



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**REDISEÑO DE UN PROCESO DE PLANIFICACIÓN Y ADQUISICIÓN DE RECURSOS
DE SUMINISTRO EN UNA EMPRESA DE RETAIL**

PROYECTO DE GRADO PARA OPTAR AL GRADO DE MAGÍSTER EN INGENIERÍA
DE NEGOCIOS CON TECNOLOGÍAS DE INFORMACIÓN

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

NICOLÁS ESTEBAN AVILÉS ESPINOZA

PROFESOR GUÍA:
SEBASTIÁN RÍOS PÉREZ

PROFESOR CO-GUÍA:
FELIPE AGUILERA VALENZUELA

COMISIÓN:
CECILIA IBARRA MENDOZA

SANTIAGO DE CHILE
2022

RESUMEN EJECUTIVO

El trabajo de rediseño de un proceso de planificación y adquisición de recursos de suministro en una empresa de retail con apoyo en técnicas de analítica de datos para la estimación de demanda es una tarea orientada a aportar directamente al Área de Operaciones de Ropas Camargo S.A (RCSA).

El proyecto consiste en el rediseño de la actividad de generar pronósticos de demanda sobre el canal digital, articulando patrones de consumo y la evaluación de información con valor de negocios, como cierre o apertura de tiendas físicas, con el fin de aportar a la toma de decisiones operacionales y promover el manejo operacional dentro de la empresa para dar soporte a la venta, y así mejorar la capacidad de cumplimiento de los pedidos realizados en el canal online.

Entre las características relevantes del desarrollo del proyecto, que cuenta como base a la Metodología de Ingeniería de Negocios (Barros, 2015), se menciona, en primer lugar, la utilización de una base de datos de venta de productos, cuyas transformación y preparación finiquitó en 96.740 filas de datos y 20 columnas. En segundo lugar, como resultado del modelo desarrollado en lenguaje computacional *Python*, se obtiene un 91% de *accuracy* en la predicción de la demanda de una semana hacia adelante. En tercer lugar, se menciona la integración exitosa de la tecnología para obtención y carga de datos a través de Azure SQL Server. Finalmente, la visualización correspondiente de resultados a través de la plataforma Power BI.

Como resultados de negocios del proyecto, se proyecta que la tecnología ayude a mejorar la capacidad operacional de la compañía, permitiéndose extender beneficios incluso sobre áreas anexas, como comercial y de marketing en la creación de promociones y fijación de precios para fortalecer la venta de categorías que proyectan una menor venta.

Junto al desarrollo del proyecto, se establece un balance económico positivo en la operación desde un sobre costo de \$8 CLP/mes por envío según estimaciones, posibilitándose un beneficio anual de más de \$92 millones CLP/año al establecer un sobre costo de \$1000 CLP/mes por cliente al contar con un servicio con mayor fiabilidad en el envío de productos debido a la mejora en la capacidad operacional posibilitada.

Se concluye con todos los objetivos específicos del proyecto conseguidos exitosamente, sobresaliendo la creación de un proceso de negocios que fortalece la planificación y adquisición de recursos de suministro resguardando resultados coherentes, combinando la inteligencia de negocios propia de los colaboradores de la con la información relevante para la toma de decisiones sobre los niveles de operación, y la compra y contratación de insumos de suministro.

AGRADECIMIENTOS

A mis padres, Daniel y Yeicy, ellos que sostienen que lo único que podrían dejarme sería mi educación, pero que han entregado muchísimo más que eso. Gracias por encaminarme hacia el futuro y haberme apoyado en todo el trayecto desde la infancia a llegar a este punto.

A mi abuela Blanca, en el cielo, y a mi abuelo Eugenio; quienes me cuidaron durante toda mi etapa universitaria, aceptaron en su casa y compartieron su cariño. Su esfuerzo, integridad y firmeza es la imagen que quiero llevar toda la vida.

A mis hermanas, Daniela y Valentina, que me han acompañado en este crecimiento, celebrando mis triunfos y abrazando en mis derrotas. Ellas que siempre me han enseñado la sensibilidad que por oficio se me olvida.

Quisiera agradecer también a las personas con las que compartí en este intenso trayecto de aprendizajes, proyectos e increíbles experiencias. Tuve la suerte de conocer a personas virtuosas en este camino que sin dudas son y serán actores principales de un mejor futuro. Amistades cercanas, amistades lejanas incluso de intercambio, grupos organizados, voluntariados, comisiones y centros de alumnos.

Finalmente, a los mentores que me han aceptado en el trayecto; especialmente a Felipe Aguilera quien posibilitó el desarrollo de este proyecto; y a Luciano Villarroel y Sebastián Ríos, quienes me recibieron desde el primer día en mi búsqueda para iniciar el postgrado y finalización de mis estudios.

Muchas gracias.

TABLA DE CONTENIDO

CAPÍTULO 1: INTRODUCCIÓN Y CONTEXTO.....	1
1.1. Antecedentes de la industria	1
1.2. Descripción general de la Empresa.....	3
1.3. Problema Identificado	6
1.4. OBJETIVOS DEL PROYECTO	8
1.4.1. Objetivo General.....	8
1.4.2. Objetivos Específicos	8
1.4.3. Resultados Esperados	9
1.5. Alcance	9
1.6. Riesgos Potenciales.....	10
CAPÍTULO 2: MARCO TEÓRICO.....	12
2.1. Metodología de Rediseño de Ingeniería de Negocios.....	12
2.2. Metodología para el Diseño de un Sistema Predictivo.....	13
2.2.1. Metodología CRISP-DM.....	13
2.2.2. Modelos de Predicción y Error	14
CAPÍTULO 3: PLANTEAMIENTO ESTRATÉGICO Y ANÁLISIS DE LA SITUACIÓN ACTUAL.....	16
3.1. Posicionamiento Estratégico.....	16
3.2. Modelo de Negocios	16
3.3. Diagnóstico de la Situación Actual	18
3.3.1. Problemas Identificados.....	18
3.3.2. Arquitectura de Procesos AS-IS	18
3.3.3. Modelamiento Detallado de Procesos AS-IS.....	19
3.4. Cuantificación del Problema	21
CAPÍTULO 4: PROPUESTA DE DISEÑO DE PROCESOS	24
4.1. Direcciones de Cambio y Alcance	24
4.2. Propuesta de Solución	25

4.2.1.	Generación de Alternativas	25
4.2.2.	Evaluación de Alternativas.....	27
4.2.3.	Arquitectura de Procesos TO-BE.....	28
4.2.4.	Modelamiento Detallado de Procesos TO-BE.....	29
CAPÍTULO 5: DISEÑO DE LÓGICA DE NEGOCIOS		32
5.1.	Preparación y Construcción del Modelo Predictivo.....	32
5.1.1.	Selección de Datos	32
5.1.2.	Exploración y Transformaciones del Set de Datos.....	33
5.1.3	Correlación de Datos	43
5.1.4	Exclusión e Inclusión de Datos.....	45
5.1.5	Set de Entrenamiento de los Datos	45
5.1.6	Parámetros Utilizados en la Ejecución de los Modelos de Predicción.....	46
5.1.7	Balanceo de Datos	47
5.2	Resultados obtenidos en Modelo de Predicción	47
5.2.3	Selección del Modelo de Predicción.....	47
CAPÍTULO 6: PROPUESTA DE APOYO TECNOLÓGICO		49
6.1	Arquitectura Tecnológica.....	49
6.1.3	Prototipo Funcional Desarrollado	49
6.1.4	Visualización de datos	50
CÁPITULO 7: PROPUESTA DE IMPLEMENTACIÓN.....		54
7.1	Gestión del Cambio.....	54
CÁPITULO 8: EVALUACIÓN DEL PROYECTO.....		56
8.1	Evaluación Técnica.....	56
8.2	Evaluación Económica.....	60
8.2.3	Definición de Beneficios y Costos.....	60
8.2.4	Flujo de Caja.....	61
CAPITULO 9: CONCLUSIONES Y TRABAJOS FUTUROS		63
9.1	Trabajos Futuros.....	64

9.1.1	Modelación Integrada	64
9.1.2	Predicción Detallada.....	64
9.1.3	Mayor Amplitud de Datos.....	65
BIBLIOGRAFÍA		66
ANEXOS		68
Anexo A: Información disponible en Azure SQL Database		68
Anexo B: Flujo de caja con valor mínimo viable.....		69

CAPÍTULO 1: INTRODUCCIÓN Y CONTEXTO

1.1. ANTECEDENTES DE LA INDUSTRIA

A nivel mundial, las compañías se han visto enfrentadas a una acelerada transformación digital en los últimos años para enfrentar las tendencias de los clientes y las exigencias de la pandemia sanitaria a raíz del COVID-19.

Entre los fenómenos observados a raíz de la situación antes descrita se destaca una creciente preferencia de canales digitales, sumados a restricciones de movilidad y presencialidad, apoyadas en la irrupción de tecnologías de información que permiten a los consumidores estar conectados constantemente a la red para desarrollar su vida social y económica entorno a dispositivos móviles. Así mismo, se observan cambios en los comportamientos de compra de ropa en consecuencia de las medidas sanitarias preventivas, potenciando segmentos de productos más informales, en preferencia de la comodidad y el acercamiento a las actividades domésticas.

Particularmente, ante el cierre de tiendas físicas, las compañías con una mayor concentración de ventas en tiendas enfrentan una presión mayor para escalar rápidamente las operaciones de comercio electrónico, debiendo hacer frente a una mayor cantidad de competidores de menor tamaño, pero con mayor experiencia en la comercialización en canales digitales; además de la adaptación operacional necesaria para asumir una participación de mercado en el comercio electrónico acorde a sus magnitudes de venta.

En la Ilustración 1 se da cuenta de esas variaciones, donde se observa un crecimiento general del comercio online del retail (e-commerce) frente al comercio del retail en su generalidad (Lever, Comité de Comercio Electrónico, 2021). También, en la Ilustración 2 se permite ver la correlación que tienen las medidas de confinamiento con el aumento de clientes online, lo que da cuenta de uno de los motivos que logra explicar este aumento del comercio electrónico.

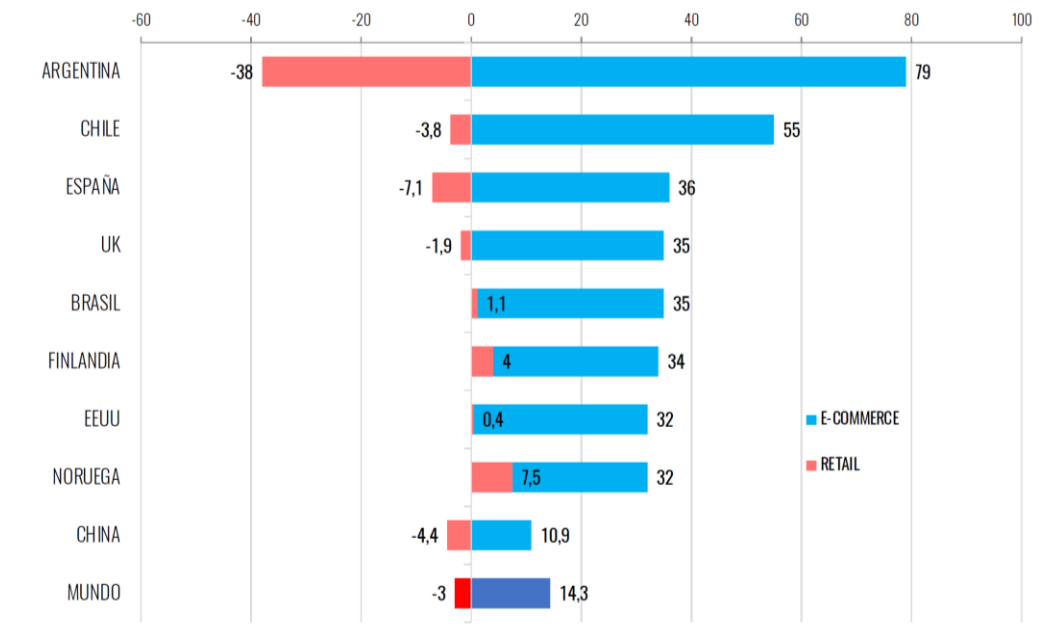


Ilustración 1: Variación porcentual de las ventas totales del comercio y del comercio online en el 2020.

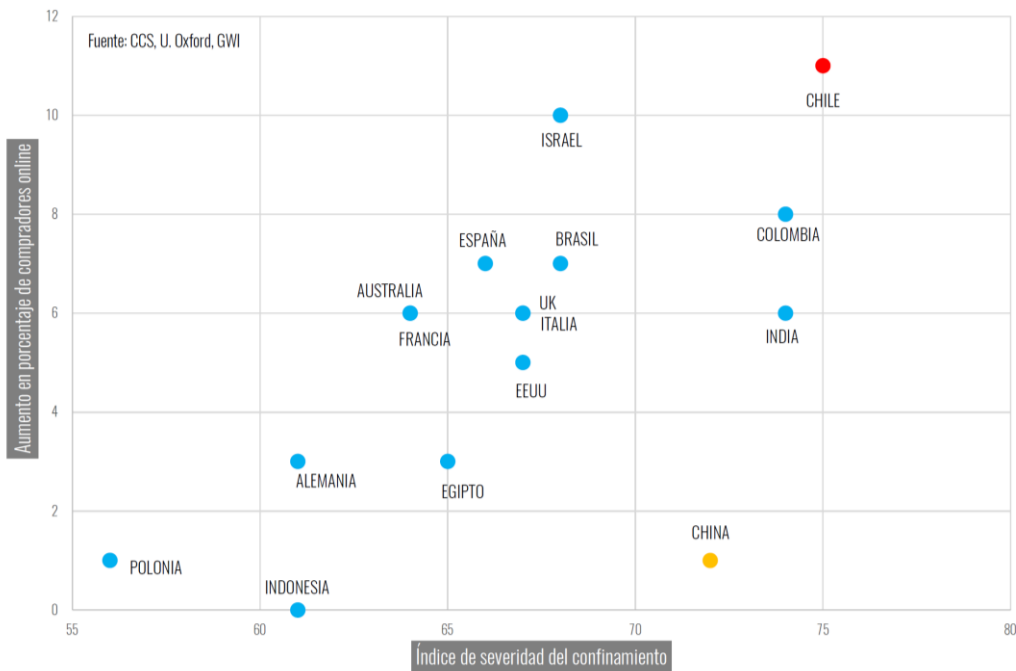


Ilustración 2: Severidad de las medidas de confinamiento versus aumento en compradores online.

Adicionalmente, en el contexto de COVID-19, se observa una recesión económica generalizada a nivel mundial y una inclinación decreciente del consumo de la población que impacta también al comercio de prendas de moda. Este último aspecto se ve bien reflejado en la Ilustración 3 para

el caso chileno (Instituto Nacional de Estadísticas, 2021), donde hasta marzo del año 2020 la tendencia de la comparación interanual para las ventas del retail sigue los años anteriores, cercana a 0 en el indicador. Pero, a partir de marzo cae drásticamente para recuperarse recién a partir de agosto del año 2020, registrando un crecimiento histórico por sobre la recesión observada en el año anterior. Este comportamiento también se ve observado en Estados Unidos, país de referencia ante el mismo indicador explicado anteriormente.

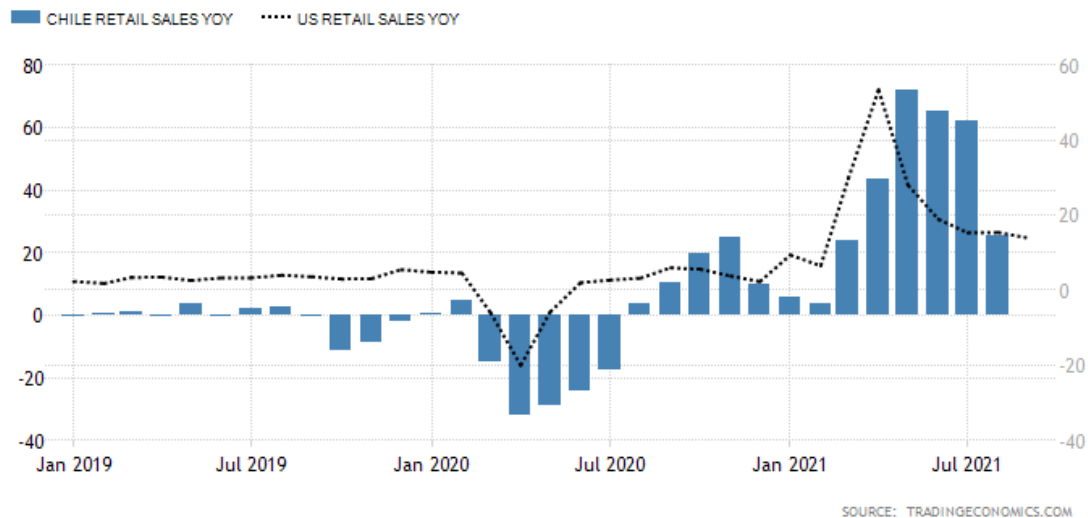


Ilustración 3: Ventas en el Retail chileno. Comparación interanual en contraste a mercado estadounidense.

Las cambiantes tendencias de compra de este mercado suponen un desafío para la industria, la que se debe ajustar a estos en términos de ventas, producción y abastecimiento para minimizar las ineficiencias y desbalances.

Como referencia al tamaño de este mercado, se permite ver que en términos netos la industria del retail ha generado más de más de 9.423 millones de dólares, como reporta el Comité de Comercio Electrónico de Chile (Lever, 2020). En particular, para la industria de la moda se cuentan con sólo estimaciones que aproximan a 4.12 millones de dólares las importaciones hacia Chile y 498 mil dólares en exportaciones, excluyendo calzado, para el año 2019 (OEC, 2021).

1.2. DESCRIPCIÓN GENERAL DE LA EMPRESA

Ropas Camargo S.A (RCSA) es una compañía con más de 60 años de historia, importadora y distribuidora de ropa multimarca cuyo presupuesto de ventas para el año 2021 es de más de \$27 mil millones; contando con la franquicia de distintas marcas de ropas internacionales en Chile.

Esta empresa cuenta con una amplia presencia a nivel nacional a través de sus canales físicos, operando 41 tiendas físicas en adición a sus clientes mayoristas, tales como Cencosud, Falabella y Ripley; y digitales, con presencia en 6 marketplace y disponiendo de 3 páginas propias de comercio electrónico diferentes. Particularmente para el canal online, la empresa comenzó la comercialización de sus productos a partir del año 2019, registrando aceleradas ventas que representan más del 30% del ingreso del negocio para el año en curso. En detalle (ver Ilustración 4), se observa que la compañía a lo largo de los últimos 3 años ha acrecentado sus ingresos en más de \$2.600 millones en el comercio electrónico, duplicando consecutivamente sus ingresos en su dominio e-commerce en estos años; y con una notoria implementación de la categoría Marketplace, registrando ventas por más del doble del valor de la categoría e-commerce para el año 2021.

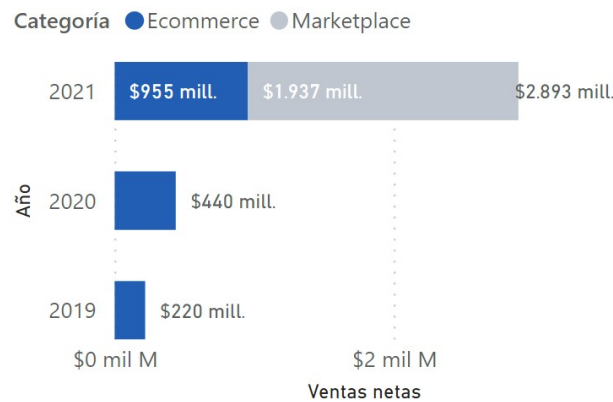


Ilustración 4: Ventas en comercio electrónico para los primeros 6 meses del año 2019, 2020 y 2021.

En cuanto a la comercialización de las diferentes marcas de ropa bajo el dominio de Ropas Camargo, se observa que esta empresa cuenta con dos nombres que comparten cerca del 90% de sus ingresos. Adicionalmente, la propuesta de valor de estas marcas se muestra complementaria, abarcando consumidores hombres y mujeres, desde los 18 hasta los 55 años, y con una gama de precios variada. A continuación, en la Ilustración 5 se resumen la proporción de ventas de las principales 4 marcas comercializadas por Ropas Camargo.

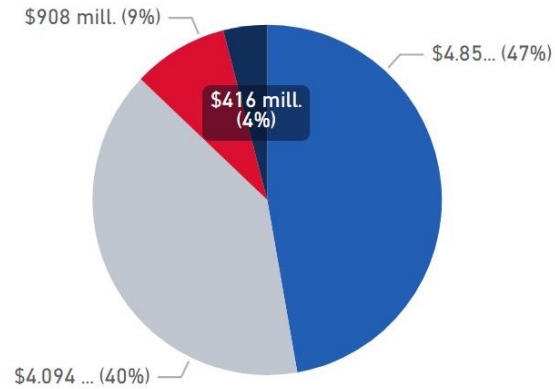


Ilustración 5: Proporción de ventas de principales 4 marcas de RCSA para los primeros 6 meses del año 2019, 2020 y 2021

Por su parte, se observan dos cadenas de valor diferenciadas. En primer lugar, se cuenta con las ventas a través de sus canales online donde destaca una oferta tanto en tiendas propias por marca, como también en outlets que favorecen la comercialización de prendas a un menor margen para rematar su stock a final de temporada. Mientras que, en segundo lugar, se sitúa el comercio electrónico. Ambas estrategias se diferencian principalmente en el acercamiento hacia los clientes, donde en el canal digital se pretende una mayor personalización al consumidor para una mejor captación, mientras que en el canal físico esta tarea resulta mayormente indiferente. Por otro lado, mientras que los canales físicos es el cliente quien se responsabiliza del transporte de la ropa; en el canal online se requiere un sistema de entrega y retiro, en caso de requerir la devolución, para el traslado de éstas. Una tercera diferencia se encuentra en los costos operacionales asociados al trabajo de comercialización en el que, para los canales físicos, permite la obtención cercana al 5% de margen; mientras que, para los canales digitales, un 25%. Estos últimos resultados se deben al costo salarial destinado al contrato de vendedores y la mantención de los establecimientos físicos, principalmente.

La Ilustración 6 muestra las dos cadenas de valor de RCSA, donde destaca unas dos primeras actividades compartidas, cuales son la importación y almacenaje de productos, anteriores a la comercialización de estos.

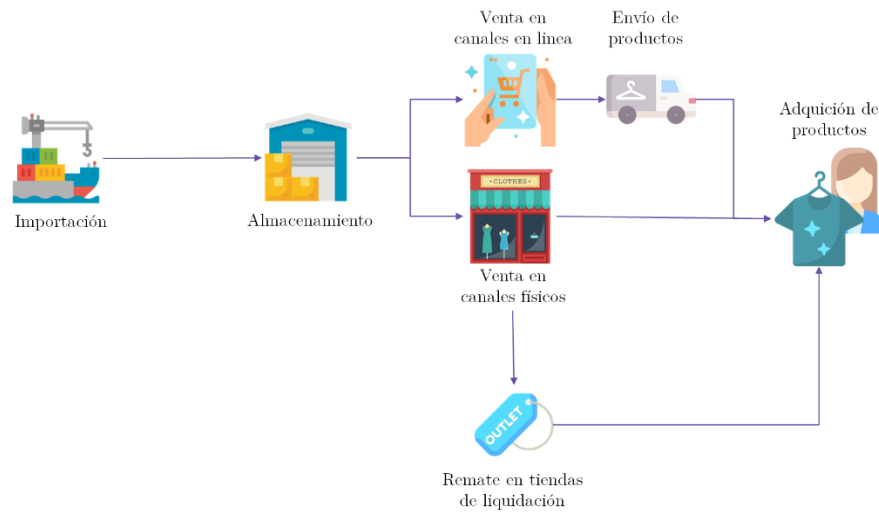


Ilustración 6: Cadenas de valor de Ropas Camargo

1.3. PROBLEMA IDENTIFICADO

En el trabajo para sostener sus resultados comerciales se han evidenciado brechas entre sus procesos operacionales, a cargo del área de operaciones, los que hoy reflejan una adaptación ineficiente a los comportamientos de compra de sus clientes, reforzados a raíz de la pandemia, para estimar su demanda y dar soporte al canal online en un contexto de COVID-19, a fin de evitar retrasos en la entrega de los pedidos y quiebres de stock, entre otras situaciones generadas.

Lo anterior tiene por origen la metodología de estimación de la compañía, la que actualmente pronostica en base a un ajuste lineal de ventas, método utilizado históricamente en su canal físico para todos sus puntos de venta. Este criterio toma como base tanto las ventas del año anterior como una estimación de crecimiento a partir del juicio experto. De esta manera se cuenta con la referencia del período anterior y un presupuesto de ventas estimado, cuyos resultados se exponen en la Ilustración 7 y la Tabla 1 para el caso del canal online.

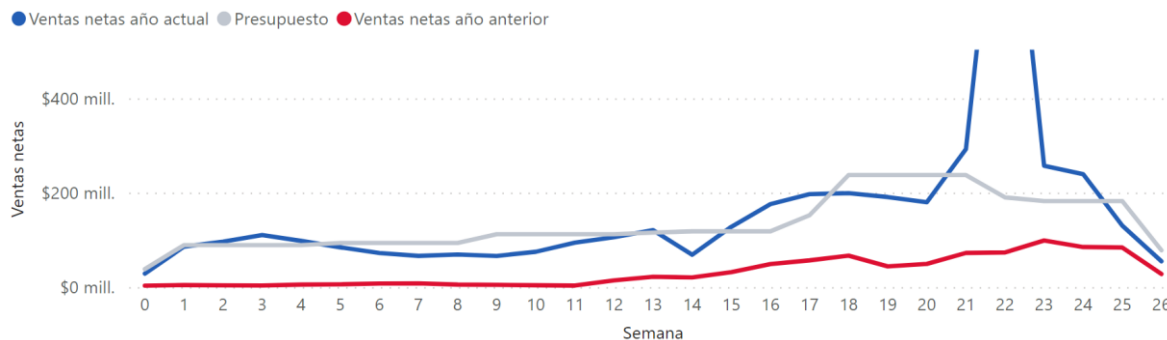


Ilustración 7: Serie de tiempo de Ventas netas Año Actual versus Presupuesto y Ventas netas del Año Anterior, para los primeros 6 meses del año 2021 en canal online.

Tabla 1: Resultados de Ventas reales versus Presupuesto y Ventas netas del Año Anterior, para los primeros 6 meses del año 2021 en canal online.

Año Mes	Ventas Netas (VN)	VN Año Anterior	2021		
			% VN c/r AA	Presupuesto (P)	% VN c/r P
Enero	\$418,9 mill.	\$19,0 mill.	▲ 2358 %	\$393,8 mill.	▲ 6 %
Febrero	\$291,9 mill.	\$27,0 mill.	▲ 1099 %	\$374,0 mill.	▼ -22 %
Marzo	\$389,2 mill.	\$31,7 mill.	▲ 1540 %	\$496,6 mill.	▼ -22 %
Abril	\$569,4 mill.	\$145,9 mill.	▲ 334 %	\$506,6 mill.	▲ 12 %
Mayo	\$1.215,8 mill.	\$260,9 mill.	▲ 400 %	\$1.051,9 mill.	▲ 16 %
Junio	\$1.448,1 mill.	\$368,8 mill.	▲ 294 %	\$781,2 mill.	▲ 85 %
TOTAL	\$4.333,3 mill.	\$853,3 mill.	437 %	\$3.604,2 mill.	20 %

Al analizar los resultados recién expuestos, se puede observar que en todos los casos las ventas del año actual se encuentran por sobre el valor del año anterior. También, que el presupuesto en base al juicio experto si bien se ajusta mejor a los valores esperados de las ventas, no es capaz de percibir los períodos de alta demanda como por ejemplo en la semana 22 y 23 del año actual, donde tomó lugar un evento de Cyber Day; produciéndose una diferencia de mínima de 6% y máxima de 85% con respecto a las ventas netas del año actual.

En este escenario, la empresa presenta brechas claras que le impiden anticiparse operacionalmente a sus compromisos comerciales. Estas consecuencias se reflejan, entre otras, en dificultades en:

- La contratación correcta de recursos humanos;
- La imposibilidad de contratación de transporte para el envío de sus productos y el manejo de stock entre sus áreas de bodegaje;
- La imprevisión de la cantidad ajustada de insumos para envío como cajas y bolsas de envío;

- La indisponibilidad de liquidez necesaria para corregir estos eventos oportunamente;
- Y, en efecto de las anteriores, el incumplimiento de promesas de compra hacia sus clientes en términos de tiempo y calidad de entrega.

1.4. OBJETIVOS DEL PROYECTO

A continuación, se describen tanto el objetivo general del proyecto, como los objetivos específicos del mismo.

1.4.1. OBJETIVO GENERAL

Rediseñar el proceso de planificación y adquisición de recursos de suministro mediante la estimación de ventas para mejorar la capacidad de cumplimiento¹ de los pedidos realizados en canales online

1.4.2. OBJETIVOS ESPECÍFICOS

- Identificar y analizar la demanda actual y las tendencias influyentes para la caracterización de ventas.
- Generar definiciones de indicadores clave de desempeño que permitan medir el resultado del proceso rediseñado.
- Generar un diseñar un proceso de analítica para la estimación de demanda dentro del proceso de planificación y adquisición de recursos de suministro.
- Utilizar analítica de datos para implementar un modelo de estimación de demanda optimizado.
- Visualizar los mejores resultados de estimación de demanda en plataforma tecnológica que esté conectado de manera automática a fuente de datos.

¹ Se define capacidad de cumplimiento como la capacidad de responder operacionalmente a la demanda de los clientes en torno a la disposición correcta de bolsas, cajas y personal para el envío de productos

1.4.3. RESULTADOS ESPERADOS

- Conjugar información disponible interna y externa a la compañía para establecer modelos de analítica de pronóstico de demanda.
- Desplegar resultados de estimación de demanda mediante herramienta de visualización en línea, Microsoft Power BI que permita coordinar ágilmente involucrados en venta de productos y procesos operacionales adyacentes, agilizando su toma de decisiones.
- Modelo de pronóstico de demanda que genera alerta anticipadamente frente a posibles alzas de consumo.
- Definición de tareas y roles dentro de la compañía para implementar un rediseño de proceso de estimación de demanda de la compañía.

1.5. ALCANCE

En cuanto al ejercicio de modelamiento, a través de la interiorización del negocio se conoce que gran parte de la demanda está determinada por los precios de los productos, su disponibilidad y el grado de promoción de estos en el momento de su venta, los que dependen del Área Comercial y de Marketing de la compañía. No obstante, el presente trabajo no integra esta consideración para generar el pronóstico de venta, utilizando un enfoque exclusivamente operacional.

Adicionalmente, se conoce que la descripción de los productos no necesariamente cuenta con un registro histórico de ventas, ya que pueden ser productos introducidos recientemente o por modificación en su identificador único (SKU). Como consecuencia de lo anterior, el presente trabajo se delimita a la anticipación de las cantidades de venta agregadas al nivel de categoría de productos, llamadas familias y líneas, cuales han tenido una mantención desde sus orígenes. Esta decisión de desarrollo también hace sentido de negocios, ya que para el Área Operacional resulta indiferente las características específicas de los productos, como talla o color, ya que las dimensiones de estos no cambian y los recursos utilizados resultan equivalentes; siendo útil las cantidades más generales únicamente. Dado este alcance, se excluye de los alcances del proyecto el tránsito de productos desde una bodega de almacenamiento a otra.

Por su parte, con respecto al trabajo de visualización de resultados, en el retail de moda se cuentan con diferentes indicadores de interés para evidenciar el desempeño del negocio. Sin embargo, el presente trabajo se limitará a entregar resultados de interés operacional sobre insumos de suministro y capacidades. Los anteriores, en base a estimaciones a partir del juicio experto y no a través de un estudio exhaustivo.

1.6. RIESGOS POTENCIALES

Dentro de las amenazas de éxito del proyecto, es posible enunciar tanto riesgos inherentes al desarrollo de una herramienta tecnológica como en la gobernabilidad del desarrollo, detallados a continuación en la Tabla 2.

Tabla 2: Matriz de Riesgos del proyecto. Probabilidad y Consecuencia en escala de 1 a 3, donde 1 es Bajo y 3 es Alto.

<i>Actividad</i>	<i>Peligro</i>	<i>Riesgo</i>	<i>Consecuencia</i>	<i>Prob.</i>	<i>Cons.</i>
R1: Predicción de demanda	Alta variabilidad en la demanda	Factores no capturados en base de datos	Imprecisión en el pronóstico de ventas y sin valor de negocio	3	3
R2: Acceso a información externa	Información externa no centralizada ni organizada necesariamente	No inclusión de variables claves para descripción de demanda	Imprecisión en el pronóstico de ventas y sin valor de negocio	2	3
R3: Continuidad del proyecto	Cambio de tecnología	Modificaciones en la arquitectura tecnológica actual	Imposibilidad de integración con sistemas de la compañía	1	3
R4: Desarrollo del proyecto	Información para desarrollo de proyecto en conocimiento experto	No involucramiento de contraparte	Retrasos y retrabajos	1	1
R5: Asignación de roles y recursos	Dependencia de disposición de compañía en favor del uso de la tecnología desarrollada	Resistencia al cambio o indeterminación de funciones	Herramienta tecnológica queda sin utilización	1	2

R6: Desarrollo del proyecto	Restricciones de seguridad a plataformas tecnológicas	de a	Las tecnologías cuentan con accesos privilegiados no disponibles	Inaccessibilidad a plataformas de la compañía	1	3
-----------------------------	---	------	--	---	---	---

Dadas las consideraciones anteriores, los resultados graduados de Probabilidad y Consecuencia dan origen al siguiente gráfico que permite evidenciar los riesgos de mayor urgencia a contener.

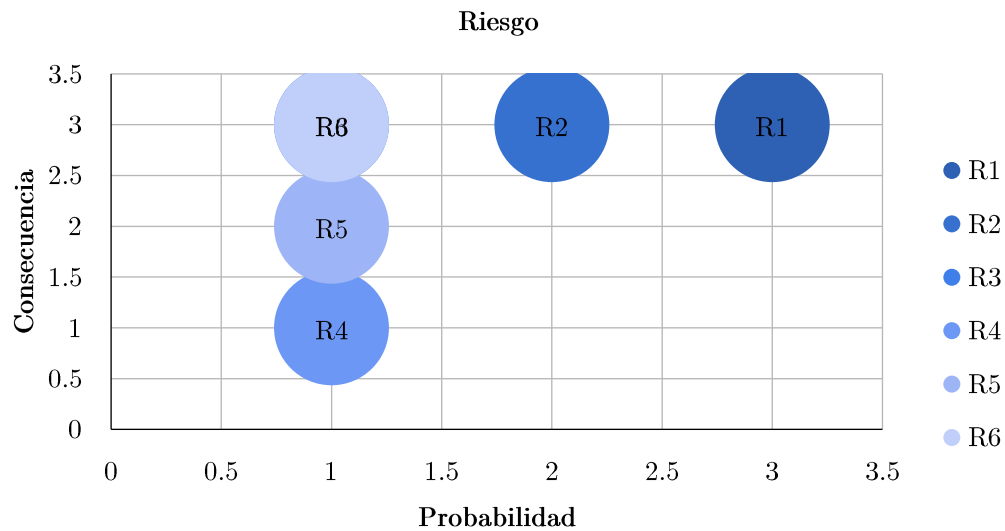


Ilustración 8: Gráfico de Matriz de Riesgos

A modo de conclusión de los resultados anteriores, se observa una gran dependencia en la predicción de demanda de la compañía (ver Ilustración 8), motivo por el cuál dentro del plazo de desarrollo de este proyecto se determina otorgar mayor tiempo a esta actividad, de forma que se puedan contener las consecuencias. También, se propone anticipar un caso en que el modelo predictivo no funcione, y que el proyecto deba depender de estadística descriptiva a modo de contener la magnitud de consecuencias.

CAPÍTULO 2: MARCO TEÓRICO

2.1. METODOLOGÍA DE REDISEÑO DE INGENIERÍA DE NEGOCIOS

Para la investigación y planteamiento del contexto de negocios se utiliza la Metodología de Ingeniería de Negocios (Barros, 2015), que permite estudiar desde el Planteamiento Estratégico (Hax, A. C., y Wilde II, D. L., 1999) de la organización hasta un Diseño de las Aplicaciones de Tecnologías de Información con apoyo en los Procesos, en adición al uso del Framework APQC (2020) para la distinción de las actividades de la compañía.

La Metodología de Ingeniería de Negocios propone 6 pasos a seguir (Barros, 2000), como se evidencia en la Ilustración 9.

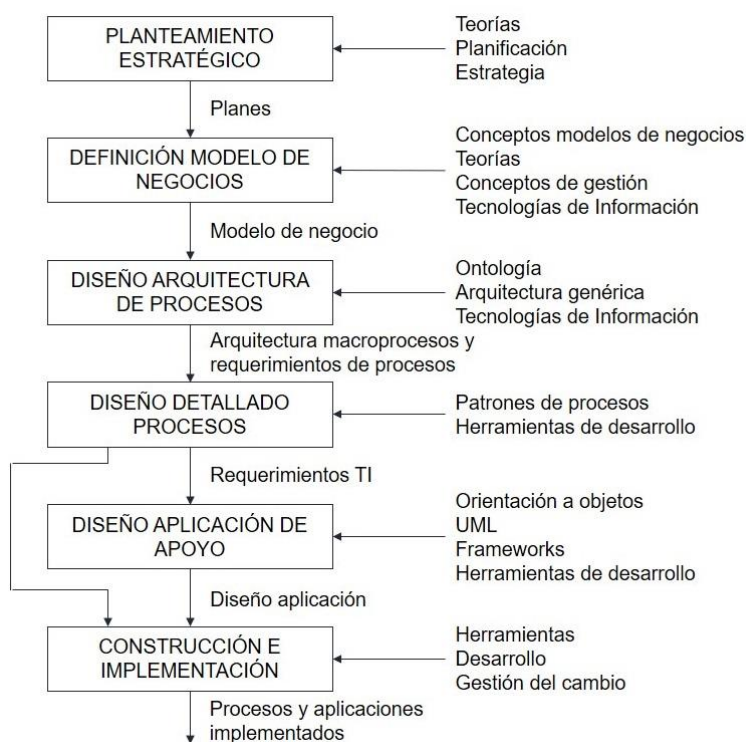


Ilustración 9: Metodología de Ingeniería de Negocios

A modo general, la metodología se puede separar en 2 grandes grupos, los que corresponden a comprensión del negocio y generación de evidencias de un problema de negocios, hasta el paso 3; y desarrollo de una solución de negocios acorde, a partir del paso 4.

2.2. METODOLOGÍA PARA EL DISEÑO DE UN SISTEMA PREDICTIVO

2.2.1. METODOLOGÍA CRISP-DM

Por su parte, para el desarrollo de una tecnología apoyada en las ciencias de datos se recurre a la utilización de la metodología CRISP-DM (Wirth, R., & Hipp, J., 2000), la que establece 6 pasos que involucran el entendimiento del negocio, entendimiento de los datos, preparación de los datos, modelamiento (Qamar, U., & Raza, M. S., 2020), evaluación y, finalmente, despliegue y visualización de datos; mientras que permite la mejora continua en una utilización cíclica de las etapas. La Ilustración 10 muestra una esquematización de la metodología CRISP-DM.

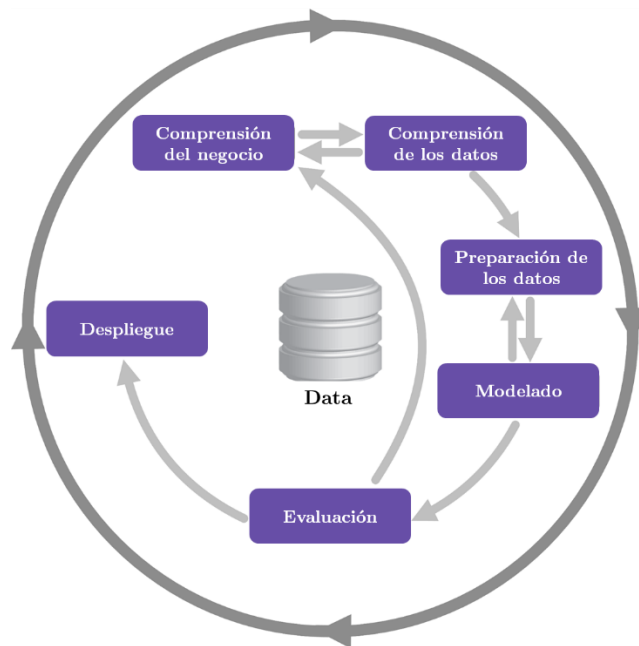


Ilustración 10: Metodología CRISP-DM

- **Comprensión del negocio:** Esta fase inicial se enfoca en entender los objetivos del proyecto en el contexto de los objetivos estratégico de la compañía.
- **Comprensión de los datos:** Esta fase de entendimiento está destinada a familiarizarse con los datos, identificar los problemas de calidad, descubrir conocimiento preliminar, y descubrir informaciones faltantes.
- **Preparación de los datos:** En esta fase se construye el conjunto final de datos a partir de los datos iniciales. Aquí se realizan las acciones de selección, limpieza, homologación y formateo de los datos.

- **Modelado:** En el modelamiento se seleccionan las técnicas y modelos pertinentes, ajustando las etiquetas o parámetros hasta llegar a un modelo que cumpla con los objetivos planteados por la compañía. También, se calibran parámetros a valores óptimos y aplican las técnicas seleccionadas.
- **Evaluación:** En esta etapa se evalúa el resultado contra los objetivos planteados anteriormente, descifrando el desempeño de las etapas anteriores y los modelos aplicados.
- **Despliegue de datos:** En esta fase se muestran los resultados con interés de negocio, persiguiendo aportar a las diferentes áreas de interés a tomar mejores y oportunas decisiones.

2.2.2. MODELOS DE PREDICCIÓN Y ERROR

Para la etapa de modelamiento existen diferentes técnicas que pueden ser implementadas en la búsqueda de un mejor rendimiento de la predicción de demanda. Entre estas se pueden nombrar los modelos de predicción y error de Series de tiempo, Random Forest (RF), K-nearest neighbour (KNN) y Support Vector Machine (SVM), las que han sido probadas con éxito en aplicaciones de pronóstico de demanda sobre el retail de ropa con distintos predictores (Fuat Şimşir, 2021)

A pesar de la basta cantidad de opciones disponibles, el estudio se enfocará exclusivamente en la utilización de un modelo de Random Forest, ya que permite la eliminación del sobreajuste; la velocidad de entrenamiento; la capacidad de lidiar con numerosas cantidades de datos; la evaluación de distintas variables según su orden de importancia; y la posibilidad de optimización a través de la optimización de parámetros. Adicionalmente, en contraste a esta selección, se utilizará la lógica actual de la compañía sobre la estimación de ventas en base al período anterior con un ajuste de un 5%.

Sobre los modelos de predicción y error seleccionados se puede detallar lo siguiente:

- 1) El modelo de predicción Random Forest se encuentra ampliamente utilizado con propósitos de clasificación; sin embargo, también genera buenos resultados para la predicción de problemas. Este método selecciona características aleatorias para crear subconjuntos de datos y optimizar la longitud del árbol y el sobreajuste, combinando la información proveniente de pequeños árboles que se crean durante el entrenamiento del algoritmo (E. Izquierdo-Verdiguier and R. Zurita-Milla, 2020).

2) La métrica de evaluación *Accuracy* es un método para cálculo de error entre el resultado real de un conjunto de datos versus el pronosticado, denotado por la siguiente fórmula

$$Accuracy = 1 - MAPE = 1 - \frac{\sum_{i=1}^n |Actual_i - Pronóstico_i|}{\frac{\sum_{i=1}^n |Actual_i|}{n}} * 100$$

Donde:

- n = corresponde al número de puntos en los cuales se calcula el error
- $Actual_i$ = corresponde al resultado real
- $Pronóstico_i$ = corresponde al resultado pronosticado

Esta medida integra el MAPE, métrica utilizada para mostrar el error porcentual absoluto medio en un pronóstico de demanda.

CAPÍTULO 3: PLANTEAMIENTO ESTRATÉGICO Y ANÁLISIS DE LA SITUACIÓN ACTUAL

3.1. POSICIONAMIENTO ESTRATÉGICO

Según el modelo de Planteamiento Estratégico propuesto por Arnaldo Hax (Hax, A. C., y Wilde II, D. L., 1999), Ropas Camargo compite en un posicionamiento estratégico de diferenciación. La explicación se fundamenta en la oferta de productos comercializados por la compañía, en el cual cada marca compite en una segmentación diferente de clientes, aun cuando los productos estén similarmente definidos. Por ejemplo, una camisa con condiciones de fabricación similares se promociona ante dos categorizaciones de negocio diferentes: ropa formal y ropa casual, estableciendo diferencias a través del mercadeo del producto fundamentalmente. Así mismo, entre las marcas propias de RCSA frente a la competencia, las propuestas de diferenciación que se logran identificar recaen en la imagen de marca que se ofrece. La Ilustración 11 muestra los modelos utilizados en la promoción de productos de las 4 diferentes marcas de la compañía, en donde se alcanza a observar que, a pesar de la semejanza de los productos comercializados entre la imagen 2 y 3, existe una diferencia en la forma de uso y la percepción de formalidad generada en la utilización de los productos.

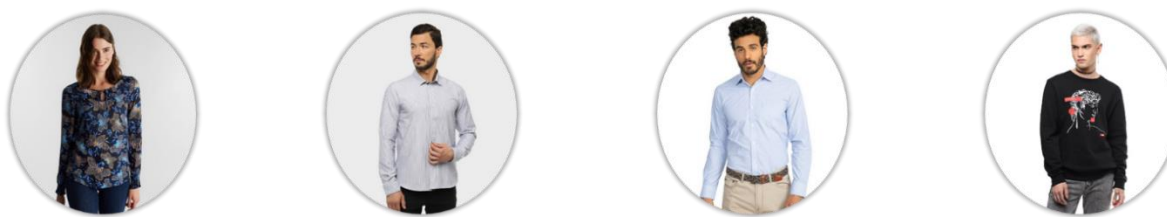


Ilustración 11: Modelos en promoción de productos de las 4 diferentes marcas de la compañía

3.2. MODELO DE NEGOCIOS

A raíz de la adaptación asumida por la compañía en el contexto de COVID-19, el modelo de negocios de la compañía se vio modificado, orientándose hacia al posicionamiento del cliente al centro de este. En este contexto las actividades claves del negocio, agrupadas en (1) Uso y análisis de datos, (2) Gestión de la cadena de suministro, (3) Generación de marketing digital, (4) Gestión de canales de ventas, y (5) Mantenimiento de relaciones comerciales; se relacionan para una mejor gestión en la venta de productos y entrega de servicios. En la Ilustración 12 se detallan las actividades claves en el modelo de negocios de la compañía.

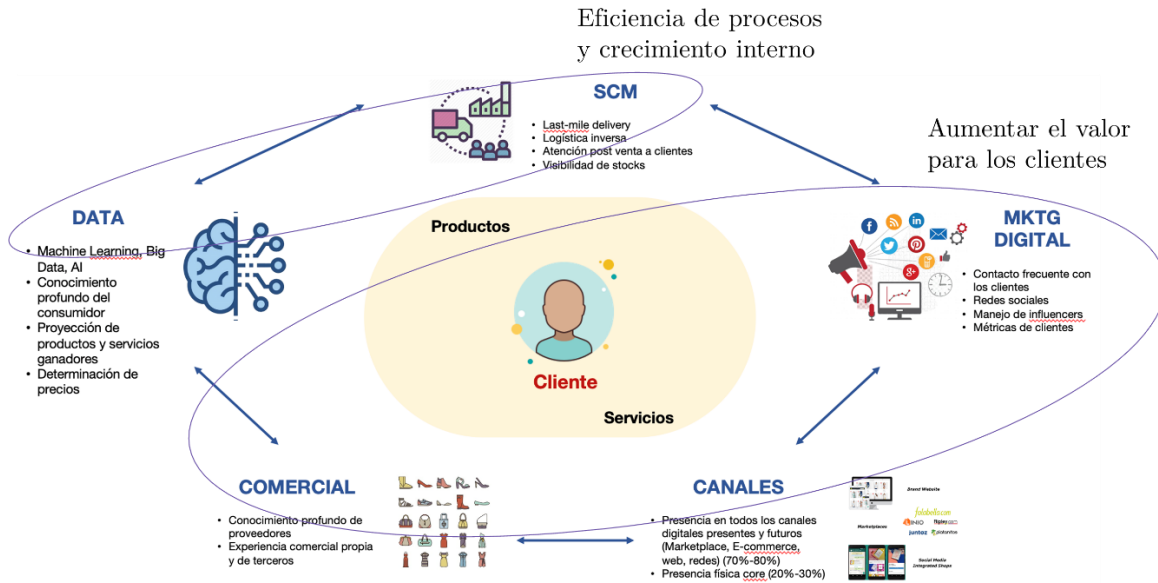


Ilustración 12: Modelo de negocios de RCSA

A través de la ilustración anterior, es posible abstraer las prioridades estratégicas de la compañía, las que inspiran los objetivos estratégicos detallados en Ilustración 13.

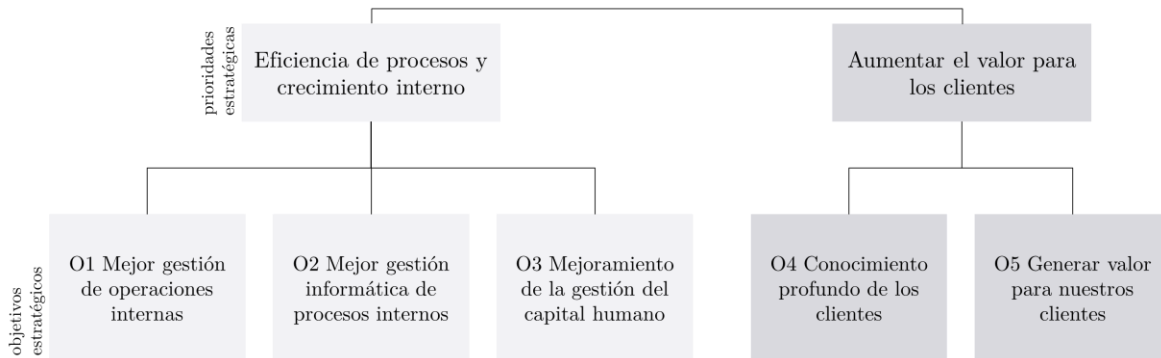


Ilustración 13: Prioridades y objetivos estratégicos de RCSA

Al respecto de la propuesta de valor hacia los clientes de la compañía, se define este como “Comercializar prendas acordes a los estilos de vida de los clientes de nuestras distintas marcas”.

Finalmente, en cuanto a la fórmula de utilidades de Ropas Camargo se destacan la generación de ingresos exclusivamente a través de la venta de sus productos, mientras que el resto de actividades para la mantención del negocio son sólo costos, entre los que se puede destacar una distinción de gastos entre los canales físicos y digitales donde, en el primer caso, los mayores costos están asociados a expensas en establecimientos y sueldos de trabajadores mientras que, en el segundo caso, en marketing digital.

3.3. DIAGNÓSTICO DE LA SITUACIÓN ACTUAL

Para la descripción de la situación actual se tomará como foco de estudio el canal online de RCSA para el período comprendido entre el primero de enero del presente año y junio de este. A continuación, se describe el problema y oportunidad identificada junto a la arquitectura de la situación AS-IS y su correspondiente modelamiento en el lenguaje BPMN².

3.3.1. PROBLEMAS IDENTIFICADOS

Para la compañía Ropas Camargo, la utilización de las herramientas actuales en el proceso de venta de productos para realizar proyecciones de demanda de su canal online no resulta en estimaciones precisas, las que generan brechas operacionales que generan incumplimiento de sus promesas de venta frente a sus clientes. Lo anterior se encuentra complementado en la sección 1.3 Problema Identificado y 3.4 Cuantificación del Problema.

3.3.2. ARQUITECTURA DE PROCESOS AS-IS

El proceso observado, donde toma lugar el rediseño propuesto en el presente estudio, es el proceso de venta de productos de las diferentes marcas de Ropas Camargo; cual es indistinto para todas estas dentro del canal online.

Dentro del proceso de venta participan cuatro áreas de la compañía: (i) Área de Marketing, (ii) Área Comercial, (iii) Área de Operaciones y TI y (iv) Área de Finanzas.

Todas las ventas son primeramente gestionadas por el Área Comercial, encargada de la constante revisión del desempeño de estas, responsabilidades dentro de las cuales se observan patrones de consumo y revisión de cumplimiento de las metas comerciales, apoyados por plataformas tecnológicas para la visualización de estos resultados. En caso de falta a las metas propuestas, se proponen nuevas iniciativas para llevar a cabo por parte del Área. Estas propuestas comerciales se ejecutan conjuntamente al Área de Marketing en caso de aprobación, quienes se encargan de entregar los mensajes en forma y contenido correcto hacia los clientes.

Ahora bien, el interés del presente trabajo se focaliza en el Área de Operaciones, cuál en particular confirma los recibimientos de órdenes de compra para el envío de los productos hacia los clientes.

² Business Process Model and Notation

En este punto del proceso se confirma la existencia de stock, insumos y transporte. En caso de inexistencia tanto de insumos como de transporte, se deben ejecutar ordenes de compra y contratación de servicios. Estos recursos son solicitados al Área de Finanzas para ejecución de compra, la cual dependiendo de la existencia de liquidez debe recurrir a financiamiento externo. Tanto la ausencia de liquidez como la inexistencia de insumos y transporte tienen tiempos de respuesta no inmediatos a considerar.

3.3.3. MODELAMIENTO DETALLADO DE PROCESOS AS-IS

La Ilustración 14 da cuenta de la situación actual en el proceso a intervenir, venta de productos de diferentes marcas a través de canal online, a través de la modelación BPMN.

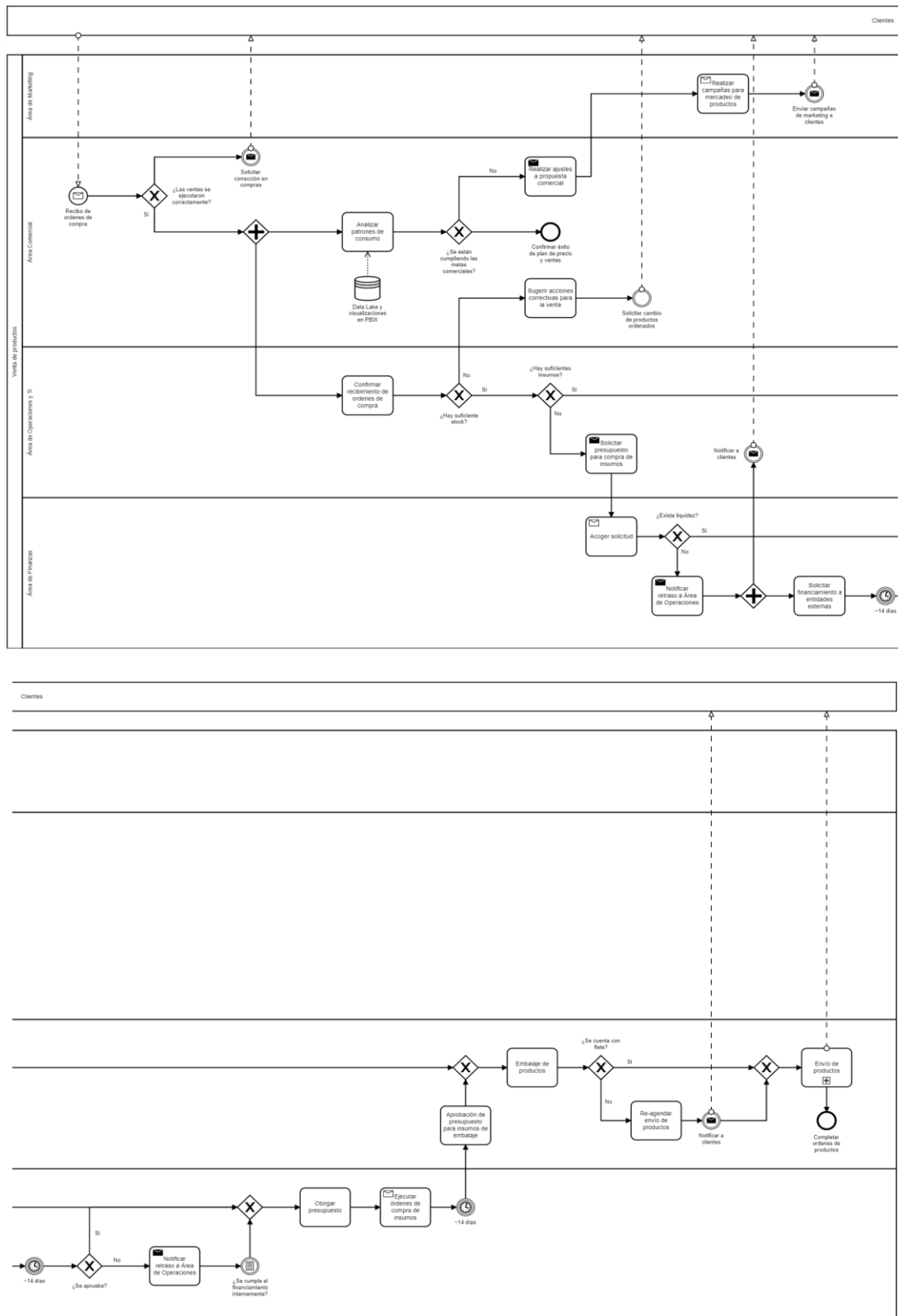


Ilustración 14: Modelación BPMN de proceso de venta de productos

También, en la Ilustración 15 se encuentra un aumento a la situación AS-IS dentro del Área Operacional y de Finanzas.

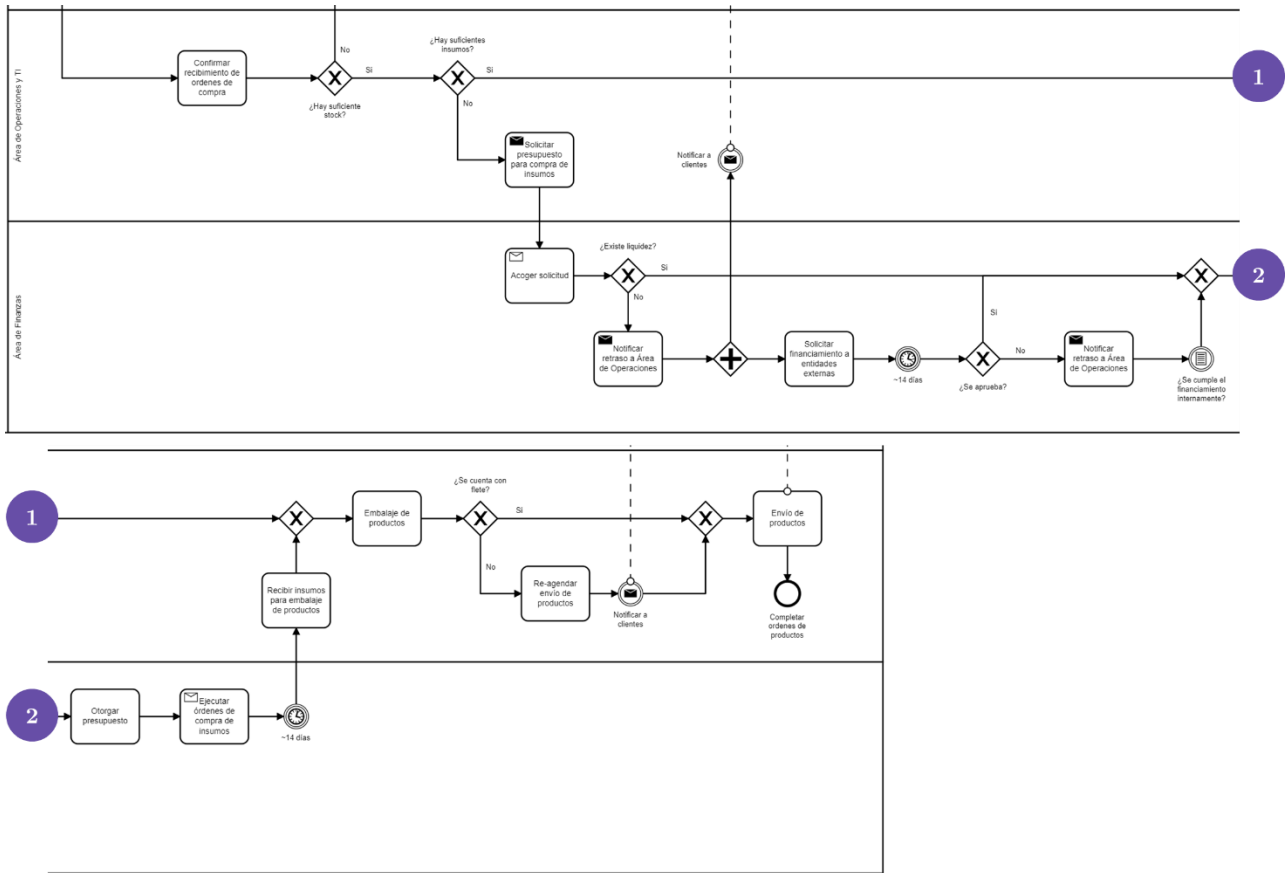


Ilustración 15: Modelación BPMN de proceso de venta de productos, Área Operacional y de Finanzas

El modelamiento de la situación actual, ilustrado anteriormente, deja entrever los tiempos de espera a raíz de las solicitudes en caso de inexistencia de insumos y transporte. En efecto, la anticipación de estos eventos puede resultar en un ahorro de estos tiempos de espera para dar soporte operacional correctamente al proceso de venta de productos de Ropas Camargo. Adicionalmente, se observa que el Área Operacional no cuenta con un proceso de observación anticipada de la demanda ni preparación previa acorde a un pronóstico de ventas.

3.4. CUANTIFICACIÓN DEL PROBLEMA

Por ejemplo, a modo agregado, las ventas netas de los primeros seis meses del año 2021 comparadas contra el presupuesto para el mismo período se diferencia en \$729 millones a favor de las ventas del año actual. En la otra mano, al realizar la misma comparación contra las ventas

netas del mismo período del año anterior, se obtiene una diferencia de más de \$3.526 millones. En la Tabla 3 se detallan los resultados anteriores, agrupados por mes.

Tabla 3: Detalle de venta neta del canal online de RCSA por mes

Año Mes	2021				
	Ventas Netas (VN)	VN Año Anterior	% VN c/r AA	Presupuesto (P)	% VN c/r P
Enero	\$418,9 mill.	\$19,0 mill.	▲ 2358 %	\$393,8 mill.	▲ 6 %
Febrero	\$291,9 mill.	\$27,0 mill.	▲ 1099 %	\$374,0 mill.	▼ -22 %
Marzo	\$389,2 mill.	\$31,7 mill.	▲ 1540 %	\$496,6 mill.	▼ -22 %
Abril	\$569,4 mill.	\$145,9 mill.	▲ 334 %	\$506,6 mill.	▲ 12 %
Mayo	\$1.215,8 mill.	\$260,9 mill.	▲ 400 %	\$1.051,9 mill.	▲ 16 %
Junio	\$1.448,1 mill.	\$368,8 mill.	▲ 294 %	\$781,2 mill.	▲ 85 %
TOTAL	\$4.333,3 mill.	\$853,3 mill.	437 %	\$3.604,2 mill.	20 %

En este escenario, RCSA al utilizar un criterio de estimación que considera un 5% más con respecto a las ventas del año anterior como de su proyección a través del presupuesto, resulta en una predicción la mayor parte insuficiente a la vista de sus resultados. Ahora bien, si se consideran estos resultados semanalmente, se observa que menos de la mitad del tiempo las ventas netas del año actual superan el presupuesto, contradiciendo la hipótesis anterior; aunque para todas las semanas las ventas del presente año superan las ventas del año anterior. La Ilustración 16 resume las características antes mencionadas.

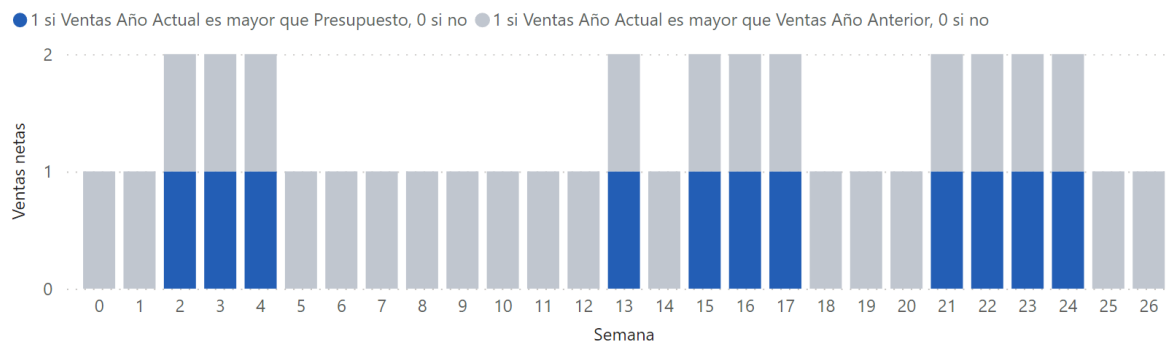


Ilustración 16: Comparación semanal de ventas netas del año actual frente a presupuesto y el año anterior.

Acumulación de valor igual a 1 si mayor y 0 si menor para ambas variables

Adicionalmente, se observa que a nivel semanal los resultados del presupuesto tienden a asemejar la venta para el año actual, no obstante, este presupuesto no se condice con situaciones extraordinarias de ventas. Particularmente, para la semana 22 del año 2021, la compañía logró ventas históricas en consideración de su participación en el evento Cyber Day de la fecha correspondiente. En este contexto, los pronósticos de venta resultan completamente diferentes, teniendo ventas netas de más de \$1.047 millones y una estimación sobre \$190 millones.

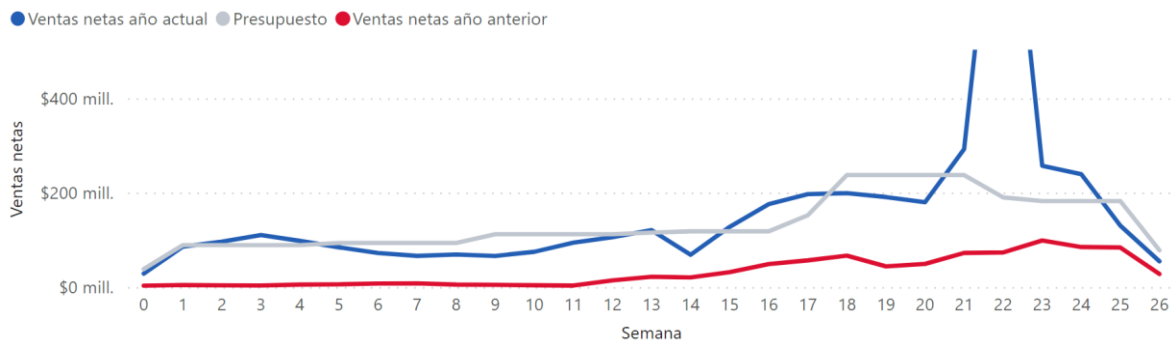


Ilustración 17: Ventas netas año actual, presupuesto y ventas netas del año anterior entre enero y junio del 2021. Límite superior del gráfico en \$500 millones

En consideración de los antecedentes mostrados anteriormente, se puede concluir que existe una disparidad efectiva entre los pronósticos desarrollados por RCSA y los resultados de la venta real. En ese sentido, a partir de la mera información de ventas históricas y el conocimiento experto de los participantes de la empresa, no se permite una estimación que reduzca los errores en el pronóstico. Esta conclusión trae como consecuencia que la compañía no se encuentre preparada para eventos de sobre demanda, limitando su capacidad de respuesta operativa; reflejándolo en aspectos como limitación de insumos para el despacho de productos, tales como cajas y bolsas de protección. La reducida disposición de fletes para la entrega en el tiempo prometido al momento de la compra. Y la improvisación en general para ante estos eventos, lo que genera ineficiencias operacionales.

CAPÍTULO 4: PROPUESTA DE DISEÑO DE PROCESOS

4.1. DIRECCIONES DE CAMBIO Y ALCANCE

Luego de describir la situación actual donde se localiza la problemática identificada de Ropas Camargo, en la sección 3.3.1 Problemas Identificados, se hace clara la necesidad de otorgar un soporte tecnológico capaz de anticipar los requerimientos operacionales de acuerdo a las ventas proyectadas; propuesta de cambio que impulsa este trabajo catalogado como *Rediseño de un proceso de planificación y adquisición de recursos de suministro en una empresa de retail*.

En el proceso de ventas es posible distinguir la falta de actividades clave para la anticipación a los requerimientos de insumos, transporte y recursos en general para el envío de los productos hacia los clientes en el canal online. En este sentido, se hace clave responder a la existencia de insumos previo a la venta de productos en sí, de forma de prever un proceso que no requiera más que un embalaje de productos y envío de productos, evitando retrasos e incumplimientos, como se logra mostrar en la Ilustración 18.

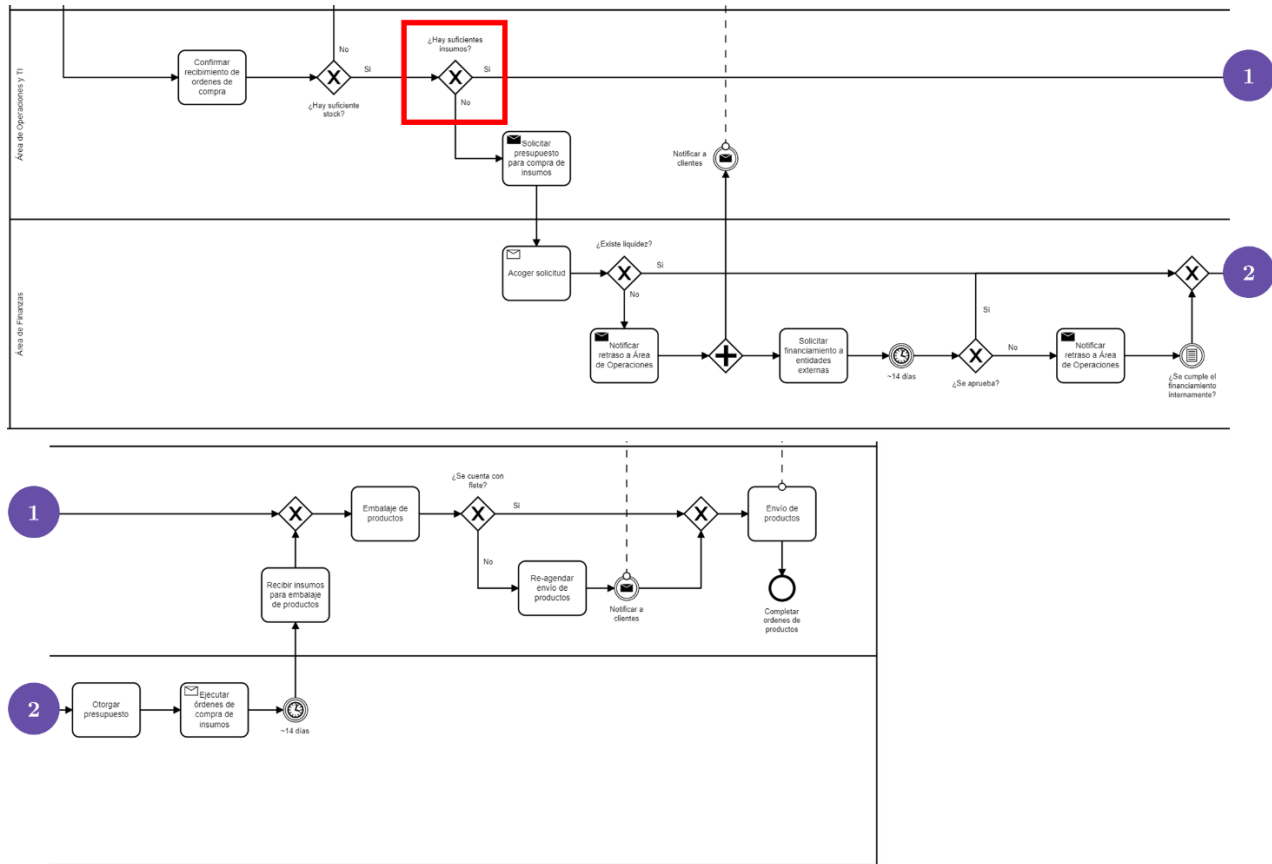


Ilustración 18: Modelación BPMN de proceso de venta de productos, Área Operacional y de Finanzas, con Identificación de Procesos Clave Faltantes.

4.2. PROPUESTA DE SOLUCIÓN

4.2.1. GENERACIÓN DE ALTERNATIVAS

Cómo se ha expuesto en las secciones más atrás, las problemáticas operacionales generadas nacen a partir de brechas de pronóstico en la demanda. Con lo anterior en cuenta, las alternativas a solución posible deben considerar un proceso de estimación de la demanda, por lo que se barajan 3 opciones principalmente:

4.2.1.1. PLANNING ANALYTICS WITH WATSON PRICING ³

Sistema proveído por IBM en su modalidad IBM Planning Analytics with Watson for supply chain planning.

En su alternativa más básica, *On Demand*, su valor asciende a 45 USD por mes. A través de esta solución se obtienen mejoras en planificación, presupuesto, pronóstico y evaluación de resultados a través de un único reporte.

Esta alternativa tiene su aplicación en la nube y provee un servicio que integra la información de las Áreas de Marketing, Comercial y de Adquisición de productos a través de una única plataforma. Como resultado de la información que se le provee al sistema, se obtiene un análisis de la demanda, cuya operación depende de la compañía.

Para la estimación de costos se toma como referencia la contratación de un Analista de Datos cuyo valor promedio de mercado es cercano a \$1.000.000 CLP por mes según *Talent.com*⁴, además de al menos \$2.000.000 CLP adicionales en estimación de costos por capacitaciones.

Como ventajas de esta solución se enuncia su alto nivel de soporte técnico provisto por el proveedor, además del amplio nivel de detalle. Sin embargo, entre sus desventajas, se enuncia el alto nivel técnico para su aplicabilidad, lo que requiere al menos un operador capacitado indisponible hoy para RCSA, además de ser necesario generar conocimiento para el uso de esta aplicación y su entendimiento a nivel general dentro de la empresa, el que hoy también está inexistente; acarreando altos costos por integración y operación.

4.2.1.2. APLICACIÓN DE CRISP-DM DENTRO DE LA COMPAÑÍA

Para la aplicación de la metodología, explicada en la sección 2.2.1 Metodología CRISP-DM, se requiere al menos un profesional de alta capacidad técnica bajo el cargo de Data Scientist, cuyas responsabilidades se extiendan desde la Ingeniería de Datos, para la manipulación de datos no organizados, hacia la Ingeniería de Negocios, para lograr integrar sus resultados dentro de los procesos de la compañía. Bajo esta descripción del cargo, los costos ascienden a \$2.350.000 CLP

³ <https://www.ibm.com/products/planning-analytics/pricing>

⁴ Plataforma tecnológica que cuenta con información salarial basados en propuestas de trabajo reales dentro de Chile y el mundo.

mensuales según *Talent.com*. Proponiendo un modelo de trabajo basado en proyectos, el proyecto en cuestión podría tomar hasta 6 meses de desarrollo bajo costos salariales, para luego solo asumir costos por soporte estimados en 2 UF por hora, requiriendo al menos 2 horas por mes.

Entre los beneficios de esta alternativa se encuentra un desarrollo especializado de alta calidad esperada para la compañía, donde la integración a sus procesos y recursos sobresale debido a la cercanía de desarrollo al conocimiento experto y procesos de la empresa. Por el contrario, entre sus desventajas resalta un alto costo y la probabilidad de falla por falta de capacidades técnicas o problemas en el desarrollo tales como la falta de datos.

4.2.1.3. *APLICACIÓN DE CRISP-DM FUERA DE LA COMPAÑÍA*

Bajo un modelo de consultoría, el proyecto tendría procesos similares que la opción anterior. Sin embargo, se destaca que la capacidad técnica de desarrollo se encuentra respaldada y con la posibilidad de ser seleccionada según la especialidad del servicio contratado. Por su parte, los costos son más flexibles y dependen del proveedor del servicio. Lo anterior constituye el principal beneficio, a diferencia de la alternativa anterior, además de contar con tiempos de desarrollo más cortos que el anterior.

4.2.2. EVALUACIÓN DE ALTERNATIVAS

Las alternativas antes mencionadas son evaluadas unas contra las otras, en un orden de 1 a 3, dónde 3 es el peor y 1 el mejor en el criterio particular. Los criterios de evaluación son el tiempo de implementación, el costo de implementación, la capacidad de integración prevista y la capacidad técnica provista, para dar un orden final ponderado. La Tabla 4: Evaluación de Alternativas resume la evaluación.

Tabla 4: Evaluación de Alternativas

Alternativa /Criterio	Tiempo de implementación	Costo implementación	Capacidad integración	Capacidad técnica	Promedio
Planning Analytics with Watson Pricing	1	2	3	3	2.25
Aplicación de CRISP-DM dentro de la Compañía	3	3	1	2	2.25
Aplicación de CRISP-DM fuera de la Compañía	2	1	2	1	1.5

Como resultado de la evaluación anterior, se observa que la tercera alternativa presenta una mejor evaluación en comparación con las alternativas anteriores, por lo que se decide la Aplicación de CRISP-DM fuera de la Compañía, solución que ayuda a la empresa en la exploración de tecnologías de forma externa y reduciendo el riesgo de desarrollo.

Si bien las primeras dos alternativas resultan menos ventajosas en el escenario actual, se recomienda dar seguimiento al desarrollo del proyecto, debido a que para ambas soluciones se cuenta con la posibilidad de desarrollar modelos integrados de pronóstico de ventas, complementando la información no estructurada de la compañía tal como la información de marketing (campañas, ofertas y promociones) y comercial (optimización de márgenes e ingresos).

4.2.3. ARQUITECTURA DE PROCESOS TO-BE

Bajo la alternativa de solución seleccionada, para el modelamiento de la situación TO-BE se toma como referencia el Framework APQC (2020). Dentro de esta metodología el proceso a intervenir se ubica en la categoría de procesos de (4.0) Entrega de Productos Físicos, dentro del grupo de proceso de (4.1) Planificación y alineación de recursos de la cadena de Suministro y (4.2) Provisión de materiales y servicios. En particular, se ajusta a las actividades (4.1.2.1) Desarrollo de un pronóstico de demanda en línea, (4.2.1.4) Empatar las necesidades a las capacidades de abastecimiento y (4.2.3) Ordenar materiales y servicios.

4.2.4. MODELAMIENTO DETALLADO DE PROCESOS TO-BE

La situación deseada TO-BE se describe a continuación en la Ilustración 19.

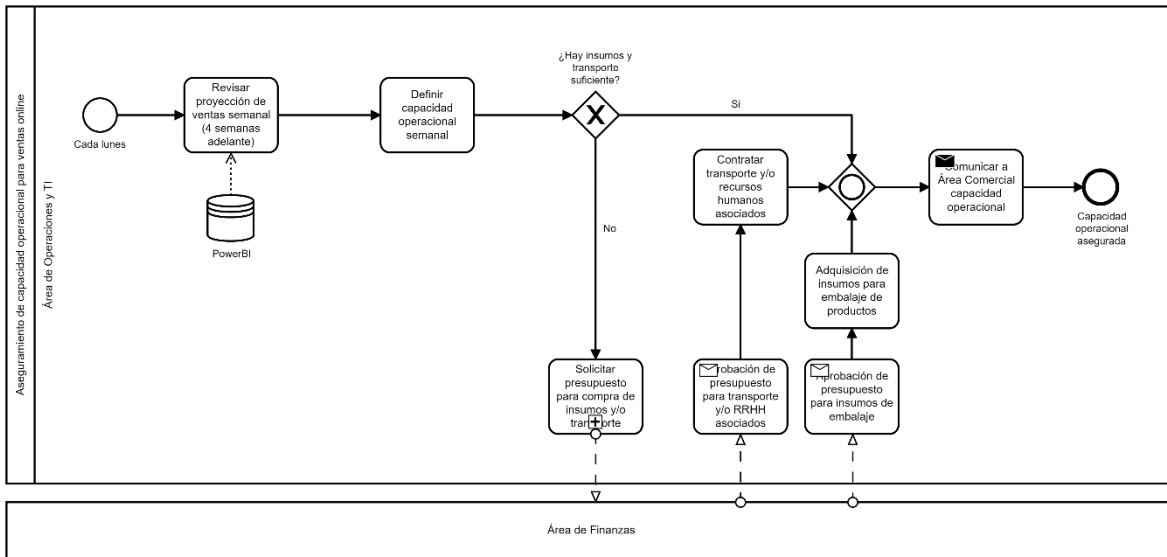


Ilustración 19: Modelación BPMN de proceso de Aseguramiento de capacidad operacional para ventas online, Área Operacional y de Finanzas

En este rediseño se destaca una anticipación de un mes para definir una capacidad operacional semanal, logrando una anticipación en tiempo y ejecución de los trabajos acorde a las necesidades definidas. En esta lógica, una vez se cuenta con la capacidad asegurada, se comunica al Área Comercial esta definición de modo que se ejecutan ventas en consideración de este resultado. El trabajo complementario propuesto permite que el Área Comercial pueda ajustar tanto sus precios y el Área de Marketing sus propuestas de publicidad para ajustarse a la realidad operacional, de forma que se comuniquen promesas de ventas hacia los clientes con un grado de certeza garantizado.

El proceso intervenido considera la utilización de una plataforma de Power BI para el Área Operacional, de forma que se logre disponer la información oportunamente para la toma de decisiones. Particularmente, este Power BI debe mostrar pronósticos de demanda junto a los requerimientos de insumo, transporte, recursos humanos y financieros en anticipación a las ventas del futuro.

A continuación, en la Ilustración 20 se muestra el proceso TO-BE enmarcado en el Framework utilizado, APQC, en el contexto de rediseño.

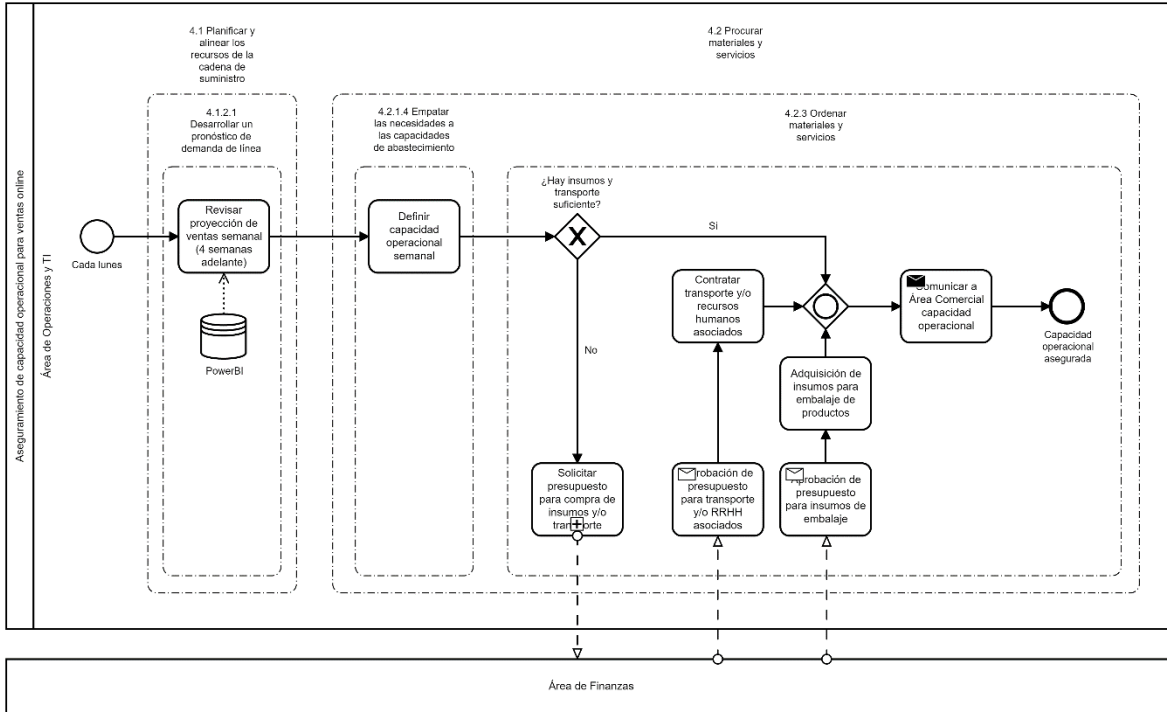


Ilustración 20: Modelación BPMN de proceso de Aseguramiento de capacidad operacional para ventas online, Área Operacional y de Finanzas. Framework APQC.

Bajo la lógica anterior, el rediseño se puede dividir en 6 pasos:

1. Cada lunes se revisa la proyección de ventas semanal para 4 semanas en adelante a través de una plataforma de visualización de datos que contemple tecnología para la generación de pronósticos de venta.
2. A partir de los resultados observados, se define una capacidad operacional para dar soporte al proceso de ventas en el canal online para 4 semanas en adelante como una decisión táctica.
3. Ante la resolución anterior, se revisan los requerimientos de insumos y transporte para envío de solicitud en caso de ser necesario.

En caso de no requerimiento, se pasa directamente al paso 6.

4. En caso de envío de solicitud por uno o ambos casos, se espera hasta recibir el presupuesto para transporte, recursos humanos asociados a la actividad, e insumos de embalaje.

5. En caso del presupuesto para transporte y/o recursos humanos asociados a la actividad de envío de productos, posterior al recibimiento se realiza la contratación según requerimiento.
6. Posterior a la ejecución de las actividades anteriores, se comunica la capacidad operacional al Área Comercial, cual se encargará de encausar el proceso de venta ante el nivel establecido; regulando la oferta y demanda de productos.
7. Se finaliza el proceso, concluyendo con una capacidad operacional asegurada para 4 semanas en adelante.

CAPÍTULO 5: DISEÑO DE LÓGICA DE NEGOCIOS

5.1. PREPARACIÓN Y CONSTRUCCIÓN DEL MODELO PREDICTIVO

El objetivo de negocio propuesto en el desarrollo de este trabajo contempla la predicción de ventas futuras a través de las ventas en canales online de la compañía, considerando el comportamiento histórico (pasado) y características exógenas a la empresa que midan el desempeño de la economía y que puedan afectar el retail de moda, mediante la programación de modelos de inteligencia artificial que sustenten el resultado más confiable.

En la presente sección se muestra el resultado final de distintas iteraciones de la metodología CRISP-DM, omitiendo parcialmente el proceso de desarrollo para dar un enfoque en su etapa concluyente.

5.1.1. SELECCIÓN DE DATOS

Los datos utilizados internos de la compañía son de propiedad de Ropas Camargo S.A, y se encuentran almacenados en el Data Lake de la empresa a través del sistema Azure de Microsoft, donde se administra la historia de ventas, márgenes y contribuciones.

Particularmente se utilizan los datos almacenados en la tabla *dbo.Ventas_Producto*, cual permite disponer de las variables a continuación.

Tabla 5: Datos disponibles en *dbo.Ventas_Producto*

Nº	Columna	Descripción
1	<i>fecha</i>	Marca de tiempo de venta de producto.
2	<i>marca</i>	Marca de producto vendido
3	<i>puntoventa_id</i>	Id asociado a punto de venta, tales como Cencosud, Ripley u otro, donde toma lugar una venta.
4	<i>familia_id</i>	Id asociado a Categoría superior de productos.
5	<i>linea_id</i>	Id asociado a Sub-categoría de productos.
6	<i>producto_id</i>	Id de producto. Contiene información de Color, Largo y Talla de prenda de ropa.
7	<i>precio</i>	Multiplicación de cantidad por precio de prenda vendida. Incluye IVA
8	<i>costo</i>	Multiplicación de cantidad por costo de prenda vendida. Incluye IVA
9	<i>precio_despacho</i>	Valor del precio de despacho asociado a la venta del producto
10	<i>cantidad</i>	Valor del precio de despacho asociado a la venta del producto

La temporalidad de los datos comprendida se encuentra entre el 07 de enero del 2019 al 31 de agosto del 2021, con un total de 1.941.239 registros.

5.1.2. EXPLORACIÓN Y TRANSFORMACIONES DEL SET DE DATOS

Los cambios y exploración de datos mostrados en esta sección se enuncian de manera enumerada, para facilitar su seguimiento.

1. Se modifican los nombres de columnas para facilitar la utilización de los datos de la manera descrita en la Tabla 6.

Tabla 6: Renombre de columnas *dbo.Ventas_Producto*

N°	Columna	Renombre de columna
1	<i>fecha</i>	<i>fecha</i>
2	<i>marca</i>	<i>marca</i>
3	<i>puntoventa_id</i>	<i>punto_venta_id</i>
4	<i>familia_id</i>	<i>Familia_id</i>
5	<i>linea_id</i>	<i>linea_id</i>
6	<i>producto_id</i>	<i>producto_id</i>
7	<i>precio</i>	<i>ventas_clp</i>
8	<i>costo</i>	<i>costos_clp</i>
9	<i>precio_despacho</i>	<i>despacho_clp</i>
10	<i>cantidad</i>	<i>ventas_n</i>

2. Se seleccionan los puntos de venta (*punto_venta_id*) específicos al mundo online, cuyos valores coinciden con los valores a continuación en la Tabla 7. Al respecto, todos los sufijos MAISA contienen hasta la venta de tan sólo una marca de la compañía, mientras que el sufijo NAISA puede incluir más de una marca. Por su parte, la categoría E-Commerce son puntos de venta exclusivos de la compañía, mientras que los Marketplace son controlados por terceros.

Tabla 7: Selección de puntos de venta específicos del mundo online.

<i>Categoría</i>	<i>ID en punto_venta_id</i>	<i>Registros</i>	<i>Primer registro</i>
E-Commerce	ARROW-MAISA	107.998	01/01/2019
	ESPRIT-NAISA	100.682	01/01/2019
	SOVIET-NAISA	12.305	16/12/2019
Marketplace	Dafiti-MAISA	44.829	27/05/2019
	Dafiti-NAISA	70.616	28/05/2020
	FALABELLA-MAISA	58.316	06/06/2020
	FalabellaMkpCl-NAISA	47.234	06/06/2020
	LINIO-MAISA	397	23/04/2021
	LINIO-NAISA	587	07/12/2020
	MERCADOLIBRE-MAISA	7.185	22/04/2021
	MercadoLibre-NAISA	11.733	11/10/2020
	PARIS-MAISA	9.712	08/04/2021
	PARIS-NAISA	6.777	11/02/2021
	Ripley-MAISA	3.243	14/04/2021
	RipleyMkpCl-NAISA	4.270	28/01/2021

Los datos muestran que una mayor porción de registros de ventas toma lugar en el canal Marketplace de la compañía (261.056 registros) que en el E-Commerce de esta (220.985 registros). Además, se muestra que, si bien la mayoría de registros coincide con la integración más temprana de puntos de ventas, 6 de estos recién tienen presencia a partir de 2021.

Se mantienen 485.884 datos posterior a esta selección.

3. Se reemplazan datos vacíos de la columna marca con el valor de *SIN_MARCA*.
4. Se excluyen outliers en variable *ventas_n*, apartando devoluciones cuyo valor es negativo y valores sobre 40 unidades de venta para evitar la dispersión de datos. Se excluyen valores de venta en variable *ventas_clp* sobre \$400.000 y con valores negativos bajo los mismos argumentos anteriores.

La Ilustración 21 muestra la distribución original de ventas posterior a la eliminación de devoluciones y ventas fuera de serie.

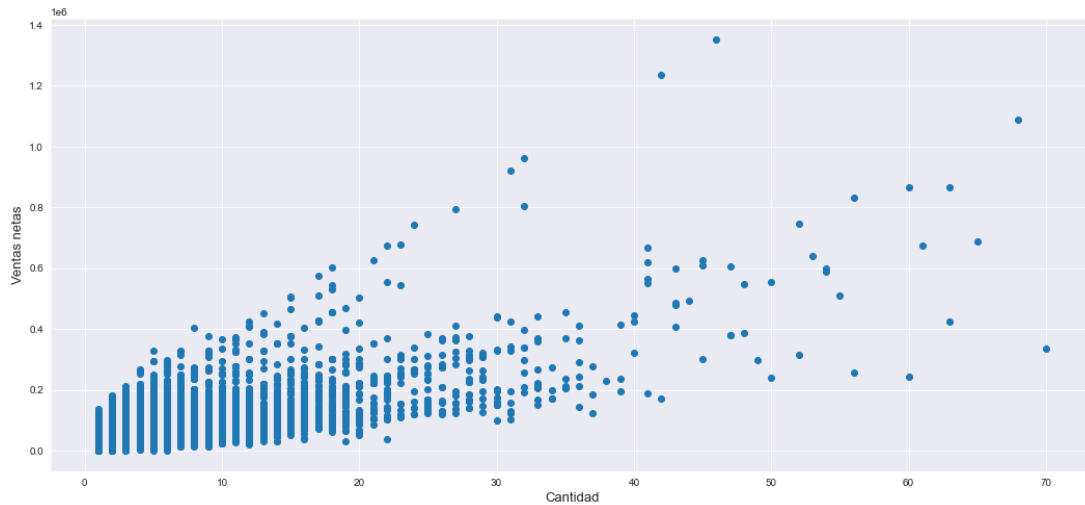


Ilustración 21: Proyección de *ventas_clp* en escala $1e6$ y *ventas_n*, en eje-y y eje-x respectivamente.

Se mantienen 485.051 datos posterior a esta selección.

5. Se excluyen valores de *Insumos* y *Consignacion* en variable *familia_id* debido a que no pertenecen a ventas del negocio. Se mantiene la misma cantidad anterior de registros posterior a esta selección.

- Se reemplazan datos vacíos en variable *familia_id* con el valor “SIN FAMILIA”. Misma transformación reemplaza los valores “SIN CATEGORIA” y SIN_FAMILIA” para su unificación.

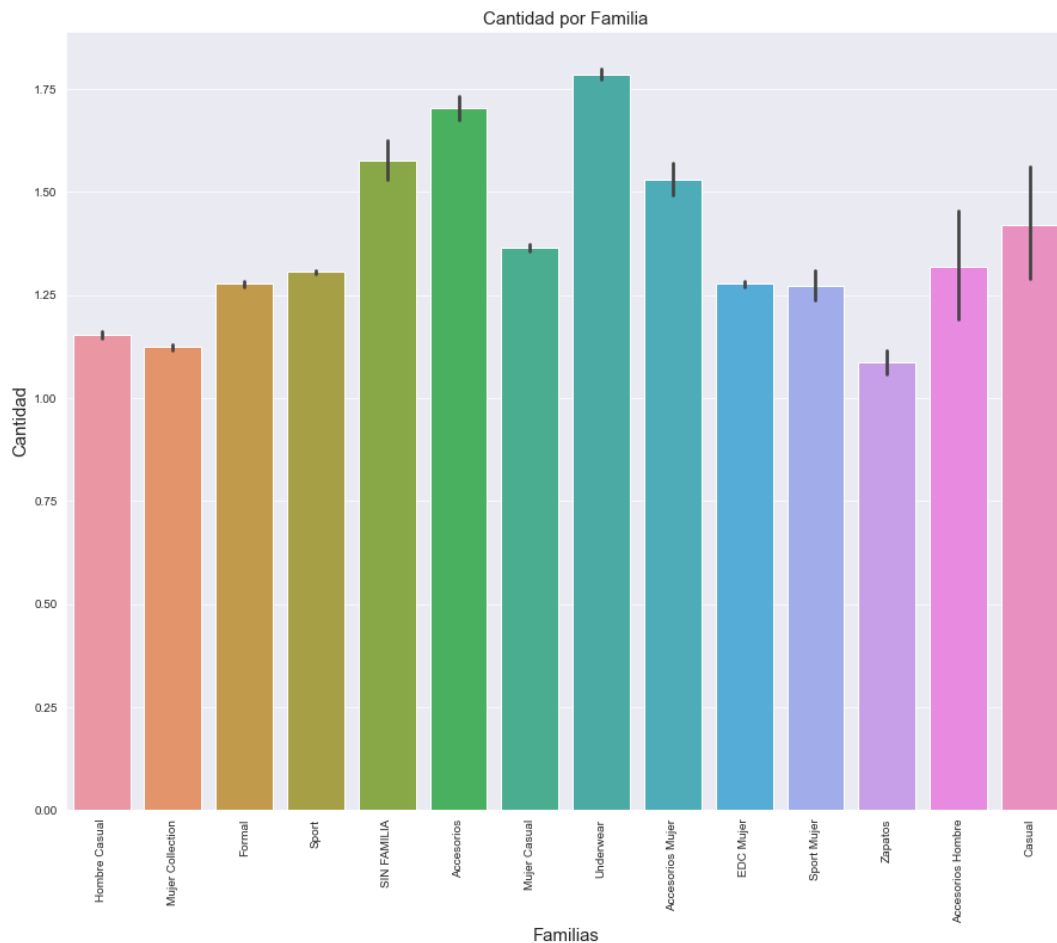


Ilustración 22: Proporción de ventas según valores en *familia_id*

En la Ilustración 22 se muestra que la cantidad de registros por valor de familia, mostrando que no existe una categoría que sea despreciable en torno a su proporción de datos dentro del total.

- Se realizan las siguientes transformaciones, descritas en la Tabla 8, a la variable *linea_id* para unificar valores con significados similares en categorías que compartan características en el tiempo.

Tabla 8: Transformación de valores sobre columna *linea_id*

Valor original	Valor destino
SIN_LINEA	SIN LINEA
Bufandas-Pañuelos	Bufandas-Pañuelos
Bufandas	Bufandas-Pañuelos
Polerón	Poleron
Polerones	Poleron
Pantalón	Pantalón
Pantalones	Pantalón
Poleras	Polera
Camisetas	Polera
Vestidos	Vestido
Vest	Vestido
Trajes de Baño	Trajes de Baño
Casacas-Abrigos-Blazers	Casaca-Abrigo-Blazer
Abrigos	Casaca-Abrigo-Blazer
Abrigo	Casaca-Abrigo-Blazer
Blusas-Camisas	Blusa-Camisa
Camisas	Blusa-Camisa
Sweaters	Sweater
Boxers	Underwear
Tangas	Underwear
Slips	Underwear
Shorts	Short
Billetera	Billeteras-Cosmetiqueros
Billeteras	Billeteras-Cosmetiqueros
Chaquetas	Chaqueta
Jockeys	Jockeys-Gorros
Gorros	Jockeys-Gorros
Faldas	Falda

La Ilustración 23 exhibe la proporción de ventas sobre la columna posterior a las transformaciones anteriores, reduciendo 56 valores diferentes a 36. Se destacan categorías de uso habitual como *Underwear*, *Pijamas*, *Bolsos*, *Calcetines*, *Billeteras-Cosmetiqueros*, *Calzas*, *Mascarillas* y *Guantes*.

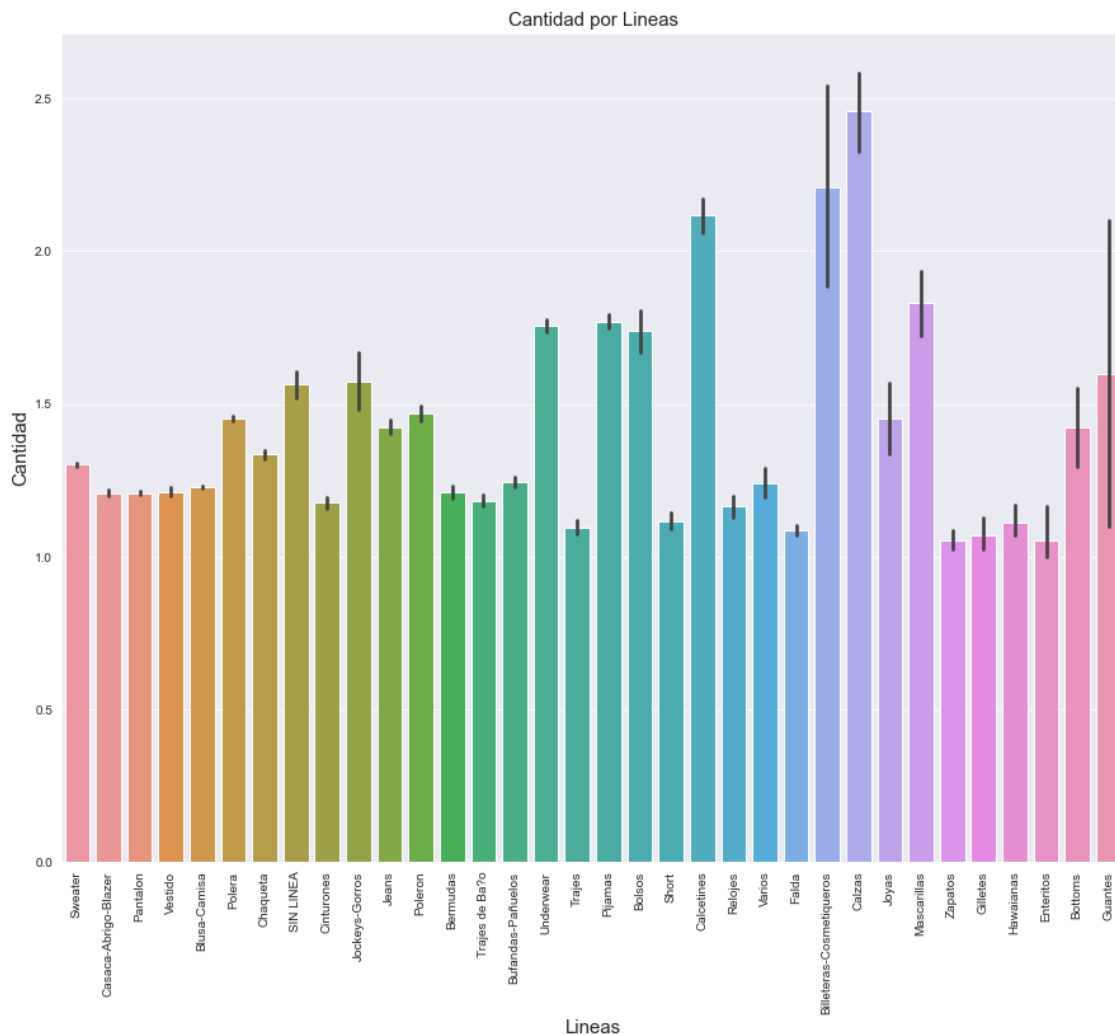


Ilustración 23: Proporción de ventas según valores en `linea_id`

8. Se eliminan variables vacías dentro de la base de datos, para descartar variables nulas, manteniéndose la cantidad de registros anterior de 485.051 filas.
9. Se elimina la columna `producto_id` de la base de datos, por su discontinuidad en el tiempo. Según la información de negocios compartida por la contraparte del proyecto dentro de la compañía, pocos productos son continuados con el mismo identificador entre temporada y otra, en ocasiones hasta siendo reemplazado productos con características completamente diferentes.
10. Se agrupan los registros en semanas, utilizando el día domingo como identificador para cada agrupación. Bajo esta estrategia, se suman todos los datos de `ventas_n`, `ventas_clp`, `costos_clp` y `despacho_clp`.

Dada la agrupación antes descrita, la base de datos se reduce a 19.136 registros de datos.

11. Se crean variables temporales que describen el año, mes, semana del año y semana del mes (columnas *year*, *month*, *week* y *week_in_month*).
12. Se crean variables numéricas sobre las variables categóricas de *marca*, *punto_venta_id*, *familia_id*, *linea_id*. Estas variables aplican la funcionalidad *LabelEncoder()* sobre las variables originales.
13. Se eliminan de la base de datos las variables *ventas_clp*, *costos_clp* y *despacho_clp* por representar valores que no se pueden proyectar en el tiempo debido a que pertenecen a un proceso de negocios anexo de la compañía, dentro de las Áreas Comerciales y de Marketing, y pueden estar sujetos a cambios inherentes al negocio.
14. Se crea columna *fecha_numerica* que tiene una representación ordinal de la variable *fecha*.

A través de los procesos anteriores, se disponen 14 columnas finales de datos, con 19.136 filas en total que representan las ventas semanales de una marca, en un punto de venta, sobre cada familia y línea. No existen valores nulos dentro de la base de datos, llamada *df*.

Por la naturaleza de la base de datos original, cual registra ventas y devoluciones de la compañía, hay períodos con discontinuidades de ventas, lo que podría afectar un modelo de predicción. Por la razón anterior se opta por llevar adelante imputación de ceros, desarrollo enumerado a continuación.

15. Se crea lista de fechas (columna *fecha*) desde primer registro en la base de datos anterior hasta la última.
16. Se agrupan las variables *marca*, *punto_venta_id*, *familia_id* y sus símiles numéricos. Como resultado, se cuentan con todas los cruces posibles en la historia de la base de datos para las columnas seleccionadas.
17. Se cruzan ambas tablas, para generar todas las posibilidades de venta en todas las fechas disponibles (llamada *df_filled*)
18. Se anexan variables de *year*, *month*, *week*, *week_in_month* y *fecha_numerica*.
19. Se cruzan ambas tablas *df* y *df_filled* para generar *df_filled_0*. Se disponen las ventas (*ventas_n*) según los registros originales.

20. Se reemplazan valores vacíos por 0 en variable $ventas_n$, para completar base de datos df_filled_0 . Con estas transformaciones, la base de datos cuenta ahora con 96.740 registros.
21. Se crea variable $ventas_nlog10$ cuyo propósito será ayudar al pronóstico de datos. Esta variable crea una transformación sobre la columna de ventas ($ventas_n$) a través de la siguiente fórmula:

$$ventas_nlog10_i = \log_{10}(2 + ventas_n_i)$$

El motivo de realizar esta transformación sobre la variable tiene que ver con normalizar la distribución de datos de la columna $ventas_n$, cuyos valores naturalmente se encuentran cercanos al valor 0, tanto por la imputación de ceros en la base de datos, como por la naturaleza del registro.

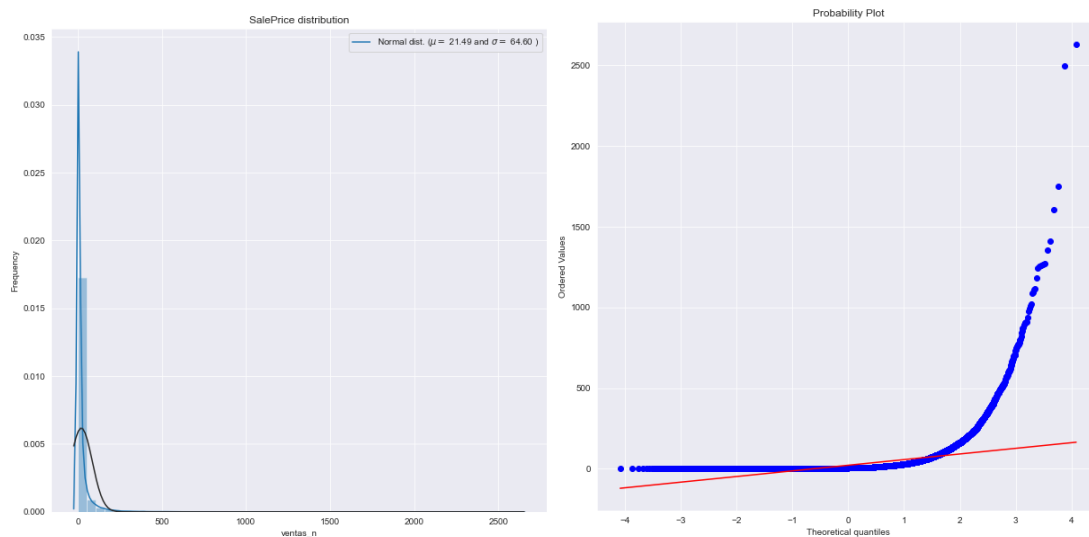


Ilustración 24: Distribución de datos y Probabilty Plot sobre columna $ventas_n$

La Ilustración 24 muestra la distribución de datos de la columna $ventas_n$ al lado izquierdo, mientras que al derecho un gráfico de probabilidades, técnica gráfica para evaluar si un conjunto de datos sigue o no una distribución determinada, como la normal.

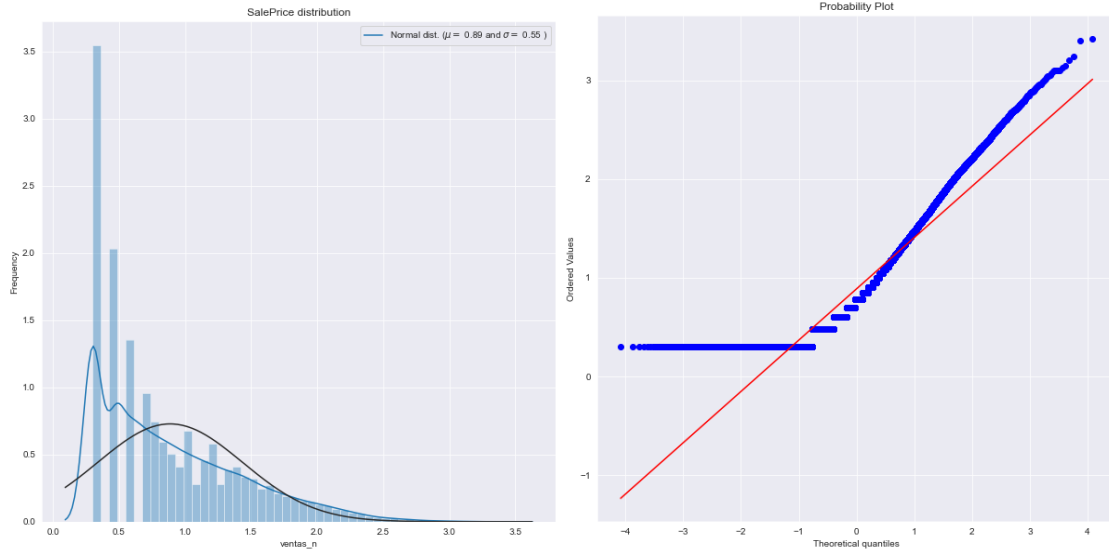


Ilustración 25: Distribución de datos y Probabilty Plot sobre columna *ventas_nlog10*

Según la teoría (Chambers, 1983) sobre esta última gráfica, la distribución de los datos (en azul) debiese seguir aproximadamente la línea recta roja, representando qué tan bien se sigue una distribución teórica normal.

Al aplicar la transformación logarítmica, se obtienen los resultados mostrados en Ilustración 25, cuyo gráfico a la derecha muestra una mejor aproximación a la distribución normal deseada.

Finalmente, la razón de sumar 2 a cada valor de *ventas_n* es asegurarse de contar con valores nulos tras la operación, de manera que cálculos futuros no sean indefinidos por llevar al denominador de una división la columna *ventas_nlog10*.

- 22. Se crean variables de desfase de 1, 2, 3 y 4 semanas con respecto a la columna *ventas_n* (*salesweek-1*, *salesweek-2*, *salesweek-3* y *salesweek-4*).

Entre la información que se excluyó del modelo de pronóstico se encuentra los casos activos semanales de COVID-19, información proporcionada por el Ministerior de Ciencias del Gobierno de Chile⁵, cuya representación se encuentra en la Ilustración 26. El motivo de su exclusión tiene que ver con la poca relación que existe, que se puede ver incluso gráficamente al compararse

⁵ <https://www.minciencia.gob.cl/covid19/>

contra la Ilustración 27, con respecto a la venta semanal de productos. La correlación calculada entre ambas variables es de 0.037.

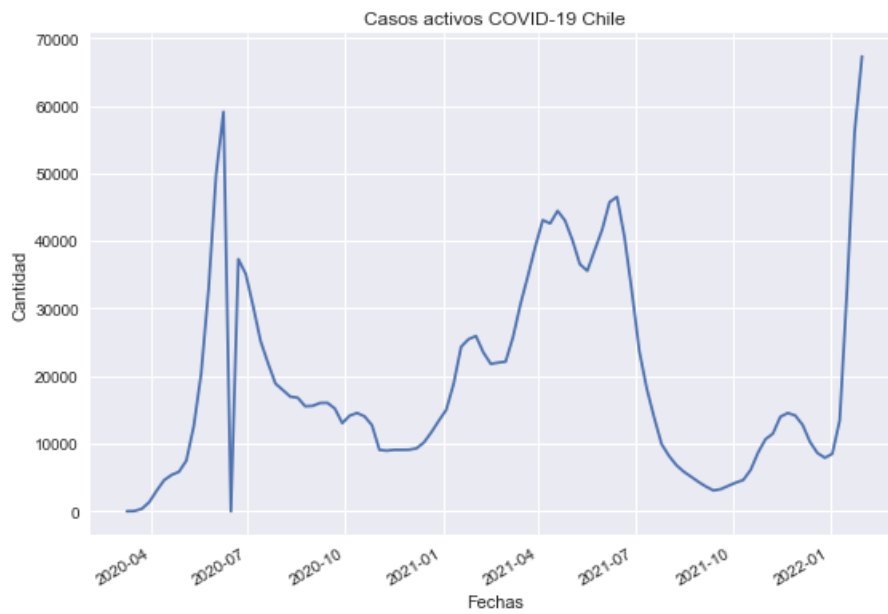


Ilustración 26: Casos activos COVID-19 Chile

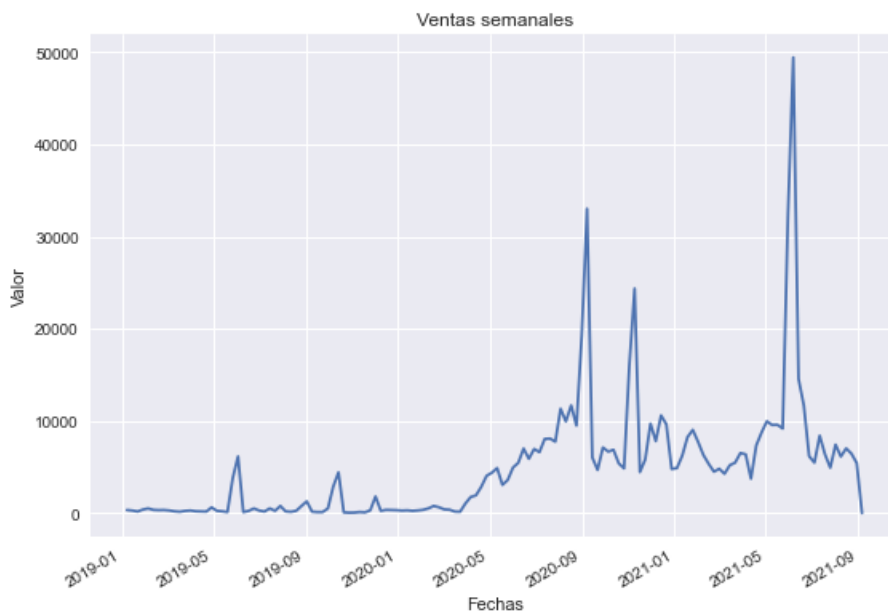


Ilustración 27: Ventas semanales en el tiempo (variable ventas_n acumulada por fecha)

Otra de las variables excluidas en el proceso de transformación y preparación de datos es la variable cantidad de puntos de venta, cuyos motivos se remontan a que no es una variable que

se tenga certeza en futuro y también la baja relación matemática con respecto a la cantidad de ventas semanales (-0.030), a pesar de tener una relación teórica fuerte con las ventas, explicada por la alza en los registros comerciales a partir de abril del año 2020 que coincide con la inclusión de nuevos puntos de venta, como se observa en la Ilustración 28.

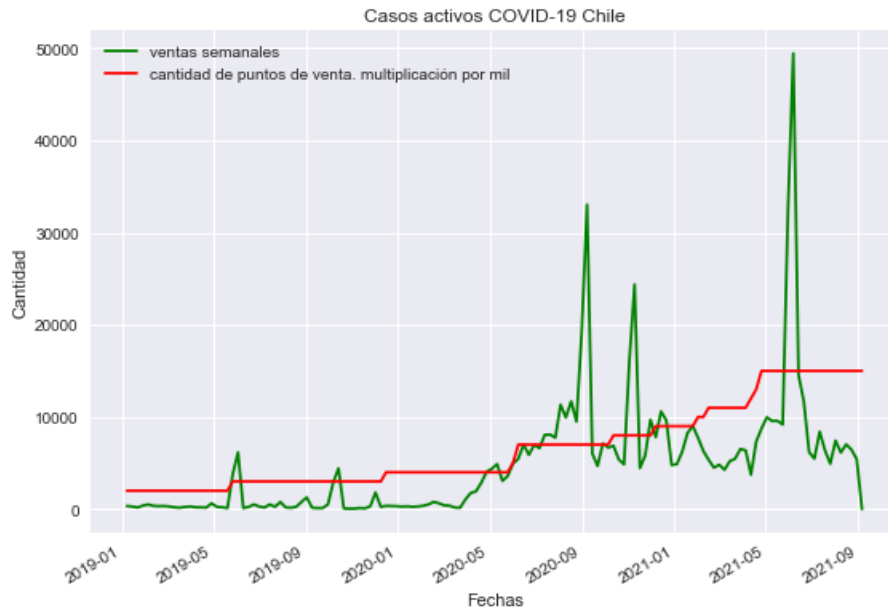


Ilustración 28: Ventas semanales y Cantidad de puntos de venta

5.1.3 CORRELACIÓN DE DATOS

A partir de las transformaciones y creación de variables descritas en la sección anterior, se procede al cálculo de correlaciones, resultándose los valores como se muestran en la Ilustración 29.

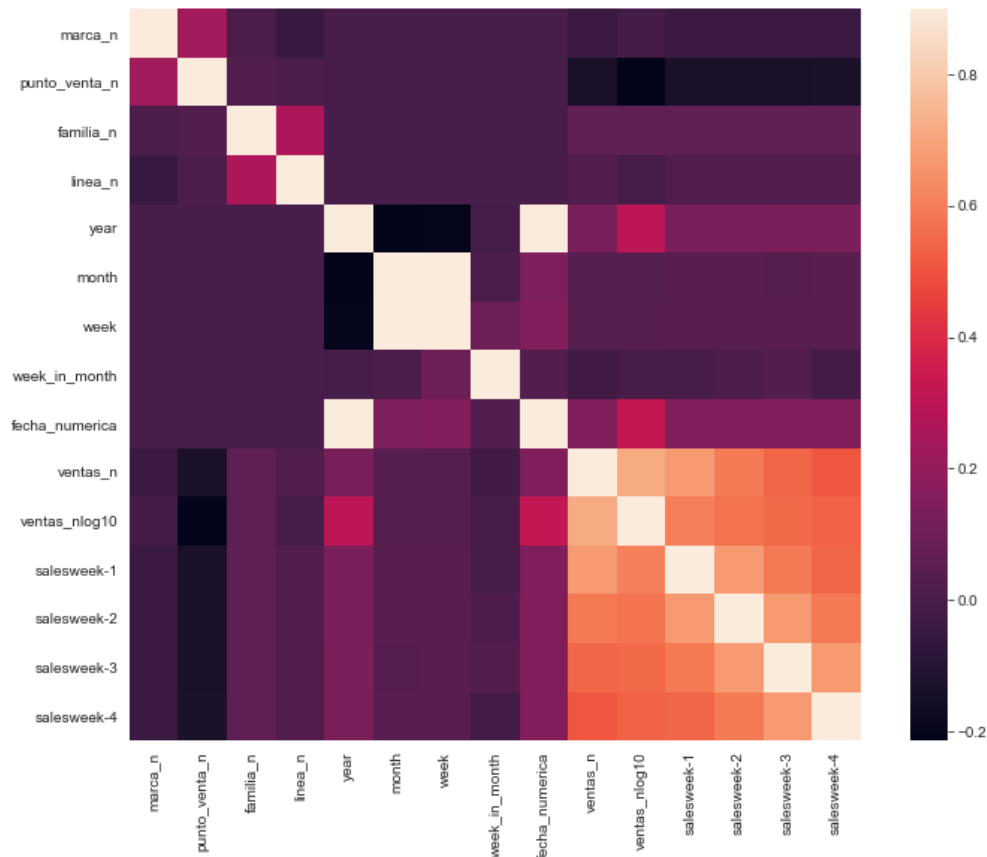


Ilustración 29: Correlación de base de datos final *df_filled_0*

Destacan las variables de desfase de ventas, cuyos valores de correlación son los mayores dentro de las columnas seleccionadas. Por su parte, se refleja que la variable con menor correlación es la variable de *punto_venta_n*. En la Tabla 9 a continuación, se resumen las correlaciones de las variables con respecto a la columna *ventas_n*.

Tabla 9: Correlaciones con respecto a variable *ventas_n*

Columna	Correlación
marca_n	-0.042
punto_venta_n	-0.137
familia_n	0.060
linea_n	0.017
year	0.129
month	0.034
week	0.035
week_in_month	-0.019
fecha_numerica	0.143

ventas_nlog10	0.713
salesweek-1	0.670
salesweek-2	0.589
salesweek-3	0.540
salesweek-4	0.506

5.1.4 EXCLUSIÓN E INCLUSIÓN DE DATOS

En relación con las modificaciones de la base de datos original descritas en la sección 5.1.1, se logra realizar un seguimiento a la cantidad de registros disponibles para el modelamiento, resumidas en la Tabla 10.

Tabla 10: Historial de transformación de datos

N°	Descripción	Registros	Columnas
1	Extracción de datos	1.941.239	10
2	Selección de puntos de venta	485.884	10
3	Eliminación de outliers	485.051	10
4	Eliminación de columna <i>producto_id</i>	485.051	9
5	Agrupación de datos a nivel de semana	19.136	9
6	Inclusión de variables <i>year</i> , <i>month</i> , <i>week</i> y <i>week_in_month</i>	19.136	13
7	Inclusión de variables numéricas <i>marca_n</i> , <i>punto_venta_n</i> , <i>familia_n</i> y <i>linea_n</i>	19.136	17
8	Eliminación de columnas <i>ventas_clp</i> , <i>costos_clp</i> y <i>despacho_clp</i>	19.136	14
9	Creación de columna <i>fecha_numerica</i>	19.136	15
10	Imputación de ceros	96.740	15
11	Transformación logarítmica de columna <i>ventas_n</i>	96.740	16
12	Creación de variables de desfase <i>salesweek-1</i> , <i>salesweek-2</i> , <i>salesweek-3</i> y <i>salesweek-4</i>	96.740	20

5.1.5 SET DE ENTRENAMIENTO DE LOS DATOS

Para la selección del set de entrenamiento de datos se utilizan funciones integradas librerías públicas de *Python*, utilizando una separación de 20% de datos para testeo y 80% para entrenamiento de manera aleatoria.

Sobre la base de datos *df_filled_0*, enunciada en la sección 5.1.1, se eliminan las columnas de texto *marca*, *punto_venta_id*, *familia_id* y *linea_id*. Además, se excluye la columna *ventas_n* por la alta dependencia de la variable a predecir *ventas_nlog10*.

Como resultado se cuenta con 77.392 registros de entrenamiento y 19.348 registros de testeo.

5.1.6 PARÁMETROS UTILIZADOS EN LA EJECUCIÓN DE LOS MODELOS DE PREDICCIÓN

Para la obtención de los parámetros de predicción se realiza un entrenamiento al modelo Random Forest a través de la optimización de grillas de parámetros utilizando la prueba y error en vista de mejorar métricas de evaluación del modelo.

Para esto se utilizan dos metodologías a través de las funciones *RandomizedSearchCV* y *GridSearchCV*. La búsqueda aleatoria y la búsqueda en cuadrícula, correspondientemente, exploran exactamente el mismo espacio de parámetros. El resultado en la configuración de parámetros es bastante similar, mientras que el tiempo de ejecución de la búsqueda aleatoria es drásticamente menor; motivo que permite probar una mayor cantidad de parámetros otorgando la posibilidad de búsqueda de un mejor resultado.

Tabla 11: Parámetros entregados para optimización a *RandomizedSearchCV*

Nombre	Parámetros
bootstrap	True, False
max_depth	10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None
max_features	auto, sqrt
min_samples_leaf	1, 2, 4
min_samples_split	2, 5, 10
n_estimators	200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000

Tabla 12: Parámetros entregados para optimización a *GridSearchCV*

Nombre	Parámetros
bootstrap	True
max_depth	80, 90, 100, 110
max_features	2, 3
min_samples_leaf	3, 4, 5
min_samples_split	8, 10, 12
n_estimators	100, 200, 300, 1000

La ejecución de la primera optimización de parámetros a través de *RandomizedSearchCV* toma alrededor de 4 horas de ejecución para una máquina de computación específica, mientras que la optimización de *GridSearchCV* toma alrededor de 2 horas para la misma máquina. Para ambas ejecuciones se toman como referencia los parámetros exhibidos en la Tabla 11 y

Tabla 12.

5.1.7 BALANCEO DE DATOS

Dentro de las técnicas utilizadas para prevenir el desbalanceo de datos se encuentra la utilización de la variable *ventas_nlog* cuyo propósito es otorgar una distribución más normal a la columna a predecir, como se describe en la sección 5.1.2. No obstante, la estadística sobre esta variable persiste en concentrarse en valores cercanos a 0.301 debido a que la variable original, *ventas_n*, es rellenada con valores 0.

Tabla 13: Estadística de columna *ventas_nlog10*

Estadístico	Valor
Promedio	0.459
Desv. Estándar	0.403
Mínimo	0.301
25%	0.301
50%	0.301
75%	0.301
Máximo	2.962

Como el caso de negocios considera relevantes todas las categorías de descripción de ventas (*marca*, *punto_venta_id*, *familia_id* y *linea_id*), resulta necesario que el modelo también reciba información sobre cuándo estas combinaciones de categorías no lograrán ventas para una fecha específica. Por lo tanto, a pesar del desbalance de datos, no se aplican más transformaciones a la variable a predecir *ventas_nlog10*.

5.2 RESULTADOS OBTENIDOS EN MODELO DE PREDICCIÓN

5.2.3 SELECCIÓN DEL MODELO DE PREDICCIÓN

Utilizando la métrica de evaluación *Accuracy* definida en la sección 2.2.2, se obtiene un 91% de *accuracy* de pronóstico en el mejor resultado de Random Forest; utilizándose los parámetros mostrados en la Tabla 14, a continuación, en el optimizador de hiper parámetros.

Tabla 14: Parámetros optimizados en modelo Random Forest

Nombre	Parámetros
bootstrap	False
max_depth	None
max_features	sqrt
min_samples_leaf	1
min_samples_split	2
n_estimators	400

Cabe mencionar que esta *accuracy* de estimación se encuentra al utilizar la optimización de parámetros sobre la información de entrenamiento y utilización sobre valores de testeo. En este sentido, esta métrica representa cuán cerca se encuentra el modelo de predecir la semana siguiente. Al ser un modelo cuya utilización de resultados es recursiva para lograr la predicción de 4 semanas en adelante, el error es acumulativo y, por tanto, su *accuracy* es esperable que sea menor.

El modelo utiliza variables categóricas, temporales y de ventas para generar el pronóstico de ventas (*ventas_nlog10*) en una semana en particular, lo que puede mostrarse como:

$$ventas_nlog10 = RandomForest(C_i, T_i, VP_i)$$

Donde estos conjuntos de variables se componen de los elementos enunciados en la Tabla 15.

Tabla 15: Variables utilizadas en modelo de pronóstico

Clasificación	Variables
Categorías (C_i)	<i>marca_n</i> <i>punto_venta_n</i> <i>familia_n</i> <i>línea_n</i>
Temporales (T_i)	<i>year</i> <i>month</i> <i>week</i> <i>week-in-month</i> <i>fecha_numerica</i>
Ventas pasadas (VP_i)	<i>salesweek-1</i> <i>salesweek-2</i> <i>salesweek-3</i> <i>salesweek-4</i>

De esta forma, para explicar la recurrencia utilizada, al tomar como ejemplo una cuarta semana de pronóstico se tiene mantienen las mismas variables de Categorías (C_i); las variables Temporales (T_i) se han adelantado 4 semanas; y las variables de Ventas pasadas (VP_i) utilizan la transformación del pronóstico de ventas ($[ventas_nlog10]^{-1}$) de la primera, segunda, tercera semana posterior a la última fecha de datos, y las ventas de la última semana de datos como datos en las columnas *salesweek-1*, *salesweek-2*, *salesweek-3* y *salesweek-4*, respectivamente.

Finalmente, como la variable pronosticada corresponde a una transformación de las ventas reales de la compañía, se debe hacer la transformación inversa:

$$ventas_n = 10^{ventasn_log10} - 2$$

Con la operación anterior se encuentran números decimales, que por conveniencia para su entendimiento han sido aproximados al entero (en el conjunto de números naturales \mathbb{N}) más próximo.

CAPÍTULO 6: PROPUESTA DE APOYO TECNOLÓGICO

6.1 ARQUITECTURA TECNOLÓGICA

6.1.3 PROTOTIPO FUNCIONAL DESARROLLADO

La solución ideada utiliza la integración con los actuales sistemas de la compañía, los que cuentan con tecnologías pagadas como Azure SQL Database y Power BI Pro Services. En este sentido, se plantea el diagrama conceptual de sistemas como se muestra en la Ilustración 30, donde se propone la obtención de datos a partir de Azure SQL Database, sistema que dispone de la información de la compañía actualizada diariamente sobre ventas, stock, estados de resultados, entre otros, según se muestra en el Anexo A: Información disponible en Azure SQL Database.

Realizada la consulta se obtienen datos de entrada para la modelación a través del lenguaje de programación *Python*, con archivo que se ejecutan diariamente y de forma automática a través de Azure Function, una funcionalidad de Azure que permite realizar acciones de forma programada.

Como resultado del proceso anterior, se obtiene los datos de salida que corresponden a transformaciones y agrupaciones de los datos de entrada, en adición a los datos obtenidos a través de la modelación desarrollada. Estos datos son automáticamente dispuestos de vuelta en el Azure SQL Database, en la utilización de la misma base de datos, de manera de no adicionar costos al proyecto, con tablas dispuesta para el propósito de este.

Finalmente, los datos son consultados a través de Power BI para exhibirlos de forma interna en la compañía y ser consultados a través de aplicaciones afines y tecnologías conectadas a la web.

A modo de representación de todo lo anteriormente descrito, en la Ilustración 30 se muestra un diagrama que conceptualiza la propuesta de solución en el contexto de integración de sistemas y tecnologías.

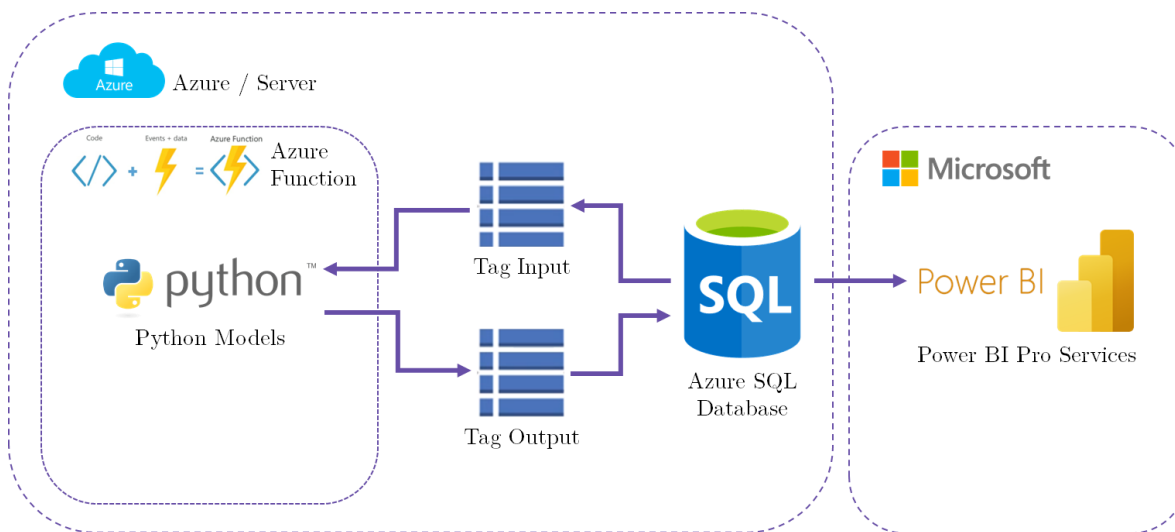


Ilustración 30: Diagrama de integración de sistemas y tecnologías.

Particularmente, para el prototipo funcional desarrollado se utiliza la última fecha sobre la data prevista, es decir el 31 de agosto del 2021. Por lo mismo, se predicen la demanda para las fechas 12 de septiembre, 19 de septiembre, 26 de septiembre y 03 de octubre del año 2021.

6.1.4 VISUALIZACIÓN DE DATOS

Como se plantea en el proceso rediseñado TO-BE, la solución debe contar con una plataforma de visualización de resultados de pronóstico que logre informar al Área de Operaciones de la venta pronosticada en el futuro. Bajo esta necesidad se plantea una solución como se visualiza en la Ilustración 31, a continuación.

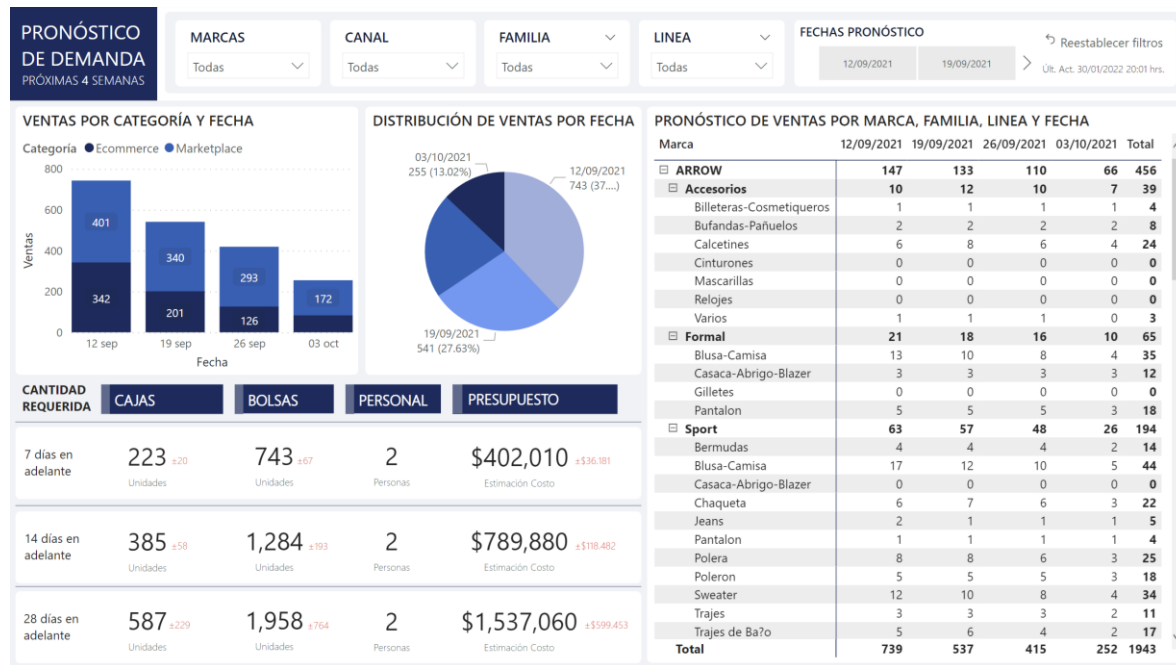


Ilustración 31: Plataforma de visualización de resultados de pronóstico en PBIX.

La propuesta anterior se logra separar en 4 sectores mostrados en la Ilustración 32, en dónde se muestra una primera área (1) que contiene los filtros interactivos con los gráficos y tablas de contenido sobre las variables de marca, canal, familia, línea y fechas; junto con la fecha de última actualización y un botón para reestablecer filtros.

En segundo lugar (2), se tiene una visualización que muestra la historia de pronósticos dentro del mes proyectado con separación de canal E-Commerce y Marketplace, además de una distribución de ventas según fecha.

Por su parte, la tercera área (3) muestra una tabla que consolida esta información según marca, familia, línea y fecha.

Finalmente, la cuarta área (4) otorga el resultado de estimación según 3 temporalidades: 7 días en adelante, 14 días en adelante y 28 días en adelante. Cada temporalidad muestra la transformación estimada respectiva hacia cajas, bolsas, personal y presupuesto requerido según los resultados de cada modelamiento. Adicionalmente, se muestra a continuación de los valores estimados el error correspondiente según el *accuracy* obtenido para cada semana: 9% hacia 1 semana en adelante y 39% para un mes en adelante.

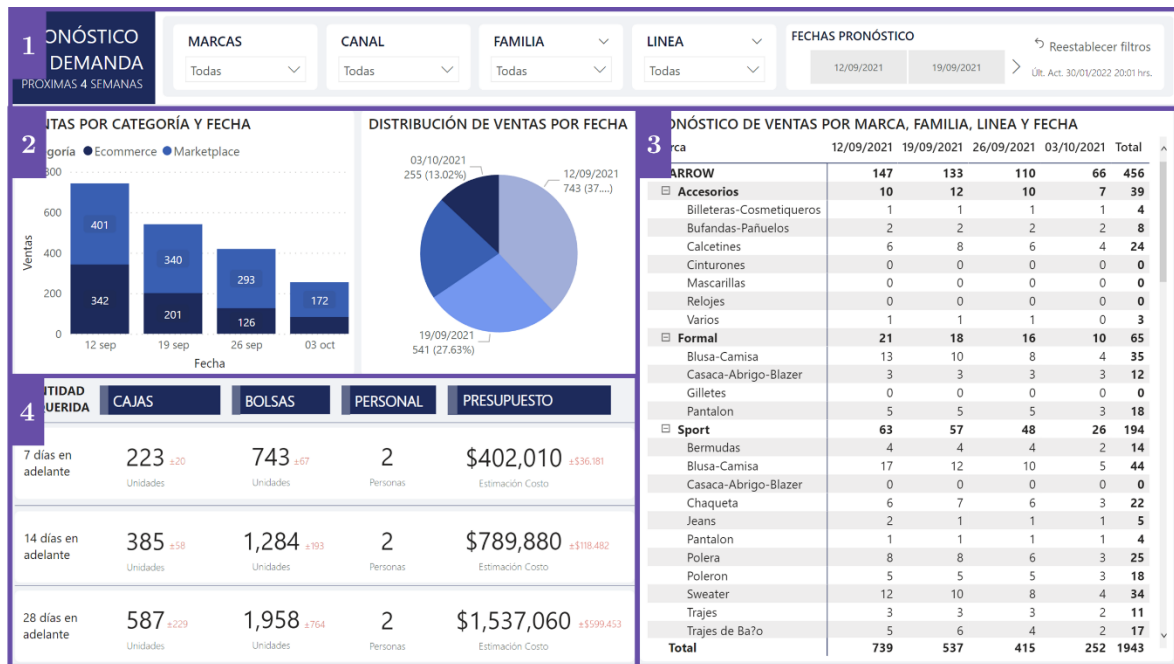


Ilustración 32: Plataforma de visualización de resultados de pronóstico en PBIXX. Separación por áreas.

Particularmente, sobre la estimación de costos se utilizan aproximaciones para su representación, descritas en la Tabla 16.

Tabla 16: Aproximaciones costos y unidades de insumo según demanda

Item	Unidades	Costo
Cajas	3 por cada 10 ventas	\$ 100 por caja
Bolsas	1 por cada venta	\$ 40 por bolsa
Personal	1 persona por cada 1.680 unidades mensuales	\$ 700.000 por persona
	1 persona por cada 840 unidades bisemanales	\$ 350.000 por persona
	1 persona por cada 420 unidades semanales	\$ 175.000 por persona

CÁPITULO 7: PROPUESTA DE IMPLEMENTACIÓN

Para consolidar el uso de la tecnología propuesta dentro de los procesos seleccionados se requieren una serie de actividades que faciliten la integración, principalmente junto a equipos humanos, de manera que se destacan las siguientes tareas a desarrollar:

- Reuniones de trabajo junto a analistas de datos y técnicos informáticos dentro de la compañía RCSA:

Explicación de beneficios de la tecnología, datos utilizados y necesidades para generar continuidad dentro la compañía del sistema desarrollado. Se propone una única instancia que reúna las personas clave.

- Reuniones con usuarios claves del proceso:

Se destaca la Gerencia de Operaciones de la compañía y todo su equipo entre los involucrados relevantes. Se intenta generar convencimiento sobre la herramienta, explicando su modo de utilización, frecuencia y advertencias sobre la confiabilidad que tiene esta tecnología para tomas de decisiones en todos los niveles. Se propone una única instancia junto a personas clave.

- Presentación e implementación final:

Se dispone oficialmente de la tecnología y explica sus modos de acceder y funcionalidades. Se propone una única instancia junto a personas claves e interesados externos.

Esta propuesta de implementación requiere de no más de 10 horas de trabajo, a través de reuniones ejecutivas y proyectadas en el plazo de un mes.

7.1 GESTIÓN DEL CAMBIO

Para poder llevar a cabo la implementación resulta conveniente asentarse en algunos tópicos que facilitan la gestión del cambio (Olguín, 2005):

- Liderazgo del proyecto: generar y establecer los liderazgos en las diferentes etapas del proyecto, empoderando a los usuarios, otorgando responsabilidades y deberes, se definen y comunican las reglas del juego. Para el caso de implementación del proyecto, se propone utilizar el liderazgo interno del Gerente de Operaciones de RCSA.

- Gestión del poder: en todo proceso de cambio se debe administrar el poder principalmente en dos puntos: generar los incentivos para mantener el proyecto vivo por parte del líder del proyecto y administrar los cambios de poderes generados por el cambio llevado a cabo. Se propone para este caso una distribución de roles consensuada, que separe la capacidad técnica de operación, del poder de cambio y convencimiento de las partes interesadas con el proyecto.
- Cambios y conservación: comunicar y registrar los cambios que se llevarán a cabo dejando explícito lo que se conservará en este proceso y lo que cambiará, condiciones del cambio y reglas del juego. Se propone comunicar exactamente cuales son los cambios, en torno a responsabilidades y posibilidades, sobre el proceso a rediseñar.
- Comunicaciones: mantener las comunicaciones vivas y expeditas es una tarea constante en el proyecto, se deben definir los canales, medios y personas oficiales para comunicar. Se debe permitir canales de comunicación claramente establecidos entre el desarrollo técnico y los usuarios de la herramienta de pronóstico de ventas.

CÁPITULO 8: EVALUACIÓN DEL PROYECTO

8.1 EVALUACIÓN TÉCNICA

Acorde a la propuesta ubicada en la sección 5.2.3, el modelo seleccionado cuenta con una métrica de precisión, *Accuracy*, de 91% en el pronóstico de una semana inmediata, generando predicciones asertivas en forma y tiempo a la realidad como se observar en la Ilustración 33, a continuación.



Ilustración 33: Comparación de ventas versus pronóstico de ventas

A modo de evaluación del pronóstico de ventas de un mes completo se utiliza el último mes de registro de ventas dentro del data set, agosto del 2021. Como resultado del pronóstico se cuenta con un *Accuracy* de 69.17%, lo que responde a las expectativas debido a que el error de pronóstico se va acumulando semana por semana. Como resultado, se muestra la Ilustración 34.

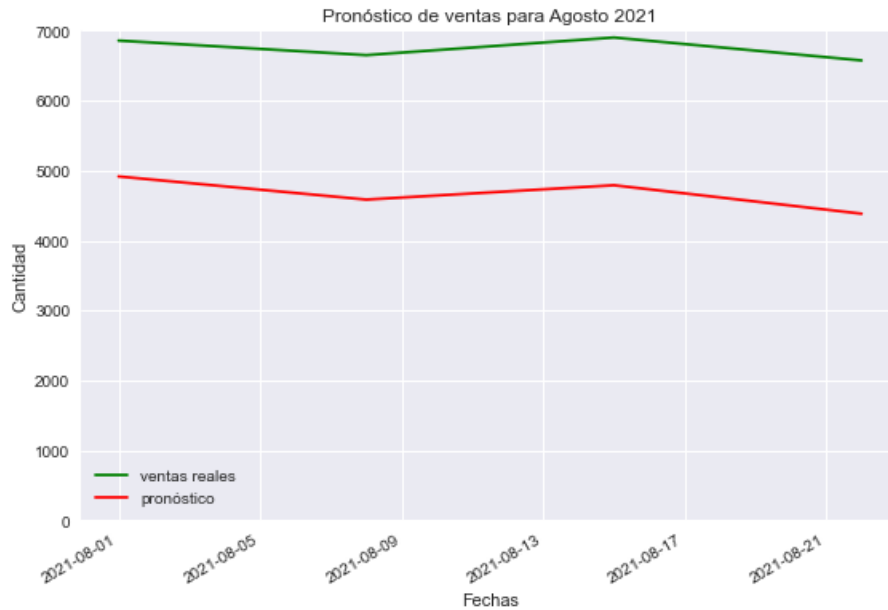


Ilustración 34: Pronóstico de ventas versus datos reales para 01, 08, 15, 22 de agosto del 2021.

Es posible observar que el pronóstico de ventas conserva la curva de datos reales, manteniéndose aproximadamente 2000 unidades por debajo de la curva real. Así también, es posible visualizar a través de la Ilustración 35 que los datos se encuentran por debajo del promedio de los datos para el año 2021. Sin embargo, esto puede hacer mayor sentido en vista del universo de datos (ver Ilustración 36), cuyo promedio se aproxima mucho más al pronóstico para esta fecha en particular.

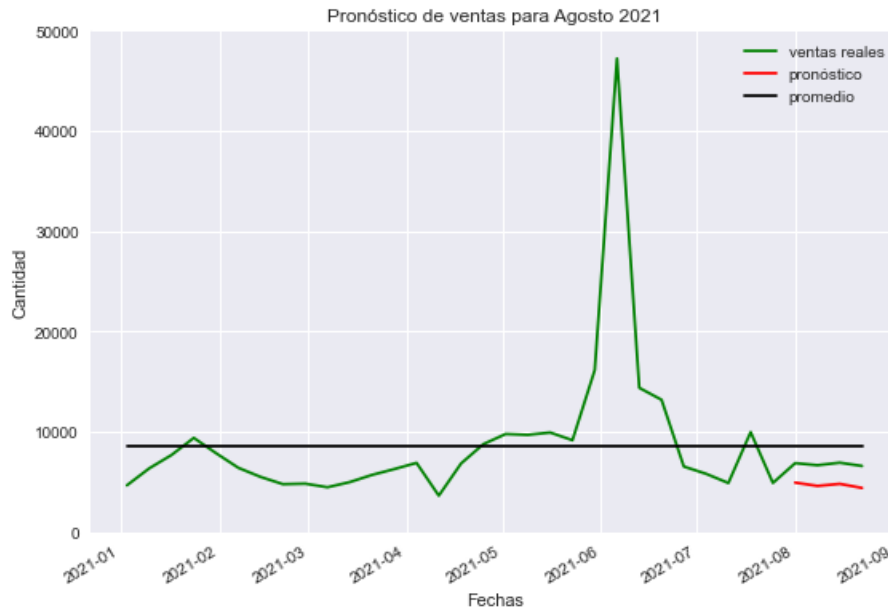


Ilustración 35: Pronóstico de ventas versus datos reales y promedio para 01, 08, 15, 22 de agosto. Perspectiva comparada frente a año 2021



Ilustración 36: Pronóstico de ventas versus datos reales y promedio para 01, 08, 15, 22 de agosto. Perspectiva comparada frente a universo de fechas

Analizar estas fechas en particular refleja realizar un análisis del peor pronóstico dentro de la serie de fechas disponible en los datos (ver Ilustración 33: Comparación de ventas versus

pronóstico de ventas), lo que a su vez refleja el cuidado que se debe tomar al momento de interpretar y hacer uso de estos resultados.

En general, se sugiere tomar como referencia la primera semana de pronóstico por sobre el resto, ya que esta cuenta con una mayor asertividad como se ha podido mostrar en los resultados obtenidos.

8.2 EVALUACIÓN ECONÓMICA

El proyecto no tiene un impacto económico directo, y sólo se pueden estimar sus ahorros (beneficios económicos) una vez en funcionamiento. No obstante, es posible fijar algunos supuestos realistas con tal de calcular el impacto económico en un plazo dado, de forma que ante datos reales el ejercicio pueda reemplazarse fácilmente.

8.2.3 DEFINICIÓN DE BENEFICIOS Y COSTOS

Definamos el aumento de la capacidad de cumplimiento como la posibilidad de responder operacionalmente a la demanda en un período dado, de manera análoga entre cajas, bolsas y personal para el envío de productos.

Supongamos que el proyecto logra mejorar la capacidad de cumplimiento en un 100% de manera inmediata desde su entrada en funcionamiento, cumplimiento las promesas de venta de la compañía en torno a la fecha de entrega de productos de manera absoluta.

Además, supongamos que por cada cliente que queda conforme con el envío de productos se percibe un beneficio económico de \$1.000 CLP/mes, precio que los clientes están dispuestos a pagar por un envío en un tiempo de 1 a 3 días después de realizada la compra con cumplimiento del 100% de seguridad.

Supongamos que cada cliente compra 4 prendas por mes, en una sola compra. Por lo que, manteniendo el promedio de ventas semanal sostenido entre 01 de enero y 31 de agosto del año 2021 de 8.623 productos, en meses de 4 semanas y 3 días se comercializan 38.188 productos y, por tanto, se cuenta con 9.547 clientes cada mes.

En vista de la cantidad de clientes y el beneficio económico que los clientes se encuentran dispuestos a pagar, se cuenta con un beneficio por mes de \$9.547.000 CLP.

Por su parte, los costos del proyecto involucran exclusivamente la contratación de tecnologías no disponibles actualmente por parte de la compañía, que como se define en la sección 6.1

Arquitectura Tecnológica, corresponde exclusivamente a la contratación de Azure Function en su versión básica⁶, cuyo valor mensual corresponde a \$18 USD o \$14.413 CLP/mes⁷.

Supongamos también un costo estimado por soporte de la tecnología de \$45.000 CLP/mes, por motivos de mantener actualizada la tecnología y atender desarrollos específicos.

Dada la naturaleza del proyecto, no se requieren inversiones iniciales; ni gasto en capital de trabajo debido a que tampoco se requiere destinar de personal exclusivo al uso y mantenimiento de la herramienta; ni inversión en activos fijos ya que se dispone de, por ejemplo, los dispositivos tecnológicos ya para hacer uso de la solución propuesta.

8.2.4 FLUJO DE CAJA

En vista de la definición de beneficios y costos anterior, el flujo de caja se muestra en la

Tabla 17 y Tabla 18. Se destaca un resultado de \$92.083.884 para el primer año de operación según los supuestos y especificaciones dadas en la sección anterior.

Cabe destacar que el resultado económico anterior se posibilita en el caso hipotético que la empresa, Ropas Camargo S.A, deseara percibir ingresos por un servicio sin retrasos. No obstante, el fin del proyecto no persigue en sí este objetivo, si se puede concluir que por cobrar un sobre precio desde \$8 CLP por cliente, el proyecto tiene un resultado económico positivo como se muestra en el Anexo B: Flujo de caja con valor mínimo viable.

⁶ <https://azure.microsoft.com/en-us/pricing/details/functions/>. Consultado al 31/01/2022

⁷ 1 USD es equivalente a \$800 CLP. Consultado valor al 31/01/2022

Tabla 17: Flujo de caja mes 1 a 6

1.0 Detalle de Ingresos	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6
1.1 Ingresos por cobranza extra de envío	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000
Cantidad	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547
Precio	\$1,000	\$1,000	\$1,000	\$1,000	\$1,000	\$1,000
1.0 Total de Ingresos	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000
1.0 Total de Ingresos después de Impuestos (IVA 19%)	\$7,733,070	\$7,733,070	\$7,733,070	\$7,733,070	\$7,733,070	\$7,733,070
2.0 Detalle de Egresos						
2.1 Servicios tecnológicos	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure Functions	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure SQL Server	\$0	\$0	\$0	\$0	\$0	\$0
Power BI Pro	\$0	\$0	\$0	\$0	\$0	\$0
2.2 Soporte de tecnología	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000
2.3 Costos Fijos	\$0	\$0	\$0	\$0	\$0	\$0
Salarios de personal	\$0	\$0	\$0	\$0	\$0	\$0
Dispositivos tecnológicos	\$0	\$0	\$0	\$0	\$0	\$0
2.0 Total de Egresos	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413
3.0 Flujo del mes	\$7,673,657	\$7,673,657	\$7,673,657	\$7,673,657	\$7,673,657	\$7,673,657
4.0 Total acumulado	\$7,673,657	\$15,347,314	\$23,020,971	\$30,694,628	\$38,368,285	\$46,041,942

Tabla 18: Flujo de caja mes 7 a 12

1.0 Detalle de Ingresos	Mes 7	Mes 8	Mes 9	Mes 10	Mes 11	Mes 12
1.1 Ingresos por cobranza extra de envío	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000
Cantidad	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547
Precio	\$1,000	\$1,000	\$1,000	\$1,000	\$1,000	\$1,000
1.0 Total de Ingresos	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000	\$9,547,000
1.0 Total de Ingresos después de Impuestos (IVA 19%)	\$7,733,070	\$7,733,070	\$7,733,070	\$7,733,070	\$7,733,070	\$7,733,070
2.0 Detalle de Egresos						
2.1 Servicios tecnológicos	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure Functions	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure SQL Server	\$0	\$0	\$0	\$0	\$0	\$0
Power BI Pro	\$0	\$0	\$0	\$0	\$0	\$0
2.2 Soporte de tecnología	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000
2.3 Costos Fijos	\$0	\$0	\$0	\$0	\$0	\$0
Salarios de personal	\$0	\$0	\$0	\$0	\$0	\$0
Dispositivos tecnológicos	\$0	\$0	\$0	\$0	\$0	\$0
2.0 Total de Egresos	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413
3.0 Flujo del mes	\$7,673,657	\$7,673,657	\$7,673,657	\$7,673,657	\$7,673,657	\$7,673,657
4.0 Total acumulado	\$53,715,599	\$61,389,256	\$69,062,913	\$76,736,570	\$84,410,227	\$92,083,884

CAPITULO 9: CONCLUSIONES Y TRABAJOS FUTUROS

El área de Business Analytics ofrece un gran potencial a aquellas organizaciones que administran grandes volúmenes de datos. Mediante la utilización de técnicas especializadas de Deep Learning o Machine Learning es posible convertir esos datos en conocimiento útil para apoyar la toma de decisiones tanto estratégicas como operacionales.

En el contexto de COVID-19, Ropas Camargo S.A no es ajena a la presión ejercida por el mercado para buscar la adaptación constante en búsqueda de generación de eficiencias en sus procesos. En este sentido, la puesta en acción de actividades con apoyo de pronósticos de demanda puede mejorar la capacidad operacional de la empresa, implicando reducción de costos y mejoras en su reputación percibida entre los clientes en el largo plazo al dar cumplimiento a sus promesas comerciales.

Se concluye que con los avances del proyecto se consiguen exitosamente los 5 objetivos propuestos:

- Se identifica y utiliza información relevante para analizar la demanda actual
- Se utilizan indicadores claves de desempeño
- Se genera un proceso de analítica para la estimación de demanda dentro del proceso de planificación y adquisición de recursos de suministro
- Se implementa un modelo, a través del método Random Forest, para la predicción de demanda
- Se visualizan los resultados del modelo en plataforma de Power BI, conectada con los servicios de Azure SQL Server

De manera general, el proyecto demuestra tener una factibilidad económica a través de una mínima suma de dinero sobre el precio de envío de los productos de \$8 CLP/mes por cliente, junto con la posibilidad de crear un negocio económicamente favorable para la compañía con una generación de ingresos de más de \$92 millones CLP/año por el cobro de \$1000 CLP/mes por cliente al ofrecer un servicio de envío de productos confiable de cara a sus consumidores.

Además, la tecnología desarrollada constituye un avance significativo para la empresa en torno a las herramientas que esta dispone para permitir un mejor rendimiento operacional, permitiéndose extender beneficios incluso sobre áreas anexas, como comercial y de marketing en

la creación de promociones y fijación de precios para fortalecer la venta de categorías que proyectan una menor venta.

No obstante, debido a la característica primeriza de la herramienta dentro de la compañía en un desarrollo acorde a metodologías de ciencia de datos, se establece exitosamente un proceso de negocios que resguarde resultados coherentes, combinando la inteligencia de negocios propia de los colaboradores de la compañía con la información relevante para la toma de decisiones sobre los niveles de operación, y la compra y contratación de insumos de suministro.

9.1 TRABAJOS FUTUROS

9.1.1 MODELACIÓN INTEGRADA

Se propone la integración de variables de precios y variables comerciales. El modelo actual tiene la posibilidad de actuar reactivamente ante las situaciones de mercado, utilizando información sobre ventas pasadas para su ajuste. En un modelo integrado de predicción de la demanda se plantea utilizar información comercial prevista por las áreas responsables ya que, en definitiva, el conocimiento de negocios indica que el factor decisivo en la demanda de productos es el precio de estos. Se puede tomar como ejemplo el caso de Cyber Day, donde dos factores son preponderantes: una mayor exposición al mercado, y clientes buscando precios bajos para su conveniencia.

Entre las variables que podrían utilizarse se resumen: Precio de productos, Promociones disponibles, Índice de visibilidad de mercado, Eventos especiales (como Cyber Days o Black Fridays), entre otras.

9.1.2 PREDICCIÓN DETALLADA

Uno de los alcances de la herramienta desarrollada corresponde al nivel de detalle de predicción que la solución ofrece. En un trabajo futuro se propone la exploración de un modelamiento predictivo a un nivel de producto para recoger estacionalidades y características específicas de estos, tales como el color y la talla, entre otras características.

Para ello, la mayor adversidad explorada dentro del proyecto tiene que ver con la continuidad de datos de este nivel, cuyos identificadores por producto son en la mayoría de los casos utilizados por temporada. Sin embargo, la posibilidad se encuentra en la exploración de tablas de información ya dispuestas en línea por la compañía sobre información detallada de productos, cuales por limitaciones de tiempo no fueron tomadas en cuenta para el desarrollo de este proyecto.

Uno de los beneficios principales de una modelación detallada a nivel de productos se encuentra en la oportunidad de ampliar el alcance del proyecto al canal online, posibilitando movimientos de productos entre las bodegas de almacenamiento y las tiendas, inclusive pudiendo optimizar los tiempos de entrega hacia los clientes.

9.1.3 MAYOR AMPLITUD DE DATOS

Entre las posibilidades para mejorar la predicción de ventas se comprende utilizar una mayor amplitud de datos para el entrenamiento de los modelos, lo que se permite naturalmente con un mayor tiempo de operación de la tecnología o al integrar información externa y complementaria al negocio.

Por ejemplo, uno de los trabajos futuros corresponde a mejorar la predicción de eventos especiales de venta como Cyber Day y Black Friday, cuya presencia en la base de datos actualizada no comprende más del 5% de los datos totales.

Entre las alternativas se encuentran el generar entrenamientos de modelos con datos de negocios que asemeje el retail de ropa, o inclusión de orígenes de datos de retail de ropa provenientes de externos, como la que pueden ofrecer otros E-Commerce y Marketplace.

Como ya se ha mencionado, el beneficio de este trabajo futuro es mejorar la exactitud de las predicciones y ampliar la fiabilidad ante eventos comerciales especiales.

BIBLIOGRAFÍA

- APQC. (21 de Enero de 2020). APQC Process Classification Framework (PCF) - Cross Industry - Excel Version 7.1.0 - Spanish. Obtenido de <https://www.apqc.org/resource-library/resource-listing/apqc-process-classification-framework-pcf-cross-industry-excel-8>
- Barros, O. (2000). *Rediseño de procesos mediante el uso de patrones*. Santiago: Dolmen.
- Barros, O. (2015). *Ingeniería de Negocios: Diseño Integrado de Servicios, sus Procesos y Apoyo TI*. Santiago: Amazon, Kindle Edition.
- CAIGG. (2016). *Implantación, mantención y actualización del proceso de gestión de riesgos en el sector público*. Obtenido de http://www.oas.org/juridico/PDFs/mesicic5_cl_insitu_DNSC_ane8.pdf.
- Chambers, J. W. (1983). *Graphical Methods for Data Analysis*. Wadsworth.
- Claudimar Pereira Da Veiga, C. R. (2014). Demand forecasting in food retail: a comparison between the Holt-. *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS*. vol. 11, 7.
- E. Izquierdo-Verdiguier and R. Zurita-Milla. (2020). An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* vol. 88., 13.
- Fuat Şimşir. (2021). Machine Learning Algorithms with Intermittent Demand Forecasting: An Application in Retail Apparel with Plenty of Predictors. *TEKSTİL VE KONFEKSİYON*. vol. 31., 12.
- Gallardo, J. (2018). *Metodología para el desarrollo de proyectos en Minería de Datos CRIPS-DM*. Obtenido de http://www.oldemarrodriguez.com/yahoo_site_admin/assets/docs/Documento_CRISP-DM.2385037
- Gravesen, J. K. (05 de noviembre de 2013). *IBM*. Obtenido de A macro-pattern for public sector systems architecture: <https://www.ibm.com/developerworks/rational/library/macro-pattern-public-sector-systems-architecture/index.html?ca=drs->

- Instituto Nacional de Estadísticas, C. (19 de Octubre de 2021). *Chile Retail Sales YoY*. Obtenido de Trading Economics: <https://tradingeconomics.com/chile/retail-sales-annual>
- IPS. (2019). *Resumen Ejecutivo Cuenta Pública 2019*. Obtenido de https://www.ips.gob.cl/servlet/internet/archivo/1421810612662/resumen_ejecutivo_2019.pdf
- Lever, G. (4 de Noviembre de 2020). *Perspectivas Económicas del sector comercio 2021. Comité del Comercio Electrónico*. Obtenido de Cámara de Comercio de Santiago: <https://www.ecommerceccs.cl/perspectivas-economicas-del-sector-comercio-2021/>
- Lever, G. (22 de Abril de 2021). *Comité de Comercio Electrónico*. Obtenido de Cámara de Comercio de Santiago: <https://www.ecommerceccs.cl/chile-perspectivas-del-comercio-electronico-2021/>
- Modernizacion. (Julio de 2019). *Agenda de Modernización del Estado*. Obtenido de Modernizacion: https://cdn.digital.gob.cl/filer_public/d3/e3/d3e3bb10-4ad2-4df8-adfa-b4ff69a658b6/agenda-de-modernizacion-del-estado.pdf
- OEC. (26 de Noviembre de 2021). *The Observatory of Economic Complexity (OEC)*. Obtenido de <https://oec.world/es/profile/country/chl#yearly-imports>
- Olguín, E. (2005). *CHESS Modelo Integral de Liderazgo y Gestión del Cambio*. Obtenido de Academia: https://www.academia.edu/38704938/CHESS_MODELO_INTEGRAL_DE_LIDERA_ZGO_Y_GESTI%C3%B3N_DEL_CAMBIO
- Qamar, U., & Raza, M. S. (2020). *Data Science Concepts and Techniques with Applications* (First ed.). Springer Singapore.
- Şimşir, İ. G. (2020). Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ANN) and support vector machines (SVM) methods. *Science Direct*, 9.
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. *4th international conference on the practical applications of knowledge discovery and data mining* (págs. 29-39). London, UK: Springer-Verlag.

ANEXOS

ANEXO A: INFORMACIÓN DISPONIBLE EN AZURE SQL DATABASE

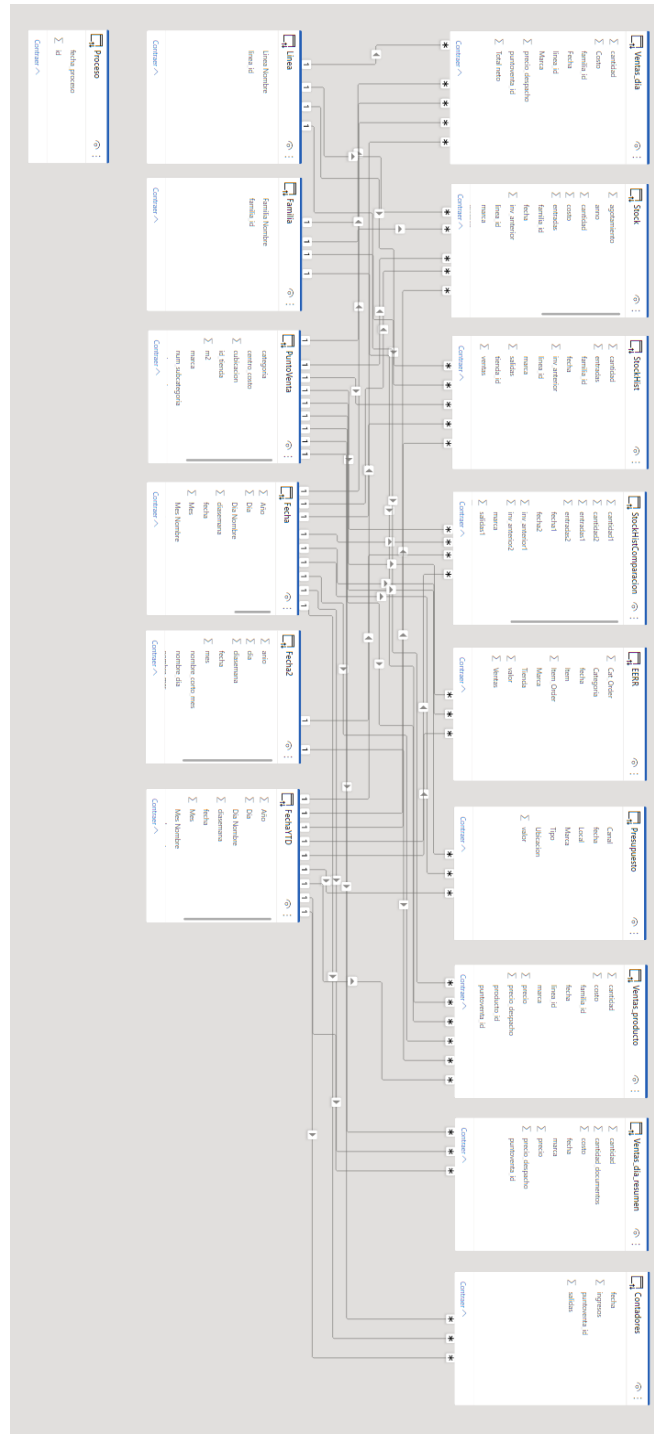


Ilustración 37: Tablas disponibles para consulta a Azure SQL Database al ingreso

ANEXO B: FLUJO DE CAJA CON VALOR MÍNIMO VIABLE

Tabla 19: Flujo de caja con valor mínimo viable. Mes 1 a 6

1.0 Detalle de Ingresos	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6
1.1 Ingresos por cobranza extra de envío	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376
Cantidad	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547
Precio	\$8	\$8	\$8	\$8	\$8	\$8
1.0 Total de Ingresos	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376
1.0 Total de Ingresos después de Impuestos (IVA 19%)	\$61,865	\$61,865	\$61,865	\$61,865	\$61,865	\$61,865
2.0 Detalle de Egresos						
2.1 Servicios tecnológicos	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure Functions	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure SQL Server	\$0	\$0	\$0	\$0	\$0	\$0
Power BI Pro	\$0	\$0	\$0	\$0	\$0	\$0
2.2 Soporte de tecnología	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000
2.3 Costos Fijos	\$0	\$0	\$0	\$0	\$0	\$0
Salarios de personal	\$0	\$0	\$0	\$0	\$0	\$0
Dispositivos tecnológicos	\$0	\$0	\$0	\$0	\$0	\$0
2.0 Total de Egresos	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413
3.0 Flujo del mes	\$2,452	\$2,452	\$2,452	\$2,452	\$2,452	\$2,452
4.0 Total acumulado	\$2,452	\$4,903	\$7,355	\$9,806	\$12,258	\$14,709

Tabla 20: Flujo de caja con valor mínimo viable. Mes 7 a 12

1.0 Detalle de Ingresos	Mes 7	Mes 8	Mes 9	Mes 10	Mes 11	Mes 12
1.1 Ingresos por cobranza extra de envío	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376
Cantidad	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547	\$9,547
Precio	\$8	\$8	\$8	\$8	\$8	\$8
1.0 Total de Ingresos	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376	\$76,376
1.0 Total de Ingresos después de Impuestos (IVA 19%)	\$61,865	\$61,865	\$61,865	\$61,865	\$61,865	\$61,865
2.0 Detalle de Egresos						
2.1 Servicios tecnológicos	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure Functions	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413	-\$14,413
Azure SQL Server	\$0	\$0	\$0	\$0	\$0	\$0
Power BI Pro	\$0	\$0	\$0	\$0	\$0	\$0
2.2 Soporte de tecnología	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000	-\$45,000
2.3 Costos Fijos	\$0	\$0	\$0	\$0	\$0	\$0
Salarios de personal	\$0	\$0	\$0	\$0	\$0	\$0
Dispositivos tecnológicos	\$0	\$0	\$0	\$0	\$0	\$0
2.0 Total de Egresos	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413	-\$59,413
3.0 Flujo del mes	\$2,452	\$2,452	\$2,452	\$2,452	\$2,452	\$2,452
4.0 Total acumulado	\$17,161	\$19,612	\$22,064	\$24,516	\$26,967	\$29,419