

UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE
APURÍMAC
FACULTAD DE INGENIERÍA
ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA INFORMÁTICA Y
SISTEMAS



**INTELIGENCIA DE NEGOCIOS USANDO PENTAHO PARA
LA GESTIÓN ACADÉMICA EN LA UNAMBA-2016**

**TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
INFORMÁTICO Y SISTEMAS**

PRESENTADO POR:

BACH. HAMELY SARMIENTO PONCE

TAMBURCO-ABANCAY-PERÚ

2018



UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE
APURÍMAC
FACULTAD DE INGENIERÍA
ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA INFORMÁTICA Y
SISTEMAS



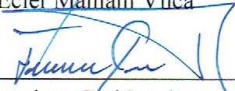
**INTELIGENCIA DE NEGOCIOS USANDO PENTAHO
PARA LA GESTIÓN ACADÉMICA EN LA UNAMBA-2016**

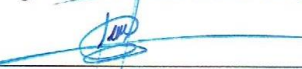
Presentado por la bachiller: Hamely Sarmiento Ponce, a la Escuela Académico Profesional de Ingeniería Informática y Sistemas, sustentado y aprobado el 18 de mayo del 2018 para obtener el título profesional de:

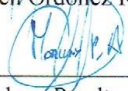
INGENIERO INFORMÁTICO Y SISTEMAS

Ante el jurado integrado por:

Presidente : 
M.Sc. Ecler Mamani Vilca

Primer miembro : 
Mg. Francisco Cari Incahuanao

Segundo miembro : 
Mg. Ezech Ordóñez Ramos

Asesor : 
Mg. Marleny Peralta Ascue

DEDICATORIA

A Dios, por iluminar mi camino y guiarme día a día para que persevere en cumplir mis metas.

A mi madre, Nalda Clara Ponce Valer por su constante apoyo y motivación para poder realizar la tesis.

A mi familia, que siempre me brindo su ayuda y me inspiro a lograr mis objetivos y a no dejar de perseverar en los estudios.

A la Universidad Nacional Micaela Bastidas de Apurímac por haberme albergado por tantos años en su casa de estudios.

AGRADECIMIENTOS

Agradezco a los docentes de la Escuela Académico Profesional de Ingeniería Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac, porque contribuyeron en mi formación académica durante mi estancia en la universidad.

A mi asesora de tesis, la Ing. Marleny Peralta Ascue, por brindarme su apoyo incondicional y motivarme a culminar mi tesis.

ÍNDICE

CAPÍTULO I	7
PLANTEAMIENTO DEL PROBLEMA	7
1.1. Definición y formulación del problema	7
1.1.1. Problema general	9
1.1.2. Problemas específicos	10
1.2. Justificación e importancia de la investigación	10
1.3. Alcances y Delimitaciones	11
CAPÍTULO II	13
OBJETIVOS	13
2.1. Objetivos	13
2.1.1. Objetivo general.....	13
2.1.2. Objetivos específicos.....	13
CAPÍTULO III	14
MARCO REFERENCIAL	14
3.1. Antecedentes	14
3.1.1. Antecedentes Internacionales	14
3.1.2. Antecedentes Nacionales	16
3.2. Marco Teórico	19
3.2.1. Inteligencia de Negocios	19
3.2.2. Gestión académica universitaria	29
3.2.3. Pentaho	31
3.2.4. Metodología HEFESTO	34
3.2.5. ISO/IEC 25012	42
3.2.6. Herramientas de inteligencia de negocios	45
3.3. Marco conceptual.....	47
CAPÍTULO IV	48
HIPÓTESIS Y VARIABLES	48
4.1. Hipótesis.....	48
4.1.1. Hipótesis general.....	48
4.1.2. Hipótesis específicas	48
4.2. Variables.....	48
4.3. Definición operacional de variables.....	49
CAPÍTULO V	50
METODOLOGÍA DE LA INVESTIGACIÓN	50
5.1. Tipo y nivel de investigación	50
5.2. Diseño y método de investigación	50
5.3. Población y muestra.....	51
5.4. Técnicas de investigación.....	51
5.4.1. Recolección de información	51
5.4.2. Procedimientos de la investigación.....	52

5.5. Plan de tratamiento de datos.....	53
CAPÍTULO VI	54
RESULTADOS.....	54
6.1. Análisis de interpretación de datos.....	54
6.1.1. Descripción de los resultados de la hipótesis	54
6.1.2. Contratación de la hipótesis	60
6.2. Resultados de la implementación de la inteligencia de negocios usando Pentaho para la gestión académica de la UNAMBA.	69
6.3. Implementación de la inteligencia de negocios	78
6.4. Implementación del Datamart usando la metodología Hefesto	79
6.4.1. Análisis de requerimientos.....	79
6.4.2. Análisis OLTP.....	84
6.4.3. Modelo lógico del datamart	94
6.4.4. Integración de datos	104
6.5. Análisis multidimensional	116
CONCLUSIONES	118
RECOMENDACIONES	120
BIBLIOGRAFÍA	121
ANEXOS	123

ÍNDICE DE TABLAS

Tabla 1: Cantidad de alumnos matriculados por carrera profesional en el 2016, sede Abancay	7
Tabla 2: Criterios de calidad en modelos conceptuales.....	42
Tabla 3: Características dependientes según ISO/IEC 25012	43
Tabla 4: Características inherentes según ISO/IEC 2012.....	44
Tabla 5: Características compartidas según ISO/IEC 25012	44
Tabla 6: Cuadro comparativo de herramientas de explotación de información de base de datos de inteligencia de negocios	46
Tabla 7: Herramientas de explotación de información de base de datos.....	46
Tabla 8: Operalización de variables	49
Tabla 9: Población.....	51
Tabla 10: Pre-prueba del número de reportes.....	54
Tabla 11: Post-prueba del número de reportes	55
Tabla 12: Pre-prueba del tiempo de consulta de respuesta	56
Tabla 13: Post-prueba del tiempo de consulta de respuesta	57
Tabla 14: Tiempo promedio de consultas antes y después de la inteligencia de negocios	57
Tabla 15: Puntajes de las alternativas de respuestas de la escala de Likert.....	58
Tabla 16: Puntajes de las respuestas de la calidad de datos del modelo de Inteligencia de negocios.....	59
Tabla 17: Resultados de modelo de calidad - Iso 25012	59
Tabla 18: Cálculos sobre el número de reportes obtenidos antes de la solución de inteligencia de negocios y después de la solución de inteligencia de negocios.....	61
Tabla 19: Cálculos sobre el tiempo de respuesta de consultas obtenidas antes de la inteligencia de negocios y después de la solución de inteligencia de negocios.....	66
Tabla 20: Herramientas utilizadas para la implementación de inteligencia de negocios	78
Tabla 21: Identificación de preguntas	79
Tabla 22: Indicadores y perspectivas.....	82
Tabla 23: Fuentes de datos del sistema de matrícula-grados y títulos.....	90
Tabla 24: Dimensión estudiante	95
Tabla 25: Dimensión sexo	95
Tabla 26: Dimensión tiempo.....	96
Tabla 27: Dimensión facultad.....	96
Tabla 28: Dimensión escuela	97
Tabla 29: Dimensión de categoría docente.....	97
Tabla 30: Dimensión de condición del docente	98
Tabla 31: Dimensión de departamento académico	98
Tabla 32: Dimensión asignatura	99
Tabla 33: Dimensión de categoría de asignatura.....	99
Tabla 34: Dimensión de fecha de colación.....	100
Tabla 35: Tabla de hechos de ingresante.....	100
Tabla 36: Tabla de hechos de matrícula de ingresante	101
Tabla 37: Tabla de hechos de docentes.....	101
Tabla 38: Tabla de hechos de grados y títulos	102
Tabla 39: Tabla de hechos de asignaturas.....	102
Tabla 40: Tabla de hechos de rendimiento académico.....	103

ÍNDICE DE GRÁFICOS

Figura 1: Arquitectura de la inteligencia de negocios	20
Figura 2: Componentes de Business Intelligences	21
Figura 3: Herramienta de Schema Workbench.....	33
Figura 4: Pasos de la Metodología Hefesto.....	34
Figura 5: Modelo Conceptual.....	37
Figura 6: Modelo Conceptual Ampliado	39
Figura 7: Diseño de tablas de dimensiones.....	40
Figura 8: Tabla de hechos.....	41
Figura 9: Disponibilidad de la información	55
Figura 10: Tiempo de respuesta de consulta	57
Figura 11: Resultados de calidad datos de modelo según ISO 25012	59
Figura 12: Resultados de muestras relacionadas de reportes con Minitab	63
Figura 13: Resultados de prueba de muestras relacionadas del tiempo de consulta en Minitab	68
Figura 14: Acceso principal al modelado de negocio.....	69
Figura 15: Reporte de número de matriculados, aprobados y desaprobados por asignatura y semestre	69
Figura 16: Cantidad de aprobados por asignaturas según semestre académico	70
Figura 17: Cantidad de estudiantes desaprobados por asignatura, según el semestre académico.....	71
Figura 18: Número de matriculados por asignatura y semestre académico	72
Figura 19: Gráfico estadístico de cantidad de matriculados por semestre académico	73
Figura 20: Total de docentes de cada escuela académico profesional por categoría	73
Figura 21: Total de docentes por género según facultad y escuela académico profesional.....	74
Figura 22: Total de graduados por escuela académica profesional de la UNAMBA según la modalidad	74
Figura 23: Total de estudiantes por sexo y grado según cada escuela académico profesional	75
Figura 24: Total de egresados por semestre académico	75
Figura 25: Gráfico estadístico de la cantidad de egresados por semestre	76
Figura 26: Número de egresados por género y año	76
Figura 27: Numero de ingresantes por semestre académico	77
Figura 28: Gráfico estadístico de número de ingresantes por semestre académico desde el 2008-I al 2016I	77
Figura 29: Número de retirados por escuela académico profesional y semestre académico	78
Figura 30: Modelado conceptual de Matriculado.....	83
Figura 31: Modelo conceptual de seguimiento académico	83
Figura 32: Modelo conceptual de docente	83
Figura 33: Modelo conceptual de grados y títulos	84
Figura 34: Modelo conceptual de Población Universitaria	84
Figura 35: Modelado de la base de datos relacional de servicios académicos	91
Figura 36: Modelado de base de datos multidimensional del datamart académico	94
Figura 37: Transformación de la dimensión asignatura	104
Figura 38: Transformación de la dimensión docente	104
Figura 39: Transformación de la dimensión estudiante.....	105
Figura 40: Transformación de la dimensión fecha de colación	105
Figura 41: Transformación de hecho asignatura.....	105
Figura 42: Transformación de la tabla de hechos grados y títulos	106
Figura 43: Transformación de la tabla de hechos rendimiento académico	107
Figura 44: Transformación de la dimensión modalidad	108
Figura 45: Transformación de la dimensión procedencia.....	108
Figura 46: Transformación de la dimensión categoría del docente	109
Figura 47: Transformación de la dimensión categoría de asignatura	109
Figura 48: Transformación de la dimensión de condición del docente.....	110
Figura 49: Transformación de la dimensión departamento académico	110
Figura 50: Transformación de la dimensión de egresado	111
Figura 51: Transformación de la dimensión escuela	111
Figura 52: Transformación de la dimensión facultad	111
Figura 53: Transformación de la dimensión grado académico	112
Figura 54: Transformación de la dimensión matricula	112
Figura 55: Transformación de la dimensión sexo.....	112

Figura 56: Transformación de la dimensión tiempo	113
Figura 57: Transformación de hecho docente.....	113
Figura 58: Transformación del hecho matricula del ingresante	114
Figura 59: Transformación del hecho población universitaria	115
Figura 60: Proceso de carga de data de las dimensiones y hechos	116
Figura 61: Estructura de cubos del modelo multidimensional.....	117

RESUMEN

El presente trabajo tiene como objetivo principal implementar una propuesta de solución de inteligencia de negocios usando Pentaho, para contribuir a mejorar el análisis de la información en la Gestión Académica de la Universidad Nacional Micaela Bastidas de Apurímac.

La UNAMBA, maneja un sistema transaccional el cual manipula datos operacionales los que por sí solos no generan información significativa para los directores, decanos, jefes de departamentos académicos ya que tienen que analizar la información para poder plantear planes de acción y estrategias de mejoras. Se propuso una solución de inteligencia de negocios para que las autoridades cuenten con una aplicación de gestión académica que proporcione información resumida, específica y precisa en tiempo real, que ayude a analizar, tomar decisiones y correcciones necesarias sobre el seguimiento y rendimiento académico de estudiantes, información de egresados, graduados y titulados, plana de docentes, y conocimiento de la población universitaria.

Para el desarrollo de la solución de inteligencia de negocios se utilizó la herramienta de Pentaho y se trabajó con la metodología Hefesto, la cual está compuesta por los siguientes pasos: análisis de requerimientos, análisis de data sources, modelo lógico e integración de datos. En la parte metodológica de investigación, el tipo de investigación utilizado es aplicada con un nivel experimental, con un diseño de investigación cuasi-experimental, método inductivo y con una muestra $n= 14$ personas.

Como resultados del trabajo de investigación, se obtuvo que el número de reportes después de la solución de Inteligencia de negocios incrementó a un 70.37% mientras que antes de la solución de inteligencia de negocios tan solo es un 29.62%. Además el tiempo de consulta es favorable ya que se redujo a un tiempo de 2.57 minutos que equivale a un 5.83% de lo que anteriormente era un promedio 41.5 minutos y que equivale al 94%.

La propuesta de solución de inteligencia de negocios no llegó a implantarse en las oficinas de las decanaturas, direcciones de las escuelas académicas profesional o jefaturas debido a que los equipos (las computadoras o laptops) no cumplen con los requerimientos mínimos de hardware que se necesitan para que soporte la herramienta de Pentaho y por la dificultad de accesibilidad a dichas oficinas.

Palabras claves: Inteligencia de negocios, Hefesto, datamart, gestión académica, Pentaho.

ABSTRACT

The main objective about this work is to implement a business intelligence solution using Pentaho, to help improve the information analysis in the Academic Management in National University Micaela Bastidas de Apurímac.

The UNAMBA manages a transactional system which manipulates operational data which alone does not generate significant information for directors, deans, heads of academic departments since they have to analyze the information to be able to propose action plans and improvement strategies. A business intelligence solution was proposed for the authorities to have an academic management application that provides summary, specific and accurate information in real time, which helps to analyze, make decisions and necessary corrections on the monitoring and academic performance of students. information on graduates, graduates and graduates, teachers' staff, and knowledge of the university population.

For the development of the business intelligence solution, the analysis and analysis tool of the Hephaestus methodology was used, which consists of the following steps: requirements analysis, data source analysis, logical model and data integration. In the methodological part of research, the type of research used is applied with an experimental level, with a quasi-experimental research design, inductive method and with a sample n equals 14 people.

As a result of the research work, it was obtained that the number of reports with Business Intelligence increased to 70.37% while without business intelligence it is only 29.62%. In addition, the consultation time is favorable since it was reduced to a time of 2.57 minutes which is equivalent to 5.83% of what was previously an average of 41.5 minutes more and which is equivalent to 94%.

The proposed business intelligence solution was not implemented in the decanaturas, addresses of professional academic schools or decanaturas because the equipment (computers or laptops) do not meet the minimum hardware requirements and the difficulty of accessibility to said offices.

Keywords:Business intelligence, Hephaestus, datamart, academic management, Pentaho.

INTRODUCCIÓN

En la actualidad, las empresas o instituciones manejan diversos tipos información las cuales son útiles siempre y cuando se le dé el uso adecuado de ésta, de no ser así, la información no se explotaría como debe ser trayendo algunos inconvenientes como: llevar a la parte gerencial a tomar decisiones incorrectas, no tener datos históricos o en tiempo real a la mano de los gerentes entre otras cosas.

La Universidad Nacional Micaela Bastidas de Apurímac cuenta con un sistema transaccional de matrículas que administra datos y los almacenan en una base de datos relacional que si bien permite obtener reportes y visualizar la información, estos son reportes tradicionales pre programados y lo muestran en dos dimensiones; a comparación de una solución de inteligencia de negocios que te muestra reportes dinámicos y gráficos estadísticos con más de dos dimensiones, logrando visualizar la información histórica de un estudiante, realizar un seguimiento académico, ver la población estudiantil, entre otros para el proceso de análisis de toma de decisiones y poder generar indicadores de estudio.

Por ello, se realizó ésta investigación para implementar la propuesta de solución de inteligencia de negocios transformando los datos en información y la información en conocimiento para optimizar el proceso de toma de decisiones y contribuir en mejorar la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac de manera que puedan realizar un análisis de datos de forma más rápida, informada y basada en hechos. Esta solución de Inteligencia de negocios permitió obtener la disponibilidad inmediata de la información a través de la generación de reportes dinámicos y gráficos estadísticos de diversos tipos (gráficos en barras, lineales, circulares, etc.) según los requerimientos de los jefes de departamento académico, decanos y directores y que las consultas obtenidas sean de manera eficiente y rápida.

Finalmente, el trabajo de investigación se dividió en 6 capítulos principales los cuales son: capítulo I, planteamiento del problema acerca de la gestión académica de la UNAMBA; capítulo II, objetivos del proyecto de investigación; capítulo III, marco referencial donde se describe los antecedentes de la investigación, el marco teórico y conceptual; capítulo IV, hipótesis y variables, capítulo V, metodología de la investigación Hefesto y capítulo VI, resultados obtenidos de las pruebas estadísticas con sus interpretaciones; conclusiones, recomendaciones y anexos del trabajo de investigación.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. Definición y formulación del problema

En la actualidad, las organizaciones manipulan diversas cantidades de información, las cuales cada vez se hacen más difíciles de manejar, y en vez de ser útil la información, en muchas circunstancias conlleva a un inadecuado uso de ésta.

La Universidad Nacional Micaela Bastidas de Apurímac (UNAMBA) cuenta con cuatro facultades las cuales son: la facultad de Ingeniería, una de las más grandes de la universidad, la facultad de Administración, la facultad de Medicina Veterinaria y Zootecnia y la facultad de Educación y Ciencias Sociales cada una de ellas con sus respectivas carreras como se muestra en el siguiente cuadro:

Tabla 1: Cantidad de alumnos matriculados por carrera profesional en el 2016, sede Abancay

Facultad	Carrera Profesional	Nº Alumnos
Ingeniería	Ingeniería Informática y Sistemas	335
	Ingeniería Agroindustrial	292
	Ingeniería de Minas	446
Administración	Administración de Empresas	482
Medicina veterinaria y Zootecnia	Medicina veterinaria y zootecnia	257
Educación y Ciencias Sociales	Ciencias políticas y gobernabilidad	269
	Educación Inicial Intercultural Bilingüe	300
		2381

Fuente: Oficina de Servicios Académicos.

Las facultades están encargadas de administrar y evaluar las escuelas que les han sido asignadas. El decano y los directores de las escuelas para poder llevar a cabo de manera eficiente su gestión, necesitan tener acceso a la información académica de los estudiantes, graduados y titulados, las cuales se encuentra administrada por la oficina de Servicios Académicos y la Unidad de Grados y Títulos.

Dichas áreas, no disponen de herramientas que les permita visualizar y analizar la información de una manera adecuada para poder tomar acciones correctivas, efectivas, lo cual dificulta el proceso de análisis para tomar decisiones. Por lo tanto, es necesario transformar la información en conocimiento, para aplicarlo en la gestión académica de manera que puedan realizar un análisis de datos de forma más rápida, informada y basada en hechos.

Las diversas facultades, necesitan elaborar reportes estadísticos, analíticos que le permitan examinar la situación académica actual de su Facultad o Escuelas Académicas Profesionales correspondientes. Las áreas en las que se enfocan sus análisis son: La población académica y sus tendencias a través del tiempo, rendimiento de los estudiantes, análisis de docentes y análisis de graduados y titulados.

La información se registra a través de un sistema transaccional u operacional denominado Sistema de Matrícula, que contiene toda la información académica del proceso de matrícula y rendimiento académico de estudiantes, cargas académicas de docentes, el cual se encuentra en la Oficina de Servicios Académicos.

Los datos operacionales por si solos no generan información significativa, asimismo se evidencia que los informes solicitados en la gestión académica son consolidados de datos numéricos y solicitados por las áreas antes mencionadas. Por ejemplo: resumen de datos de población académica, rendimiento académico, entre otros. Asimismo, para el registro de graduados y titulados utilizan una hoja electrónica Excel y la información que se requiere es resumida por ejemplo: el número de graduados y titulados por modalidad, sexo, año, entre otros.

Los análisis y reportes que se elaboran actualmente no permiten realizar comparaciones de datos históricos, pues no se aplican las herramientas necesarias.

Otra dificultad que se identificó es acerca de la elaboración del compendio estadístico por parte del área de planificación. La encargada del área de planeamiento recolecta información de los estudiantes solicitando dicha data a tres áreas, las cuales son: Oficina de grados y títulos, servicios académicos y admisión. Los datos son proporcionados en una hoja de cálculo excel, luego la encargada de la oficina tiene que comparar que coincida la información ya que servicios académicos y admisión registran la misma información y en algunas ocasiones no coincide duplicando esfuerzos. Un ejemplo clásico de esta inconsistencia es la cantidad de estudiantes que ingresan con el número de vacantes disponibles ya que hay menos vacantes y más estudiantes que ingresan. Luego de comparar la información y tener un consolidado, realiza sus gráficos estadísticos y sus cuadros de cada reporte que necesite de manera manual utilizando el software estadístico SPSS.

Hoy en día las tecnologías de información permiten poder mejorar este desempeño organizacional y poder obtener conocimiento para aplicarlo en la organización de manera que pueda orientar adecuadamente la toma de decisiones.

Finalmente, la plataforma que se utiliza para la implementación de la propuesta de solución de inteligencia de negocios es Pentaho debido a que compone un conjunto de herramientas integradas que permite realizar el análisis, el proceso extracción, transformación y carga, y los reportes de manera más didáctica.

1.1.1. Problema general

¿En qué medida la inteligencia de negocios usando Pentaho contribuye a mejorar el análisis de la información en la Gestión Académica de la Universidad Nacional Micaela Bastidas de Apurímac?

1.1.2. Problemas específicos

- a. ¿En qué medida la inteligencia de negocios usando Pentaho facilita la disponibilidad inmediata de la información analítica para la toma de decisiones de la Universidad Nacional Micaela Bastidas de Apurímac- Sede Abancay?
- b. ¿En qué medida la inteligencia de negocios usando Pentaho reduce el tiempo de consulta de la información académica de la Universidad Nacional Micaela Bastidas de Apurímac- Sede Abancay?

1.2. Justificación e importancia de la investigación

La importancia del presente trabajo consistió en implementar la propuesta de solución de inteligencia de negocios usando Pentaho, para contribuir a mejorar el análisis de la información y orientar adecuadamente el proceso de la toma de decisiones en la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac, ya que permitió explorar y obtener la información adecuada a disposición de los decanos y directores de cada facultad que la necesiten, en una forma entendible, interactiva, rápida, sencilla, incrementando la eficiencia debido a que la información se centralizó y se visualizó en una misma aplicación, convirtiéndola en útil y organizada, ahorrando tiempo y haciendo la toma de decisiones más eficiente, puesto que se obtiene respuestas más rápidas, todo ello implementado a través de cubos multidimensionales.

Se consiguió mejorar la disponibilidad inmediata de la información analítica para la toma de decisiones, reducir el tiempo de consulta de la información académica y consolidarla lo que permitió realizar los análisis de la población académica y sus tendencias a través del tiempo, rendimiento de los estudiantes, análisis de docentes y análisis de graduados y titulados, tener reportes de datos dispersos en conjunto, datos de diferentes dimensiones y eventos en el tiempo.

La herramienta que se utilizó es Pentaho debido a que es un software libre que permite crear informes y exportarlos a multitud de formatos, ayuda al usuario a ampliar la perspectiva en la toma de decisiones, ofrece al usuario final opciones de visualización en tiempo real de los datos que facilita la interacción con la herramienta y manejo de datos, dispone de aplicaciones potentes para la extracción carga y transformación de datos en diversas plataformas de base de datos.

1.3. Alcances y Delimitaciones

Alcances

- La propuesta de solución de inteligencia de negocios para la gestión académica de la UNAMBA sede Abancay, donde se consolida la información para realizar los análisis de la población académica y sus tendencias a través del tiempo, rendimiento de los estudiantes, análisis de docentes y análisis de graduados y titulados.
- La aplicación desarrollada es una propuesta de solución de inteligencia de negocios.

Delimitación Temporal

- Para el desarrollo de esta investigación se tomaron los datos del **área de servicios académicos** de la sede de Abancay desde los años 2008,2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016 y toda la información proporcionada por la **Unidad de Grados y Títulos** desde el año 2008 hasta el año 2014.

Delimitación Espacial

- Se desarrolló la Investigación en la Universidad Nacional Micaela Bastidas- sede Abancay, tomando en cuenta la facultad de ingeniería.

Delimitación teórica

- En esta investigación, para el análisis se utiliza cubos OLAP y la metodología Hefesto para la implementación de Inteligencia de Negocios.

Limitaciones

- Información desactualizada ya que la data proporciona por el área de servicios académicos es desde el semestre 2008-I hasta el semestre 2016-I.
- Los datos que corresponden al detalle de la información del estudiante, solo son de los estudiantes de la Escuela Profesional de Ingeniería Informática y Sistemas.
- La información de los docentes es desactualizada ya que en el mismo sistema transaccional de Matricula que manejan servicios académicos no se actualizó las categorías de los docentes.

CAPÍTULO II

OBJETIVOS

2.1. Objetivos

2.1.1. Objetivo general

Implementar la inteligencia de negocios usando Pentaho, para contribuir a mejorar del análisis de la información en la gestión Académica de la Universidad Nacional Micaela Bastidas de Apurímac.

2.1.2. Objetivos específicos

- a. Mejorar la disponibilidad inmediata de la información analítica con la inteligencia de negocios usando Pentaho en la Universidad Nacional Micaela Bastidas de Apurímac.
- b. Reducir el tiempo de consulta de la información académica con la inteligencia de negocios usando Pentaho de la Universidad Nacional Micaela Bastidas de Apurímac.
- c. Utilizar un conjunto de herramientas para la extracción, transformación y carga de los datos, que aseguren la consistencia de los datos dentro del Data Marts.
- d. Contar con una interfaz de acceso a cubos multidimensionales por medio de herramientas de inteligencia de negocios, que permitan consultar los reportes preestablecidos y, al mismo tiempo, facilite la elaboración de reportes dinámicos.

CAPÍTULO III

MARCO REFERENCIAL

3.1. Antecedentes

3.1.1. Antecedentes Internacionales

Castillo Rojas Wilson, Medina Fernando y Fariño Francisco (2017) en el artículo denominado: **“Data Mart para obtención de indicadores de productividad académica en una universidad”**, el cual fue publicado en la revista chilena de ingeniería afirman que las organizaciones utilizan la información para la toma de decisiones estratégicas. El Aumento de datos ha generado dos grandes problemas: Identificar los datos relevantes para dar seguimiento a la estrategia y lograr que cumplan los planes con las metas establecidas, administrar una gran cantidad de información para la toma de decisiones, utilizarla en el momento adecuado y verificar que apoye en el cumplimiento de la estrategia. Entonces, el artículo presenta y describe un PDW, para el desarrollo de un DW, que permita obtener indicadores de productividad académica de la Universidad Arturo Prat de Chile. El aporte principal considerado es la adaptación de una metodología para Datawarehouse, la cual integra diversos enfoques y técnicas tales como la especificación de requerimientos de la información utilizada en la ingeniería de software, modelo de desarrollo combinado basado en las propuestas de Kimball y Hefesto, proceso de ETL aumentado con una fase de validación de indicadores y visualizaciones integradas e interactivas para el análisis multidimensional de indicadores.

Aimacaña Quilumba Doris Eulalia (2013) en la tesis denominada “**Análisis, Diseño e Implementación de un Datamart académico usando tecnología de BI para la facultad de ingeniería, ciencias físicas y matemática**”, para optar el título de Ingeniería Informática en la Universidad Central de Ecuador. El sistema académico de la facultad de ingeniería maneja miles de registros de datos sobre estudiantes, docentes y empleados, datos que pueden ser aprovechados para la toma de decisiones, mejorando el manejo de esta información en diferentes áreas para integrar y compartir información asignándoles atributos que sirvan como base para el análisis. Por ello un Datamart aporta a la institución eliminando aquellos datos que obstaculizan la labor del análisis de la información y entregan la información que se requiera en la forma más apropiada, facilitando así el proceso de gestión de la misma. Entonces, el objetivo fue implementar un modelo de control de decisiones gerenciales tomando como aplicación la creación de Reportes e Indicadores de Control, el cual ayudará con la planificación estratégica para la toma de decisiones para el bien de la facultad; así como también se emplearán mapas estratégicos que nos permitan canalizar las metas de la organización en objetivos concretos. La investigación a realizar para este proyecto es la Investigación de Campo, que va a permitir mediante la manipulación de datos obtener información que sirva de apoyo a la facultad, permitirá conocer en qué condiciones se encuentra el manejo de la información, así como también poder determinar las mejores prácticas que se podrían aplicar para el proceso técnico de la facultad.

En conclusión, el uso de la interfaz de inteligencia de negocios de reportes de Pentaho permite un manejo intuitivo y sencillo para los usuarios finales para generar reportes y análisis acorde a las necesidades del negocio en comparación del uso de hojas de cálculo.

3.1.2. Antecedentes Nacionales

Zegarra Fuentes Gustavo Fernando (2016) en la tesis **“Solución de Inteligencia de Negocios orientada a la toma de decisiones en las operaciones mineras de extracción y metalurgia de hochschild mining”**, de la Universidad San Martín de Porres, para optar el título Ingeniero de Computación y Sistemas.

Para la recaudación de información asociada a los procesos mineros de extracción y metalurgia, se desea implementar una solución de inteligencia de negocios que permita disminuir el tiempo requerido, conocer información no evidente acerca de dichos procesos y mostrarlos en indicadores de medición de rendición. Ésta solución permitirá obtener un análisis confiable, haciendo un uso eficiente de recursos y sirve para hacer comparaciones y tomar decisiones de manera más eficaz. La investigación de la tesis tuvo como objetivo brindar soporte a la toma de decisiones empleando una solución de inteligencia de negocios de las operaciones de extracción y metalurgia. Las conclusiones que obtuvo fueron que evidenció que en el proceso de toma de decisiones a nivel de la gerencia de mineras es tener una alta disponibilidad de la información, así como una síntesis fiable de los procesos de extracción y metalurgia resultó en una mejora en la toma de decisiones que propicio la consecución de metas.

Gutiérrez Rojas Jesús (2015) en la tesis **“Implantación de Business Intelligence como mejora en la gestión académica del Área de Coordinación académica de la UPIG, Surco-2015”**, para optar el título profesional de Ingeniero de Sistemas e informático de la Universidad Peruana de Integración Global, tiene como objetivo central dar la solución al problema de las necesidades de la información para la toma de decisiones de las autoridades y coordinadores de la UPIG, realizando la implementación de la inteligencia de negocios usando la plataforma SQL SERVER 2008 R2. La implantación de inteligencia de Business Intelligence mejora la Gestión académica del área de coordinación de la UPIG, Surco 2015. Para el desarrollo, se aplicó la técnica de investigación mixta que consiste en la recopilación y tratamiento de datos. Los métodos de investigación que se utilizaron son de investigación analítico-sintéticos e inductivos-deductivos. Como principal conclusión se establece que el Business Intelligence es la habilidad para transformar los datos en información y la información en conocimiento, de forma que se puede optimizar el proceso en la toma de decisiones. Se demostró que la Inteligencia de negocios mejoró en la creación de reportes, incremento la cantidad de toma de decisiones en el mes, la UPIG obtuvo grandes beneficios al utilizar el DATAMART académico puesto que se podrá analizar el comportamiento de los estudiantes para tomar mejores decisiones en el uso de recursos, los procesos de extracción, transformación y carga son apropiados según la información requerida para los directores.

Rojas Saldivar Alejandro (2014) en la tesis **“Implementación de un data Mart como solución de inteligencia de negocios, bajo la metodología de Ralph Kimball para optimizar la toma de decisiones en el departamento de finanzas de la contraloría general de la república”**, de la Universidad San Martín de Porres, para optar el título profesional de Ingeniero de Computación y Sistemas. El problema surge porque los usuarios del Departamento de finanzas de la contraloría general de la república realizan el procedimiento de obtención de información de reportes de manera manual. Estos solicitan al departamento de tecnologías la Información semestral y anual, datos para generar reportes de presupuestos, los cuales son extraídos directamente de la base de datos y entregados en hojas de cálculo o archivos pdf, encardándose directamente de depurar, ordenar u clasificar estos datos con el fin de obtener los reportes para realizar sus labores, Por ello se realizó un proyecto que permita brindar una herramienta que automatice el procedimiento, para tomar esta información estructurada como base para que los usuarios puedan analizar y explotar la información. El objetivo principal es implementar una solución de inteligencia de negocios para el departamento de finanzas, de manera que cuenten con información sólida y confiable de alta disponibilidad, que cumplan con las necesidades de información. La hipótesis general es que la implementación de un datamart como solución de inteligencia de negocios bajo la metodología de Ralph kimball, optimizara el proceso de la toma de decisiones.

Como conclusiones, se logró identificar los procesos que permiten llevar a cabo la toma de decisiones, se analizó y selecciono los datos indispensables para el desarrollo de la solución de la inteligencia de negocios, lo que garantizó la validez y calidad de los mismos, se construyó un modelo OLAP, que permitió ejecutar las consultas a partir de la información procesada y se diseñó las interfaces mediante la herramienta QlikView para desplegar el cubo OLAP.

3.2. Marco Teórico

3.2.1. Inteligencia de Negocios

Se conoce por inteligencia de negocios al conjunto de aplicaciones, metodologías, prácticas y capacidades encaminadas a la creación y gestión de la información que contribuye a tomar mejores decisiones a los usuarios de la empresa (Conesa y Curto, 2011).

Es el conjunto de estrategias y tecnologías que nos van a ayudar a convertir los datos en información de calidad, y dicha información en conocimiento, que nos facilite en la toma de decisiones más apropiadas y en nuestra competitividad. El objetivo básico de Business Intelligence es apoyar de forma sostenible y continua a las organizaciones para mejorar su competitividad, facilitando la información necesaria para la toma de decisiones. Howard Dresner fue el primero que popularizo BI como un término paraguas para describir un conjunto de conceptos y métodos que mejoraran las tomas de decisiones, utilizando información sobre qué había sucedido (Ramos, 2011).

La inteligencia de negocios es el término que procura caracterizar una amplia variedad de tecnologías, plataformas de software especificaciones de aplicaciones y proceso. El objetivo principal es contribuir en la toma de decisiones que optimicen el desempeño dentro de la empresa y promover su ventaja competitiva en el mercado. (Ayala, 2006).

3.2.1.1. Arquitectura de Inteligencia de Negocios

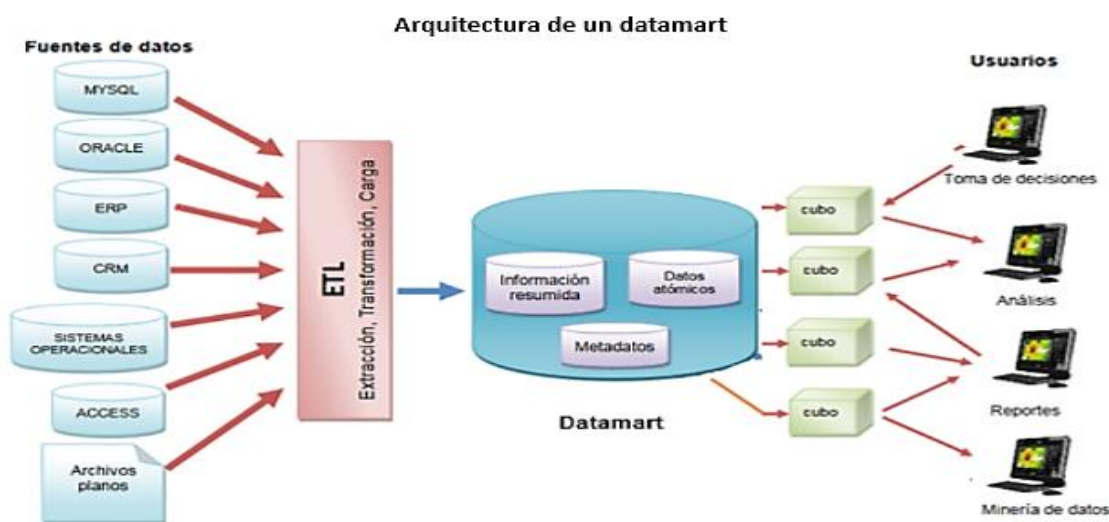


Figura 1: Arquitectura de la inteligencia de negocios
Fuente: Bernabeu (2010).

Según Bernabeu y Garcia (2017), su funcionamiento puede resumirse de la siguiente manera:

- Se extraen los datos de diferentes fuentes, bases de datos, archivos, servicios web, etc. Generalmente estos datos se encuentran en diferentes tipos de sistemas con arquitecturas y formatos variados.
- Los datos son integrados, transformados y limpiados, para luego ser cargados en el datamart.
- Con base en el datamart, se construirán los cubos multidimensionales.
- Los usuarios accederán a los cubos multidimensionales, business models (u otro tipo de estructura de datos) del datamart, utilizando diversas herramientas de consulta, exploración, análisis, reportes, etc.

3.2.1.2. Componentes de Inteligencia de Negocios

Según Cano (2007), los componentes de Business Intelligence son:

- Fuentes de información: punto de partida para alimentar la información del datawarehouse.
- Proceso ETL. Extracción, transformación y carga de los datos de datawarehouse
- Datawarehouse o almacén de datos, se busca almacenar los datos de una forma que maximice su flexibilidad, facilidad de acceso y administración.
- OLAP, debe proveer la capacidad de cálculo, consultas, funciones de planeamiento, pronóstico y análisis de escenarios de datos.
- Herramientas de visualización, permitirán el análisis y la navegación a través de los mismos.

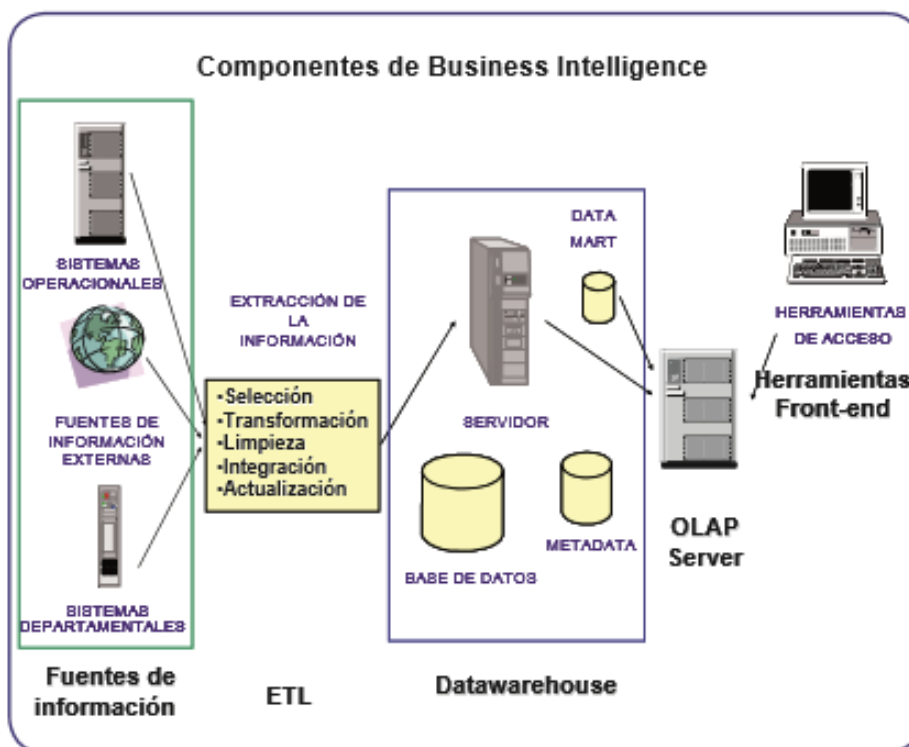


Figura 2: Componentes de Business Intelligences
Fuente: Cano (2007)

A. Fuentes de información

Cano (2007). Las fuentes de información a las que podemos acceder son:

- Básicamente, a los sistemas operacionales o transaccionales, que incluyen aplicaciones desarrolladas a medida a medida ERP, CRM, SCM, etc.
- Sistemas de información departamentales: previsiones, presupuestos, hojas de cálculo, etc.
- Fuentes de información externa son fundamentales para enriquecer la información que tenemos de nuestros clientes. En algunos casos es interesante incorporar información referente.

B. Proceso de ETL

El proceso ETL, consume entre el 60% y el 80% del tiempo de un proyecto de Business Intelligence, por lo que es el proceso clave en la vida de todo el proyecto. Cano (2007)

B.1. Extracción de los datos

El principal objetivo de la extracción es extraer tan solo aquellos datos de los sistemas transaccionales que son necesarios y prepararlos para el resto de los subprocessos de ETL. Para ello, se debe determinar las mejores fuentes de información, las de mejor calidad. Con tal finalidad, deberemos analizar las fuentes disponibles y escoger aquellas que sean mejores. (Cano, 2007)

Según Salvador Ramos (2011). Uno de los aspectos fundamentales que debemos considerar a la hora de diseñar es si optamos por una actualización completa, que es mucho más simple, o si optamos por una actualización incremental, que es lo más conveniente.

Si optamos por la primera alternativa el proceso consiste en borrar los datos y volver a cargarlos, pero si optamos por la segunda, habrá que establecer una serie de controles y técnicas, entre las que destacamos las siguientes:

- Borrado sólo de parte de los datos, por ejemplo desde una fecha, y carga de los nuevos datos.
- Comparar los datos de origen y de destino, actualizando sólo los cambios.
- Uso de triggers u otras técnicas de replicación.
- Captura de datos mediante una aplicación diseñada específicamente para ello.
- Existencia de columnas timestamp que permitan seleccionar a las filas modificadas desde la última extracción y aplicar dichos cambios al destino.

B.2. Transformación de datos

La transformación de los datos se hace partiendo de los datos una vez limpios, transformando los datos de acuerdo con las reglas de negocio y los estándares que han sido establecidos. La transformación incluye: cambios de formato, sustitución de códigos, valores derivador y agregados. (Cano, 2007)

En los procesos de transformación, es preciso asegurarnos de que los datos sean válidos, de su integridad y de su utilidad, lo que suele incluir realizar calculo y generar nuevos valores. Los datos deben ser limpios para borrar inconsistencias, disconformidades y duplicidad.

Estas transformaciones suelen conllevar cambios en el contenido de los valores de origen y creación de nuevos valores en las filas de destino. (Ramos, 2011)

B.3. Carga en el Data Warehouse

Como último paso debemos realizar el proceso de incorporar los datos al Data Warehouse y/o a los diferentes Data Marts, y a los cubos OLAP. Todo ello según la presentación, formatos definidos y periodicidad de actualización propuesta. (Ramos, 2011)

Según Cano (2007), la carga de datos es el momento en el que cargamos los datos y debemos comprobar que se ha desarrollado correctamente, ya que en caso contrario pueden llevar a decisiones erróneas a los usuarios.

C. Datawarehouse

Es un almacén o repositorio de datos que brinda una visión general, común e integrada de datos de la empresa, independiente de cómo lo vayan a utilizar después los usuarios y tiene las siguientes propiedades: es estable, coherente, fiable y cuenta con información histórica. (Conesa y Curto, 2011)

Según Ramos (2011), Un Datawarehouse es una base de datos corporativa en la que se integra información depurada de las diversas fuentes que hay en la organización. Dicha información debe ser homogénea y fiable, se almacena de forma que permita su análisis desde muy diversas perspectivas, y que a su vez de unos tiempos de respuesta óptimos.

Para ello la información se encuentra des normalizada y modela de una forma bastante diferente a los sistemas transaccionales, principalmente se utilizan los modelos de estrella y un copo de nieve. Un Data Warehouse según Bill Inmon se caracteriza por ser:

- **Orientado a temas:** los datos están organizados por temas ara facilitar el entendimiento por parte de los usuarios, de forma que todos los datos relativos a un mismo elemento de la vida real queden unidos entre sí.
- **Integrado:** los datos se deben integrar en una estructura consistente debiendo eliminarse las inconsistencias existentes entre los diversos sistemas operacionales. La información se estructura en diversos niveles de detalle para adecuarse a las necesidades de consulta de los usuarios.
- **Histórico:** los datos, que pueden ir variando a lo largo el tiempo, deben quedar reflejados de forma que al ser consultados reflejen estos cambios y no se altere la realidad que había en el momento en que se almacenaron, evitando así la problemática que ocurre en los sistemas operacionales, que reflejan solamente el estado de la actividad de negocio presente, Un data Warehouse debe almacenar los diferentes valores que toma una variable a lo largo del tiempo.
- **No volátil:** la información de un Data Warehouse, una vez introducida, debe ser sólo lectura, nunca se modifica ni se elimina, y ha de ser permanente y mantenerse para futuras consultas.

La estructura relacional de una base de datos operacional sigue las formas normales en su diseño. Un datawarehouse no debe seguir ese patrón de diseño. La idea principal es que la información sea presentada des normalizada para optimizar las consultas. Para ello debemos identificar, en el seno de nuestra organización, los procesos de negocio, las vistas para el proceso de negocio y las medidas cuantificables asociadas a los mismos. Ramos (2011)

De esta manera hablaremos de:

- Tabla de hecho: es la representación en el datawarehouse de los procesos de negocio de la organización.
- Dimensión: es la representación en el datawarehouse de una vista para un cierto proceso de negocio.
- Métrica: Son los indicadores de un proceso de negocio. Son aquellos conceptos cuantificables que permiten medir nuestro proceso de negocio.

D. OLAP

Se entiende por OLAP, al proceso analítico en línea, el método ágil y flexible para organizar datos, especialmente metadatos, sobre un objeto o jerarquía como en un sistema u organización multidimensional, y cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas e incluso informes. (Conesa y Curto, 2011)

El procesamiento analítico en línea OLAP tiene como objetivo agilizar la consulta de grandes volúmenes de información. Para ello utiliza estructuras multidimensionales, conocidas como cubos OLAP, que contienen datos pre calculados y agregados.

Estos sistemas tienen una velocidad de respuesta muy superior a los sistemas OLTP. (Ramos Salvador, 2011, p. 16)

Un cubo OLAP es un vector multidimensional, de N dimensiones, aunque por su nombre puede hacernos creer inicialmente que solo tiene tres dimensiones.

En él la información se almacena en cada una de estas dimensiones, de forma ordenada y jerarquizada, lo que nos ayuda a realizar un rápido análisis de su contenido. Una base de datos multidimensional puede contener varios de estos cubos OLAP. Por tanto, un cubo OLAP está estructurado en dimensiones, que son las diferentes perspectivas desde las que queremos analizar la información, en medidas que son los diferentes hechos con valores concretos que solicita el usuario.

E. Herramientas de Business Intelligence

Las principales herramientas de Business Intelligence según Cano (2007) son:

- Generadores de informes: Utilizadas por desarrolladores profesionales para crear informes estándar para grupos, departamentos la organización.
- Herramientas de usuario final de consultas de informes: Empleadas por usuarios finales para crear informes para ellos mismos o para otros; no requieren programación.
- Herramientas OLAP: Permiten a los usuarios finales tratar la información de forma multidimensional para explotarla desde distintas perspectivas y periodos de tiempo.

- Herramientas de Dashboard: y scorecard: Permiten a los usuarios finales ver la información crítica para el rendimiento con un simple vistazo utilizando iconos gráficos y con la posibilidad de ver más detalle para analizar la información detallada e informes, si lo desean.
- Herramientas de planificación, modelización y consolidación: Permite a los analistas y a los usuarios finales crear planes de negocios y simulaciones con la información de Business Intelligence.
- Herramientas datamining: Permiten a analistas o estadísticos de negocio crear modelos estadísticos de las actividades de negocio. El datamining es el proceso para descubrir e interpretar patrones desconocidos en la información mediante las cuales resolver problemas de negocio.

3.2.1.3. Modelo de datos multidimensionales

Ramos (2011). A la hora de modelar el Data Mart o Data Warehouse, hay que decidir cuál es el esquema más apropiado para obtener los resultados que queremos conseguir.

a. Esquema en estrella y en copo de nieve

Habitualmente, y salvo excepciones, se suele modelar la base de datos utilizando el esquema en estrella (star schema), en el que hay una única tabla central, la tabla de hechos, que contiene todas las medidas y una tabla adicional por cada una de las perspectivas desde las que queremos analizar dicha información, es decir por cada una de las dimensiones. (Ramos, 2011)

b. Esquema de copo de nieve

La otra alternativa de modelado es la utilización del esquema en copo de nieve (snowflake schema). Esta es una estructura más compleja que el esquema en estrella. La diferencia es que algunas de las dimensiones no están relacionadas directamente con la tabla de hechos, sino que se relacionan con ella a través de otras dimensiones. En este caso también tenemos una tabla de hechos, situada en el centro, que contiene todas las medidas y una o varias tablas adicionales, con un mayor nivel de normalización (Ramos, 2011)

c. Esquema de constelación

Este esquema está compuesto por una serie de esquemas de estrella, posee una tabla de hechos principal y una o más tablas de hechos auxiliares, las cuales pueden ser agregaciones de la principal. Dichas tablas están relacionadas con sus respectivas tablas de dimensiones. (Bernabeu y García, 2017).

3.2.2. Gestión académica universitaria

La gestión es un elemento determinante de la calidad del desempeño de las organizaciones; ella incide en el clima organizacional, en las formas de liderazgo y conducción institucional (Gobierno), en el aprovechamiento óptimo de los talentos, en la planificación de las tareas y la distribución del trabajo y su productividad, en la eficiencia de la administración y el rendimiento de los recursos materiales y, por cada uno de esos conceptos, para el caso de las instituciones educativas, en la calidad de los procesos educacionales.

La calidad del desempeño ha sido una preocupación permanente de los directivos, conscientes como están de la relación que guarda ésta con otras variables de gestión como la productividad y la competitividad, claros indicadores de la salud organizacional. (Blanco y Quesada, 2008).

Las universidades son organismos vivos dedicados a generar dinámicas de entendimiento y comprensión del contexto global para mejorar la calidad de vida de las personas. De ahí que se indique que, como organismos vivos, tienen una función socioeducativa innovadora y difusora de vida.

El clima organizacional se basa en las percepciones individuales, a menudo se define como los patrones recurrentes de comportamiento, actitudes y sentimientos que caracterizan la vida en la organización, y se refieren a las situaciones actuales en una organización y los vínculos entre los grupos de trabajo, los empleados y el desempeño laboral. (Campos y Dominguez, 2012)

El respeto a ese conjunto de percepciones individuales y la capacidad de mediar para lograr condiciones de entendimiento colectivas permiten una gestión académica saludable que involucra la suma de voluntades y posibilita que las personas se aventuren y se identifiquen con proyectos pedagógicos comunes que engrandezcan a la colectividad y lleven a niveles de convivencia, identidad y bienestar, necesarios para el despliegue de una gestión académica transformadora.

Las innovaciones se caracterizan por políticas claras de motivación que cautivan a la totalidad de miembros de la comunidad educativa, quienes comprenden y se identifican plenamente con el compromiso de gestar experiencias de aprendizaje y de vida que revelen y atiendan el bienestar común.

Desde esta perspectiva, la motivación supera las resistencias y los estados de comodidad. Implica una participación activa y articulada de las personas en proyectos educativos que enaltezcan y promuevan cambios en la cultura organizacional, aspecto que implica evaluación permanente. (Castillo, Flores y Cervantes, 2015)

3.2.3. Pentaho

En el mercado, Pentaho es una de las suites más completas y maduras que existen desde el año 2006. Esta herramienta sigue la estrategia open core que trata que a partir de un núcleo open source se brindan servicios y módulos mejorados. Por ese motivo existen dos versiones: la versión enterprise la que se ofrece bajo una modalidad de suscripción y la versión Community la cual es completamente gratuita (Conesa y Curto, 2011).

Los requerimientos mínimos de hardware para instalar Pentaho son los siguientes: Procesador 4 cores, RAM de 8gb (4gb para Pentaho data Integration dedicado y 4gb servidor apache tomcat) y espacio en disco de 80gb.

Algunas herramientas integradas que posee Pentaho son las siguientes:

a. Pentaho Data Integration: pdi (ex kettle)

Permite cargar datos de distintas fuentes de origen en un data warehouse o datamart para que luego la información se consolide y sea de utilidad a nivel operativo, táctico y estratégico. (Conesa y Curto, 2011)

Según Conesa y Curto (2011) las principales características de PDI son:

- Entorno gráfico orientado al desarrollo rápido y ágil basado en dos áreas: la de trabajo y la de diseño/vista.
- Multiplataforma.
- Incluye múltiples conectores a bases de datos, tanto propietarias como comerciales. Así como conectores a ficheros planos, Excel, XML u otros.
- Arquitectura extensible mediante plugins.
- Soporta uso de cluster, procesos ETL en paralelo y arquitecturas servidor maestro-esclavo.
- Completamente integrado con la suite de Pentaho
- Basado en el desarrollo de dos tipos de objetos:
 - Transformaciones: permiten definir las operaciones de transformación de datos.
 - Trabajos: permiten gestionar y administrar procesos ETL a alto nivel.

b. Schema workbench

Es una herramienta de desarrollo que permite crear, modificar y publicar un esquema de Mondrian. Es un programa java multiplataforma. Permite crear todos los objetos que soporta Mondrian: esquema, cubo, dimensiones, métricas. Tiene dos áreas: la que muestra la estructura jerárquica del esquema OLAP y la zona de edición de las propiedades de cada elemento. Presenta un menú superior para crear cubos, dimensiones, dimensiones conformadas, métricas, miembros calculados, subconjuntos (named set) y roles, así como operaciones estándar como cortar, copiar y pegar (Conesa y Curto, 2011).

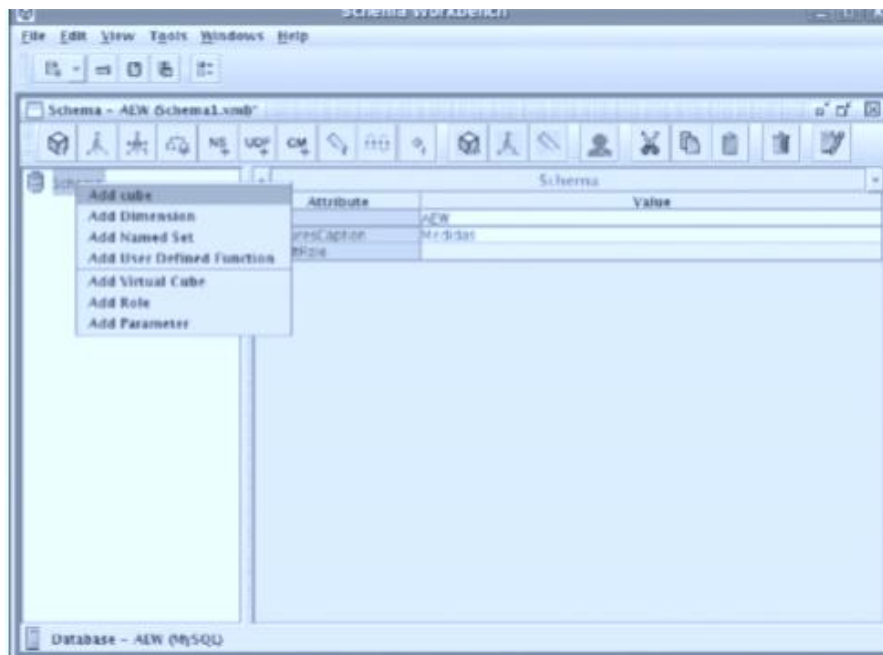


Figura 3: Herramienta de Schema Workbench
Fuente: Conesa y Curto (2011)

3.2.4. Metodología HEFESTO

Según Bernabeu (2010), HEFESTO es una metodología fundamentada en una extensa investigación, compara metodologías existentes, experiencias oportunas en procesos de elaboración del datawarehouse.

Cabe destacar que HEFESTO está en continua evolución, y se han tenido en cuenta, como gran valor agregado todos los feedbacks que han aportado quienes han utilizado esta metodología en diversos países y con diversos fines.

A continuación se puede resumir la metodología Hefesto en los siguientes pasos, a través del siguiente gráfico:



Figura 4: Pasos de la Metodología Hefesto
Fuente: Bernabeu (2010)

Una vez hecho eso, se pasará a la construcción del modelo lógico del depósito, en donde se definirá cual es el tipo de esquema que se implementará. Luego se elaborarán las tablas de dimensiones y de hechos, para luego realizar sus respectivas uniones.

Por último, utilizando técnicas de limpieza y calidad de datos, proceso ETL, etc, se definirán políticas estratégicas para la carga inicial del DW y su respectiva actualización.

3.2.4.1. Pasos y aplicación de la metodología

Según Bernabeu (2010), los pasos y la aplicación de la metodología Hefesto son:

A. Análisis de requerimientos

Se identifican los requerimientos de los usuarios a través de preguntas que expliquen los objetivos de su organización. Seguidamente, se analizan esas preguntas con la finalidad de identificar los indicadores y perspectivas que serán tomadas en cuenta para la construcción del datawarehouse o datamart.

a. Identificar preguntas

El primer paso comienza con el acopio de las necesidades de información, el cual puede llevarse a cabo a través de muy variadas y diferentes técnicas, cada una de las cuales poseen características inherentes y específicas, como por ejemplo entrevistas, cuestionarios, observaciones, etc.

El análisis de los requerimientos de los diferentes usuarios, es el punto de partida de esta metodología, ya que ellos son los que deben, en cierto modo, guiar la investigación hacia un desarrollo que refleje claramente lo que se espera del depósito de datos, en relación a sus funciones y cualidades.

El objetivo principal de esta fase, es la de obtener e identificar las necesidades de información clave de alto nivel, que es esencial para llevar a cabo las metas y estrategias de la empresa, y que facilitará una eficaz y eficiente toma de decisiones. Debe tenerse en cuenta que dicha información, es la que proveerá el soporte para desarrollar los pasos sucesivos, por lo cual, es muy importante que se preste especial atención al relevar los datos.

Una forma de asegurarse de que se ha realizado un buen análisis, es corroborar que el resultado del mismo haga explícitos los objetivos estratégicos planteados por la empresa que se está estudiando. (Bernabeu, 2010, p.89)

b. Identificar indicadores y perspectivas

Después de determinar las preguntas de negocio, se procede a descomponer cada una de ellas para descubrir los indicadores que se utilizarán y las perspectivas de análisis que intervendrán. Para ello, se debe tener en cuenta que los indicadores, para que sean realmente efectivos son, en general, valores numéricos y representan lo que se desea analizar concretamente.

En cambio, las perspectivas se refieren a los objetos mediante los cuales se quiere examinar los indicadores, con el fin de responder a las preguntas planteadas. (Bernabeu, 2010, p.90)

c. Modelo Conceptual

En esta etapa, se construirá un modelo conceptual a partir de los indicadores y perspectivas obtenidas en el paso anterior.

A través de este modelo, se podrá observar con claridad cuáles son los alcances del proyecto, para luego poder trabajar sobre ellos, además al poseer un alto nivel de definición de los datos, permite que pueda ser presentado ante los usuarios y explicado con facilidad. (Bernabeu, 2010, p.91)

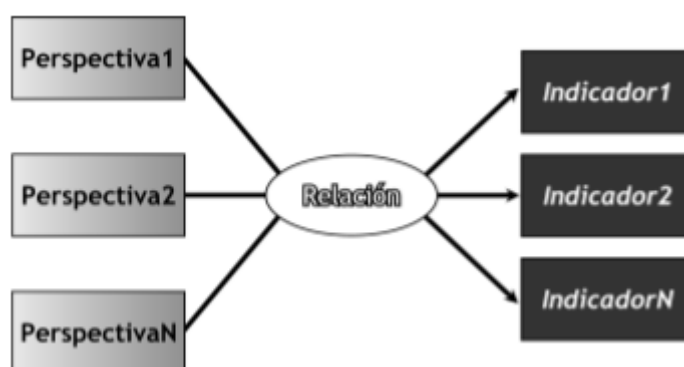


Figura 5: Modelo Conceptual
Fuente: Bernabeu (2010)

d. ANÁLISIS DE LOS OLTP

Se analizarán las fuentes OLTP para calcular los indicadores y establecer correspondencias entre el modelo conceptual creado en el paso anterior y las fuentes de datos. Luego, se definirán qué campos se incluirán en cada perspectiva. Finalmente, se ampliará el modelo conceptual con la información obtenida en este paso (Bernabeu, 2010)

a. Conformar indicadores

Explicar cómo se calculan los indicadores, definiendo los siguientes conceptos en cada uno de ellos:

Hecho/s que lo componen, con su respectiva fórmula de cálculo. Por ejemplo: Hecho1 + Hecho2.

Se utiliza función de sumarización por ejemplo: SUM, AVG, COUNT, etc.

b. Establecer correspondencias

Para identificar las correspondencias entre el modelo conceptual y las fuentes de datos, primero se debe explorar los OLTP disponibles que contengan la información requerida y sus características.

La idea es que los elementos de OLTP estén correspondidos con los elementos del modelo conceptual. (Bernabeu, 2010, p.90)

c. Nivel de granularidad

Luego de establecer las relaciones con los OLTP, se seleccionan los campos que contendrá cada perspectiva para examinar y filtrar los indicadores. Para ello, se debe presentar a los usuarios los datos de análisis disponibles para cada perspectiva. Es muy importante conocer en detalle que significa cada campo y/o valor de los datos encontrados en los OLTP, por lo cual, es conveniente investigar su sentido. (Bernabeu, 2010, p.96)

d. Modelo conceptual ampliado

Con la finalidad de graficar los resultados obtenidos en los pasos anteriores, se amplía el modelo conceptual, colocando bajo cada perspectiva los campos seleccionados y bajo cada indicador su respectiva fórmula de cálculo. (Bernabeu, 2010)

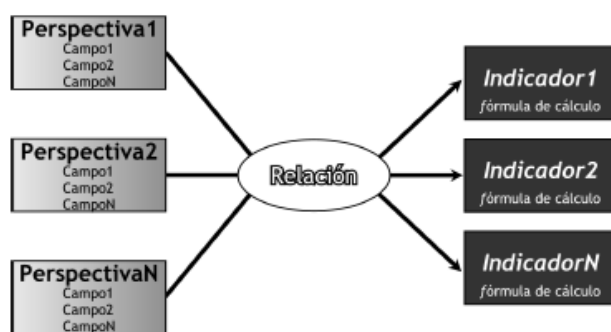


Figura 6: Modelo Conceptual Ampliado
Fuente: Bernabeu (2010)

e. Modelo lógico del datamart

a. Tipo de modelo lógico del datamart

Se debe seleccionar cuál será el tipo de esquema que se utilizará para contener la estructura del depósito de datos, que se adapte mejor a los requerimientos y necesidades de los usuarios.

Es muy importante definir objetivamente si se empleará un esquema en estrella, constelación o copo de nieve, ya que esta decisión afectará considerablemente la elaboración del modelo lógico. (Bernabeu, 2010, p.99)

b. Tablas de dimensiones

Según Bernabeu (2010) se deben diseñar las tablas de dimensiones que formaran parte del DW, para lo cual existen tres tipos de esquemas y cada perspectiva definida en modelo conceptual constituirá una tabla de dimensión. Para ello deberá tomarse cada perspectiva con sus campos relacionados y realizarse el siguiente proceso:

- Se elegirá un nombre que identifique la tabla de dimensión.
- Se añadirá un campo que represente su clave principal.
- Se redefinirán los nombres de los campos si es que no son lo suficientemente intuitivos.



Figura 7: Diseño de tablas de dimensiones
Fuente: Bernabeu (2010)

c. Tabla de hechos

Bernabeu (2010). Se definen las tablas de hechos, que son las que contendrán los hechos a través de los cuales se construirán los indicadores de estudio. Para los esquemas en estrella y copo de nieve, se realizará lo siguiente:

- Dar un nombre a la tabla de hechos que simbolice la información analizada, área de investigación, negocio enfocado, etc.
- Se define la clave primaria, que se compone de la combinación de las claves primarias de cada tabla de dimensión relacionada.

- Crear tantos campos de hechos como indicadores se hayan definido en el modelo conceptual y asignar los mismos nombres que estos.

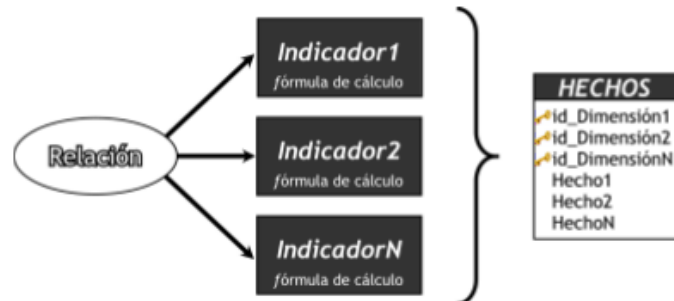


Figura 8: Tabla de hechos
Fuente: Benabeu (2010)

d. Uniones

Se realizan las uniones correspondientes entre sus tablas de dimensiones y sus tablas de hechos.

e. Integración de datos

Ya teniendo el modelo lógico, se procede a poblarlos con datos utilizando técnicas de limpieza y calidad de datos, procesos ETL; etc.

a) Carga inicial

Realizar la carga inicial al DW, poblando el modelo de datos que hemos construido, para lo cual se realizan tareas que pueden contener una lógica realmente compleja en algunos casos. Se debe evitar que el DW sea cargado con valores faltantes o anómalos, así como también se deben establecer las conexiones y restricciones para asegurar que solo se utilicen los datos de interés.

Cuando se trabaja con constelación, hay que tener presente que varias tablas de dimensiones serán compartidas con diferentes tablas de hechos.

Primero se cargarán los datos de las dimensiones y luego de las tablas de hechos, teniendo en cuenta, la correcta correspondencia entre cada elemento.

b) Actualización

Cuando se haya cargado en su totalidad el DW, se deben establecer sus políticas y estrategias de actualización o refresco de datos. Una vez realizado esto se tendrán a cabo las siguientes acciones:

- Especificar las tareas de limpieza de datos, calidad de datos, procesos ETL, que deberán realizarse para actualizar los datos del DW.
- Especificar de forma general y detallada las acciones que deberá realizar cada software.

3.2.5. ISO/IEC 25012

Gonzales, M. y Gonzales, J. (2013). La ISO/IEC para medir la calidad se divide en dos criterios:

a. Criterios de calidad en modelos conceptuales

Tabla 2: *Criterios de calidad en modelos conceptuales*

CRITERIO	DESCRIPCIÓN
Legibilidad	Son consideraciones visuales para la lectura y presentación del modelo conceptual.
Compleitud	Los requerimientos se deben de plasmar en el modelo del diseño a desarrollar, es decir cada uno de los requerimientos deben estar representados en el modelo.
Corrección	Se puede observar dos perspectivas: La sintáctica, cuando las diferentes partes de un modelo están construidas con respecto al lenguaje utilizado y la semántica, cada elemento del problema se representa haciendo uso de las estructuras adecuadas.
Minimalidad	Se considera mínimo un modelo conceptual si no existe información duplicada o redundante.
Expresividad	Intenta medir la capacidad de comunicación del modelo a nivel semántico. La realidad es representada en el modelo de manera que con sus elementos esta puede ser comprendida fácilmente.
Autoexplicación	Todos los requisitos pueden ser representados en el modelo, por lo que la lógica de negocio con respecto a los datos puede ser entendida por el modelo conceptual.
Extensibilidad	Un esquema tiene la capacidad de tolerar cambios en los requisitos y de poder adaptarse a nuevas necesidades de los usuarios de la base de datos.

Fuente: Revista facultad de Ingeniería UPTC. Miguel Gonzales Pinzón-Juan Gonzales Sonobria

b. Estándares de calidad en bases de datos

Representa los modelos de calidad para software y datos. Entonces, se establecen las características externas de calidad, que a su vez se descomponen en sub características deseables tanto en el software como en los datos, brindándose consejos prácticos sobre el uso de modelos de calidad, formulando los lineamientos para la calidad de los datos almacenados en un sistema de información con un formato estructurado.

El estándar presenta quince características de calidad que pueden ser abordadas desde los puntos de vista dependiente e inherente del sistema.

(Gonzales, M. y Gonzales, J. 2013, p.115).

Tabla 3: Características dependientes según ISO/IEC 25012

DIMENSIÓN	DESCRIPCIÓN
Disponibilidad	Es el grado en el cual el dato tiene atributos que le permiten ser recuperados por usuarios autorizados o por aplicaciones en un contexto específico de uso.
Portabilidad	Es el grado en el cual el dato tiene los atributos que le permiten ser instalado, substituido o movido de un sistema a otro conservando la calidad existente en un contexto específico de uso
Recuperabilidad	Es El grado en el cual el dato puede mantener y conservar un nivel específico de operaciones y calidad, aún en caso de falla.

Fuente: Revista facultad de Ingeniería UPTC. Miguel Gonzales Pinzón-Juan Gonzales Sonobria

Inherente, cuando la calidad de los datos se refiere a los datos en sí mismos; en general, este punto de vista está relacionado con los aspectos del dominio gestionado por expertos.

Tabla 4: Características inherentes según ISO/IEC 2012

DIMENSIÓN	DESCRIPCIÓN
Exactitud	El dato tiene atributos que representan de manera correctamente el valor del atributo intencionado de un concepto o evento en un contexto específico.
Complejidad	El dato del sujeto asociado con una entidad tiene valores para todos los atributos esperados e instancias de entidad relacionadas en un contexto específico de uso.
Consistencia	El dato tiene varios atributos que son libres de contradicción y coherentes con otros datos en un determinado contexto específico de uso.
Credibilidad	El dato posee atributos que son verdaderos y creíbles por usuarios en un contexto específico de uso.
Actualidad	El grado en el cual el dato tiene los atributos que son del periodo correcto en un contexto específico de uso.

Fuente: Revista facultad de Ingeniería UPTC. Miguel Gonzales Pinzón-Juan Gonzales Sonobria

En la siguiente tabla, se representan las características comunes a ambas perspectivas, además el estándar también define el modelo del proceso que puede ser aplicado para el uso del modelo de calidad de datos.

Tabla 5: Características compartidas según ISO/IEC 25012

DIMENSIÓN	DESCRIPCIÓN
Accesibilidad	El dato puede ser accedido dentro de un contexto específico de uso, en particular por gente que necesita soporte de tecnología o una configuración especial debido a alguna inhabilidad (incapacidad).
Conformidad	El dato tiene atributos que se adhieren a normas, convenciones o regulaciones vigentes y reglas similares relacionadas con la calidad de datos en un contexto específico de uso.
Confidencialidad	El dato tiene atributos que aseguran que solo es accesible e interpretable por usuarios autorizados en un contexto específico de uso.
Eficiencia	El dato tiene los atributos que pueden ser procesados y proporciona los niveles esperados de funcionamiento (desempeño) usando las cantidades y los tipos de recursos apropiados en un contexto específico de uso.
Precisión	El dato tiene atributos que son exactos o que proporcionan la discriminación en un contexto específico de uso.
Trazabilidad	El dato tiene atributos que proporcionan un rastro de auditoría de acceso a los datos y de cualquier cambio hecho a los datos en un contexto específico de uso.
Entendibilidad	El dato tiene atributos que le permiten ser leído e interpretado por usuarios, y es expresado en lenguajes apropiados, símbolos y unidades en un contexto específico de uso.

Fuente: Revista facultad de Ingeniería UPTC. Miguel Gonzales Pinzón-Juan Gonzales Sonobria

3.2.6. Herramientas de inteligencia de negocios

Según Torres, G & Salinas, O. (2011), se presentan diferentes herramientas de explotación de datos dentro del área de Inteligencia de negocios:

- Aricom: Herramienta multidimensional de análisis de generación de consultas sql dinámicas donde los usuarios sin tener conocimientos técnicos pueden ejecutar y realizar reportes complejos desde cualquier base relacional.
- Actue BIRT: Oferta comercial basada en código abierto
- Jaspersoft Bussiness Intelligence Suite: Conjunto de herramientas de Jaspersoft que permiten a una organización generar información basada en sus propios datos de gestión para la evaluación y toma diaria de decisiones., en forma dinámica y en línea.
- Pentaho: Plataforma de Inteligencia de negocios orientada a la solución y centrada en los procesos.
- SpagoBI: Es la Osbi Suite (Open Source Business Intelligence) que marca la diferencia en el código abierto, cubriendo las áreas analíticas de los proyectos de inteligencia de negocios, apropiada para cualquier requerimiento siendo apropiado para ambientes empresariales.
- SQL Power BI: Gama de herramientas de productividad innovadores que permiten simplificar la minería de datos, inteligencia de negocios y medición de desempeño.

Tabla 6: Cuadro comparativo de herramientas de explotación de información de base de datos de inteligencia de negocios

Módulo	Ariacom * Producto comercial	ActuateBirt *Opciones completas en su versión comercial	JasperSoft *Opciones completas en su versión comercial	Pentaho	SpagoBI	SQLPower
Servidor de Aplicación	Aplicación Conexión Web a través del uso de un complemento	Integrado *iServer	JBoss	JBoss, Tomcat	JBoss, Tomcat	Integrado
Colaboración BI	Integrado	LDAP	-	-	Dossier	Integrado * Versión Enterprise
Cuadro de mando	-	-	JFreeChart	JFreeChart	Openlaszlo	Integrado
Minería de datos	Integrado	Integrado	-	Weka	Weka	-
Sistema Manejador de Base de Datos	MS Access, MS SQL Server, Oracle y MySQL *Conexión con otras bases de datos, requiere la instalación del controlador	Informix, Oracle, PostgreSQL, DB2.	MySQL, Oracle, MSSQL Server, PostgreSQL, etc.	MySQL, Oracle, MSSQL Server, PostgreSQL, etc.	MySQL, Oracle, MSSQL Server, PostgreSQL, etc. *Conexión con otras bases a través de Drivers (archivos .jar).	Oracle, MySQL, MS SQL Server, SQL Stream, PostgreSQL, HSQLDB

Fuente: Torres & Salinas (2011)

Tabla 7: Herramientas de explotación de información de base de datos

Módulo	Ariacom * Producto comercial	ActuateBirt *Opciones completas en su versión comercial	JasperSoft *Opciones completas en su versión comercial	Pentaho	SpagoBI	SQLPower
	ODBC.					
ETL	Integrado	-	JasperETL	Pentaho Data Integration	Talend Open Studio	-
Geo-referencia	-	-	Google Maps	Google Maps	GEO	-
Cronograma de trabajo	Integrado	Integrado	Quartz	Quartz	Quartz	Integrado * Versión Enterprise
OLAP ¹	Integrado	Integrado como motor interno en tablas cruzadas	Mondrian y JPivot	Mondrian y JPivot	Mondrian y JPivot	Mondrian
Reportes	Integrado	ActuateBirtReportDesigner	JasperReport	Pentaho Report Designer, JasperReport, Birt	JasperReport, Birt	Integrado
Single SingOn (SSO) ²	Integrado	-	Acegi	CAS	CAS	-

Fuente: Torres & Salinas (2011)

3.3. Marco conceptual

- **Almacén de datos:** También conocido como datawarehouse, es un conjunto de información creada para aplicaciones de toma de decisiones.
- **Cubo OLAP:** Los cubos de procesamiento analítico en línea es una estructura de datos de una base de datos multidimensional, donde el almacenamiento físico de datos se lleva a cabo en un vector multidimensional.
- **ETL:** Es un proceso que permite extraer los datos de diferentes fuentes, transformarlos y cargarlos en otra base de datos, datamart o datawarehouse para poder analizar.
- **Hefesto:** Es una metodología, aplicada en la construcción del datawarehouse o datamart.
- **Información analítica:** se denomina así a la información que se considera para el análisis.
- **Modelo Multidimensional:** Es una disciplina de diseño que se sustenta en el modelo entidad relación y en las realidades de la ingeniería de texto y datos numéricos; se representa a través de la definición de las tablas de hechos y dimensiones.
- **Open Source:** Software de código abierto, cuyo código fuente está disponible para la modificación o mejora por cualquier persona.
- **Pentaho:** Es una plataforma compuesta de herramientas de análisis e informes integrados, siendo capaz de ejecutar reglas de negocio, expresadas en formas de proceso y entregada la información adecuada en el momento oportuno.
- **Sistema transaccional:** Es un tipo de sistema de información diseñado para recolectar, almacenar modificar y recuperar todo tipo de información que es generada por las transacciones en una organización.

CAPÍTULO IV

HIPÓTESIS Y VARIABLES

4.1. Hipótesis

4.1.1. Hipótesis general

La solución de inteligencia de negocios usando Pentaho contribuye a mejorar el análisis de la información en la Gestión Académica de la Universidad Nacional Micaela Bastidas de Apurímac.

4.1.2. Hipótesis específicas

- a. La solución de inteligencia de negocio usando Pentaho mejorará la disponibilidad inmediata de la información en el análisis para la toma de decisiones en la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac.
- b. La solución de inteligencia de negocio usando Pentaho reducirá el tiempo de consulta de la información académica de la Universidad Nacional Micaela Bastidas de Apurímac.

4.2. Variables

- **Variable independiente:** Inteligencia de negocios usando Pentaho.
- **Variable dependiente** : Gestión académica.

4.3. Definición operacional de variables

Tabla 8: Operalización de variables

VARIABLES	DIMENSIÓN	INDICADORES	INDICES
Inteligencia de Negocios Usando Pentaho Según Ramos (2011) La inteligencia de negocios es el conjunto de estrategias y tecnologías que nos van a ayudar a convertir los datos en información de calidad, y dicha información en conocimiento que nos permita una toma de decisiones más acertadas y nos ayude así a mejorar nuestra competitividad.	ISO/IEC 25012 Calidad en modelos conceptual	Legibilidad	a. Totalmente de acuerdo
			b. De acuerdo
			c. Ni de acuerdo ni en desacuerdo.
			d. En desacuerdo
			e. Totalmente en desacuerdo
		Completitud	a. Totalmente de acuerdo
			b. De acuerdo
			c. Ni de acuerdo ni en desacuerdo.
			d. En desacuerdo
			e. Totalmente en desacuerdo
		Corrección	a. Totalmente de acuerdo
			b. De acuerdo
			c. Ni de acuerdo ni en desacuerdo.
			d. En desacuerdo
			e. Totalmente en desacuerdo
		Minimalidad	a. Totalmente de acuerdo
			b. Ni de acuerdo ni en desacuerdo.
			c. Indeciso
			d. En desacuerdo
			e. Totalmente en desacuerdo
		Expresividad	a. Totalmente de acuerdo
			b. De acuerdo
			c. Ni de acuerdo ni en desacuerdo.
			d. En desacuerdo
e. Totalmente en desacuerdo			
Auto explicación	a. Totalmente de acuerdo		
	b. De acuerdo		
	c. Ni de acuerdo ni en desacuerdo.		
	d. En desacuerdo		
	e. Totalmente en desacuerdo		
Extensibilidad	a. Totalmente de acuerdo		
	b. De acuerdo		
	c. Ni de acuerdo ni en desacuerdo.		
	d. En desacuerdo		
	e. Totalmente en desacuerdo		
Gestión académica Según Blanco, I., & Quesada, V (2008). Es el elemento determinante de la calidad del desempeño de las organizaciones; incide en el clima organizacional, en las formas de liderazgo y conducción institucional (Gobierno), en el aprovechamiento óptimo de los talentos, en la planificación de las tareas y la distribución del trabajo y su productividad, en la eficiencia de la administración y el rendimiento de los recursos.	Análisis de información analítica	Información disponible	Número de reportes al mes
		Consulta de información académica	Tiempo en segundos

CAPÍTULO V

METODOLOGÍA DE LA INVESTIGACIÓN

5.1. Tipo y nivel de investigación

- a. **Tipo de investigación:** El presente estudio es una investigación aplicada, debido a que permitió determinar el nivel de contribución de la inteligencia de negocios usando Pentaho para la gestión académica de la UNAMBA-2016
- b. **Nivel de investigación:** Se pretende conocer cuáles son las mejoras que se logró con la implementación de la propuesta de solución de inteligencia de negocios, por lo que se utiliza el nivel experimental.

5.2. Diseño y método de investigación

- a. **Diseño de investigación:** El presente proyecto permitió implementar una solución de inteligencia de negocios lo cual permitió obtener metas, para lo cual se utilizó el diseño cuasi-experimental ya que la muestra es intacta es decir la muestra es igual a la población y se utilizó la pre y post prueba debido a que el grado de control es mínimo, es decir se aplicó a un mismo grupo dos tipos de situaciones distintas, antes de la solución de inteligencia de negocios y después de la solución de inteligencia de negocios, de manera más específica se usó el diseño de pre prueba y post prueba con un solo grupo, teniendo el siguiente diagrama.

G -----01-----X-----02

- b. **Método de investigación:** En el presente trabajo de investigación se utilizó el método inductivo.

5.3. Población y muestra

a. Población

Para el presente estudio, la población seleccionada está constituida por las personas que se encargan de realizar el análisis de la información, es decir está conformada por los decanos, jefes de departamentos académicos y directores de Escuelas Académicos Profesionales de la Universidad Nacional Micaela Bastidas de Apurímac-Sede Abancay.

N=14

Tabla 9: Población

PERSONAS	CANTIDAD
Decanos	4
Directores	7
Jefes de departamentos académicos	3
TOTAL	14

b. Muestra

Para la selección de la muestra se consideró que el muestreo es no probabilístico y a selección por criterio del investigador se considera la misma cantidad de la población.

n=14

5.4. Técnicas de investigación

5.4.1. Recolección de información

Se efectuó la recolección de datos e información, con el fin de dar respuesta al problema de investigación:

Las técnicas que se utilizó en la recolección de datos son:

- Entrevistas con los stakeholders como autoridades de la facultad de ingenierías, usuarios.
- Recopilación de datos e información que manejan la facultad.

El instrumento que se utilizó es:

- El cuestionario

5.4.2. Procedimientos de la investigación

Etapa 1: Definición del problema

El primer paso en esta investigación, es establecer una definición del problema con el objeto que comprender lo que se pretende resolver, es decir implementar un modelo de solución de inteligencia de negocios usando Pentaho para la gestión académica de la UNAMBA.

Etapa 2: Revisión bibliográfica

En la segunda etapa, se desarrolla la revisión bibliográfica, que abarca los aspectos conceptuales relacionados con la inteligencia de negocios.

Etapa 3: Diseño de la solución

Se desarrollara un modelo de inteligencia de negocios utilizando la herramienta Pentaho para la gestión académica de la UNAMBA, la cual consta de los siguientes pasos:

a. Análisis de datos

Se recolectó los datos del área de servicios académicos y de grados y títulos, se verificó los requerimientos de la gestión académica de la UNAMBA. Los datos conseguidos son hojas de cálculo Excel, y también el modelo relacional de la base de datos, donde se muestran las diversas tablas con sus respectivos campos.

b. Modelado dimensional

En base al modelo relacional de la base de datos, se construye el modelado dimensional.

c. Construcción del modelo

Para la construcción del modelo, se realizó el modelado dimensional y el análisis de datos para la extracción, carga y transformación de datos, identificando las variables relevantes, y planteando la utilización de una herramienta de inteligencia de negocios.

d. Visualización Reportes

Se utiliza Cubos OLAP para generar consultas y reportes. Además dichos reportes son dinámicos y poseen una interfaz agradable.

Etapas 4: Análisis y evaluación

En esta etapa se debe realizar el análisis y evaluación de los resultados o comparando los distintos modelos aplicados.

Etapas 5: Conclusiones y trabajo futuro

Finalmente se presentan las conclusiones de este trabajo y se establece algunas líneas de investigación.

5.5. Plan de tratamiento de datos

Para el tratamiento estadístico de los datos se utilizó:

- Tablas y gráficos estadísticos que permitan analizar y visualizar los resultados que se obtengan de los cuestionarios.
- Se aplicación de la prueba del estadístico t para muestras relacionadas de un grupo intacto en dos tiempos diferentes.

CAPÍTULO VI

RESULTADOS

6.1. Análisis de interpretación de datos

6.1.1. Descripción de los resultados de la hipótesis

a) Disponibilidad de la información (Número de reportes de la gestión académica)

Se llevó a cabo una prueba con un solo grupo con la finalidad de obtener el número de reportes antes y después de la solución de inteligencia de negocios, para lo cual se realizó una encuesta donde se preguntó a los directores, Jefes de departamento académico, decanos, ex decanos y ex directores si cuentan o no con diferentes tipos de reportes y gráficos estadísticos.

Pre prueba

Tabla 10: Pre-prueba del número de reportes

Usuarios	Reportes
Director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	10
Director de la Escuela Académico Profesional de Ingeniería de Minas	9
Ex director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	12
Director de la Escuela Académico Profesional de Ingeniería Agroindustrial	8
Director de la Escuela Académico Profesional de Medicina Veterinaria y Zootecnia	8
Director de la Escuela Académico Profesional de Administración de empresas	11
Director de la Escuela Académico Profesional de Educación Inicial y Bilingüe	7
Director de la Escuela Académico Profesional de Ciencias Políticas	6
Decano de la facultad de Educación y Ciencias Sociales	8
Ex decano de la facultad de Ingenierías	9
Decano de la facultad de Medicina veterinaria y Zootecnia	7
Decano de la facultad de Ingenierías	10
Jefe de Departamento Académico de Ingeniería	2
Ex Jefe Departamento de Ciencias básicas	6

Post prueba

Tabla 11: Post-prueba del número de reportes

Usuarios	Reportes
Director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	20
Director de la Escuela Académico Profesional de Ingeniería de Minas	20
Ex director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	20
Director de la Escuela Académico Profesional de Ingeniería Agroindustrial	20
Director de la Escuela Académico Profesional de Medicina Veterinaria y Zootecnia	20
Director de la Escuela Académico Profesional de Administración de empresas	20
Director de la Escuela Académico Profesional de Educación Inicial y Bilingüe	20
Director de la Escuela Académico Profesional de Ciencias Políticas	20
Decano de la facultad de Educación y Ciencias Sociales	22
Ex Decano de la facultad de Ingenierías	22
Decano de la facultad de Medicina veterinaria y Zootecnia	22
Decano de la facultad de Ingenierías	22
Jefe de Departamento Académico de Ingeniería	9
Ex Jefe de departamento de Ciencias básicas	9

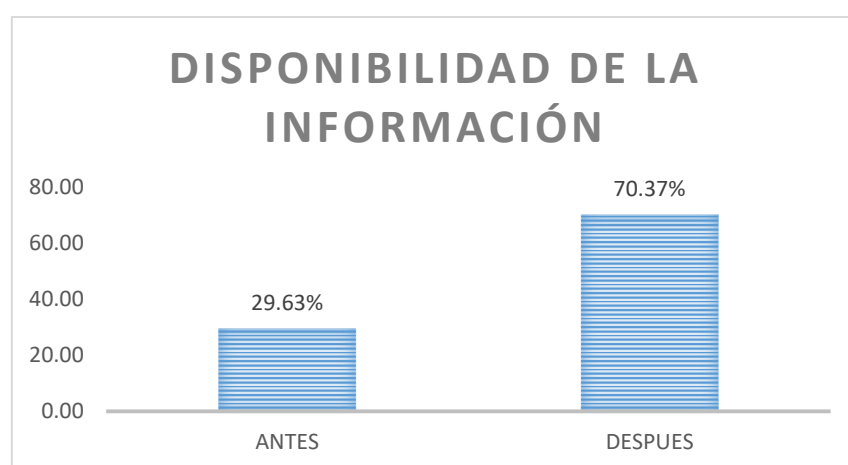


Figura 9: Disponibilidad de la información

Interpretación

Como se muestra en la figura 7, existe una diferencia entre las medias de la cantidad de reportes del antes de la solución de Inteligencia de negocios y después de utilizar la solución de Inteligencia de negocios.

Se pudo observar que el número de reportes antes era un promedio de 8 reportes que equivale al 29.62% y ahora incremento a un promedio de 19 reportes que equivalen al 70.37 % disponibles a los usuarios.

b) Tiempo de consultas de la información

Se llevó a cabo una prueba con un solo grupo con la finalidad de conocer el tiempo de respuesta de consultas antes y después de la solución de Inteligencia de negocios. Para ello, se realizó una encuesta donde se preguntó a los directores, Jefes de departamento académico, decanos, ex decanos y ex directores el tiempo de respuesta de consulta en obtener los reportes que necesitan.

Pre prueba

Tabla 12: *Pre-prueba del tiempo de consulta de respuesta*

Usuarios	Tiempo de respuesta (min)
Director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	55
Director de la Escuela Académico Profesional de Ingeniería de Minas	50
Ex director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	60
Director de la Escuela Académico Profesional de Ingeniería Agroindustrial	45
Director de la Escuela Académico Profesional de Medicina Veterinaria y Zootecnia	47
Director de la Escuela Académico Profesional de Administración de empresas	58
Director de la Escuela Académico Profesional de Educación Inicial y Bilingüe	40
Director de la Escuela Académico Profesional de Ciencias Políticas	36
Decano de la facultad de Educación y Ciencias Sociales	30
Ex decano de la facultad de Ingenierías	35
Decano de la facultad de Medicina veterinaria y Zootecnia	40
Decano de la facultad de Ingenierías	50
Jefe del Departamento Académico de Ingeniería	10
Ex Jefe del Departamento de Ciencias Básicas	25

Post prueba

Tabla 13: Post-prueba del tiempo de consulta de respuesta

Usuarios	Tiempo de respuesta (min)
Director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas	2.5
Director de la Escuela Académico Profesional de Ingeniería de Minas	2.5
Ex director de la Escuela Académico Profesional de Ingeniería Informática y Sistemas.	2.5
Director de la Escuela Académico Profesional de Ingeniería Agroindustrial	2.5
Director de la Escuela Académico Profesional de Medicina Veterinaria y Zootecnia	2.5
Director de la Escuela Académico Profesional de Administración de empresas	2.5
Director de la Escuela Académico Profesional de Educación Inicial y Bilingüe	2.5
Director de la Escuela Académico Profesional de Ciencias Políticas	2.5
Decano de la facultad de Educación y Ciencias Sociales	3.2
Ex Decano de la facultad de Ingenierías	3.2
Decano de la facultad de Medicina veterinaria y Zootecnia	3.2
Decano de la facultad de Ingenierías	3.2
Jefe de Departamento Académico de Ingeniería	1.6
Ex Jefe del Departamento de Ciencias básicas	1.6

Tabla 14: Tiempo promedio de consultas antes y después de la inteligencia de negocios

	Tiempo Promedio	%
Tiempo de consultas (Antes)	41.5	94.1683685
Tiempo de consultas (Después)	2.57	5.8316315
TOTAL	44.07	100 %

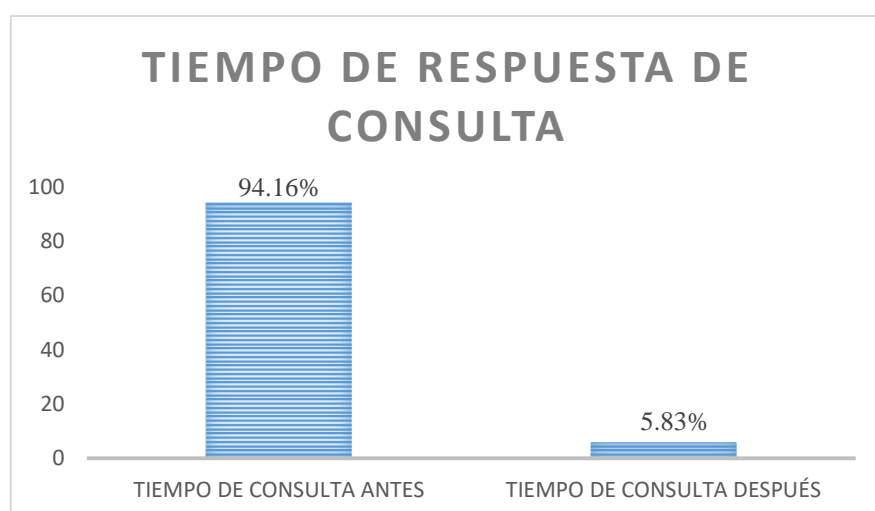


Figura 10: Tiempo de respuesta de consulta

Interpretación

Como se muestra en la figura 8, existe una diferencia entre las medias del tiempo de consulta antes de la solución de Inteligencia de negocios y después de la solución de Inteligencia de negocios. Se pudo observar que el tiempo promedio de consultas antes era 41.5 minutos que equivale a 94% y después disminuyó al 2.57 minutos que equivale a 5.83 %.

c) ISO/IEC 25012

Los atributos de calidad de modelo conceptual de datos de ISO/IEC 25012, tiene 6 características principales las cuales son: Legibilidad, completitud, corrección, minimalidad, expresividad, autoexplicación y extensibilidad. Se realizó una encuesta a expertos para determinar la calidad del modelo conceptual del aplicativo de Business Intelligence la cual se adjunta en el anexo 1, obteniendo los resultados en la tabla 14.

Para determinar la calidad del modelo conceptual, se trabajó con la escala de Likert, considerando 5 alternativas de respuesta las cuales son: Totalmente de acuerdo, de acuerdo, ni de acuerdo ni en desacuerdo, en desacuerdo, totalmente en desacuerdo, cuyas equivalencias en puntos se muestran en la siguiente tabla:

Tabla 15: Puntajes de las alternativas de respuestas de la escala de Likert

RESULTADOS	PUNTAJE
Totalmente en desacuerdo	0
En desacuerdo	1
Ni en desacuerdo ni de acuerdo	2
De acuerdo	3
Totalmente de acuerdo	4

Seguidamente, se coloca el número de las preguntas del cuestionario que se muestra en el anexo N° 2 y el número de encuestados que son 3 personas, para obtener el puntaje de cada una de sus respuestas según la escala de Likert.

Tabla 16: Puntajes de las respuestas de la calidad de datos del modelo de Inteligencia de negocios

N°	N° DE PREGUNTAS									
	1	2	3	4	5	6	7	8	9	10
1	3	3	3	3	3	3	3	3	3	2
2	3	3	3	2	2	2	3	3	3	3
3	4	4	4	3	3	3	3	4	4	3

Resultados de ISO

A continuación se muestra el resultado obtenido de la encuesta sobre calidad de uso de datos del modelo de inteligencia de negocios según la ISO 25012:

Tabla 17: Resultados de modelo de calidad - Iso 25012

RESULTADOS	Frecuencia Acumulada (FA)	%
Totalmente en desacuerdo	0	0%
En desacuerdo	0	0%
Ni en desacuerdo ni de acuerdo	4	13%
De acuerdo	21	70%
Totalmente de acuerdo	5	17%
	30	100%

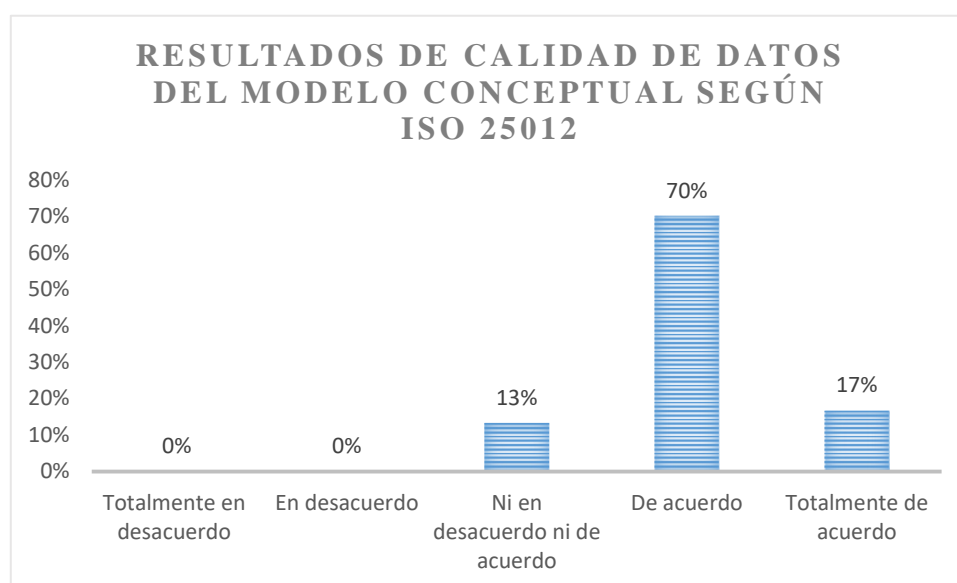


Figura 11: Resultados de calidad datos de modelo según ISO 25012

Interpretación:

En la figura 9, se muestran las 5 escalas de likert, cada uno con su frecuencia acumulada, para ello se contabilizó la cantidad de respuestas iguales por cada escala y como resultado se obtuvo que del total, de las preguntas, existe un 13 % que no está acuerdo ni en desacuerdo, un 70% que está de acuerdo y un 17% que está totalmente de acuerdo según la encuesta de calidad de modelo de inteligencia de negocios con la ISO 25012.

6.1.2. Contrastación de la hipótesis

a) Resultado de la disponibilidad de información (número de reportes)

1. Formulación de la hipótesis estadística

$H_0 : [\mu = 0]$ La disponibilidad inmediata de la información no es mayor con la solución de inteligencia de negocio usando Pentaho en el análisis para la toma de decisiones en la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac.

$H_1 : [\mu > 0]$ La disponibilidad inmediata de la información es mayor con la solución de inteligencia de negocio usando Pentaho en el análisis para la toma de decisiones en la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac.

2. Nivel de significancia

En este caso se considera el nivel de significancia $\alpha = 0.05 = 5\%$.

3. Determinación de la estadística

Si $n \geq 30$ se utiliza la “distribución normal Z”

Si $n < 30$ se utiliza la “distribución T Student”

Puesto que la disponibilidad de la información, comprende un antes y un después aplicando a una misma muestra $n = 14$, entonces se utilizó la distribución t student para muestras relacionadas.

4. Cálculo estadístico

Se aplicó la distribución T de Student de muestras relacionadas ya que se comparó las medias de dos variables de un solo grupo que contiene el número de reportes antes y después de la solución de inteligencia de negocios.

A continuación se muestra la tabla de resultados de los cálculos del número de reportes antes de la solución de Inteligencia de negocios y el número de reportes después de la solución de inteligencia de negocios.

Tabla 18: Cálculos sobre el número de reportes obtenidos antes de la solución de inteligencia de negocios y después de la solución de inteligencia de negocios

N°	x	y	$d_i = y - x$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
1	10	20	10	-0.93	0.86
2	9	20	11	0.07	0.01
3	12	20	8	-2.93	8.58
4	8	20	12	1.07	1.15
5	8	20	12	1.07	1.15
6	11	20	9	-1.93	3.72
7	7	20	13	2.07	4.29
8	6	20	14	3.07	9.43
9	8	22	14	3.07	9.43
10	9	22	13	2.07	4.29
11	7	22	15	4.07	16.58
12	10	22	12	1.07	1.15
13	2	9	7	-3.93	15.43
14	6	9	3	-7.93	62.86
	Total=113	Total=266	Total=153	Total=0	Total=138.93

Calculando la media aritmética

$$\bar{d} = \frac{\sum d_i}{n}$$

$$\bar{d} = \frac{153}{14}$$

$$\bar{d} = \mathbf{10.928}$$

Calculando la desviación típica

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n - 1}}$$

$$S_d = \sqrt{\frac{138.93}{13}}$$

$$S_d = \mathbf{3.27}$$

Calculando el error típico de la media

$$S_{\bar{d}} = \frac{S_d}{\sqrt{n}}$$

$$S_{\bar{d}} = \frac{3.27}{\sqrt{14}}$$

$$S_{\bar{d}} = \mathbf{0.874}$$

Calculando t de student de muestras relacionadas

$$t_c = \frac{\bar{d}}{S_{\bar{d}}}$$

$$t_c = \frac{10.928}{0.874}$$

$$t_c = \mathbf{12.508}$$

Donde:

y = Número de reportes después de la solución de BI

x = Número de reportes antes de la solución de BI

d_i = Diferencia por cada par de reportes

\bar{d} = Media aritmética de las diferencias

S_d = Desviación típica

$S_{\bar{d}}$ = Error estándar o típico de la media

t_c = t de student calculado

n = Tamaño de la muestra

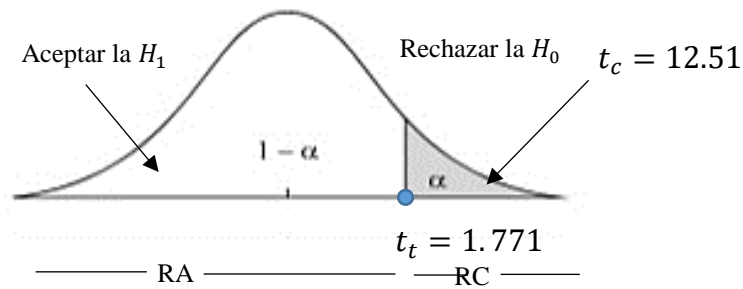
Los valores obtenidos son los mismos que se calcularon con el programa de ingeniería Minitab.

	N	Media	Desv.Est.	estándar de la media
Después	14	19.00	4.33	1.16
Antes	14	8.07	2.50	0.67
Diferencia	14	10.929	3.269	0.874

Límite inferior 95% para la diferencia media: 9.381
Prueba t de diferencia media = 0 (vs. > 0): Valor T = 12.51 Valor P = 0.000

Figura 12: Resultados de muestras relacionadas de reportes con Minitab

5. Región crítica



En caso no se tuviera el valor de probabilidad P , se halla el valor t de student de la tabla que es $t_t = 1.771$ y el valor calculado $t_c = 12.51$; entonces si $t_c \in [1.771, +\infty >$, se rechaza la hipótesis nula H_0 por lo tanto se acepta la hipótesis alterna H_1 .

Sin embargo, se observa en la figura 10 que el valor de la probabilidad de datos $P=0.000$ y el nivel de significancia es $\alpha=0.05 \Rightarrow$ si $P < \alpha$, se rechaza la hipótesis nula (H_0); por lo tanto podemos afirmar con un nivel de confianza del 95% que la disponibilidad inmediata de la información es mayor con la solución de inteligencia de negocios usando Pentaho en el análisis para la toma de decisiones en la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac; debido a que la media después de la solución de inteligencia de negocios es mayor a la media de antes ($19 > 8.07$).

b) Resultado del tiempo de consultas

1. Formulación de la hipótesis

$H_0: [\mu = 0]$ El tiempo de consulta de la información no disminuye con la solución de inteligencia de negocios usando Pentaho en la gestión académica de la UNAMBA-2016.

$H_1: [\mu < 0]$ El tiempo de consulta de la información disminuye con la solución de inteligencia de negocio usando Pentaho en la gestión académica la UNAMBA-2016.

2. Nivel de significancia

En este caso se considera el nivel de significancia $\alpha = 0.05 = 5\%$.

3. Determinación de la estadística

Si $n \geq 30$ se utiliza “distribución normal Z”

Si $n < 30$ se utiliza la “distribución T Student”

Puesto que el tiempo de consulta de la información, comprende un antes y un después aplicando a una misma muestra =14, entonces se utilizó la distribución t student para muestras relacionadas.

4. Cálculo estadístico

Se aplicó la distribución T de Student de muestras relacionadas ya que se comparó las medias de dos variables de un solo grupo que contiene el tiempo de consultas antes y después de la solución de inteligencia de negocios.

A continuación se muestra la tabla de resultados de los cálculos del tiempo de respuesta de consultas antes de la solución de Inteligencia de negocios y el tiempo de respuesta de consultas después de la solución de inteligencia de negocios.

Tabla 19: Cálculos sobre el tiempo de respuesta de consultas obtenidas antes de la inteligencia de negocios y después de la solución de inteligencia de negocios

N°	x	y	$d_i = y - x$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
1	55	2.5	-52.5	-13.57	184.18
2	50	2.5	-47.5	-8.57	73.47
3	60	2.5	-57.5	-18.57	344.90
4	45	2.5	-42.5	-3.57	12.76
5	47	2.5	-44.5	-5.57	31.04
6	58	2.5	-55.5	-16.57	274.61
7	40	2.5	-37.5	1.43	2.04
8	36	2.5	-33.5	5.43	29.07
9	30	3.2	-26.8	12.13	147.10
10	35	3.2	-31.8	7.13	50.82
11	40	3.2	-36.8	2.13	4.53
12	50	3.2	-46.8	-7.87	61.86
13	10	1.6	-8.4	30.53	931.99
14	25	1.6	-23.4	15.53	241.14
	Total=581	Total=36	Total=-545	0.00	2390.01

Calculando la media aritmética

$$\bar{d} = \frac{\sum d_i}{n}$$

$$\bar{d} = \frac{-545}{14}$$

$$\bar{d} = -38.9285714$$

Calculando la desviación típica

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n - 1}}$$

$$S_d = \sqrt{\frac{2390.01}{13}}$$

$$S_d = 13.56$$

Calculando el error típico de la media

$$S_{\bar{a}} = \frac{S_d}{\sqrt{n}}$$

$$S_{\bar{a}} = \frac{13.56}{\sqrt{14}}$$

$$S_{\bar{a}} = \mathbf{3.62}$$

Calculando t de student para muestras relacionadas

$$t_c = \frac{\bar{d}}{S_{\bar{a}}}$$

$$t_c = \frac{-38.9285714}{3.62}$$

$$t_c = \mathbf{-10.74}$$

Donde:

y = El Tiempo de consulta después de BI.

x = El Tiempo de consulta antes de BI.

d_i = Diferencia por cada par de tiempos

\bar{d} = Media aritmética de las diferencias

S_d = Desviación típica

$S_{\bar{a}}$ = Error estándar o típico de la media

t_c = T de student calculado

n = Tamaño de la muestra

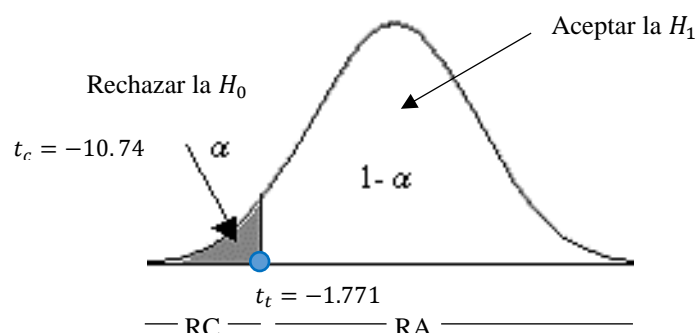
Los valores obtenidos son los mismos que se calcularon con el programa Minitab.

	N	Media	Desv.Est.	Error estándar de la media
Y	14	2.57	0.52	0.14
x	14	41.50	13.75	3.67
Diferencia	14	-38.93	13.56	3.62

Límite superior 95% para la diferencia de la media: -32.51
 Prueba t de diferencia media = 0 (vs. < 0): Valor T = -10.74 Valor P = 0.000

Figura 13: Resultados de prueba de muestras relacionadas del tiempo de consulta en Minitab

5. Región crítica



En caso no se tuviera el valor de probabilidad, se halla el valor t de student de la tabla que es $t_t = -1.771$ y el valor calculado $t_c = -10.74$; entonces si $t_c \in (-\infty, -1.771]$, se rechaza la hipótesis nula H_0 por lo tanto se acepta la hipótesis alterna H_1 .

Sin embargo, se observa en la figura 11 que el valor de la probabilidad de datos $P=0.000$ y el nivel de significancia es $\alpha=0.05 \Rightarrow$ si $P < \alpha$, se rechaza la hipótesis nula (H_0); por lo tanto podemos afirmar con un nivel de confianza del 95% que el tiempo de consulta de la información disminuye con la solución de inteligencia de negocio en la UNAMBA-2016; debido a que la media después de la solución de inteligencia de negocios es menor a la media antes ($2.51 < 41.50$).

6.2. Resultados de la implementación de la inteligencia de negocios usando Pentaho para la gestión académica de la UNAMBA.

A continuación se muestra la interfaz de inicio de sesión de la aplicación de solución de Inteligencia de negocios.



User Console

User Name: Password:

Figura 14: Acceso principal al modelado de negocio

Seguidamente, se muestran los reportes principales que responden a las preguntas o requerimientos de negocio:

Cubo Asignatura: Se tienen los siguientes reportes obtenidos del cubo de asignatura.

Available fields (9) for: Asignaturas

Find: View ▾

- ▼ Asignatura
 - 🔍 Codigo_Asignatura
 - 🔍 Asignatura
- ▼ CATEGORIA.Categoria
 - 🔍 Categoria
- ▼ Escuela
 - 🔍 Escuela
- ▼ Measures
 - 🔍 Matriculados
 - 🔍 Nro. Aprobados
 - 🔍 Nro. Desaprobados
- ▼ Tiempo
 - 🔍 Año
 - 🔍 Semestre

Layout

Rows

- ▼ Codigo_Asignatura
- ▼ Asignatura
- ▼ Semestre
- Drop Level Here

Columns

- ▼ Escuela
- Drop Level Here

Measures

- ▼ Matriculados
- ▼ Nro. Aprobados
- ▼ Nro. Desaprobados
- Drop Measure Here

			Escuela		
			INGENIERÍA INFORMÁTICA Y SISTEMAS		
Codigo_Asignatu...	Asignatura	Semestre	Matriculados	Nro. Aprobados	Nro. Desaprobados
IS304	Filosofia	2008-1	46	40	6
		2008-2	40	29	11
		2009-1	43	37	6
		2009-2	44	31	13
		2010-1	28	21	7
		2010-2	37	29	8
		2011-1	33	23	10
		2011-2	38	25	13
		2012-1	45	39	6
		2012-2	40	31	9

Figura 15: Reporte de número de matriculados, aprobados y desaprobados por asignatura y semestre

En la figura 13, se observa la cantidad de matriculados, cantidad de aprobados y desaprobadados de la escuela académico profesional de Ingeniería Informática de Sistemas, por cada asignatura desde el semestre académico 2008-1 al 2012-2.

Escuela			
INGENIERÍA INFORMÁTICA Y SISTEMAS			
Asignatura	Año	Semestre	Nro. Aprobados
Algorítmica I	2008	2008-1	41
		2008-2	37
	2009	2009-1	27
		2009-2	34
	2010	2010-1	27
		2010-2	28
	2011	2011-1	29
		2011-2	33
	2012	2012-1	34
		2012-2	39
	2013	2013-1	37
		2013-2	14
	2014	2014-1	22
		2014-2	32
	2015	2015-1	19
		2015-2	32
2016	2016-1	12	
Algorítmica II	2008	2008-1	26
		2008-2	32
	2009	2009-1	42
		2009-2	28
	2010	2010-1	26
		2010-2	17
	2011	2011-1	25
		2011-2	26
		2011-3	11
	2012	2012-1	18
		2012-2	24
		2012-3	11
	2013	2013-1	15
		2013-2	33
	2014	2014-1	19
		2014-2	26
	2015	2015-1	33
		2015-2	16
2016	2016-1	33	

Figura 16: Cantidad de aprobados por asignaturas según semestre académico

En la figura 14, se observa la cantidad de estudiantes matriculados que aprobaron cada una de las asignaturas por cada semestre académico a partir del semestre 2008-I al 2016-I de la escuela académica profesional de Ingeniería Informática y Sistemas.

			Escuela
			INGENIERÍA INFORMÁTICA Y SISTEMAS
Asignatura	Año	Semestre	Nro. Desaprobados
Algorítmica II	2008	2008-1	16
		2008-2	22
	2009	2009-1	15
		2009-2	9
	2010	2010-1	17
		2010-2	20
	2011	2011-1	15
		2011-2	19
		2011-3	14
	2012	2012-1	14
		2012-2	18
		2012-3	11
	2013	2013-1	29
		2013-2	20
	2014	2014-1	14
		2014-2	11
	2015	2015-1	5
		2015-2	10
2016	2016-1	1	

Figura 17: Cantidad de estudiantes desaprobados por asignatura, según el semestre académico

En la figura 15, se muestra el número de estudiantes que han desaprobado alguna asignatura por cada semestre académico desde el 2008-I al 2016-I de cada escuela académica profesional.

Asignatura	Semestre								
	2008-1	2008-2	2008-3	2009-1	2009-2	2009-3	2010-1	2010-2	2010-3
	Matriculados	Matriculados	Matriculados	Matriculados	Matriculados	Matriculados	Matriculados	Matriculados	Matriculados
Actividades	-	-	-	-	6	-	-	-	-
Administración Moderna	44	34	-	36	41	-	22	24	-
Administración de Tecnologías de ...	-	-	-	-	-	-	-	-	-
Algorítmica I	64	56	-	66	65	-	66	58	-
Algorítmica II	42	54	-	57	37	-	43	37	-
Algorítmica III	31	33	-	40	45	-	35	29	-
Análisis y Diseño de Sistemas de I...	-	26	-	-	29	-	41	20	-
Análisis y Diseño de Sistemas de I...	12	16	-	28	45	-	27	31	-
Arquitectura del Computador	-	-	-	15	35	-	42	25	-
Calculo I	48	65	26	48	27	24	37	35	20
Calculo II	10	27	30	42	27	25	24	18	7
Circuitos Electricos y Electronicos	-	-	-	-	-	-	35	34	-
Comercio Electrónico	-	-	-	-	-	-	-	-	-
Computación Gráfica	-	-	-	-	-	-	-	-	-
Desarrollo de Sistemas Orientado...	-	-	-	-	-	-	-	-	-
Ecología	45	49	-	47	40	-	40	48	-
Ecuaciones Diferenciales	28	19	-	30	21	28	19	-	30
Ensamblaje y Mantenimiento de C...	-	-	-	-	48	-	-	-	-
Estadística Descriptiva	-	37	46	-	22	-	37	46	-
Estadística Inferencial	-	16	12	-	31	-	16	12	-
Filosofía	-	28	37	-	33	-	28	37	-
Formulacion y Evaluacion de Proy...	-	-	33	-	37	-	-	33	-
Física I	46	28	34	14	38	46	28	34	14
Física II	10	40	20	-	21	10	40	20	-
Geografía y Recursos Naturales	-	65	67	-	64	-	65	67	-

Figura 18: Número de matriculados por asignatura y semestre académico

En la figura 16, se muestra la cantidad de estudiantes matriculados por cada Asignatura desde el semestre académico 2008-I al 2016-I.

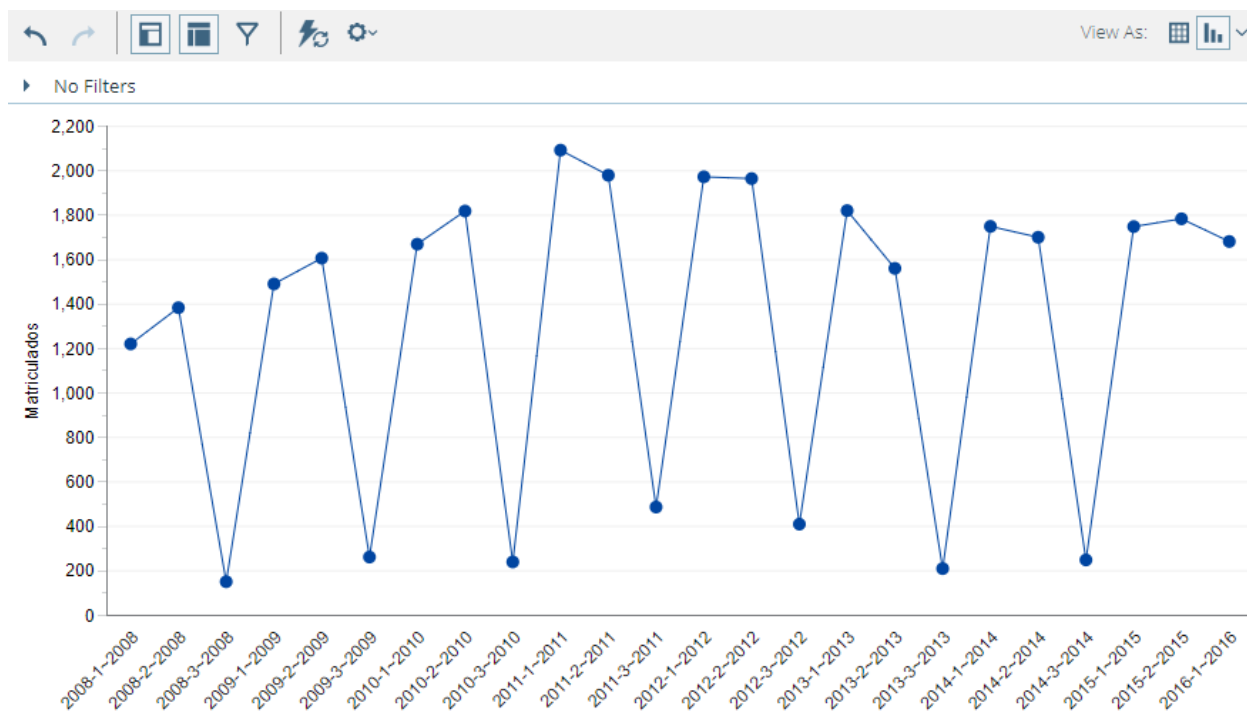


Figura 19: Gráfico estadístico de cantidad de matriculados por semestre académico

En la figura 17 desde el semestre 2008-1 al 2016-1 se observa la tendencia de la cantidad de estudiantes que han matriculado por cada semestre académico en la Universidad Nacional Micaela Bastidas de Apurímac.

Cubo Docentes: Se tienen los siguientes reportes obtenidos del cubo de docentes.

Escuela	Categoria			
	ASOCIADO	AUXILIAR	JEFE DE PRÁCTICAS	PRINCIPAL
Total	Total	Total	Total	Total
ADMINISTRACIÓN DE EMPRESAS	12	22	5	2
EDUCACIÓN INICIAL INTERCULTU...	-	-	-	1
EDUCACIÓN: ESPECIALIDAD EDUC...	6	7	-	-
EDUCACIÓN: ESPECIALIDAD MATE...	11	5	-	-
INGENIERIA CIVIL (COTABAMBAS)	-	6	-	-
INGENIERÍA AGROECOLÓGICA Y D...	-	7	-	-
INGENIERÍA AGROINDUSTRIAL	9	16	5	1
INGENIERÍA DE MINAS	11	14	5	2
INGENIERÍA DE MINAS (COTABAM...	-	2	-	-
INGENIERÍA INFORMÁTICA Y SISTE...	9	9	1	-
MEDICINA VETERINARIA Y ZOOTE...	1	21	1	1

Figura 20: Total de docentes de cada escuela académico profesional por categoría

En la figura 18, se muestra el total de docentes que hay por cada escuela académica profesional según las categorías de asociado, auxiliar y principal; también se muestra a los jefes de práctica.

Facultad	Escuela	Sexo	
		F	M
		Total	Total
FACULTAD AGROECOLOGÍA Y DESARROLLO RURAL	INGENIERÍA AGROECOLÓGICA Y DESARROLLO R...	6	1
FACULTAD DE ADMINISTRACION	ADMINISTRACIÓN DE EMPRESAS	11	30
FACULTAD DE EDUCACION	EDUCACIÓN INICIAL INTERCULTURAL BILINGÜE	1	-
	EDUCACIÓN: ESPECIALIDAD EDUCACIÓN FÍSICA ...	4	9
	EDUCACIÓN: ESPECIALIDAD MATEMÁTICA E INF...	2	14
FACULTAD DE INGENIERIA	INGENIERIA CIVIL (COTABAMBAS)	6	-
	INGENIERÍA AGROINDUSTRIAL	7	24
	INGENIERÍA DE MINAS	2	30
	INGENIERÍA DE MINAS (COTABAMBAS)	2	-
	INGENIERÍA INFORMÁTICA Y SISTEMAS	6	13
FACULTAD DE MEDICINA VETERINARIA Y ZOOTECNIA	MEDICINA VETERINARIA Y ZOOTECNIA	6	18

Figura 21: Total de docentes por género según facultad y escuela académico profesional

En la figura 19, se muestra la cantidad de docentes por sexo femenino y masculino de cada escuela académico profesional de la UNAMBA y según la facultad.

Cubo grados y títulos

Escuela	Modalidad	Total
EDUCACIÓN: ESPECIALIDAD EDUC...	AUTOMÁTICO	201
	EXAMEN DE CAPACIDAD PROFESI...	73
	SUSTENTACIÓN DE TESIS	102
EDUCACIÓN: ESPECIALIDAD MATE...	AUTOMÁTICO	148
	EXAMEN DE CAPACIDAD PROFESI...	43
	SUSTENTACIÓN DE TESIS	51
INGENIERÍA AGROECOLÓGICA Y D...	AUTOMÁTICO	36
	AUTOMÁTICO	411
INGENIERÍA AGROINDUSTRIAL	AUTOMÁTICO	411
	EXAMEN DE CAPACIDAD PROFESI...	266
	SUSTENTACIÓN DE TESIS	36
INGENIERÍA DE MINAS	AUTOMÁTICO	140
	EXAMEN DE CAPACIDAD PROFESI...	24
	SUSTENTACIÓN DE TESIS	23
INGENIERÍA INFORMÁTICA Y SISTE...	AUTOMÁTICO	115
	EXAMEN DE CAPACIDAD PROFESI...	30
	SUSTENTACIÓN DE TESIS	6
MEDICINA VETERINARIA Y ZOOTE...	AUTOMÁTICO	121
	EXAMEN DE CAPACIDAD PROFESI...	15
	SUSTENTACIÓN DE TESIS	23

Figura 22: Total de graduados por escuela académica profesional de la UNAMBA según la modalidad

En la figura 20, se muestra a la cantidad de graduados por cada escuela académico profesional de la UNAMBA, según su modalidad

		Escuela				
		ADMINISTRACIÓN DE EMPRESAS	EDUCACIÓN: ESPECIALIDAD EDUCACIÓN FÍSICA Y DANZAS	EDUCACIÓN: ESPECIALIDAD MATEMÁTICA E INFORMÁTICA	INGENIERÍA AGROECOLÓGICA Y DESARROLLO RURAL	INGENIERÍA AGROINDUSTRIAL
Sexo	Grado	Total	Total	Total	Total	Total
F	BACHILLER	247	100	49	13	166
	TITULO	191	88	25	-	110
M	BACHILLER	181	101	99	23	245
	TITULO	148	87	69	-	192

Figura 23: Total de estudiantes por sexo y grado según cada escuela académico profesional

En la imagen 21, se observa a la cantidad de estudiantes según el sexo femenino o masculino que tengan el grado de bachiller o titulado según cada escuela académica profesional de la UNAMBA

Cubo de población universitaria: Se tienen los siguientes reportes obtenidos del cubo de docentes.

		Escuela
		INGENIERÍA INFORMÁTICA Y SISTEMAS
Año	Semestre	Egresados
2011	2011-1	0
	2011-2	3
	2011-3	0
2012	2012-1	9
	2012-2	11
	2012-3	0
2013	2013-1	15
	2013-2	8
	2013-3	0
2014	2014-1	17
	2014-2	11
	2014-3	0
2015	2015-1	20
	2015-2	11
2016	2016-1	2

Figura 24: Total de egresados por semestre académico

En la imagen 22, se tiene el total de egresados por cada semestre académico desde el 2008-I al 2016-I de la escuela académico profesional de Ingeniería Informática y Sistemas de la UNAMBA

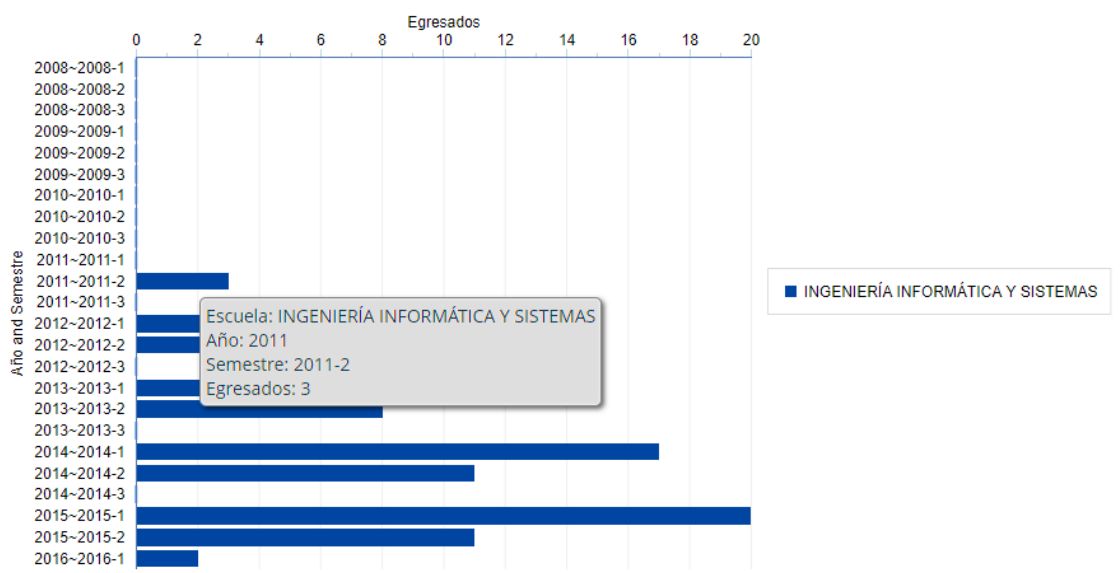


Figura 25: Gráfico estadístico de la cantidad de egresados por semestre

		Escuela
		INGENIERÍA INFORMÁTICA Y SISTEMAS
Año	Sexo	Egresados
2008	F	0
	M	0
2009	F	0
	M	0
2010	F	0
	M	0
2011	F	1
	M	2
2012	F	11
	M	9
2013	F	7
	M	16
2014	F	13
	M	15

Figura 26: Número de egresados por género y año

Cubo seguimiento de matriculados e ingresantes

		Escuela
		INGENIERÍA INFORMÁTICA Y SISTEMAS
Año	Semestre	Nro_ingresantes
2008	2008-1	49
	2008-2	50
2009	2009-1	50
	2009-2	50
2010	2010-1	51
	2010-2	50
2011	2011-1	50
	2011-2	41
2012	2012-1	52
	2012-2	50
2013	2013-1	34
	2013-2	15
2014	2014-1	51
	2014-2	49

Figura 27: Numero de ingresantes por semestre académico

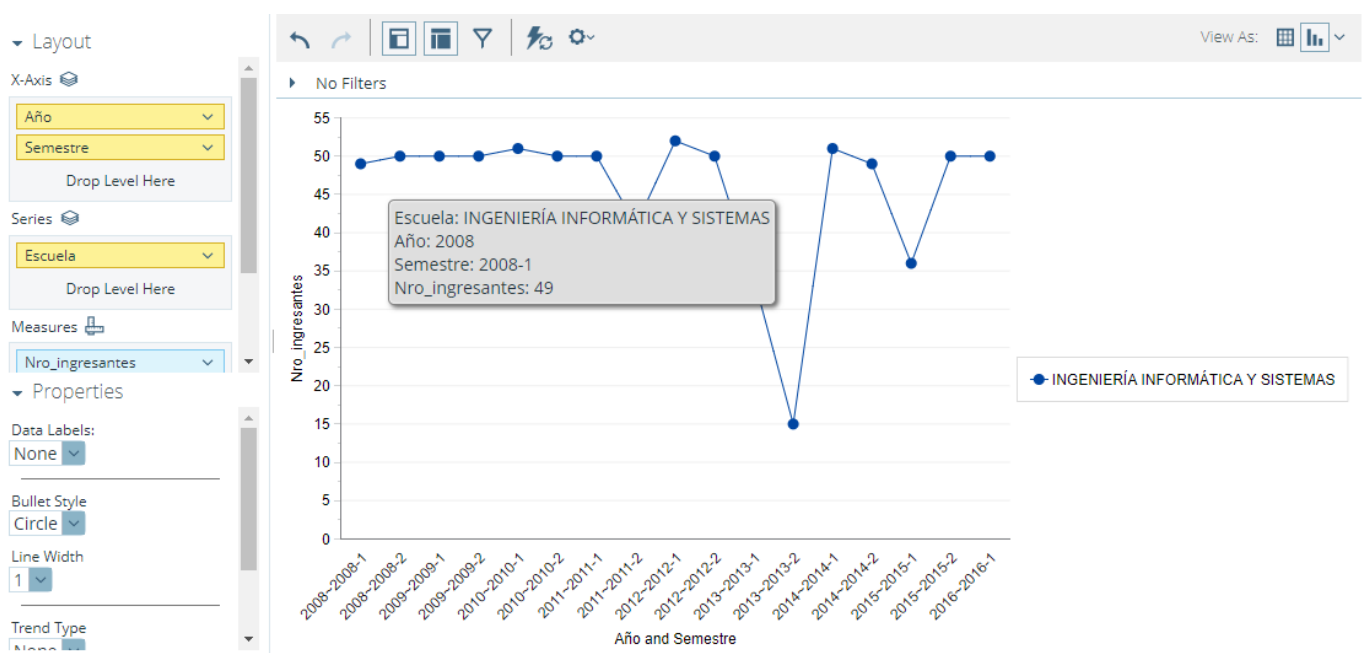


Figura 28: Gráfico estadístico de número de ingresantes por semestre académico desde el 2008-I al 2016I

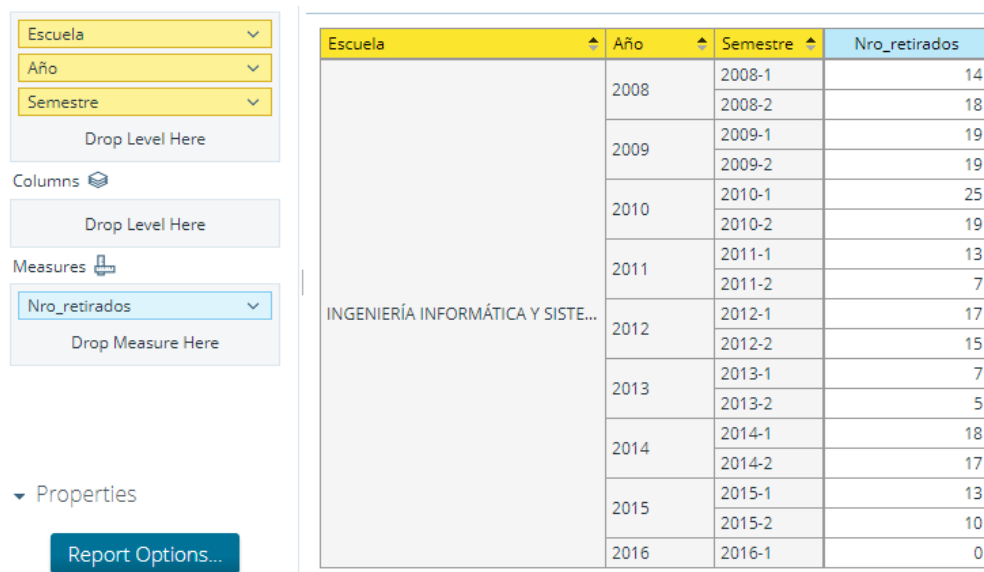


Figura 29: Número de retirados por escuela académico profesional y semestre académico

6.3. Implementación de la inteligencia de negocios

Las herramientas utilizadas para la implementación de la inteligencia de negocios en la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac fueron las que se muestran en el siguiente cuadro:

Tabla 20: Herramientas utilizadas para la implementación de inteligencia de negocios

PROCESOS	HERRAMIENTA
Plataforma	Pentaho
Proceso ETL	Pentaho Data Integration
Construcción de cubos	Pentaho schema workbench
Visualización de cubos	Pentaho bi server
Gestor de base de Datos	Mysql

6.4. Implementación del Datamart usando la metodología Hefesto

6.4.1. Análisis de requerimientos

Identificar preguntas

Tabla 21: Identificación de preguntas

Requerimientos	Preguntas
R1	¿Cuál es el número de estudiantes matriculados por sexo, tiempo, escuela académico profesional, facultad, y asignatura?
R2	¿Cuál es el número de estudiantes aprobados por sexo, tiempo, escuela académico profesional, facultad, asignatura?
R3	¿Cuál es el número de estudiantes desaprobados por sexo, tiempo, escuela académico profesional, facultad, asignatura?
R4	¿Cuál es el número de estudiantes desertores por sexo, tiempo, escuela académico profesional, facultad, asignatura?
R5	¿Cuál es el total de semestres estudiados por estudiante, sexo, tiempo, por facultad y escuela académico profesional?
R6	¿Cuál es el total de créditos aprobados por sexo, tiempo, estudiante, facultad y escuela académico profesional?
R7	¿Cuál es el total de créditos desaprobados por sexo, tiempo, estudiante, facultad y escuela académico profesional?
R8	¿Cuál es el total de asignaturas por sexo, tiempo, por estudiantes, por facultad, escuela académico profesional?
R9	¿Cuál es el total de asignaturas aprobadas por sexo, tiempo, estudiantes, facultad, escuela académico profesional?
R10	¿Cuál es el total de asignaturas desaprobadas por sexo, tiempo, por estudiantes, por facultad, por escuela?
R11	¿Cuál es el promedio ponderado por sexo, tiempo, por estudiantes, por facultad, escuela académico profesional?
R12	¿Cuál es el número de docentes que hay por cada escuela académico profesional, por sexo, tiempo, facultad?
R13	¿Cuál es el total de bachilleres por sexo, tiempo, facultad, tipo de acto, escuela y modalidad?
R14	¿Cuál es el total de titulados por sexo, tiempo, facultad, tipo de acto, escuela y modalidad?
R15	¿Cuál es el número de ingresantes por sexo, tiempo, facultad y escuela?
R16	¿Cuál es el número de matriculados por sexo, tiempo, facultad y escuela?
R17	¿Cuál es el número de egresados por sexo, tiempo, facultad y escuela?

a. Identificar indicadores y perspectivas.

Se identificó los indicadores y sus perspectivas para la gestión académica de la Universidad Nacional Micaela Bastidas de Apurímac, sede Abancay de la siguiente manera:

- *Número de estudiantes* matriculados por semestre académico según el sexo, escuela académico profesional, facultad y por asignatura, en la UNAMBA, sede Abancay.
- *Número de estudiantes* aprobados por sexo, semestre, escuela académico profesional, facultad, por asignatura, en la UNAMBA, sede Abancay.
- *Número de estudiantes* desaprobados por sexo, tiempo, escuela académico profesional, facultad, asignatura en la UNAMBA, sede Abancay.
- *Número de estudiantes* desertores por sexo, tiempo, escuela académico profesional, facultad, asignatura en la UNAMBA, sede Abancay.
- *Total de semestres* estudiados por estudiante, según sexo, tiempo, por facultad y escuela académico profesional en la UNAMBA, sede Abancay
- *Total de créditos* aprobados por estudiante según sexo, tiempo, facultad y escuela académico profesional en la UNAMBA, sede Abancay
- *Total de créditos* desaprobados por estudiante según sexo, tiempo, facultad y escuela académico profesional en la UNAMBA, sede Abancay

- ***Total de asignaturas*** por estudiantes, sexo, tiempo, facultad, escuela académico profesional en la UNAMBA, sede Abancay.
- ***Total de asignaturas*** aprobadas, por estudiantes, sexo, tiempo, facultad, escuela académico profesional en la UNAMBA, sede Abancay.
- ***Total de asignaturas*** desaprobadas por estudiantes, sexo, tiempo, facultad, escuela académico profesional en la UNAMBA, sede Abancay.
- ***Promedio ponderado*** por estudiantes, sexo, tiempo, facultad, escuela académico profesional en la UNAMBA, sede Abancay.
- ***Número de docentes*** por sexo, tiempo, escuela académico profesional en la UNAMBA, sede Abancay.
- ***Total de bachilleres*** por sexo, tiempo, facultad, tipo de acto, escuela académico profesional y modalidad en la UNAMBA, sede Abancay.
- ***Número de ingresantes*** por sexo, tiempo, facultad y escuela académico profesional en la UNAMBA, sede Abancay.
- ***Número de matriculados*** por sexo, tiempo, facultad y escuela académico profesional en la UNAMBA, sede Abancay.
- ***Número de egresados*** por sexo, tiempo, facultad y escuela académico profesional en la UNAMBA, sede Abancay.

Finalmente se tiene el siguiente cuadro resumen:

Tabla 22: Indicadores y perspectivas

Preguntas	Indicadores	Perspectivas
Número de estudiantes matriculados por sexo, tiempo, escuela académico profesional, facultad, y asignatura en la UNAMBA.	<i>Número de estudiantes matriculados</i>	Matriculados, sexo, semestre, escuela académico profesional, facultad, por asignatura.
Número de estudiantes aprobados por sexo, tiempo, escuela académico profesional, facultad, asignatura en la UNAMBA.	<i>Número de estudiantes aprobados</i>	Matriculados, sexo, semestre, escuela académico profesional, facultad, por asignatura.
Número de estudiantes desaprobados por sexo, tiempo, escuela académico profesional, facultad, asignatura en la UNAMBA	<i>Número de estudiantes desaprobados</i>	Matriculados, sexo, semestre, escuela académico profesional, facultad, por asignatura,
Número de estudiantes desertores por sexo, tiempo, escuela académico profesional, facultad, asignatura en la UNAMBA	<i>Número de estudiantes desertores</i>	Matriculados, sexo, semestre, escuela académico profesional, facultad, por asignatura,
Total de semestres estudiados por estudiante, según sexo, tiempo, por facultad y escuela	<i>Total de semestres estudiados</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Total de créditos aprobados por sexo, tiempo, estudiante, facultad y escuela.	<i>Total de créditos aprobados</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Total de créditos desaprobados por sexo, tiempo, estudiante, facultad y escuela.	<i>Total de créditos desaprobados</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Total de asignaturas por sexo, tiempo, por estudiantes, por facultad, por escuela	<i>Total de asignaturas</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Total de asignaturas aprobadas por sexo, tiempo, por estudiantes, por facultad, por escuela	<i>Total de asignaturas aprobadas</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Total de asignaturas desaprobadas por sexo, tiempo, por estudiantes, por facultad, por escuela	<i>Total de asignaturas desaprobadas</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Promedio ponderado por sexo, tiempo, por estudiantes, por facultad, por escuela	<i>Promedio ponderado</i>	Seguimiento académico, Sexo, tiempo, estudiante, facultad, escuela
Número docentes hay por cada escuela académico profesional en la UNAMBA	<i>Total de docentes</i>	Docentes, facultad, tiempo, categoría, grado académico, condición laboral, dedicación, departamento académico, sexo.
Total de bachilleres por sexo, tiempo, facultad, tipo de acto, escuela y modalidad	<i>Total de bachilleres</i>	Grados y títulos, por sexo, tiempo, facultad, tipo de acto, escuela y modalidad
Total de titulados por sexo, tiempo, facultad, tipo de acto, escuela y modalidad.	<i>Total de Titulados</i>	Grados y títulos, por sexo, tiempo, facultad, tipo de acto, escuela y modalidad
Número de ingresantes por sexo, tiempo, facultad y escuela	<i>Número de ingresantes</i>	Población universitaria, por sexo, tiempo, facultad y escuela
Número de matriculados por sexo, tiempo, facultad y escuela	<i>Número de matriculados</i>	Población universitaria, por sexo, tiempo, facultad y escuela
Número de egresados por sexo, tiempo, facultad y escuela	<i>Número de egresados</i>	Población universitaria, por sexo, tiempo, facultad y escuela
Número de docentes por sexo, tiempo, facultad y escuela	<i>Número de docentes</i>	Población universitaria, por sexo, tiempo, facultad y escuela

b. Modelo conceptual

Se realizó el modelo conceptual para poder observar de mejor manera los indicadores y perspectivas analizadas acerca de la gestión académica de la UNAMBA que contempla el área de servicios académicos, grados y títulos; los cuales se muestran a continuación:

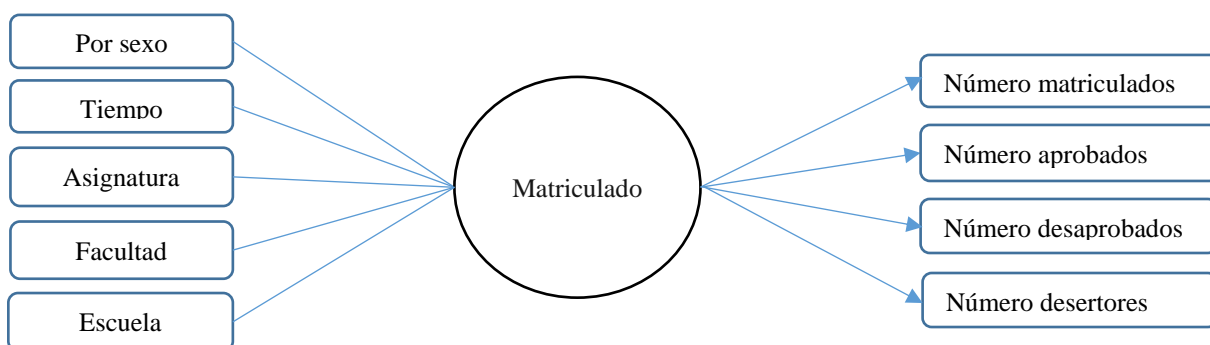


Figura 30: Modelado conceptual de Matriculado

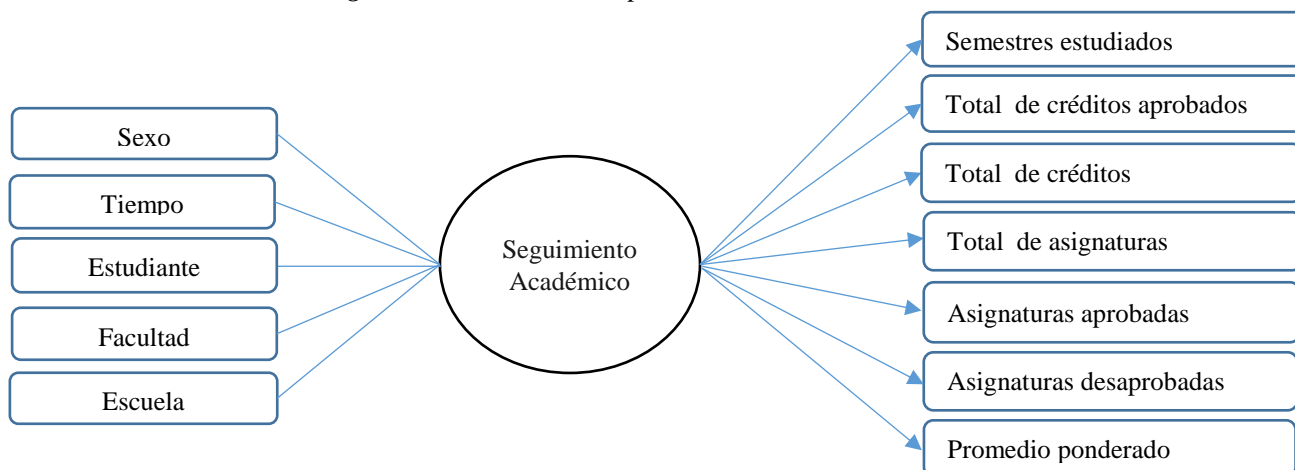


Figura 31: Modelo conceptual de seguimiento académico

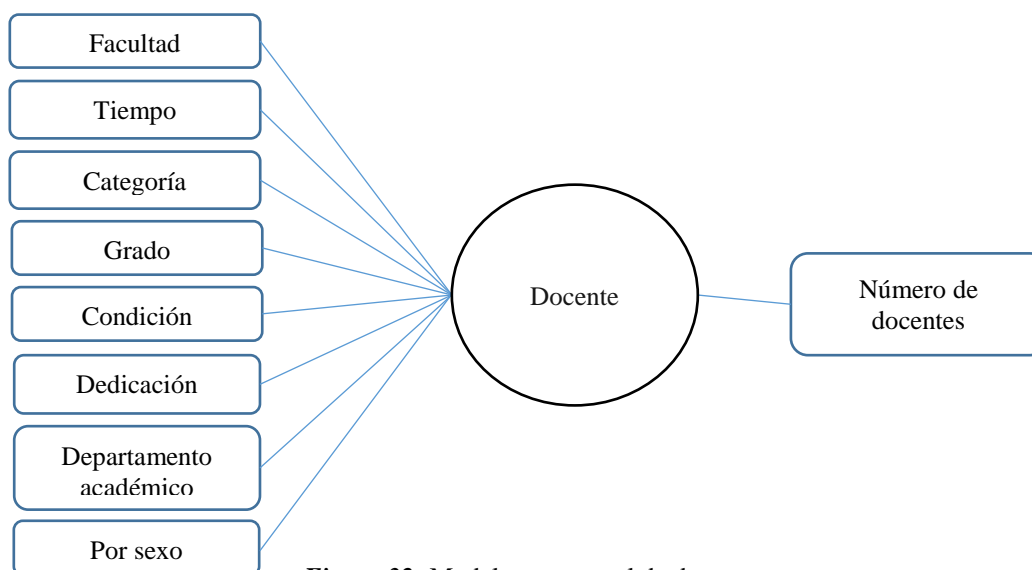


Figura 32: Modelo conceptual de docente

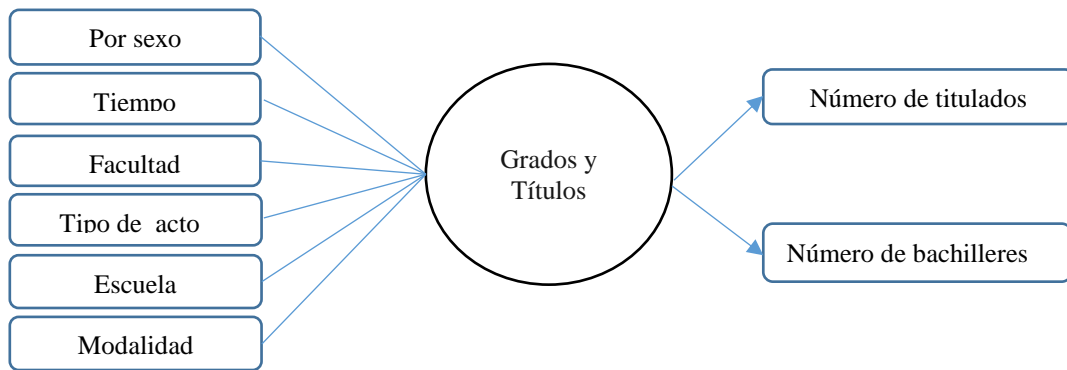


Figura 33: Modelo conceptual de grados y títulos

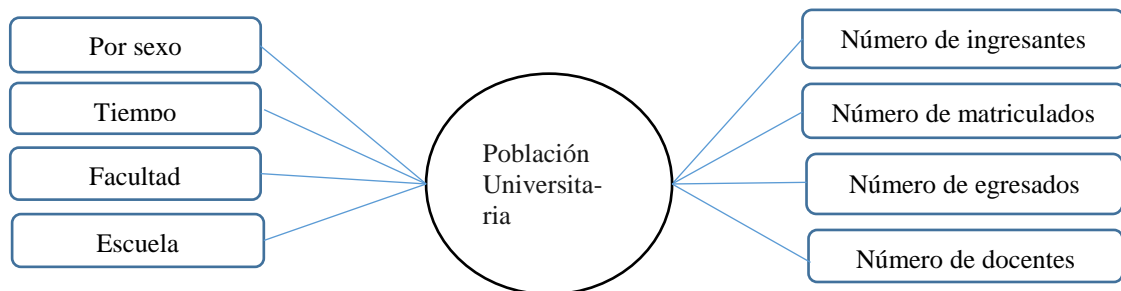


Figura 34: Modelo conceptual de Población Universitaria

6.4.2. Análisis OLTP

a. Conformar indicadores

- *“Número de estudiantes matriculados”*
 - Hechos: Estudiantes matriculados
 - Función: Count
 - Aclaración: El indicador “Número de estudiantes matriculados” representa contar los estudiantes matriculados por sexo, tiempo, asignatura, facultad y escuela académico profesional.

- ***“Número de estudiantes aprobados”***
 - Hechos: Estudiantes aprobados
 - Función: Count
 - Aclaración: El indicador “Número de estudiantes aprobados” representa contabilizar la cantidad de estudiantes matriculados que aprobaron por sexo, asignatura, tiempo, facultad y escuela académico profesional.

- ***“Número de estudiantes desaprobados”***
 - Hechos: Estudiantes desaprobados
 - Función: Count
 - Aclaración: El indicador “Número de estudiantes desaprobados” representa contar la cantidad de estudiantes que desaprobaron según el sexo, asignatura, tiempo, facultad y escuela académico profesional.

- ***“Número de estudiantes desertores”***
 - Hechos: Estudiantes desertores
 - Función: Count
 - Aclaración: El indicador “Número de estudiantes desertores” representa contar la cantidad de estudiantes que desaprobaron según el sexo, asignatura, tiempo, facultad y escuela académico profesional.

- ***“Número de semestres estudiados”***
 - Hechos: Semestres estudiados
 - Función: Count
 - Aclaración: El indicador “Número de semestres estudiados” ” representa contabilizar los créditos que se hayan aprobado por el tipo de sexo, tiempo, estudiante, facultad, escuela para realizar el seguimiento académico.

- ***“Total de créditos aprobados”***
 - Hechos: Créditos aprobados
 - Función: Count
 - Aclaración: El indicador “Total de créditos aprobados” representa contabilizar los créditos que se hayan aprobado por el tipo de sexo, tiempo, estudiante, facultad, escuela para realizar el seguimiento académico.

- ***“Total de créditos desaprobados”***
 - Hechos: Créditos desaprobados
 - Función: Count
 - Aclaración: El indicador “Total de créditos desaprobados” representa el contabilizar los créditos que se hayan aprobado por el tipo de sexo, tiempo, estudiante, facultad, escuela para realizar el seguimiento académico.

- ***“Total de asignaturas”***
 - Hechos: Total de asignaturas
 - Función: Count
 - Aclaración: El indicador “Total de asignaturas” representa contabilizar el total de asignaturas por el tipo de sexo, tiempo, estudiante, facultad, escuela para realizar el seguimiento académico.

- ***“Asignaturas aprobadas”***
 - Hechos: Asignaturas aprobadas
 - Función: Count
 - Aclaración: El indicador “Asignaturas aprobadas” representa el total de asignaturas aprobadas por tipo de sexo, tiempo, estudiante, facultad y escuela académico profesional para realizar el seguimiento académico.

- ***“Asignaturas desaprobadas”***
 - Hechos: Asignaturas desaprobadas
 - Función: count
 - Aclaración: El indicador “Asignaturas desaprobadas” representa contar el total de asignaturas por tipo de sexo, tiempo, estudiante, facultad y escuela académico profesional para realizar el seguimiento académico.

- ***“Promedio ponderado”***
 - Hechos: Promedio ponderado
 - Función: SUM
 - Aclaración: El indicador “Promedio ponderado” representa la sumatoria para obtener el promedio ponderado por tipo de sexo, tiempo, estudiante, facultad y escuela académico profesional para realizar el seguimiento académico.

- ***“Número de docentes”***
 - Hechos: Número de docentes
 - Función: COUNT
 - Aclaración: El indicador “Número de docentes” representa el total de número de docentes por facultad, tiempo, categoría grado, condición dedicación, departamento académico y sexo.

- ***“Número de bachilleres”***
 - Hechos: Número de bachilleres
 - Función: Count
 - Aclaración: El indicador “Número de bachilleres” representa contar el total de bachilleres por sexo, tiempo, facultad, tipo de acto, escuela académico profesional y modalidad de grados y títulos.

- ***“Número de titulados”***
 - Hechos: Número de titulados
 - Función: Count
 - Aclaración: El indicador “Número de titulados” representa contar el total de titulados por sexo, tiempo, facultad, tipo de acto, escuela académico profesional y modalidad de grados y títulos.

- ***“Número de ingresantes”***
 - Hechos: Número de ingresantes
 - Función: Count
 - Aclaración: El indicador “Número de ingresantes” representa contar el total de ingresantes por sexo, tiempo, facultad y escuela académico profesional.
- ***“Número de matriculados”***
 - Hechos: Número de matriculados
 - Función: Count
 - Aclaración: El indicador “Número de matriculados” representa contar el total de matriculados por sexo, tiempo, facultad y escuela académico profesional.
- ***“Número de egresados”***
 - Hechos: Número de egresados
 - Función: Count
 - Aclaración: El indicador “Número de egresados” representa contar el total de egresados por sexo, tiempo, facultad y escuela académico profesional.

b. Establecer correspondencias

Se identificó las correspondencias entre el modelo conceptual y la fuente de datos para lo cual se tuvo que extraer la información que contienen los OLPT disponibles de acuerdo a los requerimientos.

La oficina de Servicios Académicos de la UNAMBA, posee un sistema transaccional de matrícula, de los cuales se obtuvo los datos a partir del año 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, y 2016, de la gestión académica de los estudiantes. También se obtuvo información de un Excel proporcionado de la oficina de grados y títulos a partir del año 2009 al 2014.

Tabla 23: Fuentes de datos del sistema de matrícula-gradros y títulos

FUENTE	DENOMINACION DE LA FUENTE	DESCRIPCIÓN
BD Transaccional	Base de datos del sistema de matrícula de la UNAMBA	La base de datos contiene proporcionada contiene datos a partir del año 2014 (semestre académico 2014-I), al 2016 (Semestre académico 2016-I)
Hoja de cálculo de Excel	Archivo Excel grados y títulos	Está hoja contiene datos de grados y títulos a partir del año 2009 al 2014.

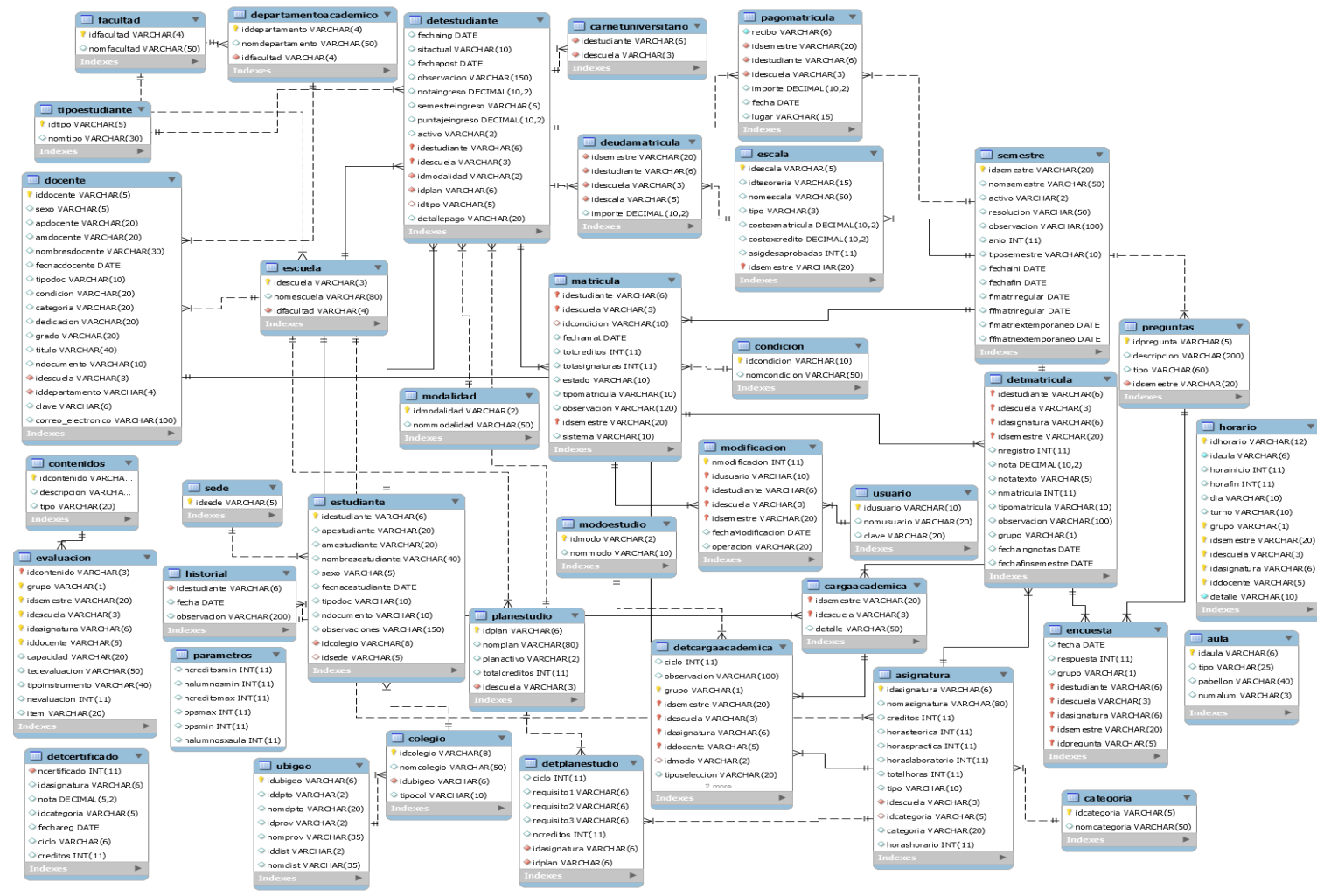


Figura 35: Modelado de la base de datos relacional de servicios académicos
Fuente: Oficina de servicios académicos de la UNAMBA

Las relaciones identificadas fueron las siguientes:

a. Para Matriculados

- El campo sexo de la tabla “Estudiante” se relaciona con la perspectiva “Sexo” (debido a que es el atributo importante para el proceso de matriculados).
- La tabla “tiempo” se relaciona con la perspectiva “Semestre”
- La tabla “asignatura” se relaciona con la perspectiva “Curso”
- La tabla “facultad” se relaciona con la perspectiva “Facultad”
- La tabla “escuela” se relaciona con la perspectiva “Escuela”

b. Para Seguimiento académico

- El campo sexo de la tabla “Estudiante” se relaciona con la perspectiva “Sexo” (debido a que es el atributo importante para el proceso de matriculados).
- La tabla “semestre” se relaciona con la perspectiva “Semestre”
- La tabla “estudiante” se relaciona con la perspectiva “Estudiante”.
- La tabla “facultad” se relaciona con la perspectiva “Facultad”.
- La tabla “escuela” se relaciona con la perspectiva “Escuela”.

c. Para Docente

- La tabla “facultad” se relaciona con la perspectiva “Facultad”
- La tabla “tiempo” se relaciona con la perspectiva “Semestre”
- La tabla “categoría” se relaciona con la perspectiva “Categoría”
- La tabla grado se relaciona con el atributo “Grado Académico”
- La tabla “condición” se relaciona con la perspectiva “Condición”
- La tabla departamento académico se relaciona con la perspectiva “Departamento Académico”

- El campo sexo de la tabla “Estudiante” se relaciona con la perspectiva “Sexo” (debido a que es el atributo importante para el proceso de matriculados).

d. Para grados y títulos

- El campo sexo de la tabla “Estudiante” se relaciona con la perspectiva “Sexo” (debido a que es el atributo importante para el proceso de matriculados).
- La tabla “tiempo” se relaciona con perspectiva “Semestre”
- La tabla “facultad” se relaciona con la perspectiva “Facultad”
- La tabla “tipo de acto” se relaciona con la perspectiva “Acto”
- La tabla modalidad se relaciona con perspectiva “Modalidad”

e. Para población universitaria

- El campo sexo de la tabla “Estudiante” se relaciona con la perspectiva “Sexo” (debido a que es el atributo importante para el proceso de matriculados).
- La tabla “tiempo” se relaciona con perspectiva “Semestre”
- La tabla “facultad” se relaciona con la perspectiva “Facultad”
- La tabla “escuela” se relaciona con la perspectiva “Escuela”.

6.4.3. Modelo lógico del datamart

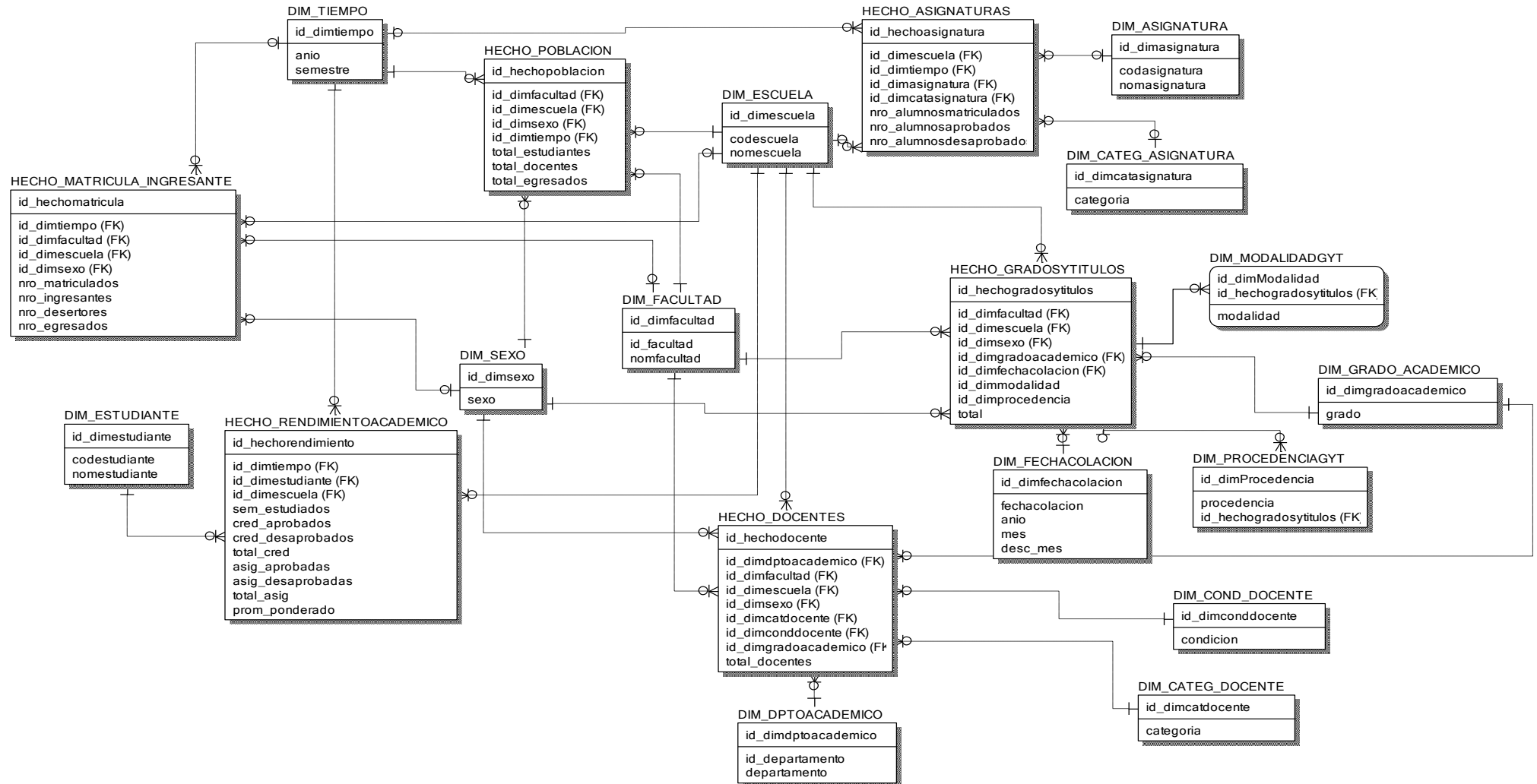


Figura 36: Modelado de base de datos multidimensional del datamart académico

Dimensión Estudiante

Tabla 24: Dimensión estudiante

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_DOCENTE		Dimensional					Registro de datos de los estudiantes		

Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimestudiante	Clave primaria que identifica a la dimensión.	INTEGER		PK		No			
Codestudiente	Código de un estudiante	CHAR	6				Estudiante	Idestudiante	VARCHAR
Nombrestudiante	Nombre completo del estudiante	VARCHAR	90				Estudiante	apestudiante, amestudiante, nombrestudiante	VARCHAR

Dimensión Sexo

Tabla 25: Dimensión sexo

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_SEXO		Dimensional					Registro de datos de sexo		

Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
id_dimsexo	Clave primaria que identifica a la dimensión sexo	INTEGER		PK		No			
Sexo	Sexo de la persona	CHAR	6				Estudiante	Sexo	VARCHAR

Dimensión tiempo

Tabla 26: Dimensión tiempo

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_TIEMPO		Dimensional					Registro de datos del tiempo		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimtiempo	Clave primaria que identifica a la dimensión tiempo.	INTEGER		PK		No			
Anio	Año	INTEGER					Semestre	Anio	INTEGER
semestre	Semestre académico	VARCHAR	7						

Dimensión facultad

Tabla 27: Dimensión facultad

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_FACULTAD		Dimensional					Registro de datos de facultad		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimfacultad	Clave primaria que identifica a la dimensión tiempo.	INTEGER		PK		No			
Id_facultad	Año	INTEGER					Facultad	idfacultad	VARCHAR
nomfacultad	Semestre académico	VARCHAR	7				Facultad	nomfacultad	VARCHAR

Dimensión escuela

Tabla 28: Dimensión escuela

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_ESCUELA		Dimensional					Registro de datos de la escuela académico profesional		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimescuela	Clave primaria que identifica a la dimensión escuela	INTEGER		PK		No			
Codescuela	Código de cada escuela académico profesional	INTEGER					Escuela	idescuela	VARCHAR
Nomescuela	Nombre de cada escuela académico profesional	VARCHAR	80				Escuela	nomescuela	VARCHAR

Dimensión categoría de docente

Tabla 29: Dimensión de categoría docente

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_CATEG_DOCENTE		Dimensional					Registro de datos de la categoría de los docentes		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimcatdocente	Clave primaria que identifica a la dimensión de la categoría del docente	INTEGER		PK		No			
Categoría	Nombre de la categoría del docente	VARCHAR	20				Docente	Categoría	VARCHAR

Dimensión condición del docente

Tabla 30: Dimensión de condición del docente

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_COND_DOCENTE		Dimensional					Registro de datos de la condición de los docentes		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimcondocente	Clave primaria que identifica a la dimensión de la condición del docente	INTEGER		PK		No			
Condición	Nombre de la condición del docente	VARCHAR	20				Docente	condicion	VARCHAR

Dimensión departamento académico

Tabla 31: Dimensión de departamento académico

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_DEPTACADEMICO		Dimensional					Registro de datos del departamento académico		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimdptoacademico	Clave primaria que identifica a la dimensión departamento académico	INTEGER		PK		No			
Iddepartamento	Código departamento	VARCHAR	6				Departamentoacademico	iddepartamento	VARCHAR
Departamento	Nombre del departamento académico	VARCHAR	50				Departamentoacademico	nomdepartamento	VARCHAR

Dimensión asignatura

Tabla 32: Dimensión asignatura

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_ASIGNATURA		Dimensional					Registro de datos de las asignaturas		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimasignatura	Clave primaria que identifica a la dimensión asignatura	INTEGER		PK		No			
Codasignatura	Código de la asignatura	CHAR	7				asignatura	idasignatura	
Nomasignatura	Nombre de la asignatura	TINYTEXT					asignatura	nomasignatura	

Dimensión de la categoría de asignatura

Tabla 33: Dimensión de categoría de asignatura

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_CAT_ASIGNATURA		Dimensional					Registro de datos de las asignaturas		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimcatasignatura	Clave primaria que identifica a la dimensión categoría de asignatura	INTEGER		PK		No			
Categoría	Nombre de la categoría de la asignatura	CHAR	7				categoría	nomcategoría	

Dimensión fecha de colación

Tabla 34: Dimensión de fecha de colación

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
DIM_FECHA_COLACION		Dimensional					Registro de datos de la tabla fecha colación		
Destino							Origen		
Columnas	Descripción	Tipo dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo dato
Id_dimfechacolacion	Clave primaria que identifica a la dimensión fecha colación	INTEGER		PK		No			
Fechacolacion	Fecha de colación	DATE	7						
Anio	Año de colación	INTEGER							
Mes	Mes de colación	INTEGER							
Desc_mes	Descripción del mes de colación	VARCHAR	10						

Tabla de hechos matricula ingresante

Tabla 35: Tabla de hechos de ingresante

TABLA		TIPO DE TABLA					DESCRIPCIÓN		
HECHO_MATRICULA_INGRESANTE		Dimensional					Tabla de hecho conformada por claves primarias de las dimensiones		
Columnas	Descripción					Tipo dato	Tamaño	Clave	Null
Id_hechomatrícula	Clave primaria que identifica a la tabla de hechos matricula ingresante					INTEGER		PK	No
Id_dimtiempo	Llave foránea de la tabla DIM_TIEMPO					INTEGER		FK	No
Id_dimfacultad	Llave foránea de la tabla DIM_FACULTAD					INTEGER		FK	No
Id_dimescuela	Llave foránea de la tabla DIM_ESCUELA					INTEGER		FK	No
Id_dimsexo	Llave foránea de la tabla DIM_SEXO					INTEGER		FK	No
Nromatriculados	Número de matriculados					INTEGER			No
Nroingresantes	Número de ingresantes					INTEGER			No
Nrodesrtores	Número de desertores					INTEGER			No
Nroegresados	Número de egresados					INTEGER			No

Tabla de hechos matricula ingresante

Tabla 36: Tabla de hechos de matrícula de ingresante

TABLA		TIPO DE TABLA	DESCRIPCIÓN			
HECHO_POBLACION_UNIVERSITARIA		Dimensional	Tabla de hecho conformada por claves primarias de las dimensiones			
Columnas	Descripción	Tipo dato	Tamaño	Clave	Null	
Id_hechopoblacion	Clave primaria que identifica a la tabla de hechos matricula ingresante	INTEGER		PK	No	
Id_dimfacultad	Llave foránea de la tabla DIM_TIEMPO	INTEGER		FK	No	
Id_dimescuela	Llave foránea de la tabla DIM_FACULTAD	INTEGER		FK	No	
Id_dimsexo	Llave foránea de la tabla DIM_ESCUELA	INTEGER		FK	No	
Id_dimtiempo	Llave foránea de la tabla DIM_SEXO	INTEGER		FK	No	
Total_estudiantes	Número de matriculados	INTEGER			No	
Total_docentes	Número de ingresantes	INTEGER			No	
Total_egresados	Número de desertores	INTEGER			No	

Tabla de hechos docentes

Tabla 37: Tabla de hechos de docentes

TABLA		TIPO DE TABLA	DESCRIPCIÓN			
HECHO_DOCENTE		Dimensional	Tabla de hecho conformada por claves primarias de las dimensiones			
Columnas	Descripción	Tipo dato	Tamaño	Clave	Null	
Id_hechodocente	Clave primaria que identifica a la tabla de hechos docente	INTEGER		PK	No	
Id_dimdptoacademico		INTEGER		FK		
Id_dimfacultad	Llave foránea de la tabla DIM_FACULTAD	INTEGER		FK	No	
Id_dimescuela	Llave foránea de la tabla DIM_ESCUELA	INTEGER		FK	No	
Id_dimsexo	Llave foránea de la tabla DIM_SEXO	INTEGER		FK	No	
Id_dimcatdocente	Llave foránea de la tabla DIM_CATEG_DOCENTE	INTEGER		FK		
Id_dimcoddocente	Llave foránea de la tabla DIM_COND_DOCENTE	INTEGER		FK		
Id_dimgradocademico	DIM_GRADO_ACADEMICO	INTEGER		FK		
Total_docentes	Total de docentes	INTEGER			No	

Tabla de hechos grados y títulos

Tabla 38: *Tabla de hechos de grados y títulos*

TABLA	TIPO DE TABLA	DESCRIPCIÓN			
HECHO_GRADOS_TITULOS	Dimensional	Tabla de hecho conformada por claves primarias de las dimensiones			
Columnas	Descripción	Tipo dato	Tamaño	Clave	Null
Id_hechogradosytitulos	Clave primaria que identifica a la tabla de hechos de grados y títulos	INTEGER		PK	No
Id_dimfacultad	Llave foránea de la tabla DIM_FACULTAD	INTEGER		FK	No
Id_dimescuela	Llave foránea de la tabla DIM_ESCUELA	INTEGER		FK	No
Id_dimsexo	Llave foránea de la tabla DIM_SEXO	INTEGER		FK	No
Id_dimgradoacademico	Llave foránea de la tabla DIM_GRADO_ACADEMICO	INTEGER		FK	No
Id_dimfechacolacion	Llave foránea de la tabla DIM_FECHA_COLACION	INTEGER		FK	No
Id_dimmodalidad	Llave foránea de la tabla DIM_MODALIDAD	INTEGER		FK	No
Id_dimprocedencia	Llave foránea de la tabla DIM_PROCEDENCIA	INTEGER		FK	No
Total	Total	INTEGER			

Tabla de hechos asignaturas

Tabla 39: *Tabla de hechos de asignaturas*

TABLA	TIPO DE TABLA	DESCRIPCIÓN			
HECHO_ASIGNATURA	Dimensional	Tabla de hecho conformada por claves primarias de las dimensiones			
Columnas	Descripción	Tipo dato	Tamaño	Clave	Null
Id_hechoasignatur	Clave primaria que identifica a la tabla de hechos de grados y títulos	INTEGER		PK	No
Id_dimescuela	Llave foránea de la tabla DIM_ESCUELA	INTEGER		FK	No
Id_dimtiempo	Llave foránea de la tabla DIM_TIEMPO	INTEGER		FK	No
Id_dimasignatura	Llave foránea de la tabla DIM_ASIGNATURA	INTEGER		FK	No
Id_dimcatisignatura	Llave foránea de la tabla DIM_CATEG_ASIGNATURA	INTEGER		FK	No
Nro_alumnosmatriculados	Número de alumnos matriculados	INTEGER			
Nro_alumnosaprobados	Número de alumnos aprobados	INTEGER			
Nro_alumnosdesaprobados	Número de alumnos desaprobados	INTEGER			

Tabla de hechos rendimiento académico

Tabla 40: *Tabla de hechos de rendimiento académico*

TABLA	TIPO DE TABLA	DESCRIPCIÓN
HECHO_RENDIMIENTO_ACADEMICO	Dimensional	Tabla de hecho conformada por claves primarias de las dimensiones

Columnas	Descripción	Tipo dato	Tamaño	Clave	Null
Id_hechorendimiento	Clave primaria que identifica a la tabla de hechos de grados y títulos	INTEGER		PK	No
Id_dimtiempo	Llave foránea de la tabla DIM_TIEMPO	INTEGER		FK	No
Id_dimestudiante	Llave foránea de la tabla DIM_ESTUDIANTE	INTEGER		FK	No
Sem_estudiados	Número de semestres estudiados	INTEGER			
Cred_aprobados	Número de créditos aprobados	INTEGER			
Cred_desaprobados	Número de créditos desaprobados	INTEGER			
Total_cred	Total de créditos	INTEGER			
Asig_aprobadas	Asignaturas aprobadas	INTEGER			
Asig_desaprobadas	Asignaturas desaprobadas	INTEGER			
Total_asig	Total de asignaturas	INTEGER			
Prom_ponderado	Promedio ponderado	NUMERIC	4,2		

6.4.4. Integración de datos

A. Diseño del proceso de ETL

Extracción

Se realizó una conexión con la base de datos `dw_academico` para el proceso de extracción

Transformación

Se realizó las transformaciones para las dimensiones y tablas de hechos.

A continuación se mostrarán cada una de ellas:

- Dim_asignatura



Figura 37: Transformación de la dimensión asignatura

Consulta de Dim_asignatura para la extracción limpia y carga de datos:

```
select idasignatura as codasignatura, nomasignatura as asignatura from asignatura;
```

- Dim_docente



Figura 38: Transformación de la dimensión docente

Consulta de Dim_docente para la extracción limpia y carga de datos:

```
select iddocente, concat(upper(nombresdocente), ' ', apdocente, ' ', amdocente) as nombresdocente from docente;
```

- **Dim_estudiante**



Figura 39: Transformación de la dimensión estudiante

Consulta de Dim_estudiante para la extracción limpieza y carga de datos:

```
select idestudiante as codestudiante, concat(nombreestudiante, '
',apestudiante,' ',amestudiante) as nomestudiante from estudiante;
```

- **Dim_fecha_colacion**

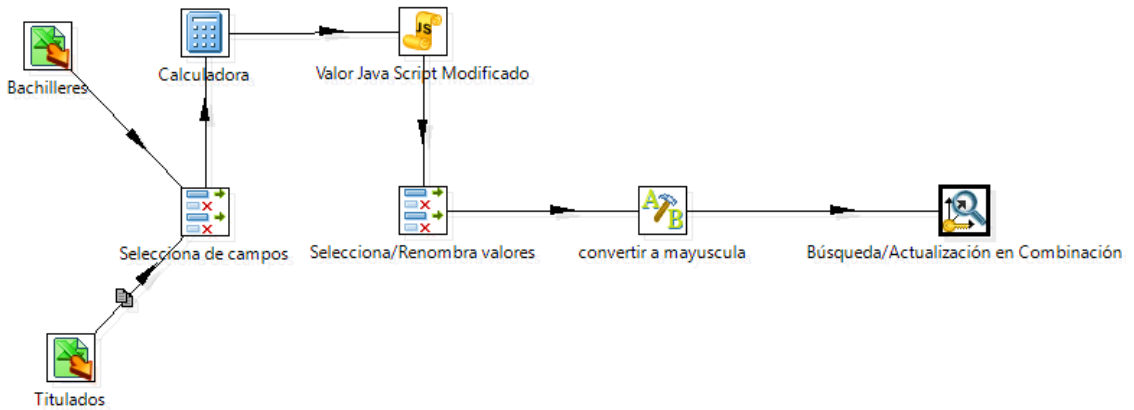


Figura 40: Transformación de la dimensión fecha de colación

- **Hecho_asignatura**

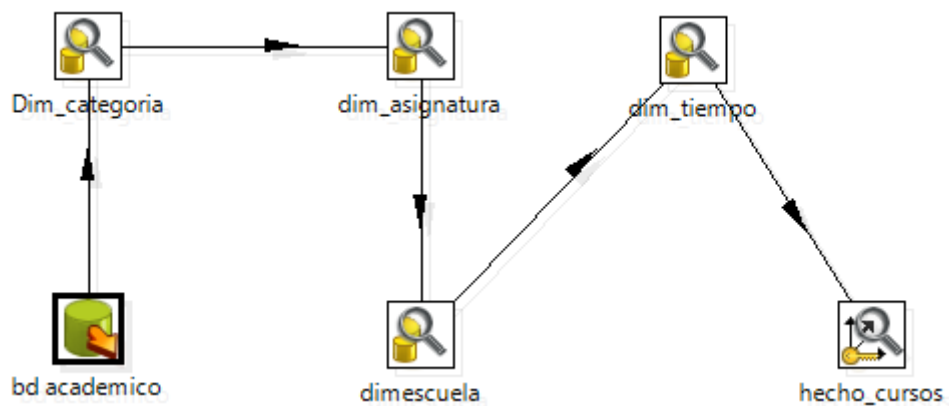


Figura 41: Transformación de hecho_asignatura

Consulta de hecho_asignatura para la extracción limpia y carga de datos:

```

select dm1.idescuela, dm1.idsemestre, al.idcategoria, dm1.idasignatura,
COUNT(*) AS nro_matriculados, nro_aprobados, nro_desaprobados FROM
detmatricula dm1 inner join asignatura al ON
dm1.idasignatura=al.idasignatura INNER JOIN ( select dm.idescuela,
dm.idsemestre, idcategoria, dm.idasignatura, count(*) AS nro_aprobados from
detmatricula dm inner join asignatura a on dm.idasignatura=a.idasignatura
where nota>10 group by dm.idescuela, dm.idsemestre, idcategoria,
dm.idasignatura) AS dm2 ON dm1.idescuela=dm2.idescuela and
dm1.idsemestre=dm2.idsemestre and dm1.idasignatura =dm2.idasignatura
-- Obtener el numero de desaprobados INNER JOIN (select dm.idescuela,
dm.idsemestre, idcategoria, dm.idasignatura, count(*) AS nro_desaprobados
from detmatricula dm inner join asignatura a on
dm.idasignatura=a.idasignatura where nota<11 group by dm.idescuela,
dm.idsemestre, idcategoria, dm.idasignatura) AS dm3 ON
dm1.idescuela=dm3.idescuela and dm1.idsemestre=dm3.idsemestre and
dm1.idasignatura =dm3.idasignatura group by dm1.idescuela, dm1.idsemestre,
al.idcategoria, dm1.idasignatura
    
```

- Hecho_grados_titulos

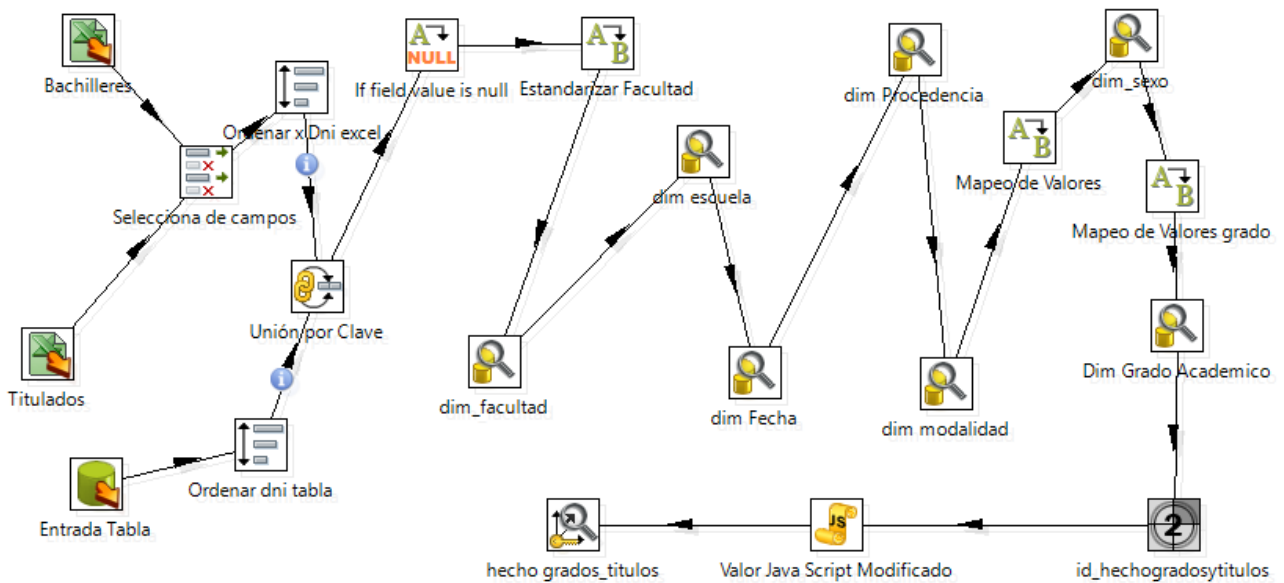


Figura 42: Transformación de la tabla de hechos grados y títulos

Consulta de Hechos_grados_titulos para la extracción limpieza y carga de datos:

```
select idestudiante, sexo from estudiante
```

- **Tabla hecho_rendimiento_academico**

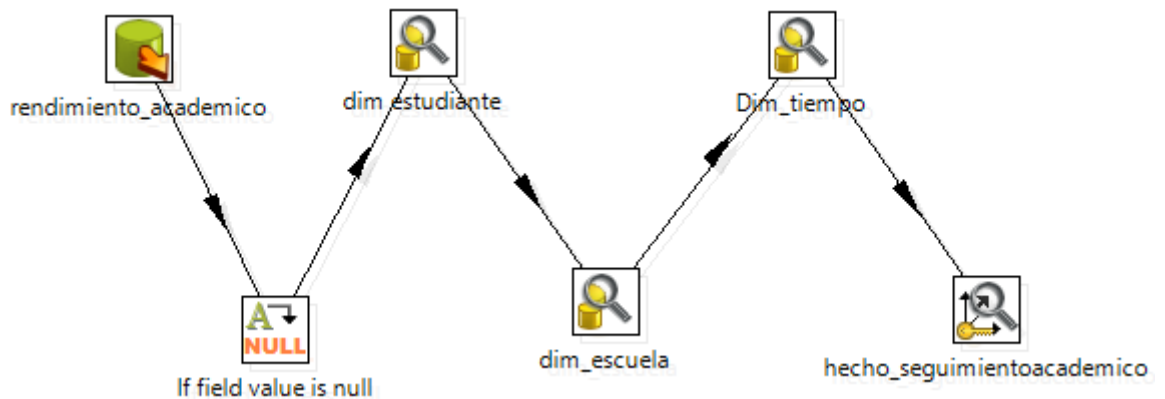


Figura 43: Transformación de la tabla de hechos rendimiento académico

Consulta de Hechos_rendimiento_academico para la extracción limpieza y carga de datos:

```

select      m.idestudiante,      de.idescuela,      m.idsemestre,
concat(nombresestudiante,'      ',apestudiante,'      ',amestudiante)      as
nomestudiante,
(select      count(m1.idsemestre)      from      matricula      m1      where
m1.idestudiante=m.idestudiante      and      m1.idsemestre<=m.idsemestre)as
sem_estudiados,

/*segunda parte*/

(select ifnull(sum(a.creditos),0) from detmatricula dml, asignatura a where
a.idasignatura=dml.idasignatura and dml.idestudiante=m.idestudiante and
dml.idsemestre=m.idsemestre and dml.nota>10) as cred_aprobados,(select
ifnull(sum(a.creditos),0) from detmatricula dml, asignatura a where
a.idasignatura=dml.idasignatura and dml.idestudiante=m.idestudiante and
dml.idsemestre=m.idsemestre and dml.nota<11) as cred_desaprobados,(select
ifnull(sum(a.creditos),0) from detmatricula dml, asignatura a where
a.idasignatura=dml.idasignatura and dml.idestudiante=m.idestudiante and
dml.idsemestre=m.idsemestre) as total_cred,(select count(dml.idasignatura)
from detmatricula dml where dml.idestudiante=m.idestudiante and
dml.idsemestre=m.idsemestre and dml.nota>10) as asig_aprobadas,(select
count(dml.idasignatura)      from      detmatricula      dml      where
dml.idestudiante=m.idestudiante      and      dml.idsemestre=m.idsemestre      and
dml.nota<11) as asig_desaprobadas,(select count(dml.idasignatura) from
detmatricula dml where dml.idestudiante=m.idestudiante and
dml.idsemestre=m.idsemestre)      as      total_asig,(select
(sum((dml.nota)*(a.creditos))/sum(a.creditos)) from detmatricula dml,
asignatura a where dml.idestudiante=m.idestudiante and
dml.idsemestre=m.idsemestre and dml.idasignatura=a.idasignatura) as
prom_ponderado from matricula m, detestudiante de, estudiante e where
m.idestudiante=e.idestudiante and e.idestudiante=de.idestudiante

```

- Dim_modalidad

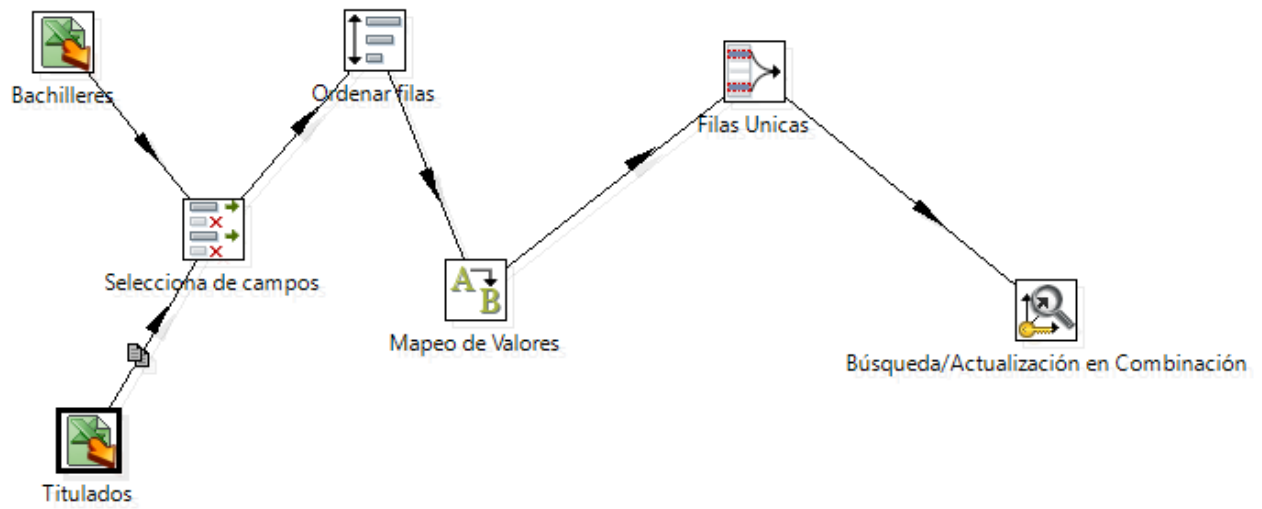


Figura 44: Transformación de la dimensión modalidad

- Dim_procedencia

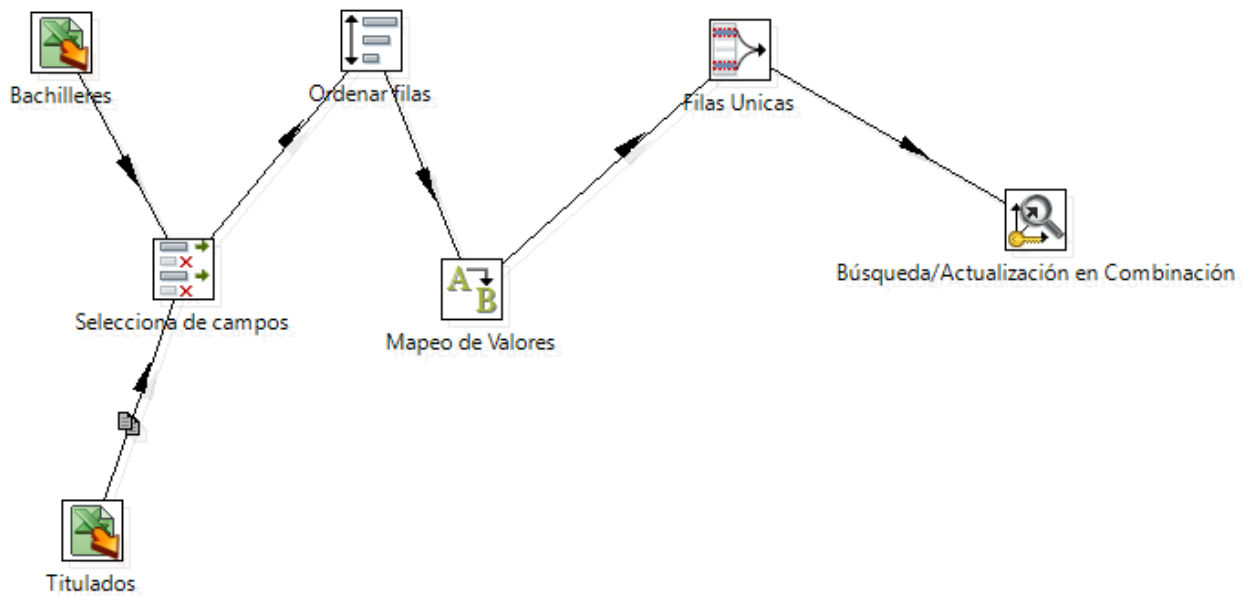


Figura 45: Transformación de la dimensión procedencia

- **Dim_categoria_docente**

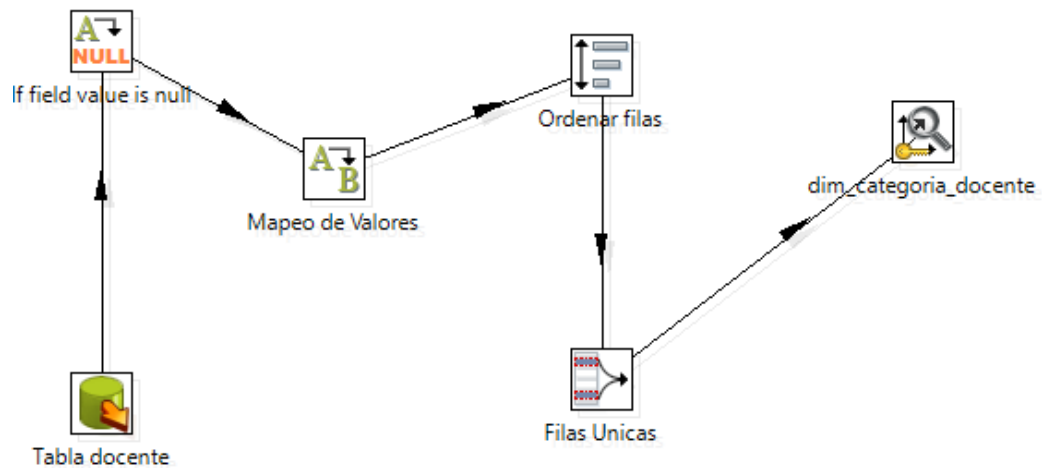


Figura 46: Transformación de la dimensión categoría del docente

Consulta de Dim_categoria_docente para la extracción limpia y carga de datos:

```
SELECT distinct categoria FROM docente
```

- **Dim_categoria_asignatura**



Figura 47: Transformación de la dimensión categoría de asignatura

Consulta de Dim_categoria_asignatura para la extracción limpia y carga de datos:

```
SELECT distinct idcategoria as nomcategoria FROM ASIGNATURA where idcategoria != ' ' ;
```

- **Dim_condicion_docente**

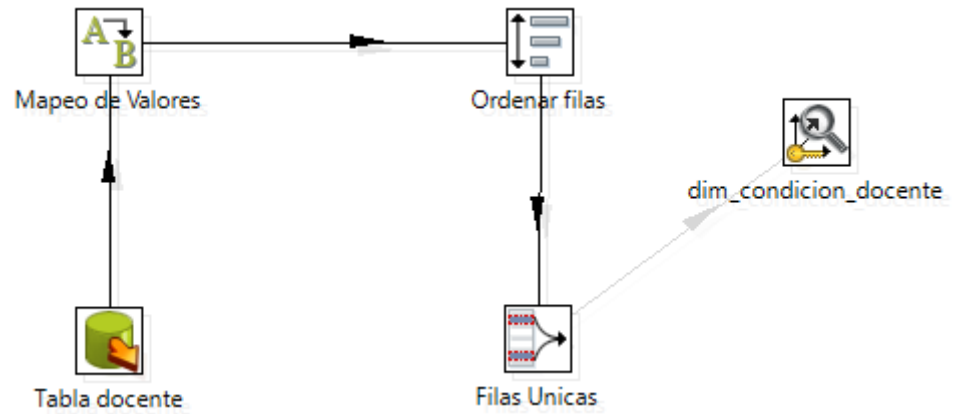


Figura 48: Transformación de la dimensión de condición del docente

Consulta de Dim_condición_docente para la extracción limpia y carga de datos:

```
SELECT distinct condicion FROM docente;
```

- **Dim_departamento_academico**

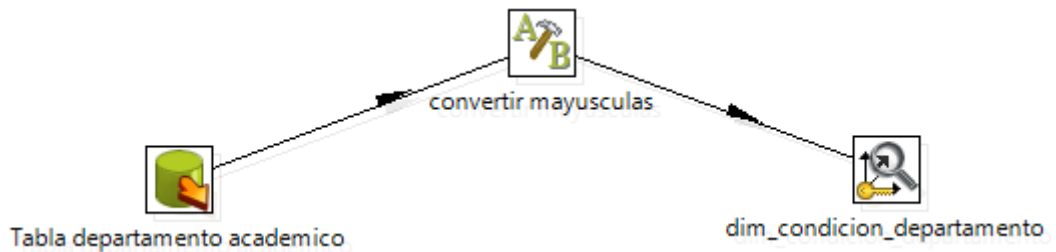


Figura 49: Transformación de la dimensión departamento académico

Consulta de Dim_departamento_academico para la extracción limpia y carga de datos:

```
SELECT iddepartamento, nomdepartamento as departamento FROM departamentoacademico;
```

- **Dim_egresado**



Figura 50: Transformación de la dimensión de egresado

Consulta de Dim_egresado para la extracción limpia y carga de datos:

```
SELECT idestudiante, condicion, año FROM V_EGRESADOSAÑO
```

- **Dim_escuela**



Figura 51: Transformación de la dimensión escuela

Consulta de Dim_escuela para la extracción limpia y carga de datos:

```
SELECT nomfacultad, idescuela as codescuela, nomescuela FROM escuela e  
inner join facultad f on e.idfacultad=f.idfacultad where f.idfacultad  
!='F108' and f.idfacultad != 'F107'
```

- **Dim_facultad**



Figura 52: Transformación de la dimensión facultad

Consulta de Dim_facultad para la extracción limpia y carga de datos:

```
SELECT idfacultad, nomfacultad FROM facultad WHERE idfacultad != 'F108' and  
idfacultad != 'F107'
```


- **Dim_grado_academico**

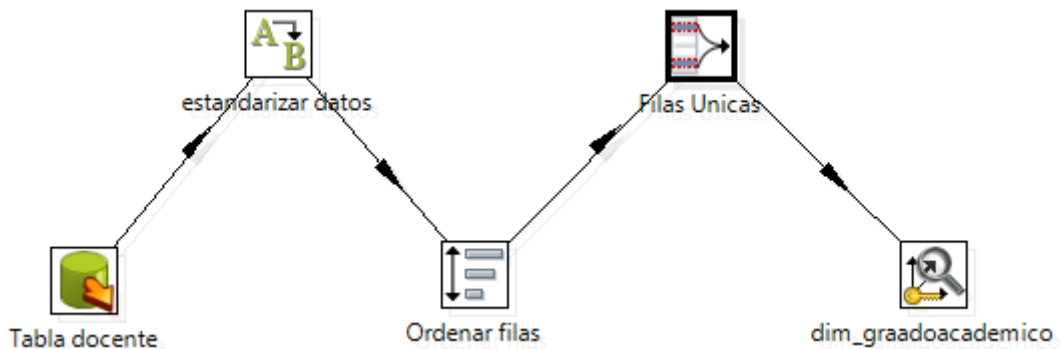


Figura 53: Transformación de la dimensión grado académico

Consulta de Dim_grado_academico para la extracción limpia y carga de datos:

```
SELECT distinct grado FROM docente Where grado != '*****' ;
```

- **Dim_matricula**



Figura 54: Transformación de la dimensión matricula

Consulta de Dim_matricula para la extracción limpia y carga de datos:

```
SELECT idestudiante, idescuela, idsemestre, fechamat FROM matricula
```

- **Dimsexo**

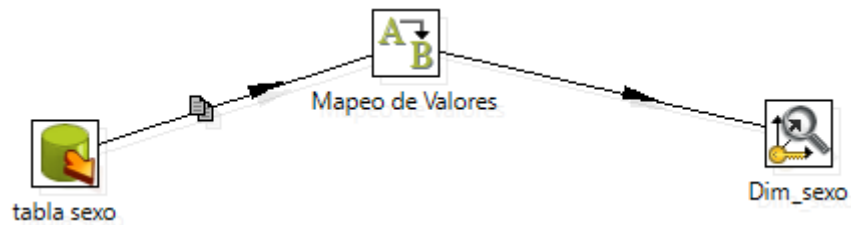


Figura 55: Transformación de la dimensión sexo

Consulta de Dimsexo para la extracción limpia y carga de datos:

```
SELECT distinct sexo FROM estudiante e inner join detestudiante de on e.idestudiante=de.idestudiante where semestreingreso between '2008-1' and '2016-1'
```

- Dim_tiempo

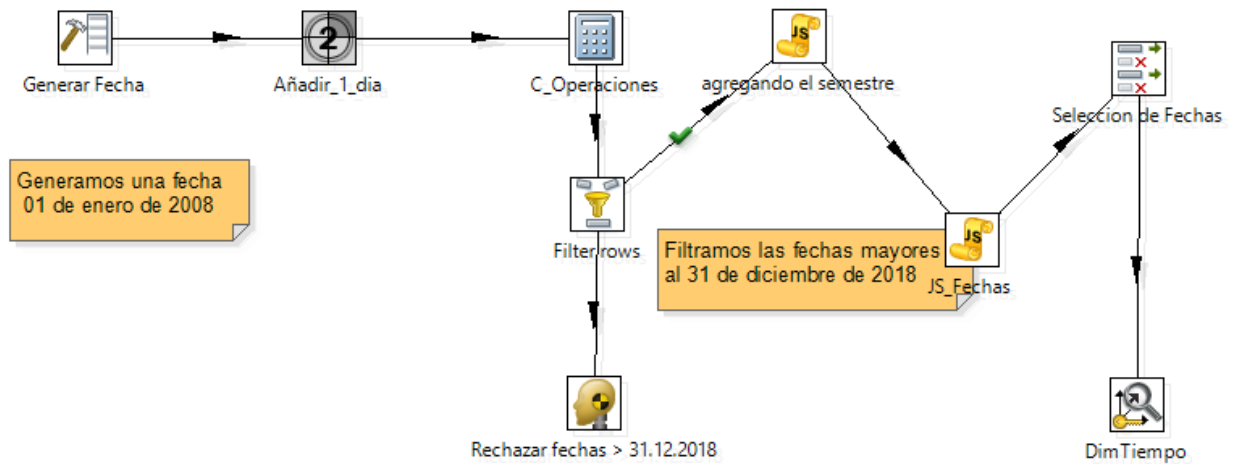


Figura 56: Transformación de la dimensión tiempo

- Hecho_docente

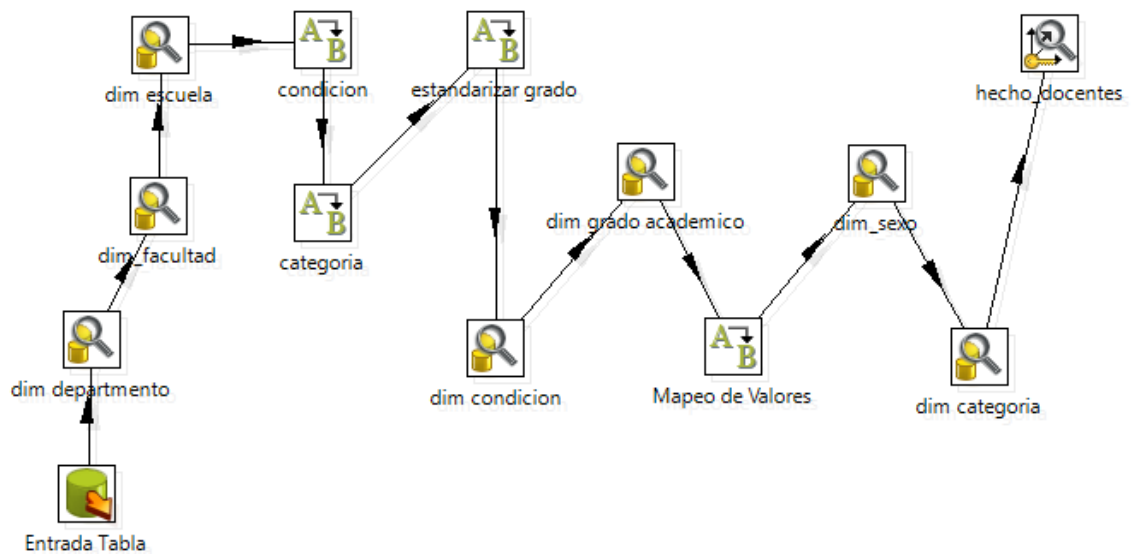


Figura 57: Transformación de hecho docente

Consulta de Hecho_docente para la extracción limpia y carga de datos:

```
select da.iddepartamento, nomfacultad, d.idescuela as codescuela, d.sexo,
d.categoria, d.condicion, d.grado, d.iddocente from docente d inner join
escuela e on e.idescuela=d.idescuela inner join facultad f on
f.idfacultad=e.idfacultad inner join departamentoacademico da on
da.iddepartamento=d.iddepartamento where f.idfacultad !='F107' and
f.idfacultad !='F108'
```

- Hecho_matricula_ingresante

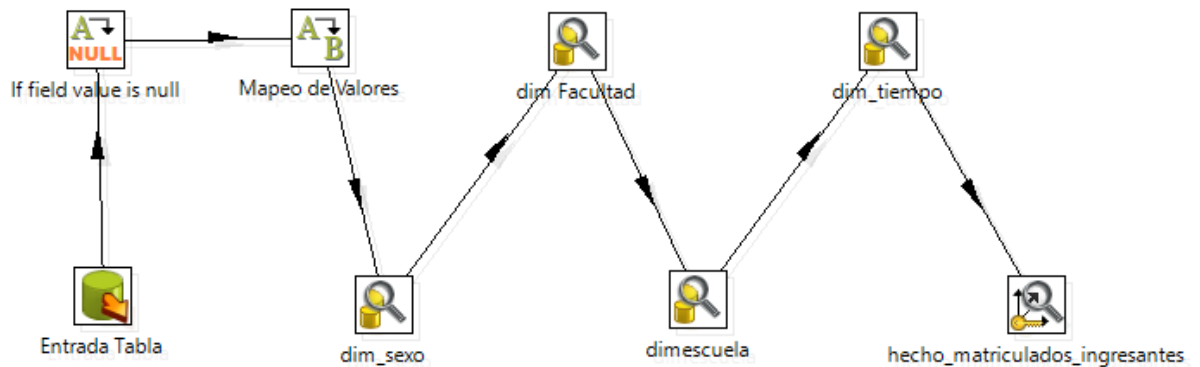


Figura 58: Transformación del hecho matricula del ingresante

Consulta de Hecho_matricula_estudiante para la extracción limpia y carga de datos:

```

Select  es1.idfacultad,  del.idescuela,  del.semestreingreso,  e1.sexo,
count(*) as nro_ingresantes,  nro_matriculados  ,  nro_egresados  ,
nro_desertores  FROM detestudiante del INNER JOIN estudiante e1
ON  del.idestudiante=e1.idestudiante  INNER JOIN escuela es1 on
es1.idescuela  =del.idescuela  INNER JOIN (SELECT  es.idfacultad,
de.idescuela,  de.semestreingreso,  e.sexo,  count(distinct m.idestudiante) as
nro_matriculados  FROM detestudiante de INNER JOIN estudiante e on
de.idestudiante=e.idestudiante  INNER JOIN matricula m on
de.idestudiante=m.idestudiante and  de.idescuela=m.idescuela inner join
escuela es on es.idescuela  =de.idescuela  group by es.idfacultad,
de.idescuela,  de.semestreingreso,  e.sexo) as de3 ON
del.idescuela=de3.idescuela and  del.semestreingreso=de3.semestreingreso
and e1.sexo=de3.sexo and es1.idfacultad=de3.idfacultad LEFT JOIN ( select
idfacultad,idescuela,  semestreingreso,  sexo,  ifnull(count(*),0) as
nro_egresados from v_egresado ve inner join estudiante e on
ve.idestudiante=e.idestudiante  group by idfacultad,idescuela,
semestreingreso,  sexo) AS de4 on del.idescuela=de4.idescuela and
del.semestreingreso=de4.semestreingreso and e1.sexo=de4.sexo and
es1.idfacultad=de4.idfacultad  LEFT JOIN (SELECT
es.idfacultad,de.idescuela,  semestreingreso,  sexo,  count(distinct
de.idestudiante) as nro_desertores FROM detestudiante de inner join
estudiante e on de.idestudiante=e.idestudiante inner join escuela es on
es.idescuela  =de.idescuela  Where activo='NO'  group by es.idfacultad,
idescuela,  semestreingreso,  sexo ) AS de2 ON de2.idescuela=del.idescuela
and  de2.semestreingreso=del.semestreingreso and e1.sexo=de2.sexo and
es1.idfacultad=de2.idfacultad where del.semestreingreso between '2008-1'
and '2016-1' group by es1.idfacultad,  del.idescuela,  del.semestreingreso,
e1.sexo;
  
```

- Hecho_población_universitaria

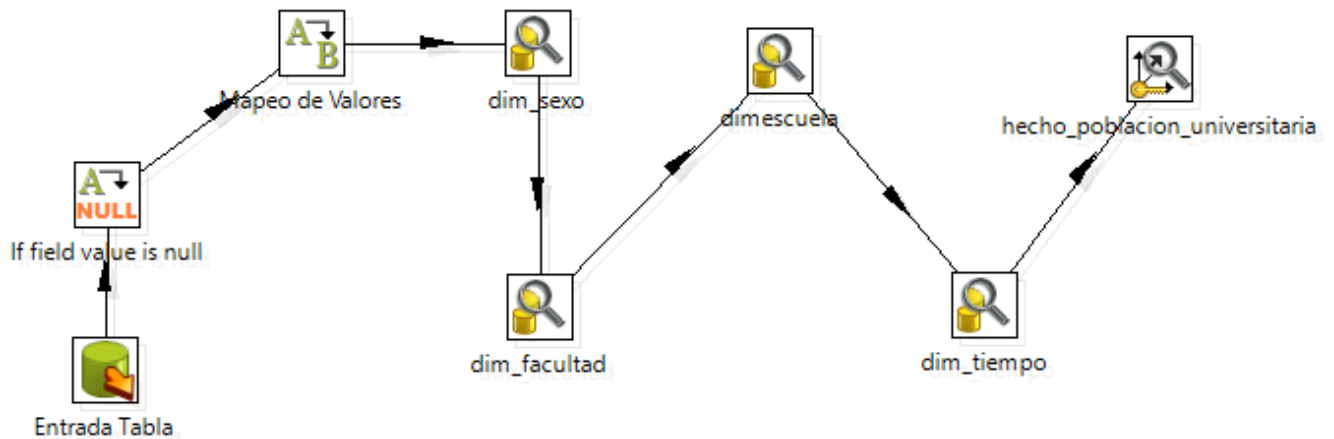


Figura 59: Transformación del hecho población universitaria

Consulta de Hecho_poblacion_universitaria para la extracción limpia y carga de datos:

```

select es.idfacultad, m.idescuela, m.idsemestre as semestre,e.sexo,
count(distinct m.idestudiante) as total_poblacion, total_egresados,
total_docentes from escuela es inner join detestudiante de on
es.idescuela=de.idescuela inner join matricula m on
de.idestudiante=m.idestudiante and de.idescuela=m.idescuela inner join
estudiante e on e.idestudiante=de.idestudiante inner join detmatricula dm
on dm.idestudiante=m.idestudiante and dm.idescuela=m.idescuela and
dm.idsemestre=m.idsemestre inner join asignatura a on
dm.idasignatura=a.idasignatura inner join detcargaacademica dca on
dca.idasignatura=a.idasignatura LEFT JOIN (select es.idfacultad,
es.idescuela, ve.semestre,e.sexo, ifnull(count(*),0) as total_egresados
from escuela es inner join detestudiante de on es.idescuela=de.idescuela
inner join v_egresadoaño1 ve on ve.idestudiante=de.idestudiante inner join
estudiante e on de.idestudiante=e.idestudiante where total_creditos>=216
group by es.idfacultad, es.idescuela, ve.semestre,sexo ) AS d1 ON
es.idfacultad=d1.idfacultad and m.idescuela=d1.idescuela and
m.idsemestre=d1.semestre and e.sexo=d1.sexo LEFT JOIN (select
es.idfacultad, dca.idescuela, dca.idsemestre, d.sexo, count( distinct
dca.iddocente) as total_docentes from detcargaacademica dca inner join
docente d on dca.iddocente=d.iddocente inner join asignatura a on
dca.idasignatura=a.idasignatura inner join detmatricula dm on
dm.idasignatura=a.idasignatura inner join matricula m on
dm.idestudiante=m.idestudiante and dm.idescuela=m.idescuela and
dm.idsemestre=m.idsemestre inner join detestudiante de on
de.idestudiante=m.idestudiante and de.idescuela=m.idescuela inner join
escuela es on es.idescuela=de.idescuela group by es.idfacultad,
dca.idescuela, dca.idsemestre, sexo) AS d2 ON es.idfacultad=d2.idfacultad
and m.idescuela=d2.idescuela and m.idsemestre=d2.idsemestre and
e.sexo=d2.sexo where dm.idsemestre between '2008-1' and '2016-1' group by
es.idfacultad, m.idescuela, m.idsemestre,sexo;

```

B. Carga: Se procede a poblar el data mart como se muestra en el siguiente

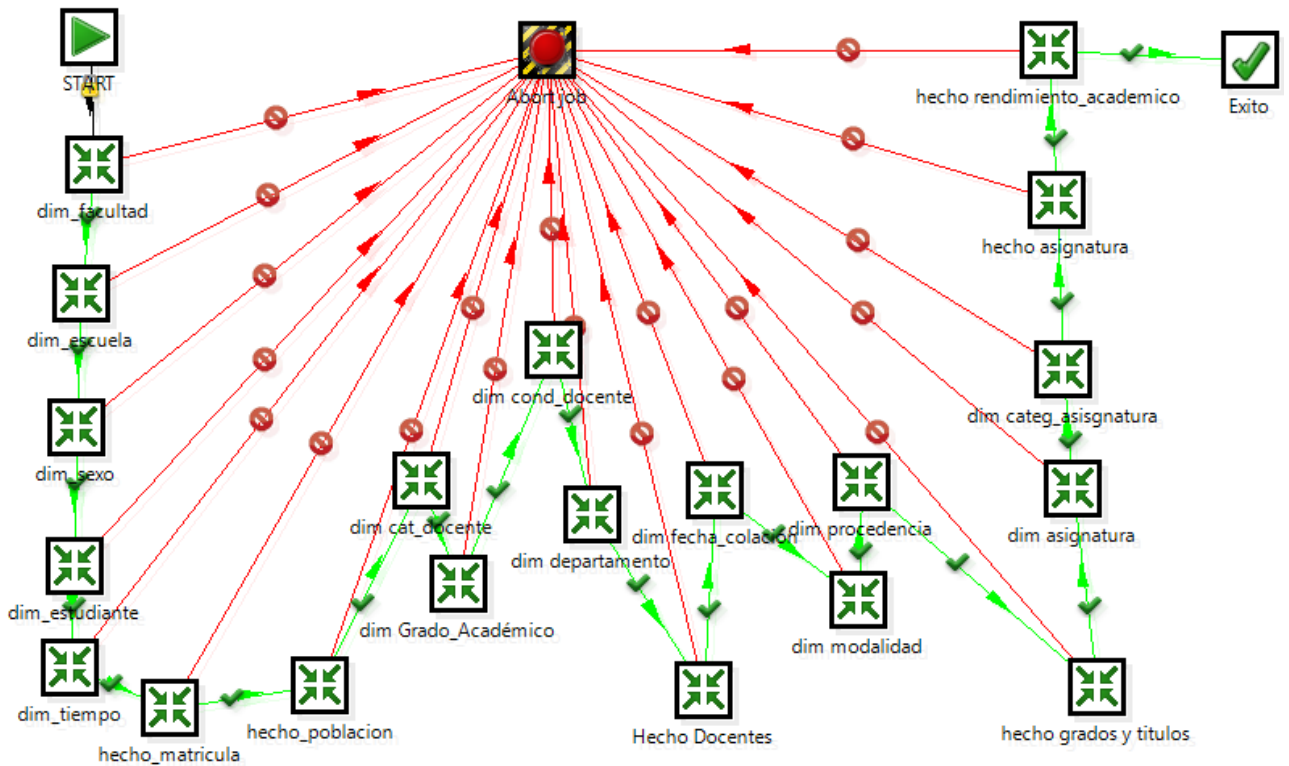


gráfico:

Figura 60: Proceso de carga de data de las dimensiones y hechos

6.5. Análisis multidimensional

Se utilizó la herramienta Schema Workbench para diseñar los de cubos multidimensionales mediante dimensiones y jerarquías lo que permitió realizar el análisis multidimensional.

El data warehouse académico, cuenta con 6 cubos los cuales son:

- Cubo de asignaturas
- Cubo de seguimiento de matriculados ingresantes
- Cubo de rendimiento académico
- Cubo de población universitaria
- Cubo de docentes
- Cubo de grados y títulos

Estructura de los cubos multidimensionales

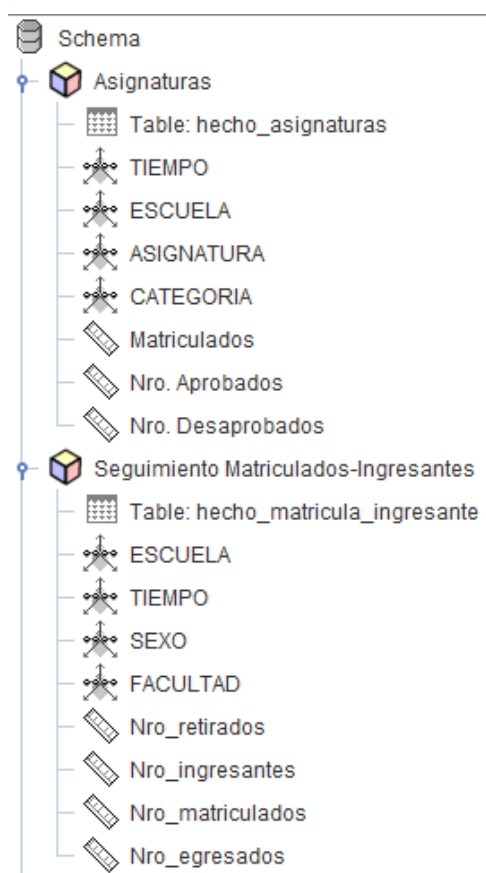


Figura 61: Estructura de cubos del modelo multidimensional

CONCLUSIONES

Al concluir el trabajo de investigación denominado “INTELIGENCIA DE NEGOCIOS USANDO PENTAHO PARA LA GESTIÓN ACADÉMICA DE LA UNAMBA 2016”, se determinó las siguientes conclusiones:

- Se implementó un modelo de solución de inteligencia de negocios el cual muestra información resumida, de manera que permita realizar un análisis de datos de forma rápida, informada, basada en hechos, que sea específica, precisa y en tiempo real, para que los decanos, directores y jefes de departamento académico puedan analizar, tomar decisiones y correcciones necesarias sobre el seguimiento y rendimiento académico de estudiantes, información de egresados, graduados y titulados, plana de docentes, y conocimiento de la población universitaria de la Universidad Nacional Micaela bastidas de Apurímac – sede Abancay.
- Se demostró que la disponibilidad de la información en la propuesta de solución de inteligencia es inmediata puesto que el número de reportes obtenidos después de la inteligencia de negocios es superior al número de reportes obtenidos antes de la solución de inteligencia de negocios; esto debido a que la aplicación centraliza, procesa y administra datos históricos con respecto a las matriculas, desempeño académico del estudiante, docentes, graduados y titulados e información de la población estudiantil a nivel de ingresantes, egresados, retirado es decir hubo un incremento del 70.37% de lo que anteriormente era un 29.62 %.
- Se confirmó que el tiempo de respuesta de consulta de los reportes de la propuesta de solución de inteligencia de negocios es menor con inteligencia de

negocios ya que representa un 5.83 %, mientras que antes de la solución de inteligencia de negocios equivale a un 94%.

- Con la aplicación de la encuesta del modelo conceptual de datos de un datamart según la ISO 25012, utilizando la escala de Likert, se demostró que del 100% que equivalen las 10 preguntas planteadas, el 13 % que no está acuerdo ni en desacuerdo, un 70% que está de acuerdo y un 17% que está totalmente de acuerdo.
- Se utilizó la metodología HEFESTO para el desarrollo del trabajo de investigación, debido a que esta metodología se puede adaptar a cualquier ciclo de vida del desarrollo del software.
- La propuesta de solución de inteligencia de negocios para la UNAMBA no se llegó a implantar debido a que las computadoras no cumplen con los requerimientos mínimos de hardware que se necesitan y por la dificultad de accesibilidad en las oficinas de las direcciones, jefaturas académicas y las decanaturas.

RECOMENDACIONES

Para futuras investigaciones acerca de inteligencia de negocios se recomienda lo siguiente:

- Ampliar la investigación con indicadores claves de desempeño a nivel académico, de investigación, responsabilidad social, biblioteca, utilizando otras técnicas de análisis de información como los dashboards.
- Para realizar un mejor análisis de la información en toda la universidad se recomienda consolidar la información académica y administrativa de la universidad implementando un Datawarehouse.
- Mantener la información actualizada en el sistema de matrícula de la universidad, así como en cada oficina.
- Se sugiere a parte de la herramienta Pentaho que es una de las herramientas más completas a nivel de funcionalidad para la explotación de base de datos de Inteligencia de negocios, utilizar otro tipo de herramientas de Inteligencia de negocios a nivel de investigación como Jaspersoft y SpagoBI, que también son completas.

BIBLIOGRAFÍA

- Aimacaña, A. (2013). *Análisis, diseño e implementación de un datamart académico usando tecnología BI para la facultad de ingeniería, ciencias físicas y matemática*. Tesis. Ingeniería Informática.
- Bernabeu, R. (2010). *Hefesto v2*. Libro
- Bernabeu, R., y Garcia, M. (2017). *Hefesto data warehousing v3*. Libro
- Blanco, I., y Quesada, V (2008). *La gestión académica, criterio clave de la calidad de la gestión de las instituciones de educación superior*, Santo Tomás Distrito Federal, México.
- Campos, A., y Dominguez, L (2012). *El clima organizacional y el compromiso institucional en las instituciones de educación superior de puerto Vallarta*. Cartagena, Colombia.
- Cano, J., (2007). *Business Intelligence: Competir con Información*. Libro publicado por ESADE, Banesto, Banesto Pyme.
- Castillo, I., Flores, L. & Cervantes, G. (2015). *Gestión académica saludable en el contexto universitario*. Costa Rica.
- Castillo, W., & Medina, F., & Fariño, F. (2017). *Data Mart para obtención de indicadores de productividad académica en una universidad*. Artículo. Revista chilena de ingeniería, Chile.
- Conesa, J., y Curto J. (2011). *Introducción al Business Intelligence*. Barcelona, España. Editorial el Ciervo 96, SA.
- Gutierrez, J; (2015). *Implantación de Business Intelligence como mejora en la gestión académica del Área de Coordinación académica de la UPIG, Surco-2015*. Tesis de Sistemas e Informática.

- Hernández, S., & Fernández, C., & Baptista, L. (2014). *Metodología de la investigación* (6et ed.). Santa Fe, México: Interamericana Editores, S.A. DE C.V.
- Marín, P., & Marx, J., & García, L. (2012). Framework para el diseño de una solución de inteligencia de negocio. Saabrücken, Alemania: Editorial Académica Española.
- Rojas, A. (2014). *Implementación de un data Mart como solución de inteligencia de negocios, bajo la metodología de Ralph Kimball para optimizar la toma de decisiones en el departamento de finanzas de la contraloría general de la república*. Tesis de Computación y Sistemas.
- Ramos, S. (2011). *Microsoft Business Intelligence*, Editado SolidQ. Libro
- Torres,G., & Salinas,O, (2011). *Herramientas de Explotación de Información de Bases de Dato*, Morelos, México.
- Zegarra, F; (2016). *Solución de inteligencia de negocios orientada a la toma de decisiones en las operaciones mineras de extracción y metalurgia de hochschild mining*. Tesis de Ingeniería de computación y Sistemas.

ANEXOS

ANEXO 1



UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE
APURÍMAC
ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA
INFORMÁTICA Y SISTEMAS



ENCUESTA

Presentación: Buenos días/tardes. La presente encuesta forma parte del proyecto de investigación de tesis denominado “**Inteligencia de negocios usando Pentaho para la gestión académica de la UNAMBA-2016**”. Agradecería unos minutos de su tiempo para contribuir en responder este cuestionario. Los datos serán anónimos y de uso exclusivo para la tesis sin utilizar para otros fines distintos de ello.

ENCUESTA DE CALIDAD DEL MODELO CONCEPTUAL DE DATOS DE UN DATAMART

Esta encuesta permite determinar la calidad del modelo conceptual de datos según el ISO/IEC 25012, de la aplicación de inteligencia de negocios.

Marque con una X la alternativa que considere la más adecuada.

Puntajes	Escalas
0	Totalmente en desacuerdo
1	En desacuerdo
2	Ni en desacuerdo ni de acuerdo
3	De acuerdo
4	Totalmente de acuerdo

Características	Preguntas	ESCALA DE LIKERT				
		Totalmente en desacuerdo	En desacuerdo	Indeciso	De acuerdo	Totalmente de acuerdo
Legibilidad	1. ¿La lectura del modelo conceptual de datos es entendible?					
	2. ¿La tipografía del modelo conceptual de datos es la más adecuada?					
Complejidad	3. ¿Se visualizan requerimientos representados en el modelo conceptual de datos?					
Corrección	4. ¿Las distintas partes del modelo conceptual de datos están construidas con la semántica correcta?					



	5. ¿Cada elemento del modelo conceptual de datos se representa haciendo uso de estructuras adecuadas?					
Minimidad	6. ¿El modelo conceptual de datos no muestra información redundante?					
	7. ¿Si se eliminará un elemento del esquema del modelo conceptual de datos se perdería la información?					
Expresividad	8. ¿Los elementos del modelo conceptual de datos pueden ser comprendidos fácilmente?					
Autoexplicación	9. ¿La lógica de negocios puede ser entendida por el modelo conceptual de datos?					
Extensibilidad	10. ¿Se identifican las partes del esquema conceptual es decir las tablas de hechos y dimensiones con sus atributos?					

ANEXO 2



**UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE
APURÍMAC
ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA
INFORMÁTICA Y SISTEMAS
ENCUESTA DIRIGIDO A EXPERTOS**



Presentación: Buenos días/tardes. La presente encuesta forma parte del proyecto de investigación de tesis denominado **“Inteligencia de negocios usando Pentaho para la gestión académica de la UNAMBA-2016”**. Agradecería unos minutos de su tiempo para contribuir en responder este cuestionario. Los datos serán anónimos y de uso exclusivo para la tesis sin utilizar para otros fines distintos de ello.

ENCUESTA

Esta encuesta permite conocer si usted cuenta o no con los reportes y gráficos Marca con una X la respuesta que considera adecuada.

1. ¿Usted cuenta con un reporte que muestre el número de estudiantes matriculados por semestre académico y asignatura? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
2. ¿Usted cuenta con un reporte que muestre el número de estudiantes matriculados por año y asignatura? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
3. ¿Usted cuenta con un reporte que muestre el número de aprobados y desaprobados por semestre académico y asignatura? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
4. ¿Usted cuenta con gráfico estadístico que muestre de cantidad de matriculados por semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese gráfico.
Sí No Tiempo: _____
5. ¿Usted cuenta con un reporte que muestre el total de docentes por cada escuela académica profesional y categoría? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
6. ¿Usted cuenta con un reporte que muestre del total de docentes por sexo según escuela académico profesional y facultad? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
7. ¿Usted cuenta con un reporte que muestre el número de docentes por condición según escuela académico profesional? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
8. ¿Usted cuenta con un reporte que muestre el número de docentes por grado académico, departamento académico y escuela académico profesional? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
9. ¿Usted cuenta con un reporte que muestre el total de graduados por año y escuela académico profesional? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____
10. ¿Usted cuenta con un reporte que muestre el total de graduados por modalidad, sexo y escuela académico profesional? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí No Tiempo: _____

11. ¿Usted cuenta con un reporte de rendimiento académico que muestre el total de asignaturas aprobadas y desaprobadas por un estudiante que pertenece a una escuela académica profesional y según el semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
12. ¿Usted cuenta con un reporte de rendimiento académico que muestre el total de créditos aprobados de un estudiante que pertenece a una escuela académica profesional, según semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
13. ¿Usted cuenta con un reporte de rendimiento académico que muestre el del total de créditos desaprobados de estudiante que pertenece a una escuela académica profesional, según semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
14. ¿Usted cuenta con un reporte de rendimiento académico que muestre el total de semestres estudiados de un estudiante según escuela académica profesional, semestre y año? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
15. ¿Usted cuenta con un reporte de rendimiento académico que muestre el promedio ponderado de un estudiante según escuela académica profesional, semestre académico y año? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
16. ¿Usted cuenta con un reporte que muestre el total de egresados por semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
17. ¿Usted cuenta con un Gráfico estadístico que muestre la cantidad de egresados por semestre De ser si su respuesta, ingrese el tiempo en que obtiene ese gráfico.
Sí _____ No _____ Tiempo: _____
18. ¿Usted cuenta con un reporte que muestre el número de egresados por género y año? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
19. ¿Usted cuenta con un reporte que muestre el número de ingresantes por semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
20. ¿Usted cuenta con un Gráfico estadístico que muestre el número de ingresantes por semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese gráfico.
Sí _____ No _____ Tiempo: _____
21. ¿Usted cuenta con un reporte que muestre el total de estudiantes por sexo y grado académico según la escuela académica profesional? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____
22. ¿Usted cuenta con un reporte que muestre el número de retirados de la escuela académico profesional y semestre académico? De ser si su respuesta, ingrese el tiempo en que obtiene ese reporte.
Sí _____ No _____ Tiempo: _____

ANEXO 3

SCRIPT DE LA BASE DE DATOS MULTIDIMENSIONAL

```
DROP DATABASE IF EXISTS DW_ACADEMICO;
```

```
CREATE DATABASE DW_ACADEMICO;
```

```
USE DW_ACADEMICO;
```

```
CREATE TABLE DIM_ESTUDIANTE
```

```
(  
  id_dimestudiante  INTEGER NOT NULL,  
  codestudiante     CHAR(6) NOT NULL,  
  nomestudiante     VARCHAR(90) NOT NULL,  
  PRIMARY KEY (id_dimestudiante)  
);
```

```
CREATE TABLE DIM_SEXO
```

```
(  
  id_dimsexo  INTEGER NOT NULL,  
  sexo       CHAR(1) NOT NULL,  
  PRIMARY KEY (id_dimsexo)  
);
```

```
CREATE TABLE DIM_TIEMPO
```

```
(  
  id_dimtiempo  INTEGER NOT NULL,  
  anio          INTEGER NULL,  
  semestre      VARCHAR(7) NULL,  
  PRIMARY KEY (id_dimtiempo)  
);
```

```
CREATE TABLE DIM_FACULTAD
```

```
(  
  id_dimfacultad  INTEGER NOT NULL,  
  idfacultad     VARCHAR(4) NULL,  
  nomfacultad    VARCHAR(50) NULL,  
  PRIMARY KEY (id_dimfacultad)  
);
```

```
CREATE TABLE DIM_ESCUELA
```

```
(  
  id_dimescuela  INTEGER NOT NULL,  
  codescuela     INTEGER NULL,  
  nomescuela     VARCHAR(80) NULL,  
  PRIMARY KEY (id_dimescuela)  
);
```



```
CREATE TABLE DIM_CATEG_DOCENTE
```

```
(  
    id_dimcatdocente INTEGER NOT NULL,  
    categoria VARCHAR(20) DEFAULT NULL,  
    PRIMARY KEY (id_dimcatdocente)  
);
```

```
CREATE TABLE DIM_COND_DOCENTE
```

```
(  
    id_dimconddocente INTEGER NOT NULL,  
    condicion VARCHAR(20) DEFAULT NULL,  
    PRIMARY KEY (id_dimconddocente)  
);
```

```
CREATE TABLE DIM_DPTOACADEMICO
```

```
(  
    id_dimdptoacademico INTEGER NOT NULL,  
    iddepartamento VARCHAR(6) NOT NULL,  
    departamento VARCHAR(50) DEFAULT NULL,  
    PRIMARY KEY (id_dimdptoacademico)  
);
```

```
CREATE TABLE DIM_GRADO_ACADEMICO
```

```
(  
    id_dimgradoacademico INTEGER NOT NULL,  
    grado VARCHAR(20) NOT NULL,  
    PRIMARY KEY (id_dimgradoacademico)  
);
```

```
CREATE TABLE DIM_ASIGNATURA
```

```
(  
    id_dimasignatura INTEGER NOT NULL,  
    codasignatura CHAR(7) NULL,  
    nomasignatura TINYTEXT NULL,  
    PRIMARY KEY (id_dimasignatura)  
);
```

```
CREATE TABLE DIM_CATEG_ASIGNATURA
```

```
(  
    id_dimcatasignatura INTEGER NOT NULL,  
    categoria VARCHAR(20) DEFAULT NULL,  
    PRIMARY KEY (id_dimcatasignatura)  
);
```

```

CREATE TABLE DIM_FECHACOLACION
(
    id_dimfechacolacion INTEGER NOT NULL,
    fechacolación DATE NOT NULL,
    año INTEGER NOT NULL,
    mes INTEGER NOT NULL,
    desc_mes VARCHAR(10) NOT NULL,
    PRIMARY KEY (id_dimfechacolacion)
);

```

```

CREATE TABLE DIM_MODALIDADGYT
(
    id_dimModalidad integer NOT NULL,
    modalidad VARCHAR(80) NOT NULL,
    PRIMARY KEY (id_dimModalidad )
);

```

```

CREATE TABLE DIM_PROCEDENCIAGYT
(
    id_dimProcedencia integer NOT NULL,
    procedencia VARCHAR(90) NOT NULL,
    PRIMARY KEY (id_dimProcedencia )
);

```

```

CREATE TABLE HECHO_MATRICULA_INGRESANTE
(
    id_hechomatricula INTEGER NOT NULL,
    id_dimtiempo INTEGER NULL,
    id_dimfacultad INTEGER NULL,
    id_dimescuela INTEGER NULL,
    id_dimsexo INTEGER NULL,
    nro_matriculados INTEGER NULL,
    nro_ingresantes INTEGER NULL,
    nro_desertores INTEGER NULL,
    nro_egresados INTEGER NULL,
    PRIMARY KEY (id_hechomatricula),
    FOREIGN KEY (id_dimtiempo) REFERENCES DIM_TIEMPO(id_dimtiempo),
    FOREIGN KEY (id_dimescuela) REFERENCES
DIM_ESCUELA(id_dimescuela),
    FOREIGN KEY (id_dimfacultad) REFERENCES
DIM_FACULTAD(id_dimfacultad),
    FOREIGN KEY (id_dimsexo) REFERENCES DIM_SEXO(id_dimsexo)
);

```

```

CREATE TABLE HECHO_POBLACION_UNIVERSITARIA
(
    id_hechopoblacion    INTEGER NOT NULL,
    id_dimfacultad       INTEGER NOT NULL,
    id_dimescuela        INTEGER NOT NULL,
    id_dimsexo           INTEGER NOT NULL,
    id_dimtiempo         INTEGER NOT NULL,
    total_estudiantes    INTEGER NOT NULL,
    total_docentes       INTEGER NOT NULL,
    total_egresados      INTEGER NOT NULL,
    PRIMARY KEY (id_hechopoblacion),
    FOREIGN KEY (id_dimfacultad) REFERENCES
DIM_FACULTAD(id_dimfacultad),
    FOREIGN KEY (id_dimescuela) REFERENCES
DIM_ESCUELA(id_dimescuela),
    FOREIGN KEY (id_dimsexo) REFERENCES DIM_SEXO(id_dimsexo),
    FOREIGN KEY (id_dimtiempo) REFERENCES DIM_TIEMPO(id_dimtiempo)
) ;

```

```

CREATE TABLE HECHO_DOCENTES
(
    id_hechodocente     INTEGER NOT NULL,
    id_dimdptoacademico INTEGER NOT NULL,
    id_dimfacultad      INTEGER NOT NULL,
    id_dimescuela       INTEGER NOT NULL,
    id_dimsexo          INTEGER NOT NULL,
    id_dimcatdocente    INTEGER NOT NULL,
    id_dimconddocente   INTEGER NOT NULL,
    id_dimgradoacademico INTEGER NOT NULL,
    total_docentes      INTEGER NOT NULL,
    PRIMARY KEY (id_hechodocente),
    FOREIGN KEY (id_dimdptoacademico) REFERENCES
DIM_DPTOACADEMICO(id_dimdptoacademico),
    FOREIGN KEY (id_dimfacultad) REFERENCES
DIM_FACULTAD(id_dimfacultad),
    FOREIGN KEY (id_dimescuela) REFERENCES
DIM_ESCUELA(id_dimescuela),
    FOREIGN KEY (id_dimsexo) REFERENCES DIM_SEXO(id_dimsexo),
    FOREIGN KEY (id_dimcatdocente) REFERENCES
DIM_CATEG_DOCENTE(id_dimcatdocente),
    FOREIGN KEY (id_dimconddocente) REFERENCES
DIM_COND_DOCENTE(id_dimconddocente),
    FOREIGN KEY (id_dimgradoacademico) REFERENCES
DIM_GRADO_ACADEMICO(id_dimgradoacademico)
) ;

```

```

-- DROP TABLE IF EXISTS HECHO_GRADOSYTITULOS;

CREATE TABLE HECHO_GRADOS_TITULOS
(
    id_hechogradosytitulos    INTEGER NOT NULL,
    id_dimfacultad            INTEGER NOT NULL,
    id_dimescuela             INTEGER NOT NULL,
    id_dimsexo                INTEGER NOT NULL,
    id_dimgradoacademico     INTEGER NOT NULL,
    id_dimfechacolacion      INTEGER NOT NULL,
    id_dimmodalidad          INTEGER NOT NULL,
    id_dimprocedencia        INTEGER NOT NULL,
    total                     INTEGER NOT NULL,
    PRIMARY KEY (id_hechogradosytitulos),
    FOREIGN KEY              (id_dimfacultad)          REFERENCES
DIM_FACULTAD(id_dimfacultad),
    FOREIGN KEY (id_dimescuela) REFERENCES DIM_ESCUELA(id_dimescuela),
    FOREIGN KEY (id_dimsexo) REFERENCES DIM_SEXO(id_dimsexo),
    FOREIGN KEY              (id_dimgradoacademico)    REFERENCES
DIM_GRADO_ACADEMICO(id_dimgradoacademico),
    FOREIGN KEY              (id_dimfechacolacion)     REFERENCES
DIM_FECHACOLACION(id_dimfechacolacion),
    FOREIGN KEY              (id_dimModalidad)        REFERENCES
DIM_MODALIDADGYT(id_dimModalidad),
    FOREIGN KEY              (id_dimprocedencia)       REFERENCES
DIM_PROCEDENCIAGYT(id_dimprocedencia)
) ;

-- DROP TABLE IF EXISTS HECHO_ASIGNATURAS;
CREATE TABLE HECHO_ASIGNATURAS
(
    id_hechoasignatura        INTEGER NOT NULL,
    id_dimescuela            INTEGER NULL,
    id_dimtiempo             INTEGER NULL,
    id_dimasignatura         INTEGER NULL,
    id_dimcatasignatura      INTEGER NULL,
    nro_alumnosmatriculados  INTEGER NULL,
    nro_alumnosaprobados    INTEGER NULL,
    nro_alumnosdesaprobados  INTEGER NULL,
    PRIMARY KEY (id_hechoasignatura),
    FOREIGN KEY (id_dimtiempo) REFERENCES DIM_TIEMPO(id_dimtiempo),
    FOREIGN KEY              (id_dimasignatura)        REFERENCES
DIM_ASIGNATURA(id_dimasignatura),
    FOREIGN KEY              (id_dimcatasignatura)    REFERENCES
DIM_CATEG_ASIGNATURA(id_dimcatasignatura),
    FOREIGN KEY              (id_dimescuela)          REFERENCES
DIM_ESCUELA(id_dimescuela)
);

```

```

CREATE TABLE HECHO_RENDIMIENTO_ACADEMICO
(
    id_hechorendimiento INTEGER NOT NULL,
    id_dimtiempo         INTEGER NOT NULL,
    id_dimestudiante     INTEGER NOT NULL,
    id_dimescuela        INTEGER NOT NULL,
    sem_estudiados       INTEGER NOT NULL,
    cred_aprobados       INTEGER NOT NULL,
    cred_desaprobados    INTEGER NOT NULL,
    total_cred           INTEGER NOT NULL,
    asig_aprobadas       INTEGER NOT NULL,
    asig_desaprobadas    INTEGER NOT NULL,
    total_asig           INTEGER NOT NULL,
    prom_ponderado       NUMERIC(4,2) NULL,
    PRIMARY KEY (id_hechorendimiento),
    FOREIGN KEY (id_dimtiempo) REFERENCES DIM_TIEMPO(id_dimtiempo),
    FOREIGN KEY (id_dimestudiante) REFERENCES
    DIM_ESTUDIANTE(id_dimestudiante),
    FOREIGN KEY (id_dimescuela) REFERENCES
    DIM_ESCUELA(id_dimescuela)
);

```

ANEXO 4

CONSTANCIA

El que suscribe, Ing Rodolfo A. Mattos Ojeda, Director de la Escuela Académica Profesional de Ingeniería de Minas, hace constar:

Que, la bachiller Hamely Sarmiento Ponce de la Escuela Académico Profesional de Ingeniería Informática y Sistemas, identificada con DNI N° 70144165 y que habiendo culminado con el desarrollado de la tesis denominada "INTELIGENCIA DE NEGOCIOS USANDO PENTAHO PARA LA GESTIÓN ACADÉMICA EN LA UNAMBA-2016", solicitó implantar en la dirección de Minas la solución de inteligencia de negocios, pero no se pudo llevar a cabo dicho proceso debido a que la PC no cumple con las características mínimas requeridas; sin embargo ella mostró y explicó de manera adecuada la implementación de la solución de inteligencia de negocios junto con sus respectivos procesos, constando que dicha propuesta de tesis contribuiría en la gestión académica de la Universidad debido a que la herramienta que propone utilizar es interactiva y fácil de comprender y aprender, ya que permite generar varios reportes dinámicos y gráficos estadísticos según nuestras necesidades, facilitando la disponibilidad de la información y ayudando en la toma de decisiones; en todo caso dicha propuesta sirve como base para implantarla en un futuro ya con las condiciones necesarias que se deban cumplir.

Sin más que decir, se expide la presente constancia, para los fines convenientes del solicitante.

Abancay, 02 de mayo del 2018.

UNIVERSIDAD NACIONAL MICAELA BASTIDAS
APURÍMAC
E.A.P. INGENIERÍA DE MINAS

Ing. Rodolfo A. Mattos Ojeda
DIRECTOR