



Facultad 2

Mercado de Datos del proceso de selección del personal de RRHH para la XETID

Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas

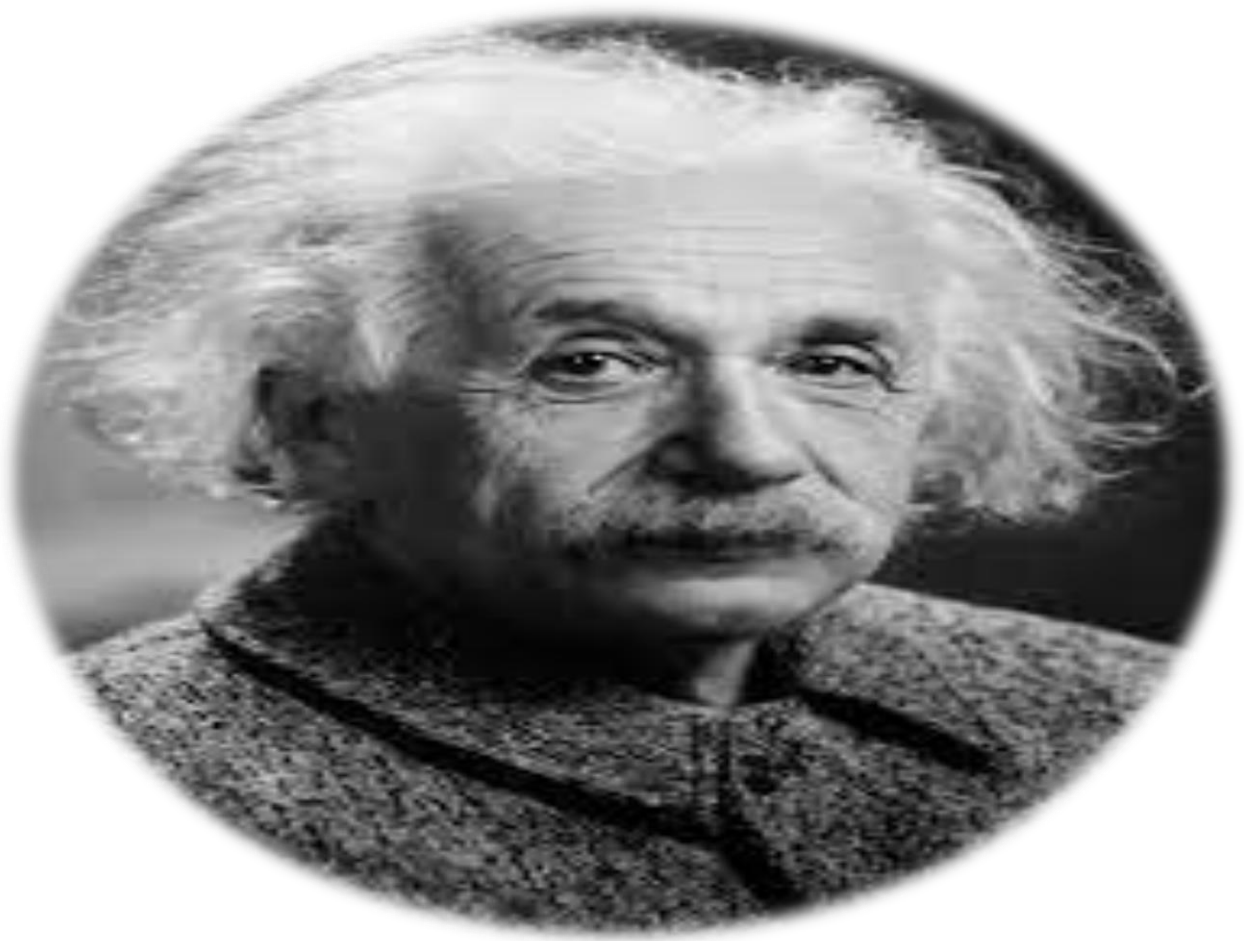
Autor: Arazay Domínguez Canos

Tutoras: MSc. Darling Darías Pérez

Ing. Yanet Quiala Cisnero

La Habana, diciembre de 2023

“Año 65 de la Revolución”



"Nunca pienses en el estudio como una obligación, sino como una oportunidad para penetrar en el bello y maravilloso mundo del saber."

Albert Einstein.

DECLARACIÓN DE AUDITORÍA

El autor del trabajo de diploma con título ***“Mercado de Datos del proceso de selección del personal de RRHH para la Xetid”*** concede a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la investigación, con carácter exclusivo. De forma similar se declara como único autor de su contenido. Para que así conste firma la presente a los 29 días del mes de noviembre del año 2023.

<Arazay Domínguez Canos>

Firma del Autor

<Darling Darías Pérez>

Firma del Tutor

<Yanet Quiala Cisnero>

Firma del Tutor

DATOS DE CONTACTO

Tutora: MSc. Darling Darias Pérez

Graduada de Ingeniera en Ciencias Informáticas. Profesora Principal de 5to año en la carrera de ICI de la Facultad 2. Posee 15 años de experiencia laboral impartiendo clases de las asignaturas de Ingeniería de Software, Gestión de Software, Base de datos, Proyecto de Investigación y Desarrollo e Introducción a las Ciencias Informáticas. De ellos, 6 años como directiva en un Centro y luego en la facultad. Subdirectora del Centro de Informática Médica: CECIM y Vicedecana de la Facultad 7. Además de tener experiencia en el rol de Analista y probadora de la calidad en dicho centro productivo. Realizó y defendió la maestría externa de Informática en Salud, en el Centro de Cibernética Aplicada a las Matemáticas. Ha participado en eventos científicos y grupos investigativos.

Correo Electrónico: ddarias@uci.cu

Tutora: Ing. Yanet Quiala Cisnero

Graduada de Ingeniera en Ciencias Informáticas. Trabaja en la Empresa de Tecnología para la Defensa (XETD). Posee 10 años de graduada y ha trabajado en línea de almacene de datos y sistemas de apoyo toma de decisiones

Correo electrónico: yqcisnero@xetid.cu

AGRADECIMIENTO

Hoy es un día muy importante para mí, ya que estoy presentando mi tesis. Me gustaría aprovechar esta oportunidad para agradecer a todas las personas que me han apoyado y ayudado en este camino.

A mis padres Aracelys y Lázaro que me han dado todo el apoyo y el amor del mundo para que yo saliera adelante en mis estudios.

A mi familia por estar siempre conmigo en las buenas y en las malas, por darme la fuerza necesaria para poseer las ganas de levantarme de cada caída y seguir a delante.

A todos mis amigos que compartieron conmigo en la universidad y estuvieron presentes para mí en todo momento, por su ayuda, disposición y preocupación.

A mis tutoras Msc. Darling Darías y Ing. Yantet Quíala quienes me han ayudado a la elaboración de la tesis siempre que lo necesite y en todo lo que estuvo a su alcance.

A los integrantes del tribunal y mi oponente que por sus consejos y recomendaciones contribuyeron a mi superación y a mejorar el trabajo realizado.

A la Revolución por todas las oportunidades que me ha brindado. Y para terminar a todas las personas que han colaborado en el desarrollo de esta tesis y me han dado de una manera u otra su ayuda en estos años de mi carrera.

DEDICATORIA

Quiero expresar mi gratitud a todas las personas que han sido parte de mi vida y han dejado una huella positiva en mi camino. A mis padres porque todo lo que soy se lo debo a ustedes, por apoyarme en todo momento y confiar siempre en mí. Aquí está el futuro de su sacrificios y amor inmenso. A mi hermano por estar siempre conmigo. A mis amigos por sus risas y aventuras. A mis profesores y mentores, por su dedicación y sabiduría que me han ayudado a crecer como persona y profesional. Y a todos los que han contribuido de alguna u otra manera en la realización de este proyecto, gracias por su colaboración y confianza en mí. ¡Gracias de todo corazón!

RESUMEN

Un Mercado de Datos permite realizar análisis estáticos de los procesos en una entidad, esta herramienta posibilita almacenar datos históricos, obtener una visión de la evolución del negocio y tomar decisiones estratégicas. El proceso de selección del personal de los Recursos Humanos de la XETID presenta inconvenientes a la hora de recopilar datos para una toma de decisión, ya que no permite tomar acciones inmediatas respecto a los principales elementos que pueden dar una mejor contratación del personal con una mejor fuerza de obra calificada. Mediante la presente investigación se desarrolló un Mercados de Datos para el proceso de selección del personal de los Recursos Humanos para la XETID, que permitió centralizar y visualizar a través de Damix la información con el uso de tablas y gráficas. Se utiliza como herramienta de modelado Visual paradigm v8.0, para la extracción, transformación y carga de los datos el Pentaho Data Integration v4.2, además se emplean herramientas de base de datos como PgAdmin y PostgreSQL para el almacenamiento. La metodología utilizada fue Metodología UCID, un híbrido entre las metodologías Hefesto y Kimball. Al concluir la investigación se obtiene la estructura del modelo de datos dimensional que comprende: las tablas de dimensiones, hechos y las medidas necesarias para la visualización de los datos. Se realizaron casos de pruebas los cuales fueron concluidos exitosamente dando como resultado un Mercado de Datos que funciona correctamente.

Palabras claves: Mercado de Datos, Recursos Humanos, Metodología, Herramienta, proceso de selección del personal

ABSTRACT

A Data Market allows static analysis of the processes in an entity. This tool makes it possible to store historical data, obtain a vision of the evolution of the business and make strategic decisions. The XETID Human Resources personnel selection process presents drawbacks when collecting data for decision-making, since it does not allow immediate actions to be taken regarding the main elements that can lead to better hiring of personnel with better qualified workforce. Through this research, a Data Market was developed for the selection process of Human Resources personnel for the XETID, which allowed centralization and visualization through Damix the information as use of tables and graphs. Visual parading v8.0 is used as a modeling tool, Pentaho Data Integration v4.2 is used for data extraction, transformation and loading, and database tools such as PgAdmin and PostgreSQL are also used for storage. The methodology used was UCID Methodology, a hybrid between the Hephaestus and Kimball methodologies. At the end of the investigation, the structure of the dimensional data model is obtained, which includes: dimension tables, facts and the measures necessary for the visualization of the data. Test cases were carried out which were successfully concluded, resulting in a Data Market that works correctly.

Keywords: Data market, Human Resources, Methodology, Tool, personnel selection process

ÍNDICE GENERAL

Declaración de auditoría	III
Datos de contacto	IV
Agradecimiento	V
Dedicatoria.....	VI
Resumen	VII
Abstract.....	VIII
Introducción	1
CAPITULO I: Fundamentación teórica sobre el objeto de estudio.....	8
I.1 Conceptos fundamentales	8
I.2 Mercado de Datos similares	10
I.2.1 Nivel internacional	10
I.2.2 Nivel nacional	10
I.2.3 Mercado de Datos en la Universidad de las Ciencias Informáticas	12
I.3 Mercado de Datos	12
I.3.1 Diseño de la Arquitectura.....	13
I.3.2 Modelado multidimensional	14
I.4 Ambiente de desarrollo	15
I.4.1 Metodología de desarrollo	15
I.4.2 Lenguaje utilizados para la construcción del Mercado de Datos	18
I.4.3 Herramienta para la construcción del Mercado de Datos.....	18
Conclusiones parciales	20
Capítulo II: Análisis y Diseño del Mercado de Datos	21
II.1 Descripción del negocio	21
II.2 Fase 1: Análisis de los requerimientos	21
II.2.1 Identificación de preguntas	21
II.2.2 Identificar indicadores y perspectivas.	22
II.2.3 Construcción del modelo conceptual	25
II.2.4 Aprobación del modelo conceptual por el cliente	25
II.3 Fase 2: Análisis de los OLTP	26
II.3.1 Determinación de indicadores.....	26
II.3.2 Establecer correspondencia	26
II.3.3 Nivel de granularidad	27
II.3.4 Modelo conceptual ampliado	27
II.3.5 Definir las reglas del negocio	29
II.4 Fase 3: Modelo lógico de DW	30
II.4.1: Definir tipo de Modelo Lógico del DW	30
II.4.2 Identificar Dimensiones.....	30
II.4.3 Identificar Hecho.....	31

II.4.4 Realizar unión entre dimensión y hecho.	31
II.4.5 Diseñar tablas y columnas físicas.....	33
Conclusiones parciales	35
Capítulo III Implementación y validación del mercado de datos	37
III.1 Fase 4: Procesos ETL	37
III.1.1 Mapeo de los datos Fuentes-A-Destino.	37
III.1.2 Cargas incrementales de los datos	38
III.1.3 Implementación de los indicadores	41
III.2 Implementación de los XML	43
III.2.1 Plataforma de Inteligencia de negocio Damix	44
III.3 Validación del Mercado de Datos	48
III.3.1 Pruebas para validar el Mercado de Datos.	49
III.3.1 Resultado de las pruebas	51
Conclusiones parciales	51
Conclusiones Generales	53
Recomendaciones	54
Referncia bibliográfica.....	54
Bibliografía.....	58
Anexos.....	62

ÍNDICE DE TABLAS

Tabla 1 Características de los sistemas analizados	11
Tabla 2 Ficha del indicador Cantidad de candidatos contratados.....	23
Tabla 3 Ficha del indicador Número de candidatos entrevistados.....	23
Tabla 4 Ficha del indicador Porcentaje de candidatos descalificados	23
Tabla 5 Ficha del indicador Cantidad de candidatos presentados.....	24
Tabla 6 Ficha del indicador Cantidad de candidatos excluidos	24
Tabla 7 Especificación de indicadores	26
Tabla 8 Nivel de granularidad	27
Tabla 9 Descripción de las perspectivas	28
Tabla 10 Descripción de los campos de las perspectivas	29
Tabla 11 Dimensiones del Mercado de Datos	30
Tabla 12 Hecho del Mercado de Datos	31
Tabla 13 Mapa Lógico de Datos	37
Tabla 14 Lista de chequeo	50

TABLA DE ILUSTRACIONES

Ilustración 1 Arquitectura de los Mercados de Datos	14
Ilustración 2 Modelo Conceptual.	25
Ilustración 3 Modelo Conceptual Ampliado	28
Ilustración 4 Unión de las dimensiones con el hecho	32
Ilustración 5 Modelo de dato físico	32
Ilustración 6 Columnas y llave primaria dimensión proceso	34
Ilustración 7 Columnas y llave primaria dimensión persona	34
Ilustración 8 Columnas y llave primaria dimensión motivo de exclusión	34
Ilustración 9 Columnas y llave primaria dimensión resultado de selección	34
Ilustración 10 Columnas y llave primaria dimensión tiempo	35
Ilustración 11 Columnas, llave primaria y llaves foráneas del cubo selección del personal ..	35
Ilustración 12 Consulta de la Dimensión Proceso	38
Ilustración 13 ETL correspondiente a la Dimensión Proceso.....	38
Ilustración 14 Consulta de la Dimensión Motivo de exclusión	39
Ilustración 15 ETL correspondiente a la Dimensión Motivo de exclusión	39
Ilustración 16 Consulta de la Dimensión Resultado de selección	39
Ilustración 17 ETL correspondiente a la Dimensión Resultado de selección	39
Ilustración 18 Consulta de la Dimensión Persona	40
Ilustración 19 ETL correspondiente a la Dimensión Persona	40
Ilustración 20 ETL correspondiente a la Dimensión Tiempo.....	40
Ilustración 21 Consulta del cubo selección del personal	41
Ilustración 22 ETL correspondiente a la Hecho Selección del personal	41
Ilustración 23 Indicador cantidad de candidatos contratados	42
Ilustración 24 Indicador número de candidatos entrevistados	42
Ilustración 25 Indicador cantidad de candidatos excluidos	42
Ilustración 26 Fragmento del archivo XML correspondiente las dimensiones.....	43
Ilustración 27 Fragmento del archivo XML correspondiente al cubo.....	44
Ilustración 28 Adicionar una fuente de datos.....	45
Ilustración 29 Componente de análisis.....	46
Ilustración 30 Tablero 1. Candidatos presentados y por ciento de descalificados	47
Ilustración 31 Tablero 2 Porcentaje de candidatos descalificados.....	47
Ilustración 32 Tablero 3 Cantidad de candidatos presentados	48
Ilustración 33 Tablero 4 Cantidad de candidatos contratados y excluidos.....	48
Ilustración 34 Resultados encontrados por iteración	51

Introducción

Las tecnologías en el mundo de hoy presentan un desarrollo científico técnico, es por ello que la informatización de la sociedad ha crecido proporcionando que aumente la capacidad de generación y almacenamiento de la información.

Las Tecnologías de la Información y la Comunicación (TIC) son cada vez más usadas para el apoyo y automatización de todas las actividades de las empresas. Gracias a ellas, las organizaciones han conseguido obtener importantes beneficios, entre los que cabe mencionar: la mejora de sus operaciones, llegada a una mayor cantidad de clientes y la optimización de sus recursos. Otros como: la apertura a nuevos mercados, un conocimiento más profundo acerca de las necesidades de la clientela para brindarles un servicio de mejor calidad y una comunicación más fluida, no sólo con sus empleados sino también con sus clientes y proveedores. (Ayala, et al., 2015)

Las TIC se han convertido a lo largo de los años en una pieza clave a la hora de gestionar una organización, tanto de sus recursos materiales como humanos y la tendencia es que siga ganando importancia exponencialmente. *“La informatización de la sociedad cubana ha sido uno de las prioridades del país. Siempre se ha considerado estratégico el uso de las TIC, no solo en instituciones y empresas, sino también al alcance del pueblo. En los últimos años se ha encontrado un contexto favorable al uso masivo de las TIC en Cuba, debido a la disminución del costo del equipamiento, mayor acceso a Internet y generalización del uso de los dispositivos móviles”.* (Gaínza Martínez, et al., 2022)

Hoy en día, el valor al talento humano es uno de los más preciados, y uno de los retos principales de las empresas es detectarlo, atraerlo, motivarlo y retenerlo. Para ello, es necesario dejar atrás los modelos tradicionales para adaptarse a las nuevas tecnologías.

Una de las áreas que ha avanzado en la informatización de sus procesos es el área de los Recursos Humanos (RRHH). La expresión de Recursos Humanos hace referencia a aquellas personas que forman parte de la organización y que contribuyen a su funcionamiento y éxito.

La digitalización de esta área, es una tendencia irrefrenable y si bien supone notablemente ventajas para las empresas, también obligan a gestionar

correctamente esta compleja transición. Los RRHH igual que el resto de los departamentos en una organización se han visto involucrado en los últimos años en el uso de los nuevos modelos digitales que se están desarrollando en el ámbito empresarial.

Entre las ventajas que proporcionan las nuevas tecnologías al departamento de los RRHH se destacan las siguientes: (Gan Estaún, 2017)

- Agiliza los procesos y evita pérdida de tiempo.
- Aumenta la calidad de las operaciones y menor riesgo de errores.
- Reducción de la carga administrativa en el departamento de los Recursos Humanos.
- Seguimiento a tiempo real y formación continua.

“La Gestión de los Recursos Humanos es primordial para las organizaciones empresariales; desde las últimas décadas se ha convertido en uno de los factores fundamentales de las políticas empresariales; su impacto tiene alcance no solo al interior de la organización, es un fenómeno que trasciende al ámbito social.” (Armijos Mayon, et al., 2019)

En Cuba, la Gestión de los Recursos Humanos en el contexto laboral está regida por el Ministerio de Trabajo y Seguridad Social (MTSS) y se integra orgánicamente con el sistema de dirección y gestión empresarial, en el decreto No 281 de 2007, dictado por el consejo de ministro. (Sounto Anido, et al., 2016).

Es importante aclarar que esta gestión es realizada por el departamento de Recursos Humanos. Su tarea principal es lograr que la organización posea la capacidad de emplear adecuadamente las herramientas con el fin de lograr los objetivos propuestos y por supuesto, sea productiva, mediante la administración adecuada de los Recursos Humanos. *“Entre sus principales funciones se encuentran: gestión administrativa de personal, reclutamiento y selección de personal, formación y desarrollo profesional, relaciones laborales, prevención de riesgos laborales, evaluación del desempeño y beneficios sociales”.* (Economipedia, 2022b)

Finalmente se afirma que una adecuada gestión de Recursos Humanos ayuda a que la empresa se vuelva más productiva, porque permite que el trabajador se integre mejor a la empresa y se comprometa a alcanzar las metas de la organización.

En la Empresa de Tecnologías de la Información para la Defensa (XETID), de origen cubano, dedicada al sector del software, la automática y las comunicaciones, existe un sistema que se encarga de gestionar los datos asociados a los procesos de Gestión de Recursos Humanos, conocido como “El Sistema Integral de Gestión de Recursos Humanos (SIGERH)”, puesto a disposición por el Ministerio de Trabajo y Seguridad Social e implementado por la empresa XETID. Cuenta con nueve módulos: estructura y comparación, registro de personas, gestión de movimientos laborales, registro y control de asistencia, evaluación, capacitación y desarrollo, idoneidad, seguridad y salud del trabajo y selección del personal. Este sistema tiene integrado un cuadro de mando diseñado e implementado por Damix el cual permite la visualización de la información.

SIGERH cuenta con una base de datos que almacena toda la información que se maneja en el departamento de los RRHH. El diseño de esta base de datos aumenta la complejidad de las consultas para poder determinar una estadística puntual y no permite que se pueda disminuir los tiempos de búsqueda relacionados.

Dentro de los procesos que se llevan a cabo por la dirección del departamento hay un proceso estratégico que es la selección del personal, el cual tiene como inconveniente: que no permite tomar acciones inmediatas respecto a los principales elementos que pueden dar una mejor contratación del personal con una mejor fuerza de obra calificada.

Propiciando varias dificultades como:

- No se encuentra disponible la información del proceso de selección del personal de una manera más organizada para que en el área de RRHH se puedan consultar los datos de manera rápida y fácil.
- No se puede acceder de forma inmediata a datos históricos del proceso de selección del personal.
- No se permite realizar comparaciones con la totalidad de los datos significativos en un determinado intervalo de tiempo, debido a la forma en que está organizada dicha información.
- El proceso de selección del personal hoy a partir de los datos existentes es más complejo y necesita disminuir el tiempo dedicado para que se pueda procesar debido a que se encuentra en un entorno que tiene muchos más datos que no hacen falta para este proceso.

- Se desconoce la tendencia que tiene la selección del personal a la hora de realizar análisis del proceso de selección del personal.

Antes la situación planteada, se identifica como **problema a resolver**: ¿Cómo facilitar el análisis del comportamiento del proceso de selección del personal de los Recursos Humanos para la XETID?

Trazando como **objetivo general**: Desarrollar un Mercado de Datos para facilitar el análisis del comportamiento en el proceso de selección del personal de los Recursos Humanos para la XETID.

Enmarcado en el **objeto de estudio**, proceso de selección del personal, delimitado por el **campo de acción**, proceso de selección del personal de los Recursos Humanos para la XETID.

Como **objetivos específicos** se definen los siguientes:

1. Construir los referentes teóricos relacionando los aspectos teóricos fundamentales que sustentan la investigación, mediante los cuales se consulta, extrae y recopila la información relevante sobre el problema a investigar.
2. Definir y emplear las herramientas informáticas para el desarrollo del Mercado de Datos del proceso de selección del personal de los RRHH en la XETID.
3. Identificar los indicadores claves de desempeño para medir el éxito del proceso de selección del personal de los RRHH en la XETID.
4. Definir la forma para medir y evaluar los resultados del proceso de selección del personal de los RRHH en la XETID.
5. Configurar los indicadores a emplear para medir el proceso de selección del personal.
6. Validar el Mercado de Datos del proceso de selección del personal mediante las pruebas de usabilidad.

Como **tareas de investigación** se definen los siguientes:

1. Evaluación del proceso de selección, utilizando diferentes técnicas como encuesta a empleados, análisis de datos de contratación y entrevistas con los responsables del proceso

2. Estudio de las herramientas informáticas a utilizar para elaborar el Mercado de Datos del proceso de selección del personal de los Recursos Humanos en la XETID
3. Estudio e identificación de los Indicadores Claves de Desempeño para medir el éxito del proceso de selección del personal
4. Configuración de los indicadores para el Mercado de Datos del proceso de selección del personal
5. Estudio de casos de prueba para garantizar el correcto funcionamiento del sistema

Posibles Resultados:

1. El diseño conceptual de los hechos y dimensiones del Mercado de Datos.
2. El Mercado de Datos del proceso de selección del personal de los RRHH para la XETID.
3. Las configuraciones de los indicadores para determinar una adecuada selección.

Beneficios:

1. Mejora en la calidad del proceso de selección del personal de los RRHH en la XETID.
2. Visualiza el comportamiento del proceso de selección del personal para ayudar a la toma de decisiones de los jefes.
3. No se genera costo a la empresa a la hora de ser utilizado.
4. Disminuye el tiempo de trabajo ya que se incorporó a SIGERH el proceso de selección del personal.

Para realizar esta investigación se utilizaron los siguientes métodos científicos:

Métodos Teóricos:

Analítico-Sintético: Se emplea para analizar los elementos bibliográficos y definiciones sobre los Mercados de Datos existentes, que realicen un trabajo similar al propuesto, con el propósito de arribar a conclusiones que sustenten la necesidad de la investigación.

Modelación: Se emplea en la confección del modelo conceptual, lógico y físico de los datos que se generan en la metodología utilizada para el desarrollo de la propuesta de solución.

Métodos Empíricos:

Revisión Bibliográfica: Consiste en la revisión de diferentes fuentes bibliográficas que aborden el tema proceso de selección de personal, su aplicación en la gestión de Recursos Humanos y los Mercado de Datos. Esta revisión permitirá identificar los conceptos claves y las mejores prácticas en la materia.

Observación: Implica la observación directa del proceso de selección del personal por parte del investigador, quien puede participar activamente en el proceso para obtener una comprensión más profunda de los indicadores utilizados y su efectividad.

Entrevista: Se realizó a los responsables del proceso de selección del personal y otros candidatos relevantes en la organización para obtener información sobre este proceso y su percepción sobre su efectividad. Puede realizarse de forma individual o grupal.

El documento está estructurado por: índices, introducción, tres capítulos, conclusiones, recomendaciones, referencias bibliográficas, bibliografías y anexos:

Capítulo I. Fundamentación teórica:

Se definen los conceptos fundamentales asociados con el objeto de estudio y el campo de acción, con el objetivo de lograr un mayor entendimiento del problema a resolver. De igual manera se realiza un estudio del proceso de selección del personal a nivel nacional e internacional. Se caracterizará la metodología UCID de desarrollo de Data Warehouse (DW) empleada. Finalmente se caracterizan las herramientas y tecnologías utilizadas.

Capítulo II. Análisis y diseño del Mercado de Datos:

Se dará una breve descripción del negocio. Se identifican perspectivas e indicadores para el desarrollo de tablas. Además se muestran los conceptos asociados al modelo conceptual y el modelo conceptual ampliado, así como la representación formal de ellos. Se realiza el modelo de datos según el estilo de diseño de almacén de datos.

Capítulo III. Implementación y validación del Mercado de Datos:

Se realizan las configuraciones e implementaciones necesarias de los indicadores del proceso de selección, así como el desarrollo de los procesos ETL. Se describen

las estrategias de pruebas a utilizar y se aplican las pruebas para poder detectar los errores existentes.

CAPITULO I: FUNDAMENTACIÓN TEÓRICA SOBRE EL OBJETO DE ESTUDIO.

Generalmente, el estudio de las bases de datos se inicia considerando las bases de datos relacionales, que son las que, de manera mayoritaria, están implantadas desde hace unos años en la industria. Este tipo de bases de datos ofrece apoyo al negocio de la organización y permite almacenar los datos y el procesamiento de la información generada día a día. Todo esto implica que están diseñadas para hacer operaciones de consulta y actualización de manera eficiente por parte de distintos usuarios. En este capítulo se abordará un nuevo tipo de bases de datos que están orientadas a ofrecer apoyo a la toma de decisiones en la organización. Se trata de los denominados almacenes de datos.

I.1 Conceptos fundamentales

A continuación, serán expuestos algunos conceptos, asociados al tema, como resultado de la consulta bibliográfica realizada para una mejor interiorización del mismo.

1.1.1 Almacén de Datos (Data Warehouse)

Sobre el tema en particular varios son los autores que lo abordan por ejemplo, (Bustos Barrera, et al., 2013) manifiesta que *(cito) “Colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo que ayuda a la toma de decisiones en la entidad en la que se utiliza”*

Otro autor lo considera *“Una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis”* (Kimball , 2008)

Sin embargo independientemente de la variedad de autores que definen el concepto para este estudio se asume el criterio de Bustos Barrera.

1.1.2 Mercado de Datos (Data Mart)

Sobre este concepto son varios los autores que lo abordan, por ejemplo, según (Wolff, 2000) *lo define como: “Un conjunto de hechos y datos organizados para soporte decisional basados en la necesidad de un área o departamento específico. Los datos son orientados a satisfacer las necesidades particulares de un departamento dado, teniendo solo sentido para el personal de ese departamento y sus datos no tienen por qué tener las mismas fuentes que los de otro Data Mart”.*

Otro autor lo considera *“Forma sencilla de un Almacén de Datos que se centra en*

un único tema o línea de negocio, como ventas, finanzas y marketing” (Oracle, 2022)

Para este estudio se asume el concepto de Oracle porque es el que se ajusta a los requerimientos del problema a resolver.

1.1.3 Indicadores claves de desempeño (KPI)

En este campo varias son las opiniones encontradas a partir de la bibliografía consultada, por ejemplo, *Según (Riaño, 2020) manifiesta que (cito) “Indicadores versátiles y fáciles de entender, que permite una diversidad de combinaciones para medir el cumplimiento de objetivos y metas de una organización o proyecto”.*

Otro autor lo define como “Herramientas de gestión ampliamente utilizadas por las empresas en todo el mundo para medir y evaluar el desempeño de sus procesos y gestionarlos de la manera más eficaz y eficiente posible” (Gesé Bordils, et al., 2021)

No obstante se asume el criterio de Gesé Bordils.

1.1.4 Selección del personal

Los autores Chiavenato y Espinosa sobre el tema en cuestión plantean que:

“La selección del personal puede definirse como el proceso de elección del individuo adecuado para el cargo, o en sentido más amplio, escoger entre los candidatos reclutados a los que más se acoplen al puesto, para ocupar los cargos existentes en la empresa.” (Chiavenato, 2006)

“Filtro para poder permitir dejar entrar a las organizaciones solo a algunas personas, y estas son las que cubren los requerimientos previamente establecidos que permiten proteger el capital intelectual de la organización.” (Espinosa, et al., 2019)

Sin embargo para esta investigación se asume el criterio de Chiavenato.

1.1.5 Recursos Humanos (RRHH)

Sobre el tema en particular varios son los autores que lo abordan por ejemplo, (Cuesta, 2005) *lo define como “Expresión utilizada para representar a las personas insertadas en una organización laboral”* Para (Machado Marquez, 2021) *“La expresión de Recursos Humanos se refiere a las personas que forman parte de las organizaciones y que desempeñan en ellas determinadas funciones para dinamizar los recursos organizacionales”*

Sin embargo independientemente de la variedad de autores que definen el concepto para este estudio se asume el criterio de Cuesta.

I.2 Mercado de Datos similares

Son muchas las empresas e instituciones que optado por la implementación de los Mercado de Datos y Almacenes de Datos como sistemas de almacenamiento tanto a nivel nacional como a nivel internacional. A continuación se realiza un estudio sobre este tema:

I.2.1 Nivel internacional

“Implementación de un Data Mart para la ayuda en la toma de decisiones de la Gerencia de Recursos Humanos” (IDMGRRHH): permite al área de los Recursos Humanos un mayor grado de agilidad en la obtención de reportes para el análisis de los datos, permitiendo además visualizar y analizar datos estadísticos y su evolución. Toda esta información se dispone mediante un fácil acceso y de forma estructurada para ayudar en la toma de decisiones de la Gerencia.

Para el proceso ETL se utilizó Pentaho Data Integration, para la estructuración de la información el esquema en estrella y se siguió la metodología sugerida por Ralph Kimball conocida como Bottom-up. El Mercado de Datos fue desarrollado en la Agencia de Recaudación Tributaria de Río Negro (ARTRN), en el año 2021.

I.2.2 Nivel nacional

“Mercado de Datos para el área de los Recursos Humanos” (MDRRHH): permite al vicedecanato de formación de la Facultad de Ciencias y Tecnologías Computacionales (FCITEC) tomar decisiones relacionadas con las actividades de los Recursos Humanos, donde se incluye lo referente a la pre Nómina para el pago a los trabajadores que laboran en dicha facultad.

Para el proceso de integración de datos se utilizó Pentaho Data Integration, para la estructuración de la información el esquema en constelación y se define como metodología a utilizar: Propuesta de Metodología para el desarrollo de Almacén de Datos en el Centro de Tecnologías de Gestión de Datos (DATEC). Esta metodología se ajusta a las condiciones y características de la producción del centro y la universidad. El Mercado de Datos fue desarrollado por el centro DATEC en conjunto con la Facultad de Ciencias y Tecnologías Computacionales (FCITEC) en el año 2017.

“Mercado de Datos en apoyo a la toma de decisiones sobre el personal docente e investigativo en el departamento de los Recursos Humanos de la UCLV” (MDTDPUCLV): permite al departamento de los Recursos Humanos contar con un Mercado de Datos que analice periódicamente la información relacionada con el departamento de los Recursos Humanos de la UCLV.

Para el proceso de integración de datos se utilizó Pentaho Data Integration, para la estructuración de la información el esquema en constelación y como metodología Kimball. El Mercado de Datos fue desarrollado por el departamento de los Recursos Humanos de la Universidad Central “Marta Abreu” de Las Villas en el año 2013.

Una vez realizado el análisis de un grupo de sistemas, tanto internacionales como nacionales, examinando sus potencialidades y características técnicas (ver Tabla 1), se puede afirmar que ninguno de los sistemas analizados satisface en su totalidad las necesidades de la investigación. Estos posibilitan el análisis y la toma de decisiones en el área de los Recursos Humanos, pero no están orientados a analizar el comportamiento referente al proceso de selección del personal. Además no se tiene la aplicación del sistema y no persiguen la metodología que se definió en la empresa. Por lo antes planteado es necesario implementar un sistema que satisfaga las necesidades de la investigación, permitiendo a los especialistas de los Recursos Humanos realizar el análisis estadístico de la información y tomar decisiones apropiadas sobre el proceso de selección del personal.

No obstante el análisis de estos sistemas de almacenamientos posibilitó un mayor acercamiento al Mercado de Datos obteniendo nuevos conocimientos sobre su diseño e implementación. Permitió conocer que para los procesos ETL se utiliza la plataforma Pentaho, la cual esta compuestas por varias herramientas para realizar los procesos ETL pero la más utilizada es Pentaho Data Integration. Además permitió conocer técnicas y herramientas a utilizar en el desarrollo del Mercado de Datos y la construcción del diseño de la propuesta de solución.

Tabla 1 Características de los sistemas analizados

Sistemas		Herramienta para el proceso ETL	Modelado multidimensional		Metodología	
			Estrella	Constelación	Propia	Kimball

Internacional	IDMGRRHH	PDI	x			x
Nacional	MDRRHH	PDI		x	x	
	MDTDPUCLV	PDI		X		X

I.2.3 Mercado de Datos en la Universidad de las Ciencias Informáticas

Entre las investigaciones consultadas se encuentra (González Piñera, 2021) la cual tiene como título *“Mercado de Datos para el proceso de perfeccionamiento educacional en Cuba”*. Esta investigación se enfoca en el uso de datos de mercados para mejorar la educación en Cuba. Además implica el análisis de estos mercados, identificar las necesidades y demandas educativas de la población cubana, así como para evaluar la eficiencia de las políticas y programas educativos en Cuba.

Otras de las investigaciones que se consulto fue (Silvero Jovel, et al., 2017) titulada: *“Mercado de Datos para el primer año de la carrera de Ingeniería en Ciencias Informáticas”*. El propósito de esta investigación fue desarrollar un Mercado de Datos que centralice la información académica y personal de los estudiantes de la FICI para disminuir las limitaciones en el proceso de toma de decisiones de los profesores y directivos debido a que la información se encuentra dispersa en varias fuentes de datos, lo que imposibilita el análisis integral de los estudiante.

I.3 Mercado de Datos

Una estructura muy particular de los Almacenes de Datos son los Mercado de Datos. El Mercado de Datos es un subconjunto de datos con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones.

Según las definiciones anteriormente dadas sobre los Mercados de Datos se puede concluir que son pequeños almacenes de datos centrados en un tema o un área de negocio específico dentro de una organización. Una de las diferencia que existe entre ellos es el ámbito de la información que contienen debido a que en los mercados es más pequeño y los datos se obtienen de un menor número de fuentes y comúnmente el tiempo de desarrollo es menor.

Entre las características de un Mercado de Datos se encuentran las siguientes: (Castillo O' Farrill, et al., 2017)

- Colección de datos que contiene información de un área de negocio específica.
- Organiza la información de tal forma que se alcanza un alto nivel de detalle.

- Permite examinar los datos desde diferentes perspectivas, logrando así un mejor control de la información.

Dentro de las ventajas de aplicar un Mercado de Datos a un negocio se han seleccionado las siguientes: (Bernabeu, 2007)

- Son simples de implementar.
- Conllevan poco tiempo de construcción y puesta en marcha.
- Permiten manejar información confidencial.
- Reflejan rápidamente sus beneficios y cualidades.
- Reducen la demanda del depósito de datos

La información que se desea gestionar en esta investigación es la referente al proceso de selección del personal de los Recursos Humanos, por lo que se hace necesario implementar un Mercado de Datos, debido a que estos cubren las necesidades de una determinada área dentro de la organización, y el costo de su uso es inferior por lo que se lleva menor tiempo para construirlo y ponerlo a funcionar; mientras que un Almacén de Datos cubre las necesidades de la organización en su conjunto, y el costo de su uso es mayor por lo que construirlo y ponerlo a funcionar llevaría más tiempo.

I.3.1 Diseño de la Arquitectura

Los Almacenes de datos están compuestos por una fuente de datos y tres subsistemas propios de este tipo de sistema que son: el subsistema de integración, el subsistema de almacenamiento y el subsistema de visualización. A continuación, se explican cada una de sus componentes: (Silvero Jovel, et al., 2017)

- **Fuente de datos:** lugar donde se extrae la información para probar el Almacén de Datos.
- **Subsistema de integración:** encargado de realizar todos los procesos de ETL donde se extrae, se limpia y se integra toda la información almacenada en las fuentes de datos a través de transformaciones.
- **Subsistema de almacenamiento:** donde se guarda toda la información luego de haber sido transformada en el subsistema de integración.
- **Subsistema de visualización:** encargado de mostrar al cliente toda la información almacenada, a través de reportes y vistas de análisis permitiendo al cliente poder analizar la información procesada.



Ilustración 1 Arquitectura de los Mercados de Datos

I.3.2 Modelado multidimensional

Debido al enfoque analítico de la tecnología de los almacenes surge una forma diferente de pensamiento y procesamiento lo que se evidencia en un modelado de bases de datos propias llamado modelo mutidimensional. Para la creación del Mercado de Datos uno de los modelos más utilizados son los modelos multidimensionales, los cuales dieron surgimiento a las bases de datos multidimensionales, garantizando una búsqueda rápida de los datos.

El Modelado multidimensional se representa a través de la definición de las tablas de hechos y dimensiones y indicadores.

Tablas de hechos: *“Las tablas de hechos son las tablas primarias en el modelo dimensional. Generalmente, almacenan medidas numéricas, las que representan valores de las dimensiones. La llave de la tabla de hecho, es una llave compuesta, debido a que se forma de la composición de las llaves primarias de las tablas dimensionales a las que está unida”.* (Castillo O' Farrill, et al., 2017)

Tablas de dimensiones: *“son elementos que contienen atributos o campos que se utilizan para restringir y agrupar los datos almacenados en una tabla de hechos cuando se realizan consultas sobre dicho datos en un entorno de Almacén de Datos o Mercado de Datos”.* (Acosta del Sol, et al., 2016)

Indicadores o medidas: *“son aquellos conceptos cuantificables que permiten medir nuestro proceso de negocio. Son características cualitativas o cuantitativas de los*

objetos que se desean analizar en las empresas. Los indicadores, siempre son numéricos, se almacenan en las tablas de hechos". (Guerra Viltres, 2020)

Las bases de datos dimensionales tienen tres variantes posibles de modelación, los cuales se mencionan a continuación:

Esquema copo de nieve(Snowflake Schema):este esquema posee una tabla de hechos que se encuentra en el centro y está rodeada de dimensiones, puede tener muchas filas, la información de las dimensiones no se encuentran normalizada esto quiere decir que para obtener información se tomaría mucho tiempo en hacerlo (Acero Calizaya, 2014). Las dimensiones van a depender de otras dimensiones(**ver anexo 2**)

Esquema de constelación(Constellation Schema):este modelo está compuesto por una serie de esquemas en estrellas, es decir, está formado por una tabla de hechos principal (HECHOS_A) y por una o más tablas de hechos auxiliares (HECHOS_B) que están relacionadas con sus respectivas tablas de dimensiones vinculándose las tablas de hechos auxiliares con algunas dimensiones asignadas a las tablas de hecho principal y también con nuevas tablas de dimensiones. (Bernabeu, 2007)(**ver anexo 3**)

Esquema estrella(Star Schema):este esquema logra que los datos multidimensionales sean representados en base relacional, las dimensiones contienen información que describen las relaciones, y la tabla de hechos contienen datos numéricos (Acero Calizaya, 2014).(**ver anexo 4**)

I.4 Ambiente de desarrollo

I.4.1 Metodología de desarrollo

La bibliografía consultada muestra que son pocos los autores que en plano nacional abordan el tema sin embargo es amplia la encontrada en el plano internacional. Como resultado del análisis de la misma para el cumplimiento del objetivo, del estudio de metodología de desarrollo se asume el definido por (Blanco-Cuaresma, 2019) que la entiende como “...una colección de documentación formal referente a los procesos, las políticas y los procedimientos que intervienen en el desarrollo del software. La finalidad de una metodología de desarrollo es garantizar la eficacia en el proceso de generación de software”.

Metodología a emplear:

Metodología UCID para el Desarrollo de Data Warehouse (DW):

En el centro de Análisis de Datos de la XETID para la implementación de Mercado de Datos y Almacén de Datos se cuenta con una metodología propia la cual fue adoptada por el centro de datos de la UCI adaptándola a las necesidades propias de UCID denominada Metodología UCID para el Desarrollo de Data Warehouse (DW). *“Esta es un híbrido entre la metodología de Hefestos y la de Kimball. Se decidió realizar un híbrido entre estas dos metodologías y no adoptar una de las dos, porque se quiere aprovechar la facilidad que brinda Hefestos de una metodología sencilla y entendible con la de Kimball de generar artefactos para la documentación a la hora de realizar el Almacén de Datos. La metodología rectora fue la de Hefestos y no la de Kimball”.* (Xetid, 2012)

Entre sus principales características se encuentran: (Xetid, 2012)

1. Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
2. Se basa en los requerimientos del usuario, por lo cual su estructura es capaz de adaptarse con facilidad y rapidez ante los cambios en el negocio.
3. Reduce la resistencia al cambio, ya que involucra al usuario final en cada etapa para que tome decisiones respecto al comportamiento y funciones del DW.
4. Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.
5. Es independiente del tipo de ciclo de vida que se emplee para contener la metodología.
6. Es independiente de las herramientas que se utilicen para su implementación.
7. Es independiente de las estructuras físicas que contengan el DW y de su respectiva distribución.
8. Cuando se culmina con una fase, los resultados obtenidos se convierten en el punto de partida para llevar a cabo el paso siguiente.
9. Se aplica tanto para Data Warehouse como para Data Mart.

Cuenta con 4 fases: (Xetid, 2012)

Fase 1: Análisis de los requerimientos

Se planifica cada una de las entrevistas que se realizarán con el cliente, luego se identificarán los requerimientos del usuario a través de preguntas que expliciten los objetivos de su organización. Se analizarán estas preguntas a fin de identificar cuáles serán los indicadores y perspectivas que serán tomadas en cuenta para la construcción del DW. Se confeccionará un modelo conceptual en donde se podrá visualizar el resultado obtenido de la entrevista y se concluye esta primera fase con la validación del mismo por el cliente donde se obtendrá un acta de validación para poder continuar con las demás fases.

Fase 2: Análisis de los OLTP

Se analizarán las fuentes OLTP (Proceso de transacciones en línea) para definir el estado general de las mismas y determinar cómo serán calculados los indicadores y para establecer las respectivas correspondencias entre el modelo conceptual creado en el paso anterior y las fuentes de datos. Luego, se definirán qué campos se incluirán en cada perspectiva. Se ampliará el modelo conceptual con la información obtenida en este paso, y se finalizará definiendo las reglas del negocio.

Fase3: Modelo lógico del DW

Se confeccionará el modelo lógico de la estructura del DW, teniendo como base el modelo conceptual que ya ha sido creado. Para ello, primero se definirá el tipo de modelo que se utilizará así como los estándares para los objetos físicos; luego se llevarán a cabo las acciones propias al caso, para diseñar las tablas de dimensiones y de hechos. Se realizarán las uniones pertinentes entre estas tablas. Y finalmente se diseñarán las tablas y columnas físicas del DW.

Fase 4: Proceso ETL

Se deberá proceder a probar con datos una vez construido el modelo lógico e implementado el mismo en el DW, a través de procesos ETL. Primeramente se realizará un mapeo de los datos fuentes hacia los destinos detallando claramente de donde salen y hacia donde se dirigen, luego se establecerán algunas restricciones y condiciones adicionales en caso que lo requiera. Una vez definidas estas

transformaciones se procede a la carga incremental de los datos hacia las dimensiones y hechos según corresponda. Finalmente se diseñará y construirá la automatización de todo el proceso.

I.4.2 Lenguaje utilizados para la construcción del Mercado de Datos

En el centro de Análisis de Datos de la XETID para la implementación de los Mercado de Datos y Almacén de Datos se utilizan los siguientes lenguajes de desarrollo, a continuación se caracterizan cada uno de ellos:

El lenguaje de Consulta Estructurado v2019 (en lo adelante SQL, por sus siglas en inglés, Structured Query Language) *“es un lenguaje de base de datos normalizado, utilizado por la mayoría de los servidores de bases de datos que manejan bases de datos relacionales u objeto-relaciones”*. (Alonso Sarría, 2016). Este lenguaje fue empleado a la hora de realizar las consultas para poblar las tablas dimensiones.

El lenguaje de Mercado Extensible v1.1 (en lo adelante XML, por sus siglas en inglés, Extensible Markup Language) *“comenzó a desarrollarse en septiembre de 1996 auspiciado por el W3C con un claro propósito: diseñar un lenguaje de marcas optimizados por internet. XML debía combinar la simplicidad de HTML con la capacidad expresiva de su procesador, SGML (Standard Generalized Markup Language)”*. (Eito Brum, 2005). Este lenguaje fue utilizado para la elaboración del archivo XML utilizado en la herramienta Damix.

I.4.2.1 Lenguaje de modelado

El Lenguaje Unificado de Modelado v2.0 (en lo adelante UML, por sus siglas en inglés, Unified Modeling Language) *“es un language de propósito general usado para especificar, visualizar y construir artefactos de sistemas de software”*. (Booch, et al., 2005).

I.4.3 Herramienta para la construcción del Mercado de Datos

En el centro de Análisis de Datos de la XETID para la implementación de Mercado de Datos y Almacén de Datos se utilizan las siguientes herramientas de desarrollo.

Después de haberse realizado una consulta bibliográfica sobre estas herramientas se asumieron las siguientes definiciones:

Damix v1.1: *“es una herramienta que permite visualizar la información resumida de una forma sencilla facilitando la toma de decisiones estratégicas mediante cuadros de mando o tableros de control que, muestra vistas en tablas dinámicas, mapas temáticos y gráficos, así como un sistema de planificación estratégicas y representa la relación causa _efecto entre los objetivos estratégicos a indicadores que se visualicen en el tablero de control”.* (Xetid, 2019). Es la plataforma de inteligencia de negocio que utiliza SIGERH para la visualización de la información a través de tableros de control.

PostgreSQL v9.4: *“Posee características técnicas que la hacen una de las bases de datos más potentes y robustas del mercado. La característica que más se han tenido en cuenta durante su desarrollo son la estabilidad, potencia, robustez, facilidad de administración e implementación de estándares. Utiliza el modelo cliente servidor y usa multiprocesos en vez de multihilos para garantizar la estabilidad de la misma .Un fallo en uno de sus procesos no afectará el resto y el sistema continua funcionando.”* (Martínez, 2013)

SQL Manager for PostgreSQL v5.9.4: *“es un motor de bases de datos relacional de código abierto”* (Urtiaga, 2020). Con esta herramienta se pudo obtener la base de datos fuente de SIGERH, de la cual se obtienen los campos para cada una de las dimensiones del Mercado de Datos.

PDI (Pentaho Data Integration) v4.2: *“es una de las herramientas de Pentaho anteriormente se llamaba Kettle y posee un servicio de integración de datos ETL (extract, transform, load), es decir, con esta herramienta se pueden extraer, transformar y cargar datos. Asimismo, se pueden diseñar flujos de datos que se ejecutan en un servidor o en procesos independientes.”* (KeepCoding, 2022). Esta herramienta fue utilizada para la realización de los procesos ETL de cada una de las dimensiones y hecho del Mercado de Datos.

PgAdmin III v1.20.0: *“es una herramienta indispensable de código abierto para gestionar y administrar PostgreSQL, es el desarrollo por una comunidad de expertos PostgreSQL en todo el mundo y está disponible en más de una docena de idiomas.”* (PostgreSQL, 2018)

I.4.3.1 Herramienta CASE

Para el modelado con UML se utilizará **Visual Paradigm for UML v8.0** por ser una herramienta UML profesional que soporta el ciclo de vida completo del desarrollo de software: análisis y diseño orientados a objetos, construcción, pruebas y despliegue. Esta herramienta nos permite modelar los artefactos arrojados por la metodología empleada como son: el modelo conceptual, lógico, físico.

Visual Paradigm *“es una de las herramientas CASE más utilizadas para el modelado de software.”* (Quintero, 2012). *“Esta herramienta cuenta con los medios necesarios para extender sus funcionalidades, pues da soporte a las extensiones de aplicación. Provee de forma libre una interfaz de programación, que permite a los desarrolladores implementar y reutilizar clases e interfaces y desarrollar funciones agregadas que son útiles para el desarrollo de software.”* (Trujillo, 2013)

Conclusiones parciales

Al finalizar este capítulo se concluyó que:

- El estudio de los principales conceptos asociados al dominio de la investigación, permitió sentar las bases teóricas de la misma.
- Se identificaron muy pocos Mercado de Datos a los RRHH, los cuales, no satisfacen las necesidades del problema planteado.
- La implementación del Mercado de Datos se realizó con los lenguajes SQL v2019, XMLv1.1 y UMLv2.0 y las herramientas, PgAdmin IIIv9.4, PDIv4.2, Damix v1.1, SQL Manager for PostgreSQL 5.9.4, Visual Paradigm for UML v8.0 y como sistema gestor de base de datos PostgreSQL v9.4.

Capítulo II: Análisis y Diseño del Mercado de Datos

En el presente capítulo se describe el proceso de análisis y diseño del Mercado de Datos del proceso de selección del personal de los Recursos Humanos para la XETID siempre tomando en consideración las etapas y pasos que brinda la metodología de desarrollo escogida en el capítulo anterior. Se hará una descripción del negocio además se identifican las perspectivas e indicadores para el desarrollo de tablas. Se muestra la representación formal del modelo conceptual y el modelo conceptual ampliado. Se identifican las reglas del negocio y se realiza el modelo de datos según el estilo de diseño de Almacén de Datos.

II.1 Descripción del negocio

La Empresa de Tecnologías de la Información para la Defensa (XETID) es una empresa dedicada al sector del software, la automática y las comunicaciones. Para una mejor gestión de los proceso en el área de los RRHH se cuenta con un sistema llamado “Sistema de Gestión Integral de Recursos Humanos” (SIGERH). Este sistema cuenta con nueve módulos: estructura y comparación, registro de personas, gestión de movimientos laborales, registro y control de asistencia, evaluación, capacitación y desarrollo, idoneidad, seguridad y salud del trabajo y selección del personal. En este sistema se trabaja toda la información que se maneja en el departamento de los Recursos Humanos en tiempo real a través de una base de datos, pero en la actualidad no se tiene cómo medir el proceso de selección del personal, pues no se realizan análisis de este proceso por la forma en que está organizada la información. En este departamento no se tienen centralizado un proceso tan importante como es el proceso de selección del personal lo que hace que las consultas sean más complejas. Además no se puede acceder de forma inmediata a datos históricos de este proceso.

II.2 Fase 1: Análisis de los requerimientos

II.2.1 Identificación de preguntas

Después de planificarse las entrevistas pertinentes (**ver anexo 5**) se identificaron las preguntas para realizárselas al cliente para obtener información sobre lo que necesitan conocer acerca del proceso de selección del personal (**ver anexo 6**), donde se obtuvo como resultado de las misma

1. Analizar la cantidad de candidatos contratados para un proceso y periodo de determinado.
2. Realizar un análisis de la cantidad de candidatos entrevistados en un periodo de tiempo.
3. Analizar la cantidad de candidatos excluidos para un proceso y periodo determinado.
4. Mostrar la cantidad de candidatos que son entrevistados para un proceso en particular.
5. Analizar la cantidad de candidatos que optaron para un proceso determinado.
6. Realizar un análisis del porcentaje que representan los candidatos descalificados en un proceso determinado.
7. Analizar el porcentaje que representan los candidatos descalificados por un determinado motivo de exclusión.

II.2.2 Identificar indicadores y perspectivas.

Una vez que se han establecido las preguntas de negocio, se debe de proceder a su descomposición para descubrir los indicadores que se utilizan y las perspectivas de análisis que intervendrán. Para ello se debe tener en cuenta que:

Indicadores

Los indicadores son valores numéricos y representan lo que se desea analizar concretamente (Bernabeu, 2007)

Perspectivas

Las perspectivas son objetos mediante los cuales se quiere examinar los indicadores, con el objetivo de responder a las preguntas planteadas. (Bernabeu, 2007)

Los indicadores obtenidos son:

1. Cantidad de candidatos contratados
2. Cantidad de candidatos presentados
3. Cantidad de candidatos excluidos
4. Número de candidatos entrevistados
5. Porcentaje de candidatos descalificados

Las perspectivas obtenidas son:

- a) Persona

- b) Proceso
- c) Tiempo
- d) Motivo de exclusión
- e) Resultado de selección

En la tabla 2 hasta la tabla 6 se representa la “Ficha de indicador” para cada uno de los indicadores identificados,

Tabla 2 Ficha del indicador Cantidad de candidatos contratados

Ficha de Indicador	
Nombre	Cantidad de candidatos contratados(Cc)
Proceso	Selección del personal
Periodicidad	Mensualmente
Objetivo	Mide la cantidad de candidatos contratados.
Descripción	Mide la cantidad de personas que han sido contratadas por una empresa durante de un periodo de tiempo determinado.
Forma de Cálculo	COUNT
Fuente de Obtención	Base de datos PostgreSQL
Perspectivas de Análisis	Tiempo, persona, proceso, motivo de exclusión, resultado de selección

Tabla 3 Ficha del indicador Número de candidatos entrevistados

Ficha de Indicador	
Nombre	Número de candidatos entrevistados(Nce)
Proceso	Selección del personal
Periodicidad	Mensualmente
Objetivo	Medir la cantidad de candidatos que han sido evaluados durante el proceso de selección.
Descripción	Mide cuantos candidatos son entrevistados para cada puesto de vacante.
Forma de Cálculo	SUM
Fuente de Obtención	Base de datos PostgreSQL
Perspectivas de Análisis	Tiempo, persona, proceso, motivo de exclusión, resultado de selección

Tabla 4 Ficha del indicador Porcentaje de candidatos descalificados

Ficha de Indicador	
Nombre	Porcentaje de candidatos descalificados (Pcd)

Proceso	Selección del personal
Periodicidad	Mensualmente
Objetivo	Evaluar la calidad de la etapa de preselección y la claridad de los requisitos del puesto.
Descripción	Este indicador mide la proporción de candidatos que no cumplen con los requisitos mínimos para el puesto y son eliminados del proceso de selección.
Forma de Cálculo	$Pcd = (Cantidad\ de\ candidatos\ excluidos / Cantidad\ de\ candidatos\ presentados) * 100$
Fuente de Obtención	Base de datos PostgreSQL
Perspectivas de Análisis	Tiempo, persona, proceso, motivo de exclusión, resultado de selección

Tabla 5 Ficha del indicador Cantidad de candidatos presentados

Ficha de Indicador	
Nombre	Cantidad de candidatos presentados(Ccp)
Proceso	Selección del personal
Periodicidad	Mensualmente
Objetivo	Mide la cantidad de personas que fueron presentados para el puesto.
Descripción	Este indicador se refiere al número total de candidatos que se han presentado para una posición vacante en un determinado periodo de tiempo.
Forma de Cálculo	COUNT
Fuente de Obtención	Base de datos PostgreSQL
Perspectivas de Análisis	Tiempo, persona, proceso, motivo de exclusión, resultado de selección

Tabla 6 Ficha del indicador Cantidad de candidatos excluidos

Ficha de Indicador	
Nombre	Cantidad de candidatos excluidos(Cce)
Proceso	Selección del personal
Periodicidad	Mensualmente
Objetivo	Mide la cantidad de personas que fueron excluidos del proceso de selección.
Descripción	Este indicador se refiere a la cantidad total de candidatos que han sido excluidos de un proceso de selección.
Forma de Cálculo	COUNT
Fuente de Obtención	Base de datos PostgreSQL
Perspectivas de Análisis	Tiempo, persona, proceso, motivo de exclusión, resultado de selección

II.2.3 Construcción del modelo conceptual

Una vez identificados los indicadores y perspectivas se confeccionó el modelo conceptual elaborado en el Visual Paradigm.

Según la metodología empleada la representación gráfica del Modelo Conceptual (**ver Anexo 7**) es la siguiente: a la izquierda las perspectivas y a la derecha los indicadores. A continuación se muestra en la ilustración 2 el diseño del modelo conceptual de esta investigación.

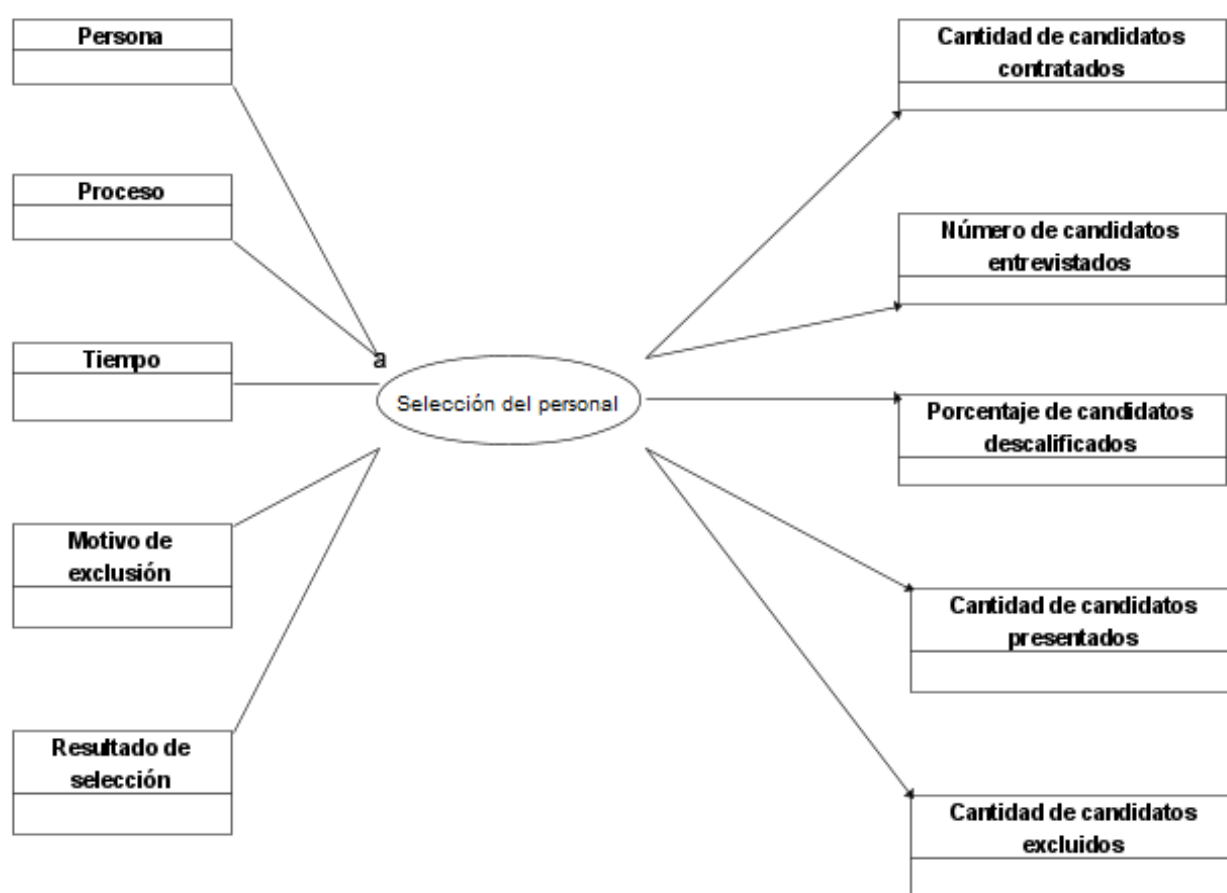


Ilustración 2 Modelo Conceptual.

II.2.4 Aprobación del modelo conceptual por el cliente

Para asegurar que el modelo conceptual se construyó correctamente y que reflejó lo que el cliente necesitaba, se aprobó el mismo en una reunión donde se expuso el Modelo Conceptual, utilizando siempre un lenguaje entendible por ambas partes.

II.3 Fase 2: Análisis de los OLTP

II.3.1 Determinación de indicadores

En este paso se debe explicar como se calculan los indicadores, definiendo los siguientes conceptos para cada uno de ellos:

- Hecho/s que lo componen, con su respectiva fórmula de cálculo. Por ejemplo: Hecho1 + Hecho2.
- Función de sumarización que se utilizará para su agregación. Por ejemplo: SUM, AVG, COUNT, etc.

Tabla 7 Especificación de indicadores

No.	Indicador	Hecho	Función de sumarización	Aclaración
1	Cantidad de candidatos contratados	Candidatos contratados	COUNT	El indicador cantidad de candidatos contratados representa la cantidad de personas que son contratadas en el proceso de selección de la XETID.
2	Cantidad de candidatos presentados	Candidatos presentados	COUNT	El indicador cantidad de candidatos presentados representa la cantidad de candidatos que son presentados para un puesto vacante en el proceso de selección de la XETID.
3	Cantidad de candidatos excluidos	Candidatos excluidos	COUNT	El indicador cantidad de candidatos excluidos representa la cantidad de candidatos que son excluidos del proceso de selección de la XETID.
4	Número de candidatos entrevistados	Candidatos entrevistados	SUM	El indicador número de candidatos entrevistados representa la sumatoria de los candidatos que son entrevistados en el proceso de selección de la XETID.
5	Porcentaje de candidatos descalificados	Candidatos descalificados	$Pcd = (cantidad \text{ de candidatos excluidos} / cantidad \text{ de candidatos presentados}) * 100$	El indicador porcentaje de candidatos descalificados representa el porcentaje de los candidatos que son excluidos del proceso de selección de la XETID

Fuente: Elaboración propia

II.3.2 Establecer correspondencia

El objetivo de este paso, es el de examinar los OLTP (Procesamiento de transacciones en línea) disponibles que contengan la información requerida, así

como también sus características, para poder identificar las correspondencias entre el modelo conceptual y la fuente de datos.

Las relaciones identificadas fueron las siguientes:

La tabla 'dat_persona' se relaciona con la perspectiva dimensión persona.

La tabla 'dat_proceso' se relaciona con la perspectiva dimensión proceso.

La tabla 'nom_motivoexclusion' se relaciona con la perspectiva dimensión motivo de exclusión.

La tabla 'dat_resultadoseleccion' se relaciona con la perspectiva dimensión resultado de seleccion.

Los campos 'desde' y 'hasta' de la tabla 'dat_proceso' se relacionan con la perspectiva dimensión tiempo.

II.3.3 Nivel de granularidad

Una vez establecidas las relaciones entre la base de datos fuente y el modelo conceptual definido y luego de determinar la forma en que se calculan los indicadores, se procede a definir los campos que contendrán cada perspectiva, ya que será a través de estos por donde se filtrarán y manipularán los diferentes indicadores. En la Tabla 8 se muestra el resultado obtenido:

Tabla 8 Nivel de granularidad

Tabla	Campo	Perspectiva
dat_persona	nombre, sexo, snombre, papel, spapel,	Persona
dat_contacto	contacto	Persona
dat_direccion	direccion	Persona
dat_candidato	contratado	Persona
nom_titulo	denominacion	Persona
dat_excluido	idexcluido	Persona
dat_proceso	denominacion	Proceso
nom_motivoseleccion	denominacion	motivo de selección
dat_resultadoseleccion	seleccion	Resultado de selección
dat_tiempo	año,día,mes,trimestre,semestre, semana,numdia,nummes,numsemestre, numtrimestre,fecha,numhora,numminuto	Tiempo

Fuente: Elaboración propia

II.3.4 Modelo conceptual ampliado

Con la utilización del método de modelación se realiza en el visual Paradigm el modelo conceptual. Para mostrar los resultados obtenidos en los pasos anteriores, se ampliará este modelo conceptual colocando debajo de cada perspectiva los

campos escogidos y debajo de cada indicador su fórmula de cálculo, mostrado en la ilustración 3.

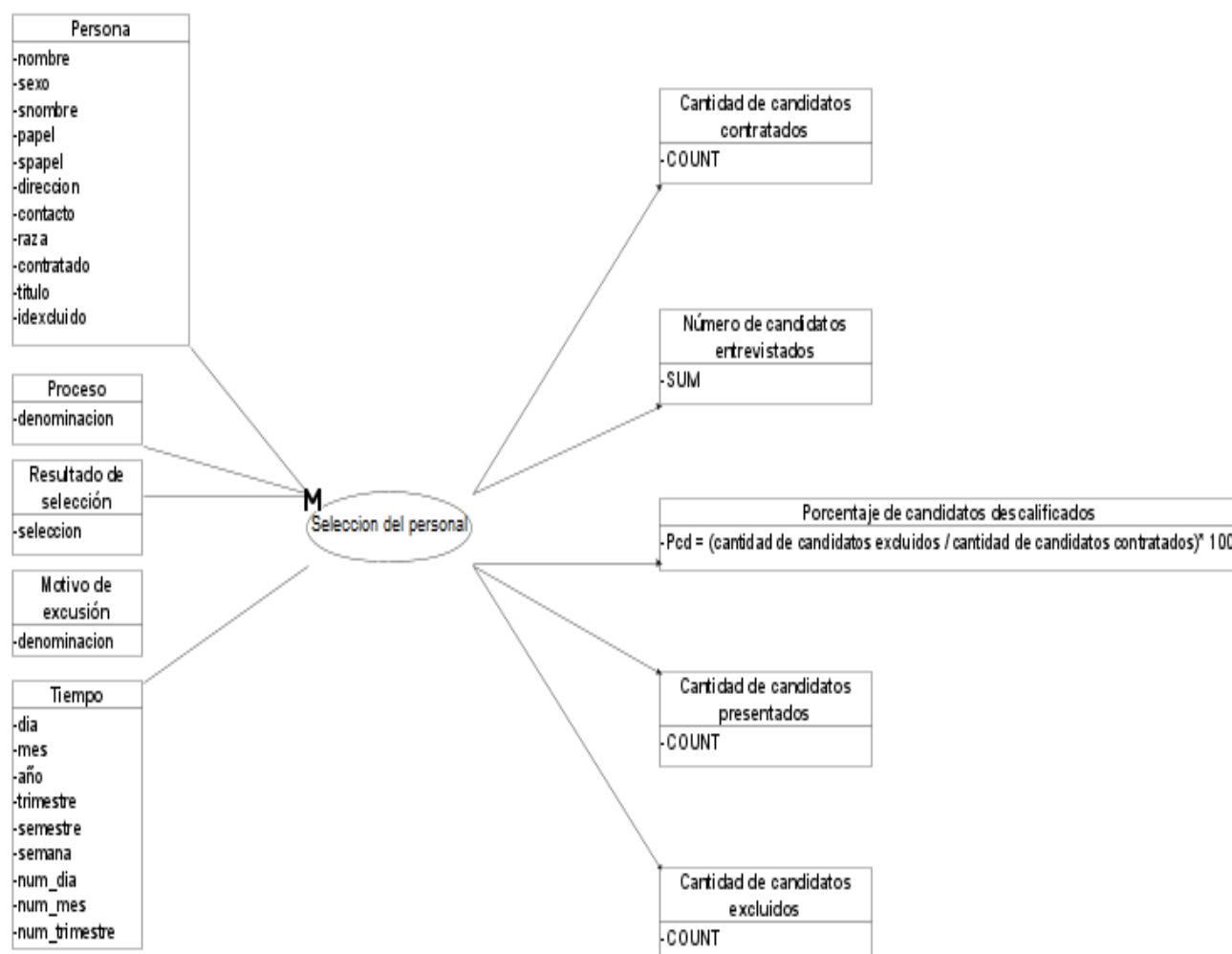


Ilustración 3 Modelo Conceptual Ampliado

A continuación se muestra una descripción detallada de cada una de las perspectivas y de cada campo mostrado en el modelo conceptual ampliado. (Ver tabla de la 9 y 10)

Tabla 9 Descripción de las perspectivas

No	Perspectiva	Descripción
1	Persona	Perspectiva donde se guarda la información correspondiente a cada trabajador.
2	Proceso	Perspectiva que guarda el nombre del proceso.
3	Motivo de exclusión	Perspectiva que se encarga de guardar el motivo por el cual un candidato es excluido del proceso.
4	Resultado de selección	Perspectiva que se encarga de guardar el resultado de la selección de las persona.
5	Tiempo	Perspectiva encargada de incorporar el desplazamiento del

		proceso en el tiempo.
--	--	-----------------------

Tabla 10 Descripción de los campos de las perspectivas

No	Perspectiva	Campo	Descripción
1	Persona	idpersona	Identificador de la persona
2	Persona	nombre	Nombre de la persona
3	Persona	sexo	Género de cada persona: 0->masculino, 1->femenino
4	Persona	snombre	Segundo nombre de la persona
5	Persona	Papel	Apellido de la persona
6	Persona	spapel	Segundo apellido de la persona
8	Persona	contacto	Contacto de la persona
9	Persona	direccion	Dirección de la persona con los valores calle, entre calle, apto etc.
10	Persona	titulo	Título de la persona
11	Persona	contratado	Denominación del candidato contratado 0->contratado, 1-> no contratado
12	Persona	idexcluido	Identificador del candidato que se encuentra excluido.
13	Proceso	idproceso	Identificador del proceso
14	Proceso	denominacion	Denominación del proceso de selección
15	Motivo de exclusión	idmotivoexclusion	Identificador del motivo de exclusión
16	Motivo de exclusión	denominacion	Denominación del motivo de exclusión
17	Resultado de selección	idresultadoseleccion	Identificador del resultado de selección
18	Resultado de selección	seleccion	Estado que se encuentra la persona en el proceso. 0->Preseleccionado, 1->Seleccionado, 2->Excluido
19	Tiempo	idtiempo	Identificador del tiempo
20	Tiempo	día	Denominación del día
21	Tiempo	semana	Número de la semana
22	Tiempo	mes	Denominación del mes
23	Tiempo	trimestre	Denominación del trimestre
24	Tiempo	semestre	Denominación del semestre
25	Tiempo	año	Número del año
26	Tiempo	numdia	Número del día
27	Tiempo	nummes	Número del mes
28	Tiempo	numtrimestre	Número del trimestre
29	Tiempo	numsemestre	Número del semestre
30	Tiempo	fecha	Fecha en el formato dd/mm/aaaa
31	Tiempo	numhora	Número de la hora
32	Tiempo	numminuto	Número del minuto

II.3.5 Definir las reglas del negocio

Reglas del negocio:

1. Los identificadores de cada dimensión deben de ser únicos.

2. Los identificadores de cada dimensión no pueden ser nulos.
3. Los valores que indiquen cantidad tiene que ser mayores o iguales a cero.

II.4 Fase 3: Modelo lógico de DW

II.4.1: Definir tipo de Modelo Lógico del DW

El **Esquema en Estrella** fue el estilo de diseño de Mercado de Datos que se utilizó porque se tiene un solo hecho y varias dimensiones que se relacionan con el hecho.

Hecho: selección del personal.

Dimensiones: persona, tiempo, proceso, motivo de exclusión, resultado de selección.

Indicadores: cantidad de candidatos contratados, cantidad de candidatos presentados, cantidad de candidatos excluidos, número de candidatos entrevistados, porcentaje de candidatos descalificados.

Características del Esquema en Estrella (Acero Calizaya, 2014)

- 1 tabla de hechos con N dimensiones relacionadas
- 100% desnormalizado
- Las dimensiones solo tienen clave primaria
- La tabla de hechos tiene llaves foráneas
- Esquema con mejor rendimiento y velocidad

II.4.2 Identificar Dimensiones

En este paso se deben diseñar las tablas de dimensiones que formarán parte del Mercado de Datos, cada perspectiva definida en el modelo conceptual constituirá una tabla de dimensión.

Se define como estándar de codificación para las dimensiones la estructura "dim_nombre", y para los campos con más de una palabra estas se unen.

Tabla 11 Dimensiones del Mercado de Datos

Dimensión	Campos
dat_persona	idpersona, nombre, sexo, direccion, contacto, titulo, contratado, idexcluido
dim_proceso	idproceso, denominacion

dim_motivoexclusion	idmotivoexclusion, denominacion
dim_resultadoseleccion	idresultadoseleccion, selecci3n
id_tiempo	idtiempo, a3o, dia, mes, trimestre, semestre, semana, numdia, nummes, numsemestre, numtrimestre, fecha, numhora, numminuto

Fuente: Elaboraci3n propia

II.4.3 Identificar Hecho

En este paso se define la tabla de hecho, que es la que contendr3 los indicadores de estudio.

Se define como est3ndar de codificaci3n para la tabla de hecho la estructura "cub_nombre".

Tabla 12 Hecho del Mercado de Datos

Hecho	Identificadores y llaves
cub_seleccionpersonal	idpersona, idproceso, idmotivoexclusion, idresultadoseleccion, idtiempo, cantidadcandidatoscontratados, candidatosexcluidos, numerocandidatosentrevistados, ,

Fuente: Elaboraci3n propia

II.4.4 Realizar uni3n entre dimensi3n y hecho.

Se realizan las uniones correspondientes entre las tablas de dimensiones y la tabla de hecho mostrando de forma detallada las relaciones entre ellas, se desarroll3 en la herramienta CASE Visual Paradigm for UML 8.0. Con esta uni3n se obtiene el modelo l3gico del Mercado de Datos que es mostrado a continuaci3n:

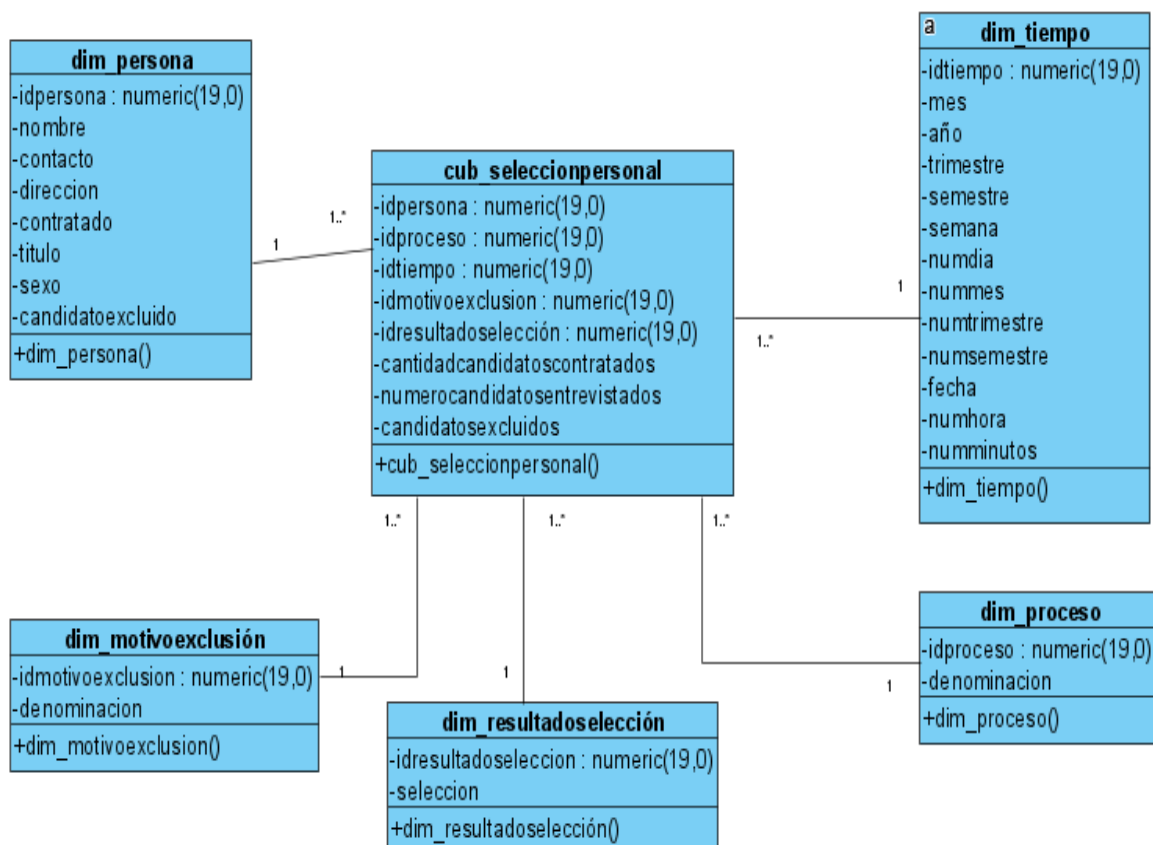


Ilustración 4 Unión de las dimensiones con el hecho

Después de obtener el modelo lógico del Mercado de Datos, se construye el modelo físico del mismo, en donde se especifican los tipos de datos de los diferentes campos. A continuación, se muestran el modelo físico obtenido.

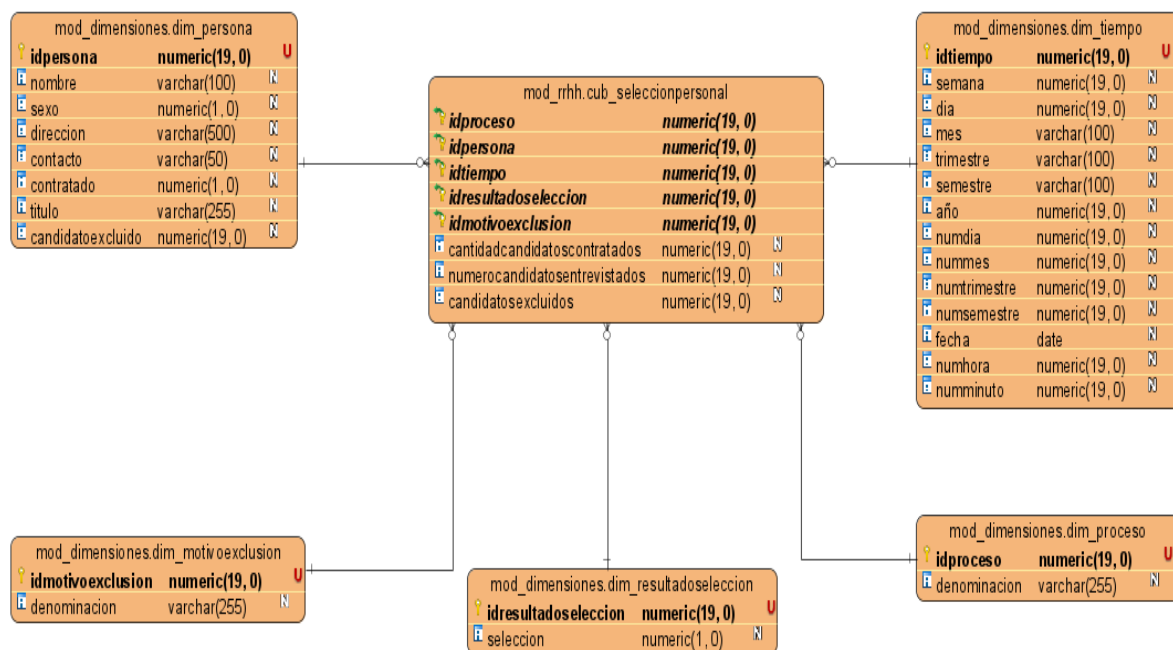


Ilustración 5 Modelo de dato físico

II.4.5 Diseñar tablas y columnas físicas

En este paso quedan diseñadas e implementadas físicamente la tabla de hecho y sus dimensiones en el Mercado de Datos según el modelo lógico.

El Mercado de Datos tiene como nombre **DwhtesisSigert**, donde se creó dos nuevos esquemas, el primero mod_dimensiones para las dimensiones y el segundo mod_rrhh para el cubo de selección del personal.

Estándares de codificación para las Dimensiones

- Las columnas si son los identificadores se nombra “iddimensión”, ejemplo idpersona, si no es un identificador se nombra “nombredelacolumna”, ejemplo dirección.
- Las llaves primarias fueron denominadas de la forma “pk_iddimensión”, ejemplo pk_idpersona.

Estándares de codificación para el Hecho

- Las columnas si son los identificadores se nombra “iddimensión”, ejemplo idpersona.
- Las medidas se escriben sin dejar espacio “nombredelamedida”, ejemplo cantidadcandidatoscontratados.
- La llave primaria fue denominada de la forma “pk_nombredelcubo_pkey”, ejemplo pk_seleccionpersonal_pkey.
- Las llaves foráneas fueron denominadas de la forma “fk_nombredeladimensión_nombredelcubo”, ejemplo fk_persona_seleccionpersonal.

A continuación se muestra la implementación de las dimensiones y el hecho con sus respectivos campos desarrollados con la herramienta PgAdmin III:

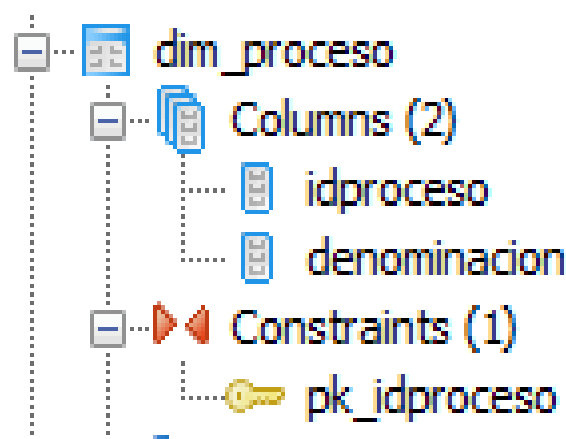


Ilustración 6 Columnas y llave primaria dimensión proceso

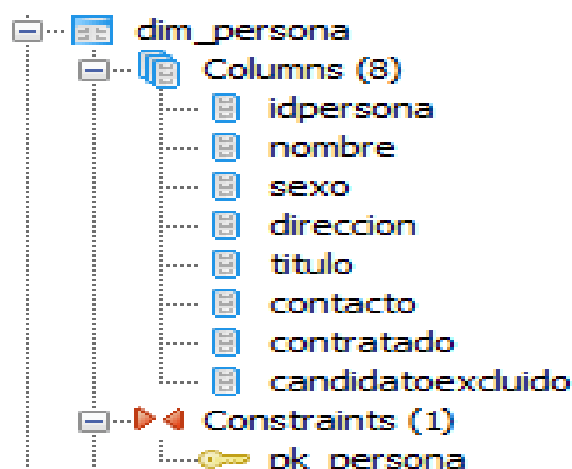


Ilustración 7 Columnas y llave primaria dimensión persona

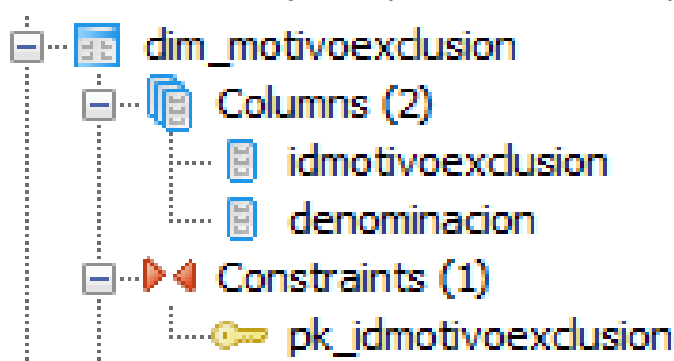


Ilustración 8 Columnas y llave primaria dimensión motivo de exclusión

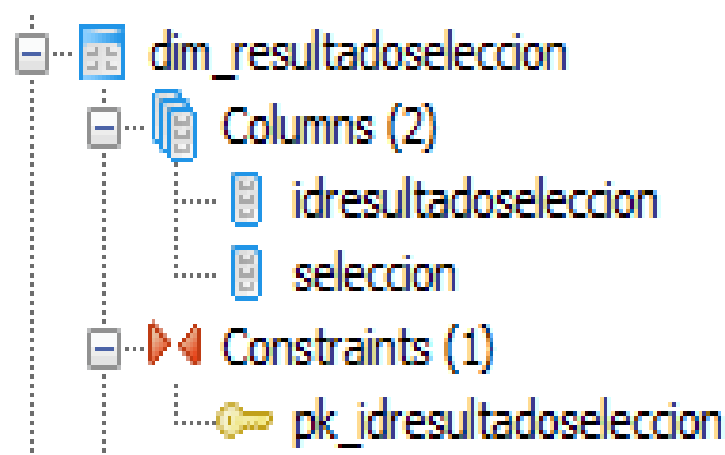


Ilustración 9 Columnas y llave primaria dimensión resultado de selección

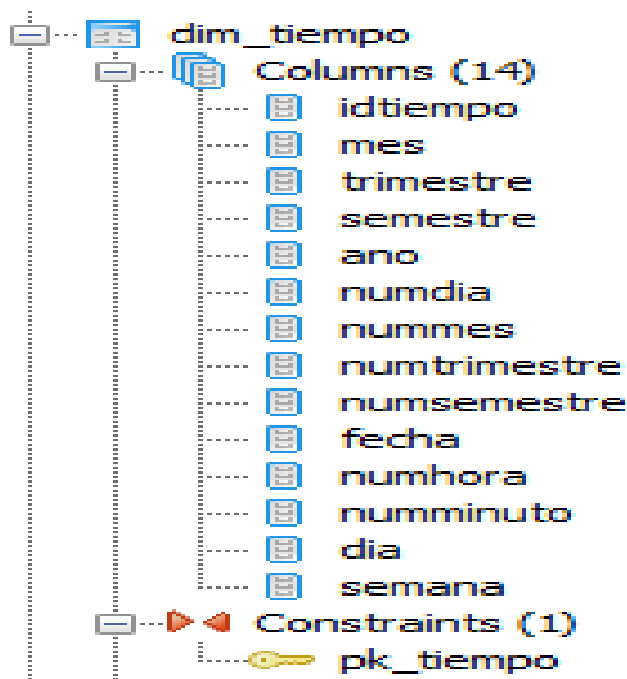


Ilustración 10 Columnas y llave primaria dimensión tiempo

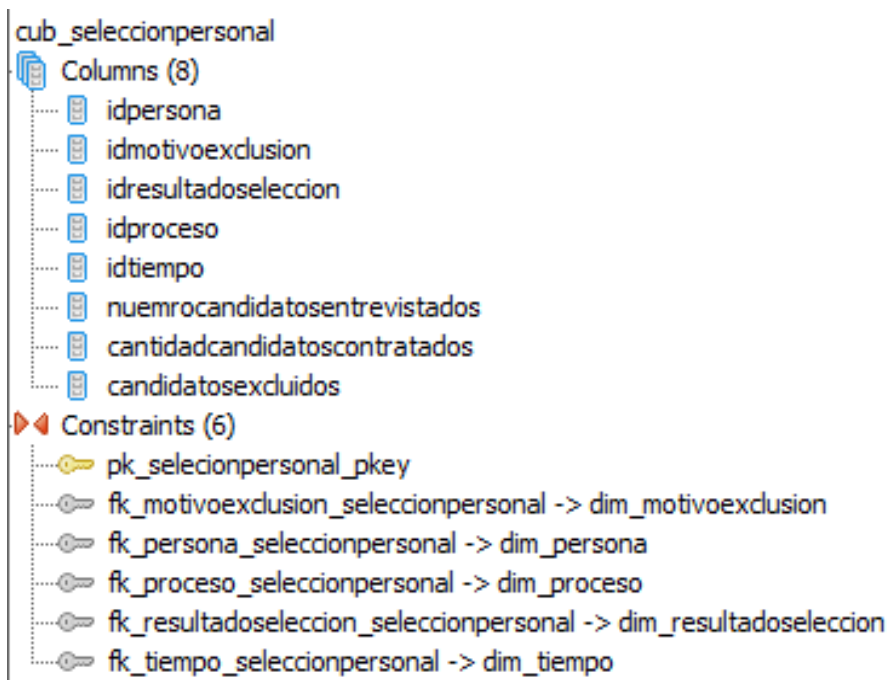


Ilustración 11 Columnas, llave primaria y llaves foráneas del cubo selección del personal

Conclusiones parciales

En este capítulo se finalizó la etapa de análisis y diseño del Mercado de Datos que va a dar solución a las interrogantes actuales del cliente, donde se concluyó que:

- El empleo de la metodología permitió organizar el desarrollo de la solución y generar los artefactos de trabajo necesarios correspondientes a la etapa de análisis y diseño.

- Se logró identificar los indicadores y perspectivas del análisis, lo que permitió elaborar el modelo conceptual de datos.
- Se representa la relación entre la tabla de hecho, las dimensiones y los indicadores identificados para el Mercado de Datos a partir del modelo dimensional.

CAPÍTULO III IMPLEMENTACIÓN Y VALIDACIÓN DEL MERCADO DE DATOS

En este capítulo se aborda todo lo referente al proceso de implementación del Mercado de Datos del proceso de selección del personal de los RRHH para la XETID, dando cumplimiento a la cuarta fase de la metodología con los procesos ETL. Se finaliza con la realización de pruebas para validar que se de solución a las necesidades del cliente.

III.1 Fase 4: Procesos ETL

III.1.1 Mapeo de los datos Fuentes-A-Destino.

Tabla 13 Mapa Lógico de Datos

Mapa Lógico de datos					
Nombre DB fuente: Sigerh			Nombre DB destino: DwhTesisSiger		
Nombre de la tabla	Nombre del campo	Tipo de dato	Nombre de la tabla	Nombre del campo	Tipo de dato
dat_persona	idpersona	numeric	mod_dimensiones.dim_persona	idpersona	numeric
dat_persona	nombre	varchar	mod_dimensiones.dim_persona	nombre	varchar
dat_persona	snombre	varchar	mod_dimensiones.dim_persona	nombre	varchar
dat_persona	papel	varchar	mod_dimensiones.dim_persona	nombre	varchar
dat_persona	spapel	varchar	mod_dimensiones.dim_persona	nombre	varchar
dat_persona	sexo	numeric	mod_dimensiones.dim_persona	sexo	numeric
dat_direccion	direccion	varchar	mod_dimensiones.dim_persona	direccion	varchar
dat_persona	contacto	numeric	mod_dimensiones.dim_persona	contacto	numeric
nom_titulo	denominacion	varchar	mod_dimensiones.dim_persona	titulo	varchar
dat_candidato	contratado	numeric	mod_dimensiones.dim_persona	contratado	numeric
dat_excluido	idexcluido	numeric	mod_dimensiones.dim_persona	candidatoexcluido	numeric
dat_proceso	idproceso	numeric	mod_dimensiones.dim_proceso	idproceso	numeric
dat_proceso	denominacion	varchar	mod_dimensiones.dim_proceso	denominacion	varchar
nom_motivoexclusión	idmotivoexclusion	numeric	mod_dimensiones.dim_motivodeexclusión	idmotivoexclusion	numeric

nom_motivoex clusión	denominacion	varchar	mod_dimensiones. dim_motivodeexcl usión	denominacion	varchar
dat_resultados elección	idresultadoselec cion	numeric	mod_dimensiones. dim_resultadosese lección	idresultadoselec cion	numeric
dat_resultados elección	seleccion	numeric	mod_dimensiones. dim_resultadosese lección	seleccion	numeric

Funete: Elaboración propia

III.1.2 Cargas incrementales de los datos

En este paso se realiza la carga de los datos almacenados en la fuente de datos para ser guardados en las dimensiones y hecho correspondientes. Para esto se llevó a cabo una serie de transformaciones y trabajos con la herramienta Pentaho Data Integration. A continuación se muestran cada una de las ETL realizadas a las dimensiones y al hecho y sus respectivas consultas:

```
SELECT id.proceso, p.denominacion
FROM mod_seleccion.dat_proceso p
```

Ilustración 12 Consulta de la Dimensión Proceso

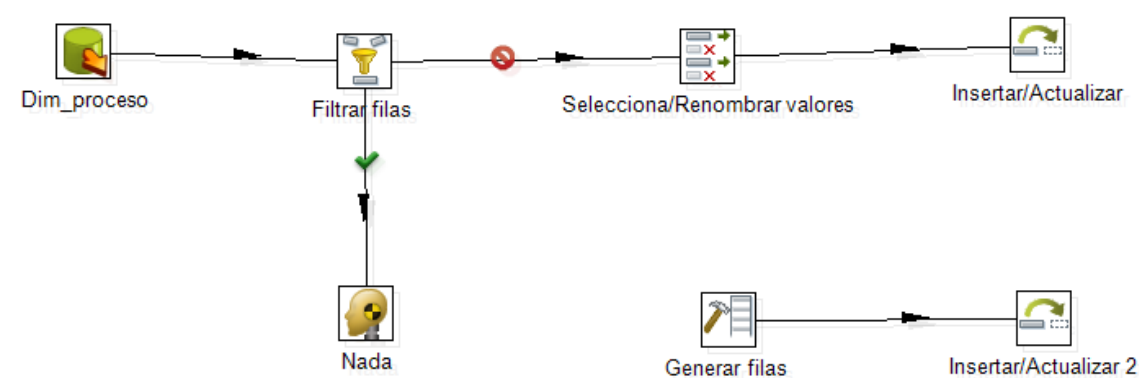


Ilustración 13 ETL correspondiente a la Dimensión Proceso

```
SELECT id.motivoexclusion, me.denominacion
FROM mod_seleccion.nom_motivoexclusion me
```

Ilustración 14 Consulta de la Dimensión Motivo de exclusión

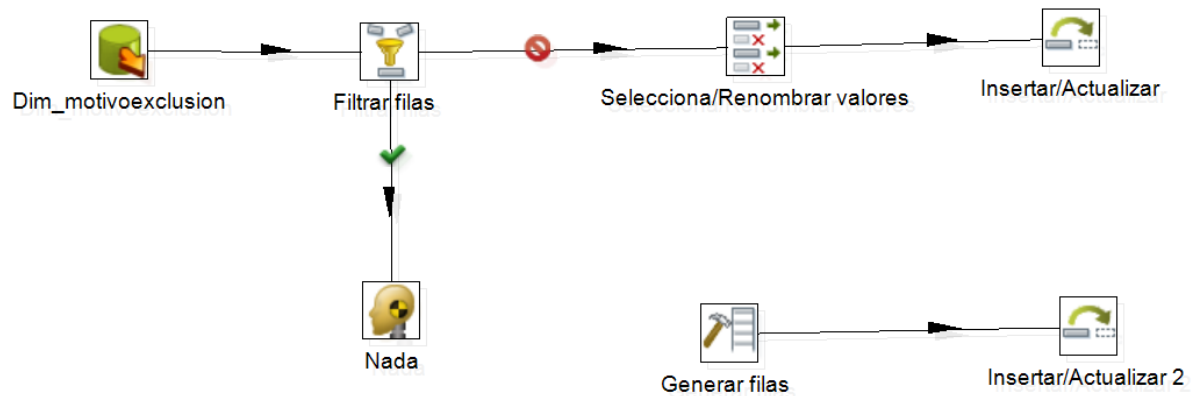


Ilustración 15 ETL correspondiente a la Dimensión Motivo de exclusión

```
SELECT re.id.resultadoseleccion, re.seleccion
FROM mod_seleccion.dat_resultadoseleccion re
where re.seleccion=0
```

Ilustración 16 Consulta de la Dimensión Resultado de selección

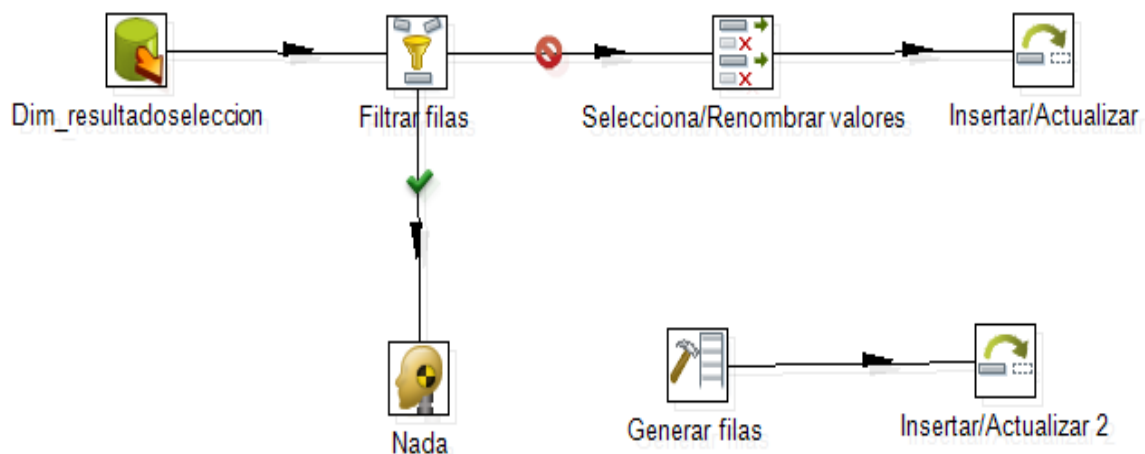


Ilustración 17 ETL correspondiente a la Dimensión Resultado de selección

```

SELECT distinct
    pe.idpersona,
    case when p.snombre is null and p.sapel is not null then concat (p.nombre||' '||p.papel||' '||p.sapel)
        when p.snombre is null and p.sapel is null then concat (p.nombre||' '||p.papel)
        when p.snombre is not null and p.sapel is null then concat (p.nombre||' '||p.snombre||' '||p.papel)
    else concat (p.nombre||' '||p.snombre||' '||p.papel||' '||p.sapel) end as nombre,
    pe.sexo,
    d.direccion,
    t.denominacion as titulo,
    c.contacto,
    sc.contratado,
    se.idexcluido as candidatoexcluido
FROM
    mod_persona.dat_persona pe
    left join mod_persona.dat_datosnacionales dn on(pe.idpersona=dn.idpersona)
    left join mod_persona.dat_direccion d on(pe.idpersona=d.idpersona)
    left join mod_persona.dat_contacto c on(pe.idpersona=c.idpersona)
    left join mod_personaext.dat_niveleducacional ne on( pe.idpersona=ne.idpersona)
    left join mod_personaext.dat_titulo t on(pe.idtitulo=t.idtitulo)
    left join mod_seleccion.dat_candidato sc on(pe.idpersona=sc.idpersona)
    left join mod_seleccion.dat_excluido se on(se.idcadidato=sc.idcandidato)

```

Ilustración 18 Consulta de la Dimensión Persona

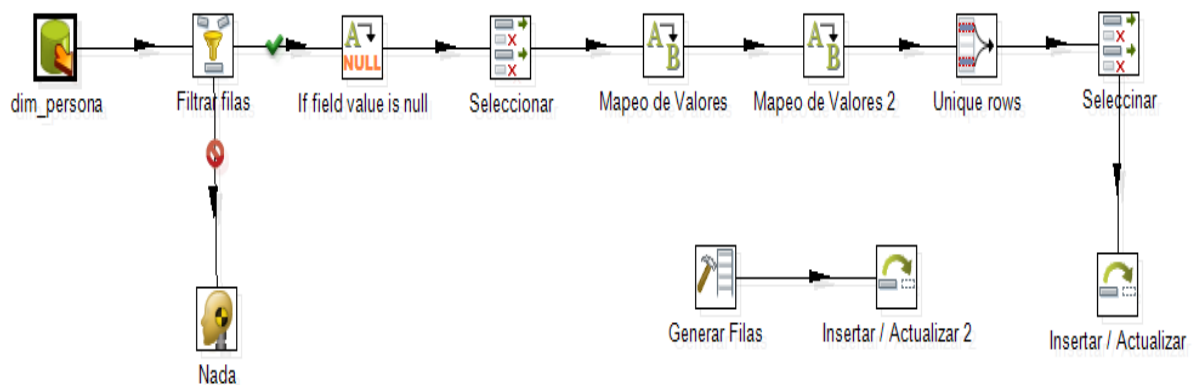


Ilustración 19 ETL correspondiente a la Dimensión Persona

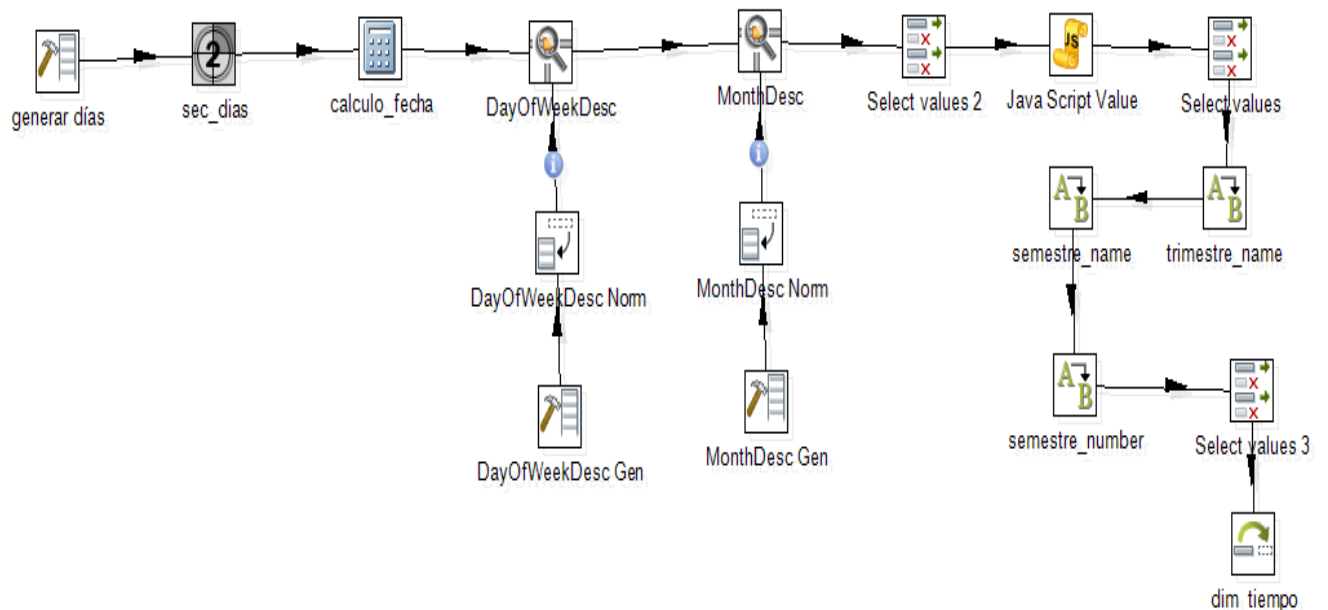


Ilustración 20 ETL correspondiente a la Dimensión Tiempo

```

SELECT
p.idproceso as procesoid,
sc.idpersona as idacandidato,
rs.idresultadoseleccion as resultadoseleccionid,
se.idmotivoexclusion as motivoexclusionid,
pe.id.persona as persona,
    case when snombre is null then concat (nombre||' '||papel||' '||sapel)
    else concat (nombre||' '||snombre||' '||papel||' '||sapel) end as nombrecandidato,
pe.sexo,
d.direccion,
c.contacto,
t.denominacion as titulo,
p.denominacion as proceso,
p.fdesde as fechainicioproceso,
p.fhasta as fechafinproceso,
n.rs.seleccion,
me.denominacion as motivoexclusion,
se.idexcluido as candidatoexcluido,
sc.contratado
FROM mod_seleccion.dat_proceso p
left join mod_seleccion.dat_candidato sc on(sc.idproceso= p.idproceso)
left join mod_seleccion.dat_resultadoseleccion rs on(rs.idcandidato=sc.idcandidato)
left join mod_seleccion.dat_excluido se on(se.idcandidato=rs.idcandidato)
left join mod_seleccion.dat_motivoexclusion me on(me.idmotivoexclusion=se.motivoexclusion)
left join mod_persona.dat_persona pe on(pe.idpersona=sc.idpersona)
left join mod_persona.dat_datosnacionales dn on(dn.idpersona=pe.idpersona)
left join mod_persona.dat_direccion d on(d.idpersona=dn.idpersona)
left join mod_persona.dat_contacto c on(dn.idpersona=c.idpersona)
left join mod_personaext.dat_niveleducacional ne on( ne.idpersona=pe.idpersona)
left join mod_personaext.dat_titulo t on(t.idtitulo=ne.idtitulo)
where pe.idpersona= sc.idpersona

```

Ilustración 21 Consulta del cubo selección del personal

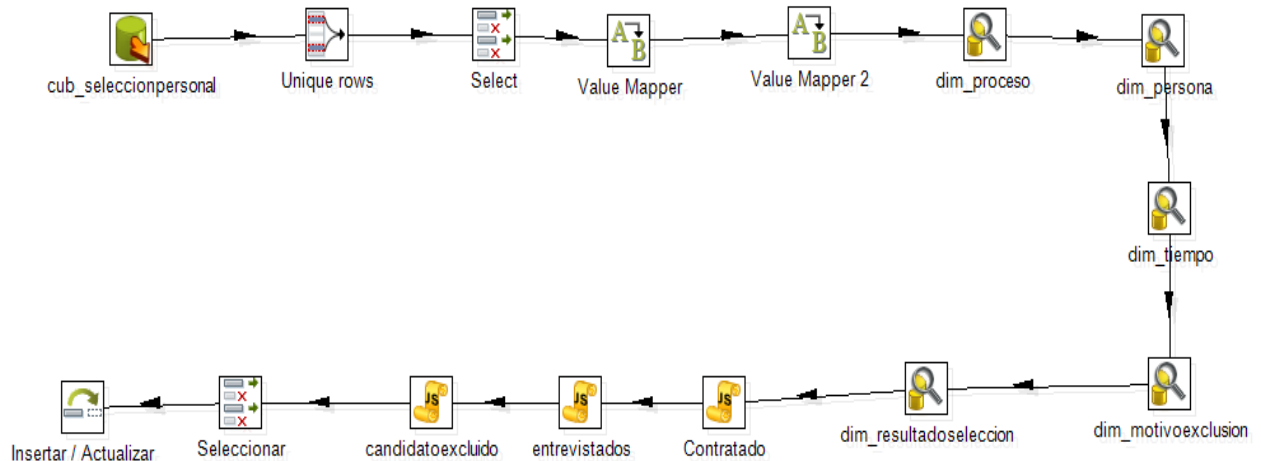


Ilustración 22 ETL correspondiente a la Hecho Selección del personal

III.1.3 Implementación de los indicadores

Se utiliza la herramienta Pentaho Data Integration v4.2, para realizar la ETL correspondiente al cubo selección del personal, representada anteriormente, donde se llevó a cabo la implementación de los indicadores a través de Modified Java Script Value, al cual se le denominó “candidato”, “entrevistado”, “candidatoexcluido”. A continuación en las ilustraciones 23, 24 y 25 se muestra el cálculo de los indicadores.

Step name

Java script:

Script 1 ✕

```
//Script here

var cantidadcandidatoscontratados=0;
if(contratado=='Contratado'){
cantidadcandidatoscontratados=1;
}
```

Ilustración 23 Indicador cantidad de candidatos contratados

Step name

Java script:

Script 1 ✕

```
//Script here

var numerocandidatosentrevistados=0;
if(seleccion=0){
numerocandidatosentrevistados=1;
}
```

Ilustración 24 Indicador número de candidatos entrevistados

Step name

Java script:

Script 1 ✕

```
//Script here

var cant=0;
if(candidatoexcluido!=0){
cant=1;
}
```

Ilustración 25 Indicador cantidad de candidatos excluidos

III.2 Implementación de los XML

Un archivo XML es un formato de archivo utilizado para almacenar y transmitir información estructurada en forma de etiqueta. Este tipo de archivos es de gran utilidad para el intercambio de datos entre diferentes sistemas y aplicaciones.

El archivo XML es una forma de mapear el Mercado de Datos (DwhtesisSigert), es decir, que las tecnologías que tiene por dentro Damix en el componente de análisis sean capaces de entender las indicaciones que se le especifican, ejemplo la dimensión persona sepa que es dim_persona y que se encuentra en el esquema mod_dimensiones, en la tabla dim_persona y que posee los campos idpersona como llave primaria y nombre, dirección, sexo, título, contacto, contratado, candidatoexcluido.

A continuación, se representa en las ilustraciones 26 y 27 un fragmento del archivo XML correspondiente a las dimensiones y el cubo respectivamente.

```
- <Dimension name="Proceso" highCardinality="false" visible="true" type="StandardDimension">
  - <Hierarchy visible="true" primaryKey="idproceso" allLevelName="Todos" allMemberName="Todos los procesos" hasAll="true">
    <Table name="dim_proceso" schema="mod_dimensiones"> </Table>
    <Level name="Proceso" visible="true" type="String" hideMemberIf="Never" levelType="Regular" uniqueMembers="true" column="denominacion"> </Level>
  </Hierarchy>
</Dimension>
- <Dimension name="Motivos de exclusión" highCardinality="false" visible="true" type="StandardDimension">
  - <Hierarchy visible="true" primaryKey="idmotivoexclusion" allLevelName="Todos" allMemberName="Todos los motivos de exclusión" hasAll="true">
    <Table name="dim_motivoexclusion" schema="mod_dimensiones"> </Table>
    <Level name="Motivo de exclusión" visible="true" type="String" hideMemberIf="Never" levelType="Regular" uniqueMembers="true" column="denominacion"> </Level>
  </Hierarchy>
</Dimension>
- <Dimension name="Resultado de selección" highCardinality="false" visible="true" type="StandardDimension">
  - <Hierarchy visible="true" primaryKey="idresultadoseleccion" allLevelName="Todos" allMemberName="Todos los resultados de selección" hasAll="true">
    <Table name="dim_resultadoseleccion" schema="mod_dimensiones"> </Table>
    <Level name="Resultado de selección" visible="true" type="String" hideMemberIf="Never" levelType="Regular" uniqueMembers="true" column="seleccion"> </Level>
  </Hierarchy>
</Dimension>
```

Ilustración 26 Fragmento del archivo XML correspondiente las dimensiones

```

- <Cube name="Selección Personal" visible="true" enabled="true" cache="true">
  <Table name="cub_seleccionpersonal" schema="mod_rrhh"> </Table>
  <DimensionUsage name="Fecha de Inicio del proceso" highCardinality="false" visible="true" foreignKey="idtiempo" source="Fecha de Inicio del proceso"> </DimensionUsage>
  <DimensionUsage name="Proceso" highCardinality="false" visible="true" foreignKey="idproceso" source="Proceso"> </DimensionUsage>
  <DimensionUsage name="Motivos de exclusión" highCardinality="false" visible="true" foreignKey="idmotivoexclusion" source="Motivos de exclusión"> </DimensionUsage>
  <DimensionUsage name="Resultado de selección" highCardinality="false" visible="true" foreignKey="idresultadoseleccion" source="Resultado de selección"> </DimensionUsage>
  <DimensionUsage name="Persona" highCardinality="false" visible="true" foreignKey="idpersona" source="Persona"> </DimensionUsage>
  <Measure name="Cantidad de candidatos contratados" visible="true" column="cantidadcandidatoscontratados" aggregator="sum" formatString="#,####"> </Measure>
  <Measure name="Numero de candidatos entrevistados" visible="true" column="numerocandidatosentrevistados" aggregator="sum" formatString="#,####"> </Measure>
  <Measure name="Cantidad de candidatos excluidos" visible="true" column="candidatoexcluido" aggregator="sum" formatString="#,####"> </Measure>
  <Measure name="Cantidad de candidatos presentados" visible="true" column="idpersona" aggregator="distinct count" formatString="#,####"> </Measure>
  <CalculatedMember name="Porcentaje de candidatos descalificados" visible="true" formatString="0.00%" dimension="Measures" formula="CASE WHEN [Measures].[Cantidad de
    candidatos excluidos] = 0 THEN 0 ELSE [Measures].[Cantidad de candidatos excluidos]*100/[Measures].[Cantidad de candidatos presentados]/ 100 END">
  </CalculatedMember>
</Cube>

```

Ilustración 27 Fragmento del archivo XML correspondiente al cubo

III.2.1 Plataforma de Inteligencia de Negocio

La herramienta que utiliza el sistema SIGERH para la visualización de la información es la Plataforma de Inteligencia de Negocios Damix. Mediante la url <https://damix-trunk.prod.xetid.cu/portal/> se da inicio a dicha plataforma. Damix cuenta con varios componentes que serán explicados a continuación.

III.2.1.1 Componente de Configuración

En el componente de configuración es el encargado de cargar el archivo XML. A continuación se definen los pasos a seguir:

Pasos para cargar el archivo XML:

1. Ir a configuración.
2. Fuente de datos.
3. Adicionar.
4. En opción: escribir importar.
5. En denominación: poner el nombre del Mercado de Datos.
6. En conexión: seleccionar donde se encuentra la base de datos.
7. En archivo: seleccionar donde se encuentra el archivo XML.
8. Aplicar.

The screenshot shows the top navigation bar of the Damix application with the following items: Inicio, Configuración, Análisis, Tablero de control, Mapa estratégico, and Notificación y aviso. The user is logged in as 'administrador'. Below the navigation bar, the 'Importar' form is displayed. It includes a dropdown menu for 'Opción:' with 'Importar' selected. There are two input fields: 'Denominación:' and 'Conexión:' (with a dropdown menu showing 'Seleccione...'). Below these is an 'Archivo:' field with '*.xml' entered and an upload icon. At the bottom of the form are three buttons: 'Cancelar', 'Aplicar', and 'Aceptar'.

Ilustración 28 Adicionar una fuente de datos

III.2.1.2 Componente de Análisis

Después de haber cargado el archivo XML se accede al componente de análisis, en este se selecciona el cubo a utilizar. Al cargar el cubo se crean las dimensiones y medidas del mismo. Este componente permite realizar análisis y mostrar resultados de los indicadores y las perspectivas, con el objetivo de extraer información relevante para la toma de decisiones.

Pasos para crear un análisis:

1. Abrir un nuevo análisis.
2. Se selecciona el cubo Selección del Personal.
3. Se selecciona las dimensiones y medidas deseadas.
4. Se guarda el análisis.
5. En categorías y en denominación: poner los nombres deseados.
6. Aceptar.

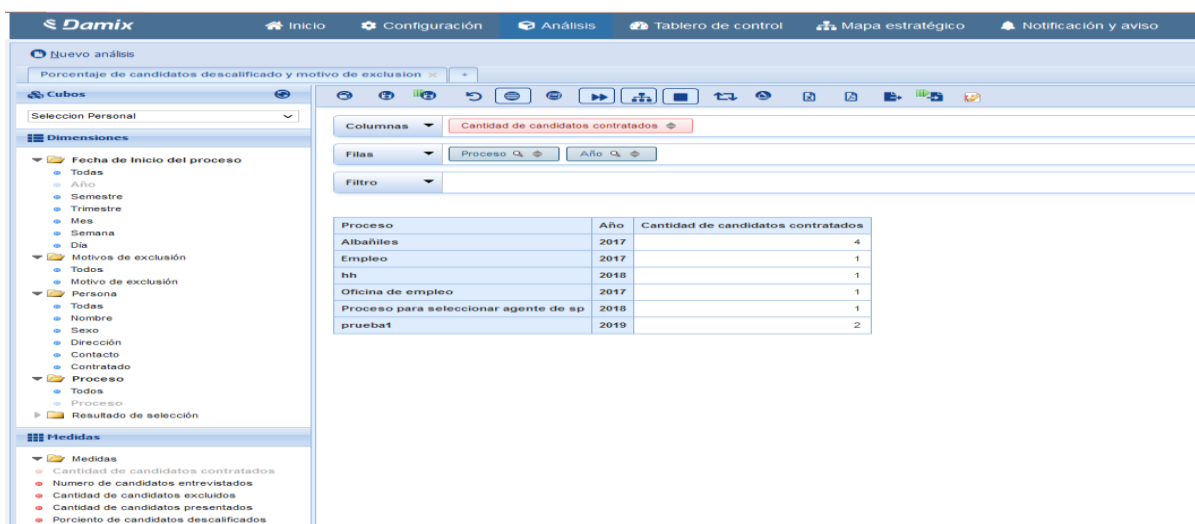


Ilustración 29 Componente de análisis

III.2.1.3 Componente de Tablero de control

Al concluir con los análisis se accede al componente de tableros de control, el cual permite realizar el diseño y visualización de dinámicos y vistosos tableros de control a través de tablas, gráficas y mapas.

Pasos para crear los tableros de control:

1. Actualizar.
2. Nuevo tablero de control.
3. Seleccionar en el menú: (tablas, gráficas (barra, línea, pastel, pirámide, reloj y radar) o mapas).
4. Arrastrar la selección al área.
5. En configuración escoger el nombre de análisis creado.
6. Opcionalmente dar clic en los tres puntos que aparecen en la parte superior derecha para darle estilo a la opción que se seleccionó en menú.
7. Guardar.

A continuación se muestran algunos de los tableros de control realizados para visualizar la información que necesita el cliente:

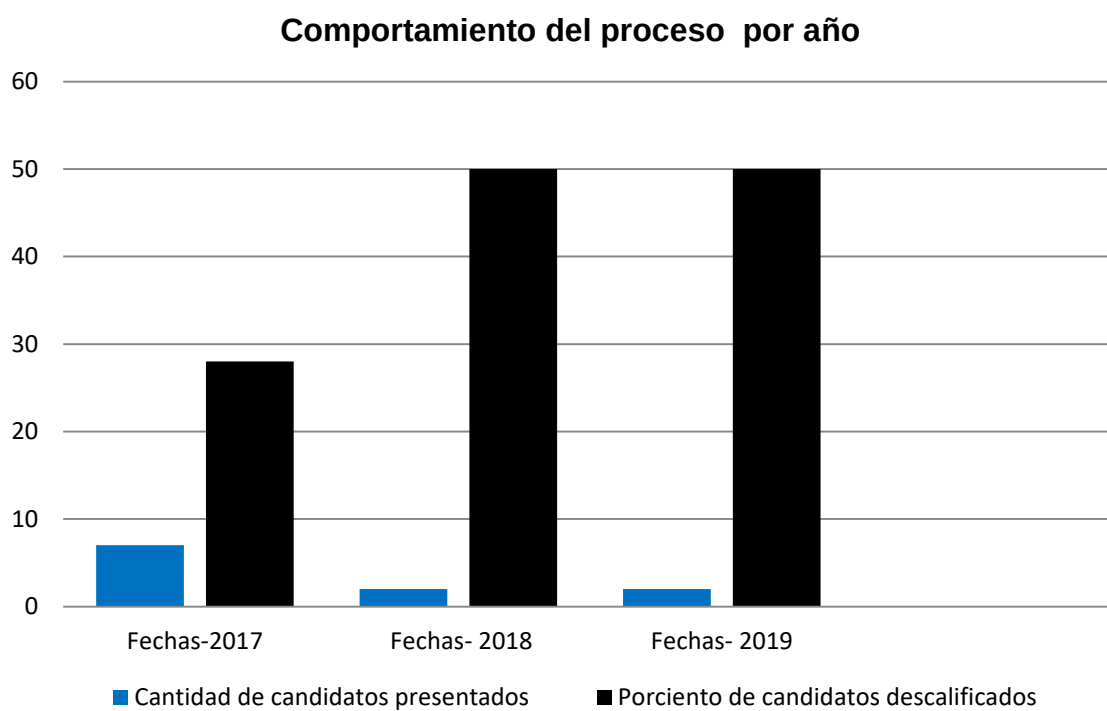


Ilustración 30 Tablero 1. Candidatos presentados y porciento de descalificados

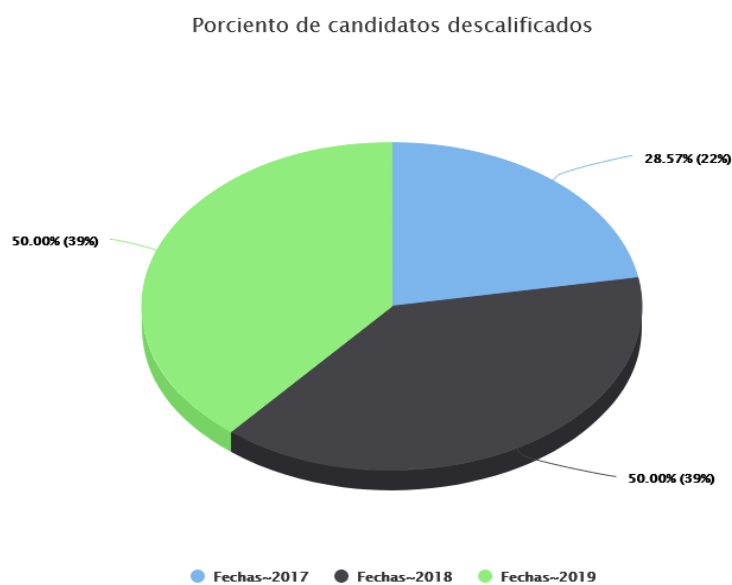


Ilustración 31 Tablero 2 Porcentaje de candidatos descalificados

candidato presentado por proceso y su nombre

		Cantidad de candidatos presentados
Albañiles	Adiel Morera Macias	1
	Lourdes Davila Carlos	1
	Mabel Brito Bartumeu	1
	Osvel Rivero López	1
	Yunaisy López Carrillo	1
Empleo	Rolando González Rodríguez	1
hh	Leonardo Alejandrez Moreno	1
Oficina de empleo	Kenia Reyes Tamayo	1
Proceso para seleccionar agente de sp	Luis Armando de la Paz Rives	1
prueba1	Marbely Castillo Villamarin	1
	Silvia García Cutiño	1

Mostrando 1 a 11 de 11 elementos

Anterior 1 Siguiente

Ilustración 32 Tablero 3 Cantidad de candidatos presentados

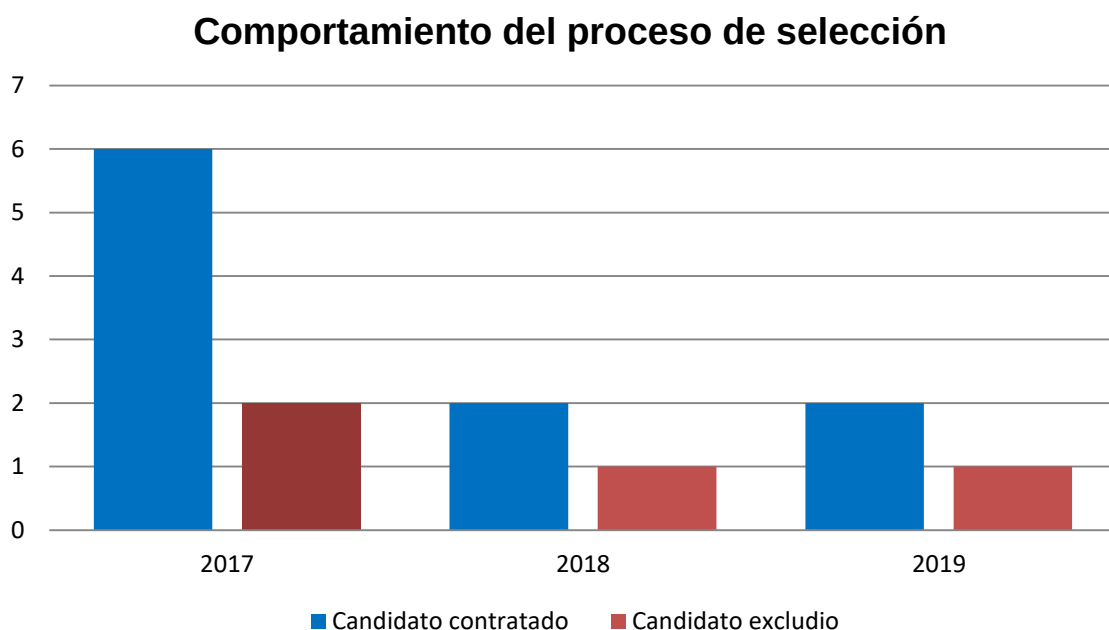


Ilustración 33 Tablero 4 Cantidad de candidatos contratados y excluidos

III.3 Validación del Mercado de Datos

La validación se refiere al proceso de asegurarse de que el sistema cumpla con los requisitos y expertativas del cliente y de los estándares establecidos. Según el autor Roger S. Pressman, la validación es “*el proceso de evaluación de un sistema o componente durante o al final del proceso de dearrollo para determiar si cumple con*

los requisitos especificados.” (Pressman, 2014). Las validaciones incluyen pruebas para garantizar que el sistema funcione correctamente y cumpla con las necesidades del cliente.

III.3.1 Pruebas para validar el Mercado de Datos.

Para cumplir con lo anteriormente planteado se realizó la siguiente prueba al Mercado de Datos:

La prueba de usabilidad se define como: una técnica de evaluación en la que los usuarios interactúan con un producto, sistema o sitio web y se observa su comportamiento. Se realizan para identificar posibles fallos en la implementación del sistema y mejorar la experiencia del usuario.

Para realizar la prueba de usabilidad al Mercado de Datos se utiliza la técnica de Evaluación Heurística, la cual consiste en realizar un diagnóstico de la interfaz, con base en una lista de chequeo de heurísticas y luego se realiza un análisis para discutir los elementos encontrados en la interfaz. La herramienta para aplicar esta técnica es mediante la lista de chequeo que a continuación, se describe:

Lista de Chequeo:

La lista de chequeo efectúa una serie de preguntas en forma de cuestionario, que sirven para verificar el grado de cumplimiento de determinadas reglas establecidas. Su objetivo principal es dar una evaluación de la calidad del producto:

Forma de uso:

Peso: Define si el criterio a evaluar es crítico o no.

Criterio de evaluación: Pregunta que permite evaluar un criterio.

Evaluación: Es la forma de evaluar el criterio en cuestión. El mismo se evalúa de 1 en caso de mal (cuando la respuesta al criterio sea “No”) y 0 en caso que el elemento revisado no presente errores (cuando la respuesta al criterio sea “Sí”).

NP (No Procede): Se usa para especificar que el criterio a evaluar no se puede aplicar en ese caso.

Observación: Especifica los señalamientos o sugerencias que quiera incluir la persona que aplica la lista de chequeo.

A continuación se muestra la lista de chequeo:

Tabla 14 Lista de chequeo

Peso	Criterio de evaluación	Evaluación	Aplicabilidad [NP]	Observaciones
Documento				
1	¿Se concluyó cada una de las fases de la metodología correctamente?	0		
2	¿Es fácil de entender lo que se ha especificado en el documento?	0		
3	¿Se generaron todos los artefactos definidos?	1		Generar los artefactos definidos
4	¿Posee una correcta ortografía?	1		Corregir los errores ortográficos existentes
Mercado de Datos				
5	¿Los subtítulos de las filas y las columnas corresponden a la descripción del mismo?	0		
6	¿Todos los campos de las dimensiones fueron llenados?	1		Faltan campos por llenar
7	¿Fueron pobladas correctamente las dimensiones y el hecho?	0		
8	¿Hay suficiente información para determinar una evaluación de otra?	0		
Damix				
9	¿El sistema tiene errores ortográficos?	1		
10	¿El proceso puede ser analizados desde todas sus medidas o indicadores?	1		
11	¿El título de la gráfica se corresponde con los datos mostrados?	0		

12	¿El tipo de gráfica seleccionado es el más representativo para mostrar la información?	1		Analizar otros tipos de gráficos
13	¿Es mayor la cantidad de candidatos contratados que en el proceso anterior?	0		
14	¿El porcentaje de candidatos descalificados fue aumentando por cada periodo de tiempo?	0		

III.3.1 Resultado de las pruebas

Se plantearon tres iteraciones para evaluar el Mercado de Datos desarrollado. A continuación se muestran la cantidad de no conformidades(Nc) por iteraciones(**ver anexo 8**)

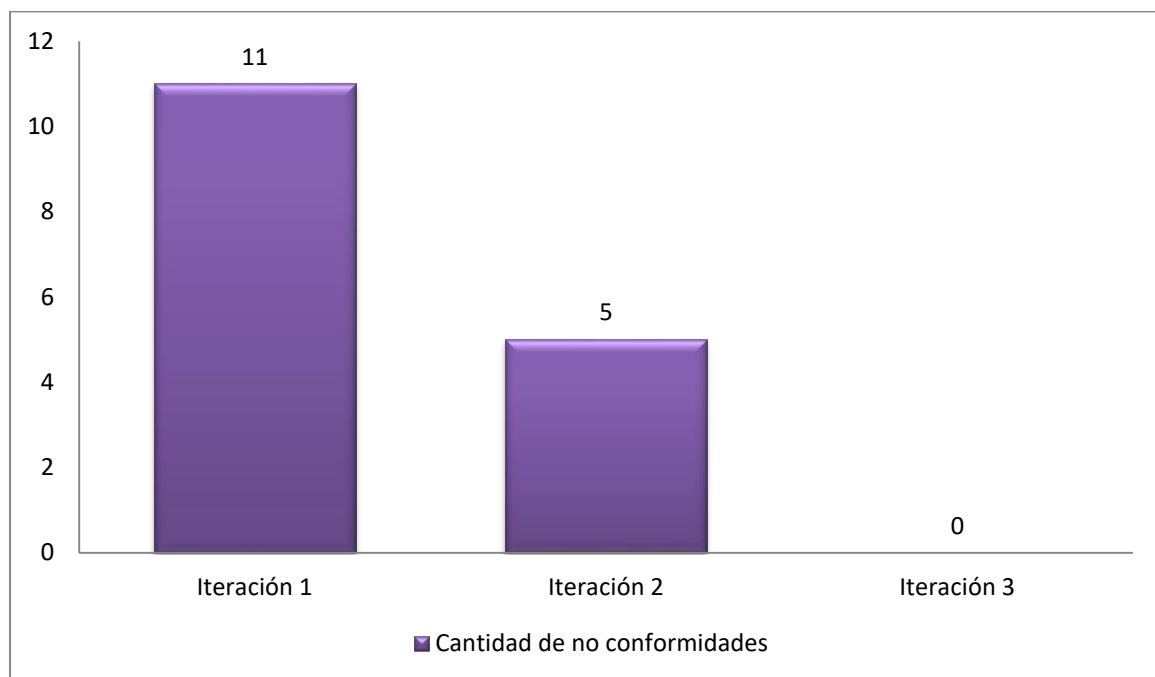


Ilustración 34 Resultados encontrados por iteración

Conclusiones parciales

Al finalizar este capítulo se concluyó que:

- La herramienta PDIv4.2 utilizada para la realización de los procesos ETL permitió la correcta población de las tablas dimensiones y el cubo del Mercado de Datos.
- Con la utilización del archivo XML, se logró la creación de los tableros de control en la herramienta Damix.
- Con la aplicación de la prueba de usabilidad se logró encontrar errores en el Mercado de Datos, y estos a su vez fueron eliminados para obtener un sistema que funcione correctamente.

CONCLUSIONES GENERALES

Al concluir la investigación para el desarrollo del Mercado de Datos del proceso de selección del personal de los RRHH para la XETID, se pudo arribar a las siguientes conclusiones:

- El uso de la metodología UCID en la investigación, permitió una construcción sencilla, y ordenada del Mercado de Datos, siendo esta la metodología utilizada para el desarrollo de Mercado de Datos y Almacén de Datos en la XETID.
- Las entrevistas realizadas permitieron identificar los principales requerimientos de información, permitiendo la obtención de los indicadores y perspectivas, que constituyó el punto de partida para el desarrollo del sistema.
- Se analizó la base de datos SIGERH, con la cual se pudo obtener los campos para cada una de las dimensiones del Mercado de Datos.
- La herramienta para los procesos ETL utilizada posibilitó la exitosa población de las tablas del Mercado de Datos.
- Se desarrolló el Mercado de Datos del proceso de selección del personal de los RRHH para la empresa XETID.
- Se logró configurar los indicadores para visualizar el comportamiento del proceso de selección del personal de los RRHH para la empresa XETID.

RECOMENDACIONES

Se recomienda:

- La implementación de futuros Mercados de Datos con características similares como, que permita medir un proceso en otras áreas, utilizando la metodología y las herramientas utilizadas en la investigación.
- Implementar nuevos indicadores en el Mercado de Datos que puedan satisfacer las necesidades del cliente.

REFERNCIA BIBLIOGRÁFICA

- Acero Calizaya, R.F. 2014.** DATAMART de Construcciones Públicas a Partir del Seace y Su Aplicación en la Toma de decisiones de las Micro y Pequeñas Emresas de la Ciudad de Puno-2013. [En línea] 2014.
- Acosta del Sol, Henry y Barroso Casanova, Juan Pablo. 2016.** Mercado de datos para el área Recurso Cama del sector de salud. [En línea] 2016. <https://repositorio.uci.cu/jspui/handle/123456789/8036>.
- Alonso Sarría, Francisco. 2016.** Programación en SQL con PostgreSQL. [En línea] 2016. <http://www.um.es/geograf/sigmur/sigpdf/postgresql.pdf>.
- Armijos Mayon, Franklin Brian, Bermúdez Burgos, Aaron Issac y Mora Sánchez, Norman Vinicio. 2019.** Gestión de Administración de los Recursos Humanos. 15 de Julio de 2019, Vol. 11, 4.
- Ayala, Evelyn y Gonzales, Santiago. 2015.** *Tecnología de la Información y las comunicaciones*. Universidad Inca Garcilaso de la Vega. Perú : s.n., 2015.
- Bernabeu, R. D. 2007.** Data Warehouse: Investigación y Sistematización de conceptos- Hefesto: Metodología propia para la construcción de un Data Warehouse. Argentina. [En línea] 2007.
- Blanco-Cuaresma, S. 2019.** Metodología de desarrollo. [En línea] 15 de diciembre de 2019. <http://www.marblestation.com/?p=644..>
- Booch, G, Rumbaugh, J y Jacobson, I. 2005.** *The Unified Modeling Language User Guide*. Addison Wesley, Reading. 2005.
- Bustos Barrera, S, y otros. 2013.** Análisi, diseño e implementación de una solución de BI para la generación de indicadores y control del desempeño, en la empresa OTECEL, utilizando la metodología Hefesto v2.0.Ecuador. 2013.
- Castillo O' Farrill, Suinny y Pérez Hernández, Yaneli. 2017.** Mercado de datos para el área de los Recursos Humanos . [En línea] Junio de 2017. <https://repositorio.uci.cu/jspui/handle/1233456789/9894>.
- Chiavenato, I. 2006.** Administración del personal. En I. Chiavenato, Administración del personal. 2006, 6.
- Cuesta, A. 2005.** *Tecnología de gestión de recursos humanos*. 2da, La Habana : Editorial Academica, 2005.
- Economipedia, E. 2022b.** Las 10 funciones del departamento de los recursos humanos. Economipedia. [En línea] 2022b. <http://economipedia.com/10-funciones-del-departamento-de-recursos-humanos.html>.

Eito Brum, Ricardo. 2005. XML y la gestión de contenido. [En línea] 2005. <http://arxiu-web.upf.edu>numero-3>.

Espinosa, W, Loáisiga, S y Sánchez , K. 2019. Proceso de selección del personal de recursos humanos en las organizaciones. [En línea] 2019. <https://repositorio.unan.edu.ni/10990/1/19664.pdf>.

Gaínza Martínez, Eberly, y otros. 2022. Gestión de Recursos Humanos en Proyectos de la Empresa de Aplicaciones Informáticas DESOFT. 28 de diciembre de 2022, Vol. 17, 1, págs. 60-77.

Gan Estaún, Marta. 2017. *Las nuevas tecnologías aplicadas a la gestión del talento. New technologies applied to talent management.* 2017.

Gesé Bordils, M, Gonzáles-Cancelas, N y Molina Serrano, B. 2021. Indicadores claves de rendimiento en terminales de contenedores y su relación con la sostenibilidad. Aplicación al sistema portuario español. [En línea] 2021. <https://doi.org/10.4067/S0718-3305221000400647>.

González Piñera, Irela. 2021. Mercado de datos para el perfeccionamiento educacional en Cuba. [En línea] 2021. <http://publicaciones.uci.cu>.

Guerra Viltres, Yoelsis Elisabeht . 2020. Cuadro de mando para la Dirección Comercial de la Xetid. [En línea] 2020.

KeepCoding, Redacción. 2022. Qué es Pentaho Data Integración? [En línea] 2022. <http://keepcoding.io/blog/que-es-pentaho-data-integration/>.

Kimball . 2008. *The Data Warehouse Lifecycle Toolkit.* [ed.] Wiley. 2, Nueva York : s.n., 2008.

Machado Marquez, Luz Edith. 2021. Planeación estratégica de Recursos Humanos. [En línea] 20 de marzo de 2021. <https://es.linkedin.com>.

Martínez, R. 2013. PostgreSQL-es. PostgreSQL-es. [En línea] 2013. <http://www.postgresql.org.es/sobre-postgresql>.

Oracle, R. 2022. Tableros de control que son, cómo diseñarlos, donde usarlo? [En línea] 2022. <https://gestion.pensemos.com/tableros-de-control-que-son-disenarlos-donde-usarlos>.

PostgreSQL. 2018. PostgreSQL. PostgreSQL. [En línea] 2018. <https://www.postgresql.org/docs/9.4/static/intro-what-is.html>.

Pressman, Roger S. 2014. *Ingeniería del software. Un enfoque práctico.* 2014.

Quintero, J. B. 2012. Un estudio comparativo de herramientas para modelado con UML. 2012, Vol. 41, 137, págs. 60-76.

Riaño, J. 2020. *Metodología para identificar y medir KPIs logísticos para el sector agroindustrial colombiano.* Universidad nacional de colombia . Colombia : s.n., 2020. Tesis de maestría .

Silvero Jovel, Yenisey y Gomzález Galá, Yenlis. 2017. Mercado de datos para el primer año de la carrera Ingeniería en Ciencias Informáticas. [En línea] 2017. <http://repositorio.uci.cu>.

Sounto Anido, Lourdes y García Rondón, Irene. 2016. Herramientas de apoyo para el Proceso de Toma de Decisiones Concernientes a la Gestión de los Recursos Humanos. 15 de julio de 2016, Vol. 2, 2, págs. 50-64.

Trujillo, A. 2013. *Extensión de la herramienta Visual Paradigm para la generación de las clases de acceso de datos con Doctrine 2.0. Tesis para optar por el título de Ingeniería en Ciencias Informáticas.* Universidas de las Ciencias Informáticas. La Habana : s.n., 2013.

Urtiaga, Gerardo G. 2020. *Administrador MySQL Y MariaDB: Aprende a administrar MySQL y MariaDB facilmente.* 2020.

Wolff, Carmen. 2000. *Implementando un DataWarehouse.* s.l. : Revista Ingeniería Informática: revista electrónica del DIICC., 2000.

Xetid. 2019. www.xetid.cu. [En línea] enero de 2019. <https://www.xetid.cu/producto/35>.

Xetid, E. 2012. Metodología UCID para el desarrollo de Data Warehouse.Habana. [En línea] 2012.

BIBLIOGRAFÍA

Acero Calizaya, R.F. 2014. DATAMART de Construcciones Públicas a Partir del Seace y Su Aplicación en la Toma de decisiones de las Micro y Pequeñas Emresas de la Ciudad de Puno-2013. [En línea] 2014.

Acosta del Sol, Henry y Barroso Casanova, Juan Pablo. 2016. Mercado de datos para el área Recurso Cama del sector de salud. [En línea] 2016. <https://repositorio.uci.cu/jspui/handle/123456789/8036>.

Ahumada Tello, E y Persuquia Velasco, J M. 2016. Inteligencia de negocio: Estrategia para el desarrollo de competitividad en empresas de base tecnológica. Contaduría y Administración. [En línea] 2016. <https://doi.org/10.1016/j.cya.2015.09.006>.

Alonso Sarría, Francisco. 2016. Programación en SQL con PostgreSQL. [En línea] 2016. <http://www.um.es/geograf/sigmur/sigpdf/postgresql.pdf>.

Armijos Mayon, Franklin Brian, Bermúdez Burgos, Aaron Issac y Mora Sánchez, Norman Vinicio. 2019. Gestión de Administración de los Recursos Humanos. 15 de Julio de 2019, Vol. 11, 4.

Ayala, Evelyn y Gonzales, Santiago. 2015. *Tecnología de la Información y las comunicaciones*. Universidad Inca Garcilaso de la Vega. Perú : s.n., 2015.

Baraybar, F A. 2011. El cuadro de Mando Integral<<Balanced Scoreal>>. 2011.

Bernabeu, R. D. 2007. Data Warehouse: Investigación y Sistematización de conceptos- Hefesto: Metodología propia para la construcción de un Data Warehouse. Argentina. [En línea] 2007.

Blanco-Cuaresma, S. 2019. Metodología de desarrollo. [En línea] 15 de diciembre de 2019. <http://www.marblestation.com/?p=644..>

Booch, G, Rumbaugh, J y Jacobson, I. 2005. *The Unified Modeling Language User Guide*.Addison Wesley, Reading. 2005.

Bustos Barrera, S, y otros. 2013. Análisi, diseño e implementación de una solución de BI para la generación de indicadores y control del desempeño, en la empresa OTECEL, utilizando la metodología Hefesto v2.0.Ecuador. 2013.

Bustos, F. G y Prats, J .T. 2013. Cuadro de mando integral. 2013.

Castillo O' Farrill, Suinny y Pérez Hernández, Yaneli. 2017. Mercado de datos para el área de los Recursos Humanos . [En línea] Junio de 2017. <https://repositorio.uci.cu/jspui/handle/1233456789/9894>.

Charry Vargas, Sebastian. 2020. Selección del personal. [En línea] 2020.

Chiavenato, I. 2006. Administración del personal. En I. Chiavenato, Administración del personal. 2006, 6.

Cuesta, A. 2005. *Tecnología de gestión de recursos humanos*. 2da, La Habana : Editorial Academica, 2005.

Dousat, Garcías. 2009. Procedimiento para la Gestión de Selección e Integración al Empleo en la Empresa Empleadora del Níquel. [En línea] 2009.

Economipedia, E. 2022b. Las 10 funciones del departamento de los recursos humanos. Economipedia. [En línea] 2022b. <http://economipedia.com/10-funciones-del-departamento-de-recursos-humanos.html>.

Eito Brum, Ricardo. 2005. XML y la gestión de contenido. [En línea] 2005. [Http://arxiu-web.upf.edu/numero-3](http://arxiu-web.upf.edu/numero-3).

Espinosa, W, Loáisiga, S y Sánchez , K. 2019. Proceso de selección del personal de recursos humanos en las organizaciones. [En línea] 2019. <https://repositorio.unan.edu.ni/10990/1/19664.pdf>.

Fernández, Manuel. 2023. Proceso de selección del personal: que fases tienen y como abordarlo. [En línea] 5 de enero de 2023. <https://www.openhr.cloud>.

Flóres-Guzman,, M H, Hernández-Aros, L y Gallego-Cossio, L C. 2015. Tableros de control como herramienta especializada: Perspectivas desde la auditoría forense. [En línea] 2015. <https://doi.org/10.11144/Javeriana.cc16-42.tche>.

Freire Llor, Cristian Joel. 2022. Evaluación del proceso de reclutamiento y selección del personal en la empresa pública de Santo Domingo EP-CONST, 2020. [En línea] Noviembre de 2022. <http://repositorio.esпам.edu.ec/handle/12000/1996>.

Gaínza Martínez, Eberly, y otros. 2022. Gestión de Recursos Humanos en Proyectos de la Empresa de Aplicaciones Informáticas DESOFT. 28 de diciembre de 2022, Vol. 17, 1, págs. 60-77.

Gan Estaún, Marta. 2017. *Las nuevas tecnologías aplicadas a la gestión del talento. New technologies applied to talent management*. 2017.

Gesé Bordils, M, Gonzáles-Cancelas, N y Molina Serrano, B. 2021. Indicadores claves de rendimiento en terminales de contenedores y su relación con la sostenibilidad. Aplicación al sistema portuario español. [En línea] 2021. <https://doi.org/10.4067/S0718-3305221000400647>.

Gómez, A A y Bautista, D W. 2010. Ingeniería de negocio: Estado de arte. Sientia et Technica. [En línea] 2010. <https://doi.org/10.22517/23447214.1803>.

González Piñera, Irela. 2021. Mercado de datos para el perfeccionamiento educacional en Cuba. [En línea] 2021. <http://publicaciones.uci.cu>.

Guerra Viltres, Yoelsis Elisabeht . 2020. Cuadro de mando para la Dirección Comercial de la Xetid. [En línea] 2020.

Hernandez-Sampieri, Roberto y Mendoza Torres, Christian Paulina. 2018. *METODOLOGÍA DE LA INVESTIGACIÓN: LAS RUTAS CUANTITATIVA, CUALITATIVA Y MIXTA*. Mexico : s.n., 2018.

Immon, W H. 1992. Building the Data Warehouse,QED Technilshin Group. [En línea] 1992.

KeepCoding, Redacción. 2022. Qué es Pentaho Data Integración? [En línea] 2022. <http://keepcoding.io/bloq/que-es-pentaho-data-integration/>.

Kimball . 2008. *The Data Warehouse Lifecycle Toolkit*. [ed.] Wiley. 2, Nueva York : s.n., 2008.

Kimball Ralph, C.J. 2004. *The Data Warehouse ETL Toolkit*. Indianapolis, United States of America: Wiley Publishing. 2004.

Lamas , Erias . 2021. Proceso de selección del personal: discriminación e igualdad de oportunidades. [En línea] 2021.

Machado Marquez, Luz Edith. 2021. Planeación estratégica de Recursos Humanos. [En línea] 20 de marzo de 2021. <https://es.linkedin.com>.

Martinez Reinoso, Olga Lidia . 2020. Procedimeinto para la gestión del proceso de reclutamiento y selección del personal en la Empresa Provincial de la Industria Alimentaria de Pinar del Rio. [En línea] 2020. <http://socielo.sld.cu>.

Martínez, L. 2015. Administración de Recursos Humanos. Cómo funciona? [En línea] 2015. [https://www.ecured.cu/Gesti%C3%B3n de recursos humanos](https://www.ecured.cu/Gesti%C3%B3n%20de%20recursos%20humanos).

Martínez, R. 2013. PostgreSQL-es. PostgreSQL-es. [En línea] 2013. <http://www.postgresql.org.es/sobre/postgresql>.

Mendez Serrano, Inalvis y Sotolongo Sánchez, María. 2022. Contribución al perfeccionamiento del proceso de selección del personal basado en competencia . [En línea] 2022. <https://apye.esceg.cu/index.php/apye/article/view/232>.

Oracle, R. 2022. Tableros de control que son,cómo diseñarlos, donde usarlo? [En línea] 2022. <https://gestion.pensemos.com/tableros-de-control-que-son-disenarlos-donde-usarlos>.

PostgreSQL. 2018. PostgreSQL.PostgreSQL. [En línea] 2018. <https://www.postgresql.org/docs/9.4/static/intro-what-is.html>.

Pressman, Roger S. 2014. *Ingeniería del software.Un enfoque práctico*. 2014.

Quintero, J. B. 2012. Un estudio comparativo de herramientas para modelado con UML. 2012, Vol. 41, 137, págs. 60-76.

Ramesh, Gopalaswamy y Desikan, Srinivasa . 2006. *Software Testing: Principles and Practices*. 2006.

Riaño, J. 2020. *Metodología para identificar y medir KPIs logísticos para el sector agroindustrial colombiano*.[.]. Universidad nacional de colombia . Colombia : s.n., 2020. Tesis de maestría .

Sani , G, Zamora, B y Moscoso, N. 2017. Diseño de manual procedimiento de mejora para la gestión del talento humano en empresa MASTERSOFT S.A Guanayil(Tesis presentada como requisito para optar por el título de Ingeniería Comercial). [En línea] 2017.

Silvero Jovel, Yenisey y Gomzález Galá, Yenlis. 2017. Mercado de datos para el primer año de la carrera Ingeniería en Ciencias Informáticas. [En línea] 2017. <http://repositorio.uci.cu>.

Sounto Anido, Lourdes y García Rondón, Irene. 2016. Herramientas de apoyo para el Proceso de Toma de Decisiones Concernientes a la Gestión de los Recursos Humanos. 15 de julio de 2016, Vol. 2, 2, págs. 50-64.

Tim, Bay. 2006. Extensible Markup Lenguaje(XML)1.0. [En línea] 16 de agosto de 2006. <http://www.w3.org/TR/REC-xml/>.

Trujillo, A. 2013. *Extensión de la herramienta Visual Paradigm para la generación de las clases de acceso de datos con Doctrine 2.0. Tesis para optar por el título de Ingeniería en Ciencias Informáticas.* Universidas de las Ciencias Informáticas. La Habana : s.n., 2013.

Urtiaga, Gerardo G. 2020. *Administrador MySQL Y MariaDB: Aprende a administrar MySQL y MariaDB facilmente.* 2020.

Vassiliadis, P, Simitsis, A y Skiadopoulos, S. 2002. Conceptual modeling for ETL processes.Proceedings of the 5th ACM international workshop on Data Warehousing and OLTP. [En línea] 2002. <https://doi.org/10.1145/583890.583893>.

Wolff, Carmen. 2000. *Implementando un DataWarehouse.* s.l.: Revista Ingeniería Informática: revista electrónica del DIICC., 2000.

Xetid. 2019. www.xetid.cu. [En línea] enero de 2019. <https://www.xetid.cu/producto/35>.

Xetid, E. 2012. Metodología UCID para el desarrollo de Data Warehouse.Habana. [En línea] 2012.

ANEXOS

Anexo 1. Entrevista 1

Se desea realizar esta entrevista para obtener información sobre el proceso de selección del personal.

Nombre del entrevistado:_____.

1. ¿Qué tiempo lleva usted laborando en la actividad de Recursos Humanos?

Menos de 1 año____ de 1 a 5 años____ más de 5 años ____

2. ¿Tiene experiencia en otras actividades dentro del centro?

No ____ Sí ____ ¿Cuál o cuáles? _____

3. ¿En qué medida le ha sido útil la experiencia que posee para realizar los análisis del proceso de selección de personal en la entidad?

4. ¿Qué documentos legales usted conoce que rigen el proceso de selección de personal en la entidad?

5. ¿Recibe indicaciones específicas para el proceso de selección de personal?

No ____ Sí ____ ¿Por qué vía? _____

6. ¿Cuál considera usted que sea el factor más relevante por el cual no se tiene un proceso sistemático para la selección de personal en la entidad?

7. ¿Por qué medio de comunicación dan a conocer las plazas vacantes que están disponibles en la empresa?

8. ¿Una vez seleccionado el personal, qué tipo de comprobaciones o pruebas realizan para evaluar su idoneidad en función de su contratación de forma indeterminada?

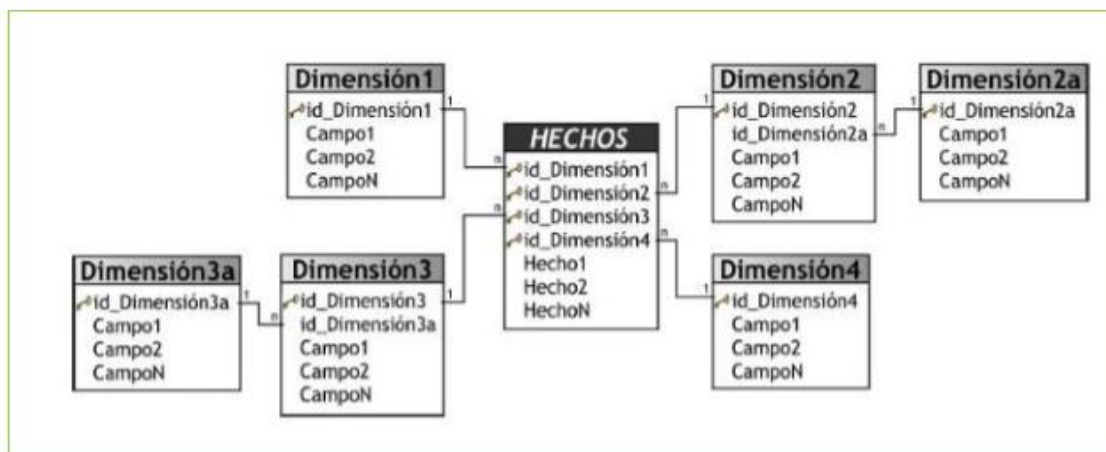
9. ¿Cuáles son los requisitos del cargo que se buscan en los nuevos empleados?

10. ¿Considera necesario que se diseñen las herramientas técnicas para medir el proceso de selección de personal en la entidad?

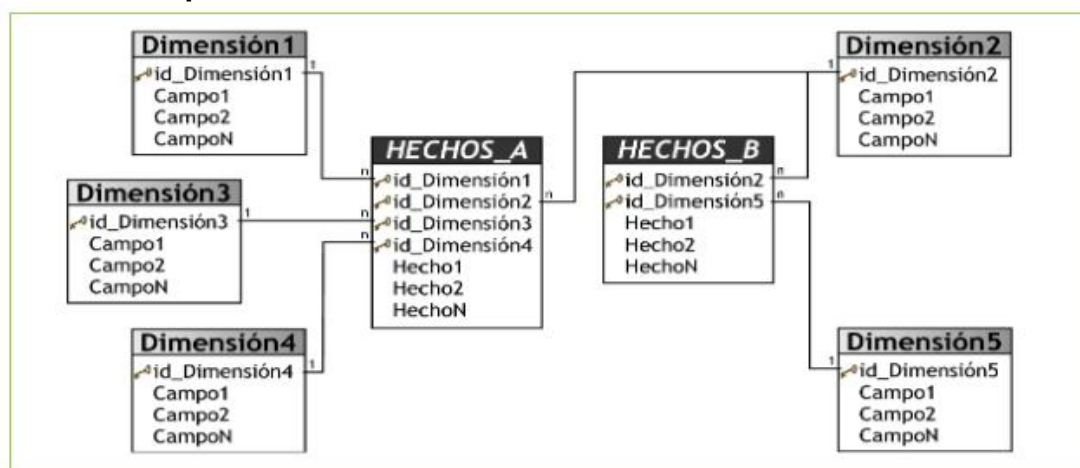
Fuente: Elaboración propia

Muchas Gracias.

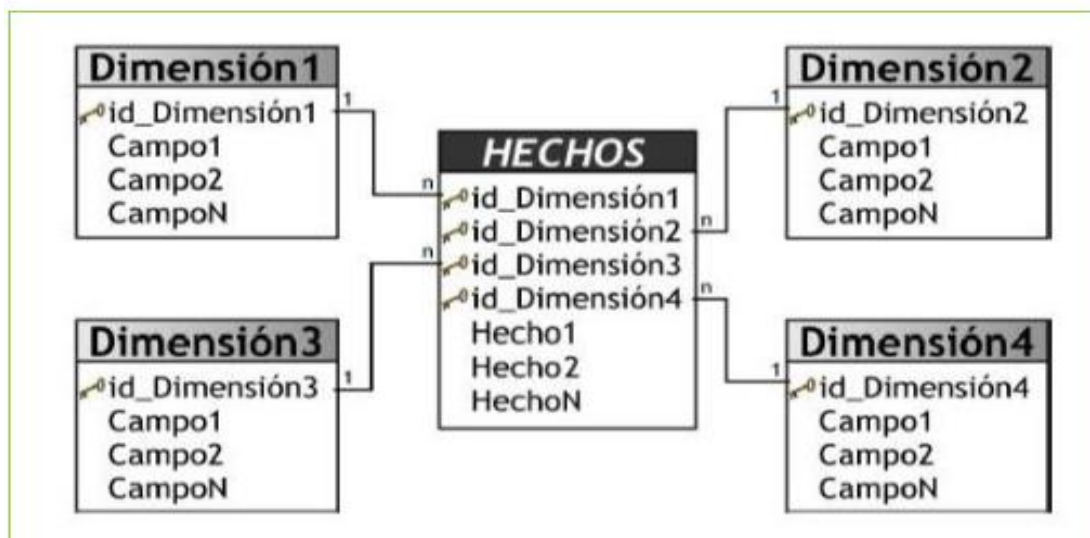
Anexo 2. Esquema copo de nieve



Anexo 3. Esquema constelación



Anexo 4. Esquema de Estrella



Anexo 5. Planificación de las entrevistas

No.	Entrevistado	Objetivo	Logro
E.1	Especialistas del departamento de los Recursos Humanos	Conocer el funcionamiento del negocio en cuestión. Además conocer acerca del sistema con que se cuenta para la gestión de los Recursos Humanos.	Tener mejor conocimiento del negocio.
E.2	Especialistas del departamento de los Recursos Humanos	Fue realizada con el objetivo de comenzar a identificar posibles indicadores y perspectivas a partir de preguntas elaboradas según las necesidades de información del cliente.	Realizar las pregunta pertinentes al cliente.
E.3	Especialistas del departamento de los Recursos Humanos	Definir el nivel de granularidad de cada una de las perspectivas identificadas para la solución.	Seleccionar los campos que integrarán cada una de las perspectivas identificadas.
E.4	Especialistas del departamento de los Recursos Humanos	Definir cómo calcular los indicadores de análisis del comportamiento del proceso de selección del personal	Obtener las fórmulas para calcular los indicadores de los hechos identificados

Anexo 6. Entrevista 2

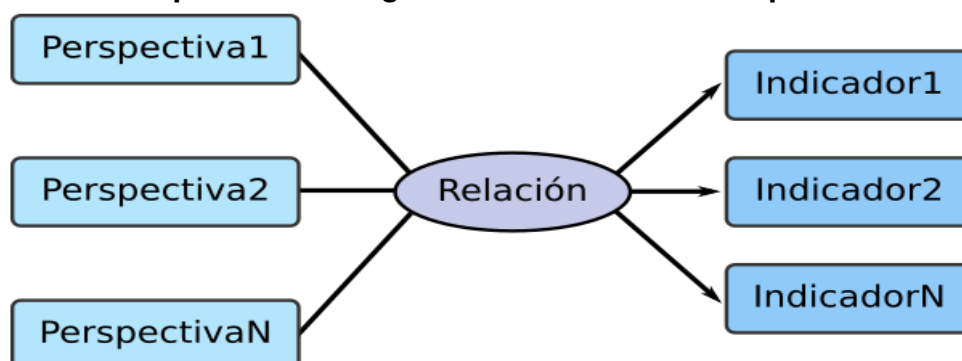
Se decea realizar una entrevista con el objetivo de comenzar a identificar los posibles indicadores y perspectivas a partir de preguntas elaboradas según las necesidades de información del cliente.

Nombre del entrevistado: _____

Preguntas realizadas al cliente

1. ¿Con cuáles indicadores se trabaja más en el área?
2. ¿Cómo se obtienen los indicadores para visualizar el proceso?
3. ¿Cuál es la frecuencia con que se actualiza la información?
4. ¿Cómo se realiza el análisis de la información en el área?
5. ¿Le gustaría que los datos se presentan mediante vistas de análisis, reportes operacionales o cuadros de mando?

Anexo 7. Representación gráfica del Modelo Conceptual



Anexo 8. No conformidades por iteraciones

Primera iteración	Segunda iteración	Tercera iteración
Documento	Docuemento	
No se generó correctamente el artefacto: Modelo lógico de datos.	Se detectaron algunos errores ortográficos en el documento.	
No se generó el artefacto: Especificación de indicadores.		
No se generó el artefacto: Ficha de indicador Número de candidatos entrevistados.		
Mercado de Datos	Mercado de Datos	
Faltaron campos por llenar en la dimensión persona.	Faltaron campos por llenar en la dimensión persona.	
Faltaron campos por llenar en la dimensión tiempo.		
Damix		
El tipo de gráfica seleccionado no era el más representativo para mostrar la información referente al comportamiento del proceso de selección.	El tipo de gráfica seleccionado no era el más representativo para mostrar la información referente al comportamiento del proceso de selección.	
El proceso no podía ser analizado desde la medida cantidad de candidatos contratados.	El proceso no podía ser analizado desde el campo nombre de la dimensión persona	
El proceso no podía ser	El sistema presenta	

analizado desde la dimensiones resultado de selección.	errores ortográficos en la dimensión resultado de selección.	
El proceso no podía ser analizado desde el campo nombre de la dimensión persona.		
El sistema presenta errores ortográficos en la dimensión resultado de selección.		
El sistema presenta errores ortográficos en la medida porcentaje de candidatos descalificados.		