



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE COMPUTACIÓN
CENTRO DE INVESTIGACIÓN EN SISTEMAS DE INFORMACIÓN

**DESARROLLO DE UNA SOLUCIÓN DE
INTELIGENCIA DE NEGOCIO QUE APOYE
A LA TOMA DE DECISIONES EN EL
SERVICIO DE ORIENTACIÓN Y CONTROL
DE ESTUDIOS DE LA FACULTAD DE
CIENCIAS**

Trabajo Especial de Grado presentado ante la ilustre
Universidad Central de Venezuela por la

Br. Emma Guevara
C.I.: V-19.819.577

Para optar por el título de
Licenciada en Computación

Tutor: Prof. Concettina Di Vasta

Caracas, Mayo de 2016

ACTA

Quienes suscriben, miembros del Jurado designado por el Consejo de Escuela de Computación, para examinar el Trabajo Especial de Grado presentado por la Br. Emma María Guevara Salas C.I. 19.819.577, con el título "DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO QUE APOYE A LA TOMA DE DECISIONES EN EL SERVICIO DE ORIENTACIÓN Y CONTROL DE ESTUDIOS DE LA FACULTAD DE CIENCIAS" a los fines de optar al título de Licenciada en Computación, dejan constancia de lo siguiente:

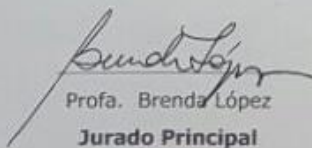
Dicho trabajo, leído por cada uno de los miembros del jurado, se fijó el día 30 de Mayo de 2016, a las 10 a.m., para que su autora lo defendiera en forma pública en el aula PB III de la Escuela de Computación, mediante una presentación oral de su contenido, luego de lo cual se respondieron las preguntas formuladas. Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobarlo con la nota de 20 puntos.

En fe de lo cual se levanta la presente Acta, en Caracas a los 30 días del mes de Mayo del año 2016.

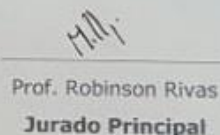


Prof. Concettina Di Vasta

Tutora



Prof. Brenda López
Jurado Principal



Prof. Robinson Rivas
Jurado Principal

RESUMEN

El objetivo de este Trabajo Especial de Grado (TEG) es desarrollar una solución de inteligencia de negocio que soporte los indicadores relacionados con el rendimiento académico y con la atención que presta el Servicio de Orientación de la Facultad de Ciencias de la Universidad Central de Venezuela con la finalidad de mejorar la prosecución estudiantil y apoyar a la toma de decisiones.

Las tecnologías utilizadas en el desarrollo de la solución de inteligencia de negocio fueron las siguientes: Pentaho Data Integration, una herramienta que permite obtener información de diversas fuentes de datos, consolidarla y cargarla en un único repositorio de datos, también se usó Pentaho Schema Workbench, un motor de análisis del negocio de código abierto que permite a organizaciones de cualquier tamaño darle a los usuarios acceso a los datos de manera interactiva y Pentaho Business Analytics Platform, una aplicación que permite transformar todos los datos en información significativa facilitando así el análisis de ésta y brindando apoyo a la toma de decisiones; también fue utilizado PostgreSQL, un sistema manejador de base de datos y como método de desarrollo de un almacén de datos, el ciclo de vida Dimensional del Negocio según (Kimball, 2002).

Palabras Clave: Inteligencia, Negocio, Pentaho, Rendimiento, Académico.

ÍNDICE GENERAL

Resumen.....	iii
Índice General	iv
Índice de Figuras.....	vi
Índice de Tablas	viii
Introducción	3
Capítulo 1: Problema de Investigación.....	4
1.1 Planteamiento del Problema.....	4
1.2 Objetivos	5
1.2.1 Objetivo General.....	5
1.2.2 Objetivos Específicos	5
1.3 Solución Propuesta.....	6
1.4 Método de Desarrollo de un Almacén de Datos Seleccionado.....	9
1.5 Justificación de la Solución Propuesta.....	11
1.6 Alcance de la Solución Propuesta	12
Capítulo 2: Marco Conceptual.....	13
2.1 Servicio de Orientación.....	13
2.2 Rendimiento Académico.....	14
2.3 Sistemas de Información.....	15
2.3.1 Definición de Sistemas de Información.....	15
2.3.2 Tipos de Sistemas de Información	15
2.3.3 Sistema OLAP	19
2.3.4 OLTP vs. OLAP	20
2.4 Inteligencia de Negocio.....	21
2.4.1 Características de una solución de inteligencia de negocio.....	21
2.4.2 Funciones de una solución de inteligencia de negocio	21
2.4.3 Arquitectura de una solución de inteligencia de negocio	22
2.4.4 Ventajas y Desventajas de una solución de inteligencia de negocio	31
2.5 Herramientas para el desarrollo de una solución de inteligencia de negocio	32
2.5.1 Pentaho Community Edition	32
2.6 Indicadores.....	34
2.6.1 Definición de Indicador	34
2.6.2 Tipos de Indicador.....	34

2.6.3	Importancia de los Indicadores	34
2.6.4	Patrones para Especificación de Indicadores	35
Capítulo 3: Marco Aplicativo.....		36
3.1	Planificación del Proyecto	36
3.2	Definición de Requisitos del Negocio.....	38
3.3	Selección de Productos de Software.....	40
3.4	Diseño Lógico.....	44
3.4.1	Diseño Lógico del Área de Almacenamiento Intermedio	44
3.4.2	Diseño Lógico del Datamart.....	46
3.5	Diseño Físico.....	48
3.5.1	Diseño Físico del Área de Almacenamiento Intermedio	48
3.5.2	Diseño Físico del Datamart.....	48
3.6	Diseño de Procesos de Extracción, Transformación y Carga	49
3.6.1	Diseño de Procesos de Extracción, Transformación y Carga del Área de Almacenamiento Intermedio.....	49
3.6.2	Diseño de Procesos de Extracción, Transformación y Carga del Datamart	51
3.6.3	Esquema de Extracción.....	53
3.7	Especificación de Aplicación de Usuario.....	54
3.7.1	Especificación de los Perfiles de Usuario	54
3.7.2	Diseño de Cubos	55
3.7.3	Reportes de la solución de inteligencia de negocio.....	57
3.8	Integración y Despliegue	74
3.9	Gestión del Proyecto	74
Conclusiones y Recomendaciones.....		79
Referencias Bibliográficas		82
Anexos.....		83
A.	Diseño Físico del Área de Almacenamiento Intermedio	83
B.	Diseño Físico del Datamart	87
C.	Diseño de Procesos de Extracción, Transformación y Carga del Área de Almacenamiento Intermedio	96
D.	Diseño de Procesos de Extracción, Transformación y Carga del Datamart.....	111

ÍNDICE DE FIGURAS

Figura 1: Arquitectura Propuesta	6
Figura 2: Ciclo de Vida Dimensional del Negocio	9
Figura 3: Adaptación del Ciclo de Vida Dimensional del Negocio de Kimball	10
Figura 4: Tipos de Sistemas de Información.....	16
Figura 5: Arquitectura de una solución de inteligencia de negocio	23
Figura 6: Ejemplo de Almacén de Datos y Datamart	25
Figura 7: Ejemplo de Tabla de Hechos.....	26
Figura 8: Ejemplo de Dimensión	27
Figura 9: Ejemplo de un Esquema Estrella	28
Figura 10: Ejemplo de un Esquema Copo de Nieve	29
Figura 11: Ejemplo de un Esquema Constelación de Estrellas	30
Figura 12: Interfaz de PostgreSQL	40
Figura 13: Interfaz de Pentaho Data Integration	41
Figura 14: Interfaz de Mondrian Schema Workbench	42
Figura 15: Interfaz de Pentaho Business Analytics Platform	43
Figura 16: Modelo Relacional del Proceso de Negocio "Servicio de Orientación"	44
Figura 17: Modelo Relacional del Proceso de Negocio "Rendimiento Académico"	45
Figura 18: Modelo Dimensional del Proceso de Negocio "Servicio de Orientación"	46
Figura 19: Modelo Dimensional del Proceso de Negocio "Rendimiento Académico"	47
Figura 20: Proceso ETC utilizado para Poblar la Tabla periodo_academico.....	50
Figura 21: Cubo "Servicio de Orientación".....	55
Figura 22: Cubo "Historial Académico"	56
Figura 23: Interfaz de inicio de sesión de la herramienta de análisis y consulta.....	57
Figura 24: Interfaz inicial de la herramienta de análisis y consulta	58
Figura 25: Cantidad de Estudiantes Atendidos por Área de Atención.....	59
Figura 26: Unidades de Crédito y Asignaturas Inscritas por Período Académico.....	60
Figura 27: Unidades de Crédito y Asignaturas Cursadas por Tipo de Examen.....	61
Figura 28: Unidades de Crédito y Asignaturas Retiradas por Licenciatura.....	62
Figura 29: Unidades de Crédito y Asignaturas Aprobadas por Periodo Académico	63
Figura 30: Unidades de Crédito y Asignaturas Aprobadas por Equivalencia por Licenciatura.....	64
Figura 31: Unidades de Crédito y Asignaturas Aprobadas en General por Nota.....	65

Figura 32: Unidades de Crédito y Asignaturas Aplazadas por Tipo de Examen	66
Figura 33: Ratio de Reprobación por Período Académico.....	67
Figura 34: Promedio General de Notas por Tipo de Examen.....	68
Figura 35: Promedio General de Notas de Asignaturas Aprobadas por Licenciatura	69
Figura 36: Promedio Ponderado de Notas por Período Académico	70
Figura 37: Promedio Ponderado de Notas de Asignaturas Aprobadas por Tipo de Examen.....	71
Figura 38: Eficiencia por Licenciatura	72
Figura 39: Efectividad por Licenciatura.....	73
Figura 40: Resultado de la encuesta a la pregunta “¿Se puede iniciar sesión y cerrar sesión de manera sencilla?”.....	76
Figura 41: Resultado de la encuesta a la pregunta “¿Se puede acceder a los reportes de manera sencilla?”.....	76
Figura 42: Resultado de la encuesta a la pregunta “¿Se utiliza un lenguaje natural y simple?”	77
Figura 43: Resultado de la encuesta a la pregunta “¿Se puede ver la información detallada fácilmente?”	77
Figura 44: Resultado de la encuesta a la pregunta “¿Los indicadores que se muestran en los reportes son de utilidad?”	78

ÍNDICE DE TABLAS

<i>Número</i>	<i>Página</i>
1. OLTP vs. OLAP	19
2. Planificación del Proyecto	34
3. Definición de Requisitos del Negocio	36

INTRODUCCIÓN

Actualmente las soluciones de inteligencia de negocio son muy importantes para las organizaciones donde se manejan grandes volúmenes de datos ya que permiten centralizarlos y analizarlos con el fin de obtener indicadores con los cuales sea posible supervisar el estado de los procesos de negocio. Las universidades no escapan de esta realidad. Para ello se tiene como objetivo de este trabajo desarrollar una solución de inteligencia de negocio que soporte los indicadores relacionados con el rendimiento académico y con la atención que presta el Servicio de Orientación de la Facultad de Ciencias de la Universidad Central de Venezuela con la finalidad de mejorar la prosecución estudiantil y apoyar a la toma de decisiones.

Este Trabajo Especial de Grado está dividido en tres capítulos; el primero de ellos corresponde al Problema de Investigación; en el cual se encuentra el planteamiento del problema, los objetivos, la justificación y el alcance de la solución propuesta así como también la adaptación más adecuada del método de desarrollo de un almacén de datos para este caso de estudio.

El segundo capítulo está relacionado con el Marco Conceptual, el cual contempla inicialmente los aspectos más importantes acerca del Servicio de Orientación y el rendimiento académico para luego explicar los elementos más relevantes de los sistemas de información y sus diferentes tipos, haciendo énfasis en los Sistemas de Procesamiento de Transacciones en Línea y los Sistemas de Procesamiento Analítico en Línea; posteriormente se toman en cuenta las características, las funciones, la arquitectura, ventajas, desventajas y las herramientas de desarrollo de una solución de inteligencia de negocio, particularmente Pentaho Community Edition. Adicionalmente, en este capítulo se exponen los conceptos básicos acerca de los indicadores.

El tercer capítulo, correspondiente al Marco Aplicativo, describe las actividades que se llevaron a cabo durante la implementación de la solución de inteligencia de negocio tales como la planificación del proyecto, la definición de los requerimientos del usuario, la selección de los productos de software a utilizar, la realización del diseño lógico y del diseño físico tanto del Área de Almacenamiento Intermedio como del Datamart, la elaboración, ejecución de los procesos de Extracción, Transformación y Carga, utilizados para poblar y mantener actualizada el Área de Almacenamiento Intermedio y el Datamart; así como también la especificación de la aplicación de usuario, la integración y despliegue y la gestión del proyecto con la finalidad de brindar apoyo a la toma de decisiones.

Por último, se exponen las conclusiones, recomendaciones, referencias bibliográficas y anexos.

Capítulo 1

PROBLEMA DE INVESTIGACIÓN

1.1 Planteamiento del Problema

Actualmente el Servicio de Orientación de la Facultad de Ciencias no cuenta con un sistema que le permita evaluar el rendimiento académico de los estudiantes con la finalidad de mejorar su prosecución estudiantil.

Es posible brindar apoyo al personal del Servicio de Orientación en el proceso de toma de decisiones, determinando cuáles son los indicadores relacionados con el desempeño del individuo en el ámbito académico y supervisando el estado de éstos a través del tiempo; ese es uno de los objetivos que se desea lograr al desarrollar la solución de inteligencia de negocio relacionada a este caso de estudio.

De igual forma, es necesario tomar en consideración los indicadores que provean información acerca de la atención que se presta a la población estudiantil en el área académica, asistencial, de prevención y vocacional.

A diferencia del Servicio de Orientación, la División de Control de Estudios (DCE) cuenta con un sistema automatizado de recolección de datos de la población estudiantil de la Facultad de Ciencias. A través del historial académico se le hace seguimiento a cada estudiante; sin embargo, actualmente no se obtiene el máximo provecho de éste, ya que se encuentra guardado en un repositorio de datos cuyo diseño lógico es transaccional, razón por la cual los indicadores relacionados con el rendimiento académico que mide la DCE, son limitados.

Se presentará información de utilidad para el Servicio de Orientación, mediante una solución de inteligencia de negocio que mostrará el estado de los indicadores asociados a los requerimientos definidos por los usuarios. Estos indicadores se encuentran especificados en la definición de requisitos que forma parte del Marco Aplicativo de este Trabajo Especial de Grado.

1.2 Objetivos

1.2.1 Objetivo General

Desarrollar una solución de inteligencia de negocio que soporte los indicadores relacionados con el rendimiento académico y con la atención que presta el Servicio de Orientación de la Facultad de Ciencias de la Universidad Central de Venezuela con la finalidad de mejorar la prosecución estudiantil y apoyar a la toma de decisiones.

1.2.2 Objetivos Específicos

Para que el objetivo general propuesto se pueda lograr, es preciso cumplir con los siguientes objetivos específicos:

- Conocer la situación actual del Servicio de Orientación y manejar la información acerca del rendimiento académico, que se encuentra disponible en la División de Control de Estudios (DCE).
- Definir los indicadores de gestión relacionados con el rendimiento estudiantil y con la labor que desempeña el Servicio de Orientación.
- Aplicar la adaptación más adecuada del Ciclo de Vida Dimensional del Negocio de (Kimball, 2002) para este caso de estudio.
- Crear y diseñar un Datamart valiéndose de la herramienta tecnológica seleccionada.
- Recopilar los datos proporcionados por la DCE y el Servicio de Orientación, transformarlos y guardarlos en el Datamart creado.
- Generar un portal web interactivo en el cual se muestren indicadores relacionados con el rendimiento académico y con la atención que presta el Servicio de Orientación de la Facultad de Ciencias.

1.3 Solución Propuesta

A continuación, en la Figura 1, se muestran los componentes que forman parte de la solución de inteligencia de negocio correspondiente a este caso de estudio.

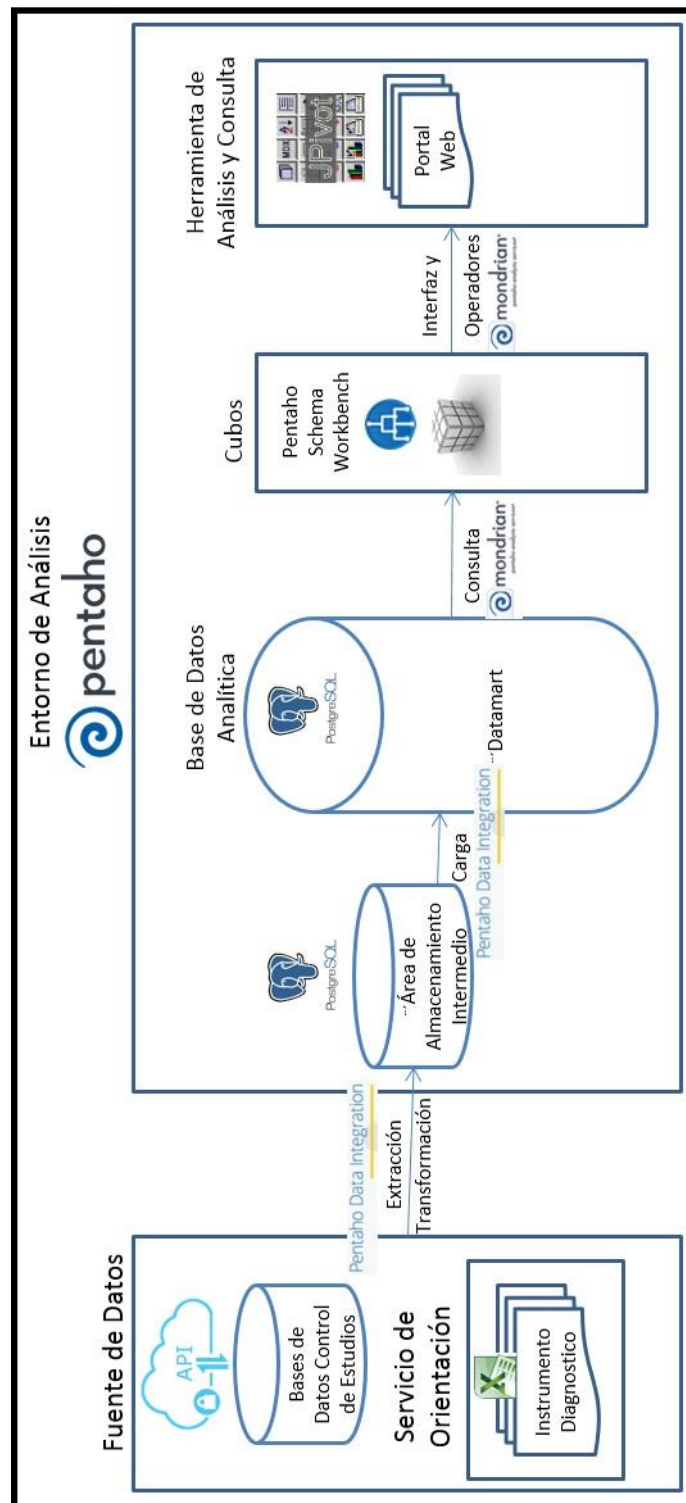


Figura 1: Arquitectura Propuesta

Los componentes de la arquitectura de la solución de inteligencia de negocio propuesta son los siguientes:

- *Fuentes de Datos*: Se cuenta con los datos relacionados con el rendimiento académico de la población estudiantil almacenados en la base de datos de la DCE, así como también, se tiene acceso a los registros de los estudiantes atendidos por el Servicio de Orientación, a través del Libro de Cálculo llamado Instrumento Diagnóstico.
- *Proceso de Extracción, Transformación y Carga (ETC)*: Con este componente se lleva a cabo la extracción de los datos que provienen de las distintas fuentes, independientemente del formato en que se encuentren, para luego integrarlos en un único repositorio que sirve como Área de Almacenamiento Intermedio y posteriormente, transformarlos de manera tal que puedan ser cargados en el Datamart al cual se harán consultas, con el fin de elaborar los reportes que le permitan al usuario analizar los resultados obtenidos y tomar decisiones en base a ellos.
- *Área de Almacenamiento Intermedio*: Esta es una base de datos relacional que sirve como repositorio de datos único, donde se guardan los registros provenientes de la DCE y del Servicio de Orientación una vez que se han realizado las tareas de limpieza, detección y eliminación de errores y registros duplicados, relleno de campos vacíos, resolución de inconsistencias, entre otras.
- *Datamart*: En esta base de datos multidimensional se guardan los registros luego de estandarizar códigos y formatos de representación, definir dominios, usar conversiones y combinaciones para crear nuevos campos, entre otras actividades. Este repositorio de datos está diseñado de tal manera que se puedan realizar consultas complejas en las cuales se manipula un gran volumen de datos.
- *Cubos*: Son estructuras dimensionales basadas en plataformas de Procesamiento Analítico en Línea (OLAP) (Kimball, 2002). Están compuestos por las dimensiones y tablas de hechos del Datamart para obtener los resultados de los indicadores asociados con el rendimiento académico de la población estudiantil y con la labor que desempeña el Servicio de Orientación. Entre los indicadores obtenidos se tienen:
 - *Cantidad de Estudiantes Atendidos*: Se refiere a los estudiantes atendidos por el Servicio de Orientación.
 - *Cantidad de Asignaturas y Unidades de Crédito (UC) Inscritas, Cursadas, Retiradas, Aprobadas, Aprobadas por Equivalencia, Aprobadas en General, Aplazadas*: Indicador

basado en el estatus de las asignaturas y las UC existentes en el historial académico del estudiante.

- *Promedio General de Notas*: Epicentro obtenido del comportamiento de las notas del estudiante, excluyendo las asignaturas cuya calificación no sea numérica.
- *Promedio Ponderado de Notas*: Valor medio, producto de la combinación de las notas y las UC, excluyendo las asignaturas cuya calificación no sea numérica.
- *Promedio General de Notas de Asignaturas Aprobadas*: Epicentro obtenido del comportamiento de las notas del estudiante, excluyendo las asignaturas aplazadas y las asignaturas aprobadas cuya calificación que no sea numérica.
- *Promedio Ponderado de Notas de Asignaturas Aprobadas*: Valor medio, producto de la combinación de las notas y las UC, excluyendo las asignaturas aplazadas y las asignaturas aprobadas con una calificación que no sea numérica.
- *Eficiencia*: Indicador obtenido a través del cociente de las UC aprobadas en general con respecto a las UC cursadas.
- *Ratio de Reprobación*: Indica cuántas asignaturas han sido reprobadas en relación a la cantidad de asignaturas que han sido aprobadas en general; en el mejor caso este indicador debe mostrar cero como resultado.
- *Efectividad*: Indicador obtenido a través del cociente, de las UC aprobadas en general con respecto a las UC inscritas, permitiendo así ver la inclinación del alumno con respecto a sus deberes académicos.
- *Herramienta de Análisis y Consulta*: Para este caso de estudio se utilizó el portal web que proporciona Pentaho Business Analytics Platform y JPivot, como herramientas de consulta, ya que permiten mostrar de manera resumida el estado de los indicadores relacionados con el rendimiento académico y con la labor que desempeña el Servicio de Orientación.

1.4 Método de Desarrollo de un Almacén de Datos Seleccionado

Se seleccionó el Método de Desarrollo de un Almacén de Datos de (Kimball, 2002), estudiado anteriormente en el seminario; éste se encuentra, a continuación, en la Figura 2.

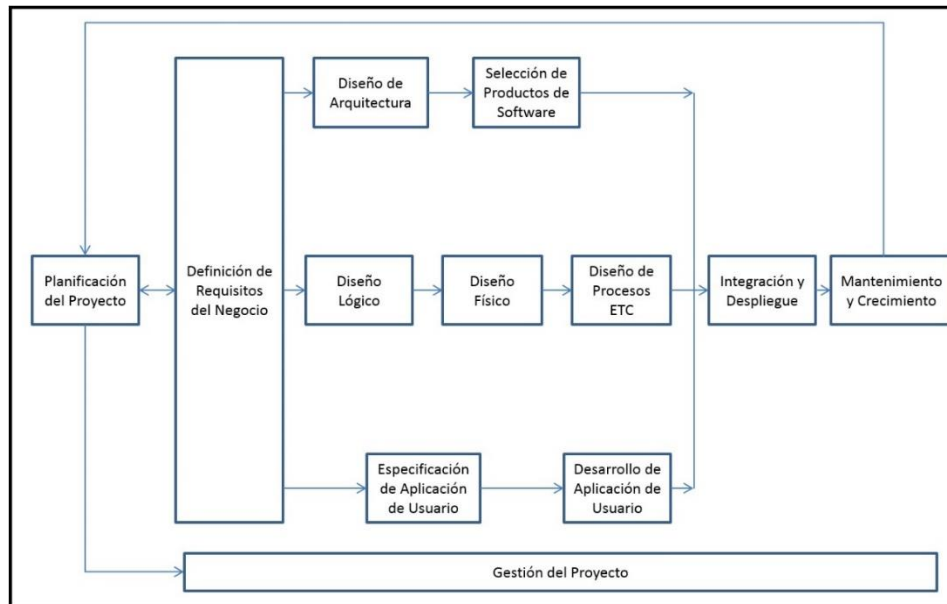


Figura 2: Ciclo de Vida Dimensional del Negocio
(Kimball, 2002)

En la Figura 3 se muestra la adaptación más adecuada del Ciclo de Vida Dimensional del Negocio para este caso de estudio en particular.

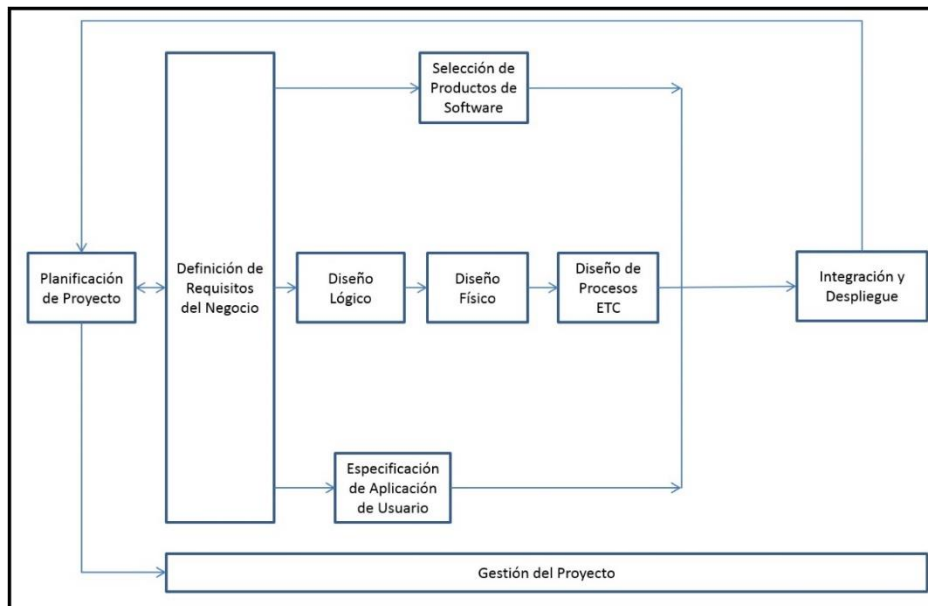


Figura 3: Adaptación del Ciclo de Vida Dimensional del Negocio de Kimball

Las etapas del Ciclo de Vida Dimensional del Negocio de (Kimball, 2002), que se aplican en este caso de estudio se explican en detalle a continuación:

- *Planificación del Proyecto*: Se toman en cuenta los recursos necesarios para realizar el proyecto así como también los procesos de negocio de la organización y los requerimientos que debe satisfacer la solución de inteligencia de negocio para determinar su alcance y así establecer una fecha tentativa para cada entregable del proyecto. Una vez que la planificación ha sido elaborada, ésta puede cambiar a lo largo del tiempo, si es preciso.
- *Definición de Requisitos del Negocio*: Se establecen los indicadores que permitirán supervisar el estado de los procesos de negocio y las reglas que rigen la interacción entre los departamentos de la organización y entre la organización y su entorno.
- *Diseño Lógico*: Se diseña el Modelo Dimensional en base a las consultas y los reportes que satisfagan las necesidades de la organización.
- *Diseño Físico*: El diseño físico de las bases de datos se enfoca en la selección de las estructuras necesarias para soportar el diseño lógico del Área de Almacenamiento Intermedio y del Datamart.
- *Diseño de Procesos de Extracción, Transformación y Carga (ETC)*: Se elaboran y se ejecutan los procesos de Extracción, Transformación y Carga con el fin de obtener los datos provenientes de las diferentes fuentes, guardarlos en un único repositorio y transformarlos en la información que será almacenada en el Datamart.

- *Selección de Productos de Software:* Se seleccionan los componentes relacionados con la solución de inteligencia de negocio propuesta como lo son el motor de base de datos, la herramienta de ETC, la herramienta de elaboración de reportes, entre otros.
- *Especificación de Aplicación de Usuario:* Se identifican los diferentes roles que se le pueden asignar a los usuarios, para determinar las vistas de la aplicación, en base al alcance de los perfiles existentes.
- *Integración y Despliegue:* Se integran todos los componentes de la solución de inteligencia de negocio propuesta, de tal manera que los usuarios puedan supervisar el funcionamiento de los procesos de negocio.
- *Gestión del Proyecto:* La gerencia del proyecto asegura que las actividades se lleven a cabo en forma correcta y sincronizada, durante todo su ciclo de vida. Para lograrlo, se supervisa el estado del proyecto verificando si se cumple con las tareas según lo estipulado en la planificación, ya que si no es así, se deben realizar los ajustes pertinentes, tomando en cuenta los requerimientos de los usuarios y los recursos disponibles para satisfacerlos.

1.5 Justificación de la Solución Propuesta

El presente TEG se enfoca en el desarrollo de una solución de inteligencia de negocio que se adapte a los requerimientos del personal del Servicio de Orientación relacionados con la cantidad de estudiantes que atiende, así como también los indicadores asociados al rendimiento académico de la población estudiantil de la Facultad de Ciencias, basándose en el historial académico del estudiante; lo cual permite a los usuarios contar con una herramienta que les brinde la oportunidad de obtener información acerca del desempeño de la población de interés en el ámbito académico y la evolución de éste a través del tiempo, desde diversas perspectivas.

Se utiliza un Datamart cuyo diseño lógico es dimensional, permitiendo así, almacenar una gran cantidad de datos relacionados con el rendimiento académico de la población estudiantil y la labor que desempeña la dependencia antes mencionada, de tal manera que el usuario no deba invertir recursos en el procesamiento de los datos y en lugar de esto pueda utilizar esos recursos en la supervisión del estado de los indicadores y en la toma de decisiones en base a ellos.

El proceso de extracción, transformación y carga de los datos desde las distintas fuentes de datos hacia el Área de Almacenamiento Intermedio y desde el Área de Almacenamiento Intermedio hacia el Datamart se realiza periódicamente de tal manera que esto no afecte el funcionamiento de

las fuentes de datos ni tampoco de la solución de inteligencia de negocio asociada a este caso de estudio, la cual cuenta con un funcionamiento sin ningún tipo de interrupción y permite a los usuarios acceder rápidamente a la información necesaria.

El cálculo de los indicadores se ejecuta de modo que los reportes puedan ser editados por los usuarios, en caso de que sea necesario, sin que esto comprometa la veracidad de los resultados obtenidos.

1.6 Alcance de la Solución Propuesta

En este trabajo se tomarán en cuenta los indicadores de rendimiento estudiantil y los que se encuentren relacionados con la labor que desempeña el Servicio de Orientación de la Facultad de Ciencias; con la finalidad de brindar apoyo a la toma de decisiones del personal de esta dependencia.

Capítulo 2

MARCO CONCEPTUAL

2.1 Servicio de Orientación

El Servicio de Orientación de la Facultad de Ciencias es una dependencia adscrita a la Coordinación Académica y tiene como objetivo fundamental asesorar al estudiante a fin de facilitarle los recursos necesarios que le permitan desenvolverse exitosamente en el medio universitario (Servicio de Orientación, 2015).

Misión

El Servicio de Orientación tiene como misión el ser una entidad prestadora de servicios a la comunidad de la Facultad de Ciencias y, por extensión, a la Universidad Central de Venezuela, con el fin de brindar ayuda psicológica y social, individual y colectiva a las personas, así como prestar asesoría técnica y profesional en áreas de su competencia para facilitar los procesos en los cuales, debido a sus características, se requiera su asistencia.

Esta dependencia se encarga de:

- Atender en consulta psicológica individual y/o grupal a los estudiantes que así lo requieran.
- Evaluar psicológicamente a consultantes y brindar recomendaciones en procesos como: cambios de facultad, normas de permanencia, reincorporaciones, entre otras.
- Asesorar a las autoridades en la toma de decisiones que involucren variables pertinentes a su labor profesional.
- Elaborar instrumentos psicométricos como escalas, cuestionarios y encuestas requeridos por la facultad, y posteriormente, la realización de los estudios estadísticos correspondientes a la aplicación de estos instrumentos.
- Participar en procesos institucionales como co-facilitador en la solución de conflictos, la evaluación de situaciones, planes de prevención en áreas de su competencia, procesos de inducción y adaptación de los estudiantes, entre otros.

Es necesario para el Servicio de Orientación conocer el rendimiento académico de la población estudiantil de la Facultad de Ciencias, a fin de brindar las herramientas que se requieren para mejorar su prosecución; por esta razón a continuación se exponen los aspectos más importantes acerca del rendimiento académico.

2.2 Rendimiento Académico

El rendimiento académico de los estudiantes es un fenómeno enmarcado en el proceso educativo, y por lo tanto reviste singular importancia como objeto de estudio. De acuerdo a esto, el rendimiento académico es definido como la relación entre lo obtenido, expresado en una apreciación objetiva y cuantitativa, es decir, puntaje o calificación o en una subjetiva y cualitativa, es decir, mediante una escala de valores y rasgos sobresalientes y el esfuerzo empleado para obtenerlo, y con ello establecer el nivel de alcance, así como también los conocimientos, habilidades y/o destrezas adquiridas, el éxito o la ausencia de éste en la escolaridad, durante un período de tiempo determinado (Rodríguez & López, 2016).

El rendimiento académico es concebido como el logro del aprendizaje obtenido por el estudiante a través de las diversas actividades planificadas por el docente, en relación a los objetivos previamente establecidos; a su vez, es definido como la calidad de la actuación del estudiante con respecto a un conjunto de conocimientos, habilidades o destrezas en una asignatura determinada como resultado de un proceso instruccional; o como el grado en que cada estudiante ha alcanzado los objetivos propuestos y las condiciones bajo las cuales se produjo ese logro.

Tomando en cuenta la población estudiantil de la Facultad de Ciencias, el rendimiento académico es un proceso de negocio en el cual se encuentran involucrados los estudiantes, docentes, la DCE y demás entidades que allí se encuentran, ya que un proceso de negocio es cualquier actividad o conjunto de actividades lógicamente relacionadas que reciben un insumo, le agregan valor a éste, utilizando los recursos de la organización, y suministran un producto a un cliente interno o externo en función de los objetivos de la organización. (Laudon, 2012).

La DCE es una unidad adscrita a la Coordinación Académica de la Facultad de Ciencias encargada de realizar los trámites académico-administrativos relacionados con el ingreso, permanencia y egreso de los estudiantes. Así como también de brindar información a los docentes y al público en general en relación a la planificación de las actividades docentes de pregrado y contribuir al mejor desarrollo de éstas. Este proceso de negocio se vale del historial académico de cada estudiante que forma parte de la población estudiantil de la Facultad de Ciencias; en él se encuentran todas las asignaturas inscritas por el estudiante y la calificación obtenida. Es utilizado por la solución de inteligencia de negocio implementada para calcular los indicadores relacionados con el rendimiento académico.

A continuación se encuentran los aspectos más relevantes que deben tomarse en cuenta al momento de elegir el sistema de información que satisfaga las necesidades de la organización y le permita evaluar el estado de los procesos de negocio que ésta realiza con la finalidad de cumplir con sus objetivos.

2.3 Sistemas de Información

2.3.1 Definición de Sistemas de Información

Un sistema de información es un conjunto de elementos relacionados entre sí que recolectan, procesan, almacenan y transmiten información para apoyar el seguimiento y control de los procesos de negocio de una organización (Laudon, 2012).

2.3.2 Tipos de Sistemas de Información

Cada sector de la organización se encarga de una serie de objetivos específicos de manera tal que, al trabajar en conjunto, el rendimiento de la organización sea óptimo. En la Figura 4 se muestran los tipos de sistemas de información.

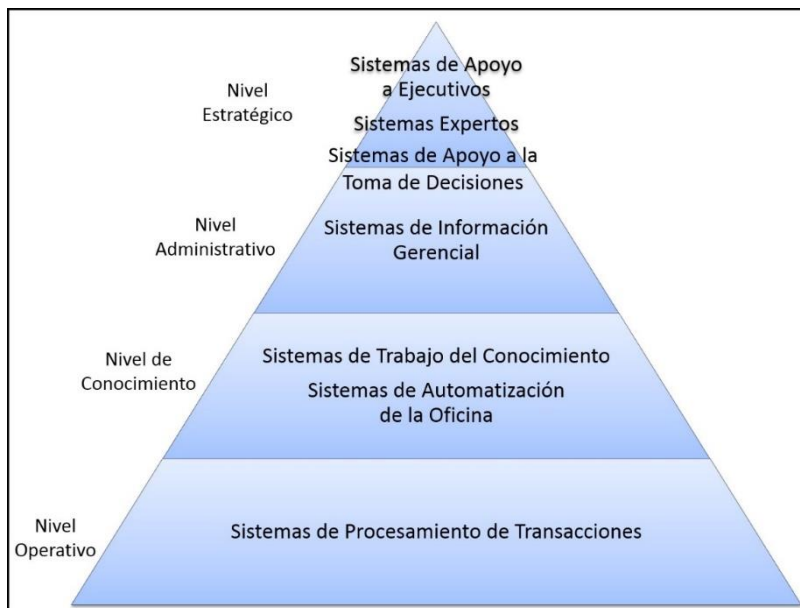


Figura 4: Tipos de Sistemas de Información
(Laudon, 2012)

A continuación se definen los diferentes tipos de sistemas de información, mediante varias clasificaciones:

- **Clasificación según los Niveles de la Organización:** Es la que permite delimitar las responsabilidades de cada departamento de la organización de la siguiente manera:
 - *Los Sistemas en el Nivel Operativo:* Supervisan las actividades y transacciones elementales de la organización. Su objetivo principal es contestar preguntas de rutina y rastrear el flujo de las transacciones pertinentes. En este nivel se utilizan los Sistemas de Procesamiento de Transacciones.
 - *Los Sistemas en el Nivel del Conocimiento:* Ayudan a controlar el flujo de documentos de la organización, utilizando los Sistemas de Automatización de la Oficina y los Sistemas de Trabajo del Conocimiento.
 - *Los Sistemas en el Nivel Administrativo:* Apoyan las actividades de seguimiento, control, toma de decisiones y administración de los empleados de la Gerencia Media de la organización, a través de informes periódicos, en lugar de información instantánea acerca de las operaciones realizadas. En este nivel se utilizan los Sistemas de Información Gerencial y los Sistemas de Apoyo a la Toma de Decisiones.

- *Los Sistemas en el Nivel Estratégico:* Apoyan las actividades de planificación a largo plazo que realizan los empleados de la Alta Gerencia de la organización, haciendo uso de los Sistemas de Apoyo a la Toma de Decisiones, los Sistemas Expertos y los Sistemas de Apoyo a Ejecutivos.
- **Clasificación según la Función que cumple en la Organización:** Es la que permite determinar qué tipos de sistemas de información satisfacen las necesidades específicas de cada organización. Los tipos de sistemas de información según la función que cumplen son los siguientes:
 - *Los Sistemas de Procesamiento de Transacciones:* Manejan grandes cantidades de datos ya que procesan y registran las transacciones diarias y rutinarias de la organización, reduciendo el tiempo en el que éstas se llevan a cabo. Una vez culminada esta tarea, se realizan reportes o informes en base a los resultados obtenidos. Este tipo de sistemas son los que le proporcionan a los administradores información actualizada acerca del funcionamiento de la organización pues, para realizar la recopilación de datos, se vale de la interacción con los usuarios y con los entes externos, así como también cuentan con un funcionamiento sin ninguna clase de interrupción. Es también conocido como Sistema de Procesamiento de Transacciones en Línea u OLTP por sus siglas en inglés; estos proveen la descripción de todas las actividades y sistemas asociados con la introducción de datos en la base de datos que sirve posteriormente como insumo para la realización de los procesos de negocio. Generalmente se utiliza en función a una base de datos relacional, aunque este enfoque también puede ser utilizado para describir cualquier entorno en el cual se lleven a cabo procesos transaccionales. La principal función de los sistemas OLTP es realizar una gran cantidad de consultas rápidamente y garantizar la integridad de los datos en un entorno de accesos múltiples, donde la eficiencia es medida a través de la cantidad de transacciones realizadas por segundo. En las bases de datos diseñadas bajo el enfoque OLTP se guardan datos detallados y actualizados.
 - *Los Sistemas de Automatización de la Oficina:* Apoyan a los trabajadores que manejan los datos, quienes por lo general, no generan conocimientos sino más bien procesan y analizan los datos antes de distribuirlos al resto de la organización o a entes externos. Entre los componentes más comunes de estos sistemas se encuentran los procesadores de texto, las hojas de cálculo, la calendarización electrónica y las comunicaciones mediante voz, correos electrónicos y videoconferencias.

- *Los Sistemas de Trabajo del Conocimiento:* Sirven de apoyo a los trabajadores profesionales como ingenieros, médicos, científicos y contadores en sus esfuerzos por crear nuevo conocimiento y compartirlo con la organización de la cual forman parte y con la sociedad.
- *Los Sistemas de Información Gerencial:* Contribuyen a la correcta interacción entre los usuarios, el hardware y el software; de manera tal que se brinde apoyo a un amplio espectro de tareas en la organización, entre ellas el análisis y la toma de decisiones, a pesar de que no existe como una estructura individual en ninguna parte de la organización. Para acceder a la información, los usuarios de este tipo de sistemas comparten una base de datos común que los ayuda a interpretar y a manipular los datos que allí se encuentran.
- *Los Sistemas de Apoyo a la Toma de Decisiones:* Son aquellos que, valiéndose de una base de datos, se caracterizan por brindar apoyo a la toma de decisiones en cada una de sus fases y por adaptarse a las necesidades específicas del usuario, es por ello que este tipo de sistemas se utilizan para implementar una solución de inteligencia de negocio.
- *Los Sistemas Expertos:* Capturan y utilizan el conocimiento de un experto para solucionar un problema específico de la organización. A diferencia de los Sistemas de Apoyo a la Toma de Decisiones que ceden al responsable la toma de la decisión definitiva, un Sistema Experto selecciona la mejor solución para un problema o una clase específica de problemas. Los componentes básicos de un Sistema Experto son la base de conocimientos, un motor de inferencia que conecta al usuario con el sistema mediante el procesamiento de consultas y la interfaz de usuario. Profesionales conocidos como Ingenieros de Conocimiento capturan la pericia de los expertos, construyen un sistema de cómputo que contiene este conocimiento experto y lo utilizan.
- *Los Sistemas de Apoyo a Ejecutivos:* Ayudan a los usuarios a organizar sus actividades relacionadas con el entorno externo mediante herramientas gráficas y de comunicaciones; así como también ayudan a resolver problemas de toma de decisiones no estructuradas mediante la creación de un ambiente en el cual tengan toda la información necesaria.

A continuación se explican con mayor nivel de detalle los Sistemas de Procesamiento Analítico en Línea u Online Analytic Processing por sus siglas en inglés OLAP utilizados en el nivel administrativo y estratégico de la organización, según la clasificación anterior.

2.3.3 Sistema OLAP

OLAP es un conjunto de procesos definidos para proveer un marco de trabajo dimensional que brinde apoyo a la toma de decisiones. El término OLAP también es utilizado para definir a los proveedores que ofrecen sistemas de bases de datos no relacionales, multidimensionales que se presten para el análisis en función de los datos almacenados en ellos (Kimball, 2002).

El objetivo principal de los sistemas OLAP es realizar una cantidad de consultas relativamente baja en comparación con los sistemas OLTP. Las consultas son complejas, es por ello que el tiempo de respuesta es una medida de eficiencia de este sistema. Las bases de datos diseñadas bajo este enfoque guardan datos históricos en un esquema multidimensional.

Debido a que este tipo de sistema se caracteriza por brindar apoyo a la toma de decisiones necesita un componente que se valga del esquema multidimensional de la base de datos para calcular indicadores; esta es una de las razones por las cuales se cuenta con cubos OLAP.

El cubo OLAP es una estructura de almacenamiento que permite realizar diferentes combinaciones de datos, para visualizar los indicadores de rendimiento de los procesos de negocio de la organización hasta un determinado nivel de detalle; a su vez permite navegar por sus dimensiones y analizar la información desde diferentes puntos de vista.

2.3.4 OLTP vs. OLAP

La Tabla 1 que se muestra a continuación contiene las principales diferencias entre los sistemas OLTP y los sistemas OLAP.

Tabla 1: OLTP vs. OLAP
(Kimball, 2002)

	Sistema OLTP	Sistema OLAP
Fuente de Datos	Datos operacionales	Varias bases de datos OLTP
Propósito de los Datos	Controlar las transacciones	Apoyar la toma de decisiones
Contenido de los Datos	Estado de los procesos de negocio	Vistas multidimensionales
Inserciones	Sencillas y rápidas.	Complejas y periódicas
Actualizaciones	Iniciadas por los usuarios finales	Complejas y periódicas
Consultas	Simple y estandarizadas	Usualmente complejas
Velocidad de Procesamiento	Generalmente muy rápida	Usualmente lenta.
Espacio Requerido	Relativamente pequeño	Grande debido a la datos históricos
Diseño de Base de Datos	Altamente normalizada.	Modelo Dimensional
Respaldo y Recuperación	Respaldos periódicos	Se utilizan los sistemas OLTP

Tanto los sistemas OLTP como los sistemas OLAP brindan beneficios a sus usuarios de diversas maneras; sin embargo, cuando éstos forman parte de la arquitectura de una solución de inteligencia de negocio, se pueden aprovechar en mayor medida las bondades de ambos sistemas. Es por ello que a continuación se explican los aspectos más importantes de una solución de inteligencia de negocio.

2.4 Inteligencia de Negocio

Debido a la situación actual de las organizaciones, es decir, el entorno competitivo y globalizado así como también los grandes volúmenes de información disponibles, se necesitan aplicaciones y procesos de recolección, almacenamiento y presentación de los datos para estudiar la naturaleza del negocio mediante indicadores, medidas y dimensiones, es decir, mediante perspectivas de análisis que ayuden a los usuarios a tomar mejores decisiones; este tipo de tecnología es conocida como una solución de inteligencia de negocio (Laudon, 2012). Su función es recuperar los datos de los sistemas operacionales, transformarlos y cargarlos en una estructura de datos dimensional, especialmente diseñada para contener los datos relacionados con el rendimiento de la organización y que al ser analizados, les brindan a los usuarios la información que requieren. Algunos de los componentes que pueden formar parte de una solución de inteligencia de negocio son los siguientes: bases de datos, aplicaciones analíticas (OLAP), herramientas que permiten realizar consultas y elaborar reportes, herramientas de Minería de Datos que se utilizan para establecer relaciones y descubrir patrones existentes en grandes bases de datos, técnicas de visualización de datos, herramientas de Extracción, Transformación y Carga, entre otros.

2.4.1 Características de una solución de inteligencia de negocio

Toda solución de inteligencia de negocio, independientemente de los procesos de negocio que posee la organización, debe ser diseñada de manera tal que cuente con las siguientes características: (Laudon, 2012)

- *Amplia*: Abarca todos los procesos de negocio de la organización.
- *Integrada*: Cuenta con un repositorio de datos y vistas multidimensionales de ellos.
- *Precisa*: Permite tomar decisiones basadas en información precisa.
- *Rápida*: Brinda fácil acceso a los datos.

2.4.2 Funciones de una solución de inteligencia de negocio

En una organización se implementa una solución de inteligencia de negocio con la finalidad de mostrar: (Laudon, 2012)

- Notificaciones de alerta.
- Reportes con un formato determinado para todos los usuarios.

- Indicadores de rendimiento de los procesos de negocio de la organización.

2.4.3 Arquitectura de una solución de inteligencia de negocio

A continuación, en la Figura 5, se muestran los componentes que forman parte de una solución de inteligencia de negocio.

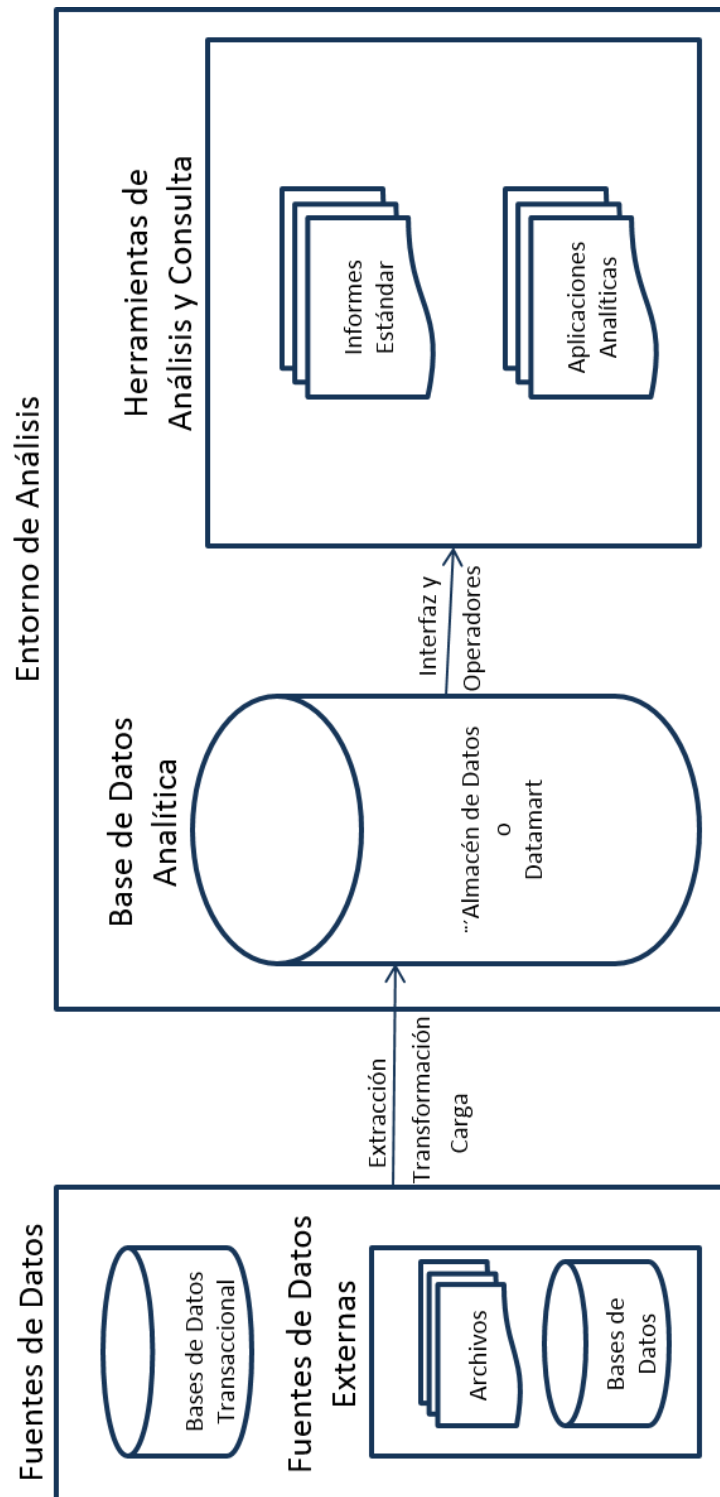


Figura 5: Arquitectura de una solución de inteligencia de negocio (Kimball, 2002)

Los componentes de la arquitectura de una solución de inteligencia de negocio, según (Kimball, 2002), son los siguientes:

- *Fuentes de Datos:* Son los documentos, archivos y bases de datos transaccionales en los cuales quedan registrados todos los datos relacionados con las operaciones de la organización y sus procesos de negocio.
- *Proceso de Extracción, Transformación y Carga (ETC):*
 - Para guardar la información perteneciente al almacén de datos se deben extraer los datos provenientes de las fuentes de datos para luego limpiarlos, transformarlos y añadirles significado y contexto; finalmente es necesario actualizar las dimensiones y las tablas de hechos. A este proceso se le conoce como Extracción, Transformación y Carga (ETC).
 - Es necesario identificar los atributos de interés y escoger la fuente de datos más idónea para cada uno de ellos, así como también es preciso unificar el significado de los datos, validar su veracidad para posteriormente definir las tareas de preprocesado y transformación con el fin de realizar la carga inicial y la sincronización.
 - Las tareas principales de preprocesado consisten en la limpieza, detección y eliminación de errores y registros duplicados, relleno de campos vacíos, resolución de inconsistencias, integración de los datos, identificación de la mejor fuente de datos para cada atributo; así como también en este conjunto de actividades se contempla la transformación de los datos para que se adapten al contexto del almacén de datos, esto implica estandarizar códigos y formatos de representación, definir dominios, usar conversiones y combinaciones para crear nuevos campos, corregir errores y generar documentación acerca de los datos que se manejan para posteriormente definir los procesos de carga y sincronización.
 - Para realizar el proceso de Extracción, Transformación y Carga correctamente se utiliza un Área de Almacenamiento Intermedio que sirve como línea divisoria entre el Área Operacional y el Almacén de Datos. En ella se realiza la estandarización y la integración de los datos, así como también se generan nuevos datos, en caso de ser necesario.
- *Base de Datos Analítica:* En esta base de datos se guardan los registros luego de estandarizar códigos y formatos de representación, definir dominios, usar conversiones y combinaciones para crear nuevos campos, corregir errores. Este repositorio de datos está diseñado de tal manera que el usuario pueda ver el estado de los indicadores desde diversas perspectivas. Generalmente, se

utiliza un Almacén de Datos o un Datamart como repositorio de datos, por lo tanto, a continuación se explican en detalle los aspectos más importantes relacionados con estas estructuras de datos.

- *Definición de Almacén de Datos:* Un almacén de datos es una colección de las transacciones realizadas, éstas son guardadas de manera tal que los datos puedan ser consultados y analizados. A su vez, un almacén de datos se define como la unión de todos los Datamarts de la organización (Kimball, 2002). Un almacén de datos es una colección de datos diseñada para dar apoyo a los procesos de toma de decisiones (Inmon, 2002).
- *Definición de Datamart:* Un Datamart es un conjunto de datos que se almacenan bajo el mismo esquema de un almacén de datos; sin embargo, su alcance es menor ya que está diseñado con una función específica que satisface las necesidades de un sector reducido de la organización (Kendall, 2005). Se refiere a una vista del Almacén de Datos orientada a solo un aspecto de la organización. Contiene mucha menos cantidad de datos que un Almacén de Datos. Un almacén de datos está relacionado con varios Datamarts que requieren hardware, software, acceso a la red y tiempo de carga adicional; cada uno de ellos utiliza el almacén de datos como fuente de datos y está dirigido a un proceso de negocio en particular de un departamento de la organización (Kimball, 2002). Un ejemplo de esto se muestra a continuación en la Figura 6. Los Datamarts son necesarios cuando las herramientas de acceso, utilizadas por los usuarios, tienen sus propias estructuras internas. Se utilizan para segmentar la información en varias plataformas, lo que brinda posible portabilidad; también sirven para controlar más fácilmente el acceso a los datos y para mejorar el tiempo de respuesta de las consultas, ya que el volumen de datos que se debe procesar es menor.

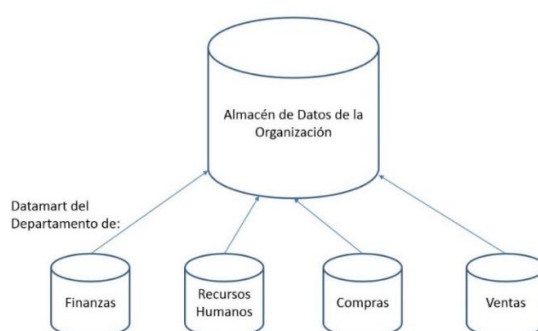


Figura 6: Ejemplo de Almacén de Datos y Datamart (Kimball, 2002)

- *Características de un Almacén de Datos:* El almacén de datos es un repositorio de datos que se distingue de otros tipos de bases de datos por las siguientes razones: (Inmon, 2002)
 - Está orientado a la información relevante de la organización

- Integrado.
 - Variable en el tiempo.
 - No volátil.
- *Modelo Dimensional*: Es un diseño lógico que permite definir la estructura de un almacén de datos. Está formado por hechos y dimensiones. Es una técnica de desarrollo para organizar los datos alrededor de los conceptos naturales de la institución y para proveer las bases para un análisis de información sofisticado. Es un modelo que se adapta a los procesos de negocio que se realicen (Kimball, 2002). Los conceptos más importantes a tomar en cuenta para diseñar un modelo dimensional son los siguientes:

- *Tabla de Hechos*: Es la tabla principal con las mediciones de rendimiento numéricas y no numéricas que son caracterizadas por una clave compuesta, donde cada elemento de ésta es una clave foránea que corresponde a una tabla de dimensión.

- En ella se encuentran los indicadores de rendimiento de los procesos de negocio de la organización.
- Cuando se agrega una nueva medida a la tabla de hechos, el campo debe figurar como no disponible, en las fechas anteriores a la adición.
- Cuando una dimensión cambia su granularidad, se debe eliminar la tabla de hechos relacionada para reconstruirla.
- La tabla de hechos puede ser transaccional, en la cual se agrega una fila cada vez que ocurre un evento o puede ser periódica, si se agrega una fila en base a un lapso de tiempo preestablecido; en ella se muestra el estado de los indicadores en un instante de tiempo.
- A continuación, se muestra en la Figura 7, un ejemplo de una Tabla de Hechos:



Figura 7: Ejemplo de Tabla de Hechos

- *Tabla de Dimensión*: Contiene información que es relativamente estática en el tiempo. Representa las diferentes formas de ver la información.

- Define las vistas que están directamente relacionadas con las medidas de las tablas de hechos. A continuación, se muestra en la Figura 8, un ejemplo de una dimensión.



Figura 8: Ejemplo de Dimensión

- Siempre existe al menos una dimensión temporal donde el nivel de especificidad o granularidad suele ser determinada por la fecha.
- Generalmente una dimensión posee 50 atributos, aproximadamente.
- Usualmente el modelo dimensional no tiene más de 15 dimensiones.
- Cuando se agrega un atributo a una dimensión, el campo debe figurar como no disponible, en las fechas anteriores a la adición.
- Cuando se agrega una dimensión al esquema dimensional, se debe incluir la clave de referencia a la tabla de hechos relacionada.
- Toda dimensión debe significar lo mismo para cada tabla de hechos con la que está relacionada.
- Una tabla de hechos puede hacer varias referencias a una misma dimensión, en este caso se dice que la dimensión juega varios roles.
- En caso de ser preciso, se crea una Dimensión Basura, en la cual se agrupan todos los atributos que no pueden organizarse coherentemente, para formar una dimensión por sí solos y se almacenan todos los posibles valores de ellos. Un ejemplo de esto es el estado de un producto en una tienda, es decir, disponible, no disponible, vendido, enviado o entregado.
- De ser necesario, se crea una Dimensión Degenerada que es un atributo en la tabla de hechos, el cual no tiene una dimensión asociada y no es una medida sino que suele ser un identificador único del registro. Un ejemplo de esto es el número de factura asociado a una venta.

- *Esquema Estrella*: Está dado por varias dimensiones y una única tabla de hechos con una relación 1:N con las tablas de dimensiones. A continuación, se muestra en la Figura 9, un ejemplo de un Esquema Estrella.

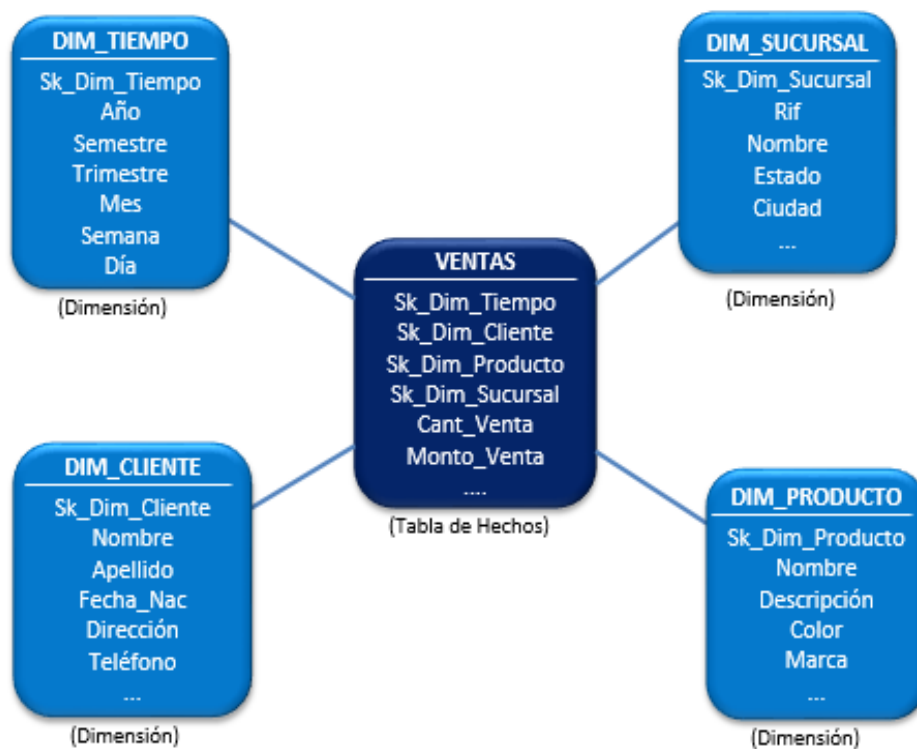


Figura 9: Ejemplo de un Esquema Estrella

▪ **Esquema Copo de Nieve:** Es un diagrama derivado del Esquema Estrella en el que las tablas de dimensión se normalizan en múltiples tablas. Por esta razón, la tabla de hechos deja de ser la única tabla del esquema que se relaciona con otras dimensiones. La diferencia entre ambos esquemas reside en la estructura de las tablas de dimensión. Para conseguir un esquema en copo de nieve se ha de tomar un esquema en estrella y conservar la tabla de hechos, centrándose únicamente en el modelado de las tablas de dimensión, que si bien en el esquema en estrella se encontraban totalmente desnormalizadas, ahora se dividen en subtablas tras un proceso de normalización. Es posible distinguir dos tipos de esquemas en copo de nieve, un snowflake completo, en el que todas las tablas de dimensión se encuentran normalizadas; mientras que en el snowflake, o snowflake parcial, sólo se lleva a cabo la normalización de algunas de ellas. En la Figura 10 se muestra un ejemplo de un Esquema Copo de Nieve.

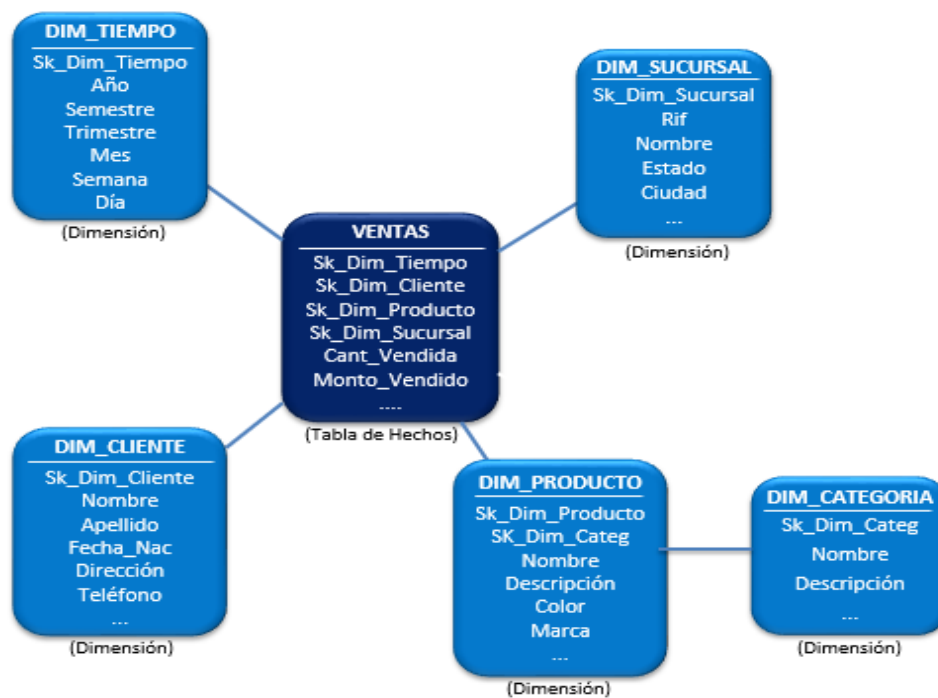


Figura 10: Ejemplo de un Esquema Copo de Nieve

- **Esquema Constelación de Estrellas:** Este modelo está compuesto por una serie de esquemas estrella, es decir, una tabla de hechos principal y una o más tablas de hechos auxiliares. Estas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones. No es necesario que las diferentes tablas de hechos compartan las mismas tablas de dimensiones, ya que, las tablas de hechos auxiliares pueden vincularse con solo algunas de las tablas de dimensiones asignadas a la tabla de hechos principal, y también pueden relacionarse con nuevas tablas de dimensiones. En la Figura 11 se muestra un ejemplo de un Esquema Constelación de Estrellas.

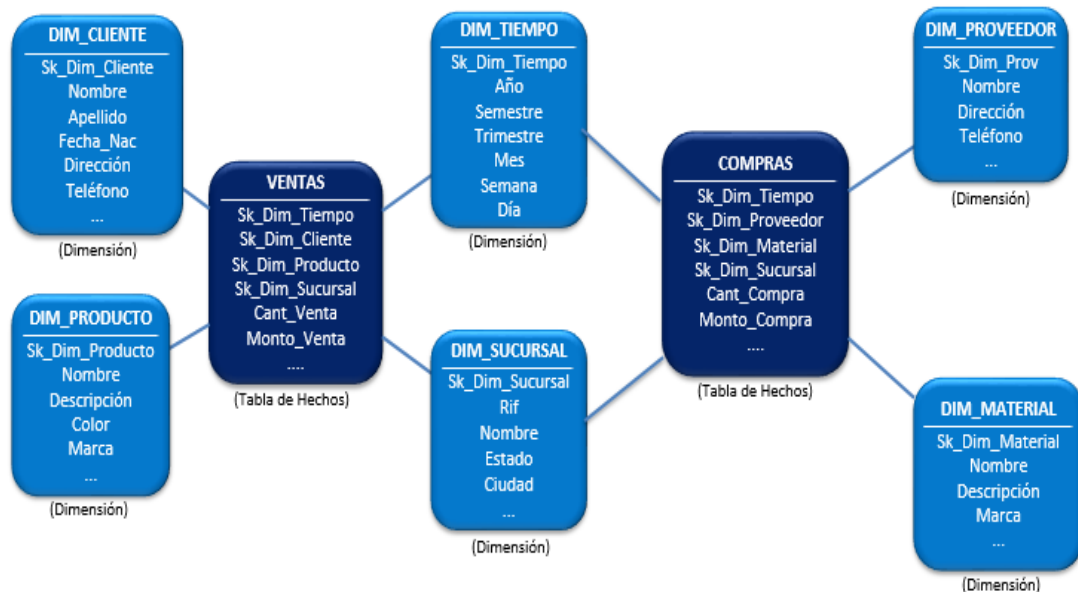


Figura 11: Ejemplo de un Esquema Constelación de Estrellas

- **Ventajas de un Almacén de Datos:** Las razones por las cuales es beneficioso crear un almacén de datos son las siguientes: (Kimball, 2002)
 - Aumenta la competitividad de la organización en el mercado.
 - Aumenta la productividad de la organización.
 - Brinda información de calidad.
 - Permite utilizar la información de mejor manera.
 - Mejora la comunicación de la organización al garantizar la disponibilidad de los datos.
- **Desventajas de un Almacén de Datos:** Usualmente no se toma en consideración, en su justa medida, el esfuerzo necesario para realizar las siguientes actividades: (Kimball, 2002)
 - Diseñar el almacén de datos.
 - Capturar, cargar y almacenar los datos en el almacén de datos.
 - Adaptarse a los continuos cambios de los requerimientos de los usuarios.

- *Herramientas de Análisis y Consulta:* Una de las principales funciones de una solución de inteligencia de negocio es facilitar el acceso a los datos. Esto se logra a través de herramientas de análisis y consulta que 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 utilizan complejos algoritmos e información del dominio. Estas aplicaciones se dividen en dos categorías, basadas en el nivel de sofisticación, llamadas informes estándar y aplicaciones analíticas.

- *Informes estándar:* Son la base del espectro de las herramientas de consulta que forman parte de una solución de inteligencia de negocio. Por lo general tienen un formato predefinido y parámetros de consulta fijos. En el caso más simple, son informes estáticos prealmacenados. Esta herramienta proporciona a los usuarios información acerca de lo que está sucediendo en un área determinada de la organización.
- *Aplicaciones analíticas:* Son más complejas que los informes estándar. Se centran en un proceso de negocio específico y resumen la experiencia acerca de cómo analizar e interpretar ese proceso de negocio. Estas aplicaciones incluyen algoritmos y modelos de minería de datos, que ayudan a identificar oportunidades o predecir el comportamiento de los procesos de negocio en base a la información que muestran. Otra característica de las aplicaciones analíticas es que el usuario puede solicitar cambios en los sistemas transaccionales, basándose en los conocimientos obtenidos como resultado del uso de la solución de inteligencia de negocio. En el otro extremo del espectro, se encuentran las aplicaciones analíticas que no pueden realizar este tipo de solicitudes y son independientes de las aplicaciones particulares de la organización.

2.4.4 Ventajas y Desventajas de una solución de inteligencia de negocio

- *Ventajas:* Los principales beneficios son: (Laudon, 2012)
 - Provee a los usuarios de la información necesaria para tomar decisiones.
 - Facilita el acceso a la información a los usuarios, según el rol que cumplen en la organización.
 - Simplifica tareas y procesos de negocio.
 - Permite crear nuevos modelos de negocio.
 - Crea nuevas relaciones entre la organización y su entorno.
- *Desventajas:* Los principales aspectos negativos son: (Laudon, 2012)

- Puede generar resistencia al cambio por parte de los usuarios que forman parte de la organización.
- El personal especializado en el área es relativamente poco.
- La base de datos de la organización se encuentra en una copia digital, vulnerable a la pérdida o al uso indebido de la información de la organización.
- Es posible que no cumpla con las expectativas de los usuarios debido a que no tenga suficiente capacidad de cómputo o que el tiempo de respuesta sea prolongado.

2.5 Herramientas para el desarrollo de una solución de inteligencia de negocio

En la siguiente sección se exponen los aspectos más importantes de Pentaho Community Edition como herramienta de software libre utilizada para implementar una solución de inteligencia de negocio.

2.5.1 Pentaho Community Edition

Es una plataforma de inteligencia de negocio orientada a la solución y centrada en procesos que incluye los principales componentes requeridos para adaptarse al funcionamiento de la organización.

Esta solución, al igual que su entorno, está basada en Java, haciéndola multiplataforma y permitiéndole satisfacer una gran variedad de necesidades de la organización ya que posee un Motor de Extracción, Transformación y Carga, un Motor OLAP, entre otras herramientas. (Pentaho, 2014).

- **Características Principales:** Pentaho se especializa en permitir a los usuarios
 - Analizar los datos.
 - Realizar informes en base a los resultados del análisis.
 - Supervisar los procesos de negocio de la organización.
 - Satisfacer las necesidades de la organización.
- **Requerimientos Mínimos de Hardware:** Se requiere que el hardware del equipo en el cual se desea utilizar la herramienta cuente con las siguientes características

- *Espacio en Disco*: 20 GB o más.
- *Memoria RAM*: 3 GB o más.
- *Procesador*: Pentium de doble núcleo o superior.
- **Requerimientos Mínimos de Software**: Es necesario que el equipo en el cual se desea utilizar la herramienta cuente con los siguientes componentes de software
 - *Sistema Operativo*:
 - Apple Macintosh OS 10.8 o superior.
 - Microsoft Windows 7 o superior.
 - Ubuntu o Ubuntu Server 12.X o superior.
 - Debian 6 o superior.
 - Java Development Kit versión 6 o 7.
 - *Navegador Web*:
 - Apple Safari 6 o superior.
 - Google Chrome 36 o superior.
 - Microsoft Internet Explorer 8 o superior.
 - Mozilla Firefox 31 o superior.
- **Ventajas**: Los beneficios más importantes de Pentaho Community Edition son los siguientes
 - Está desarrollado bajo licencias de software libre.
 - Se apoya en aplicaciones que se mantienen actualizadas.
 - Cuenta con una amplia comunidad de desarrolladores.
- **Desventajas**: Los principales aspectos negativos de Pentaho Community Edition son los siguientes
 - No posee suficientes fuentes de documentación fiables.
 - Existe una gran variedad de herramientas derivadas de Pentaho Community Edition lo cual dificulta su estandarización.

Una vez que ha sido escogida la herramienta más apropiada para desarrollar la solución de inteligencia de negocio que requiere la organización, es necesario definir los indicadores que permiten evaluar el estado de los procesos de negocio y, por ende, el nivel de cumplimiento de los objetivos previamente establecidos. Por esta razón, a continuación se exponen los aspectos más relevantes relacionados con los indicadores.

2.6 Indicadores

2.6.1 Definición de Indicador

Es un instrumento que permite medir cambios o resultados relacionados con los procesos de negocio y provee a la organización de un estándar contra el cual evaluar, estimar o demostrar el progreso que se ha logrado con respecto a metas establecidas (Beltrán, 2006).

2.6.2 Tipos de Indicador

Los indicadores se pueden clasificar de la siguiente manera: (Beltrán, 2006)

- **Indicadores de Procesos:** Están relacionados con el conjunto de actividades que forman parte del proceso. Ejemplo: Cumplimiento del Plan de Evaluación de una asignatura, durante un período.
- **Indicadores de Resultados:** Se refiere al comportamiento del proceso como un todo. Ejemplo: Cantidad de estudiantes que aprobaron una asignatura, en un período determinado.
- **Indicadores de Eficacia:** Representan el cociente entre la producción real y la esperada, independientemente de los recursos utilizados para lograrlo. Un ejemplo de esto sería un indicador que muestre a los empleados que han realizado al menos una venta en 10 días.
- **Indicadores de Eficiencia:** El concepto de eficiencia se refiere al grado de cumplimiento de los objetivos planteados, sin garantizarlo, es decir, en qué medida la organización, está cumpliendo con sus objetivos tomando en consideración los recursos con los que cuenta. Un ejemplo de esto sería un indicador que mida la cantidad de ventas que se realizan en 10 días.
- **Indicadores de Gestión:** Son medidas utilizadas para determinar el nivel de cumplimiento de los objetivos de una actividad perteneciente a un proyecto o del proyecto en sí. Por ejemplo: Administración y/o gestión del envío de productos.

2.6.3 Importancia de los Indicadores

La importancia de los indicadores radica en que son: (Mondragón, 2014)

- **Elemento de planificación:** Durante los procesos de planificación se utilizan con frecuencia los indicadores para establecer la meta u horizonte a donde se quiere llegar.
- **Estándar de seguimiento y control:** El indicador ayuda a entender o mostrar el estado del problema, ayuda a determinar la brecha entre lo planificado o esperado y el punto actual en el

que se hace la valoración o medición. Un indicador es una señal de alerta que induce a reconocer que es necesario resolver un problema.

- **Herramienta para la toma de decisiones:** Permite establecer métricas a través de las cuales se demuestre el cumplimiento de un objetivo o una meta en determinado proceso, proporcionando la información de apoyo para la toma de decisiones, así como también el planteamiento de políticas y estrategias para solucionar el problema.

2.6.4 Patrones para Especificación de Indicadores

Los indicadores pueden ser valores, unidades, índices, series estadísticas, entre otros. Deben ser relevantes, sostenibles y comprensibles, por lo tanto un indicador debe tener las siguientes características (Mondragón, 2014).

- **Nombre:** La identificación del indicador es primordial, debe ser concreto y debe definir claramente la función que cumple.
- **Forma de cálculo:** Cuando se trata de indicadores cuantitativos, se debe tener muy claro la fórmula matemática para el cálculo de su valor, lo cual implica la identificación exacta de los factores o variables que lo conforman y la manera en la que ellos se relacionan.
- **Unidades:** La manera en la que se expresa el valor de determinado indicador está dada por las unidades, las cuales son seleccionadas en función de las necesidades de la organización.
- **Glosario:** Es fundamental que el indicador se encuentre documentado, es decir, se debe especificar de manera precisa los valores que se relacionan en su cálculo. Por lo general, las organizaciones cuentan con un documento o manual de indicadores, en el cual se exponen todos los aspectos relevantes de los indicadores que se manejan.

Capítulo 3

MARCO APLICATIVO

En este capítulo se toman en cuenta las actividades llevadas a cabo durante la implementación de la solución de inteligencia de negocio que satisfacen los requerimientos del caso de estudio.

3.1 Planificación del Proyecto

La planificación es un proceso mediante el cual se eligen los pasos a seguir con el fin de modificar de manera favorable una situación determinada. Los recursos disponibles en la organización influyen directamente sobre la capacidad que ésta posee para evolucionar o resolver un problema. Por esa razón, la planificación es un elemento clave para el éxito del proyecto.

A continuación, en la Tabla 2, se muestra la planificación relacionada con este caso de estudio:

Tabla 2: Planificación del Proyecto

Paso	Descripción	Estimado
1.	Selección del tema	9 días
1.1.A	Tormenta de ideas	4 días
1.1.B	Consulta de recursos informativos disponibles	7 días
1.2	Selección de recursos necesarios para trabajar	2 días
2.	Selección de método de desarrollo	2 días
2.1.A	Definición de entregables	2 días
2.1.B	Definición de arquitectura de la solución de inteligencia de negocio	2 días
3.	Selección de las bases de datos	2 días
3.1.A	Selección de las fuentes de datos transaccionales	2 días
3.1.B	Selección de las fuentes de datos externas	2 días
4.	Elaboración de informe	5 días
5.	Diseño del proyecto	11 días
5.1.A	Formulación de indicadores	4 días
5.1.B	Prueba simulada de indicadores	2 días
5.2	Selección de indicadores	2 día
5.3	Diseño lógico de las bases de datos	1 día
5.4	Diseño físico de las bases de datos	1 día
6.	Diseño de los procesos ETC	4 días
6.1	Migración de datos	1 día
7.	Elaboración de informe	5 días
8.	Procesamiento de los datos	21 días
8.1	Aplicación de los métodos de cálculo a los datos	10 días

8.2	Diseño lógico de la presentación de los resultados	10 días
8.3.A	Generación de pruebas y resultados	1 día
8.3.B	Depuración de errores	1 día
9.	Elaboración de informe	7 días
10.	Especificación de los perfiles de usuario	3 días
11.	Elaboración de reportes	9 días
11.1.A	Pruebas de funcionamiento	1 día
11.1.B	Depuración de errores	1 día
12.	Elaboración de informe	5 días

Según la planificación solo las actividades cuyo paso está enumerado mediante el uso de números y letras, por ejemplo 11.1.A y 11.1.B, se pueden llevar a cabo de forma paralela; mientras que el resto de ellas debe realizarse de manera secuencial.

3.2 Definición de Requisitos del Negocio

En esta etapa se definen los requerimientos que les brindan a los usuarios, de la solución de inteligencia de negocio, el apoyo necesario en el proceso de toma de decisiones. Estos requisitos se definen mediante los indicadores que se muestran a continuación, en la Tabla 3.

Tabla 3: Definición de Requisitos del Negocio

Nro de Reporte	Nombre del Reporte	Indicador	Fórmula	Unidades de Medida	Frecuencia de Medición	Criterios de Clasificación	Forma de Representación
1	Atención prestada por Servicio de Orientación	Cantidad de Estudiantes Atendidos	Σ Cantidad de Estudiantes Atendidos	#	Semestral	En Total, Por Año, Por Período Académico, Por Trimestre, Por Mes, Por Fecha, Por Área de Atención, Por Tipo de Atención, Por Actividad, Por Universidad, Por Facultad, Por Escuela, Por Especialista, Por Semestre de Licenciatura, Por Año de Licenciatura, Por Género, Por Edad, Por Estado Civil, Por Estudiante	Columna Agrupada
2	Rendimiento Académico del Estudiante	Promedio General de Notas	Σ Notas / Cantidad de Asignaturas	#	Semestral	En Total, Por Año, Por Período Académico, Por Trimestre, Por Mes, Por Fecha, Por Licenciatura, Por Estudiante, Por Asignatura, Por Tipo de Examen, Por Docente	Columna Agrupada
		Promedio Ponderado de Notas	Σ (Nota*UC) / Σ UC	#	Semestral		Columna Agrupada
		Promedio General de Notas de Asignaturas Aprobadas en General	Σ Notas de Asignaturas Aprobadas / Cantidad de Asignaturas Aprobadas en General	#	Semestral		Columna Agrupada
		Promedio Ponderado de Notas de Asignaturas Aprobadas en General	Σ (Nota *UC de Asignatura Aprobada en General) / Σ UC de Asignaturas Aprobadas en General	#	Semestral		Columna Agrupada

2	Rendimiento Académico del Estudiante	Cantidad de asignaturas y UC inscritas, cursadas, retiradas, aprobadas, aprobadas por equivalencia, aprobadas en general y aplazadas	Σ Cantidad de asignaturas y UC inscritas, cursadas, retiradas, aprobadas, aprobadas por equivalencia, aprobadas en general y aplazadas	#	Semestral	En Total, Por Año, Por Período Académico, Por Trimestre, Por Mes, Por Fecha, Por Licenciatura, Por Estudiante, Por Asignatura, Por Tipo de Examen, Por Docente, Por Nota	Columna Agrupada
		Ratio de Reprobación	Cantidad de Asignaturas Aplazadas / Cantidad de Asignaturas Aprobadas en General	#	Semestral		Columna Agrupada
		Eficiencia	Σ UC de Asignaturas Aprobadas en General / Σ UC de Asignaturas Cursadas	#	Semestral		Columna Agrupada
		Efectividad	Σ UC de Asignaturas Aprobadas en General / Σ UC de Asignaturas Inscritas	#	Semestral		Columna Agrupada

3.3 Selección de Productos de Software

Las aplicaciones seleccionadas para implementar esta solución de inteligencia de negocio son las siguientes:

- **PostgreSQL:** Es un sistema de gestión de bases de datos objeto-relacional, distribuido bajo licencia de software libre y con su código fuente disponible para desarrolladores. Es el sistema de gestión de bases de datos de código abierto más potente del mercado y en sus últimas versiones no tiene nada que envidiarle a otros sistemas de gestión de bases de datos comerciales.

PostgreSQL utiliza un modelo cliente/servidor y usa multiprocesos en vez de multihilos para garantizar la estabilidad del sistema. Una falla en uno de los procesos no afectará al resto y el sistema continuará funcionando. A continuación se muestra la interfaz de la aplicación en la Figura 12.

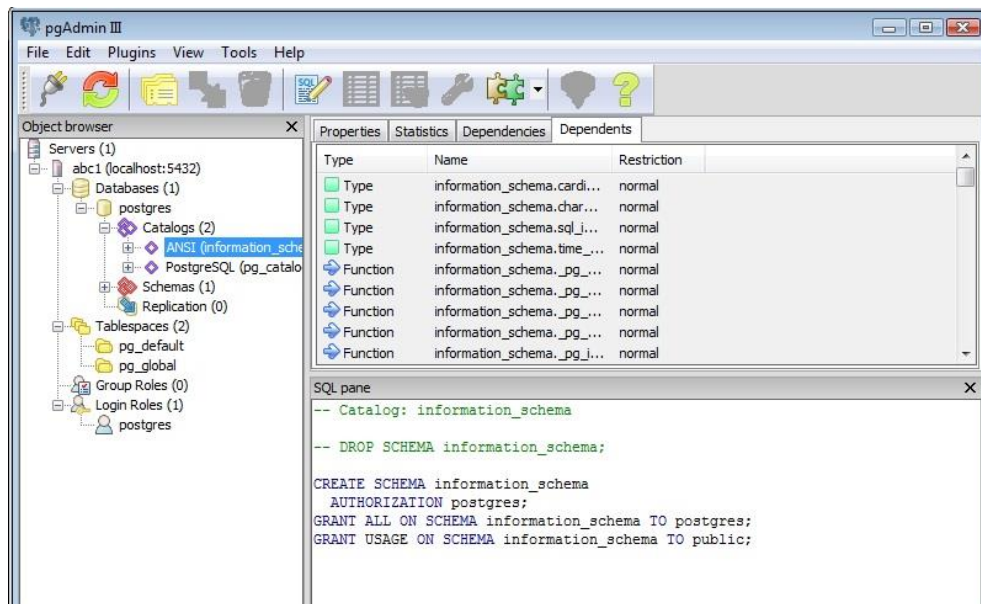


Figura 12: Interfaz de PostgreSQL
(PostgreSQL, 2015)

- **Pentaho Data Integration:** Permite extraer datos de diferentes fuentes, consolidarlos y cargarlos en un repositorio analítico. Su entorno de trabajo está formado de tal manera que la implementación y el mantenimiento de la solución de inteligencia de negocio sean más sencillos y más rápidos así como también permite al desarrollador incorporar fácilmente nuevas fuentes de información e indicadores. A continuación se muestra la interfaz de la aplicación en la Figura 13.

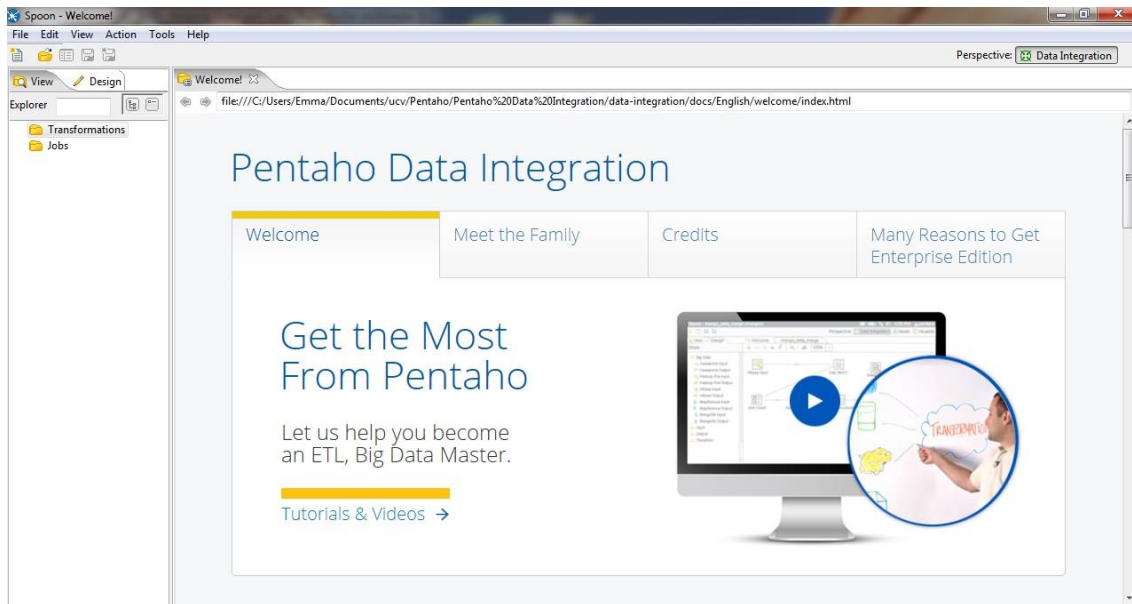


Figura 13: Interfaz de Pentaho Data Integration
(Pentaho, 2014)

• **Pentaho Schema Workbench:** Es un motor de análisis de negocio de código abierto, que permite a organizaciones de cualquier tamaño darle a los usuarios acceso a los datos de manera interactiva. Pentaho le brinda a los usuarios la oportunidad de construir una solución de inteligencia de negocio con Pentaho Schema Workbench como motor de Procesamiento Analítico en Línea, OLAP, ya que a través de él se pueden realizar consultas a la base de datos multidimensional y obtener información acerca del estado de los procesos de negocio de la organización. A continuación se muestra la interfaz de la aplicación en la Figura 14.

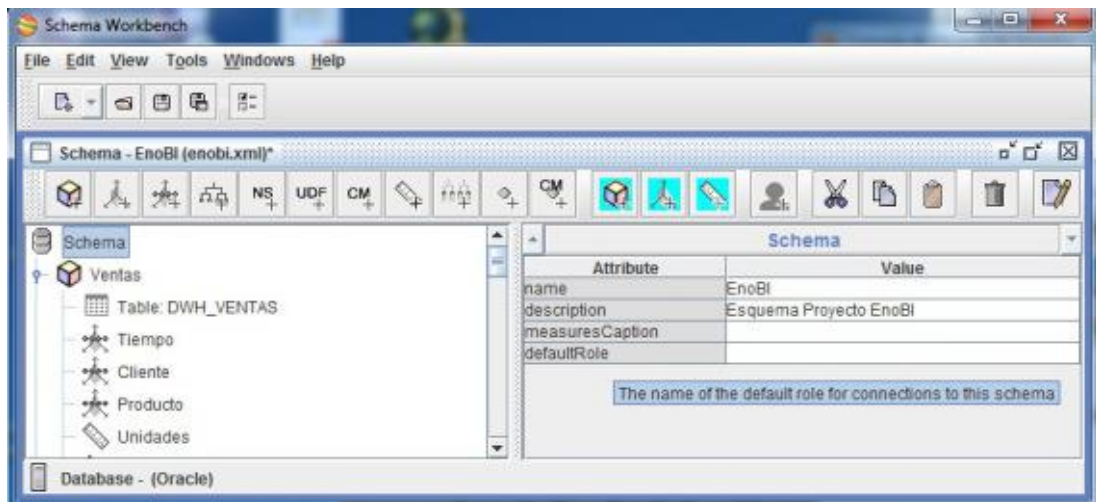


Figura 14: Interfaz de Mondrian Schema Workbench

- **Pentaho Business Analytics Platform:** Es una aplicación que permite transformar todos los datos en información significativa ya que, a través de JPivot, se utiliza para construir tablas de forma dinámica las cuales muestran los resultados de las consultas en función de los requerimientos del usuario, facilitando el análisis de la información y brindando apoyo a la toma de decisiones. A continuación se muestra la interfaz de la aplicación en la Figura 15.



Figura 15: Interfaz de Pentaho Business Analytics Platform

3.4 Diseño Lógico

En la siguiente sección se encuentra el diseño lógico tanto del Área de Almacenamiento Intermedio como del Datamart que forma parte de la solución de inteligencia de negocio relacionada con este caso de estudio.

3.4.1 Diseño Lógico del Área de Almacenamiento Intermedio

A continuación se muestra en la Figura 16 el diseño lógico del Área de Almacenamiento Intermedio que guarda la información relacionada con el proceso de negocio llamado Servicio de Orientación.

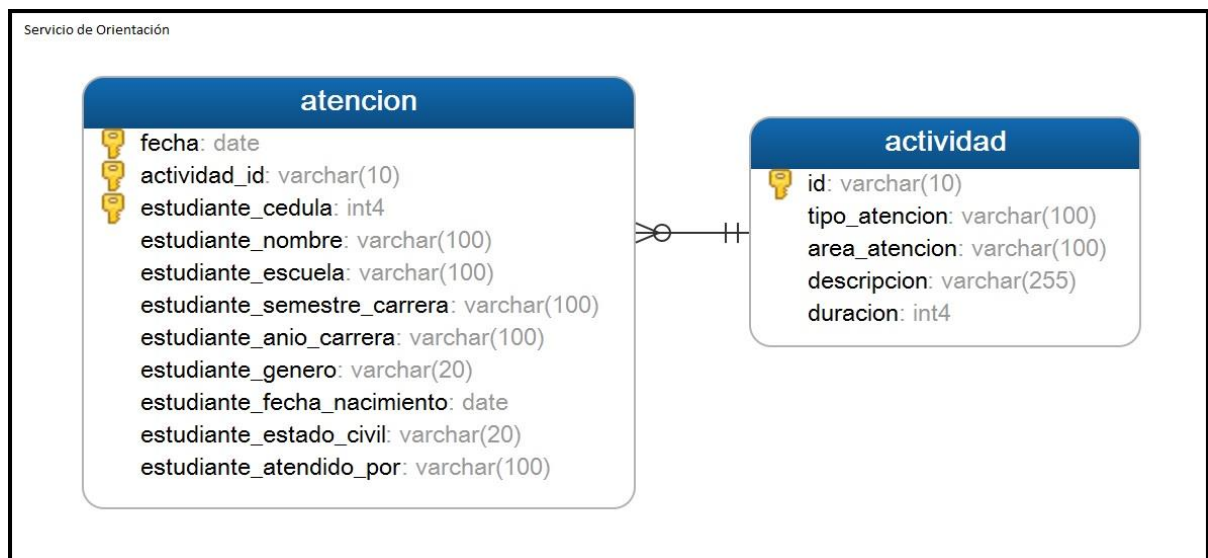


Figura 16: Modelo Relacional del Proceso de Negocio "Servicio de Orientación"

En la Figura 17 se muestra el Área de Almacenamiento Intermedio que contiene información relacionada con el proceso de negocio llamado Rendimiento Académico.

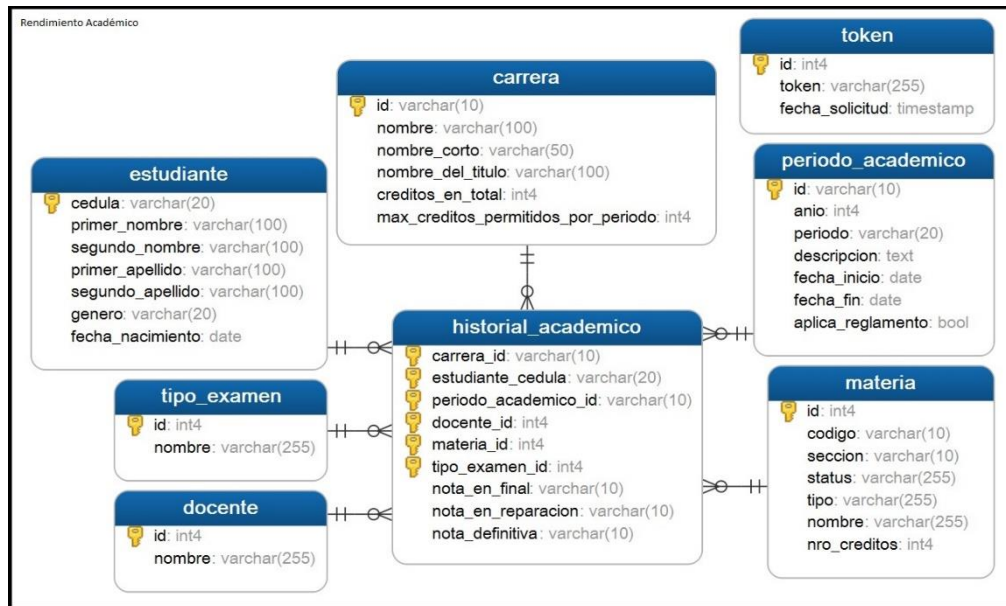


Figura 17: Modelo Relacional del Proceso de Negocio "Rendimiento Académico"

3.4.2 Diseño Lógico del Datamart

En la Figura 18, se muestra el diseño lógico del Datamart que guarda la información relacionada con la atención que presta el Servicio de Orientación a la población estudiantil.

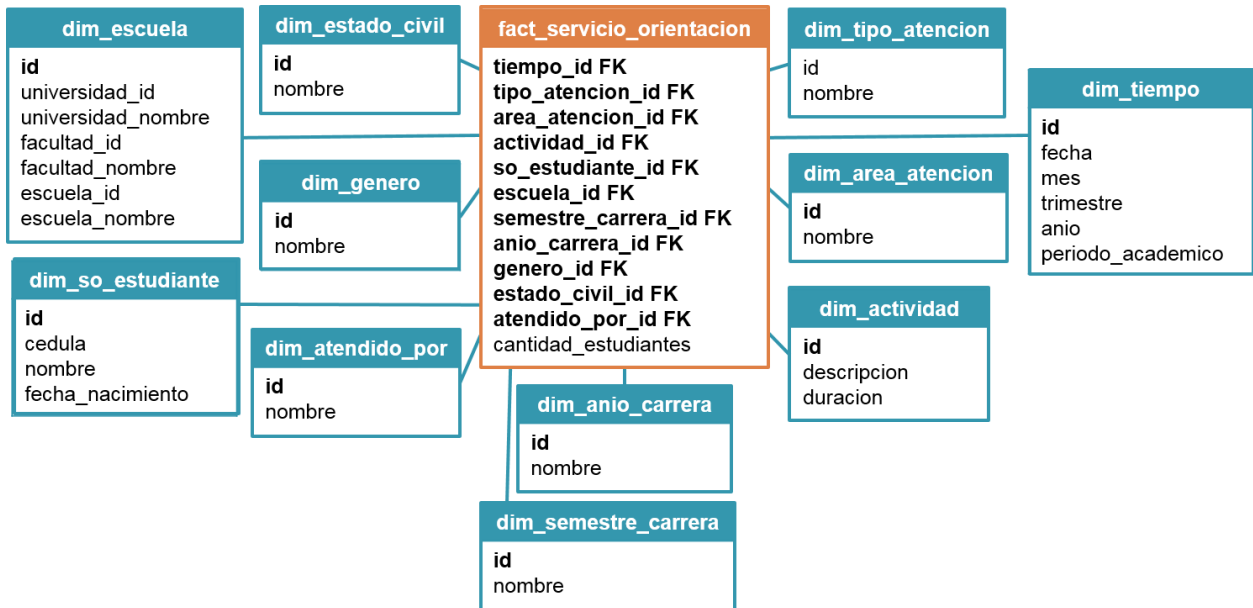


Figura 18: Modelo Dimensional del Proceso de Negocio "Servicio de Orientación"

En la Figura 19 se muestra el diseño lógico del Datamart que almacena la información relacionada con el proceso de negocio llamado Rendimiento Académico.

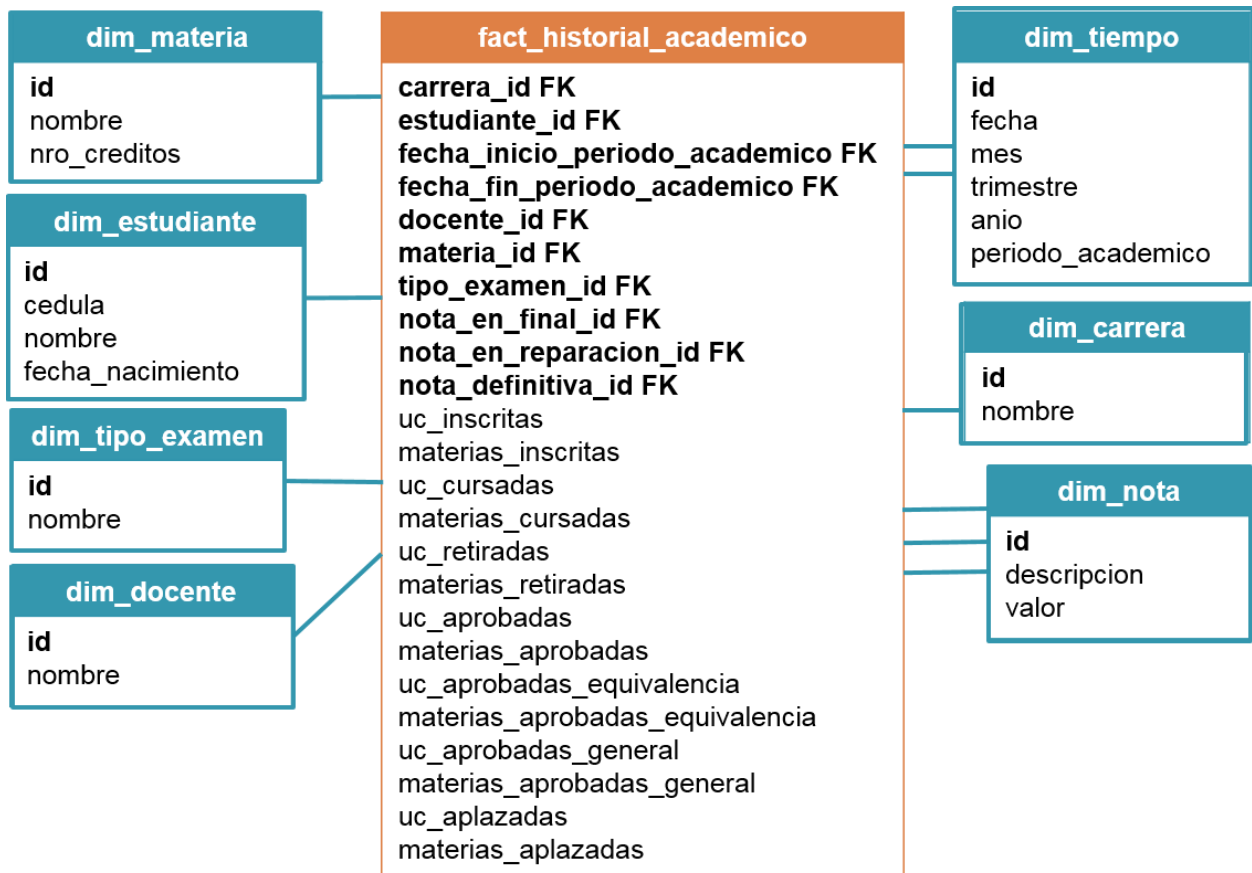


Figura 19:Modelo Dimensional del Proceso de Negocio "Rendimiento Académico"

3.5 Diseño Físico

A continuación se muestra el diseño físico del Área de Almacenamiento Intermedio y el Datamart que forma parte de la solución de inteligencia de negocio relacionada con este caso de estudio.

3.5.1 Diseño Físico del Área de Almacenamiento Intermedio

La especificación de las tablas del diseño físico del Área de Almacenamiento Intermedio se puede ver con detalle en el Anexo A. A continuación se explica la especificación de la Tabla *actividad* del proceso de negocio “Servicio de Orientación”, siguiendo este patrón:

Servicio de Orientación

Nombre de la Tabla: actividad			
Descripción: Tabla que contiene información de las actividades realizadas por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
tipo_atencion	100	character varying	Tipo de atención que se presta al estudiante durante la actividad
area_atencion	100	character varying	Área de atención relacionada con la actividad realizada
descripcion	255	character varying	Descripción de la actividad
duración	4	integer	Duración de la actividad
Relaciones:		Campos Clave: id	

3.5.2 Diseño Físico del Datamart

La especificación de las tablas del diseño físico del Datamart se puede ver con detalle en el Anexo B. A continuación se explica la especificación de la Tabla *dim_so_estudiante* del proceso de negocio “Servicio de Orientación” siguiendo el patrón mostrado:

Servicio de Orientación

Nombre de la Tabla: dim_so_estudiante			
Descripción: Dimensión que contiene información acerca de los estudiantes atendidos por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
cedula	20	character varying	Cédula del estudiante
nombre	100	character varying	Nombre del estudiante
fecha_nacimiento	-	date	Fecha de nacimiento del estudiante
Relaciones:		Campos Clave:	

3.6 Diseño de Procesos de Extracción, Transformación y Carga

A continuación se exponen los procesos de Extracción, Transformación y Carga que se llevan a cabo con el fin de poblar el Área de Almacenamiento Intermedio y el Datamart que forman parte de la solución de inteligencia de negocio relacionada con este caso de estudio.

3.6.1 Diseño de Procesos de Extracción, Transformación y Carga del Área de Almacenamiento Intermedio

El diseño de los procesos de Extracción, Transformación y Carga del Área de Almacenamiento Intermedio se puede ver con detalle en el Anexo C. Estos procesos fueron diseñados siguiendo el patrón que se muestra a continuación para el caso del Proceso *Poblar la Tabla Periodo Académico* del proceso de negocio “Rendimiento Académico”.

Rendimiento Académico

Se diseñaron nueve procesos de Extracción, Transformación y Carga con la finalidad de poblar la sección del Área de Almacenamiento Intermedio relacionada con el proceso de negocio llamado “Rendimiento Académico”.

Nombre del Proceso: Poblar la Tabla periodo_academico

Descripción: Representa la carga de la tabla periodo_academico

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	periodos.json	Contiene información acerca de los periodos académicos transcurridos a lo largo del tiempo

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	character varying(10)	Texto	No nulo	No tiene
anio		Integer	Numérico	No negativo	No tiene
periodo		character varying(20)	Texto		No tiene
descripción		character varying(255)	Texto		No tiene
fecha_inicio		date	Fecha		No tiene
fecha_fin		date	Fecha		No tiene
aplica_reglamento		boolean	Lógico		No tiene

Fuentes de Datos

Archivo JSON: periodos				
Nombre	Llave	Tipo	Formato	Consideración Importante
id		character varying(10)	Texto	
anio		Integer	Numérico	
periodo		character varying(20)	Texto	
descripción		character varying(255)	Texto	
fecha_inicio		date	Fecha	
fecha_fin		date	Fecha	
aplica_reglamento		Boolean	Lógico	

Tabla Destino

Tabla: periodo_academico		
Campo	Tipo	Correspondencia
id	character varying(10)	periodos.id
anio	Integer	periodos.anio
periodo	character varying(20)	periodos.periodo
descripción	character varying(255)	periodos.descripcion
fecha_inicio	date	periodos.fecha_inicio
fecha_fin	date	periodos.fecha_fin
aplica_reglamento	Boolean	periodos.aplica_reglamento

Proceso realizado:

- Utilizar autorización para solicitar el archivo periodos.json
- Extraer registros del archivo periodos.json
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla periodo_academico con la información correspondiente

En la Figura 20 se muestra el Proceso ETC, ejecutado a través de Pentaho Data Integration, para poblar la Tabla periodo_academico



Figura 20: Proceso ETC utilizado para Poblar la Tabla periodo_academico

3.6.2 Diseño de Procesos de Extracción, Transformación y Carga del

Datamart

El diseño de todos los procesos de Extracción, Transformación y Carga del Datamart se puede ver con detalle en el Anexo D. Seguidamente se explica el Proceso ETC *Poblar la Dimensión Estudiantes Atendidos* del proceso de negocio “Servicio de Orientación” siguiendo el patrón que se muestra a continuación:

Servicio de Orientación

Se diseñaron doce procesos de Extracción, Transformación y Carga con la finalidad de poblar la sección del Datamart relacionada con el proceso de negocio llamado “Servicio de Orientación”.

Nombre del Proceso: Poblar la Dimensión Estudiantes Atendidos

Descripción: Representa la carga de la tabla dim_so_estudiante

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
cedula	PK	character varying(20)	Texto	No nulo	No tiene
nombre		character varying(100)	Texto		No tiene
fecha_nacimiento		date	Fecha		No tiene

Fuentes de Datos

Tabla: atención				
Nombre	Llave	Tipo	Formato	Consideración Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	
estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: dim_so_estudiante		
Campo	Tipo	Correspondencia
cedula	character varying(20)	atencion.estudiante_cedula
nombre	character varying(100)	atencion.estudiante_nombre
fecha_nacimiento	date	atencion.estudiante_fecha_nacimiento

Proceso realizado:

- Extraer registros de la tabla llamada atención en la cual se encuentra la información relacionada con los estudiantes atendidos por el Servicio de Orientación
- Ordenar los registros
- Eliminar los registros repetidos
- Elegir los campos de interés, su nombre, su formato y el tipo de dato de cada uno de ellos
- Cargar los registros en la tabla dim_so_estudiante con la información correspondiente

3.6.3 Esquema de Extracción

El esquema de dependencias entre los procesos de Extracción, Transformación y Carga que se requiere para poblar el Área de Almacenamiento Intermedio y el Datamart de la solución de inteligencia de negocio asociada a este caso de estudio es el siguiente:

Frecuencia de Ejecución: Semestral		
Número	Proceso	Dependencia
1	Poblar la Tabla actividad	-
2	Poblar la Tabla atención	-
3	Solicitud de Autorización	-
4	Poblar la Tabla periodo_academico	3
5	Poblar la Tabla carrera	3
6	Poblar la Tabla estudiante	3, 5
7	Poblar la Tabla materia	3, 5, 6
8	Poblar la Tabla tipo_examen	3, 5, 6
9	Poblar la Tabla docente	3, 5, 6
10	Poblar la Tabla historial_academico	3, 5, 6
11	Solicitud de Cierre de Sesión	3
12	Poblar la Dimensión Estudiantes Atendidos	2
13	Poblar la Dimensión Tiempo	4
14	Poblar la Dimensión Tipo de Atencion	1
15	Poblar la Dimensión Área de Atención	1
16	Poblar la Dimensión Actividad	1
17	Poblar la Dimensión Escuela	2
18	Poblar la Dimensión Atendido Por	2
19	Poblar la Dimensión Semestre de Licenciatura	2
20	Poblar la Dimensión Año de Licenciatura	2
21	Poblar la Dimensión Género	2
22	Poblar la Dimensión Estado Civil	2
23	Poblar la Tabla de Hechos Servicio de Orientacion	2, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
24	Poblar la Dimensión Periodo Académico	4
25	Poblar la Dimensión Carrera	5
26	Poblar la Dimensión Estudiante	6
27	Poblar la Dimensión Materia	7
28	Poblar la Dimensión Tipo Examen	8
29	Poblar la Dimensión Docente	9
30	Poblar la Dimensión Nota	10
31	Poblar la Tabla de Hechos Historial Academico	10, 24, 25, 26, 27, 28, 29, 30

Se sugiere que la ejecución periódica de los procesos de Extracción, Transformación y Carga sea realizada por el administrador de la solución de inteligencia de negocio pues éste debe garantizar que se cuenta con los recursos necesarios para ser llevada a cabo.

3.7 Especificación de Aplicación de Usuario

3.7.1 Especificación de los Perfiles de Usuario

Los perfiles de usuario determinan los privilegios que puede tener cada usuario. Los perfiles utilizados para la solución de inteligencia de negocio que se adapta a este caso de estudio son los siguientes:

- **Administrador:** Tiene acceso a todas las herramientas que brinda la solución de inteligencia de negocio en cada una de sus etapas. Puede crear, eliminar, modificar y consultar todas las plantillas de reportes; así como también gestionar los perfiles de usuario de la solución de inteligencia de negocio y las fuentes de datos.
- **Usuario Avanzado:** Puede crear, eliminar, modificar y consultar todas las plantillas de reportes.
- **Usuario Registrado:** Puede consultar todas las plantillas de reportes.

El perfil de Usuario Avanzado y el de Usuario Registrado fueron asignados al personal del Servicio de Orientación de la Facultad de Ciencias.

3.7.2 Diseño de Cubos

Una vez que han sido construidas las dimensiones y las tablas de hechos, se debe diseñar el esquema lógico del cubo, identificando las dimensiones que relaciona, junto con sus atributos y sus jerarquías. Adicionalmente, se deben establecer los hechos a medir. Para llevar a cabo esta tarea, se utilizó el componente de Pentaho que permite diseñar cubos en base al diseño lógico del Datamart y a los requerimientos del usuario relacionados con el caso de estudio. Esta herramienta se denomina Pentaho Schema Workbench. En la solución de inteligencia de negocio relacionada con el caso de estudio se implementaron dos cubos; el primero de ellos asociado con el proceso de negocio llamado “Servicio de Orientación” y el segundo con el proceso de negocio llamado “Rendimiento Académico”.

A continuación en la Figura 21 se muestra el cubo asociado con el proceso de negocio “Servicio de Orientación” el cual contiene once dimensiones y una medida.

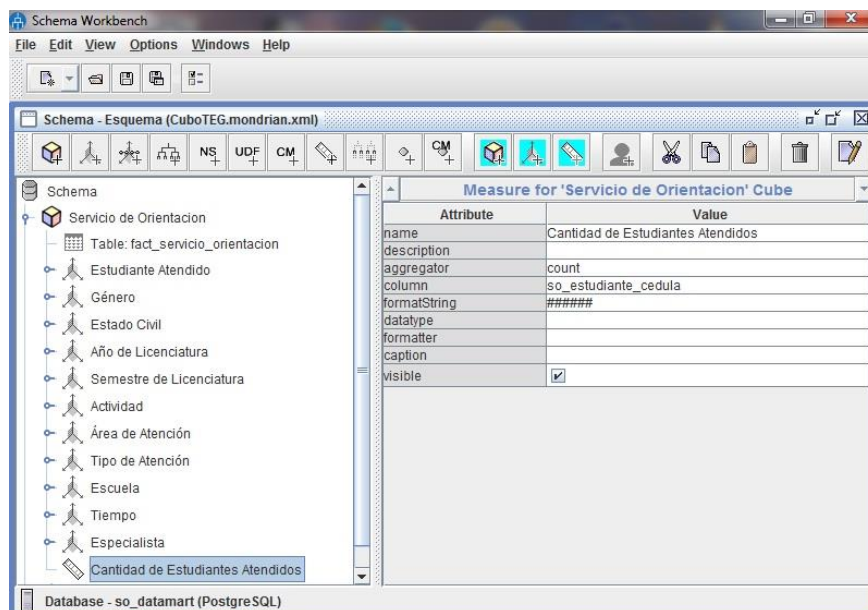


Figura 21: Cubo "Servicio de Orientación"

En la Figura 22 se muestra el cubo asociado con el proceso de negocio “Rendimiento Académico” el cual contiene siete dimensiones y catorce medidas.

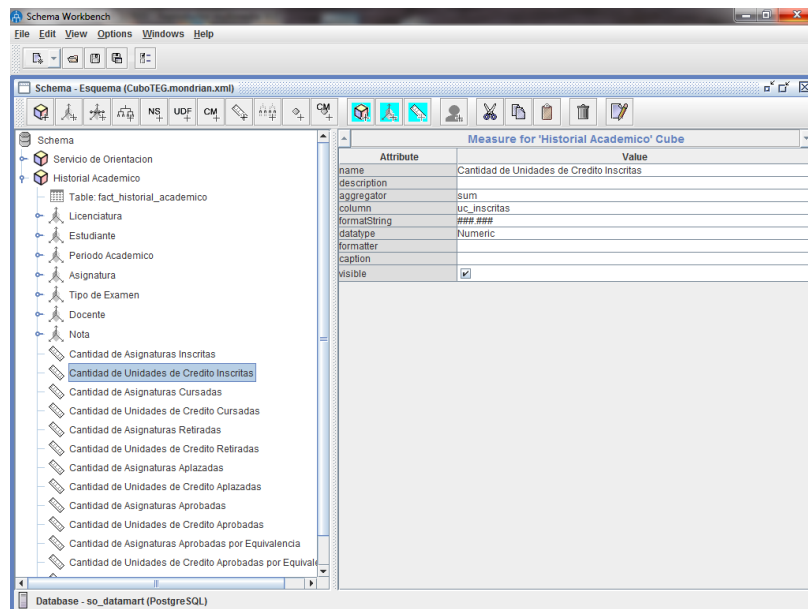


Figura 22: Cubo "Historical Académico"

3.7.3 Reportes de la solución de inteligencia de negocio

A continuación se muestran las interfaces de la herramienta de análisis y consulta que forma parte de la solución de inteligencia de negocio asociada a este caso de estudio, así como también de los reportes generados correctamente por ésta.

En la Figura 23 se muestra la interfaz de inicio de sesión de la herramienta de análisis y consulta, mediante la cual los usuarios autorizados pueden acceder a la interfaz inicial del portal web donde se encuentran los reportes y los resultados de los indicadores.

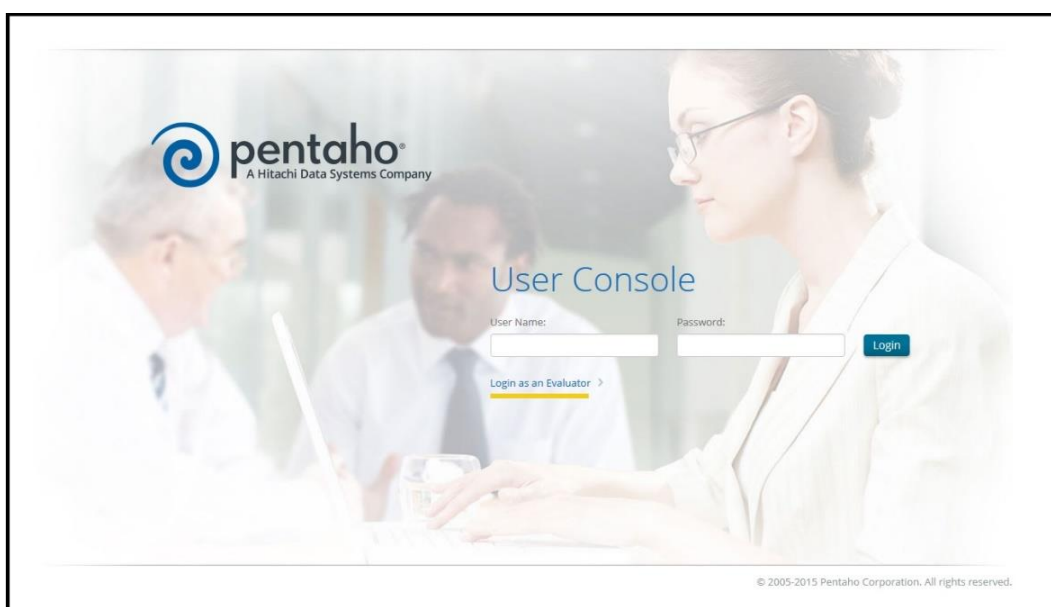


Figura 23: Interfaz de inicio de sesión de la herramienta de análisis y consulta

En la Figura 24 se muestra la interfaz inicial de la herramienta de análisis y consulta que forma parte de la solución de inteligencia de negocio mediante la cual los usuarios pueden acceder a los reportes relacionados con la labor que desempeña el Servicio de Orientación y con el rendimiento académico de la población estudiantil de la Facultad de Ciencias de la Universidad Central de Venezuela.

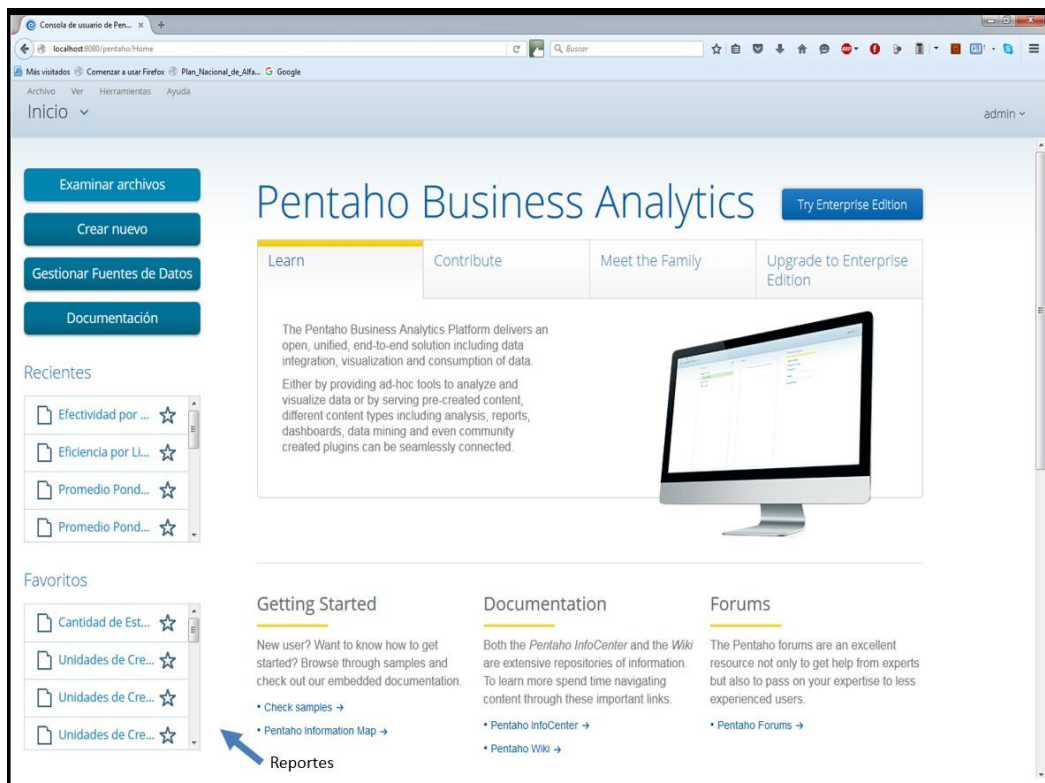


Figura 24: Interfaz inicial de la herramienta de análisis y consulta

En la Figura 25 se muestra el reporte relacionado con la cantidad de estudiantes atendidos por el Servicio de Orientación en cada una de las áreas de atención, así como también la cantidad de estudiantes atendidos en total.

Cabe destacar que este indicador ha sido calculado de manera tal que el usuario puede tener acceso a la cantidad de estudiantes atendidos pertenecientes o no a la Universidad Central de Venezuela, por Facultad, por Escuela, por año, por período académico, por trimestre, por mes, por fecha, por actividad realizada, por el tipo de atención prestada, por año y por semestre de licenciatura que cursa el estudiante, por especialista que brinda atención al estudiante, por género, por edad y por estado civil del estudiante.

En el caso particular de la columna llamada estudiante atendido, este indicador muestra como resultado la cantidad de veces en las que fue atendido el estudiante.

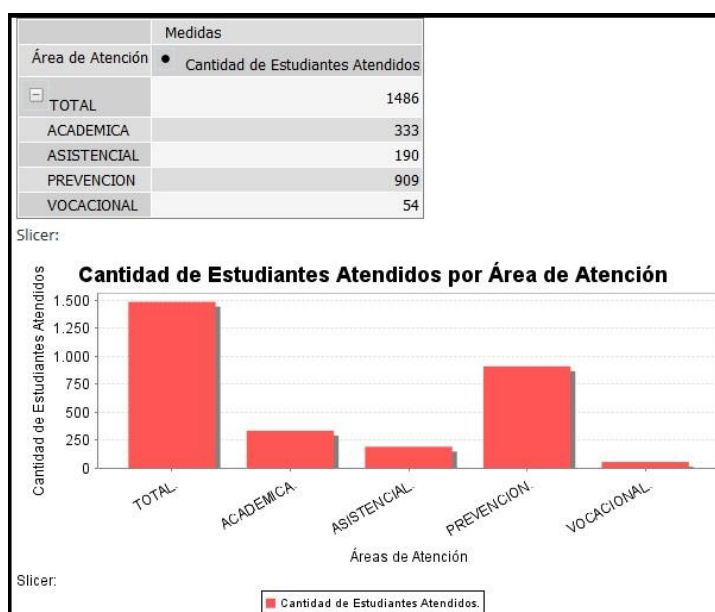


Figura 25: Cantidad de Estudiantes Atendidos por Área de Atención

En la Figura 26 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas inscritas y calificadas durante los períodos académicos 01-2012 y 02-2012 de la población estudiantil de la Facultad de Ciencias.

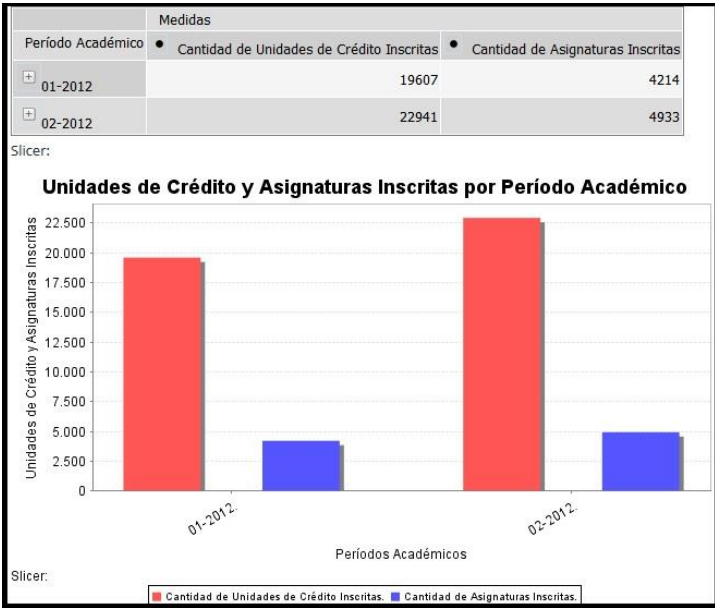


Figura 26: Unidades de Crédito y Asignaturas Inscritas por Período Académico

En la Figura 27 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas cursadas y calificadas, es decir que no fueron retiradas. Los resultados obtenidos se pueden visualizar por cada tipo de evaluación realizada así como también en total.

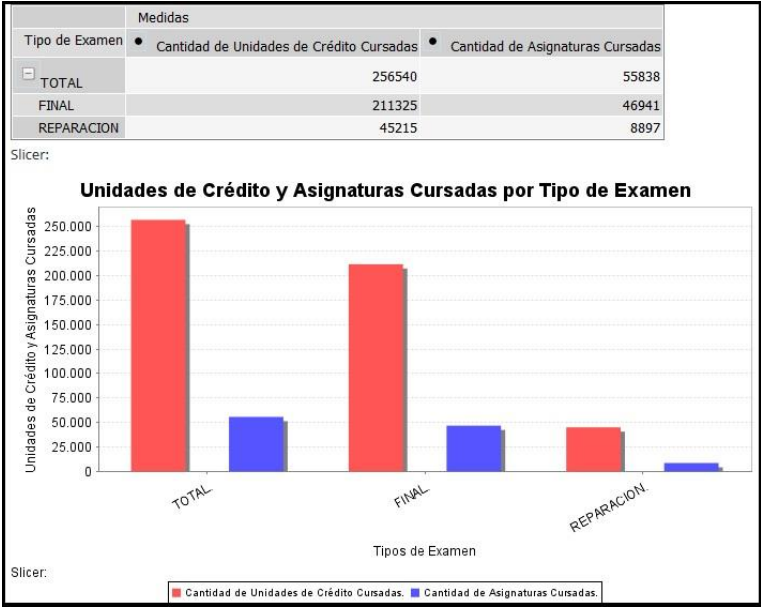


Figura 27: Unidades de Crédito y Asignaturas Cursadas por Tipo de Examen

En la Figura 28 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas retiradas por cada Licenciatura de la Facultad de Ciencias así como también de la población estudiantil en total.

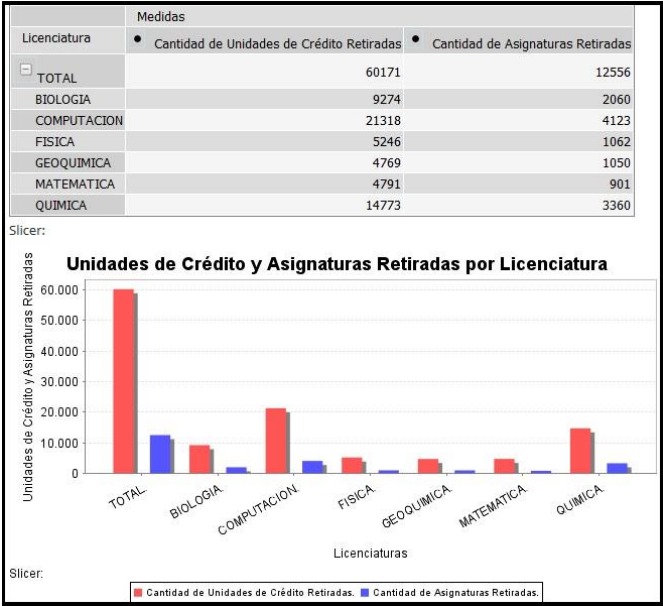


Figura 28: Unidades de Crédito y Asignaturas Retiradas por Licenciatura

En la Figura 29 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas aprobadas por el estudiante en la Facultad de Ciencias. Los resultados obtenidos se pueden visualizar por periodo académico, tomando en cuenta los periodos 01-2012 y 02-2012.

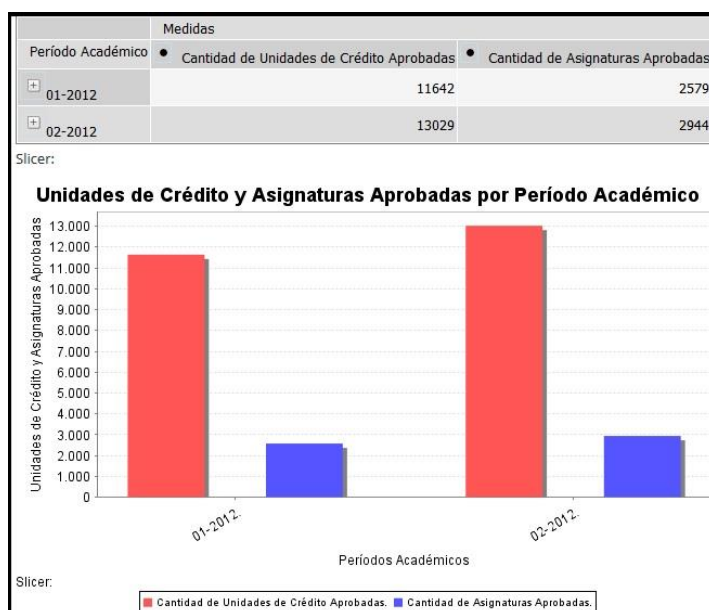


Figura 29: Unidades de Crédito y Asignaturas Aprobadas por Periodo Académico

En la Figura 30 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas aprobadas por equivalencia. Los resultados obtenidos se pueden visualizar por cada Licenciatura de la Facultad de Ciencias así como también de la población estudiantil en total.

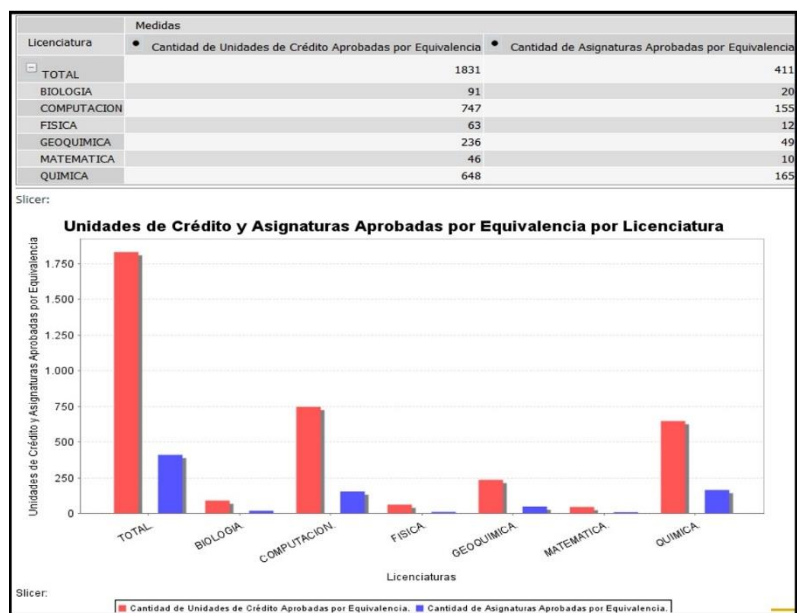


Figura 30: Unidades de Crédito y Asignaturas Aprobadas por Equivalencia por Licenciatura

En la Figura 31 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas que el estudiante ha aprobado, bien sea en la Facultad de Ciencias o por equivalencia. Los resultados obtenidos se encuentran clasificados mediante la calificación obtenida por el estudiante.

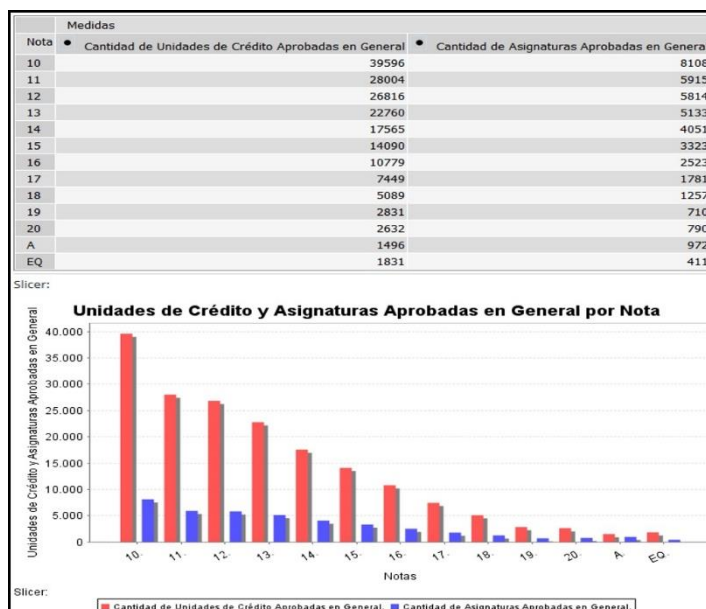


Figura 31: Unidades de Crédito y Asignaturas Aprobadas en General por Nota

En la Figura 32 se muestra el reporte relacionado con la cantidad de unidades de crédito y de asignaturas aplazadas. Los resultados obtenidos se pueden visualizar por cada tipo de evaluación realizada así como también en total.

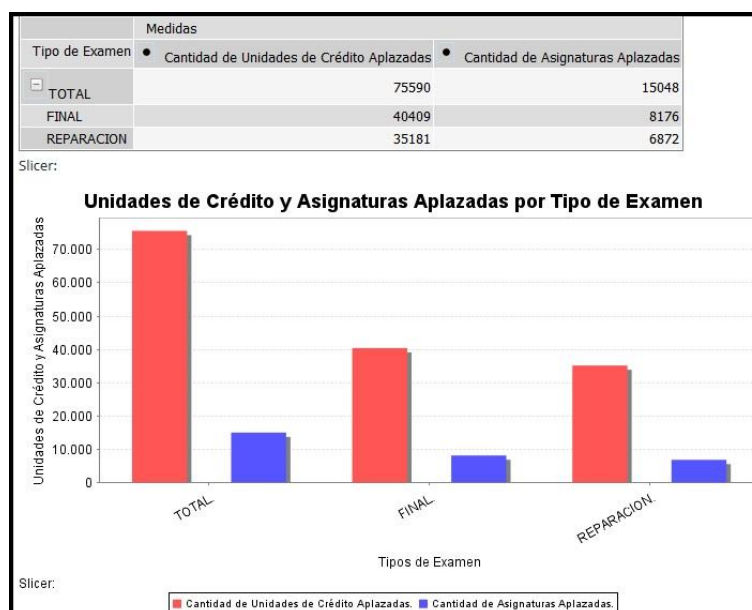


Figura 32: Unidades de Crédito y Asignaturas Aplazadas por Tipo de Examen

En la Figura 33 se muestra el reporte relacionado con el ratio de reprobación, el cual indica cuántas asignaturas han sido reprobadas por cada asignatura que ha sido aprobada. Los resultados obtenidos se pueden visualizar por periodo académico, tomando en cuenta los periodos 01-2012 y 02-2012.

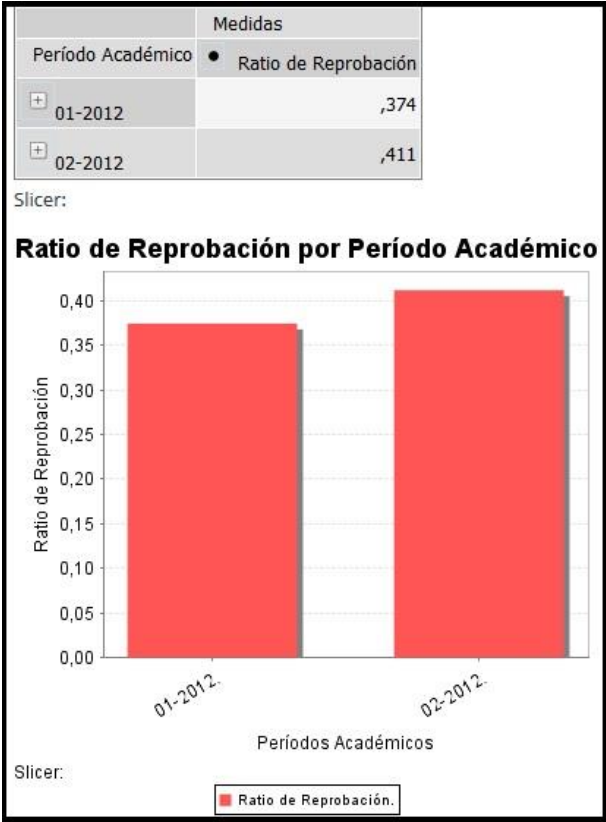


Figura 33: Ratio de Reprobación por Período Académico

En la Figura 34 se muestra el reporte relacionado con el promedio general de notas de aquellos registros cuya calificación tenga un valor numérico. Los resultados obtenidos se pueden visualizar por cada tipo de evaluación realizada así como también en total.

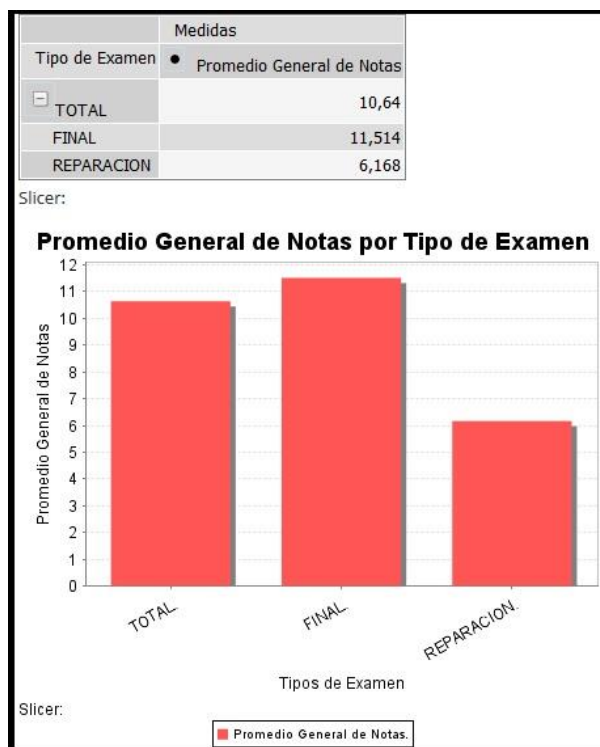


Figura 34: Promedio General de Notas por Tipo de Examen

En la Figura 35 se muestra el reporte relacionado con el promedio general de notas de asignaturas aprobadas tomando en cuenta únicamente aquellos registros cuya calificación tenga un valor numérico mayor o igual a diez. Los resultados obtenidos se pueden visualizar por cada Licenciatura así como también de la población estudiantil de la Facultad de Ciencias en total.

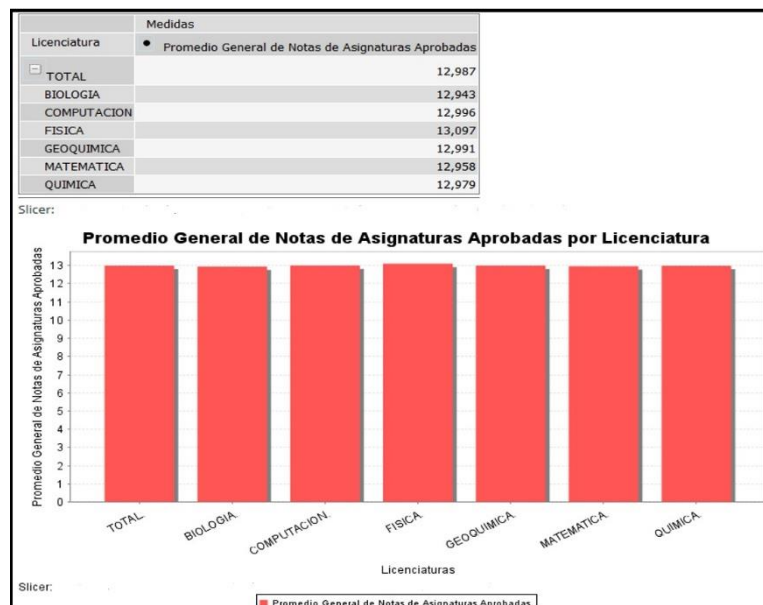


Figura 35: Promedio General de Notas de Asignaturas Aprobadas por Licenciatura

En la Figura 36 se muestra el reporte relacionado con el promedio ponderado de notas de aquellos registros cuya calificación tenga un valor numérico. Los resultados obtenidos se pueden visualizar por periodo académico, tomando en cuenta los periodos 01-2012 y 02-2012.

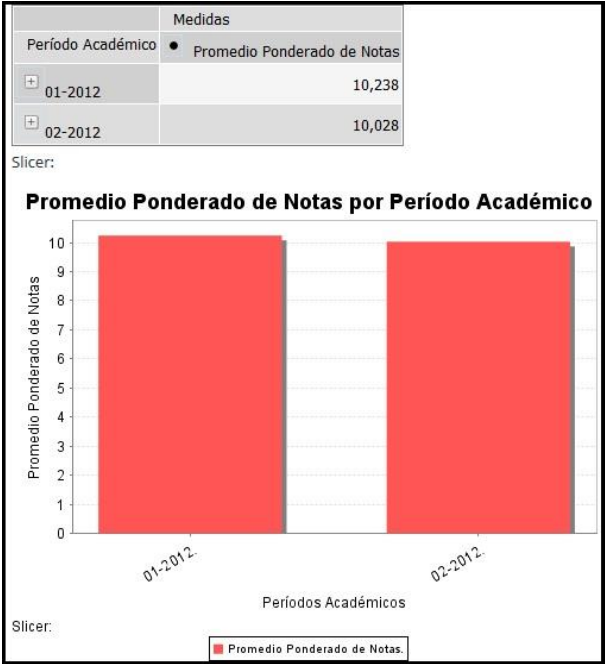


Figura 36: Promedio Ponderado de Notas por Período Académico

En la Figura 37 se muestra el reporte relacionado con el promedio ponderado de notas de asignaturas aprobadas, tomando en cuenta únicamente aquellos registros cuya calificación tenga un valor numérico mayor o igual a diez. Los resultados obtenidos se pueden visualizar por cada tipo de evaluación realizada así como también en total.

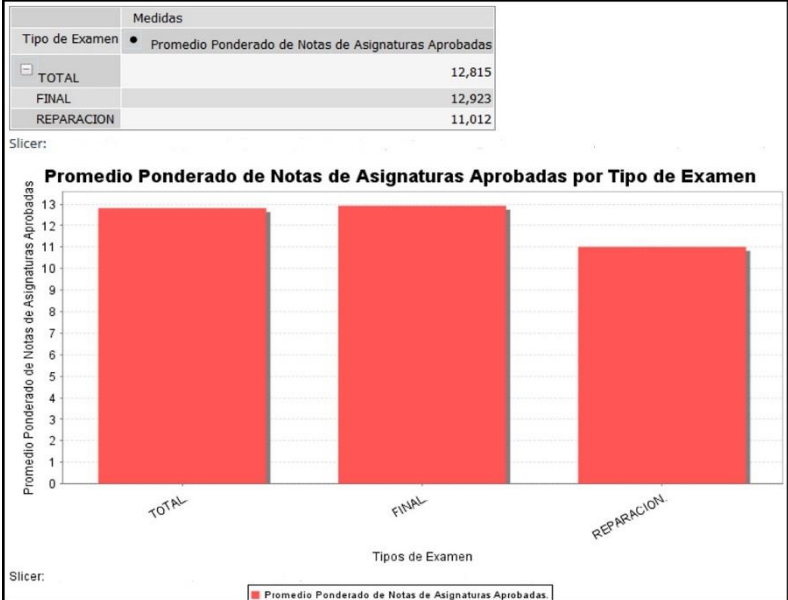


Figura 37: Promedio Ponderado de Notas de Asignaturas Aprobadas por Tipo de Examen

En la Figura 38 se muestra el reporte relacionado con la eficiencia, es decir, con la cantidad de unidades de crédito aprobadas en general en comparación con la cantidad de unidades de crédito cursadas. Los resultados obtenidos se pueden visualizar por cada Licenciatura así como también de la población estudiantil de la Facultad de Ciencias en total.

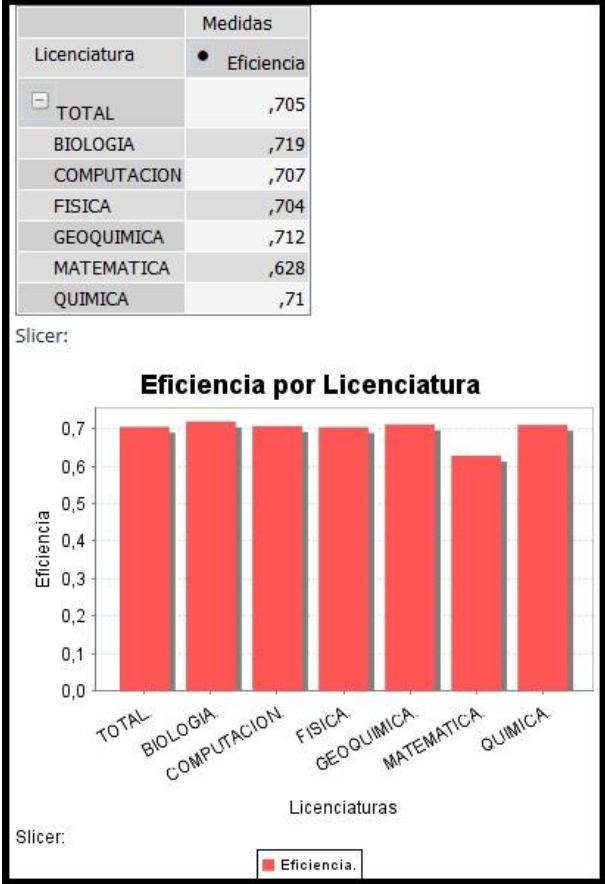


Figura 38: Eficiencia por Licenciatura

En la Figura 39 se muestra el reporte relacionado con la efectividad, es decir, con la cantidad de unidades de crédito aprobadas en general en comparación con la cantidad de unidades de crédito inscritas. Los resultados obtenidos se pueden visualizar por cada Licenciatura así como también de la población estudiantil de la Facultad de Ciencias en total.

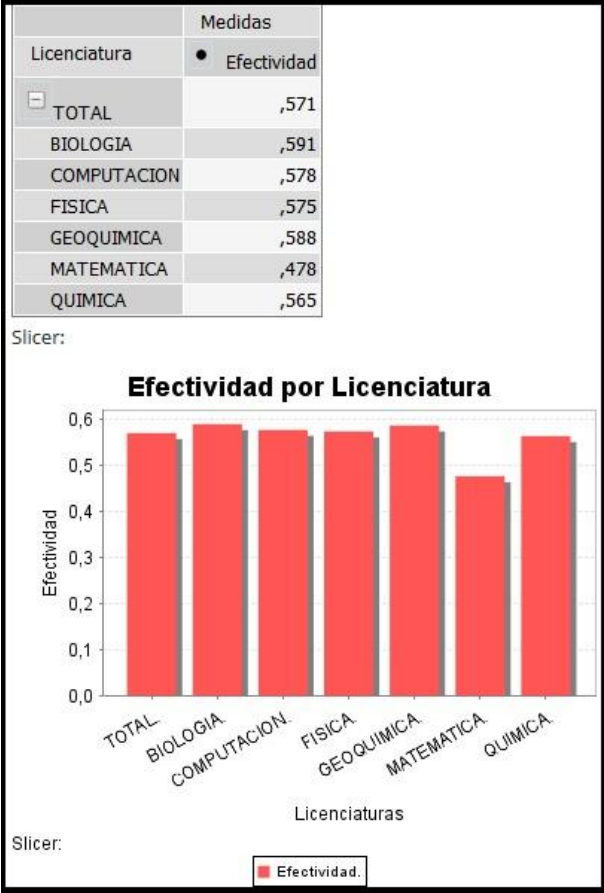


Figura 39: Efectividad por Licenciatura

3.8 Integración y Despliegue

En esta etapa del desarrollo se acopla cada componente de la solución de inteligencia de negocio de tal manera que la ejecución de los procesos de Extracción, Transformación y Carga sea periódica, con el fin de tener acceso a las diversas fuentes de datos con las que se poblará el Área de Almacenamiento Intermedio y ésta a su vez sirva como fuente de datos del Datamart.

Una vez diseñados los cubos y poblado el Datamart con los datos correspondientes, se generan los reportes que pueden mostrar el estado de los indicadores relacionados con la labor que desempeña el Servicio de Orientación y con el rendimiento académico de la población estudiantil de la Facultad de Ciencias.

El usuario final puede crear, modificar, eliminar y consultar estos reportes mediante el portal web de la solución de inteligencia de negocio, cuyo principal propósito es brindar apoyo a la toma de decisiones, a través de Pentaho Business Analytics Platform.

Con el fin de garantizar la calidad de los datos al generar los reportes asociados a los procesos de negocio, los resultados que se muestran a través de los indicadores relacionados con el rendimiento académico de la población estudiantil de la Facultad de Ciencias fueron comparados con los indicadores que se muestran en el historial académico de los estudiantes, el cual se encuentra disponible en la DCE. Asimismo, se validó la veracidad de los resultados obtenidos a través de los reportes que permiten evaluar la labor que desempeña el Servicio de Orientación; ya que se realizó un contraste entre los reportes existentes en esta dependencia y los reportes creados mediante el uso de la solución de inteligencia de negocio desarrollada.

3.9 Gestión del Proyecto

Durante el desarrollo del proyecto fue necesaria la gestión y redistribución de los recursos de tiempo disponibles con el fin de poder llevar a cabo las tareas relacionadas con el acceso a las diversas fuentes de datos, el procesamiento de éstos y su carga en el Datamart; el cálculo de los indicadores asociados con la labor que desempeña el Servicio de Orientación y con el rendimiento académico de la población estudiantil de la Facultad de Ciencias; así como también la adquisición de conocimientos acerca de los productos de software utilizados. Estas actividades requirieron un tiempo mayor del asignado inicialmente en la planificación.

Con el propósito de verificar si fueron satisfechos los requerimientos de los usuarios finales y por ende si se cumplió con sus expectativas, se tomaron en cuenta sus recomendaciones y se realizaron pruebas de aceptación a través de la siguiente encuesta:

Encuesta

Cantidad de usuarios: 3 (Personal del Servicio de Orientación)

1. ¿Se puede iniciar sesión y cerrar sesión de manera sencilla?
 - a. Totalmente de acuerdo
 - b. De acuerdo
 - c. Medianamente de acuerdo
 - d. En desacuerdo
2. ¿Se puede acceder a los reportes de manera sencilla?
 - a. Totalmente de acuerdo
 - b. De acuerdo
 - c. Medianamente de acuerdo
 - d. En desacuerdo
3. ¿Se utiliza un lenguaje natural y simple?
 - a. Totalmente de acuerdo
 - b. De acuerdo
 - c. Medianamente de acuerdo
 - d. En desacuerdo
4. ¿Se puede ver la información detallada fácilmente?
 - a. Totalmente de acuerdo
 - b. De acuerdo
 - c. Medianamente de acuerdo
 - d. En desacuerdo
5. ¿Los indicadores que se muestran en los reportes son de utilidad?
 - a. Totalmente de acuerdo
 - b. De acuerdo
 - c. Medianamente de acuerdo
 - d. En desacuerdo

Los resultados de la encuesta se muestran a continuación, en las Figuras 40, 41, 42, 43 y 44.

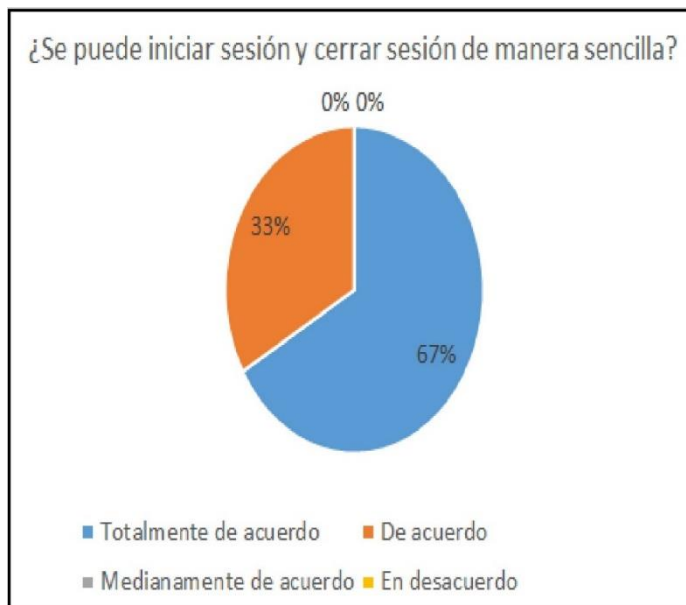


Figura 40: Resultado de la encuesta a la pregunta “¿Se puede iniciar sesión y cerrar sesión de manera sencilla?”

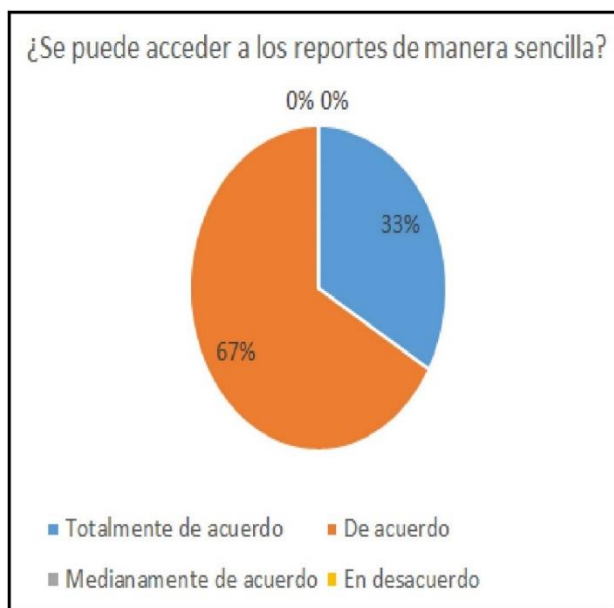


Figura 41: Resultado de la encuesta a la pregunta “¿Se puede acceder a los reportes de manera sencilla?”

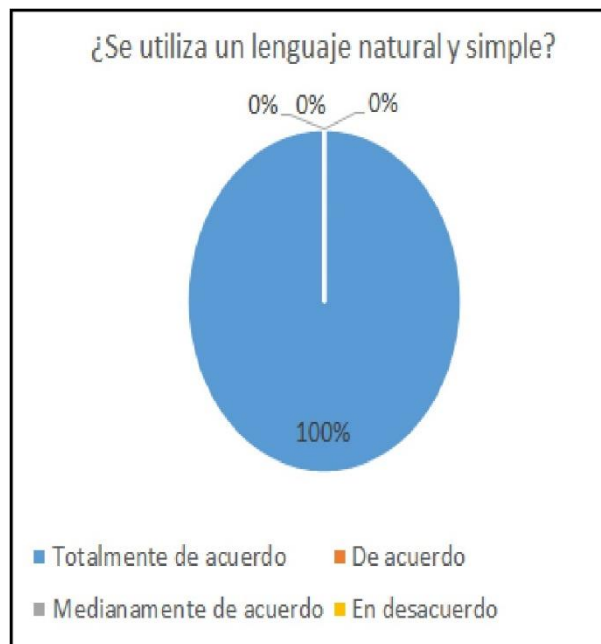


Figura 42: Resultado de la encuesta a la pregunta “¿Se utiliza un lenguaje natural y simple?”

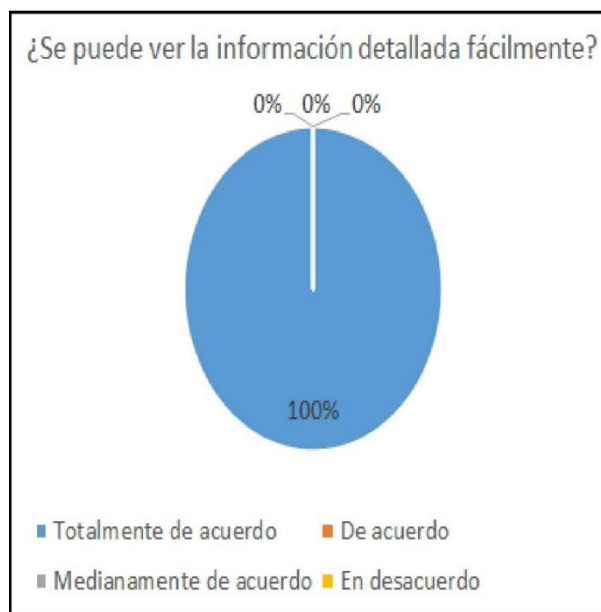


Figura 43: Resultado de la encuesta a la pregunta “¿Se puede ver la información detallada fácilmente?”



Figura 44: Resultado de la encuesta a la pregunta “¿Los indicadores que se muestran en los reportes son de utilidad?”

Mediante los resultados obtenidos en la encuesta realizada a los usuarios de la solución de inteligencia de negocio desarrollada, se determinó que ésta fue aceptada en alto porcentaje debido a su facilidad de uso y a que la información que se muestra a través de los reportes es de gran utilidad pues les brinda apoyo en el proceso de toma de decisiones.

Esto se logró, gracias a la importancia que se le dio a los requerimientos de los usuarios y a la información proporcionada por las diversas fuentes de datos; ya que en base a ello se definieron los indicadores que permiten evaluar, de la mejor manera, el rendimiento académico de la población estudiantil y la labor que desempeña el Servicio de Orientación de la Facultad de Ciencias.

CONCLUSIONES Y RECOMENDACIONES

En cuanto al problema existente en la obtención de la información acerca del rendimiento académico de los alumnos y de la atención prestada por el Servicio de Orientación a la población estudiantil de la Facultad de Ciencias, se puede concluir que la solución de inteligencia de negocio planteada brinda apoyo al proceso de toma de decisiones de los usuarios ya que permite el acceso rápido y oportuno a la información, presentada mediante diversas perspectivas de manera tal que pueda ser analizada fácilmente.

Al realizar la planificación del proyecto de desarrollo de la solución de inteligencia de negocio para el Servicio de Orientación de la Facultad de Ciencias, se determinó que gran parte del tiempo estimado para la investigación fue utilizado en la recopilación de los datos necesarios para mostrar los indicadores requeridos por los usuarios.

En cuanto al objetivo específico relacionado con realizar el diseño lógico y el diseño físico del Área de Almacenamiento Intermedio y del Datamart, se determinó la importancia de estas tareas, debido a que afectan directamente el desempeño y tiempo de respuesta de la solución de inteligencia de negocio; así como también se concluyó que es necesario hacer uso de un motor de base de datos que permita el fácil acceso a la información guardada en los repositorios de datos que forman parte de ésta. Asimismo se utilizaron herramientas de la suite de edición comunitaria de Pentaho para la elaboración y ejecución periódica de los procesos de Extracción, Transformación y Carga de las bases de datos mencionadas anteriormente y para la elaboración de los cubos que forman parte de la solución de inteligencia de negocio de este caso de estudio.

En relación al objetivo específico de definir los indicadores a través de los cuales la solución de inteligencia de negocio brinde la posibilidad de supervisar el estado de los procesos de negocio asociados a este caso de estudio, se contó con la participación activa del personal que labora en el Servicio de Orientación de la Facultad de Ciencias, lo cual permitió satisfacer los requerimientos de los usuarios.

A través del uso de la solución de inteligencia de negocio desarrollada, se logró facilitar el acceso a la información relacionada con las actividades que realiza el Servicio de Orientación y la atención que presta a la población estudiantil, así como también se permitió analizar el rendimiento académico de los estudiantes de la Facultad de Ciencias, debido a que esta solución de inteligencia

de negocio recopila y procesa los datos para luego proporcionar reportes a los usuarios, en los cuales se muestra el estado de los indicadores asociados a este caso de estudio.

Para cumplir con el objetivo específico de apoyar el proceso de toma de decisiones relacionadas con la labor que desempeña el Servicio de Orientación y con el rendimiento académico de la población estudiantil de la Facultad de Ciencias, se puede agregar que, debido a la rapidez en la generación de reportes a través del uso de la solución de inteligencia de negocio desarrollada, los usuarios pueden dedicarle una mayor cantidad de tiempo al análisis de la información mostrada en los reportes y a la creación de estrategias que permitan brindar apoyo a los estudiantes que así lo requieran y darles las herramientas que les ayuden a desenvolverse exitosamente en el medio universitario.

Esta solución de inteligencia de negocio permitió realizar el proceso de recolección de datos y cálculo de indicadores lo cual garantiza la veracidad de los resultados obtenidos. Así como también posee un portal web que facilita a los usuarios el acceso a los reportes en los cuales se muestra una mayor cantidad de indicadores de la que se tenía, antes de la implementación de esta solución. Adicionalmente permite que estos se puedan visualizar desde diferentes perspectivas, lo que trae como consecuencia que los usuarios en lugar de recopilar y procesar los datos manualmente o mediante sistemas transaccionales, se dediquen a analizar el estado en el que se encuentran los indicadores y tomar las decisiones que se consideren pertinentes, según sea el caso.

Como conclusión final, se puede añadir que el desarrollo de una solución de inteligencia de negocio relacionada con el rendimiento académico de la población estudiantil y con la labor que desempeña el Servicio de Orientación permitió evaluar estos procesos de negocio, a través de la supervisión del estado de los indicadores seleccionados; así como también permite conocer lo importante que es la existencia de herramientas que ayuden a los usuarios durante el proceso de toma de decisiones.

En base a la investigación elaborada, se plantean las siguientes recomendaciones:

- Realizar actividades mediante las cuales se den a conocer las herramientas que proporciona la suite comunitaria de Pentaho y las aplicaciones de software libre con las que se llevó a cabo el proyecto; con el propósito de facilitar el conocimiento necesario a las personas que continuarán utilizando esta solución de inteligencia de negocio.
- Mantener actualizada la documentación de esta solución de inteligencia de negocio cada vez que existan cambios en ella.
- Se recomienda que, en todo momento, exista un usuario administrador de la solución de inteligencia de negocio desarrollada cuya función será dar asistencia a los usuarios y garantizar que el Datamart se mantenga actualizado con la información necesaria para evaluar el rendimiento académico de la población estudiantil y con la labor que desempeña el Servicio de Orientación de la Facultad de Ciencias, a través de reportes e indicadores; logrando de esta manera brindar apoyo a los usuarios en el proceso de toma de decisiones.

REFERENCIAS BIBLIOGRÁFICAS

Beltrán, J. (2006). *Indicadores de Gestión: Herramientas para lograr la competitividad* (Segunda ed.). Bogotá: 3R Editores.

Inmon, B. (2002). *Building the Data Warehouse* (Primera ed.). Estados Unidos: John Wiley & Sons, Inc.

Kendall, K. E. (2005). *Análisis y Diseño de Sistemas* (Sexta ed.). México: Pearson Educación.

Kimball, R. (2002). *The Data Warehouse Toolkit. The Complete Guide to Dimensional* (Segunda ed.). Estados Unidos: John Wiley & Sons, Inc.

Laudon, K. C. (2012). *Sistemas de Información Genencial* (Decimosegunda ed.). México: Pearson Educación.

Mondragón, A. (19 de Octubre de 2014). *¿Qué son los indicadores?* Obtenido de *¿Qué son los indicadores?*:<http://www.inegi.org.mx/inegi/contenidos/espanol/prensa/contenidos/articulos/economicas/indicadores.pdf>

Oracle, C. (12 de Agosto de 2014). *Oracle para las Medianas Empresas - Productos Oracle - Business Intelligence*. Obtenido de Portal de Tecnología y Soluciones de Oracle dirigido a las Medianas Empresas: <http://www.oracle.com/es/solutions/midsize/oracle-products/business-intelligence/index.html>

Pentaho, C. (12 de Agosto de 2014). *Pentaho Suite*. Obtenido de Pentaho Suite: <http://www.pentaho.com/announcement/powering-big-data-analytics-at-scale>

PostgreSQL. (26 de Febrero de 2015). *Sobre PostgreSQL*. Obtenido de Sobre PostgreSQL: http://www.postgresql.org/es/sobre_postgresql

Rodriguez, M., & López, T. (3 de Marzo de 2016). *Logro, Internalidad, Vocación. Desempeño Académico*. Obtenido de Logro, Internalidad, Vocación. Desempeño Académico.: https://www.google.co.ve/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjJ4_KUo6XLAhWIdh4KHcNKDGYQFggaMAA&url=http%3A%2F%2Fwww.sihca.org%2Freuniones%2Ftodas%2FOctavoforo%2FMesas_de_trabajo%2F3.%2520Investigacion%2F5.%2520UCLA%2520-%2520Maritza%2520Lo

Servicio de Orientación, F. d. (9 de Diciembre de 2015). *Servicio de Orientación | Facultad de Ciencias - UCV*. Obtenido de Servicio de Orientación: <http://www.ciens.ucv.ve/ciens/servicio-de-orientacion-2/>

ANEXOS

A. Diseño Físico del Área de Almacenamiento Intermedio

A continuación se muestra la especificación de las tablas del diseño físico del Área de Almacenamiento Intermedio.

Servicio de Orientación

Nombre de la Tabla: actividad			
Descripción: Tabla que contiene información de las actividades realizadas por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
tipo_atencion	100	character varying	Tipo de atención que se presta al estudiante durante la actividad
area_atencion	100	character varying	Área de atención relacionada con la actividad realizada
descripción	255	character varying	Descripción de la actividad
duración	4	integer	Duración de la actividad
Relaciones:		Campos Clave:	
		id	

Nombre de la Tabla: atención			
Descripción: Tabla que contiene información acerca de la atención prestada por el Servicio de Orientación a los estudiantes			
Campo	Tamaño	Tipo de Dato	Descripción
fecha	-	date	Fecha en la cual fue atendido el estudiante
actividad_id	10	character varying	Identificador de la actividad en la cual participó el estudiante
estudiante_cedula	4	integer	Cédula del estudiante atendido
estudiante_nombre	100	character varying	Nombre del estudiante
estudiante_escuela	100	character varying	Escuela a la cual pertenece el estudiante
estudiante_semestre_carrera	100	character varying	Semestre que cursa el estudiante
estudiante_anio_carrera	100	character varying	Año que cursa el estudiante
estudiante_genero	20	character varying	Género del estudiante
estudiante_fecha_nacimiento	-	date	Fecha de nacimiento del estudiante
estudiante_estado_civil	20	character varying	Estado civil del estudiante
estudiante_atendido_por	100	character varying	Lic. del Servicio de Orientación que brinda atención al estudiante.
Relaciones:		Campos Clave:	
actividad_id con actividad		fecha, actividad_id, estudiante_cedula	

Rendimiento Académico

Nombre de la Tabla: carrera			
Descripción: Tabla que contiene información acerca de la carrera que cursa el estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
nombre	100	character varying	Nombre de la carrera
nombre_corto	50	character varying	Nombre corto de la carrera
nombre_del_titulo	100	character varying	Nombre del título de la carrera
creditos_en_total	4	integer	Cantidad mínima de créditos necesarios para optar por el título
max_creditos_permitidos_por_periodo	4	integer	Cantidad máxima de créditos que se le permite inscribir al estudiante durante un periodo académico
Relaciones:		Campos Clave: id	

Nombre de la Tabla: estudiante			
Descripción: Tabla que contiene información detallada del estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
cedula	20	character varying	Cédula del estudiante
primer_nombre	100	character varying	Primer nombre del estudiante
segundo_nombre	100	character varying	Segundo nombre del estudiante
primer_apellido	100	character varying	Primer apellido del estudiante
segundo_apellido	100	character varying	Segundo apellido del estudiante
genero	20	character varying	Género del estudiante
fecha_nacimiento	-	date	Fecha de nacimiento del estudiante
Relaciones:		Campos Clave: cedula	

Nombre de la Tabla: token			
Descripción: Tabla que contiene información relacionada con la autorización necesaria para solicitar los datos académicos de la población estudiantil			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	integer	Identificador único del registro
token	255	character varying	Autorización
fecha_solicitud	-	timestamp	Fecha en la cual se solicitó la autorización
Relaciones:		Campos Clave: id	

Nombre de la Tabla: periodo_academico			
Descripción: Tabla que contiene información acerca de los periodos académicos transcurridos			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
anio	4	integer	Año en el cual transcurre el periodo académico
periodo	20	character varying	Periodo asociado con el registro
descripción	-	text	Descripción del periodo académico
fecha_inicio	-	date	Fecha en la cual comienza el periodo académico
fecha_fin	-	date	Fecha en la cual culmina el periodo académico
aplica_reglamento	-	boolean	Indica si se aplica el reglamento durante el periodo académico
Relaciones:		Campos Clave: id	

Nombre de la Tabla: materia			
Descripción: Tabla que contiene información detallada de las asignaturas existentes en el historial académico de los estudiantes			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
código	10	character varying	Código de la asignatura
sección	10	character varying	Sección en la cual el estudiante inscribió la asignatura
status	255	character varying	Estatus de la asignatura inscrita por el estudiante
tipo	255	character varying	Tipo de asignatura
nombre	255	character varying	Nombre de la asignatura
nro_creditos	4	integer	Cantidad de unidades de crédito asociada a la asignatura
Relaciones:		Campos Clave: id	

Nombre de la Tabla: docente			
Descripción: Tabla que contiene información del docente que dicta la asignatura			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	integer	Identificador único del registro
nombre	255	character varying	Nombre del docente
Relaciones:		Campos Clave: id	

Nombre de la Tabla: tipo_examen			
Descripción: Tabla que contiene información del tipo de examen asociado a la asignatura			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	integer	Identificador único del registro
nombre	255	character varying	Nombre del tipo de examen
Relaciones:		Campos Clave: id	

Nombre de la Tabla: historial_academico			
Descripción: Tabla que contiene información acerca del historial académico del estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
carrera_id	10	character varying	Identificador de la carrera asociada
estudiante_cedula	20	character varying	Cedula del estudiante asociada al historial
periodo_academico_id	10	character varying	Identificador del periodo académico
docente_id	4	Integer	Identificador del docente que dictó la asignatura
materia_id	10	character varying	Identificador de la materia inscrita por el estudiante
tipo_examen_id	4	Integer	Identificador del tipo de examen asociado a la materia
nota_en_final	10	character varying	Nota obtenida en el examen final
nota_en_reparacion	10	character varying	Nota obtenida en el examen de reparación
nota_definitiva	10	character varying	Nota definitiva asociada a la materia
Relaciones carrera_id con carrera periodo_academico_id con periodo_academico estudiante_cedula con estudiante docente_id con docente materia_id con materia tipo_examen_id con tipo_examen		Campos Clave: carrera_id, periodo_academico_id, estudiante_cedula, docente_id, materia_id, tipo_examen_id	

B. Diseño Físico del Datamart

A continuación se muestra la especificación de las tablas del diseño físico del Datamart.

Servicio de Orientación

Nombre de la Tabla: dim_so_estudiante			
Descripción: Dimensión que contiene información acerca de los estudiantes atendidos por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
cedula	20	character varying	Cédula del estudiante
nombre	100	character varying	Nombre del estudiante
fecha_nacimiento	-	date	Fecha de nacimiento del estudiante
Relaciones:		Campos Clave: cedula	

Nombre de la Tabla: dim_genero			
Descripción: Dimensión que contiene información acerca del género de los estudiantes atendidos por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	20	character varying	Género del estudiante atendido
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_estado_civil			
Descripción: Dimensión que contiene información acerca del estado civil de los estudiantes atendidos por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	20	character varying	Estado civil del estudiante atendido
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_semestre_carrera			
Descripción: Dimensión que contiene información acerca del semestre de la licenciatura que cursan de los estudiantes atendidos por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	100	character varying	Semestre de la licenciatura que cursa el estudiante atendido
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_anio_carrera			
Descripción: Dimensión que contiene información acerca del año de la licenciatura que cursan de los estudiantes atendidos por el Servicio de Orientación			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	100	character varying	Año de la licenciatura que cursa el estudiante atendido
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_atendido_por			
Descripción: Dimensión que contiene información acerca del personal del Servicio de Orientación que brinda atención al estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	255	character varying	Nombre del Lic. del Servicio de Orientación que atiende al estudiante
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_tipo_atencion			
Descripción: Dimensión que contiene información acerca del tipo de atención que el Servicio de Orientación brinda al estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	255	character varying	Nombre del tipo de atención que se brinda al estudiante
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_tiempo			
Descripción: Dimensión que contiene información acerca de la fecha en la cual el Servicio de Orientación atiende al estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
fecha	-	Date	Fecha en la cual el estudiante fue atendido
año	4	Integer	Año en el cual el estudiante fue atendido
mes	4	Integer	Mes en el cual el estudiante fue atendido
dia_del_año	4	Integer	Día del año en el cual el estudiante fue atendido
dia_del_mes	4	Integer	Día del mes en el cual el estudiante fue atendido
dia_de_la_semana	4	Integer	Día de la semana en el cual el estudiante fue atendido
semana_del_año	4	Integer	Semana del año en la cual el estudiante fue atendido
dia_de_la_semana_descripcion	20	character varying	Descripción del día de la semana en el cual el estudiante fue atendido
dia_de_la_semana_corto_descripcion	10	character varying	Descripción corta del día de la semana en el cual el estudiante fue atendido
mes_descripcion	20	character varying	Descripción del mes en el cual el estudiante fue atendido
mes_corto_descripcion	10	character varying	Descripción corta del mes en el cual el estudiante fue atendido
trimestre	4	Integer	Trimestre en el cual el estudiante fue atendido
periodo_academico	10	character varying	Período académico en el cual el estudiante fue atendido
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_actividad			
Descripción: Dimensión que contiene información acerca de las actividades que realiza el Servicio de Orientación con la finalidad de brindar atención al estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
descripción	255	character varying	Descripción de la actividad realizada
duración	4	Integer	Duración de la actividad realizada
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_escuela			
Descripción: Dimensión que contiene información acerca de la escuela a la cual pertenece el estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
universidad_id	4	Integer	Identificador de la universidad a la cual pertenece el estudiante
universidad_nombre	100	character varying	Nombre de la universidad a la cual pertenece el estudiante
facultad_id	4	Integer	Identificador de la facultad a la cual pertenece el estudiante
facultad_nombre	100	character varying	Nombre de la facultad a la cual pertenece el estudiante
escuela_id	4	Integer	Identificador de la escuela a la cual pertenece el estudiante
escuela_nombre	100	character varying	Nombre de la escuela a la cual pertenece el estudiante
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_area_atencion			
Descripción: Dimensión que contiene información acerca de las áreas de atención que el Servicio de Orientación toma en cuenta			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
nombre	255	character varying	Nombre del área de atención
Relaciones:		Campos Clave: id	

Nombre de la Tabla: fact_servicio_orientacion			
Descripción: Tabla de hechos que contiene información acerca de la atención que presta el Servicio de Orientación al estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
tiempo_id	4	Integer	Identificador de la fecha en la cual el estudiante es atendido
tipo_atencion_id	4	Integer	Identificador del tipo de atención que se presta al estudiante durante una actividad
area_atencion_id	4	Integer	Identificador del área de atención en la cual está enmarcada la actividad realizada
actividad_id	10	character varying	Identificador de la actividad en la cual participa el estudiante
so_estudiante_cedula	20	character varying	Cédula del estudiante atendido
escuela_id	4	Integer	Identificador de la escuela a la cual pertenece el estudiante
semestre_carrera_id	4	Integer	Identificador del semestre de la licenciatura que cursa el estudiante
anio_carrera_id	4	Integer	Identificador del año de la licenciatura que cursa el estudiante
genero_id	4	Integer	Identificador del género del estudiante atendido
estado_civil_id	4	Integer	Identificador del estado civil del estudiante atendido
atendido_por_id	4	Integer	Identificador del Lic. del Servicio de Orientación que brinda atención al estudiante
cantidad_estudiantes	4	Integer	Campo relacionado con la cantidad de estudiantes atendidos por el Servicio de Orientación.
Relaciones:		Campos Clave:	
tiempo_id con dim_tiempo		tiempo_id, tipo_atencion_id, area_atencion_id,	
tipo_atencion_id con dim_tipo_atencion		actividad_id, so_estudiante_cedula, escuela_id,	
area_atencion_id con dim_area_atencion		semestre_carrera_id, anio_carrera_id, genero_id,	
actividad_id con dim_actividad		estado_civil_id, atendido_por_id	
so_estudiante_cedula con dim_so_estudiante			
escuela_id con dim_escuela			
semestre_carrera_id con dim_semestre_carrera			
anio_carrera_id con dim_anio_carrera			
genero_id con dim_genero			
estado_civil_id con dim_estado_civil			
atendido_por_id con dim_atendido_por			

Rendimiento Académico

Nombre de la Tabla: dim_estudiante			
Descripción: Dimensión que contiene información detallada del estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
cedula	20	character varying	Cédula del estudiante
primer_nombre	100	character varying	Primer nombre del estudiante
segundo_nombre	100	character varying	Segundo nombre del estudiante
primer_apellido	100	character varying	Primer apellido del estudiante
segundo_apellido	100	character varying	Segundo apellido del estudiante
genero	20	character varying	Género del estudiante
fecha_nacimiento	-	date	Fecha de nacimiento del estudiante
Relaciones:		Campos Clave: cedula	

Nombre de la Tabla: dim_materia			
Descripción: Dimensión que contiene información detallada de las asignaturas inscritas por los estudiantes			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
código	10	character varying	Código de la asignatura
sección	10	character varying	Sección en la cual el estudiante inscribió la asignatura
status	255	character varying	Estatus de la asignatura inscrita por el estudiante
tipo	255	character varying	Tipo de asignatura
nombre	255	character varying	Nombre de la asignatura
nro_creditos	4	integer	Cantidad de unidades de crédito asociada a la asignatura
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_tipo_examen			
Descripción: Dimensión que contiene información del tipo de examen asociado a la asignatura			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	integer	Identificador único del registro
nombre	255	character varying	Nombre del tipo de examen
Relaciones:		Campos Clave: Id	

Nombre de la Tabla: dim_carrera			
Descripción: Dimensión que contiene información acerca de la carrera que cursa el estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	10	character varying	Identificador único del registro
nombre	100	character varying	Nombre de la carrera
nombre_corto	50	character varying	Nombre corto de la carrera
nombre_del_titulo	100	character varying	Nombre del título de la carrera
creditos_en_total	4	integer	Cantidad mínima de créditos necesarios para optar por el título
max_creditos_permitidos_por_periodo	4	integer	Cantidad máxima de créditos que se le permite inscribir al estudiante durante un periodo académico
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_periodo_academico			
Descripción: Dimensión que contiene información acerca del período académico en el cual el estudiante inscribe una asignatura			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	Integer	Identificador único del registro
fecha	-	Date	Fecha asociada al período académico
anio	4	Integer	Año asociado al período académico
mes	4	Integer	Mes asociado al período académico
dia_del_anio	4	Integer	Día del año asociado al período académico
dia_del_mes	4	Integer	Día del mes asociado al período académico
dia_de_la_semana	4	Integer	Día de la semana asociado al período académico
semana_del_anio	4	Integer	Semana del año asociada al período académico
dia_de_la_semana_descripcion	20	character varying	Descripción del día de la semana asociado al período académico
dia_de_la_semana_corto_descripcion	10	character varying	Descripción corta del día de la semana asociado al período académico
mes_descripcion	20	character varying	Descripción del mes asociado al período académico
mes_corto_descripcion	10	character varying	Descripción corta del mes asociado al período académico
trimester	4	Integer	Trimestre asociado al período académico
periodo_academico	10	character varying	Descripción del período académico
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_docente			
Descripción: Dimensión que contiene información del docente que dicta la asignatura			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	integer	Identificador único del registro
nombre	255	character varying	Nombre del docente
Relaciones:		Campos Clave: id	

Nombre de la Tabla: dim_nota			
Descripción: Dimensión que contiene información de la nota obtenida por el estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
id	4	integer	Identificador único del registro
descripción	255	character varying	Descripción de la nota
valor	4	integer	Valor numérico de la nota
Relaciones:		Campos Clave: id	

Nombre de la Tabla: fact_historial_academico			
Descripción: Tabla de hechos que contiene información acerca del historial académico del estudiante			
Campo	Tamaño	Tipo de Dato	Descripción
carrera_id	10	character varying	Identificador de la carrera asociada
estudiante_cedula	20	character varying	Cedula del estudiante asociada al historial
fecha_inicio_periodo_academico_id	10	character varying	Identificador de la fecha de inicio del periodo académico
fecha_fin_periodo_academico_id	10	character varying	Identificador de la fecha de fin del periodo académico
docente_id	4	Integer	Identificador del docente que dictó la asignatura
materia_id	10	character varying	Identificador de la materia inscrita por el estudiante
tipo_examen_id	4	Integer	Identificador del tipo de examen asociado a la materia
nota_en_final_id	4	Integer	Identificador de la nota asociada a la calificación obtenida en el examen final
nota_en_reparacion_id	4	Integer	Identificador de la nota asociada a la calificación obtenida en el examen de reparación
nota_definitiva_id	4	Integer	Identificador de la nota asociada a la calificación definitiva obtenida
uc_inscritas	4	Integer	Cantidad de unidades de crédito inscritas
materias_inscritas	4	Integer	Cantidad de materias inscritas
uc_aprobadas	4	Integer	Cantidad de unidades de crédito aprobadas
materias_aprobadas	4	Integer	Cantidad de materias aprobadas
uc_aprobadas_equivalencia	4	Integer	Cantidad de unidades de crédito aprobadas por equivalencia
materias_aprobadas_equivalencia	4	Integer	Cantidad de materias aprobadas por equivalencia
uc_aprobadas_general	4	Integer	Cantidad de unidades de crédito aprobadas en general
materias_aprobadas_general	4	Integer	Cantidad de materias aprobadas en general
uc_aplazadas	4	Integer	Cantidad de unidades de crédito aplazadas
materias_aplazadas	4	Integer	Cantidad de materias aplazadas
uc_retiradas	4	Integer	Cantidad de unidades de crédito retiradas
materias_retiradas	4	Integer	Cantidad de materias retiradas
uc_cursadas	4	Integer	Cantidad de unidades de crédito cursadas
materias_cursadas	4	Integer	Cantidad de materias cursadas
Relaciones		Campos Clave:	
carrera_id con dim_carrera		carrera_id, fecha_inicio_periodo_academico_id,	
fecha_inicio_periodo_academico_id con dim_periodo_academico		fecha_fin_periodo_academico_id, estudiante_cedula, docente_id, materia_id,	
fecha_fin_periodo_academico_id con dim_periodo_academico		tipo_examen_id, nota_en_final_id, nota_en_reparacion_id, nota_definitiva_id	
estudiante_cedula con dim_estudiante			
docente_id con dim_docente			
materia_id con dim_materia			
tipo_examen_id con dim_tipo_examen			
nota_en_final_id con dim_nota			
nota_en_reparacion_id con dim_nota			
nota_definitiva_id con dim_nota			

C. Diseño de Procesos de Extracción, Transformación y Carga del Área de Almacenamiento Intermedio

A continuación se muestra el diseño de los procesos de Extracción, Transformación y Carga del Área de Almacenamiento Intermedio para los procesos de negocio “Rendimiento Académico” y “Servicio de Orientación”, sin tomar en cuenta el gráfico de éstos, debido a su nivel de complejidad; sin embargo se puede acceder a ellos utilizando Pentaho Data Integration.

Servicio de Orientación

Se diseñaron dos procesos de Extracción, Transformación y Carga con la finalidad de poblar la sección del Área de Almacenamiento Intermedio relacionada con el proceso de negocio llamado “Servicio de Orientación”.

Nombre del Proceso: Poblar la Tabla actividad

Descripción: Representa la carga de la tabla actividad

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Libro de Calculo	Instrumento Diagnóstico.xls	Contiene información acerca de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	character varying(10)	Texto	No nulo	No tiene
tipo_atencion		character varying(100)	Texto		No tiene
area_atencion		character varying(100)	Texto		No tiene
descripción		character varying(255)	Texto		No tiene
duración		Integer	Numérico	No negativo	No tiene

Fuentes de Datos

Hoja: Atención Individual				
Nombre	Llave	Tipo	Formato	Consideración Importante
Fecha		date	Fecha	
Nombre y Apellido		character varying(100)	Texto	
Semestre		Integer	Numérico	
Licenciatura		character varying(20)	Texto	
Cédula		character varying(20)	Texto	
Género		character varying(20)	Texto	
Edad		Integer	Numérico	
Estado Civil		character varying(20)	Texto	
Área de Atención		character varying(20)	Texto	
Actividad		character varying(100)	Texto	
Atendido por		character varying(100)	Texto	

Hoja: Detalle de Actividades Grupales				
Nombre	Llave	Tipo	Formato	Consideración Importante
Código		Integer	Numérico	
Área de Atención		character varying(20)	Texto	
Actividad		character varying(100)	Texto	
Duración		character varying(20)	Texto	

Tabla Destino

Tabla: actividad		
Campo	Tipo	Correspondencia
id	character varying(10)	
tipo_atencion	character varying(100)	
area_atencion	character varying(100)	
descripción	character varying(255)	
duración	Integer	

Proceso realizado:

- Extraer registros del Libro de Cálculo Instrumento Diagnóstico.xls
- Cambiar el contenido de los campos de minúscula a mayúscula
- Quitar los acentos de los campos
- Determinar el tipo de atención de cada campo
- Eliminar registros repetidos
- Determinar el código único de cada registro
- Cargar los registros en la tabla actividad

Nombre del Proceso: Poblar la Tabla atencion

Descripción: Representa la carga de la tabla atencion

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Libro de Calculo	Instrumento Diagnóstico.xls	Contiene información acerca de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
fecha	PK	date	Fecha	No nulo	No tiene
actividad_id	PK, FK	character varying(10)	Texto	No nulo	No tiene
estudiante_cedula	PK	Integer	Numérico	No nulo	No tiene
estudiante_nombre		character varying(100)	Texto		No tiene
estudiante_escuela		character varying(100)	Texto		No tiene
estudiante_semestre_carrera		character varying(100)	Texto		No tiene
estudiante_anio_carrera		character varying(100)	Texto		No tiene
estudiante_genero		character varying(20)	Texto		No tiene
estudiante_fecha_nacimiento		date	Fecha		No tiene
estudiante_estado_civil		character varying(20)	Texto		No tiene
estudiante_atendido_por		character varying(100)	Texto		No tiene

Fuentes de Datos

Hoja: Atención Individual				
Nombre	Llave	Tipo	Formato	Consideración Importante
Fecha		date	Fecha	
Nombre y Apellido		character varying(100)	Texto	
Semestre		Integer	Numérico	
Licenciatura		character varying(20)	Texto	
Cédula		character varying(20)	Texto	
Género		character varying(20)	Texto	
Edad		Integer	Numérico	
Estado Civil		character varying(20)	Texto	
Área de Atención		character varying(20)	Texto	
Actividad		character varying(100)	Texto	
Atendido por		character varying(100)	Texto	

Hoja: Atención Grupal				
Nombre	Llave	Tipo	Formato	Consideración Importante
Fecha		date	Fecha	
Área de Atención		character varying(20)	Texto	
Actividad		character varying(100)	Texto	
Código		Integer	Numérico	
Cédula		character varying(20)	Texto	
Nombre y Apellido		character varying(100)	Texto	
Licenciatura		character varying(20)	Texto	
Semestre		Integer	Numérico	
Género		character varying(20)	Texto	
Edad		Integer	Numérico	
Estado Civil		character varying(20)	Texto	
Atendido por		character varying(100)	Texto	

Tabla Destino

Tabla: atención		
Campo	Tipo	Correspondencia
fecha	date	
actividad_id	character varying(10)	
estudiante_cedula	Integer	
estudiante_nombre	character varying(100)	
estudiante_escuela	character varying(100)	
estudiante_semestre_carrera	character varying(100)	
estudiante_anio_carrera	character varying(100)	
estudiante_genero	character varying(20)	
estudiante_fecha_nacimiento	date	
estudiante_estado_civil	character varying(20)	
estudiante_atendido_por	character varying(100)	

Proceso realizado:

- Extraer registros del Libro de Cálculo Instrumento Diagnóstico.xls
- Cambiar el contenido de los campos de minúscula a mayúscula
- Quitar los acentos de los campos
- Determinar el tipo de atención de cada campo
- Eliminar los espacios en blanco de los campos
- Determinar el código único de cada registro
- Determinar la actividad en la que participó el estudiante
- Determinar el semestre de la licenciatura que cursa el estudiante
- Determinar la escuela a la que pertenece el estudiante
- Determinar el estado civil del estudiante
- Ordenar todos los registros por la actividad que realizó el estudiante y la fecha en la cual fue realizada
- Eliminar los registros repetidos
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla atención

Rendimiento Académico

Se diseñaron nueve procesos de Extracción, Transformación y Carga con el fin de poblar la sección del Área de Almacenamiento Intermedio relacionada con el proceso de negocio “Rendimiento Académico”.

Nombre del Proceso: Solicitud de Autorización

Descripción: Representa la carga de la tabla token

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	token.json	Contiene información acerca de la autorización necesaria para solicitar los datos académicos de la población estudiantil

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	integer	Numérico	No nulo	No tiene
token		character varying(255)	Texto		No tiene
fecha_solicitud		timestamp	Fecha		Fecha actual

Fuentes de Datos

Archivo JSON: token				
Nombre	Llave	Tipo	Formato	Consideración Importante
token		character varying(255)	Texto	

Tabla Destino

Tabla: token		
Campo	Tipo	Correspondencia
id	integer	
token	character varying(255)	token.token
fecha_solicitud	timestamp	

Proceso realizado:

- Solicitar autorización para acceder a los datos académicos de la población estudiantil
- Extraer registros del archivo token.json
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran

- Asignar el identificador único de cada registro
- Cargar los registros en la tabla token con la información correspondiente

Nombre del Proceso: Poblar la Tabla carrera

Descripción: Representa la carga de la tabla carrera

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	carreras.json	Contiene información acerca de las diferentes licenciaturas que puede cursar un estudiante

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	character varying(10)	Texto	No nulo	No tiene
nombre		character varying(100)	Texto		No tiene
nombre_corto		character varying(50)	Texto		No tiene
nombre_del_titulo		character varying(100)	Texto		No tiene
creditos_en_total		Integer	Numérico	No negativo	No tiene
max_creditos_permitidos_por_periodo		Integer	Numérico	No negativo	No tiene

Fuentes de Datos

Archivo JSON: carreras				
Nombre	Llave	Tipo	Formato	Consideración Importante
id		character varying(10)	Texto	
nombre		character varying(100)	Texto	
nombre_corto		character varying(50)	Texto	
nombre_del_titulo		character varying(100)	Texto	
creditos_en_total		Integer	Numérico	
max_creditos_por_periodo		Integer	Numérico	

Tabla Destino

Tabla: carrera		
Campo	Tipo	Correspondencia
id	character varying(10)	carreras.id
nombre	character varying(100)	carreras.nombre
nombre_corto	character varying(50)	carreras.nombre_corto
nombre_del_titulo	character varying(100)	carreras.nombre_del_titulo
creditos_en_total	Integer	carreras.creditos_en_total
max_creditos_permitidos_por_periodo	Integer	carreras.max_creditos_por_periodo

Proceso realizado:

- Utilizar autorización para solicitar el archivo carreras.json
- Extraer registros del archivo carreras.json para cada licenciatura
- Ordenar los registros
- Eliminar los registros repetidos
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla carrera con la información correspondiente

Nombre del Proceso: Poblar la Tabla estudiante

Descripción: Representa la carga de la tabla estudiante

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	estudiantes.json	Contiene información acerca de la población estudiantil existente

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
cedula	PK	character varying(20)	Texto	No nulo	No tiene
primer_nombre		character varying(100)	Texto		No tiene
segundo_nombre		character varying(100)	Texto		No tiene
primer_apellido		character varying(100)	Texto		No tiene
segundo_apellido		character varying(100)	Texto		No tiene
genero		character varying(20)	Texto		No tiene
fecha_nacimiento		date	Fecha		No tiene

Fuentes de Datos

Archivo JSON: estudiantes				
Nombre	Llave	Tipo	Formato	Consideración Importante
cedula		character varying(20)	Texto	
primer_nombre		character varying(100)	Texto	
segundo_nombre		character varying(100)	Texto	
primer_apellido		character varying(100)	Texto	
segundo_apellido		character varying(100)	Texto	
genero		character varying(20)	Texto	

fecha_nacimiento		date	Fecha	
------------------	--	------	-------	--

Tabla Destino

Tabla: estudiante		
Campo	Tipo	Correspondencia
cedula	character varying(20)	estudiantes..cedula
primer_nombre	character varying(100)	estudiantes..primer_nombre
segundo_nombre	character varying(100)	estudiantes..segundo_nombre
primer_apellido	character varying(100)	estudiantes..primer_apellido
segundo_apellido	character varying(100)	estudiantes..segundo_apellido
genero	character varying(20)	estudiantes..genero
fecha_nacimiento	date	estudiantes..fecha_nacimiento

Proceso realizado:

- Utilizar autorización para solicitar el archivo estudiantes.json
- Extraer registros del archivo estudiantes.json para cada licenciatura
- Ordenar los registros
- Eliminar los registros repetidos
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla estudiante con la información correspondiente

Nombre del Proceso: Poblar la Tabla materia

Descripción: Representa la carga de la tabla materia

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	historial.json	Contiene información acerca del historial académico del estudiante

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
código		character varying(10)	Texto		No tiene
sección		character varying(10)	Texto		No tiene
status		character varying(255)	Texto		No tiene
tipo		character varying(255)	Texto		No tiene
nombre		character varying(255)	Texto		No tiene
nro_creditos		Integer	Numérico	No negativo	No tiene

Fuentes de Datos

Archivo JSON: historial				
Nombre	Llave	Tipo	Formato	Consideración Importante
periodo_academico		character varying(10)	Texto	
nombre		character varying(255)	Texto	
código		character varying(10)	Texto	
sección		character varying(10)	Texto	
créditos		Integer	Numérico	
tipo_materia		character varying(255)	Texto	
docente_nombre		character varying(255)	Texto	
tipo_status_materia		character varying(255)	Texto	
tipo_examen		character varying(255)	Texto	
nota_en_final		character varying(10)	Texto	
nota_en_reparacion		character varying(10)	Texto	
nota_definitiva		character varying(10)	Texto	

Tabla Destino

Tabla: materia		
Campo	Tipo	Correspondencia
id	Integer	
código	character varying(10)	historial.codigo
sección	character varying(10)	historial.seccion
status	character varying(255)	historial.tipo_status_materia
tipo	character varying(255)	historial.tipo_materia
nombre	character varying(255)	historial.nombre
nro_creditos	Integer	historial.creditos

Proceso realizado:

- Utilizar autorización para solicitar el archivo historial.json
- Extraer registros del archivo historial.json para cada estudiante
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla materia con la información correspondiente

Nombre del Proceso: Poblar la Tabla tipo_examen

Descripción: Representa la carga de la tabla tipo_examen

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	historial.json	Contiene información acerca del historial académico del estudiante

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(255)	Texto		No tiene

Fuentes de Datos

Archivo JSON: historial				
Nombre	Llave	Tipo	Formato	Consideración Importante
periodo_academico		character varying(10)	Texto	
nombre		character varying(255)	Texto	
código		character varying(10)	Texto	
sección		character varying(10)	Texto	
créditos		Integer	Numérico	
tipo_materia		character varying(255)	Texto	
docente_nombre		character varying(255)	Texto	
tipo_status_materia		character varying(255)	Texto	
tipo_examen		character varying(255)	Texto	
nota_en_final		character varying(10)	Texto	
nota_en_reparacion		character varying(10)	Texto	
nota_definitiva		character varying(10)	Texto	

Tabla Destino

Tabla: tipo_examen		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(255)	historial.tipo_examen

Proceso realizado:

- Utilizar autorización para solicitar el archivo historial.json
- Extraer registros del archivo historial.json para cada estudiante
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato y el tipo de dato de cada uno de ellos.
- Cargar los registros en la tabla tipo_examen con la información correspondiente

Nombre del Proceso: Poblar la Tabla docente

Descripción: Representa la carga de la tabla docente

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	historial.json	Contiene información acerca del historial académico del estudiante

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(255)	Texto		No tiene

Fuentes de Datos

Archivo JSON: historial				
Nombre	Llave	Tipo	Formato	Consideración Importante
periodo_academico		character varying(10)	Texto	
nombre		character varying(255)	Texto	
código		character varying(10)	Texto	
sección		character varying(10)	Texto	
créditos		Integer	Numérico	
tipo_materia		character varying(255)	Texto	
docente_nombre		character varying(255)	Texto	
tipo_status_materia		character varying(255)	Texto	
tipo_examen		character varying(255)	Texto	
nota_en_final		character varying(10)	Texto	
nota_en_reparacion		character varying(10)	Texto	
nota_definitiva		character varying(10)	Texto	

Tabla Destino

Tabla: docente		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(255)	historial.docente_nombre

Proceso realizado:

- Utilizar autorización para solicitar el archivo historial.json
- Extraer registros del archivo historial.json para cada estudiante
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla docente con la información correspondiente

Nombre del Proceso: Poblar la Tabla historial_academico

Descripción: Representa la carga de la tabla historial_academico

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	historial.json	Contiene información acerca del historial académico del estudiante

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
carrera_id	PK, FK	character varying(10)	Texto	No nulo	No tiene
estudiante_cedula	PK, FK	character varying(20)	Texto	No nulo	No tiene
periodo_academico_id	PK, FK	character varying(10)	Texto	No nulo	No tiene
docente_id	PK, FK	Integer	Numérico	No nulo	No tiene
materia_id	PK, FK	Integer	Numérico	No nulo	No tiene
tipo_examen_id	PK, FK	Integer	Numérico	No nulo	No tiene
nota_en_final		character varying(10)	Texto		No tiene
nota_en_reparacion		character varying(10)	Texto		No tiene
nota_definitiva		character varying(10)	Texto		No tiene

Fuentes de Datos

Archivo JSON: historial				
Nombre	Llave	Tipo	Formato	Consideración Importante
periodo_academico		character varying(10)	Texto	
nombre		character varying(255)	Texto	
código		character varying(10)	Texto	
sección		character varying(10)	Texto	
créditos		Integer	Numérico	
tipo_materia		character varying(255)	Texto	
docente_nombre		character varying(255)	Texto	
tipo_status_materia		character varying(255)	Texto	
tipo_examen		character varying(255)	Texto	
nota_en_final		character varying(10)	Texto	
nota_en_reparacion		character varying(10)	Texto	
nota_definitiva		character varying(10)	Texto	

Tabla Destino

Tabla: historial_academico		
Campo	Tipo	Correspondencia
carrera_id	character varying(10)	
estudiante_cedula	character varying(20)	
periodo_academico_id	character varying(10)	
docente_id	Integer	
materia_id	Integer	
tipo_examen_id	Integer	
nota_en_final	character varying(10)	historial..nota_en_final
nota_en_reparacion	character varying(10)	historial..nota_en_reparacion
nota_definitiva	character varying(10)	historial..nota_definitiva

Proceso realizado:

- Utilizar autorización para solicitar el archivo historial.json
- Extraer registros del archivo historial.json para cada estudiante
- Determinar la licenciatura que cursa el estudiante
- Eliminar los espacios en blanco de los campos
- Ordenar los registros
- Eliminar los registros repetidos
- Encontrar las claves foráneas para cada registro en los campos donde lo amerite

- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla historial_academico con la información correspondiente

Nombre del Proceso: Solicitud de Cierre de Sesión

Descripción: Representa la eliminación de los registros de la tabla token

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Archivo JSON	token.json	Contiene información acerca de la autorización necesaria para acceder a los datos académicos de la población estudiantil

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	integer	Numérico	No nulo	No tiene
token		character varying(255)	Texto		No tiene
fecha_solicitud		timestamp	Fecha		Fecha actual

Fuentes de Datos

Archivo JSON: token				
Nombre	Llave	Tipo	Formato	Consideración Importante
token		character varying(255)	Texto	

Tabla Destino

Tabla: token		
Campo	Tipo	Correspondencia
id	integer	
token	character varying(255)	
fecha_solicitud	timestamp	

Proceso realizado:

- Utilizar autorización para solicitar el cierre de la sesión, proceso mediante el cual la autorización pierde validez.
- Eliminar registro de la tabla token

D. Diseño de Procesos de Extracción, Transformación y Carga del Datamart

A continuación se muestra el diseño de los procesos de Extracción, Transformación y Carga del Datamart para los procesos de negocio “Rendimiento Académico” y “Servicio de Orientación”, sin tomar en cuenta el gráfico de éstos debido a su nivel de complejidad; sin embargo se puede acceder a ellos utilizando Pentaho Data Integration.

Servicio de Orientación

Se diseñaron doce procesos de Extracción, Transformación y Carga con la finalidad de poblar la sección del Datamart relacionada con el proceso de negocio llamado “Servicio de Orientación”.

Nombre del Proceso: Poblar la Dimensión Tiempo

Descripción: Representa la carga de la tabla dim_tiempo

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
fecha		date	Fecha		No tiene
anio		Integer	Numérico	No negativo	No tiene
mes		Integer	Numérico	No negativo	No tiene
dia_del_anio		Integer	Numérico	No negativo	No tiene
dia_del_mes		Integer	Numérico	No negativo	No tiene
dia_de_la_semana		Integer	Numérico	No negativo	No tiene
semana_del_anio		Integer	Numérico	No negativo	No tiene
dia_de_la_semana_descripcion		character varying(20)	Texto		No tiene
dia_de_la_semana_corto_descripcion		character varying(10)	Texto		No tiene
mes_descripcion		character varying(20)	Texto		No tiene
mes_corto_descripcion		character varying(10)	Texto		No tiene
trimestre		Integer	Numérico	No negativo	No tiene
periodo_academico		character varying(10)	Texto		No tiene

Fuentes de Datos

Tabla: periodo_academico				
Nombre	Llave	Tipo	Formato	Consideración Importante
id		character varying(10)	Texto	
anio		Integer	Numérico	
periodo		character varying(20)	Texto	
descripción		character varying(255)	Texto	
fecha_inicio		date	Fecha	
fecha_fin		date	Fecha	
aplica_reglamento		boolean	Lógico	

Tabla Destino

Tabla: dim_tiempo		
Campo	Tipo	Correspondencia
id	Integer	
fecha	date	
anio	Integer	
mes	Integer	
dia_del_anio	Integer	
dia_del_mes	Integer	
dia_de_la_semana	Integer	
semana_del_anio	Integer	
dia_de_la_semana_descripcion	character varying(20)	
dia_de_la_semana_corto_descripcion	character varying(10)	
mes_descripcion	character varying(20)	
mes_corto_descripcion	character varying(10)	
trimestre	Integer	
periodo_academico	character varying(10)	periodo_academico.id

Proceso realizada:

- Generar fechas
- Generar identificador de cada registro
- Determinar el trimestre, el día de la semana y el mes correspondiente a cada registro
- Ordenar registros por fecha
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_tiempo con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Tipo de Atención

Descripción: Representa la carga de la tabla dim_tipo_atencion

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(255)	Texto		No tiene

Fuentes de Datos

Tabla: actividad				
Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	character varying(10)	Texto	
tipo_atencion		character varying(100)	Texto	
area_atencion		character varying(100)	Texto	
descripción		character varying(255)	Texto	
duración		Integer	Numérico	

Tabla Destino

Tabla: dim_tipo_atencion		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(255)	actividad.tipo_atencion

Proceso realizado:

- Extraer registros de la tabla actividad para cada tipo de atención prestada
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_tipo_atencion con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Área de Atención

Descripción: Representa la carga de la tabla dim_area_atencion

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(255)	Texto		No tiene

Fuentes de Datos

Tabla: actividad				
Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	character varying(10)	Texto	
tipo_atencion		character varying(100)	Texto	
area_atencion		character varying(100)	Texto	
descripción		character varying(255)	Texto	
duración		Integer	Numérico	

Tabla Destino

Tabla: dim_area_atencion		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(255)	actividad.area_atencion

Proceso realizado:

- Extraer registros de la tabla actividad para cada área de atención prestada
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran

- Cargar los registros en la tabla dim_area_atencion con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Actividad

Descripción: Representa la carga de la tabla dim_actividad

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	character varying(10)	Texto	No nulo	No tiene
descripción		character varying(255)	Texto		No tiene
duración		Integer	Numérico		No tiene

Fuentes de Datos

Tabla: actividad

Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	character varying(10)	Texto	
tipo_atencion		character varying(100)	Texto	
area_atencion		character varying(100)	Texto	
descripción		character varying(255)	Texto	
duración		Integer	Numérico	

Tabla Destino

Tabla: dim_actividad

Campo	Tipo	Correspondencia
id	character varying(10)	
descripción	character varying(255)	actividad.descripción
duración	Integer	actividad.duración

Proceso realizado:

- Extraer registros de la tabla actividad de cada actividad realizada
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran

- Cargar los registros en la tabla dim_actividad con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Escuela

Descripción: Representa la carga de la tabla dim_escuela

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Libro de Calculo	escuelas.xls	Contiene información acerca de las escuelas asociadas a la población estudiantil

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
universidad_id		Integer	Numérico	No negativo	No tiene
universidad_nombre		character varying(100)	Texto		No tiene
facultad_id		Integer	Numérico	No negativo	No tiene
facultad_nombre		character varying(100)	Texto		No tiene
escuela_id		Integer	Numérico	No negativo	No tiene
escuela_nombre		character varying(100)	Texto		No tiene

Fuentes de Datos

Hoja: escuelas				
Nombre	Llave	Tipo	Formato	Consideración Importante
Escuela		character varying(100)	Texto	
Facultad		character varying(100)	Texto	
Universidad		character varying(100)	Texto	

Tabla Destino

Tabla: dim_escuela		
Campo	Tipo	Correspondencia
id	Integer	
universidad_id	Integer	
universidad_nombre	character varying(100)	escuelas.Universidad
facultad_id	Integer	
facultad_nombre	character varying(100)	escuelas.Facultad
escuela_id	Integer	
escuela_nombre	character varying(100)	escuelas.Escuela

Proceso realizado:

- Extraer registros del Libro de Calculo escuelas.xls
- Ordenar los registros

- Eliminar los registros repetidos
- Generar identificador asociado a cada campo de la fuente de datos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_escuela con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Atendido Por

Descripción: Representa la carga de la tabla dim_atendido_por

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(100)	Texto		No tiene

Fuentes de Datos

Tabla: atención				
Nombre	Llave	Tipo	Formato	Consideración Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	
estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: dim_atendido_por		
Campo	Tipo	Correspondencia

id	Integer	
nombre	character varying(100)	atencion.estudiante_atendido_por

Proceso realizado:

- Extraer registros de la tabla atención de cada individuo que forma parte del personal del Servicio de Orientación
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único para cada registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_atendido_por con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Semestre de Licenciatura

Descripción: Representa la carga de la tabla dim_semestre_carrera

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(100)	Texto		No tiene

Fuentes de Datos

Tabla: atención				
Nombre	Llave	Tipo	Formato	Consideración Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	

estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: dim_semestre_carrera		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(100)	atencion.estudiante_semestre_carrera

Proceso realizado:

- Extraer registros de la tabla atención de los semestres asociados a cada licenciatura
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único para cada registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_semestre_carrera con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Año de Licenciatura

Descripción: Representa la carga de la tabla dim_anio_carrera

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(100)	Texto		No tiene

Fuentes de Datos

Tabla: atención				
Nombre	Llave	Tipo	Formato	Consideración Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	
estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: dim_anio_carrera		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(100)	atencion.estudiante_anio_carrera

Proceso realizado:

- Extraer registros de la tabla atención de los años asociados a cada licenciatura
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único para cada registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_anio_carrera con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Género

Descripción: Representa la carga de la tabla dim_genero

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(20)	Texto		No tiene

Fuentes de Datos

Tabla: atención				
Nombre	Llave	Tipo	Formato	Consideración Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	
estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: dim_genero		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(20)	atencion.estudiante_genero

Proceso realizado:

- Extraer registros de la tabla atención relacionados con el género de los estudiantes atendidos por el Servicio de Orientación
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único para cada registro

- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_genero con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Estado Civil

Descripción: Representa la carga de la tabla dim_estado_civil

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(20)	Texto		No tiene

Fuentes de Datos

Tabla: atención				
Nombre	Llave	Tipo	Formato	Consideración Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	
estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: dim_estado_civil		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(20)	atencion.estudiante_estado_civil

Proceso realizado:

- Extraer registros de la tabla atención relacionados con el estado civil de los estudiantes atendidos por el Servicio de Orientación
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único para cada registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_estado_civil con la información correspondiente

Nombre del Proceso: Poblar la Tabla de Hechos Servicio de Orientacion

Descripción: Representa la carga de la tabla fact_servicio_orientacion

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
tiempo_id	PK, FK	Integer	Numérico	No nulo	No tiene
tipo_atencion_id	PK, FK	Integer	Numérico	No nulo	No tiene
area_atencion_id	PK, FK	Integer	Numérico	No nulo	No tiene
actividad_id	PK, FK	character varying(10)	Texto	No nulo	No tiene
so_estudiante_cedula	PK, FK	character varying(20)	Texto	No nulo	No tiene
escuela_id	PK, FK	Integer	Numérico	No nulo	No tiene
semestre_carrera_id	PK, FK	Integer	Numérico	No nulo	No tiene
anio_carrera_id	PK, FK	Integer	Numérico	No nulo	No tiene
genero_id	PK, FK	Integer	Numérico	No nulo	No tiene
estado_civil_id	PK, FK	Integer	Numérico	No nulo	No tiene
atendido_por_id	PK, FK	Integer	Numérico	No nulo	No tiene
cantidad_estudiantes		Integer	Numérico		No tiene

Fuentes de Datos

Tabla: atención					
Nombre	Llave	Tipo	Formato	Consideración	

				Importante
fecha	PK	date	Fecha	
actividad_id	PK, FK	character varying(10)	Texto	
estudiante_cedula	PK	Integer	Numérico	
estudiante_nombre		character varying(100)	Texto	
estudiante_escuela		character varying(100)	Texto	
estudiante_semestre_carrera		character varying(100)	Texto	
estudiante_anio_carrera		character varying(100)	Texto	
estudiante_genero		character varying(20)	Texto	
estudiante_fecha_nacimiento		date	Fecha	
estudiante_estado_civil		character varying(20)	Texto	
estudiante_atendido_por		character varying(100)	Texto	

Tabla Destino

Tabla: fact_servicio_orientacion		
Campo	Tipo	Correspondencia
tiempo_id	Integer	
tipo_atencion_id	Integer	
area_atencion_id	Integer	
actividad_id	character varying(10)	
so_estudiante_cedula	character varying(20)	
escuela_id	Integer	
semestre_carrera_id	Integer	
anio_carrera_id	Integer	
genero_id	Integer	
estado_civil_id	Integer	
atendido_por_id	Integer	
cantidad_estudiantes	Integer	

Proceso realizado:

- Extraer registros de la tabla atención relacionados con la labor que desempeña el Servicio de Orientación
- Ordenar los registros
- Eliminar los registros repetidos
- Encontrar la clave foránea correspondiente en cada campo que lo necesite para cada registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla fact_servicio_orientacion con la información correspondiente

Rendimiento Académico

Se diseñaron ocho procesos de Extracción, Transformación y Carga con la finalidad de poblar la sección del Datamart relacionada con el proceso de negocio llamado “Rendimiento Académico”.

Nombre del Proceso: Poblar la Dimensión Periodo Académico

Descripción: Representa la carga de la tabla dim_periodo_academico

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
fecha		date	Fecha		No tiene
anio		Integer	Numérico	No negativo	No tiene
mes		Integer	Numérico	No negativo	No tiene
dia_del_anio		Integer	Numérico	No negativo	No tiene
dia_del_mes		Integer	Numérico	No negativo	No tiene
dia_de_la_semana		Integer	Numérico	No negativo	No tiene
semana_del_anio		Integer	Numérico	No negativo	No tiene
dia_de_la_semana_descripcion		character varying(20)	Texto		No tiene
dia_de_la_semana_corto_descripcion		character varying(10)	Texto		No tiene
mes_descripcion		character varying(20)	Texto		No tiene

mes_corto_descripcion		character varying(10)	Texto		No tiene
trimestre		Integer	Numérico	No negativo	No tiene
periodo_academico		character varying(10)	Texto		No tiene

Fuentes de Datos

Tabla: periodo_academico				
Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	character varying(10)	Texto	
anio		Integer	Numérico	
periodo		character varying(20)	Texto	
descripción		character varying(255)	Texto	
fecha_inicio		date	Fecha	
fecha_fin		date	Fecha	
aplica_reglamento		boolean	Lógico	

Tabla Destino

Tabla: dim_periodo_academico		
Campo	Tipo	Correspondencia
id	Integer	
fecha	date	
anio	Integer	
mes	Integer	
dia_del_anio	Integer	
dia_del_mes	Integer	
dia_de_la_semana	Integer	
semana_del_anio	Integer	
dia_de_la_semana_descripcion	character varying(20)	
dia_de_la_semana_corto_descripcion	character varying(10)	
mes_descripcion	character varying(20)	
mes_corto_descripcion	character varying(10)	
trimestre	Integer	
periodo_academico	character varying(10)	periodo_academico.id

Proceso realizado:

- Generar fechas
- Generar identificador de cada registro
- Determinar el trimestre correspondiente a cada registro
- Determinar el día de la semana correspondiente a cada registro
- Determinar el mes correspondiente a cada registro
- Ordenar registros por fecha
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran

- Cargar los registros en la tabla dim_periodo_academico con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Carrera

Descripción: Representa la carga de la tabla dim_carrera

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	character varying(10)	Texto	No nulo	No tiene
nombre		character varying(100)	Texto		No tiene
nombre_corto		character varying(50)	Texto		No tiene
nombre_del_titulo		character varying(100)	Texto		No tiene
creditos_en_total		Integer	Numérico	No negativo	No tiene
max_creditos_permitidos_por_periodo		Integer	Numérico	No negativo	No tiene

Fuentes de Datos

Tabla: carrera				
Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	character varying(10)	Texto	
nombre		character varying(100)	Texto	
nombre_corto		character varying(50)	Texto	
nombre_del_titulo		character varying(100)	Texto	
creditos_en_total		Integer	Numérico	
max_creditos_permitidos_por_periodo		Integer	Numérico	

Tabla Destino

Tabla: dim_carrera		
Campo	Tipo	Correspondencia
id	character varying(10)	carrera.id
nombre	character varying(100)	carrera.nombre

nombre_corto	character varying(50)	carrera.nombre_corto
nombre_del_titulo	character varying(100)	carrera.nombre_del_titulo
creditos_en_total	Integer	carrera.creditos_en_total
max_creditos_permitidos_por_periodo	Integer	carrera.max_creditos_permitidos_por_periodo

Proceso

- Extraer registros de la tabla carrera
- Ordenar los registros
- Eliminar los registros repetidos
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_carrera con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Estudiante

Descripción: Representa la carga de la tabla dim_estudiante

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
cedula	PK	character varying(20)	Texto	No nulo	No tiene
primer_nombre		character varying(100)	Texto		No tiene
segundo_nombre		character varying(100)	Texto		No tiene
primer_apellido		character varying(100)	Texto		No tiene
segundo_apellido		character varying(100)	Texto		No tiene
genero		character varying(20)	Texto		No tiene
fecha_nacimiento		date	Fecha		No tiene

Fuentes de Datos

Tabla: estudiante				
Nombre	Llave	Tipo	Formato	Consideración Importante
cedula	PK	character varying(20)	Texto	
primer_nombre		character varying(100)	Texto	
segundo_nombre		character varying(100)	Texto	

primer_apellido		character varying(100)	Texto	
segundo_apellido		character varying(100)	Texto	
genero		character varying(20)	Texto	
fecha_nacimiento		date	Fecha	

Tabla Destino

Tabla: dim_estudiante		
Campo	Tipo	Correspondencia
cedula	character varying(20)	estudiante.cedula
primer_nombre	character varying(100)	estudiante.primer_nombre
segundo_nombre	character varying(100)	estudiante.segundo_nombre
primer_apellido	character varying(100)	estudiante.primer_apellido
segundo_apellido	character varying(100)	estudiante.segundo_apellido
genero	character varying(20)	estudiante.genero
fecha_nacimiento	date	estudiante.fecha_nacimiento

Proceso realizado:

- Extraer registros de la tabla estudiante
- Ordenar los registros
- Eliminar los registros repetidos
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_estudiante con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Materia

Descripción: Representa la carga de la tabla dim_materia

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene

código		character varying(10)	Texto		No tiene
sección		character varying(10)	Texto		No tiene
status		character varying(255)	Texto		No tiene
tipo		character varying(255)	Texto		No tiene
nombre		character varying(255)	Texto		No tiene
nro_credits		Integer	Numérico	No negativo	No tiene

Fuentes de Datos

Tabla: materia				
Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	Integer	Numérico	
código		character varying(10)	Texto	
sección		character varying(10)	Texto	
status		character varying(255)	Texto	
tipo		character varying(255)	Texto	
nombre		character varying(255)	Texto	
nro_credits		Integer	Numérico	

Tabla Destino

Tabla: dim_materia		
Campo	Tipo	Correspondencia
id	Integer	
código	character varying(10)	materia.código
sección	character varying(10)	materia.sección
status	character varying(255)	materia.status
tipo	character varying(255)	materia.tipo
nombre	character varying(255)	materia.nombre
nro_credits	Integer	materia.nro_credits

Proceso realizado:

- Extraer registros de la tabla materia
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_materia con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Tipo Examen

Descripción: Representa la carga de la tabla dim_tipo_examen

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(255)	Texto		No tiene

Fuentes de Datos

Tabla: tipo_examen				
Nombre	Llave	Tipo	Formato	Consideración Importante
id	PK	Integer	Numérico	
nombre		character varying(255)	Texto	

Tabla Destino

Tabla: dim_tipo_examen		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(255)	tipo_examen.nombre

Proceso realizado:

- Extraer registros de la tabla tipo_examen
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_tipo_examen con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Docente

Descripción: Representa la carga de la tabla dim_docente

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
nombre		character varying(255)	Texto		No tiene

Fuentes de Datos

Tabla: docente					
Nombre	Llave	Tipo	Formato	Consideración Importante	
id	PK	Integer	Numérico		
nombre		character varying(255)	Texto		

Tabla Destino

Tabla: dim_docente		
Campo	Tipo	Correspondencia
id	Integer	
nombre	character varying(255)	docente.nombre

Proceso realizado:

- Extraer registros de la tabla docente
- Ordenar los registros
- Eliminar los registros repetidos
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_docente con la información correspondiente

Nombre del Proceso: Poblar la Dimensión Nota

Descripción: Representa la carga de la tabla dim_nota

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
id	PK	Integer	Numérico	No nulo	No tiene
descripción		character varying(10)	Texto		No tiene
valor		Integer	Numérico		No tiene

Fuentes de Datos

Tabla: historial_academico				
Nombre	Llave	Tipo	Formato	Consideración Importante
carrera_id		character varying(10)	Texto	
estudiante_cedula		character varying(20)	Texto	
periodo_academico_id		character varying(10)	Texto	
docente_id		Integer	Numérico	
materia_id		Integer	Numérico	
tipo_examen_id		Integer	Numérico	
nota_en_final		character varying(10)	Texto	
nota_en_reparacion		character varying(10)	Texto	
nota_definitiva		character varying(10)	Texto	

Tabla Destino

Tabla: dim_nota		
Campo	Tipo	Correspondencia
id	Integer	
descripción	character varying(10)	historial_academico.nota_definitiva
valor	Integer	

Proceso realizado:

- Extraer registros de la tabla historial_academico
- Ordenar los registros
- Asignar al campo llamado valor la nota correspondiente, en los casos en los cuales la nota tenga un valor numérico. En caso contrario, este campo debe estar vacío.
- Generar identificador único del registro
- Elegir los campos de interés, su nombre, su formato, el tipo de dato de cada uno de ellos y el orden en el que se encuentran
- Cargar los registros en la tabla dim_nota

Nombre del Proceso: Poblar la Tabla de Hechos Historial Academico

Descripción: Representa la carga de la tabla fact_historial_academico

Descripción de Fuentes de Datos

Tipo de Fuente	Nombre	Descripción
Base de Datos	so_area_intermedia	Contiene información acerca de la población estudiantil y de las actividades realizadas por el Servicio de Orientación

Estandarización y Limpieza de los Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
carrera_id	PK, FK	character varying	Texto	No nulo	No tiene
estudiante_cedula	PK, FK	character varying	Texto	No nulo	No tiene
fecha_inicio_periodo_academico_id	PK, FK	character varying	Texto	No nulo	No tiene
fecha_fin_periodo_academico_id	PK, FK	character varying	Texto	No nulo	No tiene
docente_id	PK, FK	Integer	Numérico	No nulo	No tiene
materia_id	PK, FK	character varying	Texto	No nulo	No tiene
tipo_examen_id	PK, FK	Integer	Numérico	No nulo	No tiene
nota_en_final_id	PK, FK	Integer	Numérico	No nulo	No tiene
nota_en_reparacion_id	PK, FK	Integer	Numérico	No nulo	No tiene

nota_definitiva_id	PK, FK	Integer	Numérico	No nulo	No tiene
uc_inscritas		Integer	Numérico		No tiene
materias_inscritas		Integer	Numérico		No tiene
uc_aprobadas		Integer	Numérico		No tiene
materias_aprobadas		Integer	Numérico		No tiene
uc_aprobadas_equivalencia		Integer	Numérico		No tiene
materias_aprobadas_equivalencia		Integer	Numérico		No tiene
uc_aprobadas_general		Integer	Numérico		No tiene
materias_aprobadas_general		Integer	Numérico		No tiene
uc_aplazadas		Integer	Numérico		No tiene
materias_aplazadas		Integer	Numérico		No tiene
uc_retiradas		Integer	Numérico		No tiene
materias_retiradas		Integer	Numérico		No tiene
uc_cursadas		Integer	Numérico		No tiene
materias_cursadas		Integer	Numérico		No tiene

Fuentes de Datos

Tabla: historial_academico				
Nombre	Llave	Tipo	Formato	Consideración Importante
carrera_id		character varying(10)	Texto	
estudiante_cedula		character varying(20)	Texto	
periodo_academico_id		character varying(10)	Texto	
docente_id		Integer	Numérico	
materia_id		Integer	Numérico	
tipo_examen_id		Integer	Numérico	
nota_en_final		character varying(10)	Texto	
nota_en_reparacion		character varying(10)	Texto	
nota_definitiva		character varying(10)	Texto	

Tabla Destino

Tabla: fact_historial_academico		
Campo	Tipo	Correspondencia
carrera_id	character varying	
estudiante_cedula	character varying	
fecha_inicio_periodo_academico_id	character varying	
fecha_fin_periodo_academico_id	character varying	
docente_id	Integer	
materia_id	character varying	
tipo_examen_id	Integer	
nota_en_final_id	Integer	
nota_en_reparacion_id	Integer	
nota_definitiva_id	Integer	
uc_inscritas	Integer	
materias_inscritas	Integer	
uc_aprobadas	Integer	
materias_aprobadas	Integer	
uc_aprobadas_equivalencia	Integer	
materias_aprobadas_equivalencia	Integer	
uc_aprobadas_general	Integer	

materias_aprobadas_general	Integer	
uc_aplazadas	Integer	
materias_aplazadas	Integer	
uc_retiradas	Integer	
materias_retiradas	Integer	
uc_cursadas	Integer	
materias_cursadas	Integer	

Proceso realizado:

- Extraer registros de la tabla historial_academico para cada estudiante
- Determinar la licenciatura que cursa el estudiante
- Encontrar la clave foránea correspondiente en cada campo que lo necesite para cada registro
- Ordenar los registros
- Utilizar la nota definitiva de cada registro existente en el historial académico del estudiante con la finalidad de tomar en cuenta la asignatura y su cantidad de unidades de crédito correspondientes, según sea el caso
- Elegir los campos de interés, su nombre, su formato y el tipo de dato de cada uno de ellos
- Cargar los registros en la tabla fact_historial_academico