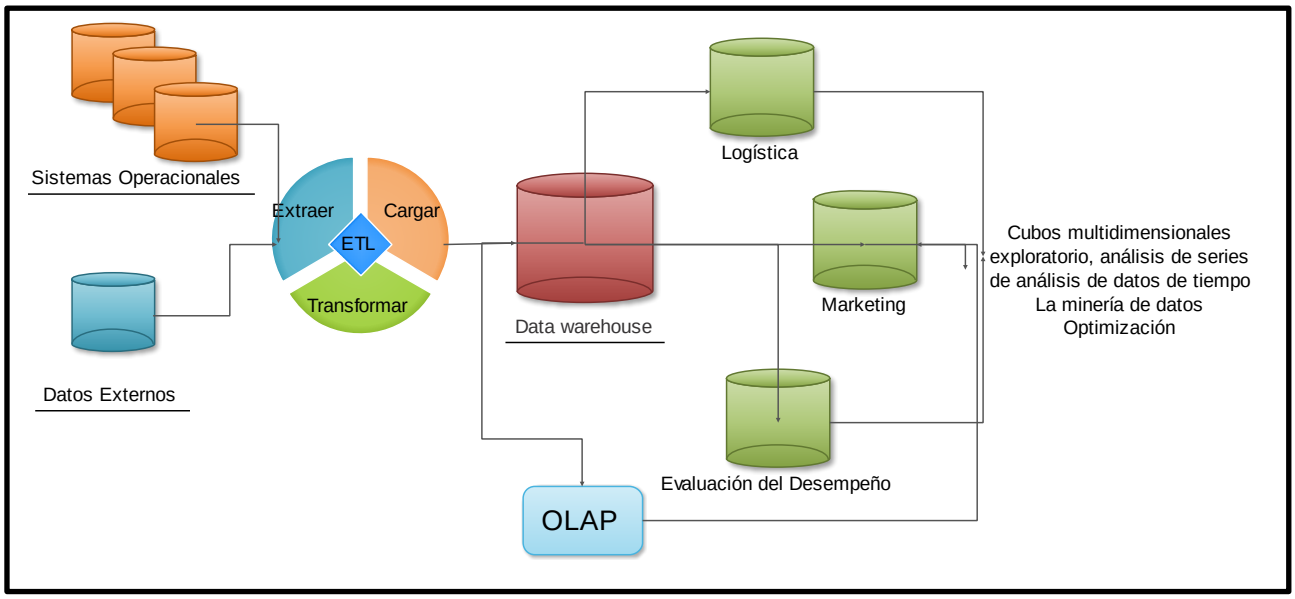
Los Data Warehouse y data Mart por lo tanto comparten el mismo marco tecnológico. Con el fin de implementar las aplicaciones de inteligencia de negocios, algunas empresas prefieren diseñar y desarrollar de manera incremental una serie de Data Mart integrados en lugar de un depósito de datos

central, con el fin de reducir el tiempo de implementación y las incertidumbres relacionadas con el proyecto.

2.5. ARQUITECTURA DATA MART

La arquitectura de un Data Mart es similar a la arquitectura de un data Warehouse, Vercellis(2009) menciona en su estudio que la arquitectura de un Data Warehouse está compuesta por todos los datos internos y externos, el proceso ETL, el repositorio del data warehouse entre otros, que se ilustran en la figura Nº 03, incluye los siguientes componentes funcionales principales.

Figura Nº 03 Arquitectura y funciones de una Data Warehouse



***Fuente**: Business Intelligence Data Mining and Optimization for Decision Making (Vercellis, Carlo)*

2.6. MODELO DIMENSIONAL

Un modelo dimensional es un sistema formal y abstracto que permite describir los datos de acuerdo con reglas y convenios predefinidos. Es formal, pues los objetos del sistema se manipulan siguiendo reglas perfectamente definidas y utilizando exclusivamente los operadores definidos en el sistema,

independientemente de lo que estos objetos y operadores puedan significar (Ullman y Widom 1999).

La arquitectura de un Data mart y/o Data Warehouse se basa en un modelo de datos llamado modelo multidimensional. Este modelo permite modelar base de datos simples y entendibles al usuario final, debido que presenta la información en un marco estándar e intuitivo que permite un acceso de alto rendimiento. Además, permite resolver con problemas planteados en sistemas transaccionales o los sistemas OLTP (On-line Transaction Processing, procesamiento transaccional en línea) generando una base de datos multidimensional que permitirá el análisis de datos utilizando herramientas OLAP (on-line Analytical Processing, procesamiento analítico en línea).

Los almacenes de datos posibilitan una visión multidimensional de enormes cantidades de datos históricos provenientes de fuentes operacionales, suministrando la información necesaria para el apoyo a los procesos de toma de decisiones de una organización (Mazón, Pardillo y Trujillo 2011, 27).

En base a lo mencionado del modelo multidimensional, se puede desglosar en dos fases para la construcción de un modelo multidimensional.

2.6.1. MODELO MULTIDIMENSIONAL CONCEPTUAL

El modelo dimensional conceptual debe representar aquellos elementos que permitan al usuario obtener la información necesaria para el apoyo a la toma de decisiones (Mazón, Pardillo y Trujillo 2011, 28). Entre los elementos o propiedades multidimensionales estructurales tenemos:

A. DIMENSIONES

Representa el contexto de análisis de un hecho… mediante una serie de atributos organizados jerárquicamente (Mazón, Pardillo y Trujillo 2011, 27). Las tablas de dimensiones representan cada uno de los ejes en un espacio multidimensional. Sus atributos son del tipo cualitativo que proporcionan el contexto en el que se obtienen las medidas en un esquema de hecho. Las dimensiones poseen jerarquías, que son varios atributos unidos mediante una relación del tipo jerárquico.

B. HECHOS

Contiene medidas de un proceso de negocio (Mazón, Pardillo y Trujillo 2011, 27). Las tablas de hechos constituyen el objeto a analizar, poseen atributos de hechos que son del tipo cuantitativo cuyos valores se obtienen por aplicación de alguna función estadística que resumen un conjunto de valores en un único valor.

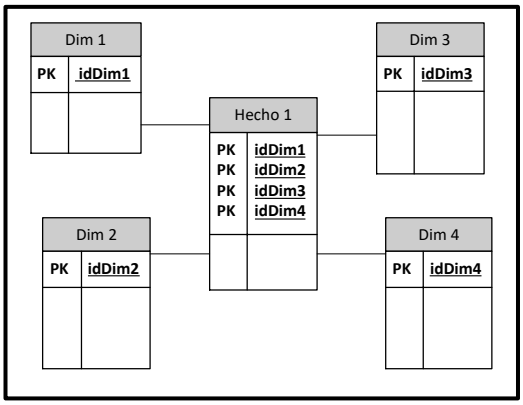
2.6.2. MODELO MULTIDIMENSIONAL LÓGICO

Dentro del modelo multidimensional lógico existe dos tecnologías, la tecnología Multidimensional – MOLAP (OLAP Multidimensional) y la tecnología Relacional – ROLAP (OLAP Relacional). Siendo la tecnología relacional la más utilizada, tal como lo menciona Ralph Kimball, quien propuso un esquema llamado esquema estrella para representar las diferentes estructuras multidimensionales haciendo uso del modelo relacional (Mazón, Pardillo y Trujillo 2011, 34). Actualmente existen tres esquemas, que se mencionan a continuación:

A. ESQUEMA ESTRELLA

El esquema estrella forma un diagrama en forma de estrella, teniendo en el centro de la estrella una o más tablas de hechos y las puntas de las estrellas a las tablas de dimensiones (Mazón, Pardillo y Trujillo 2011), como se muestra en la figura Nº 04.

Figura Nº 04. Esquema Estrella

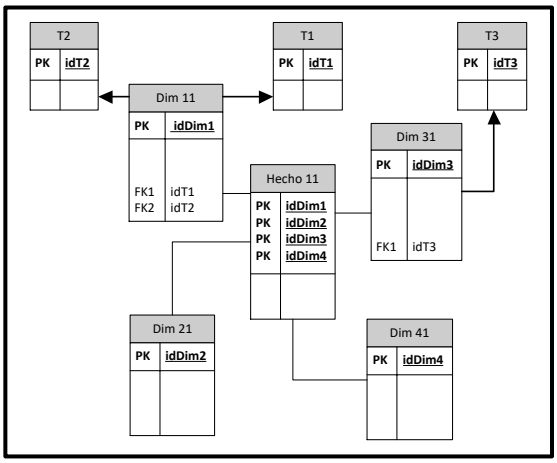


Fuente: *Elaboración Propia*

B. ESQUEMA DE COPO DE NIEVE

En el caso del esquema de copo de nieve, las tablas de dimensiones se encuentran normalizadas, es decir, cada tabla dimensional sólo contiene el nivel que es la clave primaria en la tabla y la llave foránea de su parentesco del nivel más cercano (Mazón, Pardillo y Trujillo 2011), como se muestra en la figura Nº 05.

Figura Nº 05. Esquema Copo de Nieve

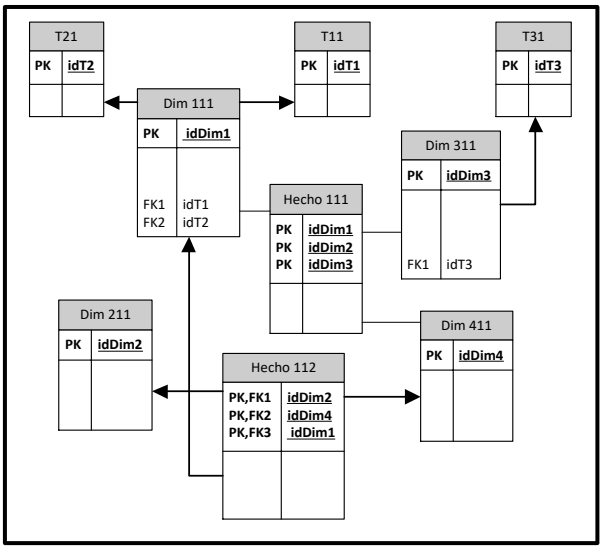


Fuente: Elaboración Propia

C. ESQUEMA DE CONSTELACIONES

En este caso es una variante del esquema estrella. En el que la diferencia radica en el que el esquema de constelaciones de hechos, además de la tabla de hechos base, se define una tabla de hechos por cada nivel de agregación. Por lo tanto, en realidad contiene más de una tabla de hechos (Mazón, Pardillo y Trujillo 2011), como se muestra en la figura Nº 06.

Figura Nº 06. Esquema Constelación de Hechos



Fuente: *Elaboración Propia*

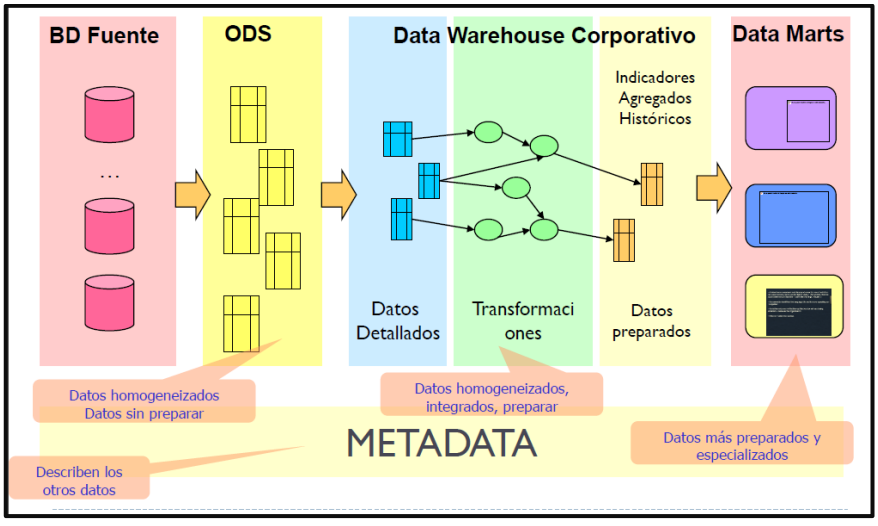
2.7. HERRAMIENTAS ETL

Las herramientas ETL (de las siglas en inglés Extraction, Transformation, Load) se encargan de las funciones de extracción de distintas fuentes de datos, sean estas transaccionales o externas, transformación, realizando tareas de limpieza y consolidación de datos y la carga del Data Warehouse o Data Mart.

Entre las principales funciones de los sistemas ETL tenemos:

- La extracción de datos.
- El filtrado de datos.
- La carga inicial al Data Warehouse o Data Mart.
- Refresco del Data Warehouse o Data Mart: operación periódica que actualiza los cambios de las fuentes externas al Data Warehouse o Data Mart.

Figura Nº 07. Proceso ETL



Fuente: *Elaboración Propia*

2.8. INGENIERÍA DE SOFTWARE

La Ingeniería de software es el establecimiento y uso de principios sólidos de la ingeniería para obtener económicamente un software confiable y que funcione de modo eficiente en maquina reales (Pressman 2006).

Es una disciplina que comprende todos los aspectos de la producción de software desde las etapas iniciales de la especificación del sistema, hasta el mantenimiento de éste después de que se utiliza (Sommerville 2005).

La ingeniería del software es una tecnología estratificada. Conformada por:

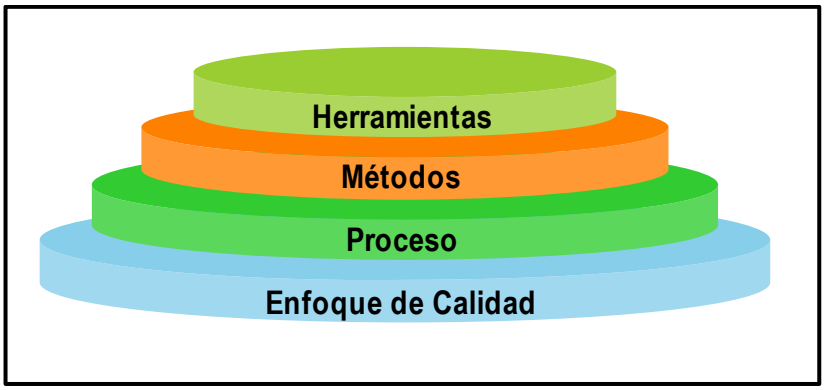
- **Los métodos de la ingeniería del software** indican cómo construir técnicamente el software. Abarcan un amplio espectro de tareas en las que se incluyen: planificación y estimación de proyectos, análisis de los requisitos del sistema y del software, diseño de estructuras de datos, arquitectura de programas y procedimientos algorítmicos, codificación, prueba, mantenimiento, documentación. En estos métodos pueden incluirse técnicas orientadas a objetos, estructuradas, de descomposición funcional, por mencionar algunas.

- **Las herramientas de la ingeniería del software** suministran un soporte automático o semiautomático para los métodos. Hoy en día existen herramientas para soportar cada uno de los métodos mencionados anteriormente. Estas herramientas son llamadas herramientas asistidas por computadora para la ingeniería de software CASE, (Computer Assisted Software Engineering).

- **Los procedimientos de la ingeniería de software** son el pegamento que junta los métodos y las herramientas y facilita el desarrollo racional del software de computadora. Los procedimientos definen la secuencia en la

que se aplican los métodos, las entregas (documentos, informes, etc.) que se requieren, los controles que ayudan a evaluar el progreso del desarrollo. Estos procedimientos, en combinación con las herramientas, pueden facilitar el primer objetivo específico que enunciamos, el cual es facilitar la comunicación.

Figura Nº 08. Ingeniería de software



Fuente: *Pressman, R. S. (2006). Ingeniería Del Software. Un enfoque práctico.*

Todos ellos basado en normas y modelos de calidad de software que tratan de evaluar y medir cómo se presenta dicha calidad en un producto, como, por ejemplo, los que emanan organizaciones como la ISO (International Standard Organization), o IEEE (Institute of Electrical and Electronic Engineers.

2.8.1. PROCESO DE INGENIERÍA DE DATA WAREHOUSE

El proceso de ingeniería de data Warehouse, también conociodo por sus siglas en inglés como DWEP (Data Warehouse Engineering Process), propuesto por Sergio Lujan-Mora y Juan Trujillo, la cual es una adaptación de la metodología RUP (Rational Unified Process) y de la herramienta UML (Unified Modeling Language) para el desarrollo de Data Warehouse y Data Mart. La cual se enfoca en cuatro fases y siete disciplinas que se detallan a continuación:

2.8.1.1. FASES

El proceso DWEP al igual que RUP divide el desarrollo en 4 fases: Concepción, elaboración, construcción y transición. Sin embargo, el DWEP presenta algunas variantes que se adapta al desarrollo de un Data Mart. (Luján Mora y Trujillo 2006)

1. FASE DE LA CONCEPCIÓN

En la fase de concepción se abarca desde la captura de las principales necesidades de los usuarios finales, poniéndose énfasis en los requerimientos funcionales y termina con la elaboración del documento de análisis de requerimientos.

2. FASE DE ELABORACIÓN

En la fase de elaboración se empieza desde el plan de proyecto, el cual contiene la secuencia de actividades a realizar. Además, se define la arquitectura del Data Mart y acaba con un esquema lógico del Data Mart.

3. FASE DE CONSTRUCCIÓN

En la fase de construcción, se implementa el Data Mart hasta obtener la primera versión operativa con datos reales. Además, se desarrollan los procesos ETL necesarios para la carga de datos.

4. FASE DE TRANSICIÓN

Por último, en la fase de transición, se pone énfasis en la detección de errores y empieza cuando el Data Mart entra en producción. Para el presente proyecto esta fase no

se llevará a cabo pues el alcance del proyecto termina con la construcción del Data Mart, es decir, en la fase de construcción.

2.8.1.2. DISCIPLINAS

Además de las mismas fases de RUP, el DWEP también propone disciplinas, pero le añade dos más que son de importancia en un proyecto de Data Mart. Sin embargo, estas disciplinas adicionales se llevan a cabo cuando el usuario ya tiene tiempo usando el Data Mart.

1. REQUERIMIENTOS

Se enfoca en las necesidades del usuario final porque los Data Mart suelen ser únicos para cada tipo de empresa. Durante esta disciplina el usuario debe especificar qué tipo de análisis y agregaciones, le serán de utilidad para generar reportes y tableros de control que le ayuden en la toma de decisiones. La captura de requerimientos se hará mediante los diagramas de caso de uso, los cuales permiten representar lo que el usuario quiere que haga el Data Mart.

2. ANÁLISIS

Empieza por definir y analizar claramente la especificación de los casos de uso encontrados en la etapa anterior. Además, se analiza los datos que servirán como fuente para el Data Mart. Se emplearán diagramas de esquemas para modelar los datos de carga y se realizan los esbozos para los procesos de ETL.

3. DISEÑO

Se define la estructura del Data Mart tanto al nivel lógico como físico. Además, se realiza un diagrama de mapeo de datos, el cual muestra la relación de cómo los datos fuentes están relacionados con los datos del modelo multidimensional.

4. IMPLEMENTACIÓN

Durante esta etapa se construye el Data Mart. La estructura física del Data Mart es construida. También, se desarrollan los procesos de extracción, transformación y carga de los datos al Data Mart.

5. PRUEBAS

Se prueba que lo implementado cumpla con las especificaciones del usuario. Se debe elaborar un plan de pruebas que permite analizar los resultados de las pruebas.

6. MANTENIMIENTO

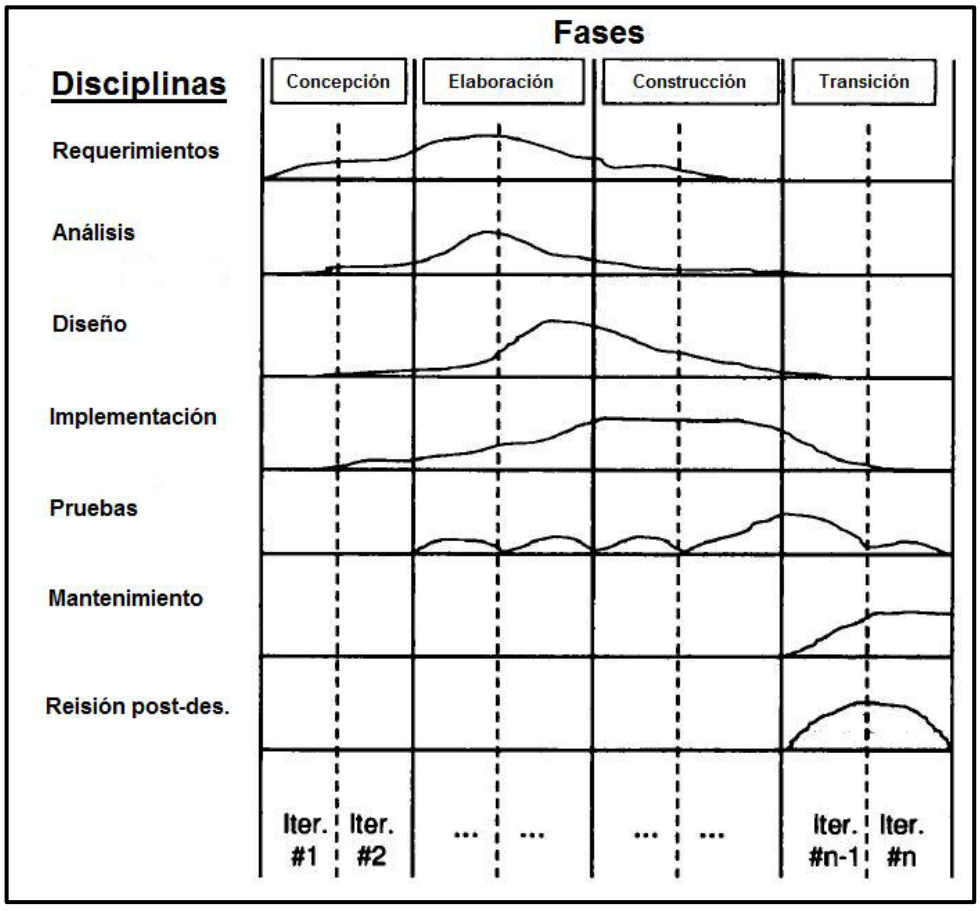
Esta etapa comienza cuando los usuarios ya están usando el Data Mart y acaba cuando el ciclo de vida del Data Mart concluya, pues durante toda su vida útil se le debe de actualizar con la información de la empresa.

7. REVISIÓN POST-DESARROLLO

Se revisa la documentación existente y se busca oportunidades de mejora al Data Mart que puede terminar generando futuros proyectos.

El DWEP plantea que el desarrollo del Data Mart está dividido en pequeños pasos o iteraciones, los cuales son más fáciles de controlar y llevar a cabo. En cada disciplina se elaborarán entregables basados en diagramas o técnicas UML.

Figura Nº 09. Las siete disciplinas propuestas se relacionan con las fases de DWEP



***Fuente**: Luján Mora y Trujillo* (2006)

2.8.1.3. ENTREGABLES

Los entregables dentro del DWEP se distribuyen según la disciplina a la que corresponde, según el estudio de Mora y Trujillo (2006), expresa los entregables mínimos para cada disciplina, tal como se muestra en el cuadro Nº 02.

Cuadro Nº 02.Entregables por disciplinas del DWEP

Disciplinas del DWEP	Entregables
Requerimientos	Diagrama de actividades Diagrama de casos de uso
Análisis	Esquema conceptual de los datos Fuente Esquema lógico de los datos fuente
Diseño	Esquema conceptual del Data Mart Diagrama de mapeo de datos
Implementación	Esquema lógico del Data Mart Esquema físico del Data Mart Proceso ETL
Pruebas	No hay
Mantenimiento	No hay

***Fuente**: Luján Mora y Trujillo* (2006)

Se detalla sintéticamente a continuación cada uno de los entregables (Diagramas) presentado por el modelo de Mora y Trujillo (2006) en el cuadro Nº 02.

1. DIAGRAMA DE ACTIVIDADES:

Muestra en orden la secuencia de actividades y las decisiones a tomar para llevar a cabo un proceso, que puede ser una actividad de negocio.

2. DIAGRAMA DE CASOS DE USO

Muestra los requerimientos funcionales o necesidades de los usuarios.

3. ESQUEMA CONCEPTUAL DE DATOS FUENTE

Muestra a nivel conceptual cómo está organizada la estructura que forma parte de la información fuente para el Data Mart.

4. ESQUEMA LÓGICO DE DATOS FUENTE

Muestra a nivel lógico la estructura de la información fuente para el Data Mart.

5. ESQUEMA CONCEPTUAL DEL DATA MART

Muestra a nivel conceptual el modelo multidimensional del Data Mart. Se observa en la figura Nº 13 que el esquema conceptual del Data Mart consta de 3 niveles. En el nivel 1 define el modelo multidimensional y cada paquete consta de un esquema estrella. En el nivel 2 se especifica cada esquema estrella del nivel anterior y muestra con qué dimensiones se relaciona cada hecho. Por último, en el nivel 3 se especifica cada dimensión u hecho presentando los niveles jerárquicos que existen en cada dimensión.

6. DIAGRAMA DE MAPEO DE DATOS

Muestra la relación entre la estructura de la información fuente del Data.

7. ESQUEMA LÓGICO DEL DATA MART

Muestra a nivel lógico el modelo multidimensional del Data Mart.

8. ESQUEMA FÍSICO DEL DATA MART

Muestra a nivel físico el modelo multidimensional del Data Mart como la configuración de los servidores y cómo está almacenado en los discos.

9. PROCESO ETL

Muestra el proceso de extracción de la información fuente, transformación de la misma y su carga en el Data Mart.

2.8.1.4. ARQUITECTURA DE UN DATA WAREHOUSE

La arquitectura de un Data warehouse y un data Mart es usualmente dividido en varios niveles de datos, en los cuales los datos de un nivel es la derivación o resultado de los datos de un nivel previo. Según Luján Mora y Trujillo (2006) consideran que "the development of a data warehouse can be structured into an integrated framework with five stages and three levels that define different diagrams for the data warehouse model", mostrado en el cuadro Nº 03.

Estados

- **Origen**, en la cual define la estructura operacional de orígenes de datos de la data warehouse, como el de los sistemas de procesos de transaccionales On-line(OLTP), orígenes de datos externos, entre otros.
- **Integracion**, en el cual define el mapa entre los orígenes de datos y el data warehouse.
- **Data warehouse**, la cual define la estructura de la data warehouse o data mart.
- **Adaptacion**, el cual define el proceso detransicion entre la data warehouse y la estructura del cliente.
- **Cliente**, la cual define estructuras especiales que son usadas por los clientes para acceder a la data warehouse, como los data marts o las aplicaciones OLAP.

Niveles

- **Conceptual**, este nivel define la data warehouse desde un punto de vista conceptual.
- **Lógico**, este nivel direcciona aspectos lógicos del diseño de la data warehouse
- **Físico**, a este nivel se define los aspectos físicos de la data warehouse.

Cuadro Nº 03. Diseño de un Data warehouse

	Origen	**Integración**	**Data Warehouse**	**Adaptación**	**Cliente**
Conceptual	**ORIGEN ESQUEMA CONCEPTUAL** Diagrama de clase Estándar UML	**DM** Diagrama de clase Perfil de mapa de datos	**DWCS** Diagrama de clase Estándar UML Perfil Multidimensional	**DM** Diagrama de clase Perfil de mapa de datos	**CCS** Diagrama de clase Estándar UML Perfil Multidimensional
Lógico	**SLS** Diagrama de clase Perfil de modelado de datos diferentes	**PROCESO ETL** Diagrama de clase Perfil ETL	**DWLS** Diagrama de clase Perfil de modelado de datos diferentes	**PROCESO DE EXPORTACION** Diagrama de clase Perfil ETL	**CLS** Diagrama de clase Perfil de modelado de datos diferentes
Físico	**SPS** Diagrama de componentes y Despliegue Perfil de despliegue de la base de datos	**DIAGRAMA DE TRANSPORTATION** Diagrama de componentes y Despliegue Perfil de despliegue de la base de datos	**DWPS** Diagrama de componentes y Despliegue Perfil de despliegue de la base de datos	**DIAGRAMA TRANPORTACION** Diagrama de componentes y Despliegue Perfil de despliegue de la base de datos	**CPS** Diagrama de componentes y Despliegue Perfil de despliegue de la base de datos

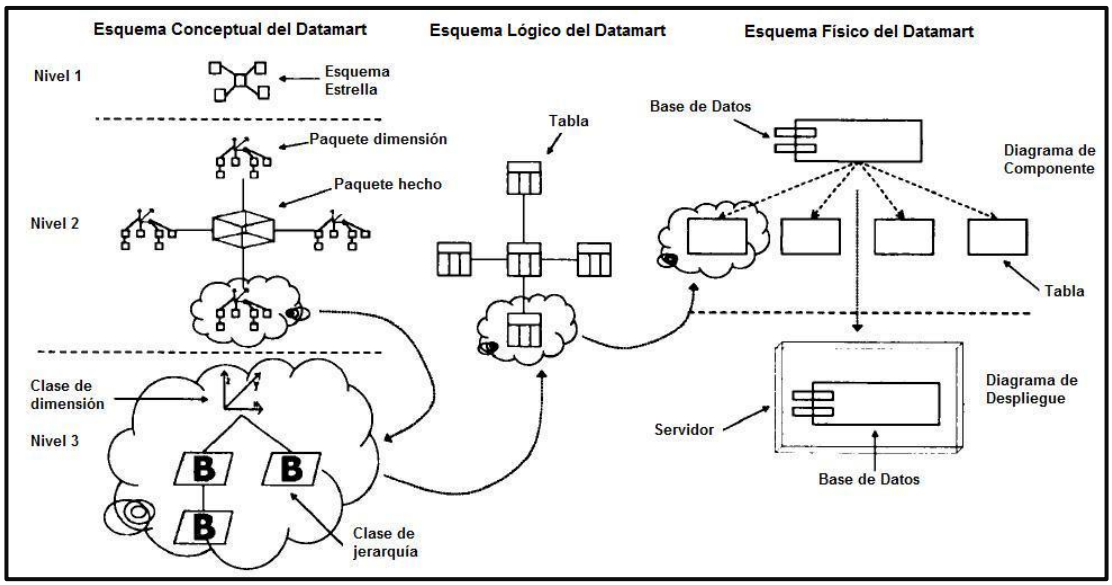
Fuente. *Lujan mora y Trujillo (2006)*

2.8.1.5. ESQUEMAS

La metodología DWEP consta de tres diferentes esquemas: el esquema conceptual, el esquema lógico y el

esquema físico, tal como es mostrado en la figura Nº 10; en cada uno de ellos con sus niveles respectivos.

Figura Nº 10. Los diferentes esquemas del Data Mart que propone la metodología DWEP: esquema conceptual, lógico y físico.



Fuente: *Lujan mora y Trujillo (2006)*

En el desarrollo de un proyecto de almacén de datos es importante la participación de los usuarios finales debido a que algunos entregables del DWEP deben contar con su participación para su elaboración.

La elección de la metodología DWEP como la metodología seleccionada para desarrollar la solución planteada está basada en las siguientes razones:

Primero, es una metodología basada en RUP (Rational Unified Process), es decir, está basada en una metodología ampliamente conocida y sobre todo propone las mejores prácticas en el desarrollo de software. Por ello, es fácil comprender las etapas que comprenden así como sus entregables. De esta manera DWEP adapta la metodología

RUP a los proyectos de desarrollo de almacenes de datos. Además, también se basa en notación UML para elaborar sus propios entregables. El emplear UML tiene las siguientes ventajas: es un lenguaje visual que permite complejidad en el modelamiento, provee flexibilidad y especialización en los diagramas, permite crear estereotipos personalizados, permite modelar dimensiones que serían difíciles con un diagrama de Entidad - Relación.

Segundo, si bien existen una gran variedad de modelos en el diseño de almacenes de datos, no existe una metodología que esté presente durante todo el desarrollo de dichos almacenes. La metodología DWEP brinda una serie de entregables y herramientas desde la etapa de levantamiento de información hasta la construcción del Data Mart cubriendo etapas importantes como el proceso ETL. Así se emplea una metodología que irá guiando al desarrollador durante todas las etapas del proyecto.

2.9. TOMA DE DECISIONES

Cada uno de nosotros toma cientos de decisiones cada día. Este rango de decisiones está dado desde las decisiones intransigentes o irrelevantes hasta las decisiones significativas o relevantes.

La toma de decisiones a nivel individual se caracteriza por el hecho de que una persona haga uso de su razonamiento y pensamiento para elegir una solución a un problema que se le presente en la vida; es decir, si una persona tiene un problema, deberá ser capaz de resolverlo individualmente tomando decisiones con ese específico motivo.

Por su parte Stoner (2003) define la toma de decisiones como “el proceso para identificar y solucionar un curso de acción para resolver un problema específico”.

Según Hellriegel, y Slocum es el “proceso de definición de problemas, recopilación de datos, generación de alternativas y selección de un curso de acción” (Don Hellriegel y Slocum 2004).

Adicionalmente Guillén Rodríguez (2012) define la toma de decisiones como un proceso mediante el cual se realiza una elección entre las opciones o formas para resolver diferentes situaciones de la vida en diferentes contextos (Guillén Rodríguez 2012, 24). Por lo mencionado la toma de decisiones importa la elección de un camino a seguir, por lo que en un estado anterior deben evaluarse alternativas de acción. Si estas últimas no están presentes, no existirá decisión (Guillén Rodríguez 2012).

Tomar una buena decisión, significa que estamos informados y tenemos la información pertinente y apropiada sobre la cual basaremos nuestras decisiones entre diversas alternativas. En algunos casos, apoyamos las decisiones a partir de datos existentes e históricos, mientras que otras veces recopilamos la información, sobre todo para un proceso de elección particular.

La información se presenta en forma de hechos, números, impresiones, gráficos, imágenes y sonidos. Tiene que ser obtenida de diversas fuentes, unido y organizado.

El proceso de organización y de examinación de la información sobre las diversas opciones es conocido como el proceso de modelado. Los modelos se crearon para ayudar a los tomadores de decisiones a entender las ramificaciones de seleccionar una opción. Los modelos pueden variar desde representaciones bastante informales a relaciones matemáticas complejas.

Actualmente en el mundo de los negocios, se utilizan modelos para ayudar a refinar nuestra comprensión de qué y cómo nuestros clientes compran para así mejorar nuestra gestión de relaciones con los clientes.

La calidad de la decisión depende de la idoneidad de la información disponible, la calidad de la información, el número de opciones, y la adecuación de los esfuerzos de modelado disponibles en el momento de la decisión (Sauter 2010). Si bien no es cierto que a más información (o incluso a más análisis) la decisión sería mejor, es cierto que a más del tipo de información requerida (y análisis) es mejor. De hecho, se podría decir que, para mejorar el proceso de elección, tenemos que mejorar los procesos de recolección y análisis de información.

Cada vez más las empresas están tratando de tomar decisiones más informadas para mejorar sus resultados. Algunos se refieren a estos esfuerzos como utilizar la información y los mejores modelos para mejorar la toma de decisiones, tales como la inteligencia de negocios (Sauter 2010). Siendo el principal objetivo reunir la información y los modelos adecuados para entender lo que está pasando en el negocio y para considerar los problemas desde múltiples perspectivas con el fin de proporcionar la mejor orientación para la toma de decisiones.

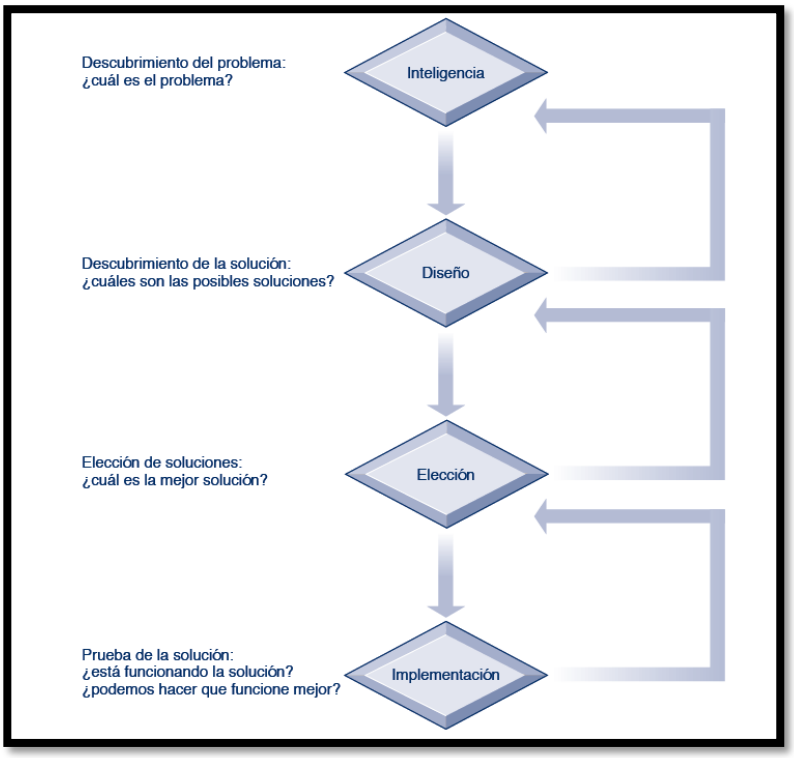
Una forma de lograr el objetivo de reunir la información y los modelos adecuados para la toma de decisión informada, es utilizar los sistemas de soporte de decisiones (DSS). Sistemas de apoyo a las decisiones son sistemas informatizados ASED que reúnen información de una variedad de fuentes, ayudar en la organización y análisis de la información, y facilitar la evaluación de las hipótesis que sustentan el uso de modelos específicos. En otras palabras, estos sistemas permiten a los tomadores de decisiones tengan acceso a los datos pertinentes a través de la organización, ya que se necesitan para tomar decisiones entre alternativas.

2.9.1. EL PROCESO DE TOMA DE DECISIONES

Para tomar una decisión existe un proceso que Simon (1960) describió en cuatro distintas etapas en la toma de decisiones: inteligencia,

diseño, elección e implementación; dicho proceso es ilustrado en la figura Nº 11 - proceso de toma de decisiones.

Figura Nº 11. Proceso de Toma de decisiones.



Fuente: *Simon* (1960)

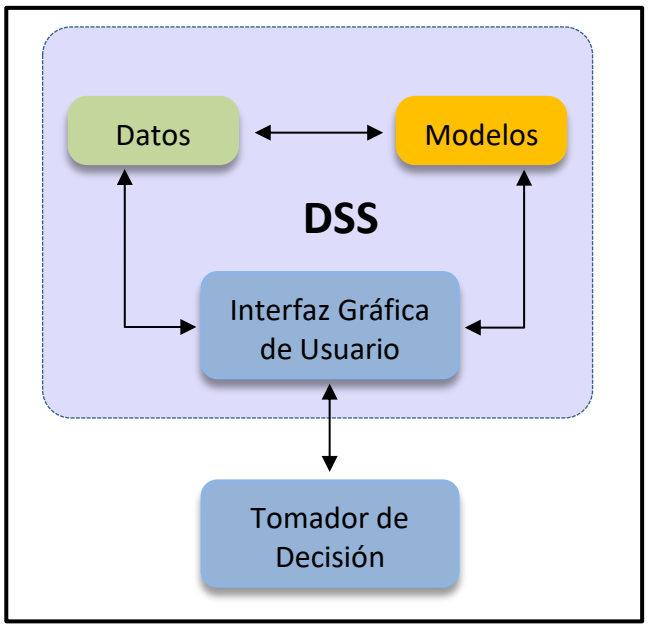
Según Simon(1960) acerca del proceso de la toma de decisiones, menciona lo siguiente: La inteligencia consiste en descubrir, identificar y comprender los problemas que ocurren en la organización: por qué existe un problema, en dónde y qué efectos tiene sobre la firma, El diseño implica identificar y explorar varias soluciones para el problema. La elección consiste en elegir una de varias alternativas de solución. La implementación implica hacer que funcione la alternativa elegida y continuar monitoreando qué tan bien funciona esa solución.

2.9.2. SISTEMAS DE SOPORTE A LA TOMA DE DECISIONES

Un sistema de soporte a la toma de decisiones, en inglés Decision Support System (DSS), ha sido definida como una aplicación interactiva basada en el computador, que combina datos y modelos matemáticos para ayudar a los tomadores de decisiones a resolver problemas complejos a los cuales se enfrentan en la gestión de las empresas y organizaciones públicas y privadas (Vercellis 2009).

Esta definición implica la tres elementos principales de DSS, las cuales son mostradas en la figura Nº 12 Datos, un repositorio de modelos matemáticos y un módulo para el manejo del interfaz gráfico de usuario, Tomador de decisiones, DSS sistema y los usuarios. De este modo, destaca el papel de DSS como el punto focal de tendencias de evolución en dos áreas distintas: por un lado, procesamiento de datos y tecnologías de la información; y por otro lado, las disciplinas abordan el estudio de los modelos y métodos matemáticos, tales como la investigación de operaciones y estadísticas.

Figura Nº 12. Estructura de los sistemas de soporte a la toma de decisiones.



***Fuente**: Elaboración Propia*

Las herramientas de análisis proporcionadas por una arquitectura de inteligencia de negocios pueden ser considerados como DSS capaces de transformar datos en información y conocimiento útil para los tomadores de decisiones (Vercellis 2009). En este sentido, DSS son un componente básico en el desarrollo de una arquitectura de inteligencia de negocios.

Los DSS permiten los tomadores de decisiones para analizar los datos generados a partir de sistemas de procesamiento de transacciones y otras fuentes de información interna fácilmente. Además, el DSS permitir el acceso a la información externa de la organización. Por último, el DSS permiten los tomadores de decisiones la capacidad de analizar la información de una manera que sea útil en una decisión en particular y proporcionará el apoyo de forma interactiva (Sauter 2010).

Por lo tanto, la disponibilidad de DSS ofrece la oportunidad de mejorar el recojo de datos y analiza los procesos asociados a la toma de decisiones. Tomando la lógica un paso más allá, la disponibilidad de DSS ofrece la oportunidad de mejorar la calidad y la capacidad de respuesta de la toma de decisiones y por lo tanto la oportunidad de mejorar la gestión de las empresas.

Dicho de otra manera, el DSS ofrece los tomadores de decisiones la posibilidad de explorar la inteligencia de negocio de una manera eficaz y oportuna (Sauter 2010).

2.9.3. HERRAMIENTAS DE TOMA DE DECISIONES

A. DATOS

Los datos representan una codificación estructurada de entidades primarias individuales, así como de las transacciones que involucran dos o más entidades primarias (Vercellis 2009).

B. INFORMACIÓN

La información es el resultado de las actividades de extracción y procesamiento llevadas a cabo en los datos, y parece significativo para los que lo reciben en un dominio específico (Vercellis 2009).

C. CONOCIMIENTO

La información se transforma en conocimiento cuando se utiliza para tomar decisiones y desarrollar las acciones correspondientes. Por lo tanto, podemos pensar en el conocimiento como un conjunto de información puesto a trabajar en un dominio específico, reforzada por la experiencia y la competencia de los tomadores de decisiones para abordar y resolver problemas complejos (Vercellis 2009).

2.9.4. GESTIÓN ACADÉMICA

La Gestión Académica se encarga de las tareas relativas a las cuestiones académicas relacionadas con los docentes y estudiantes, así como también las tareas de soporte que componen a una determinada institución.

La gestión académica se sustenta en la integración de sus componentes: la formación profesional, la investigación, extensión y proyección social, así como la cooperación nacional e internacional.

En el caso del presente proyecto se delimitara Comportamiento académico de los estudiantes, la carga académica, entre otros.

Se espera que el Data Mart responda a las consultas como:

- ¿Cuántos alumnos de la escuela profesional de Ingeniería de Sistemas e Informática, aprobaron el curso de ingeniería de software, comparativamente entre el semestre académico 2013 – I y 2013 – II?
- ¿Cuántos alumnos aprobaron el curso de “Estructura de Datos” en la escuela de Ingeniería de Sistemas con respecto a los alumnos de la escuela de Estadística e Informática en el Semestre 2013-I?
- ¿Cuántos alumnos desaprobaron la asignara de “Ingeniería de Software” con respecto en los últimos 5 semestres?
- ¿Cuántos alumnos de las diversas carreras de Facultad de Ciencias aprobaron en el semestre 2013-I?
- ¿Cuántos alumnos de las diversas carreras de Facultad de Ciencias desaprobaron en el semestre 2013-I?
- ¿Cuántos alumnos de las diversas carreras de Facultad de Ciencias aprobaron en el semestre 2013-I con respecto al semestre 2012-II?
- ¿Cuántos alumnos de la escuela de Ingeniería de Sistemas se encuentran en tercio superior en el semestre 2013-I?

- ¿Cuántos alumnos de la escuela de Ingeniería de Sistemas e Informática, se encuentran en el tercio superior en el Semestre 2013-I con respecto al semestre 2012-II?

- ¿Cuántos alumnos aprobaron por cursos de la escuela de Ingeniería de Sistemas e Informática en el semestre 2013-I con respecto al semestre 2012-II?

- ¿Cuántos alumnos de la escuela de Estadística aprobaron frente horario del curso de Estructura de Datos en el semestre 2012-II con respecto al horario del semestre 2013-I?

2.10. DEFINICIÓN DE TÉRMINOS

- **DATA WAREHOUSE**

Un Data Warehouse es un repositorio de datos disponibles para desarrollar arquitecturas de inteligencia de negocios y sistemas de apoyo a la decisión. (Vercellis 2009)

Bill Inmon define a un Data Warehouse como “un tema orientado, integrado, variante en el tiempo, colección de datos no volátil para el soporte de las decisiones de la gerencia”. (Inmon 2005)

- **DATA MART**

Una data mart es un data warehouse pero para un subconjunto pequeño de elementos de información. Estos normalmente comprenden un conjunto pequeño o una sola área, de datos diseñados para un grupo especial de usuarios (Westerman 2001)

- **ER Entity-Relationship**

  El modelo Entidad Relación fue originalmente propuesto por Peter Chen en 1976 como una forma de unificar la red y vistas de bases de datos relacionales.

  El modelo Entidad Relación es un modelo de datos conceptual que considera el verdadero mundo como entidades (con atributos) y relaciones.

- **ESQUEMA ESTRELLA**

  Un esquema estrella consiste en una tabla central(tabla hechos) y un conjunto de tablas de dimensiones (Mazón, Pardillo y Trujillo 2011).

  Esquema

- **ETL**

  Proceso de Extracción, transformación y carga, encargada del de extraer datos de Fuentes externas (comúnmente de bases de datos) para proceder luego a la transformación y carga de datos procesados. (Luján Mora y Trujillo 2006)

- **OLAP**

  Una categoría de herramientas de software que provee análisis de almacenes de datos (Luján Mora y Trujillo 2006).

- **MOLAP**

  Tecnología multidimensional que utiliza vectores o matrices multidimensionales para representar las diferentes estructuras multidimensionales (Mazón, Pardillo y Trujillo 2011).

- **ROLAP**

  OLAP Relacional, Tecnología relacional que utiliza elementos relacionales como tablas, columnas, claves primarias, ajenas, etc., para poder definir las diferentes estructuras multidimensionales (Mazón, Pardillo y Trujillo 2011).

- **RUP**

  Siglas en inglés Rational Unified Process que significan una metodología de proceso de desarrollo de software para sistemas orientados a objetos (Luján Mora y Trujillo 2006).

- **SQL**

  Un lenguaje de consulta estándar (standardized query language) para solicitar información a una base de datos. Existen diferentes dialectos de SQL porque cada vendedor de DBMS define sus extensiones (Luján Mora y Trujillo 2006).

- **UML**

  UML cuyas siglas en inglés son Unified Modeling Language Lenguaje de modelado unificado. De acuerdo a las especificaciones de UML, Lenguaje de modelado unificado es un lenguaje gráfico de visualización, específico, construido y documentado para la construcción de un software (Luján Mora y Trujillo 2006).

- **UNIFIED PROCESS**

  El proceso de desarrollo de software unificado, también conocido como UP, es un estándar de procesos de ingeniería de software industrial de los autores de UML (Luján Mora y Trujillo 2006).

- **SISTEMA DE SOPORTE DE DECISIONES**

  Sistema de información interactivo que ayuda a las personas encargadas de las tomas de decisiones a utilizar datos y modelos para resolver problemas (Luján Mora y Trujillo 2006).

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. TIPO DE ESTUDIO

Este trabajo de investigación es de tipo Aplicada, que está orientada a un problema pragmático especifico.

El tipo de investigación es Explicativa, según la profundidad, debido a que se centra en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta (Hernández Sampieri, Fernández Collado, & Baptista Lucio 2010). Teniendo en cuenta lo mencionado, se manipula la variable independiente, el Data Mart, para observar su efecto y su impacto sobre la variable dependiente, la toma de decisiones de la gestión Académica de la Facultad de Ciencias.

De acuerdo a la orientación la investigación es Aplicada, porque la presente investigación está orientada al desarrollo de un Data Mart, destinado al soporte de la Toma de Decisiones de la Facultad de Ciencias de la UNASAM.

3.2. DISEÑO DE ESTUDIO

Diseño que se empleó en el presente proyecto de investigación es Experimental con post prueba únicamente y grupo control. (Hernández Sampieri, Fernández Collado y Baptista Lucio 2010)

Cuyo diagrama de muestra a continuación:

G1 X O1

G2 — O2

G1 = Grupo-Experimental conformado por los directivos de la Escuela Académico Profesional de Ingeniería de Sistemas e Informática.

G2 = Grupo-Control conformado por los directivos de la Escuela Académico Profesional de Estadística e Informática, y Matemática.

X = Variable Independiente o manipulable, en este caso es la implementación del Data Marts.

01 = post prueba al grupo Experimental.

02 = post Prueba al grupo Control.

3.3. POBLACIÓN Y MUESTRA

3.3.1. POBLACIÓN

La población de estudio del presente proyecto de investigación estuvo conformada por la Facultad de Ciencias, la cual está compuesta por las distintas escuelas profesionales, y representadas por la plana jerárquica:

Cuadro Nº 04. Población jerárquica**.**

Población		
Código	**Personal**	**Número**
03.01	Decano de la Facultad de Ciencias	01
03.02	Jefe de departamentos Académicos	03
03.03	Director de Escuela de la escuela profesional de Ingeniería de Sistemas e Informática.	01
03.04	Director de Escuela de la escuela profesional de Estadística e Informática.	01
03.05	Director de Escuela de la escuela profesional de Matemática.	01

***Fuente**: Elaboración Propia*

3.3.2. MUESTRA

La muestra de estudio para el presente proyecto estuvo conformada por la escuela de Ingeniería de Sistemas e Informática para el Grupo experimental, y la escuela de Estadística e Informática y Matemática para el Grupo control.

3.3.3. MUESTREO

Para el presente proyecto de investigación se realizó el muestreo probabilístico estratificado con afijación proporcional, procedimiento en cual el investigador compara sus resultados entre segmentos, grupos o

nichos de la población. En la cual la distribución se hace de acuerdo con el tamaño de la población en cada estrato (Vivanco 2005).

3.4. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

La técnica consiste en visualizar o captar en forma sistemática el hecho o fenómeno que se produce en su naturaleza en función a los objetivos de su investigación preestablecidos (Arias 2011). La técnica utiliza para la recolección de datos fue la observación, y su instrumento fue la escala de estimación.

Además, se utilizó la técnica de fichaje, es una técnica que sirvió para la organización del marco teórico que viene hacer el soporte de la investigación, las mismas que se registraron en fichas textuales, fichas de resúmenes, fichas bibliográficas, entre otras.

Como instrumento se utilizó la escala de estimación que consta de 41 ítem estructurados en base a las variables, dimensiones e Indicadores relacionados a la toma de decisiones de la gestión académica de la Facultad de ciencias de la Universidad nacional Santiago Antúnez de Mayolo, cuya valoración fue dada por la escala Likert que viene a ser: Siempre (3), casi siempre (2), a veces (1) y Nunca (0).

Para la elaboración del instrumento, escala de estimación, se realizó la operacionalización de las variables con la finalidad de especificar los indicadores e ítems del cual se estructuro en el instrumento, organizando los ítems en base a las dimensiones de la variable dependiente, la toma de decisiones de la gestión académica.

3.5. TÉCNICAS DE PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN

Las técnicas de procesamiento y análisis de datos utilizados en el presente fueron las siguientes: Se utilizó las medidas de tendencia central (media, mediana y moda.), para el análisis e interpretación de datos.

Además, se utilizó la estadística inferencial, la prueba de hipótesis con t-student para la contrastación de la hipótesis.

3.6. MATERIALES

Los materiales que fueron utilizados durante el desarrollo del presente proyecto de investigación están distribuidos en dos categorías (cuadro Nº 05, cuadro Nº 06), especificando la cantidad utilizada y su unidad respectiva.

Cuadro Nº 05. Recursos materiales

1.1 MATERIALES			
Ítem	**Descripción**	**Unidad**	**Cantidad**
			Software
01.01.01	Sistema Operativo Microsoft Windows 8	Unidad	01
01.01.02	SPSS Statistics	Unidad	01
01.01.03	Microsoft SQL Server	Unidad	01
01.01.04	Php	Unidad	01
01.01.05	Rational rose	Unidad	01
01.01.06	Microsoft Visio	Unidad	01
01.01.07	Edraw Max	Unidad	01
01.01.08	Sublime text 3.0	Unidad	01
01.01.09	Laravel 4.2	Unidad	01
01.01.10	Visual Studio Business Intelligence	Unidad	01
			Otros
01.02.01	Papel Bond A4 de 75 Gr	Millares	01
01.02.02	Papel Bond A4 de 80 Gr	Millares	02
01.02.03	Fólder de manila A4	Unidad	12
01.01.04	Perforador	Unidad	01
01.02.05	Sobres manila tamaño A4	Unidad	15
01.02.06	Sobres manila tamaño Carta	Unidad	15
01.02.07	Memoria USB de 8 Gb	Unidad	02
01.02.08	CDs	Unidad	15
01.02.09	DVDs	Unidad	15
01.02.10	Clips	Caja	01
01.02.11	Cinta scotch	Unidad	05
01.02.12	Engrampadora	Unidad	01
01.02.13	Saca Grapas	Unidad	01

Ítem	Descripción	Unidad	Cantidad
			Otros
01.02.14	Lapiceros tinta seca	Unidad	04
01.02.15	Lapiceros tinta liquida	Unidad	03
01.02.16	Corrector	Unidad	01

Fuente: *Elaboración Propia*

En el cuadro Nº 06, se enfoca en todos los Equipos que fueron utilizados durante el desarrollo.

Cuadro Nº 06. Recursos en equipos

1.2 EQUIPOS			
Ítem	**Descripción**	**Unidad**	**Cantidad**
02.01	Computadora de escritorio: procesador Intel Core 2 Duo de 2.66 Ghz, memoria Ram DDR2 800 Kingston 4 Gb..	Unidad	01
02.02	Computadora portátil: procesador Intel Core i7 de 2.4 Ghz, memoria Ram DDR2 800 Kingston 8 Gb.	Unidad	01
02.03	Servidor	Unidad	01
02.04	Impresora.	Unidad	01

Fuente: *Elaboración Propia*

3.7. MÉTODOS

3.7.1. MÉTODO GENERAL

El método general que se aplicó en la presente investigación fue el método experimental que consiste en la manipulación de variable independiente, el Data Mart, para luego evaluar el impacto o el efecto que tuvo este sobre la variable dependiente, toma de decisiones de la gestión académica de la Facultad de ciencias.

3.7.2. MÉTODOS ESPECÍFICOS:

Además, entre los métodos específicos aplicados que fueron utilizados se encuentran:

El método Inductivo permitió estudiar el fenómeno de lo conocido a lo desconocido.

El método Deductivo, que es un razonamiento que consiste en tomar conclusiones generales para explicaciones particulares. El método se inicia con el análisis de los postulados, teoremas, leyes, principios, etc.; de aplicación universal y de comprobada validez, para aplicarlo a soluciones o hechos particulares. (Hernández Sampieri, Fernández Collado y Baptista Lucio 2010), que fue utilizado para contrastar las hipótesis planteadas.

El método Analítico, cuya aplicación consistió en clasificar, agrupar, comparar y diferenciar los datos de acuerdo al estudio estadístico que se desarrolló.

El método sintético, que nos permitió resumir los datos estadísticos en cuadros para realizar el posterior análisis respectivo.

3.8. OPERACIONALIZACIÓN DE VARIABLES

La operacionalización de las variables fue realizada teniendo en cuenta las variables, sus dimensiones, indicadores para la elaboración de los ítems que posteriormente fueron utilizados como insumo para el instrumento de medición. En el cuadro Nº 07 se muestra la operacionalización de variables, la cual es ampliada en el anexo Nº 05.

Cuadro Nº 07. Operacionalización de Variables

Variables	Dimensiones	Indicadores	Ítems en la Escala de Estimación
Independiente Data Mart	Integrado	Información integrada	**1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23**
		Estructura consistente	
		Niveles de detalle	
		Necesidades del Usuario	
	Temático	Datos organizados	
		Acceso fácil	
		Entendimiento del usuario final	
	No volátil	Permite el análisis	
		Toma de decisiones	
		Base de datos estable	
	Variante en el tiempo	Predice tendencias	
		Se actualiza en un periodo de tiempo	
Dependiente Toma de Decisiones de la gestión Académica	Diseño	Identifica los planes de estudio	**24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41**
		Identifica a los docentes	
		Explora a los estudiantes	
	Elección	Selecciona los datos	
		Analiza los datos	
		Elige una alternativa	

***Fuente**: Elaboración Propia*

3.9. PROCEDIMIENTO

Para la realización de la presente investigación se utilizó el diseño Experimental con post prueba para la contrastación de la hipótesis, detallado en la sección 3.2 del presente. Además de ello se utilizó la metodología DWEP, para el desarrollo del Data Mart, la cual integra modelos de casos de uso y los componentes a utilizados.

3.9.1. IDENTIFICACIÓN DE REQUERIMIENTOS

La identificación de los requerimientos es un proceso clave dentro de la construcción de un Data Mart, para ello se llevó a cabo el proceso de identificación y captura de requerimientos para lo cual se utilizó la entrevista con los stakeholders, como una técnica de recopilación de datos. Además, se utilizaron otras fuentes de datos como

son: Documentos, publicaciones, guías entre otros. Para el análisis de datos de los datos obtenidos se utilizaron las hojas de cálculo de Excel.

Después de la aplicación de la técnica de recopilación de datos, se logró identificar los requerimientos funcionales y no funcionales de que debe de cumplir la aplicación del Data Mart.

3.9.2. DISEÑO DEL DATA MART

Para el diseño de un Data Mart se utilizó la metodología DWEP, definida en la sección 2.8.1; la cual nos facilitó la realización de modelo de negocio UML y los modelos perecientes a la metodología utilizada. Ademas se definió los componentes de software, hardware que se utilizaron.

3.9.3. COMPROMETER A LOS STAKEHOLDER

Una vez definida los objetivos y estrategias, se buscó el apoyo de los principales Stakeholders tales como los directores de escuela, el jefe de la oficina general de estudio, los decanos de las diversas Facultades, entre otros.

CAPÍTULO IV

ANÁLISIS

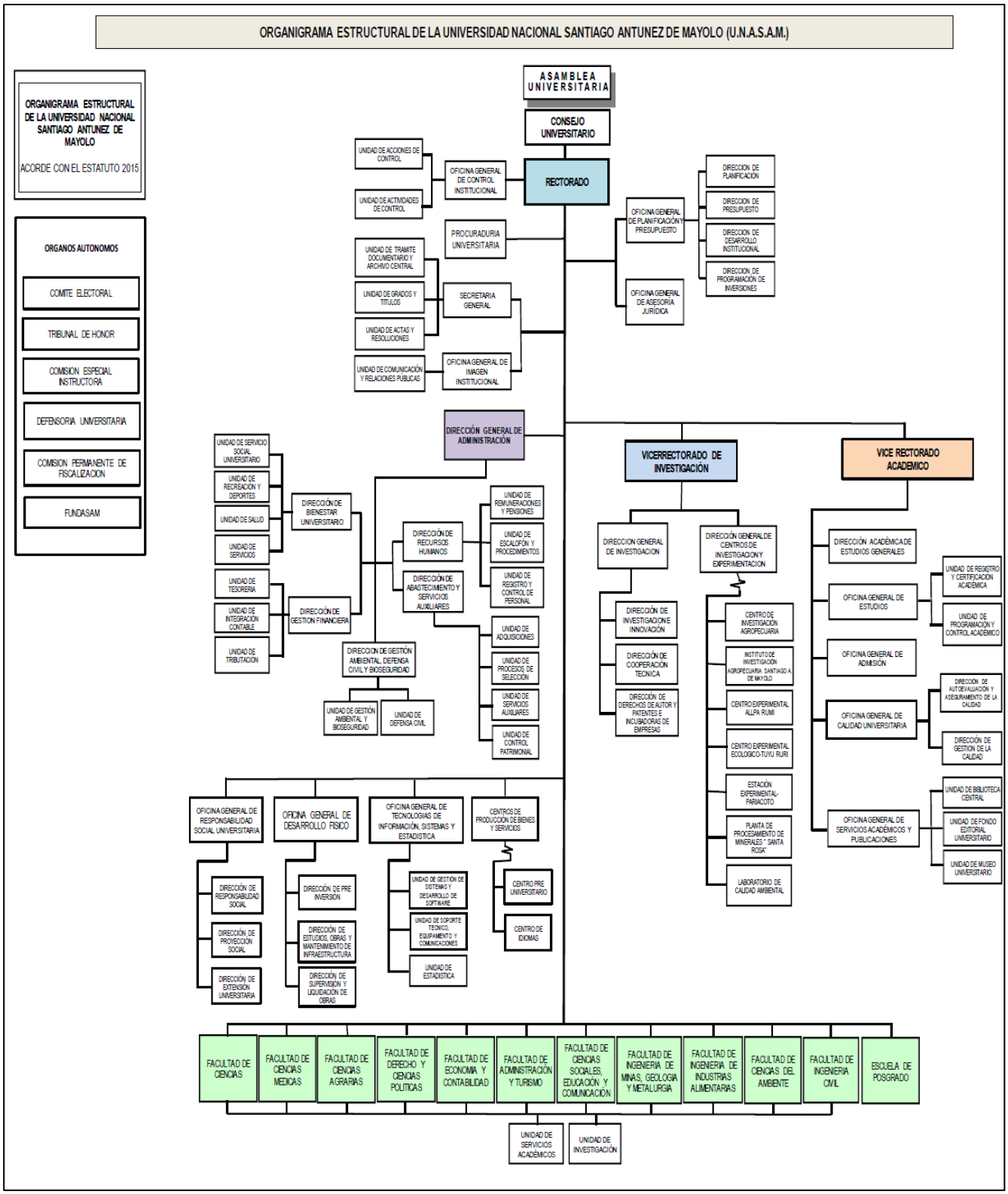
En el presente capítulo se definirá la metodología de la solución a emplear. Luego se identificarán los requerimientos funcionales y no funcionales que debe satisfacer la solución. Finalmente, se realizará el análisis dimensional que satisface a los requerimientos.

4.1. ANÁLISIS DE LA SITUACIÓN ACTUAL

4.1.1. ANÁLISIS DEL ORGANIGRAMA FUNCIONAL ESTRATÉGICO

La estructura orgánica funcional de la Universidad nacional Santiago Antúnez de Mayolo de Ancash, está estructurada en los niveles altos, por el rector, Vicerrector Académico, Vicerrectorado de Investigación, y por l diversas Facultades de quienes el decano es el principal responsable.

Figura Nº 13. Organigrama estructural de la Universidad Nacional Santiago Antúnez de Mayolo



Fuente: *Reglamento de organizaciones y funciones 2016*

Siendo la Facultad de Ciencias un órgano de línea, con unidades de servicios académicos y de investigación conformada por las direcciones de escuelas profesionales de dicha Facultad, tales como la escuela de Estadística, la escuela de Ingeniería de Sistemas e Informática y la escuela de Matemática; además de ello, conformada también por las direcciones de los Departamentos Académicos.

Siendo la Dirección de la Escuela Profesional encargada del diseño y la actualización curricular permanente de una carrera profesional, así como también de dirigir su aplicación para la formación, capacitación pertinente hasta la obtención del grado académico de bachiller y título profesional correspondiente. Dirigida por el director de la escuela profesional, quien es un docente principal. Encargado de resolver en primera instancia los problemas de los estudiantes.

Además, los Departamentos Académicos son unidades de servicio académico que reúnen a docentes de disciplinas afines, con la finalidad de estudiar, investigar y actualizar contenidos curriculares, mejorar estrategias pedagógicas y preparar los sílabos por cursos o materias, a requerimiento de las Escuelas Profesionales, además brindan servicios a las diferentes Escuelas que lo requieran. Los encargados de dirigir son los directores de departamento, es el representante del departamento académico, y la asamblea de docentes, conformada por todos los docentes ordinarios del departamento académico.

4.1.2. EVALUACIÓN DE LA CAPACIDAD INSTALADA

El personal que actualmente labora en la decanatura y las direcciones de escuela de la Facultad de ciencias y los departamentos académicos son:

Cuadro Nº 08. Personal de la Facultad de Ciencias

Código	Área	Personal	Cantidad
01.01	Decanatura	Decano de la Facultad de Ciencias	01
01.02	Decanatura	Secretaria de la Decanatura	01
01.03	Dirección de Escuela de la escuela de Estadística e Informática	Director de la Escuela de Estadística e Informática	01
01.04	Dirección de Escuela de la escuela de Estadística e Informática	Secretaria de la Escuela de Estadística e Informática	01
01.05	Dirección de Escuela de la escuela de Ingeniería de Sistemas e Informática	Director de la Escuela de Ingeniería de Sistemas e Informática	01
01.06	Dirección de Escuela de la escuela de Ingeniería de Sistemas e Informática	Secretaria de la Escuela de Ingeniería de Sistemas e Informática	01
01.07	Dirección de Escuela de la escuela de Matemática	Director de la Escuela de Matemática	01
01.08	Dirección de Escuela de la escuela de Matemática	Secretaria de la Escuela de Matemática	01
01.09	Departamento Académico de Ciencias	Director del departamento Académico de Ciencias	01
01.10	Departamento Académico de Ciencias	Secretaria del departamento Académico de Ciencias	01
01.11	Departamento Académico de Estadística y Sistemas	Director del departamento Académico de Estadística y Sistemas	01
01.12	Departamento Académico de Estadística y Sistemas	Secretaria del departamento Académico de Estadística y Sistemas	01
01.13	Departamento Académico de Sistemas y Telecomunicaciones	Director del departamento Académico de Sistemas y Telecomunicaciones	01
01.14	Departamento Académico de Sistemas y Telecomunicaciones	Secretaria del departamento Académico de Sistemas y Telecomunicaciones	01

Fuente: *Elaboración propia.*

Actualmente la Facultad de ciencias cuenta con los equipos necesarios para el cumplimiento adecuado de las labores académicas, siendo una computadora por cada uno de los cargos que son ejercidos.

4.1.3. ANÁLISIS DE LAS FORTALEZAS, OPORTUNIDADES, DEBILIDADES Y AMENAZAS

Con el fin de identificar y analizar la situación actual de la Facultad de Ciencias, se realizó un análisis FODA (Fortalezas, debilidades, Oportunidades y Amenazas) la cual permitió conocer la situación actual de la Facultad, la cual es mostrada en el cuadro Nº 09. Teniendo como base que el análisis FODA es un instrumento de planificación estratégica que permite la identificación y evaluación de las fortalezas y debilidades de la organización, así como las oportunidades y amenazas que esta presenta.

Cuadro Nº 09. Matriz de fortalezas, Oportunidades, Debilidades y Amenazas

ANÁLISIS FODA		
	FORTALEZAS	**DEBILIDADES**
INTERNAS	- Posee un presupuesto anual. - Cuenta con un personal. - Cuenta con una instalación tecnológica adecuada. - Tiene acceso a los datos del Sistema Integral de Gestión Académica.	- No todo el personal está capacitado en el uso de tecnologías de la información y la comunicación. - Los usuarios del SIGA solo pueden realizar consultas ya elaboradas de forma genérica y no especifica limitando así el análisis más detallado de los datos. - La toma de decisiones requiere mayor tiempo de lo programado. - Falta de indicadores de desempeño académico institucional. - Falta de motivación de los recursos humanos.
	OPORTUNIDADES	**AMENAZAS**
EXTERNAS	- Imagen corporativa positivamente valorada en la comunidad regional. - Buscar financiamiento para incrementar el capital de trabajo. - Ampliar los proyectos de desarrollo en la región. - Capacitaciones para el uso de las TIC por parte de la ONGEI (oficina nacional del gobierno electrónico e informático). - Publicidad electrónica utilizando las redes sociales. - Disponibilidad de equipos tecnológicos al alcance de la población.	- Falta de una cultura informática, ya que no son aprovechadas adecuadamente las nuevas TIC, así como el conocimiento sobre sus aplicaciones y la capacitación pertinente para el estudio y transmisión de los contenidos. - Fuerte competencia externa en el plano de toma decisiones efectivas. - Cambios en la legislación. - Limitado acceso al ancho de banda. - Existencia de personas con desconocimiento de tecnológico.

Fuente: *Elaboración propia*

En base al análisis FODA realizado mostrado en el Cuadro Nº 09, se plantearon estrategias de análisis FODA, ver Cuadro Nº 10.

Cuadro Nº 10. Estrategias del análisis FODA

ESTRATEGIAS DEL ANÁLISIS FODA		
EXTERNAS / **INTERNAS**	**OPORTUNIDADES**	**AMENAZAS**
	1. Imagen corporativa positivamente valorada en la comunidad regional. 2. Buscar financiamiento para incrementar el capital de trabajo. 3. Ampliar los proyectos de desarrollo en la región. 4. Capacitaciones para el uso de las TIC por parte de la ONGEI (oficina nacional del gobierno electrónico e informático). 5. Publicidad electrónica utilizando las redes sociales. 6. Disponibilidad de equipos tecnológicos al alcance de la población.	1. Falta de una cultura informática, ya que no son aprovechadas adecuadamente las nuevas TIC, así como el conocimiento sobre sus aplicaciones y la capacitación pertinente para el estudio y transmisión de los contenidos. 2. Fuerte competencia externa en el plano de toma decisiones efectivas. 3. Cambios en la legislación. 4. Limitado acceso al ancho de banda.
FORTALEZAS	**FO**	**FA**
1. Posee un presupuesto anual. 2. Cuenta con un personal. 3. Cuenta con una instalación tecnológica adecuada. 4. Tiene acceso a los datos del Sistema Integral de Gestión Académica.	- Gracias al presupuesto se podrá dar a conocer a la población acerca del trabajo realizado para así fortalecer la imagen. - Con el Presupuesto Anual asignado es posible ampliar los proyectos de desarrollo en la región. - Gracias a la instalación tecnológica con la que se cuenta se podrá seguir realizando las capacitaciones de la ONGEI. - Con el acceso al SIGA y la disponibilidad de equipos se podrá ampliar el número de personas que puedan acceder a ella desde distintos dispositivos electrónicos. - El personal podrá ser capacitado por ONGEI.	- Con el presupuesto y el personal, se podrá mejorar el proceso de toma de decisiones implementando nuevas tecnologías que den soporte a la toma de decisiones. - Con la instalación tecnológica y el presupuesto, se podrá brindar una atención más eficiente a los usuarios que lo requieran. - Con el acceso al SIGA se podrá resolver cuestiones de las personas con bajos niveles de conocimiento informático.
DEBILIDADES	**DO**	**DA**
1. No todo el personal está capacitado en el uso de tecnologías de la información y la comunicación. 2. Los usuarios del SIGA solo pueden realizar consultas ya elaboradas de forma genérica y no especifica limitando así el análisis más detallado de los datos. 3. La toma de decisiones requiere mayor tiempo de lo programado. 4. Falta de indicadores de desempeño académico institucional. 5. Falta de motivación de los recursos humanos.	- Aprovechar las capacitaciones sobre el uso de las TIC realizados por la ONGEI, para capacitar al personal de la Facultad de Ciencias. - Gracias a la disponibilidad de equipos y a la posibilidad de ampliar los proyectos, se podrá realizar un proyecto que permita a los usuarios tomadores de decisiones realizar consultas más elaboradas con el fin de que el proceso de toma de decisiones sea más ágil.	- Optimizar el proceso de toma de decisiones, utilizando los recursos mínimos. - sea en un menor tiempo. - Impulsar el desarrollo de indicadores de desempeño académico institucional; contribuyendo a la mejora de la institución. - Impulsar reconocimientos a la labor, contribuyendo a la mejora del clima organizacional y la motivación del personal.

Fuente: *Elaboración propia.*

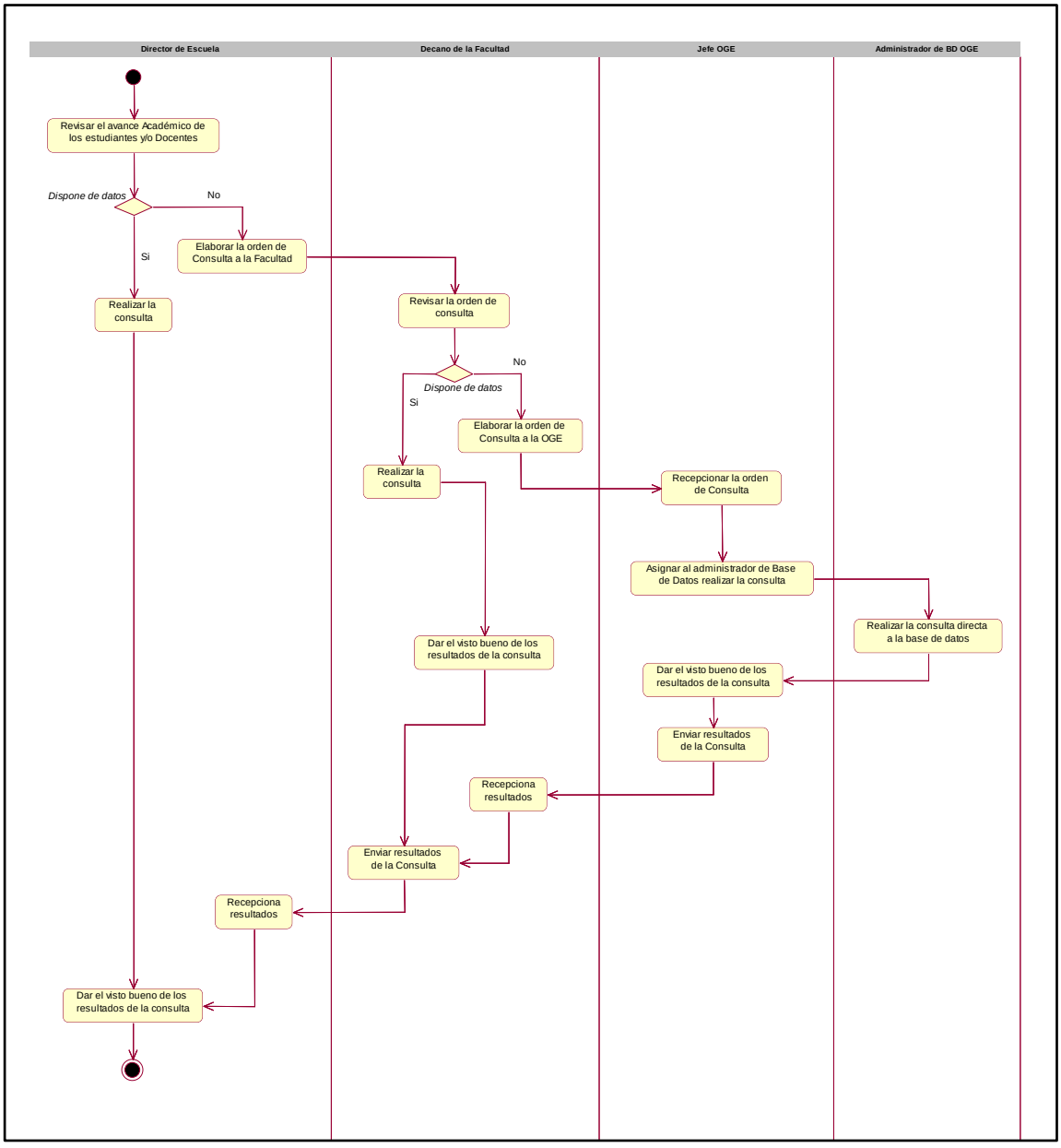
4.1.4. ANÁLISIS DE LA SITUACIÓN ACTUAL

Para comprender mejor los procesos que se realizan en la Facultad de Ciencias en la parte operativa y en la parte consecución de información de la gestión académica para el análisis para la toma de decisiones, se logró identificar lo siguiente: la Facultad de Ciencias (Direcciones de Escuela, Departamentos Académicos) realiza las siguientes acciones de registro de datos de los alumnos ingresantes, registro de matrículas en el Sistema Integral de Gestión Académica (SIGA) de los estudiantes al iniciar un semestre académico, según el cronograma emitido por la Oficina General de Estudios; recepción y envió a Oficina General de Estudios de registro auxiliares y actas de notas al finalizar un semestre, entre otras acciones a nivel opertivo.

Teniendo como base el nivel operativo, cuyos datos son almacenados en la base de datos transaccional y centralizada ITUnasam, administrada directamente por la Oficina General de Estudios de la UNASAM, para la toma de decisiones dentro de la Facultad, se convierte en un proceso tedioso que tiende a consumir tiempo, recursos y personal para absolver una consulta. Adicionalmente, es importante conocer los procesos principales para la obtención de datos dentro del negocio. Por ello, a continuación, se presentarán los procesos de consulta de datos desde la Dirección de Escuela. El Director de Escuela requiere revisar el avance académico de los estudiantes y necesariamente revisa si tiene la disponibilidad de datos, y si dispone de datos suficientes realiza la consulta, y finalmente él da el visto bueno de la consulta. En el caso de no disponibilidad de datos, el director de escuela elabora una orden de consulta a la Decanatura, quienes son los encargados de verificar dicha disponibilidad de datos, en el caso de no disponibilidad es enviado a la Oficina General de Estudios, en cuya oficina el jefe asigna al

administrador de SIGA realizar la consulta. Como es mostrado en el diagrama de actividades del Negocio.

Figura Nº 14. Diagrama de actividades del Negocio



Fuente: *Elaboración propia*

4.2. IDENTIFICACIÓN Y DESCRIPCIÓN DE REQUERIMIENTOS

La identificación y descripción de requerimientos fue el proceso clave para el desarrollo de la herramienta que mejora la toma de decisiones en la Facultad de ciencias, por ello el principal requerimiento fue que el Data mart propuesto brinde un conjunto facilidades que permitan utilizar la información disponible de la gestión académica de la Facultad de ciencias para aportar información clave en bien de mejorar la toma de decisiones. En base a esta premisa general, se desglosa los requerimientos funcionales y no funcionales que debe de cumplir la solución planteada.

4.2.1. REQUERIMIENTOS FUNCIONALES

Los requerimientos funcionales utilizados para el desarrollo de un Data Mart constituyen las necesidades de información de los usuarios finales que en este caso son los tomadores de decisiones de la gestión académica de la Facultad de ciencias, conformado por el Decano, directores de escuela de las diferentes escuelas académicas existentes en la Facultad de ciencias y los directores de departamento académico de Ciencias y de Estadística y Sistemas.

Para los tomadores de decisiones de la Facultad es imprescindible la necesidad de contar con la información histórica acerca de la gestión académica, debido a que la Facultad de Ciencias entro en funcionamiento a partir de 26 de agosto de 1994, cuenta con un gran repositorio de datos de toda la gestión académica desarrollada durante más 20 años aproximadamente. Por lo cual se ha podido obtener algunas de las necesidades de información, tales como:

- **Contar con una herramienta para facilitar la toma de decisiones,** debido a que no se cuenta con ninguna herramienta en la cual se apoye la toma de decisiones.

- **Poder acceder a la información requerida rápidamente:** el tener acceso a la información requerida en el tiempo requerido es vital para la toma de decisiones. Actualmente el Sistema de información de la gestión Académica, sistema transaccional, que utiliza la Universidad Nacional Santiago Antúnez de Mayolo depende de la administración de la oficina general de estudios, generándose un cuello de botella para el acceso a la información y consume mucho tiempo para la elaboración de reportes.

- **Contar con una herramienta fácil de usar**. El Data Mart debe contar una interfaz gráfica que permita una fácil interacción con la misma. De esta manera, para el personal no será difícil aprender a utilizarla y podrá verla como una herramienta útil.

Siendo los requerimientos funcionales identificados son los siguientes:

- Los reportes o tableros de control requeridos para el sistema son:

  - Evolución del rendimiento académico de los estudiantes de la Facultad de ciencias agrupados por modalidad de aprobación, semestre académico, por escuela académica, etc.

  - El sistema debe de mostrar el Ranking de calificaciones de los estudiantes por Facultades agrupados por condición de aprobación del último semestre.

  - Tablero de control que muestre 4 gráficos sobre información consolidada del rendimiento académico: Cantidad de matriculados en el último semestre por Facultad, rendimiento académico por Facultades en todos los semestres, Porcentaje

de aprobados versus desaprobados, Numero de Docentes registrados por Departamento académico.

– El sistema debe generar un reporte estadístico que indique la evolución de los promedios ponderados por Facultad de todos los semestres.

- El sistema debe de administrar perfiles de usuarios, debe de permitir el registro de nuevos usuarios, la actualización de datos y contraseñas.

- El sistema debe responder a las cuestiones del comportamiento académico de los estudiantes, que los tomadores de decisiones requieran, tales como: ¿Cuántos alumnos de la escuela profesional de Ingeniería de Sistemas e Informática, aprobaron el curso de ingeniería de software, comparativamente entre el semestre académico 2013 – I y 2013 – II?, ¿ Cuántos alumnos aprobaron el curso de "Estructura de Datos" en la escuela de Ingeniería de Sistemas con respecto a los alumnos de la escuela de Estadística e Informática en el Semestre 2013-I?¿ Cuántos alumnos desaprobaron la asignara de "Ingeniería de Software" con respecto en los últimos 5 semestres?, ¿Cuántos alumnos de las diversas carreras de Facultad de Ciencias aprobaron en el semestre 2013-I?, ¿Cuántos alumnos de las diversas carreras de Facultad de Ciencias desaprobaron en el semestre 2013-I?, ¿Cuántos alumnos de las diversas carreras de Facultad de Ciencias aprobaron en el semestre 2013-I con respecto al semestre 2012-II?, ¿Cuántos alumnos de la escuela de Ingeniería de Sistemas se encuentran en tercio superior en el semestre 2013-I?, ¿Cuántos alumnos de la escuela de Ingeniería de Sistemas e Informática, se encuentran en el tercio superior en el Semestre 2013-I con respecto al semestre 2012-II?,

¿Cuántos alumnos aprobaron por cursos de la escuela de Ingeniería de Sistemas e Informática en el semestre 2013-I con respecto al semestre 2012-II?, ¿Cuántos alumnos de la escuela de Estadística aprobaron frente horario del curso de Estructura de Datos en el semestre 2012-II con respecto al horario del semestre 2013-I?, entre otras cuestiones.

- El sistema debe responder a las cuestiones de los tomadores de decisiones acerca del comportamiento académico de los docentes, tales como: ¿Cuántos docentes que dictan el curso de Física I pertenecen al Departamento Académico de Ciencias? ¿Cuáles son los 5 docentes con mayores calificaciones en las encuestas del último semestre? ¿Cuáles son los departamentos académicos con mayores calificaciones de los docentes en el último semestre?, entre otras cuestiones.

4.2.2. REQUERIMIENTOS NO FUNCIONALES

Los requerimientos no funcionales identificados son los siguientes:

- El modelo multidimensional de la solución debe permitir la creación de nuevas dimensiones en el futuro.

- La herramienta de explotación seleccionada debe permitir realizar las técnicas de consulta multidimensional como Slicing, Dicing, Drilling, etc. Además, debe permitir la creación de reportes personalizados por el usuario final.

- La solución debe contener reportes y tableros de control elaborados con la herramienta de explotación seleccionada. Además de la

posibilidad de exportarlos a archivos de formatos estándar como Excel (.xls).

- la herramienta de explotación debe permitir mostrar la información del Data Mart en base a los siguientes 3 tipos: Reportes tradicionales (basados en columnas), Tablas pivote (matriz de doble entrada) y Gráficos.

- La herramienta de explotación debe poder ser accedida vía Web y tener una disponibilidad de 24 x 7.

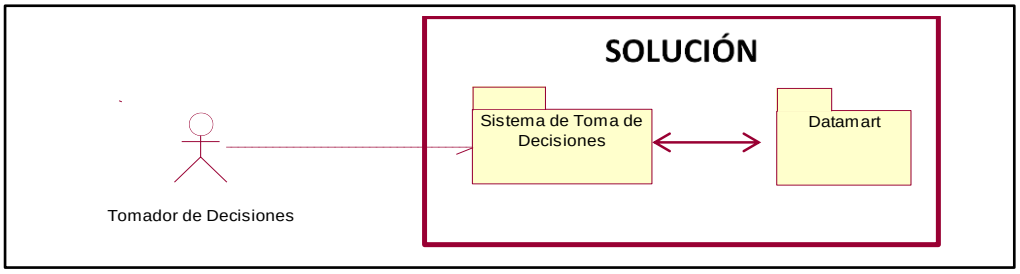
4.3. ANÁLISIS DE REQUERIMIENTOS

En este ítem se detallará el análisis de los requerimientos funcionales identificados que fueron empleados para el desarrollo de la solución.

4.3.1. DIAGRAMA DE CONTEXTO

Se tuvo en cuenta la definición de requerimientos para el planteamiento de la solución, a fin que esté satisfaga dichos requerimientos a través de su funcionalidad. En tal sentido la solución está compuesta por dos componentes, el primero por el sistema de soporte de decisiones y el otro por el Data Mart, tal como se muestra en la figura Nº 15- Diagrama de Contexto, siendo el sistema de toma de decisiones la interfaz gráfica por la cual el usuario, el tomador de decisiones, interactúa con los datos para así de esta forma realizar el análisis de la información. Siendo este sistema de toma de decisiones la herramienta de explotación de datos desarrollada en el proceso de desarrollo del presente proyecto.

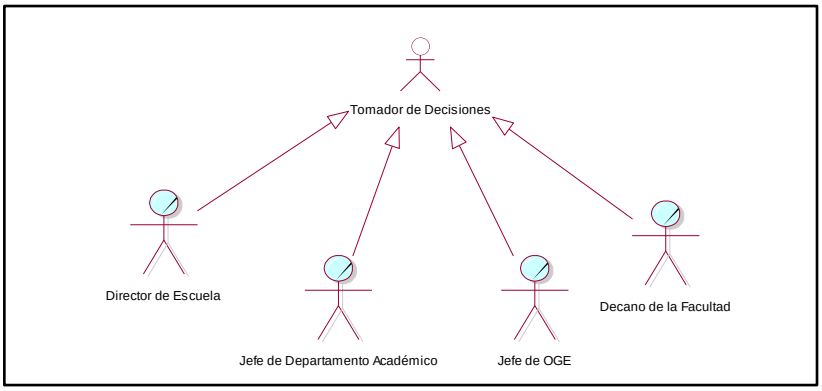
Figura Nº 15. Diagrama de Contexto



***Fuente**: Elaboración propia*

Los tipos de usuarios que interactúan con la solución confluyen en el Tomador de Decisiones, quien está conformado por el Director de Escuela, Jefe de Departamento Académico, Decano, jefe de OGE. La diferencia entre estos tipos de usuario son los permisos al acceso de la información y pantallas de la interfaz.

Figura Nº 16. Tomadores de decisiones



***Fuente:** Elaboración Propia*

4.4. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL

Después de haber realizado el análisis de la situación actual, y haber identificado las carencias de los procesos para la obtención de datos e información por parte de los tomadores de decisiones, se logró identificar que para la realización y obtención de datos académicos se realizan diversos procesos que en la mayoría son realizadas de forma manual, lo que provoca varias deficiencias y demoras en la entrega de datos solicitados.

Teniendo en cuenta que la Facultad de ciencias no cuenta con un Data Mart, que permita la realización de consultas de datos académicos de los estudiantes, docentes y mallas curriculares en los diversos semestres que viene funcionando, lo que provoca que las consultas más elaboradas requieran mayor cantidad de tiempo de lo esperado, ya que este tipo de consultas serán realizadas de forma manual utilizando recursos materiales y humanos. A ello se suma la existencia de personal con poco conocimiento de las tecnologías de información y el poco interés en cuanto a este tema.

Además, se realizó un cuadro de diagnóstico de la situación actual, mostrado en el Cuadro Nº 11, que muestra las actividades de consulta que se realizan, identificando como se llevan a cabo, que recursos utilizan y quienes realizan la actividad planteada, para ello se planteó las siguientes cuestiones ¿Cómo lo hace?, ¿Con qué lo hace? y ¿Quién lo realiza? respectivamente.

Cuadro Nº 11. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL

DIAGNÓSTICO DE LA SITUACIÓN ACTUAL				
Nº	**Actividades**	**¿Cómo lo hace?**	**¿Con qué lo hace?**	**¿Quién lo realiza?**
01	Consulta de datos académicos del estudiante	Se verifica los datos del estudiante, se realiza una consulta al SIGA-WEB. Siendo una consulta con datos comparativos, se utilizará los datos del SIGA WEB y se elaborará otra de forma manual.	Se realiza con los datos del SIGA-WEB y materiales físicos como actas.	Realizado por la secretaria de la escuela académica, directores de escuela y decanos.
02	Consulta de datos del Docente	Se verifica los datos del docente, se realiza una consulta al SIGA-WEB. Siendo una consulta con datos comparativos, se utilizará los datos del SIGA WEB, las hojas de vida y se elaborará otra de forma manual.	Se realiza con los datos del SIGA-WEB y materiales físicos como hojas de vida.	Realizado por la secretaria de la escuela académica, directores de escuela y decanos.
03	Consulta de datos curriculares de una especialidad	Se solicita la malla curricular de una especialidad, para ello se realiza una consulta al SIGA-WEB. Siendo una consulta con datos comparativos, se utilizará los datos del SIGA WEB, y las que se encuentra en el plan curricular y se elaborará otra de forma manual.	Se realiza con los datos del SIGA-WEB y materiales físicos como el plan curricular de cada especialidad.	Realizado por la secretaria de la escuela académica, directores de escuela y decanos.

Fuente: *Elaboración propia.*

CAPÍTULO V

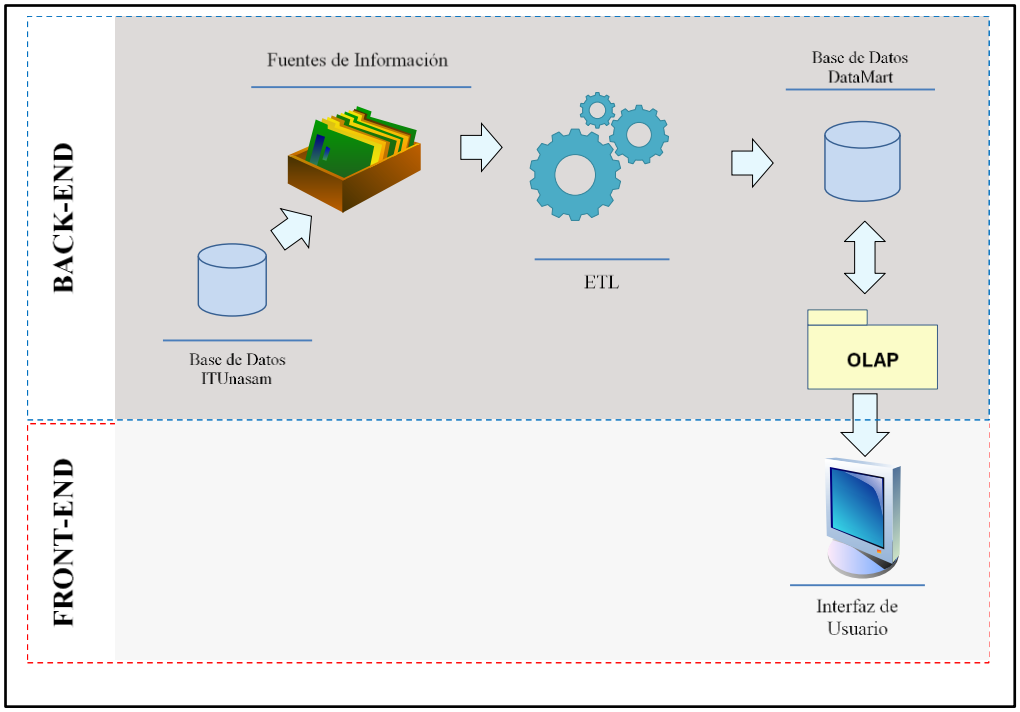
DISEÑO DE LA SOLUCIÓN

En el presente capítulo se mostrará la arquitectura que conforma la solución, para luego detallar la estructura, la funcionalidad y la interfaz de la solución

5.1. ARQUITECTURA TECNOLÓGICA DE LA SOLUCIÓN

La arquitectura que permitió la implantación del Data Mart aplicado a la Facultad de Ciencias está conformada por diversos componentes, entre las cuales se encuentra las Fuentes de Información, los componentes ETL, el Data Mart, Componentes OLAP y las interfaces de Usuario, la cual es ilustrado en la figura Nº 17 Diagrama de la arquitectura tecnológica de la solución, en la cual se subdivide entre los componentes que pertenecen al Back-End y los que pertenecen al Front-End.

Figura Nº 17. Diagrama de la arquitectura tecnológica de la solución.



Fuente: *Elaboración propia*

5.1.1. FUENTES DE INFORMACIÓN

La fuente de información principal está conformada por la base de datos ITUnasam, que contiene la información a nivel transaccional de la gestión académica de todos los estudiantes de la universidad desde sus inicios hasta la actualidad.

El diagrama de clases de la base de datos fuente se muestra en la figura Nº 18, además el diccionario de datos se muestra en el Anexo 8. Esta base de datos es la fuente principal que proporciona datos al Data Mart.

Figura Nº 18. Diagrama de Clases de la base de datos fuente

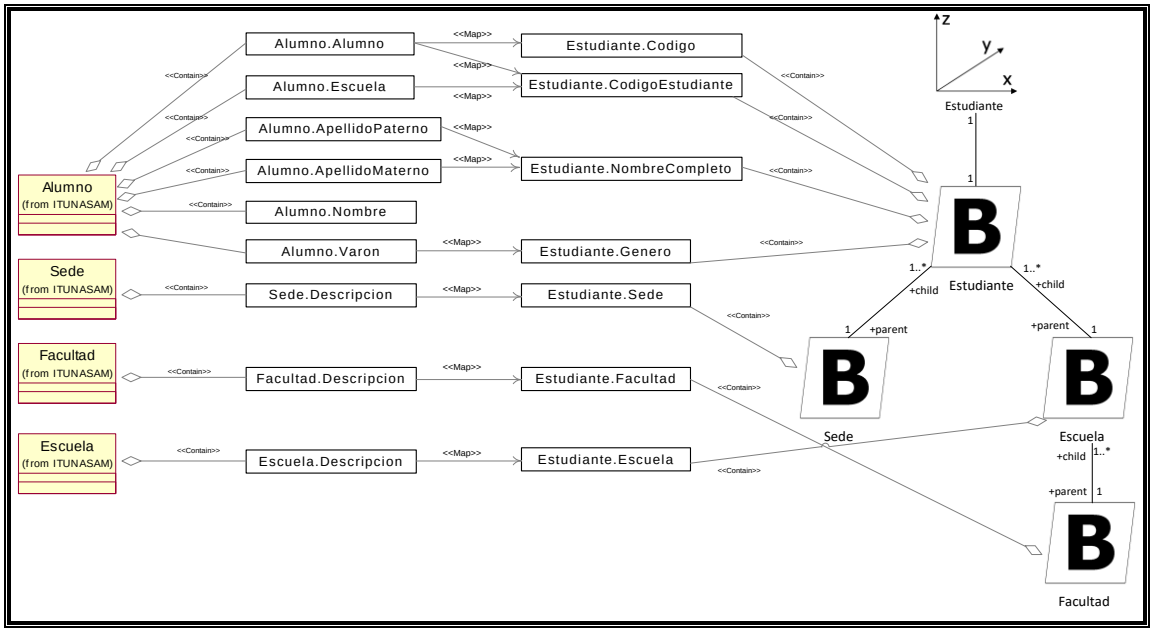
Fuente: Base datos ITUnasam

5.1.2. ETL

El componente ETL estuvo dado por los procesos inherentes para la extracción de los datos de la fuente de información, para luego transformarlo y cargarlo en la base de datos del Data Mart de la Facultad.

Para la realización del proceso ETL, fue importante realizar el mapeo de datos de la base de datos fuente con la base de datos del Data Mart, para lograr establecer las relaciones de los campos en ambas bases de datos. La metodología DWEP propone el diagrama de Mapeo de datos la cual es desarrollada en el anexo Nº 10, y uno de esos diagramas es mostrado en la figura Nº 19, en la cual se puede observar las clases del lado izquierdo que corresponde a las clases de la base de datos fuente ITUNASAM y las del lado derecho a la dimensión de Estudiante perteneciente al Data Mart.

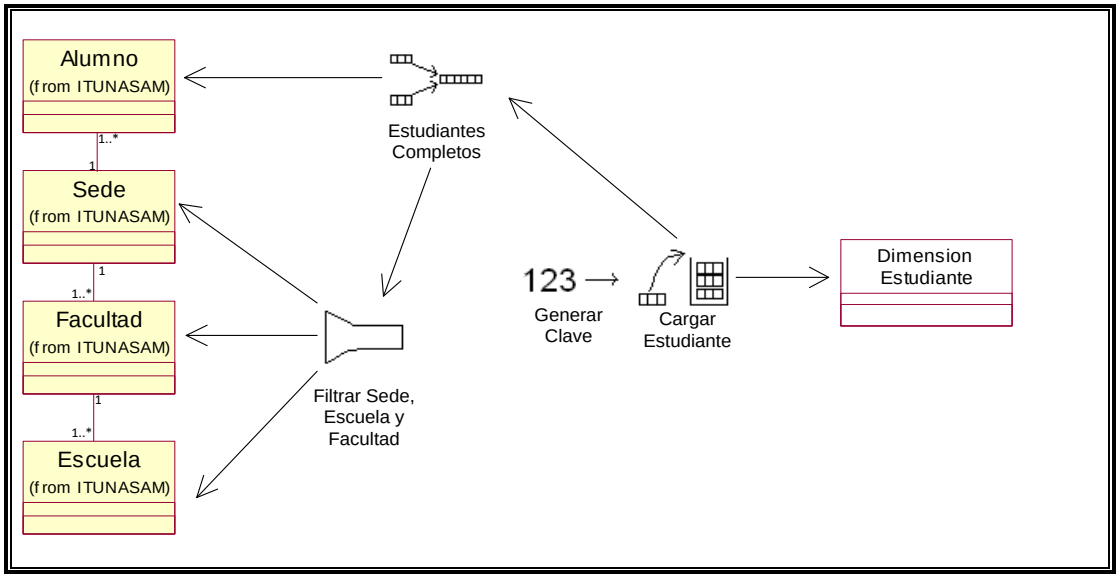
Figura Nº 19. Diagrama de Mapeo de Estudiante



Fuente: Elaboración propia

Después del mapeo de datos, fue necesario realizar el diseño del flujo del proceso ETL, con el fin de cargar cada una de las dimensiones y hechos del Data Mart teniendo en cuenta todas las variaciones y/o transformaciones de los datos que se realizaron al pasar de la base de datos fuente a la base de datos del Data Mart. La metodología DWEP propone un diagrama que define el diseño de los procesos ETL, siendo estos diagramas desarrollado en el Anexo 13. En la figura Nº 20 diagrama de proceso ETL para la carga inicial de la dimensión Estudiante, se puede observar que los datos requeridos para dicha carga son de las tablas Alumno, Sede, Facultad y escuela de la fuente de datos ITUnasam, teniendo en cuenta que se genera una nueva llave primaria y se filtran los datos por la sede, escuela y Facultad.

Figura Nº 20. Proceso ETL de la Dimensión Estudiante



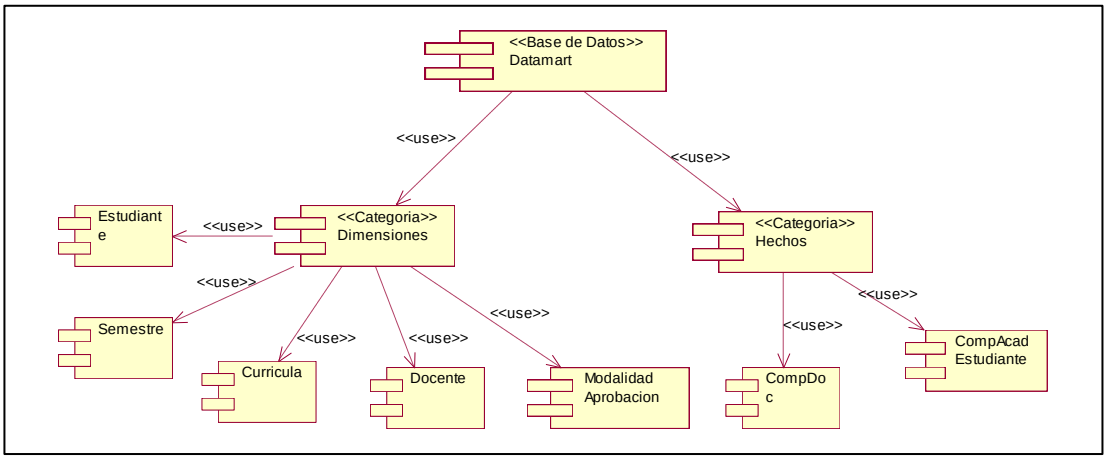
Fuente: Elaboración propia

5.1.3. DATA MART

El componente Data Mart representa a la base de datos multidimensional propiamente dicha, la cual posee una estructura adecuada que permite la explotación y el análisis de datos, que serán requeridos como fuente principal para la toma de decisiones.

Este componente es una base de datos, fue desarrollado sobre un motor de bases de Datos MSSQL tal como fue desarrollado también el componente de Fuentes de datos. Al ser un Data Mart, las tablas están divididas en dos grupos: Dimensiones y hechos. Esta división facilita agregar de forma más simple nuevas tablas en el tiempo. La representación de las tablas que conforman la base de datos del Data Mart es mostrado a través del diagrama de componente de la base de datos, figura Nº 21, la cual es ampliada en el anexo Nº 12. Esquema físico del Data Mart.

Figura Nº 21. Proceso ETL de la Dimensión Estudiante



Fuente: Elaboración propia

5.1.4. OLAP

El componente OLAP, es acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing), la que contiene los servicios OLAP que sirviéndose de la base de datos de Data Mart permite mostrar los datos en Cubos multidimensionales para su mejor análisis. En la presente se empleó consultas SQL para la creación del cubo, las cuales deben de ser transparentes para los usuarios finales.

5.1.5. INTERFAZ DE USUARIO

El componente interfaz de usuario representa la interfaz con la que interactúa el usuario final, a través de la cual es posible adquirir datos relevantes y útiles que faciliten la toma de decisiones. Para la explotación de datos se desarrolló un aplicativo web que se emplea como una herramienta de explotación de datos e interfaz de usuario, con la cual los usuarios finales interactuaran para la obtención de datos relevantes para la toma de decisiones dentro de la Facultad de ciencias.

5.2. ANÁLISIS DIMENSIONAL

En esta sección se analiza el modelo dimensional que responde a los requerimientos expuestos en la sección 4.2 del presente documento.

La metodología DWEP divide el diseño de un Data Mart en fases bien definidas, las cuales son Diseño conceptual, lógico y físico. En cada fase se genera un entregable y estos son: Esquema Conceptual del Data Mart, Esquema Lógico del Data Mart y Esquema Físico del Data Mart.

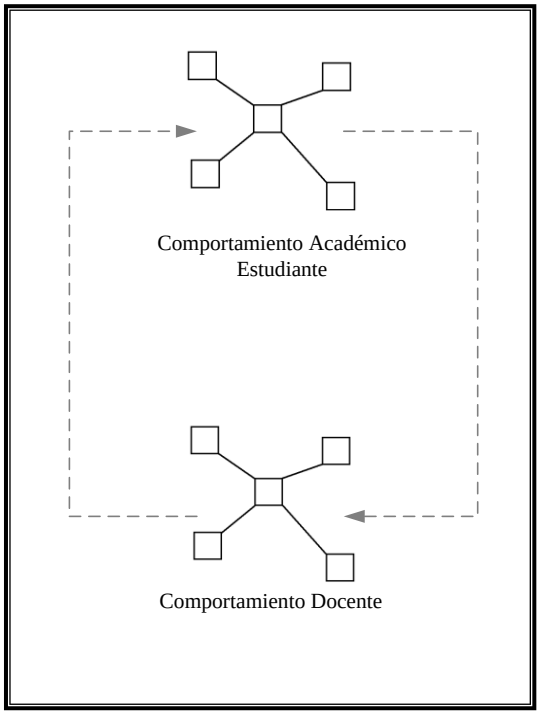
5.2.1. DISEÑO CONCEPTUAL DEL DATA MART

Para satisfacer las necesidades de información de los usuarios de la Facultad de Ciencias se definió dos esquemas de modelo estrella: Comportamiento Académico del estudiante y Comportamiento Docente. Teniendo en cuenta que cada esquema abarca un tema de análisis de información en particular organizando las variables y dimensiones necesarias para satisfacer la demanda de información.

5.2.1.1. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 1

En este nivel se muestra la relación entre los esquemas estrella que conforman la solución. Existen dos esquemas estrella en la solución. El primer esquema es el comportamiento académico estudiante el cual presenta información acerca del avance académico de los estudiantes. El segundo esquema es de comportamiento docente el cual muestra información general del avance académico de los docentes ideal para el análisis a detalle.

Figura Nº 22. Esquema Conceptual del Datamart – Nivel 1



Fuente: Elaboración propia

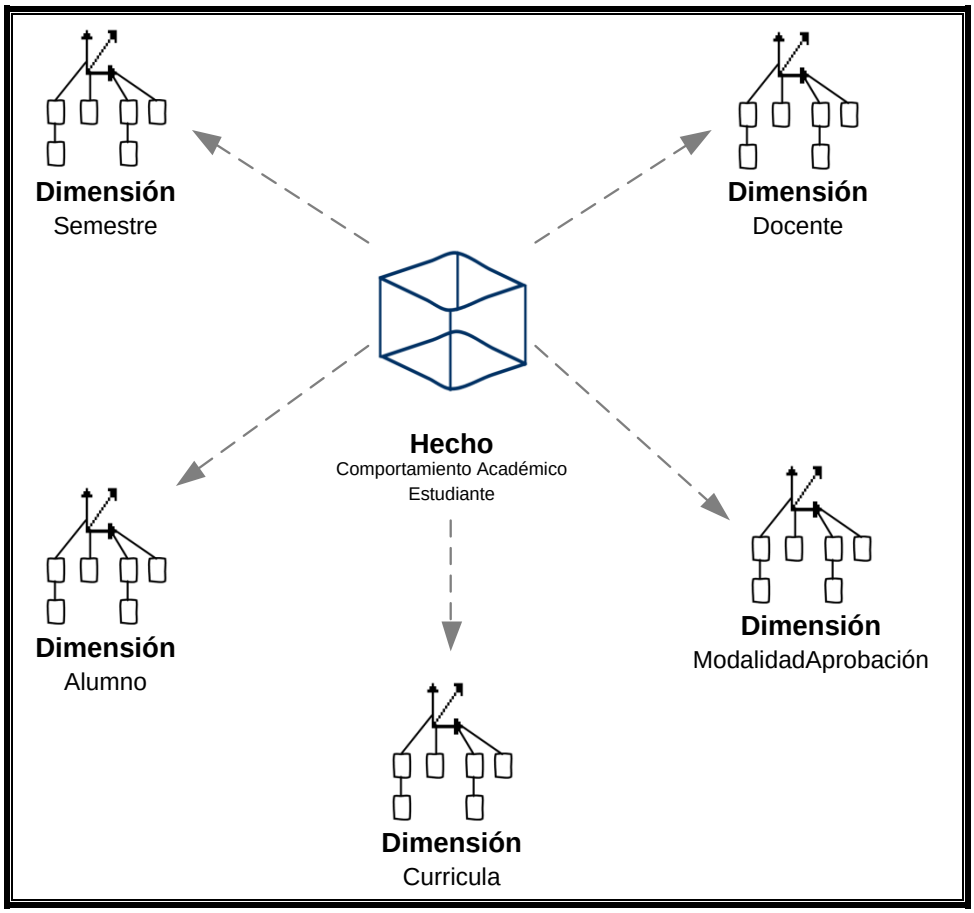
5.2.1.2. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 2

En este nivel se detalla cada uno de los esquemas presentado en el nivel 1, además se presenta las dimensiones con que se relaciona cada tabla de hecho.

A. ESQUEMA COMPORTAMIENTO ACADÉMICO DEL ESTUDIANTE

El esquema de comportamiento académico del estudiante tiene como tabla principal a la tabla de hechos Comportamiento Académico Estudiante, la cual contiene la información acerca del avance académico de los estudiantes, así como también de la malla curricular en la que está el estudiante.

Figura Nº 23. Esquema conceptual del Comportamiento Académico del estudiante

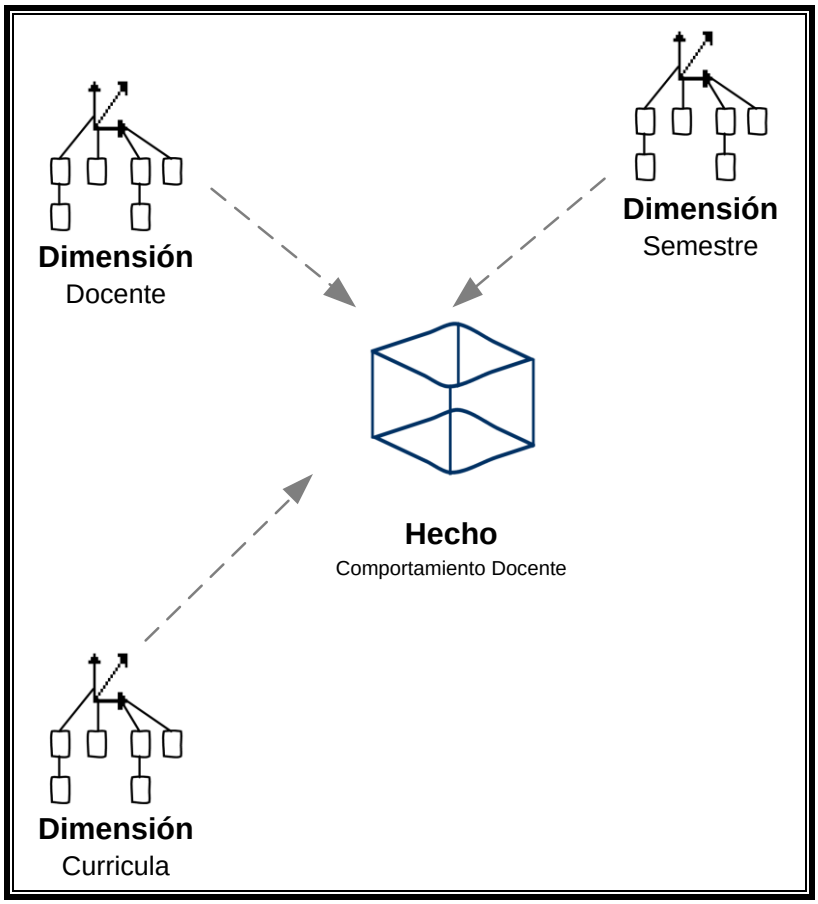


Fuente: Elaboración propia

B. ESQUEMA COMPORTAMIENTO DOCENTE

El esquema de comportamiento Docente tiene como tabla principal a la tabla de hechos Comportamiento Docente, la cual contiene la información acerca de los cursos que dicta un docente y la valoración que los estudiantes le dan a los docentes.

Figura Nº 24. Esquema conceptual del Comportamiento Docente



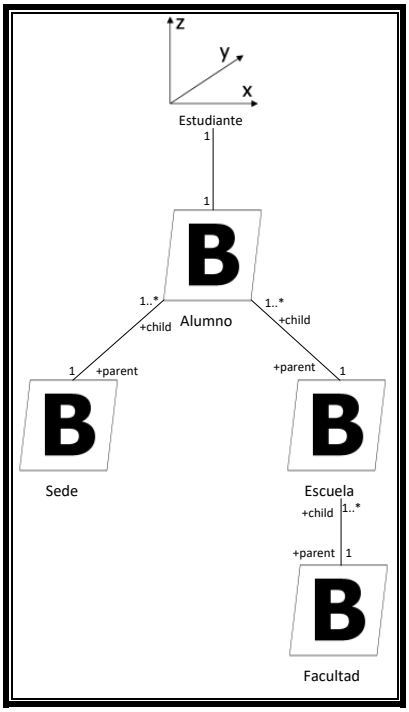
Fuente: Elaboración propia

5.2.1.3. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 3

En este nivel se desarrolla el detalle de cada dimensión presentado en el nivel anterior para poder definir la información que contiene y sobre todo la jerarquía en la información contenida por cada dimensión. La especificación de cada dimensión a nivel lógico podrá observarse en el entregable de Esquema Lógico de Data Mart.

A. DIMENSION ESTUDIANTE

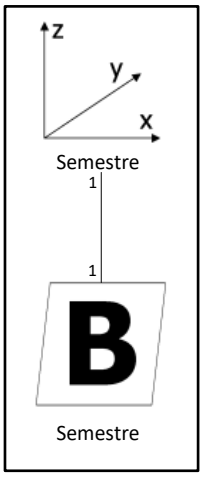
Figura Nº 25. Esquema conceptual de la Dimensión Estudiante



Fuente: Elaboración propia

B. DIMENSION SEMESTRE

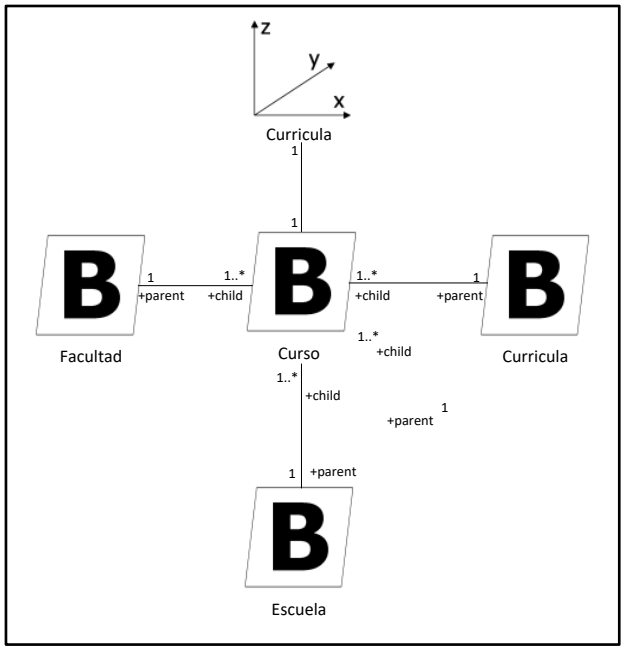
Figura Nº 26. Esquema conceptual de la Dimensión Semestre



Fuente: Elaboración propia

C. DIMENSION CURRICULA

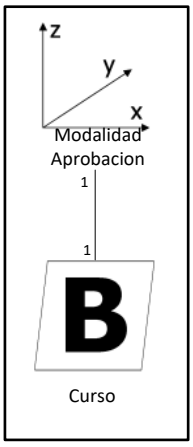
Figura Nº 27. Esquema conceptual de la Dimensión Curricula



Fuente: Elaboración propia

D. DIMENSION MODALIDAD DE APROBACION

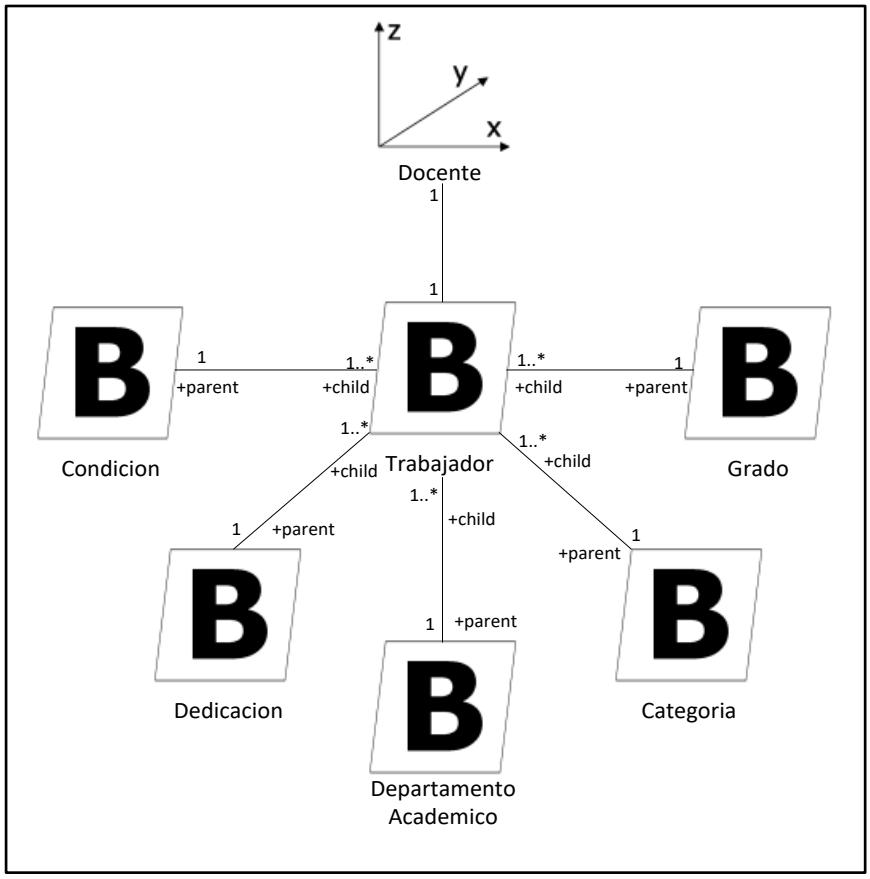
Figura Nº 28. Esquema conceptual de la Dimensión Modalidad de Aprobación



Fuente: Elaboración propia

E. DIMENSION DOCENTE

Figura Nº 29. Esquema conceptual de la Dimensión Modalidad de Aprobación



Fuente: Elaboración propia

En resumen, se puede observar en el cuadro Nº 12 Facts vs Dimensiones, en la cual se observa todos los modelos tipo de Estrella en las columnas y en las filas las dimensiones, las mismas que son empleadas en más de una tabla de Hechos.

Cuadro Nº 12. Facts vs Dimensiones

	FACTS O TABLA DE HECHOS	
DIMENSIONES	Comportamiento Académico Estudiante	Comportamiento Docente
Estudiante	**X**	
Curricula	**X**	**X**
Semestre	**X**	**X**
Modalidad de Aprobación	**X**	
Docente	**X**	**X**

Fuente: Elaboración propia

5.3. DISEÑO DE EXTRACCIÓN

El diseño de la extracción del Data Mart detalla el mapeo de datos de cómo se cargan los datos de las tablas fuentes a las tablas destino, las cuales son mostradas líneas abajo.

Cuadro Nº 13. Mapeo de Datos del Data Mart de la tabla Estudiante

Tabla	Estudiante	
Campo	**Tipo**	**Mapeo**
CodigoAlumno	VARCHAR (14)	Estudiante.CodigoAlumno
Codigo	VARCHAR (12)	Estudiante.Codigo
NombreCompleto	VARCHAR (150)	Mayuscula(Estudiante.NombreCompleto)
Sede	VARCHAR (100)	Mayuscula(Estudiante.Sede)
Facultad	VARCHAR (100)	Mayuscula(Estudiante.Facultad)
Escuela	VARCHAR (100)	Mayuscula(Estudiante.Escuela)
Genero	VARCHAR (10)	Mayuscula(Estudiante.Genero)

Fuente: Elaboración Propia

Cuadro Nº 14. Mapeo de Datos del Data Mart de la tabla Docente

Tabla	Docente	
Campo	**Tipo**	**Mapeo**
CodigoDocente	VARCHAR (6)	Docente.CodigoDocente
NombreDocente	VARCHAR (150)	Mayuscula(Docente.NombreDocente)
Dedicacion	VARCHAR (100)	Mayuscula(Docente.Dedicacion)
Categoria	VARCHAR (100)	Mayuscula(Docente.Categoria)
Condicion	VARCHAR (100)	Mayuscula(Docente.Condicion)
Departamento Academico	VARCHAR (100)	Mayuscula(Docente.Departamento Academico)
Grado	VARCHAR (100)	Mayuscula(Docente.Grado)
Profesion	VARCHAR (100)	Mayuscula(Docente.Profesion)

Fuente: Elaboración Propia

Cuadro Nº 15. Mapeo de Datos del Data Mart de la tabla Semestre

Tabla	Semestre	
Campo	**Tipo**	**Mapeo**
CodigoSemestre	varchar(6)	Semestre.CodigoSemestre
Semestre	varchar(100)	Mayuscula(Semestre.Semestre)

Fuente: Elaboración Propia

Cuadro Nº 16. Mapeo de Datos del Data Mart de la tabla ModalidadAprobacion

Tabla	ModalidadAprobacion	
Campo	**Tipo**	**Mapeo**
CodigoAprobacion	varchar(2)	ModalidadAprobacion.CodigoAprobacion
ModalidadAprobacion	varchar(100)	ModalidadAprobacion.ModalidadAprobacion

Fuente: Elaboración Propia

Cuadro Nº 17. Mapeo de Datos del Data Mart de la tabla ComAcadEstudiante

Tabla	ComAcadEstudiante	
Campo	**Tipo**	**Mapeo**
CodigoCurricula	varchar(16)	ComAcadEstudiante.CodigoCurricula
CodigoAlumno	varchar(12)	ComAcadEstudiante.CodigoAlumno
CodigoDocente	varchar(6)	ComAcadEstudiante.CodigoDocente
CodigoSemestre	varchar(6)	ComAcadEstudiante.CodigoSemestre
CodigoAprobacion	varchar(2)	ComAcadEstudiante.CodigoAprobacion
PromedioFinal	numeric(18,0)	ComAcadEstudiante.PromedioFinal
Promedio Anterior	numeric(18,0)	ComAcadEstudiante.Promedio Anterior
NotaAplazados	numeric(18,0)	ComAcadEstudiante.NotaAplazados
Condicion	varchar(100)	Mayuscula(ComAcadEstudiante.Condicion)

Fuente: Elaboración Propia

Cuadro Nº 18. Mapeo de Datos del Data Mart de la tabla CompDoc

Tabla	CompDoc	
Campo	**Tipo**	**Mapeo**
CodigoCurricula	VARCHAR (16)	CompDoc.CodigoCurricula
CodigoAlumno	VARCHAR (12)	CompDoc.CodigoAlumno
CodigoDocente	VARCHAR (6)	CompDoc.CodigoDocente
CodigoSemestre	VARCHAR (6)	CompDoc.CodigoSemestre
Seccion	CHAR (1)	CompDoc.Seccion
Encuesta	VARCHAR (100)	CompDoc.Encuesta
Pregunta	VARCHAR (200)	CompDoc.Pregunta
Valoracion	NUMERIC (18,2)	CompDoc.Valoracion
Cantidad	INT	CompDoc.Cantidad

Fuente: Elaboración Propia

Cuadro Nº 19. Mapeo de Datos del Data Mart de la tabla Curricula

Tabla	Curricula	
Campo	**Tipo**	**Mapeo**
CodigoCurricula	VARCHAR(16)	Curricula.CodigoCurricula
Facultad	VARCHAR(100)	Mayuscula(Curricula.Facultad)
Escuela	VARCHAR(100)	Mayuscula(Curricula.Escuela)
Curricula	CHAR(2)	Mayuscula(Curricula.Curricula)
NombreCurso	VARCHAR(100)	Mayuscula(Curricula.NombreCurso)
Ciclo	CHAR(4)	Mayuscula(Curricula.Ciclo)
TipoCurso	VARCHAR(100)	Mayuscula(Curricula.TipoCurso)
HorasTeorica	INT	Curricula.HorasTeorica
HorasPractica	INT	Curricula.HorasPractica

Fuente: Elaboración Propia

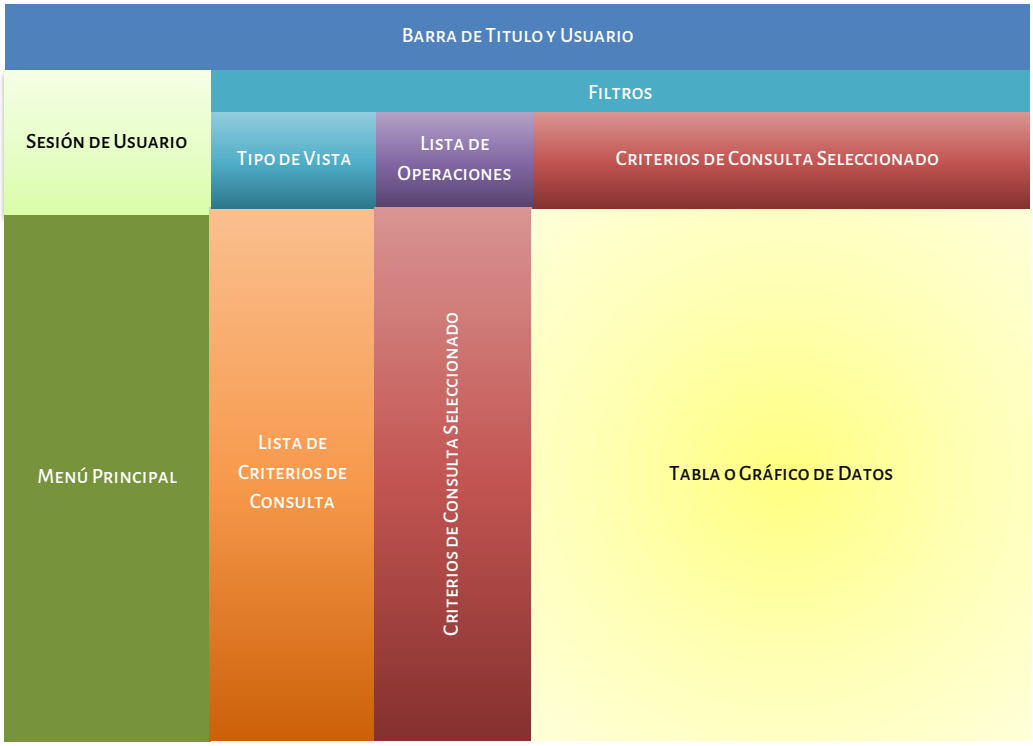
5.4. DISEÑO DE EXPLOTACIÓN

La interfaz de usuario es la herramienta clave con la cual el usuario final interactúa con el fin de obtener datos relevantes las cuales posteriormente utilizará para la toma de decisiones de la Facultad. Con este fin se desarrolló un aplicativo web que permite la explotación de datos, la cual se desarrolló utilizando herramientas actuales disponibles en el mercado.

La GUI de salida de datos del aplicativo web para la explotación, debe de permitir la fácil manipulación de criterios de consulta por dimensiones y los

valores numéricos respectivos, en tal sentido se planteó el siguiente diseño de interfaz gráfica de usuario, mostrado en la figura Nº 30, todas las páginas del aplicativo web, a excepción de la interfaz de inicio de sesión y el dashbord seguirán el patrón de diseño.

Figura Nº 30. Patrón de diseño del aplicativo web



Fuente: Elaboración Propia

- **Barra de Titulo y Usuario**, Contiene el título del aplicativo web, el logotipo de la Universidad Nacional Santiago Antúnez de Mayolo, y la configuración necesaria de Usuario.

- **Sesión de Usuario**, Durante la sesión activa se muestra el nombre del Usuario y el cargo que ocupa.

- **Menú Principal**, accede a cada ítem del menú despegable que muestra las diversas funcionalidades del aplicativo.

- **Filtros**, filtra la carga de datos según cada condicional.

- **Tipo de Vista**, selecciona el tipo de vista: Tabla, gráfico entre otras opciones de visualización.

- **Lista de Operaciones**, lista las operaciones con las cuales se pueden operacionalizar cada uno de los criterios, como por ejemplo la suma, máximo, mínimo, promedio, entre otras operaciones.

- **Lista de Criterios de Consulta**, muestra los criterios de consulta de todas las dimensiones que interactúan en un hecho.

- **Criterios de Consulta Seleccionado**, Muestra los criterios seleccionados por columna o fila.

- **Tabla o Gráfico de Datos**, operacionaliza los criterios seleccionados según las operaciones realizadas, generando así una tabla Pivote.

Además, los datos generados podrán ser migrados a una hoja de cálculo de Excel, la cual podrá ser utilizado fácilmente por el usuario final.

CAPÍTULO VI

CONSTRUCCIÓN DE LA SOLUCIÓN

En el presente capítulo se desarrolla los criterios y aspectos tomados en cuenta en la etapa de construcción. Además, la estrategia utilizada según el tipo de prueba utilizado.

6.1. CONSTRUCCIÓN

La fase de construcción está definida principalmente por la implementación de los procesos ETL debido a que los modelos de la base de dato fuente y del Data Mart ya fueron desarrolladas.

Adicionalmente, para la interfaz gráfica de usuario se emplea como herramienta de explotación un aplicativo web que fue desarrollado utilizando ingeniería de software. En esta sección se detallan las principales características tecnológicas, motores y frameworks utilizados para el desarrollo.

6.1.1. HERRAMIENTAS Y SOLUCIONES

Las herramientas utilizadas para el desarrollo del Data Mart y el aplicativo web, son detalladas líneas abajo.

6.1.1.1. BASE DE DATOS

El motor de base de datos utilizado para el desarrollo del Data Mart fue Microsoft SQL Server, que es una herramienta OLAP que permite dar soporte a las soluciones de Inteligencia de Negocios. Entre las principales ventajas de esta herramienta son:

- **Facilidad de uso**: Microsoft SQL facilita las tareas de administración de la base de datos debido a que automatizan la administración de la misma mediante una interfaz gráfica fácil de usar.

- **Autoadministración dinámica**: Automatiza muchas tareas de rutina como: los recursos de memoria, el tamaño de los archivos, etc.

- **Integración**: Permite la integración con los demás productos de la familia de Microsoft ampliamente conocidos como Windows y Microsoft Office.

- Desempeño, confiabilidad y escalabilidad.

6.1.1.2. FRAMEWORK

El framework utilizado para el desarrollo del aplicativo web fue laravel, la cual es un framework de código abierto para

desarrollar aplicaciones web y servicios web utilizando php. Las principales ventajas de este son:

- Sistema de ruteo.
- Blade y Motor de plantillas.
- Peticiones Fluent.
- Eloquent ORM.
- Basado en Composer.
- Soporte para el caché.
- Soporte para MVC, modelo vista controlador.
- Usa componentes de Symfony.

6.1.1.3. LENGUAJE DE PROGRAMACIÓN

El lenguaje de programación utilizado para el desarrollo fue PHP, Hypertext Preprocessor, la cual es un lenguaje de código abierto muy popular especialmente adecuado para el desarrollo web y que puede ser incrustado en HTML fácilmente.

6.1.2. ESPECIFICACIÓN DE CONSTRUCCIÓN

Las especificaciones de la construcción del Data Mart y el aplicativo web, definen la construcción a partir componentes o especificaciones de construcción independientes, que se corresponden con un empaquetamiento físico de los elementos del diseño de detalle, como pueden ser módulos, clases o especificaciones de interfaz.

Se generó las especificaciones necesarias para la creación de las estructuras de datos para el Data Mart, y el gestor de base de datos utilizado como se vio en el capítulo de diseño. En el cuadro Nº 20, se muestra las tareas, productos, técnicas y práctica, y los participantes

responsables de desarrollar dicha tarea todo ello según la metodología Metrica versión 3.

Cuadro Nº 20. Procedimiento de Operación y administración

Tarea		Productos	Técnicas y practicas	Participantes
DSI 8.1	Especificación del Entorno de Construcción	Especificaciones de Construcción del Sistema de Información: - Especificación del Entorno de Construcción	Análisis de requerimientos. Diagnóstico de situación Actual.	- Responsable del Proyecto - Responsable de Soporte Técnico - Responsable de Sistemas - Responsable de Seguridad
DSI 8.2	Definición de Componentes y Subsistemas de Construcción	Especificaciones de Construcción del Sistema de Información: - Descripción de Subsistemas de Construcción y Dependencias - Descripción de Componentes - Plan de Integración del Sistema de Información	Diagrama de Componentes Diagrama de Despliegue	- Responsable del Proyecto
DSI 8.3	Elaboración de Especificaciones de Construcción	Especificaciones de Construcción del Sistema de Información: - Especificación Detallada de Componentes	Diagrama de Componentes	- - Responsable del Proyecto
DSI 8.4	Elaboración de Especificaciones del Modelo Físico de Datos	Especificaciones de Construcción del Sistema de Información: - Especificación de la Estructura Física de Datos	Modelo Entidad Relación de la base de datos	- Responsable del Proyecto - Administradores de la Base de Datos

Fuente: Metodología MÉTRICA Versión 3- Construcción del Sistema de Información.

6.1.3. PROCEDIMIENTOS DE OPERACIÓN Y ADMINISTRACIÓN

El procedimiento de operación y administración del Data Mart y el aplicativo web tuvo como objetivo mantener la disponibilidad y operatividad de todos los componentes necesarios para la correcta construcción del Data Mart y el aplicativo web. Entre los componentes se encuentran las herramientas OLAP, los gestores de bases de datos, las librerías, el hardware requerido, entre otros.

Para la construcción se requirió los diversos diagramas desarrollados en los capítulos anteriores, y la especificación de construcción que representaron el insumo para el presente.

Cuadro Nº 21. Procedimiento de Operación y administración

Tarea		Productos	Técnicas y practicas	Participantes
CSI 1.1	Implantación de la Base de Datos Física.	Base de Datos del Data Mart.	-	**Equipo del Proyecto** - Administrador de Bases de Datos
CSI 1.2	Preparación del Entorno de Construcción.	Entorno de Construcción del aplicativo Web.	-	**Equipo del Proyecto** - Técnicos de Sistemas - Equipo de Operación - Administrador de Bases de Datos

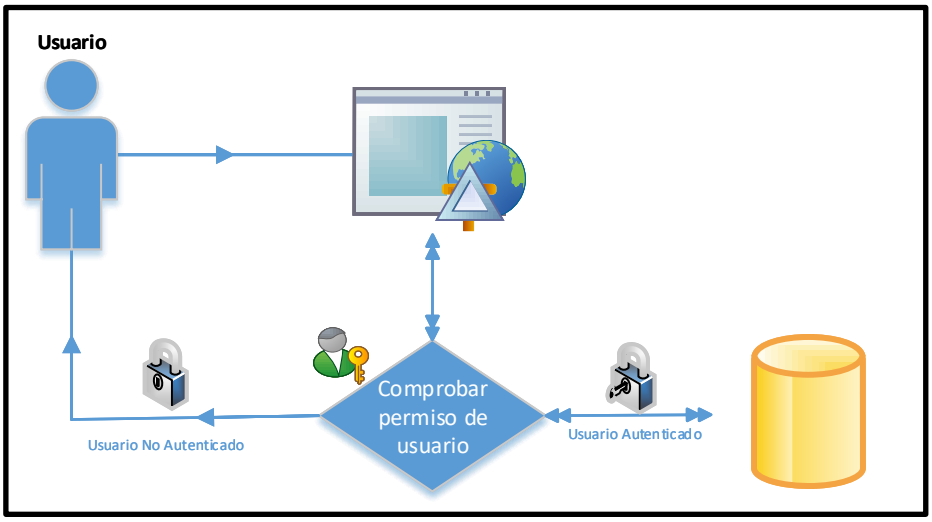
Fuente: Metodología MÉTRICA Versión 3- Construcción del Sistema de Información.

6.1.4. PROCEDIMIENTOS DE SEGURIDAD Y CONTROL DE ACCESO

Las aplicaciones web requieren cierto grado de seguridad y control de acceso, ya que estas administran datos relevantes para la organización, en tal sentido dentro de la construcción del Data Mart y el aplicativo web con el objetivo de generar procedimientos de administración, de seguridad y control de acceso a este, brindando así seguridad y confidencialidad a los datos que procesen utilizando este aplicativo.

Los procedimientos de seguridad fueron creados utilizando estándares, y normas que se detallaron en capítulos anteriores. En la figura Nº 31 muestra el flujo de la administración de Seguridad y control de acceso.

Figura Nº 31. Flujo de administración de seguridad y control de Acceso.



Fuente: Elaboración Propia.

Además, la aplicación web se desarrolló tomando en cuenta el protocolo AAA (Authentication, Authorization & Accounting), la cual realiza tres funciones la autenticación, autorización y la contabilización. La primera es encargada de validar la identidad de un usuario ante otra, la segunda es la concesión de privilegios a un determinado usuario, y la ultima es el seguimiento del consumo o acciones realizadas durante la sesión del usuario.

Cuadro Nº 22. Generación del Código de los Componentes y Procedimientos

Tarea		Productos	Técnicas y practicas	Participantes
CSI 2.1	Generación del Código de Componentes	**Producto Software:** - Código Fuente de los Componentes		**Equipo del Proyecto** - Programadores
CSI 2.2	Generación del Código de los Procedimientos de Operación y Seguridad	**Producto Software:** - Procedimientos de Operación y Administración del Sistema - Procedimientos de Seguridad y Control de Acceso	-	**Equipo del Proyecto** - Técnicos de Sistemas - Equipo de Operación - Administrador de la Base de Datos - Programadores

Fuente: Metodología MÉTRICA Versión 3- Construcción del Sistema de Información.

6.1.5. PROCEDIMIENTOS DE OPERACIÓN Y MANUALES DEL USUARIO.

Teniendo en cuenta las especificaciones de construcción, así como la construcción de procedimientos se procedió a desarrollar el Data Mart y a codificar el aplicativo web. Adicionalmente a ello se procedió a la realización de las pruebas Unitarias y de integración.

Cuadro Nº 23. Procedimiento de Operación y Manuales de Usuario

Tarea		Productos	Técnicas y practicas	Participantes
CSI 6.1	Elaboración de los Manuales de Usuario	**Producto Software** - Manuales de Usuario	-	**Equipo del Proyecto** - Administrador de Bases de Datos

Fuente: Metodología MÉTRICA Versión 3- Construcción del Sistema de Información.

6.1.6. CONSTRUCCIÓN DE LOS COMPONENTES Y PROCEDIMIENTOS DE MIGRACIÓN Y CARGA INICIAL DE DATOS

La base de datos del Data Mart requirió de los procesos ETL, y estos requerían a su vez de procedimientos de migración y carga, a partir de las especificaciones recogidas en los capítulos de análisis y diseño.

Previamente a la generación del código, se preparó la infraestructura tecnológica necesaria para realizar la codificación y las pruebas de los distintos componentes y procedimientos asociados, de acuerdo a las características del entorno de migración especificado en el plan de migración y carga inicial de datos, visto en el capítulo de diseño, para posteriormente realizar la Migración.

Cuadro Nº 24. Construcción de los Componentes y Procedimientos de Migración y Carga Inicial de Datos

Tarea		Productos	Técnicas y practicas	Participantes
CSI 6.1	Elaboración de los Manuales de Usuario	**Producto Software** - Manuales de Usuario	-	**Equipo del Proyecto** - Administrador de Bases de Datos
CSI 8.1	Preparación del Entorno de Migración y Carga Inicial de Datos	- Entorno de Migración		**Equipo del Proyecto** - Técnicos de Sistemas - Equipo de Operación - Equipo de Seguridad - Administradores de Bases de Datos
CSI 8.2	Generación del Código de los Componentes y Procedimientos de Migración y Carga Inicial de Datos	- Código Fuente de los Componentes de Migración y Carga Inicial de Datos - Procedimientos de Migración y Carga Inicial de Datos		**Equipo del Proyecto** - Programadores
CSI 8.3	Realización y Evaluación de las Pruebas de Migración y Carga Inicial de Datos	- Resultado de las pruebas de migración y carga inicial de datos - Evaluación del resultado de las pruebas de migración y carga inicial de datos	- Pruebas Unitarias - Pruebas de Integración	**Equipo del Proyecto**

Fuente: Metodología MÉTRICA Versión 3- Construcción del Sistema de Información.

6.2. PRUEBAS

Dentro del proceso de la construcción de una solución, las pruebas constituyen una de las actividades indispensables, porque a través de estas se asegurará la calidad del software. Debido a ello es necesario realizar un plan de pruebas para verificar si cumple con los requisitos planteado en capítulos anteriores.

La prueba para la presente solución de inteligencia de negocios evalúa la calidad de los datos, lo que difiere con el aplicativo web en la cual se evalúa el software y su funcionamiento, en tal sentido en el presente ítem se presenta solo las pruebas del Data Mart. En tal sentido se probaron los siguientes puntos:

- Los procesos ETL, desde la base de datos fuente hasta la base de datos del Data Mart.

- La base de datos del Data Mart.

- Herramienta OLAP – aplicativo web.

Sobre estos objetos se realizarán las siguientes pruebas empleado el método de caja negra o también conocidos como las pruebas funcionales:

- Prueba de Extracción, Transformación y Carga desde la base de datos fuente a la base de datos del Data Mart.

- Prueba de Funcionalidad en Herramienta OLAP. Se prueba que la información en cada uno de los cubos corresponda a la información de la base de datos fuente.

Cuadro Nº 25. Elementos de Prueba y criterios de aceptación

Objetos de Prueba	Proceso Involucrado	Criterio de Aceptación
Proceso de Extracción, Transformación y carga de datos del Data Mart	Extracción, transformación y carga	Los datos extraídos de la base de datos fuente deben ser los mismos que los datos cargados en la base de datos del Data Mart.
Herramienta OLAP	Explotación	La información de los cubos debe corresponder a la obtenida de la base de datos del Data Mart.

Fuente: Metodología MÉTRICA Versión 3- Construcción del Sistema de Información.

6.2.1. CASO DE PRUEBA PARA ETL ESTUDIANTE

A. Caso de prueba

Cuadro Nº 26. Caso de Prueba 1

CASO DE PRUEBA Nº01				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla ESTUDIANTE sea correctamente cargada de la tabla fuente ALUMNO.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Fecha inicio - Fecha fin				
DATOS DE ENTRADA:				
- Tabla ALUMNO - Tabla SEDE - Tabla FACULTAD - Tabla ESCUELA				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La tabla ESTUDIANTE muestre la misma cantidad de registros que la tabla ALUMNO.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla ALUMNO: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 27. Registros de la Tabla Alumno

Alumno	ApellidoPaterno	ApellidoMaterno	Nombre	Varon
00.0001.5.AC	CABELLO	CHAVEZ	JULIO CESAR	1
00.0002.N.AC	PADILLA	PALMA	LUIS MIGUEL FORTUNATO	1
00.0004.9.AC	SOLIS	VERGARAY	GUSTAVO ADOLFO	1
00.0005.3.AC	BERROCAL	CABRERA	SADRO ANGEL	1
00.0006.8.AL	GAMARRA	TAHUA	DANIEL	1
00.0007.2.AM	PACHECO	ZAMUDIO	GEDEON JUAN	0
00.0008.7.AC	PATRICIO	MEDINA	ISRAEL PEDRO	1
00.0009.1.AC	JARA	REMIGIO	LUZ GUILLERMA	1

Fuente: Base de datos ITUNASAM

Tabla SEDE: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 28. Registros de la Tabla SEDE

Sede	Descripcion
01	HUARAZ
02	BARRANCA
03	POMABAMBA
04	CHIQUIAN

Fuente: Base de datos ITUNASAM

Tabla FACULTAD: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 29. Registros de la Tabla FACULTAD

Facultad	Descripcion
01	CIENCIAS
02	CIENCIAS AGRARIAS
03	ADMINISTRACION Y TURISMO
04	INGENIERIA INDUSTRIAS ALIMENTARIAS
05	CIENCIAS DEL AMBIENTE
06	INGENIERIA MINAS GEOLOGIA Y METALURGIA
07	INGENIERIA CIVIL
08	CIENCIAS MEDICAS
10	DERECHO Y CIENCIAS POLITICAS
11	CIENCIAS SOCIALES EDUCACION Y DE LA COMUNICACION

Fuente: Base de datos ITUNASAM

Tabla ESCUELA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 30. Registros de la Tabla ESCUELA

Escuela	Descripcion	Facultad
01	AGRONOMIA	02
02	INDUSTRIAS ALIMENTARIAS	04
03	INGENIERIA AGRICOLA	02
04	ESTADISTICA E INFORMATICA	01
05	MATEMATICA	01
06	INGENIERIA AMBIENTAL	05
07	INGENIERIA SANITARIA	05
08	INGENIERIA DE MINAS	06
09	INGENIERIA CIVIL	07
10	OBSTETRICIA	08

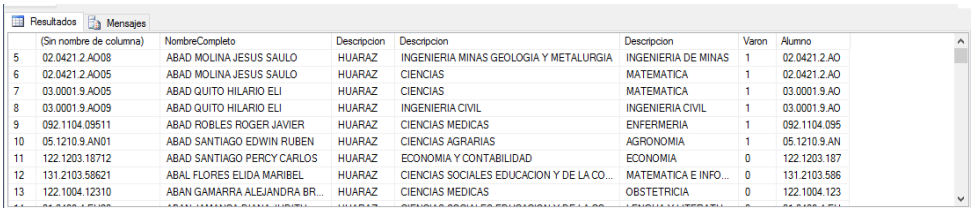
Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 32 muestra la consulta de Alumnos a la base de datos ITUNASAM, en la cual genera 32596 registros.

Figura Nº 32. Consulta a base de ITUNASAM, tabla Alumno



Resultados | Mensajes

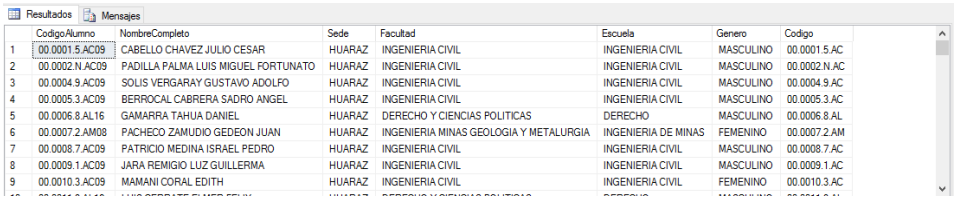
	(Sin nombre de columna)	NombreCompleto	Descripcion	Descripcion	Descripcion	Varon	Alumno
5	02.0421.2.AO08	ABAD MOLINA JESUS SAULO	HUARAZ	INGENIERIA MINAS GEOLOGIA Y METALURGIA	INGENIERIA DE MINAS	1	02.0421.2.AO
6	02.0421.2.AO05	ABAD MOLINA JESUS SAULO	HUARAZ	CIENCIAS	MATEMATICA	1	02.0421.2.AO
7	03.0001.9.AO05	ABAD QUITO HILARIO ELI	HUARAZ	CIENCIAS	MATEMATICA	1	03.0001.9.AO
8	03.0001.9.AO09	ABAD QUITO HILARIO ELI	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	1	03.0001.9.AO
9	092.1104.09511	ABAD ROBLES ROGER JAVIER	HUARAZ	CIENCIAS MEDICAS	ENFERMERIA	1	092.1104.095
10	05.1210.9.AN01	ABAD SANTIAGO EDWIN RUBEN	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	1	05.1210.9.AN
11	122.1203.18712	ABAD SANTIAGO PERCY CARLOS	HUARAZ	ECONOMIA Y CONTABILIDAD	ECONOMIA	0	122.1203.187
12	131.2103.58621	ABAL FLORES ELIDA MARIBEL	HUARAZ	CIENCIAS SOCIALES EDUCACION Y DE LA CO...	MATEMATICA E INFO...	0	131.2103.586
13	122.1004.12310	ABAN GAMARRA ALEJANDRA BR...	HUARAZ	CIENCIAS MEDICAS	OBSTETRICIA	0	122.1004.123

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 33 muestra la consulta de la dimensión Estudiante de la base de datos DATA MARTUNASAM, en la cual genera 32596 registros.

Figura Nº 33. Consulta a base de DATA MARTUNASAM, tabla Estudiante



Resultados | Mensajes

	CodigoAlumno	NombreCompleto	Sede	Facultad	Escuela	Genero	Codigo
1	00.0001.5.AC09	CABELLO CHAVEZ JULIO CESAR	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	MASCULINO	00.0001.5.AC
2	00.0002.N.AC09	PADILLA PALMA LUIS MIGUEL FORTUNATO	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	MASCULINO	00.0002.N.AC
3	00.0004.9.AC09	SOLIS VERGARAY GUSTAVO ADOLFO	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	MASCULINO	00.0004.9.AC
4	00.0005.3.AC09	BERROCAL CABRERA SADRO ANGEL	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	MASCULINO	00.0005.3.AC
5	00.0006.8.AL16	GAMARRA TAHUA DANIEL	HUARAZ	DERECHO Y CIENCIAS POLITICAS	DERECHO	MASCULINO	00.0006.8.AL
6	00.0007.2.AM08	PACHECO ZAMUDIO GEDEON JUAN	HUARAZ	INGENIERIA MINAS GEOLOGIA Y METALURGIA	INGENIERIA DE MINAS	FEMENINO	00.0007.2.AM
7	00.0008.7.AC09	PATRICIO MEDINA ISRAEL PEDRO	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	MASCULINO	00.0008.7.AC
8	00.0009.1.AC09	JARA REMIGIO LUZ GUILLERMA	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	MASCULINO	00.0009.1.AC
9	00.0010.3.AC09	MAMANI CORAL EDITH	HUARAZ	INGENIERIA CIVIL	INGENIERIA CIVIL	FEMENINO	00.0010.3.AC

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión Estudiante se comprueba que los datos son los esperados.

6.2.2. CASO DE PRUEBA PARA ETL CURRICULA

A. Caso de prueba

Cuadro Nº 31. Caso de Prueba Nº2

CASO DE PRUEBA Nº02				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla CURRICULA sea correctamente cargada de la tabla fuente CURRICULA.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Semestre de inicio - Semestre fin				
DATOS DE ENTRADA:				
- Tabla CURSOS - Tabla CURRICULA - Tabla ESCUELA - Tabla FACULTAD				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La dimensión CURRICULA muestre la misma cantidad de registros que la tabla CURRICULA.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla CURRICULA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 32. Registros de la Tabla Curricula

Curricula	Descripcion
01	CURRICULA 2005
05	CURRICULA 2007
04	CURRICULA 2007
04	CURRICULA 2008
05	CURRICULA 1997
04	CURRICULA 1998
02	CURRÍCULA 1993

Fuente: Base de datos ITUNASAM

Tabla SEDE: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 33. Registros de la Tabla SEDE

Sede	Descripcion
01	HUARAZ
02	BARRANCA
03	POMABAMBA
04	CHIQUIAN

Fuente: Base de datos ITUNASAM

Tabla CURSOS: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 34. Registros de la Tabla CURSOS

Curso	Nombre	Numero Creditos	Ciclo	Tipo Curso	Horas Teoria	Horas Practica	Horas Laboratorio	Tipo
CQ-G01	QUÍMICA GENERAL I	5	I	O	4	3	0	2
UE-K10	TALLER DE COMUNICACIÓN ORAL Y ESCRITA	2	I	O	1	2	0	2
CB-Q01	BIOLOGÍA GENERAL	4	I	O	3	3	0	2
CM-A02	MATEMÁTICA I	5	I	O	4	2	0	2
UE-L01	LENGUAJE I	3	I	O	2	2	0	2
CM-A02	MATEMATICAS I	5	I	O	4	2	0	2
BS-S01	INTRODUCCIÓN A LA INGENIERÍA SANITARIA	2	I	O	1	2	0	2
CB-Q01	BIOLOGIA GENERAL	4	I	O	3	3	0	2

Fuente: Base de datos ITUNASAM

Tabla ESCUELA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 35. Registros de la Tabla ESCUELA

Escuela	Descripcion	Facultad
01	AGRONOMIA	02
02	INDUSTRIAS ALIMENTARIAS	04
03	INGENIERIA AGRICOLA	02
04	ESTADISTICA E INFORMATICA	01
05	MATEMATICA	01
06	INGENIERIA AMBIENTAL	05
07	INGENIERIA SANITARIA	05
08	INGENIERIA DE MINAS	06

Fuente: Base de datos ITUNASAM

Tabla FACULTAD: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 36. Registros de la Tabla FACULTAD

Facultad	Descripcion
01	CIENCIAS
02	CIENCIAS AGRARIAS
03	ADMINISTRACION Y TURISMO
04	INGENIERIA INDUSTRIAS ALIMENTARIAS
05	CIENCIAS DEL AMBIENTE
06	INGENIERIA MINAS GEOLOGIA Y METALURGIA
07	INGENIERIA CIVIL
08	CIENCIAS MEDICAS
10	DERECHO Y CIENCIAS POLITICAS
11	CIENCIAS SOCIALES EDUCACION Y DE LA COMUNICACION

Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 34 muestra la consulta de Curricula a la base de datos ITUNASAM, en la cual genera 8744 registros.

Figura Nº 34. Consulta a base de ITUNASAM, tabla Curricula

Resultados | Mensajes

	(Sin nombre de columna)	(Sin nombre de columna)	(Sin nombre de columna)	Descripcion	Curricula	Nombre	Ciclo	TipoCurso	HorasTeoria	HorasPractica
1	0201*0UE-A03	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	ACTIVIDADES	I	O	0	0
2	0201*0NC-D01	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	ECONOMIA GENERAL	I	O	0	0
3	0201*0CQ-R01	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	QUIMICA GENERAL I	VI	O	0	0
4	0201*0CM-A01	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	SOCIOLOGIA GENERAL	I	O	0	0
5	0201*0UF-A01	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	BOTANICA GENERAL	VI	O	0	0
6	0201*0AS-C05	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	MANEJO Y CONSERVACIÓN DE SUELOS	X	O	0	0
7	0201*0QU-S02	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	SOCIOLOGIA GENERAL	I	O	0	0
8	0201*0AF-F07	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	BOTÁNICA SISTEMÁTICA	I	O	0	0
9	0201*0CH-111	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	LENGUAJE I	I	O	0	0
10	0201*0EP-112	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	PRINCIPIOS DE CONTABILIDAD	II	O	0	0
11	0201*0CH-202	HUARAZ	CIENCIAS AGRARIAS	AGRONOMIA	*0	SOCIOLOGIA RURAL	III	O	0	0

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 35 muestra la consulta de la dimensión Curricula de la base de datos DATA MARTUNASAM, en la cual genera 8744 registros.

Figura Nº 35. Consulta a base de DATA MARTUNASAM, tabla Curricula

Resultados | Mensajes

	CodigoCurricula	Sede	Facultad	Escuela	Curricula	NombreCurso	Ciclo	TipoCurso	HorasTeorica	HorasPract
1	010401FC-201	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	PRÁCTICAS PRE-PROFESIONALES EN ESTADÍSTICA	X	CURSO OBLIGATORIO	0	0
2	010401FC-202	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	SEMINARIO DE TÉSIS	X	CURSO OBLIGATORIO	5	3
3	010401FC-204	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	ASESORÍA Y CONSULTORIA ESTADÍSTICA	X	CURSO ELECTIVO	2	2
4	010401FC-205	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	CALIDAD TOTAL	X	CURSO ELECTIVO	2	2
5	010401FC-206	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	DECISIONES BAYESIANAS	X	CURSO ELECTIVO	2	2
6	010401FC-207	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	DEFENSA NACIONAL	X	CURSO ELECTIVO	2	2
7	010401FC-208	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	REALIDAD NACIONAL Y REGIONAL	X	CURSO ELECTIVO	2	2
8	010401FC-211	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	CÁLCULO I	I	CURSO OBLIGATORIO	4	2
9	010401FC-212	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	COMPLEMENTO DE MATEMÁTICA	I	CURSO OBLIGATORIO	4	2
10	010401FC-213	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	ESTADÍSTICA DESCRIPTIVA	I	CURSO OBLIGATORIO	4	2
11	010401FC-214	HUARAZ	CIENCIAS	ESTADISTICA E INFORMATICA	01	LENGUAJE	I	CURSO OBLIGATORIO	2	2

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión Curricula se comprueba que los datos son los esperados.

6.2.3. CASO DE PRUEBA PARA ETL DOCENTE

A. Caso de prueba

Cuadro Nº 37. Caso de prueba Nº 03

CASO DE PRUEBA Nº03				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla DOCENTE sea correctamente cargada de la tabla fuente DOCENTE.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Semestre de inicio - Semestre fin				
DATOS DE ENTRADA:				
- Tabla TRABAJADOR - Tabla DEDICACION - Tabla CATEGORIA - Tabla CONDICION - Tabla GRADO - Tabla DEPARTAMENTOACADEMICO				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La dimensión DOCENTE muestre la misma cantidad de registros que la tabla DOCENTE.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla TRABAJADOR: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 38. Registros de la Tabla TRABAJADOR

Trabajador	Apellido Paterno	Apellido Materno	Nombre	Departamento Académico	Grado
DFA028	DOMINGUEZ	FLORES	ANTONIO MARIANO	43	B
DHJ240	DAGA	HUARICANCHA	JOSE	43	B
DILA02	ITA	LUNA	ASTRID IRIS	44	B
DLG309	LOPEZ	DE GUIMARAES	DOUGLAS	47	B
DLJ003	DIAZ	LEDESMA	JOHN ALEX	46	B
DLLX02	DURAN	LEON	LIZ XIOMARA	46	B

Fuente: Base de datos ITUNASAM

Tabla DEDICACION: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 39. Registros de la Tabla DEDICACION

Dedicación	Descripción	Activo
DE	DEDICACIÓN EXCLUSIVA	1
TC	TIEMPO COMPLETO	1
TP	TIEMPO PARCIAL	1

Fuente: Base de datos ITUNASAM

Tabla CATEGORIA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 40. Registros de la Tabla CATEGORIA

Categoria	Descripcion	Abreviatura	Activo
1	ASOCIADO	A	1
2	PRINCIPAL	P	1
3	AUXILIAR	X	1
4	JEFE DE PRÁCTICAS	J	1

Fuente: Base de datos ITUNASAM

Tabla CONDICION: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 41. Registros de la Tabla CONDICION

Condicion	Descripcion	Abreviatura	Activo
1	NOMBRADO	N	1
2	CONTRATADO	C	1

Fuente: Base de datos ITUNASAM

Tabla GRADO: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 42. Registros de la Tabla GRADO

Grado	Descripcion	Activo
B	BACHILLER	1
D	DOCTORADO	1
M	MASTER	1

Fuente: Base de datos ITUNASAM

Tabla DEPARTAMENTOACADEMICO: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 43. Registros de la Tabla DEPARTAMENTOACADEMICO

Departamento Académico	Descripción	Facultad	Escuela	Activo
51	AGRONOMIA	2	1	1
38	CIENCIAS	1	5	1
39	ESTADISTICA Y SISTEMAS	1	25	1
40	CIENCIA Y TECNOLOGIA DE ALIMENTOS	4	2	1
41	INGENIERIA AGRICOLA	2	3	1
42	INGENIERIA CIVIL	7	9	1
43	INGENIERIA DE MINAS	6	8	1
44	CIENCIAS DEL AMBIENTE	5	6	1
45	ECONOMIA Y CONTABILIDAD	12	12	1
46	ADMINISTRACION Y TURISMO	3	13	1
47	CIENCIAS DE LA SALUD	8	10	1
48	DERECHO Y CIENCIAS POLITICAS	10	16	1
49	EDUCACION	11	26	1
50	CIENCIAS SOCIALES Y CIENCIAS DE LA COMUNICACION	11	24	1

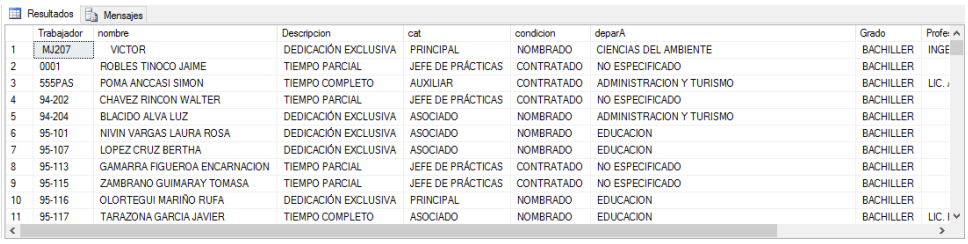
Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 36 muestra la consulta utilizando el script de carga de Docente a la base de datos ITUNASAM, en la cual genera 1431 registros.

Figura Nº 36. Consulta a base de ITUNASAM, utilizando el script de Docente



Resultados | Mensajes

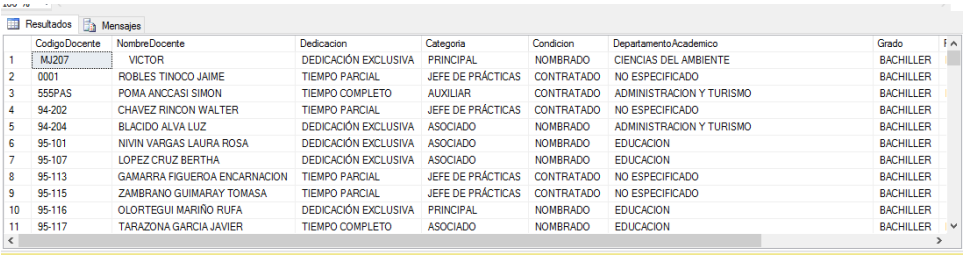
	Trabajador	nombre	Descripcion	cat	condicion	deparA	Grado	Profe
1	MJ207	VICTOR	DEDICACIÓN EXCLUSIVA	PRINCIPAL	NOMBRADO	CIENCIAS DEL AMBIENTE	BACHILLER	INGE
2	0001	ROBLES TINOCO JAIME	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
3	555PAS	POMA ANCCASI SIMON	TIEMPO COMPLETO	AUXILIAR	CONTRATADO	ADMINISTRACION Y TURISMO	BACHILLER	LIC.
4	94-202	CHAVEZ RINCON WALTER	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
5	94-204	BLACIDO ALVA LUZ	DEDICACIÓN EXCLUSIVA	ASOCIADO	NOMBRADO	ADMINISTRACION Y TURISMO	BACHILLER	
6	95-101	NIVIN VARGAS LAURA ROSA	DEDICACIÓN EXCLUSIVA	ASOCIADO	NOMBRADO	EDUCACION	BACHILLER	
7	95-107	LOPEZ CRUZ BERTHA	DEDICACIÓN EXCLUSIVA	ASOCIADO	NOMBRADO	EDUCACION	BACHILLER	
8	95-113	GAMARRA FIGUEROA ENCARNACION	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
9	95-115	ZAMBRANO GUIMARAY TOMASA	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
10	95-116	OLORTEGUI MARIÑO RUFA	DEDICACIÓN EXCLUSIVA	PRINCIPAL	NOMBRADO	EDUCACION	BACHILLER	
11	95-117	TARAZONA GARCIA JAVIER	TIEMPO COMPLETO	ASOCIADO	NOMBRADO	EDUCACION	BACHILLER	LIC.

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 37 muestra la consulta de la dimensión Docente de la base de datos DATA MARTUNASAM, en la cual genera 8744 registros.

Figura Nº 37. Consulta a base de DATA MARTUNASAM, tabla Docente



Resultados | Mensajes

	CodigoDocente	NombreDocente	Dedicacion	Categoria	Condicion	DepartamentoAcademico	Grado	F
1	MJ207	VICTOR	DEDICACIÓN EXCLUSIVA	PRINCIPAL	NOMBRADO	CIENCIAS DEL AMBIENTE	BACHILLER	
2	0001	ROBLES TINOCO JAIME	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
3	555PAS	POMA ANCCASI SIMON	TIEMPO COMPLETO	AUXILIAR	CONTRATADO	ADMINISTRACION Y TURISMO	BACHILLER	
4	94-202	CHAVEZ RINCON WALTER	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
5	94-204	BLACIDO ALVA LUZ	DEDICACIÓN EXCLUSIVA	ASOCIADO	NOMBRADO	ADMINISTRACION Y TURISMO	BACHILLER	
6	95-101	NIVIN VARGAS LAURA ROSA	DEDICACIÓN EXCLUSIVA	ASOCIADO	NOMBRADO	EDUCACION	BACHILLER	
7	95-107	LOPEZ CRUZ BERTHA	DEDICACIÓN EXCLUSIVA	ASOCIADO	NOMBRADO	EDUCACION	BACHILLER	
8	95-113	GAMARRA FIGUEROA ENCARNACION	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
9	95-115	ZAMBRANO GUIMARAY TOMASA	TIEMPO PARCIAL	JEFE DE PRÁCTICAS	CONTRATADO	NO ESPECIFICADO	BACHILLER	
10	95-116	OLORTEGUI MARIÑO RUFA	DEDICACIÓN EXCLUSIVA	PRINCIPAL	NOMBRADO	EDUCACION	BACHILLER	
11	95-117	TARAZONA GARCIA JAVIER	TIEMPO COMPLETO	ASOCIADO	NOMBRADO	EDUCACION	BACHILLER	

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión Docente se comprueba que los datos son los esperados.

6.2.4. CASO DE PRUEBA PARA ETL MODALIDAD_APROBACION

A. Caso de prueba

Cuadro Nº 44. Caso de Prueba Nº 04

CASO DE PRUEBA Nº04				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla TIPO_APROBACION sea correctamente cargada de la tabla fuente MODALIDAD_APROBACION.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Semestre de inicio - Semestre fin				
DATOS DE ENTRADA:				
Tabla TIPO_APROBACION				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La dimensión MODALIDAD_APROBACION muestre la misma cantidad de registros que la tabla TIPO_APROBACION.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla TIPO_APROBACION: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 45. Registros de la TIPO_APROBACION

TipoAprobado	Descripción
e	CONVALIDACIONES POR EQUIVALENCIAS
I	TRASLADO INTERNO
M	MATRICULA
S	EXAMEN DE SUFICIENCIA
X	TRASLADO EXTERNO

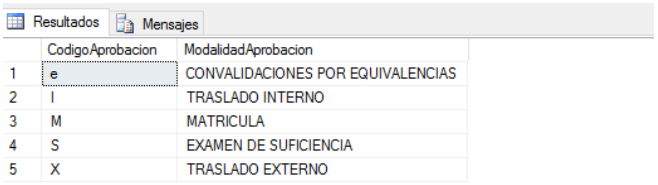
Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 38 muestra la consulta de TIPO_APROBACION a la base de datos ITUNASAM, en la cual genera 5 registros.

Figura Nº 38. Consulta a base de ITUNASAM, tabla MODALIDAD_APROBACION



Resultados | Mensajes

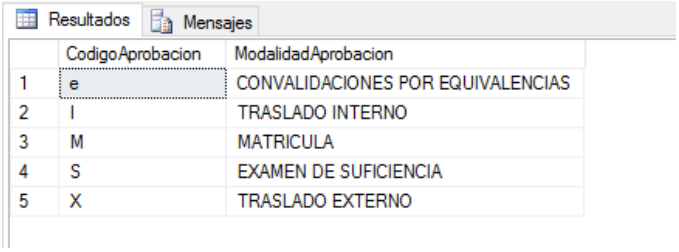
	CodigoAprobacion	ModalidadAprobacion
1	e	CONVALIDACIONES POR EQUIVALENCIAS
2	I	TRASLADO INTERNO
3	M	MATRICULA
4	S	EXAMEN DE SUFICIENCIA
5	X	TRASLADO EXTERNO

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 39 muestra la consulta de la dimensión MODALIDAD_APROBACION de la base de datos DATA MARTUNASAM, en la cual genera 5 registros.

Figura Nº 39. Consulta a base de DATA MARTUNASAM, tabla MODALIDAD_APROBACION



Resultados | Mensajes

	CodigoAprobacion	ModalidadAprobacion
1	e	CONVALIDACIONES POR EQUIVALENCIAS
2	I	TRASLADO INTERNO
3	M	MATRICULA
4	S	EXAMEN DE SUFICIENCIA
5	X	TRASLADO EXTERNO

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión MODALIDAD_APROBACION se comprueba que los datos son los esperados.

6.2.5. CASO DE PRUEBA PARA ETL SEMESTRE

A. Caso de prueba

Cuadro Nº 46. Caso de Prueba Nº 05

CASO DE PRUEBA Nº05				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla SEMESTRE sea correctamente cargada de la tabla fuente SEMESTRE.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Semestre de inicio - Semestre fin				
DATOS DE ENTRADA:				
- Tabla SEMESTRE				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La dimensión SEMESTRE muestre la misma cantidad de registros que la tabla SEMESTRE.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla SEMESTRE: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 47. Registros de la Tabla SEMESTRE

Semestre	Descripcion
2000-0	SEMESTRE 2000-0
2000-1	SEMESTRE 2000-1
2000-2	SEMESTRE 2000-2
2001-0	SEMESTRE 2001-0
2001-1	SEMESTRE 2001-1
2001-2	SEMESTRE 2001-2
2002-0	SEMESTRE 2002-0

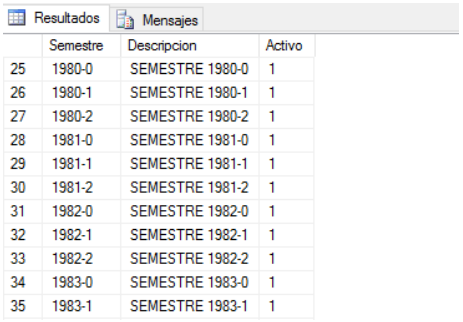
Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 40 muestra la consulta de SEMESTRE a la base de datos ITUNASAM, en la cual genera 118 filas.

Figura Nº 40. Consulta a base de ITUNASAM, tabla SEMESTRE



Resultados | Mensajes

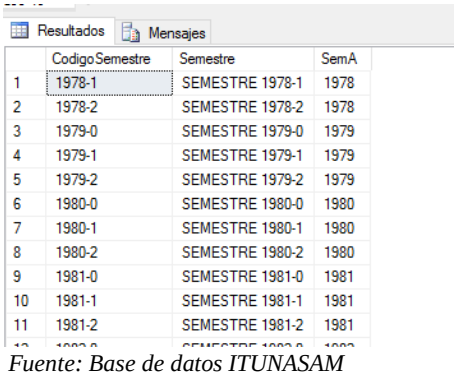
	Semestre	Descripcion	Activo
25	1980-0	SEMESTRE 1980-0	1
26	1980-1	SEMESTRE 1980-1	1
27	1980-2	SEMESTRE 1980-2	1
28	1981-0	SEMESTRE 1981-0	1
29	1981-1	SEMESTRE 1981-1	1
30	1981-2	SEMESTRE 1981-2	1
31	1982-0	SEMESTRE 1982-0	1
32	1982-1	SEMESTRE 1982-1	1
33	1982-2	SEMESTRE 1982-2	1
34	1983-0	SEMESTRE 1983-0	1
35	1983-1	SEMESTRE 1983-1	1

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 41 muestra la consulta de la dimensión SEMESTRE de la base de datos DATA MARTUNASAM, en la cual genera 118 filas.

Figura Nº 41. Consulta a base de DATA MARTUNASAM, tabla SEMESTRE



Resultados | Mensajes

	CodigoSemestre	Semestre	SemA
1	1978-1	SEMESTRE 1978-1	1978
2	1978-2	SEMESTRE 1978-2	1978
3	1979-0	SEMESTRE 1979-0	1979
4	1979-1	SEMESTRE 1979-1	1979
5	1979-2	SEMESTRE 1979-2	1979
6	1980-0	SEMESTRE 1980-0	1980
7	1980-1	SEMESTRE 1980-1	1980
8	1980-2	SEMESTRE 1980-2	1980
9	1981-0	SEMESTRE 1981-0	1981
10	1981-1	SEMESTRE 1981-1	1981
11	1981-2	SEMESTRE 1981-2	1981

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión SEMESTRE se comprueba que los datos son los esperados.

6.2.6. CASO DE PRUEBA PARA EL HECHO COMPORTAMIENTO ACADÉMICO ESTUDIANTE

A. Caso de prueba

Cuadro Nº 48. CASO DE PRUEBA Nº 06

CASO DE PRUEBA Nº06				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla **ComAcadAlumno** sea correctamente cargada de la tabla fuente **Rendimiento**.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Semestre de inicio - Semestre fin				
DATOS DE ENTRADA:				
- Tabla RENDIMIENTO				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La dimensión ComAcadAlumno muestre la misma cantidad de registros que la tabla DOCENTE.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla RENDIMIENTO: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 49. Registros de la Tabla RENDIMIENTO

Sede	Escuela	Curso	Semestre	Seccion	Alumno	TipoAprobado	Grupo	Promedio
1	1	QU-S02	1997-1	0	97.0001.1.E2	M	1	14
1	1	QU-S02	1997-1	0	97.0002.6.E2	M	1	13
1	1	QU-S02	1997-1	0	97.0003.0.E2	M	1	12
1	1	QU-S02	1997-1	0	97.0004.5.E2	M	1	12
1	1	QU-S02	1997-1	0	97.0016.7.A2	M	1	11
1	1	QU-S02	1997-1	0	97.0019.0.A2	M	1	14
1	1	QU-S02	1997-1	0	97.0025.5.A2	M	1	0

Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 42 muestra la consulta del script para el comportamiento académico a la base de datos ITUNASAM, en la cual genera 1347281 filas.

Figura Nº 42. Consulta a base de ITUNASAM, tabla Curricula

Resultados | Mensajes

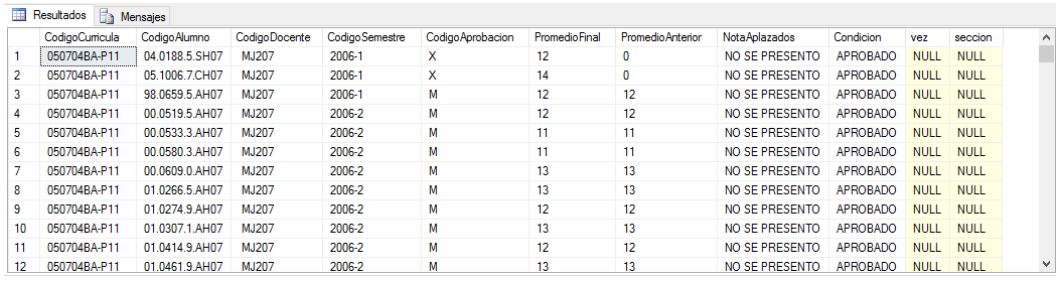
	facul	(Sin nombre de columna)	doc	Semestre	TipoAprobado	Promedio	PromedioAntes	aplazado	condicion	vez
1	112201EN-304	94.1557.6.AQ22	AUX	1996-1	M	0	0	NO SE PRESENTO	DESAPROBADO	1
2	112201EN-304	95.1225.9.AQ22	AUX	1996-1	M	13	13	NO SE PRESENTO	APROBADO	1
3	112202UH-A01	01.0076.1.AQ22	AUX	2004-2	M	12	12	NO SE PRESENTO	APROBADO	1
4	112702UE-S03	96.1524.3.AI27	AUX	1997-1	M	13	13	NO SE PRESENTO	APROBADO	1
5	112702UE-S03	96.3000.1.AI27	AUX	1997-1	M	16	16	NO SE PRESENTO	APROBADO	1
6	112702UE-S03	96.3099.7.AI27	AUX	1997-1	M	14	14	NO SE PRESENTO	APROBADO	1
7	112702UE-S03	96.3127.7.AI27	AUX	1997-1	M	15	15	NO SE PRESENTO	APROBADO	1
8	112702UE-S03	96.3132.7.AI27	AUX	1997-1	M	14	14	NO SE PRESENTO	APROBADO	1
9	112702UE-S03	96.3195.7.AI27	AUX	1997-1	M	0	0	NO SE PRESENTO	DESAPROBADO	1
10	112702UE-S03	96.3200.8.AI27	AUX	1997-1	M	13	13	NO SE PRESENTO	APROBADO	1
11	112702UE-S03	96.3953.2.UI27	AUX	1997-1	M	14	14	NO SE PRESENTO	APROBADO	1
12	112702UH-A01	96.1168.N.AI27	AUX	1999-1	M	13	13	NO SE PRESENTO	APROBADO	1

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 43 muestra la consulta de la dimensión ComAcadAlumno de la base de datos DATA MARTUNASAM, en la cual genera 1347281 filas.

Figura Nº 43. Consulta a base de DATA MARTUNASAM, tabla ComAcadAlumno



Resultados | Mensajes

	CodigoCurricula	CodigoAlumno	CodigoDocente	CodigoSemestre	CodigoAprobacion	PromedioFinal	PromedioAnterior	NotaAplazados	Condicion	vez	seccion
1	050704BA-P11	04.0188.5.SH07	MJ207	2006-1	X	12	0	NO SE PRESENTO	APROBADO	NULL	NULL
2	050704BA-P11	05.1006.7.CH07	MJ207	2006-1	X	14	0	NO SE PRESENTO	APROBADO	NULL	NULL
3	050704BA-P11	98.0659.5.AH07	MJ207	2006-1	M	12	12	NO SE PRESENTO	APROBADO	NULL	NULL
4	050704BA-P11	00.0519.5.AH07	MJ207	2006-2	M	12	12	NO SE PRESENTO	APROBADO	NULL	NULL
5	050704BA-P11	00.0533.3.AH07	MJ207	2006-2	M	11	11	NO SE PRESENTO	APROBADO	NULL	NULL
6	050704BA-P11	00.0580.3.AH07	MJ207	2006-2	M	11	11	NO SE PRESENTO	APROBADO	NULL	NULL
7	050704BA-P11	00.0609.0.AH07	MJ207	2006-2	M	13	13	NO SE PRESENTO	APROBADO	NULL	NULL
8	050704BA-P11	01.0266.5.AH07	MJ207	2006-2	M	13	13	NO SE PRESENTO	APROBADO	NULL	NULL
9	050704BA-P11	01.0274.9.AH07	MJ207	2006-2	M	12	12	NO SE PRESENTO	APROBADO	NULL	NULL
10	050704BA-P11	01.0307.1.AH07	MJ207	2006-2	M	13	13	NO SE PRESENTO	APROBADO	NULL	NULL
11	050704BA-P11	01.0414.9.AH07	MJ207	2006-2	M	12	12	NO SE PRESENTO	APROBADO	NULL	NULL
12	050704BA-P11	01.0461.9.AH07	MJ207	2006-2	M	13	13	NO SE PRESENTO	APROBADO	NULL	NULL

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión ComAcadAlumno se comprueba que los datos son los esperados.

6.2.7. CASO DE PRUEBA PARA EL HECHO COMPORTAMIENTO ACADÉMICO DOCENTE

A. Caso de prueba

Cuadro Nº 50. CASO DE PRUEBA Nº 07

CASO DE PRUEBA Nº07				
OBJETO A ANALIZAR DEL CASO DE PRUEBA				
Proceso ETL y Base de Datos del Data Mart.				
OBJETIVO DEL CASO DE PRUEBA:				
Verificar que la tabla CURRICULA sea correctamente cargada de la tabla fuente DOCENTE.				
PRE-REQUISITOS DEL CASO DE PRUEBA:				
- Disponibilidad de la BD Fuente. - BD Data Mart creada. - El script debe estar creado y listo para usarse.				
PARÁMETROS REQUERIDOS:				
- Semestre de inicio - Semestre fin				
DATOS DE ENTRADA:				
- Tabla ENCUESTA - Tabla PREGUNTA - Tabla RESPUESTA				
Nº	**FUNCIÓN A PROBAR**	**ACCIÓN**	**RESULTADO ESPERADO**	**RESULTADO OBTENIDO**
01	Cantidad de registros por tabla en la BD Fuente frente a la cantidad de registros por tabla en BD Data Mart.	Ejecutar el script	La dimensión DOCENTE muestre la misma cantidad de registros que la tabla DOCENTE.	Se cargó los datos correctamente.

Fuente: Elaboración Propia.

B. Ejecución del caso de prueba

- **Datos de Entrada:**

Tabla ENCUESTA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 51. Registros de la Tabla ENCUESTA

Encuesta	Descripcion	Semestre
1	PRIMERA ENCUESTA ALUMNOS (2007-1)	2007-1
2	SEGUNDA ENCUESTA PARA ALUMNOS (2007-2)	2007-2
3	TERCERA ENCUESTA (2008-1)	2008-1
4	CUARTA ENCUESTA (2008-2)	2008-2
5	QUINTA ENCUESTA (2009-1)	2009-1
6	SEXTA ENCUESTA (2009-2)	2009-2

Fuente: Base de datos ITUNASAM

Tabla PREGUNTA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 52. Registros de la Tabla PREGUNTA

Pregunta	Descripcion	Observacion
1	ENTREGÓ EL SILABO Y DESARROLA EL CURSO DE ACUERDO A LO PROGRAMADO	La presentación del silabo debe hacerse oportunamente al Departamento Académico y la Escuela Profesional correspondiente y durante el desarrollo del curso ceñirse de acuerdo a lo programado
2	ASISTE, INICIA Y TERMINA SUS CLASES PUNTUALMENTE	Se le recomienda llegar puntualmente al desarrollo de sus clases y cumplir con el tiempo programado según el horario correspondiente
3	PLANIFICA, IMPLEMENTE Y CONDUCE EL PROCESO DE ENSEÑANZA-APRENDIZAJE ADECUADAMENTE	Es responsabilidad del docente universitario preparar sus clases con antelación, no improvisarlas y de ese modo demostrar un adecuado desempeño
4	AL INICIO DE LA CLASE COMUNICA LOS OBJETIVOS Y/O COMPETENCIAS A LOGRAR	El docente debe explicar al inicio de la clase los contenidos y actividades que desarrollará durante la sesión de aprendizaje
5	DEMUESTRA DOMINIO AMPLIO Y PROFUNDO DE LA ASIGNATURA A SU RESPONSABILIDAD	Es necesario que usted se prepare convenientemente para un dominio amplio y profundo de los temas que desarrolla en clase
6	MANEJA ESTRATEGIAS METODOLÓGICAS ADECUADAS AL NIVEL UNIVERSITARIO	Usted debe desarrollar sus sesiones de aprendizaje empleando adecuadas estrategias, métodos y técnicas de enseñanza como: la clase magistral, el seminario, los talleres y otras adecuadas a su asignatura
7	SATISFACE LAS EXPECTATIVAS ACADÉMICAS DE LOS ALUMNOS	Para que el alumno esté satisfecho con su desempeño académico debe absolver todas sus dudas y demostrar solvencia académica en el desarrollo de su materia
8	ESTIMULA LA PARTICIPACIÓN EN CLASE, EL TRABAJO EN EQUIPO Y EL LIDERAZGO	Para el caso debe usted desarrollar las sesiones de aprendizaje empleando métodos activos y participativos que propicien la participación en clase, el trabajo en equipo y el liderazgo
9	UTILIZA MEDIOS Y MATERIALES EDUCATIVOS ADECUADOS INCLUYENDO EL SOPORTE INFORMÁTICO	Se debe incorporar el uso desde la programación del silabo, durante las tareas académicas y las sesiones de clase de material bibliográfico, e-mail, Chat, Web, equipos multimediales y otros en el desarrollo del curso
10	MOTIVA A TENER UNA ACTITUD DE INVESTIGACIÓN Y BÚSQUEDA DE INFORMACIÓN CIENTÍFICA POR DIFERENTES MEDIOS	En este ítem usted debe demostrar que se actualiza permanentemente, presentando información novedosa o en su defecto resultados de sus investigaciones o los avances de la ciencia y la tecnología

Fuente: Base de datos ITUNASAM

Tabla RESPUESTA: Se listan las columnas significativas para el proceso y una porción del número de registros.

Cuadro Nº 53. Registros de la Tabla RESPUESTA

Sede	Escuela	Curricula	Curso	Semestre	Seccion	Encuesta	Pregunta	Valoracion
1	1	3	AF-A03	2007-1	1	1	1	4
1	1	3	AF-A03	2007-1	1	1	2	4
1	1	3	AF-A03	2007-1	1	1	3	4
1	1	3	AF-A03	2007-1	1	1	4	4
1	1	3	AF-A03	2007-1	1	1	5	5
1	1	3	AF-A03	2007-1	1	1	6	5
1	1	3	AF-A03	2007-1	1	1	7	4
1	1	3	AF-A03	2007-1	1	1	8	5
1	1	3	AF-A03	2007-1	1	1	9	5
1	1	3	AF-A03	2007-1	1	1	10	4
1	1	3	AF-A03	2007-1	1	1	11	4
1	1	3	AF-A03	2007-1	1	1	12	5

Fuente: Base de datos ITUNASAM

- **Resultado final:**

  - **Tabla Fuente**

    En la figura Nº 44 muestra la consulta al script de carga de datos de Comportamiento Docente a la base de datos ITUNASAM, en la cual genera 448946 filas.

Figura Nº 44. Consulta a base de ITUNASAM, al script de Comportamiento Docente

Resultados | Mensajes

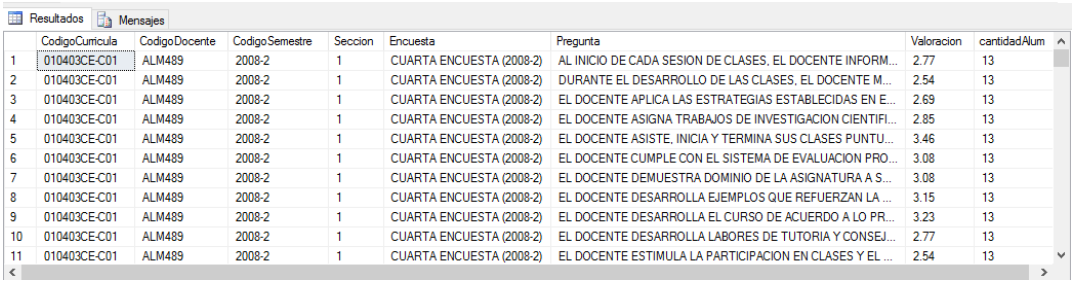
	CodigoCurricula	CodigoDocente	CodigoSemestre	Seccion	Encuesta	Pregunta	(Sin nomb
1	060802MT-T02	PFJ247	2007-1	1	PRIMERA ENCUESTA ALUMNOS (2007-1)	PRACTICA CON EL EJEMPLO LOS VALORES HUMANOS DE: RE...	3.00000
2	010403CE-E06	SMM395	2010-1	1	SEPTIMA ENCUESTA (2010-1)	EL DOCENTE DESARROLLA EL CURSO DE ACUERDO A LOS C...	3.83333
3	020304AH-R16	CVJ127	2009-1	1	QUINTA ENCUESTA (2009-1)	EL DOCENTE AYUDA Y ORIENTA MEDIANTE ACCIONES DE TU...	4.22222
4	010503CM-I01	CWM001	2010-2	1	OCTAVA ENCUESTA (2010-2)	EL DOCENTE CUMPLE CON EL DISEÑO DE EVALUACION PRO...	4.50000
5	101604DI-B02	SVL243	2008-1	1	TERCERA ENCUESTA (2008-1)	RESPECTO A TU DESEMPEÑO COMO ESTUDIANTE EN LA ASI...	3.66666
6	031304NA-C07	RAE603	2007-2	1	SEGUNDA ENCUESTA PARA ALUMNOS (2007-2)	DESARROLLA EL CURSO DE ACUERDO A LO PROGRAMADO E...	3.85714
7	112103UE-T30	PAD024	2007-2	1	SEGUNDA ENCUESTA PARA ALUMNOS (2007-2)	PRACTICA CON EL EJEMPLO LOS VALORES HUMANOS DE: RE...	4.00000
8	012502250234	GAH290	2013-2	1	DECIMOCUARTA ENCUESTA (2013-2)	EL DOCENTE ASIGNA A LOS ESTUDIANTES EL DESARROLLO ...	3.35714
9	012502250212	RFJR01	2013-1	1	DECIMOTERCERA ENCUESTA (2013-1)	EL DOCENTE FOMENTA LA IDENTIFICACION DEL ESTUDIANT...	3.90000
10	020304AM-D04	MLV036	2009-2	1	SEXTA ENCUESTA (2009-2)	EL DOCENTE REALIZA ACCIONES DE MOTIVACION PERMANE...	3.03448
11	113401340165	GVJA01	2014-2	1	DECIMOSEXTA ENCUESTA (2014-2)	EL DOCENTE REALIZA ACCIONES DE MOTIVACION PERMANE...	4.61538

Fuente: Base de datos ITUNASAM

  - **Tabla final**

    En la figura Nº 45 muestra la consulta de la dimensión Comportamiento docente de la base de datos DATA MARTUNASAM, en la cual genera 448946 filas.

Figura Nº 45. Consulta a base de DATA MARTUNASAM, tabla COMDOC



Resultados | Mensajes

	CodigoCurricula	CodigoDocente	CodigoSemestre	Seccion	Encuesta	Pregunta	Valoracion	cantidadAlum
1	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	AL INICIO DE CADA SESION DE CLASES, EL DOCENTE INFORM...	2.77	13
2	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	DURANTE EL DESARROLLO DE LAS CLASES, EL DOCENTE M...	2.54	13
3	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE APLICA LAS ESTRATEGIAS ESTABLECIDAS EN E...	2.69	13
4	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE ASIGNA TRABAJOS DE INVESTIGACION CIENTIFI...	2.85	13
5	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE ASISTE, INICIA Y TERMINA SUS CLASES PUNTU...	3.46	13
6	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE CUMPLE CON EL SISTEMA DE EVALUACION PRO...	3.08	13
7	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE DEMUESTRA DOMINIO DE LA ASIGNATURA A S...	3.08	13
8	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE DESARROLLA EJEMPLOS QUE REFUERZAN LA ...	3.15	13
9	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE DESARROLLA EL CURSO DE ACUERDO A LO PR...	3.23	13
10	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE DESARROLLA LABORES DE TUTORIA Y CONSEJ...	2.77	13
11	010403CE-C01	ALM489	2008-2	1	CUARTA ENCUESTA (2008-2)	EL DOCENTE ESTIMULA LA PARTICIPACION EN CLASES Y EL ...	2.54	13

Fuente: Base de datos ITUNASAM

Finalmente, se revisó los contenidos de la Dimensión COMDOC se comprueba que los datos son los esperados.

CAPÍTULO VII

IMPLEMENTACIÓN

En el presente capitulo se aborda la estrategia para la implantación del Data Mart, identificando los distintos recursos, infraestructura, equipamiento y redes que son requeridos para la implementación.

7.1. MONITOREO Y EVALUACIÓN DE LA SOLUCIÓN.

La solución de inteligencia de negocio planteado para la Facultad de Ciencias está conformada por diferentes elementos que conforman parte de la solución, para el correcto funcionamiento de las operaciones debe de permanecer en un estado óptimo, y para ello se planteó el monitoreo y evolución de la solución teniendo en cuenta los elementos, las políticas y reglas de procedimientos.

7.1.1. ELEMENTOS DEL MONITOREO Y EVALUACIÓN.

Para la implementación, de la solución de Inteligencia, existen elementos de monitoreo, tales como el Personal capacitado, la disponibilidad y correcto funcionamiento de las instalaciones; teniendo en cuenta el tiempo y el costo al que incurre este proceso.

7.1.2. POLÍTICAS Y REGLAS DE PROCEDIMIENTO

Las políticas y reglas de procedimientos a tomar en cuenta para el correcto funcionamiento serán planteados mensualmente y anualmente por el jefe de área inmediato tal como el decano o el jefe de OGE.

1. En las direcciones de Escuela, será responsabilidad de los directores de escuela garantizar el correcto funcionamiento de los equipos de cómputo.

2. En la oficina general de estudio, será responsabilidad del jefe de OGE solicitar el mantenimiento preventivo del equipo de cómputo.

3. El mantenimiento preventivo de los equipos se ajusta al plan anual de la oficina de informática de la Universidad Nacional Santiago Antúnez de Mayolo.

4. En caso de descubrirse fallas, el responsable del proyecto será notificado inmediatamente para efectuar la reparación que corresponda.

5. Las observaciones y resultados que no estén dentro de los parámetros esperados, se tomarán de manera directa con el administrador del Data Mart, para su conocimiento y corrección correspondiente.

6. Revisar la disponibilidad de espacio y desfragmentación de disco mensualmente en el servidor.

7. El jefe de OGE deberá establecer fechas de actividades de monitoreo periódico, con el fin de detectar problemas en el funcionamiento de este.

8. Los plazos de actualización de datos del Data Mart serán establecidos por la oficina general de estudios.

9. Revisar el estado general del sistema operativo (conexión a red, actualización de parches de seguridad), cada semana y bajo demanda.

10. En el caso de la herramienta de filtrado de sitios y contenido de Internet, éste deberá depurarse periódicamente.

11. Revisar contra la infección de virus y malware (adware, spyware, etc.) semanalmente.

7.1.3. PLAN DE MONITOREO Y EVALUACIÓN.

El proceso de monitoreo y evaluación involucra a distintos actores, siendo el principal responsable la oficina general de Estudio representado por el jefe de OGE; los principales resultados se evidenciarán durante el uso del Data Mart, en cuadro Nº 54- Plan de Monitoreo y evaluación.

Cuadro Nº 54. Plan de Monitoreo y evaluación

Lógica de Intervención	Indicador	Responsable	Actividad	Tiempo
Implementación del Data Mart.	(Número de consultas implementadas/Número de consultas programadas) x100	Jefe de Proyecto	Procesos de Migración de Datos. Proceso de evolución de datos. Levantamiento de la Base de Datos del Data Mart. Puesta en marcha de la herramienta web.	7 días. 7 días. 2 días. 5 días.
Mejora de la toma de decisiones en la gestión académica	(Número de consultas implementadas para la toma de decisiones/Número de consultas programadas para la toma de decisiones) x100	Jefe de Proyecto Directores de Escuela.	Grado de soporte a la toma de decisiones en la gestión Académica.	Permanente

Fuente: Elaboración Propia.

Tener un plan de evaluación a largo plazo ayudará a tomar decisiones futuras que tengan que ver con los criterios de evaluación, además se debe de incluir a todos los actores del proceso.

7.2. BITÁCORA Y PUESTA A PUNTO.

La implantación de la solución de inteligencia de Negocio-Data Mart se realizó correctamente teniendo en cuenta los siguientes pasos que se describen a continuación:

7.2.1. MIGRACIÓN Y CARGA INICIAL DE DATOS.

Esta actividad se lleva a cabo cuando es necesario la inicial de datos o la actualización de datos al Data Mart.

1. Se solo una vez a los usuarios del Data Mart, que podrán realizar algún tipo de consulta.

2. Solo tendrán acceso al Data Mart los usuarios que pertenezcan a la Dirección de Escuela, jefes de departamento, decanos y jefe de OGE.

3. La carga de datos se realiza después de finalizado el semestre académico.

Estas consideraciones deben de ser tomadas en cuenta al realizar la carga inicial de datos.

7.2.2. APROBACIÓN DE LA SOLUCIÓN TECNOLÓGICA

La aprobación de la solución tecnológica dependerá de la oficina de Informática de la Universidad Nacional Santiago Antúnez de Mayolo. Teniendo en cuenta ello el presente trabajo de investigación será aprobado en base a la defensa del mismo ante un jurado designado.

CAPÍTULO VIII

RESULTADOS

Con la finalidad de cumplir con los objetivos propuestos, toda la hipótesis de un trabajo de investigación debe ser válida y por ello la experimentación del trabajo de investigación se realizó en la Facultad de Ciencias en las Escuelas profesionales de Ingeniería de Sistemas e Informática, Estadística e Informática y Matemática.

Los resultados obtenidos del Post-Test, tanto del grupo experimental que está representado por la escuela de Ingeniería de Sistemas e Informática y el grupo control por la escuela de Estadística e Informática y la escuela de Matemática, para mostrar los logros obtenidos en el grupo experimental, luego de aplicar la variable independiente; así de esta manera validar la hipótesis formulada.

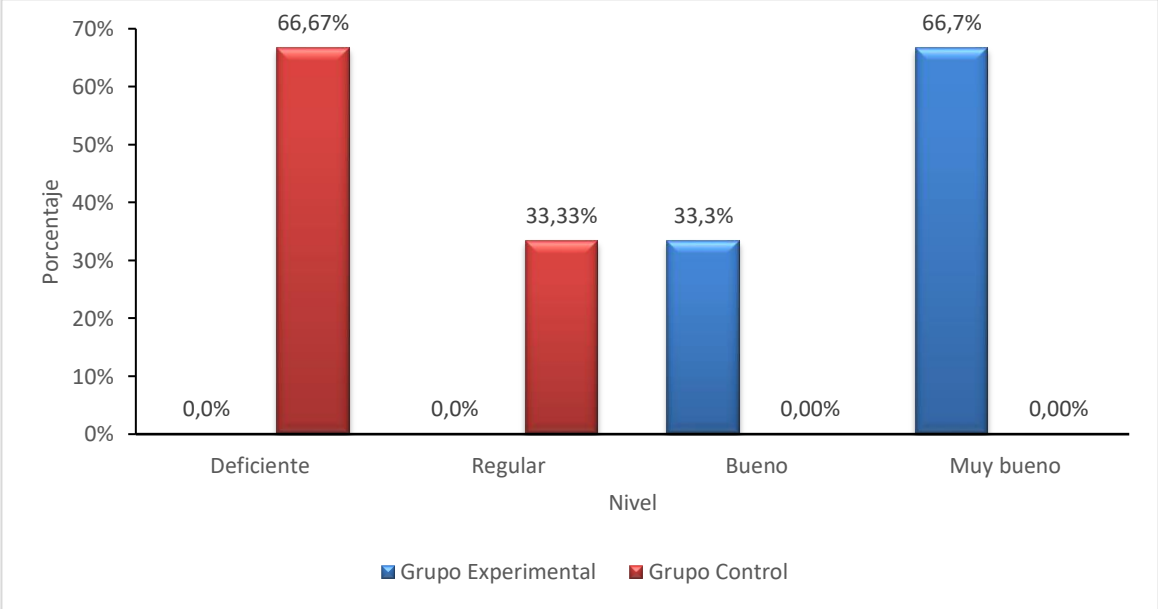
Luego de haber recolectado la información detallada sobre la aplicación del Data Mart, a través de una escala de Estimación por parte del investigador en la Facultad de ciencias de UNASAM, los datos se tabularon y se analizaron, obteniendo los resultados mostrados en los siguientes cuadros.

Cuadro Nº 55. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control.

Nivel	Intervalo	Xi	Grupo Experimental		Grupo Control	
			Absoluta f_i	Porcentual $p_i\%$	Absoluta f_i	Porcentual $p_i\%$
Deficiente	[0-30>	15	0	0%	2	66,67%
Regular	[31-61>	45	0	0%	1	33,33%
Bueno	[62-92>	75	1	33,3%	0	0%
Muy Bueno	[93-123>	105	2	66,7%	0	0%
Total			3	100%	3	100%
Media	$\overline{X}$		91		32	
Moda	M_O		98		30	
Mediana	Me		98		30	
Varianza	S^2		147		12	
Desviación	S		12,1		3,4	

Fuente: Aplicación de la escala de estimación en los grupos experimental y control.

Gráfico Nº 01. Gráfico de barras de la frecuencia porcentual del post-test del grupo experimental y control.

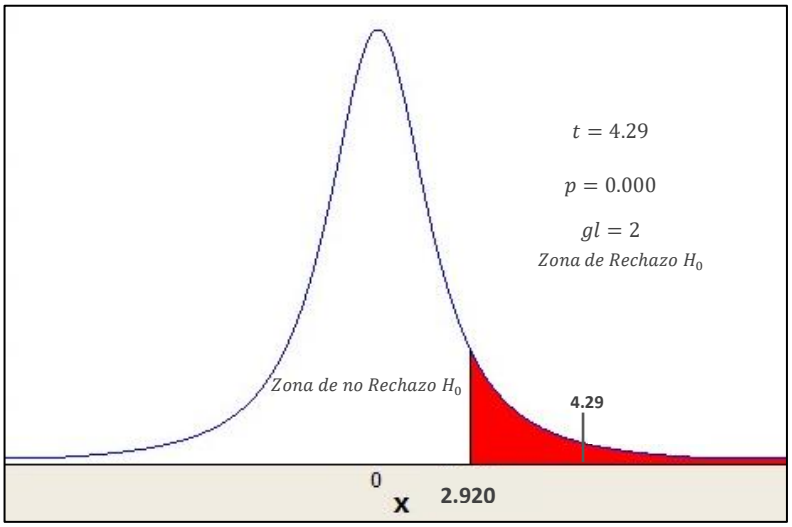


Fuente: Cuadro Nº 55. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control.

DESCRIPCIÓN

En el cuadro Nº 55 y Gráfico Nº 01, se observa que el grupo experimental está representado por un porcentaje 66.7% de los tomadores de decisiones se encuentra en un nivel muy bueno en cuanto a la toma de decisiones de la gestión académica después de la aplicación del Data Mart en comparación al grupo control, la cual está ubicada en un nivel deficiente con 66,7%, por otro lado, el 33,3% muestra un nivel regular. Con una media ($\overline{X}$) de 32.00 puntos en el post-test del grupo Control lo que significa que los tomadores de decisiones se encuentran en un nivel deficiente en cuanto a la toma de decisiones de la gestión académica, mientras media ($\overline{X}$) de 91.00 puntos en el post-test del grupo experimenta evidencia que los tomadores de decisiones se encuentran en un nivel muy Bueno en cuanto a la toma de decisiones de la gestión académica.

Gráfico Nº 02. Prueba de Hipótesis del grupo experimental.



Fuente: Cuadro Nº 55. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control.

La prueba t-student indica que $t = 4.29 > t_{(0.95,2)} = 2.920$ cae en la zona de rechazo de la hipótesis nula, y su p-valor es altamente significativo ($p = 0.000 < 0.05$) razón por la cual se afirma que el Data Mart como una solución de inteligencia

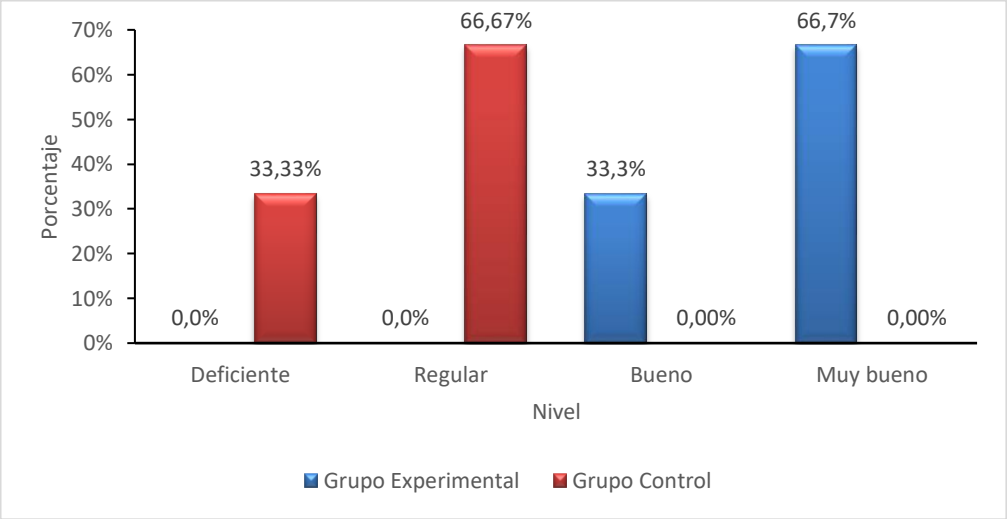
de negocios influye en la toma de decisiones de la gestión académica, la cual es mostrada en la Gráfico Nº 02.

Cuadro Nº 56. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control de la dimensión diseño

Nivel	Intervalo	Xi	Grupo Experimental		Grupo Control	
			Absoluta f_i	Porcentual $p_i\%$	Absoluta f_i	Porcentual $p_i\%$
Deficiente	[0-17>	8,5	0	0,0%	1	33,33%
Regular	[18-35>	26,5	0	0,0%	2	66,67%
Bueno	[36-53>	44,5	1	33,3%	0	0,00%
Muy Bueno	[54-69>	61,5	2	66,7%	0	0,00%
Total			3	100%	3	100%
Media		$\bar{X}$	53,0		18,3	
Mediana		Me	57,0		18,0	
Varianza		S^2	61,0		6,3	
Desviación		S	7,8		2,5	

Fuente: Aplicación de la escala de estimación en los grupos experimental y control.

Gráfico Nº 03. Gráfico de barras de la frecuencia porcentual del post-test del grupo experimental y control.



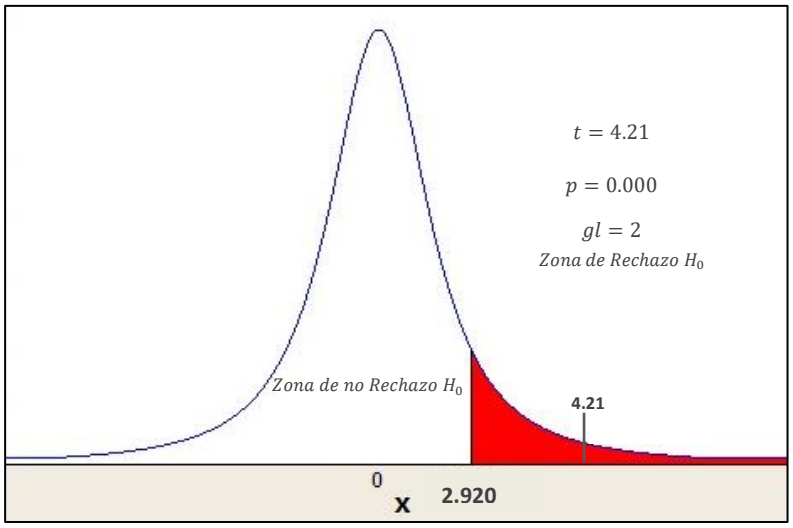
Fuente: Cuadro Nº 56. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control de la dimensión diseño.

DESCRIPCIÓN

En el cuadro N° 56 y gráfico Nº 03, se observa que el grupo experimental está representado por un porcentaje 66.7% de los tomadores de decisiones, que se encuentra en un nivel muy bueno en cuanto al diseño de las decisiones de la gestión académica después de la aplicación del Data Mart en comparación al grupo control, la cual está ubicada en un nivel regular con 66,7%, por otro lado, el 33,3% muestra un nivel Deficiente. Con una media ($\overline{X}$) de 18.3 puntos en el post-test del grupo Control lo que significa que los tomadores de decisiones se encuentran en un nivel deficiente en cuanto al diseño de la toma de decisiones de la gestión académica, mientras media ($\overline{X}$) de 53.00 puntos en el post-test del grupo experimenta evidencia que los tomadores de decisiones se encuentran en un nivel muy Bueno en cuanto a la toma de decisiones de la gestión académica.

La prueba t-student indica que $t = 4.21 > t_{(0.95,2)} = 2.920$ cae en la zona de rechazo de la hipótesis nula, y su p-valor es altamente significativo ($p = 0.000 < 0.05$) razón por la cual se afirma que el Data Mart como una solución de inteligencia de negocios influye en el diseño de las decisiones de la gestión académica.

Gráfico Nº 04. Prueba de Hipótesis del grupo experimental, dimensión Diseño.



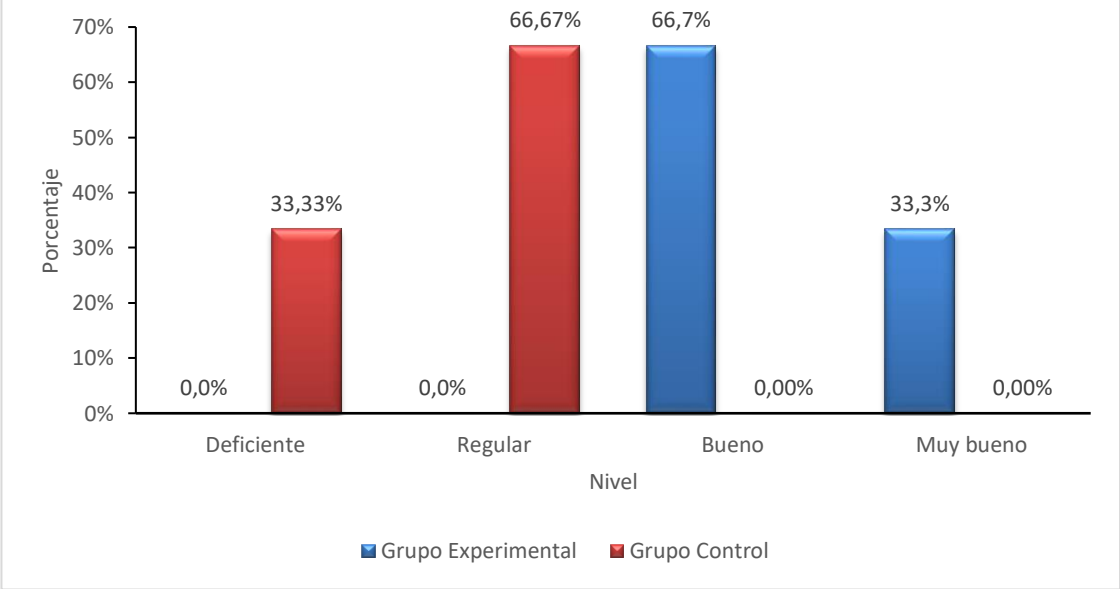
Fuente: Cuadro Nº 56. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control de la dimensión diseño.

Cuadro Nº 57. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control dimensión Elección.

Nivel	Intervalo	Xi	Grupo Experimental		Grupo Control	
			Absoluta f_i	Porcentual $p_i\%$	Absoluta f_i	Porcentual $p_i\%$
Deficiente	[0-13>	6,5	0	0,0%	1	33,33%
Regular	[14-27>	20,5	0	0,0%	2	66,67%
Bueno	[28-41>	34,5	2	66,7%	0	0,00%
Muy Bueno	[42-54>	48	1	33,3%	0	0,00%
Total			3	100%	3	100%
Media $\overline{X}$			38		13,67	
Mediana Me			40,0		14	
Varianza S^2			19,0		2,3	
Desviación S			4,4		1,53	

Fuente: Aplicación de la escala de estimación en los grupos experimental y control.

Gráfico Nº 05. Gráfico de barras de la frecuencia porcentual del post-test del grupo experimental y control dimensión Elección.



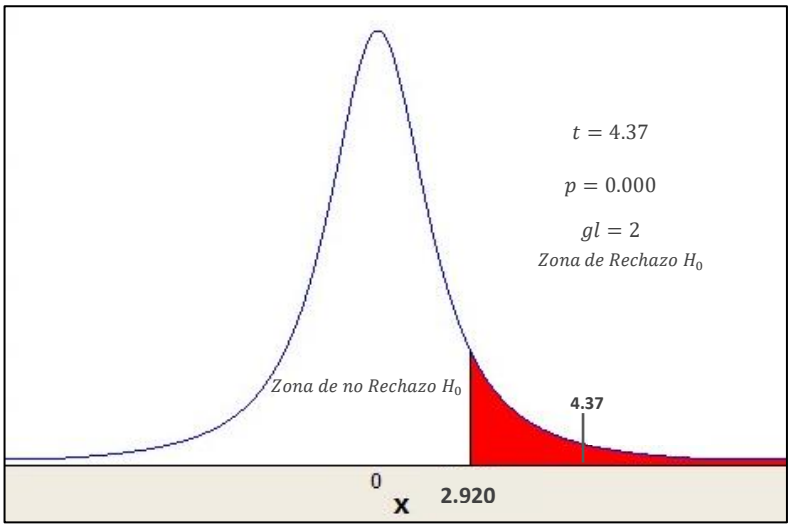
Fuente: Cuadro Nº 57. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control dimensión Elección.

DESCRIPCIÓN

En el cuadro Nº 57 y gráfico Nº 05, se observa que el grupo experimental está representado por un porcentaje 33.3% de los tomadores de decisiones se encuentra en un nivel muy bueno en cuanto a la elección de las decisiones de la gestión académica después de la aplicación del Data Mart en comparación al grupo control, la cual está ubicada en un nivel regular con 66,7%, por otro lado, el 33,3% muestra un nivel deficiente. Con una media ($\overline{X}$) de 13.00 puntos en el post-test del grupo Control lo que significa que los tomadores de decisiones se encuentran en un nivel deficiente en cuanto a la elección de las decisiones de la gestión académica, mientras media ($\overline{X}$) de 38.00 puntos en el post-test del grupo experimenta evidencia que los tomadores de decisiones se encuentran en un nivel muy Bueno en cuanto a la elección de las decisiones de la gestión académica.

La prueba t-student indica que $t = 4.37 > t_{(0.95,2)} = 2.920$ cae en la zona de rechazo de la hipótesis nula, y su p-valor es altamente significativo ($p = 0.000 < 0.05$) razón por la cual se afirma que el Data Mart como una solución de inteligencia de negocios influye en la elección de las decisiones de la gestión académica.

Gráfico Nº 06. Prueba de Hipótesis del grupo experimental, dimensión Elección.



Fuente: Cuadro Nº 57. Distribución de Frecuencias por Nivel y mediadas de tendencia central del post-test del grupo experimental y grupo control dimensión Elección.

CAPÍTULO IX

DISCUSIÓN DE RESULTADOS

Se presenta a continuación los resultados del estudio referido a la aplicación de un Data Mart, como parte de una solución de inteligencia de negocio, para mejorar el proceso de Toma de Decisiones en La Gestión Académica de Facultad de Ciencias de la UNASAM. Los datos a discutir se han ordenado de acuerdo a los objetivos del trabajo las mismas que se presenta a continuación.

Al nivel de análisis aplicado de la hipótesis general "Con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM". En el cuadro Nº 55, gráfico Nº 01 y gráfico Nº 02 se puede observar que el 66.7% de los tomadores de decisiones muestran un nivel muy bueno en cuanto a la toma de decisiones, por otro lado, el 33.3% de los muestran un nivel de bueno en cuanto a la toma de decisiones. La prueba de hipótesis p=0.000 lo que indica que se aprueba la hipótesis general,

por tal razón se afirma que con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora significativamente el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM. Según Inmin(2005) Un Data Warehouse es una colección de datos integrados, temáticos, no volátiles y variantes en el tiempo, organizados para soportar necesidades empresariales orientadas a la toma de decisiones. Mazón, Pardillo y Trujillo (2011) Los almacenes de datos posibilitan un visón multidimensional de enormes cantidades de datos históricos provenientes de fuentes operacionales, suministrando la información necesaria para el apoyo a los procesos de toma de decisiones de una organización. Corroborados con los estudios de Reyes Marroquín y Rosales Tejada (2007), sobre el desarrollo de Data Mart de información académica de estudiantes de la escuela de ciencias y sistemas de la Facultad de Ingeniería de la USAC, en la cual se concluyó en que la escuela de ciencias y sistemas se obtiene grandes beneficios al utilizarse el Data Mart académico, puesto que se podrá analizar el comportamiento de los estudiantes de la escuela y por ende se podrá tomar mejores decisiones en cuanto al uso de los recursos y el enfoque de la carrera de sistemas

En tal sentido con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora significativamente el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM

A nivel de análisis aplicado a la hipótesis específica “La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en el diseño de la toma de decisiones de la gestión académica”. En el cuadro Nº 56, gráfico Nº 03 y el gráfico Nº 04 se puede observar que el 66,7% de los tomadores de decisiones muestran un nivel muy bueno en cuanto al diseño de las decisiones de la gestión académica, por otro lado, el 33.3% de los tomadores de decisiones muestran un nivel de bueno. La prueba de hipótesis p=0.000 lo que indica que se aprueba la hipótesis especifica

Nº 01, por tal razón se afirma que, con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en el diseño de la toma de decisiones de la gestión académica. Según Vercellis (2009), Un Data Warehouse o Data Mart es un almacén de datos disponibles más importante para el desarrollo de arquitecturas de inteligencia de negocios y sistemas de soporte a la toma de decisiones. Simon (1960) menciona que el diseño implica identificar y explorar varias soluciones para el problema para poder tomar una decisión. Las mismas que fueron Corroborados con los estudios de Zambrano Alarcón (2011), sobre el Análisis, diseño e implementación de un Data Mart para el área de mantenimiento y logística de una empresa de transporte público de pasajeros, en cual concluyo en las necesidades de información del área de mantenimiento y logística de una empresa de transporte público fueron identificadas satisfactoriamente debido a que se consideró los posibles escenarios, actores y supuestos en toda empresa del mismo rubro. Esto contribuyó a identificar requerimientos claros y precisos que fueron documentados y utilizados para la construcción del modelo multidimensional.

En tal sentido La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en el diseño de la toma de decisiones de la gestión académica.

A nivel de análisis aplicado de la hipótesis específica Nº 02 “La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en la elección de decisiones de la gestión académica”. En el Cuadro Nº 57, gráfico Nº 05 y el gráfico Nº 06 se puede observar que el 33.3% de los tomadores de decisiones muestran un nivel muy bueno en cuanto a la elección de las decisiones de la gestión académica, por otro lado, el 66,7% de los tomadores de decisiones muestran un nivel de bueno. La prueba de hipótesis p=0.000 lo que indica que se aprueba la hipótesis especifica Nº02, por tal razón por la cual se afirma que, con La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye

significativamente en la elección de decisiones de la gestión académica. Guillén Rodríguez (2012) define la toma de decisiones como un proceso mediante el cual se realiza una elección entre las opciones o formas para resolver diferentes situaciones de la vida en diferentes contextos. Simon (1960) menciona que la elección consiste en elegir una de varias alternativas de solución para el problema para poder tomar una decisión. Corroborados con los estudios de Flores (2008) en su investigación titulada “Diseño, implantación y explotación de Data Mart”, los resultados del estudio mostraron que El usuario del sistema puede efectuar cualquier consulta de acuerdo a las variables y criterios de consulta del Data Warehouse así como hacer uso de múltiples variables que faciliten la proceso de la elección.

En tal sentido La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en la elección de decisiones de la gestión académica.

CONCLUSIONES

Luego del análisis de los resultados y discusión realizada sobre La implementación de un Data Mart, como parte de una solución de inteligencia de negocio para mejorar el proceso de Toma de Decisiones en La Gestión Académica de Facultad de Ciencias en la UNASAM, se llegó a las siguientes conclusiones:

1. La implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora significativamente el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM. Según los resultados el 66.7% de los tomadores de decisiones se encuentra en el nivel muy bueno y el 33.3% de los tomadores de decisiones en el nivel bueno. El p=0,000 siendo altamente significativo; por lo que queda demostrado que el Data Mart, como parte de una solución de inteligencia de negocio, se mejora significativamente el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM (ver Cuadro Nº 55, Gráfico Nº 01 y gráfico Nº 02).

2. La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en el diseño de la toma de decisiones de la gestión académica. Según los resultados el 66,7% de los tomadores de decisiones muestran un nivel muy bueno y el 33.3% en el nivel bueno. El p=0,000; por lo que queda demostrado que la influencia del Data Mart sobre diseño de las decisiones de la gestión académica (ver Cuadro Nº 56, Gráfico Nº 03 y gráfico Nº 04), razón por la cual se confirma la hipótesis específica Nº 01.

3. La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en la elección de decisiones de la gestión académica. Según los resultados el 33.3% de los tomadores de decisiones muestran un nivel muy bueno se encuentra en el nivel muy bueno y

el 66,7% de los tomadores de decisiones muestran un nivel de bueno. El p=0,000; por lo que queda demostrado la influencia del Data Mart sobre la elección de decisiones de la gestión académica (ver Cuadro Nº 57, Gráfico Nº 05 y Gráfico Nº 06), razón por la cual se confirma la hipótesis específica Nº 02.

RECOMENDACIONES

- A la Universidad Nacional Santiago Antúnez de Mayolo, impulsar el desarrollo de soluciones de inteligencia de Negocios, Data Marts u otros, para mejorar el seguimiento académico de los estudiantes u otros puntos relevantes que lo requieran.

- A la Escuela de Ingeniería de Sistemas e Informática, impulsar el desarrollo e implementación de Soluciones de Inteligencia de Negocio que contribuyan a la educación de los estudiantes de la escuela.

- Al director de Escuela de Ingeniería de Sistemas e Informática de la Universidad Nacional Santiago Antúnez de Mayolo, realizar convenios con universidades públicas y privadas e instituciones superiores públicas y privadas para que los estudiantes puedan realizar proyectos de innovación utilizando Inteligencia de Negocios que aporten conocimiento a la sociedad.

- A los estudiantes de la carrera, indagar y aplicar las soluciones de Inteligencia de Negocios como parte de su formación.

- A los futuros investigadores en campo de Inteligencia de Negocios, Se recomienda el empleo de esta metodología DWEP para el desarrollo de Data Marts o DataWarehouse, porque es una metodología robusta y muy útil que se caracteriza por ser incremental e iterativo.

REFERENCIAS BIBLIOGRÁFICAS

Ander Egg, Ezequiel. 1986. *Técnicas de investigación social.* Buenos Aires: Humanitas.

Arias, F. 2011. *El Proyecto de Investigación cientifica.* Sexta. Venezuela: Ediciones El Pasivo.

Don Hellriegel, y Jhon W. Slocum. 2004. *Comportamiento organizacional.* Mexico: Thomson.

Fischman, David. 2002. *El secreto de las siete semillas.* Lima: UPC.

Flores Chacón, Erick Giovanny, Esteban Julio Medina Rafaile, y Eddy Jesús Montañez Muñoz. 2008. *Diseño, Implementación y explotación de un Data Mart.* Tesis, Huaraz: Universidad Nacional Santiago Antúnez de Mayolo.

Gonzáles, A. M. 2002. *Aspectos éticos de la investigación.* Barcelona: Iberoamericana de Educación.

Guillén Rodríguez, Fiorelly Shirley. 2012. «Desarrollo de un Data Mart para mejorar la toma de decisiones en el área de tesorería de la Municipalidad provincial de Cajamarca.» Tesis, Universidad Privada del Norte, Cajamarca.

Hernández R., Fernández C. Baptista L. 2006. *Metodología de la Investigación. Tercera edición.* México: Mc Graw Hill Editores.

—. 2003. *Metodología de la Investigación. Tercera edición.* México: Mc Graw Hill Editores.

Hernández Sampieri, Roberto, Carlos Fernández Collado, y María del Pilar Baptista Lucio. 2010. *Metodologia de la Investigación.* Quinta. México: Mc Graw Hill.

Inmon, H. W. 2005. *Building the Data Warehouse.* Cuarta Edición. Canada: Wiley Publishing, Inc.

Juran, J. M. 1988. *Manual de control de calidad.* México: Ediciones Mc Graw Hill.

Kimball, Ralph, y Margy Ross. 2013. *The Data WareHouse Toolkit.* Third. Indianapolis, Indiana: Willey.

Landeau, R-. 2007. *Elaboración de trabajos de investigación.* Venezuela: Editorial Alfa.

Luján Mora, Sergio , y Juan Trujillo. 2006. «Applying the UML and the Unified Process to the design of Data Warehouses.» *Journal of Computer Information Systems* (Department of Software and Computing Systems University of Alicante) 30-58.

Mazón, José Norberto, Jesús Pardillo, y Juan Carlos Trujillo. 2011. *Diseño y Explotación de Almacenes de Datos.* Alicante: Editorial Club Universitario.

Pressman, Roger S. 2006. *Ingeniería del Software. Un enfoque práctico.* México: Mc Graw Hill.

Reyes Marroquín, Mario Roberto, y Pablo Augusto Rosales Tejada. 2007. *DESARROLLO DE UN DATA MART DE INFORMACIÓN ACADÉMICA DE ESTUDIANTES DE LA ESCUELA DE CIENCIAS Y SISTEMAS DE LA FACULTAD DE INGENIERÍA DE LA USAC.* Tesis, Universidad de San Carlos de Guatemala, Guatemala: Universidad de San Carlos de Guatemala, 115.

Río, Luis Méndez del. 2006. *Más allá del Business Intelligence: 16 experiencias de éxito.* Barcelona: Gestión 2000.

Rodríguez Moguel, Ernesto A. 2005. *Metodología de la Investigación.* México: Universidad Júarez Autónoma de Tabasco.

Sánchez, H. y Reyes, C. 1987. *Metodología y Diseños en la Investigación Científica.* Lima: S.C.

Sauter, Vicki Lynn. 2010. *Decision Support Systens for Busines Intelligence.* Segunda. Canada: WILEY.

Simon, H. A. 1960. *The New Science of Management Decision.* New York: Harper & Row.

Sommerville, Ian. 2005. *Ingeniería del Software.* Septima. Madrid: Pearson Addison Wesley.

Tansey, Stephen D. 2003. *Business, Information Technology and Society.* New York: Routledge.

Ullman, Jeffrey D., y Jennifer Widom. 1999. *Introducción a Los Sistemas de Bases de Datos.* Pearson Educación.

Valderrama, M. S. 2000. *Pasos para elaborar proyectos y tesis de investigación científica.* Lima: Ediciones San Marcos.

—. 2005. *Pasos para elaborar proyectos y tesis de investigación científica.* Lima: Editorial San Marcos.

Velasquez, F. A y Rey, C. N. 2007. *Metodología de la investigación científica.* Lima: Editorial San Marcos.

Vercellis, Carlo. 2009. *Business Intelligence Data Mining and Optimization for Decision Making.* Italy: Wiley.

Vivanco, Manuel. 2005. *Muestreo Estadístico Diseño y Aplicaciones.* Primera. Santiago de Chile: Universitaria.

Westerman, Paul. 2001. *Data Warehousing: Using the WaI-Mart Model.* San Diego: Academic Press.

Zambrano Alarcón, Jaime Alexander. 2011. *Análisis, diseño e implementación de un Data Mart para el área de Mantenimiento y Logística de una empresa de transporte público de pasajeros.* Tesis Pre-grado, Pontificia Universidad Católica del Perú, Lima: Tesis PUCP.

Anexos

Anexo Nº 01

Resolución de Aprobación de Proyecto

UNIVERSIDAD NACIONAL
"SANTIAGO ANTÚNEZ DE MAYOLO"
"Una nueva universidad para el desarrollo"

FACULTAD DE CIENCIAS

Resolución de Consejo de Facultad N° 109-2016-UNASAM-FC
Huaraz, 03 de Junio de 2016

Visto el acuerdo de sesión extraordinaria N° 08-2016 de Consejo de Facultad del día uno de Junio del año dos mil dieciséis, sobre Prorroga de Plan de Tesis;

CONSIDERANDO:

Que, con Resolución de Consejo de Facultad N° 083-2014-UNASAM-FC del 25 de Mayo de 2014 se APRUEBA, la Ejecución de los Proyectos de Tesis presentados por los alumnos de la Escuela Profesional de Ingeniería de Sistemas e Informática y DESIGNA los asesores respectivos de los proyectos de Tesis, según el Anexo N° 01 de dicha Resolución, como sigue:

TESISTA	TEMA	ASESOR
DURAN COLONIA Evelyn Anabel	Data Mart como parte de una solución de inteligencia de negocios, para el soporte de la toma de decisiones de la gestión académica de la Facultad de Ciencias de la UNASAM	ING. FLORES CHACÓN ERICK GIOVANNY
PURIFICACIÓN ASNATE Joseph Emerson	Sistema de información basado en realidad aumentada para la promoción gastronómica en el Callejón de Huaylas	ING. SILVA ZAPATA MIGUEL ANGEL
GAMARRA ORTIZ Joel Harley	Sistema experto para el diagnóstico y control de las plagas que afectan los cultivos de alcachofa y espárrago en Ancash	ING. SILVA ZAPATA MIGUEL ANGEL

Que, mediante Oficio N° 087-2016-UNASAM-FC-ISel/D, el Director de la Escuela Profesional de Ingeniería de Sistemas e Informática, remite adjunto la solicitud de los Bachilleres quienes están solicitando la prórroga de presentación de Tesis por motivos indicados en dicho documento, haciendo conocimiento que sí procede según el Capítulo IV, Artículo 29° del Reglamento de Grados y Títulos de la Escuela Profesional de Ingeniería de Sistemas e Informática, siendo los siguientes tesistas: DURAN COLONIA Evelyn Anabel, PURIFICACIÓN ASNATE Joseph Emerson y GAMARRA ORTIZ Joel Harley.

Que, el Consejo de Facultad en su sesión del visto, en concordancia con el Art. 157°, inc. 157.24 del Estatuto de la UNASAM, al Capítulo IV, Artículo 29° del Reglamento de Grados y Títulos de la Escuela Profesional de Ingeniería de Sistemas e Informática y en virtud de los considerandos que anteceden, acuerda por unanimidad, APROBAR la prórroga para la Presentación de Plan de Tesis de los bachilleres de la Escuela Profesional de Ingeniería de Sistemas e Informática solicitantes, por un (01) año a partir del acuerdo de Consejo de Facultad.

Que, en uso de las atribuciones conferidas y de conformidad con lo acordado en la sesión del visto del Consejo de Facultad;

SE RESUELVE:

Artículo 1°. APROBAR la **Prórroga** para la **Presentación de Plan de Tesis** de los bachilleres de la Escuela Profesional de Ingeniería de Sistemas e Informática, por un (01) año a partir del acuerdo de Consejo de Facultad, según lo siguiente:

TESISTA	TEMA	ASESOR
DURAN COLONIA Evelyn Anabel	Data Mart como parte de una solución de inteligencia de negocios, para el soporte de la toma de decisiones de la gestión académica de la Facultad de Ciencias de la UNASAM	ING. FLORES CHACÓN ERICK GIOVANNY
PURIFICACIÓN ASNATE Joseph Emerson	Sistema de información basado en realidad aumentada para la promoción gastronómica en el Callejón de Huaylas	ING. SILVA ZAPATA MIGUEL ANGEL
GAMARRA ORTIZ Joel Harley	Sistema experto para el diagnóstico y control de las plagas que afectan los cultivos de alcachofa y espárrago en Ancash	ING. SILVA ZAPATA MIGUEL ANGEL

Artículo 2°. DISPONER, a las dependencias que sean de su competencia el cumplimiento de la presente Resolución.

Regístrese, comuníquese y archívese.

UNASAM - SECRETARIA CONSEJO DE FACULTAD - FACULTAD DE CIENCIAS

M.Sc. Andrea Rosario PARI SOTO
Secretaria de Consejo de Facultad

UNIVERSIDAD NACIONAL SANTIAGO ANTUNEZ DE MAYOLO - Facultad de Ciencias - DECANATURA

M.Sc. Carlos Antonio REYES PAREJA
Decano FC-UNASAM

Dist.:DEP ISeI-Interesados-Arch.

UNIVERSIDAD NACIONAL
"SANTIAGO ANTUNEZ DE MAYOLO"
FACULTAD DE CIENCIAS

Resolución de Consejo de Facultad N° 083-2014-UNASAM-FC

Huaraz, 25 de Mayo de 2014

Visto el acuerdo de sesión extraordinaria N° 09-2014 de Consejo de Facultad del 22 de mayo de 2014, sobre aprobación de proyectos de Tesis y designación de asesores de los alumnos de la Escuela Académico Profesional de Ingeniería de Sistemas e Informática:

CONSIDERANDO:

Que, mediante Oficio Nº 049-2014-UNASAM-FC-DEAPISeI/D el Director de la Escuela Académico Profesional de Ingeniería de Sistemas e Informática, remite el Informe de la Comisión de Grados y Títulos de dicha Escuela, respecto a la Sustentación de los Proyectos de Tesis trece (13) alumnos egresados, quienes sustentaron su Proyecto de Tesis satisfactoriamente y amparado en el Reglamento de Grados y Títulos, Capítulo IV, Artículo 29º solicita emitir la resolución de Aprobación y Ejecución de los proyectos de Tesis de los alumnos en mención, así mismo la designación de los respectivos asesores de tesis.

Que, en consideración al antecedente descrito, el Consejo de Facultad acuerda por unanimidad Aprobar la Ejecución de los Proyectos de Tesis presentados por los alumnos solicitantes de la Escuela Académico Profesional de Ingeniería de Sistemas e Informática propuestos y Designar los respectivos asesores de tesis.

Estando a lo acordado por el Consejo de Facultad en su sesión del visto, en mérito al Estatuto de la UNASAM Artículo 174º, incisos (d) y (t), al Reglamento Interno de la Universidad Nacional "Santiago Antúnez de Mayolo" y en uso de las atribuciones conferidas

SE RESUELVE:

Artículo 1º. APROBAR la Ejecución de los Proyectos de Tesis presentados por los alumnos solicitantes de la Escuela Académico Profesional de Ingeniería de Sistemas e Informática, según el Anexo Nº 01.

Artículo 2º. DESIGNAR los Asesores para los proyectos de Tesis aprobados en el artículo precedente, según el Anexo Nº 01.

Artículo 3º. DISPONER, a las dependencias que sean de su competencia el cumplimiento de la presente Resolución.

Regístrese, comuníquese y archívese.

UNASAM Secretaría Consejo de Facultad FACULTAD DE CIENCIAS

MSc. Alberto M. MEDINA VILLACORTA
Secretario de Consejo de Facultad

UNIVERSIDAD NACIONAL SANTIAGO ANTUNEZ DE MAYOLO Facultad de Ciencias DECANATURA

MSc. Carlos Antonio REYES PAREJA
Decano FC-UNASAM.

Dist.: DEAPISI. Interesados. Arch.

UNIVERSIDAD NACIONAL
"SANTIAGO ANTUNEZ DE MAYOLO"
FACULTAD DE CIENCIAS

Resolución de Consejo de Facultad N° 083-2014-UNASAM-FC

Huaraz, 25 de Mayo de 2014

ANEXO Nª 01

ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

RELACION DE PROYECTOS DE TESIS Y ASESORES

TESISTA(S)	TEMA	ASESOR
Alvarado Lugo Robert Valverde Torres Jonathan	Aplicación Android como medio de prevención temprana frente al desbordamiento de cuencas y ríos en las zonas agrícolas de Ancash	Rolando Salazar Cáceres
Ciriaco Fernández Luis Geysen Garay Asencios Edward Ronald	Desarrollo de un sistema de información web de soporte para el seguimiento de egresados de la FC-UNASAM	Carlos Alva Jáuregui
Figueroa Villanueva Juan Carlos	Desarrollo de un sistema web para la gestión de información de visitantes en la red de museos y sitios arqueológicos del Ministerio de Cultura – Ancash	Rolando Salazar Cáceres
Gamarra Ortiz Joel Harley	Sistema experto para el diagnóstico y control de las plagas que afectan los cultivos de alcachofa y esparrago en Ancash	Miguel Silva Zapata
Purificación Asnate Joseph Emerson	Sistema de información basado en realidad aumentada para la promoción gastronómica en el callejón de Huaylas	Miguel Silva Zapata
Urbano García Wilder Dolores	Implementación de un sistema de información web, para optimizar el control de los procesos administrativos de la Empresa Ingeniería y Construcción Sagitario EIRL	Miguel Silva Zapata
Blas Cerda Elmer Cecilio	Implementación de un Data Warehouse para mejorar la eficiencia en el proceso de toma de decisiones académicas de la Asociación Educativa Luz y Ciencia	Miguel Silva Zapata
Guerrero Araujo James Jhojam Medina Lopez Walter Junior	Desarrollo e implementación de un sistema de soporte para optimizar los estudios de demanda social en la zona de influencia de la UNASAM	Eddy Montañez Muñoz
Mosquera Tahua Miguel	Desarrollo e implementación de un sistema de Gestión de indicadores de sostenibilidad para la evaluación del sistema agrícola de la comunidad de Huapra	Eddy Montañez Muñoz
Duran Colonia Evelyn Anabel	Data Mart como parte de una solución de inteligencia de negocios, para el soporte de la toma de desiciones de le gestión académica de la Facultad de Ciencias de la UNASAM	Erick Flores Chacón

Anexo Nº 02

Constancia de Aplicación de Proyecto

UNIVERSIDAD NACIONAL
SANTIAGO ANTÚNEZ DE MAYOLO

"Una Nueva Universidad para el Desarrollo"

ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS E INFORMATICA

"AÑO DEL BUEN SERVICIO AL CIUDADANO"

CONSTANCIA

EL QUE SUSCRIBE, DIRECTOR DE LA ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS E INFORMATICA DE LA FACULTAD DE CIENCIAS DE LA UNIVERSIDAD NACIONAL "SANTIAGO ANTUNEZ DE MAYOLO".

HACE CONSTAR:

Que, **DURAN COLONIA EVELYN ANABEL** con código N° 062.0103.121 ex alumna de la Escuela Profesional de Ingeniería de Sistemas e Informática de la Facultad de Ciencias, de la Universidad Nacional "Santiago Antúnez de Mayolo", realizó la aplicación de los Instrumentos de Medición del Proyecto de Investigación **"Data Mart como parte de una solución de de inteligencia de negocios, para el soporte de la Toma de Decisiones de la gestión Académica de la Facultad de Ciencias en la UNASAM".**

Se expide la presente a solicitud del interesado, para los fines que estime conveniente.

Huaraz, "Capital de la Amistad Internacional", **11 - Mayo - 2017.**

UNASAM DIRECTOR DE LA ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA FACULTAD DE CIENCIAS

Ing. Msc. Alberto M. MEDINA VILLACORTA
Director-Reg. C.I.P. N° 143211

AMMV/iire
C.c.:
- Arch.

Anexo N°03

Informe del Asesor

UNIVERSIDAD NACIONAL
SANTIAGO ANTÚNEZ DE MAYOLO
"Una Nueva Universidad para el Desarrollo"
Av. Centenario Nº 200 – Telefax (043) 421393 – Apartado 70
Huaraz – Ancash -Perú

INFORME Nº FC/UNASAM

SEÑOR DECANO DE LA FACULTAD DE CIENCIAS DE LA UNIVERSIDAD SANTIAGO ANTUNEZ DE MAYOLO HUARAZ.

ASUNTO : ASESORAMIENTO DE TESIS

INGENIERO : FLORES CHACÓN ERICK

DEJA CONSTANCIA QUE:

La Bach. DURAN COLONIA, Evelyn Anabel, de la carrera profesional Ingeniería de Sistemas e Informática, ha desarrollado el proyecto de investigación titulado:

DATA MART COMO PARTE DE UNA SOLUCION DE INTELIGENCIA DE NEGOCIOS, PARA EL SOPORTE DE LA TOMA DE DECISIONES DE LA GESTIÓN ACADÉMICA DE LA FACULTAD DE CIENCIAS EN LA UNASAM

Bajo mi asesoramiento, por tanto, garantizo sus fundamentos teóricos y lineamientos metodológicos, en función de la ingeniería de software.

Es todo por cuanto informo para su conocimiento y para los fines pertinentes.

Huaraz, 25 de mayo de 2017

Firma

Anexo Nº 04

Matriz de consistencia

MATRIZ DE CONSISTENCIA

TÍTULO:
"DATA MART COMO PARTE DE UNA SOLUCION DE INTELIGENCIA DE NEGOCIOS, PARA EL SOPORTE DE LA TOMA DE DECISIONES DE LA GESTIÓN ACADÉMICA DE LA FACULTAD DE CIENCIAS EN LA UNASAM"

PROBLEMA	OBJETIVOS	HIPÓTESIS	VARIABLES E INDICADORES	METODOLOGÍA
¿En qué medida la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, mejora el proceso de toma de decisiones de la gestión académica de la Facultad de Ciencias de la UNASAM?	Implementar un Data Mart, como parte de una solución de inteligencia de negocio, para mejorar el proceso de Toma de Decisiones en La Gestión Académica de Facultad de Ciencias en la UNASAM.	Con la implementación de un Data Mart, como parte de una solución de inteligencia de negocio, se mejora el proceso de Toma de Decisiones en la Gestión Académica de Facultad la Ciencias de la UNASAM.	***Variable independiente*** Data Mart ***Indicadores:*** - Información integrada - Estructura consistente - Niveles de detalle - Necesidades del Usuario - Datos organizados - Acceso fácil - Entendimiento del usuario final - Permite el análisis - Toma de decisiones - Base de datos estable - Predice tendencias - Se actualiza en un periodo de tiempo ***Variable dependiente*** Toma de Decisiones de la gestión Académica ***Indicadores:*** - Identifica los planes de estudio. - Identifica a los docentes. - Explora a los estudiantes. - Selecciona los datos. - Analiza los datos. - Elige una alternativa.	***Nivel y tipo*** Tipo de Investigación Explicativa, según la profundidad. De acuerdo a la finalidad es Aplicada ***Métodos*** Método Experimental Métodos específicos aplicados: método inductivo, el método deductivo, método sintético, método analítico. ***Diseño*** El diseño es Cuasi Experimental. G.E. ---------- x ---------- O1 G.C.1. ---------- O2 ***Técnicas e instrumentos de recolección de datos*** Técnica: Observación. Instrumento de recolección de datos: Escala de Estimación.
Problemas Específicos - ¿De qué manera la implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye en el diseño de las decisiones de la gestión académica? - ¿De qué manera la implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye en la elección de las decisiones de la gestión académica?	***Objetivos Específicos*** - Identificar y describir los requerimientos de análisis de información para el comportamiento académico de la Facultad de Ciencias de la UNASAM. - Diseñar el modelo dimensional que permita el análisis y explotación de la información identificada. - Desarrollar la estructura y funcionalidad del Data Mart, como parte de una solución de Inteligencia de Negocio. - Implementar el Data Mart, como parte de una solución de Inteligencia de Negocio. - Facilitar la toma de decisiones, en base a la Información presentada por el Data Mart.	***Hipótesis Específicos*** - La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en el diseño de las decisiones de la gestión académica. - La implementación de un Data Mart, como parte de una solución de inteligencia de negocio influye significativamente en la elección de decisiones de la gestión académica.		

Anexo Nº 05

Matriz de Operacionalización de Variables

Matriz de Operacionalización de Variables

Variables	Definición Operacional	Dimensiones	Indicadores	Ítems	Instrumentos / Niveles
				Variable Independiente	***Instrumento:*** *Escala de Estimación.* ***Niveles:*** *Siempre, Casi Siempre, A veces, Nunca*
Data **Mart**	La herramienta de Soporte de decisiones que ayudar a medir los indicadores de la toma de decisiones de la gestión académica.	Integrado	Información integrada		
			Estructura consistente		
			Niveles de detalle		
			Necesidades del Usuario		
		Temático	Datos organizados		
			Acceso fácil		
			Entendimiento del usuario final		
		No volátil	Permite el análisis		
			Toma de decisiones		
			Base de datos estable		
		Variante en el tiempo	Predice tendencias		
			Se actualiza en un periodo de tiempo		
				Variable Dependiente	
Toma de Decisiones de la gestión Académica	Determina las decisiones de la gestión académica dentro de la Facultad de Ciencias.	Diseño	Identifica los planes de estudio	Identifica la estructura del plan de estudios de una determinada carrera rápidamente	
				Explora los cursos que pertenecen a una determinada carrera eficazmente	
				Esboza las actualizaciones de los planes de estudio por carrera eficientemente	
				Examina en el tiempo la cantidad de horas	
				Localiza adecuadamente y rápidamente la cantidad de cursos electivos por carrera en una determinada Facultad	
				Reconoce el creditaje por curso	
				Identifica la cantidad de cursos por ciclo académico de una determinada carrera	
				Identifica la cantidad de horas teóricas versus las horas prácticas de un plan de estudios	
				Localiza fácilmente los planes de estudio de una determinada carrera durante un tiempo determinado.	
			Identifica a los docentes	Examina la cantidad de docentes en cada carrera rápidamente	
				Reconoce a los docentes según sus grados académicos	
				Identifica a los 10 mejores docentes según la encuesta docente durante los últimos semestres	
				Identifica los cursos dictados por un docente durante los últimos semestres	
				Identifica rápidamente el promedio de cada docente según la encuesta de los estudiantes durante los últimos semestres	
				Relaciona eficientemente la calificación obtenida con el grado académico del docente durante los últimos periodos académicos	
				Identifica la cantidad de docentes que pertenecen a un departamento académico	
				Esboza el promedio de los calificativos de los docentes durante los periodos académicos por departamento académico	

Matriz de Operacionalización de Variables

Variables	Definición Operacional	Dimensiones	Indicadores	Ítems	Instrumentos / Niveles
Toma de Decisiones de la gestión Académica	Determina las decisiones de la gestión académica dentro de la Facultad de Ciencias.	Diseño	Explora a los estudiantes	Identifica la cantidad de estudiantes en cada carrera rápidamente	***Instrumento:*** *Escala de Estimación.* ***Niveles:*** *Siempre, Casi Siempre, A veces, Nunca*
				Inspecciona el porcentaje de estudiantes aprobados en cada carrera eficazmente	
				Localiza el número de estudiantes aprobados en un determinado curso.	
				Recopila información del porcentaje aprobados de un determinado curso respecto a anteriores periodos académicos.	
				Inspecciona la cantidad de estudiantes matriculados de una determinada carrera con respecto a otras carreras	
				Identifica la fuente de información de los datos de los estudiantes rápidamente	
		Elección	Selecciona los datos	Identifica los datos de los planes de estudio según orden establecido en un tiempo mínimo.	
				Examina los datos organizados de los estudiantes rápidamente	
				Compara los datos los docentes eficazmente	
				Selecciona correctamente y eficientemente a los estudiantes del tercio superior	
				Selecciona rápidamente al estudiante con mayor promedio ponderado	
				Elige adecuadamente la apertura de cursos a programarse al inicio del semestre	
			Analiza los datos	Compara los datos de los alumnos matriculados con respecto a otros semestres en una determinada escuela rápidamente.	
				Identifica los porcentajes de los alumnos desaprobados en un determinado semestre.	
				Examina el número de estudiantes matriculados en los 10 últimos semestres rápidamente.	
				Predice tendencias de los estudiantes utilizando los datos del rendimiento académico rápidamente	
				Realiza tablas dinámicas para el análisis del rendimiento de estudiantes, docentes rápidamente	
				Realiza operaciones matemáticas para el cálculo de indicadores que las consulta requieran rápidamente.	
			Elige una alternativa	Elabora reportes de los estudiantes matriculados con los datos que posee rápidamente.	
				Identifica datos relevantes del rendimiento docente rápidamente.	
				Identifica datos relevantes del rendimiento académico rápidamente.	
				Selecciona los datos de los docentes utilizando criterios de selección rápidamente.	
				Predice la tendencia de los docentes en el tiempo utilizando los datos del rendimiento de los docentes eficazmente	
				Elabora reportes del rendimiento de los docentes utilizando con los datos que posee rápidamente.	

Anexo Nº06

Instrumento

Escala de Estimación de la Toma de Decisiones de la gestión Académica de la Facultad de Ciencias
Universidad Nacional Santiago Antúnez de Mayolo

ESCUELA		
APELLIDOS Y NOMBRES	**CARGO**	**FECHA**

Durante la aplicación de las consultas, se registró lo siguiente:

N.º	Ítems	Valoración				Tiempo
		Siempre	Casi Siempre	A veces	Nunca	
1	Identifica la estructura del plan de estudios de una determinada carrera rápidamente					
2	Explora los cursos que pertenecen a una determinada carrera eficazmente					
3	Esboza las actualizaciones de los planes de estudio por carrera eficientemente					
4	Examina en el tiempo la cantidad de horas					
5	Localiza adecuadamente y rápidamente la cantidad de cursos electivos por carrera en una determinada Facultad					
6	Reconoce el creditaje por curso.					
7	Identifica la cantidad de cursos por ciclo académico de una determinada carrera					
8	Identifica la cantidad de horas teóricas versus las horas prácticas de un plan de estudios					
9	Localiza fácilmente los planes de estudio de una determinada carrera durante un tiempo determinado.					
10	Examina la cantidad de docentes en cada carrera rápidamente					
11	Reconoce a los docentes según sus grados académicos					
12	Identifica a los 10 mejores docentes según la encuesta docente durante los últimos semestres					
13	Identifica los cursos dictados por un docente durante los últimos semestres					
14	Identifica rápidamente el promedio de cada docente según la encuesta de los estudiantes durante los últimos semestres					
15	Relaciona eficientemente la calificación obtenida con el grado académico del docente durante los últimos periodos académicos					
16	Identifica la cantidad de docentes que pertenecen a un departamento académico					
17	Esboza el promedio de los calificativos de los docentes durante los periodos académicos por departamento académico					
18	Identifica la cantidad de estudiantes en cada carrera rápidamente					
19	Inspecciona el porcentaje de estudiantes aprobados en cada carrera eficazmente					
20	Localiza el número de estudiantes aprobados en un determinado curso.					
21	Recopila información del porcentaje aprobados de un determinado curso respecto a anteriores periodos académicos.					
22	Inspecciona la cantidad de estudiantes matriculados de una determinada carrera con respecto a otras carreras					
23	Identifica la fuente de información de los datos de los estudiantes rápidamente					
24	Identifica los datos de los planes de estudio según orden establecido en un tiempo mínimo.					
25	Examina los datos organizados de los estudiantes rápidamente					
26	Compara los datos los docentes eficazmente					
27	Selecciona correctamente y eficientemente a los estudiantes del tercio superior					
28	Selecciona rápidamente al estudiante con mayor promedio ponderado					
29	Elige adecuadamente la apertura de cursos a programarse al inicio del semestre					
30	Compara los datos de los alumnos matriculados con respecto a otros semestres en una determinada escuela rápidamente.					
31	Identifica los porcentajes de los alumnos desaprobados en un determinado semestre.					
32	Examina el número de estudiantes matriculados en los 10 últimos semestres rápidamente.					
33	Predice tendencias de los estudiantes utilizando los datos del rendimiento académico rápidamente					
34	Realiza tablas dinámicas para el análisis del rendimiento de estudiantes, docentes rápidamente					
35	Realiza operaciones matemáticas para el cálculo de indicadores que las consulta requieran rápidamente.					
36	Elabora reportes de los estudiantes matriculados con los datos que posee rápidamente.					
37	Identifica datos relevantes del rendimiento docente rápidamente.					
38	Identifica datos relevantes del rendimiento académico rápidamente.					
39	Selecciona los datos de los docentes utilizando criterios de selección rápidamente.					
40	Predice la tendencia de los docentes en el tiempo utilizando los datos del rendimiento de los docentes eficazmente					
41	Elabora reportes del rendimiento de los docentes utilizando con los datos que posee rápidamente.					

Anexo Nº07

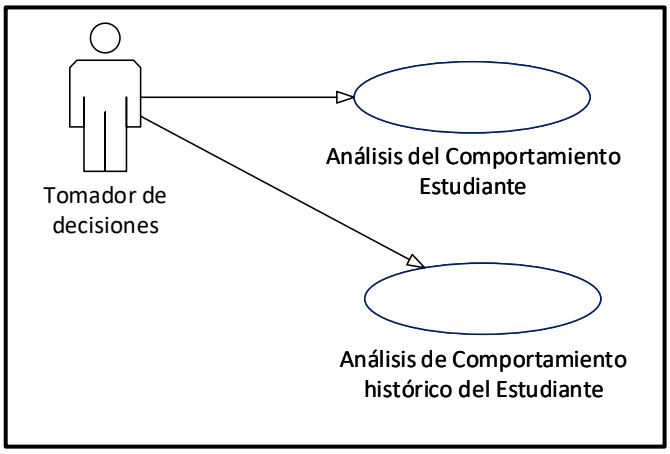
Diagrama de Casos de Uso

Diagrama de Casos de Uso

1. Paquete de Análisis del comportamiento Académico estudiante

Este paquete cubre las necesidades de información relacionadas a los mantenimientos realizados en los autobuses. En la figura Nº 01 se muestra los casos de uso que agrupan las necesidades de información relacionado a los mantenimientos.

Figura Nº 01. Paquete de Análisis del comportamiento.



Fuente: Elaboración Propia

A continuación, la especificación de cada caso de uso del paquete de análisis de comportamiento académico estudiante:

Cuadro Nº 01. Caso de Uso - Análisis del comportamiento Académico estudiante.

Caso de Uso	Análisis del comportamiento Académico estudiante.
Requerimiento que implementa	El sistema debe responder a las cuestiones del comportamiento académico de los estudiantes
Frecuencia de Actualización	Semestral.
Momento de Actualización	Después de finalizado el semestre Académico.
Dimensiones de análisis de información	• Estudiante. • Docente. • Modalidad de Aprobación. • Semestre. • Currícula.
Fórmulas de Cálculo	El actor puede seleccionar otra dimensión en cualquier momento.
Usuarios permitidos	Decano, jefes de departamento, Directores de Escuela, Secretaria de Escuela.
Cantidad / Historia de datos	Toda la historia.
Observaciones	Incluir subtotales agrupados por Semestre u otra dimensión

Fuente: Elaboración Propia

Cuadro Nº 02. Caso de Uso - Análisis del comportamiento Académico histórico del estudiante.

Caso de Uso	Análisis del comportamiento Académico histórico del estudiante.
Requerimiento que implementa	El sistema debe responder a las cuestiones del comportamiento académico de los estudiantes
Frecuencia de Actualización	Semestral.
Momento de Actualización	Después de finalizado el semestre Académico.
Dimensiones de análisis de información	• Estudiante. • Docente. • Modalidad de Aprobación. • Semestre. • Currícula.
Fórmulas de Cálculo	El actor puede seleccionar otra dimensión en cualquier momento.
Usuarios permitidos	Decano, jefes de departamento y Directores de Escuela.
Cantidad / Historia de datos	Toda la historia.
Observaciones	Incluir subtotales agrupados por Semestre u otra dimensión

Fuente: Elaboración Propia

Se presentan las especificaciones de los casos de uso según la metodología DWEP.

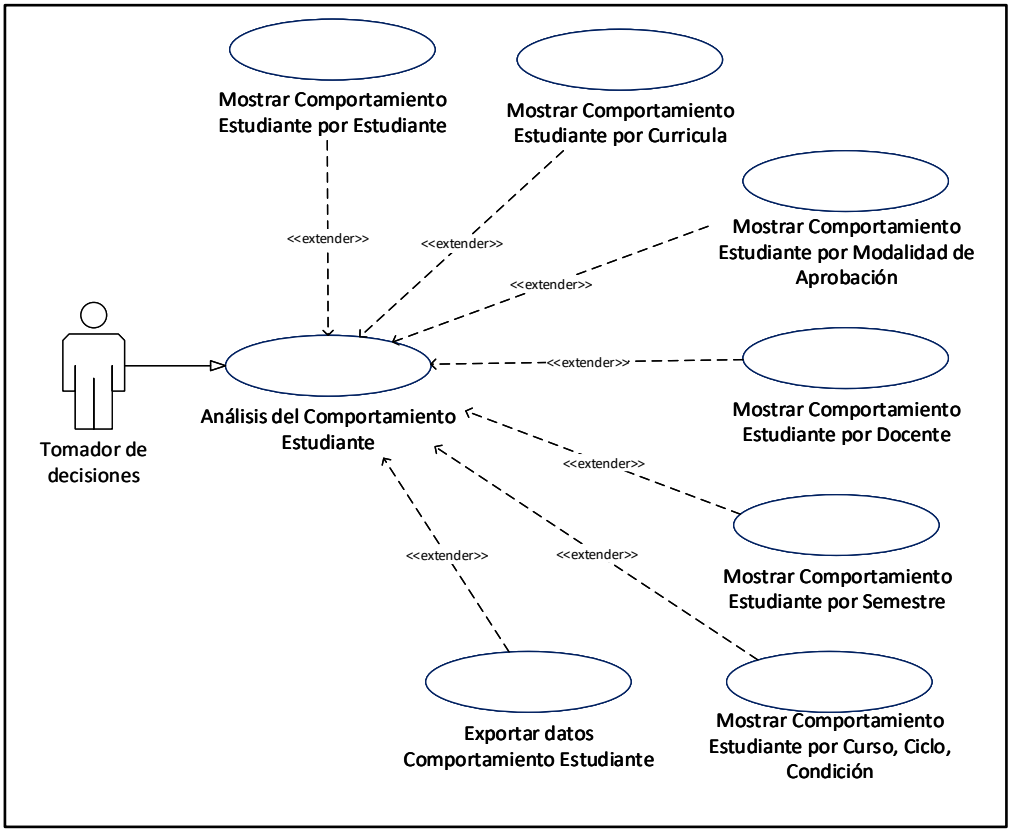
Caso de Uso Análisis del comportamiento Académico estudiante.

Cuadro Nº 03. Caso de Uso - Análisis del comportamiento Académico estudiante.

Caso de Uso	Análisis del comportamiento Académico estudiante.
ID	CU01
Actores	Decano, jefes de departamento, Directores de Escuela, Secretaria de Escuela.
Precondiciones	Haber finalizado un semestre.
Flujo de eventos	1. El actor selecciona el comportamiento académico de los estudiantes. 2. El sistema muestra las dimensiones posibles a analizar. 3. El actor selecciona alguna dimensión. 4. El sistema muestra comportamiento académico de los estudiantes agrupados por la dimensión seleccionada.
Postcondiciones	El actor puede seleccionar otra dimensión en cualquier momento.

Fuente: Elaboración Propia

Figura Nº 02. Análisis del comportamiento Estudiante.



Fuente: Elaboración Propia

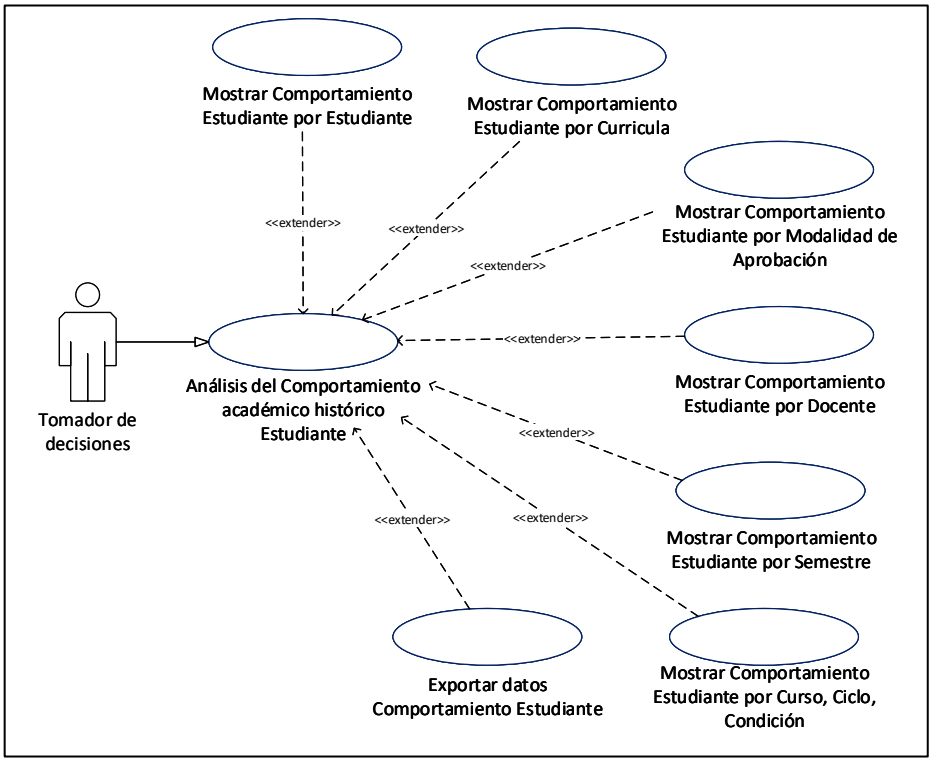
Caso de Uso Análisis del comportamiento Académico histórico del estudiante.

Cuadro Nº 04. Caso de Uso - Análisis del comportamiento Académico estudiante.

Caso de Uso	Análisis del comportamiento Académico estudiante.
ID	CU02
Actores	Decano, jefes de departamento, Directores de Escuela
Precondiciones	Haber finalizado un semestre.
Flujo de eventos	1. El actor selecciona el comportamiento académico histórico de los estudiantes. 2. El sistema muestra las dimensiones posibles a analizar. 3. El actor selecciona alguna dimensión. 4. El sistema muestra comportamiento académico histórico de los estudiantes agrupados por la dimensión seleccionada.
Postcondiciones	El actor puede seleccionar otra dimensión en cualquier momento.

Fuente: Elaboración Propia

Figura Nº 03. Análisis del comportamiento Estudiante.

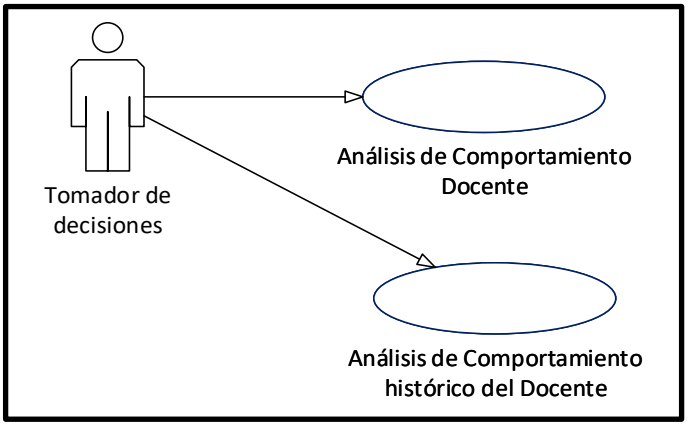


Fuente: Elaboración Propia

2. Paquete de Análisis del comportamiento Académico Docente

Este paquete cubre las necesidades de información relacionadas al análisis del comportamiento académico docente.

Figura Nº 04. Análisis del comportamiento Académico Docente.



Fuente: Elaboración Propia

A continuación, la especificación de cada caso de uso del paquete de análisis de comportamiento académico Docente:

Cuadro Nº 05. Caso de Uso - Análisis del comportamiento Académico Docente.

Caso de Uso	Análisis del comportamiento Académico Docente.
Requerimiento que implementa	El sistema debe responder a las cuestiones de los tomadores de decisiones acerca del comportamiento académico de los docentes.
Frecuencia de Actualización	Semestral.
Momento de Actualización	Después de finalizado el semestre Académico.
Dimensiones de análisis de información	• Docente. • Semestre. • Currícula.
Fórmulas de Cálculo	El actor puede seleccionar otra dimensión en cualquier momento.
Usuarios permitidos	Decano, jefes de departamento, Directores de Escuela y Secretaria de Escuela.
Cantidad / Historia de datos	Toda la historia.
Observaciones	Incluir subtotales agrupados por Semestre u otra dimensión

Fuente: Elaboración Propia

Cuadro Nº 06. Caso de Uso - Análisis del comportamiento Académico histórico del Docente.

Caso de Uso	Análisis del comportamiento Académico histórico del Docente.
Requerimiento que implementa	El sistema debe responder a las cuestiones de los tomadores de decisiones acerca del comportamiento académico de los docentes.
Frecuencia de Actualización	Semestral.
Momento de Actualización	Después de finalizado el semestre Académico.
Dimensiones de análisis de información	• Docente. • Semestre. • Currícula.
Fórmulas de Cálculo	El actor puede seleccionar otra dimensión en cualquier momento.
Usuarios permitidos	Decano, jefes de departamento y Directores de Escuela.
Cantidad / Historia de datos	Toda la historia.
Observaciones	Incluir subtotales agrupados por Semestre u otra dimensión

Fuente: Elaboración Propia

Se presentan las especificaciones de los casos de uso según la metodología DWEP.

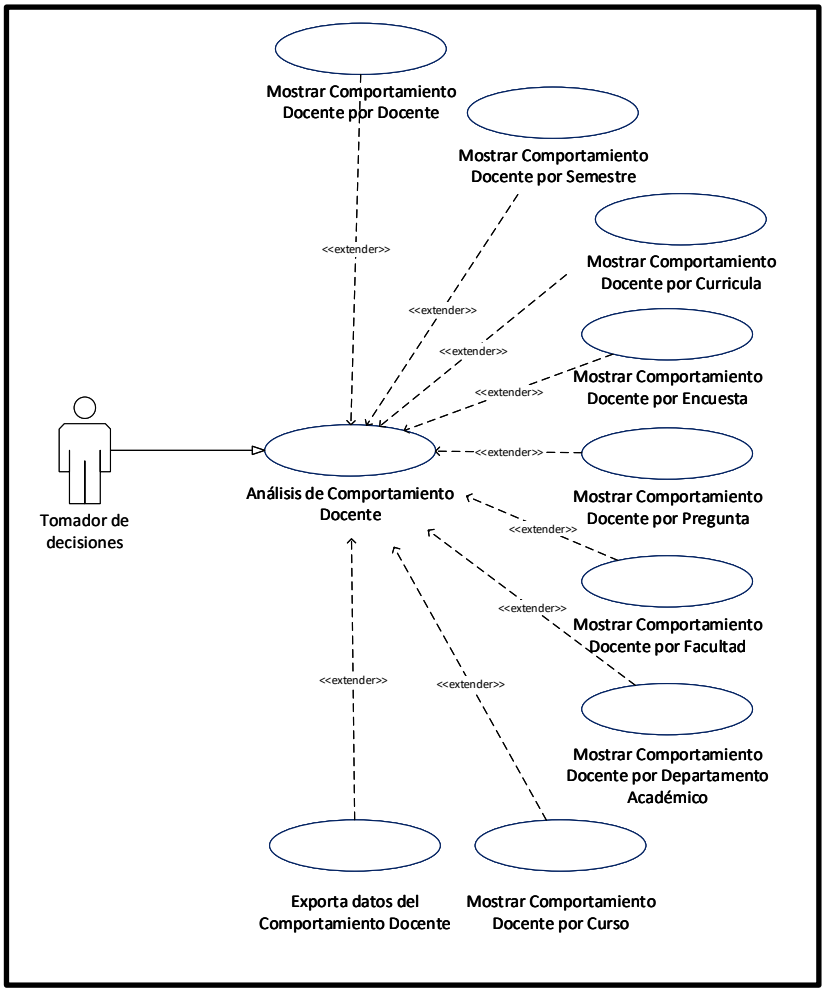
Caso de Uso Análisis del comportamiento Académico Docente.

Cuadro Nº 07. Caso de Uso - Análisis del comportamiento Académico Docente.

Caso de Uso	Análisis del comportamiento Académico Docente.
ID	CU03
Actores	Decano, jefes de departamento, Directores de Escuela, Secretaria de Escuela.
Precondiciones	Haber finalizado un semestre.
Flujo de eventos	1. El actor selecciona el comportamiento Docente. 2. El sistema muestra las dimensiones posibles a analizar. 3. El actor selecciona alguna dimensión. 4. El sistema muestra el comportamiento docente agrupados por la dimensión seleccionada.
Postcondiciones	El actor puede seleccionar otra dimensión en cualquier momento.

Fuente: Elaboración Propia

Figura Nº 05. Análisis del comportamiento Académico Docente.



Fuente: Elaboración Propia

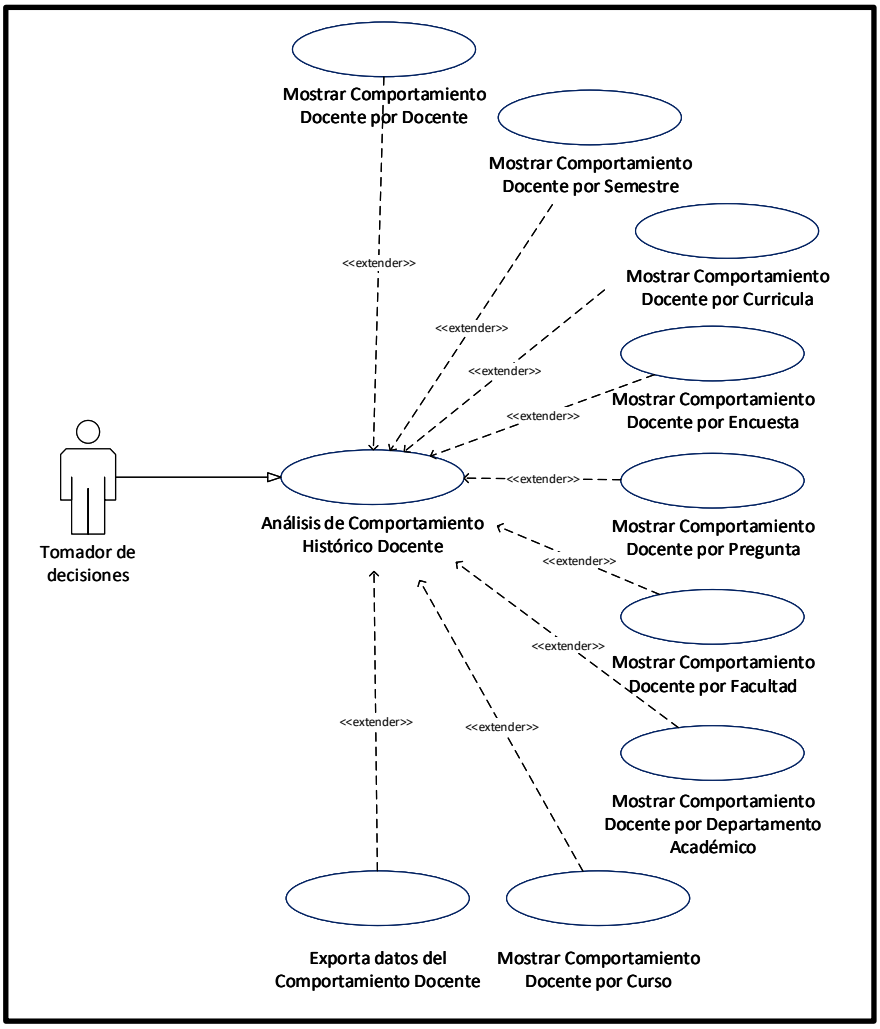
Caso de Uso Análisis del comportamiento Académico Docente.

Cuadro Nº 07. Caso de Uso - Análisis del comportamiento Académico Histórico Docente.

Caso de Uso	Análisis del comportamiento Académico Histórico Docente.
ID	CU04
Actores	Decano, jefes de departamento, Directores de Escuela.
Precondiciones	Haber finalizado un semestre.
Flujo de eventos	1. El actor selecciona el comportamiento Docente. 2. El sistema muestra las dimensiones posibles a analizar. 3. El actor selecciona alguna dimensión. 4. El sistema muestra el comportamiento docente histórico agrupados por la dimensión seleccionada.
Postcondiciones	El actor puede seleccionar otra dimensión en cualquier momento.

Fuente: Elaboración Propia

Figura Nº 06. Análisis del comportamiento Académico Docente.



Fuente: Elaboración Propia

Anexo N°8

Esquema Conceptual y Lógico de la Fuente de Datos

ESQUEMA CONCEPTUAL Y LÓGICO DE LA FUENTE DE DATOS

Según la metodología DWEP para especificar la información fuente del Data Mart es necesario realizar el Esquema Conceptual Datos Fuente (Source Conceptual Schema), el Esquema Lógico de Datos Fuente (Source Logical Schema), y el esquema físico de la fuente de Datos para lograr detallar la estructura de la información que sirvió como fuente al modelo dimensional del Data Mart de la Facultad de Ciencias.

1. ESQUEMA CONCEPTUAL DE LOS DATOS FUENTE

El esquema conceptual de los datos Fuente es mostrado en la Ilustración Nº 01, la cual está acorde a la metodología DWEP.

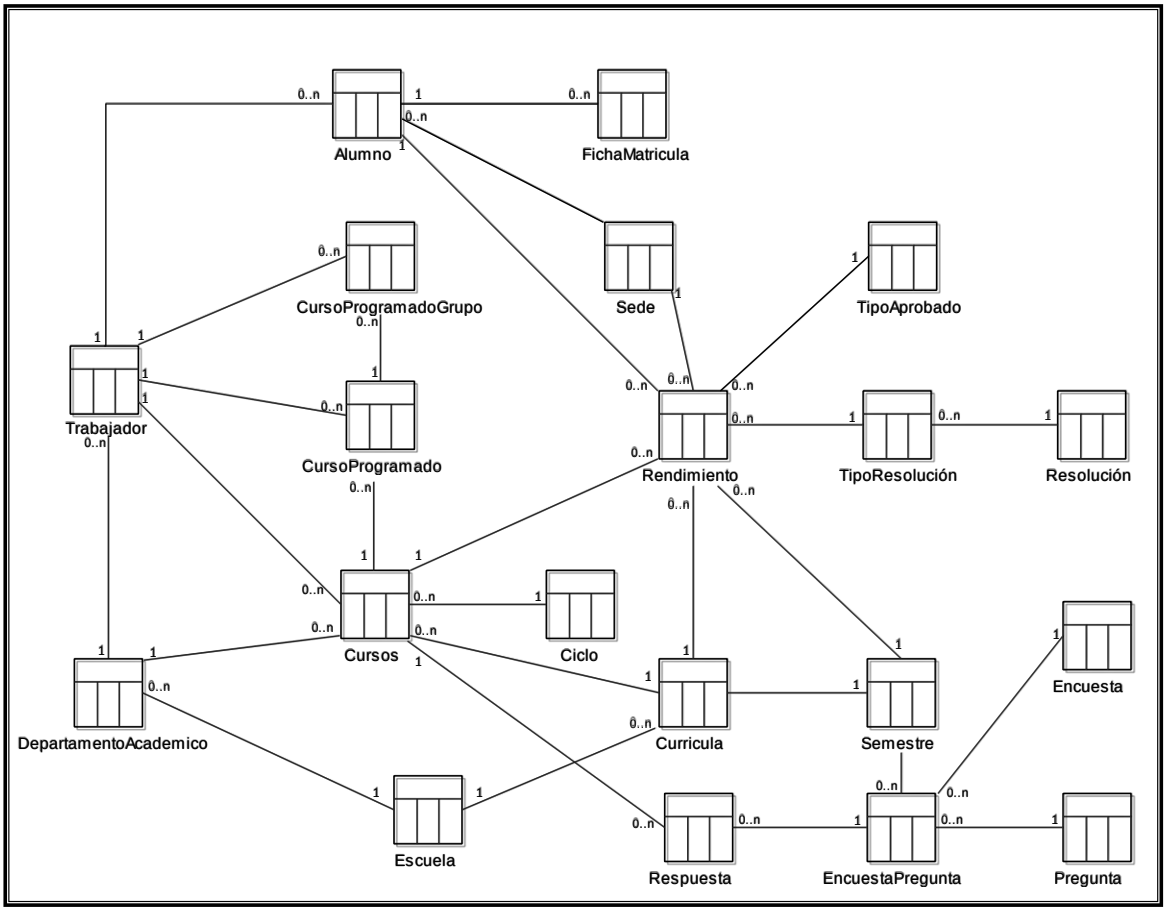
Ilustración Nº 01. Esquema Conceptual de los datos Fuente.

Fuente: Base de datos ITUNASAM

2. ESQUEMA LÓGICO DE LOS DATOS FUENTE

El esquema lógico de los datos Fuente es mostrado en la Ilustración Nº 02, la cual está acorde a la metodología DWEP.

Ilustración Nº 02. Esquema Conceptual de los datos Fuente.

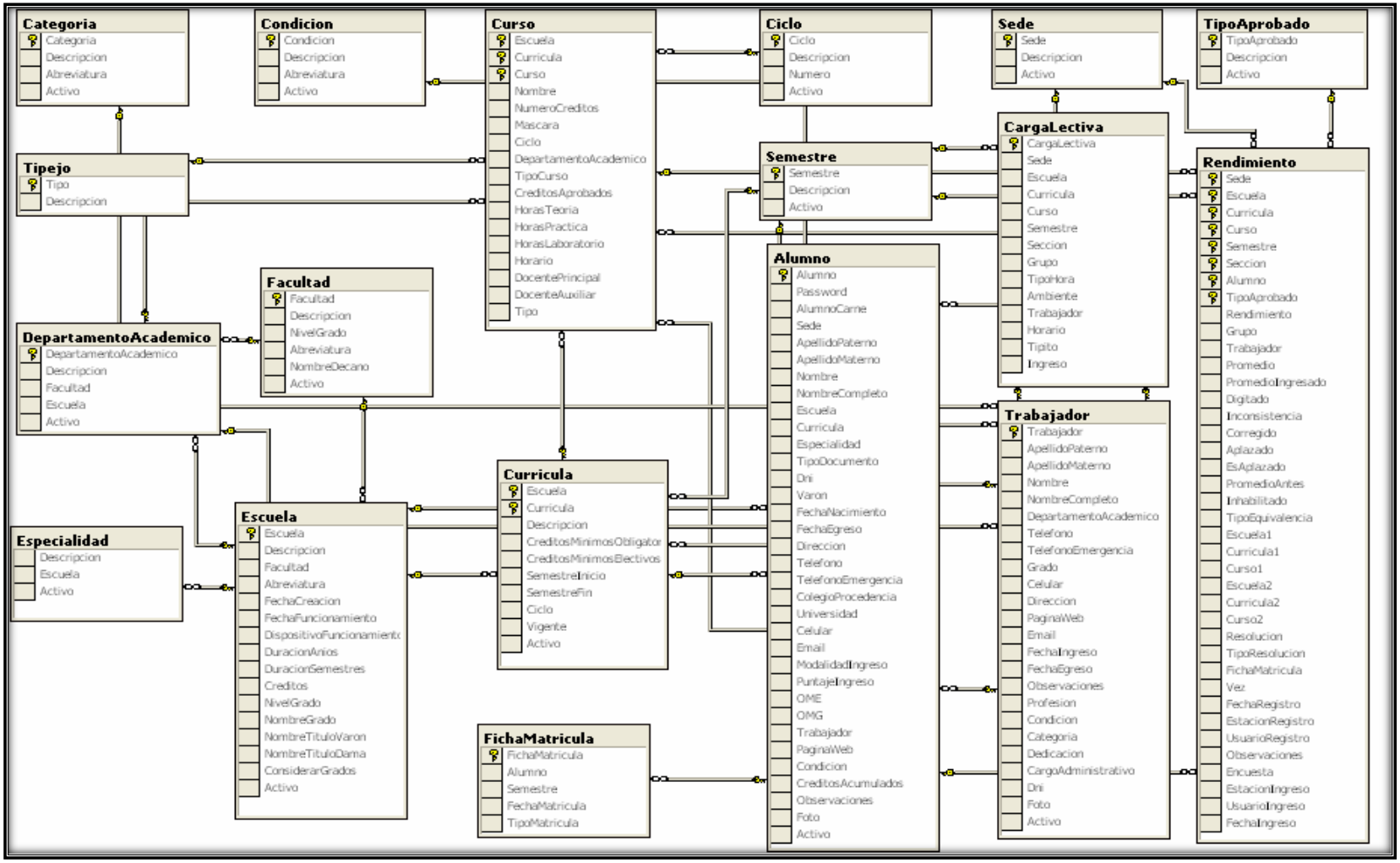


Fuente: Base de datos ITUNASAM

3. ESQUEMA FÍSICO DE LA FUENTE DE DATOS

El esquema lógico de los datos Fuente es mostrado en la Ilustración Nº 3, la cual está acorde a la metodología DWEP.

Ilustración Nº 03. Esquema Conceptual de los datos Fuente.



Fuente: Base de datos ITUNASAM

4. DICCIONARIO DE ESQUEMA LÓGICO DE DATOS FUENTE

A continuación, se especificará cada clase presentada en el esquema lógico de datos fuente.

CLASE	Alumno
DESCRIPCIÓN	Clase que representa los alumnos
	Atributos
Alumno	Código de alumno
Password	Contraseña de alumno
AlumnoCarne	Numero de carnet de alumno
Sede	Código de la sede
ApellidoPaterno	Apellido paterno del alumno
ApellidoMaterno	Apellido materno del alumno
Nombre	Nombre del alumno
Especialidad	Código de la especialidad
NombreCompleto	Nombre completo del alumno
Escuela	Código de la escuela
Curricula	Código de la currícula
TipoDocumento	Código de documento
Dni	Número de DNI
Varon	Sexo del alumno
FechaNacimiento	Fecha de nacimiento
FechaEgreso	Fecha de egreso
Direccion	Dirección del alumno
Telefono	Número de teléfono
TelefonoEmergencia	Teléfono de Emergencia
ColegioProcedencia	Código del Colegio de Procedencia
Universidad	Código de la Universidad
Celular	Numero de Celular
Email	Dirección de correo electrónico
ModalidadIngreso	Código de modalidad de Ingreso
PuntajeIngreso	Puntaje de ingreso
OME	Numero de orden de méritos en la Escuela
OMG	Numero de orden de méritos general
Trabajador	Código del trabajador
PaginaWeb	Página web del alumno
Condicion	Condición del alumno
CreditosAcumulados	Número de créditos acumulados
Observaciones	Observaciones
Foto	Foto
Activo	Activo
pago	Pago

CLASE	Ficha_Matricula
DESCRIPCIÓN	Clase que representa la ficha de Matricula
	Atributos
FichaMatricula	Código de ficha de matricula
Alumno	Código de alumno
Semestre	Código de Semestre
FechaMatricula	Fecha de matrícula
TipoMatricula	Tipo de matricula

CLASE	Trabajador
DESCRIPCIÓN	Clase que representa al Trabajador
	Atributos
Trabajador	Código de trabajador.
ApellidoPaterno	Apellido paterno del trabajador.
ApellidoMaterno	Apellido materno del trabajador.
Nombre	Nombre del trabajador.
NombreCompleto	Nombre completo del trabajador.
DepartamentoAcademico	Código de departamento académico.
Telefono	Número de teléfono.
TelefonoEmergencia	Número de teléfono de emergencia.
Grado	Grado académico del trabajador.
Celular	Numero de celular.
Direccion	Dirección.
PaginaWeb	Dirección de página web del trabajador.
Email	Dirección del correo electrónico.
FechaIngreso	Fecha de Ingreso.
FechaEgreso	Fecha de Egreso.
FechaNacimiento	Fecha de Nacimiento.
Observaciones	Observaciones.
Profesion	Profesión del trabajador.
Condicion	Código de Condición.
Categoria	Código de Categoría.
Dedicacion	Código de Dedicación.
CargoAdministrativo	Cargo Administrativo
Dni	Número de DNI.
Foto	Foto
Activo	Activo
password	Contraseña del trabajador.
PASSMD5	Contraseña encriptado

CLASE	Sede
DESCRIPCIÓN	Clase que representa la ficha de Matricula
	Atributos
Sede	Codigo de sede.
Descripcion	Nombre de la sede.
Activo	Estado de la sede.

CLASE	Curso
DESCRIPCIÓN	Clase que representa el curso.
	Atributos
Escuela	Código de Escuela.
Curricula	Código de curricula.
Curso	Código de curso
Nombre	Nombre de curso.
NumeroCreditos	Número de créditos del curso
Mascara	Mascara del curso.
Ciclo	Código del ciclo.
DepartamentoAcademico	Código del departamento académico.
TipoCurso	Código de tipo de Curso.
CreditosAprobados	Número de créditos aprobados.
HorasTeoria	Número de Horas Teóricas.
HorasPractica	Número de Horas prácticas.
HorasLaboratorio	Número de Horas de laboratorio.
Horario	Código de horario.
DocentePrincipal	Código de Docente principal.
DocenteAuxiliar	Código de Docente auxiliar.
Tipo	Código de Tipo de curso.
Sumilla	Sumilla de curso.

CLASE	CursoProgramado
DESCRIPCIÓN	Clase que representa los cursos programados en un semestre.
	Atributos
Sede	Código de Sede.
Escuela	Código de Escuela.
Curricula	Código de curricula.
Curso	Código de curso
Semestre	Código de Semestre.
Seccion	Sección
CuentaCarga	Cuenta Carga
Capacidad	Capacidad
Trabajador	Código de trabajador
TrabajadorTeoria	Código de trabajador de encargado de la teoría.
CursoProgramado	Código del Curso Programado
Silabo	Silabo
Matriculado	Número de Matriculados
Nombredocente	Nombre docente
tipo	tipo

CLASE	CursoProgramadoGrupo
DESCRIPCIÓN	Clase que representa los cursos programados con sus respectivos grupos en un semestre.
	Atributos
Sede	Código de Sede.
Escuela	Código de Escuela.
Curricula	Código de curricula.
Curso	Código de curso
Semestre	Código de Semestre.
Seccion	Sección
Grupo	Grupo
CuentaCarga	Cuenta Carga
Capacidad	Capacidad
Trabajador	Código de trabajador
CursoProgramadoGrupo	Código de Curso Programado Grupo

CLASE	Ciclo
DESCRIPCIÓN	Clase que representa el ciclo
	Atributos
Ciclo	Código de ciclo.
Descripcion	Nombre de ciclo.
Numero	Numero de ciclo.
Activo	Estado de ciclo.
Valor	Valor de ciclo.

CLASE	DepartamentoAcademico
DESCRIPCIÓN	Clase que representa el departamento académico.
	Atributos
DepartamentoAcademico	Código de departamento académico.
Descripcion	Nombre de departamento académico.
Facultad	Código de Facultad.
Escuela	Código de escuela.
Activo	Estado del departamento académico.

CLASE	Escuela
DESCRIPCIÓN	Clase que representa la Escuela
	Atributos
Escuela	Código de escuela.
Descripcion	Descripción de la escuela.

Facultad	Código Facultad.
Abreviatura	Abreviatura de la escuela.
FechaCreacion	Fecha Creación
FechaFuncionamiento	Fecha funcionamiento
DispositivoFuncionamiento	Dispositivo funcionamiento
DuracionAnios	Duración en años.
DuracionSemestres	Duración de semestres.
Creditos	Créditos.
NivelGrado	Nivel de grado.
NombreGrado	Nombre de grado.
NombreTituloVaron	Nombre del título en masculino.
NombreTituloDama	Nombre del título en femenino.
ConsiderarGrados	Considerar Grados
Activo	Estado de escuela.

CLASE	Curricula
DESCRIPCIÓN	Clase que representa la curricula.
	Atributos
Escuela	Código de escuela.
Curricula	Código de curricula.
Descripcion	Descripción de la curricula.
CreditosMinimosObligatorios	Numero de créditos mínimos obligatorios.
CreditosMinimosElectivos	Numero de créditos mínimos electivos.
SemestreInicio	Código de semestre de inicio.
SemestreFin	Código de semestre de fin.
Ciclo	Código de Ciclo
Vigente	Vigente
Activo	Estado de la curricula.

CLASE	Rendimiento
DESCRIPCIÓN	Clase que representa el rendimiento académico de los alumnos.
	Atributos
Sede	Código de sede.
Escuela	Código de escuela.
Curricula	Código de curricula.
Curso	Código de curso.
Semestre	Código de semestre.
Seccion	Sección.
Alumno	Código de alumno.
TipoAprobado	Código de tipo aprobado.
Rendimiento	Código de rendimiento.
Grupo	Grupo.
Trabajador	Código de trabajador.
Promedio	Promedio.
PromedioIngresado	Promedio ingresado
Digitado	Digitado.
Inconsistencia	Errores.
Corregido	Corregido
Aplazado	Nota de aplazados.
EsAplazado	Es aplazado.
PromedioAntes	Promedio anterior.
Inhabilitado	Inhabilitado
TipoEquivalencia	Tipo de equivalencia
Escuela1	Código de escuela primero.
Curricula1	Código de curricula primero.
Curso1	Código de curso primero.
Escuela2	Código de escuela segundo.
Curricula2	Código de curricula segundo.
Curso2	Código de curso segundo.
Resolucion	Código de resolución.
TipoResolucion	Código de Tipo de Resolución.
FichaMatricula	Código de Ficha de Matricula.

Vez	Número de veces que el alumno lleva el mismo curso.
FechaRegistro	Fecha de registro.
EstacionRegistro	Estación de registro.
UsuarioRegistro	Código de trabajador que registro.
Observaciones	Observaciones.
Encuesta	Código de encuesta.
EstacionIngreso	Estación de Ingreso
UsuarioIngreso	Código de trabajador que registro el ingreso
FechaIngreso	Fecha de ingreso
EP	EP
EF	EF
ES	ES

CLASE	TipoAprobado
DESCRIPCIÓN	Clase que representa el tipo de aprobación.
	Atributos
TipoAprobado	Código de tipo de aprobado
Descripcion	Descripción del tipo de aprobado.
Activo	Estado de tipo de aprobación.

CLASE	TipoResolucion
DESCRIPCIÓN	Clase que representa el tipo de resolución.
	Atributos
TipoResolucion	Código de tipo de resolución.
Descripcion	Descripción

CLASE	Resolucion
DESCRIPCIÓN	Clase que representa la resolución.
	Atributos
Resolucion	Código de resolución.
TipoResolucion	Tipo Resolución
ResolucionInterna	Resolución Interna
TipoGrado	Código de tipo de grado.
Alumno	Código de alumno.
Semestre	Código de semestre.
Cargo	Cargo
ApellidosAutoridad	Apellidos Autoridad.
NombresAutoridad	Nombres Autoridad.
Fecha	Fecha.
FechaActo	Fecha Acto.
FechaExpedicion	Fecha Expedición.
FechaConsejo	Fecha Consejo.
FechaResolucion	Fecha Resolución.
Firmante1	Código del trabajador que firmo.
Firmante2	Código del trabajador que firmo.
Firmante3	Código del trabajador que firmo.
Firmante4	Código del trabajador que firmo.
Observacion	Observación
Activo	Estado de la resolución.
Usuario	Código de trabajador que registro la resolución.
Estacion	Estación
FechaRegistro	Fecha Registro.

CLASE	Semestre
DESCRIPCIÓN	Clase que representa la ficha de Matricula
	Atributos
Semestre	Código de semestre.
Descripcion	Nombre de semestre.
Activo	Estado de semestre.

CLASE	Respuesta
DESCRIPCIÓN	Clase que representa la respuesta de las encuestas.
	Atributos
Sede	Código de sede.
Escuela	Código de escuela.
Curricula	Código de curricula.
Curso	Código de curso.
Semestre	Código de semestre.
Seccion	Sección.
Alumno	Código de alumno.
Encuesta	Código de Encuesta.
Pregunta	Codigo de pregunta.
Valoracion	Valoración.

CLASE	EncuestaPregunta
DESCRIPCIÓN	Clase que representa la Encuesta y preguntas.
	Atributos
Encuesta	Código de la encuesta.
Pregunta	Código de la pregunta.

CLASE	Encuesta
DESCRIPCIÓN	Clase que representa la Encuesta.
	Atributos
Encuesta	Código de la encuesta.
Descripcion	Descripción de la Encuesta.
Semestre	Semestre.
Activo	Estado de la Encuesta.

CLASE	Pregunta
DESCRIPCIÓN	Clase que representa la Pregunta.
	Atributos
Pregunta	Código de la pregunta.
Descripcion	Descripción de la pregunta
Observacion	Observación.
Activo	Estado de la pregunta.

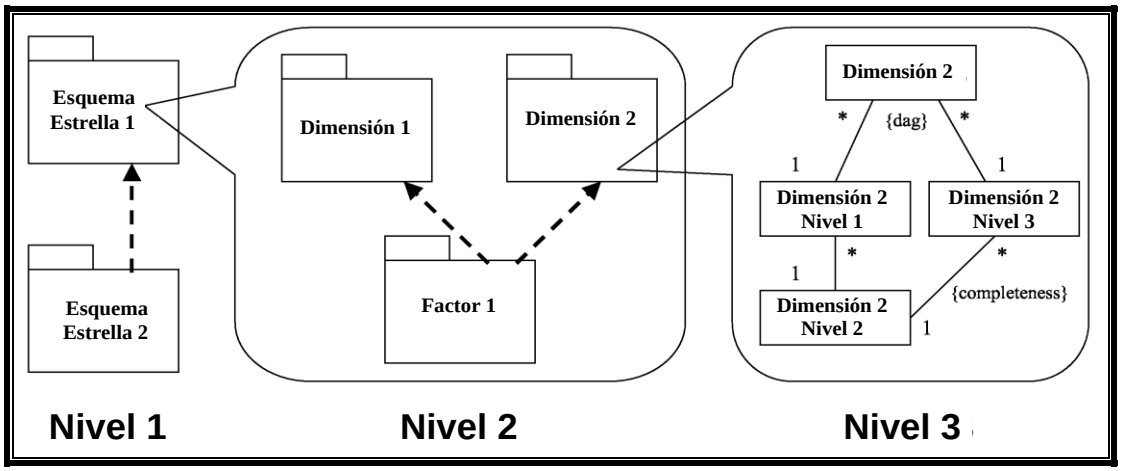
Anexo N°9

Esquema Conceptual del Data Mart

ESQUEMA CONCEPTUAL DEL DATA MART

El esquema conceptual de Data Mart se desarrolló utilizando la metodología DWEP, la cual toma como base la notación UML para representar el diseño conceptual del modelo multidimensional. El esquema conceptual de Data Mart establece el proceso de diseño en tres niveles, la cual es ilustrada en la Ilustración Nº1 y detallada línea abajo.

Ilustración Nº 1. Los tres niveles de un modelo multidimensional representados mediante paquetes de UML



Fuente: Sergio Luján

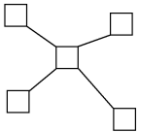
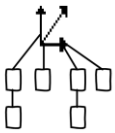
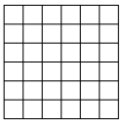
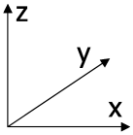
- **Nivel 1**: Definición del modelo. Un paquete representa un esquema estrella de un modelo multidimensional.

- **Nivel 2**: Definición de un esquema estrella. Un paquete representa un hecho.

- **Nivel 3**: Definición de un hecho o dimensión. Se compone de un conjunto de clases que representan los niveles jerárquicos en un paquete de dimensión o el esquema estrella completo en el caso de un paquete hecho.

La metodología DWEP propone varios estereotipos para el esquema conceptual, de las cuales se detallan en el Cuadro Nº Nº1- Conceptos multidimensionales y su representación en UML.

Cuadro Nº 1: Conceptos multidimensionales y su representación en UML

Concepto MD (Estereotipo)	Descripción	Icono
StarPackage	Paquetes de este estereotipo representan esquemas estrellas, compuestos de hechos y dimensiones; se emplea en el nivel 1	
FactPackage	Paquetes de este estereotipo representan hechos, compuestos de medidas y relacionados con las dimensions; se emplea en el nivel 2	
DimensionPackage	Paquetes de este estereotipo representan dimensiones, compuestas de jerarquías; se emplea en el nivel 2	
Fact	Clases de este estereotipo representan hechos, compuestos de medidas; se emplea en el nivel 3	
Dimension	Clases de este estereotipo representan dimensiones, compuestas de jerarquías; se emplea en el nivel 3	
Base	Clases de este estereotipo representan niveles de jerarquía en una dimensión; se emplea en el nivel 3	

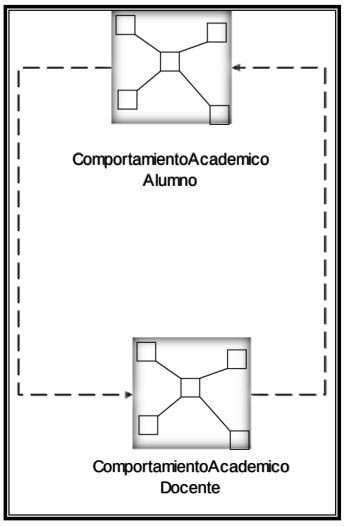
Fuente: Sergio Luján

1. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 1

En este nivel se muestra la relación entre los esquemas estrella que conforman la solución.

Existen dos esquemas estrella en la solución, mostrada en la Ilustración Nº 2. El primer esquema es el de Comportamiento Académico Estudiante el cual presenta datos ideales para el análisis a detalle del rendimiento académico de todos los estudiantes. El segundo esquema es de comportamiento Docente, la cual integra todos los datos necesarios para el análisis del rendimiento de los docentes.

Ilustración Nº 2. Esquema conceptual del Data Mart-Nivel 1



Fuente: Elaboración Propia

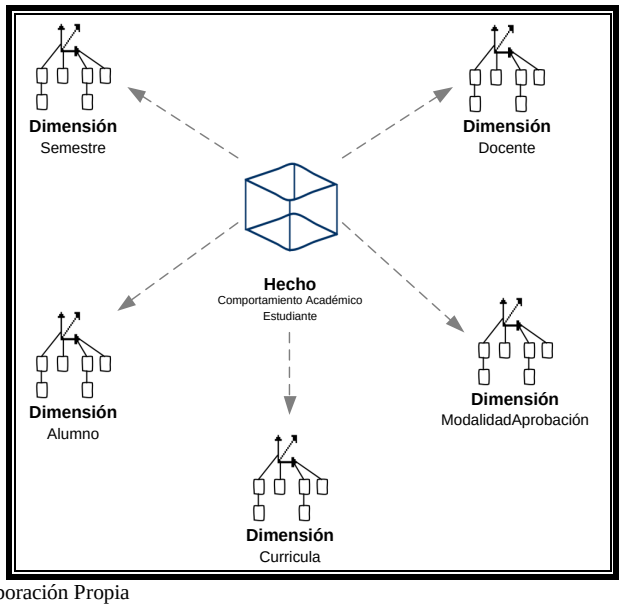
2. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 2

En este nivel se desarrolla el detalle de cada esquema presentado en el nivel 1. Se presenta las dimensiones con que se relaciona cada tabla de hecho.

A. ESQUEMA COMPORTAMIENTO ACADÉMICO ESTUDIANTE

El esquema de Comportamiento Académico Alumno tiene como tabla central a la tabla de hechos “Comportamiento Académico Alumno”, denotado por “ComAcadEstudiante”. La cual contiene la información del rendimiento académico de los alumnos de la Facultad de ciencias a detalle.

Figura Nº 03- Esquema conceptual del Comportamiento Académico del estudiante

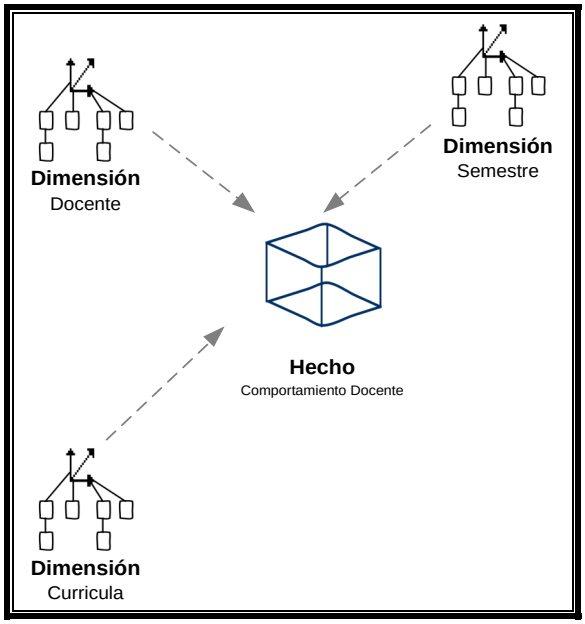


Fuente: Elaboración Propia

B. ESQUEMA COMPORTAMIENTO DOCENTE

El esquema de comportamiento Docente tiene como tabla principal a la tabla de hechos Comportamiento Docente, la cual contiene la información acerca de los cursos que dicta un docente y la valoración que los estudiantes le dan a los docentes.

Figura Nº 04- Esquema conceptual del Comportamiento Docente



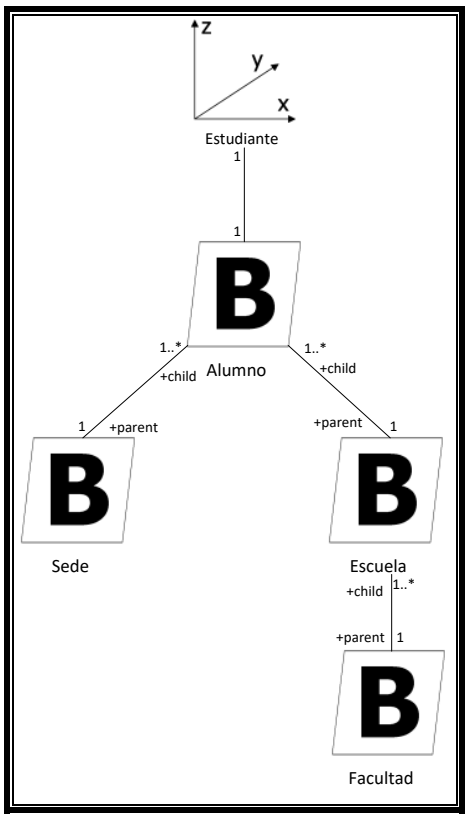
Fuente: Elaboración Propia

3. ESQUEMA CONCEPTUAL DEL DATA MART - NIVEL 3

En este nivel se desarrolla el detalle de cada dimensión presentado en el nivel anterior para poder definir la información que contiene y sobre todo la jerarquía en la información contenida por cada dimensión. La especificación de cada dimensión a nivel lógico podrá observarse en el entregable de Esquema Lógico de Data Mart.

A. DIMENSION ESTUDIANTE

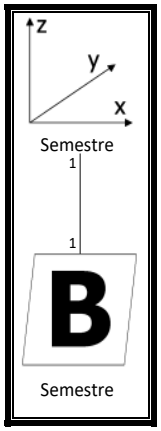
Figura Nº 05- Esquema conceptual de la Dimensión Estudiante



Fuente: Elaboración Propia

B. DIMENSION SEMESTRE

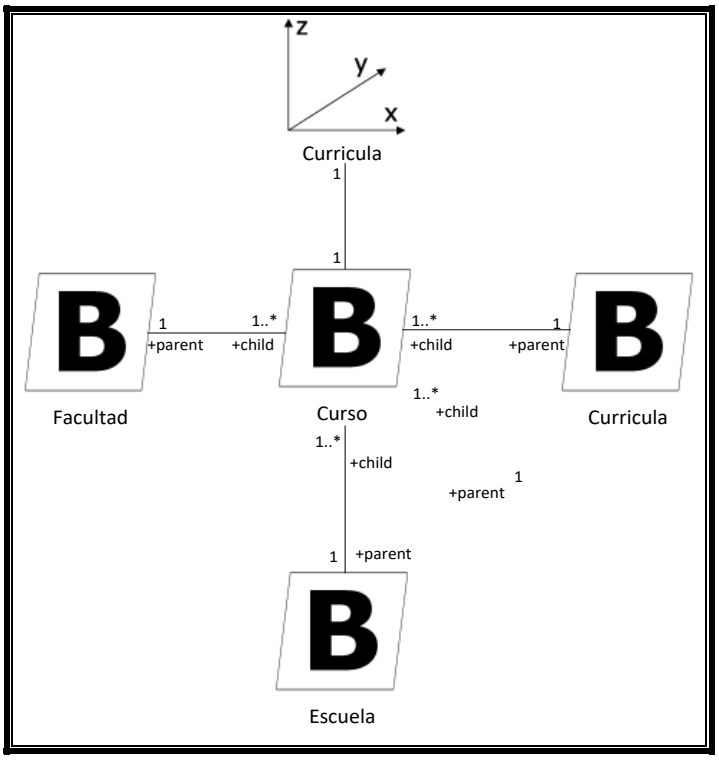
Figura Nº 06- Esquema conceptual de la Dimensión Estudiante



Fuente: Elaboración Propia

C. DIMENSION CURRICULA

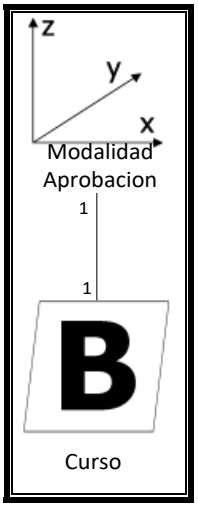
Figura Nº 07- Esquema conceptual de la Dimensión Estudiante



Fuente: Elaboración Propia

D. DIMENSION MODALIDAD DE APROBACION

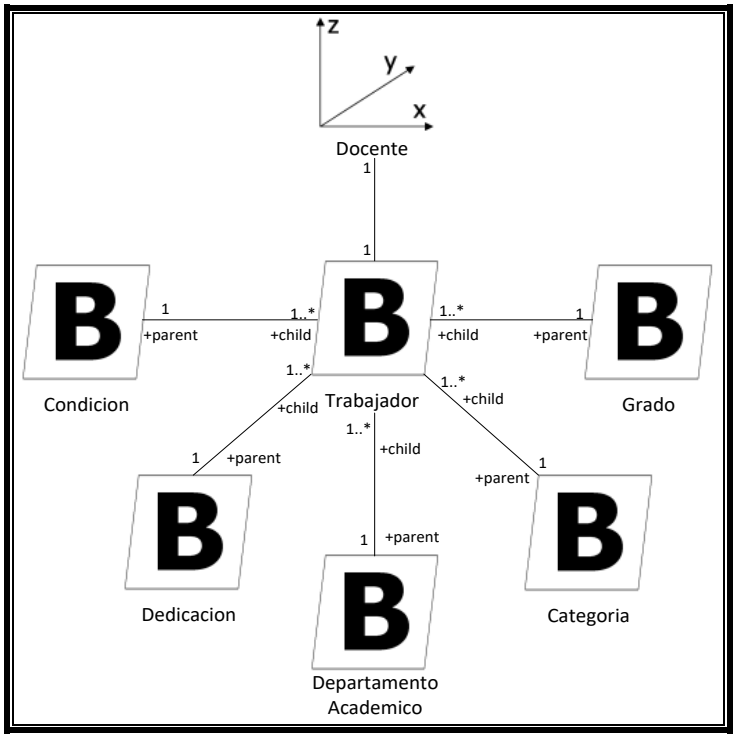
Figura Nº 08- Esquema conceptual de la Dimensión Estudiante



Fuente: Elaboración Propia

E. DIMENSION DOCENTE

Figura Nº 09- Esquema conceptual de la Dimensión Estudiante



Fuente: Elaboración Propia

4. ANÁLISIS DIMENSIONAL

4.1. ANÁLISIS DE TEMAS

A continuación, se presentan los temas que abarca el Data Mart de la Facultad de ciencias.

4.1.1. COMPORTAMIENTO ACADÉMICO ESTUDIANTE

Se debe analizar los registros del rendimiento académico (promedios finales) de los estudiantes, que se generan al finalizar un semestre académico, para lograr cumplir los objetivos de conocer el detalle del rendimiento académico de estudiantes de cada una de las Escuelas y Facultades de Universidad Santiago Antunez de Mayolo.

El resultado se mide por:

- Semestre académico.
- Ítem que puede ser un Alumno, Curricula, Escuela,Facultad.
- Tipo de Aprobación (Convalidaciones por equivalencias, Traslado interno, Matricula, Examen de suficiencia o traslado externo)
- promedio final, promedio anterior, promedio aplazados está expresado en el sistema vigesimal.

4.1.2. COMPORTAMIENTO DOCENTE

Se debe analizar los registros del rendimiento docente (valoración de las encuestas) que se generan al finalizar un semestre académico, que cumplen con el objetivo de conocer el detalle del rendimiento docente de cada una de las Escuelas y Facultades de Universidad Nacional Santiago Antúnez de Mayolo.

El resultado se mide por:

- Semestre académico.
- Ítem que puede ser un Docente, Curricula, Escuela,Facultad.
- Valoración está expresado en una escala de 0 a 4.

4.2. DIMENSIONES

Las dimensiones que conforman el Data Mart están detalladas en el Cuadro Nº Nº 2.

Cuadro Nº Nº2- Esquema conceptual de la Dimensión Estudiante

Nº	Dimensiones
1	Estudiante
2	Docente
3	Curricula
4	Semestre
5	Modalidad de Aprobación

Fuente: Elaboración Propia

4.2.1. DIMENSIÓN ESTUDIANTE

A. DESCRIPCIÓN

Dimensión que representa a los estudiantes que conforman la población estudiantil de la Universidad Nacional Santiago Antúnez de Mayolo.

B. JERARQUÍAS

Cuadro Nº 3- Esquema conceptual de la Dimensión Estudiante

Nombre del Atributo	Nivel	Contenido			
		Tipo de dato	Descripción	Formato	Valor por Defecto
CodigoAlumno	Nivel 1	varchar(14)	Codigo de Estudiante	Alfanumérico	Ninguno
Codigo	Nivel 2	varchar(12)	Codigo de Estudiante en la universidad	Alfanumérico	Ninguno
NombreCompleto	Nivel 2	varchar(150)	Nombre completo del estudiante	Mayúscula	Ninguno
Sede	Nivel 2	varchar(100)	Nombre de la sede	Mayúscula	Ninguno
Facultad	Nivel 2	varchar(100)	Nimbre de la Facultad	Mayúscula	Ninguno
Escuela	Nivel 2	varchar(100)	Nombre de la escuela	Mayúscula	Ninguno
Genero	Nivel 2	varchar(10)	Genero del estudiante	Mayuscula	'Sin dato'

Fuente: Elaboración Propia

4.2.2. DIMENSIÓN DOCENTE

A. DESCRIPCIÓN

Dimensión que representa a los Docentes que laboran en la Universidad Nacional Santiago Antúnez de Mayolo.

B. JERARQUÍAS

Cuadro Nº 4- Esquema conceptual de la Dimensión Estudiante

Nombre del Atributo	Nivel	Contenido			
		Tipo de dato	Descripción	Formato	Valor por Defecto
CodigoDocente	Nivel 1	varchar(6)	Codigo de Docente	Alfanumérico	Ninguno
NombreDocente	Nivel 1	varchar(150)	Nombre completo del Docente	Mayúscula	Ninguno
Dedicacion	Nivel 2	varchar(100)	Nombre de la dedicación del docente	Mayúscula	'Sin dato'
Categoria	Nivel 2	varchar(100)	Nombre de la categoria del docente	Mayúscula	'Sin dato'
Condicion	Nivel 2	varchar(100)	Nombre de la condición del Docente	Mayúscula	Sin dato'
Departamento Academico	Nivel 2	varchar(100)	Nombre del departamento académico	Mayúscula	Ninguno
Grado	Nivel 2	varchar(100)	Nombre del grado académico del docente	Mayúscula	'Sin dato'
Profesion	Nivel 2	varchar(100)	Nombre de la profesión del docente	Mayúscula	'Sin dato'

Fuente: Elaboración Propia

4.2.3. DIMENSIÓN CURRICULA

A. DESCRIPCIÓN

Dimensión que representa a la malla curricular de las diversas carreras profesionales de la Universidad Nacional Santiago Antúnez de Mayolo.

B. JERARQUÍAS

Cuadro Nº 5- Esquema conceptual de la Dimensión Estudiante

Nombre del Atributo	Nivel	Contenido			
		Tipo de dato	Descripción	Formato	Valor por Defecto
CodigoCurricula	Nivel 1	varchar(16)	Codigo de Currícula	Alfanumérico	Ninguno
Facultad	Nivel 2	varchar(100)	Nombre de la Facultad	Mayúscula	Ninguno
Escuela	Nivel 2	varchar(100)	Nombre de la Escuela	Mayúscula	Ninguno
Curricula	Nivel 1	char(2)	Numero de Curricula	0	Ninguno
NombreCurso	Nivel 2	varchar(100)	Nombre del curso	Mayúscula	Ninguno
Ciclo	Nivel 2	char(4)	Ciclo académico	Mayúscula	Ninguno

TipoCurso	Nivel 2	varchar(100)	Nombre del tipo de curso	Mayúscula	Ninguno
HorasTeorica	Nivel 2	int	Número de Horas teóricas	0	Ninguno
HorasPractica	Nivel 2	int	Número de Horas Prácticas	0	Ninguno

Fuente: Elaboración Propia

4.2.4. DIMENSIÓN SEMESTRE

A. DESCRIPCIÓN

Dimensión que representa a los semestres académicos que se han desarrollado en la Universidad Nacional Santiago Antúnez de Mayolo.

B. JERARQUÍAS

Cuadro Nº 6- Esquema conceptual de la Dimensión Estudiante

Nombre del Atributo	Nivel	Contenido			
		Tipo de dato	Descripción	Formato	Valor por Defecto
CodigoSemestre	Nivel 1	varchar(6)	Codigo del semestre	Alfanumérico	Ninguno
Semestre	Nivel 2	varchar(100)	Nombre del semestre	Mayúscula	Ninguno

Fuente: Elaboración Propia

4.2.5. DIMENSIÓN MODALIDAD DE APROBACIÓN

A. DESCRIPCIÓN

Dimensión que representa a la modalidad de aprobación a las cuales está sujeta la evaluación.

B. JERARQUÍAS

Cuadro Nº 7- Esquema conceptual de la Dimensión Estudiante

Nombre del Atributo	Nivel	Contenido			
		Tipo de dato	Descripción	Formato	Valor por Defecto
CodigoAprobacion	Nivel 1	varchar(2)	Código de modalidad de aprobación	Alfanumérico	Ninguno
ModalidadAprobacion	Nivel 2	varchar(100)	Nombre de la Modalidad de Aprobación	Mayúscula	Ninguno

Fuente: Elaboración Propia

4.3. Hechos versus Dimensiones

Cuadro Nº 8- Hechos versus Dimensiones

	FACTS O TABLA DE HECHOS	
DIMENSIONES	Comportamiento Académico Estudiante	Comportamiento Docente
Estudiante	**X**	
Curricula	**X**	**X**
Semestre	**X**	**X**
Modalidad de Aprobación	**X**	
Docente	**X**	**X**

Fuente: Elaboración Propia

4.3.1. HECHO COMPORTAMIENTO ACADÉMICO ESTUDIANTE

A. DESCRIPCIÓN

Hecho que representa el comportamiento académico de los estudiantes durante los diferentes semestres académicos.

B. GRANULARIDAD

Cuadro Nº 9- Granularidad de las Dimensiones del Comportamiento académico del Estudiante.

Nº	Nombre de la Dimensión	Nombre del Atributo	Descripción	Llave Primaria
1	Curricula	CodigoCurricula	Curricula utilizada	Si
2	Estudiante	CodigoAlumno	Estudiante que realiza el proceso	Si
3	Docente	CodigoDocente	Docente encargado de la evaluación	Si
4	Semestre	CodigoSemestre	Semestre en que realizó el proceso	Si
5	Modalidad Aprobación	CodigoAprobacion	Modalidad de aprobación	Si

Fuente: Elaboración Propia

C. MEDIDAS

Cuadro Nº 9- Medidas de las Dimensiones del Comportamiento académico del Estudiante.

N°	Nombre	Descripción
1	PromedioFinal	Promedio final obtenido por un estudiante en un determinado curso.

2	Promedio Anterior	Nota de aplazados obtenido por un estudiante en un determinado curso.
3	NotaAplazados	Promedio obtenido por un estudiante en un determinado curso en un semestre anterior.
4	Condicion	Condición de un estudiante respecto a un curso.

Fuente: Elaboración Propia

4.3.2. HECHO COMPORTAMIENTO DOCENTE

A. DESCRIPCIÓN

Hecho que representa el comportamiento docente que es evaluado por los estudiantes una vez durante un semestre.

B. GRANULARIDAD

Cuadro Nº 10- Granularidad de las Dimensiones del Hecho comportamiento docente.

Nº	Nombre de la Dimensión	Nombre del Atributo	Descripción	Llave Primaria
1	Curricula	CodigoCurricula	Curricula utilizada	Si
2	Docente	CodigoDocente	Docente que es evaluado	Si
3	Semestre	CodigoSemestre	Semestre en que realizó el proceso	Si
4		Seccion	Numero de sección que realizó la evaluación	Si
5		Encuesta	Nombre de la encuesta que se realiza.	Si
6		Pregunta	Pregunta que se realiza durante la encueta.	Si

Fuente: Elaboración Propia

C. MEDIDAS

Cuadro Nº 11- Medidas de las Dimensiones del Hecho comportamiento docente.

N°	Nombre	Descripción
1	Valoracion	Promedio de la valoración obtenida por pregunta.
2	Cantidad	Cantidad de estudiantes que evaluaron

Fuente: Elaboración Propia

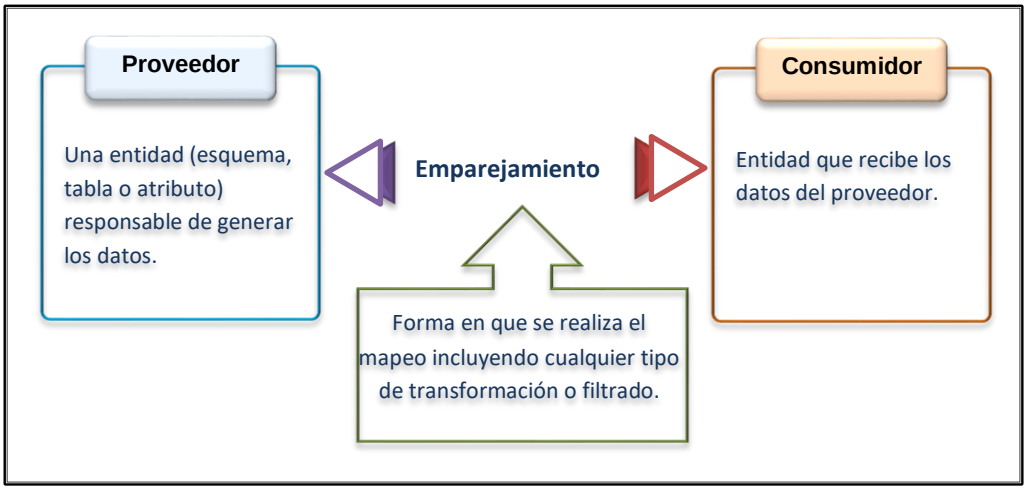
Anexo N°10

Diagrama de Mapeo de Datos

DIAGRAMA DE MAPEO DE DATOS

La metodología DWEP en su diagrama de Mapeo de Datos, basada en la notación UML, representa el flujo de datos y la captura de las interconexiones entre los elementos de diseño. Un mapeo de datos propiamente dicho se define mediante tres elementos lógicos las cuales son representados en la figura Nº 01.

Figura Nº 01. Mapeo de datos.



Fuente: Elaboración Propia.

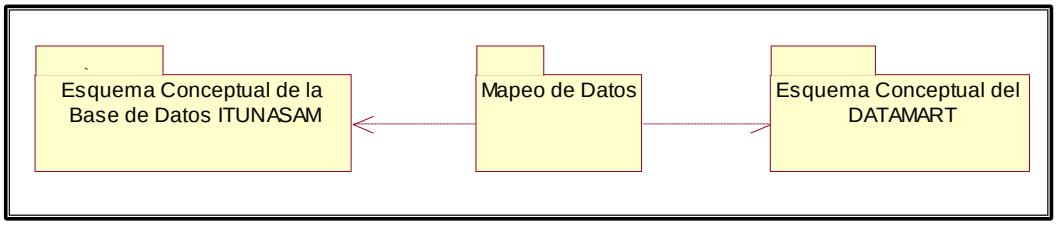
Según la metodología, el Diagrama de Mapeo de Datos propone cuatro niveles de granularidad: nivel 0, nivel 1, nivel 2 y nivel 3 o también conocidos como: el nivel de Base de datos, nivel de flujo de datos, nivel de tablas y nivel de atributos respectivamente. Al primer nivel cada esquema de almacén de datos se representa mediante un paquete, los mapeos entre los diferentes esquemas se modelan en un único paquete que encapsula todos los detalles. Al segundo nivel describe las relaciones de datos a nivel individual entre las fuentes de datos hacia los respectivos destinos. Al tercer nivel detalla todas las transformaciones intermedias que tienen lugar durante el flujo del primer nivel. El ultimo nivel detalla el mapeo a nivel de atributos.

En base a ello se elaboró los diagramas de mapeo de Datos que se muestran líneas abajo.

1. DIAGRAMA DE MAPEO DE DATOS – NIVEL 0

El diagrama de mapeo de datos a este nivel se muestra la relación de la base de datos fuente ITUNASAM con la base de datos del Data Mart a un nivel macro.

Figura Nº 02. Diagrama de Mapeo de datos nivel 0.

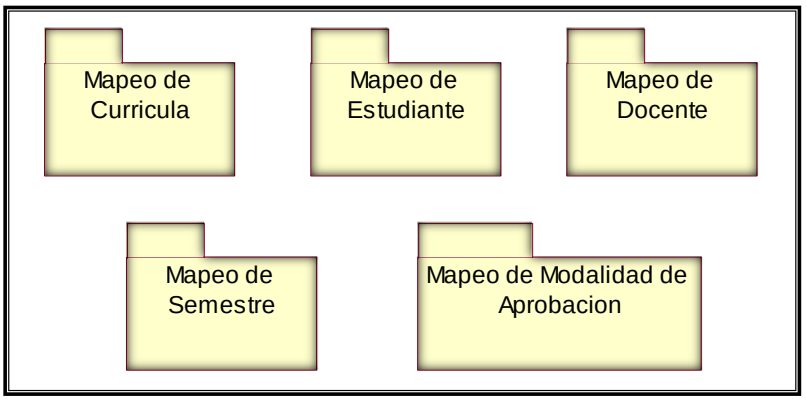


Fuente: Elaboración Propia.

2. DIAGRAMA DE MAPEO DE DATOS – NIVEL 1

A este nivel se detalla el paquete de Mapeo de Datos mostrado en el nivel anterior, figura Nª 02. Existen cinco dimensiones en el Data Mart en tal sentido el mapeo de datos para dichas dimensiones se muestra en los siguientes ítems.

Figura Nº 03. Diagrama de Mapeo de datos nivel 1.

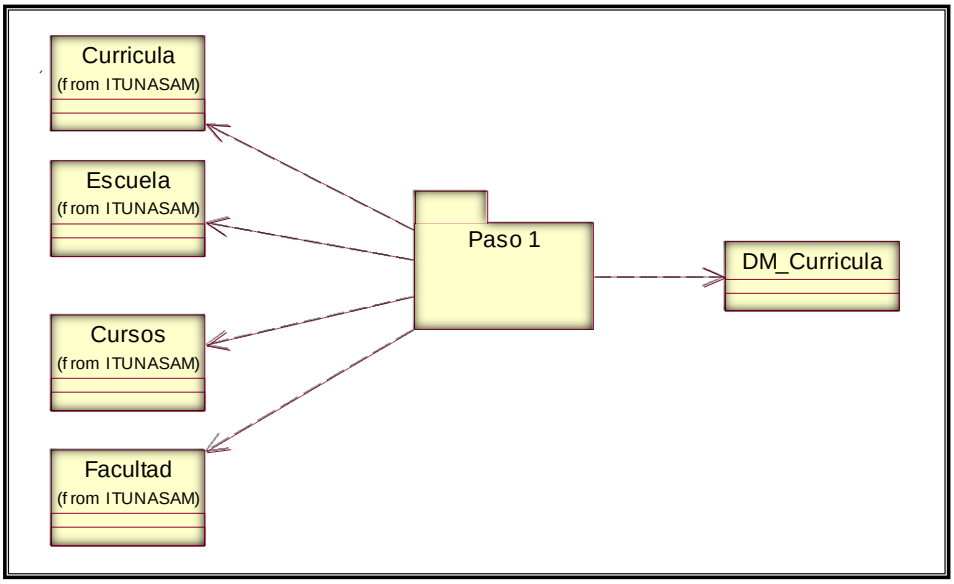


Fuente: Elaboración Propia.

3. DIAGRAMA DE MAPEO DE DATOS – NIVEL 2

3.1. MAPEO DE CURRICULA

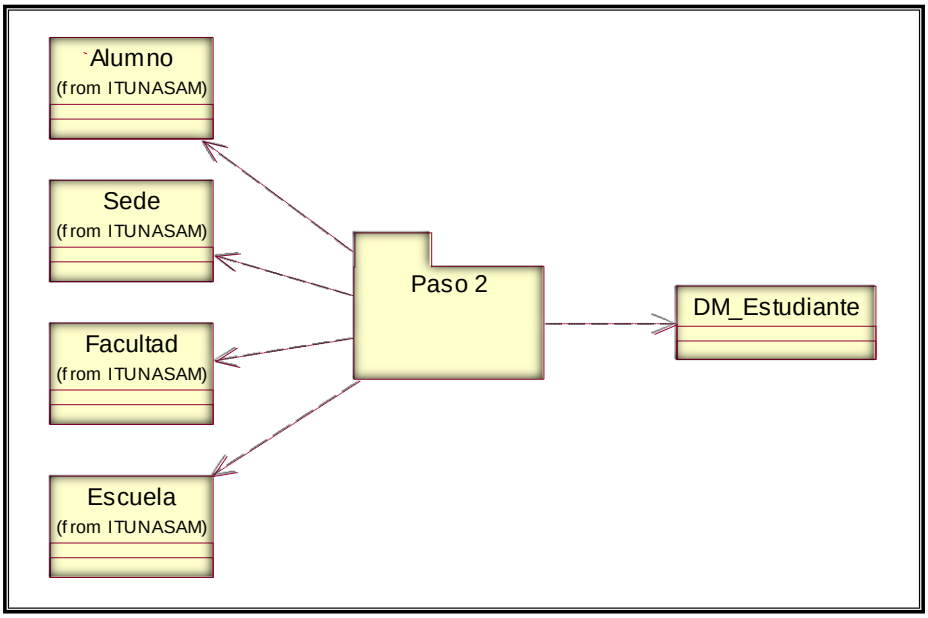
Figura Nª 04. Diagrama de Mapeo de datos nivel 2.



Fuente: Elaboración Propia.

3.2. MAPEO DE ESTUDIANTE

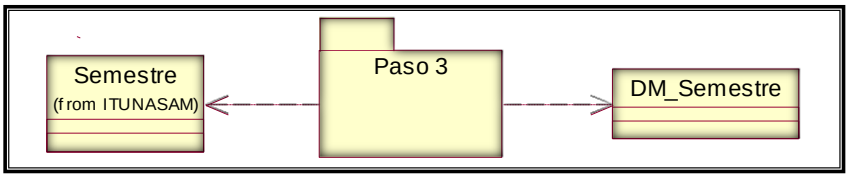
Figura Nª 05. Diagrama de Mapeo de datos de estudiante nivel 2.



Fuente: Elaboración Propia.

3.3. MAPEO DE SEMESTRE

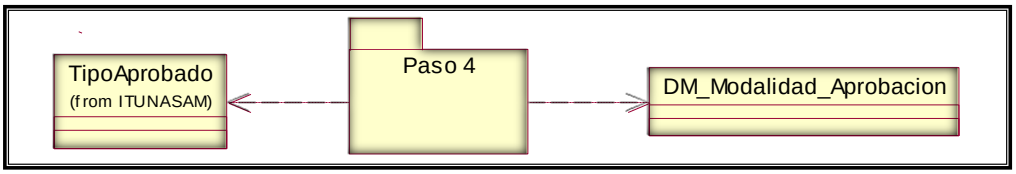
Figura Nª 06. Diagrama de Mapeo de semestre de datos nivel 2.



Fuente: Elaboración Propia.

3.4. MAPEO DE MODALIDAD DE APROBACIÓN

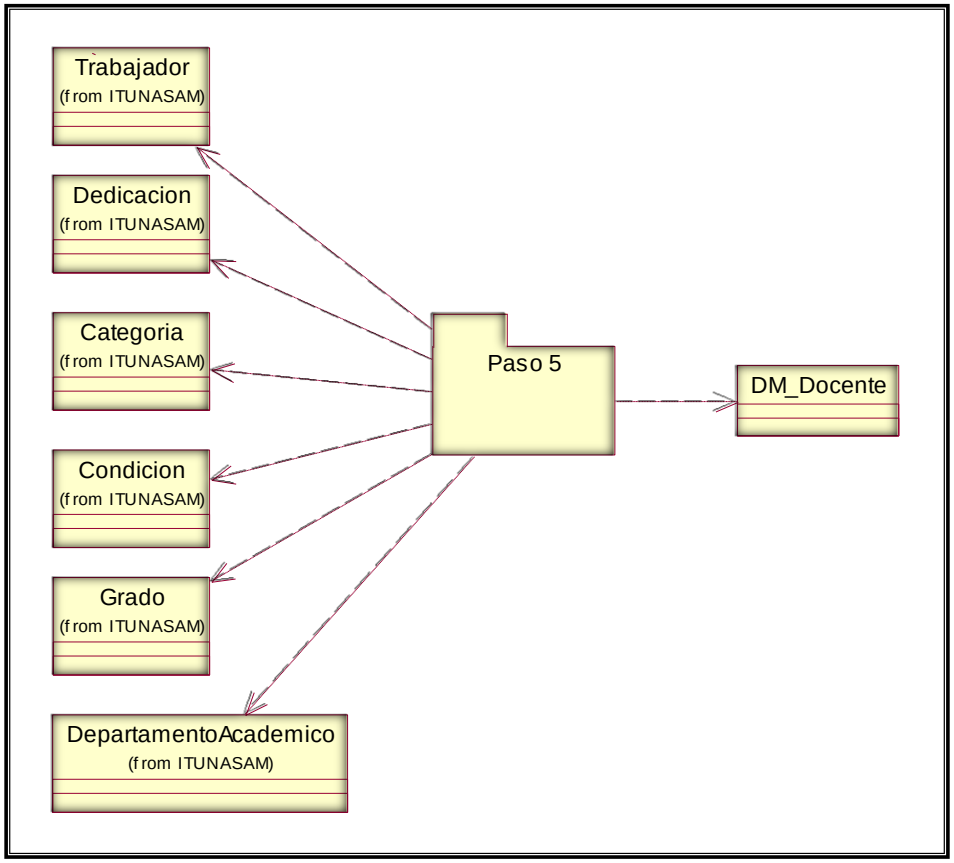
Figura Nª 07. Diagrama de Mapeo de modalidad de aprobación datos nivel 2.



Fuente: Elaboración Propia.

3.5. MAPEO DE DOCENTE

Figura Nª 08. Diagrama de Mapeo de docente de datos nivel 2.

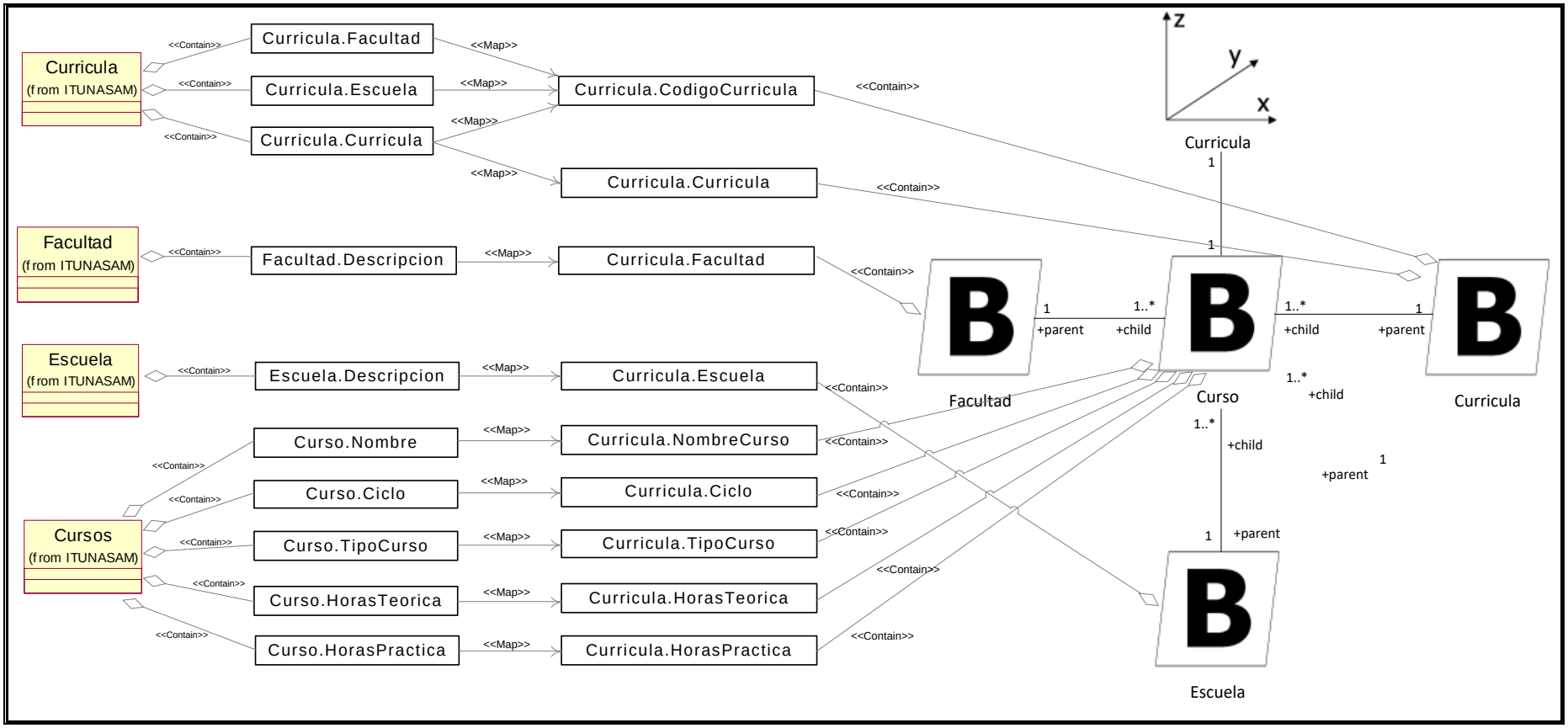


Fuente: Elaboración Propia

4. DIAGRAMA DE MAPEO DE DATOS – NIVEL 3

4.1. MAPEO DE CURRÍCULA

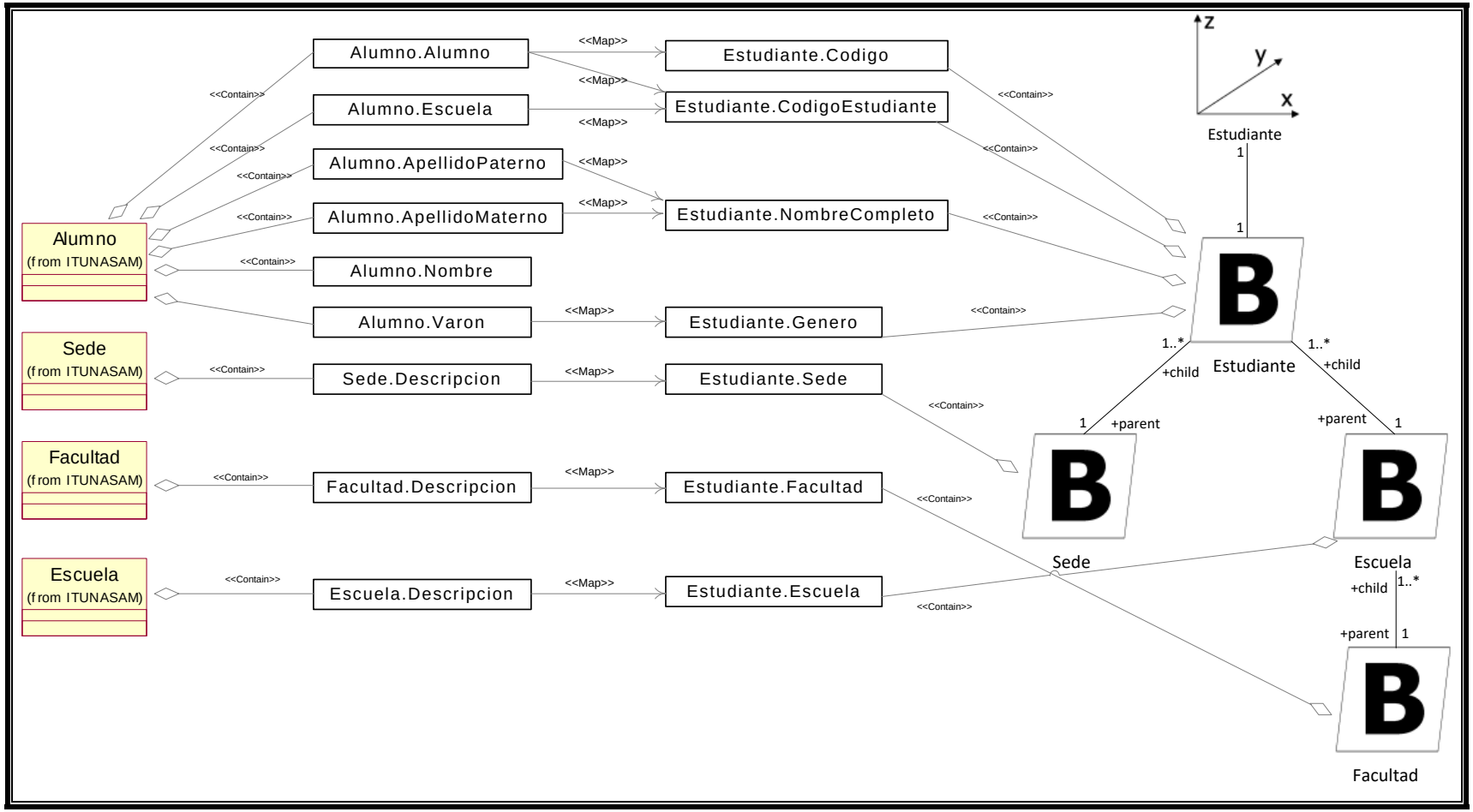
Figura Nª 09. Diagrama de Mapeo de datos de currícula nivel 3.



Fuente: Elaboración Propia.

4.2. MAPEO DE ESTUDIANTE

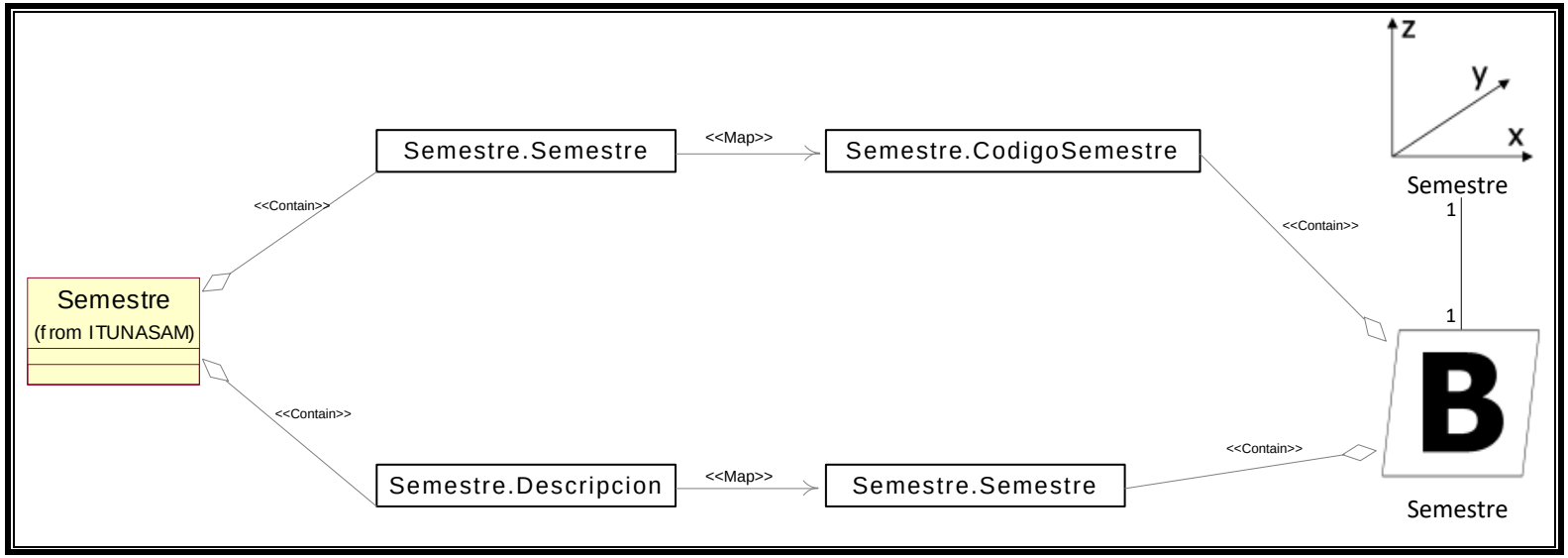
Figura Nª10. Diagrama de Mapeo de datos de estudiante nivel 3.



Fuente: Elaboración Propia.

4.3. MAPEO DE SEMESTRE

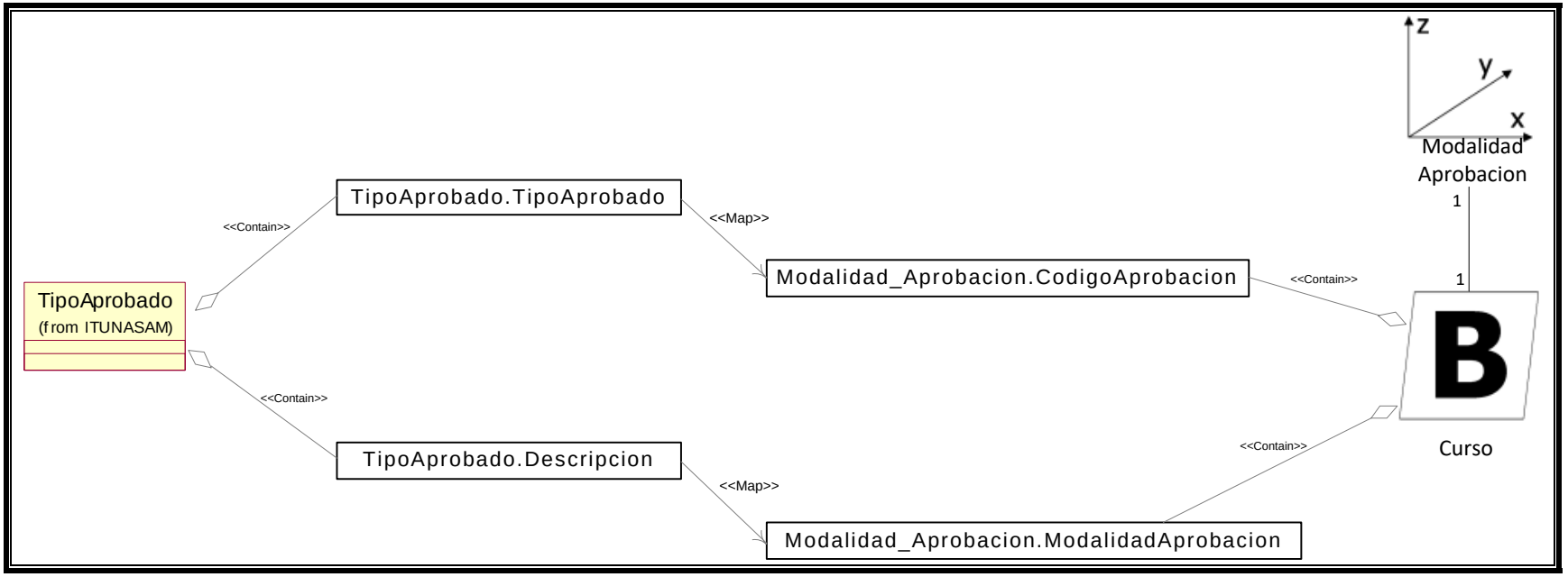
Figura Nª11. Diagrama de Mapeo de datos de semestre nivel 3.



Fuente: Elaboración Propia.

4.4. MAPEO DE MODALIDAD DE APROBACIÓN

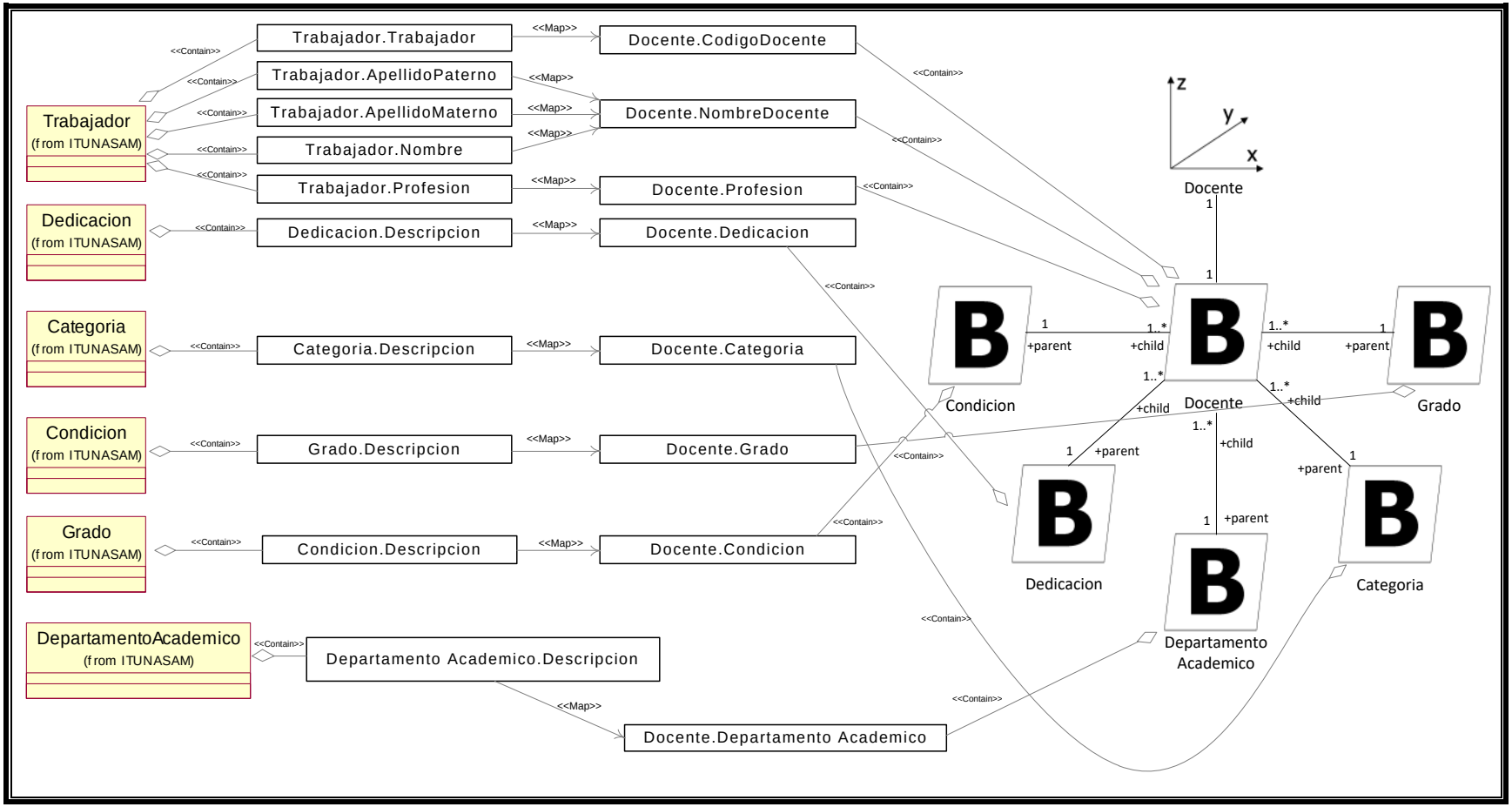
Figura Nª12. Diagrama de Mapeo de datos de modalidad de aprobación nivel 3.



Fuente: Elaboración Propia.

4.5. MAPEO DE DOCENTE

Figura Nª13. Diagrama de Mapeo de datos de docente nivel 3.



Fuente: Elaboración Propia.

Anexo Nº11

Esquema Lógico del Data Mart

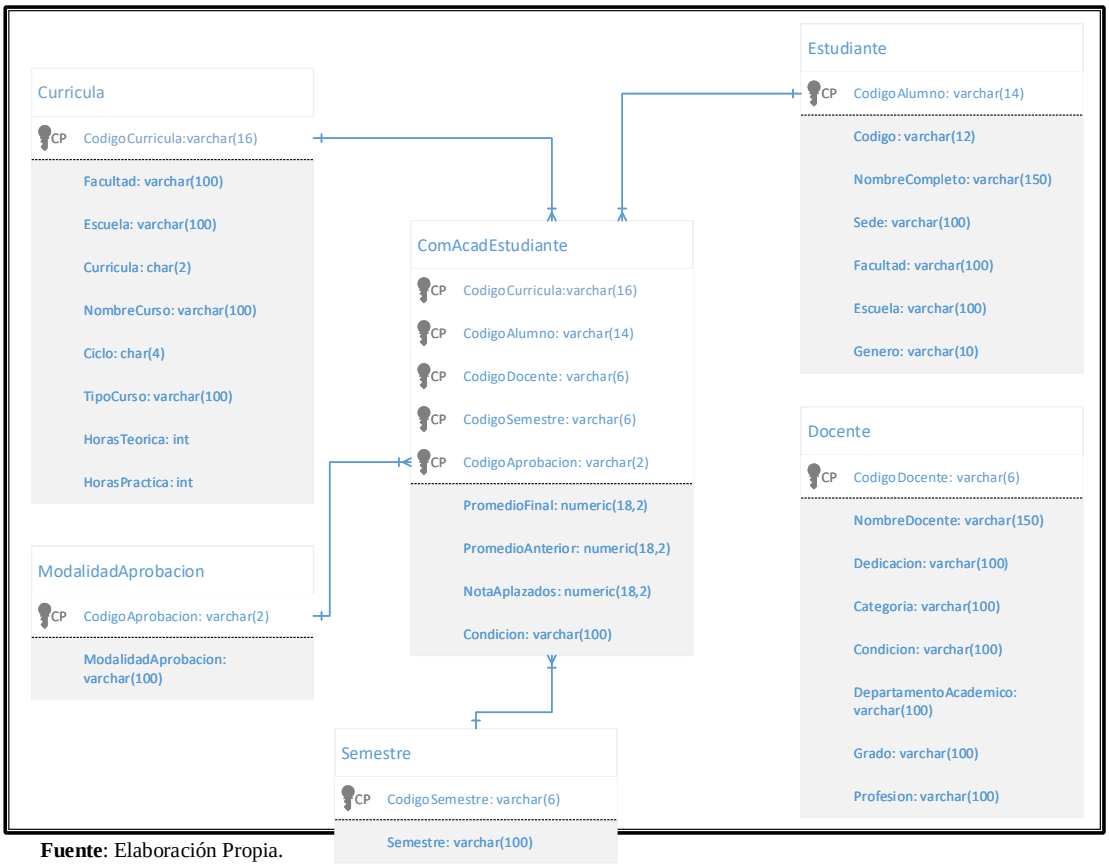
ESQUEMA LÓGICO DEL DATA MART

El esquema lógico del Data Mart permite representar las tablas de dimensiones y hechos, las cuales conforman los distintos modelos estrellas identificados en el esquema conceptual del Data Mart. En cada tabla, a este nivel, se presentan los campos que lo conforman sin obviar el tipo de dato. La metodología DWEP a este nivel propone el Esquema Lógico del Data Mart (Data Mart Logical Schema), que permite representar las tablas de dimensión y de hechos que conforman cada modelo estrella.

1. ESQUEMAS LÓGICOS DEL DATA MART

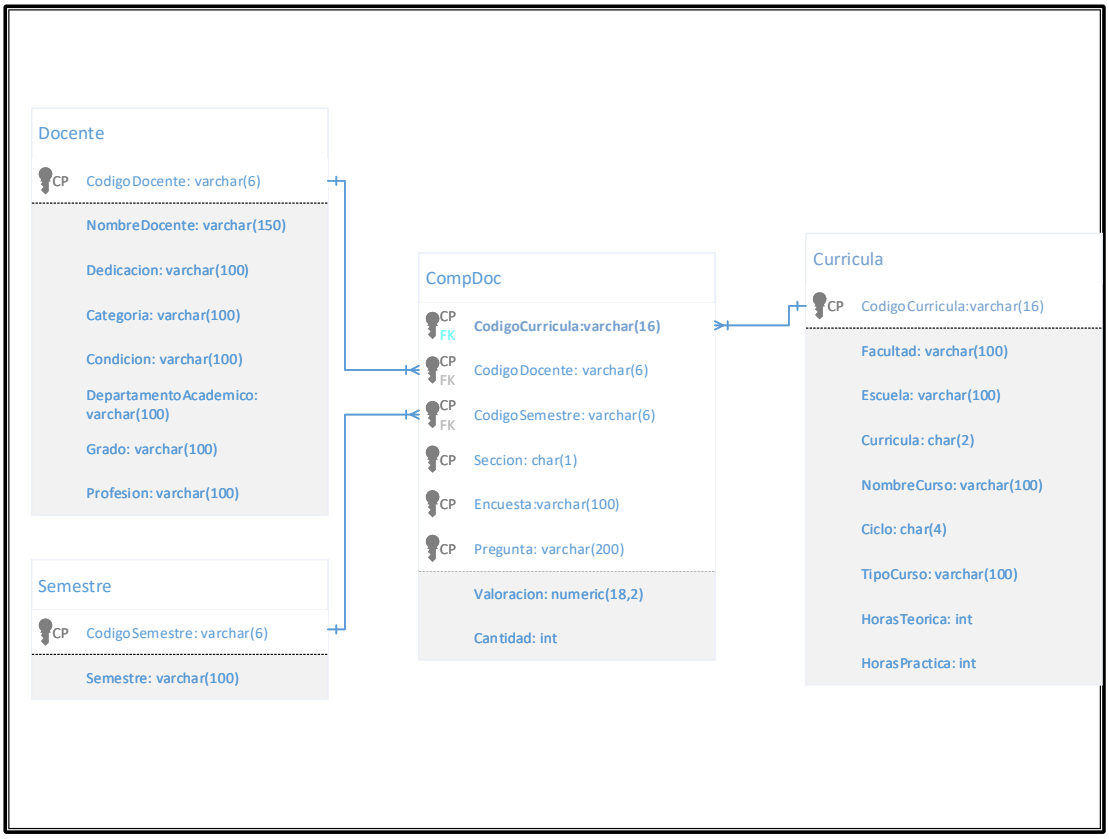
Dentro de los esquemas lógicos del Data Mart se tiene al esquema estrella del comportamiento académico estudiante y al esquema estrella del comportamiento docente.

Ilustracion Nº1. Esquema estrella del Comportamiento académico del Estudiante



Fuente: Elaboración Propia.

Ilustracion Nº2. Esquema estrella del Comportamiento Docente



Fuente: Elaboración Propia.

2. DICCIONARIO DEL ESQUEMA LÓGICO DE DATA MART

Tabla	Curricula		
Descripción	Tabla dimensión que contiene los datos de las Curriculas de la Univerisidad		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoCurricula	varchar(16)	Codigo de Currícula	32
Facultad	varchar(100)	Nombre de la Facultad	200
Escuela	varchar(100)	Nombre de la Escuela	200
Curricula	char(2)	Numero de Curricula	4
NombreCurso	varchar(100)	Nombre del curso	200
Ciclo	char(4)	Ciclo académico	8
TipoCurso	varchar(100)	Nombre del tipo de curso	200
HorasTeorica	int	Numero de Horas teóricas	4
HorasPractica	int	Numero de Horas Prácticas	4
		Tamaño de la Tabla (Bytes)	852

Tabla	Estudiante		
Descripción	Tabla dimensión que contiene los datos de los Estudiantes		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoAlumno	varchar(14)	Codigo de Estudiante	28
Codigo	varchar(12)	Codigo de Estudiante en la universidad	24
NombreCompleto	varchar(150)	Nombre completo del estudiante	300
Sede	varchar(100)	Nombre de la sede	200
Facultad	varchar(100)	Nimbre de la Facultad	200
Escuela	varchar(100)	Nombre de la escuela	200
Genero	varchar(10)	Genero del estudiante	20
		Tamaño de la Tabla (Bytes)	972

Tabla	Semestre		
Descripción	Tabla dimensión que contiene los datos de los Semestres académicos		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoSemestre	varchar(6)	Codigo del semestre	12
Semestre	varchar(100)	Nombre del semestre	200
		Tamaño de la Tabla (Bytes)	212

Tabla	ModalidadAprobacion		
Descripción	Tabla dimensión que contiene los datos de la modalidad de aprobación		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoAprobacion	varchar(2)	Código de modalidad de aprobación	4
ModalidadAprobacion	varchar(100)	Nombre de la Modalidad de Aprobación	200
		Tamaño de la Tabla (Bytes)	204

Tabla	Docente		
Descripción	Tabla dimensión que contiene los datos de los docentes		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoDocente	varchar(6)	Codigo de Docente	12
NombreDocente	varchar(150)	Nombre completo del Docente	300
Dedicacion	varchar(100)	Nombre de la dedicación del docente	200
Categoria	varchar(100)	Nombre de la categoria del docente	200
Condicion	varchar(100)	Nombre de la condición del Docente	200
Departamento Academico	varchar(100)	Nombre del departamento académico	200
Grado	varchar(100)	Nombre del grado académico del docente	200
Profesion	varchar(100)	Nombre de la profesión del docente	200
		Tamaño de la Tabla (Bytes)	1512

Tabla	ComAcadEstudiante		
Descripción	Tabla Hecho que contiene los datos del comportamiento académico de los estudiantes		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoCurricula	varchar(16)	Codigo de Currícula	32
CodigoAlumno	varchar(12)	Codigo de Estudiante	24
CodigoDocente	varchar(6)	Codigo de Docente	12

CodigoSemestre	varchar(6)	Codigo del semestre	12
CodigoAprobacion	varchar(2)	Código de modalidad de aprobación	4
PromedioFinal	numeric(18,0)	Promedio final obtenido por un estudiante en un determinado curso.	36
Promedio Anterior	numeric(18,0)	Nota de aplazados obtenido por un estudiante en un determinado curso.	4
NotaAplazados	numeric(18,0)	Promedio obtenido por un estudiante en un determinado curso en un semestre anterior.	4
Condicion	varchar(100)	Condición de un estudiante respecto a un curso.	200
		Tamaño de la Tabla (Bytes)	328

Tabla	CompDoc		
Descripción	Tabla Hecho que contiene los datos del Comportamiento Academico Docente		
Atributo	**Tipo de Dato**	**Descripción**	**Tamaño**
CodigoCurricula	varchar(16)	Codigo de Currícula	32
CodigoDocente	varchar(6)	Codigo de Docente	12
CodigoSemestre	varchar(6)	Codigo del semestre	12
Seccion	char(1)	Numero de sección que realizó la evaluación	2
Encuesta	varchar(100)	Nombre de la encuesta que se realiza.	200
Pregunta	varchar(200)	Pregunta que se realiza durante la encueta.	400
Valoracion	numeric(18,2)	Promedio de la valoración obtenida por pregunta.	4
Cantidad	int	Cantidad de estudiantes que evaluaron	4
		Tamaño de la Tabla (Bytes)	666

Anexo Nº12

Esquema Físico del Data Mart

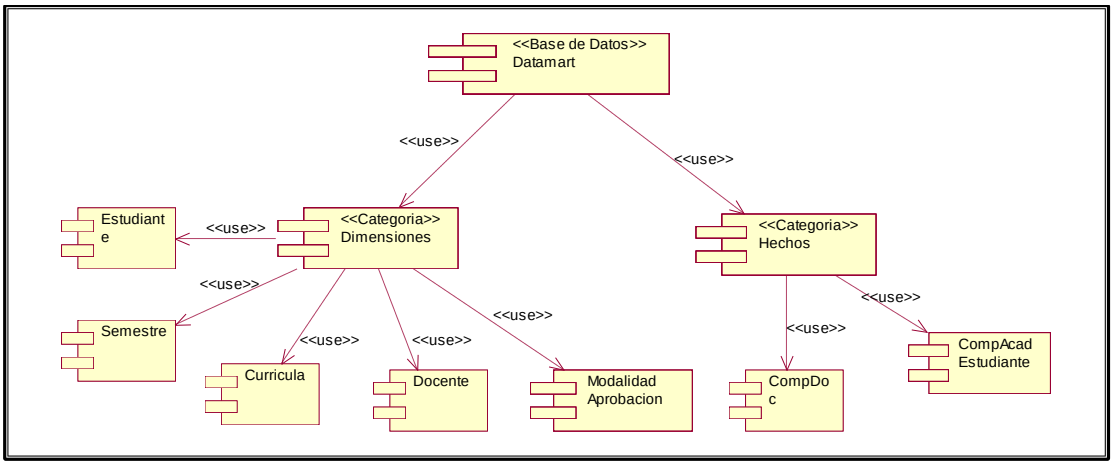
Esquema Físico del Data Mart

El esquema físico del Data Mart representa la organización física del modelo multidimensional del Data Mart. La metodología DWEP propone el diagrama de componentes y el diagrama de despliegue a este nivel.

1. Diagrama de Componentes

Este diagrama muestra la organización de la instancia de la base de datos del Data Mart, En la figura Nº 01 se muestra el diagrama de despliegue, la cual está estructurada en dos componentes: Dimensiones y Hechos, la primera contiene información relacionada a las tablas de dimensiones. La segunda contiene información de las tablas hechos, son estas las que contienen la mayor cantidad de registros.

Figura Nº 01. Diagrama de despliegue.



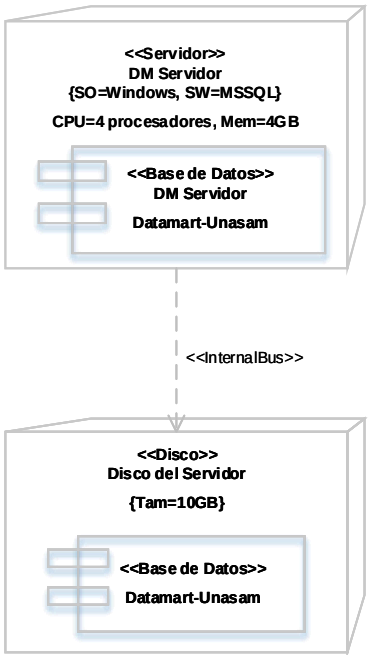
Fuente: Elaboración propia.

2. Diagrama de despliegue

El diagrama de despliegue muestra como está organizado el servidor que soportará la base de datos del Data Mart. En la figura Nº 02 se ilustra el servidor, la

cual deberá contar con un disco de un tamaño de 20GB (basada en la estimación realizada en el siguiente ítem).

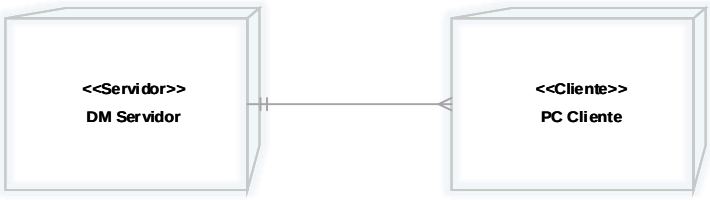
Figura Nº 02. Diagrama de despliegue del Servidor



Fuente: Elaboración Propia.

Finamente en la figura Nº 03 se muestra el diagrama de despliegue general, a la cual se conectará varias computadoras cliente las cuales solo tendrán acceso a la interfaz de usuario.

Figura Nº 03. Diagrama de despliegue



Fuente: Elaboración Propia.

3. Estimación del tamaño de la base de datos del Data Mart

Para la estimación del tamaño de la base de datos del Data Mart, se tomará en cuenta solo el tamaño de las tablas de hecho debido a que el tamaño de una tabla de dimensión es despreciable respecto a la de un hecho.

Para calcular el tamaño de las tablas de hecho, es necesario saber el tamaño de los tipos de datos que la conforman, que en este caso son los tipos de datos de MsSQL. Los tamaños en Bytes según el tipo de datos son:

Cuadro Nº 1. Tipos de datos MSSQL

Tipo de Dato	Presicion	Tamaño
numeric	1-9	5 bytes
real	7 dígitos	4 bytes
int		4 bytes
Varchar(n)		n+ 2 bytes
Char(n)		n bytes

Fuente: MSSQL.

A continuación, se muestra los cuadros de estimación de tamaño de las tablas de hecho por cada fila, para posteriormente realizar la estimación aproximada por el número de registros en el tiempo, teniendo en cuenta que para la cantidad de registros por dimensión se asume que la universidad tiene 118 semestres cumplidos y se realiza el cálculo aproximado de 10 semestres adicionales. Teniendo en cuenta el tamaño de las tablas hecho es superior que las tablas de dimensión, solo se realizó la estimación de

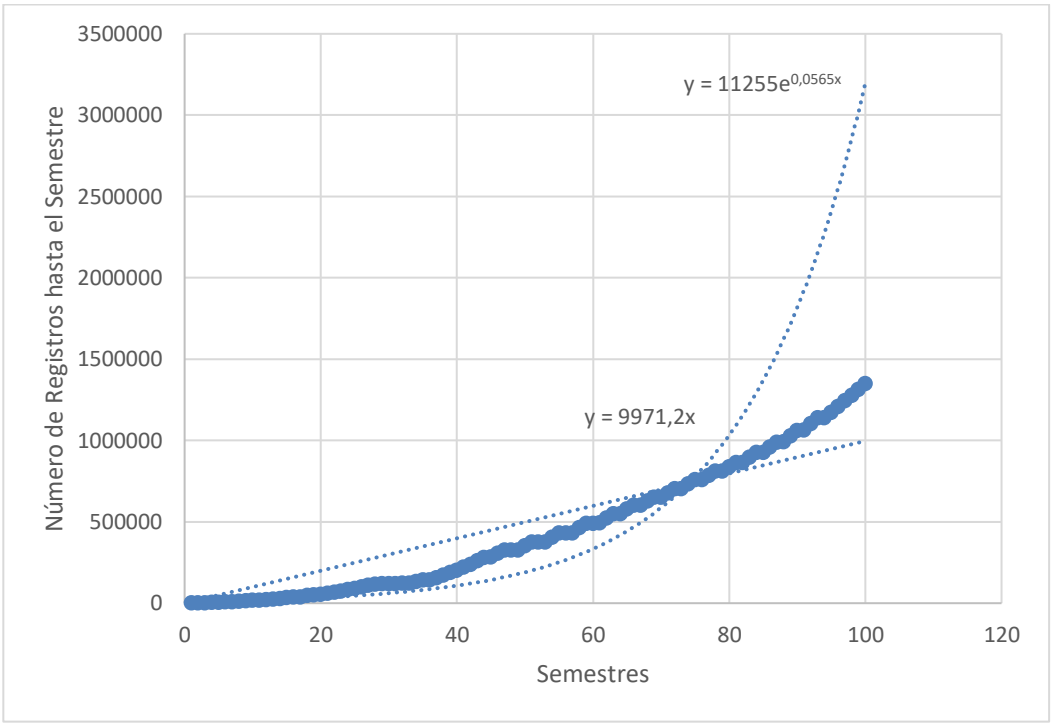
Para el Hecho: Comportamiento Academico Estuadiante

Cuadro Nº 2. Tipos de datos MSSQL

Atributos	Tipo de dato	Tamaño
CodigoCurricula	varchar(16)	18 Bytes
CodigoAlumno	varchar(12)	14 Bytes
CodigoDocente	varchar(6)	8 Bytes
CodigoSemestre	varchar(6)	8 Bytes
CodigoAprobacion	varchar(2)	4 Bytes
PromedioFinal	numeric(18,0)	5 Bytes
Promedio Anterior	numeric(18,0)	5 Bytes
NotaAplazados	numeric(18,0)	5 Bytes
Condicion	varchar(100)	102 Bytes
	Tamaño de registro por fila (Bytes):	169 Bytes

Cuadro Nº 3. Estimación del tamaño de la tabla hecho comportamiento académico Estudiante

	Semestre	Número de Registros por Semestre	Número de Registros en el Tiempo		Semestre	Número de Registros por Semestre	Número de Registros en el Tiempo
1	1978-1	13	13	51	1999-2	21697	373172
2	1978-2	736	749	52	1999-3	5	373177
3	1979-1	1562	2311	53	2000-0	2151	375328
4	1979-2	1242	3553	54	2000-1	28936	404264
5	1980-1	2009	5562	55	2000-2	25782	430046
6	1980-2	1249	6811	56	2000-3	18	430064
7	1981-1	2089	8900	57	2001-0	1848	431912
8	1981-2	1672	10572	58	2001-1	30008	461920
9	1982-1	2659	13231	59	2001-2	27965	489885
10	1982-2	2435	15666	60	2001-3	6	489891
11	1983-1	3000	18666	61	2002-0	2619	492510
12	1983-2	2958	21624	62	2002-1	29911	522421
13	1984-1	2846	24470	63	2002-2	25122	547543
14	1984-2	4112	28582	64	2003-0	1486	549029
15	1985-1	4017	32599	65	2003-1	27491	576520
16	1985-2	5879	38478	66	2003-2	24192	600712
17	1986-0	1	38479	67	2004-0	3	600715
18	1986-1	6590	45069	68	2004-1	26833	627548
19	1986-2	5851	50920	69	2004-2	23699	651247
20	1986-3	985	51905	70	2005-0	2	651249
21	1987-1	7123	59028	71	2005-1	26497	677746
22	1987-2	7430	66458	72	2005-2	26189	703935
23	1988-1	7692	74150	73	2006-0	30	703965
24	1988-2	8365	82515	74	2006-1	27477	731442
25	1989-1	7709	90224	75	2006-2	26073	757515
26	1989-2	9282	99506	76	2007-0	21	757536
27	1990-1	9247	108753	77	2007-1	26251	783787
28	1990-2	7185	115938	78	2007-2	26193	809980
29	1991-1	2636	118574	79	2008-0	30	810010
30	1991-2	7	118581	80	2008-1	27710	837720
31	1992-1	743	119324	81	2008-2	26994	864714
32	1992-2	1230	120554	82	2009-0	33	864747
33	1993-0	152	120706	83	2009-1	29891	894638
34	1993-1	10523	131229	84	2009-2	30447	925085
35	1993-2	10416	141645	85	2010-0	45	925130
36	1993-3	115	141760	86	2010-1	31794	956924
37	1994-1	12359	154119	87	2010-2	32445	989369
38	1994-2	15771	169890	88	2011-0	3232	992601
39	1995-1	17037	186927	89	2011-1	33704	1026305
40	1995-2	14854	201781	90	2011-2	34938	1061243
41	1996-1	17436	219217	91	2012-0	2870	1064113
42	1996-2	16309	235526	92	2012-1	37260	1101373
43	1997-1	24360	259886	93	2012-2	35808	1137181
44	1997-2	20429	280315	94	2013-0	33	1137214
45	1998-0	1617	281932	95	2013-1	35700	1172914
46	1998-1	22212	304144	96	2013-2	35245	1208159
47	1998-2	19758	323902	97	2014-1	35328	1243487
48	1998-3	6	323908	98	2014-2	31403	1274890
49	1999-0	2067	325975	99	2015-1	38245	1313135
50	1999-1	25500	351475	100	2015-2	34146	1347281



Cuadro Nº 4. Estimación del tamaño de la tabla hecho comportamiento académico Estudiante

N°	Semestre	Estimación Lineal	Estimación Exponencial
101	2015-1	1119461	3385688,203
102	2015-2	1132738	3582486,797
103	2016-1	1146015	3790724,627
104	2016-2	1159292	4011066,618
105	2017-1	1172569	4244216,342
106	2017-2	1185846	4490918,27
107	2018-1	1199123	4751960,147
108	2018-2	1212400	5028175,504
109	2019-1	1225677	5320446,324
110	2019-2	1238954	5629705,858
111	2020-1	1252231	5956941,602

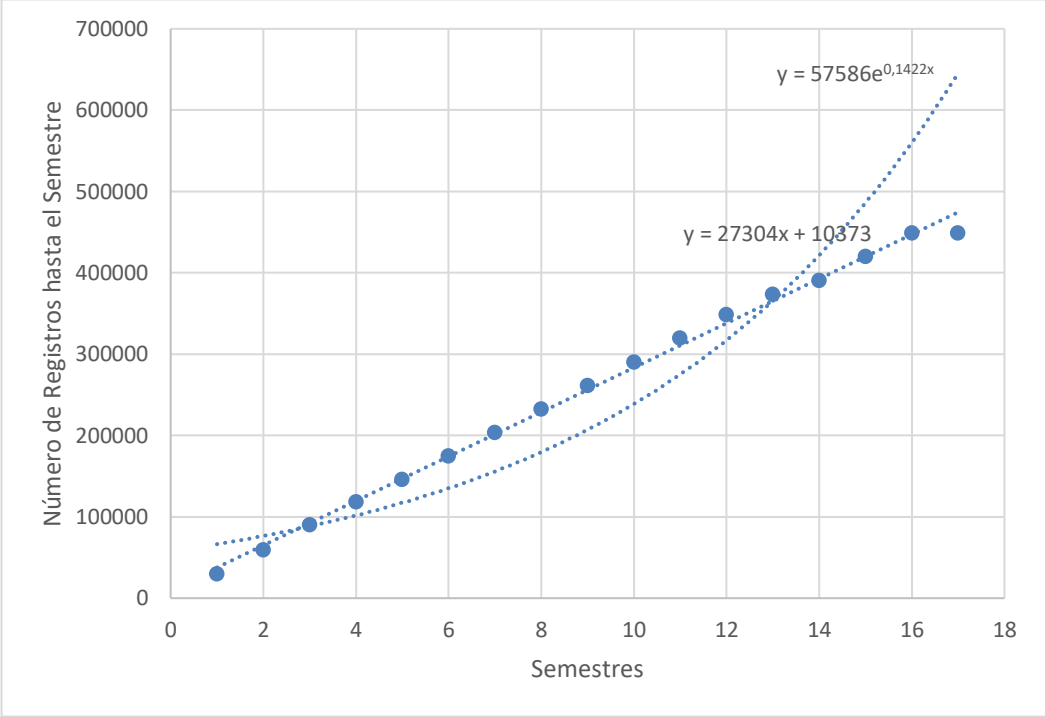
Estimación de Disco	
Tamaño de registro (Bytes):	166
Número de registro estimado	5956941,602
Tamaño en Disco	988852306
Kb	988852,306
Mb	988,852306
Gb	0,988852306

Para el Hecho comportamiento Académico Docente.

Dimensión	**Tipo de dato**	**Tamaño**
CodigoCurricula	varchar(16)	18 Bytes
CodigoDocente	varchar(6)	8 Bytes
CodigoSemestre	varchar(6)	8 Bytes
Seccion	char(1)	1 Bytes
Encuesta	varchar(100)	102 Bytes
Pregunta	varchar(200)	202 Bytes
Valoracion	numeric(18,2)	4 Bytes
Cantidad	int	4 Bytes
	Tamaño de registro por fila (Bytes):	361 Bytes

Cuadro Nº 5. Estimación del tamaño de la tabla hecho comportamiento Docente

	Semestre	**Número de Registros por Semestre**	**Número de Registros en el Tiempo**
1	2007-1	29810	29810
2	2007-2	29678	59488
3	2008-1	30668	90156
4	2008-2	27940	118096
5	2009-1	27920	146016
6	2009-2	28460	174476
7	2010-1	28780	203256
8	2010-2	28960	232216
9	2011-1	28700	260916
10	2011-2	28980	289896
11	2012-1	29960	319856
12	2012-2	28560	348416
13	2013-1	24700	373116
14	2013-2	17070	390186
15	2014-1	29880	420066
16	2014-2	28820	448886
17	2015-1	60	448946



Cuadro Nº 6. Estimación del tamaño de la tabla hecho comportamiento Docente

Nº	Semestre	Estimación Lineal	Estimación Exponencial
18	2015-1	501845	744624,0667
19	2015-2	529149	858407,9757
20	2016-1	556453	989578,8837
21	2016-2	583757	1140793,649
22	2017-1	611061	1315115,117
23	2017-2	638365	1516074,159
24	2018-1	665669	1747741,186
25	2018-2	692973	2014808,601
26	2019-1	720277	2322685,835
27	2019-2	747581	2677608,924
28	2020-1	774885	3086766,811

Estimación de Disco	
Tamaño de registro (Bytes):	361
Número de registro estimado	3086766,811
Tamaño en Disco	1114322819
Kb	1114322,819
Mb	1114,322819
Gb	1,114322819

Anexo Nº13

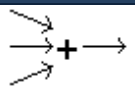
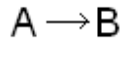
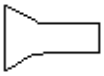
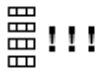
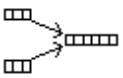
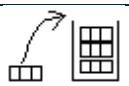
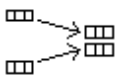
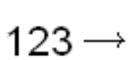
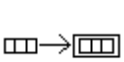
PROCESO ETL

PROCESO DE EXTRACCIÓN, TRANSFORMACIÓN Y CARGA

La metodología DWEP propone un diagrama de proceso ETL, con sus respectivos estereotipos basados en UML, con el fin de representar las operaciones dentro de este proceso tales como: la integración de los datos provenientes de las bases de datos fuentes, la transformación de los atributos, la generación de claves sustitutas, entre otras operaciones.

La metodología DWEP define ciertas operaciones y asigna un estereotipo para cada una ellas, en el Cuadro Nº Nº 01, se muestra las operaciones y/o mecanismos ETL con su respectivo uso y estereotipo.

Cuadro Nº 01. Mecanismo y estereotipos del diagrama de proceso ETL

MECANISMO ETL	DESCRIPCIÓN	ICONO
<<Aggregation>>	Agrega los datos en base a algún criterio.	
<<Conversion>>	Cambia los tipos de datos, el formato o calcula nuevos datos (atributos derivados) a partir de los datos existentes.	
<<Filter>>	Filtra los datos no deseados y verifica la calidad de los datos en base a restricciones.	
<<Incorrect>>	Redirige los registros incorrectos o descartados a un destino separado para su posterior verificación; solo se puede usar con filter, Loades y Wrapper.	
<<Join>>	Une dos fuentes de datos relacionados entre sí a través de uno o varios atributos.	
<<Loader>>	Carga los datos en el destino de un proceso ETL (en un hecho o dimensión del Datawarehouse)	
<<Log>>	Controla y registra la actividad de otro mecanismo ETL, con el fin de auditar las transformaciones realizadas.	
<<Mergue>>	Integra los datos provenientes de dos o más fuentes de datos con atributos compatibles.	
<<Surrogate>>	Genera una clave sustituta única, que se emplea para remplazar la clave empleada en las fuentes de datos	
<<Wrapper>>	Transforma una fuente de datos nativa en una fuente de datos basada en registro	

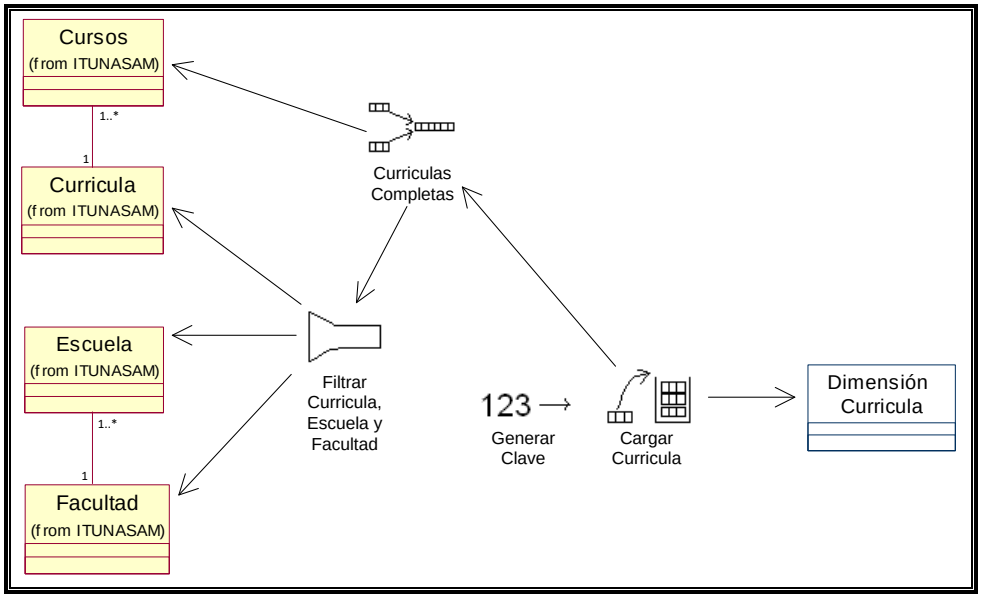
Fuente: Luján Sergio, Data Warehouse Design with UML, 2005

Los diagramas del proceso ETL fueron desarrollados según cada una de las dimensiones y los hechos del modelo multidimensional del Data Mart, ellos se detallan líneas abajo.

1. PROCESO ETL DE LA DIMENSIÓN CURRICULA

El proceso ETL de la dimensión Curricula requiere datos de las tablas Cursos, Curricula, Escuela y Facultad de la Fuente de datos ITUNASAM, con las cuales se procedió a unir los datos relacionados entre sí a través de uno o varios atributos logrando así obtener las curriculas completas que fueron filtradas previamente por la Curricula, Escuela y Facultad; posteriormente se generó la clave y se realizó la carga de datos de la Dimension Curricula, todo este proceso es mostrado en la figura Nº 01.

Figura Nª 01. Proceso ETL de la Dimensión Curricula

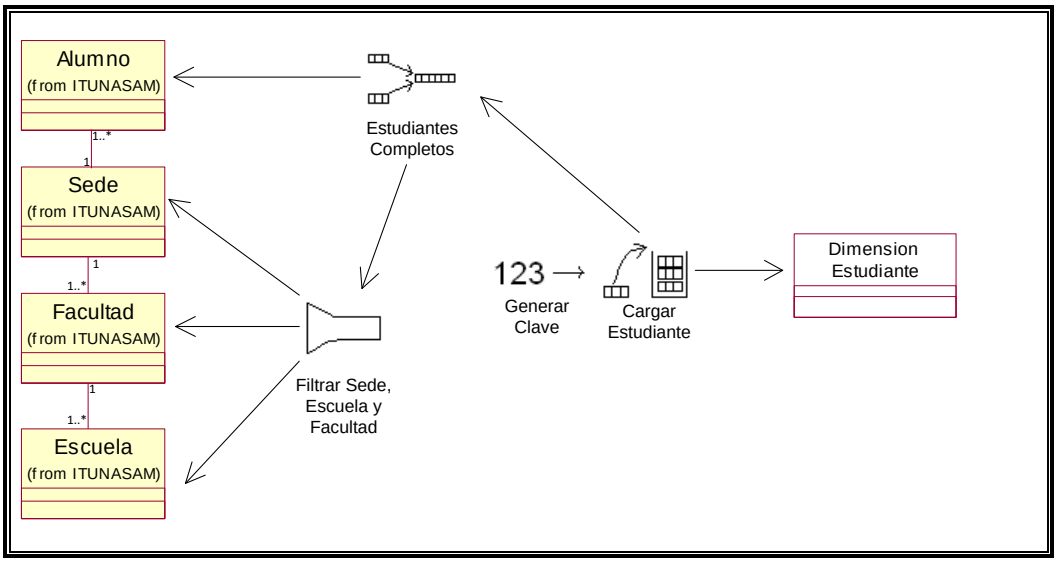


Fuente: Elaboración Propia.

2. PROCESO ETL DE LA DIMENSIÓN ESTUDIANTE

El proceso ETL de la dimensión Estudiante requiere datos de las tablas Alumno, Sede, Escuela y Facultad de la Fuente de datos ITUNASAM, con las cuales se procedió a unir los datos relacionados entre sí a través de uno o varios atributos logrando así obtener a los Estudiantes completos que fueron filtradas previamente por la Sede, Escuela y Facultad a la cual pertenecen; posteriormente se generó la clave y se realizó la carga de datos de la Dimensión Estudiante, todo este proceso es mostrado en la figura Nº 02.

Figura Nª 02. Proceso ETL de la Dimensión Estudiante

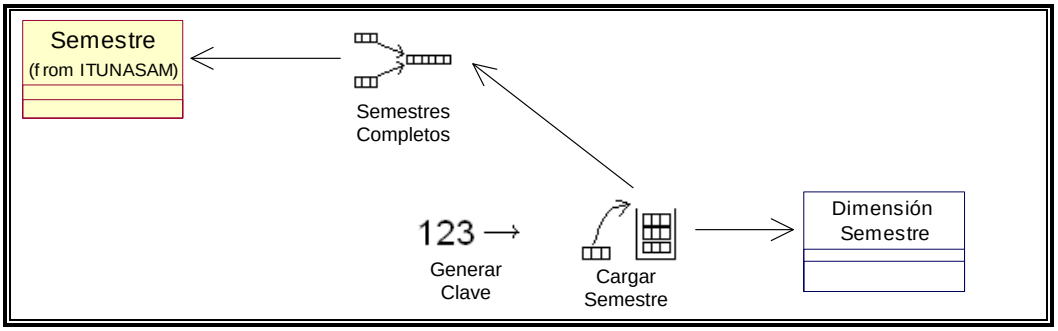


Fuente: Elaboración Propia.

3. PROCESO ETL DE LA DIMENSIÓN SEMESTRE

El proceso ETL de la dimensión Semestre requiere datos de la tabla Semestre de la Fuente de datos ITUNASAM, con las cual se procedió a unir los datos relacionados entre sí a través de los atributos, logrando así obtener a los Semestres completos; posteriormente se generó la clave y se realizó la carga de datos de la Dimensión Semestre, todo este proceso es mostrado en la figura Nº 03.

Figura Nª 03. Proceso ETL de la Dimensión Semestre



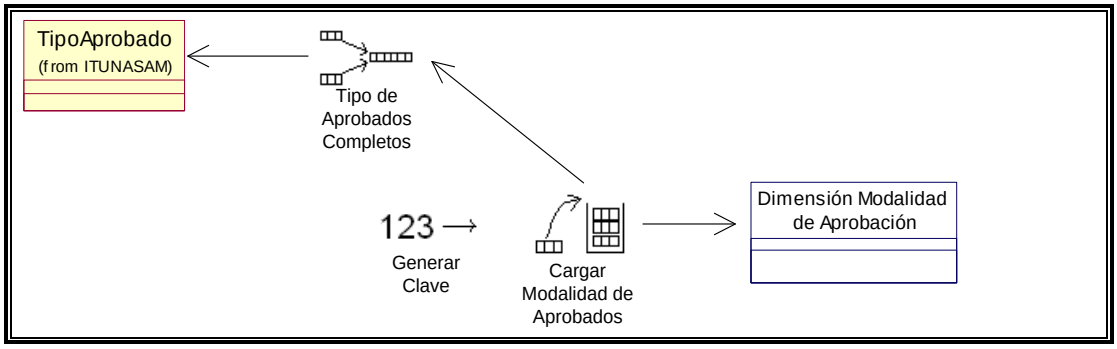
Fuente: Elaboración Propia.

4. PROCESO ETL DE LA DIMENSIÓN MODALIDAD DE APROBACIÓN

El proceso ETL de la dimensión Modalidad de Aprobación requiere datos de la tabla TipoAprobado de la Fuente de datos ITUNASAM, con las cual se procedió a unir

los datos relacionados entre sí a través de los atributos, logrando así obtener a los Tipos de Aprobados completos; posteriormente se generó la clave y se realizó la carga de datos de la Dimensión Modalidad de Aprobación, todo este proceso es mostrado en la figura Nº 04.

Figura Nª 04. Proceso ETL de la Dimensión Modalidad de Aprobación.

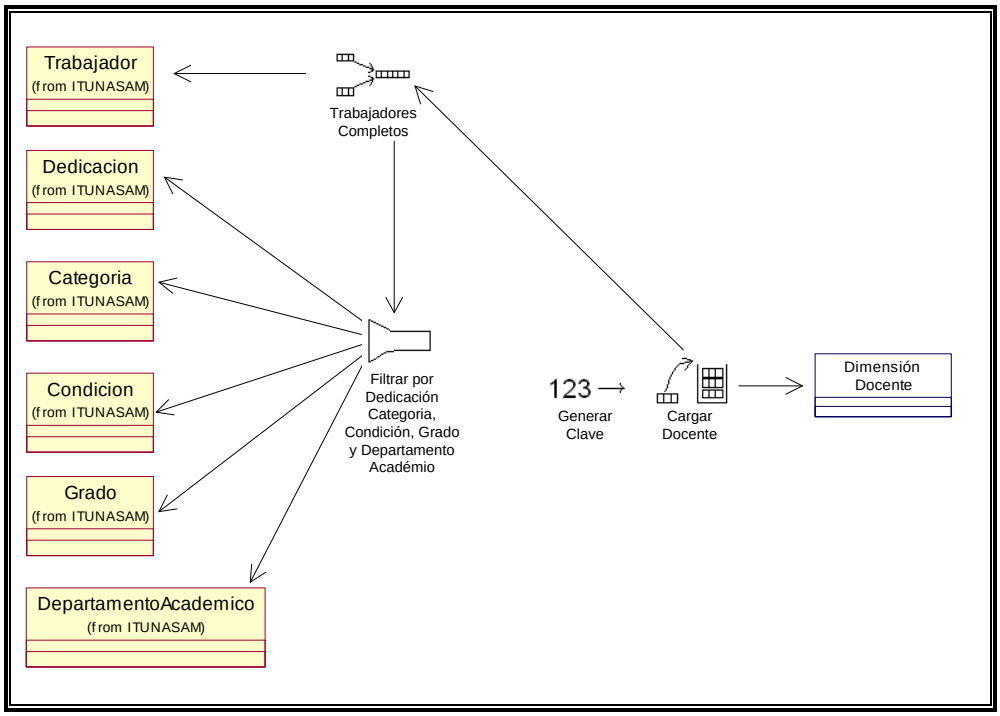


Fuente: Elaboración Propia.

5. PROCESO ETL DE LA DIMENSIÓN DOCENTE

El proceso ETL de la dimensión Docente requiere datos de las tablas Trabajador, Dedicación, Categoría, Condición, Grado y Departamento Académico de la Fuente de datos ITUNASAM, con las cuales se procedió a unir los datos relacionados entre sí a través de uno o varios atributos logrando así obtener a los Trabajadores completos que fueron filtradas previamente por la Dedicación, Categoría, condición, grado y Departamento Académico a la cual pertenecen cada uno de ellos; posteriormente se generó la clave y se realizó la carga de datos de la Dimensión Docente, todo este proceso es mostrado en la figura Nº 05.

Figura Nª 05. Proceso ETL de la Dimensión Modalidad de Aprobación.

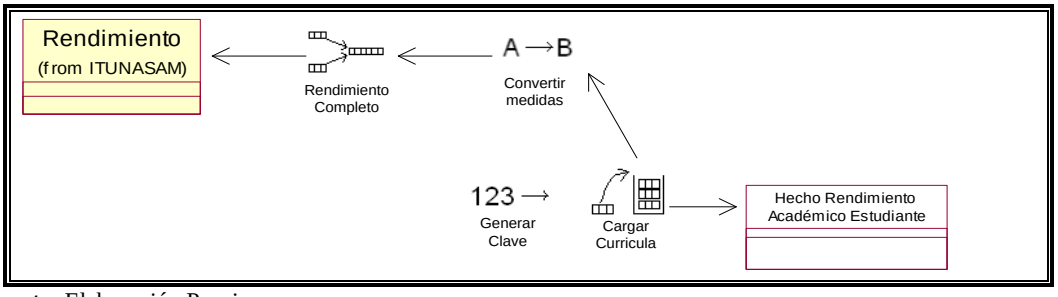


Fuente: Elaboración Propia.

6. PROCESO ETL DEL HECHO COMPORTAMIENTO ACADÉMICO ESTUDIANTE

El proceso ETL del Hecho Comportamiento Académico Estudiante requiere datos de la tabla Rendimiento de la Fuente de datos ITUNASAM, con las cual se procedió a unir los datos relacionados entre sí a través de uno o varios atributos logrando así obtener el rendimiento completo, teniendo en cuenta la conversión de medidas de los atributos; posteriormente se generó la clave y se realizó la carga de datos al hecho Comportamiento Académico Estudiante, todo este proceso es mostrado en la figura Nº 06.

Figura Nª 06. Proceso ETL del Hecho Comportamiento Académico Estudiante

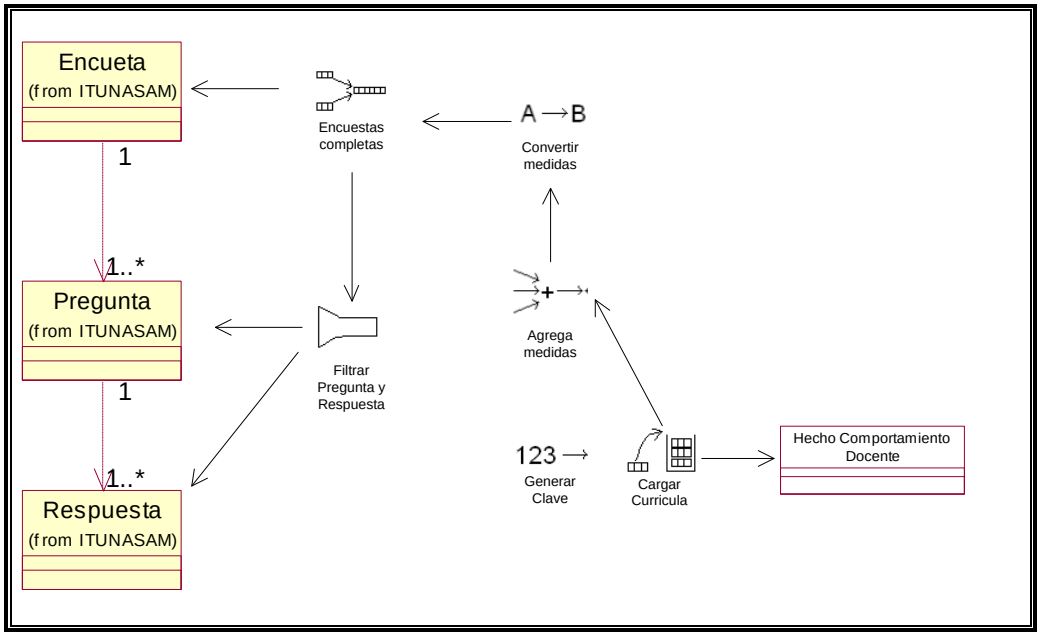


Fuente: Elaboración Propia.

7. PROCESO ETL DEL HECHO COMPORTAMIENTO ACADÉMICO DOCENTE

El proceso ETL del Hecho Comportamiento Docente requiere datos de las tablas Encuesta, Preguntas y Respuestas de las Fuente de datos ITUNASAM, con las cual se procedió a unir los datos relacionados entre sí a través de uno o varios atributos logrando así obtener las encuestas completas, que previamente fueron filtradas por la pregunta y respuesta según cada atributo. Además, se tiene que tener en cuenta la conversión de medidas de los atributos y los atributo agregados; posteriormente se generó la clave y se realizó la carga de datos al hecho Comportamiento Docente, todo este proceso es mostrado en la figura Nº 07.

Figura Nª 07. Proceso ETL del Hecho Comportamiento Docente



Fuente: Elaboración Propia.

8. ESPECIFICACIONES DE LAS TRANSFORMACIONES

8.1. TRANSFORMACIONES DE LA DIMENSIÓN CURRICULA

Representa la extracción, transformación y carga de la dimensión Curricula.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 01. Descripción de la fuente de Datos de la Dimensión Curricula

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	Curricula	Curricula académica.
	Facultad	Facultades de la Universidad
	Escuela	Escuelas académicas de la Universidad.
	Curso	Datos de los Cursos de la Universidad.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 02. Descripción de la fuente de Datos de la Dimensión Curricula

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoCurricula	PK	VARCHAR (16)	Texto	No debe ser nulo	0
Facultad		VARCHAR (100)	Texto	No debe ser nulo	0
Escuela		VARCHAR (100)	Texto	No debe ser nulo	0
Curricula		CHAR (2)	Texto	No debe ser nulo	0
NombreCurso		VARCHAR (100)	Texto	No debe ser nulo	0
Ciclo		CHAR (4)	Texto	No debe ser nulo	0
TipoCurso		VARCHAR (100)	Texto	No debe ser nulo	0
HorasTeorica		INT	Numérico	No debe ser negativo	0
HorasPractica		INT	Numérico	No debe ser negativo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 03. Descripción de la fuente de Datos de la Dimensión Curricula.

Tabla	Curricula, Facultad, Escuela , Curso			
Campos	**Llave**	**Tipo**	**Formato**	**Consideraciones**
Facultad	PK	CHAR (2)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Escuela	PK	CHAR (2)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Curricula	PK	CHAR (2)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Curso	PK	CHAR (2)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Curricula		CHAR (2)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Nombre		VARCHAR (100)	Texto	Tiene espacios en blanco
Ciclo		CHAR (4)	Texto	No tiene espacios en blanco, no tiene ceros a la izquierda
TipoCurso		CHAR (1)	Texto	No tiene espacios en blanco, no tiene ceros a la izquierda
HorasTeorica		NUMERIC (9,0)	Numérico	No considera decimales, está justificado a la derecha, no debe ser negativo
HorasPractica		NUMERIC (9,0)	Numérico	No considera decimales, está justificado a la derecha, no debe ser negativo

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 04. Descripción de la fuente de Datos de la Dimensión Curricula

Tabla	Curricula	
Campo	**Tipo**	**Mapeo**
CodigoCurricula	VARCHAR (16)	Curricula.CodigoCurricula
Facultad	VARCHAR (100)	Curricula.Facultad
Escuela	VARCHAR (100)	Curricula.Escuela
Curricula	CHAR (2)	Curricula.Curricula
NombreCurso	VARCHAR (100)	Curricula.NombreCurso
Ciclo	CHAR (4)	Curricula.Ciclo
TipoCurso	VARCHAR (100)	Curricula.TipoCurso
HorasTeorica	INT	Curricula.HorasTeorica
HorasPractica	INT	Curricula.HorasPractica
CodigoCurricula	VARCHAR (16)	Curricula.CodigoCurricula

Fuente: Elaboración propia.

8.2. TRANSFORMACIONES DE LA DIMENSIÓN ESTUDIANTE

Representa la extracción, transformación y carga de la dimensión Estudiante.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 05. Descripción de la fuente de Datos de la Dimensión Estudiante

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	Alumno	Datos del estudiante.
	Facultad	Facultades de la Universidad
	Escuela	Escuelas académicas de la Universidad.
	Sede	Datos de la sedes de la Universidad.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 06. Descripción de la fuente de Datos de la Dimensión Estudiante

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoAlumno	PK	VARCHAR (14)	Texto	No debe ser nulo	0
Codigo		VARCHAR (12)	Texto	No debe ser nulo	0
NombreCompleto		VARCHAR (150)	Texto	No debe ser nulo	0
Sede		VARCHAR (100)	Texto	No debe ser nulo	0
Facultad		VARCHAR (100)	Texto	No debe ser nulo	0
Escuela		VARCHAR (100)	Texto	No debe ser nulo	0
Genero		VARCHAR (10)	Texto	No debe ser nulo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 07. Descripción de la fuente de Datos de la Dimensión Estudiante

Tabla	Alumno, Sede, Facultad, Escuela			
Campos	**Llave**	**Tipo**	**Formato**	**Consideraciones**
Alumno	PK	CHAR (12)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Escuela		CHAR (2)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
ApellidoPaterno		CHAR (50)	Texto	Tiene espacios en blanco, no tiene ceros a la izquierda
ApellidoMaterno		CHAR (50)	Texto	Tiene espacios en blanco, no tiene ceros a la izquierda
Nombre		CHAR (50)	Texto	Tiene espacios en blanco, no tiene ceros a la izquierda
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 08. Descripción de la fuente de Datos de la Dimensión Estudiante

Tabla	Estudiante	
Campo	**Tipo**	**Mapeo**
CodigoAlumno	VARCHAR (14)	Estudiante.CodigoAlumno
Codigo	VARCHAR (12)	Estudiante.Codigo
NombreCompleto	VARCHAR (150)	Estudiante.NombreCompleto
Sede	VARCHAR (100)	Estudiante.Sede
Facultad	VARCHAR (100)	Estudiante.Facultad
Escuela	VARCHAR (100)	Estudiante.Escuela
Genero	VARCHAR (10)	Estudiante.Genero

Fuente: Elaboración propia.

8.3. TRANSFORMACIONES DE LA DIMENSIÓN SEMESTRE

Representa la extracción, transformación y carga de la dimensión Semestre.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 09. Descripción de la fuente de Datos de la Dimensión Semestre

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	Semestre	Semestres académicos.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 10. Descripción de la fuente de Datos de la Dimensión Semestre

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoSemestre	PK	VARCHAR (6)	Texto	No debe ser nulo	0
Semestre		VARCHAR (100)	Texto	No debe ser nulo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 11. Descripción de la fuente de Datos de la Dimensión Semestre

Tabla	Semestre			
Campos	**llave**	**Tipo**	**Formato**	**Consideraciones**
Semestre	PK	CHAR (6)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Descripción		VARCHAR (100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 12. Descripción de la fuente de Datos de la Dimensión Semestre

Tabla	Semestre	
Campo	**Tipo**	**Mapeo**
CodigoSemestre	VARCHAR (6)	Semestre.CodigoSemestre
Semestre	VARCHAR (100)	Semestre.Semestre

Fuente: Elaboración propia.

8.4. TRANSFORMACIONES DE LA DIMENSIÓN MODALIDAD DE APROBACIÓN

Representa la extracción, transformación y carga de la dimensión Modalidad de Aprobación.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 13. Descripción de la fuente de Datos de la Dimensión Modalidad de Aprobación

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	TipoAprobado	Tipo de aprobación.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 14. Descripción de la fuente de Datos de la Dimensión Modalidad de Aprobación

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoAprobacion	PK	VARCHAR (2)	Texto	No debe ser nulo	0
ModalidadAprobacion		VARCHAR(100)	Texto	No debe ser nulo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 15. Descripción de la fuente de Datos de la Dimensión Modalidad de Aprobación

Tabla	TipoAprobado			
Campos	**llave**	**Tipo**	**Formato**	**Consideraciones**
TipoAprobado	PK	CHAR (1)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Descripcion		VARCHAR (100)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 16. Descripción de la fuente de Datos de la Dimensión Modalidad de Aprobación

Tabla	Curricula	
Campo	**Tipo**	**Mapeo**
CodigoAprobacion	VARCHAR(2)	ModalidadAprobacion.CodigoAprobacion
ModalidadAprobacion	VARCHAR(100)	ModalidadAprobacion.ModalidadAprobacion

Fuente: Elaboración propia.

8.5. TRANSFORMACIONES DE LA DIMENSIÓN DOCENTE

Representa la extracción, transformación y carga de la dimensión Docente.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 17. Descripción de la fuente de Datos de la Dimensión Docente

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	Trabajador	Datos de los trabajadores de la Universidad.
	Dedicacion	Tiempo de dedicación a la Universidad.
	Categoria	Categoria del trabajador de la Universidad.
	Condicion	Condicion del trabajador de la Universidad
	Departamento Academico	Datos de los departamentos académicos.
	Grado	Grado académico de los trabajadores de la Universidad.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 18. Descripción de la fuente de Datos de la Dimensión Docente

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoDocente	PK	VARCHAR (6)	Texto	No debe ser nulo	0
NombreDocente		VARCHAR (150)	Texto	No debe ser nulo	0
Dedicacion		VARCHAR(100)	Texto	No debe ser nulo	0
Categoria		VARCHAR(100)	Texto	No debe ser nulo	0
Condicion		VARCHAR(100)	Texto	No debe ser nulo	0
Departamento Academico		VARCHAR(100)	Texto	No debe ser nulo	0
Grado		VARCHAR(100)	Texto	No debe ser nulo	0
Profesion		VARCHAR(100)	Texto	No debe ser nulo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 19. Descripción de la fuente de Datos de la Dimensión Docente

Tabla	Trabajador, Dedicacion, Categoria, Condicion, Departamento Academico, Grado			
Campos	**Llave**	**Tipo**	**Formato**	**Consideraciones**
Trabajador	PK	CHAR (6)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
ApellidoPaterno		VARCHAR (50)	Texto	Tiene espacios en blanco, no tiene números.
ApellidoMaterno		VARCHAR (50)	Texto	Tiene espacios en blanco, no tiene números.
Nombre		VARCHAR (50)	Texto	Tiene espacios en blanco, no tiene números.
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, no tiene números.
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, no tiene números.
Descripción		VARCHAR (100)	Texto	Tiene espacios en blanco, no tiene números.
Descripcion		VARCHAR (100)	Texto	Tiene espacios en blanco, no tiene números.
Grado		VARCHAR (100)	Texto	Tiene espacios en blanco, no tiene números.
Profesion		VARCHAR (100)	Texto	Tiene espacios en blanco, no tiene números.

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 20. Descripción de la fuente de Datos de la Dimensión Docente

Tabla	Docente	
Campo	**Tipo**	**Mapeo**
CodigoDocente	VARCHAR (6)	Docente.CodigoDocente
NombreDocente	VARCHAR (150)	Docente.NombreDocente
Dedicacion	VARCHAR(100)	Docente.Dedicacion
Categoria	VARCHAR(100)	Docente.Categoria
Condicion	VARCHAR(100)	Docente.Condicion
Departamento Academico	VARCHAR(100)	Docente.Departamento Academico

Grado	VARCHAR(100)	Docente.Grado
Profesion	VARCHAR(100)	Docente.Profesion
CodigoDocente	VARCHAR (6)	Docente.CodigoDocente
NombreDocente	VARCHAR (150)	Docente.NombreDocente

Fuente: Elaboración propia.

8.6. TRANSFORMACIONES DEL HECHO COMPORTAMIENTO ACADÉMICO ESTUDIANTE

Representa la extracción, transformación y carga del hecho Comportamiento Académico Estudiante.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 21. Descripción de la fuente de Datos del hecho Comportamiento Académico Estudiante

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	Rendimiento	Almacena el rendimiento académico de los estudiantes de la Universidad.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 22. Descripción de la fuente de Datos del hecho Comportamiento Académico Estudiante

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoCurricula	PK	VARCHAR (16)	Texto	No debe ser nulo	0
CodigoAlumno		VARCHAR (12)	Texto	No debe ser nulo	0
CodigoDocente		VARCHAR (6)	Texto	No debe ser nulo	0
CodigoSemestre		VARCHAR (6)	Texto	No debe ser nulo	0
CodigoAprobacion		VARCHAR (2)	Texto	No debe ser nulo	0
PromedioFinal		NUMERIC (18,0)	Numérico	No debe ser negativo	0
Promedio Anterior		NUMERIC (18,0)	Numérico	No debe ser negativo	0
NotaAplazados		NUMERIC (18,0)	Numérico	No debe ser negativo	0
Condicion		VARCHAR (100)	Texto	No debe ser nulo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 23. Descripción de la fuente de Datos del hecho Comportamiento Académico Estudiante

Tabla	Rendimiento			
Campos	**Llave**	**Tipo**	**Formato**	**Consideraciones**
PromedioFinal	-	NUMERIC(18,0)	Numérico	No considera decimales, está justificado a la derecha, no debe ser negativo
Promedio Anterior	-	NUMERIC(18,0)	Numérico	No considera decimales, está justificado a la derecha, no debe ser negativo
NotaAplazados	-	NUMERIC(18,0)	Numérico	No considera decimales, está justificado a la derecha, no debe ser negativo

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 24. Descripción de la fuente de Datos del hecho Comportamiento Académico Estudiante

Tabla	ComAcadEstudiante	
Campo	**Tipo**	**Mapeo**
CodigoCurricula	VARCHAR (16)	ComAcadEstudiante.CodigoCurricula
CodigoAlumno	VARCHAR (12)	ComAcadEstudiante.CodigoAlumno
CodigoDocente	VARCHAR (6)	ComAcadEstudiante.CodigoDocente
CodigoSemestre	VARCHAR (6)	ComAcadEstudiante.CodigoSemestre
CodigoAprobacion	VARCHAR (2)	ComAcadEstudiante.CodigoAprobacion
PromedioFinal	NUMERIC (18,0)	ComAcadEstudiante.PromedioFinal
Promedio Anterior	NUMERIC (18,0)	ComAcadEstudiante.Promedio Anterior
NotaAplazados	NUMERIC (18,0)	ComAcadEstudiante.NotaAplazados
Condicion	VARCHAR (100)	ComAcadEstudiante.Condicion
CodigoCurricula	VARCHAR (16)	ComAcadEstudiante.CodigoCurricula

Fuente: Elaboración propia.

8.7. TRANSFORMACIONES DEL HECHO COMPORTAMIENTO DOCENTE

Representa la extracción, transformación y carga del hecho Comportamiento Docente.

A. DESCRIPCIÓN DE LA FUENTE DE DATOS

Cuadro Nº 25. Descripción de la fuente de Datos del hecho Comportamiento Docente

Tipo de fuente	Nombre de la tabla	Descripción
ITUNASAM (MSSQL)	Encuesta	Encuestas académicas de los docentes realizados por los estudiantes de la Universidad.
	Pregunta	Preguntas de las encuestas.
	Respuesta	Respuestas de los estudiantes.

Fuente: Elaboración propia.

B. ESTANDARIZACIÓN DE DATOS

Cuadro Nº 26. Descripción de la fuente de Datos del hecho Comportamiento Docente

Nombre	Llave	Tipo	Formato	Limpieza	Valor por defecto
CodigoCurricula	PK	VARCHAR(16)	Texto	No debe ser nulo	0
CodigoDocente		VARCHAR (6)	Texto	No debe ser nulo	0
CodigoSemestre		VARCHAR (6)	Texto	No debe ser nulo	0
Seccion		CHAR (1)	Texto	No debe ser nulo	0
Encuesta		VARCHAR (100)	Texto	No debe ser nulo	0
Pregunta		VARCHAR (200)	Texto	No debe ser nulo	0
Valoracion		NUMERIC (18,2)	Numérico	No debe ser negativo, se considera dos cifras decimales	0
Cantidad		INT	Numérico	No debe ser negativo	0

Fuente: Elaboración propia.

C. FUENTE DE DATOS

Cuadro Nº 27. Descripción de la fuente de Datos del hecho Comportamiento Docente

Tabla	Respuesta, Encuesta, Pregunta			
Campos	**llave**	**Tipo**	**Formato**	**Consideraciones**
Seccion	PK	CHAR(1)	Texto	No tiene espacios en blanco, tiene ceros a la izquierda
Descripcion (Encuesta)	PK	VARCHAR(100)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Descripcion (Pregunta)	PK	VARCHAR(200)	Texto	Tiene espacios en blanco, tiene ceros a la izquierda
Valoración		CHAR(1)	Texto	No considera decimales, está justificado a la derecha, no debe ser negativo

Fuente: Elaboración propia.

D. TABLA DE DESTINO

Cuadro Nº 04. Descripción de la fuente de Datos del hecho Comportamiento Docente

Tabla	CompDoc	
Campo	**Tipo**	**Mapeo**
CodigoCurricula	VARCHAR(16)	CompDoc.CodigoCurricula
CodigoDocente	VARCHAR (6)	CompDoc.CodigoDocente
CodigoSemestre	VARCHAR (6)	CompDoc.CodigoSemestre
Seccion	CHAR (1)	CompDoc.Seccion
Encuesta	VARCHAR (100)	CompDoc.Encuesta
Pregunta	VARCHAR (200)	CompDoc.Pregunta
Valoracion	NUMERIC (18,2)	CompDoc.Valoracion
Cantidad	INT	CompDoc.Cantidad

Fuente: Elaboración propia