



POSGRADOS

MAESTRÍA EN

SOFTWARE CON MENCIÓN EN
DESARROLLO WEB Y MÓVIL

RPC-SO-34-NO.778-2021

OPCIÓN DE TITULACIÓN:

PROYECTO DE TITULACIÓN CON
COMPONENTES DE INVESTIGACIÓN
APLICADA Y/O DE DESARROLLO

TEMA:

DISEÑO DE UNA ARQUITECTURA
DE SOLUCIÓN DE INTELIGENCIA
DE NEGOCIOS APLICADO A LOS
PROCESOS ACADÉMICOS PARA LA
TOMA DE DECISIONES EN UNA
INSTITUCIÓN DE EDUCACIÓN
SUPERIOR. CASO DE ESTUDIO,
UNIVERSIDAD CENTRAL DEL
ECUADOR.

AUTOR(ES)

KLEVER JAVIER LOACHAMIN QUISILEMA

DIRECTOR:

WALTER FERNANDO GAIBOR
NARANJO, MSC.

QUITO – ECUADOR 2023

Autor(es):



Klever Javier Loachamin Quisilema

Ingeniero Informático

Candidato a Magíster en Software con Mención en Desarrollo Web y Móvil por la Universidad Politécnica Salesiana – Sede Quito.

kloachaminq@est.ups.edu.ec

Dirigido por:



Walter Fernando Gaibor Naranjo

Ingeniero de Sistemas en Computación e Informática

Magíster en Gerencia de Sistemas y Tecnologías de la Información

wgaibor@ups.edu.ec

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la libre difusión de este texto con fines académicos investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS

2023 © Universidad Politécnica Salesiana.

QUITO– ECUADOR – SUDAMÉRICA

Klever Javier Loachamin Quisilema

ARQUITECTURA DE SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS APLICADO A LOS PROCESOS ACADÉMICOS EN LA EDUCACIÓN SUPERIOR.

DEDICATORIA

Quiero dedicar de una manera muy singular el presente proyecto a mi esposa e hija, quienes son mi principal motivación para seguir adelante en mi objetivo por crecer tanto personal como profesionalmente.

También me gustaría dedicar esto a mis padres, quienes me enseñaron que, con dedicación, perseverancia y trabajo constante, es posible alcanzar los más grandes sueños y aspiraciones.

AGRADECIMIENTO

Quiero expresar un agradecimiento sincero hacia mi familia por su apoyo para llevar a cabo mi proyecto hasta su finalización.

Agradezco al Departamento de Tecnologías de la Universidad Central del Ecuador, en particular a su Director, por permitir total libertad para realizar el proyecto de investigación.

Al Ing. Walter Gaibor. MSc, quien supo guiarme en todo momento durante la planeación y desarrollo del proyecto gracias a su complaciente atención, capacidad y experiencia.

A mi esposa e hija, gracias por sus amables palabras de apoyo, su tolerancia y, lo más importante, la ayuda brindada.

TABLA DE CONTENIDO

| | |
|---|----|
| ÍNDICE DE FIGURAS | 7 |
| ÍNDICE DE TABLAS..... | 10 |
| RESUMEN..... | 12 |
| ABSTRACT..... | 13 |
| 1. INTRODUCCIÓN..... | 14 |
| 1.1 ANTECEDENTES | 15 |
| 1.2 DETERMINACIÓN DEL PROBLEMA..... | 16 |
| 1.2.1 ESPECIFICACIÓN DEL PROBLEMA | 16 |
| 1.2.2 FORMULACIÓN DEL PROBLEMA..... | 16 |
| 1.2.3 DELIMITACIÓN DEL PROBLEMA..... | 16 |
| 1.3 JUSTIFICACIÓN DEL PROBLEMA | 17 |
| 1.4 OBJETIVOS..... | 18 |
| 1.4.1 OBJETIVO GENERAL..... | 18 |
| 1.4.2 OBJETIVOS ESPECÍFICOS..... | 18 |
| 2. MARCO TEÓRICO REFERENCIAL | 19 |
| 2.1 ESTADO DEL ARTE..... | 19 |
| 2.1.1 MATRIZ DE RESUMEN DEL ESTADO DEL ARTE | 21 |
| 2.2 MARCO CONCEPTUAL..... | 26 |
| 2.2.1 EVOLUCIÓN Y CONCEPTOS DE BI..... | 26 |
| 2.2.2 CALIDAD DE LA FUENTE DE DATOS..... | 26 |
| 2.2.3 PROCESO ETL | 28 |
| 2.2.4 DATA WAREHOUSE Y DATA MARTS..... | 29 |
| 2.2.5 ARQUITECTURA DE SOLUCIÓN DE BI..... | 34 |
| 2.2.6 PROCESOS DE IMPACTO EN BI..... | 35 |
| 2.2.7 HERRAMIENTAS DE ANÁLISIS Y VISUALIZACIÓN DE DATOS EN BI..... | 36 |
| 3. DESARROLLO DEL PROYECTO | 40 |

| | | |
|--------|--|-----|
| 3.1 | UNIDAD DE ANÁLISIS..... | 40 |
| 3.2 | ANÁLISIS TÉCNICO..... | 40 |
| 3.2.1 | HARDWARE..... | 40 |
| 3.2.2 | SOFTWARE..... | 41 |
| 3.3 | METODOLOGÍA DE DESARROLLO APLICADO A LA SOLUCIÓN BI 41 | |
| 3.3.1 | FASE DE ANÁLISIS | 41 |
| 3.3.2 | FASE DE DISEÑO DEL MODELO MULTIDIMENSIONAL | 42 |
| 3.3.3 | FASE PROCESO ETL..... | 49 |
| 3.3.4 | FASE PROCESAMIENTO ANALÍTICO..... | 60 |
| 4. | RESULTADOS Y DISCUSIÓN | 83 |
| 4.1 | ANÁLISIS DE RESULTADOS..... | 83 |
| 4.2 | DASHBOARD CREADOS..... | 83 |
| 4.2.1 | MATRICULADOS POR ASIGNATURA..... | 83 |
| 4.2.2 | DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR GÉNERO Y SEMESTRE..... | 85 |
| 4.2.3 | PROMEDIO GENERAL DE LOS ESTUDIANTES POR PERIODO ACADÉMICO..... | 87 |
| 4.2.4 | DISTRIBUCIÓN DE LOS ESTUDIANTES POR NÚMERO DE MATRÍCULA..... | 88 |
| 4.2.5 | DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR PROVINCIA..... | 90 |
| 4.2.6 | DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR NACIONALIDAD..... | 91 |
| 4.2.7 | DISTRIBUCIÓN DE LOS ESTUDIANTES POR ASIGNATURA Y ESTADO..... | 92 |
| 4.2.8 | DISTRIBUCIÓN DE LAS ASIGNATURAS POR SU ESTADO Y NÚMERO DE MATRÍCULA..... | 94 |
| 4.2.9 | DISTRIBUCIÓN DE LOS ESTUDIANTES POR ESTADO CIVIL..... | 96 |
| 4.2.10 | DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR LA SITUACIÓN SOCIOECONÓMICA..... | 97 |
| 5. | CONCLUSIONES | 100 |
| | REFERENCIAS | 101 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| FIGURA 1: FASES DE LA METODOLOGÍA PDW..... | 30 |
| FIGURA 2: ESQUEMA ESTRELLA..... | 32 |
| FIGURA 3: ESQUEMA COPO DE NIEVE | 32 |
| FIGURA 4: ESQUEMA CONSTELACIÓN | 33 |
| FIGURA 5: ARQUITECTURA DE SOLUCIÓN BI..... | 34 |
| FIGURA 6: CUADRANTE MÁGICO DE GARTNER..... | 37 |
| FIGURA 7: MODELO ESTRELLA | 43 |
| FIGURA 8: MODELO LÓGICO DEL DW ESQUEMA ESTRELLA | 43 |
| FIGURA 9: TABLA DE DIMENSIÓN PERSONA..... | 44 |
| FIGURA 10: TABLA DE DIMENSIÓN MATERIA | 45 |
| FIGURA 11: TABLA DE DIMENSIÓN PERIODO ACADÉMICO | 45 |
| FIGURA 12: TABLA DE DIMENSIÓN DEPENDENCIA..... | 46 |
| FIGURA 13: TABLA DE DIMENSIÓN CARRERA..... | 46 |
| FIGURA 14: TABLA DE DIMENSIÓN TIEMPO | 47 |
| FIGURA 15: TABLA DE DIMENSIÓN ESTADO | 47 |
| FIGURA 16: TABLA DE DIMENSIÓN PROVINCIA..... | 48 |
| FIGURA 17: TABLA DE DIMENSIÓN SEMESTRE | 48 |
| FIGURA 18: TABLA DE HECHOS CALIFICACIÓN..... | 49 |
| FIGURA 19: ARCHIVO DATA_SOURCE_ACADEMICO.XLSX..... | 50 |
| FIGURA 20: TRANSFORMACIÓN TABLA DSA_CALIFICACION..... | 51 |
| FIGURA 21: TRANSFORMACIÓN TABLA DSA_ESTUDIANTE..... | 51 |
| FIGURA 22: TRANSFORMACIÓN DE TABLA DSA_FAC_CARRERA | 52 |
| FIGURA 23: TRANSFORMACIÓN DIMENSIÓN PERIODO ACADÉMICO.. | 53 |
| FIGURA 24: TRANSFORMACIÓN DIMENSIÓN PERSONA..... | 53 |
| FIGURA 25: TRANSFORMACIÓN DIMENSIÓN MATERIA..... | 54 |
| FIGURA 26: TRANSFORMACIÓN DIMENSIÓN CARRERA | 55 |
| FIGURA 27: TRANSFORMACIÓN DIMENSIÓN FACULTAD..... | 55 |
| FIGURA 28: TRANSFORMACIÓN DIMENSIÓN TIEMPO | 56 |
| FIGURA 29: TRANSFORMACIÓN DIMENSIÓN PROVINCIA..... | 57 |

| | |
|---|----|
| FIGURA 30: TRANSFORMACIÓN DIMENSIÓN ESTADO..... | 57 |
| FIGURA 31: TRANSFORMACIÓN DIMENSIÓN SEMESTRE..... | 58 |
| FIGURA 32: TRANSFORMACIÓN TABLA DE HECHOS ACADÉMICO BI...59 | |
| FIGURA 33: JOB TABLA DE HECHOS ACADÉMICO BI | 59 |
| FIGURA 34: UBICACIÓN DE ADMINTOOL..... | 60 |
| FIGURA 35: PANTALLA PRINCIPAL..... | 61 |
| FIGURA 36: CREAR NUEVO REPOSITORIO | 61 |
| FIGURA 37: ORIGEN DE DATOS..... | 62 |
| FIGURA 38: OBJETO DE METADATOS | 63 |
| FIGURA 39: ACTUALIZAR TODOS LOS RECUENTOS DE FILAS..... | 64 |
| FIGURA 40: MODELO FÍSICO..... | 65 |
| FIGURA 41: NOMBRE TABLA FÍSICA | 65 |
| FIGURA 42: MODELO FÍSICO..... | 66 |
| FIGURA 43: DIAGRAMA FÍSICO..... | 66 |
| FIGURA 44: DIAGRAMA FÍSICO CARGADO..... | 67 |
| FIGURA 45: RELACIONES ENTRE TABLAS | 68 |
| FIGURA 46: DIAGRAMA RELACIÓN ENTRE TABLAS | 69 |
| FIGURA 47: CREACIÓN DEL MODELO LÓGICO | 70 |
| FIGURA 48: UNIÓN LÓGICA | 70 |
| FIGURA 49: DIAGRAMA RELACIÓN LÓGICA | 71 |
| FIGURA 50: ÁREA TEMÁTICA | 72 |
| FIGURA 51: MODELO DE PRESENTACIÓN..... | 73 |
| FIGURA 52: CONSISTENCIA GLOBAL..... | 74 |
| FIGURA 53: RUTA DEL ARCHIVO | 75 |
| FIGURA 54: ARCHIVO DATAMODEL | 75 |
| FIGURA 55: INICIO DE SESIÓN..... | 76 |
| FIGURA 56: HOME DE ANALYTICS..... | 77 |
| FIGURA 57: SELECCIONE ÁREA TEMÁTICA | 78 |
| FIGURA 58: PANEL DE CRITERIOS..... | 78 |
| FIGURA 59: COLUMNAS SELECCIONADAS..... | 79 |
| FIGURA 60: TABLA DE RESULTADOS..... | 80 |
| FIGURA 61: TABLA DINÁMICA..... | 80 |
| FIGURA 62: GRÁFICO DE TARTA..... | 81 |
| FIGURA 63: PANEL DE CONTROL..... | 82 |

| | |
|--|----|
| FIGURA 64: ANÁLISIS MATRICULADOS POR ASIGNATURA..... | 83 |
| FIGURA 65: ANÁLISIS DE MATRICULADOS POR GÉNERO | 85 |
| FIGURA 66: ANÁLISIS PROMEDIO GENERAL POR PERÍODO ACADÉMICO | 87 |
| FIGURA 67: ANÁLISIS ESTUDIANTES MATRICULADOS POR NÚMERO DE MATRÍCULA..... | 88 |
| FIGURA 68: ANÁLISIS MATRICULADOS POR PROVINCIA..... | 90 |
| FIGURA 69: ANÁLISIS MATRICULADOS POR NACIONALIDAD..... | 91 |
| FIGURA 70: ANÁLISIS ESTUDIANTES POR ASIGNATURA Y ESTADO | 93 |
| FIGURA 71: ANÁLISIS DE ASIGNATURAS POR SU ESTADO Y NÚMERO DE MATRÍCULA..... | 94 |
| FIGURA 72: ANÁLISIS MATRICULADOS POR ESTADO CIVIL..... | 96 |
| FIGURA 73: ANÁLISIS MATRICULADOS POR LA SITUACIÓN SOCIOECONÓMICA..... | 97 |

ÍNDICE DE TABLAS

| | |
|---|----|
| TABLA 1: MATRIZ DEL ESTADO DEL ARTE | 21 |
| TABLA 2: REQUERIMIENTOS DE HARDWARE..... | 40 |
| TABLA 3: REQUERIMIENTOS DE SOFTWARE..... | 41 |

ARQUITECTURA DE
SOLUCIÓN DE
INTELIGENCIA DE
NEGOCIOS APLICADA A
LOS PROCESOS
ACADÉMICOS EN LA
EDUCACIÓN SUPERIOR

AUTOR(ES):

KLEVER JAVIER LOACHAMIN QUISILEMA

RESUMEN

El trabajo de investigación describe el proceso para alcanzar el diseño de una arquitectura de solución BI, que se adapte a las necesidades y especificaciones de la Institución de Educación Superior, para utilizar como base de futuros análisis en términos de la gestión académica. Para lo cual se elige una estructura tecnológica que involucra Oracle BI, Pentaho y una metodología de procesamiento Data Warehouse basado en los siguientes principios:

Identificación del giro de negocio: determinar los requisitos, necesidades y especificaciones de la institución.

Diseño de la arquitectura de solución: determinar una arquitectura de solución BI, basado en las tecnologías y herramientas que se utilizan para recopilar, procesar y presentar datos.

Extracción, transformación y carga de datos (ETL): analizar, procesar y almacenar datos que provienen de diferentes fuentes.

Visualización de datos: mostrar la información obtenida de una manera simple y directa, de modo que sea fácil de comprender para los usuarios finales.

Supervisión y mantenimiento: supervisar y mantener al día la solución BI, para que pueda funcionar correctamente y mejorar en el tiempo.

Palabras clave:

Inteligencia de Negocios, Diseño de arquitectura de solución, toma de decisiones, procesos académicos, procesos ETL.

ABSTRACT

The research work describes the process to achieve the design of a BI solution architecture that adapts to the needs and specifications of the Higher Education Institution, to be used as a basis for future analysis in terms of academic management. For which a technological structure is chosen that involves Oracle BI, Pentaho, and a Data Warehouse processing methodology based on the following principles:

Identification of the business line: determine the requirements, need, and specifications of the institution.

Design of the solution architecture: determine a BI solution architecture, based on the technologies and tools used to collect process, and present data.

Data extraction, transformation, and loading (ETL): analyze, process, and store data coming from different sources.

Data visualization: displaying the information obtained in a way, so that it is easy to understand for end users.

Monitoring and maintenance: monitor and keep the BI solution up to date, so that it works properly and improves over time.

Keywords:

Business Intelligence, Solution architecture design, decision making, academic processes, ETL processes.

1. INTRODUCCIÓN

“La proliferación en el desarrollo de aplicaciones web y móviles ha generado una estructura y volumen cada vez mayor, en cuanto a la información que una empresa u organización debe procesar.” (Moscoso et. al, 2019).

En una Institución de Educación Superior (IES), son los estudiantes el factor fundamental, por lo tanto, tener la capacidad de entenderlos y saber de ellos, es de vital importancia. Los sistemas académicos permiten recolectar, almacenar y procesar datos, sin embargo, estos sistemas tradicionales presentan deficiencias en el manejo de los datos, lo que se convierte en una limitante a la hora de generar información relevante para la IES.

Business Intelligence (BI) se puede utilizar con eficacia en la educación, con el fin de mejorar el análisis de los datos sobre los estudiantes y más específicamente comprender a los estudiantes con respecto a su comportamiento académico, tales como: matrículas, notas, graduados, becas, entre otros (Apraxine & Stylianou, 2017).

En este contexto, la BI permite analizar datos a través del proceso ETL (Hernandez et. al, 2021). Estos procedimientos y técnicas empleados por una empresa u organización para almacenar, procesar y analizar los datos favorecen al proceso de toma de decisiones (TD), todo esto basado en datos y cuyo enfoque son el tema de estudio de este trabajo de investigación.

La inteligencia de negocios se apoya en un grupo de tecnologías ETL, capaces de presentar información a los gerentes de una organización y permitir la TD. Sin embargo, la calidad de los datos que se almacenan no siempre es la requerida.

Por lo tanto, es imperativo desarrollar una infraestructura de información que adquiera conocimiento y facilite la experimentación con datos limpios y

consistentes en el proceso educativo (Moscoso et. al, 2019).

1.1 ANTECEDENTES

La Ley Orgánica de Educación Superior (LOES) y las disposiciones de carácter general que rigen el sistema de educación superior garantizan el derecho a una educación de calidad (CES - Consejo de Educación Superior, 2022). En cuanto a la calidad y acreditación de la educación, el Consejo de Calidad de la Educación (CACES) adopta la calidad como principio. En este contexto la acreditación institucional implica un proceso de mejora continua, bajo un ámbito de exigencia y rendición de cuentas relacionadas con toda la actividad académica de la institución.

Hamed et al. (2017), mencionan que las IES están en constante evolución en cuanto a la gestión y calidad de los procesos educativos, pero especialmente enfocados en lo académico. La autonomía reconoce a las universidades la responsabilidad académica, administrativa y financiera. En tal sentido, las IES buscan apoyarse más en los datos almacenados, invertir más en recursos tecnológicos que permitan recolectar y gestionar la información relacionada con el estudiante.

Las empresas u organizaciones utilizan la BI para obtener una vista integral de sus datos y usan esta información en la TD estratégicas que les permita contribuir a sus objetivos y mejorar sus procesos.

En este contexto, el concepto de BI puede aplicarse eficazmente a las IES, para mejorar el análisis de los datos sobre los estudiantes. Más específicamente, a través de la BI se puede analizar el comportamiento de los estudiantes a nivel académico (Apraxine & Stylianou, 2017).

Por lo expuesto anteriormente, es indispensable contar con una herramienta que también garantice la calidad de los mismos. El diseño de una arquitectura de solución de BI moderna permite tener una visión global de los datos de las IES y eliminar fallos en los procesos académicos y adaptarse rápidamente a las necesidades del mercado en el contexto educativo.

1.2 DETERMINACIÓN DEL PROBLEMA

1.2.1 ESPECIFICACIÓN DEL PROBLEMA

La necesidad de la tecnología en la sociedad abarca una variedad de contextos, entre los que se puede mencionar: la salud, la economía, los deportes, la geolocalización, la educación, entre otros. La educación es definitivamente muy importante dado que procesa demasiada información que podría ser utilizada en beneficio de la sociedad.

A propósito de la situación actual de la institución, que será el referente para la presente propuesta de investigación; la institución dispone de un sistema académico, el cual realiza todas las operaciones concernientes a los procesos educativos de la institución, entre ellas se destaca: notas, matrículas, becas, información tanto del estudiante como del docente entre otras.

El sistema académico facilita información que no siempre tiene la calidad esperada, lo que conlleva a las autoridades a tomar decisiones equivocadas para la institución. En este sentido, se puede evidenciar claramente la falta de una plataforma tecnológica especializada en el almacenamiento, procesamiento y análisis de los datos que permita definir el comportamiento académico.

1.2.2 FORMULACIÓN DEL PROBLEMA

¿Cuáles son las ventajas de aplicar una arquitectura de solución de BI en el proceso académico en la universidad?

1.2.3 DELIMITACIÓN DEL PROBLEMA

El presente proyecto contempla el análisis de los conceptos asociados con la BI, metodologías y herramientas, que faciliten el diseño de una arquitectura de solución de BI aplicado a los procesos académicos de la IES. En este sentido, la presente investigación analiza la situación actual de la institución, relacionado con el sistema académico que se emplea para el manejo de los procesos educativos.

Como parte de la investigación se deberá identificar la calidad de la fuente de datos, para posteriormente establecer los procesos ETL. Finalmente, la implementación de herramientas de visualización le permite interactuar con los datos y representar visualmente los resultados e indicadores que sirven como base para la solución de BI.

Debido a la privacidad y el gran volumen de datos almacenados, el proyecto se limita a validar el diseño de la solución, a través de una prueba de concepto aplicada a la Facultad de Ingeniería y Ciencias Aplicadas (FICA), que puede ser replicado a las distintas facultades del IES.

En este sentido, lo que se plantea es diseñar una arquitectura de BI, acorde a las necesidades y requerimientos de la universidad, mas no la aplicación; que servirá de base para futuros análisis, basado en indicadores que faciliten respuestas a las preguntas de negocio en los procesos académicos.

1.3 JUSTIFICACIÓN DEL PROBLEMA

Actualmente, la mayoría de las IES mantienen una gran cantidad de datos que no generan ningún tipo de conocimiento, sin embargo, al ser procesados estos datos se convierten en información, que luego se convierte en conocimiento. Así mismo, la presencia de las herramientas tecnológicas basadas en la BI, incrementan el potencial de las actividades de toda organización al obtener información procesada en tiempo real.

Se debe considerar que a lo largo de los años un sistema informático alberga una gran cantidad de datos y que un momento dado se vuelven inmanejables, es aquí donde las herramientas que constituyen la BI permiten simplificar dicha labor.

El aporte de la presente investigación relaciona los procesos académicos y permite tener información del rendimiento y comportamiento académico de los estudiantes y gestionar dicha información de acuerdo a sus necesidades. Además, se puede implementar modelos para predecir el aumento o disminución de los estudiantes

con el fin de proponer soluciones que minimicen pérdidas y optimizar recursos en la asignación de cupos de las materias dentro de la universidad. De esta forma, será posible identificar vacíos y debilidades institucionales en pro de mejoras.

1.4 OBJETIVOS

1.4.1 OBJETIVO GENERAL

Plantear una arquitectura de solución de BI aplicada a los procesos académicos para la TD en una IES.

1.4.2 OBJETIVOS ESPECÍFICOS

1. Determinar la calidad de la fuente de datos, de acuerdo a la situación actual de la institución.
2. Establecer las herramientas tecnológicas que mejor se adapten a la infraestructura y arquitectura de solución de BI aplicado a los procesos académicos de la institución.
3. Establecer los procesos ETL que alimentan el almacén de datos.
4. Establecer las herramientas de visualización que permitan interpretar la información resultante.
5. Validar el diseño de la solución, a través de una prueba de concepto aplicado a la FICA.

2. MARCO TEÓRICO REFERENCIAL

El alcance del presente proyecto contempla el análisis de los conceptos asociados con la BI, metodologías y herramientas que facilitan el diseño de una arquitectura de solución de BI aplicado a los procesos académicos en una IES.

2.1 ESTADO DEL ARTE

El presente proyecto está sustentado en las siguientes investigaciones que poseen la relevancia y el enfoque del tema de investigación.

Medina et al. (2018), expresan que las IES no solo requieren sistemas de información que controlen y gestionen los procesos académicos y administrativos, además necesitan de sistemas como soluciones BI que faciliten indicadores, que permitan medir el desempeño en cuanto a los procesos académicos.

Gonzaga (2021), manifiesta que no es suficiente con automatizar los procesos que dan continuidad a las operaciones dentro de la institución, sino que se precisa un procedimiento oportuno de los datos almacenados.

En Chile se realizó un estudio sobre los procesos de ingreso de los estudiantes a la Universidad, para ello Fuentes et al. (2010) utilizaron la BI para minimizar la problemática de la admisión y matrícula. Es decir, se incorporó una solución de BI, para gestionar de forma eficiente la información obtenida de dichos procesos.

En el trabajo de Hamed et al. (2017) se aplicó los conceptos de minería de datos como herramienta de BI para predecir los promedios de los estudiantes de la Universidad Internacional Árabe. El algoritmo aplicado es llamado “Máquina de vectores de soporte” que en esencia se tradujo a un proceso de asesoría académica que sigue el curso al que debe inscribirse en el siguiente nivel, en función del rendimiento académico.

En el estudio realizado por Hernández et al. (2021), los autores muestran una solución de BI enfocada al seguimiento de los créditos académicos del estudiante. Uno de los beneficios de la investigación es que se puede adaptar fácilmente a situaciones en particular. Igualmente, el uso de indicadores para en análisis datos se puede aplicar en procesos académicos tales como: rendimiento académico, deserción, egresados, becas, entre otros.

En un estudio realizado sobre los procesos denominados ETL, Stankov (2020) se planteó una arquitectura diferente a la tradicional, en la cual datos procesados se deben cargar de forma completa en las Data Marks y no de manera incremental. Para determinar si la arquitectura propuesta es eficiente se realizó una serie de experimentos; para lo que se procesaron registros en un rango 1000000-3000000 con la creación de 5, 15, y 25 Data Marks, dando como resultado que la latencia de los datos es directamente proporcional a los Data Marks creados.

Un hecho importante de las IES es el crecimiento masivo y la baja calidad de los datos, los cuales no reflejan beneficios para la gestión académica (Peralta Ascue, 2021). Frente a esta tarea tan complicada, se requieren de nuevas técnicas y herramientas capaces de procesar la enorme cantidad de datos generados en los procesos académicos, y con ello extraer conocimiento significativo para la institución.

Según, Joyanes (2019) los datos en bruto extraídos de distintas fuentes deben ser integrados y organizados, para ser analizados y posteriormente puedan servir a los encargados del proceso en la TD.

Un sistema de BI no solo debe preocuparse de mostrar información, sino también debe ayudar del usuario final, a entender la información. La BI puede ser vista como una herramienta común que se utiliza a diario, a igual que las aplicaciones como Microsoft Word y Excel, entre otros.

2.1.1 MATRIZ DE RESUMEN DEL ESTADO DEL ARTE

A continuación, se presenta la matriz concisa y moderna en la que se exponen las diferentes las referencias bibliográficas empleadas como fuentes de consulta para la investigación. Ver la tabla 1.

Tabla 1

Matriz del estado del arte

| Año y lugar de publicación | Autor(s) | Título | Objetivo/hipótesis | Idea principal | Relación con mi tema | Metodologías y técnicas que aplica | Conclusiones y/o aporte |
|----------------------------|--|--|---|--|--|-------------------------------------|---|
| 2020, Bulgaria | Stankov Ivan | Business Intelligent Systems Data Processing | Analizar un sistema de BI basado el procesamiento de datos. | El núcleo de un sistema de BI consiste en extraer, almacenar y transformar los datos disponibles. | Proceso ETL | Artículo científico, formato AIMRAD | La solución de una arquitectura de BI depende de la cantidad de datos que se quieren analizar. La latencia de los datos es mayor cuanto mayor sea el número de almacenes de datos creados y viceversa. |
| 2021, Portugal | Hernández Luz, Barrera Francisco, Mex Diana, Estrada Guadalupe | Data analysis using business intelligence in an educational context | Implementar el análisis de datos usando una arquitectura de solución de BI con un enfoque educacional. | BI como herramienta tecnológica para el seguimiento de créditos académicos. | Data Analysis, procesos ETL. | Artículo científico, formato AIMRAD | Una solución de BI es muy beneficioso en el análisis de datos en un contexto educativo. La solución de BI puede adaptarse a ámbitos educativos más específicos como: becas, egresados, titulados, entre otros, dependiendo del contexto de análisis de datos. |
| 2016, India | Kawaljeet Singh, Ravinder Singh | Designing and Examining a Data Analysis System to Improve Student Education (Case: | Discutir sobre la planificación de un almacén de datos que permita el análisis de la progresión estudiantil en una Universidad. | Una tecnología Data Warehouse que permita superar las limitaciones de los sistemas informáticos transaccionales. | Data Warehouse, procesos ETL, Metodologías Kimball y Bill. | Artículo científico, formato AIMRAD | Luego de aplicar los procesos ETL los datos se almacenan en un Data Warehouse, que permite obtener resultados sobre la progresión estudiantil (estudiantes aprobados, reprobados o u otras situaciones). |

| | | | | | | | |
|-----------------------|--|--|---|--|--|-------------------------------------|--|
| | | University Student Development System) | | | | | |
| 2017, Colombia | Fernández Marshall, Dávila Abraham, Angeleri Paula | Data quality (DQ) applied to an academic business intelligence solution: Lesson learned | Analizar los datos almacenados en el Data Warehouse(DW) utilizando estándares de DQ. | Una fuente de datos de calidad para soluciones de BI compatibles con los estándares ISO. | Procesos ETL, DQ. | Artículo científico, formato AIMRAD | En una arquitectura de BI, no solo es necesario ejecutar el proceso ETL, sino también la DQ una vez ejecutada. |
| 2019, Alicante-España | Moscoso Oswaldo, Castro Jorge | A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education | Desarrollar una arquitectura empresarial híbrida, inteligencia empresarial y sistemas analíticos para la gestión del conocimiento educativo | Big Data como herramienta tecnológica, que gestiona grandes cantidades de datos, se focaliza en la gestión del conocimiento a través del tratamiento de datos. | Data Mining, análisis de aprendizaje, procesamiento analítico en línea (OLAP). | Artículo científico, formato AIMRAD | Los avances en la ciencia y tecnología son factores críticos para el aumento actual en la cantidad, estructura y tamaño de la información en los IES, de ahí que se necesita de herramientas que permita almacenar, procesar y analizar dicha información. |

| | | | | | | | |
|------------------------------|--|---|---|---|----------------------------------|--|--|
| 2020, Alicante- España | Villegas William, Palacios Javier, Luján Sergio | A Business Intelligence Framework for Analyzing Educational Data | Establecer un Framework de BI, que permita; tratar, analizar y presentar información frente a cualquier situación relacionada con el ámbito educativo. | Un proceso que integra modelos y métodos de intercambio de datos puede conducir a un sistema. Lo mismo se puede utilizar en el modelo BI, que trata de resolver problemas educativos. | Data Mining, BI, procesos ETL | Artículo científico, formato AIMRAD | Un framework de BI combinado con técnicas y la minería de datos educativa (EDM) está en constante mejora. |
| 2017 Athens- Greece | Apraxine Dmitri, Stylianou Ermina | Business intelligence in a higher educational institution: The case of University of Nicosia | Investigar como las prácticas de BI se pueden aplicarse a la industria educativa y mejorar los distintos procesos que se manejan. | Los estudiantes son el factor fundamental de éxito de las instituciones de educación, por lo tanto, tener la capacidad de entenderlos y saber de ellos es de vital importancia. | BI | Artículo científico, formato AIMRAD | En el contexto educativo, poseer información de los estudiantes es muy valiosa, sobre todo a la hora monitorear sus fortalezas y debilidades para ayudarlos a mejorar su desempeño en lo académico. |
| 2016, Londres | Ong Vincent | Business Intelligence and Big Data Analytics for Higher Education: | Describir BI y Big Data como solución para el análisis de grandes cantidades de datos en la educación superior. | Las técnicas y modelos avanzados de Big Data se vuelven cada vez más importantes en el manejo de los datos que | BI, Big Data | Artículo científico, formato AIMRAD | Se lleva a cabo un análisis comparativo de modelos de BI de varias universidades y los resultados tienen nuevas ramificaciones para la investigación en curso. |

| | | | | | | | |
|-----------------|---|---|--|--|---|--|---|
| | | Cases from UK Higher Education Institutions | | crecen de manera exponencial. | | | |
| 2018, Chile | Medina Fernando, Fariña Francisco, Castillo Wilson | Un Data Mart (DM) para obtención de indicadores de productividad académica en una universidad | Crear un DM que permita operar con indicadores en la gestión académica. | El desarrollo de un Data Mark para los indicadores de rendimiento académico; como la docencia y la investigación. | DW, Data Mark, ETL, OLAP y metodologías de Kimball y Hefesto | Artículo científico, formato AIMRAD | La metodología utilizada en este proyecto de BI, permite concluir el desarrollo del DW de manera exitosa y oportuna, dejando en evidencia la efectividad de cada una de las etapas establecidas en el proceso de la docencia y la investigación. |
| 2020, España | Deb Rudra, Romero Oscar, Pedersen Torben, Hose Katja. | High-Level ETL for Semantic DW | Construir un ETL de alto nivel para un almacén de datos semántico. | Una parte de este trabajo es definir la formalidad semántica de las operaciones usando una ontología y demostrar matemáticamente la corrección e integridad de las construcciones ETL propuestas. | DW, Data Mark, ETL. | Artículo científico, formato AIMRAD | Se propone un proceso de integración basado en capas y un conjunto de construcciones ETL basadas en RDF de alto nivel requeridas para definir, mapear, extraer, procesar, transformar, integrar, actualizar y cargar datos semánticos. |

Nota. Matriz de resumen del estado del arte. Elaborado por: (Loachamin, 2023).

2.2 MARCO CONCEPTUAL

2.2.1.1 EVOLUCIÓN Y CONCEPTOS DE BI

A continuación, se presentan detalles sobre la evolución de los conceptos de BI, de acuerdo con las distintas fuentes consultadas. El concepto de BI no es nuevo, existe desde hace más de cincuenta años y ha pasado por diversas interpretaciones y definiciones. En 1958, Hans Peter Luhn introduce el concepto por primera vez y lo define como: "habilidad para generar interrelaciones de los hechos que permita cumplir una meta a través de una acción" (Conesa Caralt, 2012).

Un término más formal fue acuñado en 1989 por Howard Dresden, quien lo definió como: "Estrategias y técnicas para mejorar las decisiones empresariales utilizando sistemas de apoyo basados en evidencia". (Venter, 2005).

En el año de 1990, se da la proliferación del concepto de BI en múltiples aplicaciones. Y finalmente, en el año 2000, algunas plataformas comienzan a consolidar sus aplicaciones de BI. Esto también inicia una conversación sobre datos no estructurados en empresas u organizaciones.

Según Medina et al. (2018), es un conjunto de estrategias y herramientas para gestionar el análisis de datos y crear conocimiento dentro de la organización. BI se puede definir como un conjunto de aplicaciones diseñadas para recopilar, limpiar y transformar datos. Otros conceptos formales sobre BI se encuentran en (Stankov, 2020), (Fernandez et. al, 2017), (Zeng et. al, 2006).

El análisis de datos, la extracción de datos, la arquitectura y la visualización de datos se han integrado en BI en los últimos años, lo que permite a las empresas utilizar los datos para TD.

2.2.2 CALIDAD DE LA FUENTE DE DATOS

La calidad de la fuente de datos se debe evaluar en el desarrollo de un proyecto de investigación, basado en la situación actual de la organización.

En el IES que se sugirió para el estudio se encuentran disponibles varios sistemas de monitoreo por computadora. Los principales problemas con los datos obtenidos incluyen, entre otras cosas, la existencia de datos inconsistentes, campos de datos que deberían ser cero o nulos, registros duplicados y datos que contienen caracteres prohibidos.

Cuando se consultó sobre la calidad de la fuente de datos, el administrador de la base de datos de la IES manifestó:

Cualquier proyecto que involucre análisis de datos debe tener en cuenta la calidad de la fuente de datos, al ser de baja calidad puede producir resultados inexactos y sesgados, mientras que una fuente de datos de alta calidad ofrece información precisa, pertinente y actual.

Seguir un modelo de calidad de datos (DQ) en esta situación es crucial porque permite evaluar los datos que se almacenan en un DW. La norma ISO/IEC 25012 “define y analiza los requisitos DQ y la planificación de su evaluación” (Fernández et al. al., 2017).

Por lo tanto, la calidad de datos proporcionada por la institución se considera buena. En general, una fuente de datos de alta calidad debe ser precisa, relevante, completa, actualizada y consistente, es decir, que cumpla con los siguientes parámetros:

Exactitud: Una propiedad que representa el verdadero valor de una propiedad deseada en un contexto dado.

Relevancia: Los datos relevantes son aquellos que son útiles para el proyecto y pueden ayudar a tomar decisiones informadas.

Complejidad: Se refiere a la estructura y la magnitud de los datos. Una fuente de datos compleja puede requerir más tiempo y recursos para analizarla y puede ser difícil de entender.

Actualidad: La actualidad se refiere a la frescura de los datos. Los datos actualizados son aquellos que están al día y reflejan los cambios más recientes en la realidad.

Consistencia: La consistencia se refiere a la uniformidad de los datos. Los datos consistentes son aquellos que tienen la misma estructura y formato en todo el conjunto de datos.

2.2.3 PROCESO ETL

Luego de haber establecido la naturaleza de los datos, se debe comprender el principio básico de BI. Según Fernandez et. al (2017) es el conjunto de procesos ETL, que realizan la transformación de la fuente de datos en información y conocimiento útil dentro de la organización.

Una concepción más moderna del proceso ETL, de acuerdo con Camargo et al. (2016), considera la limpieza de datos como parte del proceso ETL (extraer, limpiar, transformar y cargar) que deben cumplirse antes de almacenar los datos en un DW, con el objetivo de tomar de decisiones. A continuación, se describe las fases de un proceso ETL:

Extracción: Se refiere a las distintas fuentes de datos, ya sean estructurados (sistemas de información, base de datos transaccionales o archivos con formatos definidos), no estructurados (redes sociales) y semiestructurados.

Limpieza: Luego de la extracción, se recupera datos en bruto. En este proceso inicia la DQ, para lo cual; se debe revisar la coherencia entre los datos, se aplica la eliminación de datos duplicados y finalmente se verifica que la fuente no contenga datos sucios, es decir, que no exista ausencia de valor.

Transformación: Una vez conseguido los datos limpios, consistentes y de calidad, estos datos se transforman e integran adaptándose al formato definido por el DW.

En este proceso se busca la homogeneidad e integración de datos, de modo que las consultas por parte de los usuarios sean resueltas en el menor tiempo posible.

Carga: Luego de haber conseguido los nuevos datos debidamente estructurados, el siguiente paso es cargar los datos al sistema destino, es decir, almacenar los datos en un DW.

2.2.4 DATA WAREHOUSE Y DATA MARTS

En un estudio realizado por Camargo et al. (2016) definen un Data Warehouse (DW) como “una colección de base de datos capaz de proporcionar información para la TD.”

Según Joyanes (2019) un DW, es un almacén de datos que proporciona una visión global de la organización, independientemente de cómo lo utilicen los diferentes usuarios. Un Data Mart (DM) se refiere a una base de datos más específica dentro de una organización. Son subconjuntos de datos almacenados en pequeños contenedores y categorizados porque contienen datos específicos de cada área de la organización.

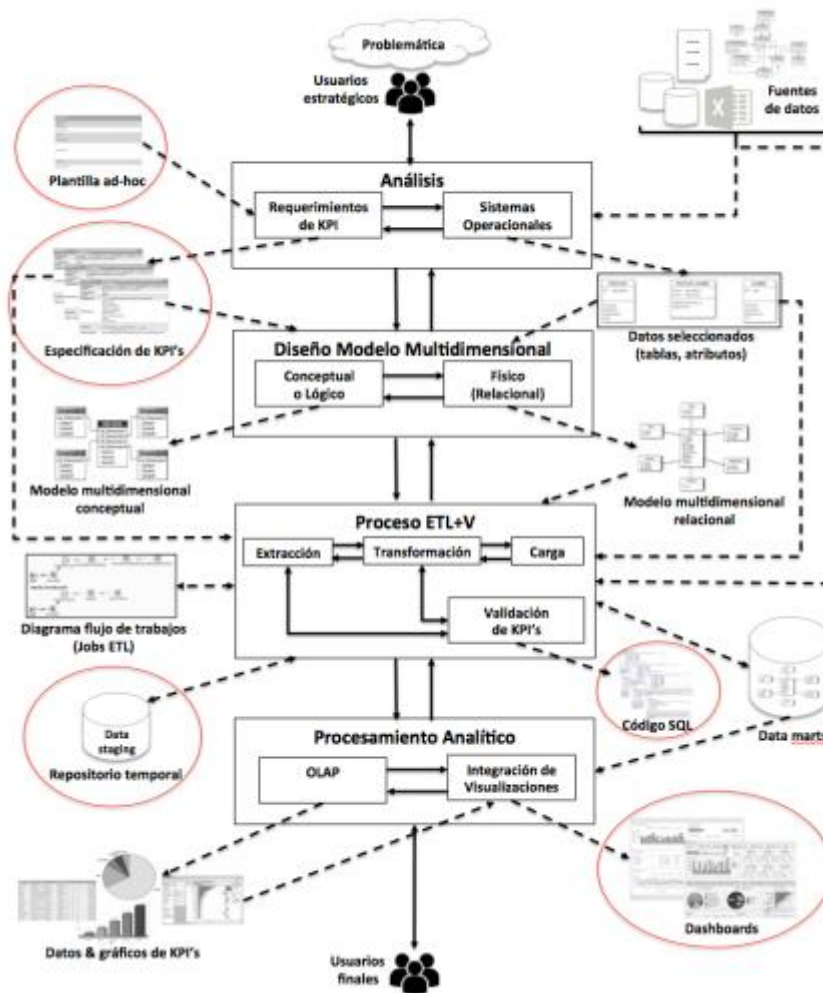
2.2.4.1 METODOLOGÍA DE INVESTIGACIÓN

Como afirman Castillo et al. (2018).” Una metodología para el desarrollo de soluciones de BI, es el proceso de planificación de datos (Data Warehouse, PDW), que incorpora distintos criterios, métodos y metodologías”.

Un PDW, combina las metodologías de Kimball (Kimball & Ross, 2011) y Hefesto (Bernabeu, 2010) para extraer las fases comunes y consolidar una metodología nueva. “El enfoque de Kimball se basa en cuatro principios: centrarse en el negocio, construir la infraestructura de información correcta, entregar de manera significativa y brindar soluciones integrales. Por otro lado, el método Hefesto consta de cinco fases: dirección y planificación, recopilación de información, procesamiento de datos, análisis y producción, y distribución.” (Gonzales & Giancarlo, 2018).

Figura 1

Fases de la metodología PDW



Nota. Estructura de una metodología de construcción de un Data Warehouse.

Elaborado por: tomado de (Joyanes Aguilar, 2019)

En la figura 1, se describe las cuatro fases de una metodología PDW de la siguiente manera:

Fase análisis: El análisis de requisitos y el análisis de sistemas transaccionales son las dos tareas que componen esta fase. La primera actividad se relaciona con las necesidades de la organización, mientras que la segunda actividad se relaciona con la fuente de datos de la organización, ya sea interna o externa. Los datos recopilados

hoy se dividen en tres categorías: estructurados (modelo relacional), no estructurados (redes sociales) y semiestructurados, que combina las dos categorías anteriores.

Fase diseño de un modelo multidimensional (MM): Las salidas de la etapa de análisis, se convierten en las entradas de esta etapa. La fase de diseño es, justamente eso, diseñar un modelo conceptual y físico de base de datos, a través de herramientas asistidas por computadora.

Fase procesó ETL: Esta etapa se crea un almacén de datos, a través de herramientas de BI como: Pentaho, Oracle BI o simplemente consultas SQL. La etapa del proceso ETL tiene mayor complejidad y demanda de más tiempo realizarla, debido a sus entradas y salidas que posee.

Fase procesamiento analítico: Corresponde a esta etapa el análisis y visualización de los resultados obtenidos a través de aplicaciones analíticas, para usuarios del nivel estratégico de la organización, y herramientas de consulta y reporte, para usuarios del nivel táctico del negocio.

2.2.4.2 MODELO MULTIDIMENSIONAL

El modelo dimensional es una composición precisa de normalización y desnormalización de base de datos, de la que se derivan los siguientes elementos:

- La granularidad se refiere a la cantidad de categorías o niveles en los que se divide la información para hacerla más fácil de manejar y comprender.
- La tabla de hechos que contiene información sobre los procesos que desea analizar en detalle.
- Las dimensiones o puntos de vista que resultaron de las necesidades y demandas de la empresa u organización.

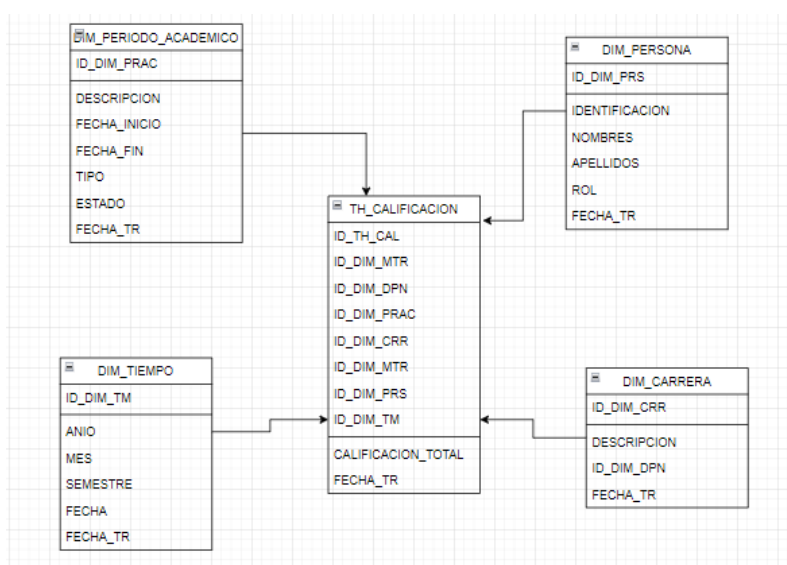
2.2.4.3 ESQUEMAS DE UN MODELO MULTIDIMENSIONAL

Para el desarrollo de un modelo multidimensional existen tres esquemas diferentes:

Esquema de estrella: Es un esquema de base de datos relacional que define dos tipos de tablas, tablas de hechos y dimensiones. Un esquema en estrella se caracteriza por la centralización de tablas de hechos en torno a sus dimensiones. Cada dimensión está conectada directamente al dispositivo central (tabla de hechos), que a su vez actúa como un concentrador de datos. Ver figura 2.

Figura 2

Esquema estrella

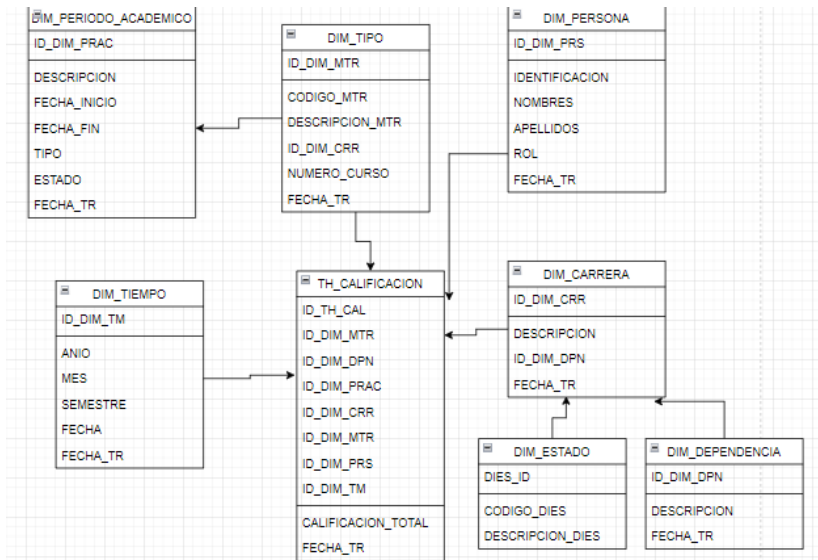


Nota. Estructura de un modelo estrella multidimensional. Elaborado por: (Loachamin, 2023).

Esquema copo de nieve: El siguiente esquema es similar al esquema de estrella, con la diferencia de que cada dimensión está conectada a otras dimensiones. Ver figura 3.

Figura 3

Esquema copo de nieve

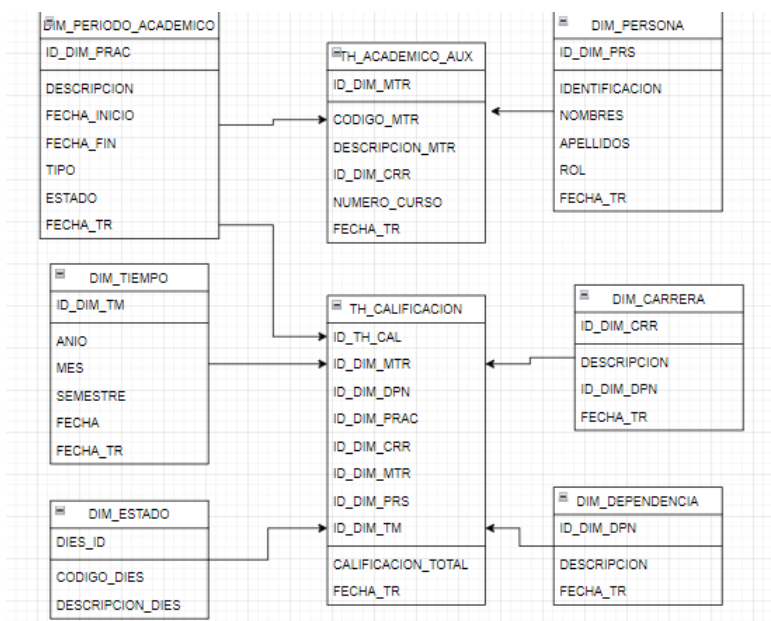


Nota. Estructura de un modelo copo de nieve multidimensional. Elaborado por: (Loachamin, 2023).

Esquema constelación: En este contexto se indica la existencia de otra tabla de hechos que se relaciona entre las dimensiones. Ver figura 4.

Figura 4

Esquema constelación



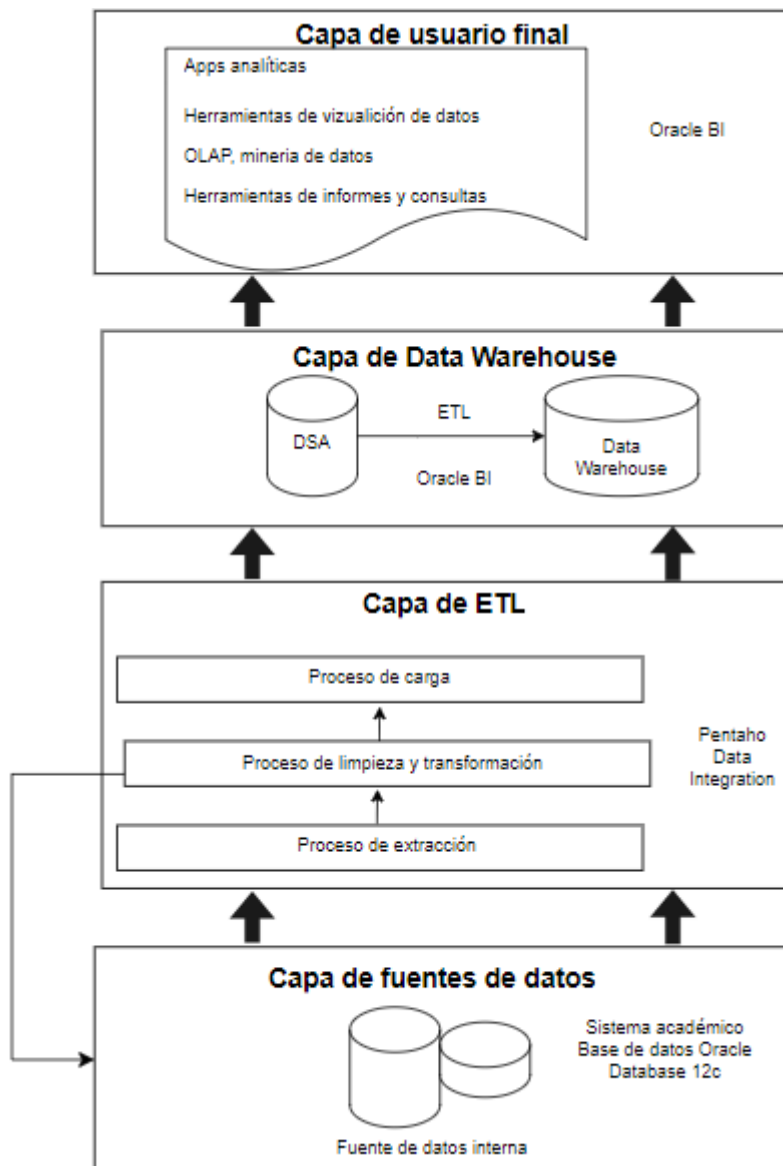
Nota. Estructura de un modelo constelación multidimensional. Elaborado por: (Loachamin, 2023).

2.2.5 ARQUITECTURA DE SOLUCIÓN DE BI

La metodología del PDW empleada en el desarrollo de la solución establece una arquitectura basada en cuatro capas: capa de la fuente de datos, capa de proceso ETL, capa de almacenamiento de datos y la capa de usuario final. Ver figura 5.

Figura 5

Arquitectura de solución BI



Nota. Estructura de un modelo de solución BI aplicando una arquitectura de cuatro capas. Elaborado por: (Loachamin, 2023).

2.2.6 PROCESOS DE IMPACTO EN BI

A continuación, se analiza dos procesos de impacto en BI y que son de vital importancia.

2.2.6.1 MINERÍA DE DATOS

La minería de datos es un proceso que se ha convertido en una tendencia de BI. Según Khan et al. (2014) definieron la minería de datos como el proceso de extraer conocimiento a partir de la detección de patrones en grandes cantidades de datos.

La minería de datos emplea técnicas y herramientas analíticas para conseguir dicho propósito, los cuales son; las supervisadas para el análisis predictivo como: redes neuronales, árboles de decisión y métodos de regresión. Y no supervisadas para el análisis descriptivo, como: método probabilístico, redes neuronales y reglas de asociación. Todas orientadas a encontrar patrones de comportamiento de un conjunto de datos y partir de ellos tomar decisiones.

En este sentido, se utilizan técnicas de minería de datos para extraer conocimiento clave, también conocido como gestión del conocimiento.

2.2.6.2 BIG DATA

Las empresas y organizaciones generan grandes cantidades de datos todos los días y necesitan herramientas para procesar esos datos y transformarlos en conocimiento significativo para su negocio.

Big Data es un proceso diseñado para emplearse conjuntamente con BI, según Joyanes (2019) el término Big Data se refiere a un conjunto enorme y complicado de datos que las herramientas tradicionales son incapaces de procesar, todo esto dentro de un rango de tiempo estimado.

Por lo tanto, el término "Big Data" se acuñó para describir una colección considerable de datos que está más allá del alcance de las herramientas convencionales (bases de datos relacionales) y como la "minería de datos" ha ganado popularidad. Sin embargo, la cantidad de datos en sí, no es el único factor crucial que debe examinarse.

Actualmente, se utilizan datos estructurados, no estructurados y semiestructurados para alimentar las fuentes de datos: datos no estructurados como por ejemplo textos, fotos y audio; datos semiestructurados como archivos de texto, archivos XML y redes sociales; datos estructurados provienen de sistemas transaccionales y bases de datos de relaciones.

Ong (2016), resume el modelo de las 4Vs (volumen, velocidad, variedad y veracidad) de Big Data de la siguiente manera:

Volumen: Es el tamaño global de los datos (ejemplo terabyte).

Velocidad: Esto se refiere al tiempo de generación de datos y la velocidad de procesamiento.

Variedad: Esto se refiere al origen de los datos, ya sean; estructurados, no estructurados y semiestructurados.

Veracidad: Se refiere a la incertidumbre de los datos. En este punto es importante conseguir datos de calidad.

En conclusión, estos datos no pueden ser procesados por simples sistemas de gestión de base de datos, sino, por el contrario, necesitan de técnicas y herramientas diferentes que permitan un procesamiento eficiente y útil.

2.2.7 HERRAMIENTAS DE ANÁLISIS Y VISUALIZACIÓN DE DATOS EN BI

Hay muchas herramientas diferentes de análisis y visualización de datos, que se dividen en las siguientes tres categorías:

- Herramientas de visualización desarrolladas por los encargados de sistemas que cuentan con estrategias para diseñar informes y tableros dinámicos de fácil interpretación – Nivel operacional.
- Las herramientas OLAP – Nivel táctico.
- Aplicaciones analíticas – Nivel estratégico de la empresa u organización.

2.2.7.1 ORACLE BI

Entre las distintas herramientas analizadas se puede mencionar a Power BI como líder en soluciones de BI, tal como se puede evidenciar en el Cuadrante Mágico de Gartner (figura 6), sin embargo, Oracle BI no deja de ser una herramienta completa y potente, con su estrategia de proporcionar capacidades únicas en una plataforma integrada que sea compatible con las tecnologías actuales.

Figura 6

Cuadrante Mágico de Gartner



Nota. Estudio de las herramientas de analítica y BI año 2022. Elaborado por: tomada de (bitec, 2022).

Oracle Business Intelligence (BI) proporciona capacidades de análisis de datos, herramientas de informes y paneles de visualización de datos, entre otros, que permiten a las organizaciones obtener información de sus datos y tomar decisiones basadas en datos.

Oracle BI es una plataforma integral con muchas características y funcionalidades diferentes. Algunas características incluyen lo siguiente:

Integración de datos: Con Oracle BI, puede combinar datos de varias fuentes, lo que facilita la consolidación de datos comerciales y ayuda a brindar a los usuarios una comprensión integral de los datos.

Modelado de datos: Es sencillo crear estructuras de datos organizadas y consistentes gracias a la colección de herramientas de modelado y definición de la herramienta.

Almacenamiento de datos: Oracle BI incluye una opción de almacenamiento de datos en su versión on-premise, y en la nube mediante Oracle Analytics Cloud (OAC). Esto permite a los usuarios almacenar más información y acceder a ella más fácilmente.

Análisis avanzado: La extracción de datos, el aprendizaje automático, el análisis predictivo y el modelado estadístico son solo algunas de las sofisticadas herramientas de análisis disponibles en la plataforma. Los usuarios pueden extraer información útil de los datos utilizando estas funciones.

Reporting: Oracle BI proporciona una variedad de opciones de informes que permiten a los usuarios diseñar informes únicos y compartirlos a través de varios canales.

Dashboarding: La plataforma permite crear paneles interactivos que presentan información de manera visual. Esto permite a los usuarios identificar patrones y tendencias de manera rápida y eficiente.

Seguridad y privacidad: Oracle BI ofrece un alto nivel de seguridad de datos y garantiza la privacidad de la información.

En pocas palabras, Oracle BI es una herramienta integral que proporciona una serie de características de vanguardia para la organización y que permite la TD basada en datos.

2.2.7.2 PENTAHO

Pentaho es una plataforma de BI que proporciona soluciones para la integración, transformación y visualización de datos. Ofrece un conjunto de análisis de datos, extracción de datos, modelado de datos, paneles y herramientas de informes. Al ser una solución de código abierto, se puede integrar fácilmente con otras herramientas comerciales.

3. DESARROLLO DEL PROYECTO

3.1 UNIDAD DE ANÁLISIS

Los datos proporcionados se derivan de un acuerdo de confidencialidad, esto debido a la sensibilidad de la información que se maneja en la IES. Estos datos permitieron establecer consideraciones relevantes y cristalizar el estudio en mención.

La metodología propuesta, denominado PDW, tiene como objetivo generar métricas de gobernanza académica para la universidad. Debido a la naturaleza del estudio, solo se consideran el análisis y el diseño, no la implementación. El método utilizado se considera cualitativo porque se centra en un caso de estudio(UCE).

En este apartado se realiza la verificación del diseño de la solución, a través de una prueba de concepto aplicado a la FICA.

3.2 ANÁLISIS TÉCNICO

La solución de BI para los procesos académicos de la institución requiere de los siguientes recursos dados, los requerimientos mínimos de hardware y software para el diseño del DW, el proceso ETL y la visualización de la información.

3.2.1 HARDWARE

Tabla 2

Requerimientos de hardware

| Requerimientos | Valor |
|-----------------------|-------------------|
| Memoria RAM | 4 GB mínimo |
| Disco Duro | 12 GB |
| Procesador | 3.6 GHz de 64 bit |

Nota. Detalle de los requerimientos de hardware. Elaborado por: (Loachamin, 2023).

3.2.2 SOFTWARE

Tabla 3

Requerimientos de software

| Requerimientos | Valor |
|-----------------------|---|
| Sistema operativo | Windows 10 pro |
| Base de datos | Oracle Data Base 12c |
| Herramientas de BI | Oracle BI 12c, Web Logic Server y Pentaho |

Nota. Detalle de los requerimientos de software. Elaborado por: (Loachamin, 2023).

3.3 METODOLOGÍA DE DESARROLLO APLICADO A LA SOLUCIÓN BI

Las cuatro fases de la metodología PDW que se utilizó para crear la solución de BI, permiten el diseño adecuado del Data Warehouse y el posterior uso de herramientas de BI como Oracle BI y Pentaho Data Integration para explotar los datos.

3.3.1 FASE DE ANÁLISIS

3.3.1.1 ANÁLISIS DE REQUERIMIENTOS

El análisis del modelo del negocio es el punto inicial del proyecto y a su vez marca el cimiento para las futuras etapas del mismo. De acuerdo con el análisis de los procesos académicos, se pudo identificar las perspectivas y variables que apoyaron en la obtención de los posibles indicadores:

- Obtener el total de estudiantes matriculados por género y periodo académico.
- Obtener el promedio total que permita cuantificar el desempeño de los estudiantes por periodo académico de toda la Facultad.
- Calcular el número total de estudiantes aprobados, reprobados y retirados por período académico.
- Obtener el promedio total de estudiantes matriculados para analizar la proyección de apertura de cursos.
- Determinar el número de matriculados por provincia.

Así también, como producto del análisis se determinó las siguientes perspectivas: persona, periodo académico, carrera, dependencia, materia, estado, tiempo, provincia y semestre.

3.3.1.2 ANÁLISIS DE LOS SISTEMAS TRANSACCIONALES

En este apartado, se analizó los sistemas informáticos tradicionales que almacenan los datos de las operaciones concernientes a los procesos educativos de la institución. En ese marco, la entrada principal la compone una fuente de datos proveniente de una base de datos Oracle 12c.

3.3.2 FASE DE DISEÑO DEL MODELO MULTIDIMENSIONAL

Luego de haber definido una arquitectura basada en la metodología del PDW, se procedió a construir un MM a partir de un modelo relacional de base de datos, además se identificó las dimensiones, los hechos y el modelo a utilizar (modelo estrella).

De acuerdo con el modelo en estrella se obtuvo una tabla de hechos; que presenta los procesos académicos a analizar, y las dimensiones; que representan las necesidades y requerimientos de la institución. En ambos casos las tablas están compuestas por tres elementos: la clave primaria, la clave foránea y las medidas que corresponden a las columnas cuantificables.

Como parte de la solución de BI para procesos académicos, se estableció un modelo con las diversas perspectivas (dimensiones) con relación al contexto de la investigación y con base en las consideraciones previas.

Figura 7

Modelo estrella



Nota. Estructura de un esquema de tipo estrella para el modelo multidimensional. Elaborado por: (Loachamin, 2023).

3.3.2.1 MODELO LÓGICO DATA WAREHOUSE

La estructura de las dimensiones y la tabla de hechos que son componentes de la solución son descritas por el modelo lógico del DW.

Figura 8

Modelo lógico del DW esquema estrella

Nota. Estructura de la dimensión persona. Elaborado por: (Loachamin, 2023).

3.3.2.3 DIMENSIÓN MATERIA

La dimensión materia almacenó toda la información relacionada con las materias de la malla curricular y agrupadas por semestres; los campos registrados fueron: el id de dimensión materia, materia, código materia, código malla, malla y semestre.

Figura 10

Tabla de dimensión materia

| DIM_MATERIA |
|-------------|
| DIMT_ID |
| COD_MATERIA |
| MATERIA |
| COD_MALLA |
| MALLA |
| SEMESTRE |

Nota. Estructura de la dimensión materia. Elaborado por: (Loachamin, 2023).

3.3.2.4 DIMENSIÓN PERIODO ACADÉMICO

La dimensión periodo académico almacenó toda la información referente a los periodos académicos cursados; los campos que se registraron fueron: el id de dimensión periodo académico, descripción, fecha inicio y fin del periodo académico.

Figura 11

Tabla de dimensión periodo académico

| DIM_PERIODO_ACADEMICO |
|-----------------------|
| DIPA_ID |
| DESCRIPCION |
| FECHA_INICIO |
| FECHA_FIN |

Nota. Estructura de la dimensión de periodo académico. Elaborado por: (Loachamin, 2023).

3.3.2.5 DIMENSIÓN FACULTAD

La dimensión facultad almacenó toda la información de las facultades de la Universidad Central, con la finalidad de algún momento se implemente una solución BI para toda la Universidad; los campos que se registraron fueron: el id de dimensión facultad, código de facultad y descripción.

Figura 12

Tabla de dimensión dependencia

| |
|--------------|
| DIM_FACULTAD |
| DIFA_ID |
| COD_FACULTAD |
| DES_FACULTAD |

Nota. Estructura de la dimensión dependencia. Elaborado por: (Loachamin, 2023).

3.3.2.6 DIMENSIÓN CARRERA

La dimensión carrera almacenó toda la información relacionada con las carreras disponibles; los campos que se registraron fueron: el id de dimensión carrera, código y descripción de carrera.

Figura 13

Tabla de dimensión carrera.

| |
|----------------|
| DIM_CARRERA |
| DICR_ID |
| CODIGO_CARRERA |
| DES_CARRERA |

Nota. Estructura de la dimensión carrera. Elaborado por: (Loachamin, 2023).

3.3.2.7 DIMENSIÓN TIEMPO

La dimensión tiempo almacenó toda la información referente al tiempo; los campos que se registraron fueron: el id de dimensión tiempo, fecha, año, mes, trimestre, día, descripción mes y día.

Figura 14

Tabla de dimensión tiempo

| DIM_TIEMPO |
|------------|
| DITM_ID |
| FECHA |
| ANIO |
| TRIMESTRE |
| MES |
| DIA |
| DES_MES |
| DES_DIA |

Nota. Estructura de la dimensión tiempo. Elaborado por: (Loachamin, 2023).

3.3.2.8 DIMENSIÓN ESTADO

La dimensión estado almacenó toda la información referente al estado del estudiante; los campos que se registraron fueron: el id de dimensión estado, código estado, descripción estado.

Figura 15

Tabla de dimensión estado

| DIM_ESTADO |
|--------------------|
| DIES_ID |
| CODIGO_ESTADO |
| DESCRIPCION_ESTADO |

Nota. Estructura de la dimensión estado. Elaborado por: (Loachamin, 2023).

3.3.2.9 DIMENSIÓN PROVINCIA

La dimensión provincia almacenó toda la información referente al lugar de procedencia del estudiante; los campos que se registraron fueron: el id de dimensión provincia, código provincia, descripción provincia.

Figura 16

Tabla de dimensión provincia

| |
|------------------|
| DIM_PROVINCIA |
| DIPV_ID |
| CODIGO_PROVINCIA |
| DES_PROVINCIA |

Nota. Estructura de la dimensión provincia. Elaborado por: (Loachamin, 2023).

3.3.2.10 DIMENSIÓN SEMESTRE

La dimensión semestre almacenó toda la información referente al número de semestres de la carrera; los campos que se registraron fueron: el id de dimensión semestre, código semestre, descripción semestre.

Figura 17

Tabla de dimensión semestre

| |
|--------------|
| DIM_SEMESTRE |
| DISM_ID |
| SEMESTRE |
| DES_SEMESTRE |

Nota. Estructura de la dimensión semestre. Elaborado por: (Loachamin, 2023).

3.3.2.11 TABLA DE HECHOS CALIFICACIÓN

La tabla de hechos denominada calificación es la tabla principal donde se almacenó, todos los indicadores de calificaciones relacionados con el estudiante, con esta información y dependiendo de lo que necesite el usuario final, se presentó los informes para la TD. La tabla de hecho tiene la siguiente estructura.

Figura 18

Tabla de hechos calificación

| TH_ACADEMICO_BI |
|------------------|
| THAC_ID |
| DIPR_ID (FK) |
| DIFA_ID (FK) |
| DICR_ID (FK) |
| DIMT_ID (FK) |
| DIPA_ID (FK) |
| DIES_ID (FK) |
| DITM_ID (FK) |
| DIPV_ID (FK) |
| NUMERO_MATRICULA |
| CREDITOS |
| NOTA_FINAL |
| ASISTENCIA_FINAL |

Nota. Estructura de la tabla de hechos denominada calificación. Elaborado por: (Loachamin, 2023).

3.3.3 FASE PROCESO ETL

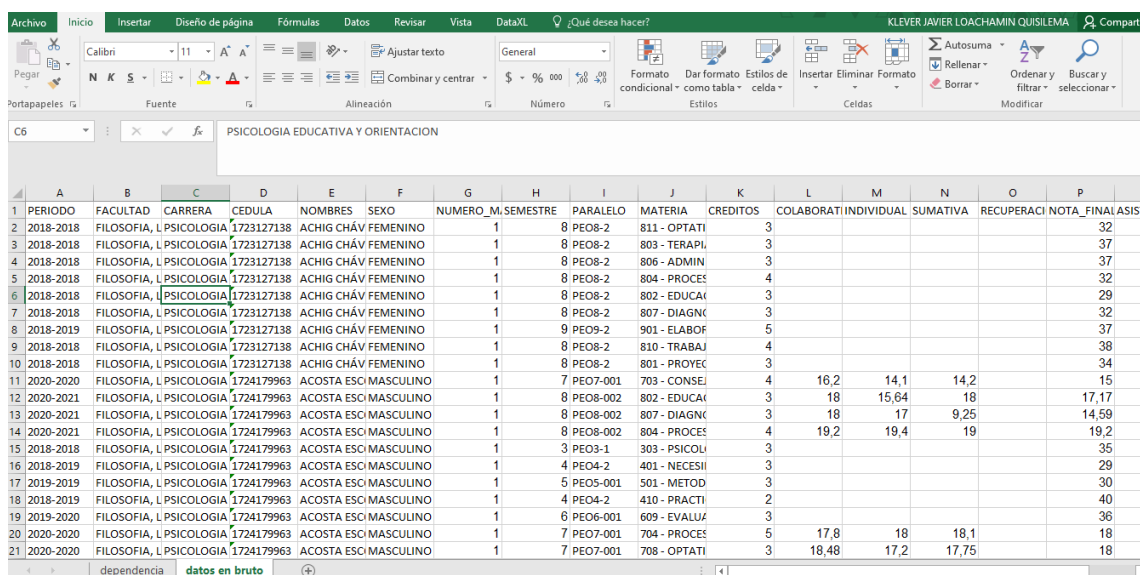
Se eligió Pentaho como la herramienta para crear el proceso ETL y Oracle BI como la herramienta de visualización de los resultados y una fuente de datos almacenada en Oracle Database. Todos los procesos incorporados insertan una columna denominada FECHA_CARGA_ETL para identificar la fecha en que se cargaron los datos.

3.3.3.1 ARCHIVO FUENTE DE DATOS

El archivo principal(data_source_academico.xlsx) proporcionado por el departamento de tecnologías de la institución para el presente estudio, contiene datos relacionados con el estudiante (carreras, materias, notas, créditos, entre otros) básicamente lo que se conoce como datos en bruto. Este archivo es la fuente de datos principal para el DW que será actualizado y una vez que haya finalizado el semestre.

Figura 19

Archivo data_source_academico.xlsx



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|----|-----------|--------------|------------|------------|-----------------------|------|----------|----------|----------|---------------|----------|------------|------------|----------|------------|------------------|
| | PERIODO | FACULTAD | CARRERA | CEDULA | NOMBRES | SEXO | NUMERO_M | SEMESTRE | PARALELO | MATERIA | CREDITOS | COLABORATI | INDIVIDUAL | SUMATIVA | RECUPERACI | NOTA_FINAL ASIST |
| 2 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 811 - OPTATI | 3 | | | | | 32 |
| 3 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 803 - TERAPI | 3 | | | | | 37 |
| 4 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 806 - ADMIN | 3 | | | | | 37 |
| 5 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 804 - PROCES | 4 | | | | | 32 |
| 6 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 802 - EDUCAI | 3 | | | | | 29 |
| 7 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 807 - DIAGNC | 3 | | | | | 32 |
| 8 | 2018-2019 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 9 | PE09-2 | 901 - ELABOF | 5 | | | | | 37 |
| 9 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 810 - TRABAJ | 4 | | | | | 38 |
| 10 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1723127138 | ACHIG CHÁV FEMENINO | | 1 | 8 | PE08-2 | 801 - PROVEC | 3 | | | | | 34 |
| 11 | 2020-2020 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 7 | PE07-001 | 703 - CONSEJ | 4 | 16,2 | 14,1 | 14,2 | | 15 |
| 12 | 2020-2021 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 8 | PE08-002 | 802 - EDUCAI | 3 | 18 | 15,64 | 18 | | 17,17 |
| 13 | 2020-2021 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 8 | PE08-002 | 807 - DIAGNC | 3 | 18 | 17 | 9,25 | | 14,59 |
| 14 | 2020-2021 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 8 | PE08-002 | 804 - PROCES | 4 | 19,2 | 19,4 | 19 | | 19,2 |
| 15 | 2018-2018 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 3 | PE03-1 | 303 - PSICOLI | 3 | | | | | 35 |
| 16 | 2018-2019 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 4 | PE04-2 | 401 - NECESII | 3 | | | | | 29 |
| 17 | 2019-2019 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 5 | PE05-001 | 501 - METOD | 3 | | | | | 30 |
| 18 | 2018-2019 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 4 | PE04-2 | 410 - PRACTI | 2 | | | | | 40 |
| 19 | 2019-2020 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 6 | PE06-001 | 609 - EVALUA | 3 | | | | | 36 |
| 20 | 2020-2020 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 7 | PE07-001 | 704 - PROCES | 5 | 17,8 | 18 | 18,1 | | 18 |
| 21 | 2020-2020 | FILOSOFIA, L | PSICOLOGIA | 1724179963 | ACOSTA ESCI MASCULINO | | 1 | 7 | PE07-001 | 708 - OPTATI | 3 | 18,48 | 17,2 | 17,75 | | 18 |

Nota. Estructura del archivo principal data_source_academico.xlsx que contiene datos en bruto. Elaborado por: (Loachamin, 2023).

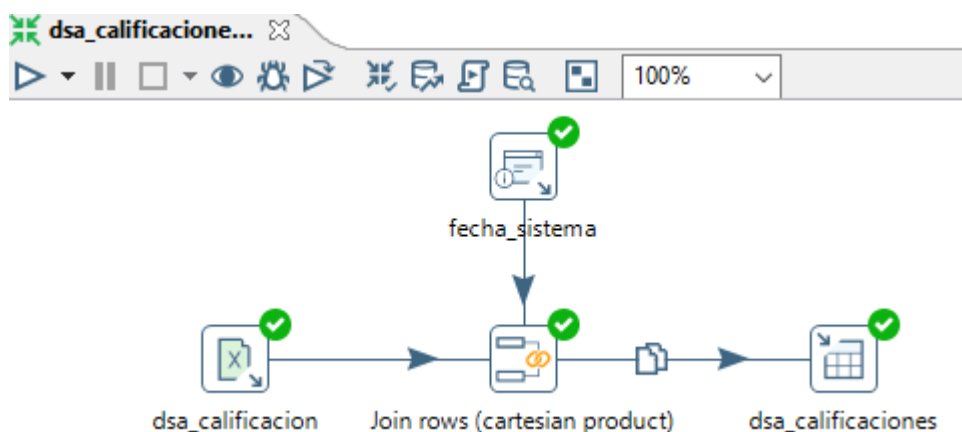
Un modelo que facilita la extracción de información de las fuentes es el área de escenario de datos (DSA), cuyo objetivo es no interferir en el funcionamiento de las fuentes y almacenar los datos en la base de datos DW. Con base en la fuente de datos proporcionada se creó el modelo DSA conformada por tres tablas definidas a continuación:

3.3.3.1.1 DSA_CALIFICACION

Este proceso se encargó de extraer los datos almacenados en el archivo data_source_academico.xlsx (hoja calificaciones) para luego guardarlos en la tabla DSA_CALIFICACION.

Figura 20

Transformación tabla dsa_calificacion



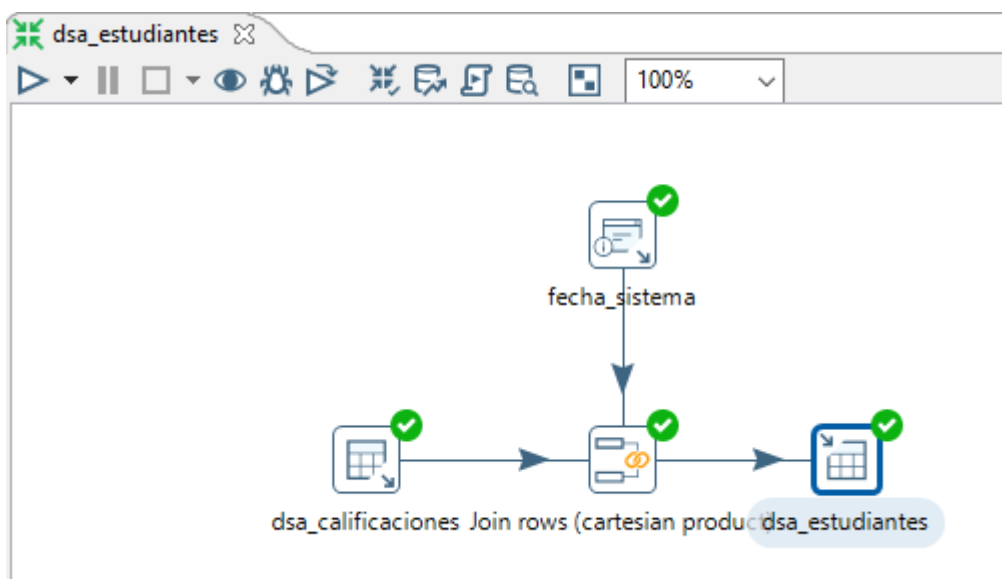
Nota. Proceso de transformación de la tabla dsa_calificacion a partir del archivo principal de la fuente de datos. Elaborado por: (Loachamin, 2023).

3.3.3.1.2 DSA_ESTUDIANTE

Este proceso se encargó de extraer los datos de los alumnos almacenados en la tabla dsa_calificacion y posteriormente almacenarlos en la tabla DSA_ESTUDIANTE.

Figura 21

Transformación tabla dsa_estudiante



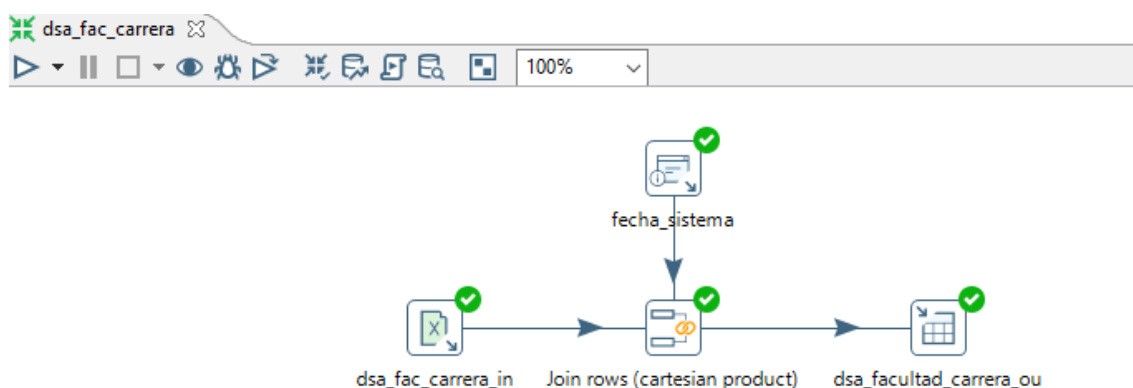
Nota. Proceso de transformación de la tabla dsa_estudiante. Elaborado por: (Loachamin, 2023).

3.3.3.1.3 DSA_FAC_CARRERA

Este proceso se encargó de extraer los datos de las facultades con sus carreras correspondientes, almacenados en el archivo data_source_academico.xlsx (hoja fac_carrera), para luego guardarlos en la tabla DSA_FAC_CARRERA.

Figura 22

Transformación de tabla dsa_fac_carrera



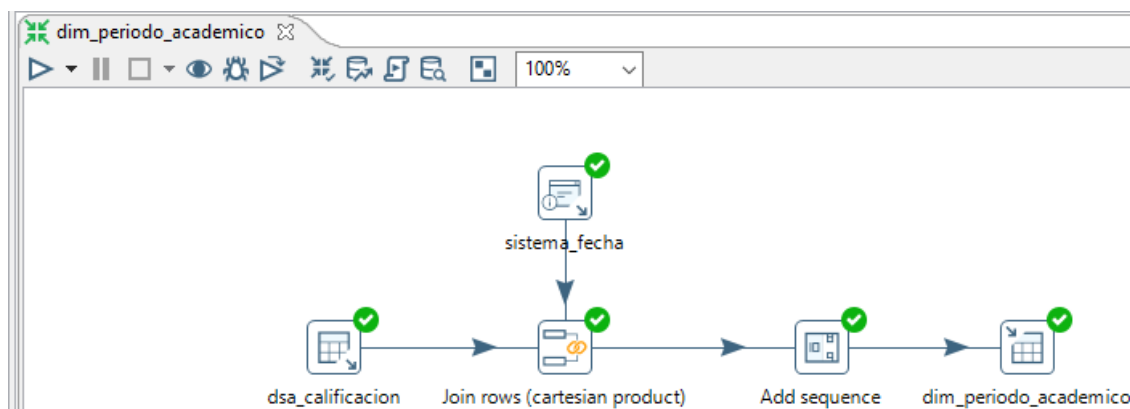
Nota. Proceso de transformación de la tabla dsa_fac_carrera a partir del archivo principal de la fuente de datos. Elaborado por: (Loachamin, 2023).

3.3.3.2 DIMENSIÓN PERIODO ACADÉMICO

Para la dimensión del periodo académico se extrajo los datos de la tabla temporal `dsa_calificacion` y se cargó en una nueva tabla llamada `DIM_PERIODO_ACADEMICO` como producto de la transformación ejecutada. En la transformación se añade un campo fecha del sistema y se aplica producto cartesiano (join rows) para luego agregar una secuencia para el número de registros que se guardarán en la tabla dimensión.

Figura 23

Transformación dimensión periodo académico



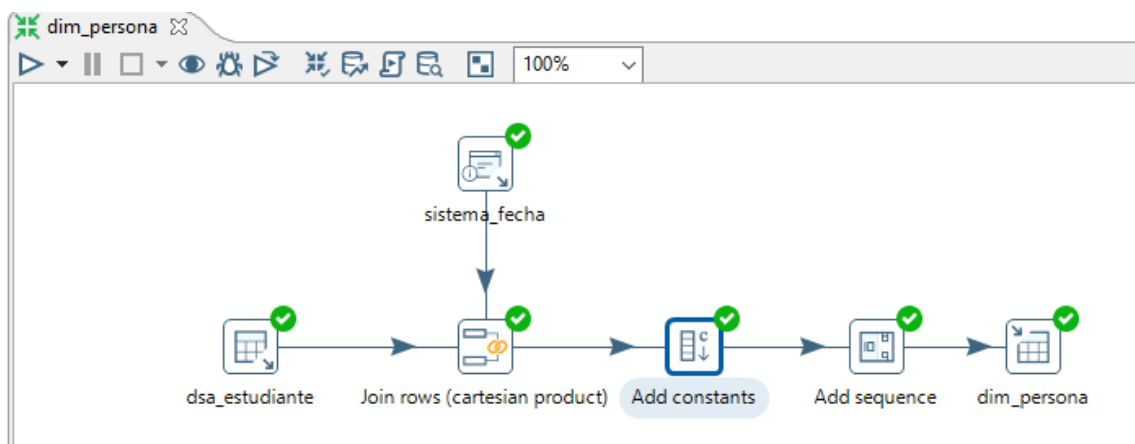
Nota. Proceso de transformación de la dimensión periodo. Elaborado por: (Loachamin, 2023).

3.3.3.3 DIMENSIÓN PERSONA

Para la dimensión persona se extrae los datos de la tabla temporal `dsa_estudiante` y se carga en una nueva tabla llamada `DIM_PERSONA` como producto de la transformación ejecutada. Dentro de este proceso se define una columna llamada `rol`, que permitió cargar tanto estudiantes como docentes para futuros estudios.

Figura 24

Transformación dimensión persona



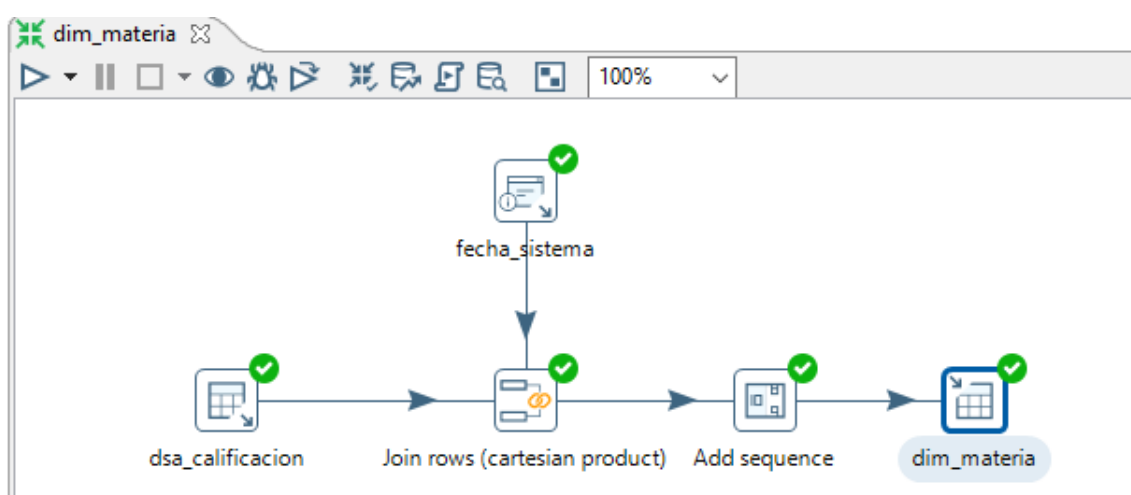
Nota. Proceso de transformación de la dimensión persona. Elaborado por: (Loachamin, 2023).

3.3.3.4 DIMENSIÓN MATERIA

Para la dimensión materia se extrae los datos de la tabla temporal dsa_calificacion y se carga en una nueva tabla llamada DIM_MATERIA como producto de la transformación ejecutada. En este proceso se quitan valores duplicados de las materias.

Figura 25

Transformación dimensión materia.



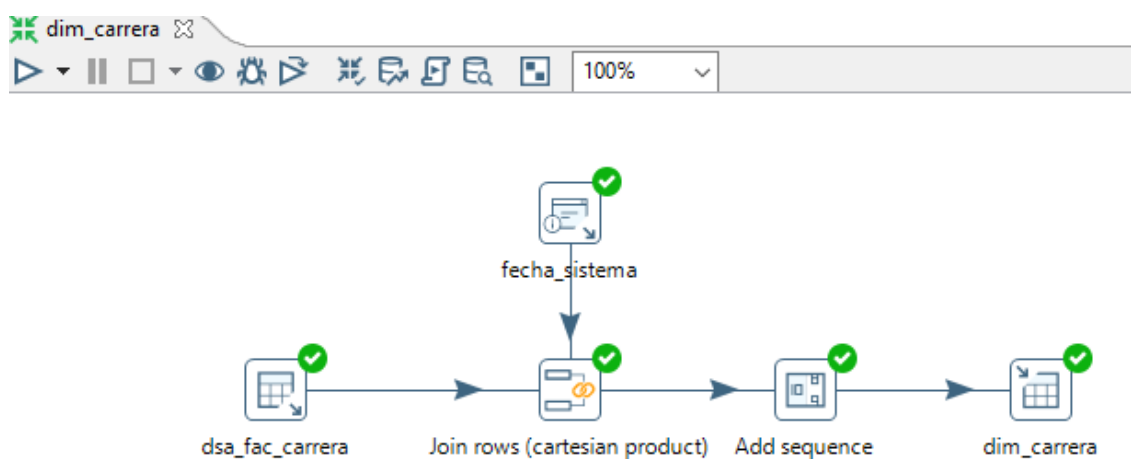
Nota. Proceso de transformación de la dimensión materia. Elaborado por: (Loachamin, 2023).

3.3.3.5 DIMENSIÓN CARRERA

Para la dimensión carrera se extrajo los datos de la tabla `dsa_fac_carrera` y se cargó en una nueva tabla llamada `DIM_CARRERA` como producto de la transformación ejecutada.

Figura 26

Transformación dimensión carrera



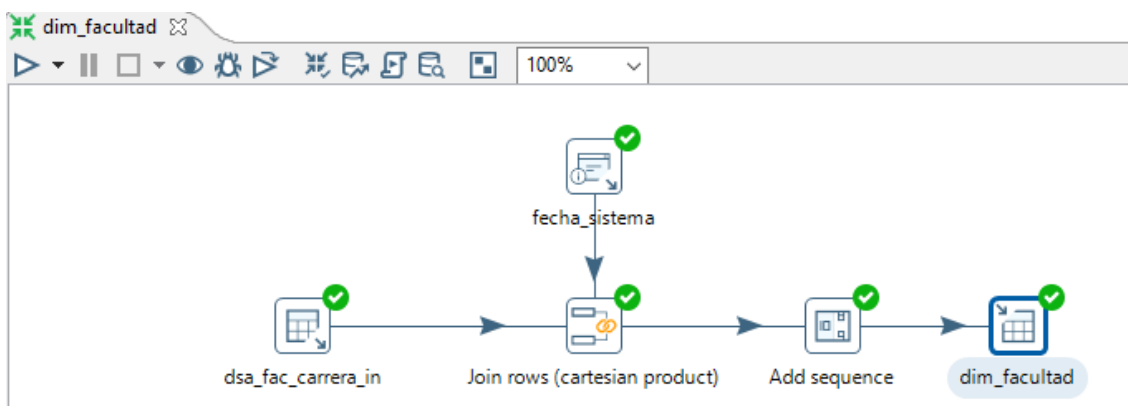
Nota. Proceso de transformación de la dimensión carrera. Elaborado por: (Loachamin, 2023).

3.3.3.6 DIMENSIÓN FACULTAD

Este proceso de transformación fue diseñado para extraer los datos de la dimensión dependencia desde la tabla `dsa_fac_carrera` y cargarlos en una nueva tabla llamada `DIM_FACULTAD` como producto de la transformación ejecutada.

Figura 27

Transformación dimensión facultad



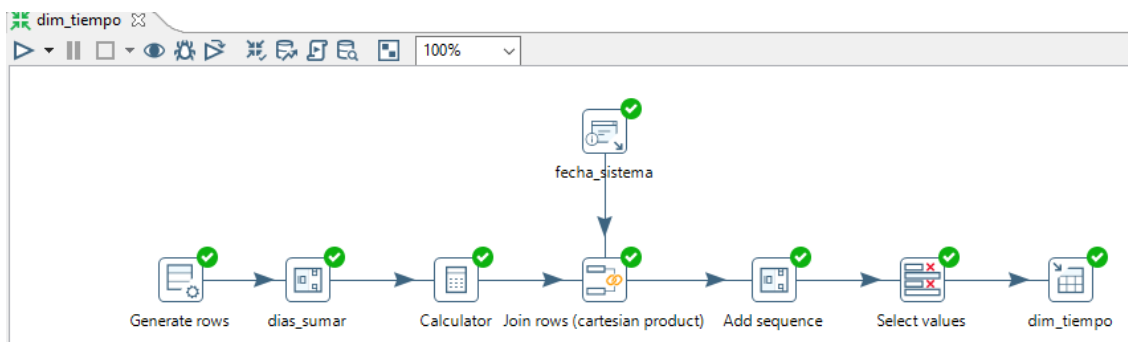
Nota. Proceso de transformación de la dimensión facultad. Elaborado por: (Loachamin, 2023).

3.3.3.7 DIMENSIÓN TIEMPO

En este proceso, la dimensión de fecha no tiene una tabla de origen y los transformadores proporcionados por la herramienta Pentaho se utilizan para generar de forma ficticia todas las fechas requeridas y almacenarlas en la tabla DIM_TIEMPO.

Figura 28

Transformación dimensión tiempo



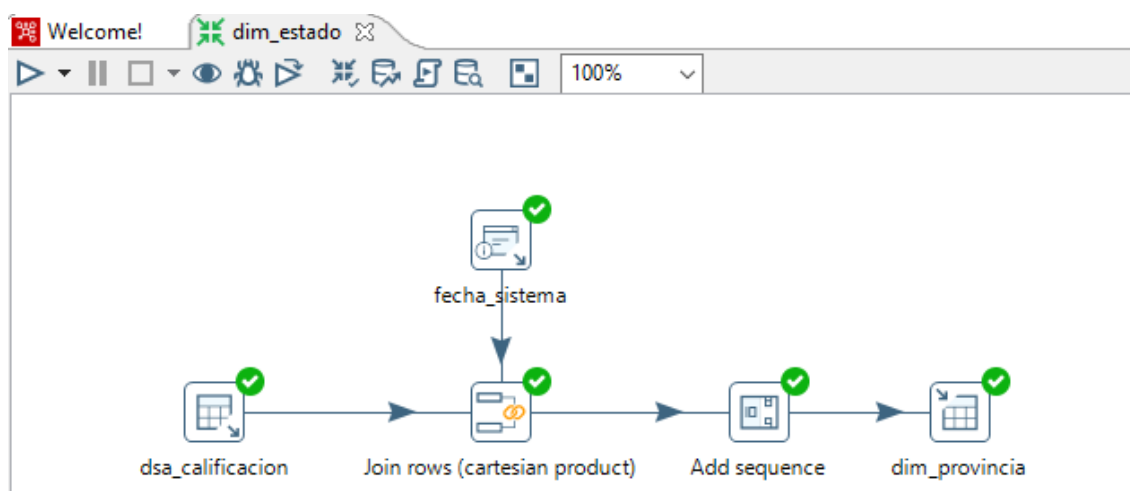
Nota. Proceso de transformación de la dimensión tiempo. Elaborado por: (Loachamin, 2023).

3.3.3.8 DIMENSIÓN PROVINCIA

Este proceso de transformación fue diseñado para extraer los datos de la dimensión dependencia desde la tabla `dsa_fac_carrera` y cargarlos en una nueva tabla llamada `DIM_PROVINCIA` como producto de la transformación ejecutada.

Figura 29

Transformación dimensión provincia



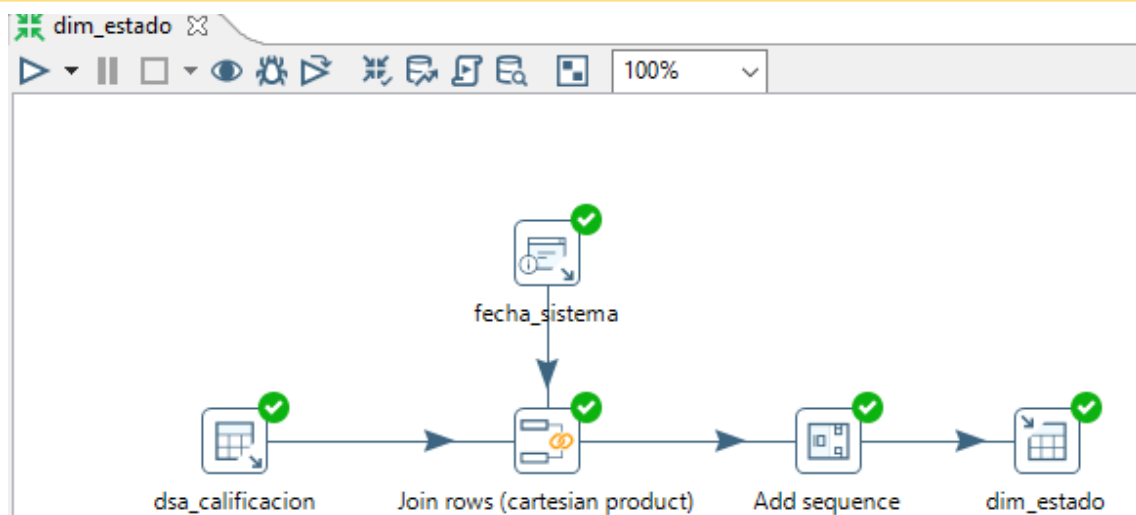
Nota. Proceso de transformación de la dimensión provincia. Elaborado por: (Loachamin, 2023).

3.3.3.9 DIMENSIÓN ESTADO

Este proceso de transformación fue diseñado para extraer los datos de la tabla temporal `dsa_calificacion`, para luego cargarlos en la tabla `DIM_ESTADO` como producto de la transformación ejecutada. En este proceso se depuran los datos duplicados.

Figura 30

Transformación dimensión estado



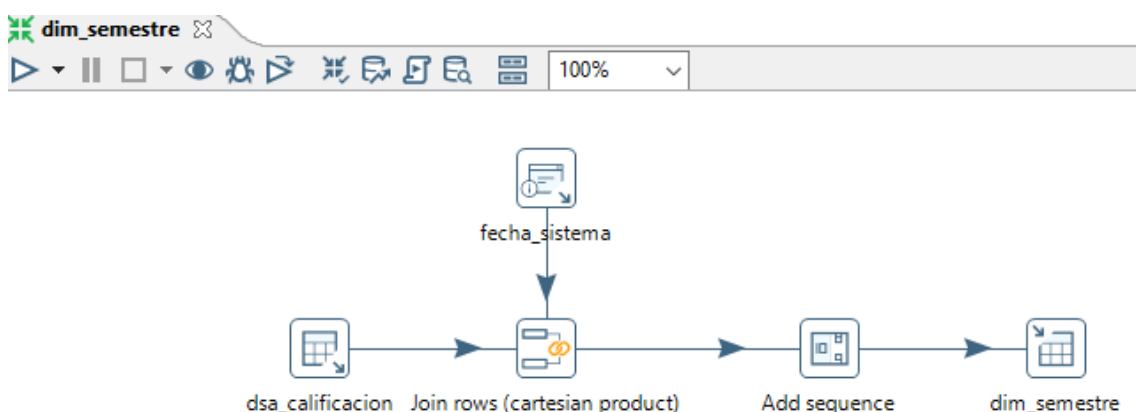
Nota. Proceso de transformación de la dimensión estado. Elaborado por: (Loachamin, 2023).

3.3.3.10 DIMENSIÓN SEMESTRE

Este proceso de transformación fue diseñado para extraer los datos de la tabla temporal *dsa_calificacion*, para luego cargarlos en la tabla *DIM_SEMESTRE* como producto de la transformación ejecutada. En este proceso que depuran los datos duplicados.

Figura 31

Transformación dimensión semestre.



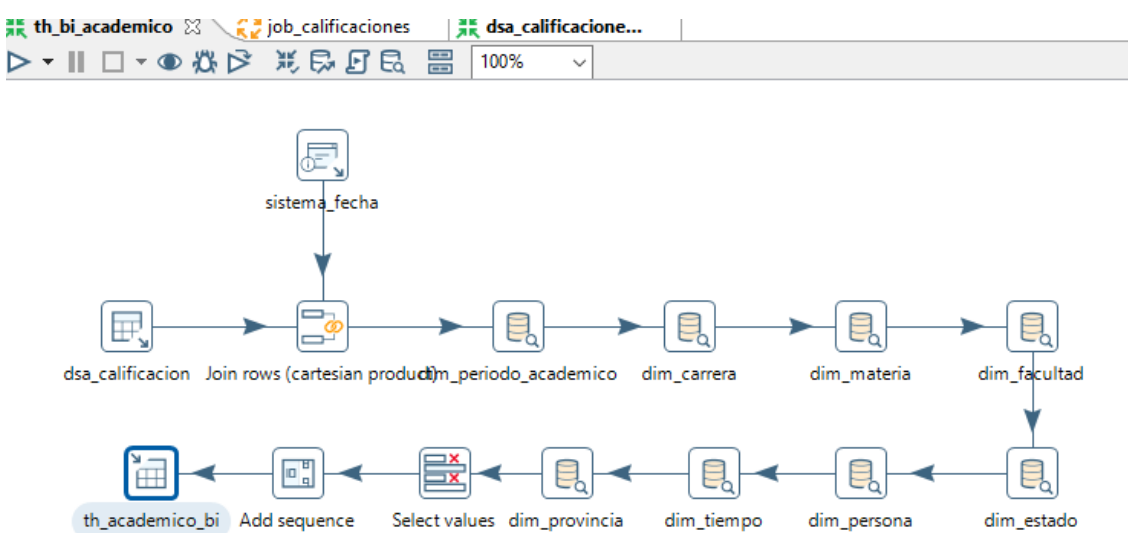
Nota. Proceso de transformación de la dimensión semestre. Elaborado por: (Loachamin, 2023).

3.3.3.11 TABLA DE HECHOS ACADÉMICO BI

Para la tabla de hechos se extrajo los datos de la tabla dsa_calificacion y se cargó en una nueva tabla llamada TH_ACADEMICO_BI como producto de la transformación ejecutada. En este proceso se obtuvieron las llaves foráneas correspondientes y los valores necesarios para ser mostrados en los reportes de TD.

Figura 32

Transformación tabla de hechos académico BI



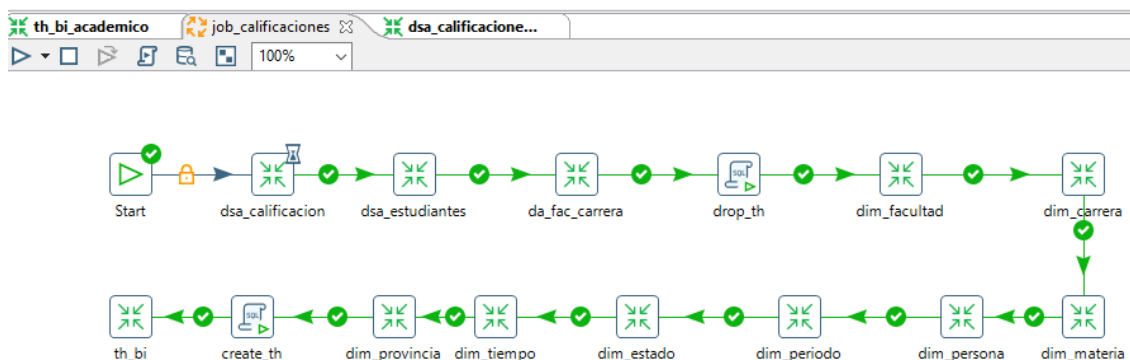
Nota. Proceso de transformación de la tabla de hechos académico bi. Elaborado por: (Loachamin, 2023).

3.3.3.12 JOB ACADÉMICO BI

Para este trabajo (Job) de Pentaho se invocó a cada una de las transformaciones anteriormente realizadas. También se ejecutó sentencias SQL para resolver el tema de los reprocesos y las cargas incrementales en la tabla de hechos.

Figura 33

Job tabla de hechos académico BI



Nota. Proceso de Job de la tabla de hechos académico bi. Elaborado por: (Loachamin, 2023).

3.3.4 FASE PROCESAMIENTO ANALÍTICO

Luego de haber concluido con proceso ELT, se realizó el proceso analítico de datos; que consiste en la construcción de los cubos OLAP, para su publicación, despliegue y finalmente la integración de visualizaciones(dashboard) con Oracle BI.

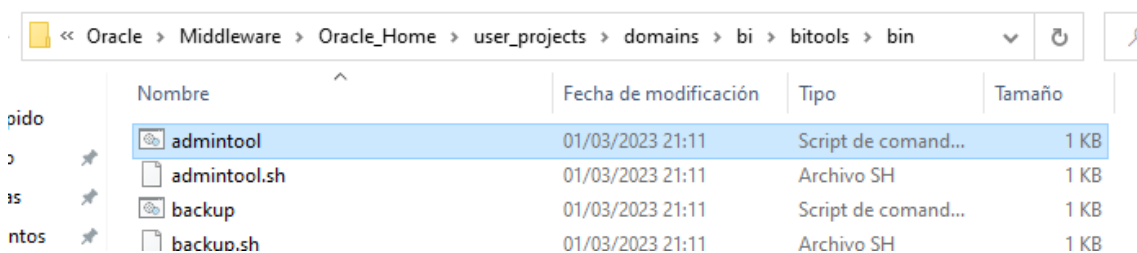
3.3.4.1 CREACIÓN DEL CUBO OLAP

A continuación, se detalla una serie de pasos que permitió la construcción del cubo OLAP de la solución de BI aplicado a los procesos académicos.

En la ruta de instalación de Oracle BI C:\Oracle\Middleware\Oracle_Home\user_projects\domains\bi\bitools\bin se ejecutó el archivo Admintool en modo administrador.

Figura 34

Ubicación de Admintool

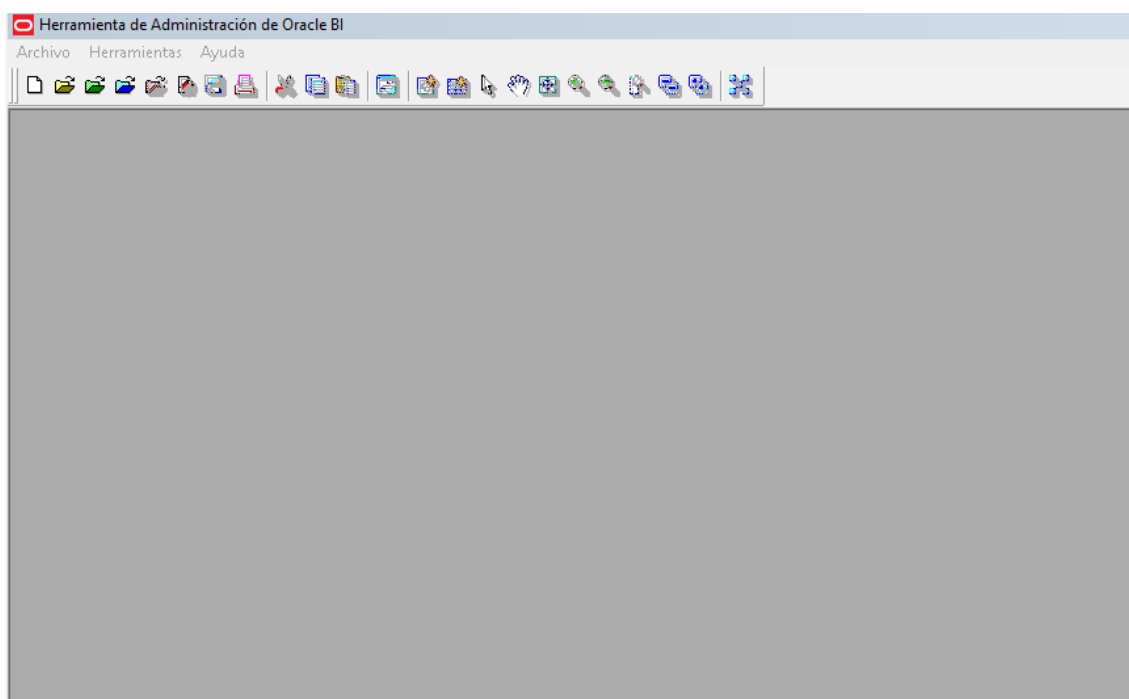


Nota. Herramienta de administración de Oracle BI. Elaborado por: (Loachamin, 2023).

En este paso se creó un nuevo archivo y se ingresó el usuario y contraseña configurados en la instalación de Oracle BI.

Figura 35

Pantalla principal



Nota. Pantalla principal de la administración de Oracle BI. Elaborado por: (Loachamin, 2023).

Para crear un nuevo repositorio se ingresó un nombre y contraseña.

Figura 36

Crear nuevo repositorio

Crear Nuevo Repositorio - Información del Repositorio

1 **Información del Repositorio**

2 Seleccionar Origen de Datos

3 Seleccionar Tipos de Metadatos

4 Seleccionar Objetos de Metadatos

5 Asignar a Modelo Lógico

6 Publicar en Almacén

Crear Repositorio: Binario Documentos XML de MDS

Nombre: Examinar...

Ubicación: Examinar...

Importar Metadatos: Sí No MDS XML

Contraseña de Repositorio:

Confirmar:

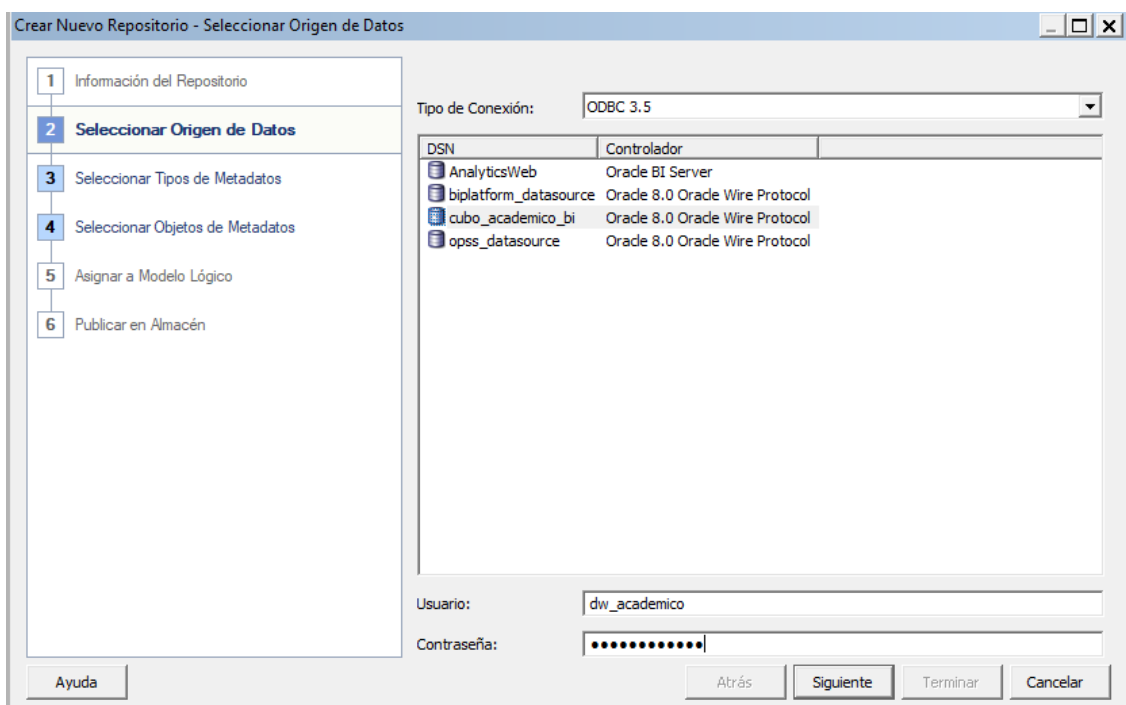
Ayuda Atrás Siguiente Terminar Cancelar

Nota. Pantalla emergente de creación de usuario y contraseña de repositorio.
Elaborado por: (Loachamin, 2023).

En este paso se seleccionó el nombre del cubo creado en la conexión ODBC, que especifica el origen de datos.

Figura 37

Origen de datos

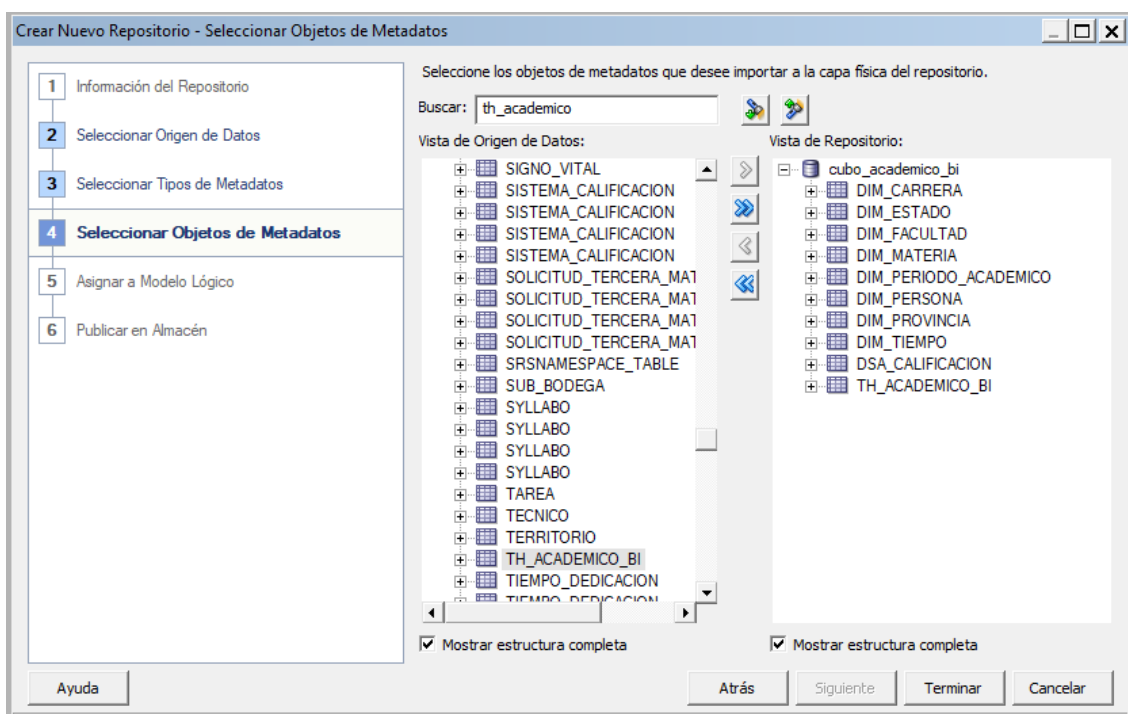


Nota. Selección del origen de datos basado en el origen ODBC. Elaborado por: (Loachamin, 2023).

En esta sección se procedió a seleccionar los metadatos. Es necesario seleccionar las tablas y vistas creadas en el proceso ETL (dimensiones y tabla de hechos).

Figura 38

Objeto de metadatos

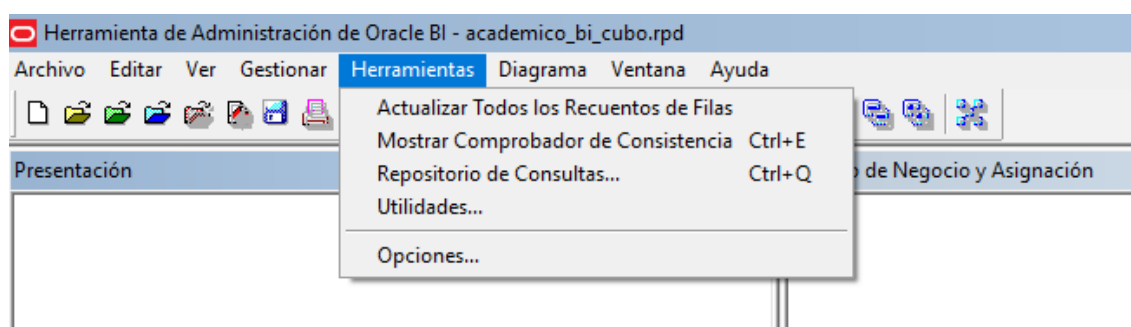


Nota. Selección de las dimensiones y tabla de hechos. Elaborado por: (Loachamin, 2023).

Una vez seleccionado las tablas, se procedió de acuerdo con la siguiente figura.

Figura 39

Actualizar todos los recuentos de filas

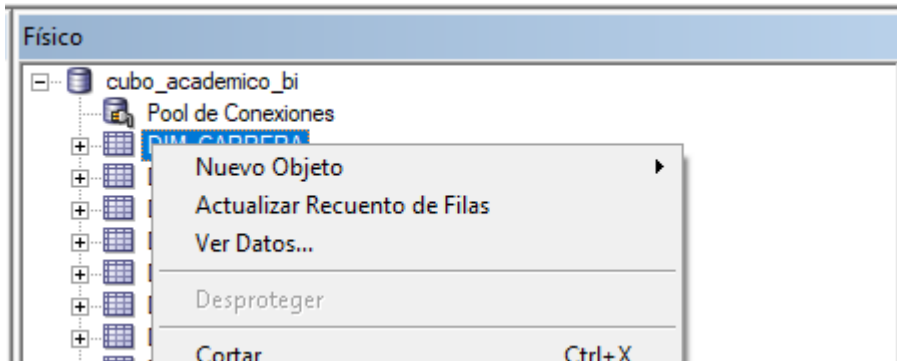


Nota. Selección de herramientas para actualizar el diagrama físico. Elaborado por: (Loachamin, 2023).

En este paso se dio clic derecho en cada una de las dimensiones extraídas y se crea un alias respetivo.

Figura 40

Modelo físico

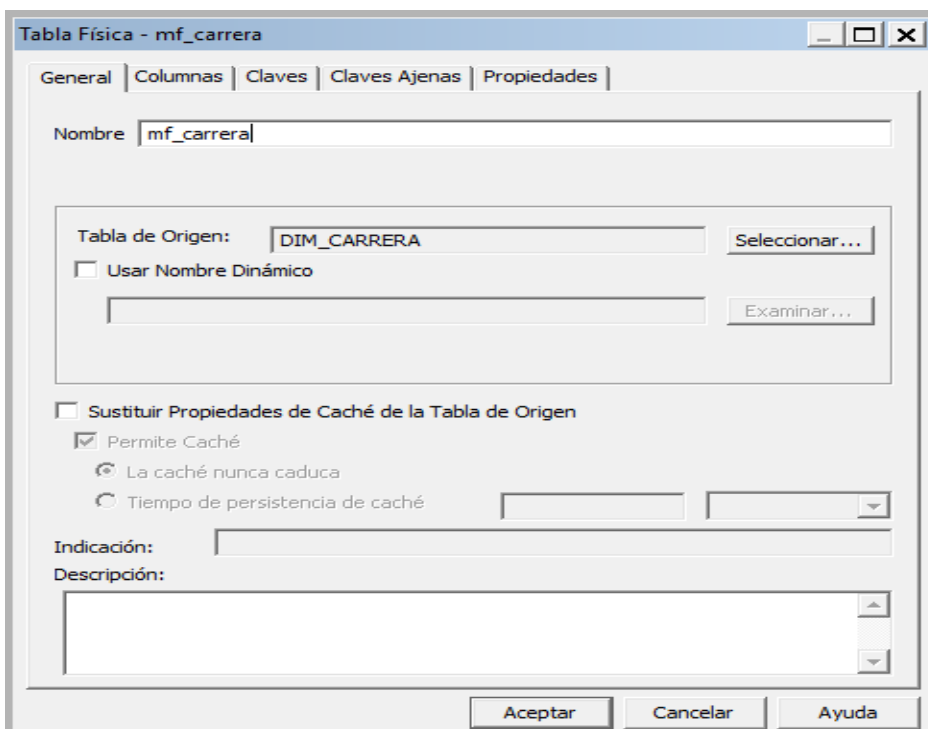


Nota. Creación del modelo físico de la solución BI. Elaborado por: (Loachamin, 2023).

Se estableció un nombre para el nuevo objeto creado. Este paso se debe repetir para cada una de las dimensiones y tabla de hechos extraídos.

Figura 41

Nombre tabla física

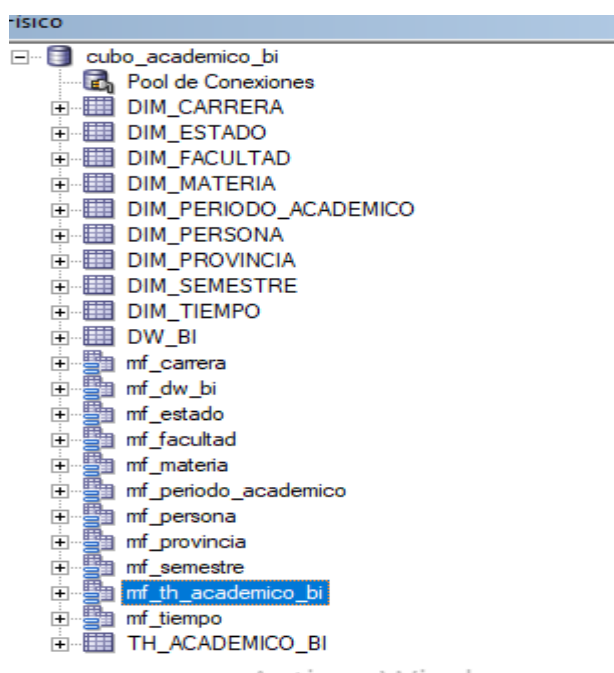


Nota. Creación y asignación de la tabla del modelo físico. Elaborado por: (Loachamin, 2023).

Una vez creado el alias de cada uno de los objetos, se seleccionó y se copió en la sección de modelo de negocio y asignación.

Figura 42

Modelo físico

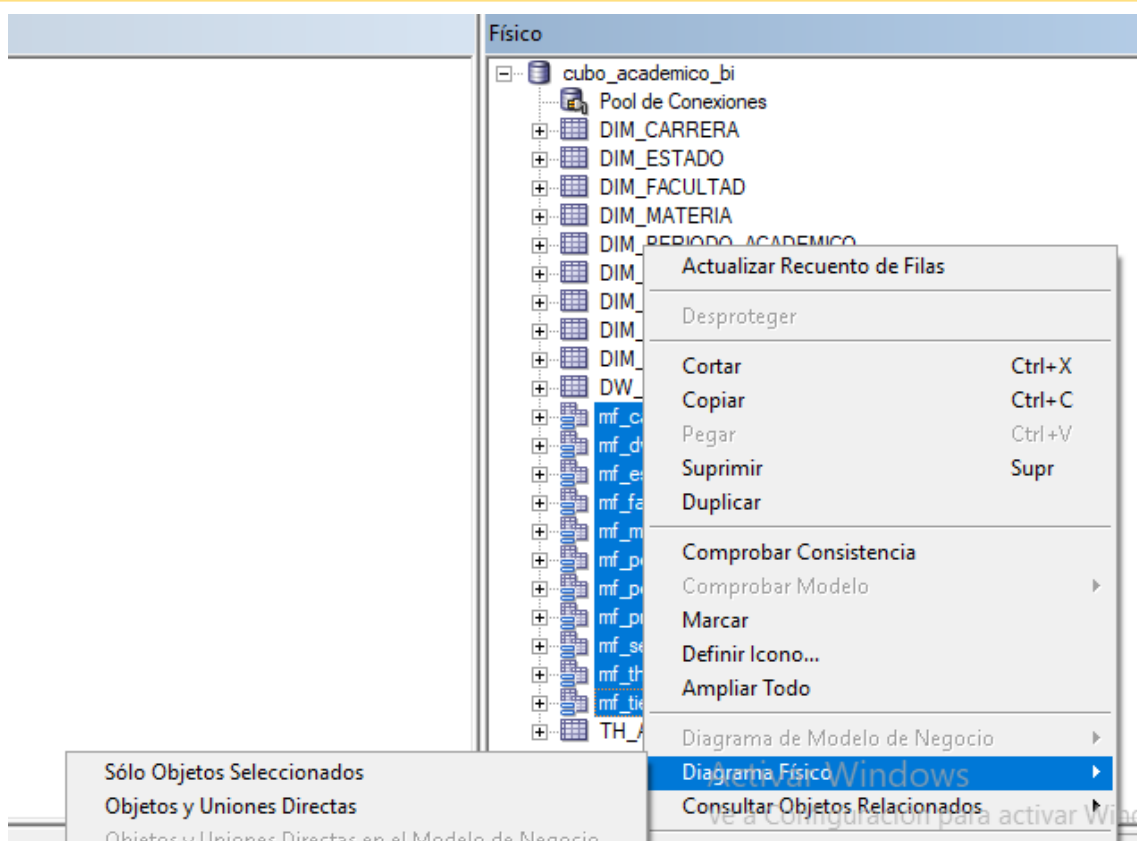


Nota. Modelo físico con todas sus dimensiones. Elaborado por: (Loachamin, 2023).

Posteriormente, se procedió a crear el diagrama físico, para lo cual se dio clic derecho en la opción Diagrama físico -> solo objetos seleccionados.

Figura 43

Diagrama físico

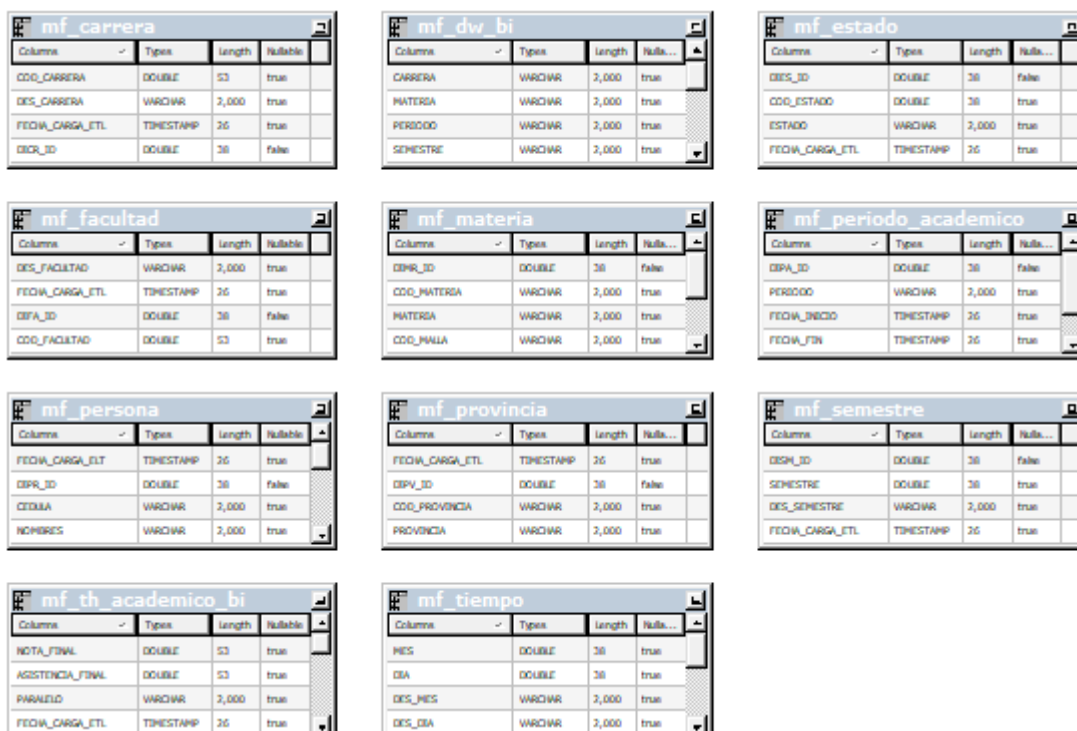


Nota. Se crea el diagrama físico de la solución BI. Elaborado por: (Loachamin, 2023).

Una vez cargado las tablas del diagrama se procedió a realizar las conexiones correspondientes. Para ello se seleccionó el icono de relación de la barra superior del menú.

Figura 44

Diagrama físico cargado



| Columna | Tipos | Length | Nullable |
|-----------------|-----------|--------|----------|
| COD_CARRERA | DOUBLE | 53 | true |
| DES_CARRERA | VARCHAR | 2,000 | true |
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |
| DECR_ID | DOUBLE | 38 | false |

| Columna | Tipos | Length | Nulla... |
|----------|---------|--------|----------|
| CARRERA | VARCHAR | 2,000 | true |
| MATERIA | VARCHAR | 2,000 | true |
| PERIODO | VARCHAR | 2,000 | true |
| SEMESTRE | VARCHAR | 2,000 | true |

| Columna | Tipos | Length | Nulla... |
|-----------------|-----------|--------|----------|
| DES_ID | DOUBLE | 38 | false |
| COD_ESTADO | DOUBLE | 38 | true |
| ESTADO | VARCHAR | 2,000 | true |
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |

| Columna | Tipos | Length | Nullable |
|-----------------|-----------|--------|----------|
| DES_FACULTAD | VARCHAR | 2,000 | true |
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |
| DIFA_ID | DOUBLE | 38 | false |
| COD_FACULTAD | DOUBLE | 53 | true |

| Columna | Tipos | Length | Nulla... |
|-------------|---------|--------|----------|
| DIRR_ID | DOUBLE | 38 | false |
| COD_MATERIA | VARCHAR | 2,000 | true |
| MATERIA | VARCHAR | 2,000 | true |
| COD_MALLA | VARCHAR | 2,000 | true |

| Columna | Tipos | Length | Nulla... |
|--------------|-----------|--------|----------|
| DIPA_ID | DOUBLE | 38 | false |
| PERIODO | VARCHAR | 2,000 | true |
| FECHA_INICIO | TIMESTAMP | 26 | true |
| FECHA_FIN | TIMESTAMP | 26 | true |

| Columna | Tipos | Length | Nullable |
|-----------------|-----------|--------|----------|
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |
| DIRR_ID | DOUBLE | 38 | false |
| CODUA | VARCHAR | 2,000 | true |
| NOMBRES | VARCHAR | 2,000 | true |

| Columna | Tipos | Length | Nulla... |
|-----------------|-----------|--------|----------|
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |
| DIRV_ID | DOUBLE | 38 | false |
| COD_PROVINCIA | VARCHAR | 2,000 | true |
| PROVINCIA | VARCHAR | 2,000 | true |

| Columna | Tipos | Length | Nulla... |
|-----------------|-----------|--------|----------|
| DIRR_ID | DOUBLE | 38 | false |
| SEMESTRE | DOUBLE | 38 | true |
| DES_SEMESTRE | VARCHAR | 2,000 | true |
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |

| Columna | Tipos | Length | Nullable |
|------------------|-----------|--------|----------|
| NOTA_FINAL | DOUBLE | 53 | true |
| ASISTENCIA_FINAL | DOUBLE | 53 | true |
| PARALELO | VARCHAR | 2,000 | true |
| FECHA_CARGA_ETL | TIMESTAMP | 26 | true |

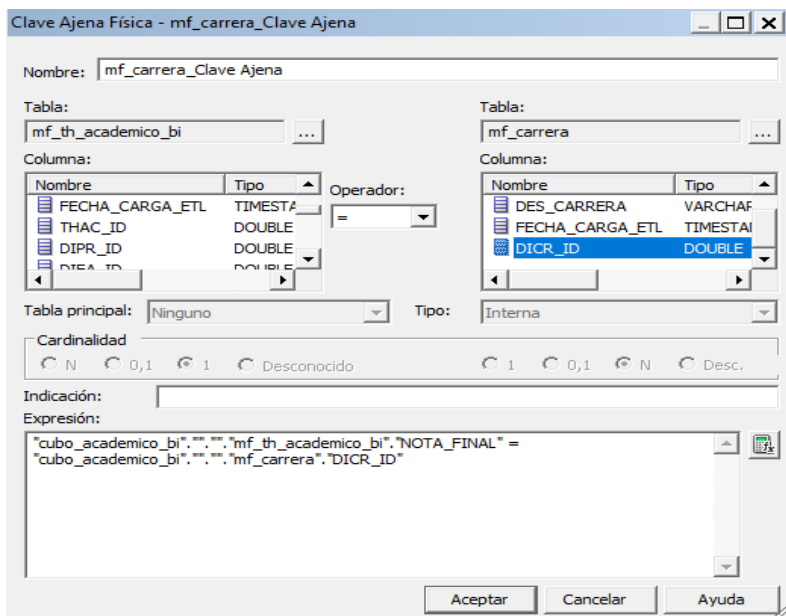
| Columna | Tipos | Length | Nulla... |
|----------|---------|--------|----------|
| MES | DOUBLE | 38 | true |
| DIAS | DOUBLE | 38 | true |
| DES_MES | VARCHAR | 2,000 | true |
| DES_DIAS | VARCHAR | 2,000 | true |

Nota. Diagrama físico cargado sin relaciones. Elaborado por: (Loachamin, 2023).

Una vez seleccionado el icono, se relacionó la tabla de origen con la tabla destino. Luego en la venta emergente se vinculó el campo de la tabla origen con el campo de la tabla destino.

Figura 45

Relaciones entre tablas

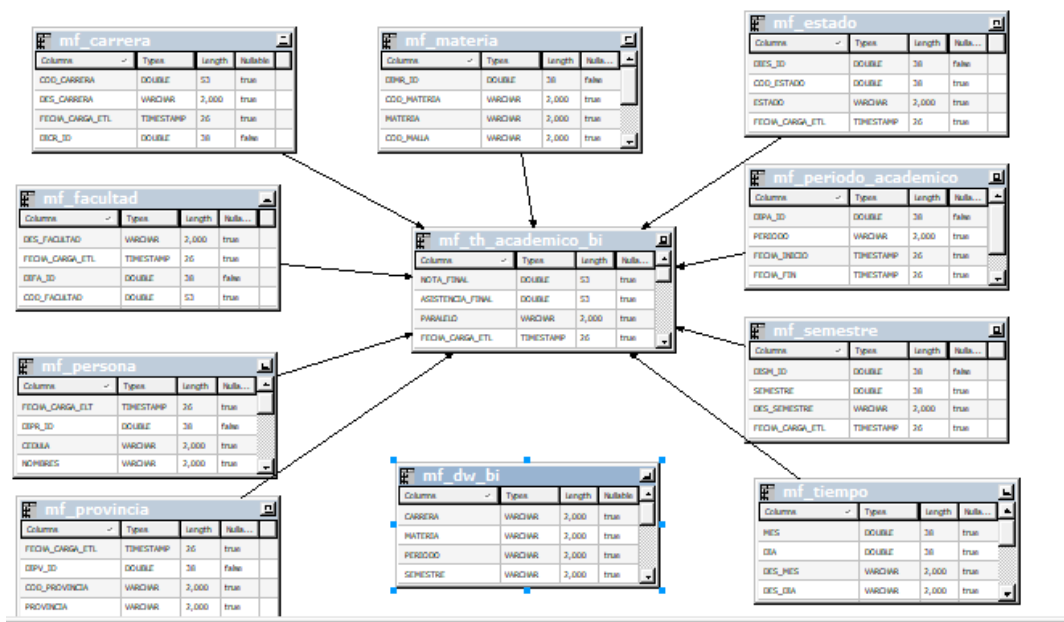


Nota. Relaciones entre la tabla origen y destino. Elaborado por: (Loachamin, 2023).

Una vez que se completó las relaciones del diagrama se visualizó de esta forma.

Figura 46

Diagrama relación entre tablas

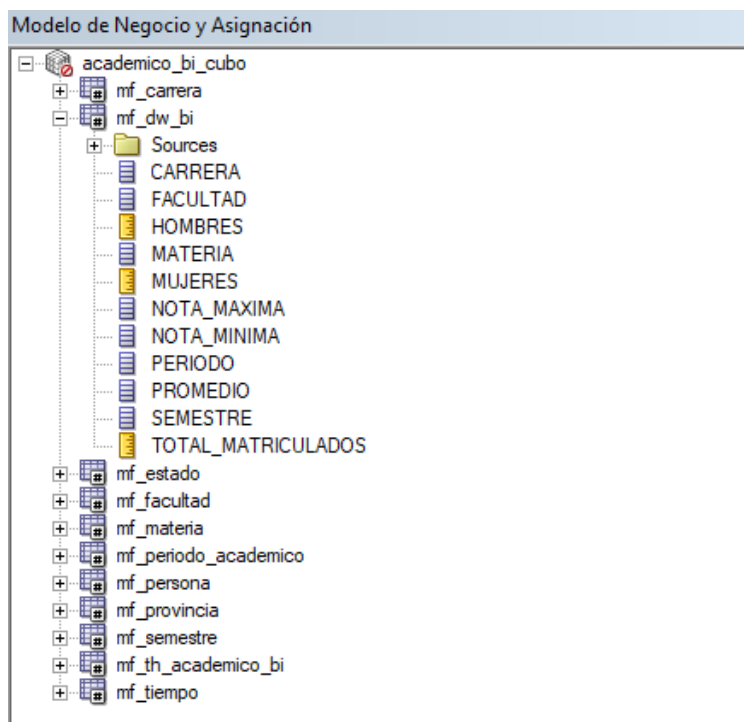


Nota. Relaciones de asignación clave ajena o foránea. Elaborado por: (Loachamin, 2023).

En esta sección, una vez cargados los objetos del diagrama físico, se seleccionó la opción de Modelo de negocio -> Diagrama completo.

Figura 47

Creación del modelo lógico

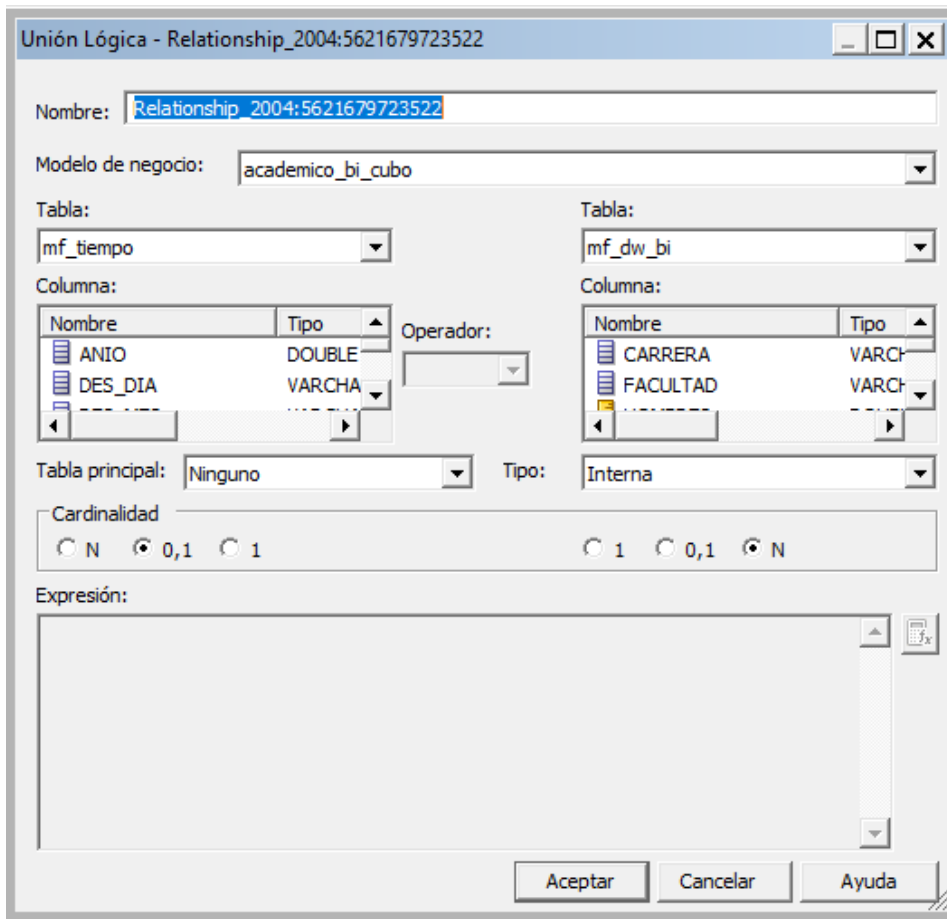


Nota. Creación del modelo lógico a partir del modelo físico. Elaborado por: (Loachamin, 2023).

Similar al diagrama físico, se creó las relaciones entre las tablas, sin embargo, en este caso no se seleccionó el campo de destino en la tabla.

Figura 48

Unión lógica

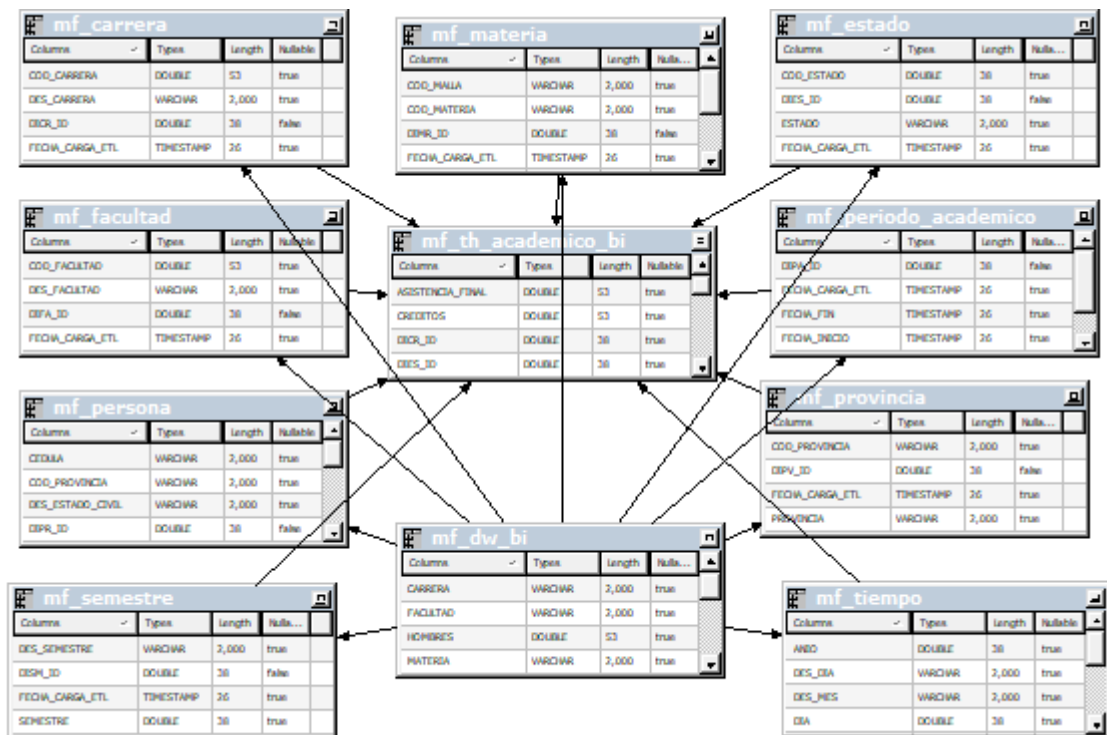


Nota. Asignación de la tabla origen hacia tabla destino. Elaborado por: (Loachamin, 2023).

Luego de haber terminado las relaciones, se logró visualizar algo así.

Figura 49

Diagrama relación lógica



Nota. Creación del diagrama de relación del modelo lógico. Elaborado por: (Loachamin, 2023).

Finalmente, se creó el área temática en la capa de presentación.

Figura 50

Área temática

Área Temática - academico_bi

General | Tablas de Presentación | Alias

Nombre: academico_bi Ppermisos...

Clave de Conversión:

Nombre Mostrado Personalizado VALUEOF(NQ_SESSION.CN_academic...

Modelo de negocio: academico_bi_cubo

Exportar claves lógicas

Columna de Hechos Implícitos

no asignado Definir... Borrar

Descripción Personalizada VALUEOF(NQ_SESSION.CD_Area_Tem...

Ocultar objeto si

Descripción:

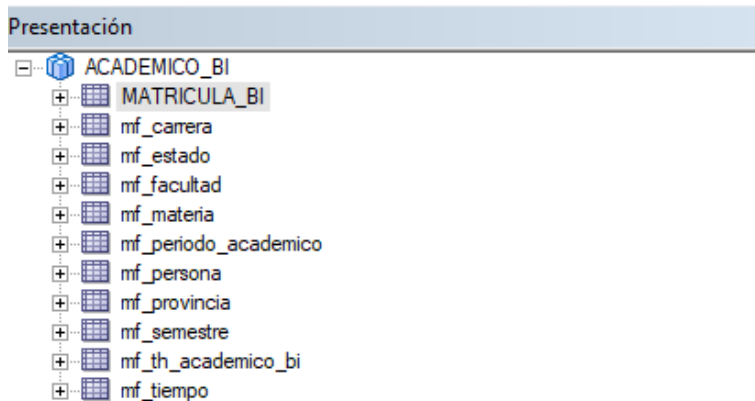
Aceptar Cancelar Ayuda

Nota. Nombre del área temática. Elaborado por: (Loachamin, 2023).

De forma similar se arrastró los objetos del modelo lógico al modelo de presentación.

Figura 51

Modelo de presentación

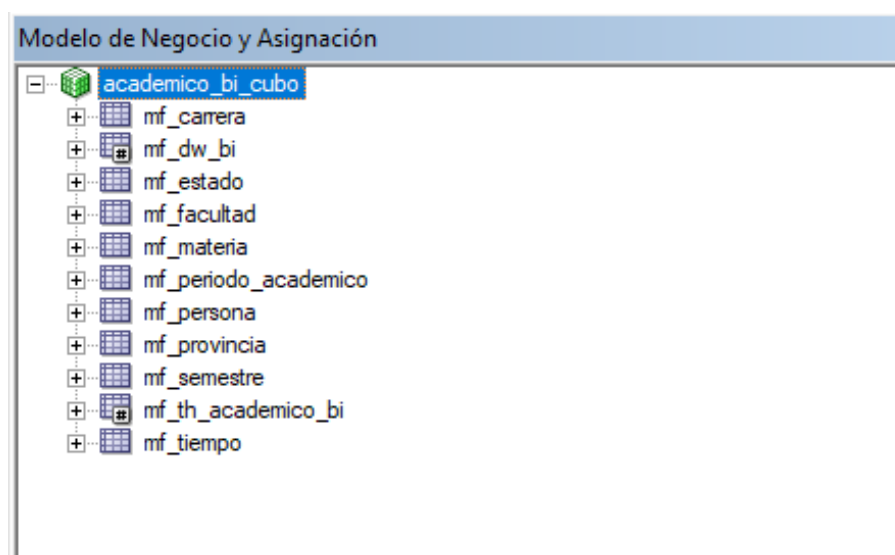


Nota. Creación del modelo de presentación. Elaborado por: (Loachamin, 2023).

El paso final fue guardar todos los cambios hechos y comprobar la consistencia global, de ser necesario se corregirá los errores encontrados, caso contrario no se podrá visualizar las tablas en el Analytics. En este caso, como todo quedó funcional, el color del cubo OLAP pasó a color verde.

Figura 52

Consistencia global



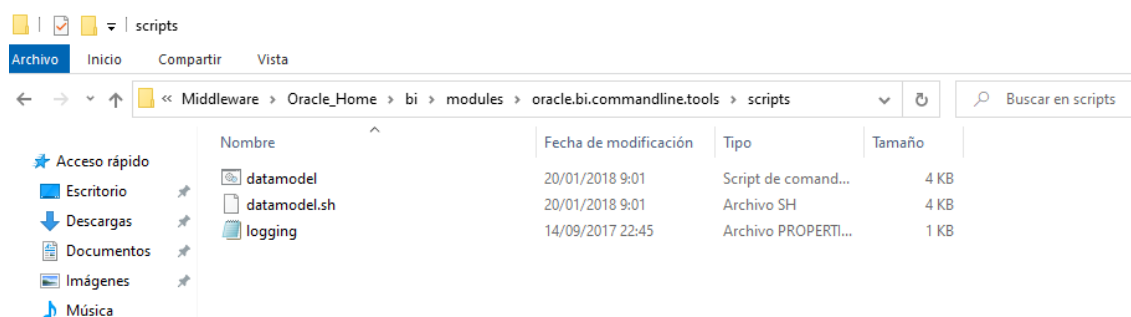
Nota. Validar la consistencia global. Elaborado por: (Loachamin, 2023).

3.3.4.2 PUBLICACIÓN DE REPOSITORIO EN ORACLE BI

Para publicar el repositorio (Cubo OLAP), se buscó la ruta siguiente:
C:\Oracle\Middleware\Oracle_Home\bi\modules\oracle.bi.commandline.tools\scripts.

Figura 53

Ruta del archivo



Nota. Navegación por el archivo datamodel. Elaborado por: (Loachamin, 2023).

Luego, se ejecutó el cmd como administrador y se cambió la ruta del directorio conforme la ruta encontrada para finalmente ingresar el siguiente comando:
datamodel.cmd upload -I [se establece la ruta/nombre del archivo.rpd] -W contraseña -U weblogic -P contraseña del weblogic -SI ssi -S localhost -N 9502.

Figura 54

Archivo datamodel

```
Administrador: Símbolo del sistema
Microsoft Windows [Versión 10.0.19045.2728]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Windows\system32>cd C:\Oracle\Middleware\Oracle_Home\bi\modules\oracle.bi.commandline.tools\scripts

C:\Oracle\Middleware\Oracle_Home\bi\modules\oracle.bi.commandline.tools\scripts>datamodel.cmd uploadrpd -I C:\Oracle\Middleware\Oracle_Home\bi\bifoundation\server\academico_bi_cubo -W Devi2010 -U weblogic -P Devi2010 -SI ssi -S localhost -N 9502
C:\Oracle\Middleware\Oracle_Home\bi\bifoundation\server\academico_bi_cubo no existe o no es un archivo.
system exit code: 1

C:\Oracle\Middleware\Oracle_Home\bi\modules\oracle.bi.commandline.tools\scripts>datamodel.cmd uploadrpd -I C:\Oracle\Middleware\Oracle_Home\bi\bifoundation\server\academico_bi_cubo.rpd -W Devi2010 -U weblogic -P Devi2010 -SI ssi -S localhost -N 9502
Instancia de servicio: ssi

Operación correcta.
La carga de RPD se ha completado correctamente.
system exit code: 0

C:\Oracle\Middleware\Oracle_Home\bi\modules\oracle.bi.commandline.tools\scripts>
```

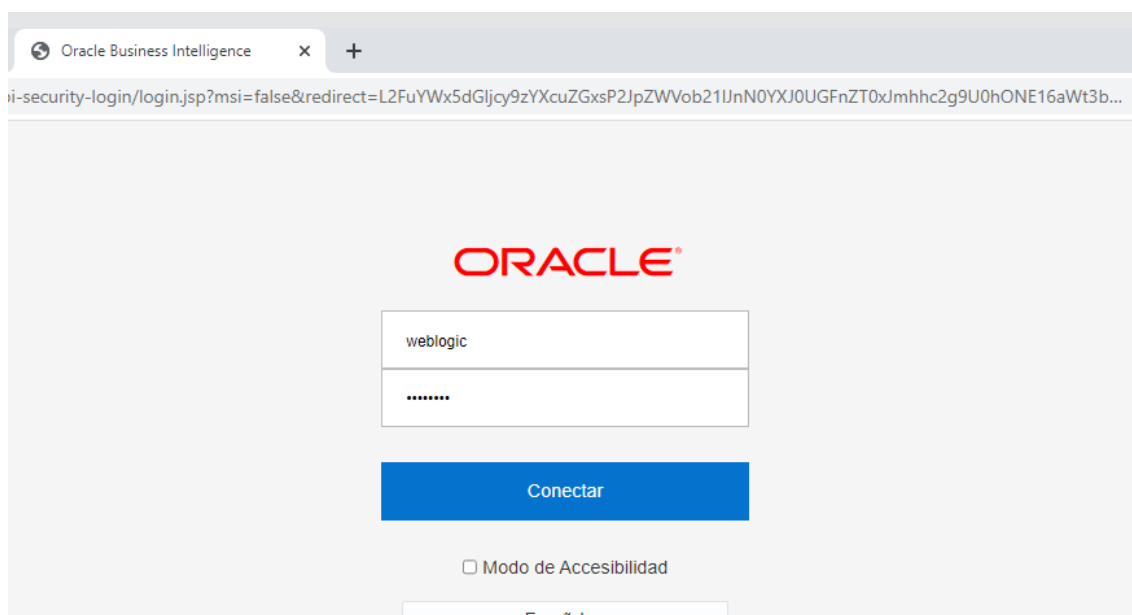
Nota. Ejecución del archivo datamodel para la publicación del cubo OLAP.
Elaborado por: (Loachamin, 2023).

3.3.4.3 CREACIÓN DE DASHBOARD

Como primer paso se ingresó en la siguiente URL <http://localhost:9502/analytics>. En esta página se procedió a realizar el login respectivo. En este punto se ingresó como usuario: weblogic y contraseña: la fijada en la instalación de la herramienta Oracle BI.

Figura 55

Inicio de sesión



Nota. Ingresar credenciales para el inicio de sesión. Elaborado por: (Loachamin, 2023).

Una vez iniciada sesión se mostró el home de la Oracle Analytics.

Figura 56

Home de Analytics

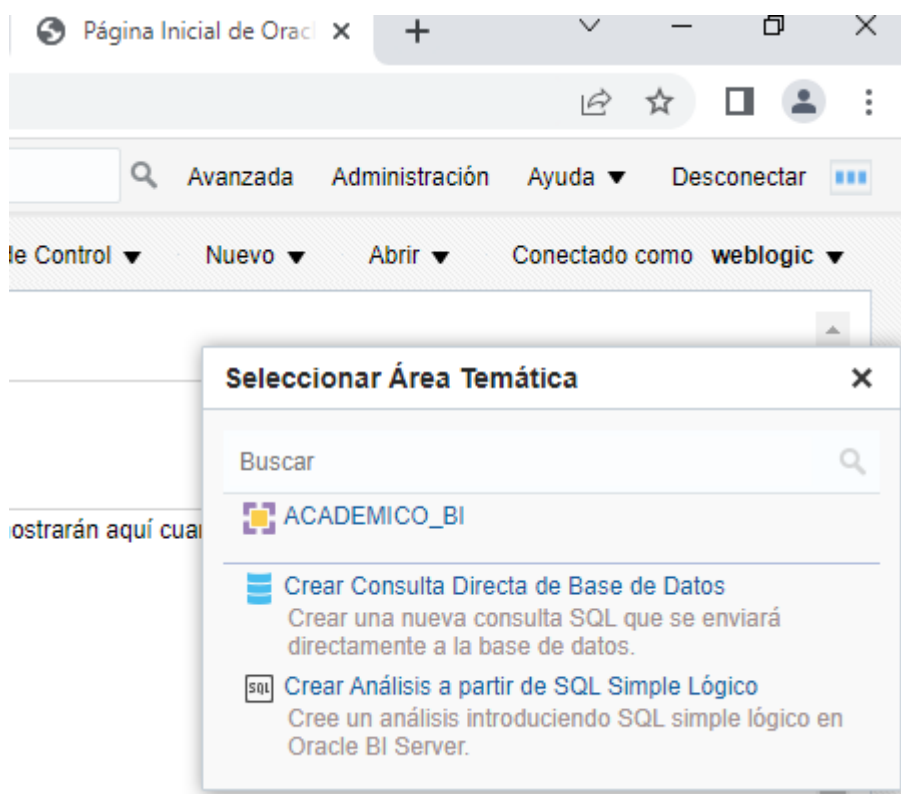


Nota. Ingreso a la pantalla principal de Analytics. Elaborado por: (Loachamin, 2023).

En parte superior izquierda del home se seleccionó nuevo-> Análisis y se encontró el área temática creada llamada "ACADEMICO_BI".

Figura 57

Seleccione área temática

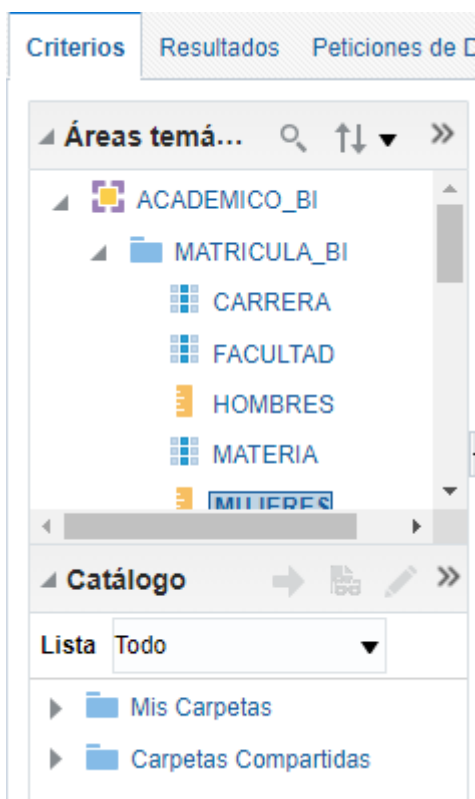


Nota. Ingreso al área temática creada en el cubo OLAP. Elaborado por: (Loachamin, 2023).

Figura 58

Panel de criterios

En este paso se desplegó la página para crear y diseñar el dashboard de la solución BI. En el panel lateral derecho se pudo visualizar el cubo OLAP juntos con sus dimensiones seleccionadas para el proceso.



Nota. Pantalla principal de inicio de análisis. Elaborado por: (Loachamin, 2023).

Figura 59

Columnas seleccionadas

En esta sección se arrastró los campos desde panel hacia al área de trabajo, de acuerdo con los requerimientos del negocio.

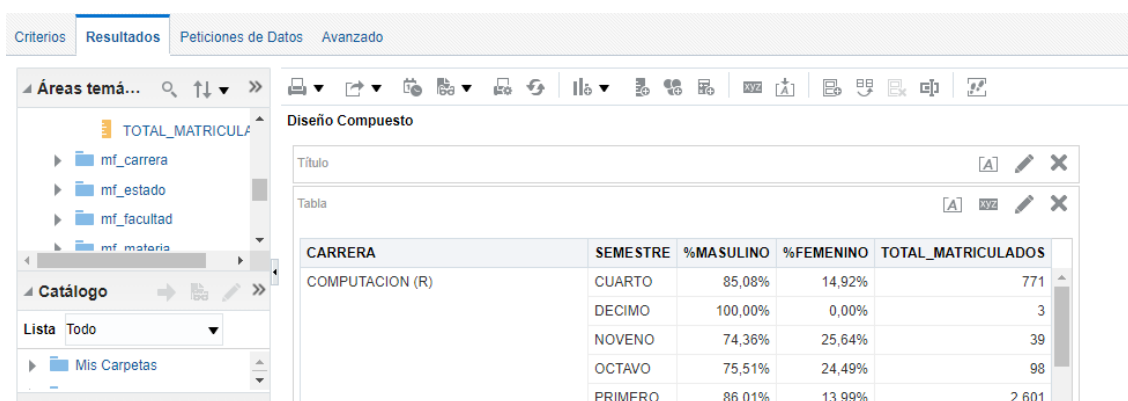


Nota. Campos para la creación del análisis de datos. Elaborado por: (Loachamin, 2023).

Una vez registrado los campos se visualizó la tabla creada dando clic en el menú de Resultados.

Figura 60

Tabla de resultados



| CARRERA | SEMESTRE | %MASULINO | %FEMENINO | TOTAL_MATRICULADOS |
|-----------------|----------|-----------|-----------|--------------------|
| COMPUTACION (R) | CUARTO | 85,08% | 14,92% | 771 |
| | DECIMO | 100,00% | 0,00% | 3 |
| | NOVENO | 74,36% | 25,64% | 39 |
| | OCTAVO | 75,51% | 24,49% | 98 |
| | PRIMERO | 86,01% | 13,99% | 2.601 |

Nota. Ejecución de la tabla de resultados. Elaborado por: (Loachamin, 2023).

Para realizar cualquier tipo de gráfica fue necesario crear una tabla dinámica con los resultados obtenidos.

Figura 61

Tabla dinámica

Diseño Compuesto

Tabla Dinámica (2)

PERIODO 2022-2023 CARRERA COMPUTACION (R)

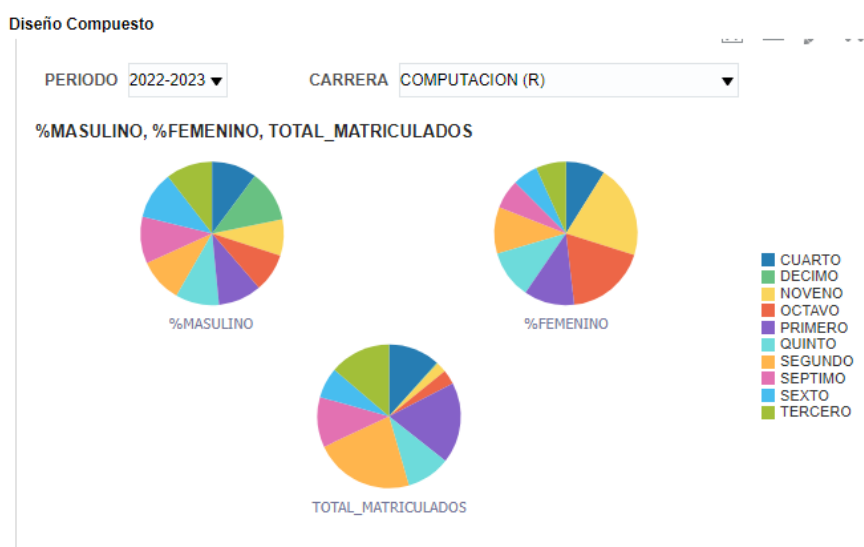
| CARRERA | SEMESTRE | %MASULINO | %FEMENINO | TOTAL_MATRICULADOS |
|-----------------|----------|-----------|-----------|--------------------|
| COMPUTACION (R) | CUARTO | 87,57% | 12,43% | 169 |
| | DECIMO | 100,00% | 0,00% | 3 |
| | NOVENO | 70,59% | 29,41% | 34 |
| | OCTAVO | 74,47% | 25,53% | 47 |
| | PRIMERO | 84,09% | 15,91% | 264 |
| | QUINTO | 84,72% | 15,28% | 144 |
| | SEGUNDO | 85,32% | 14,68% | 327 |
| | SEPTIMO | 90,85% | 9,15% | 164 |
| | SEXTO | 92,00% | 8,00% | 100 |
| | TERCERO | 90,50% | 9,50% | 200 |

Nota. Creación de la tabla dinámica. Elaborado por: (Loachamin, 2023).

Luego en el icono de gráfica se seleccionó el Gráfico-> Tarta y se obtuvo lo siguiente. Y finalmente guardar el análisis realizado.

Figura 62

Gráfico de tarta

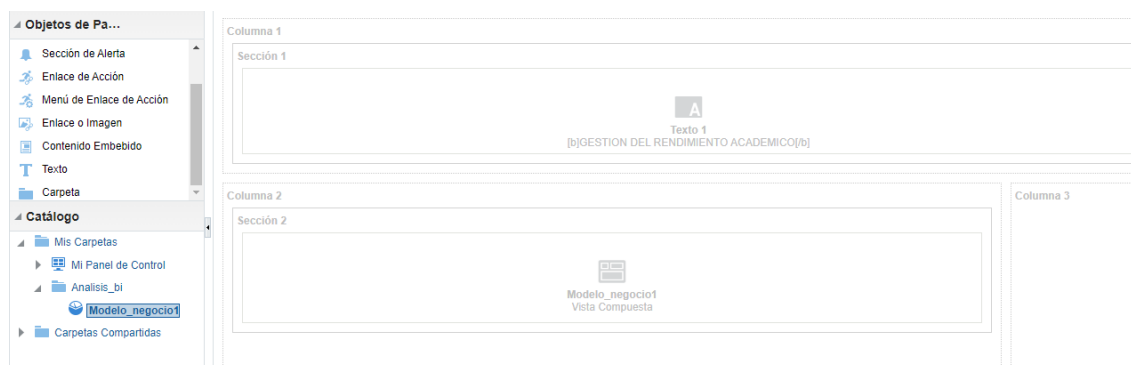


Nota. Creación del gráfico de tarta. Elaborado por: (Loachamin, 2023).

Una vez concluido el análisis, lo que se hizo fue crear el dash panel de control para agregar de manera conjunta todos los análisis realizados de acuerdo con los requisitos del negocio.

Figura 63

Panel de control



Nota. Creación del panel de control. Elaborado por: (Loachamin, 2023).

4. RESULTADOS Y DISCUSIÓN

4.1 ANÁLISIS DE RESULTADOS

Los resultados del análisis de la herramienta BI seleccionada se describen en esta sección, la validación de su efectividad a través de artículos con resultados similares al usar Oracle BI, y un ejemplo práctico de uso de esta herramienta en los procesos académicos de una IES.

De acuerdo con uno de los objetivos planteados; en la prueba de concepto se tomó como ejemplo a la FICA, por lo tanto, el modelo es fácilmente adaptable a las demás facultades y más depende de la fuente de datos obtenida. El presente modelo de BI servirá para responder las interrogantes de negocios con relación al rendimiento académico del estudiante.

4.2 DASHBOARD CREADOS

A continuación, se presenta algunos de los análisis desarrollados con la herramienta de Oracle BI.

4.2.1 MATRICULADOS POR ASIGNATURA

Este análisis muestra todos los estudiantes matriculados por asignatura, junto con la muestra, el promedio general de la asignatura, así como la nota mínima y máxima obtenida de todos los estudiantes matriculados dentro de una carrera y periodo académico específico.

Figura 64

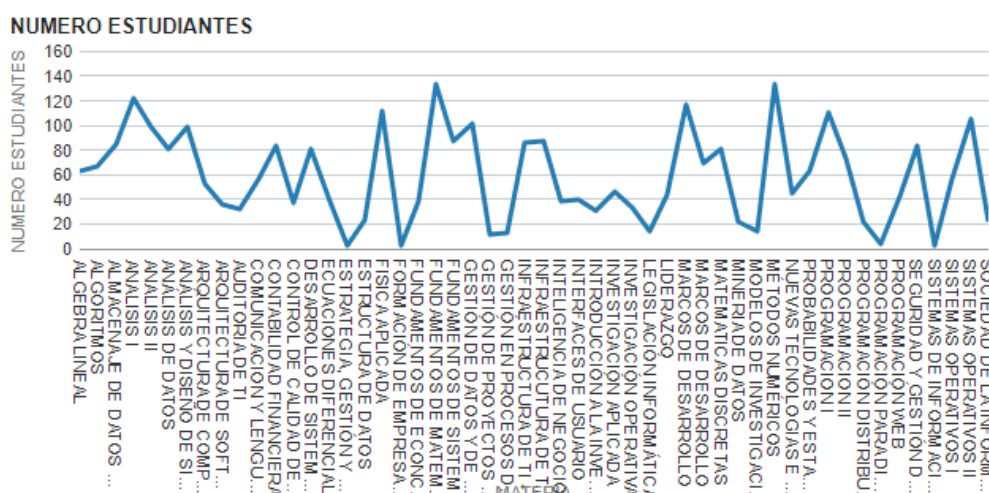
Análisis matriculados por asignatura

ESTUDIANTES MATRICULADOS POR ASIGNATURA

PERIODO CARRERA

| MATERIA | NUMERO ESTUDIANTES | PROMEDIO | MINIMA CALIFIC |
|---|--------------------|----------|----------------|
| ALGEBRA LINEAL | 63 | 4,88 | |
| ALGORITMOS | 68 | 13,69 | |
| ALMACENAJE DE DATOS Y DE LA INFORMACIÓN | 85 | 13,96 | |
| ANALISIS I | 123 | 8,94 | |
| ANALISIS II | 100 | 10,50 | |
| ANÁLISIS DE DATOS | 82 | 17,08 | |
| ANÁLISIS Y DISEÑO DE SISTEMAS | 100 | 16,10 | |
| ARQUITECTURA DE COMPUTADORES | 53 | 14,29 | |
| ARQUITECTURA DE SOFTWARE | 37 | 16,40 | |
| AUDITORIA DE TI | 33 | 19,37 | |
| COMUNICACION Y LENGUAJE | 57 | 13,37 | |
| CONTABILIDAD FINANCIERA | 84 | 14,81 | |
| CONTROL DE CALIDAD DEL SOFTWARE | 38 | 13,48 | |
| DESARROLLO DE SISTEMAS DE INFORMACIÓN | 82 | 16,13 | |

PERIODO CARRERA



Nota. Análisis de los estudiantes matriculados por asignatura con el promedio general. Elaborado por: (Loachamin, 2023).

Resultados obtenidos

Dentro del contexto de los estudiantes matriculados por asignatura se puede evidenciar claramente cuál es la asignatura con el mayor número de matriculados. Además, se puede combinar las variables para tener como resultado la asignatura

con el promedio general más alto, a su vez determinar cuál es la asignatura con la máxima y mínima calificación. Este análisis tiene como finalidad demostrar el rendimiento académico del estudiante que permite a los usuarios finales tomar decisiones inteligentes.

4.2.2 DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR GÉNERO Y SEMESTRE

Este análisis muestra el porcentaje de los estudiantes matriculados por género, junto con total de estudiantes matriculados dentro de una carrera y periodo académico específico.

Figura 65

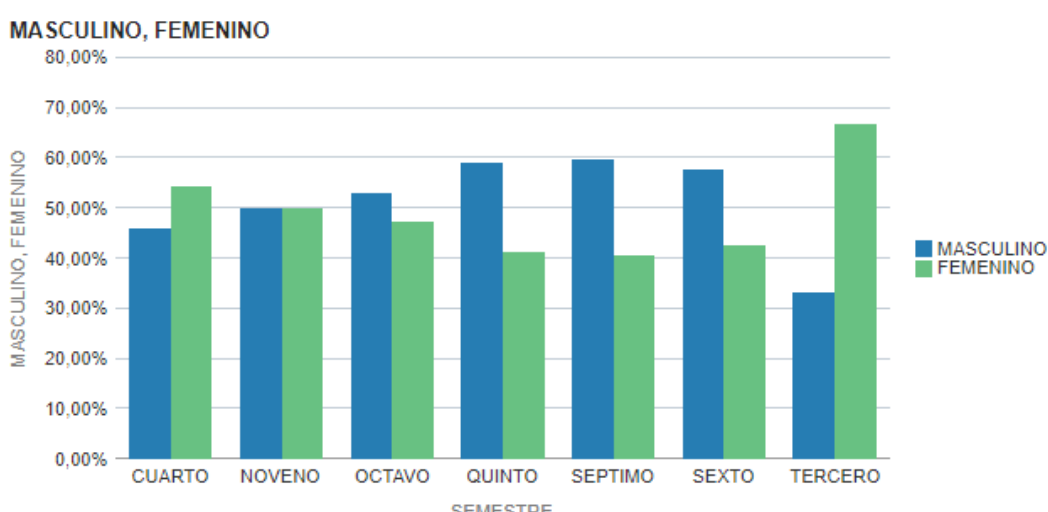
Análisis de matriculados por género

DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR GÉNERO Y SEMESTRE

PERIODO CARRERA

| SEMESTRE | TOTAL MATRICULADOS | MASCULINO | FEMENINO |
|----------|--------------------|-----------|----------|
| CUARTO | 24 | 45,83% | 54,17% |
| NOVENO | 32 | 50,00% | 50,00% |
| OCTAVO | 36 | 52,78% | 47,22% |
| QUINTO | 34 | 58,82% | 41,18% |
| SEPTIMO | 47 | 59,57% | 40,43% |
| SEXTO | 40 | 57,50% | 42,50% |
| TERCERO | 6 | 33,33% | 66,67% |

PERIODO CARRERA



Nota. Análisis de los estudiantes matriculados por semestre y género. Elaborado por: (Loachamin, 2023).

Resultados obtenidos

Con base en el número de estudiantes matriculados por género, se evidencia una baja relación de hombres versus mujeres en las carreras de la FICA, sin embargo, la carrera de Diseño Industrial cuenta con un mayor número de mujeres en la mayoría de los periodos académicos. Por lo que se pudo determinar que es una de las carreras más preferidas del género femenino en esta Facultad.

4.2.3 PROMEDIO GENERAL DE LOS ESTUDIANTES POR PERIODO ACADÉMICO

Este análisis muestra el promedio general de los estudiantes por período académico, junto con la fecha de inicio y fin de período en toda la FICA.

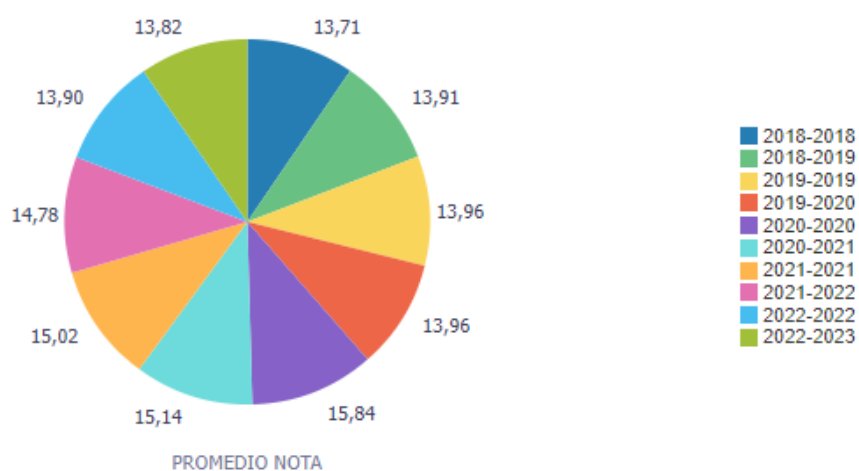
Figura 66

Análisis promedio general por período académico

PROMEDIO GENERAL POR PERIODO ACADEMICO DE TODA LA FACULTAD

| PERIODO ACADEMICO | FECHA_FIN | FECHA_INICIO | PROMEDIO NOTA |
|-------------------|--------------------|--------------------|---------------|
| 2018-2018 | 08/10/2018 0:00:00 | 01/09/2018 0:00:00 | 13,71 |
| 2018-2019 | 17/01/2019 0:00:00 | 01/06/2019 0:00:00 | 13,91 |
| 2019-2019 | 08/08/2019 0:00:00 | 02/01/2019 0:00:00 | 13,96 |
| 2019-2020 | 12/02/2020 0:00:00 | 12/08/2019 0:00:00 | 13,96 |
| 2020-2020 | 03/09/2020 0:00:00 | 08/03/2020 0:00:00 | 15,84 |
| 2020-2021 | 16/04/2021 0:00:00 | 14/10/2020 0:00:00 | 15,14 |
| 2021-2021 | 22/10/2021 0:00:00 | 30/04/2021 0:00:00 | 15,02 |
| 2021-2022 | 14/04/2022 0:00:00 | 14/10/2021 0:00:00 | 14,78 |
| 2022-2022 | 23/09/2022 0:00:00 | 19/04/2022 0:00:00 | 13,90 |
| 2022-2023 | 25/03/2023 0:00:00 | 27/09/2022 0:00:00 | 13,82 |

PROMEDIO NOTA



Nota. Análisis del promedio general de los estudiantes matriculados por periodo académico de toda la facultad. Elaborado por: (Loachamin, 2023).

Resultados obtenidos

Este análisis permite conocer la situación académica de manera general de los estudiantes con relación al rendimiento académico por período académico. La pandemia es sin duda algo terrible de recordar, sin embargo, para el análisis se pudo evidenciar que el tiempo que transcurrió la pandemia (período académico 2020-2020 hasta 2021-2022) se pudo conseguir los valores más altos en cuanto el promedio general obtenidos por los estudiantes.

4.2.4 DISTRIBUCIÓN DE LOS ESTUDIANTES POR NÚMERO DE MATRÍCULA

Este análisis muestra la distribución de los estudiantes por número de matrícula vinculados a una carrera y período académico específico dentro de la FICA.

Figura 67

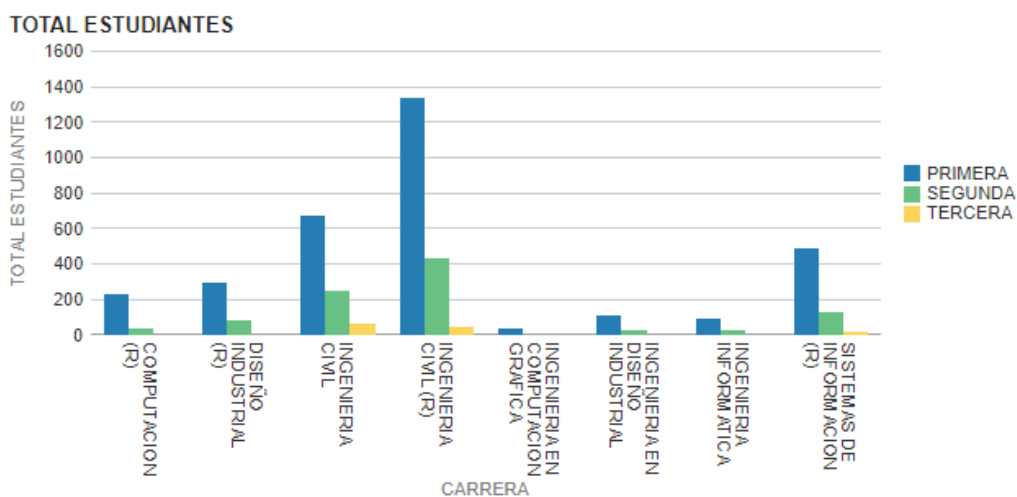
Análisis estudiantes matriculados por número de matrícula

DISTRIBUCION DE LOS ESTUDIANTES POR NUMERO DE MATRICULA

PERIODO 2021-2022 ▼

| CARRERA | PRIMERA | SEGUNDA | TERCERA |
|-----------------------------------|-------------------|-------------------|-------------------|
| | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES |
| COMPUTACION (R) | 270 | 80 | 19 |
| DISEÑO INDUSTRIAL (R) | 313 | 90 | 29 |
| INGENIERIA CIVIL | 513 | 204 | 60 |
| INGENIERIA CIVIL (R) | 1341 | 623 | 213 |
| INGENIERIA EN COMPUTACION GRAFICA | 28 | 4 | 1 |
| INGENIERIA EN DISEÑO INDUSTRIAL | 68 | 9 | 1 |
| INGENIERIA INFORMATICA | 63 | 25 | 3 |
| SISTEMAS DE INFORMACION (R) | 581 | 146 | 29 |

PERIODO 2020-2021 ▼



Nota. Distribución de los estudiantes matriculados por número de matrícula dentro de una carrera y período académico. Elaborado por: (Loachamin, 2023).

Resultados obtenidos

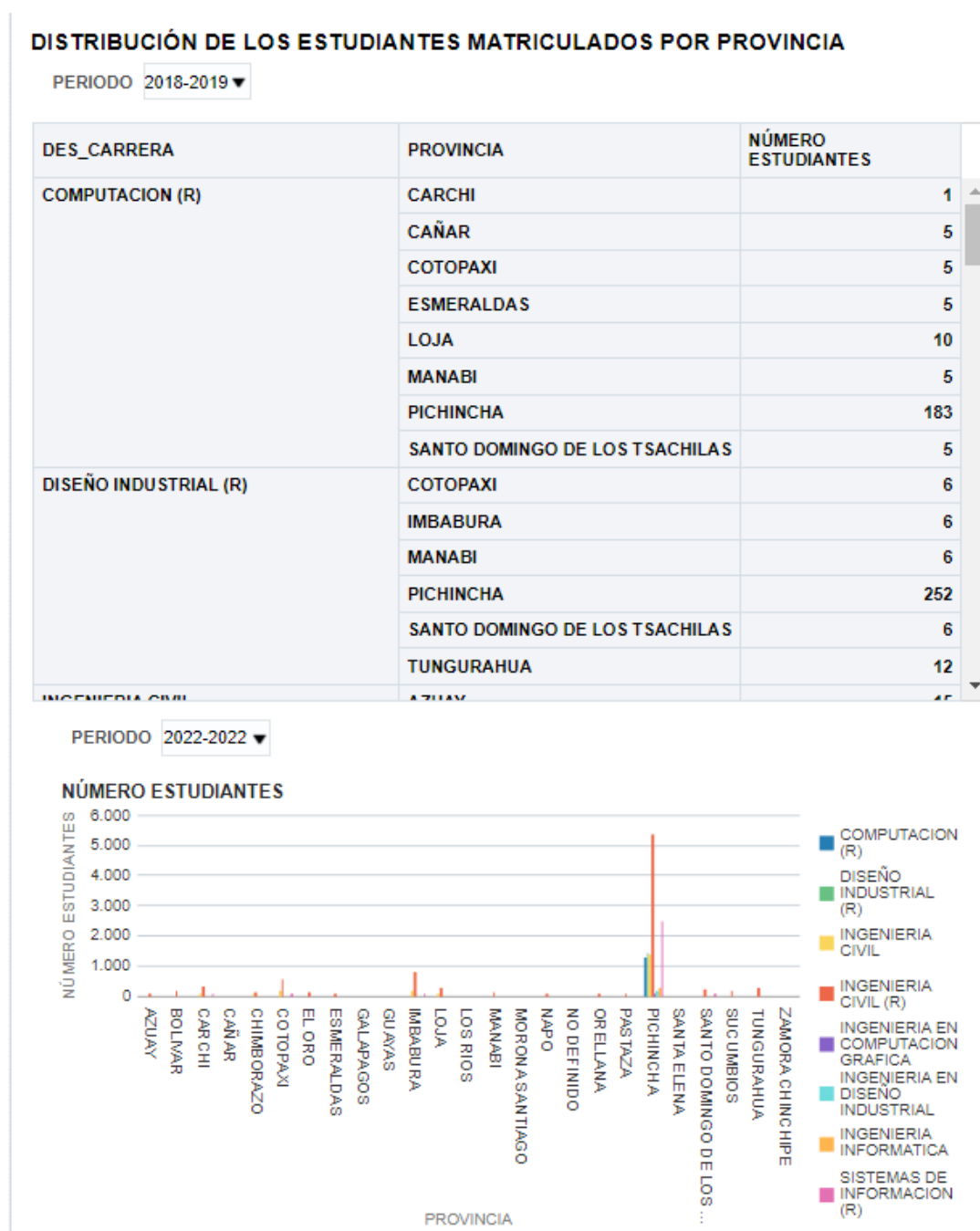
El presente análisis tiene como objetivo determinar las carreras con el mayor índice de estudiantes matriculados con segunda y tercera matrícula, a fin de tomar las mejores decisiones bajo esta problemática, así también permite tener datos históricos del comportamiento de los estudiantes que afectan al rendimiento de cada una de las unidades de formación dentro de la FICA.

4.2.5 DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR PROVINCIA.

Este análisis muestra la distribución de los estudiantes matriculados junto con el total de estudiantes y la provincia, en la carrera y período académico específico.

Figura 68

Análisis matriculados por provincia



Nota. Distribución de los estudiantes matriculados por provincia, carrera y período académico. Elaborado por: (Loachamin, 2023).

Resultados obtenidos

Si bien es cierto el mayor número de estudiantes son de la provincia de Pichincha, hay que resaltar las demás provincias como: Carchi, Cotopaxi, Imbabura, Loja, Tungurahua que también tienen un número representativo de estudiantes. Con base en el análisis realizado se pudo determinar cuáles son las provincias con el mayor número de estudiantes en las diferentes carreras que se ofrece en la FICA, además, se pudo ver qué preferían los estudiantes al momento de seleccionar una carrera.

4.2.6 DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR NACIONALIDAD.

Este análisis muestra la distribución de los estudiantes matriculados, acompañado del número total de estudiantes matriculados por nacionalidad (ecuatoriano o extranjero) en una carrera y período académico específico.

Figura 69

Análisis matriculados por nacionalidad

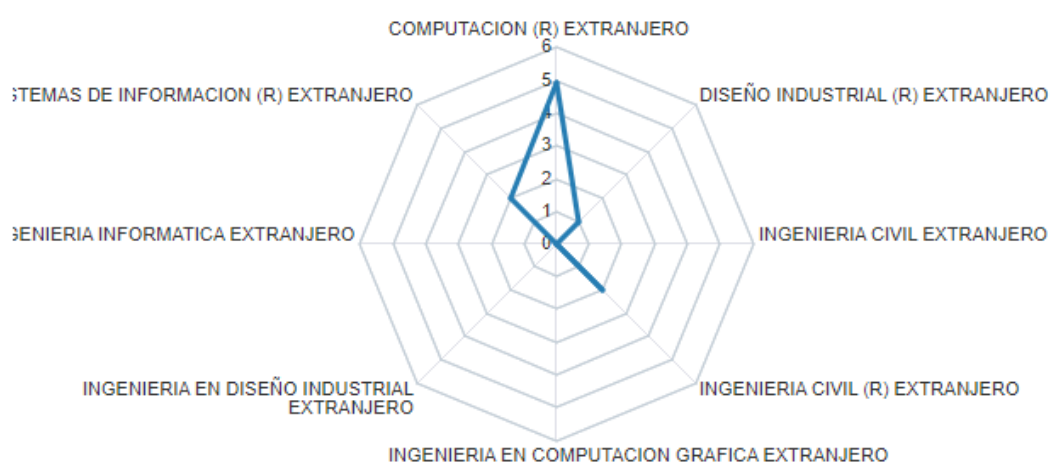
DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR NACIONALIDAD.

PERIODO 2018-2018 ▼

| CARRERA | TOTAL_MATRICULADOS | ECUATORIANO | EXTRANJERO |
|-----------------------------------|--------------------|-------------|------------|
| COMPUTACION (R) | 27 | 27 | 0 |
| INGENIERIA CIVIL | 1.370 | 1.370 | 0 |
| INGENIERIA EN COMPUTACION GRAFICA | 129 | 128 | 1 |
| INGENIERIA EN DISEÑO INDUSTRIAL | 248 | 248 | 0 |
| INGENIERIA INFORMATICA | 271 | 271 | 0 |
| SISTEMAS DE INFORMACION (R) | 75 | 75 | 0 |

PERIODO 2022-2023 ▼

EXTRANJERO



Nota. Distribución de los estudiantes matriculados por nacionalidad, carrera y período académico específico. Elaborado por: (Loachamin, 2023).

Resultados obtenidos:

Con base en los resultados del análisis, se puede determinar las profesiones que más interesan a los extranjeros y el número total de los mismos.

4.2.7 DISTRIBUCIÓN DE LOS ESTUDIANTES POR ASIGNATURA Y ESTADO.

El análisis muestra la distribución de estudiantes por asignatura, así como el número total de estudiantes por estado; aprobado, reprobado y retirado para la carrera y período académico específico.

Figura 70

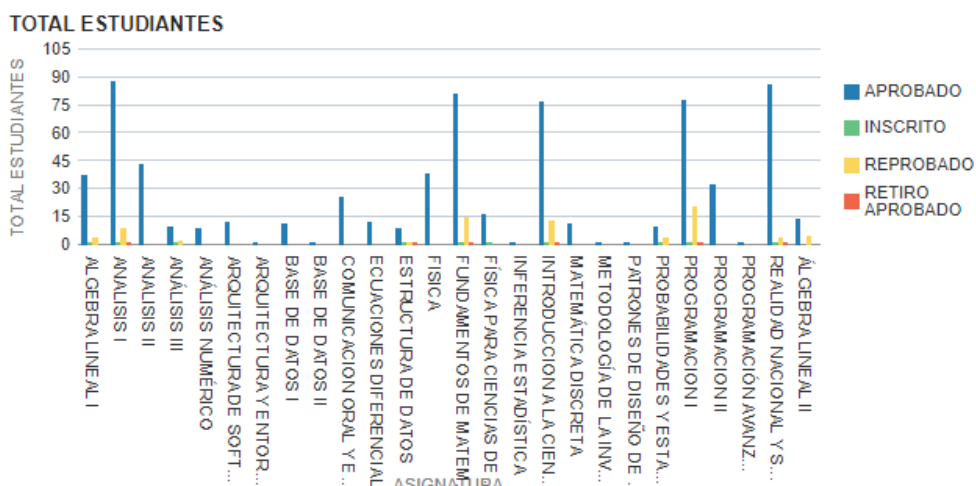
Análisis estudiantes por asignatura y estado

DISTRIBUCIÓN DE LOS ESTUDIANTES POR ASIGNATURA Y ESTADO.

PERIODO 2022-2023 DES_CARRERA SISTEMAS DE INFORMACION (R)

| ASIGNATURA | RECUPERACION | REPROBADO | RETIRO APROBADO |
|---|-------------------|-------------------|-------------------|
| | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES |
| ALGEBRA LINEAL | 5 | 51 | 3 |
| ALGORITMOS | | 1 | 5 |
| ALMACENAJE DE DATOS Y DE LA INFORMACIÓN | 1 | 15 | 1 |
| ANALISIS I | 2 | 89 | |
| ANALISIS II | 5 | 54 | 4 |
| ANÁLISIS DE DATOS | | | |
| ANÁLISIS Y DISEÑO DE SISTEMAS | | 3 | 1 |
| ARQUITECTURA DE COMPUTADORES | 2 | 10 | 7 |
| ARQUITECTURA DE SOFTWARE | | | |
| AUDITORIA DE TI | 1 | | |
| COMUNICACION Y LENGUAJE | 1 | 7 | |
| CONTABILIDAD FINANCIERA | | 10 | |

PERIODO 2020-2020 DES_CARRERA COMPUTACION (R)



Nota. Distribución de los estudiantes por asignatura y estado. Elaborado por: (Loachamin, 2023).

Resultados obtenidos:

A partir del análisis se puede determinar las asignaturas que tienen el mayor número de estudiantes reprobados, retirados u otras situaciones. Posteriormente, habrá que analizar el origen del problema con relación a la asignatura dictada, que será la clave para saber por qué y tomar la mejor decisión. Sin embargo, no todo es malo, este análisis permite tener un estadístico de la asignatura con más aprobados.

4.2.8 DISTRIBUCIÓN DE LAS ASIGNATURAS POR SU ESTADO Y NÚMERO DE MATRÍCULA.

Este análisis muestra la distribución de las asignaturas con números de matrícula: primera, segunda y tercera matrícula y estado: aprobado, reprobado y retiro aprobado en la FICA por período académico.

Figura 71

Análisis de asignaturas por su estado y número de matrícula

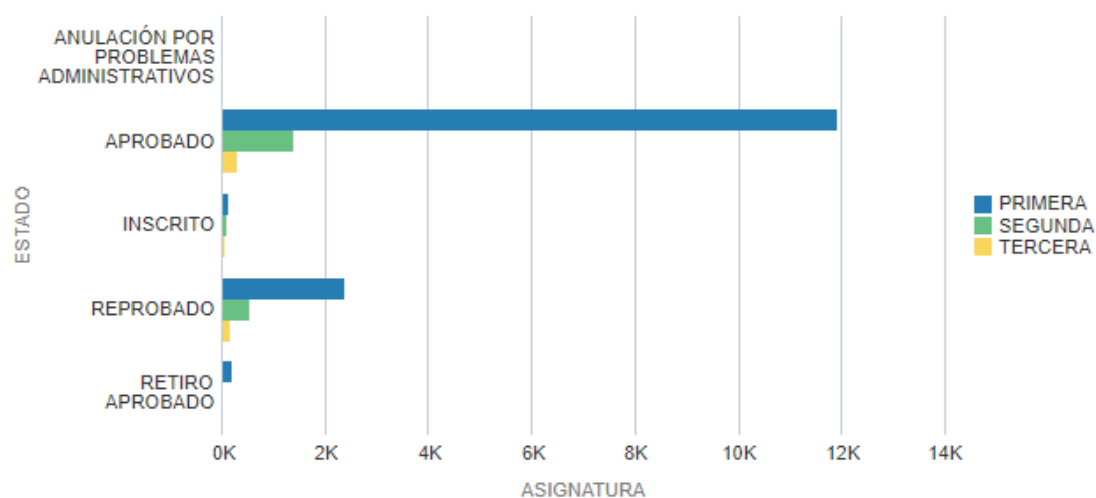
DISTRIBUCIÓN DE LAS ASIGNATURAS POR SU ESTADO.

PERIODO 2021-2022 ▼

| ESTADO | PRIMERA | SEGUNDA | TERCERA |
|---|------------|------------|------------|
| | ASIGNATURA | ASIGNATURA | ASIGNATURA |
| ANULACIÓN POR PROBLEMAS ADMINISTRATIVOS | | 6 | |
| APROBADO | 11897 | 1356 | 260 |
| INSCRITO | 99 | 59 | 43 |
| REPROBADO | 2359 | 496 | 120 |
| RETIRO APROBADO | 179 | 29 | 8 |

PERIODO 2021-2022 ▼

ASIGNATURA



Nota. Distribución de las asignaturas por número de matrícula y estado en la FICA.

Elaborado por: (Loachamin, 2023).

Resultados obtenidos:

A partir de este análisis se puede determinar el comportamiento de las asignaturas aprobadas por los estudiantes que hacen primera, segunda y hasta tercera matrícula en un período académico específico. Además, se puede medir el comportamiento en cada uno de los períodos académicos con el fin de predecir un número de estimado de estudiantes para el próximo período académico en la FICA.

4.2.9 DISTRIBUCIÓN DE LOS ESTUDIANTES POR ESTADO CIVIL.

Este análisis muestra la distribución de los estudiantes matriculados por estado civil en la carrera y período académico específico.

Figura 72

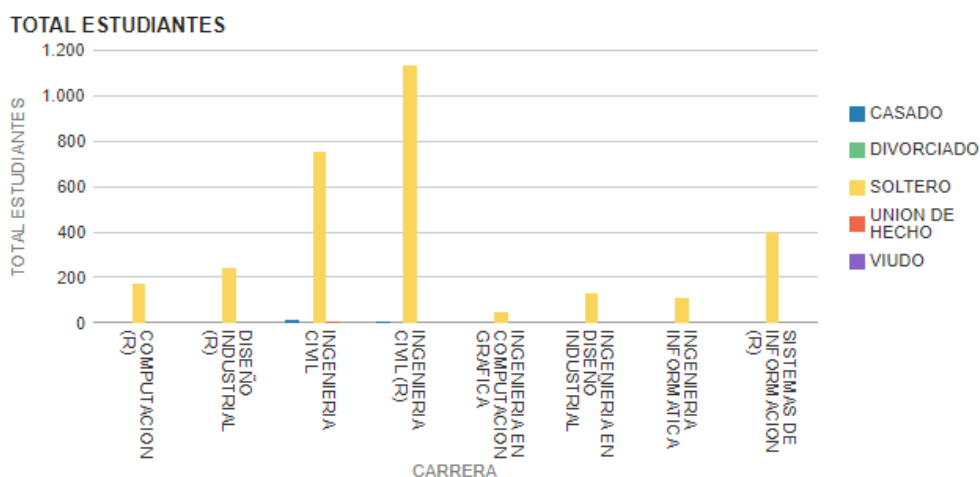
Análisis matriculados por estado civil

DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR ESTADO EN LA FACULTAD

PERIODO 2020-2020 ▼

| CARRERA | CASADO | DIVORCIADO | SEPARADO | SOLTERO |
|-----------------------------------|-------------------|-------------------|-------------------|-------------------|
| | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES | TOTAL ESTUDIANTES |
| COMPUTACION (R) | | | | 175 |
| DISEÑO INDUSTRIAL (R) | 1 | | | 244 |
| INGENIERIA CIVIL | 14 | 2 | | 758 |
| INGENIERIA CIVIL (R) | 7 | 3 | | 1138 |
| INGENIERIA EN COMPUTACION GRAFICA | 1 | | | 54 |
| INGENIERIA EN DISEÑO INDUSTRIAL | 4 | | | 133 |
| INGENIERIA INFORMATICA | 5 | | | 115 |
| SISTEMAS DE INFORMACION (R) | | | | 404 |

PERIODO 2020-2020 ▼



Nota. Distribución de los estudiantes matriculados por estado civil en la carrera y período académico específico. Elaborado por: (Loachamin, 2023).

Resultados obtenidos:

De acuerdo con el análisis y como era de esperarse, el mayor número de estudiantes son solteros, sin embargo, por alguna razón y luego de haber analizado cada uno de los períodos académicos, se pudo determinar que existe un número representativo de estudiantes casados alojados en la carrera de Ingeniería Civil, lo que hace pensar que las personas que siguen esa carrera son o están en un rango de edad superior al promedio de los estudiantes que ingresa a cursar su vida estudiantil.

4.2.10 DISTRIBUCIÓN DE LOS ESTUDIANTES MATRICULADOS POR LA SITUACIÓN SOCIOECONÓMICA.

Este análisis muestra la distribución de los estudiantes matriculados por la situación socioeconómica del estudiante, con el número de total de matriculados por período académico en la FICA.

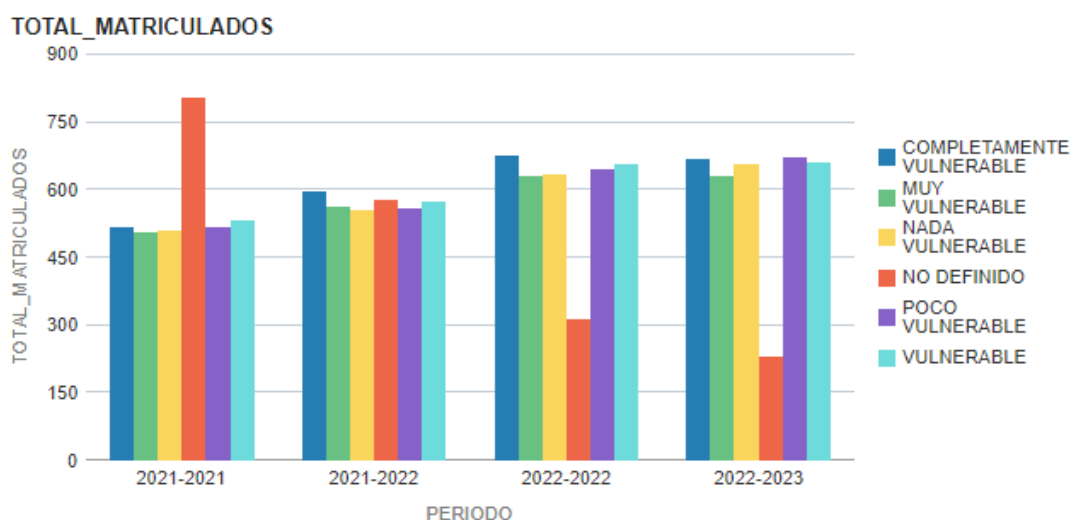
Figura 73

Análisis matriculados por la situación socioeconómica

DISTRIBUCIÓN DE LOS ESTUDIANTES POR LA SITUACIÓN SOCIO ECONÓMICA.

PERIODO 2021-2021 ▼

| VULNERABILIDAD | TOTAL_MATRICULADOS | QUINTIL |
|--------------------------|--------------------|---------|
| COMPLETAMENTE VULNERABLE | 517 | 1 |
| MUY VULNERABLE | 506 | 2 |
| NADA VULNERABLE | 509 | 5 |
| NO DEFINIDO | 803 | |
| POCO VULNERABLE | 516 | 4 |
| VULNERABLE | 533 | 3 |



Nota. Distribución de los estudiantes matriculados por la situación socioeconómica.
Elaborado por: (Loachamin, 2023).

Resultados obtenidos:

El concepto de vulnerabilidad en la institución se lo viene aplicando desde el período académico 2021-2021, el cual se refiere a las posibilidades económicas y agrupa a los estudiantes en un segmento de vulnerabilidad en cuanto a los aranceles de pago por concepto de matrículas y pérdida de la gratuidad. De modo que este análisis permite conocer la distribución de los estudiantes matriculados de acuerdo con el registro del quintil al que pertenecen los estudiantes.

Para los valores uno y dos de completamente vulnerable y muy vulnerable respectivamente, se consideran vulnerables en todo sentido y bajo este análisis

podemos conocer el grupo de estudiantes que pueden aplicar a una beca estudiantil.

5. CONCLUSIONES

- Es importante involucrar a los propietarios de los datos (fuente de datos) y a los usuarios finales en el proceso de gestión de datos. En tal sentido, la institución educativa permite el uso de datos necesarios para la validación del modelo de solución de BI, sin dejar de lado el acuerdo de confidencialidad para el uso de los mismos.
- La metodología de PWD que se utiliza en el desarrollo de la solución de BI es esencial, ya que permite mejorar la calidad de los datos, eficiencia en el procesamiento, flexibilidad en el acceso a los datos y reducir costos y tiempo en la construcción y mantenimiento del DW para respaldar TD.
- La integración de datos en la solución de BI se lo consigue a través del proceso ETL que permite extraer, transformar y cargar datos en un DW centralizado. Este almacén de datos se puede actualizar, ya que para este proceso se creó una tarea de ejecución en Pentaho y posteriormente ser utilizada para informes y análisis.
- En relación con requisitos del proyecto y el nivel de experiencia de usuario de las herramientas de BI, Oracle BI es una herramienta potente que ofrece una diversidad de funcionalidades un tanto complejas. En cualquier caso, es fundamental contar con un equipo con mucha experticia en la implementación de soluciones de BI y contar con las herramientas necesarias para finalizar el proyecto con éxito.
- Se muestra un tablero para la presentación visual de la información como resultado de la prueba de concepto realizado en la FICA. En este caso, los usuarios finales, es decir, las autoridades académicas, son quienes hacen uso de esta información para la toma de decisiones importantes.

REFERENCIAS

- Apraxine, D., & Stylianou, E. (2017). Business intelligence in a higher educational institution: The case of University of Nicosia. En *2017 IEEE Global Engineering Education Conference (EDUCON)* (págs. 1735-1746). Athens: IEEE.
- Bernabeu, R. D. (2010). Metodología para la Construcción de un. Data Warehouse.
- bitec. (2022). Obtenido de bitec: <https://www.bitec.es/noticias-bitec/cuadrante-magico-de-gartner-2022-analisis-y-business-intelligence/>
- Camargo Vega, J. J., Joyanes Aguilar, L., & Giraldo Marin, L. M. (2016). La inteligencia de negocios como una herramienta en la gestión académica -Business intelligence as a tool in academic management. *Revista Científica*.
- Castillo Rojas, W., Quispe, F. M., & Molina, F. F. (2018). Una Metodología para Procesos Data WareHousing Basada en la Experiencia. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informacìon*, 83-103.
- CES - Consejo de Educación Superior. (30 de 09 de 2022). Obtenido de CES - Consejo de Educación Superior: <https://www.ces.gob.ec/>
- Conesa Caralt, J. (2012). Introducción al business intelligence . *Introducción al business intelligence* .
- Fernandez, M., Davila, A., & Angeleri, P. (2017). Data quality applied to an academic business intelligence solution: Lesson learned. En *2017 IEEE Colombian Conference on Communications and Computing (COLCOM)* (págs. 1-6). Cartagena: IEEE.
- 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. *Ingeniare. Revista chilena de ingeniería*, 383-394.
- Garcia, D. (31 de 03 de 2022). *Cuadrante Mágico de Gartner 2022: Power BI y Qlik son líderes en Business Intelligence*. Obtenido de <https://www.bitec.es/noticias-bitec/cuadrante-magico-de-gartner-2022-analisis-y-business-intelligence/>
- Gonzaga, B. R. (2021). Inteligencia de negocios para la toma de decisiones: Un enfoque desde la dirección estratégica de instituciones educativas. *Revista Científic*, 295-312.
- Gonzales, F., & Giancarlo, D. (2018). mplementación de una solución de inteligencia de negocios utilizando la metodología {Hefesto} para las oficinas de contabilidad en universidades públicas. *Universidad Nacional Mayor de San Marcos*.
- Hamed, M., Mahmoud, T., Gómez, J. M., & Kfourri, G. (2017). Using Data Mining and Business Intelligence to Develop Decision Support Systems in Arabic Higher Education Institutions. En *Modernizing Academic Teaching and Research* (págs. 71-84). Lebanon: Springer.
- Hernandez Cruz, L. M., Barrera Lao, J. F., Mex Alvarez, D., Estrada Segovia, G., Rios, C. E., & Lopez Aban, L. A. (2021). Data Analysis using Business Intelligence in the

- educational context. En *2021 16th Iberian Conference on Information Systems and Technologies (CISTI)* (págs. 1-6). Chaves.
- Joyanes Aguilar, L. (2019). *Inteligencia de negocios y analítica de datos: una visión global de business intelligence y analytics*. Alpha Editorial.
- Khan, M. A., Gharibi, W., & Pradhan, S. K. (2014). Data mining techniques for business intelligence in educational system: A case mining. *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*. Hammamet, Tunisia: IEEE.
- Kimball, R., & Ross, M. (2011). *The data warehouse toolkit: the complete guide to dimensional modeling*. Canada: John Wiley & Sons.
- Medina Q., F., Fariña M., F., & Castillo Rojas, W. (2018). Data Mart para obtención de indicadores de productividad académica en una universidad. *Ingeniare. Revista chilena de ingeniería*, 26, 88-101.
- Moscoso Zea, O., Castro, J., Paredes Gualtor, J., & Lujan Mora, S. (2019). A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education. *IEEE Access*.
- Ong, V. K. (2016). Business Intelligence and Big Data Analytics for Higher Education: Cases from UK Higher Education Institutions. 65-75.
- Peralta Ascue, M. (2021). Calidad de Datos en Sistemas de Gestión Académica Universitaria basado en ISO/IEC 25012. *Interfases*, 65-68.
- Stankov, I. (2020). Business Intelligent Systems Data Processing. En *2020 28th National Conference with International Participation (TELECOM)* (págs. 90-93). Sofia.
- Venter, M. (2005). Business intelligence (BI) initiatives: failure and success. *Interim: Interdisciplinary Journal*, 149-163.
- Zeng, L., Lida, X., Zhongzhi, S., Maoguang, W., & Wenjuan, W. (2006). Techniques, Process, and Enterprise Solutions of Business Intelligence. En *2006 IEEE International Conference on Systems, Man and Cybernetics* (págs. 4722-4726). Taipei: IEEE.