



UNMdP - Facultad de Ingeniería  
Departamento de Informática



Desarrollo de central de información para la toma de decisiones y gestión  
de la información en empresa de artículos del hogar y deportivos  
de la Provincia de Buenos Aires

Alumnos

Andrade, Tobias

Della Rocca, Federico

Rabe Cappuccio, Gastón Andrés

Director

MBA Lic. Genin, Fernando

Proyecto final para optar por el grado de Ingeniero en Informática

Mar del Plata, 24 de junio de 2024



RINFI es desarrollado por la Biblioteca de la Facultad de Ingeniería de la Universidad Nacional de Mar del Plata.

Tiene como objetivo recopilar, organizar, gestionar, difundir y preservar documentos digitales en Ingeniería, Ciencia y Tecnología de Materiales y Ciencias Afines.

A través del Acceso Abierto, se pretende aumentar la visibilidad y el impacto de los resultados de la investigación, asumiendo las políticas y cumpliendo con los protocolos y estándares internacionales para la interoperabilidad entre repositorios



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).



UNMdP - Facultad de Ingeniería  
Departamento de Informática



Desarrollo de central de información para la toma de decisiones y gestión  
de la información en empresa de artículos del hogar y deportivos  
de la Provincia de Buenos Aires

## Alumnos

Andrade, Tobias

Della Rocca, Federico

Rabe Cappuccio, Gastón Andrés

## Director

MBA Lic. Genin, Fernando

Proyecto final para optar por el grado de Ingeniero en Informática

Mar del Plata, 24 de junio de 2024

# Índice

<b>Agradecimientos</b> .....	<b>3</b>
<b>Resumen</b> .....	<b>4</b>
<b>Introducción</b> .....	<b>5</b>
<b>Objetivos del proyecto</b> .....	<b>7</b>
Principal.....	7
Secundarios.....	7
Alcance.....	7
Entregables del proyecto.....	8
<b>Actores del sistema</b> .....	<b>10</b>
<b>Aporte del proyecto</b> .....	<b>11</b>
Demandante.....	11
Impacto esperado.....	11
Análisis FODA.....	12
Riesgos y planes de contingencia.....	13
<b>Análisis del problema</b> .....	<b>15</b>
Dominio.....	15
Problema a resolver.....	16
Flujo de trabajo.....	16
Requerimientos.....	17
Requerimientos funcionales.....	17
Requerimientos no funcionales.....	18
<b>Estimación inicial</b> .....	<b>19</b>
<b>Tecnologías</b> .....	<b>21</b>
Pentaho.....	21
Ventajas.....	22
Desventajas.....	22
Power BI.....	22
Ventajas.....	23
Desventajas.....	23
Comparación.....	24
Power BI > Pentaho.....	24
Power BI < Pentaho.....	24
Decisión de herramienta a utilizar.....	25
<b>Modelo de datos</b> .....	<b>26</b>
Modelo preliminar.....	26
Estrategia de maduración.....	26
Tablas auxiliares.....	27
Modelos finales.....	28
<b>Diseño de tableros</b> .....	<b>29</b>
Ajustes.....	29
Datos.....	29

Nuevas funcionalidades.....	29
<b>Metodologías utilizadas.....</b>	<b>31</b>
Metodología de relevamiento y análisis.....	31
Metodología de diseño y consolidación.....	32
Metodología de capacitación y creación de tableros.....	33
Metodología de implementación: extracción de datos.....	34
Metodología de pruebas de consistencia.....	35
<b>Producto.....</b>	<b>39</b>
Producto obtenido.....	39
Ventas.....	39
Cobranzas.....	44
Trabajos futuros.....	47
Soporte.....	47
<b>Memoria del proyecto.....</b>	<b>49</b>
Participación de los referentes.....	49
Cumplimiento de objetivos.....	49
Objetivo principal.....	49
Objetivos secundarios.....	50
Análisis FODA.....	51
Riesgos.....	52
Planificación esperada vs ejecución.....	52
Bitácora de trabajo.....	54
Resumen gráfico de etapas.....	55
Comparación de tiempos.....	56
División de tareas.....	56
Utilización de nuevas tecnologías.....	57
<b>Conclusiones.....</b>	<b>59</b>
<b>Apéndice.....</b>	<b>60</b>
Glosario.....	60
<b>Anexos.....</b>	<b>63</b>
Anexo I: Modelos de datos finales.....	63
Ventas.....	63
Cobranzas.....	64
Anexo II: Bitácora de trabajo detallada semanalmente.....	65
<b>Bibliografía.....</b>	<b>74</b>

# Agradecimientos

El apoyo constante de familiares y amigos fue determinante para lograr desarrollar el trabajo final de manera eficaz y en los tiempos previstos. Cabe destacar también la compañía de amigos y compañeros de la Facultad de Ingeniería, con quienes se vivieron grandes experiencias dentro de la institución en los últimos años. Extendemos nuestra más sincera gratitud a todos ellos.

También al cuerpo docente de la Facultad de Ingeniería, que colaboró con nuestro crecimiento tanto profesional como personal. Sin su enseñanza, llegar a esta instancia no hubiera sido posible. Queremos destacar al profesor Fernando Genin, director de este proyecto, quien nos brindó su apoyo incondicional a lo largo de todos estos meses de trabajo.

Por último, a los referentes funcionales, quienes confiaron en nosotros para solucionar su problemática y nos proveyeron *feedback* constante a lo largo del proyecto.

# Resumen

El objetivo principal del proyecto es optimizar la gestión y análisis de datos mediante la creación de tableros interactivos para mejorar la toma de decisiones en la empresa.

El análisis del problema incluyó un estudio detallado del dominio y el alcance del proyecto. A partir de los requerimientos funcionales y no funcionales se decidió realizar un sistema de BI (Business Intelligence). Se evaluaron las tecnologías disponibles para la creación de reportes y se optó por utilizar Power BI. Estos fueron armados con datos preprocesados que se guardaron en un esquema desnormalizado en la base de datos de la empresa.

El producto final incluyó tableros de ventas y cobranzas, proporcionando una visión clara y detallada de estos aspectos críticos del negocio. Los entregables abarcaron un sistema funcional de ETL y reportes con actualización automática en formato de tableros en Power BI.

En conclusión, se cumplió con los objetivos del proyecto al mejorar significativamente la eficiencia en la generación de reportes y facilitar el acceso a la información crítica para la toma de decisiones. Además, se sentaron las bases para futuras extensiones y mejoras en la gestión de datos de la empresa. Este trabajo no solo ha optimizado la operación actual, sino que también ha presentado nuevas oportunidades para el crecimiento y la mejora continua.

# Introducción

El acceso simplificado y ágil a la información empresarial se considera esencial. Por ende, se pretende transformar y consolidar datos provenientes de diversas fuentes para facilitar la extracción diaria de información, lo que proporciona una visión más completa y actualizada del desempeño empresarial. La implementación de tableros en Power BI aporta una dimensión visual a la información, seleccionando visualizaciones pertinentes e incorporando indicadores clave de rendimiento (*KPI*) relevantes. Esto asegura que todos los niveles de la empresa tengan acceso a informes claros y comprensibles que respalden decisiones informadas y estratégicas.

Se suma a esto el hecho de que los reportes actuales no cumplen con los estándares de claridad y comprensión necesarios, impactando negativamente tanto en la toma de decisiones como en la capacidad de respuesta a las dinámicas cambiantes del mercado. También se busca poner en práctica las habilidades de los participantes, necesarias para poder debatir ideas de diseño del esquema desnormalizado, aspecto de tableros, etc.

La modernización propuesta no solo responde a los desafíos actuales, sino que también se posiciona como una estrategia para preparar a la empresa para el futuro. Esto implica la integración de tecnologías avanzadas y metodologías innovadoras para superar las barreras operativas y garantizar una adaptación continua a un entorno empresarial en constante evolución.

A través del análisis, se buscará identificar posibles errores o inconsistencias tanto en las bases de datos como en los reportes generados a partir de ellas. Este proceso de revisión minuciosa no sólo garantizará la integridad y precisión de los datos utilizados en la toma de decisiones, sino que también permitirá corregir cualquier problema subyacente que pueda afectar la fiabilidad de los informes empresariales. De este modo, se establecerá una base sólida y confiable para el desarrollo de soluciones de análisis de datos.

Cabe destacar que se cuenta con la ayuda de los referentes funcionales, ya que ellos tienen el conocimiento absoluto de su sistema de base de datos montado en

SQL Server, y podrán hacer recomendaciones acerca de dónde sacar ciertos datos para generar reportes, aclarar dudas, brindarnos su visión según sus gustos de cómo quieren que se vean los tableros, la exactitud con la que quieren los datos, entre otras.

# Objetivos del proyecto

## Principal

Que la empresa pueda disponer de información sin errores para la toma de decisiones, que se genere en forma rápida, automática y pueda ser accedida en forma remota.

## Secundarios

Se detallan a continuación:

- Liberar recursos humanos para tareas estratégicas. Al automatizar los procesos de generación de informes, se busca liberar a los empleados de tareas operativas, permitiéndoles dedicar ese tiempo y recursos a actividades estratégicas que agreguen valor al negocio. Es importante remarcar que, hasta el momento, este proceso se hacía manualmente.
- Evitar los errores humanos que pudieran estar asociados al procesamiento manual de reportes y generación de informes. Con grandes volúmenes de datos, repetir este proceso periódicamente puede generar errores que impactarán directamente en los informes que poseen información crucial y sensible para la empresa.

## Alcance

El alcance del proyecto abarcó el desarrollo de un sistema automatizado que recopile datos de múltiples fuentes, además de la implementación de procesos de transformación de datos para limpiar, organizar y preparar la información de manera coherente y estandarizada.

Se contempló la generación de reportes de manera automatizada, con el objetivo de facilitar la toma de decisiones informadas. También se incluyó la identificación de posibles errores o inconsistencias, y la eliminación de datos duplicados e información obsoleta, de manera de asegurar la integridad y precisión de los datos utilizados en la toma de decisiones.

Por otro lado, la empresa hizo hincapié en controlar el acceso a cada reporte, para que los miembros de cada sector sólo puedan visualizar los necesarios para identificar tendencias y patrones en los datos de su propia área. Esto fue tenido en cuenta como parte del proyecto.

No entró en el alcance una capacitación para garantizar la adopción efectiva de las nuevas herramientas y procesos. Se acordó con la empresa que la interfaz de usuario debería ser intuitiva y fácil de usar.

No estuvo contemplada la organización de reuniones posteriores a fin de proyecto para extender las funcionalidades y creación de reportes, ni tampoco soporte técnico para atender consultas, resolver problemas y brindar asistencia continua. Estas tareas, que estuvieron fuera del alcance del proyecto, podrían coordinarse posteriormente dependiendo de las necesidades de la empresa.

## Entregables del proyecto

Teniendo en cuenta las etapas y los objetivos anteriores, se detallan a continuación los entregables del proyecto:

1. Diseño de una base de datos desnormalizada y consolidación de datos en tablas optimizadas: la estructura desnormalizada y optimizada simplificó la extracción y consolidación de datos. Este diseño estuvo basado en una evaluación de las fuentes de datos existentes y en los requisitos de informes estratégicos identificados. Esta desnormalización proporcionó una estructura simplificada que facilitó la extracción eficiente de datos para su posterior análisis y visualización. La consolidación de datos en tablas optimizadas garantizó que los informes generados fueran precisos, actualizados y pertinentes para la toma de decisiones empresariales. Las tablas fueron diseñadas para optimizar el acceso a la información clave.

2. Implementación de la extracción diaria de datos: se estableció un proceso automatizado que garantiza la actualización diaria de la información en el nuevo esquema de la base de datos. Este proceso se diseñó para extraer datos de todas las fuentes relevantes de manera programada y automatizada, lo que aseguró que los informes generados reflejen la información más reciente disponible.

3. Creación de tableros en Power BI: se desarrollaron tableros personalizados en Power BI diseñados específicamente para cada área de la empresa. Estos tableros presentan la información de manera clara y comprensible, utilizando una variedad de gráficos, tablas y visualizaciones interactivas. Los tableros en Power BI se diseñaron para mostrar los datos más relevantes para el usuario, proporcionando una visión integral del rendimiento empresarial. También se configuraron permisos para cada tipo de usuario según su rol en la empresa, de manera que solo pueda visualizar la información que le corresponda.

Estos entregables constituyeron un sistema integrado de Business Intelligence (BI) que busca transformar la manera en que la empresa utiliza y aprovecha sus datos. Es una base sólida para la toma de decisiones informada y ágil, cuyo fin es permitirle a la empresa adaptarse rápidamente a las cambiantes condiciones del mercado y mejorar su rendimiento general.

# Actores del sistema

- Demandante: empresa dedicada a la venta de artículos para el hogar y deportivos de la Provincia de Buenos Aires.
- Departamento de IT de la empresa en cuestión (referentes funcionales).
- Equipo formado por el director y los estudiantes.
- Empresa informática de soporte IT en BD.
- Asesor tecnológico de la empresa.
- Proveedor de *software*.

Dentro del departamento de IT, se encuentra parte de la gerencia de la empresa. Por dicha razón, este actor define tanto los requerimientos funcionales como los no funcionales.

# Aporte del proyecto

## Demandante

El proyecto fue solicitado por una empresa que vende artículos para el hogar y deportivos de la Provincia de Buenos Aires al director del proyecto. La idea surgió a partir de la gerencia de la empresa que buscaba mejorar el rendimiento en la generación de sus reportes y en el proceso de toma de decisiones.

No solo buscaban mejorar la eficiencia operativa y claridad de la información, sino también influir en la cultura empresarial promoviendo la toma de decisiones informada.

Por cuestiones de confidencialidad de la información, La Empresa solicitó figurar en modo anónimo en el presente Trabajo Final.

## Impacto esperado

En primer lugar, la automatización y optimización del proceso de generación de reportes busca liberar recursos humanos y reducir el tiempo dedicado a tareas operativas, lo que permitiría a los empleados enfocarse en sus actividades. Además, la implementación de tableros en Power BI tiene como objetivo brindar una visualización clara y comprensible de los datos, facilitando la toma de decisiones informadas en todos los niveles de la organización. Esta mejora en la eficiencia operativa y la capacidad de análisis de datos tendría un impacto positivo en la productividad, la agilidad empresarial y la capacidad para adaptarse a los cambios del mercado, lo que fortalecería la competitividad de la empresa a largo plazo.

Además, se busca abordar los desafíos presentes y también establecer una estrategia orientada hacia el futuro, tal como incorporar tecnologías y enfoques innovadores para superar los obstáculos operativos y garantizar la adaptabilidad de la empresa a medida que evoluciona el entorno empresarial.

## Análisis FODA

Previo al inicio del proyecto, se llevó a cabo un análisis FODA, que permitió obtener una visión integral de los factores internos y externos que podrían influir en el desarrollo y éxito del proyecto. Identificar y evaluar estos elementos críticos antes de la ejecución del proyecto fue fundamental para diseñar estrategias efectivas, mitigar riesgos potenciales y aprovechar al máximo las oportunidades disponibles.

### Fortalezas

- La empresa muestra una firme voluntad de adoptar la transformación digital y modernización de procesos.
- La experiencia del cliente consolidada en la venta de sus artículos proporciona un conocimiento profundo del mercado y sus necesidades.
- Se cuenta con un equipo de profesionales capacitados y familiarizados con las operaciones de la empresa.
- El proyecto no representa una carga económica significativa para la empresa.
- La presencia de una infraestructura tecnológica existente, tales como bases de datos, facilita la implementación de nuevas soluciones de Business Intelligence.
- La capacitación en nuevas tecnologías y metodologías puede potenciar el desarrollo del talento interno, mejorando la capacidad de adaptación a cambios futuros.
- La empresa dispone de datos históricos, acceso libre a los mismos y documentación de los diseños de registros.

### Oportunidades

- La inserción de herramientas de Business Intelligence puede capitalizar el crecimiento del mercado al facilitar una toma de decisiones más ágil y basada en datos.
- La adopción de tecnologías avanzadas puede proporcionar a la empresa ventajas competitivas en términos de eficiencia operativa y satisfacción del cliente.

## Debilidades

- La resistencia al cambio puede surgir entre el personal ante la introducción de nuevas tecnologías y procesos.
- La integración de nuevas soluciones con la infraestructura existente puede presentar desafíos técnicos y operativos.
- Puede ser necesario invertir en programas de capacitación adicionales para asegurar una adopción efectiva de las herramientas de Business Intelligence.
- Requiere inversión inicial para implementar la solución (costo de licencias).

## Amenazas

- La competencia puede adoptar tecnologías similares, lo que destaca la necesidad de una rápida implementación y adaptación continua.
- Problemas técnicos durante la implementación podrían causar interrupciones operativas y resistencia adicional al cambio.
- Posibles riesgos al manejar datos desnormalizados o duplicación errónea de información.
- Riesgo de interrupciones en el proceso de extracción de datos diaria.
- Baja o nula colaboración por parte de los proveedores de sistemas

## Riesgos y planes de contingencia

Se analizaron los factores que pueden poner en peligro o hacer inviable su finalización. Para cada uno, se ponderó la probabilidad de ocurrencia y el impacto que tendría en el proyecto en una escala del 1 al 3, en orden ascendente. El peso del riesgo sale del producto de ambos valores.

Si el peso resultante es mayor o igual a 6, se debe elaborar un plan de contingencia para saber cómo responder en caso de que se produzca y contrarrestar el impacto sobre el proyecto.

Riesgo	Descripción	Consecuencia	Probabilidad	Impacto	Peso
R01	Integrante abandona el proyecto	Demoras importantes en el proyecto	1	3	3
R02	Imposibilidad de coordinar reuniones con los referentes funcionales	El producto de <i>software</i> concebido no se ajusta a las necesidades del cliente	1	3	3
R03	Ausencia de uno o más miembros del equipo	Mayor trabajo para los otros miembros, además de un incremento en el tiempo estimado	3	2	6
R04	Baja o nula colaboración por parte del soporte de IT en la BD	Complicaciones para acceder y manipular la BD de la empresa. Demoras en el proyecto	2	2	4
R05	Mala planificación por parte de los integrantes	Demoras importantes en el proyecto	3	2	6
R06	Problema de compatibilidad con sistema existente	Imposibilidad de trabajar con la herramienta propuesta	1	3	3
R07	Resistencia al cambio por parte de los empleados	Los tableros no son utilizados. Fracaso del proyecto	1	3	3
R08	Aumento en el costo de las licencias de la herramienta. La empresa decide no seguir pagándolas	Los tableros no son utilizados. Fracaso del proyecto	1	3	3

En cuanto a los planes de contingencia:

- R03: las tareas asignadas al integrante ausente las deberá desarrollar fuera del horario planificado, o en su defecto se le asignan a otro integrante del proyecto de manera que se pueda cumplir con la tarea en tiempo y forma.
- R05: se trabajará más horas para cumplir con los tiempos estipulados. Si esto no fuera suficiente, se comunicará el retraso a los demandantes.

# Análisis del problema

## Dominio

En primer lugar, en el sector de la venta tanto de artículos para el hogar como de indumentaria deportiva, la competencia es intensa y está en constante evolución. Las empresas del sector enfrentan desafíos significativos en términos de gestión de inventario, optimización de precios, seguimiento de tendencias del mercado y satisfacción de los consumidores. Estos desafíos requieren una comprensión profunda de los datos relacionados con las ventas, el inventario, los clientes y otros aspectos del negocio.

Uno de los problemas principales que enfrenta la empresa es la falta de visibilidad y comprensión de los datos. Actualmente, la empresa utiliza sistemas de información dispersos y poco integrados, además de una realización manual de tareas como la extracción de los datos y la creación de reportes, lo que dificulta la recopilación y el análisis eficientes de los datos relevantes. Esto puede llevar a decisiones erróneas, ya que la probabilidad de error es grande, y también a una falta de capacidad para responder rápidamente a las demandas cambiantes del mercado.

Además, la empresa carece de herramientas adecuadas para realizar análisis avanzados de datos y crear informes de manera automática. Sin acceso fácil a información clave, como tendencias de ventas y comportamiento de los clientes, la empresa puede perder oportunidades importantes para mejorar su competitividad y rentabilidad. Otro desafío común en este sector es la gestión eficiente del inventario. La empresa puede enfrentar problemas de exceso de inventario o agotamiento de existencias, lo que puede afectar negativamente a la rentabilidad y la satisfacción del cliente. Una mejor comprensión de los patrones de demanda y las tendencias de ventas puede ayudar a optimizar la gestión del inventario y reducir los costos operativos.

## Problema a resolver

El problema a resolver está relacionado con la baja eficiencia de los procesos relacionados a la extracción de datos y generación de reportes. Además, también se destaca la falta de visibilidad y comprensión de los datos claves necesarios para tomar decisiones estratégicas y operativas.

Actualmente, la empresa enfrenta obstáculos para acceder a datos, lo que dificulta la obtención de una visión integral y actualizada del rendimiento del negocio. Este problema se manifiesta en varios aspectos críticos. En primer lugar, la falta de una central única de información dificulta la consolidación y análisis efectivo de datos relacionados con ventas, cobranzas, clientes y otros aspectos comerciales claves.

## Flujo de trabajo

A continuación se muestra el Modelo y Notación de Procesos de Negocio (BPMN, por sus siglas en inglés), que refleja cómo es la interacción con el sistema por parte de los miembros de los sectores de la empresa. El objetivo es que tanto la gerencia como los encargados y vendedores de cada sucursal tengan acceso inmediato y de manera sencilla a la información que necesiten siempre que sea correspondida.

Con respecto al control de accesos, lo maneja la misma herramienta, por ende ellos no podrán visualizar tableros que no les corresponda acceder, y por esa razón no se reflejan en este modelo.

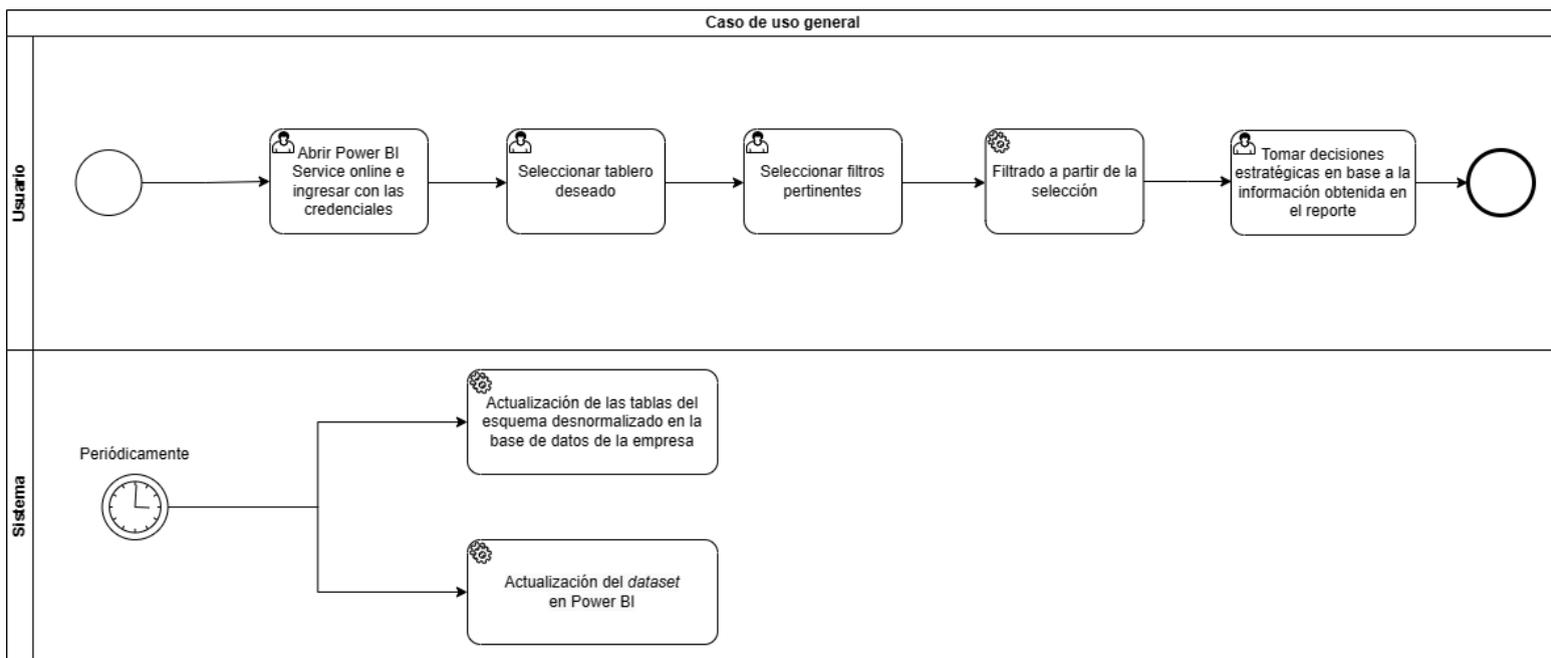


Figura 1: BPMN principal de interacción.

## Requerimientos

Los requisitos enumerados a continuación surgieron durante la fase de relevamiento, incluyendo tanto los funcionales como los no funcionales.

### Requerimientos funcionales

- RF01: la extracción de datos deberá ejecutarse de manera diaria y por la noche, fuera del horario comercial.
- RF02: los tableros deberán ser intuitivos, brindando información clara para su interpretación y análisis.
- RF03: los tableros contarán con un selector de fechas (mes/año, y en algunos casos días) que permitirá elegir hasta qué fecha se desean ver datos.
- RF04: se deberán poder aplicar filtros sobre los objetos visuales de cada tablero.
- RF05: los tableros deberán poder accederse de manera controlada; sólo serán capaces de visualizar cada tablero los usuarios que tengan el permiso de hacerlo.

- RF06: los tableros deberán actualizarse diariamente a partir de la base de datos de la empresa. La actualización deberá ser automática y no requerirá acción humana.

## Requerimientos no funcionales

- RNF01: la ejecución de los *scripts* que realizan la extracción diaria de datos no deberá demorar más de cinco minutos.
- RNF02: los tableros deberán cargar y poder visualizarse en menos de dos segundos.
- RNF03: si se aplica uno o varios filtros sobre algún objeto visual de un tablero, la información deberá ser actualizada en menos de tres segundos.

# Estimación inicial

A partir del trabajo y la metodología prevista luego de la primera reunión con el equipo, se realizó la estimación inicial de tiempos supuestos que se tardarían en realizar cada etapa por completo. La metodología de trabajo se supuso en cascada.

Se dividió el proyecto en las etapas de: relevamiento de las necesidades empresariales; análisis del modelo de la empresa; diseño de registros de las tablas del esquema desnormalizado; recolección de datos para la creación de las tablas; implementación de la extracción diaria de datos; capacitación del equipo en la tecnología seleccionada para la creación de los informes; pruebas de consistencia; y la elaboración de la documentación final. Todas estas tareas, que abarcan la totalidad del alcance del proyecto inicial, se supuso que se realizarían de manera secuencial, avanzando a la etapa siguiente únicamente luego de terminar la que estaba en curso.

Los miembros del equipo nunca habían trabajado con una tecnología de Business Intelligence, como así tampoco tenían experiencia profesional del manejo de una base de datos empresarial, sus protocolos, formas de operar, etc. Por estas razones, la estimación de tiempos de las tareas (en horas por semana), se realizó de manera aproximada, teniendo en cuenta proyectos/trabajos anteriormente realizados con otras tecnologías.

Fecha de inicio prevista: 18 de diciembre de 2023.

Fecha de finalización prevista: mayo de 2024.

N	Etapa	Semana																							
		Dic.		Enero					Febrero					Marzo					Abril				Mayo		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1	Relevamiento de las necesidades de la empresa	15	15																						
2	Análisis del modelo de datos de la empresa			15	15																				
3	Diseño de registros de las tablas del esquema desnormalizado				40	40	40	40	40																
4	Recolección de datos para la creación de las tablas									30	30	25													
5	Implementación de la extracción diaria de datos												20	10											
6	Capacitación en herramienta de BI														15	15									
7	Creación de informes en herramienta de BI															50	50	50	50	50	50	50	50		
8	Pruebas de consistencia de datos																			20	15	15	15		
9	Elaboración de la documentación final		5		5				5			5		5		5						10	10	30	

# Tecnologías

La base de datos de la empresa está montada sobre SQL Server. Para el desarrollo del proyecto, se debía decidir qué herramienta utilizar, tanto para los procesos de extracción, transformación y carga (ETL por sus siglas en inglés) como para el desarrollo de los tableros. Luego de investigar acerca de distintas tecnologías, se consideraron Pentaho y Power BI como posibles. Como no se tenía experiencia previa en el manejo de ninguna de las dos, el equipo se dedicó a investigar cada una en detalle.

## Pentaho

Es un conjunto de herramientas de *software* de código abierto que proporciona una amplia gama de funcionalidades para la integración de datos, el análisis de datos y la generación de informes. Es utilizada por organizaciones de todo el mundo para implementar soluciones de inteligencia empresarial (BI), análisis de datos, procesamiento de grandes volúmenes de datos y automatización de procesos de negocio.

Dentro de las herramientas de Pentaho se encuentran: Pentaho Data Integration que proporciona capacidades completas de ETL para extraer datos de diversas fuentes, transformarlos según las necesidades del negocio y cargarlos en destinos específicos, como almacenes de datos o sistemas de BI. También está Pentaho Reporting, que permite a los usuarios crear informes interactivos y personalizados con una variedad de gráficos, tablas y otros elementos visuales. Pentaho Dashboard Designer ofrece la posibilidad de diseñar tableros de control dinámicos y visualmente atractivos que permiten a los usuarios monitorear y analizar datos en tiempo real. Pentaho Metadata Editor, que permite a los usuarios definir y gestionar metadatos de manera centralizada, lo que facilita la comprensión y el uso de los datos en toda la organización.

Cabe recalcar que Pentaho es compatible con una amplia variedad de tecnologías y plataformas, incluidas bases de datos relacionales, bases de datos NoSQL, almacenamiento en la nube y sistemas de big data como Hadoop.

## Ventajas

- Código abierto: es una herramienta de código abierto, lo que significa que es gratuita para descargar y usar, lo que hace que el acceso a toda su funcionalidad sea gratuita.
- Amplia funcionalidad: ofrece una amplia gama de funcionalidades, desde ETL hasta análisis avanzado de datos.
- Flexibilidad y escalabilidad: permite adaptar la herramienta a las necesidades específicas de cada usuario y escalarla para manejar grandes volúmenes de datos.

## Desventajas

- Curva de aprendizaje: puede tener una curva de aprendizaje pronunciada para los usuarios nuevos, especialmente aquellos sin experiencia previa en herramientas de BI y análisis de datos.
- Documentación limitada: la documentación es limitada en comparación con otras soluciones comerciales de BI.
- Rendimiento: en entornos de gran escala y con grandes volúmenes de datos, Pentaho puede experimentar limitaciones de rendimiento en comparación con soluciones de BI comerciales de licencia paga.
- Interfaz de usuario: no presenta una interfaz de usuario intuitiva, por lo que puede traer complicaciones para la adaptación de los gerentes.

## Power BI

Power BI es una herramienta de inteligencia empresarial desarrollada por Microsoft que ofrece una amplia gama de funcionalidades para la visualización de datos, análisis de datos, creación de informes y procesos de ETL.

Esta tecnología permite crear visualizaciones interactivas y atractivas, como gráficos, tablas dinámicas, mapas y otros elementos visuales, que facilitan la comprensión de los datos y la identificación de tendencias y patrones. Además, ofrece capacidades avanzadas de análisis de datos, incluyendo el uso de medidas, cálculos DAX (Data Analysis Expressions) y funciones estadísticas para realizar

análisis profundos sobre los datos, lo que permite crear informes personalizados y dinámicos a partir de múltiples fuentes de datos. Los informes pueden ser compartidos y colaborados en tiempo real con otros usuarios dentro de la organización. Power BI ofrece capacidades de ETL a través de herramientas como Power Query y Power Automate, que permiten extraer datos de diversas fuentes, transformarlos según las necesidades del negocio y cargarlos en modelos de datos optimizados para análisis. Cabe destacar que se integra con otras herramientas de Microsoft, como Excel, SharePoint y Azure, lo que facilita la conexión y el intercambio de datos entre diferentes plataformas.

## Ventajas

- Facilidad de uso: cuenta con una interfaz intuitiva y fácil de usar que permite a los usuarios crear visualizaciones e informes sin necesidad de habilidades de programación avanzadas.
- Integración con el ecosistema de Microsoft: la integración con otras herramientas y servicios de Microsoft facilita la conexión y el intercambio de datos entre diferentes plataformas y sistemas.
- Amplia gama de visualizaciones: ofrece una amplia variedad de visualizaciones predefinidas y personalizables que permiten presentar los datos de manera efectiva y atractiva.
- Capacidad de análisis avanzado: proporciona capacidades avanzadas de análisis de datos, como cálculos DAX y funciones estadísticas, que permiten a los usuarios realizar análisis profundos sobre los datos.
- Curva de aprendizaje: Power BI es relativamente fácil de usar, únicamente puede tener una curva de aprendizaje un poco mayor a la hora de indagar en el ámbito de automatización de procesos de carga.

## Desventajas

- Costo: aunque ofrece una versión gratuita, las funcionalidades más avanzadas están disponibles en las versiones de pago, lo que puede resultar costoso para algunas organizaciones dependiendo la cantidad de licencias a contratar.

- Limitaciones en la versión gratuita: la versión gratuita tiene ciertas limitaciones en términos de almacenamiento de datos y capacidad de procesamiento, lo que puede ser restrictivo para algunas empresas.

## Comparación

### Power BI > Pentaho

Cuando una empresa ya utiliza productos de Microsoft, como Office 365, SharePoint o Azure, Power BI se integrará de manera más fluida con estas herramientas.

Para usuarios que prefieren una interfaz intuitiva y amigable, Power BI ofrece una curva de aprendizaje más rápida y una experiencia de usuario más simple. Además es ideal cuando se necesita crear visualizaciones y paneles de control de manera rápida y eficiente, sin requerir una configuración compleja, sumado a que la herramienta en sí es más fácil de implementar.

### Power BI < Pentaho

Pentaho ofrece más opciones y control sobre la integración de datos, análisis y generación de informes, lo que la hace más flexible y personalizable. Para proyectos que requieren manejar grandes volúmenes de datos y entornos empresariales complejos, ofrece una escalabilidad robusta y puede adaptarse a diferentes necesidades.

En cuanto a los procesos ETL más complejos, Pentaho Data Integration proporciona una amplia gama de funcionalidades y opciones para transformar y mover datos de manera efectiva.

Para organizaciones con presupuestos limitados o que prefieren soluciones de código abierto, Pentaho puede ser una opción más económica debido a su naturaleza de código abierto y a las versiones gratuitas disponibles.

## Decisión de herramienta a utilizar

Teniendo en cuenta las ventajas, desventajas y posibilidades que ofrece cada herramienta, el equipo decidió junto con los referentes funcionales que sería una mejor idea trabajar con Power BI.

La decisión recae sobre los factores tiempo e interfaz. Al tener una curva de aprendizaje menor y una interfaz para la generación de reportes más dinámica e intuitiva, el tiempo de diseño y desarrollo de los tableros sería mucho menor. En lo que respecta a los costos, la empresa está dispuesta a asumir el pago de las licencias, dado que no necesitan una gran cantidad. Se trata de un gasto mínimo mensual que se justifica plenamente. Con estas licencias no solo se soluciona el tema de los permisos para el acceso restringido y discriminado a la información, sino que también se permite el uso de funciones de actualización automática fundamentales para que se disparen diariamente los procesos ETL.

# Modelo de datos

## Modelo preliminar

En base a los requerimientos relevados en la primera etapa, se trabajó en las etapas de diseño de los esquemas desnormalizados. Para ello, se tuvo en cuenta la información brindada por los referentes funcionales acerca del modelo de datos de la empresa, así como también los *stored procedures* de su base de datos, que daban indicios acerca de cómo obtener o calcular ciertos parámetros necesarios para los reportes.

Así fue para las secciones de ventas y cobranzas, pero no para la de gastos, planificada en el alcance del proyecto. Luego de dos reuniones con los referentes funcionales y el asesor tecnológico de la empresa, el equipo les expresó la imposibilidad de realizar el trabajo de esta sección debido a la falta de información digitalizada, ya que la empresa carece de procesos automatizados para la carga de datos relevantes en su sistema. Por dicha razón, la empresa decidió excluir a “gastos” del desarrollo del proyecto. En el apartado “Bitácora de trabajo” se desarrollará más en profundidad esta decisión.

## Estrategia de maduración

Luego de la etapa de diseño se comenzó con la consolidación de los datos. Durante esta etapa, el foco de atención estuvo en la creación de las medidas necesarias para cada uno de los objetos visuales que conforman distintos tableros en Power BI. Para ello, fue necesario entender bien qué representaba cada campo y cómo se obtenía cada dato a representar en los reportes.

Además, fue necesario tener un *dataset* cargado en Power BI para armar tableros provisorios, de manera de testear las medidas que se iban desarrollando. Estos resultados se comparaban con los obtenidos por los referentes funcionales en sus reportes manuales. De esta manera, también se encontraron problemas asociados a campos faltantes, que fueron necesarios agregar para lograr informes óptimos.

Entre otros campos, se encontró que faltaban algunas fechas que eran fundamentales para filtrar la información a partir de la selección del usuario. En la sección de cobranzas era primordial saber de qué fecha era una transacción, cuándo vence cada cuota y cuándo fue pagada (si lo fue). Esto se debe a que los referentes funcionales deseaban ser capaces de filtrar y ver en el reporte la situación de cobranza de cuotas en un día, mes y año en específico. Para manejar transacciones pasadas como si fuesen actuales, el sistema debe comportarse como si estuviese parado en esa fecha del pasado, por ende las transacciones de fechas posteriores y sus cuotas no deberían ser tenidas en cuenta porque, en teoría, en ese momento eran inexistentes. Esto presentó un gran desafío para el equipo.

## Tablas auxiliares

Para completar la información y crear los tableros, se tuvo que recurrir a tablas auxiliares.

En la sección de ventas, la información de los objetivos tanto por sucursal como por cada vendedor no correspondía agregarse en el esquema desnormalizado armado. El esquema de ventas contiene únicamente una tabla que posee la información necesaria de cada transacción. Por ende, era necesario encontrar la forma de anexar la información acerca de los objetivos de la empresa en Power BI sin duplicar datos innecesariamente en el esquema de ventas. Esta información de objetivos se encontraba almacenada en diversas tablas del esquema principal de la base de datos de la empresa. Como la herramienta permite crear tablas a partir de consultas a una base de datos remota montada en SQL Server, se logró extraer de esta manera la información deseada.

En la sección de cobranzas, se creó una tabla auxiliar “calendario”, que sirve para luego crear un filtro visual a partir del cual el usuario elija la fecha de la información que desee ver. Esta tabla posee únicamente una columna de fecha, y fue de gran ayuda ya que las comparaciones de las fechas de transacción, vencimiento de cuota y aplicación se realizan contra esta fecha seleccionada por el usuario. De esta manera, se resolvió de manera óptima el problema mencionado anteriormente asociado a ubicar al sistema en el pasado y que se comporte como tal. Esta solución no es nueva; es ampliamente utilizada en sistemas de BI.

Por último, durante la creación de los tableros, la empresa comentó que ocasionalmente eliminan registros de su tabla de pagos de cuotas (aplicaciones). Esto afecta directamente a los reportes, ya que en la tabla de aplicaciones del esquema desnormalizado se deberían eliminar también. Las aplicaciones eliminadas pueden ser de cualquier fecha, desde una aplicación del día de ayer hasta una aplicación de tres años atrás.

El problema asociado a la eliminación de aplicaciones se debe a la automatización de la extracción diaria de datos. Como periódicamente se agregan las entradas correspondientes al día anterior, las aplicaciones eliminadas de fechas previas al día anterior no estarían consideradas en la actualización. Esto se debe al uso de la herramienta de actualización incremental en Power BI. Para considerar fechas anteriores, habría que cambiar el rango de actualización y seleccionar la opción de detectar cambios en los datos. Esta solución no fue tomada en cuenta, ya que depende de un rango fijo que se le asigne para actualizar los datos y, según los referentes funcionales, puede que se desee eliminar una aplicación de varios años atrás. Como esa solución no sería nada eficiente, para resolver este problema se acordó con los referentes funcionales crear en su esquema principal de base de datos una tabla de aplicaciones eliminadas. De esta manera, desde Power BI se importó esa tabla y, con un simple filtro, fue posible desestimar las aplicaciones eliminadas sin importar de qué fecha fueran.

## Modelos finales

Considerando todo lo mencionado anteriormente, se llegó a los modelos de datos finales (ver Anexo I), que fueron diseñados para optimizar el procesamiento y la integración de la información crucial para el proyecto.

# Diseño de tableros

El primer paso en cada etapa de creación de los tableros fue realizar todas las medidas que se necesitaban. Posteriormente, se crearon distintos objetos visuales en cada tablero, a partir de las medidas correspondientes, para visualizar los reportes de la empresa en un formato similar al que lo venían haciendo.

Una vez que los resultados eran consistentes con los reportes manuales de la empresa, se reorganizaron los elementos visuales de manera de brindar mayor claridad para su análisis. También se eligió un tema acorde a los colores de la empresa, entre otras mejoras estéticas. Posteriormente, se realizaron distintas reuniones virtuales donde los referentes manifestaron los cambios o mejoras que deseaban hacerle a los tableros actuales, y a su vez el equipo aconsejó o sugirió otras formas de visualizar la información que fueran más convenientes en algunos casos, como por ejemplo reemplazar una tabla de doble entrada por un gráfico de barras apiladas.

## Ajustes

Durante el intercambio de ideas y opiniones con los referentes, uno de los temas a tratar, además de la apariencia de los tableros, fueron las modificaciones en la información presentada. Se pueden clasificar en ajustes de:

### Datos

Estos ajustes son aquellos relacionados con los desvíos conocidos y desconocidos explicados en una sección previa de este informe. Los referentes de la empresa nos comunicaban en cada caso qué grado de especificidad y exactitud pretendían de los datos reflejados en los reportes de Power BI con respecto a los que ellos creaban manualmente.

### Nuevas funcionalidades

La capacidad de adaptación y flexibilidad de las medidas juega un papel crucial en la resolución de problemas relacionados con la visualización de datos. Las medidas

de la herramienta son polimórficas. Esto implica que las medidas pueden comportarse de manera diferente según el contexto y el tipo de gráfico en el que se utilicen, lo que permite que una misma medida proporcione resultados relevantes en distintos escenarios. Además, la capacidad para reutilizar medidas agiliza el proceso de creación y modificación de tableros, permitiendo a los usuarios adaptar rápidamente la presentación de los datos a sus necesidades específicas sin necesidad de realizar cambios complicados en la estructura subyacente de los datos.

Como consecuencia, en los tableros se lograron crear fácilmente nuevas representaciones que en sus reportes manuales eran extremadamente laboriosas, o no eran posible de hacer.

# Metodologías utilizadas

Al comenzar el proyecto, la empresa hizo hincapié en los tiempos esperados. En cuatro meses tenían que estar los tableros implementados y con los datos pertinentes. Para garantizar esto, fue fundamental una buena planificación y organización entre los integrantes del grupo.

Se comenzó con una reunión donde los referentes explicaron con detalle la idea del proyecto que tenían, el sistema de base de datos con el que contaban y qué pretendían obtener como resultado final. En primera instancia, se contempló una metodología de trabajo en cascada, donde el cliente iba a visualizar los datos y todos los tableros finalizados en última instancia. Sin embargo, luego de una segunda reunión, los referentes nos comunicaron la preferencia de dividir el trabajo en tres secciones: ventas, cobranzas y gastos. Por esta razón, el equipo optó por adoptar una metodología de trabajo ágil, priorizando la satisfacción del cliente al entregar rápidamente nuevas funcionalidades que agregan valor. Se trata de una entrega incremental, donde el producto se desarrolla y entrega de manera continua, con requerimientos que se agregan en cada iteración. Además, permite obtener una retroalimentación temprana y garantizar que el resultado se ajuste a las necesidades del cliente.

Con esta metodología, se trabajó de a una sección por vez, lo que brindó mayor eficiencia al tener la atención puesta en un tema específico y no en la generalidad del proyecto constantemente.

## Metodología de relevamiento y análisis

En diciembre del año 2023 se tuvo la primera reunión con los referentes funcionales de la empresa para dar inicio al proyecto. En una extensa charla, que incluyó la presentación del equipo, se comentaron las exigencias del proyecto y el producto esperado por el cliente. Con esta información, estaban dadas las condiciones para analizar la problemática en cuestión, hacer la estimación inicial de tiempos, analizar el modelo de base de datos con el cual se trabajaría a lo largo del proyecto, y

recabar todas las dudas existentes tanto acerca del modelo de datos como de la metodología de trabajo a seguir.

En una segunda reunión, se evacuaron dudas acerca de algunos campos de las tablas relacionadas con ventas que resultaban ambiguos o que tenían nombres poco auto descriptivos, y acerca de dónde sacar otros datos necesarios. Un punto importante de esta reunión fue plantear una metodología ágil de trabajo, donde se trabajaría con *sprints* de duración variable. Un *sprint* por cada tablero, de manera de garantizar que el cliente vaya viendo el avance del proyecto y dé un *feedback* al equipo. La duración de cada *sprint* dependería de la complejidad del tablero y de la disponibilidad horaria del equipo.

Durante el desarrollo del proyecto se continuaron haciendo consultas sobre el modelo de datos y los distintos campos de cada tabla, principalmente para entender qué representaba cada dato y qué valores podía tomar.

## Metodología de diseño y consolidación

Luego de tener en claro los requerimientos que exigía el cliente, se comenzó con la primera etapa del diseño de un esquema desnormalizado para la sección de ventas de la empresa.

Como se mencionó anteriormente, en consonancia con la metodología de trabajo adoptada para el proyecto, las etapas de diseño fueron tres, una para cada conjunto de tableros. Se trabajó secuencialmente con cada una de ellas.

En cada etapa de diseño, los miembros del equipo agruparon en un documento compartido los campos que creían necesarios para la realización del tablero en cuestión. Una vez finalizada dicha tarea, se realizaron reuniones virtuales para repasar los procedimientos almacenados de la empresa que eran los utilizados para generar sus reportes manualmente. Esto también servía de ayuda para ver qué operaciones se deben hacer para replicarlas en los campos calculados.

Teniendo armado el esquema, el paso siguiente sería garantizar que los campos del diseño final sean únicamente los necesarios. Se tienen en consideración los valores

que se necesitan para cálculos, comparaciones, muestras y la implementación de filtros.

## Metodología de capacitación y creación de tableros

La capacitación de los integrantes del equipo, previo a comenzar a desarrollar los tableros, fue de suma importancia al comenzar a trabajar con una tecnología hasta el momento desconocida. Se tomó un curso que abarcaba los conceptos básicos de la herramienta. Además, se tuvo una reunión con un exalumno del director que brindó su aporte para un mejor entendimiento del alcance y poder de Power BI. Como complemento, y ante cualquier duda, siempre se recurrió a la documentación oficial.

Posteriormente, al comenzar por el primer tablero, se encontró a la herramienta muy amigable e intuitiva, donde el trabajo de importar los datos desde SQL Server era muy sencillo. La mayor complejidad en la creación de tableros estuvo en el desarrollo de las distintas “medidas”. Estas se utilizan principalmente en tablas y gráficos para realizar cálculos y análisis sobre un conjunto de datos. Las medidas en Power BI pueden realizar una variedad de operaciones, como sumar valores, calcular promedios, contar elementos, aplicar filtros y realizar cálculos matemáticos más complejos. Además, pueden basarse en columnas específicas dentro de las tablas de datos, o pueden combinar múltiples columnas y tablas utilizando funciones de agregación y cálculo. Son lo más cercano a realizar una función en programación, donde se adapta al gráfico que la contenga y los filtros que se le apliquen para arrojar los resultados inmediatamente.

También se hizo uso de tablas auxiliares que no eran parte del esquema desnormalizado propuesto, pero que tenían información necesaria para poder completar los reportes. Por ejemplo, una tabla con los objetivos de la empresa por vendedor para cada sucursal. Esos datos no debían ser incluidos en el esquema principal.

Las etapas de desarrollo de los tableros fueron las que mayor retroalimentación constante requirieron, ya que jugaba un rol importante el grado de exactitud que requería la empresa para cada reporte, además de pedidos de cambio de diseño de

uno o varios elementos de los tableros, consejos acerca del color y forma en la que se visualizaban los datos, etcétera. Este *feedback* se realizaba mayormente por correo electrónico, ya que a fin de cada día, luego de trabajar en los tableros, el equipo subía las actualizaciones al servicio de Power BI que se accede *online*. Además, cada dos semanas se realizaba una llamada virtual para revisión de los tableros, cambios y sugerencias. En estas llamadas participaban tanto los referentes funcionales como todo el equipo.

## Metodología de implementación: extracción de datos

Esta etapa comenzó recién después de terminar los tableros de las secciones ventas y cobranzas, pese a estar previsto, según la estimación inicial, para antes de la creación de los mismos. Esto se debe a que, antes de crear los tableros, el equipo no estaba del todo familiarizado con la herramienta. Además, como desde la empresa todavía no habían abonado la licencia de Power BI, se optó por comenzar directamente con la creación de tableros y dejar la extracción y actualización de datos para después, ya que para esto último se requería la licencia del producto. Una vez que la empresa pagó la licencia, se pudo implementar la actualización incremental de los datos en Power BI. La documentación de la herramienta fue un factor clave para poder avanzar con este proceso.

Para la extracción diaria de los datos, se necesitaba realizar un proceso automatizado que actúe tanto sobre la base de datos, actualizando las tablas desnormalizadas, como también en Power BI actualizando el *dataset*. Lógicamente, no tenía sentido cargar los datos históricos de la empresa a diario, sino sólo los del último día.

Para ello, se decidió implementar: por el lado de la base de datos, un *cron job*, con un *script* de "Insert" a las tablas, ejecutado en la máquina de la empresa que dispara diariamente por la noche todos los *stored procedures* de actualización de su base de datos; y por el lado de Power BI, se usó la herramienta "Actualización Incremental", que permite programar el tiempo entre cada actualización y elegir cuántos días previos de datos incluir en el *dataset*. Entonces, cada día primero se insertan las entradas que corresponden a todas las transacciones del día anterior en las tablas del esquema desnormalizado. Luego, en Power BI, se importan las tablas

del esquema desnormalizado ya actualizado, ya que la herramienta permite importar directamente tablas desde SQL Server.

Cabe aclarar que buena parte del *script* de inserción de registros a las tablas de la base de datos se desarrolló antes de comenzar con los tableros. De esta manera, se pudo probar ejecutarlos en períodos cortos, de meses o días, y ver que se estuvieran insertando los datos deseados. Luego de cargar las tablas con datos de algunos meses, desde Power BI, al actualizar “manualmente” el *dataset*, este se cargaría con los registros de las tablas. Así, se logró tener datos de períodos cortos para las pruebas de consistencia.

## Metodología de pruebas de consistencia

Para las comparaciones, los referentes funcionales nos compartieron una carpeta en Google Drive, donde cargaban todos los reportes históricos realizados. Antes del final de cada etapa de desarrollo de los tableros, se hicieron comparaciones entre los resultados de los reportes realizados en Power BI y los generados por la empresa en Excel.

Estas comparaciones consistían en analizar, según el margen de error que los referentes funcionales consideraran, si los resultados eran precisos o no. De no serlos, se analizaban errores de implementación, tanto en el *stored procedure* para extracción de datos como en la codificación de las “medidas” en Power BI. La mayoría de los problemas encontrados se debieron a errores en la interpretación de conceptos como margen, MUP, intereses e impuestos. También fue importante diferenciar campos nulos de iguales a cero, ya que la empresa nos comentó que dicha diferencia es intencional pues tienen significados distintos.

A la hora de comparar los reportes, se tuvo en cuenta tres tipos de desvíos posibles:

- Sin desvío: los datos comparados entre reportes manuales y de Power BI dan exactamente como se esperaba, sin diferencia alguna.
- Con desvío conocido: se trata de los casos donde no coincide el resultado esperado con el obtenido en los reportes de la herramienta, pero se conoce su origen.

- Con desvío desconocido: se trata de los casos donde no coincide el resultado esperado con el obtenido, pero el origen no fue hallado y no se sabe su procedencia.

En la clasificación de “con desvío conocido”, entran los errores en la generación manual de los reportes por parte de la empresa. Por ejemplo, la comparación de los objetivos de las sucursales entre su base de datos y sus reportes no son correctas, hay un error de generación de su reporte. También se observó inconsistencia en la comparación de sus reportes mes a mes, por ejemplo en la facturación mensual entre dos meses consecutivos. El reporte de ventas hace referencia al total de la facturación del mes anterior. Si se compara ese total, con el total del reporte del mes anterior (en ese caso, mes actual), hay diferencias. Por dicho motivo, se determinó que la referencia al mes anterior en la creación de sus reportes está mal implementada.

Con respecto a la clasificación “con desvío desconocido”, únicamente se observó en el caso de los tableros de cobranzas. Los errores surgen al calcular los montos de todas las cuotas no pagadas por los clientes hasta el momento del corte, que sería la fecha seleccionada en el tablero por parte del usuario. No se pudo encontrar la fuente del error, pero los referentes funcionales dieron la aprobación de los tableros de todas formas ya que el desvío era porcentualmente muy bajo.

Se verá reflejada la corrección de los desvíos por iteración en los siguientes gráficos.

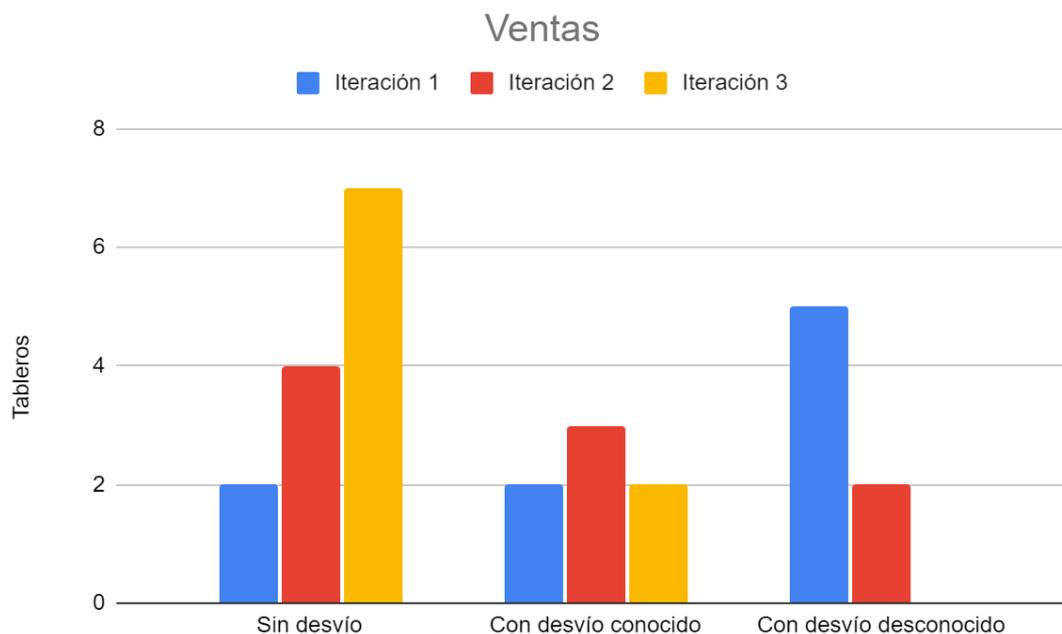


Figura 2: Iteraciones de corrección de desvíos en sección ventas.

Los desvíos conocidos en la última iteración representan los errores detectados en los tableros de la empresa ya mencionados, las comparaciones con meses anteriores y los objetivos de las sucursales. Ambos nacen de una extracción errónea de datos por su parte.

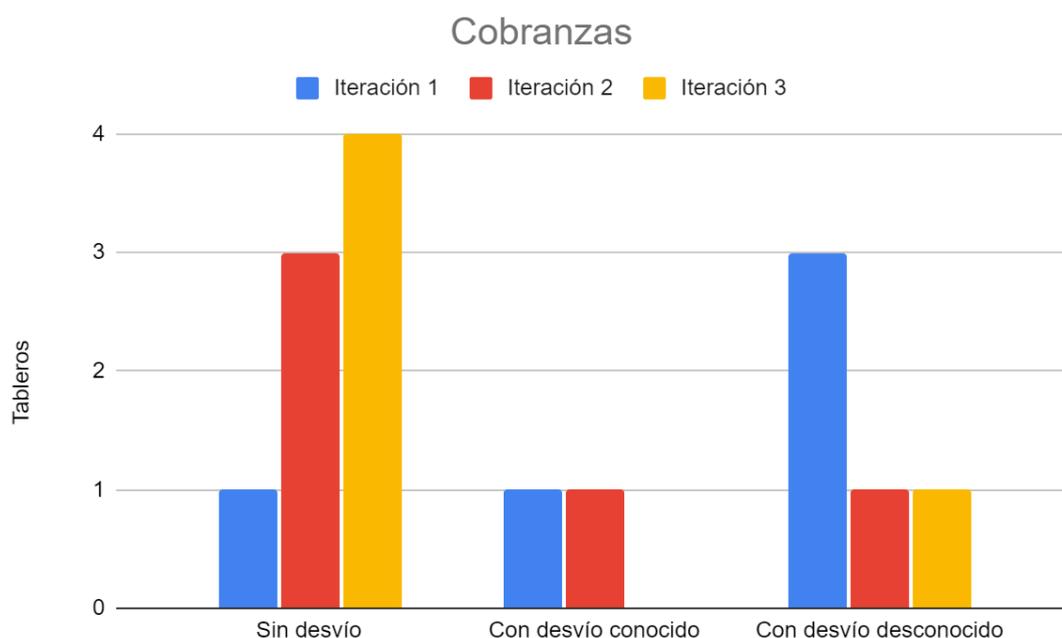


Figura 3: Iteraciones de corrección de desvíos en sección ventas.

En este caso, el desvío desconocido corresponde a las diferencias ínfimas mencionadas en el cálculo de montos de todas las cuotas no pagadas.

# Producto

## Producto obtenido

El producto desarrollado consta de un conjunto de tableros de Power BI, compuesto por un conjunto de medidas, permisos de acceso y procesos automatizados de extracción de datos a partir de un esquema de base de datos desnormalizado, el cual también se actualiza periódicamente, sin necesidad de operación humana.

El acceso por parte de los usuarios se hace a través de Power BI Service, *online*. Una vez ingresado con el usuario y contraseña de una cuenta válida de la empresa, se podrán observar los tableros según los permisos que tenga configurados.

Cabe destacar que los nombres de las sucursales, los nombres de los vendedores y todos los datos tanto de ventas como de cobranzas son ficticios, con el fin de respetar el convenio de confidencialidad que existe con la empresa.

## Ventas

- Filtro de selección de fecha con mes y año.
- Al seleccionar una sucursal cualquiera, los demás gráficos se adaptan al mismo y se actualizan.

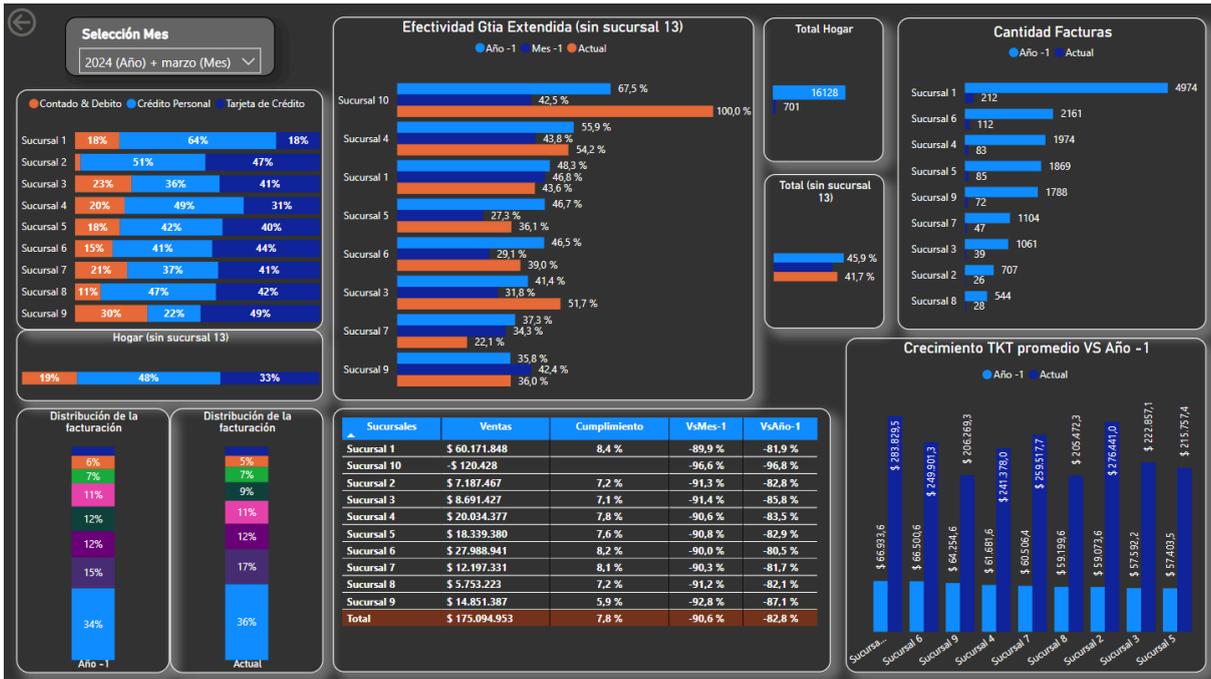


Figura 4: Tablero de gerencia, sector hogar.

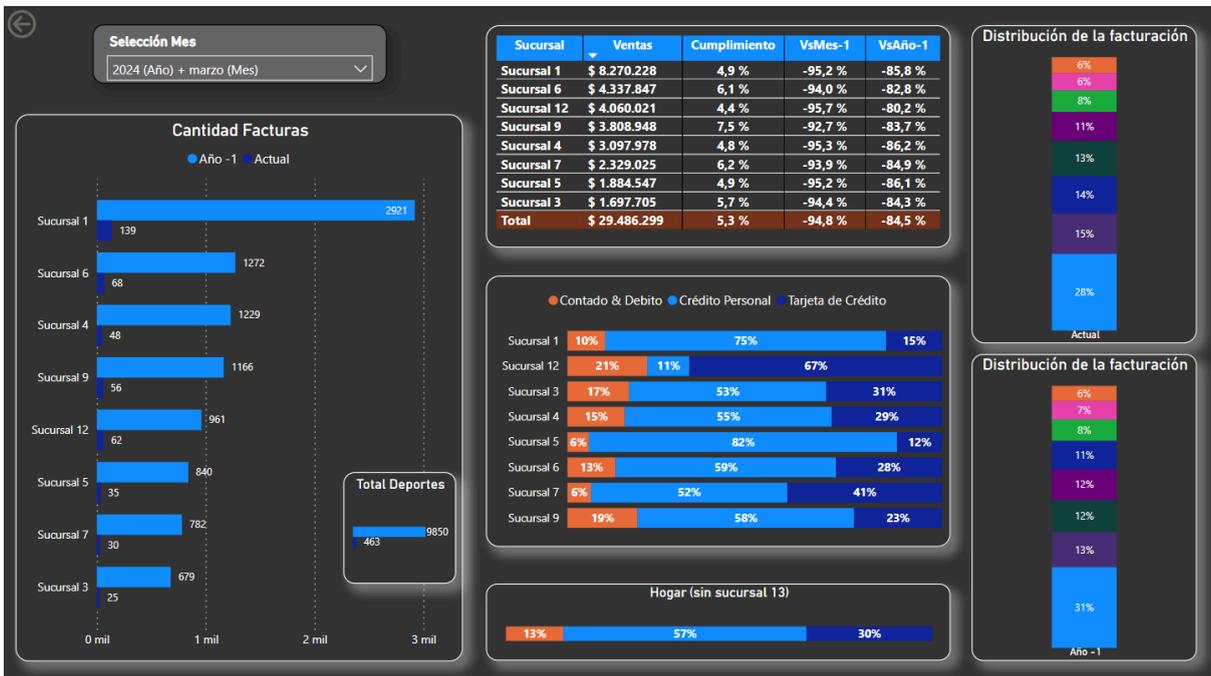


Figura 5: Tablero de gerencia, sector deporte.



Figura 6: Tablero de gerencia, ambos sectores de la compañía.

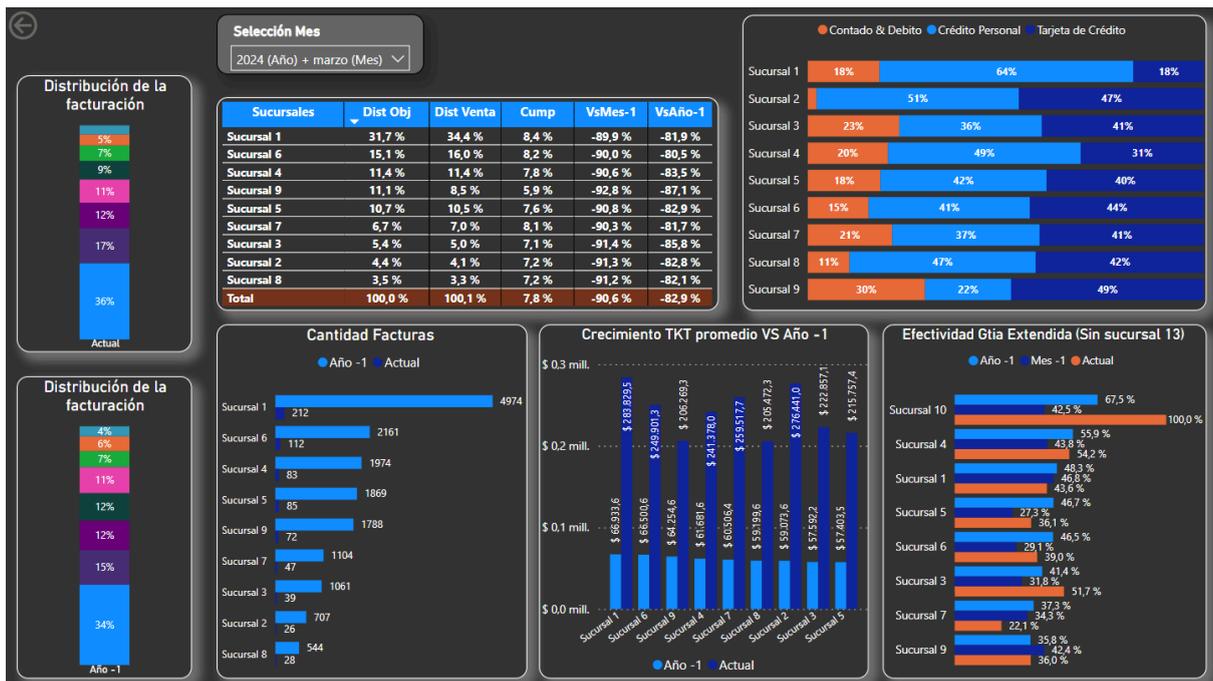


Figura 7: Tablero para gerentes de sucursales, ambos sectores de la compañía.

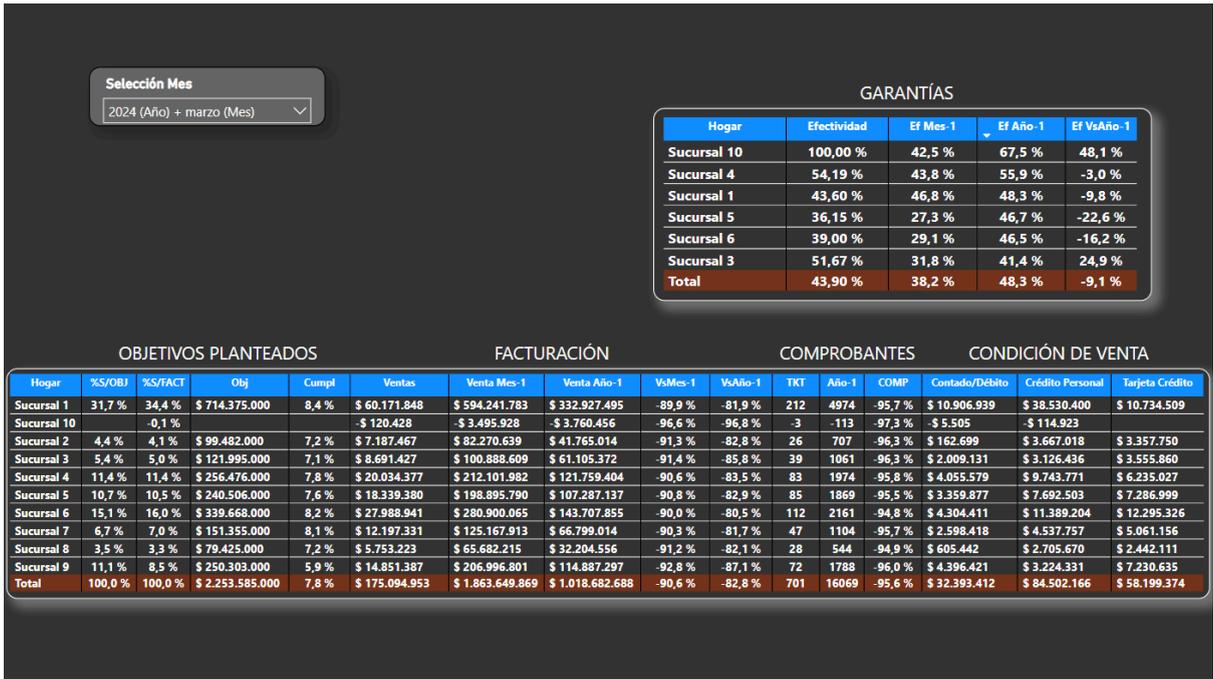


Figura 8: Tablero de información, sector hogar.

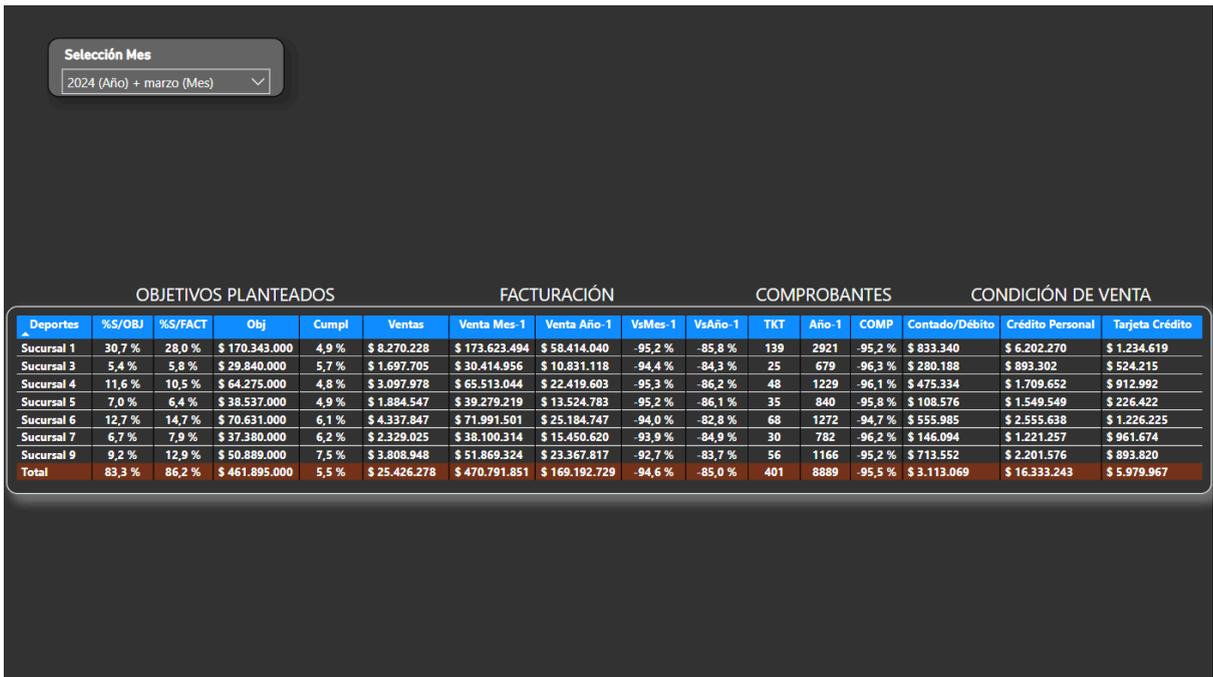


Figura 9: Tablero de información, sector deporte.

Selección Mes  
2024 (Año) + marzo (Mes)

OBJETIVOS PLANTEADOS					FACTURACIÓN				COMPROBANTES			CONDICIÓN DE VENTA			
Total Compañía	%S/OBJ	%S/FACT	Obj	Cumpl	Ventas	Venta Mes-1	Venta Año-1	VsMes-1	VsAño-1	TKT	Año-1	COMP	Contado/Débito	Crédito Personal	TarjetaCrédito
Sucursal 6	14,61 %	15,8 %	\$ 410.299.000	7,9 %	\$ 32.326.788	\$ 352.891.566	\$ 168.892.602	-90,8 %	-80,9 %	180	3433	-94,8 %	\$ 4.860.396	\$ 13.944.841	\$ 13.521.551
Sucursal 1	31,50 %	33,5 %	\$ 884.718.000	7,7 %	\$ 68.442.076	\$ 767.865.277	\$ 391.341.535	-91,1 %	-82,5 %	351	7896	-95,6 %	\$ 11.740.278	\$ 44.732.670	\$ 11.969.128
Sucursal 9	10,72 %	9,1 %	\$ 301.192.000	6,2 %	\$ 18.660.335	\$ 258.866.125	\$ 138.255.114	-92,8 %	-86,5 %	128	2954	-95,7 %	\$ 5.109.973	\$ 5.425.907	\$ 8.124.455
Sucursal 5	9,93 %	9,9 %	\$ 279.043.000	7,2 %	\$ 20.223.927	\$ 238.175.009	\$ 120.811.920	-91,5 %	-83,3 %	120	2709	-95,6 %	\$ 3.468.454	\$ 9.242.052	\$ 7.513.421
Sucursal 4	11,42 %	11,3 %	\$ 320.751.000	7,2 %	\$ 23.132.355	\$ 277.615.026	\$ 144.179.007	-91,7 %	-84,0 %	131	3202	-95,9 %	\$ 4.530.913	\$ 11.453.423	\$ 7.148.019
Sucursal 7	6,72 %	7,1 %	\$ 188.735.000	7,7 %	\$ 14.526.356	\$ 163.268.228	\$ 82.249.634	-91,1 %	-82,3 %	77	1886	-95,9 %	\$ 2.744.512	\$ 5.759.014	\$ 6.022.830
Sucursal 3	5,41 %	5,1 %	\$ 151.835.000	6,8 %	\$ 10.389.132	\$ 131.303.564	\$ 71.936.490	-92,1 %	-85,6 %	64	1740	-96,3 %	\$ 2.289.319	\$ 4.019.738	\$ 4.080.075
Sucursal 12	3,54 %	3,5 %	\$ 99.482.000	7,2 %	\$ 7.187.467	\$ 82.270.639	\$ 41.765.014	-91,3 %	-82,8 %	26	707	-96,3 %	\$ 162.699	\$ 3.667.018	\$ 3.357.750
Sucursal 12	3,32 %	2,0 %	\$ 93.282.000	4,4 %	\$ 4.060.021	\$ 94.949.645	\$ 20.564.540	-95,7 %	-80,3 %	62	960	-93,5 %	\$ 864.166	\$ 458.172	\$ 2.737.683
Sucursal 8	2,83 %	2,8 %	\$ 79.425.000	7,2 %	\$ 5.753.223	\$ 65.682.215	\$ 32.204.556	-91,2 %	-82,1 %	28	544	-94,9 %	\$ 605.442	\$ 2.705.670	\$ 2.442.111
<b>Total</b>	<b>100,00 %</b>	<b>100,1 %</b>	<b>\$ 2.808.762.000</b>	<b>7,3 %</b>	<b>\$ 204.701.680</b>	<b>\$ 2.432.887.292</b>	<b>\$ 1.212.200.413</b>	<b>-91,6 %</b>	<b>-83,1 %</b>	<b>1167</b>	<b>26031</b>	<b>-95,5 %</b>	<b>\$ 36.376.152</b>	<b>\$ 101.408.504</b>	<b>\$ 66.917.024</b>

Figura 10: Tablero de información, ambos sectores de la compañía.

Selección Mes  
2024 (Año) + marzo (Mes)

Sucursal	DEPORTES	HOGAR	TERCEROS	Total
Sucursal 1	\$ 170.343.000	\$ 714.375.000	\$ 0	\$ 884.718.000
Sucursal 6	\$ 70.631.000	\$ 339.668.000	\$ 0	\$ 410.299.000
Sucursal 4	\$ 64.275.000	\$ 256.476.000	\$ 0	\$ 320.751.000
Sucursal 9	\$ 50.889.000	\$ 250.303.000	\$ 0	\$ 301.192.000
Sucursal 5	\$ 38.537.000	\$ 240.506.000	\$ 0	\$ 279.043.000
Sucursal 7	\$ 37.380.000	\$ 151.355.000	\$ 0	\$ 188.735.000
Sucursal 3	\$ 29.840.000	\$ 121.995.000	\$ 0	\$ 151.835.000
Sucursal 2	\$ 99.482.000	\$ 0	\$ 0	\$ 99.482.000
<b>Total</b>	<b>\$ 461.895.000</b>	<b>\$ 2.174.160.000</b>	<b>\$ 0</b>	<b>\$ 2.636.055.000</b>

Vendedor	%	Part en \$	Pesos	Cumpl	Total	Cred Pers	Tarjeta	Contado	Gtia Ext	Fact
Juan Sosa	22,00 %	18,27 %	\$ 52.911.320	6,3 %	\$ 3.343.147	41,3 %	57,5 %	1,2 %	31,9 %	18
Horacio González	18,00 %	22,29 %	\$ 43.291.080	9,4 %	\$ 4.077.679	41,9 %	43,9 %	14,3 %	80,2 %	17
Gustavo Benítez	22,00 %	26,13 %	\$ 52.911.320	9,0 %	\$ 4.780.661	35,4 %	44,6 %	20,0 %	31,9 %	24
Emanuel Díaz	17,00 %	17,38 %	\$ 40.886.020	7,8 %	\$ 3.180.280	21,4 %	29,8 %	48,8 %	23,1 %	11
Agustín Pérez	20,00 %	16,52 %	\$ 48.101.200	6,3 %	\$ 3.022.227	73,4 %	16,0 %	10,6 %		16
Otros	0,00 %	0,00 %	-	0,00 %	\$ 0,00	-	-	-		0

Figura 11: Tablero de información sobre los vendedores de cada sucursal.

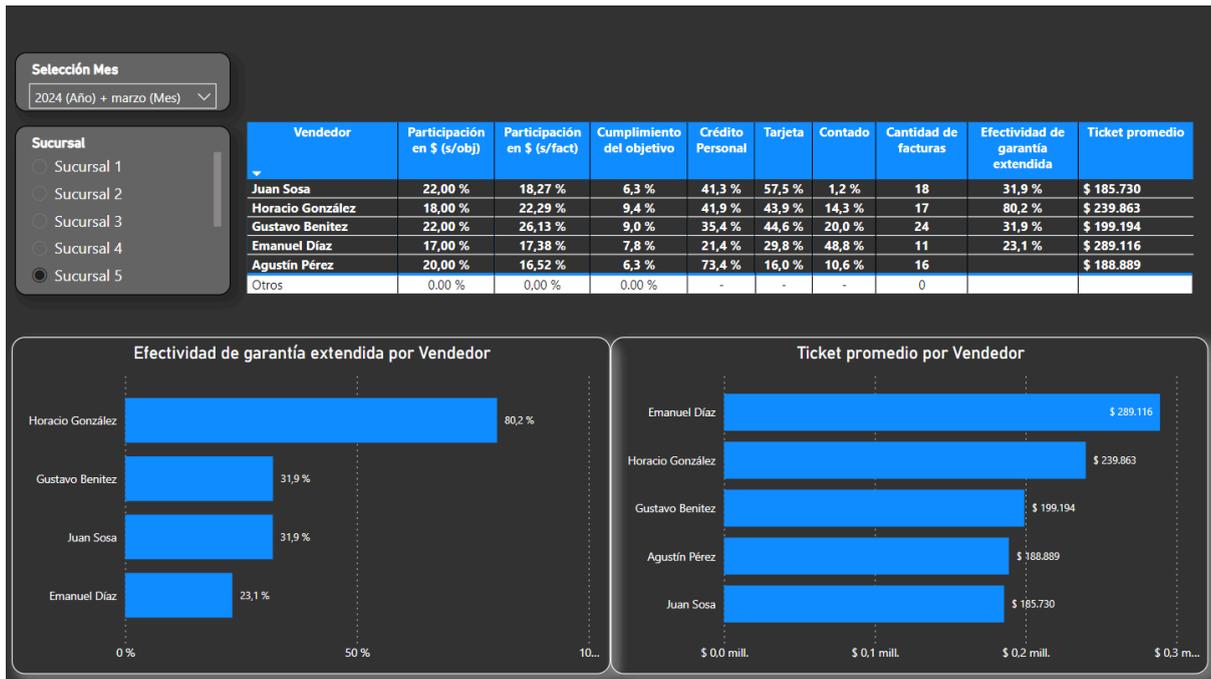


Figura 12: Tablero para gerentes de sucursales sobre vendedores.

## Cobranzas

- Filtro de selección de fecha con día, mes y año.
- Filtro de año para los gráficos.
- Filtro de sucursal y categoría para obtener los datos globales históricos.

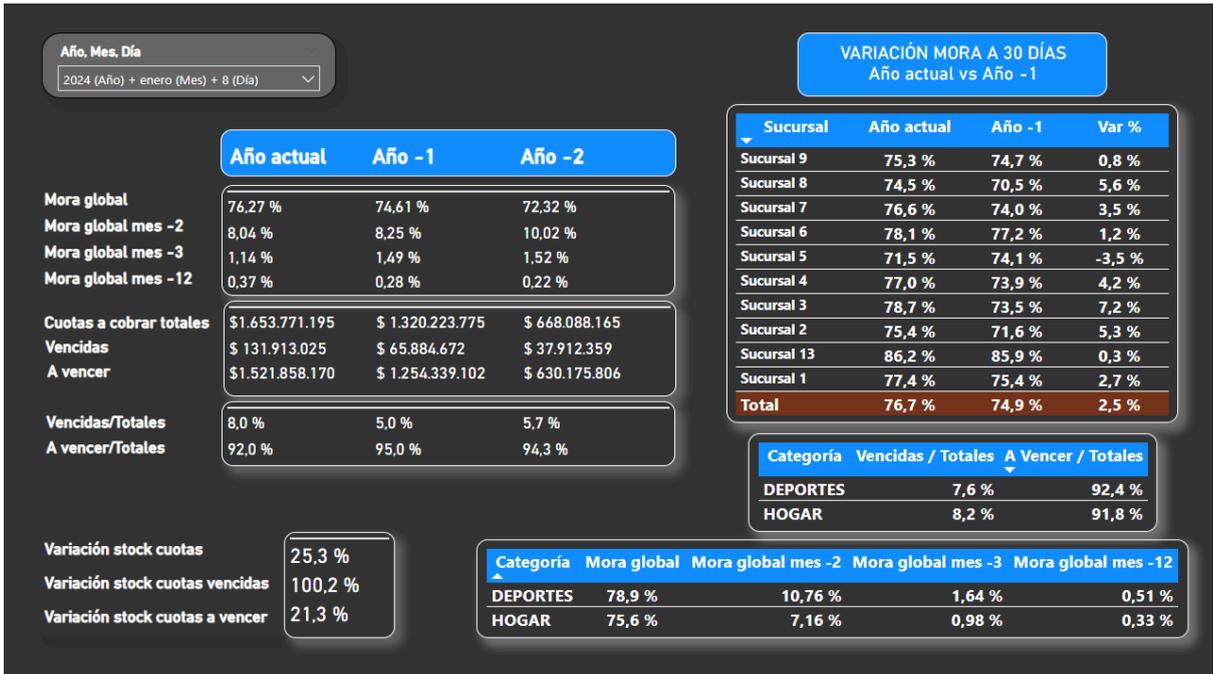


Figura 13: Tablero principal con la información relevante de mora.

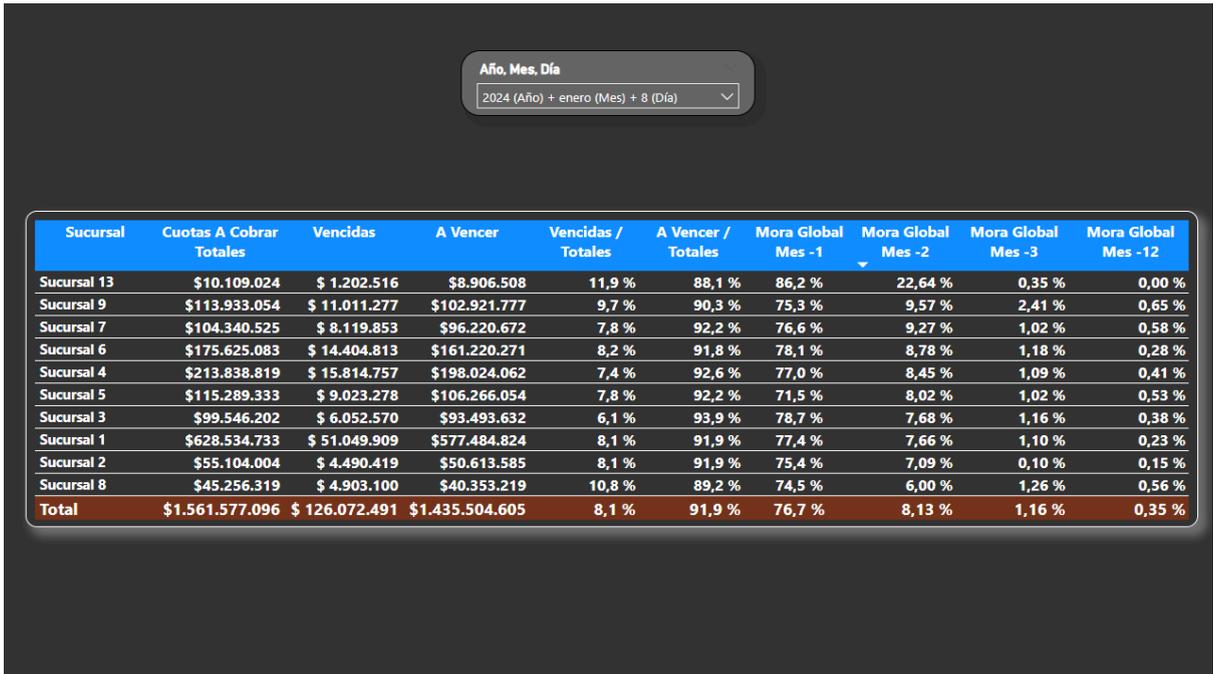


Figura 14: Tablero de información sobre cuotas vencidas/a vencer.

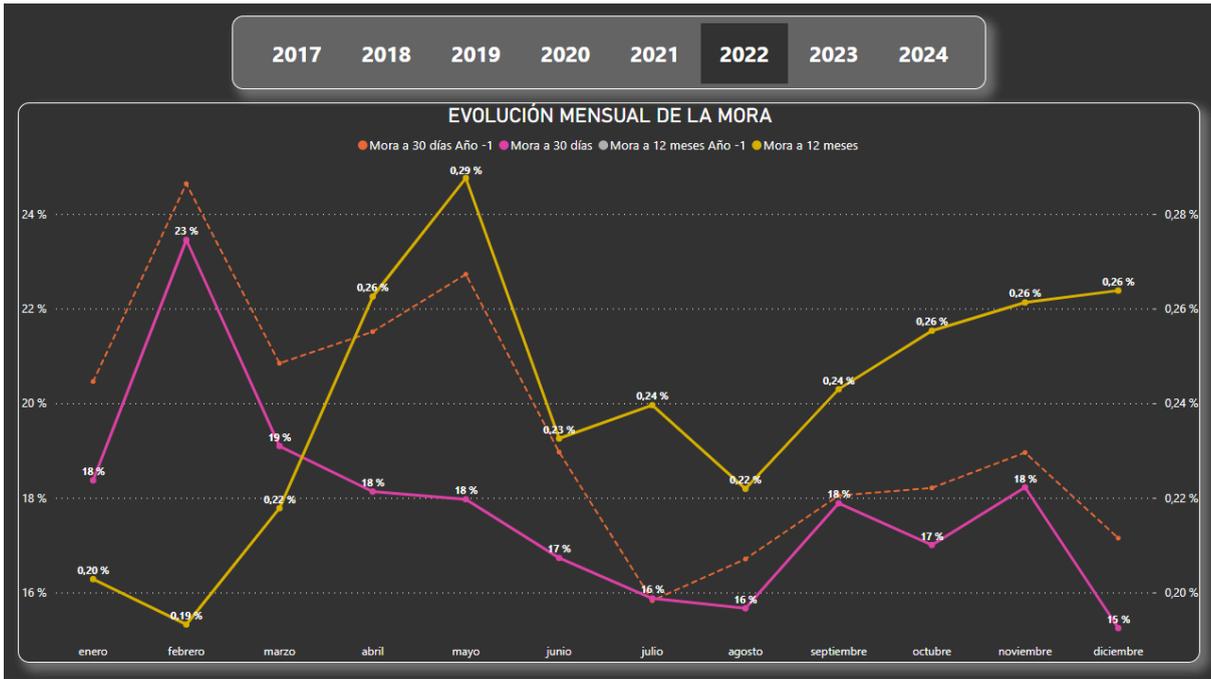


Figura 15: Tablero con gráfico de evolución mensual de la mora con filtro de año.

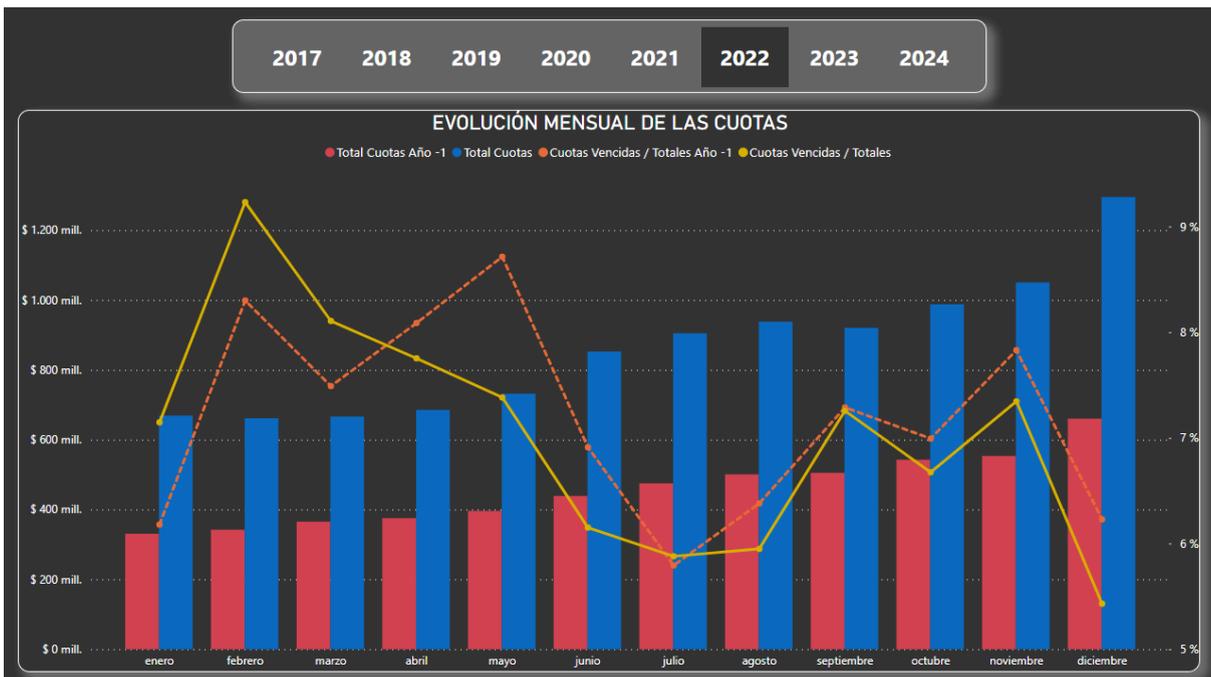


Figura 16: Tablero con gráfico de evolución mensual de las cuotas con filtro de año.

Año	Cobranza Inicial	Cobrado	Saldo	% Sobre Cza Inicial
2021	27.758.239,5	27.626.442,5	131.797,01	0,474803
2022				
enero	3.389.085,0	3.352.181,0	36.904,00	1,088907
febrero	3.187.431,8	3.149.635,8	37.796,00	1,185782
marzo	3.422.997,8	3.388.598,8	34.399,00	1,004938
abril	3.438.082,9	3.412.563,9	25.519,00	0,742245
mayo	3.614.879,5	3.608.556,5	6.323,00	0,174916
junio	3.681.150,7	3.678.190,7	2.960,00	0,080410
julio	4.526.735,0	4.522.548,0	4.187,00	0,092495
agosto	4.892.187,8	4.871.254,8	20.933,00	0,427886
septiembre	5.304.535,8	5.276.695,8	27.840,00	0,524834
octubre	5.554.176,8	5.527.685,8	26.491,00	0,476956
noviembre	6.183.009,4	6.150.143,4	32.866,00	0,531553
diciembre	6.296.125,2	6.262.236,2	33.889,00	0,538252
2023	113.545.183,4	112.767.108,0	778.075,38	0,685256
2024	70.485.343,0	31.315.236,8	39.170.106,29	55,571988

Figura 17: Tablero de información con datos clave para calcular la mora.

## Trabajos futuros

Al momento de definir el alcance del proyecto, la generación de reportes de varias secciones de la empresa quedó al margen, principalmente porque extendería considerablemente el tiempo de desarrollo. Sumado a esto, la falta de datos en la sección de gastos hizo que se postergue la generación de reportes de esta sección para un futuro.

Las tareas a realizar en un posible trabajo futuro son:

- Generación de reportes de “gastos”.
- Extensión de las funcionalidades y creación de nuevos tableros en las secciones ya trabajadas.

## Soporte

Luego de la finalización del proyecto, se acordó con la empresa estar disponibles para brindar soporte durante el primer mes de puesta en producción. Esto incluye la monitorización de las tareas automáticas establecidas para garantizar su correcto funcionamiento y eficiencia. El equipo estará dispuesto a responder cualquier

consulta que puedan tener los referentes funcionales o cualquier usuario de los tableros, proporcionando orientación y resolución de problemas según sea necesario.

# Memoria del proyecto

## Participación de los referentes

Durante el transcurso del proyecto, se estableció un flujo constante de comunicación bidireccional con los referentes funcionales de la empresa. Esta interacción no solo se limitó a la entrega de las primeras instrucciones para comenzar el trabajo, sino que también se caracterizó por un genuino interés en apoyar el progreso y el éxito del proyecto. Desde el inicio, los referentes funcionales se mostraron abiertos a escuchar las propuestas del equipo y a ofrecer su experiencia y conocimiento para enriquecer el desarrollo de la solución.

Con el paso del tiempo, se fueron organizando reuniones para revisar el avance del proyecto y discutir posibles ajustes o mejoras. Estas sesiones se convirtieron en un espacio óptimo para el intercambio de ideas y la resolución de dudas o inquietudes que pudieran surgir en el equipo de implementación. Los referentes funcionales siempre se mostraron receptivos a las sugerencias del equipo y brindaron valiosos comentarios y recomendaciones que contribuyeron significativamente al éxito del proyecto.

Su colaboración se extendió a lo largo de todo el proceso, desde la recopilación de requerimientos hasta la puesta en marcha de la solución. En cada etapa, contamos con su respaldo, lo que nos permitió avanzar de manera eficiente y superar los obstáculos que pudieran surgir en el camino, agilizando los tiempos de desarrollo de la solución. Su compromiso y dedicación se reflejaron en la calidad del trabajo realizado y en los resultados obtenidos.

## Cumplimiento de objetivos

### Objetivo principal

**Se busca implementar soluciones tecnológicas que permitan automatizar la generación de informes empresariales, optimizando los procesos involucrados y agilizando la entrega de información crucial para la toma de**

**decisiones, con la mayor claridad posible.** Con la inserción de la herramienta de inteligencia empresarial, se pudo cumplir con el objetivo. El acceso a los datos no es más un problema de la empresa, ya que se crearon los tableros y se dejaron plantillas con medidas de usos varios. Los reportes son claros y precisos; muestran la información necesaria eficazmente.

**Una herramienta fundamental para lograrlo es la creación de una central única de datos, que evitan la duplicación y su inconsistencia, eliminando redundancias para garantizar la integridad de los datos.** Se consolidaron los esquemas desnormalizados; mediante la implementación de una lógica en el proceso ETL, se aseguró que los datos provinieran de las fuentes correctas tras el análisis de la base de datos de la empresa. Este enfoque fue clave para prevenir la duplicación e inconsistencia. Estos esquemas poseen información relevante para la realización de los tableros y tienen toda la información que la gerencia de la empresa necesita visualizar.

### Objetivos secundarios

**Liberar recursos humanos para tareas estratégicas.** Se automatizaron los procesos de actualización de tableros. En consecuencia, la creación manual de los reportes, mensual o semanal, ya no es necesaria. La generación de reportes, ahora, es parte de un proceso que se realiza de manera automática diariamente. De esta manera, se liberaron los recursos humanos que estaban encargados de dicha tarea.

**Se busca simplificar y agilizar el acceso a la información empresarial, proporcionar a los usuarios acceso rápido y fácil a información empresarial relevante.** El acceso a la información se hace directamente ingresando a la aplicación web Power BI Service. Tan solo accediendo a la aplicación con el usuario de la empresa correspondiente, se podrán visualizar los tableros sobre los que ese usuario tenga permisos. Podrá aplicar filtros a gusto para visualizar el contenido que desee en cada objeto visual de cada tablero.

**Encontrar errores, si es que los hay, en el sistema de base de datos actual o en el proceso de generación de reportes.** Se lograron encontrar errores en la generación manual de reportes. Uno de los principales ejemplos es el cálculo de los objetivos de las sucursales. La información de los reportes de la empresa no

coincidía con la información que se encuentra en la base de datos, por lo que se concluyó que existía un error de extracción de los datos. Otro error encontrado fue la inconsistencia en la facturación mensual entre meses consecutivos. Todos los errores fueron contemplados y corregidos en los reportes de Power BI, y también se notificó a los referentes funcionales acerca de ello.

**Por último, se busca que el tiempo de presentación sea el más acotado posible debido a la obsolescencia y ralentización que genera el proceso manual actual ya mencionado.** Dado el poder de la herramienta elegida, basta con que el usuario seleccione los datos de qué fecha desea visualizar, e inmediatamente se aplicará el filtro correspondiente. La velocidad en que impactan los filtros en los reportes no solo se debe a la herramienta en sí, sino también a la eficiencia de las medidas implementadas por el equipo.

## Análisis FODA

Después de la implementación del proyecto, se realizó un análisis del FODA planteado inicialmente. Se observaron los siguientes cambios al respecto:

- “La empresa dispone de datos históricos, acceso libre a los mismos y documentación de los diseños de registros”: esta fortaleza fue parcialmente cierta, ya que de la sección gastos, en realidad, no se disponía de datos históricos como se pensó inicialmente.
- “La resistencia al cambio puede surgir entre el personal ante la introducción de nuevas tecnologías y procesos”: la introducción de Power BI no representó una debilidad como tal, ya que no surgió resistencia al cambio entre el personal.
- “Puede ser necesario invertir en programas de capacitación adicionales para asegurar una adopción efectiva de las herramientas de Business Intelligence”: no fue necesaria la inversión en capacitación adicional.
- “Requiere inversión inicial para implementar la solución (costo de licencias)”: esta debilidad no tuvo incidencia; la empresa decidió afrontar el costo de licencias sin problemas.

## Riesgos

En cuanto al riesgo R03 mencionado en el análisis de riesgos del proyecto, vale la pena mencionar que uno de los integrantes del equipo estuvo de viaje por una semana. Sin embargo, esta situación se sorteó de manera satisfactoria gracias al plan de contingencia oportunamente planteado.

Los demás riesgos planteados no tuvieron lugar durante el transcurso del proyecto, ni tampoco surgieron riesgos no contemplados inicialmente.

## Planificación esperada vs ejecución

Inicialmente, el cronograma se planificó bajo un esquema en cascada, donde se consideraron nueve etapas con 14 horas de trabajo semanales por integrante. El proyecto tendría una duración de 23 semanas, llegando a un total de 900 horas de trabajo totales.

En la ejecución, no fue así. El trabajo no se desarrolló con la metodología supuesta, sino que, como ya se mencionó, se utilizó una metodología ágil que consistió en *sprints* de duración variable con *feedback* constante de los referentes funcionales. Se decidió llevar a cabo el proyecto de esta manera debido a la necesidad de la empresa de ver resultados lo más pronto posible. El equipo notó que, con una metodología en cascada, no iba a ser posible tener prototipos funcionales en corto tiempo para la empresa. También, fue una gran decisión ya que acotó mucho los tiempos de desarrollo de la solución. El enfocar toda la atención de a una sección por vez permitió detectar deficiencias o errores en la implementación de manera prematura. La experiencia previa con la primera sección (ventas) fue de gran ayuda para implementar la extracción de datos y los tableros de la segunda sección (cobranzas) mucho más rápido.

Además, la extracción de datos y realización de tableros de la sección “gastos” fueron suspendidas. Como se mencionó en apartados anteriores, la información que poseía la empresa en su base de datos era deficiente, y por ende la extracción de datos relevantes resultó imposible. Como no era algo fácil de solucionar a corto plazo, decidieron suspender el trabajo en esa sección. Esto se tradujo en una

reducción de horas de trabajo del proyecto; en este caso no por cambio en la metodología de trabajo, sino debido a una disminución en el alcance del proyecto. Sin embargo, los referentes funcionales mostraron su interés en darle continuidad al proyecto en un futuro, cuando ya tuvieran solucionado el problema relacionado a la falta de información, tanto de la sección “gastos” como de nuevas secciones. En el momento que se tomó esta decisión, ya se habían terminado de implementar los tableros de la segunda sección y se estaba trabajando sobre la actualización y extracción automatizada de ventas y cobranzas.

El inicio del proyecto tuvo lugar en la fecha prevista. Sin embargo, el desarrollo del proyecto demoró tres semanas de trabajo menos que las previstas. Esta disminución se debe tanto al cambio en la metodología de trabajo, como también a la reducción del alcance del proyecto, causada por la falta de información digitalizada de los procesos referidos a la sección “gastos”.

En total se dedicaron 803 horas a lo largo de 20 semanas. A continuación, se presenta la comparación entre el cronograma estimado con el real. Los números dentro de las celdas representan la cantidad total de horas que se trabajó en ese periodo de tiempo.

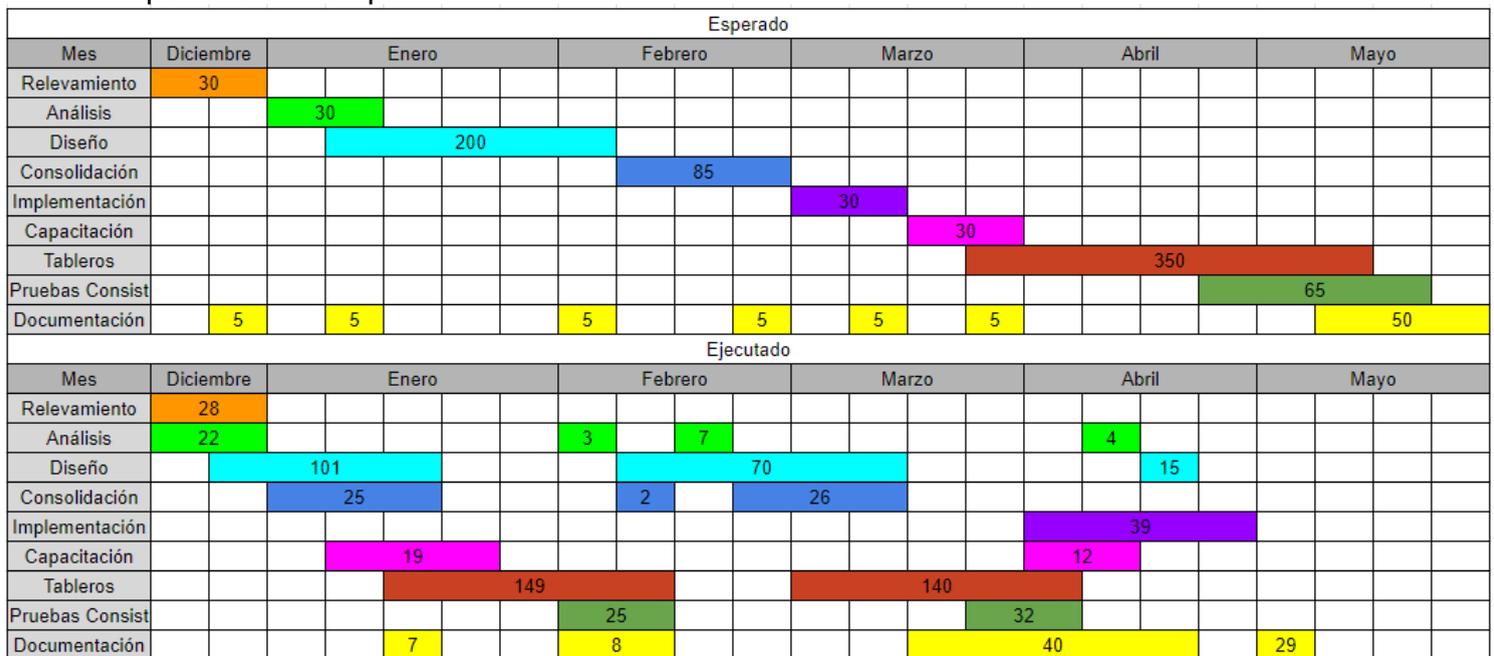


Figura 18: Diagramas de Gantt de planificación inicial estimada vs ejecución.

## Bitácora de trabajo

El proyecto comenzó con una reunión presencial donde se presentó el equipo, se anotaron los requerimientos y se revisaron los reportes actuales de la empresa. Se decidió utilizar Power BI en lugar de Pentaho debido a su facilidad de uso y mejor experiencia de usuario y se acordó crear un esquema de base de datos desnormalizado para mejorar el acceso a los datos. Se adoptó una metodología ágil con *sprints* y se definieron las secciones de la empresa a trabajar: ventas, cobranzas y gastos.

Durante las primeras semanas, el equipo trabajó en la consulta de extracción de datos para ventas. Se enfrentaron dudas conceptuales que se resolvieron en llamadas con el departamento de IT. Se realizaron ajustes para optimizar la consulta y se compararon datos con los reportes existentes, descubriendo diferencias significativas que se analizaron en reuniones virtuales con los referentes funcionales.

Se comenzaron a desarrollar los tableros en Power BI. Se realizaron reuniones periódicas con el departamento de IT para resolver dudas y realizar ajustes. A lo largo del proyecto, se mejoraron continuamente los tableros, con ajustes de medidas y objetos visuales.

Se comenzó a trabajar en la sección de cobranzas. A pesar de algunos problemas iniciales, se lograron superar y se desarrollaron tableros eficientes. Se coordinaron reuniones para presentar avances y realizar ajustes según las necesidades del cliente.

En las etapas finales, se implementó la automatización del proceso de actualización de datos en Power BI. Se configuraron actualizaciones periódicas y se realizó una breve capacitación para los referentes funcionales, quienes manifestaron interés en comprender el funcionamiento de la herramienta.

Finalmente, se verificó que las actualizaciones se ejecutaran correctamente y se completó el informe del proyecto. Se mantuvo una comunicación fluida con el soporte de IT para asegurar la correcta implementación y funcionamiento de los procesos automáticos. Se detalla la bitácora semanal en el Anexo II.

# Resumen gráfico de etapas

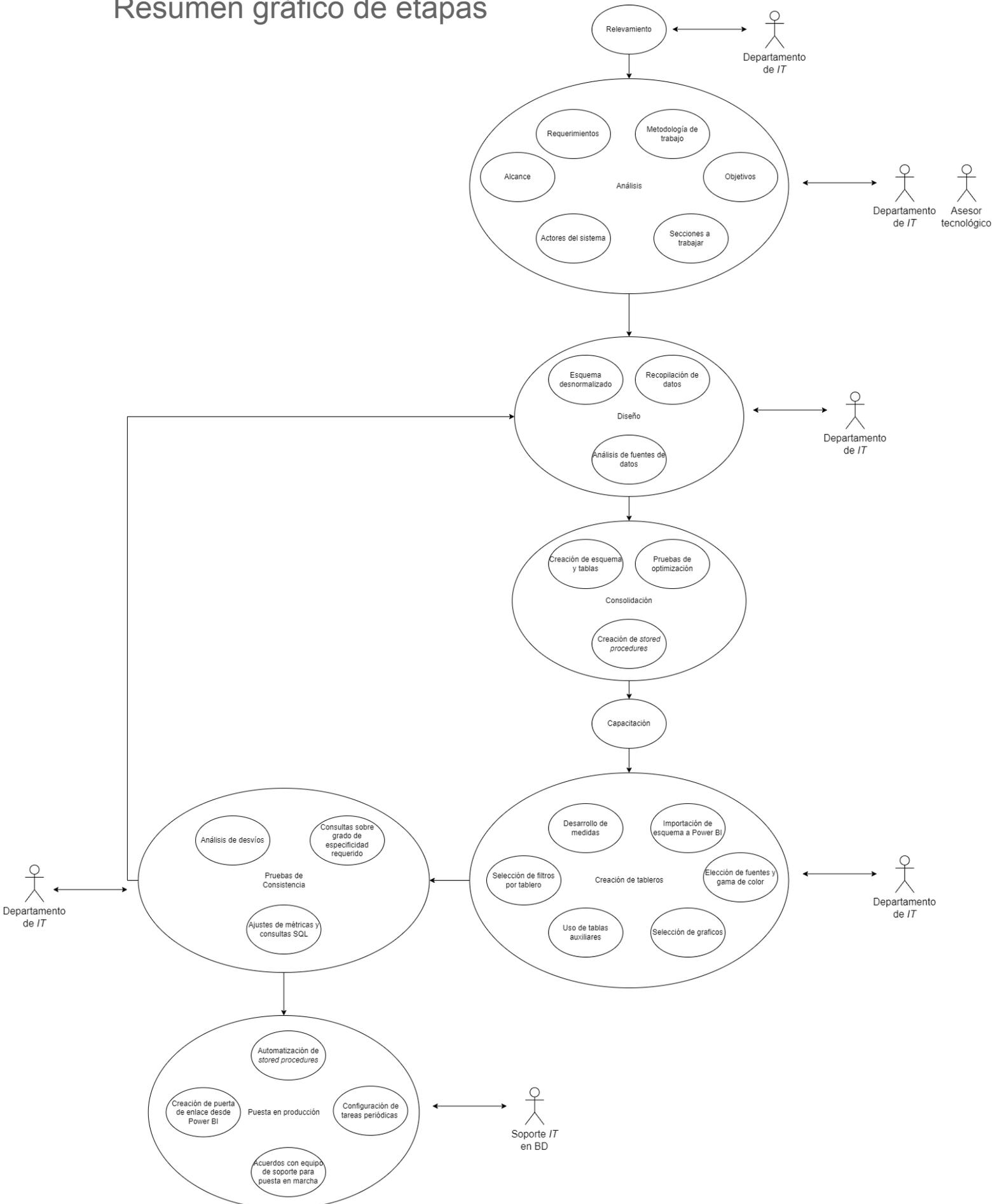


Figura 19: Resumen ilustrativo del flujo de trabajo.

## Comparación de tiempos

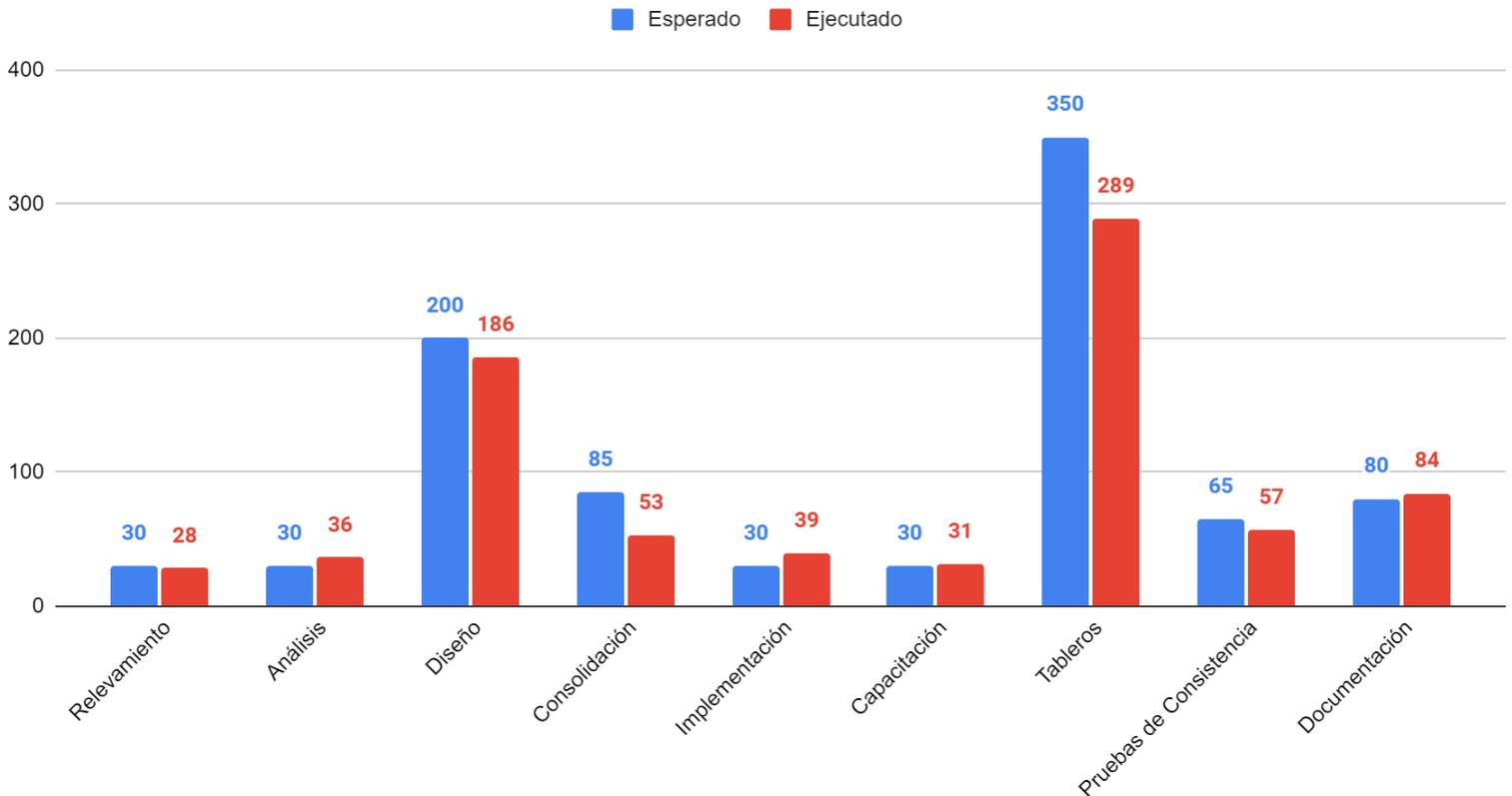


Figura 20: Comparativa de tiempos esperados vs ejecución en horas.

## División de tareas

En principio, como había una única tarea que era el armado del esquema de base de datos para la sección ventas, todo el equipo se enfocó en ello, analizando los reportes y armando las consultas SQL necesarias.

Una vez armado el esquema, el equipo completo pasó a la realización de las medidas y los tableros. Inicialmente, se plantearon en conjunto para definir la metodología interna con la que se iba a trabajar. Posteriormente, se decidió separar los distintos tableros para avanzar más rápidamente.

Cuando se estaba finalizando con la sección “ventas”, y se empezó con el armado de la sección “cobranzas”, las tareas se empezaron a dividir más. Esto se debe a que, mientras que algunos miembros se dedicaron al análisis y corrección de los

desvíos en el tablero de ventas, otros se dedicaron al análisis y desarrollo del esquema de cobranzas.

Una vez finalizados los tableros de ventas, todos los miembros del equipo se abocaron a trabajar tanto con el armado del esquema de cobranzas como con la realización de los tableros y sus medidas. En este caso, también se decidió separar los tableros entre los integrantes del equipo para avanzar de manera más rápida.

Cuando se estaba cerca de finalizar con la sección de cobranzas, se comenzó a trabajar con la automatización del proceso de actualización de datos en Power BI. Esto fue investigado principalmente por un miembro del equipo, mientras los otros terminaban los detalles pendientes en la sección de cobranzas. Luego, el equipo en conjunto se dedicó a la implementación de la automatización del proceso de actualización de datos. Primero, la referida al esquema de BI en la base de datos, y luego la referida a la actualización de datos en Power BI.

## Utilización de nuevas tecnologías

El proyecto ofrece una valiosa oportunidad de aprendizaje y desarrollo para los estudiantes. Los conocimientos aprendidos proporcionan una sólida base para comprender y trabajar con herramientas modernas de análisis de datos y bases de datos; dos áreas fundamentales en el campo de la informática.

En primer lugar, el diseño desnormalizado en SQL Server es un aspecto clave del proyecto. La desnormalización de la base de datos implica simplificar la estructura de la base de datos al reducir el número de tablas y eliminar las dependencias complejas entre ellas. Esto facilita la consulta y manipulación de datos, así como la generación de informes, ya que los datos relevantes están más cerca y son más accesibles. Se basa en aplicar técnicas de diseño de bases de datos que optimizan el rendimiento y la eficiencia de los sistemas de información empresarial.

Por otro lado, la creación de tableros en Power BI permitió a los estudiantes explorar la visualización de datos de una manera altamente interactiva y personalizable. A través de esta herramienta, se transforman datos complejos en visualizaciones claras y significativas que ayuden a los usuarios a comprender y analizar mejor la información empresarial. Los tableros ofrecen una amplia variedad de opciones de

visualización como gráficos simples, filtros de control interactivos, lo que permite desarrollar habilidades en presentación de datos de manera efectiva.

Además del aspecto técnico, este proyecto ofrece la oportunidad de adquirir experiencia práctica en trabajo en equipo, comunicación y resolución de problemas. Trabajar en conjunto en un proyecto de esta envergadura permite desarrollar habilidades de colaboración y gestión de proyectos, habilidades que son altamente valoradas en el mundo laboral.

Por último, en términos de su utilidad futura, los conocimientos adquiridos en este proyecto son transferibles y aplicables en una amplia gama de industrias y roles profesionales. La capacidad de diseñar y mantener bases de datos eficientes, así como la habilidad para crear visualizaciones de datos convincentes, son habilidades altamente demandadas en el mercado laboral actual. Además, el conocimiento de herramientas como Power BI y SQL Server abre puertas a oportunidades profesionales en el campo de la inteligencia empresarial, el análisis de datos y la consultoría tecnológica.

# Conclusiones

El objetivo principal del proyecto era implementar soluciones tecnológicas para automatizar la generación de reportes, un proceso que la empresa hacía manualmente. Esto fue logrado a través de la creación de un esquema de base de datos desnormalizado, junto con la creación de tableros y la automatización y optimización de procesos de actualización. El producto obtenido es el resultado de un arduo trabajo y colaboración entre el equipo del proyecto y los referentes funcionales de la empresa. Los tableros en Power BI proporcionan una plataforma visualmente atractiva y fácil de usar para analizar datos clave, mientras que el diseño desnormalizado en SQL Server garantiza una gestión eficiente y optimizada de la información.

Uno de los desafíos más significativos que enfrentó el equipo fue la planificación de tiempos en un proyecto de esta magnitud. Gracias a las metodologías de estimación de tiempos en proyectos de *software*, análisis y diseño de sistemas aprendidas durante la carrera, sumadas a la comunicación efectiva y compromiso entre los integrantes, se pudo llevar a cabo el proyecto dentro de los tiempos previstos. Cabe mencionar también que, pese a cambios en la metodología de trabajo, el equipo supo adaptarse sin que esto representara un problema respecto de la planificación inicial.

En conclusión, el proyecto fue llevado a cabo de manera exitosa. La experiencia y los conocimientos adquiridos son de gran utilidad para los estudiantes. Deja en evidencia la enorme importancia de la planificación y aplicación de metodologías pertinentes para la correcta ejecución de un proyecto de *software*. Además, provee a los estudiantes de habilidades prácticas en bases de datos y herramientas de BI, que son altamente demandadas en el mercado laboral actual.

# Apéndice

## Glosario

IT. En el contexto empresarial, se refiere a "Tecnología de la Información". Este término abarca todas las tecnologías, sistemas y procesos relacionados con la gestión, almacenamiento, transmisión y procesamiento de información en una organización. El departamento de IT, se encarga de administrar y mantener la infraestructura tecnológica de la empresa, incluyendo hardware, software, redes, bases de datos y sistemas de seguridad.

*Software*. Se refiere a los programas informáticos, conjuntos de instrucciones y datos que permiten a los dispositivos electrónicos realizar tareas específicas. En términos generales, es la parte no tangible de un sistema informático, que actúa como intermediario entre el hardware y el usuario, permitiendo la interacción y el funcionamiento del dispositivo según las necesidades y deseos del usuario.

ETL. Es un flujo de trabajo utilizado en la integración de datos que implica la extracción de datos desde múltiples fuentes, su transformación para cumplir con requisitos específicos y su carga en un destino final, como un almacén de datos o una base de datos. Durante la extracción, los datos se copian desde las fuentes de origen; en la transformación, se someten a limpieza, manipulación y agregación; y finalmente, en la carga, se insertan en el destino final. Este proceso es esencial para consolidar datos, garantizar su calidad y prepararlos para su análisis y uso en la toma de decisiones empresariales.

*Script*. Es un archivo de texto que contiene una serie de instrucciones o comandos que se utilizan para realizar una tarea específica en un entorno informático. Estas instrucciones pueden ser interpretadas y ejecutadas por un programa o intérprete de scripts.

Medidas. En Power BI, una medida es una expresión o fórmula que se utiliza para realizar cálculos sobre los datos de un conjunto específico. Pueden realizar operaciones matemáticas, como sumas, promedios, conteos, entre otras, y se

utilizan para generar valores agregados o calculados en un informe o visualización. Además, son dinámicas y se recalculan automáticamente en función de las interacciones del usuario con los datos, lo que garantiza que la información presentada sea siempre precisa y actualizada.

**Mora.** Se refiere al estado en el cual un cliente no cumple con el pago de una o varias de las cuotas acordadas dentro del plazo establecido en el contrato. En otras palabras, un cliente entra en mora cuando no paga una cuota en la fecha acordada.

**Desnormalización.** Es un proceso en el diseño de bases de datos que implica la incorporación de redundancia de datos o la duplicación de información en una tabla para mejorar el rendimiento de consultas y operaciones. Sacrifica la eficiencia del almacenamiento en favor de una recuperación más rápida de los datos. Esto se logra al reducir la necesidad de realizar un gran número de uniones entre tablas, lo que puede ralentizar las consultas en bases de datos normalizadas. Se usa en situaciones donde se requiere un acceso rápido a los datos y donde el espacio de almacenamiento adicional necesario para la redundancia no es una preocupación significativa.

**Cron job.** Es una tarea programada que se ejecuta automáticamente a intervalos regulares. Puede ser cualquier script, programa o comando que se desee ejecutar de manera automática en el sistema en un horario específico o en intervalos regulares.

**Insert.** En una base de datos, "insert" se refiere a la operación que se utiliza para agregar nuevos datos a una tabla. Esta operación se utiliza para crear nuevos registros en una tabla existente.

**KPI.** Por sus siglas en inglés Key Performance Indicators, son indicadores clave de rendimiento que se utilizan para medir el desempeño y el éxito de una organización, equipo o proceso en relación con sus objetivos y metas. Estos indicadores proporcionan una forma cuantitativa de evaluar el progreso hacia los objetivos establecidos y ayudan a las organizaciones a tomar decisiones informadas sobre cómo asignar recursos y mejorar el rendimiento.

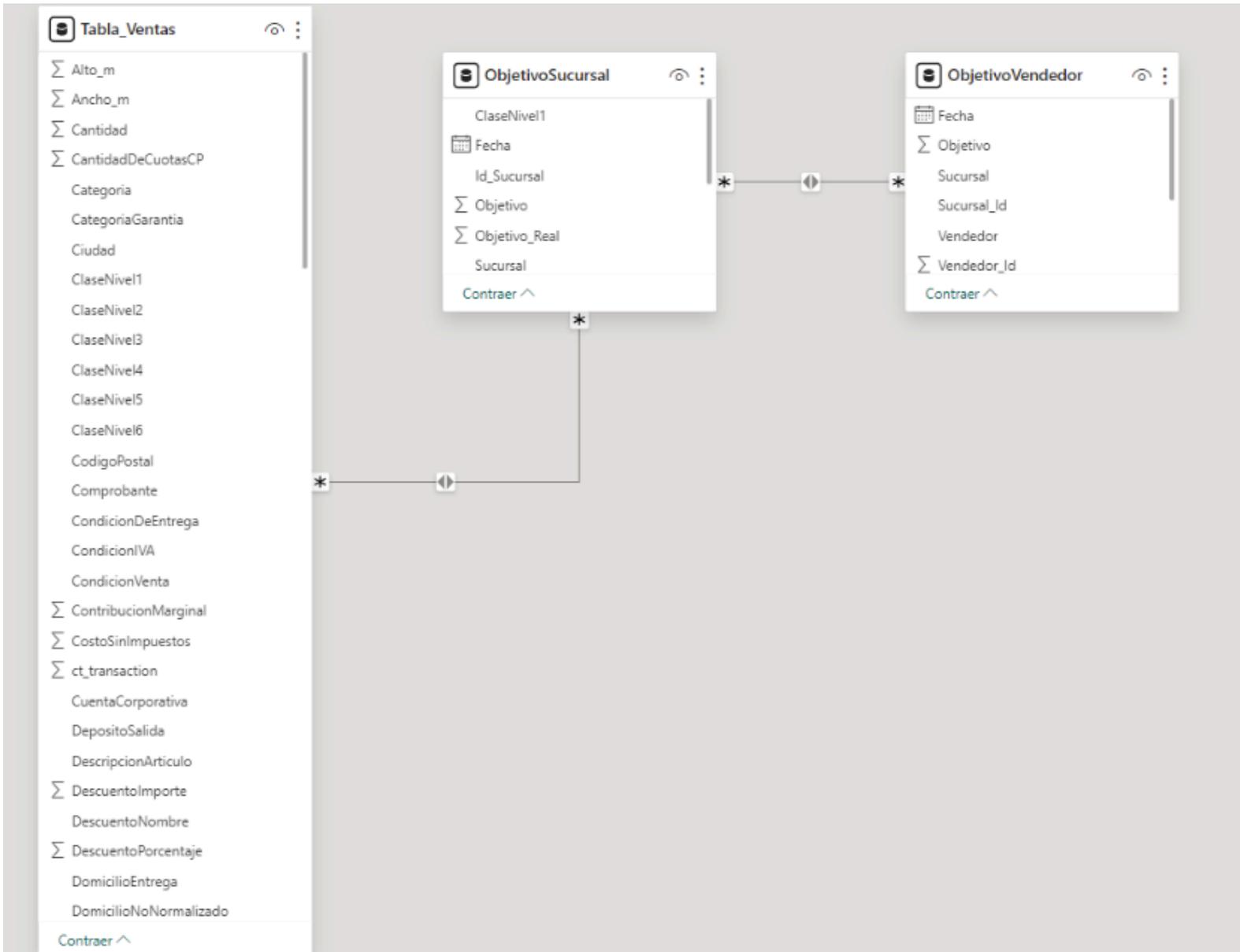
*Log.* Es un registro o registro de eventos, acciones o cambios que ocurren en un sistema. Estos registros son útiles para el seguimiento y la auditoría de actividades, así como para diagnosticar problemas y errores. Una "tabla log" se refiere específicamente a una tabla en una base de datos que se utiliza para almacenar registros de eventos o cambios.

Lógica en el proceso ETL. Se refiere a las reglas y transformaciones aplicadas durante las etapas de extracción, transformación y carga (ETL) para garantizar la calidad, integridad y consistencia de los datos.

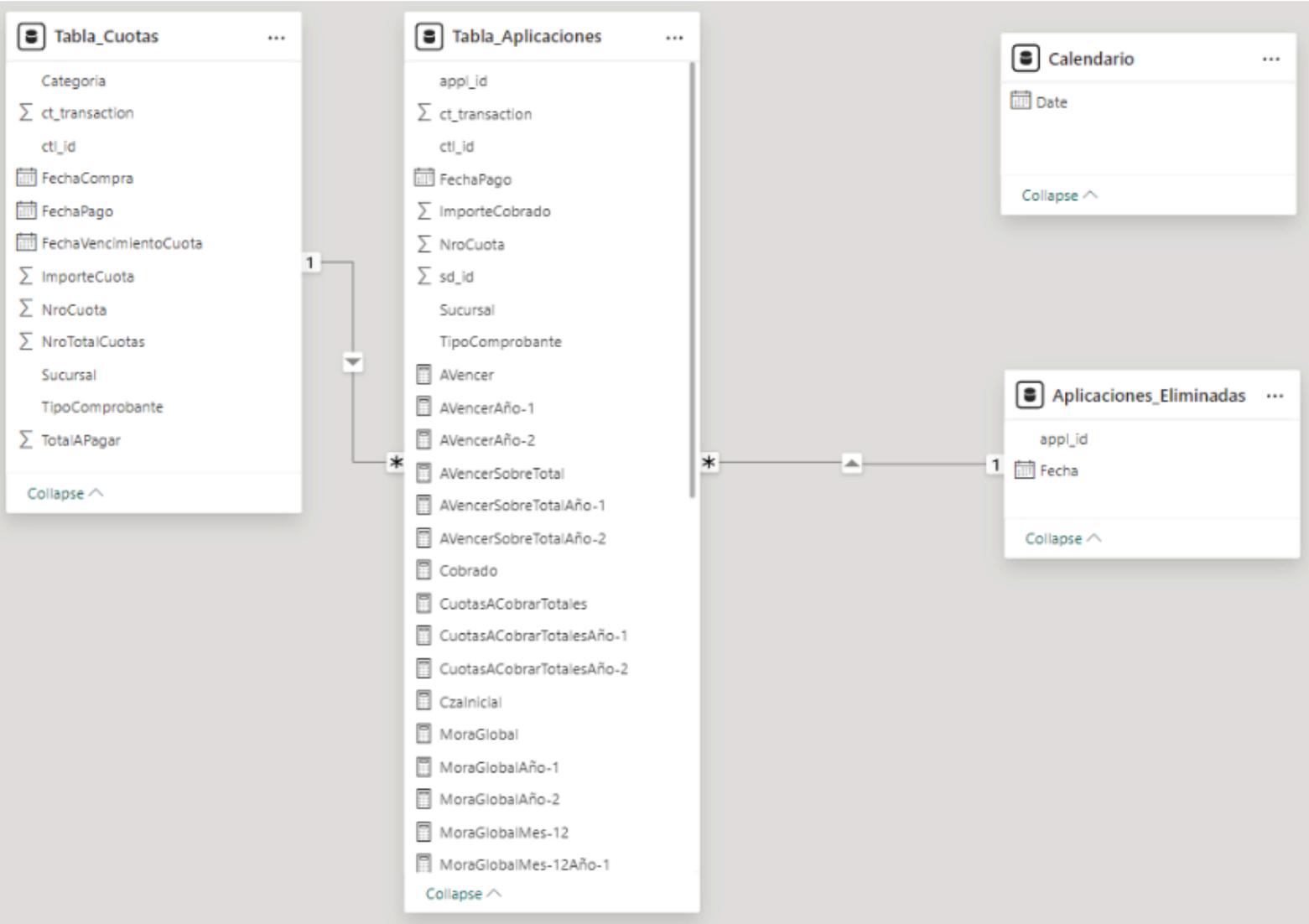
# Anexos

## Anexo I: Modelos de datos finales

### Ventas



# Cobranzas



## Anexo II: Bitácora de trabajo detallada semanalmente

- Reunión presencial. Primer día:
  - Presentación del equipo de trabajo, anotación de requerimientos básicos de la empresa, revisión de reportes actuales generados por la empresa.
  - Definición potencial de las tecnologías con las que se iba a trabajar. En principio se hablaba de Pentaho, por una cuestión de licencias. Sin embargo, se llegó a la conclusión de que iba a ser preferible trabajar con la herramienta Power BI, dado que los reportes que se iban a armar eran destinados a varios actores dentro de la empresa, cada uno con distintas habilidades para interactuar con sistemas informáticos. En ese sentido, con los tableros generados con Power BI, la experiencia de un usuario inexperto es mejor que la de Pentaho. Además, desde la empresa manifestaron que el costo de la licencia no representaría un problema.
  - Se definió que se armaría un esquema de base de datos desnormalizado para un acceso más eficiente a los datos. Para el caso de la tabla de ventas, que es la que se eligió para comenzar, la granularidad elegida para cada entrada sería cada ítem de un comprobante (factura, nota de crédito, etc).
  - Se definió la metodología de trabajo: una metodología ágil que consistiría en *sprints* de tamaño variable. También, se definieron las secciones de la empresa a trabajar: ventas, cobranzas y gastos.
- Primera semana:
  - Se comenzó trabajando con el *stored procedure* llamado 'A18 Reporte\_Tablero\_ventas', que genera un reporte que muestra las ventas por sucursal y por vendedor. Esta información es la que la empresa usaba para la creación de informes "manuales" en Excel hasta ese momento. Dicha información se buscaría replicar en tableros de Power BI posteriormente.

- Desde la empresa enviaron un listado con los distintos campos que se debían tener en cuenta para armar la tabla de ventas desnormalizada. Serían los datos de la factura en sí, y de cada ítem en particular.
- Se comenzó a analizar, a partir del reporte y de los campos necesarios, la base de datos y su estructura. Surgieron varias dudas y dificultades a la hora de entender algunos conceptos que maneja la empresa, como garantías, impuestos, notas de crédito y débito.
- Se tuvo una llamada con el departamento IT y el asesor tecnológico, que eran los que tenían los conocimientos necesarios para poder evacuar las dudas.
- Durante esta semana, fueron un problema recurrente las fallas de conexión con la base de datos a través de la VPN. Hubo interacción constante con la gente del soporte IT.
- Segunda semana:
  - Se comenzó con el armado de la consulta para la extracción de datos.
  - Surgieron dudas sobre cómo tratar los precios y sus cálculos, ya que no solo había ventas sino también devoluciones (precios y cantidades negativas).
  - Se coordinó una llamada con el departamento de IT de la empresa para evacuar dudas sobre el cálculo de precios, impuestos y garantía.
- Tercera y cuarta semana:
  - Se terminó con la primera versión de la consulta para extracción de datos.
  - Se pudo ejecutar la consulta sobre distintos meses y años. También se observó que, al hacer una cantidad considerable de *joins* y *left joins* y recuperar campos de distintas tablas, la consulta tardaba un tiempo considerable en ejecutarse para largos períodos de tiempo. También hay que destacar que la empresa registra un número importante de transacciones (ventas, devoluciones) por día.

- Se pidió al soporte de IT que otorgara los permisos necesarios a los usuarios del equipo para poder crear un esquema de base de datos nuevo y allí cargar las tablas desnormalizadas.
- Después de varios días, se obtuvo respuesta del soporte de IT y se comenzó con el desarrollo de los distintos *scripts* para la creación de tabla y carga de datos.
- Además, se fue comparando datos de distintos periodos con los datos arrojados por el reporte A18. El uso de tablas temporales fue de gran ayuda en esta instancia para evitar ejecutar el *stored procedure* A18 reiteradas veces mientras se comparaba con los datos obtenidos; esto ahorró mucho tiempo al equipo.
- Quinta semana:
  - El departamento de IT otorgó al equipo las credenciales para la conexión a Power BI. Se probó la conexión y funcionó correctamente.
  - Se ejecutó la consulta armada por el equipo en el periodo de ventas de noviembre 2023. Se comparó contra los resultados del reporte A18 del mismo periodo.
  - Se pudieron encontrar diferencias al comparar la cantidad de facturas totales y la facturación total por sección (hogar y deportes) del período que se eligió para la comparación. Esto se creía que pasaba porque, en sus reportes, si hay una misma factura con artículos de diferente sección (hogar y deportes), la transacción toma la clase del elemento de mayor valor.
  - Se consultó con los referentes funcionales dicha situación, para ver si la diferencia era significativa, y de serlo, descubrir qué está desencadenando la desviación.
  - Se modificó la consulta de extracción de datos agregando algunos campos que, aunque a la hora de mostrar un reporte no son significativos, resultan útiles para efectuar comparaciones cuando los reportes generados no coinciden con los de la empresa.

- Sexta semana:
  - Se comenzó con los tableros de la sección de ventas, aunque el desarrollo de la consulta de extracción de datos no estaba finalizado por completo, debido a que aún se buscaba cuál era la causa del desvío observado.
  - Se trabajó sobre la obtención de datos de objetivos, tanto de sucursales como de vendedores. Para ello, se recurrió a tablas auxiliares en Power BI.
  - Se coordinó una reunión con el departamento de IT para despejar dudas sobre los tableros, y mostrar los avances hasta la fecha.
- Séptima semana:
  - Se continuó con los tableros de ventas; cada miembro del equipo tomó un conjunto de tableros independiente y trabajó sobre éste, con constante comunicación en cada duda que surgía.
  - Se coordinó una reunión con el departamento de IT para mostrar avances, realizar consultas conceptuales y ajustes sobre el diseño gráfico de los tableros.
- Octava semana:
  - Se coordinó una reunión con el director para mostrar los tableros de ventas en etapa final de desarrollo, y escuchar sugerencias.
  - Se coordinó una reunión con el departamento de IT para mostrar los tableros de ventas terminados, atender a posibles ajustes y sugerencias, y planificar la continuidad del proyecto en la sección de cobranzas de la empresa.
  - Se analizaron diferencias entre sus reportes y los realizados en Power BI. La sección “hogar” presentaba pequeñas diferencias. Se definió que, tanto del lado del equipo como del departamento de IT, se estudiaría un periodo específico de datos en crudo, lo más pequeño posible, para poder encontrar por dónde vienen las diferencias.

- Novena semana:
  - Llamada con los referentes funcionales para ver la corrección de ventas. Se encontró la razón por la cual no daba bien la facturación y la cantidad de facturas. Se encontró la exactitud en todos los datos. En cuanto a la cantidad de facturas, la empresa discrimina las transacciones que son “solo flete” y el equipo no estaba al tanto de ello. Sin embargo, manifestaron que estaba bien tenerlas en cuenta, por lo que se sostuvo la implementación realizada y, en consecuencia, los nuevos reportes ya mostraban una mejora.
  - Corrección de todos los aspectos tocados en la llamada. Mayormente, ajustes de comparaciones del mes en curso respecto del mes anterior: se hace una comparación respecto de 30 días atrás, y no con el mes anterior completo.
  - Se detectaron errores en la confección de los reportes de la empresa, ya que los cálculos de facturación correspondientes al mes -1 y año -1 son incorrectos. Si se compara la facturación del mes -1 posicionado temporalmente en el mes de noviembre, con la facturación del mes actual posicionado temporalmente en el mes de octubre, deberían ser exactamente iguales. Sin embargo, en los reportes de la empresa no son iguales, mientras que en los tableros creados por el equipo sí coinciden.
  - También se detectaron errores en los objetivos de sucursales, ya que no coincidían los de la base de datos con los de los reportes en Excel de la empresa. Se notó que se extraían los datos de manera incorrecta y por ende los reportes eran erróneos. Esto fue comentado a los referentes para que estuvieran al tanto, y para que tuvieran en cuenta las diferencias existentes con los reportes de BI realizados.
- Décima y undécima semanas:
  - Reunión presencial con la referente funcional y el director, para tener una devolución de la parte de ventas y un intercambio de ideas sobre cobranzas. La referente funcional brindó información acerca de la

sección de cobranzas y los tableros que pretenden tener como resultados.

- Se planteó un esquema desnormalizado en que cada fila sea una cuota, con su cobro en caso de que se haya realizado.
- Se tuvo una reunión con el departamento de IT para realizar consultas conceptuales sobre la sección de cobranzas.
- Problemas con el esquema planteado inicialmente. La aplicación sobre una cuota puede ser modificada, debido a un posible error humano. Por este motivo, se planteó un nuevo esquema utilizando de referencia un *stored procedure* de la empresa, que trabaja con dos tablas: una para cuotas y otra para las aplicaciones sobre las cuotas.
- Duodécima semana:
  - Se terminó de definir el esquema potencial sobre el que se trabajaría: dos tablas, una de cuotas y otra de aplicaciones. Se desarrollaron las consultas para extracción de datos y carga de tablas. Posteriormente, se comenzó a trabajar con los tableros en Power BI.
  - Dificultad en la realización de los gráficos del reporte de cobranzas debido al pobre desempeño en términos de aplicación de filtros, producto de operaciones costosas computacionalmente que estaban relacionadas con comparaciones de fecha.
- Decimotercera y decimocuarta semanas:
  - Se encontró una mejor manera de realizar las comparaciones de fecha con el uso de la función EDATE.
  - Se definieron las tablas auxiliares necesarias para poder generar los reportes correctamente. Se usó una tabla calendario, para poder filtrar los datos a gusto del usuario. Y, por otro lado, se incluyó una tabla de aplicaciones eliminadas, que se extrae directamente de una tabla *log* del esquema de base de datos de la empresa.
- Decimoquinta semana:

- Reunión con los referentes funcionales. Comentaron algunas cuestiones de diseño, como por ejemplo, separar dos gráficos de un mismo tablero en dos tableros diferentes, para visualizar los datos con mayor claridad. Además, solicitaron que, para el año en curso, no se muestren los meses posteriores al actual con datos estimados, sino que sólo se muestre hasta el mes en curso.
- Se hicieron los ajustes mencionados, y se buscó optimizar las medidas implementadas en los tableros de cobranzas para mejorar la *performance*. Se eliminaron llamados a funciones innecesarios y otras operaciones costosas en términos computacionales.
- Decimosexta semana:
  - Se terminó con los tableros de cobranzas. Los tableros presentaron pequeños desvíos desconocidos en los montos de mora y cuotas a cobrar respecto de los reportes generados por la empresa. Sin embargo, los referentes funcionales consideraron dicho desvío aceptable pues era prácticamente imperceptible.
  - Se concluyó con la etapa de desarrollo de los tableros de cobranzas.
  - Se comenzó a investigar acerca de la automatización del proceso de actualización de datos desde la base de datos hacia Power BI.
  - Se comenzó a implementar la automatización del proceso de carga de datos tanto a la base de datos como a Power BI.
- Decimoséptima semana:
  - Se solicitó a la empresa que comprara la licencia “Pro” para un usuario, con el objetivo de desarrollar la automatización automática, ya que la funcionalidad “actualización incremental” solo está disponible en licencias pagas de la herramienta.
- Decimoctava semana:
  - Se comenzó a analizar qué información se debía tomar de la base de datos para la sección de gastos. Se tuvo una llamada con el

departamento IT y el asesor tecnológico, para evacuar dudas acerca de qué pretendían visualizar en sus tableros y consultar acerca de dónde extraer la información imprescindible para la creación del esquema de gastos.

- Al ver que la información necesaria no se encontraba dentro de la base de datos, se realizó otra reunión con la empresa para poder rastrear la información faltante. En la llamada nos expresaron que, luego de buscar los datos que se requieren para poder usarlos en el desarrollo del trabajo, se encontraron con que no estaba digitalizado el proceso y, por ende, la información no era parte de su base de datos de producción. Esto significó un cambio en el proyecto, ya que sin la información digitalizada, no era posible trabajar sobre esta sección. La situación fue comentada a los referentes funcionales para que decidieran cómo proseguir. Finalmente, el demandante determinó no incorporar los tableros de gastos en el desarrollo del trabajo, aunque se mostró dispuesto a documentar la información para retomarlo en un futuro.
- Se terminaron de desarrollar los *stored procedures* de actualización del esquema de BI.
- Se terminó de implementar la automatización de extracción de datos en Power BI tanto para ventas como para cobranzas. Primero, hubo que configurar en Power BI Desktop (aplicación de escritorio para crear los tableros) en cada tabla la cantidad de días hacia atrás de los que se desea extraer datos de la base de datos (uno en todos los casos, que corresponde al último día). Luego, se configuró la periodicidad de la actualización en Power BI Service (aplicación *web* para acceder a los reportes *online* como usuario final).
- Se realizó una breve capacitación a los referentes funcionales. Consistió en una explicación acerca de cómo crean las medidas, cuál es su funcionalidad y cómo se introducen en los diferentes tipos de gráficos y tablas dentro de la herramienta Power BI. Ellos manifestaron previamente interés en esta capacitación, ya que poseían

conocimientos en IT y querían entender el funcionamiento general de la herramienta.

- Decimonovena semana:
  - Se envió al soporte de IT de base de datos de la empresa un documento que detalla la periodicidad con que se deben ejecutar los *stored procedures* de actualización del esquema de BI. Además, se explica también en ese documento cómo configurar en su servidor un *gateway* de datos de Power BI, que actúe como un "puente" para la actualización de datos entre fuentes locales y el servicio de Power BI en la nube.
  - Se coordinó una llamada con el soporte de IT de la empresa para evacuar dudas respecto del documento y los pasos a seguir.
- Vigésima:
  - Una vez que el soporte de IT siguió los pasos detallados en el documento, se verificó que, accediendo al Power BI Service, efectivamente se visualicen los datos pertinentes. Esto se volvió a revisar días después, para confirmar que la actualización se estuviera ejecutando debidamente.
  - Se terminó de desarrollar el informe. Durante el desarrollo del proyecto, se fueron realizando anotaciones de manera periódica. Esto redujo considerablemente el tiempo total de desarrollo del informe.

# Bibliografía

1. Web oficial Microsoft [en línea]. Microsoft.

<https://learn.microsoft.com/es-es/power-bi/guidance/>

2. Web oficial Microsoft [en línea]. Microsoft.

<https://learn.microsoft.com/es-es/sql/sql-server/?view=sql-server-ver16>

3. Web oficial Microsoft [en línea]. Microsoft.

<https://learn.microsoft.com/en-us/power-automate/>

4. Superset [en línea]. The Free Dictionary.

<https://encyclopedia2.thefreedictionary.com/superset>