

UNIVERSIDAD NACIONAL DE CAJAMARCA



FACULTAD DE INGENIERÍA

ESCUELA ACADÉMICO PROFESIONAL
DE INGENIERÍA DE SISTEMAS

**“IMPLEMENTACION DE UN DATA MART PARA EL
SEGUIMIENTO ACADÉMICO DE LOS ESTUDIANTES EN LA
ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA DE
SISTEMAS DE LA UNIVERSIDAD NACIONAL DE CAJAMARCA”**

TESIS PARA OPTAR EL TÍTULO DE INGENIERO DE SISTEMAS

BACHILLER: DURÁN RAFAEL, LILIANA ELIZABETH

ASESOR: Ing. JAIME MEZA HUAMÁN

CAJAMARCA - PERÚ

2017

AGRADECIMIENTO

A Dios por darme la vida, y las bendiciones que necesito para ser feliz.
Gracias mamá por darme su amor, por esos abrazos y palabras cuando estaba triste, por celebrar mis triunfos y consolar mis tristezas.

Gracias papá por todo el apoyo y cariño que me ha brindado a lo largo de mi vida.

A mi asesor de tesis, Ingeniero Jaime Meza y al Ingeniero Manuel Malpica por su ayuda, apoyo y asesoramiento, además de sus enseñanzas en las aulas de nuestra alma mater.

DEDICATORIA

A mis padres por apoyarme en todo momento y confiar en mí, siempre con
demasiada paciencia

A Dios por bendecirme con salud para realizar el proyecto,

A mis amigos Anthony, Vanessa y quienes con su amistad y consejos me
ayudaron a no rendirme en el desarrollo de mi proyecto

Y a todos mi tíos que me apoyaron para escribir y concluir esta tesis.

Para ellos es esta dedicatoria de tesis, por su apoyo incondicional.

Contenido

CAPÍTULO I	1
CAPÍTULO II MARCO TEÓRICO	5
2.1. Antecedentes Teóricos de la Investigación	5
2.2. Bases Teóricas	6
2.2.1. Data Mart	6
2.2.1.1. Características del Data Mart	7
2.2.1.2. Clases de Data Mart	7
2.2.1.3. Beneficios del Data Mart	8
2.2.1.4. Tipos de Data Mart	8
2.2.1.5. Diferencias entre Data Mart y Data Warehouse [1]	9
2.2.2. Business Intelligence (BI)	10
2.2.2.1. Niveles de realización de BI (Herramientas de Explotación) [13] ..	11
2.2.2.2. Beneficios de Business Intelligence	11
2.2.2.3. Arquitectura de Business Intelligence	13
2.2.3. Base de Datos Multidimensional	14
2.2.4. Medidas o métricas	15
2.2.5. Dimensiones	15
2.2.6. Fact Table	16
2.2.7. Metodologías para el desarrollo de Data Mart	16
2.2.7.1. Metodología de Ralph Kimball	16
2.2.7.2. Metodología de Bill Inmon	20
2.2.7.3. Cuadro comparativo de Metodologías	22
2.2.8. Microsoft SQL Server Analysis Services (SSAS)	23
2.2.9. Microsoft SQL Server Integration Services (SSIS)	23
2.2.10. Power BI	24
2.3. Definición de Términos Básicos	25
2.3.1. Cubo	25
2.3.2. Datos	26
2.3.3. Información	26
2.3.4. Proceso Educativo	26
2.3.5. Rendimiento Académico	27
2.3.6. Seguimiento Académico	28
CAPÍTULO III MATERIALES Y MÉTODOS	29
3.1. Procedimiento	29
3.1.1. Planeación de Proyecto	29

3.1.1.1.	Objetivos.....	29
3.1.1.2.	Alcance	29
3.1.1.3.	Beneficios.....	30
3.1.1.4.	Entrevistas	30
3.1.1.5.	Recurso Humano.....	31
3.1.2.	Definición de requerimientos del negocio.....	31
3.1.2.1.	Requerimientos del negocio	31
3.1.2.2.	Base de datos transaccional en SQL server	35
3.1.2.3.	Diccionario de datos a nivel general de la base de datos transaccional	40
3.1.2.4.	Datos de la base de datos transaccional	41
3.1.3.	Modelado dimensional.....	46
3.1.3.1.	Elección de la dimensiones.....	46
3.1.3.2.	Dimensiones encontradas	46
3.1.3.3.	Medidas encontradas	46
3.1.3.4.	Dimensión ESTUDIANTES_DIM.....	48
3.1.3.5.	Dimensión CURSOS_DIM.....	49
3.1.3.6.	Dimensión MATRICULA_DIM	49
3.1.3.7.	Dimensión ASIGNACION DOCENTE_DIM.....	50
3.1.3.8.	Dimensión NOTA_CA_DIM.....	51
3.1.3.9.	Dimensión FACULTADS_DIM	51
3.1.3.10.	Dimensión ESCUELAS_DIM	51
3.1.3.11.	Dimensión ESPECIALIDADES_DIM	52
3.1.4.	Diseño e implementación del subsistema de ETL	52
3.1.4.1.	Flujo de datos de ESTUDIANTES_DIM	53
3.1.4.2.	Flujo de datos de CURSOS_DIM	55
3.1.4.3.	Flujo de datos de FACULTAD_DIM.....	57
3.1.4.4.	Flujo de datos de ESCUELA_DIM	59
3.1.4.5.	Flujo de datos de ESPECIALIDAD_DIM.....	61
3.1.4.6.	Flujo de datos de MATRICULA_DIM.....	63
3.1.4.7.	Flujo de datos de NOTACAB_DIM	65
3.1.4.8.	Flujo de datos de ASIGNACION_DOCENTE_DIM	67
3.1.4.9.	Flujo de datos de NOTAS_FACT	69
3.1.5.	Selección del producto.....	72
3.1.6.	Implementación	73
3.1.6.1.	Desarrollo de cubo.....	73

3.1.7.	Especificación de Aplicaciones de BI	78
3.1.8.	Diseño de la Arquitectura Técnica	82
3.1.8.1.	Capa STAGE	82
3.1.8.2.	Capa ODS (Operational Data Store)	83
3.1.8.3.	Capa DW	84
3.2.	Tratamiento, análisis de datos y presentación de resultados	85
CAPÍTULO IV ANÁLISIS Y DISCUSIÓN DE RESULTADOS		87
4.1.	Contrastación de Hipótesis	87
4.2.	Formulación de la Hipótesis Estadística	90
4.3.	Nivel de significancia	91
4.4.	Valor estadístico del procedimiento	91
4.5.	Establecer región crítica	91
4.6.	Tiempo de respuesta en el acceso a la información	92
4.7.	Toma de decisión	95
4.8.	Discusión de resultados	95
CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES		97
5.1.	Conclusiones	97
5.2.	Recomendaciones	97
Bibliografía.....		99
ANEXO 1		102
ANEXO 2		104
ANEXO 3		109
ANEXO 4		111
ANEXO 5		114
ANEXO 7		124

INDICE DE TABLAS

TABLA 1. DIFERENCIAS ENTRE DATA MART Y DATA WAREHOUSE	9
TABLA 2. KIMBALL VS INMON	22
TABLA 3. BENEFICIOS DE DATA MART	30
TABLA 4. RECURSO HUMANO	31
TABLA 5. CUADRO DE REQUERIMIENTO NRO. 1	31
TABLA 6. CUADRO DE REQUERIMIENTO NRO. 2	31
TABLA 7. CUADRO DE REQUERIMIENTO NRO. 3	32
TABLA 8. CUADRO DE REQUERIMIENTO NRO. 4	32
TABLA 9. CUADRO DE REQUERIMIENTO NRO. 5	32
TABLA 10. CUADRO DE REQUERIMIENTO NRO. 6	32
TABLA 11. CUADRO DE REQUERIMIENTO NRO. 7	33
TABLA 12. CUADRO DE REQUERIMIENTO NRO. 8	33
TABLA 13. CUADRO DE REQUERIMIENTO NRO. 9	33
TABLA 14. CUADRO DE REQUERIMIENTO NRO. 10	34
TABLA 15. CUADRO DE REQUERIMIENTO NRO. 11	34
TABLA 16. CUADRO DE REQUERIMIENTO NRO. 12	34
TABLA 17. CUADRO DE REQUERIMIENTO NRO. 13	35
TABLA 18. BASE DE DATOS TRANSACCIONAL	40
TABLA 19. DESCRIPCIÓN DE LA TABLA PERSONA	41
TABLA 20. DESCRIPCIÓN DE LA TABLA NOTAS	41
TABLA 21. DESCRIPCIÓN DE LA TABLA TRABAJADORUNC	42
TABLA 22. DESCRIPCIÓN DE LA TABLA NOTA_CAB	42
TABLA 23. DESCRIPCIÓN DE LA TABLA FACULTAD	42
TABLA 24. DESCRIPCIÓN DE LA TABLA ESCUELA	42
TABLA 25. DESCRIPCIÓN DE LA TABLA ESPECIALIDAD	43
TABLA 26. DESCRIPCIÓN DE LA TABLA MATRICULA	43
TABLA 27. DESCRIPCIÓN DE LA TABLA ESTUDIANTE	44
TABLA 28. DESCRIPCIÓN DE LA TABLA CURSO	45
TABLA 29. DESCRIPCIÓN DE LA TABLA ASIGNACION_DOCENTE	45
TABLA 30. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN ESTUDIANTES_DIM	48
TABLA 31. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN CURSOS_DIM	49
TABLA 32. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN MATRICULA_DIM	49
TABLA 33. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN ASIGNACION DOCENTE_DIM	50
TABLA 34. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN NOTA_CA_DIM	51
TABLA 35. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN MATRICULA_DIM	51
TABLA 36. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN ESCUELAS_DIM	51
TABLA 37. TABLA DE DISEÑO FÍSICO DE DIMENSIÓN ESPECIALIDADES_DIM	52
TABLA 38. RESULTADO DE ENTREVISTA	85
TABLA 39. VALORES PARA NIVEL DE SATISFACCIÓN	87
TABLA 40. PRE- TEST	88
TABLA 41. POST - TEST	89
TABLA 42. RESULTADOS PRE-TEST Y POST-TEST	90
TABLA 43. TABLA DE TIEMPO DE RESPUESTA	92

INDICE DE FIGURAS

FIGURA 1. DATA MART [9]	7
FIGURA 2. ARQUITECTURA DE BUSINESS INTELLIGENCE [16]	14
FIGURA 3. ESQUEMA ESTRELLA [17]	15
FIGURA 4. ESQUEMA COPO DE NIEVE [17]	15
FIGURA 5. CICLO DE VIDA DE KIMBALL [18]	17
FIGURA 6. ENFOQUE INMON [21]	22
FIGURA 7. POWER BI [24]	25
FIGURA 8. MALLA CURRICULAR INGENIERÍA DE SISTEMAS [31]	27
FIGURA 9. BASE DE DATOS TRANSACCIONAL EN SQL SERVER	35
FIGURA 10. CASE DE DATOS CON RELACIONES	36
FIGURA 11. TABLA PERSONA	37
FIGURA 12. TABLA NOTAS	37
FIGURA 13. TABLA TRABAJADORUNC	37
FIGURA 14. TABLA NOTA_CAB	37
FIGURA 15. TABLA FACULTAD	37
FIGURA 16. TABLA ESCUELA	38
FIGURA 17. TABLA ESPECIALIDAD	38
FIGURA 18. TABLA MATRICULA	38
FIGURA 19. TABLA ESTUDIANTE	38
FIGURA 20. TABLA CURSO	39
FIGURA 21. TABLA ASIGNACION_DOCENTE	39
FIGURA 22. DISEÑO FÍSICO DE LA DATA MART	47
FIGURA 23. DISEÑO ETL DE SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS	52
FIGURA 24. FLUJO DE DATOS DE "POBLAR ESTUDIANTE_DIM"	53
FIGURA 25. ORIGEN OLE DB DE "POBLAR ESTUDIANTE_DIM"	54
FIGURA 26. DESTINO OLE DB DE "POBLAR ESTUDIANTE_DIM"	54
FIGURA 27. MAPPING DE "POBLAR ESTUDIANTE_DIM"	55
FIGURA 28. FLUJO DE DATOS DE "POBLAR CURSO_DIM"	55
FIGURA 29. ORIGEN OLE DB DE "POBLAR CURSO_DIM"	56
FIGURA 30. DESTINO OLE DB DE "POBLAR CURSO_DIM"	56
FIGURA 31. MAPPING DE "POBLAR CURSO_DIM"	57
FIGURA 32. FLUJO DE DATOS DE "POBLAR FACULTAD_DIM"	57
FIGURA 33. ORIGEN OLE DB DE "POBLAR FACULTAD_DIM"	58
FIGURA 34. ORIGEN OLE DB DE "POBLAR FACULTAD_DIM"	58
FIGURA 35. MAPPING DE "POBLAR FACULTAD_DIM"	59
FIGURA 36. FLUJO DE DATOS DE "POBLAR ESCUELA_DIM"	59
FIGURA 37. ORIGEN OLE DB DE "POBLAR ESCUELA_DIM"	60
FIGURA 38. DESTINO OLE DB DE "POBLAR ESCUELA_DIM"	60
FIGURA 39. MAPPING DE "POBLAR ESCUELA_DIM"	61
FIGURA 40. FLUJO DE DATOS DE "POBLAR ESPECIALIDAD_DIM"	61
FIGURA 41. ORIGEN OLE DB DE "POBLAR ESPECIALIDAD_DIM"	62
FIGURA 42. DESTINO OLE DB DE "POBLAR ESPECIALIDAD_DIM"	62
FIGURA 43. MAPPING DE "POBLAR ESPECIALIDAD_DIM"	63
FIGURA 44. FLUJO DE DATOS DE "POBLAR MATRICULA_DIM"	63
FIGURA 45. ORIGEN OLE DB DE "POBLAR MATRICULA_DIM"	64
FIGURA 46. DESTINO OLE DB DE "POBLAR MATRICULA_DIM"	64
FIGURA 47. MAPPING DE "POBLAR MATRICULA_DIM"	65
FIGURA 48. FLUJO DE DATOS DE "POBLAR NOTA_CA_DIM"	65
FIGURA 49. ORIGEN OLE DB DE "POBLAR NOTA_CA_DIM"	66
FIGURA 50. DESTINO OLE DB DE "POBLAR NOTA_CA_DIM"	66
FIGURA 51. MAPPING DE "POBLAR NOTA_CA_DIM"	67
FIGURA 52. FLUJO DE DATOS DE "POBLAR ASIGNACION_DOCENT_DIM"	67
FIGURA 53. ORIGEN OLE DB DE "POBLAR ASIGNACION_DOCENT_DIM"	68
FIGURA 54. DESTINO OLE DB DE "POBLAR ASIGNACION_DOCENT_DIM"	68
FIGURA 55. MAPPING OLE DB DE "POBLAR ASIGNACION_DOCENT_DIM"	69

FIGURA 56. FLUJO DE DATOS DE “POBLAR NOTA_FACT”	69
FIGURA 57. ORIGEN OLE DB DE “POBLAR NOTA_FACT”	71
FIGURA 58. DESTINO OLE DB DE “POBLAR NOTA_FACT”	71
FIGURA 59. MAPPING DE “POBLAR NOTA_FACT”	72
FIGURA 60. ATRIBUTOS DE DIMENSIÓN “ESTUDIANTES_DIM”	73
FIGURA 61. VISTA DE ORÍGENES DE DATOS DE “ESTUDIANTES_DIM”	73
FIGURA 62. ATRIBUTOS DE DIMENSIÓN “CURSOS_DIM”	74
FIGURA 63. VISTA DE ORÍGENES DE DATOS DE “CURSOS_DIM”	74
FIGURA 64. ATRIBUTOS DE DIMENSIÓN “ESPECIALIDADES_DIM”	74
FIGURA 65. VISTA DE ORÍGENES DE DATOS DE “ESPECIALIDADES_DIM”	75
FIGURA 66. ATRIBUTOS DE DIMENSIÓN “MATRICULA_DIM”	75
FIGURA 67. VISTA DE ORÍGENES DE DATOS DE “MATRICULA_DIM”	75
FIGURA 68. ATRIBUTOS DE DIMENSIÓN “NOTA_CA_DIM”	76
FIGURA 69. VISTA DE ORÍGENES DE DATOS DE “NOTA_CA_DIM”	76
FIGURA 70. ATRIBUTOS DE DIMENSIÓN “ASIGNACION_DOCENT_DIM”	76
FIGURA 71. VISTA DE ORÍGENES DE DATOS DE “ASIGNACION_DOCENT_DIM”	77
FIGURA 72. ATRIBUTOS DE DIMENSIÓN “NOTA_FACT”	77
FIGURA 73. VISTA DE ORÍGENES DE DATOS DE “NOTA_FACT”	77
FIGURA 74. VISTA DE ORÍGENES DE DATOS DE LA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS	78
FIGURA 75. REPORTE NÚMERO DE ESTUDIANTES POR AÑO DE INGRESO	78
FIGURA 76. REPORTE NÚMERO DE VECES DE LLEVAR EL CURSO	79
FIGURA 77. REPORTE NÚMERO DE ALUMNOS APROBADOS POR CURSO	79
FIGURA 78. REPORTE NÚMERO DE ALUMNOS DESAPROBADOS POR CURSO	80
FIGURA 79. ESTADO DE ESTUDIANTE	80
FIGURA 80. NÚMERO DE ALUMNOS POR NOTA Y CURSO	81
FIGURA 81. CONSOLIDADO DE ALUMNOS	81
FIGURA 82. ESTADO DE ALUMNOS POR DOCENTE	82
FIGURA 83. EJECUCIÓN DE RECOJO DE DATOS	83
FIGURA 84. EJECUCIÓN ODS	84
FIGURA 85. EJECUCIÓN DW	85
FIGURA 86. GRÁFICO DE RESULTADOS DE ENTREVISTA	86
FIGURA 87. PRE-TEST VS POST-TEST	90
FIGURA 88. DISTRIBUCION T - STUDENT	92
FIGURA 89. GRÁFICO DEL TIEMPO DE RESPUESTA	93
FIGURA 90. DISTRIBUCIÓN T-STUDENT	94

RESUMEN

El presente proyecto de investigación se desarrolló con el objetivo de realizar el seguimiento académico de los alumnos de la Escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca y de esta manera poder apoyar en la toma de decisiones en cuanto a temas académicos con la implementación de un data mart bajo la metodología de Ralph Kimball. Se analizó el estado actual de la escuela, determinándose que si bien se cuenta con un sistema proporcionado por la Unidad Técnica de Sistemas Informáticos de la Universidad Nacional de Cajamarca, este sistema presenta algunas limitaciones en cuanto a la información disponible, de la misma manera con el acceso a esta, aspecto que le demandan exceso de tiempo al momento de realizar el análisis de la información académica y por ende la toma de decisiones.

Lo primero que se realizó fue la identificación, análisis y entendimiento de la lógica de la base de datos de la Universidad Nacional de Cajamarca, específicamente de todas las tablas involucradas en el rendimiento académico de los alumnos. Se extrajo la información necesaria para el proyecto y a continuación se realizaron los pasos de la metodología escogida, Raph Kimball, con los cuales se pudieron identificar los requerimientos de los miembros del comité de la Escuela Académico Profesional de Ingeniería de Sistemas, se diseñó el modelado dimensional con la ayuda de las herramientas BI de Microsoft, se elaboró el cubo y se construyeron los reportes con ayuda de la herramienta llamada Power BI.

Finalmente se recopiló información a través de encuestas para obtener los resultados pre – test y post – test de presentados los reportes generados con la ayuda del data mart. Se validó la hipótesis con la prueba T – Student, la implementación de un data mart permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela Académico de Ingeniería de Sistemas. Concluimos que se cumplió con el objetivo de realizar el seguimiento académico de los estudiantes para formular estrategias de mejora de manera más acertada y rápida.

Palabras clave: Seguimiento Académico, base de datos multidimensional, data mart, Inteligencia de Negocios, toma de decisiones.

ABSTRACT

The present research project was carried out with the objective of carrying out the academic follow-up of the students of the Professional Academic School of Systems Engineering of the National University of Cajamarca and in this way to be able to support in the decision making regarding academic subjects with The implementation of a datamart under the methodology of Ralph kimball. The current state of the school was analyzed, determining that although there is a system provided by the Technical Unit of Computer Systems of the National University of Cajamarca, this system presents some limitations as to the information available, in the same way with Access to this, an aspect that demands an excessive amount of time when performing the analysis of academic information and therefore decision making.

The first thing that was done was the identification, analysis and understanding of the logic of the database of the National University of Cajamarca, specifically all the tables involved in the academic performance of the students. The necessary information for the project was extracted, followed by the steps of the chosen methodology, Raph Kimball, with which it was possible to identify the requirements of the members of the committee of the Professional Academic School of Systems Engineering, dimensional modeling was designed With the help of Microsoft BI tools, the cube was built and reports were built using the tool called Power BI.

Finally, information was collected through surveys to obtain the results of the before (pre - test) and after (post - test) of the reports generated with the help of the data mart. The hypothesis was validated with the T - Student test, the implementation of a data mart will allow the monitoring and analysis of the academic information of the students of the Academic School of Systems Engineering.

Key words: Academic Tracking, multidimensional database, data mart, Business Intelligence, decisión making.

CAPÍTULO I

INTRODUCCIÓN

Desde el inicio de la era de la información las empresas necesitan explotar su mayor recurso, la información. La explotación eficiente de la información permite una rápida, acertada y oportuna toma de decisiones bajo el manejo de datos confiables. Por esta razón muchas organizaciones han logrado implementar un Data Mart que permita centralizar la información útil y necesaria para ayudar a la toma de decisiones en los siguientes niveles organizacionales: Nivel operativo, Nivel táctico y Nivel estratégico. [1]

A nivel internacional La Escuela Politécnica Nacional de Quito [2], La Universidad Carlos III de Madrid Escuela Politécnica Superior Ingeniería en Informática [3] y la Universidad de San Carlos de Guatemala [4], hacen uso de grandes cantidades de información proveniente de diferentes fuentes para poder llevar un buen funcionamiento estas escuelas apoyándose de que centraliza su información para poder tomar decisiones o identificar posibles falencias en el funcionamiento de estas, de esta manera logran “formar profesionales de alto nivel”. En la mayoría de las Universidades de Perú, no existen herramientas que ayuden a minimizar el tiempo para analizar grandes cantidades de información con mayor velocidad y precisión y de esta manera poder tomar decisiones. A nivel local veremos cómo se lleva el análisis de la información en Universidades de Cajamarca, la Universidad Privada Antonio Guillermo Urrelo (UPAGU) cuenta con un sistema de registros de alumnos y notas según el curso el cual se mantiene actualizado de acuerdo a como vayan obteniendo notas los alumnos, en este sistema también es posible ver la malla curricular y creditaje de los cursos de igual manera pasa en la Universidad Privada del Norte (UPN). En la escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca el problema principal radica en que el proceso de extracción e integración de los datos para la realización de los reportes se realiza manualmente y sistemáticamente, resultando muy complicado, provocando retraso en la atención de requerimientos haciendo que en muchos casos la entrega de la información no sea oportuna; por otro lado la elaboración de algunos reportes requiere mayor conocimiento en el manejo de las herramientas informáticas, por lo que son derivados al Unidad Técnica de Sistemas

Informáticos, lo cual ocasiona pérdida de tiempo, ya que esto demanda trámites documentarios a través de autoridades de las diferentes oficinas involucradas.

Después de describir el contexto y el problema podemos cuestionarnos, ¿La implementación de un Data Mart permitirá realizar el seguimiento académico de los estudiantes de la Escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca? La hipótesis que se plantea es que la implementación de un Data Mart sí permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela Académico de Ingeniería de Sistemas. Además que la hipótesis específica, con el uso del Data Mart se evitarán los procesos de solicitud de información a la Unidad Técnica de Sistemas Informáticos y se dispondrá de manera inmediata a los reportes con la información que se necesita del alumnado, los reportes generados a partir del Data Mart permitirán observar el rendimiento académico de los alumnos y contrastar en el tiempo, con la información mostrada en los reportes se podrá efectuar la toma de decisiones por los docentes del comité directivo para la mejora académica y se obtendrá facilidad de acceso a la información para el control y toma de decisiones de la carrera y mejorar el tiempo de respuesta.

Con los antecedentes anteriormente expuestos, la solución al problema sirve como apoyo a la Escuela Académico Profesional de Ingeniería de Sistemas para cubrir sus necesidades de mejoramiento en el manejo de la información de modo que pueda ser más eficiente la toma de decisiones organizacionales de nivel operativo y estratégico. La importancia del presente trabajo de investigación radica en agilizar y automatizar todo el proceso de gestión de la información con el objetivo de brindar un servicio de calidad con una herramienta que cumpla las expectativas de la Escuela, con ello eliminar las deficiencias de tiempo en el entorno administrativo y técnico con un manejo rápido y eficaz de la información, para ir acorde con el crecimiento de la misma. Además de poder encontrar posibles errores que se estén cometiendo lo cual afecta de manera directa al alumnado, docentes y administrativos de la escuela y tomar decisiones de acuerdo a los resultados obtenidos, de esta manera mejorar para alcanzar la acreditación académica.

Dentro del alcance se creará el Data Mart que contenga la toda la información académica, involucrada con los alumnos de la Escuela Académico Profesional de Ingeniería de Sistemas para la creación de los reportes que solicite el director

de Escuela, se implementará sobre una herramienta de Inteligencia de Negocios la solución para definir y administrar la información relacionada con los estudiantes y docentes de la escuela. Como resultado se obtendrán los modelos dimensionales que permitirán la creación del Cubo, que será la fuente de los reportes que muestren la información para la toma de decisiones estratégicas de la escuela, se entregará toda la documentación referente a este proyecto de tesis para que la escuela pueda continuar cuando crea conveniente con el proceso de implementación.

Las limitaciones que se encuentran en la investigación es la disponibilidad de obtener los datos requeridos para la construcción del Data Mart en la Unidad Técnica de Sistemas Informáticos de la Universidad Nacional de Cajamarca además que la plataforma sobre la cual se va implementar la herramienta de Inteligencia de Negocios es en base a la plataforma actual con la que trabaja la Unidad de Informática el cual es pagado. También es una limitante el hecho de que el desarrollo de la Aplicación se realizó fuera del horario laboral, lo que limitó el trato directo y constante con el usuario.

El objetivo general es implementar un Data Mart usando tecnología BI para la Escuela Académico Profesional de Ingeniería de Sistemas para poder realizar el seguimiento a los estudiantes y formular estrategias de mejora de manera más acertada y rápida. Como objetivos específicos se tiene, definir los requerimientos de parte de la Escuela Académico Profesional de Ingeniería de Sistemas en el seguimiento académico del proceso educativo, diseño y carga de datos a una estructura dimensional que permita una salida y el análisis de datos obtenidos del Data Mart, crear reportes e indicadores para la toma de decisiones de la escuela, determinar la utilidad de obtener la información consolidada del rendimiento académico de los estudiantes usando un Data Mart.

El informe de tesis está estructurado y desarrollado en 5 capítulos que a continuación se detallan:

Capítulo I, en este capítulo se presenta el problema, situación actual de la Escuela Académico Profesional de Ingeniería de Sistemas, la hipótesis, la justificación, los alcances y objetivo principal de la investigación.

Capítulo II, contiene el marco teórico (antecedentes históricos, bases teóricas y definición de términos básicos) relacionado con los requerimientos y la parte funcional de la IMPLEMENTACIÓN DE UN DATA MAR.

Capítulo III, corresponde a los materiales y métodos que se utilizaron para desarrollar el proyecto de investigación, donde se detalla y se explica la forma en que se obtuvieron los datos para luego transformarlos en información.

Capítulo IV, contiene el análisis de datos que se realizó y se explica los resultados obtenidos en la investigación. Además de la generación, análisis y presentación de reportes.

Capítulo V, contiene las conclusiones de acuerdo a los objetivos planteados y también las recomendaciones necesarias para seguir ampliando los conocimientos sobre el tema de investigación que no se incluyeron en el presente informe de tesis.

CAPÍTULO II MARCO TEÓRICO

2.1. Antecedentes Teóricos de la Investigación

Miguel Rodríguez Sanz [3] en su proyecto fin de carrera, análisis y diseño de un data mart para el seguimiento académico de los alumnos de un entorno universitario; este proyecto consiste ofrecer un marco metodológico para el desarrollo de proyectos de Data Warehouse y/o Data Mart, así como una aplicación práctica de una metodología a un caso de estudio concreto mediante la implementación de un data mart que permita responder a las necesidades de negocio del área académica de una universidad que ayudará a extraer la información acerca del paso de los alumnos por su evaluación universitaria, lo cual ha permitido obtener una mejor visión de los sucesos en el ámbito universitario así como dar apoyo a la toma de decisiones con el fin de conocer la respuesta de los diferentes alumnos a los estudios que han comenzado.

Aimacaña Quilumba Doris Eulalia [5], en su trabajo de graduación previo a la obtención del título de Ingeniero Informático, Análisis, Diseño e Implementación de una Data Mart Académico, usando tecnología BI tanto software libre, para la Facultad de Ingeniería, Ciencias Físicas y Matemática, el trabajo consiste en implementar un data mart que permita almacenar sólo la información requerida e ir eliminando aquellos datos que obstaculizan la labor del análisis y entregando la información que se requiera en la forma más apropiada, facilitando así el proceso de gestión de la misma.

A través de esto surge la promesa del manejo y control de la información, aseguran una vista única de los datos, que pueden provenir de diversas fuentes para la toma de decisiones estratégicas tanto de estudiantes como profesores, esto es muy importante para saber dónde y cómo solventar alguna falla, o mejorar las decisiones tomadas.

Tana Pasquel Gloria Estefanía [6], en su trabajo previo a la obtención del título de Ingeniería en Sistemas Computacionales, Data Mart para el Análisis de información del sistema Académico de la Universidad Técnica del Norte con Herramientas de Software Libre, en el cual la implementación de un prototipo de Data Mart para el análisis de información se hizo con la finalidad de demostrar que el uso de herramientas de software libre en la actualidad está en auge en las pequeñas y medianas empresas, el prototipo cumplió

con diferentes etapas que fueron realizadas con las distintas herramientas que ofrece la Suite de Pentaho, la información del datamart está almacenada en una base de datos Postgre SQL que también es de software libre.

A lo largo del desarrollo del trabajo de grado se pudo observar que, la implementación de datamart con herramientas de software libre es muy confiable y segura, obteniendo estabilidad al momento de integrar los datos, logrando buenos resultados.

Montalvo Barrientos, Miguel Ángel [7] , en su trabajo, Implementación de un Data mart para la Facultad Nacional de Ingeniería Caso: Ing. De Sistemas en la Universidad Técnica de Oruro, que desarrolló un producto de Data Mart independiente, soportado por una base de datos multidimensional, se ha conseguido implementar una herramienta de análisis gráfico que permita visualizar el comportamiento de un fenómeno cualquiera asociado con diferentes causas de la variación de tiempos de la permanencia estudiantil o cualquier otro fenómeno en la Carrera de Ing. de Sistemas.

Como conclusión se ha conseguido integrar la información necesaria en un Data Mart el cual es soportado por una herramienta web de análisis OLAP, que a percepción del usuario consigue una mejoría en el análisis de información.

2.2. Bases Teóricas

2.2.1. Data Mart

Un Data Mart es un subconjunto de datos derivados del Data Warehouse. Está diseñado para soportar requerimientos analíticos específicos de una determinada unidad de negocios.

Un Data Mart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un Data Mart puede ser alimentado desde los datos de un Data Warehouse, o integrar por sí mismo un compendio de distintas fuentes de información. [8]

Un Data Mart es un gran almacén de datos de una organización referente a un área, un tema o una función específica de una organización de la cual se pueden hacer consultas rápidas en un nivel más pequeño que un

Data Warehouse, es decir, es un repositorio menos ambicioso que un Data Warehouse.

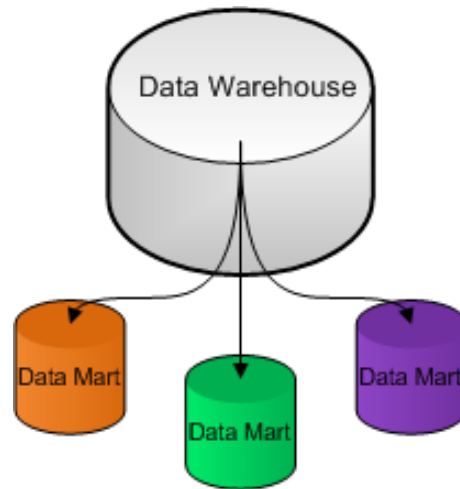


Figura 1. Data Mart [9]

2.2.1.1. Características del Data Mart

- Son poblados por usuarios finales
- Se actualizan constantemente
- Contiene información detallada
- Escalable
- Orientada al tema

2.2.1.2. Clases de Data Mart

Existen dos tipos de Data Mart:

❖ Data Mart Dependientes

Los datos que se utilizan para poblar el Data Mart provienen del Data Warehouse. Esta estrategia es particularmente apropiada cuando el Data Warehouse crece muy rápidamente y los distintos departamentos requieren sólo una pequeña porción de los datos contenidos en él.

❖ Data Mart Independientes

Los datos que se utilizan para poblar el Data Mart provienen de los sistemas operacionales y/o fuentes externas.

Problemática:

- El hecho de tener varios Data Mart independientes, sin ningún tipo de integración, puede hacer que las tareas de administración y mantenimiento se conviertan en una complicación.

- Se puede dar el caso de que algunos Data Mart necesiten los mismos datos para dar respuesta a determinadas preguntas, por lo que tenemos redundancia de datos, ya que cada Data Mart se alimenta con sus procesos de extracción y transformación propios.

2.2.1.3. Beneficios del Data Mart

- Acelera las consultas reduciendo el volumen de datos a recorrer.
- Estructura de datos para su adecuado acceso por una herramienta.
- Segmentar los datos en diferentes plataformas de hardware.
- Los costos que implica construir un Data Mart son mucho menores a un Data Warehouse. [10]
- Consultas rápidas y sencillas.
- Validación directa de la información.

2.2.1.4. Tipos de Data Mart

Data Mart OLAP (On-Line Analytical Processing)

Los sistemas OLAP son bases de datos orientadas al procesamiento analítico. Este análisis suele implicar, generalmente, la lectura de grandes cantidades de datos para llegar a extraer algún tipo de información útil, como por ejemplo: tendencias de ventas, patrones de comportamiento de los consumidores, elaboración de informes complejos, etc.

El acceso a los datos suele ser de sólo lectura. La acción más común es la consulta, con muy pocas inserciones, actualizaciones o eliminaciones.

Los datos se estructuran según las áreas de negocio, y los formatos de los datos están integrados de manera uniforme en toda la organización. El historial de datos es a largo plazo, normalmente de dos a cinco años. [11]

Se basan en los populares cubos OLAP, que se construyen agregando, según los requisitos de cada área o departamento, las dimensiones y los indicadores necesarios de cada cubo relacional. El modo de creación, explotación y mantenimiento de los cubos OLAP es muy heterogéneo, en función de la herramienta final que se utilice. [8]

Data Mart OLTP (On-Line Transactional Processing)

Los sistemas OLTP son bases de datos orientadas al procesamiento de transacciones que pueden involucrar operaciones de inserción, modificación y borrado de datos. El proceso transaccional es típico de las bases de datos operacionales.

El acceso a los datos está optimizado para tareas frecuentes de lectura y escritura, como por ejemplo, la enorme cantidad de transacciones que tienen que soportar diariamente las BD de bancos o hipermercados.

Los datos se estructuran según el nivel de la aplicación (programa de gestión a medida, ERP o CRM implantado, sistema de información departamental) el historial de datos suele limitarse a los datos actuales o recientes. [11]

2.2.1.5. Diferencias entre Data Mart y Data Warehouse [1]

Tabla 1. Diferencias entre data mart y data warehouse

	Data Warehouse	Data Mart
Alcance	Construido para satisfacer las necesidades de información de toda la organización.	Construida para satisfacer las necesidades de un área del negocio específica.
Objetivo	Diseñado para optimizar la integración y la administración de los datos fuente.	Diseñado para optimizar la entrega de información de soporte a decisiones.
Características de los datos	Administra grandes cantidades de datos.	Se concentra en administrar resúmenes y/o datos totalizados.
Pertenencia	Pertenece a toda la organización.	Pertenece al área del negocio al cual está orientado.
Administración	Es administrado por la unidad de sistemas de la organización.	Es administrado por la unidad de sistema de la unidad propietaria del Data Mart.

2.2.2. Business Intelligence (BI)

BI es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un Data Warehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones.

El proceso de Business Intelligence incluye la comunicación de los descubrimientos y efectuar los cambios. Las áreas incluyen clientes, proveedores, productos, servicios y competidores. [12]

Business Intelligence o Inteligencia de Negocios tiene como objetivo apoyar a las organizaciones para mejorar su competitividad, facilitando la información necesaria a quienes la necesitan en el momento oportuno para posibilitar una mejor y más fundamentada toma de decisiones.

Mediante el uso de tecnologías y las metodologías de BI pretendemos convertir datos en información útil y relevante, a partir de esta ser capaces de descubrir conocimiento. Posee las siguientes características:

- **Gran rigidez a la hora de extraer datos**, de manera que el usuario tiene que ceñirse a los informes predefinidos que se configuraron en el momento de la implantación, y que no siempre responden a sus dudas reales.
- **Necesidad de conocimientos técnicos**. Para la generación de nuevos informes o métricas suele resultar ineludible acudir al departamento técnico, solicitando una consulta adecuada para interrogar la base de datos.
- **Largos tiempos de respuesta**, ya que las consultas complejas de datos suelen implicar la unión de tablas operacionales de gran tamaño, lo que se traduce en una incómoda espera que dificulta la fluidez del trabajo.
- **Deterioro en el rendimiento del SI**, cuando la base de datos consultada, para generar informes o ratios de negocio, es la misma que la que soporta el operativo de la empresa, el funcionamiento del sistema puede degradarse hasta afectar y paralizar a todos los usuarios conectados.
- **Falta de integración que implica islas de datos**. Muchas organizaciones disponen de múltiples sistemas de información, incorporados en momentos distintos, para resolver problemáticas

diferentes. Sus bases de datos no suelen estar integradas, lo que implica la existencia de islas de información.

- **Datos erróneos, obsoletos o incompletos.** El tema de la calidad de los datos siempre es considerado como algo importante, pero esta labor nunca se lleva al extremo de garantizar la fiabilidad de la información aportada.
- **Problemas para adecuar la información al cargo del usuario.** No se trata de que todo el mundo tenga acceso a toda la información, sino de que tenga acceso a la información que necesita para que su trabajo sea lo más eficiente posible.
- **Ausencia de información histórica.** Los datos almacenados en los sistemas operacionales están diseñados para llevar la empresa al día, pero no permiten contrastar la situación actual con una situación retrospectiva de años atrás.

2.2.2.1. Niveles de realización de BI (Herramientas de Explotación) [13]

De acuerdo a su nivel de complejidad se pueden clasificar las soluciones de Business Intelligence en:

- Consultas e informes simples (Querys y reports).
- Cubos OLAP (On-Line Analytic Processing). Exploración, tablas dinámicas, etc.
- EIS: Soluciones que permiten visualizar, de una forma rápida y fácil, el estado de una determinada situación empresarial, presente o pasada, y que permite detectar anomalías u oportunidades.
- DSS: Aplicación informática que basándose en modelos matemáticos y mediante análisis de sensibilidad permite ayudar a la toma de decisiones.
- Data Mining o minería de datos: Pueden considerarse sistemas expertos que nos permiten “interrogar a los datos” para obtener información
- KMS: nuevas tecnologías para la gestión del conocimiento y su integración en una única plataforma.

2.2.2.2. Beneficios de Business Intelligence

- **Manejar el crecimiento:** El reto para las empresas es evolucionar,

crecer y esto significa “cambio”.

- **Control de costos:** El manejo de costos es el detonador que fuerza a muchas empresas a considerar una solución de inteligencia de negocios, para tener la capacidad de medir gastos y ver esto a un nivel de detalle que identifique la línea de negocio, producto, centro de costo, entre otras.
- **Entender mejor a los clientes:** Las empresas almacenan toneladas de información valiosa relacionada a sus clientes. El reto es transformar esta información en conocimiento y este conocimiento dirigido a una gestión comercial que represente algún tipo de ganancia para la empresa. Ventas cruzadas, ventas incrementales, retención de clientes, anticipar nuevas oportunidades, medición de efectividad de campañas e identificar los patrones de compras y/o comportamiento de los clientes hacen de la inteligencia de negocios un elemento vital para lograr los objetivos de la empresa.
- **Indicadores de gestión:** Los indicadores de desempeño me permiten representar medidas enfocadas al desempeño organizacional con la capacidad de representar la estrategia organizacional en objetivos, métricas, iniciativas y tareas dirigidas a un grupo y/o individuos en la organización. Dentro de las capacidades funcionales de los indicadores de gestión podemos mencionar: el monitoreo, análisis y la administración.

“Monitorea” lo procesos críticos de negocio y las actividades utilizando métricas que me den una alerta sobre un problema potencial o alguna gestión que debo realizar. “Analiza” la raíz de los problemas explorando la información de múltiples perspectivas en varios niveles de detalles.

“Administra” los recursos y procesos para dirigir la toma de decisiones, optimizar el desempeño. Esto nos permite tener una visión global de la empresa con la capacidad de dirigir la organización en la dirección correcta. [14]

2.2.2.3. Arquitectura de Business Intelligence

- a) Repositorios O Fuentes:** Son todos aquellos datos que reposan en diferentes fuentes o bases de datos. Ya sean tablas de Excel, archivos planos, bases de datos. Estos datos son almacenados por diferentes usuarios que manejan distintos sistemas que pueden encontrarse en diferentes lugares.
- b) Integración:** En esta etapa con la ayuda de una herramienta, se integran todos los datos definidos en el análisis y diseño de la solución BI para la empresa en particular. Aquí llamamos cada una de las fuentes y creamos un almacén de datos (Data Warehouse o Data Mart).
- c) Análisis:** Cuando ya se tiene toda la estructura de un Data Warehouse compuesta por Data Marts, estructurada y poblada. Aquí se llama al Data Warehouse y se crean los cubos. Estos se reestructuran dependiendo de los requerimientos de la organización, se crean jerarquías, se oculta información innecesaria, se hace minería de datos y se hace un análisis completo de los datos, para verificar que los mismos se están cargando consistentemente.
- d) Reportes:** Por último se diseñan los reportes donde se ilustrará toda la información de los cubos de una forma vistosa, agradable para el usuario. Teniendo un pensamiento más de diseñador gráfico, se busca la manera de mostrar la información en gráficos y tablas que se ajusten a las necesidades del cliente final, pensando quizás en colores corporativos, logos y mensajes descriptivos que ayuden de alguna manera a que el manejo de estos sea muy intuitivo.
- e) Usuario:** Es la persona que a través de distintos dispositivos puede tener acceso a todos los reportes BI, y de esta manera tomar decisiones cruciales dentro de la organización. Vale aclarar que la

persona que visualiza el reporte se puede encontrar dentro de una intranet, extranet o internet. [15]

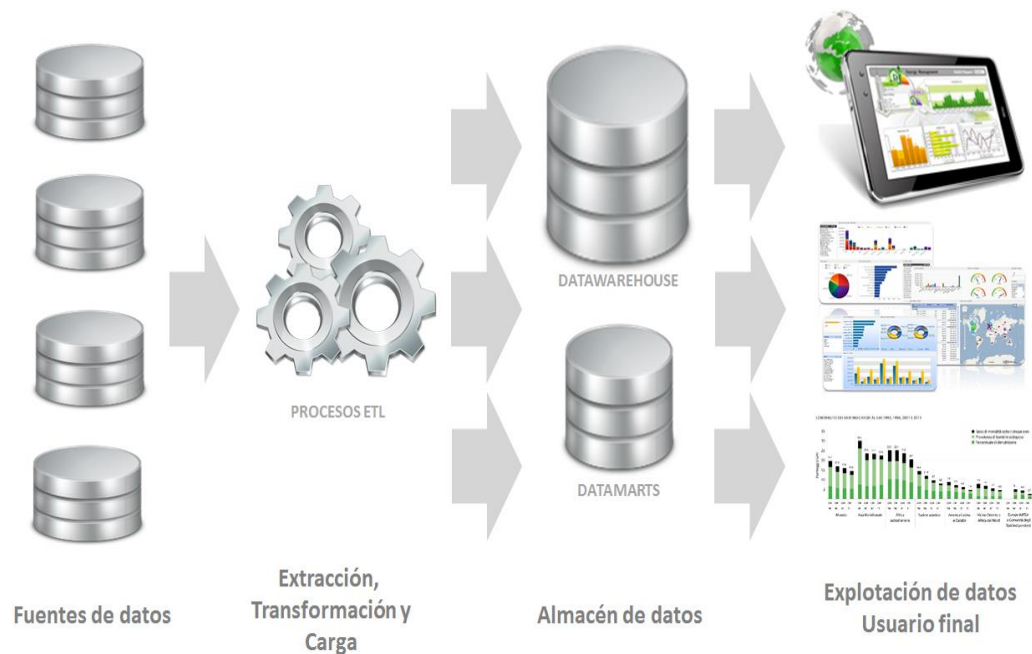


Figura 2. Arquitectura de business intelligence [16]

2.2.3. Base de Datos Multidimensional

Una base de datos multidimensional es aquella base de datos en donde su información se almacena en una estructura multidimensional, conformada por tablas de dimensiones y tablas de hechos. Se ha determinado conceptualmente dos variantes de modelamiento, puntualizando las consultas de soporte de decisión, estas son:

a) Esquema en estrella: el esquema en estrella está formado por una tabla de hechos base y de varias tablas de dimensiones relacionadas a esta, a través de las claves de relación. Es el esquema más simple de interpretar y permite optimizar los tiempos de respuesta ante las consultas de los usuarios. Este esquema, siendo el más eficiente, es soportado por la mayoría de las herramientas de consulta y análisis, y los metadatos son fáciles de documentar y mantener.

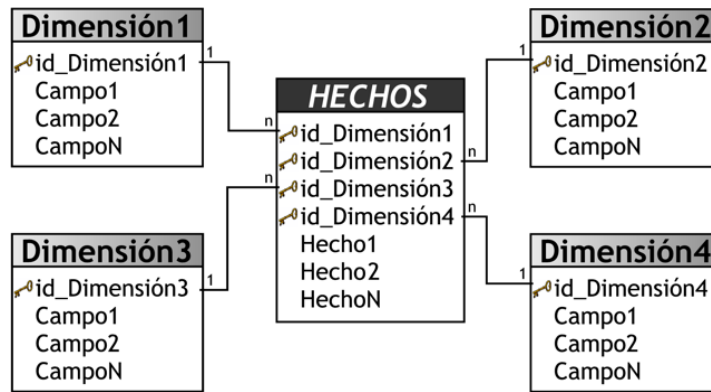


Figura 3. Esquema estrella [17]

b) Esquema copo de nieve: el esquema copo de nieve consiste en una tabla de hechos central relacionada con una o más tablas de dimensiones, las que a su vez pueden estar relacionadas con otras tablas de dimensiones. Este esquema representa el modelo en estrella pero con las dimensiones organizadas en forma de jerarquías.

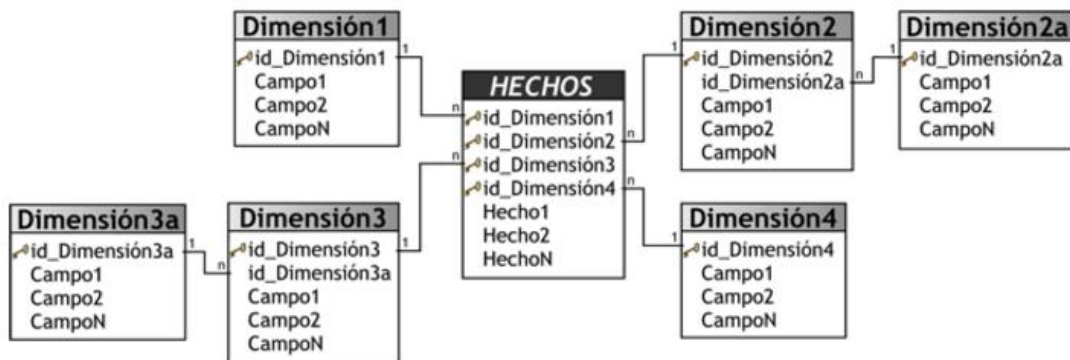


Figura 4. Esquema copo de nieve [17]

2.2.4. Medidas o métricas

Son características cualitativas o cuantitativas de los objetos que se desean analizar en las empresas u organizaciones. Las medidas cuantitativas están dadas por valores o cifras porcentuales. Por ejemplo cantidad de unidades en stock, cantidad de unidades de productos vendidos, etc.

2.2.5. Dimensiones

Son objetos del negocio con los cuales se puede analizar la tendencia y el comportamiento del mismo. Las definiciones de las dimensiones se

basan en políticas de la compañía, e indican la manera en que la organización interpreta o clasifica su información para segmentar el análisis facilitando la observación de los datos.

2.2.6. Fact Table

También llamada tabla de Hechos, es la tabla central del esquema en estrella que representa datos numéricos en el contexto de las entidades del negocio. La tabla de hechos está constituida por medidas y por foreign keys.

La fact table expresa la relación muchos a muchos entre las dimensiones dentro del modelo dimensional.

2.2.7. Metodologías para el desarrollo de Data Mart

Los padres del Data Warehousing, Bill Inmon y Ralph Kimball, proponen una arquitectura de inteligencia de negocios similar pero con diferentes variantes para la implantación de proyectos de inteligencia de negocios. Por un lado, Inmon propone una arquitectura de inteligencia de negocios llamada CIF (Corporate Information Factory o fábrica de información corporativa). Por otro lado, Kimball propone una arquitectura de inteligencia de negocios conocida como MD (Multi Dimensional o multidimensional). Las dos propuestas tienen el mismo propósito pero su principal diferencia radica en la forma de modelar los datos. A continuación un análisis de ambas metodologías.

2.2.7.1. Metodología de Ralph Kimball

Esta metodología es utilizada para la construcción de un Data Warehouse o Data Mart que no es más que, una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza.

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este, está basado en cuatro principios básicos:

- **Centrarse en el negocio:** Hay que concentrarse en la identificación de los requerimientos del negocio.
- **Construir una infraestructura de información adecuada:** Diseñar una base de información integrada, fácil de usar, de alto rendimiento

donde se reflejará la amplia gama de requerimientos de negocio identificados.

- **Realizar entregas en incrementos significativos:** Crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. En esto la metodología se parece a las metodologías ágiles de construcción de software.
- **Ofrecer la solución completa:** Proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

Podemos ver en la siguiente imagen el ciclo de vida mencionado anteriormente.

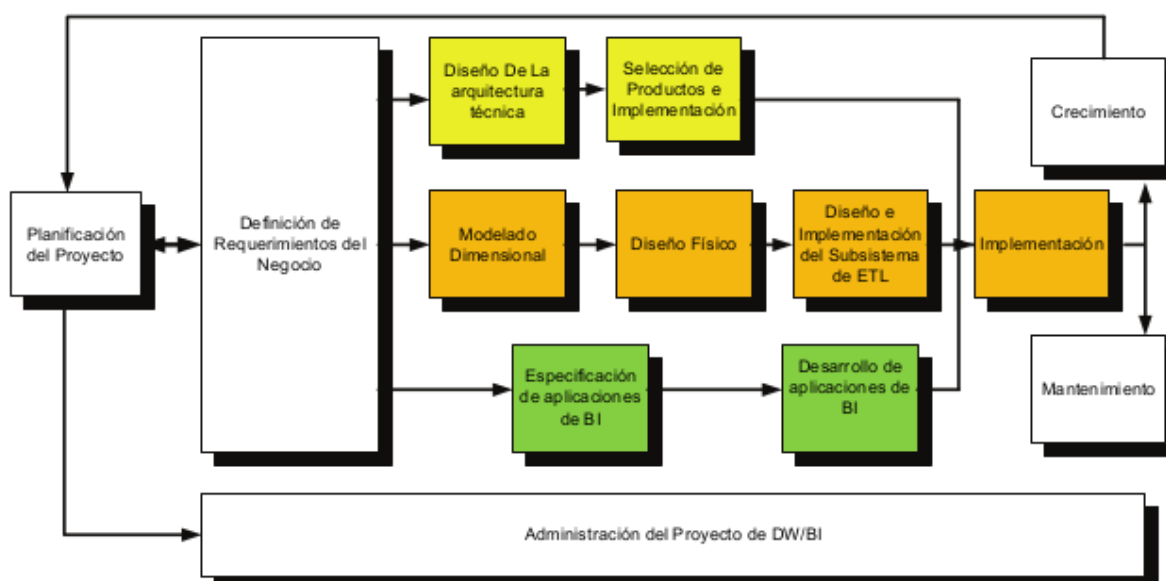


Figura 5. Ciclo de Vida de Kimball [18]

Como se puede apreciar en la figura, los Requerimientos del Negocio son el soporte inicial de las tareas subsiguientes. También tiene influencia en el plan de proyecto (puede notar la doble flecha entre la caja de definición de requerimientos y la de planificación).

Podemos también ver tres rutas o caminos que se enfocan en tres diferentes áreas [19]:

- **Tecnología (Camino Superior):** Implica tareas relacionadas con software específico, por ejemplo, Microsoft SQL Analysis Services.
- **Datos (Camino del medio):** En la misma diseñaremos e implementaremos el modelo dimensional, y desarrollaremos el subsistema de Extracción, Transformación y Carga (Extract, Transformation, and Load - ETL) para cargar el DW.
- **Aplicaciones de Inteligencia de Negocios (Camino Inferior):** En esta ruta se encuentran tareas en las que diseñamos y desarrollamos las aplicaciones de negocios para los usuarios finales.

Estas rutas se combinan cuando se instala finalmente el sistema. En la parte de debajo de la figura se muestra la actividad general de administración del proyecto. A continuación describiremos cada una de las tareas:

❖ **Planificación:** En este proceso se determina el propósito del proyecto de DW, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.

En la visión de programas y proyectos de Kimball, Proyecto, se refiere a una iteración simple del Ciclo de Vida de Kimball, desde el lanzamiento hasta el despliegue.

Esta tarea incluye las siguientes acciones típicas de un plan de proyecto:

- ✓ Definir el alcance (Entender los Requerimientos del Negocio)
- ✓ Identificar las tareas
- ✓ Programar las tareas
- ✓ Planificar el uso de los recursos
- ✓ Asignar la carga de trabajo a los recursos
- ✓ Elaboración de un documento final que representa un plan del proyecto

❖ **Análisis de requerimientos:** La definición de los requerimientos es en gran medida un proceso de entrevistar al personal de negocio y técnico, pero siempre conviene tener un poco de preparación previa. Se debe aprender tanto como se pueda sobre el negocio, los competidores, la industria y los clientes del mismo. Hay que leer todos los informes posibles de la organización; rastrear los documentos de estrategia interna; entrevistar a los empleados, analizar lo que se dice en la prensa acerca de la organización, la competencia y la industria. Se deben conocer los términos y la terminología del negocio.

❖ **Modelado dimensional:** El proceso de diseño comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados de la matriz descrita en el punto anterior. El proceso iterativo consiste en cuatro pasos:

- ✓ Elegir el Proceso de Negocio.
- ✓ Establecer el Nivel de Granularidad.
- ✓ Elegir las Dimensiones.
- ✓ Identificar medidas y las tablas de hechos.

❖ **Diseño físico:** En esta parte, intentamos contestar las siguientes preguntas:

- ✓ ¿Cómo puede determinar cuán grande será el sistema de DW?
- ✓ ¿Cuáles son los factores de uso que llevarán a una configuración más grande y más compleja?
- ✓ ¿Cómo se debe configurar el sistema?
- ✓ ¿Cuánta memoria y servidores se necesitan? ¿Qué tipo de almacenamiento y procesadores?
- ✓ ¿Cómo instalar el software en los servidores de desarrollo, prueba y producción?
- ✓ ¿Qué necesitan instalar los diferentes miembros del equipo de DW/BI en sus estaciones de trabajo?
- ✓ ¿Cómo convertir el modelo de datos lógico en un modelo de datos físicos en la base de datos relacional?
- ✓ ¿Cómo conseguir un plan de indexación inicial?
- ✓ ¿Debe usarse la partición en las tablas relacionales?

❖ **Diseño del sistema de extracción, transformación y carga (ETL):**

Es la base sobre la cual se alimenta el Data Warehouse. Si el sistema ETL se diseña adecuadamente, puede extraer los datos de los sistemas de origen de datos, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos, consolidar la información proveniente de distintos sistemas, y finalmente cargar (grabar) la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis.

❖ **Especificación y desarrollo de aplicaciones BI:** Las aplicaciones de BI son la cara visible de la inteligencia de negocios: los informes y aplicaciones de análisis proporcionan información útil a los usuarios. Las aplicaciones de BI incluyen un amplio espectro de tipos de informes y herramientas de análisis, que van desde informes simples de formato fijo a sofisticadas aplicaciones analíticas que usan complejos algoritmos e información del dominio. Kimball divide a estas aplicaciones en dos categorías basadas en el nivel de sofisticación, y les llama informes estándar y aplicaciones analíticas.

En conclusión, la metodología de Kimball proporciona una base empírica y metodológica adecuada para las implementaciones de almacenes de datos pequeños y medianos, dada su gran versatilidad y su enfoque ascendente, que permite construir los almacenes en forma escalonada. Además presenta una serie de herramientas, tales como planillas, gráficos y documentos, que proporcionan una gran ayuda para iniciarse en el ámbito de la construcción de un Data Warehouse.

2.2.7.2. Metodología de Bill Inmon

Bill Inmon ve la necesidad de transferir la información de los diferentes OLTP (Sistemas Transaccionales) de las organizaciones a un lugar centralizado donde los datos puedan ser utilizados para el análisis (sería el CIF o Corporate Information Factory). Insiste además en que ha de tener las siguientes características:

Orientado a temas.- Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.

Variante en el tiempo.- Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.

No volátil.- La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas.

Integrado.- La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes. Inmon defiende una metodología descendente (top-down) a la hora de diseñar un almacén de datos, ya que de esta forma se considerarán mejor todos los datos corporativos. En esta metodología los Data Mart se crearán después de haber terminado el Data Warehouse completo de la organización.

CIF, enfoque creado por Bill Inmon, es una metodología descendente (topdown). Se basa en la creación de un repositorio de datos corporativo como fuente de información consolidada, persistente, histórica y de calidad. Consiste como el anterior enfoque en dos fases:

Getting Data In: incluye los procesos ETL de las fuentes de origen (que incluyen Data Quality, Data Staging u otros procesos), creación del Data Warehouse Corporativo y, en caso de ser necesario, ODS y Staging Area.

Getting Information Out: procesos ETL que distribuyen la información a los diferentes Data Mart que responden necesidades diferentes (lo que llamamos Data Delivery). Las herramientas de explotación de datos atacan los Data Mart y en casos excepcionales al mismo Data Warehouse Corporativo.

Como es fácil comprender la búsqueda de un enfoque único para cada uno de los conceptos de la organización prima en este enfoque. Por ejemplo, hay veces que el concepto de cliente se entiende de forma sutilmente diferente por los diversos departamentos existentes. Bajo esta metodología se buscaría un entendimiento común. [20]



Figura 6. Enfoque Inmon [21]

2.2.7.3. Cuadro comparativo de Metodologías

A continuación, se presenta un cuadro comparativo entre la metodología de Kimball e Inmon.

Tabla 2. Kimball VS Inmon

	Kimball	Inmon
Diseño	Utiliza el enfoque "Bottom – Up"	Utiliza el enfoque "Top – Down"
Enfoque	Tiene un enfoque por procesos que son manejados por las diferentes áreas del proceso. Trata de responder necesidades específicas.	Tiene un enfoque global de toda la organización. No está basado en requerimientos específicos.
Costos	Implementar el Data Mart no presenta alto costo.	Los costos son elevados ya implica trabajo de grandes cantidades de datos.
Tiempo	Tiempo de implementación es corto.	Demanda mucho más tiempo.
Alcance	Departamentos individuales.	Toda la organización.
Modelado	Kimball plantea usar el modelamiento dimensional: esquema estrella. Identificación de dimensiones y hechos.	Inmon propone tres niveles en el modelo de datos del Data Warehouse: Alto Nivel, Nivel Medio, Nivel Bajo.

Al establecer una comparación entre las dos metodologías que son la metodología de Ralph Kimball (y su enfoque dimensional), y la metodología de Bill Inmon (y su enfoque empresarial Warehouse).

La metodología de Inmon es más apropiada para sistemas complejos, donde se quiere asegurar la perdurabilidad y consistencia de la información aunque cambien los procesos de negocio de la organización. Pero para proyectos pequeños donde se quiere asegurar la usabilidad de los usuarios, que permita un desarrollo rápido e incremental de la solución donde no se tiene claro el panorama global, el enfoque de Kimball es el más apropiado.

2.2.8. Microsoft SQL Server Analysis Services (SSAS)

Microsoft SQL Server 2005 Analysis Services (SSAS) ofrece funciones de procesamiento analítico en línea (OLAP) y minería de datos para aplicaciones de Business Intelligence. Analysis Services admite OLAP y permite diseñar, crear y administrar estructuras multidimensionales que contienen datos agregados desde otros orígenes de datos, como bases de datos relacionales. En el caso de las aplicaciones de minería de datos, Analysis Services permite diseñar, crear y visualizar modelos de minería de datos que se construyen a partir de otros orígenes de datos mediante el uso de una gran variedad de algoritmos de minería de datos estándar del sector. [22]

2.2.9. Microsoft SQL Server Integration Services (SSIS)

Microsoft Integration Services es una plataforma para la construcción de integración de datos y los datos de nivel empresarial soluciones transformaciones. Utilizar servicios de integración para resolver problemas complejos de negocios mediante la copia o la descarga de archivos, el envío de mensajes de correo electrónico en respuesta a eventos, la actualización de los almacenes de datos, limpieza y extracción de datos y la gestión de objetos y datos de SQL Server. Los paquetes pueden trabajar solos o en concierto con otros paquetes para hacer frente a las necesidades empresariales complejas. Se pueden extraer y transformar datos de una amplia variedad de fuentes, tales como archivos XML de datos, archivos planos, y fuentes de datos relacionales, y luego

cargar los datos en uno o más destinos. Integration Services incluye un amplio conjunto de tareas incorporadas y transformaciones; herramientas para la construcción de paquetes; y el servicio Integration Services para ejecutar y administrar paquetes. Puede utilizar las herramientas gráficas de Integration Services para crear soluciones sin escribir una sola línea de código; o se puede programar el modelo extensivo de objetos de Integration Services para crear paquetes mediante programación y tareas personalizadas de código y otros objetos de paquete. [23]

2.2.10. Power BI

Power BI es un conjunto de aplicaciones de análisis de negocios que permite analizar datos y compartir información. Los paneles de Power BI ofrecen a los usuarios una vista de 360 grados con sus métricas más importantes en un mismo lugar. La información se actualiza en tiempo real y está disponible en todos sus dispositivos. Con un solo clic, los usuarios pueden explorar los datos subyacentes del panel mediante herramientas intuitivas que permiten obtener respuestas fácilmente. La creación de un panel es una sencilla operación gracias a las más de 50 conexiones a conocidas aplicaciones empresariales, que se completan con paneles pregenerados y diseñados por expertos para ayudarle a ponerse en marcha rápidamente. Asimismo, puede acceder a sus datos e informes desde cualquier lugar con las aplicaciones móviles de Power BI Mobile, que se actualizan automáticamente con los cambios que se realizan en los datos.

Si es un analista de datos que proporciona informes y análisis a su organización, Power BI le permite liberar todo su potencial creativo y alcanzar la máxima productividad en sus tareas. Power BI Desktop es una herramienta de mashup de datos y creación de informes que incluye numerosas características. Combine datos de bases de datos, archivos y servicios web diversos con herramientas visuales que ayudan a comprender y corregir problemas de formato y calidad de los datos automáticamente. Con más de 20 objetos visuales integrados y una dinámica comunidad de visualizaciones personalizadas, podrá crear informes espectaculares que comuniquen su mensaje con la máxima

eficacia. Con el servicio Power BI, publique informes con seguridad en su organización y configure la actualización de datos automática para que todo el mundo disponga de la información más reciente.

Power BI puede unificar todos los datos de su organización, ya sea en la nube o localmente. Con Power BI Gateways, puede conectar bases de datos SQL Server, modelos de Analysis Services y muchos otros orígenes de datos a los mismos paneles en Power BI. Si ya cuenta con portales o aplicaciones de creación de informes, inserte los informes y paneles de Power BI para disfrutar de una experiencia unificada. [24]

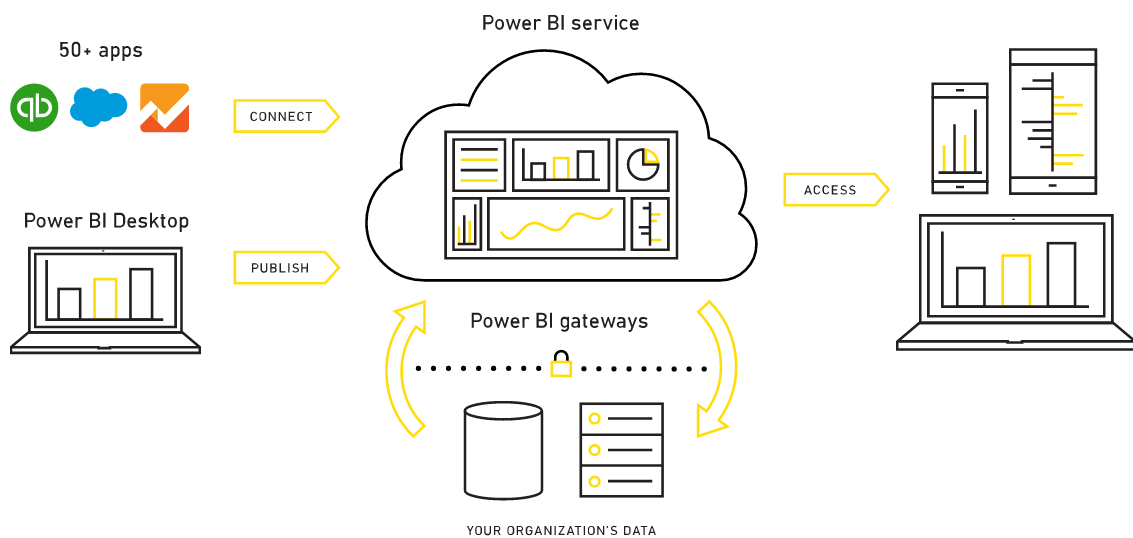


Figura 7. Power BI [24]

2.3. Definición de Términos Básicos

2.3.1. Cubo

Un cubo es una estructura multidimensional que contiene información con fines analíticos; sus componentes principales son las dimensiones y las medidas. Las dimensiones definen la estructura del cubo que se utiliza para segmentar y dividir los datos, y las medidas proporcionan valores numéricos agregados importantes para el usuario final. Como estructura lógica, un cubo permite a una aplicación cliente recuperar valores, de medidas, como si estuvieran almacenados en las celdas del cubo; las celdas se definen para cada posible valor resumido. Las celdas del cubo se definen por la intersección de miembros de dimensión y contienen los valores agregados de las medidas en esa intersección concreta. [25]

2.3.2. Datos

En el mundo de la informática, la unidad básica en la que trabajamos es definida como dato, teniendo este concepto una aserción bastante general como un caracter, una representación o un símbolo que se encuentra en forma aislada, sin un contexto determinado y sin un debido ordenamiento. Un dato es una representación simbólica (numérica, alfabética, algorítmica, espacial, etc.) de un atributo o variable cuantitativa o cualitativa. Los datos describen hechos empíricos, sucesos y entidades. Es por eso que se establece que los datos aislados tienen la necesidad de una interpretación, dándoles un contexto y un orden que permita una interpretación y un resultado único.

2.3.3. Información

Según Idalberto Chiavenato, información "es un conjunto de datos con un significado, o sea, que reduce la incertidumbre o que aumenta el conocimiento de algo. En verdad, la información es un mensaje con significado en un determinado contexto, disponible para uso inmediato y que proporciona orientación a las acciones por el hecho de reducir el margen de incertidumbre con respecto a nuestras decisiones" [26].

Para Ferrell y Hirt, la información "comprende los datos y conocimientos que se usan en la toma de decisiones" [27].

Según Czinkota y Kotabe la información "consiste en datos seleccionados y ordenados con un propósito específico" [28].

Entonces podemos concluir que los datos organizados e interpretados son considerados como Información.

2.3.4. Proceso Educativo

El proceso educativo se basa en la transmisión de valores y saberes. El proceso educativo no suele ser unidireccional, sino que es interactivo: quienes están aprendiendo, también pueden enseñar. Así el conocimiento se construye de forma social. [29]

La acción sinérgica de la Comunidad Educativa que gestiona dinámica, corresponsable y pertinentemente elementos curriculares, planificativos y administrativos, para el desarrollo integral del estudiante a continuación la malla curricular de la Escuela Académico Profesional de Ingeniería de

Sistemas, la cantidad de horas por curso y el creditaje se encuentran en el Anexo1 y Anexo 4.

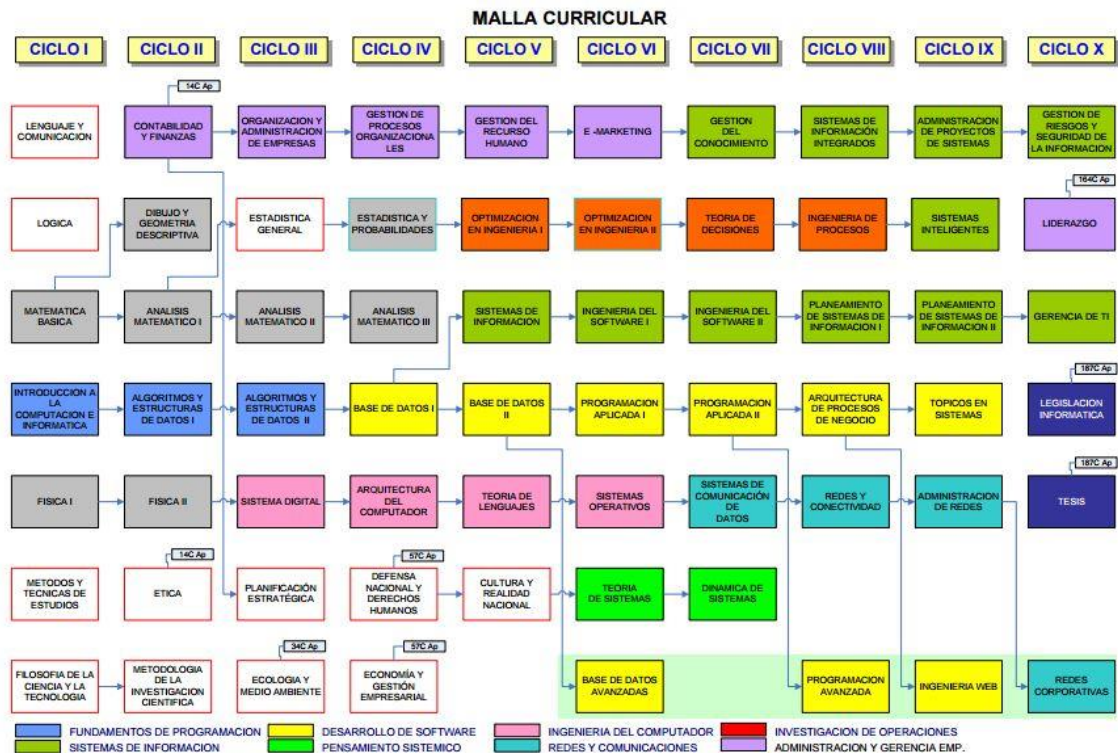


Figura 8. Malla curricular Ingeniería de Sistemas [30]

2.3.5. Rendimiento Académico

"Del latín *reddere* (restituir, pagar) el rendimiento es una relación entre lo obtenido y el esfuerzo empleado para obtenerlo. Es un nivel de éxito en la universidad, en el trabajo, etc.",

En otras palabras, el rendimiento académico es la medida de las capacidades del alumno, que expresa lo que éste ha aprendido a lo largo del proceso formativo. También supone la capacidad del alumno para responder a los estímulos educativos. En este sentido, el rendimiento académico está vinculado a la aptitud. [31]

Según Herán y Villarroel (1987). El rendimiento académico se define en forma operativa y tácita afirmando que se puede comprender el rendimiento previo como el número de veces que el estudiante ha repetido uno o más cursos.

En tanto Nováez (1986) sostiene que el rendimiento académico es el resultado obtenido por el individuo en determinada actividad académica. El concepto de rendimiento está ligado al de aptitud, y sería el resultado

de ésta, de factores volitivos, afectivos y emocionales, además de la ejercitación.

Chadwick (1979) define el rendimiento académico como la expresión de capacidades y de características psicológicas del estudiante desarrolladas y actualizadas a través del proceso de enseñanza-aprendizaje que le posibilita obtener un nivel de funcionamiento y logros académicos a lo largo de un período, año o semestre, que se sintetiza en un calificativo final (cuantitativo en la mayoría de los casos) evaluador del nivel alcanzado.

2.3.6. Seguimiento Académico

El seguimiento académico implica monitorear el desempeño estudiantil para obtener información confiable, actualizada, útil y oportuna sobre los aciertos, dificultades y problemas que se presentan en el diseño y desarrollo del plan de estudios, el cual tiene como referente los objetivos y estándares curriculares.

La información obtenida del seguimiento sirve para realizar los cambios necesarios, con el fin de promover el desarrollo de estrategias que le permitan al alumno revisar y comprender sus procesos en el aprendizaje de los diversos contenidos curriculares, entender dónde radican sus dificultades, qué tipo de contenidos se le facilitan y cómo puede mejorar su aprovechamiento académico para asumir y dirigir sus aprendizajes a lo largo de su vida. [3]

CAPÍTULO III MATERIALES Y MÉTODOS

3.1. Procedimiento

A continuación se desarrolla el proyecto de investigación siguiendo la metodología escogida, Ralph Kimball, la cual está compuesta por las siguientes fases: planificación de proyecto, definición de los requerimientos del negocio, modelado dimensional, diseño físico, diseño y desarrollo de la presentación de datos y diseño de la arquitectura técnica.

3.1.1. Planeación de Proyecto

3.1.1.1. Objetivos

- ✓ Mejorar la distribución de la información entre la Unidad Técnica de Sistemas Informáticos y la Escuela de Sistemas.
- ✓ Generar información veraz y oportuna, dirigida al usuario final en forma consolidada y representativa para el proceso de toma de decisiones.
- ✓ Integrar la información histórica de forma consistente.
- ✓ Facilitar el proceso de toma de decisiones a la Escuela Académico Profesional de Ingeniería de Sistemas.

3.1.1.2. Alcance

El presente proyecto busca ayudar a la gestión de la Escuela Académico Profesional de Ingeniería de Sistemas a través de reportes con información histórica, veraz, consolidada, permitiendo dar soporte a la toma de decisiones.

3.1.1.3. Beneficios

Tabla 3. Beneficios de data mart

	Beneficios
Humano	<ul style="list-style-type: none">- Equipo de trabajo capaz de tomar mejores decisiones.- Enriquecer el proceso de análisis de información.- Optimizar tiempos dentro de la Escuela Académico Profesional de Ingeniería de Sistemas.
Tecnológico	<ul style="list-style-type: none">- Herramientas de Información acordes con las existentes en la Unidad Técnica de Sistemas Informáticos y accesibles en toda la universidad.- Tecnología informática amigable con los usuarios.
Material	<ul style="list-style-type: none">- Reducción de tiempos en la emisión de reportes consolidados.
Información	<ul style="list-style-type: none">- Información veraz y oportuna.- Toma de decisiones adecuadas basadas en reportes dinámicos.

3.1.1.4. Entrevistas

Las entrevistas se realizaron a las siguientes personas:

- Ing. Walter Ronald Pérez Estrada
- Ing. Manuel Pérez

Se orientó principalmente en los lineamientos:

- Entender la forma de extraer reportes de la base de datos transaccional.
- Entender el funcionamiento de la base de datos transaccional.
- Comprender la forma de análisis de datos.
- Conocer las herramientas de software normalmente utilizadas.
- Ing. Manuel Malpica
- Ing. Jaime Meza

Se orientó principalmente en:

- Entender la información que utilizan para tomar decisiones.
- Entender los procesos que llevan a cabo para la toma de decisiones.
- Obtener tiempos empleados en el análisis y obtención de información.

3.1.1.5. Recurso Humano

Tabla 4. Recurso Humano

	Nombre y Apellido	Ocupación
Project Manager	Liliana Elizabeth Durán Rafael	Tesista
Project Manager	Ing. Jaime Meza Huamán	Asesor

3.1.2. Definición de requerimientos del negocio

3.1.2.1. Requerimientos del negocio

Los requerimientos presentados están enfocados en el seguimiento académico de los alumnos de la Escuela Académico Profesional de Ingeniería de Sistemas.

Tabla 5. Cuadro de requerimiento Nro. 1

Identificador	R01	Nombre	Elaborar modelo dimensional
Tipo	No Funcional		
Descripción	La solución de inteligencia de negocios permitirá implementar un modelo estrella		

Tabla 6. Cuadro de requerimiento Nro. 2

Identificador	R02	Nombre	Elaborar un cubo
Tipo	No Funcional		
Descripción	La solución de inteligencia de negocios será mediante la elaboración de un cubo		

Tabla 7. Cuadro de requerimiento Nro. 3

Identificador	R03	Nombre	Número de estudiantes por año de ingreso
Tipo	Funcional		
Descripción	La solución de inteligencia de negocios mostrará la cantidad de ingresantes a la escuela por año.		

Tabla 8. Cuadro de requerimiento Nro. 4

Identificador	R04	Nombre	Número de veces de llevar un curso
Tipo	Funcional		
Descripción	La solución de inteligencia de negocios permitirá ver cuántas veces un alumno ha llevado un curso por año de ingreso y año de estudios.		

Tabla 9. Cuadro de requerimiento Nro. 5

Identificador	R05	Nombre	Número de alumnos aprobados por curso
Tipo	Funcional		
Descripción	La solución de inteligencia de negocios mostrará el número de alumnos aprobados en un curso por año de ingreso, utilizando de guía el anexo 2.		

Tabla 10. Cuadro de requerimiento Nro. 6

Identificador	R06	Nombre	Número de alumnos desaprobados por curso
Tipo	Funcional		
Descripción	La solución de inteligencia de negocios facilitará saber cuántos alumnos desaprobaron un curso por año de ingreso, utilizando de guía el anexo 2.		

Tabla 11. Cuadro de requerimiento Nro. 7

Identificador	R07	Nombre	Estado del estudiante
Tipo	Funcional		
Descripción	La solución de inteligencia de negocios mostrará el alumno, cursos y el número de veces que ha llevado cada curso por año de ingreso y año de estudios, utilizando de referencia el anexo 3.		

Tabla 12. Cuadro de requerimiento Nro. 8

Identificador	R08	Nombre	Número de alumnos por nota y curso
Tipo	Funcional		
Descripción	La solución de inteligencia de negocios mostrará el número de alumnos en las diferentes notas por curso por año de ingreso.		

Tabla 13. Cuadro de requerimiento Nro. 9

Identificador	R09	Nombre	Consolidado de alumnos
Tipo	Funcional		
Descripción	La solución de inteligencia de mostrará la cantidad de alumnos desaprobados, aprobados e inhabilitados por curso y periodo, anexo 5.		

Tabla 14. Cuadro de requerimiento Nro. 10

Identificador	R10	Nombre	Estado de alumnos por docente
Tipo	Funcional		
Descripción	La solución mostrará la cantidad de alumnos desaprobados, aprobados e inhabilitados por docente, año de ingreso y año de estudios, utilizando como referencia el anexo 2.		

Tabla 15. Cuadro de requerimiento Nro. 11

Identificador	R11	Nombre	Software
Tipo	No Funcional		
Descripción	La solución de inteligencia de negocios será implementada con software utilizado en la universidad, herramientas Microsoft.		

Tabla 16. Cuadro de requerimiento Nro. 12

Identificador	R12	Nombre	Software
Tipo	No Funcional		
Descripción	Windows 7 y Windows server 2008 R2 o posterior, .NET 4.5, internet explorer 9 o posterior o google chrome.		

Tabla 17. Cuadro de requerimiento Nro. 13

Identificador	R13	Nombre	Hardware
Tipo	No Funcional		
Descripción	Los requisitos minimos son memoria RAM de al menos 1GB disponible, la pantalla de al menos 1440x900 o 1600x900 y CPU de 1Ghzo superior.		

3.1.2.2. Base de datos transaccional en SQL server

La base de datos mostrada a continuación es tal cual fue proporcionada:

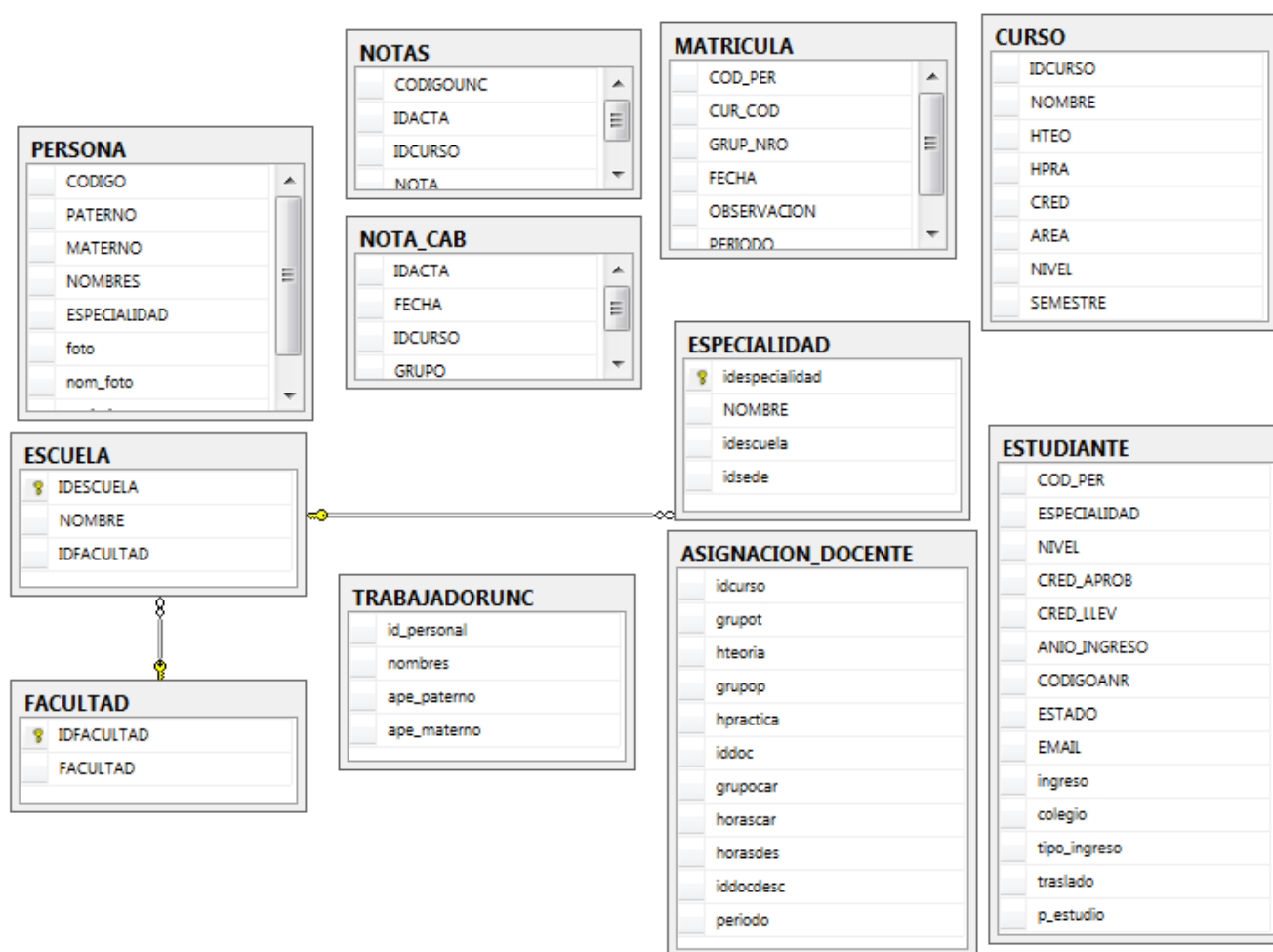


Figura 9. Base de datos transaccional en SQL Server

Base de datos con relaciones encontradas:

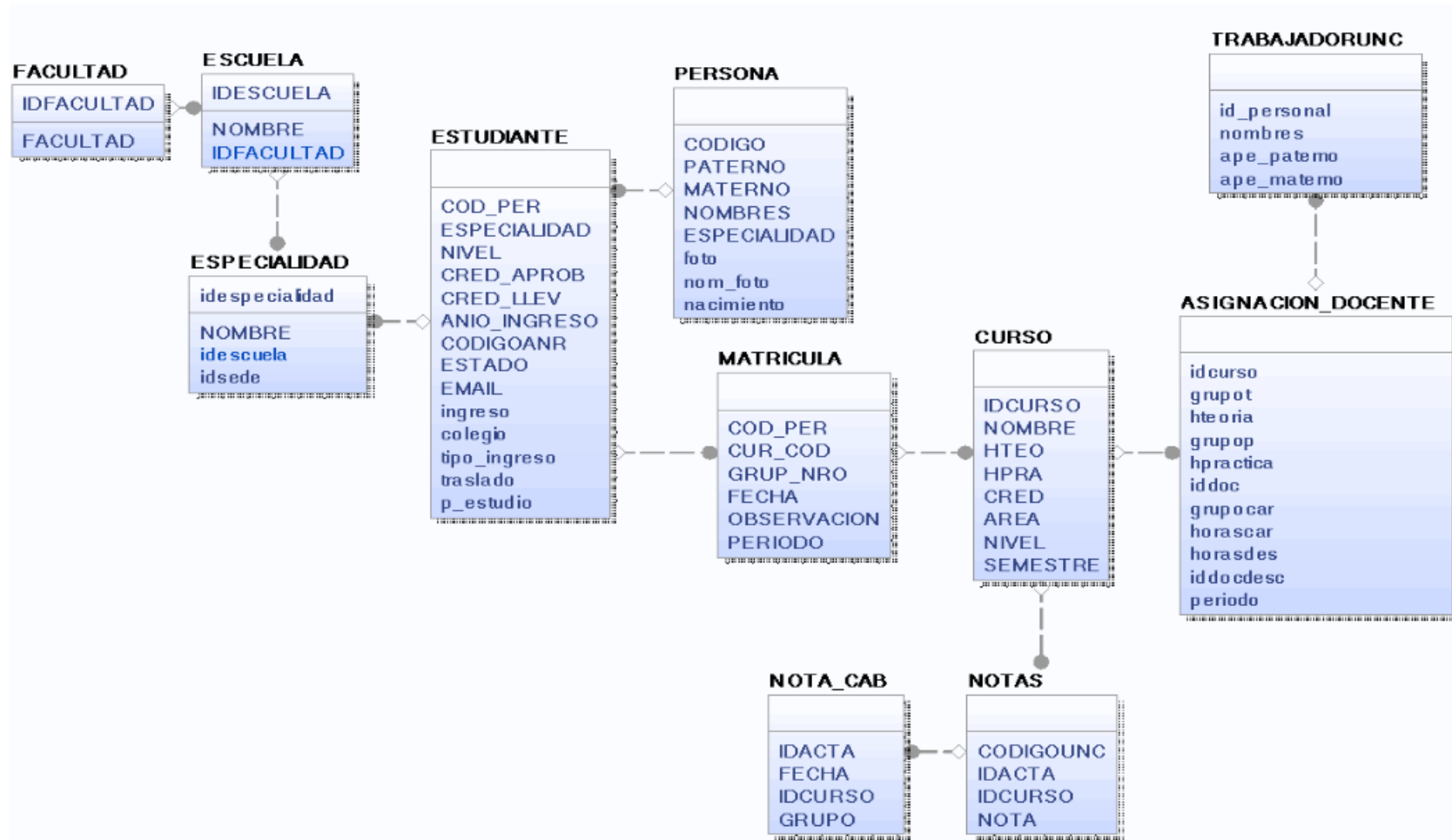


Figura 10. Case de datos con relaciones

PERSONA	
<input type="checkbox"/>	CODIGO
<input type="checkbox"/>	PATERNO
<input type="checkbox"/>	MATERNO
<input type="checkbox"/>	NOMBRES
<input type="checkbox"/>	ESPECIALIDAD
<input type="checkbox"/>	foto
<input type="checkbox"/>	nom_foto
<input type="checkbox"/>	nacimiento

Figura 11. Tabla PERSONA

NOTAS	
<input type="checkbox"/>	CODIGOUNC
<input type="checkbox"/>	IDACTA
<input type="checkbox"/>	IDCURSO
<input type="checkbox"/>	NOTA

Figura 12. Tabla NOTAS

TRABAJADORUNC	
<input type="checkbox"/>	id_personal
<input type="checkbox"/>	nombres
<input type="checkbox"/>	ape_paterno
<input type="checkbox"/>	ape_materno

Figura 13. Tabla TRABAJADORUNC

NOTA_CAB	
<input type="checkbox"/>	IDACTA
<input type="checkbox"/>	FECHA
<input type="checkbox"/>	IDCURSO
<input type="checkbox"/>	GRUPO

Figura 14. Tabla NOTA_CAB


FACULTAD	
<input type="checkbox"/>	 IDFACULTAD
<input type="checkbox"/>	FACULTAD

Figura 15. Tabla FACULTAD

ESCUELA	
	IDESCUELA
	NOMBRE
	IDFACULTAD

Figura 16. Tabla ESCUELA

ESPECIALIDAD	
	idespecialidad
	NOMBRE
	idescuela
	idsede

Figura 17. Tabla ESPECIALIDAD

MATRICULA	
	COD_PER
	CUR_COD
	GRUP_NRO
	FECHA
	OBSERVACION
	PERIODO

Figura 18. Tabla MATRICULA

ESTUDIANTE	
	COD_PER
	ESPECIALIDAD
	NIVEL
	CRED_APROB
	CRED_LLEV
	ANIO_INGRESO
	CODIGOANR
	ESTADO
	EMAIL
	ingreso
	colegio
	tipo_ingreso
	traslado
	p_estudio

Figura 19. Tabla ESTUDIANTE

CURSO	
	IDCURSO
	NOMBRE
	HTEO
	HPRA
	CRED
	AREA
	NIVEL
	SEMESTRE

Figura 20. Tabla CURSO

ASIGNACION_DOCENTE	
	idcurso
	grupot
	hteoria
	grupop
	hpractica
	iddoc
	grupocar
	horascar
	horasdes
	iddocdesc
	periodo

Figura 21. Tabla ASIGNACION_DOCENTE

3.1.2.3. Diccionario de datos a nivel general de la base de datos transaccional

Tabla 18. Base de datos transaccional

Tabla	Descripción
PERSONA	Se registran todos los alumnos ingresantes a la Universidad Nacional de Cajamarca.
NOTAS	Se registran todas las notas que tiene cada curso.
TRABAJADORUNC	Se registran todos los trabajadores de la Universidad Nacional de Cajamarca.
NOTA_CAB	Se registran las actas en la que figuran las notas de los cursos.
FACULTAD	Se registran todas las facultades de la Universidad Nacional de Cajamarca.
ESCUELA	Se registran todas las escuelas de la Universidad Nacional de Cajamarca.
ESPECIALIDAD	Se registran todas las especialidades de la Universidad Nacional de Cajamarca.
MATRICULA	Se registran los alumnos matriculados en un curso en diferentes periodos.
ESTUDIANTE	Se registran todos los estudiantes de la Universidad Nacional de Cajamarca, código y especialidad.
CURSO	Se registran todos los cursos existentes en las diferentes escuelas de la Universidad Nacional de Cajamarca.
ASIGNACION_DOCENTE	Se registran los docentes, los cursos que dictan y los diferentes grupos que tienen asignados.

3.1.2.4. Datos de la base de datos transaccional

A continuación un análisis de los datos de las tablas de la base de datos transaccional.

Tabla 19. Descripción de la tabla PERSONA

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
CODIGO	Varchar(8)	Not Null	Código de la persona	PK	-
PATERNO	Varchar(50)	Null	Apellido paterno	-	-
MATERNO	Varchar(50)	Null	Apellido materno	-	-
NOMBRES	Varchar(50)	Null	Nombres completos	-	-
ESPECIALIDAD	Varchar(2)	Null	Número de especialidad	-	-
FOTO	image	Null	Foto de la persona	-	-
NOM_FOTO	Varchar(50)	Null		-	-
NACIMIENTO	smalldatetime	Null	Fecha de nacimiento	-	-

Tabla 20. Descripción de la tabla NOTAS

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
CODIGOUNC	Varchar(8)	Null	Código de estudiante	-	-
IDACTA	Varchar(10)	Null	Código de acta	-	-
IDCURSO	Varchar(6)	Null	Código de curso	-	-
NOTA	Varchar(2)	Null	Nota de cada curso	-	-

Tabla 21. Descripción de la tabla TRABAJADORUNC

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
ID_PERSONAL	Char(6)	Null	Código de trabajador	PK	-
NOMBRES	Varchar(200)	Null	Nombres completos	-	-
APE_PATERNO	Varchar(60)	Null	Apellido paterno	-	-
APE_MATERNO	Varchar(60)	Null	Apellido materno	-	-

Tabla 22. Descripción de la tabla NOTA_CAB

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
IDACTA	Varchar(10)	Not Null	Código de acta	PK	-
FECHA	smalldatetime	Null	Fecha de registro de acta	-	-
IDCURSO	Varchar(6)	Null	Código del curso	-	FK
GRUPO	Varchar(6)	Null	Grupo de asistencia	-	-

Tabla 23. Descripción de la tabla FACULTAD

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
IDFACULTAD	Varchar(2)	Not null	Código de facultad	PK	-
FACULTAD	Varchar(50)	Null	Nombre de la facultad	-	-

Tabla 24. Descripción de la tabla ESCUELA

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
IDESCUELA	Varchar(2)	Not Null	Código de escuela	PK	-
NOMBRE	Varchar(40)	Null	Nombre de la escuela	-	-
IDFACULTAD	Varchar(2)	Null	Código de facultad	-	FK

Tabla 25. Descripción de la tabla ESPECIALIDAD

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
IDESPECIALIDAD	Varchar(2)	Not Null	Código de especialidad	PK	-
NOMBRE	Varchar(50)	Null	Nombre de especialidad	-	-
IDESCUELA	Varchar(2)	Null	Código de escuela	-	FK
IDSEDE	Varchar(2)	Null	Código de sede	-	-

Tabla 26. Descripción de la tabla MATRICULA

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
COD_PER	Varchar(8)	Not Null	código de persona	-	FK
CUR_COD	Varchar(6)	Not Null	código de curso	-	FK
GRUP_NRO	Varchar(5)	Null	Grupo al que se matriculó	-	-
FECHA	Smalldatetime	Null	Fecha de matricula	-	-
OBSERVACION	Varchar(15)	Null		-	-
PERIODO	Varchar(10)	Not Null	Periodo académico	-	-

Tabla 27. Descripción de la tabla ESTUDIANTE

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
COD_PER	Varchar(8)	Not Null	Código de estudiante	PK	-
ESPECIALIDAD	Varchar(2)	Not Null	Código de especialidad	-	-
NIVEL	Int	Null	Año de estudios	-	-
CRED_APROB	Int	Null	Créditos aprobados	-	-
CRED_LLEV	Int	Null	Créditos llevados	-	-
ANIO_INGRESO	Varchar(50)	Null	Año de ingreso	-	-
CODIGOANR	Varchar(10)	Null	Código ANR	-	-
ESTADO	Char(2)	Null	-	-	-
EMAIL	Varchar(50)	Null	Email universitario	-	-
INGRESO	Int	Null	-	-	-
COLEGIO	Char(1)	Null	-	-	-
TIPO_INGRESO	Char(2)	Null	-	-	-
TRASLADO	Int	Null	-	-	-
P_ESTUDIO	Char(4)	Null	Plan de estudios	-	-

Tabla 28. Descripción de la tabla CURSO

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
IDCURSO	Varchar(6)	Not Null	Código de Curso	PK	-
NOMBRE	Varchar(50)	Null	Nombre de curso	-	-
HTEO	Float	Null	Horas de teoría	-	-
HPRA	Float	Null	Horas de practica	-	-
CRED	Float	Null	Creditos del curso	-	-
AREA	Varchar(50)	Null	-	-	-
NIVEL	Int	Null	Año de estudios al que pertenece el curso.	-	-
SEMESTRE	Varchar(1)	Null	Semestre en que se dicta el curso	-	-

Tabla 29. Descripción de la tabla ASIGNACION_DOCENTE

Nombre Columna	Tipo de Dato	Null Option	Descripción	PK	FK
IDCURSO	Char(6)	Not Null	Código de curso	PK	-
GRUPOT	Varchar(3)	Null	Grupo de teoría	-	-
HTEORIA	Int	Null	Horas de teoría	-	-
GRUPOP	Varchar(3)	Null	Grupo de práctica	-	-
HPRACTICA	Int	Null	Horas de práctica	-	-
IDDOC	Char(6)	Not Null	Código de docente	-	-
GRUPOCAR	Varchar(3)	Not Null	Carga en cada grupo	-	-
HORASCAR	Int	Null	Horas de carga	-	-
HORASDES	Int	Null	-	-	-
IDDOCDESC	Char(6)	Null	-	-	-
PERIODO	Varchar(5)	Not Null	Periodo de la asignación de docente	-	-

3.1.3. Modelado dimensional

Habiendo realizado el análisis de las entrevistas y los requerimientos, a continuación se identificaron las medidas y dimensiones orientadas a analizar la información en sus diferentes niveles.

3.1.3.1. Elección de la dimensiones

Para determinar las dimensiones con las que va a contar el Data Mart, se empezó identificando las variables de análisis por las cuales el usuario suele solicitar sus reportes. Entre las más destacadas se encontraron:

- Notas de los alumnos
- Información de alumnos
- Docentes
- Cursos dictados
- Ciclos académicos

Se agrupan las variables entre ellas, generalmente cada una es una característica o atributo de alguna entidad importante que podría ser una dimensión.

3.1.3.2. Dimensiones encontradas

Luego del análisis anterior se concluye que las dimensiones que conforman el Data Mart son:

- ASIGNACION_DOCENTE_DIM
- CURSOS_DIM
- ESCUELAS_DIM
- ESPECIALIDADES_DIM
- ESTUDIANTES_DIM
- FACULTAD_DIM
- MATRICULA_DIM
- NOTA_CA_DIM

3.1.3.3. Medidas encontradas

De acuerdo al análisis realizado se encontraron las siguientes medidas.

- Número de veces que se ha llevado un curso.
- Número de alumnos por curso.

Diseño del modelo dimensional

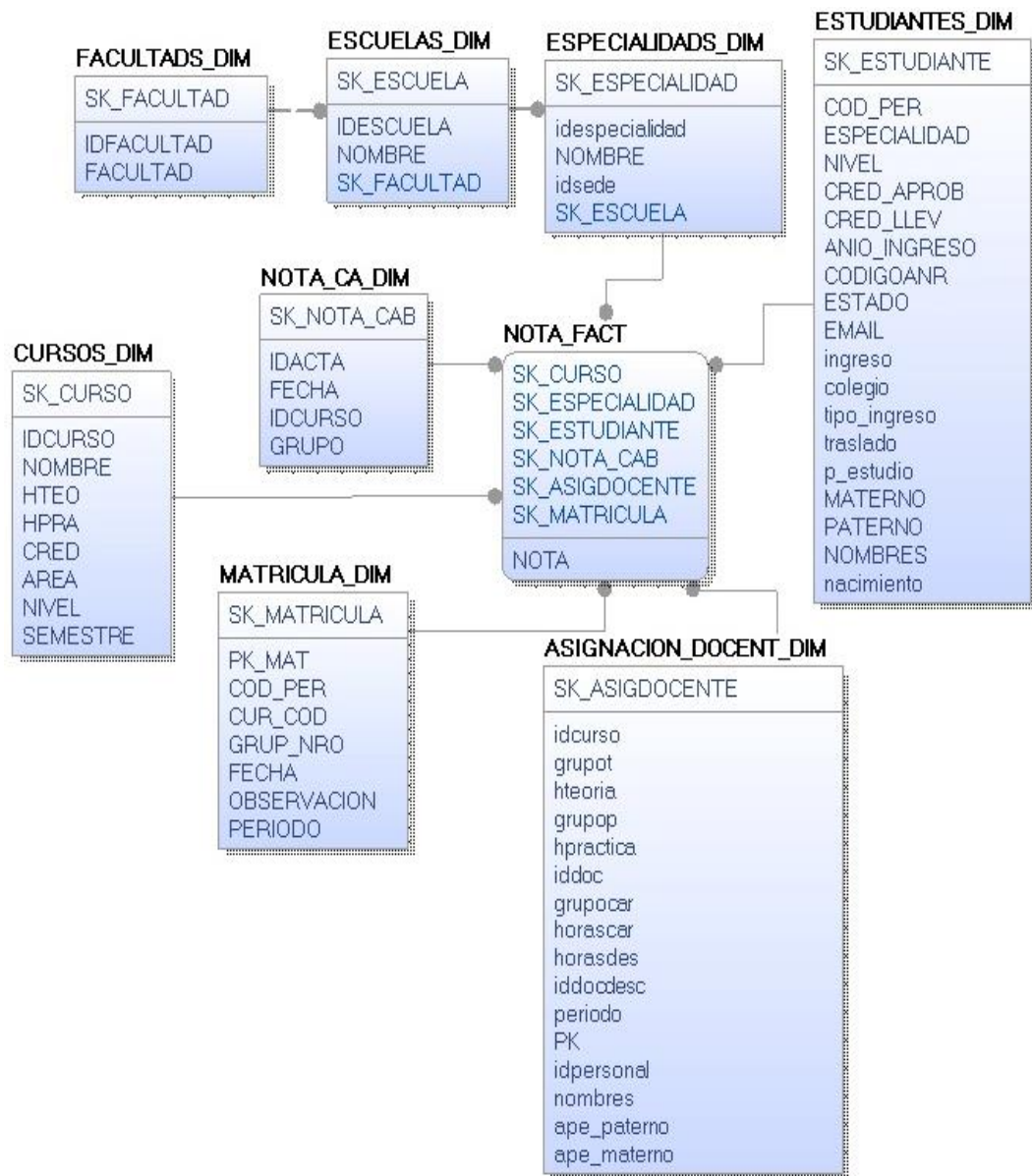


Figura 22. Diseño físico de la Data Mart

3.1.3.4. Dimensión ESTUDIANTES_DIM

La dimensión ESTUDIANTES_DIM está conformada por:

Tabla 30. Tabla de diseño físico de dimensión ESTUDIANTES_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_ESTUDIANTE	Int	Clave
PATERNO	Varchar(50)	Apellido paterno
MATERNO	Varchar(50)	Apellido materno
NOMBRES	Varchar(50)	Nombres completos
ESPECIALIDAD	Varchar(2)	Número de especialidad
COD_PER	Varchar(8)	Código de alumno
NIVEL	Int	Año de estudios del alumno
NACIMIENTO	smalldatetime	Fecha de nacimiento
CRED_APROB	Int	Créditos aprobados
CRED_LLEV	Int	Créditos llevados
ANIO_INGRESO	Varchar(50)	Año de ingreso a la UNC
CODIGOANR	Varchar(10)	Código ANR de estudiante
ESTADO	Char(2)	-
EMAIL	Varchar(50)	Correo institucional de estudiante
INGRESO	Int	-
COLEGIO	Char(1)	Colegio de procedencia
TIPO_INGRESO	Char(2)	-
TRASLADO	Int	-
P_ESTUDIO	Char(4)	Plan de estudios que cursa el estudiante

3.1.3.5. Dimensión CURSOS_DIM

La dimensión CURSOS_DIM está conformada por:

Tabla 31. Tabla de diseño físico de dimensión CURSOS_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_CURSO	Int	Clave
IDCURSO	Varchar(6)	Código de Curso
NOMBRE	Varchar(50)	Nombre de curso
HTEO	Float	Horas de teoría
HPRA	Float	Horas de practica
CRED	Float	Creditos del curso
AREA	Varchar(50)	-
NIVEL	Int	Año de estudios al que pertenece el curso.
SEMESTRE	Varchar(1)	Semestre en que se dicta el curso

3.1.3.6. Dimensión MATRICULA_DIM

La dimensión MATRICULA_DIM está conformada por:

Tabla 32. Tabla de diseño físico de dimensión MATRICULA_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_MATRICULA	Int	Clave
COD_PER	Varchar(8)	código de persona
CUR_COD	Varchar(6)	código de curso
GRUP_NRO	Varchar(5)	Grupo al que se matriculó
FECHA	Smalldatetime	Fecha de matricula
OBSERVACION	Varchar(15)	
PERIODO	Varchar(10)	Periodo académico

3.1.3.7. Dimensión ASIGNACION DOCENTE_DIM

La dimensión ASIGNACION DOCENTE_DIM está conformada por:

Tabla 33. Tabla de diseño físico de dimensión ASIGNACION DOCENTE_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_ASIGDOCENTE	Int	Clave
IDCURSO	Char(6)	Código de curso
GRUPOT	Varchar(3)	Grupo de teoría
HTEORIA	Int	Horas de teoría
GRUPOP	Varchar(3)	Grupo de práctica
HPRACTICA	Int	Horas de práctica
IDDOC	Char(6)	Código de docente
GRUPOCAR	Varchar(3)	Carga en cada grupo
HORASCAR	Int	Horas de carga
HORASDES	Int	-
IDDOCDESC	Char(6)	-
PERIODO	Varchar(5)	Periodo de la asignación de docente
PK	Varchar(20)	Código generado por concatenación
IDPERSONAL	Char(6)	Código de trabajador
NOMBRES	Varchar(200)	Nombres completos
APE_PATERO	Varchar(60)	Apellido paterno
APE_MATERNO	Varchar(60)	Apellido materno

3.1.3.8. Dimensión NOTA_CA_DIM

La dimensión NOTA_CA_DIM está conformada por:

Tabla 34. Tabla de diseño físico de dimensión NOTA_CA_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_NOTA_CAB	Int	Clave
IDACTA	Varchar(10)	Código de acta
FECHA	smalldatetime	Fecha de registro de acta
IDCURSO	Varchar(6)	Código del curso
GRUPO	Varchar(6)	Grupo en el que se puso la nota

3.1.3.9. Dimensión FACULTADS_DIM

La dimensión MATRICULA_DIM está conformada por:

Tabla 35. Tabla de diseño físico de dimensión MATRICULA_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_FACULTAD	Int	Clave
IDFACULTAD	Varchar(2)	Código de facultad
FACULTAD	Varchar(50)	Nombre de la facultad

3.1.3.10. Dimensión ESCUELAS_DIM

La dimensión ESCUELAS_DIM está conformada por:

Tabla 36. Tabla de diseño físico de dimensión ESCUELAS_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_ESCUELA	Int	Clave
IDESCUELA	Varchar(2)	Código de escuela
NOMBRE	Varchar(40)	Nombre de la escuela
SK_FACULTAD	Int	Foreign key

3.1.3.11. Dimensión ESPECIALIDADES_DIM

La dimensión ESPECIALIDADES_DIM está conformada por:

Tabla 37. Tabla de diseño físico de dimensión ESPECIALIDADES_DIM

Nombre Columna	Tipo de Dato	Descripción
SK_ESPECIALIDAD	Int	Clave
IDESPECIALIDAD	Varchar(2)	Código de especialidad
NOMBRE	Varchar(50)	Nombre de especialidad
SK_ESCUELA	Varchar(2)	Foreing key
IDSEDE	Varchar(2)	Código de sede

3.1.4. Diseño e implementación del subsistema de ETL

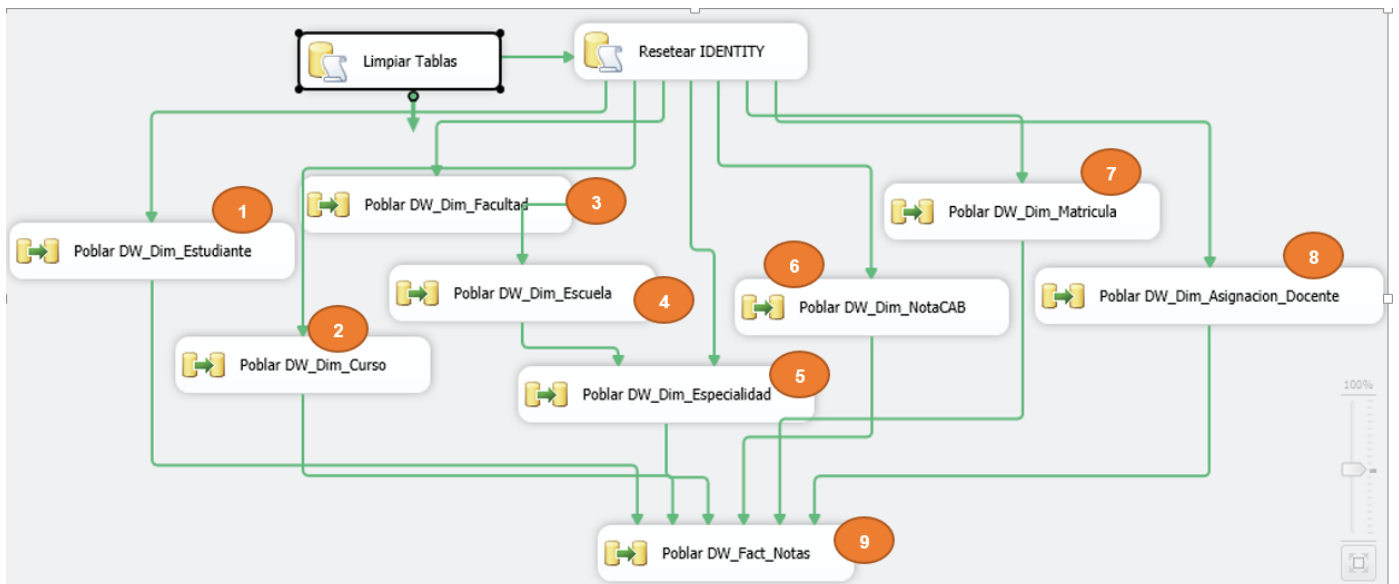


Figura 23. Diseño ETL de solución de inteligencia de negocios

Para la realización del proceso ETL de las dimensiones y fact table, se utilizó la herramienta INTEGRATION SERVICES que nos brinda SQL SERVER en su versión 2012.

3.1.4.1. Flujo de datos de ESTUDIANTES_DIM



Figura 24. Flujo de datos de "Poblar ESTUDIANTE_DIM"

Consulta

```
SELECT    E.COD_PER, E.ESPECIALIDAD, E.NIVEL,  
E.CRED_APROB, E.CRED_LLEV, E.ANIO_INGRESO,  
E.CODIGOANR, E.ESTADO, E.EMAIL, E.INGRESO,  
E.TIPO_INGRESO, E.TRASLADO, E.P_ESTUDIO, P.PATERNNO,  
P.MATERNNO, P.NOMBRES, P.NACIMIENTO  
  
        FROM  ESTUDIANTE AS E INNER JOIN PERSONA AS P  
  
        ON E.COD_PER = P.CODIGO  
  
        WHERE E.ANIO_INGRESO between 10 and 15
```

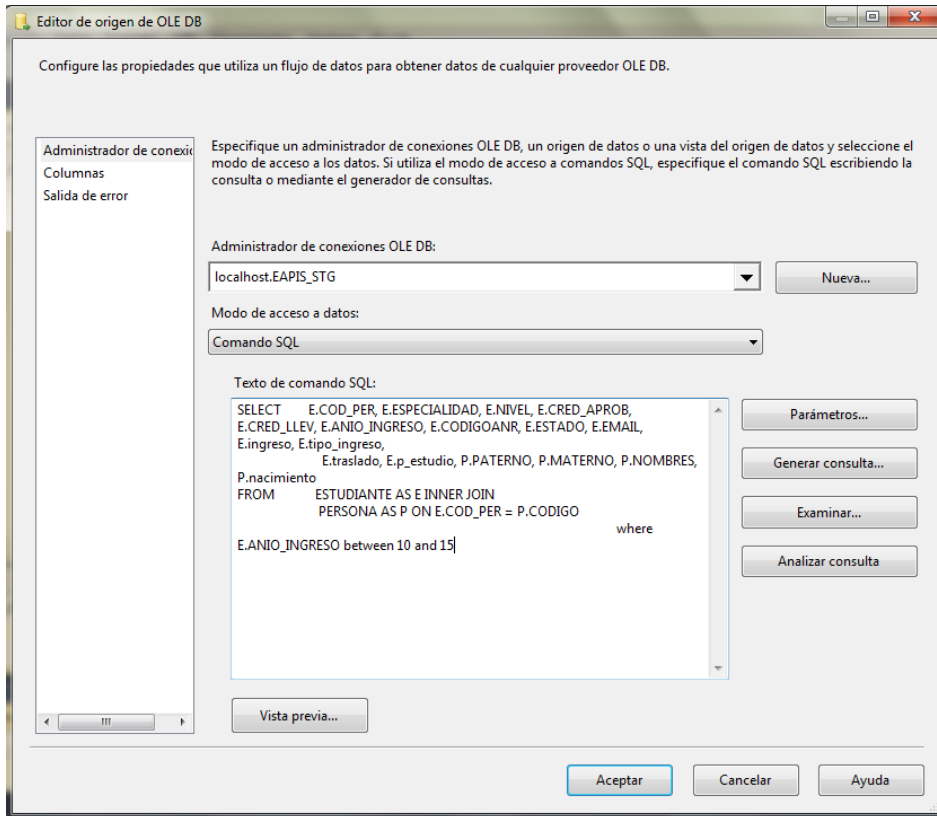



Figura 25. Origen OLE DB de "Poblar ESTUDIANTE_DIM"

Destino TransaccionalMart [dbo].[DIM_ESTUDIANTE]

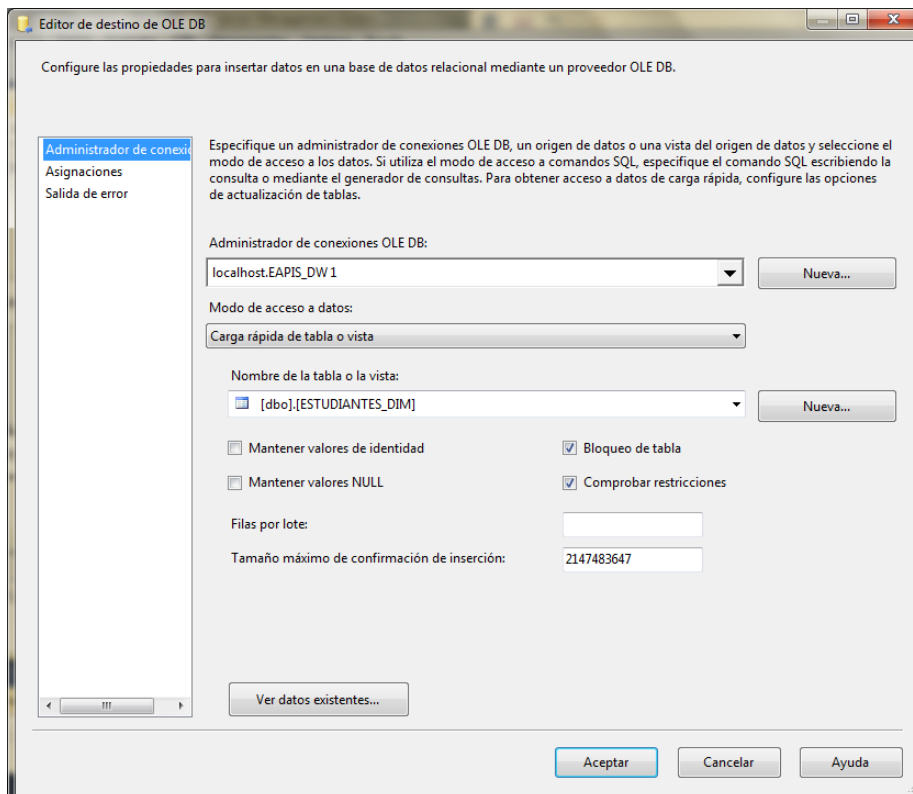


Figura 26. Destino OLE DB de "Poblar ESTUDIANTE_DIM"

Transformación

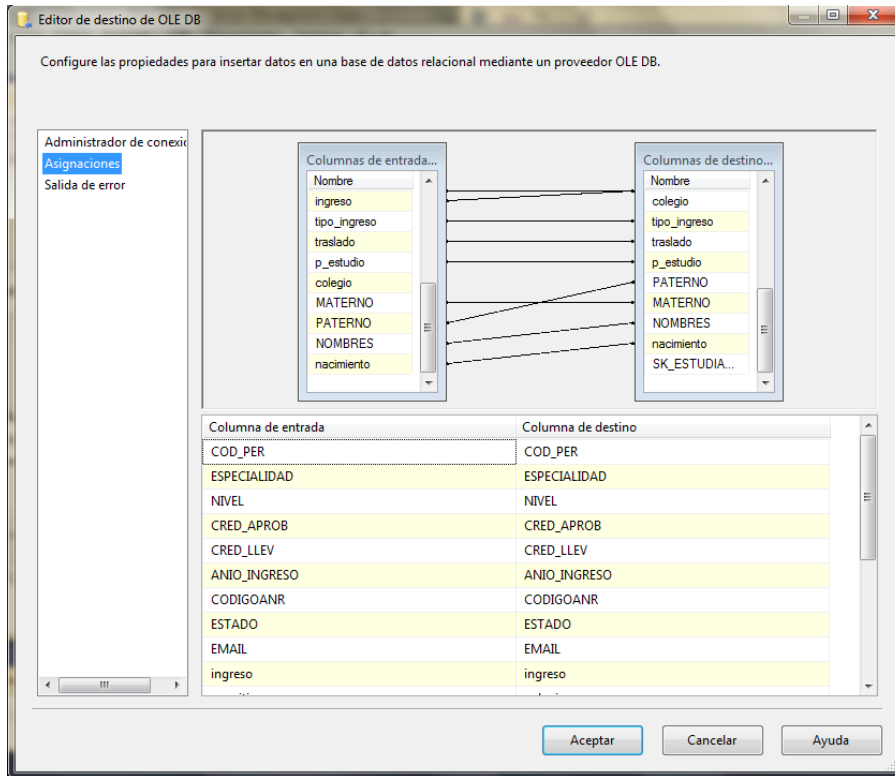


Figura 27. Mapping de "Poblar ESTUDIANTE_DIM"

3.1.4.2. Flujo de datos de CURSOS_DIM

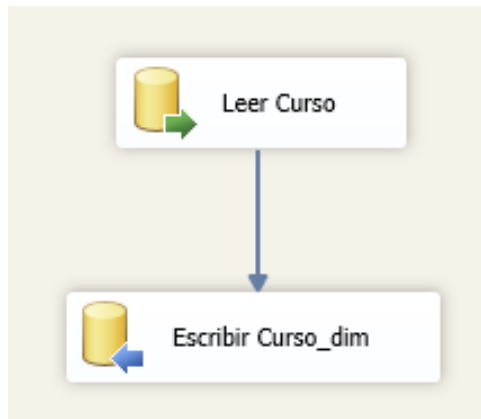


Figura 28. Flujo de datos de "Poblar CURSO_DIM"

Consulta

```
SELECT IDCURSO, NOMBRE, HTEO, HPRA, CRED, AREA,
NIVEL, SEMESTRE
FROM CURSO
```

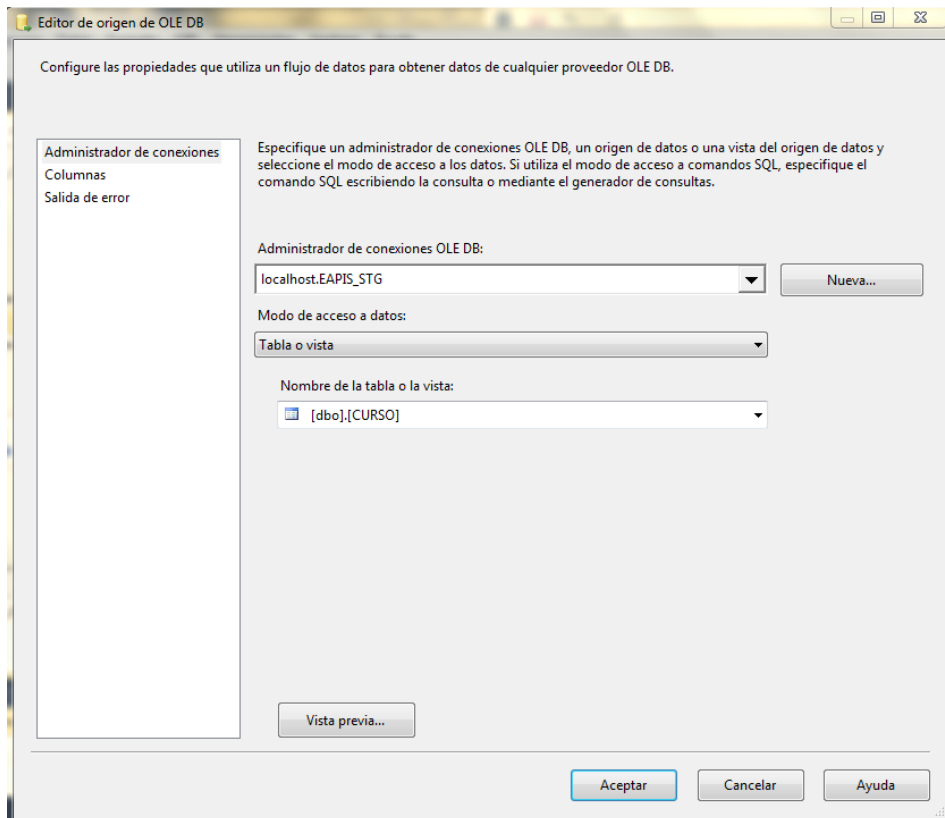


Figura 29. Origen OLE DB de "Poblar CURSO_DIM"

Destino Transaccional Mart [dbo].[CURSOS_DIM]

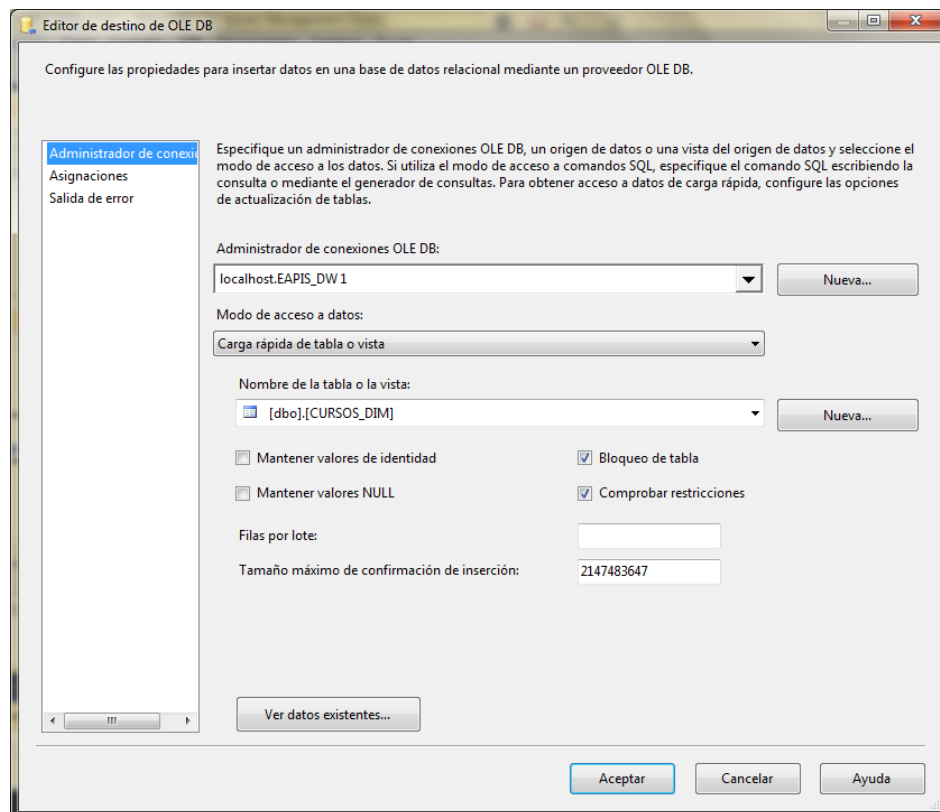


Figura 30. Destino OLE DB de "Poblar CURSO_DIM"

Transformación

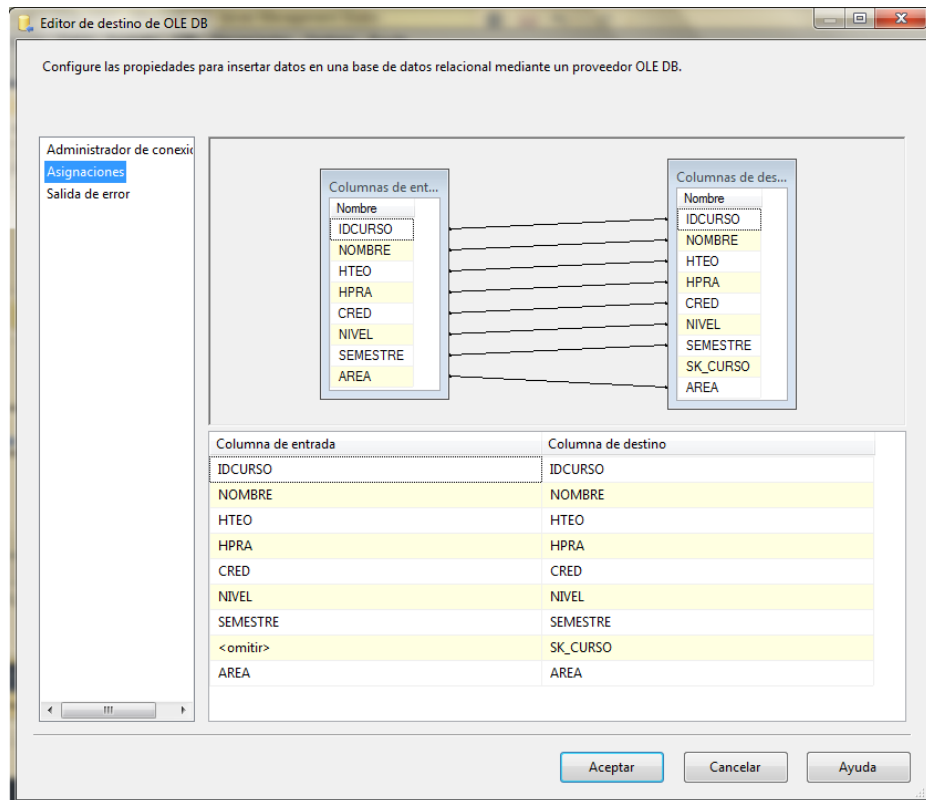


Figura 31. Mapping de "Poblar CURSO_DIM"

3.1.4.3. Flujo de datos de FACULTAD_DIM

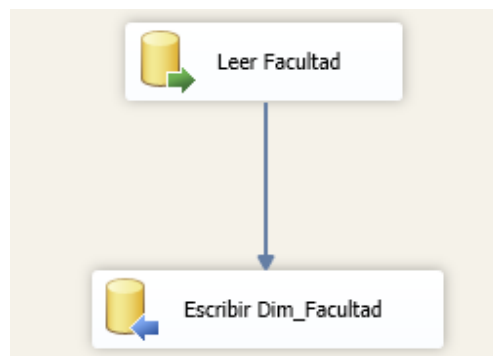


Figura 32. Flujo de datos de "Poblar FACULTAD_DIM"

Consulta
SELECT IDFACULTAD, FACULTAD FROM FACULTAD

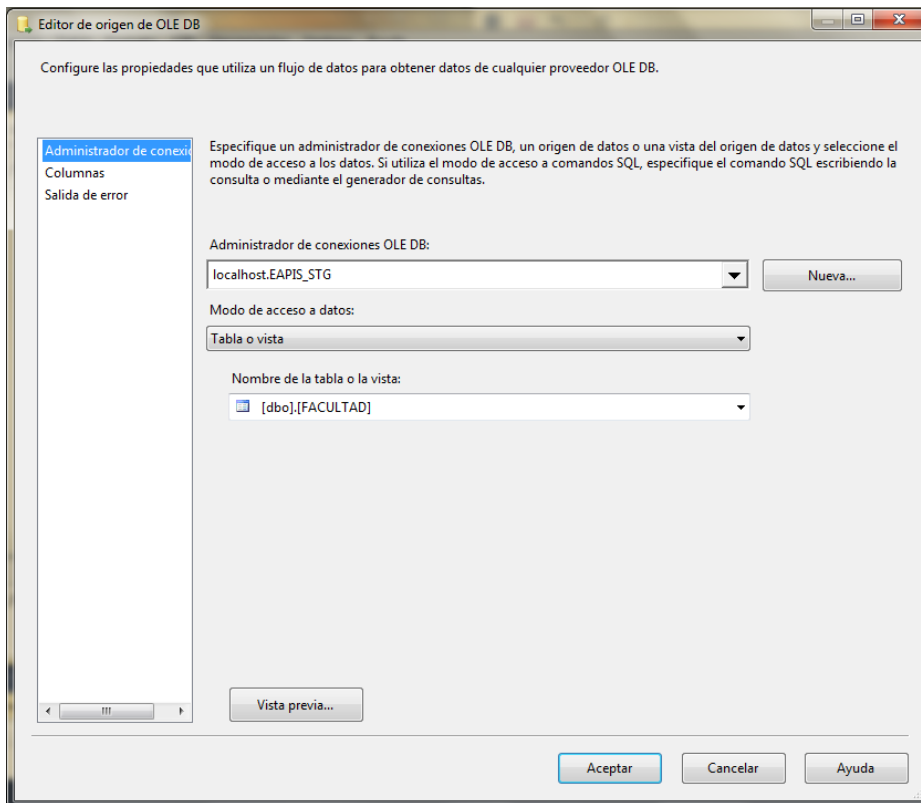


Figura 33. Origen OLE DB de “Poblar FACULTAD_DIM”

Destino TransaccionalMart

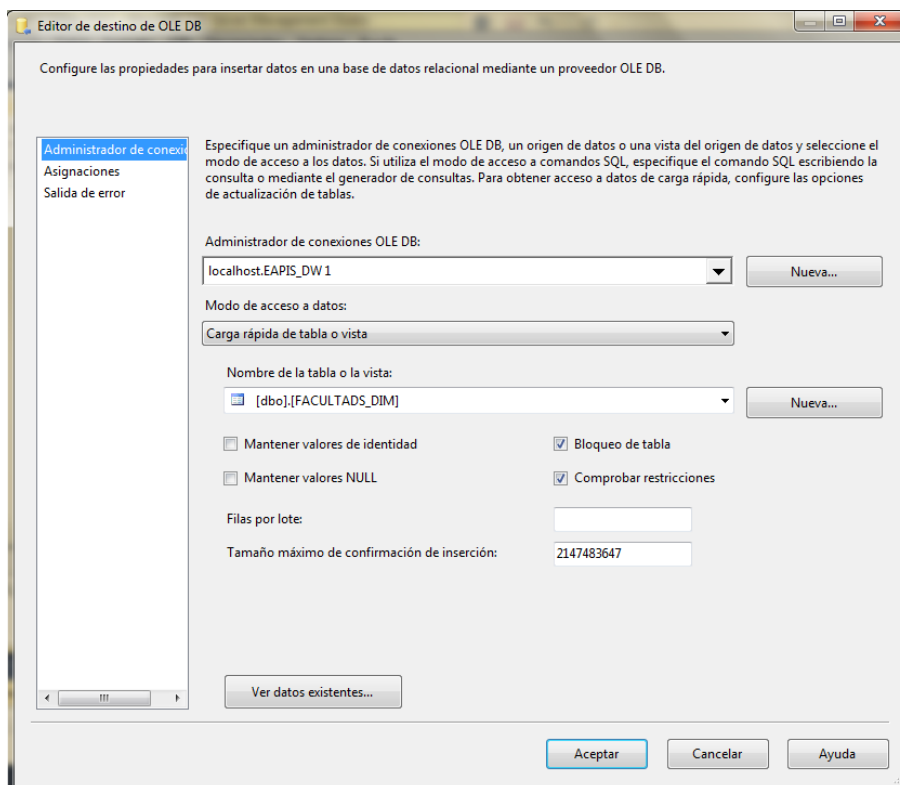


Figura 34. Origen OLE DB de “Poblar FACULTAD_DIM”

Transformación

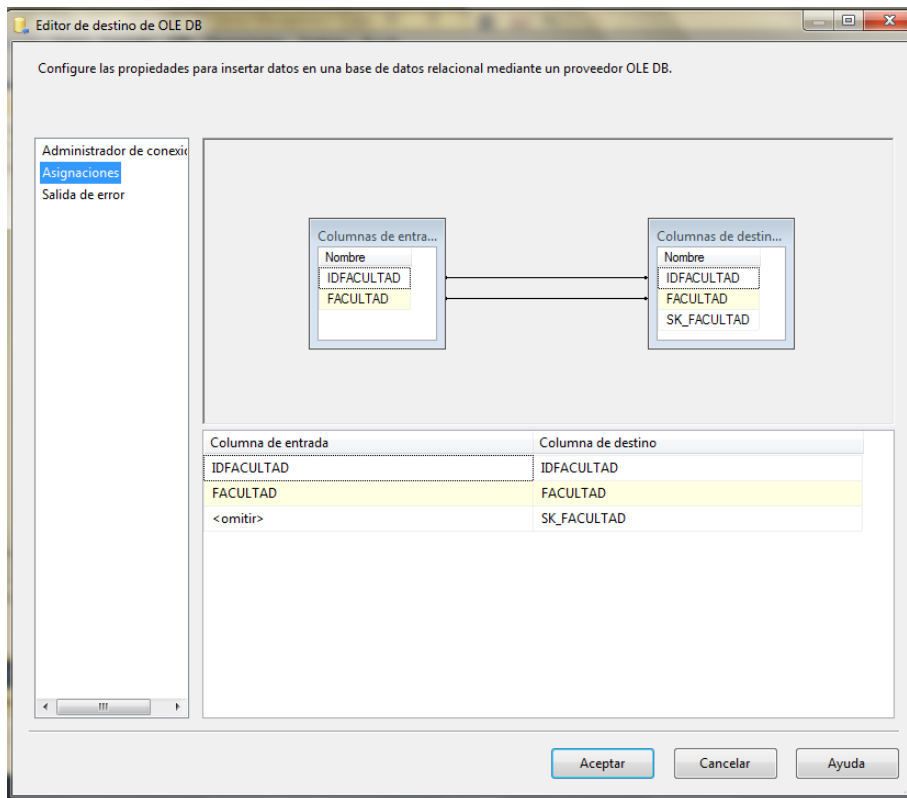


Figura 35. Mapping de "Poblar FACULTAD_DIM"

3.1.4.4. Flujo de datos de ESCUELA_DIM

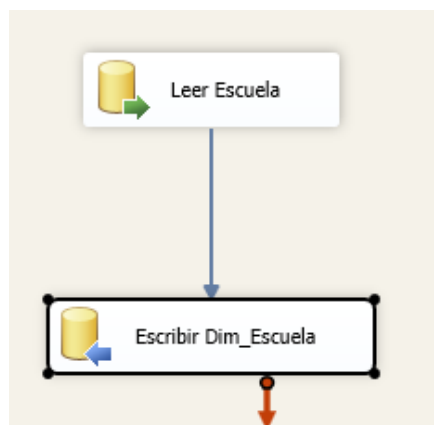


Figura 36. Flujo de datos de "Poblar ESCUELA_DIM"

Consulta

```
SELECT E.IDESCUELA, E.NOMBRE, F.IDFACULTAD  
FROM ESCUELA E INNER JOIN EAPIS_ODS..DIM_FACULTAD F  
ON E.IDFACULTAD=F.IDFACULTAD
```

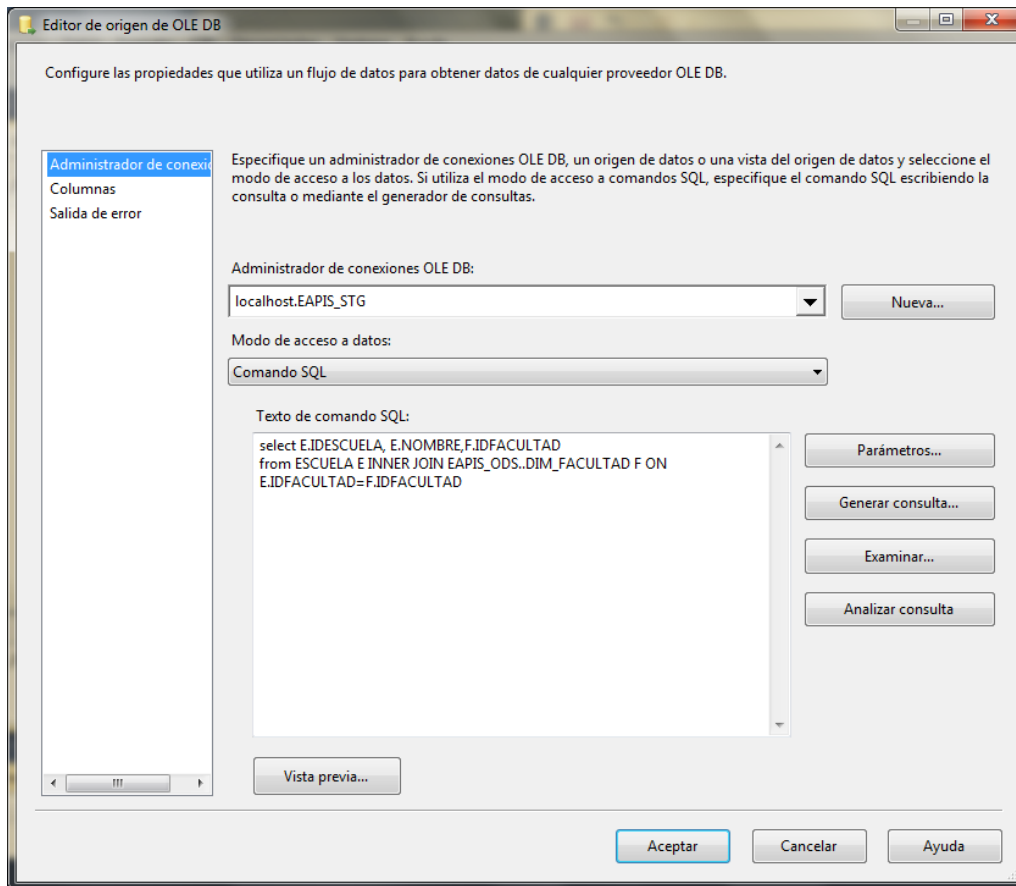


Figura 37. Origen OLE DB de "Poblar ESCUELA_DIM"

Destino TransaccionalMart

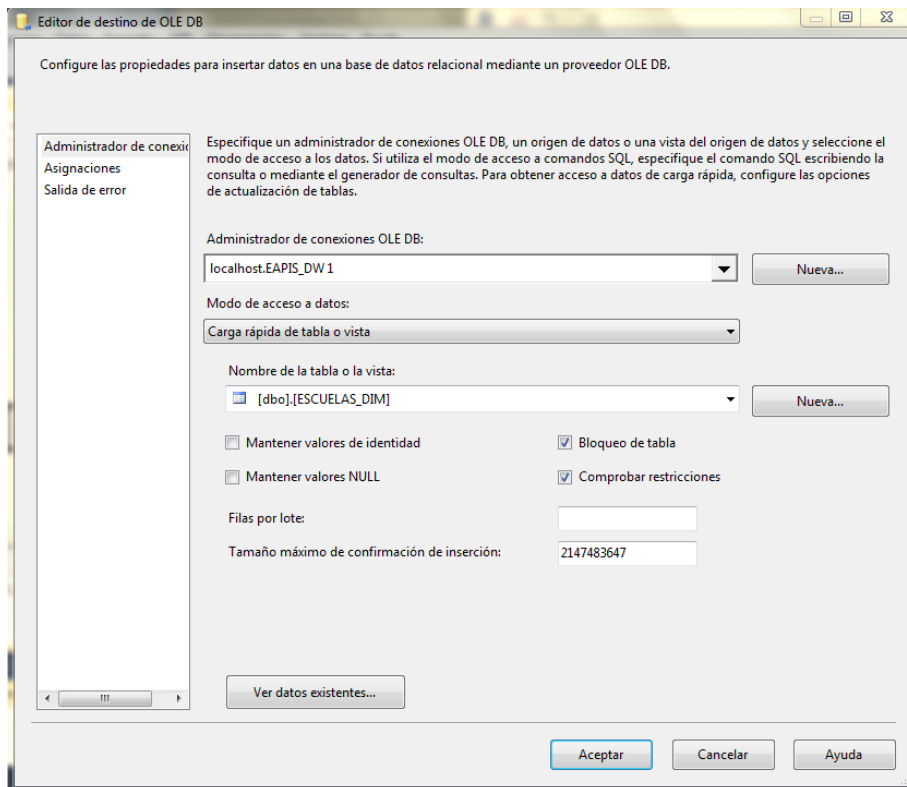


Figura 38. Destino OLE DB de "Poblar ESCUELA_DIM"

Transformación

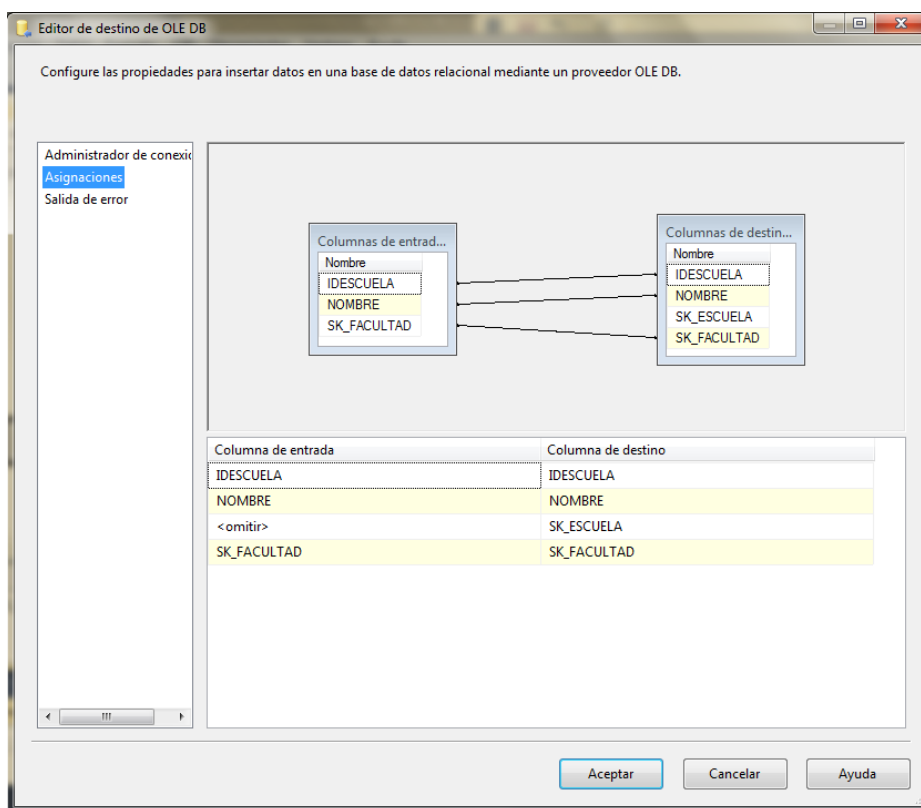


Figura 39. Mapping de "Poblar ESCUELA_DIM"

3.1.4.5. Flujo de datos de ESPECIALIDAD_DIM

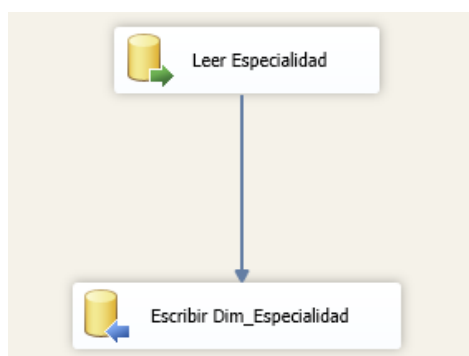


Figura 40. Flujo de datos de "Poblar ESPECIALIDAD_DIM"

Consulta

```
SELECT ESP.IDESPECIALIDAD, ESP.NOMBRE, ESP.IDSEDE,  
ESC.IDESCUELA  
FROM ESPECIALIDAD ESP INNER JOIN EAPIS_ODS..DIM_ESCUELA  
ESC ON ESP.IDESCUELA=ESC.IDESCUELA
```

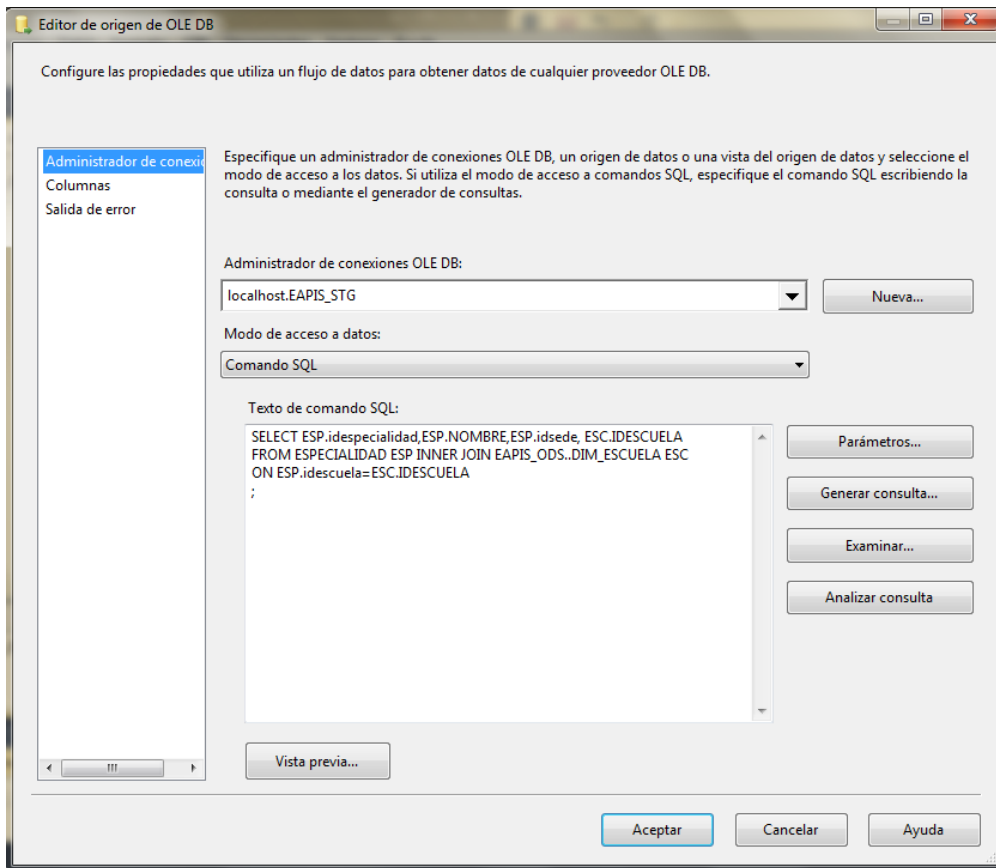



Figura 41. Origen OLE DB de "Poblar ESPECIALIDAD_DIM"

Destino TransaccionalMart

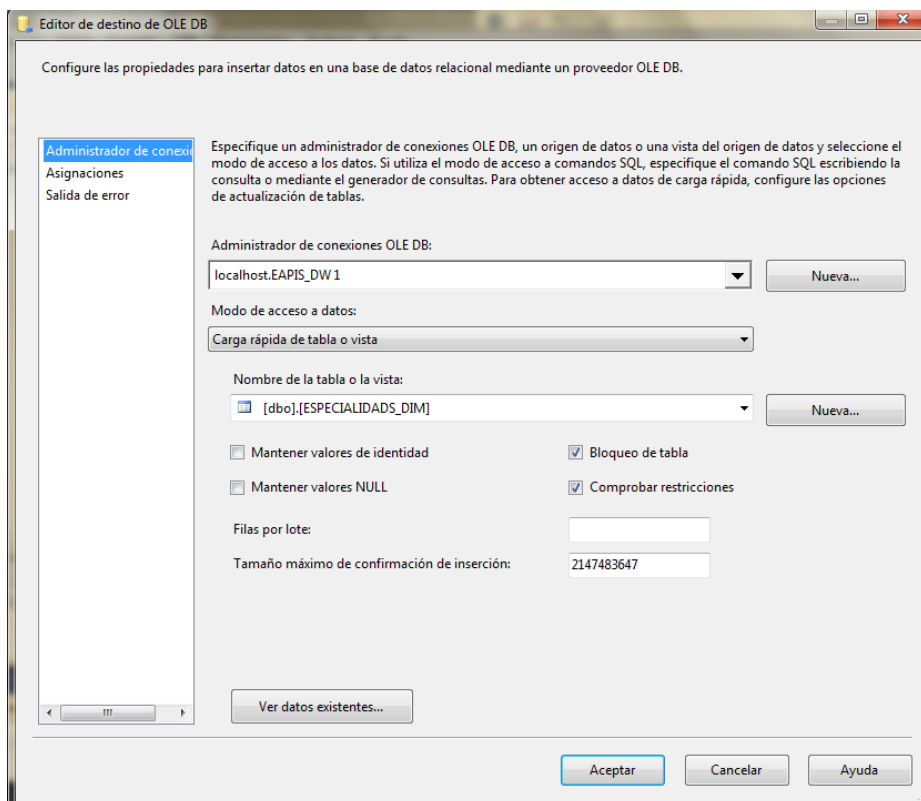


Figura 42. Destino OLE DB de "Poblar ESPECIALIDAD_DIM"

Transformación

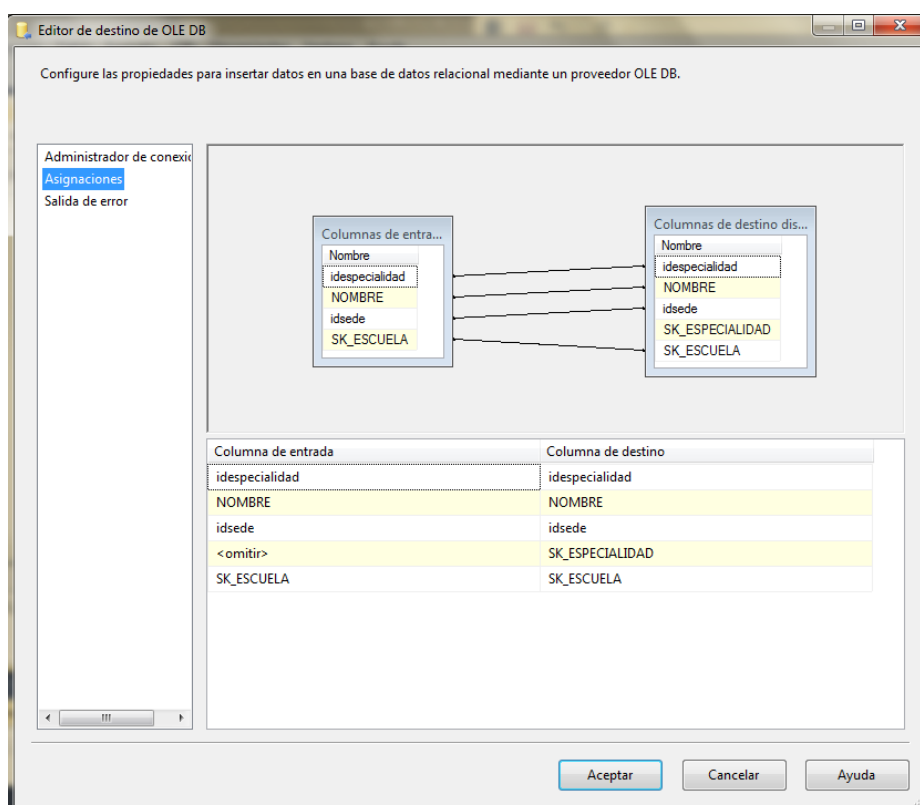


Figura 43. Mapping de "Poblar ESPECIALIDAD_DIM"

3.1.4.6. Flujo de datos de MATRICULA_DIM

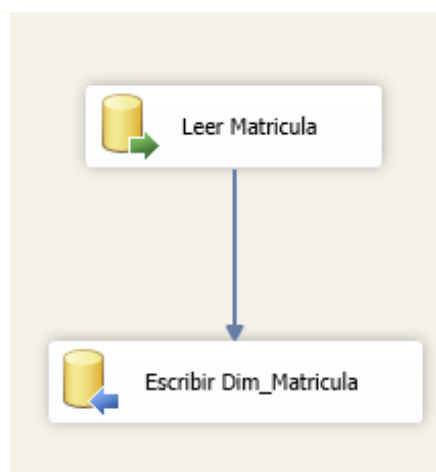


Figura 44. Flujo de datos de "Poblar MATRICULA_DIM"

Consulta

```
SELECT COD_PER+CUR_COD+PERIODO PK_MAT,  
COD_PER, CUR_COD, GRUP_NRO,FECHA,OBSERVACION,PERIODO  
FROM MATRICULA
```

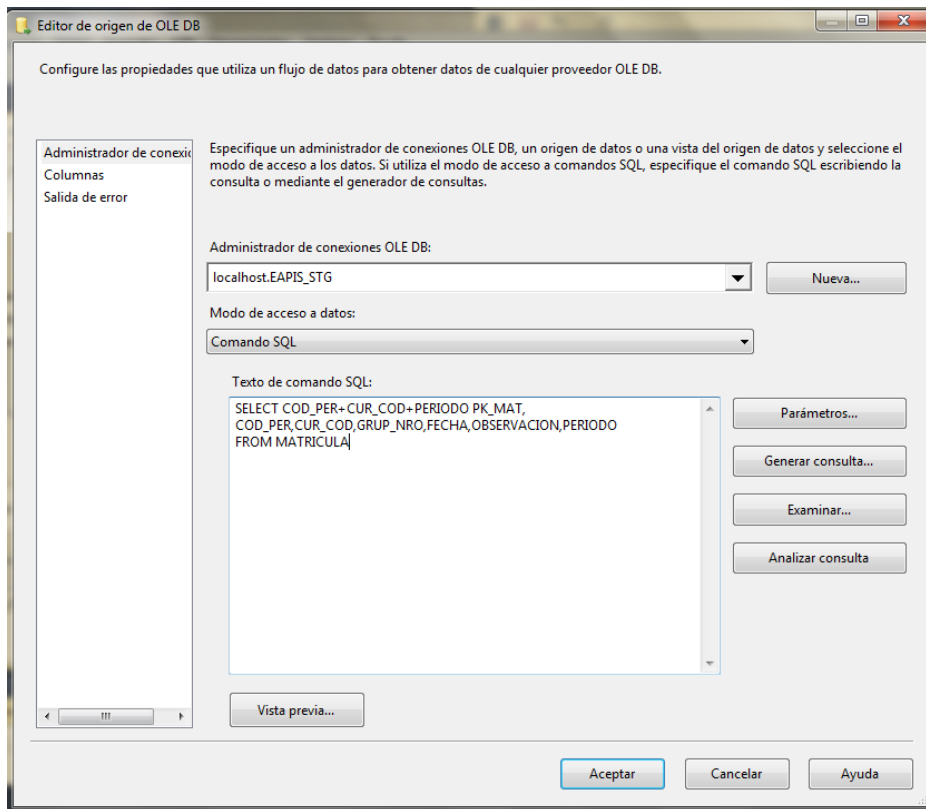


Figura 45. Origen OLE DB de "Poblar MATRICULA_DIM"

Destino TransaccionalMart

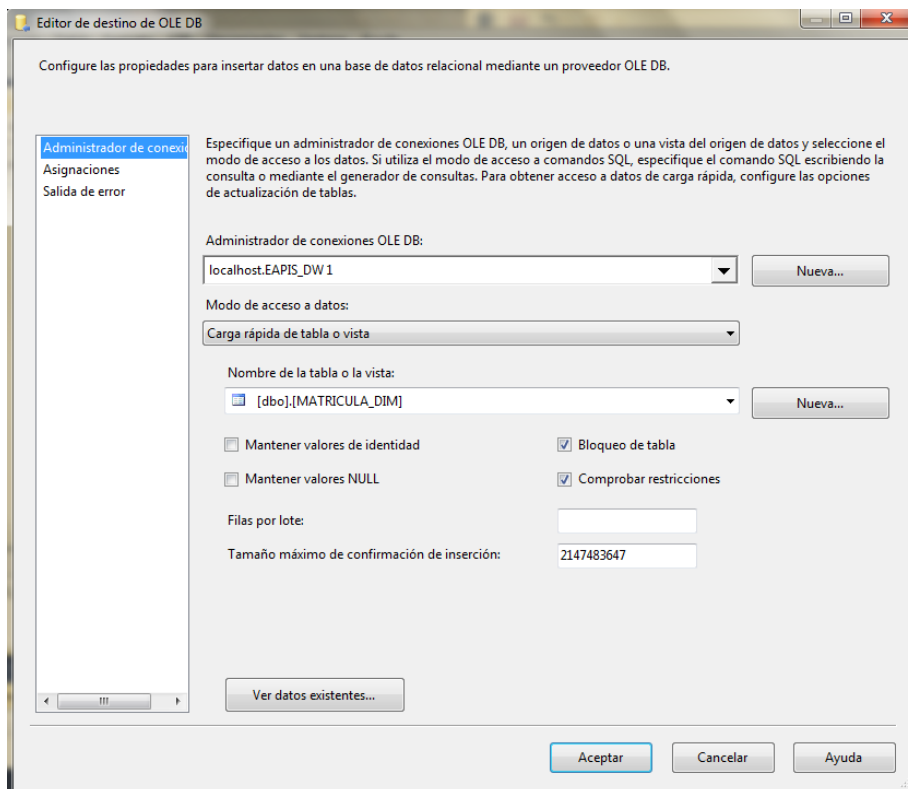


Figura 46. Destino OLE DB de "Poblar MATRICULA_DIM"

Transformación

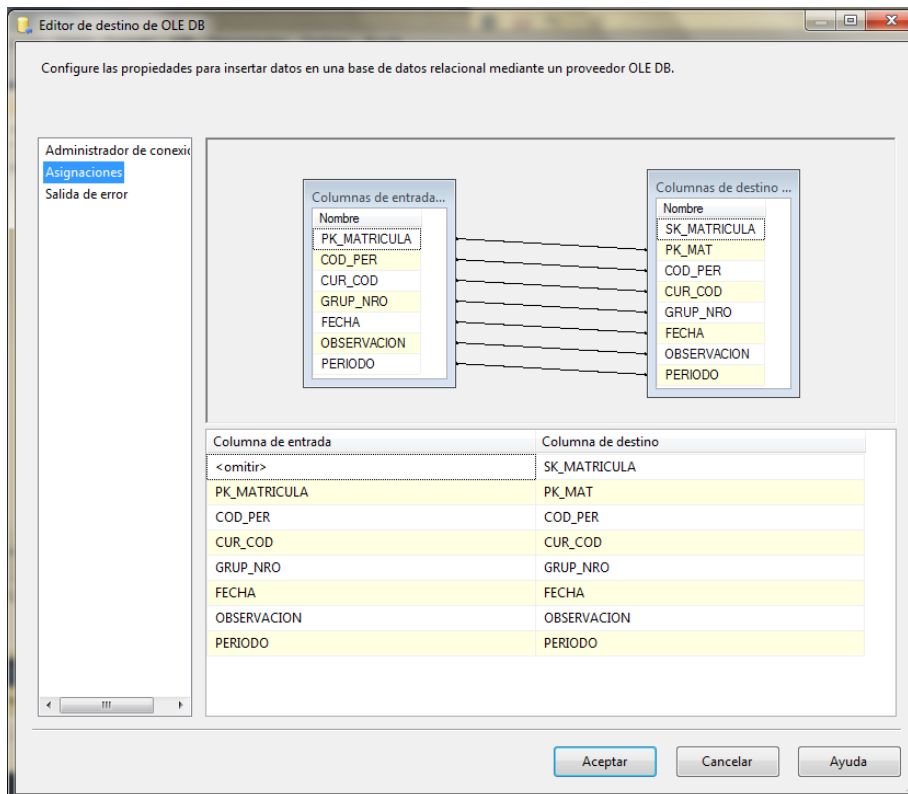


Figura 47. Mapping de "Poblar MATRICULA_DIM"

3.1.4.7. Flujo de datos de NOTACAB_DIM

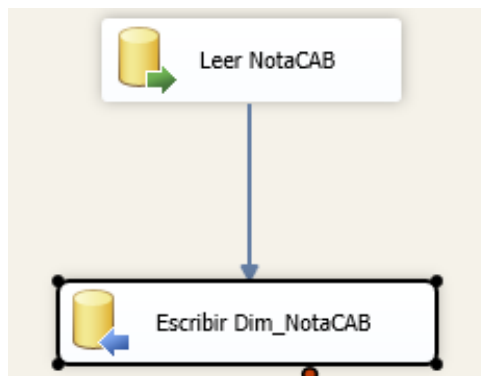


Figura 48. Flujo de datos de "Poblar NOTA_CA_DIM"

Consulta

```
SELECT * FROM NOTA_CAB
WHERE FECHA BETWEEN '20100101'AND '20151231'
```

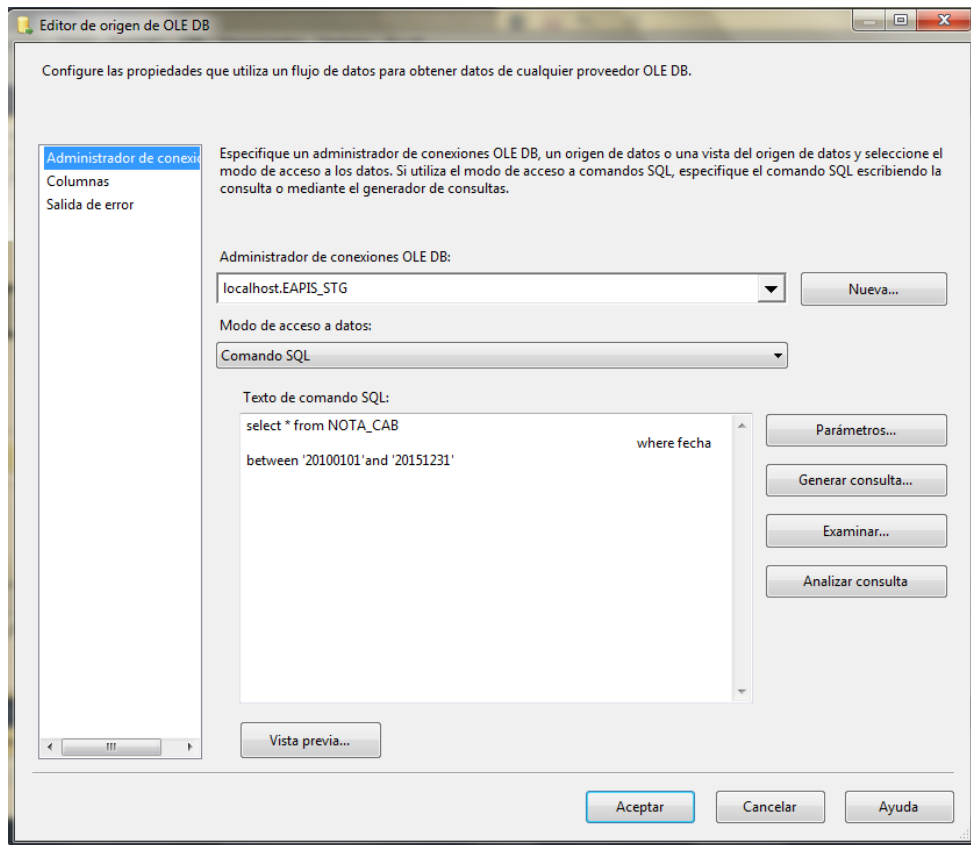


Figura 49. Origen OLE DB de "Poblar NOTA_CA_DIM"

Destino TransaccionalMart

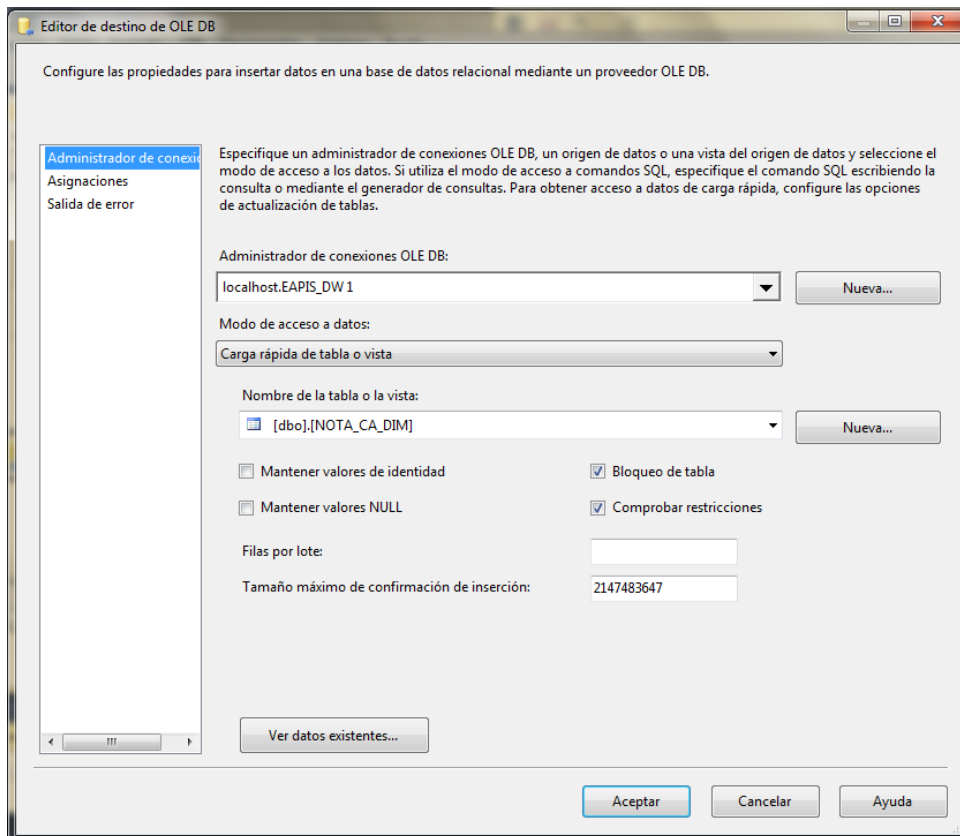


Figura 50. Destino OLE DB de "Poblar NOTA_CA_DIM"

Transformación

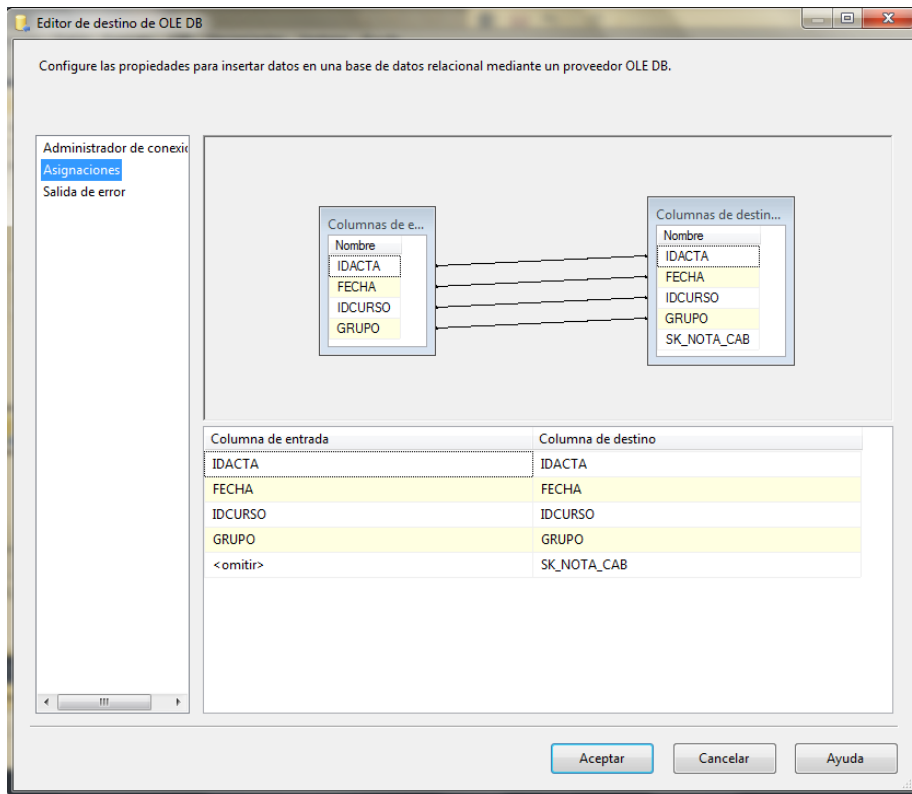


Figura 51. Mapping de "Poblar NOTA_CA_DIM"

3.1.4.8. Flujo de datos de ASIGNACION_DOCENTE_DIM

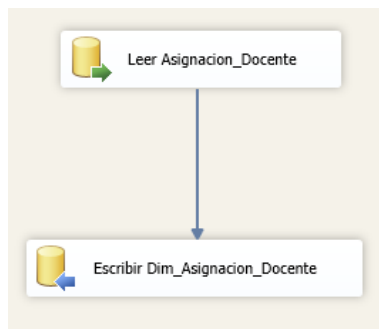


Figura 52. Flujo de datos de "Poblar ASIGNACION_DOCENTE_DIM"

Consulta

```

SELECT DISTINCT TU.ID_PERSONAL, TU.NOMBRES, U.APE_PATERNO,
TU.APE_MATERNO, AD.GRUPOCAR,AD.GRUPOP,AD.GRUPOT,
AD.HORASCAR,AD.HORASDES,AD.HPRACTICA, AD.HTEORIA,
AD.IDCURSO, AD.IDDOCDESC, AD.PERIODO, AD.IDCURSO + AD.IDDOC
+ AD.GRUPOCAR + AD.PERIODO AS PK, AD.IDDOC
FROM ASIGNACION_DOCENTE AD INNER JOIN TRABAJADORUNC TU
ON AD.IDDOC=TU.ID_PERSONAL
  
```

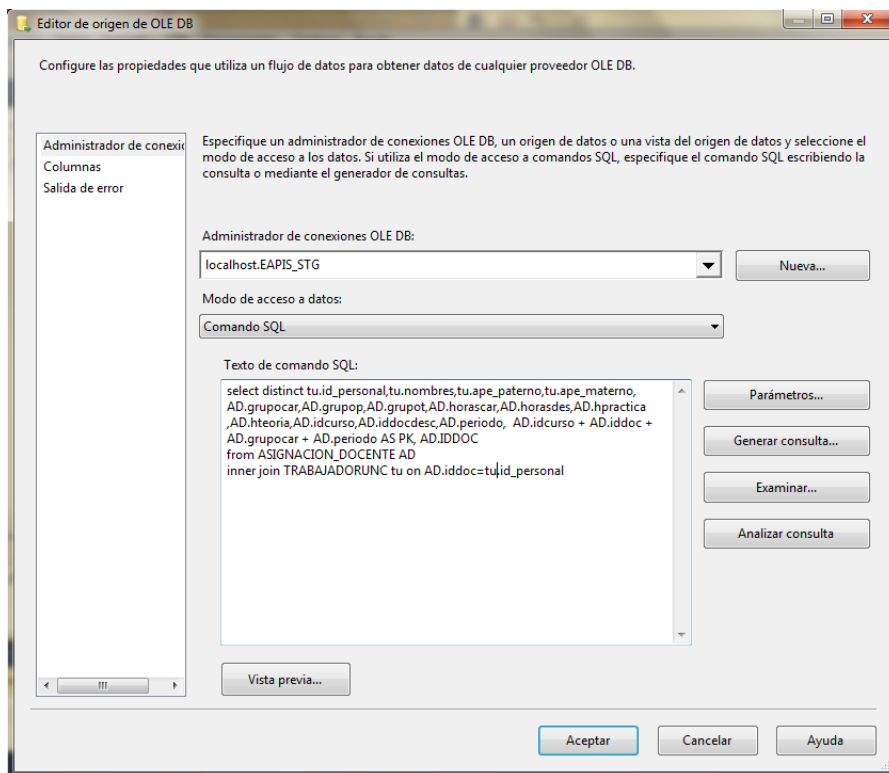


Figura 53. Origen OLE DB de "Poblar ASIGNACION_DOCENT_DIM"

Destino Transaccional Mart

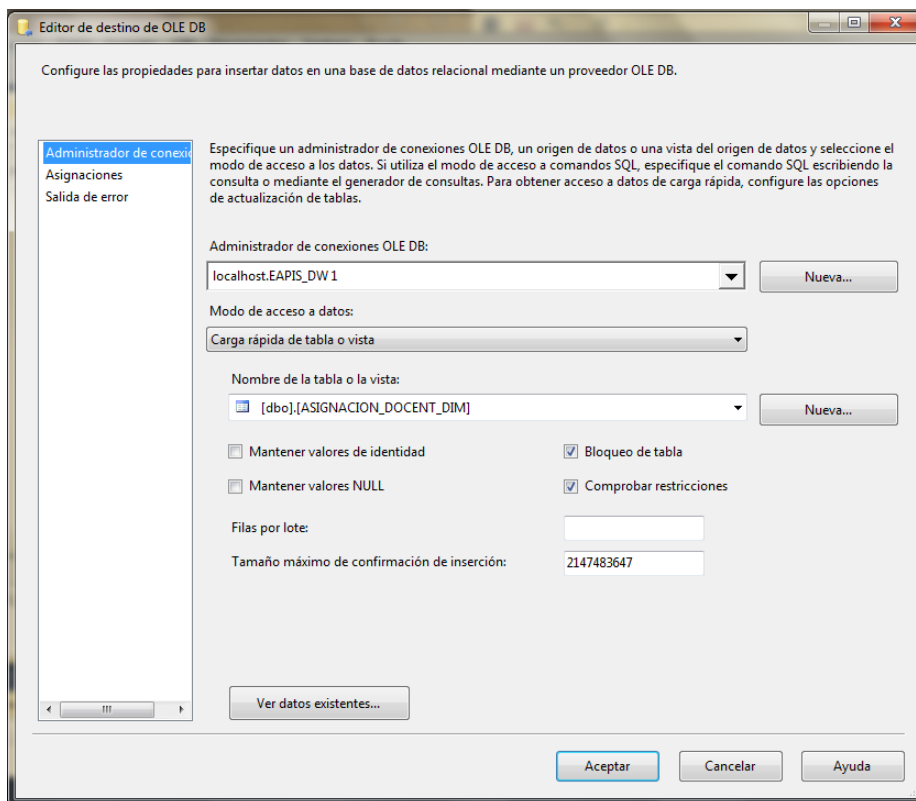


Figura 54. Destino OLE DB de "Poblar ASIGNACION_DOCENT_DIM"

Transformación

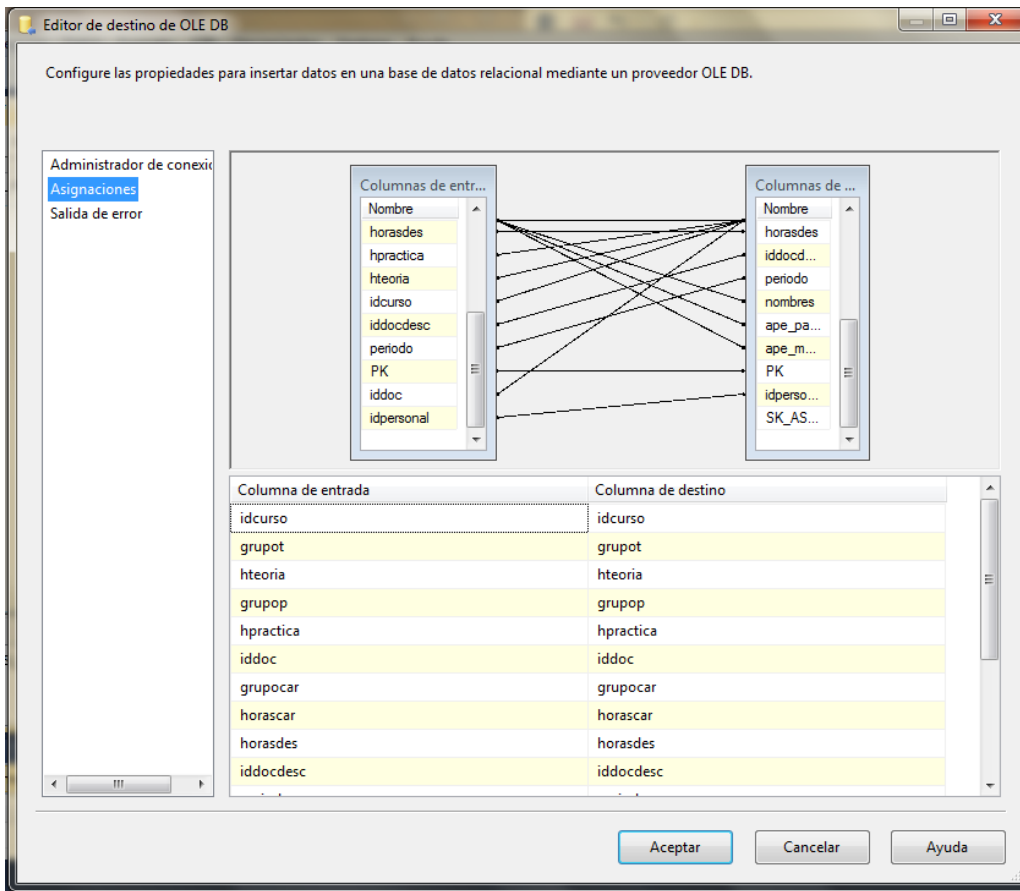


Figura 55. Mapping OLE DB de "Poblar ASIGNACION_DOCENT_DIM"

3.1.4.9. Flujo de datos de NOTAS_FACT

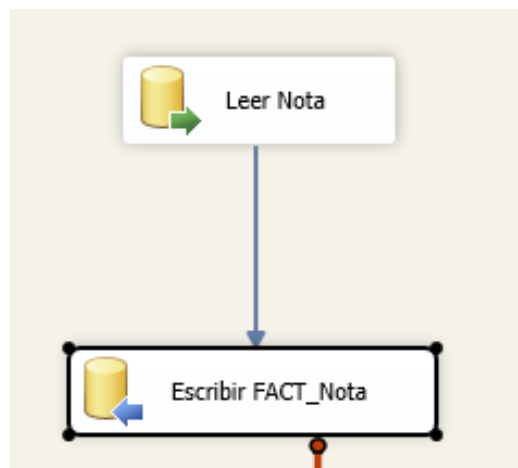


Figura 56. Flujo de datos de "Poblar NOTA_FACT"

Consultar

```
SELECT DISTINCT EST.COD_PER, AD.IDCURSO + AD.IDDOC +  
AD.GRUPOCAR + AD.PERIODO AS PK, ESP.IDESPECIALIDAD,  
CUR.IDCURSO, NC.IDACTA,  
MAT.COD_PER+MAT.CUR_COD+MAT.PERIODO AS PK_MAT, N.NOTA  
FROM      PERSONA AS PER INNER JOIN  
ESTUDIANTE AS EST ON PER.CODIGO = EST.COD_PER INNER JOIN  
MATRICULA AS MAT ON MAT.COD_PER = EST.COD_PER INNER JOIN  
NOTAS AS N ON N.IDCURSO = MAT.CUR_COD AND MAT.COD_PER =  
N.CODIGOUNC INNER JOIN  
CURSO AS CUR ON CUR.IDCURSO = N.IDCURSO INNER JOIN  
ASIGNACION_DOCENTE AS AD ON AD.IDCURSO = CUR.IDCURSO AND  
MAT.PERIODO = AD.PERIODO AND (MAT.GRUP_NRO = AD.GRUPOT OR  
MAT.GRUP_NRO = AD.GRUPOP OR  
MAT.GRUP_NRO = AD.GRUPOCAR) INNER JOIN  
TRABAJADORUNC AS TRA ON TRA.ID_PERSONAL = AD.IDDOC INNER  
JOIN NOTA_CAB AS NC ON NC.IDACTA = N.IDACTA AND NC.IDCURSO =  
N.IDCURSO AND CUR.IDCURSO = NC.IDCURSO INNER JOIN  
ESPECIALIDAD AS ESP ON ESP.IDESPECIALIDAD = EST.ESPECIALIDAD  
WHERE     EST.ANIO_INGRESO BETWEEN 10 AND 15 AND NC.FECHA  
BETWEEN '20100101'AND '20151231'AND (N.NOTA <> '00')
```

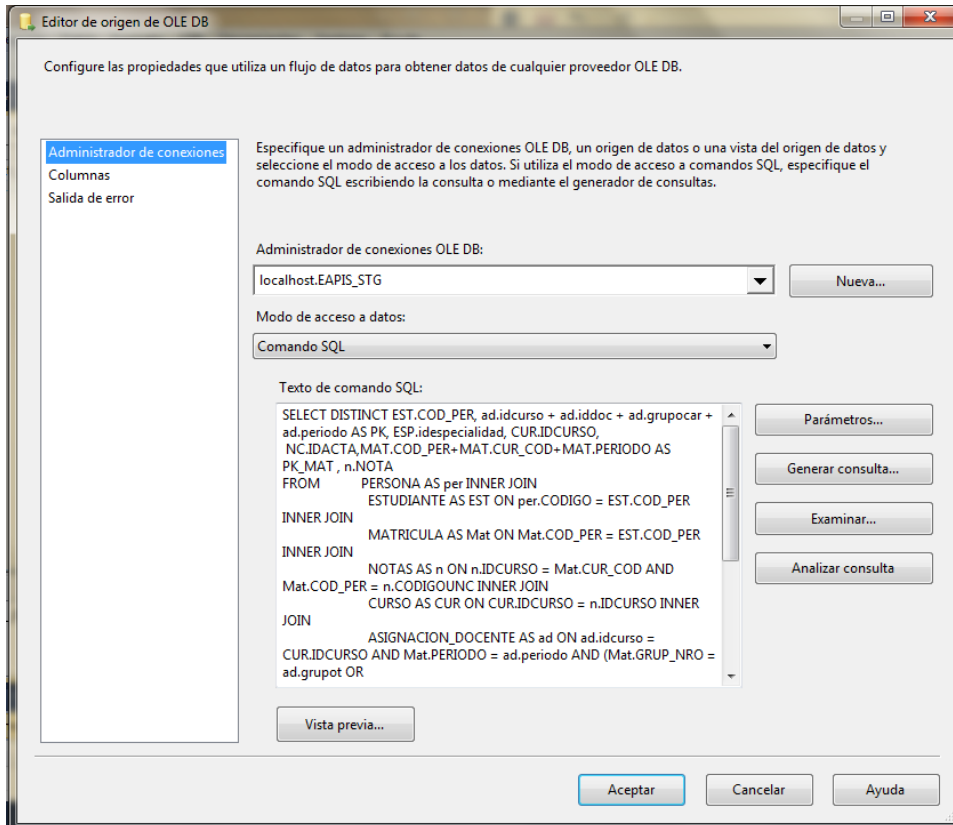


Figura 57. Origen OLE DB de "Poblar NOTA_FACT"

Destino TransaccionalMart

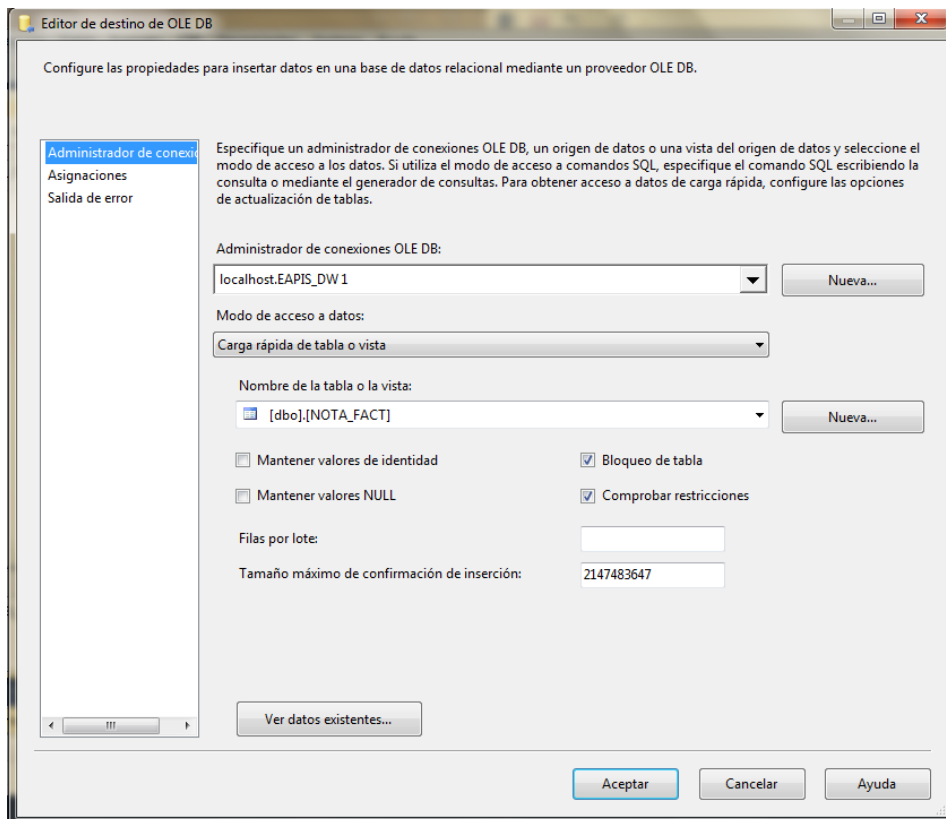


Figura 58. Destino OLE DB de "Poblar NOTA_FACT"

Transformación

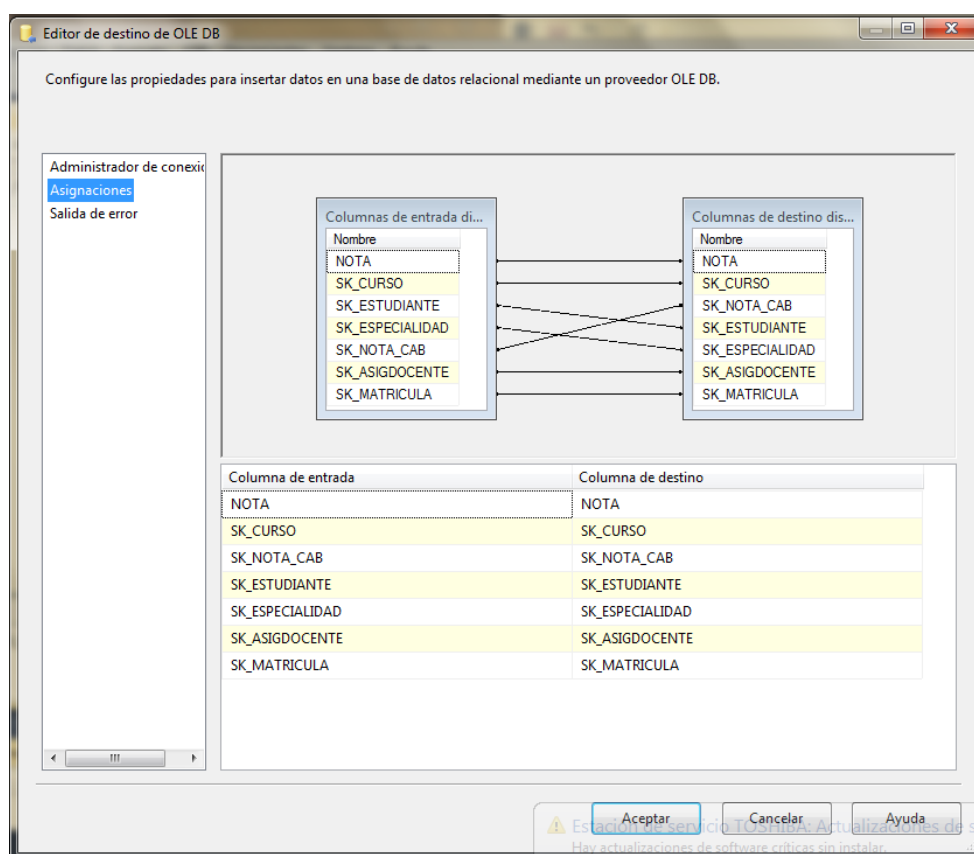


Figura 59. Mapping de "Poblar NOTA_FACT"

3.1.5. Selección del producto

La Unidad Técnica de Sistemas Informáticos de la Universidad Nacional de Cajamarca cuenta con herramientas de inteligencia de negocios de Microsoft y Oracle.

Debido a la familiaridad con que cuenta la Unidad Técnica de Sistemas Informáticos con el uso de estas herramientas de inteligencia de negocios de Microsoft (Integration Services, Analysis Services, SQL Server) y las características de hardware con las que se cuenta, se utilizaron dichas herramientas para el desarrollo e implementación del proyecto.

3.1.6. Implementación

3.1.6.1. Desarrollo de cubo

❖ ESTUDIANTE_DIM

Atributos

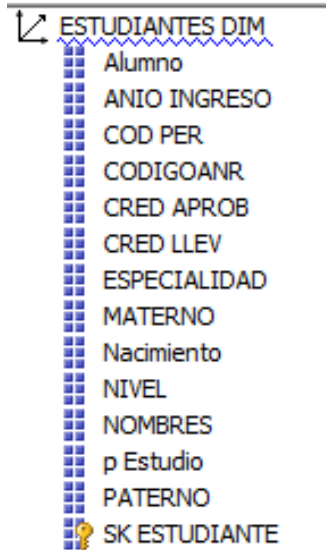


Figura 60. Atributos de dimensión "ESTUDIANTES_DIM"

Vista de orígenes de datos

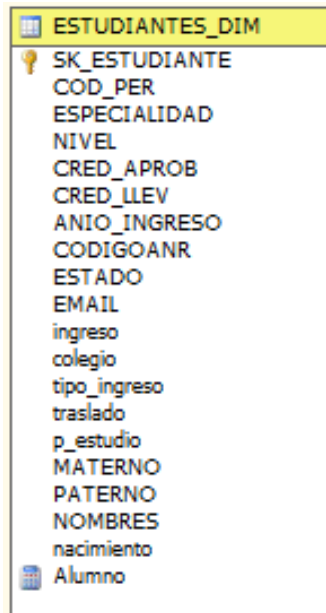


Figura 61. Vista de orígenes de datos de "ESTUDIANTES_DIM"

❖ CURSO_DIM

Atributos

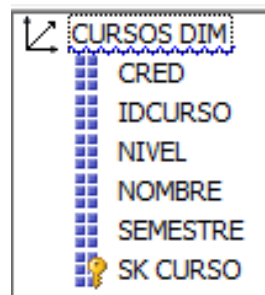


Figura 62. Atributos de dimensión "CURSOS_DIM"

Vista de orígenes de datos

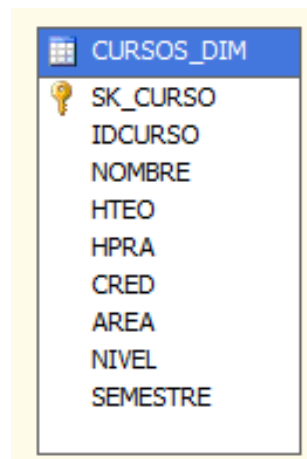


Figura 63. Vista de orígenes de datos de "CURSOS_DIM"

❖ ESPECIALIDAD_DIM

Atributos

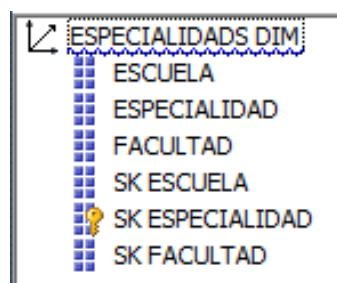


Figura 64. Atributos de dimensión "ESPECIALIDADES_DIM"

Vista de orígenes de datos

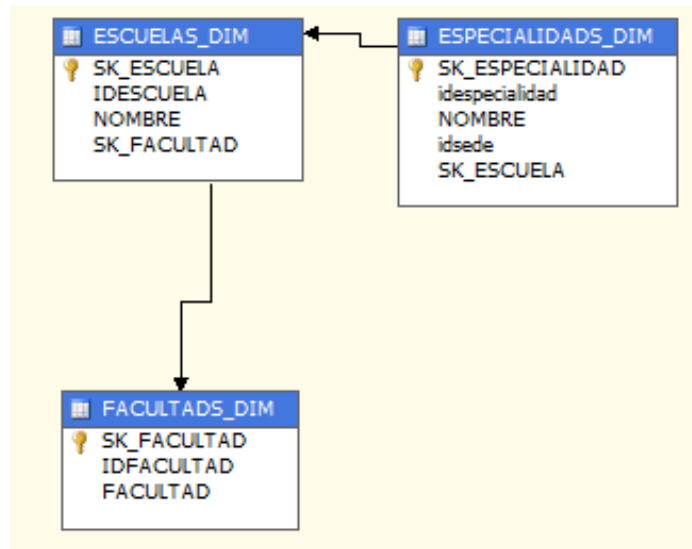


Figura 65. Vista de orígenes de datos de "ESPECIALIDADES_DIM"

❖ MATRICULA_DIM

Atributos



Figura 66. Atributos de dimensión "MATRICULA_DIM"

Vista de orígenes de datos

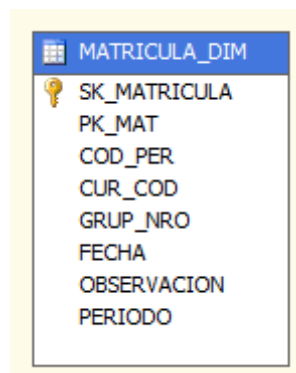


Figura 67. Vista de orígenes de datos de "MATRICULA_DIM"

❖ NOTA_CA_DIM

Atributos

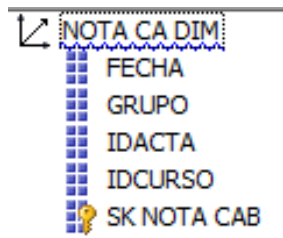


Figura 68. Atributos de dimensión "NOTA_CA_DIM"

Vista de orígenes de datos

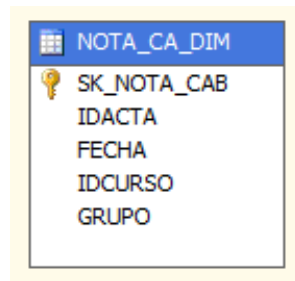


Figura 69. Vista de orígenes de datos de "NOTA_CA_DIM"

❖ ASIGNACION_DOCENTE_DIM

Atributos

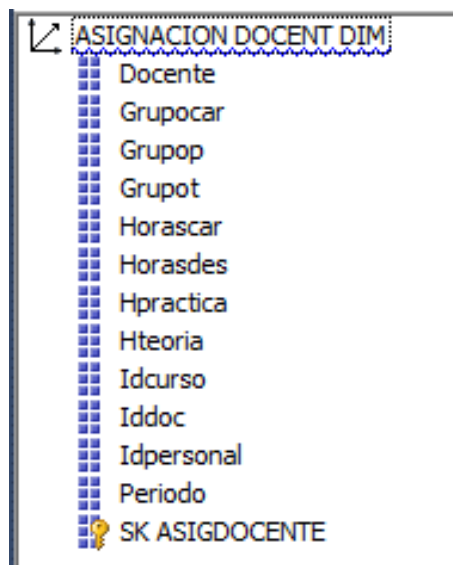


Figura 70. Atributos de dimensión "ASIGNACION_DOCENTE_DIM"

Vista de orígenes de datos

ASIGNACION_DOCENT_DIM	
SK_ASIGDOCENTE	
idcurso	
grupot	
hteoria	
grupop	
hpractica	
iddoc	
grupocar	
horascar	
horasdes	
iddocdesc	
periodo	
PK	
idpersonal	
nombres	
ape_paterno	
ape_materno	
Docente	

Figura 71. Vista de orígenes de datos de "ASIGNACION_DOCENT_DIM"

❖ NOTA_FACT

Atributos

NOTA_FACT	
NOTA	
NOTA_FACT - SK ESPECIALIDAD	
SK ASIGDOCENTE	
SK CURSO	
SK ESPECIALIDAD	
SK ESTUDIANTE	
SK MATRICULA	
SK NOTA CAB	

Figura 72. Atributos de dimensión "NOTA_FACT"

Vista de orígenes de datos

NOTA_FACT	
SK_ESPECIALIDAD	
SK_NOTA_CAB	
SK_CURSO	
SK_ASIGDOCENTE	
SK_ESTUDIANTE	
SK_MATRICULA	
NOTA	

Figura 73. Vista de orígenes de datos de "NOTA_FACT"

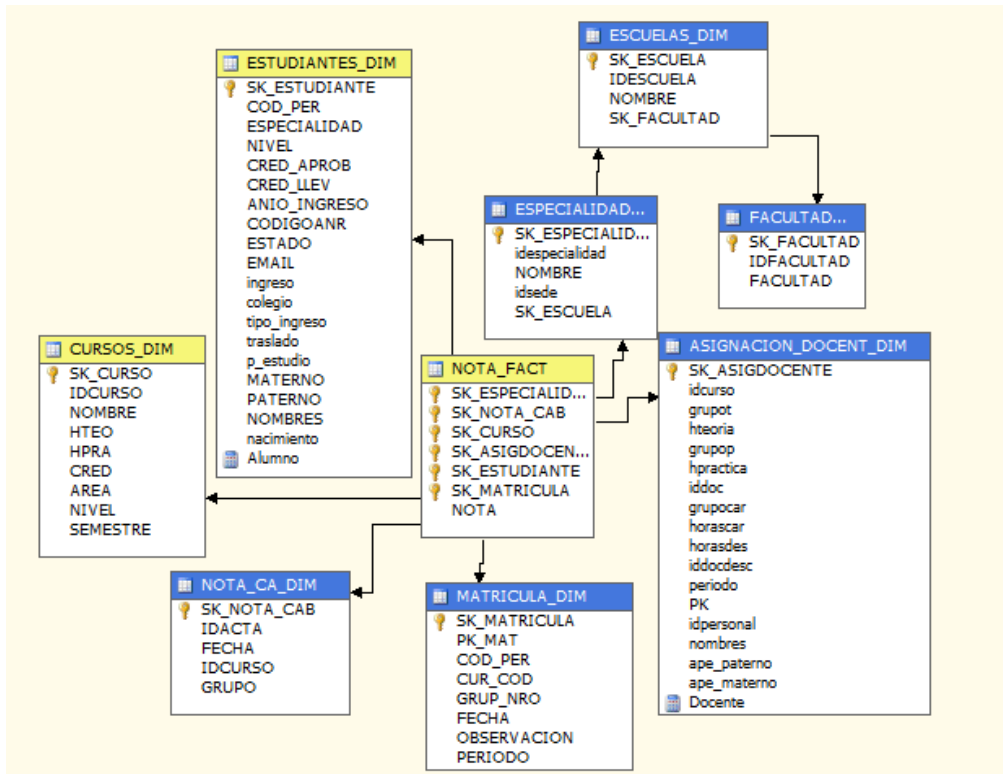


Figura 74. Vista de orígenes de datos de la solución de inteligencia de negocios

3.1.7. Especificación de Aplicaciones de BI

Descripción: el reporte mostrado responde al requerimiento R-03[Número de estudiantes por año de ingreso]

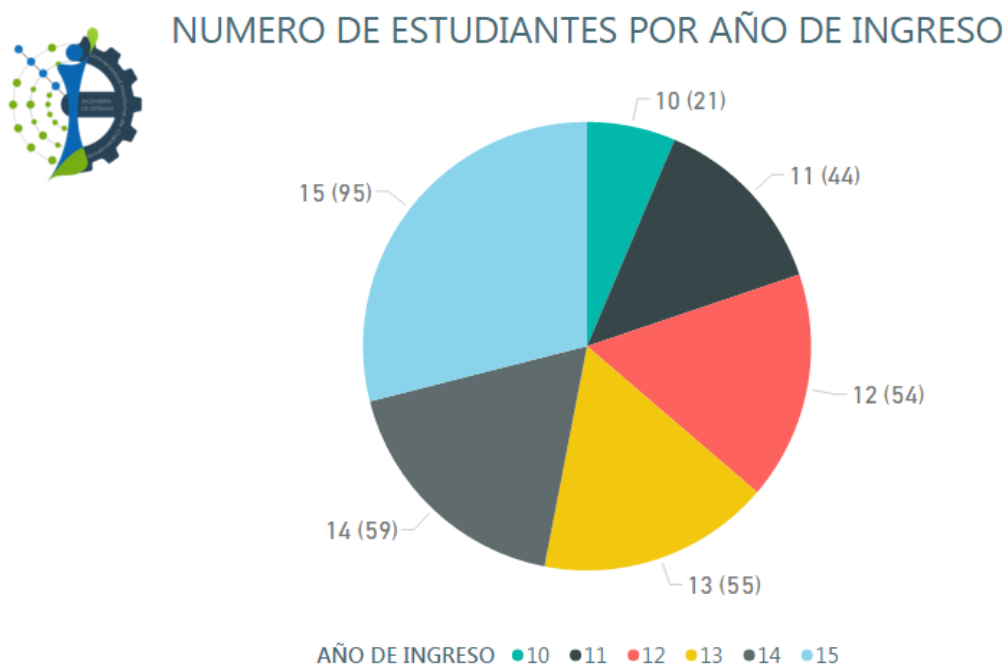


Figura 75. Reporte Número de estudiantes por año de ingreso

Descripción: el reporte mostrado responde al requerimiento R-04[Número de veces de llevar un curso]

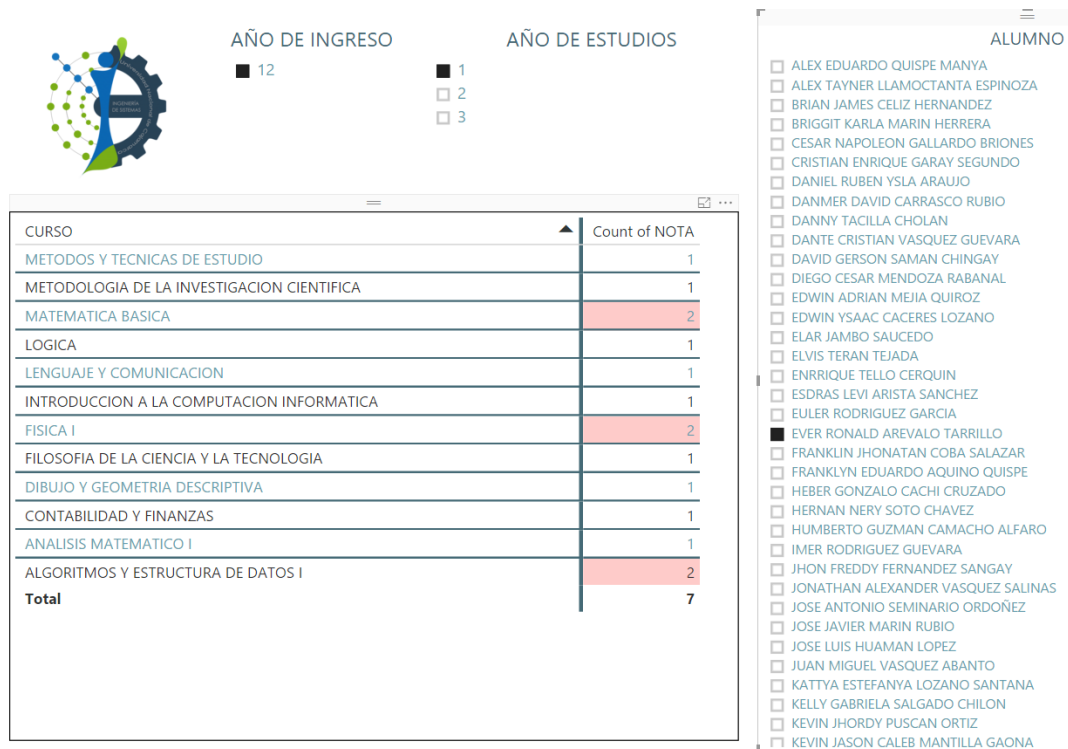


Figura 76. Reporte Número de veces de llevar el curso

Descripción: el reporte mostrado responde al requerimiento R-05 [Número de alumnos aprobados por curso]

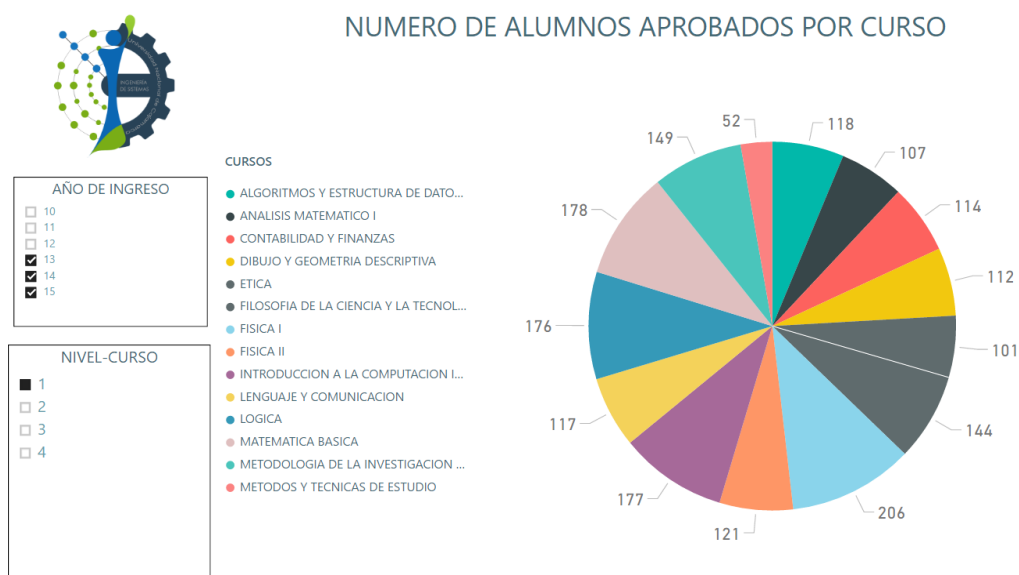


Figura 77. Reporte Número de alumnos aprobados por curso

Descripción: el reporte mostrado responde al requerimiento R-06[Número de alumnos desaprobados por curso]

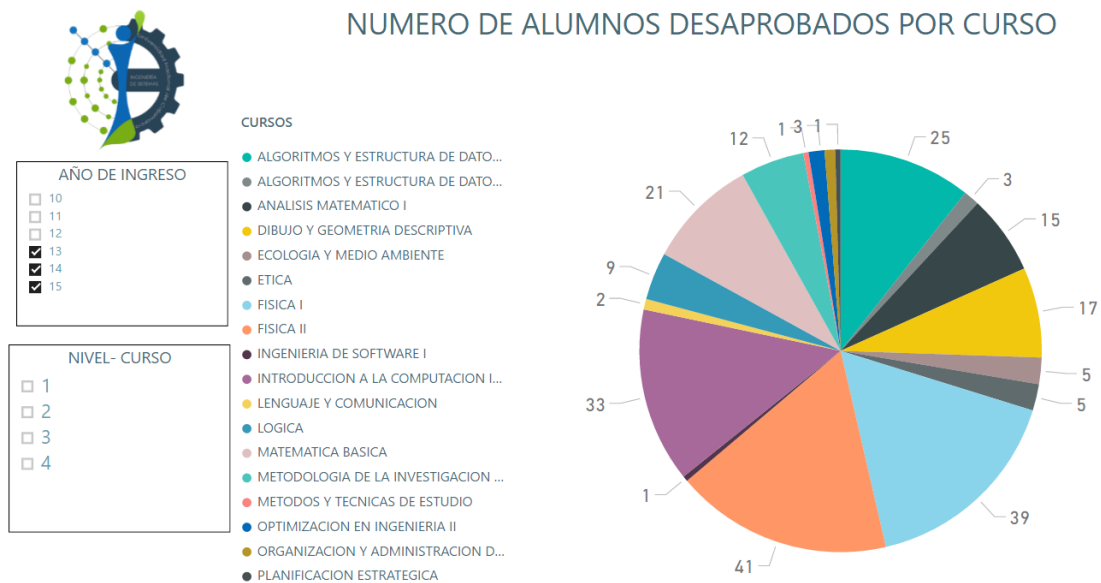


Figura 78. Reporte Número de alumnos desaprobados por curso

Descripción: el reporte mostrado responde al requerimiento R-07[Estado de estudiante en cada curso]

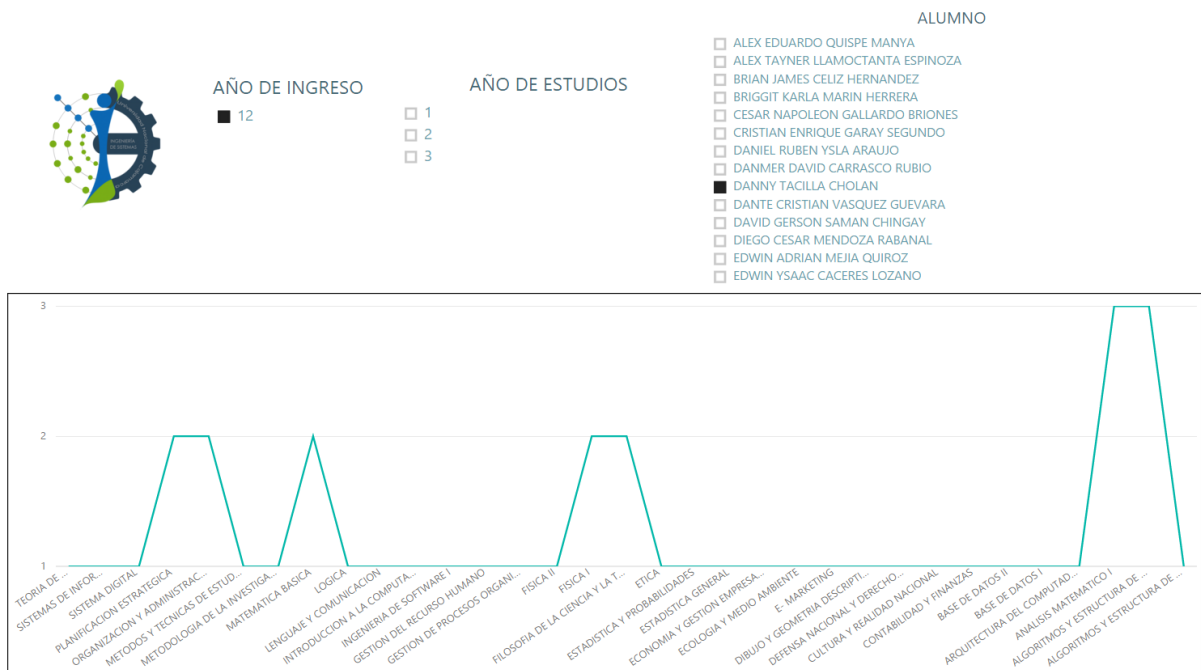


Figura 79. Estado de Estudiante

Descripción: el reporte mostrado responde al requerimiento R-08 [Número de alumnos por nota y curso]

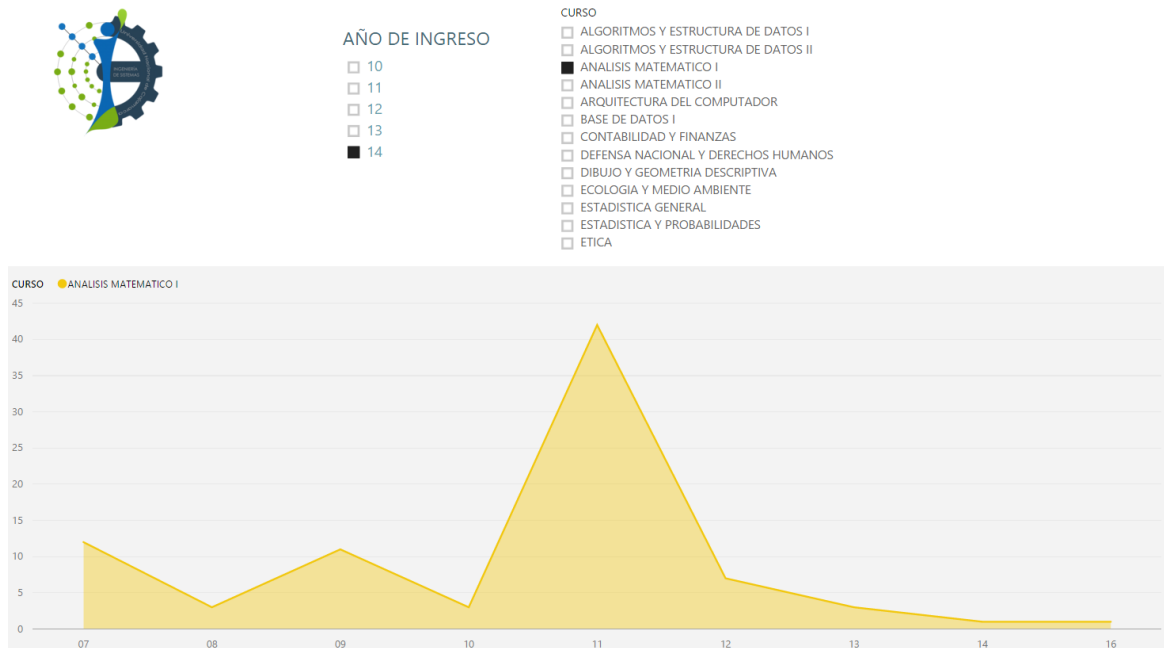


Figura 80. Número de alumnos por nota y curso

Descripción: el reporte mostrado responde al requerimiento R-09 [Consolidado de alumnos]



Figura 81. Consolidado de alumnos

Descripción: el reporte mostrado responde al requerimiento R-10 [Estado de alumnos por docente]



Figura 82. Estado de alumnos por docente

3.1.8. Diseño de la Arquitectura Técnica

La arquitectura presentada está basada en una de las mejores prácticas en datawarehousing. Está compuesta por tres capas a través de las cuales se garantiza la integridad de la información sumada con limpieza de datos. Para la extracción se dispuso de una fuente de datos SQL SEVER.

3.1.8.1. Capa STAGE

Es un área temporal donde se recogen los datos que se necesitan de los sistemas origen. Se recogen los datos estrictamente necesarios para las cargas, y se aplica el mínimo de transformaciones a los mismos. No se aplican restricciones de integridad ni se utilizan claves, los datos se tratan como si las tablas fueran ficheros planos. De esta manera se minimiza la afectación a los sistemas origen, la carga es lo más rápida posible para minimizar la ventana horaria necesaria, y se reduce también al mínimo la posibilidad de error. Una vez que los datos están traspasados, el DWH se independiza de los sistemas origen hasta la siguiente carga. Lo único que se suele añadir es algún campo que almacene la fecha de la carga.

Obviamente estos datos no van a dar servicio a ninguna aplicación de reporting, son datos temporales que una vez hayan cumplido su función

serán eliminados, de hecho en el esquema lógico de la arquitectura muchas veces no aparece, ya que su función es meramente operativa. [17]

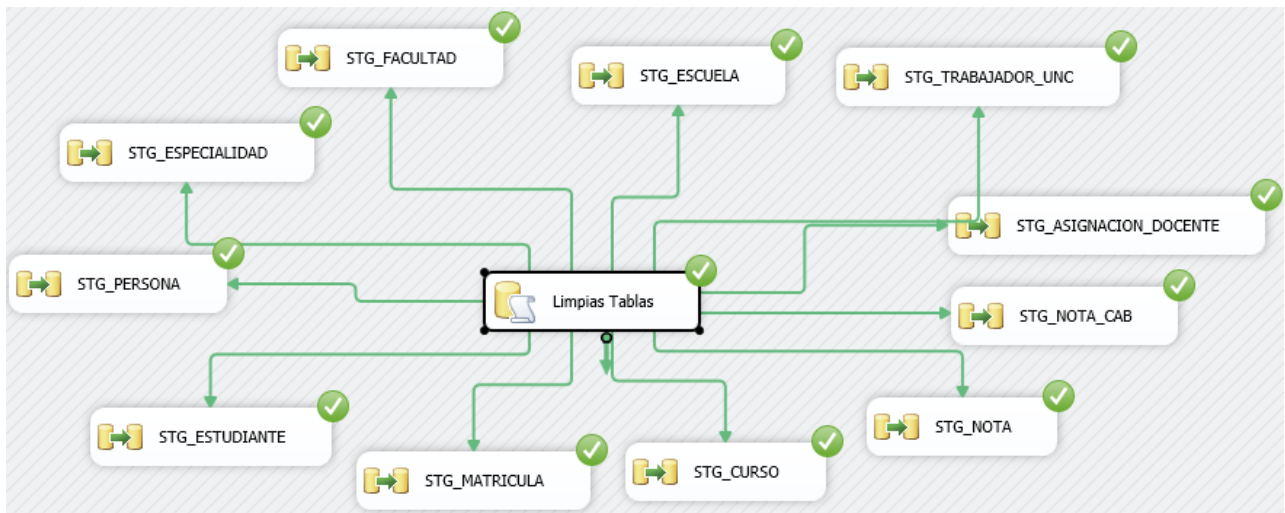


Figura 83. Ejecución de recojo de datos

3.1.8.2. Capa ODS (Operational Data Store)

Los datos del ODS se recogen de la Stage Area, y aquí sí que se realizan transformaciones, limpieza de datos y controles de integridad referencial para que los datos estén perfectamente integrados en el modelo relacional normalizado

Hay que tener en cuenta que la actualización de los datos del ODS no va a ser instantánea, los cambios en los datos de los sistemas origen no se verán reflejados hasta que finalice la carga correspondiente. Es decir, que se irán actualizando los datos cada cierto tiempo, cosa que hay que explicar a los usuarios, porque los informes que se lancen contra el ODS casi nunca podrán estar tan 'al minuto' como los que existan en el sistema origen. Lo que sí se puede hacer es definir una mayor frecuencia de carga para el ODS que para el Almacén Corporativo. Si es necesario, se puede refrescar el ODS cada 15 minutos, y el resto cada día, por ejemplo. [17]

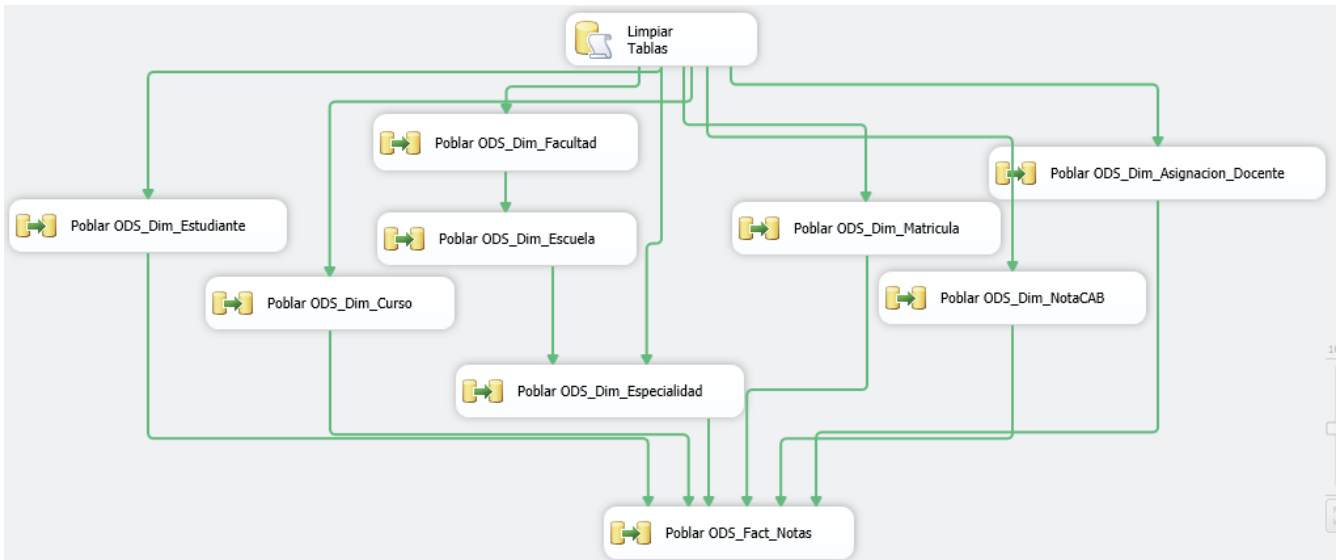


Figura 84. Ejecución ODS

3.1.8.3. Capa DW

Los Data Mart se diseñan con estructura multidimensional, cada objeto de análisis es una tabla de hechos enlazada con diversas tablas de dimensiones. Si se diseñan siguiendo el Modelo en Estrella habrá prácticamente una tabla para cada dimensión, es la versión más desnormalizada.

Este área puede residir en la misma base de datos que las demás si la herramienta de explotación es de tipo ROLAP, o también puede crearse ya fuera de la BD, en la estructura de datos propia que generan las aplicaciones de tipo MOLAP, más conocida como los cubos multidimensionales. [17]

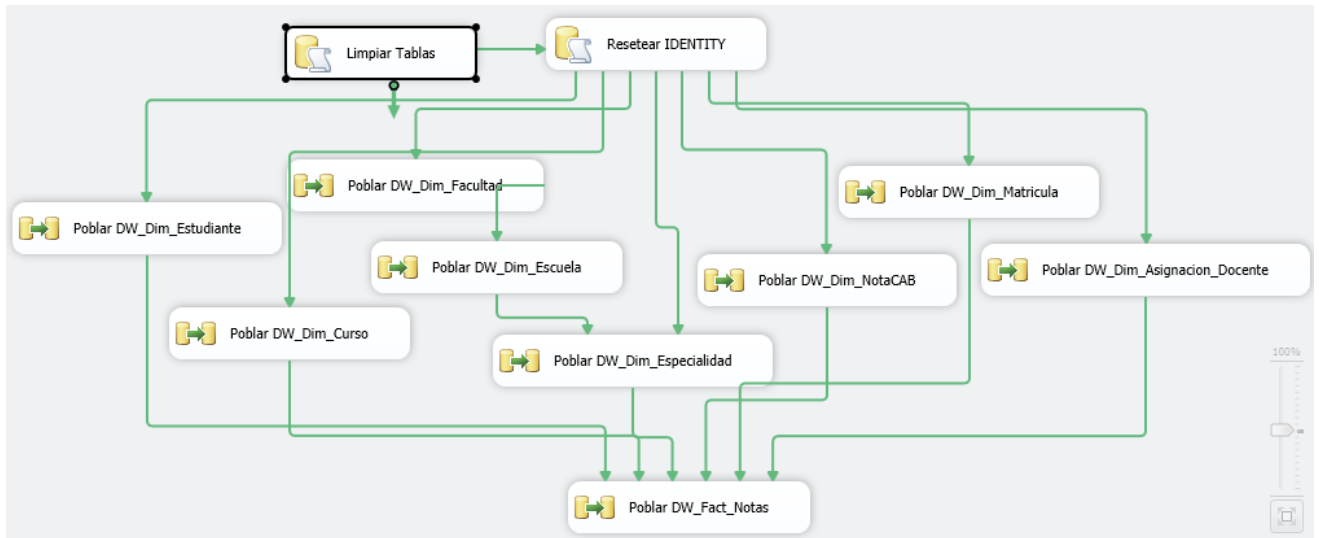


Figura 85. Ejecución DW

3.2. Tratamiento, análisis de datos y presentación de resultados

A continuación se presentan los resultados de la investigación en base a la información recogida mediante las técnicas e instrumentos de recolección de datos, las que se muestran mediante cuadros y gráficos de acuerdo a la hipótesis de trabajo. Para evaluar la variable independiente se aplicó una encuesta a cada uno de los miembros actuales del comité de la Escuela Académico Profesional de Ingeniería de Sistemas, en total fueron tres los encuestados, obteniendo el siguiente resultado:

❖ Resultado de la entrevista realizada

En el ANEXO 1 podemos ver el modelo de entrevista realizada para conocer el procedimiento e información utilizada en la EAPIS. Donde sí es 1, no es 2 y no sabe 3.

Tabla 38. Resultado de Entrevista

	Cuenta c/reportes	acceso directo	acceso total	horas extras	Dep. encargado	inf. centralizada	inf. Util	formato adecuado
Miembro 1	si	si	no	si	no	no	no	no sabe
Miembro 2	si	No sabe	no	si	no	no	no	no
Miembro 3	si	si	no	si	no	no	no	no

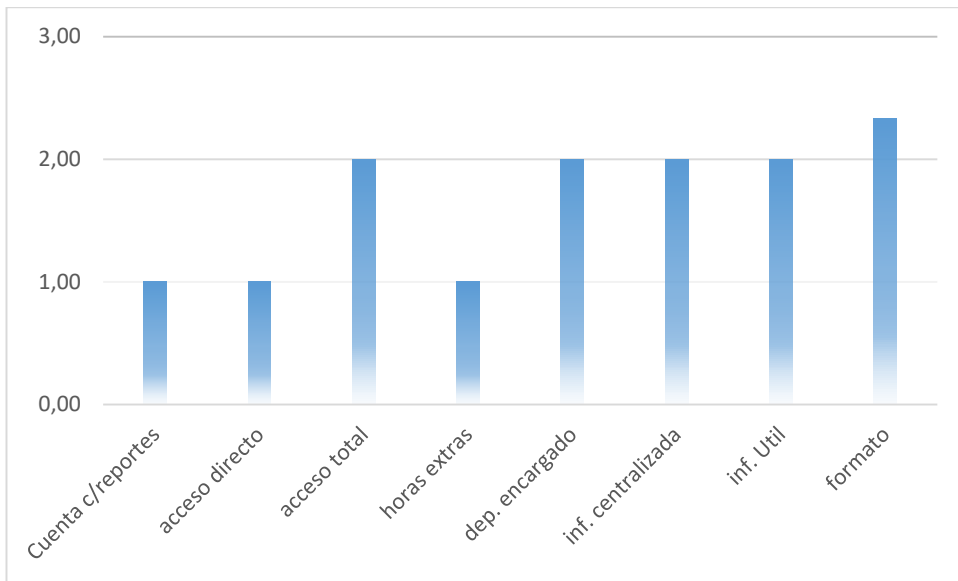


Figura 86. Gráfico de resultados de Entrevista

Como se puede observar la escuela de sistemas cuenta con un sistema al cual se permite el acceso a reportes no en su totalidad, en un formato que no es completamente dinámico, lo cual en ocasiones involucra horas extras para poder tomar decisiones en el tiempo establecido.

CAPÍTULO IV ANÁLISIS Y DISCUSIÓN DE RESULTADOS

4.1. Contratación de Hipótesis

La contrastación de la hipótesis se realizó mediante la prueba T de Student en la cual hay dos momentos uno antes (Pre-Test) y otro después (Pos-Test), con ello podemos entender que el primer momento sirve para entender los cambios que se suscitan en el segundo momento luego de aplicar la variable independiente.

A través del cual podremos aceptar o rechazar la hipótesis planteada. Para llevarla a cabo se identificaron los siguientes indicadores:

- ❖ Número de alumnos aprobados y desaprobados por curso.
- ❖ Número de veces que un alumno ha llevado un curso.
- ❖ Tiempo promedio utilizado en la construcción de reportes.
- ❖ Grado de dificultad para la generación de reportes.

Para contrastar la hipótesis se aplicó una encuesta los miembros del comité de la EAPIS. Ha sido tabulada, de manera que se calculen los resultados obtenidos de acuerdo a los rangos que se presentan a continuación.

En la Tabla podemos ver los valores para evaluar el nivel de satisfacción de los indicadores, antes mencionados, para el seguimiento académico.

Tabla 39. Valores para nivel de satisfacción

Valor	0	1	2
Nombre	NS	N	S
Nivel de Satisfacción	No sabe	NO	Si

Se tiene que:

$$PT_i = \sum_{j=1}^8 (F_{ij} * P_j)$$

Dónde:

PT_i = Puntaje Total de la pregunta i - esima

F_{ij} = Frecuencia j – esima de la pregunta i

P_j = Peso j – esima

El cálculo del promedio ponderado por cada pregunta sería:

$$\overline{PP}_1 = \frac{PT1}{n}$$

\overline{PP}_1 = Promedio de Puntaje Total de la pregunta i -esima

n = 3 usuarios

Tabla 40. Pre- test

Item	Preguntas	NS	N	S	Total	Puntaje Promedio
		0	1	2		
1	Cuenta con un sistema dedicado a generar reportes	0	0	3	3	1,67
2	Facilidad de uso del sistema actual	1	1	1	3	1,00
3	Accede directamente a la información académica	1	2	0	3	0,67
4	Posee acceso total a la información académica	0	3	0	3	1,00
5	Dedica horas extras al análisis de la información	0	0	3	3	1,33
6	Pérdida de tiempo por falta de disponibilidad de información	0	0	3	3	2,00
7	Esta de acuerdo con el encargado de la información	1	1	1	3	1,33
8	La información está centralizada	0	3	0	3	1,00
9	La información permite medir el desenvolvimiento académico	0	3	0	3	1,00
10	Esta de acuerdo al formato	0	2	1	3	1,33

Tabla 41. Post - test

Item	Preguntas	NS	N	S	Total	Puntaje Promedio
		0	1	2		
1	Cuenta con un sistema dedicado a generar reportes	0	0	3	3	2,00
2	Facilidad de uso del sistema actual	0	0	3	3	2,00
3	Accede directamente a la información académica	0	0	3	3	2,00
4	Posee acceso total a la información académica	0	0	3	3	2,00
5	Dedica horas extras al análisis de la información	0	3	0	3	0,67
6	Pérdida de tiempo por falta de disponibilidad de información	0	3	0	3	1,00
7	Esta de acuerdo con el encargado de la información	0	0	3	3	1,67
8	La información está centralizada	0	0	3	3	2,00
9	La información permite medir el desenvolvimiento académico	0	0	3	3	2,00
10	Esta de acuerdo al formato	0	0	3	3	1,33

A continuación la tabla de contrastación de resultados de las pruebas realizadas Pre y Post Test.

Tabla 42. Resultados Pre-test y Post-test

Pregunta	Pre-test (X _A)	Post-test (X _D)	d=(X _D -X _A)	(d-d')	(d-d') ²
1	1,67	2,00	0,33	-0,10	0,01
2	1,00	2,00	1,00	0,57	0,32
3	0,67	2,00	1,33	0,90	0,82
4	1,00	2,00	1,00	0,57	0,32
5	1,33	0,67	-0,67	-1,10	1,20
6	2,00	1,00	-1,00	-1,43	2,04
7	1,33	1,67	0,33	-0,10	0,01
8	1,00	2,00	1,00	0,57	0,32
9	1,00	2,00	1,00	0,57	0,32
10	1,33	1,33	0,00	-0,43	0,18
N=10	12,33	16,67	4,33		5,57

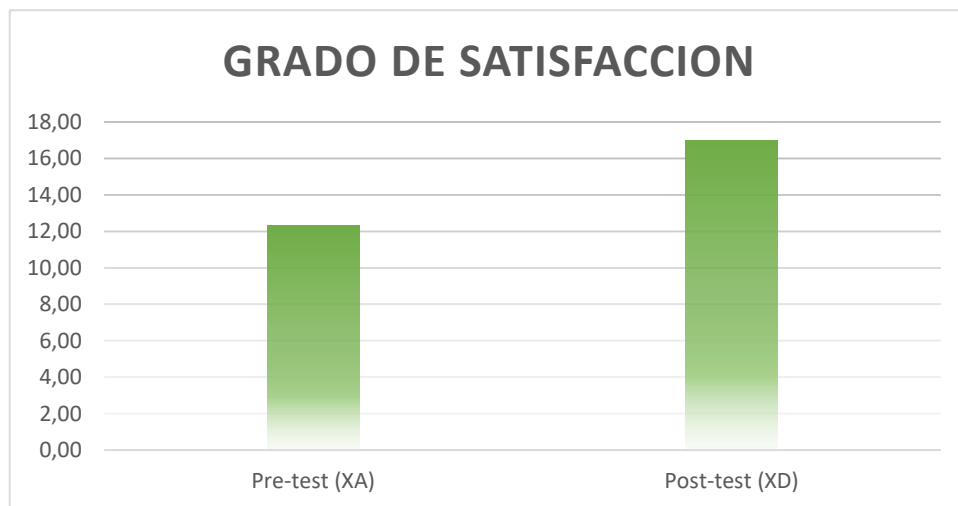


Figura 87. Pre-test vs Post-test

4.2. Formulación de la Hipótesis Estadística

❖ **Hipótesis Nula (H₀):**

La implementación de un data mart no permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela Académico de Ingeniería de Sistemas.

$$H_0: \mu_D \geq \mu_A; \mu_D - \mu_A \geq 0$$

❖ **Hipótesis alternative (Ha):**

La implementación de un data mart permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela Académico de Ingeniería de Sistemas.

$$H_0: \mu_D < \mu_A; \mu_D - \mu_A < 0$$

4.3. Nivel de significancia

En esta investigación se utilizará el nivel de significancia igual a 5%, ya que se trata de un proyecto de investigación, por lo tanto:

$$\alpha = 0.05$$

4.4. Valor estadístico del procedimiento

$$d' = \frac{4.33}{10} = 0.43$$

$$\sigma d = \sqrt{\frac{5.57}{9}} = 0.78$$

$$t = \frac{0.43}{\frac{0.78}{\sqrt{10}}} = 1.79$$

De acuerdo a la tabla t-student el valor de la probabilidad p del valor estadístico del procedimiento es el siguiente:

$$\rho(t < 1.79) = 0.02$$

4.5. Establecer región crítica

gl: Grados de libertad (N-1)

tt: Distribución t de student

$$gl = 9$$

$$tt_{0.05;9} = 1.833$$

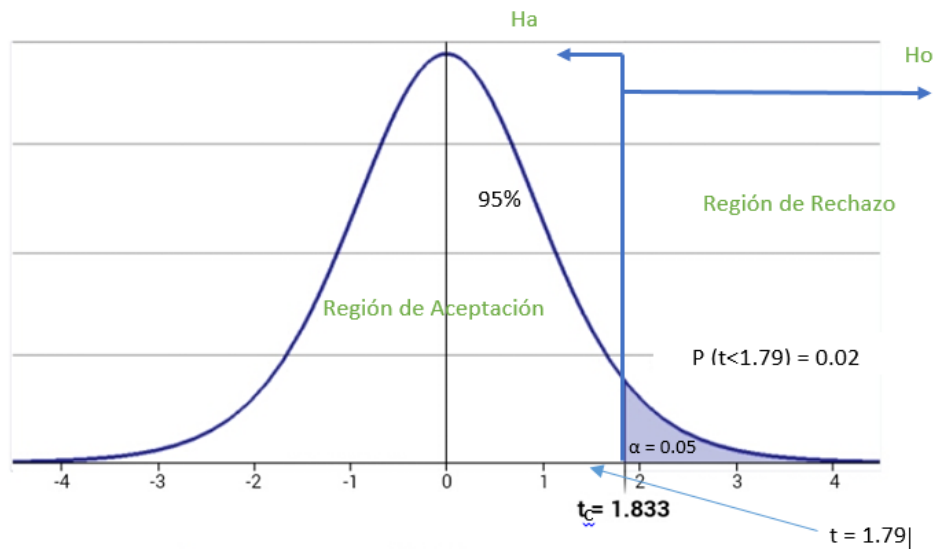


Figura 88. Distribucion T - Student

De acuerdo a lo obtenido:

$$t < t_c = 1.79 < 1.833$$

$$p < \alpha = 0.02 < 0.05$$

Como son verdaderas las dos proposiciones, entonces rechazamos la hipótesis H_0 , y aceptamos H_a .

4.6. Tiempo de respuesta en el acceso a la información

Tabla 43. Tabla de tiempo de respuesta

Reporte	Tiempo hrs		$d=(X_D-X_A)$	$(d-d')$	$(d-d')^2$
	Pre-test (X_A)	Post-test (X_D)			
1	24	0,01	-23,99	-11,38	129,41
2	10	0,01	-9,99	2,62	6,88
3	5	0,02	-4,98	7,63	58,21
4	8	0,02	-7,98	4,63	21,41
5	11	0,03	-10,98	1,64	2,67
6	9	0,02	-8,98	3,63	13,19
7	10	0,02	-9,98	2,63	6,91
8	24	0,02	-23,98	-11,37	129,35
N=8	101	0,15	-100,85		368,04

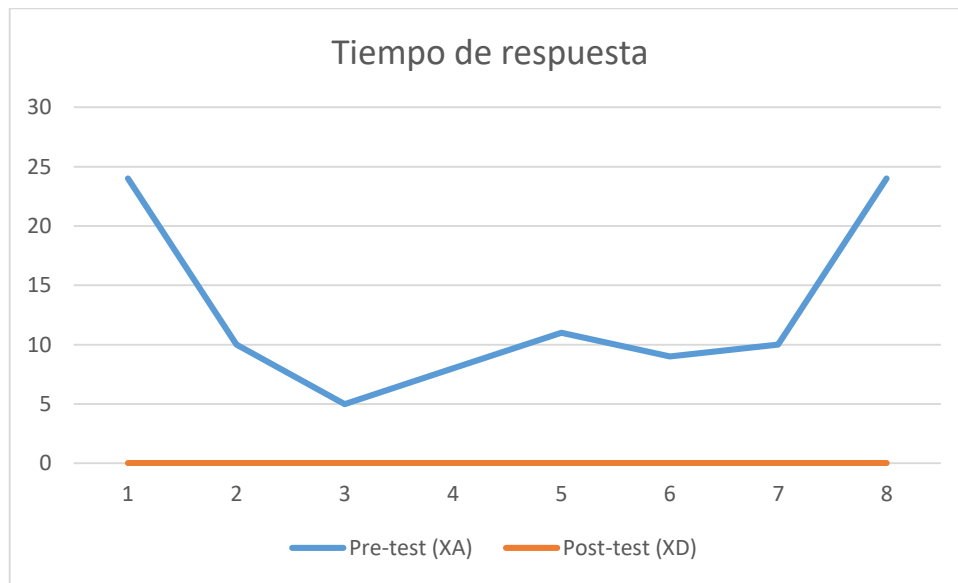


Figura 89. Gráfico del tiempo de respuesta

❖ Formulación de Hipótesis

▪ Hipótesis Nula (H₀):

El uso del data Mart no disminuye el tiempo de respuesta en el acceso a la información para la toma de decisiones en la Escuela Académico de Ingeniería de Sistemas.

$$H_0: \mu_D \geq \mu_A; \mu_D - \mu_A \geq 0$$

▪ Hipótesis alternative (H_a):

El uso del data Mart disminuye el tiempo de respuesta en el acceso a la información para la toma de decisiones en la Escuela Académico de Ingeniería de Sistemas.

$$H_0: \mu_D < \mu_A; \mu_D - \mu_A < 0$$

❖ Nivel de significancia

$$\alpha = 0.05$$

❖ Valor estadístico del procedimiento

$$d' = \frac{-100.85}{8} = -12,61$$

$$\sigma d = \sqrt{\frac{368.04}{7}} = 7.25$$

$$t = \frac{-12.61}{\frac{7.25}{\sqrt{8}}} = -4.93$$

De acuerdo a la tabla t-student el valor de la probabilidad p del valor estadístico del procedimiento es el siguiente:

$$\rho(t < -4.93) = 0.0008$$

❖ **Establecer región crítica**

gl: Grados de libertad (N-1)

tt: Distribución t de student

$$gl = 7$$

$$tt_{0.05;7} = 1.895$$

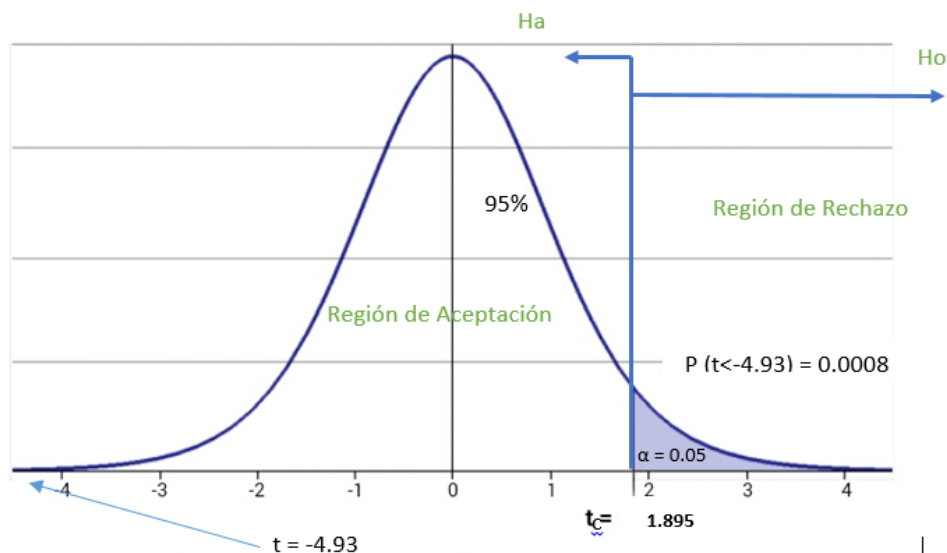


Figura 90. Distribución t-student

De acuerdo a lo obtenido:

$$t < tt = -4.93 < 1.895$$

$$p < \alpha = 0.0008 < 0.05$$

Como son verdaderas las dos proposiciones, entonces rechazamos la hipótesis H_0 , y aceptamos H_a .

4.7. Toma de decisión

Aceptamos la hipótesis alternativa y diremos que se puede ratificar que la implementación de un data mart permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela Académico de Ingeniería de Sistemas, con un nivel de confianza del 95% y un nivel de significancia del 5%.

Los datos confirman la hipótesis parcialmente porque se ha inicializado un cambio y una mejora en la manera en la que se realiza el seguimiento académico de los alumnos de la Escuela Académico Profesional de Ingeniería de Sistemas, sin embargo todo esto depende de la escuela para que los cambios se efectúen y se respeten.

4.8. Discusión de resultados

Con respecto a la comparación de los resultados obtenidos en el presente trabajo de investigación, con los resultados de algunos de los antecedentes Teóricos indicados anteriormente, se encuentran las siguientes semejanzas y diferencias:

Miguel Rodríguez Sanz en su proyecto fin de carrera, análisis y diseño de un Data Mart para el seguimiento académico de los alumnos de un entorno universitario.

Semejanzas: El proyecto ayuda a extraer la información universitaria de los alumnos, así como dar apoyo a la toma de decisiones con el fin de conocer la respuesta de los diferentes alumnos a los estudios que han comenzado.

Diferencias: En el proyecto se muestra un marco metodológico para el desarrollo de proyectos de Data Warehouse y/o Data Mart, así como la aplicación está enfocada a una universidad privada, vista como un negocio.

Aimacaña Quilumba Doris Eulalia, en su trabajo de graduación previo a la obtención del título de Ingeniero Informático, Análisis, Diseño e Implementación de una Data Mart Académico, usando tecnología de BI tanto software libre.

Semejanzas: la implementación del proyecto apoya la toma de decisiones estratégicas tanto de estudiantes como profesores, esto es muy importante para saber dónde y cómo solventar alguna falla, o mejorar las decisiones tomadas.

Diferencias: el proyecto se desarrolla con uso del software libre además que el Data Mart que permita almacenar sólo la información requerida e ir eliminando aquellos datos que obstaculizan la labor del análisis.

Tana Pasquel Gloria Estefanía, en su trabajo previo a la obtención del título de Ingeniería en Sistemas Computacionales, Data Mart para el Análisis de información del sistema Académico de la Universidad Técnica del Norte con Herramientas de Software Libre.

Semejanzas: el desarrollo de la investigación demuestra que es importante el uso de las tecnologías BI para apoyar la toma de decisiones en todo tipo de ámbitos incluso en la educación.

Diferencias: la implementación del Data Mart se hizo con la finalidad de demostrar que el uso de herramientas de software libre en la actualidad está en auge en las pequeñas y medianas empresas además el proyecto se desarrolló con software libre.

Montalvo Barrientos, Miguel Ángel, en su trabajo, Implementación de un Data Mart para la Facultad Nacional de Ingeniería Caso: Ing. De Sistemas en la Universidad Técnica de Oruro.

Semejanzas: con la implementación del Data Mart con una base de datos multidimensional se ha logrado integrar la información necesaria y la mejoría en el análisis de información.

Diferencias: se ha implementado en una herramienta web de análisis OLAP.

CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

Se cumplió con el objetivo de implementar un Data Mart para realizar el seguimiento académico de los estudiantes de la Escuela Académico Profesional de Ingeniería de Sistemas, para de esta manera alcanzar la posibilidad de formular estrategias de mejora de manera más acertada y rápida.

Se han logrado satisfacer los requerimientos de la Escuela Académico Profesional de Ingeniería de Sistemas en cuanto al seguimiento académico del proceso educativo de esta escuela, obtenidos a través de reuniones con los miembros del comité de toma de decisiones los cuales fueron descritos en la metodología empleada.

El diseño y carga de datos proporcionados por la Unidad de informática, en una estructura dimensional, identificando las tablas hechos y las dimensiones, si permitió obtener una salida de información, datos obtenidos del Data Mart, los cuales pueden ser analizados para apoyar la toma de decisiones de la Escuela académico Profesional de Ingeniería de Sistemas.

Los reportes fueron construidos con la información proporcionada, según el comité, para la toma de decisiones de la escuela; para esta construcción se utilizó la herramienta BI llamada Power BI, los cuales has sido aprobados por el comité.

Determinar la utilidad de obtener la información consolidada de los estudiantes usando un Data Mar permite:

- ❖ Reducir el tiempo de espera por la información necesaria para la toma de decisiones en cuanto al rendimiento académico de los alumnos.
- ❖ Mejorar los formatos en los que la información es presentada.

5.2. Recomendaciones

Es recomendable tener en cuenta para la implementación de futuras versiones del Data Mart, las siguientes funcionalidades:

- ❖ Se debe realizar un backup de la información y del Data Mart, ante posibles dificultades, esto servirá para tener mayor seguridad.

- ❖ La seguridad es un aspecto importante cuando de notas de los alumnos hablamos, por tal motivo es importante cuidar seguridad en el acceso de personal autorizado a la base de datos así como el acceso restringido a diferentes usuarios de los reportes y la información aquí utilizada.
- ❖ El éxito de la implementación del Data Mart depende del uso constante de los miembros del comité antes de tomar las decisiones en la EAPIS.
- ❖ Es importante que el Data Mart sea aceptado por la Unidad Técnica de Sistemas informáticos para asegurar que la información se actualice a medida que la base de datos se actualiza.
- ❖ Se recomienda que se sanee la información de la base de datos de la Unidad Técnica de Sistemas Informáticos.

Bibliografía

- [1] L. P. P. Julio Yalan Castillo, "Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso," *REVISTA DE INVESTIGACION DE SISTEMAS E INFORMATICA*, vol. 10, no. 1, p. 54, ENERO - JUNIO 2013. Lima-Peru.
- [2] G. L. J. E. y. V. A. J. d. Carmen, "DATA WAREHOUSE PARA EL ANÁLISIS ACADÉMICO DE LA ESCUELA POLITECNICA NACIONAL," 2007. Quito-Ecuador.
- [3] M. R. Sanz, "ANÁLISIS Y DISEÑO DE UN DATA MART PARA EL SEGUIMIENTO ACADÉMICO DE ALUMNOS EN UN ENTORNO UNIVERSITARIO," . Madrid- España. 2010.
- [4] M. R. R. M. Y. P. A. R. Tejada, "DESARROLLO DE UN DATAMART DE INFORMACIÓN ACADÉMICA DE ESTUDIANTES DE LA ESCUELA DE CIENCIAS Y SISTEMAS DE LA FACULTAD DE INGENIERÍA DE LA USAC," 2007. Guatemala.
- [5] A. Q. D. Eulalia, "ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE UN DATA MART ACADÉMICO USANDO TECNOLOGÍA DE BI PARA LA FACULTAD DE INGENIERÍA, CIENCIAS FÍSICAS Y MATEMÁTICA.," Quito- Ecuador, 2013.
- [6] G. E. T. Pasquel, "Datamart para el Análisis de Información del Sistema Académico de la Universidad Técnica del Norte con Herramientas de Software Libre," Ibarra - Ecuador, 2014.
- [7] M. A. M. Barrientos, "Implementación de un Data Mart para la Facultad Nacional de Ingeniería Caso: Carrera Ing. de Sistemas," . Oruro - Bolivia. 2010.
- [8] S. e. I. d. N. S. L., Enero 2007-2016. [Online]. Available: http://www.sinnexus.com/business_intelligence/datamart.aspx.
- [9] [Online]. Available: http://mydaneeblog.blogspot.pe/2011_03_01_archive.html.
- [10] T. S. a. M. P. R. C. Q. R. Gustavo Hernandez, "Slide Share," 13 Junio 2012. [Online]. Available: <http://es.slideshare.net/GustavoHernandez10/data-mart>.
- [11] "BUYTO," 2009. [Online]. Available: <http://www.buyto.es/general-business-intelligence/almacenamiento-de-datos-datawarehouse-datamart-en-business-intelligence>. [Accessed 20 Mayo 2016].
- [12] J. L. Cano, "Business Intelligence: Competir con Información".
- [13] "DATAPRIX," [Online]. Available: <http://www.dataprix.com/blogs/respinosamilla/qu-business-intelligence>. [Accessed 20 Mayo 2016].
- [14] Oracle. [Online]. Available: http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf. [Accessed 20 Mayo 2016].
- [15] J. A. Velez, "IT SYNERGY," 2016. [Online]. Available: <http://blogs.itsynergy.co/jvelezc/2011/02/18/arquitectura-de-inteligencia-de-negocios/>. [Accessed 21 Mayo 2016].
- [16] "MIRA," [Online]. Available: <http://blog.mirai-advisory.com/?p=461>. [Accessed 08 09 2016].
- [17] "Dataprix," [Online]. Available: <http://www.dataprix.com/arquitectura-data-warehouse-areas-datos-nuestro-almacen-corporativo>. [Accessed 12 08 2016].
- [18] W. 2.0, "Inteligencia de Negocios," 30 Enero 2014. [Online]. Available: <http://inteligenciadenegociosval.blogspot.pe/2014/01/metodologia-de-kimball.html>. [Accessed 2016 09 11].
- [19] D. Brito, "Blogger.com," 26 enero 2014. [Online]. Available: <http://inteligenciadenegociosdiegobrito.blogspot.pe/2014/01/la-metodologia-de-kimball.html>. [Accessed 2016 09 11].

- [20] Josep, "Blogger.com," 12 Enero 2009. [Online]. Available: <http://bi-businessintelligence.blogspot.pe/2009/01/cif-vs-md-dos-enfoques-clsicos-en-el.html>. [Accessed 2016 09 11].
- [21] "El rincon del BI," 19 Abril 2010. [Online]. Available: <https://churriwifi.wordpress.com/2010/04/19/15-2-ampliacion-conceptos-del-modelado-dimensional/>. [Accessed 2016 09 11].
- [22] "technet," [Online]. Available: [https://technet.microsoft.com/es-es/library/ms175609\(v=sql.90\).aspx](https://technet.microsoft.com/es-es/library/ms175609(v=sql.90).aspx). [Accessed 11 09 2016].
- [23] Microsoft, "Microsoft," [Online]. Available: <https://msdn.microsoft.com/en-us/library/ms141026.aspx>. [Accessed 11 09 2016].
- [24] Microsoft, "POWER BI," [Online]. Available: <https://powerbi.microsoft.com/es-es/what-is-power-bi/>. [Accessed 11 09 2016].
- [25] Microsoft, "Microsoft," Microsoft, 2016. [Online]. Available: <https://msdn.microsoft.com/es-pe/library/ms175641.aspx>. [Accessed 10 10 2016].
- [26] C. Idalberto, Introducción a la Teoría General de la Administración, Séptima edición ed., McGraw-Hill Interamericana, 2006.
- [27] F. O. C. y. H. Geoffrey, «Introducción a los Negocios en un Mundo Cambiante», Cuarta Edicion ed., McGraw-Hill Interamericana, 2004.
- [28] C. M. y. K. Masaaki, «Administración de Mercadotecnia», Segunda Edicion ed., International Thomson Editores, 2001.
- [29] Febrero 2016. [Online]. Available: <http://definicion.de/proceso-educativo/#ixzz40IWuBwl>.
- [30] E. P. d. I. d. Sistemas, "Curricula de estudios 2007," Cajamarca, 2007.
- [31] Enero 2016. [Online]. Available: http://www.ecured.cu/Rendimiento_acad%C3%A9mico.
- [32] [Online]. Available: http://revistaeconomia.puj.edu.co/html/articulos/Numero_3/9.pdf..
- [33] G. G. Olgún. [Online]. Available: <http://www.eiq.cl/cchiq2009/resumenes/C/CP1.pdf>.
- [34] Anonimo, Enero 2016. [Online]. Available: <http://www.mastermagazine.info/termino/5366.php>.
- [35] Anonimo, 2009. [Online]. Available: www.sap.com/spain/contactsap.
- [36] A. Q. D. Eulalia, "ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE UN DATA MART ACADÉMICO USANDO TECNOLOGÍA DE BI PARA LA FACULTAD DE INGENIERÍA, CIENCIAS FÍSICAS Y MATEMÁTICA," Quito, 2013.
- [37] I. G. Camejo, 2000. [Online]. Available: http://www.bvs.sld.cu/revistas/aci/vol8_3_00/aci05300.pdf.
- [38] "Universidad Nacional de Cajamarca," [Online]. Available: www.unc.edu.pe. [Accessed 19 07 2016].
- [39] U. d. Granada. [Online]. Available: <http://elvex.ugr.es/idbis/db/docs/design/2-requirements.pdf>. [Accessed 10 08 2016].
- [40] Slide, "Slide," [Online]. Available: https://eva.fing.edu.uy/pluginfile.php/58481/mod_resource/content/1/pavan-teorico04-requerimientos.pdf. [Accessed 10 08 2016].
- [41] Slide, "Slide," [Online]. Available: https://eva.fing.edu.uy/pluginfile.php/58481/mod_resource/content/1/pavan-teorico04-requerimientos.pdf. [Accessed 10 08 2016].
- [42] DATAprix. [Online]. Available: <http://www.dataprix.com/forum/2010/04/ods-staging-area>. [Accessed 12 08 2016].
- [43] A. Azuaje, "Metodología de Kimball.," caracas, 2014.

- [44] M. L. E. Velasquez, "SlideShare," 15 01 2014. [Online]. Available: <http://es.slideshare.net/PTAaTLANTICO/seguimiento-a-procesos-academicos-2>. [Accessed 09 10 2016].
- [45] L. G. S. I. y. otros, "Orientación Educativa," Jimdo, [Online]. Available: <http://orientacioneducativa04.jimdo.com/tutoria/seguimiento-del-proceso-de-los-alumnos/>. [Accessed 10 10 2016].
- [46] B. R. M. A. Huaman, "“IMPACTO DE LA IMPLEMENTACIÓN DE DATA MARTS EN LAS ÁREAS DE COMPRAS Y VENTAS DE LA EMPRESA COMUNICACIONES SAGPERÚ MÓVILES SRL”, " Cajamarca, 2016.
- [47] S. Ramos, Data Warehouse, Data Marts y Modelos Dimensionales. Un pilar fundamental para la Toma de Decisiones, España, 2016.

ANEXO 1

1. ¿La EAPIS cuenta actualmente con un sistema dedicado a generar reportes con la información académica de los alumnos de la misma?
 - a) Si
 - b) No
 - c) No sabe
2. ¿El sistema con el que se cuenta actualmente es de fácil uso?
 - a) Si
 - b) No
 - c) No sabe
3. ¿accede Ud. Directamente a la información académica de la EAPIS?
 - a) Si
 - b) No
 - c) No sabe
4. ¿Posee Ud. Total acceso a la información académica de la EAPIS?
 - a) Si
 - b) No
 - c) No sabe
5. ¿Dedica Ud. Horas extras al análisis de documentos e informes relacionados con la parte académica de los alumnos de la EAPIS?
 - a) Si
 - b) No
 - c) No sabe
6. ¿Ha perdido Ud. tiempo por la falta de disponibilidad de información?
 - a) Si
 - b) No
 - c) No sabe
7. ¿Considera factible o necesaria la instalación de componentes de inteligencia de negocios dentro de la dirección de la EAPIS?
 - a) Si
 - b) No
 - c) No sabe

8. En cuanto a la información académica de la EAPIS, ¿Está de acuerdo con la persona o departamento encargado?
- a) Si
 - b) No
 - c) No sabe
9. ¿La información académica de la EAPIS se encuentra centralizada o distribuida en diferentes repositorios?
- a) Si
 - b) No
 - c) No sabe
10. ¿La información académica brindada le permiten a Ud. medir el desenvolvimiento académico de los alumnos de la EAPIS?
- a) Si
 - b) No
 - c) No sabe
11. En cuanto a la entrega de información y reportes, ¿Esta Ud. De acuerdo en el formato en que recibe la información?
- a) Si
 - b) No
 - c) No sabe

INGENIERIA DE SISTEMAS ESCUELA ACADÉMICO PROFESIONAL ABRIL PROCES DE NEGOCIO ASIGNATURA		INGENIERIA DE SISTEMAS ESCUELA ACADÉMICO PROFESIONAL MAYO PROCES DE NEGOCIO ASIGNATURA		INGENIERIA DE SISTEMAS ESCUELA ACADÉMICO PROFESIONAL JUNIO PROCES DE NEGOCIO ASIGNATURA	
SÁNCHEZ CHÁVEZ ROGER M. Apellidos y Nombres del Docente		SÁNCHEZ CHÁVEZ ROGER M. Apellidos y Nombres del Docente		SÁNCHEZ CHÁVEZ ROGER M. Apellidos y Nombres del Docente	
Fecha	TEMA DESARROLLADO	Fecha	TEMA DESARROLLADO	Fecha	TEMA DESARROLLADO
Mes: SEPTIEMBRE		Mes: OCTUBRE		Mes: DICIEMBRE	
02	Introducción al curso Contenido de vídeo	05	Nivel de madurez BPM	14	Segunda Práctica Calificada
07	Introducción a BPM Remanentes de madurez	07	Modelos de Medición de BPM	16	Introducción y alcance de medición
09	Organización del equipo Asesistente de BPM	12	Asesoría y sustentación de BPM	21	Desempeño de BPM en medición
14	Implementación de BPM	14	Modelos en caso de sus mediciones BPM	23	Construcción de los datos en JIRA
16	Uso de los scripts de BPM 2.0	19	Paradigma orientado a servicios	28	Aplicación del paradigma en sus caso
21	Identificación de procesos principales	21		29	Implementación de datos en JIRA - Proyecto
23		26			
28	Captura de datos principales en un BPM	28			
30	Práctica Práctica Calificada				

Nacional de Cajamarca

LA UNIVERSIDAD PERUANA

Ley 14015 del 13 de Febrero de 1962

CAJAMARCA - PERÚ

Resolución Rectoral N° 245-2016-UNC.

Cajamarca, 23 de agosto del 2016.

con Registro N° 0626-2016-SG-UNC. de fecha 23 de agosto del 2016. Rector de la Universidad Nacional de Cajamarca. y:

en el último párrafo del Artículo 18°, señala que: "Cada una de las facultades de la Universidad, en el marco de su autonomía, de gobierno, académico, administrativo y económico. Las facultades operan en el marco de la Constitución y de la Ley";

Artículo 220, establece: "El Estado reconoce la autonomía universitaria. La autonomía se ejerce en conformidad con lo establecido en la Constitución, la Ley y demás normas legales. Los regímenes: Normativo, De gobierno, Académico, Administrativo y Económico";

La Universidad Nacional de Cajamarca, establece que la Oficina de Sistemas Informáticos y Plataformas Virtuales, diseña, planifica, ejecuta, actualiza, supervisa los sistemas y académicos de la Universidad. Implementa plataformas virtuales que establecen en el Estatuto, Reglamento General y a través de las unidades administrativas que el Reglamento

4-2016-UTSI. de fecha 12 de agosto del 2016, suscrito por el Rector de la Universidad Nacional de Cajamarca, General de Sistemas Informáticos y Plataformas Virtuales, autorice el uso del Sistema Informático Académico en la

1, es un sistema de almacenamiento y procesamiento de datos de la Universidad Nacional de Cajamarca, permitirá agilizar y asegurar el correcto funcionamiento académico, optimizando el uso de los recursos disponibles;

2, emitir la Resolución correspondiente, a través de la cual se autorice el uso del Sistema Informático Académico en la

3, las facultades conferidas por el Artículo 62° de la Ley Universitaria

4, el uso del SISTEMA INFORMÁTICO ACADÉMICO-SIA de la Universidad Nacional de Cajamarca.

5, por la presente Resolución al Rectorado, Vicerrectorado de Investigación y Presupuesto, Calidad Educativa y Acreditación y las Unidades Académicas Virtuales, para los fines pertinentes.

Regístrese, comuníquese y archívese.



Cs. VICTOR ELI RODRIGUEZ LESCANO
SECRETARIO GENERAL

ESPECIFICACIONES	N°	%	% Evaluados
MATRICULADOS	23	100	
APROBADOS	19	83	100
DESAPROBADOS	0	0	0
INHABILITADOS	4	17	
TOTAL	23	100	100%

FIRMA DEL DOCENTE

V° B°

Jefe del Dpto. Académico

Director de la Escuela Académica Profesional

ANEXO 3

ANEXO 3
Nacional De Cajamarca

REPORTE ACADEMICO DE NOTAS

Unidad Técnica de Registro y Matricula

SPECIALIDAD : INGENIERIA DE SISTEMAS

CODIGO UNC : 09110022

CODIGO ANR : 2009110022

ALUMNO : LUDENA CHAVEZ JEAN OMAR

Idcurso	Nota	Cred	Asignatura	Nº acta	Sem	Fecha	Observ.
PRIMER AÑO							
11Q145	(16)	3	LENGUAJE Y COMUNICACION	2010111028	1	27/08/2010	
11Q146	(13)	3	LOGICA	2010111030	1	27/08/2010	
11Q147	09	3	MATEMATICA BASICA	2010111032	1	27/08/2010	
11Q147	(11)	3	MATEMATICA BASICA	2011111033	1	22/07/2011	
11Q148	(13)	3	INTRODUCCION A LA COMPUTACION INFORMATICA	2010111025	1	27/08/2010	
11Q149	(14)	3	FISICA I	2010111017	1	27/08/2010	
11Q150	(16)	3	METODOS Y TECNICAS DE ESTUDIO	2010111035	1	27/08/2010	
11Q151	(15)	3	FILOSOFIA DE LA CIENCIA Y LA TECNOLOGIA	2010111014	1	27/08/2010	
11Q152	(17)	4	ALGORITMOS Y ESTRUCTURA DE DATOS	2010411004	2	31/12/2010	
11Q153	(13)	4	ANALISIS MATEMATICO I	2011411005	2	16/12/2011	
11Q154	(14)	3	CONTABILIDAD Y FINANZAS	2010411016	2	31/12/2010	
11Q155	(12)	3	DIBUJO Y GEOMETRIA DESCRIPTIVA	2011411019	2	16/12/2011	
11Q156	(14)	3	ETICA	2010411027	2	31/12/2010	
11Q157	(13)	3	FISICA II	2010411029	2	31/12/2010	
11Q158	(15)	3	METODOLOGIA DE LA INVESTIGACION CIENTIFICA	2010411040	2	31/12/2010	
SEGUNDO AÑO							
11Q159	(13)	4	ALGORITMOS Y ESTRUCTURA DE DATOS II	2011111002	1	22/07/2011	
11Q160	(11)	4	ANALISIS MATEMATICO II	2012111006	1	26/08/2012	
11Q161	(15)	3	ECOLOGIA Y MEDIO AMBIENTE	2011111013	1	22/07/2011	
11Q162	(12)	3	ESTADISTICA GENERAL	2012111014	1	25/08/2012	
11Q163	(15)	3	ORGANIZACION Y ADMINISTRACION DE EMPRESAS	2011111036	1	22/07/2011	
11Q164	(15)	3	PLANIFICACION ESTRATEGICA	2011111040	1	22/07/2011	
11Q165	(16)	3	SISTEMA DIGITAL	2011111047	1	22/07/2011	
11Q166	(11)	4	ANALISIS MATEMATICO III	2015411009	2	23/12/2015	
11Q167	(14)	3	ARQUITECTURA DEL COMPUTADOR	2011411010	2	16/12/2011	
11Q168	(13)	4	BASE DE DATOS I	2011411014	2	16/12/2011	
11Q169	(14)	3	DEFENSA NACIONAL Y DERECHOS HUMANOS	2012411016	2	22/12/2012	
11Q170	(14)	3	ECONOMIA Y GESTION EMPRESARIAL	2013411021	2	20/12/2013	
11Q170	09	3	ECONOMIA Y GESTION EMPRESARIAL	2012411021	2	22/12/2012	
11Q171	(11)	3	ESTADISTICA Y PROBABILIDADES	2012411023	2	22/12/2012	
11Q172	(11)	3	GESTION DE PROCESOS ORGANIZACIONALES	2011411029	2	16/12/2011	
TERCER AÑO							
11Q173	(11)	4	BASE DE DATOS II	2012111007	1	25/08/2012	
11Q174	(14)	3	CULTURA Y REALIDAD NACIONAL	2013111009	1	09/08/2013	
11Q175	(14)	3	GESTION DEL RECURSO HUMANO	2012111020	1	25/08/2012	
11Q176	(11)	3	OPTIMIZACION EN INGENIERIA I	2013111035	1	09/08/2013	
11Q177	(12)	4	SISTEMAS DE INFORMACION	2012111054	1	25/08/2012	
11Q178	(14)	3	TEORIA DE LENGUAJES	2012111059	1	25/08/2012	
11Q179	(12)	4	INGENIERIA DE SOFTWARE I	2012411032	2	22/12/2012	
11Q180	(16)	3	E-MARKETING	2012411019	2	22/12/2012	
11Q181	(11)	3	OPTIMIZACION EN INGENIERIA II	2013411044	2	20/12/2013	
11Q182	(18)	4	PROGRAMACION APLICADA I	2013311008	2	08/03/2013	
11Q182	09	4	PROGRAMACION APLICADA I	2012411042	2	22/12/2012	
11Q183	(13)	3	SISTEMAS OPERATIVOS	2012411052	2	22/12/2012	
11Q184	(12)	3	TEORIA DE SISTEMAS	2014411056	2	19/12/2014	
11Q206	(14)	3	BASE DE DATOS AVANZADAS (E)	2012411011	2	22/12/2012	
CUARTO AÑO							
11Q185	(11)	3	DINAMICA DE SISTEMAS	2015111013	1	07/08/2015	

SPECIALIDAD : INGENIERIA DE SISTEMAS

CODIGOUNC : 09110022

CODIGOANR : 2009110022

ALUMNO : LUDEÑA CHAVEZ JEAN OMAR

Idcurso	Nota	Cred	Asignatura	Nº acta	Sem	Fecha	Obsarv.
CUARTO AÑO							
11Q186	(14)	3	GESTION DEL CONOCIMIENTO	2013111018	1	09/08/2013	
11Q187	(11)	4	INGENIERIA DE SOFTWARE II	2013111021	1	09/08/2013	
11Q188	(13)	4	PROGRAMACION APLICADA II	2013111041	1	09/08/2013	
11Q189	(11)	4	SISTEMA DE COMUNICACION DE DATOS	2013111045	1	09/08/2013	
11Q190	(16)	3	TEORIA DE DECISIONES	2014111053	1	01/08/2014	
11Q191	(14)	4	ARQUITECTURA DE PROCESOS DE NEGOCIO	2013411010	2	20/12/2013	
11Q192	(12)	3	INGENIERIA DE PROCESOS	2014411037	2	19/12/2014	
11Q193	(12)	4	PLANEAMIENTO DE SISTEMAS DE INFORMACION I	2013411046	2	20/12/2013	
11Q194	(16)	4	REDES Y CONECTIVIDAD	2013411052	2	20/12/2013	
11Q195	(12)	3	SISTEMAS DE INFORMACION INTEGRADOS	2013411055	2	20/12/2013	
11Q207	(17)	3	PROGRAMACION AVANZADA (E)	2014411049	2	19/12/2014	
QUINTO AÑO							
11Q196	(14)	3	ADMINISTRACION DE PROYECTOS DE SISTEMAS	2014111001	1	01/08/2014	
11Q197	(13)	4	ADMINISTRACION DE REDES	2014111003	1	01/08/2014	
11Q198	(13)	4	PLANEAMIENTO DE SISTEMAS DE INFORMACION II	2014111042	1	01/08/2014	
11Q199	(14)	3	SISTEMAS INTELIGENTES	2015111054	1	07/08/2015	
11Q200	(13)	4	TOPICOS EN SISTEMAS	2014111057	1	01/08/2014	
11Q208	(11)	3	INGENIERIA WEB (E)	2014111026	1	01/08/2014	
11Q201	(12)	3	GERENCIA EN TI	2014411031	2	19/12/2014	
11Q202	(13)	4	GESTION DE RIESGOS Y SEGURIDAD DE LA INFORMACION	2014411035	2	19/12/2014	
11Q203	(12)	3	LEGISLACION INFORMATICA	2015411042	2	23/12/2015	
11Q204	(15)	3	LIDERAZGO	2015411044	2	23/12/2015	
11Q205	(15)	3	TESIS	2015411066	2	23/12/2015	
11Q209	(15)	3	REDES CORPORATIVAS (E)	2014411051	2	19/12/2014	

Resultados Académicos:

Total de Créditos Llevados : 225
 Total de Créditos Aprobados : 215
 Promedio Global : 13.182

NOTA:

Alumno:- VERIFICAR SUS NOTAS BAJO SU RESPONSABILIDAD
 CUALQUIER ERROR QUE CONSIDERE INFORMAR A LA
 UNIDAD TECNICA DE REGISTRO Y MATRICULA
 CANCELAR RECIBO POR DERECHO DE RECORD DE NOTAS

Decanato:- LA PERSONA ENCARGADA DE LA ELABORACION DE CERTIFICADOS
 DE ESTUDIOS, BAJO RESPONSABILIDAD REALIZARA LA VERIFICACION DE
 LAS NOTAS CON LAS ACTAS DEFINITIVAS.

ESCUELA ACADÉMICO PROFESIONAL DE INGENIERIA DE SISTEMAS
Facultad de Ingeniería
RELACION DE ASIGNATURAS (S.A-2014-11)

PRIMER AÑO

SEGUNDO SEMESTRE						Nº Secciones y grupos		DOCENTE
Cod.	Asignatura	HT	HP	TH	Dpto. Académico	Teoría	Práctica	
	Contabilidad y Finanzas	2	2	4	Contabilidad			
	Contabilidad y Finanzas	2	2	4	Contabilidad			
	Dibujo y Geometría Descriptiva	2	2	4	DAOI			
	Dibujo y Geometría Descriptiva	2	2	4	DAOI			
	Análisis Matemático I	3	3	6	Matemáticas			
	Análisis Matemático I	3	3	6	Matemáticas			
	Algoritmos y Estructura de Datos I	3	3	6	DASEI			
	Algoritmos y Estructura de Datos I	3	3	6	DASEI			
	Algoritmos y Estructura de Datos I	3	3	6	DASEI			
	Física II	2	3	5	Física			
	Física	2	2	4	derecho			
	Física	2	2	4	derecho			
	Metodología de la Investigación Científica	2	3	5	CC. SS.			
	Metodología de la Investigación Científica	2	3	5	CC. SS.			

SEGUNDO AÑO						Nº Secciones y grupos		DOCENTE
Cod.	Asignatura	HT	HP	TH	Dpto. Académico	Teoría	Práctica	
	Gestión de Procesos Organizacionales	2	2	4	DASEI			
	Gestión de Procesos Organizacionales	2	2	4	DASEI			
	Estadística y Probabilidades	2	3	5	DASEI			
	Estadística y Probabilidades	2	3	5	DASEI			
	Análisis Matemático III	3	3	6	Matemáticas			
	Análisis Matemático III	3	3	6	Matemáticas			

Base de Datos I	3	3	6	DASEI			
Arquitectura del Computador	2	3	5	DASEI			
Defensa Nacional y Derechos Humanos	2	2	4	Derecho			
Economía y Gestión Empresarial	2	2	4	Economía			

TERCER AÑO

SEGUNDO SEMESTRE

Cod.	Asignatura	HT	HP	TH	Dpto. Académico	Nº Secciones y grupos		DOCENTE
						Teoría	Práctica	
	e - Marketing	2	2	4	DASEI			
	Optimización en Ingeniería II	2	3	5	DASEI			
	Ingeniería de Software I	3	3	6	DASEI			
	Ingeniería de Software I	3	3	6	DASEI			
	Programación Aplicada I	3	3	6	DASEI			
	Programación Aplicada I	3	3	6	DASEI			
	Sistemas Operativos	2	3	5	DASEI			
	Sistemas Operativos	2	3	5	DASEI			
	Teoría de Sistemas	2	2	4	DASEI			
	Base de Datos Avanzados	2	2	4	DASEI			
	Base de Datos Avanzados	2	2	4	DASEI			

CUARTO AÑO

SEGUNDO SEMESTRE

Cod.	Asignatura	HT	HP	TH	Dpto. Académico	Nº Secciones y grupos		DOCENTE
						Teoría	Práctica	
	Sistemas de Información Integrados	2	3	5	DASEI			
	Sistemas de Información Integrados	2	3	5	DASEI			
	Ingeniería de Procesos	2	3	5	DASEI			
	Ingeniería de Procesos	2	3	5	DASEI			
	Planamiento de Sistemas de Información I	3	3	6	DASEI			
	Planamiento de Sistemas de Información I	3	3	6	DASEI			
	Arquitectura de Procesos de Negocio	3	3	6	DASEI			
	Arquitectura de Procesos de Negocio	3	3	6	DASEI			
	Redes Y Conectividad	3	3	6	DASEI			

Redes y Conectividad	3	3	6	DASEI			
Programación Avanzada	2	2	4	DASEI			

QUINTO AÑO

SEGUNDO SEMESTRE

Cod.	Asignatura	HT	HP	TH	Dpto. Académico	N° Secciones y grupos		DOCENTE
						Teoría	Práctica	
	Gestión de Riesgos y Segur. de la Inform.	3	3	6	DASEI			
	Liderazgo	2	2	4	Contabilidad			
	Gerencia en TI	2	3	5	DASEI			
	Legislación Informática	2	2	4	Derecho			
	Tesis	2	3	5	DASEI			
	Redes Corporativas	2	2	4	DASEI			

ANEXO 5

CONSOLIDADO DE ALUMNOS(APROB. DESAPROB. INHAB.) POR CURSO

Universidad Nacional de Cajamarca
Oficina Técnica de Registro Central

Fecha Rep. 05/03/2015 09:36:03

Escuela: **11 INGENIERIA DE SISTEMAS**

Período 2014 - II

			NIVAPROBADOS	DESAPROB.	INHAB.	TOTAL	
1	ALGORITMOS Y ESTRUCTURA DE DATOS I	11Q152	1	51 69.86%	19 26.03%	3 4.11%	73 100%
2	ANALISIS MATEMATICO I	11Q153	1	26 44.07%	8 13.56%	25 42.37%	59 100%
3	CONTABILIDAD Y FINANZAS	11Q154	1	52 86.67%	2 3.33%	6 10.00%	60 100%
4	DIBUJO Y GEOMETRIA DESCRIPTIVA	11Q155	1	45 65.22%	15 21.74%	9 13.04%	69 100%
5	ETICA	11Q156	1	45 72.58%	9 14.52%	8 12.90%	62 100%
6	FISICA II	11Q157	1	38 40.00%	41 43.16%	16 16.84%	95 100%
7	METODOLOGIA DE LA INVESTIGACION CIENTI	11Q158	1	41 67.21%	15 24.59%	5 8.20%	61 100%
8	ANALISIS MATEMATICO II	11Q160	2	22 73.33%	1 3.33%	7 23.33%	30 100%
9	ANALISIS MATEMATICO III	11Q166	2	44 86.27%	0 0.00%	7 13.73%	51 100%
10	ARQUITECTURA DEL COMPUTADOR	11Q167	2	31 93.94%	0 0.00%	2 6.06%	33 100%
11	BASE DE DATOS I	11Q168	2	37 86.05%	2 4.65%	4 9.30%	43 100%
12	DEFENSA NACIONAL Y DERECHOS HUMANOS	11Q169	2	50 92.59%	1 1.85%	3 5.56%	54 100%
13	ECONOMIA Y GESTION EMPRESARIAL	11Q170	2	59 93.65%	2 3.17%	2 3.17%	63 100%
14	ESTADISTICA Y PROBABILIDADES	11Q171	2	44 86.27%	0 0.00%	7 13.73%	51 100%
15	GESTION DE PROCESOS ORGANIZACIONALES	11Q172	2	55 93.22%	0 0.00%	4 6.78%	59 100%
16	BASE DE DATOS AVANZADAS	11Q206	3	24 80.00%	3 10.00%	3 10.00%	30 100%
17	E- MARKETING	11Q180	3	43 89.58%	3 6.25%	2 4.17%	48 100%
18	INGENIERIA DE SOFTWARE I	11Q179	3	33 68.75%	6 12.50%	9 18.75%	48 100%
19	OPTIMIZACION EN INGENIERIA II	11Q181	3	47 94.00%	2 4.00%	1 2.00%	50 100%
20	PROGRAMACION APLICADA I	11Q182	3	23 60.53%	7 18.42%	8 21.05%	38 100%
21	SISTEMAS OPERATIVOS	11Q183	3	33 75.00%	3 6.82%	8 18.18%	44 100%
22	TEORIA DE SISTEMAS	11Q184	3	38 95.00%	1 2.50%	1 2.50%	40 100%
23	ARQUITECTURA DE PROCESOS DE NEGOCIO	11Q191	4	37 97.37%	1 2.63%	0 0.00%	38 100%
24	INGENIERIA DE PROCESOS	11Q192	4	38 86.36%	1 2.27%	5 11.36%	44 100%
25	PLANEAMIENTO DE SISTEMAS DE INFORMACI	11Q193	4	24 77.42%	5 16.13%	2 6.45%	31 100%
26	PROGRAMACION AVANZADA	11Q207	4	40 88.89%	3 6.67%	2 4.44%	45 100%
27	REDES Y CONECTIVIDAD	11Q194	4	42 100.00%	0 0.00%	0 0.00%	42 100%
28	SISTEMAS DE INFORMACION INTEGRADOS	11Q195	4	29 93.55%	1 3.23%	1 3.23%	31 100%
29	GERENCIA EN TI	11Q201	5	49 96.08%	1 1.96%	1 1.96%	51 100%
30	GESTION DE RIESGOS Y SEGURIDAD DE LA IA	11Q202	5	49 76.56%	10 15.63%	5 7.81%	64 100%
31	LEGISLACION INFORMATICA	11Q203	5	47 97.92%	0 0.00%	1 2.08%	48 100%

Nota: Los cursos remarcados en el reporte , son los mayores o iguales al 50% de Desaprobados

CONSOLIDADO DE ALUMNOS (APROB. DESAPROB. INHAB.) POR CURSO

Universidad Nacional de Cajamarca
Oficina Técnica de Registro Central

Fecha Rep. 05/03/2015 09:36:03

Escuela: **11 INGENIERIA DE SISTEMAS**

Periodo 2014 - II

			NIV APROBADOS	DESAPROB.	INHAB.	TOTAL					
32	LIDERAZGO	11Q204	5	32	100.00%	0	0.00%	0	0.00%	32	100%
33	REDES CORPORATIVAS	11Q209	5	28	100.00%	0	0.00%	0	0.00%	28	100%
34	TESIS	11Q205	5	42	93.33%	0	0.00%	3	6.67%	45	100%

Nota: Los cursos remarcados en el reporte , son los mayores o iguales al 50% de Desaprobados

Pag. 2 de 2

CONSOLIDADO DE ALUMNOS(APROB. DESAPROB. INHAB.) POR CURSO

Universidad Nacional de Cajamarca
Oficina Técnica de Registro Central

Fecha Rep. 05/03/2015 09:36:25

Escuela: **11 INGENIERIA DE SISTEMAS**

Periodo 2014 - I

		NIVAPROBADOS		DESAPROB.		INHAB.		TOTAL			
1	FILOSOFIA DE LA CIENCIA Y LA TECNOLOGIA	11Q151	1	55	91.67%	0	0.00%	5	8.33%	60	100%
2	FISICA I	11Q149	1	56	62.92%	19	21.35%	14	15.73%	89	100%
3	INTRODUCCION A LA COMPUTACION INFORMA	11Q148	1	45	64.29%	16	22.86%	9	12.86%	70	100%
4	LENGUAJE Y COMUNICACION	11Q145	1	55	80.88%	1	1.47%	12	17.65%	68	100%
5	LOGICA	11Q146	1	56	77.78%	8	11.11%	8	11.11%	72	100%
6	MATEMATICA BASICA	11Q147	1	44	55.70%	24	30.38%	11	13.92%	79	100%
7	METODOS Y TECNICAS DE ESTUDIO	11Q150	1	52	88.14%	0	0.00%	7	11.86%	59	100%
8	ALGORITMOS Y ESTRUCTURA DE DATOS II	11Q159	2	40	70.18%	11	19.30%	6	10.53%	57	100%
9	ANALISIS MATEMATICO II	11Q160	2	29	48.33%	12	20.00%	19	31.67%	60	100%
10	ECOLOGIA Y MEDIO AMBIENTE	11Q161	2	47	85.45%	1	1.82%	7	12.73%	55	100%
11	ESTADISTICA GENERAL	11Q162	2	47	94.00%	0	0.00%	3	6.00%	50	100%
12	ORGANIZACION Y ADMINISTRACION DE EMPR	11Q163	2	54	85.71%	5	7.94%	4	6.35%	63	100%
13	PLANIFICACION ESTRATEGICA	11Q164	2	56	90.32%	3	4.84%	3	4.84%	62	100%
14	SISTEMA DIGITAL	11Q165	2	32	100.00%	0	0.00%	0	0.00%	32	100%
15	BASE DE DATOS II	11Q173	3	33	71.74%	10	21.74%	3	6.52%	46	100%
16	CULTURA Y REALIDAD NACIONAL	11Q174	3	29	85.29%	1	2.94%	4	11.76%	34	100%
17	GESTION DEL RECURSO HUMANO	11Q175	3	43	97.73%	0	0.00%	1	2.27%	44	100%
18	OPTIMIZACION EN INGENIERIA I	11Q176	3	49	92.45%	1	1.89%	3	5.66%	53	100%
19	SISTEMAS DE INFORMACION	11Q177	3	39	95.12%	2	4.88%	0	0.00%	41	100%
20	TEORIA DE LENGUAJES	11Q178	3	35	68.63%	5	9.80%	11	21.57%	51	100%
21	DINAMICA DE SISTEMAS	11Q185	4	38	92.68%	0	0.00%	3	7.32%	41	100%
22	GESTION DEL CONOCIMIENTO	11Q186	4	31	93.94%	0	0.00%	2	6.06%	33	100%
23	INGENIERIA DE SOFTWARE II	11Q187	4	24	75.00%	4	12.50%	4	12.50%	32	100%
24	PROGRAMACION APLICADA II	11Q188	4	39	82.98%	3	6.38%	5	10.64%	47	100%
25	SISTEMA DE COMUNICACION DE DATOS	11Q189	4	42	97.67%	0	0.00%	1	2.33%	43	100%
26	TEORIA DE DECISIONES	11Q190	4	33	97.06%	0	0.00%	1	2.94%	34	100%
27	ADMINISTRACION DE PROYECTOS DE SISTEM	11Q196	5	51	98.08%	0	0.00%	1	1.92%	52	100%
28	ADMINISTRACION DE REDES	11Q197	5	32	84.21%	1	2.63%	5	13.16%	38	100%
29	INGENIERIA WEB	11Q208	5	32	88.89%	2	5.56%	2	5.56%	36	100%
30	PLANEAMIENTO DE SISTEMAS DE INFORMACI	11Q198	5	47	83.93%	8	14.29%	1	1.79%	56	100%
31	SISTEMAS INTELIGENTES	11Q199	5	45	100.00%	0	0.00%	0	0.00%	45	100%

Nota: Los cursos remarcados en el reporte , son los mayores o iguales al 50% de Desaprobados

CONSOLIDADO DE ALUMNOS(APROB. DESAPROB. INHAB.) POR CURSO

Universidad Nacional de Cajamarca
Oficina Técnica de Registro Central

Fecha Rep. 05/03/2015 09:36:25

Escuela: **11 INGENIERIA DE SISTEMAS**

Periodo 2014 - I

		NIV	APROBADOS	DESAPROB.	INHAB.	TOTAL					
32	TOPICOS EN SISTEMAS	11Q200	5	38	84.44%	4	8.89%	3	6.67%	45	100%

Nota: Los cursos remarcados en el reporte , son los mayores o iguales al 50% de Desaprobados

Pag. 2 de 2

Plan de Estudios

Cursos obligatorios Primer Año: Primer semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Lenguaje y Comunicación	2	2	4	3	Idiomas y Literatura	No tiene
Lógica	2	3	5	3	Matemáticas	No tiene
Matemática Básica	2	3	5	3	Matemáticas	No tiene
Introducción a la Computación Informática	2	3	5	3	DASEI	No tiene
Física I	2	3	5	3	Física	No tiene
Métodos y Técnicas de Estudio	2	2	4	3	Ciencias de la Educación	No tiene
Filosofía de la Ciencia y la Tecnología	2	2	4	3	Ciencias Sociales	No tiene
Totales	14	18	32	21		

Cursos obligatorios Primer Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Contabilidad y Finanzas	2	2	4	3	Ciencias Contables y Administrativas	14 Créditos Aprobados
Dibujo y Geometría Descriptiva	2	2	4	3	Ciencias de la Ingeniería	Matemática Básica
Análisis Matemático I	3	3	6	4	Matemáticas	Matemática Básica
Algoritmos y Estructura de Datos I	3	3	6	4	DASEI	Introducción a la Computación Informática
Física II	2	3	5	3	Física	Física I
Ética	2	2	4	3	Derecho	14 Créditos Aprobados
Metodología de la Investigación Científica	2	3	5	3	Ciencias Sociales	Filosofía de la Ciencia y la Tecnología
Totales	16	18	34	23		

Cursos obligatorios Segundo Año: Primer Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Organización y Administración de Empresas	2	2	4	3	Ciencias Contables y Administrativas	Contabilidad y Finanzas
Estadística General	2	3	5	3	DASEI	Análisis Matemático I
Análisis Matemático II	3	3	6	4	Matemáticas	Análisis Matemático I
Algoritmos y Estructura de Datos II	3	3	6	4	DASEI	Algoritmos y Estructura de Datos I
Sistema Digital	2	3	5	3	DASEI	Física II

Planificación Estratégica	2	2	4	3	Ciencias Económicas	Contabilidad y Finanzas
Ecología y Medio Ambiente	2	2	4	3	Ciencias Biológicas	34 Créditos Aprobados
Totales	16	18	34	23		

Cursos obligatorios Segundo Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Gestión de Procesos Organizacionales	2	2	4	3	DASEI	Organización y Administración de Empresas
Estadística y Probabilidades	2	3	5	3	DASEI	Estadística General
Análisis Matemático III	3	3	6	4	Matemáticas	Análisis Matemático II
Base de Datos I	3	3	6	4	DASEI	Algoritmos y Estructura de Datos II
Arquitectura del Computador	2	3	5	3	DASEI	Sistema Digital
Defensa Nacional y Derechos Humanos	2	2	4	3	Derecho	57 Créditos Aprobados
Economía y Gestión Empresarial	2	2	4	3	Ciencias Contables y Administrativas	57 Créditos Aprobados
Totales	16	18	34	23		

Cursos obligatorios Tercer Año: Primer Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Gestión del Recurso Humano	2	2	4	3	Ciencias Contables y Administrativas	Gestión de Procesos Organizacionales
Optimización en Ingeniería I	2	3	5	3	DASEI	Estadística y Probabilidades
Sistemas de Información	3	2	5	4	DASEI	Base de Datos I
Base de Datos II	3	3	6	4	DASEI	Base de Datos I
Teoría de Lenguajes	2	3	5	3	DASEI	Arquitectura del Computador
Cultura y Realidad Nacional	2	2	4	3	Ciencias Sociales	67 créditos aprobados
Totales	14	15	29	20		

Cursos obligatorios Tercer Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
e - Marketing	2	2	4	3	DASEI	Gestión del Recurso Humano
Optimización en Ingeniería II	2	3	5	3	DASEI	Optimización en Ingeniería I
Ingeniería de Software I	3	3	6	4	DASEI	Sistemas de Información

Escuela Académico Profesional de Ingeniería de Sistemas - UNC

Programación Aplicada I	3	3	6	4	DASEI	Base de Datos II
Sistemas Operativos	2	3	5	3	DASEI	Teoría de Lenguajes
Teoría de Sistemas	2	2	4	3	DASEI	Cultura y Realidad Nacional
Totales	14	16	30	20		

Curso Electivo Tercer Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Base de Datos Avanzadas	2	2	4	3	DASEI	Base de Datos II
Totales	2	2	4	3		

Cursos obligatorios Cuarto Año: Primer Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Gestión del Conocimiento	2	2	4	3	DASEI	e - Marketing
Teoría de Decisiones	2	3	5	3	DASEI	Optimización en Ingeniería II
Ingeniería del Software II	3	3	6	4	DASEI	Ingeniería de Software I
Programación Aplicada II	3	3	6	4	DASEI	Programación Aplicada I
Sistemas de Comunicación de Datos	3	2	5	4	DASEI	Sistemas Operativos
Dinámica de sistemas	2	3	5	3	DASEI	Teoría de Sistemas
Totales	15	16	31	21		

Cursos Obligatorios Cuarto Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Sistemas de Información Integrados	2	3	5	3	DASEI	Gestión del Conocimiento
Ingeniería de Procesos	2	3	5	3	DASEI	Teoría de Decisiones
Planeamiento de sistemas de Información I	3	3	6	4	DASEI	Ingeniería del Software II
Arquitectura de Procesos de negocio	3	3	6	4	DASEI	Programación Aplicada II
Redes y Conectividad	3	3	6	4	DASEI	Sistemas de Comunicación de Datos
Totales	13	15	28	18		

Curso Electivo Cuarto Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Programación Avanzada	2	2	4	3	DASEI	Programación Aplicada II
Totales	2	2	4	3		

Cursos Obligatorios Quinto Año: Primer Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Administración de Proyectos de Sistemas	2	3	5	3	DASEI	Sistemas de Información Integrados
Sistemas Inteligentes	2	3	5	3	DASEI	Ingeniería de Procesos
Planeamiento de Sistemas de Información II	3	3	6	4	DASEI	Planeamiento de sistemas de Información I
Tópicos en Sistemas	3	3	6	4	DASEI	Arquitectura de Procesos de negocio
Administración de Redes	3	3	6	4	DASEI	Redes y Conectividad
Totales	14	14	27	18		

Curso Electivo Quinto Año: Primer Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Ingeniería Web	2	2	4	3	DASEI	Arquitectura de Procesos de negocio
Totales	2	2	4	3		

Cursos obligatorios Quinto Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Gestión de riesgos y Seguridad de la Información	3	3	6	4	DASEI	Administración de Proyectos de Sistemas
Liderazgo	2	2	4	3	Ciencias Contables y Administrativas	164 Créditos Aprobados
Gerencia en TI	2	3	5	3	DASEI	Planeamiento de Sistemas de Información II
Legislación Informática	2	2	4	3	Derecho	187 Créditos Aprobados
Tesis	2	3	5	3	DASEI	187 Créditos Aprobados
Totales	11	13	24	16		

Curso Electivo Quinto Año: Segundo Semestre						
Curso	HT	HP	TH	CR	Dpto. Académico	Prerrequisito
Redes Corporativas	2	2	4	3	DASEI	Administración de Redes
Totales	2	2	4	3		

CREDITOS	TOTAL
Créditos obligatorios	203
Créditos electivos	12
Total de créditos	215

ANEXO 6

R-10 [Estado de alumnos por docente]

SANDRA CECILIA RODRIGUEZ AVILA

APROBADOS



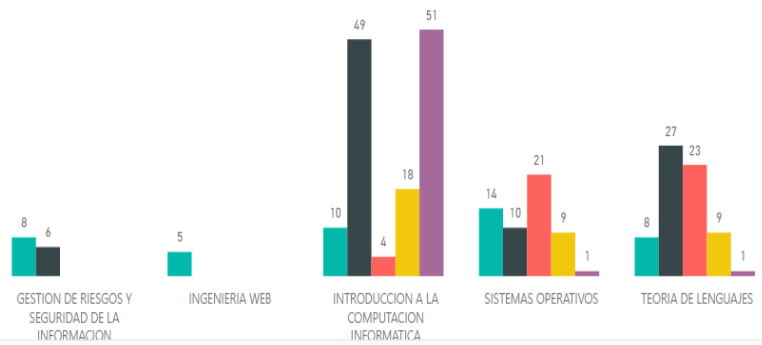
PROMOCION ● 10 ● 11 ● 12 ● 13 ● 14

DOCENTE

- AIDE YANET HUAMAN RIVERA
- ALEJANDRO ANGULO BADA
- ALEJANDRO CHACHA CORTEZ
- ALEJANDRO QUISPE SILVA
- AMALIA DELICIA DEL SAGRARIO FERNANDEZ VARGAS
- ANA CECILIA SARMIENTO LONGO
- ARACELLI MARIBEL POEMAPE GRADOS
- CARLOS ALFONSO PEREZ CERNA
- CARLOS ENRIQUE APARICIO ARTEAGA
- CARLOS ENRIQUE MORENO HUAMAN
- CARLOS JESUS KOO LABRIN
- CARMEN GLORIA CASTILLO DIAZ

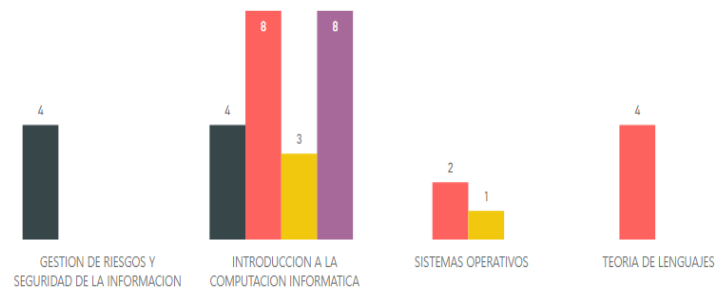
CURSO

- GESTION DE RIESGOS Y SEGURIDAD DE LA INFORMACION
- INGENIERIA WEB
- INTRODUCCION A LA COMPUTACION INFORMATICA
- SISTEMAS OPERATIVOS
- TEORIA DE LENGUAJES



DESAPROBADOS

PROMOCION ● 11 ● 12 ● 13 ● 14



CAPÍTULO I INTRODUCCIÓN

1.1. Planteamiento del problema

Desde el inicio de la era de la información las empresas necesitan explotar su mayor recurso, la información.

La explotación eficiente de la información permite una rápida, acertada y oportuna toma de decisiones bajo el manejo de datos confiables. Por esta razón muchas organizaciones han logrado implementar un Data Mart que permita centralizarla información útil y necesaria para ayudar a la toma de decisiones en los siguientes niveles organizacionales: Nivel operativo, Nivel táctico y Nivel estratégico.[1]

A nivel internacional La Escuela Politécnica Nacional de Quito[2], La Universidad Carlos III de Madrid Escuela Politécnica Superior Ingeniería en Informática[3] y la Universidad de San Carlos de Guatemala[4], hacen uso de grandes cantidades de información proveniente de diferentes fuentes para poder llevar un buen funcionamiento estas escuelas apoyándose de que centraliza su información para poder tomar decisiones o identificar posibles falencias en el funcionamiento de estas, de esta manera logran “formar profesionales de alto nivel”.

En la mayoría de las Universidades de Perú, no existen herramientas que ayuden a minimizar el tiempo para analizar grandes cantidades de información con mayor velocidad y precisión y de esta manera poder tomar decisiones.

A nivel local veremos cómo se lleva el análisis de la información en Universidades de Cajamarca, la Universidad Privada Antonio Guillermo Urrelo (UPAGU) cuenta con un sistema de registros de alumnos y notas según el curso el cual se mantiene actualizado de acuerdo a como vayan obteniendo notas los alumnos, en este sistema también es posible ver la malla curricular y creditaje de los cursos de igual manera pasa en la Universidad Privada del Norte (UPN).

En la escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca el problema principal radica en que el

proceso de extracción e integración de los datos para la realización de los reportes se realiza manualmente y sistemáticamente, resultando muy complicado, provocando retraso en la atención de requerimientos haciendo que en muchos casos la entrega de la información no sea oportuna; por otro lado la elaboración de algunos reportes requiere mayor conocimiento en el manejo de las herramientas informáticas, por lo que son derivados al Unidad Técnica de Sistemas Informáticos, ocasionan pérdida de tiempo ya que esto demanda tramites documentarios a través de autoridades de las diferentes oficinas involucradas.

1.2. Formulación del problema

¿La implementación de un Data Mart permitirá realizar el seguimiento académico de los estudiantes de la Escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca?

1.3. Justificación de la investigación

Con los antecedentes anteriormente expuestos, la solución al problema sirve como apoyo a la Escuela Académico Profesional de Ingeniería de Sistemas para cubrir sus necesidades de mejoramiento en el manejo de la información de modo que pueda ser más eficiente la toma de decisiones organizacionales de nivel operativo y estratégico.

La importancia del presente trabajo de investigación radica en agilizar y automatizar todo el proceso de gestión de la información con el objetivo de brindar un servicio de calidad con una herramienta que cumpla las expectativas de la Escuela, con ello eliminar las deficiencias en tiempo en el entorno administrativo y técnico con un manejo rápido y eficaz de la información, para ir acorde con el crecimiento de la misma.

Además de poder encontrar posibles errores que se estén cometiendo lo cual afecta de manera directa al alumnado, docentes y administrativos de la Escuela y tomar decisiones de acuerdo a los resultados obtenidos, de esta manera mejorar para alcanzar la acreditación académica.

1.4. Alcances o delimitación de la investigación

1.4.1. Alcances

- Se creará el Data Mart que contenga la toda la información académica, involucrada, con los alumnos de la Escuela Académico

Profesional de Ingeniería de Sistemas para la creación de los reportes que solicite el director de Escuela.

- Implementar sobre la herramienta de Inteligencia de Negocios la solución para definir y administrar la información relacionada con los estudiantes y docentes de la Escuela.
- Como resultado se obtendrán los modelos dimensionales que permitirán la creación de Cubos, que serán la fuente de los reportes que muestren la información para la toma de decisiones estratégicas de la Escuela.
- Se entregará toda la documentación referente a este proyecto de tesis para que la Escuela pueda continuar cuando crea conveniente con el proceso de implementación.

1.4.2. Limitaciones

- La disponibilidad de obtener los datos requeridos para la construcción del Data Mart en la Unidad Técnica de Sistemas Informáticos de la Universidad Nacional de Cajamarca.
- La plataforma sobre la cual se va implementar la herramienta de Inteligencia de Negocios es en base a la plataforma actual con la que trabaja la Unidad de Informática el cual es pagado.
- El desarrollo de la Aplicación se realizó fuera del horario laboral, lo que limitó el trato directo y constante con el usuario.

1.5. Objetivos

Objetivo general

Implementar un Data Mart usando tecnología BI para la Escuela Académico Profesional de Ingeniería de Sistemas para realizar el seguimiento a los estudiantes y formular estrategias de mejora de manera más acertada y rápida.

Objetivos específicos

- Definir los requerimientos de parte de la Escuela Académico Profesional de Ingeniería de Sistemas en el seguimiento académico del proceso educativo.
- Diseñar y cargar de datos una estructura dimensional que permita una salida del análisis de datos obtenidos del Data Mart.

- Crear reportes e indicadores para la toma de decisiones de la escuela.
- Determinar la utilidad de obtener la información consolidada del rendimiento académico de los estudiantes usando un Data Mart.

CAPÍTULO II MARCO TEÓRICO

2.4. Antecedentes Teóricos

Miguel Rodríguez Sanz (2010) en su proyecto fin de carrera, análisis y diseño de un data mart para el seguimiento académico de los alumnos de un entorno universitario; este proyecto consiste ofrecer un marco metodológico para el desarrollo de proyectos de Data Ware house y/o Data Mart, así como una aplicación práctica de una metodología a un caso de estudio concreto mediante la implementación de un data mart que permita responder a las necesidades de negocio del área académica de una universidad que ayudará a extraer la información acerca del paso de los alumnos por su evaluación universitaria, lo cual ha permitido obtener una mejor visión de los sucesos en el ámbito universitario así como dar apoyo a la toma de decisiones con el fin de conocer la respuesta de los diferentes alumnos a los estudios que han comenzado.[3]

Aimacaña Quilumba Doris Eulalia (2013), en su trabajo de graduación previo a la obtención del título de Ingeniero Informático, Análisis, Diseño e Implementación de una Data Mart Académico, usando tecnología de BI tanto software libre, para la Facultad de Ingeniería, Ciencias Físicas y Matemática, el trabajo consiste en implementar un data mart que permita almacenar sólo la información requerida e ir eliminando aquellos datos que obstaculizan la labor del análisis y entregando la información que se requiera en la forma más apropiada, facilitando así el proceso de gestión de la misma.

A través de esto surge la promesa del manejo y control de la información, aseguran una vista única de los datos, que pueden provenir de diversas fuentes para la toma de decisiones estratégicas tanto de estudiantes como profesores, esto es muy importante para saber dónde y cómo solventar alguna falla, o mejorar las decisiones tomadas.[5]

Tana Pasquel Gloria Estefanía (2014), en su trabajo previo a la obtención del título de Ingeniería en Sistemas Computacionales,

Datamart para el Análisis de información del sistema Académico de la Universidad Técnica del Norte con Herramientas de Software Libre, en el cual la implementación de un prototipo de datamart para el análisis de información se lo hizo con la finalidad de demostrar que el uso de herramientas de software libre en la actualidad está en auge en las pequeñas y medianas empresas, el prototipo cumplió con diferentes etapas que fueron realizadas con las distintas herramientas que ofrece la Suite de Pentaho, la información del datamart está almacenada en una base de datos Postgre SQL que también es de software libre.

A lo largo del desarrollo del trabajo de grado se pudo observar que, la implementación de datamart con herramientas de software libre es muy confiable y segura, obteniendo estabilidad al momento de integrar los datos, logrando buenos resultados.[6]

Montalvo Barrientos, Miguel Ángel (2010), en su trabajo, Implementación de un Data mart para la Facultad Nacional de Ingeniería Caso: Ing. De Sistemas en la Universidad Técnica de Oruro, que desarrolló un producto de Data Mart independiente, soportado por una base de datos multidimensional, se ha conseguido implementar una herramienta de análisis gráfico que permita visualizar el comportamiento de un fenómeno cualquiera asociado con diferentes causas de la variación de tiempos de la permanencia estudiantil o cualquier otro fenómeno en la Carrera de Ing. de Sistemas.

Como conclusión se ha conseguido integrar la información necesaria en un Data Mart el cual es soportado por una herramienta web de análisis OLAP, que a percepción del usuario consigue una mejoría en el análisis de información.[7]

2.5. Bases Teóricas

2.5.1. Data Mart

Un Datamart es un subconjunto de datos derivado del Data Warehouse. Está diseñado para soportar requerimientos analíticos específicos de una determinada unidad de negocios.

Un Datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de

datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un datamart puede ser alimentado desde los datos de un data warehouse, o integrar por sí mismo un compendio de distintas fuentes de información.[8]

Un Data mart es un gran almacén de datos de una organización referente a un área, un tema o una función específica de una organización de la cual se pueden hacer consultas rápidas en un nivel más pequeño que un data warehouse, es decir, es un repositorio menos ambicioso que un DWH.

2.5.2. Business Intelligence (BI)

BI es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un data warehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones.

El proceso de Business Intelligence incluye la comunicación de los descubrimientos y efectuar los cambios.

Las áreas incluyen clientes, proveedores, productos, servicios y competidores.[9]

Business Intelligence o Inteligencia de Negocios tiene como objetivo apoyar a las organizaciones para mejorar su competitividad, facilitando la información necesaria a quienes la necesitan en el momento oportuno para posibilitar una mejor y más fundamentada toma de decisiones.

Mediante el uso de tecnologías y las metodologías de BI pretendemos convertir datos en información útil y relevante, a partir de esta ser capaces de descubrir conocimiento.

Posee las siguientes características:

- **Accesibilidad a la información:** Los datos son la fuente principal de este concepto. Lo primero que deben garantizar las herramientas y técnicas será el acceso de los usuarios a los datos con independencia de la procedencia de éstos.
- **Apoyo en la toma de decisiones:** Se busca ir más allá en la presentación de la información, de manera que los usuarios

tengan acceso a herramientas de análisis que les permitan seleccionar y manipular sólo aquellos datos que les interesen.

2.5.3. Rendimiento Académico

"Del latín reddere (restituir, pagar) el rendimiento es una relación entre lo obtenido y el esfuerzo empleado para obtenerlo. Es un nivel de éxito en la universidad, en el trabajo, etc.",

En otras palabras, el rendimiento académico es una medida de las capacidades del alumno, que expresa lo que éste ha aprendido a lo largo del proceso formativo. También supone la capacidad del alumno para responder a los estímulos educativos. En este sentido, el rendimiento académico está vinculado a la aptitud.[10]

Según Herán y Villarroel (1987). El rendimiento académico se define en forma operativa y tácita afirmando que se puede comprender el rendimiento previo como el número de veces que el estudiante ha repetido uno o más cursos.

En tanto Nováez (1986) sostiene que el rendimiento académico es el resultado obtenido por el individuo en determinada actividad académica. El concepto de rendimiento está ligado al de aptitud, y sería el resultado de ésta, de factores volitivos, afectivos y emocionales, además de la ejercitación.

Chadwick (1979) define el rendimiento académico como la expresión de capacidades y de características psicológicas del estudiante desarrolladas y actualizadas a través del proceso de enseñanza-aprendizaje que le posibilita obtener un nivel de funcionamiento y logros académicos a lo largo de un período, año o semestre, que se sintetiza en un calificativo final (cuantitativo en la mayoría de los casos) evaluador del nivel alcanzado.

2.5.4. Datos

En el mundo de la informática, la unidad básica en la que trabajamos es definida como Dato, teniendo este concepto una aserción bastante general como un caracter, una Representación

o un Símbolo que se encuentra en forma aislada, sin un contexto determinado y sin un debido ordenamiento.

Un dato es una representación simbólica (numérica, alfabética, algorítmica, espacial, etc.) de un atributo o variable cuantitativa o cualitativa. Los datos describen hechos empíricos, sucesos y entidades.

Es por eso que se establece que los datos aislados tienen la necesidad de una interpretación, dándoles un contexto y un orden que permita una interpretación y un resultado único.

2.5.5. Información

Según Idalberto Chiavenato, información "es un conjunto de datos con un significado, o sea, que reduce la incertidumbre o que aumenta el conocimiento de algo. En verdad, la información es un mensaje con significado en un determinado contexto, disponible para uso inmediato y que proporciona orientación a las acciones por el hecho de reducir el margen de incertidumbre con respecto a nuestras decisiones" [11].

Para Ferrell y Hirt, la información "comprende los datos y conocimientos que se usan en la toma de decisiones"[12].

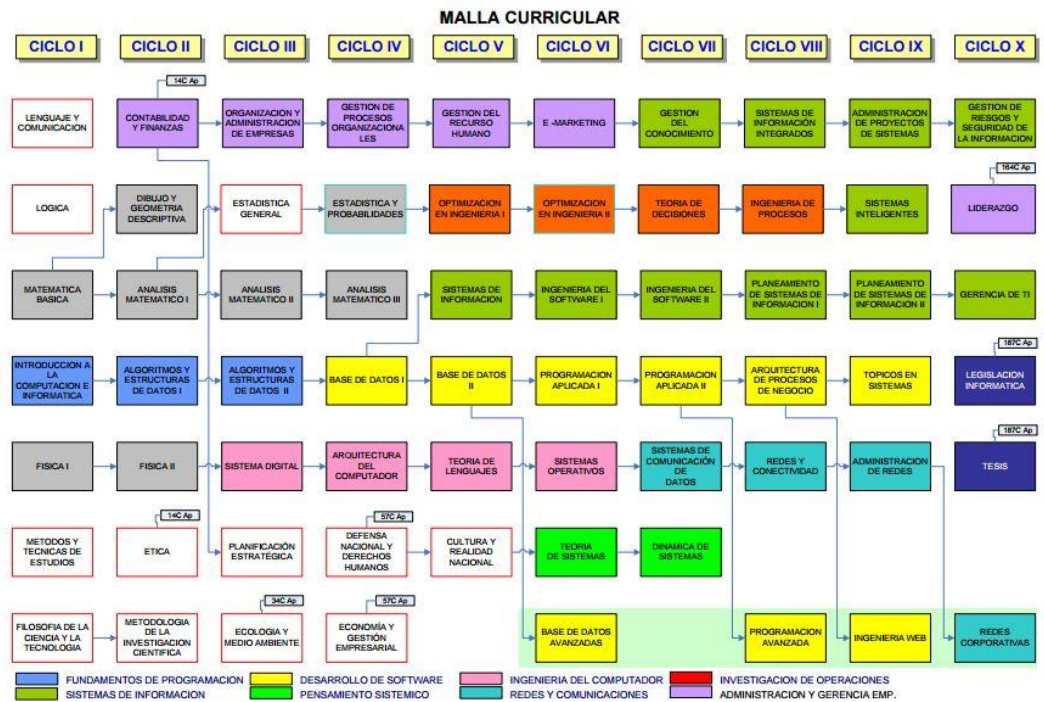
Según Czinkota y Kotabe la información "consiste en datos seleccionados y ordenados con un propósito específico"[13].

Entonces podemos concluir que los datos organizados e interpretados son considerados como Información.

2.5.6. Proceso Educativo

El proceso educativo se basa en la transmisión de valores y saberes. El proceso educativo no suele ser unidireccional, sino que es interactivo: quienes están aprendiendo, también pueden enseñar. Así el conocimiento se construye de forma social.[14]

La acción sinérgica de la Comunidad Educativa que gestiona dinámica, corresponsable y pertinentemente elementos curriculares, planificativos y administrativos, para el desarrollo integral del estudiante a continuación la malla curricular de la Escuela Académico Profesional de Ingeniería de Sistemas.



CAPÍTULO III HIPÓTESIS DE INVESTIGACIÓN

3.1. Hipótesis general

La implementación de un data mart permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela Académico de Ingeniería de Sistemas.

3.2. Hipótesis específica

Con el uso del data mart se evitarán los procesos de solicitud de información a la Unidad Técnica de Sistemas Informáticos.

Con el uso del data mart se dispondrá de manera inmediata a los reportes con la información que se necesita del alumnado.

Los reportes generados a partir del data mart permitirán observar el rendimiento académico de los alumnos y contrastar en el tiempo.

Con la información mostrada en los reportes se podrá efectuar la toma de decisiones por los docentes del comité directivo para la mejora académica.

Con el Data mart se obtendrá facilidad de acceso a la información para el control y toma de decisiones de la carrera y mejor tiempo de respuesta.

3.3. Definición de variables

3.3.1. Variable Independiente

Implementación de un Data Mart

3.3.2. Variable Dependiente

Seguimiento Académico de estudiantes

3.4. Operacionalización de variables

Variables	Indicadores	Escala de Medición
Variable Dependiente Seguimiento Académico de estudiantes	Número de alumnos aprobados y desaprobados por curso.	Razón
	Número de veces que un alumno ha llevado un curso.	Razón
Variable Independiente Implementación de un Data Mart	Grado de dificultad para la generación de reportes.	Ordinal
	Tiempo promedio utilizado en la construcción de reportes.	Razón

3.5. Matriz de consistencia

IMPLEMENTACION DE UN DATA MART PARA EL SEGUIMIENTO ACADÉMICO DE LOS ESTUDIANTES EN LA ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD NACIONAL DE CAJAMARCA			
DEFINICIÓN DEL PROBLEMA	OBJETIVOS DE LA INVESTIGACIÓN	HIPÓTESIS DE LA INVESTIGACIÓN	VARIABLES DE ESTUDIO
<p>Problema Principal ¿La implementación de un Data Mart sí permitirá realizar el seguimiento académico de los estudiantes de la Escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca?</p>	<p>Objetivo General Implementar un Data Mart usando tecnología BI para la Escuela Académico Profesional de Ingeniería de Sistemas para que pueda realizar el seguimiento a los estudiantes y formular estrategias de mejora de manera más acertada y rápida.</p> <p>Objetivos Específicos</p> <ul style="list-style-type: none"> - Definir una estructura dimensional que permita una salida del análisis de datos obtenidos del Data Mart. - Crear reportes e indicadores para la toma de decisiones de la escuela. - Determinar la utilidad de obtener la información consolidada del rendimiento académico de los estudiantes usando un Data Mart. 	<p>Hipótesis General La implementación de un data mart si permitirá realizar el seguimiento y análisis de la información académica de los alumnos de la Escuela de Ingeniería de Sistemas.</p>	<p>Variable Independiente: Implementación de un Data Mart</p> <p>Variable Dependiente: Seguimiento Académico de estudiantes</p>

CAPÍTULO IV METODOLOGÍA

4.1. Tipo, nivel, diseño y método de investigación

4.1.1. Tipo de Investigación

Aplicada, en razón a que se implementará un data mart para realizar el seguimiento académico de los estudiantes de la Escuela Académico Profesional de Ingeniería de Sistemas lo que apoyará a la toma de decisiones.

4.1.2. Nivel de Investigación

Correlacional, debido a que la variable independiente, el data mart, y la variable dependiente, seguimiento académico, se correlacionan entre sí, es decir el grado en que las variaciones que sufre un factor se corresponden con las que experimenta el otro.

4.1.3. Diseño de Investigación

Experimental, debido a que se manipula intencionalmente la variable independiente para observar los efectos en la variable dependiente. Utilizaremos el Pre – Experimental. Consiste en administrar un estímulo o tratamiento a un grupo y después aplicar una medición en una o más variables para observar cual es el nivel del grupo en estas variables.

4.1.4. Método de Investigación

El desarrollo de la investigación se realizará con los métodos Inductivo – Deductivo observando rasgos generales y específicos del problema, así como también el Método de análisis para poder establecer la relación causa – efecto entre los elementos que compone el objeto de investigación.

4.2. Población de estudio

Para esta investigación la población estará conformada por las cuatro últimas promociones de la Escuela Académico Profesional de Ingeniería de Sistemas de la Universidad Nacional de Cajamarca.

4.3. Muestra

La muestra es igual a 181 alumnos de la Escuela de Académico Profesional de Ingeniería de Sistemas que fue obtenida mediante la fórmula de muestra.

ERROR (e)	3.0%
TAMAÑO POBLACIÓN (N)	218
NIVEL DE CONFIANZA (Z)	95%

TAMAÑO DE LA MUESTRA = 181

$$\frac{N * (\alpha_c * 0,5)^2}{1 + (e^2 * (N - 1))}$$

4.4. Unidad de análisis

La unidad de análisis será la información procesada para el seguimiento académico de los alumnos de la Escuela Académico Profesional de Ingeniería de Sistemas.

4.5. Técnicas e instrumentos de recolección de datos

TÉCNICA	INSTRUMENTO	FUENTE	OBJETIVO
Encuesta	Cuestionario	Administrativos de la Escuela Académico Profesional de Ingeniería de Sistemas	Determinar la satisfacción del usuario, con respecto a la información procesada por el datamart. Determinar el grado de dificultad en la generación de reportes.
Observación Directa	Procedimientos	Fuentes de información para la gestión académica de los estudiantes.	Identificar los tiempos de acceso a la información. Establecer los tiempos de respuesta.

4.6. Análisis e interpretación de datos

Para procesar los datos, en el presente trabajo de investigación, se hará uso de la Distribución Normal.

$$X \rightarrow N(\mu, \sigma)$$

$$Z \rightarrow N(0,1)$$

$$P(X \leq \alpha) = P\left(Z \leq \frac{\alpha - \mu}{\sigma}\right)$$

CAPÍTULO V ASPECTOS ADMINISTRATIVOS

5.1. Recursos y presupuesto

5.1.1. Recursos

Para el desarrollo de la presente investigación se requieren los siguientes recursos:

Recursos Humanos

Recurso	Unidad de Medida
Bach. Durán Rafael, Liliana Elizabeth	Tesista
Ing. Jaime Meza Huamán	Asesor

Materiales e Insumos

Recurso	Unidad de Medida
Lapiceros	Unidad
Papel bond A4	Millar
Cartucho Negro y Color	Unidad
Folder Manila	Unidad
Papel A3	Millar
DVD	Unidad

Hardware y Software

Clasificación	Recurso
Hardware	Laptop
	Impresora
Software	CA Erwin Data Modeler
	ReportingServices
	SQL Server

Servicios

Servicios	Unidad de Medida
Internet	Horas
Consumo Eléctrico	Kw/hora x 5 meses
Movilidad	Días
Anillado	Unidad
Empastado	Unidad

5.1.2. Presupuesto

Clasificación	Recurso	Concepto/Unidad Medida	Cantidad	Precio Unitario S/.	Total S/.
Recursos Humanos	Bach. Durán Rafael, Liliana Elizabeth	Tesista	5	0.00	0.00
	Ing. Jaime Meza Huamán	Asesor	5	0.00	0.00
	Total S/.				0.00
Materiales e Insumos	Lapiceros	Unidad	10	0.50	5.00
	Papel bond A4	Millar	5	20.00	100.00
	Cartucho Negro y Color	Unidad	1	80.00	80.00
	Folder Manila	Unidad	10	0.50	5.00
	Papel A3	Unidad	10	1.00	10.00
	DVD	Unidad	25	1.00	25.00
	Total S/.				225.00
Servicios	Internet	Horas	400	1.00	400.00
	Consumo Eléctrico	Kw/hora x 5 meses	400	0.80	320.00
	Movilidad	Días	150	2.00	300.00
	Anillado	Unidad	10	3.00	30.00
	Empastado	Unidad	4	3.50	14.00
	Total S/.				1064.00
Hardware	Laptop	Unidad	1	2000.00	2000.00
	Impresora	Unidad	1	580.00	580.00
	Total S/.				3688.00
TOTAL S/.					4977.00

5.2. Financiamiento

El presente trabajo de investigación será asumido por el investigador en su totalidad.