

**UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE MECÁNICA ELÉCTRICA, ELECTRÓNICA
Y SISTEMAS
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**



**INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE
NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA
INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE
ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018
TESIS**

PRESENTADA POR:

YOEL DIOMEDEZ ROMERO ALMONTE

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO DE SISTEMAS

PUNO – PERÚ

2018

UNIVERSIDAD NACIONAL DEL ALTIPLANO - PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y
SISTEMAS
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

**INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE
NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN,
EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE
LENGUAS EXTRANJERAS Y NATIVAS 2018**

TESIS PRESENTADO POR:

YOEL DIOMEDEZ ROMERO ALMONTE

PARA OPTAR EL TITULO PROFESIONAL DE:

INGENIERO DE SISTEMAS



APROBADO POR EL JURADO REVISOR CONFORMADO POR:

PRESIDENTE:



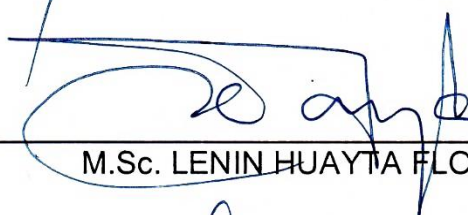
Mg. CARLOS BORIS SOSA MAYDANA

PRIMER MIEMBRO:



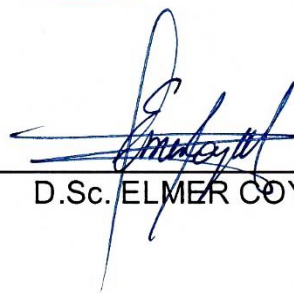
M.Sc. PABLO CESAR TAPIA CATACORA

SEGUNDO MIEMBRO:



M.Sc. LENIN HUAYTA FLORES

DIRECTOR / ASESOR:



D.Sc. ELMER COYLA IDME

ÁREA: Sistemas de Información

TEMA: Ingeniería de Software, Bases de Datos e Inteligencia de Negocios

FECHA DE SUSTENTACIÓN 31 DE DICIEMBRE DEL 2018

Dedicatoria

A mi madre.

Por hacerme el regalo más grande, más fuerte,
haberme regalado todo lo que tienes.

Por buscarme un lugar donde fuera valiente,
para ser feliz, conmigo mismo.

Agradecimientos

A la Universidad Nacional del Altiplano – Puno
y a la Escuela Profesional de Ingeniería de
Sistemas por haber sido el alma mater de mi
educación superior.

Al Centro de Estudios de Lenguas Extranjeras
y Nativas por abrirme las puertas para realizar
el presente trabajo de investigación.

Índice General

Dedicatoria	3
Agradecimientos	4
Índice de Figuras	8
Índice de Tablas	10
Índice de Acrónimos	11
Resumen	12
Abstract	13
CAPÍTULO I	14
INTRODUCCIÓN	14
1.1. Problema	16
1.2. Justificación	16
1.3. Objetivos	17
1.3.1. General	17
1.3.2. Específicos	18
1.4. Hipótesis	18
1.5. Variables e Indicadores	18
1.5.1. Independiente	18
1.5.2. Dependiente	18
1.5.3. Indicadores	18
1.6. Limitaciones	20
CAPÍTULO II	21
REVISIÓN DE LITERATURA	21
2.1. Antecedentes	21
2.2. Gestión de la Información	26
2.2.1. Indicadores	27
2.3. Inteligencia de Negocios	29
2.4. Proceso de Extracción, Transformación y Carga	32
2.5. Data Mart	34
2.5.1. Tabla de Hechos	36
2.5.2. Dimensiones	37
2.5.3. Medidas	38
2.6. Data Warehouse	38
2.7. Tecnologías OLTP Y OLAP	39
2.7.1. Sistemas OLTP	40
2.7.2. Sistemas OLAP	41

2.8.	Pentaho	42
2.9.	Dashboard	44
2.10.	Metodología de Bill Inmon	45
2.11.	Metodología de Ralph Kimball.....	46
CAPÍTULO III		48
MATERIALES Y MÉTODOS		48
3.1.	Análisis del Negocio y Obtención de Requerimientos e Indicadores	48
3.1.1.	Proceso del Negocio Académico	51
3.1.2.	Requerimientos Funcionales	53
3.1.3.	Requerimientos No Funcionales	54
3.1.4.	Estado Actual de los Indicadores.....	55
3.2.	Identificación de las Fuentes de Datos	55
3.2.1.	Fuente y Destino de los Datos	56
3.3.	Diseño del Data Mart.....	59
3.3.1.	Antigüedad de los Datos.....	60
3.3.2.	Definición de Dimensiones	60
3.3.3.	Jerarquías Analíticas y Granularidad	62
3.3.4.	Claves Sustitutas y Primarias	63
3.3.5.	Definición de la Tabla de Hechos.....	66
3.3.6.	Traslado del Modelo Dimensional al Modelo Lógico.....	67
3.4.	Implementación de Soluciones para la Visualización del Data Mart	70
3.4.1.	Carga de Dimensiones.....	70
3.4.2.	Carga de la Tabla de Hechos	71
3.4.3.	Carga de Dimensiones y Tabla de Hechos.....	72
3.4.4.	Gestión del Cubo OLAP	72
3.4.5.	Reportes.....	74
3.4.6.	Dashboard	82
CAPÍTULO IV.....		86
RESULTADOS Y DISCUSIÓN		86
3.4.	Población y Muestra	86
3.5.	Análisis e Interpretación de Resultados	87
3.5.1.	Resultados Genéricos	87
3.5.2.	Resultados Específicos	88
3.6.	Nivel de Confianza.....	89
3.7.	Contraste de la Hipótesis	89
3.7.1.	Contraste para la Exactitud de la Información.....	89

3.7.2. Contraste para el Tiempo En Realizar Los Reportes	91
3.7.3. Contraste para el Número de Reportes Solicitados por Ciclo	93
3.7.4. Contraste para la Comprensión de los Reportes	94
3.8. Discusión	96
CONCLUSIONES.....	97
RECOMENDACIONES.....	98
REFERENCIAS BIBLIOGRÁFICAS	99
ANEXOS	101

Índice de Figuras

Figura 2.1: Elementos de BI.....	31
Figura 2.2: Esquema del Modelo Estrella	36
Figura 2.3: Interacción de los Sistemas OLTP y OLAP	40
Figura 2.4: Arquitectura de Bill Inmon	46
Figura 2.5: Ciclo de Vida de la Metodología de Kimball	47
Figura 3.1: Ubicación del CELEN.....	49
Figura 3.2: Principales Competidores del CELEN	50
Figura 3.3: Tablas de la Base de Datos Transaccional.....	55
Figura 3.4: Tablas para la Base de Datos Multidimensional	56
Figura 3.5: Fuente Tabla Estudiantes - Destino Dimensión Estudiante.....	56
Figura 3.6: Fuente Tabla Docentes - Destino Dimensión Instructor.....	57
Figura 3.7: Fuente Tabla Especialidades - Destino Dimensión Especialidad	57
Figura 3.8: Fuente Tabla Cursos - Destino Dimensión Curso	58
Figura 3.9: Fuente Tabla Horarios - Destino Dimensión Horario	58
Figura 3.10: Fuente Tabla Registro de Notas - Destino Dimensión Tiempo	59
Figura 3.11: Mapeo de Datos para la Dimensión Estudiante	60
Figura 3.12: Mapeo de Datos para la Dimensión Docente.....	61
Figura 3.13: Mapeo de Datos para la Dimensión Tiempo.....	61
Figura 3.14: Mapeo de Datos para la Dimensión Horario.....	61
Figura 3.15: Mapeo de Datos para la Dimensión Especialidad	62
Figura 3.16: Mapeo de Datos para la Dimensión Curso	62
Figura 3.17: Modelo Estrella, Relaciones entre Dimensiones y la Tabla de Hechos.....	67
Figura 3.18: Modelo Lógico Dimensional	69
Figura 3.19: Entorno de la Herramienta Gráfica PDI	70
Figura 3.20: Carga de Datos para la Dimensión Estudiante.....	70
Figura 3.21: Carga de Datos para la Dimensión Instructor	70
Figura 3.22: Carga de Datos para la Dimensión Especialidad	71
Figura 3.23: Carga de Datos para la Dimensión Curso.....	71
Figura 3.24: Carga de Datos para la Dimensión Horario	71
Figura 3.25: Carga de Datos para la Dimensión Tiempo	71
Figura 3.26: Carga de Datos para la Tabla de Hechos	71
Figura 3.27: Carga de Dimensiones y Tabla de Hechos	72
Figura 3.28: Entorno de la Herramienta PSW	72
Figura 3.29: Diseño del Cubo OLAP CELEN.....	73
Figura 3.30: Publicación del Esquema OLAP	73
Figura 3.31: Entorno del Plugin Saiku Analytics	74
Figura 3.32: Porcentaje de Cursos tomados por Especialidad.....	80
Figura 3.33: Cantidad de Horarios por Periodo.....	80
Figura 3.34: Cantidad de Instructores por Sexo	81
Figura 3.35: Cantidad de Cursos por Año	81
Figura 3.36: Interfaz de la Herramienta CDE.....	82
Figura 3.37: Listas desplegable para el filtrado de Estudiante, Curso y Año.....	82
Figura 3.38: Gráfico de Líneas - Rendimiento del Estudiante por Año.....	82
Figura 3.39: Gráfico de Barras - Rendimiento del Estudiante por Curso.....	83
Figura 3.40: Gráfico Circular - Horarios del Estudiante por Año	83
Figura 3.41: Gráfico Circular - Genero de los Docentes del Alumno por Año	84
Figura 3.42: Tabla de Datos - Datos completos del Estudiante.....	84

Figura 3.43: Vista General del Dashboard Principal	85
Figura 4.1: Distribución de Probabilidad para la Exactitud de la Información	90
Figura 4.2: Ingreso de Datos para realizar la Prueba-t.....	90
Figura 4.3: Distribución de la Probabilidad para el Tiempo en Realizar los Reportes	92
Figura 4.4: Distribución de la Probabilidad para el Número de Reportes Solicitados por Ciclo.	93
Figura 4.5: Distribución de la Probabilidad para la Comprensión de los Reportes	95

Índice de Tablas

Tabla 1.1 Indicadores para la Variable Independiente	18
Tabla 1.2: Indicadores para la Variable Dependiente	19
Tabla 1.3: Operacionalización de la Variable Independiente	19
Tabla 1.4: Operacionalización de la Variable Independiente	20
Tabla 2.1: Diferencias entre el Esquema Estrella y Copo de Nieve	36
Tabla 2.2: Diferencias entre Bases de Datos Transaccionales y Almacenes de Datos	39
Tabla 2.3: Comparativa de Softwares de Inteligencia de Negocios.....	44
Tabla 3.1: Cursos y Duración por Idioma	48
Tabla 3.2: Horarios por Turno	49
Tabla 3.3: Criterios de Evaluación para el Idioma Ingles	52
Tabla 3.4: Lista de Requerimientos Funcionales.....	53
Tabla 3.5: Calificación y Significado de la Prioridad y Dificultad de los Requerimientos.....	54
Tabla 3.6: Datos Actuales de los Indicadores.....	55
Tabla 3.7: Definición de Dimensiones	60
Tabla 3.8: Jerarquía y Granularidad de la Dimensión Estudiante.....	62
Tabla 3.9: Jerarquía y Granularidad de la Dimensión Docente	63
Tabla 3.10: Jerarquía y Granularidad de la Dimensión Tiempo.....	63
Tabla 3.11: Jerarquía y Granularidad de la Dimensión Horario.....	63
Tabla 3.12: Jerarquía y Granularidad de la Dimensión Especialidad	63
Tabla 3.13: Jerarquía y Granularidad de la Dimensión Curso.....	63
Tabla 3.14: Dimensión Estudiante con Atributos Jerárquicos.....	64
Tabla 3.15: Dimensión Docente con Atributos Jerárquicos	64
Tabla 3.16: Dimensión Tiempo con Atributos Jerárquicos	65
Tabla 3.17: Dimensión Horario con Atributos Jerárquicos	65
Tabla 3.18: Dimensión Especialidad con Atributos Jerárquicos	65
Tabla 3.19: Dimensión Curso con Atributos Jerárquicos.....	66
Tabla 3.20: Granularidad de la Tabla de Hechos.....	66
Tabla 3.21: Modelo Dimensional Curso	67
Tabla 3.22: Modelo Dimensional Instructor.....	68
Tabla 3.23: Modelo Dimensional Horario.....	68
Tabla 3.24: Modelo Dimensional Especialidad	68
Tabla 3.25: Modelo Dimensional Estudiante	68
Tabla 3.26: Modelo Dimensional Tiempo.....	68
Tabla 3.27: Modelo Dimensional Récord Académico	69
Tabla 3.28: Reporte del Récord Académico de un Estudiante por Especialidad y Curso	74
Tabla 3.29: Reporte del Récord Académico de un Estudiante por Curso y Docente.....	75
Tabla 3.30: Reporte del Récord Académico de un Estudiante por Tiempo (Año - Mes).....	77
Tabla 3.31: Reporte del Récord Académico de un Estudiante por Periodo y Horario.....	78
Tabla 3.32: Reporte del Récord Académico de un Estudiante por Trimestre	79
Tabla 4.1: Resultados Generales de la Pre-Test (GC) y Post-Prueba (GE).....	88
Tabla 4.2: Resultados Específicos Pre-Test (GC) y Post-Test (GE).....	88
Tabla 4.3: Resultados de la Prueba-t para la Exactitud de la Información.....	91
Tabla 4.4: Resultados de la Prueba-t para el Tiempo en Realizar los Reportes	92
Tabla 4.5: Resultados de la Prueba-t para el Número de Reportes por Ciclo.....	94
Tabla 4.6: Resultados de la Prueba-t para la Comprensión de los Reportes	96

Índice de Acrónimos

BI	: Business Intelligence
CDE	: Community Dashboard Editor
CELEN	: Centro de Estudios de Lenguas Extranjeras y Nativas
DM	: Data Mart
DW	: Data Warehouse
ETL	: Extract Transform and Load
GC	: Grupo de Control
GE	: Grupo Experimental
OLAP	: On-Line Analytical Processing
OLTP	: On-Line Transaction Processing
PDI	: Pentaho Data Integration
PSW	: Pentaho Schema Workbench
TI	: Tecnologías de Información
UNA	: Universidad Nacional del Altiplano

Resumen

El presente proyecto incorporó un conjunto de elementos de inteligencia de negocios que apoyaron a mejorar la gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas de la Universidad Nacional del Altiplano Puno.

El objetivo de este proyecto fue incorporar elementos de inteligencia de negocios para proporcionar calidad a la presentación de los datos y que a partir de hechos e información argumentada sirva como un apoyo a la toma de decisiones; iniciando con el análisis y obtención de requerimientos e indicadores del negocio, identificación de las distintas fuentes de datos, creación de reportes y un dashboard a partir de indicadores cuantitativos.

Para el desarrollo del proyecto se utilizó la metodología y el enfoque planteado por Ralph Kimball¹; como herramientas de software, la suite de Inteligencia de Negocios Pentaho, para evitar costos de licencias.

Teniendo como resultados el incremento de la exactitud de la información de un 60 a un 93 por ciento, la generación de un mayor número de reportes, de 3 a 8 por ciclo, la reducción del tiempo de elaboración de reportes, de 62 a 30 minutos y una mayor comprensión de reportes, de un 66 a un 86 por ciento.

Palabras Clave: Área académica, gestión de la información, inteligencia de negocios.

¹ Es un autor en el tema de almacenamiento de datos e inteligencia empresarial. Su metodología, también conocida como modelado dimensional o la metodología de Kimball, se ha convertido en el estándar en el área de soporte de decisiones. (Véase más en la sección 2.11.)

Abstract

The present project incorporated a set of business intelligence elements that supported improving information management, in the Puno National University's Center for Foreign and Native Language Studies' academic area.

The aim of this project was to incorporate elements of business intelligence to provide quality to the presentation of data and that from facts and information reasoned serve as a support for decision-making; starting with the analysis and obtaining requirements and indicators of the business, identification of the different sources of data, creation of reports and a dashboard based on quantitative indicators.

For the development of the project, the methodology and approach proposed by Ralph Kimball will be used; as software tools, the Pentaho Business Intelligence suite, to avoid licensing costs.

Taking as a result the increase of the accuracy of the information from 60 to 93 percent, the generation of a greater number of reports from 3 to 8 per cycle, the reduction of the time of elaboration of reports of 62 30 minutes and the increase in the understanding of reports from 66 to 86 percent.

Key Words: Academic area, business intelligence, information management.

CAPÍTULO I

INTRODUCCIÓN

Hoy en día, las organizaciones están comprendiendo la importancia de la gestión de la información y las ventajas competitivas que implica su uso. Este proceso de gestión consiste en lograr de una manera eficiente el análisis de distintos tipos de datos de la empresa y su entorno, a través de la explotación de la información por medio de las tecnologías de la información (TI), facilitando la adaptación de aplicaciones para la Inteligencia de Negocios. El éxito de una organización, depende fundamentalmente de las acertadas decisiones por parte de la administración, la cual, si es competente, es probable que alcance los objetivos trazados, ya que sobre ella recae la responsabilidad de gestionar todos los recursos tanto humanos, económicos y materiales.

También es un apoyo primordial al momento de tomar decisiones con la información obtenida, en el momento y lugar correcto, lo que les permite aumentar la efectividad de la organización, incrementando su progreso y bienestar.

La hipótesis que se quiere demostrar es, si la incorporación de elementos de inteligencia de negocios, mejorará la gestión de la información en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas (CELEN).

La metodología utilizada para la incorporación de los elementos de inteligencia de negocios conduce a una solución completa en una cantidad de tiempo relativamente pequeña. Además, debido a la gran cantidad de documentación que se puede encontrar y a los numerosos ejemplos aportados en diferentes entornos, permite encontrar una respuesta a casi todas las preguntas que puedan surgir, sobre todo cuando no se dispone de la experiencia previa necesaria.

Las limitaciones encontradas en la fase de la incorporación de los elementos de inteligencia de negocios fue el tiempo en la limpieza de base de datos y el tiempo limitado de la Alta Gerencia por eso hubo un retraso en la realización de las entrevistas.

Con el propósito de hacer más entendible la presente tesis, se ha dividido la misma en cuatro capítulos, cuyos contenidos son los siguientes:

En el **Capítulo I: Introducción**. Se describe todo lo referente al planteamiento del problema, justificación, objetivos, hipótesis, variables e indicadores así como las limitaciones.

En el **Capítulo II: Revisión de la Literatura**. Se describe los antecedentes, teniendo como referencias tesis, libros, artículos científicos, y la parte teórica de la tesis, la validación del marco teórico relacionado con la metodología y modelos que se están usando para el desarrollo de la tesis.

En el **Capítulo III: Materiales y Métodos**. Esta es la parte más importante de la tesis, ya que, se describe la parte de la incorporación de elementos de inteligencia de negocios usando la Metodología de Ralph Kimball además de técnicas e instrumentos de recolección de información.

En el **Capítulo IV: Resultados y Discusión**. Se realiza la prueba empírica para la recopilación, análisis e interpretación de los resultados obtenidos. En primer lugar, se describe la población y muestra, luego el tipo de muestra y nivel de confianza. También se muestra el análisis de la Pre-Test (GC) y Post-Test (GE). Los datos se visualizan en tablas, las cuales, al término de este capítulo, serán analizados para la contrastación de la hipótesis.

Finalmente, se presenta las conclusiones a la cuales se llegaron, recomendaciones para futuras investigaciones, referencias bibliográficas y anexos.

1.1. Problema

El Centro de Estudios de Lenguas Extranjeras y Nativas de la Universidad Nacional del Altiplano Puno mensualmente registra y archiva información del desempeño académico. Esta información comprende diferentes capacidades tales como hablar, leer, escuchar, escribir, además de la asistencia y participación en clase, estas capacidades son evaluados por el docente durante y al terminar cada ciclo. El problema reside en que todos estos datos se quedan almacenados sin ser aprovechados.

La ineficaz gestión de esta información genera que los docentes no puedan adaptar sus metodologías a la clase de estudiantes con los que se enfrenta cada mes, pues está dado que tanto los estudiantes y docentes migran de grupos de enseñanza mensualmente, por lo cual es necesaria la consulta de estos datos de forma adecuada.

Para darle una solución al problema se propone incorporar elementos de inteligencia de negocios que nos permite almacenar sólo la información requerida e ir eliminando aquellos datos que obstaculizan la labor del análisis y entregando la información que se requiera en la forma más apropiada, facilitando así el proceso de gestión de la misma.

1.2. Justificación

La Inteligencia de negocios hace referencia al uso de datos dentro de una organización siendo un conjunto de estrategias enfocadas a la administración y creación de conocimiento para facilitar la toma de decisiones.

En la actualidad los datos son lo más importante para una organización, por ende, el adecuado manejo es importante generando procesamientos y almacenamientos para ser monitoreados de forma rápida. Las diferentes tecnologías han facilitado un mejor aprovechamiento de la información brindando seguridad y estabilidad. En toda organización se requiere tomar decisiones por lo que resulta importante mantener los recursos de información bien organizados, estructurados e integrados.

Las estrategias para el manejo de información son variantes por los avances tecnológicos, y es necesario desarrollar modelos que permitan la administración y gestión de los datos que ayuden a las organizaciones a ser más competitivas generando estrategias sencillas para todos los usuarios, aprovechando toda la información generada y teniendo un conocimiento más amplio de ellas. El buen funcionamiento de una organización se basa en la utilización de sistemas de información de inteligencia mediante herramientas como Data Warehouse, Data Mart y técnicas encargadas de extraer y preparar los datos para luego ser cargados y almacenados; las herramientas se encuentran divididas en dos segmentos: plataformas de Inteligencia de Negocios, que se usan para analizar o crear aplicaciones fáciles con pocos datos, y herramientas de Inteligencia de Negocios empresariales, utilizadas cuando hay muchos datos con diferentes requerimientos, reportes y vistas.

1.3. Objetivos

1.3.1. General

Incorporar elementos de inteligencia de negocios, para mejorar la gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas.

1.3.2. Específicos

- Analizar la naturaleza del negocio, obteniendo los requerimientos e indicadores.
- Identificar las distintas fuentes de datos que serán la base para el diseño de un Data Mart.
- Diseñar un Data Mart.
- Implementar soluciones para la visualización y análisis de los datos contenidos en el Data Mart.

1.4. Hipótesis

La incorporación de elementos de inteligencia de negocios, mejora la gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas

1.5. Variables e Indicadores

1.5.1. Independiente

La incorporación de elementos de inteligencia de negocios.

1.5.2. Dependiente

Gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas.

1.5.3. Indicadores

A. Conceptualización

Variable Independiente: La incorporación elementos de inteligencia de negocios.

Tabla 1.1 Indicadores para la Variable Independiente

Indicadores	Descripción
Ausencia	Cuando es No , los elementos de inteligencia de negocios son inexistentes y aún nos encontramos en la situación actual del problema.
Presencia	Cuando es Sí , se incorporó los elementos de inteligencia de negocios, esperando obtener mejorar la gestión de la información

Elaborado por el equipo de trabajo

Variable Dependiente: Gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas.

Tabla 1.2: Indicadores para la Variable Dependiente

Indicadores	Descripción
Exactitud de la información	Es el número de personas que creen que la información es completa.
Tiempo en realizar los reportes	Es el tiempo que se utiliza para realizar los reportes.
Número de reportes solicitados por ciclo	Es el número de reportes solicitados por el Coordinador académico y/o instructores.
Comprensión de los reportes	Es el entendimiento de los reportes.

Elaborado por el equipo de trabajo

B. Operacionalización

Variable Independiente: La incorporación elementos de inteligencia de negocios.

Tabla 1.3: Operacionalización de la Variable Independiente

Indicadores	Índice
Ausencia	No
Presencia	Sí

Elaborado por el equipo de trabajo

Variable Dependiente: Gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas.

Tabla 1.4: Operacionalización de la Variable Independiente

Indicadores	Índice
Exactitud de la información	Completa, Incompleta
Tiempo en realizar los reportes	Minutos
Número de reportes solicitados por ciclo	# de reportes solicitados /ciclo
Comprensión de los reportes	Comprensible, Incomprensible

Elaborado por el equipo de trabajo

1.6. Limitaciones

- La presente investigación no contó con mucho acceso a la información sobre el Área de Académica del CELEN.
- El tiempo para la limpieza de la base de datos del Área Académica del CELEN.

CAPÍTULO II

REVISIÓN DE LITERATURA

2.1. Antecedentes

Para consolidar el presente Proyecto de Investigación se toma en consideración los siguientes Trabajos de investigación relacionados al tema los que se mencionan a continuación.

(Fuentes & Pinto, 2010: 383) En su trabajo de tesis titulado “Incorporación de Elementos de Inteligencia de Negocios en el Proceso de Admisión y Matrícula de una Universidad Chilena”. Describe un proceso orientado a la incorporación de elementos de inteligencia de negocios (Business Intelligence - BI) en la Universidad de Tarapacá (UTA), Arica, Chile. Con este fin, se implementó un Data Mart (DM) centrado en el área de Admisión y Matrícula de la Vicerrectoría Académica.

Su desarrollo requirió de la realización de actividades tales como la obtención de los requerimientos del negocio, la investigación del indicador clave de rendimiento (KPI) del área, el análisis de las distintas fuentes de información interna y el desarrollo de un modelado dimensional basado en el esquema estrella de Kimball. Para la correcta implementación e integración de este repositorio de datos se debió realizar un proceso de extracción, transformación y carga (ETL) a partir de dos fuentes de datos.

La creación de este DM permitió que los usuarios de la Vicerrectoría Académica pudieran visualizar la información que requerían a través de

herramientas de procesamiento analítico en línea (OLAP). Complementándose, además, con herramientas para la generación de reportes y herramientas para la creación de dashboards. La integración de estos elementos conformó una plataforma de inteligencia de negocios, que permite dar soporte a los requerimientos de información y análisis asociados al proceso de admisión y matrícula.

(Arenas & Montes, 2017: 9) En su trabajo de tesis titulado “Inteligencia de Negocios aplicada a los Procesos de Autoevaluación de la Universidad de Manizales”. El proyecto contiene un conjunto de procedimientos y técnicas, que, desde la inteligencia de negocios, apoyan los procesos de autoevaluación institucional de la Universidad de Manizales.

El objetivo de este proyecto fue diseñar una solución que proporcione calidad a la presentación de los datos y que a partir de hechos e información argumentada sirva como un apoyo a la toma de decisiones, iniciando con el levantamiento de la información, análisis de fuentes de datos, creación de los reportes o informes diseñados a partir de los indicadores cuantitativos que permitirán la toma de decisiones e identificación de necesidades o fortalezas a lo largo de los procesos de autoevaluación que se definen continuamente por la institución.

Al tener la información y los datos conectados correctamente se obtuvo informes y comportamientos que representan y gestionan los grandes volúmenes de información.

(Catañeda Vásquez 2015: 4) En su trabajo de tesis titulado “Desarrollo de Business Intelligence, basado en la Metodología de Ralph Kimball, para mejorar

el Proceso de Toma de Decisiones en el Área de Admisión de la Universidad Autónoma del Perú”. La investigación trató sobre la mejora de un proceso de negocios, específicamente sobre el pobre desempeño del Proceso de Toma de Decisiones del Área de Admisión de la Universidad Autónoma del Perú, en lo que corresponde a: Porcentaje de exactitud de información, tiempo en realizar los reportes, porcentaje de malas decisiones tomadas, número de reportes solicitados por ciclo, comprensión de los reportes, utilizando Business Intelligence.

Para darle una solución al problema se propuso realizar un Business Intelligence que nos permite almacenar sólo la información requerida e ir eliminando aquellos datos que obstaculizan la labor del análisis y entregando la información que se requiera en la forma más apropiada, facilitando así el proceso de gestión de la misma.

Business Intelligence surge con la promesa del manejo y control de la información, asegura una vista única de los datos, que pueden provenir de diversas fuentes. Existen diferentes herramientas BI tanto software libre como propietario, para esta solución de la Universidad se usó software con licencia y después de un análisis en cuanto a herramientas de este tipo se utilizó SQL Server 2014.

(Guillén Quisca, 2016: 13) En su trabajo de tesis titulado “Sistema de Soporte de Decisiones con Tecnología Data Warehouse para la Gestión de la Información de la Empresa Mallku Import SAC – Juliaca”. La investigación tuvo por finalidad el desarrollo de un Sistema de Soporte de Decisiones con Tecnología Data Warehouse que optimice la Gestión de la Información de la Empresa Mallku Import SAC, la empresa en mención tiene su centro de operaciones en la ciudad de Juliaca, región de Puno y se dedica al rubro de la importación de máquinas

remanufacturadas como son fotocopiadoras, impresoras, repuestos, insumos y accesorios; donde sus operaciones diarias son la venta de estos productos. Encontrando como uno de los principales problemas la mala gestión de la información, el desaprovechamiento de información almacenada a diario, no generando información a partir de esto.

Por lo cual se planteó un sistema de soporte de decisiones basado en inteligencia de negocios que tendrá como fin solucionar este problema. Para el desarrollo del software se utilizó la metodología de Moss Larissa que es basada en Inteligencia de Negocios de Roadmap y el enfoque de Ralph Kimball para la construcción del Data Warehouse y como herramientas de software se utilizó la suite de Inteligencia de Negocios “Pentaho” que viene a ser un conjunto de programas libres para generar inteligencia empresarial escrita en Java con un ambiente de implementación basado en Java, para evitar costos altísimos en las licencias de software.

Finalmente se tiene como principal conclusión, que el sistema desarrollado ha optimizado la gestión de la información en un más del 83%, logrando validar y cumplir con el objetivo principal de la investigación.

(Quispe Paco, 2012: 9). “Sistema de Inteligencia de Negocios para el Soporte de Toma de Decisiones en el Área de Nutrición y Aprendizaje Infantil del Programa Nacional Wawa Wasi sede Puno”. El proyecto de tesis tuvo por objetivo dar soporte a la toma de decisiones en el área de nutrición y aprendizaje infantil en niñas y niños en primera infancia del Programa Nacional Wawa Wasi Sede Puno, para cumplir con objetivos específicos según los “Lineamientos Técnicos para la Atención Integral en Salud en el Servicio Wawa Wasi”.

Los estudios como antecedentes con relación a este tipo de proyecto con el pasar de los años han ido cobrando importancia, ya que un sistema de inteligencia de negocios ayuda grandemente en el soporte de toma de decisiones. El sistema de inteligencia de negocios dará el soporte de toma de decisiones en el área de nutrición y aprendizaje infantil del Programa Nacional Wawa Wasi Sede Puno.

Se logró determinar cuál es la cantidad de niñas y niños en situación de desnutrición aguda (peso/talla) y desnutrición crónica (talla/edad), así como la estimulación en Aprendizaje Infantil temprano no se está logrando los cuales son factores determinantes en el desarrollo del niño de la primera infancia, para tomar las decisiones adecuadas recurriendo a las líneas de acción para su correcto desarrollo, de acuerdo a los lineamientos que posee en el Programa Nacional Wawa Wasi. Se logró implementar el Sistema de Inteligencia de Negocios, para el soporte de toma de decisiones, el cual brinda alertas oportunas para la toma de decisiones y propuesta de mejora continua de la calidad de servicio que el Programa Nacional Wawa Wasi ofrece.

(Acero Calizaya, 2013: 10) “Data Mart de Contrataciones Públicas a partir del SEACE, y su aplicación en la Toma de Decisiones de las Micro y Pequeñas Empresas de la ciudad de Puno”. El trabajo investigación tuvo como objetivo general determinar el grado de influencia de la aplicación de la Data Mart de Contrataciones públicas en la toma de decisiones de las MYPEs.

Durante el desarrollo de la investigación se abordó los conceptos de análisis, exploración, información histórica, procesos selección, contratación de bienes y servicios.

Se utilizó la metodología HEFESTO para el desarrollo del Data Mart, desde los requerimientos hasta la difusión de los resultados, permitiendo que los usuarios finales (MYPEs) analicen y exploren la información de manera sencilla e intuitiva. El Data Mart de Contrataciones públicas contiene el comportamiento de los precios, cantidades, montos adjudicados de contratación de bienes y servicios por entidad, montos adjudicados por proveedor (MYPE), la información es presentado en función al tiempo.

Para el despliegue de la solución Data Mart se utilizó la herramienta Pentaho con capacidad de personalizar reportes de acuerdo a las necesidades del negocio y MySQL como motor de base de datos.

La investigación siguió el diseño de investigación pre test y post test, donde se demostró estadísticamente la hipótesis planteado y como resultado final se obtuvo que la aplicación del Data Mart de Contrataciones públicas a partir del SEACE apoya significativamente la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno.

2.2. Gestión de la Información

Gestión de la información es la denominación convencional de un conjunto de procesos por los cuales se controla el ciclo de vida de la información, desde su obtención (por creación o captura), hasta su disposición final (su archivo o eliminación). Tales procesos también comprenden la extracción, combinación, depuración y distribución de la información a los interesados. El objetivo de la

gestión de la información es garantizar la integridad, disponibilidad y confidencialidad de la información. (Arévalo, 2018: 5)

La Gestión de la Información mantiene una estrecha relación con la disciplina de la Gestión del Conocimiento en el contexto organizacional. Los objetivos de la Gestión de Información se centran en aquellos procesos relacionados con el almacenamiento, el tratamiento y la difusión del conocimiento explícito que se encuentra representado en los documentos. Sin embargo, en este contexto, la Gestión del Conocimiento iría un poco más allá que la Gestión de la Información. Ésta se encargaría de convertir todo el conocimiento en conocimiento corporativo y de difundirlo adecuadamente. Se ocuparía, principalmente, de las decisiones pragmáticas y estratégicas relativas a la creación, la identificación, la captura, el almacenamiento y la difusión el conocimiento integrado en una organización. Y, el desarrollo de estas operaciones se implementaría en sintonía con la dimensión humana de esos procesos y respetando y rediseñando los elementos organizativos necesarios. (Pérez-Montoro, 2009: 1)

2.2.1. Indicadores

Existen una serie de conceptos de indicador; no obstante, el concepto de indicador varía de unos autores a otros. Para unos cumple una función informativa, para otros es un instrumento en la toma de decisiones, o bien cumple una función de evaluación.

La definición más simple de indicador corresponde a la identificación de una magnitud numérica referida a un evento, que pone en evidencia la intensidad, situación o evolución del mismo. De su

tratamiento es posible establecer la explicación, evolución y predicción de un fenómeno estudiado. Constituye un instrumento que pueden permitir evaluar de manera objetiva aspectos particulares del proceso de ejecución o de los resultados de un programa o proyecto a través de mediciones de carácter cualitativo o cuantitativo. (Contreras, 2018: 6)

Características

Para definir un buen Indicador, es importante desarrollar un criterio para la selección de los indicadores que deberán controlarse en forma continua, ya que el seguimiento tiene un alto costo cuando no está soportado por un verdadero beneficio. Para esto se puede utilizar una sencilla técnica que consiste en responder cuatro preguntas básicas:

- ¿Es fácil de medir?
- ¿Se mide rápidamente?
- ¿Proporciona información relevante en pocas palabras?
- ¿Se grafica fácilmente?

Funciones Básicas

- Simplificación
- Cuantificación
- Comunicación

Dimensiones

- Cualitativas: Descripción de la variable
- Cuantitativas: Expresión porcentual, numérica, promedio, número absoluto

Representación

- Cifra absoluta: Refleja características particulares en un momento dado (Por ejemplo: tiempo en realizar reportes o número de reportes solicitados en un ciclo)
- Porcentajes: Se refiere fundamentalmente a aspectos de distribución (Por ejemplo: Porcentaje de la exactitud de la información).
- Promedios y Otras Medidas estadísticas: Representa comportamientos típicos (Por ejemplo: Comprensión de reportes).

Los indicadores requieren ser construidos tomando en consideración el espacio, el contexto, la especificidad, etc. Pueden ser:

- Simples: Cuando se refieren a una sola variable.
- Complejos: Cuando se refieren a una relación de dos o más variables.

2.3. Inteligencia de Negocios

Una interesante definición para inteligencia de negocios o BI, por sus siglas en inglés, según el Data Warehouse Institute, lo define como la combinación de tecnología, herramientas y procesos que me permiten transformar datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial. La inteligencia de negocios debe ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa

y la capacidad de tomar buenas decisiones para así obtener mejores resultados.
(Oracle, 2018: 6)

Es el Proceso mediante el cual las organizaciones segmentan, unifican, jerarquizan y definen contextos sobre la información, con la finalidad de identificar la información más relevante y representativa para la organización, que, junto con sus correspondientes herramientas, hacen centro en el análisis de la información para la correcta toma de decisiones que le permita a la organización cumplir con los objetivos de negocios. (Nima, 2009: 84)

Ventajas

- Crear escenarios favorables tendientes a lograr una óptima toma de decisiones.
- Ayuda a mejorar la eficiencia operacional. Al reestructurar la información con esta herramienta se reducen los tiempos de procesamiento y respuesta de las bases de datos, ya que todo está en un solo punto y organizado para cada área con los datos más representativos que el negocio requiere.
- El usuario final no necesita tener conocimientos para la generación de nuevos informes.
- Evita los largos tiempos de respuesta en las consultas de datos complejos que implican la unión de tablas operacionales de gran tamaño, lo que a su vez representa un tiempo de espera que obstruye la fluidez del trabajo.
- Evita datos erróneos, obsoletos, ya que éstos han sido previamente analizados y depurados, lo que garantiza calidad y fiabilidad de la información obtenida. (Curto, 2010: 20)

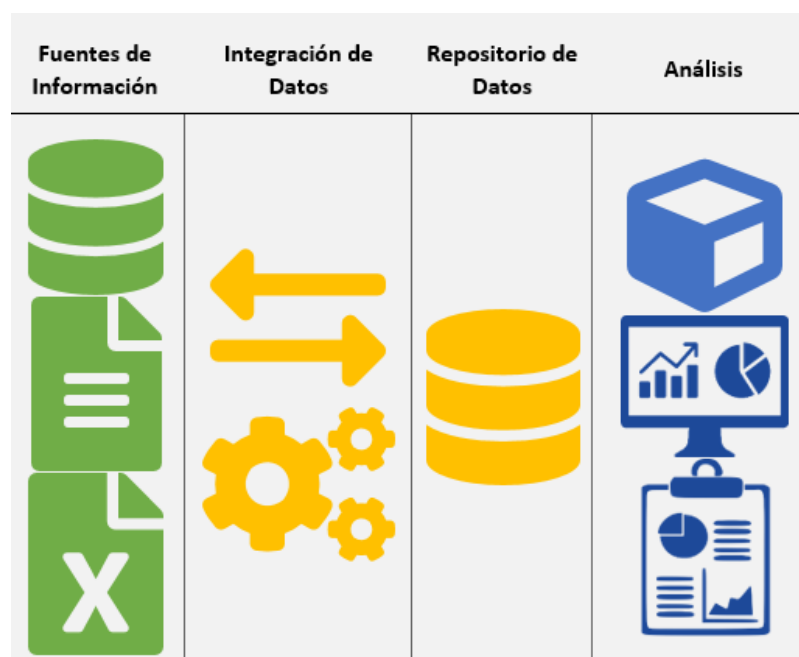
Desventajas

- Saturación en los Data Warehouse ya que crecen los datos de manera desproporcionada porque los técnicos del área no consiguen decir no a las excesivas demandas de los usuarios.
- Extracción, transformación y carga de la información cuando los datos de origen no están limpios, existe duplicidad, caracteres erróneos, lo cual implica un Proceso más costoso, de mayor tamaño y con menor rendimiento. (Nima, 2009: 84)

Elementos

Los elementos de Inteligencia de Negocios están compuestos de diferentes tecnologías que se integran para formar una solución empresarial. Los elementos están orientados a la transformación de los datos en información oportuna y confiable para el usuario final. (Kimball & Ross, 2013: 37)

Figura 2.1: Elementos de BI



FUENTE: (Fuentes & Valdivia, 2010: 384)

Funciones de la base de Datos Relacional

Son base de datos que almacena la información en forma de tabla en dos Dimensiones, creando en forma de uniones, relaciones entre estas tablas; también llamada simplemente relacional.

- Se encarga de realizar la depuración y homogenización de los datos.
- Almacenamiento de datos (hasta el nivel de detalle máximo y recogiendo datos actuales e históricos).

Funciones de la Base de Datos Multidimensional

Son bases de datos ideadas para desarrollar aplicaciones muy concretas, como creación de cubos OLAP (Procesamiento Analítico en Línea). La diferencia entre las bases de datos relacionales es a nivel conceptual; en las bases de datos multidimensionales los campos o atributos de una tabla pueden ser de dos tipos, o bien representan Dimensiones de la tabla, o bien representan métricas que se desean estudiar. (Kimball & Ross, 2013: 40)

- Velocidad de acceso y consulta.
- Capacidad de análisis desde varios puntos de vista (Dimensiones de análisis).

2.4. Proceso de Extracción, Transformación y Carga

El proceso ETL (Extract, Transform and Load, por sus siglas en inglés) proporciona a las empresas y organizaciones la posibilidad de mover, cambiar y unificar sus datos a partir de diversas fuentes para que sean cargados en otra fuente de datos. De esta manera se trata del proceso que permite a las organizaciones

mover datos desde múltiples fuentes, reformatearlos y cargarlos en otra base de datos (denominada Data Mart o Data Warehouse) con el objeto de analizarlos. También pueden ser enviados a otro sistema operacional para apoyar un proceso de negocio. (PowerData, 2018: 1)

Extraer: Para la extracción es conveniente tener en cuenta el manejo adecuado de las diferentes fuentes de datos, conocer su estructura y los diferentes componentes y especificaciones necesarias para que sean conectadas de forma efectiva. Otro punto que hay que tener en cuenta es que a la hora de realizar la tarea de extracción es que la misma cause un impacto mínimo en los sistemas de origen. Si la cantidad de datos a extraer es muy elevada, el sistema se puede ralentizar, o incluso colapsarse, por lo que las grandes operaciones de extracción se suelen realizar en momentos donde el impacto sobre el sistema sea el mínimo posible. (Basantes & López, 2012: 18)

Transformar: Para la etapa de transformación, es importante un ajuste y limpieza en los datos, con el fin de tener estos de forma precisa, completa, y coherente. Es decir, se tiene una limpieza en los datos inicialmente, seguida de su transformación, para que finalmente estén totalmente integrados y compatibles.

En la tarea de transformación se aplican una serie de funciones sobre los datos extraídos al objeto de convertirlos en datos preparados para su carga. Algunas fuentes de datos tan solo requerirán mínimas transformaciones, mientras que otras necesitarán de un gran número de ellas. Entre las operaciones de transformación podemos encontrar las siguientes:

- Obtención de valores calculados.
- Generación de nuevos campos.

- División de la información.
- Unión de datos de múltiples fuentes

Cargar: Como su nombre lo indica, en esta fase son cargados los datos que fueron extraídos y transformados con anterioridad, y de esta manera, son escritos en las estructuras dimensionales para ser accedidos por los usuarios finales y las aplicaciones. Así mismo y dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos. Los data Warehouse mantienen un historial de los registros de manera que se pueda hacer una auditoría de los mismos y disponer de un rastro de toda la historia de un valor a lo largo del tiempo. (Bernabeu, 2010: 26)

2.5. Data Mart

Un Data Mart (DM) es un conjunto de datos estructurados que contiene la información referente a un área, tema o función en particular. Se caracterizan por disponer de la información para analizarla al detalle desde todas las perspectivas que afecten un proceso. Los DMs pueden ser de diversas bases de datos dependiendo del tipo de análisis que se quiera desarrollar. Un DM es considerado un elemento de un Data Warehouse (DW), pero con información de un área en específico de la organización. (Curto, 2010: 33)

Ventajas

- Un Data Mart soporta menos usuarios que un Data Warehouse, dado esto se puede optimizar para recuperar más rápidamente los datos que necesitan los usuarios.

- Menores cantidades de datos implican que se procesen antes, tanto las cargas de datos como las consultas.
- La aplicación cliente, que pide la consulta es independiente del servidor que la procesa y del servidor de bases de datos que almacenan la información.
- Los costos que implica la construcción de un Data Mart son mucho menores a los de la implementación de un Data Warehouse.

Desventajas

- No permite el manejo de grandes volúmenes de información por lo que muchas veces se debe recurrir a un conjunto de Data Mart para cubrir todas las necesidades de información de la empresa. (Basantes & López, 2012: 25)

Esquema de Estrella

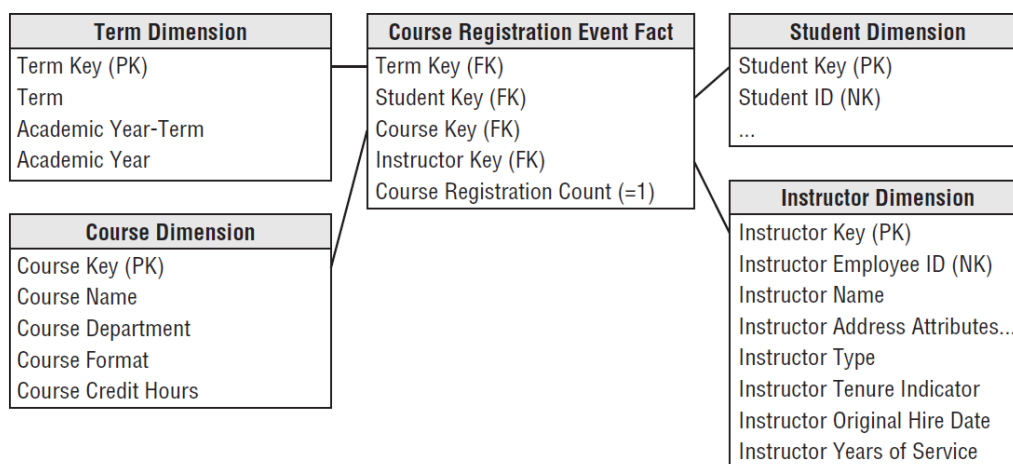
“Para facilitar el análisis, el DM organiza los datos en una estructura llamada Esquema de Estrella. Esta estructura está compuesta por una tabla central (tabla de hechos) y un conjunto de tablas organizadas alrededor de ésta (tabla de dimensiones) (Kimball & Ross, 2013: 40).”

En las puntas de la estrella se encuentran las tablas de dimensión que contiene los atributos de las aperturas que interesan al negocio que se pueden utilizar como criterios de filtro y son relativamente pequeñas. Cada tabla de dimensión se vincula con la tabla de hechos por un identificador. Las características de un esquema de estrella son: (Curto, 2010: 95)

- El centro de la estrella es la tabla de hecho.

- Los puntos de la estrella son las tablas de dimensiones.
- Cada esquema está compuesto por una sola tabla de hechos.
- Generalmente es un esquema totalmente des normalizado, pudiendo estar parcialmente normalizado en las tablas de dimensiones.

Figura 2.2: Esquema del Modelo Estrella



FUENTE: (Kimball & Ross, 2013: 331)

Tabla 2.1: Diferencias entre el Esquema Estrella y Copo de Nieve

	Esquema Estrella	Copo Nieve
Almacenamiento	Aumenta el espacio	Ahorra el espacio
Consultas	Mejora el Rendimiento	Aumenta la cantidad de uniones entre tabla provocando bajo rendimiento.
Cantidad de Tablas	Menor	Mayor

FUENTE: (Catañeda, 2015: 27)

2.5.1. Tabla de Hechos

“La tabla de hechos es la tabla primaria del modelo dimensional, y contiene los valores del negocio que se desea analizar, cada tabla de hechos

contiene las claves externas que se relacionan con sus respectivas tablas de dimensiones” (Bernabeu, 2010: 30).

El modelo dimensional divide el mundo de los datos en dos grandes tipos: las medidas y las dimensiones de estas medidas. Las medidas, siempre son numéricas, se almacenan en las tablas de hechos y las dimensiones que son textuales se almacenan en las tablas de dimensiones.

Un hecho es un concepto de interés primario para el Proceso de toma de decisiones, corresponde a eventos que ocurren dinámicamente en el negocio de la empresa.

2.5.2. Dimensiones

Las dimensiones son las que organizan los datos en función de un área de interés para los usuarios, cada dimensión describe un aspecto del negocio y proporciona el acceso intuitivo y simple a datos, una dimensión provee al usuario de un gran número de combinaciones e intersecciones. (Kimball & Ross, 2013: 46)

Las tablas de dimensiones son las compañeras de las tablas de hechos, cada dimensión se define por su clave primaria que sirve para mantener la integridad referencial en la tabla de hechos a la que se relaciona. Un cubo requiere que se defina al menos una dimensión en su empresa.

2.5.3. Medidas

Las medidas dentro de la construcción de un DM son los valores de datos que se analizan. Estas medidas pueden ser cuantitativas o numéricas, dentro de la tabla de hechos, permitiendo analizarlos ya que estos valores son la base de las cuales el usuario puede realizar cálculos. (Curto, 2010: 140)

2.6. Data Warehouse

Un Data Warehouse (DW) no es más que la unión de los almacenes de datos departamentales (DM), integrado, no volátil y variable en el tiempo que facilita tener una visión general de los datos de una organización facilitando tener la información estable, coherente, fiable e información histórica. (Kimball & Ross, 2013: 31)

Ventajas

- Se puede ejecutar o procesar una gran cantidad de información. Rapidez y flexibilidad al momento de acceder a la información.
- Mejora la toma de decisiones y el aumento de la productividad en una empresa, puesto que permite conocer los resultados, sean estos positivos o negativos, evaluando de esta forma los tiempos de respuesta y los costos de operaciones.
- Es fiable la comunicación entre los diferentes departamentos de la empresa.

Desventajas

- El DW realiza una reestructuración de los sistemas operacionales, lo que implica altos costos.
- Es necesario de aplicaciones, sistemas, y almacenamientos específicos.
- Presenta revisiones continuas de los modelos de datos, objetos, transacciones, lo que provoca que el diseño sea complejo.

Tabla 2.2: Diferencias entre Bases de Datos Transaccionales y Almacenes de Datos

Parámetros	Base de Datos	Almacén de Datos
Propósito	Operaciones diarias. Soporte a las aplicaciones.	Recuperación de información, informes, análisis y minería de datos.
Tipo de Datos	Datos de funcionamiento de la organización.	Datos útiles para el análisis, la sumarización, etc.
Características de los Datos	Datos de funcionamiento, cambiantes, internos.	Datos históricos, datos internos y externos, datos descriptivos.
Modelo de Datos	Datos normalizados.	Datos en estrella, en copo de nieve, parcialmente desnormalizados. Multidimensionales.
Número y Tipo de Usuarios	Cientos/miles: aplicaciones, operarios, administrador de la base de datos.	Directores, ejecutivos, analistas.
Acceso	SQL. Lectura y escritura.	SQL y herramientas propias (slice & dice, drill, roll, pivot). Lectura.

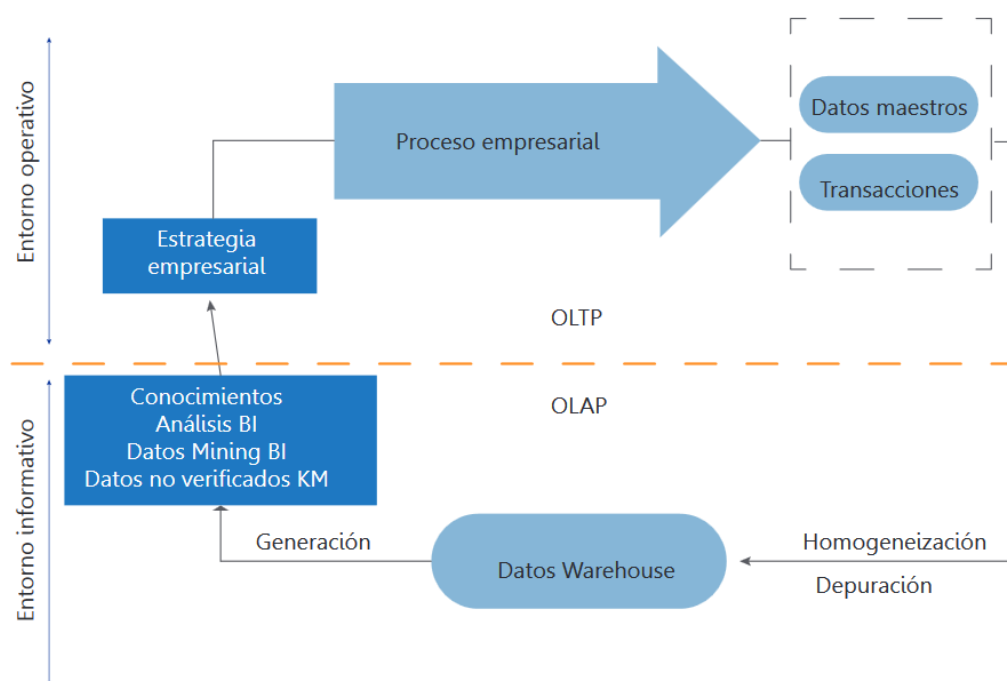
FUENTE: (Catañeda, 2015: 23)

2.7. Tecnologías OLTP Y OLAP

La plataforma de Inteligencia de Negocios permite a las compañías ampliar la información para que solo los datos necesarios sean utilizados con fines específicos. A través de plataformas de Inteligencia de Negocios, una empresa

puede gestionar todo y analizar los datos vitales necesarios para su funcionamiento, estos datos podrían orientar a las empresas sobre adónde dirigir sus esfuerzos corporativos en términos de mercado, decisiones emergentes, tendencias de la industria y responder a las exigencias y especificaciones del cliente. Por ello, tener una solución para Inteligencia de Negocios hoy en día, esencial en cualquier negocio o empresa. (Murillo & Cáceres, 2013: 128)

Figura 2.3: Interacción de los Sistemas OLTP y OLAP



FUENTE: (Murillo & Cáceres, 2013: 128)

2.7.1. Sistemas OLTP

“Estos tipos de sistemas de Procesamiento Transaccional en Línea son los sistemas operacionales que capturan las transacciones de los Procesos de una empresa, y las persisten en estructuras relacionales llamadas base de datos” (Rivadera, 2010: 56).

Características

- Realizar transacciones en tiempo real del Proceso de un negocio.
- Estos sistemas son los responsables del mantenimiento de los datos, ya sea agregando, realizando, actualizaciones o bien eliminándolos.
- Las estructuras de datos deben estar depuradas para validar la entrada de los mismos, y rechazarlos si no cumplen con determinadas reglas.

2.7.2. Sistemas OLAP

“Los sistemas de Procesamiento Analítico en Línea proporcionan una alternativa a los sistemas transaccionales, ofreciendo una visión de los datos orientada hacia el análisis y una rápida y flexible navegación de estos” (Salaki & Waworuntu, 2016: 123).

Características

- Los cubos de OLAP almacenan varios niveles de datos conformados por estructuras altamente optimizadas que responden a las expectativas de negocios de la empresa.
- Los sistemas OLAP están preparados para realizar informes complejos de una manera simple.
- Los usuarios pueden cambiar fácilmente las filas, las columnas, y las páginas en informes de OLAP, pudiendo leer la información de la manera que se crea más conveniente para el análisis.

2.8. Pentaho

Pentaho es un software libre bajo varias licencias, entre ellas la licencia GPLv2. Esta licencia implica que se puede usar, copiar, distribuir y modificar gratuitamente. Esto es un punto a favor de este software ya que posibilita la implantación de un sistema BI bajo un desembolso nulo en licencias de software. Hay que tener en cuenta que los precios de licencias de software privativo para este tipo de soluciones de Business Intelligence son caras debido a su complejo desarrollo. No cualquier empresa se puede permitir la adquisición de este tipo de licencias al empezar su actividad en este campo del análisis de datos. (Acero: 2014, 54)

Además de ofrecer la versión gratuita, Pentaho Community Edition, los creadores de Pentaho ofrecen una versión de pago llamada Pentaho BI Suite Enterprise Edition. Esta versión es más completa que la que ofrece con licencia de código abierto (Open Source), y destaca por tener un servicio de soporte al usuario para resolver cualquier tipo de incidencia con el software. Pentaho ha sido creado bajo las plataformas libres de Java y MySQL. Ambas plataformas estaban en auge y cada vez más gente hacia uso de ellas. Según el ranking de los lenguajes de programación de 2011 elaborada por el portal de tecnología Eweek, el lenguaje de programación JAVA fue el lenguaje más usado en aquel entonces. (Durá: 2011, 14)

La comunidad de Pentaho es muy amplia. Todo software antes de ser liberado oficialmente, cuando aún está en la fase beta, pasa por una serie de pruebas realizadas por la comunidad de Pentaho de forma altruista. Usuarios de la aplicación que reportan fallos y ponen a prueba todo tipo de opciones disponibles del software. Además, se cuenta con una wiki en constante actualización que

orienta a los usuarios en posibles problemas o dudas que pudieran tener a la hora del uso o implantación del software. (Guillén: 2017: 60)

Pentaho no es una sola aplicación, sino es una recopilación de programas específicos según las funciones ofertadas (Anexo 8). Pentaho oferta una serie de programas para realizar las siguientes funciones:

- **Informes:** permite crear a los usuarios y personalizar informes además de exportar en multitud de formatos.
- **Análisis:** a través de las potentes herramientas ayuda al usuario a ampliar la perspectiva y mejorar la toma de decisiones en el negocio.
- **Cuadros de mando (dashboards):** ofrece al usuario final opciones de visualización en tiempo real de los datos, gestionados por 'displays' interactivos y últimas tecnologías multimedia para facilitar la interacción con la herramienta y manejo de datos.
- **Integración de datos:** dispone de aplicaciones potentes para la extracción, transformación y carga de datos en diversas plataformas de bases de datos.

Todos los programas desarrollados por Pentaho son creados y diseñados con independencia del resto y en esquema modular. Con ello se logra descargar, instalar o hacer uso de algunas de las aplicaciones con independencia del resto, haciendo posible una posible integración con otros programas ajenos a la Suite.

La solución Pentaho es una alternativa a las soluciones propietarias tradicionales más completas: Business Objects, Cognos, Microstrategy, Microsoft, IBM, etc., por lo que incluye todos aquellos componentes que se pueden encontrar en las soluciones Business Intelligence (BI) propietarias más avanzadas. (Quispe, 2012: 23)

Tabla 2.3: Comparativa de Softwares de Inteligencia de Negocios

Características	SAP Business Objects	Microsoft SQL Server Business Intelligence	pentaho A Hitachi Group Company
¿Forma parte de una plataforma integrada de Inteligencia de Negocios?	Sí	No	Sí
¿La herramienta de extracción posee una interface gráfica de uso?	Sí	No	Sí
¿Soporta diversos tipos de bases de datos?	No	Sí	Sí
¿Permite cargas desde ficheros Excel, XML y planos?	Sí	No	Sí
¿Requiere una fácil instalación de la herramienta?	Sí	Sí	No
¿La plataforma posee herramientas de reporting, herramientas de consultas y análisis?	Sí	No	Sí
¿Se encuentra fácilmente consultoras de sistemas para la herramienta?	Sí	Sí	Sí

FUENTE: (Catañeda, 2015: 32)

2.9. Dashboard

No solo OLAP ayuda a visualizar la información del negocio, existen diversas técnicas que facilitan el análisis de los datos DM. Estas incluyen reportes de datos y dashboards, siendo este último una gran herramienta para visualizar el estado actual de indicadores cuantitativos del entorno de la organización en tiempo real. (Fuentes & Valdivia, 2010: 384)

“Los dashboards realizados ofrecen gráficos en tiempo real y diagramas basados en pronósticos sobre el futuro de la organización. La potencialidad y velocidad de los dashboards hace de ellos un componente popular y poderoso en el entorno ejecutivo” (Marcus, 2006: 48).

2.10. Metodología de Bill Inmon

Según Bill Inmon, propone que un DW es el conjunto de datos orientados a temas específicos de la organización, es decir deberán ir cambiando con el transcurso del tiempo, pero dichos datos, no deben ser volátiles (no es posible eliminar, ni modificar la información). Al momento de generar un cambio en los datos, se debe de efectuar de una manera que, quede reflejado el cambio que se efectuó y mantener siempre la integridad de la información, se espera que todo resultado del proceso y transformación de la información por medio de herramientas de BI, sea un firme apoyo para los tomadores de decisiones al momento de orientar una estrategia de negocio. (Alfaro, Medrano, & Nolasco, 2016: 37)

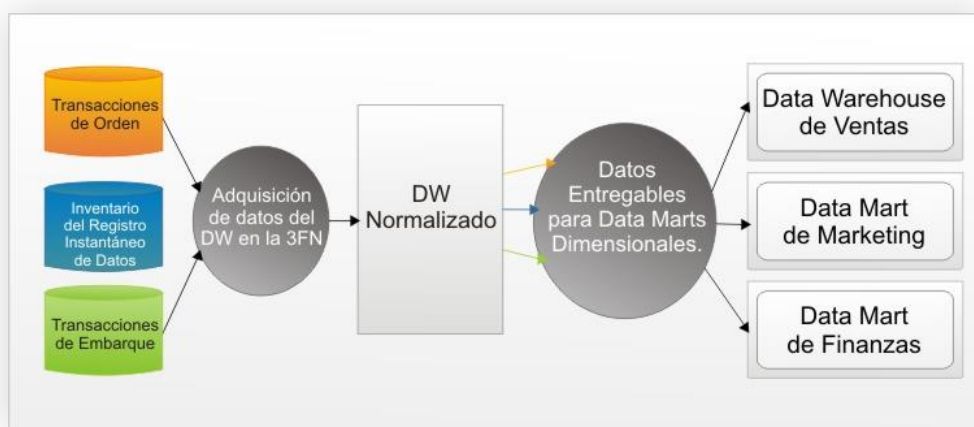
Bill Inmon, considera que debido al exceso y sobrecarga de información que se está manejando en un DW, estas bases de datos deben estar aisladas y solo manipularse para procesos de BI. Toda información que se almacena en un DW o DM deberá ser procesada y normalizada antes. La recomendación sobre la información, deberá estar a su máximo detalle posible, con lo cual se logre cubrir todas las necesidades de cada departamento, la importancia a la necesidad de transferir la información de los diferentes OLTP de las organizaciones a un lugar centralizado donde los datos sean utilizados para el análisis. (Espinosa, 2010: 1)

Características

- **Orientado a temas:** Los datos en la base de datos están estructurados de forma que los datos estén relacionados.
- **Integrado:** La base de datos contiene todos los datos de los sistemas operacionales de la organización, con la característica de ser consistentes.

- **No volátil:** Los datos no se modifica ni se elimina, almacenado los datos, éste se convierte en información de sólo lectura.
- **Variante en el tiempo:** Los datos que sufren un cambio a través del tiempo, deben ser registrados para que los informes generados reflejen esas variaciones.

Figura 2.4: Arquitectura de Bill Inmon



FUENTE: (Espinosa, 2010: 3)

2.11. Metodología de Ralph Kimball

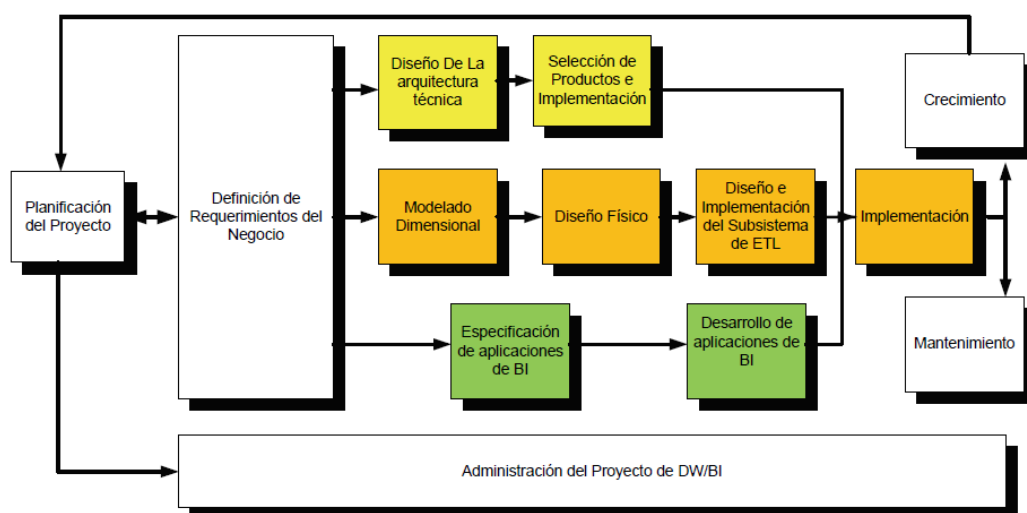
La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio. Este ciclo de vida del proyecto de DW, está basado en cuatro principios básicos:

- **Centrarse en el negocio:** Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- **Construir una infraestructura de información adecuada:** Diseñar una base de información única, integrada, fácil de usar, de alto rendimiento

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

- **Realizar entregas en incrementos significativos:** crear el almacén de datos en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos. En esto la metodología se parece a las metodologías ágiles de construcción de software.
- **Ofrecer la solución completa:** proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación. (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008: 17)

Figura 2.5: Ciclo de Vida de la Metodología de Kimball



FUENTE: (Rivadera, 2010: 59)

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. Análisis del Negocio y Obtención de Requerimientos e Indicadores

El Centro de Estudios de Lenguas Extranjeras y Nativas de la Universidad Nacional del Altiplano de Puno (CELEN de la UNA Puno) ha sido creado de acuerdo al Estatuto de la UNA, aprobado con R.R. N°475-85-UNA, con fecha 09 de mayo de 1985. Desde 1999 viene prestando servicios a la comunidad en el Edificio Universitario del Parque Pino.

Actualmente la directora del Centro de Estudios de Lenguas Extranjeras y Nativas es la Dra. Delicia Vilma Gonzales Arestegui, cuenta con 49 Docentes y 05 Trabajadores Administrativos. Este Centro ofrece al público en general la enseñanza de los idiomas extranjeros: inglés, francés, portugués, italiano, alemán, e idiomas nativos como el quechua y aimara (Anexo 6).

Tabla 3.1: Cursos y Duración por Idioma

Idioma	Duración	Detalle
AIMARA	12 MESES	<ul style="list-style-type: none"> – 4 BÁSICOS – 4 INTERMEDIOS – 4 AVANZADOS
ALEMÁN	18 MESES	<ul style="list-style-type: none"> – 6 BÁSICOS – 6 INTERMEDIOS – 6 SUPERIORES
FRANCÉS	20 MESES	<ul style="list-style-type: none"> – 5 ELEMENTALES – 5 INTERMEDIOS – 5 SUPERIORES – 5 SUPERIORES
INGLÉS	33 MESES	<ul style="list-style-type: none"> – 12 BÁSICOS – 06 PRE-INTERMEDIOS – 05 INTERMEDIOS – 05 SUPERIORES – 05 AVANZADOS

ITALIANO	14 MESES	<ul style="list-style-type: none"> - 5 BÁSICOS - 5 INTERMEDIOS - 4 AVANZADOS
PORTUGUÉS	12 MESES	<ul style="list-style-type: none"> - 3 BÁSICOS - 2 INTERMEDIOS - 2 AVANZADOS - 3 AVANZADOS SUPERIORES - 2 PERFECCIONAMIENTOS
QUECHUA	12 MESES	<ul style="list-style-type: none"> - 4 BÁSICOS - 4 INTERMEDIOS - 4 AVANZADOS

FUENTE: CELEN

Tabla 3.2: Horarios por Turno

Turno	Horario
MAÑANAS	- 06:30 - 08:00 A.M.
	- 08:00 - 09:30 A.M.
	- 09:30 - 11:00 A.M.
	- 11:00 - 12:30 P.M.
TARDES	- 15:00 - 16:30 P.M.
	- 16:30 - 18:00 P.M.
NOCHE	- 18:00 - 19:30 P.M.
	- 19:30 - 21:00 P.M.
	- 21:00 - 22:30 P.M.

FUENTE: CELEN

Dirección y Ubicación: Jr. Lima Nº 317 / Esq. Parque Pino – Puno, Perú.

Figura 3.1: Ubicación del CELEN



FUENTE: Google Maps

Misión

Ser la primera institución en el Altiplano en brindar servicios académicos en la enseñanza de idiomas nativos y extranjeros, contribuyendo a la preparación de profesionales que pertenecen a diversos grupos económico-sociales, mediante una solvencia funcional de calidad activa para una comunicación multilingüe e intercultural.

Visión

Ser una institución autónoma y de mayor aceptación en el mercado, orientada a convertirse en la Facultad de Idiomas de la UNA Puno. Que será una puerta de acceso al mundo en base a convenios interinstitucionales, con prestigio y reconocimiento social, brindando opciones de especialización con estándares de acreditación internacional.

Competidores

Figura 3.2: Principales Competidores del CELEN



Elaborado por el equipo de trabajo

3.1.1. Proceso del Negocio Académico

Matriculas

A) Alumnos Nuevos

Se considera Alumno Nuevo, a aquel que por primera vez se matricula al CELEN de la UNA Puno.

- Pago en el Banco de la Nación por Tasas Educativas CELEN-UNA, para estudiantes universitarios de la UNA Puno el monto es de S/ 30.60 y para estudiantes Particulares es de S/ 45.60.
- Registro de matrícula vía internet todos los idiomas, niveles y horarios solo para estudiantes regulares de la UNA Puno (unapvirtual.unap.edu.pe/siu/idiomas/).
- Entrega de baucher de pago a ventanillas del CELEN en horario de atención.

También se considera Alumno Nuevo a aquel estudiante que proviene de otras instituciones debidamente acreditado, debiéndose someter a un EXAMEN DE NIVELACIÓN, realizando previo pago de S/ 15.00 para dicho Examen. (Deberá seguir los mismos procedimientos señalados anteriormente)

B) Alumnos Reingresantes y No Secuenciales

Se considera Alumno Reingresante a aquel estudiante que ha dejado de estudiar por más de 03 y menor a 06 meses, y deberá realizar REFRESCAMIENTO en el último ciclo aprobado, debiéndolo aprobar obligatoriamente para continuar sus estudios.

Alumno No Secuencial, es aquel estudiante que ha dejado de estudiar por más de 06 meses, y deberá someterse a un EXAMEN DE SUFICIENCIA (Examen Escrito y Oral), realizando previo pago de S/ 15.00 para dicho examen, cuya nota final es solo referente para reubicación mas no se considerará en actas; el puntaje final indica:

- De 80 a 100 puntos: Promovido al ciclo inmediato superior.
- De 60 a 79 puntos: Refreshamiento del último ciclo aprobado.
- De 40 a 59 puntos: Refreshamiento de los dos últimos ciclos aprobados.
- De 01 a 39 puntos: Refreshamiento de los tres últimos ciclos aprobados.

Registro de Notas

Los alumnos matriculados en un ciclo regular para el idioma ingles son evaluación según los siguientes criterios:

Tabla 3.3: Criterios de Evaluación para el Idioma Ingles

Criterio		Puntaje
ASISTENCIA		5
PUNTUACIÓN ESCRITA	EXAMEN PARCIAL	25
	EXAMEN FINAL	25
	WORKBOOK	5
PUNTUACIÓN ORAL	EXAMEN ORAL	20
	PARTICIPACIÓN EN CLASE	5
	PLATAFORMA VIRTUAL	15
PUNTUACIÓN FINAL		100

FUENTE: CELEN

Solo los alumnos que al menos hayan cumplido con el 75% de asistencia durante el ciclo matriculado y además hayan obtenido un puntaje final (Final Score) mayor o igual a 75 puntos serán promovidos al siguiente ciclo según

corresponda. Para los alumnos con menos del 75% de asistencia tendrán la oportunidad de justificar sus inasistencias previa solicitud antes de culminar el ciclo. Los alumnos que hayan obtenido un puntaje final menor a 75 puntos tendrán que repetir el ciclo sin lugar a reclamos.

3.1.2. Requerimientos Funcionales

Para comprender de mejor los requerimientos del CELEN, y en particular de su área de académica, se realizó un análisis e investigación de su entorno. Este proceso permitió adquirir los requerimientos del negocio y sus indicadores. Además, se definieron las distintas fuentes de información que fueron utilizadas. Todas estas tareas tuvieron como fin el comprender el ámbito del negocio para la correcta implementación del esquema dimensional (Anexo 9).

Tabla 3.4: Lista de Requerimientos Funcionales

Requerimiento	Prioridad	Dificultad
<ul style="list-style-type: none"> – Se requiere mostrar su código del CELEN, código de la UNA Puno en caso fuera estudiante, sexo, nombres y apellidos. – Obtener el detalle del último registro del estudiante. – Realizar un seguimiento histórico del rendimiento de cada estudiante por año. – Realizar un detalle de estos puntajes del estudiante por curso año. – Reportes del estudiante por Especialidad, Curso, Docente y Horario 	1	3
<ul style="list-style-type: none"> – Se requiere conocer su código de docente, sexo, nombres y apellidos. 	2	1
<ul style="list-style-type: none"> – Se requiere conocer la(s) especialidades y cursos tomados. 	1	2

– Se requiere conocer la cantidad de horarios tomados por año.	3	1
– Se requiere mostrar la información por año, trimestre, mes.	1	2

Elaborado por el equipo de trabajo

Donde:

Tabla 3.5: Calificación y Significado de la Prioridad y Dificultad de los Requerimientos

Calificación	Significado (Prioridad)	Significado (Dificultad)
1	Importante	Fácil
2	Regular	Regular
3	Prescindible	Difícil

Elaborado por el equipo de trabajo

3.1.3. Requerimientos No Funcionales

La seguridad del sistema es un factor crítico pues la información que se maneja es sensible es decir representa información acerca del rendimiento académico del estudiante y no debe ser manipulada por usuarios no autorizados. Como el software que se utilizará es de código abierto, no se incurrirá en gastos por licencias de software; sin embargo, se requiere un servidor de aplicaciones, un servidor de base de datos y un dominio.

La página de administración de Pentaho, herramienta de BI, consiste en un (user console) aplicativo web que se utilizará para análisis del DM, es desde allí que el usuario con suficiente privilegio generará los reportes y realizará la evaluación de las estrategias.

Exportación de reportes a archivos de formatos estándar como archivo de PDF (.pdf) o Excel (.xls) y PNG (.png) o JPG (.jpg) en caso de gráficos.

Mostrar la información del DM en base a los siguientes 2 tipos: Reportes tradicionales (basados en columnas), y gráficos estadísticos.

El sistema al ser de consultas y análisis no impactará a las operaciones diarias que se realizan en el CELEN. El dashboard debe ser intuitivo y sin demasiados controles además debe adaptarse a cualquier tipo de pantalla de un dispositivo.

3.1.4. Estado Actual de los Indicadores

Tabla 3.6: Datos Actuales de los Indicadores

Indicadores		Datos de Pre-Test (GC)
Exactitud de la Información	Completa	60.0 %
	Incompleta	40.0 %
Tiempo en realizar los reportes	Promedio	62 minutos
Número de reportes solicitados por ciclo	Promedio	3 reportes
Comprensión de los reportes	Comprensible	66.7 %
	Incomprensible	33.3 %

Elaborado por el equipo de trabajo

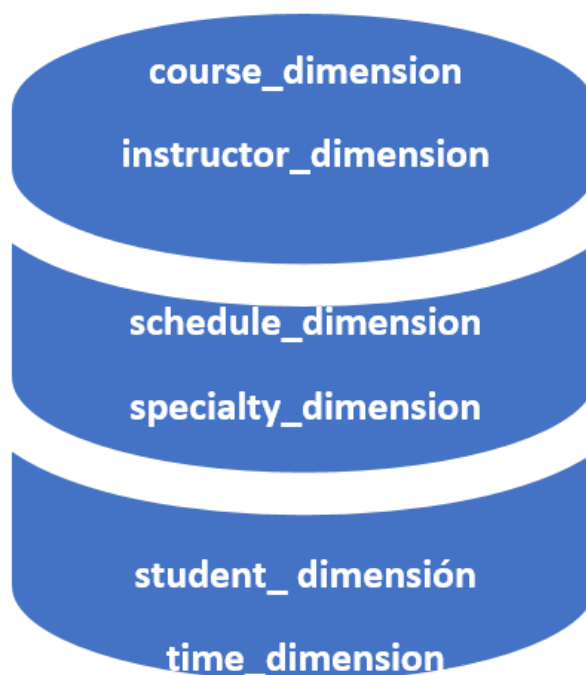
3.2. Identificación de las Fuentes de Datos

Figura 3.3: Tablas de la Base de Datos Transaccional



Elaborado por el equipo de trabajo

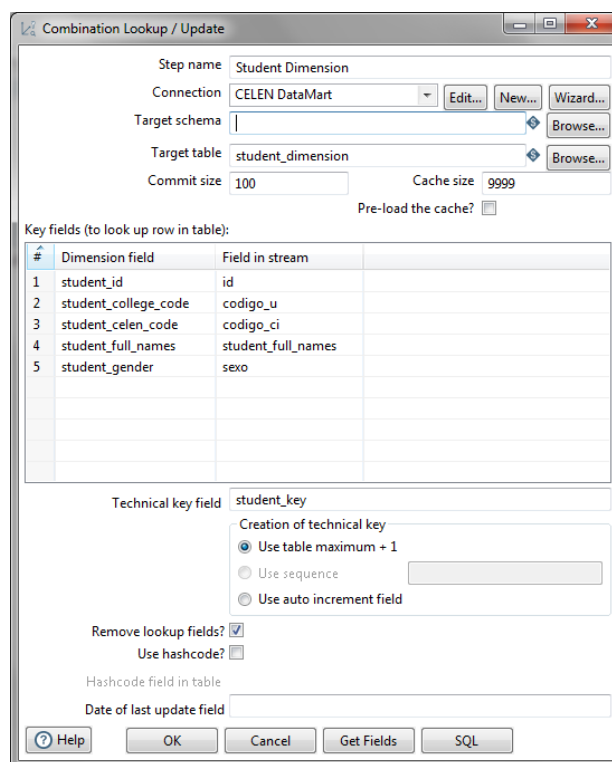
Figura 3.4: Tablas para la Base de Datos Multidimensional



Elaborado por el equipo de trabajo

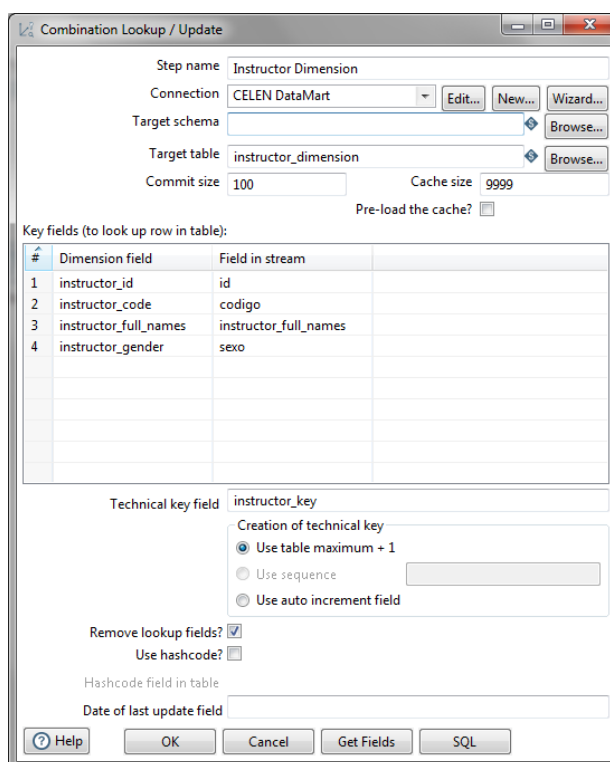
3.2.1. Fuente y Destino de los Datos

Figura 3.5: Fuente Tabla Estudiantes - Destino Dimensión Estudiante



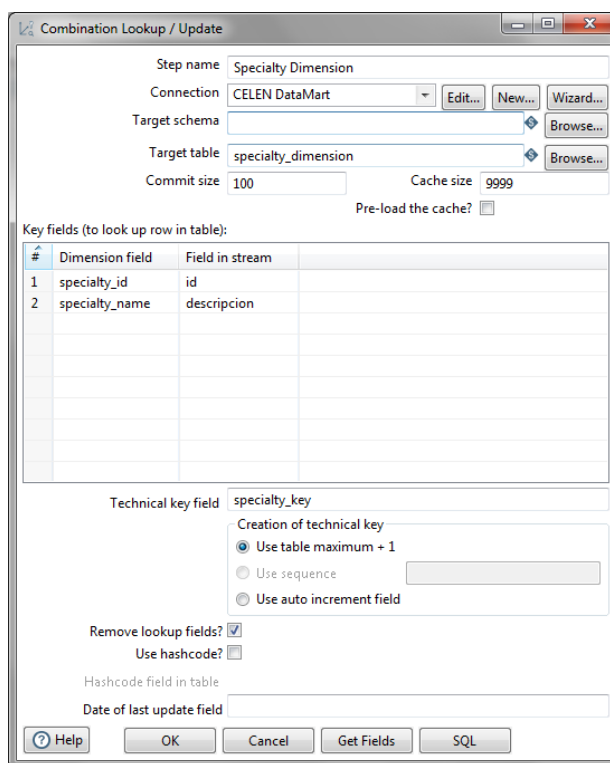
Elaborado por el equipo de trabajo

Figura 3.6: Fuente Tabla Docentes - Destino Dimensión Instructor



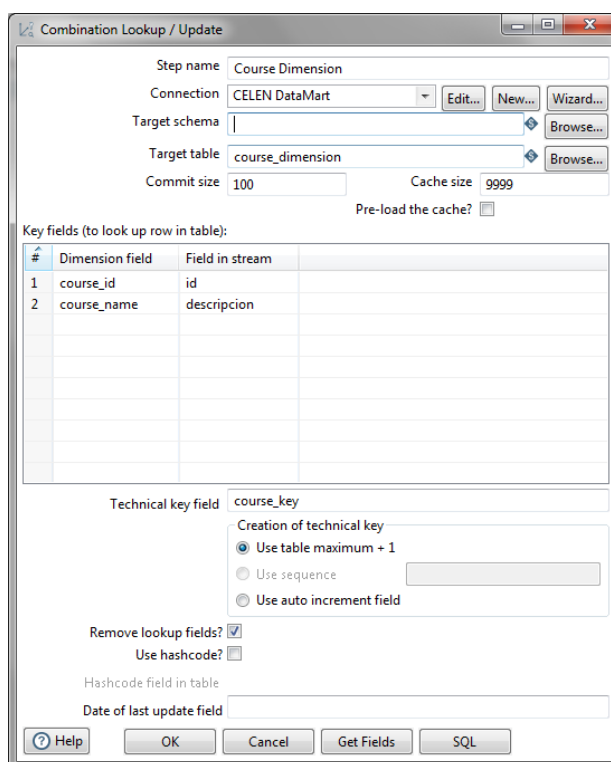
Elaborado por el equipo de trabajo

Figura 3.7: Fuente Tabla Especialidades - Destino Dimensión Especialidad



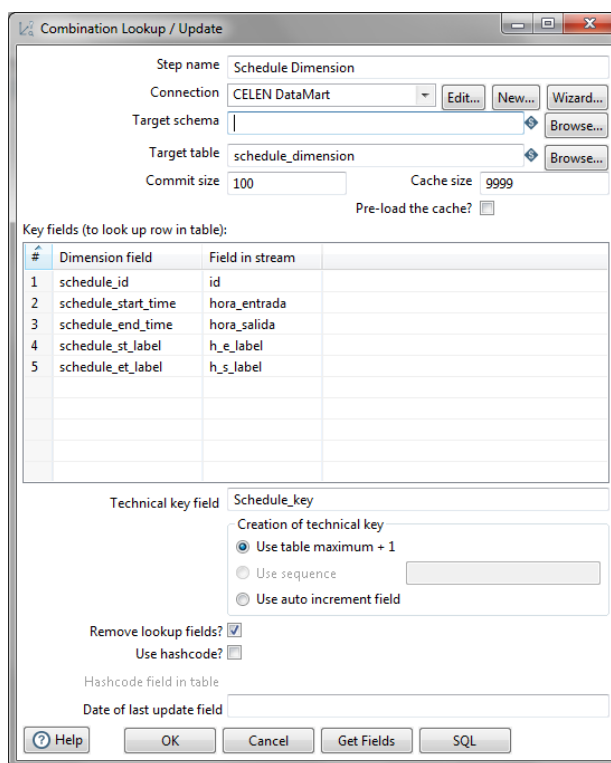
Elaborado por el equipo de trabajo

Figura 3.8: Fuente Tabla Cursos - Destino Dimensión Curso



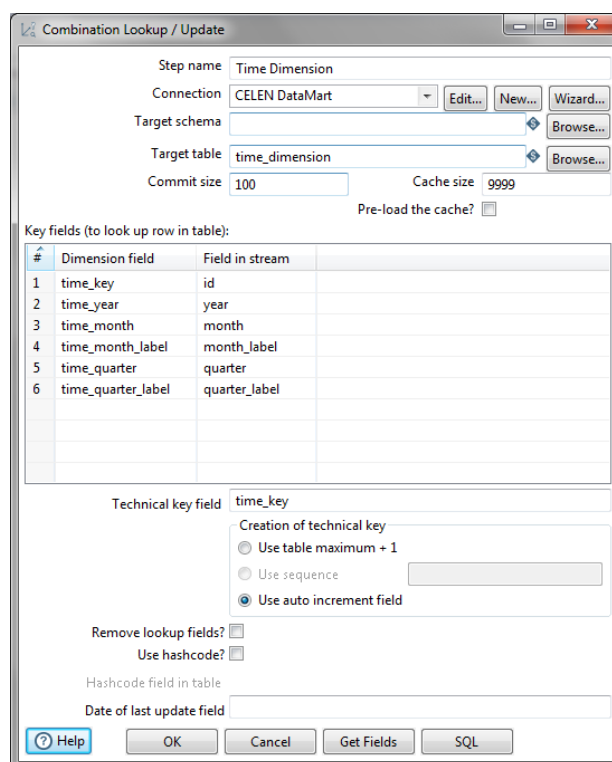
Elaborado por el equipo de trabajo

Figura 3.9: Fuente Tabla Horarios - Destino Dimensión Horario



Elaborado por el equipo de trabajo

Figura 3.10: Fuente Tabla Registro de Notas - Destino Dimensión Tiempo



Elaborado por el equipo de trabajo

3.3. Diseño del Data Mart

Para el desarrollo de la construcción de una solución de Inteligencia de Negocios se sigue la Metodología estudiada de Ralph Kimball, dado que establece claros procesos para todo el ciclo del desarrollo del proyecto y garantiza la calidad y eficiencia de la solución de inteligencia de negocios.

El modelo utilizado para soportar esta solución se basó en un esquema estrella. En general, un esquema estrella está constituido por un conjunto de tablas que representan dimensiones y una tabla de hechos. Las dimensiones sirven para detallar los valores que se encuentran asociados a la tabla de hechos que registran las medidas o los estados sobre un acontecimiento o una situación.

3.3.1. Antigüedad de los Datos

La antigüedad de los datos es tomada según la fecha de la tabla registro de notas de la base de datos transaccional, de las cuales se obtuvo desde el año 2013 hasta el 2018.

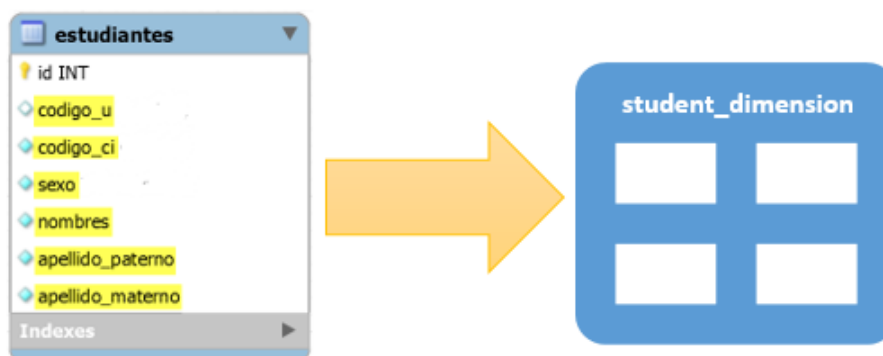
3.3.2. Definición de Dimensiones

Tabla 3.7: Definición de Dimensiones

Dimensiones	Descripción
student_dimension	Contiene la información básica del estudiante tales como código del CELEN, código de la UNA Puno, sexo, nombres y apellidos.
instructor_dimension	Contiene la información básica del instructor o docente tales como código, sexo, nombres y apellidos.
time_dimension	Contiene la información del periodo que existe en el año, los semestres y meses.
schedule_dimension	Contiene la información de los horarios existentes durante un ciclo.
course_dimension	Contiene la descripción del curso y su código.
specialty_dimension	Contiene la descripción de la especialidad y su código.

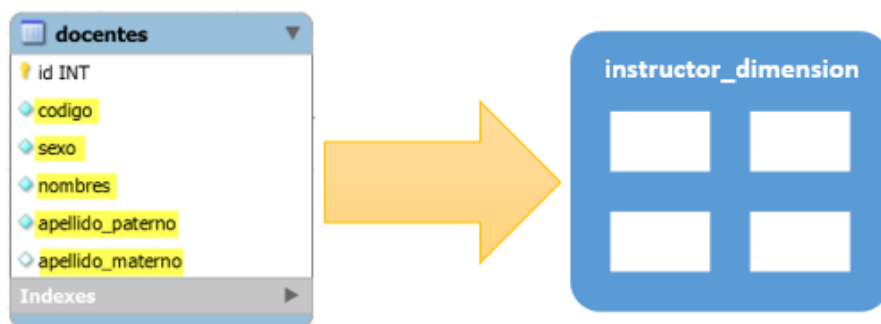
Elaborado por el equipo de trabajo

Figura 3.11: Mapeo de Datos para la Dimensión Estudiante



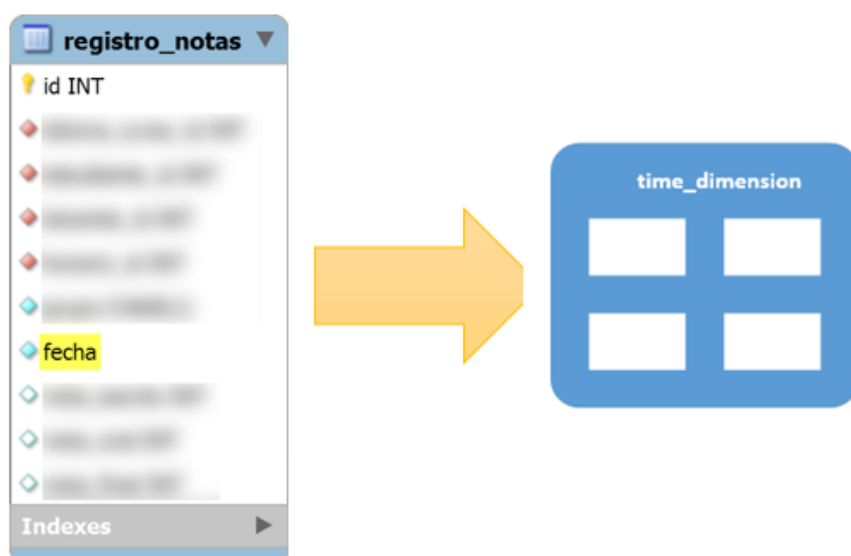
Elaborado por el equipo de trabajo

Figura 3.12: Mapeo de Datos para la Dimensión Docente



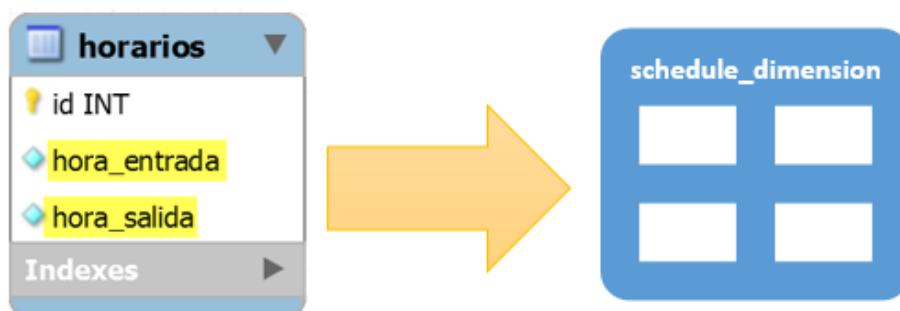
Elaborado por el equipo de trabajo

Figura 3.13: Mapeo de Datos para la Dimensión Tiempo



Elaborado por el equipo de trabajo

Figura 3.14: Mapeo de Datos para la Dimensión Horario



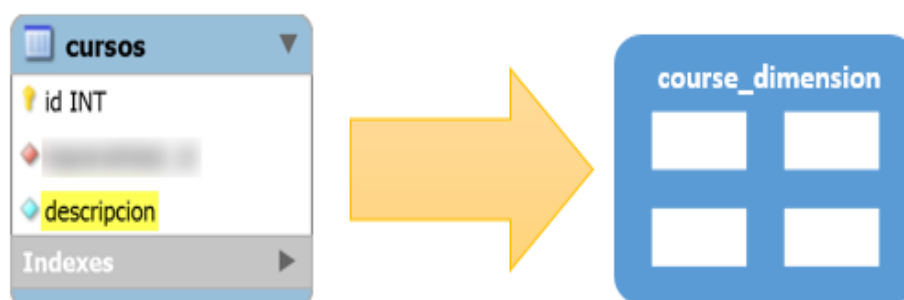
Elaborado por el equipo de trabajo

Figura 3.15: Mapeo de Datos para la Dimensión Especialidad



Elaborado por el equipo de trabajo

Figura 3.16: Mapeo de Datos para la Dimensión Curso



Elaborado por el equipo de trabajo

3.3.3. Jerarquías Analíticas y Granularidad

La granularidad significa especificar el nivel de detalle. La elección de la granularidad depende de los requerimientos del negocio y lo que es posible a partir de los datos actuales. La sugerencia general es comenzar a diseñar la solución de BI al mayor nivel de detalle posible, ya que se podría luego realizar agrupamientos al nivel deseado.

Tabla 3.8: Jerarquía y Granularidad de la Dimensión Estudiante

Nivel	Atributo	Descripción
1	student_college_code	Código de la UNA Puno
2	student_celen_code	Código del CELEN
3	student_full_names	Nombres y Apellidos
4	student_gender	Sexo

Elaborado por el equipo de trabajo

Tabla 3.9: Jerarquía y Granularidad de la Dimensión Docente

Nivel	Atributo	Descripción
1	instructor_code	Código
2	instructor_full_names	Nombres y Apellidos
3	instructor_gender	Sexo

Elaborado por el equipo de trabajo

Tabla 3.10: Jerarquía y Granularidad de la Dimensión Tiempo

Nivel	Atributo	Descripción
1	time_year	Año
2	time_month	Mes
3	time_month_label	Etiqueta del mes
4	time_quarter	Trimestre
5	time_quarter_label	Etiqueta del trimestre

Elaborado por el equipo de trabajo

Tabla 3.11: Jerarquía y Granularidad de la Dimensión Horario

Nivel	Atributo	Descripción
1	schedule_start_time	Hora de Inicio
2	schedule_st_label	Etiqueta de la hora inicio
3	schedule_end_time	Hora de Salida
4	schedule_et_label	Etiqueta de la hora salida

Elaborado por el equipo de trabajo

Tabla 3.12: Jerarquía y Granularidad de la Dimensión Especialidad

Nivel	Atributo	Descripción
1	specialty_name	Nombre de la especialidad

Elaborado por el equipo de trabajo

Tabla 3.13: Jerarquía y Granularidad de la Dimensión Curso

Nivel	Atributo	Descripción
1	course_name	Nombre del curso

Elaborado por el equipo de trabajo

3.3.4. Claves Sustitutas y Primarias

Una tabla dimensional está diseñada con una columna que sirve como llave primaria única. Esta llave primaria no puede ser la misma que la llave primaria natural de la base de datos transaccional porque las claves naturales para una

dimensión pueden ser creadas por más de una fuente, y estas claves naturales pueden ser incompatibles o mal administrados. Cada tabla de dimensional tiene una sola columna de clave primaria. Esta clave principal está incrustada como una clave externa en cualquier tabla de hechos asociada donde el contexto descriptivo de la fila de dimensión es exactamente correcto para esa fila de la tabla de hechos. Las tablas de dimensiones son generalmente planas, desnormalizadas con muchos atributos de texto de baja cardinalidad.

Tabla 3.14: Dimensión Estudiante con Atributos Jerárquicos

Atributo	Descripción	Formato	Llave Primaria
student_key	Llave primaria de la dimensión estudiante	Número entero	Sí
student_id	Llave primaria natural de la base de datos transaccional	Número entero	No
student_college_code	Código de la UNA Puno	Cadena de caracteres	No
student_celen_code	Código del CELEN	Cadena de caracteres	No
student_full_names	Nombres y Apellidos	Cadena de caracteres	No
student_gender	Sexo	Cadena de caracteres	No

Elaborado por el equipo de trabajo

Tabla 3.15: Dimensión Docente con Atributos Jerárquicos

Atributo	Descripción	Formato	Llave Primaria
instructor_key	Llave primaria de la dimensión docente	Número entero	Sí
instructor_id	Llave primaria natural de la base de datos transaccional	Número entero	No
instructor_code	Código	Cadena de caracteres	No
instructor_full_names	Nombres y Apellidos	Cadena de caracteres	No

instructor_gender	Sexo	Cadena de caracteres	No
-------------------	------	----------------------	----

Elaborado por el equipo de trabajo

Tabla 3.16: Dimensión Tiempo con Atributos Jerárquicos

Atributo	Descripción	Formato	Llave Primaria
time_key	Llave primaria de la dimensión periodo	Número entero	Sí
time_year	Año	Número entero	No
time_month	Mes	Número entero	No
time_month_label	Etiqueta del mes	Cadena de caracteres	No
time_quarter	Trimestre	Cadena de caracteres	No
time_quarter_label	Etiqueta del trimestre	Cadena de caracteres	No

Elaborado por el equipo de trabajo

Tabla 3.17: Dimensión Horario con Atributos Jerárquicos

Atributo	Descripción	Formato	Llave Primaria
schedule_key	Llave primaria de la dimensión horario	Número entero	Sí
schedule_id	Llave primaria natural de la base de datos transaccional	Número entero	No
schedule_start_time	Hora de Inicio	Hora	No
schedule_st_label	Etiqueta de la hora inicio	Texto	No
schedule_end_time	Hora de Salida	Hora	No
schedule_et_label	Etiqueta de la hora salida	Texto	no

Elaborado por el equipo de trabajo

Tabla 3.18: Dimensión Especialidad con Atributos Jerárquicos

Atributo	Descripción	Formato	Llave Primaria
specialty_key	Llave primaria de la dimensión especialidad	Número entero	Sí
specialty_id	Llave primaria natural de la base de datos transaccional	Número entero	No
specialty_name	Nombre de la especialidad	Cadena de caracteres	No

Elaborado por el equipo de trabajo

Tabla 3.19: Dimensión Curso con Atributos Jerárquicos

Atributo	Descripción	Formato	Llave Primaria
course_key	Llave primaria de la dimensión curso	Número entero	Sí
course_id	Llave primaria natural de la base de datos transaccional	Número entero	No
course_name	Nombre del curso	Cadena de caracteres	No

Elaborado por el equipo de trabajo

3.3.5. Definición de la Tabla de Hechos

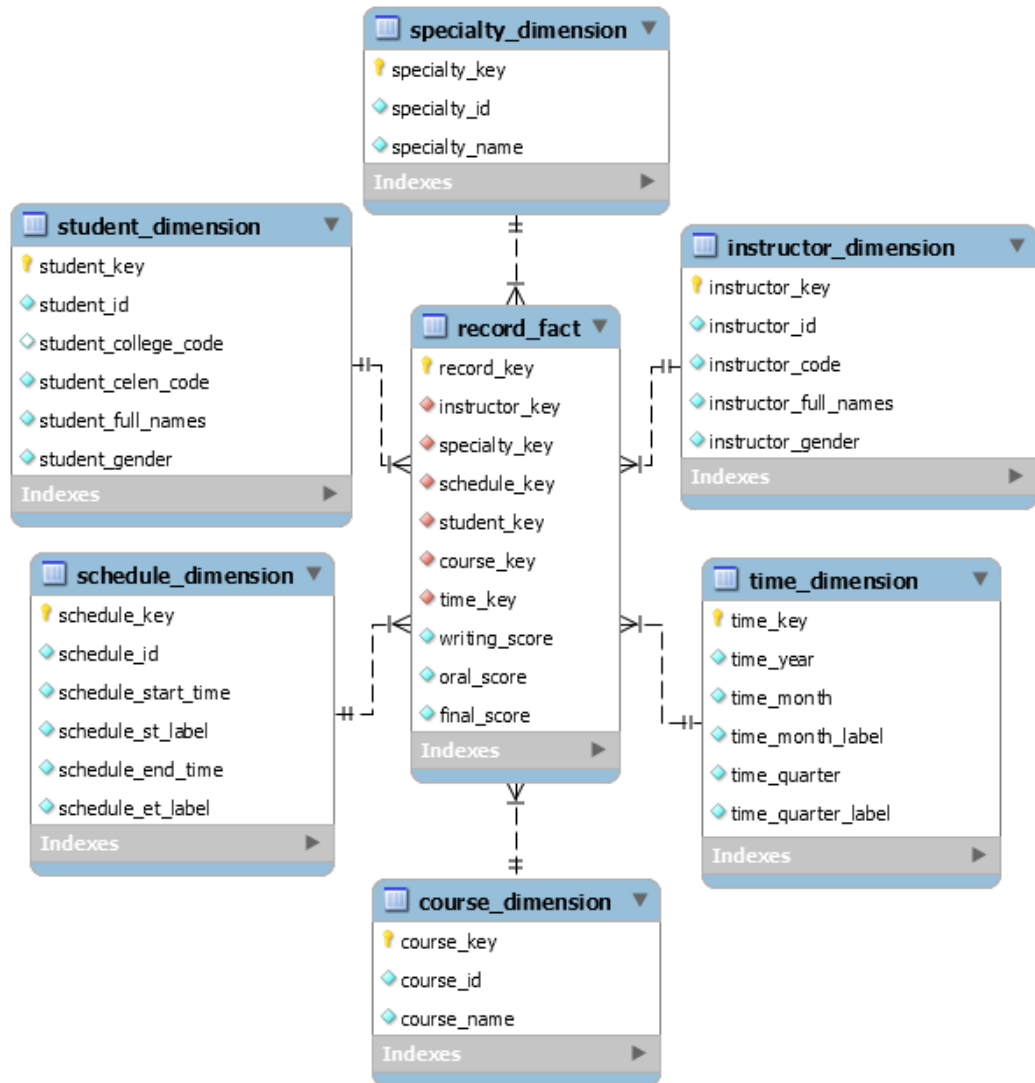
Contiene las llaves de las Dimensiones y las medidas que fueron levantadas en las etapas anteriores del proyecto de acuerdo a las especificaciones del equipo de requerimiento.

Tabla 3.20: Granularidad de la Tabla de Hechos

Atributo	Descripción	Formato	Llave Primaria
record_key	Llave primaria de la tabla de hechos.	Número entero	Sí
instructor_key	Llave primaria de la dimensión Instructor	Número entero	No
specialty_key	Llave primaria de la dimensión Especialidad	Número entero	No
schedule_key	Llave primaria de la dimensión Horario	Número entero	No
student_key	Llave primaria de la dimensión Instructor	Número entero	No
course_key	Llave primaria de la dimensión Curso	Número entero	No
time_key	Llave primaria de la dimensión Periodo	Número entero	No
writing_score	Puntaje Escrito del estudiante	Número entero	No
oral score	Puntaje Oral del estudiante	Número entero	No
final_score	Puntaje Final del estudiante	Número entero	No

Elaborado por el equipo de trabajo

Figura 3.17: Modelo Estrella, Relaciones entre Dimensiones y la Tabla de Hechos



Elaborado por el equipo de trabajo

3.3.6. Traslado del Modelo Dimensional al Modelo Lógico

El modelo de datos lógicos describe los datos con el mayor detalle posible, independientemente de cómo se implementarán físicamente en la base de datos.

Tabla 3.21: Modelo Dimensional Curso

Column	Type	Null
course_key (Primary)	int(10)	No
course_id	int(11)	No
course_name	varchar(100)	No

Elaborado por el equipo de trabajo

Tabla 3.22: Modelo Dimensional Instructor

Column	Type	Null
instructor_key (Primary)	int(10)	No
instructor_id	int(11)	No
instructor_code	varchar(7)	No
instructor_full_names	varchar(255)	No
instructor_gender	enum('MASCULINO', 'FEMENINO')	No

Elaborado por el equipo de trabajo

Tabla 3.23: Modelo Dimensional Horario

Column	Type	Null
schedule_key (Primary)	int(10)	No
schedule_id	int(11)	No
schedule_start_time	time	No
schedule_st_label	tinytext	No
schedule_end_time	time	No
schedule_et_label	tinytext	No

Elaborado por el equipo de trabajo

Tabla 3.24: Modelo Dimensional Especialidad

Column	Type	Null
specialty_key (Primary)	int(10)	No
specialty_id	int(11)	No
specialty_name	varchar(100)	No

Elaborado por el equipo de trabajo

Tabla 3.25: Modelo Dimensional Estudiante

Column	Type	Null
student_key (Primary)	int(10)	No
student_id	int(11)	No
student_college_code	char(6)	Yes
student_celen_code	char(7)	No
student_full_names	varchar(255)	No
student_gender	enum('MASCULINO', 'FEMENINO')	No

Elaborado por el equipo de trabajo

Tabla 3.26: Modelo Dimensional Tiempo

Column	Type	Null
time_key (Primary)	int(10)	No
time_year	int(11)	No
time_month	int(11)	No
time_month_label	tinytext	No
time_quarter	int(11)	No

time_quarter_label	char(2)	No
--------------------	---------	----

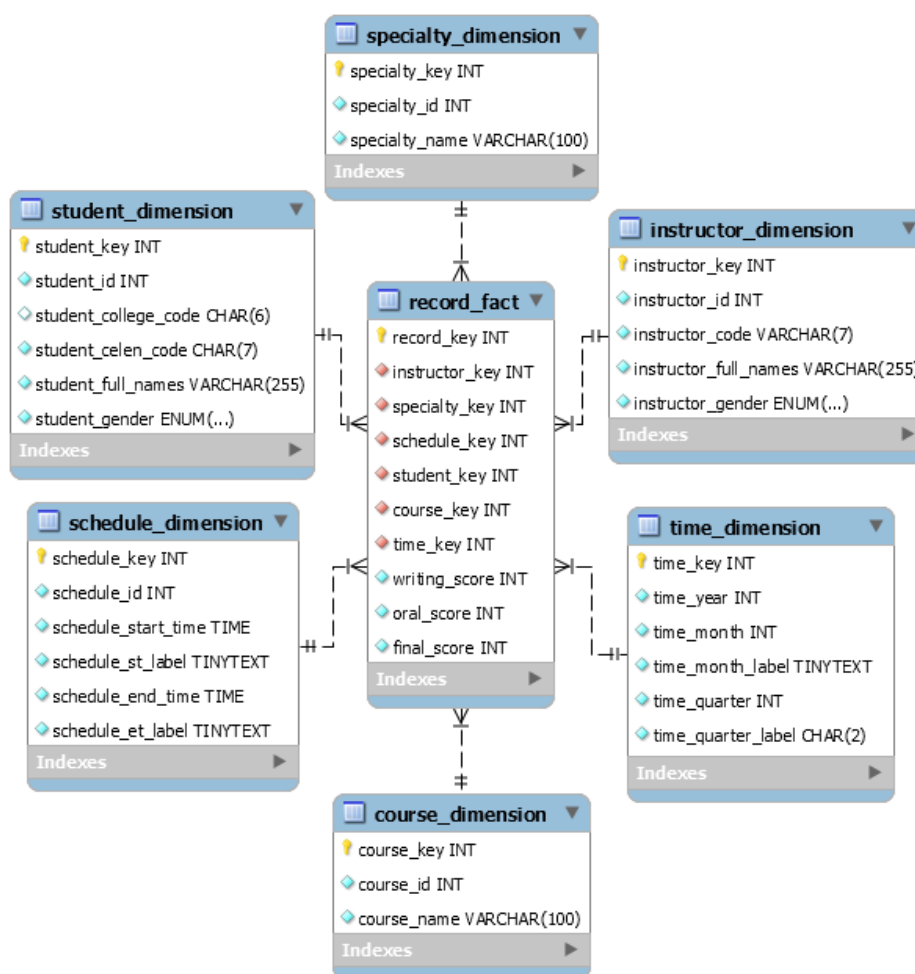
Elaborado por el equipo de trabajo

Tabla 3.27: Modelo Dimensional Récord Académico

Column	Type	Null
record_key (Primary)	int(11)	No
instructor_key	int(10)	No
specialty_key	int(10)	No
schedule_key	int(10)	No
student_key	int(10)	No
course_key	int(10)	No
time_key	int(10)	No
writing_score	int(11)	No
oral_score	int(11)	No
final_score	int(11)	No

Elaborado por el equipo de trabajo

Figura 3.18: Modelo Lógico Dimensional

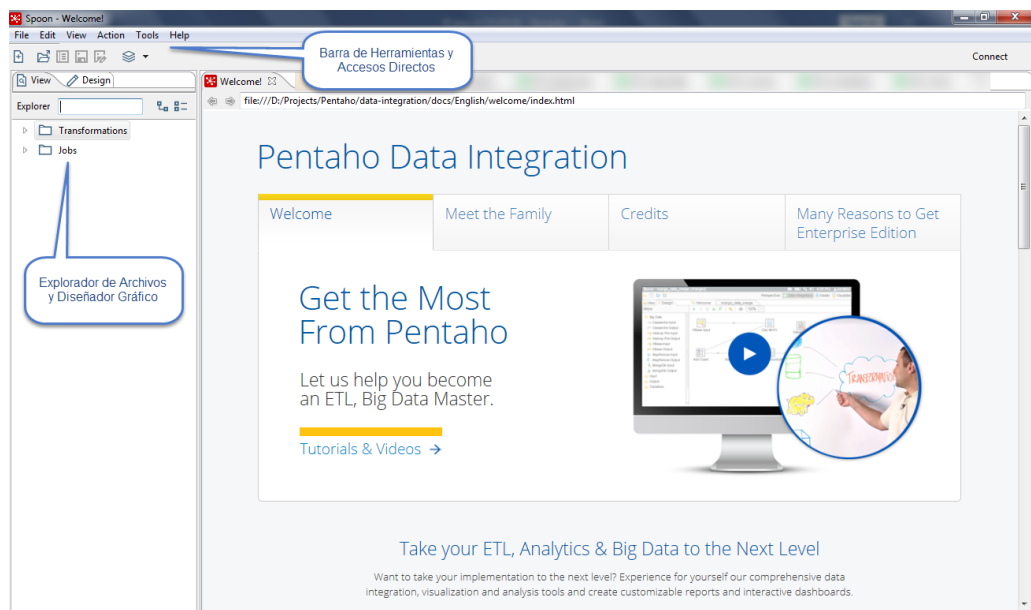


Elaborado por el equipo de trabajo

3.4. Implementación de Soluciones para la Visualización del Data Mart

La herramienta ETL que se usó es Pentaho Data Integration (Spoon) (Anexo 8) para lo cual se desarrolla los paquetes para la extracción, transformación y carga de los datos con origen en el OLTP y con destino en el OLAP.

Figura 3.19: Entorno de la Herramienta Gráfica PDI



Elaborado por el equipo de trabajo

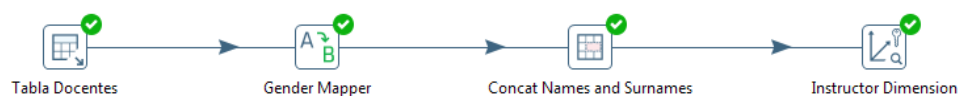
3.4.1. Carga de Dimensiones

Figura 3.20: Carga de Datos para la Dimensión Estudiante



Elaborado por el equipo de trabajo

Figura 3.21: Carga de Datos para la Dimensión Instructor



Elaborado por el equipo de trabajo

Figura 3.22: Carga de Datos para la Dimensión Especialidad



Elaborado por el equipo de trabajo

Figura 3.23: Carga de Datos para la Dimensión Curso



Elaborado por el equipo de trabajo

Figura 3.24: Carga de Datos para la Dimensión Horario



Elaborado por el equipo de trabajo

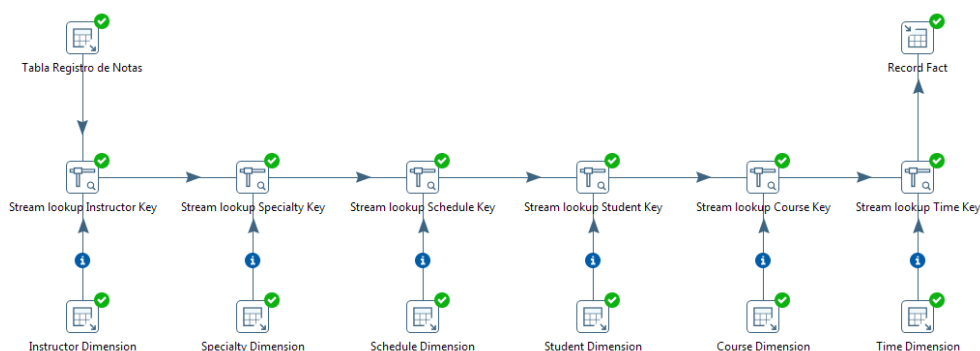
Figura 3.25: Carga de Datos para la Dimensión Tiempo



Elaborado por el equipo de trabajo

3.4.2. Carga de la Tabla de Hechos

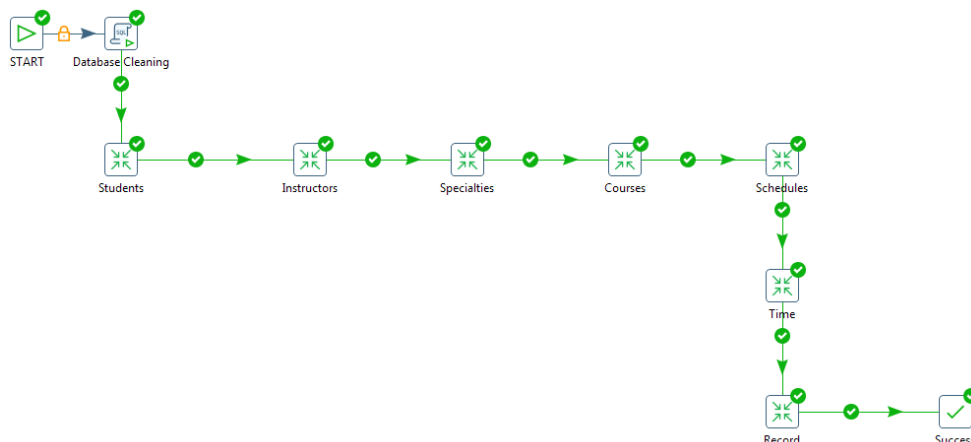
Figura 3.26: Carga de Datos para la Tabla de Hechos



Elaborado por el equipo de trabajo

3.4.3. Carga de Dimensiones y Tabla de Hechos

Figura 3.27: Carga de Dimensiones y Tabla de Hechos

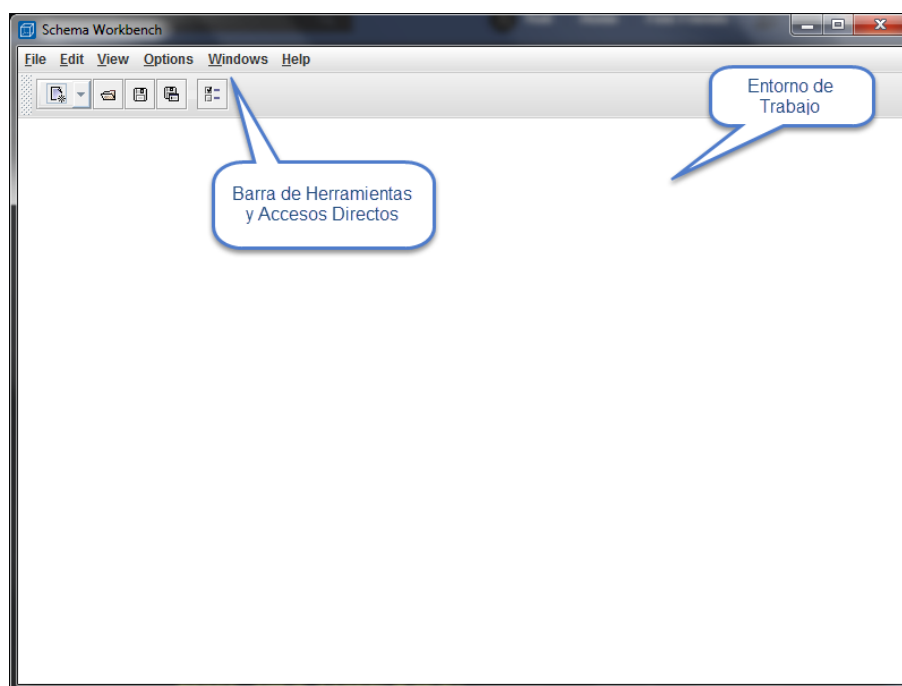


Elaborado por el equipo de trabajo

3.4.4. Gestión del Cubo OLAP

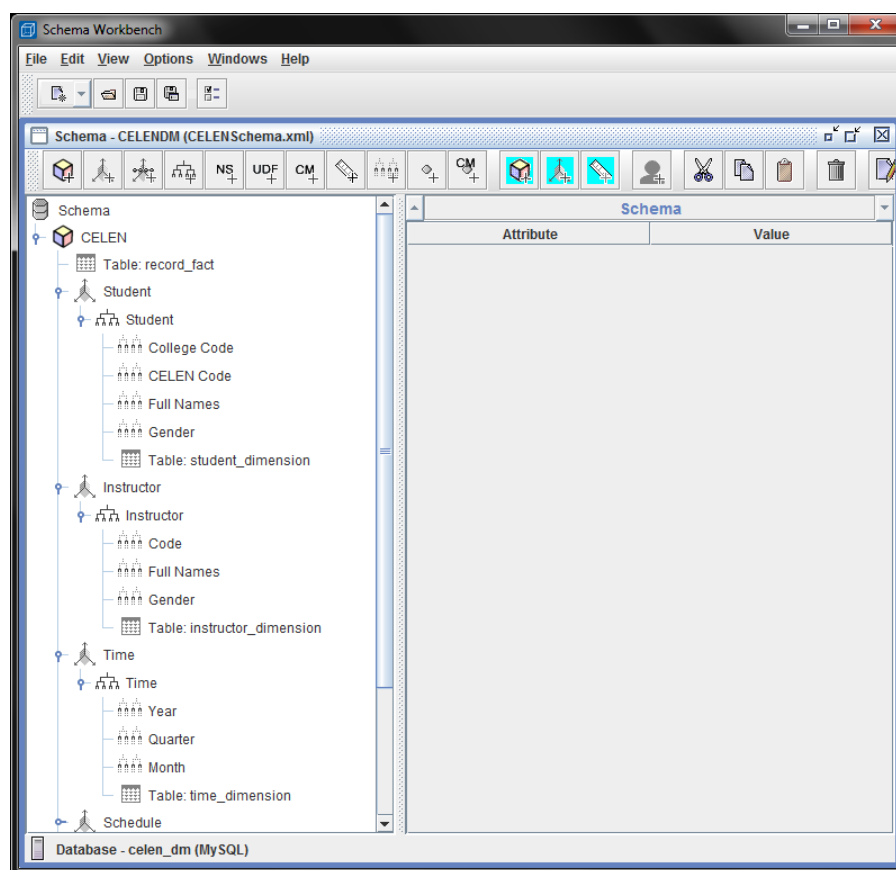
Para la creación y carga del cubo, se ingresó al entorno de Pentaho Schema Workbench (PSW) para la creación y diseño del cubo OLAP.

Figura 3.28: Entorno de la Herramienta PSW



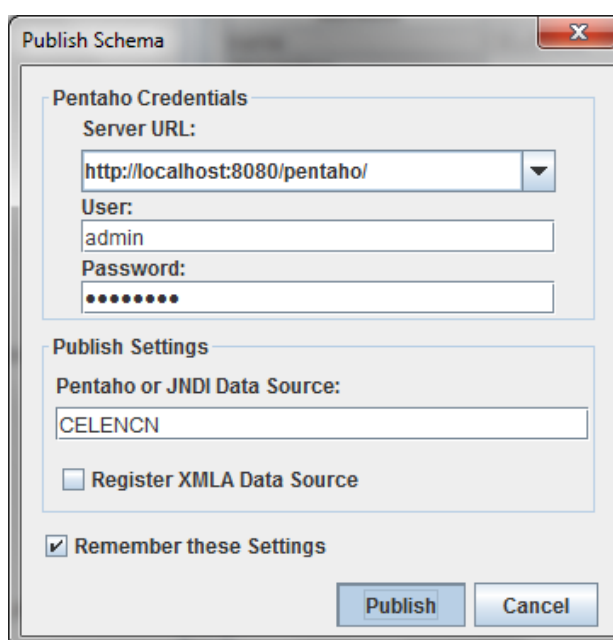
Elaborado por el equipo de trabajo

Figura 3.29: Diseño del Cubo OLAP CELEN



Elaborado por el equipo de trabajo

Figura 3.30: Publicación del Esquema OLAP



Elaborado por el equipo de trabajo

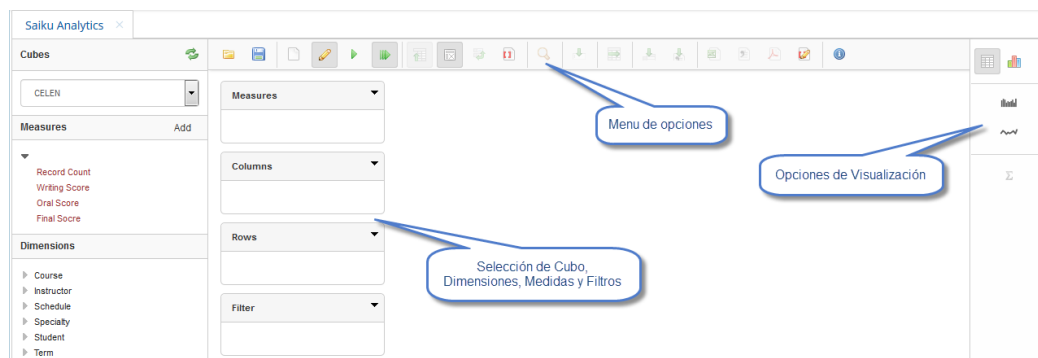
3.4.5. Reportes

En esta fase se construye la vista de los reportes que serán usados por los usuarios finales, dichos reportes están basados en la información que ha proporcionado.

Para la generación de reportes se han considerado desarrollar consultas por estudiante ya que es esencial el seguimiento de su récord académico a lo largo del tiempo.

Por otro lado, la herramienta para la generación de estos reportes se utilizó el plugin Saiku Analytics para Pentaho Business Analytics (BA) Server (Anexo 5).

Figura 3.31: Entorno del Plugin Saiku Analytics



Elaborado por el equipo de trabajo

Tabla 3.28: Reporte del Récord Académico de un Estudiante por Especialidad y Curso

Student/Full Names		SANTOS HAÑARI MAMANI		
Specialty/Name	Course/Name	Writing Score	Oral Score	Final Score
AMERICAN ENGLISH FILE	BASICO I	90	96	93
	BASICO II	92	92	92
	BASICO III	96	95	96
	BASICO IV	87	93	90
	BASICO IX	82	82	82
	BASICO V	86	80	83
	BASICO VI	87	85	86

	BASICO VII	88	92	90
	BASICO VIII	100	95	98
	BASICO X	85	78	82
	BASICO XI	85	80	83
	BASICO XII	81	80	81
	INTERMEDIO I	97	98	98
	INTERMEDIO II	12	10	11
AMERICAN ENGLISH FILE 2ED	INTERMEDIO I	87	84	86
	INTERMEDIO II	93	89	91
	INTERMEDIO III	84	88	86
	INTERMEDIO IV	84	92	88
	INTERMEDIO V	79	91	85
	PRE INTERMEDIO II	38	38	38
	PRE INTERMEDIO III	76	92	84
	PRE INTERMEDIO IV	91	98	95
	PRE INTERMEDIO V	99	94	97
	PRE INTERMEDIO VI	94	97	96
	SUPERIOR I	83	89	86
	SUPERIOR II	92	89	91
	SUPERIOR III	90	88	89
	SUPERIOR IV	48	45	47
	SUPERIOR V	0	0	0

Elaborado por el equipo de trabajo

Tabla 3.29: Reporte del Récord Académico de un Estudiante por Curso y Docente

Student/Full Names		SANTOS HAÑARI MAMANI		
Course/Name	Instructor/Full Names	Writing Score	Oral Score	Final Score
BASICO I	DANA EVELYN MEZA LAURA	90	96	93
BASICO II	CLAUDIA CANDELARIA CUENTAS ALVARADO	92	92	92
BASICO III	ÁNGEL JAVIER QUISPE CARITA	96	95	96
BASICO IV	GUIDO JOSÉ CATACORA PINAZO	87	93	90
BASICO IX	HERMELINDA SALAS SAAVEDRA	82	82	82
BASICO V	YOCELYN RITA FERNANDEZ SALAS	86	80	83
BASICO VI	ELIDA KARINA LARRAÑAGA GALINDO	87	85	86

BASICO VII	ALEJANDRO SERRUTO BELLIDO	88	92	90
BASICO VIII	JAIME ENRIQUE MALMA ZARAVIA	100	95	98
BASICO X	ELDER MAGUIÑA CUTIPA	89	89	89
	ELARD MURILLO TICONA	81	67	74
BASICO XI	ELIDA KARINA LARRAÑAGA GALINDO	85	80	83
BASICO XII	CLAUDIA CANDELARIA CUENTAS ALVARADO	81	80	81
INTERMEDIO I	CLAUDIA CANDELARIA CUENTAS ALVARADO	87	84	86
	ELARD MURILLO TICONA	97	98	98
INTERMEDIO II	ELDER MAGUIÑA CUTIPA	93	89	91
	JUAN JOSÉ MEMBRILLO APARICIO	12	10	11
INTERMEDIO III	DAVID ALARCON HURTADO	84	88	86
INTERMEDIO IV	LAURA ENRIQUETA TORRES LLERENA	84	92	88
INTERMEDIO V	JAIME ENRIQUE MALMA ZARAVIA	79	91	85
PRE INTERMEDIO II	LAURA ENRIQUETA TORRES LLERENA	76	76	76
	HORARACIO ERICK ANGULO MENDOZA	0	0	0
PRE INTERMEDIO III	HENRY RIGOBERTO CHOQUEHUANCA FLORES	76	92	84
PRE INTERMEDIO IV	HERMELINDA SALAS SAAVEDRA	91	98	95
PRE INTERMEDIO V	ELARD MURILLO TICONA	99	94	97
PRE INTERMEDIO VI	YUVITZA PUMAYALI QUILCA	94	97	96

SUPERIOR I	CAROLA EMILIA ROCA AVILA	83	89	86
SUPERIOR II	ELIZABETH JESÚS VALVERDE HERRERA	92	89	91
SUPERIOR III	NORMA AYME CHOQUE TICONA	90	88	89
SUPERIOR IV	ALBERT EZEQUIEL CARI ARAGON	0	0	0
	YUVITZA PUMAYALI QUILCA	96	90	93
SUPERIOR V	ZOCORRO LICELY AGUILAR PORTUGAL	0	0	0

Elaborado por el equipo de trabajo

Tabla 3.30: Reporte del Récord Académico de un Estudiante por Tiempo
(Año - Mes)

Student/Full Names	SANTOS HAÑARI MAMANI			
Year	Month	Writing Score	Oral Score	Final Socre
2013	April	90	96	93
	May	92	92	92
	June	96	95	96
	July	87	93	90
	August	86	80	83
	September	87	85	86
	October	88	92	90
	November	100	95	98
	December	82	82	82
2014	January	81	67	74
	March	89	89	89
	April	85	80	83
	August	81	80	81
	October	97	98	98
	November	12	10	11
2016	June	0	0	0
	July	76	76	76
	August	76	92	84
	September	91	98	95
	October	99	94	97
	November	94	97	96
	December	87	84	86
2017	January	93	89	91
	February	84	88	86
	March	84	92	88
	April	79	91	85
	May	83	89	86

	June	92	89	91
	July	90	88	89
	August	96	90	93
	September	0	0	0
2018	July	0	0	0

Elaborado por el equipo de trabajo

Tabla 3.31: Reporte del Récord Académico de un Estudiante por Periodo y Horario

Student/Full Names			SANTOS HAÑARI MAMANI			
Year	Month	Start Time	End Time	Writing Score	Oral Score	Final Score
2013	April	06:00 PM	07:30 PM	90	96	93
	May	06:00 PM	07:30 PM	92	92	92
	June	04:30 PM	06:00 PM	96	95	96
	July	04:30 PM	06:00 PM	87	93	90
	August	04:30 PM	06:00 PM	86	80	83
	September	06:00 PM	07:30 PM	87	85	86
	October	06:00 PM	07:30 PM	88	92	90
	November	06:00 PM	07:30 PM	100	95	98
	December	06:00 PM	07:30 PM	82	82	82
2014	January	06:00 PM	07:30 PM	81	67	74
	March	06:00 PM	07:30 PM	89	89	89
	April	06:00 PM	07:30 PM	85	80	83
	August	06:00 PM	07:30 PM	81	80	81
	October	06:00 PM	07:30 PM	97	98	98
	November	07:30 PM	09:00 PM	12	10	11
2016	June	06:30 AM	08:00 AM	0	0	0
	July	06:00 PM	07:30 PM	76	76	76
	August	06:00 PM	07:30 PM	76	92	84
	September	06:00 PM	07:30 PM	91	98	95
	October	06:00 PM	07:30 PM	99	94	97
	November	06:00 PM	07:30 PM	94	97	96
	December	06:00 PM	07:30 PM	87	84	86
2017	January	06:00 PM	07:30 PM	93	89	91
	February	06:00 PM	07:30 PM	84	88	86
	March	06:00 PM	07:30 PM	84	92	88
	April	06:00 PM	07:30 PM	79	91	85
	May	06:00 PM	07:30 PM	83	89	86
	June	06:00 PM	07:30 PM	92	89	91
	July	06:00 PM	07:30 PM	90	88	89
	August	06:00 PM	07:30 PM	96	90	93
	September	06:00 PM	07:30 PM	0	0	0
2018	July	06:00 PM	07:30 PM	0	0	0

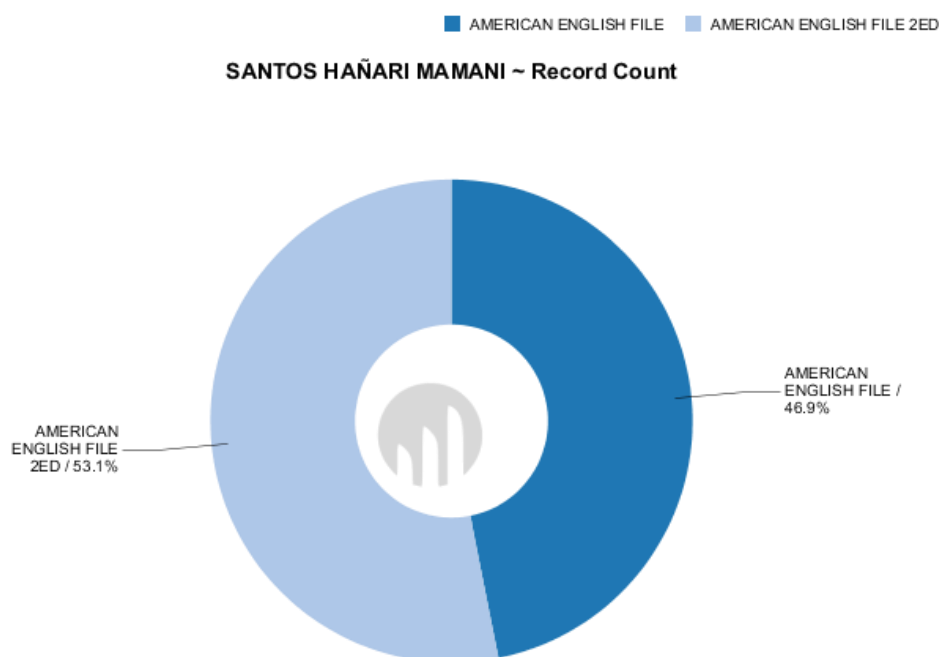
Elaborado por el equipo de trabajo

Tabla 3.32: Reporte del Récord Académico de un Estudiante por Trimestre

Student/Full Names		SANTOS HAÑARI MAMANI			
Year	Quarter	Name	Writing Score	Oral Score	Final Score
2013	Q2	BASICO I	90	96	93
		BASICO II	92	92	92
		BASICO III	96	95	96
	Q3	BASICO IV	87	93	90
		BASICO V	86	80	83
		BASICO VI	87	85	86
	Q4	BASICO IX	82	82	82
		BASICO VII	88	92	90
		BASICO VIII	100	95	98
2014	Q1	BASICO X	85	78	82
	Q2	BASICO XI	85	80	83
	Q3	BASICO XII	81	80	81
	Q4	INTERMEDIO I	97	98	98
		INTERMEDIO II	12	10	11
2016	Q2	PRE INTERMEDIO II	0	0	0
	Q3	PRE INTERMEDIO II	76	76	76
		PRE INTERMEDIO III	76	92	84
	Q4	PRE INTERMEDIO IV	91	98	95
		INTERMEDIO I	87	84	86
	Q4	PRE INTERMEDIO V	99	94	97
		PRE INTERMEDIO VI	94	97	96
2017	Q1	INTERMEDIO II	93	89	91
		INTERMEDIO III	84	88	86
		INTERMEDIO IV	84	92	88
	Q2	INTERMEDIO V	79	91	85
		SUPERIOR I	83	89	86
		SUPERIOR II	92	89	91
	Q3	SUPERIOR III	90	88	89
		SUPERIOR IV	96	90	93
		SUPERIOR V	0	0	0
2018	Q3	SUPERIOR IV	0	0	0

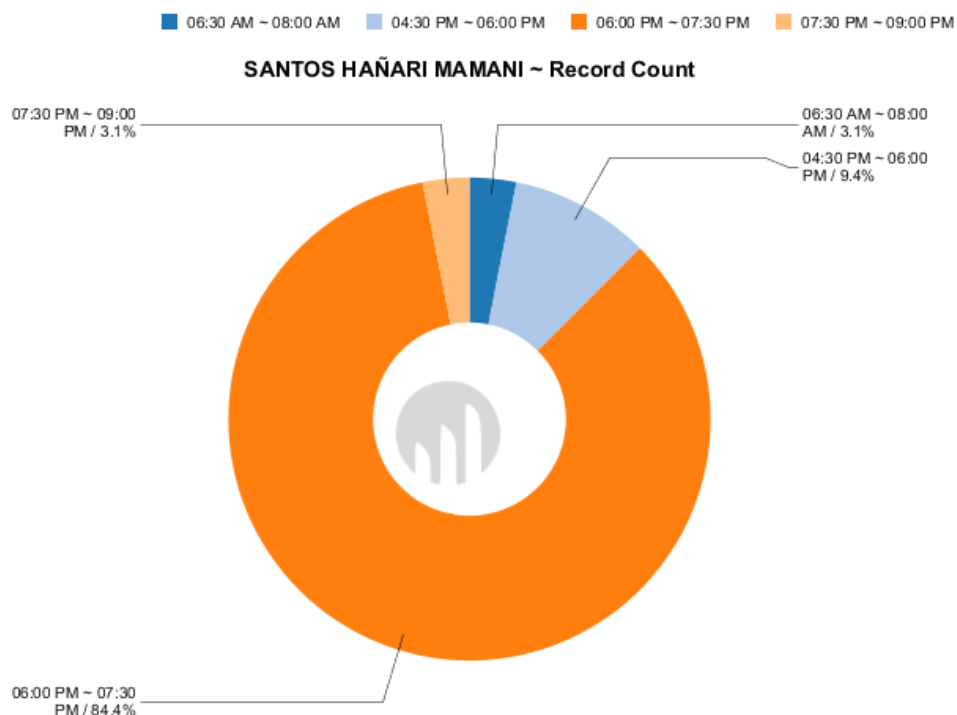
Elaborado por el equipo de trabajo

Figura 3.32: Porcentaje de Cursos tomados por Especialidad



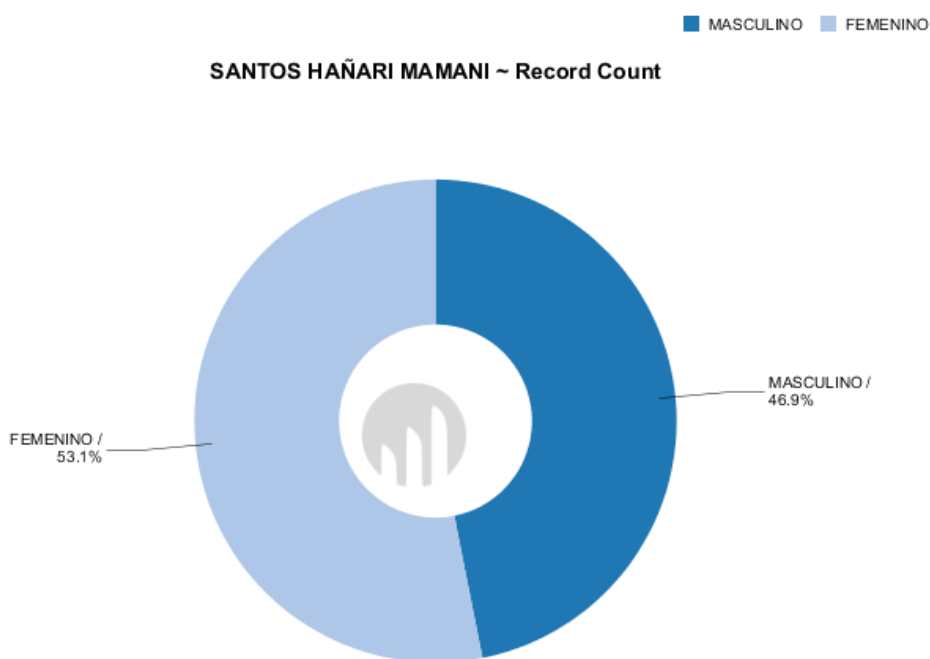
Elaborado por el equipo de trabajo

Figura 3.33: Cantidad de Horarios por Periodo



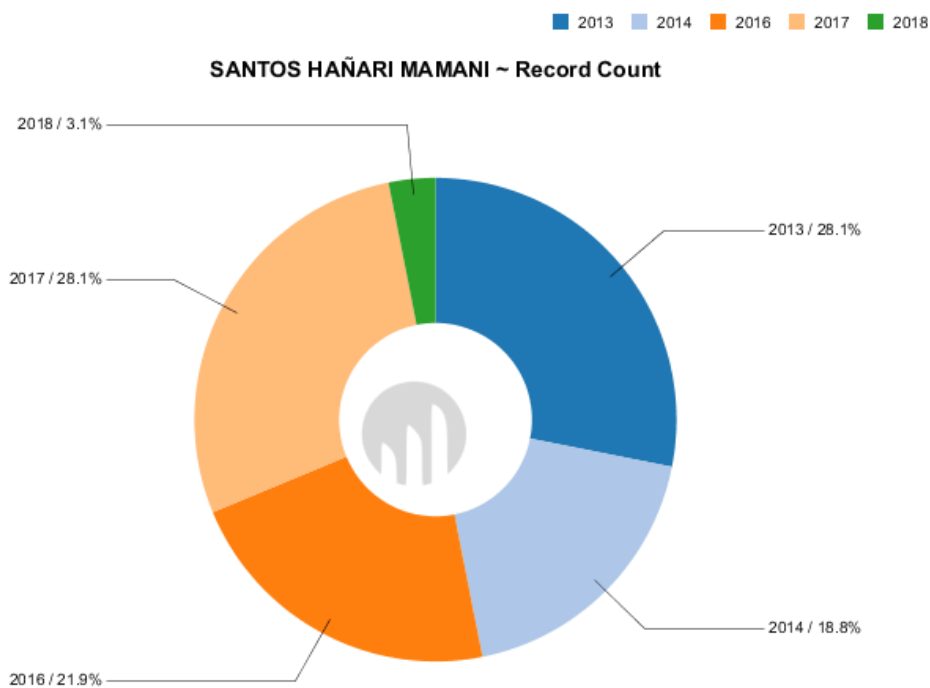
Elaborado por el equipo de trabajo

Figura 3.34: Cantidad de Instructores por Sexo



Elaborado por el equipo de trabajo

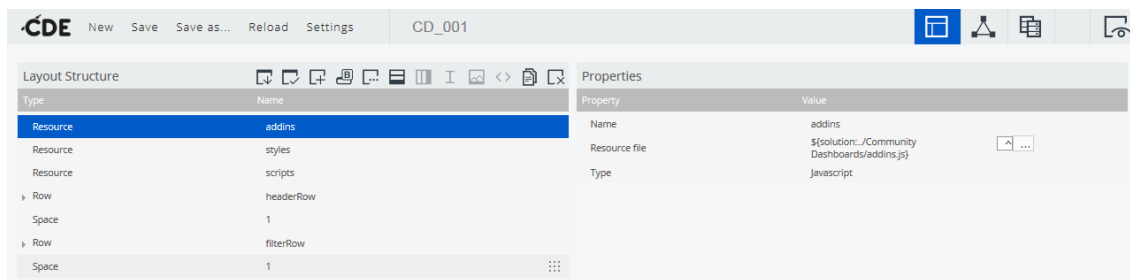
Figura 3.35: Cantidad de Cursos por Año



Elaborado por el equipo de trabajo

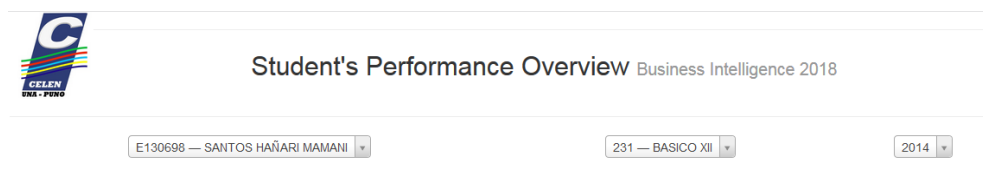
3.4.6. Dashboard

Figura 3.36: Interfaz de la Herramienta CDE



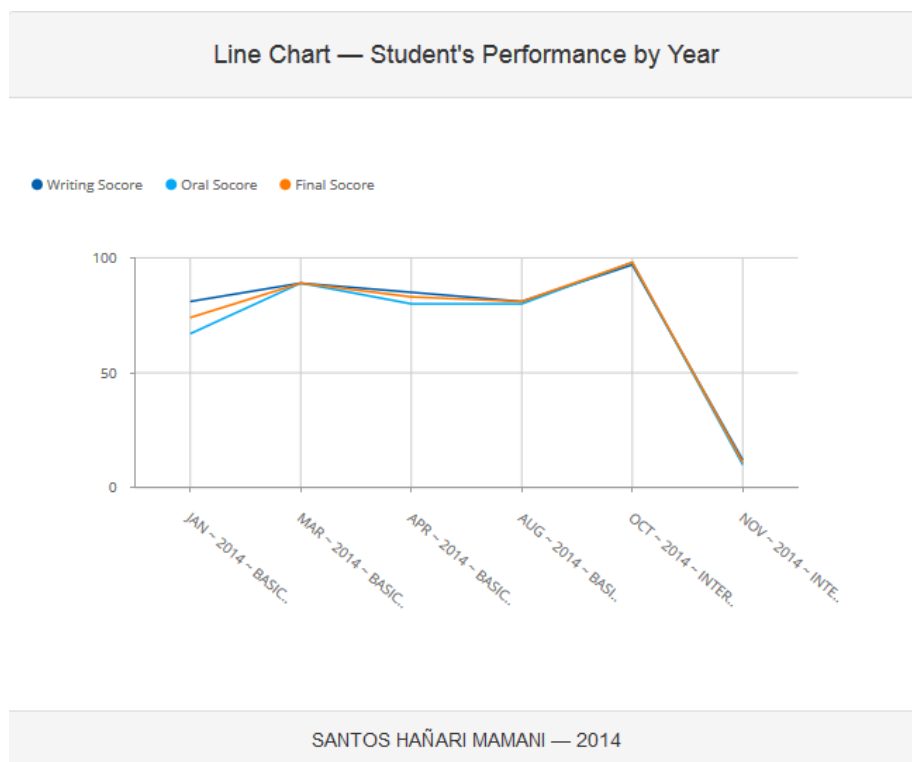
Elaborado por el equipo de trabajo

Figura 3.37: Listas desplegable para el filtrado de Estudiante, Curso y Año



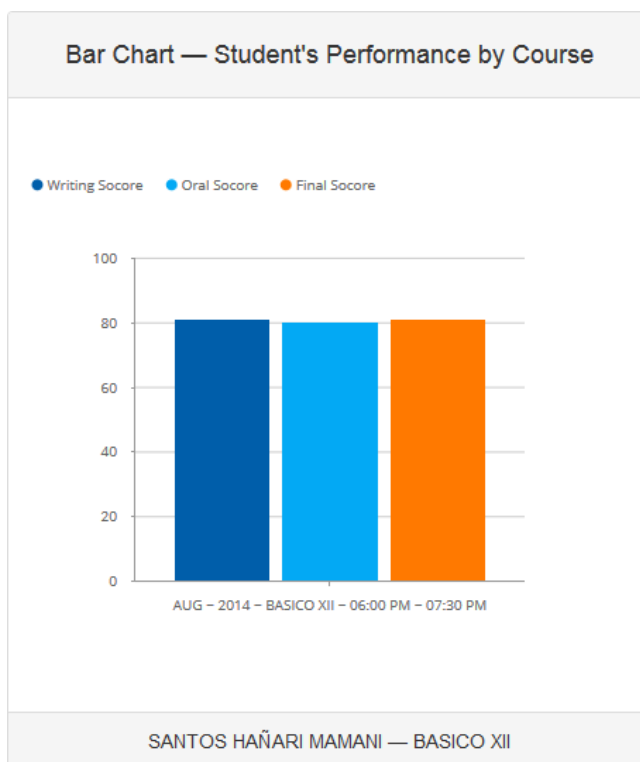
Elaborado por el equipo de trabajo

Figura 3.38: Gráfico de Líneas - Rendimiento del Estudiante por Año



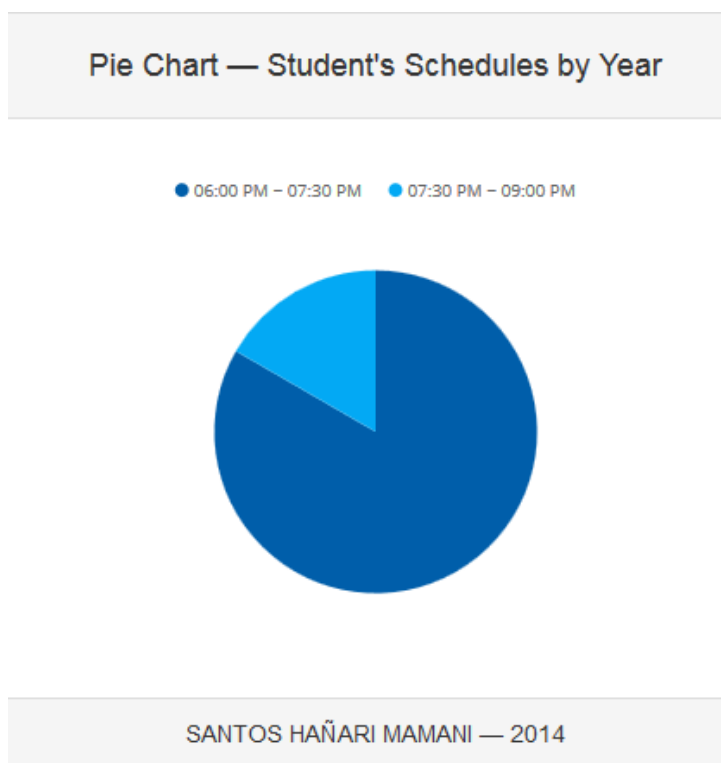
Elaborado por el equipo de trabajo

Figura 3.39: Gráfico de Barras - Rendimiento del Estudiante por Curso



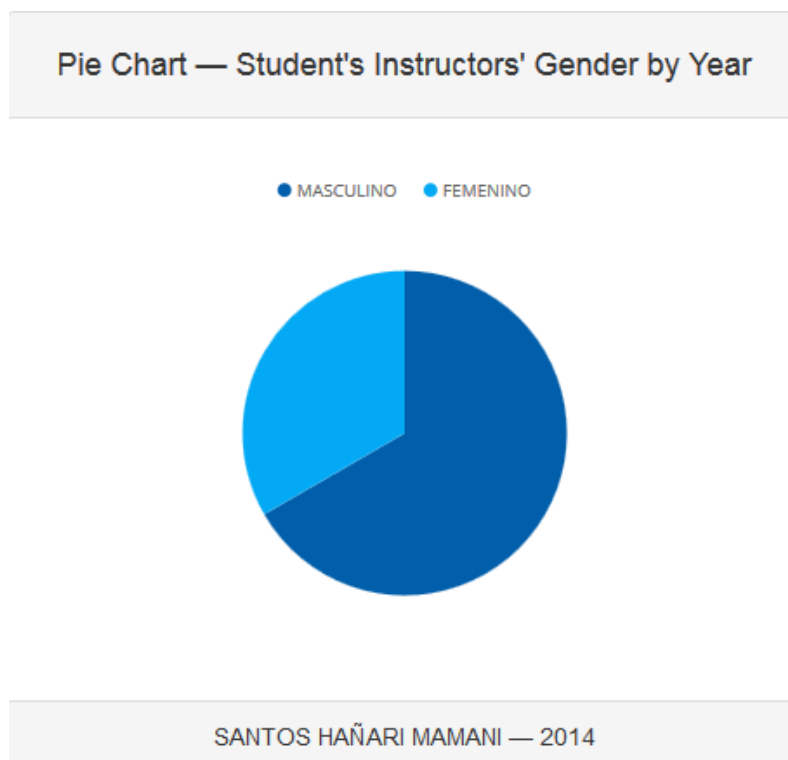
Elaborado por el equipo de trabajo

Figura 3.40: Gráfico Circular - Horarios del Estudiante por Año



Elaborado por el equipo de trabajo

Figura 3.41: Gráfico Circular - Genero de los Docentes del Alumno por Año



Elaborado por el equipo de trabajo

Figura 3.42: Tabla de Datos - Datos completos del Estudiante

Data Table — Full Student's Data

Show 10 entries Search:

Year	Month	Specialty	Course	Instructor	Schedule	Writing	Oral	Final
2013	APRIL	AMERICAN ENGLISH FILE	BASICO I	DANA EVELYN MEZA LAURA	06:00 PM - 07:30 PM	90	96	93
2013	MAY	AMERICAN ENGLISH FILE	BASICO II	CLAUDIA CANDELARIA CUENTAS ALVARADO	06:00 PM - 07:30 PM	92	92	92
2013	JUNE	AMERICAN ENGLISH FILE	BASICO III	ÁNGEL JAVIER QUISPE CARITA	04:30 PM - 06:00 PM	96	95	96
2013	JULY	AMERICAN ENGLISH FILE	BASICO IV	GUIDO JOSÉ CATACTORA PINAZO	04:30 PM - 06:00 PM	87	93	90
2013	AUGUST	AMERICAN ENGLISH FILE	BASICO V	YOCELYN RITA FERNANDEZ SALAS	04:30 PM - 06:00 PM	86	80	83
2013	SEPTEMBER	AMERICAN ENGLISH FILE	BASICO VI	ELIDA KARINA LARRAÑAGA GALINDO	06:00 PM - 07:30 PM	87	85	86
2013	OCTOBER	AMERICAN ENGLISH FILE	BASICO VII	ALEJANDRO SERRUTO BELLIDO	06:00 PM - 07:30 PM	88	92	90
2013	NOVEMBER	AMERICAN ENGLISH FILE	BASICO VIII	JAIME ENRIQUE MALMA ZARAVIA	06:00 PM - 07:30 PM	100	95	98
2013	DECEMBER	AMERICAN ENGLISH FILE	BASICO IX	HERMELINDA SALAS SAAVEDRA	06:00 PM - 07:30 PM	82	82	82
2014	JANUARY	AMERICAN ENGLISH FILE	BASICO X	ELARD MURILLO TICONA	06:00 PM - 07:30 PM	81	67	74

Showing 1 to 10 of 32 entries Previous 1 2 3 4 Next

E130698 — SANTOS HAÑARI MAMANI

Elaborado por el equipo de trabajo

Figura 3.43: Vista General del Dashboard Principal



Elaborado por el equipo de trabajo

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

3.4. Población y Muestra

Se identifica como unidad de análisis a las personas involucradas en la gestión de la información del área académica del Centro de Estudios de Lenguas Extranjeras y Nativas de la Universidad Nacional del Altiplano – Puno para lo cual se contó con 4 miembros del personal administrativo y 21 instructores del idioma inglés. Teniendo en total una población de 25 personas.

Determinamos la muestra según la fórmula:

$$n = \frac{n'}{1 + \left(\frac{n'}{N}\right)} \quad (\text{Ec.4.1})$$

$$n' = \frac{S^2}{V^2} \quad (\text{Ec.4.2})$$

$$S^2 = p(1 - p) \quad (\text{Ec.4.3})$$

FUENTE: (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2014, pp.170-191)

- S^2 = varianza de la muestra expresada como la probabilidad de ocurrencia de y .
- y = valor promedio de una variable
- V = varianza de la población. Su definición es (S_e) el cuadrado del error estándar.
- S_e = error estándar (0.05)
- N = población
- n' = muestra.

$$s^2 = 0.9 (1 - 0.9) = 0.09$$

$$V = 0.05^2 = 0.0025$$

$$n' = \frac{0.09}{0.0025} = 36$$
$$n = \frac{36}{1 + \left(\frac{36}{25}\right)} = 14.75 = 15$$

Es decir que, para la investigación, se necesita una muestra de 15 personas para evaluar la incorporación de elementos de inteligencia de negocios.

3.5. Análisis e Interpretación de Resultados

La prueba T Student se utiliza para contrastar hipótesis sobre medias en poblaciones con distribución normal. También proporciona resultados aproximados para los contrastes de medias en muestras suficientemente grandes cuando estas poblaciones no se distribuyen normalmente. Para conocer si se puede suponer que los datos siguen una distribución normal, se pueden realizar diversos contrastes llamados de bondad de ajuste, de los cuales el más usado es la prueba de Kolmogorov. A menudo, la prueba de Kolmogorov es referida erróneamente como prueba de Kolmogorov-Smirnov, ya que en realidad esta última, sirve para contrastar si dos poblaciones tienen la misma distribución. (Sealey, 2018: 2)

La prueba de T Student para muestras independientes, se utiliza para comparar las medias de un mismo grupo en diferentes etapas, por ejemplo, para las comparaciones de los resultados de una prueba Pre-Test (GC) y Post-Test (GE).

3.5.1. Resultados Genéricos

A continuación, se muestra los resultados obtenidos después de la Post-Test (GE), en comparación con los datos previamente obtenidos en el Pre-Test (GC) (Anexo 1).

Tabla 4.1: Resultados Generales de la Pre-Test (GC) y Post-Prueba (GE)

Indicadores		Pre-Test (GC)	Post-Test (GE)
Exactitud de la Información	Completa	60.0 %	93.3 %
	Incompleta	40.0 %	6.7 %
Tiempo en realizar los reportes	Promedio	62 minutos	30.6 minutos
Número de reportes solicitados por ciclo	Promedio	3 reportes	8 reportes
Comprensión de los reportes	Comprensible	66.7 %	86.7 %
	Incomprensible	33.3 %	13.3 %

Elaborado por el equipo de trabajo

3.5.2. Resultados Específicos

Tabla 4.2: Resultados Específicos Pre-Test (GC) y Post-Test (GE)

#	Exactitud de la información		Tiempo en realizar los reportes (min.)		Número de reportes solicitados por ciclo (reportes/ciclo)		Comprensión de los reportes	
	Pre-Test (GC)	Post-Test (GE)	Pre-Test (GC)	Post-Test (GE)	Pre-Test (GC)	Post-Test (GE)	Pre-Test (GC)	Post-Test (GE)
1	Incompleta	Incompleta	70	45	4	7	Incomprensible	Incomprensible
2	Incompleta	Completa	60	35	3	8	Incomprensible	Incomprensible
3	Incompleta	Completa	70	45	3	7	Incomprensible	Comprensible
4	Incompleta	Completa	75	50	4	5	Incomprensible	Comprensible
5	Incompleta	Completa	60	30	2	8	Incomprensible	Comprensible
6	Incompleta	Completa	80	35	4	10	Incomprensible	Comprensible
7	Incompleta	Completa	60.5	20	3	8	Incomprensible	Comprensible
8	Incompleta	Completa	90	30	4	10	Incomprensible	Comprensible
9	Incompleta	Completa	75	20	3	10	Incomprensible	Comprensible
10	Completa	Completa	60	25	2	10	Incomprensible	Comprensible
11	Completa	Completa	60.5	23	3	10	Comprensible	Comprensible
12	Completa	Completa	80	32	4	9	Comprensible	Comprensible
13	Completa	Completa	60	20	3	10	Comprensible	Comprensible
14	Completa	Completa	55	30	2	10	Comprensible	Comprensible
15	Completa	Completa	45	20	2	8	Comprensible	Comprensible

Elaborado por el equipo de trabajo

3.6. Nivel de Confianza

El nivel de confianza será de 95%, es decir, se tiene un margen de error de 5%

3.7. Contraste de la Hipótesis

En este punto se realizó la contrastación de las muestras Pre-Test (GC) y Post-Test (GE) de los indicadores definidos en los puntos anteriores. Los planteamientos de la hipótesis de estos indicadores se detallan a continuación:

3.7.1. Contraste para la Exactitud de la Información

A) Planteamiento de la Hipótesis

Se debe validar el impacto que tiene la incorporación de elementos de inteligencia de negocios en el porcentaje de Exactitud de Información en la Gestión de la Información en el Área de Académica del CELEN de la UNA Puno, llevado a cabo en la muestra. Se realiza una medición antes de la incorporación de los elementos de BI (Pre-Test (GC)) y otra después de la incorporación de los elementos de BI (Post-Test (GE)). Para cuantificar los índices de medición del indicador de la exactitud, en esta investigación se decidió medirla de forma binaria (1 o 0) completa o incompleta.

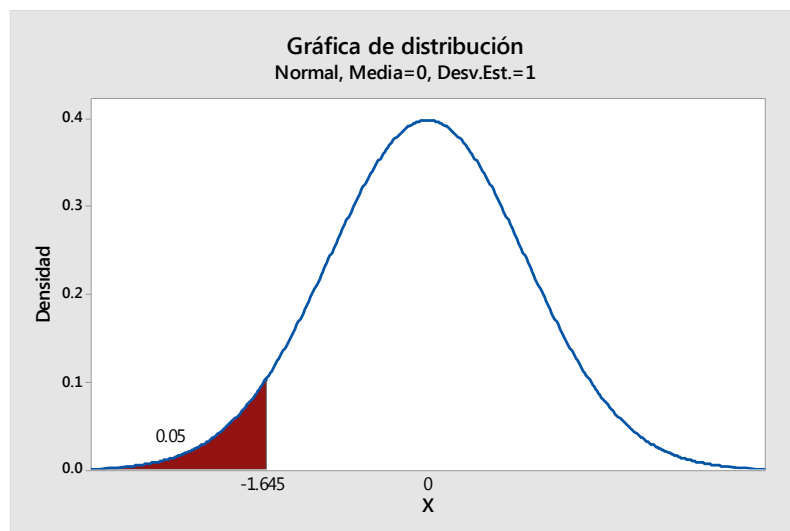
Pre-Test (GC)	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Post-Test (GE)	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1

μ_1 = Media de la exactitud de la información en el Pre-Test (GC)
 μ_2 = Media de la exactitud de la información en el Post-Test (GE)

$$\begin{aligned}
 H_0 &: \mu_1 \geq \mu_2 \\
 H_a &: \mu_1 < \mu_2
 \end{aligned}
 \tag{Ec.4.4}$$

B) Criterios de Decisión

Figura 4.1: Distribución de Probabilidad para la Exactitud de la Información

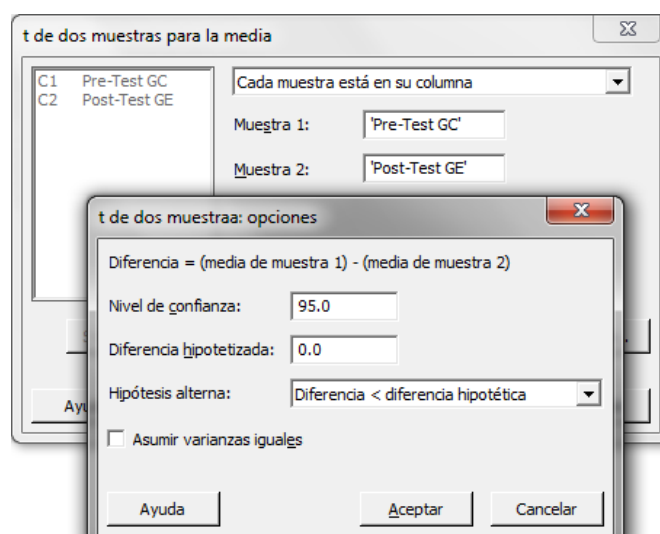


Elaborado por el equipo de trabajo

C) Cálculo: Prueba-T de dos Muestras para la Media

Para visualizar los resultados de la contrastación, primero se selecciona los datos de la muestra Pre-Test (GC) y luego los datos de la muestra Post-Test (GE), con un nivel de confianza del 95%.

Figura 4.2: Ingreso de Datos para realizar la Prueba-t



Elaborado por el equipo de trabajo

Tabla 4.3: Resultados de la Prueba-t para la Exactitud de la Información

	Pre-Test (GC)	Post-Test (GE)
Media	0.4	0.93
Desviación estándar	0.5	0.25
Observaciones	15	15
Diferencia hipotética de las medias	-0.533	
Valor p	0.0008340728223332	
Valor t	-3.63	

Elaborado por el equipo de trabajo

D) Decisión Estadística

Puesto que el valor $p = 0,000834 < \alpha = 0,05$, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H_0), y la hipótesis alterna (H_a) es cierta. La prueba resultó ser significativa.

3.7.2. Contraste para el Tiempo En Realizar Los Reportes

A) Planteamiento de la Hipótesis

Se debe validar el impacto que tiene la incorporación de elementos de inteligencia de negocios en el tiempo en realizar los reportes para la Gestión de la Información en el Área de Académica del CELEN de la UNA Puno, llevado a cabo en la muestra. Se realiza una medición antes de la incorporación de los elementos de BI (Pre-Test (GC)) y otra después de la incorporación de los elementos de BI (Post-Test (GE)).

Pre-Test (GC) (min.)	70	60	70	75	60	80	60.5	90	75	60	60.5	80	60	55	45
Post-Test (GE) (min.)	45	35	45	50	30	35	20	30	20	25	23	32	20	30	20

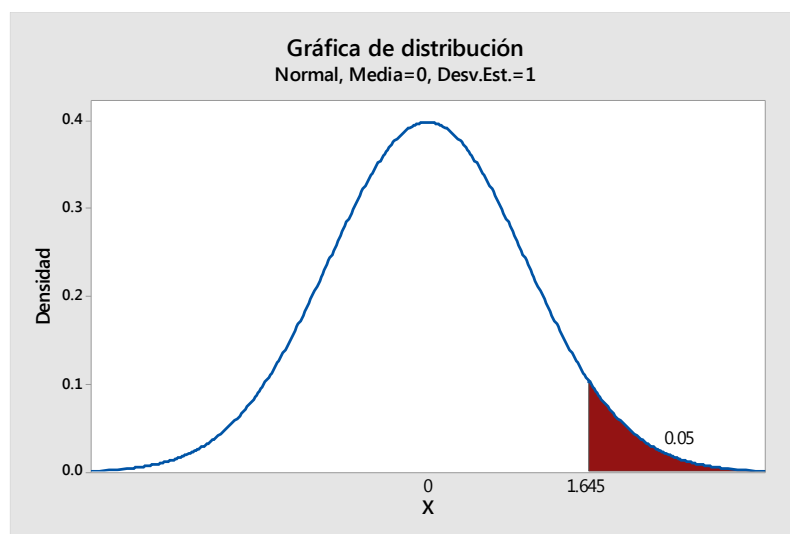
μ_1 = Media del tiempo en realizar los reportes en el Pre-Test (GC)

μ_2 = Media del tiempo en realizar los reportes en el Post-Prueba (GE)

$$\begin{aligned}
 H_0 &: \mu_1 \leq \mu_2 \\
 H_a &: \mu_1 > \mu_2
 \end{aligned}
 \tag{Ec.4.5}$$

B) Criterios de Decisión

Figura 4.3: Distribución de la Probabilidad para el Tiempo en Realizar los Reportes



Elaborado por el equipo de trabajo

C) Cálculo: Prueba-T de dos Muestras para la Media

Para visualizar los resultados de la contrastación, primero se selecciona los datos de la muestra Pre-Test (GC) y luego los datos de la muestra Post-Test (GE), con un nivel de confianza del 95%.

Tabla 4.4: Resultados de la Prueba-t para el Tiempo en Realizar los Reportes

	Pre-Test (GC)	Post-Test (GE)
Media	66.7	30.67
Desviación estándar	11.7	9.90
Observaciones	15	15
Diferencia hipotética de las medias	31.4	
Valor p	0.0000000005046664	
Valor t	9.11	

Elaborado por el equipo de trabajo

D) Decisión Estadística

Puesto que el valor $p = 0,00000000050 < \alpha = 0,05$, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H_0), y la hipótesis alterna (H_a) es cierta. La prueba resultó ser significativa.

3.7.3. Contraste para el Número de Reportes Solicitados por Ciclo

A) Planteamiento de la Hipótesis

Se debe validar el impacto que tiene la incorporación de elementos de inteligencia de negocios el número de reportes solicitados por ciclo en la Gestión de la Información en el Área de Académica del CELEN de la UNA Puno, llevado a cabo en la muestra. Se realiza una medición antes de la incorporación de los elementos de BI (Pre-Test (GC)) y otra después de la incorporación de los elementos de BI (Post-Test (GE)).

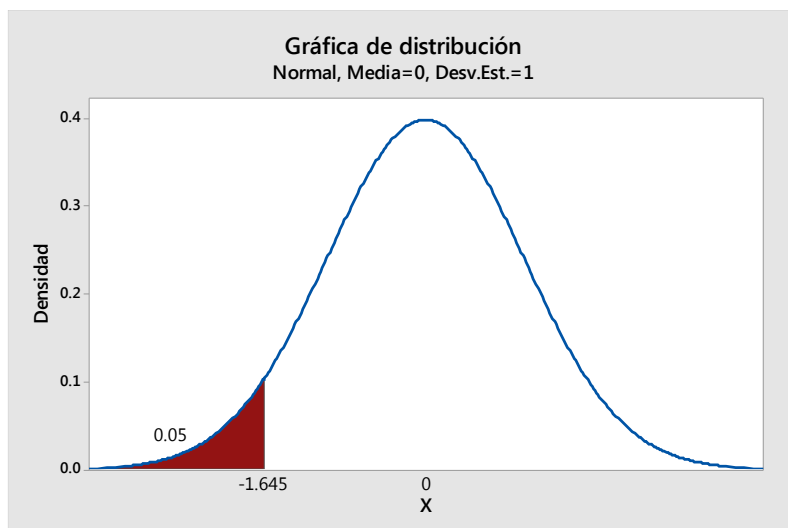
Pre-Test (GC) (reporte/ciclo)	4	3	3	4	2	4	3	4	3	2	3	4	3	2	2
Post-Test (GE) (reporte/ciclo)	7	8	7	5	8	10	8	10	10	10	10	9	10	10	8

μ_1 = Media del número de reportes solicitados por ciclo en el Pre-Test (GC)
 μ_2 = Media del número de reportes solicitados por ciclo en el Post-Test (GE)

$$\begin{aligned}
 H_0 &: \mu_1 \geq \mu_2 \\
 H_a &: \mu_1 < \mu_2
 \end{aligned}
 \tag{Ec.4.6}$$

B) Criterios de Decisión

Figura 4.4: Distribución de la Probabilidad para el Número de Reportes Solicitados por Ciclo



Elaborado por el equipo de trabajo

C) Cálculo: Prueba-T de dos Muestras para la Media

Para visualizar los resultados de la contrastación, primero se selecciona los datos de la muestra Pre-Test (GC) y luego los datos de la muestra Post-Test (GE), con un nivel de confianza del 95%.

Tabla 4.5: Resultados de la Prueba-t para el Número de Reportes por Ciclo

	Pre-Test (GC)	Post-Test (GE)
Media	3.067	0.79
Desviación estándar	8.67	1.54
Observaciones	15	15
Diferencia hipotética de las medias	-5.600	
Valor p	0.0000000000175403	
Valor de t	-12.48	

Elaborado por el equipo de trabajo

D) Decisión Estadística

Puesto que el valor $p = 0,000000000017 < \alpha = 0,05$, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H_0), y la hipótesis alterna (H_a) es cierta. La prueba resultó ser significativa.

3.7.4. Contraste para la Comprensión de los Reportes

A) Planteamiento de la Hipótesis

Se debe validar el impacto que tiene la incorporación de elementos de inteligencia de negocios en la comprensión de los reportes emitidos para la Gestión de la Información en el Área de Académica del CELEN de la UNA Puno, llevado a cabo en la muestra. Se realiza una medición antes de la incorporación de los elementos de BI (Pre-Test (GC)) y otra después de la incorporación de los elementos de BI (Post-Test (GE)). Para cuantificar los índices de medición del indicador de la comprensión, en esta investigación se decidió medirla de forma binaria (1 o 0) Comprensible o incomprensible.

Pre-Test (GC)	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Post-Test (GE)	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

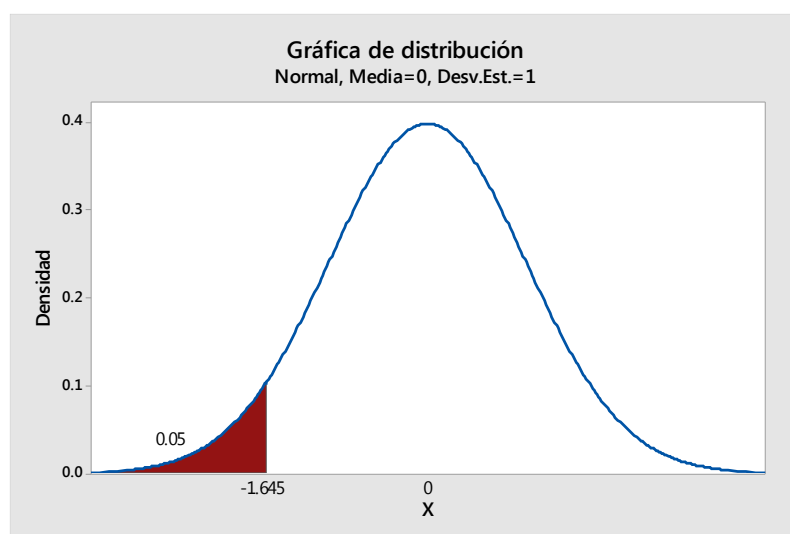
μ_1 = Media de la comprensión los reportes en el Pre-Test (GC)

μ_2 = Media de la comprensión los reportes en el Post-Test (GE)

$$\begin{aligned}
 H_0 &: \mu_1 \geq \mu_2 \\
 H_a &: \mu_1 < \mu_2
 \end{aligned}
 \tag{Ec.4.7}$$

B) Criterios de Decisión

Figura 4.5: Distribución de la Probabilidad para la Comprensión de los Reportes



Elaborado por el equipo de trabajo

C) Cálculo: Prueba-T de dos Muestras para la Media

Para visualizar los resultados de la contrastación, primero se selecciona los datos de la muestra Pre-Test (GC) y luego los datos de la muestra Post-Test (GE), con un nivel de confianza del 95%.

Tabla 4.6: Resultados de la Prueba-t para la Comprensión de los Reportes

	Pre-Test (GC)	Post-Test (GE)
Media	0.333	0.867
Desviación estándar	0.488	0.352
Observaciones	15	15
Diferencia hipotética de las medias	-0.533	
Valor p	0.0010423626710508	
Valor t	-3.43	

Elaborado por el equipo de trabajo

D) Decisión Estadística

Puesto que el valor $p = 0,0010423 < \alpha = 0,05$, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H_0), y la hipótesis alterna (H_a) es cierta. La prueba resultó ser significativa.

3.8. Discusión

Dada las pruebas realizadas y los resultados obtenidos por cada uno de los indicadores en el punto anterior se puede confirmar y aceptar la hipótesis general de que la incorporación de elementos de inteligencia de negocios, mejora la gestión de la información, en el área académica del Centro de Estudios de Lenguas Extranjeras y Nativas.

CONCLUSIONES

PRIMERO: El análisis de la naturaleza del negocio en el área académica del CELEN de la UNA Puno condujeron a la obtención de los requerimientos e indicadores claves del negocio tales como la exactitud de la información, tiempo en realizar reportes, número de reportes solicitados por ciclo y la comprensión de los mismos.

SEGUNDO: Se comprueba que, mejoró la gestión de la información en el área académica del CELEN de la UNA Puno mediante la incorporación de elementos de Inteligencia de Negocios, utilizando la Metodología de Ralph Kimball, teniendo como resultado el incremento de la exactitud de la información de un 60 a un 93 por ciento.

TERCERO: Es evidente que, la identificación de las diferentes fuentes de datos permitió sentar una base para el diseño de un Data Mart y la explotación del Cubo OLAP ha permitido la generación de un mayor número de reportes, de 3 a 8 por ciclo.

CUARTO: Se observa, que la incorporación de inteligencia de negocios para el área académica del CELEN de la UNA Puno ayudó a la reducción del tiempo de elaboración de reportes, de 62 a 30 minutos.

QUINTO: Es notorio que, la implementación de soluciones para la visualización y análisis de los datos contenidos en el Data Mart permite un manejo intuitivo y sencillo a los usuarios finales que condujeron una mayor comprensión de reportes, de un 66 a un 86 por ciento

RECOMENDACIONES

PRIMERO: Para alcanzar la efectividad de la gestión de la información se debe establecer una cultura organizacional, para ello directores académicos deben saber comprometer e involucrar al personal hacia nuevos objetivos de la institución las cuales debería estar enfocado a la inteligencia de negocios.

SEGUNDO: Se recomienda para futuras versiones del proyecto, la ampliación de más indicadores del proceso académicos, así como la adición de los demás idiomas a parte del inglés y seguir incorporando elementos de Inteligencia de Negocios para el CELEN de la UNA Puno.

TERCERO: Se recomienda elaborar el diccionario de datos con la finalidad de minimizar el tiempo de entender la funcionalidad del cubo OLAP y dar seguimiento a la carga de información hacia el Data Mart, a fin de que se mantenga actualizado y sea realmente de beneficio para los usuarios finales y tomadores de decisiones.

CUARTO: Finalmente se aconseja, realizar capacitación a los usuarios finales y tomadores de decisiones en cuanto al uso de las herramientas y al contenido del del Data Mart implementad

REFERENCIAS BIBLIOGRÁFICAS

- Acero Calizaya, R. F. (2013). *Data Mart de Contrataciones Públicas a partir del SEACE, y su aplicación en la Toma de Decisiones de las Micro y Pequeñas Empresas de la ciudad de Puno*. Puno: Universidad Nacional del Altiplano.
- Alfaro Barrera, O. J., Medrano Artola, O. A., & Nolasco Díaz, J. I. (2016). *Análisis de Inteligencia de Negocios desde la Perspectiva de los Creadores del Concepto*. El Salvador: Universidad Dr. José Matías Delgado.
- Arenas López, M. C., & Gómez Montes, A. M. (2017). *Inteligencia de Negocios aplicada a los Procesos de Autoevaluación de la Universidad de Manizales*. Manizales: Universidad de Manizales.
- Arévalo, J. A. (21 de diciembre de 2018). *Gestión de la Información, Gestión de Contenidos y Conocimiento*. Obtenido de E-prints: http://eprints.rclis.org/11273/1/Jornadas_GRUPO_SIOU.pdf
- Basantes Espinoza, G. P., & López Galarza, D. E. (2012). *Estudio de la Aplicación de Inteligencia de Negocios en los Procesos Académicos*. Guayaquil: Universidad Politécnica Salesiana.
- Bernabeu, R. D. (2010). *DATA WAREHOUSING: Investigación y Sistematización de Conceptos*. Córdoba: Bernabeu.
- Catañeda Vásquez, A. (2015). *Desarrollo de Business Intelligence, basado en la Metodología de Ralph Kimball, para mejorar el Proceso de Toma de Decisiones en el Área de Admisión de la Universidad Autónoma del Perú*. Lima: Universidad Autónoma del Perú.
- Contreras Contreras, F. (14 de diciembre de 2018). *Indicadores de Gestión en Unidades de Información*. Obtenido de E-Prints: http://eprints.rclis.org/7008/1/1_10.pdf
- Curto Díaz, J. (2010). *Introducción al Business Intelligence*. Barcelona: UOC.
- Durá, S. S. (2011). *Sistemas de Información BI: Estado Actual y Herramientas de Software Libre*. 1-19.
- Espinosa, R. (19 de abril de 2010). *Kimball vs Inmon. Ampliación de conceptos del Modelado Dimensional*. Obtenido de El Rincón del BI. Kimball vs Inmon: <https://churriwifi.wordpress.com/2010/04/19/15-2-ampliacion-conceptos-del-modelado-dimencional/>
- Fuentes Tapia, L., & Valdivia Pinto, R. (2010). *Incorporación de Elementos de Inteligencia de Negocios en el Proceso de Admisión y Matrícula de una Universidad Chilena*. *Revista Chilena de Ingeniería*, vol. 18 N° 3, 383-394.
- Guillén Quisca, R. N. (2016). *Sistema de Soporte de Decisiones con Tecnología Data Warehouse para la Gestión de la Información de la Empresa Mallku Import SAC – Juliaca*. Puno: Universidad Nacional del Altiplano.


- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la Investigación Científica*. México D.F.: McGraw-Hill Education.
- Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition*. Indiana: John Wiley & Sons, Inc.
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2008). *The Data Warehouse Lifecycle Toolkit, Second Edition*. Indiana: Wiley Publishing, Inc.
- Marcus, A. (2006). Dashboard In Your Future. *The Art of Prototyping*. Vol. 13, Issue 1, 48.
- Murillo Junco, M. J., & Cáceres Castellanos, G. (2013). Business intelligence y la toma de decisiones financieras: una aproximación teórica. *Revista Logos, Ciencia & Tecnología*, vol. 5, núm. 1, 119-138.
- Nima Ramos, J. D. (2009). SOLUCIONES DE INTELIGENCIA DE NEGOCIOS A SU ALCANCE: FUNDAMENTOS Y CASOS DE APLICACIÓN. *Contribuciones a la Economía*, 84.
- Oracle. (10 de diciembre de 2018). *¿Qué es Inteligencia de Negocios?* Obtenido de Oracle:
https://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf
- Pérez-Montoro Gutiérrez, M. (10 de diciembre de 2009). *Gestión de la Información*. Obtenido de glossariumBITri: <http://glossarium.bitrum.unileon.es/Home/gestion-de-la-informacion>
- PowerData. (10 de diciembre de 2018). *Procesos ETL: Definición, Características, Beneficios y Retos*. Obtenido de PowerData: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/312584/Procesos-ETL-Definicion-Caracteristicas-Beneficios-y-Retos>
- Quispe Paco, E. Á. (2012). *Sistema de Inteligencia de Negocios para el Soporte de Toma de Decisiones en el Área de Nutrición y Aprendizaje Infantil del Programa Nacional Wawa Wasi sede Puno*. Puno: Universidad Nacional del Altiplano.
- Rivadera, G. (2010). La Metodología de Kimball para el Diseño de Almacenes de Datos. *Cuadernos de la Facultad N° 5*, 56-71.
- Salaki, & Waworuntu. (2016). Extract Transformation Loading from OLTP to OLAP data using Pentaho Data Integration. *IOP Conference Series: Materials Science and Engineering*, 123-132.
- Sealey Gosset, W. (23 de Noviembre de 2018). *UPTC Colombia*. Obtenido de Prueba T Student : <http://virtual.uptc.edu.co/ova/estadistica/docs/libros/tstudent.pdf>

ANEXOS


Anexo 1: Encuestas Post-Test (GE)



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 35 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 8 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprendible
- B) Incomprendible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 45 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 7 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprendible
- B) Incomprendible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 45 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 7 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprendible
- B) Incomprendible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 30 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 8 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprendible
- B) Incomprendible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 50 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 5 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Compreensible
- B) Incompreensible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 35 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 10 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Compreensible
- B) Incompreensible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 30 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 10 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Compreensible
- B) Incompreensible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 20 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 10 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Compreensible
- B) Incompreensible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 20 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 8 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprensible
- B) Incomprensible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 20 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 8 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprensible
- B) Incomprensible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 30 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 9 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Compreensible
- B) Incompreensible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 20 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 10 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Compreensible
- B) Incompreensible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 20 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 8 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprendible
- B) Incomprendible

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
- B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 30 minuto(s)
- B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 10 reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprendible
- B) Incomprendible



UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y
SISTEMAS
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

ENCUESTA POST-TEST

Instrucciones:

El presente instrumento forma parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018".

Sea tan amable de leer cuidadosamente las preguntas subrayar o completar una alternativa según sea el caso, de acuerdo a su criterio.

1. ¿Cómo calificaría la exactitud de la información actual del Área Académica?

- A) Completa
 B) Incompleta

2. ¿Qué tiempo estimado toma realizar reportes del Área Académica?

- A) 20..... minuto(s)
B) hora(s)

3. ¿Cuántos reportes por ciclo son solicitados en Área Académica?

- A) 8..... reporte(s)

4. ¿El nivel de abstracción de los reportes emitidos son?

- A) Comprensible
 B) Incomprensible

Anexo 2: Formato del Cuestionario Requerimientos Técnico

UNIVERSIDAD NACIONAL DEL ALTIPLANO PUNO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

**CUESTIONARIO TÉCNICO****Instrucciones:**

Este cuestionario formará parte del Proyecto de Investigación de Tesis titulado: "INCORPORACIÓN DE ELEMENTOS DE INTELIGENCIA DE NEGOCIOS, PARA MEJORAR LA GESTIÓN DE LA INFORMACIÓN, EN EL ÁREA ACADÉMICA DEL CENTRO DE ESTUDIOS DE LENGUAS EXTRANJERAS Y NATIVAS 2018". Sea tan amable de leer cuidadosamente las preguntas y subraye una respuesta. De la veracidad de los datos depende de la realización de este Proyecto. Muchas gracias por su colaboración.

Objetivos:

Determinar los factores que inciden en el correcto uso de la Información y la Gestión de la misma.

1. ¿El CELEN cuenta con servidores dedicados para el almacenamiento de la información?

A) Sí B) No C) No sabe

2. ¿El Sistema Académico bajo que Gestor de Base de Datos funciona?

A) MySQL B) Oracle C) MS SQL 2008 R2 D) MS Access E) MS Excel

3. ¿Actualmente la infraestructura del CELEN cuántos años tiene de capacidad para almacenar información?

A) De 3 a 5 años B) De 5 a 7 años C) Más de 7 años

4. ¿Bajo qué tipo de plataforma está hecho el sistema académico?

A) Licencia B) Open Source

5. ¿Cuentan con herramientas para generar informes estadísticos?

A) Sí B) No C) No sabe

6. ¿Cuentan con herramientas para generar reportes de manera que la información pueda ser procesada de diferentes maneras?

A) Sí B) No C) No sabe

Si su respuesta fue afirmativa conteste la siguiente pregunta.

¿Cuáles son esas herramientas?
.....

7. ¿Cree usted que el CELEN cuenta con herramientas especializadas para la mejor la Gestión de la Información al momento de procesarla?

A) Sí B) No C) No sabe

8. ¿De qué manera el usuario visualiza la información?

A) Aplicación de escritorio B) Web C) Intranet D) Otras

9. ¿Sabe usted que es Inteligencia de Negocios?

A) Sí B) No

10. ¿Inteligencia de Negocios, lo podría definir cómo?

- A) Proceso para definir, estructurar, contextualizar y explotar la información de una Empresa y tomar las mejores decisiones.
B) Parametrizar la información de una Empresa y tomar las mejores decisiones.
C) Ninguna de las anteriores.

11. ¿Tiene conocimientos sobre alguna herramienta de Inteligencia de Negocios?

A) Sí B) No C) No sabe

Anexo 3: Instalación de Controladores JDBC

Para conectarse a una base de datos, incluida la base de datos Repositorio BA o Repositorio DI, deberá descargar e instalar un controlador JDBC en los lugares apropiados para los componentes de Pentaho, así como en el servidor de aplicaciones web que contiene los servidores de Pentaho.

1. Descargue un controlador JAR JDBC Driver de su proveedor de base de datos o un desarrollador de controladores de terceros.
2. Copie el JAR del controlador JDBC que acaba de descargar en el directorio pentaho / jdbc-distribution.

Lista de productos y ubicaciones correspondientes para controladores JDBC

Server or Design Tool	Directory
Business Analytics (BA) Server	pentaho/server/biserver-ee/tomcat/lib
Data Integration (DI) Server	pentaho/server/data-integration-server/tomcat/lib
Pentaho Data Integration (Spoon)	pentaho/design-tools/data-integration/lib
Pentaho Report Designer (PRD)	pentaho/design-tools/report-designer/lib/jdbc
Pentaho Aggregation Designer (PAD)	pentaho/design-tools/aggregation-designer/drivers
Pentaho Schema Workbench (PSW)	pentaho/design-tools/schema-workbench/drivers
Pentaho Metadata Editor (PME)	pentaho/design-tools/metadata-editor/libext/JDBC

FUENTE: <https://help.pentaho.com/Documentation/5.4/OD0/160/030>

Anexo 4: Instalación y Configuración de Java JDK para Pentaho en Windows

1. Descargue e instale Java JDK según el sistema operativo y la versión del mismo:

Java SE Development Kit 8u191

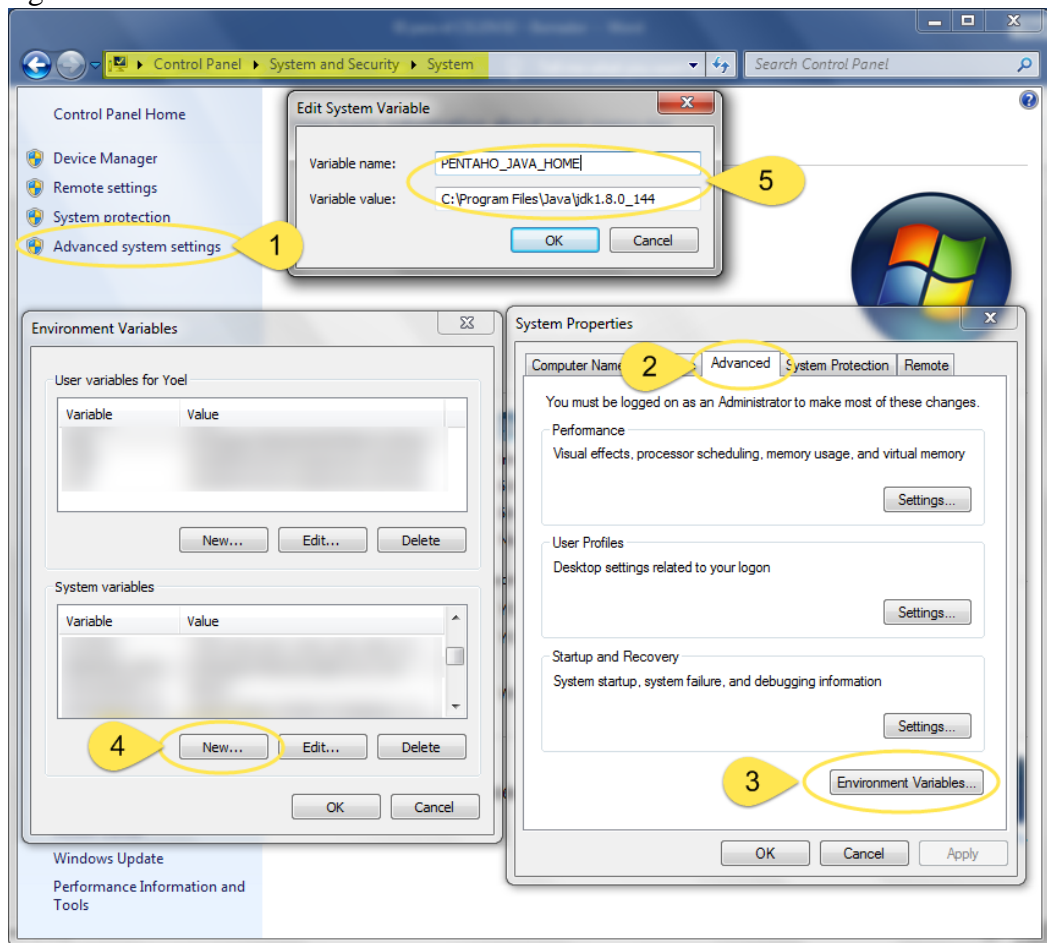
You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this software.

Thank you for accepting the Oracle Binary Code License Agreement for Java SE; you may now download this software.

Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	72.97 MB	jdk-8u191-linux-arm32-vfp-hflt.tar.gz
Linux ARM 64 Hard Float ABI	69.92 MB	jdk-8u191-linux-arm64-vfp-hflt.tar.gz
Linux x86	170.89 MB	jdk-8u191-linux-i586.rpm
Linux x86	185.69 MB	jdk-8u191-linux-i586.tar.gz
Linux x64	167.99 MB	jdk-8u191-linux-x64.rpm
Linux x64	182.87 MB	jdk-8u191-linux-x64.tar.gz
Mac OS X x64	245.92 MB	jdk-8u191-macosx-x64.dmg
Solaris SPARC 64-bit (SVR4 package)	133.04 MB	jdk-8u191-solaris-sparcv9.tar.Z
Solaris SPARC 64-bit	94.28 MB	jdk-8u191-solaris-sparcv9.tar.gz
Solaris x64 (SVR4 package)	134.04 MB	jdk-8u191-solaris-x64.tar.Z
Solaris x64	92.13 MB	jdk-8u191-solaris-x64.tar.gz
Windows x86	197.34 MB	jdk-8u191-windows-i586.exe
Windows x64	207.22 MB	jdk-8u191-windows-x64.exe

FUENTE: <https://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>

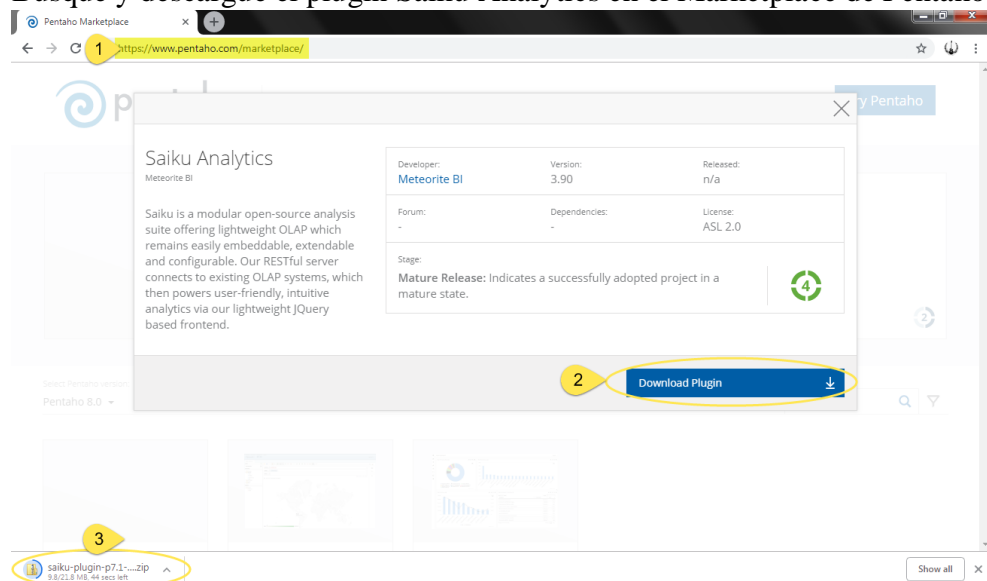
2. Agregue la variable de entorno PENTAHO_JAVA_HOME como sigue en la figura:



Elaborado por el equipo de trabajo

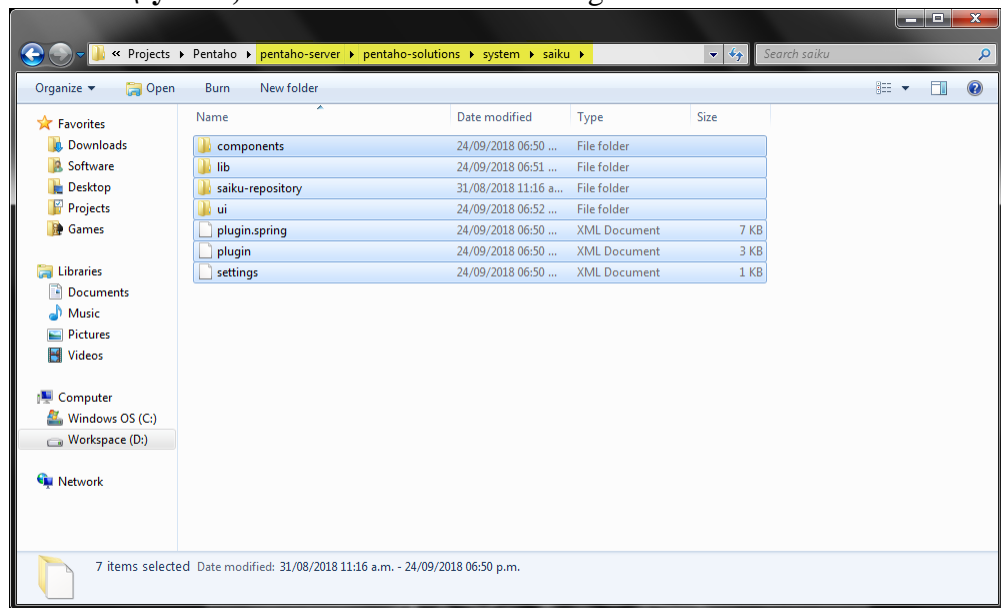
Anexo 4: Instalación del Plugin Saiku Analytics para Pentaho

1. Busque y descargue el plugin Saiku Analytics en el Marketplace de Pentaho



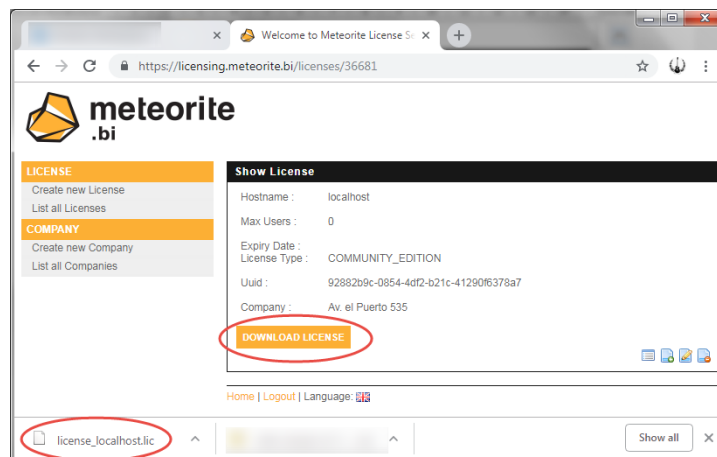
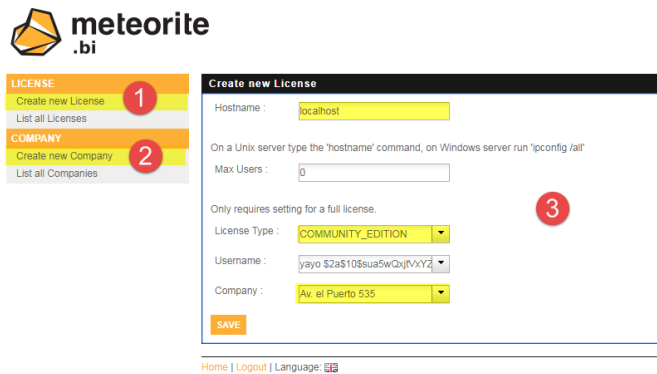
FUENTE: <https://www.pentaho.com/marketplace/>

- Extraiga los archivos en el directorio `\pentaho-server\pentaho-solutions\system`, como se muestra en la imagen:



Elaborado por el equipo de trabajo

- Generar, descargar e instalar la licencia para Saiku Analytics (previamente se debe realizar el registro para poder acceder y generar una licencia).



FUENTE: <https://licensing.meteorite.bi>,

- Copiar y pegar el archivo generado y descargado en la raíz de la carpeta del plugin.

Name	Date modified	Type	Size
components	24/09/2018 06:50 ...	File folder	
lib	24/09/2018 06:51 ...	File folder	
saiku-repository	31/08/2018 11:16 a...	File folder	
ui	24/09/2018 06:52 ...	File folder	
license.lic	24/09/2018 05:17 ...	VisualStudio.lic.bc...	1 KB
plugin.spring	24/09/2018 06:50 ...	XML Document	7 KB
plugin	24/09/2018 06:50 ...	XML Document	3 KB
settings	24/09/2018 06:50 ...	XML Document	1 KB

Elaborado por el equipo de trabajo

Anexo 5: Lista del Personal Administrativo y Docente del CELEN

1. Personal Administrativo

Nombres y Apellidos	Cargo
DRA. DELICIA VILMA GONZALES ARESTEGUI	DIRECTORA
PROF. ANTONIETA DELGADO	ADMINISTRADORA
ING. NESTOR HUANCA FRIAS	COORDINADOR ACADÉMICO
PROF. MARCO ANTONIO ATENCIO PAREDES	APOYO COORDINACIÓN
SRA. MILAGROS NUÑEZ LOZA	SECRETARIA

FUENTE: CELEN

2. Instructores de CELEN

#	Docente	Idioma
1	AGUILAR PORTUGAL ZOCORRO LICELY	INGLÉS
2	ALARCON HURTADO DAVID	INGLÉS
3	ALFARO CARO BRAHIN JIMMY	INGLÉS
4	ALVAREZ PAMPA SERGIO	INGLÉS
5	ANGULO MENDOZA HORARACIO ERICK	INGLÉS
6	CARI ARAGON ALBERT EZEQUIEL	INGLÉS
7	CATACORA PINAZO GUIDO JOSÉ	INGLÉS
8	CHOQUE TICONA NORMA AYME	INGLÉS
9	CHOQUEHUANCA FLORES HENRY RIGOBERTO	INGLÉS
10	CUENTAS ALVARADO CLAUDIA CANDELARIA	INGLÉS
11	FERNANDEZ SALAS YOCELYN RITA	INGLÉS
12	FERRO GONZALES CHRISTOPHER LEE	INGLÉS
13	HOLGUIN ORDOÑO ROLANDO	INGLÉS
14	LARRAÑAGA GALINDO ELIDA KARINA	INGLÉS

15	MAGUIÑA CUTIPA ELDER	INGLÉS
16	MALMA ZARAVIA JAIME ENRIQUE	INGLÉS
17	MEMBRILLO APARICIO JUAN JOSÉ	INGLÉS
18	MEZA LAURA DANA EVELYN	INGLÉS
19	MURILLO TICONA ELARD	INGLÉS
20	OLARTE TORRES JORGE EDUARDO	INGLÉS
21	PONCE PEÑALOZA ALCIDES PONCIANO	INGLÉS
22	PUMAYALI QUILCA YUVITZA	INGLÉS
23	QUISPE CARITA ÁNGEL JAVIER	INGLÉS
24	ROCA AVILA CAROLA EMILIA	INGLÉS
25	RUIZ MARQUEZ HILDUARA ROSARIO	INGLÉS
26	SALAS SAAVEDRA HERMELINDA	INGLÉS
27	SERRUTO BELLIDO ALEJANDRO	INGLÉS
28	SOSA CORNEJO MERCEDES LISBETH	INGLÉS
29	TORRES LLERENA LAURA ENRIQUETA	INGLÉS
30	VALVERDE HERRERA ELIZABETH JESÚS	INGLÉS
31	VEIGA DELGADO CARLOS JOSÉ	INGLÉS
32	ALVES NELSON	PORTUGUÉS
33	CHOQUE ZAPANA DAVID REYNALDO	PORTUGUÉS
34	HUALPA RODAS ANDRÉS RODRIGO	PORTUGUÉS
35	QUISPE CARITA ÁNGEL JAVIER	PORTUGUÉS
36	SARAIVA DA SILVA ANA LOURDES	PORTUGUÉS
37	CASTRO VILLEGAS ROSA MARÍA	FRANCES
38	CARI ARAGON ALBERT EZEQUIEL	ITALIANO
39	MEZA LAURA DANA EVELYN	ITALIANO
40	ANDERS MAYA NATALIE	ALEMÁN
41	MAMANI ARUQUIPA HUGO	QUECHUA
42	MANGO MAMANI BENJAMÍN	QUECHUA
43	VALERIANO ANAHUI GUILLERMO	QUECHUA
44	JIMENEZ MITA MIRIAM	AIMARA
45	MAMANI CONDORI JULIÁN	AIMARA

FUENTE: CELEN

ANEXO 6: Script SQL para la Creación del Data Mart

```

1. -- MySQL Workbench Forward Engineering
2.
3. SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0;
4. SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0;
5. SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='TRADITIONAL,ALLOW_INVALID_DATES';
6.
7. -----
8. -- Schema celen_dm
9. -----
10.
11. CREATE SCHEMA IF NOT EXISTS `celen_dm` DEFAULT CHARACTER SET utf8 ;
12. USE `celen_dm` ;
13.
14.

```

```

15. -----
16. -- Table `celen_dm`.`student_dimension`
17. -----
18. CREATE TABLE IF NOT EXISTS `celen_dm`.`student_dimension` (
19.   `student_key` INT UNSIGNED NOT NULL AUTO_INCREMENT,
20.   `student_id` INT NOT NULL,
21.   `student_college_code` CHAR(6) NULL,
22.   `student_celen_code` CHAR(7) NOT NULL,
23.   `student_full_names` VARCHAR(255) NOT NULL,
24.   `student_gender` ENUM('MASCULINO', 'FEMENINO') NOT NULL,
25.   PRIMARY KEY (`student_key`))
26. ENGINE = InnoDB;
27.
28. -----
29. -- Table `celen_dm`.`instructor_dimension`
30. -----
31. CREATE TABLE IF NOT EXISTS `celen_dm`.`instructor_dimension` (
32.   `instructor_key` INT UNSIGNED NOT NULL AUTO_INCREMENT,
33.   `instructor_id` INT NOT NULL,
34.   `instructor_code` VARCHAR(7) NOT NULL,
35.   `instructor_full_names` VARCHAR(255) NOT NULL,
36.   `instructor_gender` ENUM('MASCULINO', 'FEMENINO') NOT NULL,
37.   PRIMARY KEY (`instructor_key`))
38. ENGINE = InnoDB;
39.
40. -----
41. -- Table `celen_dm`.`course_dimension`
42. -----
43. CREATE TABLE IF NOT EXISTS `celen_dm`.`course_dimension` (
44.   `course_key` INT UNSIGNED NOT NULL AUTO_INCREMENT,
45.   `course_id` INT NOT NULL,
46.   `course_name` VARCHAR(100) NOT NULL,
47.   PRIMARY KEY (`course_key`))
48. ENGINE = InnoDB;
49.
50. -----
51. -- Table `celen_dm`.`time_dimension`
52. -----
53. CREATE TABLE IF NOT EXISTS `celen_dm`.`time_dimension` (
54.   `time_key` INT UNSIGNED NOT NULL AUTO_INCREMENT,
55.   `time_year` INT NOT NULL,
56.   `time_month` INT NOT NULL,
57.   `time_month_label` TINYTEXT NOT NULL,
58.   `time_quarter` INT NOT NULL,
59.   `time_quarter_label` CHAR(2) NOT NULL,
60.   PRIMARY KEY (`time_key`))
61. ENGINE = InnoDB;
62.
63. -----
64. -- Table `celen_dm`.`specialty_dimension`
65. -----
66. CREATE TABLE IF NOT EXISTS `celen_dm`.`specialty_dimension` (
67.   `specialty_key` INT UNSIGNED NOT NULL AUTO_INCREMENT,
68.   `specialty_id` INT NOT NULL,
69.   `specialty_name` VARCHAR(100) NOT NULL,
70.   PRIMARY KEY (`specialty_key`))
71. ENGINE = InnoDB;
72.
73. -----
74. -- Table `celen_dm`.`schedule_dimension`
75. -----
76. CREATE TABLE IF NOT EXISTS `celen_dm`.`schedule_dimension` (
77.   `schedule_key` INT UNSIGNED NOT NULL AUTO_INCREMENT,
78.   `schedule_id` INT NOT NULL,
79.   `schedule_start_time` TIME NOT NULL,
80.   `schedule_st_label` TINYTEXT NOT NULL,

```

```



81. `schedule_end_time` TIME NOT NULL,
82. `schedule_et_label` TINYTEXT NOT NULL,
83. PRIMARY KEY (`schedule_key`)
84. ENGINE = InnoDB;
85.
86. -----
87. -- Table `celen_dm`.`record_fact`
88. -----
89. CREATE TABLE IF NOT EXISTS `celen_dm`.`record_fact` (
90. `record_key` INT NOT NULL AUTO_INCREMENT,
91. `instructor_key` INT UNSIGNED NOT NULL,
92. `specialty_key` INT UNSIGNED NOT NULL,
93. `schedule_key` INT UNSIGNED NOT NULL,
94. `student_key` INT UNSIGNED NOT NULL,
95. `course_key` INT UNSIGNED NOT NULL,
96. `time_key` INT UNSIGNED NOT NULL,
97. `writing_score` INT NOT NULL,
98. `oral_score` INT NOT NULL,
99. `final_score` INT NOT NULL,
100. INDEX `fk_record_fact_student_dimension_idx` (`student_key` ASC),
101. INDEX `fk_record_fact_instructor_dimension_idx` (`instructor_key` ASC)
102. ,
103. INDEX `fk_record_fact_course_dimension_idx` (`course_key` ASC),
104. INDEX `fk_record_fact_time_dimension_idx` (`time_key` ASC),
105. INDEX `fk_record_fact_specialty_idx` (`specialty_key` ASC),
106. INDEX `fk_record_fact_schedule_idx` (`schedule_key` ASC),
107. PRIMARY KEY (`record_key`),
108. CONSTRAINT `fk_record_fact_student_dimension`
109. FOREIGN KEY (`student_key`)
110. REFERENCES `celen_dm`.`student_dimension` (`student_key`)
111. ON DELETE CASCADE
112. ON UPDATE NO ACTION,
113. CONSTRAINT `fk_record_fact_instructor_dimension`
114. FOREIGN KEY (`instructor_key`)
115. REFERENCES `celen_dm`.`instructor_dimension` (`instructor_key`)
116. ON DELETE CASCADE
117. ON UPDATE NO ACTION,
118. CONSTRAINT `fk_record_fact_course_dimension`
119. FOREIGN KEY (`course_key`)
120. REFERENCES `celen_dm`.`course_dimension` (`course_key`)
121. ON DELETE CASCADE
122. ON UPDATE NO ACTION,
123. CONSTRAINT `fk_record_fact_time_dimension`
124. FOREIGN KEY (`time_key`)
125. REFERENCES `celen_dm`.`time_dimension` (`time_key`)
126. ON DELETE CASCADE
127. ON UPDATE NO ACTION,
128. CONSTRAINT `fk_record_fact_specialty_dimesion`
129. FOREIGN KEY (`specialty_key`)
130. REFERENCES `celen_dm`.`specialty_dimension` (`specialty_key`)
131. ON DELETE CASCADE
132. ON UPDATE NO ACTION,
133. CONSTRAINT `fk_record_fact_schedule_dimension`
134. FOREIGN KEY (`schedule_key`)
135. REFERENCES `celen_dm`.`schedule_dimension` (`schedule_key`)
136. ON DELETE CASCADE
137. ON UPDATE NO ACTION)
138. ENGINE = InnoDB;
139.
140. SET SQL_MODE=@OLD_SQL_MODE;
141. SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;
142. SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;

```

Anexo 7: Software de Inteligencia de Negocios y Otros

Nombre del Software	Rol del Producto
Pentaho Data Integration (PDI)	Spoon es el diseñador gráfico de transformación y trabajo asociado con la suite de integración de datos Pentaho, también conocido como el proyecto Kettle.
Pentaho Schema Workbench (PSW)	Pentaho Schema Workbench (PSW) es una herramienta gráfica: <ul style="list-style-type: none"> ▪ Para crear esquemas de Mondrian. ▪ Para publicar esquemas en el servidor Pentaho.
Pentaho Business Analytics (BA) Server	Es un Host de contenidos creado por Pentaho y creado por el usuario. El servidor BA se incluye con la Consola de usuario, que es una interfaz basada en web que le permite administrar el servidor. La Consola de usuario también le permite crear contenido de análisis empresarial, visualizar y programar informes utilizando una interfaz basada en web y administrar la seguridad de Pentaho.
Pentaho Community Dashboard Editor (CDE)	El Community Dashboard Editor (CDE) y sus tecnologías subyacentes (CDF, CDA y CCC) permiten un rápido desarrollo y despliegue de los paneles de Pentaho CTools. La herramienta CDE se creó para simplificar los procesos de creación, diseño y representación de los tableros de CTools. Es una herramienta potente y completa que integra a la perfección la interfaz de usuario con las fuentes de datos y los componentes personalizados.
Saiku Analytics	Saiku es una suite de análisis modular de código abierto que ofrece OLAP ligero que se puede integrar, ampliar y configurar fácilmente. Su servidor RESTful se conecta a los sistemas OLAP existentes, que a su vez impulsan análisis intuitivos y fáciles de usar a través de una interfaz liviana basada en JQuery.

Elaborado por el equipo de trabajo

Logo del Software	Descripción del Producto
	XAMPP es una distribución Apache completamente gratuita y fácil de instalar que contiene MariaDB, PHP y Perl. El paquete de código abierto XAMPP se ha configurado para que sea increíblemente fácil de instalar y usar.
	MySQL Workbench es una herramienta visual unificada para arquitectos de bases de datos, desarrolladores y administradores de bases de datos. MySQL Workbench proporciona modelado de datos, desarrollo de SQL y herramientas de administración integrales para la configuración del servidor, la administración de usuarios, las copias de respaldo y mucho más. MySQL Workbench está disponible en Windows, Linux y Mac OS X.



Minitab es un programa de computadora diseñado para ejecutar funciones estadísticas básicas y avanzadas. Combina lo amigable del uso de Microsoft Excel con la capacidad de ejecución de análisis estadísticos.

Elaborado por el equipo de trabajo

Anexo 8: Entrevista Personal Área Académica

Nombres del Entrevistado(a): ING. NESTOR HUANCA FRIAS

Cargo del Entrevistado(a): COORDINADOR ACADÉMICO

¿Quiénes son sus clientes?

Mis clientes potenciales son los interesados en estudiar uno o más idiomas que ofrecemos: inglés, alemán, italiano, francés, portugués aimara y quechua. Especialmente son los alumnos de la Universidad Nacional del Altiplano y el idioma con mayor demanda es el inglés.

¿En qué consiste su trabajo en el área?

La función principal es la de brindar información precisa y pertinente acerca de los beneficios de estudiar en el CELEN.

¿Qué retos se tienen que cumplir en el área?

El reto es cumplir con las metas propuestas por la dirección y que cada estudiante sea matriculado en el idioma de su elección de acuerdo a los horarios que tenemos disponibles.

¿El sistema de información actual le ayuda a la Gestión de la Información?

En términos del registro de notas del estudiante, el sistema si me ayuda a obtener y procesar la información que necesito para cada estudiante.

¿Cuáles son las limitaciones del Sistema de Información?

Las limitaciones se presentan cuando queremos hacer reportes estadísticos por estudiante matriculado en cada ciclo, el sistema no filtra comparaciones con periodos anteriores de cada curso de un idioma.

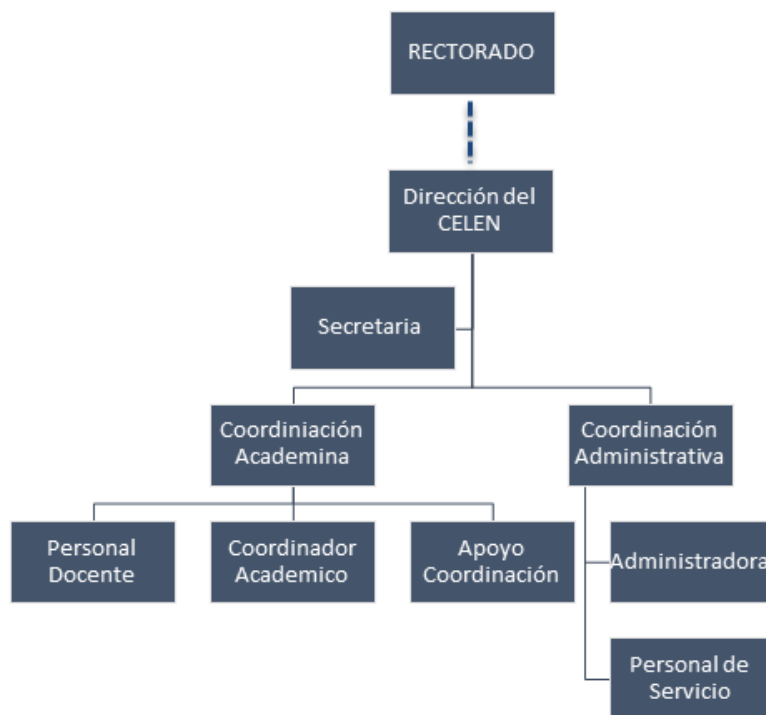
¿Tiene algún problema o dificultad en la Gestión de la Información?

El problema del área académica son las diferentes solicitudes que tenemos que enviar para poder acceder a la información de un o más estudiantes.

¿Qué es lo que requiere en el Sistema de Información para que te ayude a mejorar la Gestión de la Información del área Académica?

Para que el sistema de búsqueda sea más ágil, se podría implementar un registro histórico de los estudiantes según los cursos del idioma(s) en el que este matriculado ejemplo, los docentes nos solicitan el estado del estudiante es decir con que puntajes llegó hasta el ciclo actual y/o los puntajes obtenidos anteriormente, además en que horarios estuvo y con qué docentes ha cursado los diferentes ciclos.

Anexo 9: Organigrama del CELEN



FUENTE: CELEN