

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

DISEÑO E IMPLEMENTACIÓN DE ARQUITECTURA DE DATOS GERENCIAL,
PARA VISUALIZAR LA PRODUCCIÓN DE LA CARTERA DE SERVICIOS
OFERTADA POR EL HOSPITAL IESS DE QUEVEDO.

PROYECTO DE TITULACIÓN

Previo la obtención del Título de:

Magister en Ciencias de Datos

Presentado por:

Kelly Jazmín Márquez Morán

GUAYAQUIL - ECUADOR

Año: 2022

DEDICATORIA

El presente trabajo de investigación lo dedico a Dios, por darme fuerza para culminar esta meta tan anhelada.

A mis padres, mis hijos quienes con su amor, comprensión, paciencia y esfuerzo me permitieron llegar a cumplir hoy un sueño tan anhelado, gracias por inculcar en mí el ejemplo de esfuerzo y valentía, de no temer a las adversidades y retos presentados y poniendo a Dios siempre por delante.

Finalmente, a todas las personas que compartieron sus conocimientos, nos apoyaron e hicieron que el trabajo se realice con éxito.

.

AGRADECIMIENTOS

Al culminar nuestros estudios de Maestría en Ciencia de Datos quiero agradecer a Dios por todas las cosas buenas que han ocurrido a lo largo de mi paso por la maestría.

También quiero agradecer a mi familia, quienes han sido el pilar fundamental en esta etapa de mi vida, y que, gracias a ellos, a mi esfuerzo y dedicación se ha logrado cumplir esta meta.

También agradezco a la Escuela Superior Politécnica a sus docentes y tutora de maestría Dra. Vanessa Cedeño, por la dedicación, tiempo y por compartir sus conocimientos e ideas ganadas a lo largo de su vida profesional, por todo el apoyo que ha brindado para que este trabajo concluya de la mejor manera.

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, me corresponde conforme al reglamento de propiedad intelectual de la institución; Kelly Jazmín Márquez Morán y doy mi consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”

Kelly Márquez Morán
Autora

COMITÉ EVALUADOR

.....
Vanessa Cedeño Mieles, P.h.D.

PROFESOR TUTOR

.....
Nombre del Profesor

PROFESOR EVALUADOR

RESUMEN

La Arquitectura de Datos identifica los procesos, sistemas y operativa necesaria para identificar, acceder, interoperar y gestionar los datos de gobierno. El objetivo del trabajo es diseñar una arquitectura de datos gerencial para su implementación con el fin de una adecuada visualización de la producción de la cartera de servicios ofertada por el hospital del IESS de Quevedo. Utilizando una minería de datos CRISP-DM, la investigación es predominante cuantitativa y de tipo no experimental. Se aplica una planificación, desarrollo y control del proceso; se debe crear un lenguaje, una base de datos y creación de informes. Concluyendo que el software de visualización planteado puede mejorar la forma en la que se examinan los datos del Hospital, permitiendo que el encargado de realizar estos reportes pueda tener un control total de los datos obtenidos, mismos que serán presentados con tablas y gráficos estadísticos lo cual permitirá la obtención de datos reales, claros y organizados.

Palabras Clave: arquitectura de datos, hospital, minería de datos.

ABSTRACT

The Data Architecture identifies the processes, systems and operations necessary to identify, access, interoperate and manage government data. The objective of the work is to design a management data architecture for its implementation in order to adequately visualize the production of the service portfolio offered by the IESS de Quevedo hospital. Using a CRISP-DM data mining, the research is predominantly quantitative and non-experimental. Planning, development and control of the process is applied; a language, a database and reporting must be created. Concluding that the proposed visualization software can improve the way in which the Hospital data is examined, allowing the person in charge of making these reports to have full control of the data obtained, which will be presented with statistical tables and graphs, which It will allow obtaining real, clear and organized data.

Keywords: data architecture, hospital, data mining.

ÍNDICE GENERAL

COMITÉ EVALUADOR.....	5
RESUMEN.....	2
<i>ABSTRACT</i>	3
ÍNDICE GENERAL.....	4
ÍNDICE DE FIGURAS.....	6
ÍNDICE DE TABLAS.....	¡Error! Marcador no definido.
INTRODUCCIÓN.....	8
CAPÍTULO 1: PLANTEAMIENTO DE LA PROBLEMÁTICA.....	9
1.1 Descripción del problema.....	9
1.2 Justificación.....	10
1.3 Solución propuesta.....	11
1.4 Objetivos.....	11
1.4.1 Objetivo General.....	11
1.4.2 Objetivos Específicos.....	11
1.5 Metodología.....	12
1.5.1 Enfoque de la investigación.....	12
1.5.2 Alcance de la investigación.....	12
1.5.3 Diseño de la investigación.....	12
1.5.4 Métodos y técnicas.....	12
1.5.5 Estrategia metodológica operativa.....	13
1.5.5.1 Fase 1. Planificación.....	13
1.5.5.2 Fase 2. Desarrollo de la solución.....	13
1.5.5.3 Fase 3. Control.....	14
CAPÍTULO 2: ESTADO DEL ARTE.....	15
2.1 Fundamentos del Problema.....	15
2.1.1 Arquitectura de datos.....	15
2.1.2 Gestión de la información.....	16

2.2 Fuente de datos relacionados al problema	17
2.2.1 MIS AS-400	17
2.3 Soluciones de analítica y aprendizaje relacionadas al problema	19
2.3.1 Sistemas de información	19
2.3.2 Extracción del conocimiento	22
2.3.3 Fases del proceso de extracción del conocimiento	22
2.3.3.1 Recolección de datos	22
2.3.3.2 Pre procesamiento y transformación de datos	23
2.3.3.3 Análisis de datos	23
2.3.4 Sistemas de soporte a la decisión (DSS)	24
2.4 Librerías y software a utilizar	26
2.5 Trabajos Relacionados	26
CAPÍTULO 3: DISEÑO E IMPLEMENTACIÓN.....	28
3.1 Recopilación	28
3.2 Transformación.....	31
3.3 Distribución.....	38
3.4 Consumo	42
CAPÍTULO 4: ANÁLISIS DE RESULTADOS.....	46
4.1 Puesta en marcha y funcionamiento.....	46
4.2 Análisis de datos.....	49
CONCLUSIONES	52
RECOMENDACIONES	53
BIBLIOGRAFÍA.....	54

ÍNDICE DE FIGURAS

Figura 1. Matriz producción mensual consulta externa	10
Figura 2. Minería CRISP-DM	13
Figura 3. Módulos AS-400	18
Figura 4. Descargas AS-400.....	18
Figura 5. Flujo de Integración AS-400	19
Figura 6. Modelo general de un sistema de información	20
Figura 7. Xampp	28
Figura 8. Localhost	29
Figura 9. Phpmyadmin.....	29
Figura 10. Colocar nombre	30
Figura 11. SQL Server 21	31
Figura 12. Navicat.....	32
Figura 13. Creación de conexión	32
Figura 14. BD analisismgst.....	33
Figura 15. Borrado DSA.....	34
Figura 16. Carga del archivo excel	34
Figura 17. Carga de base de datos.....	35
Figura 18. Abrir base de datos creada.....	36
Figura 19. Estructura base de datos	37
Figura 20. Codificación	38
Figura 21. MySQL.....	39
Figura 22. Power BI	40
Figura 23. Obtención datos de otro origen.....	40
Figura 24. Base de datos.....	41
Figura 25. Datos del servidor	41
Figura 26. Credenciales.....	42
Figura 27. Proceso ETL	43
Figura 28. Base datos creada.....	44
Figura 29. Análisis	45
Figura 30. Carga de información.....	46
Figura 31. Tablas	47

Figura 32. Informe.....	47
Figura 33. Informes generados	48
Figura 34. Informe consulta externa	49
Figura 35. Informe hospitalización	50
Figura 36. Informe emergencia.....	51
Figura 37. Informe facturación	51

INTRODUCCIÓN

Hoy en día, la tecnología es tan avanzada que la automatización y el control de procesos se han vuelto imprescindibles en muchos campos, incluido el médico, para permitir una gestión más eficiente de los diagnósticos y las historias clínicas de las personas (Gastesi, 2019).

Actualmente, la demanda de servicios de salud en el hogar está aumentando, lo que resulta en una mayor cantidad de información de los pacientes que buscan atención (Salvatierra & Vasquez, 2022), esta información se ha convertido en un recurso necesario y muy importante para la gestión ya que es una herramienta clave para apoyar a los gerentes en la toma de decisiones siempre y cuando esta información sea ágil y oportuna (Cadmen, 2019).

El Instituto Ecuatoriano de Seguridad Social (IEES) maneja gran cantidad de información acerca de la atención médica que prestan a sus afiliados, es por ello que utilizan un sistema gerencial (MIS) AS400 (servidor multiusuario) con el objetivo de gestionar toda la información de pacientes, profesionales, bodega y parte financiera. Los datos ingresados al sistema pueden ser descargados dentro de la misma aplicación (IESS, 2019), el gran problema existe al momento de querer obtener datos esenciales para el análisis y toma de decisiones, actualmente los datos que nos arroja el sistema (.xls) contienen matrices de información confusas y difíciles de entender a simple vista, la idea es poder contar con información clara y efectiva permitiendo el ahorro de tiempo y recursos (Larrea, 2016).

Por lo antes mencionado es de gran importancia para el establecimiento IEES contar con un software que permita el control de información adecuado, para la toma de decisiones más eficaces por parte de gerencia, para conseguir la integración y el aprovechamiento de todos los recursos con los que cuenta esta casa de salud.

CAPÍTULO 1: PLANTEAMIENTO DE LA PROBLEMÁTICA

1.1 Descripción del problema

Desde el año 2017, el Hospital General de Quevedo, servicio médico de 2do nivel del Instituto Ecuatoriano de Seguridad Social (IESS), abre sus puertas para atender a las personas, a los asegurados de la Provincia Médica Los y al público integral. la Red de Higiene. de Ríos. El hospital proporciona seguro de salud a más de 350.000 personas aseguradas. Los servicios incluyen medicina general, medicina familiar, medicina preventiva, medicina interna, medicina general, pediatría y cirugía vascular. Hemos completado una cartera de 24 especialidades, entre pediatría y ginecología, atendiendo una media de 650 pacientes diarios (IESS, 2020).

El hospital trabaja con el sistema AS400 que es un servidor multiusuario, de gama media alta, creado por IBM para todo tipo de empresas y usuarios departamentales, el sistema tiene su propia interfaz, controlada por menús, que utiliza terminales e incluye su propio sistema operativo OS/400, integra herramientas propias para garantizar su funcionamiento: compiladores, comunicaciones, bases de datos, hardware, software, seguridad. Este sistema se emplea a nivel nacional en todas las dependencias del IESS (IESS, 2019) .

Debido a la cantidad de pacientes atendidos diariamente y los diferentes servicios que se brindan, en esta casa hospitalaria se generan grandes volúmenes de datos: prescripciones, historiales de pacientes, pruebas diagnósticas, miles de datos generados cada día y que forman una parte inevitable de las estadísticas necesarias para el análisis de la gestión hospitalaria (Salud, 2018).

Si bien es cierto que el sistema AS400 es utilizado en el hospital, no existe una herramienta que presente indicadores a nivel gerencial, donde se pueda reflejar la carga laboral de los profesionales médicos y los montos facturados por cada servicio (Larrea, 2016). Toda la información requerida sobre la gestión hospitalaria se encuentra desagregada en diferentes archivos, generados por el sistema AS400 a partir de los datos que maneja. Para acceder a esos archivos se realiza una descarga del sistema y se obtienen archivos en formato (.xls). Como se puede observar en la Figura 1.1, este es un breve ejemplo de los archivos que son presentados al momento que la gerencia

pide un reporte actualizado y consolidado de la información, mismos que son extraídos a partir de los archivos descargados del sistema AS400.

I. E. S. S. DIRECCION DEL SEGURO GENERAL DE SALUD INDIVIDUAL Y FAMILIAR														Población Afiliada Estimada		181,637	
PRODUCCION DE CONSULTA EXTERNA														Promedios Días Laborados		20	
UNIDAD MÉDICA		HOSPITAL GENERAL QUEVEDO															
AÑO	2021	MES	MAYO														
CONSULTA EXTERNA	TOTAL CONSULTA POR PROGRAMA			ESPECIALIDADES	PRIMERAS CONSULTAS EN EL AÑO	SEG. SALUD	RIESGO TRABAJO	VOLUNTAR.	CONYUGE	MENOR 18 AÑOS	JUBILD.	S.S.C.	MONTEPIO	NO AFILIADO	TOTAL	HORAS LABORADAS	
PRIMERA CONSULTA EN AÑO	5,043			ALERGOLOGIA											-		
ACTIVOS				ALTO RIESGO OBSTETRICO											-		
SEG. DE SALUD IND.FAM.	4,993			ANESTESIOLOGIA											-		
SEG. RIESGOS DEL TRABAJO	-			CALIFICACION MEDICA											-		
VOLUNTARIOS	478			CARDIOLOGIA	88	82		28	8	5	76	40	5	-	323	88	
ST(1)	5,471			CIRUGIA CARDIOTORAC.											-		
CONYUGE	577			CIRUGIA GENERAL	1	1		-				2		-	3	3	
HUOS <6 AÑOS	1,545			CIRUGIA MAXILO FACIAL											-		
ST(2)	2,122			CIRUGIA ONCOLOGICA											-		
JUBILADO	1,051			CIRUGIA PEDIATRICA	16				44					1	45	9	
S.S.C.	717			CIRUGIA PLASTICA											-		
MONTEPIO	167			CIRUGIA VASCULAR	103	80		9	8	1	9	16	7	-	140	41	
ST(3)	1,935			DERMATOLOGIA	372	282		27	27	193	45	55	7	-	616	223	
NO AFILIADO	13			ENDOCRINOLOGIA											-		
TOTAL(ST(1)+ST(2)+ST(3)+C22)	9,541			GASTROENTEROLOGIA	232	188		21	23	7	53	55	5	-	332	96	
SEXO				GERIATRIA	138	40		21	8		79	65	8	-	221	61	
MASCULINO	4,557			GINECOLOGIA	467	405		16	95	40	11	69	5	2	645	22	
FEMENINO	4,984			HEMATOLOGIA											-		
				HEMODINAMIA											-		
				INFECTOLOGIA											-		
				RECETAS PRESCRITAS											-		

Figura 1. Matriz producción mensual consulta externa

Fuente: Tomado de Informe de producción del Hospital General Quevedo- Provincia de Los Ríos, Ecuador (2021)

Esta es la información que presentan a los directivos, ellos no la pueden interpretar de manera fácil o intuitiva, y requiere un proceso adicional para generar nuevos reportes. Para obtener este análisis de toda la unidad médica se requiere que al menos 6 áreas realicen individualmente su informe de resultados a partir de las matrices. De no implementar una arquitectura de datos el Hospital se quedará sin información real y, el sólo revisar el wholesale, no necesariamente refleja lo que está pasando en el mismo, igualmente no tendría información para poder medir la verdadera gestión de cada área, dicha información se refiere a los importes, además de controlar la situación financiera de cada uno de ellos.

1.2 Justificación

Actualmente el Hospital General del IESS de Quevedo existen numerosas fuentes de datos que contienen una gran cantidad de información valiosa relacionada con los pacientes y los servicios que se ofrecen. Pero actualmente el proceso de la información

es tan engorroso que se dificulta su análisis e interpretación. La idea es poder brindar presentar la información generada de forma fácil y oportuna, con ello se tendrá una visión integral de los datos del hospital. También permite la mejora continua al dar a los hospitales control sobre todos los datos que procesan. Los procesos de gestión de la información correctamente implementados alientan a las instituciones a incluir análisis de indicadores específicos para apoyar la toma de decisiones basadas en datos reales y bien procesados. El proceso de investigación demuestra la importancia de la arquitectura de datos en una organización, por lo que se considera factible su aplicación.

1.3 Solución propuesta

Se propone el desarrollo de un software que utilice técnicas de Big data y provea de inteligencia y significado a los datos, para hacerlos relevantes y de fácil interpretación. Implica que se pueda analizar todo ese volumen de información y que los directivos tendrán acceso a una herramienta de uso práctico y sencillo, lo que supone una ventaja competitiva para una toma de decisiones enfocada a ofrecer un mejor servicio de salud, hacer uso óptimo de los insumos y planificar estratégicamente los recursos del hospital.

La aplicación también proporciona efectividad y eficiencia en el manejo del flujo de información, facilita el procesamiento de los grandes volúmenes de datos y permite a los directivos ejecutar las actividades relacionadas a este proceso con menos esfuerzo y tiempo. Además, esta propuesta es novedosa debido a que sería la primera herramienta de este tipo que tendría el hospital para obtener información sobre su gestión de una manera actualizada, rápida y confiable.

1.4 Objetivos

1.4.1 Objetivo General

Diseñar una arquitectura de datos gerencial para su implementación con el fin de una adecuada visualización de la producción de la cartera de servicios ofertada por el hospital del IESS de Quevedo.

1.4.2 Objetivos Específicos

- Reconocer el panorama actual de los datos y herramientas tecnológicas que brindan soluciones a la información originada por el Hospital IESS de Quevedo.

- Analizar la producción de datos que ingresan a: emergencia, consulta externa y egresos del hospital IEES de Quevedo durante los años 2017-2021. Con el fin de una utilidad y su practicidad a la hora de tomar decisiones.
- Diseñar un software de arquitectura de datos para la carga, procesamiento y transformación de los archivos descargados del sistema AS-400 durante los periodos comprendidos 2017-2021.
- Implementar una interfaz (power bi) de usuario que permita la visualización de los datos generados por el Hospital IEES durante el periodo 2017 hasta el 2021 a través de tablas y gráficos estadísticos. Facilitando así el acceso a información tanto detallada como de su universo.

1.5 Metodología

1.5.1 Enfoque de la investigación

El enfoque de esta investigación es predominantemente cuantitativo, al analizarse datos numéricos, estadísticos y financieros (Hernández, Fernández, & Baptista, 2016). Esto que permite tener una visión clara sobre los requerimientos del sector para la aplicación de la propuesta planteada en el estudio.

1.5.2 Alcance de la investigación

El alcance de esta investigación es descriptivo, es decir, se trata de explicar las características del objeto de estudio en sus características principales, a fin de proponer una solución específica a la situación negativa detectada (Hernández, Fernández, & Baptista, 2016).

1.5.3 Diseño de la investigación

El diseño de esta investigación es no experimental, por cuanto no se modifica intencionalmente ninguna variable, sino que se examina la situación de las mismas, evaluando sus características. Además, la investigación es transversal, o sea, se aplican los instrumentos de medición en un momento único (Hernández, Fernández, & Baptista, 2016).

1.5.4 Métodos y técnicas

Los métodos empleados en esta investigación son:

Análisis. Permite examinar las características principales del objeto de estudio, señalando los detalles positivos y negativos (Tamayo y Tamayo, 2017). Esto permite plantear medidas específicas para mejorar la situación del sector.

1.5.5 Estrategia metodológica operativa

La metodología a utilizarse para la minería de datos es la CRISP-DM

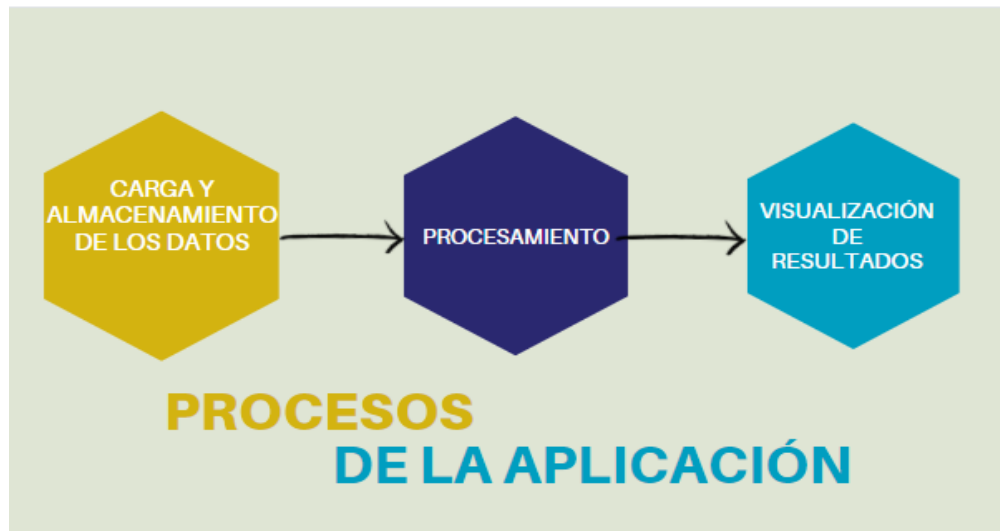


Figura 2. Minería CRISP-DM

Fuente: (Burriss, 2004)

1.5.5.1 Fase 1. Planificación

En esta fase se detalla todo lo necesario para la implementación del proyecto desde los objetivos hasta los alcances; todo esto sale con base en la información que se tiene de la entidad y lo que espera cubrir en este caso la exposición de la producción de la cartera de servicios ofertada por el Hospital IESS de Quevedo e indicadores que ayuden en la toma de decisiones.

1.5.5.2 Fase 2. Desarrollo de la solución

En esta fase están incluidos los formatos de archivos como referencia para ser incluidos en una base de datos DSA, con ella se construye un esquema ETL transaccional, este tipo de esquema relaciona la información para extraerla, transformarla y cargarla. Finalmente, convertir la información en útil por medio de un modelo de cubo de información este tipo de estructura rompe los esquemas de la ETL y hace un análisis rápido de los datos permitiendo al usuario un acceso categórico y fácil.

1.5.5.3 Fase 3. Control

La técnica que complementa a la primera es la del diseño de KPIs, esta consiste en identificar los Stake Holders, definir metas y objetivos para en los posterior dar seguimiento a la consecución de resultados. Con esta base de interpretación se logró obtener datos reales dentro de la situación de la institución, para definir indicadores que permitan un análisis minucioso y apoyen en la toma de decisiones para el aumento de la rentabilidad.

CAPÍTULO 2: ESTADO DEL ARTE

2.1 Fundamentos del Problema

2.1.1 Arquitectura de datos

Es necesario iniciar con la inteligencia de negocios, esta se define como la combinación de tecnologías, herramientas y procesos que le permiten convertir los datos almacenados en información, esa información en conocimiento y ese conocimiento en un plan o estrategia. Debe ser parte de su estrategia comercial para que pueda optimizar la utilización de los recursos, monitorear el cumplimiento de los objetivos comerciales y tomar las decisiones correctas para lograr mejores resultados (Conesa y Curto, 2010).

Dentro de cualquier organización o empresa la información es uno de los pilares más importantes al momento de buscar mejoras y mayores proyecciones, representa básicamente la idea central de lo que es y lo que busca está creando caminos válidos para tomar acción si alguna eventualidad ocurre, al gestionarla de manera adecuada se convierte en una pieza estratégica a la hora de tomar decisiones.

Así, se da pie a la Inteligencia de Negocios (BI: Business Intelligence), como miembro de una estrategia de gestión empresarial, su dinámica gira en torno a la tecnología como herramienta de tratamiento de información (Rosado y Rico, 2010). De manera muy general se advierte que el BI incluye una variedad de metodologías, especializadas en la creación, análisis y gestión de la información con el propósito de identificar los principales indicadores, mismos que deben tener la característica de medibles para que el progreso de una empresa sea evidenciable, es decir, son alertas que ayudan a prevenir, esclarecer e intervenir al momento de existir un problema.

Por lo tanto, al hablar de BI de manera paralela o complementaria aparecen disciplinas para análisis de datos, una de las más relevante es la DM o Minería de datos, esta, se define como la fase del proceso de descubrimiento de información en una metodología que agrupa procesos no comunes de identificación de patrones válidos, novedosos, potencialmente útiles en bases de datos (KDD: Knowledge Discovery in Databases) (Rosado y Rico, 2010). Para contextualizar el panorama de mejor manera se debe tomar en cuenta a la gestión de la información, en el siguiente apartado se desarrolla este tema.

2.1.2 Gestión de la información

A continuación, habla de la gestión de la información como referencia para la investigación. Hace veinte años, la gestión se asumía en campos relacionados con la información. Hoy en día, todas las organizaciones tienen claro que la información es un recurso imprescindible si quieren ser competitivas y mejorar la calidad. También lo hace la satisfacción del cliente por operar en el mercado global. Como disciplina revolucionaria dentro de la sociedad de la información, el propio proceso de gestión de la información surgió en la década de 1980, replanteando el mercado laboral de los profesionales de la información y demandando nuevos tipos de profesionales con importantes responsabilidades en el diseño y desarrollo de la información, cuyo reconocimiento es bien conocido. sistema dentro de la organización.

Brown (2017), Lo describe como un estudio interdisciplinario de las propiedades y el comportamiento de la información, el poder que determina el flujo de información cuando se usa y los métodos manuales y mecánicos de procesamiento de información para almacenar, recuperar y distribuir información de manera más efectiva. ciencia científica De esta forma, se amplió el ámbito y área de aplicación, y se destacó cada vez más su importancia en el desarrollo sostenible de todo tipo de organizaciones, caracterizadas por el impacto del avance tecnológico en los modelos de procesamiento y difusión de la información. Su principal representante es la informática, seguida de las telecomunicaciones. Ambas referencias para acceder a la información desde cualquier parte del mundo, en menor tiempo ya costos más razonables, aumentando su efectividad.

Especialmente en la década de los 80, la gestión de la información se consolidó como una innovación en la práctica informática actual. Aunque es similar a otras disciplinas de procesamiento como la biblioteconomía, la documentación, el archivo y la biblioteconomía, no debe confundirse ni considerarse ciencia de la información. La gestión de los recursos de información proporciona mecanismos que ayudan a las organizaciones a adquirir, generar y transmitir datos e información con calidad y precisión al menor costo posible y en el tiempo suficiente para lograr las metas establecidas.

Con lo dicho, se evidencia que estrictamente la información se la define como recurso crucial para la consecución de los objetivos de cualquier organización, por tal motivo se establece que debe ser abordado con criterio y táctica, surgiendo así, los sistemas y

redes de información como estructuras básicas para procesar y aplicar la información, con el tiempo se transformaron en los componentes puntales de la gestión de la información.

En este aspecto, cabe mencionar términos como el de información, este de manera global se define como “el mensaje que genera un emisor destinado a un receptor en un sistema de comunicación en un soporte perdurable” (York, 1994), puntualizando dentro del contexto de investigación se debe destacar dentro de la información su organización, almacenamiento, la difusión y recuperación de la información.

Antes de abordar el tema de los sistemas de información es importante diferenciar ciertas definiciones indispensables (Burriss, 2004):

- Datos, se trata de una medición objetiva.
- Información, conjunto de datos relacionados e interpretados.
- Conocimiento, conjunto de información desarrollada, que permite prever y planificar.

2.2 Fuente de datos relacionados al problema

2.2.1 MIS AS-400

El sistema AS-400 creado por IBM gestiona el sistema operativo en base a objetos y las llamadas bibliotecas. Este sistema se llama OS/400 y su base de datos es DBS/400. Esta base de datos admite datos para aplicaciones atribuidas a ALS. Un sistema completo y potente, la principal ventaja de AS-400 es la inclusión de control de tareas para lograr un mejor rendimiento del sistema (Gastesi, 2019). Este programa, utilizado actualmente por la administración de salud del IEES, proporciona información sobre los siguientes módulos procesados de forma independiente.



Figura 3. Módulos AS-400

Fuente: (IESS, 2019)

Es sistema permite realizar reportes de las actividades que realizan los profesionales estos reportes son generados en un archivo plano (.xlsx) como se muestra en la siguiente ilustración.

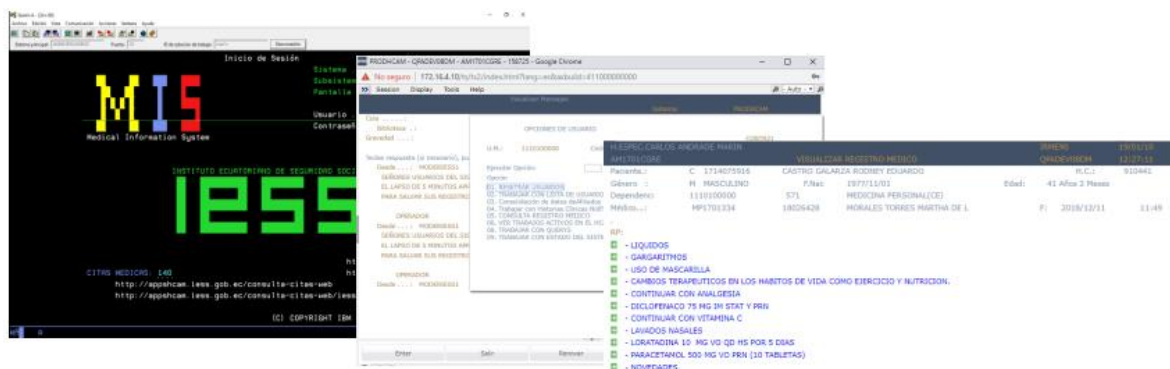


Figura 4. Descargas AS-400

Fuente: (IESS, 2019)



Figura 5. Flujo de Integración AS-400

2.3 Soluciones de analítica y aprendizaje relacionadas al problema

2.3.1 Sistemas de información

Técnicamente, un sistema de información se define como un conjunto de componentes interrelacionados que recopilan, procesan, almacenan y distribuyen información para apoyar la toma de decisiones y los procesos de control dentro de una organización. También puede ayudar a los gerentes y trabajadores del conocimiento a analizar problemas, visualizar problemas complejos y desarrollar nuevos productos (Laudon y Laudon, 2012).

Los sistemas de información contienen información relacionada con información sobre personas, lugares y cosas importantes dentro de una organización, o información como datos modelados de manera significativa y útil para los humanos. Más bien, los datos son un flujo de elementos sin procesar que representan eventos que ocurren en una organización o entorno físico antes de que se organicen e interpreten para la comprensión y el uso humano (Laudon y Laudon, 2012).

En las empresas, este sistema tiene como objetivo recopilar la información necesaria y proporcionarla a los miembros de la empresa para su uso en la toma de decisiones, la gestión estratégica o la implementación de las decisiones tomadas durante la transformación requerida (Baquerizo y Cevallos, 2017).

Los sistemas de información que complementan las definiciones establecidas son conjuntos representativos de ideas, enfoques, alcances, matices y diferentes perspectivas (ya veces incluso las más opuestas) juntas. Esto lleva a diseñar supuestos discriminantes para la implementación de sistemas de la actividad humana con el fin de enfrentar adecuadamente los inconvenientes que puedan presentarse en el mundo en general. Todo esto sugiere que la determinación de objetivos es un componente fundamental de los sistemas de información (Tramulas, 1997).

Hay muchos tipos diferentes de sistemas, pero la mayoría puede representarse mediante un modelo que consta de cinco bloques básicos: elementos de entrada, elementos de salida, secciones de transformación, mecanismos de control y objetivos, como se ilustra en la Figura 6.

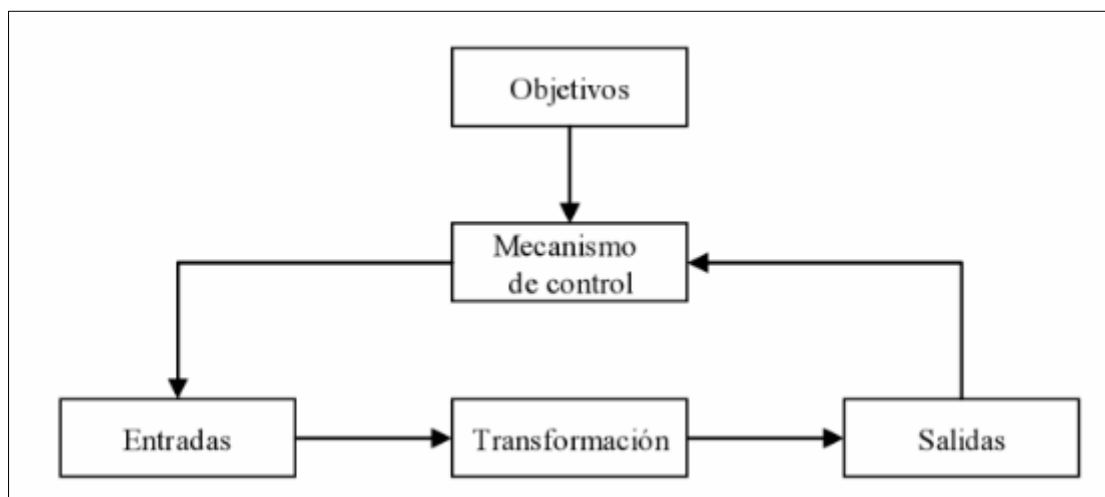


Figura 6. Modelo general de un sistema de información

Fuente: (Fernández, 2006)

La figura muestra los recursos que acceden al sistema, el concepto básico que gira en torno a esto es el proceso de transformación. En este sentido, todas las acciones a llevarse a cabo son el resultado de la interacción de actividades necesarias para transformar las entradas en salidas, lo que supone un grado de interconectividad entre sus componentes (entidades) para definir los recursos del sistema (Senn, 1992).

La transformación que se menciona debe tener a disposición ciertos mecanismos para que se asegure un control de lo realizado, aquí surgen las medidas de desempeño donde las más populares son la toma de decisiones como apoyo al control de mando. Por esta razón, el sistema primero utiliza información interna (de la misma empresa u organización) para coordinar las acciones realizadas dentro del entorno. Los sistemas

de apoyo a la decisión permiten la evaluación de estrategias para la introducción de nuevos productos o la evaluación de diferentes alternativas a lo largo del tiempo (Decisiones Estructuradas POCO).

Una de las funciones que deben realizar los sistemas de información es la colaboración en la toma de decisiones. Dependiendo de dónde se toman las decisiones en la pirámide jerárquica, los sistemas de información se pueden clasificar como estratégicos, de gestión u operativos. Los sistemas de apoyo a la decisión extraen datos de los sistemas de procesamiento de datos y de los sistemas de información de gestión, pero también utilizan fuentes externas a la organización que brindan información sobre competidores, clientes, mercados, proveedores, etc. (Fernández, 2006).

Los sistemas enfatizan una relación simbiótica entre los actores: usuarios, información y herramientas. Manejar la información juntos. Esto nos permite ver que la mayoría de los sistemas de información son mixtos en el sentido de que combinan procedimientos manuales y automatizados (Devlin, 1991). A partir de datos relacionados con las operaciones de la empresa, el sistema modifica las condiciones iniciales para simular resultados y, al variar los parámetros iniciales, los gerentes pueden simular resultados con base en eventos actuales y pasados en la organización y el entorno.

Estos sistemas resultan útiles en todos los niveles de una organización y áreas funcionales, desde la alta gerencia hasta los equipos de operaciones, estos están estrechamente ligados con la toma de decisiones en toda la pirámide organizacional, ya sean decisiones estructuradas, semiestructuradas o no estructuradas, los sistemas de información pueden mejorar la calidad de la toma de decisiones, por ejemplo al decidir quién es apto para otorgarle un crédito, para volver eficientes a los procesos repetitivos y muy estructurados de algún área, o para hacer estimaciones sobre los posibles efectos al realizar una inversión de riesgo.

Entonces, la información que se encuentra dentro de los sistemas sugiere una variedad de procesos sobre todo cuando se busca una gestión del conocimiento, para efectos de lo dicho en apartados siguientes se aborda el proceso de la extracción del conocimiento en bases de datos, esto abre paso a la minería de datos, llegando al DSS y finalmente a la inteligencia de negocio. Con base en lo expuesto, se explica a continuación el proceso de extracción del conocimiento que integra la minería de datos y todo lo que implica el análisis de información.

2.3.2 Extracción del conocimiento

Knowledge Discovery in Databases (KDD) es el proceso crítico de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles en los datos.

Todo lo que se extraiga en el proceso necesita contar con acreditación, para lo que se exponen ciertas propiedades para afianzar la premisa, así (Burch y Grudnitsky, 1992):

- Activo. El modelo o muestra encontrado debe tener suficiente precisión al usar nuevos datos.
- El conocimiento nuevo, extraído, tenía que ser conocido antes de que pudiera ser adquirido.
- Útil. El conocimiento adquirido debe permitirnos mejorar el sistema o tomar decisiones que aporten algún beneficio al usuario.
- Comprensible, pero por lo demás difícil de interpretar, verificar y aplicar.

Además, cuenta con fases por las que este conocimiento debe pasar para que sea útil, mismas que se detallan a continuación.

2.3.3 Fases del proceso de extracción del conocimiento

2.3.3.1 Recolección de datos

Hoy en día, hay innumerables tipos de información que se pueden encontrar en la vida diaria, como redes sociales, supermercados y negocios, y hay innumerables formas de obtenerlos, y esa información se puede obtener en forma de datos. Al mismo tiempo, estos archivos se pueden ver de varias maneras. Ejemplos: texto, imágenes, videos, archivos XML, JSON, etc. Esta información puede enviarse de diversas formas, incluidos correo electrónico, redes sociales, mensajería instantánea, aplicaciones móviles y otros medios.

Con el advenimiento de Internet y las computadoras, diariamente se generan infinitas cantidades de información sobre una variedad de temas, lo que complica la tarea de encontrar datos relevantes para ayudar a resolver problemas como la toma de decisiones. Se pueden aplicar varios criterios para determinar si la información que

encuentra es útil. Los datos deben ser confiables y precisos, deben provenir de fuentes confiables, deben ser precisos, relevantes y reflejar el uso previsto. Deben ser consistentes y deben permanecer consistentes. Finalmente, deben ser cíclicos, ya que el hecho de que se publiquen sobre una base cíclica específica agrega valor (Ramesh, 2010).

La recopilación de datos consiste en definir varias fuentes de datos que contienen datos sin procesar en varios formatos estructurados, semiestructurados y no estructurados, como: B.: Bases de datos relacionales, archivos de Excel, formato CSV plano, sistemas de informes de transacciones ERP, CRM, SCM, etc. Fuentes de datos externas como redes sociales, sistemas de información geográfica y sistemas gubernamentales. Datos en formato JSON o XML obtenidos de servicios web u otros sistemas (Mazon, 2018).

2.3.3.2 Pre procesamiento y transformación de datos

En esta fase, los datos se procesan, transforman y generan nuevas variables a partir de las que ya están disponibles. Todo esto es para estructurar los datos a medida que procedemos a utilizar técnicas de minería (Ruiz, 2018). Establece la preparación y limpieza de datos extraídos de diversas fuentes que son referenciados en fases posteriores. En este punto, se tienden a utilizar varias estrategias para manejar datos faltantes o vacíos fuera de los límites para que la estructura de datos sea finalmente adecuada para transformaciones posteriores (Ruiz, 2018).

2.3.3.3 Análisis de datos

El análisis de datos es la parte que examina los datos y saca conclusiones sobre la información que crea valor de alguna manera o enriquece las operaciones de la empresa. Las empresas suelen utilizar los análisis para tomar mejores decisiones comerciales basadas en ellos. También se utiliza en la ciencia para probar o refutar modelos y teorías existentes. El propósito principal de llevar a cabo procesos de análisis de datos generalmente es ayudar a las empresas a tomar mejores decisiones comerciales al permitir que los científicos y otros usuarios de datos analicen grandes volúmenes de datos transaccionales. Se puede utilizar para estadísticas, toma de decisiones, prevención de riesgos y extracción de negocios. inteligencia, etc.

Primero, como puede ver, el análisis de datos es similar a la extracción, con la diferencia de que al análisis se le da un propósito y un enfoque, es decir, el análisis se realiza de acuerdo con el propósito deseado. Por otro lado, el proceso típico de extracción no

distingue entre lo que es útil y lo que no, o cuál de esta información representa un valor particular. Este análisis se puede realizar utilizando herramientas de datos comunes en el marco de disciplinas analíticas avanzadas como el análisis predictivo y la minería de datos (García, 2018).

Las técnicas de análisis de datos que utilizan las empresas pueden ser muy diversas en función del tipo de datos con los que se trabaja, algunas de las técnicas más utilizadas:

El propósito de la visualización de datos, informes de visualización es extraer instantáneamente información relevante de los datos almacenados para brindar agilidad e información para contribuir a la toma de decisiones. Los informes son especialmente útiles para socializar informes históricos sobre el comportamiento de los indicadores, o para socializar información agregada de varias fuentes. Algunos informes se utilizan para analizar resultados pasados y predecir acciones para mejorar el desempeño de los indicadores (Reyes, 2015).

Data mining, minería de Datos es un término genérico que engloba resultados de investigación, técnicas y herramientas usadas para extraer información útil de grandes bases de datos (García y Molina, 2012).

Árbol de decisiones, las metodologías de aprendizaje supervisado incluyen el aprendizaje del árbol de decisiones. Las representaciones utilizadas para describir los conceptos adquiridos son árboles de decisión, que consisten en representaciones relativamente simples del conocimiento y cuyos procedimientos utilizados para el aprendizaje son más sencillos que los de los sistemas que utilizan lenguajes representacionales más potentes. Redes semánticas, representaciones en lógica de primer orden, etc.

2.3.4 Sistemas de soporte a la decisión (DSS)

Hay muchos tipos de sistemas de información, desde sistemas transaccionales que automatizan los procesos operativos y ahorran personal, hasta sistemas estratégicos y sistemas que brindan información para respaldar los procesos de toma de decisiones. Este último incluye sistemas para sistemas de información ejecutiva (EIS), sistemas de soporte de decisiones (DSS) y sistemas expertos. Estos se detallan específicamente para respaldar el estudio de antecedentes DSS.

Los sistemas de soporte de decisiones (DSS o Decision Support Systems) pertenecen a la clase superior de los sistemas de información computarizados. Los sistemas de soporte de decisiones se enfocan en respaldar la toma de decisiones en todas las etapas, pero si bien la toma de decisiones en sí sigue siendo solo para el usuario, los sistemas de soporte de decisiones son más adecuados para individuos o grupos de usuarios. A veces denominados sistemas centrados en la inteligencia empresarial (Kendall y Julie, 2011).

Los sistemas de apoyo a la decisión se utilizan para resolver problemas no estructurados (problemas que son impredecibles y la información necesaria para resolverlos no está disponible) o problemas semiestructurados (Fernández, 2006). Si desea resolver problemas no estructurados, su sistema de información debe ser altamente flexible (para adaptarse a cualquier tipo de situación) y tener una gran cantidad de herramientas analíticas que permitan una investigación analítica suficiente. Para lograr la flexibilidad necesaria para resolver estos problemas, los sistemas de soporte de decisiones deben ofrecer un alto grado de interactividad entre el usuario y el sistema.

Dada la complejidad de los procesos de procesamiento de la información y los diversos grados o niveles en los que se pueden estructurar los datos y los procesos, deben existir diferentes categorías de DSS que puedan cubrir toda la información que necesita una organización, según el problema que se presente (Lapiedra, Devece, y Guiral, 2011).

La toma de decisiones es el resultado de un proceso en el que se eligen entre diferentes alternativas. Un sistema de soporte de decisiones (DSS) es un sistema informático que apoya el proceso de toma de decisiones, generando datos y modelos, estimando y ayudando a los tomadores de decisiones a recopilar información, generar opciones y tomar decisiones Evaluación y/o comparación sistemática de alternativas.

DSS es una herramienta para tomar decisiones efectivas con diferentes enfoques y diferentes actividades. Se pueden utilizar para obtener información que indique los elementos más destacados del problema y las relaciones entre ellos, y para identificar, desarrollar y comunicar los cursos de acción disponibles y las alternativas de decisión (Ramesh, 2010).

2.4 Librerías y software a utilizar

Software, XAMPP es un paquete de software libre, que consiste principalmente en el sistema de gestión de bases de datos MySQL, el servidor web Apache y los intérpretes para lenguajes de script PHP y Perl. El nombre es en realidad un acrónimo: X, Apache, MariaDB/MySQL, PHP, Perl. Base de datos relacional, la misma se guarda en un servidor con una instalación de SQL Server 201.

2.5 Trabajos Relacionados

A continuación, se detalla varios trabajos relacionados al tema de investigación realizado al sector de la salud los cuales tratan de dar solución a problemas sobre gestión de información y toma de decisiones por parte de los directivos.

- a. Un estudiante de Ingeniería de Universidad Pontificia Católica del Perú realizó un trabajo de tesis titulado “Análisis, diseño e implementación de un DataWarehouse de soporte de decisiones para un hospital del sistema de salud público” (Data Warehouse permite procesar información mediante su sólida plataforma de datos históricos integrados para realizar análisis). Como resultado de este estudio tenemos la implementación de un sistema de soporte de decisiones PENTAHO, el cual permite a los usuarios interactuar con los datos en un nivel alto con el objetivo de ayudar a toma decisiones con base a la información obtenida.

El proyecto de tesis de Villanueva dio resultado impactando de manera favorable tanto a los pacientes como a los trabajadores del centro de salud. Mejorando la satisfacción de los pacientes y la productividad de los empleados y además de tener una base para explotar mejor la información del centro de salud. (Villanueva, 2008)

- b. Este trabajo fue realizado en Lima Perú titulado “Implementación de un sistema de información gerencial para la mejora en la toma de decisiones del área de recursos humanos del ministerio de salud aplicando Business Intelligence”, esta investigación tiene un enfoque cuantitativo no experimental, fue realizada a 25 expedientes de recursos humano en el área de salud, utilizo herramientas como SPSS (procesamiento de datos), este proyecto se realizó con la finalidad de contribuir al conocimiento efectivo sobre el uso de los sistemas de información, como herramienta de apoyo a la toma de decisiones y en los procesos de registro, consolidación y consulta el mismo que da como conclusión que el proceso de

información gerencial si mejorara en la toma de decisiones en el área de recursos humanos del Ministerio de Salud. (CACERES & LOPEZ, 2018)

- c. Dos estudiantes de la Universidad Tecnológica del Perú, realizaron una investigación la cual tiene como título “Implementación de un sistema de información gerencial para mejorar la toma de decisiones en el área logística de un centro de salud en Lima”, la metodología que se utiliza en este proyecto como marco es PMBOK para gestionar el proyecto y la metodología RUP para elaboración del software y como herramientas tecnológicas se utiliza base de datos MySQL Workbench, lenguaje de programación JAVA Enterprise Edition, y el entorno de desarrollo (IDE) IntelliJ Idea y Servidor Tomcat. Se concluyó con el proyecto afirmando que los sistemas de información gerencial son una pieza fundamental para los gerentes, puesto que es un gran soporte a la hora de tomar decisiones y más aún si se trata de la logística de los centros de salud, ya que es una parte fundamental para el buen servicio a los pacientes (Pinglo & Ramos, 2018)

CAPÍTULO 3: DISEÑO E IMPLEMENTACIÓN

3.1 Recopilación

Los datos a utilizar se obtienen del sistema AS400 utilizado por el IESS a nivel nacional estos son descargados y visualizados en hojas de cálculo. Estos archivos son generados y actualizados de forma diaria y se descargan de forma mensual para ser analizados. Cada archivo contiene información de todos los datos ingresados en el AS-400 durante el periodo desde el 2017 – 2021. Existe más de 40 variables tanto cualitativas como cuantitativas algunas de estas serán tomadas en cuenta para la investigación mientras que algunas serán desechadas durante la depuración y limpieza de los datos para luego hacer su análisis y transformación.

- En xampp se crea la BD, por medio de PHPMYADMIN
- Luego de Instalar XAMPP se corre la consola y se pone en *start mysql* y *apache*.

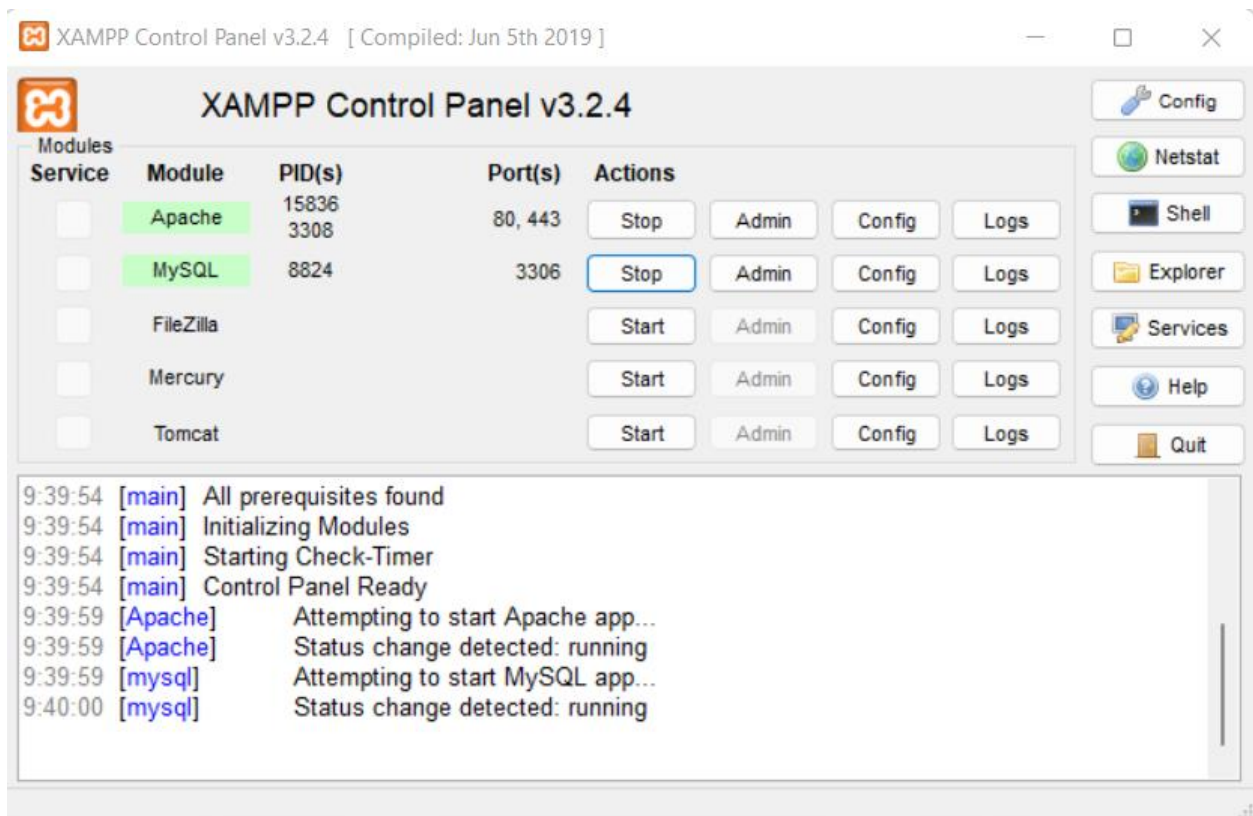


Figura 7. Xampp

- Luego se dirige a un explorador de internet y se coloca *localhost*



Figura 8. Localhost

- Luego se da un click en *phpmyadmin*

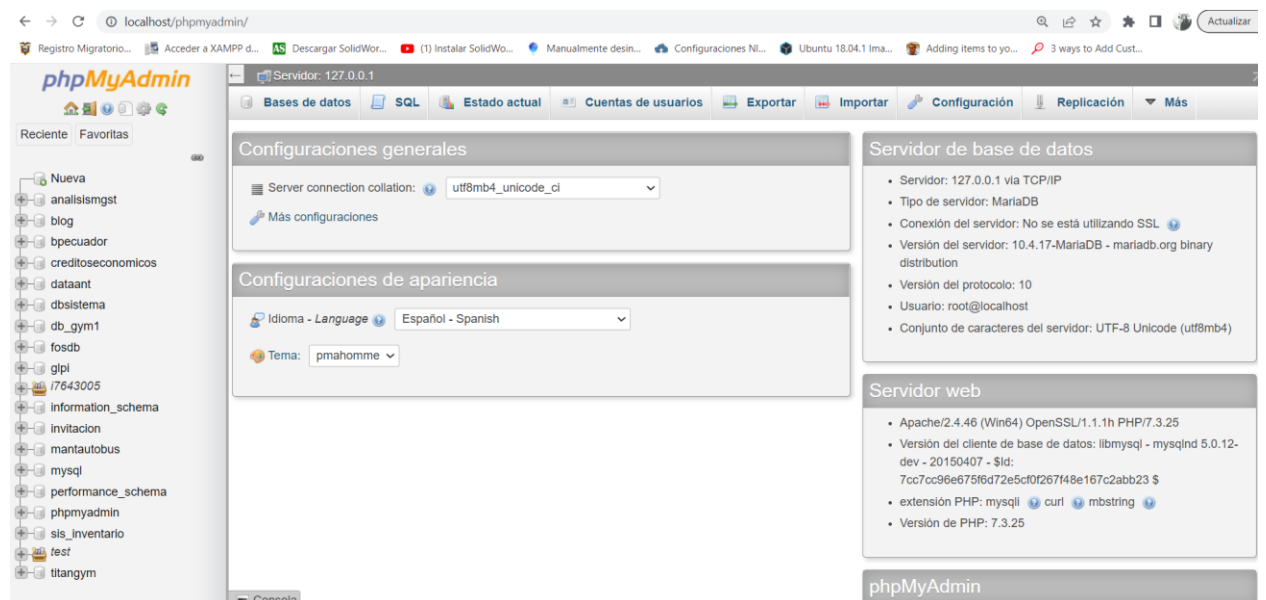


Figura 9. Phpmyadmin

- Se da un click en nueva y se coloca un nombre para dar click en crear

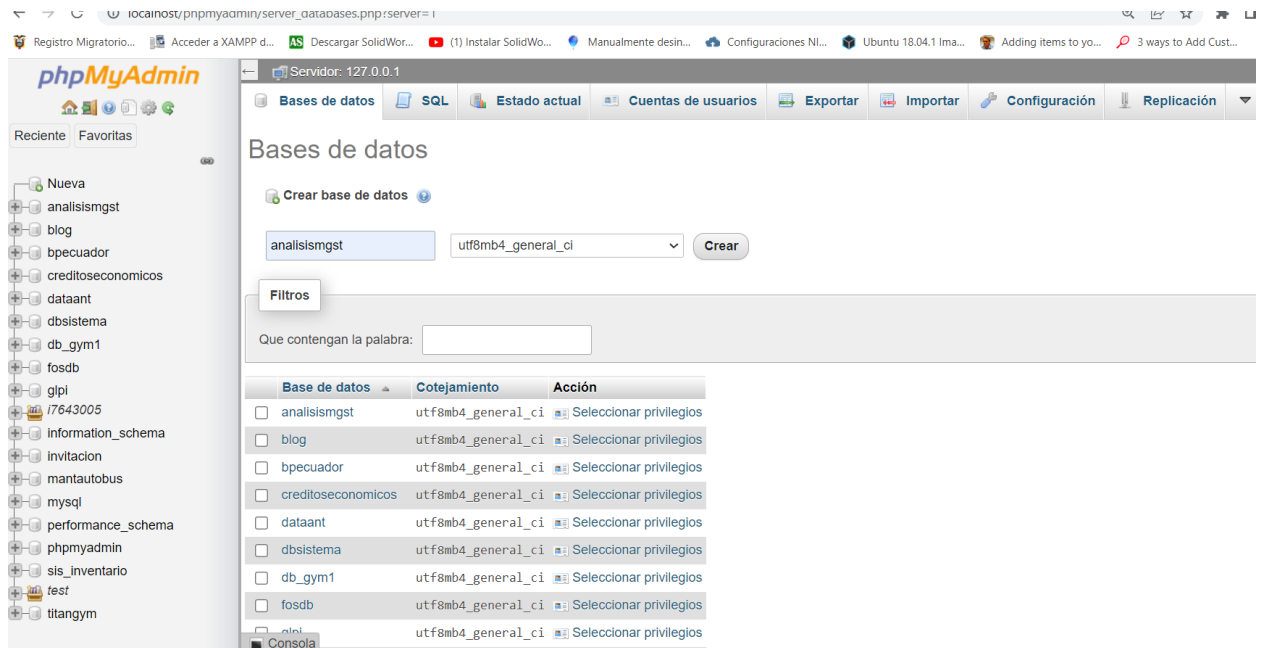


Figura 10. Colocar nombre

Nota. Este es el manejador de base datos usado para el hospital; es posible crear varias dependiendo el tipo de ordenador y sus características.

3.2 Transformación

La información receptada se almacena en una base de datos relacional, la misma se guarda en un servidor con una instalación de SQL Server 201.

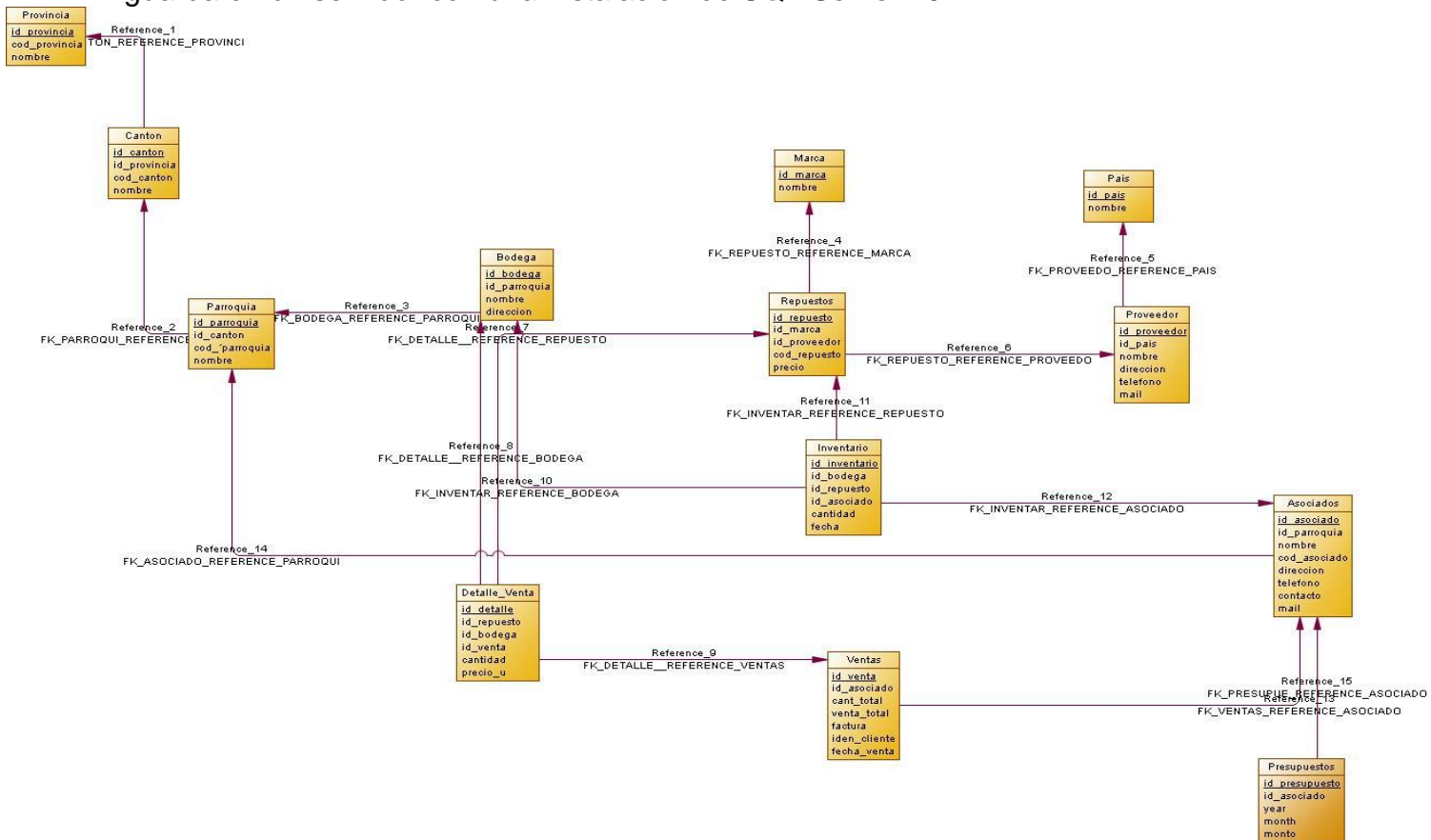


Figura 11. SQL Server 21

Navicat Premium 15, es una herramienta de desarrollo de bases de datos de múltiples conexiones la cual le permite conectar varios tipos de bases de datos dentro de una sola aplicación: MySQL cree un acceso rápido y fácil a todas sus bases de datos, todo de una vez.

- Una vez creada la Base de datos se utiliza el navicat para importar los datos de *excel* para lo cual es necesario crear una conexión con los siguientes datos:
- Se dirige a conexión Mysql

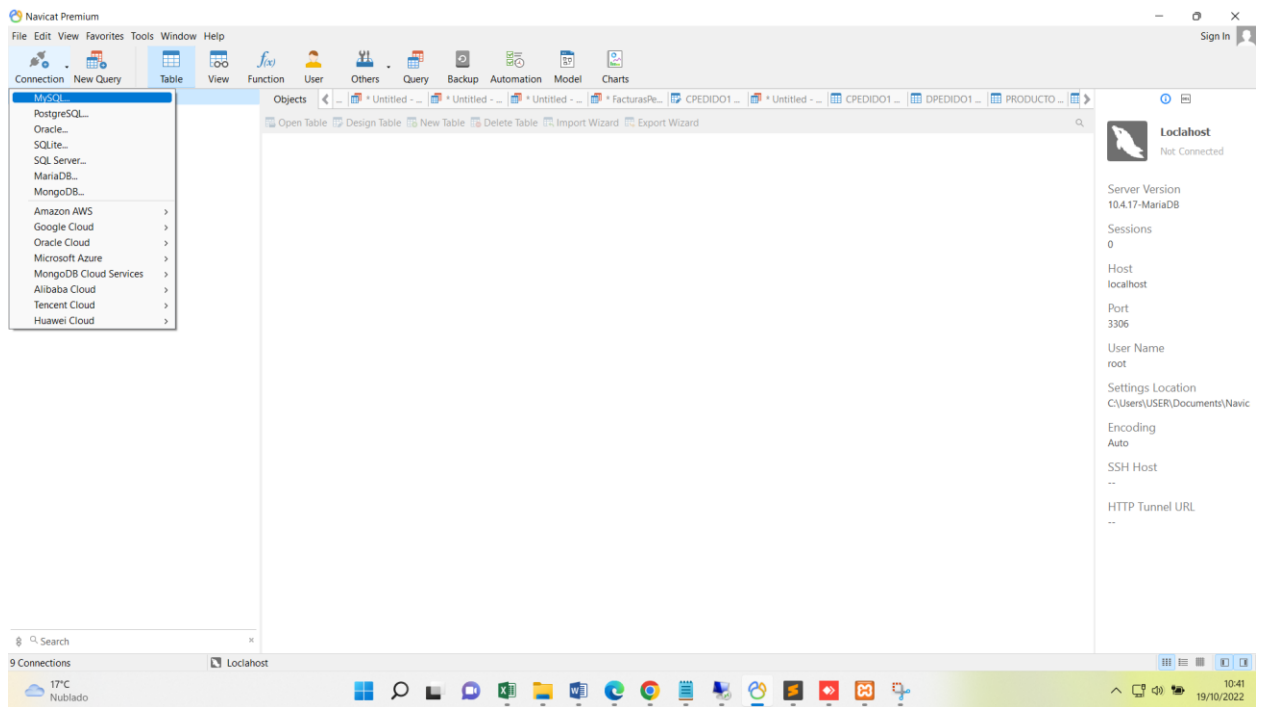


Figura 12. Navicat

- Se crea la conexión con los siguientes datos

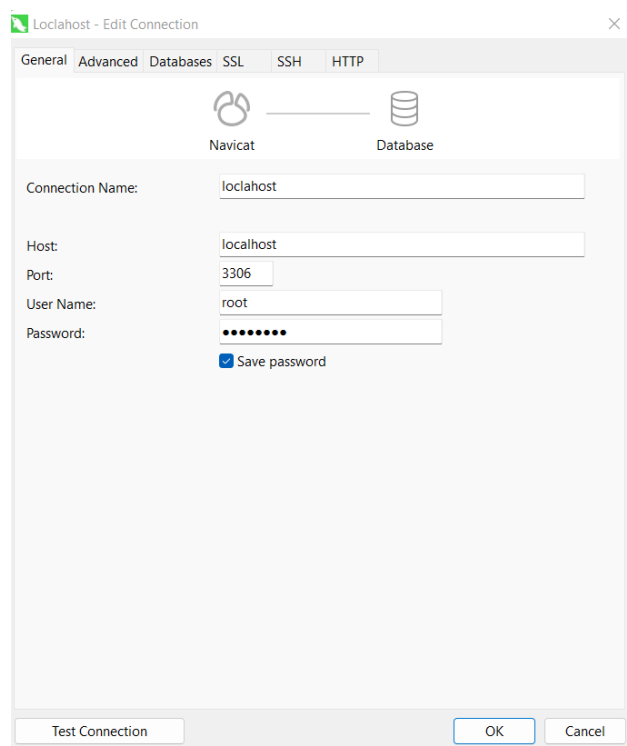


Figura 13. Creación de conexión

- Ahora se da un doble click sobre la BD analisismgst para crear las tablas importando I0061 información dada en databases, cabe indicar que cada archivo de Excel será una tabla.

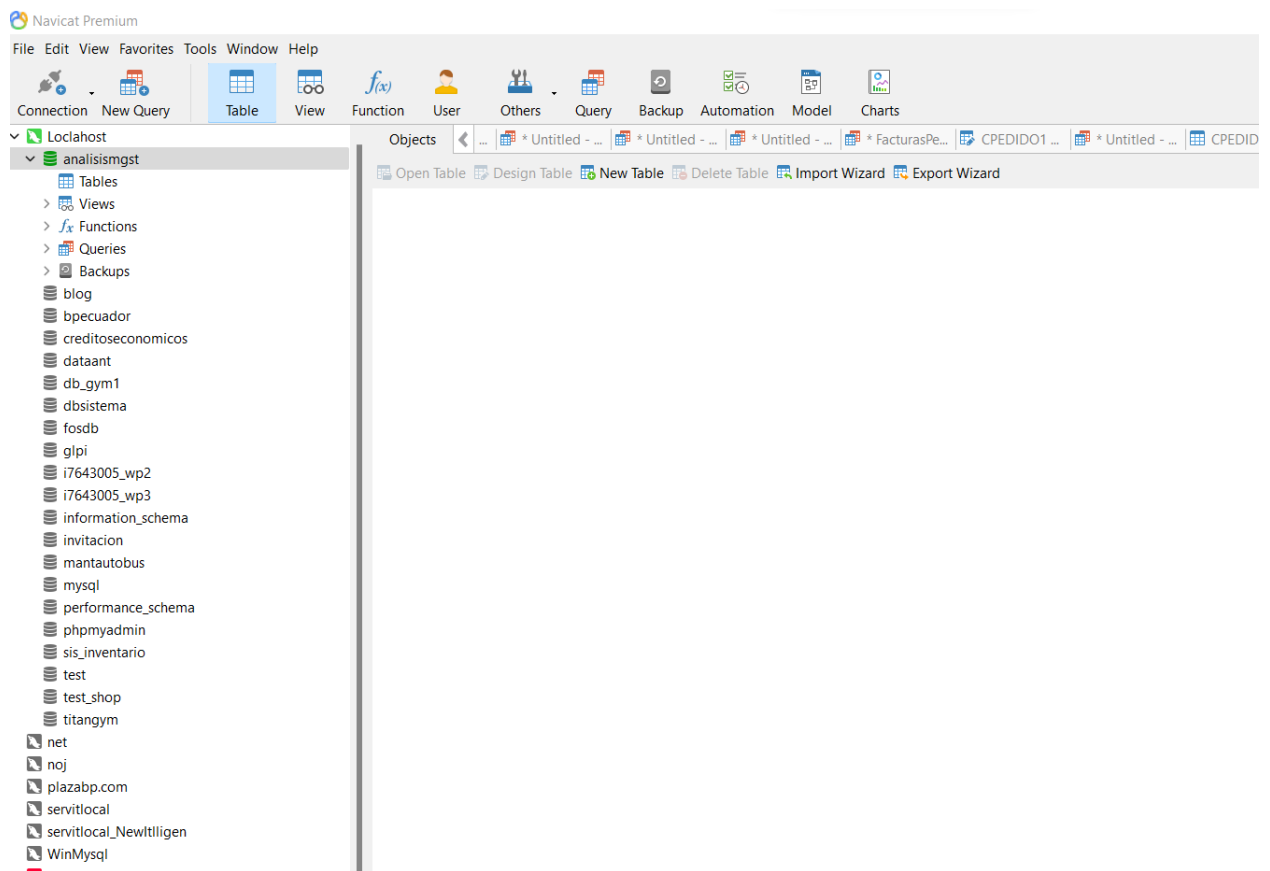


Figura 14. BD analisismgst

ETL base

Seguido a la carga de base de datos, se procede con un proceso ETL, establecido en 3 fases para lo que se obtienen las figuras posteriores .

- Borrado de tabla DSA

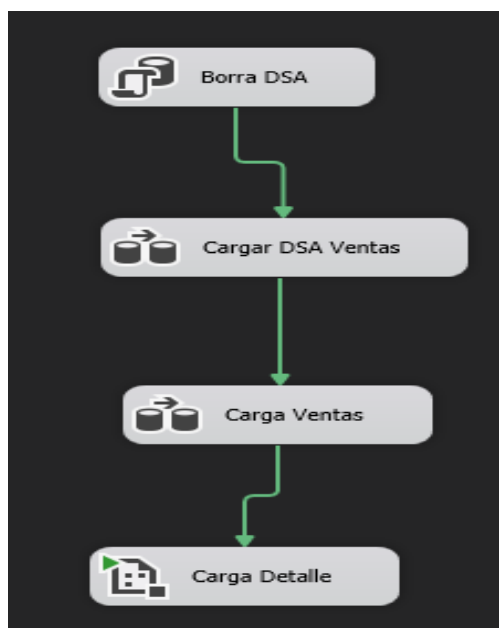


Figura 15. Borrado DSA

- Carga del archivo Excel a la tabla DSA

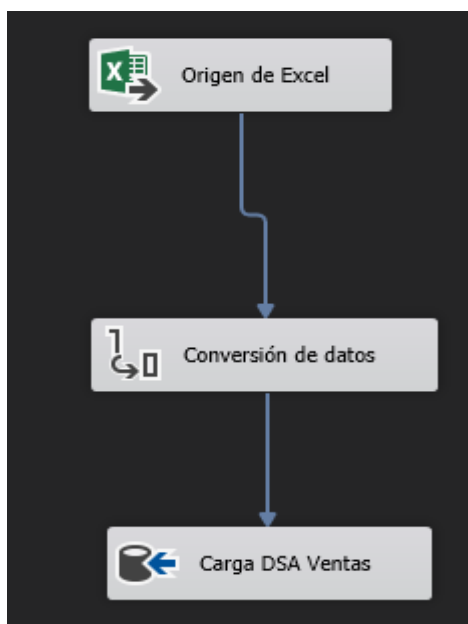


Figura 16. Carga del archivo excel

- Carga de DSA a la base de datos

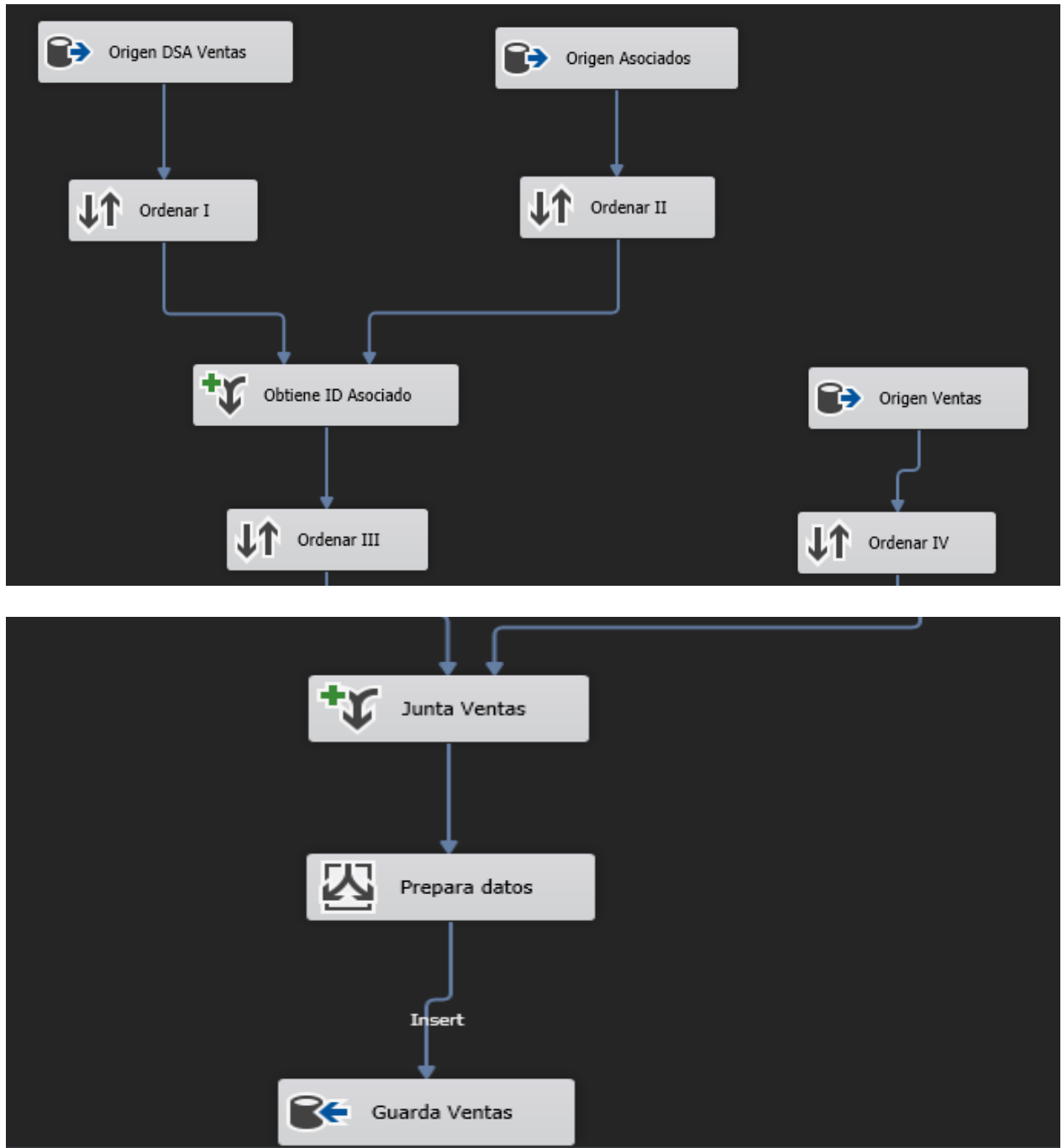


Figura 17. Carga de base de datos

- Una vez ingresado es posible conectarse a la base de datos, se da un doble click sobre localhost y se abre la conexión con la base de datos creada

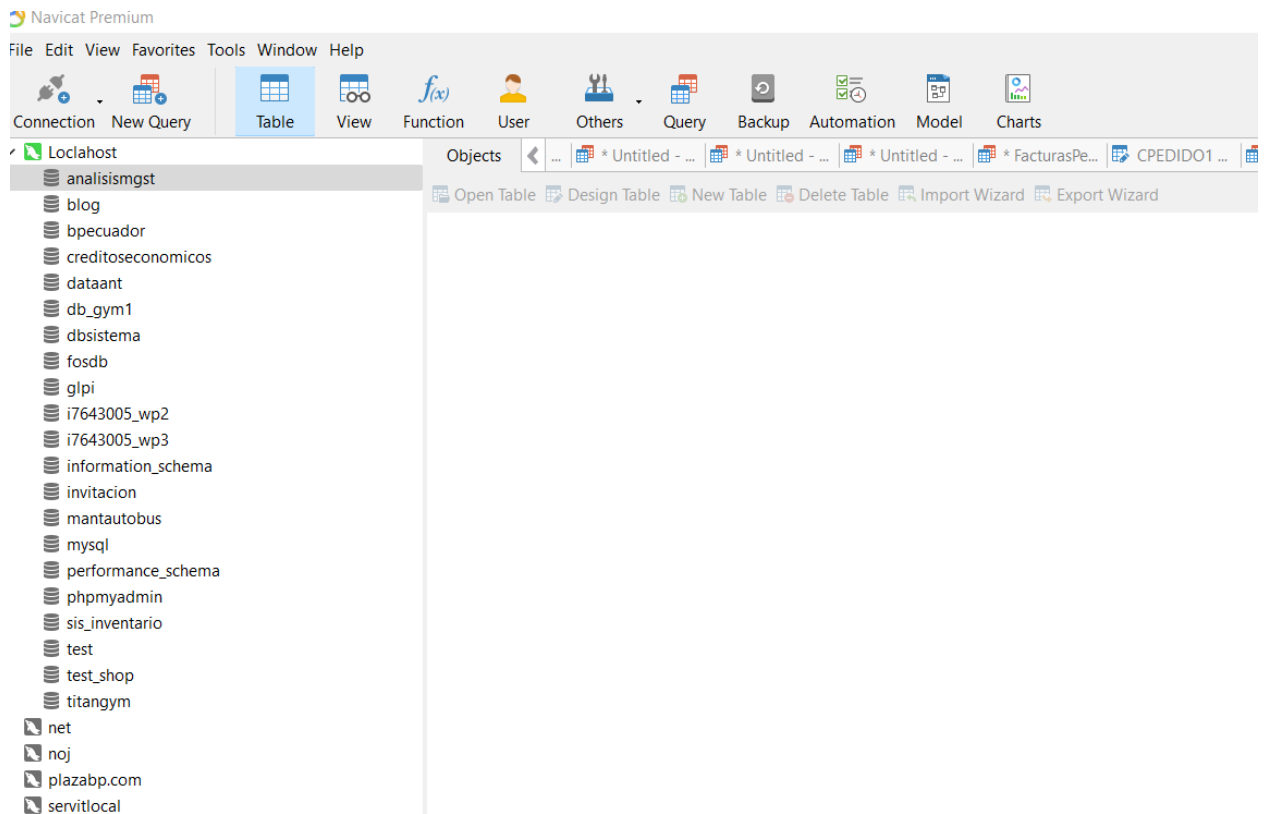


Figura 18. Abrir base de datos creada

- Luego se realiza el proceso desde importación del resto de años hasta completar todos los archivos. Escoger el nombre de campo que coincida en la tabla clic en siguiente y start, nuevamente esto hacer con todos los archivos de los demás años para tener en una sola tabla toda la información disponible.
- Luego de Importar toda la información, es posible sacar un *backup* de la Data y la Estructura de la Base de datos.

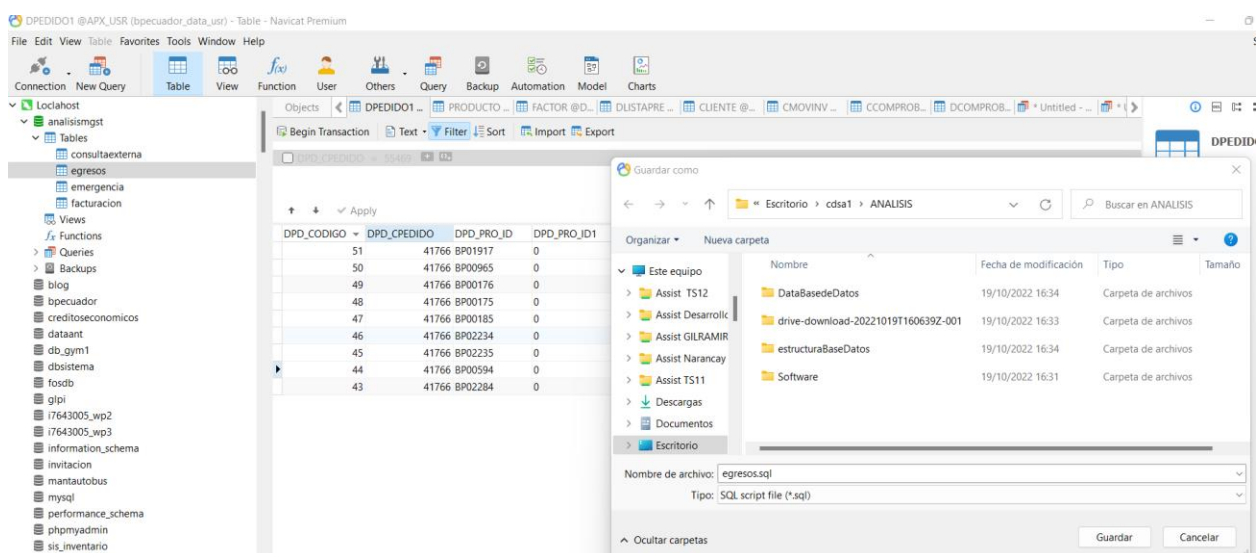
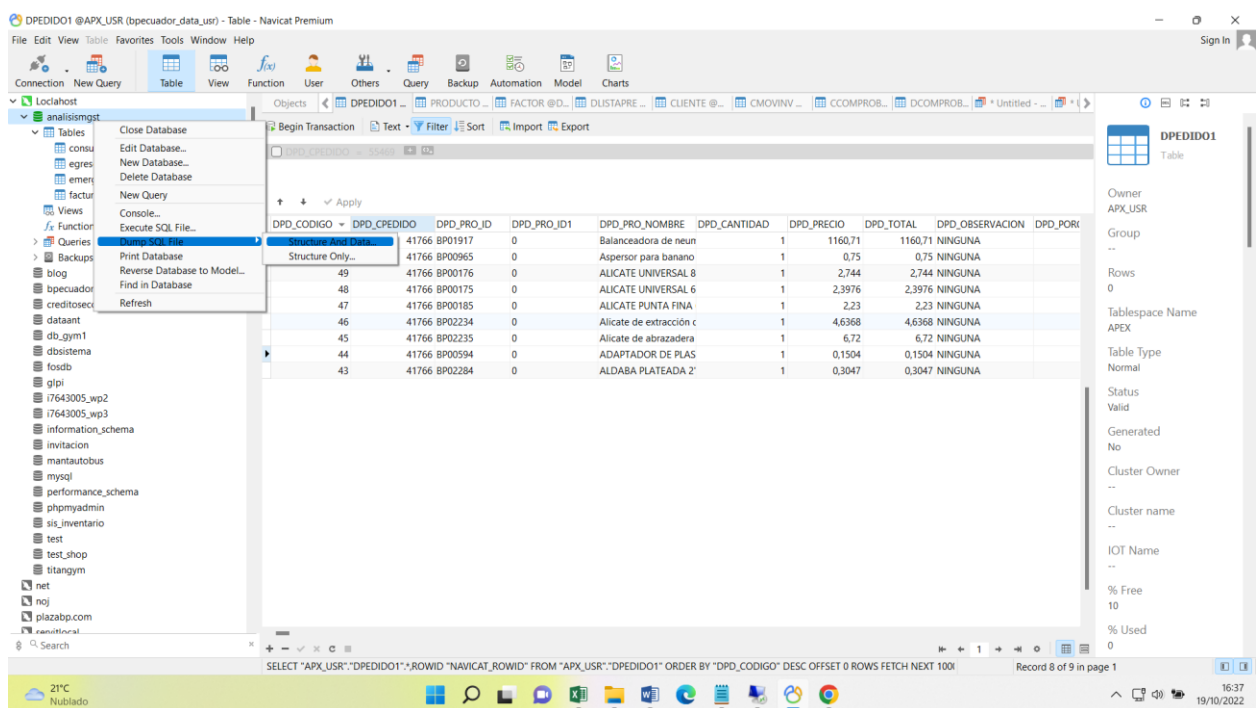


Figura 19. Estructura base de datos

Nota: Al escoger *structure* y *data* se realiza un backup de todo los tipos de datos y la data insertada. Al escoger *structure* solo se realiza un backup de la estructura. El formato del archivo generado es “ejm estructuraegresos.sql”, y se puede visualizar con un programa llamado sublime text que están dentro del software.

```
16
17 SET NAMES utf8mb4;
18 SET FOREIGN_KEY_CHECKS = 0;
19
20 -----
21 -- Table structure for egresos
22 -----
23 DROP TABLE IF EXISTS `egresos`;
24 CREATE TABLE `egresos` (
25   `HINUHI` int NULL DEFAULT NULL,
26   `IAFCED` int NULL DEFAULT NULL,
27   `IAFTAF` varchar(5) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
28   `IAFNOM` varchar(150) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
29   `IAFSEX` varchar(5) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
30   `IAFFNA` date NULL DEFAULT NULL,
31   `HIFEIN` date NULL DEFAULT NULL,
32   `HIHOIN` int NULL DEFAULT NULL,
33   `HIFEAL` date NULL DEFAULT NULL,
34   `HIHOAL` int NULL DEFAULT NULL,
35   `DIAS` int NULL DEFAULT NULL,
36   `HIMOAL` varchar(5) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
37   `HISVOR` int NULL DEFAULT NULL,
38   `ICDESA` varchar(150) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
39   `HIHBOR` int NULL DEFAULT NULL,
40   `HICMOR` int NULL DEFAULT NULL,
41   `HISVAC` int NULL DEFAULT NULL,
42   `HIHBAC` int NULL DEFAULT NULL,
43   `HICMAC` int NULL DEFAULT NULL,
44   `HIESCL` varchar(5) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
45   `HIESIN` varchar(5) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
46   `HIDIAG` varchar(5) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
47   `HIMEOR` int NULL DEFAULT NULL,
48   `HIMETR` int NULL DEFAULT NULL,
49   `HIRES` varchar(20) CHARACTER SET utf8mb4 COLLATE utf8mb4_general_ci NULL DEFAULT NULL,
50   `EDAD` double(4, 2) NULL DEFAULT NULL,
51   `1a` int NULL DEFAULT NULL,
```

Figura 20. Codificación

3.3 Distribución

Una vez procesados los datos dentro de una BD de mysql, es hora de conectar la herramienta de análisis de datos con el software POWERBI.

- Para poder tener conexión con la Base de Datos creada, se debe instalar un conector mysql-installer-web-community-8.0.31.0.msi.

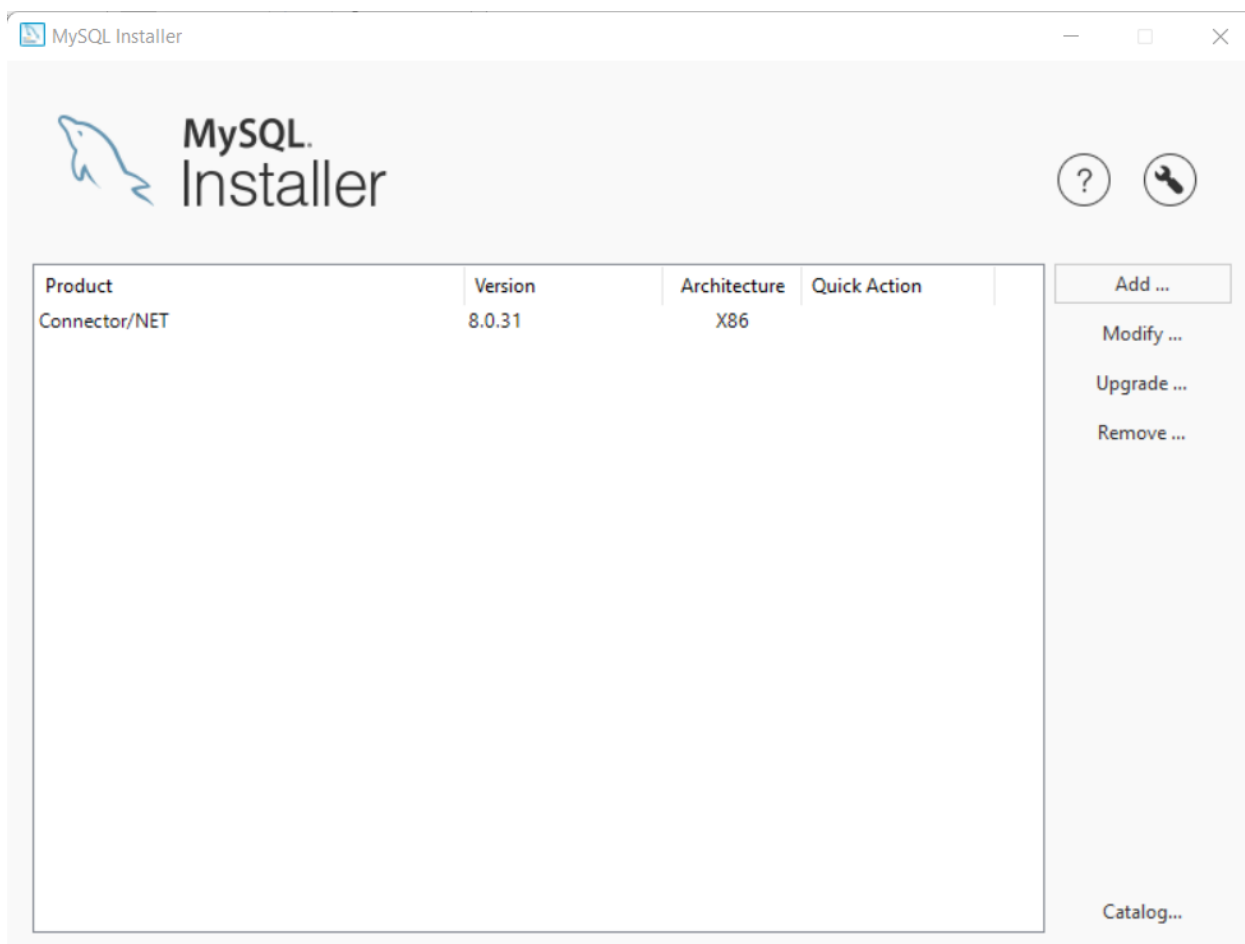


Figura 21. MySQL

Microsoft Power BI Desktop, se creó para los analistas. Combina visualizaciones interactivas de última generación con consultas y modelado de datos líderes en la industria integrados. Cree y publique sus informes en Power BI. Facilita a otras personas información fundamental puntual, en cualquier momento y desde cualquier lugar.

- Una vez instalado, se debe abrir el software PowerBi, y procedemos a conectarnos.

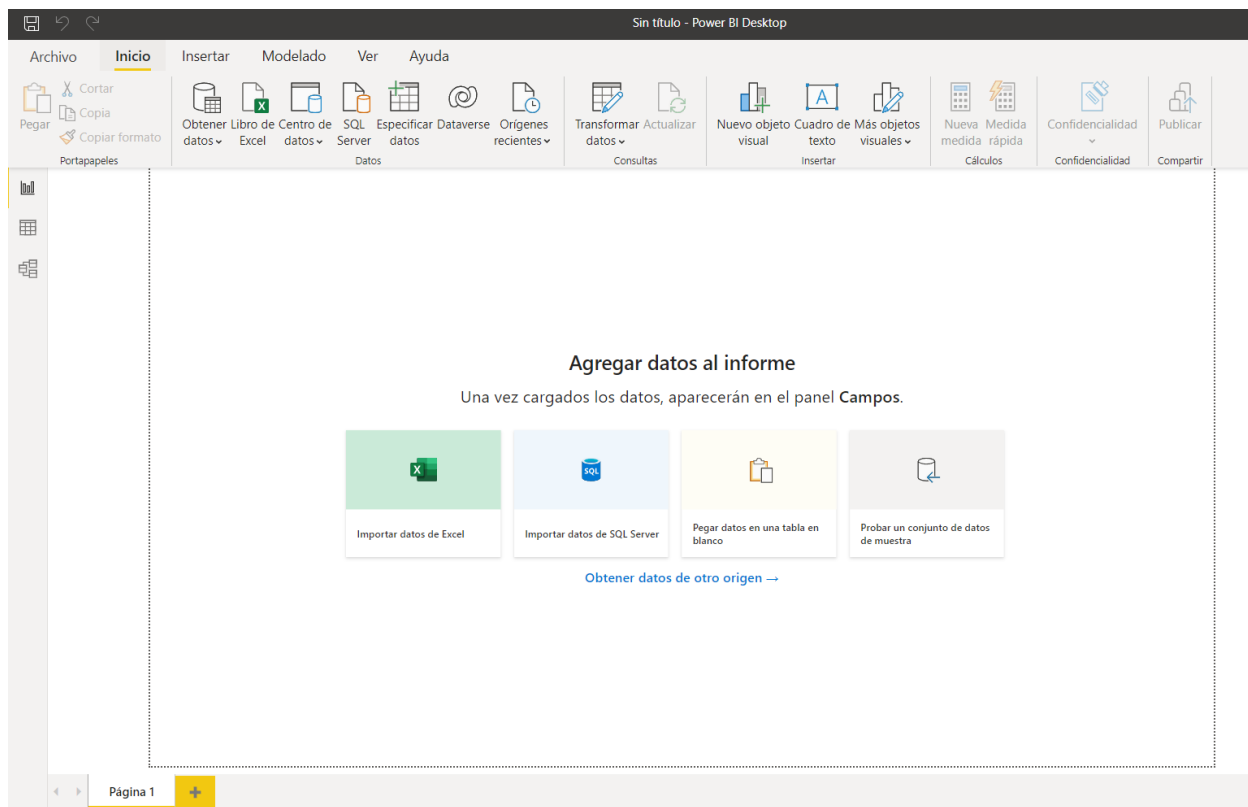


Figura 22. Power BI

- Se da click en obtener datos de otro origen

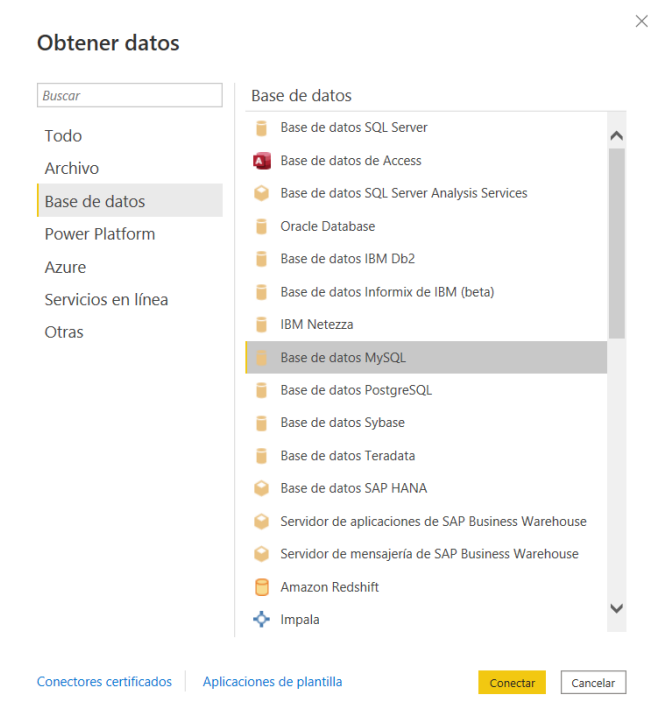


Figura 23. Obtención datos de otro origen

- Se escoge Base de Datos-> Base de Datos Mysql y click en Conectar

Figura 24. Base de datos

- Seguido se ingresan los datos del servidor instalado localmente:
Servidor: localhost
Base de Datos: analisismgst

Figura 25. Datos del servidor

On Premise Data gateway, puerta de enlace de datos local de Microsoft, mantiene los paneles y sus informes actualizados mediante la conexión a los orígenes de datos locales, sin necesidad de desplazar los datos.

- Se clickea en Usar Credenciales alternativas e ingresamos los datos:
Nombre de Usuario: root
Contraseña: root o la contraseña que se haya colocado a la hora de instalar el software Xampp; click en conectar.

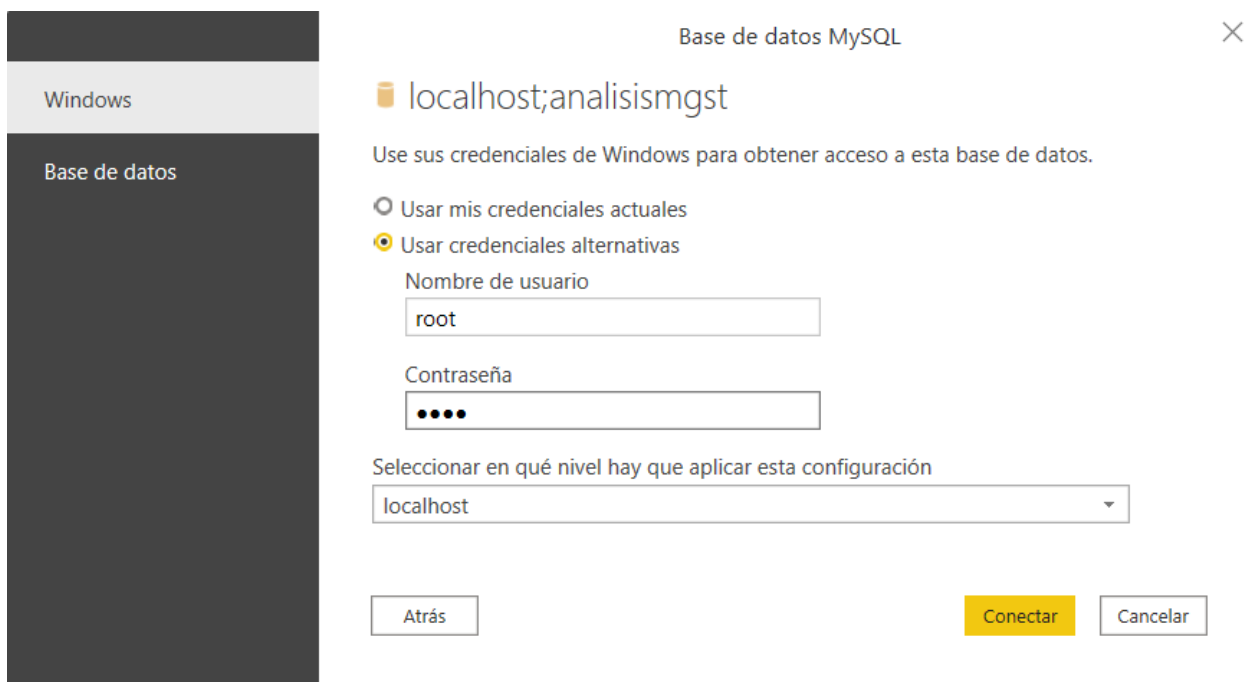


Figura 26. Credenciales

3.4 Consumo

A continuación, se describen las fases del desarrollo del proceso ETL que fue implementado para el sistema.

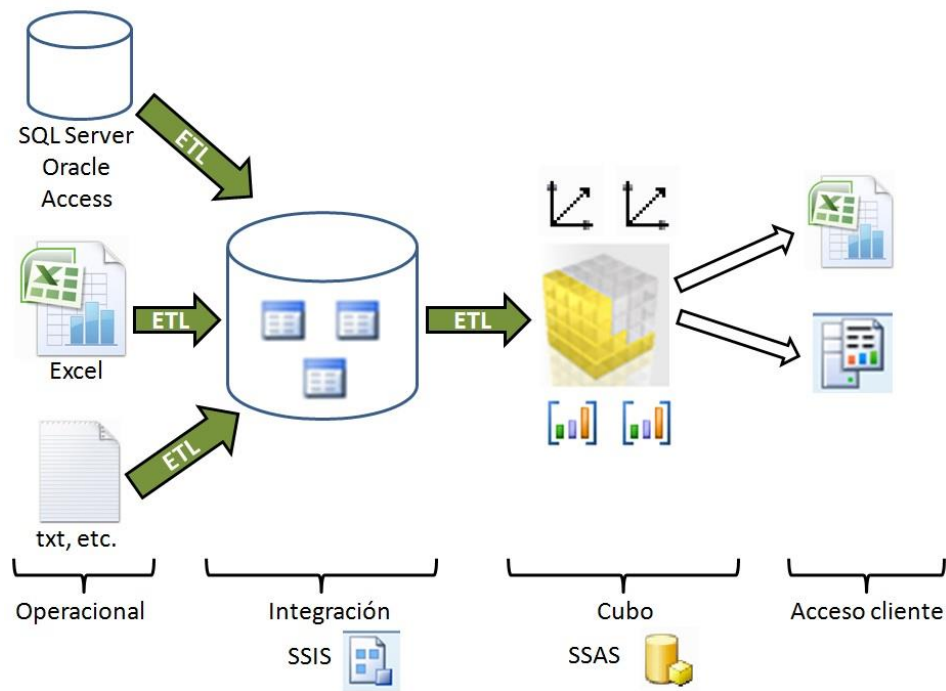


Figura 27. Proceso ETL

1. El sistema extrae la información requerida de la página principal y es almacenada en un archivo con extensión txt para después ser depositado a un directorio específico. Una vez depositado el archivo en el directorio se realiza la carga de información en una base de datos origen para iniciar el proceso de extracción. La información es separada para obtener el layout.
2. Proceso de extracción, la información que se encuentra en la base de datos origen es seleccionada, extraída por medio de SQL y mantenida en memoria para después ser procesada en la fase de transformación del ETL. El proceso de extracción ejecuta después de que ha finalizado la extracción del archivo de texto origen y la carga de la información hacia la base de datos origen, con la finalidad de mantener los datos en memoria.
3. Proceso de transformación: se necesitaba fragmentar ciertos datos, esto se logró con un módulo JAVA y el uso de expresiones regulares.
4. Proceso de carga: una vez que termina el proceso de transformación, lo que sigue es la actualización y carga a una base de datos destino, en esta parte se hace un llenado automático de aquellos campos que no se encuentran informados. El proceso finaliza con la generación de un archivo con extensión .xls el cual es funcional para realizar cálculos y análisis correspondientes.

5. La finalidad de la implementación del sistema es obtener la información necesaria y funcional para la estrategia del hospital.
6. Ahora aparece toda la Base de datos que se creo

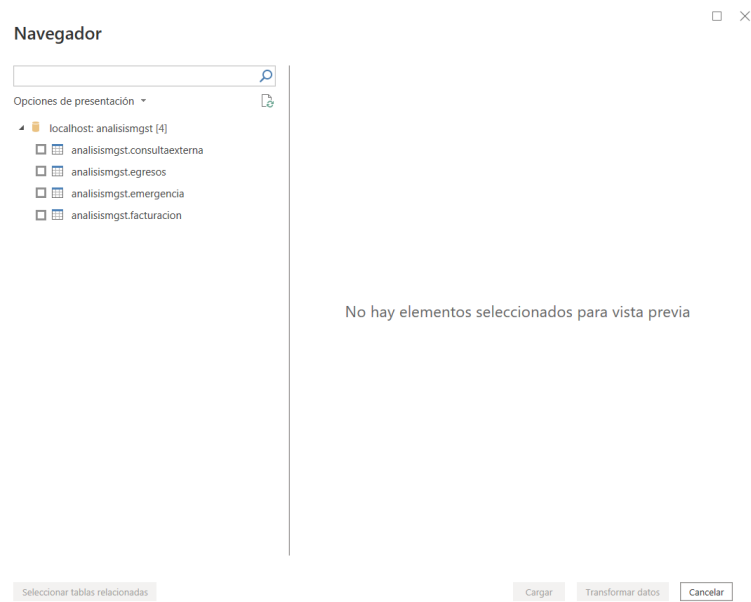


Figura 28. Base datos creada

Nota: se eligen todos porque se van a crear las hojas de datos a partir de la información de esas tablas, luego se da click en cargar.

Navegador

localhost: analisismgst [4]

- analisismgst.consultaexterna
- analisismgst.egresos
- analisismgst.emergencia
- analisismgst.facturacion

analisismgst.facturacion

DEPENDENCIA	SNA1	FECHA	TIPO SEG	CEDULA	PACIEN
143	2017186133	31/1/2017	AG	1208881357H1	CHANGI
312	2017186133	31/1/2017	AG	1208881357H1	CHANGI
312	2017186133	31/1/2017	AG	1208881357H1	CHANGI
143	2017186133	31/1/2017	AG	1208881357H1	CHANGI
143	2017186133	31/1/2017	AG	1208881357H1	CHANGI
143	2017186133	31/1/2017	AG	1208881357H1	CHANGI
143	2017186133	31/1/2017	AG	1208881357H1	CHANGI
142	2017186141	25/1/2017	SG	1201254909	ANDALL
312	2017186141	25/1/2017	SG	1201254909	ANDALL
312	2017186141	25/1/2017	SG	1201254909	ANDALL
142	2017186141	25/1/2017	SG	1201254909	ANDALL
142	2017186141	25/1/2017	SG	1201254909	ANDALL
142	2017186141	25/1/2017	SG	1201254909	ANDALL
142	2017186141	25/1/2017	SG	1201254909	ANDALL
142	2017186141	26/1/2017	SG	1201254909	ANDALL

i Los datos de la vista previa se han truncado debido a límites de tamaño.

Seleccionar tablas relacionadas

Cargar

Transformar datos

Cancelar

Figura 29. Análisis

CAPÍTULO 4: ANÁLISIS DE RESULTADOS

4.1 Puesta en marcha y funcionamiento

En cuanto al proceso, se analizaron las transacciones del Hospital desde el 1 de enero de 2017 hasta el 31 de diciembre de 2021. El último proceso y no por ello menos importante, es el encargado de visualizar los resultados obtenidos del procesamiento de los datos. Los resultados presentan la información relevante que responde a la producción de consulta externa, atenciones en área de emergencia, hospitalización y datos sobre facturación. En cada una de estas categorías se presenta la información dividida por indicadores relevantes. El indicador que todas comparten en común es el periodo de análisis que será tomado desde el 2017 hasta el 2021. Las consultas serán presentadas a través de datos y gráficos estadísticos. Para luego, poder definir indicadores de gestión que midan la eficiencia del sistema de servicio, este análisis se organizó partiendo de un resumen general de las transacciones que fue profundizando hasta obtener descubrimientos específicos.

- Se va a empezar a cargar la información al *power bi* para procesarla

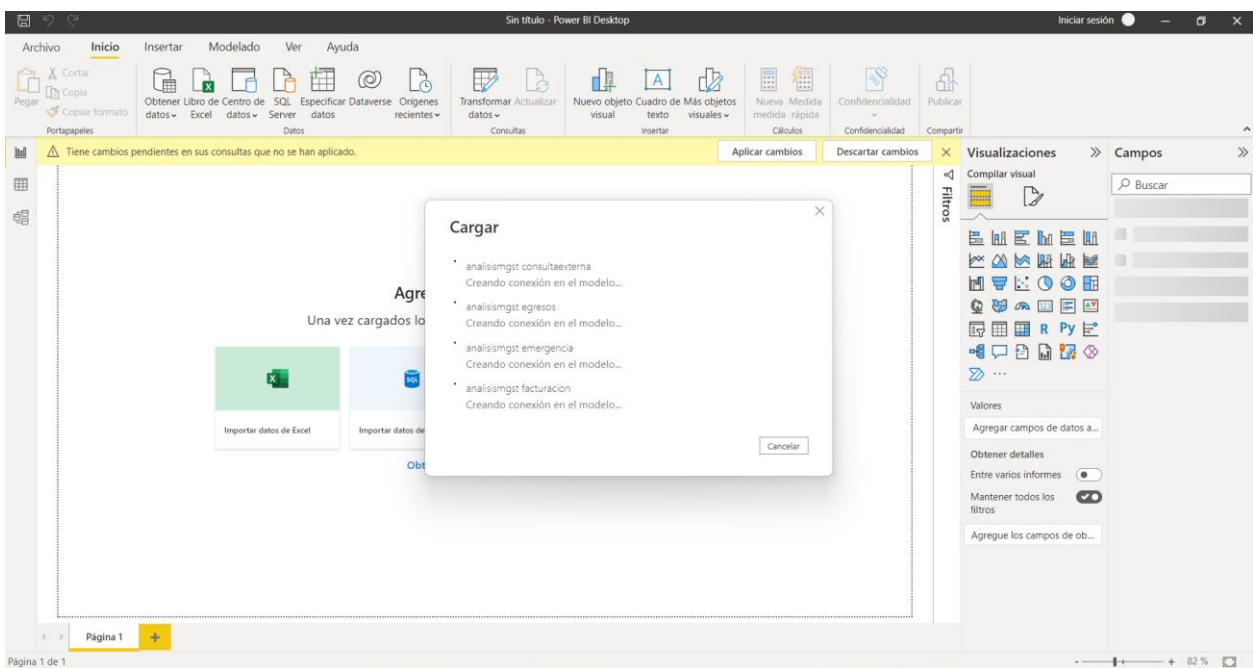


Figura 30. Carga de información

- Una vez leído todos los datos, aparecen las tablas cargadas para poder darle sentido a la información



Figura 31. Tablas

- Se guarda la hoja de datos para tener ya cargado toda la data y se le coloca un *analismgst.pbix* .

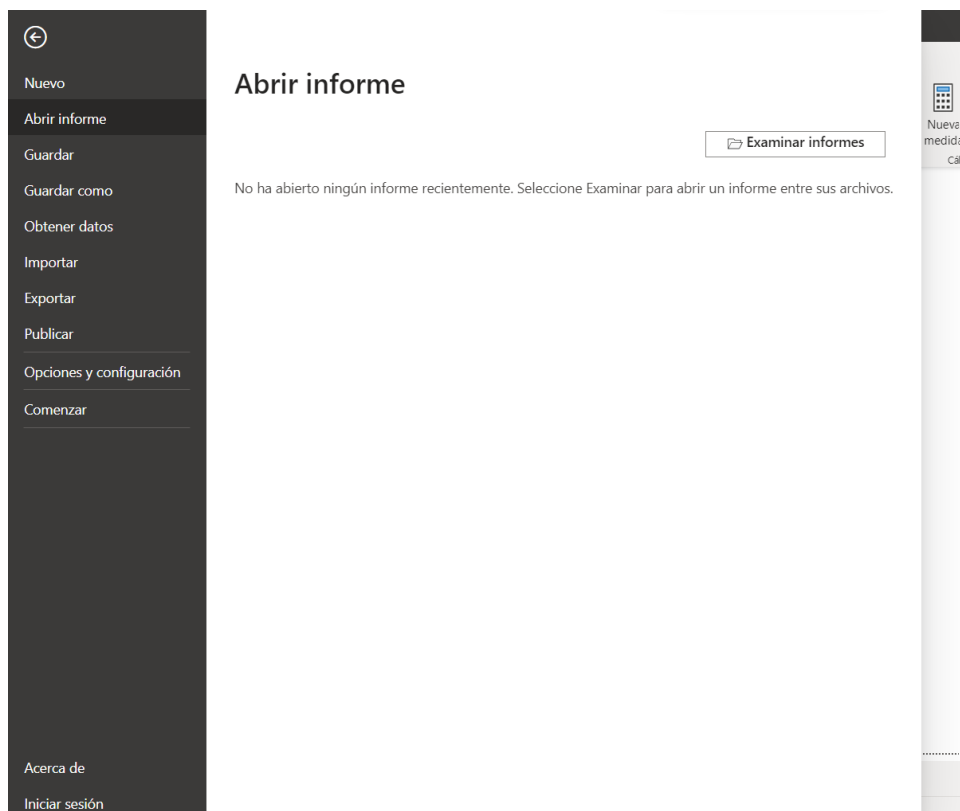


Figura 32. Informe

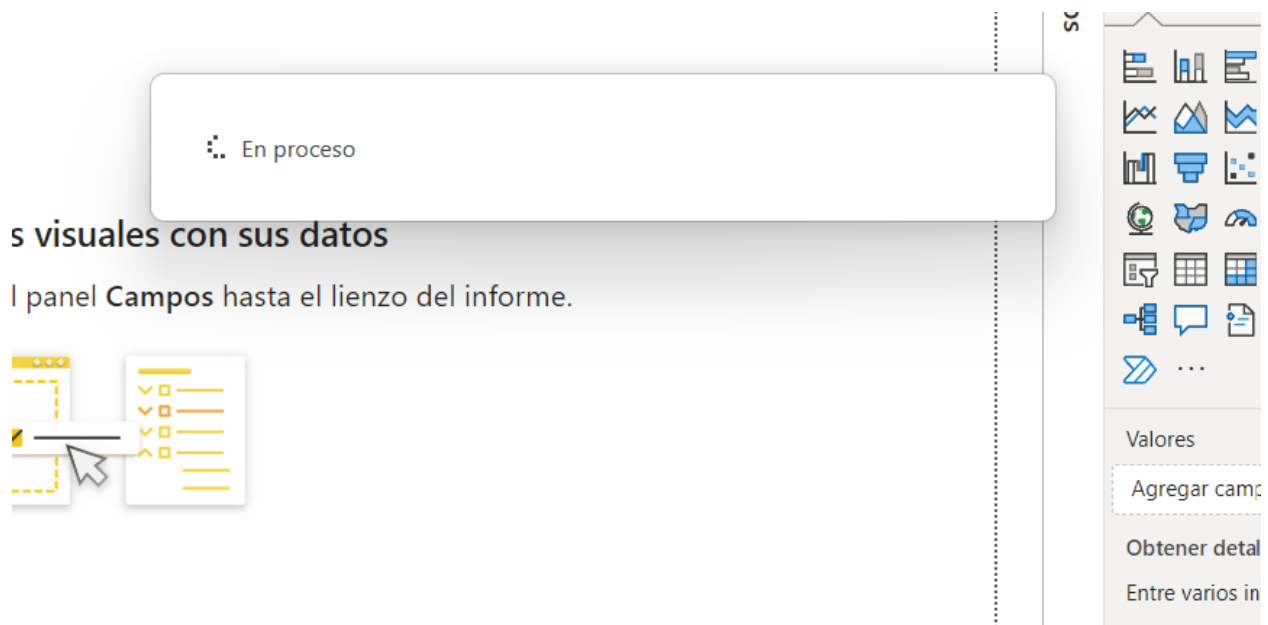
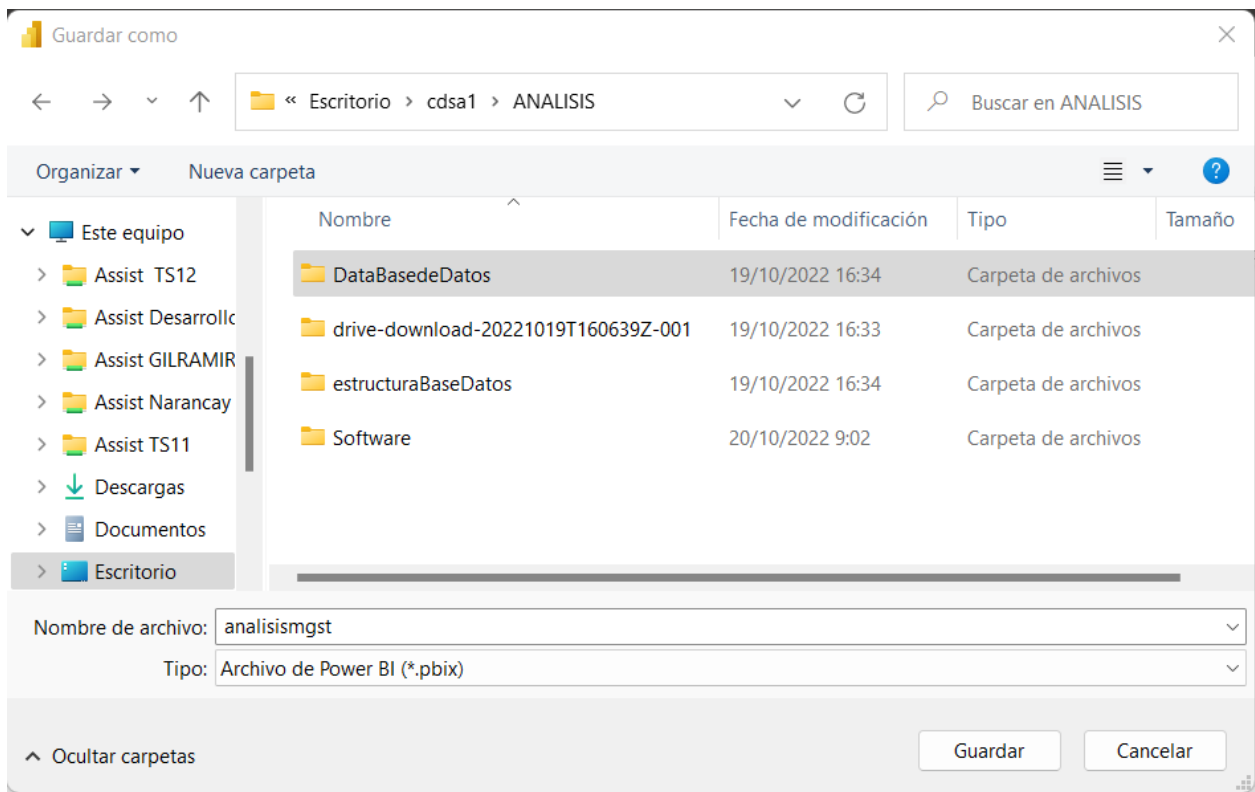


Figura 33. Informes generados

4.2 Análisis de datos

Se tienen en cuenta los totales de las citas en cada área de estudio.

Consulta Externa

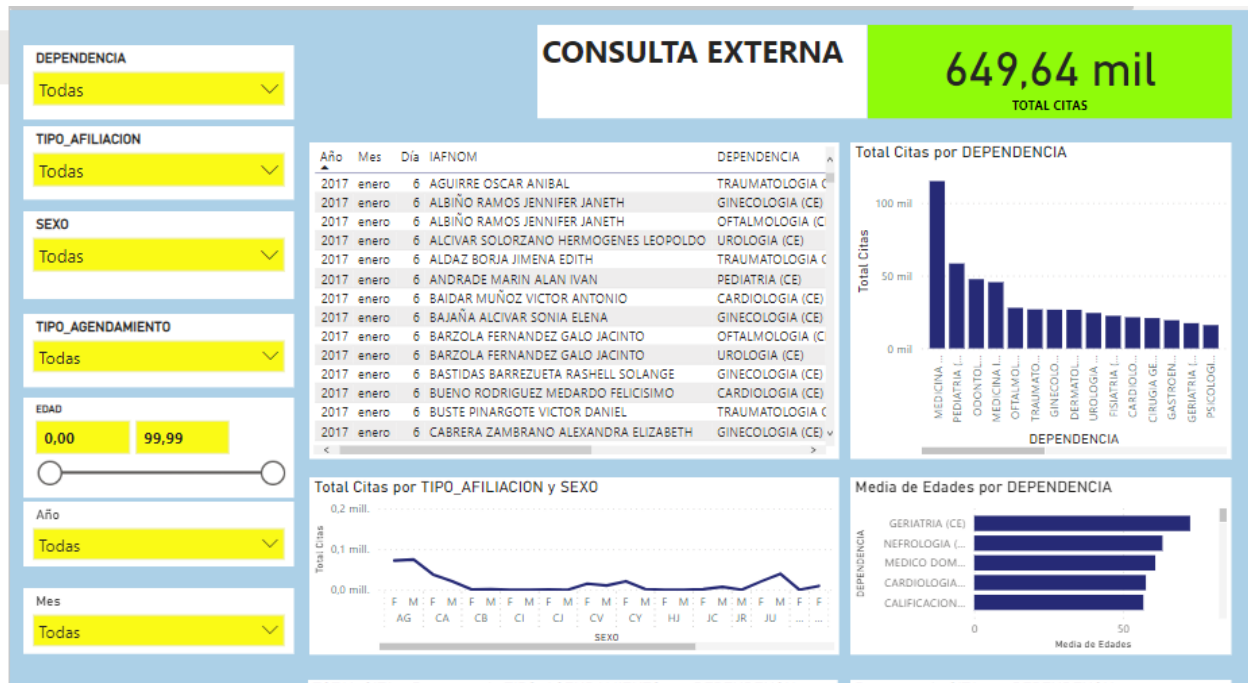


Figura 34. Informe consulta externa

Los resultados se explican empezando con un resumen general de las transacciones, los montos recaudados y el tipo de usuarios (dependencia), el total de transacciones (citas) en los 4 años para consulta externa es 649,64 mil.

Hospitalización

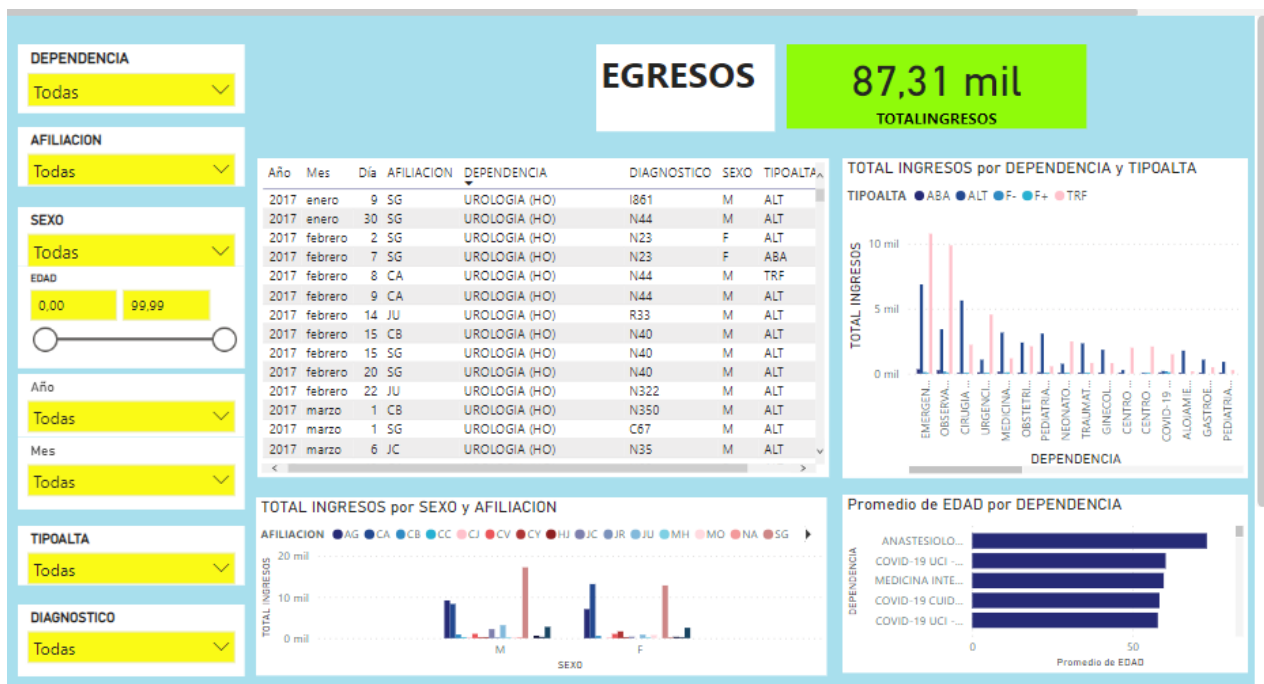


Figura 35. Informe hospitalización

Los resultados se explican empezando con un resumen general de las transacciones, los montos recaudados y el tipo de usuarios (dependencia), el total de transacciones (citas) en los 4 años para hospitalización es 87,31 mil.

Emergencia

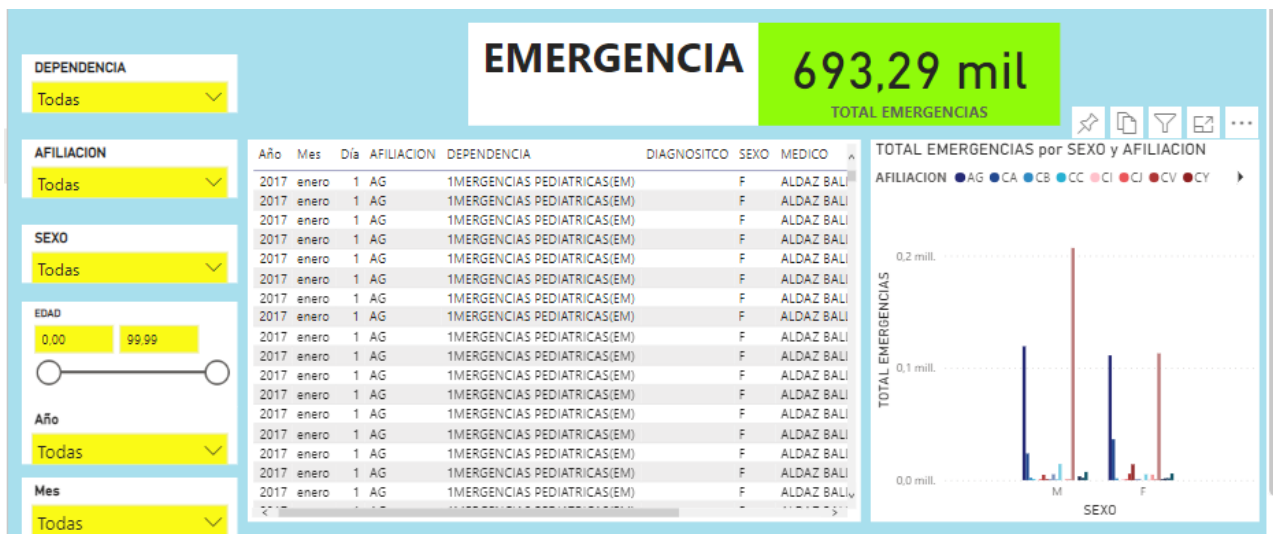


Figura 36. Informe emergencia

Finalmente, los resultados se explican empezando con un resumen general de las transacciones, los montos recaudados y el tipo de usuarios (dependencia), el total de transacciones (citas) en los 4 años para emergencia es 693,29 mil.



Figura 37. Informe facturación

De forma general, estos procesos representan una recaudación de \$35,79mill.

CONCLUSIONES

De acuerdo con los resultados obtenidos, es posible concluir:

- A partir del análisis y procesamiento de datos se ha logrado extraer características relevantes de las atenciones de los pacientes, lo que permitirá visualizar los resultados de forma más fácil y eficaz ayudando a generar una alta precisión en la toma de decisiones por parte de los profesionales encargados
- Con la implementación del sistema realizado se podrá acopiar todos los datos en un solo archivo y evitará realizar el análisis de los datos de forma aislada como se venía realizando anteriormente además permitirá la optimización de procesos que, al momento son realizados manualmente o mediante el uso de sistemas utilitarios.
- Este software de visualización podrá mejorar la forma en la que se examinan los datos del Hospital, permitiendo que el encargado de realizar estos reportes pueda tener un control total de los datos obtenidos, mismos que serán presentados con tablas y gráficos estadísticos lo cual permitirá la obtención de datos reales, claros y organizados.

RECOMENDACIONES

- Una vez que se cuente con los datos generados en el año 2022 se recomienda agregar al sistema ya que actualmente no existen datos relacionados a este año.
- En proyección futura se recomienda adjuntar este programa como herramienta para futuro donde todos los centros de atención hospitalaria puedan acceder a la información generada de esta manera las decisiones de forma global.

BIBLIOGRAFÍA

- Arcila Calderón, C., Barbosa Caro, E., & Cabezuelo Lorenzo, F. (2016). *Técnicas big data: análisis de textos a gran escala para la investigación científica y periódica*. Obtenido de <http://eprints.rclis.org/34193/>
- Arevalo Tenelema, M. (2019). *Implementación de un prototipo de sistema de alerta para conductores distraídos y somnolientos de*. Obtenido de <http://dspace.esPOCH.edu.ec/bitstream/123456789/13667/1/108T0308.pdf>
- Botero Villada, M. (Mayo de 2019). *Estudio comparativo de metodologías para minería de datos*. Obtenido de https://www.scipedia.com/public/Botero_Villada_2019a
- Brown, P. N., & Marden, J. R. (2017). The benefit of perversity in taxation mechanisms for distributed routing. *IEEE 56th Annual Conference on Decision and Control (CDC)*. Melbourne.
- Burris, M. W. (2004). *Assessing the Benefits and Costs of ITS: Making the Business Case for ITS. Electronic Toll Collection and Variable Pricing*. Texas: Sponsored by the California Department of Transportation.
- CACERES, M., & LOPEZ, F. (2018). *SISTEMA DE INFORMACIÓN GERENCIAL PARA LA MEJORA DE LA TOMA DE DECISIONES DEL AREA DE RECURSOS HUMANOS DEL MINISTERIO DE SALUD APLICANDO BUSINESS INTELLIGENCE*. Obtenido de <https://repositorio.utelesup.edu.pe/handle/UTELESUP/1282>
- Cadmen, J. (2019). *Propuesta para el Desarrollo de un modelo de Arquitectura Empresarial para el gobierno centralizado del canton Morona*. Obtenido de https://repositorio.uta.edu.ec/bitstream/123456789/30564/1/Tesis_t1649msi.pdf
- Challenger, I., Díaz, Y., & Becerra, R. (Junio de 2014). *El lenguaje de programación Python*. Obtenido de <https://www.redalyc.org/pdf/1815/181531232001.pdf>
- Devlin, K. (1991). *Logic and Information*. Cambridge University Press, 16.
- Fernández, V. (2006). *Desarrollo de Sistemas de Información*. Barcelona: Edicions UPC.

- Fernández, Y., & Díaz, Y. (Abril de 2012). *Patrón Modelo-Vista-Controlador*. Obtenido de <https://revistatelematica.cujae.edu.cu/index.php/tele/article/view/15/10>
- García, A. (2018). Análisis de los Modelos de Inteligencia de Negocios basados en Big Data en las Pymes del Ecuador Analysis of Business Intelligence Models based on Big Data in SMEs in Ecuador. *Revista Ciencia y Tecnología*, 46-56.
- Gastesi, R. (marzo de 2019). *Estudio comparativo del sistema informático AS-400 del IESS con el REDACCA en los agendacion de fichas y en el control de fichas medicas*. Obtenido de <http://dspace.utb.edu.ec/bitstream/handle/49000/5599/-E-UTB-FAFI-SIST-000137.pdf?sequence=1&isAllowed=y>
- Guerrero, J., & Tupia, F. (10 de Junio de 2019). *Arquitectura Empresarial para el proceso de compra de prestaciones de salud*. Obtenido de https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/626013/tupia_vf.pdf?sequence=3&isAllowed=y
- Hernández, R., Fernández, C., & Baptista, P. (2016). *Metodología de la investigación*. México: McGraw Hill.
- IEES. (s.f.). UNIDADES MÉDICAS I.E.S.S. Ecuador.
- IESS. (Noviembre de 2019). *Sistema de Gestión Médica*. Recuperado el 15 de Octubre de 2021, de <https://www.flacso.edu.ec/flax15/graficos/Rodney.pdf>
- IESS. (2020). *SE AMPLÍA LA CARTERA DE SERVICIOS DEL HOSPITAL GENERAL QUEVEDO*. Quevedo.
- LaPlante, A. (2019). *Getting Started with Machine Learning in the Cloud*. Sebastopol: O'Reilly Media.
- Larrea, L. (26 de Enero de 2016). Manual MIS_AS-400. Quito, Pichincha, Ecuador.
- Laudon, K., & Laudon, J. (2016). *Sistemas de Información Gerencial*. Obtenido de http://cotana.informatica.edu.bo/downloads/Id-Sistemas_de_informacion_gerencial_14%20edicion.pdf
- Mazon, B. (2018). *Inteligencia de negocios en el sector agropecuario*. Machala: Universidad Técnica de Machala.

Megasalud, Chile ; Área Informática Médica, Hospital Italiano de Buenos Aires, Argentina . (2008). *Rediseño del Sistema de Información de una Red Integrada de Salud de cobertura nacional en Chile, Proyecto SIRIS*. Obtenido de https://www.hospitalitaliano.org.ar/multimedia/archivos/servicios_attachs/4821.pdf

Moya, D., Tapia, L., Albán, M., & Rodríguez, G. (2020). Un enfoque de Machine Learning en el desarrollo de un Sistema Recomendadores para procesos de enseñanza. *Revista Ibérica de Sistemas e Tecnologías de Información*, 816-827.

Oracle. (2019). *Why are Data Lakes the Future of Big Data?* Obtenido de https://www.oracle.com/lad/a/ocom/docs/why-are-data-lakes-the-future-of-big-data-infographic.pdf?source=:ad:pas:go:dg:a_lad:71700000086979734-58700007414859471-p66509544640:RC_LAMK200526P00014C0001:DataManagementPAN

Oracle. (2020). *What´s the difference between data lake, database and a data warehouse?* Obtenido de https://www.oracle.com/lad/a/ocom/docs/database/difference-between-data-lake-data-warehouse.pdf?source=:ad:pas:go:dg:a_lad:71700000086979734-58700007414859471-p66509544640:RC_LAMK200526P00014C0001:DataManagementPAN

Oracle. (2021). *¿Qué es el big data?* Obtenido de https://www.oracle.com/lad/big-data/what-is-big-data/?source=:ad:pas:go:dg:a_lad:71700000086979734-58700007414859471-p66509544640:RC_LAMK200526P00014C0001:DataManagementPAN

Pastor, K. (25 de Septiembre de 2020). *ReactJS + Python Flask on Heroku. Making a Framework for API Development and Deployment*. Recuperado el 14 de Octubre de 2021, de <https://towardsdatascience.com/reactjs-python-flask-on-heroku-2a308272886a>

Pinglo, E., & Ramos, W. (2018). *Implementación de un sistema de información gerencial para mejorar la toma de decisiones en el área logística de un centro de salud en Lima*. Obtenido de <https://repositorio.utp.edu.pe/bitstream/handle/20.500.12867/2787/Erick%20Pingl>

o_Wilmer%20Ramos_Trabajo%20de%20Investigacion_Bachiller_2018.pdf?sequence=1&isAllowed=y

Python Software Foundation. (2022). *Python 3.10.5 documentation*. Obtenido de <https://docs.python.org/3/>

Ramesh, E. (2010). *Decision Support And Business Intelligence Systems*. Prentice Hall.

Reyes, Y. (2015). La inteligencia de negocio como apoyo a la toma de decisiones en el ámbito académico (Business Intelligence as decision support system in academic environment). *Revista Internacional de Gestión del Conocimiento y la Tecnología (GECONTEC)*, 67-73.

Ruiz, A. (2018). *MINERIA DE DATOS EN REDES SOCIALES PARA PYMES*. Jaén, Andalucía: Universidad de Jaén.

Salesforce. (2007). *Heroku*. Obtenido de <https://www.heroku.com/>

Salud, O. P. (OCTUBRE de 2018). *INTELIGENCIA ARTIFICIAL EN SALUD PUBLICA* . Obtenido de https://www.paho.org/ish/images/toolkit/IS4HCC_InteligenciaArtificial.pdf

Salvatierra, C., & Vasquez, M. (25 de 05 de 2022). *Propuesta de una arquitectura empresarial para una institucion de servicios de salud*. Obtenido de https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/624082/Salvatierra_CC.pdf?sequence=5

Secretaría General de Comunicación de la Presidencia. (2017). *El Hospital General de Quevedo refleja el manejo eficiente, transparente y responsable del IESS*. Quevedo.

Senn, J. (1992). *Análisis y diseño de sistemas de información*. México: McGraw-Hill.

Tamayo y Tamayo, M. (2017). *El proceso de la investigación científica*. México: Limusa-Noriega.

Timarán Pereira, S., Hernández Arteaga, I., Caicedo Zambrano, S., Hidalgo Troya, A., & Alvarado Pérez, J. (2016). *El proceso de descubrimiento de conocimiento en bases de datos*. Obtenido de <https://ediciones.ucc.edu.co/index.php/ucc/catalog/download/36/40/230?inline=1>

Tramulas, J. (1997). Los sistemas de información: una reflexión sobre información, sistema y documentación. *Revista General de Información y Documentación*, 207-229.

Villanueva, Á. (2008). *título: Análisis, diseño e implementación de un DataWarehouse de soporte de decisiones para un hospital del sistema de salud público*. Obtenido de https://tesis.pucp.edu.pe/repositorio/bitstream/handle/20.500.12404/333/VILLANUEVA_%c3%81LVARO_AN%c3%81LISIS_DISE%c3%91O_E_IMPLEMENTACION_DE_UN_DATAWAREHOUSE_DE_SOPORTE_DE_DECISIONES_PARA_UN_HOSPITAL_DEL_SISTEMA_DE_SALUD_P%c3%9aBLICO.pdf?sequence=1&i

York, J. (1994). *CALITIVIDAD*. Malaga: AG Library.