

UNIVERSIDAD PRIVADA ANTENOR ORREGO
FACULTAD DE INGENIERÍA
ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN Y
DE SISTEMAS



TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE INGENIERO DE
COMPUTACION Y SISTEMAS

DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE
NEGOCIOS PARA LA MEJORA EN EL ANÁLISIS DE INFORMACIÓN
EN LAS ÁREAS DE ADMISIÓN Y LABORATORIO DEL CENTRO
MÉDICO PACIFICO DEL NORTE DE LA CIUDAD DE TRUJILLO
USANDO LA ARQUITECTURA TECNOLÓGICA DE PENTAHO BI Y
LA METODOLOGÍA DE LARISSA MOSS

Línea de Investigación:

Gestión de Datos y de Información.

AUTORES: Br. CUZCO ODAGUIRI JOSE KONOSUKE

Br. UGAZ VITTERI FIORELLA MILDRED

ASESOR: Ing. AGUSTIN EDUARDO ULLON RAMIREZ

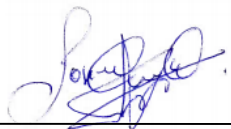
TRUJILLO - PERÚ

2019

Fecha de Sustentación:27/10/20

**“DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE
NEGOCIOS PARA LA MEJORA EN EL ANÁLISIS DE
INFORMACIÓN EN LAS ÁREAS DE ADMISIÓN Y
LABORATORIO DEL CENTRO MÉDICO PACIFICO DEL NORTE
DE LA CIUDAD DE TRUJILLO USANDO LA ARQUITECTURA
TECNOLÓGICA DE PENTAHO BI Y LA METODOLOGÍA DE
LARISSA MOSS”**

Elaborado por:



Br. Cuzco Odaguiri José Konosuke



Br. Ugaz Vitteri Fiorella Mildred

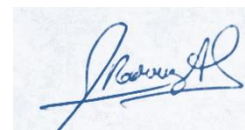
Aprobada por:



Ing. Edward Fernando Castillo Robles
Presidente
CIP: 192352



Ing. Heber Gerson Abanto Cabrera
Secretario
CIP: 106421



Ing. Silvia Ana Rodríguez Aguirre
Vocal
CIP: 107615



Ing. Agustín Eduardo Ullón Ramírez
Asesor
CIP: 137602

PRESENTACIÓN

Señores Miembros del Jurado:

Cumpliendo los requisitos descritos en el reglamento de grados y Títulos de la Universidad Privada Antenor Orrego y el Reglamento Interno de la Escuela Profesional de Ingeniería de Computación y Sistemas, se pone a disposición el presente Trabajo de Tesis titulado: “**DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA LA MEJORA EN EL ANÁLISIS DE INFORMACIÓN EN LAS ÁREAS DE ADMISIÓN Y LABORATORIO DEL CENTRO MÉDICO PACIFICO DEL NORTE DE LA CIUDAD DE TRUJILLO USANDO LA ARQUITECTURA TECNOLÓGICA DE PENTAHO BI Y LA METODOLOGÍA DE LARISSA MOSS**” para obtener el Título Profesional de Ingeniero de Computación y Sistemas.

El presente trabajo ha sido desarrollado tomando en cuenta el marco de referencia de los lineamientos establecidos por la Facultad de Ingeniería, la Escuela Profesional de Ingeniería de Computación y Sistemas, además de los conocimientos que se adquirieron durante nuestra formación profesional.

Los autores.

DEDICATORIA

Con mucho amor y cariño le dedico a mi familia y en especial a mi abuela que está en el cielo, todo mi esfuerzo y trabajo puesto para la realización de esta tesis.

Br. CUZCO ODAGUIRI JOSE KONOSUKE

A mis padres quienes siempre han sido un ejemplo. A mis profesores por el apoyo y la incansable paciencia. A mis hijos por comprenderme a pesar de no poder acompañarlos en todo momento.

Br. UGAZ VITTERI FIORELLA MILDRED

AGRADECIMIENTO

Se agradece de manera especial el apoyo del personal del del centro médico Pacifico del Norte quienes nos apoyaron y nos brindaron todas las facilidades para obtener la información necesaria para el desarrollo del presente trabajo de tesis.

También agradecemos a nuestro asesor al Ing. Agustín Ullón, por su apoyo y asesoría en el desarrollo y término del presente trabajo de tesis.

También se agradece el apoyo de los amigos que compartieron aula con nosotros en gran parte de nuestra carrera y que siempre estuvieron pendiente de nuestro avance y logro de nuestro trabajo, así como los concejos que nos brindaron.

Muchas Gracias.

Los autores.

RESUMEN
**“DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS
PARA LA MEJORA EN EL ANÁLISIS DE INFORMACIÓN EN LAS ÁREAS
DE ADMISIÓN Y LABORATORIO DEL CENTRO MÉDICO PACIFICO DEL
NORTE DE LA CIUDAD DE TRUJILLO USANDO LA ARQUITECTURA
TECNOLÓGICA DE PENTAHO BI Y LA METODOLOGÍA DE LARISSA
MOSS”**

Por:

Br. Cuzco Odaguiri José Konosuke

Br. Ugaz Vitteri Fiorella Mildred

El Centro Médico Pacífico del Norte es una empresa dedicada a prestación de servicios de salud ocupacional y asistencial. La clínica requiere mejorar el análisis de la información para poder tomar decisiones más rápidas y mejorando la creatividad que se emplee, permitiendo la reducción de tiempo en la obtención de informes solicitadas por la gerencia, que servirán para tomar decisiones en el momento necesario, lo cual le permitirán tomar ventajas sobre el mercado y así tener un mejor panorama del estado de su estado actual para poder realizar una mejor toma de decisiones de sus áreas de áreas de Admisión y Laboratorio utilizando la inteligencia de negocios como herramienta tecnológica.

El objetivo principal del trabajo es “Mejorar en el análisis de información para el soporte en la toma de decisiones en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo usando la arquitectura tecnológica de Pentaho BI y basándose en parte de la metodología de Larissa Moss”. Teniendo como objetivos específicos: Obtener y analizar los requerimientos de las áreas críticas del negocio, Desarrollar el proceso ETL utilizando como herramienta Pentaho BI, Diseñar la estructura de que tendrán las dimensiones y cubos pertinentes para crear el diseño de vista de análisis y la navegabilidad que tendrá la información., Utilizar la herramienta Pentaho Report Designer para crear y diseñar los reportes correspondientes al Modelo implementado. Evaluar los resultados de los informes que muestra la solución BI implementada.

ABSTRACT

"DEVELOPMENT OF A BUSINESS INTELLIGENCE SOLUTION FOR IMPROVEMENT IN INFORMATION ANALYSIS IN THE ADMISSION AND LABORATORY AREAS OF THE PACIFIC MEDICAL CENTER OF THE CITY OF TRUJILLO USING THE TECHNOLOGICAL ARCHITECTURE OF PENTAHO BI AND THE LARISSA MOSS METHODOLOGY"

By:

Br. Cuzco Odaguiri José Konosuke

Br. Ugaz Vitteri Fiorella Mildred

The North Pacific Medical Center is a company dedicated to providing occupational and health care services. The clinic needs to improve the analysis of the information in order to make faster decisions and improve the creativity that is used, allowing the reduction of time in obtaining reports requested by management, which will serve to make decisions at the necessary time, which They will allow you to take advantages over the market and thus have a better overview of the state of your current state in order to make a better decision making in your areas of Admission and Laboratory areas using business intelligence as a technological tool.

The main objective of the work is to “Improve the analysis of information for support in decision-making in the Admission and Laboratory areas of the North Pacific Medical Center of the city of Trujillo using the technological architecture of Pentaho BI and based on part of the methodology of Larissa Moss”. Having as specific objectives: Obtain and analyze the requirements of the critical areas of the business, Develop the ETL process using as a Pentaho BI tool, Design the structure that will have the relevant dimensions and cubes to create the analysis view design and the navigability that will have the information., Use the Pentaho Report Designer tool to create and design the reports corresponding to the Model implemented. Evaluate the results of the reports shown by the implemented BI solution.

ÍNDICE DE CONTENIDO

PRESENTACIÓN	ii
DEDICATORIA	iii
AGRADECIMIENTO	iv
RESUMEN	v
ABSTRACT	vi
ÍNDICE DE CONTENIDO	vii
INDICE DE FIGURAS	x
INDICE DE TABLAS	xi
1. INTRODUCCION	01
1.1. Planteamiento del problema	01
1.2. Delimitación del problema	02
1.3. Características problemáticas.....	02
1.4. Definición del problema	04
1.5. Formulación del problema.....	04
1.6. Formulación del hipótesis.....	05
1.7. Objetivos del estudio	05
1.8. Justificación del estudio.....	05
1.8.1. Importancia de la investigación.....	05
1.8.2. Viabilidad de la investigación.....	06
1.8.3. Aportes.....	07
2. MARCO TEÓRICO	08
2.1. ANTECEDENTES.....	08
2.2. DEFINICIONES.....	11
2.2.1. INTELIGENCIA DE NEGOCIOS.....	11
2.2.2. VENTAJAS DE LA INTELIGENCIA DE NEGOCIOS	12
2.2.3. OLAP	13
2.2.4. TIPOS DE SISTEMAS OLAP	13
2.2.5. DATAWAREHOUSE	16
2.2.6. DATAMART.....	17
2.2.7. ETL	18
2.2.8. SISTEMA DE SOPORTE A LAS DECISIONES	19

2.2.9. FUNCIÓN Y CARACTERÍSTICAS DEL DSS	20
2.2.10. PENTAHO	21
2.3. METODOLOGIA PARA EL DESARROLLO DEL PROYECTO.....	23
3. MATERIALES Y METODOS.....	31
3.1. Material.....	31
3.1.1. Población.....	31
3.1.2. Muestra.....	31
3.1.3. Unidad de análisis.....	31
3.2. Método.....	31
3.2.1. Tipo de investigación.....	31
3.2.2. Diseño de Investigación.....	31
3.2.3. Variables de estudio y Operacionalización.....	32
3.2.4. Técnicas e instrumentos de recolección de datos.....	33
3.2.5. Técnicas de procesamiento y análisis de datos.....	33
4. RESULTADOS: APLICACIÓN DE LA METODOLOGIA.....	34
4.1. JUSTIFICACION.....	34
4.1.1. EVALUACIÓN DEL CASO DEL NEGOCIO.....	34
4.2. PLANIFICACIÓN.....	37
4.2.1. EVALUACIÓN DE INFRAESTRUCTURA DE LA EMPRESA.....	37
4.2.2. PLANIFICACIÓN DEL PROYECTO.....	38
4.3. ANÁLISIS DEL NEGOCIO.....	39
4.3.1. DEFINICIÓN DE LOS REQUISITOS DEL PROYECTO.....	39
4.3.1.1. Descripción de los Stakeholders.....	40
4.3.1.2. Requerimientos del Usuario Final.....	42
4.3.1.3. Requerimientos no funcionales	44
4.3.1.4. Análisis de los requerimientos	45
4.3.1.5. Identificar Medidas y Perspectivas de Análisis	49
4.3.1.6. Modelo conceptual del BI	49
4.3.2. ANÁLISIS DE DATOS.....	51
4.3.3. PROTOTIPO DE LA APLICACIÓN.....	53
4.4. DISEÑO.....	60
4.4.1. DISEÑO DE LA BASE DE DATOS.....	60
4.4.2. DISEÑO EXTRAER/TRANSFORMAR/CARGAR (ETL).....	65

4.5.	CONSTRUCCIÓN.....	68
4.5.1.	CONSTRUCCIÓN DE LAS TABLAS DEL BI EN LA BD MySQL	68
4.5.2.	PROCESO EXTRAER/TRANSFORMAR/CARGAR (ETL).....	73
4.5.3.	CREACION DE LA SOLUCION DE BI EN MICROSOFT AZURE	93
4.5.4.	DESARROLLO DE APLICACIONES.....	94
4.6.	GENERACIÓN DE REPORTES.....	106
4.7.	DESPLIEGUE.....	113
5.	DISCUSION DE RESULTADOS	114
6.	CONCLUSIONES	124
7.	RECOMENDACIONES.....	125
8.	REFERENCIAS BIBLIOGRAFICAS.....	126
	ANEXOS.....	129

INDICE DE FIGURAS

Figura 1: Esquema estrella	15
Figura 2: Esquema copo de nieve.....	16
Figura 3. Proceso ETL	19
Figura 4: Etapas de la Metodología	23
Figura 5: Etapas de la Metodología	30
Figura 6: Organigrama de la empresa	35
Figura 7: Proceso a desarrollar	36
Figura 8: Infraestructura tecnológica de la empresa	38
Figura 9: Modelo conceptual	51
Figura 10: Diagrama BD	52
Figura 11: Diagrama de Tabla de Hechos Admisión	55
Figura 12: Diagrama de Tabla de Hecho Laboratorio	56
Figura 13: Esquema Lógico de la Tabla de Hechos Admisión	58
Figura 14: Esquema Lógico de la Tabla de Hechos Laboratorio	59
Figura 15: Diagrama Multidimensional.....	59
Figura 16: Pasos de los Workflows.....	66
Figura 17: Diagrama Workflow con Restricciones de Precedencia	67
Figura 18: Creación base de datos en Azure	91
Figura 19: Diagrama de despliegue Detallado	110

INDICE DE TABLAS

Tabla 1: Diseño de la Investigación	32
Tabla 2: Operacionalización de las variables.....	32
Tabla 3: Lista de requerimientos	44
Tabla 4: Detalle de las claves de las dimensiones.....	56
Tabla 5: Detalle de las Medidas de las Tablas de Hechos.....	57
Tabla 6: Nombre Estándares para la Tabla de Hechos y sus Dimensiones.....	63
Tabla 7: Software para el proceso de Diseño del BI	92
Tabla 8: Rango de grado de valoración	113
Tabla 9: Evaluación de los indicadores de la hipótesis.	115
Tabla B1. Jefe de Admisión - Laboratorio y Gerente	127

1. INTRODUCCION

1.1. Planteamiento del problema

Actualmente la Tecnología de la Información se ha convertido en una herramienta clave en el proceso de desarrollo continuo dentro de las empresas. La competitividad del mercado y la globalización de la industria plantean un reto mayor dentro de toda organización pues se hace necesaria la innovación y el planeamiento estratégico que permita a la empresa trascender con un producto, proceso o servicio diferenciado.

El manejo eficiente y eficaz de los datos es un componente indispensable para la toma de decisiones que genera un conocimiento importante para cualquier organización en este contexto es donde aparecen conceptos como el de inteligencia de negocios que busca facilitar a las empresas a la recolección de su información corporativa, de la presentación de sus resultados, en resumen, ayudar a la empresa a comprender el significado de los datos que diariamente almacena en sus sistemas logísticos para que éstas puedan tomar decisiones que ayuden a crecer su rentabilidad .

Una empresa que no tome la inteligencia de negocios como una solución a sus problemas y mejora de procesos es una empresa que quedaría varada frente a un mercado competitivo, que aprovecha de manera inteligente los recursos tecnológicos y las tecnologías de información que existen hoy en día.

El Centro Médico Pacifico del Norte es una empresa dedicada a prestación de servicios de salud ocupacional y asistencial, constituida por profesionales altamente calificados que garantizan idoneidad en los resultados de los exámenes médicos ocupacionales, los cuales cumplen con los estándares y criterios de calidad exigidos por las normas legales; complementando sus servicios con la evaluación diagnóstica y tratamiento de enfermedades de sus pacientes.

La clínica requiere mejorar el análisis de la información para poder tomar decisiones más rápidas y mejorando la creatividad que se emplee, permitiendo la

reducción de tiempo en la obtención de informes solicitadas por la gerencia, que servirán para tomar decisiones en el momento necesario, lo cual le permitirán tomar ventajas sobre el mercado y así tener un mejor panorama del estado de su estado actual para poder realizar una mejor toma de decisiones de sus áreas de áreas de Admisión y Laboratorio utilizando la inteligencia de negocios como herramienta tecnológica ya que esta carece de conocimiento para mejorar sus procesos y aumentar su rentabilidad.

En el área de Admisión, así como en laboratorio, algunos de sus problemas son los siguientes: no se conoce con exactitud la cantidad total de exámenes médicos que se generan en un determinado mes, o indicados por un médico, o por tipo de examen. Otros de los problemas es que tampoco existe una comparativa de exámenes médicos hechos a pacientes en diferentes meses, ni la cantidad de consultas médicas realizadas, ni el promedio de pacientes particulares o por empresa atendidos en diferentes meses.

La prioridad de esta organización es mejorar la gestión de estas áreas con el fin de que la gerencia tenga indicadores con resultados que les permita tomar mejores decisiones y así mejorar la gestión y los procesos de las áreas interesadas.

1.2. Delimitación del problema

El proyecto se realizará analizando la realidad en la que se encuentra las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte mejorando el análisis de la información para dar soporte en la toma de decisiones.

1.3. Características problemáticas

Las características problemáticas nos dan la siguiente información:

- ✓ **Demora en los tiempos de ejecución o solicitudes de reportes de los sistemas actuales sobre información que afecta al proceso de las áreas de laboratorio y Admisión de la clínica.** No existe ningún sistema automatizado que pueda brindar estadísticas, reportes y proyecciones sobre los temas analíticos del proceso por lo que el tratar de hacerlo con los sistemas actuales

que cuenta la institución genera demora en atender dichas solicitudes, los cuales se realizan de 2 a 3 días.

- ✓ **Reprocesamiento de la información.** El reprocesar la información proporcionada por los sistemas actuales se ha convertido en un desafío por los tomadores de decisiones de la institución ya que estas generan limitantes que ocasionan que se organice la información por separado. Los tomadores de decisiones se demoran hasta 2 días en reprocesar información.

- ✓ **Falta de conocimiento de su propia data.** El área de TI de la clínica no explota adecuadamente la información que provienen de sus fuentes transaccionales, por lo que solo entregan reportes estáticos y difíciles de entender para un proceso de toma de decisiones de la línea ejecutiva de la clínica especialmente a los encargados de las áreas de admisión y laboratorio.

- ✓ **Procesos internos del área no están claramente definidos.** El director, así como las personas que toman decisiones no cuentan con la información oportuna y relevante para el análisis de la problemática respecto a lo que sucede en el proceso de admisión y de laboratorio, esto trae consigo que algunas actividades no se realicen adecuadamente, por ejemplo, como conocer el número de pacientes con un tipo de examen médicos específico y esto se debe a que no se cuenta con información rápida y confiable para el correcto análisis de la evolución de los procesos del área de admisión.

- ✓ **Costos por tomar decisiones fuera de tiempo.** La institución no cuenta con soluciones tecnológicas que den estadísticas y reportes en línea de cómo se está desarrollando la atención de sus pacientes. El no contar con estas soluciones generan una pérdida económica por no saber si están funcionando adecuadamente.

- ✓ **Falta de personal capacitado en análisis de información.** La clínica no cuenta con un área o personal específico que se encargue de tareas de analizar información procedente de las fuentes transaccionales y el área de TI no realiza este trabajo, por lo que se cae en el reprocesamiento de información y en otras

ocasiones en obtener información irrelevante para el área de admisión y laboratorio.

Estas limitaciones hacen que la información que se obtiene a través de los sistemas transaccionales sean solo datos en un repositorio sin ningún valor. Los análisis van perdiendo valor por el tiempo que toma en organizarlos manualmente, por la manipulación de los datos y posibles errores humanos que ello ocasiona.

- ✓ **Falta de herramienta que apoye a la toma de decisiones.** También la falta de una herramienta de apoyo a la toma de decisiones hace que estas sean lentas y fuera de tiempo, por lo que es necesario contar con una solución de BI para solucionar esta falta.

En este contexto el presente trabajo de tesis propone implementar una solución de inteligencia de negocios como una herramienta que se ajuste a los requerimientos reales de la institución añadiendo el análisis dimensional y los indicadores necesarios para el aprovechamiento de la información de las personas involucradas en las tomas de decisiones de institución.

1.4. Definición del problema

En el Centro Médico Pacífico del Norte se tiene de necesidad de contar con una herramienta que permita tener un mejor análisis de información en las áreas de Admisión y Laboratorio y que le permita dar respuestas inmediatas a las preguntas o requerimientos de la organización y poder tomar mejores decisiones en dicha área.

1.5. Formulación del problema

¿Cómo mejorar el análisis de la información para el soporte en la toma de decisiones en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte usando los Sistemas de Información?

1.6. Formulación de la hipótesis

El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss para la mejora en el análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo.

1.7. Objetivos del estudio

El **Objetivo general** es:

Mejorar el análisis de información para el soporte en la toma de decisiones en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo usando la arquitectura tecnológica de Pentaho BI y basándose en parte de la metodología de Larissa Moss.

Los **objetivos específicos** son los siguientes:

1. Obtener y analizar los requerimientos de las áreas críticas del negocio.
2. Desarrollar el proceso ETL utilizando como herramienta Pentaho BI
3. Diseñar la estructura de que tendrán las dimensiones y cubos pertinentes para crear el diseño de vista de análisis y la navegabilidad que tendrá la información.
4. Utilizar la herramienta Pentaho Report Designer para crear y diseñar los reportes correspondientes al Modelo implementado.

1.8. Justificación del estudio

1.8.1. Importancia de la investigación

- La solución de inteligencia de negocios que se realizara permitirá recolectar, depurar y transformar los datos que se encuentren en la base de datos transaccional de la institución evitando la redundancia de datos.
- La importancia de tener en la institución una solución de inteligencia de negocios es que este englobara los sistemas donde se pueda obtener información de distintas áreas de donde se requiera indicadores para tomar una decisión y así crear una ventaja competitiva en el mercado.

- La solución de inteligencia de negocios representa un apoyo de gran importancia al momento de realizar informes y reportes ya que el personal encargado de realizar la tarea de brindar los reportes prioriza realizar esta tarea a tiempo para poder brindar la información a la gerencia para poder definir las estrategias y las decisiones que tomaran los gerentes.
- Tener una buena consistencia en los datos es un factor muy importante hoy en día para cualquier empresa sea empresa grande o una pyme ya que de acuerdo a esto se podrá tener una mejor fiabilidad de la información que se procesara para la toma de decisiones.
- Los datos estadísticos que se generarán de la herramienta de inteligencia de negocios permitirán determinar medidas de cambio y mejora provenientes de los datos procesados que podrá brindar una realidad más clara a la situación actual del negocio.

1.8.2. Viabilidad de la investigación

- Es viable porque el equipo de trabajo cuenta con el conocimiento para poder realizar el proyecto que se ha propuesto contando con la experiencia en conocimientos del asesor en el tema.
- Es factible porque se cuenta con el acceso a los datos y la información de la institución y conocimiento de los procesos siendo primordial esto para el desarrollo de la herramienta de BI.
- El objetivo del proyecto es implementar indicadores que ayude en el soporte de las decisiones que generara la reducción del tiempo en obtener informes rápidamente, tener una personalización de realizar los informes de acuerdo a los requerimientos de la empresa y así poder brindárselas a la gerencia para la toma de decisiones.
- El tomar decisiones correctas basadas en la información integra que nos brinda la herramienta de BI generara la optimización de costos y crecimiento de la empresa por ello la inversión que se realizara es justificada.

1.8.3. Aportes

El desarrollo de esta investigación generará considerables beneficios a la institución, entre los cuales tenemos:

- Información correcta y oportuna para que la institución pueda tomar decisiones acertadas.
- Confianza en las personas, especialmente en los tomadores de decisiones que trabajan con esta institución.
- Celeridad en la obtención de información sobre los que está sucediendo en los procesos de admisión y de laboratorio de la clínica.
- Aporte en conocimientos adquiridos en este estudio que permitirá desarrollar proyectos similares en alguna otra institución, facilitando la implementación de un sistema de soporte de decisiones basados en Inteligencia de negocios.

2. MARCO TEÓRICO

2.1. ANTECEDENTES

- **Autores:** Sánchez Guevara, Omar Antonio

Título de Investigación: Modelo de inteligencia de negocio para la toma de decisiones en la empresa San Roque S.A., Trujillo 2014 UPAO

Descripción:

El presente trabajo muestra una propuesta de Modelo de Inteligencia de Negocios que asegure una mejora en el proceso de ventas, que permita transformar información clave del negocio en acciones concretas traduciéndose en beneficios tangibles que les proporcione una ventaja competitiva a los tomadores de decisiones. Para lograrlo se identificaron y compararon las características más relevantes de las principales metodologías para el desarrollo de modelos de inteligencia de negocios tanto tradicionales como de la actual tendencia de BI denominada Big Data, luego se propuso un Modelo de Inteligencia de Negocios que cuenta con las siguientes fases: planificación, modelo del negocio, análisis, diseño, construcción e implementación. Para el despliegue del modelo se utilizó la plataforma BI Pentaho EE versión 4.8 y los componentes de Big Data: Hadoop y Hive. Así se implementó el Modelo de Inteligencia de Negocio en el área de ventas de la empresa San Roque S.A. Se analizaron los datos obtenidos a través de la implementación del modelo, con lo que se demostró mejoras en los indicadores definidos para su evaluación.

- **Autores:** Joao Danilo Chavez Yrigoyen

Título de Investigación: Análisis, Diseño e Implementación una Solución de Inteligencia de Negocios orientada a Controlar Los Procesos de Generación y emisión del DNI (Documento Nacional De Identidad) en el Reniec (Registro Nacional de Identificación y Estado Civil), Lima 2014 PUCP.

Descripción:

El presente trabajo tiene como propuesta una solución de Inteligencia de Negocios como alternativa para manejar la gran cantidad de información que posee

RENIEC. Mediante el uso de la herramienta se organizará adecuadamente la información generada en cada una de las unidades orgánicas (Gerencias, Jefaturas Regionales y otras) que gestionan el proceso de generación y emisión del DNI, y de ésta forma se brindarán indicadores que son de mucha utilidad a las diferentes unidades orgánicas. De esta manera las personas que deseen acceder a estos indicadores solo deberán utilizar la herramienta. Teniendo como objetivos: Realizar el análisis y el modelado de los procesos de generación y emisión del DNI, Diseñar la arquitectura lógica de la solución y la configuración de los datos a usar, Realizar la construcción de los procesos ETL para la extracción, transformación y carga de datos desde un sistema transacción a un sistema de inteligencia de negocios, Desarrollar los diferentes reportes para la explotación de la información del Data Mart y Realizar la prueba de la herramienta configurada con la data de RENIEC. Teniendo como principal conclusión que los esquemas de análisis, gráficos de barra, reportes elaborados permitieron mostrar los indicadores que brindan información relevante a los usuarios respectivos. Y la forma dinámica en que se pueden ir modificando ayuda mucho al análisis de los datos.

- **Autores:** Analí Del Rosario López Palacios

Título de Investigación: Implementación de una Solución de Inteligencia de Negocios basado en el Algoritmo de Serie Temporal para la Mejora del Proceso de Toma de Decisiones Gerenciales en una empresa Comercial, Chiclayo 2015
UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO

Descripción:

El presente trabajo de investigación tuvo como objetivo mejorar el proceso de toma de decisiones de la empresa “El Ofertón SAC”. Dentro de los problemas identificados se encontró que los reportes solicitados por la gerencia eran entregados con demoras debido a que el personal encargado de realizar dichos reportes tomaba mucho tiempo en la culminación de los mismos, generando adicionalmente, un costo para la empresa. Los reportes generados presentaban un nivel de calidad de la información bajo, por lo que varias veces los clientes solicitaban productos que no tenían en stock mientras existían otros productos que durante grandes periodos de tiempo no tenían movimientos de ventas, generando

que la gerencia tuviera un nivel de satisfacción bajo por la información brindada en los reportes.

La implementación de la Solución de Inteligencia de Negocios basado en el algoritmo de Serie Temporal ha mejorado significativamente el proceso de toma decisiones en la empresa “El Ofertón SAC”. La solución redujo en un 96,25% el tiempo promedio en la elaboración de informes solicitados por la gerencia, se incrementó el nivel de calidad de la información solicitada de 1,8667 a 4,600 y se incrementó el nivel de satisfacción sobre la información solicitada de 2,2000 a 4,5333, ambos basados en una escala de 5 puntos del estilo Likert.

- **Autores:** Rosales Barreto, Ciro; Cruz Palacios, Jessy

Título de Investigación: Implementación de un sistema web business intelligence para la gestión de indicadores en la empresa Crea a través de la plataforma Pentaho, Lima 2015 UPC

Descripción:

La propuesta de este trabajo tiene como objetivo Implementar un sistema web que permita analizar los datos de los indicadores estratégicos de Crea+, con la finalidad de obtener información analítica que ayude a la toma de decisiones. Para lograr esto el trabajo propone los siguientes objetivos específicos: Identificar y modelar los procesos que realizan la creación de informes para los indicadores de Crea+, Implementar una aplicación web que realice el análisis de los datos contenidos en una base de datos dimensional; creando informes con semáforos y cuadros de mando; para finalmente hacer el Despliegue de la aplicación web en servidores de la empresa ITEXpert.

La solución permite que los datos se encuentren en un repositorio único. Con lo cual los usuarios pueden acceder a datos históricos. Adicionalmente, el acceso a los reportes es en tiempo real. Además permite reducir tiempos de espera al estar automatizada la creación de reportes.

- **Autores:** Angeles Pacheco, Vicente Miguel

Título de Investigación: Aplicativo Datamart y La Agilización de la Toma de Decisiones en el Departamento de Farmacia del Hospital Eleazar Guzmán Barrón

– Nuevo Chimbote, Chimbote 2015 UCV.

Descripción:

El presente trabajo permite agilizar el proceso de toma de decisiones en del departamento de farmacia de dicho hospital, a través de herramientas y tecnologías de inteligencia de negocios, que convertirán datos en información útil. El aplicativo es desarrollado bajo la realización de cada uno de las etapas de la metodología de Ralph Kimball. El aplicativo propuesto muestra todos los requerimientos y necesidades del área de farmacia, asegurándose que el proceso de toma de decisiones sea de manera oportuna y eficaz. Para llevarlo a cabo se planteó: disminuir el tiempo en la elaboración de reportes administrativos, disminuir el tiempo en la búsqueda de información personalizada de farmacia e incrementar el nivel de satisfacción en el proceso de toma de decisiones de los directivos de farmacia.

2.2. DEFINICIONES

2.2.1. INTELIGENCIA DE NEGOCIOS

Según el autor (Cano, 2007) El objetivo básico de la Business Intelligence es apoyar de forma sostenible y continuada a las organizaciones para mejorar su competitividad, facilitando la información necesaria para la toma de decisiones. El primero que acuñó el término fue Howard Dresner que, cuando era consultor de Gartner, popularizó Business Intelligence o BI como un término paraguas para describir un conjunto de conceptos y métodos que mejoraran la toma de decisiones, utilizando información sobre que había sucedido (hechos). Mediante el uso de tecnologías y las metodologías de Business Intelligence pretendemos convertir datos en información y a partir de la información ser capaces de descubrir conocimiento. (Cano, 2007)

Los procesos de inteligencia de negocios se han convertido en una parte esencial de la gestión empresarial, ya que nos proporcionan información de mucha importancia para una organización:

Nos permiten saber con exactitud ¿quiénes son nuestros clientes?

- ¿dónde están los clientes?
- ¿cómo son los clientes?
- ¿qué pautas de comportamiento tienen?
- ¿qué es lo que quieren los clientes?

Así mismo Permiten conocer el funcionamiento de los procesos internos de una organización. Dan respuesta a preguntas importantes como:

- ¿cuánto cuesta cada uno de los productos o servicios que ofrece la empresa?
- ¿cuáles de esos productos o servicios son realmente rentables?
- ¿dónde está perdiendo la empresa?
- ¿cuánto cuesta conseguir cada cliente?
- ¿cuáles departamentos de la empresa funcionan correctamente?
- ¿qué servicios puede externalizar la empresa?

2.2.2. VENTAJAS DE LA INTELIGENCIA DE NEGOCIOS

La adquisición de un sistema de BI debe proporcionar las siguientes ventajas: (Rio, 2006)

- Una plataforma de tecnología integrada que se añadiría a las inversiones ya realizadas por una organización, para proporcionar información de alta calidad a cada computadora o servidor de cada departamento de una empresa, añadiendo valor en cada paso del proceso y proporcionando una versión única de la realidad.
- Acceso amplio y potenciado a las capacidades de análisis ya conocidas que ayudan a conocer el pasado de una organización para controlar y comunicar el presente y predecir el futuro con fiabilidad.
- Interfaces de usuario personalizadas que “se adapten a cada tarea”, concebida y diseñadas para todos los niveles de

experiencia y patrones de uso de los usuarios de la información (como expertos, ejecutivos, gerentes, tecnólogos, etc.).

- Una gama de soluciones para satisfacer la demanda de información y generación de informes de diferentes sectores de actividad (servicios financieros, fabricación, telecomunicaciones, salud, etc.) y de toda la empresa (gestión del rendimiento empresarial, inteligencia de clientes, inteligencia financiera, gestión del capital humano, inteligencia de la cadena de suministro, etc.)

2.2.3. OLAP

El procesamiento analítico en línea permite obtener acceso a datos organizados y agregados de orígenes de datos empresariales, organiza subconjuntos de datos con una estructura multidimensional de manera que represente un significado especial o responda a una pregunta en particular (Roussel, 2006) Estas herramientas soportan el análisis interactivo de la información de resumen, soportando muchas tareas de agrupación de datos que no pueden realizarse empleando las facilidades básicas de agregación y agrupamiento .

2.2.4. Tipos de sistemas OLAP

Tradicionalmente, este sistema se clasifica según las siguientes categorías:

- ROLAP. Implementación que almacena los datos en un motor relacional. Típicamente, los datos son detallados, evitando las agregaciones y las tablas se encuentran normalizadas.
- MOLAP. Esta implementación almacena los datos en una base de datos multidimensional. Para optimizar los tiempos de

respuesta, el resumen de la información es usualmente calculado por adelantado.

- HOLAP. Almacena algunos datos en un motor relacional y otros en una base de datos multidimensional

A. **Esquema estrella:** Consiste en estructurar la información en procesos, vistas y métricas recordando a una estrella. Es decir, tendremos una visión multidimensional de un proceso que medimos a través de unas métricas. A nivel de diseño, consiste en una tabla de hechos (lo que en los libros encontraremos como **fact table**) en el centro para el hecho objeto de análisis y una o varias tablas de dimensión (dimensión table) por cada dimensión de análisis que participa de la descripción de ese hecho. En la tabla de hecho encontramos los atributos destinados a medir (cuantificar) el hecho: sus métricas. Mientras, en las tablas de dimensión, los atributos se destinan a elementos de nivel (que representan los distintos niveles de las jerarquías de dimensión) y a atributos de dimensión (encargados de la descripción de estos elementos de nivel). En el esquema en estrella la tabla de hechos es la única tabla del esquema que tiene múltiples join que la conectan con otras tablas (foreign keys hacia otras tablas). El resto de tablas del esquema (tablas de dimensión) únicamente hacen join con esta tabla de hechos. Las tablas de dimensión se encuentran además totalmente desnormalizadas, es decir, toda la información referente a una dimensión se almacena en la misma tabla. (Curto, 2007)

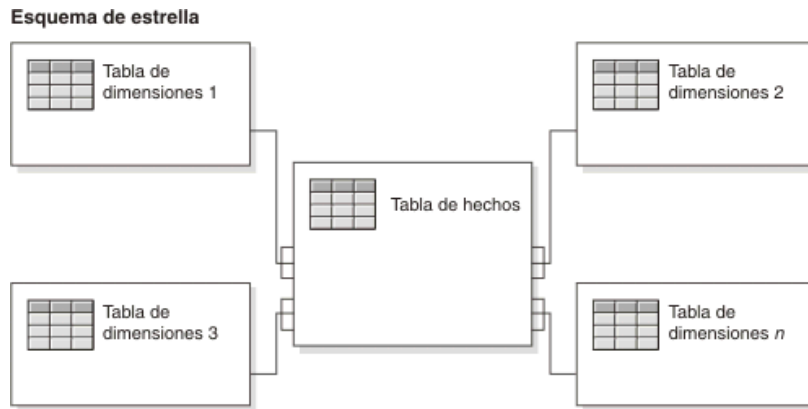


Figura 1 . Esquema estrella

Fuente (IBM, 2017)

B. Esquema copo de nieve: El esquema en copo de nieve es un esquema de representación derivado del esquema en estrella, en el que las tablas de dimensión se normalizan en múltiples tablas. Por esta razón, la tabla de hechos deja de ser la única tabla del esquema que se relaciona con otras tablas, y aparecen nuevas join gracias a que las dimensiones de análisis se representan ahora en tablas de dimensión normalizadas. En la estructura dimensional normalizada, la tabla que representa el nivel base de la dimensión es la que hace join directamente con la tabla de hechos. La diferencia entre ambos esquemas (star y snowflake) reside entonces en la estructura de las tablas de dimensión. Para conseguir un esquema en copo de nieve se ha de tomar un esquema en estrella y conservar la tabla de hechos, centrándose únicamente en el modelado de las tablas de dimensión, que, si bien en el esquema en estrella se encontraban totalmente normalizadas, ahora se dividen en subtablas tras un proceso de normalización. Es posible distinguir dos tipos de esquemas en copo de nieve, un snowflake completo (en el que todas las tablas de dimensión en el esquema en estrella aparecen ahora normalizadas en el snowflake) o un snowflake parcial (sólo se lleva a cabo la normalización de algunas de ellas).

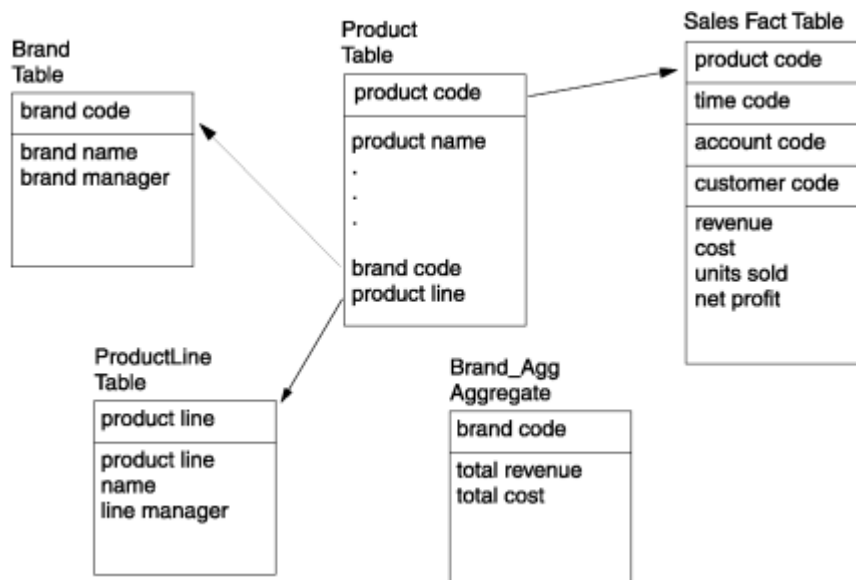


Figura 2. Esquema copo de nieve

Fuente (IBM, 2017)

2.2.5. DATAWAREHOUSE

Es el proceso de extraer datos de distintas aplicaciones (internas y externas), para que una vez depurados y especialmente estructurados sean almacenados en un depósito de datos consolidado para el análisis del negocio. Requiere una combinación de metodologías, técnicas, hardware y los componentes de software que proporcionan en conjunto la infraestructura para soportar el proceso de información (Stackowiak, Rayman, & Rick, 2007) La estructura que se defina debe reflejar las necesidades y características del negocio, sus departamentos, equipos de trabajo y directivos, esto permitirá responder a interrogantes generados al tratar de tomar las decisiones (Ian H. witten, 2000) y con el tiempo se va convirtiendo en la memoria corporativa (Wang, 2009) ; describiendo el pasado y el presente de la empresa. Data Warehouse desglosa, resume, ordena y compara, pero no descubre, ni predice. (Flores, 2004) .

Para la construcción de un Data Warehouse se establecen tres etapas; la primera está dedicada a examinar el esquema Entidad Relación de la base de datos operacional, generando los esquemas multidimensionales candidatos.

La segunda etapa, consiste en recoger los requisitos de usuario por medio de entrevistas, para obtener información acerca de las necesidades de análisis de estos, y la tercera etapa, contrasta la información obtenida en la segunda etapa, con los esquemas multidimensional candidatos formados en la primera etapa generando así, una solución que refleja los requisitos de usuario (sanchez, 2008)

Por otra parte implementar una solución de este tipo, ocasiona un costo que no todas las organizaciones están dispuestas a pagar (debido a sus capacidades de inversión), es por eso que los promotores del proyecto dentro de la empresa deben persuadir a los directivos y compañeros de trabajo, una buena alternativa de hacerlo es mediante el uso de técnicas administrativas, que permitan conocer a los directivos como se puede establecer el retorno de la inversión del proyecto equiparando inversión contra beneficios. (Arturo L, 2001)

2.2.6. DATAMART

Un Datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un datamart puede ser alimentado desde los datos de un Datawarehouse, o integrar por sí mismo un compendio de distintas fuentes de información. Por tanto, para crear el datamart de un área funcional de la empresa es preciso encontrar la estructura óptima para el análisis de su información, estructura que

puede estar montada sobre una base de datos OLTP, como el propio Datawarehouse, o sobre una base de datos OLAP. La designación de una u otra dependerá de los datos, los requisitos y las características específicas de cada departamento.

2.2.7. ETL

- **Extracción:** En este proceso se extraerán los datos desde múltiples fuentes y se los integra. Primero se identificarán las fuentes, luego se seleccionará las fuentes a usar, y finalmente se unen los datos de las fuentes para ser extraídos.
- **Transformación:** Conjunto de técnicas y herramientas para cambiar los datos extraídos de las diferentes fuentes, es decir quitar y añadir información, corregir si hay errores y uniformizar los datos (asignar un formato estándar); en otras palabras, refinarlo y agregarles calidad.
- **Carga:** En este proceso se Suben los datos transformados al almacén de datos a través de métodos, ya que esto controlará la Actualización de datos de acuerdo al tiempo programado para Su elaboración. Finalizado el proceso de Carga se debería tener en el Data Warehouse datos de calidad, los cuales servirán de soporte para la toma de decisiones.

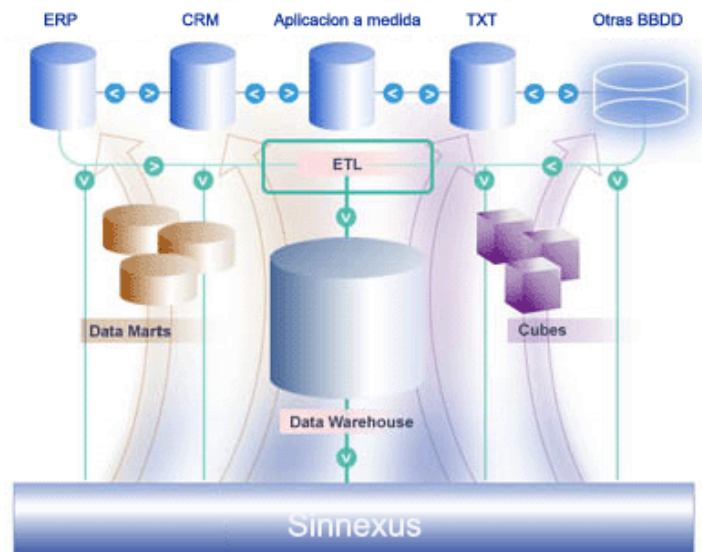


Figura 3. Proceso ETL

Fuente (Sinnexus, 2014)

2.2.8. SISTEMA DE SOPORTE A LAS DECISIONES DSS

En la empresa no todas las decisiones son de carácter recurrente, sino que algunas se presentan muy pocas veces o incluso una sola vez. Los dss son instrumentos para abordar problemas de definición o estructuración menos precisa, y de carácter más esporádico. Los sistemas de apoyo a la decisión ayudan a los directivos que deben tomar decisiones no estructuradas. Una decisión se considera no estructurada si no existen procedimientos claros para tomarla y tampoco es posible identificar, con antelación, todos los factores que deben considerarse en la decisión. Hay que decir que todos los sistemas de información sirven de apoyo a la toma de decisiones, aunque sea de forma indirecta. Los dss son el tipo de sistemas de información desarrollados expresamente para servir de apoyo en el proceso de toma de decisiones. Estos sistemas facilitan un diálogo con el usuario que está considerando soluciones alternativas a un problema, y el sistema proporciona modelos construidos para la presentación de la información y acceso a bases de datos. Los sistemas de apoyo a la decisión son interactivos y su objetivo es la ampliación del razonamiento humano en la resolución de

problemas particulares de toma de decisiones no estructuradas. Este tipo de sistemas se centra en los procesos de decisión y deberá proporcionar de forma fácil, rápida y exacta hechos importantes relacionados con la decisión a tomar y facilitando el acceso interactivo a medios de tratamiento que se utilizan creativamente y que permiten explorar las distintas posibilidades, suministrando las informaciones necesarias para responder a los problemas planteados. Al utilizar un dss, un directivo considerará un número posible de escenarios con el razonamiento: ¿Qué pasaría si?. Por ejemplo, un directivo que estuviera estudiando fijar el precio de un nuevo producto puede utilizar un sistema de apoyo a la decisión en el área marketing. El sistema contendrá un modelo que relacione varios factores, tales como el precio del producto, el coste de los materiales, los gastos de promoción, que afecten a la estimación de beneficios de las ventas del producto para un período de unos cinco años. Variando el precio del producto en el modelo, el directivo puede comparar resultados previstos y, entonces, seleccionar un precio (Alcami, Carañana, & Herrando, 2009).

2.2.9. FUNCIÓN Y CARACTERÍSTICAS DEL DSS

El DSS es una de las herramientas más emblemáticas del Business Intelligence ya que, entre otras propiedades, permiten resolver gran parte de las limitaciones de los programas de gestión. Estas son algunas de sus características principales:

Informes dinámicos, flexibles e interactivos, de manera que el usuario no tenga que ceñirse a los listados predefinidos que se configuraron en el momento de la implantación, y que no siempre responden a sus dudas reales. No requiere conocimientos técnicos. Un usuario no técnico puede crear nuevos gráficos e informes y navegar entre ellos, haciendo drag&drop o drill through. Por tanto, para examinar la información disponible o crear nuevas métricas no es imprescindible buscar auxilio en el departamento de informática (Alcami, Carañana, & Herrando, 2009).

- **Integración entre todos los sistemas/departamentos de la compañía.**
El proceso de ETL previo a la implantación de un Sistema de Soporte a la Decisión garantiza la calidad y la integración de los datos entre las diferentes unidades de la empresa. Existe lo que se llama: integridad referencial absoluta.
- **Drag&drop:** Es una expresión informática que se refiere a la acción de mover con el ratón objetos de una ventana a otra o entre partes de una misma ventana.
- **Rapidez en el tiempo de respuesta:** ya que la base de datos subyacente suele ser un Datawarehouse corporativo o un datamart, con modelos de datos en estrella o copo de nieve. Este tipo de bases de datos están optimizadas para el análisis de grandes volúmenes de información.
- **Drill Through:** Es el análisis de los datos que va desde un cubo OLAP en la base de datos relacional.
- **Disponibilidad de información histórica:** En estos sistemas está a la orden del día comparar los datos actuales con información de otros períodos históricos de la compañía, con el fin de analizar tendencias, fijar la evolución de parámetros de negocio... etc.

2.2.10. Pentaho BI

Pentaho es una herramienta de Business Intelligence desarrollada bajo la filosofía del software libre para la gestión y toma de decisiones empresariales. Es una plataforma compuesta de diferentes programas que satisfacen los requisitos de BI. Ofreciendo soluciones para la gestión y análisis de la información, incluyendo el análisis multidimensional OLAP, presentación de informes, minería de datos y creación de cuadros de mando para el usuario. La plataforma ha sido desarrollada bajo el lenguaje de programación Java y tiene un ambiente de implementación también basado en Java, haciendo así que Pentaho sea una solución muy flexible al cubrir una alta gama de necesidades empresariales. (Gravitar, 2017)

Características de Pentaho BI :

Pentaho Business Intelligence abarca las siguientes áreas de reporte:

- **Pentaho reporting:** Pentaho Reporting es un potente generador de informes: Permite la distribución de los resultados del análisis en múltiples formatos.
- **Pentaho Data Integration:** Muchas organizaciones tienen información disponible en aplicaciones y base de datos separados. Pentaho Data Integration abre, limpia e integra esta valiosa información y la pone en manos del usuario. Provee una consistencia, una sola versión de todos los recursos de información, que es uno de los más grandes desafíos para las organizaciones TI hoy en día. Pentaho Data Integration permite una poderosa ETL (Extracción, Transformación y Carga). El uso de kettle permite evitar grandes cargas de trabajo manual frecuentemente difícil de mantener y de desplegar.
- **Pentaho BI Plataforma:** La plataforma de Pentaho BI provee de servicios críticos incluyendo programación, seguridad, integración automatización y flujo de trabajo. Proporcionando habilidades a los usuarios finales de Pentaho y proviendo un lugar central para administrar y mantener el despliegue de la empresa BI (Pentaho, 2017)
 - Integración con procesos de negocio.
 - Administra y programa reportes.
 - Administra seguridad de usuarios.

Componentes de Pentaho

- ✓ Plataforma 100% J2EE, asegurando la escalabilidad, integración y portabilidad.
- ✓ Servidor: puede correr en servidores compatibles con J2EE como JBOSS AS, WebSphere, Tomcat, WebLogic y Oracle AS.
- ✓ Base de datos: vía JDBC, IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR Teradata, Firebird.
- ✓ Sistema operativo: no hay dependencia. Lenguaje interpretado.

- ✓ Lenguaje de programación: Java, JavaScript, JSP, XSL (XSLT/XPath/XSL-FO).
- ✓ Interfaz de desarrollo: Java SWT, Eclipse, Web-based.
- ✓ Repositorio de datos basado en XML.
- ✓ Todos los componentes están expuestos vía Web Services para facilitar la integración con Arquitecturas Orientadas a Servicios (SOA). (Pentaho, 2017)

2.3. METODOLOGIA PARA EL DESARROLLO DEL PROYECTO

Existen diversas tecnologías y metodologías para el desarrollo de un BI. Nosotros optamos por la metodología de Larissa Moss:

2.3.1. BUSINESS INTELLIGENCE ROADMAP - LARISSA MOSS & SHAKU ATRE

(Moss & Atre, 2003) Casi todo tipo proyectos de ingeniería, como ingeniería estructural o ingeniería software pasa por seis etapas entre el inicio y ejecución.

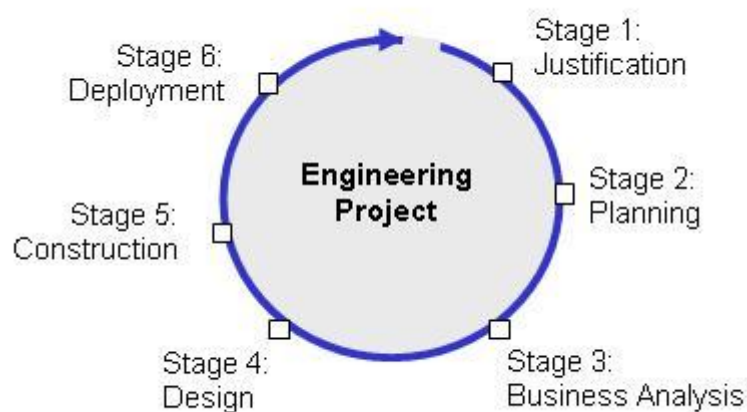


Figura 4: Etapas de la Metodología (Moss & Atre, 2003)

Como lo muestra la figura, las etapas de ingeniería son iterativas. Una vez

desplegado un producto se mejora continuamente, estas mejoras se basan en la retroalimentación de los grupos del negocio que utiliza el producto. Cada iteración produce la realización de un nuevo producto (versión) el cual evoluciona y madura.

- ✓ **Etapa de Justificación:** evaluar las necesidades del negocio que dan origen al proyecto de ingeniería.
- ✓ **Etapa de Planificación:** desarrollar planes estratégicos y tácticos, que establecen cómo se desplegara y llevará a cabo el proyecto.
- ✓ **Etapa de Análisis del negocio:** realizar un análisis detallado de los problemas y oportunidades del negocio para adquirir una comprensión sólida de los requisitos y llegar a una solución (producto).
- ✓ **Etapa de Diseño:** concebir un producto que resuelva el problema de negocio o provea oportunidades de negocio.
- ✓ **Etapa de Construcción:** construir el producto, el cual debe proporcionar un retorno de la inversión dentro de un lapso de tiempo definido.
- ✓ **Etapa de Implementación:** implementar el producto final, y medir su efectividad para determinar si la solución no cumple, cumple o excede con el retorno de la inversión esperado.

Dentro de cada etapa, ciertos pasos son llevados a cabo para que el proyecto llegue a su conclusión. En el plan de trabajo de BI, se describen 16 pasos para el desarrollo de estas seis etapas, como se indica a continuación.

Etapa de justificación

Paso 1: Evaluación del caso de negocio: se define el problema o la oportunidad del negocio, y se propone una solución de BI. Cada lanzamiento de una aplicación de BI debe justificar su costo y definir claramente sus beneficios, o la solución de un problema de negocio o el aprovechamiento de una oportunidad de negocio.

Etapa de planificación

Paso 2: Evaluación de infraestructura de la empresa: Ya que las aplicaciones de BI son iniciativas de toda la organización, esta debe crear una infraestructura para apoyarlas. Algunos componentes de la infraestructura pueden estar ya en su lugar, antes de que el primer proyecto de BI este en marcha. Otros componentes de la infraestructura pueden ser desarrollados con el tiempo, como parte de los proyectos de BI. Una infraestructura de la organización tiene dos componentes:

La infraestructura técnica: que incluye hardware, software, middleware, sistemas de gestión de bases de datos, sistemas operativos, componentes de red, repositorios de metadatos, utilidades, etc.

Infraestructura no técnica: que incluye estándares de metadatos, estándares de minería de datos, el modelo lógico empresarial (en evolución), metodologías, directrices, procedimientos de prueba, control de cambios y procesos, procedimientos para tareas administrativas y resolución de problemas, entre otros.

Paso 3: Planificación de Proyectos: los proyectos de BI son extremadamente dinámicos. Los cambios en el personal, en el presupuesto, en la tecnología, en los representantes del negocio y los patrocinadores, pueden afectar seriamente el éxito del proyecto. Por lo tanto, la planificación del proyecto debe ser detallada, y el progreso efectivo debe ser observado de cerca y reportado.

Etapa de análisis del negocio

Paso 4: Definición de requisitos del proyecto: administrar el alcance del proyecto es una de las tareas más difíciles en el transcurso del proyecto de BI. La necesidad de tener todo al instante es difícil de

reducir, pero que se reduzca esta necesidad es uno de los aspectos más importantes en la negociación de los requisitos para cada entrega. Los integrantes de los equipos del proyecto deben saber que los requisitos cambian durante todo el ciclo de desarrollo, y los directivos deben conocer más sobre las posibilidades y las limitaciones de la tecnología de BI durante el desarrollo del proyecto.

Paso 5: Análisis de Datos: el mayor desafío de todos los proyectos de BI es la calidad de los datos de origen. Los malos hábitos desarrollados en las últimas décadas son difíciles de romper, y los daños provenientes de estos resultan muy caros, consumen mucho tiempo, y es tedioso encontrarlos y corregirlos. Además, el análisis de datos en el pasado se limitaba a la vista de una línea de negocio y nunca fue consolidada o conciliada con otros puntos de vista de la organización. Este paso requiere un porcentaje significativo del tiempo dedicado al calendario del proyecto completo.

Paso 6: Prototipo de la aplicación: el análisis de los resultados funcionales, que solía ser llamado análisis del sistema, se logra mediante los prototipos, por lo que se puede combinar con el diseño de aplicaciones. Las nuevas herramientas y lenguajes de programación permiten a los desarrolladores probar o refutar con relativa rapidez un concepto o una idea. Los prototipos también permiten a los empresarios ver el potencial y los límites de la tecnología, lo que les da la oportunidad de ajustar los requisitos del proyecto y sus expectativas.

Paso 7: Análisis de repositorio de metadatos (Datawarehouse): tener más herramientas significa tener más metadatos técnicos, además de los metadatos del negocio que suelen ser capturados mediante la ingeniería de software asistida por un ordenador de modelado de

herramientas (CASE). Los metadatos técnicos necesitan ser asignados a los metadatos del negocio, y todos los metadatos deben ser almacenados en un repositorio de metadatos, estos últimos, pueden ser con licencia (comprados) o construidos. Los requisitos para que los tipos de datos sean capturados y almacenados, deben ser documentados en un modelo lógico de metadatos. Cuando se tienen las licencias de un producto de repositorio de metadatos, los requisitos documentados en este modelo lógico de metadatos deben ser comparados con el modelo metadatos del proveedor, si lo proporciona. Además, los requisitos para la entrega de los metadatos a la comunidad empresarial tienen que ser analizados.

Etapa de diseño

Paso 8: Diseño de bases de datos: uno o más objetivos de la base de datos de BI es almacenar de forma general y detallada los datos del negocio, dependiendo de las exigencias de la comunidad empresarial. No todos los requisitos de información son estratégicos y no todos son multidimensionales. Los esquemas de diseño de bases de datos deben coincidir con los requisitos de acceso a la información de la comunidad empresarial.

Paso 9: Diseño Extraer/Transformar/Cargar (ETL): el proceso ETL es el más complicado de todo el proyecto de BI, también es el menos glamoroso. Las ventanas de procesamiento ETL (ventanas de proceso por lotes) usualmente son pequeñas, sin embargo, debido a la mala calidad de la fuente de datos por lo general requiere mucho tiempo para ejecutar la transformación y los programas de limpieza. Acabar el proceso de ETL dentro del calendario previsto es un desafío para la mayoría de las organizaciones.

Paso 10: Diseño del repositorio de metadatos (Datawarehouse): si un repositorio de metadatos es comprado, lo más probable es que tenga que ser mejorado con características que fueron documentadas en el modelo lógico de metadatos, pero estas no se proporciona con el producto. Si se está construyendo un repositorio de metadatos, se debe tomar la decisión de si se diseña el repositorio de metadatos de la base de datos basado en entidad-relación u orientado a objetos. En cualquier caso, el diseño tiene que cumplir los requisitos del modelo lógico de metadatos.

Etapas de construcción

Paso 11: Desarrollo Extraer/Transformar/Cargar (ETL): muchas herramientas están disponibles para el proceso de ETL, algunas son sofisticadas y otras sencillas. Dependiendo de los requisitos para la limpieza y transformación de datos desarrollados en el paso 5, Análisis de Datos y en el Paso 9, Diseño ETL, una herramienta de ETL puede o no ser la mejor solución. En cualquier caso, se requiere con frecuencia el pre-procesamiento de los datos y la creación de ampliaciones para complementar las capacidades de la herramienta de ETL.

Paso 12: Desarrollo de Aplicaciones: una vez que el prototipo concretó los requisitos funcionales, el verdadero desarrollo del acceso y el análisis de la aplicación puede empezar. El desarrollo de la aplicación puede ser una simple cuestión de la finalización de un prototipo operativo, o puede ser un esfuerzo de desarrollo que esté más involucrado con diferentes y más robustas herramientas de acceso y análisis. En ambos casos las actividades de desarrollo de aplicación front-end son realizadas generalmente en paralelo con las actividades de desarrollo de ETL back-end y el desarrollo del repositorio de metadatos.

Paso 13: Minería de datos: muchas organizaciones no utilizan el ambiente de BI en toda su extensión. Las aplicaciones de BI a menudo son limitadas a prescribir informes, algunos de los cuales incluso no son los nuevos tipos de informes, pero reemplazan los informes viejos. El retorno de la inversión real proviene de la información oculta en los datos de la organización, que sólo se puede descubrir con las herramientas de minería de datos.

Paso 14: Desarrollo del repositorio de metadatos: si se toma la decisión de construir un repositorio de metadatos en lugar de comprarlo, un equipo independiente se debe encargar del proceso de desarrollo. Esto se convierte en un sub-proyecto considerable en el proyecto global de BI.

Etapa de despliegue

Paso 15: Implementación: una vez el equipo ha probado a fondo todos los componentes de la aplicación de BI, libera las bases de datos y aplicaciones. La formación está prevista para todo el personal del negocio y para otras personas que también utilizaran la aplicación de BI y el repositorio de metadatos. Las funciones de soporte que comienzan, incluyen operaciones desde mesa de ayuda, mantenimiento de las bases de datos de destino de BI, programación y ejecución de trabajos por lotes ETL, monitoreo del desempeño y puesta a punto de bases de datos.

No es necesario realizar los pasos de desarrollo en secuencia, la mayoría de los equipos de proyecto los lleva a cabo en paralelo. No hay un orden natural de la progresión de una etapa de ingeniería a otra, sin embargo, existen ciertas dependencias entre algunas de las fases de desarrollo. Los pasos que se encuentran en el diagrama, de la figura N°5, unos sobre otros, se puede realizar de forma simultánea, mientras que los pasos que aparecen a la

derecha o a la izquierda de cada uno, se llevan a cabo de manera relativamente lineal (con menos coincidencia) a causa de sus dependencias.

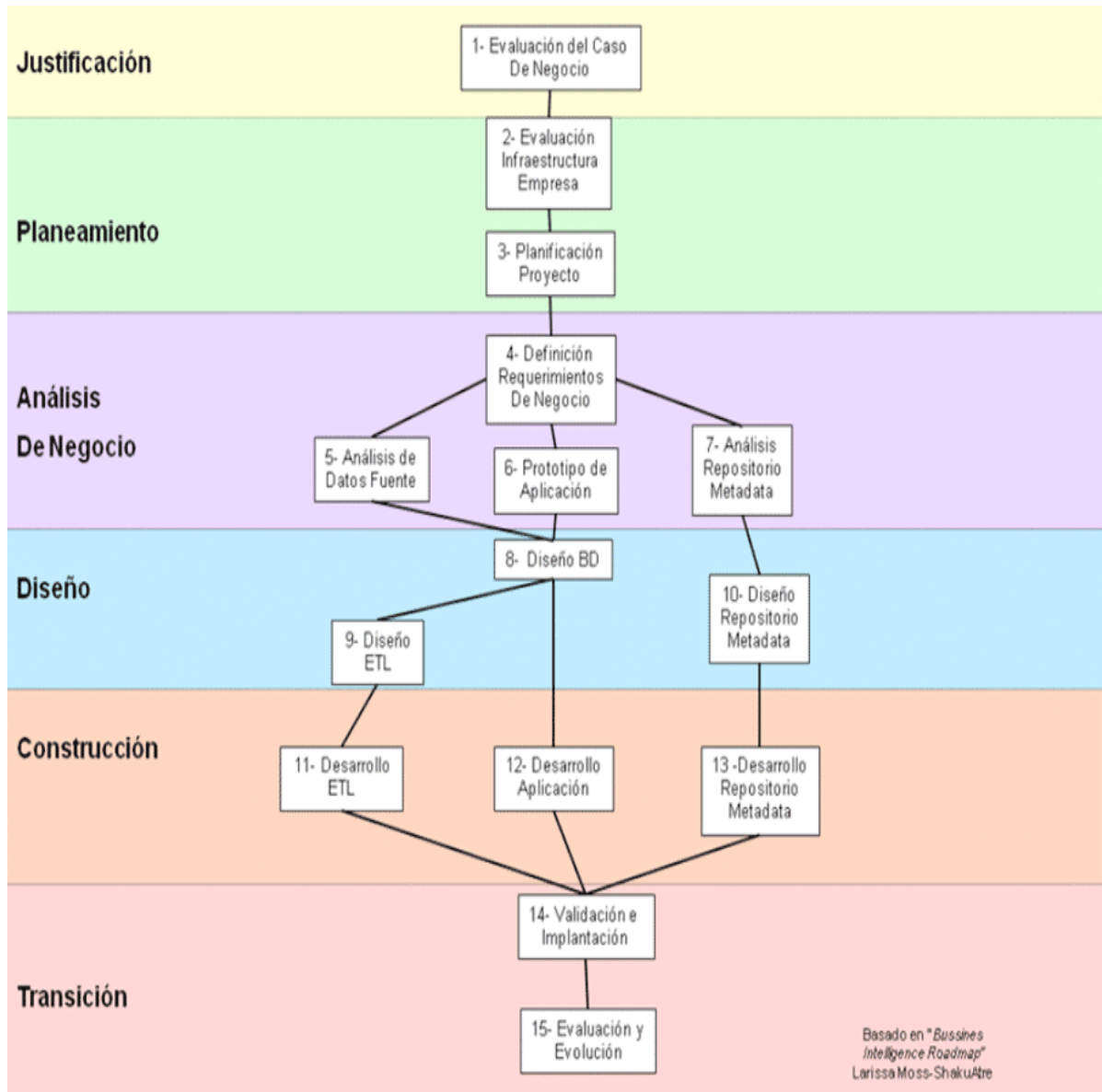


Figura 5: Etapas de la Metodología (Moss & Atre, 2003)

3. MATERIALES Y METODOS

3.1. MATERIAL

3.1.1. Población

Los tomadores de decisiones en las áreas de Admisión y Laboratorio, así como el director del Centro Médico Pacífico del Norte. (Director, Jefe de Admisión y de Laboratorios)

3.1.2. Muestra

Reportes utilizados por los tomadores de decisiones de las áreas de Admisión y Laboratorio.

3.1.3. Unidad de análisis

Información para dar soporte a la toma de decisiones del Área de Admisión y de laboratorio del Centro Médico Pacífico del Norte.

3.2. MÉTODO

3.2.1. Tipo de investigación

Aplicada.

3.2.2. Nivel de Investigación

Investigación Tecnológico - Profesional

3.2.3. Diseño de la Investigación

Se determinó usar el Diseño Pre-experimental:

Diseño del modelo pre-experimental	G -> O₁ -> X -> O₂
G (Grupo a investigar)	Personas que toman las decisiones.

Diseño del modelo pre-experimental	G -> O₁ -> X -> O₂
X (Tratamiento)	Aplicación de la solución de BI
O (Observación)	O ₁ : Observación pre-test
	O ₂ : Observación post-test

Tabla 1: Diseño de la Investigación

3.2.4. Variables de estudio y Operacionalización

- Independiente (VI): “Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss”.
- Dependiente (VD): “Análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacifico del Norte de la ciudad de Trujillo”.

3.1.3. Operacionalización de las variables

Tabla 2: Operacionalización de las variables

Variable	Dimensión	Indicador	Unidad de medida	Instrumento de Investigación
VI	Utilidad de la Herramienta	No. de informes generados	No. de informes	No. de informes
VD	Oportunidad	Tiempo en proporcionar la información	Segundos	Observación
	Satisfacción	Grado de satisfacción del proceso	% Satisfacción	Hoja resumen de grado de satisfacción

3.1.4. Técnicas e instrumentos de recolección de datos

3.1.4.1. Técnicas

Entre las técnicas para recopilar información dentro de este proceso investigativo podemos mencionar:

- ✓ La entrevista es una indagación directa con una persona directamente relacionada con el problema.
- ✓ El análisis documental que permitirá obtener información de para la documentación del proyecto y desarrollo de la solución.
- ✓ Observación.

3.1.4.2. Instrumentos

Se utilizarán los siguientes instrumentos en las técnicas mencionadas anteriormente:

- ✓ Cuestionarios.
- ✓ Ficha de Observación.

3.1.5. Técnicas de procesamiento y análisis de datos

3.1.5.1. Procesamiento de datos

Después de realizar el cuestionario a los tomadores de decisiones, “los datos serán ingresados en el Ms Excel” para la realización de análisis descriptivo. Las respuestas que “se obtengan se presentarán en tablas y gráficos dinámicos”, permitiendo conocer los resultados numéricos y porcentuales de las variables en estudio de acuerdo los objetivos.

3.1.5.2. Análisis de datos

Para el análisis de los datos “se utilizarán cuadros estadísticos descriptivos” (Pruebas hipótesis nula y alternativa).

Las Pruebas estadísticas de realizarán “con la Distribución de t de Student”, (por el tamaño de la muestra) ya que la información obtenida será analizada y mostrada por medio de cuadros y gráficos.

4. RESULTADOS: APLICACIÓN DE LA METODOLOGIA

4.1. JUSTIFICACION

4.1.3. EVALUACIÓN DEL CASO DEL NEGOCIO

El Centro Médico Pacífico del Norte es una empresa dedicada a prestación de servicios de salud ocupacional y asistencial, constituida por profesionales altamente calificados que garantizan idoneidad en los resultados de los exámenes médicos ocupacionales, los cuales cumplen con los estándares y criterios de calidad exigidos por las normas legales; complementando sus servicios con la evaluación diagnóstica y tratamiento de enfermedades de sus pacientes.

- ✓ **Razón Social:** CENTRO MEDICO PACIFICO DEL NORTE S.A.C.

- ✓ **Ubicación:** CAL. LOS CIPRESSES NRO. 462 URB. CALIFORNIA - TRUJILLO - VICTOR LARCO HERRERA.

- ✓ **Rubro Económico:** OTRAS ACTIV.RELAC. CON SALUD HUMANA con CIU: 85193

- ✓ **Clientes:**
 - DANPER TRUJILLO S.A.C.
 - AGUALIMA S.A.C.
 - MUNICIPALIDAD DISTRITAL DE VICTOR LARCO HERRERA
 - COMPAÑIA MINERA PODEROSA S. A
 - BBVA BANCO CONTINENTAL
 - ADECCO CONSULTING S.A.
 - EXSA S A
 - COMPAÑIA MINERA RIO CHICAMA S.A.C
 - GALLETERA DEL NORTE S.A.
 - MINERA CHINALCO PERÚ S.A.

✓ **Visión:** Ser una Empresa Líder en Salud Ocupacional y Asistencial reconocida a nivel Local y Regional, con sistemas administrativos integrados, completamente automatizados, precisos, eficientes y versátiles; equipos de alta tecnología e infraestructura acorde con nuestros servicios.

✓ **Misión:** Brindar servicios de salud ocupacional y asistencial al sector privado, institucional y población en general a través de una atención personalizada con calidad y calidez, un alto nivel profesional y humano, para lograr el mejoramiento continuo de la calidad de vida de nuestros clientes.

✓ **Organigrama:**

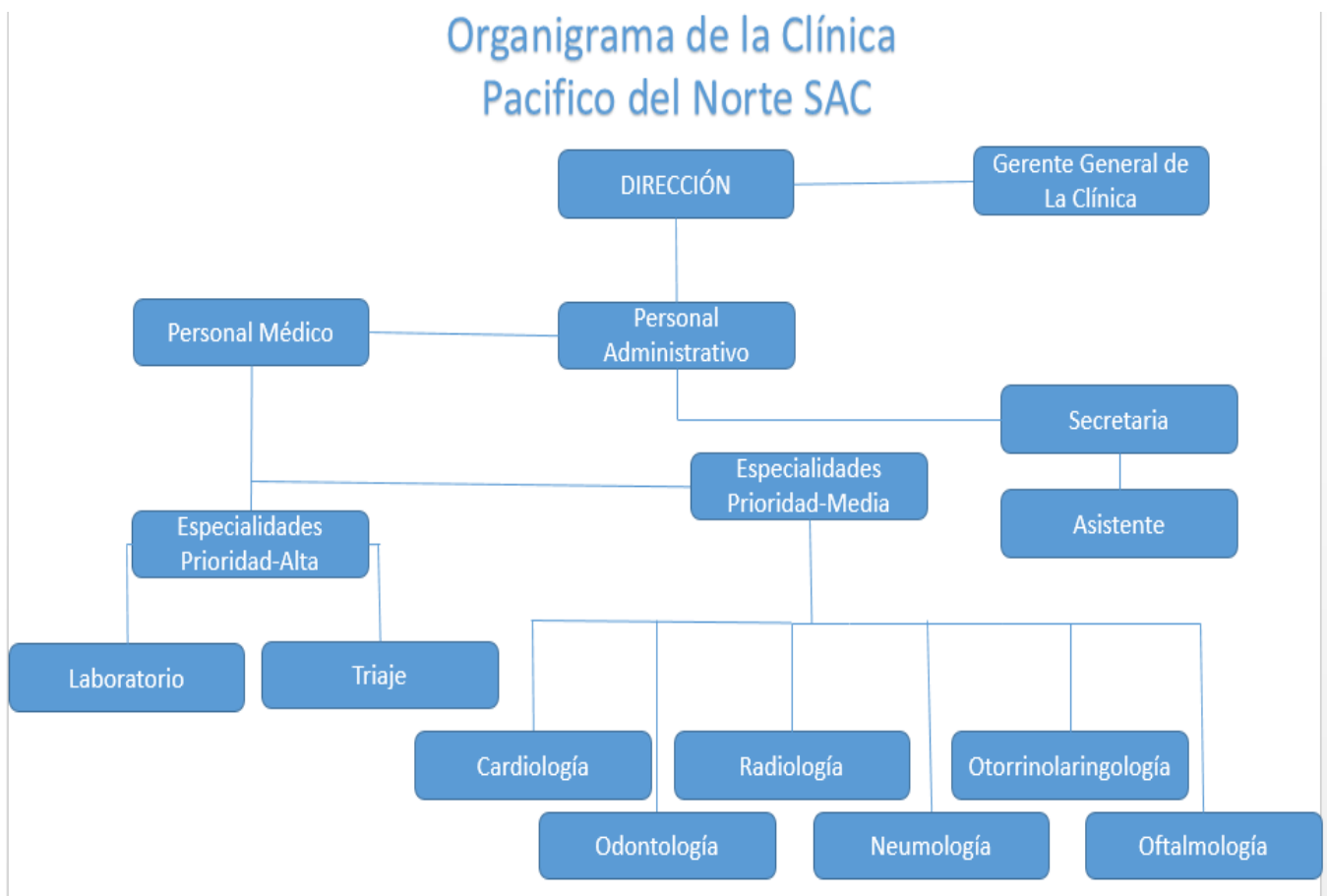


Figura 6: Organigrama de la empresa

✓ Proceso a Desarrollar de Laboratorio:

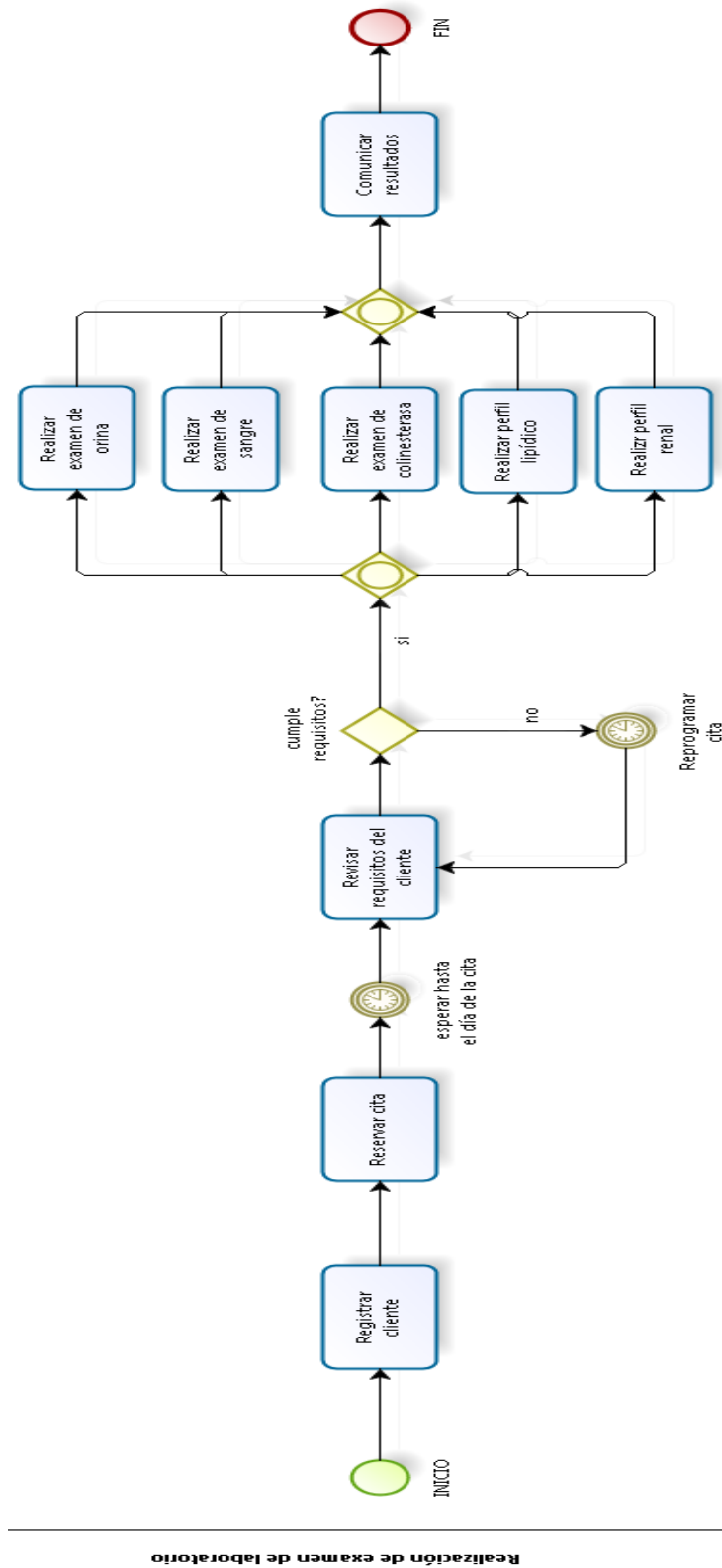


Figura 7: Proceso a desarrollar (Proporcionado por la empresa)

4.2. PLANIFICACIÓN

4.2.1. EVALUACIÓN DE INFRAESTRUCTURA DE LA EMPRESA

a. Infraestructura técnica

El “área de admisión y laboratorio” del centro médico donde se dearrollará la solución de inteligencia de negocios tiene los siguientes recursos tecnológicos:

Características de Recursos: hardware y software

HARDWARE

- ✓ 06 computadoras de escritorio con las siguientes características:
 - Procesador Intel Corei5 Lenovo M720s
 - Memoria Ram 8 GB
 - DDR4 1TB SATA
 - Video y audio integrado
 - Lenovo Monitor LI2215s LED 21.5"
 - Lector y Grabador de CD/DVD - LG
 - Mouse y teclado
- ✓ 03 laptops con las siguientes características:
 - Lenovo
 - Corei5
 - Ram 8GB
 - DD 1 TB
 - Lector de CD/DVD
- ✓ 01 Fotocopiadora Brother Dcp-l2540dw.

SOFTWARE

- ✓ Sistema Operativo Windows 10
- ✓ Microsoft Office 2016
- ✓ Adobe Reader DC
- ✓ Google Chrome

La información fue proporcionada por la empresa y la evaluación para su respectiva evaluación en la solución de Business Intelligence, donde se ve la necesidad de adquirir en la parte de hardware un servidor donde se aloje información del sistema BI. En cuanto al software, se realizarán instalaciones que permitan que el sistema BI implemente correctamente. El software a utilizar esta bajo la arquitectura tecnológica de Pentaho BI.

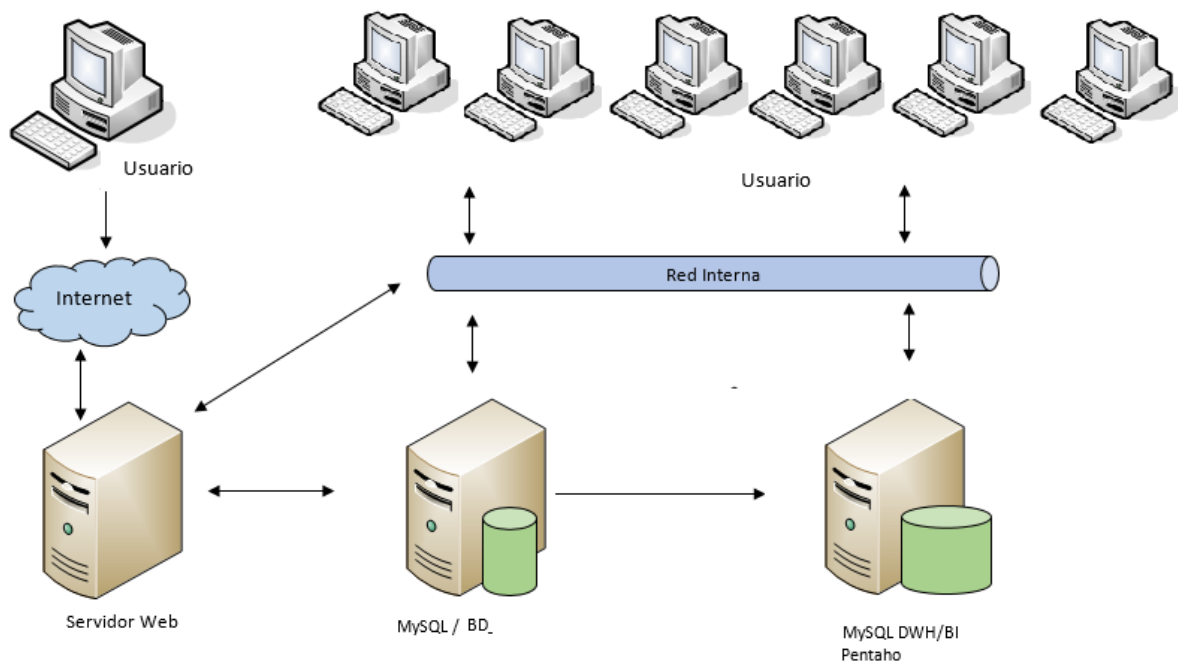


Figura 8: Infraestructura tecnológica de la empresa

b. Infraestructura no técnica

La información obtenida de la administración de la clínica es que no existe un sistema avanzado y tampoco algún tipo de Sistema de soporte de decisiones. En la institución esta sería la primera vez que se implementa un Sistema de soporte de decisiones.

4.2.2. PLANIFICACIÓN DEL PROYECTO

Para realizar una correcta planificación del desarrollo del proyecto se ve por conveniente desarrollar las fases correspondiente a la metodología: (Moss & Atre, 2003)

- ✓ Etapa de Justificación.
- ✓ Etapa de Planificación.
- ✓ Etapa de Análisis del negocio
- ✓ Etapa de Diseño
- ✓ Etapa de Construcción
- ✓ Etapa de Implementación

4.3. ANÁLISIS DEL NEGOCIO

4.3.1. DEFINICIÓN DE LOS REQUISITOS DEL PROYECTO

Según la metodología la implementación de una solución de BI “es un proceso complejo, es por esto que deben usarse las mejores prácticas existentes. Una de ellas es el método propuesto por la metodología, el cual se utilizará para desarrollar el trabajo, adaptándola al caso específico de este proyecto” (Moss & Atre, 2003).

El proyecto se realizara analizando la realidad en la que se encuentra las áreas de Admisión y Laboratorio del Centro Médico Pacifico del Norte para la generación de conocimiento que ayuden a mejorar la toma de decisiones del área.

Es importante definir el escenario de uso empresarial de la organización respecto a la solución.

4.3.1.1. Descripción de los Stakeholders

✓ Personal Involucrado en el Proyecto

Nombre	Representa	Rol
Gerente General	Personal encargado de Planificar, organizar y dirigir la administración de la institución, de acuerdo a normas y políticas establecidas.	<ul style="list-style-type: none"> ▪ “Ejercer la representación legal de la empresa, tomando las decisiones definitivas de acuerdo a las atribuciones conferidas”. ▪ “Controlar el funcionamiento operativo de la institución, responsabilizándose de la situación económico-financiera de la empresa, tomando decisiones definitivas en provecho del cumplimiento de los objetivos de la empresa”. ▪ “Ejercer el control interno de todas las operaciones y actividades que se realizan en la empresa”.
Analista de Sistemas	Encargado del desarrollo de aplicaciones en lo que respecta a su diseño y obtención de los algoritmos, así como de analizar las posibles utilidades y modificaciones	<ul style="list-style-type: none"> ▪ Conocimiento del paradigma tradicional de la ingeniería del software y del tradicional ciclo de vida del software en cascada. ▪ Modelado funcional: Diagrama de flujo de datos, diagrama de estado, etc.

	necesarias de los sistemas operativos para una mayor eficacia de un sistema informático	<ul style="list-style-type: none"> ▪ Modelado de datos y sus técnicas: Diagrama entidad-relación, modelo relacional, etc. ▪ Conocimiento de la tecnología: arquitectura de software, bases de datos, etc.
Administrador de la base de datos	Operador que maneja los datos y lo mantiene actualizado.	➤ Maneja el acceso de los usuarios al Sistema de Información y la seguridad de los datos, además de la actualización de los datos.

✓ **Usuarios la Solución de BI**

Nombre	Representa	Stakeholder
Gerente	Encargado de Planificar, organizar y dirigir la administración de la empresa.	Gerente
Jefe de Admisión y laboratorio	Responsable de la eficaz administración del área de admisión y laboratorio y que permitan el eficiente funcionamiento de las operaciones de la institución.	Jefe de Admisión y Laboratorio
Administrador del BI	Persona que se encarga de la seguridad y mantenimiento del BI.	Administrador de la Solución de BI

Con la información anterior se realizó la obtención de requerimiento basándonos principalmente en el objetivo principal y área crítica de la institución, siendo el área de Admisión y Laboratorio la prioridad para la solución de BI, donde se realizó lo siguiente:

- ✓ Se realizaron reuniones y entrevistas a los tomadores de decisiones según el cargo y responsabilidades dentro del área. Se utilizó un cuestionario con preguntas dirigidas hacia el interés de la investigación. (Anexo A).
- ✓ Se revisó los procesos que se desarrollan en la institución para tener una mejor visión de lo que se desea y como esto puede dar un beneficio a la empresa.
- ✓ Se obtuvo acceso a la fuente de datos y registros que maneja la organización actualmente, accediendo a la base de datos transaccional que se encuentra en MySQL, el cual será analizado y procesado, para que luego sea de utilidad para el desarrollo de la solución.

La implementación de la solución de inteligencia de negocios permite obtener información correcta y oportuna para que la institución pueda tomar decisiones acertadas, además de la celeridad en la obtención de información sobre los que está sucediendo en los procesos de admisión y de laboratorio de la clínica.

4.3.1.2. Requerimientos del Usuario Final

Para determinar los requerimientos del área de admisión y laboratorio de la empresa se realizaron entrevistas a los usuarios que van a utilizar la solución de BI, en donde se especifica los requerimientos del personal implicado en el área los cuales fueron recopilados de las reuniones llevadas a cabo:

Área de Admisión:

- R1 ¿Cuál es la cantidad total de exámenes médicos que se generan en un determinado mes?
- R2 ¿Cuál es la cantidad de exámenes médicos por medico en un mes determinado?
- R3 ¿Cuál es la cantidad de exámenes médicos generados por tipo y empresa en un mes determinado?
- R4 ¿Cuál es la cantidad de exámenes médicos que se realizan por cada paciente en un determinado mes?
- R5 ¿Cuál es la cantidad total de consultas médicas realizadas por paciente en un determinado mes?
- R6 ¿Cuál es la cantidad media de pacientes particulares por día?
- R7 ¿Cuál es la cantidad y tipo de exámenes médicos por empresa en un mes determinado?

Área de Laboratorio:

- R8 ¿Cuáles son los exámenes médicos que menos se realizan en un mes determinado?
- R9 ¿Cuál es la cantidad de pacientes que tuvieron un examen médico en un mes determinado?
- R10 ¿Cuál es la cantidad de pacientes particulares que pasaron examen médico de laboratorio en un mes determinado?

NRO	CONSULTAS	UNIDAD
01	¿Cuál es la cantidad total de exámenes médicos que se generan en un determinado mes?	Cantidad de reportes
02	¿Cuál es la cantidad de exámenes médicos por medico en un mes determinado?	Cantidad de reportes
03	¿Cuál es la cantidad de exámenes médicos generados por tipo y empresa en un mes determinado?	Cantidad de reportes
04	¿Cuál es la cantidad de exámenes médicos que se realizan por cada paciente en un determinado mes?	Cantidad de reportes
05	¿Cuál es la cantidad total de consultas médicas realizadas por paciente en un determinado mes?	Cantidad de consultas
06	¿Cuál es la cantidad media de pacientes particulares por día?	Cantidad de pacientes
07	¿Cuál es la utilidad por mes de un tipo de examen médico en un mes determinado?	Porcentaje Utilidad
08	¿Cuáles son los exámenes médicos que menos se realizan en un mes determinado?	Cantidad de exámenes médicos
09	¿Cuál es la cantidad de pacientes que tuvieron un examen médico en un mes determinado?	Cantidad de pacientes particulares
10	¿Cuál es la cantidad de pacientes particulares que pasaron examen de laboratorio en un mes determinado?	Cantidad de pacientes particulares

Tabla 3: Lista de requerimientos

4.3.1.3. Requerimientos no funcionales

- ✓ Para asegurar la integridad de la información, cada usuario que tenga acceso al mismo deberá tener el privilegio correspondiente y solo ese, de tal manera que no pueda manipular información no autorizada.

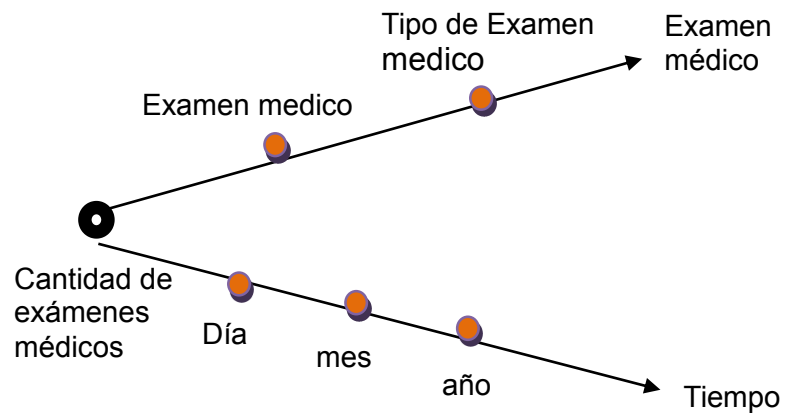
- ✓ Usabilidad, el proceso en el que el usuario genera la información que requiere será lo suficientemente sencillo para personas no afines a la informática y sus aplicaciones.
- ✓ El sistema tendrá que ser lo suficientemente escalable en la medida que la información vaya creciendo con el tiempo y pueda seguir cumpliendo su funcionalidad.
- ✓ El sistema no ha de almacenar o procesar información sensible del usuario o que en cuyo caso no haya proporcionado libremente.
- ✓ El software debe ser construido bajo las herramientas de desarrollador de open source, lo cual incluye MySQL, Herramientas de Bi de Pentaho.
- ✓ El BI debe ser construido bajo la metodología Larisa Moss.

4.3.1.4. Análisis de los requerimientos

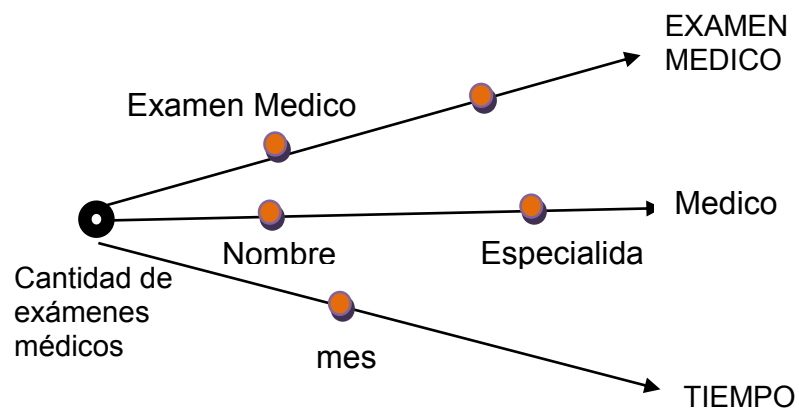
En esta parte se hará un análisis de cada requerimiento del usuario final y definiendo algunas tablas de la Base de Datos Operacional de la empresa, que actuarían como dimensión dentro del BI. También se definirá las posibles medidas dentro del BI. Con la finalidad que el sistema brinde apoyo a la toma de decisiones.

ÁREA DE ADMISIÓN

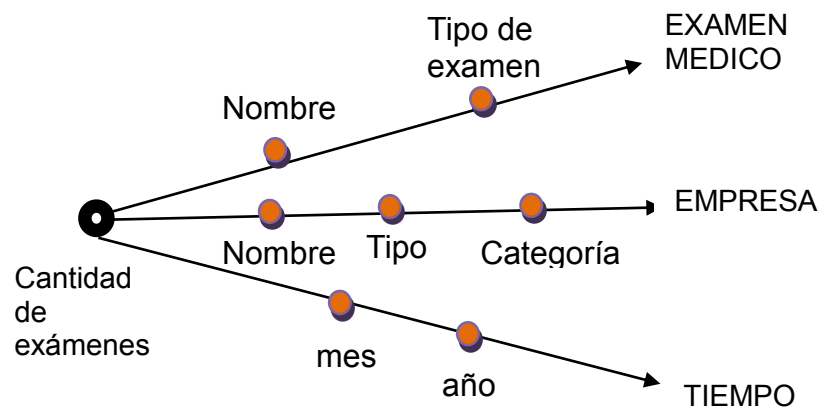
R1 ¿Cuál es la cantidad total de exámenes médicos que se generan en un determinado mes?



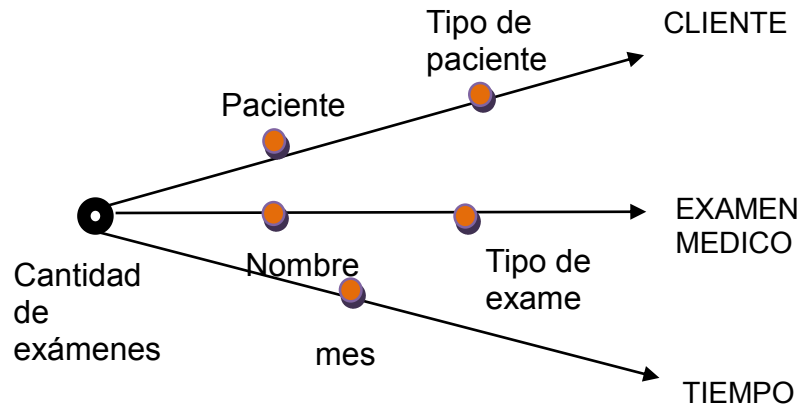
R2 ¿Cuál es la cantidad de exámenes médicos por medico en un mes determinado?



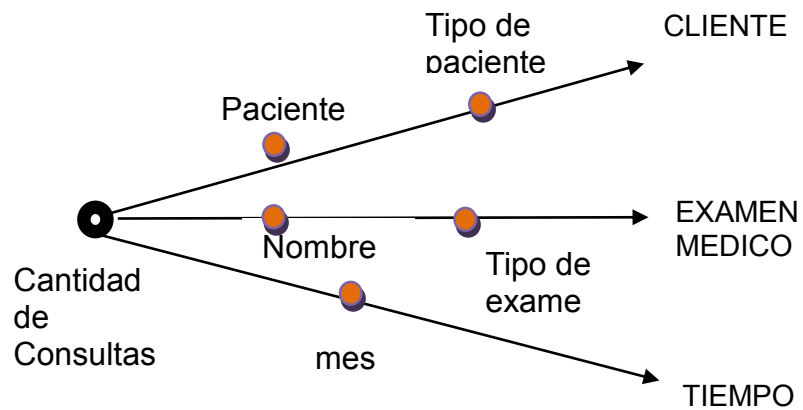
R3 ¿Cuál es la cantidad de exámenes médicos generados por tipo y empresa en un mes determinado?



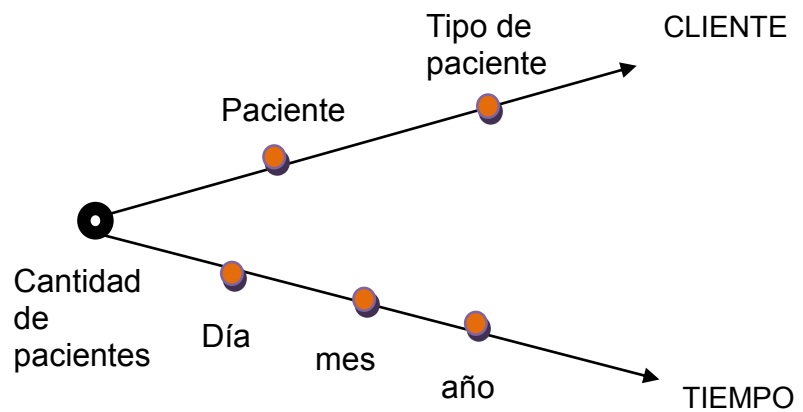
R4 ¿Cuál es la cantidad de exámenes médicos que se realizan por cada paciente en un determinado mes?



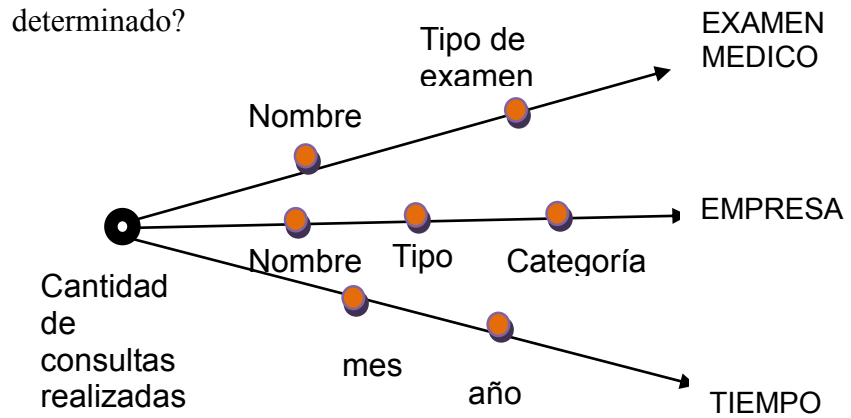
R5 ¿Cuál es la cantidad total de consultas médicas realizadas por paciente en un determinado mes?



R6 ¿Cuál es la cantidad media de pacientes particulares por día?

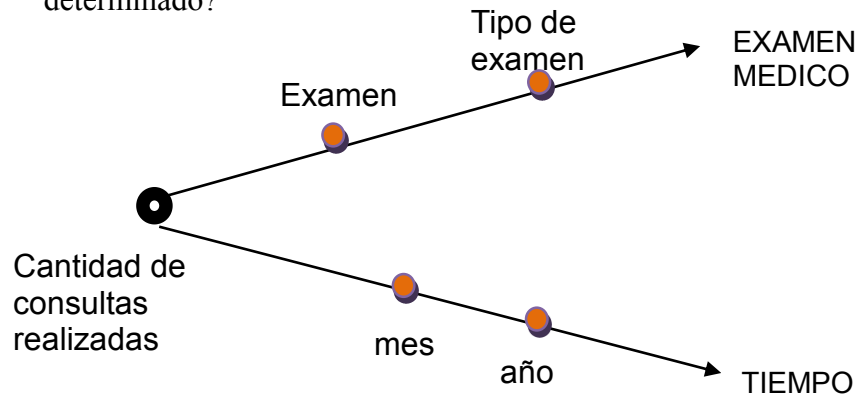


R7 ¿Cuál es la cantidad y tipo de exámenes médicos por empresa en un mes determinado?

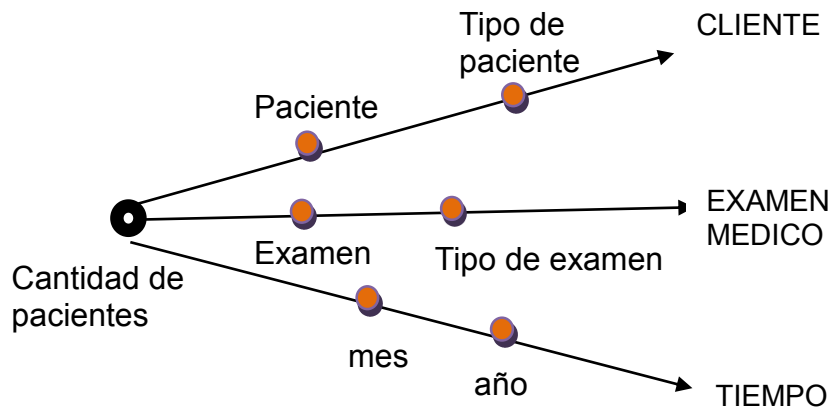


ÁREA DE LABORATORIO

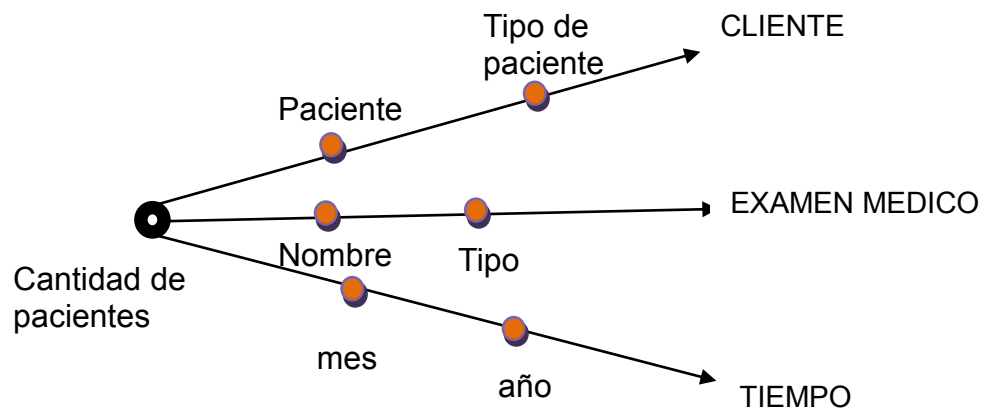
R8 ¿Cuáles son los exámenes médicos que menos se realizan en un mes determinado?



R9 ¿Cuál es la cantidad de pacientes que tuvieron un examen médico en un mes determinado?



R10 ¿Cuál es la cantidad de pacientes particulares que pasaron examen médico de laboratorio en un mes determinado?



4.3.1.5. Identificar Medidas y Perspectivas de Análisis

MEDIDAS

- ✓ Cantidad de exámenes médicos
- ✓ Cantidad de Consultas medicas
- ✓ Cantidad de pacientes

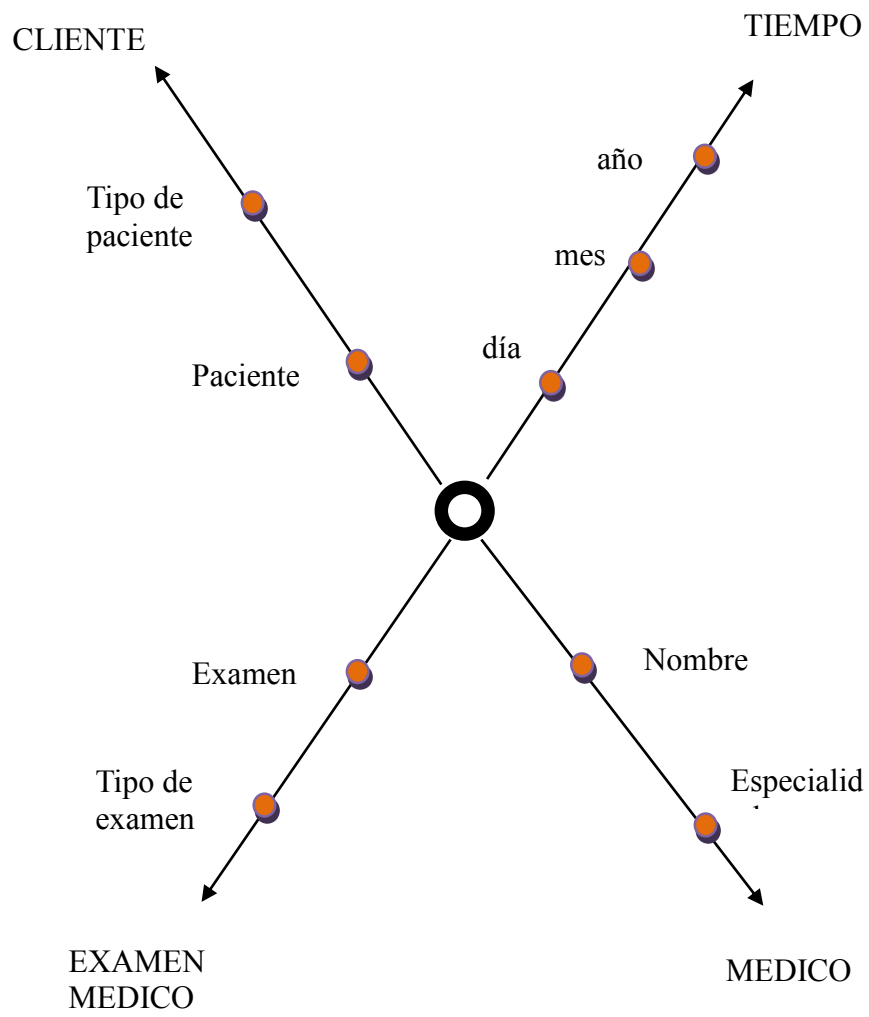
PERSPECTIVAS

- ✓ TIEMPO
- ✓ EMPLEADO
- ✓ CLIENTE
- ✓ EXAMEN MEDICO
- ✓ LABORATORIO

4.3.1.6. Modelo conceptual del BI

Considerando las perspectivas e indicadores analizadas y mostradas en el paso anterior por los requerimientos funcionales del BI que permite demostrar con mayor facilidad el alcance del proyecto a continuación se presentara el modelo que se trabajara en el proyecto

Modelo Startnet Admisión



Modelo Startnet Laboratorio

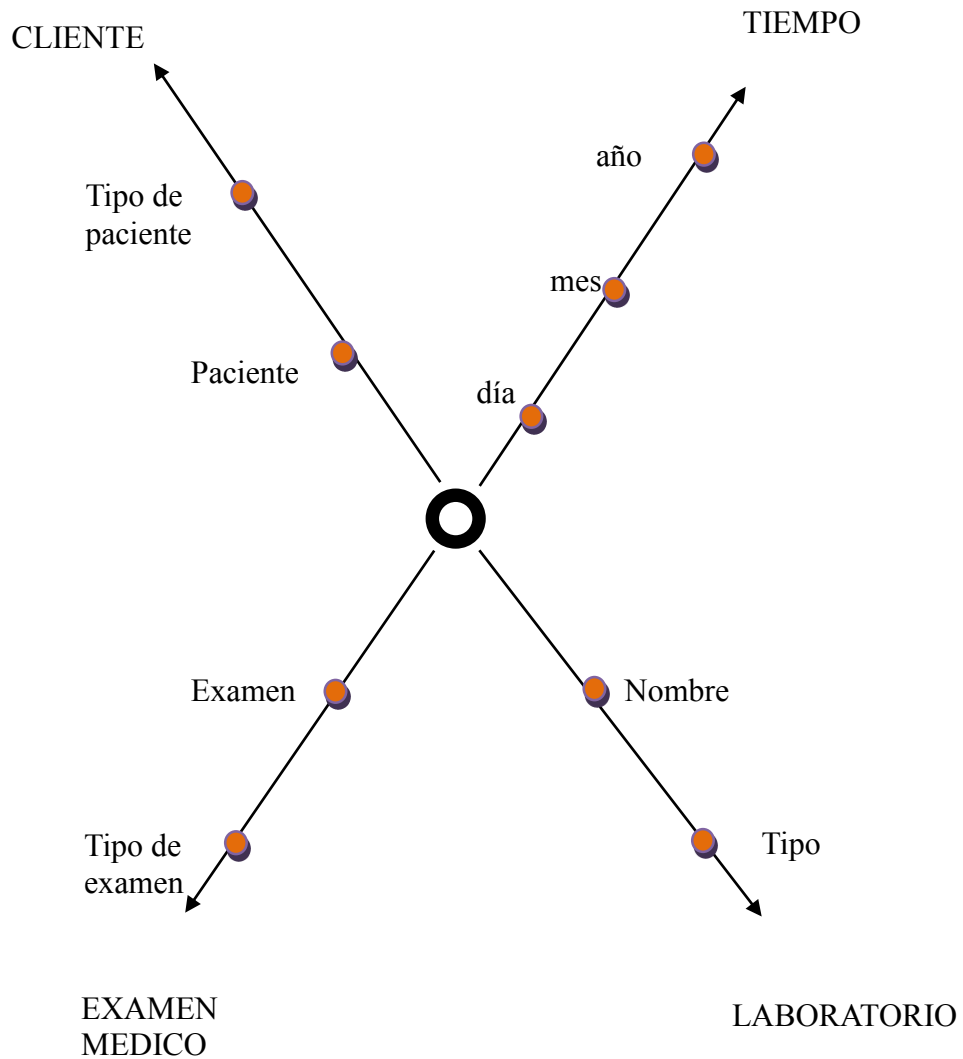


Figura 9: Modelo conceptual

4.3.2. ANÁLISIS DE DATOS

Para esta etapa, se pudo obtener la fuente de datos desde donde se van a extraer la información relevante para la toma de decisiones en la empresa,

Base de Datos Origen:

- ✓ En este diagrama se muestran las tablas que se relacionan con el área de Admisión, así como con el área Laboratorio y que son de interés para el BI. La

mayoría de estas tablas servirán para crear las consultas para el poblamiento de nuestro BI.

BASE DE DATOS

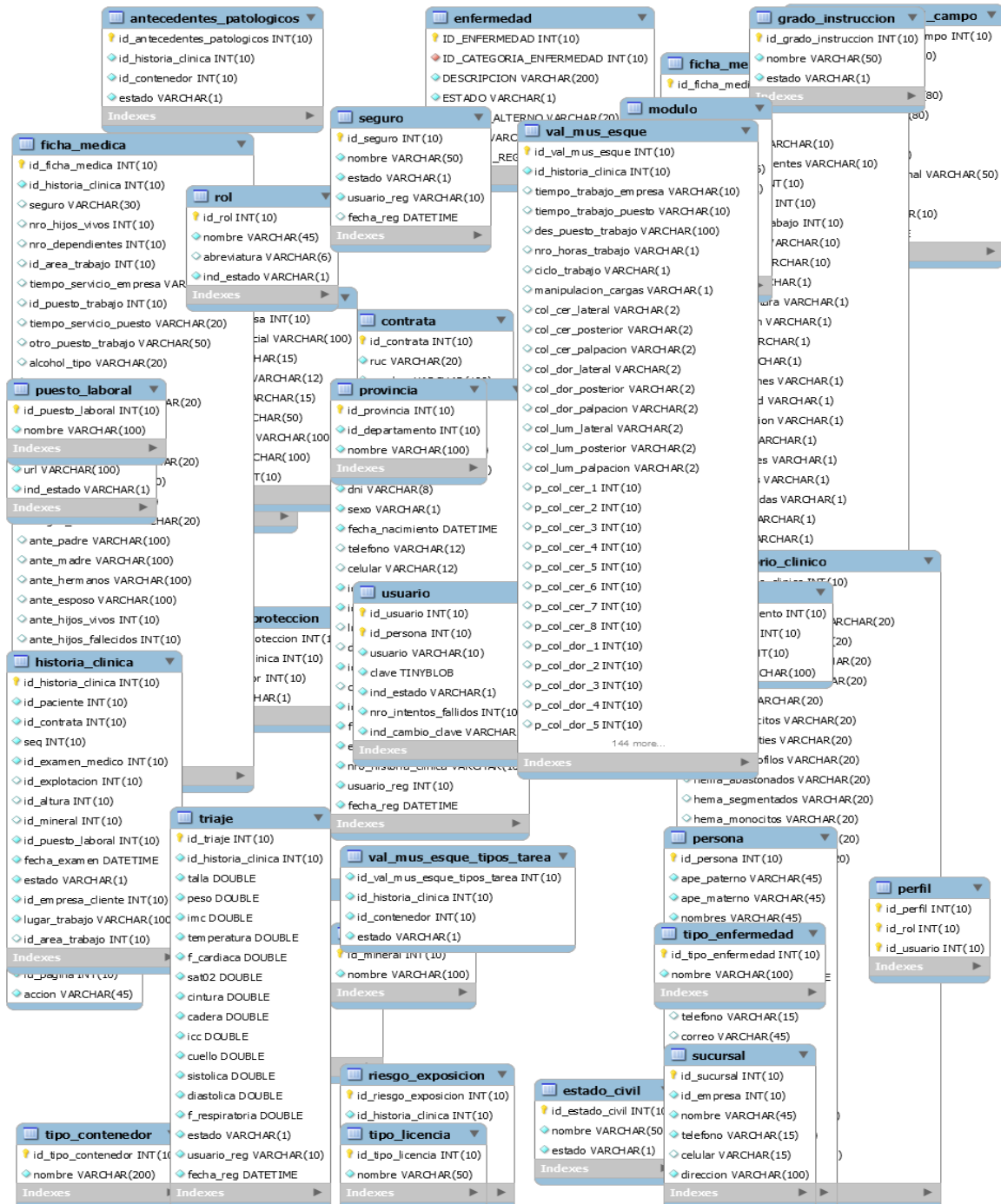


Figura 10: Diagrama BD

4.3.3. PROTOTIPO DE LA APLICACIÓN

A continuación se mostrará la implementación de la tabla de hecho y dimensiones para la solución de BI:

Identificando los Componentes del Modelo:

- a) Elementos de la Solución de BI
- Tablas de Hechos: Admisión / Laboratorio
 - Listado de Dimensiones:
 - ✓ TIEMPO
 - ✓ EMPLEADO
 - ✓ CLIENTE
 - ✓ EXAMEN MEDICO
 - ✓ LABORATORIO

- b) Marcado de las Intersecciones:

Procedemos entonces a organizar ordenadamente las filas y columnas en una tabla tablas, dándole una forma de matriz que nos permita establecer las intersecciones donde se presente una dimensión, de tal forma se relacione a un BI

DIMENSIONES	Áreas	
	Área de Admisión	Área de Laboratorio
TIEMPO	X	X
MÉDICO		X
EXAMEN MÉDICO		X
PACIENTE	X	X

c) Objetivo de la Tabla de Hechos

TABLA DE HECHOS	OBJETIVO
Admisión	“Tener control de pacientes atendidos”.
Laboratorio	“tener un control sobre el tipo de enfermedades encontradas”.

d) Dimensiones:

DIMENSIONES DE LA TABLA DE HECHOS ADMISION

TABLA DE HECHOS	OBJETIVOS	DIMENSIONES
ADMISION	“Tener control de pacientes atendidos”.	Examen Medico Empleado Tiempo Cliente

DIMENSIONES DE LA TABLA DE HECHOS LABORATORIO

TABLA DE HECHOS	OBJETIVOS	DIMENSIONES
LABORATORIO	“Conocer Índice de enfermedades de exámenes realizados por tipo, por empresa en un determinado tiempo”.	Tiempo Cliente Laboratorio

e) Hechos o Medidas

TABLA DE HECHOS	HECHOS
ADMISION	<ul style="list-style-type: none"> • Cantidad de exámenes médicos • Cantidad de Consultas medicas • Cantidad de pacientes
LABORATORIO	<ul style="list-style-type: none"> • Cantidad de exámenes • Cantidad de pacientes

DIAGRAMAS DE LA TABLA DE HECHOS

DIAGRAMA DE LA TABLA DE HECHOS ADMISION

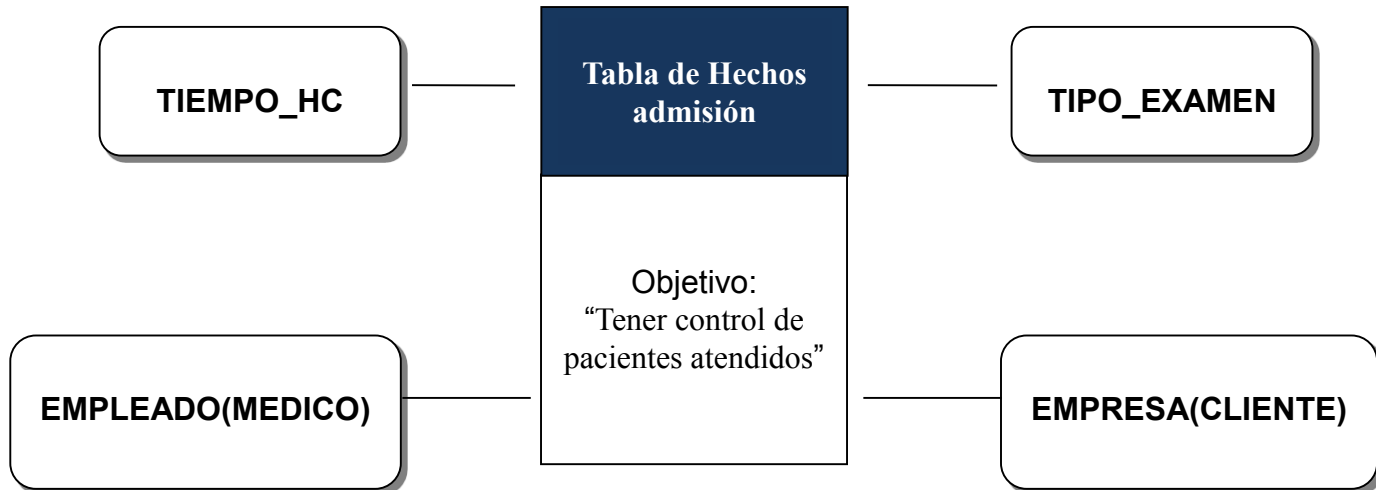


Figura 11: Diagrama de Tabla de Hechos Admisión

DIAGRAMA DE LA TABLA DE HECHOS LABORATORIO

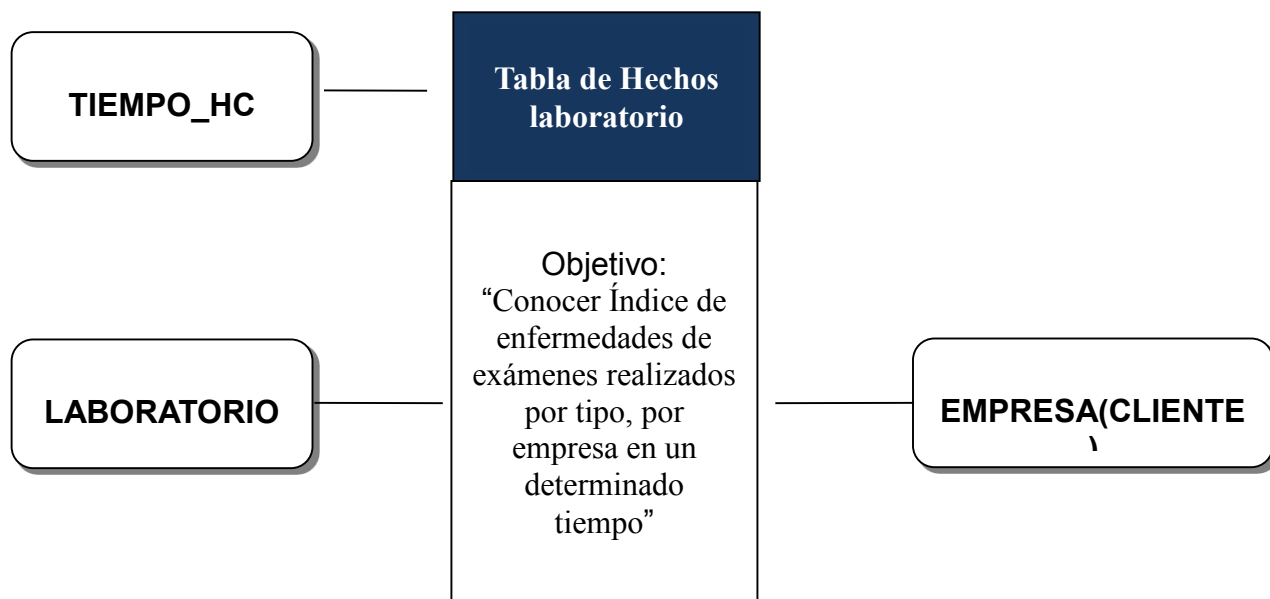


Figura 12: Diagrama de Tabla de Hecho Laboratorio.

DETALLE DE LAS TABLAS DE HECHOS

DETALLE DE LAS CLAVES DE LAS DIMENSIONES

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Dimensión Tiempo Admisión	DIM_ID	Llave primaria única para la Dimensión Tiempo Admisión
Dimensión Cliente(Empresa)	ID_DIM_CLIENTE	Llave primaria única para la Dimensión Cliente(empresa)
Dimensión LABORATORIO	ID_ECHO_LAB	Llave primaria única para la Dimensión laboratorio

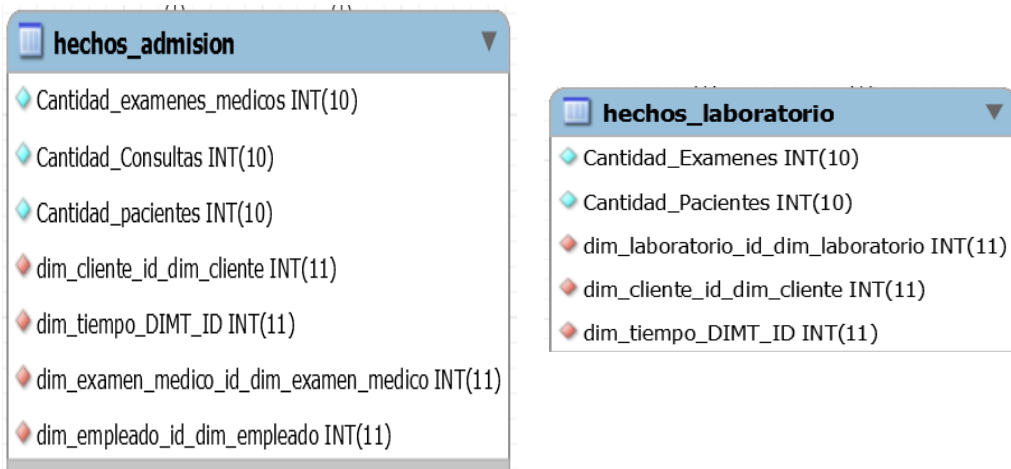
Tabla 4: Detalle de las claves de las dimensiones

DETALLE DE LAS MEDIDAS DE LAS TABLAS DE HECHO

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Hecho Admisión	Cantidad de exámenes médicos	Cantidad de exámenes médicos
	Cantidad de Consultas	Cantidad de Consultas
	Cantidad de pacientes	Cantidad de pacientes
Hecho Laboratorio	Cantidad de exámenes	Cantidad de exámenes
	Cantidad de pacientes	Cantidad de pacientes

Tabla 5: Detalle de las Medidas de las Tablas de Hechos

Por consiguiente, de las tres tablas completadas anteriormente se obtienen las siguientes tablas de hechos:



ESQUEMA ESTRELLA

Orientamos la solución de Inteligencia de negocios al Esquema Estrella para una mejor visualización del Modelo para el Diseño Físico, donde se conoce el número de Tablas de Hechos y las dimensiones asociada a esta solución.

- ✓ Componente: TABLA DE HECHOS ADMISION
- ✓ Componente: TABLA DE HECHOS LABORATORIO
- ✓ Componente: TABLA DE DIMENSION TIEMPO
- ✓ Componente: TABLA DE DIMENSIÓN CLIENTE
- ✓ Componente: TABLA DE DIMENSIÓN EMPLEADO
- ✓ Componente: TABLA DE DIMENSIÓN LABORATORIO
- ✓ Componente: TABLA DE DIMENSIÓN EXAM_MEDICO

Figura 13: Esquema Lógico de la Tabla de Hechos Admisión

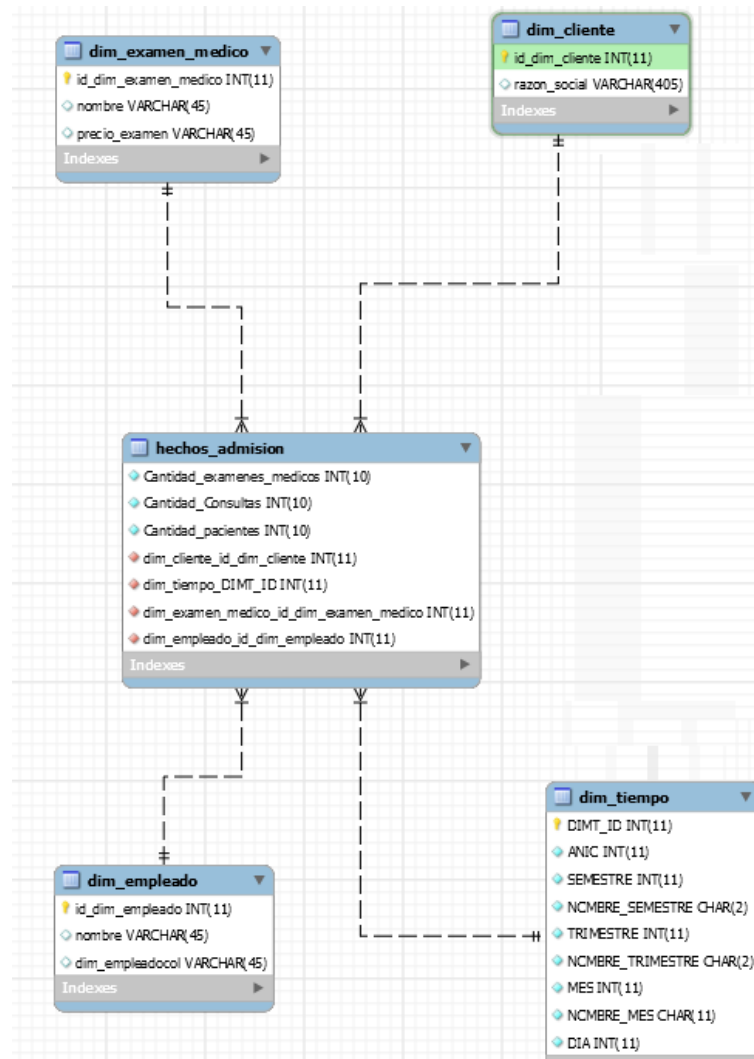


Figura 14: Esquema Lógico de la Tabla de Hechos Laboratorio

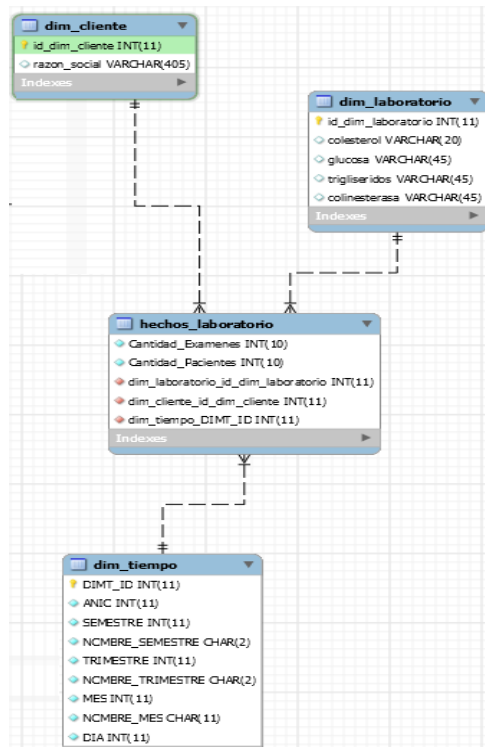
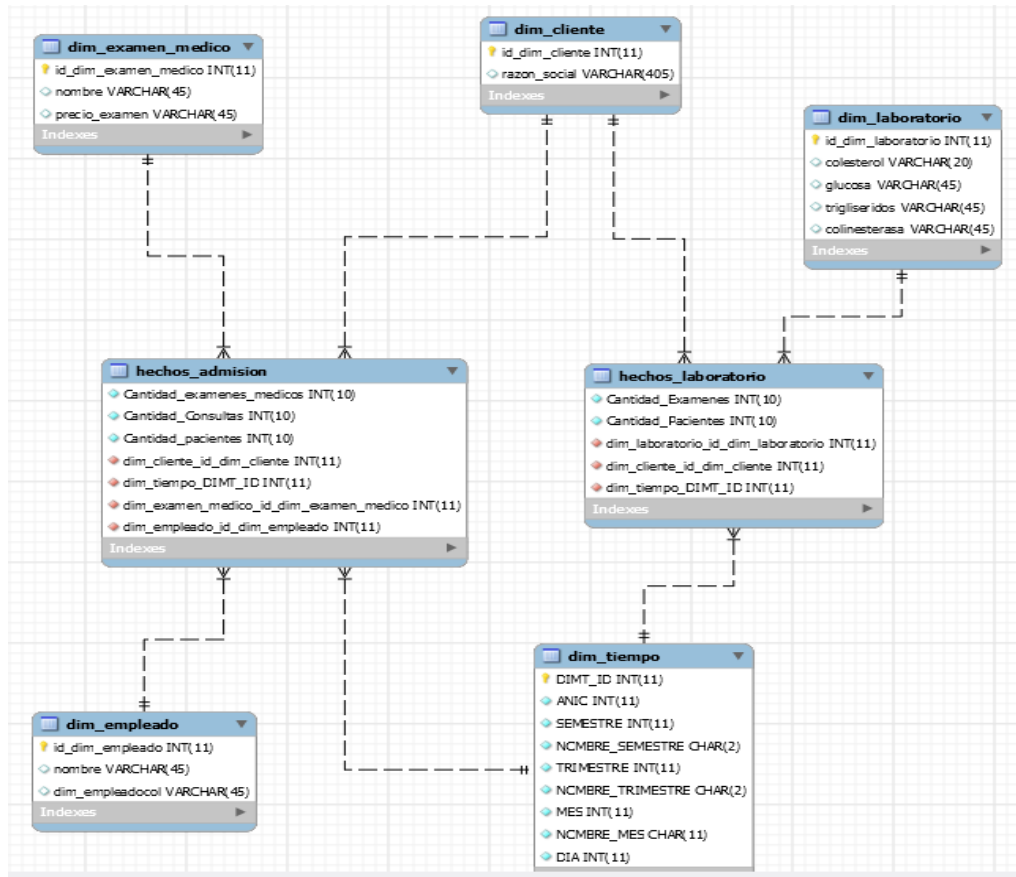


Figura 15: Diagrama Multidimensional

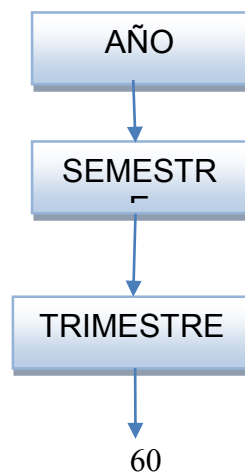


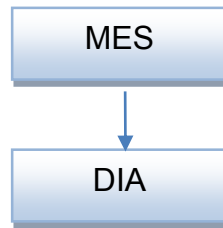
4.4. DISEÑO

4.4.1. DISEÑO DE LA BASE DE DATOS

De acuerdo a nuestras Dimensiones elegidas anteriormente construimos los detalles y sus jerarquías de cada una de ellas. Las mismas que sirven para las 2 tablas de hechos.

Dimensión Tiempo





Jerarquía de los Atributos de la Dimensión Tiempo

Detalle de Dimensión Tiempo Admisión/Laboratorio

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Año	Año	No Actualizar	2015,2016,...
Semestre	Semestre	No Actualizar	Julio, Agosto,...
Nombre semestre	Nombre_semestre	No Actualizar	2015,2016,...
Trimestre	Trimestre	No Actualizar	Julio, Agosto,...
Nombre trimestre	Nombre_trimestre	No Actualizar	2015,2016,...
Mes	Mes	No Actualizar	1,2,...12
Nombre de mes	Nombre_mes	No Actualizar	Julio, Agosto,...
Dia	Dia	No Actualizar	

Dimensión Empleado

Detalle de Dimensión Empleado

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre Empleado	Representa el nombre del Empleado.	No Actualizar	Jouverth Custodio Torres
DNI del empleado	Representado por el DNI del empleado	No Actualizar	73183329

Dimensión Cliente

Detalle de Dimensión Cliente

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Razón social(empresa)	Representa la razón social del cliente	No Actualizar	ADECCO CONSULTING S.A

Dimensión Examen medico

Detalle de Dimensión Examen medico

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre Examen medico	Representa el nombre del examen médico.	No Actualizar	Pre-ocupacional
Precio_examen	Representa el precio de cada examen medico	Actualizar	120.00

Dimensión Examen laboratorio

Detalle de Dimensión Laboratorio

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
colesterol	Representa el valor de colesterol	No Actualizar	85.5
glucosa	Representa el valor de glucosa	Actualizar	80.5
colinesterasa	Representa el valor de colinesterasa	No Actualizar	86.0
triglicéridos	Representa el valor de triglicéridos	Actualizar	98.30

Modificar nombres a nombres estándar, si fuera necesario. Para nuestro BI se realizaron los siguientes cambios:

**NOMBRES ESTÁNDARES PARA LAS TABLAS HECHOS Y
DIMENSIONES**

DISEÑO LÓGICO	DISEÑO FÍSICO
Tabla de Hecho Admision	Hechos_Admission
Tabla de Hecho Laboratorio	Hechos_Laboratorio
Dimensión Cliente	Dim_Cliente
Dimensión Laboratorio	Dim_Laboratorio
Dimensión Empleado	Dim_Empleado
Dimensión Examen Medico	Dim_Examen_Medico
Dimensión Tiempo	Dim_Tiempo

Tabla 6: Nombre Estándares para la Tabla de Hechos y sus Dimensiones

**NOMBRES ESTÁNDARES PARA LOS ATRIBUTOS DE LAS
DIMENSIONES**

NOMBRE DE LA TABLA	DISEÑO LÓGICO	DISEÑO FÍSICO
Dimensión Examen Medico	id_dim_examen_medico	id_dim_examen_medico
Dimensión Examen	precio	precio

Medico		
Dimensión Examen Medico	precio_examen	precio_examen
Dimensión Laboratorio	id_dim_laboratorio	id_dim_laboratorio
Dimensión Laboratorio	colesterol	colesterol
Dimensión Laboratorio	glucosa	glucosa
Dimensión Laboratorio	trigliceridos	trigliceridos
Dimensión Laboratorio	colinesterasa	colinesterasa
Dimensión Cliente	id_dim_cliente	id_dim_cliente
Dimensión Cliente	razon_social	razon_social
Dimensión Tiempo	DIMT_ID	DIMT_ID
Dimensión Tiempo	ANIO	ANIO
Dimensión Tiempo	SEMESTRE	SEMESTRE
Dimensión Tiempo	NOMBRE_SEMESTRE	NOMBRE_SEMESTRE
Dimensión Tiempo	TRIMESTRE	TRIMESTRE
Dimensión Tiempo	NOMBRE_TRIMESTRE	NOMBRE_TRIMESTRE
Dimensión Tiempo	MES	MES
Dimensión Tiempo	NOMBRE_MES	NOMBRE_MES
Dimensión Tiempo	DIA	DIA
Dimensión Empleado	id_dim_empleado	id_dim_empleado
Dimensión Empleado	nombre	Nombre
Dimensión Empleado	dim_empleadocol	dim_empleadocol

NOMBRES ESTÁNDARES PARA LOS ATRIBUTOS DE LAS TABLAS DE

HECHOS

NOMBRE DE LA TABLA	DISEÑO LÓGICO	DISEÑO FÍSICO
Tabla de Hecho Admisión	precio_examen	precio_examen
Tabla de Hecho Admisión	monto_total	monto_total
Tabla de Hecho Admisión	cantidad_examen_tipo	cantidad_examen_tipo
Tabla de Hecho Laboratorio	indc_colesterol	indc_colesterol
Tabla de Hecho Laboratorio	indc_colesterol	indc_colesterol
Tabla de Hecho Laboratorio	indc_trigliceridos	indc_trigliceridos
Tabla de Hecho Laboratorio	indc_colinesterasa	indc_colinesterasa

4.4.2. DISEÑO EXTRAER/TRANSFORMAR/CARGAR (ETL)

Para ingresar los datos a la solución de BI se debe poblar cada Tabla de la Base de Datos. Como se conoce la estructura de cada dimensión y tabla de hecho, se debe seguir los siguientes pasos para el poblamiento:

- ✓ Definición de los Pasos de Transformación
- ✓ Definición de los Flujos de trabajo (workflow).
- ✓ Restricciones de Precedencia
- ✓ Diagrama Workflow con las Restricciones de Precedencia.

a) Definición de los Pasos de Transformación

Para un correcto Poblamiento de los datos debemos definir los Pasos de Transformación, con la unidad de trabajo como parte de un proceso de

transformación. Para poblar el BI se tiene que realizar los siguientes pasos:

Paso 1: Limpiando Dimensiones, consiste en eliminar los datos de todas las Dimensiones y Tablas de Hechos, paso que nos permite asegurar de que no pueda existir algún dato que se pudiera duplicar.

Paso 2: Poblamiento de la Dimensión Tiempo, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Tiempo.

Paso 3: Poblamiento de la Dimensión Cliente, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Cliente

Paso 4: Poblamiento de la Dimensión empleado, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Empleado

Paso 5: Poblamiento de la Dimensión Laboratorio, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Laboratorio.

Paso 6: Poblamiento de la Dimensión Examen Medico, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Examen Medico.

Paso 7: Poblamiento de las Tabla de Hechos.

b) Definición de los Pasos de los Workflows

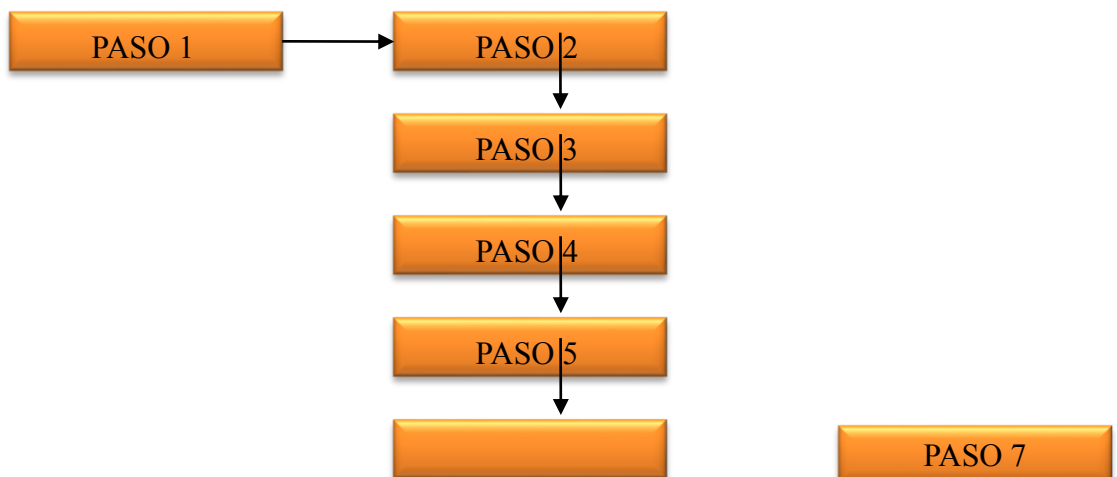
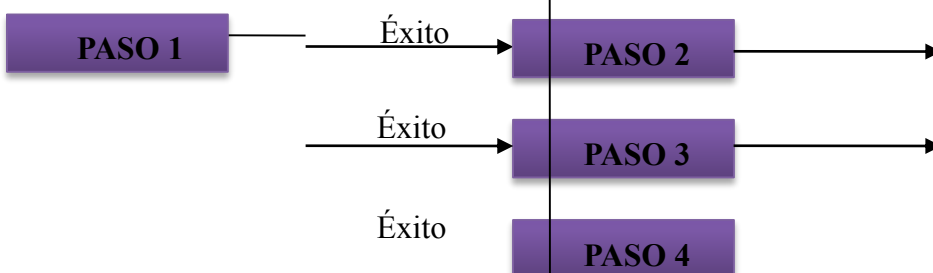


Figura 16: Pasos de los Workflows

c) Restricciones de Precedencia:

- ✓ La Limpieza de las Dimensiones (Paso 1) debe realizarse al inicio del proceso.
- ✓ El Poblamiento de la Dimensión Tiempo (Paso 2) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- ✓ El Poblamiento de la Dimensión Cliente (Paso 3) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- ✓ El Poblamiento de la Dimensión Empleado (Paso 4) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- ✓ El Poblamiento de la Dimensión Laboratorio (Paso 5) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- ✓ El Poblamiento de la Dimensión Examen Medico (Paso 6) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- ✓ El Poblamiento de las Tablas de Hechos (Paso 7) debe realizarse sólo cuando se tenga la seguridad de que el paso 2,3,4,5,6,7,8,9 se hallan ejecutado con éxito.

Una vez conocido las relaciones de precedencias diagramaremos el workflows que se necesitará realizar al construir el paquete de poblamiento del BI, para Centro médico.



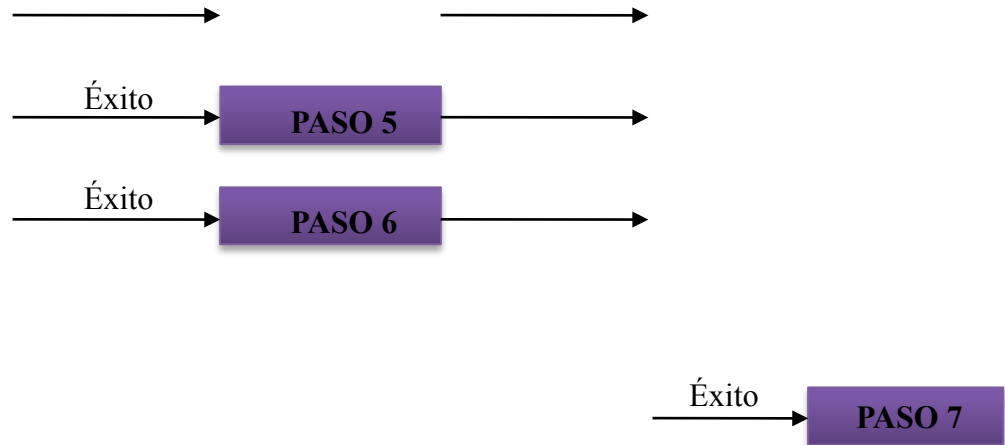


Figura 17: Diagrama Workflow con Restricciones de Precedencia

4.5. CONSTRUCCIÓN

4.5.1. CONSTRUCCIÓN DE LAS TABLAS DE LA SOLUCION DE INTELIGENCIA DE NEGOCIOS EN LA BASE DE DATOS MySQL

DIM CLIENTE

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_dim_cliente	INT(11)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
razon_social	VARCHAR(405)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

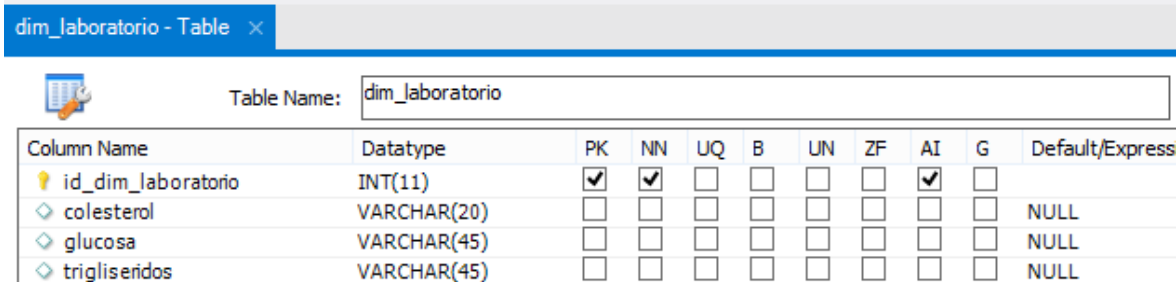
Tabla de la Dimensión CLIENTE

DIM TIEMPO

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default
DIMT_ID	INT(11)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
ANIO	INT(11)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
SEMESTRE	INT(11)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
NOMBRE_SEMESTRE	CHAR(2)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Tabla de la Dimensión Tiempo

DIM LABORATORIO



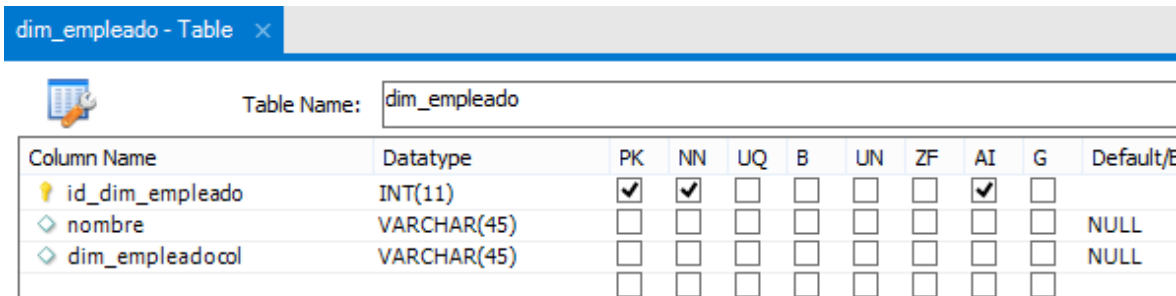
dim_laboratorio - Table ×

Table Name: dim_laboratorio

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Express
id_dim_laboratorio	INT(11)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
colesteroles	VARCHAR(20)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
glucosa	VARCHAR(45)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
trigliceridos	VARCHAR(45)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

Tabla de la Dimensión LABORATORIO

DIM EMPLEADO



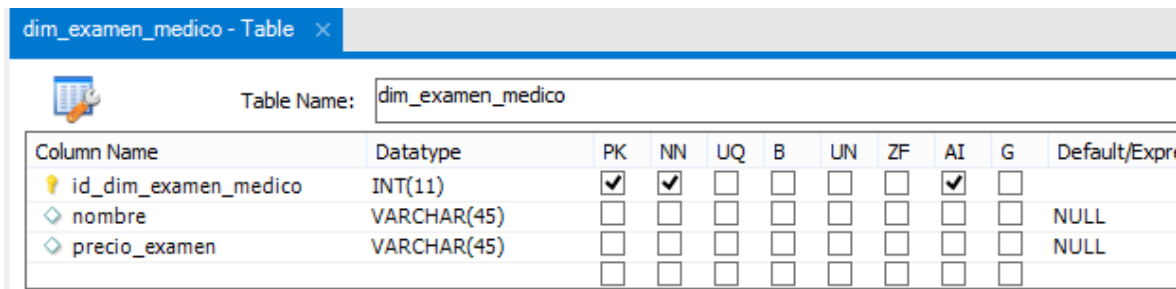
dim_empleado - Table ×

Table Name: dim_empleado

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Express
id_dim_empleado	INT(11)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
nombre	VARCHAR(45)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
dim_empleadocol	VARCHAR(45)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

Tabla de la Dimensión EMPLEADO

DIM EXAMEN MEDICO



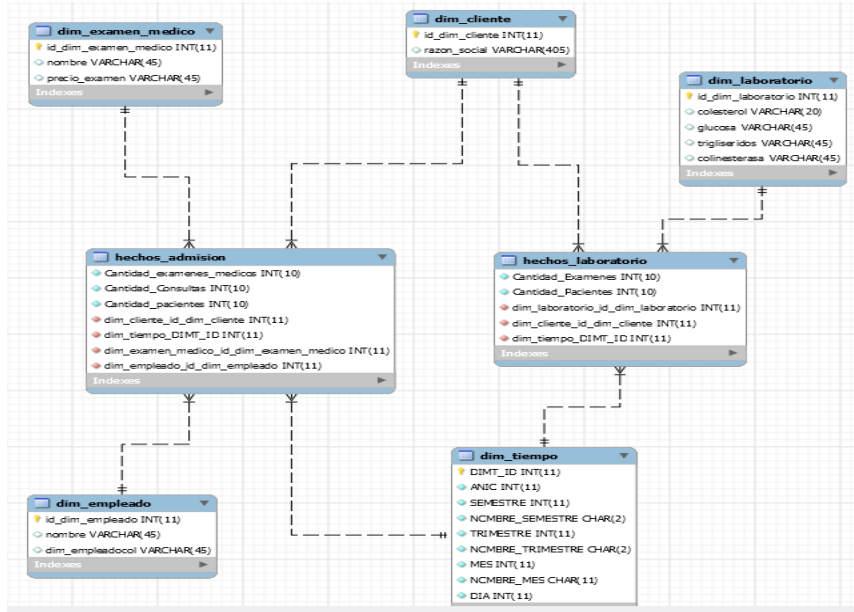
dim_examen_medico - Table ×

Table Name: dim_examen_medico

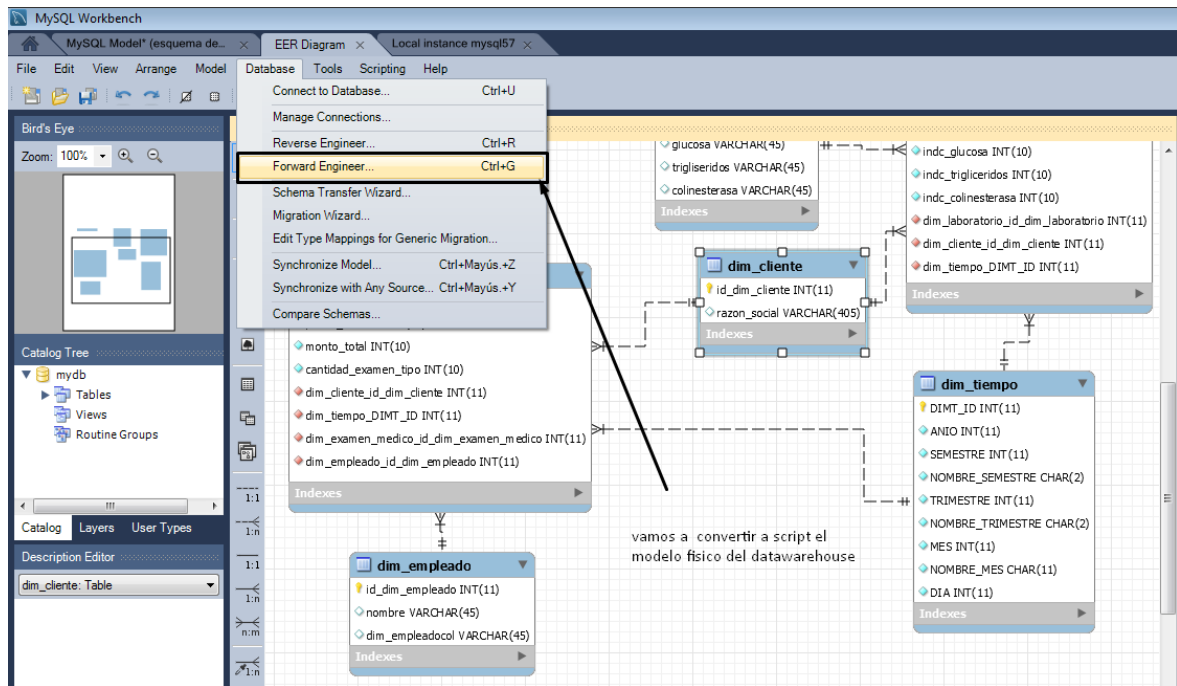
Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Express
id_dim_examen_medico	INT(11)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
nombre	VARCHAR(45)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL
precio_examen	VARCHAR(45)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL

Tabla de la Dimensión EXAMEN MEDICO

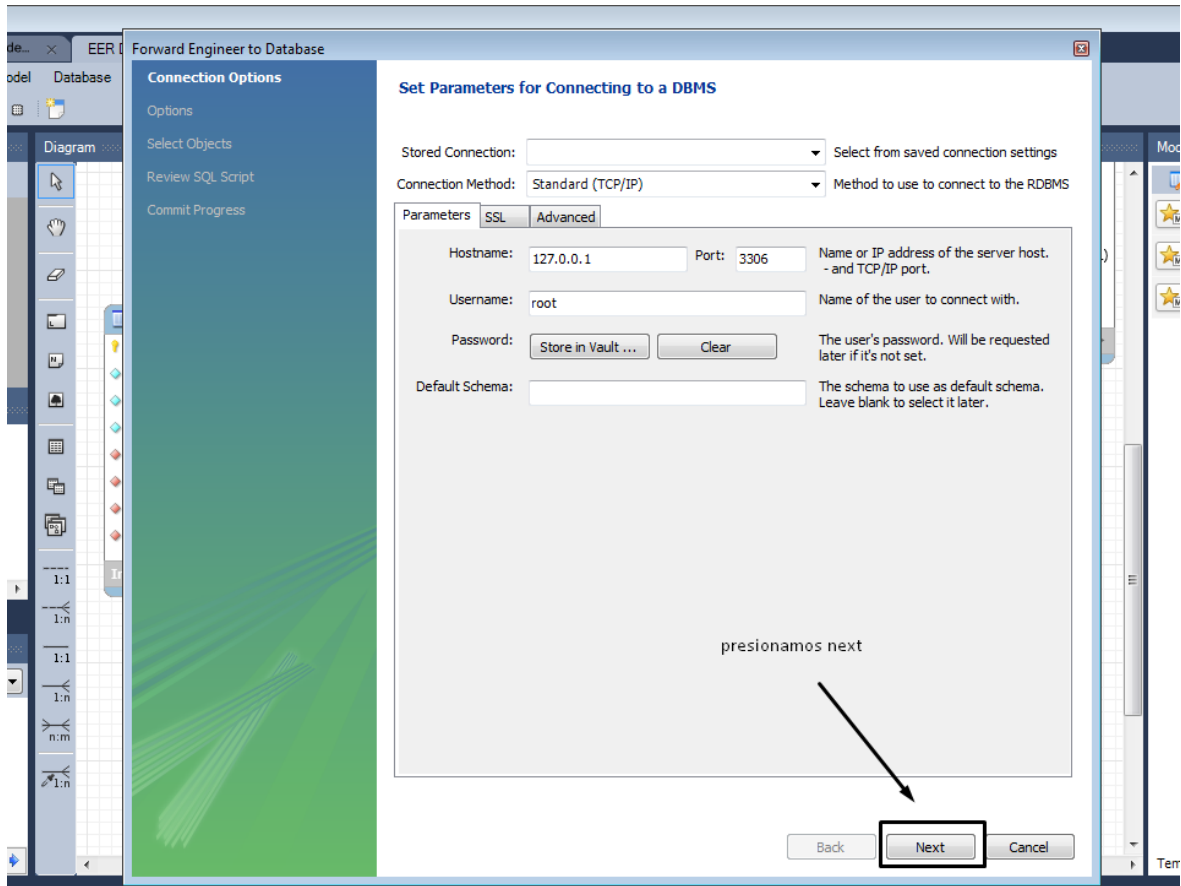
Primero diseñaremos en la herramienta workbench para luego hacer una forward engineer



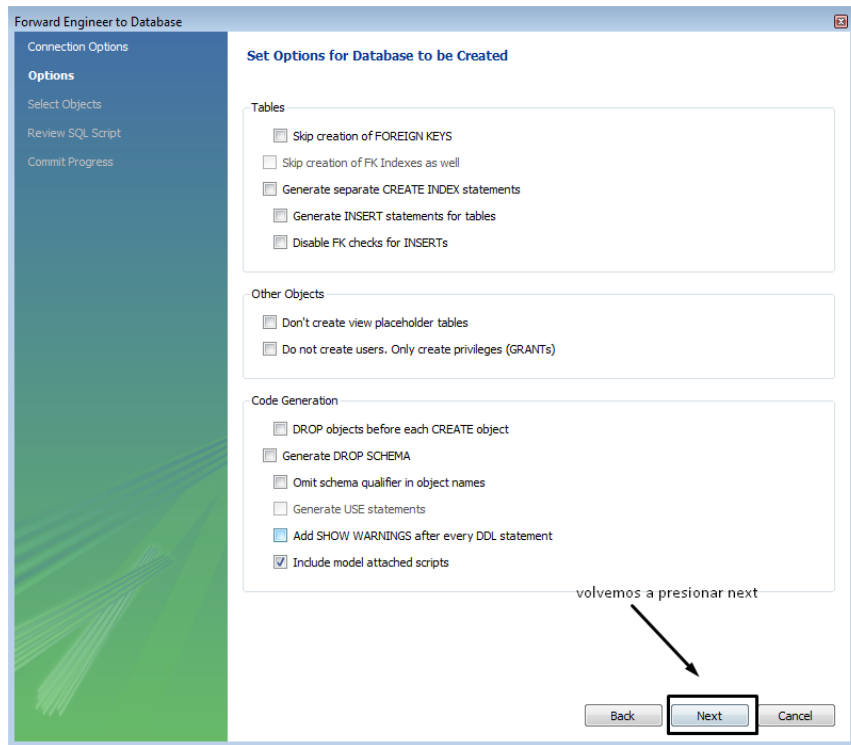
Primer paso



Segundo paso



Tercer paso



Cuarto paso

Forward Engineer to Database

Connection Options

Options

Select Objects

Review SQL Script

Commit Progress

Select Objects to Forward Engineer

To exclude objects of a specific type from the SQL Export, disable the corresponding checkbox. Press and add objects or patterns to the ignore list to exclude them from the export.

Export MySQL Table Objects
7 Total Objects, 7 Selected

Export MySQL View Objects

Export User Objects
0 Total Objects, 0 Selected

Connect to MySQL Server

Please enter password for the following service:

Service: Mysql@127.0.0.1:3306

User: root

Password:

Save password in vault

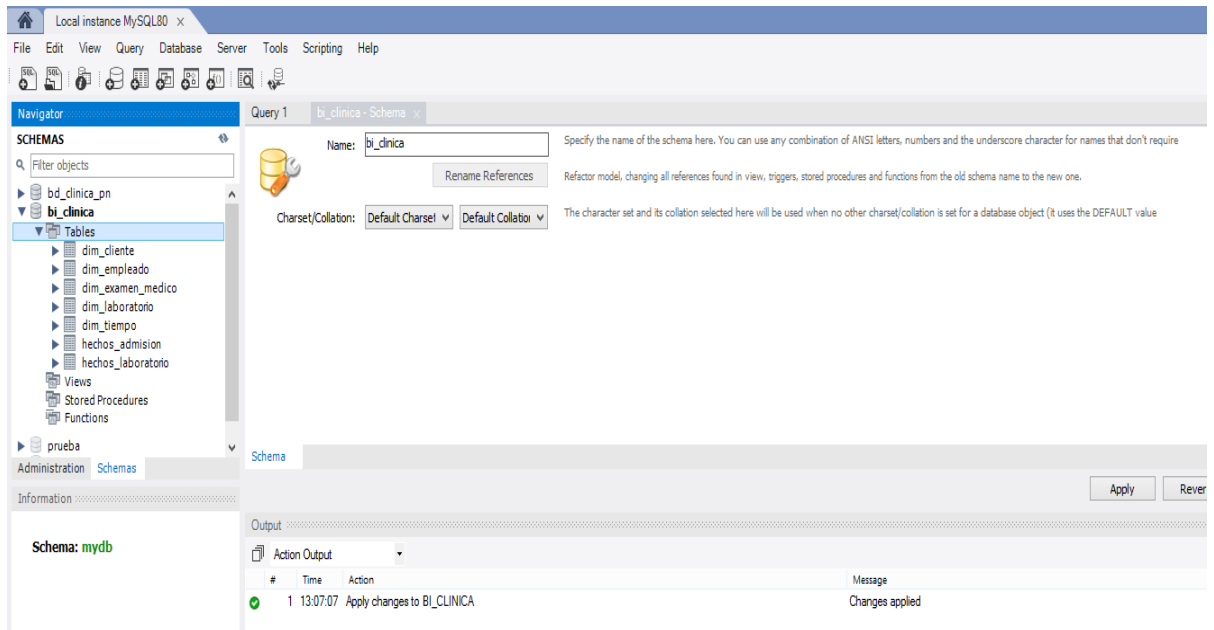
OK Cancel

establecemos la conexión

Finalmente tenemos el script para ejecutar

```
19 -----
20 • CREATE TABLE IF NOT EXISTS `mydb`.`dim_cliente` (
21   `id_dim_cliente` INT(11) NOT NULL AUTO_INCREMENT,
22   `razon_social` VARCHAR(405) NULL DEFAULT NULL,
23   PRIMARY KEY (`id_dim_cliente`))
24 ENGINE = InnoDB
25 AUTO_INCREMENT = 706
26 DEFAULT CHARACTER SET = utf8;
27
28 -----
29
30 -- Table `mydb`.`dim_empleado`
31 -----
32 • CREATE TABLE IF NOT EXISTS `mydb`.`dim_empleado` (
33   `id_dim_empleado` INT(11) NOT NULL AUTO_INCREMENT,
34   `nombre` VARCHAR(45) NULL DEFAULT NULL,
35   `dim_empleadocol` VARCHAR(45) NULL DEFAULT NULL,
36   PRIMARY KEY (`id_dim_empleado`))
37 ENGINE = InnoDB
```

ejecutamos el script y ya tenemos mybd(dw)



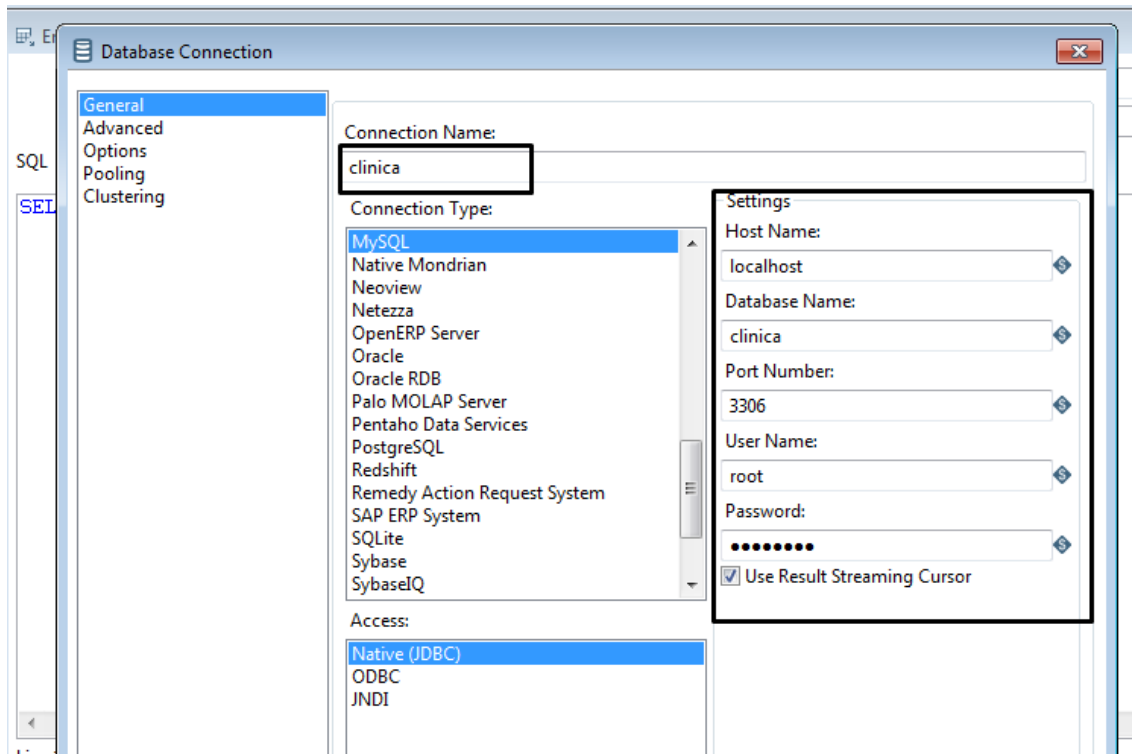
4.5.2. PROCESO EXTRAER – TRANSFORMAR - CARGAR DATOS A LA SOLUCION DE BI (ETL)

a) Crear un Proyecto de Integración en la Herramienta Pentaho Data Integration

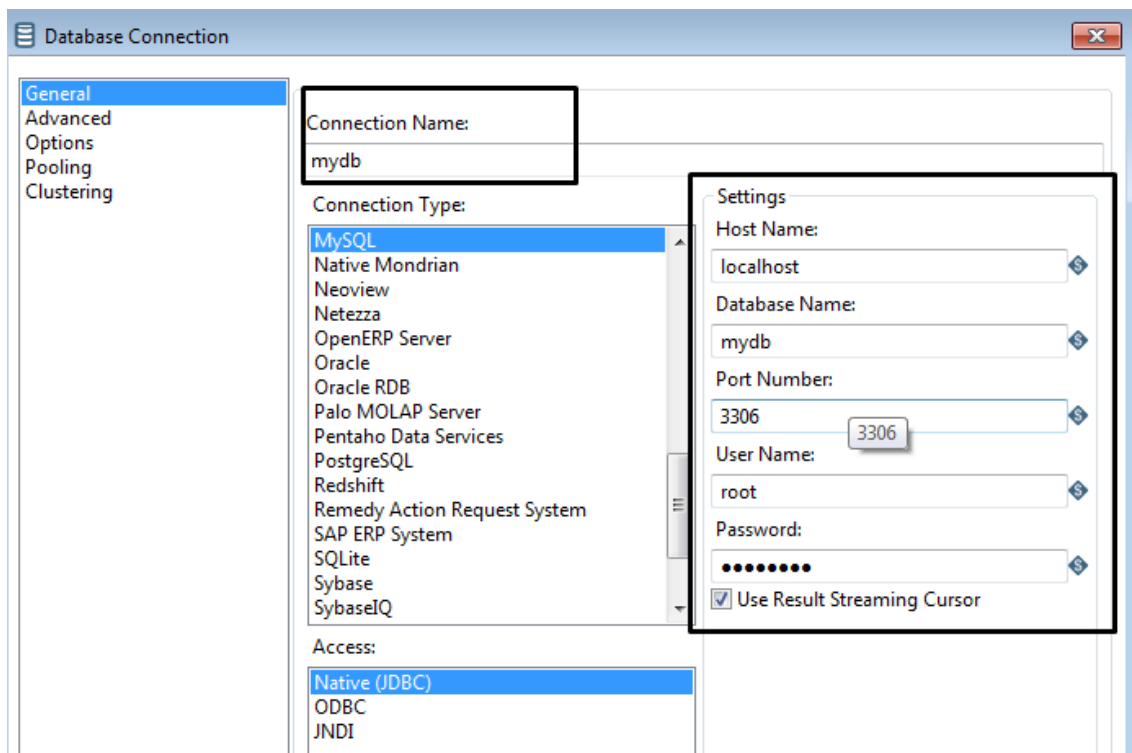
La herramienta Pentaho Data Integration va permitir crear un proyecto para realizar todo el proceso ETL de una manera gráfica generando un flujo de control para el paquete y también los flujos de datos para desarrollar dicho trabajo.

Primero establecemos las conexiones tanto a la base de datos relacional como al BI.

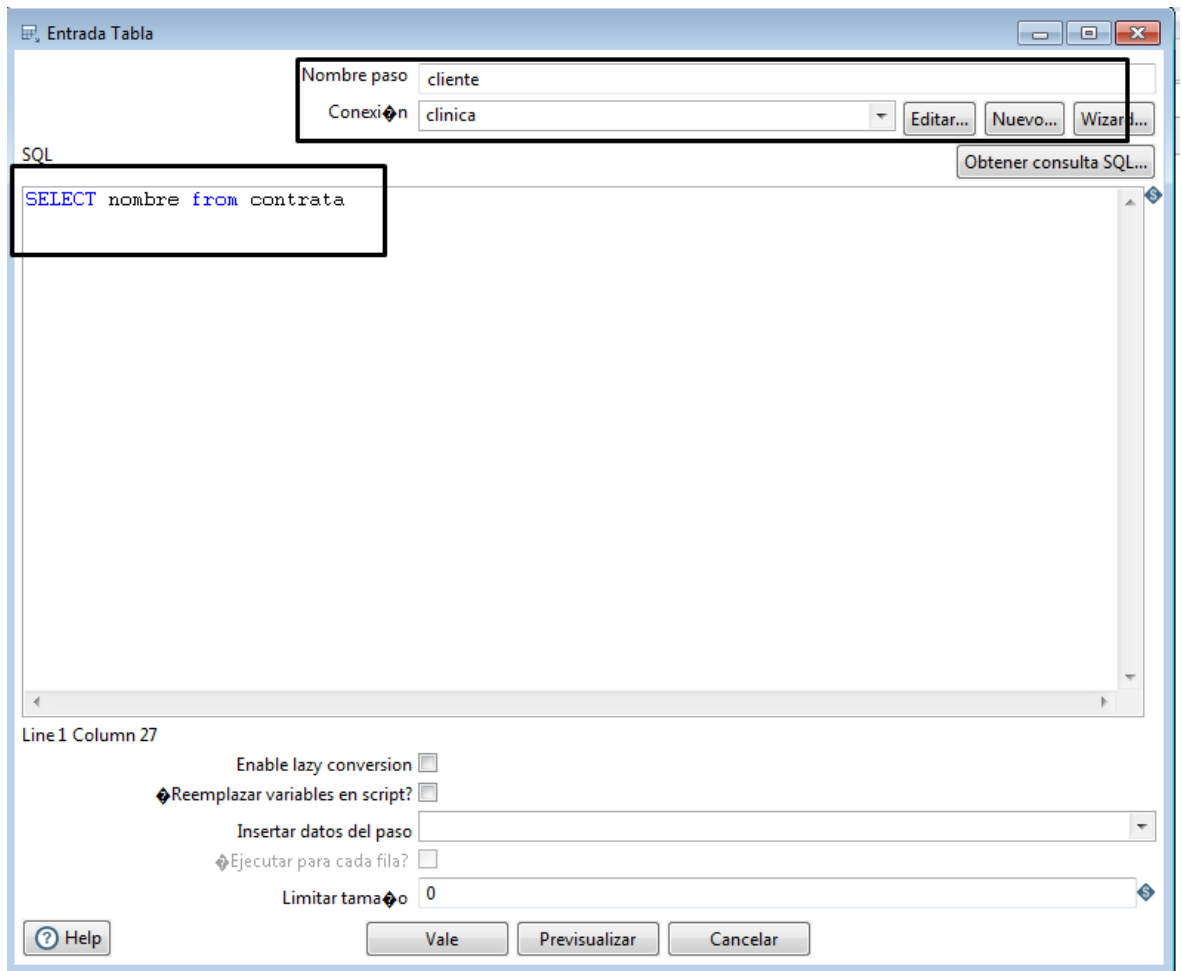
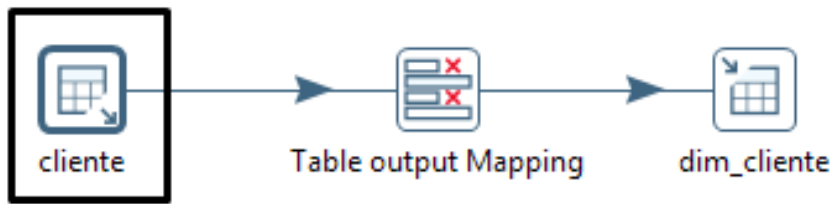
Conexión a la base de datos relacional



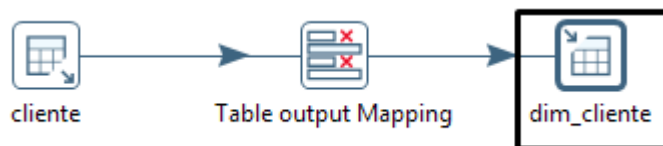
Conexión al BI



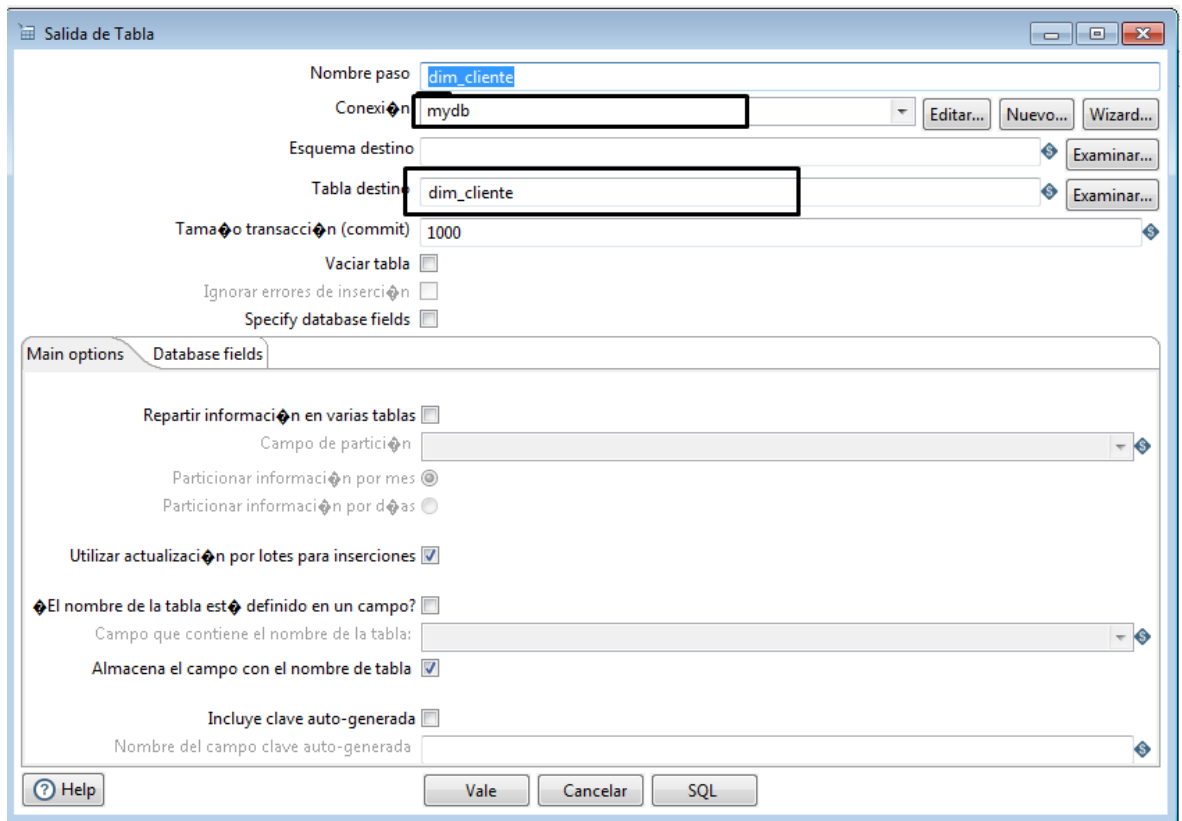
Luego para la dimensión cliente establecemos los datos a tomar en cuenta para llenar dicha dimensión siendo tomada de la base de datos transaccional.



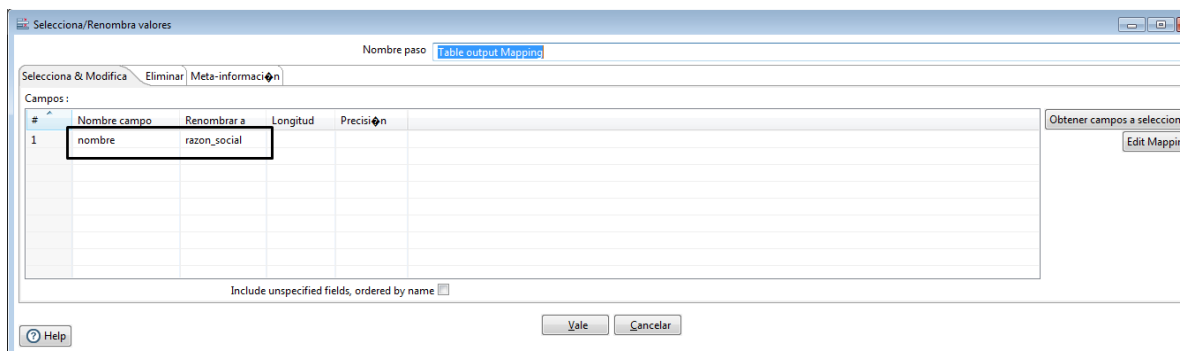
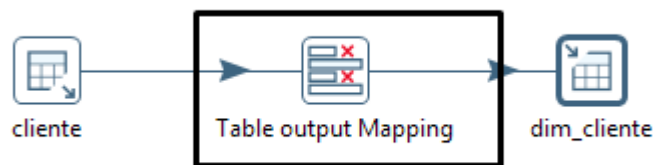
Posteriormente indicamos a que tabla de la data warehouse está dirigiéndose los datos tomados de la transaccional.



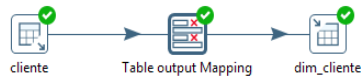
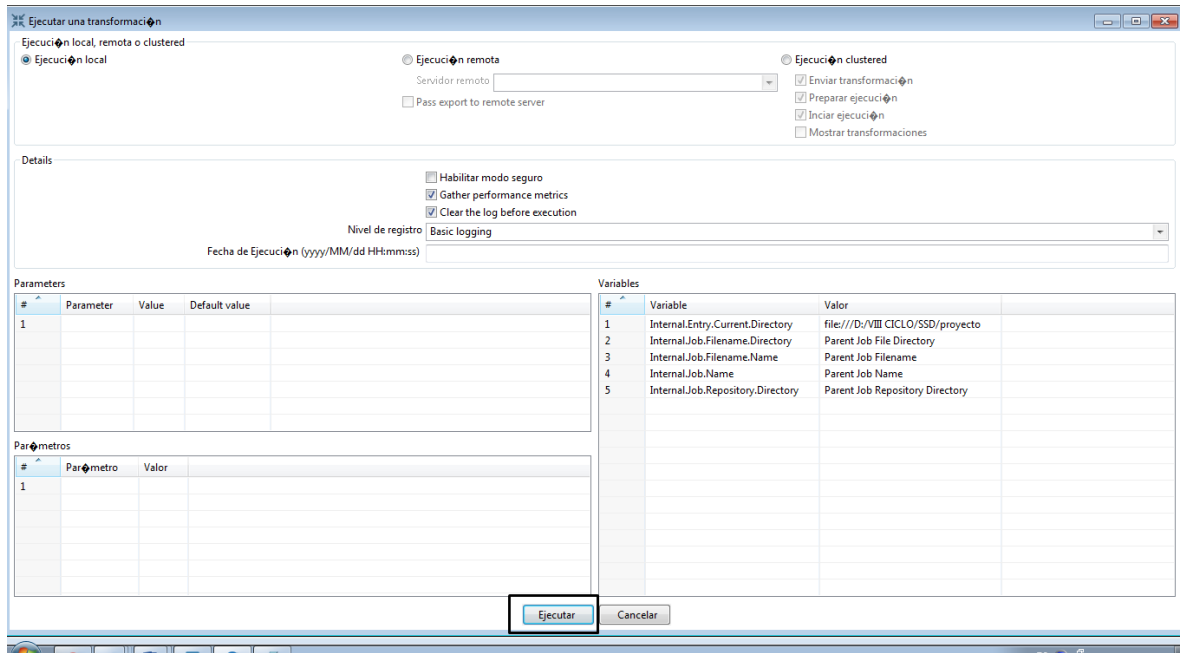
Se indica la tabla dimensional y la conexión a utilizar



Luego hacemos el mapeo de componentes



Procedemos a ejecutar y cargar la dimensión cliente

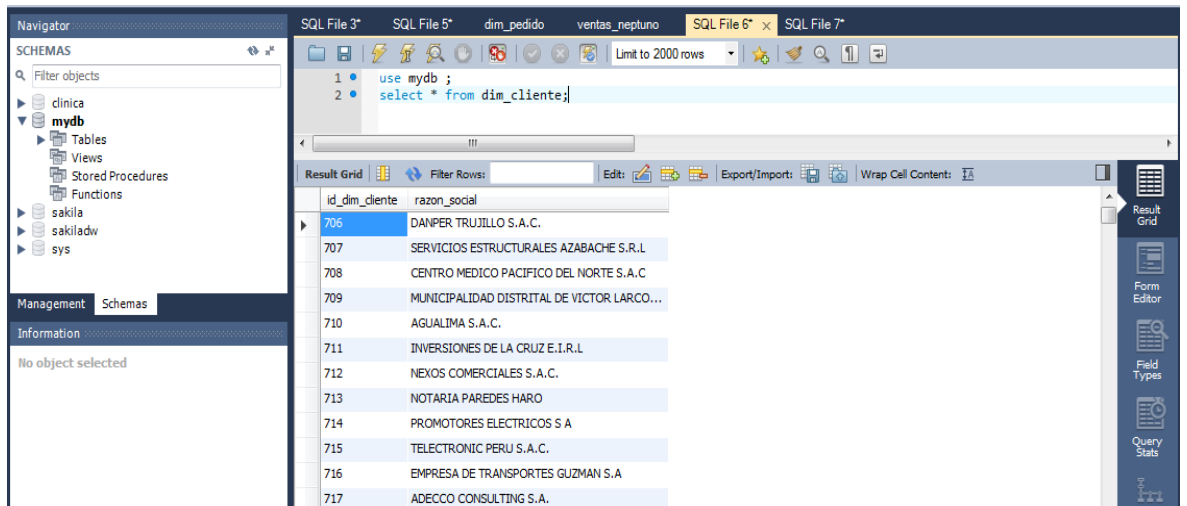


Execution Results

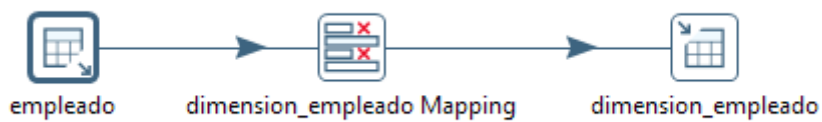
Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo	Tiempo	Velocidad (r/s)
1	cliente	0	0	235	235	0	0	0	0	Finalizado	0.1s	2.500
2	Table output Mapping	0	235	235	0	0	0	0	0	Finalizado	0.1s	2.398
3	dim_cliente	0	235	235	0	235	0	0	0	Finalizado	0.4s	651

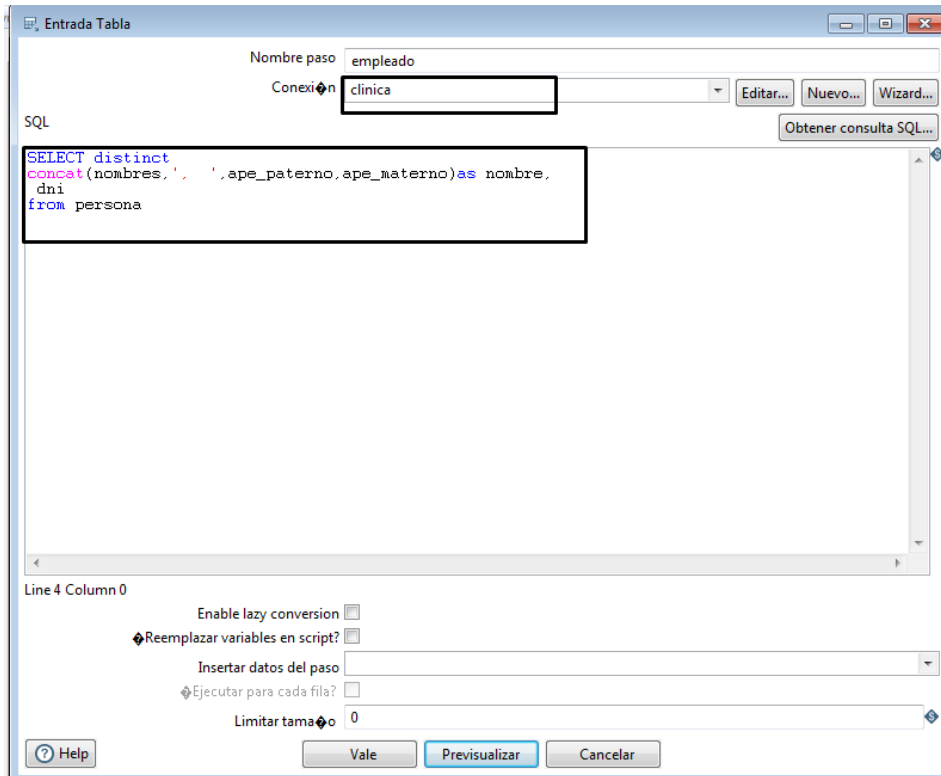
Y podemos observar la dimensión cliente llena de datos



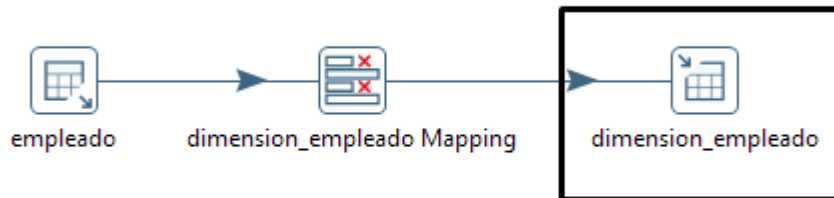
Dimensión Empleado

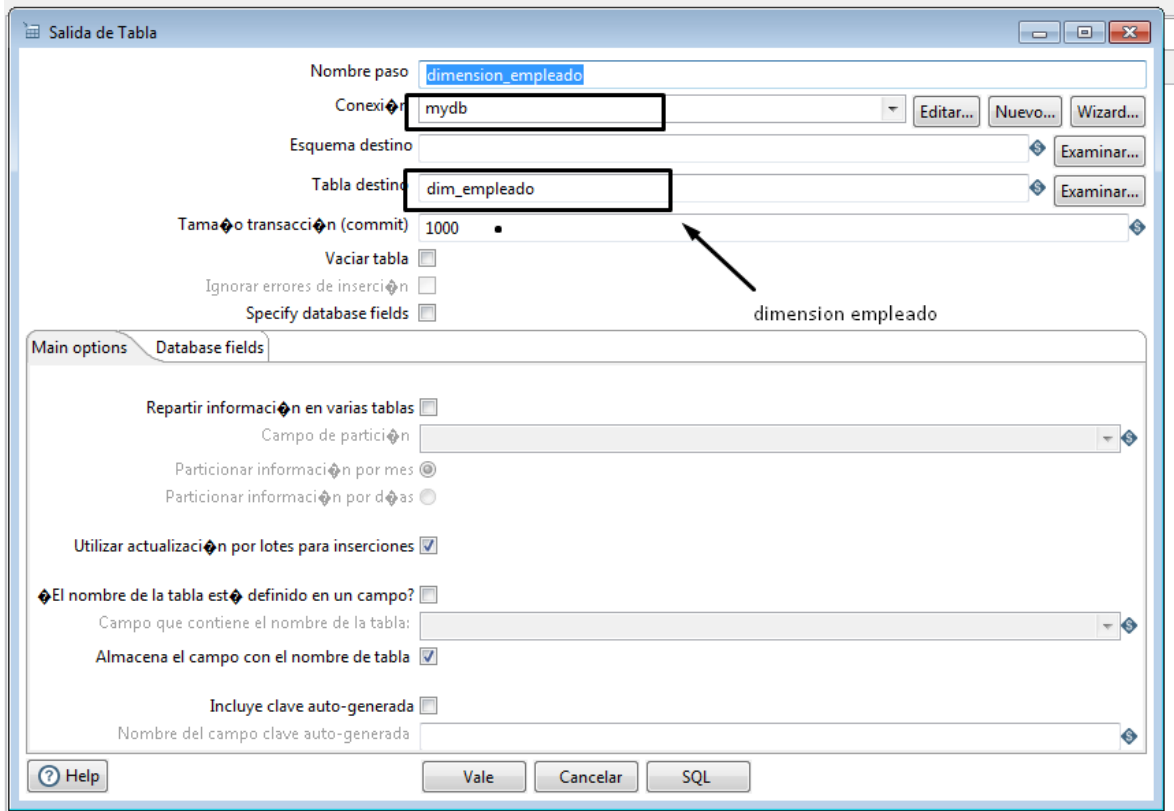


Seleccionamos la conexión e indicamos la sentencia para seleccionar los datos de la base de datos transaccional

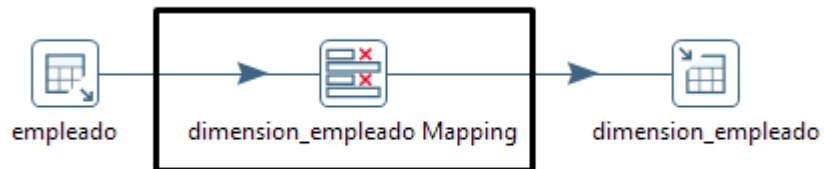


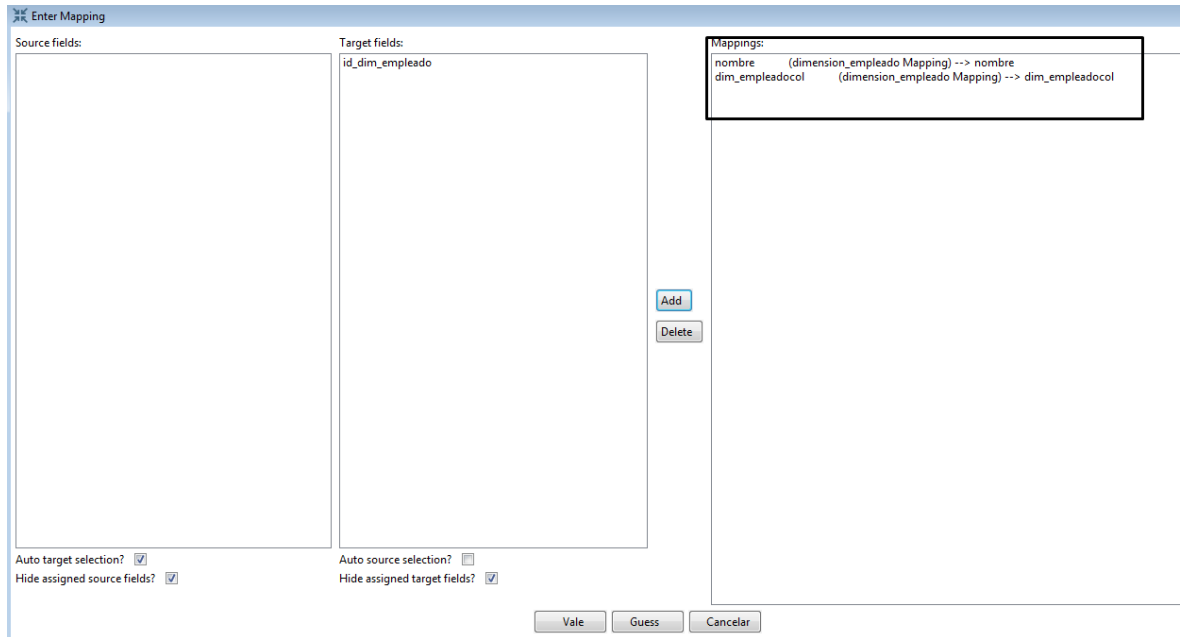
Indicamos a que dimensión esta direccionado estos datos tomados anteriormente





Luego procedimos a hacer el mapeo correspondiente





Si siguiendo con el proceso ejecutamos para cargar la dimensión empleado



Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo	Tiempo	Vel
1	empleado	0	0	36	36	0	0	0	0	Finalizado	0.1s	
2	dimension_empleado Mapping	0	36	36	0	0	0	0	0	Finalizado	0.1s	
3	dimension_empleado	0	36	36	0	36	0	0	0	Finalizado	0.3s	

Y observamos los datos en la dimensión empleado

```

1 • use mydb ;
2 • select * from dim_empleado;

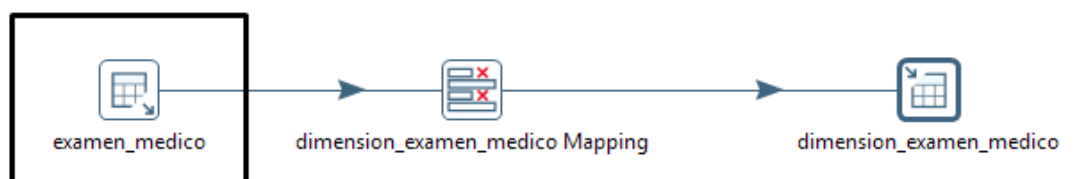
```

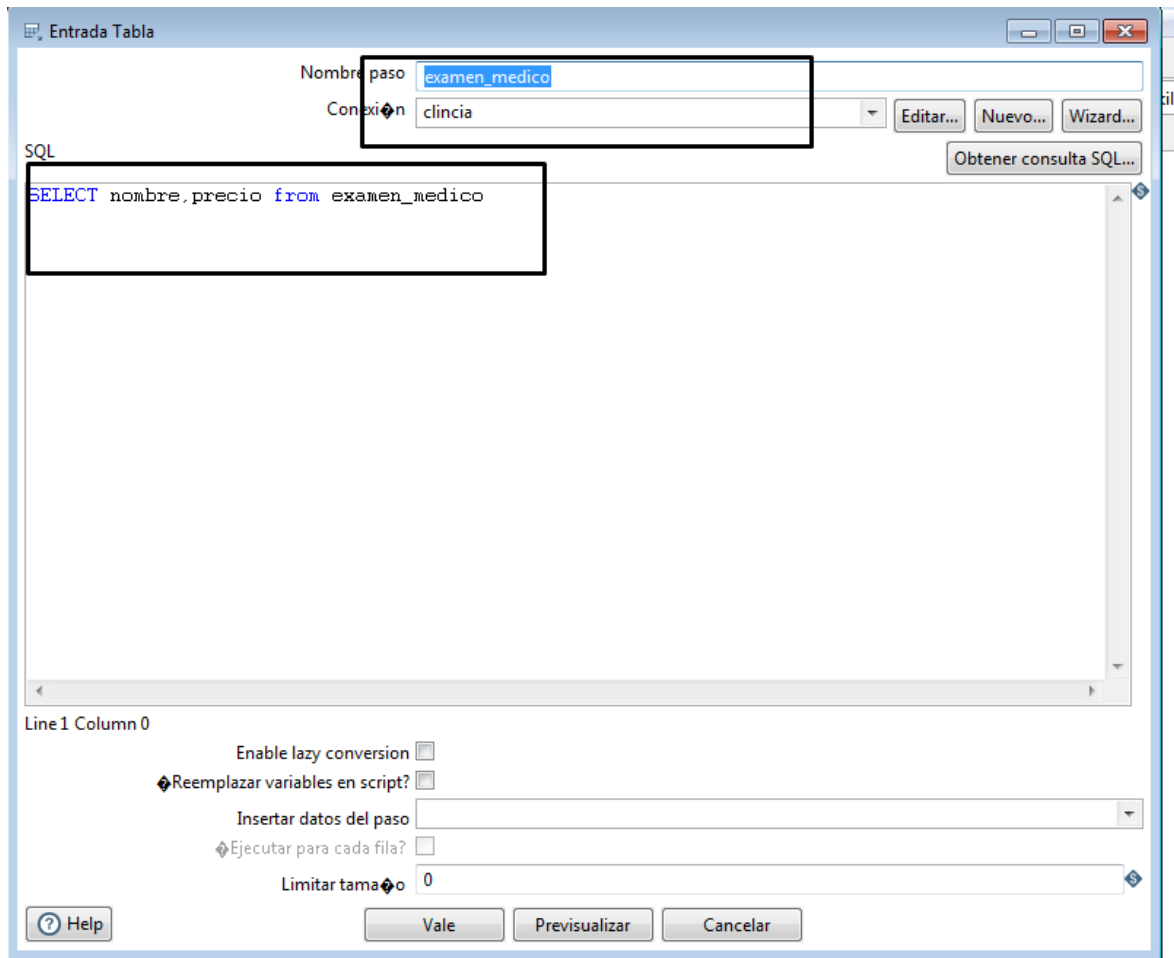
id_dim_empleado	nombre	dim_empleadocol
1	VICTOR ALEXANDER, JAVESACRAMENTO	45909624
2	JESSICA, ENCOGUTIERREZ	43074143
3	NORA, SALAZARN	40204365
4	INGRID, MEDINA TUESTA	45400122
5	ANGEL , PEREDA.	46450501
6	LILLIANA, GAMARRAGARCIA	18104441
7	SILVIA, MARINGOICOCHEA	45217699
8	FRANK, VALVERDECRUZ	41850887
9	ROCIO , HERRERAE	42022814
10	MARIA ISABEL, PAZPEÑA	42599665
11	JOVERTH DENIS, CUSTODIOTORRES	73183329
12	GERALDINE KIMAYRA, VASQUEZRODRI...	71001311

Dimensión examen medico

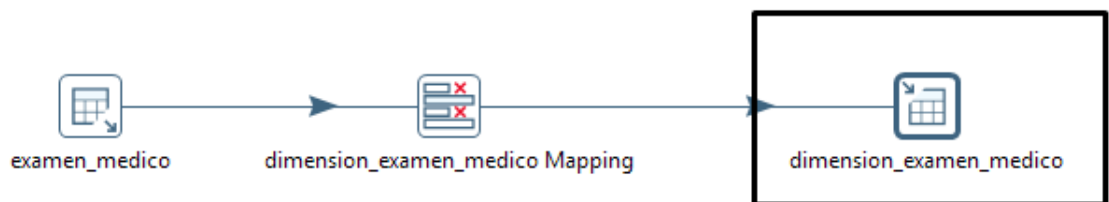


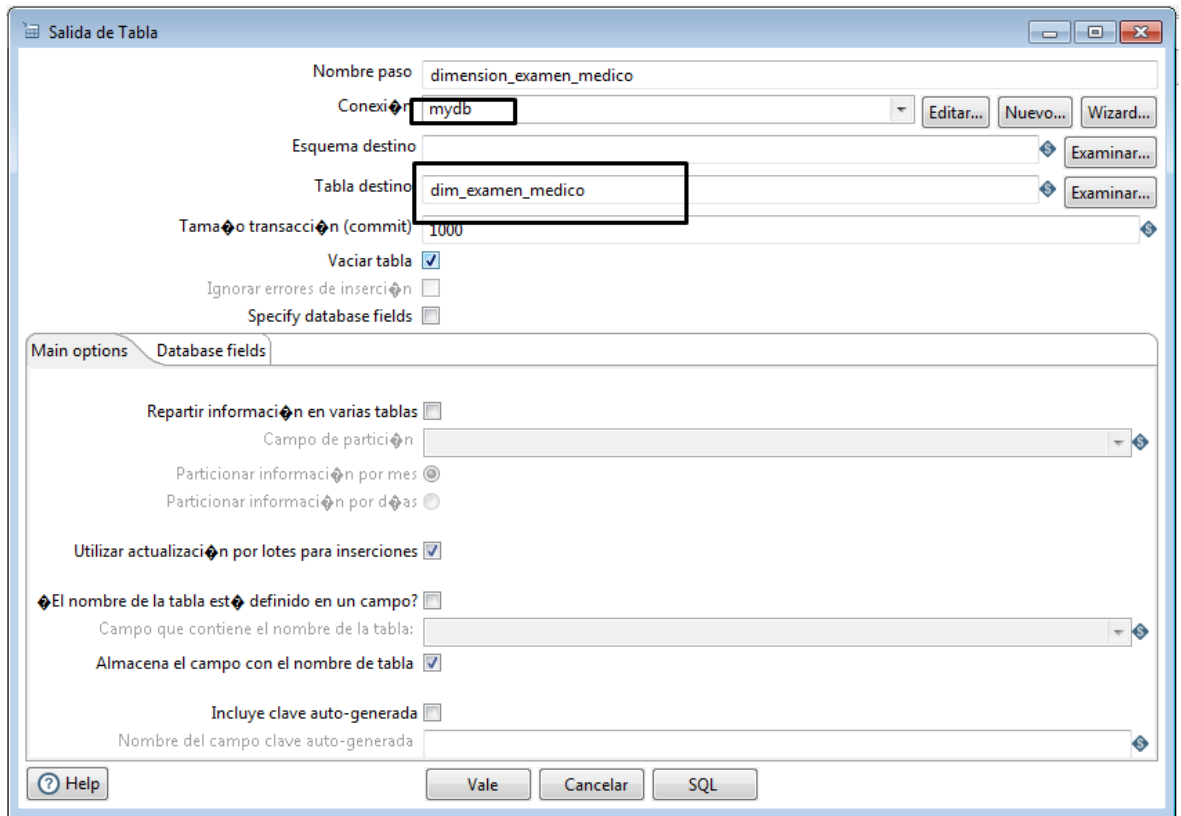
Establecemos la conexión y la consulta para extraer los datos requeridos para el datawarehouse como se muestra a continuación



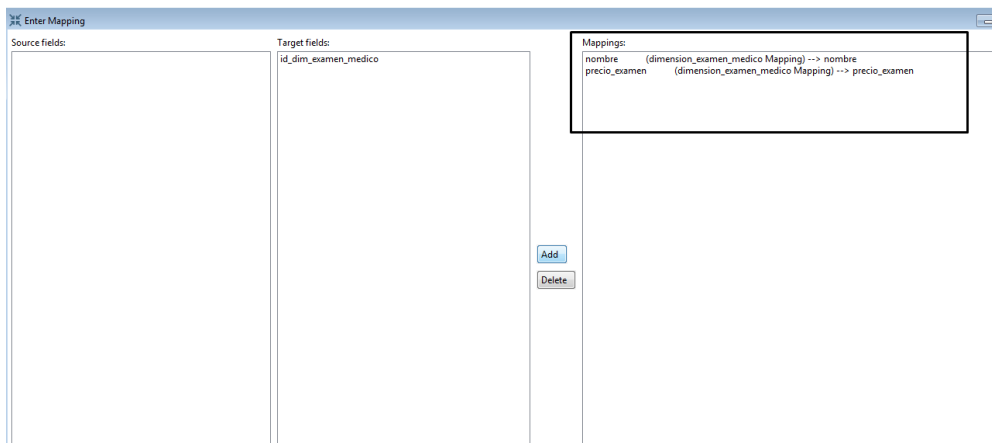
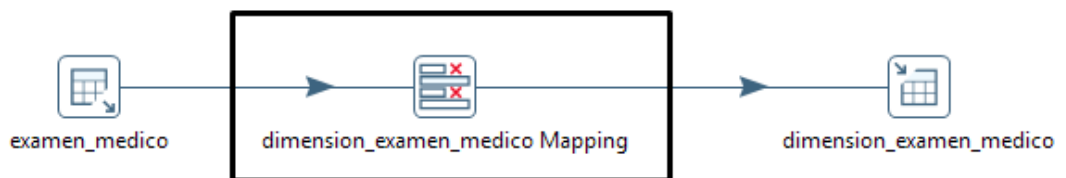


Luego indicamos a que tabla dimensional va dirigida

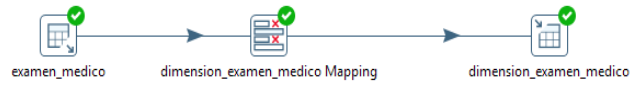




y luego proseguimos a hacer el mapeo correspondiente



Procedemos a ejecutar la carga de la dimensión examen medico



Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo	Tiempo
1	examen_medico	0	0	11	11	0	0	0	0	Finalizado	0.0s
2	dimension_examen_medico Mapping	0	11	11	0	0	0	0	0	Finalizado	0.0s
3	dimension_examen_medico	0	11	0	0	0	0	0	0	Ejecutando	0.0s

Y comprobamos que realmente se haya cargado la dimensión

SQL File 3* | SQL File 5* | dim_pedido | ventas_neptuno | SQL File 6* | SQL File 7* x | historia_clinica

```

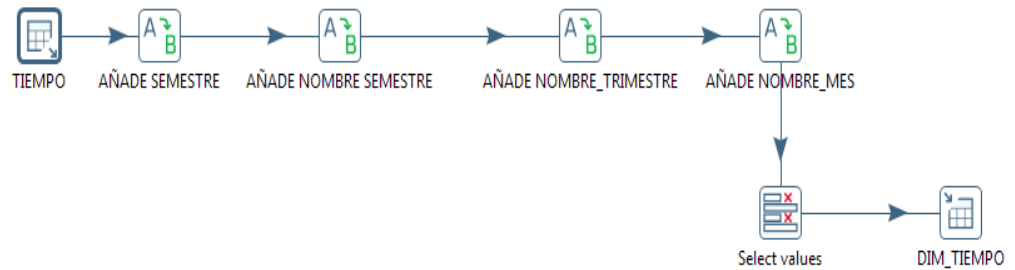
1 • use mydb ;
2 • select * from dim_examen_medico
  
```

Limit to 2000 rows

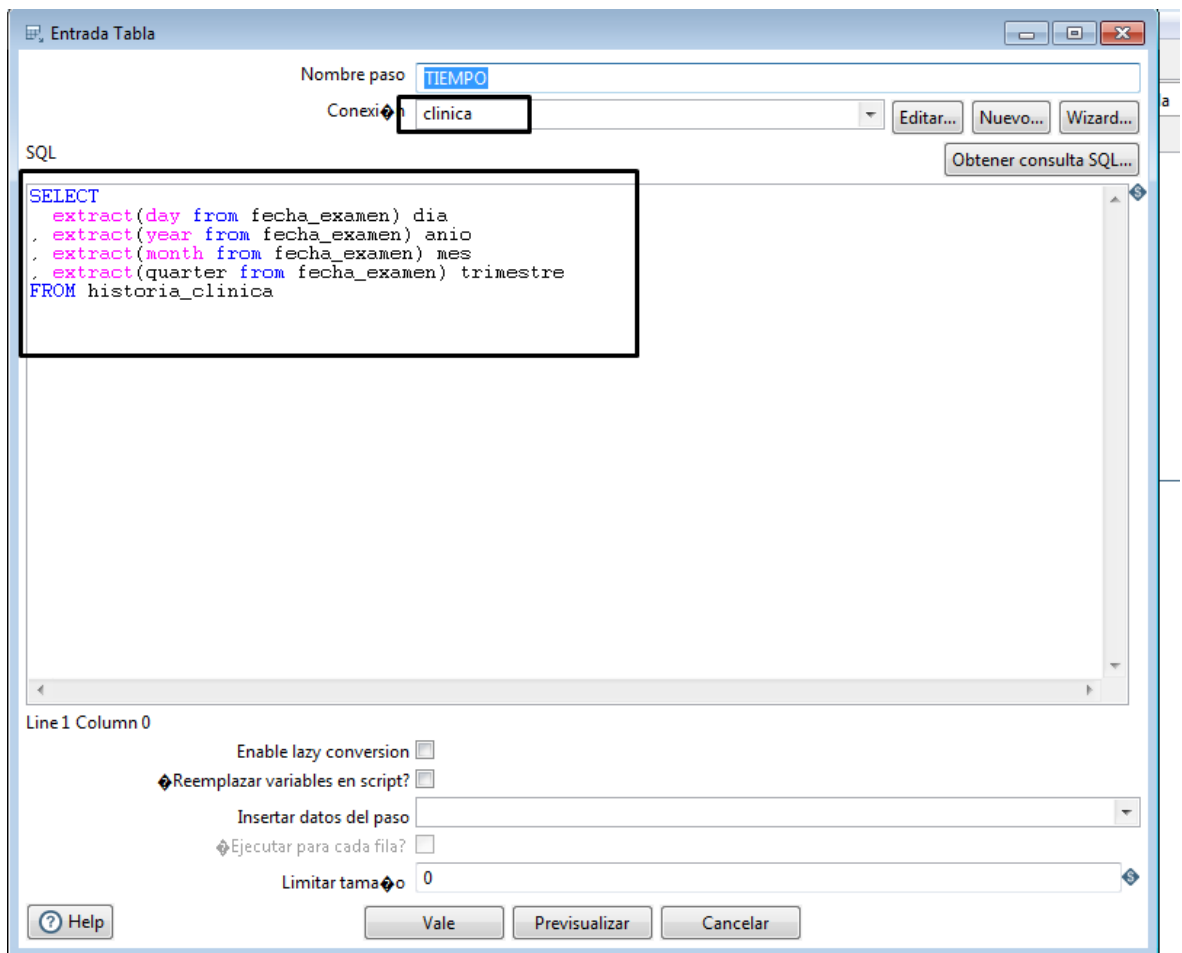
Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: IA

	id_dim_examen_medico	nombre	precio_examen
▶	1	PRE OCUPACIONAL	150
	2	PERIÓDICO	120
	3	RETIRO	120
	4	REUBICACIÓN	100
	5

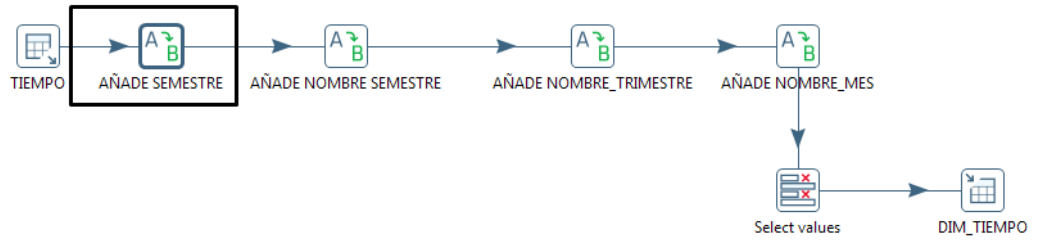
Dimension tiempo



Establecemos la conexión y la sentencias sql a utilizar para poder tener los datos requeridos para la tabla dimensional



Hacemos la primera transformación



Mapeo de Valores

Nombre de paso : AÑADE SEMESTRE

Nombre de campo origen : trimestre

Nombre de campo destino : semestre

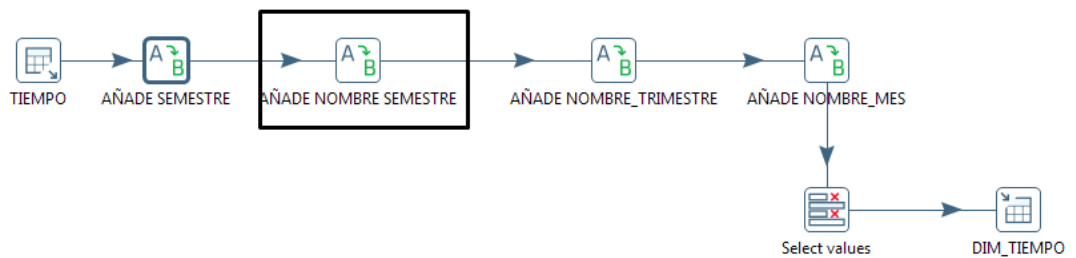
Default upon non-matching :

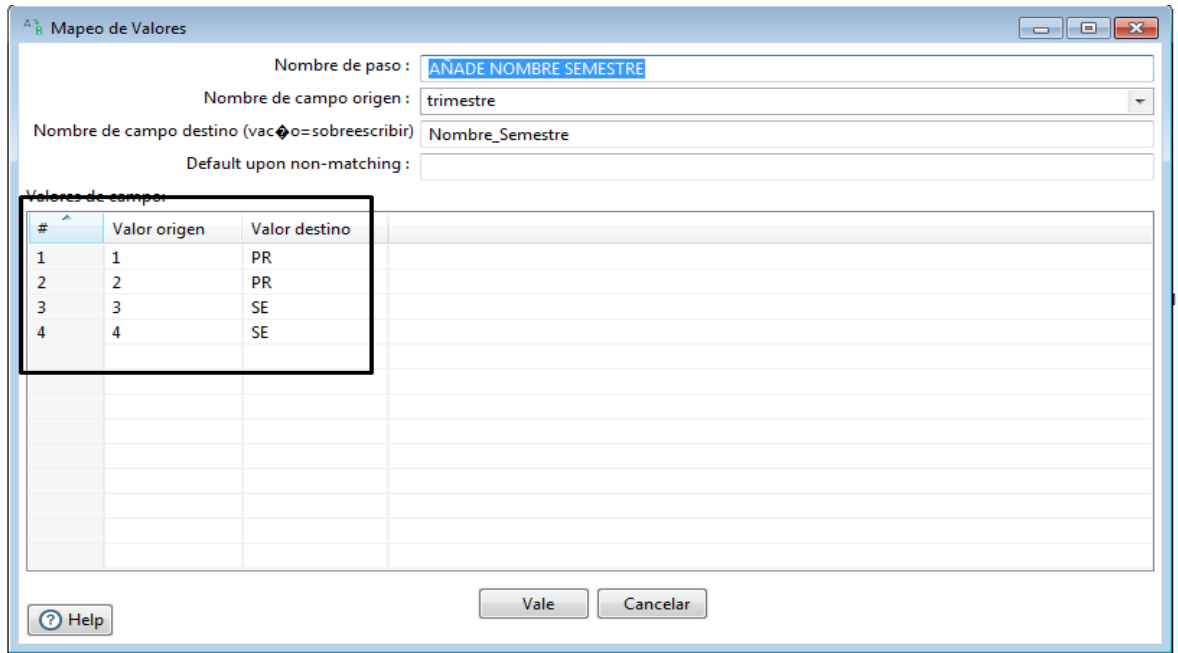
Valores de campo:

#	Valor origen	Valor destino
1	1	1
2	2	1
3	3	2
4	4	2

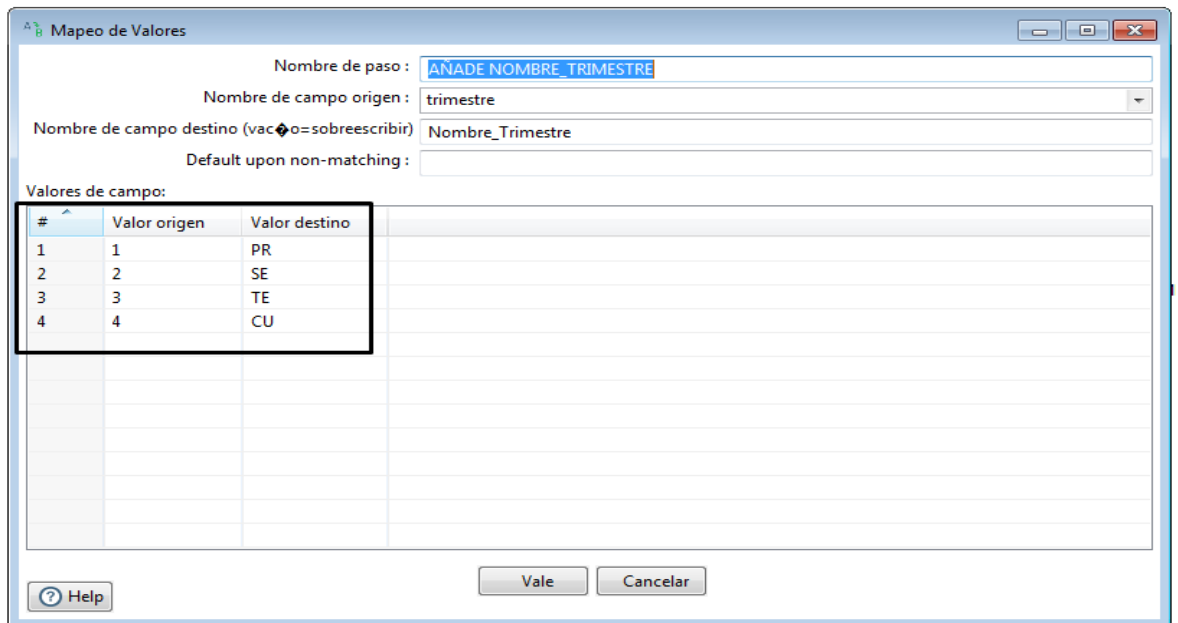
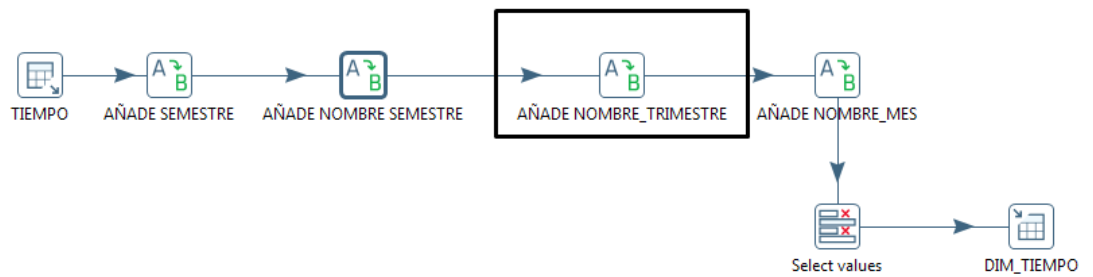
Help Vale Cancelar

Hacemos la segunda transformación

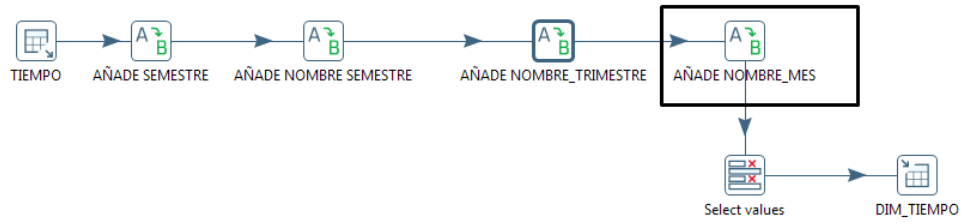




Hacemos la tercera transformación



Cuarta transformación



Mapeo de Valores

Nombre de paso: **AÑADE NOMBRE_MES**

Nombre de campo origen: mes

Nombre de campo destino (vacío=sobrescribir): NombreMes

Default upon non-matching:

#	Valor origen	Valor destino
1	1	ENERO
2	2	FEBRERO
3	3	MARZO
4	4	ABRIL
5	5	MAYO
6	6	JUNIO
7	7	JULIO
8	8	AGOSTO
9	9	SEPTIEMBRE
10	10	OCTUBRE
11	11	NOVIEMBRE
12	12	DICIEMBRE

Vale Cancelar

Determinamos la dimensión a la cual va direccionado

Salida de Tabla

Nombre paso: **DIM_TIEMPO**

Conexión: mydb

Esquema destino:

Tabla destino: dim_tiempo

Tamaño de transacción (commit): 1000

Vaciar tabla:

Ignorar errores de inserción:

Specify database fields:

Main options Database fields

Repartir información en varias tablas:

Campo de partición:

Particionar información por mes:

Particionar información por días:

Utilizar actualización por lotes para inserciones:

El nombre de la tabla está definido en un campo?:

Campo que contiene el nombre de la tabla:

Almacena el campo con el nombre de tabla:

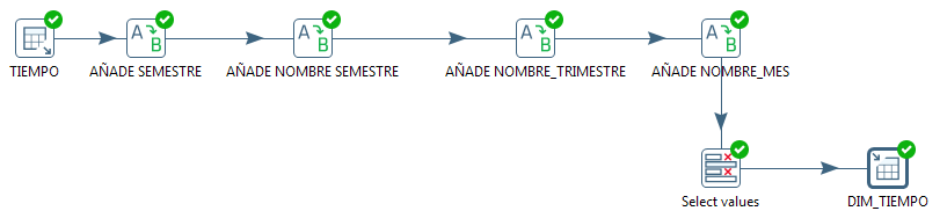
Incluye clave auto-generada:

Nombre del campo clave auto-generada:

Vale Cancelar SQL

Y el mapeo respectivo

Ejecutamos para poblar la dimensión tiempo



Execution Results

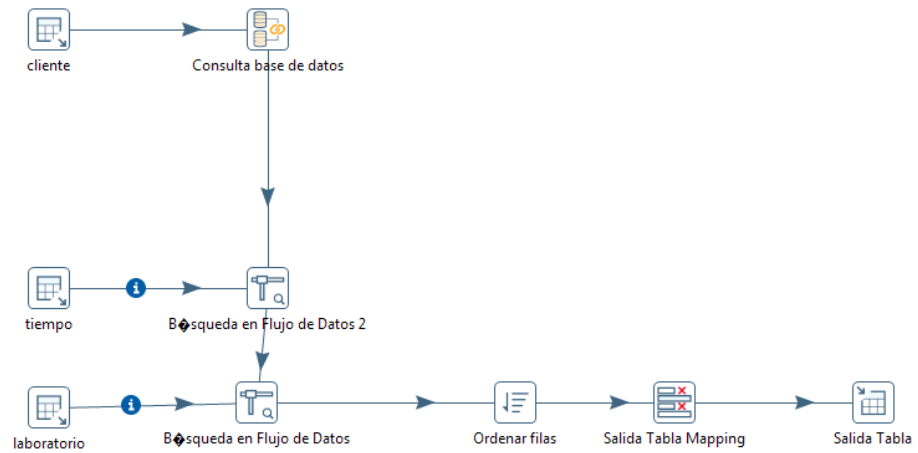
Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo	Tiem
1	TIEMPO	0	0	9108	9108	0	0	0	0	Finalizado	0.
2	AÑADE SEMESTRE	0	9108	9108	0	0	0	0	0	Finalizado	0.
3	AÑADE NOMBRE SEMESTRE	0	9108	9108	0	0	0	0	0	Finalizado	0.
4	AÑADE NOMBRE_TRIMESTRE	0	9108	9108	0	0	0	0	0	Finalizado	0.
5	AÑADE NOMBRE_MES	0	9108	9108	0	0	0	0	0	Finalizado	0.

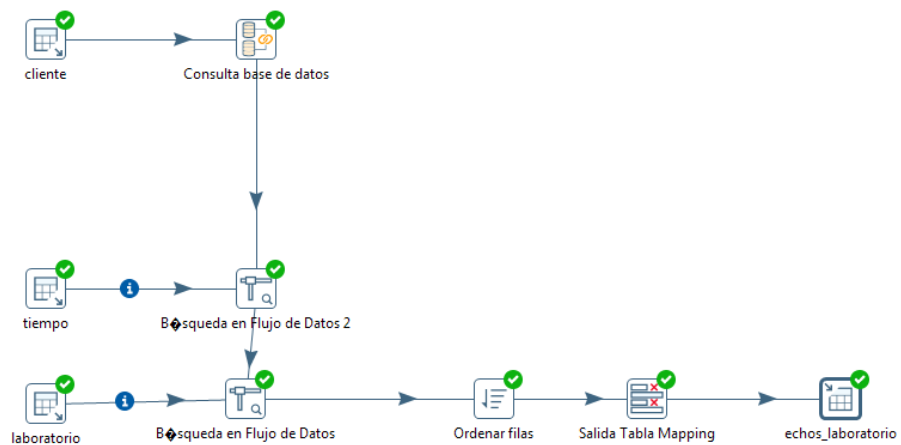
Y verificamos si es que pobló dicha dimensión

	DIMT_ID	ANIO	SEMESTRE	NOMBRE_SEMESTRE	TRIMESTRE	NOMBRE_TRIMESTRE	MES	NOMBRE_MES	DIA
▶	1	2016	2	SE	3	TE	7	JULIO	11
	2	2016	1	PR	2	SE	5	MAYO	17
	3	2016	1	PR	2	SE	5	MAYO	17
	4	2016	2	SE	3	TE	7	JULIO	14

Hechos laboratorio



Ejecutamos



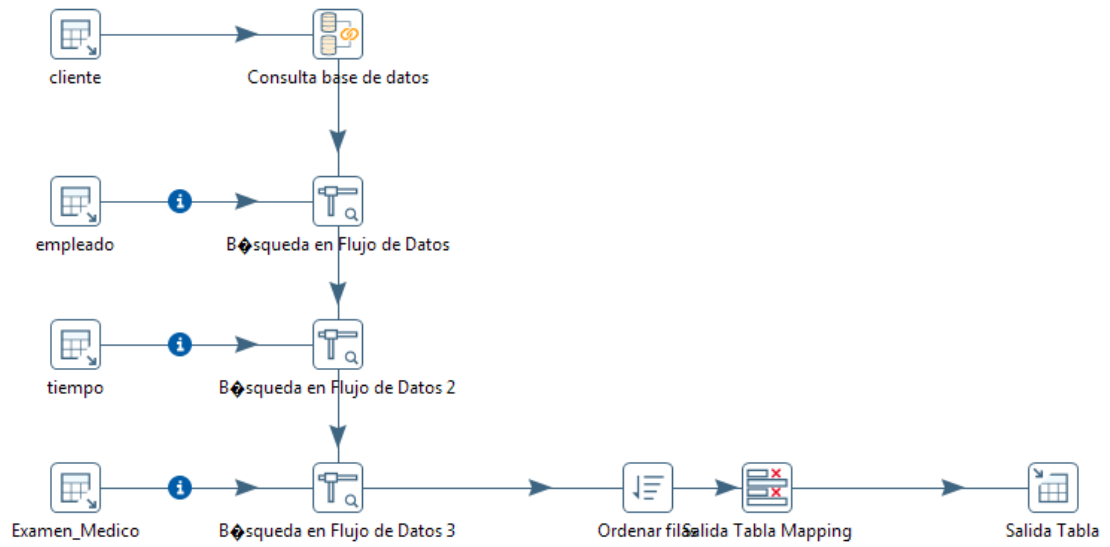
Luego verificamos si a poblado la tabla de hechos laboratorio

SQL Query Browser - Connection: root@localhost:3306 / clinica

SELECT * FROM mydb.hechos_laboratorio h;

id_hechos_lab	indc_cholesterol	indc_glucosa	indc_hglicoides	indc_colinester...	dim_laboratorio...	dim_cliente_id...	dim_tiempo_DI...
5142	0	0	0	0	7709	2169	36432
5143	0	0	0	0	7709	2160	36432
5144	0	0	0	0	7709	2158	36384
5145	0	0	0	0	7709	2143	36083
5146	0	0	0	0	7709	2132	36432
5147	0	0	0	0	7709	2132	36384
5148	0	0	0	0	7709	2125	35959
5149	0	0	0	0	7709	2102	33049
5150	0	0	0	0	7709	2100	33049
5151	0	0	0	0	7709	2098	33049
5152	0	0	0	0	7709	2097	33049
5153	0	0	0	0	7709	2096	34964
5154	0	0	0	0	7709	2093	31537
5155	0	0	0	0	7709	2083	31537
5156	0	0	0	0	7709	2083	31537
5157	0	0	0	0	7709	2066	30680
5158	0	0	0	0	7709	2034	36432
5159	0	0	0	0	7709	2034	36432
5160	0	0	0	0	7709	2034	36432
5161	0	0	0	0	7709	2032	30680
5162	0	0	0	0	7709	2026	33049
5163	0	0	0	0	7709	2026	33049
5164	0	0	0	0	7709	2026	31537
5165	0	0	0	0	7709	2026	30680
5166	0	0	0	0	7709	2026	30680
5167	0	0	0	0	7709	2026	30680
5168	0	0	0	0	7709	2026	33080
5169	0	0	0	0	7709	2026	33080
5170	0	0	0	0	7709	2026	33080

Finalmente la tabla de hechos de admisión



Y verificamos que se haya poblado la tabla de hechos

MySQL Query Browser - Connection: root@localhost:3306 / clinica

File Edit View Query Script Tools Window Help

Go back Next Refresh

SELECT * FROM mydb.hechos_admision h;

Execute Stop

Resultset 1 Script 1 Resultset 2 Script 2 Resultset 3

id_hechos_...	precio_examen	monto_total	dim_cliente_id_...	dim_tiempo_DI...	dim_examen_m...	dim_empleado_...
9056	150	150	2175	36432	34	159
9057	150	150	2175	36432	34	159
9058	150	150	2174	36432	34	159
9059	75	75	2174	36384	40	159
9060	120	120	2173	36432	35	159
9061	120	120	2173	36432	35	159
9062	120	120	2173	36384	35	159
9063	120	120	2173	36384	35	159
9064	120	120	2173	36384	35	159
9065	120	120	2173	36384	35	159
9066	120	120	2173	36384	35	159
9067	120	120	2173	36384	35	159
9068	120	120	2173	36384	35	159
9069	120	120	2173	36384	35	159
9070	120	120	2173	36384	35	159
9071	150	150	2172	36384	34	159
9072	150	150	2171	36384	34	159
9073	150	150	2170	36432	34	159
9074	150	150	2170	36432	34	159
9075	150	150	2170	36432	34	159
9076	150	150	2170	36432	34	159
9077	150	150	2170	36432	34	159
9078	150	150	2170	36432	34	159
9079	150	150	2170	36432	34	159
9080	150	150	2170	36432	34	159
9081	150	150	2170	36432	34	159
.....

Schemata Bookmarks History

- ferrari
- ferrari_dw
- icecream
- icecream_dw
- information_schema
- inventioite
- ivpedu_sge
- mydb
 - dim_cliente
 - dim_empleado
 - dim_examen_medico
 - dim_laboratorio
 - dim_tiempo
 - hechos_admision
 - hechos_laboratorio
- mysql
- netcomputer_dw
- peniarga_typographic
- performance_schema

Syntax Functions Params Trx

- Data Definition Statements
- Data Manipulation Statements
- MySQL Utility Statements
- MySQL Transactional and Locking ...
- Database Administration Statements
- Replication Statements
- SQL Syntax for Prepared Statements

4.5.3. CREACIÓN DE LA SOLUCIÓN DE BI EN MICROSOFT AZURE

PASO1: Ingresando a Microsoft Azure y la Creación de la base de datos

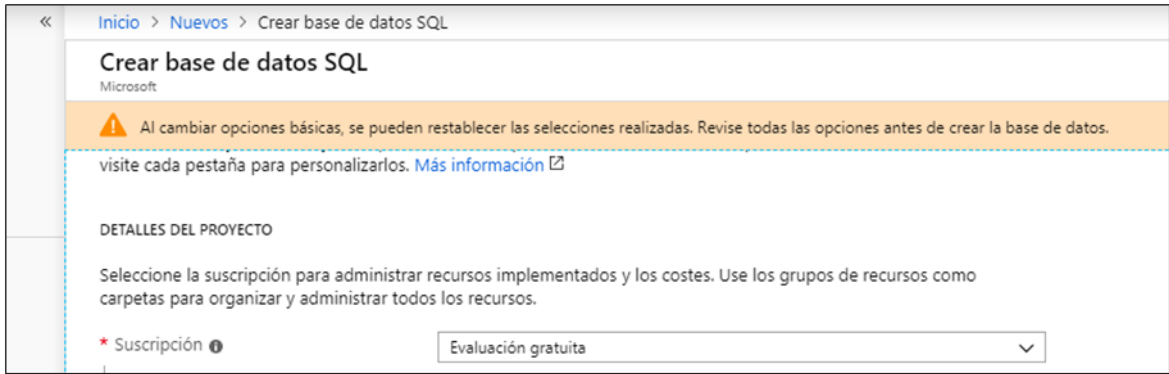
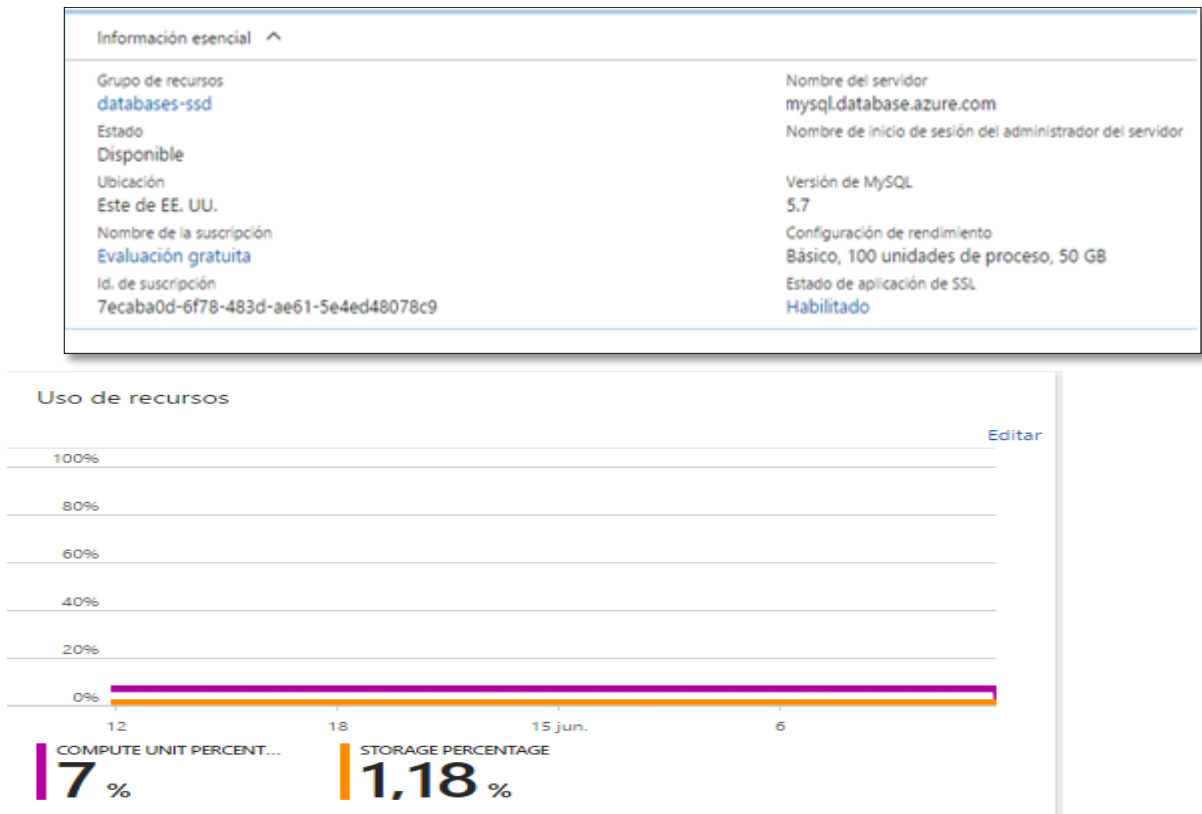


Figura 18: Creación base de datos en Azure

PASO 2: Verificación del estado de la base de datos



Bases de datos MySQL

2 Bases de datos

NOMBRE

clinica

clinica-datamart

4.5.4. DESARROLLO DE APLICACIONES

Para la parte de implementación en la tabla siguiente se detalla las herramientas a utilizar y los procesos a ejecutar.

Tabla 7: Software para el proceso de Diseño del BI

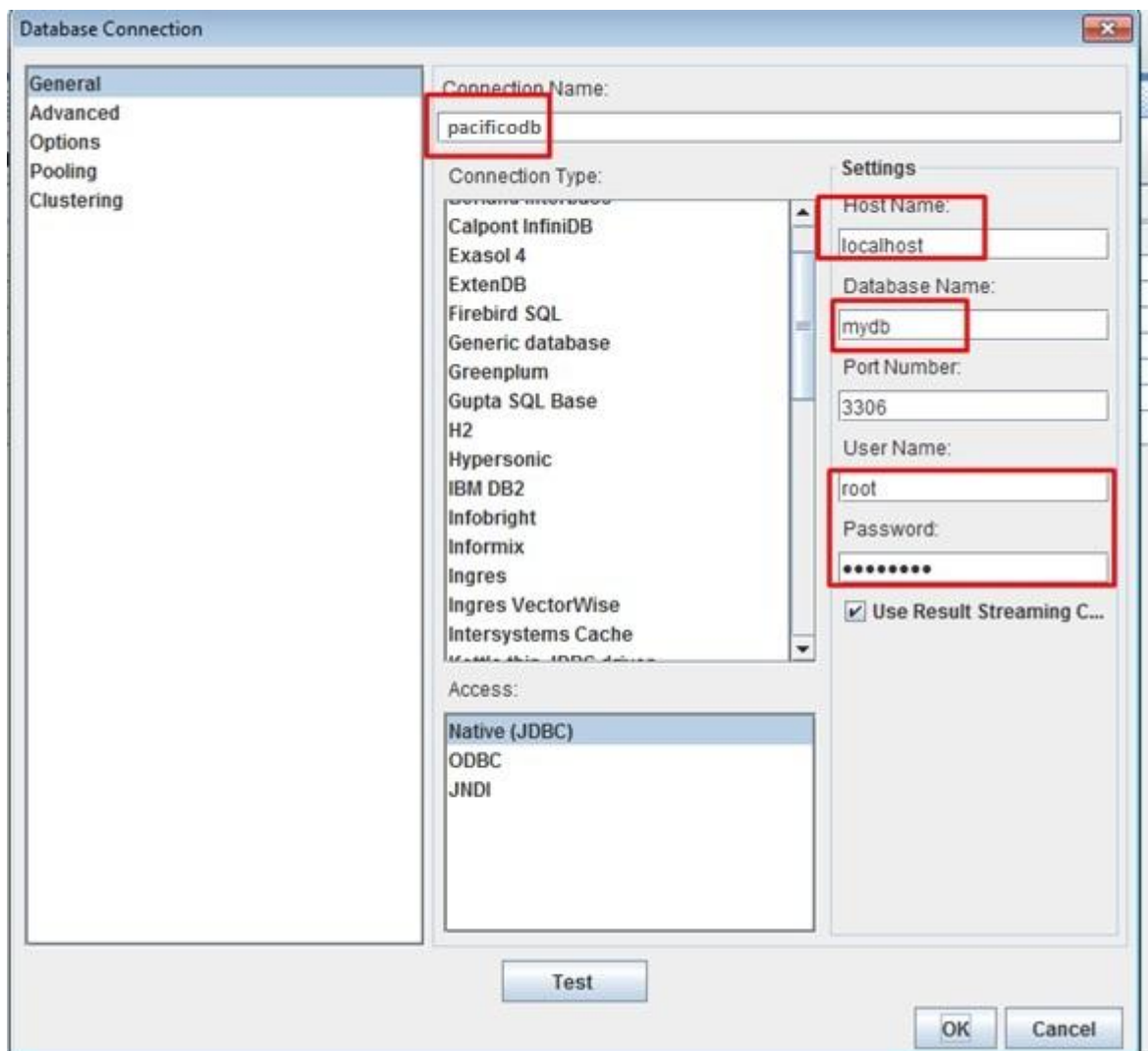
Componente del Proceso	Herramienta
Construcción de Interfaces	Pentaho report designer
Cubos OLAP	Pentaho Cube Designer
Poblamiento ETL	Pentaho Data Integration

ESPECIFICACIÓN DE LA APLICACIÓN DEL USUARIO FINAL

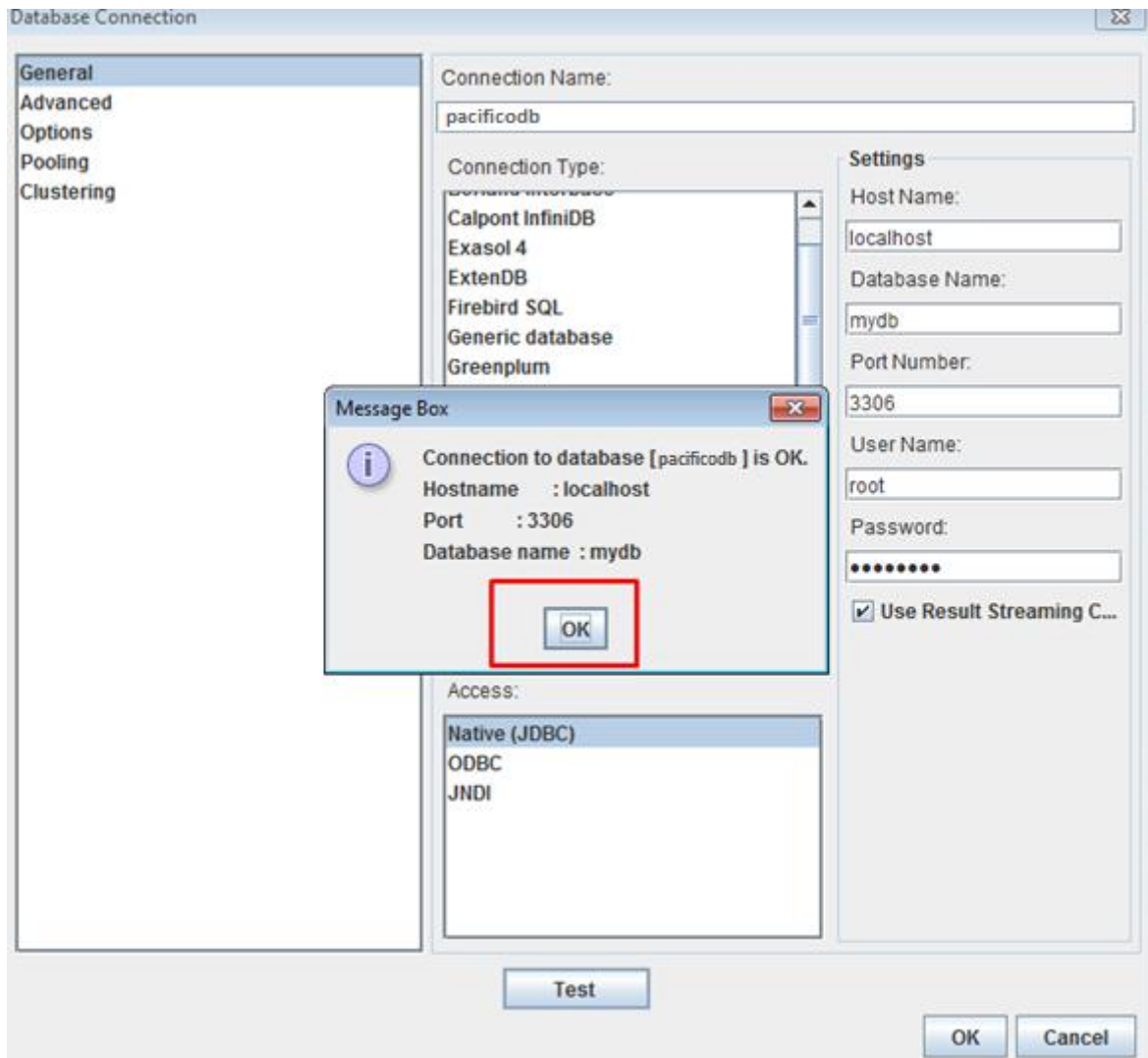
✓ Construcción del Cubo

Estructura de Cubo: Utilizando Schema workbench

Para eso primero creamos la conexión en la cual especificamos a que base de datos tenemos la conexión el servidor y las credenciales para dicho server



Y probamos que la conexión este establecida



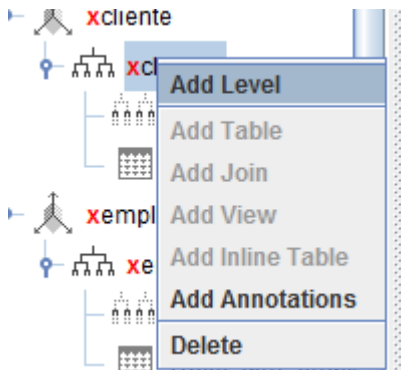
Luego empezamos a crear las dimensiones en la cual la primera es la dimensión cliente.

Shared Dimension	
Attribute	Value
name	cliente
description	
foreignKey	
type	StandardDimension
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

Seguidamente colocamos la tabla a la cual pertenece dicha dimensión

Table for 'cliente' Hierarchy	
Attribute	Value
schema	
name	dim_cliente
alias	

Posteriormente agregamos un nivel



Y posteriormente colocamos los atributos de dicho nivel

Level for 'cliente' Hierarchy	
Attribute	Value
name	cliente
description	
table	dim_cliente
column	id_dim_cliente
nameColumn	razon_social
parentColumn	
nullParentValue	
ordinalColumn	
type	String
internalType	
uniqueMembers	<input checked="" type="checkbox"/>
levelType	Regular
hideMemberif	Never
approxRowCount	
caption	
captionColumn	
formatter	
visible	<input checked="" type="checkbox"/>

En segundo lugar, pasamos a la dimensión empleado

Shared Dimension	
Attribute	Value
name	empleado
description	
foreignKey	
type	StandardDimension
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

Agregamos la tabla

Table for 'empleado' Hierarchy	
Attribute	Value
schema	
name	dim_empleado
alias	

Y luego el nivel

Level for 'empleado' Hierarchy	
Attribute	Value
name	empleado
description	
table	dim_empleado
column	id_dim_empleado
nameColumn	nombre
parentColumn	
nullParentValue	
ordinalColumn	
type	String
internalType	
uniqueMembers	<input checked="" type="checkbox"/>
levelType	Regular
hideMemberIf	Never
approxRowCount	
caption	
captionColumn	
formatter	
visible	<input checked="" type="checkbox"/>

Como tercera dimensión tenemos a examen médico

Shared Dimension	
Attribute	Value
name	examen_medico
description	
foreignKey	
type	StandardDimension
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

Agregamos la tabla correspondiente a dicha dimensión

Table for 'examen_medico' Hierarchy	
Attribute	Value
schema	
name	dim_examen_medico
alias	

Así mismo agregamos un nivel

Level for 'examen_medico' Hierarchy	
Attribute	Value
name	examen_medico
description	
table	dim_examen_medico
column	id_dim_examen_medico
nameColumn	nombre
parentColumn	
nullParentValue	
ordinalColumn	
type	String
internalType	
uniqueMembers	<input checked="" type="checkbox"/>
levelType	Regular
hideMemberIf	Never
approxRowCount	
caption	
captionColumn	
formatter	
visible	<input checked="" type="checkbox"/>

Y pasamos a la dimensión tiempo

Shared Dimension	
Attribute	Value
name	tiempo
description	
foreignKey	
type	StandardDimension
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

Agregamos la tabla

Table for 'tiempo' Hierarchy	
Attribute	Value
schema	
name	dim_tiempo
alias	

Luego un nivel

Attribute	
name	mes
description	
table	dim_tiempo
column	MES
nameColumn	NOMBRE_MES
parentColumn	
nullParentValue	
ordinalColumn	
type	String
internalType	
uniqueMembers	<input checked="" type="checkbox"/>
levelType	Regular
hideMemberIf	Never
approxRowCount	
caption	
captionColumn	
formatter	
visible	<input checked="" type="checkbox"/>

De igual manera para la dimensión laboratorio

Shared Dimension	
Attribute	Value
name	dimension_laboratorio
description	
foreignKey	
type	StandardDimension
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

Agregamos la tabla

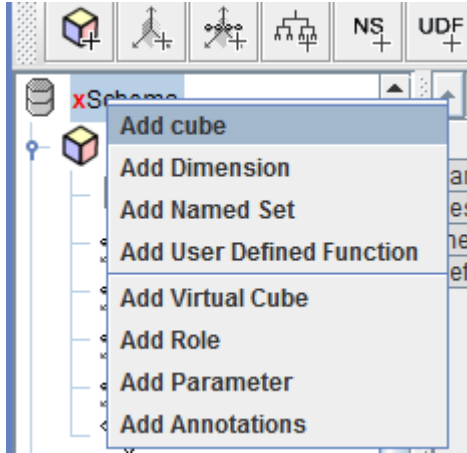
Table for 'New Hierarchy 0' Hierarchy	
Attribute	Value
schema	
name	dim_laboratorio
alias	

Y posteriormente un nivel

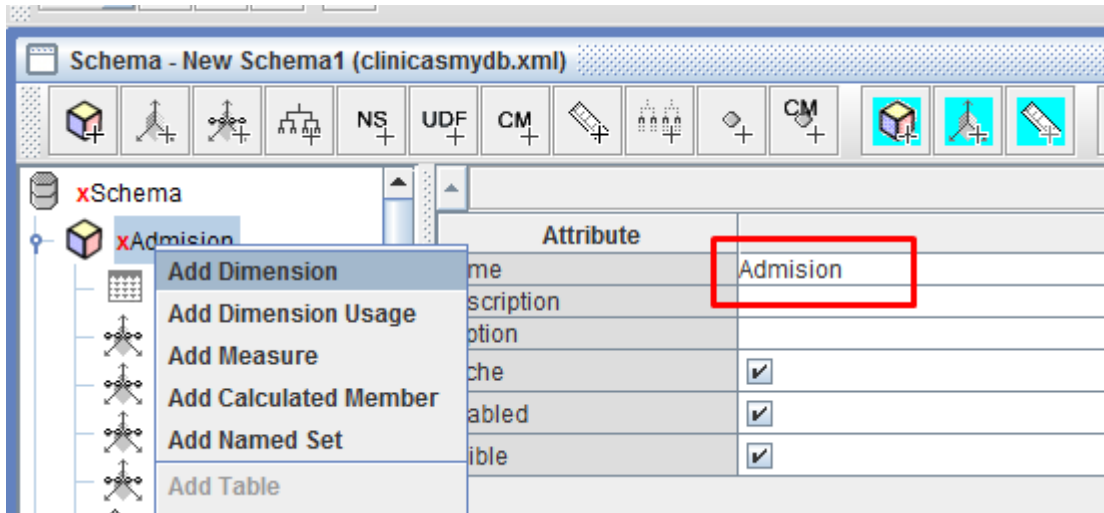
Level for 'New Hierarchy 0' Hierarchy	
Attribute	Value
name	laboratorio
description	
table	dim_laboratorio
column	id_dim_laboratorio
nameColumn	colesterol
parentColumn	
nullParentValue	
ordinalColumn	
type	String
internalType	
uniqueMembers	<input checked="" type="checkbox"/>
levelType	Regular
hideMemberf	Never
approxRowCount	
caption	
captionColumn	
formatter	
visible	<input checked="" type="checkbox"/>

CREACIÓN DE LOS CUBOS

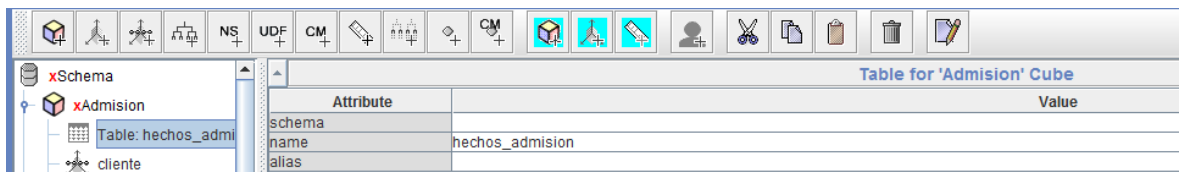
Creamos el cubo luego de haberse creado las dimensiones mostradas en el apartado anterior



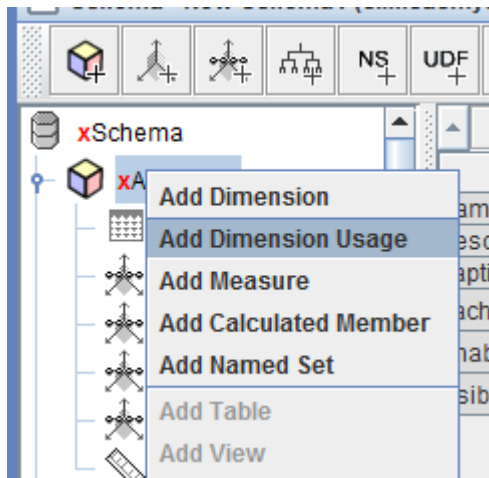
Se utilizará 2 cubos uno para admisión y otra para laboratorio ahora empezaremos por Admisión



Luego procedemos a agregarle la tabla y las dimensiones así que empezamos con la tabla.



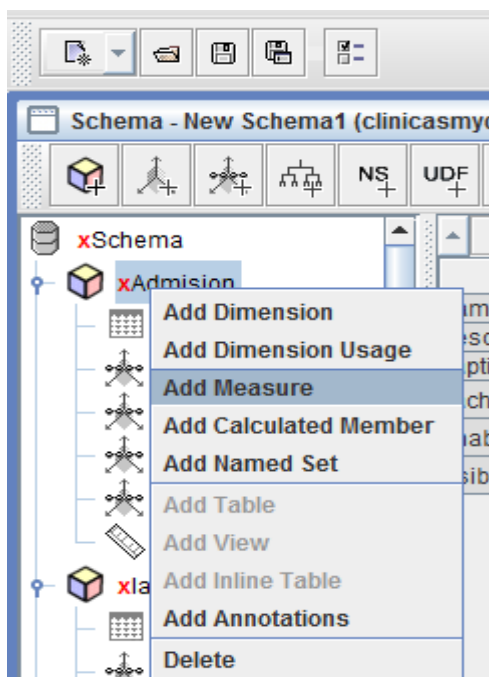
Y como ya fueron creadas las dimensiones ahora solo las referenciaremos.

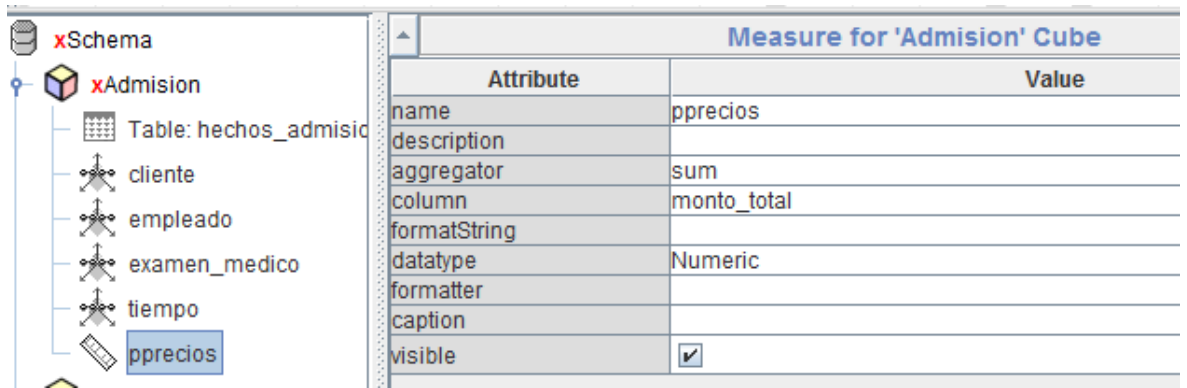


Posteriormente indicamos las referencias

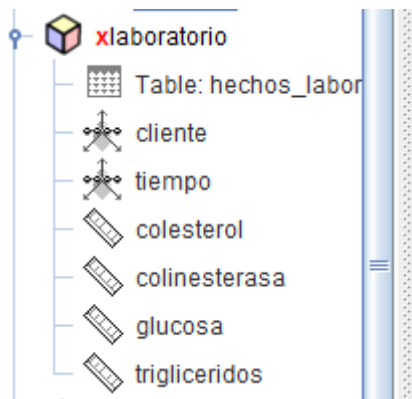
Dimension Usage for 'Admision' Cube	
Attribute	Value
name	cliente
foreignKey	dim_cliente_id_dim_cliente
source	cliente
level	
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

De la misma manera para todas las dimensiones y al final se le agrega una regla o medida

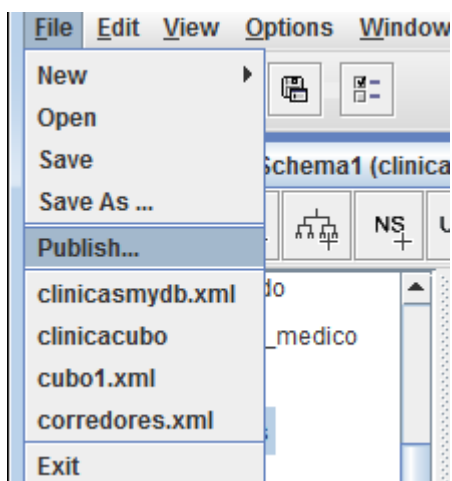




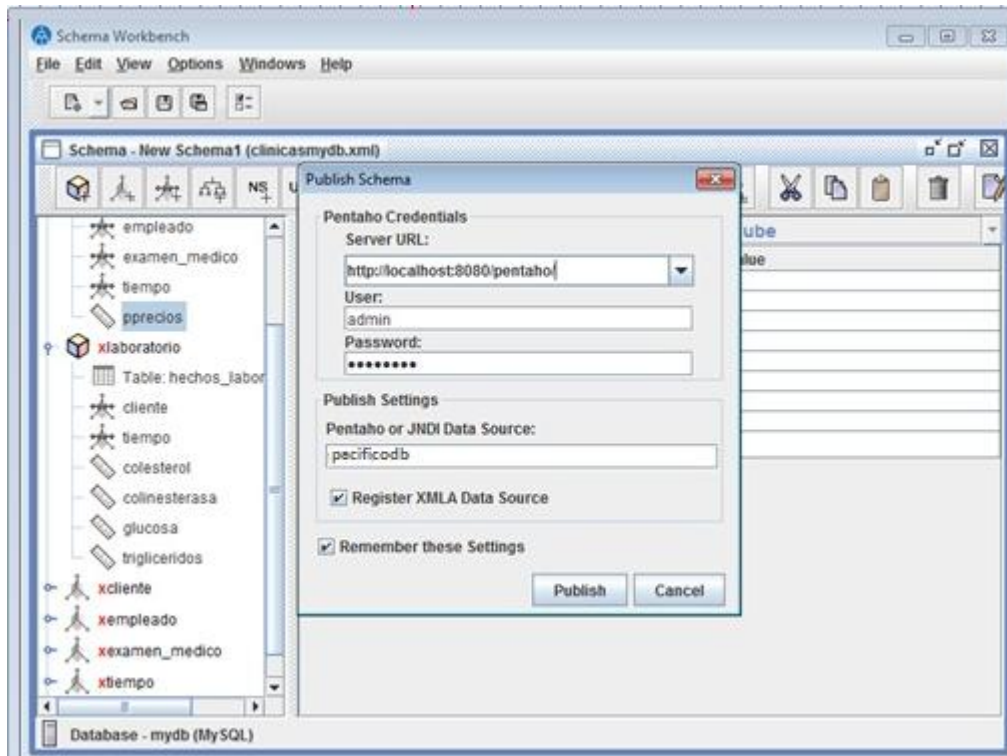
De la misma manera sucede con el cubo laboratorio



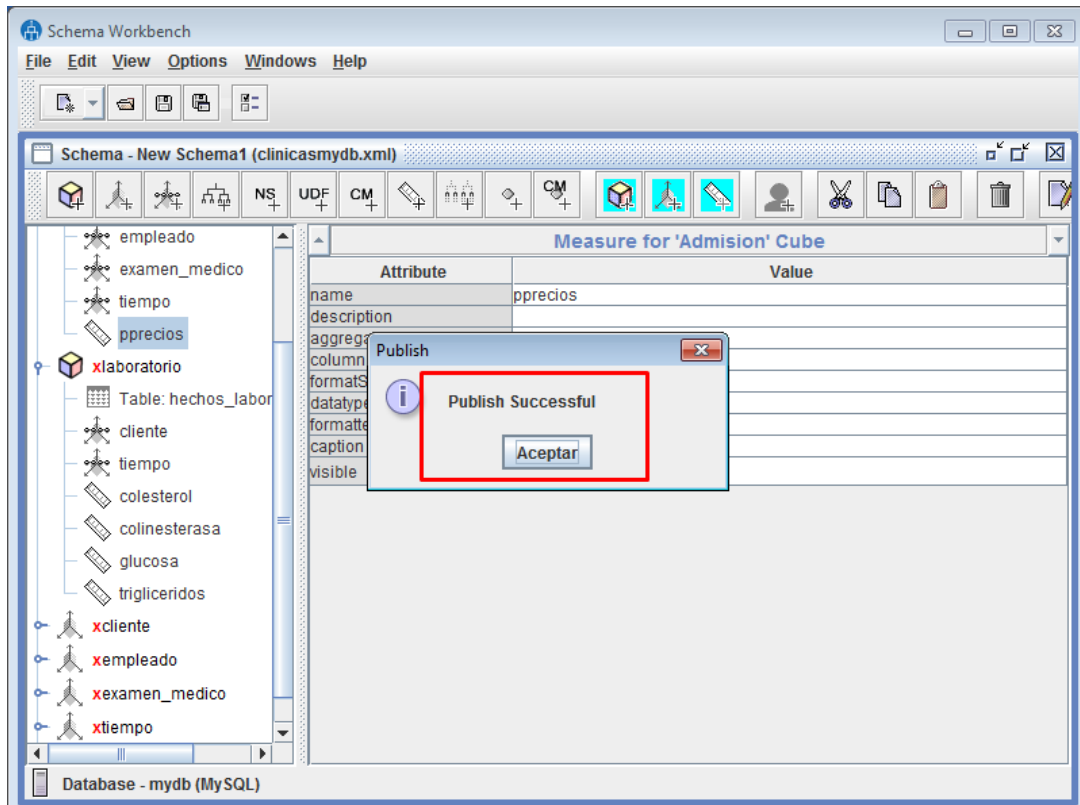
Luego de haber echo los cubos procedemos a publicar dichos cubos



Indicamos las credenciales necesarias

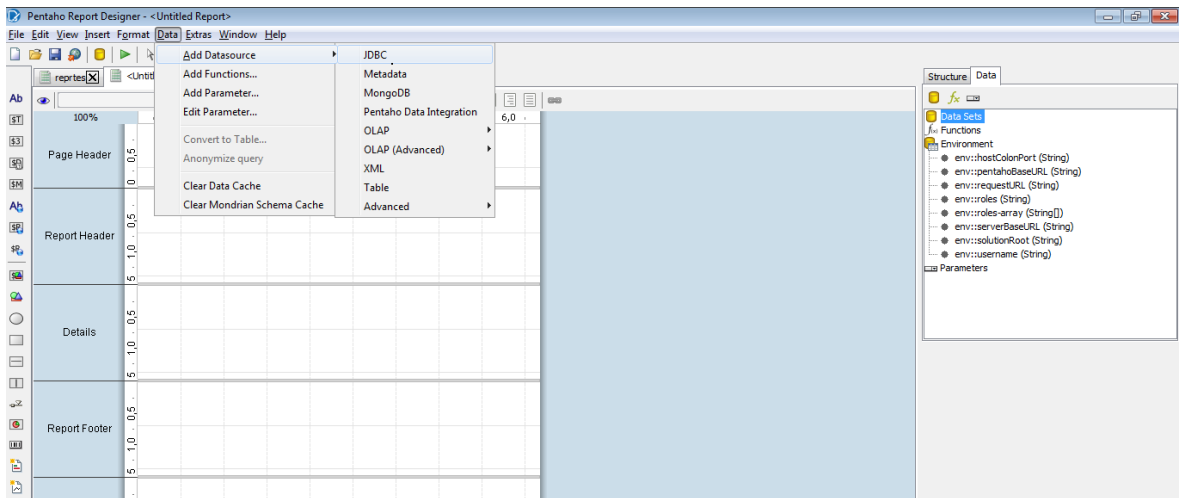


Finalmente muestra que la publicación fue realizada con éxito

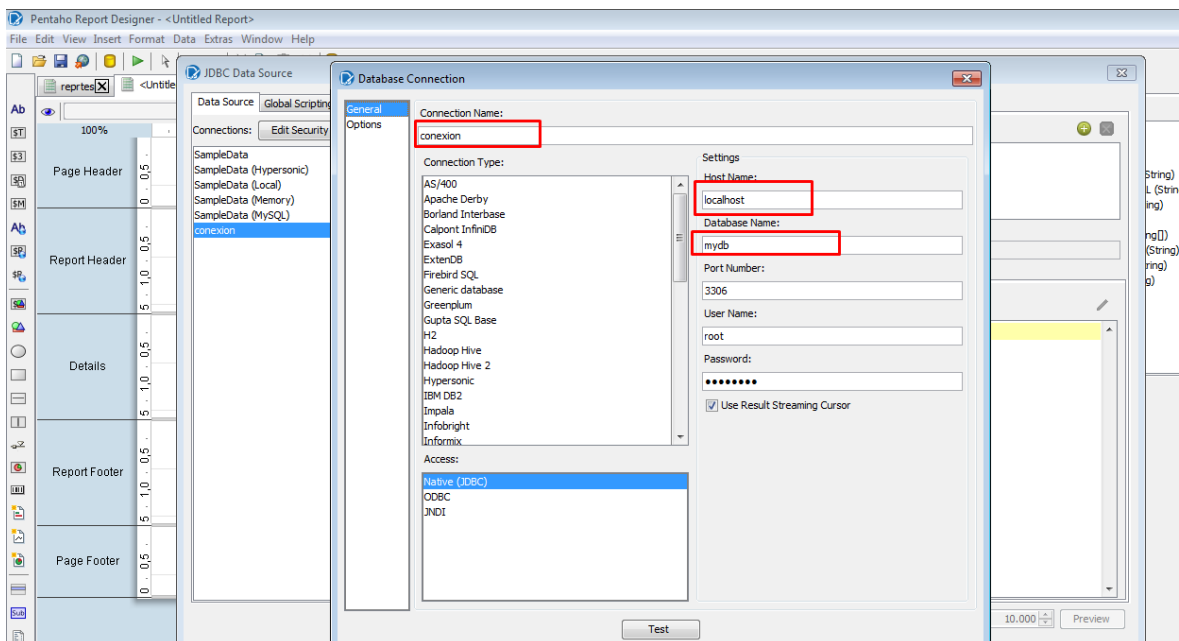


4.6. GENERACIÓN DE REPORTE

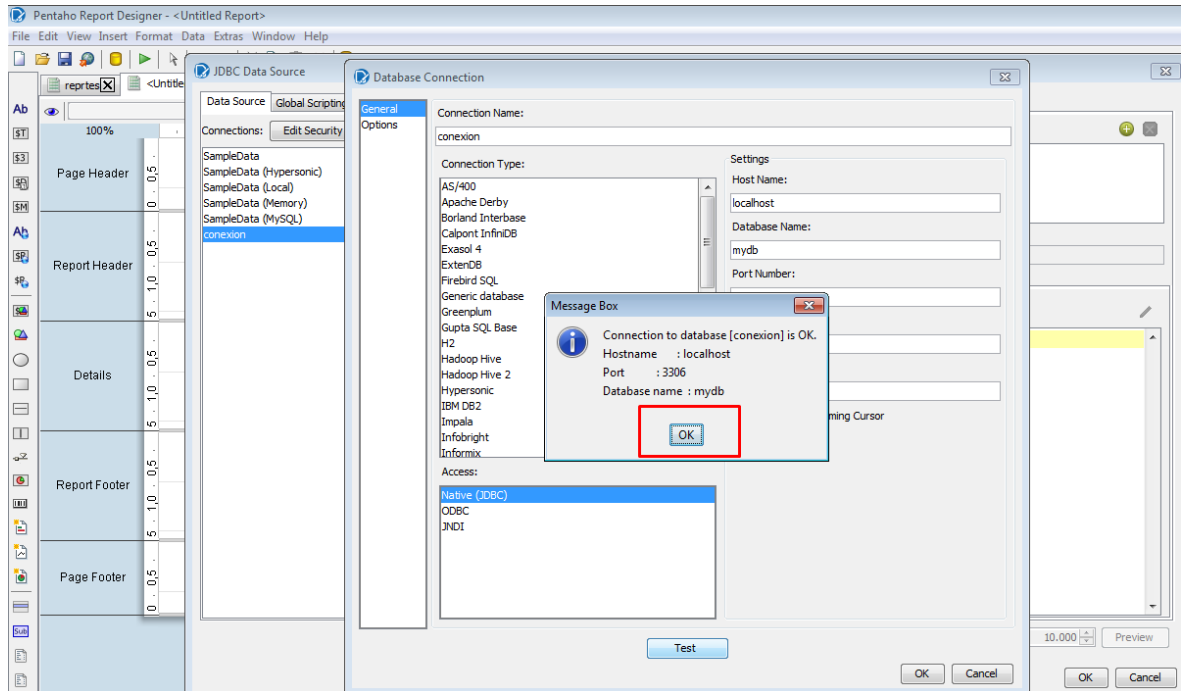
Para los reportes utilizaremos pentaho report designer



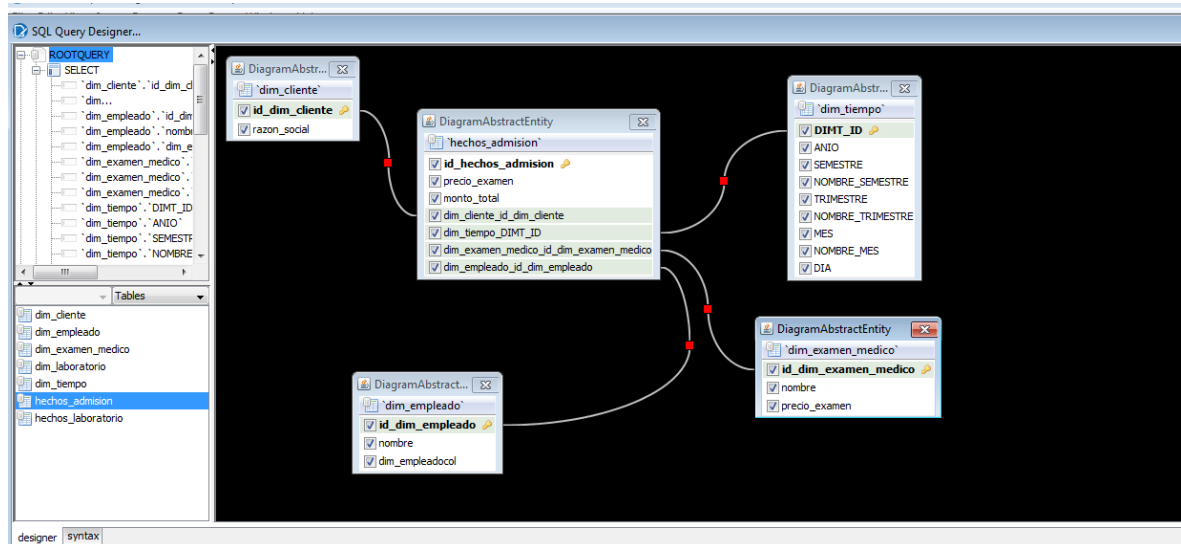
Nos aseguramos con establecer una conexión con los credenciales necesarios



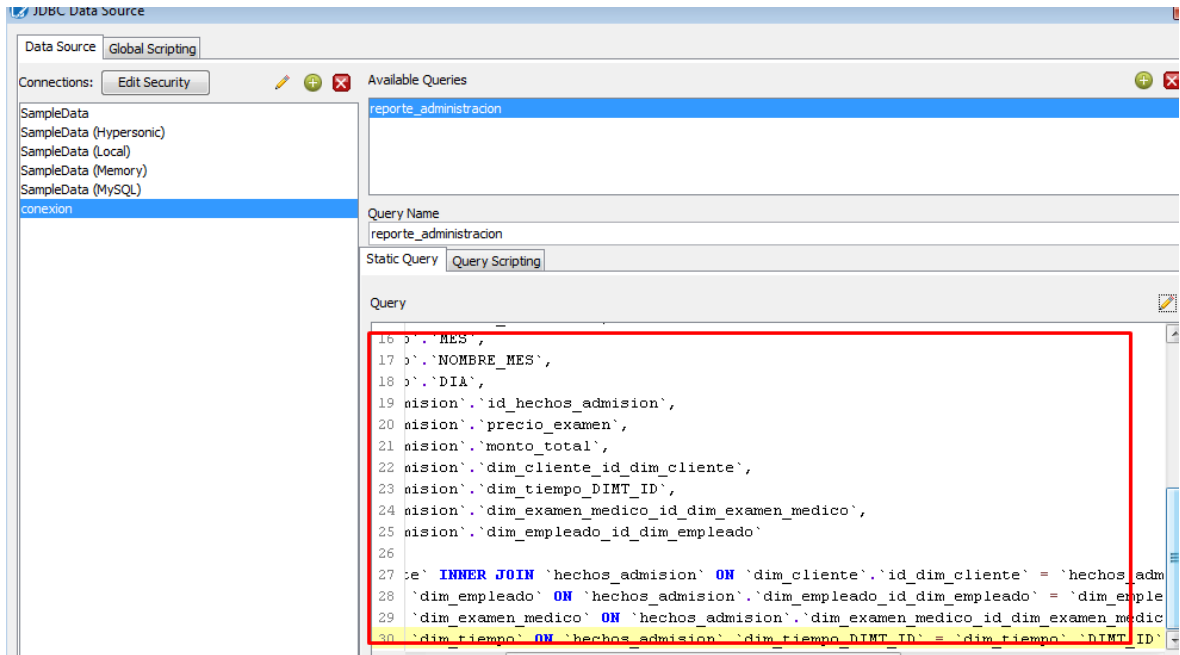
y probamos si dicha conexión está realmente bien



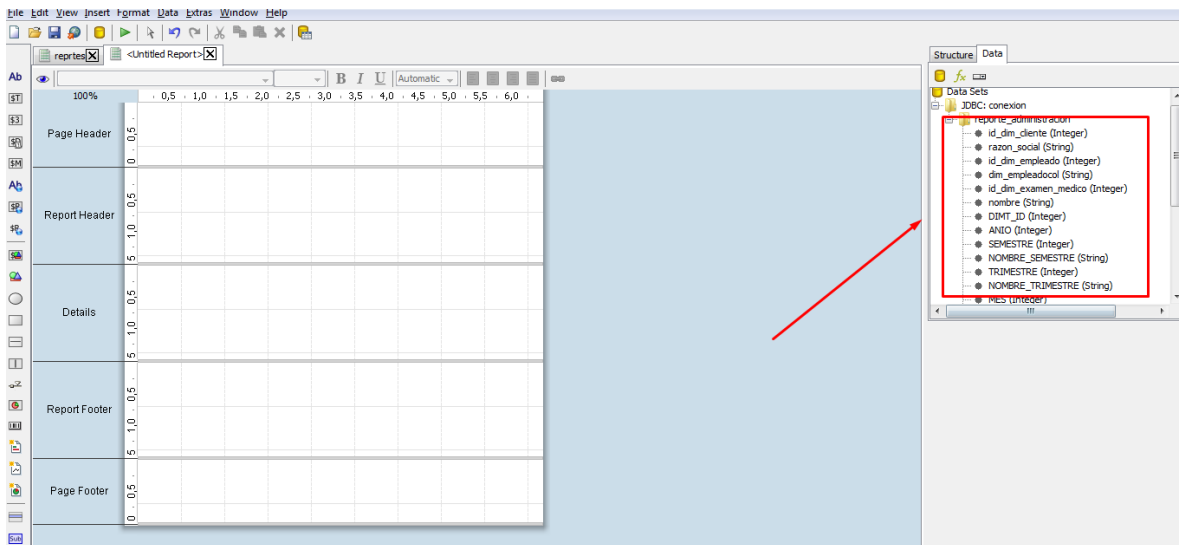
Escogemos las dimensiones a utilizarse



Del cual se genera la consulta



En la parte superior derecha muestra todos aquellos campos seleccionados




REPORTES DEL CUBO ADMISION

Resultado del reporte cantidad de pacientes atendidos que pertenecen a una empresa en un determinado mes: caso empresa Adecco Perú S.A para el mes de febrero.

empresa_cliente: AGUALIMA S.A.C.

mes_clinica: NOVIEMBRE

Auto-Update on selection



REPORTE GENERAL DE ADMISIÓN

Cantidad de pacientes atendidos en un determinado mes

Reporte del mes de: NOVIEMBRE


Empresa Cliente	Cantidad de pacientes	Monto total en el mes
AGUALIMA S.A.C.	20	3.000

↑
cantidad facturada para la empresa

Caso del mes de noviembre

empresa_deinte ADECCO PERU S.A.
mes_clinica NOVIEMBRE

Auto-Update on selection




REPORTE GENERAL DE ADMISIÓN

Cantidad de pacientes atendidos en un determinado mes

Reporte del mes de: NOVIEMBRE


Empresa Cliente	Cantidad de pacientes	Monto total en el mes
ADECCO PERU S.A.	5	750

Reporte total de pacientes atendidos en un mes

 100%

clinica_mess SETPTIEMBRE

Auto-Update on selection



REPORTE GENERAL DE ADMISIÓN

Cantidad de pacientes atendidos en un determinado mes

Reporte del mes de: SETPTIEMBRE


Cantidad de pacientes atendidos 1.874

Reporte para cubrir el requerimiento cantidad de pacientes atendidos por un médico en un determinado mes

medico MARIA ISABEL, PAZPEÑA

mes NOVIEMBRE

Auto-Update on selection



REPORTE GENERAL DE ADMISIÓN

Cantidad exámenes atendidos por un médico

Médico: MARIA ISABEL, PAZPEÑA

Mes: NOVIEMBRE


Cantidad de pacientes 334

Reporte de cantidad y precio por tipo de exámenes

cliente DANPER TRUJILLO S.A.C.

tipo_examen PRE OCUPACIONAL

Auto-Update on selection



REPORTE GENERAL DE ADMISIÓN

Cantidad exámenes portipo y empresa

Cliente: DANPER TRUJILLO S.A.C.

Tipo de examen: PRE OCUPACION ..

Cantidad de exámenes:	Monto total
7.202	1.080.300

Reporte de tipo de examen de una empresa en un mes determinado

empresa

tipo

mes

Auto-Update on selection



REPORTE GENERAL DE ADMISIÓN

Cantidad exámenes por tipo y empresa en un mes determinado

Cliente: DANPER TRUJILLO S.A.C.
 Tipo de examen: PRE OCUPACIONAL
 Mes: JULIO

Cantidad de exámenes:	Monto total
553	82.950


Reporte de cantidad de tipos de exámenes en laboratorio por empresa en un determinado mes

empresa

tipo

mes

Auto-Update on selection



REPORTE GENERAL DE ADMISIÓN

Cantidad exámenes por tipo y empresa en un mes determinado

Cliente: DANPER TRUJILLO S.A.C.
 Tipo de examen: PRE OCUPACIONAL
 Mes: SETPTIEMBRE

Cantidad de exámenes:	total
Colesterol	4
colinesterasa	4
triglicéridos	4
glucosa	4

4.7. DESPLIEGUE

4.7.1. IMPLEMENTACIÓN

Después de haber probado cada componente del proyecto, se comienza a instalar el motor de la base de datos y se programa un entrenamiento para todos los usuarios de la solución. En esta etapa es donde se da inicio a las funciones de soporte y mantenimiento de la base de datos, se programa y corre los ETL, se monitorea el comportamiento del sistema y también se afina la base de datos.

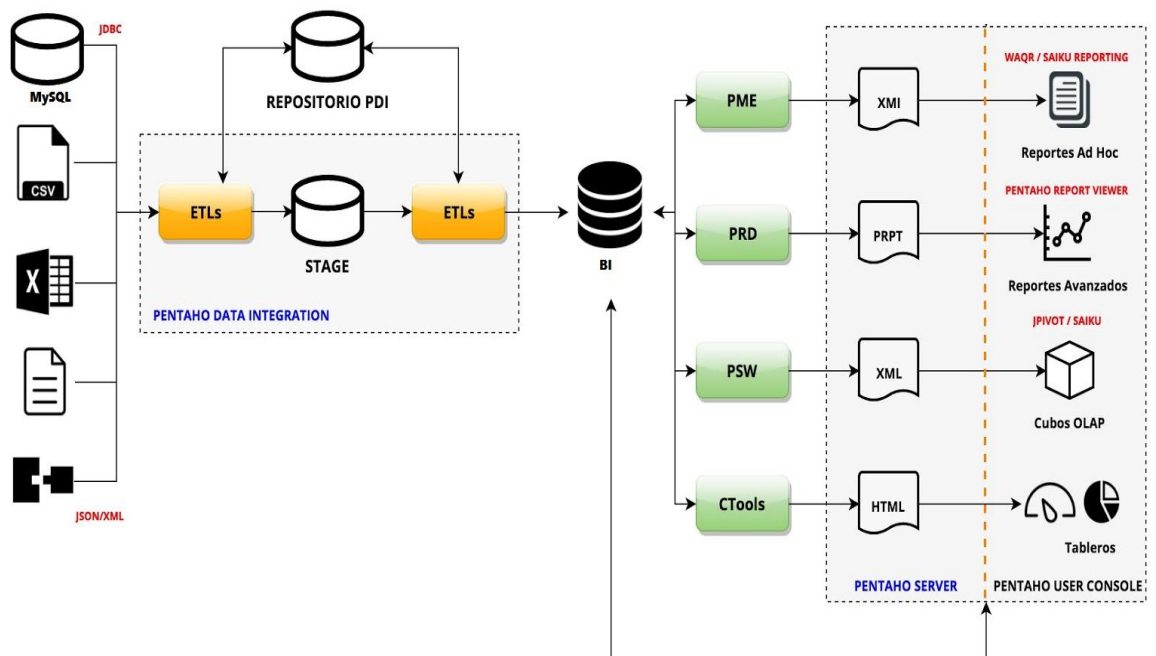


Figura 19: Diagrama de despliegue Detallado

5. DISCUSIÓN DE RESULTADOS

Para la contrastación de la hipótesis se ha considerado lo siguiente:

Formulación del Problema

¿Cómo mejorar el análisis de la información para el soporte en la toma de decisiones en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte usando los Sistemas de Información?

Hipótesis

El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss mejora en el análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo.

Luego se definen las variables que intervienen en la veracidad o falsedad de la hipótesis:

- ✓ Independiente (VI): Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss.
- ✓ Dependiente (VD): Análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo.

Población y muestra.

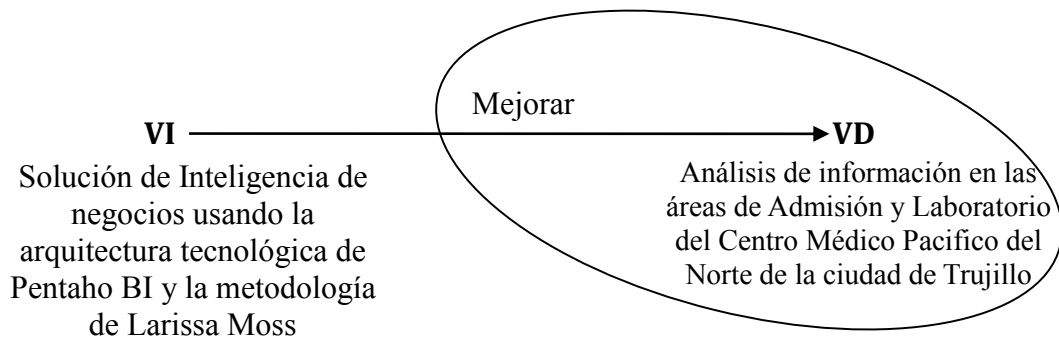
Población

Los tomadores de decisiones en las áreas de Admisión y Laboratorio, así como el director del Centro Médico Pacífico del Norte. (Director, Jefe de Admisión y de Laboratorios)

Muestra

Reportes utilizados por los tomadores de decisiones de las áreas de Admisión y Laboratorio.

5.2. MANERA PRESENCIAL



5.3. DISEÑO PREEXPIMENTAL PRE-PRUEBA Y POST-PRUEBA

PRE-PRUEBA (O₁): Es la medición previa de X a G

POST-PRUEBA (O₂): Corresponde a la nueva medición de X a G

Se determinó usar el Diseño PreExperimental Pre-Prueba y Post-Prueba, porque nuestra hipótesis se adecua a este diseño. Este diseño experimenta con un solo grupo de sujetos el cual es medido a través de un cuestionario antes y después de presentar el estímulo (BI). Este diseño se presenta de la siguiente manera:

G O₁ X O₂

Donde:

X: Tratamiento, estímulo (BI)

O: Medición a sujetos (Cuestionario)

G: Grupo de sujetos (Empleados)

El espacio de la muestra que se tomó para la medición de los indicadores de la hipótesis, correspondió al total de personas que operarán el BI, siendo estos 2; a estas personas se le aplicó un cuestionario, antes de interactuar con el BI (O₁) y después de interactuar con el mismo (O₂).

Al concluir la investigación se establecen las diferencias entre O₁ y O₂ para determinar si hay o no incremento en los resultados obtenidos.

5.3.1. CÁLCULO DE LOS INDICADORES DE LA HIPÓTESIS

Para el cálculo de los indicadores de la hipótesis en la solución de BI Propuesto (BIP) propuesto y el Sistema Actual (SA), se realizó un cuestionario (Ver Anexo B) donde se evaluó a los usuarios luego de haber interactuado con el BI. Los valores que los usuarios dieron a las respuestas del cuestionario fueron aplicados según el siguiente Rango de valoración:

RANGO	GRADO DE VALORACIÓN
1	Desacuerdo
2	Regular
3	Bueno
4	Muy Bueno
5	Excelente

Tabla 8: Rango de grado de valoración

ESCALA DE VALORACIÓN TOTAL	
Inadecuado	15-45
Adecuado	46- 75

5.3.2. APLICACIÓN DEL RANGO DE VALORACIÓN A LOS INDICADORES DE LA HIPÓTESIS

Los valores aplicados a los indicadores de la hipótesis tanto para el sistema Actual como para el BI propuesto se muestran en la siguiente tabla:

Evaluación de los indicadores de la hipótesis:

N°	INDICADORES	VALORACIÓN					\bar{X}
		1	2	3	4	5	
1	La carga de datos al BI es la suficiente o requerida				1	1	4.5
2	Los reportes obtenidos son claros en sus respuestas				1	1	4.5
3	Los reportes obtenidos son los requeridos u óptimo				1	1	4.5
4	La variedad de informes de evaluación proporcionados por el BI cubren con las expectativas de los stakeholders.				1	1	4.5
5	Con el BI es posible obtener reportes aplicando un número de cruces de requerimientos.					2	5
6	El grado de acceso al sistema es lo requerido				1	1	4.5
7	Se puede acceder al sistema desde varios tipos de dispositivos				1	1	4.5
8	La interface de reportes es intuitiva, amigable y sencilla				1	1	4.5
9	La implementación del BI es una decisión acertada.				1	1	4.5
10	La información presentada por el BI cumple con el criterio de exactitud				2		4
11	El personal se encuentra satisfecha con el cambio del modo anterior de proceso de toma de decisiones al actual con el BI					2	5
12	La información presentada apoya al proceso de toma de decisiones del área de admisión y laboratorio.					2	5
$\sum \bar{X}$							55

Donde: $X = (\text{Valor Valoración} * \text{Número de empleados respondieron en nivel valoración}) / 2$

Tabla 9: Evaluación de los indicadores de la hipótesis.

Interpretación: De acuerdo a la escala de valoración definida para esta ficha de observación, se determina que la Solución del BI es el **adecuado**, por ser sumatoria de los promedios 55 y superior a 45.

5.3.3. ANÁLISIS ESTADÍSTICO PARA LA PRUEBA PRESENCIAL DE LA HIPÓTESIS

Paso 1: Planteamiento de hipótesis.

$$H_0 : O_1 \geq O_2$$

$$H_1 : O_2 \geq O_1$$

Dónde:

H₀ es la hipótesis Nula: “El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss **no mejora** en el análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo.”

H₁ es la hipótesis Alternativa: “El Desarrollo de una Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss **mejora** en el análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacífico del Norte de la ciudad de Trujillo”

Paso 2: Nivel de significancia.

Para todo valor de probabilidad igual o menor que 0.05, se acepta H₁ y se rechaza H₀. $\alpha = 0,05$.

Paso 3: Prueba estadística.

Debido a que la muestra es $n = 2$, y por ende menor a 30, se aplicó la prueba estadística t-student, en esta prueba estadística se exige dependencia entre ambas, en las que hay dos momentos uno antes y otro después. Con ello se da a entender que, en el primer período, las

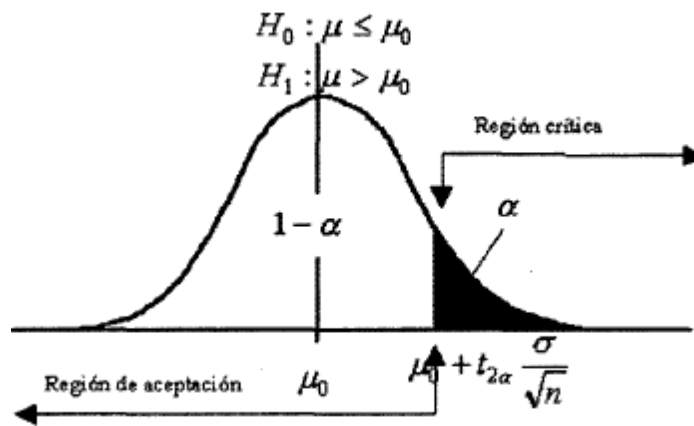
observaciones servirán de control o testigo, para conocer los cambios que se susciten después de aplicar una variable experimental.

Paso 4: Zona de rechazo.

Para todo valor de probabilidad mayor que 0.05, se acepta H_0 y se rechaza H_1 .

Si la $t_c > t_t$ se rechaza H_0 y se acepta H_1 .

Dónde: t_c es la t calculada y t_t es la t de tabla



Paso 5: Calculo de t_t y t_c

Calculo de la t de tabla t_t

$t_t (95\%, 2) = 2,92 \quad \rightarrow \text{Ver Anexo C.}$

Calculo de la t calculado t_c

$$\bar{D} = \frac{\sum D}{n}, \delta = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n - 1}}, t_c = \frac{\bar{D}}{\frac{\delta}{\sqrt{n}}}$$

Donde:

- t_c : T calculado.
- δ : Desviación estándar
- n : Tamaño de la muestra
- \bar{D} : Valor promedio o media aritmética de las diferencias entre los momentos antes y después.

Para el cálculo del valor de t calculado

Para el cálculo del valor T calculado se realizó un cuestionario (Ver Anexo C) donde se evaluó el grado de satisfacción a los usuarios luego de haber interactuado con el BI.

Los valores que los usuarios dieron a las respuestas del cuestionario fueron aplicados según el rango de satisfacción que muestran en la siguiente tabla:

RANGO	GRADO DE SATISFACCIÓN
0 – 2.5	Insatisfecho
2.5 – 5.0	Medianamente Satisfecho
5.0 – 7.5	Satisfecho
7.5 – 10.0	Muy Satisfecho

EVALUACIÓN DEL GRADO DE SATISFACCIÓN DE LOS USUARIOS:

Nº	INDICADORES	Media Pre U ₁	Media Post U ₂	D= (U ₂ - U ₁)	(Di - \bar{D})	(Di - \bar{D}) ²
1	¿Cuál es la cantidad total de exámenes médicos que se generan en un determinado mes?	2.0	9.0	7	0.7	0.49

2	¿Cuál es la cantidad de exámenes médicos por medico en un mes determinado?	2.0	9.0	7	0.7	0.49
3	¿Cuál es la cantidad de exámenes médicos generados por tipo y empresa en un mes determinado?	2.0	9.0	7	0.7	0.49
4	¿Cuál es la cantidad de exámenes médicos que se realizan por cada paciente en un determinado mes?	3.0	9.0	6	-0.3	0.09
5	¿Cuál es la cantidad total de consultas médicas realizadas por paciente en un determinado mes?	2.0	7.0	5	-1.3	1.69
6	¿Cuál es la cantidad media de pacientes particulares por día?	2.0	9.0	7	0.7	0.49
7	¿Cuál es la utilidad por mes de un tipo de examen médico en un mes determinado?	2.0	9.0	7	0.7	0.49
8	¿Cuáles son los exámenes médicos que menos se realizan en un mes determinado?	2.0	8.0	6	-0.3	0.09
9	¿Cuál es la cantidad de pacientes que tuvieron un examen médico en un mes determinado?	4.0	9.0	5	-1.3	1.69
10	¿Cuál es la cantidad de pacientes particulares que pasaron examen de laboratorio en un mes determinado?	3.0	9.0	6	-0.3	0.09

$N = 10$; $\sum D = 63$; $\bar{D} = 6.3$; $\sum (D_i - \bar{D})^2 = 6.1$; $\delta = 0.82$; $\sqrt{n} = 3.16$

$$t_c = \frac{\bar{D}}{\frac{\delta}{\sqrt{n}}}$$

$$t_c = 24.23$$

Interpretación: Como $t_c > t_t$, se acepta la hipótesis alternativa, entendiéndose que una “Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss **si mejora** el análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacifico del Norte de la ciudad de Trujillo”.

5.4. CUADRO DE LA COMPARACIÓN DE TIEMPO DE DEMORA EN LA EJECUCIÓN DE LAS CONSULTAS.

NRO	CONSULTAS	SISTEMA OLTP	SOLUCIÓN DE BI
1	¿Cuál es la cantidad total de exámenes médicos que se generan en un determinado mes?	Se debe de procesar 01 horas = 3 600 seg	02 seg.
2	¿Cuál es la cantidad de exámenes médicos por medico en un mes determinado?	Se debe de procesar 0.5 horas = 1 800 seg	03 seg.
3	¿Cuál es la cantidad de exámenes médicos generados por tipo y empresa en un mes determinado?	Se debe de procesar 01 horas = 3 600 seg	03 seg.
4	¿Cuál es la cantidad de exámenes médicos que se realizan por cada paciente en un determinado mes?	Se debe de procesar 01 horas = 3 600 seg	03 seg.
5	¿Cuál es la cantidad total de consultas médicas realizadas por paciente en un determinado mes?	Se debe de procesar 01 horas = 3 600 seg	02 seg
6	¿Cuál es la cantidad media de pacientes particulares por día?	Se debe de procesar 01 horas = 3 600 seg	02 seg
7	¿Cuál es la utilidad por mes de un tipo de examen médico en un mes determinado?	Se debe de procesar 01 horas = 3 600 seg	02 seg
8	¿Cuáles son los exámenes médicos que menos se realizan en un mes determinado?	Se debe de procesar 0.5 horas = 1 800 seg	03 seg
9	¿Cuál es la cantidad de pacientes que tuvieron un examen médico en un mes	Se debe de procesar 02 horas = 7 200 seg	03 seg

	determinado?		
10	¿Cuál es la cantidad de pacientes particulares que pasaron examen de laboratorio en un mes determinado?	Se debe de procesar 02 horas = 7 200 seg	03 seg
		Fuente: La empresa	Fuente: Solución BI

Resumen:

NRO CONSULTAS	SISTEMA OLTP	SOLUCIÓN DE BI
10	39600 segundos	26 segundos
<p>Interpretación: una “Solución de Inteligencia de negocios usando la arquitectura tecnológica de Pentaho BI y la metodología de Larissa Moss si mejora el análisis de información en las áreas de Admisión y Laboratorio del Centro Médico Pacifico del Norte de la ciudad de Trujillo” respecto al tiempo de obtener resultados a las consultas de 39600 segundos a 26 segundos.</p>		

6. CONCLUSIONES

- Se determinó que el alcance del proyecto está en el Área de Admisión y Laboratorio de la institución, basándose en las entrevistas hechas a los stakeholders se determinó que esta es el área crítica para el desarrollo de la empresa, recolectando 10 requerimientos principales que el jefe de Admisión y Laboratorio necesitan como soporte a la toma de decisiones.
- Se realizó el análisis de cada requerimiento a través del análisis de Molde de Consultas y se logró identificar la jerarquía de los datos en cada consulta hecha por el usuario determinando el esquema estrella que ayudará al análisis de datos. Al realizar el diseño se identificaron 02 tablas hechos y 5 dimensiones.
- La construcción del BI se realizó utilizando las herramientas de MySQL Server, y como herramienta ETL Pentaho Data Integration. Para la construcción de los cubos se utilizó Pentaho Cube Designer, implementando el modelo estrella para ventas y distribución.
- Se implementaron 10 reportes basados en los requerimientos del usuario a fin de garantizar el correcto funcionamiento del BI, estas implementaciones se realizaron utilizando Pentaho Report Designer.
- Se mejoró el tiempo en obtener los resultados a las consultas o requerimientos de los tomadores de decisiones de las áreas de 39600 segundos a 26 segundos.

7. RECOMENDACIONES

- Dirigir los proyectos de TI hacia los lineamientos estratégicos de la institución y realizar una evaluación periódica de los procesos de Tecnologías de Información de la empresa.
- Realizar un Planeamiento Estratégico de Tecnologías de Información respecto a nuevas tecnologías en el análisis de datos y llevar el proyecto de BI a la Nube.
- Mantener una relación estrecha con los usuarios involucrados en el área de admisión y laboratorio porque ellos siempre nos tendrán al día de sus necesidades y nos brindarán mayor información y ayudarán a identificar los puntos necesarios para el desarrollo de proyectos futuros.
- Para proyectos futuros en analítica de datos se recomienda el uso de metodologías ágiles y la integración con más servicios que ofrece la nube de Microsoft Azure como herramientas de trabajo para mejorar las versiones del proyecto de BI, creando los informes dinámicos en herramientas como Power BI o Qlikview.

8. REFERENCIAS BIBLIOGRÁFICAS

- Alcami, R. I., Carañana, C. D., & Herrando, J. G. (2009). Introducción a la gestión de sistemas de información en la empresa. Universitat Jaume.
- Arturo L, C. C. (2001). GUÍA PARA OBTENER EL RETORNO A LA INVERSIÓN EN PROYECTOS DE DATAWAREHOUSE .
- Cano, J. L. (2007). BUSINESS INTELLIGENCE: competir con información. Madrid: Fundación Banesto.
- Curto, J. (2007). Introducción al Business Intelligence . España: UOC.
- Flores, R. (2004). Aplicación de minería de datos en ambiente universitario.
- Gravitar. (8 de 5 de 2017). <http://gravitar.biz/pentaho/>. Obtenido de <http://gravitar.biz/pentaho/>
- Ian H. Witten, E. F. (2000). Data Mining Practical Machine Learning Tools and Techniques. San Francisco: Academic Press.
- IBM. (29 de 04 de 2017). IBM Knowledge Center. Obtenido de https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.0/com.ibm.datatools dimensional.ui.doc/topics/c_dm_star_schemas.html
- Moss, L., & Atre, S. (2003). Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Boston: Personal Education.
- Pentaho. (8 de 5 de 2017). pentaho. Obtenido de <http://www.pentaho.com/>
- Rio, L. M. (2006). Más allá del Business Intelligence. Barcelona: Gestión 2000.
- Roussel, G. (2006). Decision Support Systems Serving the Company : The Secrets to a Successful Project.
- Sanchez, L. Z. (2008). Metodología para el Diseño. Valencia.

Sinnexus. (20 de 04 de 2014). sinnexus. Obtenido de http://www.sinnexus.com/business_intelligence/datawarehouse.aspx

Stackowiak, R., Rayman, J., & Rick, R. (2007). Oracle Data Warehousing and Business Intelligence. wiley.

Wang, j. (2009). Encyclopedia of Data Warehousing and Mining, Second Edition.

ANEXOS

ANEXO A

CUESTIONARIO PARA ENTREVISTA N° 1: Director de Clínica

1. LAS RESPONSABILIDADES

- Describe su área y su relación con el resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

2. LOS OBJETIVOS COMERCIALES Y PROBLEMAS

- ¿Cuáles son los objetivos de su área?
- ¿Qué usted está tratando de lograr con estos objetivos?
- ¿Cuáles de estos objetivos son su prioridad para alcanzar sus metas dentro de su organización?
- ¿Cuáles son sus factores críticos de éxito?
- ¿Cómo usted sabe que usted está haciendo bien?
- ¿Qué tan menudo usted mide los factores de éxito importantes?
- ¿De los departamentos que funcionan? ¿cuáles son cruciales para asegurar que los factores de éxito importantes se logren?
- ¿Qué roles cumplen estos departamentos?
- ¿Cómo ellos trabajan para asegurar el éxito junto?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy dentro de su función? Y ¿Cuál es el impacto en la organización?
- ¿Cómo usted identifica sus problemas en su Área o sabe que usted se dirige hacia el problema?

3. ANÁLISIS DE LOS REQUISITOS

- En el análisis de los datos ¿Qué papel juega las decisiones que usted y otros gerentes toman en la ejecución del negocio?

- ¿Qué información importante se exige a hacer o a apoyar las decisiones que usted hace en el proceso de lograr sus metas y superar los obstáculos? ¿Cómo usted consigue esta información hoy?
- ¿Está allí otra información que no está disponible a usted hoy y que usted cree tendría el impacto significativo en ayudar a encontrar sus metas?
- ¿Están allí los cuellos de botella específicos a llegar a la información?
- ¿Qué informes usted usa actualmente?
- ¿Qué datos en el informe son importantes?
- ¿Cómo usted usa la información?
- ¿Qué problemas encuentra en estos informes?
- ¿Cuánta información histórica se requiere?
- ¿Si el informe fuera dinámico, en que lo haría diferentemente?
- ¿Qué capacidades analíticas le gustaría tener?
- ¿Qué oportunidades existen para mejorar dramáticamente su organización basándose en el acceso mejorado de la información?
- ¿Cuál es el impacto financiero usted piensa que tendría?

CUESTIONARIO PARA ENTREVISTA N° 2: Jefe de admisión y laboratorio

1. LAS RESPONSABILIDADES

- Describe su organización y su relación con el resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

2. LOS OBJETIVOS COMERCIALES Y PROBLEMAS

- ¿Cuáles son los objetivos de en el desempeño de su función?
- ¿Qué usted está tratando lograr con estos objetivos?
- ¿Cuáles de estos objetivos son sus prioridades para alcanzar sus metas dentro de su función?
- ¿Cuáles son sus factores críticos de éxito?
- ¿Qué tan menudo usted mide los factores de éxito importantes?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy?
- ¿Qué le impide cubrir sus objetivos?
- ¿Cuál es el impacto de estos problemas en la organización?

- ¿Cómo usted identifica sus problemas en su organización o sabe que usted se dirige hacia el problema?
- ¿Cómo es la relación que tiene con el Gerente?

3. ANÁLISIS LOS REQUISITOS

- ¿Qué tipo de análisis rutinario usted realiza actualmente? ¿Qué datos se usa?
- ¿Cómo usted consigue los datos actualmente?
- ¿Qué usted hace una vez con la información que usted obtiene?
- ¿Qué informes usted usa actualmente?
- ¿Qué datos en el informe es importante?
- ¿Cómo usted usa la información?
- ¿Qué oportunidades existen para mejorar dramáticamente su negocio basándose en el acceso mejorado de la información?

ANEXO B

CUESTIONARIO DIRIGIDO: Jefe de admisión – Laboratorio y Director

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
¿Cuál es la cantidad total de exámenes médicos que se generan en un determinado mes?											
¿Cuál es la cantidad de exámenes médicos por medico en un mes determinado?											
¿Cuál es la cantidad de exámenes médicos generados por tipo y empresa en un mes determinado?											
¿Cuál es la cantidad de exámenes médicos que se realizan por cada paciente en un determinado mes?											
¿Cuál es la cantidad total de consultas médicas realizadas por paciente en un determinado mes?											
¿Cuál es la cantidad media de pacientes particulares por día?											
¿Cuál es la utilidad por mes de un tipo de examen médico en un mes determinado?											
¿Cuáles son los exámenes médicos que menos se realizan en un mes determinado?											
¿Cuál es la cantidad de pacientes que tuvieron un examen médico en un mes determinado?											
¿Cuál es la cantidad de pacientes particulares que pasaron examen de laboratorio en un mes determinado?											

Tabla B1. Jefe de Admisión - Laboratorio y Director

ANEXO C

Tabla t-Student



Grados de libertad	0.25	0.1	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.8165	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.7407	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.6800	1.3007	1.6794	2.0141	2.4121	2.6896
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6800