

Tipo de documento: Tesis de maestría

Master in Management + Analytics

Pronóstico de demanda automatizado para empresa distribuidora de alimentos. Comparación de técnicas de pronóstico avanzadas para mejorar la precisión del modelo de pronóstico de demanda.

Autoría: Pacher, Matías Ezequiel

Fecha de defensa de la tesis: 2023

¿Cómo citar este trabajo?

Pacher, M. (2023) "Pronóstico de demanda automatizado para empresa distribuidora de alimentos. Comparación de técnicas de pronóstico avanzadas para mejorar la precisión del modelo de pronóstico de demanda.". [*Tesis de maestría. Universidad Torcuato Di Tella*]. Repositorio Digital Universidad Torcuato Di Tella. <https://repositorio.utdt.edu/handle/20.500.13098/12037>

El presente documento se encuentra alojado en el Repositorio Digital de la Universidad Torcuato Di Tella bajo una licencia Creative Commons Atribución-No Comercial-Compartir Igual 2.5 Argentina (CC BY-NC-SA 2.5 AR)
Dirección: <https://repositorio.utdt.edu>



UNIVERSIDAD
TORCUATO DI TELLA

MASTER IN MANAGEMENT + ANALYTICS

Pronóstico de demanda automatizado para empresa distribuidora de alimentos

Comparación de técnicas de pronóstico avanzadas para mejorar la precisión del
modelo de pronóstico de demanda.

Matias Ezequiel Pacher

Tutora: Teresita Ines Cancelo

Resumen

La predicción precisa de la demanda es fundamental para la gestión eficiente de la cadena de suministro de cualquier empresa, y en particular, para una empresa distribuidora de alimentos.

En este proyecto de tesis, se propone desarrollar un modelo de pronóstico de demanda más preciso y automatizado para esta empresa, que actualmente utiliza un método de media móvil con una gran componente manual.

Se compararán dos técnicas de pronóstico avanzadas: Prophet y ARIMA, y se evaluará su eficacia en la predicción de la demanda de la empresa.

El objetivo es proporcionar una solución más precisa y automatizada, lo que permitirá mejorar la toma de decisiones sobre la planificación de la cadena de suministro y la gestión de inventarios.

Los resultados de este proyecto de tesis tienen el potencial de mejorar significativamente la eficiencia y rentabilidad de la empresa distribuidora de alimentos en el mercado altamente competitivo de la industria alimentaria.

Abstract

Precise demand forecasting is essential for the efficient management of the supply chain of any company, and in particular, for a food distribution company.

In this thesis project, it is proposed to develop a more accurate and automated demand forecasting model for this company, which currently uses a moving average method with a large manual component.

Two advanced forecasting techniques will be compared: Prophet and ARIMA, and their effectiveness in predicting company demand will be evaluated.

The goal is to provide a more accurate and automated solution, which will enable better decision-making on supply chain planning and inventory management.

The results of this thesis project have the potential to significantly improve the efficiency and profitability of the food distribution company in the highly competitive market of the food industry.

Índice de Contenido.

I. Introducción.....	7
1. Contexto.....	7
2. Problema.....	9
3. Objetivos de la investigación.....	9
II. Marco Teórico.....	11
1. Pronósticos.....	11
1. Horizonte de tiempo.....	11
1.1. Corto plazo.....	11
1.2. Mediano plazo.....	11
1.3. Largo plazo.....	11
2. Procedimiento.....	12
2.1. Cualitativo.....	12
2.2. Cuantitativos.....	12
3. Nivel.....	13
3.1. Estratégico.....	13
3.2. Táctico.....	13
3.3. Operativo.....	13
4. Agrupación.....	14
4.1. Agregado.....	14
4.2. Productos.....	15
2. Patrones en series de tiempo.....	17
1. Modelos predictivos cuantitativos.....	18
1. Univariados.....	18
1.1. Métodos de Suavización.....	18
1.2. Modelos de Descomposición de Series de Tiempo.....	19
2. Modelos Multivariados.....	22
2.1. Métodos Lineales.....	22
2.2. Modelos No Lineales y de Aprendizaje Automático.....	23
3. Consideraciones sobre el Overfitting y el Costo Computacional.....	23
3.1. Sobreajuste (Overfitting):.....	24
3.2. Optimización y Calibración:.....	25
3.3. Costo Computacional:.....	25
4. Significatividad en los resultados.....	26
4.1. Independiente o pareada.....	27
4.2. Paramétrica o no paramétrica.....	27
5. Estado del arte.....	27
5.1. Evolución de utilización de técnicas de pronósticos de ventas (1984-2006).....	27

5.2. Nuevos métodos.....	29
5.3. Consideraciones sobre el estado actual de los metodos de pronostico.....	30
III. Metodología.....	32
1. Datos.....	32
1.1. Forma y transformación del dataset.....	32
1.2. Análisis descriptivo.....	34
1.3. Separación de entrenamiento y testeo.....	46
1.3.1. Tipo de pronóstico Fuera de Muestra (Out of Sample).....	47
2. Método utilizado por la empresa.....	47
3. Métodos propuestos.....	48
3.1. Prophet	48
3.2. ARIMA.....	51
4. Diseño e implementación.....	53
4.1. Métrica de error.....	53
4.2. Métrica de éxito.....	54
4.3. Búsqueda de grilla.....	54
4.4. Tipo de pronóstico Fuera de Muestra (Out of Sample).....	56
4.5. Tratamiento de pandemia COVID 19.....	57
IV. Resultados.....	58
1. Tabla comparativa de resultados.....	58
2. Significancia estadística de la diferencia en los errores.....	59
3. Gráficos.....	60
3.1. Prophet.....	61
3.2. ARIMA.....	67
V. Implementación y calibración.....	73
VI. Conclusiones.....	74
VII. Futuras líneas de investigación.....	75
Referencias.....	76

Índice de Figuras

Figura 1. Ventas totales por mes medidas en unidades monetarias.....	35
Figura 2. Mapa de calor de ventas totales por mes medidas en unidades monetarias.....	36
Figura 3. Variación de precio de los 11 principales productos.....	37
Figura 4. Variación de precio de los 11 principales productos deflactados.....	38
Figura 5. Ventas totales por mes medidas en unidades monetarias deflactadas.....	39
Figura 6. Mapa de calor de ventas totales por mes medidas en unidades monetarias deflactados.....	40

Figura 7. Ventas totales por mes medidas en unidades físicas.....	40
Figura 8. Productos más vendidos.....	41
Figura 9. Distribución mensual de los 11 productos más vendidos.....	42
Figura 10. Distribución de unidades vendidas por mes.....	43
Figura 11. Unidades vendidas totales por día.....	44
Figura 12. Distribución de ventas diarias de 10 productos más vendidos.....	45
Figura 13. Top 3 productos vendidos por día de la semana.....	46
Figura 14. Componentes de pronóstico Prophet.....	61
Figura 15. Comparación de ventas reales y predichas por Prophet.....	63
Figura 16. Relación de errores entre modelo Prophet y Benchmark.....	64
Figura 17. Relación entre dinero inmovilizado Prophet y Benchmark.....	65
Figura 18. Componentes del modelo ARIMA.....	67
Figura 20. Relación de errores entre modelo ARIMA y Benchmark.....	70
Figura 21. Relación entre dinero inmovilizado ARIMA y Benchmark.....	71

Índice de Tablas.

Tabla 1. Tabla comparativa de errores de los 10 primeros productos de la lista.....	53
---	----

I. Introducción

1. Contexto.

La empresa es una distribuidora de productos alimenticios de la ciudad de Parana, Entre Rios; enfocados en locales gastronómicos. Posee más de 15 años en el mercado y su propuesta está compuesta por artículos refrigerados y no refrigerados. Dentro de los mismos, hay fechas de vencimiento dispares, siendo algunas muy extensas y otras de un par de semanas.

El marco en el que se enmarca la empresa tiene dos caras. Esto se debe a su carácter de intermediario en la cadena de abastecimiento, en cuanto a los comercios que atiende y en cuanto a la industria alimenticia de la cual se abastece.

En relación a los comercios que atiende, se ha observado un desarrollo interesante del sector gastronómico desde que inicio la empresa, en parte por ser la capital de la provincia y ver el cambio que viven muchas ciudades en las que pasan de ser un centro productivo a ofrecer mayor cantidad de servicios terciarios, y la creación de diferentes opciones orientadas al consumo, en gran medida alrededor del ocio, el entretenimiento, la cultura y el turismo (Zukin, 1998), constituyéndose así en centros en los cuales se instalan (y desde los cuales se transmiten) tendencias en cuanto a nuevos estilos de vida y consumo (Lew et al., 2014). Estos nuevos habitantes de las ciudades, entonces, que forman parte de un mundo globalizado, observan las tendencias mundiales y de ciudades más grandes, requiriendo, entre estas, propuestas variadas que atraviesan las distintas comidas del día.

En relación a los proveedores de la empresa, los mismos no escapan a la inestabilidad que atraviesa el país, imposibilitando en mayor o menor medida la certeza acerca de si los productos demandados van a ser recibidos o no. Nuevos regímenes de exportación pueden generar el direccionamiento de materias primas hacia el exterior, mientras que trabas a las importaciones

pueden paralizar la producción de ciertos productos indeterminadamente. Además, las inclemencias del clima pueden generar periodos de sequías o inundaciones que disminuyen la producción, como así también su aumento de precios y en algunos casos redireccionamiento de materias primas hacia productos más rentables.

Este desacople entre oferta y demanda se ve incrementado en Argentina, pero no es algo exclusivamente propio del país. Según la Food and Agriculture Organization (FAO), la demanda de los productos perecederos es constante a lo largo del tiempo, no siendo así su producción (Contreras Juarez et al., 2016), lo que trae a consideración la posibilidad de emplearlo como “una estrategia para diferir la oferta de producto hasta que el mercado se encuentre desabastecido y de esta manera obtener mejores precios” (López Camelo, 2003).

Por supuesto que la contracara de esta oportunidad es la posibilidad de quedarse largo (sobrestock) de mercadería, lo que provocaría en algunos productos, en especial aquellos con fecha de vencimiento corta, a la práctica de venderlos a un precio igual o menor al costo, o, en el peor caso, a la pérdida de los mismos. Además, los productos refrigerados tienen menos lugar de guardado que aquellos no refrigerados, por lo cual también existe un costo de oportunidad de conservación en un lugar escaso.

Por otro lado, en sus competidores se pueden diferenciar aquellos que son similares a la empresa, que distribuyen diversas marcas sin exclusividad con ninguna, y aquellas que sí la tienen por lo que solo pueden vender productos de cierta compañía. Dentro de estos, la empresa compra productos que mantiene en stock para atender inmediatamente los pedidos de los clientes, a diferencia de otros que reciben los productos que previamente tienen ya encargados y vendidos, pero que resigna la rapidez al atender la demanda.

En esta breve introducción se ha tratado de reflejar algunas consideraciones que reflejan la importancia de una correcta previsión de las cantidades solicitadas, siendo fundamental para ello un correcto pronóstico de ventas.

2. Problema.

El problema que busca abordar la tesis es el hecho de que la predicción de ventas solo se basa en las unidades vendidas en el periodo inmediatamente anterior que busca predecir, el cual en este caso es de 31 días. Esto puede tener sentido siempre y cuando la demanda sea estable en el tiempo, lo cual es difícil. Además, se abandona la posibilidad de incorporar datos más antiguos que puedan capturar patrones más complejos a través del tiempo.

Lo señalado anteriormente puede culminar en pérdidas de ventas por quedarse corto en mercadería; o inmovilización de dinero o vencimiento de productos por quedarse largo.

La falta de automatización de este proceso provoca pérdidas de tiempo en tareas repetitivas, y constituye un problema a la hora de evaluar la posibilidad de la incorporación de nuevos productos, ya que eleva el costo de la realización de las predicciones para ese número creciente de SKU's.

De hecho, el número de SKU's se ha mantenido constante en los últimos años producto de las dificultades que plantea la no posibilidad de escalar este proceso.

3. Objetivos de la investigación.

El objetivo será generar y evaluar nuevos modelos con nuevas técnicas de predicción de demanda, que incorporen los datos históricos de unidades vendidas de una ventana de tiempo mayor.

Estas técnicas serán insertadas en el proceso de toma de decisión en forma previa a determinar las unidades por comprar, ya que el proceso de compra de mercadería comienza con la generación del pronóstico de demanda.

La solución planteada tendrá, en un primer momento, la forma de un script en el que los datos serán cargados y se obtendrán los pronósticos.

Finalmente, la tesis se posiciona como un trabajo novedoso en cuanto desarrollo de un pronóstico sobre bases cuantitativas para buscar otorgar nuevas herramientas a una empresa local, atendiendo la experiencia y conocimiento de los miembros de la organización. De esta manera se buscará diferenciarla de su competencia, y que, eventualmente, le sirva como plataforma para el avance en la adopción de más y mejores tecnologías de información en su trabajo cotidiano.

II. Marco Teórico.

1. Pronósticos.

Una buena manera de definir los pronósticos es “realizar un enunciado sobre el valor futuro de una variable de interés, fundamentado ya sea por el análisis de datos históricos disponibles, por el juicio de expertos en el tema o por una combinación de ambas cosas” (Montemayor Gallegos, 2014).

1. Horizonte de tiempo.

Se pueden comenzar diferenciando los pronósticos según su horizonte de tiempo, en corto, mediano y largo plazo.

1.1. Corto plazo.

Los de corto plazo son normalmente los más certeros, y buscan predecir las unidades a producir o vender, para saber cuanta mercadería y/o insumos adquirir, y tienen un plazo no mayor a 3 meses.

1.2. Mediano plazo.

Los de mediano plazo buscan predecir periodos más largos de tiempo, teniendo mayor agregación y margen de error, normalmente en forma de presupuestos, yendo de los 3 meses a los 3 años.

1.3. Largo plazo.

Aquellos de largo plazo tienen que ver con predicciones al abrir una fábrica o introducir nuevas tecnologías que cambian notoriamente la empresa, y cuyos efectos perduran más de un periodo.

2. Procedimiento.

Hanke (2010) menciona la diferencia existente en el procedimiento para llegar al pronóstico, en cualitativos, como aquellos en los que no se utilizan datos y el pronóstico se realiza en base al juicio del pronosticador; y aquellos cuantitativos, en los que si se utilizan datos pasados para la confección de los pronósticos, al aplicarseles métodos matemáticos que incorporan la información pasada.

2.1. Cualitativo.

Los primeros son los más extendidos, y se conocen como “Managerial Judgement” (juicio gerencial), llevado adelante por los gerentes o dueños de las empresas. También pueden incorporar elementos cuantitativos básicos, como datos históricos y complementarlo con su propia experiencia. Aunque útil, no es eficiente ni humanamente capaz llevar adelante esta técnica con una cantidad creciente de SKU’s. A pesar de esto, es el método más extendido según un artículo del International Journal of Forecasting (Fildes et al., 2019), que trae a consideración estudios realizados por Peterson (1993) y McCarthy et al.(2006).

2.2. Cuantitativos.

Los métodos cuantitativos son aquellos que se basan en modelos computacionales con base matemática y estadística. Según el tipo de datos que se posean hay una gran variedad de opciones y parámetros a utilizar, para que puedan recrear de la forma más exacta posible la realidad.

Dentro de los métodos cuantitativos, se puede diferenciar aquellos univariados y multivariados. Los univariados son aquellos en el que el valor de la variable a predecir depende

exclusivamente de sus valores pasados; mientras que en los multivariados la variable a predecir depende, además, de otras variables bajo control.

La idea en la confección de todo buen pronóstico es unir métodos cualitativos, en cuanto a experiencia y conocimiento de los gerentes o dueños acerca del mercado y los productos, con métodos cuantitativos que tomen esa información para otorgar mejores predicciones a la vez que escala la cantidad de demanda de muchos productos.

3. Nivel.

Los pronósticos apoyan decisiones a lo largo de la organización, en sentido estratégico, táctico y operacional (Fildes et al., 2019).

3.1. Estratégico.

Estratégicamente, al planificar decisiones en un horizonte de tiempo largo teniendo en cuenta cambios en el mercado, tecnologías, regulaciones, etc.

3.2. Táctico.

Tácticamente, se apoyan las decisiones del nivel estratégico, pero el cómo debe ser determinado. Aquí se pueden encontrar políticas de ventas, producción, compras, comercialización, entre otras.

3.3. Operativo.

A nivel operativo se encuentran las decisiones con el horizonte temporal más corto, como las decisiones relativas al nivel de compras y producción, cumpliendo requerimientos de los clientes a la vez que se reducen costos.

4. Agrupación.

Los pronósticos también pueden diferenciarse según la agrupación en la que se los trabaje, sea en forma agregada o por producto (Fildes et al., 2019).

4.1. Agregado.

A nivel agregado, no se tienen en cuenta diferencias entre productos o promociones, uniéndolas en base a un criterio temporal. Naturalmente, son medidas en forma de ganancia total en vez de unidades de productos. Se pueden separar por mercado, cadena o canal, o tipo o locación de tienda.

- Mercado: A nivel de mercado posee la utilidad de, en un periodo de tiempo mensual, cuatrimestral o anual, conocer la evolución de las condiciones de mercado y sus efectos en su negocio. La agregación puede ser efectuada en todo un país, región, industria o categoría. (Alon et al., 2001).

Los resultados que se obtienen normalmente poseen efectos de tendencias, variaciones estacionales, correlaciones y cambios de regímenes económicos, al poseer un tiempo prolongado que capture inflación, crecimiento económico y eventos inesperados.

- Cadena o Canal: A nivel de cadena o canal son útiles para el manejo de stocks de la compañía.

En la primera, para guiar sus decisiones de abastecimiento y manejo de inventarios;

En la segunda, relativas a apertura de nuevos canales de comercialización, como online; y su respectivas consecuencias en el apartado financiero de la empresa. Normalmente tienen una ventana temporal de un año.

- Tiendas: A nivel de tiendas, su importancia reside en el hecho de que los clientes de diferentes localizaciones geográficas pueden ser muy diferentes en preferencias y costumbres, y, aunque no lo fuesen, pueden verse afectadas por eventos particulares, clima, competidores, turismo, etcétera. Decisiones en este nivel son frecuentes al estimar insumos, recursos, promociones, precios, de las tiendas ya emplazadas, y las inversiones y proyecciones de ventas estimadas para colocar nuevas tiendas.

4.2. Productos.

A nivel de productos, cuando se busca predecir para un horizonte de tiempo corto, dado que se estiman varias series de tiempo individuales por cada producto o categoría, y se busca reducir el error lo más que se pueda, para evitar faltantes o sobre stocks que puedan generar incumplimientos con clientes o inmovilización innecesaria de capital. Por lo expresado anteriormente, tiene sentido considerar no solo el pronóstico que minimice alguna técnica de error sino también trabajar con cuantiles o intervalos de confianza, en orden de reducir las posibilidades de faltantes o sobrantes, según sea más costoso para la empresa.

Consideraciones importantes en este nivel tienen que ver con las características de los datos que se poseen y los factores que influyen la demanda, pudiendo mencionar la diferencia entre ventas y demanda, estacionalidad, eventos de calendario, clima, campañas de marketing y promociones, opiniones.

- Estacionalidades: Un producto puede tener múltiples estacionalidades, pudiendo ser mensuales, semanales, diarias e incluso horarias.
- Eventos: Eventos de calendario determinados generan mayores ventas de ciertos productos, como platos típicos en festividades como Navidad o Pascuas, o en fechas

patrias, y también picos de consumos de algunos productos por eventos deportivos, recitales, etcétera.

- Clima: El clima también afecta qué productos son consumidos. Bebidas o platos consumidos en verano son diferentes a los de invierno, no obstante, dentro de estas estaciones pueden haber días con climas opuestos que puedan cambiar las opciones elegidas.
- Promociones y marketing: Promociones y marketing también deben ser consideradas, en especial en pronósticos a corto plazo, en los que los efectos de descuentos pueden cambiar decisiones de compra establecidas, como la compra de un producto costoso por una reducción de precio que lo deja igual que uno de inferior calidad.
- Opiniones: Por último, las opiniones acerca de ciertos productos siempre existieron y guiaron las elecciones de los consumidores, pero eran difíciles de medir. Hoy, se pueden incorporar más fácilmente a través de opiniones en internet, o mediciones de *engagement* en redes sociales.

Al predecir con este nivel, se deben tener en cuenta tres dimensiones: el tiempo, producto y la posición en la cadena de suministro.

- Tiempo: las decisiones operativas tendrán un periodo de tiempo de pronóstico más corto, incrementándose en aquellas tácticas y estratégicas. Por ejemplo, la predicción de cuánta mercadería solicitar para la próxima semana será realizada todas las semanas, mientras que una estimación de ventas anuales para decidir una compra de un nuevo vehículo se realizará solo una vez.
- Producto: hay tres niveles: SKU, marca y categoría. De menor a mayor agregación, a nivel de SKU se predice para cada producto en particular, y tienen un horizonte de tiempo

de días o semanas. A nivel de marca, se juntan distintos productos de la misma, buscando capturar efectos de promociones o preferencias a través de las distintas marcas. A nivel de categoría, se reúnen según las clases determinadas productos de marcas diferentes, siendo posible desagregarse en subcategorías, a fin de apoyar decisiones como cuánto personal o espacio físico es necesario para cada una. Niveles de marcas o categorías tienen un horizonte de planeación de semanas o meses.

- Cadena de suministro: las empresas deben predecir según cada tienda, centro de distribución, fábrica, etcétera, que posean.

2. Patrones en series de tiempo.

Una parte importante previa a la selección de qué modelo utilizar en el pronóstico es el reconocimiento si los datos exhiben algún patrón.

Las mismas pueden ser clasificadas en horizontal o estacionaria, tendencia, estacionales y cíclicas.

- Horizontales o estacionarias: Aquellas en las cuales los datos recolectados se mantienen alrededor de su media.
- Tendencia: Aquellas en las que los datos tienen periodos marcados de crecimiento o decrecimiento a lo largo de varios periodos.
- Estacionales: Aquellas en las que la tendencia de los datos se ve modificada en cierto momento durante el día, semana, mes o año, repitiéndose continuamente en cada periodo.
- Cíclicos: Aquellas en las cuales los datos fluctúan alrededor de una tendencia. Sus patrones son inestables, y rara vez se repiten en intervalos de tiempo fijos. Normalmente, pueden describir ciclos económicos.

1. Modelos predictivos cuantitativos

Hay una gran cantidad de modelos predictivos utilizados para la generación de pronósticos dentro de aquellos que son cuantitativos. Una diferenciación que es muy utilizada es la previamente señalada entre univariados y multivariados.

1. Univariados

Aquellos que predicen el valor de la variable en función de sus valores pasados. Por ello, son denominados también modelos de análisis de series de tiempo. Estudios realizados (Ali et al., 2009) encontraron que estas técnicas producen buenos resultados en periodos sin promociones particulares, preferentemente debiendo ser adoptados para modelos agregados o para productos con bajas promociones o elasticidad precio de la demanda (Fildes et al., 2019).

Entre los posibles tratamientos para series de tiempo univariadas, se encuentran los Métodos de Suavización y los Métodos de Descomposición de Series de Tiempo (Montemayor Gallegos, 2014).

1.1. Métodos de Suavización

Toman la siguiente forma matemática:

$$y_{(t+1)} = f(y_{(t)}, y_{(t-1)}, y_{(t-2)}, \dots, y_{(t-k)}) \quad (1)$$

Utilizando como base el patrón histórico de los datos de la serie para realizar los pronósticos. Los modelos univariados de Suavización más utilizados son: Promedios Medios Móviles, Promedios Móviles Ponderados, Suavización Exponencial.

- Promedios Medios Móviles.

Este método va reemplazando la última observación cada vez que ingresa una nueva y obtiene un nuevo promedio. Todas las observaciones reciben el mismo peso. El resultado será el pronóstico.

$$SMA_t = \frac{y_t + y_{t-1} + \dots + y_{t-k}}{k} \quad (2)$$

- Promedios Medios Ponderados: Este método es igual al Promedio Medio Móvil a excepción de que los pesos de las observaciones no son todos iguales, ya que para cierto conjunto de datos se les otorga mayor influencia. Según el juicio del pronosticador, decidirá si le otorga más a las observaciones antiguas o nuevas, acorde a lo que crea que representa mejor el futuro.

$$\bar{x} = \frac{\sum_{i=1}^n \omega_i x_i}{\sum_{i=1}^n \omega_i} \quad (3)$$

Ambos son utilizados en datos que exhiben patrones estacionarios o de tendencia.

- Suavización Exponencial: Es un caso especial del Promedio Móvil Ponderado, ya que las observaciones tienen distinto peso. Pero, en este caso, solo se establece un peso, que es para la más reciente. A medida de que van aumentando los datos, los más longevos van perdiendo cada vez más influencia.

$$s_t = \alpha x_t + (1 - \alpha) s_{t-1}, \quad t > 0 \quad (4)$$

Son utilizados en datos con patrones de tendencia y estacionales.

1.2. Modelos de Descomposición de Series de Tiempo.

Toman la siguiente forma matemática:

$$y_t = f(E_t, T_t, C_t, I_t) \quad (5)$$

Parten de la idea de que cualquier variable proveniente de una serie de tiempo puede ser descompuesta en los componentes de estacionalidad, tendencia, ciclicidad y aleatoriedad, estimando por separado cada uno de estos patrones, y agregandolos para efectuar el pronóstico de forma Aditiva o Multiplicativa.

- Modelo Aditivo: es aquel que trata los valores de la serie como una suma de los componentes, funcionando mejor cuando la variabilidad de la misma es aproximadamente la misma, cercana al centro.

Utilizando este modelo, todos los componentes deben estar expresados en las mismas unidades de la serie.

El componente estacional tiene un efecto que no parte del valor de la tendencia, por lo que posee oscilaciones de amplitud fija.

Toma la siguiente forma:

$$y_t = E_t + T_t + C_t + I_t \quad (6)$$

- Modelo Multiplicativo: es aquel que trata los valores de la serie como el producto de los componentes, funcionando mejor cuando existe variabilidad en la misma.

En este modelo, solo un componente puede tener las mismas unidades de la serie, eligiendo frecuentemente la de tendencia.

El componente estacional tendrá un efecto proporcional al de la tendencia en el mismo sentido.

Toma la siguiente forma:

$$y_t = E_t * T_t * C_t * I_t \quad (7)$$

Aunque no son los únicos esquemas posibles, son los más utilizados, y la mayoría de los modelos utilizan alguno. Tampoco hay una regla general para elegir la utilización de uno u otro, si bien la variabilidad en la tendencia es utilizada, es fácil igualar cuando no lo son mediante la transformación logarítmica de los componentes.

Por ello, una segunda opción es mediante el análisis de magnitudes de efectos estacionales.

Por ejemplo, comparar la estacionalidad mensual de distintos años de una serie. La comparación puede ser realizada mediante el cociente o la diferencia.

Si se usa un esquema multiplicativo, los cocientes serán parecidos, mientras que las diferencias no.

Si se usa un esquema aditivo, los cocientes serán disímiles, mientras que las diferencias no.

De esta manera, se pueden comparar ambas formas para analizar cuál es la más homogénea, siendo el Coeficiente de Variación de Pearson una buena forma de realizar esta tarea.

Un método muy utilizado dentro de esta categoría es ARIMA (Autoregressive Integrated Moving Average, por sus siglas en inglés) con Metodología Box-Jenkins.

El componente AR (p) está determinado por los valores actuales de la serie de datos; el componente MA (q) proviene de los valores actuales de error aleatorio correlacionados con los valores del pasado. Finalmente, el modelo supone que los datos no son estacionales; y son

estacionarios en media y varianza, pero si no lo fuesen, ingresa el término I (d) para subsanar esa falta a través de diferenciación.

2. Modelos Multivariados.

Son aquellos que proponen que la variable bajo estudio (dependiente) puede ser predecida a través de los valores de otras variables (independientes o explicativas), proporcionando la ventaja de poder identificar qué factores influyen en la variable estudiada.

Dentro de las técnicas que pueden ser utilizadas en los Modelos Multivariados, se las puede diferenciar según sean Métodos Lineales o No Lineales.

2.1. Métodos Lineales.

Los más utilizados son los Modelos de Regresión Lineales, los cuales requieren que los parámetros sean lineales, por ende, cada término del modelo debe ser aditivo y sólo uno puede multiplicar el término. Se limita la ecuación a una sola forma, la cual puede variar según sea una Regresión Simple, en la que solo se busca determinar el efecto de una variable independiente, o Múltiple, que admite más de una variable independiente.

Éstos modelos tienen la ventaja de ser simples, fáciles de comprender y rápidos. Esta última consideración los hace especialmente atractivos con un número creciente de SKU por pronosticar

$$\text{Simple: } Y = \beta_0 + \beta_i X_i + u_i \quad (8)$$

$$\text{Múltiple: } Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + u_k \quad (9)$$

En estos modelos, también se puede considerar que la variable endógena se presente con rezagos como variable explicativa, de manera de generalizar aún más la especificación. Esto permite capturar posibles efectos rezagados y considerar la dependencia temporal en los datos. Esto puede mejorar la capacidad del modelo para capturar relaciones dinámicas y patrones a lo largo del tiempo, lo que puede resultar en pronósticos más precisos y robustos.

$$\text{Simple: } Y = \beta_0 + \beta_i X_i + \gamma_n Y(t - n) + u_i \quad (10)$$

Múltiple:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \gamma_1 Y(t - 1) + \gamma_2 Y(t - 2) + \dots + \gamma_n Y(t - n) + u_k \quad (11)$$

2.2. Modelos No Lineales y de Aprendizaje Automático

Dentro de estas opciones se pueden encontrar Modelos de Regresión No Lineales y Algoritmos de Aprendizaje Automático, como Redes Neuronales, Máquinas de Vector Soporte, Árboles de Decisión, entre otros.

La diferencia con respecto a los Modelos Lineales es que se pueden aproximar con funciones de manera no lineal para extraer información de los datos, lo que puede derivar en mejores pronósticos. El caso de la industria de la moda, donde la demanda es muy volátil y es difícil determinar patrones (Choi et al., 2014), es un buen ejemplo de caso de éxito de estos enfoques.

3. Consideraciones sobre el Overfitting y el Costo Computacional

En el desarrollo de modelos de predicción, dos aspectos cruciales a tener en cuenta son el riesgo de sobreajuste (overfitting) y el costo computacional asociado. En este apartado, se

abordarán estas consideraciones en relación con los diferentes modelos de pronóstico previamente discutidos, destacando la importancia de aplicar técnicas de validación y optimización para garantizar la calidad de las predicciones y la eficiencia del proceso computacional.

3.1. Sobreajuste (Overfitting):

El sobreajuste ocurre cuando un modelo se ajusta demasiado a los datos de entrenamiento y pierde su capacidad de generalización a nuevos datos. Esto sucede cuando el modelo captura tanto los patrones genuinos como el ruido aleatorio presente en los datos de entrenamiento. Como resultado, el modelo puede proporcionar predicciones inexactas cuando se aplica a nuevos conjuntos de datos. Para evitar el sobreajuste, es importante encontrar un equilibrio entre la complejidad del modelo y su capacidad para capturar patrones genuinos sin ajustarse en exceso a las fluctuaciones aleatorias.

Entrenamiento y Testeo (Train & Test): El enfoque entrenamiento y testeo implica dividir el conjunto de datos en dos partes: uno para el entrenamiento del modelo y otro para evaluar su rendimiento. El modelo se entrena utilizando los datos de entrenamiento y luego se evalúa utilizando los datos de prueba, que no se utilizaron durante el entrenamiento. Esto permite medir cómo se desempeña el modelo en datos no vistos previamente.

Validación cruzada (Cross-validation): La validación cruzada es una técnica que busca superar las limitaciones del enfoque de entrenamiento y testeo al realizar múltiples divisiones de los datos en conjuntos de entrenamiento y prueba. Por ejemplo, en la validación cruzada de k iteraciones (k -fold cross-validation), los datos se dividen en k subconjuntos o "folds". Luego, el modelo se entrena y evalúa k veces, cada vez utilizando un subconjunto diferente como datos de

prueba y el resto como datos de entrenamiento. Los resultados se promedian para obtener una estimación más precisa del rendimiento del modelo.

Validación fuera de muestra (Out-of-sample): La validación fuera de muestra implica retener un conjunto adicional de datos que no se utiliza durante ninguna etapa del proceso de entrenamiento o calibración del modelo.

3.2. Optimización y Calibración:

Además de las técnicas de validación, es importante realizar una optimización y calibración adecuadas de los modelos de predicción. Esto implica ajustar los parámetros del modelo de manera que se obtenga el mejor rendimiento posible en términos de precisión y generalización. Para ello, se pueden aplicar métodos como la búsqueda sistemática de hiperparámetros, utilizando técnicas de optimización, y evaluar el rendimiento del modelo en conjuntos de validación. Estos enfoques permiten evitar el sobreajuste al encontrar la configuración óptima del modelo y, a su vez, mejorar la calidad de las predicciones.

3.3. Costo Computacional:

El costo computacional es otro aspecto relevante en el desarrollo de modelos de pronóstico, especialmente cuando se trata de un número creciente de SKU o se utilizan algoritmos de aprendizaje automático más complejos. La aplicación de estos modelos puede requerir un tiempo de cómputo considerable y recursos computacionales significativos. Por lo tanto, es esencial evaluar cuidadosamente los recursos disponibles y considerar enfoques computacionalmente eficientes, como así también la búsqueda de optimización de los procesos de cálculo y la implementación de estrategias de paralelización cuando sea posible. La consideración del costo

computacional permite garantizar una implementación práctica y escalable de los modelos de predicción a implementar.

4. Significatividad en los resultados.

Cuando se comparan distintos métodos, es útil evaluar si las diferencias observadas entre ambos análisis son estadísticamente significativas, en pos de buscar validar los resultados obtenidos.

La significancia estadística es “una medida de la probabilidad de que la hipótesis nula sea cierta, en comparación con el nivel aceptable de incertidumbre con respecto a la respuesta verdadera” (Tenny & Abdelgawad, 2019).

Dicho de otro modo, brinda una medida cuantitativa de la probabilidad de que las diferencias observadas sean resultado de la variabilidad aleatoria en lugar de ser representativas de una diferencia real entre las variables analizadas.

Para ello, se determina una “hipótesis a ser desaprobada, llamada hipótesis nula, que típicamente es la declaración inversa de la hipótesis”(Tenny & Abdelgawad, 2019).

Como el investigador nunca estará totalmente seguro de su hipótesis, se determina un nivel de significancia, que representa la probabilidad de que esté equivocado.

Normalmente, el nivel de significancia elegido es de un 5%, que representa que el investigador espera que su hipótesis sea correcta el 95% de las veces, y acepta una posibilidad de error del 5%.

Para conocer el resultado, se debe llevar adelante un test de significatividad que arroja un valor llamado p-valor o p-value, el cual “muestra la probabilidad que, bajo un modelo estadístico

específico, un resumen estadístico de los datos (p. ej., la diferencia media de la muestra entre dos grupos comparados) sea igual o más extremo que su valor observado” (Tenny & Abdelgawad, 2019). Así, si el p-valor es menor al nivel de significatividad elegido se puede rechazar la hipótesis nula.

Para determinar qué test realizar se debe prestar atención, en general, a dos condiciones de los datos.

4.1. Independiente o pareada.

Pueden ser independientes, en la que los individuos o datos en las muestras sean distintos; o pareadas, en las que los individuos o datos en las muestras sean los mismos, pero las condiciones a las que han sido expuestos son diferentes.

4.2. Paramétrica o no paramétrica.

Paramétricas, en las que se asume un determinado tipo de distribución de los datos, como puede ser la distribución normal; o no paramétricas, en las que no se asume un determinado tipo de distribución en los datos.

5. Estado del arte.

5.1. Evolución de utilización de técnicas de pronósticos de ventas (1984-2006).

Para contrastar la información teórica señalada anteriormente, en lo referido al uso de los diversos modelos de pronósticos en el ámbito empresarial, se trae a colación el material del artículo *The Evolution of Sales Forecasting Management: A 20-Year Longitudinal Study of Forecasting Practices* (McCarthy et al., 2006).

La información acerca del uso de los diversos métodos que nuclea el artículo proviene de encuestas realizadas en tres momentos diferentes. Primero, el artículo “Familiarity, application, and performance of sales forecasting techniques” (Mentzer & Cox, 1984) y, luego, “Forecasting technique familiarity, satisfaction, usage, and application” (Mentzer & Kahn, 1995), para finalmente, a los 20 años de la primera publicación, los autores llevar a cabo su propia encuesta.

La misma fue obtenida mediante una muestra aleatoria de ejecutivos involucrados en el proceso de pronóstico, a través un cuestionario enviado por email. Se obtuvieron 86 respuestas, de las cuales el 5% correspondió a CEOs, 12% a vicepresidentes, 28% a directores, 50% a administradores, y 5% a analistas. La industria con más representación fue la de bienes de consumo inmediato.

En lo referido a Métodos Cuantitativos utilizados en nivel de SKU, se obtuvieron las siguientes conclusiones:

En la primera encuesta, observaron que los más utilizados eran Análisis de Línea de Tendencia, Medias Móviles, y Regresión.

En la segunda encuesta, observaron que los más utilizados eran Suavización Exponencial, Regresión, y Medias Móviles empatadas con Análisis de Línea de Tendencia.

Finalmente, en el artículo que las resume, observaron que los más utilizados eran Suavización Exponencial, Regresión empatada con Medias Móviles, y Análisis de Línea de Tendencia.

En cuanto a la familiaridad de los encuestados con los distintos métodos de pronósticos, observadas a lo largo de las encuestas, se advierte que Medias Móviles es la más familiar,

seguida por Proyección Lineal, Suavización Exponencial, Regresión y Análisis de Línea de Tendencia.

En cuanto a la satisfacción, el Suavizado Exponencial sigue siendo la mejor, seguida del Análisis de Línea de Tendencia. La Regresión, anteriormente la más satisfactoria, ahora tiene una calificación más baja, mientras que la fuerza de ventas vio un aumento en los niveles de satisfacción.

En general, los niveles de satisfacción con las técnicas de pronóstico parecen estar disminuyendo, pero los niveles de insatisfacción también han disminuido, lo que indica un cambio hacia una comprensión más madura de la gestión de pronósticos de ventas como un proceso en lugar de un enfoque basado en técnicas.

5.2. Nuevos métodos

Para conocer el estado más actual de las técnicas de pronóstico utilizadas a la fecha, se realizó una investigación de las principales publicaciones en el ámbito de la predicción en *retail*.

En ese sentido, se arribó a la información que proporcionó una de las competiciones mundiales más reconocidas en este ámbito, como es la M5 Competition. Esta es la 5ta edición de dicha contienda, cuyo propósito es arribar a los mejores resultados en cuanto a pronósticos, en sentido teórico y práctico.

En particular, la M5 Competition se “enfocó en el pronóstico de ventas en *retail* con el objetivo de producir el pronóstico más acertado para 42.480 series de tiempo, que representan las ventas de la empresa de *retail* más grande del mundo, Walmart, así como para proporcionar las estimaciones más precisas sobre la incerteza de estos pronósticos” (Makridakis et al., 2022).

Por mucho años, “fue empíricamente comprobado que los métodos simples han sido tan acertados como aquellos complejos o estadísticamente sofisticados” (Makridakis et al., 2020).

Esta última edición fue especial ya que fue la primera en la que los 5 primeros lugares fueron puramente algoritmos de aprendizaje supervisado, performando significativamente mejor que los métodos estadísticos previos, así como aquellas combinaciones de estos dos (Makridakis et al., 2022).

En concreto, los 5 primeros lugares fueron conformados por variantes y combinaciones de técnicas como:

- LightGBM: Utiliza árboles de decisión y el método de gradient boosting como base .
- Neural Networks: Consisten en una red interconectada de unidades (neuronas), que procesan y transmiten información emulando el comportamiento del cerebro humano.
- N-BEATS: Consiste en modelos de redes neuronales que utilizan una arquitectura especializada para capturar patrones y tendencias en los datos de series de tiempo.

5.3. Consideraciones sobre el estado actual de los metodos de pronostico

Los avances computacionales han hecho más populares y accesibles las técnicas de pronósticos clásicas, como los métodos estadísticos. Estos han continuado performando de gran manera, por sí solos o combinados con otros modelos de aprendizaje supervisado.

Así también han generado nuevas posibilidades, como el poder generar predicciones con modelos que soportan grandes volúmenes de datos y que pueden incorporar un gran número de variables externas.

Finalmente, más que una única receta, el método elegido dependerá de las características de la serie de tiempo a predecir, como así también del poder de cómputo y conocimientos del pronosticador para elegir la mejor alternativa posible.

III. Metodología

1. Datos

Los datos provienen del software de facturación de la empresa, el cual permite acceder a diferentes tipos de reportes según la información que se quiera obtener.

En este caso, serán utilizados los reportes de unidades vendidas por producto.

La ventana de tiempo de la que se cuentan con datos irán desde el 06/2019 (Junio 2019) hasta 01/2023 (Enero 2023), en frecuencia diaria ($T = 1.309$ observaciones temporales)

En cuanto a las dimensiones del dataset, son de 150 artículos.

La variable relevada es la cantidad de ventas pronosticada, en función de las ventas pasadas.

1.1. Forma y transformación del dataset.

El reporte del software de facturación de la empresa puede mostrar, como máximo, una ventana de tiempo de 31 días. De manera que, por cada mes, se obtiene un reporte.

El sistema puede emitirlo en diversos formatos, como pdf, txt u hoja de cálculo. Se decidió generarlos en este último, y posteriormente transformarlo a valores separados por comas (csv, por sus siglas en inglés).

Las columnas que posee el reporte son, además del nombre de los artículos y las unidades vendidas por cada día del mes, las siguientes:

- Código: Código interno de los productos de la empresa.
- Fecha Inicio: Fecha de inicio del reporte.

- Fecha Fin: Fecha de finalización del reporte.
- Ventas: En unidades monetarias, obtenidos por la multiplicación del precio unitario por la cantidad vendida.
- Costo de Ventas : En unidades monetarias, obtenidos por la multiplicación del costo unitario por la cantidad vendida.
- Costo de Venta Unitario: En unidades monetarias, el costo unitario de cada producto.
- Existencias: Cantidad de cada producto al final del periodo reportado.

Con el dataset en esta forma se realizaron los gráficos que se podrán ver en la próxima sección, mediante R Studio, entorno de programación en lenguaje R.

Este lenguaje fue utilizado para esa tarea ya que es conocido por su amplia gama de bibliotecas y paquetes especializados en visualización de datos, destacándose entre ellas *ggplot2*. De manera sencilla, genera gráficos de excelente calidad con una gran flexibilidad para su manipulación y personalización, lo que permite explorar y comunicar eficazmente los patrones y las características de los datos.

Luego, a través de Spyder, entorno de programación en lenguaje Python, se eliminaron las columnas que no iban a ser utilizadas nuevamente, siendo estas el Código, Fecha Inicio, Fecha Fin, Ventas (en unidades monetarias totales), Costo de Ventas (en unidades monetarias totales), Costo de Venta Unitario, Existencias.

Este lenguaje fue utilizado durante la mayor parte del trabajo, por su facilidad de uso y su capacidad para manipular grandes conjuntos de datos, a la vez de ser ampliamente utilizado en el ámbito de la ciencia de datos, gracias a su rica colección de bibliotecas especializadas que

implementan de manera eficiente y sencilla la mayoría de los algoritmos de predicción comúnmente utilizados.

Posteriormente, se reordeno el data frame para que muestre, por cada artículo, una columna con la fecha y otra con las cantidades vendidas.

Con respecto a la presencia de datos faltantes en algunos artículos, se determinó que los mismos tenían origen en la no disponibilidad para la venta de los mismos en ese momento.

Esta situación no era de magnitud considerable, y el tratamiento elegido fue reemplazarlos con 0, lo cual es sensato ya que representa la no venta de ninguna unidad de ese producto.

Finalmente, se predijeron los artículos activos en el set de testeo, ya que de predecir todos los productos, incluidos los que no tenían valores en este, provocaría la imposibilidad de ser sujetos a evaluación. Aquí tampoco fueron de magnitud considerable la cantidad de artículos que no estaban activos, siendo específicamente solo 11 casos.

1.2. Análisis descriptivo.

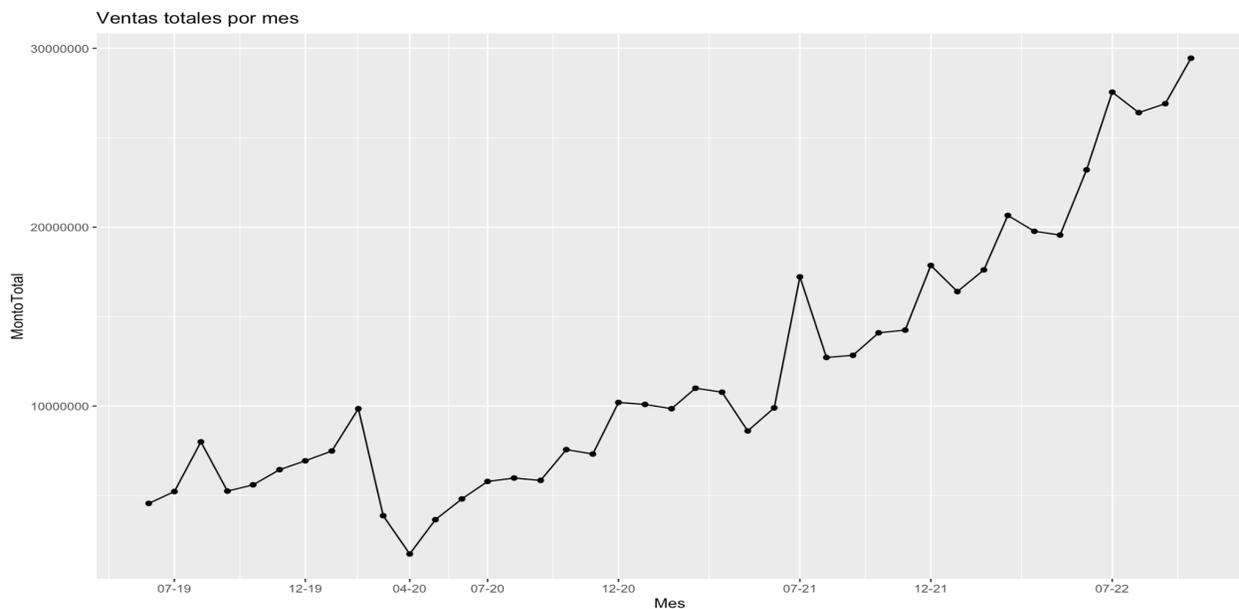
Para efectuar este trabajo son bastante útiles las representaciones gráficas de los datos, ya que nos permiten conocer patrones, distribuciones, características y tendencias de los mismos.

Empezamos observando un gráfico sencillo que muestra las ventas totales de todos los artículos, en términos monetarios, por cada mes, expresados en pesos argentinos corrientes. Los cortes en el Eje X fueron realizados cada 7 meses, coincidiendo con los meses de Julio y Diciembre, ya que son los meses que a priori poseen una mayor demanda por cuestiones estacionales. Una vez realizado el gráfico, vemos que esta idea inicial tiene sentido ya que se

observan picos de consumo en los mismos. La tendencia creciente que se ve reflejada puede suponer una buena marcha del negocio, no obstante, el contexto inflacionario del país también colabora en ese aspecto.

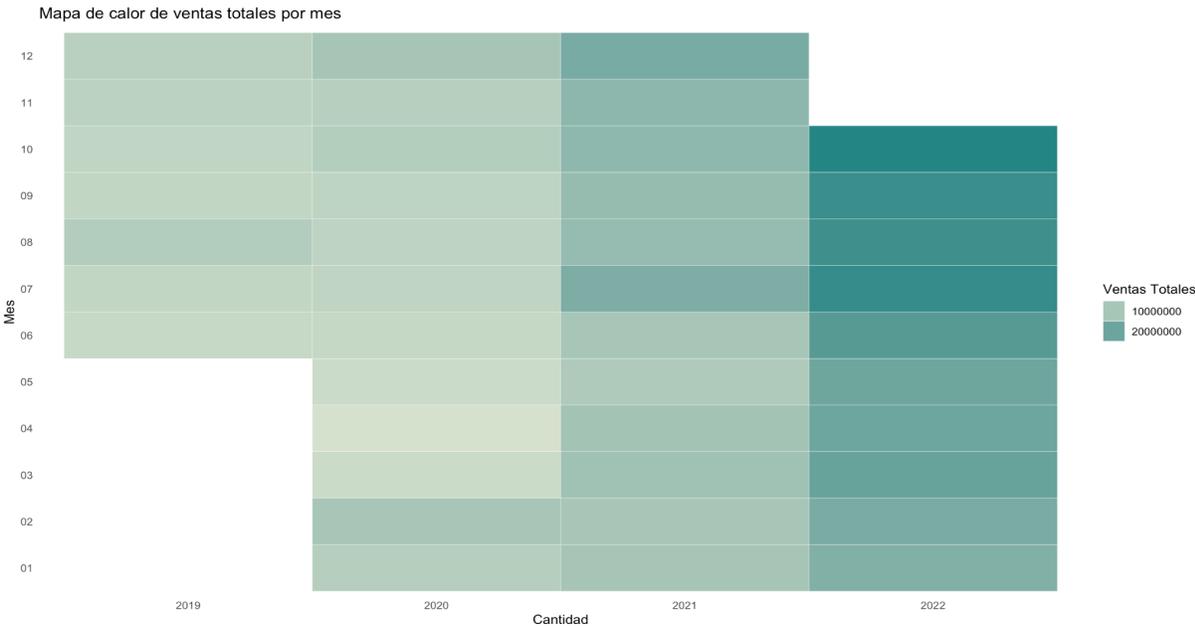
Por último, se marcó en especial el mes de Abril de 2020 para comprender la brusca caída en las ventas producto de la restricción gubernamental por la pandemia de COVID-19.

Figura 1. Ventas totales por mes medidas en unidades monetarias.



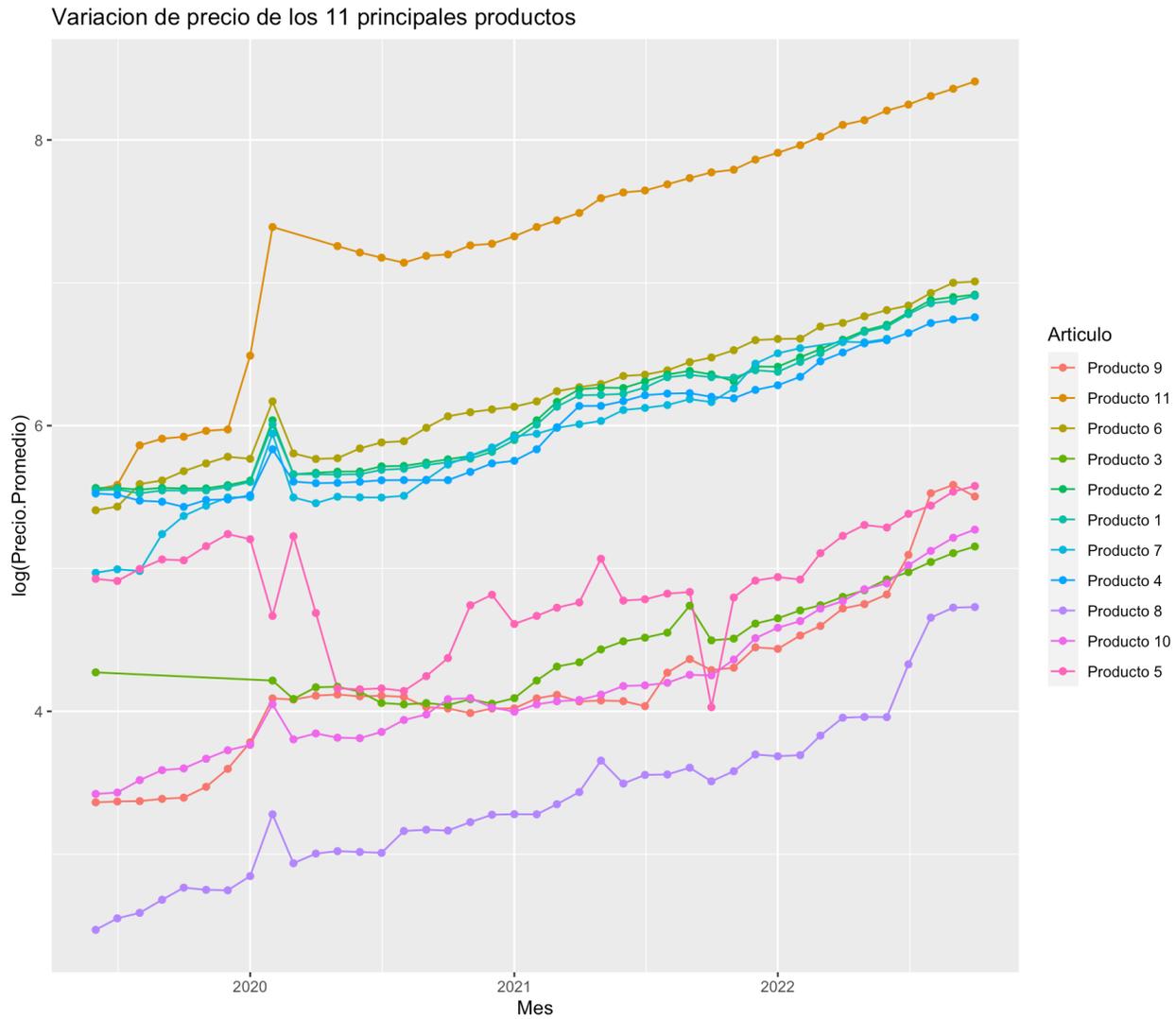
Otra forma de observar lo anterior es a través de un mapa de calor. En el mismo se puede apreciar, los meses de cada año, en donde a medida que los colores se acentúan, el valor monetario de las ventas crecen.

Figura 2. Mapa de calor de ventas totales por mes medidas en unidades monetarias.



Para finalizar el análisis de valores monetarios, observamos la variación en precios de los 11 productos más vendidos. Antes de graficar se convirtió la variable de Precio a logaritmo para facilitar la observación.

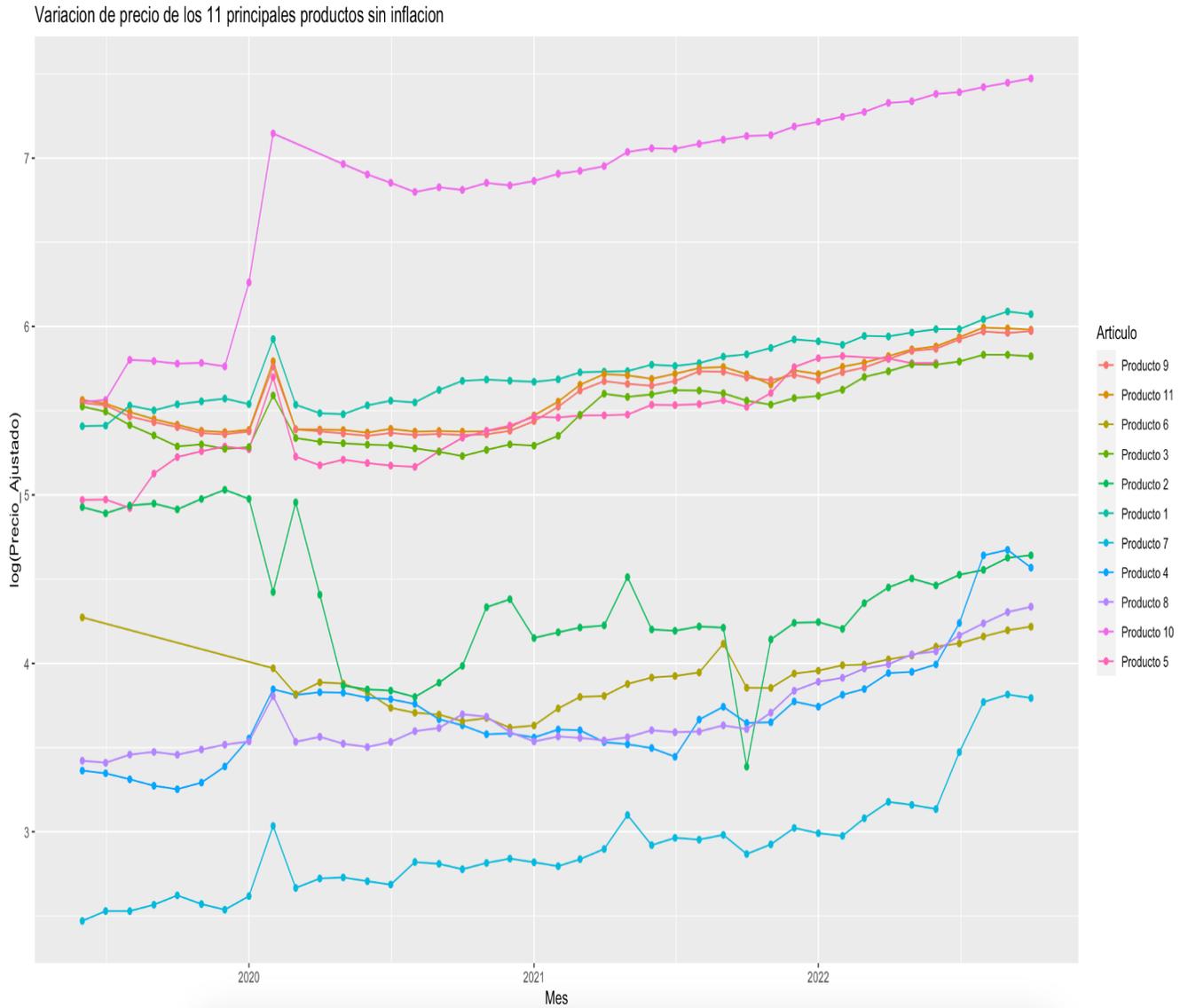
Figura 3. Variación de precio de los 11 principales productos.



A nivel de ventas de unidades de productos, este gráfico agregado similar a la Figura 1 muestra una tendencia alcista, de menor magnitud, pero señalando que, más allá del hecho que el componente inflacionario de los datos que incrementa los valores, también aumenta por la mayor cantidad de unidades vendidas.

No obstante, para profundizar el análisis, se deflactó la serie de precios tomando como base Junio 2019, utilizando el Índice de Precios al Consumidor Nivel General publicado por el Instituto Nacional de Estadísticas y Censos de la República Argentina (INDEC)¹.

Figura 4. Variación de precio de los 11 principales productos deflactados.



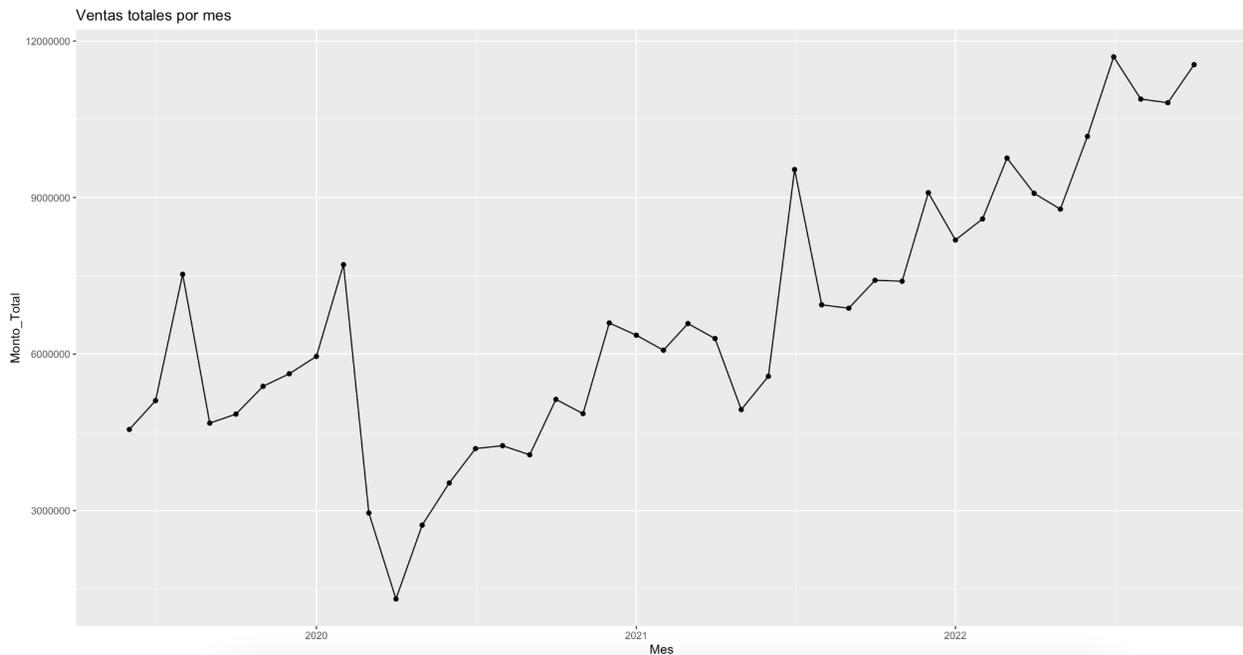
Quitando la inflación, es interesante observar que la variación de precio de los 11 productos más vendidos no es igual que en forma nominal. Por ejemplo, ahora es el Producto 10

¹ <https://www.indec.gob.ar/indec/web/Nivel4-Tema-3-5-31>

el que más variación tuvo, reemplazando al Producto 11. Esta información revela cambios relevantes en la clasificación de los productos en términos de su evolución de precios, lo cual proporciona insights valiosos, como cambios en el mercado, rentabilidad, planificación de precios, estrategias, etc., para la toma de decisiones y la comprensión del comportamiento del mercado.

En el próximo gráfico, se puede observar que la tendencia alcista sin el efecto inflacionario no es tan marcada, pero se observa un crecimiento, que será corroborado nuevamente en el gráfico que muestra solo la evolución de unidades vendidas.

Figura 5. Ventas totales por mes medidas en unidades monetarias deflactadas.



Por último observamos un nuevo mapa de calor de unidades monetarias vendidas, pero sin inflación. La tendencia alcista se mantiene más allá de haber removido el componente inflacionario.

Figura 6. Mapa de calor de ventas totales por mes medidas en unidades monetarias deflactados.

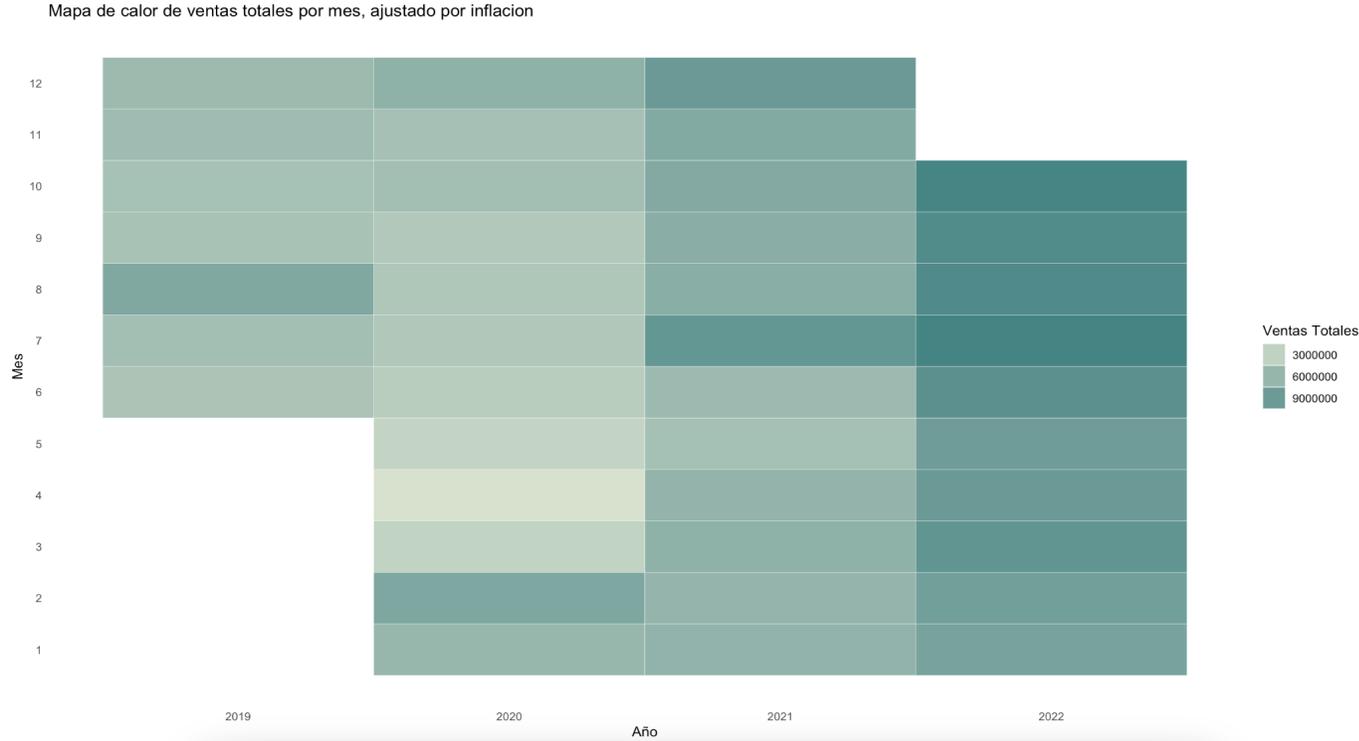
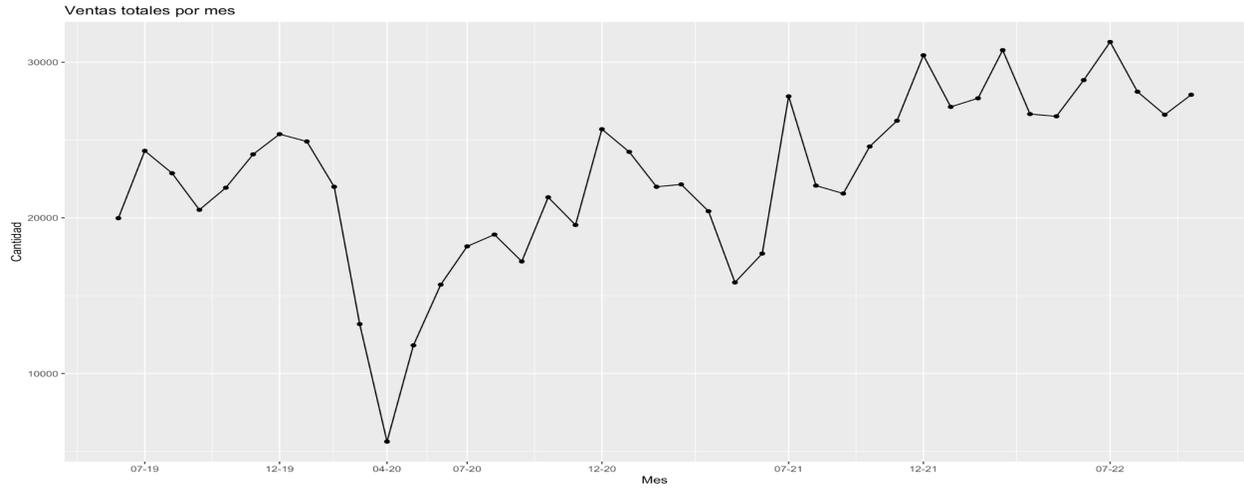
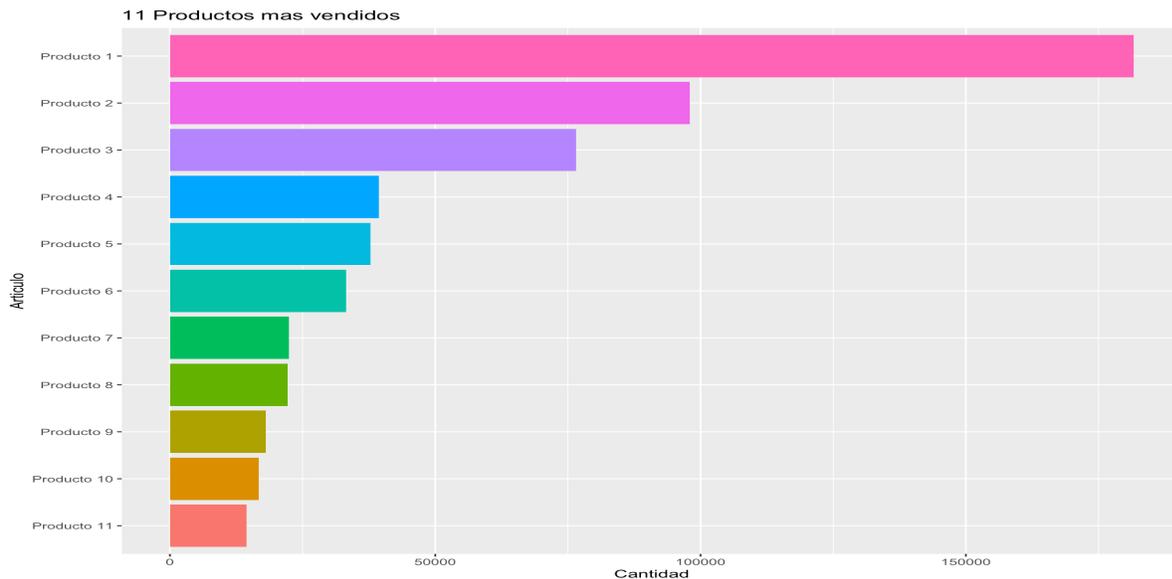


Figura 7. Ventas totales por mes medidas en unidades físicas.



Por el lado de los productos más vendidos en el periodo de tiempo trabajado, se puede observar el siguiente gráfico:

Figura 8. Productos más vendidos.

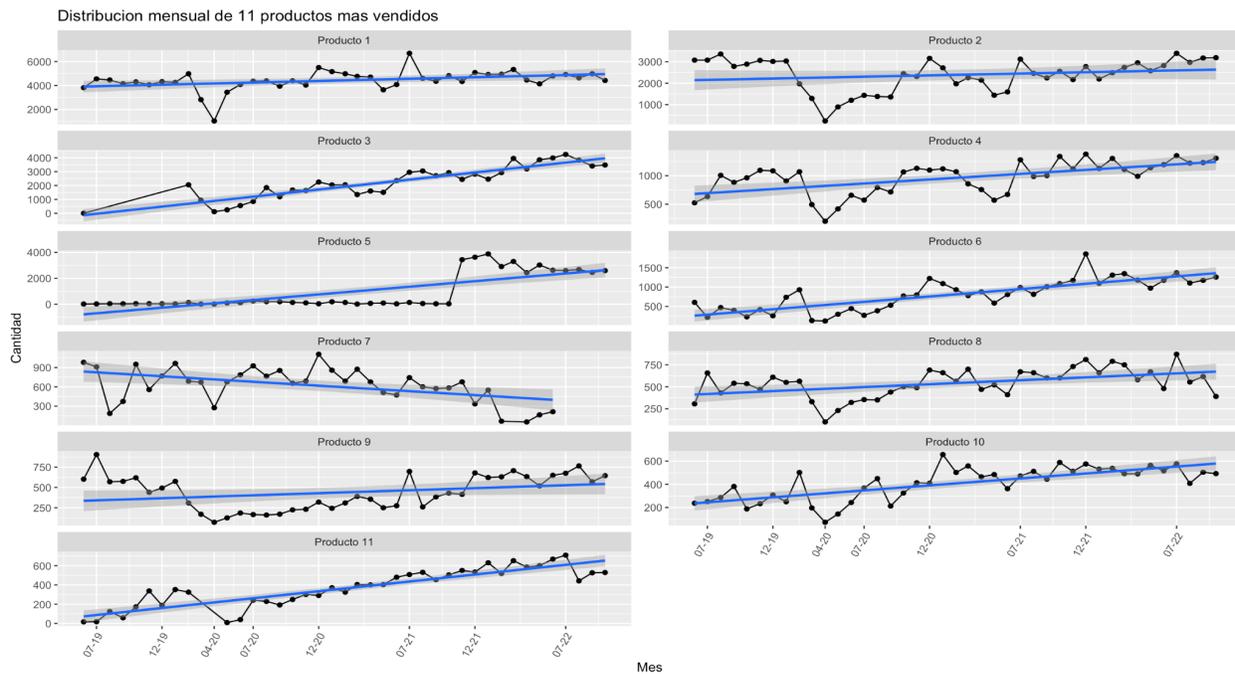


En base a lo anterior, seguimos hurgando para conocer la distribución temporal de las ventas de estos productos, y sus tendencias individuales. Para facilitar esa tarea, añadimos líneas de tendencias por cada uno.

Sin considerar el mes de pandemia, podemos observar que los dos primeros productos mantienen una cantidad constante, con un leve aumento.

Los demás productos, a excepción del número 7, han mostrado una tendencia alcista. A diferencia de los dos primeros, son menos constantes y tienen más variaciones mensuales.

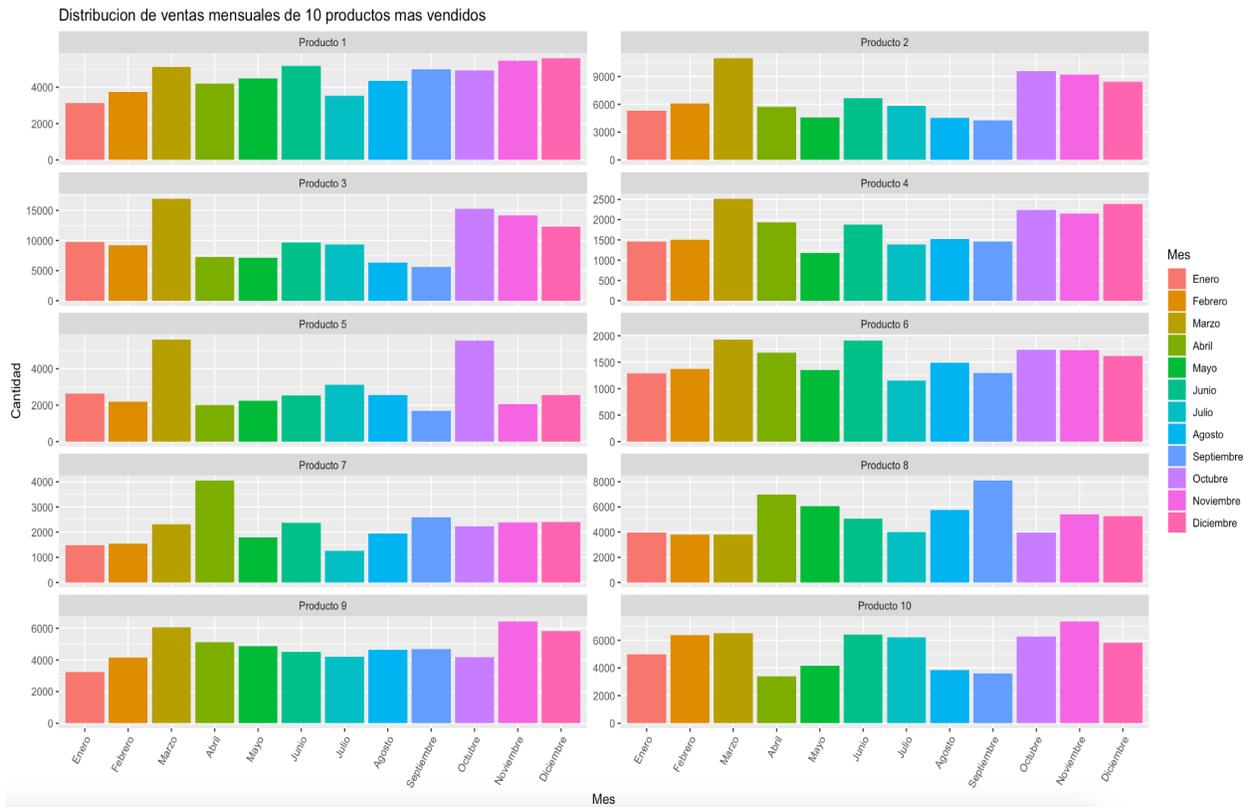
Figura 9. Distribución mensual de los 11 productos más vendidos.



La distribución anterior también puede observarse al agrupar las ventas totales de los productos más vendidos por cada mes.

Los productos 1, 4, 6 y 9 mantienen un nivel constante de ventas a lo largo del año. Los demás, a excepción del 8, observan mesetas en los meses de otoño e invierno.

Figura 10. Distribución de unidades vendidas por mes.



Datos diarios.

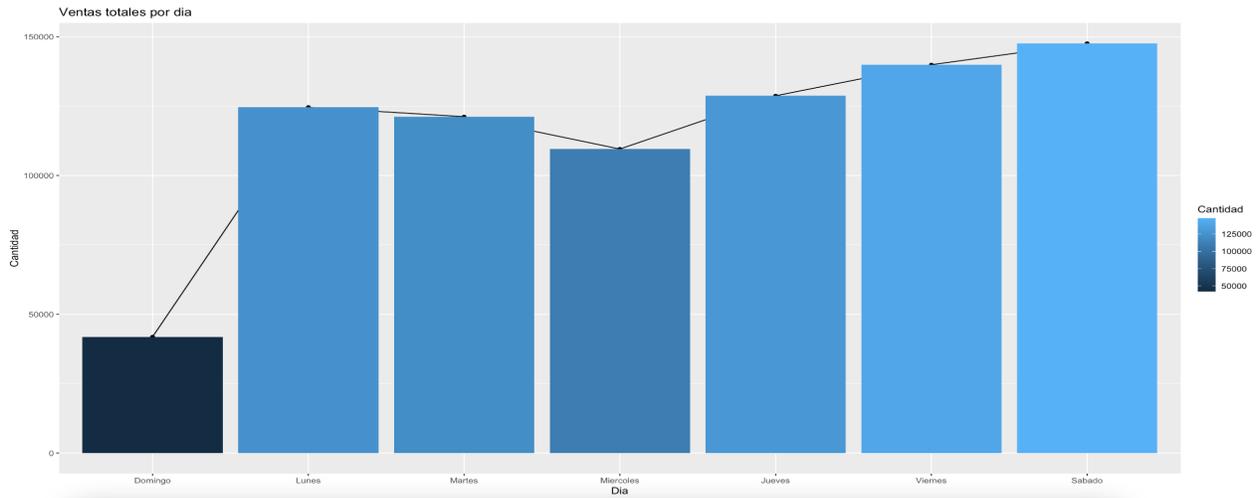
Al utilizar los datos en forma diaria, se puede seguir obteniendo información relevante.

Para empezar, se analiza la distribución de unidades vendidas totales según los distintos días de la semana. La información obtenida se puede separar en dos momentos, posterior y previo al fin de semana.

Los días lunes y martes los clientes reponen parte de la materia prima vendida el fin de semana. El miércoles es el día de menor actividad, para luego comenzar una escalada a partir del jueves que alcanza su punto máximo el día sábado. Esto tiene sentido ya que la mayor actividad comercial de los clientes es obtenida, en general, en ese periodo.

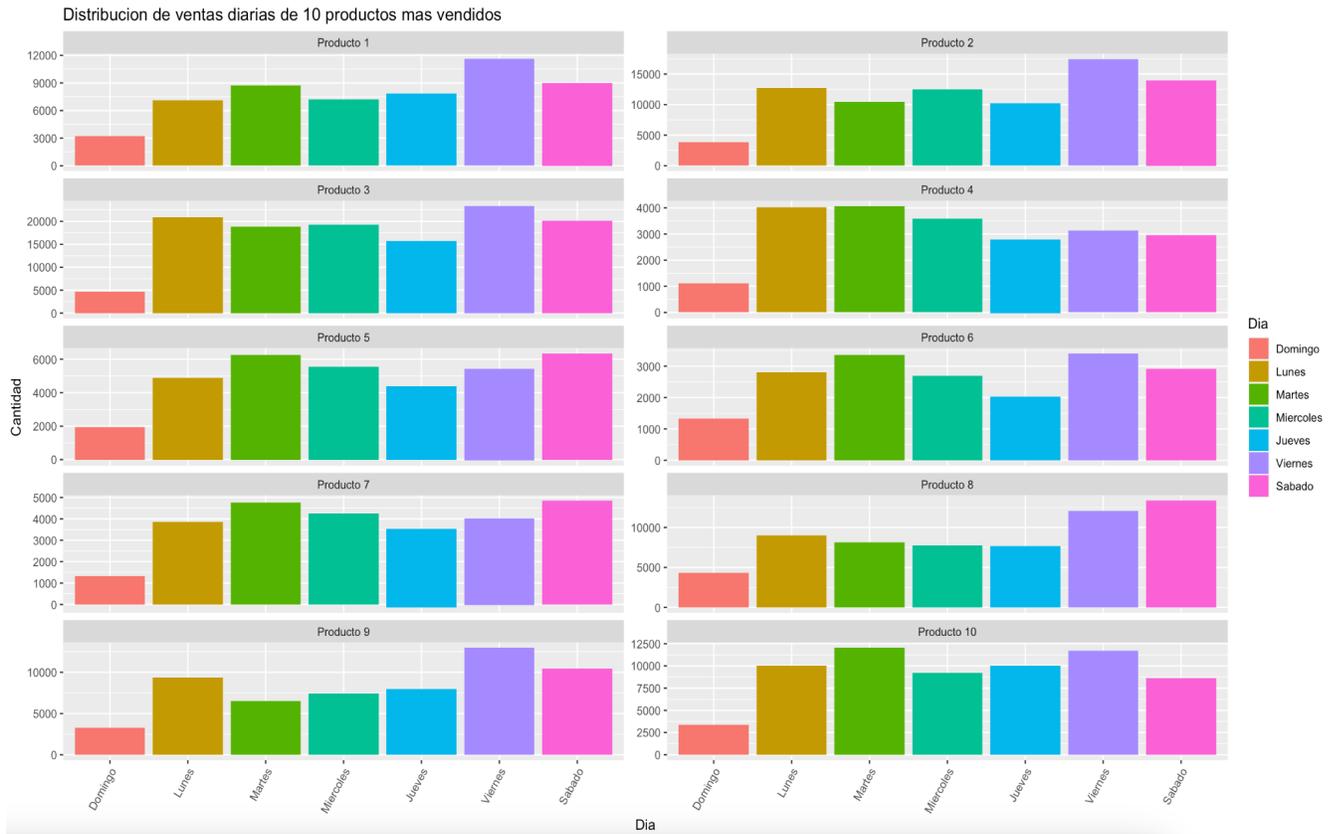
Párrafo aparte merece el día domingo, en el cual la empresa no está abierta, pero posee observaciones por pedidos puntuales de mercadería.

Figura 11. Unidades vendidas totales por día.



Similar al gráfico que mostraba la evolución mensual de las ventas del top 10 de productos más vendidos, se lo réplica pero por día de la semana.

Figura 12. Distribución de ventas diarias de 10 productos más vendidos.

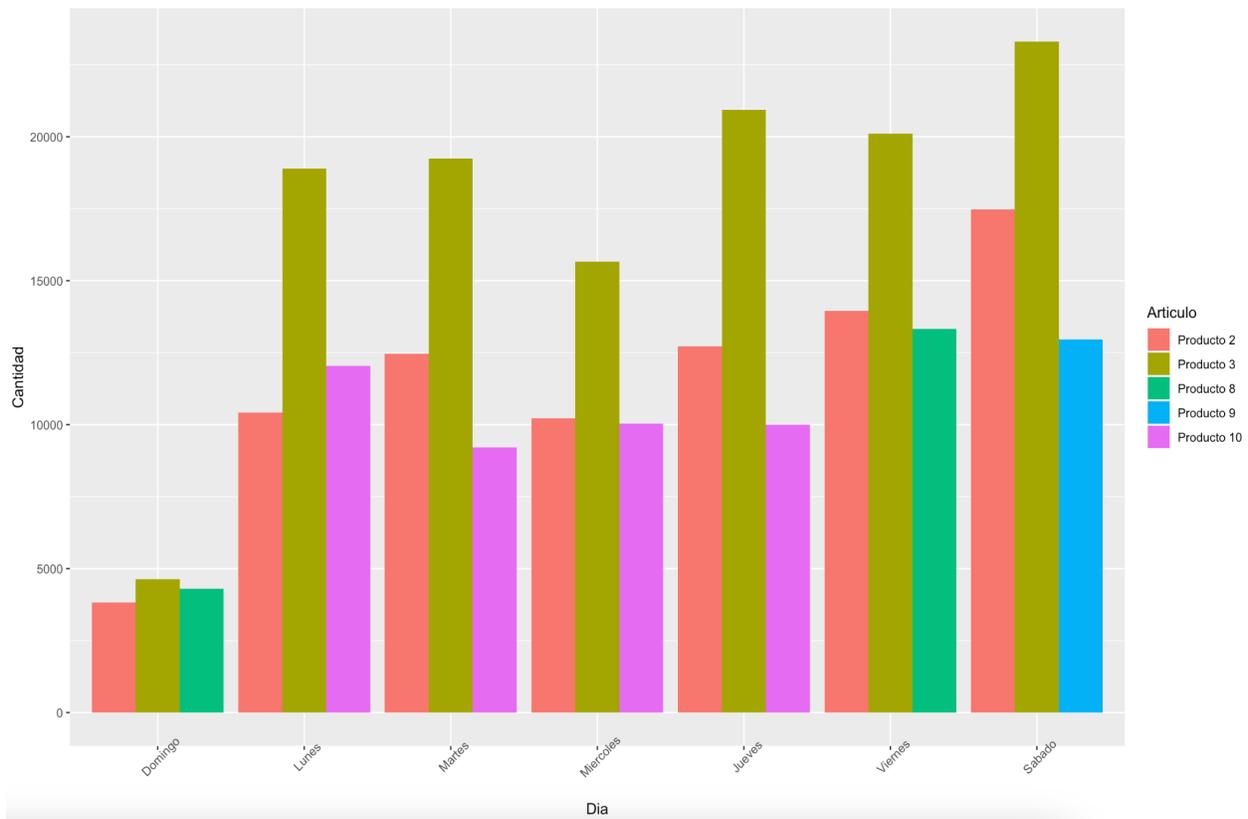


Luego, por cada día se muestran los 3 productos más vendidos.

Los productos 2 y 3 se mantienen siempre en el top 3 de los más vendidos. El producto 3 conserva siempre el primer lugar, mientras que el 2 alterna entre la segunda y tercera posición.

Lunes, martes, miércoles y jueves el producto 10 ingresa en el top, mientras que el sábado es el único día que cuenta con el producto 9 en el podio.

Figura 13. Top 3 productos vendidos por día de la semana.



1.3. Separación de entrenamiento y testeo.

La separación del conjunto de datos que se posee en entrenamiento y testeo tiene como finalidad que el modelo pueda, con los datos de entrenamiento, aprender en dicha etapa de la mejor manera los patrones implícitos en los mismos. Luego, una vez que el modelo haya sido entrenado, realizará las predicciones, las cuales serán comparadas con el conjunto de datos de testeo, para conocer qué tan bien predice el modelo.

Específicamente, los datos de entrenamiento comprenden desde el 06/2019 (Junio 2019) hasta 12/2022 (Diciembre 2022), mientras que los datos de test son los del mes de 01/2023 (Enero 2023).

1.3.1. Tipo de pronóstico Fuera de Muestra (Out of Sample)

De lo anterior se desprende el tipo de pronóstico utilizado, mediante la técnica conocida como Fuera de Muestra.

La misma consiste en que el modelo se entrene únicamente con los datos disponibles hasta el último punto del conjunto de entrenamiento (a diferencia del método Paso a Paso o *Step by Step*, que genera las predicciones secuencialmente utilizando las predicciones anteriores como entrada, inclusive las del conjunto de prueba) para luego hacer las predicciones futuras, de manera de evaluar la capacidad del modelo para generalizar y hacer predicciones precisas sobre datos nuevos y desconocidos.

2. Método utilizado por la empresa.

Actualmente, la empresa realiza sus proyecciones de ventas con la técnica de Promedios Medios Móviles. La misma consiste en, para cada artículo, obtener las ventas en una ventana de tiempo determinada, que normalmente se ubica en los 31 días anteriores. Una vez obtenido ese total se divide por la ventana de tiempo, y se obtiene la estimación de ventas de un día. Luego, procede a multiplicar ese valor por la cantidad de días que se quiere pronosticar, que usualmente son de unos 14 días.

Los datos obtenidos surgen del sistema de facturación de la empresa, el cual emite un informe con las unidades vendidas. Luego, los siguientes cálculos son efectuados manualmente por el gerente o dueño de la empresa.

Este método, aunque sencillo, es muy efectivo para la empresa, ya que no tiene grandes cambios en la demanda de sus productos. Sin embargo, es una buena oportunidad de desarrollar nuevos métodos que exploren la capacidad de mejora de pronóstico, en especial al buscar mayor

robustez ante el eventual cambio en los patrones de demanda y un ajuste más preciso ante variaciones estacionales y de tendencias. Asimismo, se buscará evitar utilizar el factor humano en tareas repetitivas y de poco valor agregado, como es el caso actual.

3. Métodos propuestos

3.1. Prophet ².

El primer método propuesto para llevar adelante las predicciones de ventas es el desarrollado por los científicos de datos de Facebook, llamado Prophet, el cual es un modelo de series de tiempo que busca ser flexible para un amplio rango de series de negocios, configurable por personas no expertas que conozcan algo del proceso de generación de datos pero no tanto de series de tiempo y modelos (Taylor & Letham, 2017).

El acercamiento se lleva adelante mediante el conocimiento de series de tiempo pronosticadas en Facebook, así como también los desafíos que envuelven los pronósticos en escala.

El modelo utiliza un modelo de serie de tiempo descomponible (Harvey & Peters, 1990) con tres componentes principales: tendencia, estacionalidad y días festivos. La ecuación resultante es la siguiente:

$$y(t) = g(t) + s(t) + h(t) + e_t \quad (12)$$

Siendo $g(t)$ la función de tendencia que modela cambios no periódicos en el valor de la serie de tiempo; $s(t)$ representando cambios periódicos (por ejemplo, estacionalidad semanal y/o

² <https://facebook.github.io/prophet/>

anual); $h(t)$ representando los efectos de días festivos. Finalmente, e_t es el término de error que representa cualquier cambio idiosincrásico que no se adapte al modelo.

La misma es similar a la del Modelo Aditivo Generalizado (GAM, por sus siglas en inglés), (Hastie & Tibshirani, 1987), una clase de modelos de regresión con potenciales suavizantes no lineales aplicados a los regresores. Para hacerlo, reemplaza los parámetros de una Regresión Lineal con una función que permite relaciones no lineales.

La función encargada de que esta tarea sea realizable es *spline*, las cuales, en conjunto, forman el GAM. Mediante la misma, los parámetros son reemplazados de la siguiente manera:

$$Z = s_0x_0 + s_1x_1 + \dots + s_nx_n \quad (13)$$

s representa el peso, igual que B . Ambas son expansiones bases, siendo s :

$$s(x) = \sum_1^k B_k b_k(x) \quad (14)$$

Esto permite una gran flexibilidad y menor linealidad, al permitir k pesos y funciones por cada variable.

La estacionalidad puede ser modelada de forma Aditiva, el cual utiliza el mismo acercamiento tomado por el método de Suavizado Exponencial (Gardner, 1985), o de forma Multiplicativa, en el cual el efecto estacional es un factor que multiplica $g(t)$, y se obtiene a través de una transformación logarítmica.

La formulación de GAM posee la ventaja de descomponer fácilmente y acomodar nuevos componentes cuando sea necesario, por ejemplo, cuando una nueva forma de estacionalidad es detectada.

El pronóstico es trabajado como un ejercicio en el que se fitea una curva, lo cual es muy diferente a otros modelos de series de tiempo que se basan en la estructura temporal de los datos. Como resultado, le otorga flexibilidad al acomodar estacionalidades con múltiples periodos, no se necesita eliminar outliers y/o interpolar valores faltantes.

En lo que respecta a costo computacional, el ajuste es realizado rápidamente, pudiendo usar dos formas: Backfitting (Breiman & Friedman, 1985) o L-BFGS (Byrd et al., 1995).

Las tendencias no lineales son ajustadas con estacionalidad anual, semanal y diaria, además de los efectos de vacaciones o días festivos. Trabaja mejor con series de tiempo con mucha estacionalidad y mucha data histórica; y es robusto a datos faltantes, cambios de tendencias y outliers (*Prophet | Forecasting at Scale.*, n.d.).

El acercamiento propuesto por Prophet, explicado en su paper, es novedoso. El mismo se propone permitir la participación del especialista en el negocio de aplicar sus conocimientos a través de un pequeño set de parámetros y opciones intuitivas, a la vez que mantiene totalmente automatizada la parte estadística y computacional (Taylor & Letham, 2017).

En particular, cada componente de la fórmula admite ajustes según la experiencias y/o preferencias de las personas encargadas de la confección del software (Hills, 2020).

Por lo expresado anteriormente, se opta por llevar adelante las predicciones a través de Prophet, al priorizar la capacidad del modelo de poseer un complejo desarrollo matemático que corre por detrás, fiteando parámetros de forma automática, pero, no obstante, dejando lugar a la participación del pronosticador en cuanto a la adición de diferentes variables que tienen que ver con las cuestiones de negocios.

3.2. ARIMA.

El modelo ARIMA (Autoregressive Integrated Moving Average, por sus siglas en inglés) es una técnica estadística utilizada para analizar y predecir series de tiempo. Este modelo combina tres componentes: el modelo autorregresivo (AR), el modelo de media móvil (MA) y la integración (I), para capturar las propiedades estacionarias y no estacionarias de la serie de tiempo.

El modelo autorregresivo (AR) se refiere a la relación lineal entre una observación y un número específico de observaciones anteriores. El modelo AR se puede expresar matemáticamente de la siguiente manera:

$$y_t = C + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (15)$$

donde y_t es el valor de la serie en el tiempo t , C es una constante, ϕ_1 a ϕ_p son los coeficientes de autorregresión, y ε_t es el término de error aleatorio en el tiempo t .

El modelo de media móvil (MA) se refiere a la relación lineal entre una observación y un número específico de errores de predicción anteriores. El modelo MA se puede expresar matemáticamente de la siguiente manera:

$$y_t = C + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (16)$$

donde y_t es el valor de la serie en el tiempo t , C es una constante, θ_1 a θ_q son los coeficientes de la media móvil, y ε_t es el término de error aleatorio en el tiempo t .

El término de integración (I) se refiere a la necesidad de diferenciar la serie de tiempo para que sea estacionaria. La estacionariedad se refiere a la propiedad de que la media y la

varianza de la serie de tiempo son constantes a lo largo del tiempo. Si la serie de tiempo no es estacionaria, se pueden aplicar diferenciaciones para lograr la estacionariedad. Por ejemplo, la primera diferencia se puede expresar como:

$$\Delta Y_t = Y_t - Y_{t-1}$$

donde ΔY es la primera diferencia entre el valor de la serie en el tiempo t y el valor de la serie en el tiempo $t - 1$.

En el modelo ARIMA, se combinan los componentes AR, MA e I para crear un modelo que puede modelar tanto las tendencias como la variabilidad aleatoria en la serie de tiempo. El modelo ARIMA se puede expresar matemáticamente de la siguiente manera:

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (17)$$

Dentro de las ventajas que motivaron su implementación se puede mencionar el hecho de que es un método ampliamente utilizado para pronósticos de series temporales, en donde se destaca por su buen desempeño a corto plazo, permitiendo obtener predicciones precisas en un horizonte temporal cercano.

Otra ventaja significativa es su capacidad para modelar datos que presentan tendencias y estacionalidades. Siendo estos últimos especialmente comunes en series mensuales, ARIMA resulta particularmente adecuado para su análisis y pronóstico; incluso a pesar de que la frecuencia de los datos sea diaria, ya que, como se pudo observar en la parte de gráficos, muchos poseen tendencias mensuales claras.

4. Diseño e implementación.

4.1. Métrica de error.

Tan importante como la predicción es la evaluación del modelo. Para ello, hay una gran variedad de métricas que comparan los valores predichos con los reales.

Estas diferentes métricas tienen ventajas y desventajas, y su elección depende del modelo utilizado, la escala de los productos, el tiempo del pronóstico, la importancia del sobre stock o bajo stock, entre otras.

En relación al problema particular, se debe señalar las siguientes características para poder elegir un buen método de medición del error:

- Diferentes escalas: hay productos vendidos por unidades y por kilogramo.
- Diferentes demandas: hay productos con demandas de grandes cantidades y otras con muy pocas.

Por lo anteriormente señalado, la elección es el Error Porcentual Absoluto Ponderado (WAPE, por sus siglas en inglés).

$$WAPE = \frac{\sum_{t=1}^n |A_t - F_t|}{\sum_{t=1}^n |A_t|} \quad (18)$$

Esta métrica se obtiene de realizar la sumatoria de las diferencias entre el pronóstico y el valor real, dividido por la sumatoria de los valores reales de ventas. De esta forma, se pondera atendiendo la demanda total de cada producto, expresado como porcentaje. Esto tiene dos ventajas:

- Permite comparar productos con diferentes demandas, ya que compara el error de cada producto respecto a sus ventas.
- Al expresar el error en porcentaje no se tiene problema en comparar productos con distintas unidades.

4.2. Métrica de éxito.

El éxito o no del modelo estará dado por si obtiene una menor tasa de error medido con metodología WAPE que el que poseen las predicciones del modelo actual de la empresa.

4.3. Búsqueda de grilla.

La búsqueda de grilla es un procedimiento habitual para buscar las mejores configuraciones de hiperparametros en modelos de aprendizaje automático.

Consiste en generar una lista con determinados valores susceptibles de ser utilizados en cierto hiper parámetro del modelo; luego, mediante un bucle se realiza la predicción y medición para cada combinación y se guardan los resultados para, finalmente, y luego de haber iterado por todas las posibles combinaciones, elegir la que posea la mejor métrica que se buscaba alcanzar.

Si bien existen métodos de búsqueda predefinidos, en este caso no fueron utilizados.

Prophet.

Los parámetros optimizados en la grilla correspondientes a Prophet son los siguientes:

- `changepoint_prior_scale`: 0.01, 0.05 (default), 0.1.
- `seasonality_mode`: additive (default), multiplicative.

changeoint_prior_scale: Este parámetro se utiliza para controlar la flexibilidad del modelo en la identificación de puntos de cambio en la tendencia de la serie de tiempo. Un valor mayor de changeoint_prior_scale significa que el modelo será más flexible para ajustar cambios en la tendencia, mientras que un valor menor lo hace menos flexible. Los valores predeterminados de este parámetro son 0.05, pero en la grilla de parámetros se probarán, además, valores de 0.01 y 0.1 para encontrar el mejor ajuste para los datos.

seasonality_mode: Este parámetro se utiliza para especificar si la estacionalidad en la serie de tiempo es aditiva o multiplicativa. Si se especifica "additive", se asume que la estacionalidad tiene un efecto lineal en la serie de tiempo. Si se especifica "multiplicative", se asume que la estacionalidad tiene un efecto exponencial en la serie de tiempo. El valor predeterminado de este parámetro es "additive", pero en la grilla de parámetros se probará también "multiplicative" para encontrar el mejor ajuste para los datos.

La técnica a optimizar fue el error, encontrando el menor con la configuración de 0.1, multiplicative.

ARIMA.

Los parámetros optimizados en la grilla correspondientes a ARIMA son los siguientes:

- p: 0, 1, 2.
- d: 0, 1.
- q: 0, 1, 2.

p: El parámetro "p" se refiere al orden del componente de autorregresión (AR) del modelo. El orden del componente AR se refiere al número de valores de la serie de tiempo previos que se utilizan para predecir el valor actual. Por lo tanto, el parámetro "p" se refiere al

número de valores previos que se utilizarán en el modelo. En la grilla de parámetros mencionada en la pregunta, se probarán valores de 0, 1 y 2 para el parámetro "p".

d: El parámetro "d" se refiere al número de veces que se debe diferenciar la serie de tiempo para hacerla estacionaria. La diferenciación se utiliza para eliminar tendencias no deseadas y hacer que la serie de tiempo sea más estacionaria. En la grilla de parámetros mencionada en la pregunta, se probarán valores de 0 y 1 para el parámetro "d".

q: El parámetro "q" se refiere al orden del componente de media móvil (MA) del modelo. El orden del componente MA se refiere al número de valores de error previos que se utilizan para predecir el valor actual. Por lo tanto, el parámetro "q" se refiere al número de errores previos que se utilizarán en el modelo. En la grilla de parámetros mencionada en la pregunta, se probarán valores de 0, 1 y 2 para el parámetro "q".

La técnica a optimizar fue el error, encontrando el menor con la configuración de 2, 1, 2.

La grilla reveló cómo mejores opciones aquellos términos autorregresivos "p" y de diferenciación "q" más elevados, por lo que se intentó aumentar los mismos a 3 para evaluar si esto contribuía a una mejora en las predicciones del modelo. La configuración resultante de esta prueba fue 3,1,3. Sin embargo, el rendimiento fue mejor con la configuración originalmente descubierta de 2,1,2.

4.4. Tipo de pronóstico Fuera de Muestra (Out of Sample)

Para especificar al modelo el tipo de pronóstico deseado, primero se separaron los datos en entrenamiento y prueba, y luego se realizaron las predicciones con los métodos específicos de cada modelo.

En Prophet, mediante el llamado a *predict()*, de la librería que lleva el nombre del modelo; mientras que en ARIMA es mediante el llamado a *forecast()*, de la librería *statsmodel*.

4.5. Tratamiento de pandemia COVID 19

No se realizaron tratamientos especiales en los modelos para incorporar los efectos de la pandemia COVID 19.

La incorporación de dicho efecto al modelo debe realizarse a través de variables exógenas que requieren supuestos, recopilación y procesamiento de datos adicionales, como la cantidad de muertes o contagios, variables categóricas para indicar si cierta fecha se trataba de un momento de cuarentena, entre otros. Sin embargo, no hay una técnica clara para modelar de manera óptima sus consecuencias en las series de tiempo, y usualmente varían según cada caso en particular y del modelo predictivo utilizado.

En este marco, se priorizo la tarea de pronosticar la demanda con un nivel de error similar o mejor al modelo actualmente empleado por la empresa. Al haber conseguido dicho fin, tal y como se podrá apreciar en la próxima sección de Resultados, se optó por utilizar sólo los patrones históricos, sin incluir factores externos.

IV. Resultados.

1. Tabla comparativa de resultados.

A continuación se presenta una tabla resumida con los 10 primeros productos, que muestra, para cada producto:

- Cantidad Real: Cantidad vendida.
- y_{ARIMA} : Valor predicho por el modelo ARIMA propuesto
- $WAPE_{ARIMA}$: Error del modelo ARIMA propuesto.
- $y_{Prophet}$: Valor predicho por el modelo Prophet propuesto.
- $WAPE_{Prophet}$: Error del modelo Prophet propuesto.
- $y_{Benchmark}$: Valor predicho por el modelo de la empresa.
- $WAPE_{Benchmark}$: Error del modelo de la empresa.

Tabla 1. Tabla comparativa de errores de los 10 primeros productos de la lista.

Productos	Cantidad Real	y_{ARIMA}	$WAPE_{ARIMA}$	$y_{Prophet}$	$WAPE_{Prophet}$	$y_{Benchmark}$	$WAPE_{Benchmark}$
Producto 1	76,00	41,09	0,46	40,77	0,46	22	0,71
Producto 2	153,00	145,55	0,05	157,53	0,03	140	0,08
Producto 3	31,00	26,03	0,16	27,78	0,1	25	0,19
Producto 4	57,00	52,95	0,07	57,51	0,01	52	0,09
Producto 5	12,00	14,69	0,22	16,27	0,36	13	0,08
Producto 6	8,00	12,20	0,53	12,08	0,51	11	0,38
Producto 7	39,00	32,86	0,16	28,46	0,27	30	0,23
Producto 8	17,00	22,78	0,34	23,77	0,4	29	0,71
Producto 9	56,00	52,58	0,06	44,22	0,21	58	0,04
Producto 10	20,00	24,43	0,22	25,49	0,27	29	0,45
...

ERRORES TOTALES		0,272		0,336		0,359
-----------------	--	-------	--	-------	--	-------

Estos valores son los totales del mes de Enero.

El error medido con técnica WAPE obtenido, con respecto al modelo actual de la empresa, es 2,3 puntos porcentuales menor con el modelo propuesto de Prophet, y 8,7 puntos porcentuales menor con el modelo propuesto de ARIMA, existiendo una diferencia de 6,4 puntos porcentuales entre estos dos últimos.

Aunque un diseño de grilla de hiperparametros distinta puedan arrojar resultados diferentes, la mejor performance de ARIMA puede tener base en el hecho de que es un modelo que se adapta de manera rápida y precisa a series temporales con componentes estacionales regulares y estables en el tiempo, al detectar patrones con frecuencia fija y regular. Por otro lado, Prophet puede proporcionar mejores resultados en series temporales con patrones estacionales más complejos y no estacionarios, especialmente cuando la amplitud de la estacionalidad varía en el tiempo.

2. Significancia estadística de la diferencia en los errores

Para complementar el análisis se llevó adelante un análisis de significancia estadística para determinar si las diferencias observadas en los errores de los dos modelos son reales o provenientes del azar.

Para ello, primero se llevó adelante una prueba para determinar si la muestra sigue un patrón paramétrico o no paramétrico. Se utilizó la Prueba de Normalidad Shapiro Wilk (Shapiro & Wilk, 1965), en la cual la hipótesis nula es que es una muestra normalmente distribuida, de la librería *scipy.stats*, de Python.

La misma arrojó los siguientes valores:

- Estadístico de prueba: 0.97
- p-valor: 0.0035

Con un nivel de significancia del 5%, y un p-valor de 0.0035, se puede rechazar la hipótesis nula de que la muestra está normalmente distribuida.

Sabiendo que la muestra es no paramétrica, y dado que es una muestra pareada ya que las diferencias se obtuvieron sobre el mismo conjunto de artículos a los que se les aplicaron métodos predictivos diferentes, se decidió usar, para evaluar la significatividad de las diferencias observadas, el test de Wilcoxon (Wilcoxon, 1945).

En este, la hipótesis nula es que no existen diferencias (en términos de tendencia central) entre ambos grupos.

Se utilizó la implementación de la librería *scipy.stats*, de Python, que arrojó los siguientes valores:

- Estadístico de prueba: 3554
- Valor p: 0.0006

Con un nivel de significancia del 5%, y un p-valor de 0.0006, se puede rechazar la hipótesis nula de que no existen diferencias entre las muestras, por lo que las diferencias son estadísticamente significativas.

3. Gráficos.

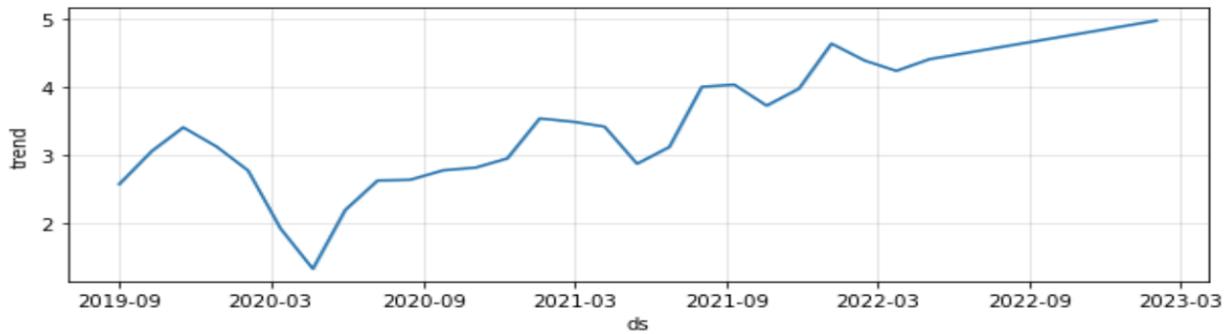
Antes de presentar los diversos gráficos, explicaré ciertos términos comunes que serán utilizados en las próximas páginas.

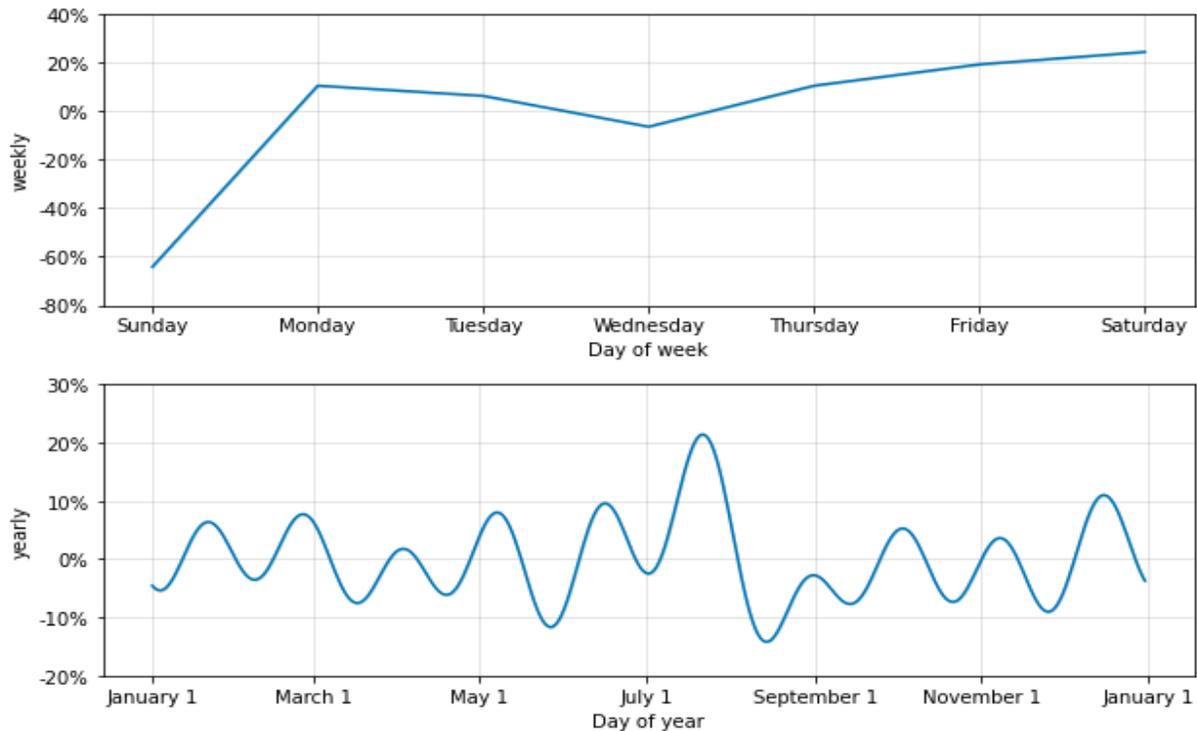
La mercadería inmovilizada hace referencia a aquella disponible al final del mes que no fue vendida, proveniente de la diferencia entre el valor de ventas predicho y las ventas reales.

El stockout fue obtenido mediante la diferencia de las ventas reales y las predichas, en aquellos casos en los que las ventas reales fueron mayores que las predichas se considera que se incurrió en faltante de mercadería para hacer frente a la demanda.

3.1. Prophet.

Figura 14. Componentes de pronóstico Prophet.





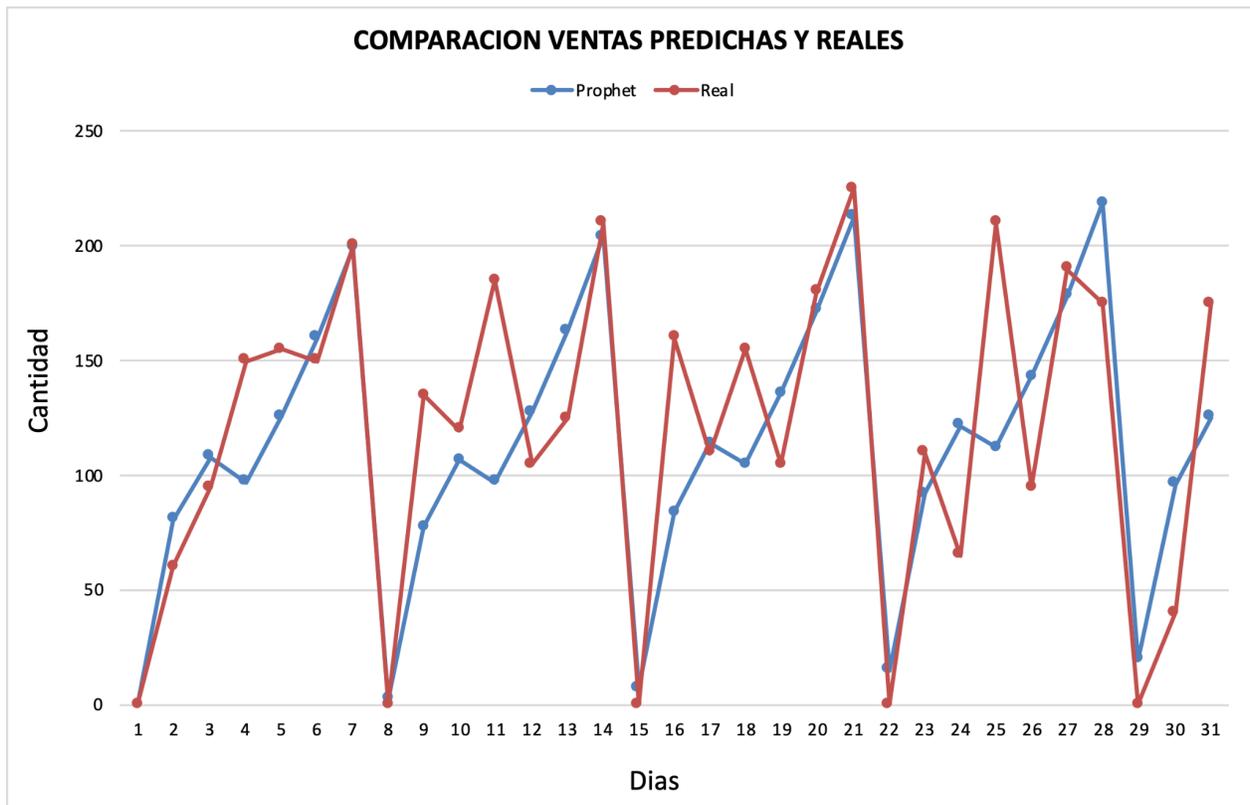
El gráfico divide los componentes en tendencia, estacionalidad semanal y anual.

La tendencia muestra una caída que tiene su expresión máxima en Abril 2020, con la pandemia de Coronavirus. A partir de allí, siempre fue positiva.

El componente semanal refleja información similar a los gráficos previamente observados. El lunes las ventas crecen, tienen un pequeño descenso del martes al miércoles, y luego consiguen su máximo el día sábado.

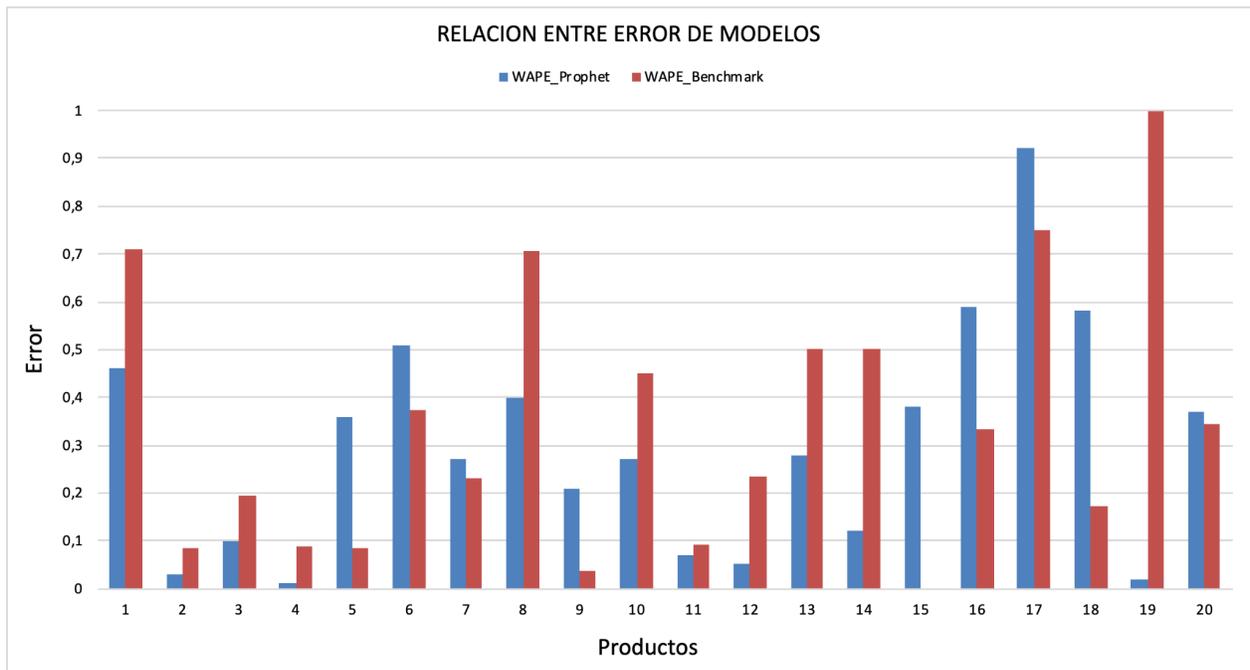
El componente anual muestra el pico en Julio, seguido de los meses de Diciembre, Enero y Febrero, lo cual es lógico ya que coinciden con las etapas de receso invernal y con los meses de verano.

Figura 15. Comparación de ventas reales y predichas por Prophet.



En este gráfico se puede observar un ejemplo de cómo Prophet predice la cantidad demandada de un artículo, para cada día del mes de Enero, representada por la línea azul; mientras que la línea roja representa la demanda real. Se observa como Prophet entiende muy bien las tendencias de ventas de cada día según, por ejemplo, el momento de la semana o del mes.

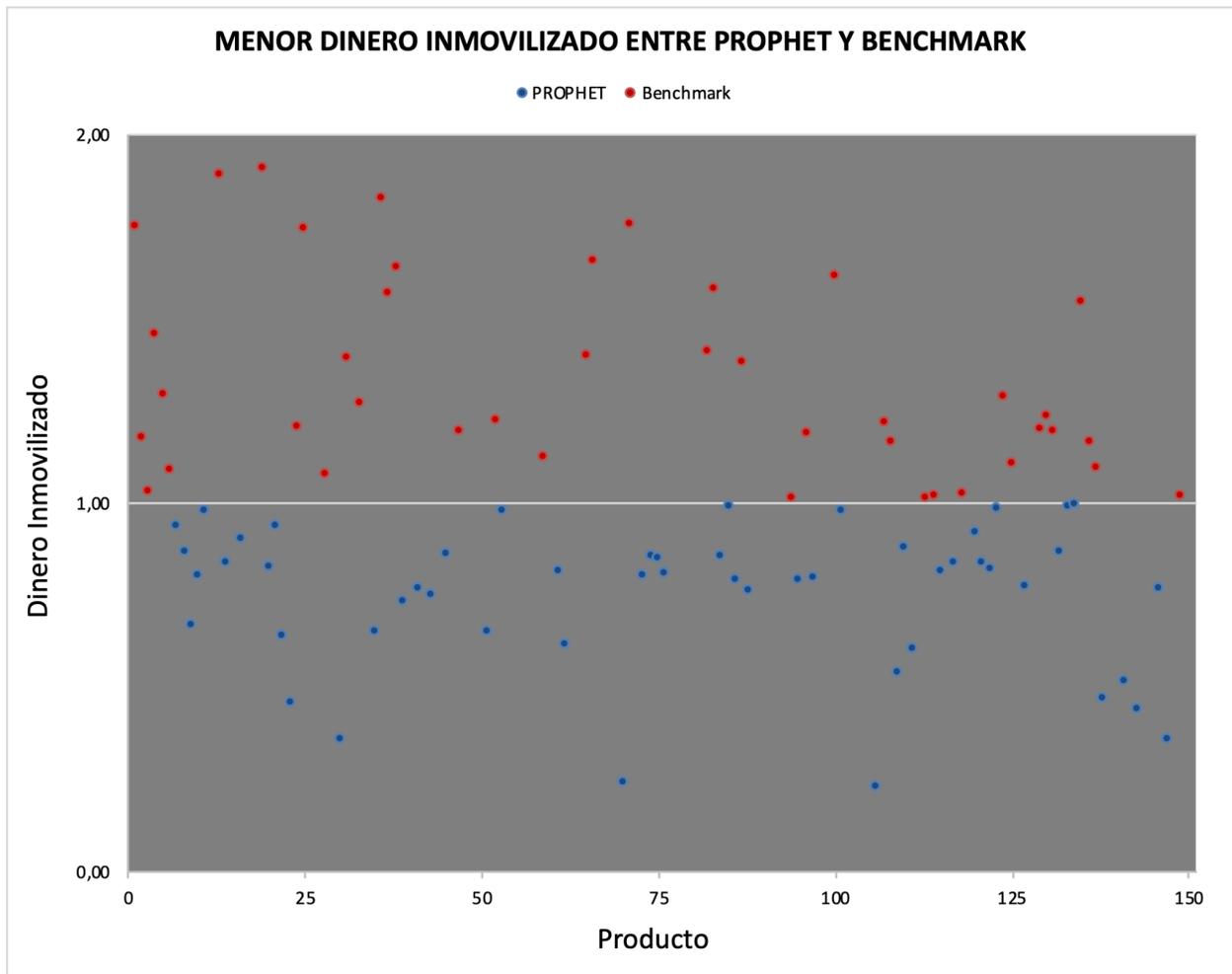
Figura 16. Relación de errores entre modelo Prophet y Benchmark.



Mediante un gráfico de columnas agrupadas, cada una representa el error de cada modelo, y se agrupan según qué artículo se trate. Las barras rojas representan el error del modelo de la empresa y las azules el de Prophet. La magnitud de cada error está representada por la altura de la columna.

Se utilizaron los primeros 20 productos para priorizar la visualización.

Figura 17. Relación entre dinero inmovilizado Prophet y Benchmark.



Este gráfico destaca los artículos en los que la predicción del modelo actual de la empresa resultó en una cantidad menor de mercadería inmovilizada al final del mes, indicados por puntos de color rojo; mientras que los puntos de color azul resaltan los artículos para los cuales la predicción del modelo Prophet fue más precisa. A medida que los puntos se alejan del 1, la magnitud de la cantidad de mercadería inmovilizada al final del mes es mayor.

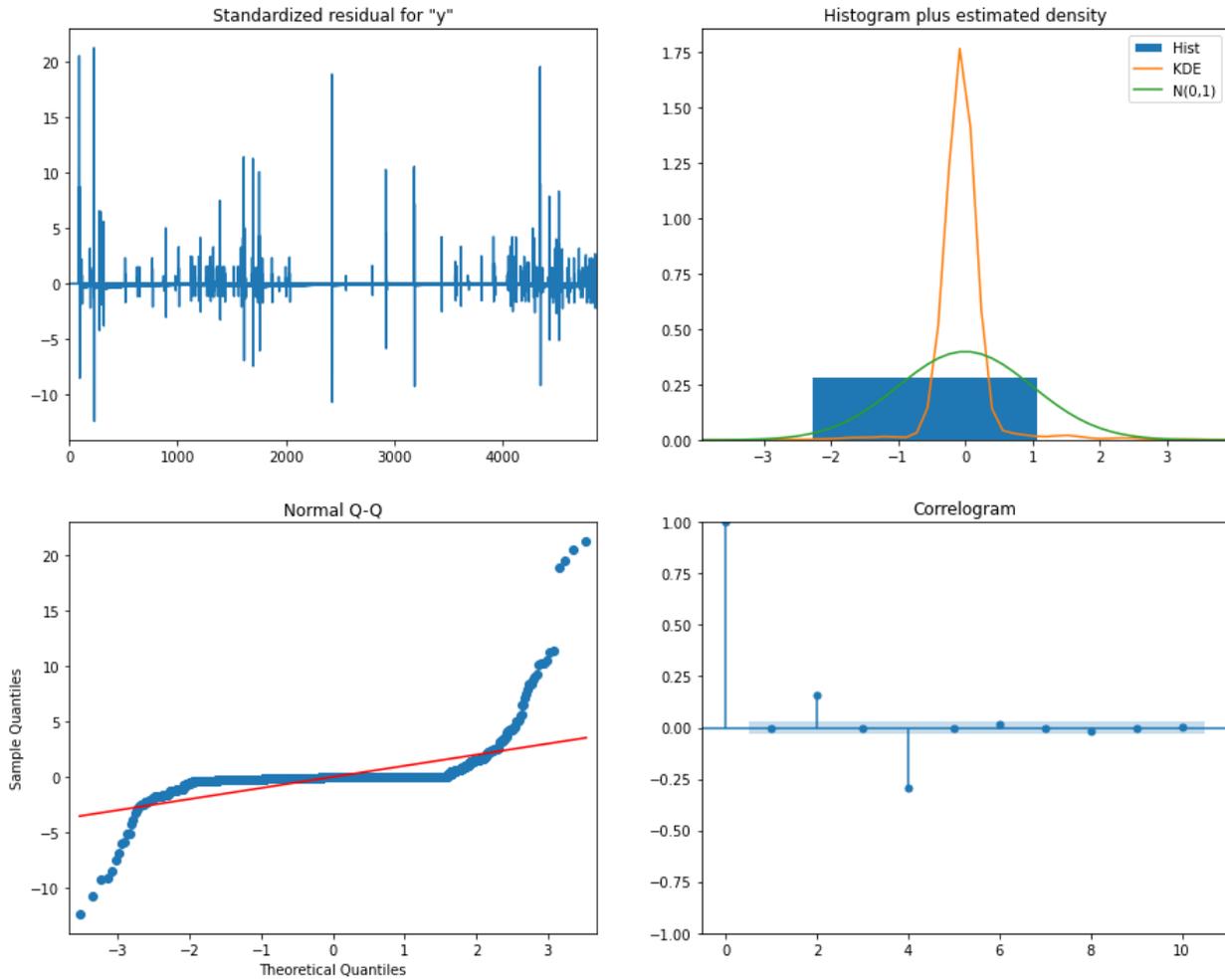
Al proveer una representación clara de la precisión de los modelos y permitir una comparación rápida entre el modelo actual y el modelo Prophet, la información proporcionada por el gráfico puede ayudar a la empresa a optimizar su gestión de inventario. Al identificar los

artículos con una mayor cantidad de mercadería inmovilizada al final del mes (puntos alejados del 1), la empresa puede enfocar sus esfuerzos en reducir estos excedentes, lo que puede conducir a una mayor eficiencia y reducción de costos.

Por último, el modelo utilizado por la empresa superó en este rubro a Prophet, ya que permitió ahorrar un 2,12 puntos porcentuales en términos monetarios la mercadería inmovilizada. Además, este ahorro no lo hizo caer en una mayor cantidad de stockouts de mercadería, ya que solo lo hizo para 22 artículos, frente a los 30 productos en los que lo hizo Prophet.

3.2. ARIMA.

Figura 18. Componentes del modelo ARIMA.

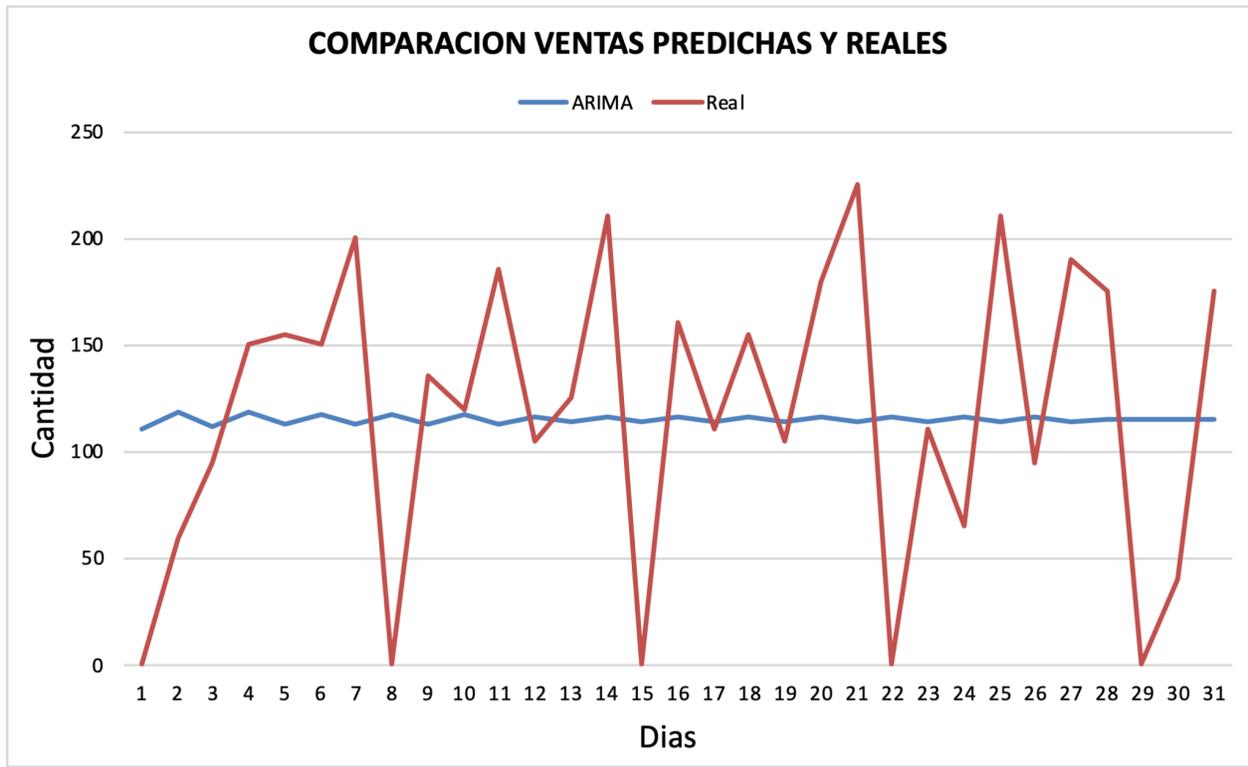


- Standardized residual for y (Residuales estandarizados para y): Este gráfico muestra los residuos estandarizados, que son los residuos divididos por la desviación estándar de los residuos. Los residuos estandarizados se utilizan para identificar valores atípicos y para verificar la homocedasticidad de los residuos. Si los residuos estandarizados tienen una varianza constante, deberían estar distribuidos de manera uniforme en torno a cero. Se puede observar que este es mayoritariamente el caso.

- Normal Q-Q Plot (Gráfico cuantil-cuantil normal): Este gráfico compara los cuantiles de la muestra de residuos con los cuantiles de una distribución normal teórica. Si los residuos se distribuyen normalmente, los puntos en el gráfico seguirán aproximadamente una línea recta. Los puntos que se desvían de la línea recta indican desviaciones de la normalidad en los residuos. En este caso, la mayor parte se ubica sobre esta línea.
- Histogram plot estimated density (Histograma de la densidad estimada): Este gráfico muestra un histograma de los residuos con una curva de densidad estimada superpuesta (KDE). Si los residuos se distribuyen normalmente, la curva de densidad debería ser simétrica y centrada en torno a cero, lo cual es cierto en este caso.
- Correlogram (Correlograma): Este gráfico muestra los coeficientes de correlación entre los residuos en función de la diferencia en el tiempo entre las observaciones. Si los residuos son aleatorios, los coeficientes de correlación deberían ser cercanos a cero y no significativos. En este caso, los rezagos 2 y 4 son significativos, por lo cual el modelo no estaría capturando de la mejor manera la dinámica de los datos en esos puntos.

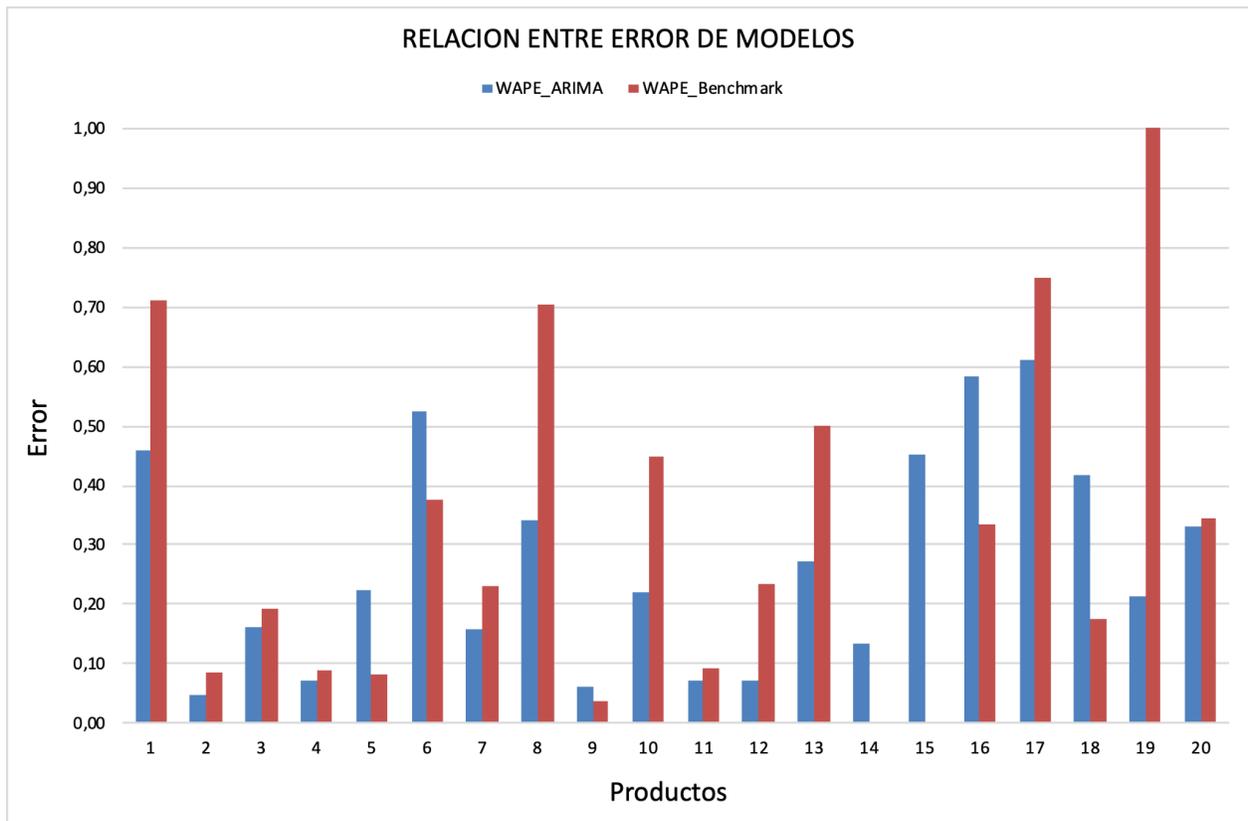
Los gráficos anteriores son interesantes a los efectos de entender si los resultados del modelo siguen los parámetros señalados, ya que poseer diferencias significativas podría no ser el modelo adecuado. En este caso se observa que en líneas generales se ajusta a las formas normales.

Figura 19. Comparación de ventas reales y predichas por ARIMA.



En este gráfico se puede observar un ejemplo de cómo ARIMA predice la cantidad demandada de un artículo, para cada día del mes de Enero, representada por la línea azul; mientras que la línea roja representa la demanda real. Se observa como ARIMA no varía la cantidad predicha como Prophet, sino que tiende a la estacionariedad.

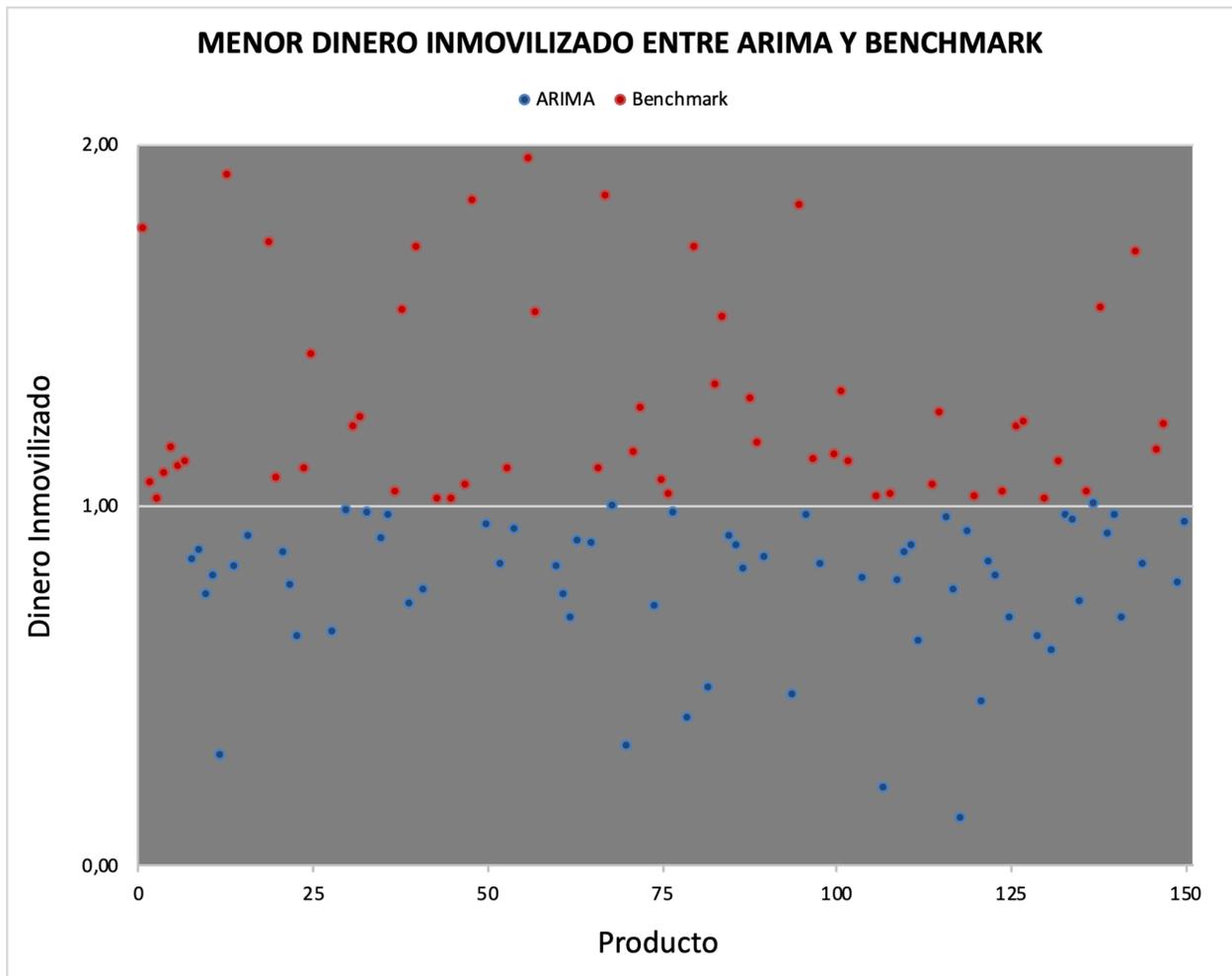
Figura 20. Relación de errores entre modelo ARIMA y Benchmark.



Mediante un gráfico de columnas agrupadas, cada una representa el error de cada modelo, y se agrupan según qué artículo se trate. Las barras rojas representan el error del modelo de la empresa y las azules el de ARIMA. La magnitud de cada error está representada por la altura de la columna.

El error del benchmark de los productos 14 y 15 fue muy cercano a 0, lo que explica esa inexistencia en la columna roja.

Figura 21. Relación entre dinero inmovilizado ARIMA y Benchmark.



Este gráfico destaca los artículos en los que la predicción del modelo actual de la empresa resultó en una cantidad menor de mercadería inmovilizada al final del mes, indicados por puntos de color rojo; mientras que los puntos de color azul resaltan los artículos para los cuales la predicción del modelo ARIMA fue más precisa. A medida que los puntos se alejan del 1, la magnitud de la cantidad de mercadería inmovilizada al final del mes es mayor.

Al proveer una representación clara de la precisión de los modelos y permitir una comparación rápida entre el modelo actual y el modelo ARIMA, la información proporcionada por el gráfico puede ayudar a la empresa a optimizar su gestión de inventario. Al identificar los

artículos con una mayor cantidad de mercadería inmovilizada al final del mes (puntos alejados del 1), la empresa puede enfocar sus esfuerzos en reducir estos excedentes, lo que puede conducir a una mayor eficiencia y reducción de costos.

Por último, el modelo ARIMA propuesta por la empresa superó aquí también al modelo de la empresa, ya que permitió ahorrar un 6,59 puntos porcentuales en términos monetarios la mercadería inmovilizada. En el rubro de stockouts, ARIMA igualó la cantidad del modelo de la empresa, cayendo en el mismo en 22 oportunidades.

V. Implementación y calibración.

La implementación propuesta consiste en un script en lenguaje de programación Python, en el que se carguen los datos históricos hasta la última semana de la que se posee información.

Primero, esta última semana será utilizada en el reentrenamiento del modelo, como datos de test; y luego, se la incorporará como dato de entrenamiento para el próximo periodo a pronosticar. Se ejecuta el código y se obtienen las predicciones.

En cuanto a la calibración, cada una ventana de tiempo determinada, que se propone en 3 meses, se puede generar nuevamente la búsqueda de los mejores hiperparametros para el modelo ARIMA. Esto con el objetivo de evaluar si estos siguen siendo óptimos o, por el contrario, deben ser sujetos a ajuste.

VI. Conclusiones

El objetivo del presente trabajo fue desarrollar un modelo de pronóstico de demanda automático, con poca intervención humana y que sea escalable. No obstante, el método actual de la empresa, aunque no cumplía con estos objetivos, si era muy efectivo en cuanto a los resultados que arrojaba.

Es por ello que no solo fue necesario buscar algún algoritmo que permitiera automatizar el proceso, sino que además se debieron probar diferentes configuraciones de hiperparamétricos en distintos modelos de predicción.

Finalmente, los resultados fueron abordados desde una perspectiva integral. Un pronóstico más preciso de la demanda es solo una parte, pero no suficiente. Luego, se deben analizar otras variables, como la cantidad de dinero inmovilizado en mercadería y la cantidad de veces que se cayó en stockout, por ejemplo.

Así, bajo la primera condición, Prophet y ARIMA serían opciones equivalentes. Sin embargo, al analizar las demás variables, se observó que ARIMA tenía una ventaja no solo respecto a la precisión del pronóstico, sino también en el ahorro de mercadería inmovilizada y la disminución de stockouts. Por lo tanto, el método propuesto es ARIMA.

VII. Futuras líneas de investigación.

Como futuras actualizaciones al presente trabajo, se podrían trabajar en la exploración de nuevos métodos de pronósticos y la evaluación de otras variables que pudieran resultar importantes a la hora de realizar la predicción de demanda.

Con respecto a un tópico no abarcado en el presente trabajo, se podría evaluar los mejores métodos de pronóstico para predecir la demanda de nuevos productos. Un posible acercamiento podría ser el agrupamiento por líneas de producto, enmarcar el nuevo producto en una de estas, y hacer el pronóstico en base a ella.

Referencias

Ali, O. G., Sayin, S., van Woensel, T., & Fransoo, J. (2009). SKU demand forecasting in the presence of promotions. *Expert Systems with Applications*. doi:10.1016/j.eswa.2009.04.052

Alon, I., Qi, M., & Safdowsky, R. J. (2001, Mayo). Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, 8(3), 147-156. [https://doi.org/10.1016/S0969-6989\(00\)00011-4](https://doi.org/10.1016/S0969-6989(00)00011-4)

Breiman, L., & Friedman, J. H. (1985). Estimating optimal transformations for multiple regression and correlations. *Journal of the American Statistical Association*, 80(391), 614-619. <https://doi.org/10.2307/2288477>

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16(5), 1190-1208. <https://doi.org/10.1137/0916069>

Choi, T.-M., Hui, C.L., Liu, N., Ng, S.-F., & Yu, Y. (2014, Marzo). Fast fashion sales forecasting with limited data and time. *Decision Support Systems*, 59, 84-92. <https://doi.org/10.1016/j.dss.2013.10.008>

Fildes, R., Ma, S., & Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*, 38(4), 1283 - 1318. <https://doi.org/10.1016/j.ijforecast.2019.06.004>

Gardner, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1), 1-28. <https://doi.org/10.1002/for.3980040103>

Hanke, J. (2010). *Pronosticos en Los Negocios* (9na ed.). Pearson Educación.

Harvey, A. C., & Peters, S. (1990). Estimation Procedures for Structural Time Series Models. *Journal of Forecasting*, 9, 89-108.

http://www.stat.yale.edu/~lc436/papers/Harvey_Peters1990.pdf

Hastie, T., & Tibshirani, R. (1987, Jun.). Generalized Additive Models. *Journal of the American Statistical Association*, 82(398), 371-386. <https://doi.org/10.2307/2289439>

Hills, J. (2020, August 22). *Facebook Prophet. (Almost) everything you should know to...* | by Moto DEI | *The Startup*. Medium. Retrieved December 31, 2022, from <https://medium.com/swlh/facebook-prophet-426421f7e331>

McCarthy, T. M., Davis, D. F., Golicic, S. L., & Mentzer, J. T. (2006). The Evolution of Sales Forecasting Management: A 20-Year Longitudinal Study of Forecasting Practices. *Journal of Forecasting*, 25(5), 303-324. doi.org/10.1002/for.989

Mentzer, J. T., & Cox, J. E. (1984). Familiarity, application, and performance of sales forecasting techniques. *Journal of Forecasting*, 3(1), 27-36. doi.org/10.1002/for.3980030104

Mentzer, J. T., & Kahn, K. B. (1995). Forecasting technique familiarity, satisfaction, usage, and application. *Journal of Forecasting*, 14(5), 465-476. doi.org/10.1002/for.3980140506

Montemayor Gallegos, J. E. (2014). *Métodos de pronósticos para negocios*. Instituto Tecnológico y de Estudios Superiores de Monterrey. <http://hdl.handle.net/11285/621230>

Prophet | Forecasting at scale. (n.d.). Meta Open Source. Retrieved December 27, 2022, from <https://facebook.github.io/prophet/>

Taylor, S., & Letham, B. (2017). *Forecasting at scale*. PeerJ Preprints. <https://doi.org/10.7287/peerj.preprints.3190v2>

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36, 54-76.

doi.org/10.1016/j.ijforecast.2019.04.014

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38, 1346-1364.

doi.org/10.1016/j.ijforecast.2021.11.013

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). The M5 competition: Background, organization, and implementation. *International Journal of Forecasting*, 38(4),

1325-1336. doi.org/10.1016/j.ijforecast.2021.07.007