



INSTITUTO TECNOLÓGICO SUPERIOR DE XALAPA

TÍTULO DEL PROYECTO

**Sistema de información para la gestión estratégica en los
Servicios de Salud de Veracruz**

**Opción de Titulación:
I Tesis profesional**

Que como requisito parcial para la obtención de grado de

Maestría en Sistemas Computacionales

Presenta:

María del Rosario Suazo García

**No. De Control:
197003606**

**Director
Dra. Virginia Lagunes Barradas**

**Co- Director
Dr. Alberto Téllez Valero**

Xalapa-Enríquez, Veracruz, noviembre del 2022

Dedicatoria

Mamá Toña,

Por ser mi mentora de vida y estar siempre cuando lo necesité.

Gracias por todo.

Víctor,

Por ser mi mentor profesional, por enseñarme e inculcar en aprender cosas nuevas, pero, sobre todo que se adquiere más cuando el conocimiento se comparte.

Gracias.

Agradecimientos

*Pensó que si alguien le había ayudado
tenía que sentirse agradecida y decírselo.
Murmuró: gracias.
Gonzalo Torrente Ballester.*

Gracias a todas aquellas personas que siempre estuvieron de alguna manera presente y haber hecho posible concluir esta tesis. De manera especial a:

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada durante los dos años de estudio y en especial al Instituto Tecnológico Superior de Xalapa, por facilitar y colaborar en el desarrollo de este proyecto.

A mis directores de tesis, Dra. Virginia Lagunes Barradas y Dr. Alberto Téllez Valero, por su apoyo en el desarrollo del tema de investigación.

A mi comité tutorial en especial al Dr. Salvador Herrera Velarde, por su tiempo dedicado, consejos y sugerencias para la mejora de este trabajo.

A mis padres y Padrino, porque con su ejemplo y actitud resiliente ante la vida, forjaron mi carácter, mis valores e infundieron en mi ser el trabajo, la paciencia y la fortaleza ante la adversidad.

A mis hermanos, porque nuestras diferencias nos complementan y enriquecen el camino hacia nuestros objetivos comunes, siempre con entereza y unidad.

A mí mamá Toña, porque su abrazo eterno y su fe inquebrantable me acompañan, me sostienen y están allí bajo toda circunstancia.

A Víctor, por su profesionalismo y mostrarme que la sed de aprendizaje sólo se sacia al compartir el conocimiento.

Gracias a Chela, Chris, Fer, Celia, Lore y Xóchitl por escucharme, enseñarme y brindarme su apoyo incondicional, pero sobre todo por su amistad.

A todos y a cada uno de ellos que se interesaron en este proyecto. ¡Muchas Gracias!

Índice detallado

Índice de Figuras.....	5
Listado de Tablas.....	6
Resumen	7
Listado de Abreviaturas	8
Introducción	9
Capítulo I. Desarrollo metodológico	11
1.1 Planteamiento del problema	11
1.2 Objetivos de la investigación.....	13
1.2.1 General.....	13
1.2.2 Específicos	13
1.3 Justificación	13
1.4 Alcances y limitaciones.....	15
1.5 Hipótesis.....	16
1.5.1 Preguntas de investigación.....	16
1.5.2 Variables:	16
1.6 Tipo de investigación.....	17
1.7 Estado del arte.....	17
Capítulo II: Marco de referencia	21
2.1 Servicios de Salud del de Veracruz.....	22
2.1.2 Marco normativo en materia de salud	24
2.2 Inteligencia de negocios.....	25
2.2.1 Proceso de toma de decisiones	25
2.3.1 Bases de datos multidimensionales	27
2.3.2 Tecnologías OLAP	28
2.3.3 Cubos OLAP	30
2.3.4 Esquemas de modelos multidimensionales	31
2.3.5 Sistema de Almacén de Datos.....	34
2.3.6 Extraer Transformar y Cargar	35
2.4 Metodologías para el desarrollo del Almacén de datos.....	36
2.5 Visualizaciones y procesamiento analítico	37

Capítulo III: Diseño	38
3.1 Material y métodos	38
3.2 Metodología	39
3.3 Técnicas e instrumentos para el diseño del almacén de datos	40
3.3.1 Fase 1: Identificar eventos	41
3.3.2 Fase 2: Clasificar eventos	43
3.3.3 Fase 3: Matriz de eventos	44
3.3.4 Fase 4: Esquema multidimensional	50
3.4 ETL	51
3.5 Integración y generación de visualizaciones de consultas	52
3.5.1 Fase 1: Planificar	53
3.5.2 Fase 2: Diseño y construcción	54
3.5.3 Fase 3: Implementación y mantenimiento	54
Capítulo IV: Resultados	55
4.1 Diseño de solución BI en Visual Studio	55
4.1.1 Creación de la base de datos	55
4.1.2 Creación de la solución	57
4.1.4 Creación de dimensiones	58
4.1.5 Esquema copo de nieve	59
4.2 Diseño y ejecución del ETL	60
4.3 Integración y generación de visualizaciones	62
4.4 Evaluación de la usabilidad en el Tablero de datos de ETV	66
Conclusiones	68
Trabajo futuro	69
Referencias	71
Anexos	76
Anexo 1: Entrevista	76
Anexo 2: Matriz o tabla de eventos	78
Anexo 3: Descripción de tablas modelo relacional	79
Anexo 4: Diseño de instrumento de usabilidad	83

Índice de Figuras

Figura 1. Organigrama de la Dirección de Planeación y Desarrollo	22
Figura 2. Flujo de información de notificación de vigilancia epidemiológica.....	24
Figura 3. Proceso de toma de decisiones en la Inteligencia de negocios.....	26
Figura 4. Ejemplo de estructura del Cubo OLAP.....	31
Figura 5. Esquema copo de estrella	32
Figura 6. Esquema copo de nieve.....	33
Figura 7. Esquema de constelación.....	33
Figura 8. Arquitectura de un almacén de datos	34
Figura 9. Proceso ETL	35
Figura 10. Metodología para la propuesta de inteligencia de negocios.....	40
Figura 11. Modelado lógico del diseño dimensional mediante técnicas BEAM.....	41
Figura 12. Propuesta de Tabla de eventos BEAM.....	49
Figura 13. Esquema de copo de nieve propuesto	50
Figura 14. Diseño de ETL de ETV.....	51
Figura 15. Diseño de visualizaciones de datos.....	53
Figura 16. Restauración de base de datos.....	56
Figura 17. Consultas SQL	57
Figura 18. Vista del modelo multidimensional.....	58
Figura 19. Esquema multidimensional copo de nieve.....	60
Figura 20. Proceso de ETL de las ETV.....	61
Figura 21. Procesamiento de consultas al Cubo OLAP.....	62
Figura 22. Jerarquía del modelo multidimensional para datos epidemiológicos.....	63
Figura 23. Tablero casos de dengue por años.....	64
Figura 24. Tablero de datos de casos de dengue geográficos del Estado de Veracruz	65
Figura 25. Tablero de datos de casos de dengue por jurisdicción.....	65
Figura 26. Formato matriz de eventos BEAM.....	78

Listado de Tablas

Tabla 1: Cuadro comparativo de base de datos transaccionales y multidimensionales _____	28
Tabla 2. Tabla comparativa de tecnologías OLAP _____	29
Tabla 3. Tabla de comparación de metodologías _____	36
Tabla 4. Roles y departamentos _____	42
Tabla 5. Requerimientos de 7W's _____	44
Tabla 6. Pregunta del evento 1 _____	45
Tabla 7. Pregunta del evento 2 _____	45
Tabla 8. Pregunta del evento 3 _____	46
Tabla 9. Pregunta del evento 4 _____	46
Tabla 10. Pregunta del evento 5 _____	46
Tabla 11. Pregunta de evento 6 _____	47
Tabla 12. Tabla de dimensiones _____	47
Tabla 13. Jerarquías _____	48
Tabla 14 Tabla de descripción dimensiones _____	58
Tabla 15. Tabla de descripción de los participantes en instrumento de usabilidad _____	66
Tabla 16 Puntuaciones a las preguntas del cuestionario de evaluación de la satisfacción del usuario __	67
Tabla 17. Descripción detallada de la dimensión de unidades médicas _____	79
Tabla 18. Descripción detallada de la dimensión años _____	80
Tabla 19. Descripción detallada de la dimensión de género _____	80
Tabla 20. Descripción detallada de la dimensión grupo de diagnóstico _____	80
Tabla 21. Descripción detallada de la dimensión grupo de edad _____	80
Tabla 22. Descripción detallada de la dimensión meses _____	81
Tabla 23. Descripción detallada de la dimensión padecimientos _____	81
Tabla 24. Descripción detallada de la dimensión semanas epidemiológicas _____	81
Tabla 25. Descripción detallada de la tabla de hechos _____	82
Tabla 26. Cuestionario de usabilidad _____	83

Resumen

La información es un gran recurso en las empresas e instituciones debido a que con esta se puede generar conocimiento y tomar decisiones para mejorar procesos y así obtener una ventaja competitiva.

En este ámbito las herramientas de inteligencia de negocios permiten a las organizaciones descubrir o conocer sus datos de manera rápida y dinámica, con el fin de consultar miles de registros y se puede obtener un análisis de datos retrospectivo.

Así surge esta investigación, la cual propone la implementación de un sistema de información para la gestión estratégica en los Servicios de Salud de Veracruz, con la finalidad de implementar un tablero de análisis de datos a través de herramientas de inteligencia de Negocios (BI). El objetivo específico del proyecto consiste en apoyar a los usuarios para la consulta de forma interactiva y rápida los casos de dengue en el Estado de Veracruz y, al mismo tiempo apoye a la toma de decisiones respecto al control de Enfermedades Transmitidas por Vector. De igual manera, se plantea realizar un único repositorio de datos (Data warehouse) para que los usuarios consulten la información a través de sus dispositivos móviles y les facilite la elaboración de sus propios informes.

Por último se plantea un instrumento de usabilidad que permita evaluar si la propuesta de la herramienta de análisis de datos es óptima para la consulta de información.

Listado de Abreviaturas

A continuación se incluye un listado de las abreviaturas utilizadas en el documento por orden alfabético

BEAM	Business Event Analysis and Modeling
BI	Business Intelligence
CIE-10	Clasificación internacional de enfermedades, 10ª edición
DGE	Dirección General de Epidemiología
DW	Data Warehouse
ETL	Extract, Transform and Load
ETV	Enfermedades Transmitidas por Vector
HOLAP	Hybrid OnLine Analytic Processing
MOLAP	Multidimensional online analytical processing
NM	Normas Mexicanas
OLAP	On-Line Analytical Processing
OLTP	OnLine Transaction Processing
OMS	Organización Mundial de la Salud
OPS	Organización Panamericana de la Salud
ROLAP	Relational OnLine Analytic Processing
SE	Semana Epidemiológica
SESVR	Servicios de Salud de Veracruz
SINAVE	Sistema Nacional de Vigilancia Epidemiológica
SIRES	Sistemas de Información de Registro Electrónico para la Salud
SQL	Structured Query Language o Lenguaje de consultas estructuradas
SSDT	SQL Server Data Tools
SSIS	SQL Server Integration Services
SSRR	SQL Server Reporting Services
SUAVE	Sistema Único Automatizado para la Vigilancia Epidemiológica
SUIVE	Sistema Único de Información para la Vigilancia Epidemiológica

Introducción

El *Aedes aegypti* es un mosquito responsable de transmitir enfermedades epidemiológicas a la población como el dengue, zika y chikungunya. Este mosquito se propaga con gran facilidad debido a factores climatológicos haciéndolo presente en gran parte del mundo. En particular el dengue es una enfermedad febril que forma parte del grupo de las 16 Enfermedades Transmitidas por Vector, de acuerdo con la Clasificación Internacional de Enfermedades (CIE-10) y es un problema de salud pública en México y en el Estado de Veracruz.

En la presente investigación se plantea el uso de diferentes herramientas tecnológicas de inteligencia de negocios que permita a los usuarios de los Servicios de Salud de Veracruz consultar información del dengue a través de un tablero de datos y, a la vez, apoye a encontrar estrategias de su control y erradicación. A la fecha el sistema de Servicios de Salud de Veracruz no cuenta con un repositorio de datos homogéneo. Actualmente la generación de informes se realiza a través de archivos PDF y hojas de cálculo, por lo que resulta complejo para los usuarios.

En esta tesis se plantea la necesidad de estandarizar el análisis y generación de reportes a través de herramientas de Inteligencia de Negocios. El presente documento se encuentra estructurado de la siguiente manera.

En el capítulo uno se detalla el planteamiento del problema y los objetivos de investigación, así como la búsqueda de trabajos similares realizados en otros países.

El capítulo dos presenta la investigación documental referente a las características de las herramientas y metodologías para realizar este trabajo. De igual manera se presenta la propuesta de las visualizaciones de los datos epidemiológicos.

El capítulo tres describe los materiales, métodos y la metodología que se proponen para lograr implementar las herramientas de inteligencia de negocios y desarrollar un modelo multidimensional que permita consultar los casos de dengue a través de tableros dinámicos y accesibles.

El capítulo cuatro describe la propuesta de implementación de un tablero de datos como herramienta de análisis de información de las Enfermedades Transmitidas por Vector, en el caso específico del dengue.

Las conclusiones y trabajo futuro se detallan al final del documento. El trabajo de desarrollo contiene 4 anexos. El primero comprende el formato de entrevistas realizada a los usuarios de

los Servicios de Salud de Veracruz. El anexo 2 tiene el formato de una matriz o tabla de eventos propuesta para el desarrollo de modelos multidimensionales. El anexo 3 contiene la descripción de tablas para la elaboración del modelo de la base de datos y por último el anexo 4, considera un el diseño de un instrumento de usabilidad, que permite evaluar la satisfacción de la implementación de la herramienta de datos.

Capítulo I. Desarrollo metodológico

En este capítulo se describe la problemática de estudio referente a las enfermedades transmitidas por vector y la complejidad de generar reportes por parte de los usuarios del sector salud. También se plantea el objetivo general, el cual, a grandes rasgos, consiste en el desarrollo de un Sistema de información para la gestión estratégica en Servicios de Salud de Veracruz.

En la siguiente sección se plantea el objeto de estudio y la razón por la cual su estudio, así como el tipo de investigación que se realizó. De igual manera se presentan resultados de la búsqueda literaria de los estudios relacionados con la implementación de inteligencia de negocios, desarrollo de sistemas o visualización de datos de enfermedades epidemiológicas.

1.1 Planteamiento del problema

La información es considerada un recurso de gran valor con la que cuentan las instituciones públicas y privadas, pero si no es administrada de manera adecuada y puesta a disposición de todos los niveles jerárquicos de la organización para tomar decisiones, simplemente deja de tener valor (Pineda & Diaz, 2006). Desde ese enfoque informático, el sector salud siempre origina datos cuya captura, intercambio, almacenamiento, consulta y análisis de estos apoyan la toma de decisiones en materia de salubridad. De allí surge el objeto de estudio de este proyecto de investigación: resaltar la implementación de herramientas de inteligencia de negocios para obtener una gestión apropiada de la información a través de un sistema estratégico igual de eficaz y oportuno.

Dicho sector cuenta con diversos sistemas y procesos de información a través de los cuales integra y controla la prestación de sus servicios e implementa medidas, estrategias y técnicas relacionadas con la detección de contingencias de salud. Por lo que, a partir de la producción de esta información, los tomadores de decisiones pueden alertar sobre contingencias epidemiológicas, contribuir a su prevención y control, y de esa manera evitar que se conviertan en problemas mayores de salud pública (Cervantes, 2014).

Esto último es de gran importancia, ya que la detección de enfermedades constituye el punto de partida, junto con la notificación de casos, de una serie de actividades efectuadas para establecer una vigilancia epidemiológica oportuna (Secretaría de Salud, 2017). En ese contexto, la institución gubernamental veracruzana que, dentro de sus múltiples funciones, precisamente identifica riesgos, eventos o situaciones de vigilancia epidemiológica, con el

propósito de implementar las correspondientes acciones de prevención y control, y reducir de ese modo su impacto en la salud de la población, es Servicios de Salud de Veracruz (SESVER, 2016).

En la actualidad, SESVER forma parte del Sistema Nacional de Vigilancia Epidemiológica (SINAVE): este es un conjunto de herramientas tecnológicas, métodos, estrategias, lineamientos y procesos que garantiza información oportuna para la vigilancia epidemiológica; además, se encuentra estructurado jerárquicamente. También posee múltiples herramientas técnicas, dos de ellas dedicadas a la recolección de datos epidemiológicos: el Sistema Único Automatizado de Vigilancia Epidemiológica (SUAVE) y el Sistema Único de Información para la Vigilancia Epidemiológica (SUIVE). La primera herramienta es un sistema automatizado que captura información confiable a través del segundo instrumento, un formato semanal que recolecta manualmente información mediante diagnósticos médicos integrados en unidades médicas y de acuerdo con la NOM-017-SSA2-2012. Ambos instrumentos producen información de consulta e interés público que se concentra en el Boletín Epidemiológico semanal y los cubos dinámicos.

En la actualidad SESVER emplea diferentes herramientas tecnológicas para visualizar y analizar información sanitaria, tales como archivos en PDF y archivos en hojas de cálculo de Excel. Estos formatos representan obstáculos para la búsqueda de información y la subsecuente elaboración de reportes.

La primera herramienta es el Boletín Epidemiológico —documento electrónico no manipulable con información sobre casos nuevos de cerca de 147 enfermedades o padecimientos sujetos a vigilancia epidemiológica, publicado en el sitio web del Gobierno federal cada miércoles—, debido a que en caso de necesitar SESVER datos históricos semanales o anuales para analizarlos y a partir de ellos tomar decisiones en materia de vigilancia epidemiológica, debe realizar una “farragosa” búsqueda secuencial de los archivos y descargar cada uno de ellos (DGE, 2020b).

En segundo lugar, SESVER hace uso de los cubos dinámicos exportables a Excel, los cuales analizan los datos a partir de estructuras multidimensionales. Estos requieren de una configuración específica en el equipo de cómputo, aunado a instalar complementos en el navegador web para obtener una mejor visualización gráfica de los datos. Estas dificultades expresadas por los usuarios confirman que, dependiendo del volumen de datos, de su actualización y de la infraestructura del equipo de cómputo con que cuentan, el procedimiento de

consulta y análisis de información puede resultar complejo y lento a la hora de manipular la información o generar reportes.

1.2 Objetivos de la investigación

1.2.1 General

Desarrollar un sistema de información estratégico para Servicios de Salud del Estado de Veracruz (SESVER), con la finalidad de optimizar su análisis de datos y apoyar su toma oportuna de decisiones en materia de salud.

1.2.2 Específicos

- Analizar los procesos de consulta de información en SESVER para identificar las variables de construcción del sistema estratégico conforme las normas oficiales mexicanas NOM-035-SSA3-2012, NOM-024-SSA3-2012 y NOM-017-SSA2-2012.
- Implementar herramientas de inteligencia de negocios para la optimización de datos en apoyo a la toma de decisiones en materia de salud en SESVER.
- Agilizar la toma de decisiones a través del uso de tableros dinámicos prediseñados y personalizados.
- Evaluar la usabilidad del sistema de información para la gestión estratégica, priorizando los tiempos de consulta y la generación de reportes.

1.3 Justificación

Las Enfermedades Transmitidas por Vector (ETV) son aquellas causadas por un agente vivo (insecto o mosquito) que ingiere microorganismos de la sangre de sus huéspedes mediante picaduras: dengue y paludismo son los ejemplos más comunes. En los últimos 30 años el incremento considerable del número de casos de estos padecimientos en zonas urbanas y semiurbanas de nuestro país lo ha convertido en un problema predominantemente de salud pública (OMS, 2017a).

López-Gatell (2018) expresó que el dengue es una enfermedad viral que ha prevalecido durante décadas en el panorama de salud pública nacional, a la cual se le ha intentado encontrar una respuesta mediante el control de los mosquitos causantes y reducir su transmisión, sin embargo, al momento todavía es un problema de salud que consume demasiados recursos

financieros, tecnológicos, humanos y materiales (Academia Nacional de Medicina de México, 2018)

El dengue pertenece al grupo de los 16 padecimientos de las ETV, por ende, está dentro del radar del SINAVE, se trasmite a los seres humanos por un piquete del mosquito vector, cuyo nombre científico es *Aedes aegypti*. Este mismo agente, como señala la OMS puede completar su ciclo de vida (de huevo a adulto) en un período de 7-10 días, y como adultos generalmente viven de 4 a 6 semanas. Asimismo, el dengue se clasifica en cuatro serotipos o microorganismos: DENV-1, DENV-2, DENV-3 y DEN-V 4 (OPS, 2020).

El ciclo de vida de los mosquitos causantes del dengue varía acorde a las características climáticas de cada país, y principalmente pueden vivir en climas tropicales con humedad y oscuridad; normalmente no sobreviven en temperaturas frías o calurosas (CENEPRACE, 2020). Sin embargo, el *Aedes aegypti* es capaz de adaptarse a los cambios climáticos y posee una gran habilidad de guarecerse en microhábitats que les propician una propagación y proliferación más constante. De acuerdo con la OMS (2020b) este mosquito “está en casi todos los países excepto Canadá y Chile”.

En 1970 se reportaron los primeros casos de incidencia del dengue a nivel mundial en nueve países. En la actualidad, debido a su rápida adaptación climatológica y dispersión geográfica el *Aedes aegypti* se encuentra presente en 128 países (Hernández-Ávila, 2013). En el año 2019 se reportaron más de cuatro millones de casos epidemiológicos, incluidas más de mil defunciones (DeAntonio et. al, 2021). En el 2020 los países que registraron el mayor número de casos de dengue fueron: Brasil, Paraguay, Bolivia, Argentina y Colombia. De igual manera, Honduras, México y Nicaragua reportaron altas tasas de incidencia de dengue (OMS, 2020b).

La historia del dengue en México comenzó en 1978 cuando se reportaron en el Estado de Chiapas los primeros pacientes con tal padecimiento febril (Ochoa et. al, 2015). Durante el período del 2008 a 2011, se notificaron casos de dengue en 30 de las 32 entidades federativas de México (Hernández-Ávila, 2013). Su costo anual en la población mexicana se estimó en alrededor de US \$170 millones de dólares (Undurraga et. al, 2015). De acuerdo con el informe del Sistema Nacional de Vigilancia Epidemiológica, en el 2019, Veracruz fue una de las entidades con mayor incidencia de casos confirmados (DGE, 2021c). Por lo que a la fecha es uno de los Estados de la República Mexicana con alta endemicidad para dengue, inclusive se tiene documentada la circulación de sus cuatro serotipos (Ortiz & Ayala, 2017).

El anterior panorama epidemiológico y endémico de la entidad veracruzana vislumbra lo fundamental de una gestión adecuada de la información que genera y maneja una institución como SESVER, la cual debe responder a circunstancias cambiantes y mostrar resultados en el menor tiempo posible a sus usuarios con mecanismos o estrategias que propicien actividades de prevención y control de asuntos epidemiológicos de gran impacto como lo es el dengue.

El objeto de la presente investigación es la implementación de una plataforma de inteligencia de negocios interactiva, gráfica y dinámica; más un correspondiente proceso analítico, que apoye la toma de decisiones de los usuarios de SESVER de realizar acciones de prevención y control de los riesgos sanitarios del dengue. Ésta permitirá a los usuarios consultar acontecimientos históricos de los casos del dengue e integrar información de dieciocho años proveniente de los municipios de Veracruz, a través de tableros gráficos dinámicos, tendientes a facilitar al usuario la elaboración y difusión de sus propios informes. De igual manera, se pretende simplificar el acceso a un solo repositorio de datos para agilizar la consulta de información desde diversos dispositivos, incluidos los móviles.

1.4 Alcances y limitaciones

El presente proyecto consiste en la implementación de un sistema de información dirigido hacia la gestión estratégica en SESVER, especialmente en la Subdirección de Análisis e Integración de Información en Salud, que permita optimizar la toma de decisiones a través de un proceso analítico y a la par, considere los acontecimientos de ETV en 947 unidades médicas y 11 jurisdicciones municipales de Veracruz durante un período del 2003 a 2020.

Se presupone que el oportuno análisis de la información a partir de la integración de herramientas de inteligencia de negocios y la elaboración de tableros dinámicos permitirá a los usuarios llevar a cabo estrategias y emplear técnicas sobre el control de vectores, en el caso particular de este trabajo, del dengue; además de que proporcionará la visualización de reportes históricos y analíticos prediseñados del mismo.

No obstante, una limitante para alcanzar la anterior premisa radica en que la comprensión de algunos datos estadísticos e indicadores planteados por este trabajo de investigación requiere del asesoramiento de un experto (epidemiólogo). Aunado a la escasa documentación acerca de la arquitectura empleada por las herramientas de inteligencia de negocios en materia sanitaria para las ETV; una que permita el apoyo a la toma de decisiones de manera oportuna y realizar un análisis eficiente de vigilancia epidemiológica.

1.5 Hipótesis

La tesis plantea como hipótesis que el modelo de inteligencia de negocios para la gestión estratégica de información en Servicios de Salud de Veracruz (SESERVER) optimizará su proceso de análisis, monitoreo y toma de decisiones respecto a enfermedades transmitidas por vector al mismo tiempo, se piensa que generará servicios oportunos de atención a la salud de la población del estado.

1.5.1 Preguntas de investigación

Las preguntas de investigación se formulan de acuerdo con las herramientas y técnicas para el desarrollo de este proyecto que, permita la visualización de los datos de manera dinámica:

- a) ¿Cuáles serán las herramientas de inteligencia de negocios que permitirán construir un sistema de información para la gestión estratégica en SESERVER que apoye la toma de decisiones en materia de salud para las enfermedades transmitidas por vector?

- b) ¿Cómo deberán diseñarse los tableros dinámicos y sus pruebas de usabilidad para optimizar los tiempos de consulta y la generación de reportes referentes a las Enfermedades Transmitidas por Vector?

1.5.2 Variables:

El planteamiento de las variables de investigación para el proyecto de tesis se formulan de acuerdo con el planteamiento del problema, en las que se clasifican de la siguiente manera:

- 1) Dependientes:
 - a. Uso de la información para la toma de decisiones a través de tableros dinámicos que permitan la generación de reportes históricos.
 - b. Visualización de datos a partir de la creación del repositorio de datos (cubo OLAP), el cual muestre el número de casos de dengue, padecimientos, jurisdicciones, unidades médicas, grupo de edad, semana epidemiológica y género.
 - c. Funcionalidad de los tableros dinámicos.
 - d. Reducción del tiempo de procesamiento de la información.

- 2) Independientes:

- a. Plataforma de inteligencia de negocios que permita la consulta de información de datos históricos del dengue en Veracruz.
- b. Generación de tableros dinámicos.

1.6 Tipo de investigación

El presente proyecto utiliza los métodos de la investigación documental o teórica, que implicó la búsqueda de normas oficiales mexicanas, manuales de organización, programas de operación, lineamientos y publicaciones que traten temas relacionados con la vigilancia epidemiológica de las ETV y sus respectivos almacenes de datos. Por ende, analizará los procesos de SESVER a través de manuales de organización y operación para posteriormente diseñar y desarrollar un almacén de datos adecuado que cumpla con las normas establecidas en materia de salud.

Además, por la naturaleza de los datos, se utilizó un enfoque cuantitativo aplicada a la elaboración de la estructura del almacén de datos o Data Warehouse (DW), que permitió determinar un conjunto de técnicas e instrumentos para la implementación de herramientas de inteligencia de negocios. Estas herramientas y metodologías ágiles implicaron el diseño del DW. Se identificaron las variables de construcción, extracción, transformación y carga de datos, así como, en el procesamiento analítico en línea. El objetivo es que los usuarios de SESVER realicen consultas de manera sencilla, rápida y oportuna mediante la visualización de los datos. El análisis de la información se realizó mediante el método de observación y descriptivo basado en la experiencia práctica de los usuarios de la Subdirección de Análisis e Integración de la Información en Salud, perteneciente a SESVER.

1.7 Estado del arte

Los sistemas de información actuales evolucionan enormemente con el transcurrir de cada año; hoy en día se concentran en mostrar, coadyuvar y automatizar datos que apoyan la toma de decisiones de las empresas. La Organización Panamericana de la Salud, indicaba que “la única preocupación con respecto a los sistemas de información es que éstos deben apoyar a la salud y a los servicios sanitarios, y la finalidad de la tecnología debe consistir en favorecer el sistema de información correspondiente” (PAHO, 1999).

Los sistemas computarizados relacionados con el sector salud se orientan al entendimiento de sus necesidades; se encargan de obtener, organizar, analizar y compartir información indispensable que conlleve a mejores decisiones respecto a la vigilancia epidemiológica, y la evaluación de los servicios o programas de intervención en materia de salud, a fin de aumentar la calidad de cada rubro. No obstante, los sistemas necesitan estar vinculados con los requerimientos de los usuarios, de la misma manera en que los servicios de salud registran eventos de importancia epidemiológica.

Una tarea imprescindible al momento de establecer todo sistema de información en el sector salud es determinar cuáles son los problemas y factores que repercuten. Por ello a continuación se describen algunas investigaciones, tesis e informes en torno a la inteligencia de negocios en el sector salud.

En el marco de la Convención Internacional de Salud Pública “Cuba-Salud 2018” se propuso, la implementación de un almacén de datos para la vigilancia epidemiológica en la provincia de Santiago de Cuba. Una de las problemáticas identificadas fue el gran volumen de datos sanitarios y estadísticos manejados por su Departamento de Vigilancia Epidemiológica y Análisis en Salud. La información era tratada de forma semi-manual por medio de documentos en Word, tablas en Excel y diversos sistemas estadísticos independientes, lo que ocasionaba la dispersión de la información en distintos sistemas, hacía lento e ineficiente el análisis de los datos y provocaba la poca efectividad de las acciones y decisiones sanitarias. De ese modo, dicho proyecto buscaba garantizar una mejor organización, integración y acceso de la información con el empleo de la metodología Hefesto y la *suite* Pentaho BI. Además, su planteamiento permitió integrar en un solo sistema los cálculos estadísticos, el análisis de los datos y los reportes correspondientes, a la par que propició un minucioso control epidemiológico (Quintana *et al.*, 2018).

Por otra parte, Ortiz y Ayala (2017) idearon un proyecto de inteligencia de negocios para la empresa ecuatoriana Farmaenlace Cía. Ltda., dado que no contaba con un sistema de información con indicadores de desempeño (KPI) establecidos que apoyaran el proceso de toma de decisiones. Motivo por el cual una vez implementada la propuesta se tuvo alcance a los indicadores antes mencionados, necesarios para el análisis situacional y como apoyo en la toma de decisiones.

Durante la novena Conferencia Internacional de Ciencia, Ciencias Sociales, Ingeniería y Energía (I-SEC-2018) se presentó un estudio auspiciado por la Universidad Kasem Bundien en

Tailandia. Tal trabajo contenía un análisis de datos espaciales basado en un modelo multidimensional de un esquema de copo de nieve, diseñado a partir de herramientas de inteligencia de negocios con código abierto. Asimismo, se efectuó mediante una aplicación web de mapeo interactivo de datos de los casos de dengue, lo que ayudó a facilitar la supervisión y el control efectivos de los brotes de enfermedades. La evaluación de los usuarios mostró resultados favorables, pues estaban satisfechos con el uso de un sistema del más alto nivel en términos de efectividad (Asavasuthirakul, 2018).

En el estudio estadístico de Jaya y Folmer (2020), se realizó un modelo que consistía en el Mapeo espacio - temporal bayesiano del riesgo relativo de enfermedad del dengue en Bandung, Indonesia. Este modelo contiene datos de 30 distritos en la ciudad de Bandung en un periodo de enero de 2009 a diciembre de 2016. Se compararon modelos estadísticos en que se incluyeron datos climatológicos para el pronóstico. Sin embargo, se observó que las condiciones ambientales y socioeconómicas aumentaron los casos de incidencia del dengue en los distritos del noroeste, además, los factores que observaron es que cada distrito tiene un patrón temporal diferente, lo que no permite realizar una la predicción asertiva.

En la investigación de Wibowo et al. (2020), se planteó un sistema con herramientas de inteligencia de negocios que permitía predecir el número de casos de dengue en un período próximo y así reducían la mortalidad. Comprobaron que con un tablero gráfico (mapas) como apoyo de los sistemas de decisiones la información se volvía más rápida y permitía a la Oficina de Salud de Puskesmas, Indonesia, preparar mejor sus instalaciones médicas. El panel de control de tal propuesta mostró un pronóstico y mapa de distribución de pacientes, así como una predicción con número de pacientes, datos climatológicos (lluvia, temperatura y humedad), estado de salud de los pacientes, así como los resultados de investigaciones epidemiológicas.

Por otra parte, en el artículo de investigación de Kiang et. al (2021) consideraron importante la incorporación de datos de movilidad humana para la mejora de las previsiones de la fiebre del dengue en Tailandia. Por lo que, realizaron los pronósticos de series de tiempo (integrando datos epidemiológicos del programa de vigilancia del dengue) con modelos de movilidad generados a partir de datos de teléfonos móviles de aproximadamente 11 millones de suscriptores. En dicho análisis, mostraron que las provincias geográficamente distantes están fuertemente conectadas por viajes y tienen una incidencia de dengue altamente correlacionada que las provincias débilmente conectadas de la misma distancia; de igual manera, determinaron que las provincias del norte del país son más difíciles de pronosticar con confianza que las del

sur, independientemente de la elección del modelo, y que los resultados de los diferentes modelos pueden estar vinculados a factores demográficos y sociales.

Por último, Nuraini et al (2021) realizó un estudio de un modelo matemático referente a las predicciones del dengue basado en el clima, de Semarang, Indonesia. Dicho modelo fue integrado por factores meteorológicos en el modelo de infección de enfermedades de transmisión del dengue (modelo huésped-vector); sin embargo, observaron que el clima está asociado positivamente tanto con la incidencia como con la existencia de vectores. En sus resultados determinaron, que el modelo matemático propuesto no ofrece predicciones a largo plazo que puedan verse influenciadas de la correlación entre la infección y los factores climatológicos. Además, menciona que el estudio de investigación cuenta con limitaciones, como por ejemplo los datos de incidencia de dengue obtenidos solo realizan el número total de casos hospitalizados. Sin embargo, los detalles de la clasificación, por ejemplo, género, estructura de edad, ocupación y dirección, no están disponibles. De igual manera los factores climáticos no siempre permiten determinar si disminuye o aumenta el dengue, por lo que, las variables climáticas por sí solas no pueden ser un buen predictor del riesgo futuro de dengue en un tiempo de observación prolongado, para predecir la incidencia del dengue.

Capítulo II: Marco de referencia

Para realizar un sistema de información para las empresas es necesario conocer cuáles son los criterios requeridos por el nivel operativo. De igual manera es sustancial determinar las normas que establece la organización de acuerdo con los procesos para recolectar, analizar, validar y visualizar la información; estos criterios normalmente se basan en los procedimientos administrativos y legales, cuyo objetivo es determinar qué información es relevante que permita satisfacer las necesidades y apoye a tomar decisiones que le permitan tener una ventaja o mejorar sus propios procesos.

En este capítulo se identificarán las actividades y procesos que realiza SESVER para la notificación de casos nuevos de vigilancia epidemiológica de las enfermedades transmitidas por vector, así como, la normatividad vigente relativa a la captura de datos y análisis de información. Así mismo se identificarás las técnicas adecuadas a fin de transformar, almacenar y visualizar la información de manera dinámica.

2.1 Servicios de Salud del de Veracruz

Los Servicios de Salud de Veracruz (SEVER) es la institución responsable de realizar la notificación periódica de las ETV y de casos nuevos de enfermedades epidemiológicas a través del SINAVE. Su estructura orgánica está integrada por ocho direcciones: general, salud pública, atención médica, protección contra riesgos sanitarios, planeación y desarrollo, infraestructura de salud, administrativa y jurídica y diez órganos descentralizados (SESVER, 2016).

Nuestro objeto de estudio es la Subdirección de Análisis e Integración de Información de Salud que pertenece a la Dirección de Planeación y Desarrollo (ver Figura 1); en esta área los usuarios consultan información de las enfermedades o padecimientos de vectores. Posteriormente realizan la difusión de informes institucionales de 147 enfermedades sujetas a vigilancia epidemiológica de las 947 unidades médicas, permitiendo a los directivos la toma de decisiones y contribuyendo a la optimización de los recursos en congruencia con las políticas federales y estatales de modernización administrativa (SESVER, 2016).

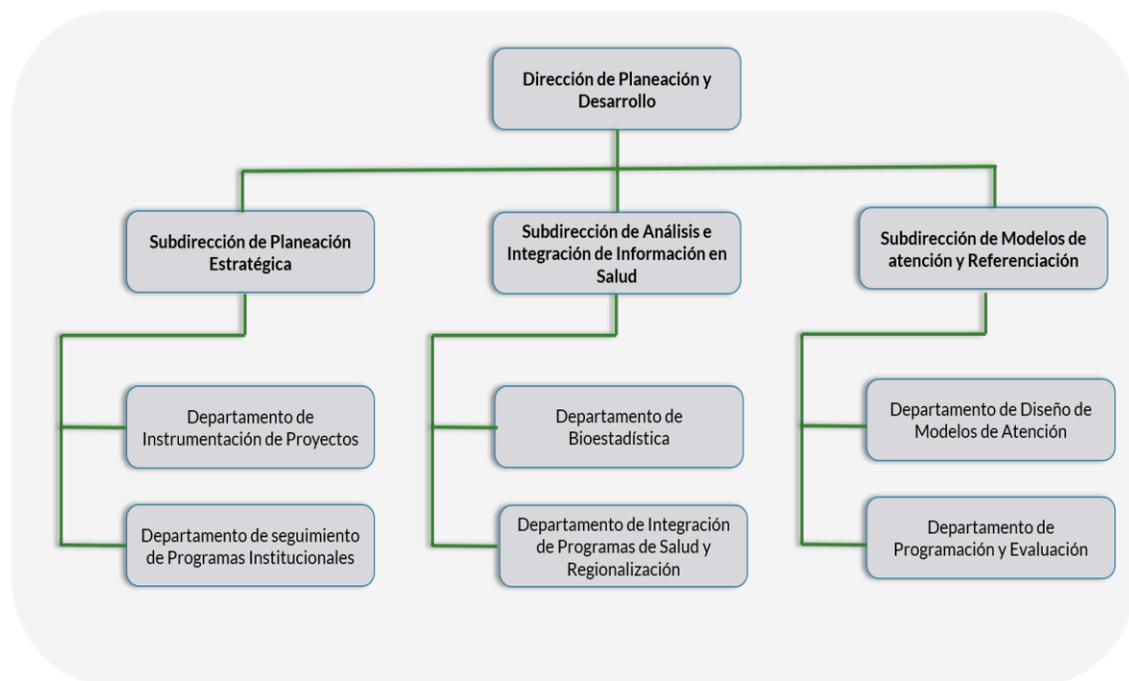


Figura 1. Organigrama de la Dirección de Planeación y Desarrollo

Como se muestra en la Figura 2 el proceso de flujo de información para que los usuarios de la subdirección consulten información y generen reportes a través de las herramientas y técnicas del SINAVE, se realiza a través de 2 niveles jerárquicos:

- 1) **Nivel estatal:** se encuentran las 947 unidades médicas u hospitalarias y 11 jurisdicciones de los 212 municipios. En otros términos, los servicios de salud (jurisdicciones) informan a la autoridad sanitaria (federal) sobre la atención de eventos de vigilancia epidemiológica a través del formato SUIVE-2014 en Excel (o formularios en papel), en que se registran datos estadísticos numéricos.

- 2) **Nivel federal:** en este nivel se cumplen tres funciones principales:
 - a) Concentrar información reportada por las unidades médicas, de todos los estados de la República Mexicana, a través de sus plataformas tecnológicas de recolección de datos.

 - b) Normar las funciones para la vigilancia epidemiológica y de laboratorio de la ETV en el país, además, de asesorar las actividades de vigilancia epidemiológica en todos los niveles operativo.

 - c) Difundir la información estadística a través de cubos dinámicos y del Boletín Epidemiológico. Este es un informe que se publica semanalmente en la página oficial de la Dirección General de Epidemiología (DGE) de manera electrónica y se mantienen los informes históricos para la consulta de usuarios.

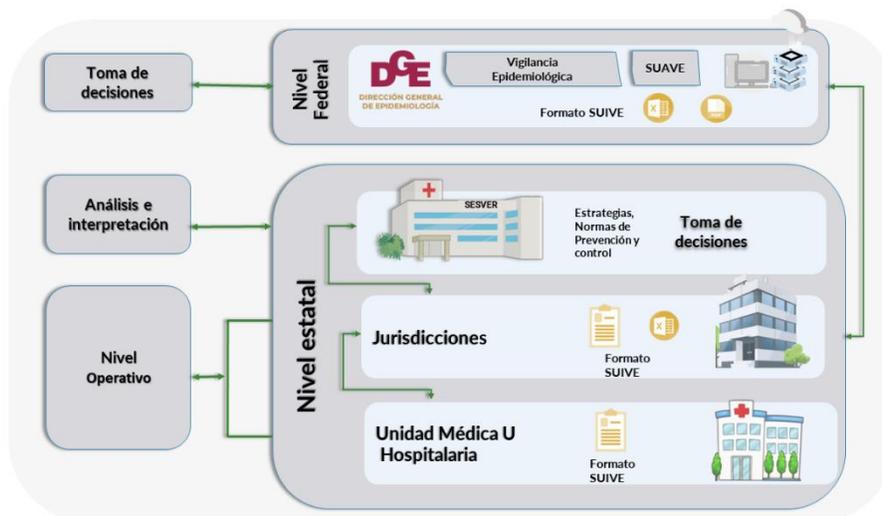


Figura 2. Flujo de información de notificación de vigilancia epidemiológica

Nota. Proceso de SINAVE. Adaptado de (SESVER, 2016).

2.1.2 Marco normativo en materia de salud

La Servicios de Salud a través de la Dirección General de Información en Salud se rige por Normas Oficiales Mexicanas (NM) que, son acuerdos y leyes en materia de salud. Estas permiten realizar regulaciones con la finalidad de establecer los procesos o servicios que de alguna manera permiten la seguridad, evitar dañar la salud humana y el riesgo. (Secretaría de Salud, 2017). Las normas aplicables de referencia para el desarrollo del proyecto se describen a continuación:

- **NOM-035-SSA3-2012:** establece los criterios en los que se establecen la forma en integrar o procesar la información; así como la manera en divulgar la información en salud a través de las herramientas tecnológicas establecidas por el gobierno federal. Así mismo, muestra información estadísticas relevante y prioritaria para que apoye los procesos de toma de decisiones (DGE, 2021c).
- **NOM-024-SSA3-2012:** define los criterios, mecanismos y especificaciones técnicas para el intercambio de información entre Sistemas de Información de Registro Electrónico para la Salud (SIRES), entre los que, se encuentran el Expediente Clínico Electrónico (DGE, 2021c).
- **NOM-017-SSA2-2012:** indica el proceso del SINAVE, en el que muestra los datos que se obtienen del sistema nacional de salud local, jurisdiccional, estatal y federal.

Posteriormente evalúa mediante un análisis las condiciones de salud de la población en vigilancia epidemiológica. Lo anterior permite que la Secretaría de Salud establezca las políticas, normas y programas de salud pública, así como, la información que debe difundirse a través de reportes impresos o electrónicos que favorezcan la accesibilidad los datos (DGE, 2021c)

2.2 Inteligencia de negocios

Actualmente las organizaciones se encuentran en constante cambio, por lo que crean y usan la información de sus datos, con el fin de realizar actividades que con lleven al conocimiento y análisis de información y generen una ventaja competitiva, con respecto al rubro de su actividad empresarial. Como dicen Ahumada y Perusquia (2016), el problema de una organización sin conocimiento es que no se podría organizar a sí misma. A partir de la gestión del conocimiento, surge el concepto de inteligencia de negocios (Business Intelligence, BI). Se llama así a la solución o 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 (Ahumada-Tello et al., 2012).

Pineda y Diaz (2006) indicaban en su trabajo que una solución de Inteligencia negocios, es una plataforma integral de análisis que proporciona accesibilidad a los usuarios a los datos e información. De igual manera, esta plataforma permite a los usuarios examinar los datos de diversas maneras que conllevan a encontrar estrategias comerciales, de control, prevención o tomar decisiones que impliquen una mejora.

Todo objetivo de la inteligencia de negocios o herramientas consiste en transformar datos de la organización en información, para posteriormente generar conocimiento de manera homogénea y confiable optimizando recursos (materiales, tecnológico, económico, humanos) tiempo y toma decisiones.

2.2.1 Proceso de toma de decisiones

Las decisiones que toman las organizaciones deben contar con buena información a partir de sus datos. Estas pueden ser oportunas o insuficientes, debido a que depende de la necesidad o el valor de los datos y criterio la empresa, pues guardan una relación directa, por lo que se debe disponer con ella de manera pertinente. La información es un recurso primordial de la organización, debido a que puede transformar un problema en una decisión coherente y

fundamentada; la cual consiguen tener más probabilidades de una mejora continua o de prevención. Pineda y Díaz (2006) indican que “es primordial comprender la forma como el recurso de la información se relaciona con los datos y con el conocimiento”.

El proceso de toma de decisiones se basa en tres puntos primordiales: los datos se convierten en información y posteriormente en conocimiento como se muestra en la Figura 3:

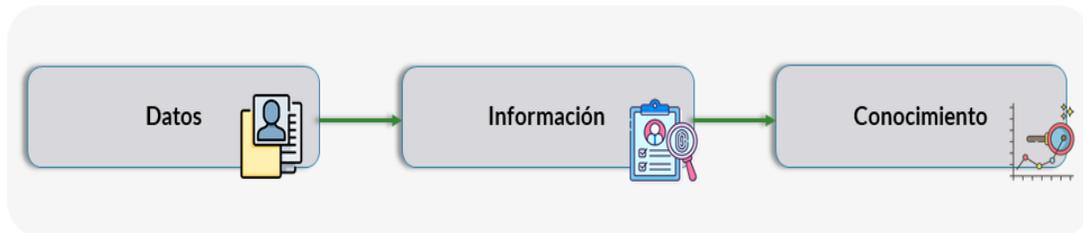


Figura 3. Proceso de toma de decisiones en la Inteligencia de negocios

- **Datos:** son elementos en bruto que por sí solos no dicen nada o no tienen significado para la empresa. Son la mínima unidad para brindar información. Se agrupan o relacionan entre sí. Por ejemplo, pueden representar: la edad, género, padecimiento, municipio, jurisdicción, fecha de ingreso y egreso hospitalario, etc. Los datos provienen de diferentes orígenes para luego procesarlos (cálculos aplicados sobre ellos) o condensarlos (resumidos) (Guillén et al 2015).
- **Información:** es el conjunto de datos transformados, agrupados o relacionados con un significado específico para la empresa. Los datos se convierten en información cuando los usuarios le agregan un valor o sentido mediante la contextualización de estos, es decir, conocer para qué propósito han sido recolectados (De la Puente, 2010). Por lo que un conjunto de datos relacionados de manera lógica y coherente como un grupo de edad o género en una jurisdicción de ETV se convierte en información y sobre todo si se clasifican por grupo de enfermedades o padecimientos.
- **Conocimiento:** se obtiene de la información, así como ésta obtiene del conjunto de datos. El conocimiento es el conjunto de la información de manera organizada que ha sido incorporada a través de estrategias de análisis y evaluada por los directivos que conllevan a un beneficio de la empresa (Guillén et al, 2015). El conocimiento se genera a partir de que la información se pone en un contexto general de la organización y que en un

momento permita el apoyo a la toma de decisiones. Por ejemplo, en SESVER se relaciona con el grupo de diagnósticos o padecimientos o conocer cuales jurisdicciones cuentan con mayor incidencia de casos de ETV; incluso se puede obtener información a detalle por unidad médica, municipio.

2.3 Herramientas de Inteligencia de Negocios

El propósito de las herramientas o técnicas de inteligencia de negocios es coadyuvar a generar conocimiento a las empresas para que permitan administrar y crear el conocimiento mediante el análisis de información (Ahumada-Tello et al, 2012).

2.3.1 Bases de datos multidimensionales

En un simposio en California del año 1963, fue acuñado por primera vez el término base de datos. Es una colección de datos relacionados de manera organizada y estructurada. Feliciano (2016) expresa que debe ser diseñada con una finalidad específica para la empresa y debe estar organizada de manera lógica coherente. Asimismo, los datos pueden ser compartidos por con múltiples usuarios y aplicaciones. Sin embargo, los datos deben mantener fiabilidad y seguridad y los datos deben ser únicos para minimizar la redundancia y maximizar la independencia en su utilización (Feliciano et. al, 2016).

El diseño de una base de datos multidimensional (Data Warehouse) se realiza mediante una técnica o modelo en el que se relacionan hechos y dimensiones. Las tablas de hechos se elaboran dependiendo de las necesidades de la empresa a través del análisis de lo que sucede en la organización. Estas son independientes entre sí y en su esquema de diseño se relacionan las dimensiones (tablas que permiten el detalle de la información).

En este modelo se conservan los valores numéricos de la organización. Cada valor numérico (o dato) corresponde con una intersección con las dimensiones y generalmente se trata de cantidades numéricas. Estas características permiten transformar grandes volúmenes de información en campos definidos y accesibilidad a los usuarios de manera oportuna e inmediata (Feliciano et al, 2016).

En la tabla 1 se muestra la comparación entre base de datos transaccionales y multidimensionales, en otros términos, las diferencias de un modelo transaccional OLTP y modelo de análisis de datos OLAP (Corr & Stagnitto, 2013).

Tabla 1: Cuadro comparativo de base de datos transaccionales y multidimensionales

Criterios	Base de datos OLTP	OLAP Data Warehouse
Propósito	Se ejecuta de manera individual	Evalúa múltiples procesos empresariales.
Tipo de transacción	Insertar, seleccionar, modificar, eliminar	Seleccionar
Estilo de transacción	Predefinido: predecible, estable	Ad-hoc: impredecible, volátil
Optimizado para	Eficiencia de actualización y consistencia de escritura.	Rendimiento y facilidad de consultas.
Frecuencia de actualización	En tiempo real: cuándo ocurren eventos empresariales.	Periódicamente, (diariamente) a través de ETL (extraer, transformar, cargar) programado.
Actualizar concurrencia	Alto	Bajo
Acceso a datos históricos	Períodos actuales y recientes	Actual + años de historia (histórico)
Selección de criterios	Preciso	Difuso y amplio
Comparación	Poco frecuente	Frecuente
Complejidad de consulta	Bajo	Alto
Tablas/uniones por transacción	Pocos (1-3)	Muchos (10+)
Filas por transacción	Decenas	Millones
Transacciones por día	Millones	Miles
Volumen de datos	Gigabytes-Terabytes	Terabytes-Petabytes (muchas fuentes, historia)
Datos	Principalmente datos detallados en bruto	Modelo dimensional
Diagrama del modelo de datos	Diagrama Entidad-Relación (ER)	Esquema de estrella, copo de nieve o híbrido

Nota. Adaptado de (Corr & Stagnitto, 2013)

2.3.2 Tecnologías OLAP

La tecnología OLAP es un conjunto de herramientas y técnicas que facilita el análisis de información en línea de un almacén de datos, conocido por sus siglas en inglés como DW (Data Warehouse). Este proporciona al usuario respuestas rápidas de consultas analíticas complejas, apoya en la toma de decisiones y presenta los datos a los usuarios a través de un modelo intuitivo y natural (Tamayo & Moreno, 2006).

Esta tecnología se usa para el modelado lógico de las bases de datos multidimensionales: permite manejar, administrar, operar y consultar grandes volúmenes de datos; se clasifica en tres categorías de acuerdo con su estructura y funcionamiento como se muestra en la Tabla 2 de comparación de tecnologías OLAP.

Tabla 2. Tabla comparativa de tecnologías OLAP

Criterios	ROLAP	MOLAP	HOLAP
Definición	Relational OnLine Analytic Processing, ROLAP. Procesamiento Relacional en línea.	Multidimensional OnLine Analytic Processing, MOLAP. Procesamiento Analítico en Línea	Hybrid OnLine Analytic Processing, HOLAP. Procesamiento Analítico en Línea híbrido
Arquitectura	3 niveles (base de datos relacional, motor de consultas multidimensionales y procesamiento analítico)	2 niveles (base de datos y procesamiento analítico).	Combina ambas arquitecturas
Almacenamiento	Se almacena en la tabla principal. Los datos se organizan en tablas (filas y columnas).	Se almacenan en la base de datos registrada de manera multidimensional. Se almacenan en cubos.	Se almacenan en las bases de datos relacionales. Se organizan multidimensionalmente
Métodos de extracción y estructura	Se obtienen de un repositorio principal; es escalable y soporta muchas dimensiones,	Se obtienen de una base de datos propietaria; está limitado a tener 10 dimensiones debido a la complejidad	Se obtienen de base de datos relacionales
Volumen de datos	Grandes volúmenes de datos	Conjuntos de datos reducidos.	Grandes volúmenes de datos.
Tecnología	Utiliza consultas SQL complejas para obtener datos del almacén principal	Crea cubos de información pre calculados y prefabricados	Crea cubos para su consulta, sin embargo, puede causar redundancia de datos
Tiempo de respuesta y rendimiento	Máximo, es de menor rendimiento con MOLAP	Mínimo; Es de mayor rendimiento frente a ROLAP	Mínimo
Esquema	Estrella, copo de nieve y constelación	Estrella, copo de nieve, costelación	Estrella, copo de nieve, costelación
Consultas complejas	Limitación, estas se realizan en la memoria caché. Consultas SQL, lenguaje MDX	Diseñado para consultas complejas, consultas con lenguaje MDX	Lenguaje SQL y MDX

Nota. Tomado de referencia (Tamayo & Moreno, 2006)

2.3.3 Cubos OLAP

Un cubo OLAP es una herramienta que permite el análisis de datos de manera dimensional. Los usuarios pueden obtener información desde varios puntos o a nivel de detalle lo que facilita el análisis de información y visualización de los datos. En otros términos, es una bodega de datos, que mejora la consulta de información de manera rápida debido a su estructura y arquitectura.

El cubo OLAP tiene la característica de que los datos pueden estar de forma agregada y a la hora de procesarlo se puede consultar información en un tiempo de respuesta rápida debido a su navegación. Este permite crear medidas, dimensiones y hechos; su navegación puede contener hasta 64 dimensiones (Microsoft, 2019). Las medidas (measures) son valores numéricos que surgen de la información; las dimensiones (dimensions) son un conjunto de datos que permiten la descripción o características que definen los datos y los hechos corresponden a la existencia de las medidas y las dimensiones.

Otra característica particular de los cubos es la jerarquía de agregación, por ejemplo, una jerarquía de tiempo puede darse por año, estación, mes, semana, día, hora. De igual manera el cubo OLAP tiene formas que permiten consultar la información: a) Drill-down, los usuarios pueden analizar los datos en un nivel de detalle, es decir diferente de resumen. b) Drill-through: permite los usuarios a consultar información a un nivel de detalle para un valor dado (Microsoft, 2019).

La Figura 4 muestra un ejemplo de estructura o representación visual del cubo. Cada arista del cubo representa una dimensión que puede contener miles de registros.

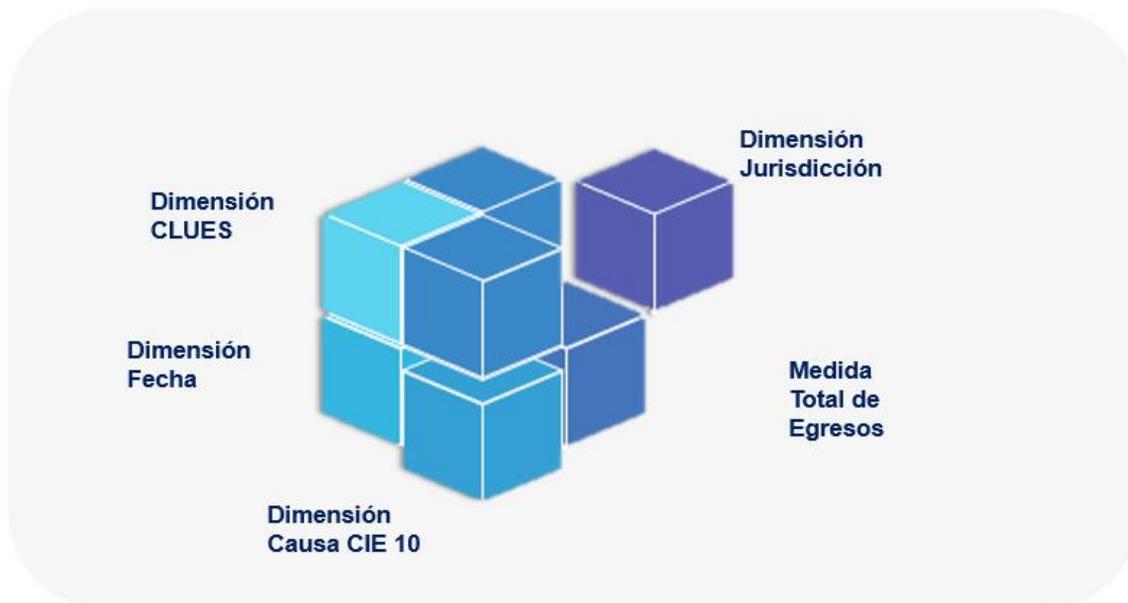


Figura 4. Ejemplo de estructura del Cubo OLAP

2.3.4 Esquemas de modelos multidimensionales

El diseño lógico del modelo multidimensional se representa mediante los siguientes esquemas: estrella, copo de nieve y constelación. El nombre se debe a su representación visual de cada uno de ellos.

En un modelo multidimensional y principalmente en los esquemas se manejan los términos tabla de hechos y dimensiones. Una tabla de hechos por sus siglas en inglés (Fact tables) es la tabla central del esquema: representa datos claves, contiene las medidas y su objetivo es evitar la redundancia de los datos. Esta tabla puede contener millones de registros.

Las dimensiones son tablas desnormalizadas, es decir presentan todos los atributos de la tabla (granularidad). Contienen una llave foránea o clave subrogada numérica, esta debe relacionarse con la tabla de hechos. Las medidas son valores numéricos que se encuentran asociados a un elemento de cada dimensión.

- **Esquema de Estrella:** es sencillo de interpretar y realizar. Su nombre se debe a que su estructura simula una estrella como se muestra en la Figura 5. Está integrado por 4 componentes: una tabla de hechos, una o más tablas de dimensiones, atributos y jerarquías de atributos. Su característica principal es la tabla de hechos y dimensiones se

relacionan entre sí 1: N (de uno a muchos), representado por su nomenclatura de normalización de las bases de datos, es decir, la relación de dimensiones y tabla de hechos se representa mediante llaves foráneas y no hay relación entre dimensiones. Otra característica es que sus tablas se encuentran desnormalizadas con relación a sus campos (este término se refiere a la granularidad de los datos en de la primera forma normal). Una desventaja es que puede presentar redundancia de datos.

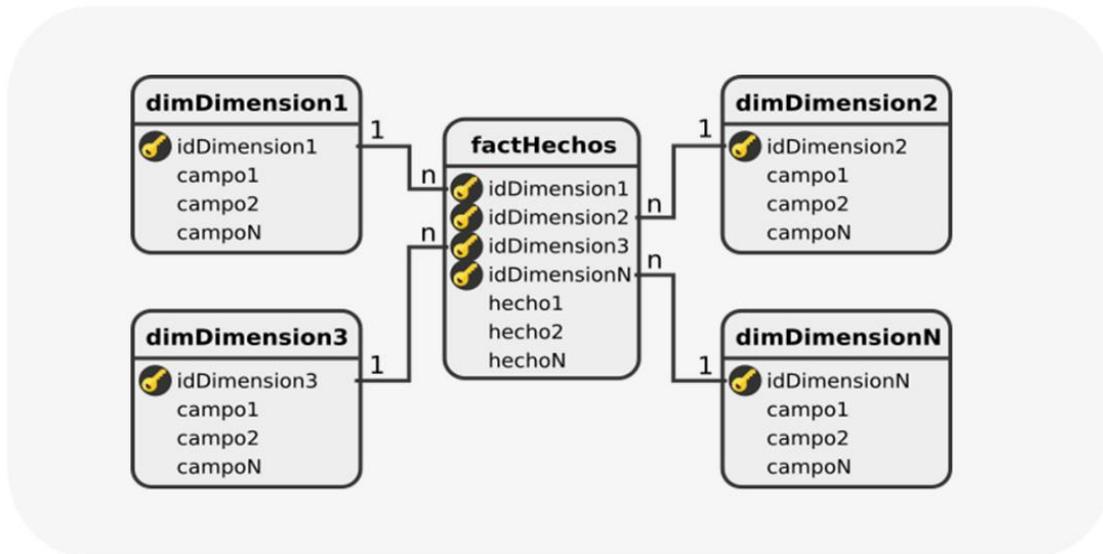


Figura 5. Esquema copo de estrella

Nota. Figura tomada del sitio web (Bernabeu & García, troyanx, 2021a)

- **Esquema copo de nieve:** es una variación del esquema de estrella y de igual manera está compuesto por tabla de hechos, dimensiones, atributos y jerarquías (Figura 6). Sin embargo, en este esquema se puede relacionar entre dimensiones; cada una de ellas almacena las jerarquías de atributos y en esta se usa la normalización de la tercera forma normal representado con las siglas 3FN.

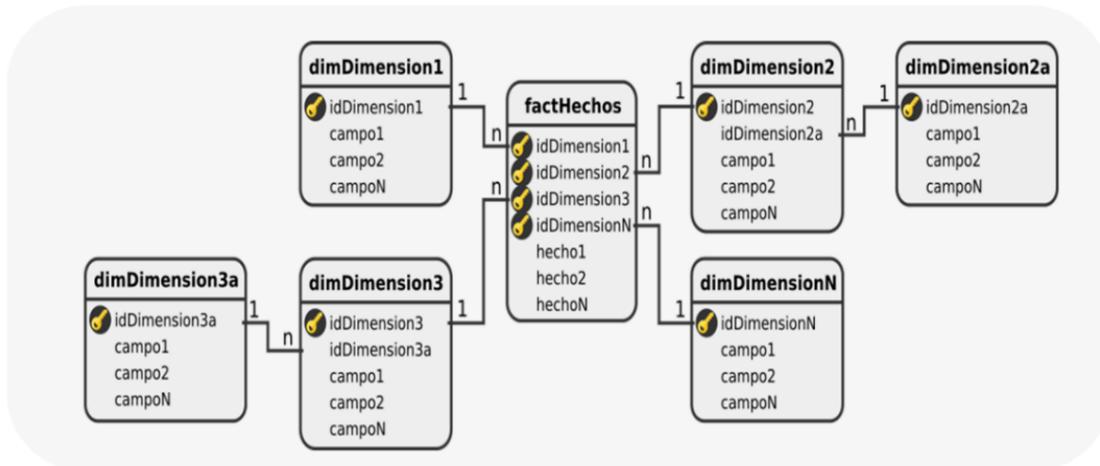


Figura 6. Esquema copo de nieve

Nota. Figura tomada del sitio web (Bernabeu & García, troyanx, 2021b)

- **Esquema constelación:** es también una modificación del modelo estrella (Figura 7). Sin embargo, está compuesto por 5 componentes: una tabla de hechos principal, tabla de hechos secundaria, una o más tablas de dimensiones, atributos y jerarquías de atributos. La tabla de hechos (principal y secundaria) no necesariamente comparten las mismas dimensiones. Las relaciones entre dimensiones son a partir de la primera forma norma 1: N (de uno a muchos). La característica de este esquema es que es útil para evitar la redundancia o la explosión de datos.

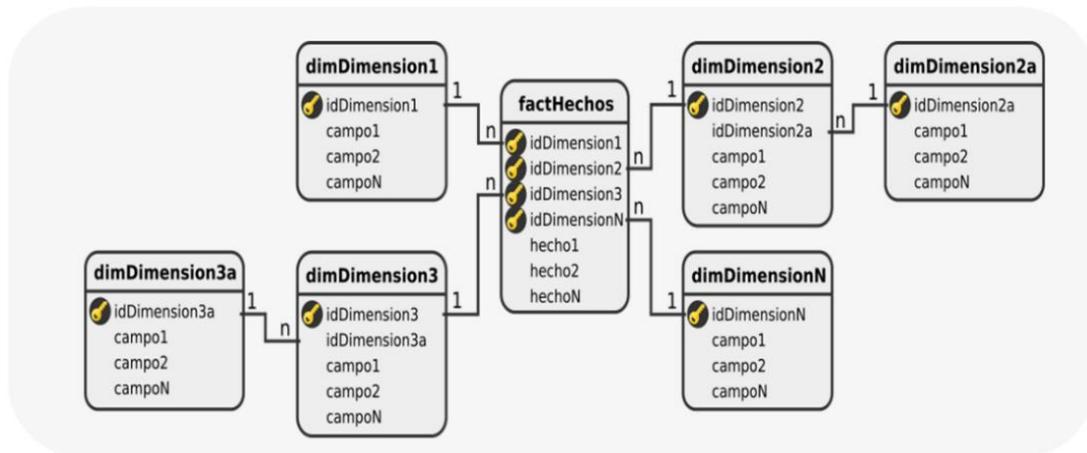


Figura 7. Esquema de constelación

Nota. Figura tomada del sitio web (Bernabeu & García, troyanx, 2021c)

2.3.5 Sistema de Almacén de Datos

El almacén de datos o DW, es un repositorio proveniente de una o varias fuentes de datos y este contiene volúmenes de datos que integran y optimizan la información histórica de meses o años. Una de sus características es que facilita la consulta de los datos y posteriormente proporciona un análisis de datos actuales y comparaciones con datos históricos para el apoyo a la toma de decisiones de la organización.

Las fuentes de datos se clasifican en estructuradas (csv, cadena de texto, números, hojas de cálculo), no estructuradas (archivos PDF, Word) y operacionales que provienen de una base de datos transaccional (OLTP). Los DW filtran, procesan y almacenan datos, proporcionando una solución estratégica a la empresa. Estos permiten a los miembros de la empresa consultar información consolidada e histórica para desarrollar una estrategia o ventaja competitiva (Tamayo & Moreno, 2006).

Los DW tienen una estructura de base de datos multidimensional que contienen una serie de atributos agrupados en dimensiones que, a su vez, hacen parte de un esquema. Las características que debe cumplir un DW es la integración de la información proveniente de varias fuentes de datos. Entre sus características de desarrollo es que no debe ser volátil y tener una variación de tiempo. El DW se alimenta de un proceso ETL (Extract, Transform, Load) que es una acción para extraer fuentes de datos, posteriormente mediante este proceso deben ser limpiados y estructurarlos de acuerdo con la estructura del esquema y posteriormente cargarlos al modelo multidimensional. En la Figura 8 se muestra la arquitectura de un DW.

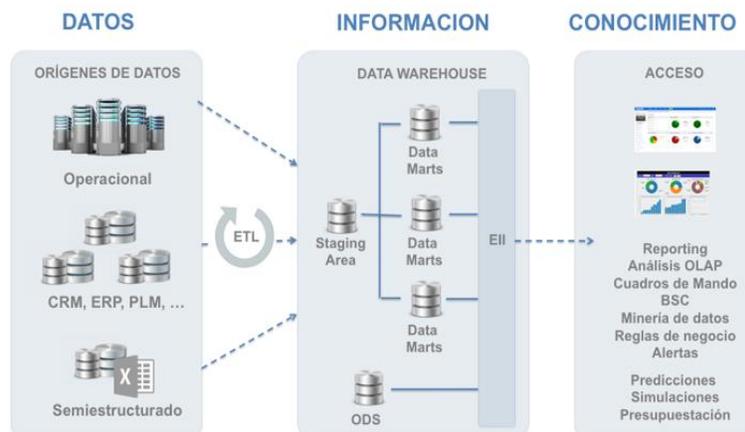


Figura 8. Arquitectura de un almacén de datos

Nota. Imagen tomada del sitio web (Calvo, 2016)

2.3.6 Extraer Transformar y Cargar

Los procesos de extracción, transformación y carga de datos (ETL), se encargan de las actividades clave o específicas de las bases de datos. El ETL permiten transferir los datos de una fuente a otra. Este término está relacionado con los procesos de la construcción del almacén de datos, o Data Warehouse. El ETL es la actividad con mayor costo, tiempo, y recursos. Duque dice que esta “tarea está asociada a la unificación de datos provenientes de diferentes fuentes, con estructuras y formatos variantes” (Ahumada, 2016; Duque et al, 2016).

Como se muestra en la Figura 9, el proceso de extracción de datos se lleva a cabo mediante las fuentes de datos como sistemas transaccionales, ficheros, CSV, etc., que posteriormente en el proceso de transformación consiste limpiar los datos y tratarlos. Por ejemplo, quitar aquellos que vengan vacíos o no definidos; o en su defecto que no estén representados del acuerdo con el tipo de dato. La carga corresponde al proceso de “subida” de información al modelo multidimensional.

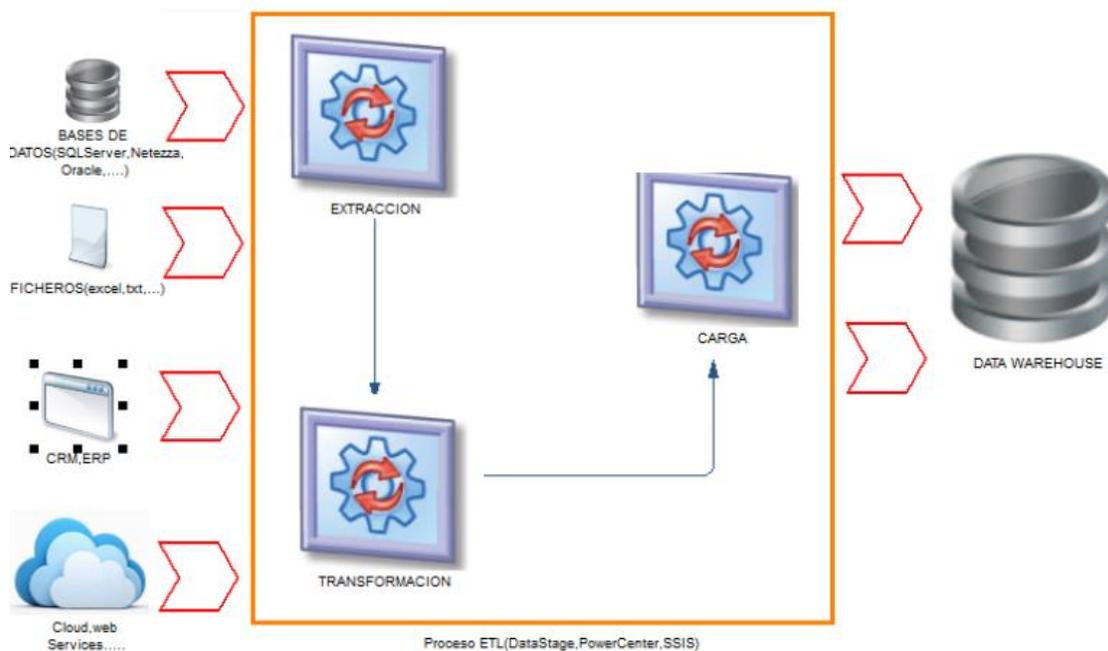


Figura 9. Proceso ETL

Nota. Figura tomada del sitio (Latabernadelbi, 2021)

2.4 Metodologías para el desarrollo del Almacén de datos

Para el diseño del almacén de datos se hace uso de metodologías. Estas proponen técnicas y herramientas para el diseño y desarrollo; lo que permite obtener una arquitectura del DW. Cada metodología una tiene características particulares que las hacen diferentes. En la tabla 3 se realiza una comparativa de las metodologías de Inmmon, Kimball y Corr (Mercado et el., 2018).

Tabla 3. Tabla de comparación de metodologías

Criterio	Inmmon (tradicional)	Kimball (Tradicional)	Corr (Agile)
Ámbito de aplicación	Propietario del producto	Director del proyecto	Equipo
Aproximación	Descendente	Ascendente	Ascendente
Objetivo	Ofrecer una solución técnica sólida basado en métodos probados	Proporcionar una solución que facilite a los usuarios finales la consulta de los datos	Responder al cambio y a las necesidades de los usuarios
Requerimientos de datos	Para toda la empresa	Proceso del negocio	Requerimientos individuales (KPIs)
Modelo de datos	Forma normalizada (3NF)	Modelo dimensional (estrella o copo de nieve)	Modelo dimensional (estrella o copo de nieve)
Tiempo	Mayor tiempo de arranque	Menor tiempo de arranque	Mínimo tiempo de arranque
Capacidad de respuesta al cambio	Alto	Medio	Alto
Participación del Usuario final	Mínimo	Depende del proyecto	Alto

Nota. Adaptado de (Mercado, et al., 2018).

La metodología ágil de Lawrence Corr, permite el diseño, análisis y modelamiento de eventos de negocio. Por sus siglas en inglés (Business Event Analysis and Modeling) BEAM, por lo que identifica las variables a fin de crear el almacén de datos del sistema estratégico. Su proceso consiste en una serie de técnicas y herramientas interactivas y colaborativas. Con ello busca generar la extracción, transformación y carga de estos datos en el almacén de datos para posteriormente presentarlos de manera visual e interactiva (Corr & Stagnitto, 2013).

En esta metodología se combinan técnicas de análisis y diseño que garantizan cumplir con requerimientos relacionados con eventos de negocios y modelado de datos con el fin de diseñar las bases de datos. BEAM involucra a las partes interesadas para definir las variables del almacén de datos a partir de sus narraciones que definen los detalles dimensionales de las actividades de la organización. Posteriormente se determinan las jerarquías de los datos y se obtienen las medidas (cantidades) que deben ser tomadas en cuenta. Con el propósito de obtener estas historias de datos, se hace uso de una técnica basada en 7 preguntas a los usuarios, utilizando un marco basado en las 7W (¿quién? ¿qué? ¿dónde? ¿cuándo? ¿cuántas? ¿por qué? y ¿cómo?). BEAM tiene varios diagramas que permiten realizar el análisis y diseño del almacén: tabla BEAM, gráfico de jerarquía, línea de tiempo, matriz de eventos y esquema de estrella mejorado (Mercado et al., 2018).

2.5 Visualizaciones y procesamiento analítico

La visualización de datos es el último paso en la entrega de una solución de inteligencia de negocios. Consiste en plasmar los datos que se generaron en una imagen, conocida como tablero o visualización, por lo que se convierte en una herramienta que comunica el mensaje. Según Cole Nussbaumer Knaflic “cuando se necesita comunicar algo a alguien mediante el uso de datos, existen maneras efectivas de hacerlo”. (Amador, 2018)

La visualización de los datos contiene al almacén de datos que corresponde a la explotación y análisis de la información; por ello se hace uso de herramientas de la tecnología OLAP. En este concepto se introduce el término de tableros de control o dashboard, lo que facilita a los usuarios finales interactuar con los datos a partir de una interfaz gráfica y de manera interactiva y generar sus propias consultas (Castillo et al., 2018).

Actualmente, existen múltiples herramientas que apoyan a la visualización de los datos de una empresa, lo que permite obtener información de manera rápida y gráfica. Por ejemplo, Gartner, Inc., es una empresa consultora y de investigación de las tecnologías de la información, que cada año publica un cuadrante mágico, lo que facilita a las empresas tener una mejor visión de las características de la tecnología que contrata e implementa. Esta consultora utiliza cuadrantes mágicos en los que clasifica de acuerdo con ciertas categorías: líderes, retadores o aspirantes, visionarios, jugadores de nicho.

Capítulo III: Diseño

En el presente capítulo se realizó la selección de técnicas e instrumentos que permiten el diseño del almacén de datos para el desarrollo y aplicación de la solución de inteligencia de negocios, además del diseño del proceso de extracción, transformación y carga de datos con el fin de realizar procesamiento analítico en línea. El análisis se llevó a cabo mediante el método de observación y descriptivo, basado en la búsqueda de información documental, normas oficiales mexicanas (MX), manuales, lineamientos y en la experiencia práctica en la Subdirección de Análisis e Integración de la Información en Salud.

Finalmente se exponen las técnicas de BEAM que apoyan a la recolección de datos para determinar los requerimientos de los usuarios de SESVER. De igual manera, se determinó la visualización de los datos mediante tableros dinámicos lo que, agrega valor a la empresa. Por consiguiente, se detallan las fuentes de información, instrumentos, población, técnicas y tipo de investigación.

3.1 Material y métodos

En esta sección se definieron los materiales y métodos que se usaron en el desarrollo del Sistema de Información para la Gestión Estratégica en los Servicios de Salud de Veracruz.

Fuentes de información: contienen datos principales y sirven como un acercamiento para comprender el negocio, es decir, permiten identificar actividades y procesos importantes dentro de la empresa. De la información que se obtiene se adquiere un conocimiento a fin de entender mejor la problemática del negocio. Para este caso de estudio se determinaron que las fuentes primarias se componen de los sistemas informáticos clasificados en dos rubros: a) sistemas dinámicos (cubos dinámicos en Excel) y b) estáticos (el boletín epidemiológico), así como de las entrevistas realizadas a los usuarios.

Tiempo: se tomó información estadística de un periodo de dieciocho años comprendido del 2003 al 2020.

Espacio: la solución de inteligencia de negocios se implementará en la Subdirección de Análisis e Integración de la Información en Salud.

Criterios de inclusión: solo se incluyeron los padecimientos del grupo de las 16 enfermedades transmitidas por vector (dengue), de las 947 unidades médicas de los 212 municipios del Estado de Veracruz y de la institución de servicios de salud.

Criterios de exclusión: únicamente se tomó en cuenta los casos notificados por primera vez a través del SUIVE del Estado de Veracruz, correspondiente a las unidades médicas de salud activas y los padecimientos de ETV (dengue con signos de alarma, dengue sin signos de alarma, dengue grave).

Tamaño de muestra: los datos con los que se trabajaron corresponden a los padecimientos transmitidos por el vector *Aedes aegypti* de las 11 jurisdicciones del Estado de Veracruz y los 947 establecimientos activos de salud, con un aproximado de 8 586 048 registros.

3.2 Metodología

La metodología BEAM indica que en cada interacción los stakeholder son parte sustancial para el diseño y vista del modelo dimensional. El modelado de datos lógico y físico para el diseño del almacén de datos se realizó mediante la aplicación de técnicas de BEAM que permiten identificar los elementos. En la Figura 10 se muestran los procesos y actividades que se emplean en el desarrollo de esta metodología. En cada iteración se llevan a cabo las actividades relacionadas al diseño, construcción del DW y visualización de los datos, donde se implementaron técnicas y herramientas de BI, para llevar a cabo el Sistema de Información para la gestión estratégica en SESVER.

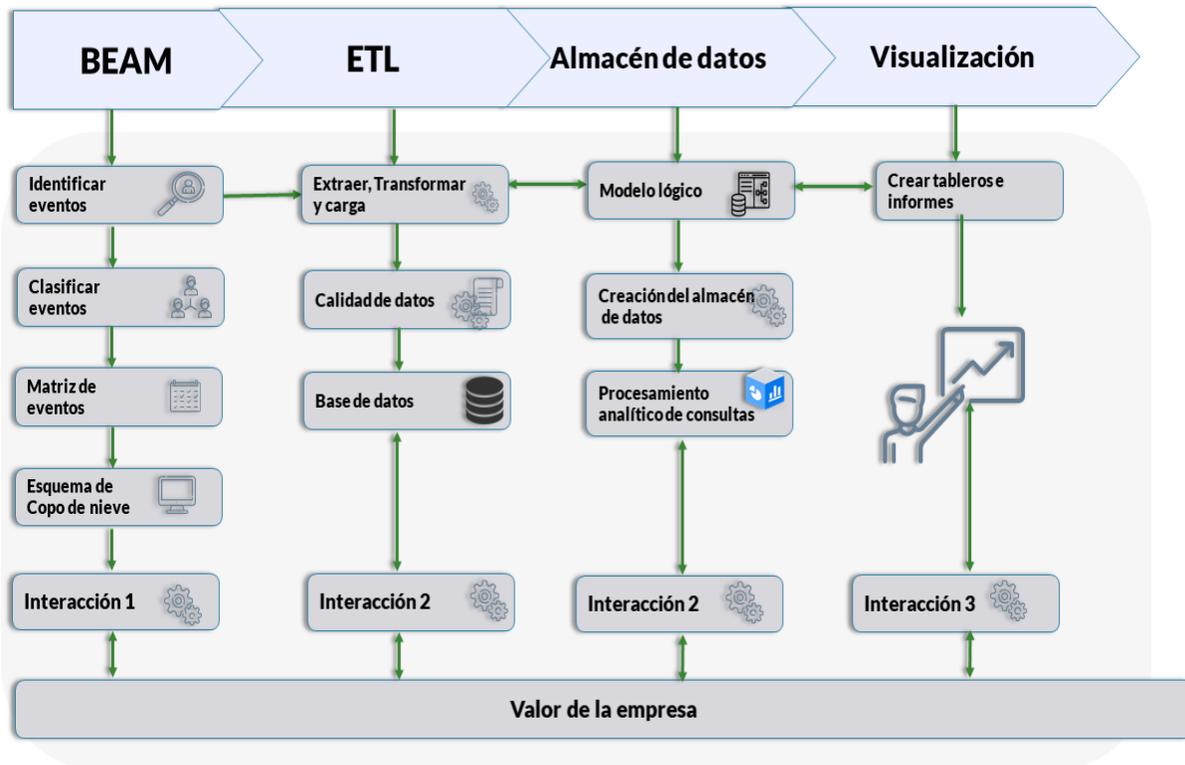


Figura 10. Metodología para la propuesta de inteligencia de negocios

El objeto de estudio y las herramientas de recolección se explican a detalle en el capítulo 4 de resultados, debido a que forman parte de la implementación de la solución de inteligencia de negocios.

3.3 Técnicas e instrumentos para el diseño del almacén de datos

En el siguiente apartado se describen las técnicas e instrumentos para el diseño del almacén de datos de acuerdo con la metodología ágil BEAM, la cual consiste en identificar una serie de eventos y acciones, que posteriormente se clasifican e identifican en tablas (dimensiones y hechos) y finalmente realiza el diseño del esquema de copo de nieve para el almacén de datos. En la Figura 11 se puede observar las técnicas de BEAM empleadas en la propuesta de solución de inteligencia de negocios. Consta de cuatro fases, nueve procesos y cuatro documentos o entregables para el diseño del almacén de datos.

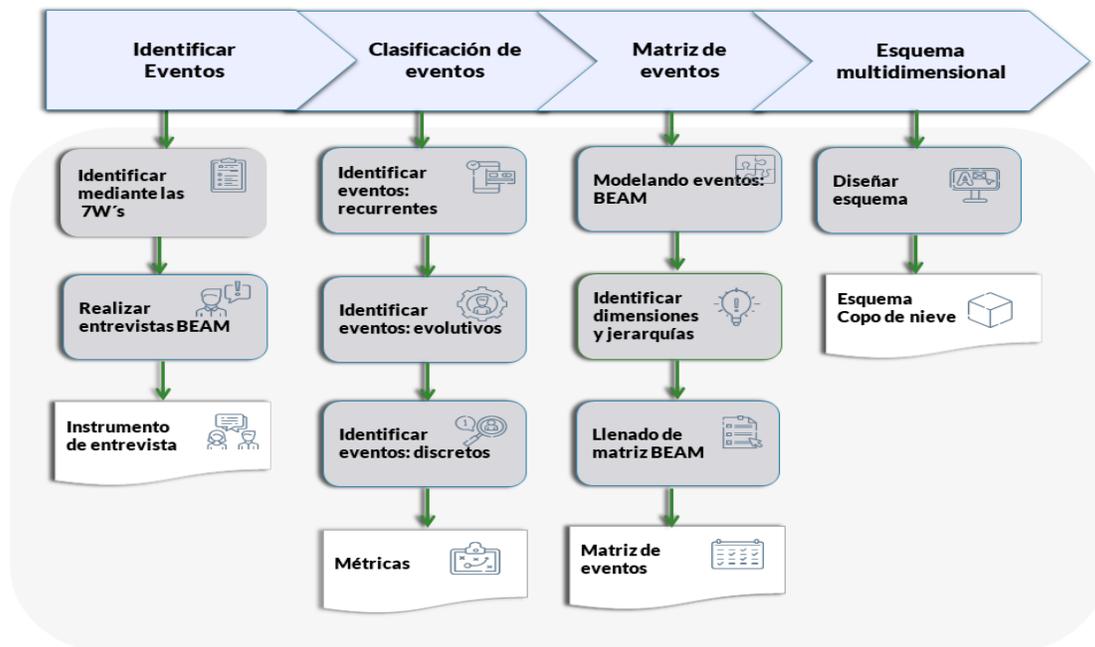


Figura 11. Modelado lógico del diseño dimensional mediante técnicas BEAM.

3.3.1 Fase 1: Identificar eventos

BEAM define los eventos de una organización como las acciones individuales realizadas por personas, departamentos o áreas durante la ejecución de procesos empresariales. Esto conlleva a la comprensión del negocio. Esta técnica proporciona a los usuarios contar sus historias de eventos, lo que permite conocer los detalles medibles de la información (a nivel atómico) para la construcción de los sistemas DW (Corr & Stagnitto, 2013).

Asimismo, se identificaron en esta fase las preguntas enfocadas a determinar las actividades que realizan los perfiles de usuarios de la Subdirección para consultar información de ETV. Dichas acciones, deben especificar el sujeto que las lleva a cabo. Así que, con el fin de descubrir el primer evento basados en la premisa de que la Subdirección elabora informes de los padecimientos de dengue por un periodo específico de tiempo, se determinó la primera pregunta: ¿Cuántos casos de dengue son notificados en la semana?

Identificar 7W's

BEAM utiliza las 7W's para descubrir y modelar los requisitos de datos como detalles de eventos. Cada detalle del evento necesita que las partes interesadas respondan a uno de los siete tipos

de W (preguntas) que, en combinación de ambas forman las historias de los eventos (Corr & Stagnitto, 2013). Con base en el análisis documental se determinó que la subdirección lleva a cabo reportes o consultas de información de los tipos de casos dengue, por género, grupo de edad y lugar. Por lo que se plantean las siguientes preguntas.

1. ¿Cuántos casos de ETV hay en un año, semana epidemiológica, mes, género, grupo de edad, padecimiento, grupo de padecimiento, unidad médica?
2. ¿En qué (año, mes, semana epidemiológica) se presentan más casos de ETV?
3. ¿Dónde hay casos con mayor incidencia ETV?
4. ¿Qué grupo de edad presenta mayor incidencia en casos de ETV?
5. ¿Qué género es el que presenta mayor incidencia en casos de ETV?
6. ¿Qué padecimientos de las ETV son los que tienen mayor incidencia en el estado de Veracruz?

Entrevista BEAM

Para el diseño de la solución de BI y como parte del levantamiento de requerimientos, se utilizó la técnica de sesión de BEAM, la cual consiste en realizar una entrevista basada en las 7W's a los usuarios con datos de interés que procesan información y toman decisiones. Para mayor detalle, la guía de dicha entrevista se encuentra en el anexo 2.

Se identificaron tres tipos de usuarios fundamentales en el desarrollo de procesos del negocio: a) usuarios que consultan información con el fin de generar reportes (semanales y/o anuales) para toma de decisiones; b) usuarios que toman decisiones estratégicamente; c) usuarios encargados de concentrar la información en sistemas informáticos.

La entrevista se realizó a los usuarios que cumplen estos roles en la subdirección antes mencionada. Dichos usuarios se muestran en la tabla 4, detallando el rol y las acciones que realizan. El objetivo de las entrevistas fue obtener una estructura de la tabla de eventos de BEAM.

Tabla 4. Roles y departamentos

Rol	Descripción	Ejemplo de consulta de datos
Jefe del departamento	Toma decisiones estratégicamente, analiza la información y procesos que se realizan en el área. Se encarga de integrar la información de todos los sistemas de salud, para posteriormente realizar los reportes a las áreas correspondientes. Además, el departamento	Corto plazo: informe anual del número de casos de dengue por jurisdicción, para

	atiende las consultas o peticiones que se generan de los usuarios de otras áreas de las enfermedades transmitidas por vector.	determinar medidas de control y prevención.
Administrador de base de datos	Procesa la información de los sistemas informáticos de salud, así como de alojarlos en un repositorio de datos, para que posteriormente se realicen informes.	Mediano plazo: determinar que unidades médicas no han reportado información.
Desarrollador Analítico (DAMA International, 2017)	Estos usuarios son los que solicitan reportes estadísticos para los tomadores de decisiones o para el público.	

3.3.2 Fase 2: Clasificar eventos

BEAM clasifica los eventos empresariales en tres tipos de historias: discretas, evolutivas y recurrentes, en función de cómo se desarrollan sus historias con respecto al tiempo. Los eventos discretos ocurren en un momento dado y pueden estar desconectados (en gran medida) entre sí. Se producen de forma imprevisible. Los eventos recurrentes que representan los lapsos de tiempo se producen de forma escalonada a intervalos predecibles. Los eventos evolutivos representan periodos irregulares periodos de tiempo. Sus historias pueden no haber "terminado" (Corr & Stagnitto, 2013).

De acuerdo con el análisis de las fuentes de información y de los procesos que se llevan a cabo, las unidades médicas notifican semanalmente (cada miércoles), los casos de vigilancia epidemiológica, y reportan información a los sistemas electrónicos que posteriormente se contabilizan de manera semanal, mensual o anual. Por lo anterior, los eventos se pueden clasificar como eventos recurrentes.

De igual manera, se determinó que el número de casos de ETV son la única medida o métrica. Estos son reportados por las unidades a través del SUAVE e identifican a la tabla hechos. Cabe mencionar que la granularidad mínima de tiempo para la consulta de incidencias de ETV es la semana epidemiológica SE (corresponde a un periodo de 7 días lo que permite la comparación de eventos epidemiológicos sucedidos en determinado año o históricos). Por otro lado, la granularidad de las unidades médicas se podrá mostrar en las jerarquías que garantiza la búsqueda a detalle (municipio, jurisdicción y localidad).

3.3.3 Fase 3: Matriz de eventos

Bajo esta estructura, se registran los eventos en secuencias de la cadena de valor y promueven la definición y reutilización de las dimensiones conformadas a través de los modelos dimensionales. Se utilizan en lugar de los diagramas de entidad-relación de alto nivel para proporcionar visiones generales de almacenes de datos. La matriz de eventos constituye la integración de información a través de los datos que garantizan la creación de las dimensiones. Esta puede emplearse de acuerdo con las necesidades del esquema para creación del modelo dimensional, es decir, puede contener todos los eventos de la organización o solo los eventos de interés (Corr & Stagnitto, 2013).

Para el diseño de la tabla de eventos o tabla BEAM, primero se recopilaron los datos a través de la obtención de información documental y entrevistas. Posteriormente, se analizaron con el fin de iniciar esta fase y responder a las 7 preguntas (7w's). En la tabla 5 se muestra los requerimientos a partir de la pregunta, su descripción y los datos de donde se obtendrá la información.

Tabla 5. Requerimientos de 7W's

Pregunta	7W's	
	Descripción	Respuesta
¿Quién?	Personas u organizaciones	Unidad médica, institución
¿Qué?	Cosas (productos o servicios)	Grupo de edad, género, padecimiento, grupo de padecimiento
¿Cuándo?	Tiempo	Semana epidemiológica, mes, año
¿Dónde?	Localización	Localidad, municipio, jurisdicción
¿Cuántos?	Transacciones	Número de casos de dengue
¿Por qué?	Razón y casualidad	Notificación de casos epidemiológicos
¿Cómo?	Medidas e indicadores	Medidas e indicadores

Actividad 1: Modelando dimensiones BEAM

Esta actividad consiste en identificar los eventos que formarán parte de la matriz de BEAM. Estos se identifican a partir de las 7ws y posteriormente apoyan en la documentación de la tabla de

BEAM. En las siguientes tablas se describen los eventos identificados y con un ejemplo de la consulta.

Evento 1: ¿Cuántos casos de ETV hay en una semana/mes/año, género, grupo de edad, padecimiento, grupo de padecimiento, unidad médica?

Ejemplo del evento 1 ¿Cuántos casos de ETV hay en un año de acuerdo con el género y grupo de edad en la jurisdicción de Coahuila de Zaragoza?

Tabla 6. Pregunta del evento 1

7ws	¿Cuántos?	¿Qué?	¿Dónde?	¿Cuándo?
	Número de casos	Grupo de Edad/Género/Padecimiento	Unidad Médica/jurisdicción/localidad/municipio	Periodo, Año, mes, Semana epidemiológica
Ejemplo	62, 000	De 20 a 25 años/ Femenino/ Dengue grave	Coahuila de Zaragoza	2010-2020

Evento 2: ¿En qué (año, mes, semana epidemiológica) se presentan más casos de ETV en el estado de Veracruz?

Ejemplo del evento 2 ¿En qué año (2003-2020) se presentan más casos de ETV en el estado de Veracruz?

Tabla 7. Pregunta del evento 2

7ws	¿Cuántos?	¿Qué?	¿Dónde?	¿Cuándo?
	Número de casos	Padecimiento	Estado	Periodo
Ejemplo	62, 000	Dengue grave, Dengue no grave y Dengue S/C	Veracruz	2003-2020

Evento 3: ¿En qué lugar (unidad médica/jurisdicción/localidad/municipio) se presentan mayores casos de incidencia de las ETV?

Ejemplo del evento 3: ¿En año 2018 qué municipio de la jurisdicción de Coatzacoalcos presenta mayores casos de influencia de los padecimientos del dengue?

Tabla 8. Pregunta del evento 3

7ws	¿Cuántos?	¿Qué?	¿Dónde?	¿Cuándo?
	Número de casos	Grupo de Edad/Género/Padecimiento	Jurisdicción	Año
Ejemplo	222	De 25-a 44 años/ Femenino/ Dengue grave	Coatzacoalcos	2018

Evento 4: ¿Qué grupo de edad es el que presenta mayor número de casos de incidencia de ETV en el 2018?

Ejemplo de pregunta evento 4: ¿Qué grupo de edad es el que presenta mayor número de casos de incidencia de las ETV?

Tabla 9. Pregunta del evento 4

7ws	¿Cuántos?	¿Qué?	¿Dónde?	¿Cuándo?
	Número de casos	Grupo de Edad/Padecimiento	Unidad Médica/jurisdicción/ localidad/municipio	Año
Ejemplo	44	De 20 a 25 años/ Dengue grave	Coatzacoalcos	2018

Evento 5: ¿Qué género es el que presenta mayor número de casos de incidencia de ETV en el 2018?

Ejemplo de pregunta del evento 5 ¿Qué género masculino o femenino es el que presenta mayor incidencia en casos de ETV en la jurisdicción de Coatzacoalcos?

Tabla 10. Pregunta del evento 5

7ws	¿Cuántos?	¿Qué?	¿Dónde?	¿Cuándo?
	Número de casos	Género/Padecimiento	Unidad Médica/jurisdicción/ localidad/municipio	Año
Ejemplo	222	Femenino/ Dengue grave	Coatzacoalcos	2018

Evento 6: ¿Qué padecimientos de ETV son las que tienen mayor número de casos de incidencia en el estado de Veracruz en el 2018?

Ejemplo de pregunta del evento 6 ¿Qué padecimientos de ETV son los que presenta mayor incidencia en casos de ETV en el 2018?

Tabla 11. Pregunta de evento 6

7ws	¿Cuántos?	¿Qué?	¿Dónde?	¿Cuándo?
	Número de casos	Padecimiento	Estado	Año
Ejemplo	2741	Dengue grave, dengue no grave, dengue S/C	Veracruz	2018

Actividad 2: Identificar dimensiones

La fase de *identificación de dimensiones* consiste en encontrar sustantivos que son fundamentales dentro de los procesos de negocios que la empresa lleva a cabo. A continuación, en la Tabla 12, se enlistan los nombres identificados (sustantivos) en forma de tabla, es decir el nombre de “*table*” de la base de datos:

Tabla 12. Tabla de dimensiones

Sustantivo	Descripción
Unidad médica	Unidad de servicio de salud que registra semanalmente el número de casos de dengue a través del formato SUIVE.
Grupo de edad	Rango de edad en el que se encuentra la persona que padece dengue, este sustantivo despliega doce grupos de edad.
Género	Este sustantivo despliega sexo femenino y masculino.
Semana epidemiológica	Rango de tiempo establecido por el SINAVE para reportar el registro del formato SUIVE-1-2014 a nivel jurisdiccional.
Mes	Rango de tiempo donde es factible medir incidencia de casos de dengue.
Año	Rango de tiempo donde es factible medir incidencia de casos de dengue.
Padecimiento	El padecimiento registrado para este caso es el dengue.
Grupo de diagnósticos	Contiene los padecimientos que pertenecen al grupo de diagnóstico clasificados de acuerdo con CIE-10.
Grupo de diagnósticos Nom17	Contiene el grupo de padecimientos que corresponden a los diagnósticos de la Nom 17.

Actividad 2: Identificar jerarquías

Esta actividad consiste en *identificar las jerarquías*, es decir, el nivel de detalle que puede contener el almacén de datos. Por lo que en la Tabla 13 se presentan el nombre de las tres jerarquías identificadas en el cubo dinámico que se analizó anteriormente.

Tabla 13. Jerarquías

Sustantivo	Descripción
Jerarquía por jurisdicción	Contiene el nivel de detalle las unidades médicas agrupadas a la jurisdicción que corresponde.
Jerarquía por municipio	Contiene el nivel de detalle las unidades médicas agrupadas al municipio que corresponde.
Jerarquía por tipología	Contiene el nivel de detalle de las unidades médicas agrupadas por tipología que corresponde.

Actividad 3: Llenado de matriz de eventos BEAM

El llenado de la *matriz de eventos* debe basarse en los eventos (procesos) y sus dimensiones antes obtenidas, así como la secuencia de estos. En los siguientes apartados se indica la forma en la que se llenó la matriz.

- **Secuencia de eventos:** este apartado no se llena de forma alfabética, sino que se toma en cuenta el valor que cada evento posee dentro del proceso de negocio en el cual se desempeña y cada uno se identifica con su pregunta principal. De igual manera se determina que, no es necesario ordenar los eventos de mayor a menor importancia, debido a que todos los eventos pertenecen a consultas con el mismo nivel de valor para SESVER.
- **Secuencia de Procesos:** en esta parte se registran las dimensiones pertenecientes a los eventos. Usualmente se aplica una sangría cada vez que se registra un evento evolutivo. Sin embargo, todos estos se presentan de forma recurrente, por lo cual no es necesario aplicarla. En la Figura 12 se muestra cómo fueron registradas las dimensiones de los eventos junto el desglose de estos.

EVENTOS	Valor/Secuencia/ tiempo	Num Dimensión	Importancia	Unidad Estimada	Institución Médica	Grupo de diagnóstico	Patecimiento	Grupo de Edad	Genero	Semana	Caso de dengue	Mes	Año	Depto. Regionalización	Subdirección	Bioestadística	stakeholder				
					¿Quién?	¿Dónde?	¿Qué?			¿Cuándo?			¿cuántos?								
						100	100					100	100	100	100						
Importancia																					
Estimado																					
¿Cuántos casos de ETV hay en un año, semana epidemiológica, mes, genero, grupo de edad, padecimiento, grupo de padecimiento, unidad médica?	1	1	100	1													✓	✓	✓		
¿En qué (año, mes, semana epidemiológica) se presentan más casos de ETV en el estado de Veracruz?	2	7	100	1	✓				✓	✓		✓	✓	✓			✓	✓	✓		
¿Dónde (lugar) hay casos con mayor incidencia ETV?	3	3	100	1	✓				✓								✓	✓	✓		
¿Qué grupo de edad es presenta mayor incidencia en casos de ETV?	4	3	100	1					✓		✓						✓	✓	✓		
¿Qué género es el que presenta mayor incidencia en casos de ETV?	5	3	100	1					✓	✓							✓	✓	✓		
6.¿Qué padecimientos de las ETV son los que tienen mayor incidencia en el estado de Veracruz?	6	3	100	1	✓				✓								✓	✓	✓		
Total de eventos					6	3	0		0	2	1	1	1	1	6	0	6	6	6	0	0

Figura 12. Propuesta de Tabla de eventos BEAM

Nota. En la tabla de eventos a partir de las 7ws se trata de identificar las dimensiones que puede contener el esquema. En este caso particular cada dimensión esta identificada por un color. Para el caso de las preguntas ¿Quién y dónde? se encuentran en una sola dimensión debido al análisis del cubo dinámico y la granularidad de los datos.

3.3.4 Fase 4: Esquema multidimensional

En este último paso, la metodología BEAM propone el esquema de estrella como parte de una mejor optimización de procesamiento de consultas. Sin embargo, de acuerdo con la tabla de eventos de BEAM, se determinó que la propuesta del diseño de almacenamiento multidimensional para SESVER sea un copo de nieve, pues se puede utilizar para agilizar la obtención de análisis referentes al ETV en el Estado de Veracruz, debido a que las unidades médicas registran información en sus sistemas federalizados y los padecimientos se encuentran clasificados mediante 16 grupos. Las enfermedades del dengue (dengue grave, con signos de alarma y no grave) pertenecen a un grupo de ETV. En la Figura 13 se muestra el esquema de copo de nieve propuesto que consta 9 dimensiones y una tabla de hechos central.

Del mismo modo se observó en las fuentes de información y de las entrevistas realizadas a los usuarios, que el sistema informático para el registro de las ETV tuvo una reclasificación de los grupos de diagnósticos. Por consiguiente, se debe tomó en cuenta las últimas dos dimensiones (grupo de diagnósticos y grupo de diagnósticos Nom17) con el fin de explotar al máximo los datos del DW, por lo que se puede obtener una tener una mayor integración con los demás padecimientos, involucrar más áreas y procesos en materia de salud.

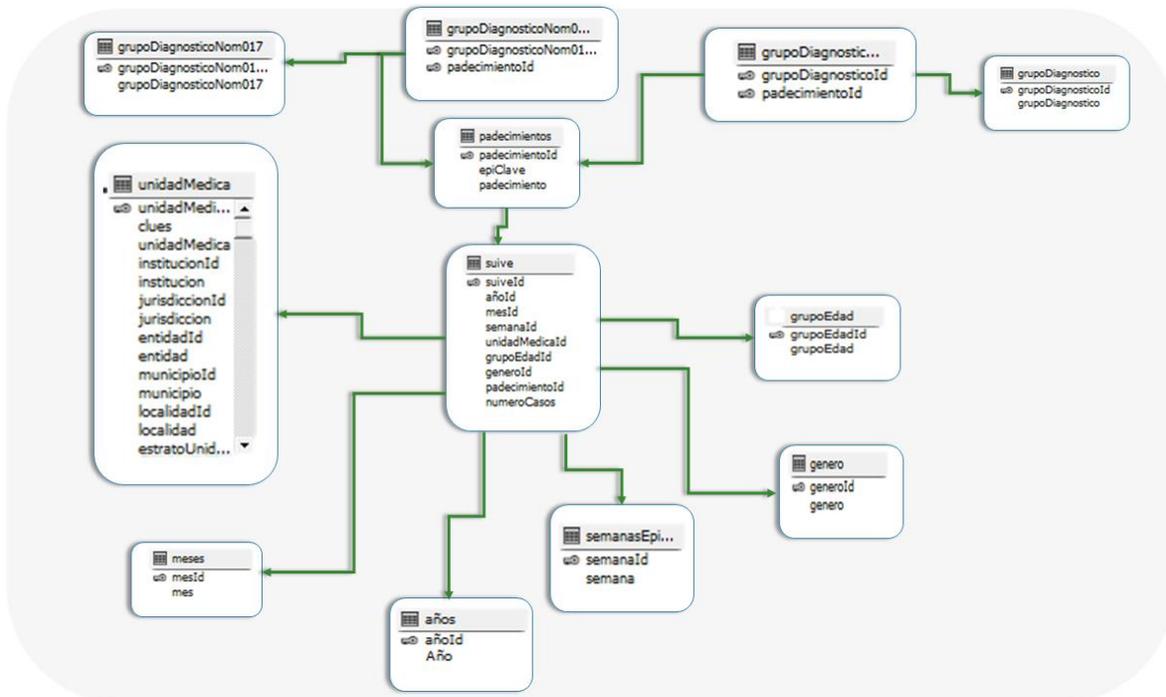


Figura 13. Esquema de copo de nieve propuesto

3.4 ETL

La creación del esquema del DW permitió identificar las fuentes de datos y determinar un proceso ETL que garantice la extracción de los datos y organizar la información permitiendo que los datos obtenidos sean confiables y listos para ser utilizados por el análisis. La integración de estos abre la posibilidad de tener una visión diferente en la visualización de los datos. Por lo que, en este apartado, se realizará una propuesta del ETL que se conectará a la plataforma del sistema federalizado.

En la Figura 14 se muestran las técnicas para la realización del ETL de las ETV, por lo que se proponen tres procesos consistentes: el primero crear el ETL, luego verificar la calidad de los datos y por último, almacenar la información en un repositorio.

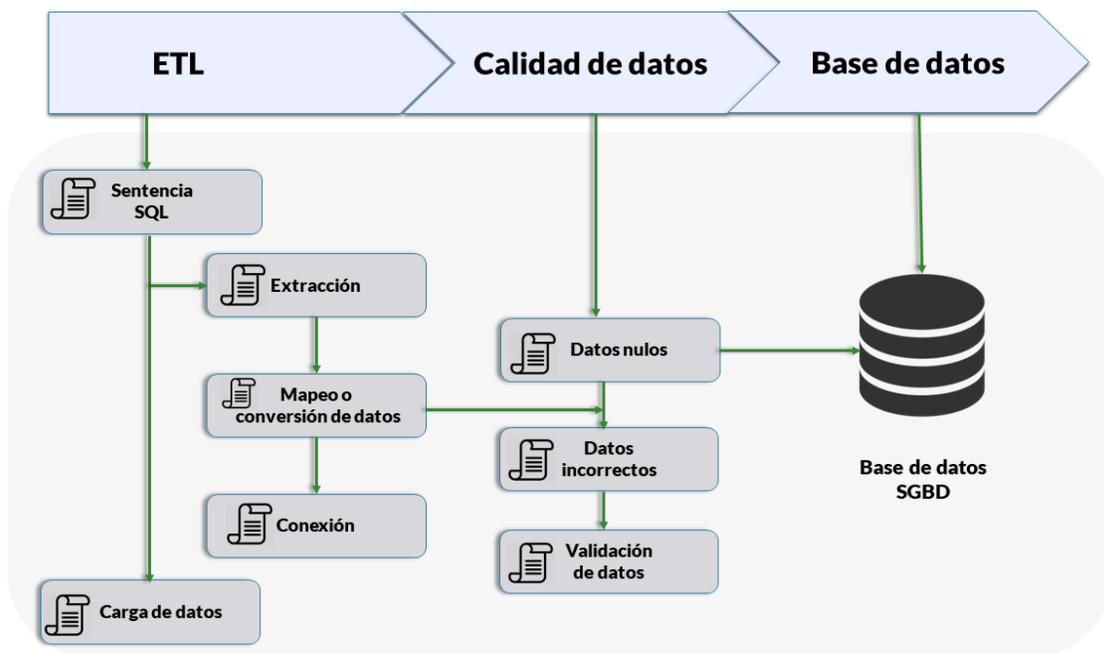


Figura 14. Diseño de ETL de ETV

La Figura 14 muestra el proceso del ETL que se realizó para la extracción transformación y carga de datos que permitió la creación del cubo de datos. Primero se aplicó un proceso de limpieza de datos (truncar). La razón por la cual se hace esta actividad es porque de acuerdo con las entrevistas, los usuarios indican que, las unidades médicas realizan actualizaciones al sistema sobre los casos de dengue (semanal, mensual o por año). Es decir, una unidad médica pudo no reportar incidencias en el mes de enero, sin embargo, en el mes de febrero puede

realizar una modificación y reportar casos anteriores. Este proceso tiene que truncar la base de datos relacional que esta creada en SQL server database.

Actividad 2: Extracción de los datos

Las fuentes primarias fueron proporcionadas por SESVER. El repositorio de datos contiene información estadística de los padecimientos del dengue de las 11 jurisdicciones y un total de 947 unidades médicas. Posteriormente se creó en un nuevo repositorio en el motor de base de datos, a fin de llevar a cabo una validación de datos y luego se realiza una conexión OLE DB mediante una consulta o comando SQL y facilitó la creación del nuevo cubo OLAP.

Actividad 3: Transformación de datos

Esta actividad consistió en analizar los tipos de datos que contiene el repositorio relacional por lo que se llevó a cabo un mapeo de los datos. De igual manera se realizó un script de validación que permitió identificar si existen valores nulos y si la secuencia de los registros es correcta.

Actividad 4: Carga de datos

Posterior a la extracción de los datos a través de la sentencia SQL, se realizó el proceso de carga de datos mediante otra sentencia o *query*, que consistió en elegir la fuente destino en donde se alojaron los datos que permitieron realizar la calidad de datos descritos en la actividad 5.

Actividad 5: Calidad de datos

La norma ISO/IEC 25012 define el concepto de calidad de datos como el nivel en que los datos satisfacen las necesidades de los *stakeholders* (Calabrese, et al, 2019). Por lo cual, una vez que la información se encontró alojada, se analizaron los datos con el objetivo que deben ser confiables y no contengan valores nulos o fuera de rango; incluso que las unidades médicas no se encuentren definidas (nuevas en el catálogo). Por lo anterior descrito se llevó a cabo un script en SQL o un procedimiento almacenado que contenía dichas instrucciones.

3.5 Integración y generación de visualizaciones de consultas

Como parte del objetivo de este proyecto se integró el cubo de datos con una herramienta de inteligencia de negocios llamada Power BI, que permite visualizar las consultas de los usuarios de manera gráfica y dinámica de acuerdo con los requerimientos y problemas encontrados.

En la Figura 15 se muestran los procesos determinados para el diseño y construcción del *dashboard*. En ella se plasman tres fases en donde primero se identifican los tipos de gráficos empleados, las unidades de medida y estilos (colores, logotipos y fuente). La segunda fase consiste en crear los tableros e informes con la herramienta de visualización, así como, definir filtros de información a través de las jerarquías de datos y por último implementar en SESVER el tablero.

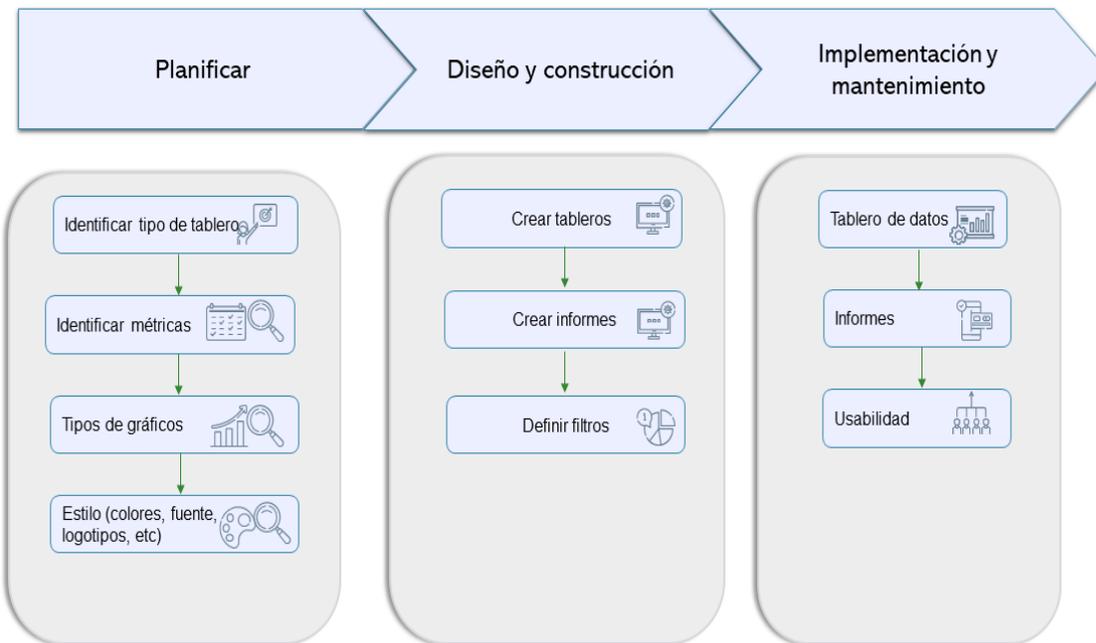


Figura 15. Diseño de visualizaciones de datos

Nota. Imagen adaptada (Martínez Daniel, 2017).

3.5.1 Fase 1: Planificar

En esta fase se llevaron a cabo cuatro actividades que consistieron en identificar el tipo de tablero y los usuarios de SESVER. Por lo consiguiente se realizaron las siguientes preguntas:

1. ¿Para quién va dirigido?
2. ¿Qué es lo que se va a presentar?
3. ¿Cuál es el nivel de detalle?

Además, es conveniente precisar que los tipos de gráficos fueron diseñados para su fácil interpretación, además se incluyó un formato proporcionado por la subdirección debido a que

SESVER cuenta con un manual de identidad, por lo que se agregaron los cintillos o logos institucionales del tablero.

3.5.2 Fase 2: Diseño y construcción

Implica diseñar y construir el tablero de datos de acuerdo con los objetos visuales, por lo cual se optó por seleccionar los gráficos de barras, pastel, cuantificadores, mapa, entre otros, usando la herramienta de visualización; de igual manera se seleccionan los filtros conforme a las jerarquías. Otro aspecto realizado, es la conexión de la herramienta con el modelo multidimensional, por lo que se configura con los instrumentos de la arquitectura de la creación del DW.

3.5.3 Fase 3: Implementación y mantenimiento

Esta etapa consiste en implementar el tablero de datos en SESVER de acuerdo con la arquitectura tecnológica y las técnicas detalladas en el capítulo 2. Sin embargo, es importante evaluar los resultados y si la propuesta es viable, pero sobre todo si responde a la problemática planteada en el capítulo 1. Esta fase se realiza el instrumento de usabilidad cuyo objetivo es evaluar la usabilidad del tablero de datos presentado. En el contexto de la presente investigación, se elaboró un cuestionario compuesto por preguntas sobre usabilidad (Anexo 3), dirigido a usuarios de la Subdirección de Análisis e Integración de Información en Salud y del Departamento de Integración de Programas de Salud, quienes son las personas que poseen los conocimientos necesarios.

Capítulo IV: Resultados

En el presente capítulo se presentan los resultados obtenidos de la implementación del sistema estratégico para SESVER, referente a la problemática identificada en el capítulo 1.

De acuerdo con la metodología propuesta en este proyecto las dos primeras etapas se realizaron con las técnicas de BEAM con el fin de permitir la comprensión del negocio. En las siguientes etapas se hizo uso de herramientas de BI de Microsoft: SQL Data Tool (SSDT), para el desarrollo de base de datos relacionales; SQL Server, a fin de construir el modelo de datos de análisis services. La ventaja principal del uso de estas herramientas es permitir desarrollar con mayor facilidad los procesos de ETL y el procesamiento analítico del cubo de datos.

De igual manera, se determinó que la herramienta para la creación de tableros y visualización es Power BI Desktop de Microsoft, que posee todas las posibilidades de análisis detalladas en este trabajo. Este tipo de tecnología permite al usuario interactuar con los datos utilizando filtros y visualizaciones dinámicas de tal manera que pueda realizar un análisis integral obteniendo sus propias conclusiones.

4.1 Diseño de solución BI en Visual Studio

En esta fase se procedió a la creación de la solución de BI en Visual Studio 2019. Esta solución contiene la integración de las herramientas de SQL Server Data Tools (SSDT), con el fin de desarrollar la base de datos relacionales Analysis Services (SSAS), Integration Services (SSIS) y Reporting Services (SSRS), con el fin de construir un modelo multidimensional. La ventaja principal del SSDT es que permite desarrollar con mayor facilidad los procesos de ETL y diseño del cubo de datos.

4.1.1 Creación de la base de datos

SSDT es una herramienta de desarrollo BI, diseñada para permitir un ambiente interactivo en el entorno integrado de desarrollo Visual Studio. Las características principales es la creación de proyectos basados en inteligencia de negocios, tales como: SSAS, SSIS, SSRS que permiten crear código en el editor de Visual Studio (Microsoft, 2021).

En esta fase primero se creó un repositorio de datos en el motor de SQL Server con la información proporcionada por SESVER. Después se restauró el repositorio para posteriormente analizar y verificar los datos de acuerdo con los requerimientos establecidos. Se combinaron

algunas técnicas de mapeo de datos realizadas en lenguaje SQL, para cerciorarse que no existieran valores nulos o fuera de rangos, es decir se verificó mediante sentencias SQL la validación de datos fueran correctos. Se determinó el tipo de datos de cada tabla (véase el anexo 3). En la Figura 16 se muestra que la restauración de la base de datos fue correcta.

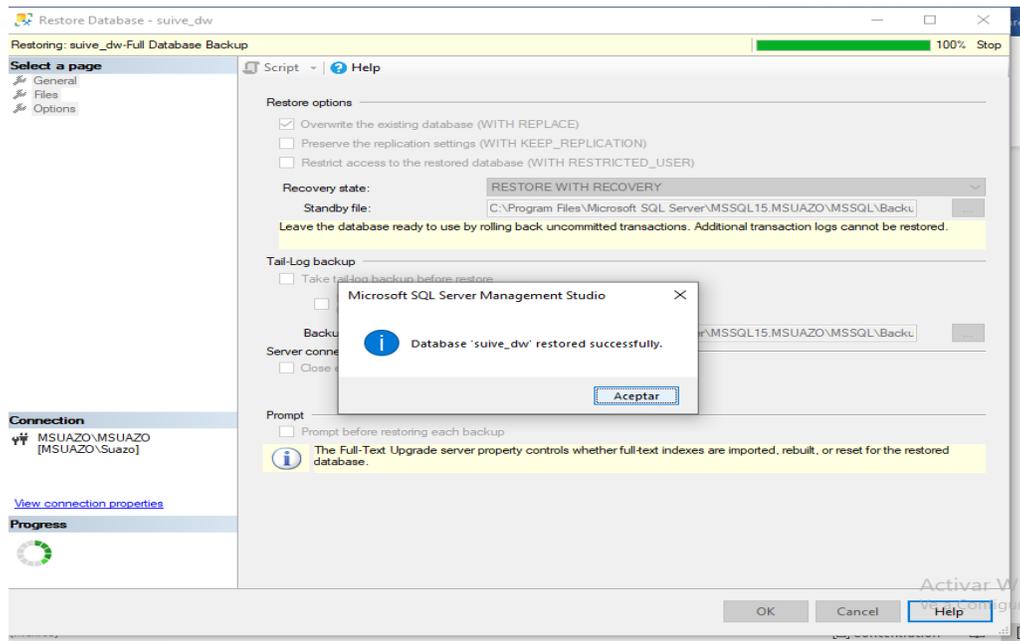
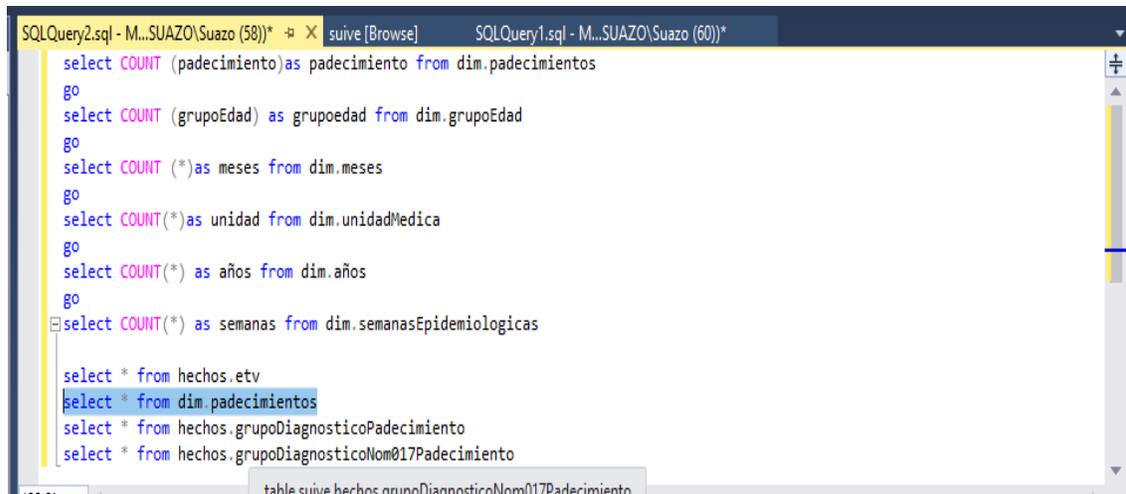


Figura 16. Restauración de base de datos

La Figura 17 muestra parte del script SQL en el que se realiza una consulta al repositorio de datos y la información reportada de la DGIS con el fin de que coincida. La principal razón para llevar a cabo este paso es que las unidades médicas realizan constantes cambios y puede haber discrepancia en la información. Otro motivo es para garantizar que los registros se restauraron sin pérdida de información. Se verifica que no haya registros con valores negativos o fuera de rango.



```
SQLQuery2.sql - M...SUAZO\Suazo (58))* X suive [Browse] SQLQuery1.sql - M...SUAZO\Suazo (60))*
select COUNT (padecimiento) as padecimiento from dim.padecimientos
go
select COUNT (grupoEdad) as grupoedad from dim.grupoEdad
go
select COUNT (*) as meses from dim.meses
go
select COUNT(*) as unidad from dim.unidadMedica
go
select COUNT(*) as años from dim.años
go
select COUNT(*) as semanas from dim.semanasEpidemiologicas
select * from hechos.etv
select * from dim padecimientos
select * from hechos.grupoDiagnosticoPadecimiento
select * from hechos.grupoDiagnosticoNom017Padecimiento
```

Figura 17. Consultas SQL

4.1.2 Creación de la solución

En esta fase se procedió a la creación de la solución de BI en Visual Studio. Esta plataforma de desarrollo contiene múltiples herramientas para crear aplicaciones web y servicios en la nube. En el IDE se creó un nuevo proyecto Analysis Services multidimensional. Este nos permitió la creación del cubo y el proceso del ETL. Posteriormente se llevó a cabo la conexión OLEDB (*Object Linking and Embedding for Databases*) a los orígenes de datos, que en este caso es la base de datos que se describió en el tema anterior.

4.1.3 Vista de esquema multidimensional

La Figura 18 muestra la lista con el nombre del objetos y tipo de la base de datos relacional que debe seleccionarse para el modelo dimensional. En este caso el programa detecta el nombre de la tabla como objeto disponible y determina si el objeto puede seleccionarse como tabla de dimensión o hechos. Por lo que en esta fase se eligieron 8 tablas con sus respectivos atributos: años, género, grupo de diagnóstico, grupo de diagnóstico nom 17, grupo edad, meses, padecimientos, semanas epidemiológicas y unidad médica que conforman las dimensiones y la tabla de suive: tabla de hechos que contiene las relaciones y unidades de medida.

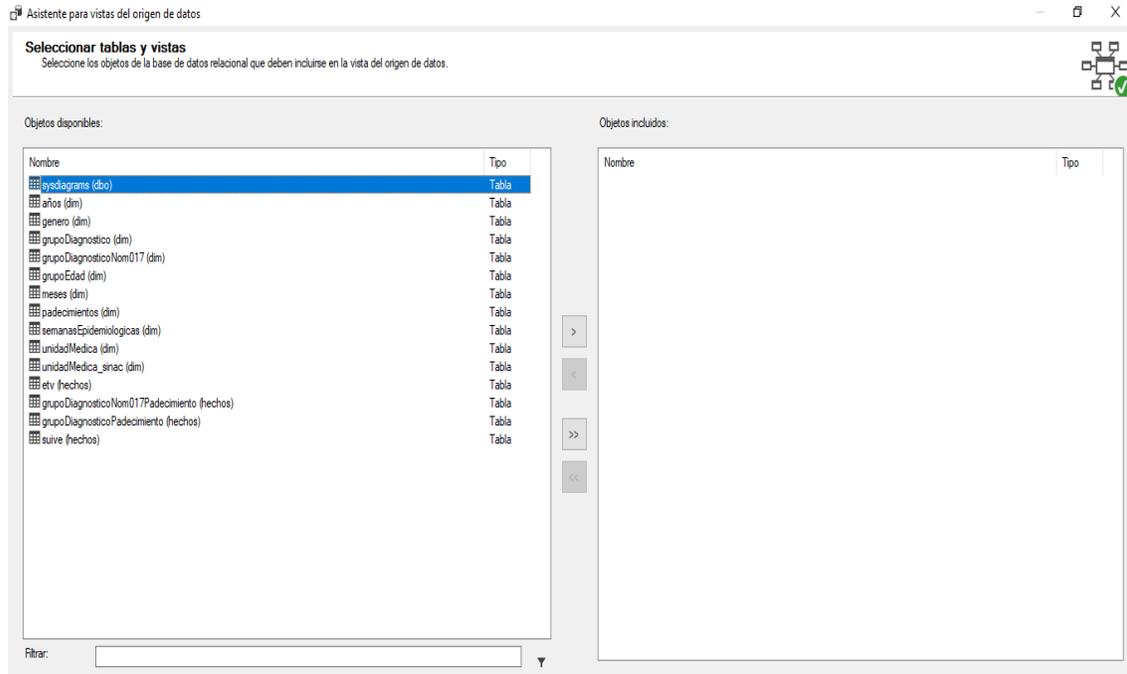


Figura 18. Vista del modelo multidimensional

4.1.4 Creación de dimensiones

Las dimensiones fueron creadas en la solución, posteriormente se identificaron los atributos, tipos de datos y la llave subrogada. Esta es una llave que permitirá la relación entre las tablas de dimensiones o tabla de hechos, sin embargo, debe cumplir las características de ser numérica y no estar asociada a un atributo de la dimensión. En la tabla 1 se describe cada uno de los atributos de la tabla de dimensiones.

Tabla 14 Tabla de descripción dimensiones

Dimensión	Descripción
Años	Esta dimensión cuenta con dos atributos y agrupa los años del 2003 al 2020- Está compuesto por 4 dígitos, ejemplo 2020
Género	La dimensión cuenta con dos atributos y entre sus registros solo hay dos opciones la letra M, que se refiere a mujer y letra H que se refiere a hombre. Esta dimensión permite identificar por grupo los padecimientos del dengue.
Grupo diagnóstico	Contiene el grupo de padecimientos registrados en la CIE-10. Los padecimientos se clasifican en 16 grupos.

Grupo diagnóstico nom 17	Contiene atributos relacionados a los padecimientos de acuerdo con la norma de salud 17. Que es sobre la notificación semanal de casos de vigilancia.
Grupo de edad	Agrupar la clasificación del grupo de edad, de acuerdo con el formato F-114.
Meses	La dimensión contiene dos atributos y está relacionado con los meses del año.
Padecimientos	Contiene los padecimientos y están clasificados de acuerdo con el grupo de diagnóstico.
Semanas epidemiológicas	Contiene las semanas epidemiológicas o de notificación por parte de las unidades médicas de los casos de dengue.
Unidad médica	Contiene todos los datos representativos de éstas y clasificados de acuerdo con la jurisdicción. Se encuentra también datos de localización como lo es longitud y latitud.

Para mayor detalle, la descripción de cada dimensión (tabla) con sus atributos, tipo y longitud, se encuentra en el anexo 3.

4.1.5 Esquema copo de nieve

Una vez analizadas las técnicas de la metodología BEAM, se realizó el diseño lógico (esquema) del almacén de datos mediante la herramienta de Visual Studio. La estructura del esquema fue copo de nieve, debido a que existe una relación entre dimensiones o normalización.

En la Figura 19 se expone la integración del modelo dimensional creado el cual está integrado por 12 tablas clasificadas en: 9 tablas de dimensiones (ver tabla 14), 1 tabla central de hechos (llamada suive) y 2 tablas de relaciones entre dimensiones (grupo de diagnóstico y grupo de diagnóstico nom). La razón por la cual se genera esta relación es porque los usuarios de SESVER requieren consultar los padecimientos por grupo.

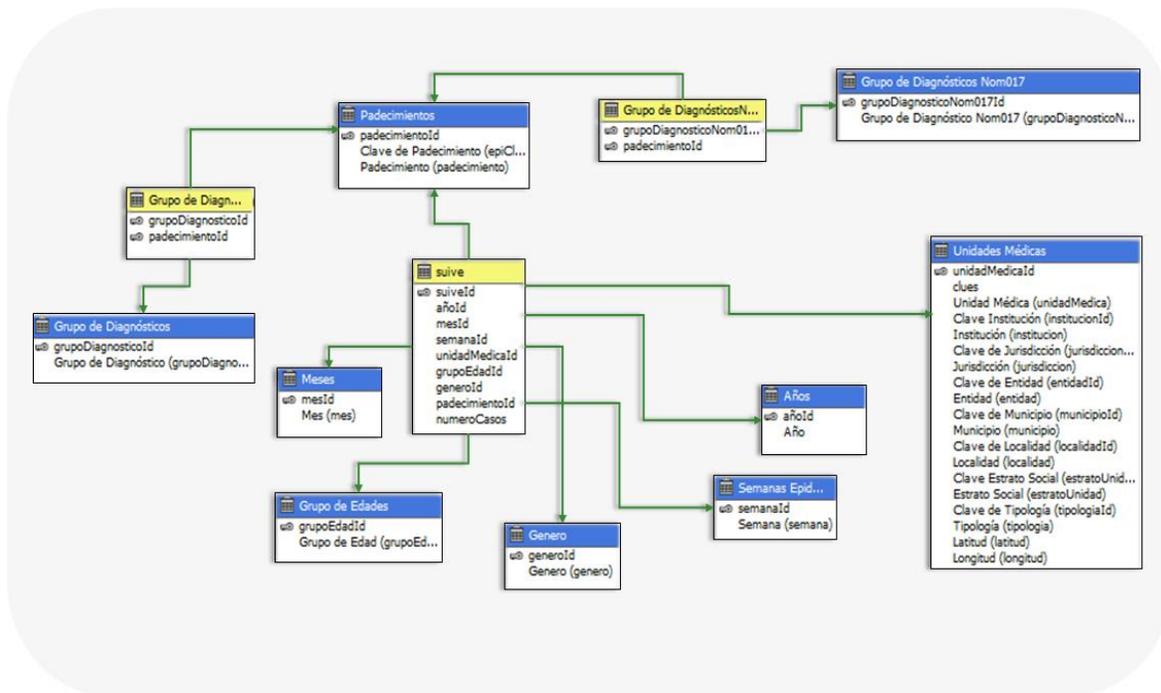


Figura 19. Esquema multidimensional copo de nieve

El modelo dimensional (cubo OLAP) creado en la herramienta Visual Studio, representa a las dimensiones con la etiqueta color azul y en amarillo la tabla de hechos y las relaciones entre dimensiones.

4.2 Diseño y ejecución del ETL

Para realizar el proceso ETL, se utilizó la herramienta de Microsoft SSIS de Visual Studio 2019. El flujo de control es ejecutado por dos procesos: el primero (extracción) realiza una limpieza de los datos que se ejecuta mediante una consulta SQL, la cual elimina los datos que se encuentran alojados en el repositorio de datos y el siguiente conlleva a los procesos de conversión (transformación) o mapeo y carga de datos, a la fuente destino donde se alojarán (SSIS).

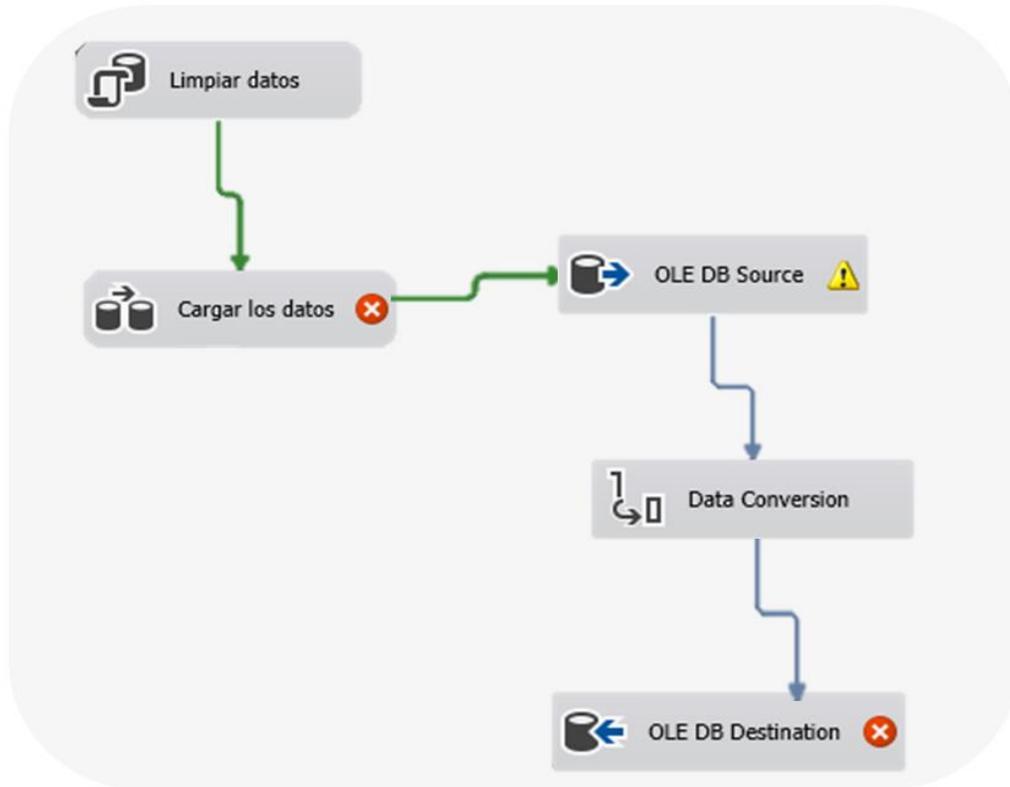


Figura 20. Proceso de ETL de las ETV

La Figura 20 muestra el flujo del proceso de extracción de datos, en el cual lleva a cabo una conexión OLEDB al gestor del repositorio mediante una consulta SQL. Posteriormente se realiza una transformación de estos, debido a que son elementos provenientes de un cubo dinámico y se debe realizar la conversión a los tipos de datos que contendrá el nuevo repositorio (origen), es decir, la consulta de estos se ejecuta en lenguaje DAX y se transforma a expresiones SQL y finalmente, se crea un enlace a la base de datos destino.

4.3 Creación del cubo de datos y procesamiento de consultas

Con el objeto de que los datos puedan ser analizados a través de la creación de consultas multidimensionales, se necesita una estructura de datos determinada. Esta estructura la proporcionan los cubos multidimensionales, los cuales permiten el acceso flexible a los datos, para explorar y analizar sus relaciones y consiguientes resultados.

Una vez que se ejecuta la solución y se carga el almacén de datos, se crea un cubo de datos, que de ahora en adelante será llamado cubo de casos. Para este fin se utilizó la herramienta SQL Server Analysis Services (SSAS).

Una vez creado el modelo dimensional de las ETV (cubo OLAP), se realizan las pruebas. Para esto se utiliza la herramienta de analysis services. Dicha herramienta permite procesar las consultas del usuario mediante el cubo de datos de acuerdo con los criterios de inclusión y exclusión. En la Figura 21 se presenta un ejemplo de la consulta: ¿cuántos casos de dengue notificados por años en SESVER hay en el Estado de Veracruz?

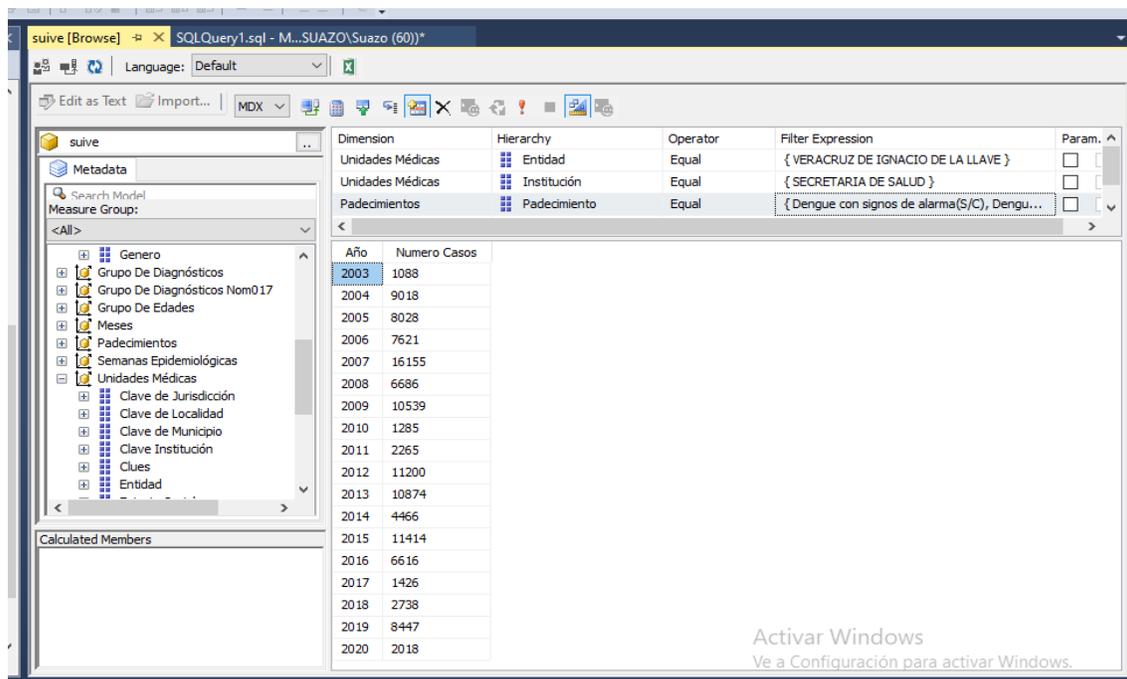


Figura 21. Procesamiento de consultas al Cubo OLAP

4.3 Integración y generación de visualizaciones

Los denominados *dashboard* son la forma de comunicar a los usuarios del nivel estratégico datos estadísticos de información relevante presentada a partir de elementos gráficos sencillos. Su objetivo consiste en apoyar la toma de decisiones o generar ventajas competitivas en la empresa. Dado lo anterior, se propuso la integración de datos mediante la implementación de una herramienta de inteligencia de negocios. Se determinó usar Microsoft Power BI Desktop para visualizar e integrar el cubo de datos. Esta tecnología posee todos los elementos visuales y gráficos y permite a los usuarios diseñar sus propios tableros dinámicos e informes que apoyen

a la toma de decisiones en materia de salud, con un excelente tiempo de respuesta en el procesamiento de consultas. Para su elaboración, primero se seleccionaron los datos de acuerdo con el esquema y jerarquía del modelo dimensional, como la forma en la que se puede acceder a la información de manera detallada o profundizada al DW.

En la Figura 22 se pueden observar las tres jerarquías propuestas del modelo multidimensional, las cuales contienen a nivel de detalle las unidades médicas agrupadas por: a) jurisdicción, b) municipio y c) tipología.



Figura 22. Jerarquía del modelo multidimensional para datos epidemiológicos.

Por último, se crean las visualizaciones requeridas por los usuarios. Para esto se utilizan el programa Power BI el cual se conecta al almacén de datos para acceder a la información y generar las visualizaciones.

En la Figura 23 se muestra el resultado de la propuesta del tablero de datos que contiene gráficos a partir de las consultas realizadas de Analysis Services. En el tablero se puede observar el número de casos de dengue por año del 2003 al 2020. También se pueden apreciar los casos de dengue del 2020, con un total de 2, 053. Pánuco fue la Jurisdicción con mayor número de casos (548) y Orizaba la de menor incidencia (3). Asimismo, se observa que el grupo de edad

más afectado por el dengue es el de 25 a 44 años, mientras el 43.55 % de los casos contagiados corresponde al sexo masculino y el 56.45 % al femenino.

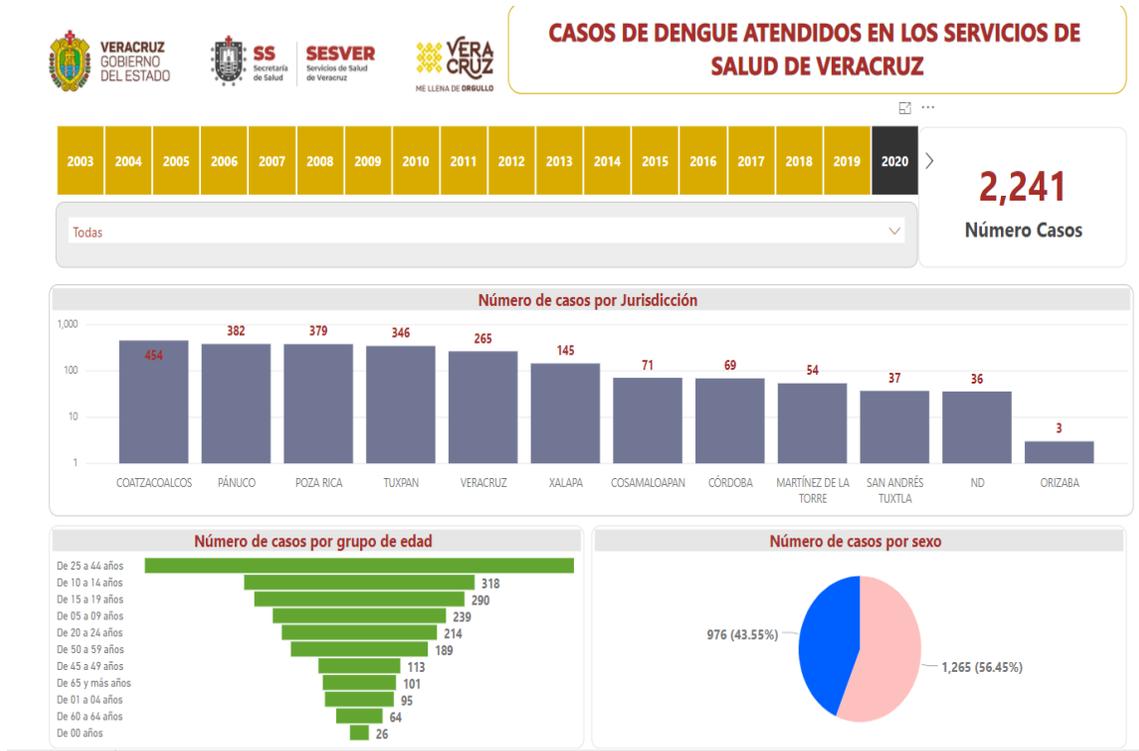


Figura 23. Tablero casos de dengue por años.

En la Figura 24 se muestra un segundo tablero propuesto que contiene objetos visuales geográficos y cuantificadores. En la imagen se muestra que en el año 2020 hay un registro de 2 053 casos de dengue reportados (dengue con signos de alarma, dengue grave y dengue no grave) del Estado. El 43.9% son hombres y el 56.09% son mujeres. También se puede observar que el municipio de Pánuco presenta un mayor número de incidencias (548) en comparación con Orizaba que presentó 3 casos.

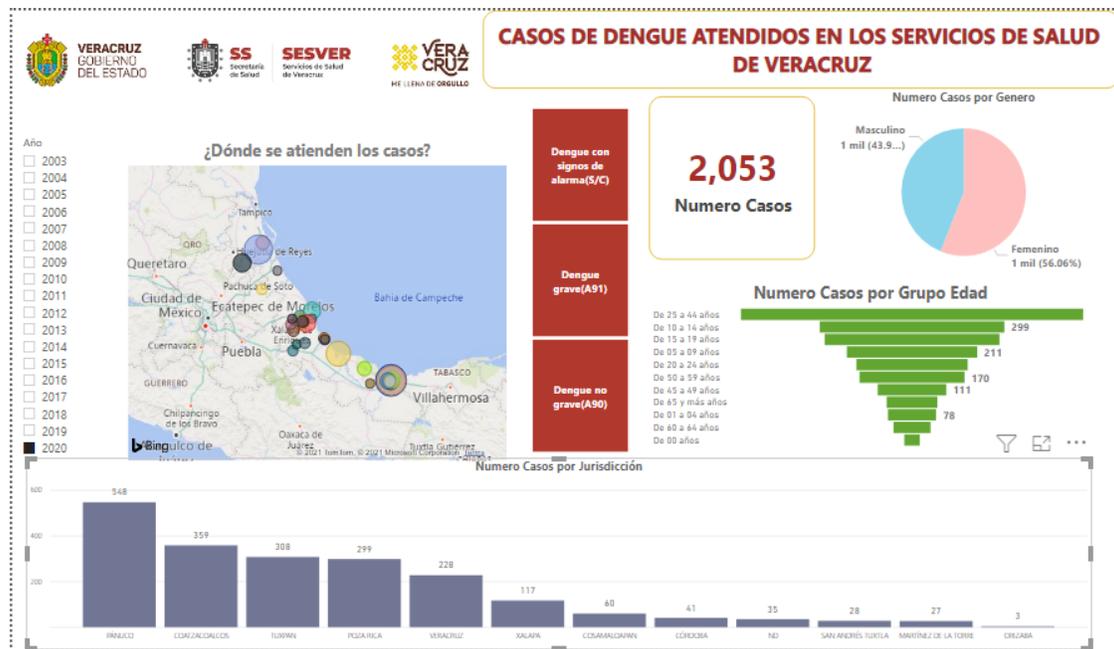


Figura 24. Tablero de datos de casos de dengue geográficos del Estado de Veracruz

En el siguiente gráfico (Figura 25) se muestra un mapa de Veracruz que se encuentra dividido en las 11 jurisdicciones. De aquí se observa que durante el periodo (2003-2020) se reportaron 107, 992 casos en todo el estado y Orizaba fue la jurisdicción que reportó un número de casos menor a 1000. El grupo de edad con mayor contagio es de 25 a 44 años (554).

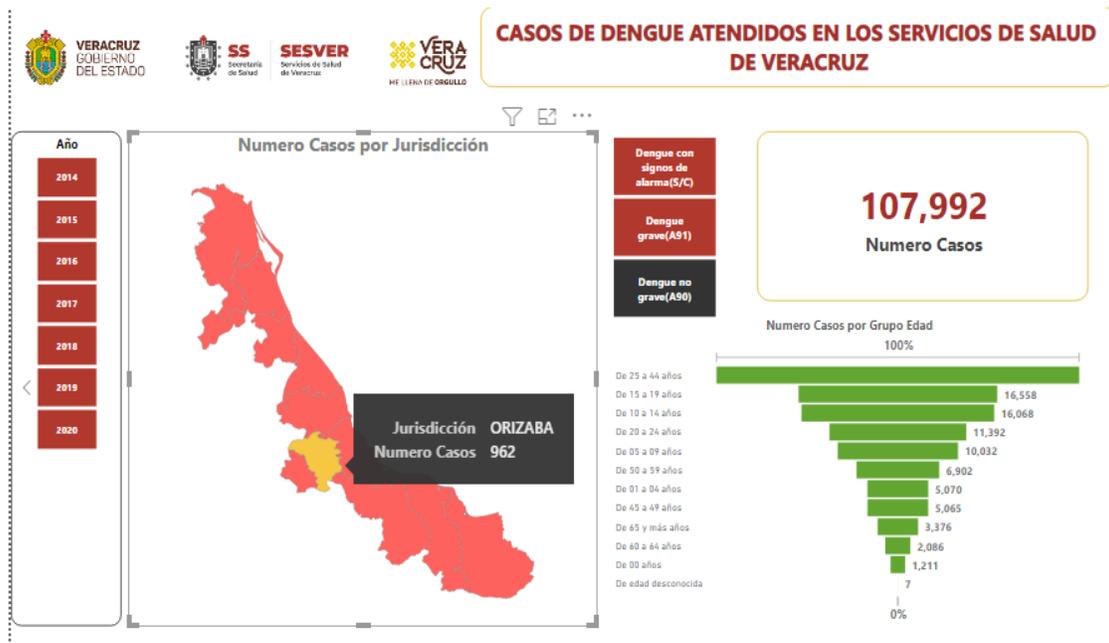


Figura 25. Tablero de datos de casos de dengue por jurisdicción

4.4 Evaluación de la usabilidad en el Tablero de datos de ETV

En principio, la usabilidad no es un término o concepto que se use directamente a los sistemas web, tecnología o aplicaciones; también se refiere a todo aquello (objeto o herramienta) que pueda ser utilizado por una persona o usuario para el fin que fue creado. Como definición se podría decir que la usabilidad es la característica de facilidad de uso que tienen los usuarios para la búsqueda de información o navegación en un artefacto (página web, sistema, aplicación), o de igual manera la consulta de datos bajo ciertas condiciones determinadas (Mascaraque, 2010).

Para llevar a cabo la evaluación de usabilidad del tablero de datos de las ETV. Se diseñó un instrumento de usabilidad. El cuestionario fue aplicado a los usuarios de SESVER quienes conocen los datos, así como quienes generan y consultan la información. El instrumento consta de 13 preguntas sobre aspectos de accesibilidad, usabilidad (satisfacción) y diseño. Para ello, los usuarios encuestados interactuaron con el tablero de datos por dos semanas y posteriormente respondieron el cuestionario aplicado (ver anexo 4).

Los usuarios respondieron la encuesta a una valoración sobre dichas preguntas, empleando las categorías: 1–Totalmente en desacuerdo, 2– En desacuerdo, 3–Indiferente, 4– De acuerdo y 5–Totalmente de acuerdo. Dicho cuestionario fue realizado a través de un formulario web habilitado. En la Tabla 15 se muestran las características de los participantes a los que se aplicó el cuestionario.

Rol	Número
Jefe del departamento	1
Administrador de base de datos	1
Desarrollador analítico	1

Tabla 15. Tabla de descripción de los participantes en instrumento de usabilidad

Los resultados de usabilidad se presentan en la tabla 16. En esta se encuentra la puntuación y se puede observar que tuvo una excelente aceptación. La puntuación media del cuestionario (accesibilidad, satisfacción y diseño) es de 0.8 y su desviación estándar, que permite el grado de confiabilidad fue de 0.1. El tema más crítico fue el tamaño y el estilo de la letra dentro

de los gráficos que se les hace muy pequeño. Incluso observaron que la imagen de los filtros (drill-down), puede ser confusa al principio debido a que es difícil de identificar.

Tabla 16 Puntuaciones a las preguntas del cuestionario de evaluación de la satisfacción del usuario

Pregunta	Media	D.E
1. Creo que me gustaría usar el tablero con frecuencia	0.5	0.3
2. Pensé que el tablero era fácil de usar	1.0	0
3. Creo que necesitaría ayuda de un técnico para poder usar el tablero.	0.5	0.3
4. Encontré las funciones del tablero estaban integradas adecuadamente	1.0	0
5. Creo que hay mucha inconsistencia en el tablero.	0.5	0.3
6. La interacción con el tablero me resulto agradable.	1.0	0
7. El uso del tablero es confuso.	0.5	0.3
8. El sistema facilita los informes sobre el dengue para apoyar la toma de decisiones de los directivos.	1.0	0
9. El tablero muestra gráficos sencillos y de fácil uso del dengue.	1.0	0
10. El tiempo en actualizar la información resulto adecuado.	1.0	0
11. Las funcionalidades del tablero me resultaron fáciles de identificar.	1.0	0
12. En cuestión visual el tablero muestra imágenes de tamaño y letra adecuada.	0.3	0.2
13. Hay elementos que distraen fácilmente	1.0	0
Promedio	0.8	0.1

En los comentarios extra, los participantes solicitaron la posibilidad de consultar más datos de ETV, por ejemplo, Sika y Chikungunya, así como la posibilidad de ellos mismos interactuar con las herramientas para diseñar sus propios tableros.

Conclusiones

El dengue es una problemática en la República Mexicana debido a que genera grandes costos económicos y uso de recursos. El gobierno federal ha desarrollado estrategias y acciones a través del Programa Sectorial de Salud 2020-2024, que consisten en aminorar la propagación e incidencia de las ETV a través de controles epidemiológicos (Diario Oficial de la Federación, 2020) .

Las fuentes de información son un aspecto primario que permiten realizar de forma rápida y sencilla los informes o consultas acerca de la incidencia actual e histórico de los casos de dengue en el Estado de Veracruz. Para SESVER, es fundamental que exista un único repositorio de datos epidemiológicos de manera homogénea, implementado a través de herramientas de inteligencia de negocios que utilicen un tablero de datos dinámico.

La metodología BEAM propone el esquema de estrella del almacén de datos como parte de una mejor optimización de procesamiento de consultas. Sin embargo, de acuerdo con la tabla de eventos de BEAM, el diseño lógico del modelo multidimensional y para agilizar el análisis de las enfermedades epidemiológicas en el Estado, se determinó que la propuesta del esquema de almacenamiento en SESVER sea un “Copo de nieve”. Las unidades médicas registran información federal y los padecimientos se encuentran clasificados en 16 grupos y 147 padecimientos, incluyendo al grupo de ETV, por lo que este esquema es el más conveniente.

El uso de los tableros permitió visualizar las unidades médicas, género, grupo de edad y/o jurisdicciones que presentan mayores incidencias a través de la línea del tiempo. Esto permite a SESVER analizar valores históricos y tener una referencia de la evolución del dengue, colaborando así con la generación de estrategias de control, prevención y seguimiento más acertadas, optimizando la reducción o asignación de materiales, personal e infraestructura.

Trabajo futuro

En la actualidad los sistemas de información tienen un gran interés en generar conocimiento en las empresas y que estas conozcan el valor de sus datos para que puedan tomar decisiones que les permitan generar ventajas o estrategias competitivas.

Conocer los datos del dengue e identificar las zonas con mayor riesgo facilitaría a los directivos del sector salud generar estrategias de control epidemiológico permitiendo en un futuro ahorrar costos económicos, humanos, materiales y tecnológicos. En algunos países como la India y Tailandia, tienen mayor preocupación por encontrar modelos que les permitan realizar predicciones de datos epidemiológicos a través del clima, demográficos, económicos y movilidad, con el objetivo de reducir la tasa de incidencia del dengue.

Por lo tanto, en este trabajo de investigación se pretende abrir la posibilidad de realizar trabajos futuros de investigación que permitan buscar una mayor optimización de los datos de las ETV y que estos respondan ante situaciones de: a) control de vectores, b) realizar mecanismos o estrategias que propicien actividades de prevención y control de asuntos epidemiológicos de gran impacto como lo es el dengue, c) reducir la tasa de incidencia del dengue.

1. Wibowo, et al (2020) en Indonesia han mostrado en sus investigaciones que el clima es un factor considerable para el contagio del dengue debido a las características de este, sin embargo, Jaya y Folmer (2020) mencionan que es importante recabar datos demográficos y de movilidad humana y del mosquito. En Tailandia Kian et al (2021), incluyeron en su investigación datos de teléfonos móviles para realizar predicciones, pero determinaron que existe una relación con los factores demográficos y sociales de la población. Neraini et. al (2021), proponen en su trabajo de investigación en Indonesia realizar predicciones de contagios del dengue, pero menciona que es importante contar los datos personales de expedientes médicos de los pacientes, para que, se pueda realizar una buena predicción de la incidencia del dengue.
2. El empleo de herramientas tecnológicas permite una mejor comprensión de la información para los tomadores de decisión, por lo que se recomienda explorar otro tipo de elementos, metodologías y técnicas que generen un buen resultado.

3. Analizar modelos de predicción que permitan identificar los probables casos de incidencia del dengue en municipios o jurisdicciones. Jaya y Folmer (2020) proponen analizar modelos de predicción que permitan identificar los probables contagios con mayor exactitud.

Referencias

- Academia Nacional de Medicina de México. (18 de junio de 2018). 2. Dr. Hugo López-Gatell *Vacunas humanas contra el dengue [Archivo de vídeo]*. Recuperado el 12 de julio de 2020, de <https://www.youtube.com/watch?v=Q-KIPZRZh5Q>
- Ahumada, E., y Perusquia, J. M. (2016). Inteligencia de negocios: estrategia para el desarrollo de competitividad en empresas de base tecnológica. *Contaduría y Administración*, 61(1), 127-158. doi:<https://doi.org/10.1016/j.cya.2015.09.006>
- Ahumada-Tello, E., Zarate, E., Plascencia, I., & Perusquia, J. (2012). Modelo de competitividad basado en el conocimiento: el caso de las pymes del sector de tecnologías de información en Baja California. *Revista Internacional Administración & Finanzas*, 5(4), 13-27. Obtenido de <https://ssrn.com/abstract=1954073>
- Alcocer, J. (2019). *Caminando a la Excelencia, Segundo Trimestre*. México: Salud.
- Amador, J. (2018). *Desarrollo de una solución de inteligencia de negocios para el proceso de gestión de casos del área de planes de servicio*. Costa Rica: [Tesis Licenciatura, Tecnológico de Costa Rica]. Obtenido de <https://repositoriotec.tec.ac.cr/handle/2238/11058>
- Asavasuthirakul, A. (2018). Spatial Data Warehouse and interactive mapping application for supporting dengue fever surveillance. *Kasem Bundit Engineering Journal*, 8, 327-340.
- Aznárez, F., Renzo, G., y Islas, A. (2015). *Dvelop Report Plus: Generador dinámico de cubos OLAP y reportes con soporte para dispositivos móviles*. Uruguay: Universidad ORT Uruguay.
- Bernabeu, D., y García, M. (2021a). *troyanx*. Recuperado el 22 de 04 de 2021, de <https://troyanx.com/Hefesto/estrella.html>
- Bernabeu, D., y García, M. (2021b). *troyanx*. Recuperado el 22 de 04 de 2021, de <https://troyanx.com/Hefesto/copo-de-nieve.html>
- Bernabeu, D., y García, M. (2021c). *troyanx*. Recuperado el 22 de 04 de 2021, de <https://troyanx.com/Hefesto/constelacion.html>
- Calabrese, J., Esponda, S., Pasini, A., Boracchia, M., & Pesado, P. (2019). Guía para evaluar calidad de datos basada en ISO/IEC 25012. *XXV Congreso Argentino de Ciencias de la Computación*, 694-706.
- Calvo, D. (27 de 07 de 2016). *Diego Calvo*. Obtenido de <http://www.diegocalvo.es/data-warehouse/>
- Castillo-Rojas, W., Medina-Quispe, F., y Molina, F. (2018). Una Metodología para Procesos Data Warehousing Basada en la Experiencia. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, 26, 83-103. doi:<https://doi.org/10.17013/risti.26.83-103>
- Cedeño Trujillo, A. (2006). Modelo Multidimenciona. *Ingeniería Industrial*, 27(1), 15-18. Obtenido de <https://www.redalyc.org/articulo.oa?id=360433560009>

- CENEPRACE. (17 de julio de 2020). *Enfermedades transmitidas por vector*. Obtenido de <https://www.gob.mx/salud/cenaprece>
- Cervantes, A. (2014). *Epidemiología de la fiebre por dengue en Xalapa, Veracruz del año 2008 al 2011*. Xalapa: UV.
- Chaudhuri, S. (1997). An Overview of Data Warehousing and OLAP Technology. *Appears in ACM Sigmod Record*, 26(1), 65-74. doi:<https://doi.org/10.1145/248603.248616>
- CONAVE. (2012). *Manual de Procedimientos para el Comité de Vigilancia Epidemiológica*. México: Secretaria de Salud.
- Corr, L., y Stagnitto, J. (2013). *Agile Data Warehouse Desing*. United Kingdom, Reino Unido: Decisionone Press.
- De la Puente, M. (2010). *Gestión del conocimiento y Minería de datos*. Buenos Aires: Consultora de Ciencias de la Información.
- DeAntonio, R., Amaya-Tapia, G., Ibarra-Nieto, G., Huerta, G., Damaso, S., & Guignard, A. &. (2021). Incidence of dengue illness in Mexican people aged 6 months to 50 years old: A prospective cohort study conducted in Jalisco. *PLOS ONE*, 16(5), e0250253. doi:<https://doi.org/10.1371/journal.pone.0250253>
- DGE. (2019a). *Dirección General de Epidemiología*. Mexico.
- DGE. (17 de 07 de 2020b). *Dirección General de Epidemiología*. Recuperado el 20 de 05 de 2021, de <https://www.gob.mx/salud/acciones-y-programas/direccion-general-de-epidemiologia-boletin-epidemiologico>
- DGE. (18 de enero de 2021c). *Dirección General de Epidemiología*. Recuperado el 18 de 03 de 2021, de <https://www.gob.mx/salud/acciones-y-programas/direccion-general-de-epidemiologia>
- Diario Oficial de la Federación. (17 de 08 de 2020). *Programa Sectorial de Salud 2020-2024*. Obtenido de https://dof.gob.mx/nota_detalle.php?codigo=5598474&fecha=17/08/2020
- Duque, N., Hernández, E., Pérez, N., Arroyave, A., & Espinosa, D. (2016). Modelo para el proceso de Extracción, Transformación y Carga de datos, una aplicación con datos ambientales. *Cencia e Ingeniería Neogranadina*, 26(2), 95-109. doi:<https://doi.org/10.18359/rcin.1799>
- Feliciano, A. C. (2016). Procesamiento Analítico con Minería de Datos. *Iberoamericana de las Ciencias Computacionales e Informática*, 5(9), 22-43. doi:<https://doi.org/10.23913/reci.v5i9.40>
- García Pérez, C., y Alfonso Aguilar, P. (2013). Vigilancia epidemiológica en salud. *Archivo Medico de Camagüey*.
- Gómez Díaz, D. (2005). Gestión del conocimiento y su importancia en las organizaciones. *Ingeniería Industrial*, 37-46.

- Guillén, M. A. (2015). Una revisión de la Cadena Datos-Información-Conocimiento desde el Pragmatismo de Peirce. *Documentación de las Ciencias de la Información*, 38, 153-177. doi:https://doi.org/10.5209/rev_dcin.2015.v38.50814
- Hernández-Ávila, J. E.-L.-C. (2013). Nation-wide, web-based, geographic information system for the integrated surveillance and control of dengue fever in Mexico. *PLOS ONE*, 8(8), e70231. doi:<https://doi.org/10.1371/journal.pone.0070231>
- Jaya, I.G.N.M., y Folmer, H. (2020). Bayesian spatiotemporal mapping of relative dengue disease risk in Bandung, Indonesia. *Journal of Geographical Systems*, 22, 105-142. doi:<https://doi.org/10.1007/s10109-019-00311-4>
- Kiang, M. V.-M. (2021). Incorporating human mobility data improves forecasts of Dengue fever in Thailand. *Scientific Reports*, 11(1), 923. doi:<https://doi.org/10.1038/s41598-020-79438-0>
- Latabernadelbi. (2021). *La Taberna del BI*. Obtenido de <https://latabernadelbi.wordpress.com/etl-extraction-transformation-load/>
- Leonard, E., y Catro, Y. (2013). Metodologías para desarrollar Almacén de Datos. *Revista de Arquitectura e Ingeniería*, 1-12.
- Ley 54 que crea el Organismo público descentralizado Servicios de Salud de Veracruz*. (2015). Xalapa, Veracruz: Gaceta del Estado.
- Mariño, S., y Alonzo, P. (2014). Implementación de SCRUM en el diseño del Implementación de SCRUM en el diseño del. *Scientia Et Technica*, 413-418.
- Martínez, M. D. (2012). Distribución, abundancia y composición de la comunidad de dípteros hematófagos vectores de enfermedades en Castilla-La Mancha: riesgos para la salud pública y la sanidad animal .
- Mascaraque, E. (2010). Accesibilidad vs usabilidad web: evaluación y correlación. *Investigación Bibliotecológica: archivonomía, bibliotecología e información*, 23(48), 61-103. doi:<https://doi.org/10.22201/iibi.0187358xp.2009.48.16970>
- Mercado, K., Pérez, C., López, L., Caro, K., Castro, L., & Rodríguez, L. (2018). Agile Dimensional Model for a Data Warehouse Implementation in a Software Developer Company. *Research in Computing Science*, 147(3), 27-34.
- Microsoft. (05 de 06 de 2019). *Overview of Service Manger OLAP cubes for advanced analytics*. Recuperado el 21 de 08 de 2021, de <https://docs.microsoft.com/en-us/system-center/scsm/olap-cubes-overview?view=sc-sm-2019>
- Microsoft. (21 de 08 de 2021). *SQL Server Data Tools (SSDT) for Visual Studio*. Recuperado el 01 de 09 de 2021, de <https://docs.microsoft.com/es-mx/>
- Morales, A., y Cuevas, R. (2016). Procesamiento Analítico con Minería de Datos. *Iberoamericana de las Ciencias Computacionales e Informática*, 5(9).
- Narro, J., y Gómez, H. (1995). *El Dengue en México, Un problema Prioritario en México de Salud Pública*. México: Salud Pública de México.

- Nerea, M., y Reinoso, A. (2016). Herramientas Basadas en Business Inteligencia (BI) para la Toma de Decisiones en el ámbito de la Gestión Universitaria. *Revista Tecnología y desarrollo*.
- Nuraini, N. F. (2021). Climate-based dengue model in Semarang, Indonesia: Predictions and descriptive analysis. *Infectious Disease Modelling, Volume 6*, 598-611. doi:<https://doi.org/10.1016/j.idm.2021.03.005>.
- Ochoa, R., Casanova, M., y Díaz, M. (2015). Análisis sobre el dengue, su agente transmisor y estrategias de prevención y control. *Archivo Medico de Camagüey, 19(2)*, 189-202. Obtenido de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1025-02552015000200013&lng=es&tlng=es.
- OMS. (2017a). *Respuesta Mundial para el Control de Vectores 2017-2030*. Organización Mundial de la Salud. Obtenido de https://www.who.int/malaria/areas/vector_control/Draft-WHO-GVCR-2017-2030-esp.pdf
- OMS. (20 de junio de 2020b). *Dengue y dengue grave*. Obtenido de <https://www.who.int/es/news-room/fact-sheets/detail/dengue-and-severe-dengue#>
- OPS. (2019). *Actualización Epidemiológica Dengue*. Washington: Organización Panamericana de la Salud.
- OPS. (18 de 06 de 2020). *Organización Panamericana de la Salud*. Obtenido de <https://www.paho.org/es/temas/dengue>
- Ortiz, J., y Ayala, J. (2017). *Inteligencia de Negocios Aplicado a Indicadores Claves de Desempeño (KPI's) para apoyar la Toma de Decisiones en la Gestión del Centro de Distribución de Famaenlace CIA. LTDA*. Ecuador: [Tesis maestría, Universidad Técnica del Norte]. Obtenido de <http://repositorio.utn.edu.ec/handle/123456789/7687>
- PAHO. (1999). *El establecimiento de sistemas de información en servicios de atención de salud*. Washington: Pan American Health Organization.
- Pineda, O., y Diaz, J. (2006). *¿Qué hace que un proyecto de "Business Intelligence" sea exitoso?* Medellín, Colombia: Universidad EAFIT.
- Quintana, Y., Goire, M., y Cantero, A. (2018). Almacén de datos para la vigilancia epidemiológica en la provincia Santiago de Cuba. *Convección Internacional de Salud*.
- Quiroz, J. (2003). El modelo relacional de bases de datos. *Boletín de Informática Núm. 6*, 53-61.
- Ramez, E., & Shamkant B., N. (2007). *Fundamentos de Base de Datos*. Madrid: Pearson.
- Ramírez. (diciembre-marzo de 2008). Modelo de datos del sistema de relaciones interempresariales: RIE. *Revista Ciencia e Ingeniería, 29(1)*, 41-46. Recuperado el 19 de 07 de 2020
- Rodríguez, J. (2002). Las enfermedades transmitidas por vector en México. *Revista de la Facultad de Medicina, 126-141*.

- Rodríguez, Y. (2015). La Gestión de Información y del Conocimiento para la toma de decisiones. *Bibliotecas anales de investigación*, 150-163.
- Ruiz, F., y Vilca, R. (2016). *Desarrollo de una solución de Inteligencia de Negocios en la Oficina General de Registros y Servicios Academicos de la Univesidad Nacional de la Amazonua Peruana*. PERU: Universiad Nacional de la Amazonia Peruana.
- Secretaría de Salud. (2014). *Manual de procedimientos Estandarizados para la notificación para la Notificación Convencional de Casos Nuevos de Enfermedad*. México: Salud.
- Secretaría de Salud. (2017). *Manual de Procedimientos Estandarizados para Vigilancia Epidemiológica de las Enfermedades Trasmítidas por Vector*. México: Secretaria de Salud.
- SESVVER. (2016). *Manual General de Servicios de Salud de Veracruz*. Xalapa: SESVER. Obtenido de <https://www.ssaver.gob.mx/rhumanos/files/2018/10/MANUAL-GENERAL-DE-ORGANIZACI%C3%93N-DE-SERVICIOS-DE-SALUD-DE-VERACRUZ.pdf>
- SESVVER, S. D. (02 de 07 de 2020). SESVER. Obtenido de <https://www.ssaver.gob.mx/funcionario/files/2016/07/ESTRUCTURAS-AUTORIZADAS-2016-1.jpg>
- Tamayo, M., y Moreno, F. (2006). Análisis del modelo de almacenamiento MOLAP frente al modelo de almacenamiento ROLAP. *Ingeniería e Investigación*, 26(6), 135-142. Obtenido de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-56092006000300016&lng=en&nrm=iso. ISSN 0120-5609.
- Torres Gonzales, B., y Martín, R. (2016). *Propuesta de Business Intelligence para mejorar el proceso de toma de decisiones en los programas presupuestales del Hospital Santa Rosa*. Lima, Perú.
- Torres Sánchez, M., Espinosa Cervantes, Y., Simón Cuevas, A., García Hernández , L., & Simón Cuevas, A. (2015). Modelación Tabular: una alternativa sugerente para el análisis de los datos. *Ciencias de la Información*, 3-10.
- Undurraga, E. A.-C.-C.-V.-G.-M.-C. (2015). Economic and disease burden of dengue in Mexico. *PLoS neglected tropical diseases*, 9(3), e0003547. doi:<https://doi.org/10.1371/journal.pntd.0003547>
- Villares Pazmiño, J. D. (2012). *Las Aplicaciones OLAP y su importancia en el Soporte de Decisiones Gerenciales en los Procesos de Compras*. Ambato, Ecuador.
- Wibowo, R. P., Anggraeni, W., Arifiyah, T., Riksakomara, E., Samopa, F., Pujiadi, P., Zehroh, S. A., y Lestari, N. A. (2020). Business Intelligence Development in Distributed Information Systems to Visualized Predicting and Give Recommendation for Handling Dengue Hemorrhagic Fever. *Journal of Information Systems Engineering*, 6(1), 55-69. doi:<https://doi.org/10.20473/jisebi.6.1.55-69>
- Zambrano, C., y Álvarez, C. (2014). *Propuesta para la implementación de una solución de Negocios para una cooperativa del sector financiero*. Colombia: Universidad EAFIT.

Anexos

Anexo 1: Entrevista

Presentación inicial

La siguiente entrevista es para el desarrollo del Sistema de Información para la gestión estratégica en Servicios de Salud de Veracruz que permita la visualización interactiva de las enfermedades transmitidas por vector (dengue). Los datos personales obtenidos se protegerán mediante la ley de protección de datos personales. La información que se obtenga será utilizada con fines académicos.

1. Sobre Actividades:

- a. ¿Cómo describiría un día típico en el contexto de sus actividades? (p.ej. para realizar reportes epidemiológicos)
- b. ¿Qué aplicaciones o herramientas usa para para realizar sus actividades? ¿Cómo las usa? ¿requiere de alguna área externa?
- c. ¿Cuáles son los resultados de esas actividades? ¿Puede contar algunos ejemplos?
- d. ¿Qué influye en la forma en que realiza esas tareas o actividades? (p.ej. que factores tecnológicos, logísticos) ¿Qué considera éxito en ese contexto?
- e. ¿Cómo planifica las actividades? ¿Cómo afecta el tiempo disponible a la organización de las actividades?
- f. ¿Utiliza algunas herramientas o fuentes de información para organizarse en función del tiempo?

2. Sobre la elaboración de los reportes

- a. ¿Realiza o genera reportes epidemiológicos? ¿Cuál es la función o por qué genera los reportes epidemiológicos?
- b. ¿Qué tipos de reportes son los que genera? ¿Con que frecuencia los realiza?
- c. ¿Para quién genera los reportes?
- d. ¿Cuál es la cantidad de datos que presenta en un reporte?
- e. ¿Qué tareas o actividades realizar para generar reportes epidemiológicos? ¿Cuáles son los problemas que encuentra?
- f. ¿Qué tipos de datos son los que los utiliza para generar sus reportes?
- g. ¿Utiliza graficas para la generación de sus reportes? ¿de qué tipo?

- h. ¿Hay alguna información o datos que requiera de alguna área externa para generar sus reportes?
 - i. ¿Hace uso de alguna fórmula u operación matemática para realizar sus operaciones?
- 3. Sobre propuesta:**
- a. ¿Qué recomendaciones daría para que la generación de reportes sea satisfactoria?
 - b. ¿Qué elementos (datos) considera importantes que deberían estar en la generación de reportes?
- 4. Cierre:**
- a. **Muchas gracias por su tiempo empleado y por la información brindada.**

Anexo 2: Matriz o tabla de eventos

EVENT	Value/Time Sequence	Dimension Count	Importance	Estimate												
					who	what	where	why & how	stakeholder group							
Importance	Estimate			0												
		1	0													
		2	0													
		3	0													
		4	0													
		5	0													
		6	0													
		7	0													

Figura 26. Formato matriz de eventos BEAM

Anexo 3: Descripción de tablas modelo relacional

Tabla 17. Descripción detallada de la dimensión de unidades médicas

Unidad Medica			
Nombre campo	Descripción	Tipo Variable	Tipo de datos y longitud
UnidadmédicaID	Es la llave primaria de la tabla	Numérica	Smallint, not null
CLUES	Catálogo de Clave Única de Establecimientos de Salud, la CLUES está constituida inicialmente por dos caracteres alfanuméricos que identifican la Entidad Federativa, posteriormente tres caracteres que corresponden a la Institución, después cinco números que identifican al establecimiento y finalmente un número que es un dígito verificador.	Nominal	Varchar (11) not null
Nombre de unidad médica	Nombre de la unidad médica con base al Catálogo de Clave Única de Establecimientos de Salud.	Nominal	Varchar (150) not null
JurisdiccionID	Identificador de jurisdicción.	Numérica	Int not null
jurisdicción	Corresponde a la unidad técnico-administrativa, que cuenta con recursos y facultades para otorgar atención médica a la población no asegurada, con el propósito de conducir adecuadamente las acciones del sector en su área de influencia.	Nominal	Varchar(50) not null
Municipioid	Identificador de municipio	Numérica	Int not null
Municipio	Corresponde al nombre de la división territorial o administrativa determinada por el catálogo de localidades y municipios establecida por el INEGI.	Nominal	Varchar(150) not null
LocalidadId	Identificador de localidad.	Numérica	Int not null
Localidad	Corresponde al nombre de la división territorial o administrativa determinada por el catálogo de localidades y municipios establecida por el INEGI	Nominal	Varchar(150)
EstratoID	Identificador de estrato.	Numérica	Int not null
EstratoUnidad	Clasificación de la unidad médica (urbano o rural)	Nominal	Varchar(50) not null
Tipología	Clasificación de unidad médica (rural, urbana o no especificada)	Nominal	Varchar(50) not null
Latitud	La latitud proporciona la localización de un lugar.	Numérica	Float not null
Longitud	La longitud proporciona la localización de un lugar.	Numérica	Float not null

Tabla 18. Descripción detallada de la dimensión años

Años			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
Añold	Llave primaria de la tabla	Numérica	smallint
Años	Corresponde al número de año compuesto por 4 dígitos.	Nominal	Varchar (4)

Tabla 19. Descripción detallada de la dimensión de género

Género			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
Generold	Llave primaria de la tabla	Numérica	Smallint not null
Genero	Con base a la referencia de las letras F y M. Corresponde a la clave que corresponda al sexo según condición biológica de nacimiento del paciente: M. Hombre o F- Mujer.	Nominal	Varchar (30) not null

Tabla 20. Descripción detallada de la dimensión grupo de diagnóstico

Grupo de diagnóstico			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
Grupold	Llave primaria de la tabla	Numérica	Smallint not null
GrupoDiganostico	Corresponde a la clasificación del grupo de enfermedades de vigilancia epidemiológica:	Nominal	Varchar (100) not null

Tabla 21. Descripción detallada de la dimensión grupo de edad

Grupo de edad			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
Grupodeedadld	Llave primaria de la tabla	Numérica	Smallint not null

Grupoedad	Conjunto de casos dentro de un rango de edades determinadas, debido a que la edad es una variable demográfica, se realiza de acuerdo con la pirámide de edades en un lapso de 4 años, < de 1 año, 1-4, 5-9, 10-14, 15-19, 20-24, 25-44, 45-49, 50-59, 60-64, 65 y >, Ignoran De igual manera hay un grupo de aquellos que ignoran la edad.	Nominal	Varchar(50) not null
-----------	---	---------	-------------------------

Tabla 22. Descripción detallada de la dimensión meses

Grupo de meses			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
MesesId	Llave primaria de la tabla	Numérico	Smallint not null
Meses	Corresponde al número de mes compuesto por 2 dígitos.	Nominal	Varchar (50) not null

Tabla 23. Descripción detallada de la dimensión padecimientos

Padecimientos			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
Padecimientold	Llave primaria de la tabla	Numérico	Smallint not null
Padecimiento	Contiene el nombre de todos los padecimientos identificados de acuerdo con la Clasificación CIE	Nominal	Varchar (150) not null
Epiclave	Corresponde a la clave de las enfermedades con la que se identifican las enfermedades en el SUAVE (Sistema Único Automatizado para la Vigilancia Epidemiológica).	Numérico	Varchar (50) not null

Tabla 24. Descripción detallada de la dimensión semanas epidemiológicas

Semanas epidemiológicas			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
Semanald	Llave primaria de la tabla	Numérico	Smallint not null

Semana	Corresponde al periodo determinado en el que se agrupan los padecimientos o eventos epidemiológicos. A este periodo se le conoce semana epidemiológica.	Nominal	Varchar(10) not null
---------------	---	---------	-------------------------

Tabla 25. Descripción detallada de la tabla de hechos

Hechos			
Nombre campo	Descripción	Tipo variable	Tipo de datos y longitud
suiveld	Llave primaria de la tabla	Numérico	Int, Not null
añold	Llave foránea de la tabla año	Numérico	Smallint, not null
mesld	Llave foránea de la tabla meses	Numérico	Smallint, not null
semanald	Llave foránea de la tabla semana	Numérico	Smallint, not null
grupoedadld	Llave foránea de la tabla grupo de edad	Numérico	Smallint, not null
generold	Llave foránea de la tabla genero	Numérico	Smallint, not null
padecimientoID	Llave foránea de la tabla padecimiento	Numérico	Smallint, not null
Casos	Sumatoria del total de casos	Numérico	Int, not null

Anexo 4: Diseño de instrumento de usabilidad

Responda al siguiente cuestionario, para ello, debe considerar las siguientes categorías:

1–Totalmente en desacuerdo,

2– En desacuerdo,

3–Indiferente,

4–De acuerdo y

5–Totalmente de acuerdo.

Tabla 26. Cuestionario de usabilidad

Pregunta	1	2	3	4	5
1. Creo que me gustaría usar el tablero con frecuencia					
2. Pensé que el tablero era fácil de usar					
3. Creo que necesitaría ayuda de un técnico para poder usar el tablero.					
4. Encontré las funciones del tablero estaban integradas adecuadamente					
5. Creo que hay mucha inconsistencia en el tablero.					
6. La interacción con el tablero me resulto agradable.					
7. El uso del tablero es confuso.					
8. El sistema facilita los informes sobre el dengue para apoyar la toma de decisiones de los directivos.					
9. El tablero muestra gráficos sencillos y de fácil uso del dengue.					
10. El tiempo en actualizar la información resulto adecuado.					
11. Las funcionalidades del tablero me resultaron fáciles de identificar.					
12. En cuestión visual el tablero muestra imágenes de tamaño y letra adecuada.					
13. Hay elementos que distraen fácilmente					
Observaciones					