



UNIVERSIDAD CATÓLICA DE CUENCA

Comunidad Educativa al Servicio del Pueblo

UNIDAD ACADÉMICA DE INFORMÁTICA, CIENCIAS DE LA COMPUTACIÓN E
INNOVACIÓN TECNOLÓGICA

CARRERA DE INGENIERÍA EN SISTEMAS DE INFORMACIÓN

**CONSTRUCCION DE UNA SOLUCION BUSINESS INTELLIGENCE EN LOS
MODULOS DE HISTORIAS CLINICAS Y ODONTOLOGICAS DEL
PROYECTO “JUNTOS POR UNA SONRISA”**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL
TÍTULO DE INGENIERO EN SISTEMAS DE INFORMACIÓN**

AUTOR: HERMAN LEONEL BARRAZUETA CALLE

JOSE LUIS LEMA ESPINOZA

DIRECTOR: ING. OLGER ANTONIO CAJAMARCA CRIOLLO MGS.

AZOGUES-ECUADOR

2023

DIOS, PATRIA, CULTURA Y DESARROLLO

Declaratoria de Autoría y Responsabilidad

Herman Leonel Barraqueta Calle portador de la cédula de ciudadanía N° **0350012191**. Declaro ser el autor de la obra: “CONSTRUCCION DE UNA SOLUCION BUSINESS INTELLIGENCE EN LOS MODULOS DE HISTORIAS CLINICAS Y ODONTOLOGICAS DEL PROYECTO “JUNTOS POR UNA SONRISA””, sobre la cual me hago responsable sobre las opiniones, versiones e ideas expresadas. Declaro que la misma ha sido elaborada respetando los derechos de propiedad intelectual de terceros y eximo a la Universidad Católica de Cuenca sobre cualquier reclamación que pudiera existir al respecto. Declaro finalmente que mi obra ha sido realizada cumpliendo con todos los requisitos legales, éticos y bioéticos de investigación, que la misma no incumple con la normativa nacional e internacional en el área específica de investigación, sobre la que también me responsabilizo y eximo a la Universidad Católica de Cuenca de toda reclamación al respecto.

Azogues, **20 de marzo de 2023**

F:



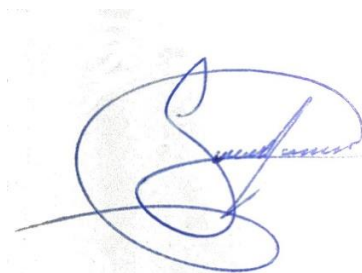
Herman Leonel Barraqueta Calle

C.I. 0350012191

Declaratoria de Autoría y Responsabilidad

José Luis Lema Espinoza portador de la cédula de ciudadanía N° **0302731120**. Declaro ser el autor de la obra: “CONSTRUCCION DE UNA SOLUCION BUSINESS INTELLIGENCE EN LOS MODULOS DE HISTORIAS CLINICAS Y ODONTOLOGICAS DEL PROYECTO “JUNTOS POR UNA SONRISA””, sobre la cual me hago responsable sobre las opiniones, versiones e ideas expresadas. Declaro que la misma ha sido elaborada respetando los derechos de propiedad intelectual de terceros y eximo a la Universidad Católica de Cuenca sobre cualquier reclamación que pudiera existir al respecto. Declaro finalmente que mi obra ha sido realizada cumpliendo con todos los requisitos legales, éticos y bioéticos de investigación, que la misma no incumple con la normativa nacional e internacional en el área específica de investigación, sobre la que también me responsabilizo y eximo a la Universidad Católica de Cuenca de toda reclamación al respecto.

Azogues, **20 de marzo de 2023**



F:

Jose Luis Lema Espinoza

C.I. 0302731120

UNIDAD ACADÉMICA DE TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN

CERTIFICACIÓN DEL DIRECTOR DE TESIS

Ing Olger Antonio Cajamarca Criollo Mgs.

DOCENTE DE LA CARRERA DE INGENIERIA EN SISTEMAS DE INFOMACION

De mi consideración:

Certifico que el presente trabajo de titulación denominado: "**CONSTRUCCION DE UNA SOLUCION BUSINESS INTELLIGENCE EN LOS MODULOS DE HISTORIAS CLINICAS Y ODONTOLÓGICAS DEL PROYECTO "JUNTOS POR UNA SONRISA"**", realizado por: **Herman Leonel Barrazueta Calle, Jose Luis Lema Espinoza**, con documentos de identidad: **0350012191, 0302731120**, previo a la obtención del título de **INGENIERO EN SISTEMAS DE INFORMACIÓN** ha sido asesorado, orientado, revisado y supervisado durante su ejecución, bajo mi tutoría en todo el proceso, por lo que certifico que el presente documento, fue desarrollado siguiendo los parámetros del método científico, se sujeta a las normas éticas de investigación que exige la Universidad Católica de Cuenca, por lo que está expedito para su presentación y sustentación ante el respectivo tribunal.

Azogues, 20 de marzo de 2023



Ing. Olger Antonio Cajamarca Criollo Mgs.
0301432035
DIRECTOR

DEDICATORIA

Quiero dedicar esta tesis a mis queridos padres, quienes han sido mi guía y apoyo incondicional en todo momento. Gracias por ser mi inspiración, por enseñarme los valores más importantes en la vida y por siempre estar ahí para mí, incluso en los momentos más difíciles.

Mamá y papá, su amor, dedicación y sacrificio han hecho posible que hoy celebre este importante logro en mi vida. Desde que era niño, ustedes me han enseñado la importancia del esfuerzo y la perseverancia, y gracias a eso, he podido alcanzar este objetivo académico.

No tengo palabras para expresar lo agradecido que estoy por todo lo que han hecho por mí. Su amor incondicional y su constante apoyo me han dado la fuerza necesaria para seguir adelante, incluso cuando las cosas parecían difíciles. Gracias por creer en mí, por motivarme y por ser mi mayor inspiración en todo momento.

Este trabajo de titulación está dedicado a ustedes, mis amados padres, como una muestra de mi amor y agradecimiento por todo lo que han hecho por mí. Espero poder devolverles todo ese amor y apoyo que me han brindado a lo largo de mi vida. Los quiero con todo mi corazón.

Herman Leonel Barraqueta Calle.

AGRADECIMIENTO

Quiero expresar mi agradecimiento a Dios quien ha sido mi fortaleza y refugio en todo momento, gracias por haberme dado la oportunidad de culminar esta etapa en mi vida y por haberme dado la sabiduría y el conocimiento necesarios para llevar a cabo este proyecto con éxito.

Expreso mi mayor agradecimiento a mis padres Herman Barrazueta Silva y Julia Calle quienes han sido mi mayor apoyo, guía y ejemplo a seguir, gracias por haberme dado la vida y por haberme brindado su amor incondicional desde el primer día. Gracias por haberme enseñado los valores más importantes de la vida, como la perseverancia, el esfuerzo y la dedicación, y por haberme apoyado en cada uno de mis proyectos y metas. Este logro es un reflejo de su amor y su dedicación en mi vida, y sé que, sin su ayuda, no estaría aquí celebrando este importante logro en mi vida.

Agradezco a mis hermanos Alexander y Cinthya, gracias por estar siempre a mi lado, por brindarme su tiempo, sus consejos y su apoyo incondicional en todo momento. Por haber sido mis cómplices, mis amigos y mis confidentes en cada etapa de mi vida. Su amor y cariño han sido una de las mayores bendiciones en mi vida, y estoy muy agradecido por tenerlos como hermanos.

Además, quiero agradecer a mi compañero y amigo José Lema quien ha sido pieza fundamental en este proceso, quiero agradecerle por su amistad, su apoyo constante, dedicación y compromiso nada de esto hubiera sido posible sin su aportes y esfuerzo. Su amistad ha sido importante en este proceso y espero poder seguir contando con su apoyo en los nuevos desafíos que se presenten en el futuro.

Finalmente Agradezco al Ing. Antonio Cajamarca quien ha sido un guía invaluable en la realización de este proyecto. Gracias a su orientación, su apoyo y su sabiduría, he podido avanzar en este camino y alcanzar los objetivos que me propuse. Su paciencia y dedicación en cada etapa de la tesis han sido fundamentales para obtener los resultados obtenidos.

Gracias a todos y cada uno de los mencionados. ¡Este logro es también de ustedes!

Herman Leonel Barrazueta Calle.

DEDICATORIA

Esta tesis es el resultado de años de esfuerzo, dedicación y perseverancia. Pero nada de esto habría sido posible sin el amor, el apoyo y el aliento incondicional de mis padres. Desde el primer día, ustedes han sido mi fuente de inspiración y motivación, me han enseñado los valores de la honestidad, la humildad y el trabajo duro. Gracias por haber estado a mi lado durante cada paso del camino, por escucharme, aconsejarme y celebrar mis logros. Espero que esta tesis sea una pequeña muestra de mi gratitud y amor hacia ustedes.

Durante todo este tiempo de estudio y trabajo, ustedes han sido mi roca, mi sostén y mi motivación. Han estado ahí para escucharme cuando necesitaba desahogarme, para aconsejarme cuando no sabía qué camino tomar y para celebrar mis logros y avances. Gracias por creer en mí, incluso cuando yo mismo dudaba de mis capacidades.

Además, quiero agradecerles por haberme dado la vida, por haberme enseñado los valores y principios que me han guiado a lo largo de este camino, y por haberme apoyado en cada una de mis decisiones. Vuestra dedicación y amor incondicional han sido el motor que me ha impulsado a seguir adelante, incluso en los momentos más difíciles.

Esta tesis es, en parte, vuestra tesis también. Espero que, al leer estas líneas, sientan el orgullo que yo siento al haber llegado hasta aquí y al saber que ustedes han sido una pieza clave en este logro. Les dedico este trabajo con todo mi amor y agradecimiento por ser los mejores padres que alguien podría desear. Gracias por todo.

José Luis Lema Espinoza

AGRADECIMIENTO

Quiero agradecer de manera especial a Dios, por brindarme el conocimiento y fuerza necesarios para llegar a culminar este largo camino.

Agradezco a mis amados padres, Noemi Espinoza y José Alberto Lema por todo el apoyo, amor y sacrificio que han ofrecido a lo largo de mi vida y, en especial, durante la realización de este trabajo. Agradezco por todas las veces que han sacrificado sus propios intereses y deseos para ayudarme en este camino, por todas las veces que me han escuchado con paciencia y cariño, gracias por enseñarme los valores y principios que me han guiado en la vida, por inculcarme la importancia del esfuerzo, la perseverancia y la dedicación. Este logro no habría sido posible sin su amor, su guía y su ejemplo a seguir.

Gracias también a mi querida hermana Aurora Lema por demostrarme su amor y cariño toda la vida y en especial en los momentos que más necesite, te agradezco profundamente por ser en mi vida una gran hermana, consejera y amiga.

A mi compañero de tesis, gran amigo y apoyo incondicional, Herman Barraqueta le extiendo mis mas sinceros agradecimientos por apoyarme en las largas noches de desvelo, en los momentos difíciles y estresantes, por estar ahí cuando lo necesitaba no solo en la elaboración de este trabajo si no también por poder contar con él en las buenas y malas, le agradezco por todo y espero nuestra amistad perdure siempre.

Quiero dedicar un espacio especial en esta tesis para agradecer a mi novia Flor Méndez por todo su apoyo, amor y comprensión durante la realización de este trabajo. Su presencia ha sido una verdadera bendición en mi vida y en este proceso de investigación. Su paciencia, motivación y amor incondicional me han permitido sobrellevar los momentos difíciles, superar los obstáculos y celebrar los logros alcanzados.

Finalmente, quiero extender mis más profundos agradecimientos a nuestro tutor de tesis, Ing. Antonio Cajamarca por habernos apoyado en cada paso de este arduo proceso, por su guía y conocimientos, le agradezco mucho y le auguro muchos éxitos es su vida.

José Luis Lema Espinoza

RESUMEN

La presente investigación se centra en la implementación de una solución de Business Intelligence para el proyecto “JUNTOS POR UNA SONRISA” con el fin de proporcionar conocimiento útil para su análisis. El proceso de implementación incluye la identificación de requerimientos a través de reuniones con los directivos de medicina y odontología, la determinación del alcance, el tratamiento de datos y la construcción de la solución de BI.

Una vez identificados los requerimientos, se procede a la extracción de datos útiles para el análisis y la construcción de un almacén de datos basado en la metodología Hefesto que alimentará la solución de BI. Para ello, se llevan a cabo procesos ETL sobre el sistema informático JPUS para garantizar la integridad de los datos obtenidos del mismo.

Finalmente, se construye la solución de BI con la herramienta Microsoft Power BI, la cual proporcionará una visión completa del proyecto presentando dashboards dinámicos para los módulos de odontología y medicina, permitiendo la identificación de oportunidades y áreas de mejora mediante el uso de diferentes tipos de gráficos según los requerimientos obtenidos.

La implementación exitosa de una solución de BI requiere un enfoque integral que incluya la fundamentación teórica, la identificación de metodologías y herramientas adecuadas, la ejecución de procesos técnicos y la implementación de soluciones personalizadas para las necesidades específicas del proyecto. La solución de BI ayudará a los responsables de vinculación a tomar mejores decisiones y aumentará la eficiencia del proyecto al permitir un análisis más detallado de los datos médicos y odontológicos.

Palabras clave: Hefesto, inteligencia de negocio, ETL, Power BI, almacén de datos

ABSTRACT

This research focuses on implementing a Business Intelligence (BI) solution for the “TOGETHER FOR A SMILE” project to provide valuable insights for analysis. The implementation process includes identifying requirements through meetings with medical and dental managers, scoping, data processing, and building the BI solution.

Once the requirements were identified, data extraction was conducted for analysis, and a data warehouse was developed based on the Hephaestus methodology to feed the BI solution. For this purpose, ETL processes are performed in the JPUS computer system to guarantee the integrity of its data.

Finally, the BI solution is made with the Microsoft Power BI tool, which will provide a complete view of the project by presenting dynamic dashboards for the dentistry and medicine modules, making it possible to identify opportunities and areas for improvement through the use of different types of graphics according to the requirements obtained.

Successful implementation of a BI solution requires a comprehensive approach that includes theoretical background, identification of appropriate methodologies and tools, execution of technical processes, and implementation of customized solutions for the project's specific needs. The BI solution will help linkage managers make better decisions and increase project efficiency by enabling more detailed medical and dental data analysis.

Keywords: Hephaestus, business intelligence, ETL, Power BI, data warehouse

ÍNDICE DE CONTENIDO

1. CAPÍTULO 1.....	16
1.1. INTRODUCCION	16
1.2. ANTECEDENTES	16
1.3. DESCRIPCIÓN DEL PROBLEMA.....	17
1.4. UBICACIÓN DEL PROBLEMA.....	17
1.5. DELIMITACIÓN DEL PROBLEMA	18
1.6. FORMULACIÓN DEL PROBLEMA	18
1.7. OBJETIVO GENERAL	18
1.8. OBJETIVOS ESPECÍFICOS	18
1.9. JUSTIFICACIÓN	19
1.10. ALCANCE	19
1.11. ESTADO DEL ARTE	20
2. CAPÍTULO 2 FUNDAMENTACIÓN TEÓRICA.....	25
2.1. DATOS.....	25
2.2. INFORMACION.....	25
2.3. CONOCIMIENTO	25
2.4. INDICADORES KPI.....	25
2.5. SISTEMAS DE INFORMACIÓN	26
2.5.1. OLAP	26
2.5.2. ROLAP	27
2.5.3. MOLAP	27
2.5.4. HOLAP	27
2.5.5. CUBOS MULTIDIMENSIONALES	27
2.5.5.1. Elementos de un cubo multidimensional	28
2.5.6. TIPOS DE MODELOS MULTIDIMENSIONALES.....	28
2.5.6.1. Esquema estrella	28
2.5.6.2. Esquema copo de nieve	28
2.5.6.3. Esquema de constelación	28
2.6. DATAWAREHOUSE	29
2.6.1. PROCESOS ETL.....	29
2.6.1.1. Extracción	29
2.6.1.2. Transformación	30
2.6.1.3. Carga	30
2.6.1.4. Como funciona una ETL.....	30
2.6.2. DATA MART	31
2.6.3. MINERIA DE DATOS.....	31
2.6.4. BASE DE DATOS RELACIONAL	31
2.7. INTELIGENCIA DE NEGOCIOS	32
2.8. CUADRO DE MANDO INTEGRAL.....	32
2.9. ANALISIS COMPARATIVO DE HERRAMIENTAS DE BI.....	33
2.9.1. POWER BI.....	33
2.9.2. TABLEAU	34
2.9.3. ORACLE ANALYTICS CLOUD	34
2.9.4. CUADRO COMPARATIVO DE HERRAMIENTAS BI.....	35
2.10. ANALISIS COMPARATIVOS DE BASES DE DATOS RELACIONALES.....	38
2.10.1. MySQL.....	38
2.10.2. PostgreSQL	38
2.10.3. Microsoft SQL Server.....	39

2.10.4. Oracle.....	39
2.10.5. CUADRO COMPARATIVO DE SISTEMAS GESTORES DE BASES DE DATOS	39
2.11. ANALISIS COMPARATIVOS DE HERRAMIENTAS PARA ETL.....	41
2.11.1. KNIME	42
2.11.2. TALEND	42
2.11.3. QLIK	42
2.11.4. SQL SERVER INTEGRATION SERVICES.....	42
2.11.5. CUADRO COMPARATIVO DE HERRAMIENTAS PARA ETL	42
2.12. ANALISIS COMPARATIVOS DE METODOLOGIAS DE DATA WAREHOUSE.....	45
2.12.1. Bill Inmon.....	45
2.12.2. Metodología Ralph Kimball	46
2.12.3. Metodología Hefesto.....	47
2.12.4. Metodología SAS Rapid Data Warehouse	49
2.12.5. Lenguaje C#	50
2.12.6. Arquitectura .NET	50
3. CAPÍTULO 3 DESARROLLO DE LA SOLUCIÓN BUSINESS INTELLIGENCE.....	51
3.1. INTRODUCCIÓN	51
3.2. MATERIALES Y METODOS	52
3.3. FASE 1: ANALISIS DE REQUERIMIENTOS.....	52
3.3.1. IDENTIFICAR PREGUNTAS DE NEGOCIO.....	52
3.3.2. MEDICINA	53
3.3.3. ODONTOLOGIA	53
3.3.4. INDICADORES Y PERSPECTIVAS	53
3.3.5. INDICADORES MEDICINA.....	53
3.3.6. PERSPECTIVAS MEDICINA.....	54
3.3.7. INDICADORES ODONTOLOGIA.....	54
3.3.8. PERSPECTIVAS ODONTOLOGIA.....	54
3.3.9. MODELO CONCEPTUAL.....	55
3.4. FASE 2: ANALISIS DE DATA SOURCES	55
3.4.1.1. HECHOS E INDICADORES.....	56
3.4.1.1.1. HECHOS E INDICADORES MEDICINA	56
3.4.1.1.2. HECHOS E INDICADORES ODONTOLOGIA	57
3.4.1.2. MAPEO	57
3.4.1.2.1. PERSPECTIVAS	58
3.4.1.2.2. INDICADORES.....	59
3.4.1.2.2.1 INDICADORES MEDICINA	59
3.4.1.2.2.2 INDICADORES ODONTOLOGIA	60
3.4.1.3. GRANULARIDAD.....	60
3.4.1.4. MODELO CONCEPTUAL AMPLIADO	61
3.5. FASE 3 MODELO LOGICO DEL DW.....	62
3.5.1.1. TIPOLOGIA.....	62
3.5.1.2. TABLA DE DIMENSIONES	62
3.5.1.3. TABLA DE HECHOS.....	64
3.5.1.4. UNIONES.....	66
3.6. FASE 4 INTEGRACION DE DATOS.....	67
3.6.1.1. CARGA INICIAL.....	67
3.6.1.2. ACTUALIZACION.....	68
3.7. CUBO MULTIDIMENSIONAL.....	72
3.7.1.1. CONEXIÓN DE BASE DE DATOS DE ORIGEN.....	73

3.7.1.2. CONEXIÓN DE BASE DE DATOS DEL DATA WARE HOUSE	74
3.7.1.3. LIMPIEZA DEL DATA WAREHOUSE	75
3.7.1.4. CARGA DE LA DIMENSION PERSONA.....	75
3.7.1.5. CARGA DE LA DIMENSION UBICACIÓN.....	76
3.7.1.6. CARGA DE LA DIMENSION TIEMPO MEDICINA	76
3.7.1.7. CARGA DE LA DIMENSION TIEMPO ODONTOLOGIA	77
3.7.1.8. CARGA DE PERSONAS EN LA TABLA DE HECHOS DE MEDICINA	77
3.7.1.9. CARGA DE UBICACIONES EN LA TABLA DE HECHOS DE MEDICINA..	77
3.7.1.10. CARGA DE TIEMPO EN HECHOS DE MEDICINA.....	78
3.7.1.11. CARGA DE PESO EN HECHOS DE MEDICINA.....	78
3.7.1.12. CARGA DE DURACION DE EMBARAZO EN HECHOS MEDICINA.....	79
3.7.1.13. CARGA DE TOMÓ SENO Y TOMÓ FORMULA EN HECHOS MEDICINA	79
3.7.1.14. CARGA DE TIPO DE ALIMENTACION EN HECHOS MEDICINA.....	80
3.7.1.15. CARGA DE MALNUTRICION EN HECHOS MEDICINA	80
3.7.1.16. CARGA DE PERSONAS EN HECHOS ODONTOLOGIA.....	80
3.7.1.17. CARGA DE UBICACIÓN EN HECHOS ODONTOLOGIA.....	81
3.7.1.18. CARGA DE TIEMPO EN HECHOS ODONTOLOGIA.....	81
3.7.1.19. CARGA DE DIAGNOSTICO EN HECHOS ODONTOLOGIA.....	82
3.7.1.20. CARGA DE TRATAMIENTO EN HECHOS ODONTOLOGIA.....	82
3.7.1.21. CARGA DE CPO EN HECHOS ODONTOLOGIA.....	82
3.7.1.22. CARGA DE PROBLEMAS BUCALES EN HECHOS ODONTOLOGIA.....	83
3.7.1.23. MODELO FINAL DEL ETL	83
3.7.2. DISEÑO DEL BI.....	84
3.7.2.1. SOLUCIÓN	85
3.7.2.1.1. INFORME MEDICINA:.....	87
3.7.2.1.1.1 TARJETA TOTAL PERSONAS ATENDIDOS.....	87
3.7.2.1.1.2 TARJETA TOTAL PACIENTES ATENDIDOS MEDICINA.	87
3.7.2.1.1.3 TARJETA PROM EDADES.	88
3.7.2.1.1.4 PERSONAS ATENDIDAS POR UBICACIÓN.....	88
3.7.2.1.1.5 TASA DE PESO	88
3.7.2.1.1.6 MALNUTRICIÓN	89
3.7.2.1.1.7 MALNUTRICIÓN POR UBICACIÓN.....	90
3.7.2.1.1.8 ALIMENTOS CONSUMIDOS	91
3.7.2.1.1.9 CONSUMO DE LECHE MATERNA Y FORMULA	92
3.7.2.1.1.10 NEONATOS PREMATUROS	92
3.7.2.1.2. INFORME ODONTOLOGIA:.....	93
3.7.2.1.2.1 PACIENTES ATENDIDOS ODONTOLOGIA	93
3.7.2.1.2.2 PROMEDIO DE EDADES	93
3.7.2.1.2.3 PACIENTES ATENDIDOS POR CANTON Y AÑO.....	94
3.7.2.1.2.4 PERSONAS ATENDIDAS POR UBICACIÓN.....	95
3.7.2.1.2.5 AREAS DE TRATAMIENTOS	95
3.7.2.1.2.6 TRATAMIENTOS POR CANTON.....	96
3.7.2.1.2.7 DIAGNOSTICOS COMUNES POR UBICACIÓN.....	97
3.7.2.1.2.8 PROMEDIO CPO POR UBICACIÓN Y SEXO.....	97
3.7.2.1.2.9 SALUD BUCAL	99
CONCLUSIONES	99
RECOMENDACIONES.....	100

BIBLIOGRAFÍA:.....	101
ANEXOS	104

LISTA DE FIGURAS

Figura 1. Estructura OLAP.....	26
Figura 2. Proceso ETL.....	30
Figura 3. Flujo para la generación del conocimiento.....	32
Figura 4. Magic Quadrant for Analytics and Business Intelligence Platforms	38
Figura 5 Arquitectura de un DW según Bill Inmon	46
Figura 6. Ciclo de vida, metodología Kimball	47
Figura 7. Ciclo de vida, metodología Hefesto	48
Figura n 8 Fases de la metodología SAS.....	49
Figura 9: Modelo conceptual.....	55
Figura 10. MODELO CONCEPTUAL BASE DE DATOS OPERATIVA	58
Figura 11 modelo conceptual ampliado	61
Figura 12 Dimensión Persona	63
Figura 13 Dimensión Ubicación	63
Figura 14 Dimensión Tiempo Medicina.....	64
Figura 15 Hechos Odontología	66
Figura 16 Modelo Logico del Dw	66
Figura 17 Proceso ETL principal.....	67
figura 18. Archivo .Bat para la actualización del proceso ETL.....	68
Figura 19. Tarea programada dentro del servidor.....	68
Figura 20. Actualización Dashboard de Medicina.	69
Ilustración 21. Actualización Dashboard de Odontología.....	69
figura 22. SQL Executor para vaciar el data warehouse.	70
Figura 23. Carga de datos en las dimensiones del data warehouse	71
Figura 24. Carga de datos en la tabla de hechos de medicina.....	71
Figura 25. Carga de datos en la tabla de Hechos de Odontología.....	72
Figura 26 Cubo multidimensional de medicina.....	72
Figura 27 cubo multidimensional de odontología	73
Figura 28 Obtención de data source.....	74
Figura 29 Conexión de data warehouse	75
Figura 30 vaciado de base de datos final.....	75
Figura 31 Carga de dimensión persona	76
Figura 32 Carga de dimensión Ubicación	76
Figura 33 Carga de dimensión tiempo medicina	76
Figura 34 Carga de dimensión tiempo odontología.....	77
Figura 35 Carga de personas medicina en hechos medicina.....	77
Figura 36 Carga de ubicación en hechos medicina	78
Figura 37 Carga de tiempo en hechos medicina.....	78
Figura 38 Carga de peso en hechos medicina.....	79
Figura 39 Carga de duración de embarazo en hechos medicina	79
Figura 40 tomo seno y tomo formula en hechos medicina	79
Figura 41 Carga de tipo de alimentación en hechos medicina	80
Figura 42 Carga de mal nutrición en hechos medicina	80
Figura 43 Carga de personas en hechos odontología.....	81
Figura 44 Carga de ubicación en hechos odontología	81
Figura 45 Carga de tiempo en hechos odontología	82
Figura 46 Carga de diagnóstico en hechos medicina.....	82
Figura 47 Carga de tratamiento en hechos odontología	82
Figura 48 Carga de CPO en hechos odontología	83

Figura 49 Carga de problemas bucales en hechos odontología	83
Figura 50 Modelo ETL final.....	84
Figura 51 Conexión de Power BI a MSSQL.....	85
Figura 52. Carga de datos medicina a Power BI.....	86
Figura 53. Carga de datos medicina a Power BI.....	86
Figura 54. Total, Personas atendidos.	87
Figura 55. Total, Pacientes atendidos medicina.....	87
Figura 56. Prom edades.	88
Figura 57. Personas atendidas por ubicación.	88
Figura 58. Tasa de peso.....	89
Figura 59. MALNUTRICIÓN	90
Figura 60. Malnutrición por ubicación	90
Figura 61. Alimentos consumidos.....	91
Figura 62. Consumo de leche materna y formula.	92
Figura 63. Neonatos prematuros	92
Figura 64 Pacientes atendidos en odontología	93
Figura 65 Promedio de edades en los pacientes	93
Figura 66 Matriz de pacientes por cantón y año	94
Figura 67 Treemap de personas atendidas por parroquia	95
Figura 68 Grafico de tratamientos.....	95
Figura 69 Mapa coroplético de tratamientos por cantón	97
Figura 70 Grafico de barras de diagnósticos comunes por ubicación	97
Figura 71 Grafico de cintas de promedio de CPO por ubicación y sexo	98
Figura 72 Grafico de columnas de salud bucal	99

LISTA DE TABLAS

TABLA 1. COMPARATIVA DE HERRAMIENTAS BI	36
TABLA 2. CUADRO COMPARATIVO DE SISTEMAS GESTORES DE BASES DE DATOS	40
TABLA 3. CUADRO COMPARATIVO DE HERRAMIENTAS PARA ETL.....	44
TABLA 4. METHODOLOGY SAS RAPID DATA WAREHOUSE	50

LISTA DE ANEXOS

ANEXOS	104
ANEXO A NODOS DEL MODELO ETL EN KNIME	105
ANEXO B DASHBOARDS DE MEDICINA Y ODONTOLOGIA EN POWER BI.....	116
ANEXO C ACTUALIZACION DE ETL A TRAVES DE COMANDOS BASH	121
ANEXO D ACTUALIZACION DE ORIGEN DE DATOS EN POWER BI.....	121
ANEXO E PUBLICACION DE DASHBOARDS EN LA PAGINA DE JPUS.....	123

1. CAPÍTULO 1

1.1. INTRODUCCION

La gestión de datos y la toma de decisiones basada en información se han convertido en aspectos críticos para la supervivencia y el éxito de las organizaciones en la actualidad. En este contexto, Business Intelligence (BI) emerge como una solución para transformar los datos en conocimiento y mejorar el proceso de toma de decisiones. El proyecto “JUNTOS POR UNA SONRISA” (JPUS), no es la excepción, y se enfrenta a la necesidad de contar con una solución Business Intelligence que le permita obtener un conocimiento lucrativo que pueda ser útil para el mismo.

Por lo tanto, el propósito del presente estudio tiene como objetivo investigar y analizar las principales herramientas y técnicas de BI que se utilizan actualmente en el entorno, así como evaluar su efectividad en la mejora de la toma de decisiones y la gestión de datos en las organizaciones y/o proyectos. Además, se pretende analizar las datos y situaciones en las que han implementado soluciones de BI y evaluar los resultados obtenidos a través de su experiencia.

Para lograr esto, se realizó levantamiento de requerimientos a través de reuniones virtuales y presenciales con los distintos directivos de medicina y odontología. De esta manera, se detalló y depurado las distintas preguntas del modelo de negocio planteadas. Posteriormente, se determinará el alcance de la solución, extrayendo los requisitos funcionales y especificando el origen de los datos a analizar, la arquitectura de la solución BI, el modelado de datos OLAP, el modelado de hechos y dimensiones, entre otros.

Se pretende que los resultados de esta solución puedan contribuir a la comprensión del valor de las soluciones de Business Intelligence en la toma de decisiones dentro de proyectos universitarios, y proporcionar recomendaciones útiles para su implementación.

1.2. ANTECEDENTES

En la provincia del Cañar, en la ciudad de Azogues se ejecuta el proyecto pluridisciplinario “Juntos por una sonrisa”, el mismo que tiene como objetivo general la promoción y prevención de riesgos nutricionales en niños de 0 a 12 años de edad en la provincia realizando triadas a las distintas parroquias y cantones, prestando sus servicios de

medicina general y odontológicos para el público en general de manera gratuita, determinando la funcionalidad familiar y las competencias parentales entre otros procesos, este proceso se organiza con la dirección de los tutores de vinculación con la sociedad de la Universidad Católica de Cuenca junto con la Prefectura del Cañar y pretende contribuir a la disminución de los riesgos de enfermedades relacionadas al estado de nutricional de los pacientes.

El registro de fichas médicas y odontológicas de pacientes atendidos por los participantes del proyecto se realizaba a través del llenado manual de hojas con la información del paciente, para cada uno de los casos. Actualmente “JUNTOS POR UNA SONRISA” mantiene un sistema informático que ayuda al llenado de sus historias clínicas mediante dicho software que se encuentra alojado en un servidor de la Universidad Católica de Cuenca campus Azogues, que almacena su información en una base de datos relacional para su mejor funcionamiento.

1.3. DESCRIPCIÓN DEL PROBLEMA

En la actualidad, las organizaciones han tratado de adquirir valor empleando diversas estrategias administrativas para considerarse más competitivas. Es así que, la Facultad de Medicina y Odontología de la Universidad Católica de Cuenca, a través del proyecto de vinculación “JUNTOS POR UNA SONRISA”, por medio de la carrera de ingeniería en sistemas de la información pudieron automatizar las fichas y registros de las historias clínicas de sus pacientes que anteriormente lo hacían de forma física y actualmente lo hacen en un sistema informático, sin embargo la información vertida en el software no es de utilidad para los administrativos debido a que los registros solo son visualizados de manera global y estática por lo que no aportan valor analítico para los directivos.; además, se considera necesario para futuras investigaciones que puedan realizar los administrativos del proyecto, cubriendo la necesidad de obtener y compartir información, que pueda ser fácilmente analizada, permitiendo ahorrar tiempo y recursos para el análisis de datos y toma de decisiones.

1.4. UBICACIÓN DEL PROBLEMA

El proyecto "Juntos por una sonrisa", una iniciativa conjunta entre la Universidad Católica de Cuenca, campus Azogues y la Prefectura del Cañar, se ha enfrentado a un desafío importante: la falta de cuadros de mando integral para el acceso de los directivos de vinculación que manejen información sobre problemas de malnutrición y salud bucal. Es necesario encontrar una solución efectiva para desarrollar una herramienta de inteligencia de

negocios que permita una visión más amplia y detallada de la situación actual con alcance en los diferentes cantones y parroquias de la provincia del Cañar, con el fin de tomar medidas preventivas y proactivas para prevenir los diferentes problemas de salud.

1.5. DELIMITACIÓN DEL PROBLEMA

El proyecto “Juntos por una sonrisa” apoyará a la Universidad Católica de Cuenca a través de un sistema informático que aporta los datos para una solución de inteligencia de negocios que permitirá el análisis de datos, control de desempeño, generación de indicadores y la correcta toma de decisiones, que facilitará así el estudio requerido por los responsables del sistema, en función a sus necesidades.

1.6. FORMULACIÓN DEL PROBLEMA

La información existente en el servidor de la base de datos del sistema “JPUS” en su versión anterior no permite realizar un estudio estadístico adecuado, ni aportan a la toma de decisiones, siendo así se presenta la siguiente interrogante: ¿Cómo puede ayudar la construcción de una solución de Business Intelligence al proyecto “JUNTOS POR UNA SONRISA” en los módulos de medicina y odontología?

1.7. OBJETIVO GENERAL

Implementar una solución de Business Intelligence para el sistema informático "Juntos por una sonrisa", que permita la generación de indicadores, control de desempeño, mejorar el análisis de datos y la toma de decisiones en la gestión de información médica y odontológica.

1.8. OBJETIVOS ESPECÍFICOS

1. Fundamentar teóricamente los conceptos de Business Intelligence, consultando fuentes de datos bibliográficos.
2. Identificar la metodología y utilizar herramientas para el soporte de decisiones basadas en un Data Warehouse.
3. Ejecutar el proceso ETL de la base de datos relacional del software “JUNTOS POR UNA SONRISA” para migrar hacia el almacén de datos.
4. Implementar la solución de Business Intelligence en el software “JUNTOS POR UNA SONRISA”.

1.9. JUSTIFICACIÓN

Durante el transcurso en dos años de ejecución del proyecto “JUNTOS POR UNA SONRISA” se ha identificado la necesidad de proveer de atención médica y odontológica para la atención a pacientes de recursos económicos desfavorables; es por esto que, las visitas in situ se planifican sin el conocimiento adecuado sobre los puntos geográficos vulnerables que necesiten más atención. Con un mejor análisis de la información sobre enfermedades, patologías, y padecimientos comunes que se identifiquen, resultado de las estrategias generadas por los mandos directivos de las instituciones involucradas en el proyecto, se analizarán datos relacionados a los ámbitos sociales, económicos y de salubridad.

El sistema informático actualmente no cuenta con un medio que ayude en el análisis de datos para la generación de indicadores, control de desempeño y cuadros de mando integral dinámicos, es por eso que mediante la creación de una solución BI, ayudaremos a interpretar de mejor manera la información registrada en el sistema informático, haciendo que los mandos directivos tengan una imagen clara de los datos almacenados en las bases de datos operativas para convertirlos en información que genere conocimiento y permita tomar mejores decisiones mediante el empleo de metodologías, aplicaciones y tecnologías a través de técnicas analíticas de extracción, brindando así una visión más amplia sobre las triadas clínicas, dando paso a investigaciones sobre mal nutrición y enfermedades bucales, misma que aporta a la mejora de calidad de vida de los habitantes de la provincia del Cañar, donde se ha identificado estos problemas en diversos sectores de la población, además de ahorrar tiempo y destinar los esfuerzos de manera más eficiente a las distintas ubicaciones de la provincia dependiendo de aquellas que requieran una atención más urgente, esto se verán reflejados en la generación de informes, actuando en beneficio de los directivos de las instituciones benefactoras, por lo que es importante el uso de herramientas de gestión de datos para realizar una mejor seguimiento de los pacientes afectados y optimizar el uso de los recursos disponibles para el tratamiento de estas enfermedades.

1.10. ALCANCE

El proyecto incluye las fases de análisis, diseño e implementación de una solución de Business Intelligence a través de un Dashboard dinámico para la toma de decisiones, análisis de datos médicos y odontológicos del sistema informático. Previo a la construcción del almacén de datos y el proceso ETL, se realizará una comparativa entre distintas metodologías y herramientas, posterior al análisis y selección, se creará el modelo respectivo para ejecutar

el proceso de extracción, transformación y carga en el Data Warehouse, que estará alojado en un servidor perteneciente a la Universidad Católica de Cuenca.

A continuación, se creará un cubo multidimensional, se aplicarán los procesos de BI con sus respectivas herramientas y de esta forma se podrán generar indicadores, controles de desempeño, informes, cuadros de mando integral dinámicos, etc. El objetivo final del proyecto es utilizar los resultados obtenidos a través del análisis de datos médicos y odontológicos para proporcionar información valiosa que pueda ser utilizada en la toma de decisiones. Para lograr este objetivo, se realizará una evaluación exhaustiva de los datos recopilados. Mismos que se utilizarán para informar a los responsables de la planificación y gestión de políticas de salud sobre las medidas necesarias para mejorar la atención médica y odontológica.

Este proceso será desarrollado en dos módulos independientes (medicina y odontología) debido a los distintos requerimientos que presentan estas dos ramas de la salud y bienestar, proceso que se realizó mediante un conector JDBC hacia la base de datos operativa, permitiendo extraer la información necesaria para el almacén de datos con los distintos niveles de granularidad, como por ejemplo en la dimensión tiempo se usa año, mes y día, dando paso a una perspectiva más amplia en función de lo que se requiera visualizar en los distintos dashboards del software de Power BI esto permitirá obtener mejores y más precisos resultados, además de futuras proyecciones basadas en el conocimiento obtenido en la solución implementada.

1.11. ESTADO DEL ARTE

En el año 2020, en el Hospital Regional Ernesto Torres Galdames, de la ciudad de Iquique, en la región de Tarapacá-Chile, se realizó la siguiente investigación, titulada: “Elementos de Inteligencia de Negocios para el Apoyo de Toma de Decisiones en la Unidad De Docencia Del Hospital Regional Ernesto Torres Galdames”, en el cual se implementó una solución en base al uso de herramientas de Inteligencia de Negocios. Se consideró un análisis de los procesos de la Unidad de Docencia, para conocer su modelo de procesos, realizar la selección de las herramientas implementar, establecer los procesos de Extracción, Transformación y Carga, y la presentación de los datos reportes. En base a la Metodología de Kimball se construyó un Data Mart, cuyo uso permitió a la Unidad de Docencia observar tendencias, cantidades de estudiantes registrados y, analizar Y coordinar de mejor forma los cupos de prácticas profesionales [1].

En el año 2017, en la universidad técnica del norte de la ciudad de Ibarra-Ecuador se realizó la investigación denominada: “INTELIGENCIA DE NEGOCIOS APLICADO A INDICADORES CLAVES DE DESEMPEÑO (KPI’s) PARA APOYAR LA TOMA DE DECISIONES EN LA GESTIÓN DEL CENTRO DE DISTRIBUCIÓN DE FARMAENLACE CÍA. LTDA.”. En este sentido, este proyecto se presentó como una alternativa con la cual se podrá atender los requerimientos de información más importantes del Centro de Distribución y permitirá tener al alcance los indicadores claves de desempeño (KPI) necesarios para el análisis situacional y de apoyo para la toma de decisiones. Para efectos de la investigación se realizaron revisiones bibliográficas de los principales conceptos de Inteligencia de Negocios e Indicadores Claves de desempeño. Para conocer las necesidades referentes a indicadores de desempeño y la obtención de información se realizaron entrevistas a los responsables de las áreas del CEDIS, este acercamiento permitió orientar el desarrollo del proyecto acorde a los requerimientos de los usuarios. La implementación de la solución de Inteligencia de Negocios permitió transformar los datos generados en el CEDIS en conocimiento útil para dar soporte a la toma de decisiones [2].

En el año 2019 se presentó una investigación llamada “Propuesta de una Solución de Business Intelligence para el Monitoreo y Control de Gestantes de Alto Riesgo en el Hospital Vitarte” que consiste en desarrollar un estudio de los principales factores asociados a morbilidad materna que determinan la consideración de una paciente como gestante de alto riesgo. Con esta información se logró proponer una solución tecnológica e indicadores de gestión que permiten monitorear y controlar este grupo de pacientes que son atendidas en el Hospital Vitarte de Lima, Perú. La investigación de esta tesis se basa principalmente en la propuesta del modelo de Molina et al. (1998) sobre los factores de riesgo presentes durante la etapa de embarazo, en el momento del parto y puerperio en madres adolescentes, así como investigaciones publicadas de la OMS, OPS, SCIELO, PUBMED, entre otros. Los resultados obtenidos permitieron proponer una solución de Business Intelligence que ayude al monitoreo y control de las gestantes de alto riesgo que presentan estos factores asociados a morbilidad, mediante la elaboración de un Dashboard de gestión a ser utilizado por los tomadores de decisiones del Hospital Vitarte. El Dashboard permitirá obtener información histórica y en tiempo real de las gestantes de alto riesgo, así como la evolución de los casos que presentaron morbilidad de las gestantes que tuvieron los factores de alto riesgo previamente indicados. Esta herramienta tecnológica ayudará a tomar decisiones como la realización de campañas

médicas en zonas específicas, así como monitorear a cada una de las gestantes de alto riesgo según el factor de riesgo que presente[3].

En el año 2019, en la escuela politécnica nacional en Quito-Ecuador, se desarrolló la tesis denominada: “Desarrollo de un sistema web de Inteligencia de Negocios para el manejo de indicadores socioeconómicos en el área de trabajo social del Cuerpo de Bomberos de Quito”. En tal virtud este trabajo establece la implementación de un software de Inteligencia de Negocios, conformada por el desarrollo de una aplicación llamada ficha digital, aplicando la metodología SCRUM, que, a través de cuatro formularios, permitirán la recolección de información personal de los funcionarios que laboran en el CBDMQ. Se obtiene de esta manera una fuente de información que se almacenará en el Sistema Gestor de Base de Datos (SGBD) PostgreSQL, para su posterior procesamiento mediante la creación de procedimientos almacenados y vistas. Consecutivamente, se procede con el desarrollo de una aplicación de Inteligencia de Negocios a través de dos metodologías CRISP-DM y Kimball. Se comienza con la creación de un data Warehouse, almacenando la información extraída del SGBD de dos bases de datos distintas, que permita la depuración, transformación, visualización de la información. Finalmente, para la gestión de la información se realizarán diversas hojas corporativas o Dashboards a través del software QlikSense, permitiendo la visualización de una gran cantidad de información, por medio de gráficos, tablas, entre otros. El producto obtenido son reportes que facilitan la visualización y comprensión de problemas socioeconómicos de los funcionarios del CBDMQ, que serán la base para una adecuada toma de decisiones por parte del área de Trabajo Social del CBDMQ[4].

En el año 2021, en la universidad Cesar Vallejo Lima-Perú, se desarrolló la tesis denominada: “Business Intelligence para la gestión de ventas de productos odontológicos: caso VERO DENT” Esta tesis fue desarrollada con la metodología HEFESTO ya que es una metodología que se acomoda a las necesidades y etapas del proyecto, también es rápida en tiempos de entrega y adaptabilidad a cambio de los usuarios, para el desarrollo visual y base de datos SQL Server y la interfaz de usuario Power BI a fin de automatizar las tareas diarias del personal que se encarga del proceso. La población se determinó a 40 documentos de ventas, agrupadas en 26 fichas de registro. Y con un tamaño de muestra de 40 documentos de ventas, estratificados por días. Por lo tanto, la muestra queda conformada en 26 fichas de registro. El tipo de muestreo es aleatorio simple y la técnica de recolección de datos fue el fichaje y el instrumento la ficha de registro. Los cuales fueron validados por expertos. La implementación del Business Intelligence permitió incrementar la rotación de stock en un mes

de 10% a 21% incrementando en un 11% y un crecimiento de ventas de 11% a 56% incrementando en un 45%. Finalmente se puede determinar que el BI cumplió con los objetivos planteados en la presente investigación, mejorando el proceso de gestión de ventas para la venta de productos odontológicos: Caso Vero Dent, ya que cuenta con toda la información relevante para el análisis de ventas [5].

En el año 2020 en Cuenca- Ecuador se realizó la investigación denominada “Business Intelligence aplicado al sector Salud” Se tuvo por objetivo analizar el Business Intelligence aplicado al sector de salud basado en una base de datos universal. La investigación fue de tipo descriptiva con diseño de campo no experimental, recopilándose información por medio de 690 funcionarios de la rama médica y odontológica, pública y privada de la ciudad de Cuenca – Ecuador. Un 59.42% piensa que es muy necesario el archivo de los datos. Normalmente en los centros de salud públicos o privados esta información tiene un almacenamiento físico, por cuanto utilizan formularios de papel para la recopilación. La importancia de crear una solución de inteligencia de negocios, permitiendo que existan alternativas de cómo modernizar la parte administrativa-tecnológica-social, de esta manera generar procesos más rápidos, ágiles, que permitan también respuestas efectivas por parte de los expertos de salud [6],

En el año 2019, en la universidad católica santo toribio de Mogrovejo se realizó el trabajo de investigación titulado como solución basada en inteligencia de negocios para apoyar a la toma de decisiones en el área de ventas de una empresa comercial de la ciudad de Chiclayo la cual tiene como objetivo apoyar en la toma de decisiones del área de ventas de una empresa comercial, de este modo hacer frente a problemáticas, como: información de baja calidad para mejorar dichas decisiones, datos duplicados, toma de decisiones de último momento, demora en la entrega de reportes, siendo estos tabulares y poco entendibles. Asimismo, el desconocimiento de las preferencias del cliente impide la realización de recomendaciones o promociones. Con esta finalidad, se planteó el desarrollo de una solución basada en Inteligencia de Negocios, utilizando el algoritmo de clustering. El tipo de investigación fue cuasi experimental debido a que la solución se aplicó en un solo caso de estudio. Se realizó una preprueba y una posprueba sin grupo control con el fin de observar los cambios presentados. El tipo de muestreo fue no probabilístico y como instrumentos de recolección de datos se hizo uso de pruebas de sistema y cuestionarios; se utilizaron técnicas como la observación, entrevista y encuesta. Se aplicó la metodología de Kimball, la que se complementó con la metodología CRISP-DM. Como resultados, se obtuvieron reportes interactivos sobre el comportamiento de las ventas, se consiguió reducir el tiempo promedio

de información a un 94.47% y se realizó la segmentación de clientes identificando 14 grupos a fin de elaborar mejores estrategias de mercado. En conclusión, se alcanzó el objetivo; los ejecutivos obtuvieron reportes interactivos sin demora alguna y con la información necesaria que consiga ayudar a la toma de decisiones, logrando así su satisfacción [7].

En el año 2020, en el país de Ecuador se realizó un estudio denominado: Business Intelligence aplicado al sector Salud, que tuvo por objetivo analizar el Business Intelligence aplicado al sector de salud basado en una base de datos universal. La investigación fue de tipo descriptiva con diseño de campo no experimental, recopilándose información por medio de 690 funcionarios de la rama médica y odontológica, pública y privada de la ciudad de Cuenca – Ecuador. Un 59.42% piensa que es muy necesario el archivo de los datos. Normalmente en los centros de salud públicos o privados esta información tiene un almacenamiento físico, por cuanto utilizan formularios de papel para la recopilación. La importancia de crear una solución de inteligencia de negocios, permitiendo que existan alternativas de cómo modernizar la parte administrativa-tecnológica-social, de esta manera generar procesos más rápidos, ágiles, que permitan también respuestas efectivas por parte de los expertos de salud [8].

En el apartado referente al estado del arte, se citan distintas problemáticas presentadas con relación a la inteligencia de negocios, donde al no tener un correcto uso de información vertida en distintas organizaciones, genera un desperdicio de recursos y tiempo. Por lo que, al aplicar una solución BI, se expande el abanico de posibilidades al brindar una información convertida en conocimiento, solucionando así una problemática general de muchas organizaciones y ayudando a la toma de decisiones [1].

CONTRIBUCIONES

En el presente trabajo de titulación, se presenta una solución Business Intelligence para los módulos de medicina y odontología del sistema del proyecto “Juntos Por Una Sonrisa”. Esta tiene como finalidad presentar un cuadro de mando dinámico dirigido al mando administrativo del proyecto ya mencionado, que contribuirá al objetivo establecido, como resultado su punto de vista médico se verá beneficiado al momento de tomar decisiones en cuanto a malnutrición y problemas bucales se refiere, pues mediante los gráficos pueden analizar la posibilidad de aumentar la atención en ciertas ubicaciones y también centrarse en los problemas de salud más reincidentes en los pacientes; presentando indicadores cruciales

como patologías comunes o lugares a los cuales se les debe brindar más atención, control de desempeño y la generación de indicadores, entre otros.

2. CAPÍTULO 2 FUNDAMENTACIÓN TEÓRICA

En el presente capítulo se expondrán las bases teóricas de las tecnologías usadas para el desarrollo del proyecto, esta información recopilada servirá de soporte para la revisión y entendimiento de: Inteligencia de negocios, análisis de datos, Data Warehouse, bases de datos relacionales y el empleo de herramientas específicas para soluciones BI (Power BI, Tableau, Oracle BI, Domo, etc.)

2.1. DATOS

Según [2] Datos hace referencia a: “el elemento constitutivo del conocimiento. Comprenden hechos, representaciones o los mecanismos por los cuales nos es posible medir e identificar algún aspecto de nuestro mundo.”

2.2. INFORMACION

Según [2] información es: “es una interpretación de los datos basada en un cambio de las condiciones y en el paso del tiempo.”

2.3. CONOCIMIENTO

[2] Define al conocimiento como: “información organizada dentro de un marco conceptual como lo puede ser: una visión del mundo, un concepto, un principio, una teoría o cualquier otra base de la necesaria abstracción conceptual que nos permite comprender nuestro entorno, mejorar la capacidad para resolver problemas y tomar decisiones.”

2.4. INDICADORES KPI

[3] Menciona que los KPI son: “indicadores claves de rendimiento, los cuales ayudan a conocer el desempeño de un proceso, o persona, dentro de las empresas, estos indicadores son aplicados para conocer si sus procesos son adecuados, al igual que eficientes, para poder cumplir su objetivo.”

2.5. SISTEMAS DE INFORMACIÓN

2.5.1. OLAP

Según [4] OLAP es: “es una solución utilizada en el campo de la inteligencia empresarial (o Business Intelligence) cuyo objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras multidimensionales (o cubos OLAP) que contiene información resumida de grandes bases de datos.”

Se usa en informes de negocios de ventas, marketing, informes de dirección, minería de datos y áreas similares. Sin embargo, OLAP tiene ciertas limitaciones: Requiere un esquema estrella o copo de nieve para la reestructuración de datos

- Tiene un número limitado de dimensiones
- Difícil acceso a datos transaccionales en el cubo OLAP
- La actualización de OLAP requiere cambios en todo el cubo

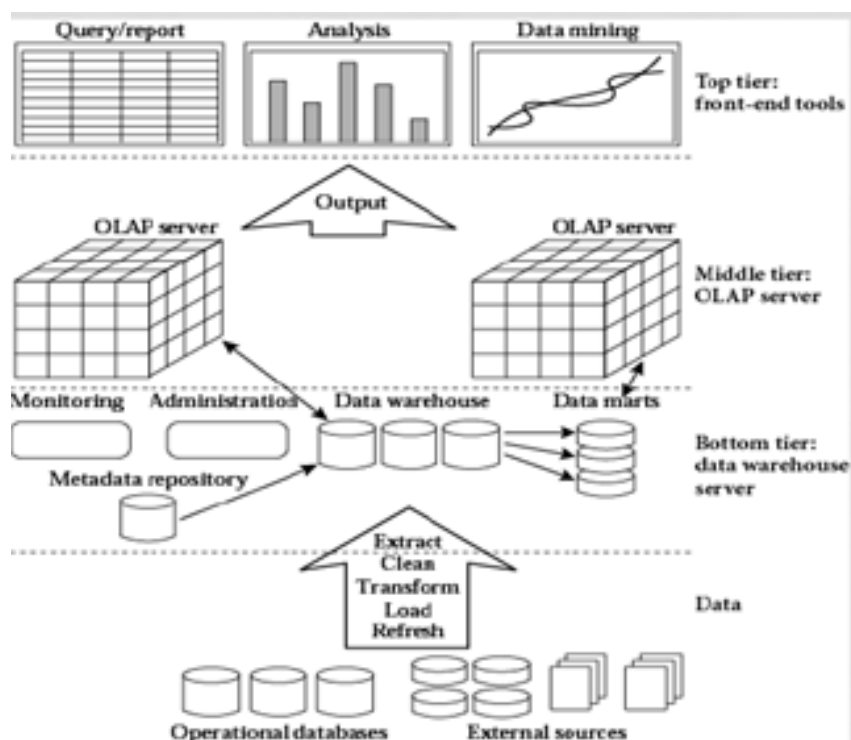


Figura 1. Estructura OLAP

Fuente: <https://www.reci.org.mx/index.php/reci/article/view/40/176>

Existe también un desglose de subcategorías de cubos multidimensionales basados en OLAP:

2.5.2. ROLAP

Según [5] ROLAP (procesamiento analítico en línea relacional) consiste en: “almacenar datos en columnas y filas, y recuperarlos como información a pedido a través de consultas enviadas por el usuario.”

ROLAP permite acceder a través de consultas SQL complejas para calcular información, manejando grandes volúmenes de datos, pero cuanto más grandes sean los datos, más lentos serán los tiempos de procesamiento. Debido a que estas consultas se llevan a cabo bajo demanda, no se requiere de almacenamiento y cálculo previo, sin embargo, su más grande limitación yace en el rendimiento y escalabilidad resultantes de grandes datos almacenados en tablas.

2.5.3. MOLAP

Según [5] MOLAP (procesamiento analítico en línea multidimensional) consiste en: “un cubo multidimensional que accede a los datos almacenados a través de varias combinaciones. Los datos se calculan previamente, se resumen previamente y se almacenan.”

Este enfoque hace uso de varios pequeños cubos previamente calculados que en conjunto forman un hipercono, lo cual a diferencia de OLAP permite usar varios cubos, además de agilizar el procesamiento de datos.

2.5.4. HOLAP

Según [5] HOLAP (procesamiento analítico en línea híbrido): “HOLAP conecta atributos de MOLAP y ROLAP. Dado que HOLAP implica almacenar parte de sus datos en un almacén ROLAP y otra parte en un almacén MOLAP, los desarrolladores obtienen los beneficios de ambos.”

Es decir, la información se almacena tanto en bases de datos relacionales como multidimensionales, lo que permite escoger la forma de acceder a dicha información según que aplicación OLAP sea la más adecuada para el tratamiento solicitado, además permite más flexibilidad para el manejo de datos.

2.5.5. CUBOS MULTIDIMENSIONALES

Según la definición citada por [6] “Un Cubo OLAP es una base de datos multidimensional orientada al proceso analítico, que se utiliza para hacer análisis sobre

grandes cantidades de información”. Un cubo multidimensional es aquel que manipula grandes cantidades de información, los datos se clasifican según las necesidades de la organización para obtener o visualizar la información esto les podrá servir para toma de decisiones dentro de la organización.

2.5.5.1. Elementos de un cubo multidimensional

- Base de Datos. – Alojamiento de la información de la organización
- Dimensiones. -Incorporan los objetos en los cuales una empresa pretende conservar sus datos constituidos
 - Hechos. - . Están conformados por las medidas, que establecen las dimensiones. Cada hecho individual está coligado a un segmento de cada dimensión[6].
 - Medidas. – Son los atributos medibles numéricamente que se encuentran ligados a los hechos que son lo que realmente se miden[6].

2.5.6. TIPOS DE MODELOS MULTIDIMENSIONALES

2.5.6.1. Esquema estrella

Este esquema consta de una tabla de hechos y dos o más tablas de dimensiones que se relacionan cada uno de ellos con la tabla de hechos definida dentro del esquema.

2.5.6.2. Esquema copo de nieve

Este esquema cuenta con las mismas características mencionadas en el esquema estrella con la singularidad de que una dimensión puede tener ligada otra dimensión esto se hace con el fin de poder dividir los datos de las dimensiones para suministrar un esquema que respalde los requerimientos de diseño[6].

2.5.6.3. Esquema de constelación

Según [6] el esquema constelación está conformado por varios esquemas estrellas, las dimensiones se hallan en el centro del modelo y se encuentran relacionadas con sus respectivas tablas de referencia.

2.6. DATAWAREHOUSE

Según [7] Data Warehouse “Es la combinación de conceptos y tecnologías que procesan la información para el análisis y la consulta en lugar de para el procesamiento de transacciones.” Se deposita información histórica de varias fuentes de datos para comprenderla mejor, lo que permite a la dirección contestar a interrogantes clave de negocio o meramente ejecutar decisiones mejores para escalar su posición en el mercado. La base fundamental para generar inteligencia de negocio es un almacén de datos, conocido como Data Warehouse. Este almacén se encarga de centralizar, reforzar y concentrar los datos generados por una empresa, permitiendo su acceso y exploración de manera eficiente. A través del análisis de los datos almacenados, es posible tomar decisiones estratégicas y obtener información relevante para el negocio. Es importante destacar que un Data Warehouse se construye específicamente para ejecutar consultas e indagaciones sobre las acciones de la organización y no para los procesos que se llevan a cabo en ella. Asimismo, permite almacenar datos históricos, organizar y clasificar la información necesaria para el procesamiento analítico e informático, con el objetivo de responder a las preguntas de negocio y ofrecer una forma intuitiva de interpretar los resultados.

2.6.1. PROCESOS ETL

Según [8] ETL es: “un conjunto de programas encargados de la extracción, transformación y carga de datos. Es uno de los componentes más importantes de una estrategia de bodega de datos en una compañía, puesto que debe asegurar la calidad de datos final de la bodega de datos.” Con el fin de que estos datos finales puedan ser usados para la toma de decisiones estratégicas, estos deben pasar por varias fases de validación antes de ser entregados a su sistema de destino, este proceso requiere mucho detalle, incluyendo la limpieza, estandarización y aprobación de los requerimientos de negocio. Todo este proceso llevado a cabo en el ETL debe ser lo menos intrusivo sobre los sistemas en los cuales acceda a la información, siendo así, este debe consumir una mínima cantidad de recursos sobre el origen de datos.

2.6.1.1. Extracción

Según [9] la extracción de datos es: “el proceso de leer datos de una o varias bases de datos. En este punto, los datos son recogidos para su posterior tratamiento y pueden tener su origen en diferentes fuentes.” Es la primera fase del ETL y sus fuentes de datos van desde

bases de datos operativas dentro de un sistema, un ERP, CRM o incluso archivos planos, cualquier información que sea útil dentro de los requerimientos se usara para la extracción.

2.6.1.2. Transformación

Según [9] la transformación es: “el proceso que convierte el formato de los datos extraídos en el paso previo a otro formato que necesitemos y que pueda ser almacenado en la base de datos de destino.” En esta fase también se pueden filtrar los datos o incluso unirlos con otras fuentes, se pueden hacer varias combinaciones útiles para la siguiente etapa del proceso.

2.6.1.3. Carga

Según [9] la carga es: “el proceso de almacenamiento de los datos transformados en la base de datos de destino.” Juntando las dos etapas anteriores, se culmina en la carga de datos en una nueva fuente funcional y estandarizada para un uso concreto dependiendo de las necesidades del negocio.

2.6.1.4. Como funciona una ETL

En cierta empresa se dispone de tres fuentes de datos distintas y se requiere juntarlas con un formato adecuado para posteriormente almacenarlas en una base de datos analítica, así podemos demostrarlo en la siguiente figura:

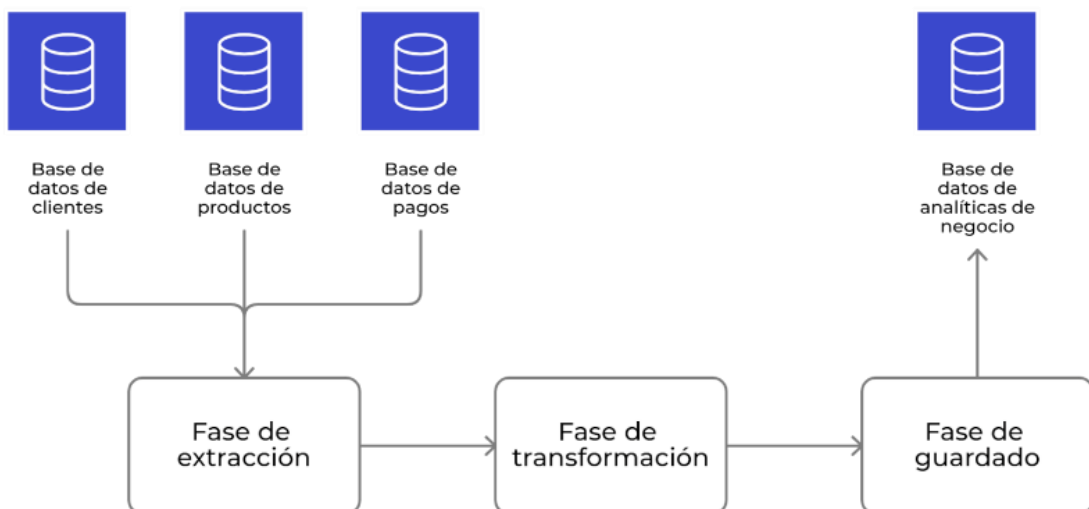


Figura 2. Proceso ETL

Fuente: <https://geekytheory.com/que-es-una-etl-y-como-funciona/>

2.6.2. DATA MART

Según Ralph Kimball, “Un Data mart es un conjunto de datos flexible, idealmente basado en el nivel de granularidad mayor que sea posible, presentado en un modelo dimensional que es capaz de comportarse bien ante cualquier consulta del usuario. En su definición más sencilla, un data mart representa un único proceso de negocio”.

2.6.3. MINERIA DE DATOS

Según [10] la minería de datos: “consiste en un conjunto de metodologías estadísticas y computacionales que, junto a un enfoque desde las ciencias de la conducta, permite el análisis de datos y la elaboración de modelos matemáticos descriptivos y predictivos de la conducta del consumidor.”

El Data Mining y sus metodologías pueden ser aplicados a innumerables áreas, que por ejemplo ayudan a determinar si un cliente es bueno o malo, fechas donde hay más ventas, temporadas donde se perciben más gastos, etc. Así podremos distinguir la variedad de comportamientos del mundo que nos rodea y predecir futuros eventos a través de la segmentación por grupos de individuos que muestren comportamientos similares, como el ejemplo anterior del cliente malo, el cual se puede determinar revisando su historial crediticio, el cual se visualiza por lo general en haciendo uso de métodos estadísticos cuantificables y en base a esto un cliente puede recibir un puntaje alto o bajo que determinara como bueno o malo.

2.6.4. BASE DE DATOS RELACIONAL

Según [11] se define por base de datos a: “una recopilación organizada de información o datos estructurados, que normalmente se almacena de forma electrónica en un sistema informático.” Por lo general las bases de datos relacionales son manejadas por un software gestor de bases de datos abreviado (DBMS). Conjuntamente, los datos almacenados y el DBMS, junto con los aplicativos asociados a ellos, se denominan sistema de bases de datos, simplificado normalmente a base de datos relacional. Los datos de bases de datos en operación se suelen utilizar como estructuras de tuplas repartidas en tablas para aumentar la eficacia del procesamiento y la consulta de registros. Así, se puede acceder, gestionar, cambiar, renovar, controlar y establecer fácilmente los datos.

2.7. INTELIGENCIA DE NEGOCIOS

La definición fue “usada por primera vez en 1958 por Hans Peter Luhn, puede resumirse como la adquisición (recopilación) y uso de conocimiento basado en hechos con la finalidad de mejorar la estrategia de negocio y las ventajas tácticas en el mercado”[12]. Actualmente la inteligencia de negocios está un poco más ligada al factor tecnológico, pues su uso se ve agilizado gracias a las distintas herramientas y avances tecnológicos disponibles en el mercado.

La inteligencia de negocios desde la perspectiva actual es un medio usado por muchas empresas a nivel global para mejorar su toma de decisiones y darle un valor agregado al negocio, a través del discernimiento facilitado de manera rápida y fiable que proporciona una gran versatilidad al hacer uso de la información histórica almacenada en bases de datos, este proceso se realiza a través del flujo para la generación de conocimiento:

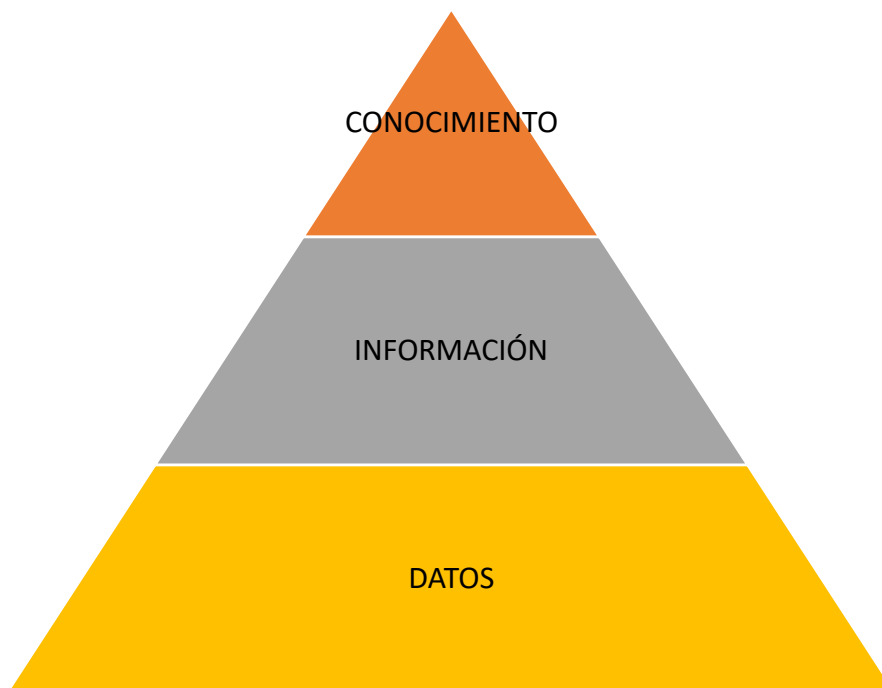


Figura 3. Flujo para la generación del conocimiento

Fuente: los autores

2.8. CUADRO DE MANDO INTEGRAL

Según [13] El cuadro de mando integral hace referencia a: “una herramienta de gestión que ayuda a la toma de decisiones directivas al proporcionar información periódica sobre el nivel de cumplimiento de los objetivos previamente establecidos mediante indicadores.” Los

indicadores cumplen la tarea de recopilar información financiera y no financiera, lo que da soporte en la determinación de equilibrio entre las decisiones de una empresa y las líneas estratégicas.

El cuadro de mando integral tiene un enfoque dinámico, permitiendo apoyar continuamente en la toma de decisiones, comunicación y seguimiento.

Según [14] “El CMI integra cuatro perspectivas o áreas claves y las relaciona con la misión, visión y objetivos. Las cuatro perspectivas son la financiera, la de formación, los procesos internos y la relación con los clientes / usuarios.”

Existen diferentes tipos de CMI [21]:

- 1) CMI operativos: útiles para la gestión del cambio (innovaciones en la organización) en periodos breves de tiempo
- 2) CMI estratégicos: definen los objetivos básicos de la organización en relación en su misión y visión a largo plazo.
- 3) CMI departamentales: específicos para un área de la organización: financiera, dirección, recursos humanos, etc.
- 4) CMI organizativos: definidos según los niveles de responsabilidad. [21]

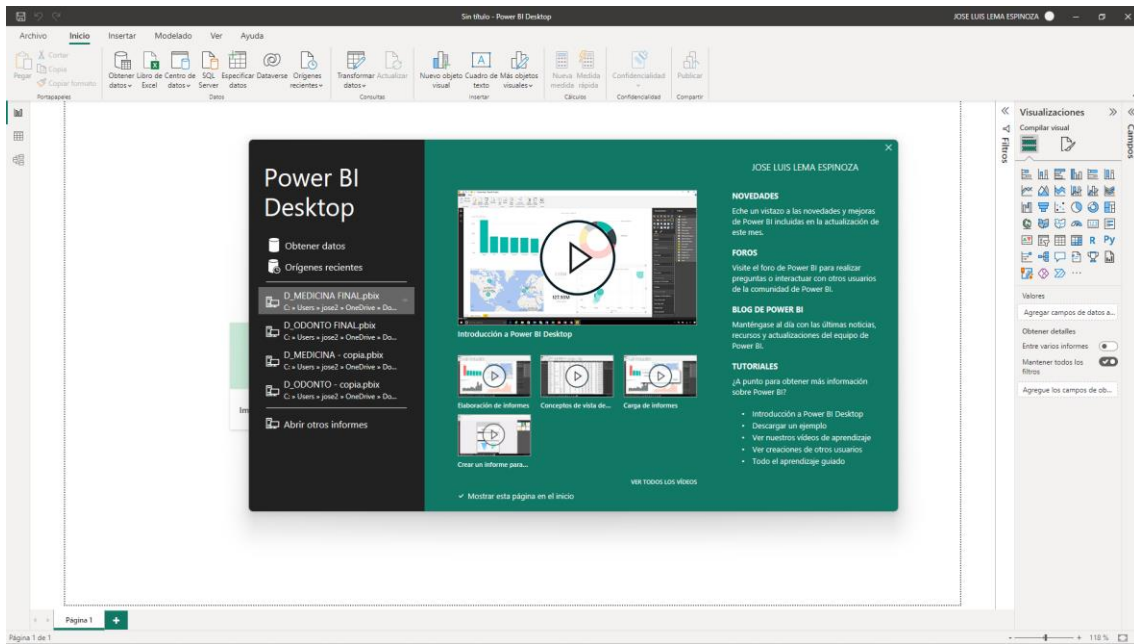
2.9. ANALISIS COMPARATIVO DE HERRAMIENTAS DE BI

Es importante realizar una evaluación exhaustiva de las diferentes soluciones de software que existen en el mercado para ayudar al proyecto a visualizar y analizar sus datos. Actualmente, el mercado de BI está en constante evolución, con una gran variedad de opciones disponibles que prometen ofrecer una amplia gama de funcionalidades y características. Estas herramientas son utilizadas para ayudar a tomar decisiones informadas basadas en datos, y, por lo tanto, es crucial tener una comprensión clara de las diferentes opciones disponibles. En este análisis comparativo, se examinarán las características de las herramientas de BI más populares del mercado, para ayudar a seleccionar la solución que mejor se adapte a las necesidades y objetivos del proyecto.

2.9.1. POWER BI

Según [15] Power BI se define como: “herramienta de análisis que ayuda en la generación de informes, la extracción de datos y la visualización de datos para proporcionar información empresarial.”

Según Gartner [16] en su informe de "Cuadrante Mágico para plataformas de análisis y Business Intelligence" (revisar figura 4), Gartner posicionó a Microsoft Power BI como líder en la categoría de plataformas de análisis y BI, que ofrece una excelente combinación de facilidad de uso y funcionalidad, lo que la convierte en una opción popular entre las empresas de diferentes tamaños y sectores. También se destaca la capacidad de la herramienta para integrarse con otras aplicaciones de Microsoft y con terceros, lo que facilita la colaboración y la interoperabilidad con otras soluciones de software.



2.9.2. TABLEAU

Según [17] Tableau es: “una solución de análisis y visualización de datos que ayuda a las empresas a tomar decisiones comerciales basadas en datos. Combina información de una amplia gama de fuentes para brindar información procesable en tiempo real.”

En la figura 4 según [16] el informe de Gartner, Tableau destaca su facilidad de uso y su enfoque centrado en el usuario, cuenta también con una amplia gama de características y funcionalidades avanzadas, y una fuerte atención al usuario final. Por lo tanto, es una opción popular y recomendada para empresas que buscan una solución de análisis de datos y BI altamente funcional y fácil de usar.

2.9.3. ORACLE ANALYTICS CLOUD

Según [18] Oracle Analytics Cloud es: “una solución impulsada por IA que proporciona sólidas funciones de informes y análisis para empresas de todos los tamaños.”

Según [16] el informe del "Cuadrante Mágico para plataformas de análisis y Business Intelligence" de Gartner que se visualiza en la figura 4 Oracle Analytics Cloud se encuentra posicionado en la categoría de competidores en crecimiento ya que ha mejorado significativamente en términos de facilidad de uso y rendimiento en los últimos años. Sin embargo, el informe señala que la plataforma todavía tiene algunas limitaciones en cuanto a la personalización de informes y la flexibilidad de las visualizaciones de datos, en general análisis de datos sólida y en crecimiento, con una amplia variedad de características y funcionalidades.

2.9.4. CUADRO COMPARATIVO DE HERRAMIENTAS BI

Puntuación de 100%: Esta función cuenta con un soporte integral listo para usar con capacidades líderes en la industria y está disponible inmediatamente después de la instalación, sin necesidad de módulos adicionales, integraciones o desarrollo personalizado.

Puntuación de 85%: Esta característica es moderadamente compatible desde el primer momento y está disponible inmediatamente después de la instalación, sin necesidad de módulos adicionales, integraciones o desarrollo personalizado.

Puntuación de 70%: Esta función no está disponible directamente en el software, pero se puede lograr utilizando otras funciones integradas o cualquier otra solución, sin ningún costo adicional.

Puntuación de 60%: Esta característica está disponible a través de módulos o productos adicionales del proveedor a un costo adicional.

Puntuación de 50%: Esta característica está disponible a través de integraciones adicionales, complementos o aplicaciones de mercado de un proveedor externo a un costo adicional.

Puntuación de 25%: Esta función no está integrada y no se puede agregar mediante la compra de módulos o integraciones adicionales, pero se puede desarrollar de forma personalizada utilizando las API, las bibliotecas, las extensiones y el marco de desarrollo admitido por el software, con o sin costo adicional.

Puntuaciones 0: Esta característica no es compatible.

A continuación, se tomaron las características más relevantes y en las cuales estas herramientas se diferencian.

COMPARATIVA DE HERRAMIENTAS BI	POWER BI	TABLEAU	ORACLE
Minería de texto [análisis de texto] La solución permite indagar y estudiar grandes cantidades de datos de texto no organizados e identificar conceptos, patrones, temas, palabras clave y otros atributos en los datos.	60%	0%	60%
Gestión de Metadatos y Catálogo de Datos La solución proporciona una herramienta de administración de metadatos para centralizar los metadatos.	100%	100%	60%
Reescrituras seguras La solución admite la actualización segura de múltiples fuentes de datos en tiempo real mediante el inicio de procesos de back-end en la aplicación de integración, desde el contexto analítico.	100%	25%	100%
Segmentación y análisis de cohortes La solución puede segmentar los datos a lo largo de muchas dimensiones, en función de criterios específicos, y crear grupos de datos para un análisis más detallado.	100%	100%	100%
Análisis de series de tiempo y pronóstico Ayuda a pronosticar tendencias futuras basadas en datos y tendencias pasadas y presentes utilizando métodos de pronóstico.	100%	100%	100%
Análisis de datos avanzado usando Python y R La solución admite la manipulación y el análisis de datos avanzados y sofisticados mediante el uso de bibliotecas y paquetes de lenguajes de programación Python y R.	100%	100%	100%
Bases de datos relacionales La solución se conecta a bases de datos relacionales como Oracle, SQL Server, MySQL, PostgreSQL, Sybase ASE, IBM DB2, etc.	100%	100%	100%
Herramientas de integración de datos La solución se integra con herramientas populares de integración de datos como Informática, etc.	100%	100%	100%
Plataformas ERP La solución se conecta a los sistemas ERP populares de SAP, Oracle, etc.	85%	50%	85%
Plataformas de mensajería empresarial La solución se conecta a plataformas de mensajería empresarial populares como Slack.	100%	50%	0%
Seguridad integrada La solución admite la importación y herencia de la configuración y las especificaciones de seguridad de un sistema de seguridad externo.	85%	0%	85%
PUNTAJE TOTAL	94%	68%	81%

TABLA 1. COMPARATIVA DE HERRAMIENTAS BI.

La compañía SelectHub elabora una comparación detallada entre las principales herramientas de BI según distintas características propuestas en la tabla1 que han sido tomadas en cuenta en base a las necesidades de nuestro caso de estudio para el proyecto “JUNTOS POR UNA

SONRISA”. Las características más importantes que hacen que Power BI se adapte mejor a las necesidades del proyecto son:

- **Facilidad de uso:** Power BI tiene una interfaz intuitiva y fácil de usar, lo que la hace accesible para usuarios con diferentes niveles de experiencia en análisis de datos. Esto significa que los usuarios pueden crear informes y visualizaciones de datos de manera más rápida y eficiente, sin necesidad de habilidades de programación o codificación complejas.
- **Integración con otras herramientas de Microsoft:** Power BI está estrechamente integrado con otras herramientas de Microsoft, como Excel, SharePoint y Teams, lo que permite una mayor colaboración y eficiencia en el trabajo. Además, la integración con Azure y otras herramientas de nube de Microsoft permite una mayor flexibilidad y escalabilidad.
- **Amplia variedad de conectores de datos:** Power BI tiene una amplia variedad de conectores de datos para diferentes tipos de fuentes de datos, lo que permite a los usuarios conectarse a múltiples fuentes de datos y combinarlos en un solo informe o visualización.
- **Visualizaciones y gráficos avanzados:** Power BI tiene una amplia variedad de opciones de visualización y gráficos avanzados, lo que permite a los usuarios crear visualizaciones de datos altamente personalizadas y atractivas.
- **Comunidad y soporte:** Power BI tiene una gran comunidad de usuarios y una amplia variedad de recursos de soporte y aprendizaje en línea, lo que facilita la resolución de problemas y la mejora de habilidades.

La figura 4 publicada por Gartner Inc. proyecta una visión general sobre los resultados del estudio en el campo de inteligencia de negocios, en el cual podemos destacar el cuadrante de líderes, que tiene a la cabeza a la compañía de Microsoft y Tableau.



Figura 4. Magic Quadrant for Analytics and Business Intelligence Platforms

Fuente: [16]

2.10. ANALISIS COMPARATIVOS DE BASES DE DATOS RELACIONALES

2.10.1. MySQL

Según [19] MySQL es: “Es un SGBD multihilo y multiusuario utilizado en la gran parte de las páginas web actuales. Además, es el más usado en aplicaciones creadas como software libre.”

2.10.2. PostgreSQL

Según PostgreSQL es: “un sistema gestor de base de datos relacional está orientado a objetos y es libre, publicado bajo la licencia BSD.”

2.10.3. Microsoft SQL Server

Según Microsoft SQL Server es: “un sistema gestor de bases de datos relacionales basado en el lenguaje Transact-SQL, capaz de poner a disposición de muchos usuarios grandes cantidades de datos de manera simultánea.”

2.10.4. Oracle

Según Oracle es: “el SGBD por excelencia para el mundo empresarial, considerado siempre como el más completo y robusto, destacando por: Soporte de transacciones, estabilidad y Escalabilidad y multiplataforma.”

2.10.5. CUADRO COMPARATIVO DE SISTEMAS GESTORES DE BASES DE DATOS

La elección de un sistema gestor de bases de datos adecuado es fundamental para cualquier empresa que necesita automatizar procesos de extracción y distribución de información. Un sistema adecuado puede mejorar la eficiencia, la seguridad y la calidad de los datos, mientras que un sistema inadecuado puede ser costoso y limitar la capacidad de la empresa para almacenar y acceder a sus datos de manera eficiente. Es por ello que esta comparación ayuda a evaluar diferentes sistemas en función de las necesidades específicas, como la escalabilidad, la seguridad, el rendimiento y la facilidad de uso, y tomar una decisión informada sobre qué sistema es el más adecuado para el proyecto.

CUADRO COMPARATIVO DE SISTEMAS GESTORES DE BASES DE DATOS

SGBD	CARACTERISTICAS	VENTAJAS	DESVENTAJAS
MYSQL	Amplio subconjunto del lenguaje SQL. Disponibilidad en gran cantidad de plataformas y sistemas. Posibilidad de selección de mecanismos de almacenamiento	Es de código abierto, tiene velocidad al realizar operaciones en las bases de datos, facilidad de configuración e instalación y soporta variedad de S.O.	Un gran porcentaje de sus utilidades no están documentadas
POSTGRE SQL	Tiene la extensión POSTGIS para bases de datos espaciales.	Código abierto y gratuito, multiplataforma. Gran volumen de datos. Transacciones, disparadores y afirmaciones.	Respuesta lenta. Requiere hardware. No es intuitivo.
MICROSOFT SQL SERVER	Soporte de transacciones. Soporta procedimientos almacenados. Incluye también un entorno gráfico de administración que permite el uso de comandos DDL y DML gráficamente. Permite trabajar en modo cliente/servidor.	Registra transacciones. Es la interfaz de acceso de OLE, DB y ADO	Bloqueos a nivel de páginas.
ORACLE	Entorno cliente/servidor. Gestión de grandes bases de datos, Usuarios concurrentes. Alto rendimiento en transacciones. Sistemas de alta disponibilidad. Disponibilidad controlada de los datos de las aplicaciones.	Es un sistema muy completo, tiene un soporte muy aceptable, ya que puede ejecutarse en cualquier sistema operativo.	Es elevado su costo y el coste de información

Tabla 2. CUADRO COMPARATIVO DE SISTEMAS GESTORES DE BASES DE DATOS

Nota: Se muestran las características entre las principales bases de datos en el mercado.[20].

Existen características y beneficios que hacen que Microsoft SQL Server sea una opción popular y recomendada para el proyecto. Mismas que se detallan a continuación:

- Integración con otras herramientas de Microsoft: Microsoft SQL Server está estrechamente integrado con otras herramientas de Microsoft, como Power BI, Excel, SharePoint y Teams, lo que permite una mayor colaboración y eficiencia en el trabajo. Además, la integración con Azure y otras herramientas de nube de Microsoft permite una mayor flexibilidad y escalabilidad.
- Seguridad avanzada: Microsoft SQL Server tiene un conjunto de características de seguridad avanzadas, como la encriptación de datos, la autenticación de usuarios y la auditoría de acceso, que ayudan a proteger los datos de la empresa y cumplir con las normativas y regulaciones de seguridad.
- Escalabilidad y rendimiento: Microsoft SQL Server tiene la capacidad de manejar grandes volúmenes de datos y altos niveles de tráfico de usuarios sin comprometer el rendimiento. Además, Microsoft SQL Server tiene la capacidad de escalar horizontal y verticalmente, lo que permite a las empresas aumentar la capacidad y el rendimiento según sea necesario.
- Herramientas de desarrollo avanzadas: Microsoft SQL Server ofrece un conjunto de herramientas de desarrollo avanzadas, como Visual Studio y SQL Server Management Studio, que permiten a los desarrolladores crear y mantener bases de datos de manera eficiente.
- Soporte y comunidad: Microsoft cuenta con una amplia comunidad de usuarios y una amplia variedad de recursos de soporte y aprendizaje en línea, lo que facilita la resolución de problemas y la mejora de habilidades.

2.11. ANALISIS COMPARATIVOS DE HERRAMIENTAS PARA ETL

Una herramienta ETL debe garantizar que los datos se integren correctamente y que la información almacenada sea precisa, confiable y esté disponible para su análisis, además de ayudar a ahorrar tiempo y recursos al automatizar los procesos de integración de datos. Por ello se debe evaluar las diferentes herramientas en función de la complejidad de las fuentes de datos, la cantidad de datos que deben procesarse, la frecuencia de los procesos de integración de datos y la escalabilidad. La evaluación de herramientas ETL también puede incluir la revisión de características adicionales como la capacidad de programación, la facilidad de uso, la compatibilidad con diferentes sistemas operativos y la integración con otras herramientas de BI.

2.11.1. KNIME

Según [21] Knime es: “una plataforma de analítica de datos de código abierto que permite a los usuarios realizar una variedad de tareas de análisis de datos, como la limpieza y preparación de datos, la visualización de datos y la generación de informes.”

2.11.2. TALEND

Menciona [22] que Talend es: “una plataforma de gestión e integración de datos de código abierto que permite la ingesta, transformación y mapeo de big data a nivel empresarial. El proveedor proporciona conectividad entre redes, calidad de datos y gestión de datos maestros en un solo centro unificado: Data Fabric.”

2.11.3. QLIK

Según [23] QLIK es: “un conjunto de soluciones de gestión de datos de extremo a extremo que ofrece análisis a través de flujos de trabajo orquestados.”

2.11.4. SQL SERVER INTEGRATION SERVICES

Según [24] SQL SERVER INTEGRATION SERVICES es: “una plataforma para crear paquetes de integración y transformación de datos. Está diseñado para extraer y transformar datos de archivos XML, archivos planos y fuentes de datos relacionales, y luego cargarlos en repositorios de destino.”

2.11.5. CUADRO COMPARATIVO DE HERRAMIENTAS PARA ETL

Se realizará una comparación entre diferentes herramientas utilizadas para el proceso de extracción, transformación y carga de información (ETL). Para ello, se analizarán tres de las principales herramientas disponibles en el mercado: KNIME, QLIK y Servicios de Integración de SQL Server de Microsoft. Se tomarán en cuenta diversas características comunes entre las herramientas. A continuación, se asignará una puntuación a cada herramienta en función de su desempeño en cada una de estas características. El objetivo principal de esta comparación es proporcionar una visión detallada de las fortalezas y debilidades de cada herramienta, a fin de ayudar a tomar una decisión informada al seleccionar la herramienta ETL que mejor se adapte a nuestras necesidades.

Puntuación de 100%: Esta función cuenta con un soporte integral listo para usar con capacidades líderes en la industria y está disponible inmediatamente después de la instalación, sin necesidad de módulos adicionales, integraciones o desarrollo personalizado.

Puntuación de 85%: Esta característica es moderadamente compatible desde el primer momento y está disponible inmediatamente después de la instalación, sin necesidad de módulos adicionales, integraciones o desarrollo personalizado.

Puntuación de 70%: Esta función no está disponible directamente en el software, pero se puede lograr utilizando otras funciones integradas o cualquier otra solución, sin ningún costo adicional.

Puntuación de 60%: Esta característica está disponible a través de módulos o productos adicionales del proveedor a un costo adicional.

Puntuación de 50%: Esta característica está disponible a través de integraciones adicionales, complementos o aplicaciones de mercado de un proveedor externo a un costo adicional.

Puntuación de 25%: Esta función no está integrada y no se puede agregar mediante la compra de módulos o integraciones adicionales, pero se puede desarrollar de forma personalizada utilizando las API, las bibliotecas, las extensiones y el marco de desarrollo admitido por el software, con o sin costo adicional.

Puntuaciones 0: Esta característica no es compatible.

A continuación, se tomaron las características más relevantes y en las cuales estas herramientas se diferencian.

CUADRO COMPARATIVO DE HERRAMIENTAS PARA ETL

CARACTERISTICAS/PRODUCTOS	KNIME	QLIK	Servicios de integración de SQL Server de Microsoft
<i>Conectores de software funcionales del cliente</i>	100%	50%	50%
<i>Conectores de software de comercio electrónico</i>	100%	50%	50%
<i>Conectores de software de gestión de proyectos</i>	85%	100%	100%
<i>Conectores de bases de datos relacionales</i>	100%	50%	50%
<i>Conectores de software de análisis web</i>	100%	100%	100%
<i>Conversiones de tipos de datos</i>	100%	100%	100%
<i>Conectores de software de análisis web</i>	100%	100%	100%
TOTAL PINTAJE	98%	75%	71%

TABLA 3. CUADRO COMPARATIVO DE HERRAMIENTAS PARA ETL

En base a los resultado obtenidos en la tabla 3, se concluye que KNIME cumple con las características necesarias para incluirse como herramienta destinada al ETL del proyecto, pues cuenta con una interfaz gráfica de usuario que permite a los usuarios crear flujos de trabajo de ETL arrastrando y soltando nodos preconstruidos, mismos que permiten acceder a una gran cantidad de datos para ser procesados lo que hace que sea fácil para los usuarios crear flujos de trabajo complejos sin tener que escribir código. Además, KNIME es una herramienta de código abierto, lo que significa que es gratuito para usar y personalizar. Tiene una gran comunidad de usuarios y desarrolladores que han creado muchas extensiones y plugins que amplían la funcionalidad de la herramienta.

2.12. ANALISIS COMPARATIVOS DE METODOLOGIAS DE DATA WAREHOUSE

Las metodologías presentadas a continuación ofrecen distintos enfoques para el diseño, desarrollo e implementación de un data warehouse, por lo que es necesario realizar una evaluación comparativa para determinar cuál es la más adecuada para las necesidades y objetivos del proyecto, donde se identifican sus fortalezas y debilidades, y se determina cuál de ellas es la mejor opción para el propósito en cuestión. Algunos factores para considerar en este análisis pueden incluir la complejidad del proyecto, los recursos disponibles, las necesidades de los administrativos y los objetivos a largo plazo. Además, este análisis también puede ayudar a identificar áreas de mejora y oportunidades para optimizar el proceso de construcción del data warehouse.

2.12.1. Bill Inmon

Según [25] Bill Inmon: “fue uno de los primeros autores en escribir sobre el tema de los almacenes de datos, para él es necesaria la transferencia de la información de los OLTP de la empresa a un solo lugar.” La metodología Inmon debe cumplir con una serie de características [25]:

- **Orientación temática:** Para garantizar la cohesión de los datos almacenados, es necesario organizarlos por temas relacionados en el almacén de datos.
- **Variantes en el tiempo:** Se registran los cambios que se producen en los datos a lo largo del tiempo para que los informes generados reflejen esas variaciones.
- **No volatilidad:** La información almacenada no se modifica ni elimina, una vez que se guarda un dato, se convierte en información de solo lectura y se mantiene para consultas futuras.
- **Integración de datos:** El almacén de datos contiene información de todos los sistemas operativos de la organización y estos datos deben ser consistentes entre sí.[25].

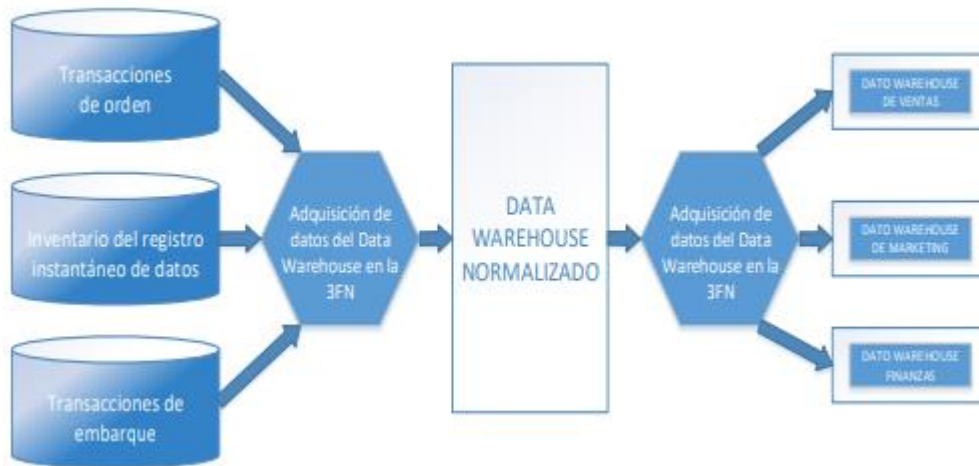


Figura 5 Arquitectura de un DW según Bill Inmon

Fuente: [25]

Al desarrollar un Data Warehouse mediante la metodología Inmon, se consideran los datos de forma descendente, debido a que de esta manera se tiene una perspectiva más amplia, además los Data marts se crean luego de haber culminado el Data Warehouse.

2.12.2. Metodología Ralph Kimball

Se denomina también modelo dimensional, basado en el ciclo de vida dimensional del negocio y es una de las más usadas al momento de construir un Data Warehouse. Para llevarlo a cabo se construyen tablas relacionales con el objetivo de mejorar la toma de decisiones, en base a las consultas realizadas en una base de datos relacional que están ligadas a los procesos de negocio.

Según [26] el modelo dimensional es: “una técnica de diseño lógico que tiene como objetivo presentar los datos dentro de un marco de trabajo estándar e intuitivo, para permitir su acceso con un alto rendimiento.”

Este ciclo de vida del proyecto de DW, está basado en cuatro principios básicos [32]:

- Centrarse en el negocio: Para asegurar el éxito en el desarrollo de una solución efectiva de inteligencia de negocios, es importante centrarse en identificar los requisitos del negocio y su valor asociado. Construir relaciones sólidas con el negocio y mejorar las habilidades de análisis y consultoría puede ayudar a lograr este objetivo [32].

- También es crucial establecer una infraestructura de información adecuada, que incluya un almacén de datos bien diseñado, integrado y de alto rendimiento que refleje los diversos requisitos del negocio identificados en la organización. El desarrollo del almacén de datos debe realizarse en entregas incrementales, con cada entrega basada en el valor del negocio de los elementos identificados, similar a las metodologías ágiles de desarrollo de software [32].
- Por último, para proporcionar valor a los usuarios de negocios, es importante entregar una solución completa que incluya un almacén de datos sólido, bien diseñado y probado, así como herramientas de consulta ad hoc, aplicaciones de informes y análisis avanzado, capacitación, soporte, sitio web y documentación[32].



Figura 6. Ciclo de vida, metodología Kimball

Fuente: [25]

2.12.3. Metodología Hefesto

Según [31] la metodología Hefesto esta principalmente fundamentada en una amplia exploración y cotejo entre metodologías existentes, se toma en cuenta además la perspectiva propia quien realiza los procesos de construcción de almacenes de datos. Cuando se ejecuta un Data Warehouse, es importante tener en cuenta que la metodología utilizada no debe ser demasiado larga ni compleja, ya que esto podría obstaculizar su desarrollo. En consecuencia, para el proyecto Hefestos se han definido cuatro pasos que deben seguirse en el proceso. [31]:

- Análisis de requerimientos
- Análisis de los OLTP

- Modelo lógico del Data Warehouse
- E integración de datos

Al determinar las necesidades de información clave de los usuarios del sistema, identificar las fuentes de datos y sus indicadores específicos, y diseñar el modelo de datos, se pueden establecer claramente los objetivos y los resultados esperados para cada fase de la metodología. Además, la estructura del Data Warehouse es altamente adaptable ya que se construye de acuerdo a los requisitos de los usuarios.[31]

La metodología de Hefestos posee las características listadas a continuación:

- Al identificar las necesidades clave de información por parte de los usuarios del sistema, establecer las fuentes de datos y sus indicadores específicos, y diseñar el modelo de datos, se pueden diferenciar claramente los objetivos y los resultados esperados en cada fase de la metodología. Además, la estructura del Data Warehouse es altamente adaptable, ya que se construye en función de las necesidades y requerimientos de los usuarios [31].
- Se reduce la resistencia al cambio por parte de los usuarios finales, ya que se involucran en cada etapa para determinar el comportamiento y las funciones que se incorporan en el diseño del Data Warehouse. Los modelos conceptuales y lógicos que se implementan son de fácil comprensión y análisis [31].
- La metodología es independiente de las estructuras físicas y de la distribución que contienen el Data Warehouse. Además, tanto el tipo de ciclo de vida como las herramientas utilizadas para construir el Data Warehouse son independientes de la metodología utilizada. Los resultados obtenidos al final de cada fase se convierten en un nuevo punto de partida para el siguiente paso, aplicándose tanto al Data Warehouse como a los Data Marts [31].



Figura 7. Ciclo de vida, metodología Hefesto

Fuente: [25]

2.12.4. Metodología SAS Rapid Data Warehouse

Según [27] la Rapid Warehousing Methodology es: “una metodología iterativa, y está basada en el desarrollo incremental del proyecto de Data Warehouse dividido en cinco fases.”

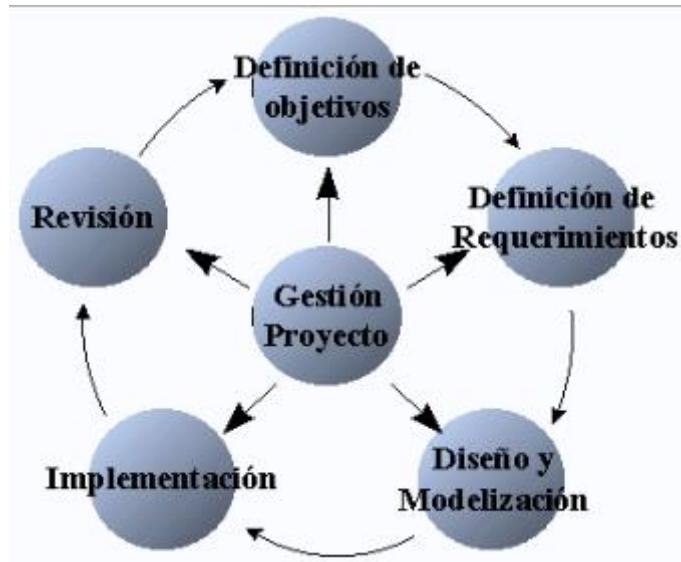


Figura n 8 Fases de la metodología SAS

Fuente: [27]

Fases	Índices
Requerimientos	Determinación de requerimientos Características de la organización Análisis de usuarios [31].
Estrategia de proyecto	Cuantificar el tiempo para cumplir la necesidad (corto mediano-largo plazo) Definir ventajas y desventajas.
Planificación del proyecto	Cronograma de actividades Análisis de riesgo Definición de responsabilidades y grupo de trabajo.
Selección de la tecnología	Entorno actual tecnológico de la organización Definición de tecnologías
Diseño del sistema de información	Determinación del modelo de información Diseño de la interfaz de usuario
Elaboración del sistema de información	Análisis de requerimientos (indicadores) Modelo conceptual Análisis OLTP Conformar indicadores Nivel de granularidad Nivel conceptual ampliado Modelo lógico de la estructura del DW Tablas de dimensiones Tablas de hechos Procesos ETL

Tabla 4. Methodology SAS Rapid Data Warehouse

Nota: Se exponen las fases de la metodología SAS con sus respectivos índices[25].

2.12.5. Lenguaje C#

Según [28] C# es: “un lenguaje de programación moderno, orientado a objetos y con seguridad de tipos. C# permite a los desarrolladores crear muchos tipos de aplicaciones sólidas y seguras que se ejecutan en .NET. C# tiene sus raíces en la familia de lenguajes C.” Está dirigido a la programación orientada a objetos, permitiendo a los programadores tener un flujo de trabajo basado en distintas prácticas de diseño de software, esto permite definir los distintos tipos de comportamientos y manejarlos desde la abstracción de cada individuo.

2.12.6. Arquitectura .NET

Según [28] .NET es: “un sistema de ejecución virtual llamado Common Language Runtime (CLR) y un conjunto de bibliotecas de clases. El CLR es la implementación por parte de Microsoft de la infraestructura de lenguaje común (CLI), un estándar internacional. La CLI

es la base para crear entornos de ejecución y desarrollo en los que los lenguajes y las bibliotecas funcionen juntos sin problemas.”

3. CAPÍTULO 3 DESARROLLO DE LA SOLUCIÓN BUSINESS INTELLIGENCE

3.1. INTRODUCCIÓN

La intención por la que se construye una solución de inteligencia de negocio para el proyecto "JPUS" es ofrecer una toma de decisiones intuitiva y automatizar los procesos de extracción y distribución de información. En el presente capítulo, se tomará información ligada a distintos KPI's que expongan la situación actual del proyecto y que ayuden con futuras proyecciones a los responsables de vinculación a tomar mejores decisiones.

En el presente capítulo, se enfocará en los KPI (Indicadores Clave de Rendimiento) para obtener información sobre la situación actual del proyecto. La información obtenida será utilizada para proyectar el futuro del proyecto y tomar mejores decisiones. La solución de Business Intelligence permitirá a los responsables del proyecto tener una visión clara de la situación actual, identificar áreas de mejora y tomar medidas para mejorar el rendimiento del proyecto.

En resumen, la construcción de la solución de Business Intelligence para el proyecto JPUS tiene como objetivo proporcionar información útil para el análisis y la toma de decisiones, utilizando KPI específicos para evaluar el rendimiento de los procesos de vinculación y relaciones con los clientes. Esto permitirá a los responsables del proyecto tomar mejores decisiones y mejorar la gestión del proyecto.

Para el desarrollo de la solución y dando cumplimiento a la metodología Hefesto, primero se realizó un levantamiento de requerimientos a través de reuniones virtuales y presenciales, con los distintos directivos de medicina y odontología, detallando y depurando las distintas preguntas planteadas por parte de los directivos de vinculación. Posteriormente se determinará el alcance de la solución, cabe la pena indicar que aquí se extrajo los requisitos funcionales, especificando el origen de los datos a analizar, la arquitectura de la solución BI, el modelado de datos OLAP, el modelado de hechos y dimensiones, etc.

La parte final del desarrollo corresponde al tratamiento de datos, a partir de la información registrada en la base de datos operativa se procede con la extracción de datos

útiles para el análisis y construcción de un almacén de datos que alimentará la solución BI. Una vez constituido el DW y los procesos de transformación de datos necesarios, se llevará a cabo la construcción de la solución con la herramienta Microsoft power BI.

3.2. MATERIALES Y METODOS

Esta investigación usa información recopilada de distintas fuentes bibliográficas confiables, mismas que fundamentan el marco teórico. Previo al desarrollo, en el capítulo 1 se estudió distintas herramientas para la realización del proceso ETL, conexión con la base de datos, creación de métricas, indicadores, creación de Dashboard dinámicos, etc.

Las herramientas de Knime y Microsoft Power BI que fueron seleccionadas en el proceso de evaluación del capítulo 2 son las que mejor se acomodan para la construcción de la solución, es por eso que fueron usadas en la construcción de la solución BI, lo que nos permitió aplicar el método inductivo para interpretar los datos obtenidos; y el deductivo para analizar la información empleada.

3.3. FASE 1: ANALISIS DE REQUERIMIENTOS

Para la obtención de información necesaria, se realizaron entrevistas, encuestas y reuniones, con el fin de analizar las necesidades del proyecto “JPUS”, para identificar perspectivas e indicadores del modelo de inteligencia de negocios a realizarse, identificando las siguientes necesidades:

3.3.1. IDENTIFICAR PREGUNTAS DE NEGOCIO

En este apartado, nos enfocaremos en identificar las preguntas de negocio que serán abordadas en la solución de BI. Para lograr esto, consideraremos aquellas preguntas que sean relevantes para identificar las perspectivas y los indicadores clave que permitirán la construcción de dicha solución.

SIMBOLOGIA: Las **PERSPECTIVAS** se encontrarán con el formato de texto en negrita, mientras que los INDICADORES se encontrarán con el formato de subrayado al texto, dentro del presente estudio.

A continuación, se presentan las preguntas de negocio que hemos obtuvieron tanto como para medicina como para odontología:

3.3.2. MEDICINA

- Promedio de paciente atendidos por **cantón y año (ubicación y tiempo)**
- Promedio de peso de acuerdo a la **ubicación y edad**
- Índice de neonatos prematuros (*corresponde a menores de 37 semanas*)
- Promedio de consumo de leche materna de acuerdo por ubicación y genero
- Promedio de tipo alimentación por pacientes **y edad**
- Diagnósticos comunes por **ubicación y tiempo**
- Promedio de pacientes con problemas de malnutrición de acuerdo al **género**

3.3.3. ODONTOLOGIA

- Promedio de pacientes atendidos por **cantón y año (ubicación y tiempo)**
- Patologías más comunes por **genero**
- Índice CPO por **edad y ubicación**
- Tasa de diagnósticos comunes por **ubicación**
- Tasa de tratamientos comunes por **ubicación**
- Total, de pacientes atendidos por **sector** en determinado **tiempo**
- Índice de mal oclusiones por **edad y genero**
- Índice de enfermedad periodontal por **edad y genero**
- Índice de fluorosis por **edad y genero**

3.3.4. INDICADORES Y PERSPECTIVAS

Después de haber completado el proceso de identificación de preguntas de negocio en función de los requisitos de la información, se llevó a cabo la identificación de los siguientes indicadores:

3.3.5. INDICADORES MEDICINA

Después del análisis de las preguntas de negocio de medicina se obtuvieron los siguientes indicadores:

- Promedio de peso
- Promedio de edad
- Índice de neonatos prematuros
- Promedio de consumo de leche materna

- Promedio de consumo de leche en formula
- Promedio tipo alimentación
- Diagnósticos comunes
- Promedio de pacientes con problemas de malnutrición
- Total, de pacientes atendidos

3.3.6. PERSPECTIVAS MEDICINA

Seguido al análisis de las preguntas de negocio de medicina se obtuvieron las siguientes perspectivas:

- Ubicación
- Pacientes
 - Edad
 - Genero
- Tiempo Medicina
- Departamento

3.3.7. INDICADORES ODONTOLOGIA

Después del análisis de las preguntas de negocio de odontología se identificó los siguientes indicadores:

- Promedio de edad
- Patologías comunes
- Índice CPO
- Tasa de diagnósticos y tratamientos comunes
- Total, de pacientes atendidos
- Índice de problemas bucales
- Índice de mal oclusiones
- Índice enfermedad periodontal
- Índice de fluorosis
- Total, de pacientes atendidos

3.3.8. PERSPECTIVAS ODONTOLOGIA

Posteriormente del análisis de las preguntas de negocio de odontología se consiguieron las siguientes perspectivas:

- Ubicación
- Pacientes
 - Edad
 - Genero
- Tiempo Odontología
- Departamento

3.3.9. MODELO CONCEPTUAL

De acuerdo con los indicadores y perspectivas planteados en los puntos anteriores se elaboró el siguiente modelo conceptual:

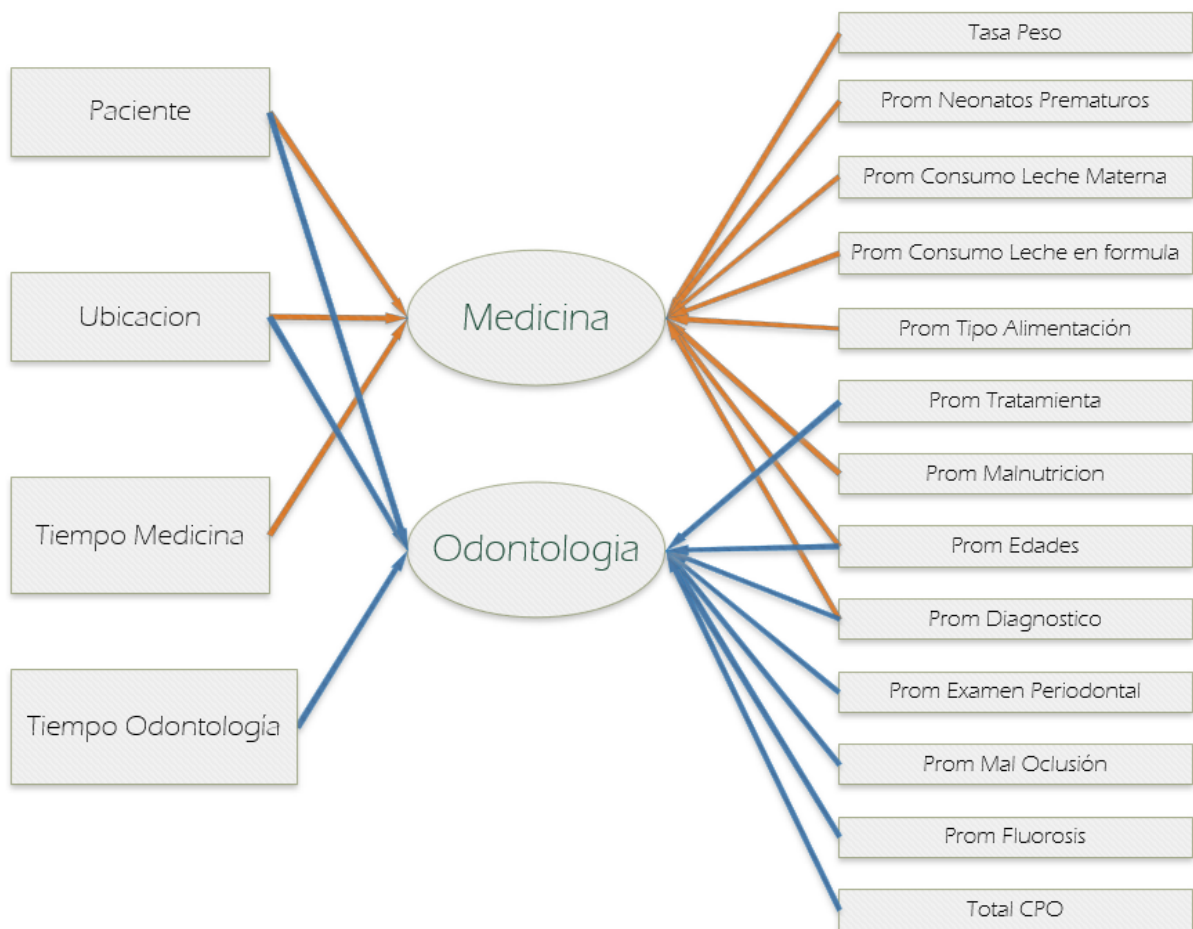


Figura 9: Modelo conceptual

Fuente: Los autores.

3.4. FASE 2: ANALISIS DE DATA SOURCES

Los datos necesarios para la construcción del DW son tomados de la base de datos operativa del sistema JPUS, en base a los cuales se han definido los KPIs, se definieron las

métricas y granularidad El modelo conceptual es la base para el desarrollo de la solución de BI.

3.4.1.1. HECHOS E INDICADORES

Acorde a los puntos anteriores, se definen los siguientes indicadores con sus correspondientes funciones:

3.4.1.1.1. HECHOS E INDICADORES MEDICINA

- Indicador promedio de edades
 - Hechos: promedio de edades
 - Función de agregación: AVG
- Indicador Promedio de peso
 - Hechos: Promedio de peso
 - Función de agregación: AVG
- Indicadores neonatos prematuros
 - Hechos: Cantidad de neonatos prematuros
 - Función de agregación: SUM
- Indicador promedio consumo de leche materna
 - Hechos: Promedio de consumo de leche materna
 - Función de agregación: AVG
- Indicador promedio consumo de leche en formula
 - Hechos: Promedio consumo leche en formula
 - Función de agregación: AVG
- Indicador Promedio de alimentos consumidos
 - Hechos: Promedio de alimentos consumidos
 - Función de agregación: AVG
- Indicador diagnósticos comunes
 - Hechos: diagnósticos comunes
 - Función de agregación: AVG
- Indicador promedio de pacientes con problemas de malnutrición
 - Hechos: promedio de pacientes con problemas de malnutrición
 - Función de agregación: AVG
- Indicador promedio de pacientes

- Hechos: promedio de pacientes
- Función de agregación: AVG

3.4.1.1.2. HECHOS E INDICADORES ODONTOLOGIA

- Indicador promedio de edades
 - Hechos: promedio de edades
 - Función de agregación: AVG
- Indicador patologías más comunes
 - Hechos: patologías más comunes
 - Función de agregación: AVG
- Indicador CPO
 - Hechos: Índice CPO
 - Función de agregación: AVG
- Indicador mal oclusiones
 - Hechos: Índice de mal oclusiones
 - Función de agregación: AVG
- Indicador promedio de pacientes
 - Hechos: promedio de pacientes
 - Función de agregación: AVG

3.4.1.2. MAPEO

En el Data Source, el proceso correspondiente que se lleva a cabo para el registro de fichas clínicas está representado por el siguiente Diagrama de Entidad Relación (representa la información a través de Entidades, Relaciones, Cardinalidades, Claves, Atributos y Jerarquías de generalización):

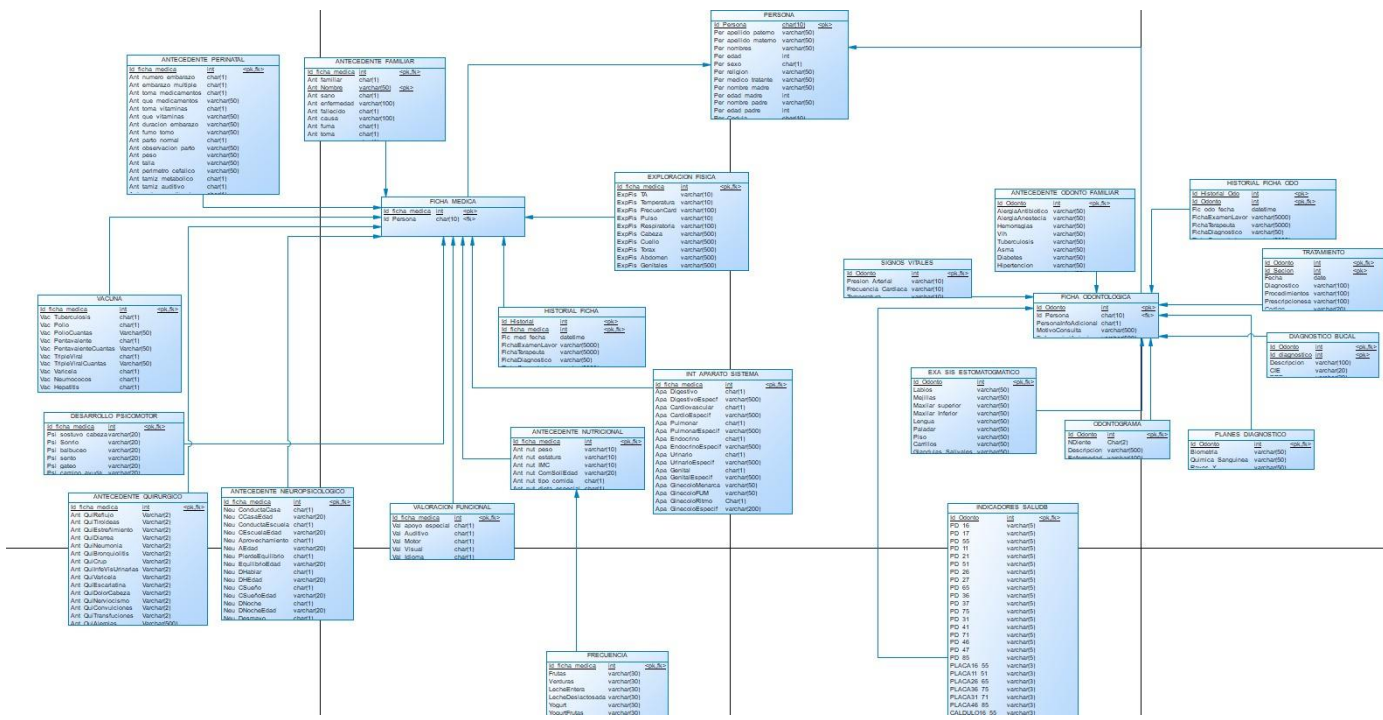


Figura 10. MODELO CONCEPTUAL BASE DE DATOS OPERATIVA

Fuente: Los Autores

A continuación, se muestra la correspondencia del mapeo entre los dos modelos:

3.4.1.2.1. PERSPECTIVAS

La perspectiva **Paciente** se relaciona con las siguientes tablas:

- Persona

La perspectiva **Tiempo** se relaciona con las siguientes tablas:

- FICHA_MEDICA
- FICHA_ODONTOLOGICA

La perspectiva **Ubicación** se relaciona con las siguientes tablas:

- Persona
- Parroquias
- Catones

3.4.1.2.2. INDICADORES

3.4.1.2.2.1 INDICADORES MEDICINA

El indicador de peso se relaciona con las siguientes tablas:

- Ant_nutricional
 - Atributo: Ant_Nut_Peso (AVG)

El indicador de neonatos prematuros se relaciona con las siguientes tablas:

- Antecedente perinatal
 - Atributo: Ant_Duracion_Embarazo (SUM)

Indicador promedio consumo de leche materna se relaciona con las siguientes tablas:

- Antecedente perinatal
 - Atributo: Ant_Tomo_Seno (AVG)

Indicador promedio consumo de leche en formula se relaciona con las siguientes tablas:

- Antecedente perinatal
 - Atributo: Ant_Tomo_Formula (AVG)

Indicador de tipo de alimentación se relaciona con las siguientes tablas:

- ANTECEDENTE_NUTRICIONAL
 - Atributos: Ant_nut_tipo_comida (AVG)

Indicadores diagnósticos comunes medicina se relaciona con las siguientes tablas:

- ANTECEDENTE_NUTRICIONAL
 - Atributo: Ant_nut_peso, Ant_nut_estatura(AVG)

Indicador promedio de pacientes con problemas de malnutrición se relaciona con las siguientes tablas:

- Antecedente_Nutricional
 - Atributos: Ant_Nut_Imc (AVG)

3.4.1.2.2 INDICADORES ODONTOLOGIA

Indicador Total de pacientes se relaciona con las siguientes tablas:

- Ficha_Odontológica
 - Atributos: Id_ Odonto (SUM)

Indicador patologías más comunes se relaciona con las siguientes tablas:

- Tratamiento
 - Atributo: Diagnostico (AVG)

Indicador CPO se relaciona con las siguientes tablas:

- Indicadores_SaludB
 - Atributo: Total_CPO(AVG)

Indicador promedio mal oclusiones se relaciona con las siguientes tablas:

- Indicadores_SaludB
 - Atributo: Mal_oclusión

Indicador promedio enfermedad periodontal se relaciona con las siguientes tablas:

- Indicadores_SaludB
 - Atributo: Enf_Periodontal

Indicadores diagnósticos comunes se relaciona con las siguientes tablas:

- Tratamiento
 - Atributo: Diagnostico

Indicadores tratamientos comunes se relaciona con las siguientes tablas:

- Tratamiento
 - Atributo:Codigo_Procedimiento

3.4.1.3. GRANULARIDAD

Tras el análisis realizado al modelo propuesto se determinó el nivel de detalle por cada perspectiva que se va a tomar en cuenta en la solución:

Perspectiva Paciente:

Paciente_Edad de la tabla Persona atributo Per_Edad que identifica la edad del paciente.

Paciente_Genero de la tabla Persona atributo Per_Sexo que identifica el sexo del paciente.

Perspectiva Ubicación:

Nombre_Canton de la tabla Persona atributo Per_Canton que identifica el cantón.

Nombre_Parroquia de la tabla Persona atributo Per_Parroquia que identifica la parroquia.

Perspectiva Tiempo Medicina:

Los campos Anio, Mes, Dia se obtendrán de la tabla Ficha_medica atributo FechaCreacion.

Perspectiva Tiempo Odontología:

Los campos Anio, Mes, Dia se obtendrán de la tabla Ficha_odontologica atributo FechaCreacion

3.4.1.4. MODELO CONCEPTUAL AMPLIADO

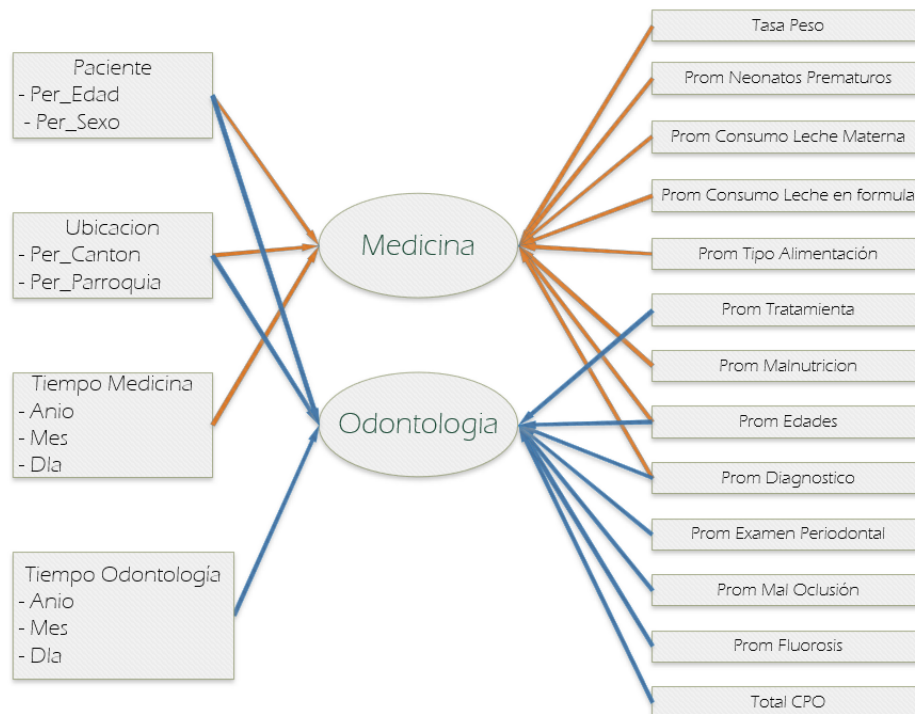


Figura 11. Modelo conceptual ampliado

Fuente: Los autores

3.5. FASE 3 MODELO LOGICO DEL DW

Esta parte del desarrollo corresponde al modelo lógico final del data warehouse antes de realizar los procesos ETL.

3.5.1.1. TIPOLOGIA

Se empleará el modelo tipo constelación para la construcción del modelo lógico, El esquema de constelación es más complejo que otras arquitecturas debido a que contiene múltiples tablas de hechos. Esta solución permite que las tablas de dimensiones puedan ser compartidas entre más de una tabla de hechos. La arquitectura de constelación ofrece mucha flexibilidad, ya que las distintas tablas de hechos se asignan a las dimensiones relevantes para cada uno de ellos. Esta flexibilidad resulta útil cuando los hechos se asignan a un nivel de una dimensión y otros hechos a otro nivel de detalle de la misma dimensión.

Este modelo se adecua a las necesidades del proyecto, por ejemplo, los hechos se dividen en dos tablas independientes para seccionar los campos pertenecientes a medicina y odontología, que comparten dimensiones de persona y ubicación, pero que usan dos tablas de tiempo autónomas, además las medidas no pueden ser las mismas, optando por separarlas específicamente para cada caso clínico, permitiendo visualizar la información requerida de forma que no haya interferencia entre datos.

3.5.1.2. TABLA DE DIMENSIONES

En base al modelo conceptual presentado anteriormente en la figura 11, se definen las siguientes dimensiones:

Perspectiva Paciente:

- La dimensión se denominará Persona
- Su llave primaria es: Persona_Id
- El campo Per_Edad cambia a Persona_Edad
- El campo Per_Sexo cambia a Persona_Sexo

La siguiente figura muestra la dimensión Paciente

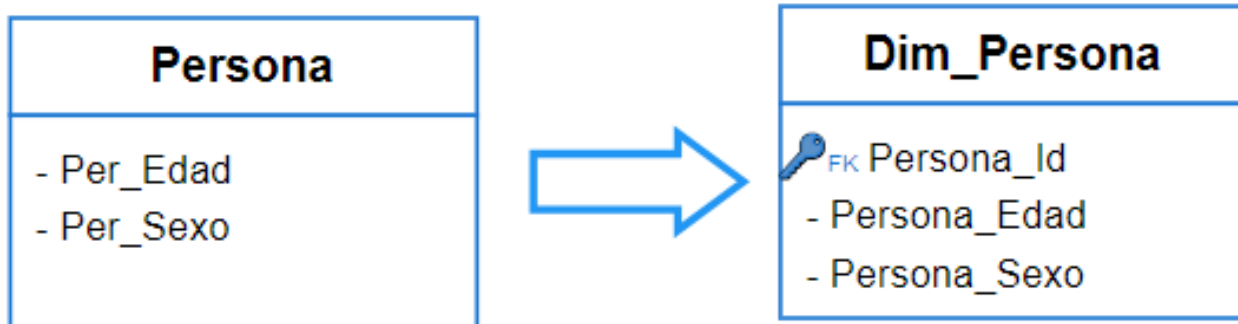


Figura 12 Dimensión Persona

Fuente: Los autores

Perspectiva Ubicación:

- La dimensión se denominará Ubicación
- Su llave primaria es: Ubicación_Id
- El campo Per_Canton cambia a Nombre_Canton
- El campo Per_Parroquia cambia a Nombre_Parroquia

La siguiente figura muestra la dimensión Ubicación:

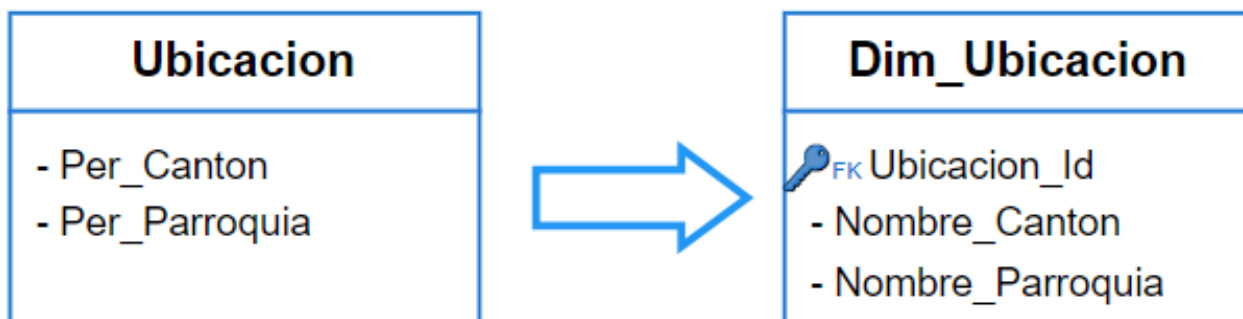


Figura 13 Dimensión Ubicación

Fuente: Los Autores

Perspectiva Tiempo_Med:

- La dimensión se denominará Tiempo:
- Su llave primaria es: Tiempo_Id_Med
- El campo Ficha_med_Fecha se separa en:
 - Año
 - Mes

- Dia

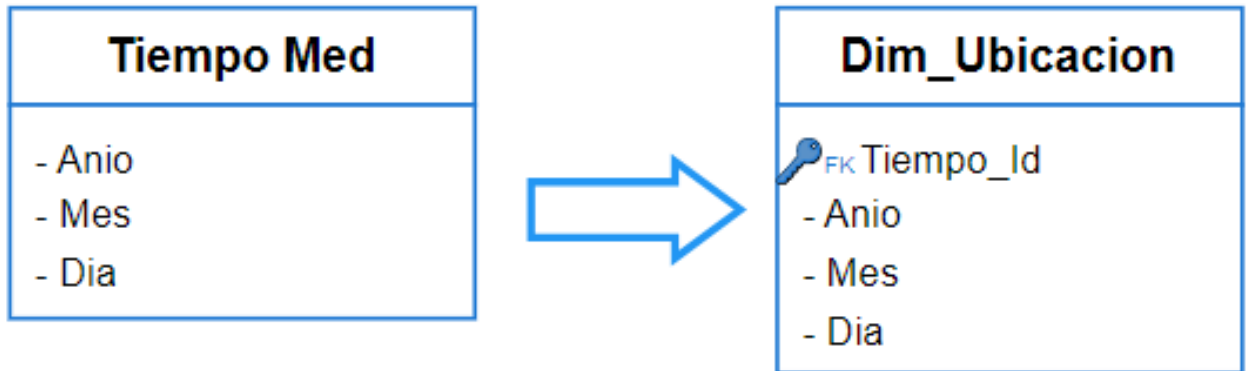


Figura 14. Dimensión Tiempo Medicina

Fuente: Los autores

Perspectiva Tiempo_Odo:

- La dimensión se denominará Tiempo:
- Su llave primaria es: Tiempo_Id_Odo
- El campo Ficha_med_Fecha se separa en:
 - Anio
 - Mes
 - Dia

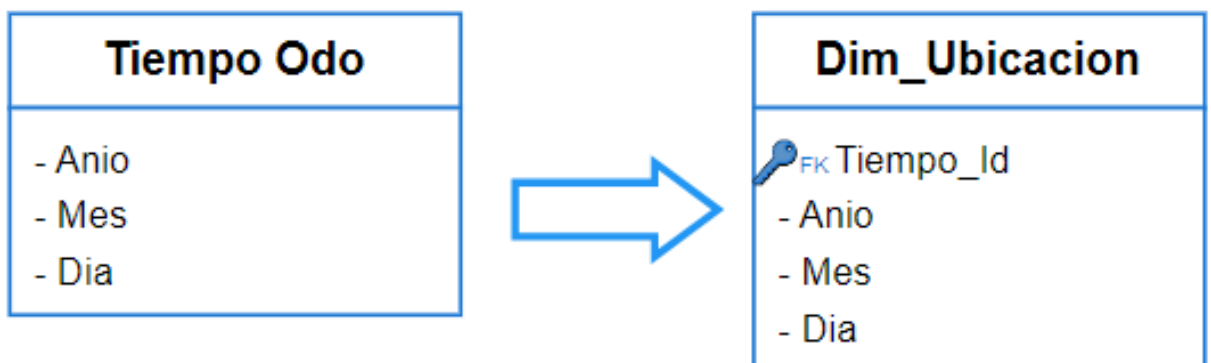


Figura 15 Dimensión Tiempo Odontología

Fuente: Los autores

3.5.1.3. TABLA DE HECHOS

Dentro de la arquitectura del Data Warehouse se determinó la necesidad de crear dos tablas de hechos que llevaran los nombres de Medicina_Hechos y Odontologia_Hechos,

segmentando específicamente la información para medicina y odontología correspondientemente, a partir de esto se tienen las siguientes tablas de hechos:

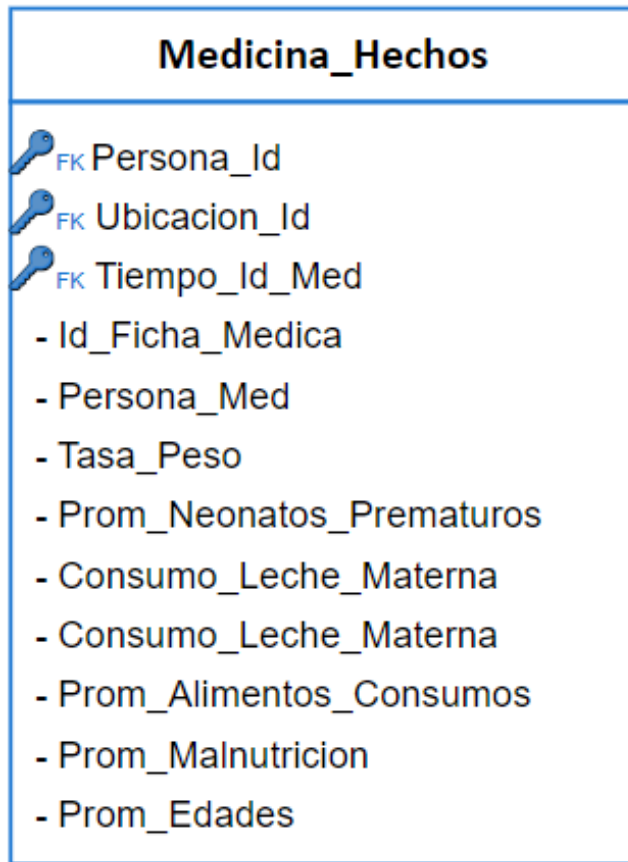


Figura 16 Hechos Medicina

Fuente: Los autores

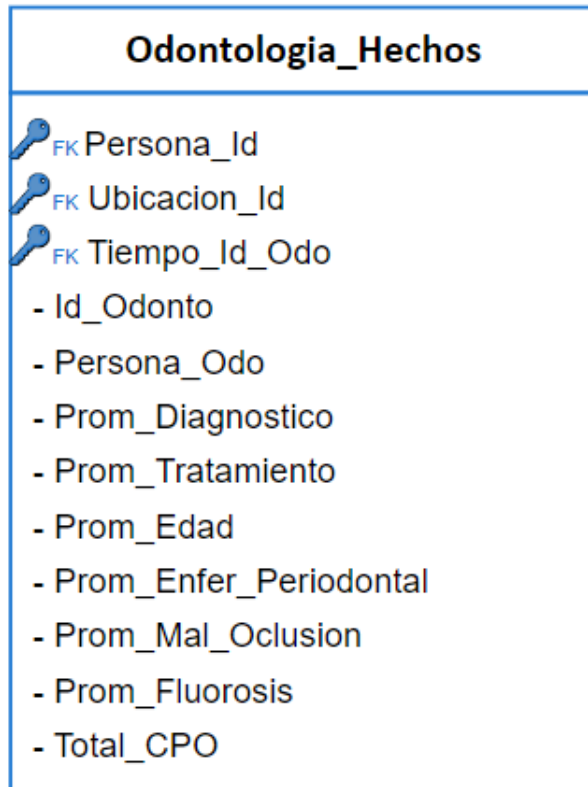


Figura 15 Hechos Odontología

Fuente: Los autores

3.5.1.4. UNIONES

El modelo del data warehouse adjunto quedaría de la siguiente manera:

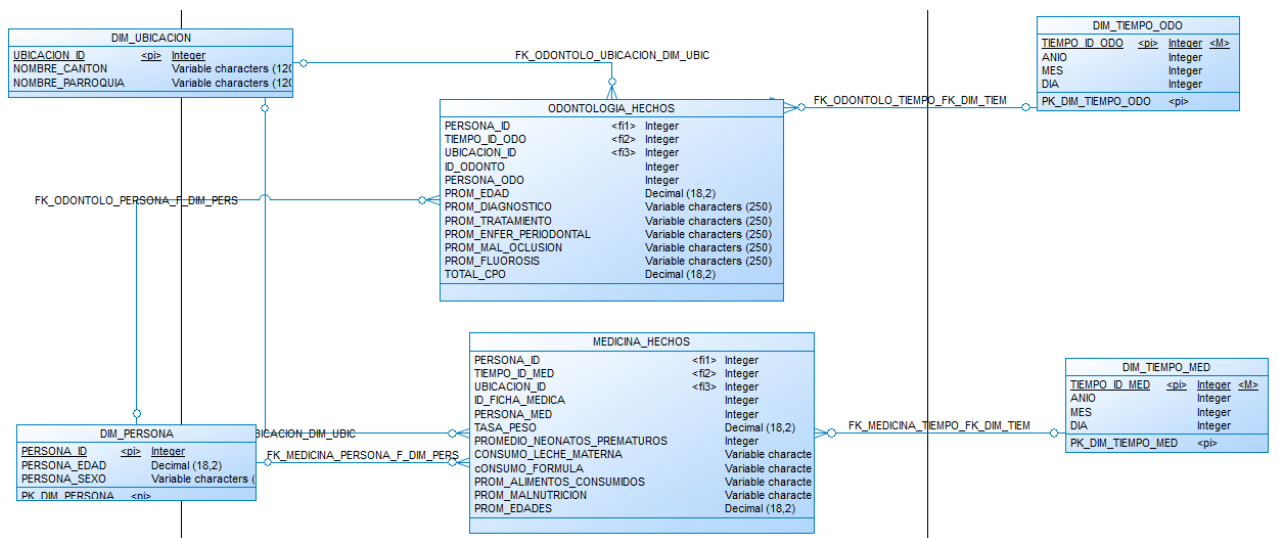


Figura 16 Modelo Logico del Dw

Fuente: Los Autores

3.6. FASE 4 INTEGRACION DE DATOS

Una vez construido el Modelo Lógico, se deberá proceder a poblarlo con datos, utilizando técnicas de limpieza y calidad de datos, procesos ETL, etc.

Luego se definirán las reglas y políticas de actualización, así como también los procesos que la llevarán a cabo.

3.6.1.1. CARGA INICIAL

Este paso se encargará de poblar con datos el modelo del data warehouse, asegurando la limpieza y calidad de la información a verter, ya que se usan varias tablas tanto para dimensiones como para hechos, este proceso lo va a llevar a cabo el software Knime Analytics Plataforma facilitando la organización del ETL así como la integridad.

En medida de lo posible se debe evitar cargar datos nulls, vacíos o que no aporten integridad, por ende, se debe establecer una serie de condiciones y restricciones específicas para cada caso, asegurando solamente el paso de datos de interés.

Primero, se cargarán los datos de las dimensiones y luego los de las tablas de Hechos, esto debido a la existencia de claves foráneas, evitando problemas rechazo por parte del sistema gestor de base de datos, en este paso, es importante registrar en detalle las acciones realizadas en Knime. Por lo general, los sistemas ETL trabajan con pasos y relaciones, donde cada paso realiza una tarea específica del proceso ETL y cada relación indica la dirección en la que debe fluir el conjunto de datos. Es decir, se debe comenzar por lo más general y avanzar hacia lo más específico para obtener una visión general y detallada del proceso.

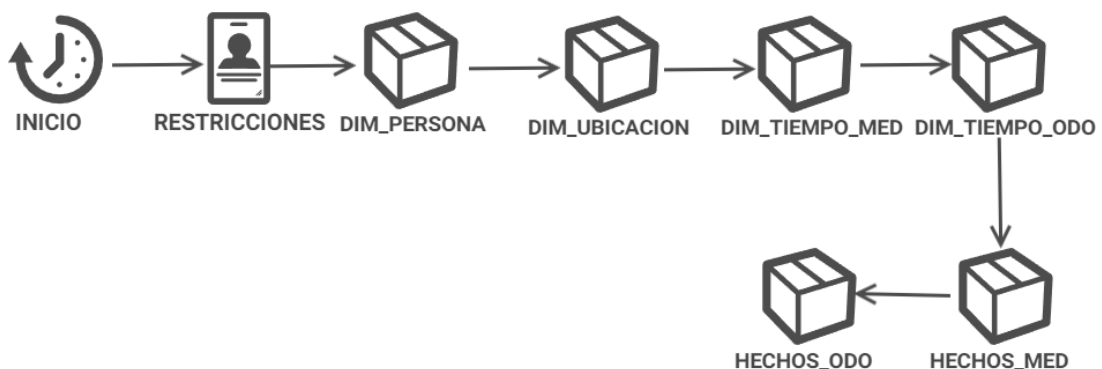


Figura 17 Proceso ETL principal

Fuente: Los autores

3.6.1.2. ACTUALIZACION

Cuando se haya ejecutado la carga inicial del DW, se deben establecer las políticas y estrategias de actualización periódica.

Entonces, se deben llevar a cabo las siguientes acciones:

- La información se refrescará: todos los días a las 00:00hs mediante una tarea programada dentro del servidor de producción.

```
"C:\Program Files\KNIME\knime.exe" -nosave -consoleLog -noexit -nosplash -reset -application  
|prg.knime.product.KNIME_BATCH_APPLICATION -workflowDir="C:\Users\jose2\knime-workspace\ETL_FINAL"
```

Figura 18. Archivo. Bat para la actualización del proceso ETL

Fuente: Los autores

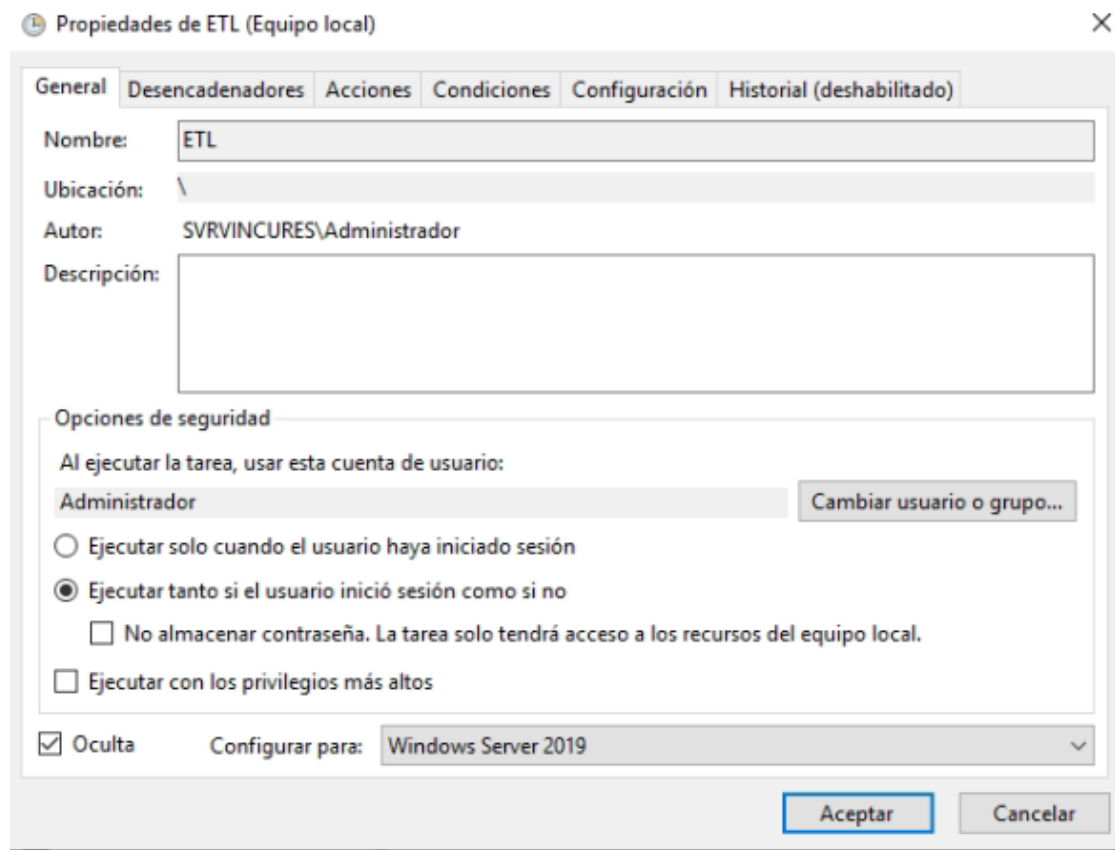


Figura 19. Tarea programada dentro del servidor.

Fuente: Los autores

Actualización programada

Mantener los datos actualizados

Configure una programación de actualización de datos para importar datos del origen de datos al conjunto de datos. [Más información](#)

Activar

Frecuencia de actualización

Diaría

Zona horaria

(UTC-05:00) Bogotá, Lima, Quito

Hora

1 00 a. m

[Agregar otra hora](#)

Destinatario del envío de notificaciones de los errores de actualización

Propietario del conjunto de datos

Estos contactos:

Escriba las direcciones de correo electrónico

Aplicar

Descartar

Figura 20. Actualización Dashboard de Medicina.

Fuente: Los autores

Actualización programada

Mantener los datos actualizados

Configure una programación de actualización de datos para importar datos del origen de datos al conjunto de datos. [Más información](#)

Activar

Frecuencia de actualización

Diaría

Zona horaria

(UTC-05:00) Bogotá, Lima, Quito

Hora

1 00 a. m

[Agregar otra hora](#)

Destinatario del envío de notificaciones de los errores de actualización

Propietario del conjunto de datos

Estos contactos:

Escriba las direcciones de correo electrónico

Aplicar

Descartar

Figura 21. Actualización Dashboard de Odontología.

Fuente: Los autores

- Se ejecutará siempre primero el vaciado de la base de datos del data warehouse.

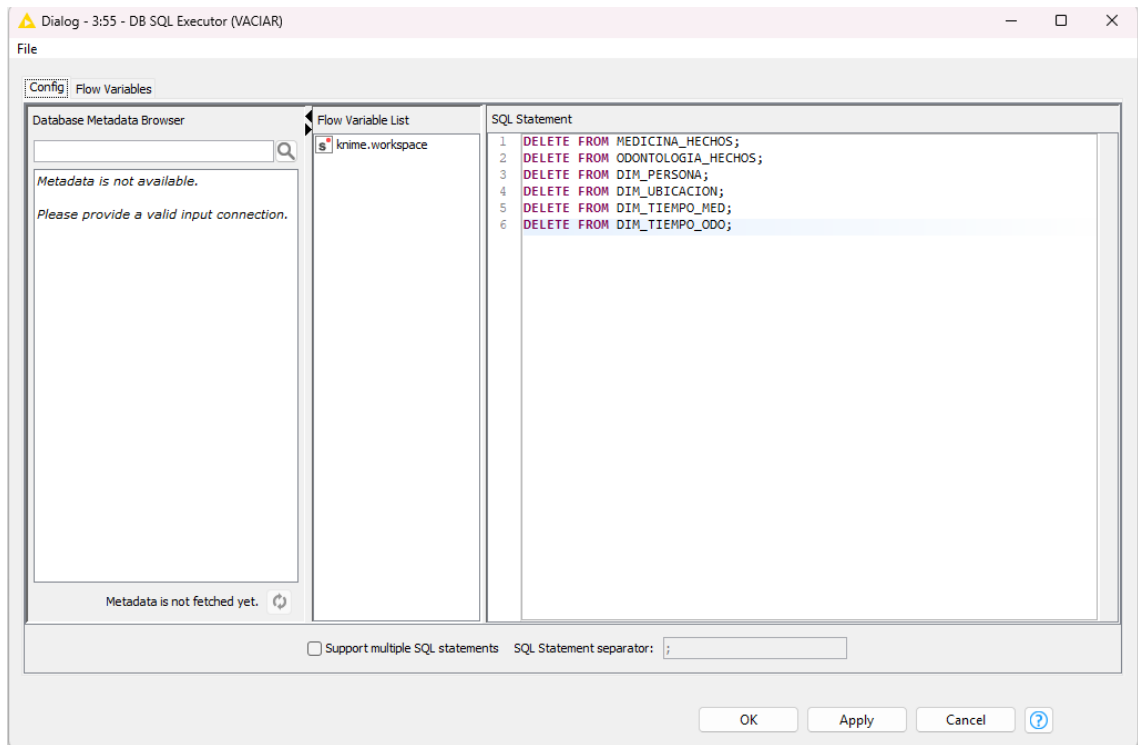


figura 22. SQL Executor para vaciar el data warehouse.

Fuente: Los autores

- Los datos de las tablas de Dimensiones serán cargados siempre en su totalidad.

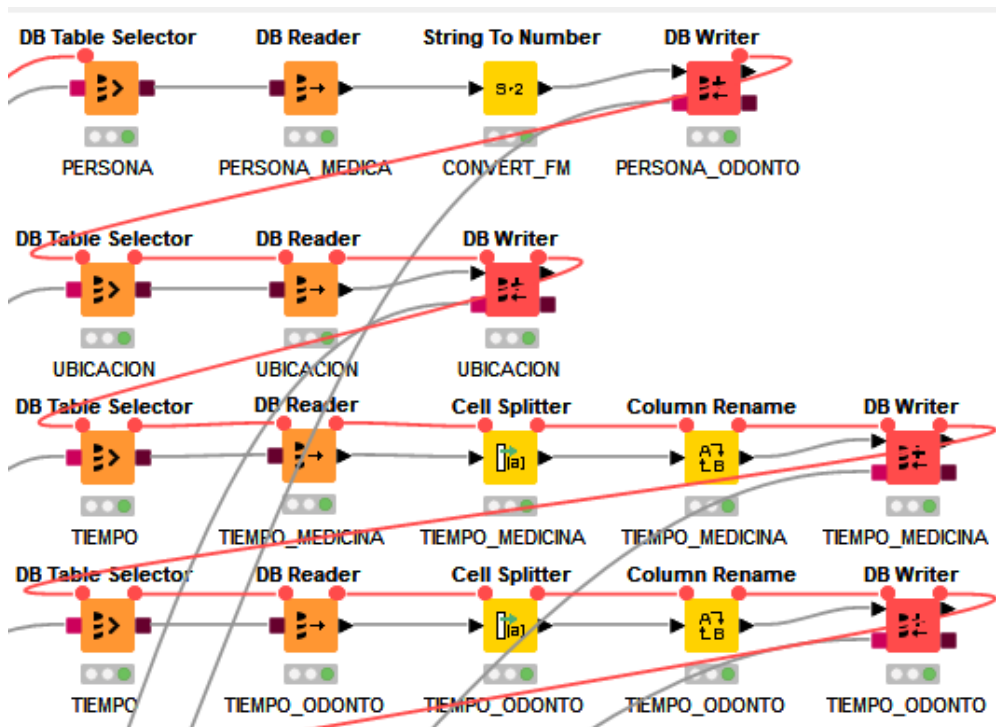


Figura 23. Carga de datos en las dimensiones del data warehouse

Fuente: Los autores

- Los datos de las tablas de Hechos serán cargados siempre en su totalidad

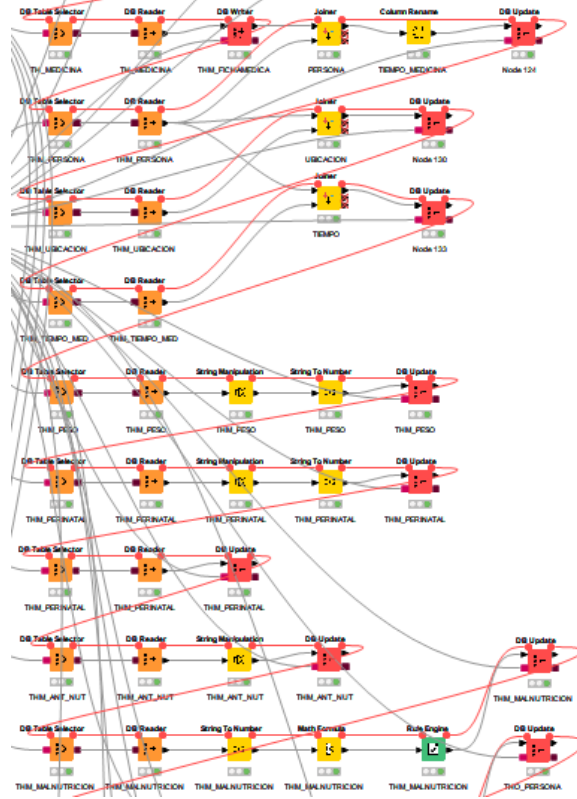


Figura 24. Carga de datos en la tabla de hechos de medicina

Fuente: Los autores

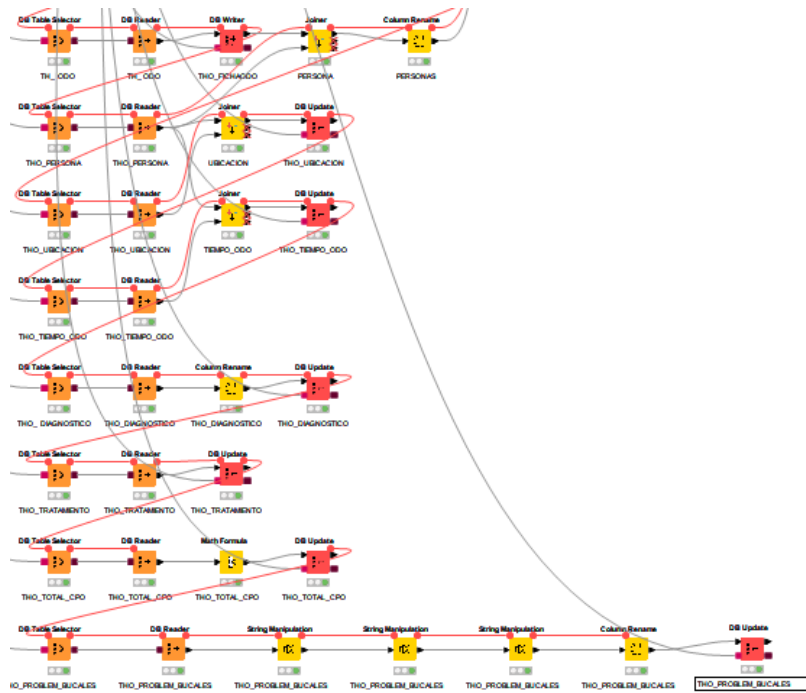


Figura 25. Carga de datos en la tabla de Hechos de Odontología

Fuente: Los autores

3.7. CUBO MULTIDIMENSIONAL

Una vez finalizado el proceso ETL, se cargan los datos en el software de Microsoft Power BI segmentando de dos formas tanto para medicina y odontología, dando paso a la creación de dos cubos multidimensionales.

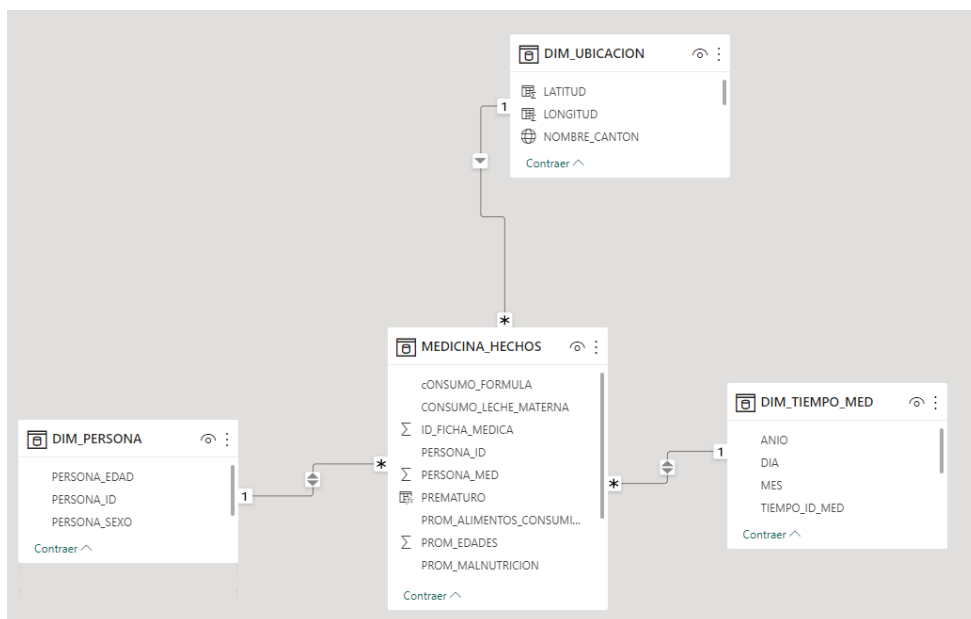


Figura 26 Cubo multidimensional de medicina

Fuente: Los autores

Como se observa en la figura medicina está separado y funciona de manera independiente sobre los datos de odontología, permitiendo un análisis a profundidad, mayor análisis en los datos y mejores tiempos de respuesta cuando se actualice el dashboard general. De la misma forma tenemos:

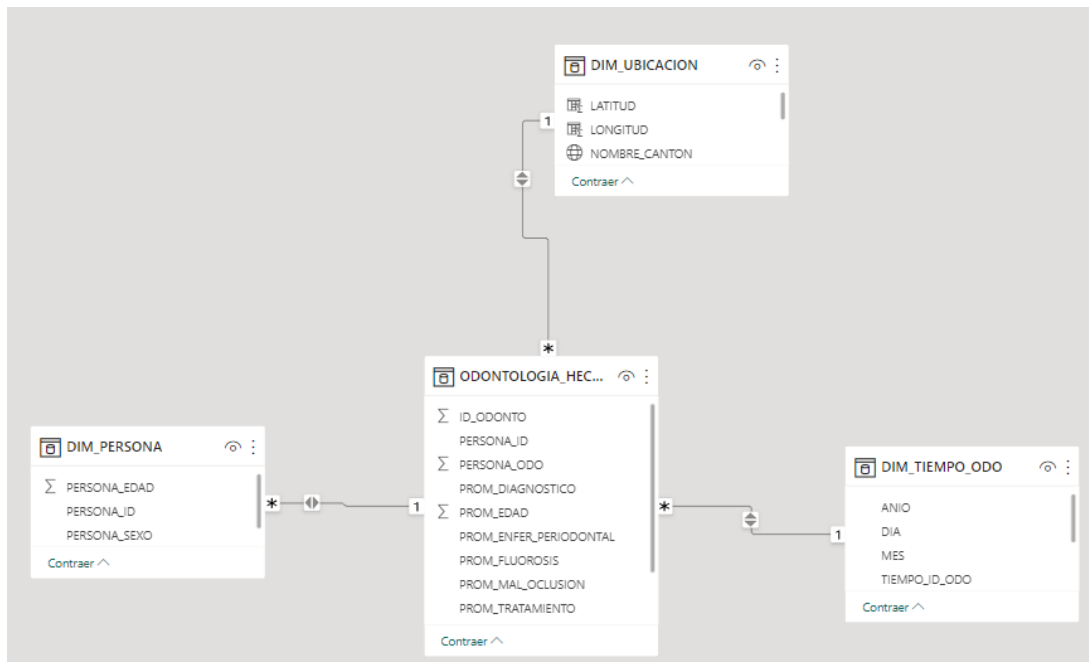


Figura 27 cubo multidimensional de odontología

Fuente: Los autores

3.7.1.1. CONEXIÓN DE BASE DE DATOS DE ORIGEN

Como primer paso del proceso ETL se obtiene el data source (origen de datos), en este caso corresponde a la base de datos operacional del proyecto JPUS, alojado en el motor de base de datos proveído por Microsoft, mismo que se maneja con un driver del tipo JDBC (Java Database Connectivity) que es una interfase de acceso a bases de datos SQL que proporciona un acceso uniforme a la base de datos.

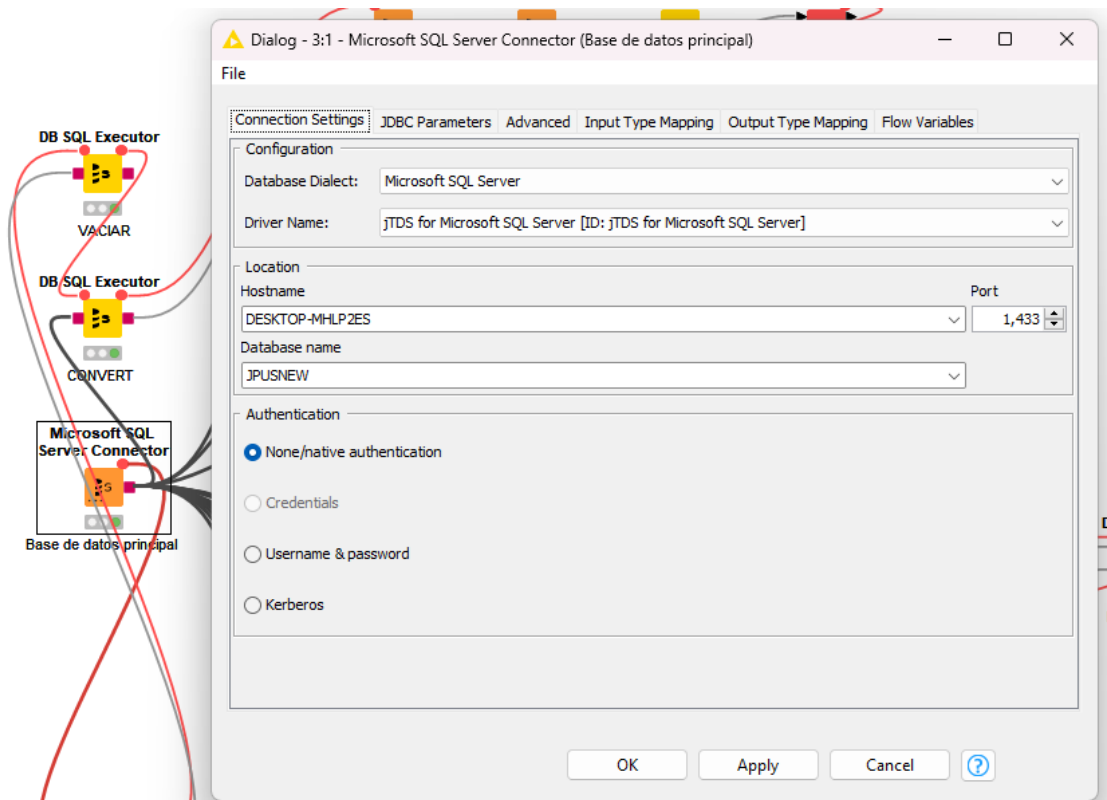


Figura 28 Obtención de data source

Fuente: los autores

3.7.1.2. CONEXIÓN DE BASE DE DATOS DEL DATA WARE HOUSE

De igual forma se obtiene la conexión hacia la base de datos de destino de la data warehouse, alojada igualmente el motor de base de datos de Microsoft

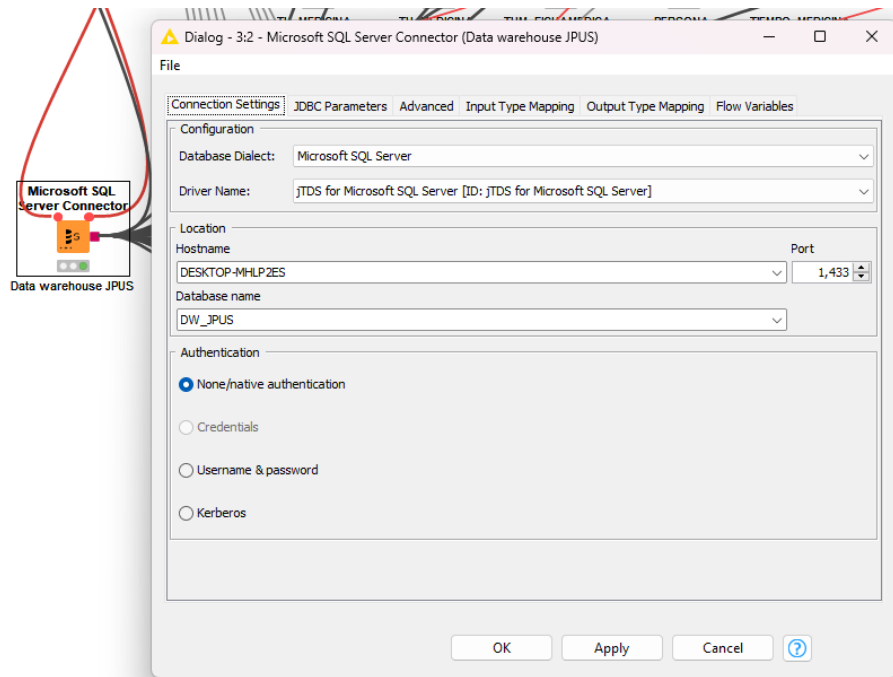


Figura 29 Conexión de data warehouse

Fuente: Los autores

3.7.1.3. LIMPIEZA DEL DATA WAREHOUSE

Como parte de las buenas prácticas realizadas en el proceso ETL, se añade un nodo SQL Executor que vacía la base de datos final cada que se ejecuta una nueva carga, evitando errores de duplicidad de información, así como errores en la carga por llaves primarias y foráneas.

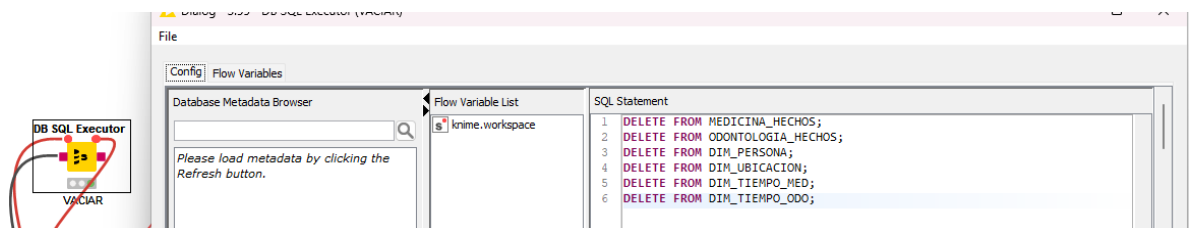


Figura 30 vaciado de base de datos final

Fuente: Los autores

3.7.1.4. CARGA DE LA DIMENSION PERSONA

En la carga de la dimensión se selecciona la tabla persona de la base de datos operacional, se seleccionan los campos de ID, edad y sexo, se transforma la edad de varchar a decimal para poder realizar cálculos una vez cargado en la dimensión.

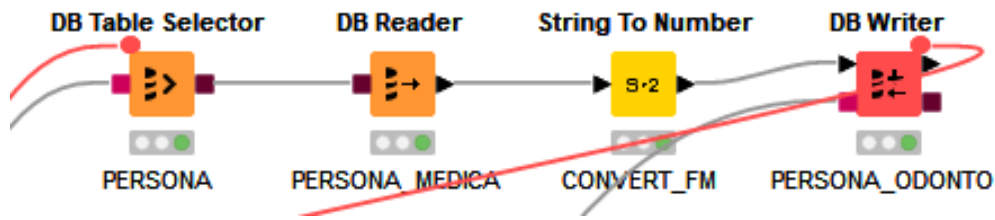


Figura 31 Carga de dimensión persona

Fuente: los autores

3.7.1.5. CARGA DE LA DIMENSION UBICACIÓN

Se selecciona la tabla ubicación para su carga en la dimensión, con los campos de ID, cantón y parroquia, que son renombrados en la ejecución de la sentencia SELECT para no tener inconvenientes a la hora de cargarlos en la base de datos de destino.

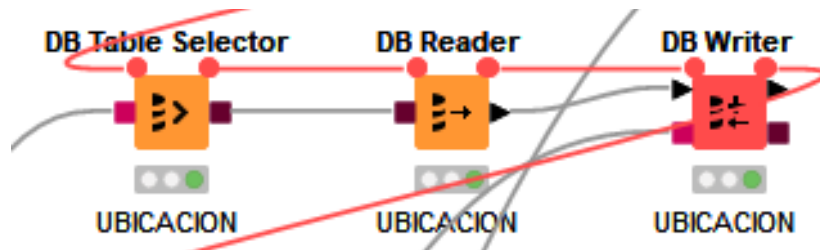


Figura 32 Carga de dimensión Ubicación

Fuente: Los autores

3.7.1.6. CARGA DE LA DIMENSION TIEMPO MEDICINA

La carga de la dimensión tiempo para medicina realiza la separación de las fechas en año, mes y día, con el fin de establecer la granularidad correspondiente, proceso que se realiza usando el nodo de cell splitter, así mismo se usa el column rename para renombrar las nuevas columnas agregadas.

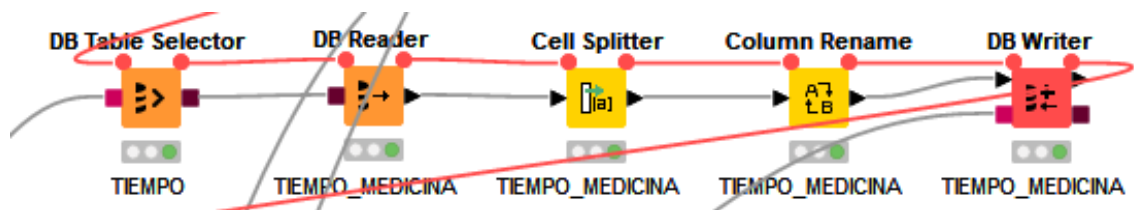


Figura 33 Carga de dimensión tiempo medicina

Fuente: los autores

3.7.1.7. CARGA DE LA DIMENSION TIEMPO ODONTOLOGIA

De la misma forma que en la dimensión anterior se realiza el proceso de división del campo de fecha y su renombramiento para poder establecer la granularidad.

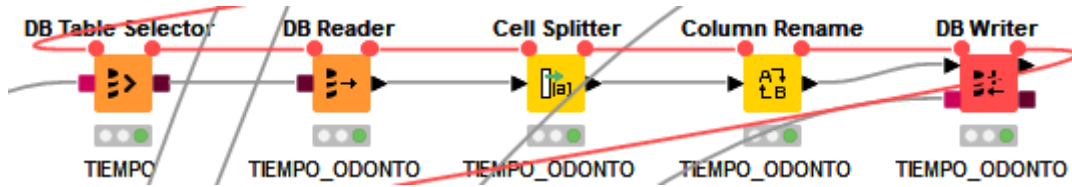


Figura 34 Carga de dimensión tiempo odontología

Fuente: Los autores

3.7.1.8. CARGA DE PERSONAS EN LA TABLA DE HECHOS DE MEDICINA

Para la tabla de hechos de medicina primero se selecciona a todos los pacientes, se cargan en la tabla de hechos, posteriormente con estos datos se usa un nodo joiner que se compara con la dimensión persona y establece que personas de dicha dimensión se cargan, así mismo se carga el campo de edades que será usado para realizar medidas.

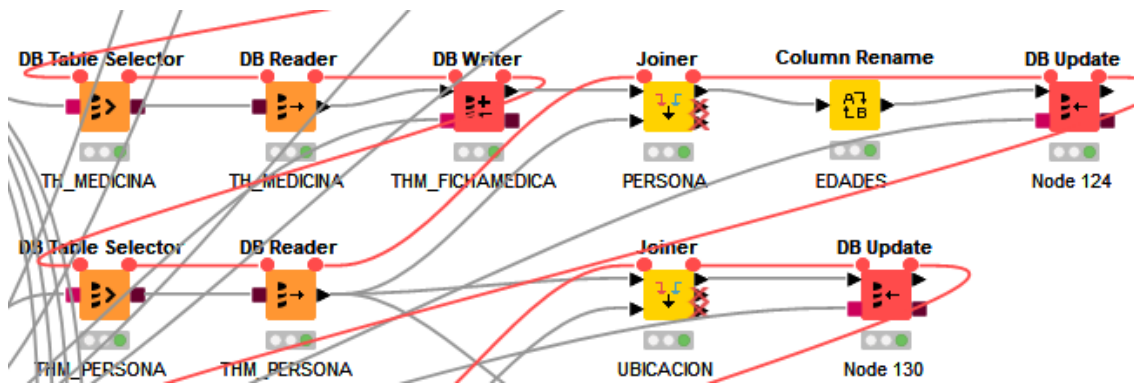


Figura 35 Carga de personas medicina en hechos medicina

Fuente: los autores

3.7.1.9. CARGA DE UBICACIONES EN LA TABLA DE HECHOS DE MEDICINA

Para la carga de ubicación como llave foránea en la tabla de hechos se compara con el nodo joiner de persona para saber a qué columnas pertenece ubicación, evitando así duplicidad y falta de integridad en los datos cargados.

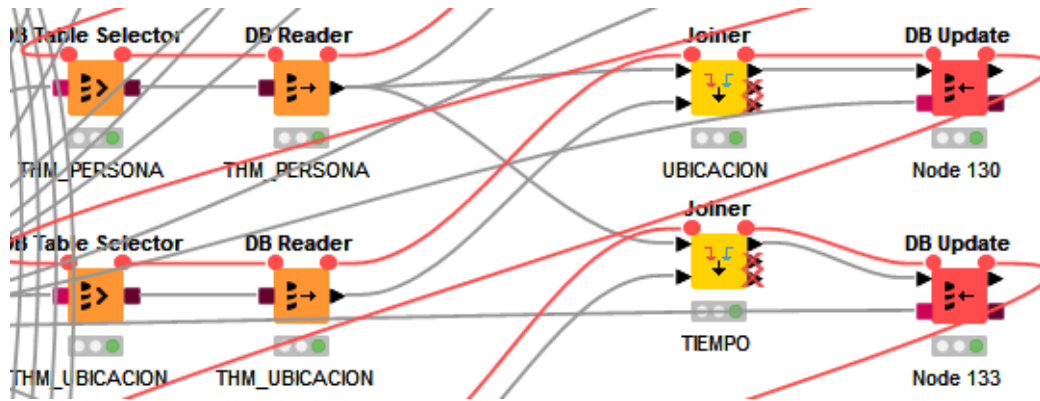


Figura 36 Carga de ubicación en hechos medicina

Fuente: Los autores

3.7.1.10. CARGA DE TIEMPO EN HECHOS DE MEDICINA

Para la carga de esta última dimensión, de igual forma se compara con un nodo joiner para saber a que columna corresponde la llave foránea, garantizando la carga exitosa de los datos.

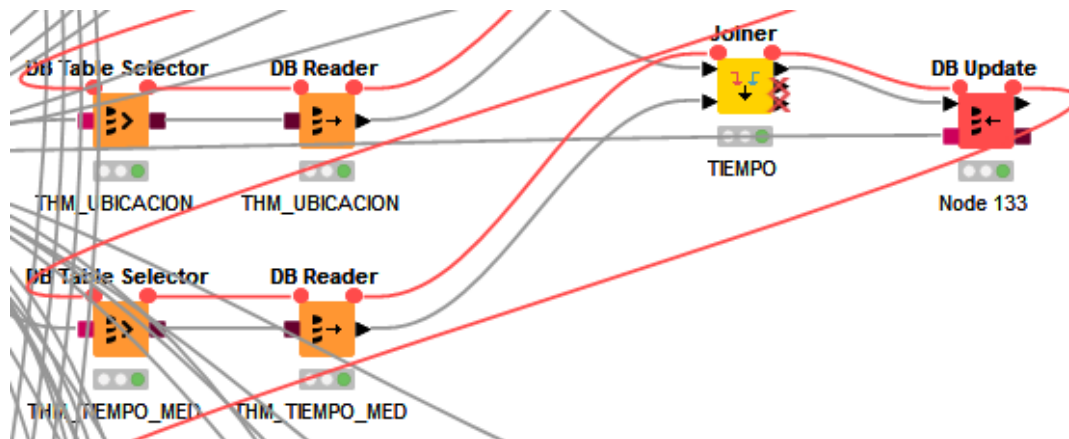


Figura 37 Carga de tiempo en hechos medicina

Fuente: Los autores

3.7.1.11. CARGA DE PESO EN HECHOS DE MEDICINA

Este campo corresponde a las medidas que se realizarán para la visualización de datos en dashboards, por ello para garantizar la integridad de la información, a través del nodo string manipulation se eliminan chars en el campo de peso, para luego poder transformarlo a decimal y colocarlo en la tabla de hechos.

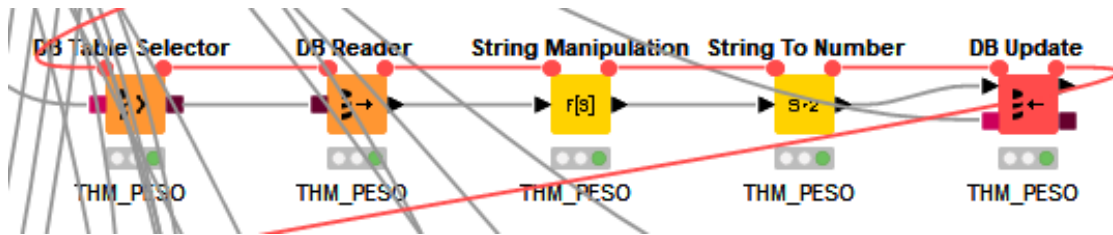


Figura 38 Carga de peso en hechos medicina

Fuente: Los autores

3.7.1.12. CARGA DE DURACION DE EMBARAZO EN HECHOS MEDICINA

Para la carga de este campo que se usara para determinar la cantidad de neonatos prematuros, se selecciona el campo de duración de embarazo, se realiza un joiner entre la base de datos operacional y la base de datos del data warehouse para saber la correspondencia de las semanas de embarazo con los distintos pacientes, se realiza también la remoción de chars para garantizar la integridad de la información, se convierte en decimal y finalmente se carga en la tabla de hechos.

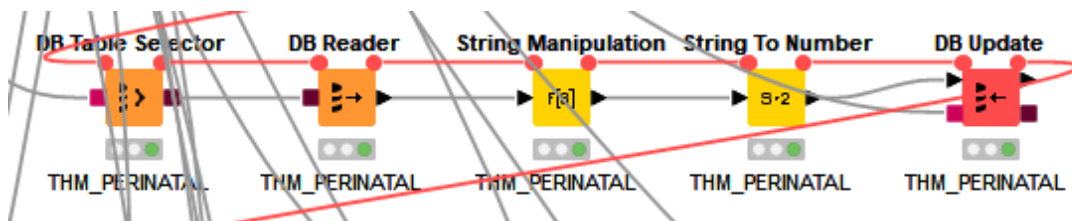


Figura 39 Carga de duración de embarazo en hechos medicina

Fuente: Los autores

3.7.1.13. CARGA DE TOMÓ SENO Y TOMÓ FORMULA EN HECHOS MEDICINA

En la carga de este campo se realiza un join entre la entre la base de datos operacional y la base de datos del data warehouse para corresponder las columnas con sus respectivas llaves foráneas y se carga en la tabla de hechos.

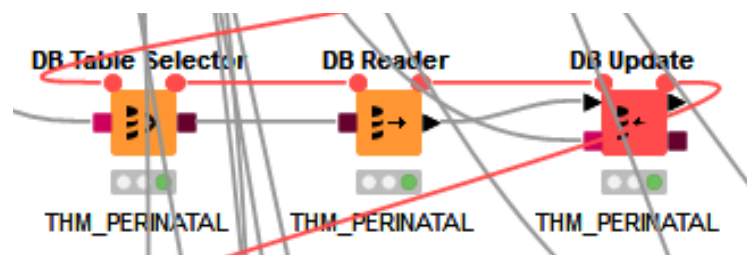


Figura 40 tomo seno y tomo formula en hechos medicina

Fuente: Los autores

3.7.1.14. CARGA DE TIPO DE ALIMENTACION EN HECHOS MEDICINA

Para este campo, primero en el nodo de SELECT se cambia las siglas por números para identificar el tipo de alimentación, posterior a esto a través de un nodo de string manipulation se transforman los números en palabras identificables para cargarlos en la tabla de hechos.

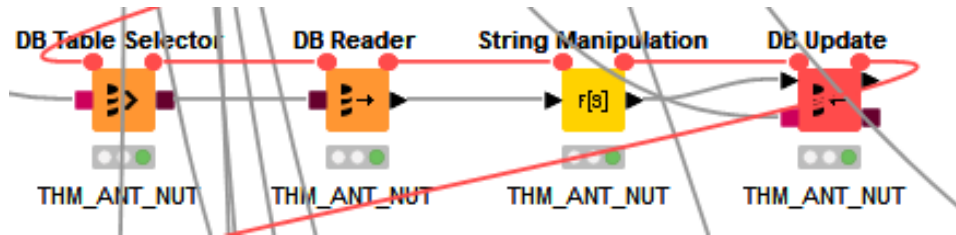


Figura 41 Carga de tipo de alimentación en hechos medicina

Fuente: Los autores

3.7.1.15. CARGA DE MALNUTRICION EN HECHOS MEDICINA

Se selecciona la tabla de antecedente nutricional, de la que se usa estatura, peso e IMC, debido a que estos campos son del tipo varchar, primero se los pasa por un nodo string to number para poder hacer el cálculo del índice de masa corporal y por último a través de un nodo rule engine comparar con varios if anidados que tipo de enfermedad de mal nutrición padece la persona.

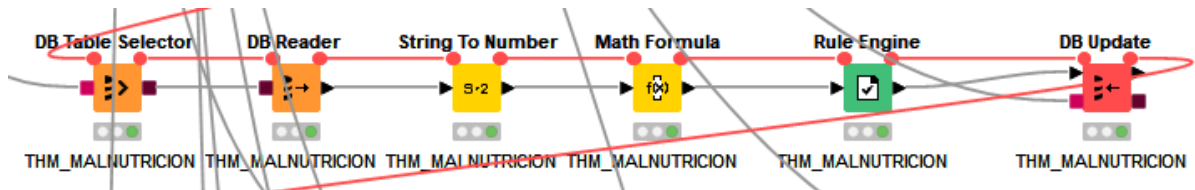


Figura 42 Carga de mal nutrición en hechos medicina

Fuente: Los autores

3.7.1.16. CARGA DE PERSONAS EN HECHOS ODONTOLOGIA

Para la carga inicial en la tabla de hechos de odontología se selecciona la tabla de ficha odontológica y la dimensión de personas, junto con un join se identifica las columnas correspondientes para cargar en los hechos.

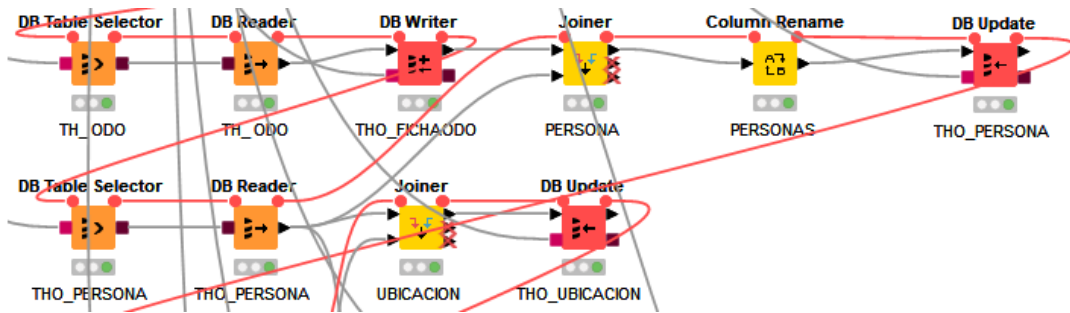


Figura 43 Carga de personas en hechos odontología

Fuente: Los autores

3.7.1.17. CARGA DE UBICACIÓN EN HECHOS ODONTOLOGIA

Para la carga de ubicación en la tabla de hechos se realiza un join entre la dimensión ubicación y hechos odontología para corresponder las llaves foráneas a las distintas columnas, evitando errores de gar y duplicidad de datos.

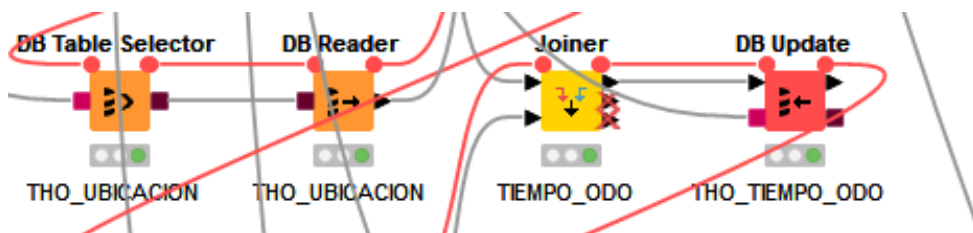


Figura 44 Carga de ubicación en hechos odontología

Fuente: Los autores

3.7.1.18. CARGA DE TIEMPO EN HECHOS ODONTOLOGIA

Para la carga de la dimensión tiempo como llave foránea en la tabla de hechos se realiza un join que comprara a que persona pertenece cada fecha, cargando así de manera eficaz la información a usar.

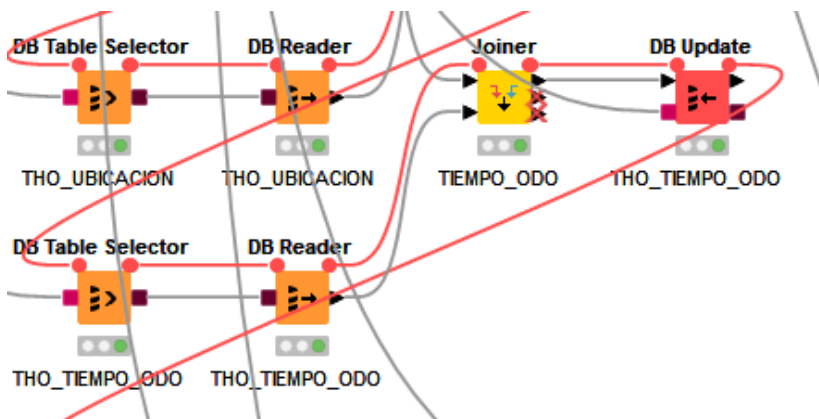


Figura 45 Carga de tiempo en hechos odontología

Fuente: Los autores

3.7.1.19. CARGA DE DIAGNOSTICO EN HECHOS ODONTOLOGIA

Se selecciona la tabla tratamiento, posteriormente se renombra para evitar errores en la carga de datos, este proceso se realiza con un nodo column rename.

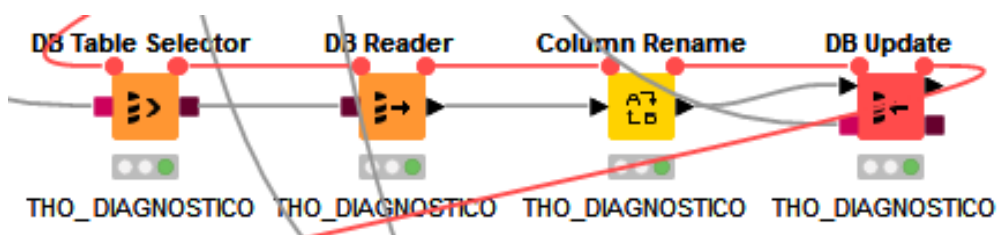


Figura 46 Carga de diagnóstico en hechos medicina

Fuente: Los autores

3.7.1.20. CARGA DE TRATAMIENTO EN HECHOS ODONTOLOGIA

En la carga de este campo se selecciona el código de tratamiento de su correspondiente tabal para cárgalos en los hechos.

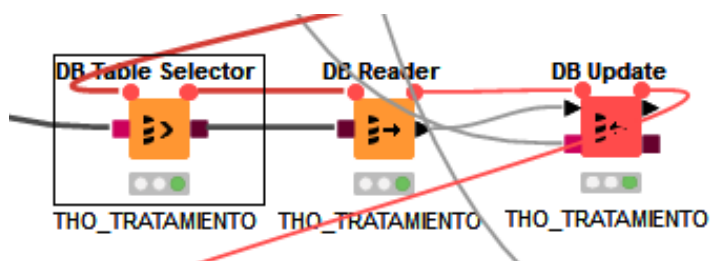


Figura 47 Carga de tratamiento en hechos odontología

Fuente: Los autores

3.7.1.21. CARGA DE CPO EN HECHOS ODONTOLOGIA

Para este campo es necesario seleccionar DP, DC, DO, y usarlos en un nodo de math fórmula para emplear el cálculo correspondiente y cárgalo en los hechos.

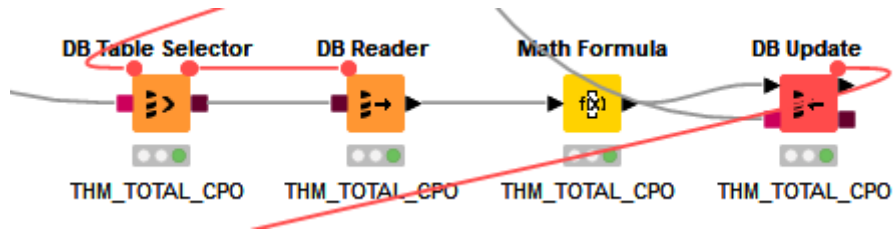


Figura 48 Carga de CPO en hechos odontología

Fuente: Los autores

3.7.1.22. CARGA DE PROBLEMAS BUCALES EN HECHOS ODONTOLOGIA

Varios campos corresponden a problemas bucales, como: Mal oclusión, fluorosis y enfermedad periodontal, todos los mencionados son transformados en nodos de string manipulation para transformar sus siglas en palabras, por último, se renombran las columnas y se cargan en los hechos.

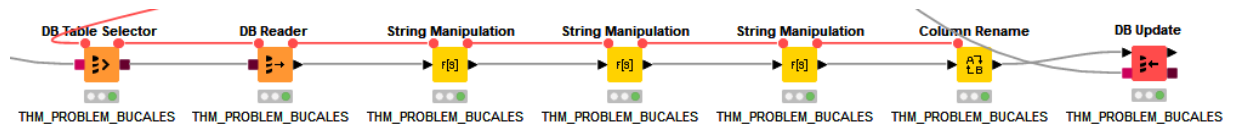


Figura 49 Carga de problemas bucales en hechos odontología

Fuente: los autores

3.7.1.23. MODELO FINAL DEL ETL

El modelo final del ETL corresponde a la ejecución secuencial y ordenada de todo el flujo de trabajo, debido a que, si no se usa la secuencia definida en la figura 40, al momento de hacer la carga nos mostrarán varios errores en llaves foráneas y datos inexistentes, por ello no se debe alterar dicha secuencia para el correcto funcionamiento del Work Flow.

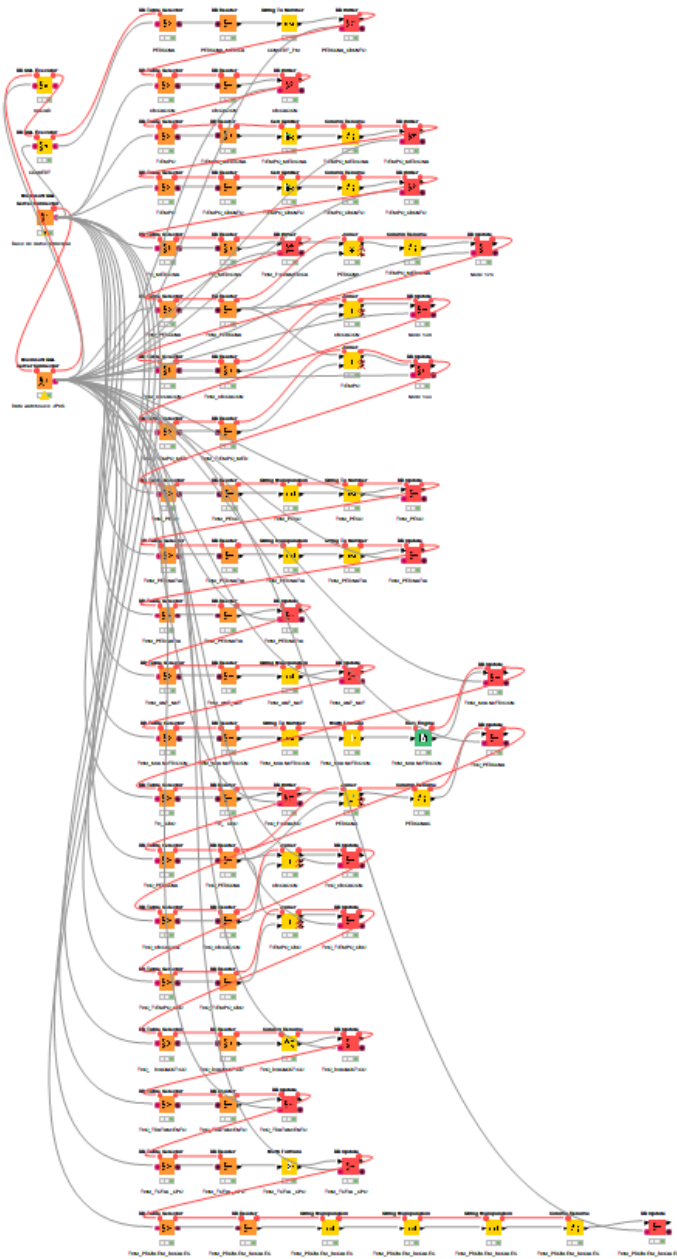


Figura 50 Modelo ETL final

Fuente: los autores

3.7.2. DISEÑO DEL BI

La solución de inteligencia de negocios aborda la visualización gráfica y análisis de los datos plasmados de manera general.

Los reportes generados dentro de los dashboards presentados para la solución BI en los módulos de medicina y odontología cumplen con los requerimientos presentados por parte

de los directivos del proyecto, es decir están basados en las preguntas de negocio, adaptados para facilitar el entendimiento y análisis de los gráficos.

3.7.2.1. SOLUCIÓN

La construcción de una solución de Business Intelligence resulta de gran ayuda para el proyecto "JUNTOS POR UNA SONRISA" en los módulos de medicina y odontología, ya que esta solución puede permitir una mejor gestión y análisis de los datos relacionados con los pacientes y los tratamientos dentales y médicos que se proporcionan. Al consolidar estos datos en una sola plataforma, los profesionales de la salud pueden obtener una toma de decisiones intuitiva, una visión más completa y detallada de la información del paciente, permitiendo mejorar la precisión del diagnóstico y el tratamiento.

Para la creación de gráficos se utilizó la herramienta de Power BI, como primer paso se crea la conexión de base de datos seleccionando la opción de obtención de datos SQL Server además de especificar el servidor y Base de datos de donde se obtendrá datos para la creación de los Dashboards.



Figura 51 Conexión de Power BI a MSSQL

Fuente: los autores.

Después de establecer la conexión de datos, se procede a la selección de las dimensiones y tablas de hechos que se utilizarán en la creación y visualización de los diferentes gráficos. Este proceso se realiza de forma independiente para cada uno de los dashboards creados, uno para Medicina y otro para Odontología, en función de las necesidades específicas de los administrativos del proyecto.

PERSONA_ID	TIEMPO_ID_MED	UBICACION_ID	ID_FICHA_MEDICA	PERSONA_MED	TASA_PESO	PROMEDIO_NEURATOS_PREMATUROS	CONSUMO_LECHE_MATERNA	CONSUMO_FORMULA	FROM_ALIMENTOS_CONSUMIDOS	FROM_MALNUTRICION	FROM_EDADES	PREMEDI	
2	2	2	2	2	69.2	2	40 S	S	VARIAO	DEGRADE ACCEPTABLE	10	NORMA	
5	5	5	5	5	28.5	5	38 S	N	VARIAO	DEGRADE MODERADA	11	PREMA	
6	6	6	6	6	25.7	6	41 S	N	VARIAO	DEGRADE MODERADA	9	NORMA	
7	7	7	7	7	27	7	40 N	N	VARIAO	DEGRADE SEVERA	11	NORMA	
8	8	8	8	8	28.1	8	40 S	N	VARIAO	DEGRADE ACCEPTABLE	8	NORMA	
9	9	9	9	9	27	9	40 S	S	VARIAO	OBESIDAD TIPO I	12	NORMA	
10	10	10	7	7	35.4	10	38 S	S	VARIAO	DEGRADE SEVERA	11	NORMA	
11	11	11	8	8	34.5	11	40 N	N	VARIAO	PESO NORMAL	11	NORMA	
12	12	12	9	9	28	12	37 S	S	VARIAO	DEGRADE SEVERA	9	PREMA	
13	13	13	10	10	34	13	34	40 S	N	VARIAO	DEGRADE MODERADA	8	NORMA
14	14	14	11	11	42.6	14	38 S	S	VARIAO	OBESIDAD TIPO II	10	NORMA	
15	15	15	12	12	28	15	28	40 S	N	VARIAO	DEGRADE ACCEPTABLE	8	NORMA
16	16	16	13	13	37.9	16	37 S	N	VARIAO	DEGRADE SEVERA	8	PREMA	
17	17	17	14	14	27	17	28	38 N	N	VARIAO	DEGRADE ACCEPTABLE	7	PREMA
18	18	18	15	15	27.4	18	5	40 S	N	VARIAO	DEGRADE SEVERA	10	PREMA
19	19	19	16	16	20.5	19	38 S	N	VARIAO	DEGRADE SEVERA	7	NORMA	
20	20	20	17	17	34.4	20	40 N	N	VARIAO	DEGRADE MODERADA	14	NORMA	
21	21	21	18	18	34.5	21	5	40 S	N	VARIAO	DEGRADE ACCEPTABLE	12	PREMA
22	22	22	19	19	33.4	22	40 S	N	VARIAO	DEGRADE MODERADA	4	NORMA	
23	23	23	20	20	31	23	39 S	N	VARIAO	OTROS	10	NORMA	
24	24	24	21	21	27.7	24	40 S	N	VARIAO	DEGRADE SEVERA	8	NORMA	
25	25	25	22	22	27.7	25	28.9	40 S	N	VARIAO	DEGRADE SEVERA	10	NORMA
26	26	26	23	23	28.9	26	38.9	40 S	N	VARIAO	DEGRADE SEVERA	8	NORMA
27	27	27	24	24	24.7	27	40 S	N	VARIAO	DEGRADE SEVERA	2.8	NORMA	
28	28	28	25	25	32.9	28	40 S	N	VARIAO	DEGRADE SEVERA	10	NORMA	
29	29	29	26	26	38.5	29	40 S	N	VARIAO	DEGRADE SEVERA	5	NORMA	
30	30	30	27	27	28.9	30	40 S	N	VARIAO	DEGRADE SEVERA	9	NORMA	
31	31	31	28	28	45.7	31	38 S	N	VARIAO	PESO NORMAL	11	NORMA	
32	32	32	29	29	34.2	32	40 S	N	VARIAO	DEGRADE MODERADA	5	NORMA	
33	33	33	30	30	38	33	38 S	N	VARIAO	PESO NORMAL	7	NORMA	
34	34	34	31	31	34	34	40 S	S	VARIAO	DEGRADE MODERADA	0.02	NORMA	
35	35	35	32	32	38	35	40 N	S	VARIAO	DEGRADE ACCEPTABLE	0.02	NORMA	
36	36	36	33	33	9.7	36	40 S	S	VARIAO	DEGRADE SEVERA	0.02	NORMA	
37	37	37	34	34	34	37	40 S	N	VARIAO	DEGRADE SEVERA	2.2	NORMA	
38	38	38	35	35	24.7	38	38 N	N	VARIAO	DEGRADE SEVERA	9	PREMA	
39	39	39	36	36	64.3	39	33 S	S	VARIAO	OBESIDAD TIPO I	11	PREMA	
40	40	40	37	37	40	39 S	37 S	S	VARIAO	DEGRADE ACCEPTABLE	0.02	PREMA	
41	41	41	38	38	37.8	41	38 S	N	VARIAO	PESO NORMAL	8	NORMA	
42	42	42	39	39	17.8	42	40 S	S	VARIAO	DEGRADE MODERADA	4	NORMA	

Figura 52. Carga de datos medicina a Power BI

Fuente: los autores.

PERSONA_ID	TIEMPO_ID_ODO	UBICACION_ID	ID_ODONTO	PERSONA_ODO	FROM_ESDAD	FROM_DIAGNOSTICO	FROM_TRATAMIENTO	FROM_ENFER_PERIODONTAL	FROM_MAL_OCLUSION	FROM_FLUOROSIS	TOTA
284	284	284	220	284	7	K20	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
286	286	286	221	286	7		D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
288	288	288	222	288	12		D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
4	4	4	223	4	12		D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
280	280	280	228	280	33	FLUOROSIS	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
292	292	292	229	292	34	FLUORACION	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
287	287	287	230	287	13	FLUORACION	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
290	290	290	232	290	14	PLACA BLANDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
303	303	303	234	303	14	PLACA BLANDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
296	296	296	235	296	14	PLACA BLANDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
293	293	293	236	293	14	EXTRACCION DIENTE TEMPORAL	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
304	304	304	237	304	13	PLACA BLANDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
300	300	300	238	300	14	PLACA BLANDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
302	302	302	239	302	14	PLACA BLANDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
305	305	305	240	305	13	K305	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
307	307	307	241	307	12	K306	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
308	308	308	242	308	8	CARIES EN OCLUSAL	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
284	284	284	243	284	14	GINGIVITIS ASOCIADA A PLACA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
288	288	288	244	288	14	GINGIVITIS AGUDA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
283	283	283	245	283	13	GINGIVITIS	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
313	313	313	246	313	14	K302 K301	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
314	314	314	247	314	12	ALTERACIONES EN LA FORMACION DENTARIA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
315	315	315	248	315	11	CARIES DE DENTINA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
316	316	316	249	316	10	CARIES DE DENTINA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
319	319	319	250	319	10	CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
320	320	320	251	320	12	CARIES DE LA DENTINA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
323	323	323	252	323	11	CARIES DE DENTINA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
324	324	324	253	324	11	CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
325	325	325	254	325	12	CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
326	326	326	255	326	11	CARIES EN ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
327	327	327	256	327	10	GINGIVITIS ASOCIADA A PLACA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
328	328	328	257	328	11	GINGIVITIS ASOCIADA A PLACA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
329	329	329	258	329	12	CARIES DE DENTINA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
331	331	331	259	331	10	CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
332	332	332	260	332	10	CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
333	333	333	261	333	10	GINGIVITIS Y CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
334	334	334	262	334	11	GINGIVITIS Y CARIES DE DENTINA	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	
335	335	335	263	335	13	CARIES DE ESMALTE	D0120-EVALUACION PERIODICA ORAL	LEVE	ANGLE I	LEVE	

Figura 53. Carga de datos medicina a Power BI

Fuente: los autores.

Una vez establecida la conexión de datos, se procede a generar informes con diversos gráficos para su visualización y análisis. En este sentido, se han creado dos informes: uno para Medicina y otro para Odontología. En cada uno de ellos se presentan diferentes tipos de

gráficos, que permiten una mejor comprensión de la información, los cuales se detallan a continuación:

3.7.2.1.1. INFORME MEDICINA:

3.7.2.1.1.1 TARJETA TOTAL PERSONAS ATENDIDOS



Figura 54. Total, Personas atendidos.

Fuente: los autores.

La tarjeta TOTAL PERSONAS ATENDIDAS refleja el número total de personas registradas en el sistema informático del proyecto hasta la fecha del 6 de febrero de 2023, alcanzando un total de 718 personas atendidas. Es importante destacar que esta cifra puede variar, ya que el proyecto continúa en crecimiento y se siguen registrando datos. Cabe señalar que esta tarjeta es compartida entre los módulos de medicina y odontología, por lo que el número total que se muestra es la suma de pacientes atendidos en ambos módulos.

3.7.2.1.1.2 TARJETA TOTAL PACIENTES ATENDIDOS MEDICINA.



Figura 55. Total, Pacientes atendidos medicina.

Fuente: los autores.

En esta tarjeta se observa el total de pacientes atendidos por el módulo de medicina indicando que hasta la fecha se han ingresado en el sistema informático 147 fichas médicas.

3.7.2.1.1.3 TARJETA PROM EDADES.

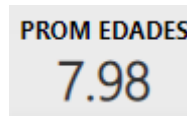


Figura 56. Prom edades.

Fuente: los autores.

Se puede apreciar un promedio general de edades dentro de los pacientes atendidos de las fichas ingresadas en medicina dando un promedio de 7.98

3.7.2.1.1.4 PERSONAS ATENDIDAS POR UBICACIÓN

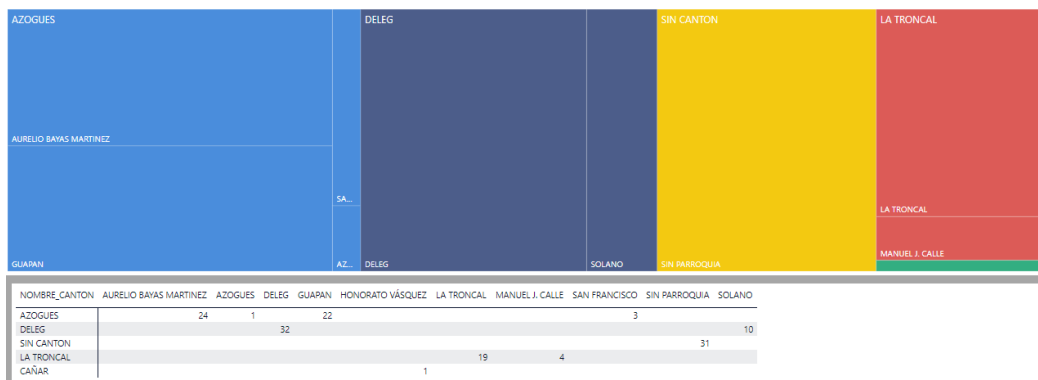


Figura 57. Personas atendidas por ubicación.

Fuente: los autores.

En el Treemap se puede observar la distribución de los pacientes atendidos según su ubicación geográfica (cantón y parroquia con granularidad fina de las parroquias que pertenecen a cada caso), de los pacientes atendidos por ubicación: 50 en el cantón Azogues, 42 en el cantón Déleg, 23 en el cantón La Troncal, 1 en el cantón Cañar y 31 que no registra ubicación.

3.7.2.1.1.5 TASA DE PESO

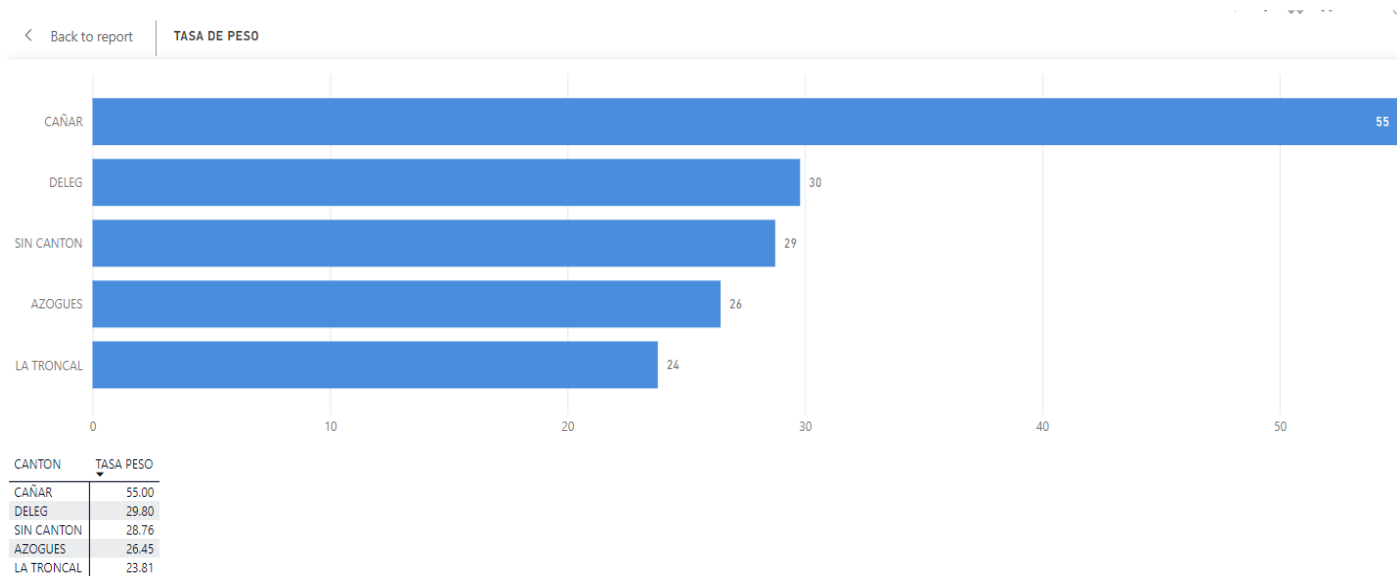


Figura 58. Tasa de peso

Fuente: los autores.

Dentro del grafico de barras apiladas se puede observar la tasa de peso por ubicación de los pacientes atendidos e ingresados dentro del sistema informático dando la parroquia de Honorato Vásquez una tasa de 55 KG siendo la tasa de peso más alta, y la parroquia San Francisco siendo la última de la lista con un peso de 8 KG hasta la fecha (13 de febrero del 2023) tomando en cuenta que los datos pueden variar debido al ingreso constante de fichas médicas.

3.7.2.1.1.6 MALNUTRICIÓN

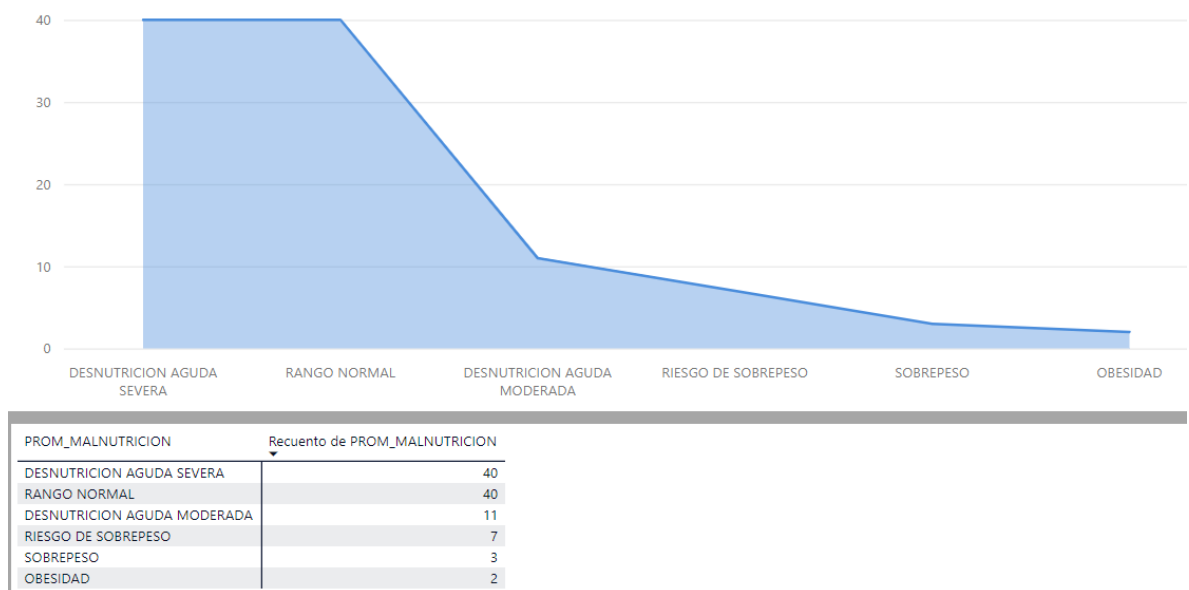


Figura 59. MALNUTRICIÓN

Fuente: Los autores

El gráfico de barras apiladas muestra el total de pacientes con problemas de malnutrición clasificándolos por desnutrición aguda severa con un total de 40, desnutrición aguda moderada dando un total de 11, rango normal con un total de 40, riesgo de sobrepeso con un total de 7, sobrepeso con un total de 3 y obesidad con un total de 2 pacientes, siendo rango normal y desnutrición aguda severa los problemas de malnutrición comunes dentro de los datos analizados.

3.7.2.1.1.7 MALNUTRICIÓN POR UBICACIÓN

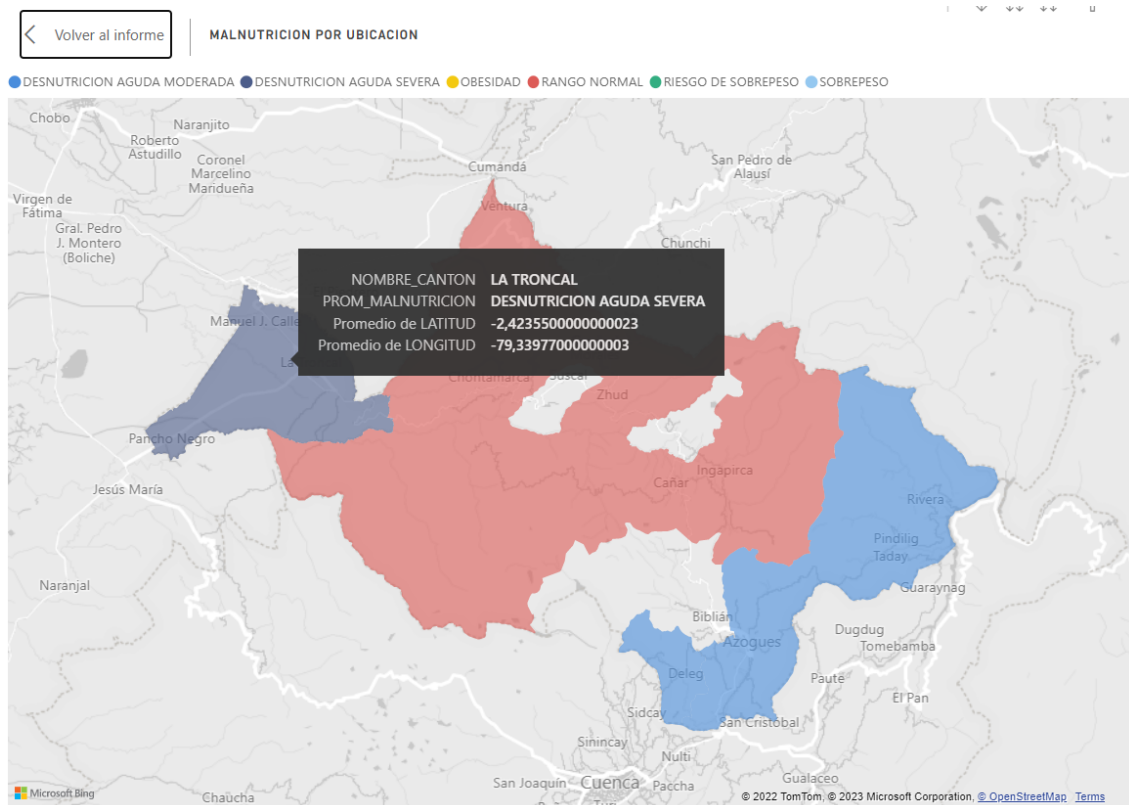


Figura 60. Malnutrición por ubicación

Fuente: Los autores.

El Mapa Coroplético permite visualizar el diagnóstico más común por su ubicación geográfica. Según los datos registrados, en el cantón Azogues, el problema más común de malnutrición es la desnutrición aguda moderada. En el cantón Biblián no se han registrado datos. En Cañar, el diagnóstico muestra un rango normal. En el cantón Deleg, el problema más común es la desnutrición aguda moderada, mientras que, en La Troncal, se ha identificado

el problema de malnutrición de desnutrición aguda severa. Este gráfico nos permite apreciar de manera visual los datos geográficos y sus correspondientes diagnósticos.

3.7.2.1.1.8 ALIMENTOS CONSUMIDOS

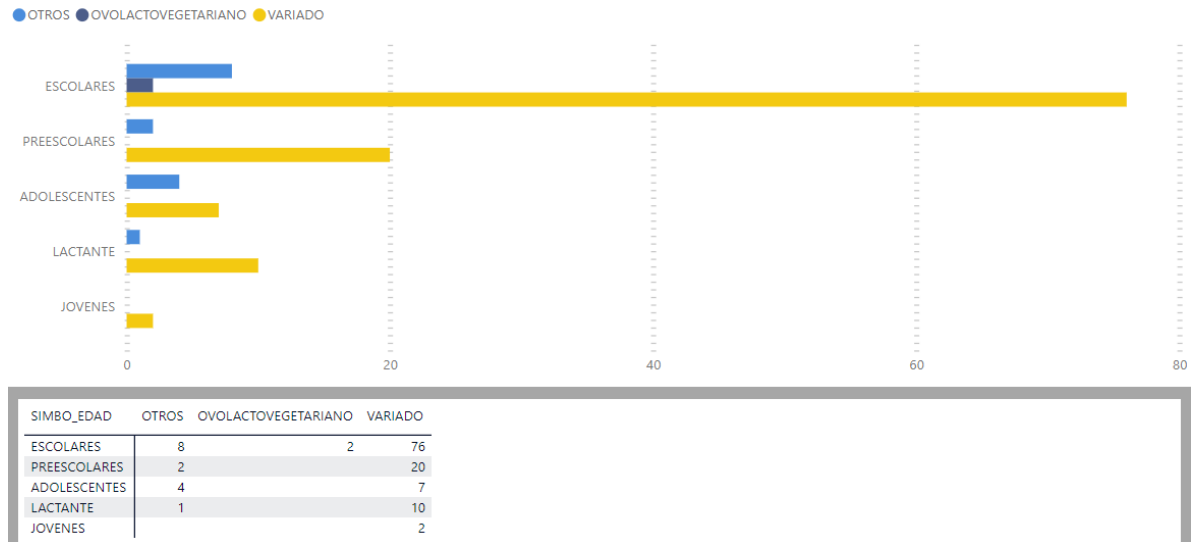


Figura 61. Alimentos consumidos.

Fuente: los autores.

El gráfico de barras agrupadas indica los alimentos consumidos por edad estando clasificada en lactantes, escolares, preescolares, adolescentes y adultos según los requerimientos de los administrativos del proyecto para el cual se desarrolló la solución BI dando este gráfico como los alimentos consumidos comunes: alimentos variados, ovolactovegetarianos y otros.

3.7.2.1.1.9 CONSUMO DE LECHE MATERNA Y FORMULA

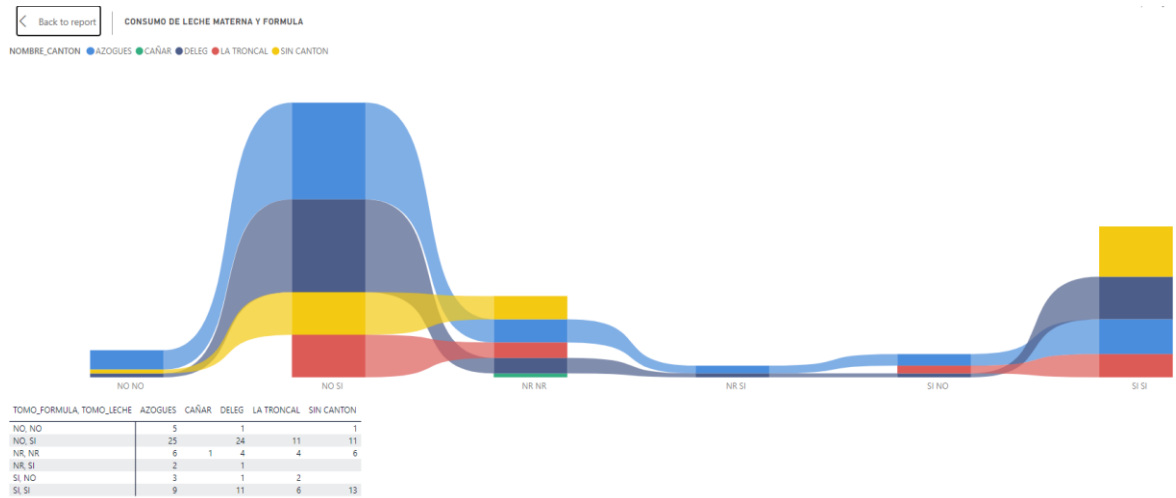


Figura 62. Consumo de leche materna y formula.

Fuente: Los autores.

En el gráfico de cintas se puede apreciar una comparación entre los campos de "tomo leche materna" y "tomo fórmula", presentando varias posibilidades: "NO, NO", es decir, no tomó ni leche materna ni fórmula; "NO, SI", que indica que no tomó leche materna pero sí fórmula; "NR, NR", en donde no se especifica si tomó leche materna o fórmula; "NR, SI", en donde no se especifica si tomó leche materna pero sí fórmula; "SI, NO", que indica que tomó leche materna pero no fórmula; y finalmente, "SI, SI", que indica que tomó tanto leche materna como fórmula. Además, estos datos se han agrupado por cantones para facilitar la comprensión y la abstracción de información presentada.

3.7.2.1.1.10 NEONATOS PREMATUROS

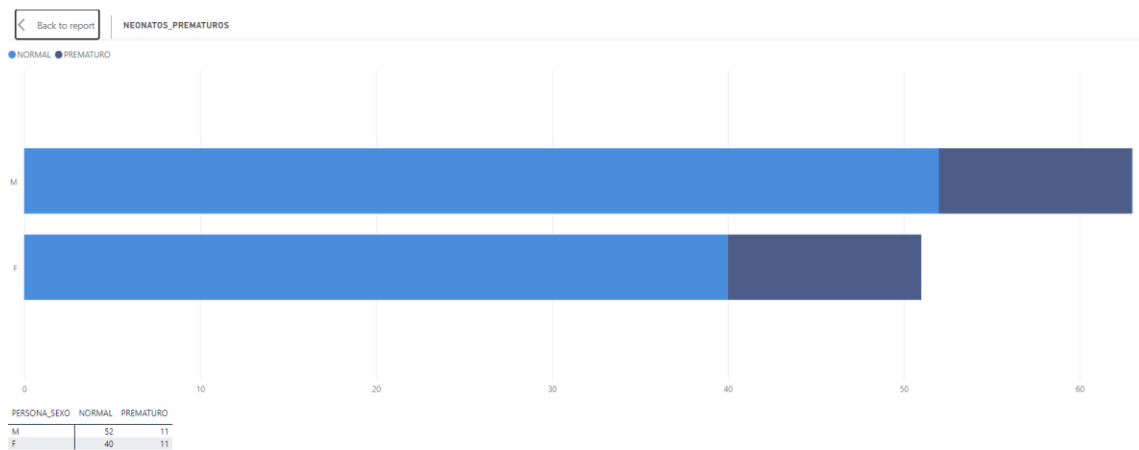


Figura 63. Neonatos prematuros

Fuente: Los autores.

Como ultimo grafico dentro del Dashboard de medicina, encontramos un gráfico de barras que representa la cantidad de neonatos normales y prematuros, segmentados por género. Esta información se obtuvo a través del proceso ETL. De acuerdo con el gráfico, se puede observar que hay 52 neonatos normales masculinos, 11 neonatos prematuros masculinos, 40 neonatos normales femeninos y 11 neonatos prematuros femeninos. Este gráfico proporciona una visión detallada y desglosada de los datos sobre neonatos atendidos en el proyecto.

3.7.2.1.2. INFORME ODONTOLOGIA:

3.7.2.1.2.1 PACIENTES ATENDIDOS ODONTOLOGIA



Figura 64 Pacientes atendidos en odontología

Fuente: Los autores

La figura presenta la cantidad total de pacientes que han sido atendidos por el departamento de odontología. Se puede observar que el número de pacientes ingresados y atendidos asciende a 472.

3.7.2.1.2.2 PROMEDIO DE EDADES



Figura 65 Promedio de edades en los pacientes

Fuente: Los autores

La figura muestra el promedio de edad en el que fueron atendidos los pacientes, en este caso es 11,99, es decir que la media es de 12 años aproximadamente, correspondiente a pacientes escolares.

3.7.2.1.2.3 PACIENTES ATENDIDOS POR CANTON Y AÑO

NOMBRE_CANTON	2021	2022	2023	Total	
AZOGUES	10	37	183	33	263
BIBLIAN	5	3			8
CAÑAR		7			7
DELEG	1	49	47		97
LA TRONCAL		8	25	16	49
SIN CANTON	1		40	7	48
Total	17	104	295	56	472

Figura 66 Matriz de pacientes por cantón y año

Fuente: Los autores

Este grafico permite visualizar a forma de matriz la segmentación existente por los distintos cantones de la provincia de cañar por los años en los cuales fueron atendidos los pacientes, mostrando también los totales por tiempo, se puede apreciar que la mayor cantidad atendida yace en el cantón Azogues con un total de 263, además que el año en el que más se dio consulta fue en el 2022.

3.7.2.1.2.4 PERSONAS ATENDIDAS POR UBICACIÓN

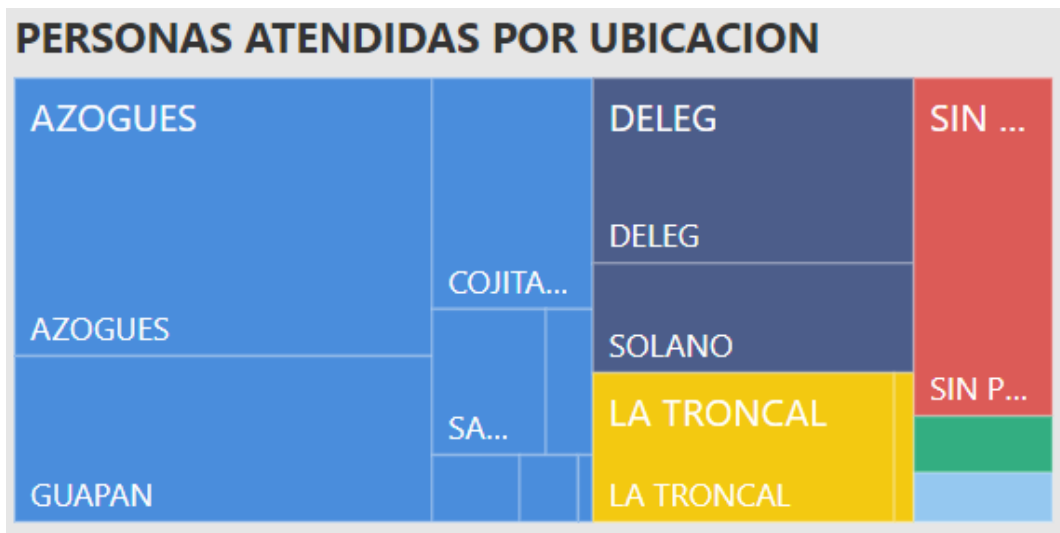


Figura 67 Treemap de personas atendidas por parroquia

Fuente: Los autores

El treemap expone con una granularidad más específica el total de personas que asistieron a consulta odontológica en las diferentes parroquias, separadas por colores que identifican el cantón al que pertenecen, por ejemplo, el cantón Azogues posee el color azul, el cantón La Troncal el color amarillo, Déleg morado, Biblián verde, Cañar celeste y aquellos que fueron ingresados sin ubicación poseen color rojo para sus parroquias.

3.7.2.1.2.5 AREAS DE TRATAMIENTOS

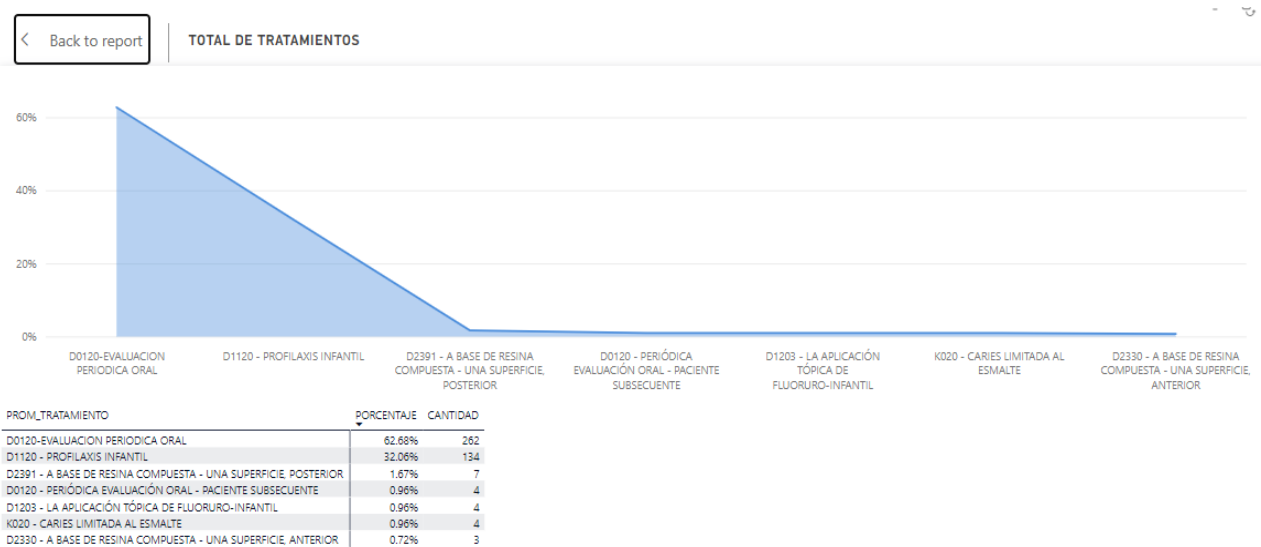


Figura 68 Grafico de tratamientos

Fuente: los autores

A través de los códigos establecidos para tratamientos, se define cuales se han realizado frecuentemente, la figura presenta predominancia con el código D0120 la evaluación periódica oral en el pico más alto con un total de 262.

3.7.2.1.2.6 TRATAMIENTOS POR CANTON

El mapa está diseñado de tal manera que cada código de paciente tiene una ubicación exacta en el mapa y se muestra mediante diferentes tonalidades de color. Por ejemplo, los códigos correspondientes a pacientes con ciertas características de salud pueden tener un color más oscuro, mientras que los códigos correspondientes a pacientes con otras características de salud pueden tener un color más claro.

Esto permite a los usuarios identificar fácilmente los tratamientos adecuados para cada grupo de pacientes y, por lo tanto, tomar decisiones informadas sobre la atención médica que se debe proporcionar. Además, la tonalidad del mapa puede variar según las selecciones y filtros que se apliquen, lo que significa que los usuarios pueden ajustar el mapa para mostrar solo la información relevante para una situación particular.

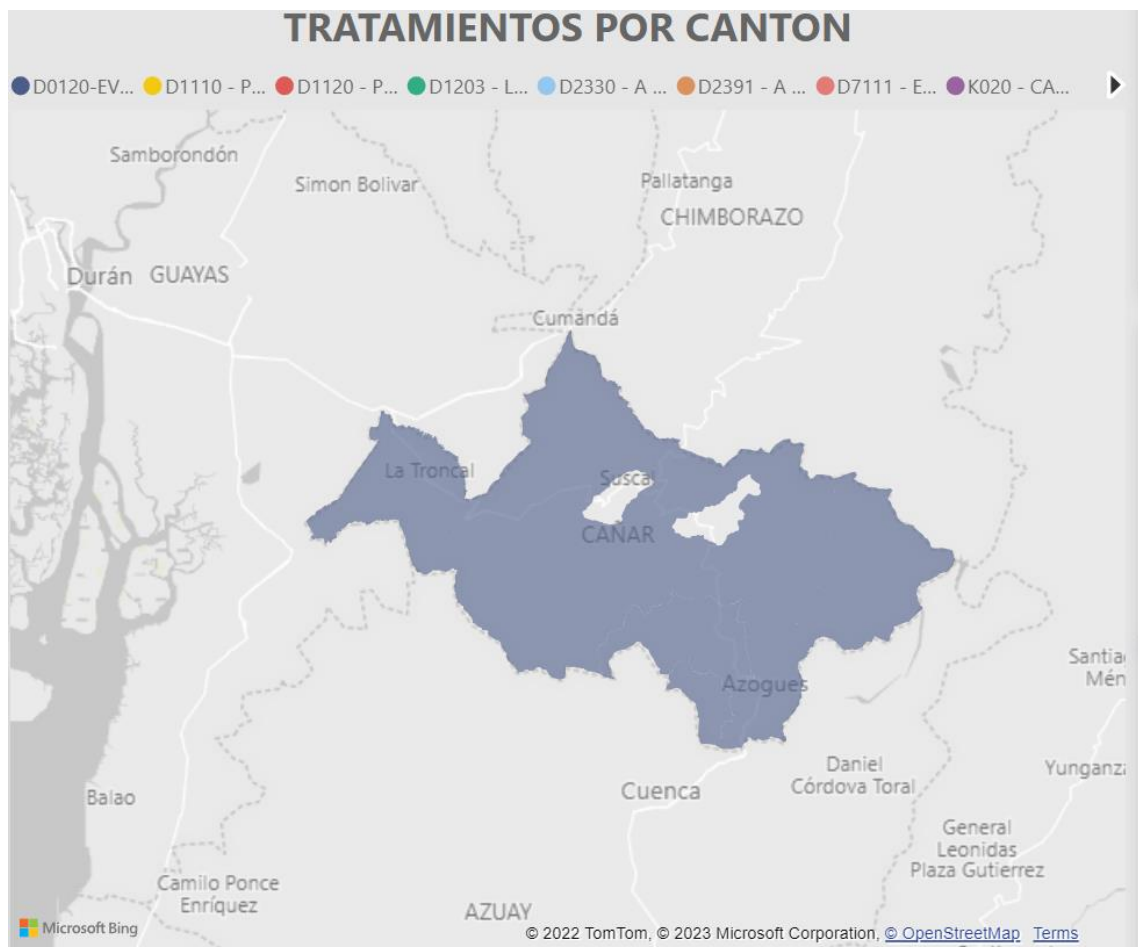


Figura 69. Mapa coroplético de tratamientos por cantón.

Fuente: Los autores

3.7.2.1.2.7 DIAGNOSTICOS COMUNES POR UBICACIÓN

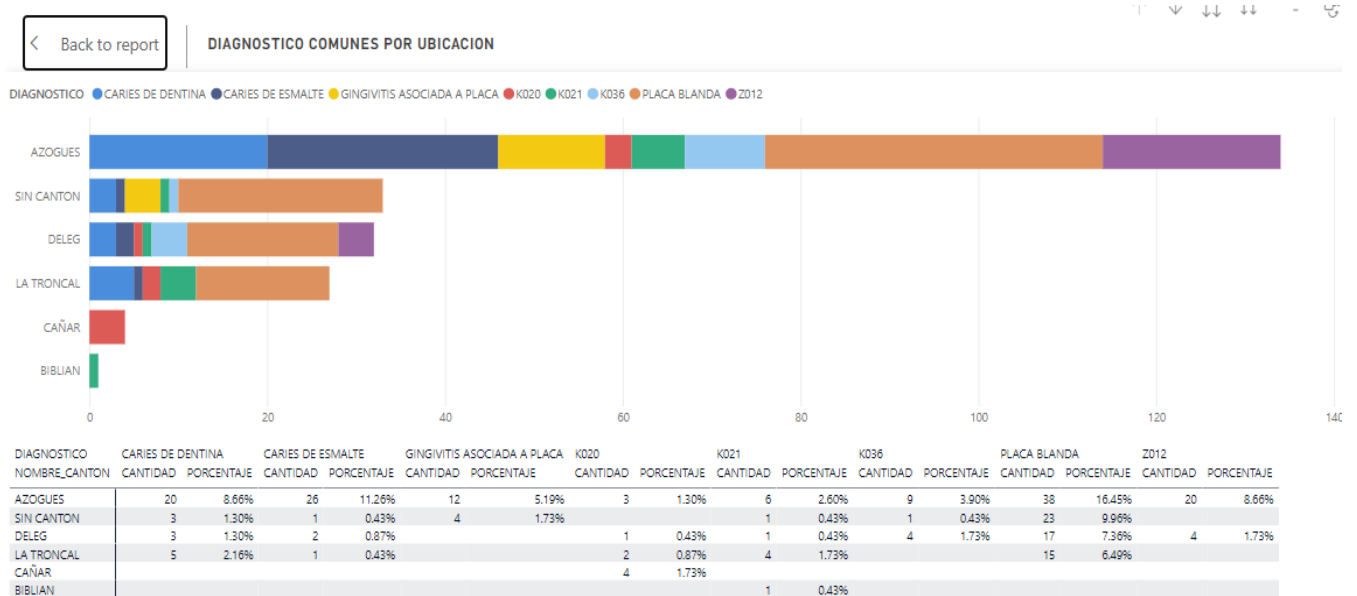


Figura 70. Gráfico de barras de diagnósticos comunes por ubicación.

Fuente: Los autores

Esta figura permite visualizar cuales diagnósticos son los más comunes por ubicación en la provincia del Cañar, como ejemplo en el cantón Azogues la caries de esmalte posee un el total más alto con 25 diagnósticos confirmados.

3.7.2.1.2.8 PROMEDIO CPO POR UBICACIÓN Y SEXO

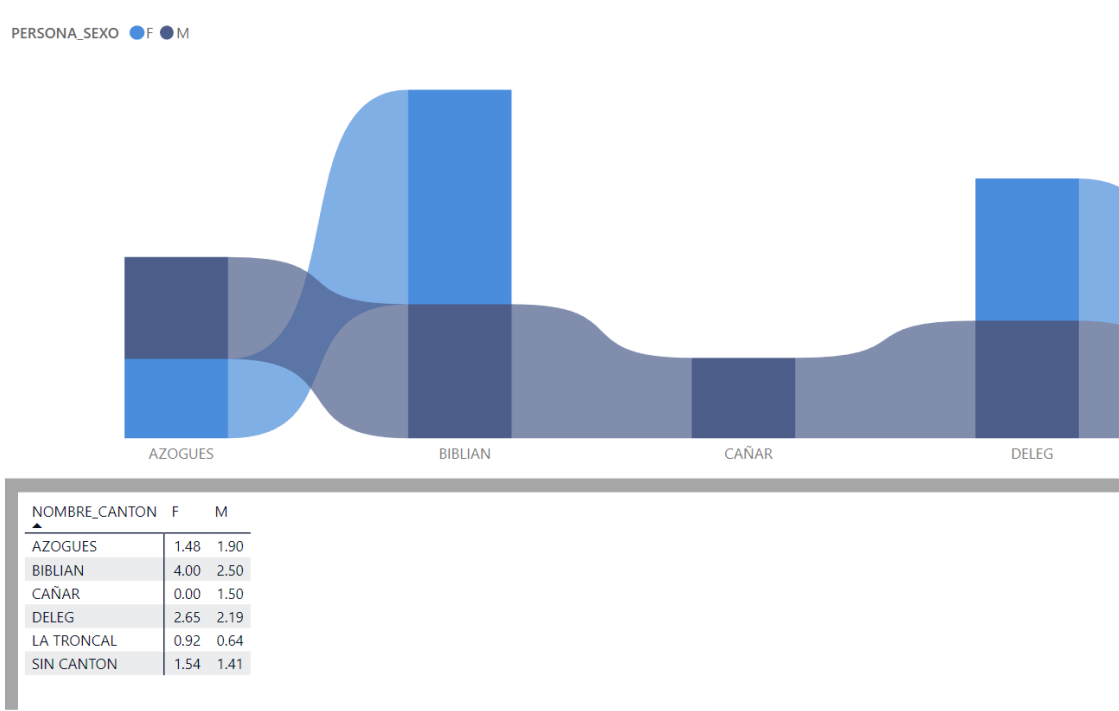


Figura 71. Gráfico de cintas de promedio de CPO por ubicación y sexo.

Fuente: Los autores

El promedio de CPO (Cariado, Perdido, Obturado) se muestra por ubicación y sexo, obteniendo como resultado que en el cantón Azogues el promedio en hombres es de 1.90 y en mujeres 1.48. En el cantón Biblián, el promedio en hombres es de 2.5 y en mujeres 4. Cañar no tiene datos registrados para mujeres, y para hombres el promedio es de 1.5. En Déleg, el promedio en hombres es de 2.19 y en mujeres 2.65. En La Troncal, el promedio en hombres es de 0.64 y en mujeres 0.92. Además, hay un promedio de 1.41 en hombres y 1.54 en mujeres que no tienen una ubicación registrada.

3.7.2.1.2.9 SALUD BUCAL

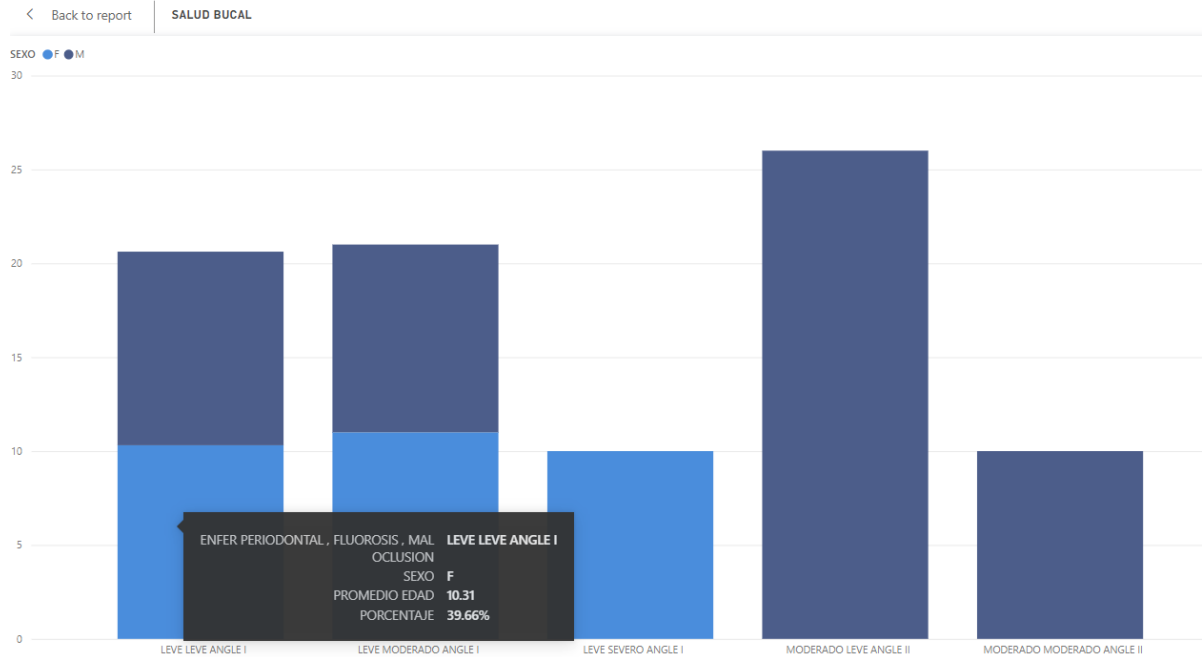


Figura 72. Gráfico de columnas de salud bucal.

Fuente: Los autores

El gráfico de columnas apiladas permite visualizar la enfermedad periodontal, fluorosis y mal oclusión como promedio por edad y sexo, y en qué grado afectan cada uno, de acuerdo a lo establecido por la especialidad de odontología. Esto permite analizar la salud bucal general de los pacientes, y se ha encontrado que la Enfermedad Periodontal Leve, la Fluorosis Leve y la Mal oclusión en Angle I son las condiciones más comunes, con un 39.66% de los datos obtenidos en mujeres y un 57.47% en hombres.

CONCLUSIONES

Investigando conceptos, metodologías y herramientas de Business Intelligence, se pudo fundamentar teóricamente el presente trabajo, lo que permitió conocer las técnicas de análisis de datos y visualización que se aplicó en el presente trabajo empleado a la toma de decisiones y la calidad de la atención médica y odontológica.

Se pudo concluir, a través de las distintas comparativas, que emplear Power BI como aplicación de inteligencia de negocios se acomoda a las necesidades del proyecto por su fácil adaptabilidad, Microsoft SQL Server por su escalabilidad, seguridad, rendimiento e integración con otras tecnologías de Microsoft, Knime por sus nodos predeterminados que

facilitan el flujo de trabajo, y la metodología Hefesto debido a su enfoque en la entrega de resultados valiosos, colaboración y comunicación efectiva. Todo esto permitió ayudar a garantizar que el Data Warehouse cumpla con los requisitos del trabajo y proporcione información valiosa para la toma de decisiones intuitivas.

Se evidencio que, el proceso ETL es esencial para migrar la información de la base de datos relacional del sistema informático "JUNTOS POR UNA SONRISA" hacia el almacén de datos. Entre las tareas para la extracción, transformación y carga de datos, se usó la herramienta Knime que aseguro de manera eficiente la calidad e integridad de datos, mediante operaciones de limpieza y estandarización.

Se implementó una solución de Business Intelligence en el software "JUNTOS POR UNA SONRISA", lo que permitió generar beneficios a través de diferentes gráficos estadísticos, como mapas coropléticos, gráficos de barras, matrices e indicadores. Esta implementación ha abierto la posibilidad de realizar investigaciones sobre problemas relacionados con la malnutrición y la salud bucal, ya que los cuadros de mando dinámicos impulsados por la toma de decisiones permiten realizar análisis detallados. Además de mejorar la eficiencia y efectividad del proyecto.

RECOMENDACIONES

- Es fundamental comprender que los datos precisos y completos son esenciales para el éxito de cualquier sistema de Business Intelligence. Por lo tanto, nuestra recomendación sería verificar la exactitud de los datos ingresados, mantener una consistencia en la forma en que se registran los datos, actualizar los registros regularmente y asegurarse de que los datos estén completos y sean relevantes para los análisis que se deseen realizar.
- Se recomienda que se considere la inclusión de información sobre otros campos de los pacientes en el cuadro de mando general. Esta información puede proporcionar una visión más completa de la salud de los pacientes y, por lo tanto, permitir un análisis más profundo y preciso de la población de pacientes en su conjunto. Al incluir campos como antecedentes médicos previos, historial familiar, condiciones crónicas y hábitos de vida, el cuadro de mando general puede ofrecer una visión más completa del perfil de salud de cada paciente, permitiendo a los profesionales de la salud tomar decisiones más informadas y personalizadas.
- Se recomienda utilizar el dashboard general para aprender y obtener conocimiento sobre las fichas médicas y odontológicas en la provincia. En particular, se debe enfatizar el

uso del cuadro de mando integral para monitorear los casos de malnutrición y enfermedades bucales, ya que estas son áreas críticas que requieren una atención y seguimiento constantes.

- Para futuros cambios que requieran los administrativos del proyecto se recomienda seguir usando la metodología Hefesto y el esquema constelación, esto se debe a que esta arquitectura de BI es simple y se adapta perfectamente al proyecto, lo que permite un posterior análisis de datos más eficiente y efectivo. Se debe considerar también la implementación de procesos de control de calidad, fomentar la colaboración y el intercambio de conocimientos entre los diferentes actores involucrados en el proyecto.

BIBLIOGRAFÍA:

- [1] V. Vilca, J. Luis, N. Chamorro, J. Isaac, and L. De, "FACULTAD DE INGENIERÍA Y ARQUITECTURA ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS Business Intelligence para la gestión de ventas de productos odontológicos: caso VERO DENT," 2021.
- [2] E. P. Madrid, "LOS INSUMOS INVISIBLES DE DECISIÓN: DATOS, INFORMACIÓN Y CONOCIMIENTO," *Nº*, vol. 11, pp. 183–196, 2008.
- [3] G. Martínez, "Vista de Importancia De Los KPI De Logística En Las PYMES," *Publicaciones e investigación Unad*, vol. 15, no. 2, 2021, Accessed: Jul. 10, 2022. [Online]. Available: <https://hemeroteca.unad.edu.co/index.php/publicaciones-e-investigacion/article/view/5548/5285>
- [4] M. A. Romero and J. R. García, "COMPARISON OF OPTIONS FOR INTELLIGENCE BUSINESS IN MAJOR SYSTEMS MANAGERS MARKET DATABASES," *REVISTA ECONOMÍA & ADMINISTRACIÓN*, vol. 6, no. COMPARACIÓN DE OPCIONES PARA INTELIGENCIA DE NEGOCIOS EN LOS PRINCIPALES SISTEMAS GESTORES DE BASES DE DATOS DEL MERCADO, pp. 5–20, 2016.
- [5] SISENSE, "OLAP, ROLAP, MOLAP, HOLAP." <https://www.sisense.com/glossary/olap/> (accessed Jul. 05, 2022).
- [6] T. de Maestría, G. Gonzalo, V. Sánchez, U. Autonoma, and D. E. Bucaramanga, "DISEÑO Y CONSTRUCCIÓN DE UN CUBO MULTIDIMENSIONAL OLAP USANDO SOFTWARE LIBRE: ESTUDIO DE CASO SECTOR COMERCIAL REPUESTOS AUTOMOTRIZ CÉSAR AUGUSTO POVEDA FLORÉZ," 2018.
- [7] C. V. R. E. M. C. J. M. Feliciano Morales Angelino, "Procesamiento Analítico con Minería de Datos," *Revista Iberoamericana de las ciencias computacionales e informática*, vol. 5, no. Procesamiento Analítico con Minería de Datos, 2016, Accessed: Jul. 05, 2022. [Online]. Available: <https://www.reci.org.mx/index.php/reci/article/view/40/176>
- [8] Sepúlveda Aguirre Jovany, *RETOS DE LA INVESTIGACIÓN EN INGENIERÍA DE SISTEMAS: APLICACIONES, HERRAMIENTAS Y DESARROLLOS*, 1st ed., vol. 1. Medellin: Editorial Coruniamericana, 2018.

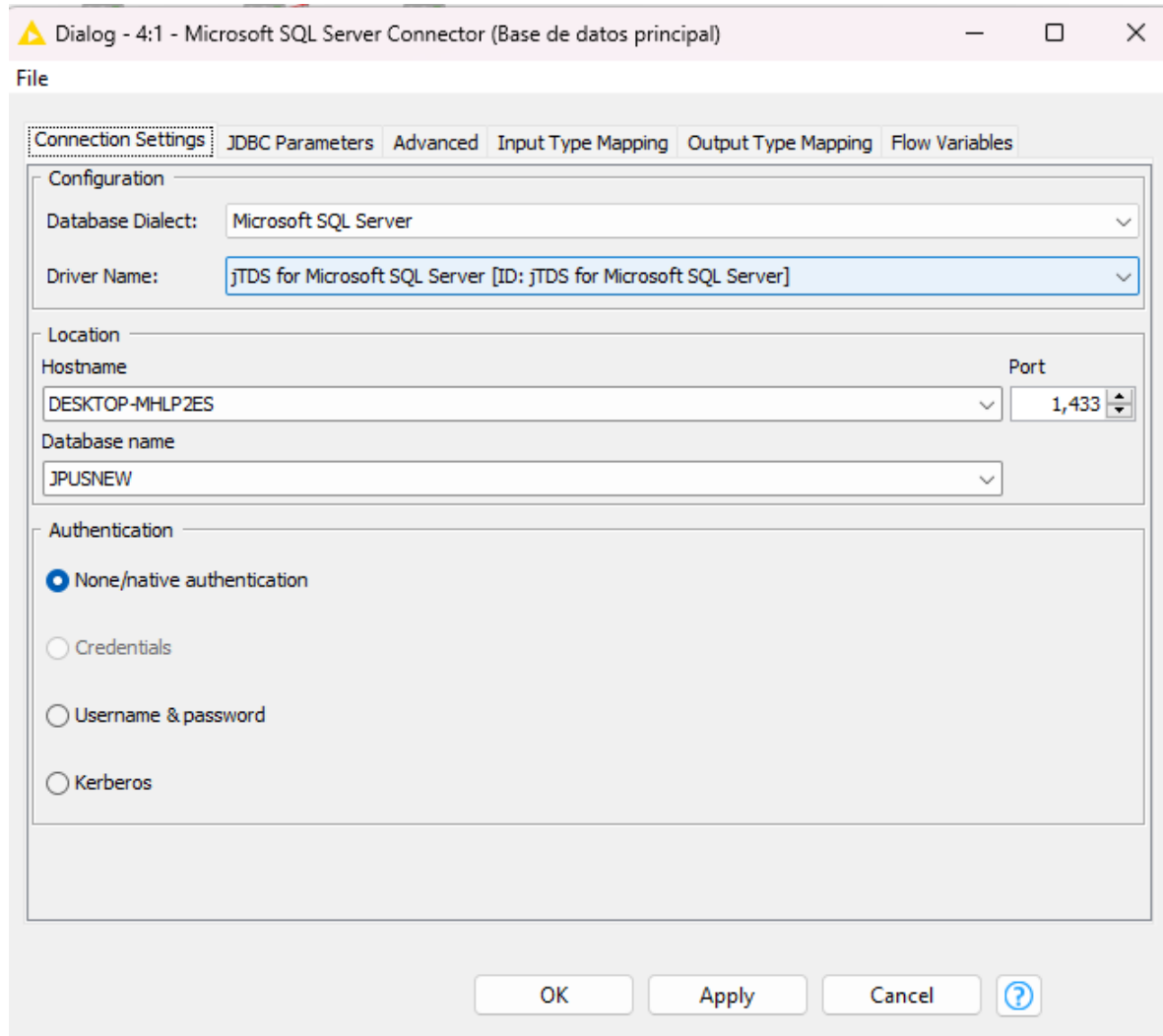
- [9] “Qué es una ETL y cómo funciona.” <https://geekytheory.com/que-es-una-etl-y-como-funciona/> (accessed Jul. 06, 2022).
- [10] C. Palma, W. Palma, and R. Pérez, *Dat a Mining El arte de Anticipar*, 1st ed., vol. 1. Santiago de Chile: Estadísticos e-Books & Papers, 2009.
- [11] “Qué es una base de datos | Oracle México.” <https://www.oracle.com/mx/database/what-is-database/> (accessed Jul. 05, 2022).
- [12] M. A. Romero and J. R. García, “COMPARISON OF OPTIONS FOR INTELLIGENCE BUSINESS IN MAJOR SYSTEMS MANAGERS MARKET DATABASES,” *REVISTA ECONOMÍA & ADMINISTRACIÓN*, vol. 6, no. COMPARACIÓN DE OPCIONES PARA INTELIGENCIA DE NEGOCIOS EN LOS PRINCIPALES SISTEMAS GESTORES DE BASES DE DATOS DEL MERCADO, pp. 5–20, 2016.
- [13] A. Leal, J. Pablo, and B. Luque, “El cuadro de mando integral,” *Fundacion para la motivacion de los recursos humanos*, 2005.
- [14] Altair, “El cuadro de mando integral,” *Economía 3 Colecciones*, p. 86, 2005, Accessed: Jul. 07, 2022. [Online]. Available: www.altair-consultores.com%5Cnaltair@altair-consultores.com
- [15] SelectHub, “Precios de Power BI, demostración, reseñas, características.” <https://www.selecthub.com/business-intelligence-tools/power-bi/> (accessed Jul. 07, 2022).
- [16] “Gartner Magic Quadrant & Critical Capabilities | Gartner.” <https://www.gartner.com/en/research/magic-quadrant> (accessed Mar. 07, 2023).
- [17] SelectHub, “Tableau Pricing, Demo, Reviews, Features.” <https://www.selecthub.com/business-intelligence-tools/tableau/> (accessed Jul. 07, 2022).
- [18] SelectHub, “Precios de Oracle Analytics Cloud, demostración, reseñas, características.” <https://www.selecthub.com/business-intelligence-tools/oracle-analytics-cloud/> (accessed Jul. 07, 2022).
- [19] Marín Rafael, “Los gestores de bases de datos (SGBD) más usados,” *Revista Digital INESEM*. 2019. Accessed: Jul. 07, 2022. [Online]. Available: <https://www.inesem.es/revistadigital/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>
- [20] Carreón Iglesias David Alberto, “Comparación de Gestores de Bases de Datos,” 2016.
- [21] Select hub, “Talend Pricing, Demo, Reviews, Features.” <https://www.selecthub.com/etl-tools/talend/> (accessed Jul. 11, 2022).
- [22] Select hub, “Qlik Data Integration Pricing, Demo, Reviews, Features.” <https://www.selecthub.com/etl-tools/qlik-data-integration/> (accessed Jul. 11, 2022).
- [23] Select Hub, “SQL Server Integration Services Pricing, Demo, Reviews, Features.” <https://www.selecthub.com/etl-tools/sql-server-integration-services/> (accessed Jul. 11, 2022).
- [24] G. E. Silva Peñafiel, V. M. Zapata Yáñez, K. P. Morales Guamán, and L. M. Toaquiza Padilla, “Análisis de metodologías para desarrollar Data Warehouse aplicado a la toma de decisiones,” *Ciencia Digital*, vol. 3, no. 3.4., pp. 397–418, Sep. 2019, doi: 10.33262/cienciadigital.v3i3.4..922.

- [25] “Inteligencia de Negocio: Ciclo de vida de Ralph Kimball.”
<http://luisleonin.blogspot.com/2014/02/ciclo-de-vida-de-ralph-kimball.html>
(accessed Jul. 17, 2022).
- [26] Espinosa Roberto, “Fases en la implantación de un sistema DW. Metodología para la construcción de un DW.” <https://churriwifi.wordpress.com/2009/12/05/5-fases-en-la-implantacion-de-un-sistema-dw-metodologia-para-la-construccion-de-un-dw/>
(accessed Jul. 17, 2022).
- [27] Microsoft, “Un recorrido por C#: descripción general | Documentos de Microsoft,” 2022. <https://docs.microsoft.com/en-us/dotnet/csharp/tour-of-csharp/> (accessed Aug. 14, 2022).

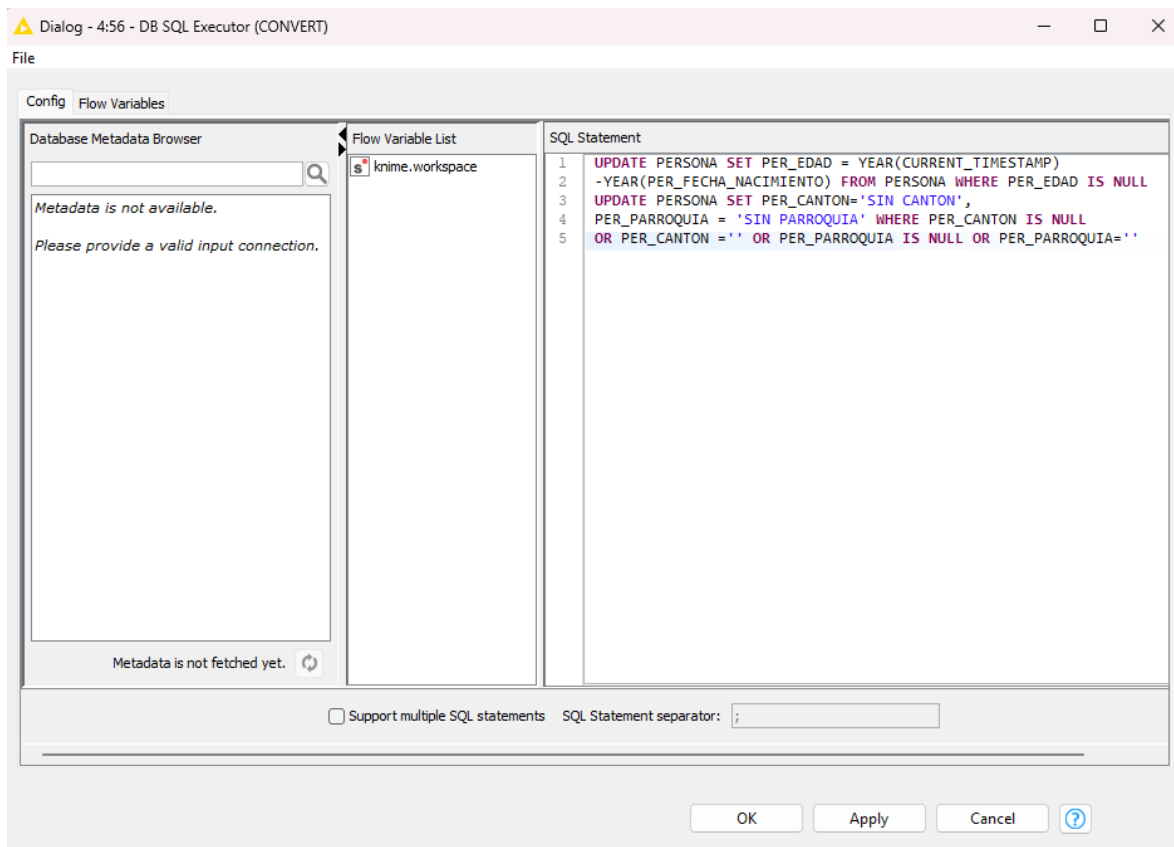
ANEXOS

4. ANEXO A NODOS DEL MODELO ETL EN KNIME

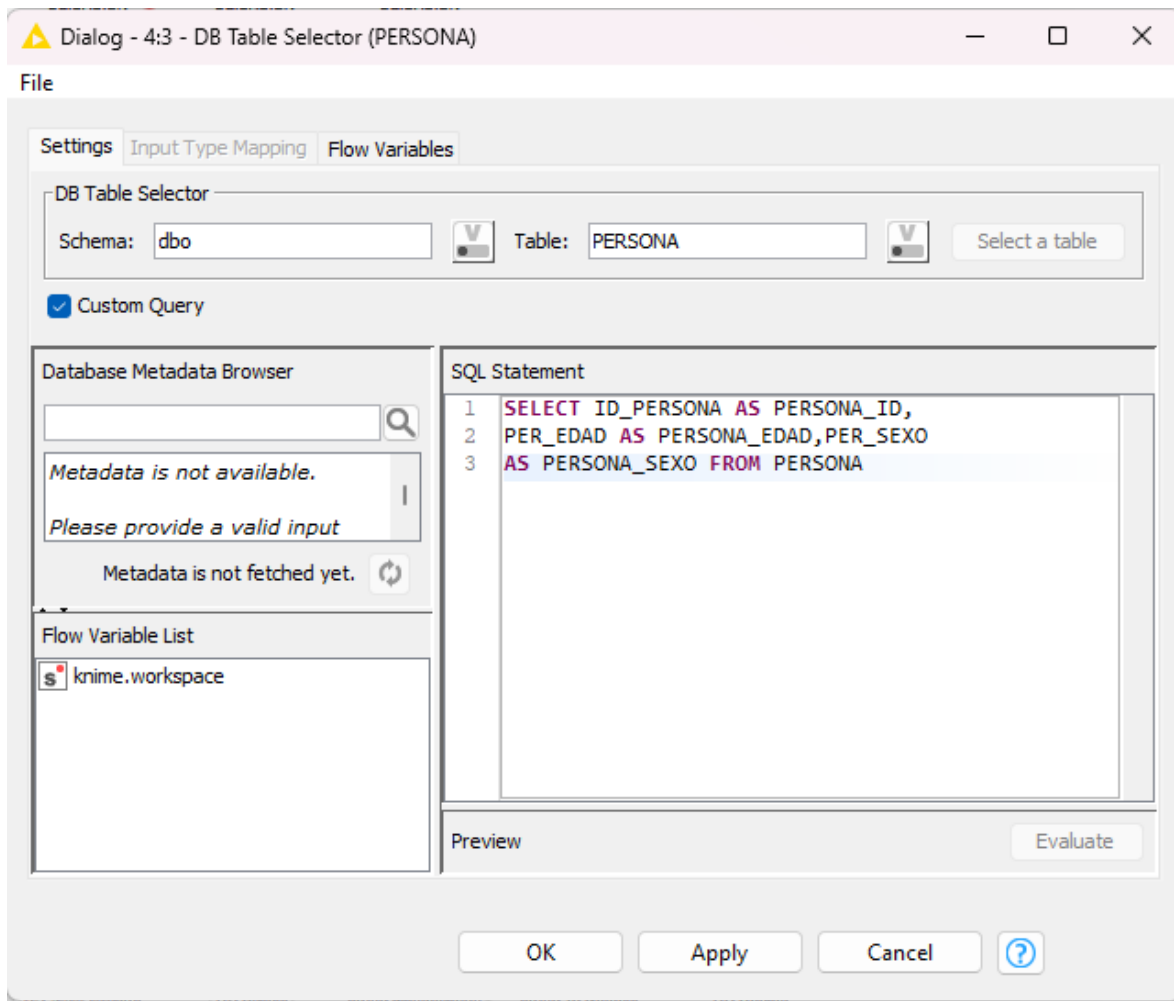
A continuación, se visualiza los nodos usados en el modelado del Data Warehouse



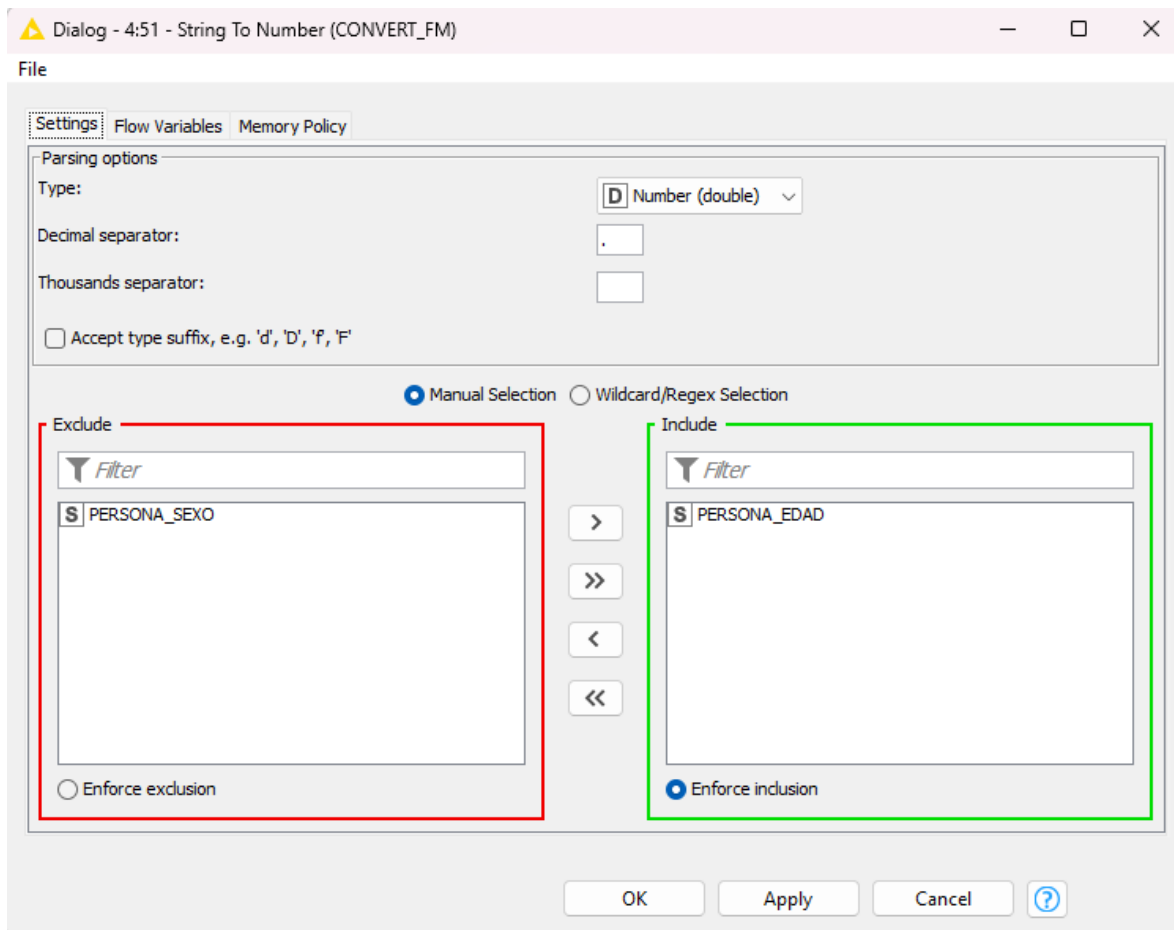
Microsoft SQL Server Connector, se usa para establecer la conexión con la base de datos operacional y la del almacén de datos



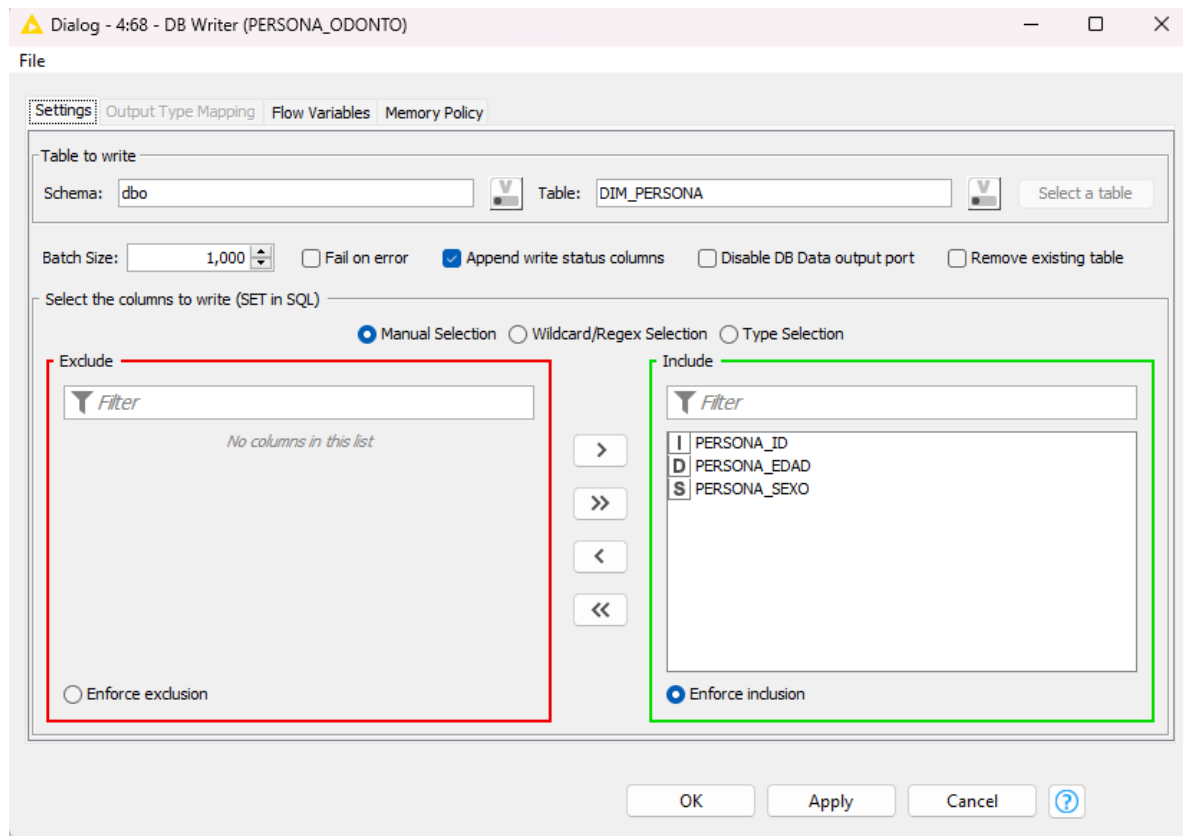
DB SQL Executor, se utiliza para ejecutar sentencias determinadas en lenguaje SQL para actualizar datos de las tablas



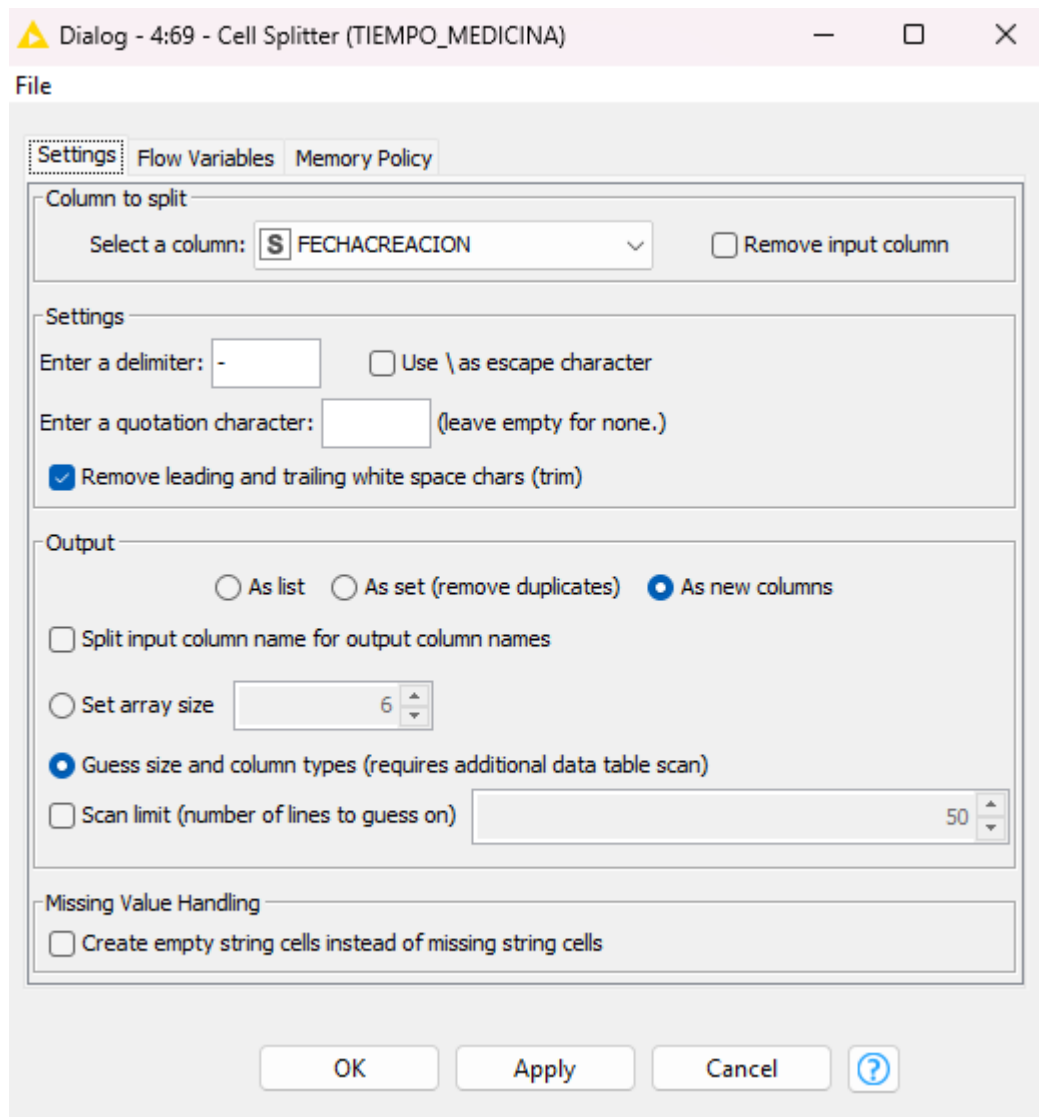
DB Table Selector, permite seleccionar una tabla que se usara para procesos posteriores, debe estar conectada primero a una base de datos para extraer información



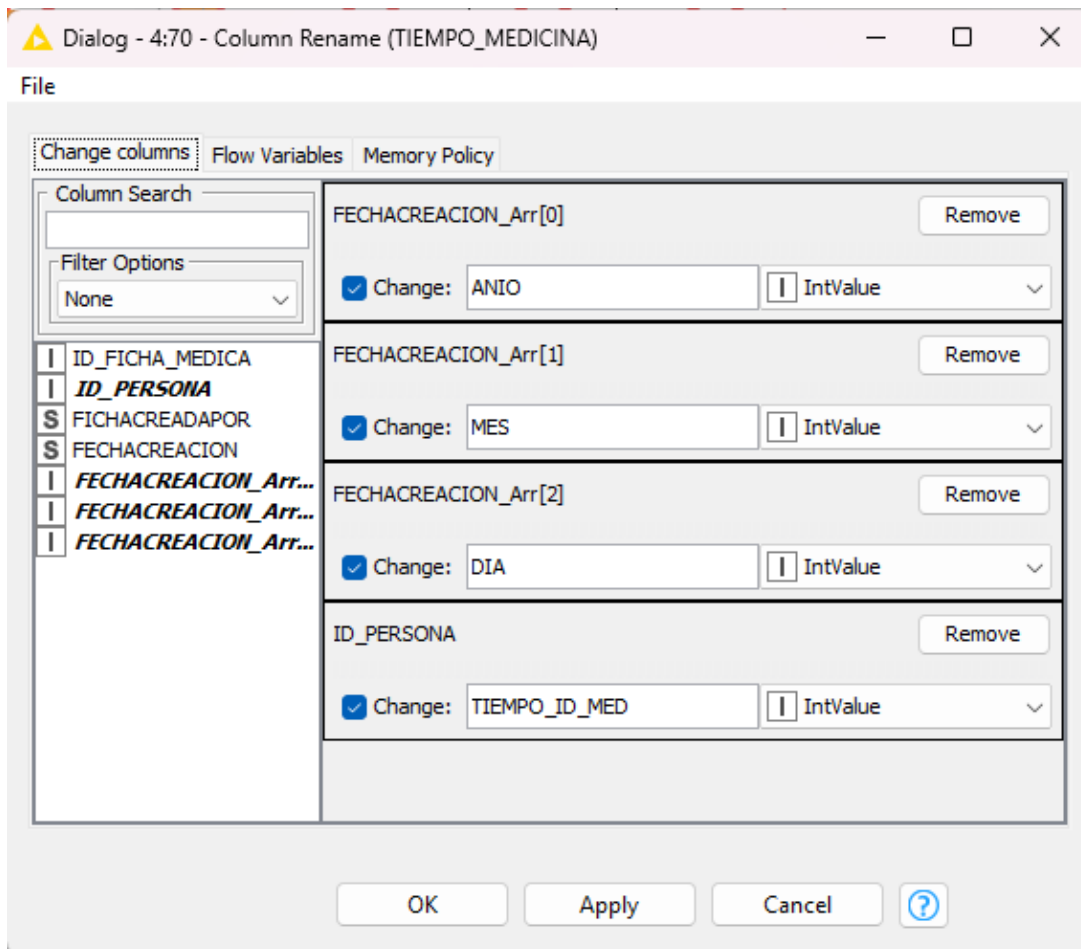
String to Number, convierte una cadena de texto en un valor numérico, permite seleccionar el punto o la coma como separador decimal.



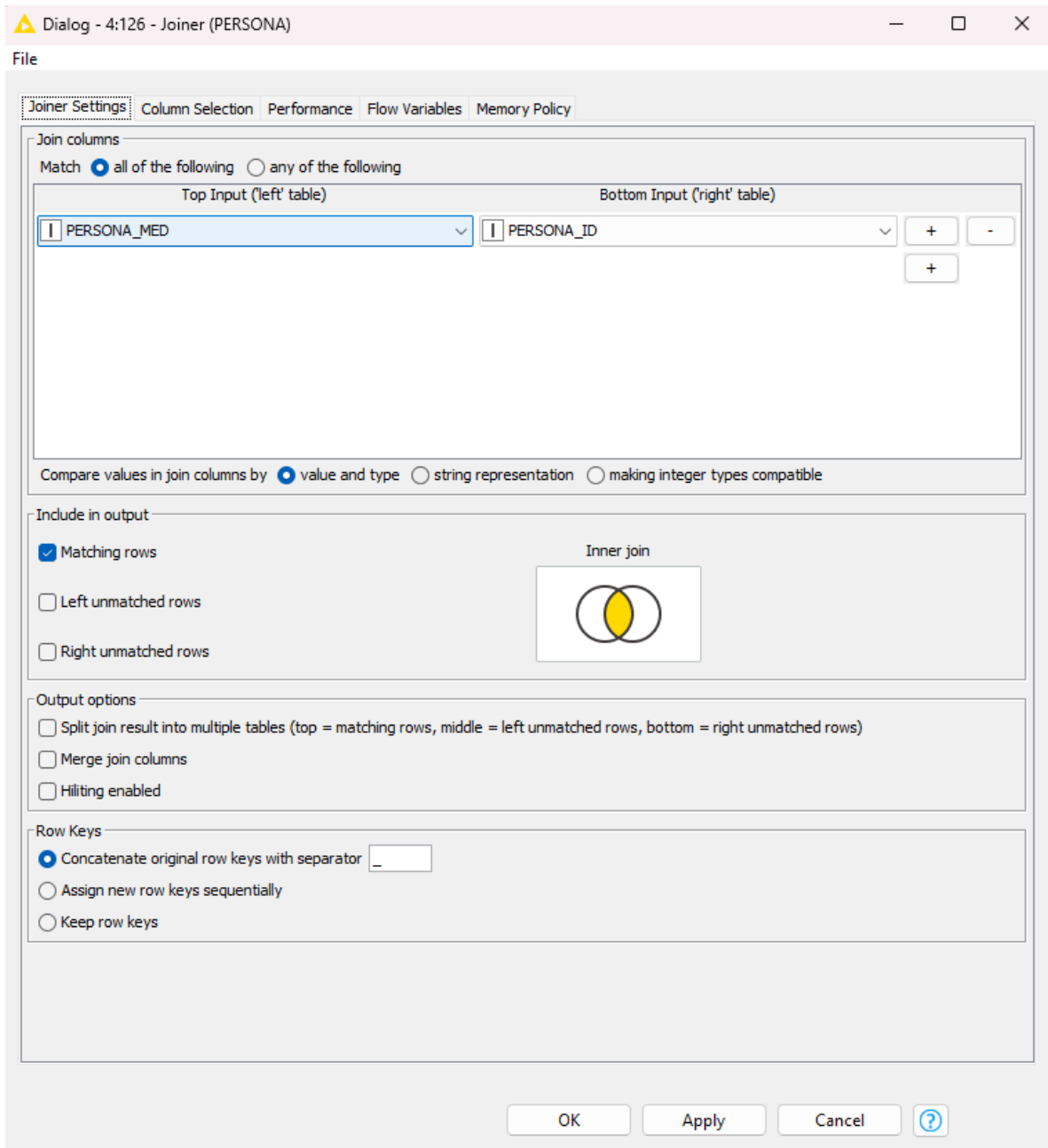
DB Writer cumple la función de insertar las filas seleccionadas a una tabla en una nueva base de datos.



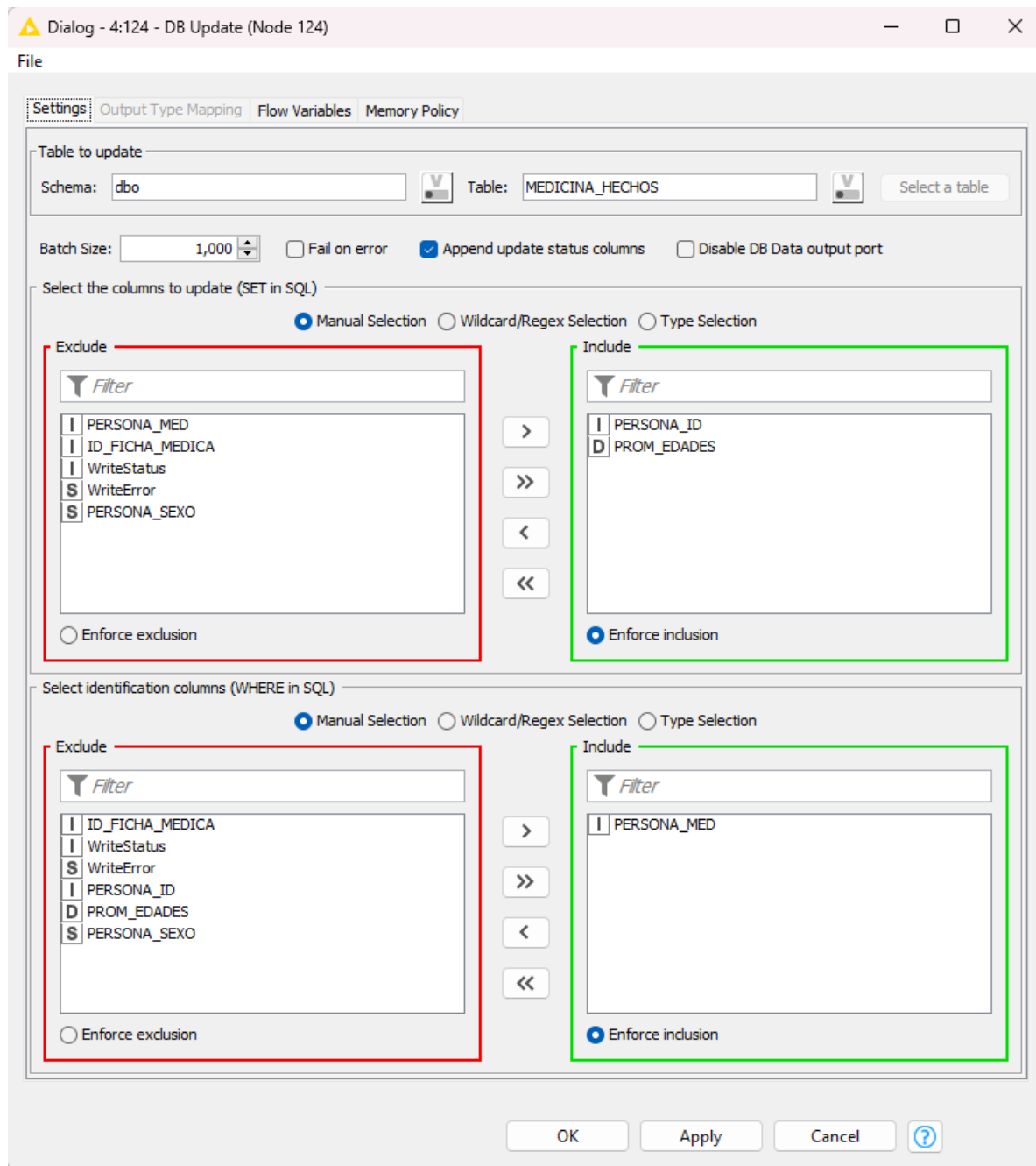
Cell Splitter a través de un delimitador establecido por el usuario permite separar el contenido de una columna y pasarlo a varias para tratarlos como nuevos datos.



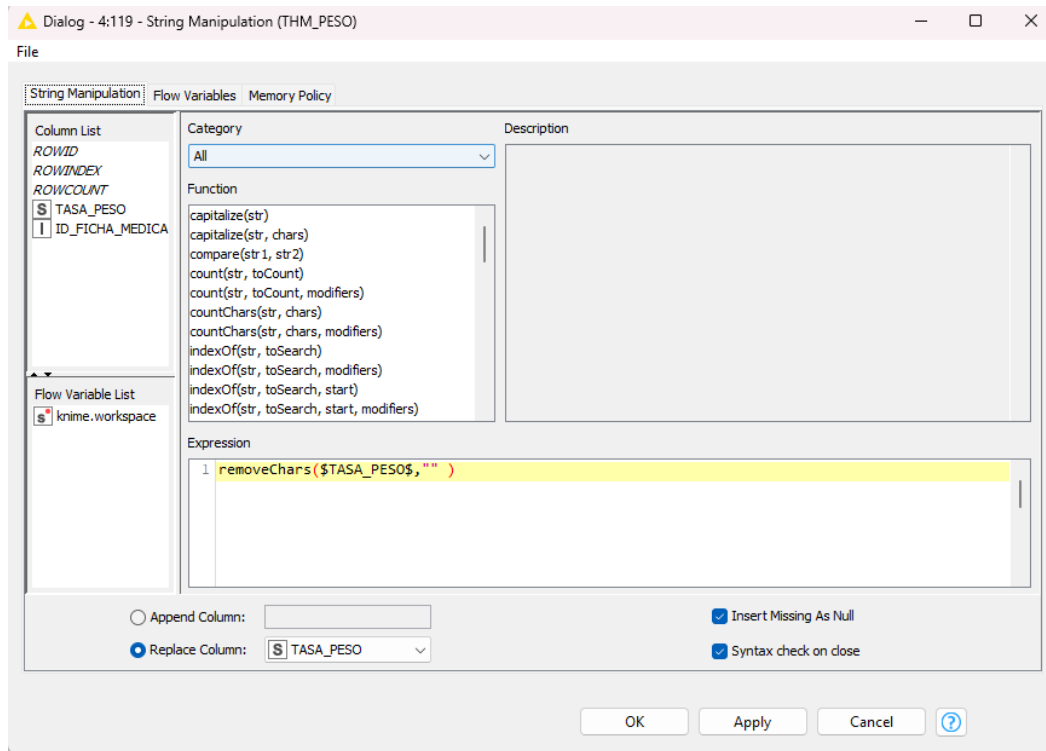
Column Rename renombra las columnas seleccionadas para evitar inconvenientes al momento de cargar los datos en la etapa final.



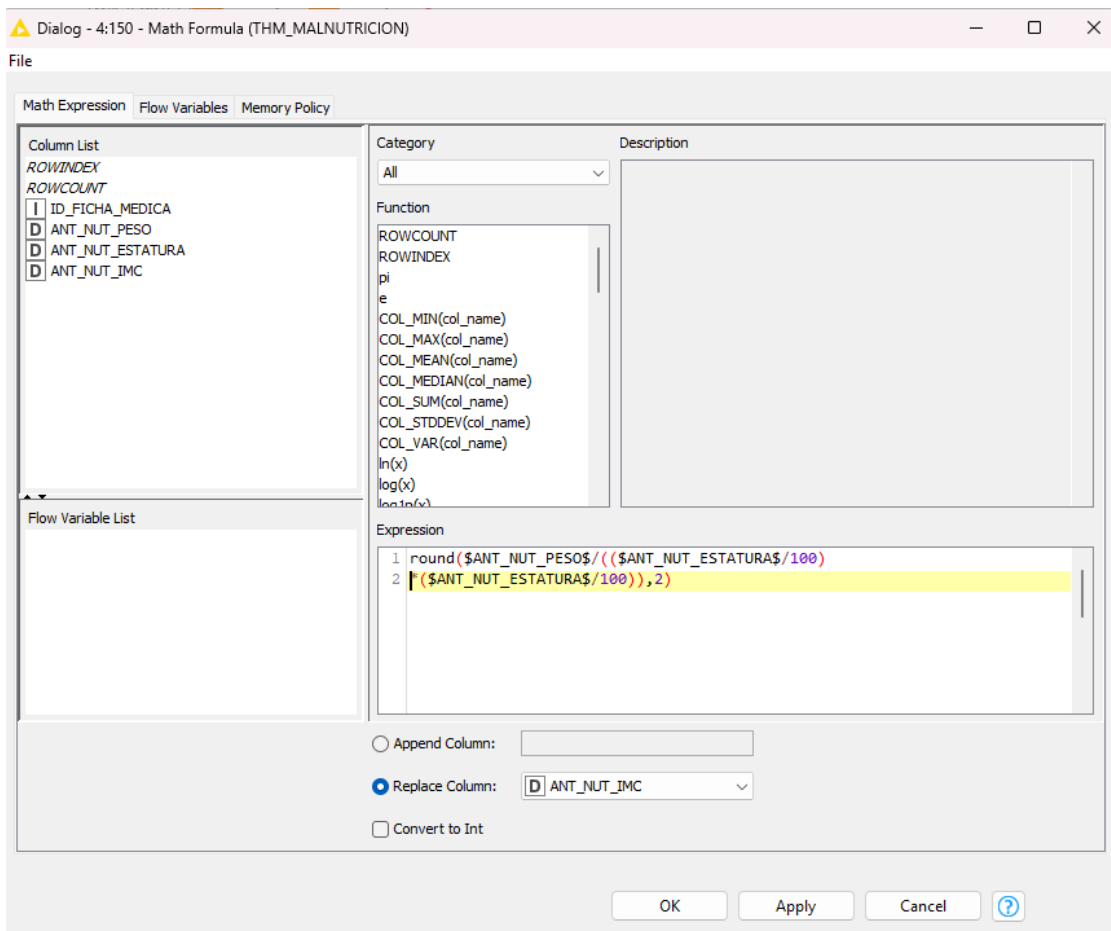
Joiner permite la unión de dos tablas relacionadas por una columna en común, con el resultado se facilita la depuración de información que se cargara



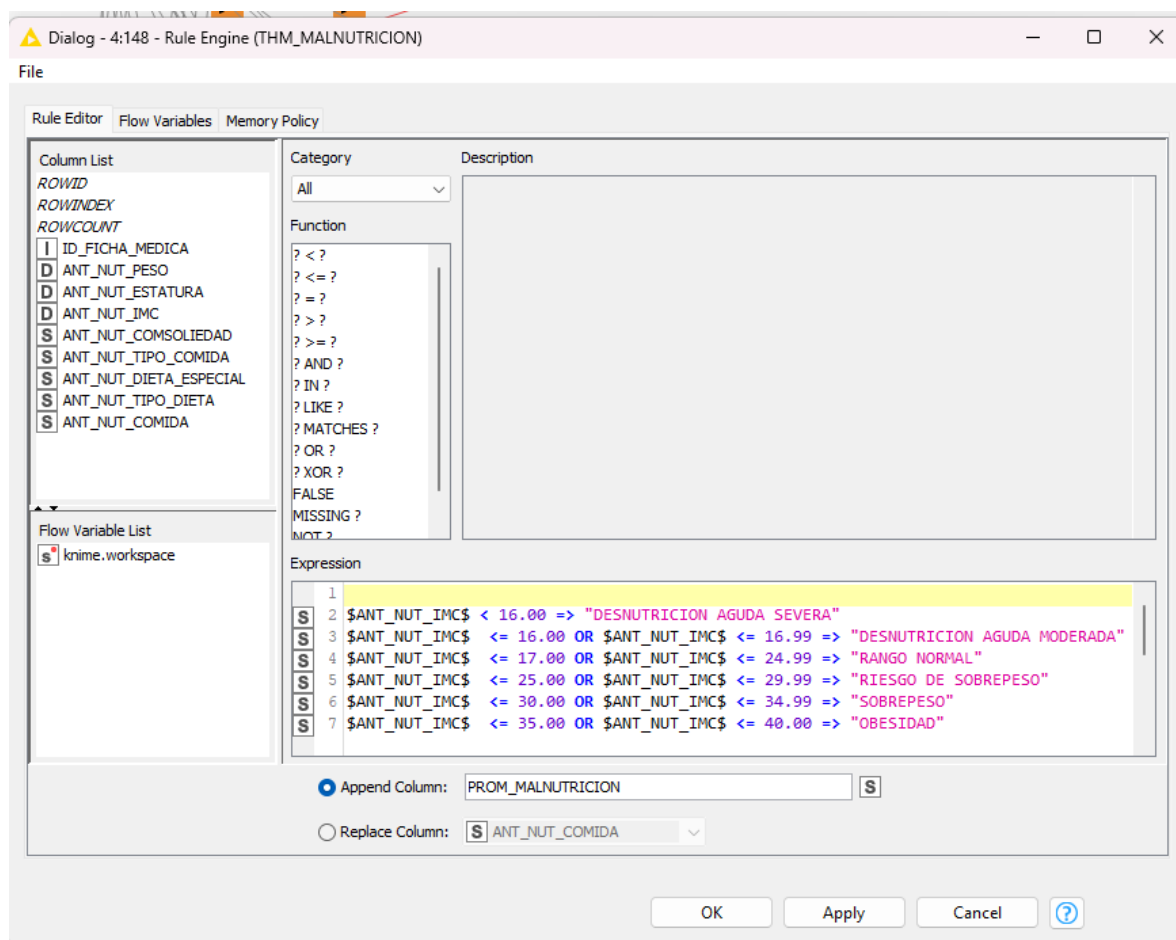
DB Update actualiza la información de una tabla, ejecutando una sentencia SQL usando una columna como Where y otra como SET que será donde se actualiza.



String Manipulation funciona como un editor de cadenas, permitiendo realizar distintas operaciones a una variable string, por ejemplo, eliminar espacios o insertar mas texto.

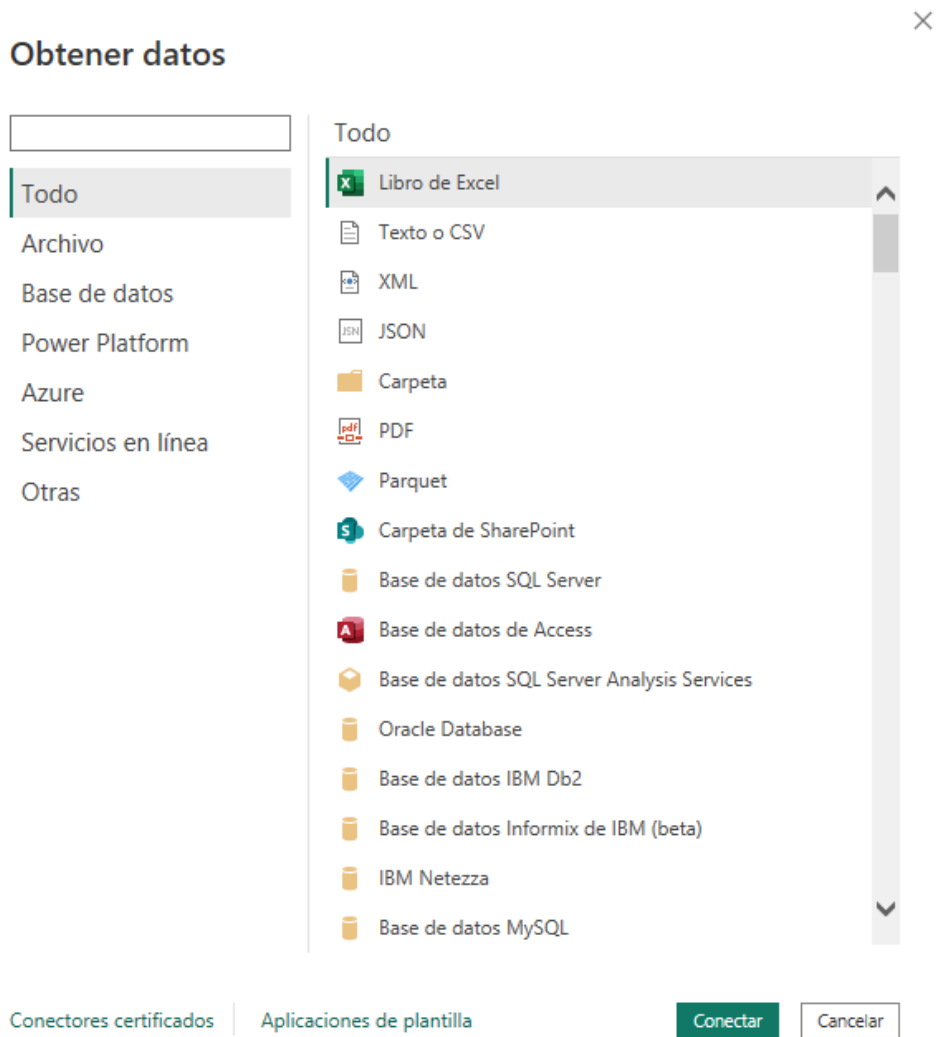


Math Formula realiza operaciones matemáticas que se guardan en una nueva columna o reemplazan los datos existentes.



Rule Engine evalúa que se cumplan ciertas condiciones establecidas para dar paso a nuevos datos, usa los operadores lógicos y algunos agregados por la herramienta.

5. ANEXO B DASHBOARDS DE MEDICINA Y ODONTOLOGIA EN POWER BI



La herramienta Power BI permite obtener datos desde diversas fuentes como se muestra en la figura

Base de datos SQL Server

Servidor ⓘ

Base de datos (opcional)

Modo Conectividad de datos ⓘ

Importar

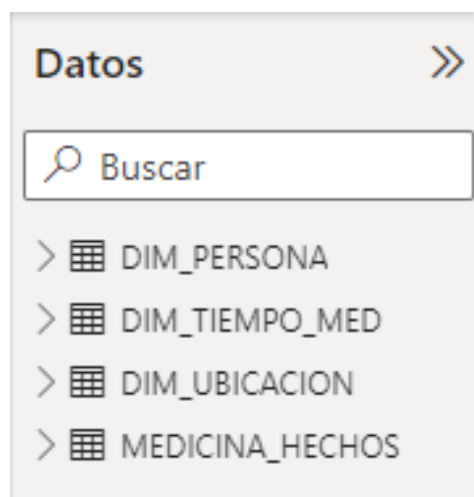
DirectQuery

▸ Opciones avanzadas

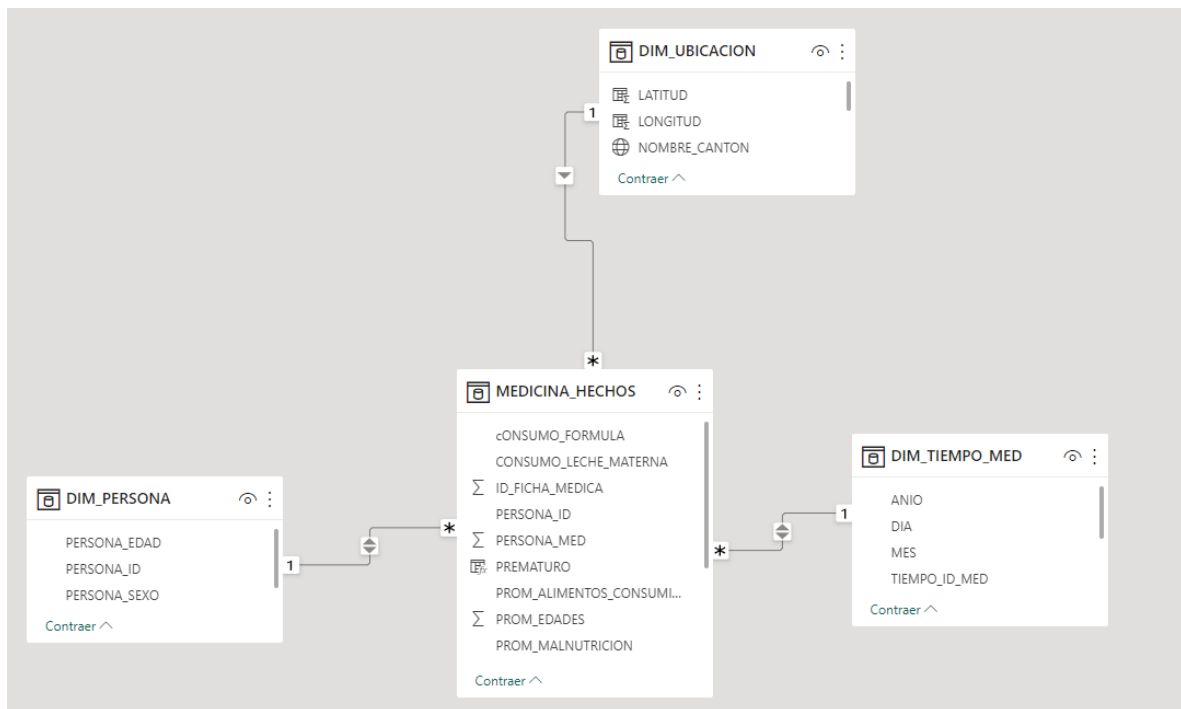
Aceptar

Cancelar

Conexión de base de datos SQL Server, requiere un servidor y una base de datos de manera opcional para extraer información.



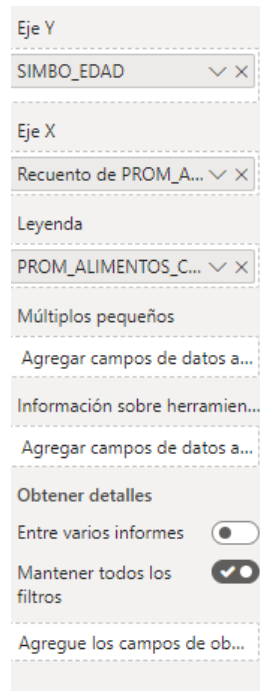
Una vez elegido el origen de datos se cargarán las tablas correspondientes, mismas que se usan para realizar medidas y gráficos.



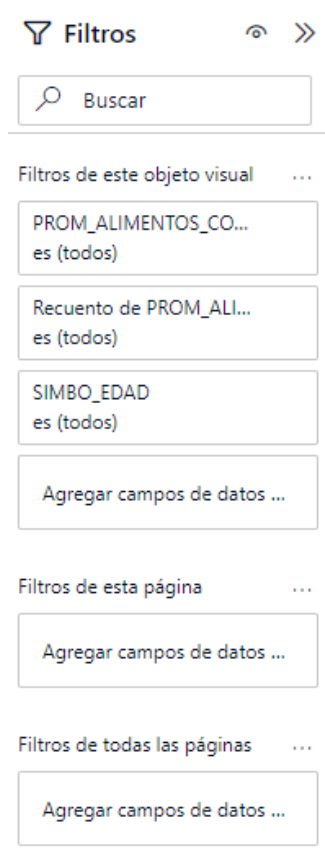
La herramienta también permite visualizar de mejor manera el modelado del origen de datos con sus columnas y relaciones.

PERSONA_ID	PERSONA_EDAD	PERSONA_SEXO	SIMBO_EDAD
2	10	M	ESCOLARES
14	10	M	ESCOLARES
18	10	F	ESCOLARES
23	10	M	ESCOLARES
27	10	F	ESCOLARES
31	10	F	ESCOLARES
59	10	F	ESCOLARES
60	10	F	ESCOLARES
74	10	F	ESCOLARES
80	10	F	ESCOLARES
85	10	F	ESCOLARES
88	10	F	ESCOLARES
97	10	F	ESCOLARES
98	10	M	ESCOLARES
104	10	F	ESCOLARES
114	10	M	ESCOLARES
116	10	M	ESCOLARES
128	10	F	ESCOLARES
129	10	M	ESCOLARES
166	10	M	ESCOLARES
171	10	F	ESCOLARES
176	10	M	ESCOLARES
181	10	F	ESCOLARES
182	10	F	ESCOLARES
193	10	F	ESCOLARES
200	10	F	ESCOLARES
206	10	M	ESCOLARES
210	10	M	ESCOLARES
236	10	M	ESCOLARES
251	10	M	ESCOLARES
253	10	F	ESCOLARES

Se tiene la vista de datos, que muestra de manera textual las columnas con su correspondiente información.



En este apartado se escogen las columnas que darán paso al gráfico seleccionado, donde se puede detallar los ejes, leyendas, categorías e información.



Los filtros de datos permiten evitar cierta información irrelevante en el dibujado de gráficos como lo es los datos en blanco o categorías que no son necesarias.

6. ANEXO C ACTUALIZACION DE ETL A TRAVES DE COMANDOS BASH

```
"C:\Program Files\KNIME\knime.exe" -nosave -consoleLog -noexit -nosplash -reset -application org.knime.product.KNIME_BATCH_APPLICATION -workflowDir="C:\Users\jose2\knime-workspace\ETL_FINAL"
```

El archivo bat abre el programa Knime, ejecuta el workflow específico donde se encuentra el ETL del data warehouse.

Nombre	Estado	Desencadenadores	Hora próxima ejecución	Hora última ejecución	Resultado de última ejecución	Autor
ETL	Listo	Se definieron varios desencadenadores	2/23/2023 7:59:39 PM	2/22/2023 7:59:40 PM	(0xC000013A)	DESKTOP-...
GoogleUpda...	Listo	Se definieron varios desencadenadores	2/22/2023 10:30:21 PM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	
GoogleUpda...	Listo	A las 10:30 PM todos los días - Tras desencadenarse, repetir cada 1 hora durante 1 día.	2/22/2023 9:30:21 PM	2/22/2023 8:53:45 PM	The operation completed successfully. (0x0)	
MicrosoftEd...	Listo	Se definieron varios desencadenadores	2/23/2023 7:00:00 PM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	
MicrosoftEd...	Listo	A las 6:30 PM todos los días - Tras desencadenarse, repetir cada 1 hora durante 1 día.	2/22/2023 9:30:00 PM	2/22/2023 8:53:45 PM	The operation completed successfully. (0x0)	
NvDriverUp...	Listo	A las 12:25 PM todos los días	2/23/2023 12:25:10 PM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	NVIDIA Co
NVIDIA GeF...	Listo	Al producirse un evento - Registro: Application, origen: NVIDIA GeForce Experience SelfUpdate Source, id. de evento: 0		11/30/1999 12:00:00 AM	The task has not yet run. (0x41303)	NVIDIA Co
NvNodeLau...	Listo	Al iniciar la sesión un usuario - Tras desencadenarse, repetir cada 1.00:00:00 indefinidamente.		2/22/2023 7:48:07 PM	The operation completed successfully. (0x0)	NVIDIA Co
NvProfileUp...	Listo	A las 12:25 PM todos los días	2/23/2023 12:25:07 PM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	NVIDIA Co
NvProfileUp...	Listo	Al iniciar la sesión un usuario		2/22/2023 7:50:07 PM	The operation completed successfully. (0x0)	NVIDIA Co
NvTmRep_C...	Listo	A las 12:25 PM todos los días	2/23/2023 12:25:10 PM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	NVIDIA Co
NvTmRep_C...	Listo	A las 6:25 PM todos los días	2/23/2023 6:25:10 PM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	NVIDIA Co
NvTmRep_C...	Listo	A las 12:25 AM todos los días	2/23/2023 12:25:10 AM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	NVIDIA Co
NvTmRep_C...	Listo	A las 6:25 AM todos los días	2/23/2023 6:25:10 AM	2/22/2023 7:53:44 PM	The operation completed successfully. (0x0)	NVIDIA Co
OneDrive Pe...	Listo	A las 7:00 PM el 5/1/1992 - Tras desencadenarse, repetir cada 1.00:00:00 indefinidamente.	2/23/2023 10:27:11 PM	11/30/1999 12:00:00 AM	The task has not yet run. (0x41303)	Microsoft
OneDrive Re...	Listo	A las 8:48 PM el 2/21/2023 - Tras desencadenarse, repetir cada 1.00:00:00 indefinidamente.	2/23/2023 8:48:37 PM	2/22/2023 8:48:38 PM	The operation completed successfully. (0x0)	Microsoft

En el programador de tareas se sube el archivo del tipo Bat, se elige una hora a ejecutar y el intervalo de tiempo en donde se hará.

7. ANEXO D ACTUALIZACION DE ORIGEN DE DATOS EN POWER BI

4 Conexión de puerta de enlace

Para usar una puerta de enlace de datos, asegúrese de que el equipo está en línea y de que el origen de datos se agrega en [Administrar conexiones y puertas de enlace](#). Si está usando una puerta de enlace de datos local (modo estándar), seleccione los orígenes de datos correspondientes y haga clic en Aplicar.

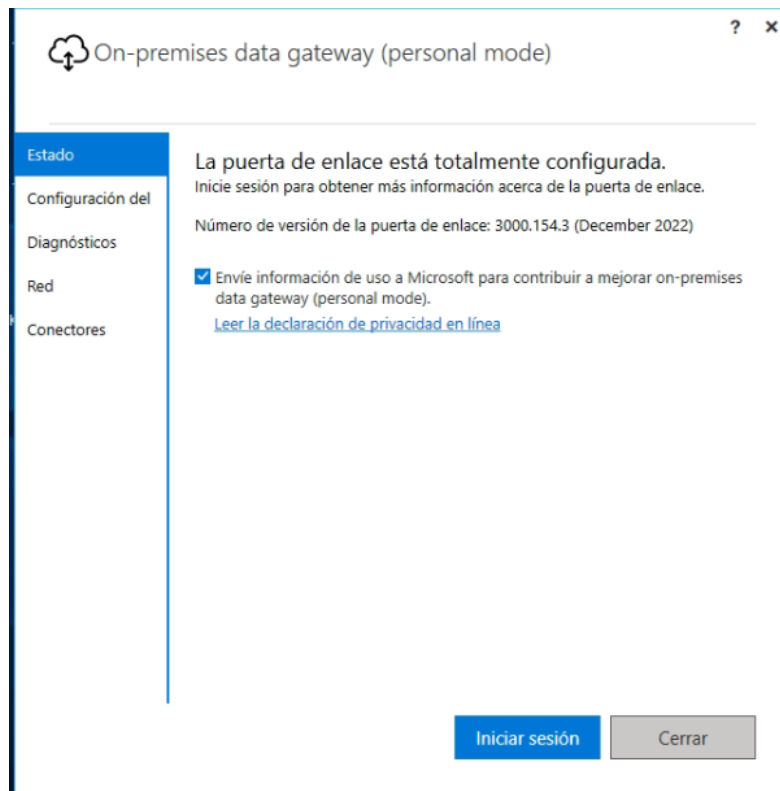
Use una puerta de enlace de datos local o VNet

Activar

Puerta de enlace	Departamento	Información de contacto	Estado	Acciones
Puerta de enlace personal			Ejecutando en SVRVINCURES	

Aplicar Descartar

Power BI app en la nube se conecta a través de una puerta de enlace con el servidor en donde está la data.



On-premises data Gateway establece la conexión entre el servidor y la nube de Microsoft power BI

Actualización programada

Mantener los datos actualizados

Configure una programación de actualización de datos para importar datos del origen de datos al conjunto de datos. [Más información](#)

Activar

Frecuencia de actualización

Diaria

Zona horaria

(UTC-05:00) Bogotá, Lima, Quito

Hora

1 00 a. m.

[Agregar otra hora](#)

Destinatario del envío de notificaciones de los errores de actualización

Propietario del conjunto de datos

Estos contactos:

Escriba las direcciones de correo electrónico

Aplicar

Descartar

La actualización programada establece una hora en la que el conjunto de datos se actualizara en caso de que existan nuevos registros.

8. ANEXO E PUBLICACION DE DASHBOARDS EN LA PAGINA DE JPUS

Código para insertar

Vínculo que puede enviar por correo electrónico

<https://app.powerbi.com/view?r=eyJrIjojOGQyNjBk> Copia

Código HTML que puede pegar en un sitio web

`<iframe title="D_MEDICINA FINAL" width="600" h` Copia

Tamaño

Imagen de marcador de posición

Cargar Eliminar

Página predeterminada

Power BI permite crear etiquetas HTML para insertar en cualquier página de manera pública y que los dashboards sean visualizados sin necesidad de descargar ningún complemento.

```

1 <asp:Page Title="" Language="C#" MasterPageFile="~/Principal.Master" AutoEventWireup="true" Co
2 <asp:Content ID="Content1" ContentPlaceHolderID="head" runat="server">
3 </asp:Content>
4 <asp:Content ID="Content2" ContentPlaceHolderID="PHContenido" runat="server">
5 <div class="row-cols-lg-1" style="background-color: rgba(255, 0, 0, 0.5)">
6 <iframe title="Report Section" width="1024" height="1060"
7 src="https://app.powerbi.com/view?r=eyJrIjojNzZlMGUyOjUwMTZlNiNSO0NTc5LTlhMDk0kTN2I5
8 frameborder="0" allowFullScreen="true"></iframe>
9 </div>
10 </asp:Content>
11

```

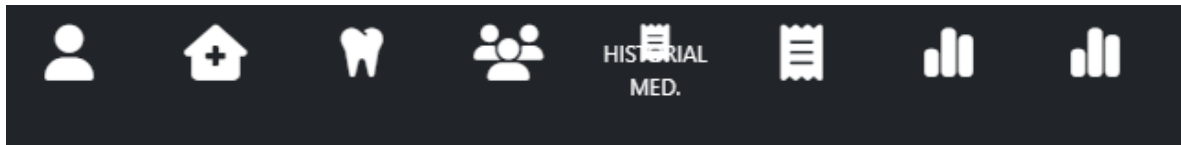
Para hacer uso de la etiqueta donde se encuentra el dashboard, se crea una nueva página de acceso único para administrativos y docentes del proyecto donde se incluye la información mencionada

```

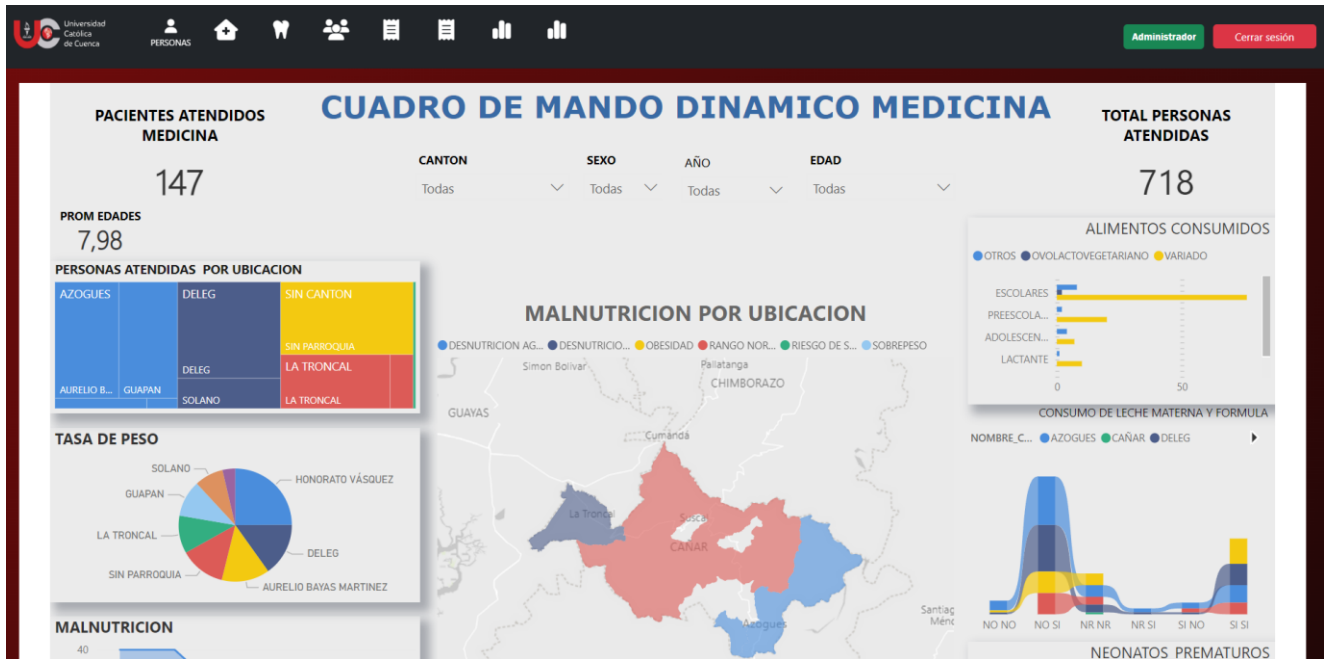
</li>
<li class="lista nav-item">
<a runat="server" href="~/Sistema/Dashboards/Dashboard_Med.aspx">
<span class="icono"><i class="fa-solid fa-chart-simple"></i></span>
<span class="titulo">DASHBOARD MEDICINA</span>
</a>
</li>
<li class="lista nav-item">
<a runat="server" href="~/Sistema/Dashboards/Dashboard_Odo.aspx">
<span class="icono"><i class="fa-solid fa-chart-simple"></i></span>
<span class="titulo">DASHBOARD ODONTOLOGIA</span>
</a>
</li>

```

En el código del menú principal se agrega las páginas correspondientes a los dashboards de medicina y odontología.



El usuario final visualiza el menú con nuevas opciones que dirigen al cuadro de mando integral



Dentro de la página, los gráficos se visualizan sin necesidad de redirigir a webs externas, ni hacer uso de complementos para acceder al sitio web se puede acceder mediante el siguiente link:

<http://45.182.119.100:50378/Default.aspx>

Herman Leonel Barrazueta Calle portador de la cédula de ciudadanía N° **0350012191**. En calidad de autor y titular de los derechos patrimoniales del trabajo de titulación **“CONSTRUCCION DE UNA SOLUCION BUSINESS INTELLIGENCE EN LOS MODULOS DE HISTORIAS CLINICAS Y ODONTOLOGICAS DEL PROYECTO “JUNTOS POR UNA SONRISA””** de conformidad a lo establecido en el artículo 114 Código Orgánico de la Economía Social de los Conocimientos, Creatividad e Innovación, reconozco a favor de la Universidad Católica de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos y no comerciales. Autorizo además a la Universidad Católica de Cuenca, para que realice la publicación de éste trabajo de titulación en el Repositorio Institucional de conformidad a lo dispuesto en el artículo 144 de la Ley Orgánica de Educación Superior.

Azogues, 20 de marzo de 2023

F: 

Herman Leonel Barrazueta Calle

C.I. 0350012191

José Luis Lema Espinoza portador de la cédula de ciudadanía N° **0302731120**. En calidad de autor y titular de los derechos patrimoniales del trabajo de titulación **“CONSTRUCCION DE UNA SOLUCION BUSINESS INTELLIGENCE EN LOS MODULOS DE HISTORIAS CLINICAS Y ODONTOLOGICAS DEL PROYECTO “JUNTOS POR UNA SONRISA””** de conformidad a lo establecido en el artículo 114 Código Orgánico de la Economía Social de los Conocimientos, Creatividad e Innovación, reconozco a favor de la Universidad Católica de Cuenca una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos y no comerciales. Autorizo además a la Universidad Católica de Cuenca, para que realice la publicación de éste trabajo de titulación en el Repositorio Institucional de conformidad a lo dispuesto en el artículo 144 de la Ley Orgánica de Educación Superior.

Azogues, 20 de marzo de 2023



F:

José Luis Lema Espinoza

C.I. 0302731120