



**Universidad Central de Venezuela**  
Facultad de Ciencias  
Escuela de Computación



**SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA GESTIÓN DE UN  
SISTEMA AUTOMATIZADO DE VOTACIÓN**

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela por la Br.

Daimys Carolina Tarazón Simancas

Para optar por el título de Licenciada en Computación.

Tutores:

Profa. Concettina Di Vasta

Prof. Dedaniel Urribarri

Caracas, Octubre 2019



Universidad Central de Venezuela  
Facultad de Ciencias  
Escuela de Computación  
Centro de Investigación en Sistemas de Información

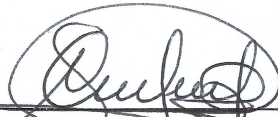



### ACTA DE VEREDICTO


Quienes suscriben, miembros del Jurado designado por el Consejo de Escuela de Computación, para examinar el Trabajo Especial de Grado presentado por la Br. Daimys Carolina Tarazón Simancas C.I. 21.130.596, con el título "SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA GESTIÓN DE UN SISTEMA AUTOMATIZADO DE VOTACIÓN" a los fines de optar al título de Licenciada en Computación, dejan constancia de lo siguiente:

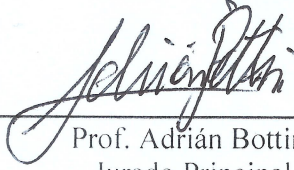
Dicho trabajo, leído por cada uno de los miembros del jurado, se fijó el día 16 de Octubre de 2019, a las 3:00 p.m., para que su autora lo defendiera en forma pública en la Sala de videoconferencias del Centro de Computación, mediante una presentación oral de su contenido, luego de lo cual respondió las preguntas formuladas. Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobarlo con la nota de 19 puntos.

En fe de lo cual se levanta la presente Acta, en Caracas a los 16 días del mes de Octubre del año 2019.

  
\_\_\_\_\_  
Prof. Concettina Di Vasta  
Tutor

  
\_\_\_\_\_  
Prof. Dedaniel Urribarrí  
Tutor

  
\_\_\_\_\_  
Prof. Robinson Rivas  
Jurado Principal

  
\_\_\_\_\_  
Prof. Adrián Bottini  
Jurado Principal

UNIVERSIDAD CENTRAL DE VENEZUELA

FACULTAD DE CIENCIAS

ESCUELA DE COMPUTACIÓN

## **SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA GESTIÓN DE UN SISTEMA AUTOMATIZADO DE VOTACIÓN**

**Autor:** Br. Daimys Carolina Tarazón Simancas.

**Tutores:** Profa. Concettina Di Vasta,

Prof. Dedaniel Urribarri

**Fecha:** Octubre, 2019

# **Resumen**

En la actualidad el éxito de las organizaciones y las aplicaciones que éstas desarrollan dependen de saber aprovechar sus recursos intangibles, es decir, la información y el conocimiento que se puede obtener de los datos generados para un acertado proceso de toma de decisiones. Sin embargo, no todas las organizaciones cuentan con las herramientas necesarias para manipular estos datos y convertirla en información valiosa, de manera que necesitan apoyo de un sistema donde se combinen las herramientas informáticas y se enfatice la Inteligencia de Negocio. Por tal motivo, este trabajo tiene como objetivo desarrollar un Sistema de Inteligencia de Negocio que permita capturar, almacenar, procesar, analizar y mostrar de manera eficiente los datos generados en un sistema automatizado de votación, mejorando el apoyo al proceso de toma de decisiones. Así pues, para el desarrollo de la solución, se utilizó una adaptación de la metodología Scrum y el ciclo de vida de Ralph Kimball, así como una suite de herramientas tales como Oracle, Pentaho Data Integration y Tableau, para la implementación del sistema.

**Palabras Claves:** Business Intelligence, Data Warehouse, Data Mart, Sistema Automatizado de Votación, BI, DW.

# Dedicatoria

Este trabajo y todo el esfuerzo se lo dedico a mi mamá porque gracias a ella he logrado llegar a ser la mujer que soy, una buena persona y buena profesional. Le agradezco enormemente por todo el tiempo y amor que me ha dedicado, así como también por acompañar y guiar mis pasos en todo momento. Por sus trasnochos conmigo durante las entregas de proyectos y sus madrugadas para que estuviese a tiempo esperando el transporte de la universidad. Por sus maravillosos abrazos reconfortantes cada vez que tenía una dificultad, por ser mí ejemplo a seguir y esperarme con los brazos abiertos al salir de un examen importante o una presentación diciéndome siempre que el apoyo es lo más importante. Por todo eso y mucho más, Te Amo Mamá.

# Agradecimientos

A mi mamá, por ser la que guía mis pasos con todo el amor del mundo y la que me dio todas las enseñanzas para convertirme en una buena persona, con valores y saber resolver mis problemas. Por apoyarme siempre en todas las decisiones de mi vida y acompañarme en esta hermosa etapa de mi vida.

A mi esposo, por todo su apoyo y constante amor. Por la paciencia que me tuvo en los momentos de estrés y buenos consejos para continuar adelante. Así como también por consentirme tanto y hacerme reír.

A mi tutora Concettina, por confiar en mí, creerme capaz, apoyarme, aconsejarme y darme ánimos en los momentos difíciles. Así como también por todos los conocimientos que me transmitió a lo largo de la carrera.

A mi tutor Dedaniel, por confiar en mí, brindarme su apoyo y un espacio en un proyecto para llevar a cabo esta actividad, y por darme ánimos para continuar adelante a pesar de las dificultades.

A mis amigos de la universidad y compañeros de trabajo por los abrazos recibidos, hacerme reír en cada momento de estrés y animarme a continuar con mis actividades.

A todos los profesores y personas que conocí durante mi trayecto en la universidad, porque cada uno me dejó una importante lección y conocimiento tanto de crecimiento profesional como personal.

Y a Dios, por darme salud y que todo se haya dado en el momento adecuado y bajo las condiciones indicadas para que salga bien.

# Índice General

Resumen .....	3
Introducción .....	12
Problema de Investigación .....	14
1.1. Planteamiento del Problema .....	14
1.2. Objetivos del Trabajo Especial de Grado .....	16
1.2.1. Objetivo General .....	16
1.2.2. Objetivos Específicos .....	16
1.3. Solución Propuesta .....	16
1.4. Justificación de la Solución .....	17
1.5. Alcance de la Solución .....	18
Marco Conceptual .....	19
2.1. Sistema Electoral .....	19
2.2. Total Vote Tracker .....	19
2.3. Sistema de Información .....	22
2.3.1. Tipos de Sistemas de Información .....	22
2.3.2. OLAP .....	23
2.3.3. OLTP vs. OLAP .....	23
2.4. Inteligencia de Negocio .....	25
2.4.1. Características .....	25
2.4.2. Arquitectura de Inteligencia de Negocio .....	26
2.5. Indicador .....	33
2.5.1. Objetivos de los Indicadores .....	34
2.5.2. Tipos de Indicadores .....	34
2.6. Oracle Database .....	35

2.7.	Herramientas para el desarrollo de una Solución de Inteligencia de Negocio .....	36
2.7.1.	Tableau .....	36
2.7.2.	Qlik .....	37
2.7.3.	Microsoft Power BI .....	37
	Marco Metodológico .....	39
3.1.	SCRUM .....	39
3.2.	Ciclo de Vida Dimensional de Kimball .....	44
	Marco Aplicativo .....	54
4.1.	Fase I: Inicio del Proyecto .....	54
4.1.1.	Planificación del Proyecto .....	55
4.1.2.	Análisis de la Base de Datos Relacional .....	56
4.1.3.	Definición de los Indicadores de Gestión .....	57
4.1.4.	Selección de Productos de Software .....	58
4.2.	Fase II: Desarrollo .....	60
4.2.1.	Modelo dimensional .....	61
4.2.2.	Diseño Físico .....	64
4.2.3.	Desarrollo de los datos simulados .....	64
4.2.4.	Diseño y desarrollo del ETL .....	65
4.2.5.	Desarrollo de la aplicación BI .....	66
4.3.	Fase III: Finalización .....	80
4.3.1.	Prueba de aceptación .....	81
	Resultados Obtenidos .....	88
	Conclusiones y Recomendaciones .....	89
	Referencias .....	91
	Anexos .....	94

A.	Diccionario de Datos del Data Warehouse.....	94
B.	Tabla de Indicadores.....	104
C.	Instrumento de aceptación – Módulo de reportes de TVT.....	105



# Índice de Ilustraciones

Ilustración 1: Arquitectura Propuesta.....	17
Ilustración 2: Suite de aplicaciones de TVT .....	20
Ilustración 3: Flujo de Procesos de TVT.....	21
Ilustración 4: Tipos de Sistemas de Información .....	23
Ilustración 5: Arquitectura de una Solución de BI.....	26
Ilustración 6: Ejemplo de Granularidad .....	28
Ilustración 7: Ejemplo de Jerarquía.....	29
Ilustración 8: Ejemplo de Tabla de Hechos .....	29
Ilustración 9: Ejemplo de Dimensión.....	30
Ilustración 10: Ejemplo de un Modelo Dimensional con Esquema Estrella.....	31
Ilustración 11: Ejemplo de un Esquema Copo de Nieve.....	32
Ilustración 12: Ejemplo de un Esquema de Constelación de Estrellas .....	32
Ilustración 13: Fases del método Scrum .....	43
Ilustración 14: Ciclo de Vida Dimensional de Kimball .....	44
Ilustración 15: Adaptación de la metodología Scrum y el Ciclo de Kimball.....	54
Ilustración 16: Arquitectura de la Solución BI .....	58
Ilustración 17: Interfaz de Pentaho Data Integration .....	59
Ilustración 18: Interfaz de Tableau .....	60
Ilustración 19: Modelo Dimensional del Proceso de Negocio: “Historial de Encuestas” .....	61
Ilustración 20: Modelo Dimensional del Proceso de Negocio: “Totalización de Votos” .....	62
Ilustración 21: Modelo Dimensional del Proceso de Negocio: “Historial de Incidencias” .....	63
Ilustración 22: Modelo Dimensional del Proceso de Negocio: “Captura de Actas”.....	64
Ilustración 23: ETL para la Dimensión de Medios de Votación.....	65
Ilustración 24: ETL para la Tabla de Hechos del Proceso de Negocios: “Captura de Acta” .....	65
Ilustración 25: ETL para la Dimensión Tiempo .....	66
Ilustración 26: Job para el proceso de ETL.....	66
Ilustración 27: Ejemplo de conexión con la fuente de datos desde Tableau.....	67
Ilustración 28: Esquema estrella - Actas.....	67
Ilustración 29: Esquema estrella – Incidencias .....	68

Ilustración 30: Esquema estrella – Encuestas .....	68
Ilustración 31: Esquema estrella – Votos.....	68
Ilustración 32: Dashboard de Encuestas – Total de Respuestas .....	70
Ilustración 33: Dashboard de Encuestas – Total de Encuestados por Región .....	71
Ilustración 34: Dashboard de Encuestas – Total de Encuestados por un Encuestador .....	72
Ilustración 35: Dashboard de Votos – Total de Votos por Medio .....	73
Ilustración 36: Dashboard de Votos - Total de Votos.....	74
Ilustración 37: Dashboard de Votos – Afluencia .....	75
Ilustración 38: Dashboard de Incidencias .....	76
Ilustración 39: Dashboard de Actas – Captura de Actas por Región.....	77
Ilustración 40: Dashboard de Actas – Captura de Actas por Estatus.....	78
Ilustración 41: Dashboard de Actas – Captura de Actas por Usuario.....	79
Ilustración 42: Prueba de Aceptación – Pregunta 1 .....	82
Ilustración 43: Prueba de Aceptación – Pregunta 2 .....	82
Ilustración 44: Prueba de Aceptación – Pregunta 3 .....	83
Ilustración 45: Prueba de Aceptación – Pregunta 4 .....	83
Ilustración 46: Prueba de Aceptación – Pregunta 5 .....	84
Ilustración 47: Prueba de Aceptación – Pregunta 6 .....	84
Ilustración 48: Prueba de Aceptación – Pregunta 7 .....	85
Ilustración 49: Prueba de Aceptación – Pregunta 8 .....	85
Ilustración 50: Prueba de Aceptación – Pregunta 9 .....	86
Ilustración 51: Prueba de Aceptación – Pregunta 10 .....	86
Ilustración 52: Prueba de Aceptación – Pregunta 11 .....	87
Ilustración 53: Prueba de Aceptación – Pregunta 12 .....	87

# Índice de Tablas

Tabla 1: Descripción de las funcionalidades de TVT .....	21
Tabla 2: OLTP vs. OLAP.....	24
Tabla 3: Procesos de Scrum .....	43
Tabla 4: Planificación del Proyecto .....	55
Tabla 5: Iteraciones Scrum para el desarrollo de la solución BI.....	60
Tabla 6: Preguntas de la Prueba de Aceptación .....	81
Tabla 7: Tabla Dimensión Pregunta.....	94
Tabla 8: Tabla Dimensión Encuestado .....	94
Tabla 9: Tabla Dimensión Opción .....	95
Tabla 10: Tabla Dimensión Tiempo .....	95
Tabla 11: Tabla Dimensión Encuesta .....	96
Tabla 12: Tabla Dimensión Entidad .....	96
Tabla 13: Tabla Dimensión Usuario .....	97
Tabla 14: Tabla Dimensión Perfil.....	98
Tabla 15: Tabla Dimensión Incidencia .....	98
Tabla 16: Tabla Dimensión Elección.....	99
Tabla 17: Tabla Dimensión Organización .....	100
Tabla 18: Tabla Dimensión Postulante .....	101
Tabla 19: Tabla Dimensión Medio .....	101
Tabla 20: Tabla Dimensión Acta .....	102
Tabla 21: Tabla de Hechos Encuesta .....	102
Tabla 22: Tabla de Hechos Voto.....	103
Tabla 23: Tabla de Indicadores .....	104

# Introducción

Hoy en día, muchas organizaciones han comenzado a entender que el término Business Intelligence (BI) o Inteligencia de Negocio, hace referencia al uso de estrategias y herramientas que sirven para la toma de decisiones, ya que permite transformar los datos en información útil para la organización y, con ello, generar conocimiento para tomar mejores decisiones.

Al mismo tiempo, uno de los puntos que se debe tomar en cuenta es que BI está enfocado en cubrir necesidades de información de orden táctico (Gerentes y/o jefes) y estratégico (Alta Dirección). Generalmente, estos usuarios deben decidir en tiempos muy cortos y para ello requieren de información confiable y oportuna que les permitan tomar decisiones favorables para la organización.

Por consiguiente, cuando se decide implementar una solución BI no solo se refiere a automatizar los reportes, transformar o consolidar los datos. Una buena implementación de BI ayudará a las organizaciones a lograr ventaja competitiva, mejorar el servicio al cliente, mejorar la productividad y hacer que el negocio sea rentable a través de cambios pensados con base en información.

Probablemente, las organizaciones producen grandes cantidades de datos que provienen de aplicaciones, sistemas transaccionales, archivos Excel, entre otros. Pero al estar en forma aislada, el personal pasa mayor tiempo elaborando reportes manuales en lugar de estar analizando los principales indicadores del negocio, por lo que el mayor problema es que no se logra obtener el provecho de todo ese volumen de datos de forma rápida y oportuna para tomar algunas decisiones tales como: mejorar las ventas, cumplir con los objetivos estratégicos, incrementar el número de clientes, entre otros.

Es por ello que surge la necesidad de contar con un repositorio central de datos, llamado Data Warehouse (DW), que consolide todo el conjunto de datos de la organización a través de procesos de integración que nos ayuden a extraer, transformar y cargar los datos. Esto nos ayuda a reducir trabajos operativos, dado que las personas ya no tienen que consolidar datos de forma manual. Sino más bien podrán extraer los datos, de forma mucho más rápida, para generar información relevante para el negocio.

Por otra parte, hay muchos sistemas desarrollados en el mundo que han sido justamente criticados por defectos de diseño en lo que concierne a la extracción de los datos para la adquisición de conocimiento y toma de decisiones. Como caso de ejemplo, tenemos los sistemas automatizados de votación, donde el foco consiste en automatizar todos o parte de estos procesos: registro del voto del elector, escrutinio, totalización y adjudicación, utilizando Tecnologías de Información y Comunicaciones (TIC).

Sin embargo, no tienen en consideración la importancia del manejo de estos datos al finalizar el proceso para la correcta y oportuna toma de decisiones. Por lo que con este trabajo se desea poder suministrar una manera de poder llevar a cabo esto, contando con la información adecuada en el tiempo adecuado para soportar toda la gestión de las operaciones de la aplicación de forma ágil y rápida.

Este trabajo está estructurado en capítulos y a continuación, se hace un resumen de cada uno de ellos:

**Capítulo 1: Problema de Investigación.** Describe la situación actual y las dificultades asociadas al problema para el cual se propone una solución. Además, se plantea el objetivo general y los objetivos específicos, la justificación y el alcance del Trabajo Especial de Grado (TEG).

**Capítulo 2: Marco Conceptual.** Contiene las bases teóricas conceptuales que dan soporte al desarrollo de una solución de Inteligencia de Negocio para la gestión de un sistema automatizado de votación. Aquí se contemplan los conceptos asociados a BI, sistema electoral y DW, así como un estudio y análisis de las herramientas que son utilizadas para su desarrollo.

**Capítulo 3: Marco Metodológico.** Describe las metodologías que son utilizadas durante el desarrollo de la solución de Inteligencia de Negocio. La primera de ellas es Scrum, que se utiliza en el proceso de diseño y construcción. La segunda, es el ciclo de vida de Ralph Kimball que es incorporado en las iteraciones de Scrum y es utilizado en el desarrollo del módulo analítico.

**Capítulo 4: Marco Aplicativo.** Explica de manera detallada cada una de las fases de trabajo que se aplican para el análisis, diseño y desarrollo de la solución propuesta, haciendo una adaptación de los métodos de desarrollo seleccionados para el caso de estudio de este TEG.

Para finalizar se presentan las conclusiones y recomendaciones, así como las referencias bibliográficas que se utilizan durante todo el desarrollo del presente TEG.

# CAPÍTULO 1

## Problema de Investigación

En este capítulo se lleva a cabo la descripción y planteamiento del problema, acompañado de los objetivos, la solución propuesta, la justificación y el alcance para el caso de estudio de un sistema automatizado de votación. Es importante mencionar que todos los conceptos utilizados serán descritos con mayor detalle en el próximo capítulo.

### 1.1. Planteamiento del Problema

En la actualidad, los sistemas electorales han pasado de ser sistemas manuales a sistemas automatizados, tomando auge en muchos países desde aproximadamente dos décadas. Cada uno de estos sistemas se ajusta a los requerimientos legales y tecnológicos de acuerdo con las características específicas de cada región o nación.

En consecuencia, los sistemas automatizados de votación son variados, involucrando desde el proceso de autenticación e identificación de los votantes, la actualización del censo electoral, hasta incluso el proceso de votación, conteo y consolidación de resultados en su totalidad. Por lo tanto, dichos sistemas ofrecen una visión completa de lo que transcurre en cada etapa del proceso electoral y eso se logra a través del uso de TIC.

De esta manera, la solución de sistemas automatizados que ofrecen algunas compañías para el voto a distancia o voto remoto permite a los organismos electorales servir y atraer a un electorado cada vez más móvil y disperso. Asimismo, los sistemas electrónicos correctamente implementados, pueden proporcionar integridad de los datos, mantener la confidencialidad y asegurar alta disponibilidad de este.

Adicionalmente, se sabe que la gestión de los procesos electorales es muy compleja. En consecuencia, una de sus principales debilidades es que se apoyan en estadísticas básicas para la generación de sus reportes, dejando de lado indicadores de gestión que permiten obtener información valiosa que puede ser de gran aporte para la toma de decisiones durante el proceso electoral.

A lo largo de distintos proyectos y la experiencia que se ha obtenido en estos trabajos, se determinó que se tiene la necesidad de obtener información adicional que no estaba previamente analizada ni pensada en la solución, en tal sentido se plantea usar BI para mejorar el producto y cumplir con los requerimientos de la alta gerencia de los entes electorales.

Por tal motivo, se elige como caso de estudio, el sistema automatizado de votación denominado Total Vote Tracker (TVT). Éste se define como una suite de aplicaciones para ser ejecutadas en Smartphone, Tablet y Laptops, con la finalidad de registrar, transmitir y totalizar los resultados de procesos electorales dentro de una organización, industria, cooperativa, localidad o país, de forma rápida y segura.

Adicionalmente, TVT cuenta con varias funcionalidades, tales como: captura, transmisión, digitalización, validación de actas, totalización de resultados, publicación de actas y reportes estadísticos.

Sin embargo, la funcionalidad de reportes actualmente es desarrollada bajo un esquema manual, donde un equipo de desarrollo debe realizar consultas muy complejas y costosas sobre la base de datos transaccional, lo que para algunos reportes ha provocado que la lógica de programación se tenga que manejar a nivel de la base de datos, el back-end y el front-ent.

Adicionalmente, esto ocasiona que el despliegue de los reportes dentro de la aplicación para la visualización del usuario final tarde mucho tiempo en cargar la información, debido a la cantidad de transacciones que al mismo tiempo se encuentran ejecutándose dentro de la misma base de datos. De igual modo existen otros inconvenientes como lo son:

- Los reportes actuales no son los más adecuados para la alta gerencia, debido a que se manejan datos muy básicos.
- Las consultas son más complejas, lo que requiere de una gran inversión de tiempo y recursos para el desarrollo de los mismos.
- Las peticiones para la construcción de los reportes son mucho más lentas, ya que tiene millones de peticiones transaccionales por segundo a las mismas tablas.
- Dificultad para la toma de decisiones de data histórica de forma fácil y rápida.

- Todos los datos se encuentran almacenados en la misma base de datos transaccional, lo cual causa que el modelo se torne más complejo.

El módulo de reportes actual se denomina Reportes Operacionales, generados con base a métricas especificadas por medio código fuente o desarrollo, basándose en resultados analíticos previamente.

## **1.2. Objetivos del Trabajo Especial de Grado**

### **1.2.1. Objetivo General**

Desarrollar una solución de Inteligencia de Negocio que permita generar indicadores asociados al proceso de votación, para mejorar la toma de decisiones y cumplir con las expectativas de las organizaciones electorales.

### **1.2.2. Objetivos Específicos**

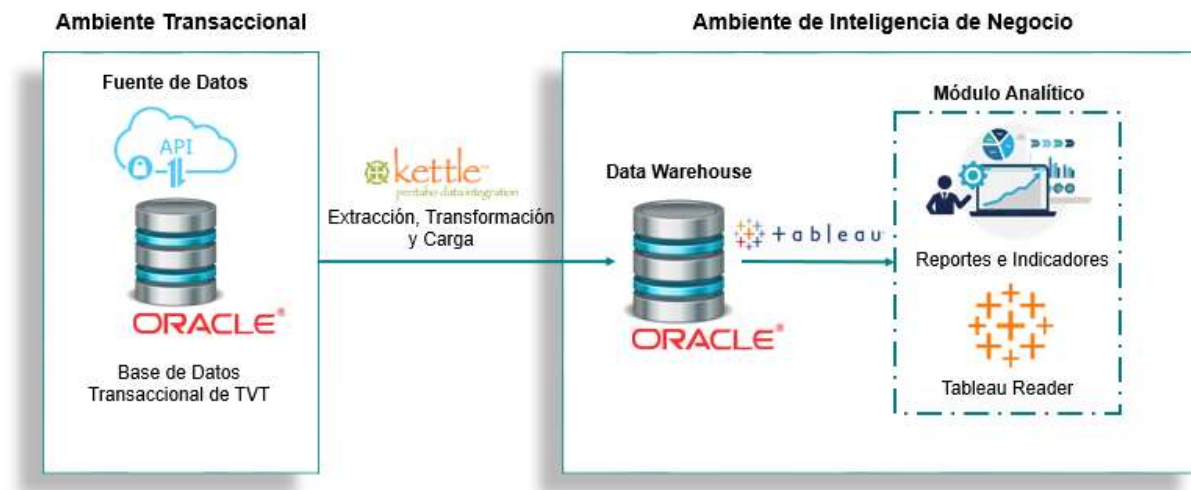
1. Planificar el proyecto basándose en la metodología escogida para el desarrollo de este TEG.
2. Analizar la Base de Datos Relacional implementada en Oracle, que sirve como fuente de datos para la solución de Inteligencia de Negocio.
3. Definir los indicadores relacionados con los requerimientos funcionales del sistema automatizado de votación evaluados previamente.
4. Seleccionar e instalar los productos necesarios para el desarrollo e implementación de la solución de Inteligencia de Negocio.
5. Crear, diseñar e implementar el modelo dimensional, que se avoque a los indicadores definidos con anterioridad.
6. Simular la data que se utiliza como insumo de los reportes de la solución.
7. Realizar el proceso de extracción, transformación y carga de los datos (ETL).
8. Generar un dashboard interactivo para mostrar los reportes relacionados con los indicadores levantados con anterioridad para la toma de decisiones.
9. Realizar las pruebas de aceptación del dashboard.

## **1.3. Solución Propuesta**

Por todo lo anteriormente expuesto en el planteamiento del problema, se propone una solución de Inteligencia de Negocio, que permita generar indicadores para el área de análisis de información



de un sistema automatizado de votación, permitiendo mejorar el proceso de la toma de decisiones, así como también la generación y visualización de reportes estadísticos, cuya arquitectura se puede observar a continuación (Ver ilustración 1).



**Ilustración 1: Arquitectura Propuesta**

De esta manera, tal como se visualiza en la ilustración 1, la solución está orientada a extraer los datos del sistema transaccional y transformarlos para incorporarlos en un DW, el cual almacena toda la data histórica, agregando un alto valor estratégico para la toma de decisiones.

Adicionalmente, se presenta la información al usuario de forma confiable y dinámica, promoviendo lo siguiente:

- Definición clara de indicadores y sus dimensiones.
- Presentación de datos al usuario de una forma que estos puedan ser interpretados de forma rápida y de manera sencilla para la comprensión y precisión del reporte.

## **1.4. Justificación de la Solución**

Basado en la solución propuesta, se considera necesario incluir una solución BI, con el fin de mejorar la funcionalidad de reportes y/o indicadores del sistema automatizado de votación, teniendo en cuenta las limitaciones que se tienen actualmente para generar reportes a la medida y con ello poder mejorar la toma de decisiones durante un proceso electoral.

De esta manera, se cuenta con un conjunto de procesos de negocio que aportarán información confiable y oportuna, que les permitan tomar decisiones tanto para la organización como para la

mejora a futuro del producto. En otras palabras, se puede contar con la visualización de totalizaciones, porcentajes y otras medidas a través de gráficos y tablas dinámicas, a fin de proporcionar al personal un mejor tiempo de análisis de los indicadores del negocio y mejorar la toma de decisiones.

## **1.5. Alcance de la Solución**

La solución propuesta es realizada en ambientes controlados, tomando como caso de estudio el sistema automatizado de votación denominado Total Vote Tracker (TVT). Los datos que sirven como suministro para el BI a implementar son simulados, para garantizar la confidencialidad de los datos. El trabajo consta de la implementación de un BI, sin necesidad de estar integrado con la aplicación. Así mismo, cuenta con una base de datos relacional que contiene todos los datos necesarios para alimentar el DW.

Dado que la data es simulada, también forma parte de este trabajo desarrollar todos los casos de ejemplos basados en la data de Venezuela. Finalmente, todo este desarrollo se lleva a cabo en un ambiente controlado, local y virtualizado, con el sistema operativo (SO) Windows 7 Professional x64bits.

# CAPÍTULO 2

## Marco Conceptual

En este capítulo se describen conceptos, métodos y tecnologías que sirven como base teórica para el desarrollo de la solución BI propuesta, cuyo propósito es la mejora de la funcionalidad de reportes de un sistema automatizado de votación para permitir una mejor toma de decisiones. A su vez, sirve de base para abordar el próximo capítulo, donde se describe la metodología a utilizar para el desarrollo de este trabajo.

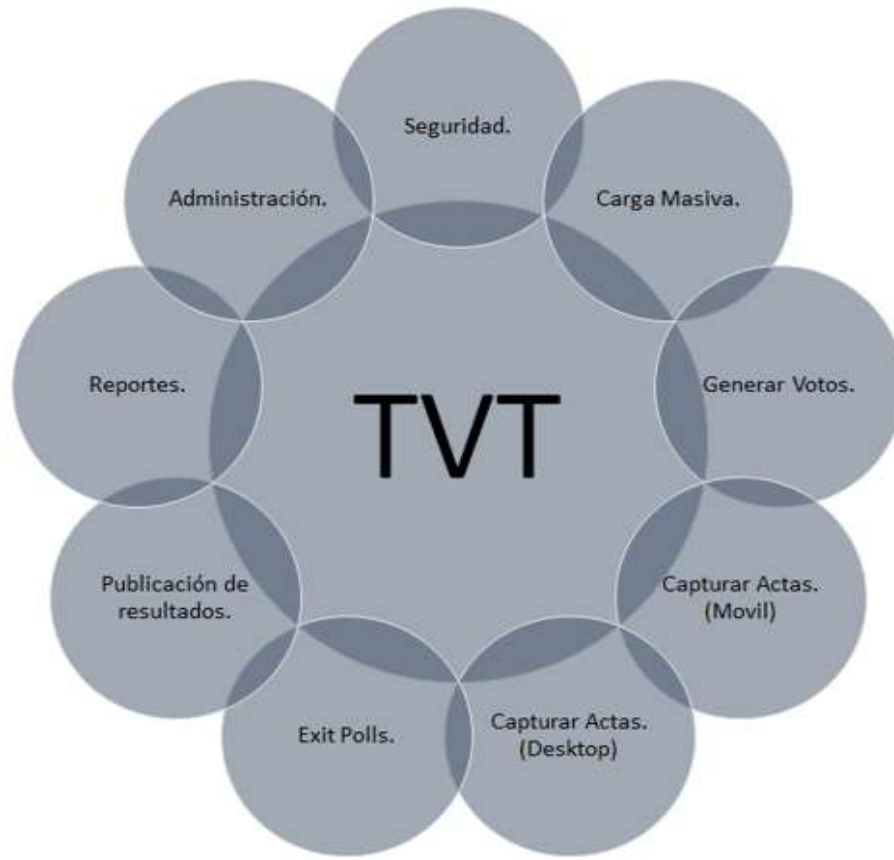
### 2.1. Sistema Electoral

Se define como proceso electoral a la forma legal y pacífica para designar gobernantes mediante la participación de la ciudadanía y la interacción entre partidos y grupos políticos. Una de las características fundamentales de las sociedades democráticas es la realización de elecciones libres. Se puede decir que en muchos países, la mayoría de los ciudadanos cree que sólo se relaciona con la política cuando participa en las elecciones.

Las elecciones son necesarias por una razón muy importante: es imposible que una sociedad entera pueda tomar decisiones de manera unánime, adecuada y oportuna. Se requiere por tanto de dar cierto poder de decisión a un individuo o a un grupo de ellos, para que, como representantes del pueblo, gobiernen en favor de los intereses de toda la población. [1]

### 2.2. Total Vote Tracker

Total Vote Tracker (TVT) es una suite de aplicaciones diseñada para ser ejecutada en Smartphone, tabletas y laptops, con la finalidad de registrar, transmitir y totalizar los resultados de procesos de elecciones de las diferentes contiendas dentro de una empresa, industria, cooperativa, localidad o país, de forma rápida y segura. Por consiguiente, las aplicaciones que comprenden este proyecto son:



**Ilustración 2: Suite de aplicaciones de TVT**

Así mismo, esta suite de aplicaciones cuenta con las siguientes características:

- Consultas y verificación de votos: se puede visualizar la tendencia de los votos, según las opciones ingresadas.
- Accesibilidad y Usabilidad: el diseño es sencillo e intuitivo, el cual proporciona facilidad en la utilización del mismo.
- Plataforma transparente, auditable con niveles de trazabilidad: el registro, verificación y validez de las actas de escrutinio se puede realizar en tiempo real.
- Transmisión segura de los datos: viaja por un canal seguro, sus datos son cifrados al momento de realizar la transmisión y descifrados en la recepción, de igual modo dispone de una autenticación del usuario autorizado para las operaciones.

Por lo tanto, el flujo de procesos principal en la suite se define tal como se visualiza en la ilustración 3.



**Ilustración 3: Flujo de Procesos de TVT**

Adicionalmente, en la tabla 1 se puede observar una breve descripción de las funcionalidades que posee la suite de aplicaciones de TVT.

**Tabla 1: Descripción de las funcionalidades de TVT**

Número	Título	Descripción
1	Gestión de Usuarios.	Realiza la administración de los diferentes usuarios por perfiles.
2	Configurar Elecciones.	Administra los datos de una elección en específico.
3	Registrar Geo-Localización.	Registra las diferentes localizaciones que la elección requiere.
4	Gestión de Partidos Políticos.	Registra los partidos políticos de una elección si es de índole político.
5	Gestión de Candidatos.	Registra los postulantes de una elección.
6	Gestionar Alianzas Políticas.	Administra las alianzas políticas de una oferta electoral.
7	Carga Masiva.	Módulo que permite cargar grandes cantidades de datos. Ej. Los datos de los usuarios, categorías, partidos y candidatos.
8	Configurar medio de votación	Configura los distintos dispositivos a través de los cuales se puede realizar la elección.
9	Registrar Votos.	Registra, contabiliza y transmite los votos al servidor central.
10	Capturar Actas.	Permite la captura de las actas de los centros electorales.
11	Digitalizar Actas.	Permite la transcripción de los votos obtenidos por los candidatos.
12	Verificar Actas.	Permite verificar los datos de las imágenes de las actas contra las cifras digitalizadas.
13	Resultados de la votación.	Se refiere a los principales indicadores de una elección.
14	Publicar Actas.	Visualizar las actas de votación
15	Registrar Exit Poll.	Permite el ingreso de la información de las encuestas.
16	Reportes.	Gestiona resultados estadísticos de la información contenida en el sistema.
17	Multilinguaje.	Permite cambiar de idioma la aplicación.
18	Soporte.	Permite brindar soporte de los sistemas en una elección.

## 2.3. Sistema de Información

Un sistema de información es un conjunto de datos que interactúan entre sí con un fin común. De esta manera pueden ayudar a administrar, recolectar, procesar, almacenar y distribuir información relevante para los procesos fundamentales y las particularidades de cada organización.

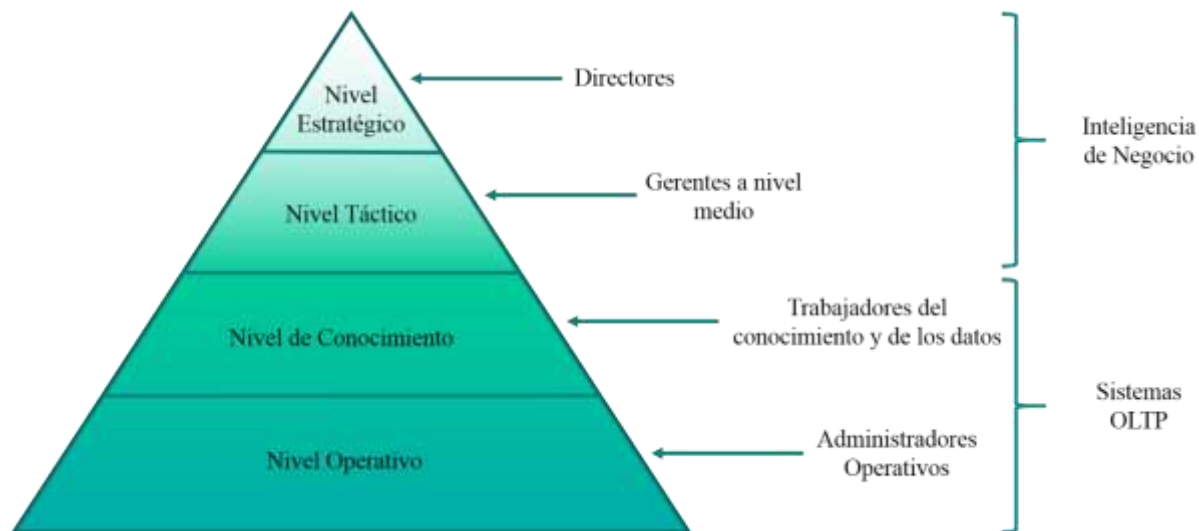
La importancia de un sistema de información radica en la eficiencia en la correlación de una gran cantidad de datos ingresados a través de procesos diseñados para cada área con el objetivo de producir información valiosa para la posterior toma de decisiones. [2]

### 2.3.1. Tipos de Sistemas de Información

En la cultura organizacional, existen varios tipos de sistemas de información según el nivel operacional en que se utilicen. Algunos de los sistemas más comunes se detallan a continuación. [2]

- **Sistemas de Procesamiento de transacciones:** (Traditional processing system – TPS) son los sistemas empresariales básicos que sirven al nivel operacional de la organización, destinados a procesar grandes volúmenes de información alimentando grandes bases de datos.
- **Sistemas expertos o basados en el conocimiento:** (Knowledge working system – KWS) Se encuentran en el nivel operativo, selecciona la mejor solución para el problema presentado.
- **Sistemas para la gestión de Información:** (Management information system – MIS) son los sistemas que van dirigidos al nivel administrativo, su principal funcionalidad es gestionar y elaborar informes periódicos.
- **Sistemas para la toma de decisiones:** (Decision support system – DSS) se encuentra en el nivel estratégico, se destaca por su diseño e inteligencia que permite una adecuada selección e implementación de proyectos.

A continuación, se muestra una ilustración de jerarquía de los tipos de sistema de información, indicando a qué personal va dirigido cada nivel (Ver Ilustración 4).



**Ilustración 4: Tipos de Sistemas de Información**

### 2.3.2. OLAP

Se define OLAP como el procesamiento analítico en línea, que permite a los usuarios extraer de forma fácil y selectiva los datos, como también verlos desde diferentes puntos de vista. Para facilitar este tipo de análisis, los datos OLAP se almacenan en una base de datos multidimensional. Considerando que una base de datos relacional puede ser pensada como de dos dimensiones, en cambio una base de datos multidimensional, considera cada atributo de datos (tal como producto, región geográfica de ventas, y período de tiempo) como una “dimensión” separada.

El software OLAP puede localizar la intersección de las dimensiones (todos los productos vendidos en la región oriental por encima de un precio determinado durante un cierto periodo de tiempo) y mostrarla. Atributos tales como períodos de tiempo se pueden desglosar en sub- atributos.

Una base de datos OLAP no necesita ser tan grande como un Data Warehouse, ya que no todos los datos transaccionales se necesitan para el análisis de tendencias. Usando Conectividad de Base de datos Abierta (ODBC), los datos se pueden importar de bases de datos relacionales existentes para crear una base de datos multidimensional para OLAP. [3]

### 2.3.3. OLTP vs. OLAP

A continuación, se muestra en la Tabla 2 las principales diferencias entre los sistemas OLTP y los sistemas OLAP [4].

**Tabla 2: OLTP vs. OLAP [4] [5]**

<b>Características</b>	<b>OLTP</b>	<b>OLAP</b>
<b>Fuente de Datos</b>	Almacena hechos transaccionales.	Almacena hechos del negocio.
<b>Propósito de los datos</b>	BD para el control de los procesos de la operación (información).	BD para auditoría y toma de decisiones (conocimiento).
<b>Contenido de los datos</b>	Se almacenan datos de cada transacción. Información de soporte al día a día (operacional).	Información histórica para ser analizada. Integra datos para generar información relevante.
<b>Inserciones</b>	El usuario agrega, modifica, elimina, consulta.	El usuario solo consulta.
<b>Actualizaciones</b>	Gran cantidad de transacciones por día.	Se maneja solo una transacción que contiene cientos de registros de manera periódica.
<b>Consultas</b>	Simple y estandarizadas.	Usualmente complejas.
<b>Espacio Requerido</b>	Pequeño a mediano. Depende del negocio.	Muy grande, debido a los datos históricos.
<b>Diseño de Base de Datos</b>	El diseño de la base de datos es normalizado.	El diseño de la base de datos es desnormalizado.

Tal como se ha podido observar en la tabla, tanto los sistemas OLTP como los sistemas OLAP brindan beneficios a sus usuarios de diversas maneras. Sin embargo, cuando se trata de una solución analítica para la toma de decisiones, es necesario considerar que se deben integrar diferentes componentes. Es por ello, que a continuación se explican los aspectos más importantes de una solución de Inteligencia de Negocio [5].



## 2.4. Inteligencia de Negocio

La Inteligencia de Negocio, o mejor conocido por sus siglas en inglés BI (Business Intelligence), según el Data Warehouse Institute, es definido como la combinación de tecnología, herramientas y procesos que permiten transformar los datos en información y a partir de la información ser capaces de descubrir conocimientos basados en hechos, donde finalmente esta información irá dirigida a un plan o una estrategia comercial [6].

La Inteligencia de Negocio debe ser parte de la estrategia empresarial, esta le permite a la organización optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la misma y la capacidad de tomar buenas decisiones para así obtener mejores resultados.

En general, se puede decir que BI es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un Data Warehouse), para descubrir tendencias o patrones, a partir de los cuales se derivan ideas y se extraen conclusiones.

### 2.4.1. Características

Según lo explica Cano en su libro “Business Intelligence: Competir con Información.” [7], los sistemas BI deben cumplir con las siguientes características:

- ✓ **Apoyo a la toma de decisiones:** consiste en organizar y presentar los datos relevantes para que sirvan como soporte a la hora de tomar decisiones. Esto implica la utilización de tecnologías y técnicas de análisis con el fin de obtener información relevante para el dueño o gerentes del negocio.
- ✓ **Visión unificada de los datos:** todos los datos deben estar localizados en un único repositorio de datos, sin importar el tipo de datos o la fuente de donde provenga, para así dar la sensación de que los datos están centralizados.
- ✓ **Creación personalizada de informes y consultas:** permite el desarrollo de consultas y reportes a la medida sobre información contenida en los Almacenes de Datos.
- ✓ **Vistas gráficas e interactivas para la representación de información analítica:** a través de cuadros de mandos integrales y estratégicos se facilita la visualización de los indicadores de negocio.
- ✓ **Capacidad de procesamiento de grandes volúmenes de datos:** las soluciones de BI permiten realizar consultas comparando los datos actuales con los históricos.

## 2.4.2. Arquitectura de Inteligencia de Negocio

Las soluciones de BI se construyen sobre una arquitectura general básica, que está compuesta por cinco (5) capas: Datos Fuente, Integración, Base Analítica, Capa Lógica y Visualización. En la Ilustración 5 se puede observar con mayor detalle. [8]

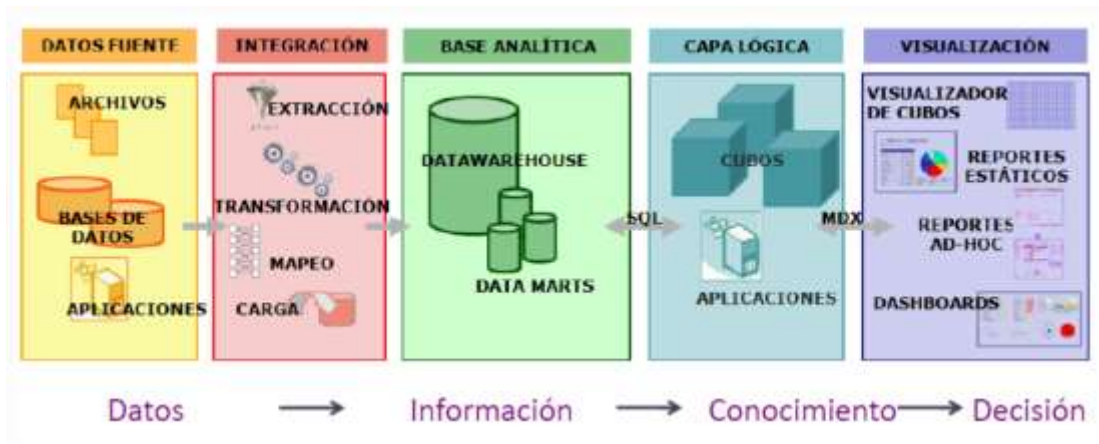


Ilustración 5: Arquitectura de una Solución de BI. Fuente: [8]

### Datos Fuente

El primer recuadro representa las Fuentes de Datos (Hojas de Excel, archivos csv, texto plano, Bases de Datos transaccionales, cubos, archivos xml, entre otros.). De aquí se extraen múltiples datos que requieren una transformación estructural para perfeccionar su análisis y estudio. Las fuentes de datos pueden tener un origen variado:

- ✓ Datos operacionales procedentes de bases de datos corporativas.
- ✓ Generalmente procedentes de ambientes transaccionales OLTP.
- ✓ Datos generados por sistemas de información.
- ✓ Datos privados.
- ✓ Datos externos (Internet, Bases de Datos comerciales, de clientes, información económica, entre otros.)

### Integración

Esta capa es la encargada de la integración de los datos desde las múltiples fuentes existentes para la construcción del Data Warehouse y los Data Mart de la base analítica. Se utilizan rutinas para Extracción, Transformación y Carga (ETL). Dicho proceso de ETL suele apoyarse en una bodega de datos intermedia que interactúa como un área de paso en donde son descargados los datos,

crudos y sin transformaciones, una vez que sean extraídos desde los sistemas fuente transaccionales. Este proceso es en el que se definen de las fuentes heterogéneas que campos se van a utilizar, si necesitan algún tipo de modificación y/o transformación y donde se quieren ubicar esos datos, este proceso se le conoce comúnmente como “mapping”.

## **Base Analítica**

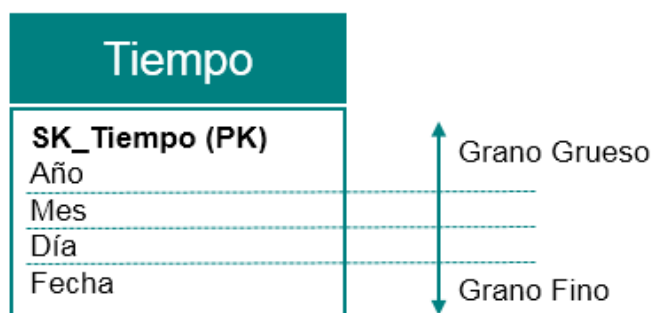
La base analítica es el núcleo del sistema. Puede contener diversos componentes tales como el área intermedia o de integración de los datos y el almacén de datos. La información que fue transformada en el área de integración debe ser almacenada en un Data Warehouse o repositorio de datos, en donde es posible administrar y monitorear los procesos o consultas del sistema, y que a la vez está relacionado con la construcción de los Data Mart. A continuación, se explica con detalle los aspectos más importantes relacionados con estas estructuras de datos.

- **Definición de Data Warehouse:** Según R. Kimball [2], un Data Warehouse (DW) se define como “una copia de la data transaccional específicamente estructurada para consultas y análisis” donde además especifica que se debe iniciar con la construcción de los Data Mart, los cuales define como un subconjunto de datos de un Data Warehouse.
- **Definición de Data Mart:** Un Data Mart es un subconjunto de datos de un Data Warehouse o dicho de otra forma, es un Data Warehouse de alcance limitado. Según Kimball y Ross [2], es una base de datos departamental, orientada al análisis, almacenamiento e integración de los datos de un área de negocio específica. Esto quiere decir que un Data Mart posee la misma funcionalidad de un Data Warehouse. Así mismo, puede ser alimentado desde los datos de un Data Warehouse, o integrar por sí mismo un compendio de distintas fuentes de información. Estos pequeños almacenes de datos pueden ser una alternativa para organizaciones medianas que no pueden afrontar costos de implementar un sistema tan amplio de almacenamiento de data. Por lo general están estructurados en modelos dimensionales de estrella o copos de nieve.
- **Área Intermedia:** Es un espacio temporal y de carácter volátil. En él, se realiza una primera extracción rápida de las fuentes de datos y se almacenan temporalmente mientras se analizan, limpian, mejoran y, posteriormente, se carga al Data Warehouse. Este trabajo es necesario realizarlo pues los datos que se tienen en los sistemas transaccionales no siempre están preparados para la toma de decisiones. A dicho espacio se le conoce

comúnmente como Staging Area, o área intermedia y los usuarios finales no hacen consultas sobre los datos alojados en esta área. [9].

- **Modelo Dimensional:** El modelo dimensional se describe por primera vez en el año 1996 por Ralph Kimball [10], como propuesta para el diseño del Data Warehouse. Es un diseño lógico que pretende representar los hechos del negocio. Este modelo está optimizado para llevar a cabo consultas con un alto rendimiento. Cada dato de negocio puede representarse en un cubo con dos o más dimensiones. Un modelo dimensional permite estructurar los datos alrededor de ciertos conceptos naturales para la organización, y así proveer las bases para un análisis intuitivo de la información que permita tomar decisiones de importancia para dicha organización. [11]. Los elementos básicos de un modelo dimensional son los siguientes:

- ✓ **Granularidad:** se entiende como el nivel de detalle (mínimo) con el que se almacenarán los datos en la tabla de hechos de acuerdo con la actividad a modelar. La granularidad determina el significado de las tuplas de la tabla de hechos. (Ver Ilustración 6).



**Ilustración 6: Ejemplo de Granularidad**

- ✓ **Jerarquía:** una jerarquía define la posición relativa de un atributo con respecto a otros pertenecientes a la misma dimensión. Representados bajo una relación de tipo jerárquica o bien conocida como forma de árbol, donde los atributos serán progresivamente más detallados si se recorre de manera descendente el árbol hasta llegar a las hojas, quienes son los que tienen mayor nivel de detalle. (Ver Ilustración 7).



**Ilustración 7: Ejemplo de Jerarquía**

- ✓ **Hecho:** es el resultado medible por parte de la organización y es el punto central para la toma de decisiones. Es una medida cuantitativa o cualitativa del negocio. Es esencial que el hecho posea aspectos dinámicos con el fin de poder observar la evolución a través del tiempo del mismo. Físicamente, se representan dentro de una tabla central llamada Tabla de Hechos, éstas poseen una clave primaria compuesta, formada por claves foráneas provenientes de las dimensiones, poseen uno o más campos llamados medidas o hechos, que se hacen sobre la granularidad de dicha tabla. Ejemplos de hechos: cantidad vendida, monto cobrado, costo, precio, entre otros. (Ver Ilustración 8).



**Ilustración 8: Ejemplo de Tabla de Hechos**

- ✓ **Dimensión:** son aquellas tablas que acompañan a la tabla de hechos y le dan contexto a la información. Contiene un conjunto de atributos con una clave única (subrogada<sup>1</sup>) que corresponde con uno de los componentes de la tabla de hechos. Los aspectos más relevantes de estas tablas son:

<sup>1</sup> En un almacén de datos debe evitarse el uso de las claves del sistema operacional como clave primaria, es por eso que las claves de las dimensiones deben ser generadas artificialmente, llamándose así claves subrogadas. Claves de tipo entero (4 bytes) son suficientes para las dimensiones de cualquier tamaño.

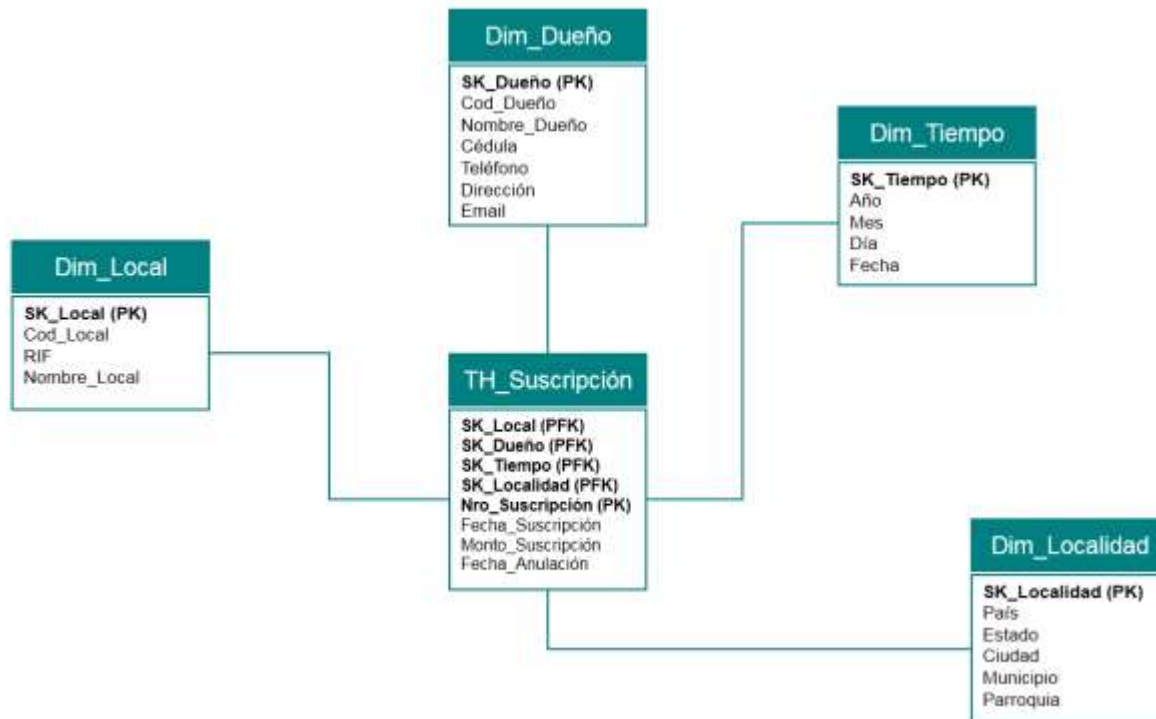
- Contiene información textual descriptiva del negocio.
- Tienen muchas columnas o atributos que describen los hechos del negocio.
- Los atributos sirven como fuente para las restricciones de las consultas, agrupaciones, etiquetas de los análisis realizados sobre el almacén de datos.
- Son los “por”: ventas por año, por mes, por marca, por producto.
- Los atributos de la dimensión surgen del sistema operacional y son los más relevantes para el análisis de los datos.
- La información puede encontrarse desnormalizada.

A continuación se muestra en la Ilustración 9, un ejemplo de tabla Dimensión con las características anteriormente mencionadas.



**Ilustración 9: Ejemplo de Dimensión**

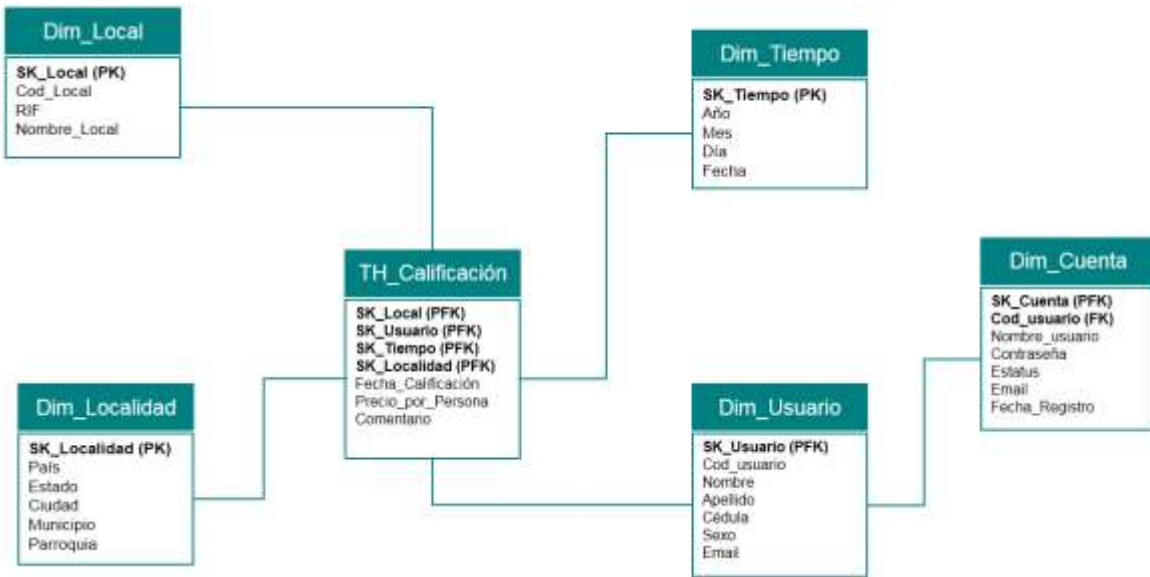
- ✓ **Esquemas:** los esquemas del modelo dimensional de un almacén de datos se entienden como la colección u organización de las tablas que conforman a éste. Los tipos de esquemas más comunes son:
  - **Estrella:** es el esquema más simple. Forma una estrella con puntos radiales desde el centro. El centro de la estrella está conformado por una tabla de hechos, y en los extremos de ésta se encuentran una o más tablas dimensión, las cuales están conectadas a la tabla central de hechos. No pueden existir relaciones entre las tablas dimensión. Las tablas dimensión que conforman el esquema estrella se encuentran desnormalizadas parcialmente, mientras que las tablas de hechos se encuentran normalizadas, esto permite minimizar el número de uniones existentes. En la Ilustración 10 se puede observar un ejemplo.



**Ilustración 10: Ejemplo de un Modelo Dimensional con Esquema Estrella**

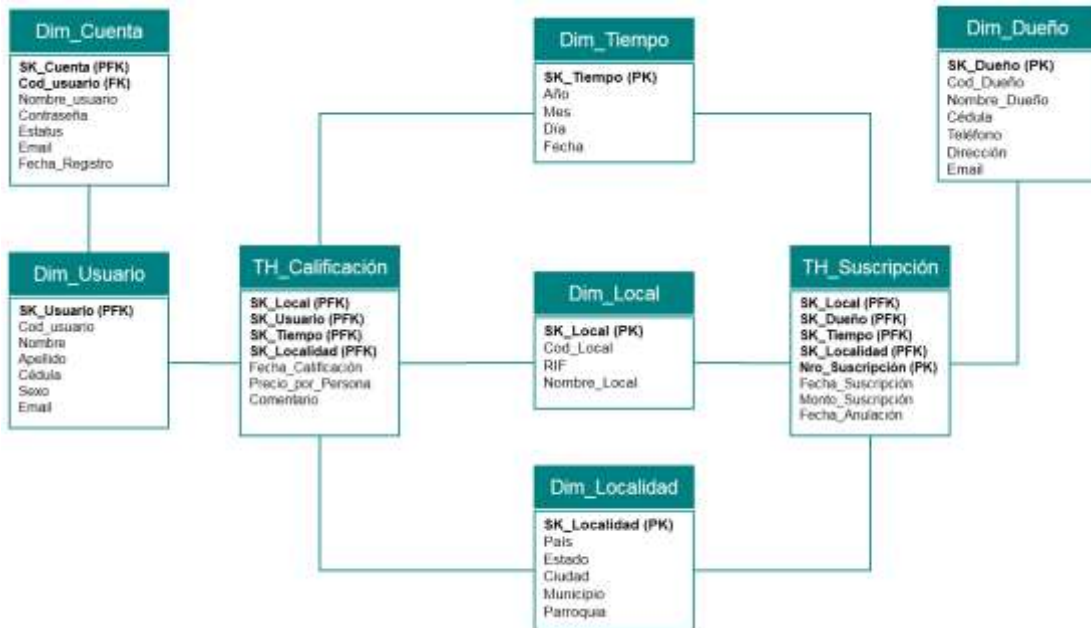
- **Copo de Nieve:** es una variación del esquema estrella, que incorpora la normalización de las dimensiones. En cada punta de la estrella (dimensión) se generan otras dimensiones, lo que se asemeja a un copo de nieve. Las tablas dimensión se encuentran normalizadas en 3FN<sup>2</sup>. Generalmente se utilizan cuando las tablas dimensiones son muy grandes o muy complejas y es muy difícil representar los datos en un esquema estrella. En la Ilustración 11 se puede observar un ejemplo.

<sup>2</sup> 3FN (Tercera Forma Normal): se utiliza para la normalización de las Base de Datos y cumple con tres condiciones: la tabla está en la segunda forma normal (2FN), ningún atributo no-primario de la tabla es dependiente transitivamente de una clave primaria, y es una relación que no incluye ningún atributo clave.



**Ilustración 11: Ejemplo de un Esquema Copo de Nieve**

- **Constelación de estrellas:** este modelo está compuesto por una serie de esquemas en estrella. Está formado por una tabla de hechos principal y una serie de tablas de hechos agregadas o auxiliares, no necesariamente están relacionadas con las mismas dimensiones de la principal sino que pueden relacionarse con un subconjunto de ellas o con nuevas dimensiones. (Ver Ilustración 12).



**Ilustración 12: Ejemplo de un Esquema de Constelación de Estrellas**



## **Capa Lógica**

Contiene el corazón del sistema. Aquí es donde se almacenan las reglas de negocio. Estas reglas de negocio no afectan a las reglas de transformación de datos, pero afectan a lo que luego se puede visualizar en los informes.

Es un conjunto de subrutinas que regulan la acción del usuario enviándolas a la capa de repositorio de datos. Esta información es enviada y encapsulada en entidades de negocio

## **Visualización**

Esta capa es la encargada de desplegar los análisis sobre los cubos mediante sus diversos cortes de información, cruces y filtros. Todos los datos almacenados se exploran a partir de herramientas de visualización de la información, las cuales permiten el desarrollo de reportes, análisis, cuadros de mando, alertas y diferentes instrumentos que se llevan hasta los usuarios para dar soporte a sus decisiones y así proporcionar soluciones de BI mucho más completas. Es decir, la interface de acceso a usuarios permite interaccionar con los datos, representar de forma gráfica con aquellos resultados de las consultas y los indicadores de gestión que fueron construidos.

## **2.5. Indicador**

Un indicador es una expresión cualitativa o cuantitativa observable, que permite describir características, comportamientos o fenómenos de la realidad permitiendo evaluar el desempeño y su evolución en el tiempo. Los indicadores tienen como objetivo generar información útil para mejorar el proceso de toma de decisiones, monitorear, evaluar el desempeño de acuerdo al cumplimiento de acuerdos y compromisos, cuantificar los cambios, identificar problemas y oportunidades, medir comportamientos, facilitar la delegación en las personas, identificar iniciativas y acciones necesarias. [17]

A través de los indicadores se puede observar:

- ✓ Evolución de una variable.
- ✓ Establecimiento de una relación entre variables.
- ✓ Comparación con períodos anteriores, productos similares o una meta o compromiso.

Un indicador se rige por el acróstico “SMART” (Specific – Measurable – Actionable – Relevance – Timely) que significa que debe proporcionar información Específica y relevante para no sesgar

su análisis, debe ser oportuno y permitir la toma de decisiones y el control efectivo, el indicador es objetivo y preciso, no comporta juicios de valor en su resultado y debe estar parametrizado en un tiempo definido.

### **2.5.1. Objetivos de los Indicadores**

Los objetivos principales para todo indicador se definen de la siguiente manera: [17]

- a) Generar información útil para mejorar el proceso de toma de decisiones.
- b) Monitorear y evaluar el desempeño conforme al cumplimiento de acuerdos y compromisos.
- c) Cuantificar los cambios.
- d) Identificar problemas y oportunidades.
- e) Medir comportamientos y facilitar la delegación en las personas.
- f) Identificar iniciativas y acciones necesarias.

### **2.5.2. Tipos de Indicadores**

Así mismo, existen muchos tipos de clasificaciones comunes en la teoría de indicadores. Se debe saber discernir entre indicadores de cumplimiento, de evaluación, de eficiencia, de eficacia e indicadores de gestión. Es por ello que a continuación se describen brevemente:

#### **Indicadores de Cumplimiento**

Los indicadores de cumplimiento están relacionados con las razones que indican el grado de consecución de tareas y/o trabajos, es decir, o la conclusión de una tarea. Ejemplo: cumplimiento del programa de pedidos. [18]

#### **Indicadores de Evaluación**

Los indicadores de evaluación están relacionados con las razones y/o los métodos que ayudan a identificar nuestras fortalezas, debilidades y oportunidades de mejora. Es decir, tiene que ver con el rendimiento que se obtiene de una tarea, trabajo o proceso. Ejemplo: evaluación del proceso de gestión de pedidos. [18]

#### **Indicadores Cuantitativos**

Los indicadores cuantitativos se refieren a la representación numérica de la realidad, su característica más importante es que, al encontrarse valores diferentes, estos pueden ordenarse de forma ascendente o descendente. [17]

## **Indicadores Cualitativos**

Su característica principal es que su resultado se refiere a una escala de cualidades. Es un instrumento que permite tener en cuenta la heterogeneidad, amenazas y oportunidades del entorno organizacional y/o territorial. Además, permiten evaluar, con un enfoque de planeación estratégica, la capacidad de gestión de la dirección y demás niveles de la organización. [17]

## **Indicadores de Eficacia**

Miden el grado de cumplimiento de los objetivos propuestos, se enfocan en el Qué se debe hacer. Para este indicador se deben conocer y definir los requerimientos del cliente del proceso contra lo que él espera. [17]

## **Indicadores de Eficiencia**

Miden el nivel de ejecución del proceso, se concentran en el Cómo se hicieron las cosas y miden el rendimiento de los recursos utilizados en un proceso. [17]

## **Indicadores de Efectividad**

Este concepto involucra la eficiencia y la eficacia, es decir, el logro de los resultados programados en el tiempo y con los costos más razonables posibles. Supone hacer lo correcto con gran exactitud y sin ningún desperdicio de tiempo o dinero. [17]

## **2.6. Oracle Database**

Oracle database (Oracle DB) es un sistema de gestión de base de datos relacional (Relational Data Base Management System), desarrollado por Oracle Corporation. Fue desarrollado originalmente en 1977 por Lawrence Ellison y otros desarrolladores, Oracle DB es uno de los motores de base de datos relacional más confiables y más utilizado. Es considerado uno de los sistemas de base de datos más completos destacando el soporte de transacciones, la escalabilidad, la estabilidad y el soporte multiplataforma. Aunque su dominio en el mercado de servidores empresariales ha sido casi total hasta hace poco, recientemente sufre la competencia del Microsoft SQL Server de Microsoft y de la oferta de otros RDBMS con licencia libre como PostgreSQL, MySQL o Firebird. Las últimas versiones de Oracle han sido certificadas para poder trabajar bajo GNU/Linux. [24]

## **2.7. Herramientas para el desarrollo de una Solución de Inteligencia de Negocio**

Después de hacer un estudio del cuadrante mágico de Gartner para las plataformas de Inteligencia de Negocio, que fue publicado en Febrero del 2018, se tomó en cuenta para el análisis de la elección de la herramienta tecnológica, Tableau, Qlik y Microsoft por ser las que se encuentran en cuadrante de los líderes.

### **2.7.1. Tableau**

Tableau es un software de Inteligencia de Negocio, propietario, con opción a licenciamiento estudiantil por un año y una versión de prueba de 14 días para todo público. Es una herramienta de análisis de datos que ayuda a mejorar y acelerar la toma de decisiones en una organización y se caracteriza por “lograr que los datos sean comprensibles para las personas comunes y corrientes”.

Así mismo, destaca por ser muy fácil de utilizar, intuitiva y permitir ver todos los cambios a medida que se van realizando. También, es capaz de combinar datos de múltiples fuentes y publicarlos en el portal web de la organización. Además, se orienta a la optimización del uso de la memoria disponible. Es posible filtrar, ordenar y analizar los datos en tiempo real durante una reunión, permitiendo responder a preguntas en el mismo momento que surjan.

Tableau se ha mantenido liderando el cuadro mágico de Gartner durante seis (6) años consecutivos. Esta herramienta se basa en el lenguaje VizQL, que es un lenguaje visual de consulta que transforma acciones de arrastrar y soltar en consultas de datos y, luego, muestra esos datos en forma visual.

Posee conexiones con Base de Datos relacionales y no relacionales. Disponible únicamente para sistemas operativos Windows 7 o una versión más reciente (64bits) y Mac - OSX 10.11 o una versión más reciente. También en entornos virtuales como Citrix (sin transmisión), Microsoft Hyper-V, Parallels, VMWare, Microsoft Azure y Amazon EC2.

Estos productos están habilitados para Unicode y son compatibles con datos almacenados en cualquier idioma. La interfaz de usuario y la documentación correspondiente se encuentran disponibles en inglés, francés, alemán, español, portugués (Brasil), japonés, coreano y chino simplificado. [12]

### **2.7.2. Qlik**

Qlik es una compañía de software que cuenta con una suite de productos dedicados al análisis de los datos. Esta plataforma se ofrece por licenciamientos pagos para soporte y productos Enterprise. Así mismo, ofrece una versión Open Source completa para uso personal y de pequeñas empresas.

La plataforma ha sido desarrollada para ser intuitiva y fácil de usar. El acceso instantáneo a los datos, la flexibilidad en el uso y la actualización en tiempo real son los tres grandes valores para un uso eficaz. Además, permite utilizar datos de cualquier fuente, desde archivos de Excel hasta todo tipo de Base de Datos. [13]

Más aún, está diseñado para aprovechar el poder de las soluciones basadas en contenedores y las arquitecturas de micro servicio, como Docker. Destacado por el tratamiento columnar de la información permite optimizar la consulta al no procesar tablas completas, consiguiendo el análisis en menor tiempo. Los datos son procesados en la memoria RAM de los servidores Qlik, lo que lo convierte en la manera actual más rápida de procesar información. [14]

### **2.7.3. Microsoft Power BI**

Power BI es un conjunto de herramientas de análisis empresarial que permite interactuar con una cantidad masiva de datos dentro de Excel. Power BI se complementa con la potencia de SharePoint y ofrece capacidades para colaborar y dar acceso a la información de manera segura y en un entorno de configuración. [15]

Power BI cuenta con un grupo de funcionalidades que sirven de complemento para Excel:

- ✓ Power Pivot, es el motor que hace funcionar todos los cálculos que se muestran visualmente de forma interactiva en los informes generados.
- ✓ Power Query, transforma los datos brutos en tablas útiles para poder ser utilizados por Power Pivot.
- ✓ Power View, es la herramienta con la que se puede visualizar los datos, explorándolos de una manera muy sencilla e intuitiva.
- ✓ Power Map, es la herramienta de visualización de datos geoespaciales en 3D.

Adicionalmente, cuenta con servicios basados en la nube tales como: Refresh, Share and Manage data source, Lenguaje natural de consultas Q&A, Dashboard y Apps nativas para móviles y tablets. [16]

De esta manera, Power BI ofrece cinco (5) modalidades de uso:

- ✓ Power BI Desktop: es gratuito para la integración de los orígenes de datos, modelización y construcción de sus paneles e informes, seleccionando y arrastrando los objetos visuales que desee, para la representación, explotación y compartición de la información.
- ✓ Power BI Mobile: se puede descargar la APP gratuita para la visualización de los paneles e informes desde dispositivos móviles ya sean Android o iOS.
- ✓ Power On-line: se puede acceder a través de una dirección web a los paneles e informes, que ha sido proporcionada en el momento de la publicación. Se puede consultar, editar y volver a publicar, desde cualquier ordenador, con conexión a Internet.
- ✓ Power BI Pro: Es la licencia para administradores y usuarios que necesitan integrar grandes cantidades de datos y consumirlos cada hora.
- ✓ Power BI Embedded: El servicio más avanzado permitirá administrar grandes cantidades de datos, generados en una industria para ser por ejemplo publicados en servicios web o integrar datos de elementos externos del tipo Internet of Things (IoT) accediendo a un nivel superior de conocimiento de información.

Una vez analizadas y probadas las tres herramientas, se decidió trabajar con Tableau por ser la que mejor se adapta a los requerimientos del negocio.

# CAPÍTULO 3

## Marco Metodológico

En relación con lo que se explica en el Capítulo 2, en el presente capítulo se describe la metodología Scrum y el Ciclo de Vida Dimensional de Kimball. Ambas metodologías, adaptadas a este caso de estudio permiten sustentar los objetivos específicos e interrogantes del TEG en su marco aplicativo que se encuentra descrito en el próximo capítulo.

Este TEG se lleva a cabo mediante Scrum, por ser una metodología ágil que permite controlar y planificar todo el proyecto. En una de las iteraciones definidas en este TEG, se incorpora el Ciclo de Vida Dimensional de Kimball, que sirve como base teórica para el desarrollo de una solución BI que permita generar indicadores asociados al proceso de votación, para mejorar la toma de decisiones y cumplir con las expectativas de las organizaciones electorales.

### 3.1. SCRUM

En la década de los 80s, Hirotaka Takeuchi e Ikujiro Nonaka definieron una estrategia de desarrollo de producto flexible e incluyente donde el equipo de desarrollo trabaja en unidad para alcanzar un objetivo común. De esta manera, basaron su enfoque en casos de estudio de diversas industrias de fabricación y propusieron que el desarrollo no debe ser como una carrera de relevos secuencial, sino análogo al juego de rugby, donde el equipo trabaja en conjunto. [19]

Ken Schwaber y Jeff Sutherland desarrollaron el concepto de Scrum definiéndolo como un framework adaptable, iterativo, rápido, flexible y eficaz, para ofrecer un valor considerable en forma rápida a lo largo del proyecto. Así mismo, el framework está estructurado de tal manera que es compatible con el desarrollo de productos y servicios en todo tipo de industrias y en cualquier tipo de proyecto, independientemente de su complejidad.

Scrum tiene muchas características, entre ellas resaltan dos: la primera es que el desarrollo del software se realiza mediante iteraciones, denominadas Sprints, cuya duración es de dos (2) a cuatro (4) semanas. El resultado de cada sprint es un incremento ejecutable que se muestra al propietario del producto. La segunda característica importante son las reuniones a lo largo de del proyecto,

entre ellas destaca la reunión diaria de quince (15) minutos del equipo de desarrollo para la coordinación e integración.

### 3.1.1. Roles de Scrum

Así mismo, en la guía de SCRUMstudy [19] se explica que dicho método tiene dos (2) aspectos fundamentales a diferenciar: los actores y las acciones, donde los actores son los que ejecutan las acciones. De esta manera, los actores tienen roles y responsabilidades que son importantes para garantizar la implementación exitosa de los proyectos Scrum. En consecuencia, los roles se definen en dos grandes categorías:

- **Roles Centrales:** son aquellos que se requieren obligatoriamente para crear el producto del proyecto, están comprometidos con el proyecto y por último son los responsables del éxito de cada sprint del proyecto y del proyecto en su totalidad.
- **Roles No Centrales:** son aquellos que no son necesariamente obligatorios para el proyecto Scrum, y pueden incluir miembros de los equipos que tengan interés en el proyecto, pero que no tienen ninguna función formal en el equipo del proyecto. Pueden interactuar con el equipo, pero no son responsables del éxito del proyecto.

De lo anteriormente mencionado, también se explica que existen tres (3) roles centrales en Scrum que en última instancia tienen la responsabilidad de cumplir con los objetivos de del proyecto. Por lo tanto, los roles centrales son el Product Owner, el Scrum Master y el equipo de Scrum. Sin embargo, es importante tener en cuenta que de estos tres roles, ninguno tiene autoridad sobre los otros.

- **Product Owner:** es la persona responsable de maximizar el valor del negocio para el proyecto. Este rol es el responsable de articular los requisitos del cliente y de mantener la justificación del negocio del proyecto. En resumen, representa la voz del cliente.
- **Scrum Master:** el Scrum Master guía, facilita y enseña las prácticas de Scrum a todos los participantes en el proyecto, elimina los impedimentos que enfrenta el equipo y se asegura que todos estén siguiendo el proceso Scrum.
- **Equipo Scrum:** es un grupo o equipo de personas responsables de atender los requerimientos del negocio especificados por el Product Owner, de estimar las historias de usuarios y de la creación final de los entregables del proyecto.



En cuanto a los roles no centrales, se explica que son aquellos que no son obligatoriamente necesarios para el proyecto. Sin embargo, es importante tener conocimiento de éstos, ya que podrían desempeñar un rol importante en algunos casos.

- **Stakeholder(s):** es un término colectivo que incluye a clientes, usuarios y patrocinadores. Influyen en el proyecto a lo largo del desarrollo. Los stakeholders también pueden desempeñar un rol importante en los procesos de Scrum, por ejemplo: desarrollando épics, creando el Backlog priorizado del producto, realizando la planificación del lanzamiento y la retrospectiva del sprint.
- **Vendedores:** incluyen a los individuos u organizaciones que ofrecen productos y servicios que no están dentro de las competencias básicas de la organización del proyecto.
- **Scrum Guidance Body:** es un rol opcional, aunque altamente recomendado para formalizar las prácticas organizacionales relacionadas a Scrum. Por lo general, se compone de un grupo de documentos y/o un grupo de expertos que normalmente están involucrados en definir los objetivos relacionados con la calidad, regulaciones gubernamentales, seguridad y otros parámetros claves de la organización. También ayuda a captar las mejores prácticas que deben utilizarse en todos los proyectos Scrum de la organización.

### 3.1.2. Eventos de Scrum

Normalmente regulan y minimiza la necesidad de llevar a cabo reuniones no planeadas. Todos estos eventos son de tiempo limitado, lo que significa que tienen una duración máxima. En el momento que un evento comienza, tiene una duración fija que no puede cambiarse y una vez que se cumple el objetivo, todos los eventos restantes pueden finalizarse. Esto asegura que se gasta el tiempo mínimo posible en proceso. [20]

Cada evento da la oportunidad de inspeccionar o adaptar algo y ha sido diseñado para permitir una transparencia radical. Así pues, existen cuatro (4) eventos formales programados en el método Scrum que se describen a continuación.

### Planificación del Sprint

Cualquier trabajo que desee realizarse bajo esta metodología debe planificarse, ya que permite unir al equipo y crear un trabajo colaborativo. La planificación del sprint tiene una duración concreta

de ocho (8) horas como máximo para un Sprint de un mes. Cuando el Sprint dura menos, también el evento tiende a ser más corto. El Scrum Master enseñará al equipo a mantenerse en la duración establecida.

En la planificación se responden las preguntas sobre qué se puede liberar en el incremento del siguiente Sprint, y cómo se conseguirá realizar el trabajo que se va a liberar.

### **Scrum Diario**

Es una tarea iterativa que se realiza con todo el equipo de desarrollo, se lleva a cabo diariamente durante el periodo del sprint. En ella se busca que los miembros del equipo se actualicen el uno al otro referente a sus progresos y sobre los impedimentos, obstáculos o riesgos que impidan el normal avance, así como también las planificaciones de las tareas para el día. [19]

### **Revisión del Sprint**

En este proceso, se proporciona una demostración de los entregables al Product Owner y a los stakeholders relevantes, sólo si éstos cumplen con los criterios de aceptación definidos. El equipo de Scrum tiene un aproximado de dos (2) horas para mostrar los resultados del sprint, con el fin de asegurar la aprobación y aceptación del Product Owner. [21]

### **Retrospectiva del Sprint**

En este proceso, el Scrum Master y el equipo Scrum se reúnen para analizar las lecciones aprendidas durante todo el Sprint. Esta información se documenta en forma de lecciones aprendidas que pueden aplicarse en futuros sprints. Frecuentemente, como resultado de esta discusión, puede haber mejoras aceptadas o recomendaciones actualizadas por parte del Scrum Guidance Body. [19]

### **3.1.3. Procesos de Scrum**

Los procesos de Scrum abordan actividades específicas y el flujo de un proyecto de Scrum. En total hay diecinueve (19) procesos fundamentales que se aplican a todos los proyectos. Estos procesos se agrupan en cinco (5) fases tal como se muestran en la Tabla 3:

**Tabla 3: Procesos de Scrum [19]**

Fase	Procesos fundamentales de Scrum
Inicio	<ol style="list-style-type: none"> <li>1. Crear la visión del proyecto.</li> <li>2. Identificar al Scrum Master y al Stakeholder(s).</li> <li>3. Formar equipos Scrum.</li> <li>4. Desarrollar épicas.</li> <li>5. Crear el Backlog priorizado del producto.</li> <li>6. Realizar la planificación del lanzamiento.</li> </ol>
Planificación y Estimación	<ol style="list-style-type: none"> <li>7. Crear historias de usuario.</li> <li>8. Estimar historias de usuario.</li> <li>9. Comprometer historias de usuario.</li> <li>10. Identificar tareas.</li> <li>11. Estimar tareas.</li> <li>12. Crear el Sprint Backlog.</li> </ol>
Implementación	<ol style="list-style-type: none"> <li>13. Crear entregables.</li> <li>14. Realizar Daily.</li> <li>15. Refinar el Backlog Priorizado del producto.</li> </ol>
Revisión y Retrospectiva	<ol style="list-style-type: none"> <li>16. Demostrar y validar el Sprint.</li> <li>17. Retrospectiva del Sprint.</li> </ol>
Lanzamiento	<ol style="list-style-type: none"> <li>18. Enviar entregables.</li> <li>19. Retrospectiva del proyecto.</li> </ol>

En la Ilustración 13, se puede visualizar las fases del método Scrum:



**Ilustración 13: Fases del método Scrum [19]**

### 3.2. Ciclo de Vida Dimensional de Kimball

En su libro, *“The Data Warehouse Toolkit”* [10], el autor presenta y explica la metodología que hace uso de las técnicas para la construcción de un Data Warehouse. Se basa en la interpretación de la frase “Ciclo de Vida”, o mejor conocido por sus siglas en inglés, KCL (Kimball Life Cycle), que describe un flujo de tareas de alto nivel requeridas para el diseño, desarrollo e implementación de un Data Warehouse.

Kimball define que un Data Warehouse no es más que la unión de todos los Data Mart de una entidad, es decir, presenta un método de almacenamiento de datos Bottom Up en el que los Data Mart individuales ofrecen vistas finas de los datos de la organización que podrían ser combinados en un almacén de datos. Para lograr diseñar un Data Mart o un Data Warehouse propone la metodología que se muestra a continuación (Ver ilustración 14).

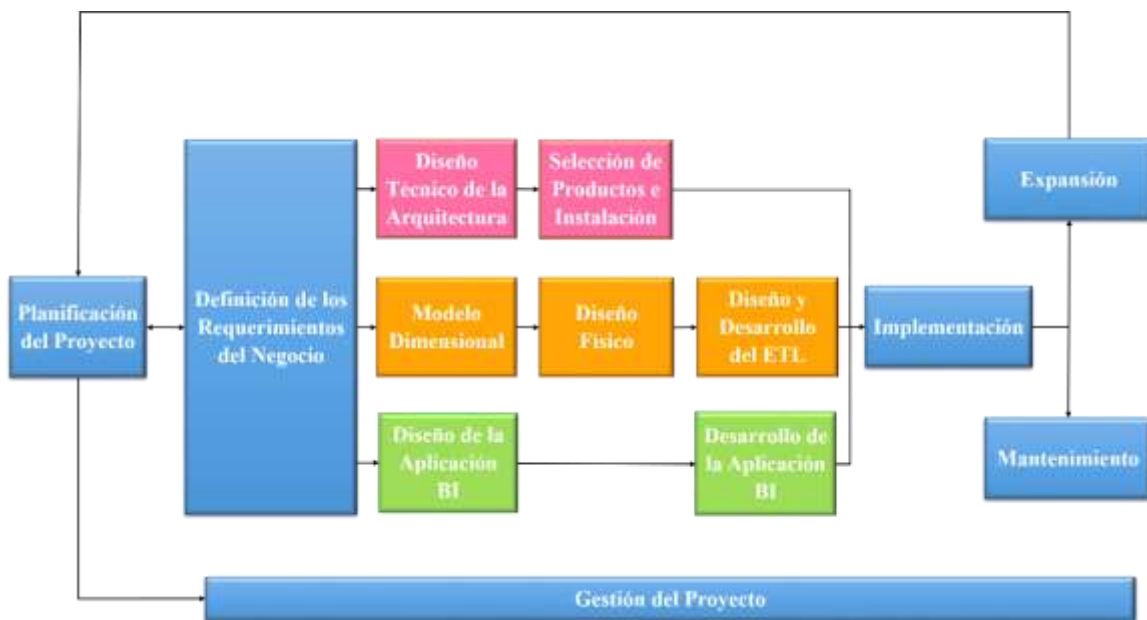


Ilustración 14: Ciclo de Vida Dimensional de Kimball

Esta estructura no posee una línea temporal absoluta, por el contrario se presentan tres rutas paralelas que parten de la definición de los requerimientos que se enfocan en tres áreas diferentes:

- **Línea Tecnológica:** es el camino superior. Implica tareas relacionadas con software específico.

- **Línea de Datos:** es el camino central. En este se diseña e implementa el modelo dimensional, y se desarrolla el subsistema de Extracción, Transformación y Carga (ETL – Extract, Transformation and Load) que alimenta el Data Warehouse.
- **Línea de Aplicaciones:** es el camino inferior. En esta ruta se encuentran las tareas para diseñar y desarrollar las aplicaciones de negocios para los usuarios finales.

A continuación, se describe con detalle las tareas del ciclo de vida de Kimball que permitirán formar parte de la metodología para este trabajo especial de grado:

## **Comienzo del Ciclo**

### **3.2.1. Planificación del Proyecto**

Según lo indica Rivadera [22], este es el primer paso que debe llevarse a cabo para poder iniciar el desarrollo del Data Warehouse. En este proceso se determina cuál es la situación actual de la organización y cuál será el propósito del proyecto de DW/BI, sus objetivos específicos y alcances del mismo, los principales riesgos y una aproximación inicial a las necesidades de información. Visto de otra forma, se refiere a una iteración simple del ciclo de vida de Kimball desde el lanzamiento hasta el despliegue. Esta tarea incluye las siguientes acciones típicas de un plan de proyecto:

- Definir el alcance (entender los requerimientos del negocio).
- Identificar las tareas.
- Programar las tareas.
- Planificar el uso de los recursos.
- Asignar la carga de trabajo a los recursos.
- Elaboración de un documento final que representa un plan del proyecto.

El plan de proyecto resultante identifica todas las tareas asociadas con el ciclo de vida del Data Warehouse e identifica las partes involucradas. (Kimball, R. y Ross, M., 2013).

### **3.2.2. Definición de los requerimientos del negocio**

En esta fase se realiza una interacción directa con la organización. Se debe aprender tanto como se pueda sobre el negocio, personal, clientes y competidores del mismo. Todo este

análisis se realiza por medio de los documentos internos, entrevistas con el personal y un estudio externo de la organización. [10]

Parte del proceso de preparación es averiguar a quien se debe realmente entrevistar. Esto normalmente implica examinar cuidadosamente el organigrama de la empresa. Hay básicamente cuatro grupos de personas con las que se debe hablar:

- El directivo responsable de la toma de decisiones estratégicas.
- Los administradores intermediarios y de negocio responsables de explorar alternativas estratégicas y aplicar decisiones.
- El personal de sistemas, si existe.
- El personal que se entrevista por razones políticas.

Los diseñadores de los Data Warehouse deben entender los factores claves que guían al negocio para determinar efectivamente los requerimientos y traducirlos en consideraciones de diseño apropiadas, pues son la base para las tres etapas paralelas subsiguientes, focalizadas en la tecnología, los datos y las aplicaciones, por lo cual es altamente crítica y el diseño es el centro de atención del Business Dimensional Lifecycle (BDL).

Esta fase tiene mucha influencia sobre la Planificación del Proyecto, es por eso que se tiene una doble flecha entre ambas cajas. De aquí se derivan las tres rutas que se comentaban anteriormente: línea tecnológica, línea de datos y línea de aplicaciones.

## **Línea Tecnológica**

### **3.2.3. Diseño Técnico de la Arquitectura**

En esta etapa se deben definir los planos que permitirán contar con un diseño integral que tome en cuenta los aspectos técnicos y elementos del Data Warehouse. Estos elementos son representados por medio de modelos que van en diferentes niveles de detalle mostrando los requerimientos inmediatos [23]. Para este diseño se siguen los pasos descritos a continuación:

- a) Establecer la fuerza de la arquitectura. Para ello es conveniente elegir a tres (3) personas en el diseño: un arquitecto técnico, un diseñador del área de intermedia y el desarrollador de aplicaciones.

- b) Recolectar los requerimientos del negocio. Esto se hace de acuerdo a las necesidades críticas del negocio, como pueden ser tiempos, disponibilidad, rendimiento, etc.
- c) Documentar los requerimientos de arquitectura. Lo cual se realizará a partir de las entrevistas, enfocado en los aspectos que pudieran impactar en la arquitectura.
- d) Clasificar los requerimientos de acuerdo a los datos de staging, acceso de datos, metadata e infraestructura para desarrollar un modelo de arquitectura de alto nivel.
- e) Determinar la fase de implementación de la Arquitectura. Estableciendo prioridades para la implementación de las definiciones hechas de acuerdo a los requerimientos del negocio.
- f) Realizar la documentación técnica de la Arquitectura. Éste debe contener la información necesaria para que se lleve a cabo la implementación del Data Warehouse.
- g) Revisión y finalización de la arquitectura técnica. La cual debe ser distribuida por los miembros de TI y el negocio con la finalidad de que sea retroalimentado este plan y quede completo para su validación.

### **3.2.4. Selección de Productos e Instalación**

De acuerdo a lo establecido en el diseño técnico de la arquitectura [23], se debe escoger un producto que se adapte lo mejor posible a las necesidades del plan. Para hacer una buena selección del mismo, se pueden seguir los siguientes pasos:

- a) Realizar una matriz de evaluación con las características más relevantes según el plan que se definió.
- b) Hacer una búsqueda en el mercado de todos los productos que cumplan con las características que se reflejaron en la matriz de evaluación.
- c) Reducir opciones al mínimo para realizar evaluaciones detalladas.
- d) Requerir pruebas de conceptos (POC<sup>3</sup>).
- e) Seleccionar el producto y negociar.
- f) Continuar con el ciclo de vida del proyecto.

---

<sup>3</sup> POC: Proof of Concept

## Línea de Datos

### 3.2.5. Modelado Dimensional

El modelado dimensional, según su creador Ralph Kimball [10], es el punto donde se debe hacer un análisis exhaustivo de los datos generados por el proceso. Si bien se realizó una auditoría de alto nivel durante la definición de los requerimientos del negocio, se necesita profundizar en lo esencial para evaluar la granularidad, la consistencia histórica, los valores válidos y la disponibilidad de los atributos.

Cada modelo dimensional contiene una tabla con una clave compuesta llamada tabla de hechos y un conjunto de tablas más pequeñas llamadas dimensiones. Cada tabla dimensión tiene una clave primaria simple, que a su vez forma parte de la clave compuesta de la tabla de hechos.

La creación de un modelo dimensional es un proceso dinámico y altamente iterativo. El proceso de diseño comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados de la matriz descrita en el punto anterior. [22]. Este proceso consiste en cuatro pasos:

- **Elegir el proceso de negocio:** consiste en elegir el área a modelar. Esta es una decisión de la dirección y depende fundamentalmente del análisis de requerimientos y de los temas analíticos anotados en la etapa anterior.
- **Establecer el nivel de granularidad:** en este punto se debe enfocar en el nivel de detalle con el cual será desarrollado el Data Warehouse. La granularidad depende de los requerimientos solicitados por el negocio y lo que es posible a partir de los datos actuales. La sugerencia general es comenzar a diseñar el Data Warehouse al mayor nivel de detalle posible, ya que se podrían realizar agrupamientos posteriores, al nivel deseado.
- **Elegir las dimensiones:** las dimensiones surgen naturalmente de las discusiones del equipo en conjunto con la elección del nivel de granularidad y de la matriz de procesos/dimensiones que se realiza en la definición de los requerimientos del negocio. Las tablas de dimensiones tienen un conjunto de atributos que brindan una perspectiva o forma de análisis sobre una medida en una tabla de hechos. Una forma de identificar las tablas de dimensiones es que sus atributos son posibles candidatos para ser encabezados en los informes, tablas pivote, cubos, o cualquier forma de visualización, unidimensional o multidimensional.



- **Identificar medidas y las tablas de hechos:** Este paso consiste en identificarlas medidas que surgen de los procesos de negocios. Una medida es un atributo (campo) de una tabla que se desea analizar, sumando o agrupando sus datos y usando los criterios de corte conocidos como dimensiones. Las medidas habitualmente se vinculan con el nivel de granularidad del paso 2, y se encuentran tablas que denominamos tablas de hechos. Cada tabla de hechos tiene como atributos una o más medidas de un proceso organizacional, de acuerdo a los requerimientos. Un registro contiene una medida expresada en números, como ser cantidad, tiempo, dinero, etc., sobre la cual se desea realizar una operación de agregación (promedio, conteo, suma, etc.) en función de una o más dimensiones. La granularidad, en este punto, es el nivel de detalle que posee cada registro de una tabla de hechos.

### 3.2.6. Diseño Físico

El diseño físico se centra en la selección de las estructuras necesarias que permitan soportar el diseño lógico. Alguno de los elementos principales de este proceso son la definición de nombres y datos específicos del ambiente de la base de datos basándose en convenciones estándares. Tal como lo explica Rivadera [22], en esta fase se intentan resolver preguntas basándose en los siguientes planteamientos:

- Identificar cuáles son los factores de uso que llevarán a una configuración más grande y más compleja.
- Definir cómo se debe configurar el sistema. Por ejemplo, especificar cuánta memoria, servidores, tipo de almacenamiento y procesadores se necesitan.
- Definir si las estaciones de trabajo para el equipo de DW/BI necesitan instalar algún componente específico.
- Plantear una forma adecuada de convertir el modelo de datos lógico en un modelo de datos físico en la Base de Datos Relacional.
- Plantar si es necesario usar la partición en las tablas relacionales.
- Definir cómo conseguir un plan de indexación inicial.

### **3.2.7. Diseño y Desarrollo del ETL**

El sistema de Extracción Transformación y Carga (ETL) es la base sobre la cual se alimenta el Data Warehouse. Si éste se diseña adecuadamente, se puede extraer los datos de los sistemas de origen, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos, consolidar la información proveniente de distintos sistemas, y finalmente cargar (grabar) la información en el Data Warehouse en un formato acorde para la utilización por parte de las herramientas de análisis. [22]

En otras palabras, esta es la capa encargada de la integración de los datos desde las múltiples fuentes existentes para la construcción del Data Warehouse y los Data Marts de la base analítica. Las rutinas ETL tienen la función de manipular los datos de las distintas fuentes de datos a fin de integrarlos todos al Data Warehouse según formatos y estructuras que requieren. Aquí también se limpian los datos, removiendo problemas transaccionales y asegurando la consistencia de los mismos.

Los procesos ETL se pueden describir de la siguiente manera:

#### **Extracción**

La primera etapa consiste en extraer los datos de su origen o fuente. Cada una de las fuentes de datos identificadas tiene sus propias características y atributos, que serán manejados dentro del proceso. Por lo tanto, la función principal de esta etapa es dejar un formato preparado para dar inicio al proceso de transformación.

Para la extracción es conveniente tener en cuenta el manejo adecuado de las diferentes fuentes de datos, conocer su estructura y los diferentes componentes y especificaciones necesarias para que sean conectadas de forma efectiva.

Otro punto que hay que tener en cuenta es que si la cantidad de datos a extraer es muy elevada, el sistema se puede ralentizar, o incluso colapsarse, por lo que las grandes operaciones de extracción se suelen realizar en horas donde el impacto sobre el sistema sea el mínimo posible.

Entonces, podemos puntualizar esta etapa de la siguiente manera:

- Se considera la consolidación de los datos desde las diversas fuentes que usan diferentes estructuras y formatos.

- Establece el acceso a las fuentes de datos.
- Extrae eficientemente cada una de las fuentes de datos, con el mejor procesamiento posible, buscando solo los datos requeridos.
- Se pueden almacenar datos temporalmente en un área especializada para su transformación e integración con otras fuentes.

## **Transformación**

En esta fase es importante realizar un ajuste y limpieza de los datos, con el fin de tener estos de forma precisa, completa y coherente. Es decir, se aplican una serie de reglas o funciones a los datos extraídos desde la fuente para generar los datos a cargar al destino final (el Data Warehouse o Data Mart).

Algunos datos requerirán muy poca o ninguna manipulación. En otros casos, se requerirá una serie de pasos de transformación para satisfacer las necesidades técnicas y del negocio de la base analítica.

Algunos ejemplos de transformaciones son la generación de nuevos valores ( $\text{monto\_venta} * \text{cantidad} * \text{precio\_unitario}$ ), filtrar, ordenar, unir datos de múltiples fuentes, generar claves sustitutas, etc.

## **Carga**

Este proceso puede variar dependiendo de los requerimientos de la organización. Algunos Data Warehouse sobrescribirán la información con información acumulada. Otros, o incluso partes del mismo, pueden agregar información históricamente.

Las ventanas de tiempo y el alcance para reemplazar o concatenar datos son decisiones estratégicas de diseño que dependerán del tiempo disponible y las necesidades del negocio.

En la fase de carga, al interactuar con la base de datos, aplican las restricciones definidas en el esquema de la Base de Datos, lo que también contribuye al desempeño global de la calidad de los datos en el proceso de ETL.

## **Línea de Aplicación**

### **3.2.8. Diseño de la aplicación BI**

No todos los usuarios del Data Warehouse necesitan el mismo nivel de análisis. Es por ello que en esta etapa se identifican los diferentes roles o perfiles de usuarios para determinar los diferentes tipos de aplicaciones necesarias en base al alcance de los diferentes perfiles (gerencial, analista del negocio, vendedor, etc.) [10]

### **3.2.9. Desarrollo de la Aplicación BI**

Siguiendo a la especificación de las aplicaciones para usuarios finales, el desarrollo de las mismas involucra configuraciones del metadata<sup>4</sup> y construcción de reportes específicos. Una vez que se ha cumplido con todos los pasos de la especificación y se tiene la posibilidad de trabajar con algunos datos de prueba, comienza el desarrollo de la aplicación. [10]

## **Finalización del Ciclo**

### **3.2.10. Implementación**

En esta fase de la metodología es necesario contemplar todos aquellos factores que durante fases iniciales y de diseño no fueron contempladas. Se debe realizar una última revisión sobre el diseño lógico, determinando cuáles hechos deben ser agregados con respecto a las dimensiones, para mejorar el rendimiento del Data Warehouse.

Es importante culminar el diseño físico, pues a partir del estado de éste, se debe realizar un corte preliminar en el tamaño de la base de datos y su respectiva tasa de crecimiento. Al mismo tiempo, cuando se tienen definidas todas las tablas, se debe hacer un corte para la estrategia de establecimiento de índices. A partir de este momento se puede dar por concluido el diseño físico y se puede iniciar realmente el proceso de implementación pues se cuenta con suficiente información acerca de lo que ocurre en la base de datos con sus mínimos detalles.

---

<sup>4</sup> Se refiere a la información acerca de las estructuras de datos contenidas en las bases de datos. También es conocido por la frase “Son las “enzimas” que ayudan a digerir los datos”, o simplemente datos que describen datos.

### **3.2.11. Mantenimiento**

En esta fase se realizan tareas de soporte técnico, como por ejemplo, resolver problemas técnicos presentados con la plataforma, manejo de la data, respaldo, entre otros. También es un área donde se puede educar al personal que interactúe con el producto. Dicho de otra manera, se realiza la verificación del estado de salud del proyecto. En esta etapa se pueden realizar informes con las posibles mejoras que requiera el sistema, pasando de esta manera a la fase de expansión. [23].

### **3.2.12. Expansión**

Se refiere a las posibles mejoras e innovaciones que puede sufrir el modelo. Permite la evolución del mismo enfocándose en la experiencia del usuario final. Dicha expansión puede darse, porque se requiera un nuevo módulo o porque existan mejoras que hacer sobre el proyecto ya existente.

## **Gestión del Proyecto**

### **3.2.13. Gestión del Proyecto**

Asegura que las actividades del ciclo del Data Warehouse se lleven en forma correlativa y sincronizada. Como lo indica el diagrama, la gestión del proyecto acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra el monitoreo del estado del proyecto y la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos. [10].

# CAPÍTULO 4

## Marco Aplicativo

Luego de haber mencionado todos los conceptos en el capítulo 2 y la metodología planteada en el capítulo 3 de este documento, en este capítulo se explican todas las actividades que se siguieron para la implementación de la solución BI que satisfacen los requerimientos del caso de estudio.

Haciendo uso se la metodología SCRUM, este proyecto se divide en tres (3) fases: Inicio, Desarrollo y Finalización. En una de las iteraciones del desarrollo, se integra el Ciclo de Vida Dimensional de Kimball, que sirve de base para el desarrollo de la solución de Inteligencia de Negocio. En la Ilustración 15, se presenta la adaptación realizada entre el método Scrum y el Ciclo de Vida de Kimball.

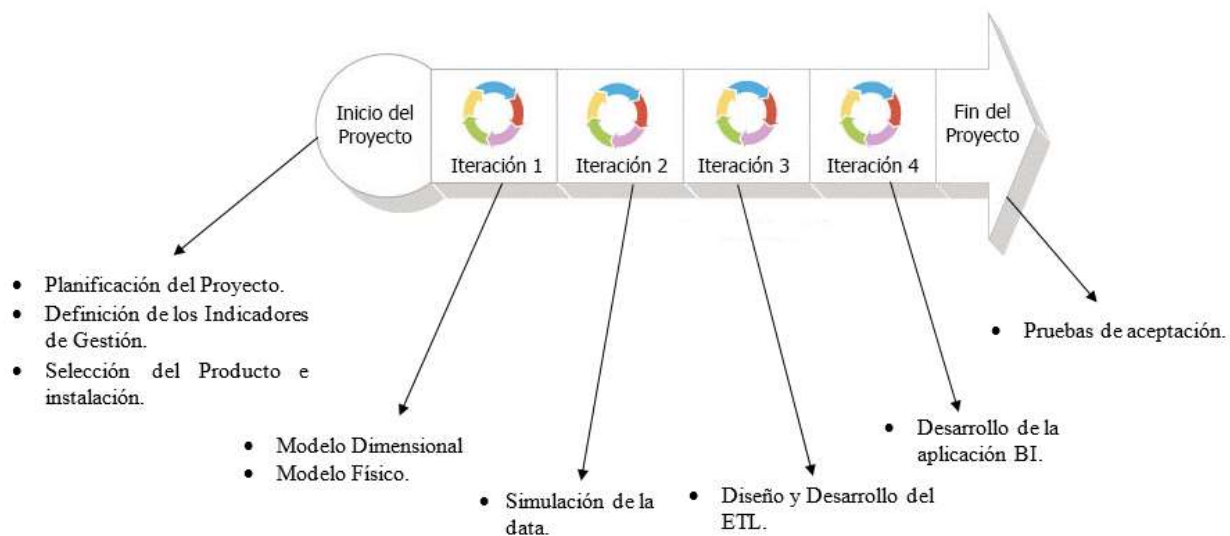


Ilustración 15: Adaptación de la metodología Scrum y el Ciclo de Kimball

### 4.1. Fase I: Inicio del Proyecto

Abarca los objetivos del uno (1) al cuatro (4) descritos en la sección 1.2.2 de este documento. A continuación, se detalla un plan de proyecto, el análisis de la BD relacional, los indicadores de gestión y la selección de los productos a utilizar durante el desarrollo de este TEG. Por lo tanto, se divide en cuatro (4) sub-fases:

#### 4.1.1. Planificación del Proyecto

En esta etapa es definido el equipo de desarrollo, los requerimientos del negocio, los tiempos a trabajar, los ambientes de implementación y las complejidades de cada actividad. Así mismo, se definen los pasos a seguir en cada etapa del proyecto y la arquitectura tecnológica a emplear. Los recursos disponibles en la organización influyen directamente sobre la capacidad que ésta posee para evolucionar o resolver un problema. Por esa razón, la planificación es la clave para el éxito del proyecto.

A continuación, en la Tabla 4 se muestra la planificación del proyecto, lo que permite completar el objetivo específico número uno (1).

**Tabla 4: Planificación del Proyecto**

Pasos	Descripción	Estimado
1.	Selección del tema	1 día
1.1.A	Tormenta de ideas	2 días
1.1.B	Consulta de recursos informativos disponibles	3 días
1.2	Selección de recursos disponibles para trabajar	2 días
2.	Selección del método de desarrollo	2 días
2.1.A	Definición de entregables	1 día
2.1.B	Definición de la arquitectura de la solución de Inteligencia de Negocio	1 día
3.	Selección de la Base de Datos	1 día
4.	Análisis de las fuentes de datos transaccionales	5 días
5.	Diseño del proyecto	5 días
6.	Formulación de indicadores	5 días
6.1	Selección de indicadores	2 días
7.	Simulación de la data	20 días
8.	Diseño lógico de la Base de datos	10 días
8.1.	Diseño físico de la base de datos	10 días
9.	Diseño de los ETL	20 días
10.	Diseño lógico de la presentación de resultados	2 días
11.	Elaboración de informe	15 días

En la planificación, solo las actividades que se encuentran enumeradas mediante el uso de números y letras, por ejemplo 1.1.A y 1.1.B, se pueden llevar a cabo de forma paralela. Mientras que el resto de ellas se deben realizar de forma secuencial.

#### **4.1.2. Análisis de la Base de Datos Relacional**

En este punto, para dar cumplimiento al objetivo específico número dos (2), se procede a analizar la base de datos relacional que sirve como fuente de datos para el desarrollo e implementación de la solución de Inteligencia de Negocio que fue propuesta anteriormente.

En principio, se observa que esta BD se encuentra implementada en OracleDB 12C release 2, donde todas sus tablas se encuentran en un solo esquema denominado TVT. La misma, cuenta con un usuario de acceso a este esquema, el cual posee los privilegios básicos para la manipulación de los datos.

Dicho esquema cuenta con un conjunto de tablas que conforman el modelo de seguridad de TVT, las cuales permiten almacenar las credenciales de los usuarios del sistema, así como asignación de perfiles con asignación de acceso sobre los distintos módulos del sistema.

También, cuenta con el modelo de configuración de la elección, donde además permite configurar el tipo de elección a trabajar. Por ejemplo: políticas, educacionales, deportivas, cooperativas, premiaciones, entre otras. Y dependiendo del evento requerido se puede configurar el comprobante de voto impreso o que solo se almacene digitalmente.

Por otro lado, se tiene configuración de encuestas, donde se puede armar la misma paso a paso, definiendo las preguntas de tipo abierta o cerrada y las opciones para cada una de ellas. Además, permite almacenar el resultado de cada una de las encuestas aplicadas.

Adicionalmente, posee un conjunto de tablas para almacenar las entidades o regiones donde será ejecutada la elección en cuestión. Dichas entidades, se encuentran bajo un esquema de jerarquía y permite asociar casi todos los procesos de TVT con una localidad.

También, contiene una tabla destinada a almacenar las incidencias del sistema, asociadas a cada uno de los usuarios. Con el fin de llevar el control y brindar soporte técnico sobre la aplicación a los usuarios del sistema.



De la misma manera, otro grupo de tablas se encuentran destinadas a almacenar los datos de los postulantes de la elección y la información de partidos políticos y sus alianzas en caso de que se requiera.

Finalmente contiene un conjunto de tablas que se avocan al proceso de captura, transmisión, transcripción y escrutinio de actas, para así cumplir con los requerimientos de los dieciocho (18) módulos de TVT que fueron descritos en el marco conceptual.

Así mismo, la BD cuenta con algunas funciones y paquetes que suministra los datos necesarios a cada módulo de la aplicación.

#### **4.1.3. Definición de los Indicadores de Gestión**

Luego del análisis previo, es importante tener claro que un sistema automatizado de votación puede ser muy amplio de estudiar, por lo que es posible levantar tantos indicadores como se desee de cada módulo. Sin embargo, como este TEG tiene un alcance de desarrollo limitado, se escogió un pequeño conjunto de indicadores generales y sencillos para abarcar los módulos del sistema más relevantes y así poder mostrar el beneficio de la solución BI.

En consecuencia, a fin de cumplir con el objetivo específico número tres (3), previamente analizando la BD relacional y los módulos del sistema, se definen los indicadores de negocio que se muestran a continuación:

1. Total de respuestas por opción de cada pregunta.
2. Total de encuestas por región.
3. Porcentaje de encuestas aplicadas por un encuestador.
4. Rendimiento de un usuario encuestador.
5. Porcentaje de votos por medio de votación.
6. Total de votos por partido y por postulante.
7. Afluencia de votación.
8. Cantidad de incidencias.
9. Total de actas capturadas por región.
10. Porcentaje de actas según su estatus.
11. Rendimiento de un usuario capturador.

En otras palabras, se puede observar como los indicadores 1, 2, 3 y 4 tienen relación directa con el módulo de encuestas. De la misma forma, los indicadores 5, 6 y 7 mantienen una relación directa con los votos almacenados y los medios configurados en la aplicación. Así mismo, el indicador 8, se centra en el módulo de soporte y finalmente, los indicadores 9, 10 y 11 con el proceso de captura, transmisión, digitalización y escrutinio del acta.

La idea central es poder agrupar toda la información que se almacena en la Base de Datos transaccional y mediante los indicadores definidos poder pasar al diseño de la Base de Datos dimensional.

Adicionalmente, los detalles de cada uno de los indicadores mencionados se pueden consultar en la tabla 23 que se encuentra en el anexo B de este documento.

#### 4.1.4. Selección de Productos de Software

Las aplicaciones seleccionadas para implementar la solución de Inteligencia de Negocio son las que se mencionan a continuación y pueden ser observadas en la Ilustración 16.

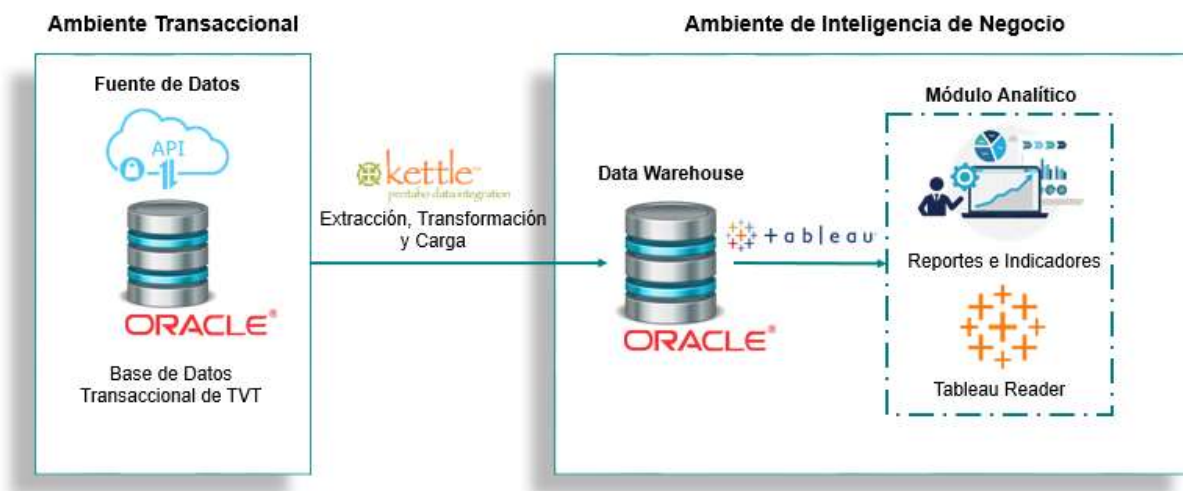
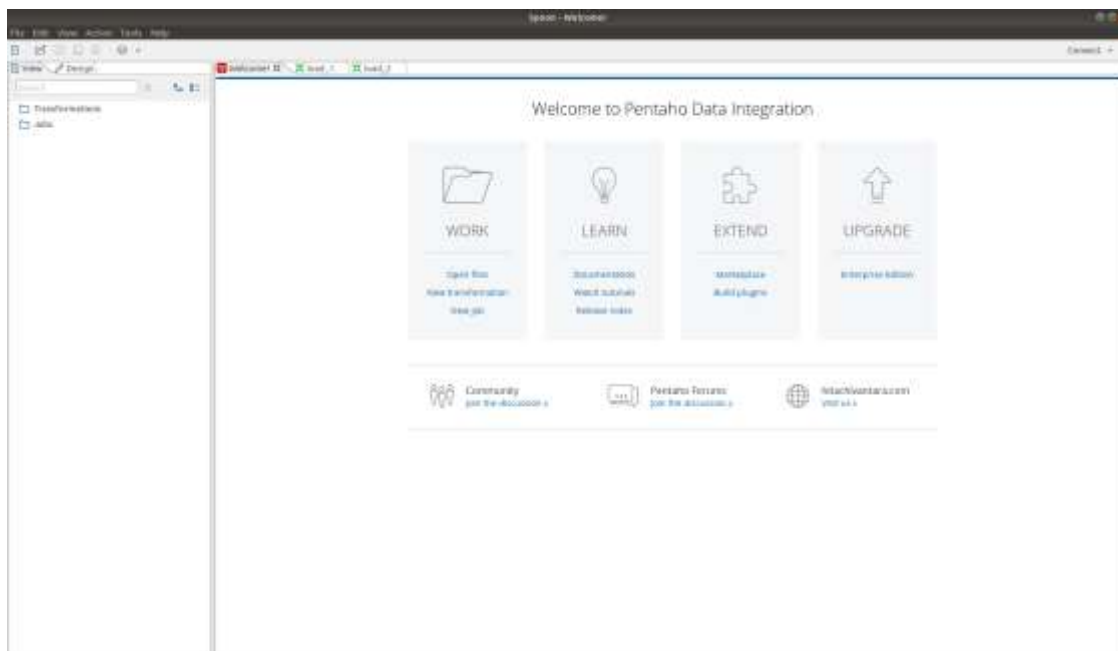


Ilustración 16: Arquitectura de la Solución BI

- **Oracle Database:** elegida por ser uno de los motores de base de datos relacional más confiables y más utilizado. Es considerado uno de los sistemas de base de datos más completos destacando el soporte de transacciones, la escalabilidad, la estabilidad y el soporte multiplataforma. Adicionalmente, las últimas versiones de Oracle han sido certificadas para poder trabajar bajo GNU/Linux.

- **Pentaho Data Integration:** elegida por permitir la extracción de los datos de diferentes fuentes, así como consolidarlos y cargarlos en un repositorio analítico. Su entorno de trabajo está diseñado de tal manera que permite que la implementación y el mantenimiento de la solución de Inteligencia de Negocio sea más sencillo y más rápido. De la misma manera, permite que el desarrollador pueda incorporar nuevas fuentes de información e indicadores. A continuación, se muestra la interfaz de la aplicación en la Ilustración 17.



**Ilustración 17: Interfaz de Pentaho Data Integration**

- **Tableau:** elegido por ser un software de análisis de datos con una excelente capa de visualización y presentación, considerado por muchos como uno de los mejores programas para la presentación visual de datos y con muy alta clasificación en la facilidad de uso. Tableau permite arrastrar y soltar grandes cantidades de datos en un "lienzo" digital y realizar gráficas al instante. La idea detrás de cómo funciona Tableau es que es más fácil de manipular lo que está pasando en la interfaz para que se pueda ver lo que se está haciendo bien y, por extensión, lo que se está haciendo mal. Tableau ha sido definido como una forma más avanzada de la tabla dinámica. A continuación, se muestra una ilustración de la interfaz de la aplicación. (Ver Ilustración 18).



**Ilustración 18: Interfaz de Tableau**

Así, una vez elegidas todas las herramientas necesarias para el desarrollo de la solución BI, se procede a realizar la instalación de los mismos, permitiendo cumplir con el objetivo específico número cuatro (4).

## 4.2. Fase II: Desarrollo

En esta fase se abarcan los objetivos específicos cinco (5), seis (6), siete (7) y ocho (8) definidos para el desarrollo de la solución de Inteligencia de Negocio y el tiempo estimado de cada una de las iteraciones. Basado en los tiempos propuestos por el método Scrum, se define de 1 a 4 semanas por sprint.

En total se realizan cuatro (4) iteraciones para cumplir con la fase del desarrollo. En la Tabla 5 se pueden observar cada una de ellas.

**Tabla 5: Iteraciones Scrum para el desarrollo de la solución BI**

Iteración	Descripción	Estimado
1	Diseño lógico o modelamiento dimensional. Diseño Físico.	2 Semana
2	Simulación de la data.	2 Semanas
3	Diseño y desarrollo del ETL.	2 Semanas
4	Desarrollo de la aplicación BI.	2 Semanas

## 4.2.1. Modelo dimensional

En esta sección se encuentra detallado el diseño lógico del Data Warehouse que forma parte de la solución BI. Basado en los indicadores levantados, surgieron 4 procesos de negocio, los cuales serán descritos a continuación de manera individual para mayor entendimiento.

### Diseño lógico del Proceso de Negocio: Encuesta

Este proceso de negocio se encarga de evaluar cómo va la tendencia de las votaciones basándose en las encuestas diseñadas previamente en la aplicación. A través de este modelo se puede realizar el cálculo de cuántas encuestas han sido aplicadas en una determinada región, así como también la cantidad de respuestas para cada una de las preguntas configuradas y adicionalmente se puede evaluar el rendimiento de los usuarios encuestadores. A continuación, se muestra en la ilustración 19 de dicho modelo.

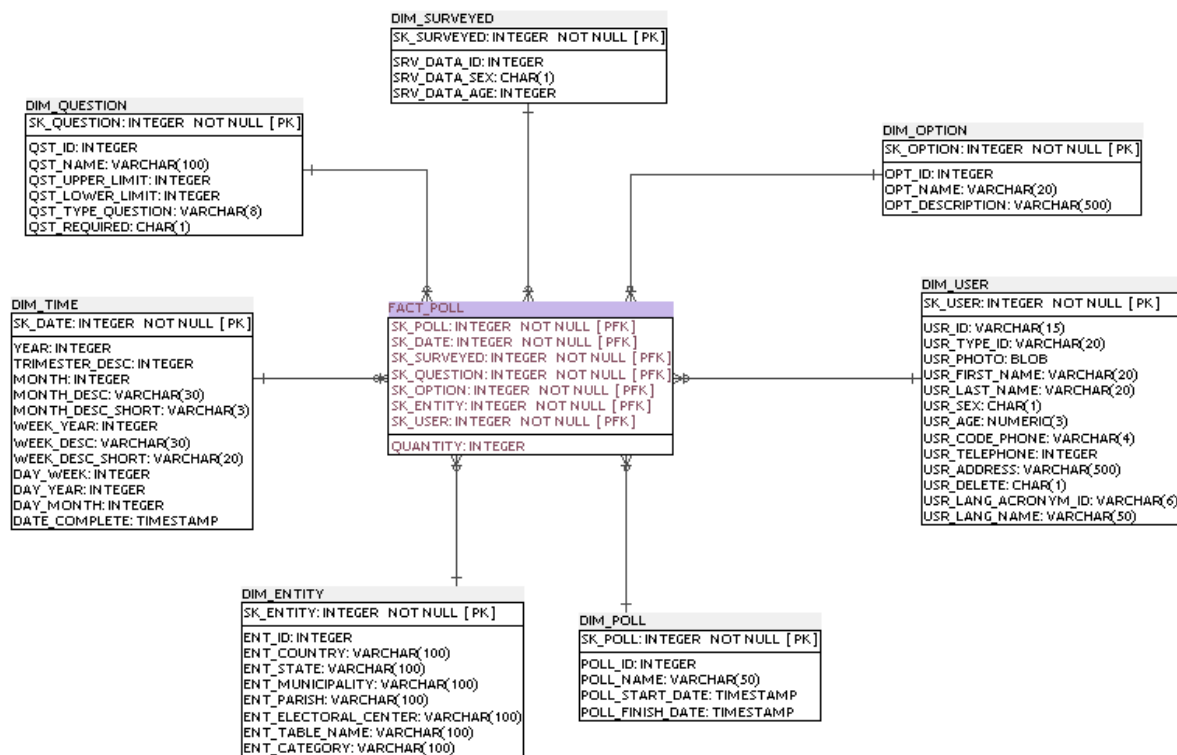


Ilustración 19: Modelo Dimensional del Proceso de Negocio: “Historial de Encuestas”

## Diseño lógico del Proceso de Negocio: Totalización de Votos

En este proceso se busca realizar un conteo de los votos para emitir los reportes correspondientes en la aplicación. Dichos votos deben ser contados por candidato y por partido político, así como también deben ser agrupados por región. Los reportes que pertenecen a este proceso de negocio deben comprender los porcentajes de votos por medio de votación y la afluencia de votación para referencias futuras. A continuación, se muestra el modelo que corresponde a este módulo del proceso BI. (Ver ilustración 20)

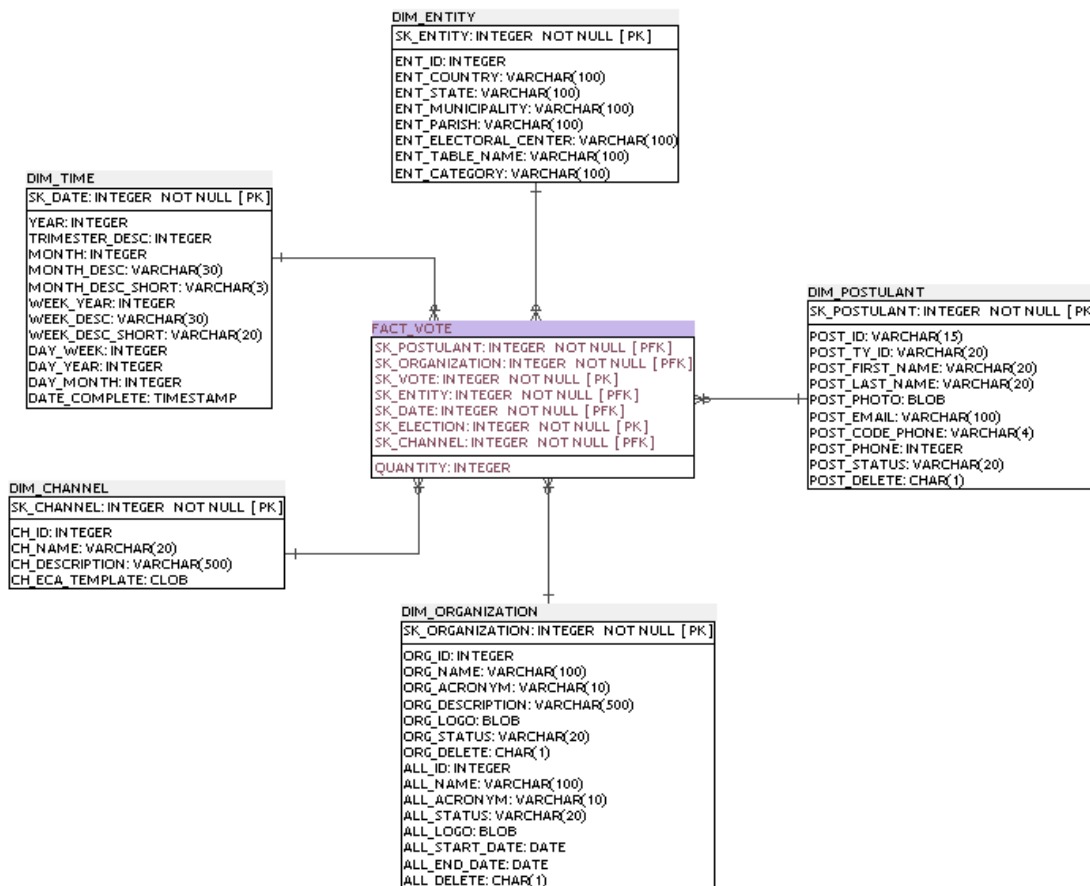
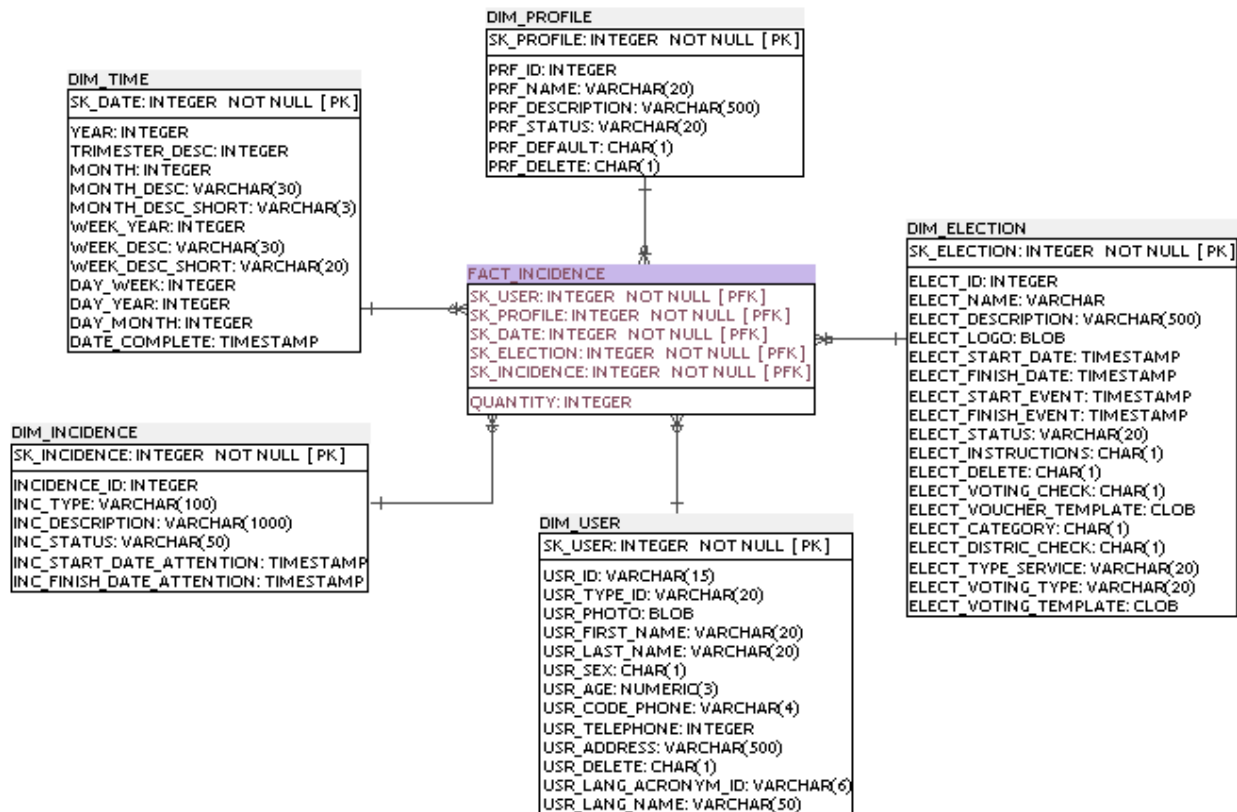


Ilustración 20: Modelo Dimensional del Proceso de Negocio: “Totalización de Votos”

## Diseño lógico del Proceso de Negocio: Incidencias en el sistema

Este proceso de negocio depende del área de soporte técnico y está totalmente dedicado a los dueños del producto, de tal manera que puedan observar en qué puntos deben hacer refuerzos o mejoras al sistema. Los reportes contemplados para este proceso se enfocan en visualizar la

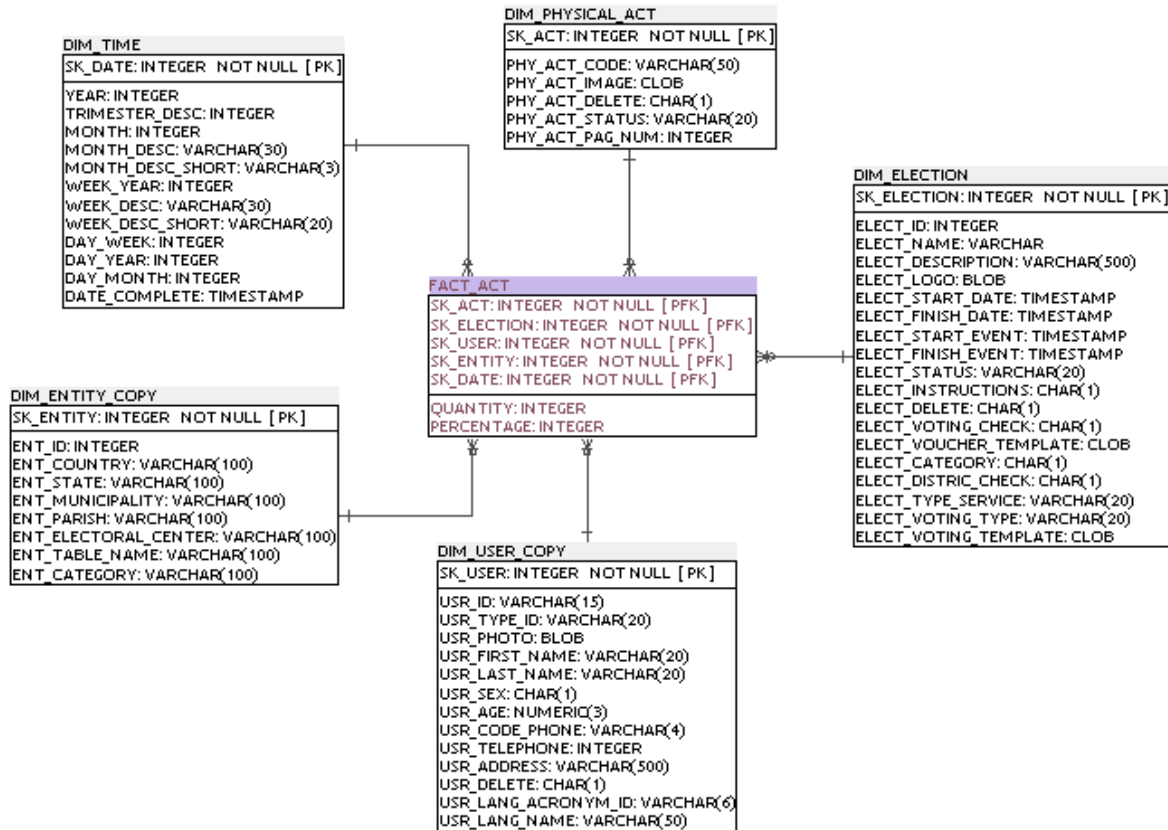
cantidad de incidencias que se tuvieron durante una elección o totalización de votos. Así, surge el siguiente modelo dimensional. (Ver ilustración 21)



**Ilustración 21: Modelo Dimensional del Proceso de Negocio: “Historial de Incidencias”**

## Diseño lógico del Proceso de Negocio: Escaneo de Actas

Es quizá uno de los procesos más relevantes, ya que está dedicado al proceso de captura de las actas, siendo éste uno de los módulos con mayor impacto del sistema. Así pues, los indicadores que se tomaron en cuenta fueron: total de actas capturadas por región, porcentaje de actas según su estatus y rendimiento de un usuario con perfil capturador. De esta manera, con base en estos requerimientos se plantea el siguiente modelo dimensional. (Ver ilustración 21)



**Ilustración 22: Modelo Dimensional del Proceso de Negocio: “Captura de Actas”**

#### 4.2.2. Diseño Físico

Partiendo de cada uno de los modelos descritos anteriormente, los cuales conforman en conjunto el modelo dimensional que da base al desarrollo de la solución de Inteligencia de Negocio, se implementa un único esquema llamado Dimensional Total Vote Tracker (DTVT). Así, se utiliza el SMDB Oracle que es manipulado a través de la herramienta SQL Developer.

Este esquema es el Data Warehouse de dicha solución y está compuesto por dieciséis (16) dimensiones más cinco (5) tablas de hechos. En el anexo A de este documento puede encontrar el diccionario de datos que describe detalladamente cada uno de las tablas que lo conforman.

#### 4.2.3. Desarrollo de los datos simulados

Con el fin de resguardar cualquier información confidencial de los clientes, es necesario simular los datos que deberá contener la BD transaccional y que posteriormente a través del proceso ETL, llenará la BD dimensional.



Es importante acotar que este proceso se realiza de manera manual, teniendo en cuenta todos los detalles de la lógica entre las tablas. Por tal motivo, es posible que no se refleje una cantidad grande de datos en los gráficos ya que no todos los datos pueden ser generados de manera aleatoria por un programa.

#### 4.2.4. Diseño y desarrollo del ETL

Una vez que ya se encuentra el diseño físico del Data Warehouse, se procede a diseñar y construir lo que será el proceso de extracción, transformación y carga de los datos con el apoyo de la herramienta Pentaho Data Integration. De manera individual, se construyen las transformaciones que realizan la carga de cada una de las dimensiones y tablas de hechos. Finalmente, una vez que se tienen todas las transformaciones individuales, éstas son unidas con un Job. En total se construyen veintiún (21) transformaciones y un (1) Job.

Las transformaciones de las dimensiones cuentan todas con la misma estructura. Se basan en un paso llamado “Table\_input” y en un paso llamado “Dimension\_lookup”. En la Ilustración 23 se muestra el ejemplo del ETL para una de las dimensiones.



**Ilustración 23: ETL para la Dimensión de Medios de votación**

Las transformaciones que corresponden a las tablas de hechos trabajan de igual manera con un “Table\_input” pero en lugar de tener un “Dimension lookup”, trabajan con un paso llamado “insert/update”. En la ilustración 24 se muestra el ejemplo del ETL para una tabla de hechos.



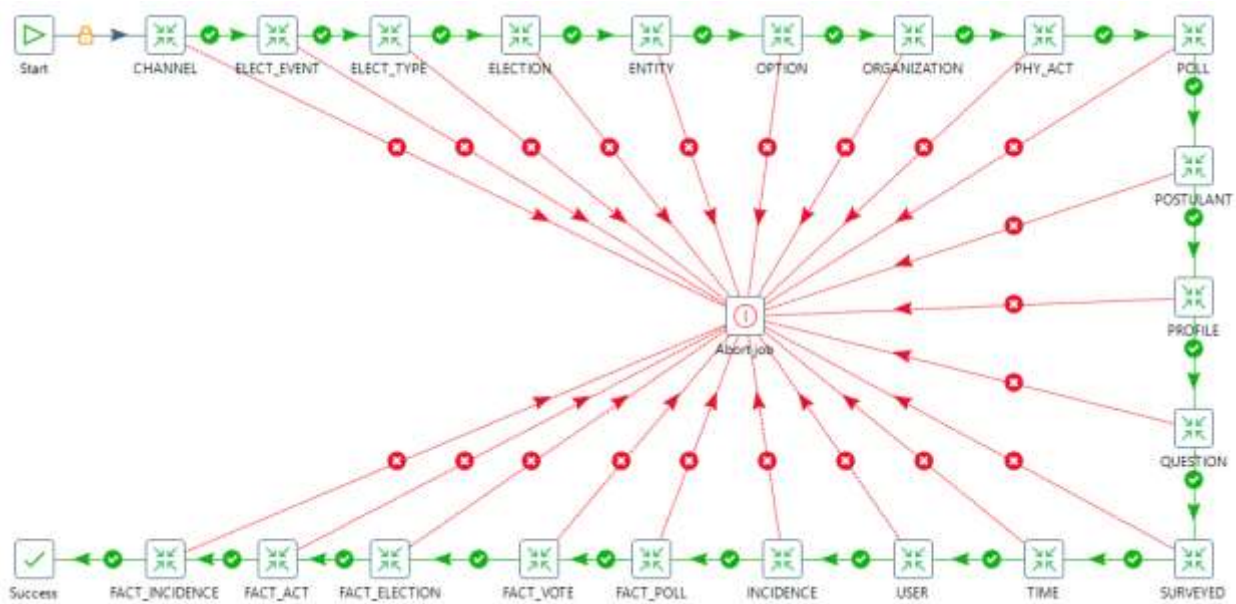
**Ilustración 24: ETL para la Tabla de Hechos del Proceso de Negocios: “Captura de Acta”**

Adicionalmente, hizo falta contemplar en la dimensión tiempo las horas y los minutos para los cálculos del sistema. Quedando la dimensión tiempo como se muestra en la Ilustración 25.



**Ilustración 25: ETL para la Dimensión Tiempo**

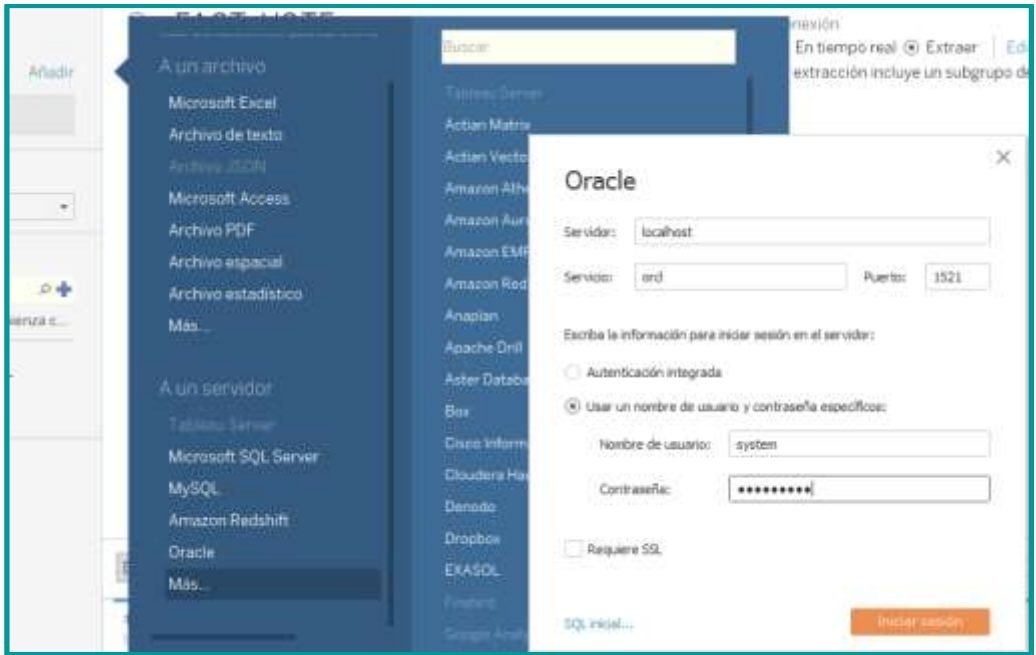
Finalmente, se crea el Job que une todas las transformaciones. Este consta con un punto de partida (Start) y un punto de finalización exitosa (Success), de la misma manera se contempla trabajar con “Abort Job” para el manejo de errores. Así pues, el Job queda de la siguiente manera. (Ver Ilustración 26).



**Ilustración 26: Job para el proceso de ETL**

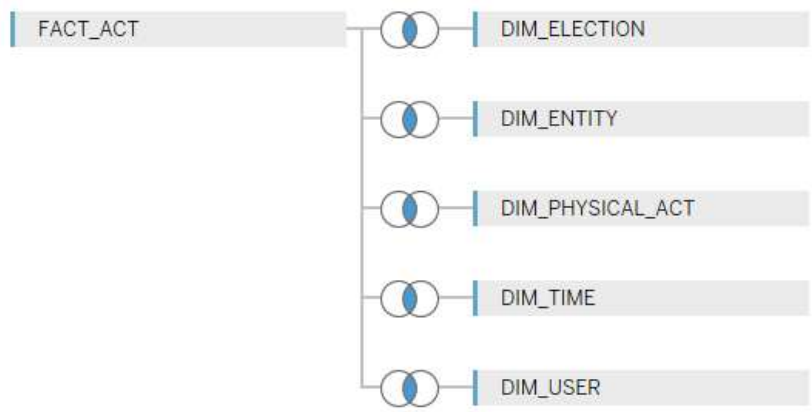
#### 4.2.5. Desarrollo de la aplicación BI

El primer paso para el desarrollo de esta actividad es establecer la conexión entre Oracle y Tableau, esto se logra directamente en Tableau, mediante la pestaña “Fuentes de Datos”. (Ver Ilustración 27).

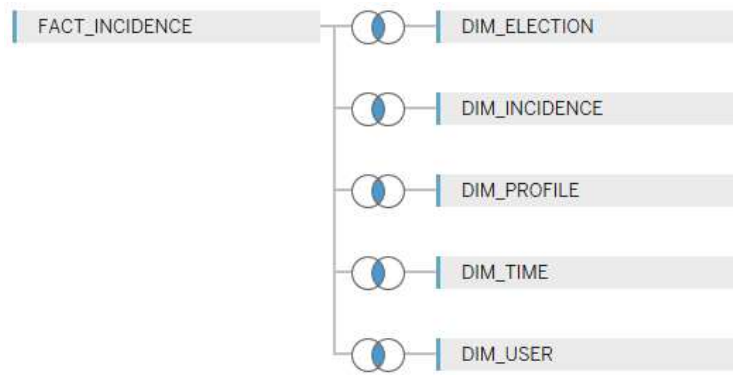


**Ilustración 27: Ejemplo de conexión con la fuente de datos desde Tableau**

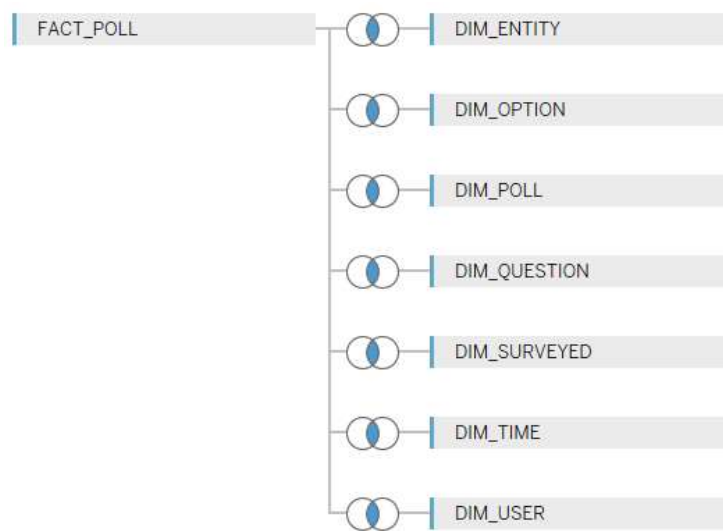
Para poder representar en Tableau el esquema de constelación de estrellas, diseñado e implementado en las iteraciones anteriores, se debe añadir una fuente de datos por cada estrella (proceso de negocio), con el fin de que no se crucen mal los datos de cada estrella pero se pueda trabajar en conjunto todos los procesos de negocio. A continuación, en las Ilustraciones 28 - 31, se muestra como queda la unión de cada dimensión con su tabla de hechos.



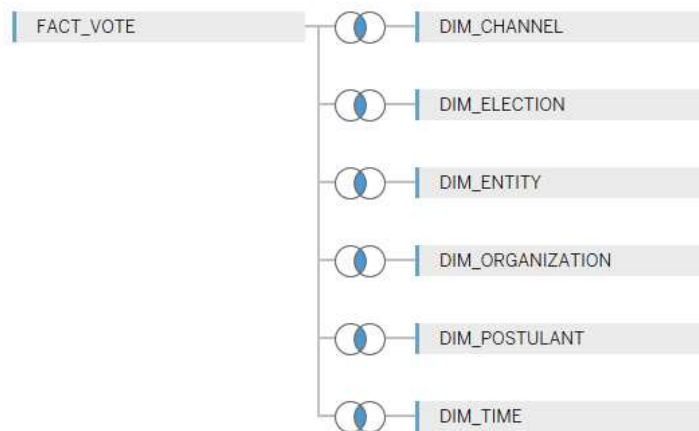
**Ilustración 28: Esquema estrella - Actas**



**Ilustración 29: Esquema estrella – Incidencias**



**Ilustración 30: Esquema estrella – Encuestas**



**Ilustración 31: Esquema estrella – Votos**

La conexión tiene dos modos, “en tiempo real” o “extraer”, para efectos de este TEG, se eligió la opción extraer, para poder utilizar posteriormente la aplicación de visualización “Tableau Reader”. Para un ambiente empresarial, puede utilizarse sin ningún problema la opción “en tiempo real”, lo que facilitará la actualización del BI cada vez que ingresen nuevos datos en modelo dimensional.

Una vez que ya se tiene establecida la conexión a base de datos y todas las fuentes de datos añadidas, se pasa a la pestaña “hoja de trabajo” donde comienza el diseño de los gráficos correspondientes a los indicadores levantados. Cada hoja de trabajo corresponde a un gráfico y un conjunto de hojas forman uno o más dashboard.

En cada hoja, se trabajan los filtros, leyendas, cálculos de totales o porcentajes basados en las dimensiones del modelo.

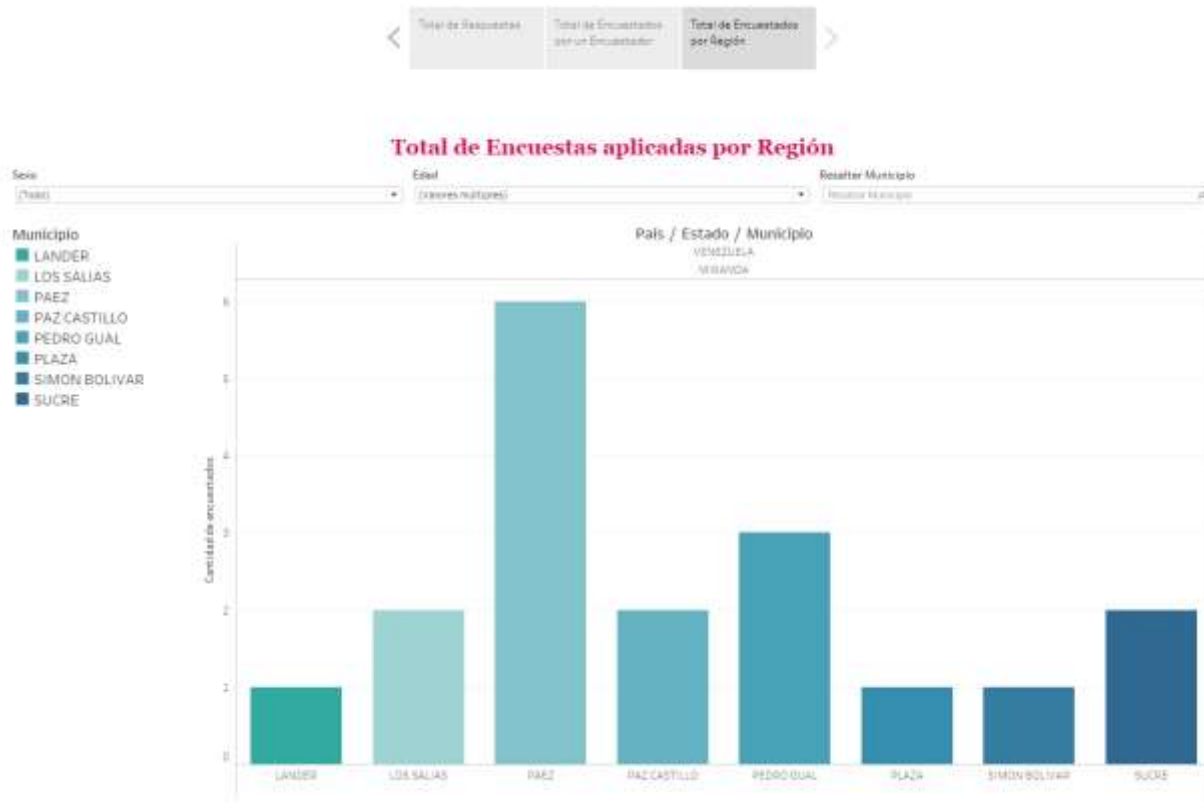
Continuando con el diseño de la aplicación BI, se diseña y crea cada uno de los gráficos que corresponden a los indicadores previamente establecidos. Estos gráficos son recreados en lo que Tableau denomina “hoja de trabajo”. Una vez que se tienen las hojas de trabajo completas para un proceso de negocio determinado, se unen todas en un dashboard. Por lo que, finalmente se tienen cuatro (4) dashboard disponibles en esta aplicación. A continuación se muestran y describen cada uno de los dashboard.

## Proceso de Negocio: Encuesta



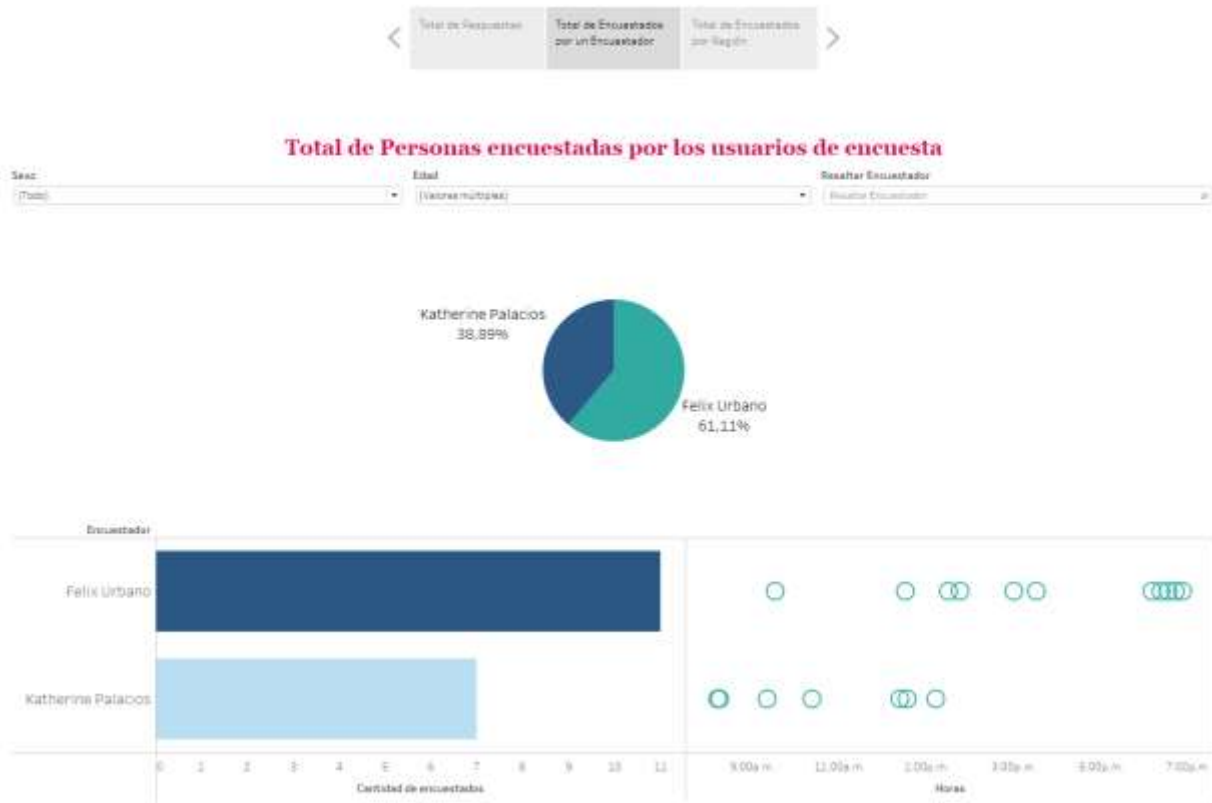
**Ilustración 32: Dashboard de Encuestas – Total de Respuestas**

En la ilustración 32, se puede visualizar el dashboard que corresponde al indicador número uno (1) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el total de respuestas por opción de cada pregunta. De esta manera, se trabaja con dos tablas dinámicas donde se pueden contraer o expandir las opciones. Adicionalmente, se presentan dos modalidades para la tabla, en forma de barras y numérico, con el fin de darle opciones de análisis al usuario. Así mismo, se contemplan tres (3) filtros: sexo, edad y resaltar opción (de cada pregunta).



**Ilustración 33: Dashboard de Encuestas – Total de Encuestados por Región**

En la ilustración 33, se puede visualizar el dashboard que corresponde al indicador número dos (2) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el total de encuestas por región. De esta manera, se trabaja con un gráfico de barras donde se pueden contraer o expandir la jerarquía de las regiones. Adicionalmente, se contemplan tres (3) filtros: sexo, edad y resaltar municipio.



**Ilustración 34: Dashboard de Encuestas – Total de Encuestados por un Encuestador**

En la ilustración 34, se puede visualizar el dashboard que corresponde al indicador número tres (3) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el porcentaje de encuestas aplicadas por un encuestador. De esta manera, se tiene un gráfico de torta en la parte superior, donde se puede visualizar el porcentaje por encuestador.

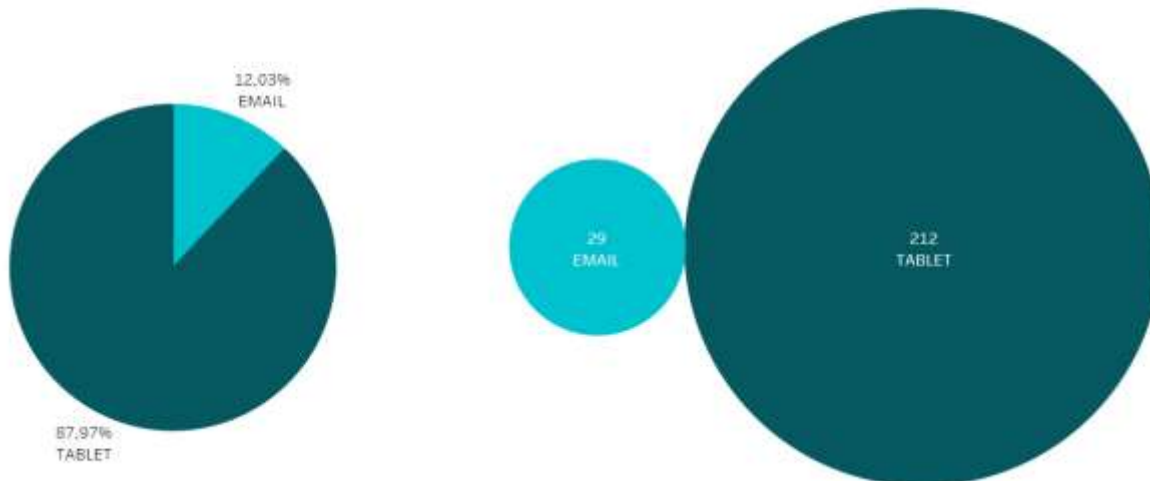
Así mismo, en la parte inferior, se puede visualizar un gráfico de barras combinado con uno de círculos que corresponde al indicador número cuatro (4) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el rendimiento de un usuario encuestador, por lo que se puede visualizar la cantidad de personas encuestadas y las horas en las que se encuestaron más personas. Adicionalmente, para ambos gráficos, se contemplan tres (3) filtros: sexo, edad y resaltar encuestador.



## Proceso de Negocio: Totalización de Votos

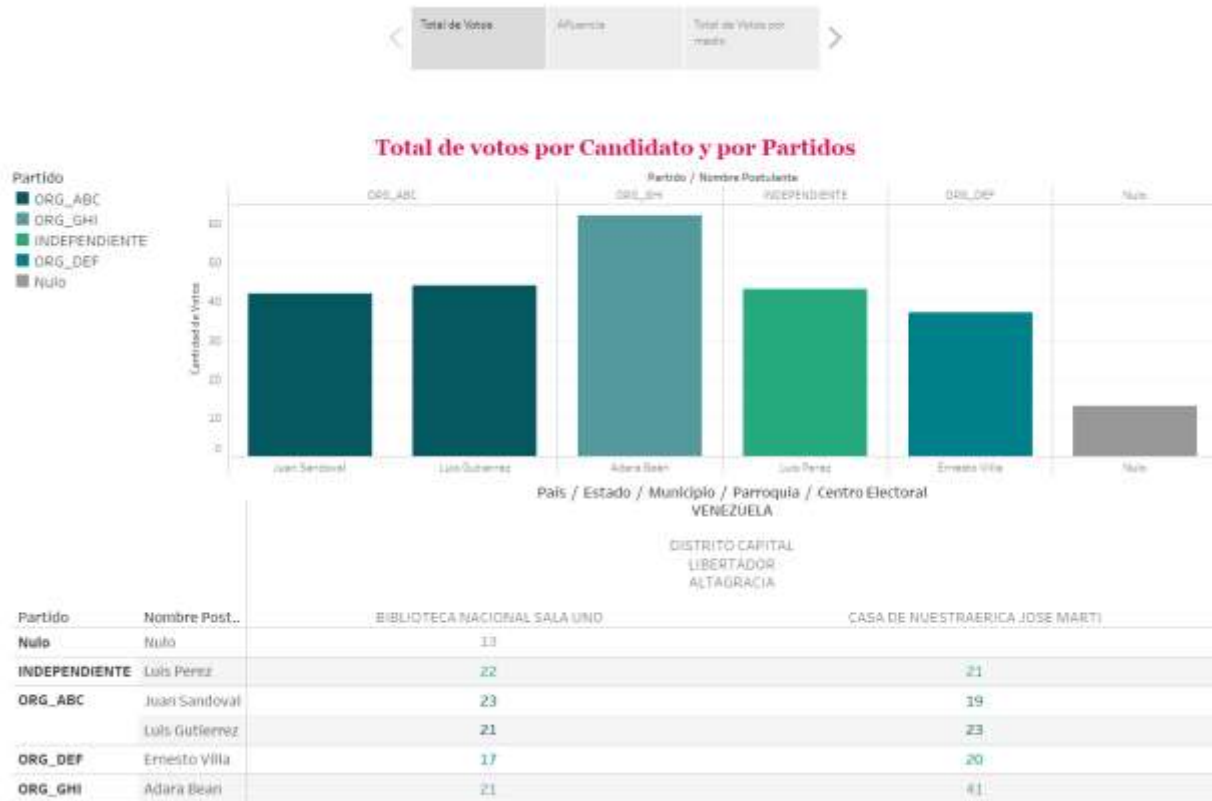


### Cantidad de Votos por Medio de votación



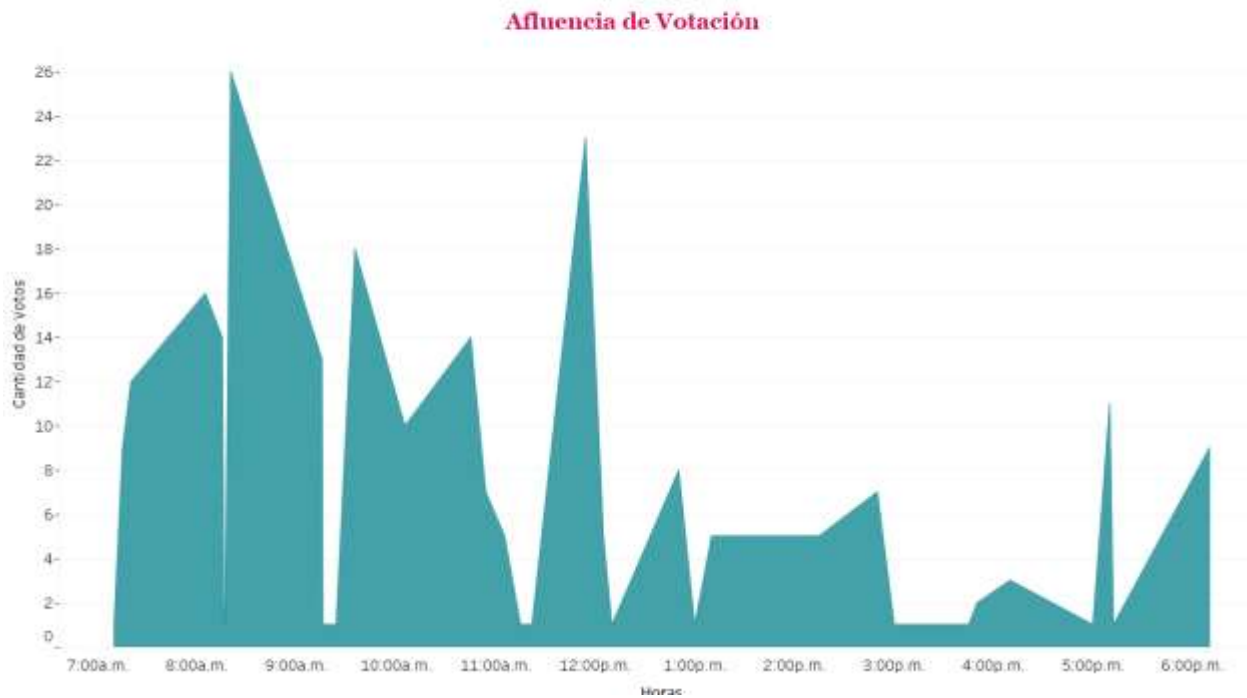
**Ilustración 35: Dashboard de Votos – Total de Votos por Medio**

En la ilustración 35, se puede visualizar el dashboard que corresponde al indicador número cinco (5) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el porcentaje de votos por medio de votación. De esta manera, se trabaja con un gráfico de torta a la izquierda y un gráfico de círculos a la derecha, donde se puede observar tanto el porcentaje como la cantidad de votos por medio respectivamente.



**Ilustración 36: Dashboard de Votos - Total de Votos**

En la ilustración 36, se puede visualizar el dashboard que corresponde al indicador número seis (6) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el total de votos por partido y por postulante. De esta manera, se trabaja con dos tablas dinámicas, una que muestra los resultados en barras (parte superior) y otra que muestra los resultados en números (parte inferior), donde se puede expandir y contraer la jerarquía Partido - Candidato y la jerarquía de región.



**Ilustración 37: Dashboard de Votos – Afluencia**

En la ilustración 37, se puede visualizar el dashboard que corresponde al indicador número siete (7) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar la afluencia de votación. De esta manera, se trabaja con un gráfico de área, el cual refleja la cantidad de personas que fueron a ejercer su derecho al voto en un periodo de tiempo determinado.

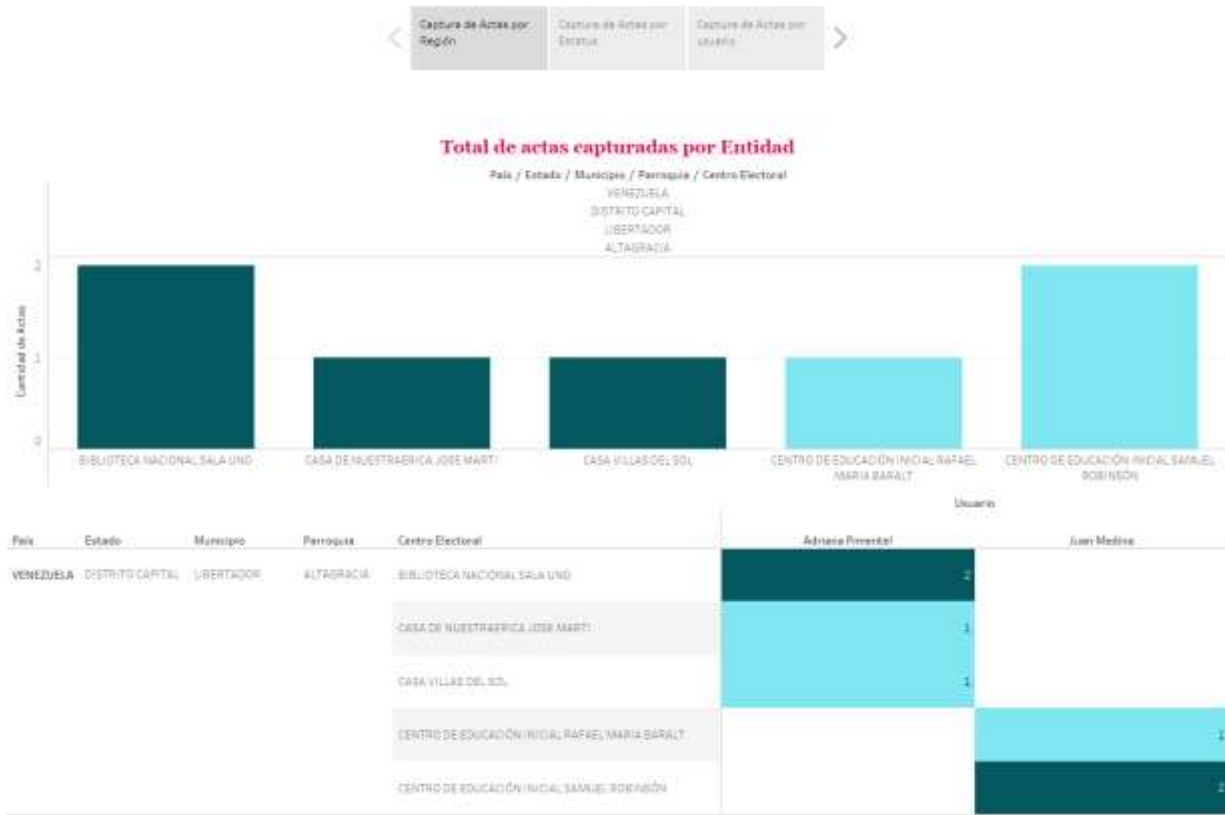
## Proceso de Negocio: Incidencias en el sistema



Ilustración 38: Dashboard de Incidencias

En la ilustración 38, se puede visualizar el dashboard que corresponde al indicador número ocho (8) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar la cantidad de incidencias. De esta manera, se trabaja con una tabla que muestra unas barras del lado izquierdo donde refleja la cantidad de incidencias que se le presentó a cada usuario y unos círculos al lado derecho, donde indica en qué periodo de tiempo se le presentó cada incidencia a cada usuario. Adicionalmente, se presentan en la parte superior una leyenda de colores para cada perfil de usuario y tres (3) filtros: perfil, tipo de incidencia y año.

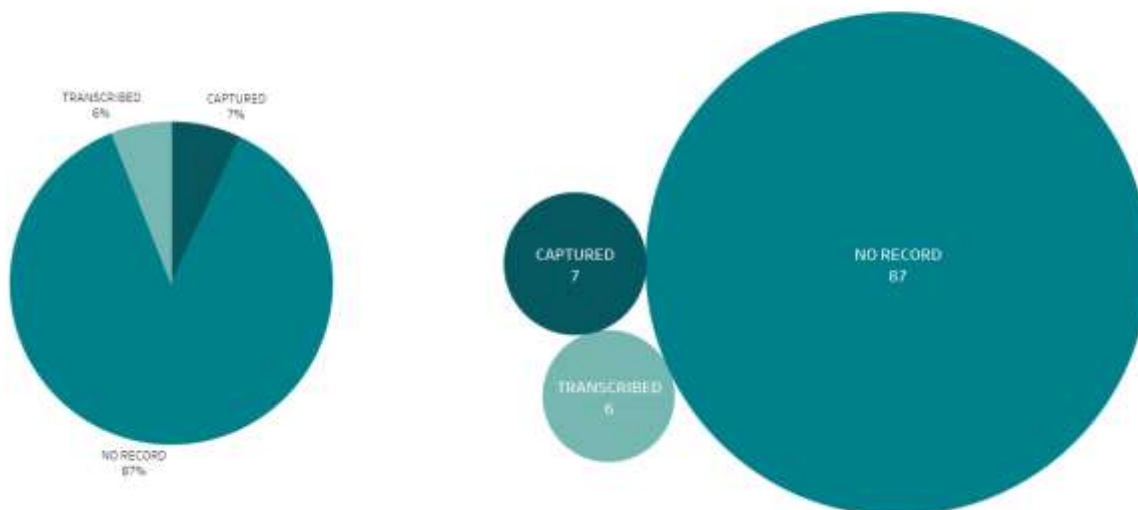
## Proceso de Negocio: Escaneo de Actas



**Ilustración 39: Dashboard de Actas – Captura de Actas por Región**

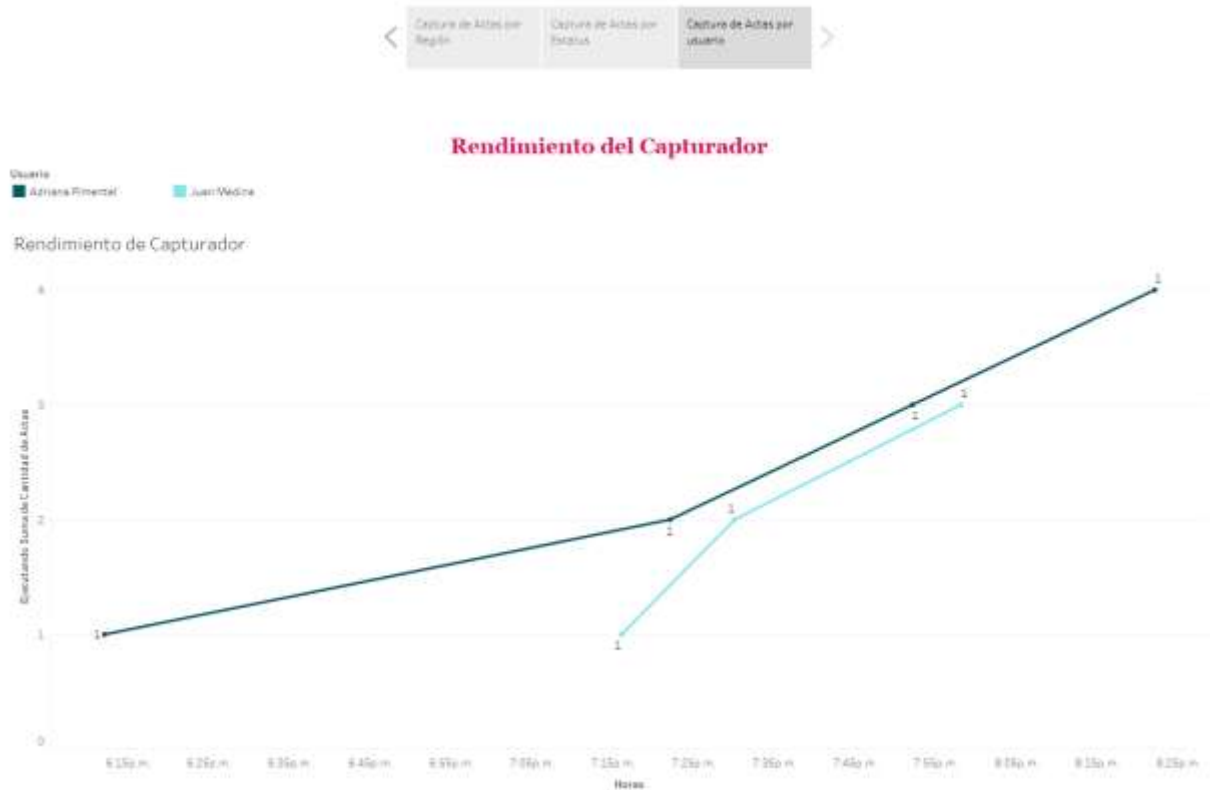
En la ilustración 39, se puede visualizar el dashboard que corresponde al indicador número nueve (9) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el total de actas capturadas por región. De esta manera, se trabaja con dos tablas dinámicas, en la parte superior se visualiza la cantidad de actas por región y en la parte inferior se incorpora la cantidad de actas capturadas por cada usuario.

**Total de Actas por estatus**



**Ilustración 40: Dashboard de Actas – Captura de Actas por Estatus**

En la ilustración 40, se puede visualizar el dashboard que corresponde al indicador número diez (10) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el porcentaje de actas según su estatus. De esta manera, se trabaja con un gráfico de torta a la izquierda y un gráfico de círculos a la derecha, donde se puede observar tanto el porcentaje como la cantidad de actas capturadas por estatus respectivamente.



**Ilustración 41: Dashboard de Actas – Captura de Actas por Usuario**

En la ilustración 41, se puede visualizar el dashboard que corresponde al indicador número once (11) definido en la sección 4.1.2 de este documento. El indicador consiste en mostrar el rendimiento de un usuario capturador. De esta manera, se trabaja con un gráfico de línea que refleja la cantidad de actas que ha capturado un usuario a lo largo del tiempo, dejando visible cuál de ellos destaca por mejor rendimiento. En la parte superior del dashboard se muestra una leyenda de los nombres de usuarios.

### **4.3. Fase III: Finalización**

Una vez finalizado el desarrollo e implementación de la solución de Inteligencia de Negocio, se procede a realizar las pruebas de calidad de datos y las pruebas de aceptación que permitan validar el correcto funcionamiento de la solución y así poder realizar la entrega al usuario final, cumpliendo de esta manera con el objetivo número nueve (9). Asimismo, se realiza una inducción al personal que recibe esta funcionalidad para ser integrada con el sistema.

Para empezar, se realiza una prueba de la calidad de los datos. A través de sentencias SQL se realiza la captura de los datos que se encuentran almacenados en las tablas del modelo transaccional y los mismos son comparados con el resultado que se muestra en cada uno de los dashboard de la solución. Así, se puede mencionar como ejemplo de las mismas, la cantidad de respuestas para cada una de las preguntas configuradas, el total de votos para un candidato, entre otros.

Posteriormente, se aplica una prueba de aceptación a un grupo de siete (7) personas que pertenecen al área de computación y a un grupo de tres (3) personas que no pertenecen al área pero que se les explica el proceso, con el fin de poder certificar que los reportes cumplen con los objetivos del proceso, es un mecanismo de toma de decisiones y satisface a los usuarios finales.

Para hacer estas pruebas se crea un instrumento de evaluación y se define un período de una semana. En cada una de estas encuestas se presentan una serie de preguntas que cuentan con las siguientes opciones de respuestas, cada una con una ponderación del uno (1) al cinco (5):

- Completamente en desacuerdo (1).
- En desacuerdo (2).
- Ni de acuerdo, ni en desacuerdo (3).
- De acuerdo (4).
- Totalmente de acuerdo (5).



Dicho instrumento puede ser visualizado en el anexo D de este documento. A continuación, se presentan los resultados obtenidos en la aplicación de estas pruebas.

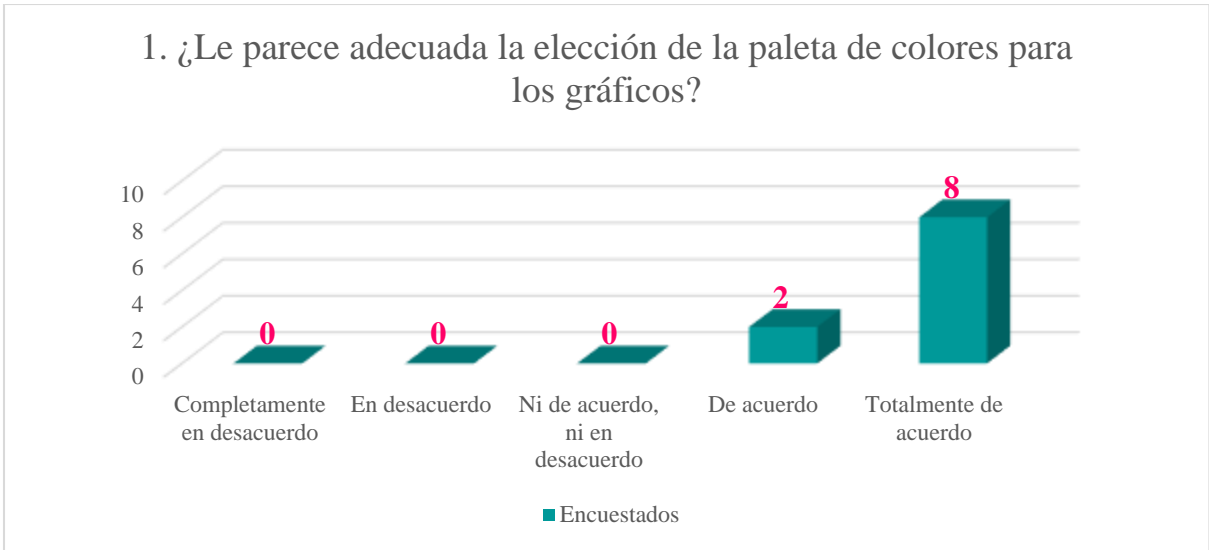
### 4.3.1. Prueba de aceptación

El instrumento de evaluación que se realizó para evaluar el desarrollo del módulo de reportes de este proyecto consiste en una encuesta de seis (6) sencillas preguntas, las cuales pueden ser observadas en la Tabla 6.

**Tabla 6: Preguntas de la Prueba de Aceptación**

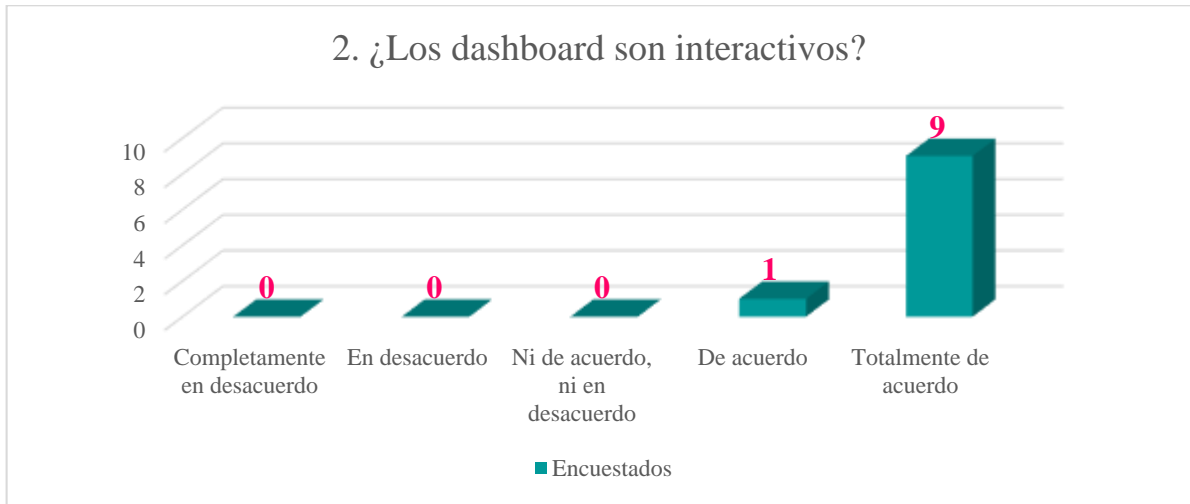
Número	Pregunta
1	¿Le parece adecuada la elección de la paleta de colores para los gráficos?
2	¿Los dashboard son interactivos?
3	¿Considera que los dashboard se encuentran equilibrados?
4	¿Le parece intuitivo la navegación entre los dashboard?
5	¿Es suficiente la información que se presenta en los gráficos para identificar las tendencias?
6	¿Los textos usados en el contenido son lo suficientemente descriptivos?
7	¿Le parece adecuada el orden de visualización de la información?
8	A simple vista, ¿le queda claro cuál es el objetivo de cada reporte?
9	¿Se contempla en los reportes los indicadores descritos en la matriz?
10	¿Cree usted que este módulo le ayudará a tomar mejores decisiones en los procesos de votación?
11	¿Es fácil de usar?
12	¿Recomendaría la adaptación de este módulo para softwares similares?

Tal como se menciona anteriormente, dicha prueba fue aplicada a una población de diez (10) personas, las cuales un grupo siete (7) de ellas pertenecen al área de computación y tres (3) no pertenecen al área. A continuación, se procede a mostrar los resultados para cada una de las mismas y una breve conclusión.



**Ilustración 42: Prueba de Aceptación – Pregunta 1**

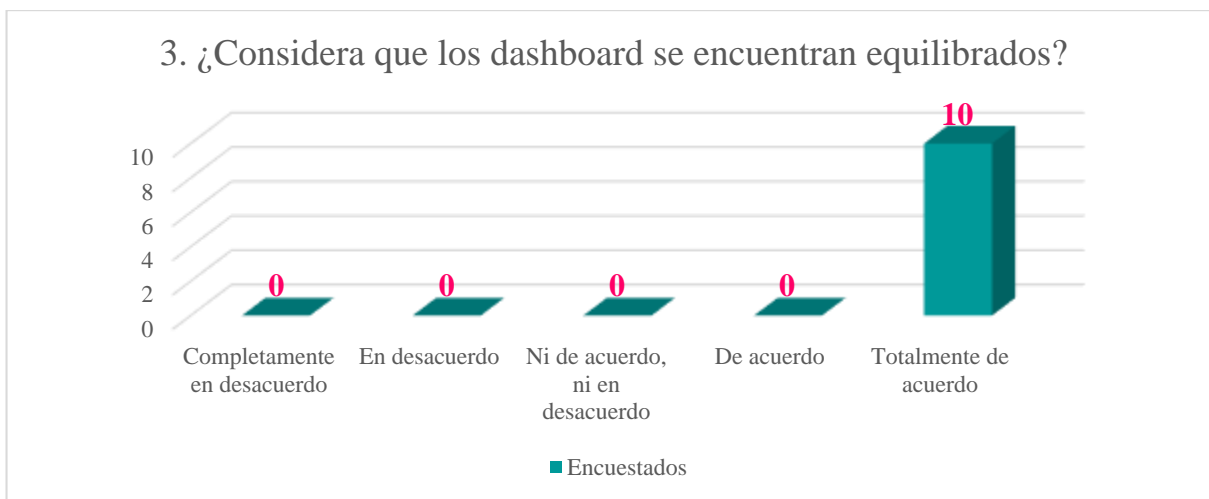
En la Ilustración 42, se observa que un 80% de las personas encuestadas estuvieron conformes con la paleta de colores que se eligió para el módulo de reportes, al responder “Totalmente de acuerdo” y 20 % de los mismos, estuvo “De acuerdo”. La paleta de colores se basó en los colores de la organización: Magenta, Turquesa, Aguamarina y Cian. Sólo para uno de los gráficos se decidió dejar una paleta variada para facilitar el entendimiento de la información que se presenta en él (Reportes de Incidencias).



**Ilustración 43: Prueba de Aceptación – Pregunta 2**

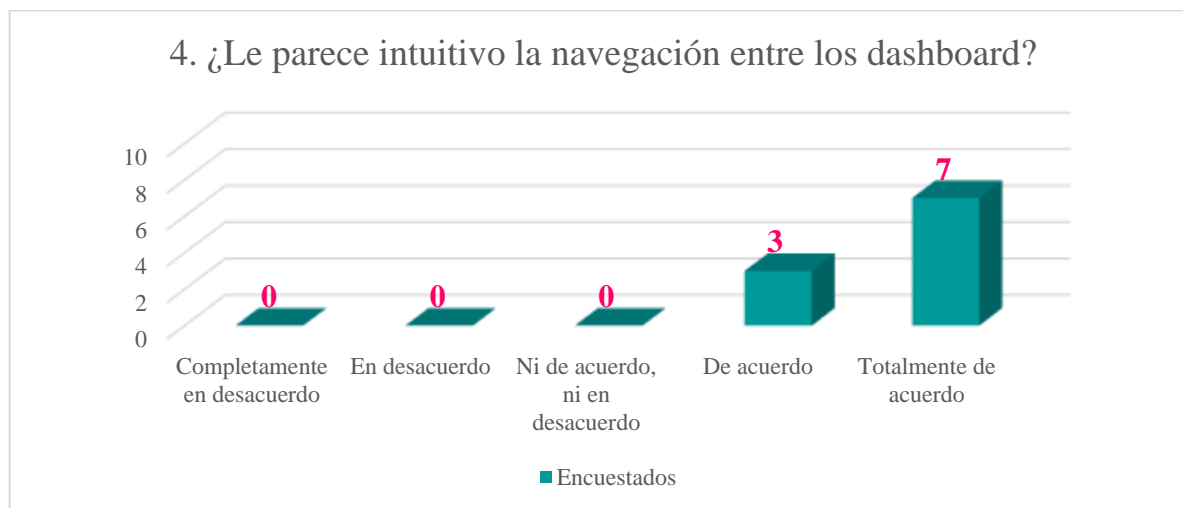
En la Ilustración 43, se observa que un 90% de las personas encuestadas consideran que los reportes son interactivos, ya que pueden seleccionar y resaltar un grupo de datos, así como también

uno de los títulos de los gráficos es capaz de adaptarse a lo que se escoja en el filtro. Por tal motivo se considera que no hace falta realizar cambios en esta funcionalidad.



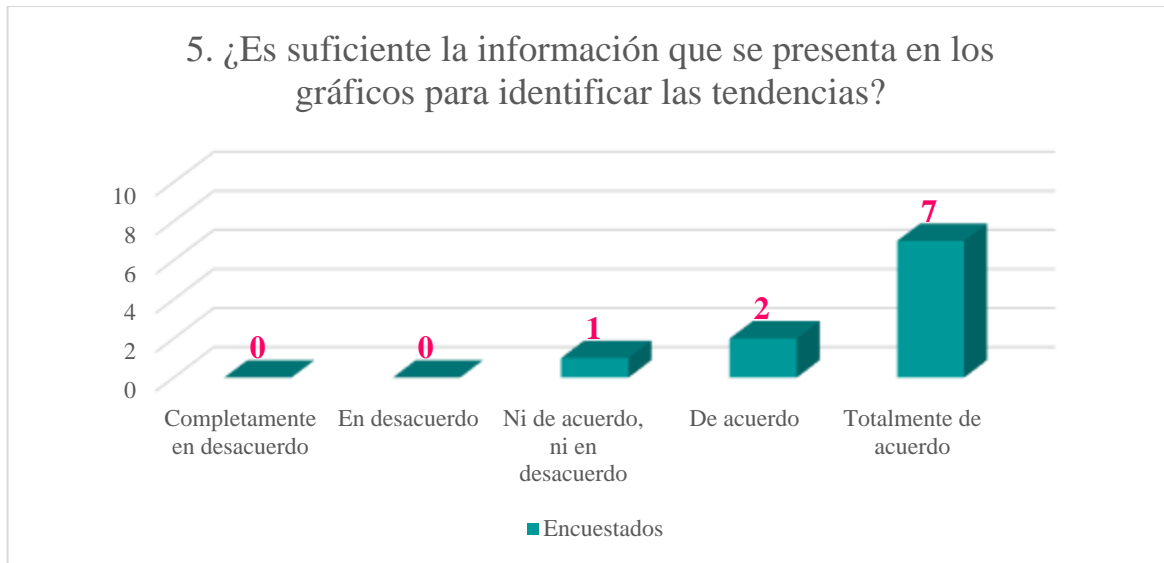
**Ilustración 44: Prueba de Aceptación – Pregunta 3**

En la Ilustración 44, se observa que el 100% de las personas encuestadas consideran que el módulo de reportes está completamente ordenado y armonioso a la vista de cualquier usuario, al responder “Totalmente de acuerdo” por lo que se considera que no debe haber cambios en esta funcionalidad.



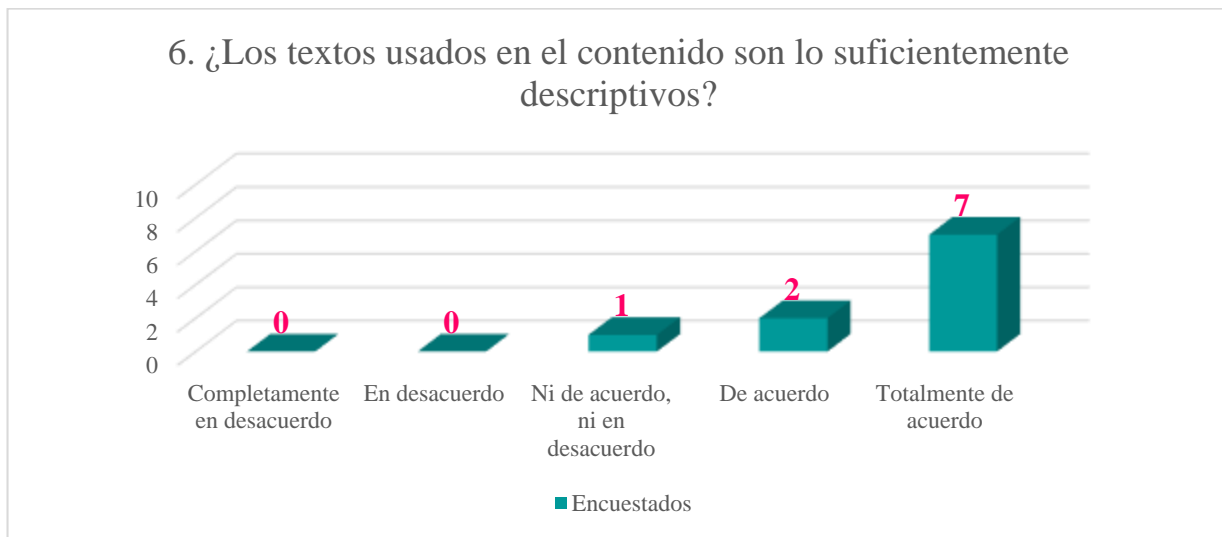
**Ilustración 45: Prueba de Aceptación – Pregunta 4**

En la Ilustración 45, se observa que un 70% de las personas encuestadas consideran que el módulo de reportes es completamente intuitivo para cualquier usuario, al responder “Totalmente de acuerdo”. Así mismo, al 30% de las personas encuestadas se encuentran “De acuerdo” con la navegación, sugiriendo que pueden ser resaltados con otros colores la navegación.



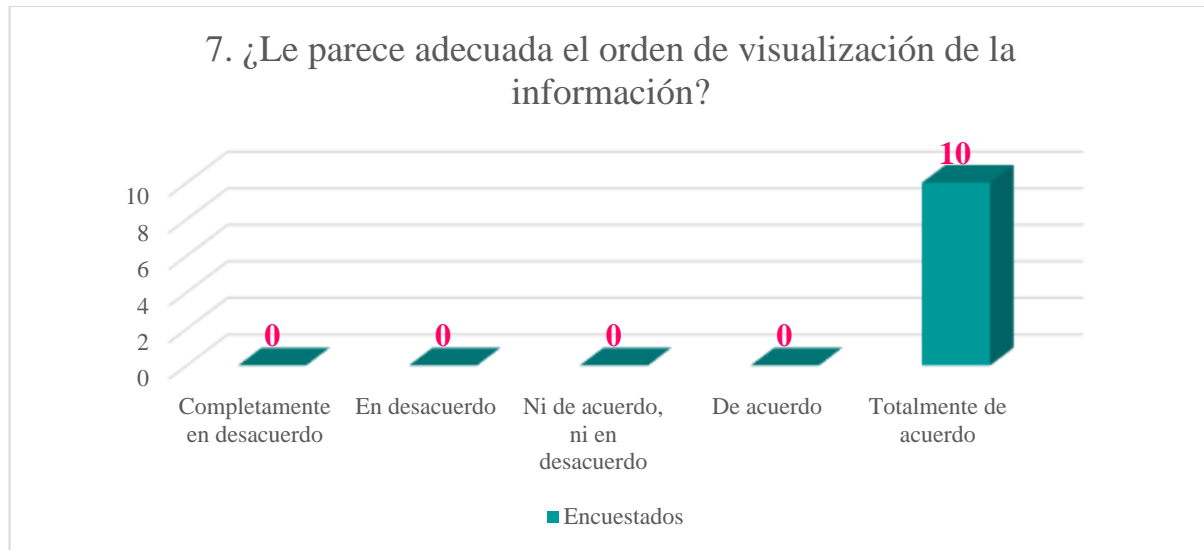
**Ilustración 46: Prueba de Aceptación – Pregunta 5**

En la Ilustración 46, se observa que un 70% de las personas encuestadas consideran que el módulo de reportes es completamente claro para identificar las tendencias, al responder “Totalmente de acuerdo”. Así mismo, al 20% de las personas encuestadas se encuentran “De acuerdo” y sólo un 10% sugiere mayor información en algunos gráficos.



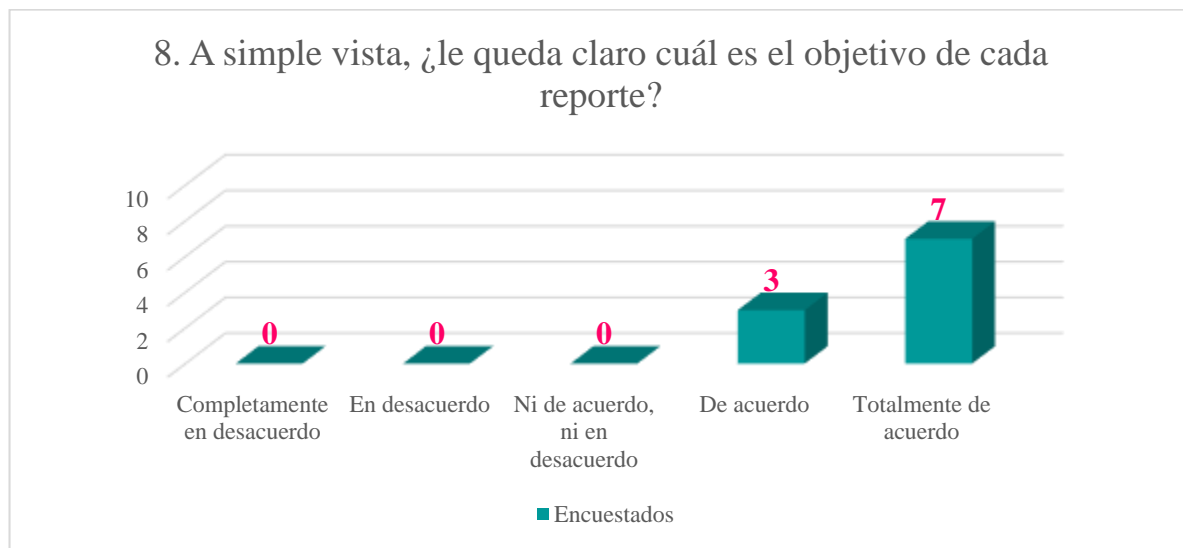
**Ilustración 47: Prueba de Aceptación – Pregunta 6**

En la Ilustración 47, se observa que un 70% de las personas encuestadas consideran que los textos usados en el contenido son completamente descriptivos, al responder “Totalmente de acuerdo”. Así mismo, al 20% de las personas encuestadas se encuentran “De acuerdo” y sólo un 10% sugiere mayor texto en algunos gráficos.



**Ilustración 48: Prueba de Aceptación – Pregunta 7**

En la Ilustración 48, se observa que al 100% de las personas encuestadas consideran que la información presentada en cada uno de los dashboard es completamente ordenada, al responder “Totalmente de acuerdo”. Por lo que se considera que no debe haber cambios en esta funcionalidad.



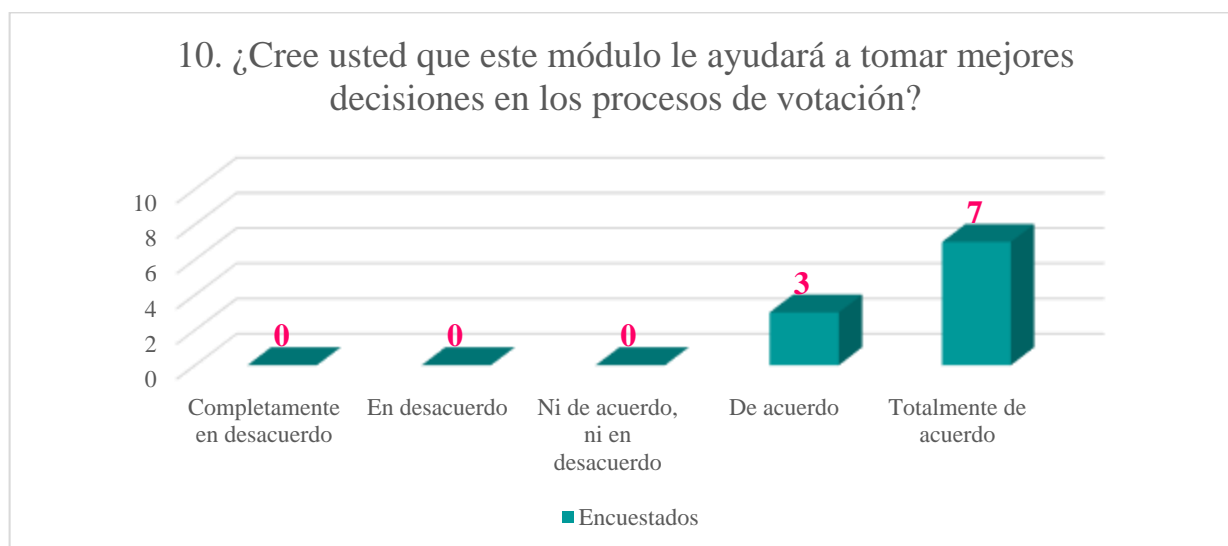
**Ilustración 49: Prueba de Aceptación – Pregunta 8**

En la Ilustración 49, se observa que un 70% de las personas encuestadas les quedó completamente claro el objetivo de cada reporte, al responder “Totalmente de acuerdo”. Por otro lado, al 30% hizo falta ampliarles un poco más el contexto del proyecto.



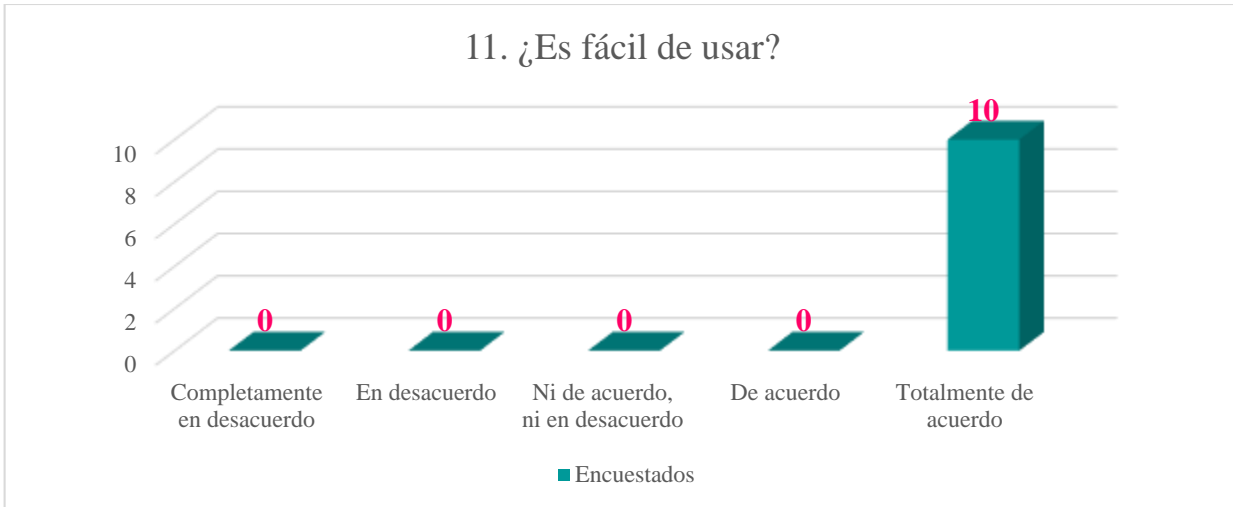
**Ilustración 50: Prueba de Aceptación – Pregunta 9**

En la Ilustración 50, se observa que al 100% de las personas encuestadas considera que los reportes presentados se corresponden exactamente con los indicadores descritos previamente. Por tal motivo, todos respondieron estar “Totalmente de acuerdo”.



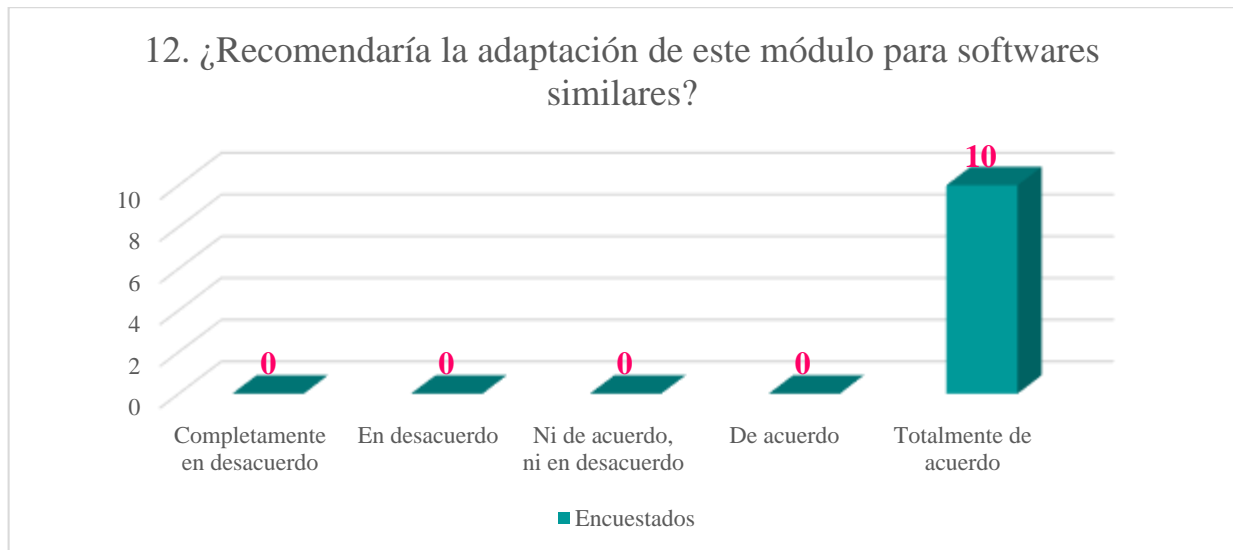
**Ilustración 51: Prueba de Aceptación – Pregunta 10**

En la Ilustración 51, se observa que un 70% de las personas encuestadas considera que los reportes presentados servirán de mucha ayuda en la toma de decisiones durante un proceso de votación. Adicionalmente, un 30% opina lo mismo, pero con la variante de sugerir nuevos reportes para ampliar esa toma de decisiones.



**Ilustración 52: Prueba de Aceptación – Pregunta 11**

En la Ilustración 52, se observa que el 100% de las personas encuestadas considera que el módulo es totalmente sencillo de usar y amigable. Por lo que no hace falta hacer cambios en este aspecto.



**Ilustración 53: Prueba de Aceptación – Pregunta 12**

En la Ilustración 53, se observa que el 100% de las personas encuestadas recomendarían este módulo para softwares similares. Por lo que se puede deducir que el TEG ha cumplido con los objetivos propuestos.

# Resultados Obtenidos

Una vez que se ha realizado el desarrollo e implementación de la solución BI y se ha aplicado tanto la prueba de calidad de los datos como la prueba de aceptación se obtienen los siguientes resultados.

En primer lugar, se observa que los datos pasaron por el proceso de extracción, transformación y carga de manera exitosa, quedando los mismos distribuidos de una manera más adecuada en el modelo dimensional, con el objetivo de procesar los reportes deseados con mayor facilidad. Así mismo, se observa que los cálculos de los reportes corresponden satisfactoriamente con lo que se tenía en el modelo transaccional.

También, se resuelve el problema de tener que desarrollar manualmente cada uno de los reportes, teniendo que realizar consultas complejas a la base de datos transaccional y en ocasiones teniendo que llevar la lógica de programación a un nivel superior. En lugar de esto, ahora la herramienta es la que se encarga de realizar las consultas, permitiendo que el desarrollador pueda armar fácilmente un reporte con la opción de Tableau “arrastrar y soltar” las dimensiones, mejorando así el rendimiento.

Por otro lado, la información es más sencilla de procesar por parte de los usuarios finales, lo que permite dedicar mayor tiempo a producir conocimiento en lugar de esperar largos periodos por la aplicación para la visualización de un reporte. De esta manera se logra mejorar la toma de decisiones.

Por último, esta solución también permite que los reportes sean modulares y adaptados en la aplicación en las combinaciones deseadas, de manera que este módulo de TVT puede ser fácilmente ensamblado para cada cliente de acuerdo a sus necesidades.

Concluyendo de esta manera, que los resultados obtenidos aportan grandes beneficios para el funcionamiento y mantenimiento de la aplicación.



# Conclusiones y Recomendaciones

Luego de haber completado el presente TEG, el cual consistió en desarrollar una solución de Inteligencia de Negocio que permitiera generar indicadores asociados al proceso de votación, para mejorar la toma de decisiones y cumplir con las expectativas de las organizaciones electorales, se puede concluir lo siguiente.

En primer lugar, se realizó la planificación del proyecto, donde se eligió trabajar bajo la metodología Scrum. De esta manera, los objetivos planteados se pudieron cumplir de manera ágil, incremental, iterativa y estructurada. Así mismo, se realizó una adaptación con el fin de poder incorporar en una de sus iteraciones el ciclo de vida de Raph Kimball, el cual sirvió como guía para la construcción de la solución BI.

Luego, se realizó el análisis del contenido de la Base de Datos Relacional que se encontraba implementada en Oracle Enterprise Edition 12C. Dicha BD era la fuente de datos destinada para la solución de Inteligencia de Negocio. Adicionalmente, se analizó todo el documento de requerimientos funcionales del proyecto y se acordaron reuniones con el cliente. Como resultado, fueron definidos los indicadores de gestión que servirían de base para el diseño del Data Warehouse.

Por otro lado, se hizo un cuadro comparativo con las características de los productos del mercado relacionados a este TEG, con el fin de poder determinar cuáles serían los más adecuados para el desarrollo de la solución propuesta. De esta manera, se eligió trabajar con OracleDB 12C para el modelo dimensional, Pentaho Data Integration para el proceso ETL y Tableau para la construcción y visualización del BI.

Además, se realizó el diseño, desarrollo e implementación del modelo dimensional en Oracle Enterprise Edition 12C. Dicho modelo consta de cuatro (4) procesos de negocio que fueron definidos y analizados de la misma manera con el cliente durante las reuniones definidas por la metodología Scrum.

Con respecto a la data, para resguardar la confidencialidad de los clientes, la BD transaccional fue actualizada con data simulada. La data fue extraída de diferentes fuentes de Internet, como por ejemplo el CNE para todo lo que corresponde a los centros electorales ubicados en Venezuela.

Otros datos como los nombres de usuarios fueron generados al azar, de manera tal que no se encuentra ningún dato personal embebido en el desarrollo.

También, partiendo de los modelos de Base de datos, se hizo el diseño y desarrollo del ETL a través de la herramienta Pentaho Data Integration. Cada dimensión posee su propia transformación, al igual que las tablas de hecho y finalmente un “Job” que los ejecuta todos en secuencia.

Seguidamente, se generó un dashboard interactivo para mostrar los indicadores previamente definidos, con las funcionalidades necesarias y gráficos llamativos e intuitivos para permitir mejorar la toma de decisiones. El diseño y visualización de los reportes fue realizado con Tableau desktop.

Para terminar, se realizaron las pruebas de aceptación las cuales permitieron obtener opiniones de las personas asociadas al área de computación y las que no lo son, de tal manera de conseguir una visión bastante equilibrada con respecto a la usabilidad, diseño, armonía y el apoyo que este módulo brinda al sistema automatizado de votación elegido como caso de estudio.

Finalmente, se puede concluir que se abarcaron de manera satisfactoria todos los objetivos específicos de este TEG y en consecuencia, se cumplió con el objetivo general generando grandes beneficios para la aplicación, siendo éste desarrollo totalmente escalable, así como también el análisis de resultados de manera rápida, efectiva y objetiva.

## **Recomendaciones**

Dada la experiencia obtenida a lo largo de este periodo de desarrollo, se proponen las siguientes recomendaciones para trabajos futuros y para mejoras del sistema automatizado de votación:

- Separar por módulos y esquemas el modelo de datos transaccional para un mejor mantenimiento y escalabilidad del sistema.
- Añadir otros procesos de negocio que sirvan de apoyo a los dueños del negocio a identificar las cuentas por cobrar y la gestión del proyecto.
- Si se decide integrar esta solución al proyecto, se deben considerar medidas de seguridad en cuanto a las actividades realizadas por cada usuario en cada uno de los módulos.
- Plantear otros procesos de negocio que apoyen las necesidades de los entes electorales.

# Referencias

- [1] Cursosinea, «cursosinea,» [En línea]. Available: [http://www.cursosinea.conevyt.org.mx/cursos/mexico/contenidos/recursos/revista/4\\_1.htm](http://www.cursosinea.conevyt.org.mx/cursos/mexico/contenidos/recursos/revista/4_1.htm). [Último acceso: 15 Marzo 2019].
- [2] C. Chen, "Sistema de información," 21 05 2019. [Online]. Available: <https://www.significados.com/sistema-de-informacion/>. [Accessed 13 07 2019].
- [3] M. Rouse, «Tech Target,» Enero 2015. [En línea]. Available: <https://searchdatacenter.techtarget.com/es/definicion/Definicion-de-OLAP-procesamiento-analitico-en-linea>. [Último acceso: 23 Diciembre 2018].
- [4] E. Guevara, "DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO QUE APOYE A LA TOMA DE DECISIONES EN EL SERVICIO DE ORIENTACIÓN Y CONTROL DE ESTUDIOS DE LA FACULTAD DE CIENCIAS," Caracas, 2016.
- [5] V. A. T. Zambrano, "DESARROLLO DE UN SISTEMA DE INFORMACIÓN PARA LA ADMINISTRACIÓN Y GESTIÓN DE CONTRATOS DE ARRENDAMIENTO DE CENTROS COMERCIALES," Caracas, 2017.
- [6] Oracle Corporation, «Whitepaper - Que es Inteligencia de Negocio,» 2008. [En línea]. Available: [http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529\\_esa.pdf](http://www.oracle.com/ocom/groups/public/@otn/documents/webcontent/317529_esa.pdf).
- [7] Cano, «<http://itemsweb.esade.edu>,» 2007. [En línea]. Available: [http://itemsweb.esade.edu/biblioteca/archivo/Business\\_Intelligence\\_competir\\_con\\_informacion.pdf](http://itemsweb.esade.edu/biblioteca/archivo/Business_Intelligence_competir_con_informacion.pdf). [Último acceso: 10 10 2018].
- [8] B. López, «Clase 2: Arquitectura general de Soluciones BI,» Caracas, 2016.

- [9] F. García, «La Inteligencia de Negocio y el Datawarehouse,» 2 Marzo 2016. [En línea]. Available:  
<https://community.toadworld.com/platforms/oracle/b/weblog/archive/2016/02/03/la-inteligencia-de-negocios-y-el-datawarehouse>. [Último acceso: 12 Febrero 2018].
- [10] R. Kimball y M. Ross, The Data Warehouse Toolkit, Canada: John Wiley and Sons, Inc., 2002.
- [11] B. López, «Clase 5 – Capa analítica: Datawarehouse/Datamarts,» Caracas, 2016.
- [12] Tableau Software, «[www.tableau.com](http://www.tableau.com),» 2003-2018. [En línea]. Available:  
<https://www.tableau.com/es-es/products/techspecs>. [Último acceso: 8 10 2018].
- [13] QlikTech, «qlik,» 1993-2018. [En línea]. Available: <https://www.qlik.com/es-es/>. [Último acceso: 9 Octubre 2018].
- [14] Mine-class, «5 cosas que me gustan de Qlik Sense,» 25 Agosto 2017. [En línea]. Available:  
<https://mine-class.com/qliksense/>. [Último acceso: 10 Octubre 2018].
- [15] Microsoft, «powerbi,» 2018. [En línea]. Available: <https://powerbi.microsoft.com/es-es/desktop/>. [Último acceso: 11 Octubre 2018].
- [16] MakeSoft Technologies, «[makesoft](http://www.makesoft.com),» [En línea]. Available:  
<https://www.makesoft.es/es/productos/power-bi/>. [Último acceso: 11 Octubre 2018].
- [17] B. López, «Clase 3 - Indicadores de Gestión,» Caracas, 2016.
- [18] J. Camejo, «Indicadores de gestión ¿Qué son y por qué usarlos?,» 28 November 2012. [En línea]. Available: <https://www.gestiopolis.com/indicadores-de-gestion-que-son-y-por-que-usarlos/>.
- [19] SCRUMstudy™, A Guide to the Scrum Body Of Knowledge (SBOK™Guide) – 3rd Edition, 3rd ed., Arizona: SCRUMstudy™, una marca de VMEdU, Inc., 2017, p. 429.

- [20] L. Gonçalves, «QUÉ ES LA METODOLOGÍA SCRUM, TODO LO QUE NECESITAS SABER,» 25 Enero 2019. [En línea]. Available: <https://luis-goncalves.com/es/que-es-la-metodologia-scrum/>. [Último acceso: 10 Marzo 2019].
- [21] «Qué es SCRUM y los roles en SCRUM,» 2017. [En línea]. Available: <https://platzi.com/blog/que-es-scrum-y-los-roles-en-scrum/>. [Último acceso: 12 04 2019].
- [22] G. Rivadera, «La metodología de Kimball para el diseño de almacenes de datos (Data warehouses),» Instituto Tecnológico de Buenos Aires (ITBA), Argentina., Noviembre 2010. [En línea]. Available: <http://www1.ucasal.edu.ar/htm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>. [Último acceso: 12 Febrero 2018].
- [23] P. Gutiérrez, «Metodología de uso de Herramientas de Inteligencia de Negocios como estrategia para aumentar la productividad y competitividad de una PyME,» México, D.F., 2012.
- [24] I. Julio, «¿Que es Oracle?,» 02 Octubre 2015. [En línea]. Available: <https://grupo4herramientasinformatica.blogspot.com/2015/10/que-es-la-oracle.html>. [Último acceso: 12 04 2019].
- [25] Arimetrics, «Qué es Tableau,» [En línea]. Available: <https://www.arimetrics.com/glosario-digital/tableau>. [Último acceso: 12 04 2019].
- [26] M. Ospina, «Diseño de Bases de Datos,» Venezuela, 2009.
- [27] M. Ospina, «Fundamentos y Conceptos Básicos de Bases de Datos,» Venezuela, 2009.

# Anexos

## A. Diccionario de Datos del Data Warehouse

**Tabla 7: Tabla Dimensión Pregunta**

<b>Nombre de la tabla:</b> DIM_QUESTION			
<b>Descripción:</b> Almacena la información de las preguntas. Tanto las preguntas de las encuestas como las preguntas de la elección.			
Campo	Tamaño	Tipo de Dato	Descripción
SK_QUESTION	38	NUMBER	Clave subrogada de la dimensión
QST_ID	38	NUMBER	Clave primaria del modelo transaccional
QST_NAME	100	VARCHAR2	Nombre de la pregunta
QST_UPPER_LIMIT	38	NUMBER	Límite superior de opciones que puede seleccionar una persona para esta pregunta
QST_LOWER_LIMIT	38	NUMBER	Límite inferior de opciones que puede seleccionar una persona para esta pregunta
QST_TYPE_QUESTION	8	VARCHAR2	Indica el tipo de pregunta (Cerrada - Abierta)
QST_REQUIRED	1	CHAR	Indica si es una pregunta obligatoria de responder

**Tabla 8: Tabla Dimensión Encuestado**

<b>Nombre de la tabla:</b> DIM_SURVEYED			
<b>Descripción:</b> Almacena la información del encuestado. El sexo y la edad.			
Campo	Tamaño	Tipo de Dato	Descripción
SK_SURVEYED	38	NUMBER	Clave subrogada de la dimensión
SRV_DATA_ID	38	NUMBER	Clave primaria del modelo transaccional
SRV_DATA_SEX	1	CHAR	Sexo de la persona encuestada
SRV_DATA_AGE	38	NUMBER	Edad de la persona encuestada

**Tabla 9: Tabla Dimensión Opción**

<b>Nombre de la tabla:</b> DIM_OPTION			
<b>Descripción:</b> Almacena la información del encuestado. El sexo y la edad.			
Campo	Tamaño	Tipo de Dato	Descripción
SK_OPTION	38	NUMBER	Clave subrogada de la dimensión
OPT_ID	38	NUMBER	Clave primaria del modelo transaccional
OPT_NAME	20	VARCHAR2	Nombre de la opción

**Tabla 10: Tabla Dimensión Tiempo**

<b>Nombre de la tabla:</b> DIM_TIME			
<b>Descripción:</b> Almacena la información del tiempo. Tiene la jerarquía Año - Trimestre - Mes - Semana - Día			
Campo	Tamaño	Tipo de Dato	Descripción
SK_DATE	38	NUMBER	Clave subrogada de la dimensión
YEAR	38	NUMBER	Año
TRIMESTER_DESC	20	VARCHAR2	Trimestre
MONTH	28	NUMBER	Número del mes
MONTH_DESC	30	VARCHAR2	Mes en letras
MONTH_DESC_SHORT	3	VARCHAR2	Abreviación del mes
WEEK_YEAR	38	NUMBER	Semana del año
WEEK_DESC	30	VARCHAR2	Día de la semana
WEEK_DESC_SHORT	20	VARCHAR2	Abreviación del día de la semana
DAY_WEEK	38	NUMBER	Día de la semana en número
DAY_YEAR	38	NUMBER	Día del año
DAY_MONTH	38	NUMBER	Día del mes
DATE_COMPLETE	-	TIMESTAMP	Fecha completa

**Tabla 11: Tabla Dimensión Encuesta**

<b>Nombre de la tabla:</b> DIM_POLL			
<b>Descripción:</b> Almacena la información de la encuesta			
Campo	Tamaño	Tipo de Dato	Descripción
SK_POLL	38	NUMBER	Clave subrogada de la dimensión
POLL_ID	38	NUMBER	Clave primaria del modelo transaccional
POLL_NAME	100	VARCHAR2	Nombre de la encuesta
POLL_START_DATE	100	VARCHAR2	Fecha de inicio para aplicar la encuesta
POLL_FINSH_DATE	100	VARCHAR2	Fecha fin para aplicar la encuesta

**Tabla 12: Tabla Dimensión Entidad**

<b>Nombre de la tabla:</b> DIM_ENTITY			
<b>Descripción:</b> Almacena la información de las regiones			
Campo	Tamaño	Tipo de Dato	Descripción
SK_ENTITY	38	NUMBER	Clave subrogada de la dimensión
ENT_ID	38	NUMBER	Clave primaria del modelo transaccional
ENT_COUNTRY	100	VARCHAR2	País
ENT_STATE	100	VARCHAR2	Estado
ENT_MUNICIPALITY	100	VARCHAR2	Municipio
ENT_PARISH	100	VARCHAR2	Parroquia
ENT_ELECTORAL_CENTER	100	VARCHAR2	Centro Electoral
ENT_TABLE_NAME	100	VARCHAR2	Mesa
ENT_CATEGORY	100	VARCHAR2	Categoría de la región



**Tabla 13: Tabla Dimensión Usuario**

<b>Nombre de la tabla:</b> DIM_USER			
<b>Descripción:</b> Almacena la información del usuario			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
SK_USER	38	NUMBER	Clave subrogada de la dimensión
USR_ID	38	NUMBER	Clave primaria del modelo transaccional
USR_TYPE_ID	15	VARCHAR2	Tipo de identificación
USR_PHOTO	-	BLOB	Foto tipo carnet del usuario
USR_FIRST_NAME	20	VARCHAR2	Nombre del usuario
USR_LAST_NAME	20	VARCHAR2	Apellido del usuario
USR_SEX	1	CHAR	Sexo del usuario
USR_AGE	3	NUMBER	Edad del usuario cuando se incorporó en el sistema
USR_CODE_PHONE	4	VARCHAR2	Código del país
USR_TELEPHONE	38	NUMBER	Teléfono
USR_ADDRESS	500	VARCHAR2	Dirección completa del usuario
USR_DELETE	1	CHAR	Indica la eliminación lógica del usuario
USR_LANG_ACRONYM_ID	6	VARCHAR2	Acrónimo del lenguaje por defecto que escogió el usuario
USR_LANG_NAME	50	VARCHAR2	Nombre completo del lenguaje por defecto que escogió el usuario

**Tabla 14: Tabla Dimensión Perfil**

<b>Nombre de la tabla:</b> DIM_PROFILE			
<b>Descripción:</b> Almacena la información del perfil del usuario			
Campo	Tamaño	Tipo de Dato	Descripción
SK_PROFILE	38	NUMBER	Clave subrogada de la dimensión
PRF_ID	38	NUMBER	Clave primaria del modelo transaccional
PRF_NAME	20	VARCHAR2	Nombre del perfil
PRF_DESCRIPTION	500	VARCHAR2	Descripción
PRF_STATUS	20	VARCHAR2	Estatus del perfil
PRF_DEFAULT	1	CHAR	Indica si es un perfil que viene cargado en el sistema por defecto
PRF_DELETE	1	CHAR	Indica la eliminación lógica del usuario

**Tabla 15: Tabla Dimensión Incidencia**

<b>Nombre de la tabla:</b> DIM_INCIDENCE			
<b>Descripción:</b> Almacena la información con respecto a las incidencias que se le presentan a un usuario			
Campo	Tamaño	Tipo de Dato	Descripción
SK_INCIDENCE	38	NUMBER	Clave subrogada de la dimensión
INCIDENCE_ID	38	NUMBER	Clave primaria del modelo transaccional
INC_TYPE	100	VARCHAR2	Tipo de Incidencia
INC_DESCRIPTION	1000	VARCHAR2	Descripción
INC_STATUS	50	VARCHAR2	Estatus de la incidencia
INC_START_DATE_ATTENTION	-	TIMESTAMP	Fecha de inicio de atención de la incidencia
INC_FINISH_DATE_ATTENTION	-	TIMESTAMP	Fecha de cierre de la incidencia

**Tabla 16: Tabla Dimensión Elección**

<b>Nombre de la tabla:</b> DIM_ELECTION			
<b>Descripción:</b> Almacena la información con respecto a la elección			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
SK_ELECTION	38	NUMBER	Clave subrogada de la dimensión
ELECT_ID	38	NUMBER	Clave primaria del modelo transaccional
ELECT_NAME	20	VARCHAR2	Nombre de la elección
ELECT_DESCRIPTION	500	VARCHAR2	Descripción de la elección
ELECT_LOGO		BLOB	Imagen de identificación
ELECT_START_DATE		TIMESTAMP	Fecha de Inicio de configuración
ELECT_FINISH_DATE		TIMESTAMP	Fecha de la culminación de la configuración
ELECT_START_EVENT		TIMESTAMP	Fecha de inicio del evento
ELECT_FINISH_EVENT		TIMESTAMP	Fecha de culminación del evento
ELECT_STATUS	20	VARCHAR2	Estatus en el que se encuentra el evento (Configuración, Activa, Escrutinio, Terminada, Inactiva)
ELECT_INSTRUCTIONS	1	CHAR	Indica si se desean desplegar instrucciones para el evento
ELECT_DELETE	1	CHAR	Eliminación Lógica
ELECT_VOTING_CHECK	1	CHAR	Indica si la elección requiere comprobante de voto.
ELECT_VOUCHER_TEMPLATE		CLOB	Almacena la plantilla del comprobante del voto
ELECT_CATEGORY	1	CHAR	Indica si tiene una ubicación específica
ELECT_DISTRIC_CHECK	1	CHAR	Indica si tiene una ubicación específica
ELECT_TYPE_SERVICE	20	VARCHAR2	Indica si es una elección o un conteo rápido
ELECT_VOTING_TYPE	20	VARCHAR2	Voto lista o Uninominal
ELECT_VOTING_TEMPLATE		CLOB	Plantilla del cartón de la elección para la votación

**Tabla 17: Tabla Dimensión Organización**

<b>Nombre de la tabla:</b> DIM_ORGANIZATION			
<b>Descripción:</b> Almacena la información de la organización y sus alianzas si las tiene			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
SK_ORGANIZATION	38	NUMBER	Clave subrogada de la dimensión
ORG_ID	38	NUMBER	Clave primaria del modelo transaccional
ORG_NAME	100	VARCHAR2	Nombre de la organización
ORG_ACRONYM	10	VARCHAR2	Acrónimo de la organización
ORG_DESCRIPTION	500	VARCHAR2	Descripción breve ( misión, visión)
ORG_LOGO	-	BLOB	Imagen que lo identifica
ORG_STATUS	20	VARCHAR2	Indica si se encuentra activo o inhabilitado
ORG_DELETE	1	CHAR	Eliminación lógica
ALL_ID	38	NUMBER	Clave primaria del modelo transaccional
ALL_NAME	100	VARCHAR2	Nombre de la alianza, si existe
ALL_ACRONYM	10	VARCHAR2	Acrónimo de la alianza
ALL_STATUS	20	VARCHAR2	Indica si se encuentra activo o inhabilitado
ALL_LOGO	-	BLOB	Imagen que identifica la alianza
ALL_START_DATE	-	DATE	Fecha de inicio de la alianza
ALL_END_DATE	-	DATE	Fecha de finalización de la alianza
ALL_DELETE	1	CHAR	Eliminación lógica

**Tabla 18: Tabla Dimensión Postulante**

<b>Nombre de la tabla:</b> DIM_POSTULANT			
<b>Descripción:</b> Almacena la información de los postulantes			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
SK_POSTULANT	38	NUMBER	Clave subrogada de la dimensión
POST_ID	15	VARCHAR2	Identificación del postulante
POST_TY_ID	20	VARCHAR2	Tipo de identificación
POST_FIRST_NAME	20	VARCHAR2	Nombre del postulante
POST_LAST_NAME	20	VARCHAR2	Apellido del postulante
POST_PHOTO	-	BLOB	Foto tipo carnet del postulante
POST_EMAIL	100	VARCHAR2	Email del postulante
POST_CODE_PHONE	4	NUMBER	Código de teléfono del país
POST_PHONE	38	NUMBER	Teléfono del postulante
POST_STATUS	20	VARCHAR2	Indica si el postulante se encuentra activo o inactivo
POST_DELETE	1	CHAR	Eliminación lógica

**Tabla 19: Tabla Dimensión Medio**

<b>Nombre de la tabla:</b> DIM_CHANNEL			
<b>Descripción:</b> Almacena la información de los medios de votación			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
SK_CHANNEL	38	NUMBER	Clave subrogada de la dimensión
CH_ID	38	NUMBER	Clave primaria del modelo transaccional
CH_NAME	20	VARCHAR2	Nombre del medio de votación
CH_DESCRIPTION	500	VARCHAR2	Breve descripción del medio
CH_ECA_TEMPLATE	-	CLOB	Template en caso de ser de tipo WEBMAIL

**Tabla 20: Tabla Dimensión Acta**

<b>Nombre de la tabla:</b> DIM_PHYSICAL_ACT			
<b>Descripción:</b> Almacena la información de las actas			
Campo	Tamaño	Tipo de Dato	Descripción
SK_ACT	38	NUMBER	Clave subrogada de la dimensión
PHY_ACT_CODE	50	VARCHAR2	Clave primaria del modelo transaccional
PHY_ACT_IMAGE		CLOB	Captura del acta
PHY_ACT_DELETE	1	CHAR	Eliminación lógica
PHY_ACT_STATUS	20	VARCHAR2	Estatus del acta
PHY_ACT_PAG_NUM	38	INTERGER	Cantidad de páginas de un acta

**Tabla 21: Tabla de Hechos Encuesta**

<b>Nombre de la tabla:</b> FACT_POLL			
<b>Descripción:</b> Tabla de hechos para el requerimiento de negocio de encuestas.			
Campo	Tamaño	Tipo de Dato	Descripción
SK_POLL	38	NUMBER	Clave subrogada de la dimensión de encuestas
SK_DATE	38	NUMBER	Clave subrogada de la dimensión tiempo
SK_SURVEYED	38	NUMBER	Clave subrogada de la dimensión de encuestados
SK_QUESTION	38	NUMBER	Clave subrogada de la dimensión de preguntas
SK_OPTION	38	NUMBER	Clave subrogada de la dimensión de opciones
SK_ENTITY	38	NUMBER	Clave subrogada de la dimensión de entidad
SK_USER	38	NUMBER	Clave subrogada de la dimensión de usuarios
QUANTITY	38	NUMBER	Campo que permite hacer los cálculos

**Tabla 22: Tabla de Hechos Voto**

<b>Nombre de la tabla:</b> FACT_VOTE			
<b>Descripción:</b> Tabla de hechos para el requerimiento de negocio de votos			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
SK_POSTULANT	38	NUMBER	Clave subrogada de la dimensión de postulantes
SK_ORGANIZATION	38	NUMBER	Clave subrogada de la dimensión organización
SK_VOTE	38	NUMBER	Clave subrogada de la dimensión votos
SK_ENTITY	38	NUMBER	Clave subrogada de la dimensión de entidad
SK_DATE	38	NUMBER	Clave subrogada de la dimensión tiempo
SK_ELECTION	38	NUMBER	Clave subrogada de la dimensión de elección
SK_CHANNEL	38	NUMBER	Clave subrogada de la dimensión de medios de votación
QUANTITY	38	NUMBER	Campo que permite hacer los cálculos

## B. Tabla de Indicadores

Tabla 23: Tabla de Indicadores

#	Objetivo	Nombre del Indicador	Calculo del indicador (Fórmula)	Unidad de Medida (Bs. #, %)	Frecuencia de Medición	Criterios de Clasificación	Forma de Representación
<b>Proceso de Negocio: Encuesta</b>							
1	Calcular el total de respuestas de cada una de las opciones que conforman las preguntas asociadas a una encuesta.	Total de respuestas por opción de cada pregunta.	$\sum [Nro. de respuestas por opción]$	#	Hora	Opción / Pregunta / Sexo / Edad	Gráfico de Barras / Tabla Dinámica
2	Calcular el total de encuestas respondidas en una región determinada.	Total de encuestas por región.	$\sum [Nro. de encuestas por entidad]$	#	Hora	Región / Sexo / Edad	Gráfico de Barras / Tabla Dinámica
3	Calcular el porcentaje de encuestas que aplicó un encuestador, según el sexo y/o la edad.	Porcentaje de encuestas aplicadas por un encuestador.	$\sum [Nro. de encuestas por usuario]$	%	Hora	Edad / Sexo / Usuario	Gráfico de Torta
4	Calcular la cantidad de encuestas que aplicó un encuestador en un tiempo.	Rendimiento de un usuario encuestador	$\sum [Nro. de encuestas por usuario]$	#	Hora	Edad / Sexo / Usuario	Gráfico de Barras / Gráfico de puntos
<b>Proceso de Negocio: Votos</b>							
5	Calcular el porcentaje de votos para una elección según el medio de votación utilizado.	Porcentaje de Votos por Medio de Votación.	$\frac{\sum [Nro. de votantes]}{\sum [Nro. votos por medio]} * 0.01$	%	Hora	Medio / Elección	Gráfico de Torta
6	Calcular la cantidad de votos que fueron obtenidos por partido y por postulante	Total de Votos por Partido y por Postulante.	$\sum [Nro. de votos por partido y por usuario]$	#	Hora	Partido / Postulante	Gráfico de Barras / Tabla Dinámica
7	Calcular la(s) hora(s) de mayor afluencia de votación	Afluencia de Votación	$\sum [Nro. de eventos de votación]$	#	Hora	Voto / Elección / Tiempo	Gráfico de área
<b>Proceso de Negocio: Incidencias</b>							
8	Calcular la cantidad de incidencias de acuerdo a un tipo	Cantidad de Incidencias	$\sum [Nro. de soportes (atendidos, pendientes o cerrados)]$	#	Mensual	Incidencias / Tiempo / Elección / Perfil / Tipo	Gráfico de Barras / Tabla Dinámica
<b>Proceso de Negocio: Actas</b>							
9	Calcular el total de actas almacenadas en la base de datos para una elección y una entidad específica.	Total de actas capturadas por región	$\sum [Nro. de actas]$	#	Hora	Región / Acta / Elección	Gráfico de Barras / Tabla Dinámica
10	Calcular el porcentaje de las actas almacenadas según su estatus para una elección y una entidad específica.	Porcentaje de actas según su estatus.	$\frac{\sum [Nro. de actas] (estatus)}{(Total de actas capturadas)} * 100$	%	Hora	Estatus / Elección	Gráfico de Torta
11	Calcular la cantidad de actas que un usuario con el perfil de Capturador pudo procesar	Rendimiento de un usuario capturador	$\sum [Nro. de cuentas de usuarios]$	#	Hora	Actas / Tiempo / Usuario	Gráfico de líneas



## C. Instrumento de aceptación – Módulo de reportes de TVT



Universidad Central de Venezuela

Facultad de Ciencias  
Escuela de Computación



El presente cuestionario tiene como objetivo conocer su opinión con respecto al módulo de reportes de un sistema automatizado de votación; dicha opinión será utilizada como referencia en el Trabajo Especial de Grado: SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA GESTIÓN DE UN SISTEMA AUTOMATIZADO DE VOTACIÓN. Se agradece de antemano su colaboración.

---

Este cuestionario es entregado de forma digital. En cada una de las siguientes preguntas, seleccione la respuesta que mejor se adecúe a su opinión sobre la importancia del asunto en cuestión. La escala para cada una de las opciones es la siguiente: Totalmente de acuerdo = 5; De acuerdo = 4; Ni de acuerdo, ni en desacuerdo = 3; En desacuerdo = 2; Completamente en desacuerdo = 1.

1. ¿Le parece adecuada la elección de la paleta de colores para los gráficos?  
Seleccione una opción.
2. ¿Los dashboard son interactivos?  
Seleccione una opción.
3. ¿Considera que los dashboard se encuentran equilibrados?  
Seleccione una opción.
4. ¿Le parece intuitivo la navegación entre los dashboard?  
Seleccione una opción.
5. ¿Es suficiente la información que se presenta en los gráficos para identificar las tendencias?  
Seleccione una opción.
6. ¿Los textos usados en el contenido son lo suficientemente descriptivos?

Seleccione una opción.

7. ¿Le parece adecuada el orden de visualización de la información?

Seleccione una opción.

8. A simple vista, ¿le queda claro cuál es el objetivo de cada reporte?

Seleccione una opción.

9. ¿Se contempla en los reportes los indicadores descritos en la matriz?

Seleccione una opción.

10. ¿Cree usted que este módulo le ayudará a tomar mejores decisiones en los procesos de votación?

Seleccione una opción.

11. ¿Es fácil de usar?

Seleccione una opción.

12. ¿Recomendaría la adaptación de este módulo para softwares similares?

Seleccione una opción.