

UNIVERSIDAD NACIONAL SAN CRISTÓBAL DE HUAMANGA

FACULTAD DE INGENIERÍA DE MINAS, GEOLOGÍA Y CIVIL

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



**“DATA MART PARA EL ESTADO DE CONTROL DE CRECIMIENTO Y
DESARROLLO PARA EL ESTADO NUTRICIONAL DE NIÑOS Y NIÑAS DE LA
RED DE SALUD HUAMANGA, AYACUCHO 2019”**

Tesis presentada por : Bach. Adan Ccoyllo Aguilar

Para optar el título profesional de : Ingeniero de sistemas

Tipo de Investigación : Observacional, Retrospectiva y transversal

Área de Investigación : Inteligencia de Negocios

Asesor : Mag. Ing. Hubner Janampa Patilla

Ayacucho - Perú

2020

DEDICATORIA

A Dios, por darme vida, salud y la oportunidad de cumplir mis objetivos y permitido vivir en esta hermosa ciudad lleno de alegría y emociones en las cuales aprendí el arte del estudio.

A mis padres, por su apoyo fundamental a cada momento, por su paciencia, por los valores que me inculcaron para que pudiera ser una persona de bien, por sus buenos consejos, y sobre todo por ese amor que nunca me faltó al lado de ellos.

AGRADECIMIENTO

A Dios.

A mis dos hermanos menores, por tomarme como ejemplo y seguir el camino del estudio.

A la Universidad Nacional de San Cristóbal de Huamanga por medio de ello a la Escuela Profesional de Ingeniería de Sistemas y a sus docentes quienes con sus conocimientos impartidos hicieron que aprendiera y me formara en el mundo del conocimiento.

A sus trabajadores de la Red de Salud de Huamanga, por haberme colaborado en mi formación profesional y sobre todo por tenerme paciencia.

CONTENIDO

DEDICATORIA.....	i
AGRADECIMIENTO.....	ii
CONTENIDO.....	iii
RESUMEN.....	vi
INTRODUCCIÓN.....	vii

CAPITULO I

PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1. DIAGNOSTICO Y ENUNCIADO DEL PROBLEMA	1
1.2. DEFINICIÓN DEL PROBLEMA DE INVESTIGACIÓN	2
1.3. OBJETIVOS DE LA INVESTIGACIÓN	2
1.4. JUSTIFICACIÓN Y DELIMITACIÓN DE LA INVESTIGACIÓN.....	2
1.4.1. IMPORTANCIA DEL TEMA	2
1.4.2. JUSTIFICACIÓN.....	3
1.4.3. DELIMITACIÓN.....	3

CAPITULO II

REVISIÓN DE LA LITERATURA

2.1. ANTECEDENTES DE LA INVESTIGACIÓN	4
2.2. MARCO TEÓRICO	5
2.2.2. ESTADO NUTRICIONAL.....	7
2.2.3. RED DE SALUD HUAMANGA.....	9
2.2.4. INTELIGENCIA DE NEGOCIOS	9
2.2.5. DATA WEREHOUSE	12
2.2.6. DATA MART	13
2.2.7. VENTAJAS Y DESVENTAJAS	15

2.2.8.	DIFERENCIA ENTRE DATA MART Y DATA WEREHOUSE.....	15
2.2.9.	METODOLOGÍA RALPH KIMBALL	17
2.2.10.	PROCESAMIENTO ANALÍTICO EN LÍNEA OLAP (ON LINE ANALYTICAL PROCESSING)	28
2.2.11.	SISTEMA GESTOR DE BASES DE DATOS (SGBD).....	32
2.2.12.	POBLACIÓN	35
2.2.13.	CENSO.....	35

CAPITULO III

METODOLOGÍA DE LA INVESTIGACIÓN

3.1.	TIPO Y NIVEL DE LA INVESTIGACIÓN	36
3.2.	DISEÑO DE LA INVESTIGACIÓN.....	37
3.3.	POBLACIÓN Y MUESTRA	37
3.4.	VARIABLES E INDICADORES	37
3.5.	TÉCNICAS E INSTRUMENTOS PARA RECOLECTAR INFORMACIÓN	39
3.6.	HERRAMIENTAS PARA EL TRATAMIENTO DE DATOS	39
3.7.	TÉCNICAS PARA APLICAR LA METODOLOGÍA KIMBALL.....	40

CAPITULO IV

RESULTADO DE LA INVESTIGACIÓN

4.1.	RESULTADOS APLICANDO LA METODOLOGÍA RALPH KIMBALL.....	45
4.1.1.	PLANIFICACIÓN DEL PROYECTO	45
4.1.2.	DEFINICIÓN DE LOS REQUERIMIENTOS	45
4.1.3.	DISEÑO DE LA ARQUITECTURA TÉCNICA	47
4.1.4.	SELECCIÓN DE PRODUCTO	48
4.1.5.	MODELO DIMENSIONAL	49
4.1.6.	DISEÑO DEL MODELO FÍSICO.....	67
4.1.7.	DISEÑO E IMPLEMENTACIÓN DEL SUB SISTEMA DE ETL.....	76
4.1.8.	IMPLEMENTACIÓN	102

4.1.9.	ESPECIFICACIÓN DE APLICACIONES PARA USUARIOS FINALES	108
4.1.10.	DESARROLLO DE APLICACIONES PARA USUARIOS FINALES	109
4.1.11.	MANTENIMIENTO Y CRECIMIENTO.....	122

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

5.1.	CONCLUSIONES.....	123
5.2.	RECOMENDACIONES	124
	BIBLIOGRAFÍA	125

ANEXOS

RESUMEN

La unidad de Estadística, informática y Telecomunicaciones de la Red de Salud Huamanga cuenta con un Software llamado HisMinsa, mediante el cual se hace el llenado de atenciones que se realiza en los diferentes puestos de salud, dicha información es utilizada por la unidad para el análisis y procesamiento de data, para las altas gerencias de la Red, que estas a su vez brinda la información a la dirección regional de salud de Ayacucho (DIRESA), para las diferentes tomas de decisiones.

El estudio se realizó con la base de datos de la Red de Salud de Huamanga, el tipo de investigación es observacional, retrospectivo y transversal con el nivel de investigación descriptiva.

Se implementará un Data Mart que tiene como objetivo principal construir el Data Mart para el estado de control de crecimiento y desarrollo que ayude al monitoreo del estado nutricional de los niños y niñas de la Red de Salud Huamanga, Ayacucho. Usando la metodología Ralph Kimball, que incluye una serie de tareas como primera tarea es la planificación del proyecto, seguido por la definición de los requerimientos del negocio, y centrarnos en la tarea de diseño e implantación del subsistema ETL, en la cual se podrá realizar la extracción, transformación y carga. Se utilizará SQL Server, y se podrá visualizar los resultados en el software de Power BI, para las tomas de decisiones.

Como una de las principales conclusiones se tiene que de acuerdo al marco teórico del capítulo II, sección 2.2.9, se construye el Data Mart de acuerdo a la metodología de Ralph Kimball definida sus técnicas en el capítulo III, sección 3.7, la cual nos permitió lograr los objetivos trazados, la cual se puede mostrar en el capítulo IV, gráficos 4.1. al 4.3, que brinda toda la información del Control de crecimiento y desarrollo

Palabra clave: Control de Crecimiento y Desarrollo, Estado Nutricional, Data Mart, Metodología Ralph Kimball, Proceso ETL.

INTRODUCCIÓN

La Red de Salud de Huamanga es una institución prestadora de servicios de salud en las diferentes localidades de la región que están dentro de su jurisdicción, para ello cuenta con las Micro redes y Establecimientos de salud, y se genera una gran cantidad de data con las diferentes atenciones que realiza a los pacientes de todas las edades y genero a lo largo de todo año en curso, a lo que principalmente pone énfasis es a los niños y niñas con sus diferentes atenciones que se realiza durante toda la etapa de desarrollo de vida del niño, entre uno de las atenciones se encuentra el Control de Crecimiento y Desarrollo, siendo este uno de los principales atenciones que se realiza al niño y niña de la provincia de huamanga, por ello se desea realizar un Data Mart del control de crecimiento y desarrollo para así saber cuántos atenciones se ha realizado a cada niño o niña, utilizando la metodología de Ralph Kimball.

Mi motivación para la implementación del Data Mart, se debe a que la Red de Salud Huamanga cuenta con una gran cantidad de data, que le llega de los diferentes lugares, que aún no están procesadas, a la cual brindarle una herramienta para procesar la data y tener una información rápida, para las diferentes tomas de decisiones, para ello se plantea el problema principal ¿ Como el Data Mart para el estado de control de crecimiento y desarrollo ayuda al monitoreo del estado nutricional de los niños y niñas de la Red de Salud Huamanga.

Los objetivos específicos son: a) Diseñar el Data Mart para el Estado de Control de Crecimiento que ayude al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho 2019. B) Diseñar el Data Mart para el Estado de Control de Desarrollo que ayude al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho 2019.

CAPITULO I

PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1. DIAGNOSTICO Y ENUNCIADO DEL PROBLEMA

La etapa infantil es considerada la más importante, para el desarrollo de toda la etapa de vida del ser humano, por lo tanto se debe realizar un seguimiento de manera adecuada y oportuna la evolución y desarrollo de los niños y niñas, siendo uno de los principales deberes del estado de velar por su desarrollo y preocuparse por su bienestar a través de instituciones, entidades protectoras de estrategias sanitarias y los programas de salud que son implementados para su atención, y una de las instituciones es la Red de Salud Huamanga, ellas mediante los centros de Salud velan a los niños y niñas de la región de Ayacucho. Las atenciones que se realiza al niño y niña, cuentan con uno de las principales atenciones que es el Control de Crecimiento y Desarrollo (CRED) cuya finalidad de la atención es contribuir en el crecimiento, desarrollo y estado nutricional, así detectar oportunamente las enfermedades como alteraciones o trastornos que pudieran acoger al niño o niña.

En la actualidad no se cuenta con datos exactos y visibles sobre cada atención que debe llevar cada niño y niña, o simplemente no se sabe exactamente cuantas atenciones tiene y tubo según su edad correspondiente, esto poniendo la deserción al control CRED por parte de sus padres del niño y convirtiéndolo en un problema para el personal de salud, debido a que no se cuenta con la información necesaria, esto originando mayor número de atenciones o menor número de atenciones según lo que le corresponde al niño y niña.

La unidad encargada de procesar la información de las diferentes atenciones es la unidad de Estadística, Informática y Telecomunicaciones, cuya recepción de datos lo realiza mediante el sistema web HisMinsa, cuyas atenciones que fueron realizadas por el personal de salud de manera diaria de los diferentes centros de salud, convirtiéndolo en una gran cantidad de datos para procesar, dentro del proceso de datos que realiza la unidad se ha encontrado ambigüedad en el procesamiento de datos, duplicidad de información y entregas tardías, esto convirtiéndolo en problema a la hora de tomar decisiones.

1.2. DEFINICIÓN DEL PROBLEMA DE INVESTIGACIÓN

PROBLEMA PRINCIPAL

¿Cómo el Data Mart para el Estado de Control de Crecimiento y Desarrollo ayuda al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho 2019?

PROBLEMA ESPECÍFICO

- a) ¿Cómo el Estado de Control de Crecimiento ayuda al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho, 2019?
- b) ¿Cómo el Estado de Control de Desarrollo ayuda al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho, 2019?

1.3. OBJETIVOS DE LA INVESTIGACIÓN

OBJETIVO PRINCIPAL

Construir el Data Mart para el Estado de Control de Crecimiento y Desarrollo que ayude al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho, 2019.

OBJETIVOS ESPECÍFICOS

- a) Diseñar el Data Mart para el Estado de Control de Crecimiento que ayude al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho, 2019.
- b) Diseñar el Data Mart para el Estado de Control de Desarrollo que ayude al monitoreo del Estado Nutricional de los Niños y Niñas de la Red de Salud Huamanga, Ayacucho, 2019.

1.4. JUSTIFICACIÓN Y DELIMITACIÓN DE LA INVESTIGACIÓN

1.4.1. IMPORTANCIA DEL TEMA

Como importancia técnica se busca contribuir a las instituciones a adoptar el desarrollo de inteligencia de negocios como los data werehouse o la Data Mart adecuando a sus necesidades para tener una información clara y precisa a la mano para tomar las mejores

decisiones, para su desarrollo se busca mostrar el desarrollo, análisis y diseño de un Data Mart con la metodología de Kimball.

1.4.2. JUSTIFICACIÓN

En la actualidad la cantidad de datos fue creciendo y estos a la vez convirtiéndose en el principal activo de cualquier organización, y la Red de Salud de Huamanga siendo una organización prestadora de servicio de salud genera gran cantidad de información mediante las atenciones que realiza en sus centros de salud, para la toma de decisiones eficientes se basa en una gestión adecuada del conocimiento y uno de los pilares de esta son los almacenes de datos, pero la Red de Salud Huamanga no cuenta con ello, es por ello que se plantea desarrollar un Data Mart para poder tener datos integrados consistentes, confiables y así poder poner en disposición de usuarios autorizados de forma oportuna para la toma de decisiones.

El presente trabajo de investigación pretende brindar información confiable y oportuna a la Red de Salud Huamanga, para poder optar estrategias, acuerdos, seguimiento y cumplimiento sobre las atenciones de las etapas de desarrollo de vida del niño y niña, sobre todo el control CRED, la cual es el centro de la investigación, esto permitiendo detectar las diferentes enfermedades de manera temprana y tomar las mejores decisiones.

1.4.3. DELIMITACIÓN

La investigación se realizará en la ciudad de Ayacucho, abarcando las atenciones que se realizan a los niños y niñas de 0 a 11 años de edad, en los diferentes establecimientos de salud, que son pertenecientes a la Red de Salud Huamanga,

CAPITULO II

REVISIÓN DE LA LITERATURA

2.1. ANTECEDENTES DE LA INVESTIGACIÓN

Chávez (2017) en su tesis sobre Asistir a las consultas de Crecimiento y Desarrollo (CRED) plantea que es importante para un niño permitiendo así identificar de manera temprana si existe algún riesgo que pondría en peligro el desarrollo normal de la condición física y emocional del infante.

Campos (2018) en su tesis sobre asistir a los controles de CRED, nos dice que son muy importantes para el niño(a), ya que permite identificar oportunamente cualquier riesgo, alteraciones que pongan en peligro el crecimiento y desarrollo normal de él.

Huamán y Huingo (2016) en su tesis sobre el control de crecimiento y desarrollo del menor de un año, teniendo como objetivo principal evaluar el cumplimiento de la norma técnica de la atención integral del niño, así concluyendo que el mayor tiempo se los destina a los registros de datos.

Rojas (2018) en su tesis de investigación habla como desde la implementación de un Datamart se logró obtener información táctica y oportuna sobre las prestaciones de servicios de salud en adolescentes, como prestación preventiva, prestación recuperativa y prestación de rehabilitación. Además, la facilidad de realizar reportes personalizados orientados al análisis de datos y la toma de decisiones.

Duran (2017) en su tesis de investigación concluye que se logró hacer un seguimiento académico a los alumnos de ingeniería de sistemas de la universidad nacional de Cajamarca, mediante un Datamart bajo la metodología de Ralph Kimball, esto apoyando mejor a las tomas de decisiones.

2.2. MARCO TEÓRICO

2.2.1. CONTROL DE CRECIMIENTO Y DESARROLLO (CRED)

Según el Ministerio de Salud (MINSA, 2005), define como conjunto de actividades periódicas y sistemáticas desarrolladas por el profesional ya sea enfermera o médico, con el objetivo de vigilar de manera adecuada y oportuna el crecimiento y desarrollo de la niña y el niño; detectar de manera precoz y oportuna riesgos, alteraciones o trastornos, así como la presencia de enfermedades, facilitando su diagnóstico e intervención oportuna disminuyendo deficiencias y discapacidades.

Para el Ministerio de Salud (MINSA), “El control de Crecimiento y desarrollo (CRED) ofrece la oportunidad de brindar la consejería como un espacio de comunicación entre el personal de salud y la madre o cuidador para apoyar prácticas saludables” (2011, p. 52)

CRECIMIENTO

Según MINSA (2011), define como el proceso de incremento de la masa corporal de un ser vivo, que se produce por el aumento en el número de células (hiperplasia) o de su tamaño (hipertrofia). Es un proceso que está regulado por factores nutricionales, socioeconómicos, culturales, emocionales, genéticos y neuroendocrinos. Se mide por medio de las variables antropométricas: peso, talla, perímetro cefálico.

Para Torres, “Movimiento de la materia viva que se desplaza en el tiempo y en el espacio. El crecimiento es sólo la manifestación de la capacidad de síntesis de un organismo y de cada una de sus células” (2002, p. 92).

Según Ferri, El crecimiento indica en cantidad “Es el resultado de la división celular y de la síntesis de proteínas. Se refleja por el aumento del tamaño y peso, del conjunto o de cualquiera de sus partes” (1996, p. 32)

DESARROLLO

Según MINSA (2011), define como proceso dinámico por el cual los seres vivos logran mayor capacidad funcional de sus sistemas a través de fenómenos de maduración, diferenciación e integración de sus funciones, en aspectos como el biológico, psicológico, cognoscitivo, nutricional, sexual, ecológico, cultural, ético y social. Se encuentra influenciado por factores genéticos, culturales y ambientales.

Para Ferri. “El desarrollo es definido como un aumento progresivo de habilidades y capacidad para funcionar, indica un cambio cualitativo en el funcionamiento del niño y niña” (1996, p. 2)

“El desarrollo implica la diferenciación y madurez de las células y se refiere a la adquisición de destrezas y habilidades en varias etapas de la vida” (Torres, 2002, p. 54).

A. PERIODICIDAD DEL CONTROL DE CRECIMIENTO Y DESARROLLO DE LA NIÑA Y NIÑO MENOR DE 5 AÑOS.

Tabla 1

Periodo de controles del niño y niña menores de 5 años

EDAD	CONCENTRACIÓN	PERIODICIDAD
Recién nacido	2	7 y días de vida
De 01 a 11 meses	11	1m, 2m, 3m, 4m, 5m, 6m, 7m, 8m, y 9m, 10m, 11m
De 12 a 23 meses	6	12m,14m, 16m, 18m, 20m, 22 meses
De 24 a 59 mes	12 (4 por año)	24m, 27m, 30m, 33m, 36m, 42m, 45m, 48m, 51m, 54m y 57mes

Fuente: MINSA, 2011

Tabla 2

Control de crecimiento y desarrollo (Z001)

GRUPO DE EDAD	NÚMERO DE CONTROLES											Total
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	
De 1d a 7d												
De 8d a 14d												
De 15d a 21d												
De 22d a +d												
Total Recién Nacido												
29 días - 11 meses												
01 año												
02 años												
03 años												
04 años												
05 a 11 años												

Fuente: MINSA, 2011

2.2.2. ESTADO NUTRICIONAL

Según Minsa (2011), el estado nutricional se evalúa y monitorea el crecimiento utilizando las medidas antropométricas (peso, longitud, talla y perímetro cefálico) las mismas que son tomadas en cada contacto de la niña o niño con los servicios de salud y son comparadas con los patrones de referencia vigentes, determinando la tendencia del crecimiento.

Para la Organización Mundial de la Salud (OMS). “Es la condición del organismo que resulta de la relación entre las necesidades nutritivas individuales y la ingestión, absorción y utilización de los nutrientes contenidos en los alimentos” (s, f, p. 129)

Tabla 3

Codificación del estado nutricional.

PUNTO DE CORTE	PESO POR EDAD	CIE10	PESO POR TALLA	CIE 10	TALLA PARA EDAD	CIE10
Desviación Estándar	Clasificación		Clasificación		Clasificación	
>+3			Obesidad	E669		
>+2	Sobrepeso	E660	Sobrepeso	E660	Alto	E344
+2 a -2	Normal	Z006	Normal	Z006	Normal	Z006
<-2 a -3	Desnutrición	E440	Desnutrición Aguda	E440	Talla Baja	E45X
< -3			Desnutrición Severa	E43X		

Fuente: MINSA, 2012

Tabla 4

Estado nutricional (Z006)

GRUPO DE EDAD	NÚMERO DE EVALUACIONES ANTROPOMÉTRICAS											
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	TOTAL
De 1d a 7d												
De 8d a 14d												
De 15d a 21d												
De 22d a +d												
Total, Recién Nacido												
29 días - 11 meses												
01 año												
02 años												
03 años												
04 años												
05 a 11 años												

Fuente: MINSA, 2012

2.2.3. RED DE SALUD HUAMANGA

Para la Dirección Regional de Salud de Ayacucho (DIRESA, 2012), define como el conjunto de establecimientos y servicios de salud, de diferentes niveles de complejidad y capacidad de resolución, interrelacionados por una red vial y corredores sociales, articulados funcional y administrativamente, cuya combinación de recursos y complementariedad de servicios asegura la provisión de un conjunto de atenciones prioritarias de salud, en función de las necesidades de la población.

A. MICRO RED

La Micro Red es el conjunto de establecimientos de salud del primer nivel de atención cuya articulación funcional, según criterios de accesibilidad y ámbito geográfico, facilita la organización de la prestación de servicios de salud. La Micro Red es la superficie territorial dentro de la cual se puede dar con facilidad las relaciones entre la población y el establecimiento de salud o entre establecimientos, debido a que presenta facilidad para la comunicación y el transporte dentro de su territorio. (Para la Red de Salud Huamanga, 2012, p. 32)

B. CENTRO DE SALUD

Para la Red de Salud Huamanga “Los centros de salud son lugares que brindan cuidados para la salud. Que son acreditados por la autoridad competente y registrados en la Superintendencia Nacional de Aseguramiento en Salud, autorizados para brindar los servicios de salud correspondientes a su nivel de atención” (2012, p. 45)

C. UNIDAD DE ESTADÍSTICA, INFORMÁTICA Y TELECOMUNICACIONES

Red de Salud Huamanga, define como una unidad de apoyo a la gerencia. “Encargada de lograr que la Red de Salud Huamanga provea la información estadística de salud y el soporte informático, mecanización e integración de los sistemas de información requeridos para los procesos organizacionales” (2012, p. 118)

2.2.4. INTELIGENCIA DE NEGOCIOS

Según Ahumada, Perusquia (2014) a partir de la gestión del conocimiento, surge el concepto de inteligencia de negocios (Business Intelligence, inteligencia empresarial o

inteligencia de negocios); se llama así al conjunto de estrategias, acciones y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa.

Para Oracle, La inteligencia de negocios debe ser parte de la estrategia empresarial, ya que le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y aumenta la capacidad de tomar buenas decisiones para así obtener mejores resultados. (2016, p. 2)

Según Rotaeché, Business Intelligence es el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa. (2007, p. 52)

Según Cano (2007). Business Intelligence es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un data warehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones.

Según Cutro (s.f.), Business Intelligence es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios.

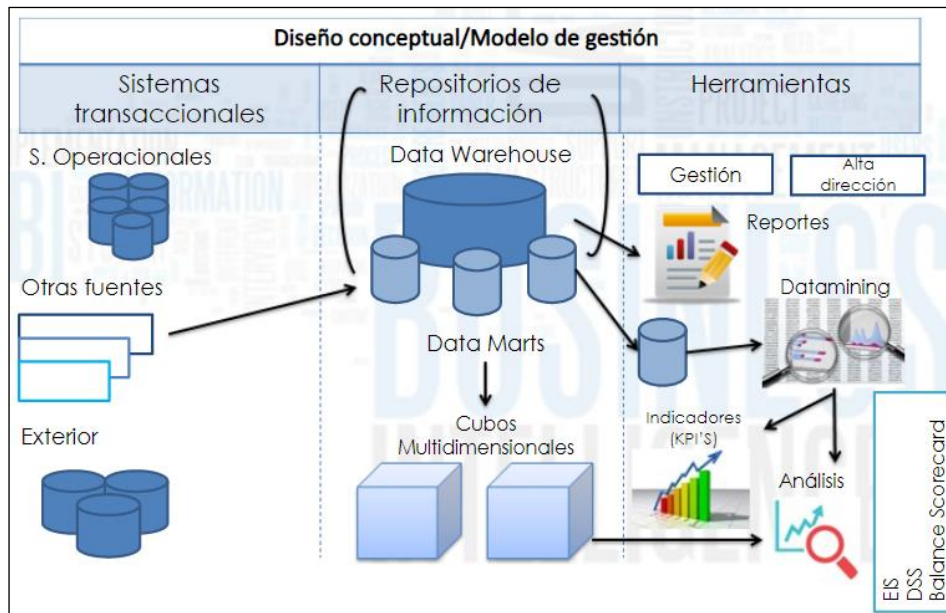


Figura 2.1. Arquitectura de una solución de Business Intelligence (Coronel 2015)

Acciones, lograr conclusiones más precisas y llegar a decisiones efectivas y oportunas. Por lo tanto, podemos concluir que la principal ventaja derivada de la adopción de un sistema de inteligencia empresarial se encuentra en la mayor efectividad del proceso de toma de decisiones. (Wiley, 2009)

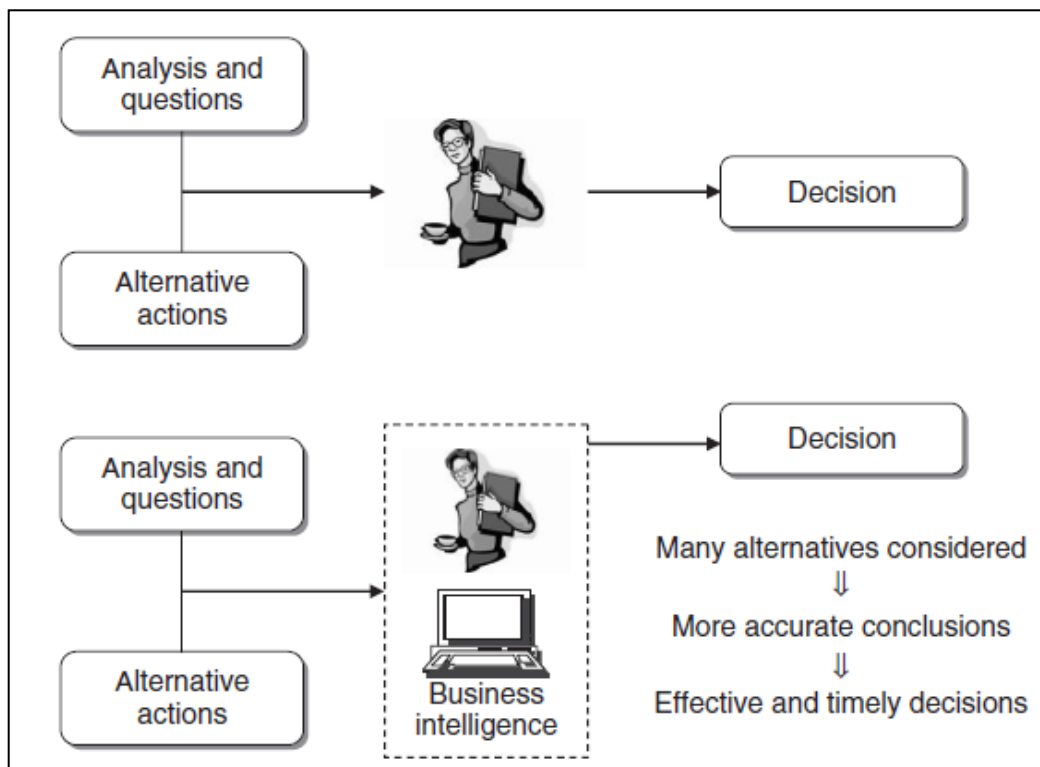


Figura 2.2: Beneficios de un sistema de inteligencia de negocios (Wiley, 2009)

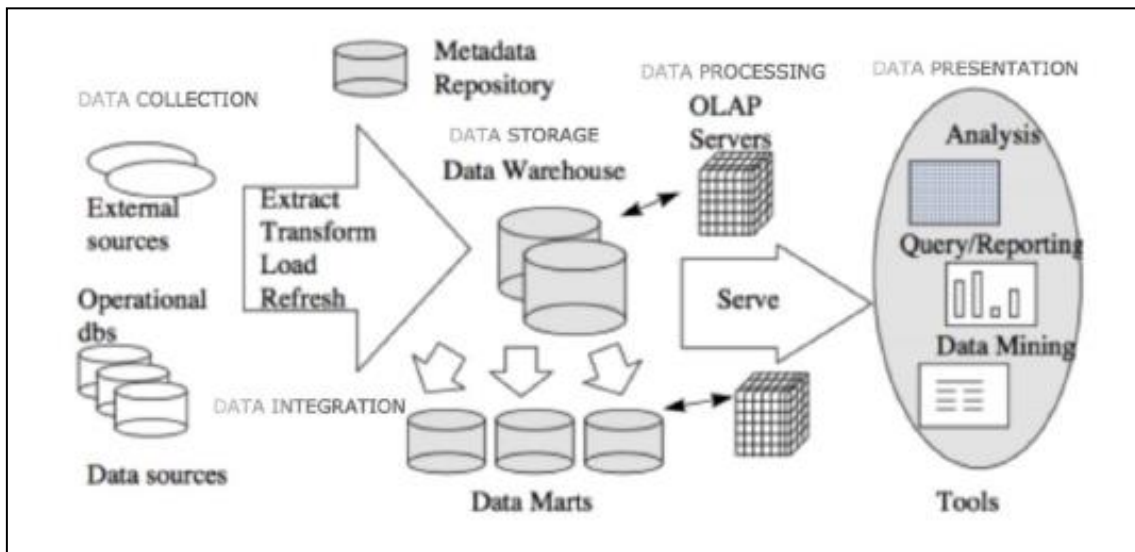


Figura 2.3: Proceso de inteligencia de negocios (Wang, J. y Chen, 2005)

2.2.5. DATA WEREHOUSE

Según Conesa y Curto (2011), una data warehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización, independientemente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios, con las propiedades siguientes: estable, coherente, fiable y con información histórica.

Para Cano. “Un data warehouse es una colección de información creada para soportar las aplicaciones de toma de decisiones” (2007, p. 24).

Según Lopez (2002), el data warehouse es una colección de datos orientados al tema, integrados, no volátiles e historiadados, organizados para el apoyo de un proceso de ayuda a la decisión.

Según Kimball y Caserte, Un Data warehouse es una copia de los datos transaccionales específicamente estructurada para la consulta y el análisis en el procesamiento de la información. (2004, p. 63)

Según Inmon (2007), afirma que el Data warehouse es un modelado y análisis de datos para los tomadores de decisiones, que son útiles en el apoyo y soporte, el cual garantiza la

confiabilidad y provee operaciones lógicas estructurales de los datos procesados a nivel organizacional.

“Conjunto de tecnologías de soporte a la toma de decisión, cuyo objeto es que quien trabaja con los conocimientos (ejecutivo, director, analista) pueda tomar decisiones de manera más rápida y eficaz” (Chaudhuri y Dayal, 1997, p. 518).

2.2.6. DATA MART

Para Conesa y Curto (2011), un Data Mart es un subconjunto de datos del data warehouse cuyo objetivo es responder a una necesidad, con una población de usuaria específica. Los datos están estructurados en modelo estrella o copo de nieve. Un Data Mart puede ser independiente o dependiente de un data warehouse.

Según Tana (2014), son almacenes de datos más pequeños que una data werehouse que por lo general son utilizados por las pequeñas y medianas empresas a diferencia de los cubos de información que nos ayudan a visualizar la información clasificada.

Un Data mart es como un almacén de datos especializado, orientado a un tema, integrado, volátil y variante en el tiempo para apoyar un subconjunto específico de decisiones de administración. (Date, 2001)

Para Moss y Atre. “Un Data Mart es un subconjunto de datos derivado del Data Warehouse (DWH). Está diseñado para soportar requerimientos analíticos específicos de una determinada unidad de negocios. Es un repositorio menos ambicioso que un DWH” (2003, p. 45).

Según Silva, 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 (2007, p. 45)

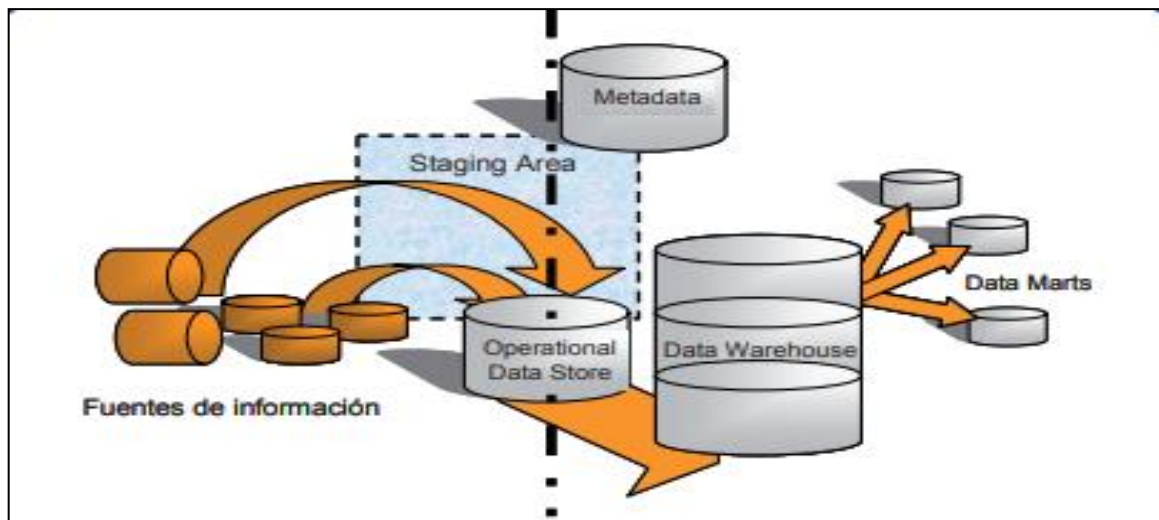


Figura 2.4: Explicación de una Data Mart (Cano, 2007)

2.2.6.1. CLASIFICACIÓN DEL DATA MART

Según Inmon (2002) Los clasifica en tres tipos de Data Mart las cuales son:

a) DATA MART DEPENDIENTE

“Los Datamart dependientes son aquellos que reciben los datos desde una Data Warehouse. En este tipo de DataMart la fuente de los datos es única” (Inmon, 2002, p. 75)

b) DATAMART INDEPENDIENTE

“Los Datamart llamados independientes son aquellos que toman sus datos directamente desde los sistemas transaccionales y no dependen de otros” (Inmon, 2002, p. 75)

c) DATA MAR HÍBRIDO

“Los Data Mart híbridos permite admitir englobar las fuentes de datos de un Data warehouse corporativo con otras fuentes de datos tales como sistemas transaccionales y/o operacionales” (Inmon, 2002, p. 75)

2.2.6.2. TIPOS DE DATA MART

Según Inmon (2002), existen dos tipos de Data Marts.

➤ **DATA MART OLAP**

“Se basan en los populares cubos OLAP, que se construyen agregando, según los requisitos de cada área o departamento, las dimensiones y los indicadores necesarios de cada cubo relacional” (Inmon, 2002, p. 76)

➤ **DATA MART OLTP**

“Pueden basarse en un simple extracto de la data warehouse, no obstante, lo común es introducir mejoras en su rendimiento, las operaciones más usuales aprovechando las características particulares de cada área de la empresa” (Inmon, 2002, p. 75)

2.2.7. VENTAJAS Y DESVENTAJAS

VENTAJAS

Para Nader (2002), los principales beneficios al utilizar una Data Mart son:

- a) Dado que un Data Mart soporta menos usuarios que una data Warehouse se puede optimizar para recuperación más rápida los datos que necesitan los usuarios.
- b) Menores cantidades de datos implica que se procesan antes, tanto las cargas de datos como las consultas.
- c) Las peticiones pueden acotarse al área o red sirve esos datos, sin afectar al resto de los usuarios.
- d) Las aplicaciones cliente, pide las consultas es independiente del servidor que la procesa y del servidor de bases de datos que almacenan la información.
- e) Los costos que implica la construcción de un Data Mart son mucho menos a los de la implementación de un data Warehouse

DESVENTAJAS

Según Vizuite y Yela, (2006), “no permite el manejo de grandes volúmenes de información por lo que muchas veces se debe recurrir a un conjunto de Data Marts para cubrir todas las necesidades de información de la empresa”

2.2.8. DIFERENCIA ENTRE DATA MART Y DATA WEREHOUSE

Según Acosta (2019) da un cuadro de las principales diferencias entre un Data Mart y un data werehouse.

Tabla 5

Principales Diferencias entre un Data Mart y Data warehouse

CRITERIO	DATA WAREHOUSE	DATA MART
Uso	Ayuda a tomar una decisión estratégica.	Ayuda a tomar decisiones tácticas para el negocio.
Objetivo	El objetivo principal es proporcionar un entorno integrado y una imagen coherente de la empresa en un momento determinado.	Utilizado principalmente en una división de negocios a nivel de departamento.
Modelo dimensional	Puede o no puede usarse en un modelo dimensional. Sin embargo, puede alimentar modelos dimensionales.	Se construye enfocado en un modelo dimensional usando un esquema de inicio.
Gestión de datos	Incluye una gran área de la corporación, por lo que se tarda mucho tiempo en procesarla.	Son fáciles de usar, diseñar e implementar, ya que solo puede manejar pequeñas cantidades de datos.
Foco	Se enfoca ampliamente en todos los departamentos. Es posible que incluso pueda representar a toda la empresa.	Está orientado a un área de negocio y se utiliza a nivel de departamento.
Tipo de datos	Los datos almacenados siempre ofrecen más detalle en comparación con Data Mart.	Están contruidos para grupos de usuarios particulares. Por lo tanto, los datos son cortos y limitados.
Normalización	Los almacenes modernos están en su mayoría des normalizados para proporcionar consultas de datos más rápidas y un buen rendimiento de lectura	No hay preferencia entre una estructura normalizada o des normalizada.

Almacenamiento de datos	Diseñado para almacenar datos de decisiones de toda la empresa, no solo datos de marketing.	Modelado dimensional y diseño de esquema en estrella empleado para optimizar el rendimiento de la capa de acceso.
Valor de los datos	Solo lectura desde el punto de vista de los usuarios finales.	Datos transaccionales agrupados alimentados directamente desde el Data Warehouse.
Alcance	Es más útil ya que puede traer información de cualquier departamento.	Contiene datos, de un departamento específico de una empresa. Tiene un uso limitado
Fuente	Los datos provienen de muchas fuentes.	Los datos provienen de muy pocas fuentes.
Tamaño	El tamaño del Data Warehouse puede variar de 100 GB a más de un TB.	El tamaño de Data Mart es inferior a 100 GB.
Tiempo de implementación	El proceso de implementación puede extenderse de meses a años.	El proceso de implementación de Data Mart está restringido a unos pocos meses.

Fuente: Acosta, 2019

2.2.9. METODOLOGÍA RALPH KIMBALL

Según Rodríguez “Ralph Kimball es el autor considerado como el "Gurú" del DWH junto con Bill Inmon. Su metodología se ha convertido en el estándar de facto en el área de apoyo a las decisiones empresariales” (2010, p. 154)

Según Rivadera (s.f.), la Metodología Kimball, es una metodología empleada para la construcción de un almacén de datos (data warehouse) que no es más que, una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza.

Según Matallana y Vivanco (2014), es una #metodología orientada al diseño de base de datos que almacena la información que servirá como apoyo a la toma de decisiones”

Según Rivadera (s. f), la metodología se basa en lo que “Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle)”. Este ciclo de vida del proyecto, está basado en cuatro principios básicos:

- **Centrarse en el negocio:** Concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y crear relaciones sólidas con el negocio.
- **Construir una infraestructura** de información adecuada: Diseñar una base de información única, integrada, fácil de usar, de alto rendimiento.
- **Realizar entregas en incrementos significativos:** Crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses.
- **Ofrecer la solución completa:** Proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios.

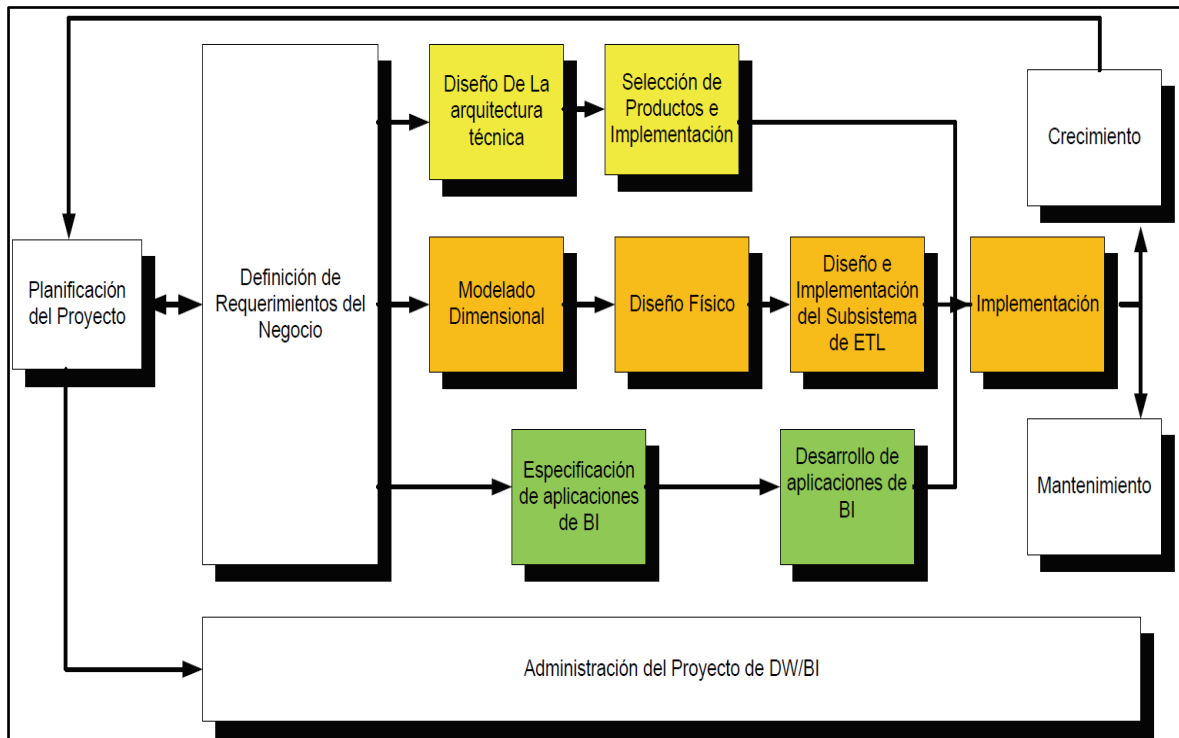


Figura 2.5. Tareas de la metodología de Kimball, denominada Business Dimensional Lifecycle (Riadera, s. f)

2.2.9.1. TAREAS DE LA METODOLOGÍA DE KIMBALL

A. PLANIFICACIÓN DEL PROYECTO

Según Rivadera (s.f), se determina el propósito del proyecto, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.

Según Guillen (2012), “La planificación busca identificar la definición y el alcance del proyecto de data warehouse, el cual se focaliza sobre recursos, perfiles, tareas, duraciones y secuencialidad”

B. DEFINICIÓN DE LOS REQUERIMIENTOS DEL NEGOCIO

Según Rivadera (s.f.) la “definición de los requerimientos es en gran medida un proceso de entrevistar al personal de negocio, Se debe aprender tanto como se pueda sobre el negocio, los competidores, la industria y los clientes del mismo”

Según Guillen (2012), “La definición de los requerimientos del negocio determina los datos necesarios para cumplir los requerimientos analíticos de los usuarios. Diseñar los modelos de datos para soportar estos análisis requiere un enfoque diferente al usado en los sistemas operacionales”

C. MODELO DIMENSIONAL

Según Rivadera (s.f.) La creación de un modelo dimensional es un proceso dinámico y altamente iterativo, y el proceso iterativo consiste en cuatro pasos.

1. Elegir el proceso de negocio (Elección del Data Mart)
2. Elección del objetivo de las tablas de hechos.
3. Elegir las dimensiones.
4. Identificar medidas y las tablas de hechos

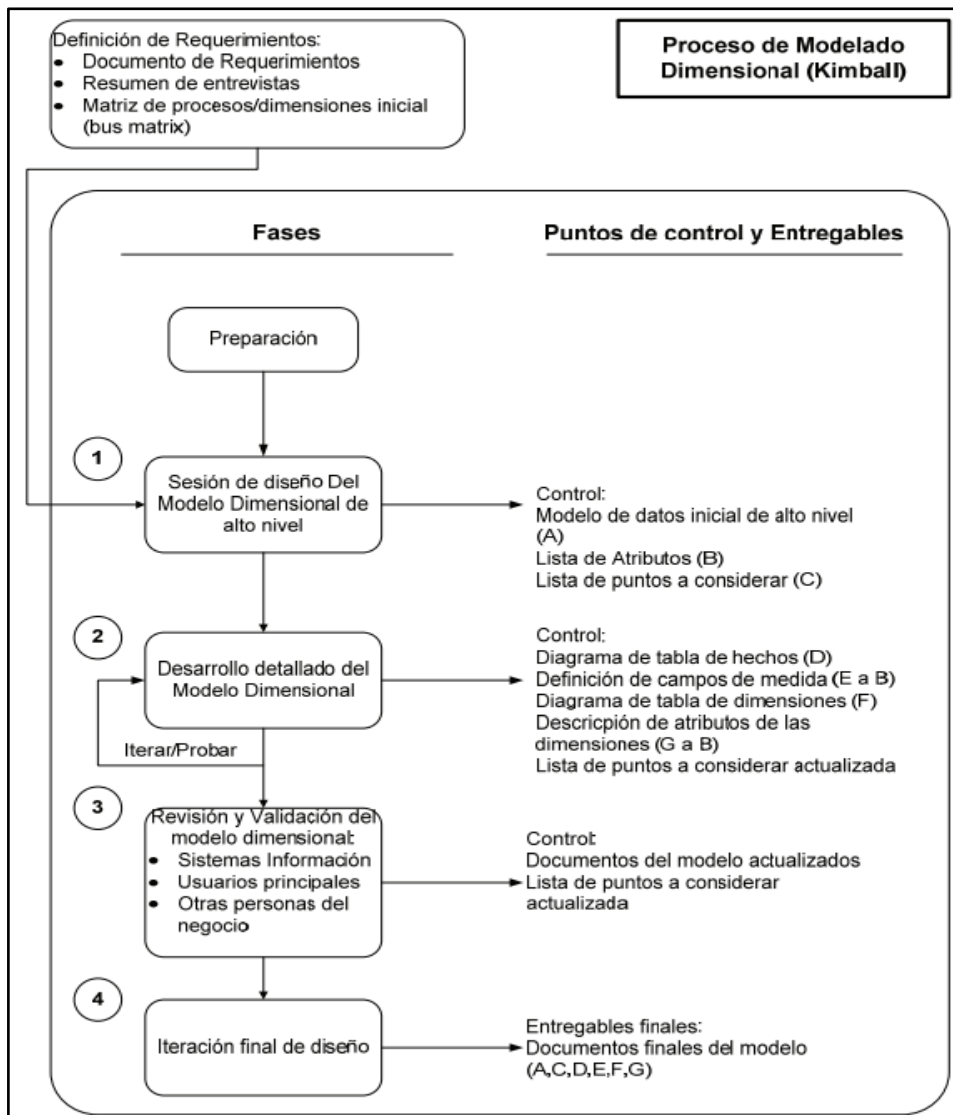


Figura 2.6. Diagrama de flujo del proceso dimensional de Kimball (Rivadera, s.f.)

Según Sullcarayme y Mamani, comienza con una matriz donde se determina la dimensionalidad de cada indicador y luego se especifican los diferentes grados de detalle dentro de cada concepto del negocio, así como la granularidad de cada indicador y las diferentes jerarquías que dan forma al modelo dimensional del negocio (2017, p. 145)

a. MODELO ESTRELLA

Según Cedeño (2005), es un paradigma de modelado que tiene un solo objeto en el medio conectado con varios objetos de manera radial el objeto en el centro de la estrella es la tabla de hechos y los objetos conectados a ella son las tablas de dimensiones. Las tablas de dimensiones solo se relacionan con la tabla de hechos, es decir, no existen relaciones entre estas.

Según Ullman y Widow (1999) define al esquema estrella “forma un diagrama en forma de estrella teniendo en el centro de la estrella una o más tablas de hechos y las puntas de las estrellas a las tablas de dimensiones”

Según Rojas (2014), un “modelo estrella es un modo de representar datos multidimensionales en una base de datos relacional. Las tablas dimensiones guardan información descriptiva acerca de sus miembros y sus relaciones”. Mientras, que las tablas de hechos almacenan datos de negocio

Según Wolff, en general, el modelo multidimensional también se conoce con el nombre de esquema estrella, pues su estructura base es similar: una tabla central y un conjunto de tablas que la atienden radialmente. Este modelo entonces, resulta ser asimétrico, pues hay una tabla dominante en el centro con varias conexiones a las otras tablas. Las tablas Lock-up tienen sólo la conexión a la tabla fact y ninguna más (s. f. p. 32)

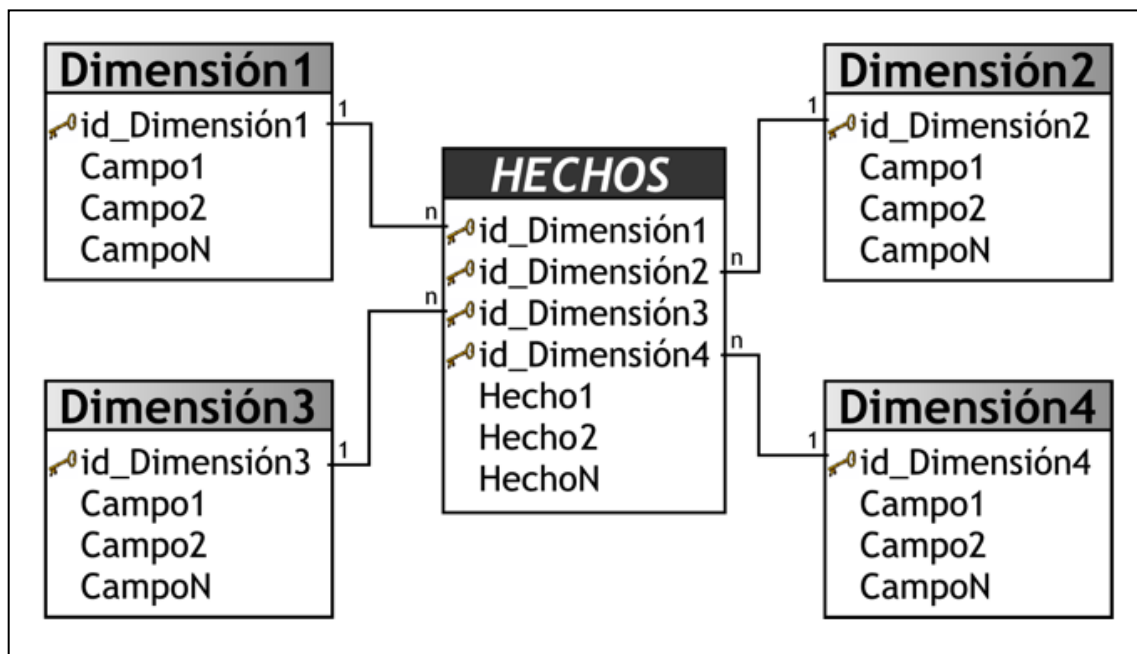


Figura 2.7. Modelo estrella (Bernabéu, 2010)

b. MODELO COPO DE NIEVE

Según Ramos (2016), esta es una estructura más compleja que el esquema en estrella. La diferencia es que algunas de las dimensiones no están relacionadas directamente con la tabla de hechos, sino que se relacionan con ella a través de otras dimensiones. En este

caso también tenemos una tabla de hechos, situada en el centro, que contiene todas las medidas y una o varias tablas adicionales, con un mayor nivel de normalización.

Según Rojas (2014), los esquemas de copos de nieve contienen una tabla de hechos centrales para el tema y numerosas tablas de dimensiones para la información descriptiva sobre las dimensiones del tema. La tabla hechos puede contener varios millones de filas. La información a la que se tiene acceso con más frecuencia se agrega previamente y se resume para mejorar aún más, el rendimiento.

Según Cedeño (2005), el modelo copo de nieve es una extensión del esquema estrella en donde cada una de las puntas de la estrella puede dividirse en más puntas. En esta forma de esquema, las tablas de dimensión pueden tener relación con otras tablas de dimensión.

Según Bernabéu (2010), el modelo copo de nieve representa una extensión del modelo en estrella cuando las tablas de dimensiones se organizan en jerarquías de dimensiones, es decir existe una tabla de hechos central que está relacionada con una o más tablas de dimensiones, quienes a su vez pueden estar relacionadas o no con una o más tablas de dimensiones.

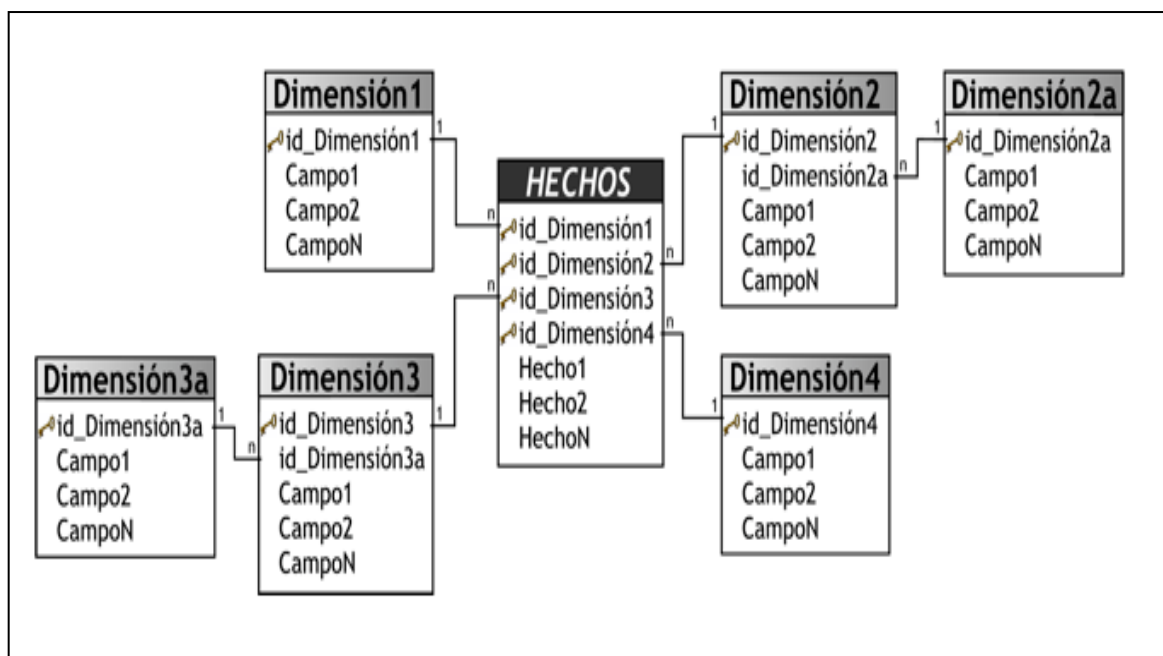


Figura 2.8. Modelo Copo de nieve (Bernabéu, 2010)

TABLA DE HECHOS

Según Kimball y Ross (2013), la tabla de hechos en un “modelo dimensional que almacena las mediciones de los rendimientos resultantes de los eventos, de los procesos de negocio de una organización”

Según Rojas y Zamudio (2016) “El término hecho representa una medida del negocio. Cada fila de una tabla de hechos corresponde a un evento de medición. Los datos de cada fila están en un nivel específico de detalle”

TABLA DE DIMENSIONES

Según Kimball y Ross (2013), las tablas de dimensiones son compañeros integrales de una tabla de hechos. “Las tablas de dimensiones contienen el contexto textual asociado a un evento de la medición de los procesos del negocio, describen el qué, quién, dónde, cuándo cómo y por qué” (Rojas y Zamudio, 2016, p. 43).

D. DISEÑO FÍSICO

Según Rivadera (s. f), se debe concentrar en las siguientes preguntas.

- ¿Cómo puede determinar cuán grande será el sistema de DW/BI?
- ¿Cuáles son los factores de uso que llevarán a una configuración más grande y más compleja?
- ¿Cómo se debe configurar el sistema?
- ¿Cuánta memoria y servidores se necesitan? ¿Qué tipo de almacenamiento y procesadores?
- ¿Cómo instalar el software en los servidores de desarrollo, prueba y producción?
- ¿Qué necesitan instalar los diferentes miembros del equipo de DW/BI en sus estaciones de trabajo?
- ¿Cómo convertir el modelo de datos lógico en un modelo de datos físicos en la base de datos relacional?
- ¿Cómo conseguir un plan de indexación inicial?
- ¿Debe usarse la partición en las tablas relacionales?

Según Robles y Girón (2015), algunos “elementos principales del diseño físico son la definición de convenciones estándares de nombres específicos del ambiente de la base de datos”

E. DISEÑO E IMPLEMENTACIÓN DEL SUBSISTEMA ETL

Según Rivadera (s.f.), en esta parte se “realiza el ETL, la base sobre la cual se alimentará el DW, provenientes de los distintos sistemas y finalmente cargar la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis”

Según Guillen (2012), se definen como “procesos de extracción a aquellos requeridos para obtener los datos que permitirán efectuar la carga del modelo físico, los procesos de carga de datos son necesarios para poblar el Data Warehouse”.

Según Pacco (2013), esta tarea tiene como “principales actividades la extracción, transformación y carga (ETL), estas actividades son altamente críticas ya que tienen que ver con la materia prima Data Warehouse”.

Según Bustamante, Galvis y Gomez (2013), la función del “proceso ETL es de extraer, limpiar, transformar, resumir y formatear los datos que se almacenaran en la bodega de datos, el proceso se puede componer en tres subprocesos o componentes: componente extracción, transformación y carga”

Según Muñoz y Mozón (s.f) Un proceso “ETL bien diseñado extrae datos de las fuentes de datos, hace cumplir estándares de calidad de datos, a fin de que los datos puedan ser utilizados por los desarrolladores para las aplicaciones y los usuarios finales puedan tomar decisiones estratégicas”

Kimball y Caserta, afirman que un sistema ETL extrae los datos desde los sistemas fuente, verifica los estándares de calidad y consistencia que los datos requieren para ser procesados y la entrega de tal manera que a partir de ellos se pueda construir las herramientas para la toma de decisiones. Dicho proceso no solo se refiere al transporte de los datos de una estructura a otra, por el contrario, durante este proceso se agrega valor a los datos en las siguientes actividades: a) Elimina errores y corrige datos faltantes, b) Integra diferentes

fuentes de información y c) Estructura los datos para que puedan ser usadas por las herramientas de análisis (2004, p. 77)

Para Henao, el objetivo de ETL es extraer, transformar y cargar los datos directamente a una base de datos la cual contiene tablas especialmente creadas para los datos de paso o temporales, estas tablas se conocen como Tablas Dimensionales y Tablas de Hecho, las cuales pueden ser actualizadas en cada proceso de carga. (2008, p. 89)

A. FACES DEL PROCESO ETL

“La administración de bodegas de datos requiere de un procesamiento para garantizar la veracidad, integridad y centralización de los datos cuando existen diversas fuentes de información, haciendo necesario utilizar aplicativos especializados para la Extracción, Transformación y Carga de datos (ETL)”. (Duque y Hernández, 2016, p. 72)

1. E: EXTRAER

Etapa donde se extraen y se obtienen los datos de las distintas fuentes de datos tanto internas como externas de la empresa. La extracción selecciona los datos en un formato listo para iniciar el proceso de transformación. (Silva, 2017, p. 35)

- a) **Análisis previo de las necesidades.** “se evalúan las necesidades concretas de la organización en cuanto a movimiento y transformación de datos. se empieza con el tratamiento de unos datos concretos según una necesidad puntual y luego se realizan ampliaciones a medida que se van necesitando”.
- b) **Identificación de archivos.** “se identifica de qué tipo son y en qué formato se encuentran los sistemas fuente. Lo habitual es que los datos de los sistemas de origen provengan de formatos distintos, que tendrán que ser fusionados.”
- c) **Extracción de los datos.** “En función de las necesidades detectadas, se procede a la extracción en sí de dichos datos”

2. T: TRANSFORMAR:

“La fase de transformación consiste en la aplicación de una serie de funciones o reglas de negocio sobre los datos extraídos para convertirlos en datos que, a continuación, serán cargados en la nueva fuente”.

3. L: CARGA:

“La última parte de los procesos ETL es la fase de carga, el momento en el cual los datos procedentes de la fase de transformación son cargados en el sistema de destino”

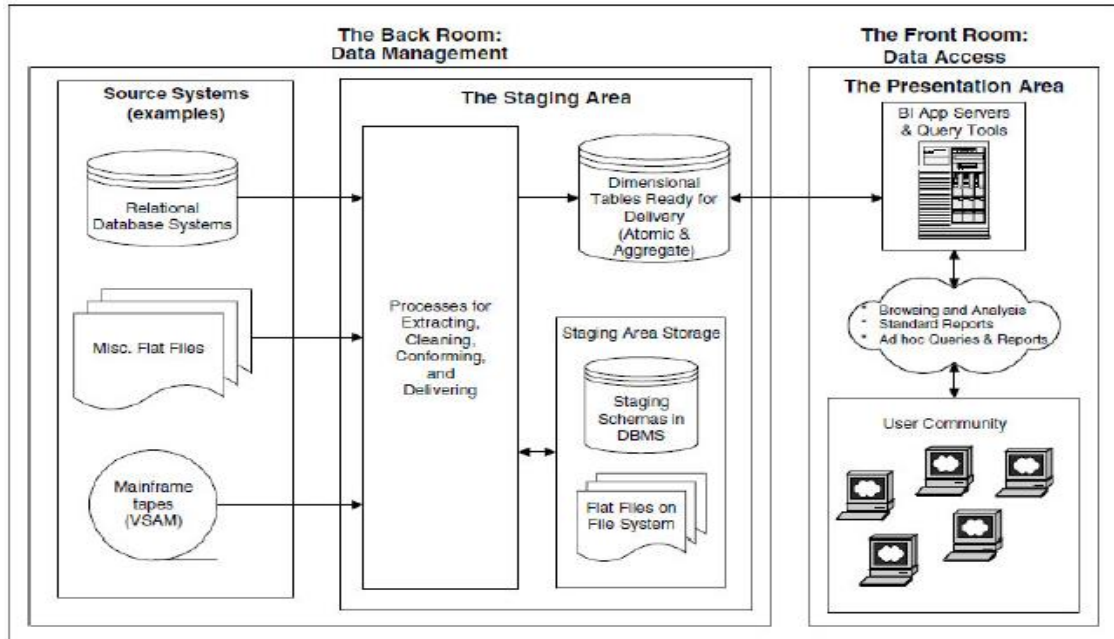


Figura 2. 9: División entre gestión de datos y capa de presentación de un Data warehouse (Kimball y Caserta, 2004).

F. DISEÑO DE ARQUITECTURA TÉCNICA

Según Guillen (2012), el diseño de arquitectura técnica cuenta con tres factores: los requerimientos del negocio, los actuales ambientes técnicos y las directrices técnicas estratégicas futuras planificadas.

“En esta fase se deben tener en cuenta tres factores: los requerimientos de negocio, los actuales entornos técnicos, y las directrices técnicas y estratégicas futuras planificadas por la compañía, lo que permitirá establecer el diseño de la arquitectura técnica del entorno del Data Warehouse” (Pacco, 2013, p. 45)

G. SELECCIÓN DE PRODUCTOS E IMPLEMENTACIÓN

Según Guillen (2012), utilizando el “diseño de arquitectura técnica, es necesario evaluar y seleccionar componentes específicos, como ser la plataforma de hardware, el motor de base de datos, la herramienta de ETL y herramientas de acceso”.

Según Pacco (2013), “evalúa y selecciona cuales son los componentes necesarios específicos de la arquitectura (plataforma de hardware, motor de la BD, herramienta de ETL. Luego se realiza la instalación de los componentes previamente evaluados y seleccionado”.

H. ESPECIFICACIÓN DE APLICACIONES DE BUSINESS INTELLIGENCE

Según Guillen (2012), se “identifican los diferentes roles o perfiles de usuarios para determinar los diferentes tipos de aplicaciones necesarias en base al alcance de los diferentes perfiles” (gerencial, analista del negocio, vendedor, etc.)

I. DESARROLLO DE APLICACIONES DE BUSINESS INTELLIGENCE

Según Rivadera (s.f), “las aplicaciones de BI incluyen un amplio espectro de tipos de informes y herramientas de análisis, que van desde informes simples de formato fijo a sofisticadas aplicaciones analíticas que usan complejos algoritmos e información del dominio”.

J. IMPLEMENTACIÓN

Según Guillen (2012), “la implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales para su correcto funcionamiento se encuentra la capacitación, el soporte técnico, la comunicación”

Según Morales, la implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Existen varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación y las estrategias de feedback. (2019, p. 132)

K. MANTENIMIENTO Y CRECIMIENTO

Según Guillen, Data Warehouse es un proceso, de etapas bien definidas con comienzo y fin, pero de naturaleza espiral, pues acompaña a la evolución de la organización durante toda su historia de la organización, como también afirma Kimball, si se ha utilizado el ciclo de vida, el data warehouse está preparado para evolucionar y crecer. (2012, p. 62)

“Se basa en la necesidad de continuar con las actualizaciones de forma constante para así lograr la evolución de las metas por conseguir” (Pacco, 2013, p. 48)

“Para administrar el entorno de DW actual es importante enfocarse en los usuarios de negocio, los cuales son el motivo de su existencia, medir y proyectar su éxito” (Morales, 2019, p. 61)

L. GERENCIAMIENTO DEL PROYECTO

Según Guillen (2012), “el gerenciamiento del proyecto asegura que las actividades del ciclo de vida dimensional se lleven en forma sincronizada”.

2.2.10. PROCESAMIENTO ANALÍTICO EN LÍNEA OLAP (ON LINE ANALYTICAL PROCESSING)

Según Bernabéu (2010), “OLAP es el acrónimo para Procesamiento Analítico en Línea (On-Line Analytical Processing) que representa una forma de ver los datos resumidos históricos en estructuras multidimensionales (conocidos como cubos OLAP) que permitan el acceso rápido a los mismos”.

Según Tapia (2006), define como el “análisis multidimensional e interactivo de la información de negocios a escala empresarial. El análisis multidimensional consiste en combinar distintas áreas de la organización, y así ubicar ciertos tipos de información que revelen el comportamiento del negocio”.

Según Laudon, OLAP soporta el análisis de datos multidimensionales, el cual permite a los usuarios ver los mismos datos en diferentes formas utilizando múltiples dimensiones, obtener respuestas en línea a preguntas específicas en un lapso de tiempo sumamente rápido aun cuando los datos están almacenados en base de datos sumamente grande. OLAP representa las relaciones entre los datos y cubos dentro de cubos de datos para permitir un análisis de datos más complejos. (2008, p. 21)

Según Leme (2004), “OLAP es un conjunto de herramientas direccionadas a los usuarios finales para acceder a las informaciones del Data Warehouse y que así sea posible crear informes que auxilien la toma de decisión”.

OLAP Son aplicaciones que se encargan de analizar datos del negocio para generar información táctica y estratégica que sirve de soporte para la toma de decisiones logrando

su máxima eficiencia y flexibilidad operando sobre Bases de datos multidimensionales. Se basan en los cubos OLAP, que se construyen agregando, según los requisitos de cada área o departamento, las dimensiones y los indicadores necesarios de cada cubo relacional (Laudon y Laudon, 2008, p. 65).

Según Colaco y Turban (2009), afirman que las tecnologías “OLAP es la multidimensionalidad de los datos. Esto significa que los datos son modelados en una estructura conocida como cubo, que en cada dimensión representa temas como, por ejemplo, producto, cliente y tiempo”

Según Pinheiro (2008), “menciona que la principal ventaja del OLAP es permitir que los datos sean ordenados como a cada gerente les gustaría verlos y no por la visión de los profesionales de las TIC”

Según Vitt y Luckevich (2002), OLAP proporciona un modelo de datos intuitivo y conceptual, para que los usuarios que no tengan experiencia como analistas puedan comprender y relacionar los datos mostrados. Este modelo es llamado análisis multidimensional, siendo habilitado para ver los datos a través de múltiples filtros, o dimensiones, Existen variaciones de OLAP según la cantidad de datos y la eficiencia requerida. OLAP, no se recomienda para consultas complejas y que recorran muchas tablas. (2002, p. 35)

A. MOLAP (Multidimensional online analytical processing).

Según Vitt y Luckevich (2002), “MOLAP Brinda el mayor rendimiento de recuperación de información; porque los datos son estructurados y colocados en niveles especiales que se encuentran en un servidor central”.

Según Ibarra (2006), “los datos se encuentran almacenados en una estructura multidimensional, para optimizar los tiempos de respuesta, el resumen de la información es usualmente calculado por adelantado, que estos valores pre calculados son la base de las ganancias de desempeño”.

Para Moss y Atre. Su premisa base es el pre cálculo de sus estructuras, de modo que siempre se hacen las mismas consultas sobre una misma estructura, se carece de una base de datos

relacional. Cada vez que un sistema cambia algún sector de sus estructuras vuelve a cargar gran parte de sus dimensiones lo cual lo hace poco dinámico, pero a su vez con una gran capacidad de respuesta (2003, p. 125).

El sistema MOLAP permite almacenar físicamente los datos en estructuras multidimensionales de manera que la representación externa y la interna coincidan, disponiendo de estructuras de almacenamiento específicas (Arrays) y técnicas de comparación de datos que favorecen el rendimiento del Data Warehouse (Bernabéu, 2010, p. 263).

Según Torres (2007) el MOLAP tiene las siguientes ventajas y desventajas:

Ventajas

- a) Excelente performance: los cubos MOLAP son construidos para tener una rápida recuperación de datos y esta optimizado para operaciones.
- b) Puede realizar cálculos complejos: ya que todos los cálculos han sido pre generados cuando el cubo se crea. Por lo tanto, los cálculos complejos se almacenan y regresan su resultado rápidamente.
- c) Extracto eficaz de los datos alcanzados con la pre- estructuración de datos agregados.

Desventajas

- a) El paso de proceso (carga de los datos) puede ser absolutamente muy largo, especialmente en volúmenes grandes de los datos. Esto es remediado generalmente haciendo solamente el proceso incremental, es decir, procesando solamente los datos que ha cambiado (generalmente los nuevos datos).
- b) Limitado en el monto de datos a ser manejados. Porque todos los cálculos son construidos cuando se genera el cubo, no es posible incluir grandes cantidades de datos en el cubo en sí mismo. Esto no quiere decir que los datos del cubo no deriven de una gran cantidad de datos. Si es posible, pero en este caso, solo la información de alto nivel puede ser incluida en este.
- c) El acercamiento de MOLAP introduce redundancia de datos.

B. ROLAP (Relational online analytical processing)

“En los sistemas ROLAP, los cubos multidimensionales se generan dinámicamente al instante de realizar las diferentes consultas, haciendo de esta manera el manejo de cubos transparente” (Bernabeu, 2010, p. 35).

Según Vitt y Luckevich (2002), permite tomar ventaja de uno de sus más grandes beneficios, el almacenamiento de inmensas cantidades de datos. El rendimiento de recuperación de la información para ROLAP frecuentemente no es tan rápido como otras opciones de almacenamiento. ROLAP es recomendado para consultas pesadas que no se usan muy a menudo.

“Acrónimo de Relational Online Analytical Processing, almacena los datos en un motor relacional logrando una mejor flexibilidad mediante los tipos de análisis disponibles, tener menor tiempo de respuesta para la elaboración de reportes, análisis de una enorme cantidad de datos”. Se implementa sobre tablas físicas diseñadas siguiendo un modelo en estrella o copo de nieve (Tamayo, M. y Moreno F., 2006).

“ROLAP se utiliza para ahorrar espacio de almacenamiento cuando se trabaja con grandes conjuntos de datos que se consultan con poca frecuencia; por ejemplo, datos exclusivamente históricos” (Kimball y Ross, 2013).

Según Torres (2007) “el ROLAP tiene las siguientes ventajas y desventajas”:

Ventajas

- a) Almacena la información en una base de datos relacional
- b) Utiliza índices de mapas de bits.
- c) Utiliza índices de Join.
- d) Posee optimizadores de consultas.
- e) Cuenta con extensiones de SQL (drill- up, drill- down, etc).

Desventajas

- a) Performance bajo. Ya que ROLAP es esencialmente múltiples Querys de SQL en la base de datos relacional, el tiempo de respuesta se alarga entre el tamaño de la Base de Datos, mientras sea más grande será más lenta.

- b) Limitada funcionalidad SQL. Ya que la tecnología ROLAP utiliza básicamente sentencias SQL o queries de la Base de Datos relacional, y SQL no aporta todas las necesidades de consultas multidimensionales, ROLAP son limitadas a lo que el lenguaje Base de Datos soporte. Se ha desarrollado últimamente herramientas externas que permiten utilizar formulación más compleja que pueda cubrir parte de estas deficiencias.

C. HOLAP (Hybrid online analytical processing).

Según Kimball y Ross (2013), afirman que “HOLAP hace que las agregaciones se almacenen en una estructura multidimensional, y los datos a nivel de detalle, en una base de datos relacional como lo hace el almacenamiento ROLAP, obteniendo así el rendimiento de las consultas”.

Según Vitt y Luckevich, “es un híbrido entre MOLAP y ROLAP, HOLAP no es realmente un modo diferente de almacenamiento de datos”. Más bien es la habilidad para diseminar los datos a través de bases de datos relacionales y multidimensionales con la finalidad de obtener lo mejor de ambos sistemas (2002, p. 91).

Acrónimo Hybrid Online Analytical Process, almacena datos con las dos técnicas anteriores, utilizando MOLAP que ofrece análisis sobre los datos agregados, métricas o indicadores precalculados y ROLAP que ofrece escalabilidad, cálculo en tiempo real de reportes requeridos por usuarios, concurrencia y administración madura de los datos. (Tomayo y Moreno, 2006, p.55)

Según Bernabeu (2010), menciona que “HOLAP constituye un sistema híbrido entre MOLAP Y ROLAP, que combina estas dos implementaciones para almacenar algunos datos en un motor relacional y otros en una base de datos multidimensional”.

2.2.11. SISTEMA GESTOR DE BASES DE DATOS (SGBD)

Nevado (s.f.), un SGBD tiene como funciones principales: la descripción, manipulación y control, que en conjunto permiten: a) Definir los datos a los distintos niveles de abstracción (físico, lógico y externo), b) Manipular los datos en la base de datos, permitiendo insertar, modificar, borrar y consultar los datos, c) Mantener la integridad de la base de datos, y d) Controlar de la privacidad y seguridad de los datos en la base de datos.

Para Silberschatz y Korth. Los Sistemas Gestores de Bases de Datos (SGBD), también llamado DBMS (Data Base Management System) como una colección de datos relacionados entre sí, estructurados y organizados, y un conjunto de programas que acceden y gestionan esos datos, y esos datos se denomina bases de datos (2006, p. 15)

Para Cobo. “La existencia de los lenguajes denominados algebra relacional y cálculo relacional, son la base del éxito comercial de los SGBDR (Sistemas gestores de Base de Datos Relacionales)” (s.f., p. 49).

Para Arreguin. “Es un conjunto de datos interrelacionados, recolectados, que satisfacen las necesidades de información de una comunidad determinada de usuarios, estos son los que dieron a la creación de numerosas bases de datos” (2006, p. 9).

Según Sánchez (2012), “un sistema gestor de base de datos es el software que permite gestionar bases de datos, ocultando la física de la misma y permitiendo manejarla desde un nivel más conceptual”.

A. BASE DE DATOS

Según Gutiérrez (s. f) lo define como un “almacén de datos relacionados con diferentes modos de organización. Una base de datos representa algunos aspectos del mundo real, aquellos que le interesan al usuario”.

Según Opper y Sheldon, a pesar de la falta de una definición absoluta, la mayoría de las fuentes coinciden en que una base de datos, por lo menos, es una colección de datos organizada en un formato estructurado que es definido como metadatos que describe esa estructura. Puede pensar en los metadatos como información sobre los datos almacenados, que define cómo se almacenan éstos en una base de datos. (2009, p. 124)

Según Batai (1994), “las bases de datos presentan una serie de características propias, las cuales son: 1) independencia de datos, 2) eficiencia, 3) Normalización. Estas definen a una base de datos relacional consistente dentro de una arquitectura”.

Según Gómez (2013), “es una colección de datos, donde los datos están lógicamente relacionados entre sí, tienen una definición y descripción comunes y están estructurados de

una forma particular. Una base de datos es también un modelo del mundo real y, como tal, debe poder servir para toda una gama de usos y aplicaciones”.

Según McLeod (2000), “afirma que la base de datos está formada por tablas en las cuales se almacenan los datos. Así mismo, las tablas están formadas por columnas y filas a las cuales se les conoce como campos y registros”.

Una base de datos de un Sistema de Información es la representación integrada de los conjuntos de entidades instancia correspondiente a las diferentes entidades tipo del Sistema de Información y de sus interrelaciones. Esta representación informática (o conjunto estructurado de datos) debe poder ser utilizada de forma compartida por muchos usuarios de distintos tipos. (Camps et al, 2007, p. 21).

B. BASE DE DATOS RELACIONAL

“Una colección de tablas a cada una de las cuales se asigna un nombre único. Cada tabla tiene una estructura similar” (Korth y Silberschatz, 2006, p. 12).

De acuerdo a Osorio (2008), es un paradigma que se ha adoptado en las tecnologías de la información, ninguno como el modelo relacional de las bases de datos se ha consolidado de una manera tan categórica y unánime, pudiéndose decir que la actual orientación a objetos debe su éxito a la consolidación de este modelo en la implementación de las bases de datos.

De acuerdo a Heurtel (2009), una base de datos relacional presenta una organización de los datos basada en el modelo relacional desarrollado en 1970 por Edgar Codd. Es la estructura más extendida actualmente; En una base de datos relacional, los datos se organizan en tablas enlazadas de manera lógica. Una tabla incluye columnas (o campos) que describen una fila (o registro). La relación entre las tablas se establece mediante una columna.

Según Quiroz (2003), afirma que una “base de datos relacional se utiliza para modelar problemas reales y administrar datos dinámicamente. Es una forma sencilla de representar los datos en este modelo y es la representación por medio de tablas, una para cada relación”.

(IBM, 1920), “Es un conjunto de tablas y se manipula de acuerdo con el modelo de datos relacional. Contiene un conjunto de objetos que se utilizan para almacenar y gestionar los datos, así como para acceder a los mismos”.

2.2.12. POBLACIÓN

Para Chávez (2007), “la población es el universo de estudio de la investigación, sobre el cual se pretende generalizar los resultados, constituida por características o estratos que le permiten distinguir los sujetos, unos de otros”.

Según Tamayo y Tamayo (1997), “La población se define como la totalidad del fenómeno a estudiar donde las unidades de población poseen una característica común la cual se estudia y da origen a los datos de la investigación”.

“Una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones” (Hernández et al. 2014, p. 170)

2.2.13. CENSO

Para Billy y Ango, “Consiste en obtener mediciones del número total de elementos que componen la población (Teórica o muestreada). La realización del censo es factible cuando la población es pequeña o se dispone entre otros, de recursos económicos, lógicos” (2018, p. 254)

Para Hernández et al. “No siempre, pero en la mayoría de las situaciones sí realizamos el estudio en una muestra. Sólo cuando queremos efectuar un censo debemos incluir todos los casos (personas, animales, plantas, objetos) del universo o la población” (2014, p. 172).

CAPITULO III

METODOLOGÍA DE LA INVESTIGACIÓN

3.1. TIPO Y NIVEL DE LA INVESTIGACIÓN

B. TIPO DE INVESTIGACIÓN

Para Supo (2015), “en los estudios observacionales no existe intervención de ningún tipo por parte del investigador, de manera que los datos observados y la información consignado refleja la evolución natural de los eventos”. Por esta consideración la investigación es de tipo observacional.

Para Supo (2015), los estudios retrospectivos “utilizan datos que se obtienen de registros preexistentes, datos que provienen de mediciones en donde el investigador no tuvo participación alguna. A este tipo de información se le suele llamar datos secundarios”. Por esta consideración la investigación es de tipo retrospectivo.

Para Supo (2015), en un estudio transversal todas las variables (incluyendo la variable de estudio) “son medidas en una sola ocasión bajo esta condición, si realizamos comparaciones entre estas mediciones se les suele llamar entre muestras independientes, aunque el nombre correcto sería entre grupos independientes”. Por esta consideración la investigación es de tipo transversal.

C. NIVEL DE INVESTIGACIÓN

Según Bernal (2010), la investigación descriptiva es la capacidad para seleccionar las características fundamentales del objeto de estudio y su descripción detallada. Esta investigación se soporta fundamentalmente en técnicas como la encuesta, la entrevista, la observación y la revisión documental. Algunos temas de investigación descriptiva son: a) estudios de carácter diagnóstico, b) diseños de guías, modelos, productos, prototipos, etc., c) estudio de mercado, d) estudio de descripción de conducta, e) estudio de tiempo y movimiento.

Para Supo (2015), “la investigación descriptiva es el primero de los niveles cuantitativos. Aquí, el análisis estadístico es indispensable a la hora de hacer cálculos para completar los

objetivos de estudio. Los estudios descriptivos son “univariados” por qué el análisis de datos se centra en una sola variable”.

Para Hernandez, et al. (2014) Los estudios descriptivos buscan especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que se someta a un análisis. Es decir, únicamente pretenden medir o recoger información de manera independiente o conjunta sobre los conceptos o las variables a las que se refieren, esto es, su objetivo no es indicar como se relacionan estas. Por esta consideración el nivel de investigación es descriptiva.

3.2. DISEÑO DE LA INVESTIGACIÓN

La investigación está enmarcada en el diseño no experimental, puesto que el caso de estudio no amerita la manipulación de las variables. Asimismo, se ha considerado el diseño transversal, ya que la recolección de datos se realizará en un solo momento.

La información que se necesita para el estudio, se ha recolectado de la base de datos y registros digitales existentes para el año 2018 y parte del 2019 en la Red de Salud Huamanga, esta información se procesara mediante la metodología de Kimball y los resultados obtenidos serán presentadas mediante indicadores para el control CRED.

3.3. POBLACIÓN Y MUESTRA

POBLACIÓN

La población está conformada por todas las atenciones de control de crecimiento y desarrollo, de los niños y niñas de 0 a 11 años, de todos los establecimientos de salud, pertenecientes a la Red de Salud Huamanga, durante el año 2019.

MUESTRA

No existe muestra, porque se utiliza toda la población, siendo un censo.

3.4. VARIABLES E INDICADORES

3.4.1. DEFINICIÓN CONCEPTUAL DE LAS VARIABLES

VARIABLES

Estado nutricional. – Es el estado de salud de una persona, de cómo se encuentra de acorde a su edad o etapa de desarrolla de vida, como también el estado nutricional es el encargado

de que una persona se encuentre sano o enfermo en lo largo de su etapa de vida, y donde una persona se tiene que tener mejor alimentación y encontrarse sano y fuerte es en la etapa de la infancia, ya que mientras se tenga mejor alimentación en esta etapa de la infancia se podrá ver mejor en la etapa de la adultez.

Red de salud Huamanga. – Es una Institución encargada de velar y controlar a los centros de salud, organizándolas por Micro Redes y por territorio a lo largo de la provincia de huamanga, está a la vez brindara información de las diferentes atenciones que se está realizando a la población y la dirección regional de salud.

VARIABLES DESCRIPTIVAS

Estado de control de Crecimiento. – Proceso mediante el cual se hace una serie de mediciones al niño o niña, acorde a su edad, toma de medidas talla peso.

Estado de control de Desarrollo. – Proceso de mediciones que se realiza al niño o niña, de acuerdo a su edad, mediante el cual se visualiza su maduración y estado nutricional.

Unidad de estadística, Informática y Telecomunicaciones. – Unidad que centraliza la información proveniente de las micro redes y Centros de Salud, encargada de brindar información a la Red de Salud Huamanga.

3.4.2. DEFINICIÓN OPERACIONAL DE LAS VARIABLES

VARIABLES DE INTERÉS

X: Estado Nutricional

Y: Red de Salud de Huamanga.

VARIABLES DESCRIPTIVAS

X1: Estado de Control de Crecimiento.

X2: Estado de Control de Desarrollo.

Y1: Unidad de Estadística, Informática y Telecomunicaciones.

La operacionalización de las variables, se muestra en el anexo A.

3.5. TÉCNICAS E INSTRUMENTOS PARA RECOLECTAR INFORMACIÓN

3.5.1. TÉCNICAS

Se utiliza la técnica de análisis documental, para recolectar datos en relación a la variable de control de crecimiento y desarrollo y a las variables descriptivas estado de control de crecimiento.

3.5.2. INSTRUMENTOS

Se utilizó la ficha de análisis documental que se denominó, ficha de análisis de la base de datos. Ficha que permite obtener datos a partir de la base de datos y a partir del cubo, sobre las atenciones de control, se muestra en el anexo B.

3.6. HERRAMIENTAS PARA EL TRATAMIENTO DE DATOS

Las herramientas tecnológicas que se utilizó, pertenecen a la línea de Microsoft, debido a que los usuarios están más familiarizados con dichas tecnologías, que por lo general a los usuarios les ofrece una alta gama de tecnologías que facilitan en la elaboración de proyectos, enfocados a una línea en específico, esto con la intención de optimizar el tiempo de desarrollo. Es por ello que se seleccionó las tecnologías enfocadas al proceso de desarrollo de Data Mart que nos permite consolidar información. La cual se muestra en la tabla 6

Tabla 6

Herramientas tecnológicas para el tratamiento de datos

SOFTWARE	FABRICANTE	SERVICIO
Windows 10 pro	Microsoft Corporation	Sistema operativo para escritorio, da la facilidad de instalar herramientas de desarrollo, siendo la última versión y la más estable a la actualidad de la línea de sistemas operativos producida por Microsoft Corporation,
Sql server business intelligence development studio 2012	Microsoft Corporation	Es el entorno que utilizará para desarrollar cubos de Procesamiento analítico en línea (OLAP) y modelos de minería de datos en SQL Server Analysis Services.

Microsoft Sql Server 2012	Microsoft Corporation	Siendo una de las versiones más utilizadas del sistema operativo Windows para servidores, ofrece el marco para instalar el servidor web y el servidor de datos.
Visual Studio 2012	Microsoft Corporation.	Es un Entorno Integrado de Desarrollo para Sistemas operativos Windows. Soporta varios lenguajes de programación.
Power BI	Microsoft Corporation	Es una solución de análisis empresarial que permite visualizar los datos y compartir información con toda la organización, o insertarla en su aplicación o sitio web.
Microsoft Excel	Microsoft Corporation	Es una hoja de cálculo desarrollada por Microsoft cuyo trabajo está en base a datos numéricos.
Microsoft Project	Microsoft Corporation	Es un software de administración de proyectos para asistir a administradores de proyectos en el desarrollo de planes, asignación de recursos a tareas, dar seguimiento al progreso, administrar presupuesto y analizar cargas de trabajo.

Fuente: Elaboración propia.

3.7. TÉCNICAS PARA APLICAR LA METODOLOGÍA KIMBALL

Como se observa la revisión literaria desarrollada en el capítulo II, sección 2.2.9, formula el proceso, que se considera las fases para el desarrollo del Data Mart que aplica la metodología de Kimball, se muestra en las siguientes tablas.

Tabla 7

Planificación del proyecto del Data Mart

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Planificar las tareas a realizar en la construcción del Data Mart.	Cronograma de actividades.	Disponibilidad de recursos.	Analista
		Complejidad del proyecto	

Fuente: Robles y Girón, 2015

Tabla 8

Identificar los requerimientos.

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Definir requerimientos funcionales y no funcionales	Resumen de requerimientos funcionales y no funcionales	Entrevista con los expertos del negocio	Analista. Experto del negocio. Usuario.
		Entrevista con los usuarios,	
		Análisis documental de los procesos del negocio.	

Fuente: Robles y Girón, 2015

Tabla 9

Diseño de la arquitectura técnica

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Diseñar la arquitectura técnica	Arquitectura técnica	Revisión de los requisitos funcionales.	Analista Diseñador
		Revisión del entorno tecnológico.	

Fuente: Robles y Girón, 2015

Tabla 10

Selección de producto

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Selección de producto	Herramientas	Software	Analista
		Hardware	Desarrollador

Fuente: Robles y Girón, 2015

Tabla 11

Modelo dimensional.

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Realizar el modelo dimensional	Modelo dimensional	Analizar el proceso de negocio escogido	Diseñador
		Elección del objetivo de las tablas de hechos	Analista
Diseñar las tablas de dimensiones	Diagrama de las tablas de dimensiones	Definir nombre a la tabla de dimensión Añadir un campo de la clave principal Redefinir nombres de campos	Analista de negocio, desarrollador
Realizar las tablas de hechos	Diagrama de las tablas de hechos	Definir nombre a la tabla de hechos Definir la clave primaria Añadir campos de hechos como indicadores	Analista de negocio, desarrollador
Hacer uniones	Diagrama del esquema de uniones	Esquema estrella	desarrollador
Realizar el modelo lógico	Diagrama del modelo lógico	Esquema estrella	Desarrollador

Fuente: Robles y Girón, 2015

Tabla 12

Diseño físico

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Realizar el modelo físico.	Diseño físico	Convertir el modelo lógico a modelo físico.	Diseñador Analista
		Generar SQL a partir del modelo físico.	

Fuente: Robles y Girón, 2015

Tabla 13.

Proceso ETL

TAREA	ENTREGABLE	TÉCNICA	RESPONSABLE
Extracción, transformación de la carga inicial	Carga del almacén intermedio	Limpieza de datos Procesos ETL	Desarrollador
Actualización de las tablas	Datos cargados en el Data Mart	Carga de todas las tablas y dimensiones	Analista de negocio, desarrollador
Creación de cubos multidimensionales	Cubo multidimensional	Creación de indicadores, atributos y jerarquías.	Desarrollador

Fuente: Robles y Girón, 2015

Tabla 14

Implementación

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Implementación	Implementación del cubo multidimensional	Creación de indicadores, atributos y jerarquías.	Desarrollador

Fuente: Robles y Girón, 2015

Tabla 15.

Especificación de aplicaciones para usuarios finales

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Especificación de aplicación para usuarios finales	Roles	Identificación de la parte operativa y estratégica	Analista

Fuente: Robles y Girón, 2015

Tabla 16

Mantenimiento y crecimiento

TAREA	ARTEFACTO	TÉCNICA	RESPONSABLE
Mantenimiento y crecimiento	Mantenimiento y crecimiento	Identificación de nuevos requerimientos	Analista Desarrollador Especialista del negocio

Fuente: Robles y Girón, 2015

CAPITULO IV

RESULTADO DE LA INVESTIGACIÓN

4.1. RESULTADOS APLICANDO LA METODOLOGÍA RALPH KIMBALL

4.1.1. PLANIFICACIÓN DEL PROYECTO

Como primera tarea la metodología nos pide realizar un plan de trabajo la cual nos permitirá seguir de manera adecuada la metodología y cada una de las tareas, para lo cual se realiza el cronograma de actividades que se muestran en el anexo C.

4.1.2. DEFINICIÓN DE LOS REQUERIMIENTOS

Para definir los requerimientos se realiza una entrevista con el responsable de la unidad de estadista, informática y telecomunicaciones (Stakeholders), para poder recopilar los requerimientos y aclarar todas las dudas que tengan, centrándonos siempre en la necesidad principal que es el control CRED, para ello se elabora una guía de entrevista con una serie de preguntas que nos permitirá generar los requerimientos funcionales finales, se presenta en el anexo D.

4.1.2.1. REQUERIMIENTOS FUNCIONALES

Tabla 17.

Lista de requerimientos.

ÍTEM	PREGUNTA	FINALIDAD
1	¿Cuántos niños y niñas llevan el control CRED durante el año?	Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos
2	¿Número de atenciones de control CRED en recién nacidos durante el año?	Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos
3	¿Número de atenciones de Control CRED, en niños y niñas de 01-11 meses?	Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos

- | | | |
|-----------|--|--|
| 4 | ¿Número de atenciones de Control CRED, en niños y niñas de 1 año – 1 año con 11 meses? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 5 | ¿Número de atenciones de Control CRED, en niños y niñas de 2 años – 2 años con 11 meses? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 6 | ¿Número de atenciones de Control CRED, en niños y niñas de 3 años – 3 años con 11 meses? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 7 | ¿Número de atenciones de Control CRED, en niños y niñas de 4 años – 4 años con 11 meses? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 8 | ¿Número de atenciones de Control CRED, en niños y niñas mayores de 5 años? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 9 | ¿Durante el año el personal de salud a cuantos niños y niñas ha atendido? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 10 | ¿Cuántos niños y niñas terminaron su control CRED, durante el año? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 11 | ¿Qué centro de salud tiene en mayor cantidad de niños y niñas mayores de 5 años? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 12 | ¿En qué fechas fueron atendidos los niños y niñas y por qué profesional? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 13 | ¿Cuántos niños y niñas completaron sus atenciones de Control CRED, según su edad? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |
| 14 | ¿Cuántos niños o niñas cuentan con buena salud según el estado nutricional? | Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos |

- 15 ¿Qué puesto de salud muestra a más niños con mejor estado nutricional? Conociendo la cantidad podrá realizar la monitorización de indicadores y emitir reportes oportunos

Fuente: Elaboración propia.

4.1.2.2. REQUERIMIENTOS NO FUNCIONALES

- Brindar la mejor información correspondiente a los usuarios, en el menor tiempo posible.
- El Data Mart debe funcionar sobre la plataforma de Microsoft Windows 10, Windows server 2012, con el motor de base de datos de SQL Server 2012, Analysis Services, para la construcción de cubos.
- Utilización de la tecnología de Microsoft Power BI, para la elaboración de los reportes hacia los usuarios finales.

4.1.3. DISEÑO DE LA ARQUITECTURA TÉCNICA

NIVEL DE DATOS

Para el nivel de datos es preferible empezar analizando los datos de fuentes, requerimientos y procesos que maneja la institución, con ello se crea un tipo de base de datos y la estructura de tablas.

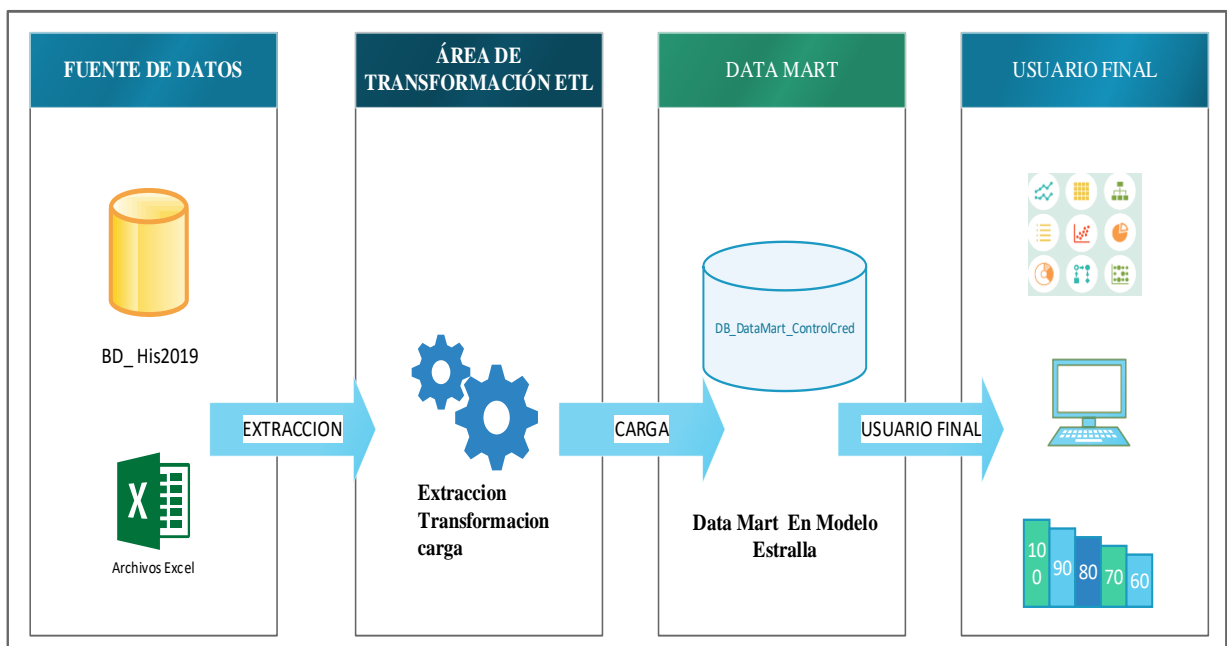


Figura 4.1: Arquitectura técnica

La arquitectura técnica está compuesta por los siguientes niveles:

1. **Fuente de datos:** fuentes principales con la que contamos es la Base de datos His2019, siendo solo una sola tabla que tiene como datos las atenciones de todos los centros de salud que se descarga del sistema web HisMinsa con la que cuenta la institución, y las otras fuentes son obtenidas de los archivos Excel.
2. **Área de transformación:** En donde se realiza el proceso ETL, donde tiene que extraer los datos obtenidos de las diferentes fuentes de datos que nos brinda la institución, prosiguiendo la transformación de los datos, finalmente la carga al Data Mart.
3. **Data mart:** Sera nuestra área especializada de datos, a la cual nos permitirá realizar consultas y reportes necesarios y así cumplir con los requerimientos del usuario.
4. **Usuario Final:** Es la parte final del proyecto, que es la entrega de los reportes del proyecto que serán visualizados por el usuario final, quien podrá interactuar según sus necesidades, ya que será entregado en Microsoft Power BI.

4.1.4. SELECCIÓN DE PRODUCTO

HARDWARE

La base de datos y los servidores de OLAP deben de tener las siguientes configuraciones en función al hardware,

- Memoria Ram 8 GB
- Disco duro 1 TB
- Procesador Intel Core i5 2.70 GHz

SOFTWARE

Para la parte del software se necesitará los siguiente:

Tabla 18

Software para el proceso de diseño de Data Mart

COMPONENTE DEL PROCESO	HERRAMIENTA
Proceso ETL	SQL server Busines Intelligence
Data Mart	SQL Server 2012
Cubos	Analysis Service
Construcción de interfaces	Power BI
Modelo Lógico y físico	Erwin Data Modeler

Fuente: Elaboración propia

4.1.5. MODELO DIMENSIONAL

4.1.5.1. REALIZACIÓN DEL MODELO DIMENSIONAL

En esta parte del modelo dimencional se mostrara los componentes del modelo, desarrollando el metodo de los cuatro pasos.

A. ELECCIÓN DEL DATA MART

En este paso por ende ya se tiene el Data Mart elegido que es el Data Mart del control de crecimiento y desarrollo de los niñas y niños. De todos los centros de salud que están afiliados o son pertenecientes a la red de Salud de Huamanga, que es perteneciente a la provincia de Huamanga, esto no nos limita que se puede realizarse otros Data Marts, como por ejemplo se podría realizar un Data Mart para lo que es la Anemia en niños menores de 5 años, entre otros, pero nos centraremos más en el control CRED, debido que gracias a ellos se puede saber el seguimiento de atenciones para su estado nutricional de cada niño y niña.

Debido a que no se cuenta con una base de datos general de la Red de Salud de Huamanga se procede a su construcción y modelado de base de datos, y elección adecuada de dimensiones y hechos, para ello nos basaremos y ayudaremos con los requisitos funcionales, ya previamente definidos en la Tabla 4.1, de capítulo IV.

Tabla 19.

Lista de Indicadores

Nº	INDICADORES
IND 1	Número Total de atenciones de niños y niñas en los diferentes centros de salud.
IND 2	Numero de atenciones de control CRED, en recién nacidos.
IND 3	Numero de atenciones de control CRED en niños y niñas de 1 mes a 11 meses de edad.
IND 4	Numero de atenciones de control CRED en niños y niñas de 1 año a 1 año con 11 meses de edad.
IND 5	Numero de atenciones de control CRED en niños y niñas de 2 años a 2 años con 11 meses de edad.
IND 6	Numero de atenciones de control CRED en niños y niñas de 3 años a 3 años con 11 meses de edad.

IND 7	Numero de atenciones de control CRED en niños y niñas de 4 años a 4 años con 11 meses de edad.
IND 8	Numero de atenciones de control CRED, en niños y niñas mayores de 5 años de edad.
IND 9	Ganancia inadecuada de peso del niño o niña.
IND 10	Numero de Atenciones por estado nutricional.
IND 11	Recién nacido peso extremadamente bajo.
IND 12	Recién nacido con peso muy bajo.
IND 13	Recién nacido con peso bajo al nacer.
IND 14	Niños o niñas con desnutrición aguda.
IND 15	Niños o niñas con desnutrición severa.
IND 16	Niño o niña con sobre peso.
IND 17	Niño o niña con tamizaje de anemia.

Fuente: Elaboración propia

Tabla 20

Lista de perspectivas

N°	Perspectivas
1	Tiempo
2	Sistema de Información Sanitaria (HIS)
3	Ubigeo
4	Centro de Salud
5	Paciente (niño y niña)
6	Profesional Salud

Fuente: Elaboración Propia

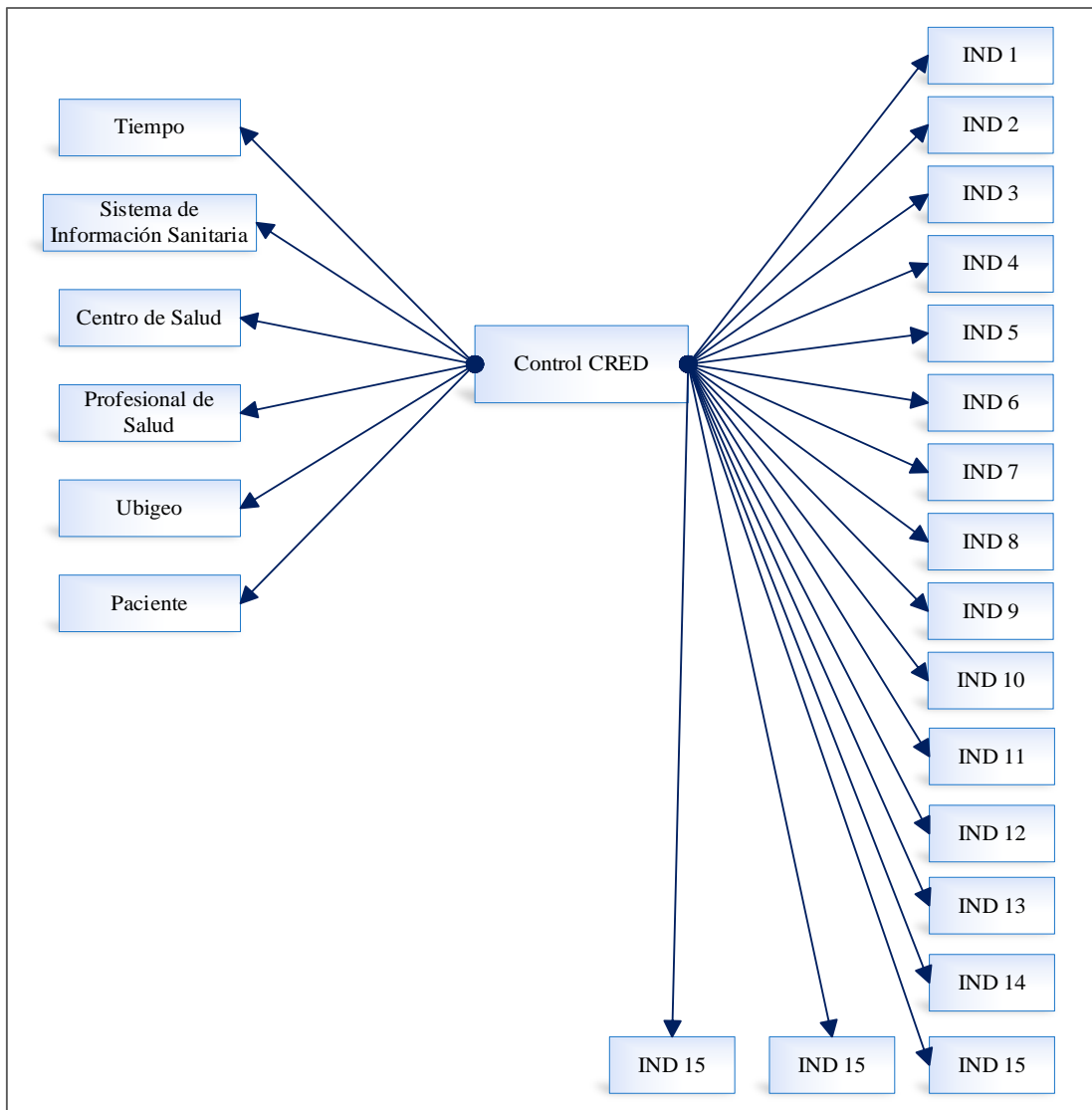


Figura 4.2: Diagrama del modelo conceptual inicial

B. ELECCIÓN DEL OBJETIVO DE LAS TABLAS DE HECHOS.

En esta parte consiste en declarar claramente cuál será el objetivo de la tabla de hechos, es decir se tiene que definir claramente que registros tendrá la tabla de hechos del Data Mart.

Tabla 21

Objetivo de la tabla de hecho

TABLA DE HECHO	OBJETIVO
Control CRED	“Tener las atenciones del control CRED en los niños y niñas en sus diferentes edades”

Fuente: Elaboración propia.

C. ELECCIÓN DE LAS DIMENSIONES

En esta parte de la metodología Kimball nos dice que se tiene que elegir adecuadamente que dimensiones serán que acompañen a la tabla de hecho, para ello se muestra en la tabla 4.3.

Tabla 22

Dimensiones de la tabla de hecho

TABLA DE HECHO	OBJETIVO	DIMENSIONES
Control Cred	“Tener las atenciones del control CRED en los niños y niñas en sus diferentes edades”	- Tiempo - Profesional de salud - Ubigeo - Centro de Salud - Paciente (niños y niñas) - Sistema de Información Sanitaria (HIS)

Fuente: Elaboración propia.

Para la elección de dimensiones, se ha tenido que elegir a la necesidad del Data Mart, y modelado adecuadamente, definidos previamente en las tablas 4.2. y 4.3. en la sección del Capítulo IV.

D. ELECCIÓN DE HECHO (FACT)

El objetivo de cada tabla de hecho, permite definir los hechos y hace claro el alcance que estos deben tener, en nuestro caso contamos con una sola tabla de hecho que es el control CRED.

Tabla 23.

Elección de hechos

TABLA DE HECHO	HECHOS
Control CRED	<ul style="list-style-type: none">- Total de atenciones- Estado nutricional- Recién nacidos- Niños de un año- Niños de dos años- Niños de 3 años- Niños de 4 años- Niños de 5 años- Niños mayores de 5 años- Ganancia inadecuada de peso- Recién nacido peso extremadamente bajo- Recién nacido peso muy bajo- Recién nacido con peso muy bajo al nacer- Desnutrición aguda- Desnutrición severa- Sobrepeso- Tamizaje de anemia

Fuente: Elaboración propia.

4.1.5.2. DIAGRAMA DE LAS TABLA DE DIMENSIONES

En esta etapa de acuerdo a las dimensiones ya elegidas anteriormente se elaborará de manera detallada y sus jerarquías de cada uno de ellas.

i. DIMENSIÓN TIEMPO

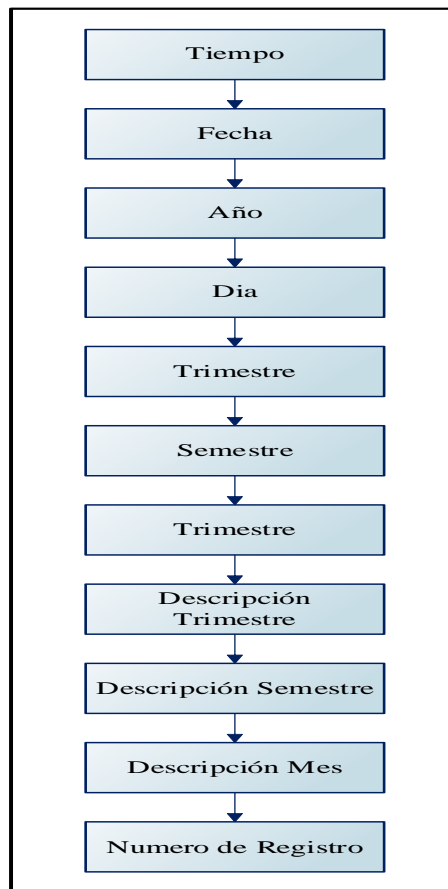


Figura 4.3. Jerarquía de atributos dimensión del tiempo.

Tabla 24:

Detalle de dimensión tiempo.

NOMBRE DEL ATRIBUTO	DESCRIPCIÓN DEL ATRIBUTO
Tiempo key	Representa la llave principal de la dimensión tiempo.
Tiempo fecha	Representa la fecha que fueron atendidas cada paciente.
Tiempo año	Representa el año que fue atendida cada paciente.
Tiempo mes	Representa el mes que fue atendida cada paciente.
Tiempo día	Representa el día que fue atendida cada paciente.
Tiempo trimestre	Representa a que trimestre del año pertenece cada atención.
Tiempo semestre	Representa a que semestre del año pertenece cada atención.
Tiempo descripción trimestre	Representa a la descripción del trimestre.

Tiempo semestre	descripción	Representa a la descripción del semestre.
Tiempo mes	descripción	Representa la descripción de cada mes del año.
Paciente registro	número de	Representa al número de atención que fue registrado en el sistema.

Fuente: elaboración propia.

ii. DIMENSIÓN CENTRO DE SALUD

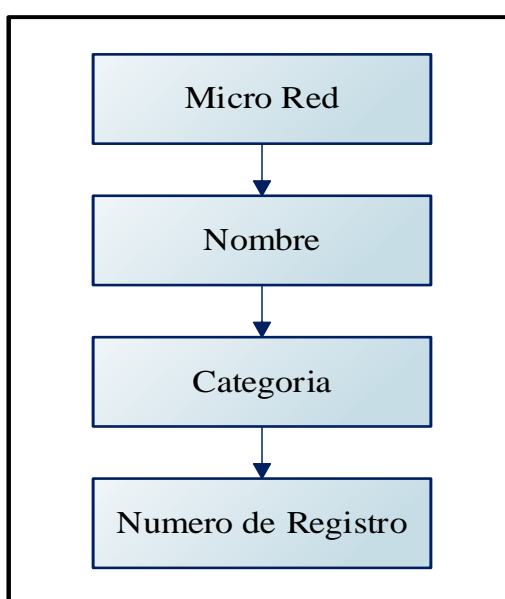


Figura 4.4: Jerarquía de atributos de la dimensión centro de salud

Tabla 25

Detalle dimensión centro de salud.

NOMBRE DEL ATRIBUTO	DESCRIPCIÓN DEL ATRIBUTO
Centro salud key	Representa la llave principal de la dimensión.
Centro salud micro red	Representa el nombre de la micro red a al que pertenece el centro de salud.
Centro salud nombre	Representa el nombre del centro de salud.
Centro salud categoría	Representa al número de categoría que pertenece el centro de salud.

Número de registro	Representa al número de atención que fue registrado en el sistema.
--------------------	--

Fuente: Elaboración propia

iii. DIMENSIÓN PROFESIONAL SALUD

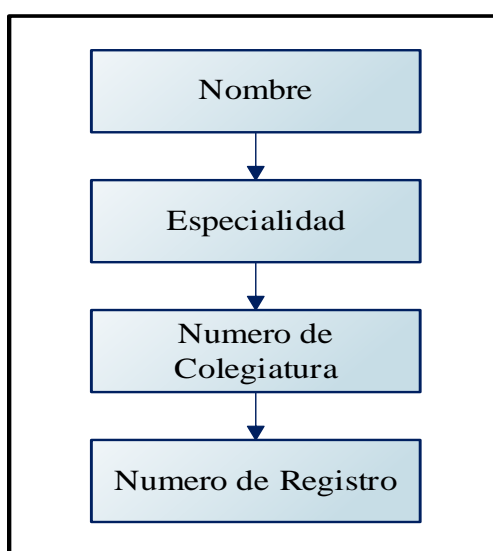


Figura 4.5. Jerarquía de atributos de la dimensión profesional salud

Tabla 26

Detalle dimensión profesional salud

NOMBRE DEL ATRIBUTO	DESCRIPCIÓN DEL ATRIBUTO
Profesional salud key	Representa la llave principal de la dimensión.
Profesional Nombre	Representa el nombre completo del profesional de salud.
Profesional Especialidad	Representa su especialidad del profesional, el nombre de su carrera profesional.
Profesional número de colegiatura	Representa el número de colegiatura, que le fue asignado por su respectiva colegiatura.
Número de registro	Representa al número de atención que fue registrado en el sistema.

Fuente: Elaboración propia

iv. DIMENSIÓN PACIENTE

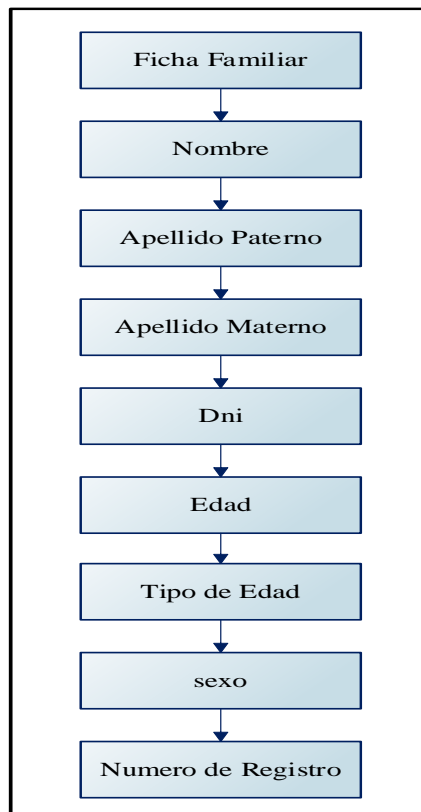


Figura 4.6. Jerarquía de atributos de la dimensión paciente

Tabla 27

Detalle dimensión paciente

NOMBRE DEL ATRIBUTO	DESCRIPCIÓN DEL ATRIBUTO
Paciente key	Representa la llave principal de la dimensión.
Paciente ficha familiar	Representa el nombre completo del profesional de salud.
Paciente Nombre	Representa el nombre paciente
Paciente apellido paterno	Representa el apellido paterno del paciente.
Paciente apellido materno	Representa el apellido materno del paciente.
Paciente DNI	Representa el número de DNI del paciente.
Paciente edad	Representa la edad del paciente.

Paciente tipo edad	Represente el tipo de edad del paciente ya sea días meses años.
Paciente sexo	Representa el sexo del paciente.
Número de registro	Representa al número de atención que fue registrado en el sistema.

Fuente: Elaboración propia

v. DIMENSIÓN UBIGEO

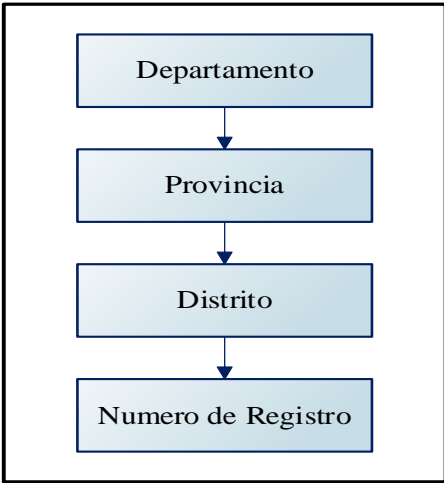


Figura 4.7. Jerarquía de atributos de la dimensión ubigeo

Tabla 28.

Detalle dimensión ubigeo

NOMBRE DEL ATRIBUTO	DESCRIPCIÓN DEL ATRIBUTO
ubigeo key	Representa la llave principal de la dimensión.
Departamento	Representa el nombre del departamento
Provincia	Representa el nombre de la provincia
Distrito	Representa el nombre del distrito
Número de registro	Representa al número de atención que fue registrado en el sistema.

Fuente: Elaboración propia

vi. **DIMENSIÓN HIS**

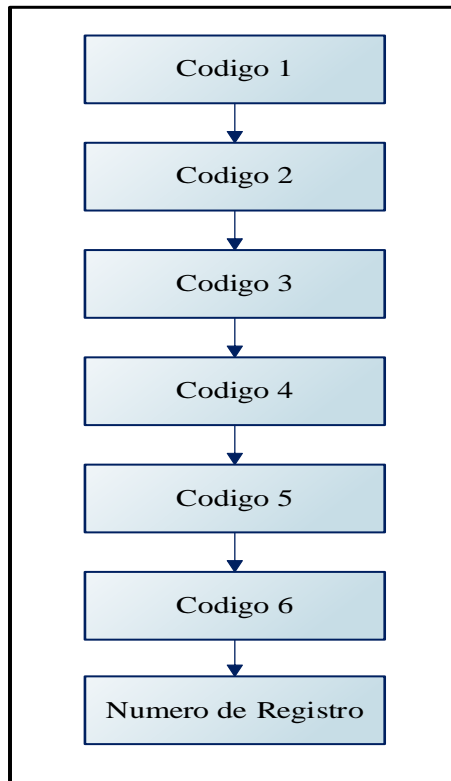


Figura 4.8. Jerarquía de atributos de la dimensión HIS

Tabla 29

Detalle dimensión HIS

NOMBRE DEL ATRIBUTO	DESCRIPCIÓN DEL ATRIBUTO
His key	Representa la llave principal de la dimensión
His código 1	Representa el código de atención.
His código 2	Representa el código de atención.
His código 3	Representa el código de atención.
His código 4	Representa el código de atención.
His código 5	Representa el código de atención.
His código 6	Representa el código de atención.
Número de registro	Representa al número de atención que fue registrado en el sistema.

Fuente: Elaboración propia

4.1.5.3. DIAGRAMA DE LA TABLA DE HECHO

Para esta etapa de la metodología se realizará un diagrama lógico para la tabla hecho ya completada, establecer su objetivo y mostrar todas las dimensiones conectadas a la tabla de hechos.

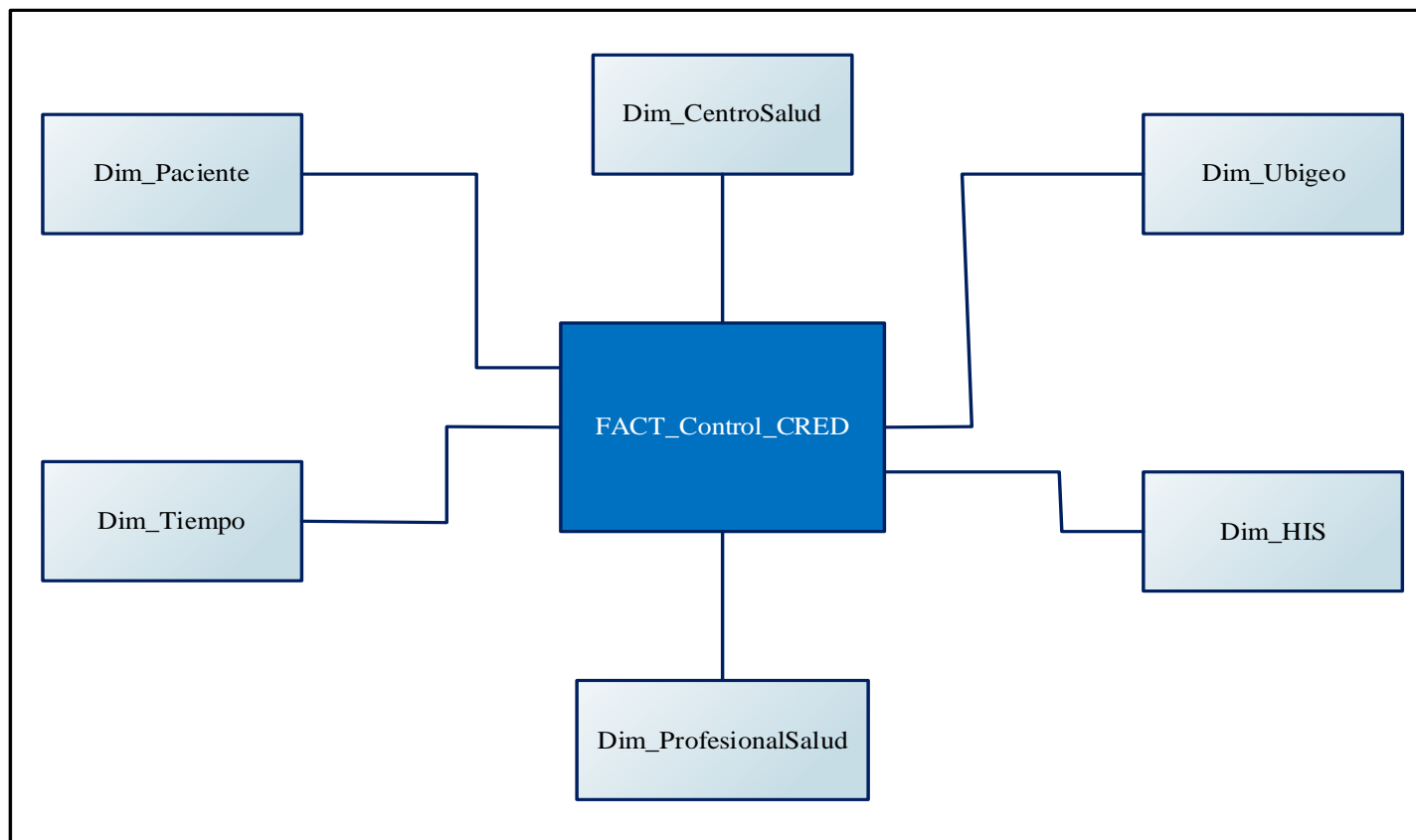


Figura 4.9: Diagrama de la tabla de hecho control CRED

DETALLE DE LA TABLA DE HECHO

Tabla 30.

Detalle de las llaves de las dimensiones

NOMBRE DE LA TABLA	NOMBRE DE LA COLUMNA	DESCRIPCIÓN DE LA COLUMNA
Dimensión Paciente	Paciente_Key	Llave primaria de la dimensión paciente, está convirtiéndose en llave foránea en la tabla hecho.
Dimensión Centro salud	Centro_Salud_Key	Llave primaria de la dimensión centro salud, está convirtiéndose en llave foránea en la tabla hecho.
Dimensión profesional salud	Profesional_Salud_Key	Llave primaria de la dimensión profesional salud, está convirtiéndose en llave foránea en la tabla hecho.
Dimensión Tiempo	Tiempo_Key	Llave primaria de la dimensión tiempo, está convirtiéndose en llave foránea en la tabla hecho.
Dimensión His	His_Key	Llave primaria de la dimensión HIS, está convirtiéndose en llave foránea en la tabla hecho.
Dimensión Ubigeo	Ubigeo_Key	Llave primaria de la dimensión Ubigeo, está convirtiéndose en llave foránea en la tabla hecho.

Fuente: Elaboración propia

Tabla 31*Detalle de las medidas de la tabla de hecho fact_Control_CRED*

NOMBRE DE LA TABLA	NOMBRE DE LA COLUMNA	DESCRIPCIÓN DE LA COLUMNA
Fact Control CRED	Cred total atenciones	Total de atenciones de control cred por parte de la red de Salud de Huamanga
	Cred estado nutricional	Control cred en el estado nutricional.
	Cred ganancia inadecuada de peso	Es la subida de peso del niño de manera inadecuada acorde a su edad.
	Cred numero de recién nacidos	Control cred en recién nacidos
	Cred niños 1 año	Control cred en niños de 1 año
	Cred niños 2 años	Control cred en niños de 2 años
	Cred niños 3 años	Control cred en niños de 3 años
	Cred niños 4 años	Control cred en niños de 4 años
	Cred niños 5 años	Control cred en niños de 5 años
	Cred RN con peso extremadamente bajo	Es el peso inadecuado del niño, siendo muy bajo acorde a su edad.
	Cred RN con peso muy bajo	Es el peso inadecuado a su desarrollo del recién nacido
	Cred RN con peso bajo al nacer	Es cuando el niño nace con un peso bajo acorde a los pesos estimados.
	Cred desnutrición aguda	Es cuando el peso y talla es inadecuado acorde a su edad de desarrollo
	Cred desnutrición severa	El paciente se encuentra desnutrido y de manera severa, impidiendo este su desarrollo
	Cred Sobre peso	Mide el peso exagerado acorde a su edad del paciente.
Cred Tamizaje de anemia	Mide si el paciente se encuentra con anemia o no.	

Fuente: Elaboración propia

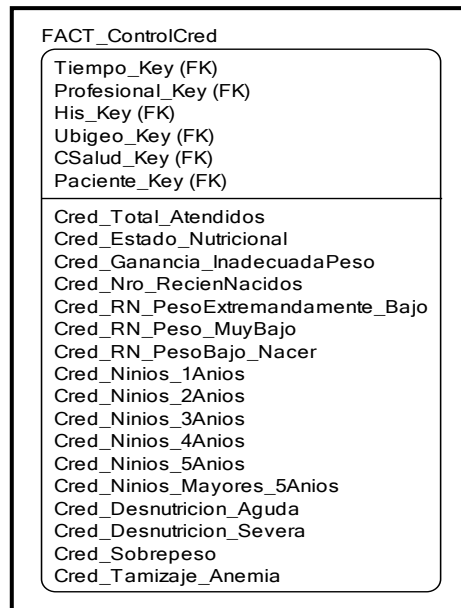


Figura 4.10. Tabla de hecho del Data Mart

4.1.5.4. DIAGRAMA DEL MODELO LÓGICO

Ya teniendo las tablas de dimensiones y la tabla hechos asociadas con cada una de ellas mediante las llaves principales, o bien llamadas llaves primarias, estas ya más orientadas al Data Mart que se quiere realizar, para una mejor visualización se presenta en un modelo lógico de esquema estrella.

Para un mejor entendimiento se realiza un listado de todos los componentes que intervienen en el esquema estrella.

MODELO DEL ESQUEMA ESTRELLA

- Componente: TABLA DE DIMENSIÓN TIEMPO
- Componente: TABLA DE DIMENSIÓN PROFESIONAL SALUD
- Componente: TABLA DE DIMENSIÓN CENTRO DE SALUD
- Componente: TABLA DE DIMENSIÓN PACIENTE
- Componente: TABLA DE DIMENSIÓN HIS
- Componente: TABLA DE DIMENSIÓN UBIGEO
- Componente: TABLA DE HECHO FACT CONTROL CRED

LLAVES PRIMARIAS

- Tiempo_Key
- CSalud_Key
- Paciente_Key
- His_Key
- Ubigeo_Key
- Profesional_Key

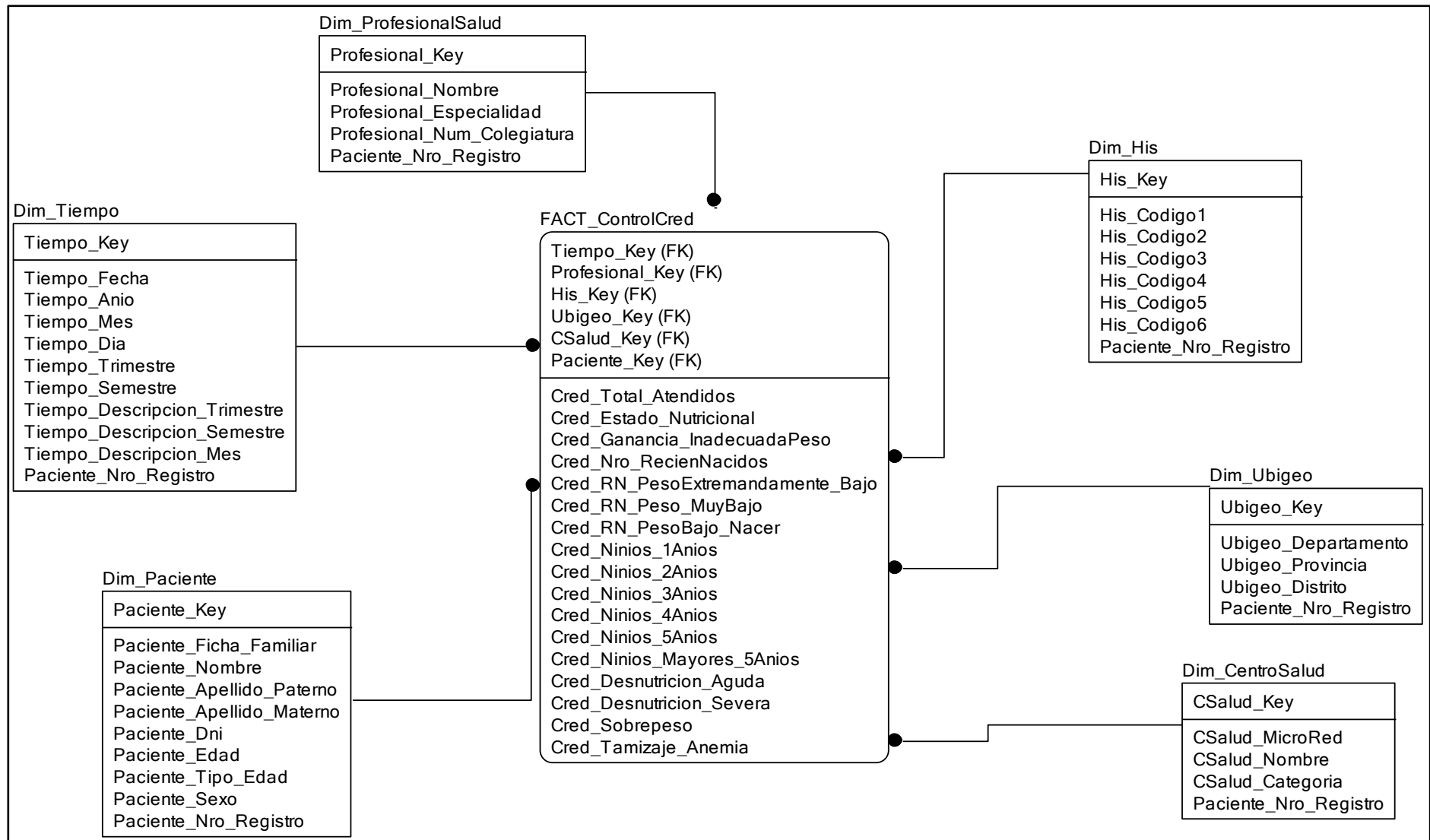


Figura 4.11: Modelo lógico en esquema estrella del Data Mart

4.1.6. DISEÑO DEL MODELO FÍSICO

Ya teniendo el modelo lógico en el esquema estrella, con todas las dimensiones relacionadas hacia a la tabla de hechos, se procede a realizar el modelo físico, que consiste en poner el tamaño de cada una de las variables de cada dimensión, que estas se pondrán ya en el motor de bases de datos que es el SQL Server 2012.

i. Tabla: Dim_CentroSalud

Tabla 32.

Tipo de datos para la Dim_CentroSalud

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
CSalud_Key	Int	
CSalud_MicroRed	Nvarchar	20
CSalud_Nombre	Nvarchar	70
CSsalud_Ctegoria	Nvarchar	6
Paciente_Nro_Registro	Nvarchar	8

Fuente: Elaboración propia

ii. Tabla: Dim_ProfesionalSalud

Tabla 33.

Tipo de datos para la Dim_ProfesionalSalud

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
Profesional_Key	Int	
Profesional_Nombre	Nvarchar	100
Profesional_Especialidad	Nvarchar	50
Profesional_Num_Colegiatura	Nvarchar	15
Paciente_Nro_Registro	Nvarchar	8

Fuente: Elaboración propia

iii. Tabla: Dim_Ubigeo

Tabla 34.

Tipo de datos para la Dim_ProfesionalSalud

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
Ubigeo_Key	Int	
Ubigeo_Departamento	Nvarchar	30
Ubigeo_Provincia	Nvarchar	30
Ubigeo_Distrito	Nvarchar	45
Paciente_Nro_Registro	Nvarchar	8

Fuente: Elaboración propia

iv. Tabla: Dim_Tiempo

Tabla 35.

Tipo de datos para la Dim_Tiempo

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
Tiempo_Key	Int	
Tiempo_Fecha	Datetime	
Tiempo_Anio	Int	
Tiempo_Mes	Int	
Tiempo_Dia	Int	
Tiempo_Trimestre	Nvarchar	20
Tiempo_Semestre	Nvarchar	20
Tiempo_Descripcion_Trimestre	Nvarchar	25
Tiempo_Descripcion_Semestre	Nvarchar	25
Tiempo_Descripcion_Mes	Nvarchar	25
Paciente_Nro_Registro	Nvarchar	8

Fuente: Elaboración propia

v. **Tabla: Dim_Paciente**

Tabla 36

Tipo de datos para la Dim_Paciente

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
Paciente_Key		
Paciente_Ficha_Familiar	Nvarchar	10
Paciente_Nombre	Nvarchar	25
Paciente_Apellido_Paterno	Nvarchar	50
Paciente_Apellido_Materno	Nvarchar	25
Paciente_Dni	Nvarchar	20
Paciente_Edad	Nvarchar	3
Paciente_Tipo_Edad	Nvarchar	6
Paciente_Sexo	Nvarchar	2
Paciente_Nro_Registro	Nvarchar	8

Fuente: Elaboración propia

vi. **Tabla: Dim_HIS**

Tabla 37.

Tipo de datos para la Dim_HIS

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
His_Key		
His_Codigo1	Nvarchar	6
His_Codigo2	Nvarchar	6
His_Codigo3	Nvarchar	6
His_Codigo4	Nvarchar	6
His_Codigo5	Nvarchar	6
His_Codigo6	Nvarchar	6
Paciente_Nro_Registro	Nvarchar	8

Fuente: Elaboración propia

vii. **Tabla: FACT_ControlCred**

Tabla 38.

Tipo de datos para la FACT_ControlCred

NOMBRE DE CAMPO	TIPO DE VARIABLE	LONGITUD
Cred_Total_Atendidos	Int	
Cred_Estado_Nutricional	Int	
Cred_Ganancia_InadecuadaPeso	Int	
Cred_Nro_RecienNacidos	Int	
Cred_Ninios_1Anios	Int	
Cred_Ninios_2Anios	Int	
Cred_Ninios_3Anios	Int	
Cred_Ninios_4Anios	Int	
Cred_Ninios_5Anios	Int	
Cred_Ninios_Mayores_5Anios	Int	
Cred_Rn_PesoExtremadamente_bajo	Int	
Cred_Peso_MuyBajo	Int	
Cred_PesoBajo_Nacer	Int	
Cred_Desnutricion_Aguda	Int	
Cred_Desnutricion_Severa	Int	
Cred_Sobrepeso	Int	
Cred_Tamizaje_Anemia	Int	

Fuente: Elaboración propia

Por consiguiente, ya contando con el modelo lógico y la estructura del modelo físico, se obtiene el siguiente modelo de bases de datos físico para el Data Mart.

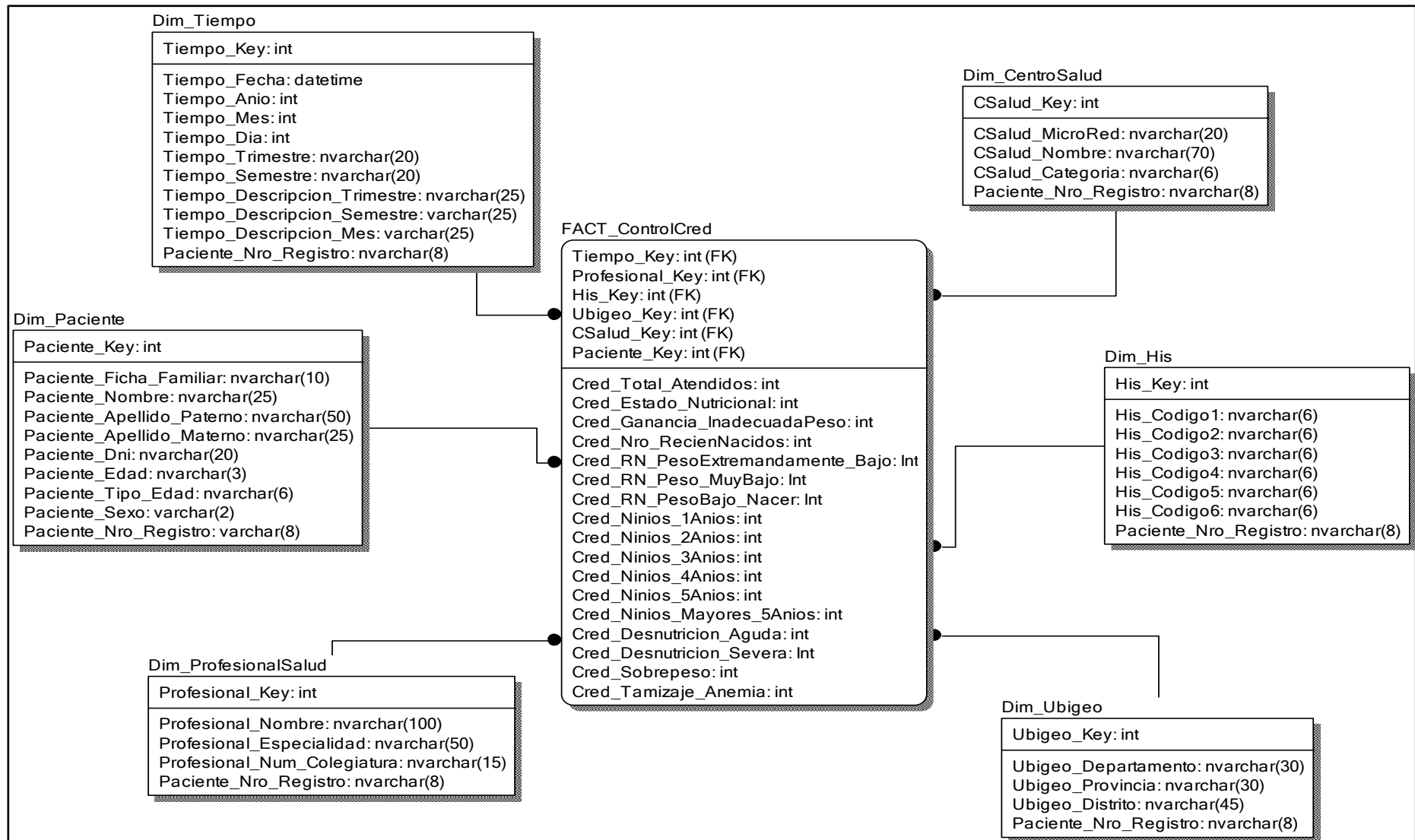
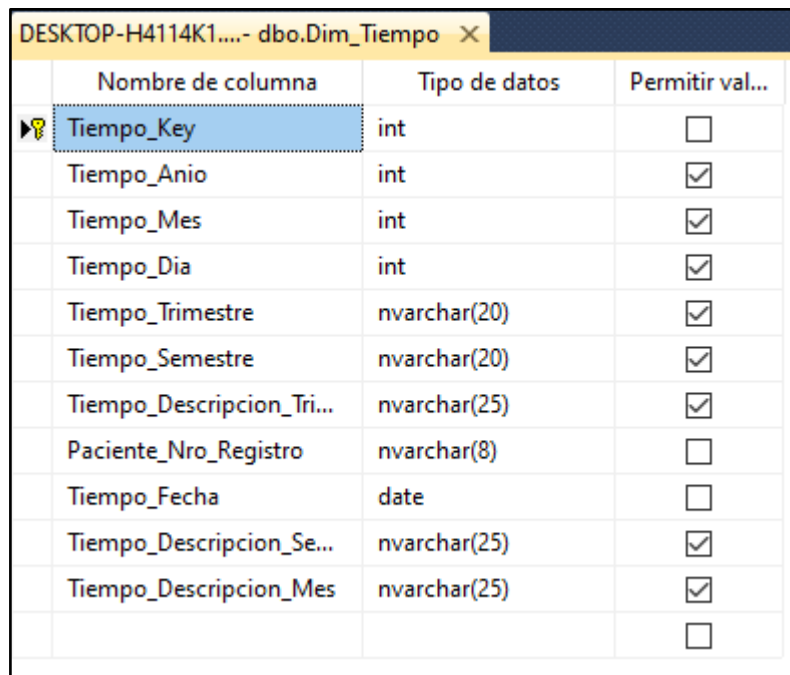


Figura 4.12. Diseño Físico de la base de datos del Data Mart

4.1.6.1. CONSTRUCCIÓN DE LAS TABLAS Y LA BASE DE DATOS EN SQL

Después de haber realizado el diseño físico de las tablas pasamos a la etapa de construcción de dichas tablas, y pasamos al motor de base de datos que es el SQL Server 2012.

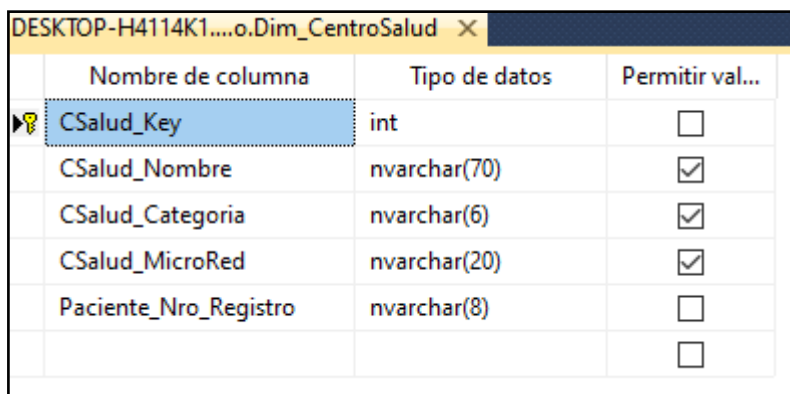
A. Tabla: Dim_Tiempo



Nombre de columna	Tipo de datos	Permitir val...
Tiempo_Key	int	<input type="checkbox"/>
Tiempo_Anio	int	<input checked="" type="checkbox"/>
Tiempo_Mes	int	<input checked="" type="checkbox"/>
Tiempo_Dia	int	<input checked="" type="checkbox"/>
Tiempo_Trimestre	nvarchar(20)	<input checked="" type="checkbox"/>
Tiempo_Semestre	nvarchar(20)	<input checked="" type="checkbox"/>
Tiempo_Descripcion_Tri...	nvarchar(25)	<input checked="" type="checkbox"/>
Paciente_Nro_Registro	nvarchar(8)	<input type="checkbox"/>
Tiempo_Fecha	date	<input type="checkbox"/>
Tiempo_Descripcion_Se...	nvarchar(25)	<input checked="" type="checkbox"/>
Tiempo_Descripcion_Mes	nvarchar(25)	<input checked="" type="checkbox"/>
		<input type="checkbox"/>

Figura 4.13. Tabla Dim_Tiempo

B. Dim_CentroSalud



Nombre de columna	Tipo de datos	Permitir val...
CSalud_Key	int	<input type="checkbox"/>
CSalud_Nombre	nvarchar(70)	<input checked="" type="checkbox"/>
CSalud_Categoria	nvarchar(6)	<input checked="" type="checkbox"/>
CSalud_MicroRed	nvarchar(20)	<input checked="" type="checkbox"/>
Paciente_Nro_Registro	nvarchar(8)	<input type="checkbox"/>
		<input type="checkbox"/>

Figura 4.14: Tabla Dim_CentroSalud

C. Dim_ProfesionalSalud

Nombre de columna	Tipo de datos	Permitir val...
Profesional_Key	int	<input type="checkbox"/>
Profesional_Nombre	nvarchar(100)	<input checked="" type="checkbox"/>
Profesional_Especialidad	nvarchar(50)	<input checked="" type="checkbox"/>
Profesional_Num_Colegi...	nvarchar(15)	<input checked="" type="checkbox"/>
Paciente_Nro_Registro	nvarchar(8)	<input type="checkbox"/>
		<input type="checkbox"/>

Figura 4.15. Tabla Dim_ProfesionalSalud

D. Dim_Paciente

Nombre de columna	Tipo de datos	Permitir val...
Paciente_Key	int	<input type="checkbox"/>
Paciente_Nombre	nvarchar(25)	<input checked="" type="checkbox"/>
Paciente_Apellido_Paterno	nvarchar(50)	<input checked="" type="checkbox"/>
Paciente_Apellido_Mater...	nvarchar(25)	<input checked="" type="checkbox"/>
Paciente_Dni	nvarchar(20)	<input type="checkbox"/>
Paciente_Edad	nvarchar(3)	<input checked="" type="checkbox"/>
Paciente_Ficha_Familiar	nvarchar(10)	<input type="checkbox"/>
Paciente_Nro_Registro	nvarchar(8)	<input type="checkbox"/>
Paciente_Tipo_Edad	nvarchar(6)	<input checked="" type="checkbox"/>
Paciente_Sexo	nvarchar(2)	<input checked="" type="checkbox"/>
		<input type="checkbox"/>

Figura 4.16. Tabla Dim_Paciente

E. Dim_Ubigeo

Nombre de columna	Tipo de datos	Permitir val...
Ubigeo_Key	int	<input type="checkbox"/>
Ubigeo_Departamento	nvarchar(30)	<input checked="" type="checkbox"/>
Ubigeo_Provincia	nvarchar(30)	<input checked="" type="checkbox"/>
Paciente_Nro_Registro	nvarchar(8)	<input type="checkbox"/>
Ubigeo_Distrito	nvarchar(45)	<input checked="" type="checkbox"/>
		<input type="checkbox"/>

Figura 4.17. Tabla Dim_Ubigeo

F. Dim_His

Nombre de columna	Tipo de datos	Permitir val...
His_Key	int	<input type="checkbox"/>
His_Codigo1	nvarchar(6)	<input checked="" type="checkbox"/>
His_Codigo2	nvarchar(6)	<input checked="" type="checkbox"/>
His_Codigo3	nvarchar(6)	<input checked="" type="checkbox"/>
His_Codigo4	nvarchar(6)	<input checked="" type="checkbox"/>
His_Codigo5	nvarchar(6)	<input checked="" type="checkbox"/>
His_Codigo6	nvarchar(6)	<input checked="" type="checkbox"/>
Paciente_Nro_Registro	nvarchar(8)	<input type="checkbox"/>
		<input type="checkbox"/>

Figura 4.18. Tabla Dim_His

G. Fact_ControlCred

Nombre de columna	Tipo de datos	Permitir val...
Cred_Ninios_2Anios	int	<input checked="" type="checkbox"/>
Tiempo_Key	int	<input type="checkbox"/>
Profesional_Key	int	<input type="checkbox"/>
His_Key	int	<input type="checkbox"/>
Ubigeo_Key	int	<input type="checkbox"/>
CSalud_Key	int	<input type="checkbox"/>
Paciente_Key	int	<input type="checkbox"/>
Cred_Total_CentroSalud	int	<input checked="" type="checkbox"/>
Cred_Estado_Nutricional	int	<input checked="" type="checkbox"/>
Cred_Total_MicroRed	int	<input checked="" type="checkbox"/>
Cred_Nro_RecienNacidos	int	<input checked="" type="checkbox"/>
Cred_Ninios_1Anios	int	<input checked="" type="checkbox"/>
Cred_Ninios_3Anios	int	<input checked="" type="checkbox"/>
Cred_Ninios_4Anios	int	<input checked="" type="checkbox"/>
Cred_Ninios_5Anios	int	<input checked="" type="checkbox"/>
Cred_Ninios_Mayores_5...	int	<input checked="" type="checkbox"/>
Cred_TotalAtendios	int	<input checked="" type="checkbox"/>
		<input type="checkbox"/>

Figura 4.19. Tabla Fact_ControlCred

Una vez construido todas las tablas para el Data Mart procedemos con la construcción total del diagrama de base de datos, en SQL Server 2012.

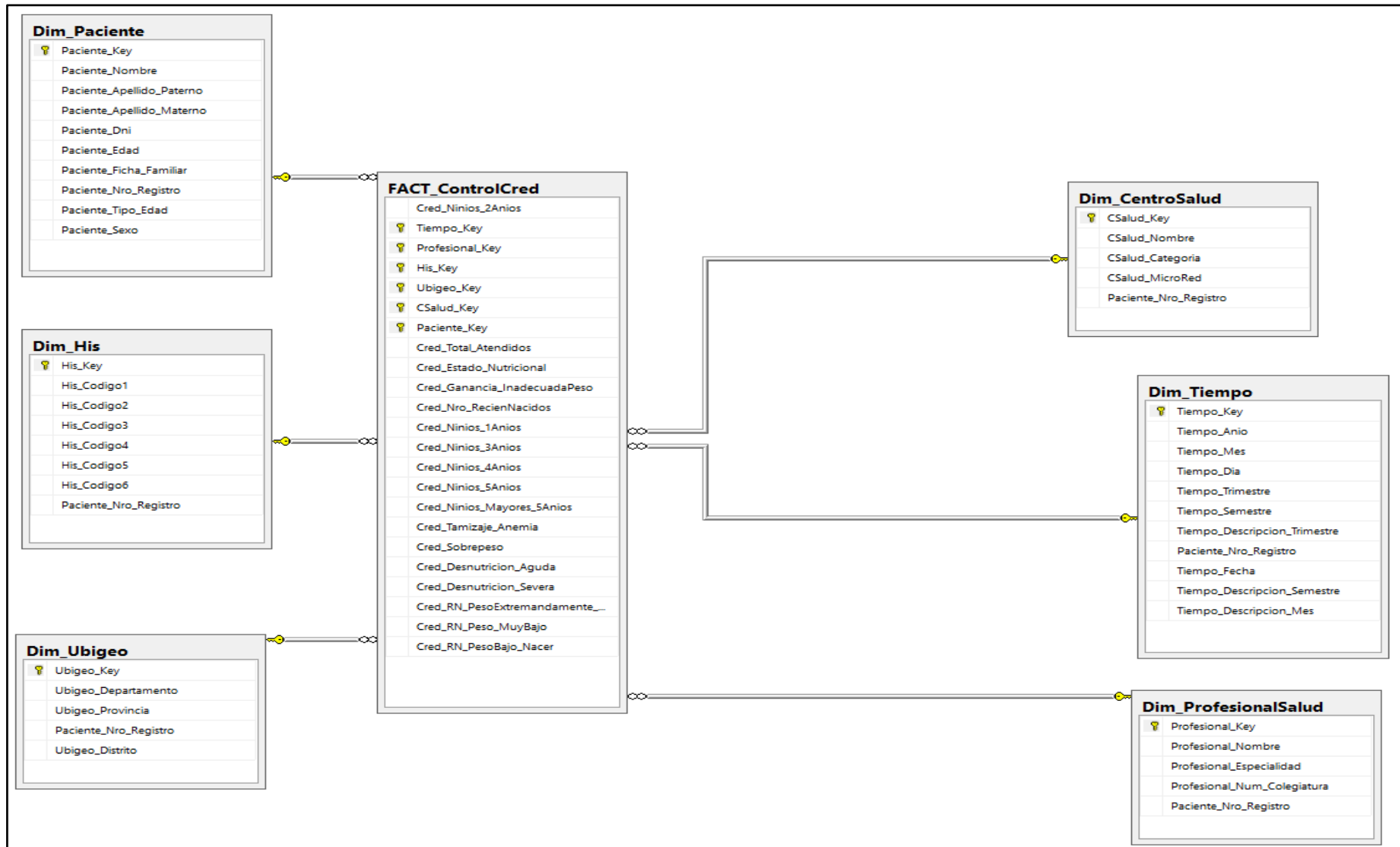


Figura 4.20. Modelo físico del Data Mart en el motor de base de datos SQL Server

4.1.7. DISEÑO E IMPLEMENTACIÓN DEL SUB SISTEMA DE ETL

4.1.7.1. FASES DE PROCESO ETL

La metodología Ralph Kimball, nos dice que para completar la construcción del Data Mart, se tiene que poblar la base de datos, como ya se conoce y se tiene el modelo físico ya montado en el motor de bases de datos que es el SQL Server, se tiene que seguir las fases del proceso ETL, para ello se realiza con la herramienta de Integración Services.

Los pasos de la fase del proceso ETL, detallamos a continuación:

Paso 1: Limpiar tablas. Consiste en borrar todos los registros de las tablas, relaciones, dimensiones y hecho de la base de datos del Data Mart, el cual nos permite asegurar de que nos pueda existir algunos datos.

Paso 2: Poblamiento de dimensión His. Se ejecuta el proceso de cargar de la dimensión con una serie de sentencia de Sql, que se detallara más adelante.

Paso 3: Poblamiento de dimensión paciente. Se ejecuta el proceso de carga de la dimensión paciente con una serie de sentencia SQL, se detallará más adelante.

Paso 4: Poblamiento de dimensión profesional salud. Se ejecuta el proceso de carga de la dimensión profesional salud con una serie de sentencia SQL, se detallará más adelante.

Paso 5: Poblamiento de dimensión centro salud. Se ejecuta el proceso de carga de la dimensión centro salud con una serie de sentencia SQL, se detallará más adelante.

Paso 6: Poblamiento de dimensión ubigeo. Se ejecuta el proceso de carga de la dimensión centro salud con una serie de sentencia SQL, se detallará más adelante.

Paso 7: Poblamiento de dimensión tiempo. Se ejecuta el proceso de carga de la dimensión centro salud con una serie de sentencia SQL, se detallará más adelante.

Paso 8: Completar el poblamiento de dimensión tiempo. Se ejecuta proceso de carga de la dimensión tiempo, para lograr poblar por completo la dimensión.

Paso 9: poblamiento de la tabla hecho.

- a. **Cred_Total_CentroSalud:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Total_CentroSalud.
- b. **Cred_Estado_Nutricional:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Estado_Nutricional.
- c. **Cred_Total_Atendidos_MicroRed:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Total_Atendidos_MicroRed.
- d. **Cred_Nro_RecienNacidos:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Nro_RecienNacidos.
- e. **Cred_Ninios_1Anios:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Ninios_1Anios.
- f. **Cred_Ninios_2Anios:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Ninios_1Anios.
- g. **Cred_Ninios_3Anios:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Ninios_1Anios.
- h. **Cred_Ninios_4Anios:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Ninios_1Anios.
- i. **Cred_Ninios_5Anios:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Ninios_1Anios.
- j. **Cred_Ninios_Mayores_5Anios:** Se ejecuta el proceso de la carga de la tabla de hecho Cred_Ninios_Mayores_5Anios.

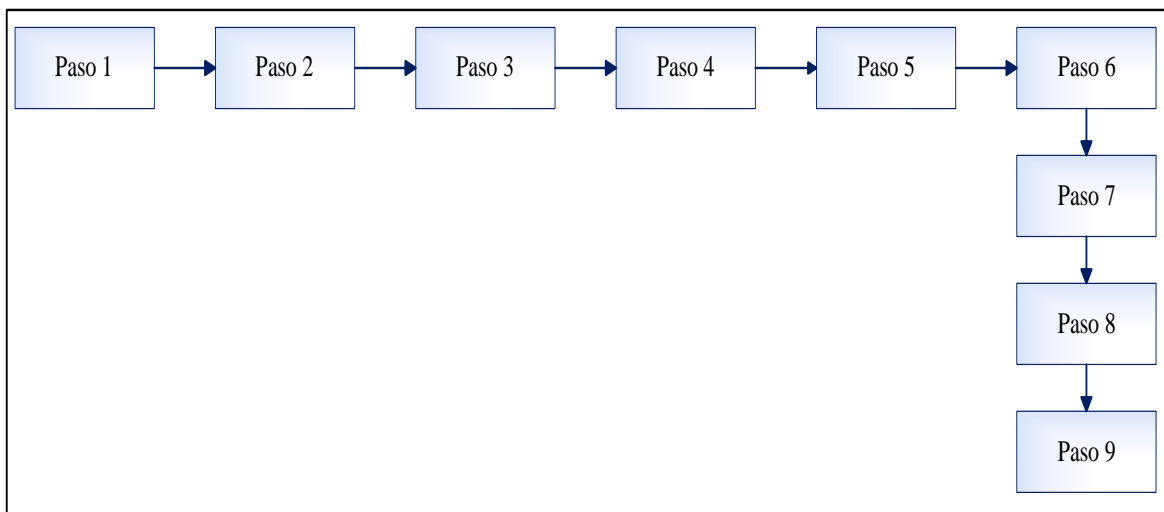


Figura 4.21. diagrama de pasos del proceso ETL.

Restricciones de precedencia

- a. La limpieza de las dimensiones (Paso 1) tiene que realizarse al inicio del proceso.
- b. El poblamiento de la dimensión His (Paso 2) tiene que realizarse solo cuando se tenga seguridad de que el paso 1 se ha completado adecuadamente y con éxito.
- c. El poblamiento de la dimensión paciente (Paso 3) tiene que realizarse cuando se tenga seguridad de que el paso 1 se ha completado adecuadamente y con éxito.
- d. El poblamiento de la dimensión profesional salud (Paso 4) tiene que realizarse cuando se tenga seguridad de que el paso 1 se ha completado adecuadamente y con éxito.
- e. El poblamiento de la dimensión Centro Salud (Paso 5) tiene que realizarse cuando se tenga seguridad de que el paso 1 se ha completado adecuadamente y con éxito.
- f. El poblamiento de la dimensión Ubigeo (Paso 6) tiene que realizarse cuando se tenga seguridad de que el paso 1 se ha completado adecuadamente y con éxito.
- g. El poblamiento de la dimensión Tiempo (Paso 7) tiene que realizarse cuando se tenga seguridad de que el paso 1 se ha completado adecuadamente y con éxito.
- h. El paso de completar carga de tiempo (Paso 8) se tiene que realizarse cuando se tenga seguridad de que el paso 7 se haya completado adecuadamente y con éxito.
- i. El poblamiento de la tabla de echo control CRED (paso 9) se tiene que realizar cuando los pasos 2,3,4,5,6,7,8 hayan culminado con éxito.

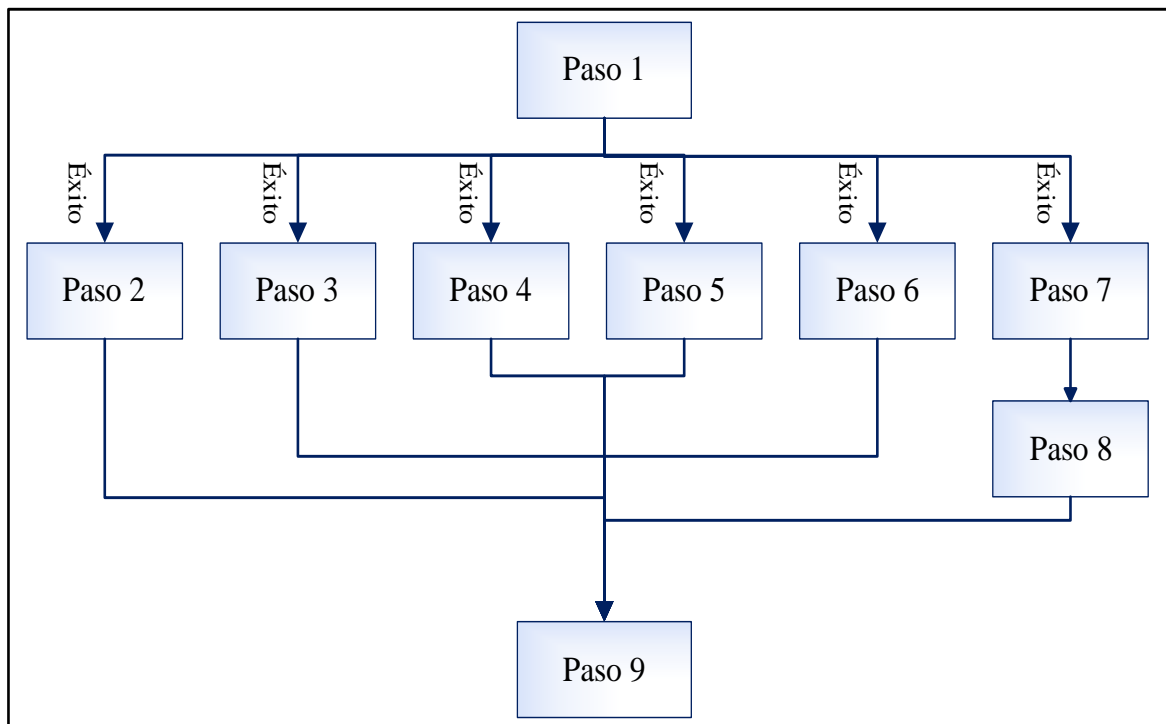


Figura 4.22. Diagrama de pasos con restricciones de precedencia

4.1.7.2. SENTENCIAS DE CADA PASO DEL PROCESO ETL

CARGA INICIAL

Para la carga del modelo se procede a llenar cada uno de las dimensiones como ya mencionado anterior mente siguiendo los pasos y restricciones de precedencia.

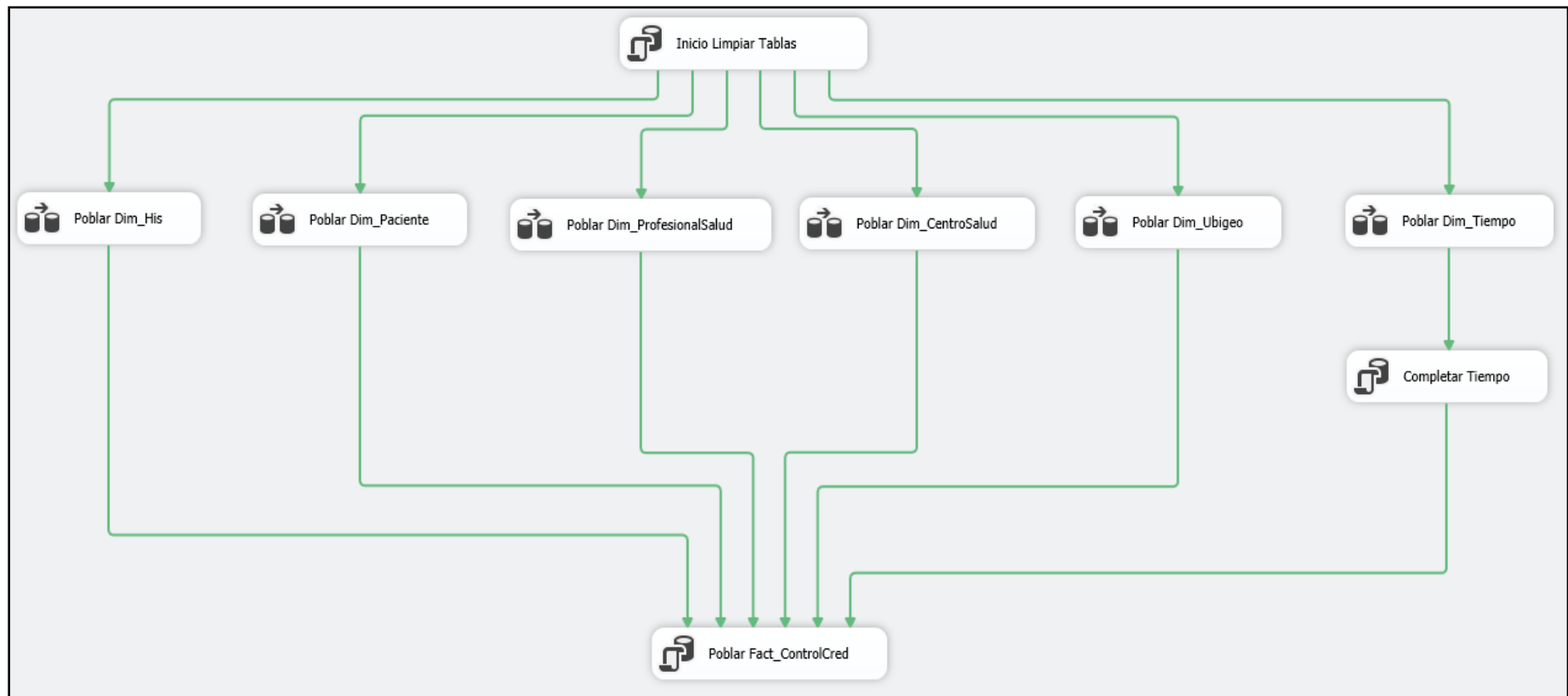


Figura 4.23. Proceso ETL para la carga inicial

1. LIMPIADO DE DIMENSIONES

```
DELETE FROM Dim_Tiempo;  
DELETE FROM Dim_Centrosalud;  
DELETE FROM Dim_His;  
DELETE FROM Dim_Paciente;  
DELETE FROM Dim_Profesionalsalud;  
DELETE FROM Dim_Ubigeo;  
TRUNCATE TABLE Fact_Controlcred;
```

Figura 4.24. Sentencia del proceso ETL para la limpieza de toda la base de datos de BD_DataMart_Cred

2. POBLAMIENTO DE LA DIMENSIÓN HIS

A. ORIGEN DE DATOS

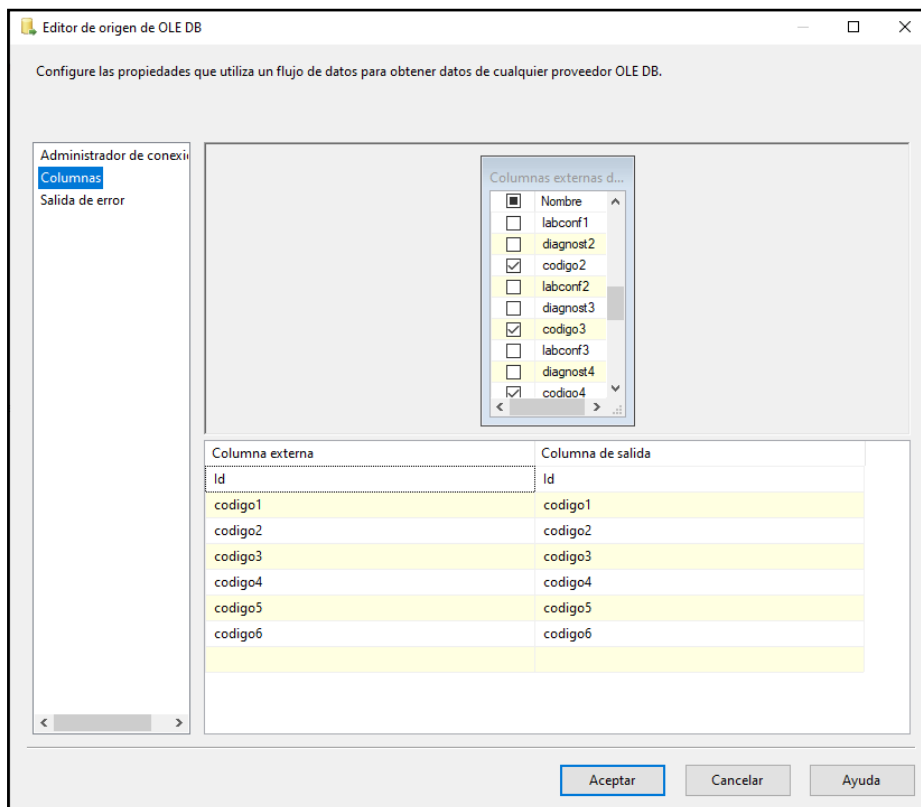


Figura 4.25. Origen de datos para la dimensión His

B. CONVERSIÓN DE DATOS

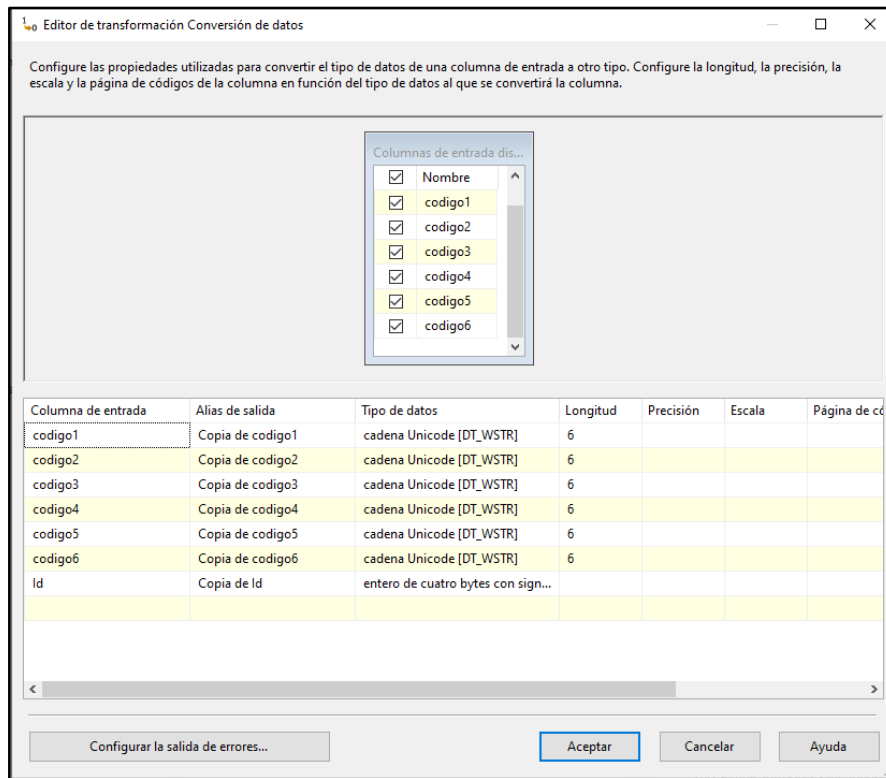


Figura 4.26. Transformación de datos para la dimensión His

C. POBLAR DIMENSIÓN HIS

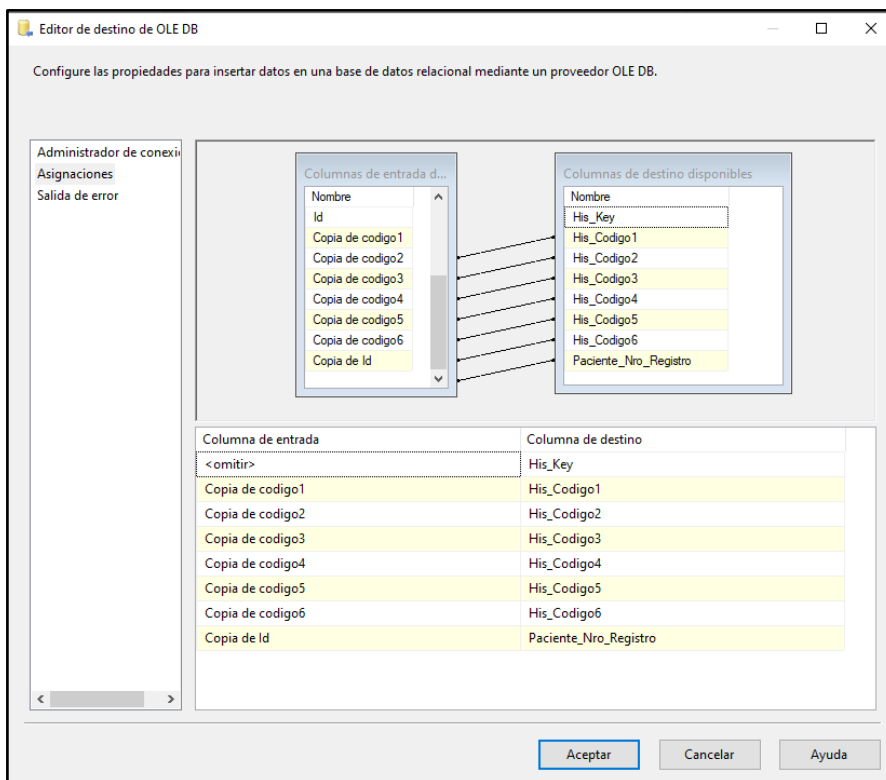


Figura 4.27. Carga de datos a la tabla Dim_His

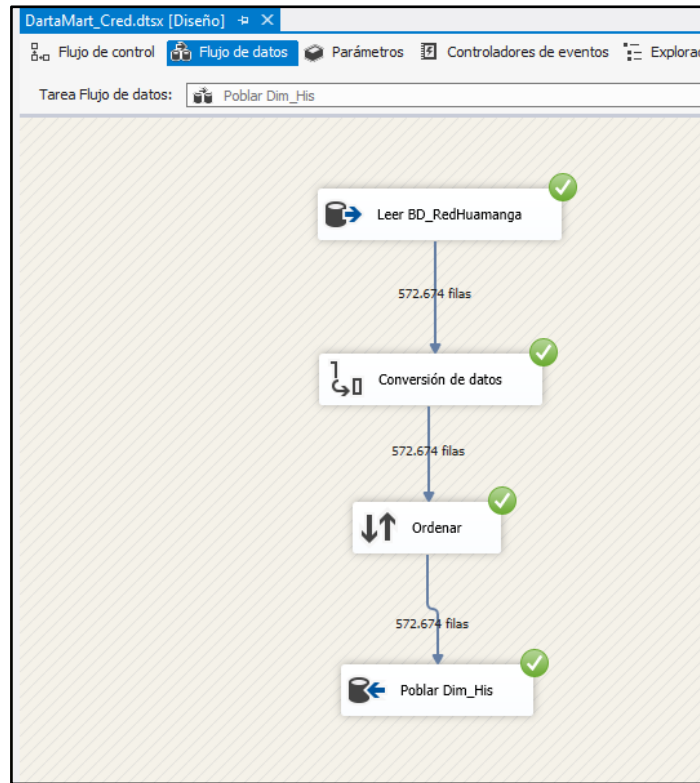


Figura 4.28. Proceso ETL culminada para la dimensión His

3. POBLAMIENTO DE LA DIMENSIÓN PACIENTE

A. ORÍGENES DE DATOS

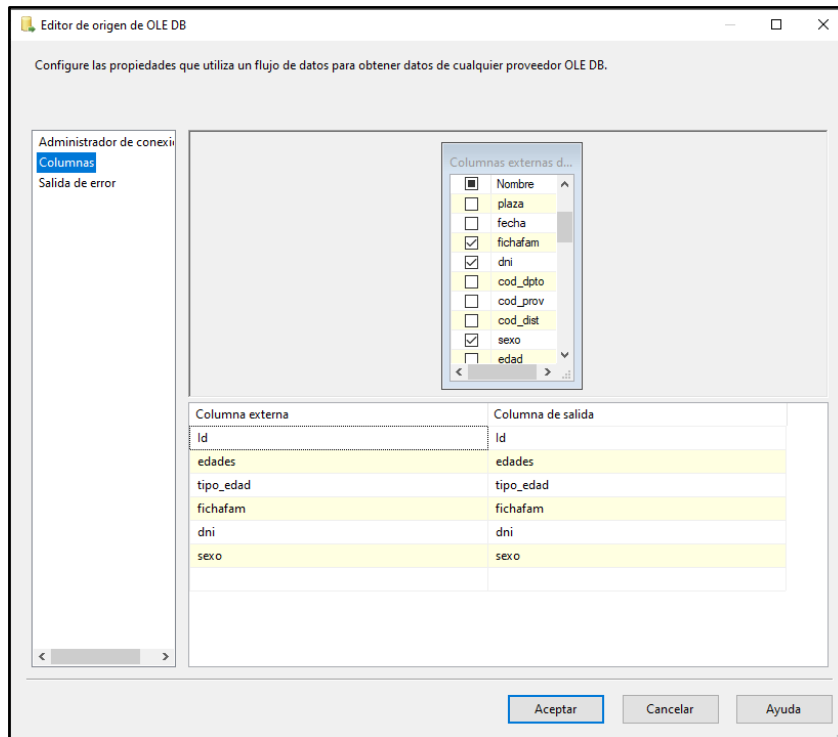


Figura 4.29. Orígenes de datos de la dimensión Paciente

B. CONVERSIÓN DE DATOS

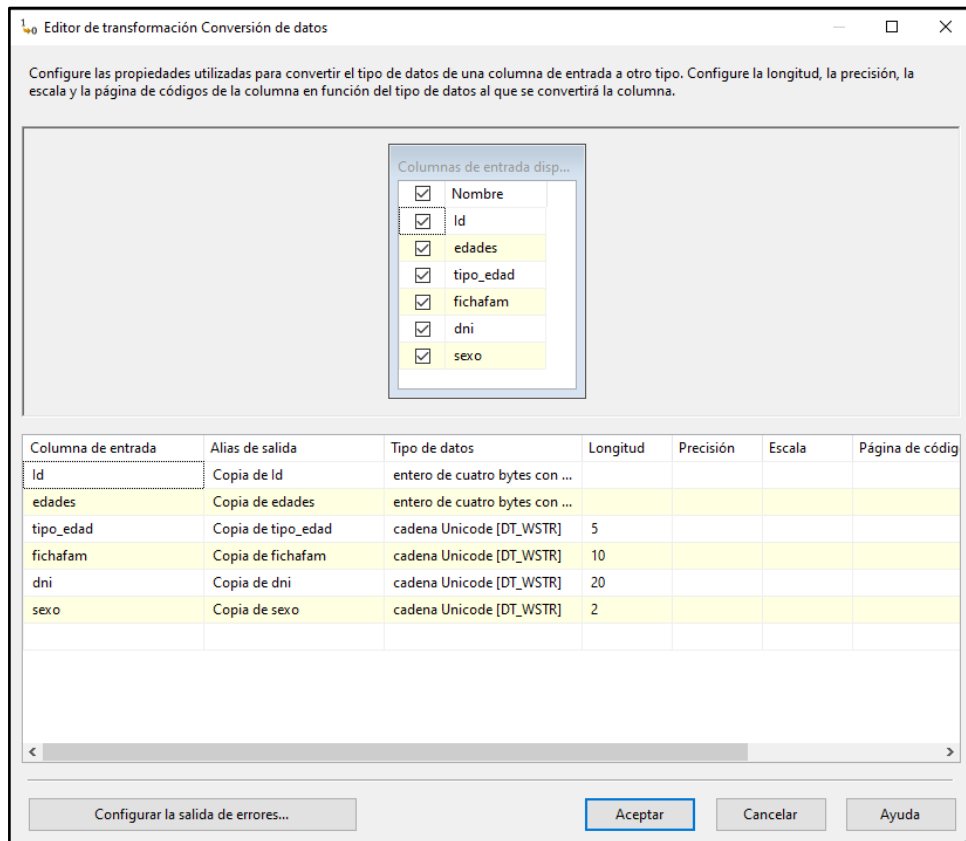


Figura 4.30. Transformación de datos para la dimensión paciente

C. POBLAR DIMENSIÓN PACIENTE

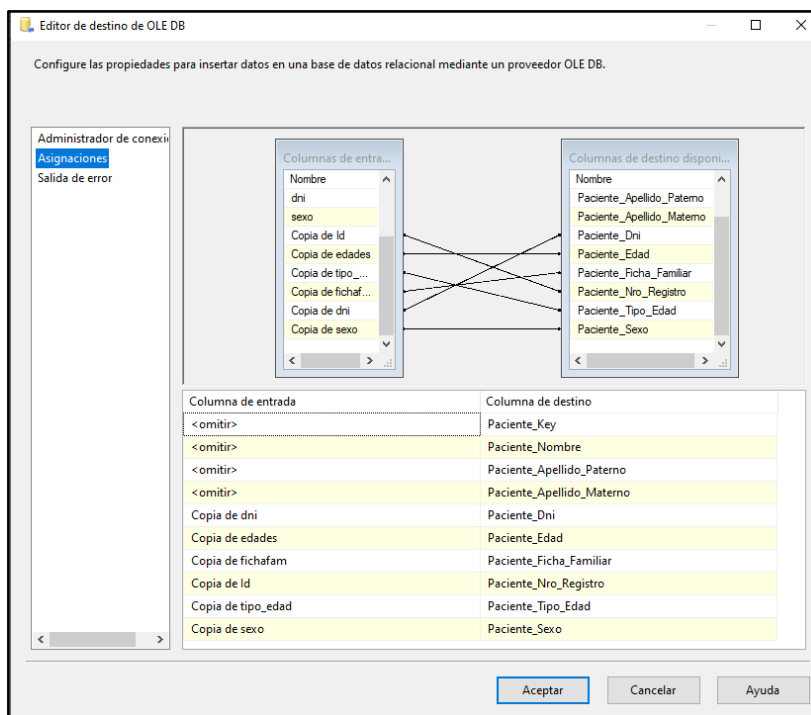


Figura 4.31. Carga de datos a la tabla Dim_Paciente

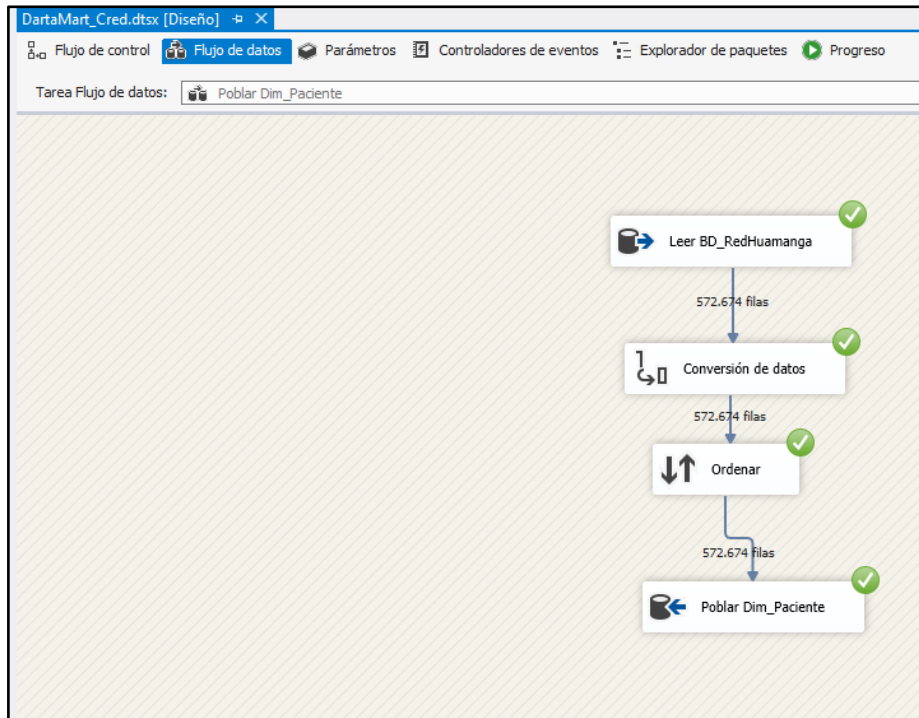


Figura 4.32. Proceso ETL culminado para la dimensión paciente

4. POBLAR DIMENSIÓN PROFESIONAL SALUD

A. ORIGEN DE DATOS

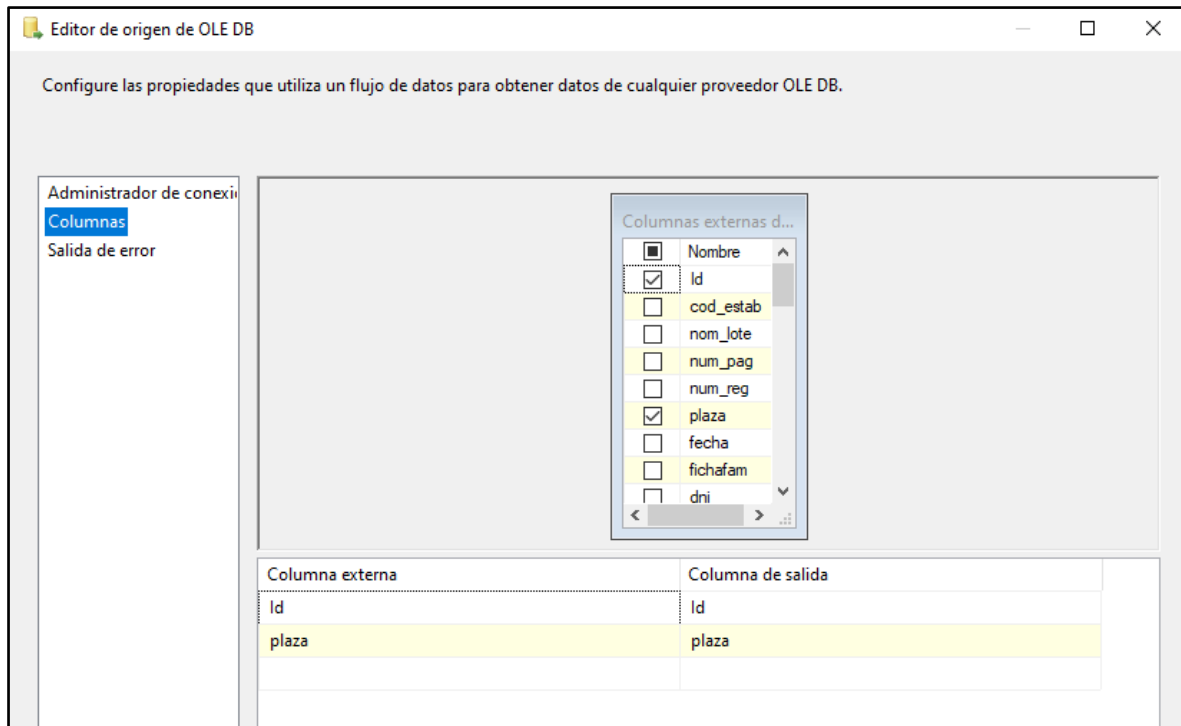


Figura 4.33. Origen de datos para la dimensión profesional salud de la BD_His2018

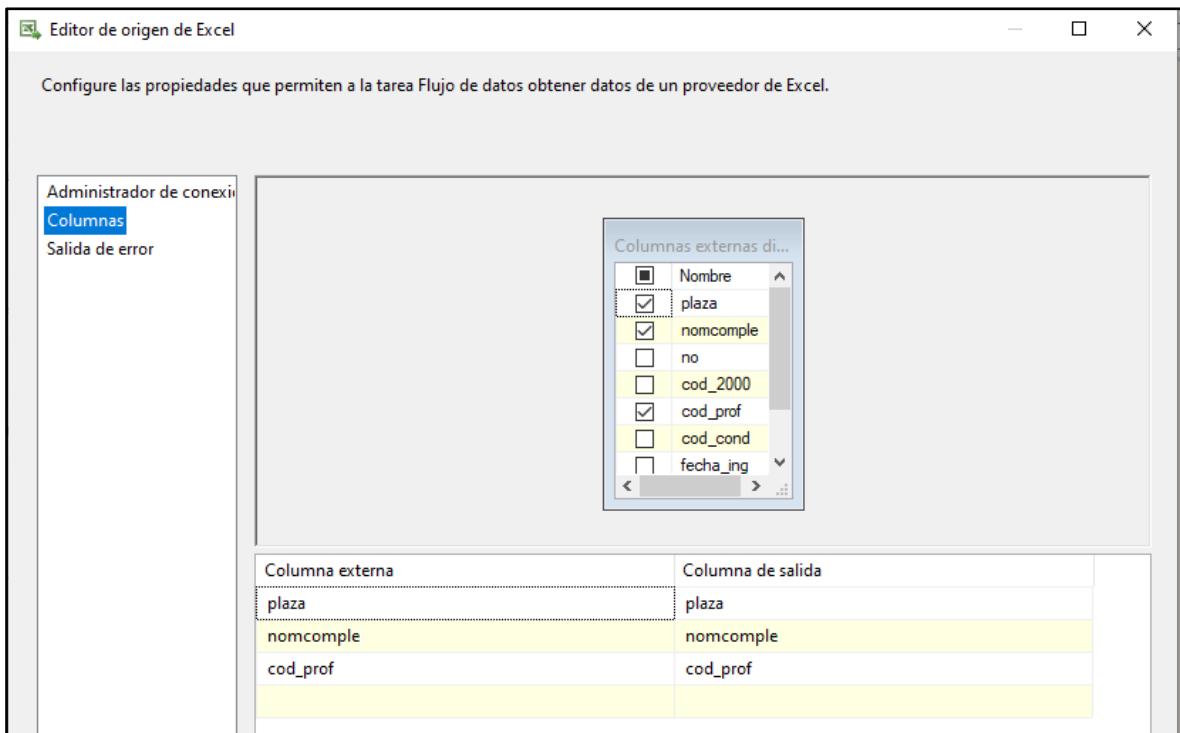


Figura 4.34. Origen de datos para la dimensión profesional salud de archivo Excel

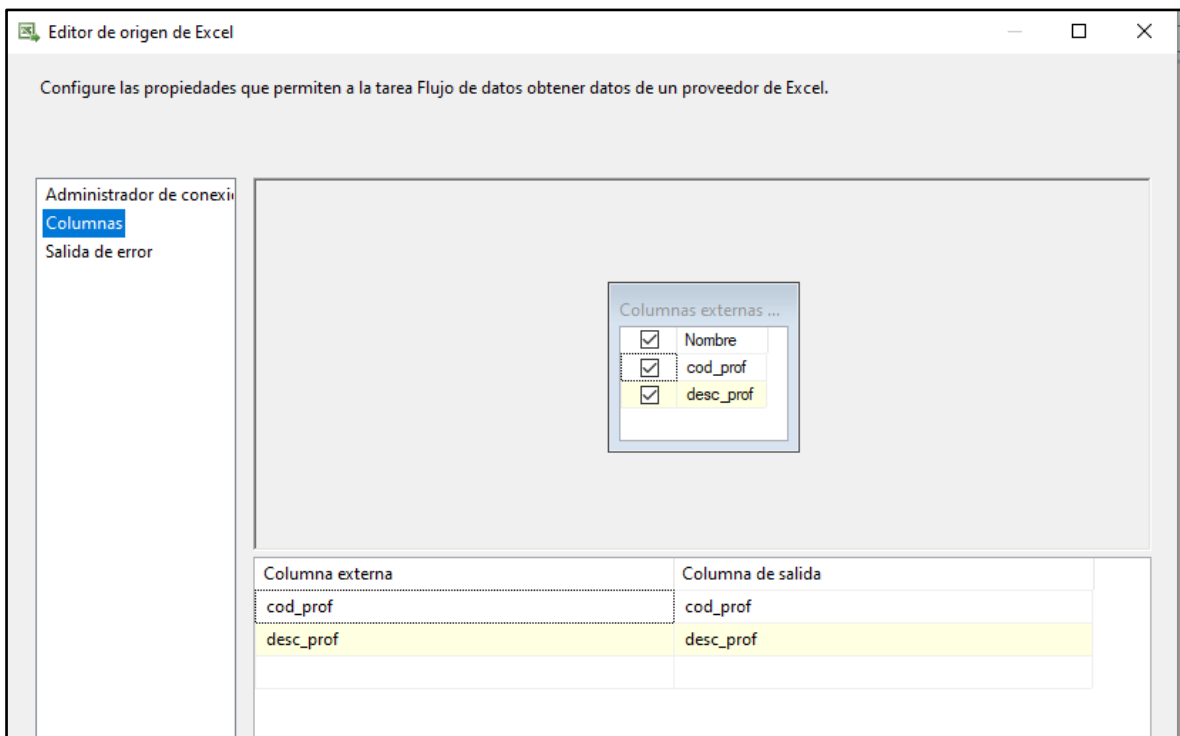


Figura 4.35. Origen de datos para la dimensión profesional salud de archivo Excel

B. CONVERSIÓN DE DATOS

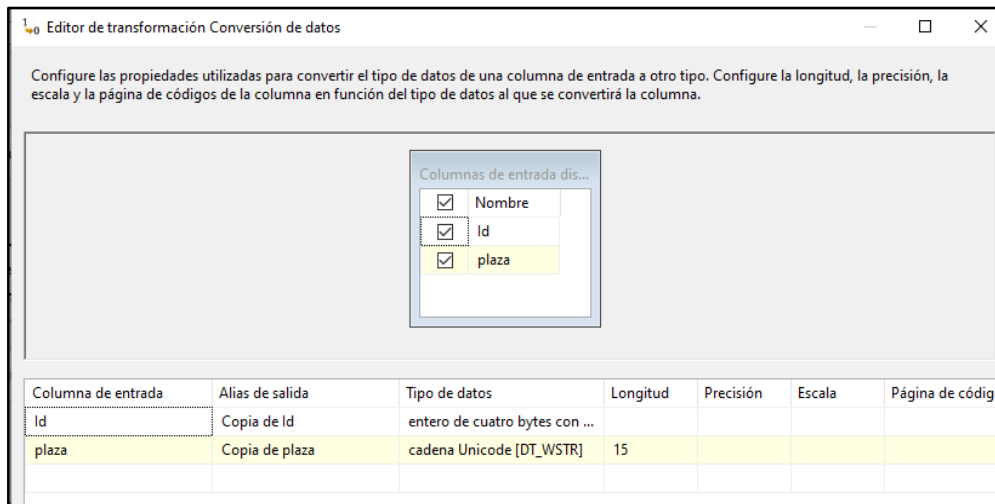


Figura 4.36. Transformación de datos del origen BD_His2018

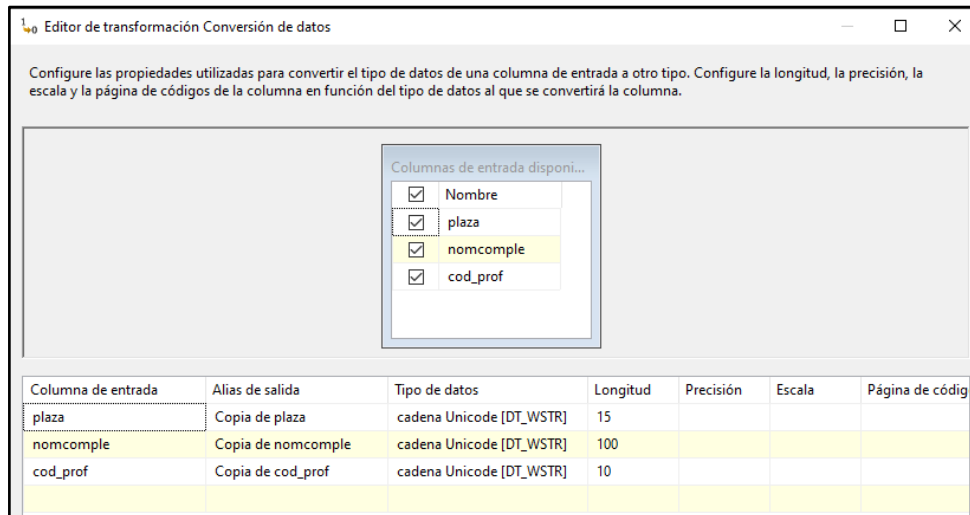


Figura 4.37. Transformación de datos del origen Excel

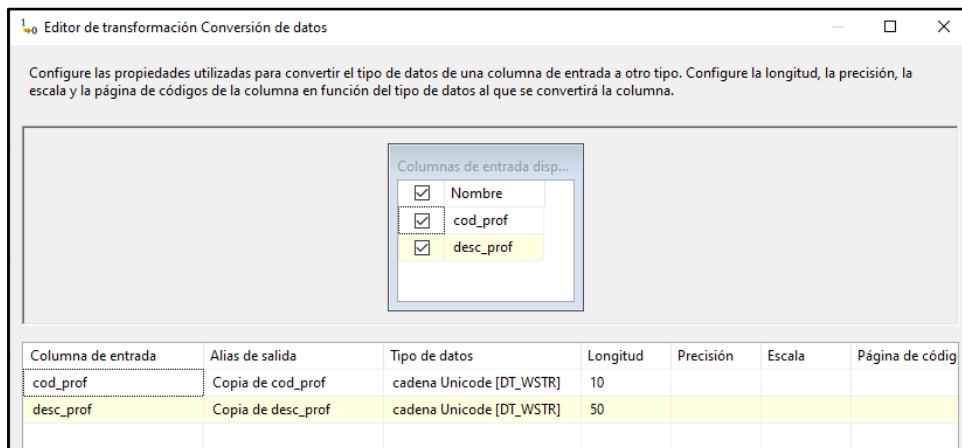


Figura 4.38. Transformación de datos origen Excel

C. MESCLA DE DATOS

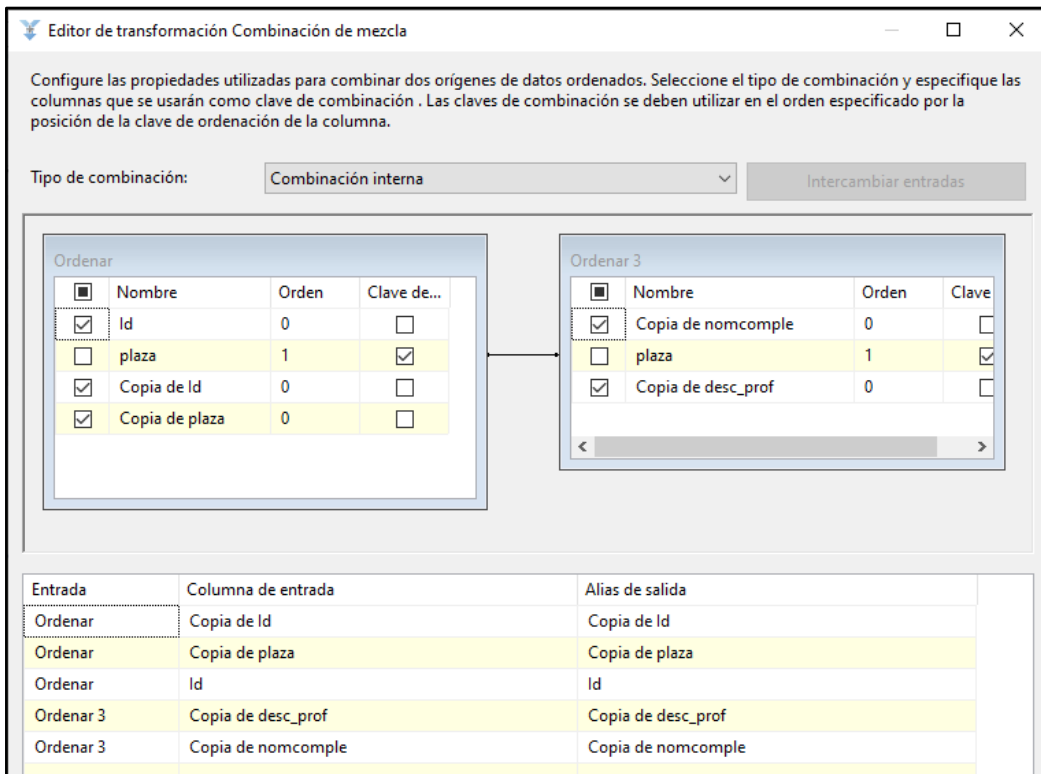


Figura 4.39. Mezcla de datos para la tabla de dimensión Profesional salud

D. POBLAR DIMENSIÓN PROFESIONAL SALUD

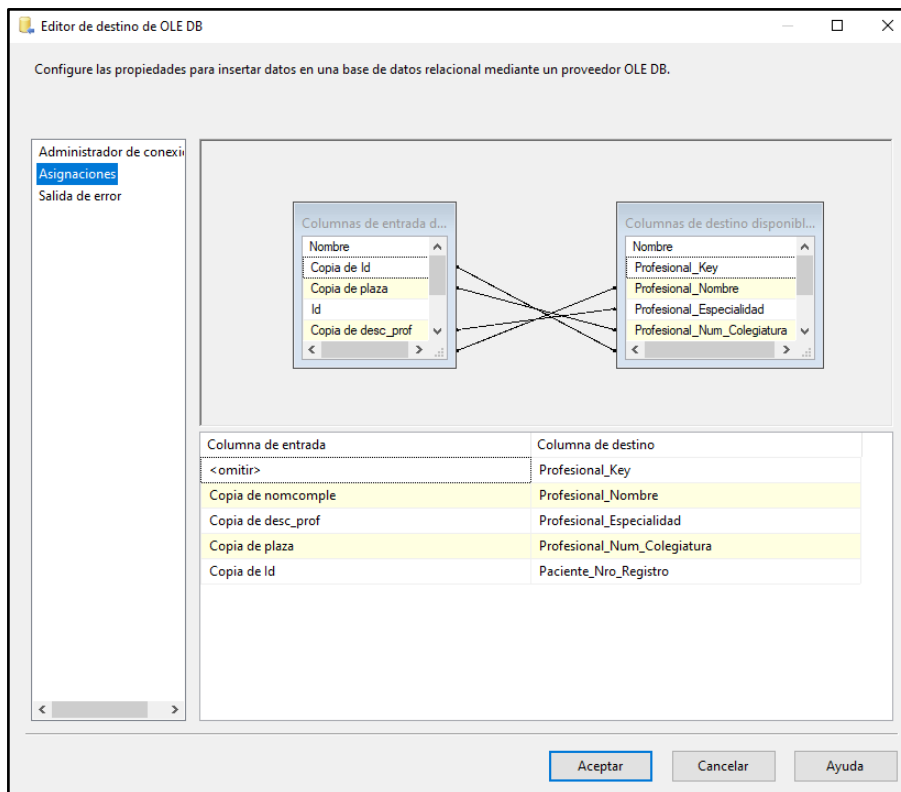


Figura 4.40. Carga de datos a la tabla dim_ProfesionalSalud

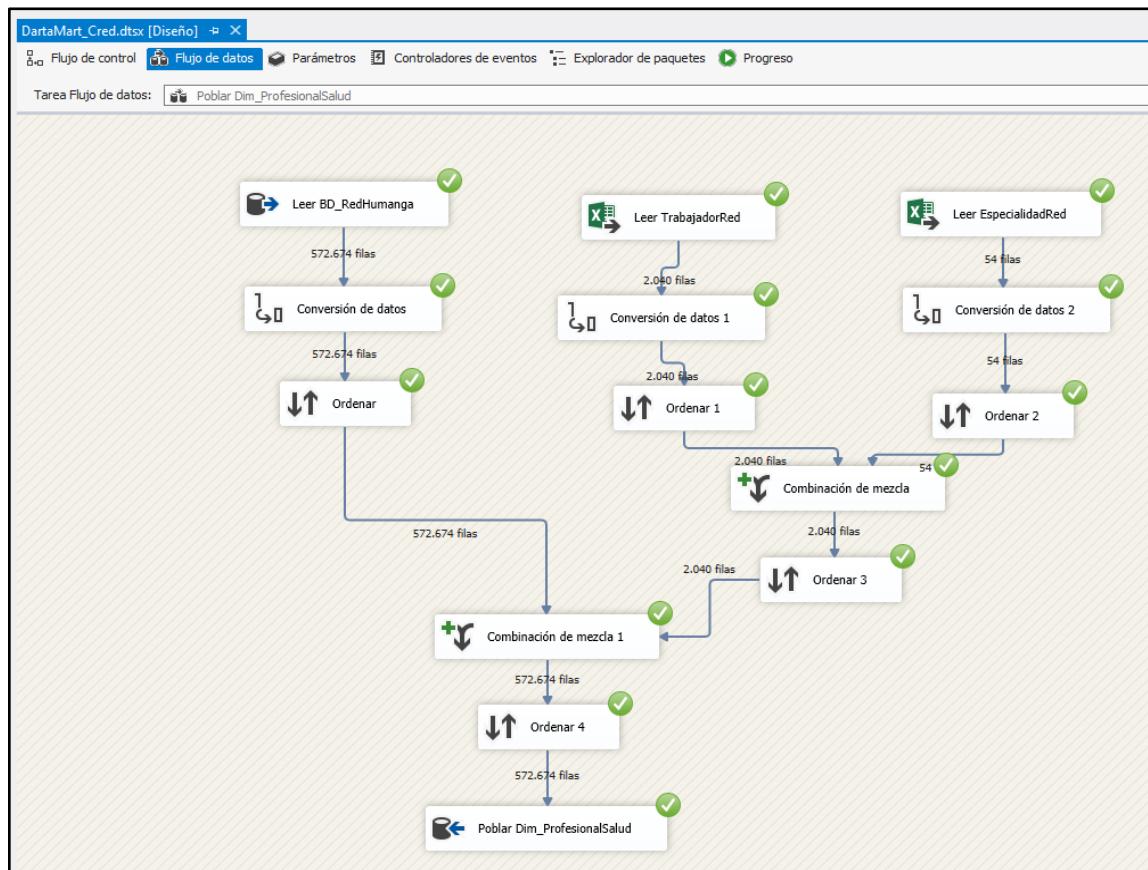


Figura 4.41. Proceso ETL culminado para la dimensión profesional salud

5. POBLAR DIMENSIÓN CENTRO SALUD

A. ORIGEN DE DATOS

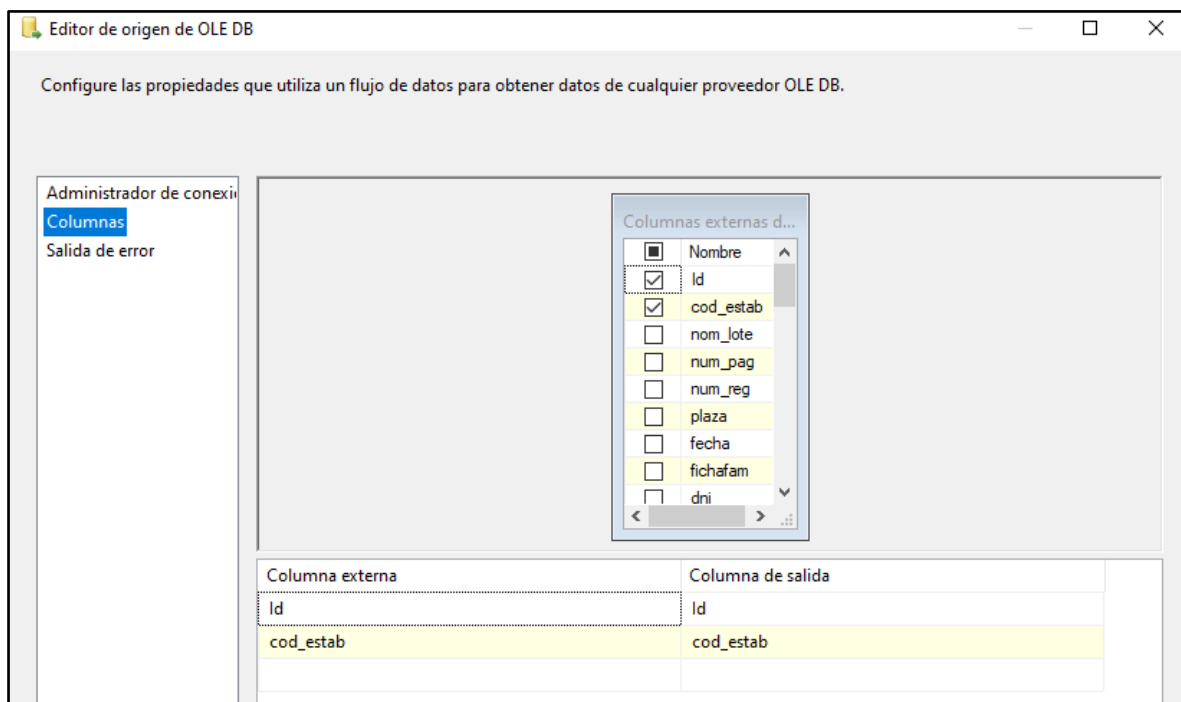


Figura 4.42. Origen de datos para la dimensión centro salud desde BD_His2018

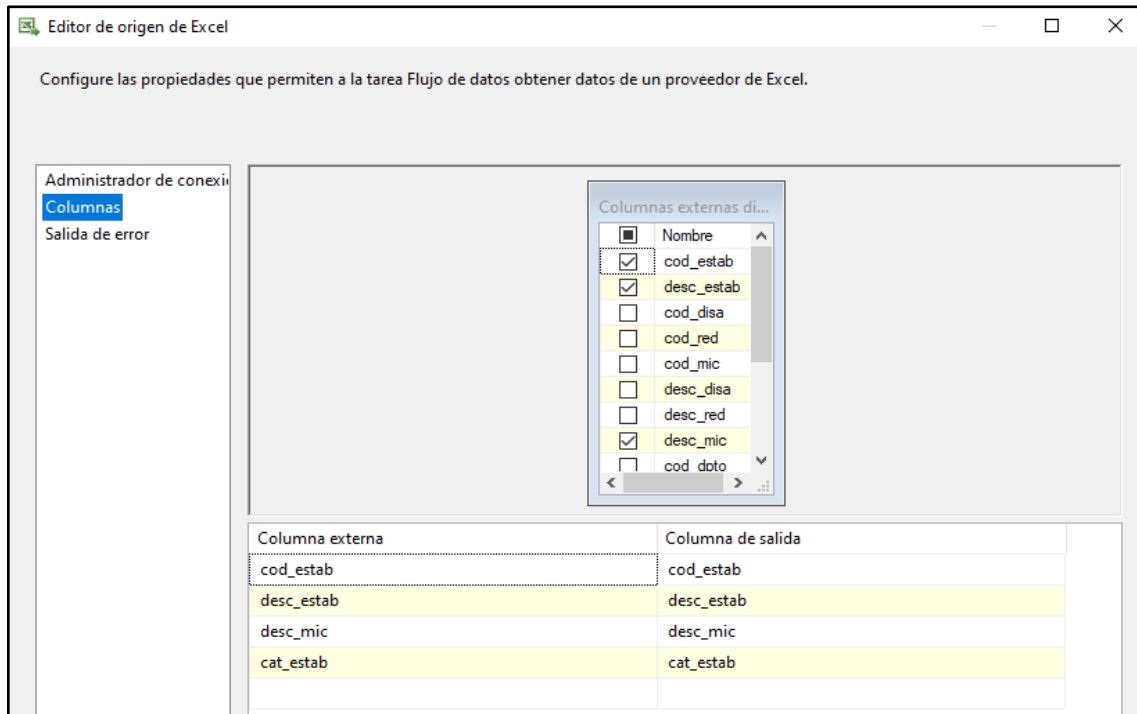


Figura 4.43. Origen de datos para la dimensión centro salud desde un archivo Excel

B. CONVERSIÓN DE DATOS

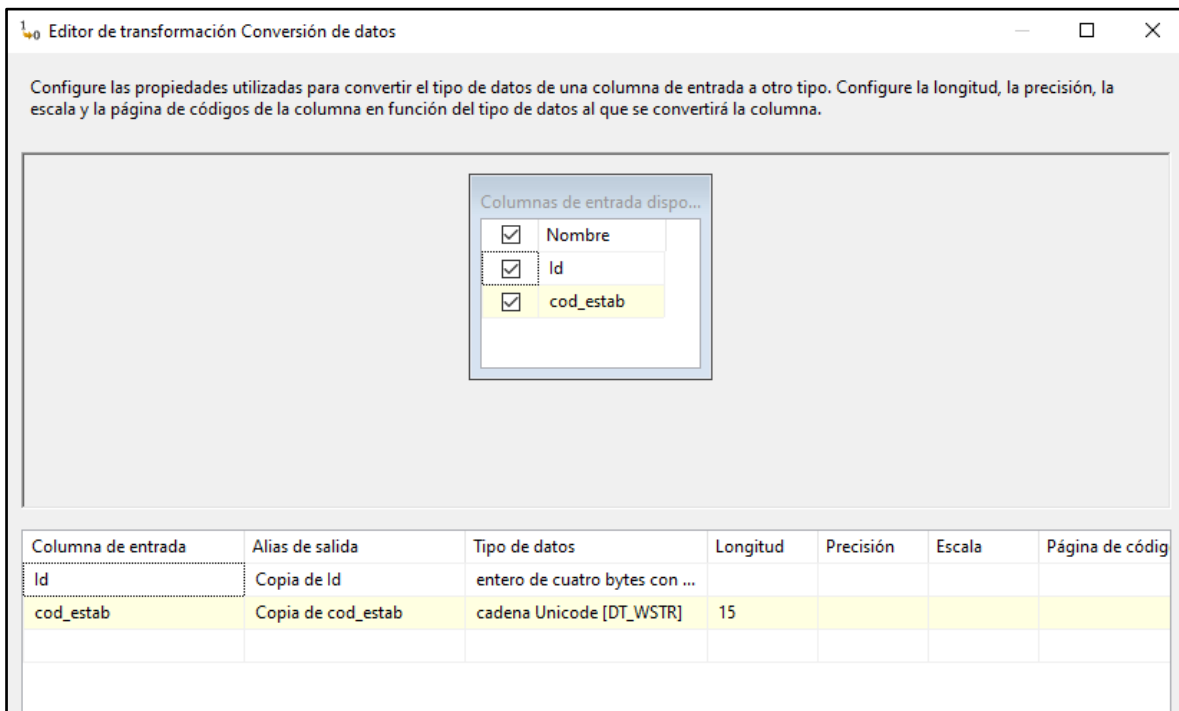


Figura 4.44. Transformación de datos proveniente de origen BD_His 2018

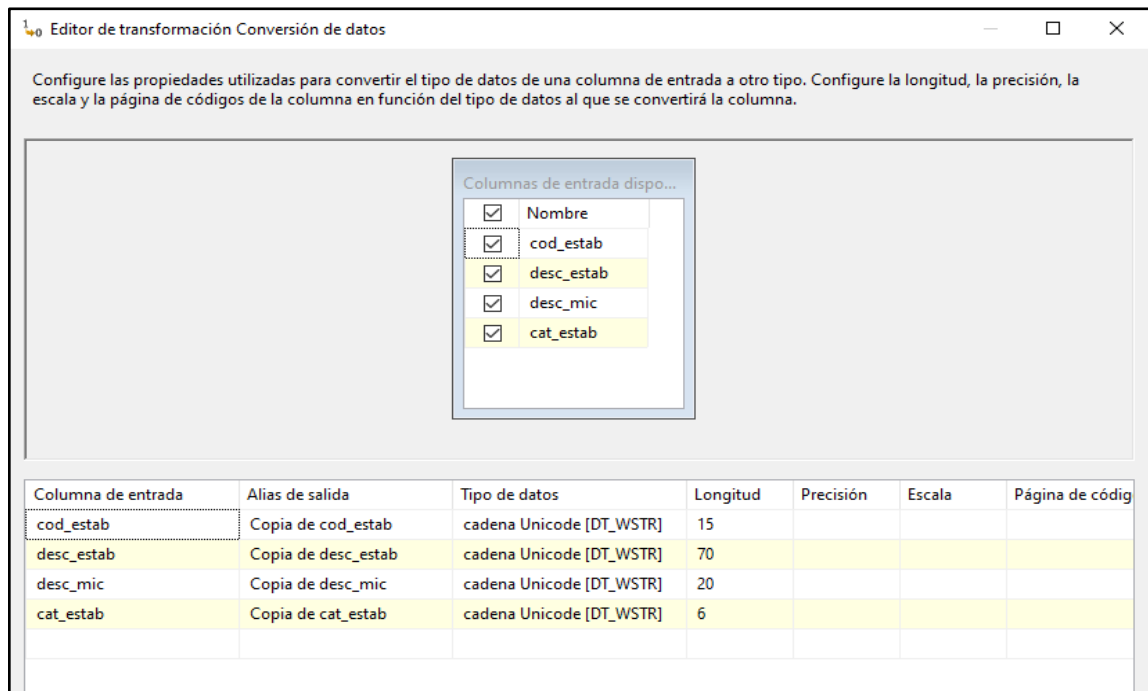


Figura 4.45. Transformación de datos proveniente de origen Excel

C. MEZCLA DE DATOS

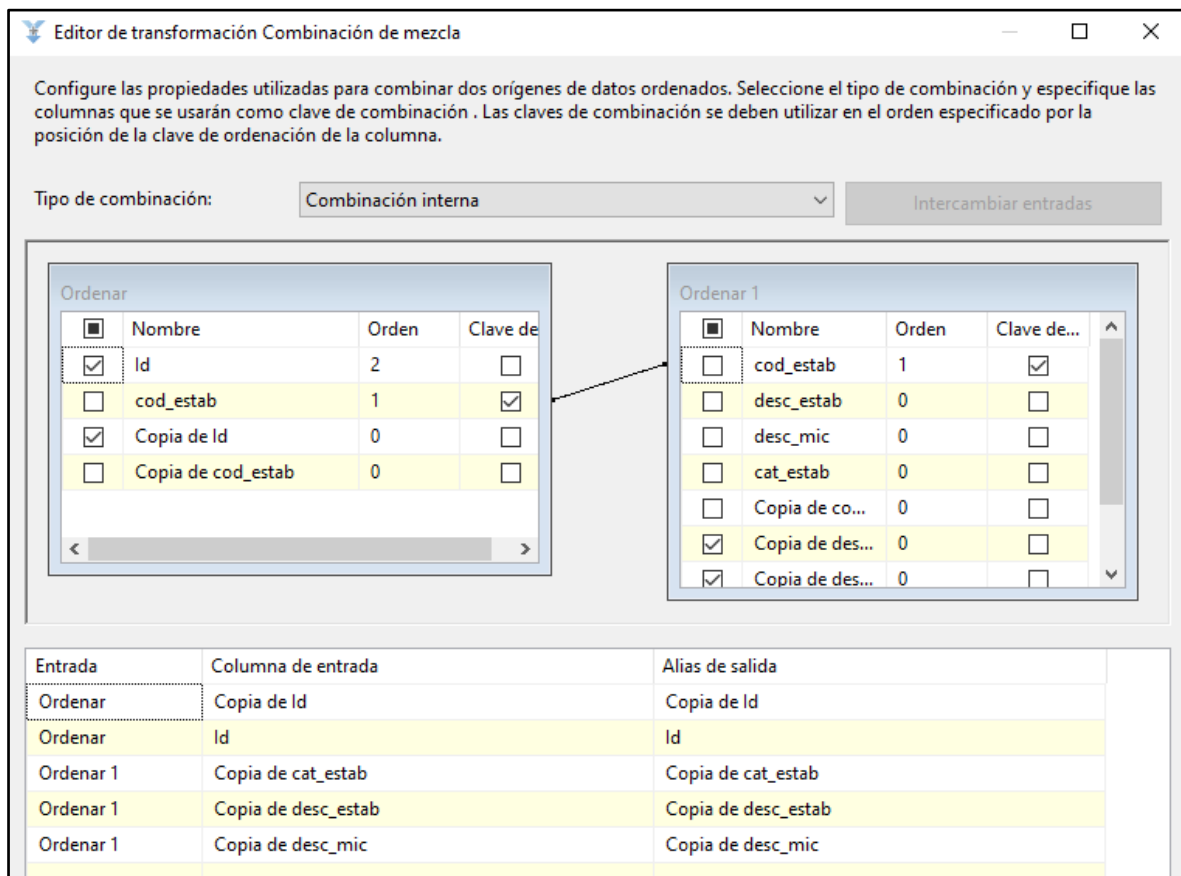


Figura 4.46. Mezcla de datos provenientes de BD_His2018 y Archivo Excel

D. POBLAR DIMENSIÓN CENTRO SALUD

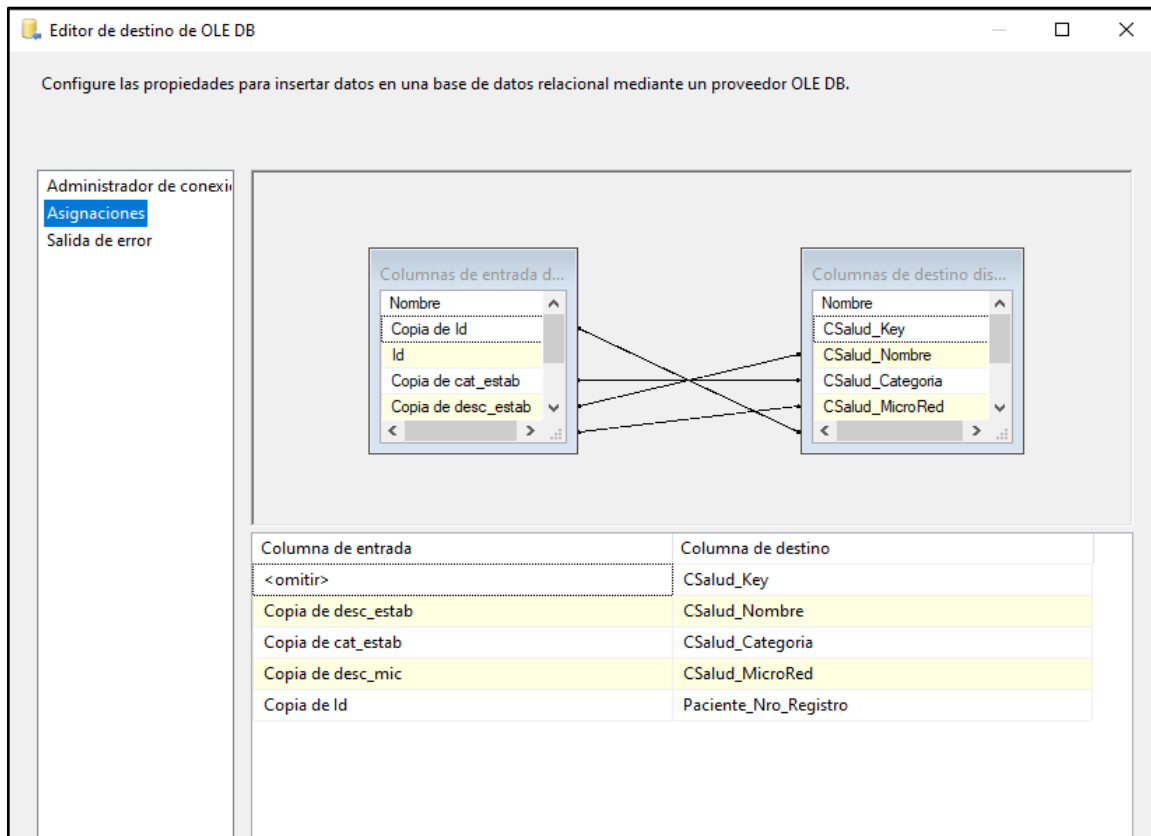


Figura 4.47. Carga de datos a la dimensión centro salud.

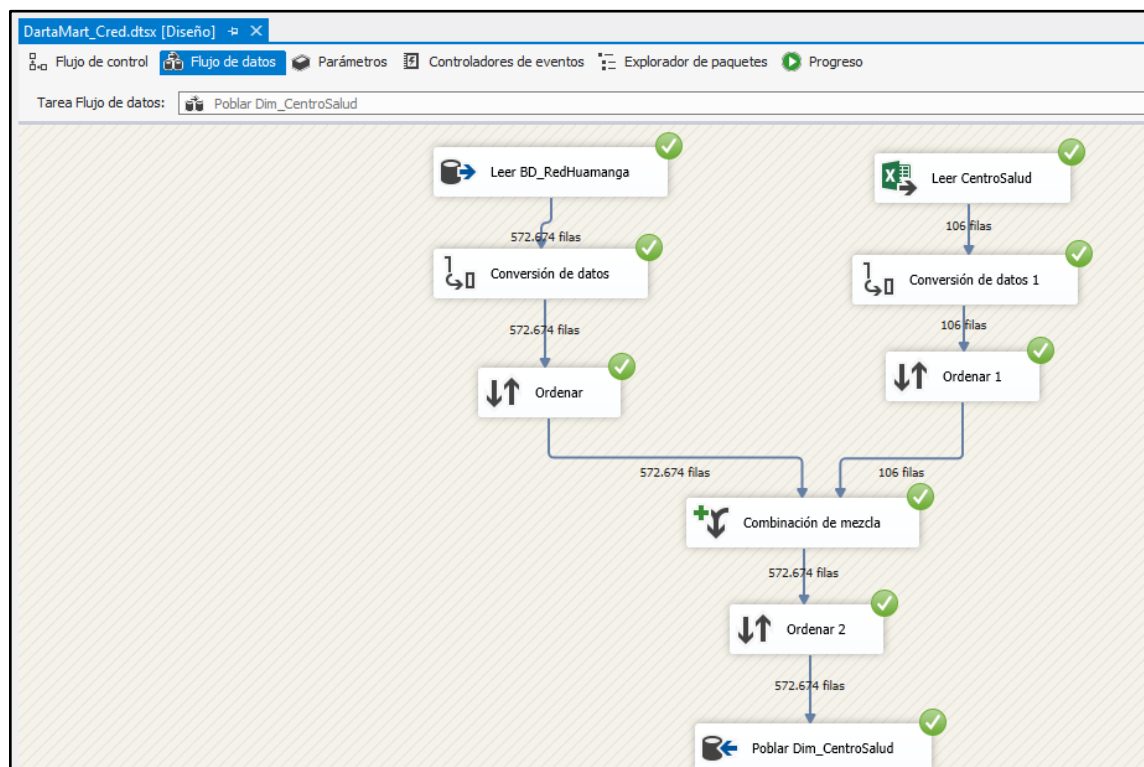


Figura 4.48. Proceso ETL culminado para la dimensión centro salud.

6. POBLAR DIMENSIÓN UBIGEO

A. ORIGEN DE DATOS

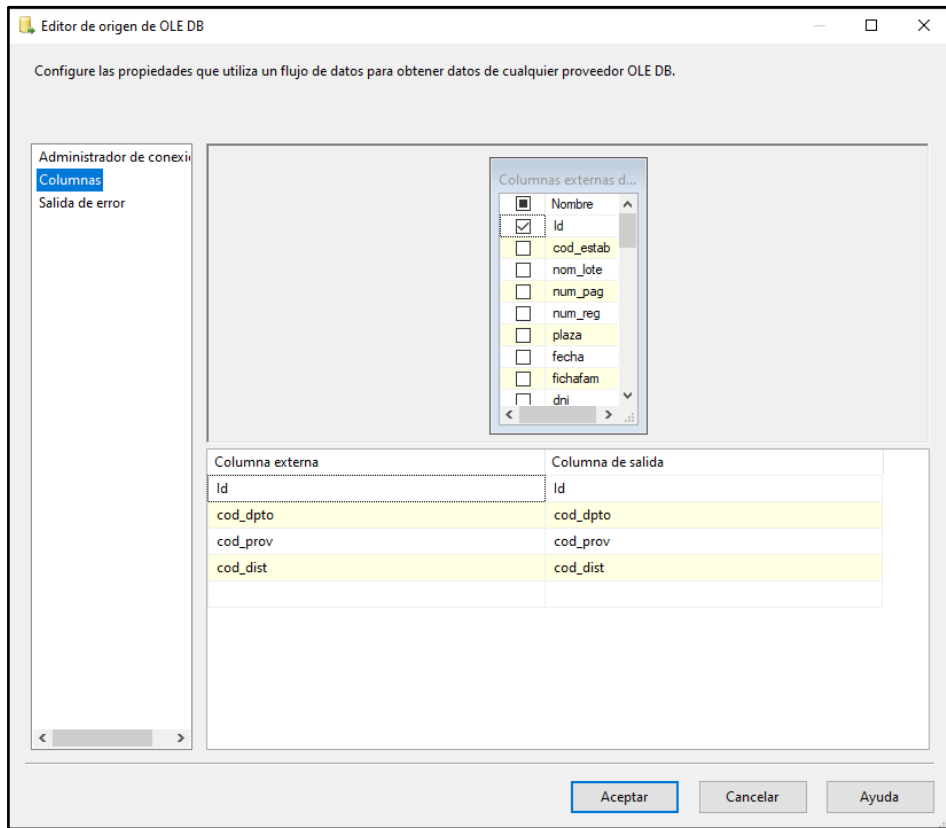


Figura 4.49. Origen de datos para la dimensión ubigeo desde la BD_His218.

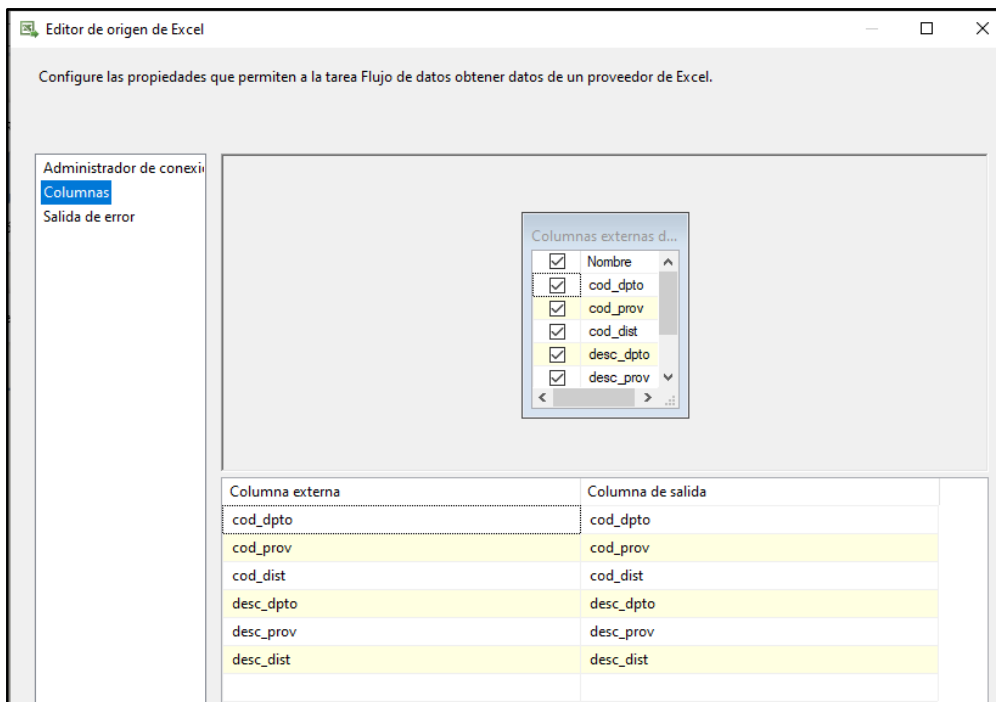


Figura 4.50. Origen de datos para la dimensión ubigeo desde un archivo Excel.

B. CONVERSIÓN DE DATOS

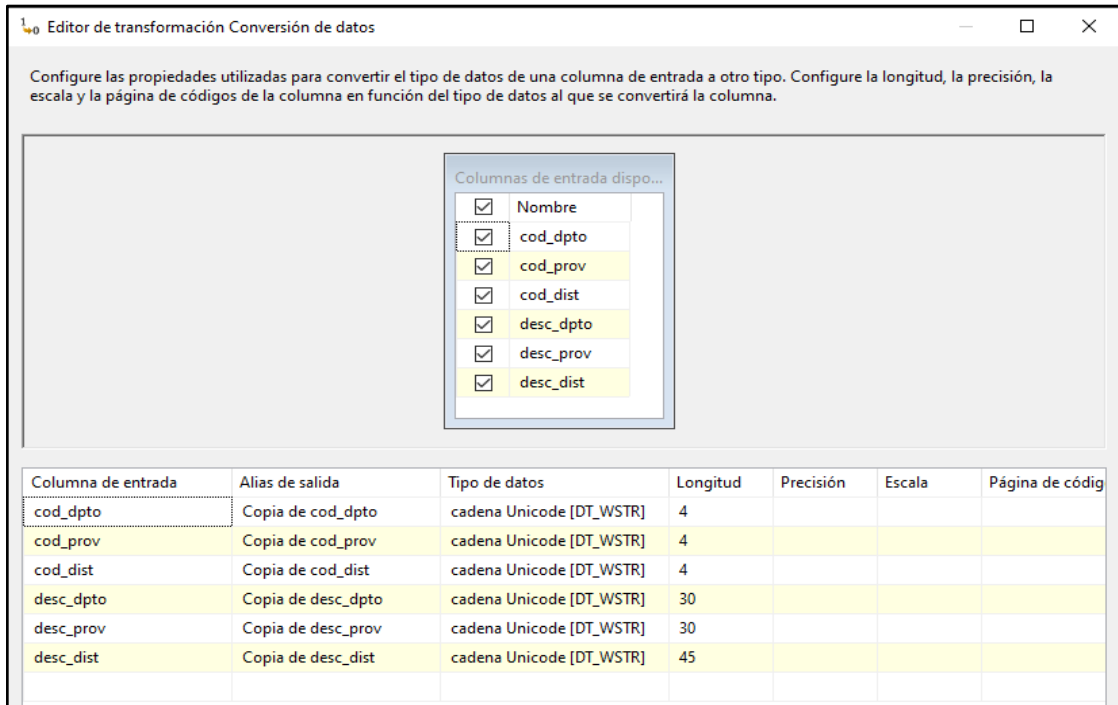


Figura 4.51. Transformación de datos para la dimensión ubigeo

C. POBLAMIENTO DE DATOS

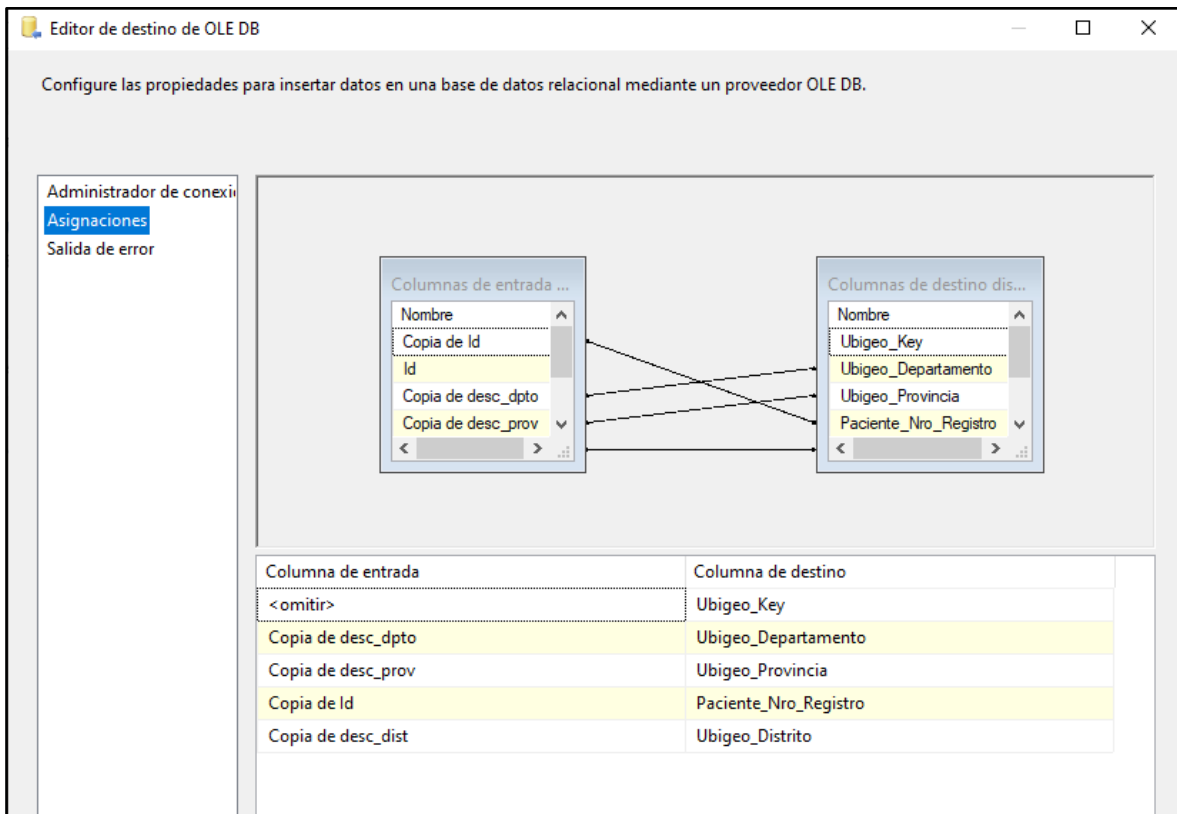


Figura 4.52. Carga de datos para la dimensión ubigeo.

7. POBLAR DIMENSIÓN TIEMPO

A. ORIGEN DE DATOS

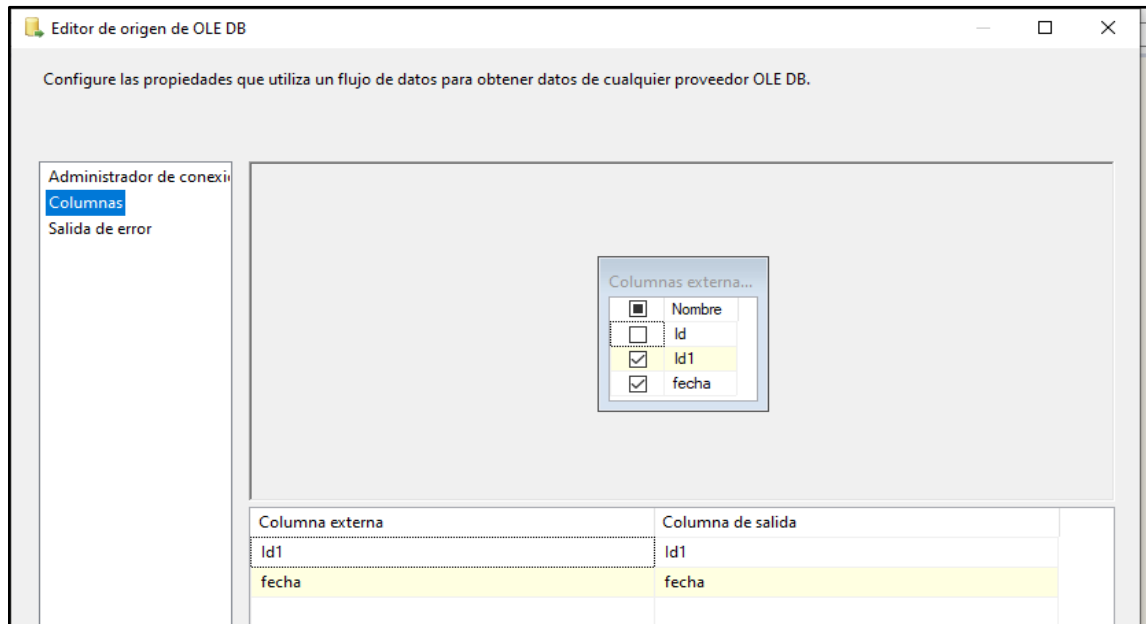


Figura 4.53. Origen de datos para la dimensión tiempo.

B. CONVERSIÓN DE DATOS

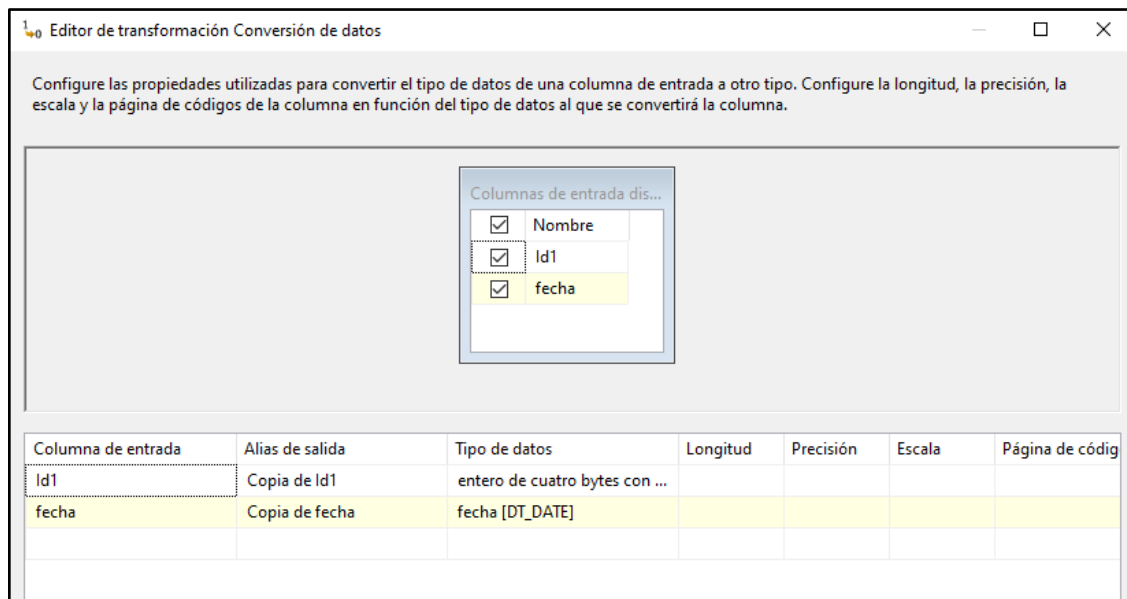


Figura 4.54. Transformación de datos para la dimensión tiempo.

C. POBLAMIENTO DE DATOS

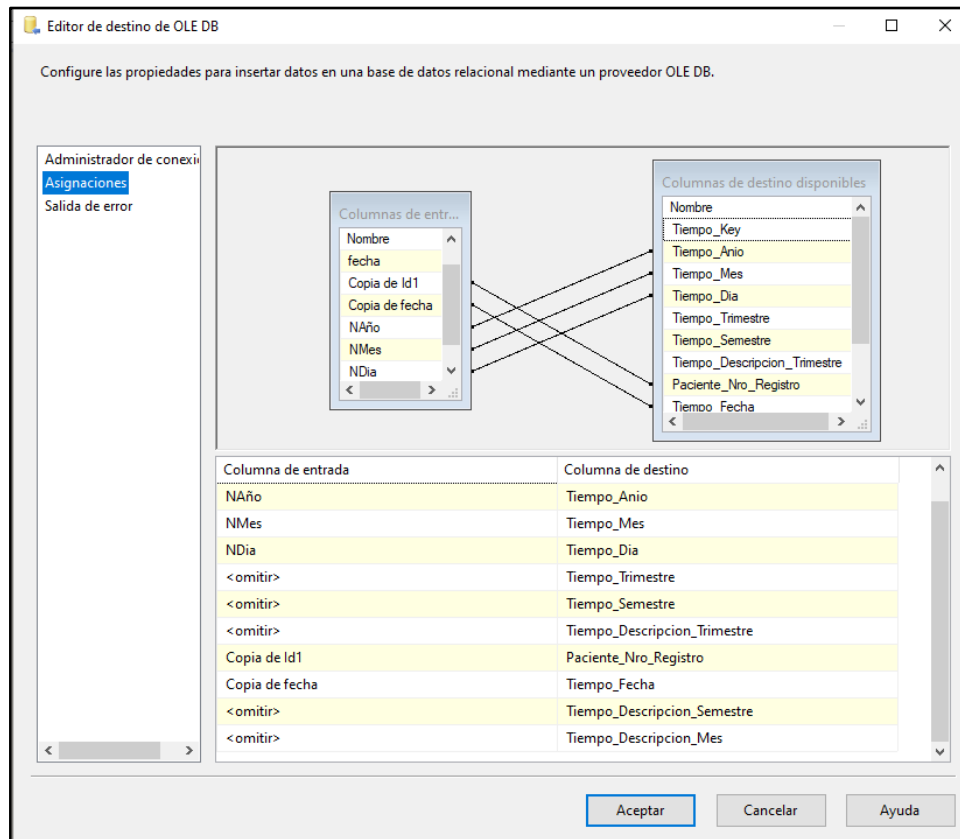


Figura 4.55. Carga de datos a la tabla dim_Tiempo.

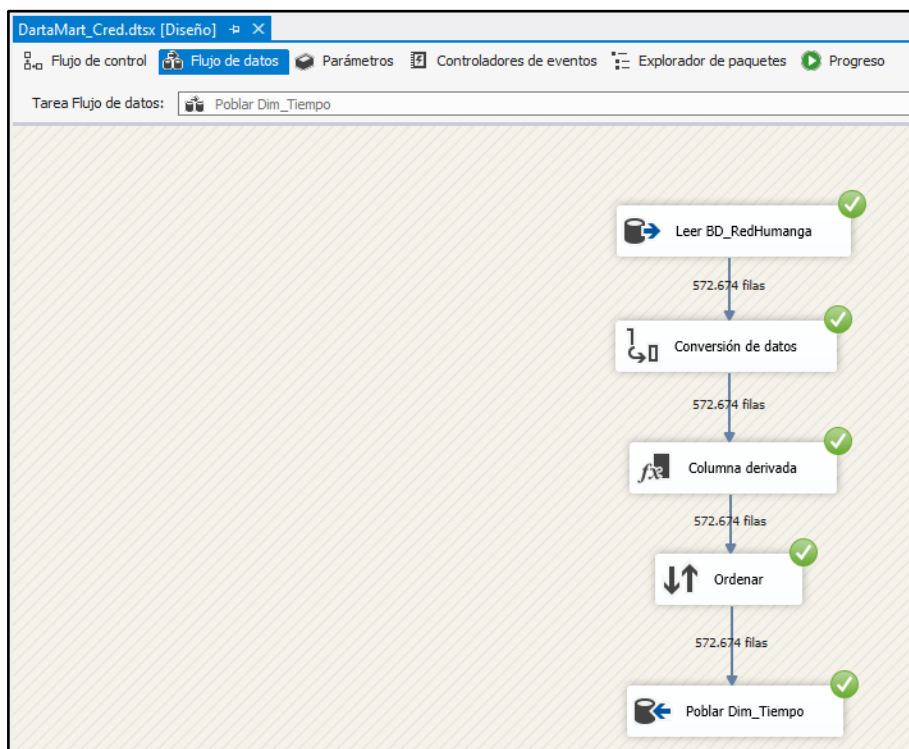


Figura 4.56. Proceso ETL culminado para la dimensión Tiempo

8. COMPLETAR EL POBLAMIENTO DE LA DIMENSIÓN TIEMPO

```
USE [BD_DatMartCRED]
GO
/***** Object: StoredProcedure [dbo].[sp_CargarTiempo]  Script Date: 30/04/2020
18:36:08 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO

CREATE procedure [dbo].[sp_CargarTiempo]
as
begin
update DIM_Tiempo set DIM_Tiempo.Tiempo_Trimestre=cast(case
when Tiempo_Mes in('01','02','03') then '01'
when Tiempo_Mes in('04','05','06') then '02'
when Tiempo_Mes in('07','08','09') then '03'
when Tiempo_Mes in('10','11','12') then '04'
end as nvarchar), DIM_Tiempo.Tiempo_Semestre=cast(case
when Tiempo_Mes in('01','02','03','04','05','06') then '01'
when Tiempo_Mes in('07','08','09','10','11','12') then '02'
end as nvarchar),DIM_Tiempo.Tiempo_Descripcion_Trimestre =cast(case
when Tiempo_Mes in('01','02','03') then 'PRIMER TRIMESTRE'
when Tiempo_Mes in('04','05','06') then 'SEGUNDO TRIMESTRE'
when Tiempo_Mes in('07','08','09') then 'TERCER TRIMESTRE'
when Tiempo_Mes in('10','11','12') then 'CUARTO TRIMESTRE'
end as nvarchar),DIM_Tiempo.Tiempo_Descripcion_Semestre =cast(case
when Tiempo_Mes in('01','02','03','04','05','06') then 'PRIMER SEMESTRE'
when Tiempo_Mes in('07','08','09','10','11','12') then 'SEGUNDO SEMESTRE'
end as nvarchar), Dim_Tiempo.Tiempo_Descripcion_Mes= cast (case
when Tiempo_Mes in ('1') then 'Enero'
```

```

when Tiempo_Mes in ('2') then 'Febrero'
when Tiempo_Mes in ('3') then 'Marzo'
when Tiempo_Mes in ('4') then 'Abril'
when Tiempo_Mes in ('5') then 'Mayo'
when Tiempo_Mes in ('6') then 'Junio'
when Tiempo_Mes in ('7') then 'Julio'
when Tiempo_Mes in ('8') then 'Agosto'

```

Figura 4.57. Proceso ETL para culminar la carga de la dimensión tiempo

9. POBLAR LA TABLA DE HECHO FACT_CONTROL_CRED

```

USE [DMCRED]
GO
/***** Object: StoredProcedure [dbo].[Sp_cargar_FactCred]*****/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE procedure [dbo].[Sp_cargar_FactCred]
as
begin

insert into FAC_ControlCred (His_Key, Paciente_Key,
                            Ubigeo_Key,
                            CSalud_Key,
                            Profesional_Key,
                            Tiempo_Key,

                            [Cred_TotalAtendios],[Cred_Estado_Nutricional],
                            [Cred_Nro_RecienNacidos],
                            [Cred_Ninios_1Anios],[Cred_Ninios_2Anios],
                            [Cred_Ninios_3Anios],[Cred_Ninios_4Anios],
                            [Cred_Ninios_5Anios],
                            [Cred_Ninios_Mayores_5Anios],
                            [Cred_Ganancia_InadecuadaPeso],

                            [Cred_RN_PesoExtremadamente_Bajo],[Cred_RN_Peso_MuyBajo],[Cred_RN_PesoBajo_
Nacer],
                            [Cred_Desnutricion_Severa],[Cred_Desnutricion_Aguda],
                            [Cred_SobrePeso],[Cred_Tamizaje_Anemia])

select
Hs.[His_Key],          pa.[Paciente_Key],Ub.[Ubigeo_Key],          Cs.[CSalud_Key],
Pr.[Profesional_Key], Tm.[Tiempo_Key],

```



```

CAST (case
when Hs.his_codigo1 ='Z001' or Hs.His_Codigo2 ='Z001' or Hs.his_codigo3 ='Z001'
or Hs.his_codigo4 ='Z001'
and Hs.his_codigo5 ='Z001' or Hs.his_codigo6 ='Z001' THEN 1 --CONTROLCREC
else 0
End AS int ),

CAST (CASE
WHEN Hs.His_Codigo1 = 'Z006' or Hs.His_Codigo2 = 'Z006' or Hs.His_Codigo2 =
'Z006' or Hs.His_Codigo2 = 'Z006'
or Hs.His_Codigo2 = 'Z006' or Hs.His_Codigo2 = 'Z006' THEN '1' -- ESTADO
NURICIONAL
else 0
END AS INT),-- CENTRO SALUD

CAST (CASE
when (pa.Paciente_Edad > 0 and pa.paciente_edad < 30 and
pa.Paciente_Tipo_Edad='Dias' and (Hs.His_Codigo1 ='Z001' or Hs.His_Codigo2 =
'Z001' or Hs.His_Codigo3 = 'Z001'
or Hs.His_Codigo4 = 'Z001' or Hs.His_Codigo5 = 'Z001' or Hs.His_Codigo6 = 'Z001'))
) then 1 -- RECIEN NACIDOS
else 0
End as int),

CAST (CASE
when (pa.Paciente_Edad > 0 and pa.paciente_edad<12 and
pa.Paciente_Tipo_Edad='Meses'and (Hs.His_Codigo1 ='Z001' or Hs.His_Codigo2 =
'Z001' or Hs.His_Codigo3 = 'Z001'
or Hs.His_Codigo4 = 'Z001' or Hs.His_Codigo5 = 'Z001' or Hs.His_Codigo6 =
'Z001')) then 1 -- NIÑOS DE UN AÑO
else 0
End as int),

CAST (CASE
when (pa.Paciente_Edad = '2' and pa.Paciente_Tipo_Edad='Años'and (Hs.His_Codigo1
='Z001' or Hs.His_Codigo2 = 'Z001' or Hs.His_Codigo3 = 'Z001'
or Hs.His_Codigo4 = 'Z001' or Hs.His_Codigo5 = 'Z001' or Hs.His_Codigo6 =
'Z001')) then 1 -- NIÑOS DE 2 AÑOS

```

```

CAST (CASE
when (pa.Paciente_Edad = '5' and pa.Paciente_Tipo_Edad='Años'and
(Hs.His_Codigo1 ='Z001' or Hs.His_Codigo2 = 'Z001' or Hs.His_Codigo3 = 'Z001'
or Hs.His_Codigo4 = 'Z001' or Hs.His_Codigo5 = 'Z001' or Hs.His_Codigo6 =
'Z001')) then 1 -- NIÑOS DE 5 AÑOS
else 0
End as int),

CAST (CASE
when (pa.Paciente_Edad >'5' and pa.Paciente_Edad <'12' and
pa.Paciente_Tipo_Edad='Años'and (Hs.His_Codigo1 = 'Z001' or Hs.His_Codigo2 =
'Z001' or Hs.His_Codigo3 = 'Z001'
or Hs.His_Codigo4 = 'Z001' or Hs.His_Codigo5 = 'Z001' or Hs.His_Codigo6 =
'Z001')) then 1 -- NIÑOS MAYORES DE 5 AÑOS
else 0
End as int),

CAST (CASE
when Hs.his_codigo1 ='Z724'or Hs.His_Codigo2 ='Z724' or Hs.his_codigo3 ='Z724'
or Hs.his_codigo4 ='Z724'
and Hs.his_codigo5 ='Z724' or Hs.his_codigo6 ='Z724'THEN 1 --GANANCIA INADECUADA
DE PESO
else 0
End AS int ),

CAST (CASE
when Hs.his_codigo1 ='P070'or Hs.His_Codigo2 ='P070' or Hs.his_codigo3 ='P070'
or Hs.his_codigo4 ='P070'
and Hs.his_codigo5 ='P070' or Hs.his_codigo6 ='P070'THEN 1 --RN PESO
EXTREMADAMENTE BAJO
else 0
End AS int ),

CAST (CASE
when Hs.his_codigo1 ='P0711'or Hs.His_Codigo2 ='P0711' or Hs.his_codigo3
='P0711' or Hs.his_codigo4 ='P0711'
and Hs.his_codigo5 ='P0711' or Hs.his_codigo6 ='P0711'THEN 1 --RN MUY BAJO PESO
else 0
End AS int ),

CAST (CASE
when Hs.his_codigo1 ='P0712'or Hs.His_Codigo2 ='P0712' or Hs.his_codigo3
='P0712' or Hs.his_codigo4 ='P0712'
and Hs.his_codigo5 ='P0712' or Hs.his_codigo6 ='P0712'THEN 1 --RN BAJO PESO AL
NACER
else 0
End AS int ),

CAST (CASE
when Hs.his_codigo1 ='E43X'or Hs.His_Codigo2 ='E43X' or Hs.his_codigo3 ='E43X'
or Hs.his_codigo4 ='E43X'
and Hs.his_codigo5 ='E43X' or Hs.his_codigo6 ='E43X'THEN 1 --DESNUTRICION SEVERA
else 0
End AS int ),

```

```

CAST (CASE
when Hs.his_codigo1 ='E440'or Hs.His_Codigo2 ='E440' or Hs.his_codigo3
='E440' or Hs.his_codigo4 ='E440'
and Hs.his_codigo5 ='E440' or Hs.his_codigo6 ='E440'THEN 1 --DESNUTRICION
AGUDA
else 0
End AS int ),

CAST (CASE
when Hs.his_codigo1 ='E660'or Hs.His_Codigo2 ='E660' or Hs.his_codigo3
='E660' or Hs.his_codigo4 ='E660'
and Hs.his_codigo5 ='E660' or Hs.his_codigo6 ='E660'THEN 1 --SOBREPESO
else 0
End AS int ),

CAST (CASE
when Hs.his_codigo1 ='Z017'or Hs.His_Codigo2 ='Z017' or Hs.his_codigo3
='Z017' or Hs.his_codigo4 ='Z017'
and Hs.his_codigo5 ='Z017' or Hs.his_codigo6 ='Z017'THEN 1 --SOBREPESO
else 0

End AS int )
from Dim_His hs

INNER JOIN Dim_CentroSalud Cs ON
Cs.PACIENTE_NRO_REGISTRO=Hs.PACIENTE_NRO_REGISTRO
INNER JOIN DIM_Paciente Pa ON Pa.Paciente_Nro_Registro=
Hs.Paciente_Nro_Registro
INNER JOIN Dim_Ubigeo ub ON ub.Paciente_Nro_Registro=
Hs.Paciente_Nro_Registro
INNER JOIN Dim_ProfesionalSalud pr ON pr.Paciente_Nro_Registro=
Hs.Paciente_Nro_Registro
INNER JOIN Dim_Tiempo Tm ON Tm.Paciente_Nro_Registro=
Hs.Paciente_Nro_Registro
END

```

Figura 4.58. Proceso ETL para cargar FACT_ControlCred

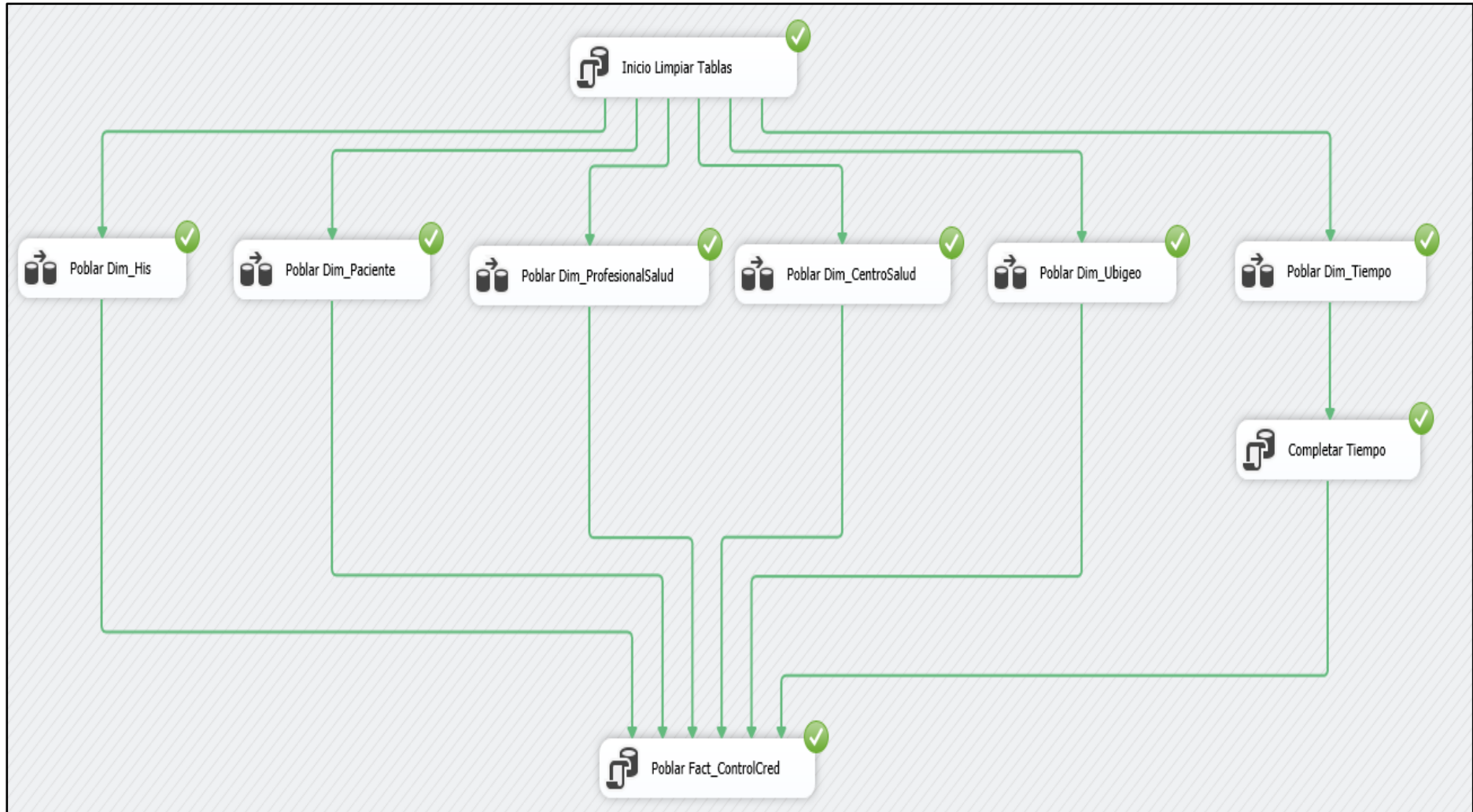


Figura 4.59. Proceso ETL Completado

4.1.8. IMPLEMENTACIÓN

IMPLEMENTACIÓN DEL CUBO MULTIDIMENSIONAL

Para la implementación del cubo se utilizará la herramienta que nos proporciona el Visual Studio 2012 que es el Analysis Service a la vez es parte del SQL Server, esto nos permitirá implementar de manera adecuado el cubo.

CONFIGURACIÓN DE ORÍGENES DE DATOS

Para poder realizar la conexión necesitamos realizar una serie de configuraciones, que este nos permitirá la comunicación entre el servidor y el Analysis Service, así para poder implementar el Data Mart que es BD_DataMart_ControlCred.

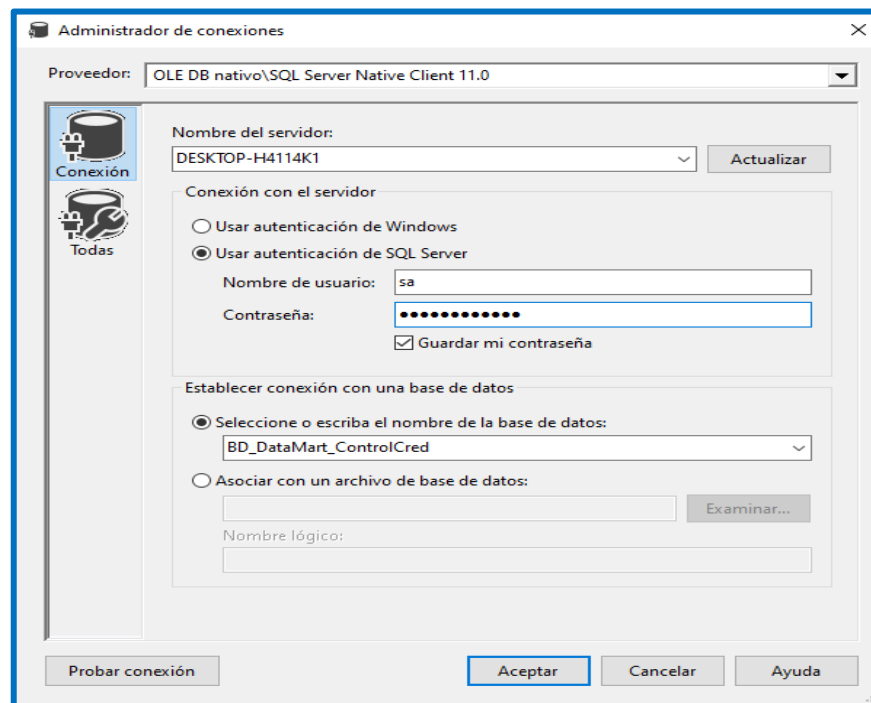


Figura 4.61. Administración de conexión

CONSTRUCCIÓN DE LA VISTA DE DATOS

En esta parte el asistente solicitará las tablas que formarán parte de nuestra vista, y el asistente nos mostrará las tablas de dimensiones como también la tabla de hecho. Que formarán el cubo, el asistente permitirá la unión de dimensiones y la tabla Fact_ControlCred.

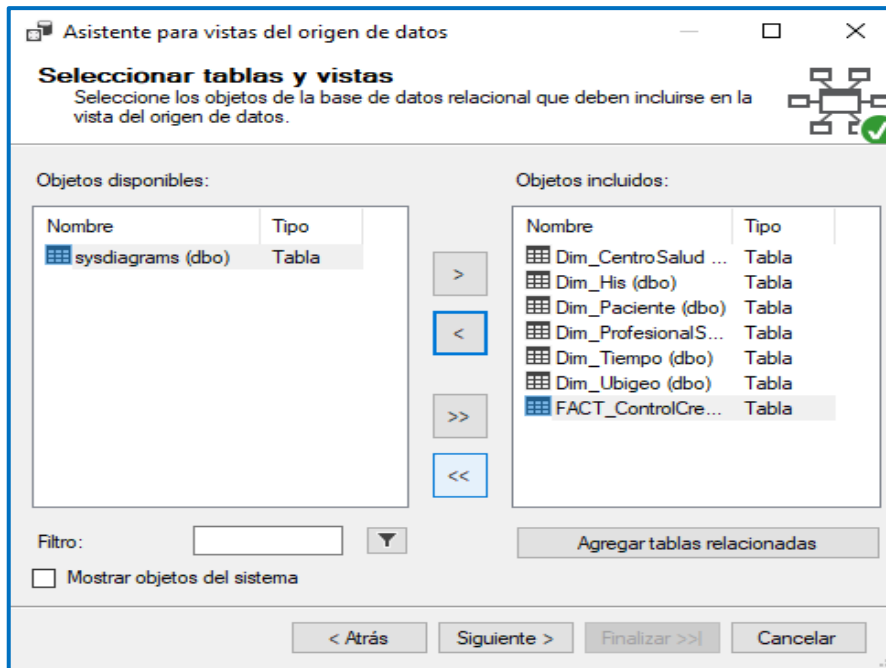


Figura 4.62. Selección de tablas y vistas

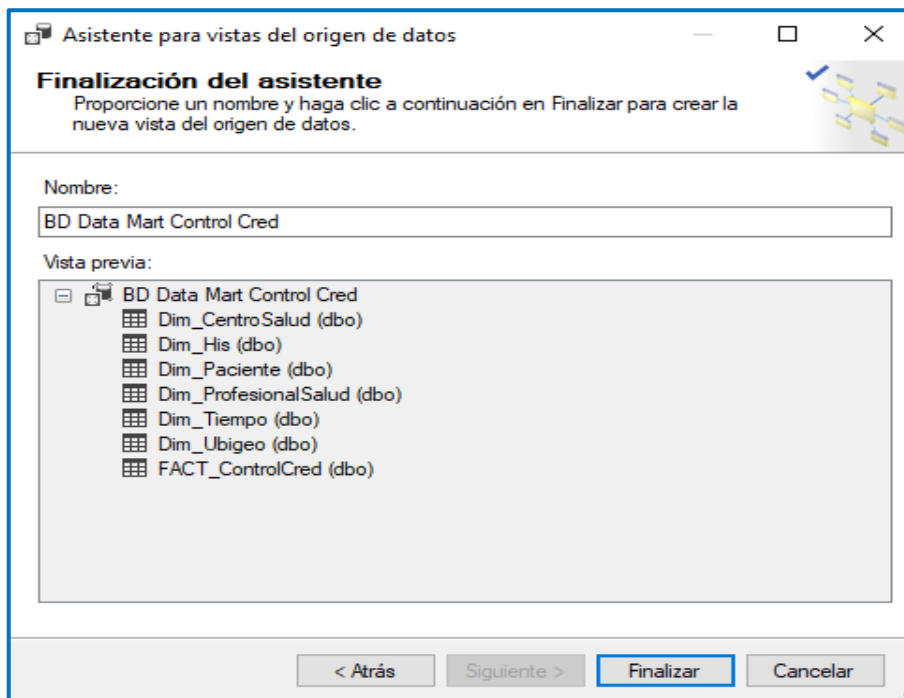


Figura 4.63. Tablas de dimensiones y tabla de hecho

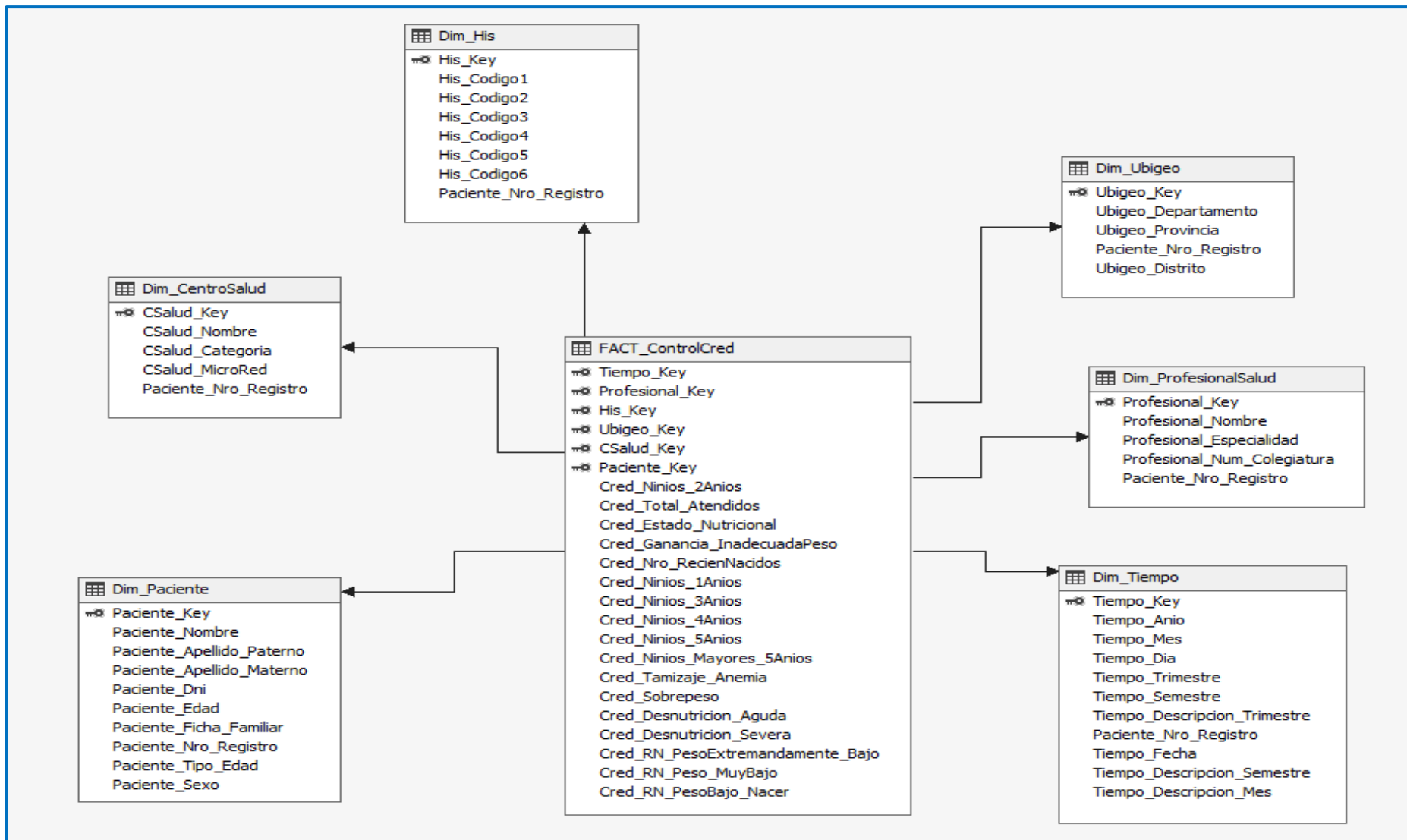


Figura 4.64. vista del origen de datos con el esquema de uniones

La Figura 4.64. Nos muestra la primera vista creada, donde se observa que se encuentra las diferentes tablas de dimensiones y tabla de hechos que es el Fact_ControlCred, que forman parte la primera vista, como así también las uniones entre la tabla de hecho y las tablas de dimensiones bajo el modelo estrella.

CONSTRUCCIÓN DEL CUBO MULTIDIMENSIONAL

Para la creación del cubo se necesita cargar la selección de tabla dimensiones y hechos obtenidas en nuestra primera vista, al seleccionar la tabla principal nos acogerá todas las tablas de dimensiones relacionadas a ella ya que están bien realizadas en caso contrario no se realizará.

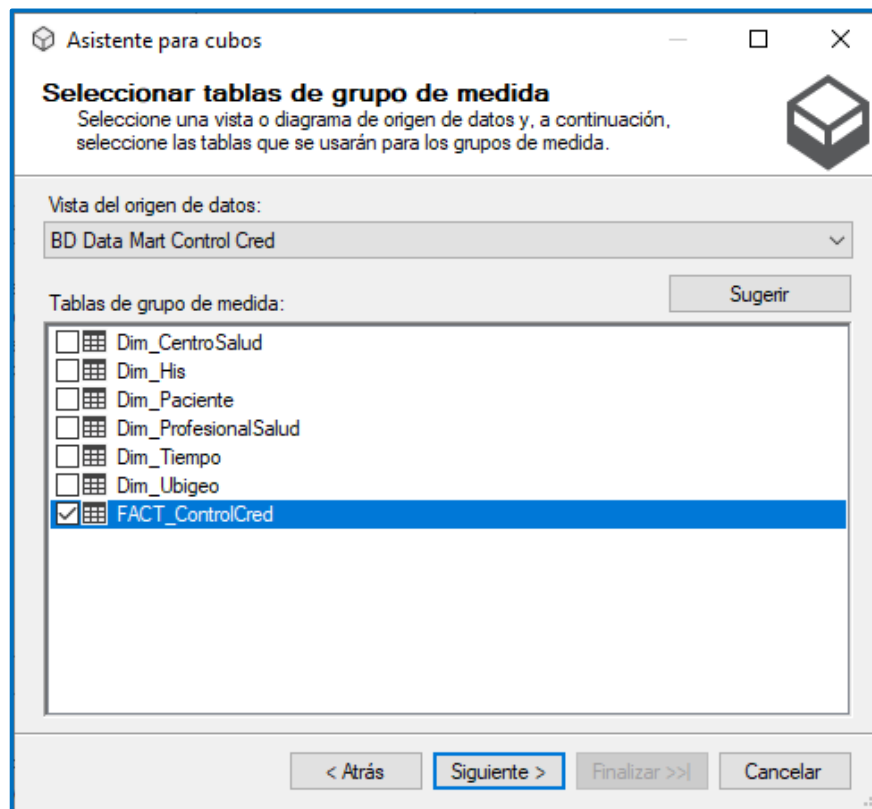


Figura 4.65. Selección de la tabla principal

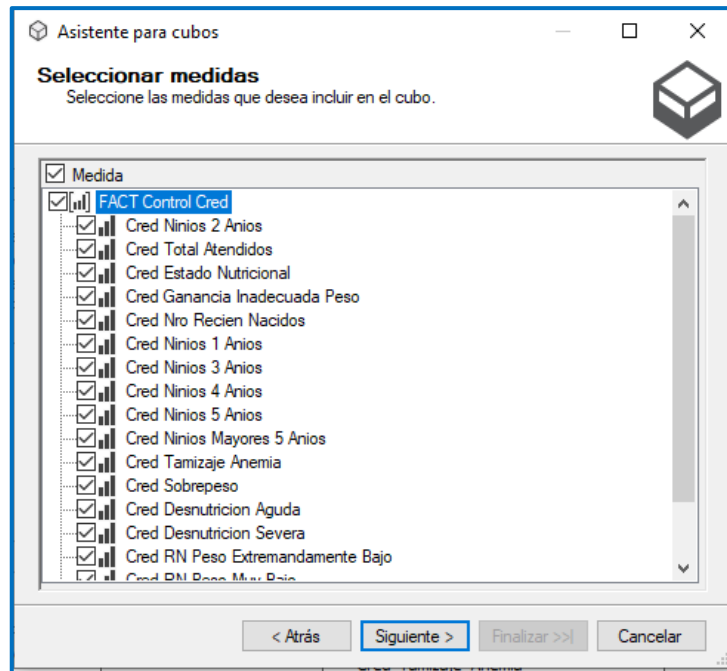


Figura 4.66. Tabla principal con el grupo de medida

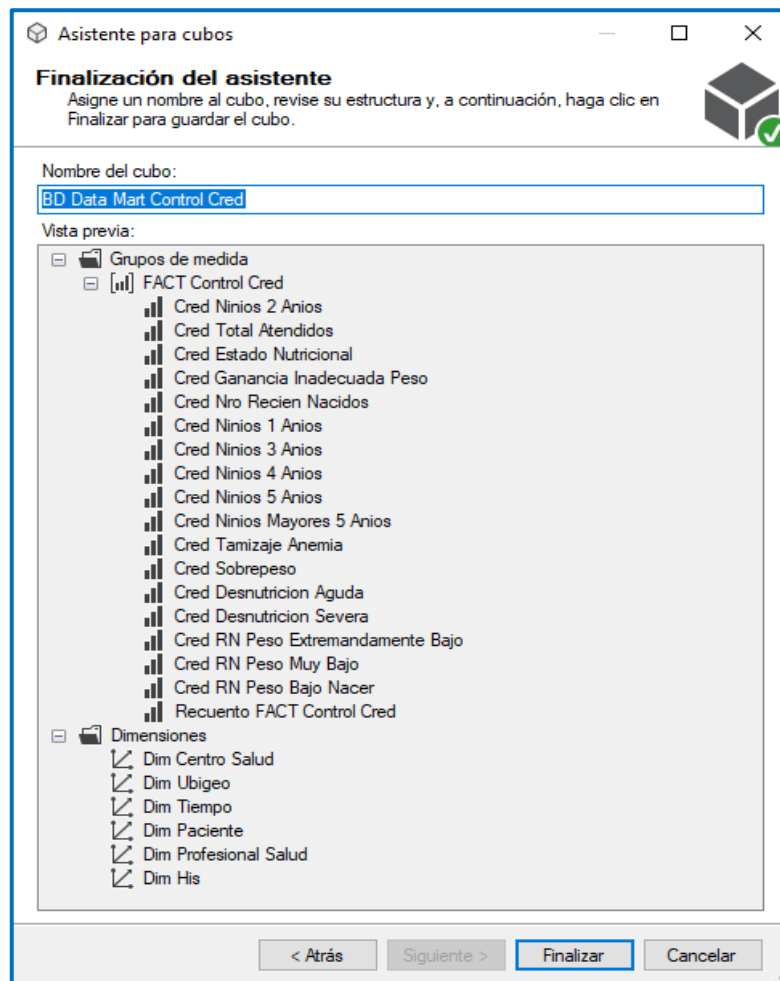


Figura 4.67. Creación del cubo con el grupo de medida y las dimensiones relacionadas.

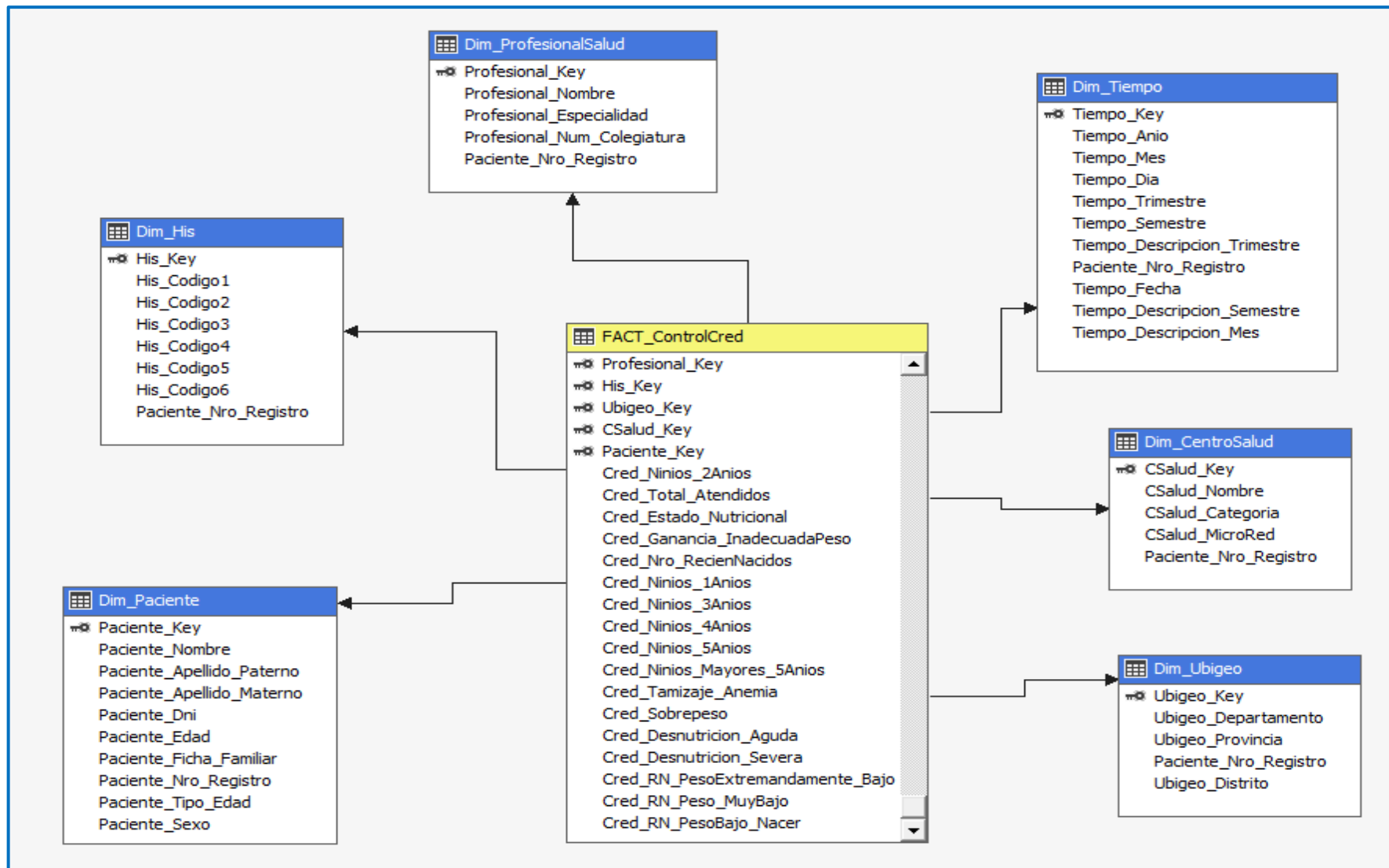


Figura 4.68. Configura ración final del cubo Fact_ControlCred.

4.1.9. ESPECIFICACIÓN DE APLICACIONES PARA USUARIOS FINALES

Como siendo una de las tareas de la mitología, nos dice una descripción con mayor detalle de roles y procedimientos del negocio, para ello es necesario identificar a los actores principales en la toma de decisiones, esto nos permitirá las responsabilidades específicas de cada rol, por ello se identificó a la parte operativa y la parte estratégica, siendo estos los usuarios finales.

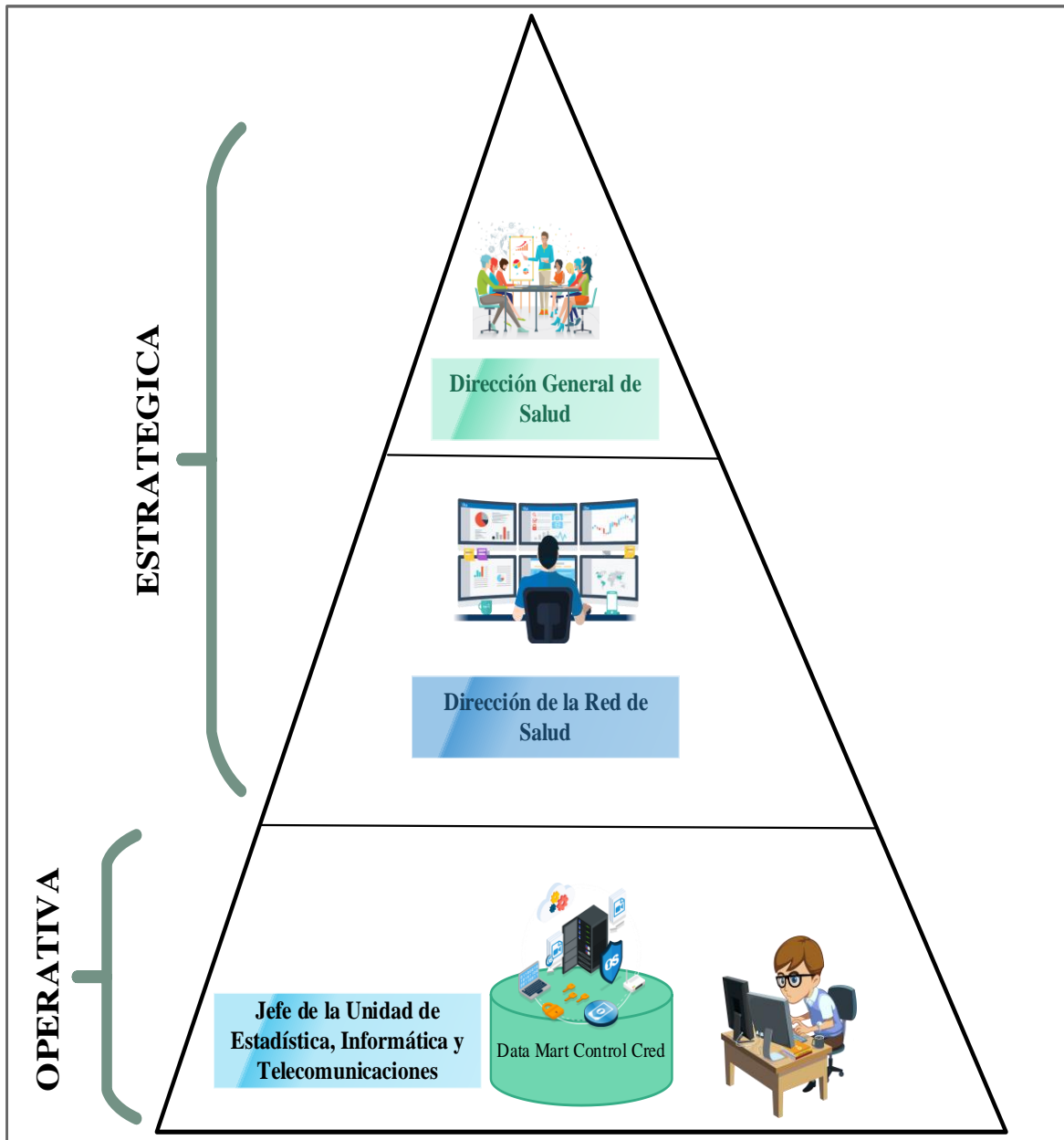


Figura 4.69. Identificación de roles.

4.1.10. DESARROLLO DE APLICACIONES PARA USUARIOS FINALES

Ya teniendo toda la parte de desarrollada hasta el cubo multidimensional, lo que procede ahora es el desarrollo de los reportes para el usuario final que sean fáciles de entender y tenga la mejor vista posible, para ello se realizara con la herramienta de Power BI.

CONEXIÓN CON POWER BI

En esta parte se realizará las configuraciones necesarias para poder realizar la conexión entre en Análisis Service que contiene el cubo del Data Mart Control Cred y el Power BI, la cual nos permitirá visualizar los gráficos necesarios.

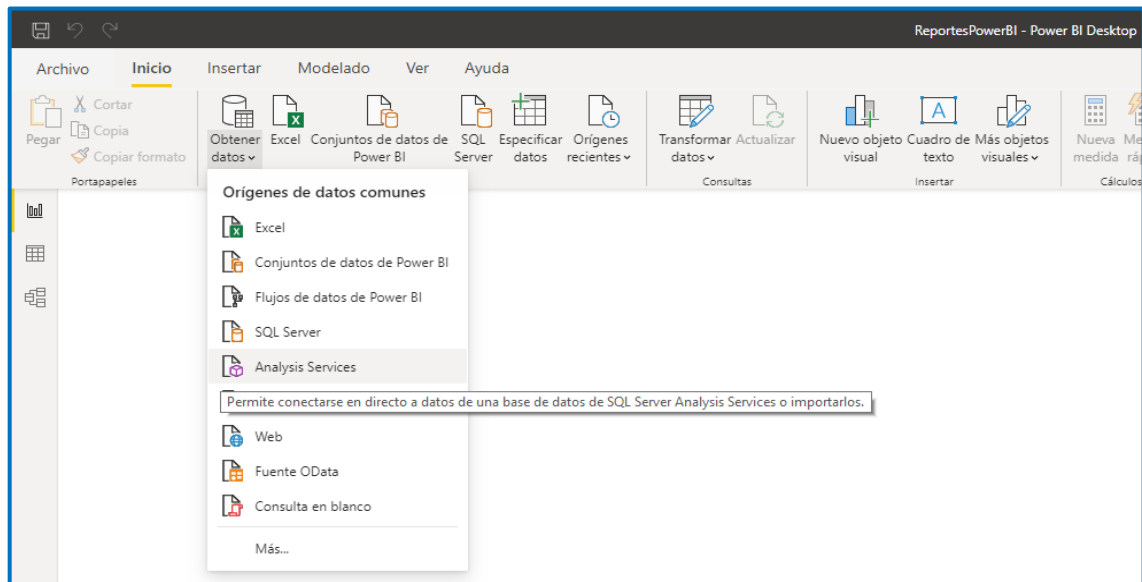


Figura 4.70. Conexión entre Analysis Service y Power BI

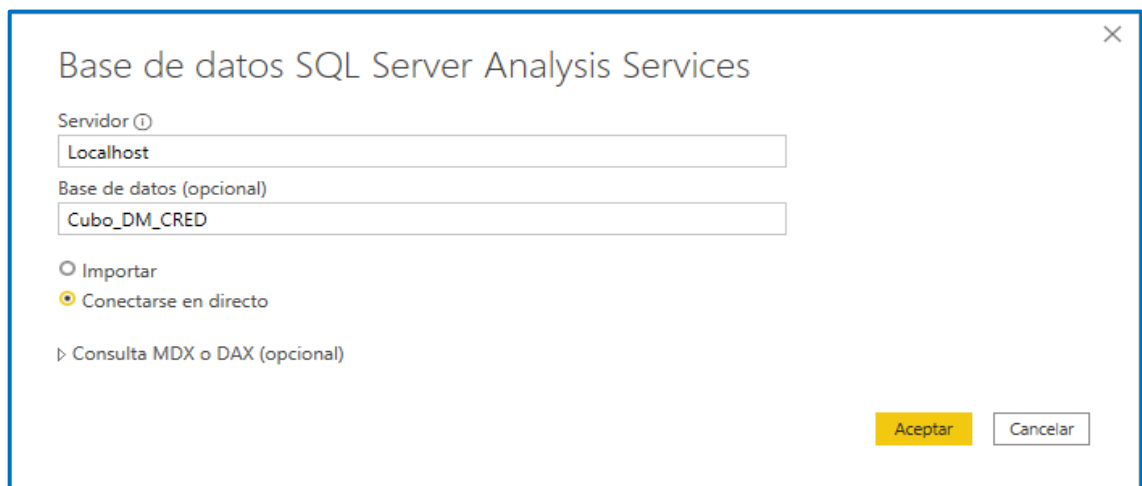


Figura 4.71. Elección del servidor y el cubo de Analysis Service

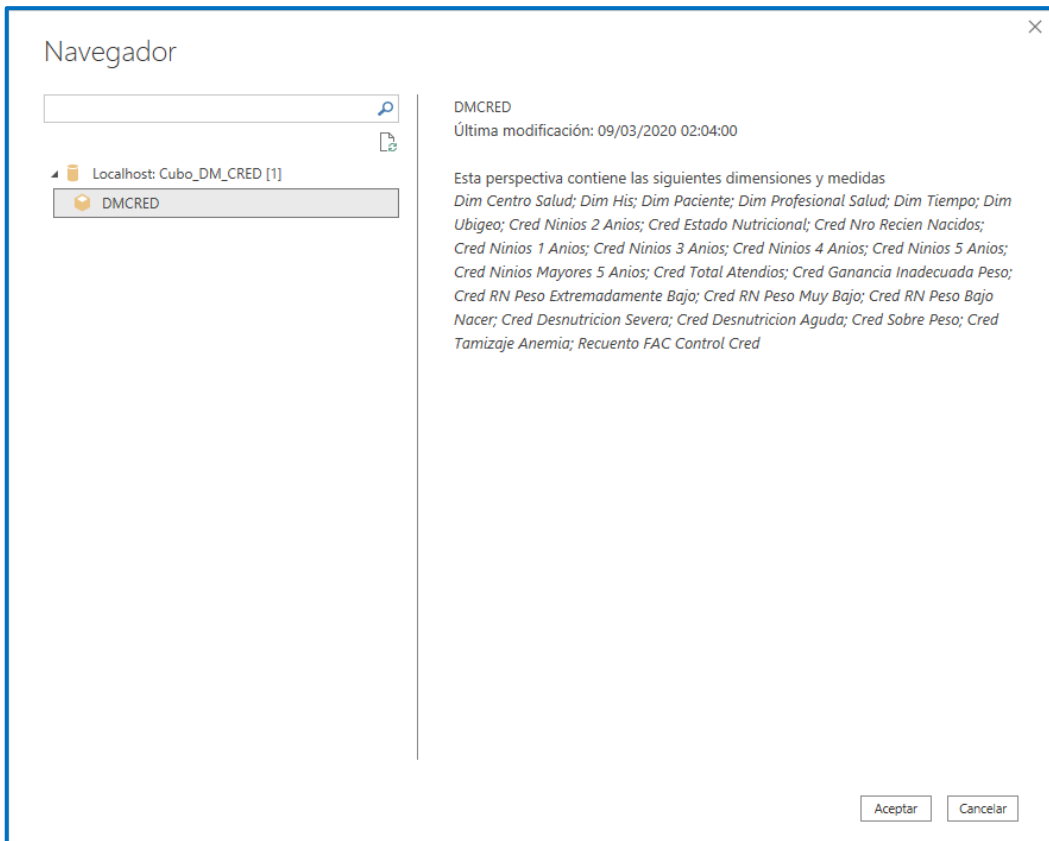


Figura 4.72. Selección de Cubo del Data Mart.

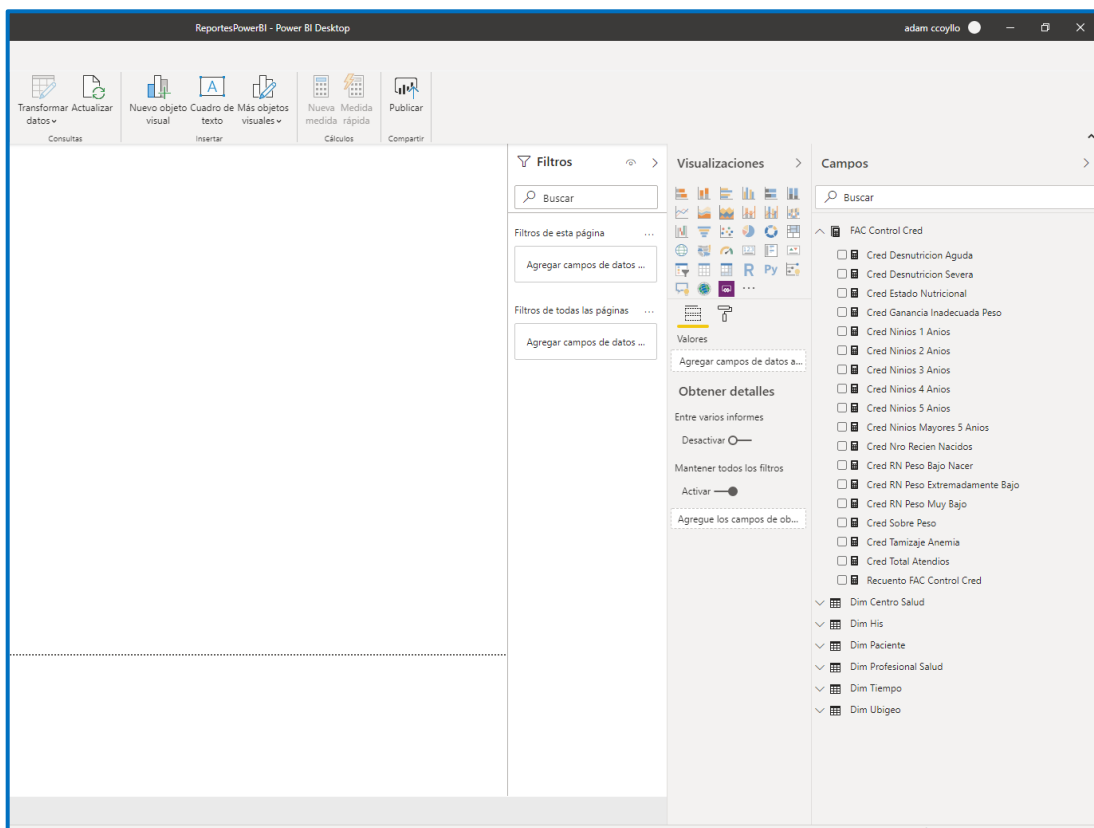


Figura 4.73. Cubo del Data Mart importado al entorno Power BI.

RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN POWER BI POR EL INDICADOR DE CONTROL CRED EN NIÑOS Y NIÑAS POR EL INDICADOR DE ESTADO DE CONTROL DE CRECIMIENTO

Gráfico 1. Control de crecimiento y desarrollo por edad en meses.



Gráfico 2. Control Cred por micro red, centro de Salud, Edades y género.

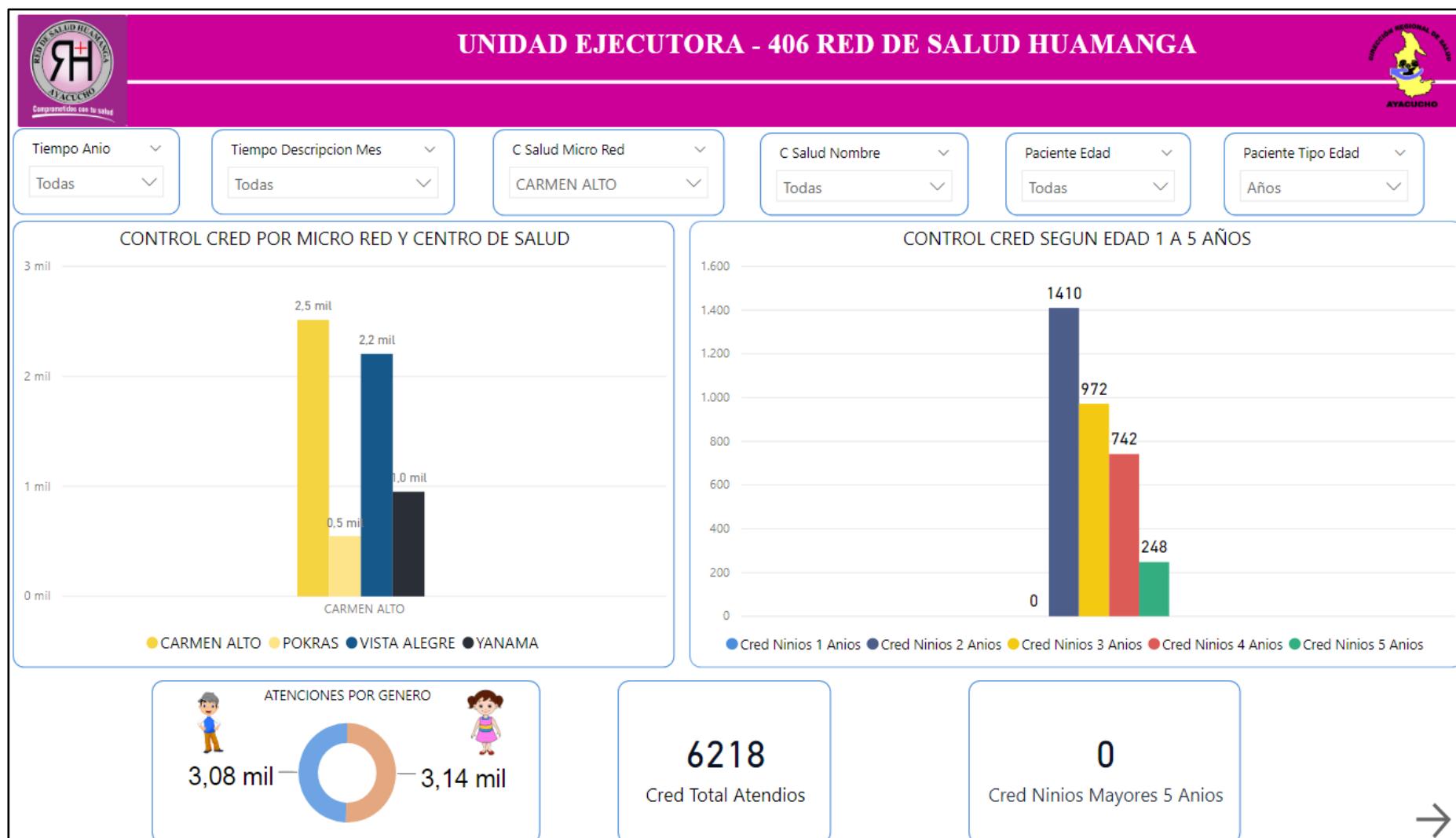
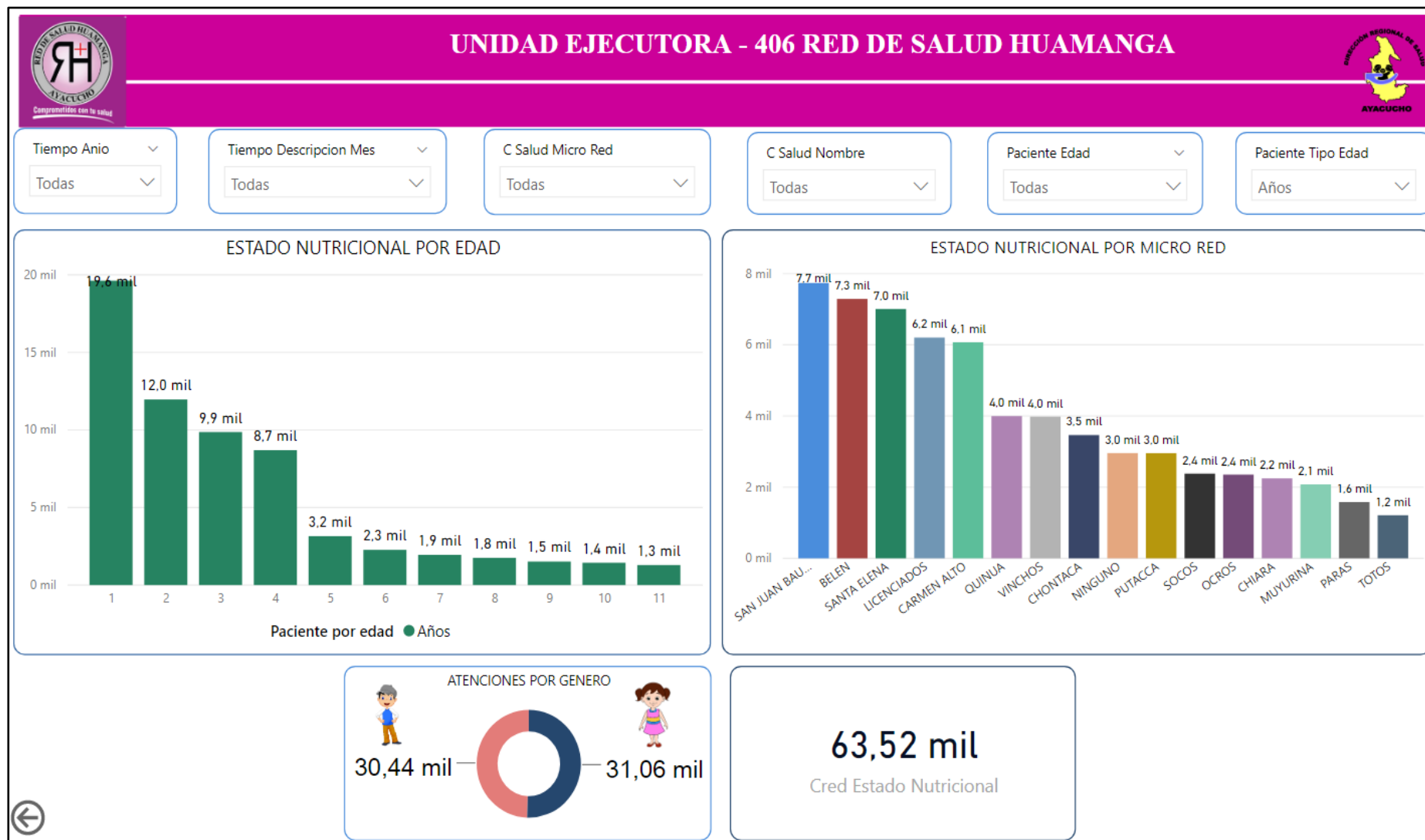


Gráfico 3. Estado nutricional por edad y Micro redes.



RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EL POWER BI POR EL INDICADOR DE ESTADO DE DESARROLLO.

Gráfico 4. Desnutrición aguda y severa por género y micro red.

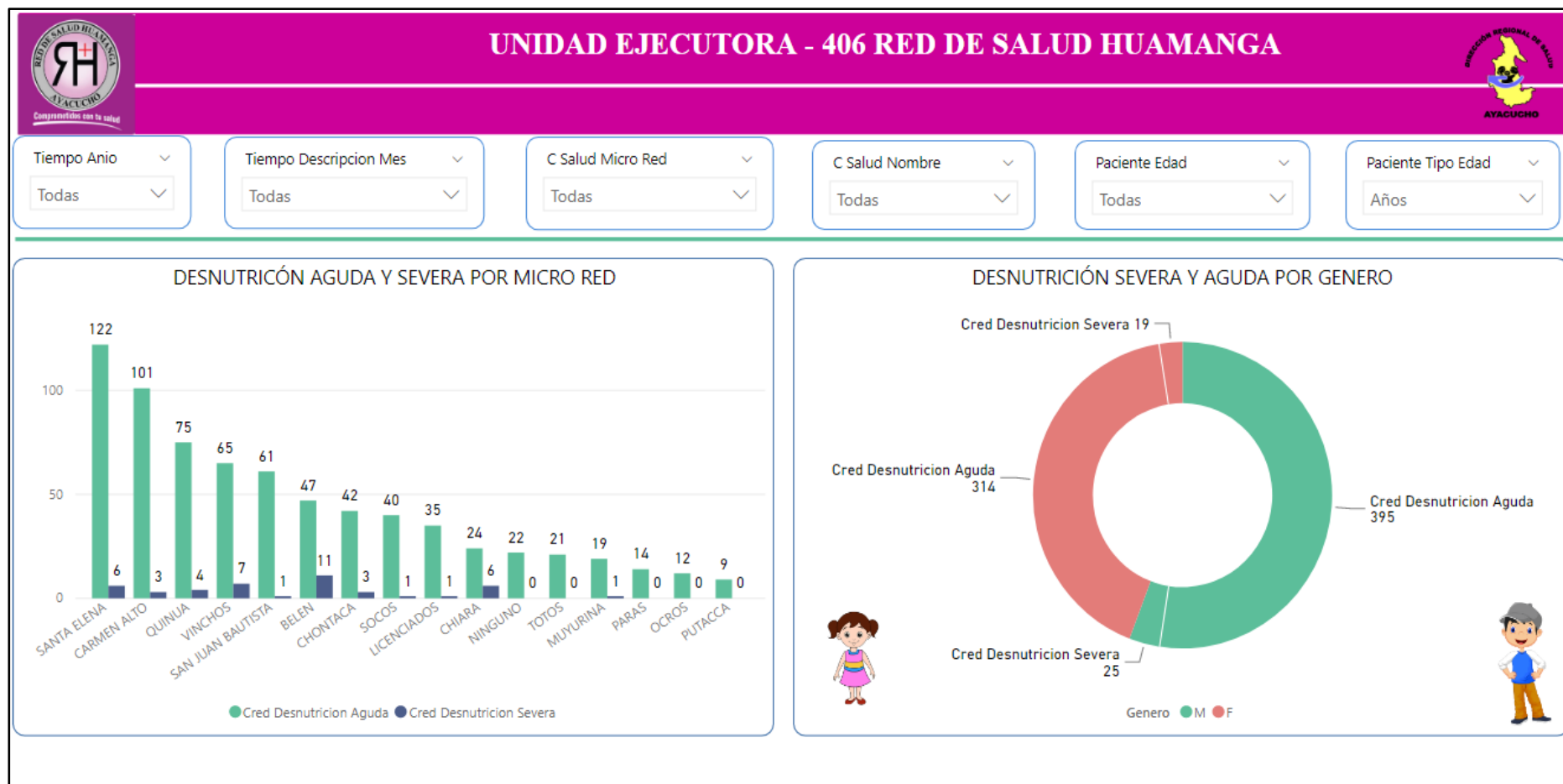


Gráfico 5. Peso bajo al nacer y peso extremadamente bajo en recién nacidos por micro red.

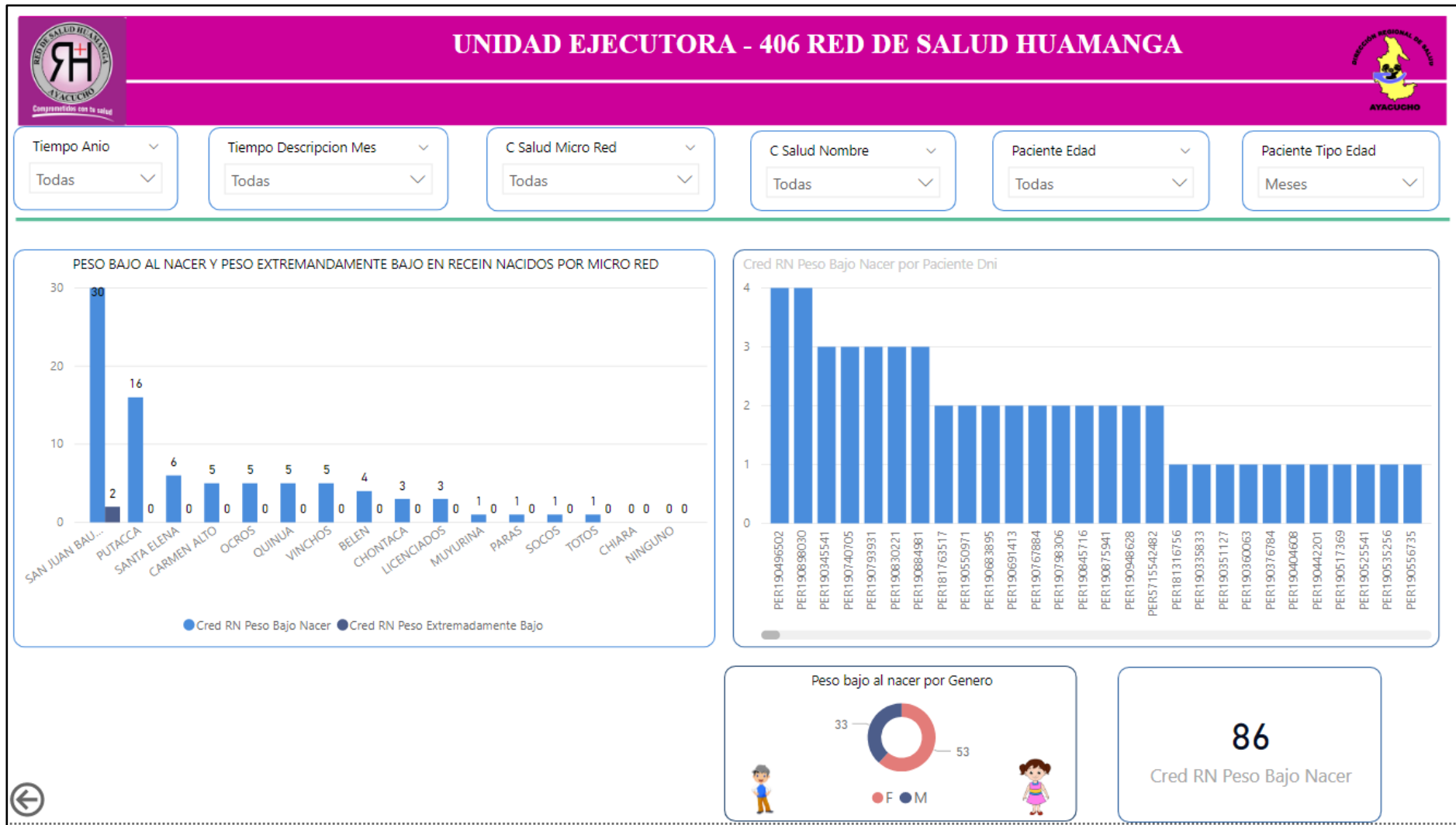
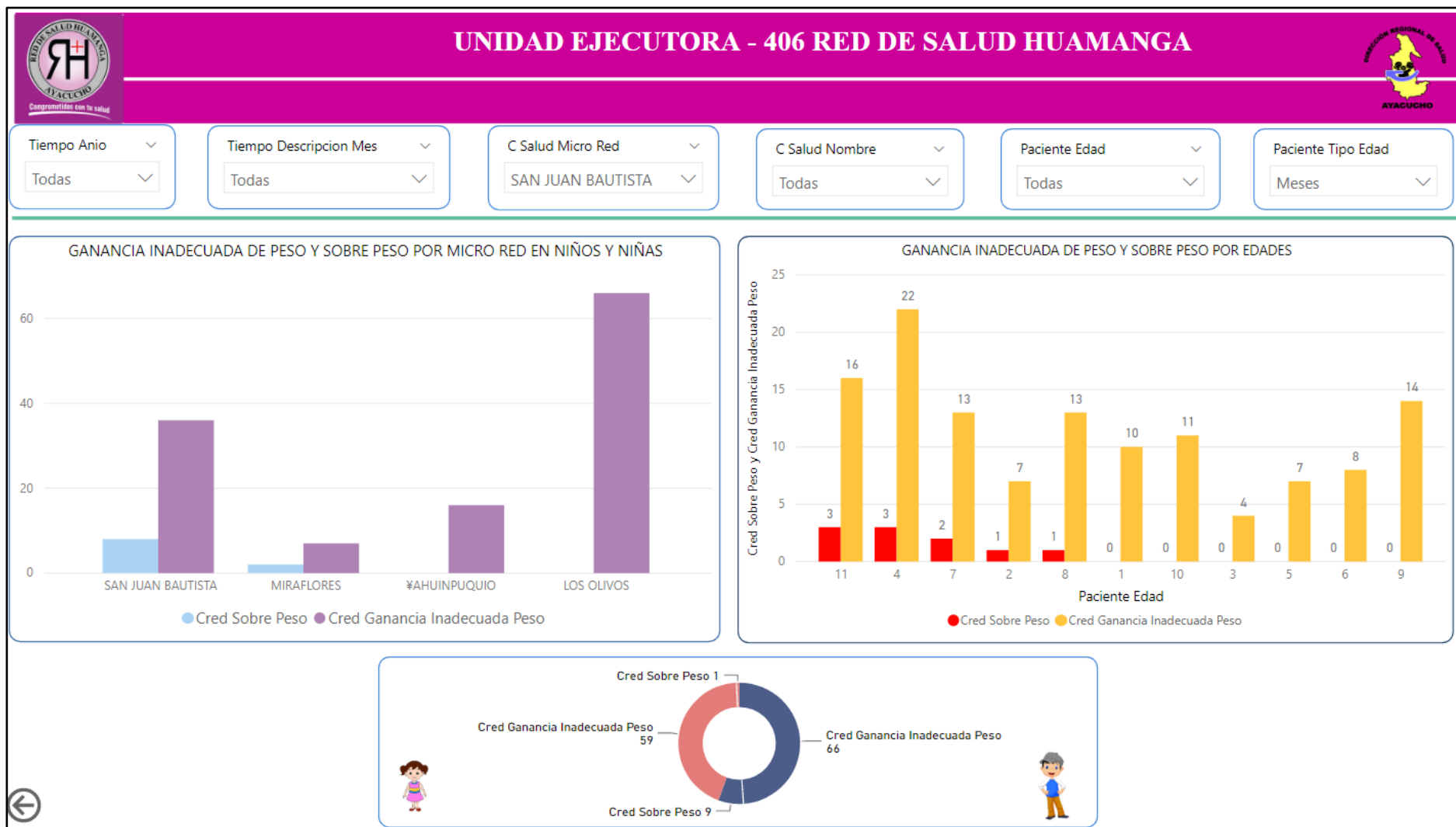


Gráfico 6. Ganancia inadecuada de peso y sobre peso por micro redes y edades.



RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EL POWER BI POR LA VARIABLE DE UNIDAD DE RED DE SALUD DE HUAMANGA.

Gráfico 7. Control Cred por Micro redes y género.

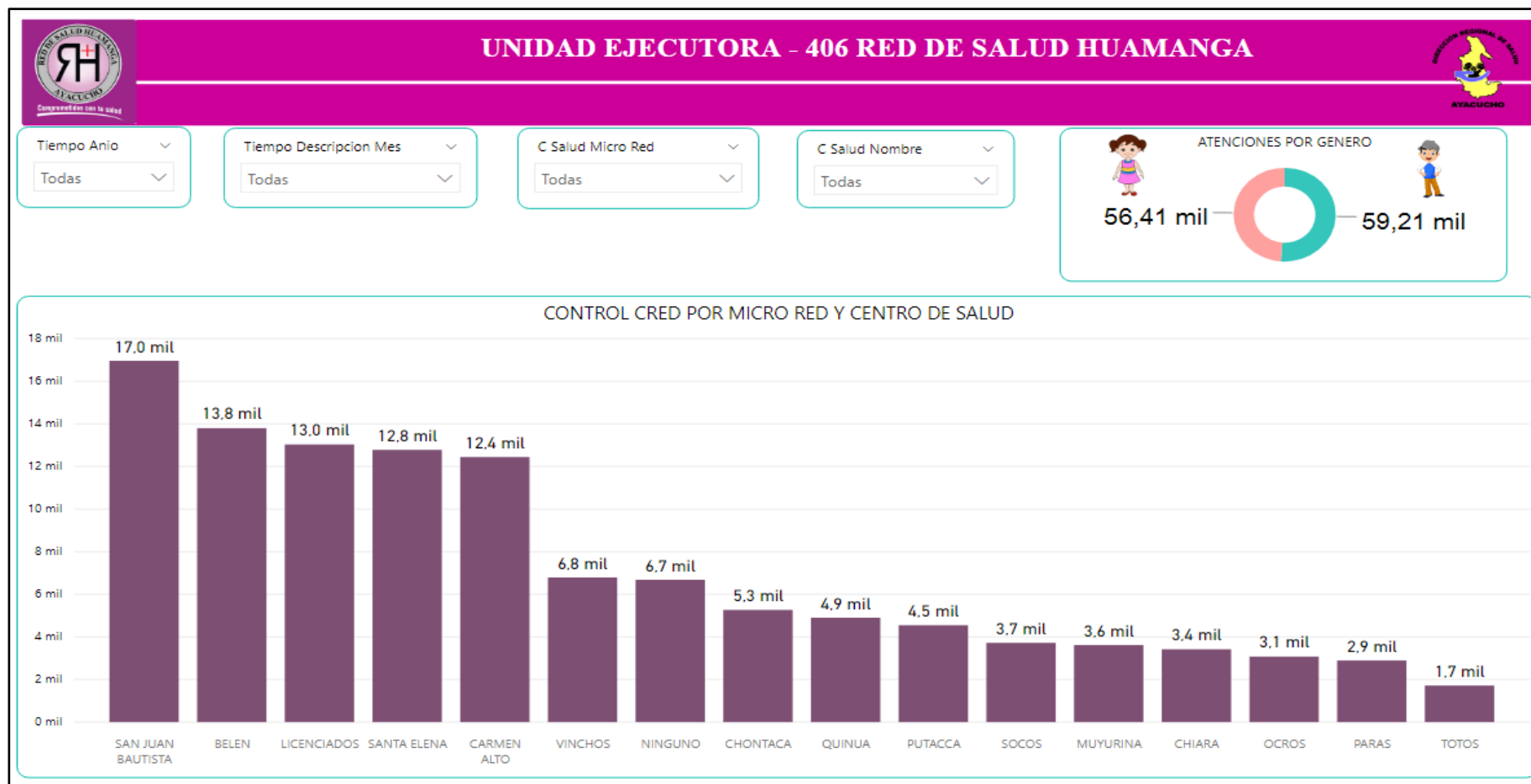
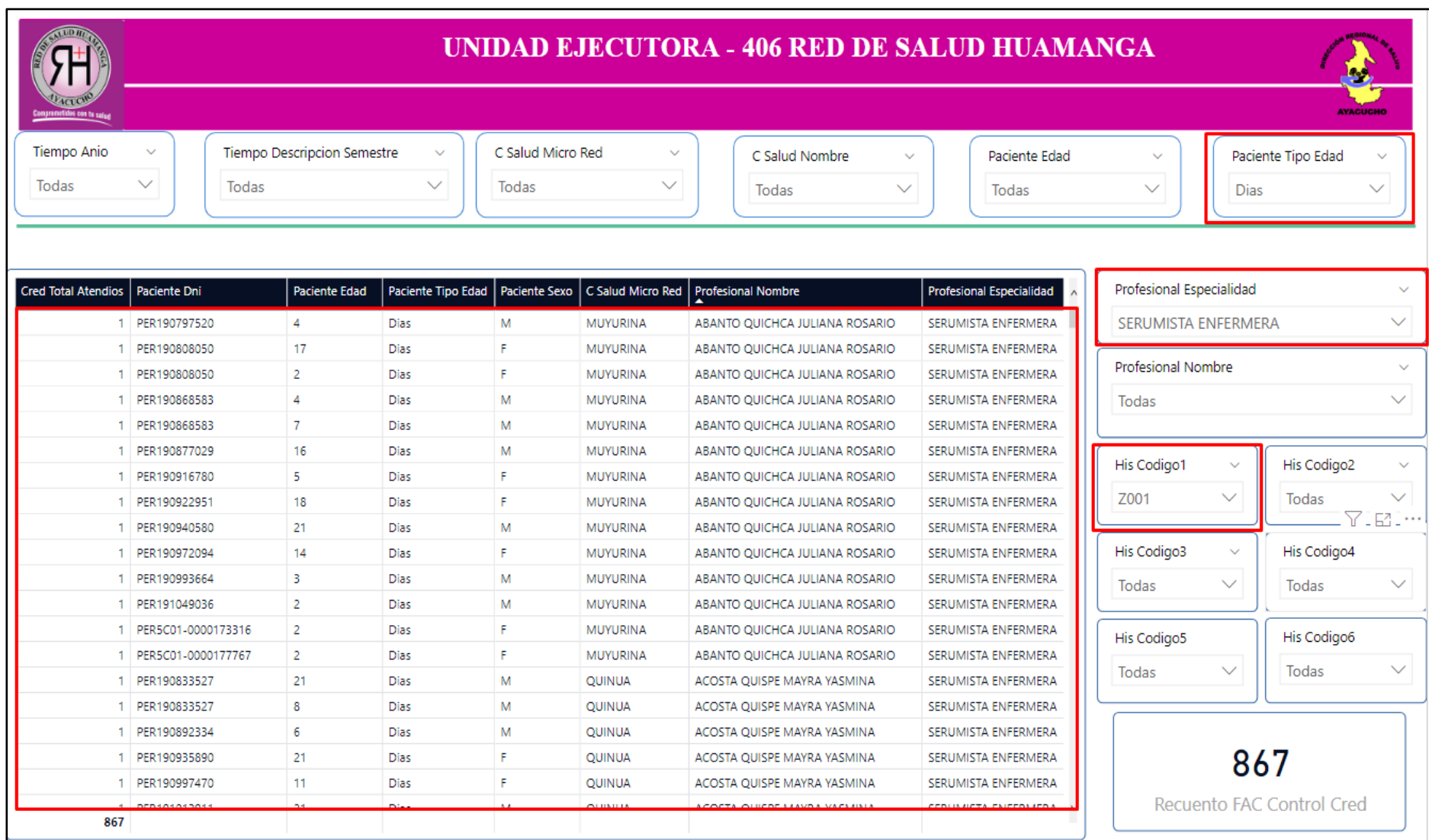


Gráfico 8. Control de crecimiento y desarrollo por especialidad y recién nacidos.



RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EL POWER BI POR EL INDICADOR DE ESTADO DE CONTROL DE CRECIMIENTO Y DESARROLLO

Gráfico 9. Tamizaje de Anemia Control de crecimiento y desarrollo.

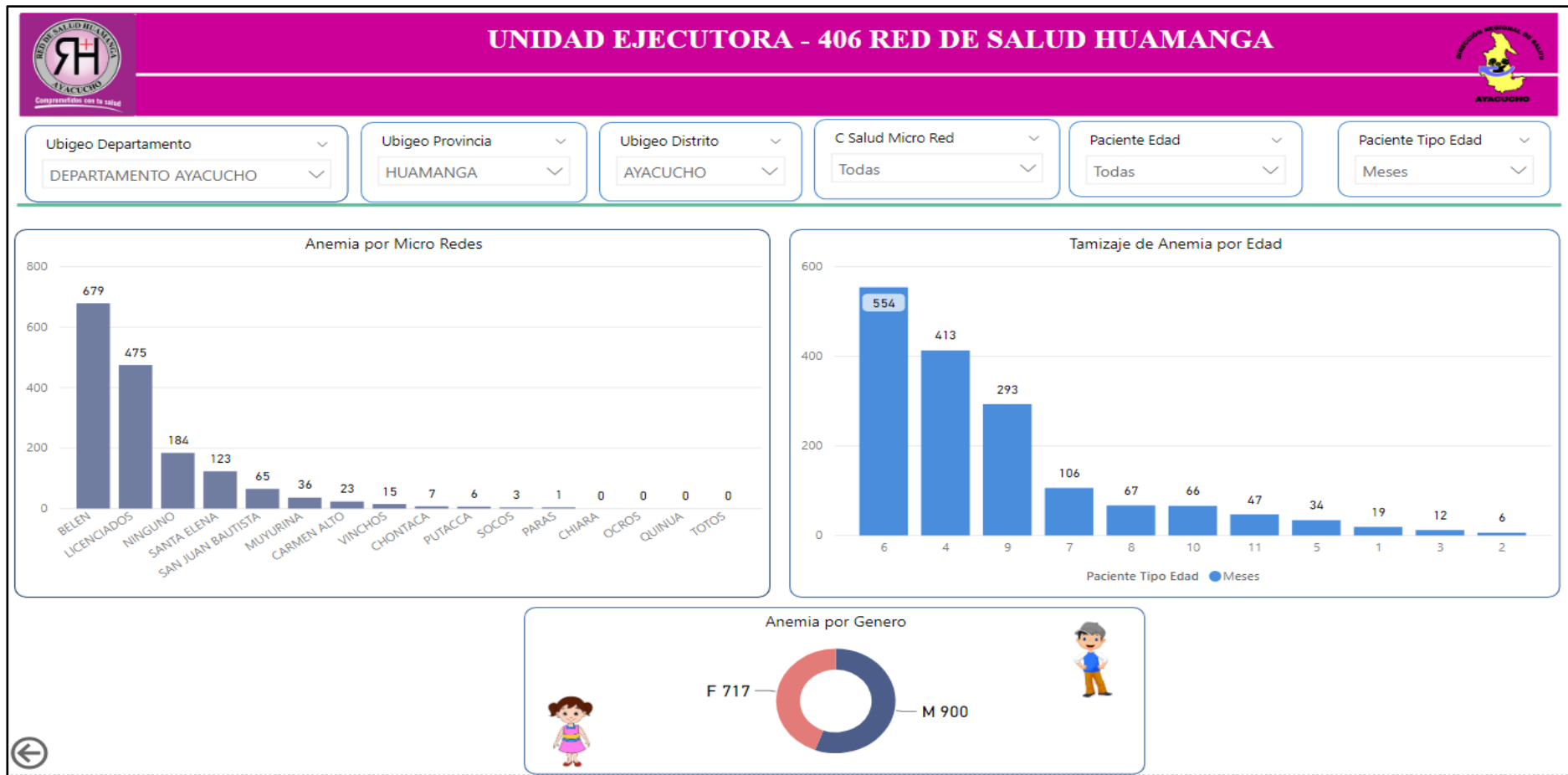


Gráfico 10. Atenciones Control Cred por Centro de Salud, visto desde una página web.

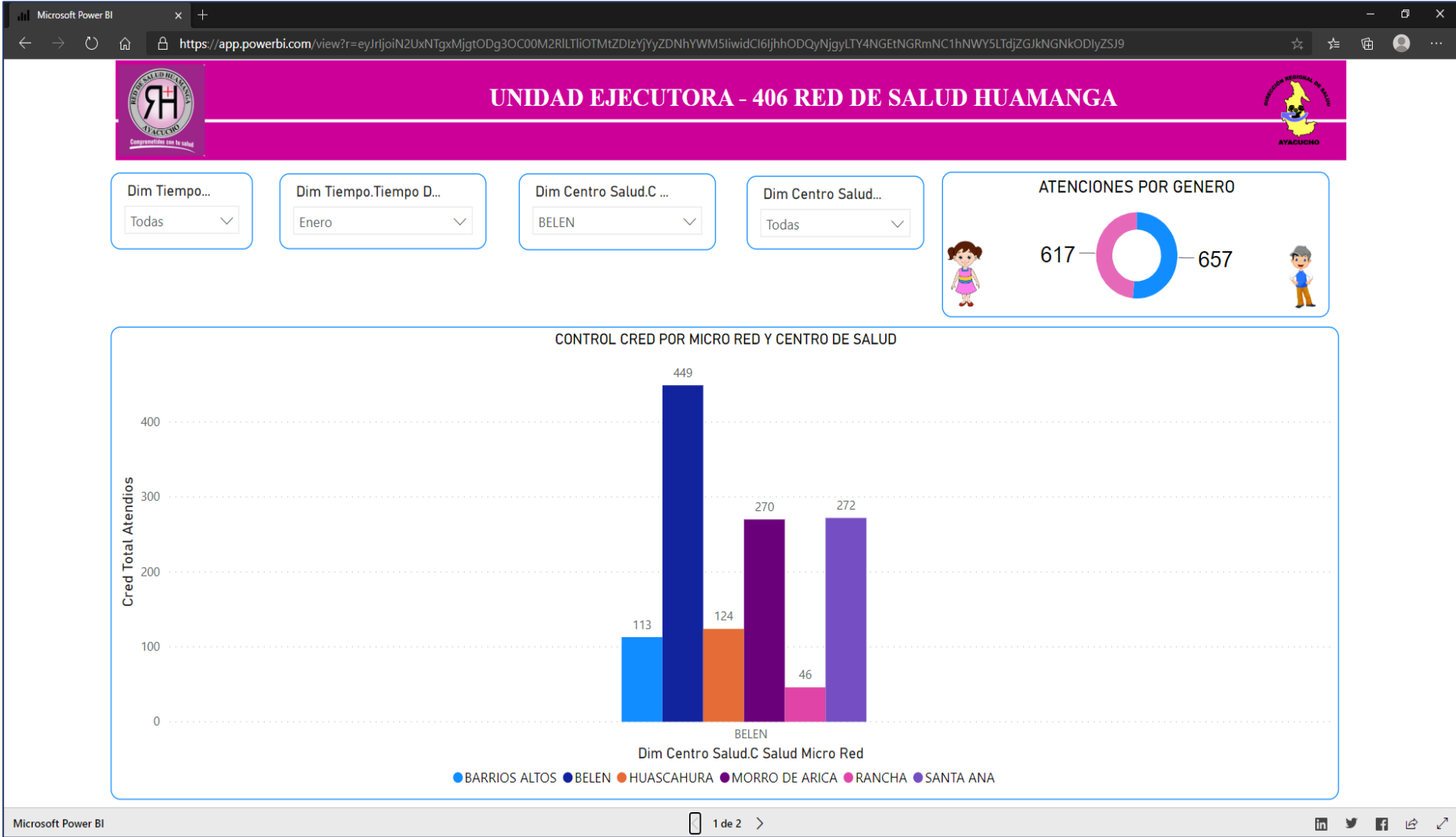


Gráfico 11. Atenciones de control CRED por Micro Redes, visto desde la página web.



4.1.11. MANTENIMIENTO Y CRECIMIENTO

MANTENIMIENTO

Con respecto al mantenimiento, para el Data Mart se vio conveniente realizar backup cada fin de mes, esto debido que los datos se registran en el sistema HisMinsa es diario y cada fin de mes se realiza los reportes necesarios, y este backup se realizara de manera completa (Full).

CRECIMIENTO

El tema de crecimiento del Data Mart se ve como una mejora a su funcionamiento para su mayor explotación del sistema de inteligencia de negocios los cuales podrían ser:

- Incorporación de datos de los años anteriores y posteriores.
- Actualización en el nivel analítico, nuevos hechos de alto nivel.
- Creación de nuevas dimensiones para un mayor análisis, acorde a los requerimientos que podrían surgir en la institución.
- Creación de una nueva tabla de hechos, relacionada a la tabla actual, esto dependiendo de los nuevos reportes que se fueran necesitando a medida que pase el tiempo, esto siendo mayor beneficio para la institución.

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

- a) De acuerdo al marco teórico del capítulo II, sección 2.2.9, se construye el Data Mart de acuerdo a la metodología de Ralph Kimball definida sus técnicas en el capítulo III, sección 3.7, la cual nos permitió lograr los objetivos trazados, la cual se puede mostrar en el capítulo IV, gráficos 4.1. al 4.3, que brinda toda la información del Control de crecimiento y desarrollo
- b) De acuerdo al marco teórico del capítulo II, sección 2.2.9, se construye el Data Mart de acuerdo a la metodología de Ralph Kimball definida sus técnicas en el capítulo III, sección 3.7, la cual nos permitió lograr los objetivos trazados, la cual se puede mostrar en el capítulo IV, gráficos 4.4 al 4.6, que brinda la información del estado de control de desarrollo.
- c) De acuerdo al marco teórico del capítulo II, sección 2.2.9, se construye el Data Mart de acuerdo a la metodología de Ralph Kimball definida sus técnicas en el capítulo III, sección 3.7, la cual nos permitió lograr los objetivos trazados, la cual se puede mostrar en el capítulo IV, gráficos 4.7y 4.8, que brinda la información de acuerdo a las Micro Redes a los Centros de Salud que están afiliadas a la Red de salud Huamanga.
- d) También se puede concluir que en el capítulo IV, Grafico 4.9. muestra la información del tamizaje de anemia, que es una de las enfermedades responsables para el desarrollo adecuado del niño o niña de la región de Ayacucho, que este grafico muestra con mayor numero en los niños de 6 meses de edad.

5.2. RECOMENDACIONES

- a) Se recomienda a futuros estudiantes que tengan interés en el mismo tipo de investigación, a realizar un Data Mart para la anemia, para la misma institución que es la Red de Salud Huamanga, utilizando la metodología de Ralph Kimball, o uno similar que es la metodología de Hefesto, siendo este también una metodología para la construcción de Data Warehouse.
- b) Se debe realizar la investigación del mismo tema de control CRED, a nivel de todas las redes de salud de la región de Ayacucho siendo este de mayor importancia para las tomas de decisiones por parte de las instituciones competentes que es la dirección regional de salud (DIRESA).
- c) Se recomienda mostrar la información del Data Mart Control CRED en la página principal de la Red de Salud de Huamanga. Para dar conocimiento a la población.

BIBLIOGRAFÍA

Ahumado, E. y Perusquia, J. (2014) *Inteligencia de negocios: Estrategia para el desarrollo de competitividad en empresa de base tecnológica*. Universidad Autónoma de Baja California, Mexico.

Alexis Cedeño T. (2005) *Modelo Multidimensional*. Instituto Superior Politécnico José Antonio Echevarría Cujae. La Habana, Cuba.

Arreguin, J. (2006). Bases de Datos SGMA. Recuperado de la página web <https://www.coursehero.com/file/31824061/Historia-BDpdf/>

Acosta, M. (2019). Descubre la principal diferencia entre Data Mart y data warehouse. Recuperado el 24 de noviembre del 2019 de <https://revistadigital.inesem.es/informatica-y-tics/diferencia-entre-data-mart-y-data-warehouse/>

Bernabeu, R. D. (2010), *Metodología Hefesto (5taEd.)*. Córdoba, Argentina: Editorial Tierra del sur.

Bustamante, A. Galvis, E. y Gomez, L. (2013). *Técnicas de modelado de procesos ETL: una revisión de alternativa y su aplicación en un proyecto de desarrollo de soluciones de BI*. Universidad Tecnológica de Pereira. Pereira Colombia.

Billy, J. Ango, H. (2018) *Diseño del proyecto de investigación científica*. Editorial San Marcos. Lima Perú.

Cutro, A. (s.f) *Tendencias de BI Business Intelligence*. Lima, Perú.

Coronel, C. (2015). *Implementación de una solución de inteligencia de negocios en la mesa de servicios Tecnológicos IBM- UTPL*. Tesis de pregrado. Universidad Técnica Particular de Loja. Loja Ecuador.

Conesca, J. y Curto, J. (2011) *Introducción al Business Intelligence*. Edt. El Ciervo 96, S.A. Barcelona, España.

Cobo, A. (s.f.). *Base de datos relacionales: Teoría y práctica* (1ª ed.). Madrid, España: Visión Libros.

Quiroz, J. (2003) El modelo relacional de bases de datos. Recuperado de http://www.doanalytics.net/Documents/Modelo_Relacional.pdf

Camps, R. Casillas, L. Costal, D. (2007). *Bases de Datos*. (2ª ed.). UOC (Formación de posgrado). Barcelona España.

Campos, M. (2018), Factores de incumplimiento del control de crecimiento y desarrollo del niño (a) menor de 1 año en el centro de salud santa rosa de pachacutec ventanilla. *Tesis de pregrado*. Universidad Cesar Vallejo. Lima Perú.

Chávez, D. (2017), *Factores sociales y culturales de las madres de 3 años con inasistencia al control cred centro de salud la flor, Carabayllo*. Tesis de Pregrado. Universidad San Juan Bautista. Lima Perú.

Chaudhuri, S. y Dayal, U (1997). *An overview of data warehousing and OLAP technology*. ACM Sigmond, recuperado de <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/sigrecord.pdf>

Cano, J. (2007) *Business Intelligence: Competir con información*. Escuela Banespyme.

Colaco, M. Turban, A. (2009) *Los impactos de Business Intelligence en la Gestión del Área Comercial*. La Habana Cuba.

Duran, L. (2017). *Implementación de un Data mart para el seguimiento académico de los estudiantes en la escuela académica profesional de ingeniería de sistemas de la Universidad Nacional de Cajamarca*. Tesis de Pregrado. Universidad Nacional de Cajamarca. Cajamarca Perú.

Date, C. (2001). *Introducción a los Sistemas de Bases de Datos*. Séptima Edición. Pearson Educación. México.

Dirección Regional de Salud (2012), Unidades de Gestión Territoriales de Salud, Recuperado de <http://www.Redhuamanga.gob.pe>

Duque, D. y Hernández, J. (2016) *Modelo para el proceso de extracción, transformación y carga en bodegas de datos. una aplicación con datos ambientales* (Tesis de Pregrado) Universidad Militar Nueva Granada. Nueva Granda.

Ferri, R. (1996). *Introducción a los cuidados de Enfermería Infantil*. Alicante España.

Gutiérrez, A. (s. f) *Bases de Datos*. Centro Cultural Itaca S.C.

Gómez, C. (2013) *Bases de Datos* (1ª ed.). México. Publidisa Mexicana S.A.

Guillen, F. (2012) *Desarrollo de un datamart para mejorar la toma de decisiones en el área de tesorería de la municipalidad de Cajamarca*. (Tesis de pregrado) Universidad del Norte. Cajamarca, Perú.

Hernández S, Fernández, C. y Baptista, P. (2014) *Metodología de la investigación* (6ta Ed.). México, D.F., México: McGraw Hill Interamericana.

Henao, D. (2008) *Como Disminuir los Riesgos de los procesos de ETL en el proyecto de inteligencia de negocios en una empresa de transporte*. (Tesis de pregrado) Universidad EAFIT. Medellín Colombia.

Humana, E. y Huingo, C. (2016) Control de crecimiento y desarrollo de menor de un año por el profesional de enfermería en el centro de salud Pachacutec. *Tesis de pregrado*. Universidad Privada Antonio Guillen Urrelo. Cajamarca. Perú.

Inmon, B. (2002), *Building the Data Warehouse* (3ra Ed.). Toronto. Canadá: Wiley

Ibarra, M. (2006). *Procesamiento Analítico en línea Olap*. (Tesis de Pregrado). Universidad Nacional del Norte. Argentina.

Kimball, R. y Ross, M. (2013). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling* (2da edición). Toronto. Canadá: Wiley.

Kimball, R. y Caserte, J. (2004). *The data Warehouse ETL toolkit*, Indianapolis, Wiley. Canada.

Leme, F. (2004) *Business Intillegence no Microsoft Excel*. Rio de Janeiro Brasil.

Laudon, K. y Laudon, J. (2008). *Sistemas de Información Gerencial* (10ma Ed.). México, D.F., México: Pearson Educación.

Lopez, B. (2002) *Guía para la construcción de un Data Warehouse*. (Tesis de Magister) Universidad Autónoma de nuevo León, Mexico.

Ministerio de Salud. (2005). *Manual de procedimientos para el control de crecimiento y desarrollo de la niña y el niño*. Lima.

Ministerio de Salud. (2011). *Norma técnica de salud para el control del crecimiento y desarrollo de la niña y el niño menos de cinco años*. Lima: Biblioteca Nacional del Perú.

Moss, L. y Atre, S. (2003) *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. Boston. Inglaterra: Addison Wesley.

Morales, S. (2019). *Metodología para procesos de inteligencia de negocios con mejoras en la extracción y transformación de fuentes de datos, orientado a la toma de decisiones*. (Tesis de Doctorado) Universidad de Alicante. España.

Muñoz, L. y Mazon, J. (s.f) *ETL Modeling Conceptual for Data Werehouse: A Systematic Mapping Study*. Universidad Tecnológica de Panamá. Panamá.

Nevado, V. (s.f.). *Introducción a las bases de datos relacionales*. Madrid, España: Visión Libros.

Nader, J. (2002). *Sistema de Apoyo gerencial Universitario*. Tesis de postgrado. Instituto Tecnológico de Buenos Aires. Buenos Aires. Argentina.

Oppel, A. y Sheldon, R. (2009). *Fundamentos de SQL* (3ª ed.) McGraw – Hill/ Interamericana editores S.A. Impreso en Mexico.

Oracle (2016) Business Intelligence, Stándard Edition one Tutorial.

Organismo Mundial de la Salud (s.f). Glosario de términos.
<http://www.fao.org/3/am401s/am401s07.pdf>

Pacco, R. (2013). *Sistema de gestión financiera basado en sistemas de información ejecutiva y modelo kimball para vicerrectorado académica de la universidad peruana unión*. (Tesis de Pregrado). Universidad Peruana Unión. Lima Perú.

Rojas, G. (2018). *Data mart para información táctica de prestación de salud en adolescentes, región Ayacucho*. Tesis de pregrado. Universidad Nacional de San Cristóbal de Huamanga. Ayacucho Perú.

Red de Salud Huamanga (2012) *Manual de Organización y funciones (Mof)* Red de Salud de Huamanga. Ayacucho, Perú.

Rotaeché, C. (2007) Business Intelligence, Ibermatica.

Rojas, D. y Zamudio, L. (2016), *Implementación de inteligencia de negocios, utilizando la metodología de Ralph kimball, en la toma de decisiones en el área de ventas. empresa sid sac*. (Tesis de Pregrado), Lima Perú.

Rivadera, G. (s.f) *Metodología Kimball para el diseño de almacenes de datos (Data Warehouse)* Lima, Perú.

Robles, B. y Girón, L. (2015) *Desarrollo de modelos, aplicaciones y arquitecturas de sistemas para toma de decisiones y dirección estratégica* (Tesis pregrado) Universidad Privada Antenor Orrego. Trujillo, Perú.

Ramos, S. (2016) *Data Warehouse, Datamarts y Modelos Dimensionales, un pilar fundamental para la toma de decisiones*. Editorial Soliq Global SA. España.

Rojas, S. (2014) *Implementación de un Datamart como solución de Inteligencia de Negocios, Bajo la Metodología Kimball para optimizar la toma de decisiones en el departamento de finanzas de contraloría general de la república* (Tesis de pregrado) Universidad San Martín de Porres. Lima, Perú.

Silva, S. (2017). *Análisis, diseño e implementación de un Datamart que garantice una adecuada toma de decisiones en el área de ventas en la empresa promed EIRL*. (Tesis de Pregrado). Universidad Peruana de las Américas. Lima Perú.

Silberschatz, A. y Korth, H. (2006). *Fundamentos de bases de datos* (5ª ed.) McGraw – Hill/ Interamericana de España, S.A.U. Madrid España.

Sánchez, J. (2012). *Instalación y Configuración de Bases de Datos usando Oracle* 11g. Recuperado de <http://creativecommons.org/licenses/by-nc-sa/3.0/deed.es>

Sullcarayme, R. y Mamani A. (2017). *Implementación de business intelligence, utilizando la metodología de ralph kimball, para el proceso de toma de decisiones de las compras en la empresa edipesa s.a* (Tesis de pregrado) Universidad Autónoma del Perú. Lima Perú

Tamayo y Tamayo, M. (1997) *El Proceso de la Investigación científica*. México D.F., México: Editorial Limusa S.A.

Tamayo, M. y Moreno, F. (2006) *Análisis del modelo de almacenamiento MOLAP frente al modelo de almacenamiento ROLAP*. Ingeniería e Investigación. Recuperado el 30 de Setiembre del 2018, de <http://www.redalyc.org/html/643/64326317/>.

Torres, L. (2007). *Business Intelligence*. Recuperado de la página web <http://www.gravitar.biz/index.php/bi/bi-terminologia-1>

Tapia, I. (2006). *Una metodología para sectorizar pacientes en el consumo de medicamentos aplicando Datamart y Datamining en un hospital nacional*. (Tesis de Pregrado). Universidad Nacional Mayor de San Marcos. Lima Perú.

Tana, R. (2014). *Datamart para el análisis de información del sistema académico de la universidad técnica del norte con herramienta de software*. Universidad del Norte. Ecuador.

Torres, A. (2002). *Revista Mexicana de Medicina y Rehabilitación - Crecimiento y Desarrollo*. México.

Ullman, J.D. y Widom, J. (1999). *Introducción a los Sistemas de Bases de Datos*. México D.F, México: Editorial PPH.

Vizuite, M. y Yela, C. (2006). *Análisis, diseño e implementación de un Data mart para el área de sismología del departamento de geofísica de la escuela politécnica nacional*. (Tesis de Pregrado). Escuela politécnica Nacional. Quito Ecuador.

Vitt, E. y Luckevich, M. (2002). *Business Intelligence Técnicas de análisis para la toma de decisiones estratégicas* (1era Ed.). Madrid. España: McGraw-Hill/Interamericana.

Wolf, C. (s.f) *Modelamiento Multidimensional*. Universidad de Concepción, Chile.

Wiley. J. (2009). *Business Intelligence Data Mining and Optimization for Decision Making*. British Library. India

.

ANEXO A. OPERACIONALIZACIÓN DE VARIABLES

OPERACIONALIZACIÓN DE VARIABLES			
VARIABLE	VARIABLE DESCRIPTIVA	INDICADOR	ÍTEM
X: Estado Nutricional	Estado de Crecimiento	Atenciones CRED	¿Qué cantidad de niños y niñas llevan el control de crecimiento?
			¿Qué cantidad de niños y niñas cumplieron sus atenciones según su edad?
	Estado de Desarrollo	Desnutrición	¿Qué cantidad de desnutrición existe según el género?
			Nutrición
		Anemia	
			Recién Nacidos
Y: Red de salud Huamanga	Unidad de estadística, informática y telecomunicaciones.	Establecimientos de Salud	¿Qué cantidad de atenciones de control CRED se realizó en los diferentes establecimientos de salud?
			¿Qué cantidad de atenciones de control CRED en niños y niñas en los diferentes
	Micro redes	Micro Redes?	
		¿Qué personal de salud realizó las atenciones CRED y de que establecimientos?	

ANEXO B: FICHA PARA ANÁLISIS DE LA BASE DE DATOS

FICHA N°1:

¿Cuántos niños y niñas fueron atendidos durante el año, mayores 1 año y menores de 12 años?

SELECT * FROM [dbo].[His2019] where edad>1 AND edad<112

FICHA N°2:

¿Cuántos niños y niñas fueron atendidas durante el año, menores de un año?

SELECT * FROM [dbo].[His2019] where edad>50 AND edad<62

FICHA N°3:

¿Cuántas atenciones de control cred se realizó durante el año?

SELECT * FROM [dbo].[His2019] WHERE codigo1='Z001' OR codigo2='Z001' OR codigo3='Z001' OR codigo4='Z001'OR codigo5='Z001' OR codigo6='Z001' OR codigo2='Z001'

FICHA N°4:

¿Cuántas atenciones de estado nutricional se realizó durante el año?

SELECT * FROM [dbo].[His2019] WHERE codigo1='Z006' OR codigo2='Z006' OR codigo3='Z006' OR codigo4='Z006'OR codigo5='Z006 'OR codigo6='Z006' OR codigo2='Z006'

FICHA N°5:

¿Cantidad de atenciones de control de crecimiento y desarrollo?

Dimensión	Jerarquía	Operador	Expresión de filtro
Dim His	His Codigo1	Contiene	Z001
<Seleccionar dimensión>			
Cred Total Atendios			
115580			

FICHA N°6:

¿Cantidad de atenciones de control de crecimiento por meses?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
Dim Paciente	Paciente Edad	Igual	{All}	<input type="checkbox"/>
<Seleccionar dimensión>				

Tiempo	Descripcion Mes	Cred Total Atendios
	Abril	9678
	Agosto	10222
	Diciembre	4236
	Enero	10214
	Febrero	9902
	Julio	10081
	Junio	10191
	Marzo	10244
	Mayo	10696
	Noviembre	9812
	Octubre	10453
	Setiembre	9890

FICHA N°7:

¿Atención de niños y niñas, mayores de 1 año y menores de 12 años, por Micro Redes?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
Dim Centro Salud	C Salud Micro Red	Igual	{All}	<input type="checkbox"/>
<Seleccionar dimensión>				

Cred Niños 1 Anios	Cred Niños 2 Anios	Cred Niños 3 Anios	Cred Niños 4 Anios	Cred Niños 5 Anios	Cred Niños Mayores 5 Anios
41743	12482	10036	8635	2486	0

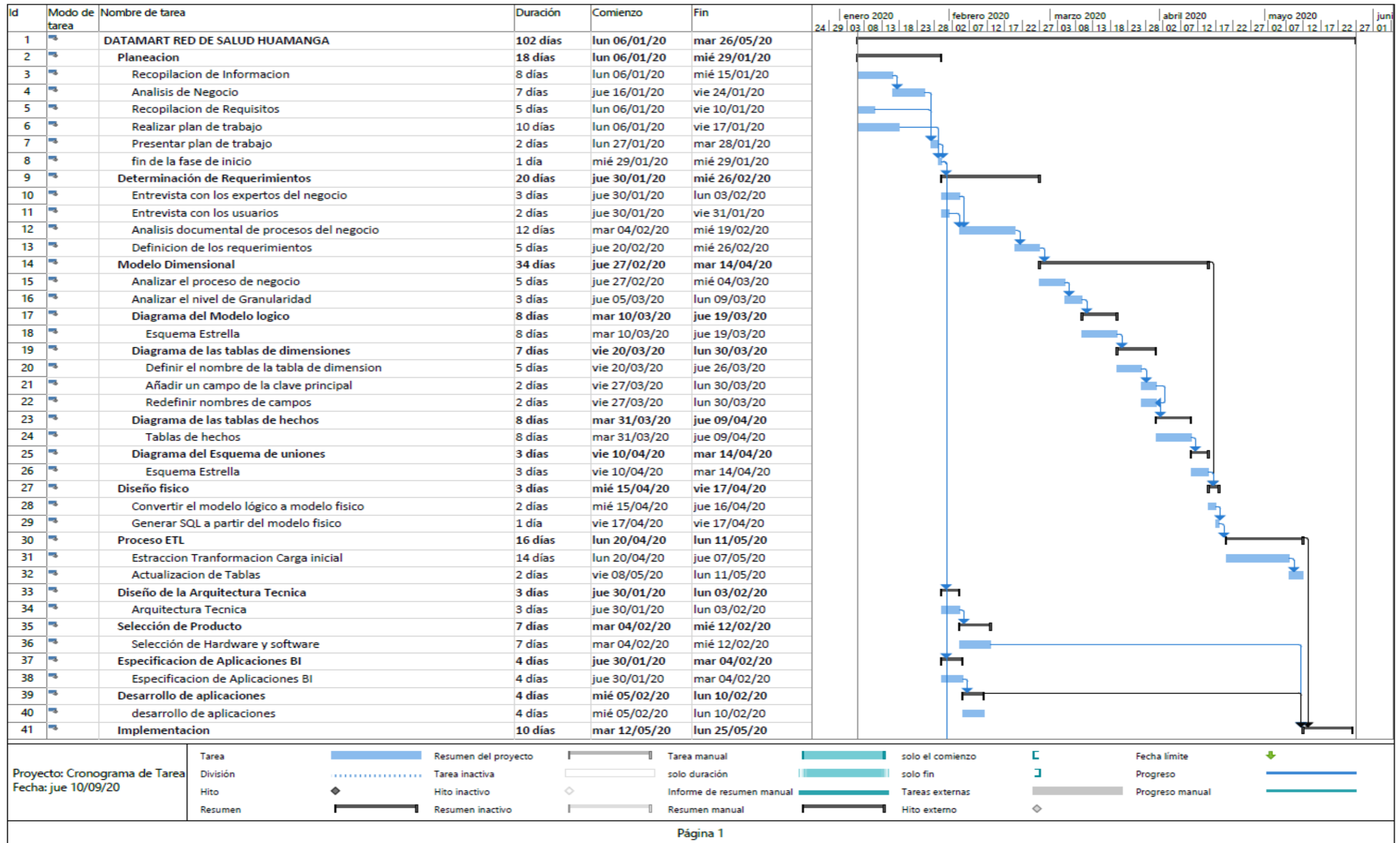
FICHA N°8:

¿Atención de niños y niñas por estado de control de desarrollo por edades?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
Dim His	His Codigo2	Igual	{All}	<input type="checkbox"/>
<Seleccionar dimensión>				

C Salud Nombre	Cred Tamizaje Anemia
YAHUJINPUQUIO	312
ACOSVINCHOS	665
YALPALLACCTA	95
ACOCRO	335
ALLPACHACA	121
ARIZONA	402
ATACOCHA	68
BARRIOS ALTOS	506
BELEN	1342
BELLAVISTA	73
CARMEN ALTO	993
CATALINAYOC	367
CCACCCAMARCA	167

ANEXO C. PLAN DE TRABAJO DE DATA MART CONTROL CRED



Id	Modo de tarea	Nombre de tarea	Duración	Comienzo	Fin	Gantt Chart						
						enero 2020	febrero 2020	marzo 2020	abril 2020	mayo 2020	junio 2020	
42		creacion del Cubo	8 días	mar 12/05/20	jue 21/05/20							
43		Prueba de reporte de requerimientos	2 días	mié 20/05/20	jue 21/05/20							
44		cierre de proyecto	2 días	vie 22/05/20	lun 25/05/20							
45		Administracion del Proyecto	84 días	jue 30/01/20	mar 26/05/20							
46		Administracion del Proyecto	84 días	jue 30/01/20	mar 26/05/20							
47		Mantenimiento y crecimiento	1 día	mar 26/05/20	mar 26/05/20							
48		Mantenimiento	1 día	mar 26/05/20	mar 26/05/20							
49		crecimiento	1 día	mar 26/05/20	mar 26/05/20							

Proyecto: Cronograma de Tarea
 Fecha: jue 10/09/20

Tarea		Resumen del proyecto		Tarea manual		solo el comienzo		Fecha límite	
División		Tarea inactiva		solo duración		solo fin		Progreso	
Hito		Hito inactivo		Informe de resumen manual		Tareas externas		Progreso manual	
Resumen		Resumen inactivo		Resumen manual		Hito externo			

ANEXO D. ENTREVISTA Y CUESTIONARIO

CUESTIONARIO PARA ENTREVISTAR AL JEFE DE LA UNIDAD

1. LAS RESPONSABILIDADES

- Describa la responsabilidad principal de la unidad.
- Describa la relación de su unidad con el resto de las unidades

2. OBJETIVOS DE LA UNIDAD

- ¿Cuál es el objetivo de la unidad?
- ¿Cuáles son sus factores de riesgo?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy dentro de sus funciones?
- ¿con que frecuencia usted mide los factores de éxito de su unidad?
- Análisis de requisitos
- ¿Qué información usa usted actualmente?
- ¿Qué datos son importantes actualmente para su unidad?
- ¿Cuánta información maneja actualmente, de manera semanal y mensual?
- ¿Con respecto a los niños y niñas que información maneja?
- ¿Cuánta información mantiene actualmente en el tema de atenciones?
- ¿Las otras gerencias que información necesitan actualmente?

**UNSCH**FACULTAD DE
INGENIERÍA
DE MINAS, GEOLOGÍA Y CIVIL**“Año de la Universalización de la Salud”****ACTA DE SUSTENTACIÓN DE TESIS N° 034-2020-FIMGC**

En la ciudad de Ayacucho, en cumplimiento a la **Resolución Decanal N° 240-2020-FIMGC-D**, siendo los trece días del mes de octubre del 2020, a horas 11.00 a.m.; se reunieron los jurados del acto de sustentación, en el Auditorium virtual google meet del Campus Universitario de la Universidad Nacional de San Cristóbal de Huamanga.

Siendo el Jurado de la sustentación de tesis compuesto por el Presidente el, **Dr. Ing. Efraín Elías PORRAS FLORES**, Jurado la **Mg. Ing. Celia E. MARTÍNEZ CÓRDOVA**, Jurado – Asesor el **Mg. Ing. Hubner JANAMPA PATILLA**, y Secretario del proceso **Ing. Christian LEZAMA CUELLAR**, con el objetivo de recepcionar la sustentación de la tesis denominada **“DATA MART PARA EL ESTADO DE CONTROL DE CRECIMIENTO Y DESARROLLO PARA EL ESTADO NUTRICIONAL DE NIÑOS Y NIÑAS DE LA RED DE SALUD HUAMANGA, AYACUCHO 2019”**, sustentado por el Bach. **Adan CCOYLLO AGUILAR**, bachiller en Ingeniería de Sistemas.

El Jurado luego de haber recepcionado la sustentación de la tesis y realizado las preguntas, el sustentante al haber dado respuesta a las preguntas, y el Jurado haber deliberado; califica con la nota aprobatoria de **14 (catorce)**.

En fe de lo cual, se firma la presente acta, por los miembros integrantes del proceso de sustentación.

Firmado digitalmente por Dr.
Ing. Efraín Elías Porras Flores
Fecha: 2020.10.21 10:50:11
-05'00'**Dr. Ing. Efraín Elías PORRAS FLORES**
Presidente**Mg. Ing. Celia E. MARTÍNEZ CÓRDOVA**
Jurado**Mg. Ing. Hubner JANAMPA PATILLA**
Jurado - Asesor**Ing. Christian LEZAMA CUELLAR**
Secretario del Procesoc.c.:
Bach. Adan CCOYLLO AGUILAR
Jurados (4)
Archivo



UNSCH

FACULTAD DE
INGENIERÍA
DE MINAS, GEOLOGÍA Y CIVIL

CONSTANCIA DE ORIGINALIDAD DE TRABAJO DE INVESTIGACIÓN

El que suscribe; responsable verificador de originalidad de trabajos de tesis de pregrado en segunda instancia para las Escuelas Profesionales de la Facultad de Ingeniería de Minas, Geología y Civil; en cumplimiento a la Resolución de Consejo Universitario N° 039-2021-UNSCH-CU, Reglamento de Originalidad de Trabajos de Investigación de la UNSCH y Resolución Decanal N° 158-2021-FIMGC-UNSCH-D, deja constancia que:

- Apellidos y Nombres del Bach. : Ccoyllo Aguilar Adan
- Escuela Profesional : Ingeniería De Sistemas
- Título de la Tesis : Data MART para el estado de control de crecimiento y desarrollo para el estado nutricional de niños y niñas de la Red de Salud Huamanga, Ayacucho 2019.
- Evaluación de la originalidad : 9 % de similitud

Por tanto, según los artículos 12, 13 y 17 del Reglamento de Originalidad de Trabajos de Investigación, **es procedente otorgar la constancia de originalidad** para los fines que crea conveniente.

Ayacucho, 05 de junio del 2021

Firmado digitalmente por Mg. Ing.
Johnny Henry Ccatamayo Barrios
Fecha: 2021.06.05 05:31:37 -05'00'

Mg. Ing. Ccatamayo Barrios Johnny Henry
Verificador de originalidad de trabajos de tesis de pregrado de la FIMGC

Numero de constancia: 041-2021-FIMGC.

DATA MART PARA EL ESTADO
DE CONTROL DE CRECIMIENTO
Y DESARROLLO PARA EL
ESTADO NUTRICIONAL DE
NIÑOS Y NIÑAS DE LA RED DE
SALUD HUAMANGA,
AYACUCHO 2019

por Adan Ccoyllo Aguilar

Fecha de entrega: 03-jun-2021 07:16p.m. (UTC-0500)

Identificador de la entrega: 1599972204

Nombre del archivo: TESIS_CCOYLLO_AGUILAR_ADAN.docx (8.68M)

Total de palabras: 15723

Total de caracteres: 84914

DATA MART PARA EL ESTADO DE CONTROL DE CRECIMIENTO Y DESARROLLO PARA EL ESTADO NUTRICIONAL DE NIÑOS Y NIÑAS DE LA RED DE SALUD HUAMANGA, AYACUCHO 2019

INFORME DE ORIGINALIDAD

9%

INDICE DE SIMILITUD

9%

FUENTES DE INTERNET

0%

PUBLICACIONES

4%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1

repositorio.upao.edu.pe

Fuente de Internet

3%

2

repositorio.unsch.edu.pe

Fuente de Internet

3%

3

Submitted to Universidad Nacional de San Cristóbal de Huamanga

Trabajo del estudiante

2%

4

cybertesis.unmsm.edu.pe

Fuente de Internet

<1%

5

repositorio.ucv.edu.pe

Fuente de Internet

<1%

6

repositorio.uladech.edu.pe

Fuente de Internet

<1%

7

repositorio.autonoma.edu.pe

Fuente de Internet

<1%

8

repositorio.udl.edu.pe

Fuente de Internet

<1%

Excluir citas

Activo

Excluir coincidencias < 30 words

Excluir bibliografía

Activo