

**Tipo de documento:** Tesis de maestría

*Master in Management + Analytics*

# Modelos de forecasting y optimización de compra de insumos para el mantenimiento de surtidores

Autoría: *Luque, Macarena Melina*

Año académico: 2023

## ¿Cómo citar este trabajo?

Luque, M. (2023) "Modelos de forecasting y optimización de compra de insumos para el mantenimiento de surtidores". [*Tesis de maestría. Universidad Torcuato Di Tella*]. Repositorio Digital Universidad Torcuato Di Tella

<https://repositorio.utdt.edu/handle/20.500.13098/12101>

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>



## Master in Management + Analytics

“Modelos de forecasting y optimización de compra de insumos para el mantenimiento de surtidores”

Macarena Melina Luque

Tutor: Javier Marengo

---

Mayo 2023

## Resumen

Este trabajo final de maestría se centra en mejorar los procesos de compra y gestión de insumos de una PYME argentina especializada en el servicio técnico integral de mantenimiento de surtidores de combustible líquido.

Los insumos son componentes que desempeñan un papel fundamental en el correcto funcionamiento de los surtidores que las estaciones de servicio atendidas por el negocio poseen. Es por este motivo que resulta crucial gestionar de manera óptima la adquisición de cada uno de los insumos. Para ello, en el presente trabajo, se entrenan distintos algoritmos de aprendizaje supervisado de tipo ARIMA con el objetivo de predecir la demanda de los cinco insumos más utilizados por el negocio. Posteriormente, se desarrolla un modelo de optimización fundamentado en programación lineal para determinar la cantidad y el mes de adquisición de dichos insumos con el objetivo de minimizar sus costos de compra.

En primer lugar, se analiza información histórica brindada por la empresa para identificar los cinco insumos más demandados por sus clientes. En segundo lugar, se entrenan cinco modelos de machine learning de tipo SARIMA para cada uno de los insumos seleccionados. Estos modelos permiten predecir la cantidad de insumos que los clientes demandarán durante el período comprendido entre julio 2023 y junio 2024. Los resultados de los modelos SARIMA se utilizan como componente de un modelo de optimización de tipo lot sizing para minimizar los costos de adquisición de los cinco insumos.

Luego de plantear tres posibles escenarios de fluctuación de demanda, se recomienda que la empresa adquiera todos los componentes durante los primeros cinco meses del periodo analizado. Esto permitirá optimizar el proceso de compra, reducir costos innecesarios y maximizar el rendimiento del capital disponible.

Este trabajo destaca la importancia de combinar los conocimientos del negocio con las herramientas de análisis de datos para que la empresa pueda tomar decisiones fundamentadas en información y así lograr optimizar sus procesos de compra, minimizando sus costos y maximizando sus beneficios.

## Abstract

This thesis focuses on improving the procurement and supply management processes of an Argentine SME specialized in providing integral technical services for the maintenance of liquid fuel dispensers.

The spare parts are components that play a fundamental role in the proper functioning of the dispensers owned by the service stations served by the business. Therefore, it is crucial to manage the acquisition of each spare part optimally. In this thesis, various supervised learning algorithms, such as ARIMA models, are trained to predict the demand of the five components most used by the business. Subsequently, a linear programming-based optimization model is developed to determine the quantity and month of acquisition for each component with the aim of minimizing total purchase cost.

Firstly, historical information provided by the company is analyzed to identify the five spare parts most demanded by the customers. Secondly, five SARIMA machine learning models are trained for each selected component. These models allow predicting the quantity of each component that customers will demand between July 2023 - June 2024. SARIMA models' outputs are used as inputs of a lot sizing optimization model to minimize the acquisition costs of the five spare parts.

After considering three possible demand fluctuation scenarios, it is recommended that the company acquires all five spare parts during the first five months of the analyzed period. This approach will optimize the purchasing process, reduce unnecessary costs, and maximize the utilization of available capital.

This thesis highlights the importance of combining business knowledge with data analysis tools to enable the company to make data-driven decisions and achieve optimization of its purchasing processes, minimizing costs and maximizing benefits.

## Índice

1. Capítulo 1. Introducción.....	6
1.1 Contexto.....	6
1.2 Problema.....	8
1.3 Objetivo.....	8
2. Capítulo 2. Datos.....	9
2.1 Descripción de las bases de datos.....	9
2.1.1 Diagrama Entidad/Relación (DER).....	11
2.2 Creación de la tabla “df_incidencias”.....	11
2.3 Análisis exploratorio de los datos.....	12
2.3.1 Exploración Macro.....	12
2.3.2 Zoom in Insumos.....	14
2.4 Input final del negocio y elección de insumos a predecir.....	17
2.4.1 Zoom in Insumos Consumibles.....	18
2.4.2 Zoom in Estaciones de Servicio.....	22
2.5 Análisis Exploratorio de Datos – Conclusión final.....	24
3. Capítulo 3. Forecasting.....	25
3.1 Modelo ARIMA.....	25
3.1.1 Modelo ARIMA: Introducción.....	25
3.1.2 Modelo ARIMA: Ejemplos.....	25
3.1.3 Modelos ARIMA: Aplicación.....	27
3.2 Análisis de estacionalidad.....	27
3.2.1 Análisis de estacionalidad: Resultados.....	27
3.3 Selección Mejor Modelo ARIMA.....	28
3.3.1 Método Grid Search.....	28
3.4 Modelos ARIMA: Extensiones y variantes.....	29
3.4.1 Modelo ARIMAX.....	29
3.4.2 Modelo ARIMAX: Aplicación.....	30
3.4.3 Modelos Arima Estacionales (SARIMA).....	30
3.4.4 Modelos SARIMA: Aplicación.....	31
3.4.5 Modelos SARIMAX.....	31
3.4.6 Modelos SARIMAX: Aplicación.....	31
3.5 Mejor Modelo MSE.....	32
3.6 Desempeño in sample vs. out of sample.....	32
4. Capítulo 4. Optimización de compra de insumos.....	41
4.1 Describiendo el problema de optimización.....	41

4.1.1 Función objetivo.....	41
4.1.2 Componentes de la función objetivo .....	41
4.1.3 Restricciones .....	41
4.2 Profundizando el modelo de optimización .....	42
4.2.1 Modelo Final.....	46
4.2.2 Presunciones .....	47
4.3 Escenarios.....	47
4.3.1 Recomendación.....	47
5. Capítulo 5. Conclusión.....	51
5.1 Recomendación final.....	51
5.2 Síntesis final: Conocimientos adquiridos, limitaciones y futuras mejoras.....	51
Bibliografía .....	53
Apéndices .....	54
Apéndice A .....	54
Apéndice B .....	55
Apéndice C.....	56
Apéndice D .....	56
Apéndice E.....	57

# 1. Capítulo 1. Introducción

## 1.1 Contexto

“M Surtidores” es una empresa familiar de tipo PYME que se dedica al servicio técnico integral de mantenimiento de surtidores de combustible líquido. Con un staff de 30 empleados conformado por técnicos mecánicos, electrónicos y personal administrativo, la empresa es liderada por la familia de G.M., un inmigrante italiano que llegó a Argentina a mitad de siglo pasado buscando una oportunidad de progreso en nuestro país. Con más de 70 años de historia en el rubro, la empresa ha pasado de generación en generación, y actualmente es gestionada por los dos hijos y tres nietos de G.M.

El principal cliente de “M Surtidores” es YPF, la empresa argentina líder en la distribución de combustible dentro del territorio argentino. Como proveedor de servicios, la empresa acude a las estaciones de servicio de YPF todos los días para atender cualquier problema técnico en sus equipos. Además, la empresa también brinda sus servicios a otras estaciones de servicio particulares, aunque en menor medida.

Dentro de los servicios brindados por la empresa se encuentra el mantenimiento, la instalación, reparación y venta de surtidores nuevos y reacondicionados. El staff de especialistas mecánicos y técnicos electrónicos está capacitado para resolver todo tipo de problemas que puedan llegar a presentarse en un surtidor, desde la reparación electrónica de un display, pasando por el cambio de “pico” y “manguera”, hasta la pintura que recubre el equipo. “M Surtidores” cuenta con certificación ISO para poder brindar estos servicios a todas las marcas y modelos de surtidores homologados a lo largo y ancho del territorio argentino. La empresa logra llegar a cada rincón del país a través de sus contratistas localizados en los puntos estratégicos de cada una de las 23 provincias argentinas.

Las instalaciones de “M Surtidores” se encuentran en el Gran Buenos Aires, allí conviven su depósito de surtidores, su depósito de insumos, su taller electrónico y mecánico junto con sus oficinas administrativas.

En un día habitual de trabajo, temprano por la mañana el equipo administrativo recibe distintas solicitudes de reparación o mantenimiento de las estaciones de servicio a las cuales atiende, esta solicitud la empresa la denomina “incidencia”. Luego de analizar rápidamente todas las incidencias obtenidas, el equipo administrativo las distribuye entre cada uno de los mecánicos de calle, es decir, aquellos mecánicos que se dedican solamente a la atención del surtidor en la estación de servicio asignada. Junto con la asignación de la incidencia, a cada mecánico se le brindan los insumos y herramientas necesarias para acudir a la asistencia de la estación de servicio. Entre los cinco insumos más utilizados por “M Surtidores” se encuentran los picos, los acoples, las mangueras y los filtros.



*Ilustración 1: InsCod 1 - Pico OPW 11 A*



*Ilustración 2: InsCod 22 - Acople Articulado OPW 3 / 4*



*Ilustración 1: InsCod 1057 - Manguera 5/8"*



*Ilustración 4: InsCod 1074 - Filtro 3/4*



*Ilustración 5: InsCod 1086 - Filtro 1"*

Luego de finalizado el trabajo, los mecánicos de calle vuelven a la empresa ese mismo día (o directamente el día posterior) y devuelven el detalle de la incidencia indicando qué tipo de trabajo realizaron y cuántos insumos utilizaron en el servicio brindado. El equipo administrativo alimenta la base de datos donde llevan registro de todas las incidencias atendidas con los comentarios del equipo mecánico de calle.

El proceso de atención de incidencias se repite diariamente e involucra uno de los servicios más importantes que posee la empresa que es la atención casi inmediata para la reparación y mantenimiento de los equipos de surtidores que poseen sus clientes. Como "M Surtidores" desea mejorar continuamente su nivel de servicio, la empresa está buscando profesionalizar la compra de sus insumos y tomar decisiones de negocio fundamentadas en datos concretos. En

este sentido, el presente trabajo final de maestría tiene como objetivo mejorar sus procesos de compra y gestión de insumos que son de suma importancia para la correcta gestión de las incidencias. Actualmente, todas las decisiones de negocio son tomadas en base a la experiencia de los gerentes, por lo que la empresa busca contar con información y datos concretos que le permitan tomar decisiones más informadas y acertadas.

## 1.2 Problema

“M Surtidores” cuenta con el registro detallado de las incidencias (mantenimientos correctivos y preventivos) que sus mecánicos de calle llevan a cabo diariamente. Actualmente, la empresa no ha estado utilizando esta información para tomar decisiones sobre sus compras de insumos, lo que podría tener un impacto significativo en su rendimiento financiero. Actualmente, los insumos que son utilizados para atender a las incidencias son comprados por la empresa a discreción de la gerencia y son utilizados como reserva de valor.

Esta práctica podría representar un costo de oportunidad para el capital de “M Surtidores” ya que podrían estar comprando más insumos de los verdaderamente demandados. Además, la empresa podría estar asignando de manera no eficiente su capital, que se vería erróneamente destinado a la compra de stock y no a la inversión de distintos activos financieros. Resulta importante que la empresa utilice el capital disponible de manera eficiente para otras inversiones que le permitan generar mayores beneficios.

## 1.3 Objetivo

Para solventar este problema de toma de decisión sin información, en la presente tesis se expone un modelo de machine learning que predice la cantidad y tipo de insumos que la empresa demandará en los próximos meses. Este modelo de machine learning se fundamenta en el análisis de los datos de incidencias pasadas para identificar patrones y tendencias en la demanda de insumos. El output del modelo es utilizado como componente de un modelo de optimización de tipo lot sizing que tiene como objetivo minimizar los costos que afronta la empresa al momento de abastecerse de insumos para atender su demanda. El modelo de optimización devuelve el mix de compra de insumos a lo largo del periodo julio 2023 – junio 2024 que la empresa debe llevar a cabo para cubrir su demanda al menor costo posible. En este sentido, el modelo de optimización informa qué porción de los beneficios mensuales de la empresa deben ser destinados a la compra mensual de insumos, dándole la oportunidad al negocio de visualizar el capital libre disponibilidad con el que contará luego de haber comprado los insumos necesarios para satisfacer la demanda. “M surtidores” puede así identificar el saldo disponible para ser destinado a la compra de activos financieros o a la inversión de otros proyectos. Con esta información, la empresa optimiza su proceso de compra de insumos, reduciendo los costos innecesarios y maximizando el rendimiento de su capital.

## 2. Capítulo 2. Datos

### 2.1 Descripción de las bases de datos

La empresa brindó una base de datos relacional compuesta por cinco tablas. Dicha base de datos es utilizada para llevar a cabo la etapa de exploración de información.

#### Tabla Incidencias

Esta tabla es la fuente principal del presente trabajo. La misma contiene todas las incidencias registradas por “M Surtidores” en el periodo febrero 2000 – noviembre 2022. Dichas incidencias ascienden a un total de 154361 registros. En cada registro se detalla a qué estación de servicio se acudió y qué tipo de trabajo se realizó. En la Tabla 1 se muestran los atributos más significativos de esta base de datos.

Tabla 1. Tabla Incidencias

Column	Description	Dtype	
OSNro	Código identificador de la incidencia.	int64	Clave Primaria
OSDet	Descripción del trabajo que se realizó en la estación de servicio a la cual se acudió.	object	
OSFec	Fecha en la cual se acude a resolver la incidencia.	datetime64[ns]	
OSIni	Hora en la que el técnico acudió a la estación de servicio.	object	
OSFin	Hora en la que el técnico terminó el servicio.	object	
OSPer	Cantidad de técnicos que acudieron a resolver la incidencia	int64	
ESCod	Código identificador de la estación de servicio donde ocurrió la incidencia.	object	Clave Foránea
ConCod	Código identificador del mecánico que acudió a resolver la incidencia.	int64	
CorrecPrev	Si la incidencia fue correctiva o preventiva.	object	

#### Tabla Incidencias\_insumos

Esta tabla está relacionada con la tabla “Incidencias” a través de la clave foránea “OSNro”. En cada registro de esta tabla se detalla qué insumo se utilizó en cada incidencia indicada en la tabla “Incidencias”. Si bien, todos los insumos de la tabla “Incidencias\_insumos” están relacionados con la tabla “Incidencias”, no todas las incidencias requieren de insumos, en algunos casos, la incidencia solo requiere que el mecánico de calle se presente en la estación de servicio para reparar un surtidor sin necesidad de reemplazar alguno de sus componentes. En la Tabla 2 detallan los atributos que componen a la tabla “Incidencias\_insumos”.

Tabla 2. Incidencias\_insumos

Column	Description	Dtype	
OSNro	Código identificador de la incidencia.	int64	Clave foránea
InsCod	Código del insumo que fue utilizado en la incidencia.	object	Clave foránea
Cantidad	Cantidad de insumo utilizado en la incidencia OSNro.	float64	
Precio	Precio de venta del insumo.	float64	

### Tabla Estaciones de Servicio

Esta tabla contiene el detalle de las estaciones de servicio que “M Surtidores” atiende. “ESCod” es el código identificador de la estación de servicio y actúa como clave primaria en esta tabla. En la Tabla 3 se visualizan los atributos de mayor interés.

Tabla 3. Estaciones de Servicio

Column	Description	Dtype
ESCod	Código identificador de la estación de servicio.	object
ESNom	Nombre de la estación de servicio.	object
ESDir	Dirección.	object
ESLoc	Localidad.	object
ESCantSurt	Cantidad de surtidores en la estación de servicio.	int64
ZonCod	Código identificador de la zona en la cual se encuentra la estación de servicio.	int64
EsProvincia	Provincia.	object

Clave  
Primaria

### Tabla Equipos

La tabla “Equipos” contiene un registro por cada equipo que se encuentra en cada una de las estaciones de servicio detalladas en la tabla “Estaciones de Servicio”. En la Tabla 4 se presentan los atributos más destacados en esta base.

Tabla 4. Equipos

Column	Description	Dtype
ESCod	Código identificador de la estación de servicio.	object
EquiCod	Código identificador del equipo que se encuentra en la estación de servicio.	object
EquiMar	Marca del equipo.	object
EquiMod	Modelo del equipo.	object
EquiSer1	Serie del equipo.	object

Clave  
foránea

### Tabla Insumos

Esta tabla contiene más información acerca de los insumos que vende y repara “M Surtidores”. “InsCod” es la clave primaria que relaciona esta tabla con las tablas “incidencias\_insumos” e “incidencias”. La Tabla 5 describe los principales atributos.

Tabla 5. Insumos

Column	Description	Dtype
InsCod	Código identificador del insumo.	object
InsDesCorta	Descripción del insumo.	object
InsPrecio	Precio del insumo.	float64
InsUM	Unidad de medida del insumo.	object
RubCod	Código de clasificación del insumo.	object

Clave  
primaria

### 2.1.1 Diagrama Entidad/Relación (DER)

A fines de visualizar con más precisión la relación entre las cinco tablas, se confeccionó un diagrama DER que puede ser consultado en el *apéndice A*.

### 2.2 Creación de la tabla “df\_incidencias”

Las tablas descritas en el punto anterior se combinaron a través de sus claves primarias y foráneas para poder dar paso al análisis exploratorio de las incidencias, sus insumos asociados y estaciones de servicio atendidas.

La base de datos “df\_incidencias” combina las tablas “Incidencias” e “Incidencias\_insumos” a través de la clave primaria “OSNro”. Al momento de combinar las tablas, se encontraron 72826 registros nulos en las columnas pertenecientes a la base de datos “Incidencias\_insumos”. Esta cifra representa el 54% de los registros de la tabla combinada. La explicación de este fenómeno está dada por el hecho de que no todas las incidencias requieren de insumos. Por este motivo se decidió reemplazar estas nulidades con el valor “0”. De esta manera, “InsCod” = 0 indica que la incidencia no necesitó de un insumo para ser atendida. Para poder identificar con más claridad esta característica del registro se confeccionó un atributo que indica TRUE si la incidencia lleva insumo e indica FALSE en caso contrario.

Tabla 6. Atributo “Requiere\_Insumos”

Column	Dtype
Requiere_Insumos	bool

Adicionalmente, para poder llevar a cabo una mejor visualización de los datos, se crearon las columnas detalladas en la Tabla 7. Estos atributos representan, el día, el mes y el año en que se asistió a la incidencia en cuestión.

Tabla 7. Atributos "Day"- "Month" - "Year"

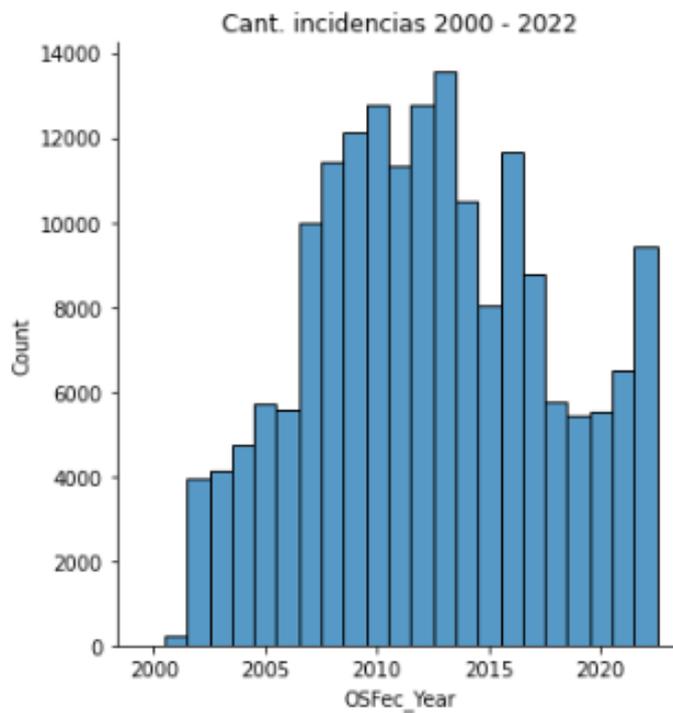
Column	Dtype
OSFec_Day	Int64
OSFec_Month	Int64
OSFec_Year	Int64

## 2.3 Análisis exploratorio de los datos

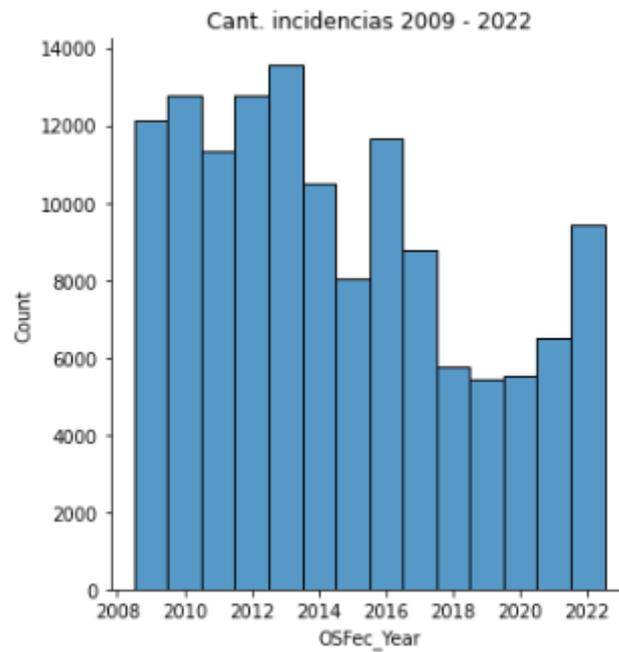
### 2.3.1 Exploración Macro

Al realizar la primera consulta sobre la tabla "df\_incidencias" se puede ver que en los primeros años (periodo 2000 - 2006) las incidencias son menores comparadas con los demás años. En este punto se decidió consultar con el negocio sobre dicha anomalía. "M surtidores" informó que los datos recolectados hasta el año 2008 inclusive no son del todo confiables, es por esto que se decidió descartar dichos registros.

Gráfico 1. Cantidad de incidencias en el periodo 2000 – 2022

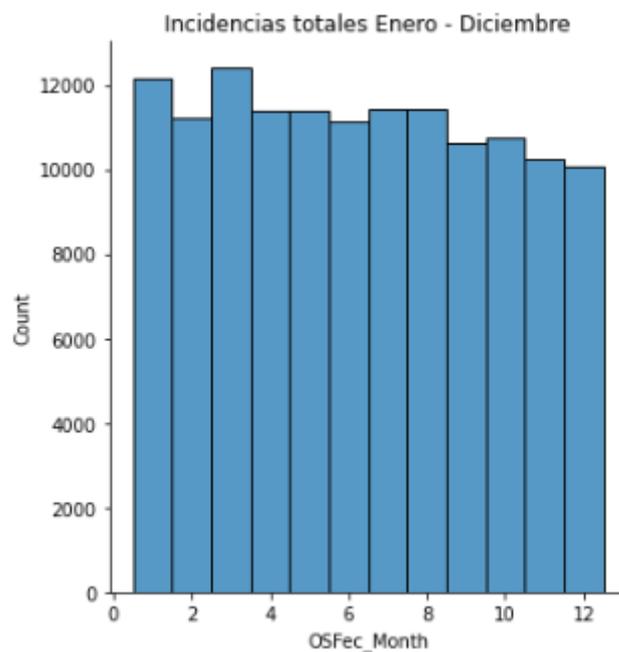


**Gráfico 2.** Cantidad de incidencias en el periodo 2009 - 2022



Con la tabla “df\_incidencias” comprendida en el periodo 2009-2022, se procedió a explorar el comportamiento de las incidencias a lo largo de los meses del año.

**Gráfico 3.** Cantidad de incidencias totales en el periodo Enero - Diciembre

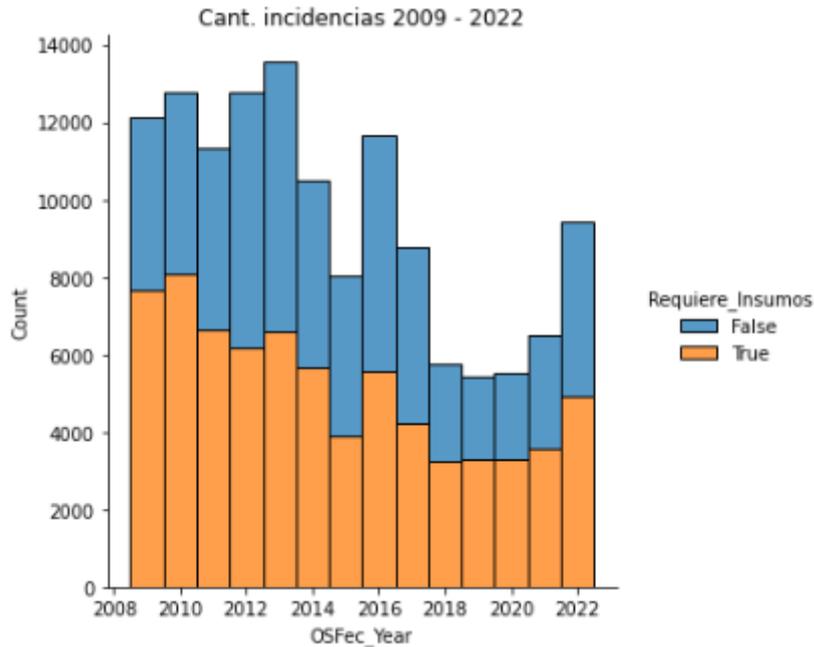


En el Gráfico 3 se puede apreciar una mayor cantidad de incidencias en los primeros 8 meses del año, identificando como meses más fuertes a enero y marzo.

### 2.3.2 Zoom in Insumos

Avanzando en la exploración de los datos, se buscó discriminar las incidencias que requerían insumos de las que no.

**Gráfico 4.** Cantidad de incidencias en el periodo 2009 – 2022 diferenciadas por el requerimiento de insumos



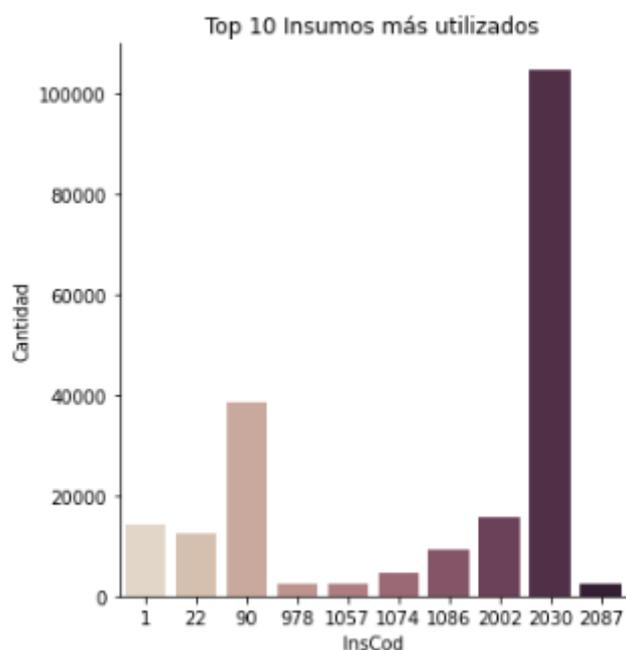
El atributo “Requiere\_Insumos” arroja que el 54% de los registros totales utilizó algún insumo. Por el contrario, el 45% de los datos no requirió de estos materiales.

En este punto, se decidió dejar de lado los registros que indicaban “FALSE” para continuar la exploración de los datos centrados en el comportamiento de las incidencias que requerían de insumos. De esta manera, se agruparon los datos de la tabla “df\_incidencias[“Requiere\_Insumos”] == TRUE” para confeccionar un ranking de insumos más utilizados por “M Surtidores”. Se agregó un atributo denominado “Cantidad\_%” para mostrar las cantidades empleadas de cada insumo en relación al total de insumos utilizados en toda la tabla. En la Tabla 8 se muestran los insumos más utilizados ordenados de mayor a menor.

**Tabla 8.** Top 10 insumos V1

InsCod	Cantidad	Cantidad_%	InsDesCorta
2030	104630	38%	PURGA SUMP RISER
90	38547	14%	BOMBA DE PURGAR
2002	15477	6%	PURGA SUMP DISPENSER
1	14144	5%	PICO OPW 11A
22	12563	5%	ACOPLE ARTICULADO OPW 3/4
1086	9047	3%	FILTRO 1"
1074	4610	2%	FILTRO 3/4
1057	2561	1%	MANGUERA 5/8" GOOD YEAR
978	2471	1%	MANGUERA DUNLOP 5/8 X 5MTS
2087	2279	1%	PURGA PRODUCTO

Gráfico 5. Top Ten Insumos más utilizados V1



De la Tabla 8 resulta interesante resaltar que los tres primeros insumos del ranking representan el 55% de los insumos utilizados y que cada uno de ellos hace referencia al término “Purgar” en su descripción. Con esta información, se decide consultar al negocio sobre el caso para obtener *insights* adicionales. “M Surtidores” indica que todos los códigos que en su descripción se refieren al término “Purgar” aluden al servicio de purga que se realiza a los tanques de combustibles de las estaciones que atienden. La unidad de medida de estos “Insumos” está expresada en litros y es por eso que presenta una distorsión grande en los datos. En este punto se decide quitar de la tabla “df\_incidencias” a aquellos “InsCod” que hacen referencia a este servicio. De esta manera se vuelve a identificar un nuevo ranking.

Tabla 9. Top Ten Insumos más utilizados V2

InsCod	Cantidad	Cantidad_ %	InsDesCorta
1	14144	12%	PICO OPW 11A
22	12563	11%	ACOPLE ARTICULADO OPW 3/4
1086	9047	8%	FILTRO 1"
1074	4610	4%	FILTRO 3/4
1057	2561	2%	MANGUERA 5/8" GOOD YEAR
978	2471	2%	MANGUERA DUNLOP 5/8 X 5MTS
85	2121	2%	ABRAZADERA DE MANGUERA
3	2015	2%	PICO OPW 11A/11B REPARADO
994	1781	2%	MANGUERA GOODYEAR 5/8 X 5MTS
38	1524	1%	VALVULA BREAKAWAY OPW 3/4

Al visualizar nuevamente el ranking, se puede notar que el código número tres que indica “PICO OPW 11A/11B REPARADO” parece estar refiriendo a un servicio. A su vez, se puede

presumir que tanto el código número 1057 como el código número 994 refieren al mismo insumo.

Con estas dos hipótesis se vuelve a consultar al negocio. “M surtidores” valida ambas hipótesis, haciendo que nuevamente se reduzcan los “InsCod” de la base de datos “df\_incidencias”.

Con estos ajustes en los datos se procede a realizar un nuevo ranking.

**Tabla 10.** Top Ten Insumos más utilizados V3

InsCod	Cantidad	Cantidad_%	InsDesCorta
1	14144	13%	PICO OPW 11A
22	12563	11%	ACOPLE ARTICULADO OPW 3/4
1086	9047	8%	FILTRO 1"
1074	4610	4%	FILTRO 3/4
1057	4342	4%	MANGUERA 5/8" GOOD YEAR
978	2471	2%	MANGUERA DUNLOP 5/8 X 5MTS
85	2121	2%	ABRAZADERA DE MANGUERA
38	1524	1%	VALVULA BREAKAWAY OPW 3/4
106	1458	1%	FILTRO UDAD. SELLADA 1"
69	1409	1%	CORREA A23

En el nuevo ranking se puede ver que el código 106 refiere al mismo insumo que el código 1086. Luego de confirmarlo con “M Surtidores” se lleva a cabo la unificación de ambos códigos. Finalmente, el top 10 incluyendo la unificación mencionada se visualiza en la Tabla 11.

**Tabla 11.** Top Ten Insumos más utilizados V4

InsCod	Cantidad	Cantidad_%	InsDesCorta
1	14144	13%	PICO OPW 11A
22	12563	11%	ACOPLE ARTICULADO OPW 3/4
1086	10505	9%	FILTRO 1"
1074	4610	4%	FILTRO 3/4
1057	4342	4%	MANGUERA 5/8" GOOD YEAR
978	2471	2%	MANGUERA DUNLOP 5/8 X 5MTS
85	2121	2%	ABRAZADERA DE MANGUERA
38	1524	1%	VALVULA BREAKAWAY OPW 3/4
69	1409	1%	CORREA A23
4	1388	1%	PUNTERA OPW 11A

La Tabla 11 evidencia que el 41% de los insumos se compone de solo cinco materiales.

## 2.4 Input final del negocio y elección de insumos a predecir

En este punto se decidió mostrar al negocio el ranking de insumos más utilizados para poder incorporar sus comentarios finales. “M Surtidores” indicó que tanto el código 1057 (MANGUERA 5/8" GOOD YEAR) como el código 978 (MANGUERA DUNLOP 5/8 X 5MTS) son utilizados para atender el mismo tipo de incidencia y que la diferencia entre ambos insumos radica en la marca del repuesto a utilizar. De esta manera se procedió a unificar ambos códigos en el “InsCod” 1057.

*Tabla 12. Top Ten Insumos más utilizados versión definitiva*

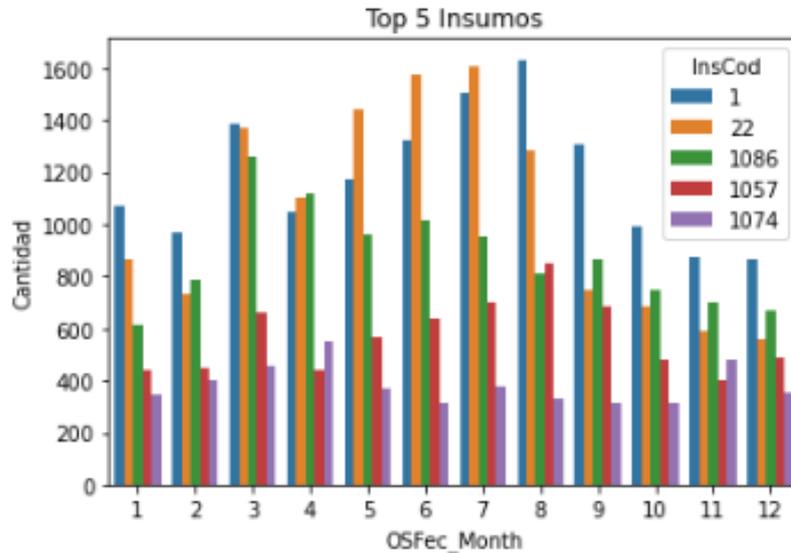
<b>InsCod</b>	<b>Cantidad</b>	<b>Cantidad_%</b>	<b>InsDesCorta</b>
1	14144	13%	PICO OPW 11A
22	12563	11%	ACOPLE ARTICULADO OPW 3/4
1086	10505	9%	FILTRO 1"
1057	6813	6%	MANGUERA 5/8" GOOD YEAR
1074	4610	4%	FILTRO 3/4
85	2121	2%	ABRAZADERA DE MANGUERA
38	1524	1%	VALVULA BREAKAWAY OPW ¾
69	1409	1%	CORREA A23
4	1388	1%	PUNTERA OPW 11A
36	1386	1%	BREAK A WAY (VALVULA 3/4")

Finalmente, a partir del input brindado por el negocio y por los datos obtenidos en el primer análisis, se decide seguir trabajando solo con los cinco insumos más utilizados por “M Surtidores”. Estos cinco insumos son los que se predicen y posteriormente se optimizan en el presente trabajo. A partir de este momento nos referiremos a estos cinco insumos como “consumibles”.

### 2.4.1 Zoom in Insumos Consumibles

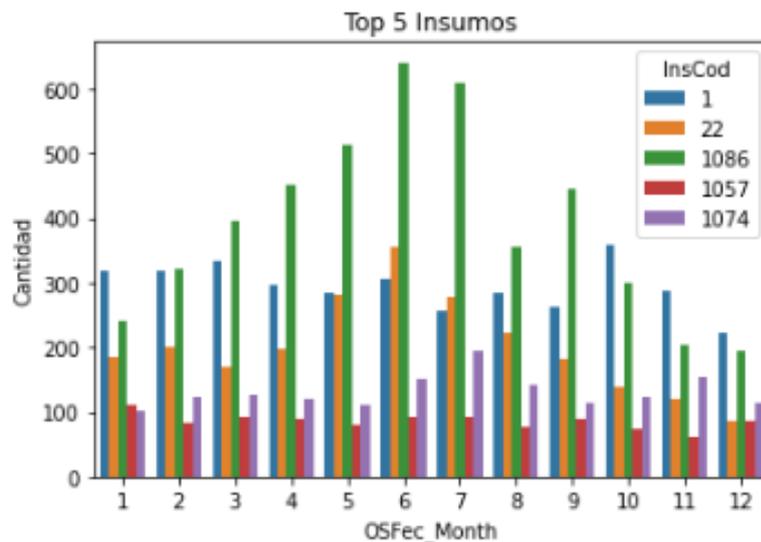
Con los insumos más utilizados siendo identificados en el punto anterior, se continuó con un análisis de comportamiento para los primeros cinco materiales del ranking.

Gráfico 6. Top Five Insumos más utilizados. Distribución anual



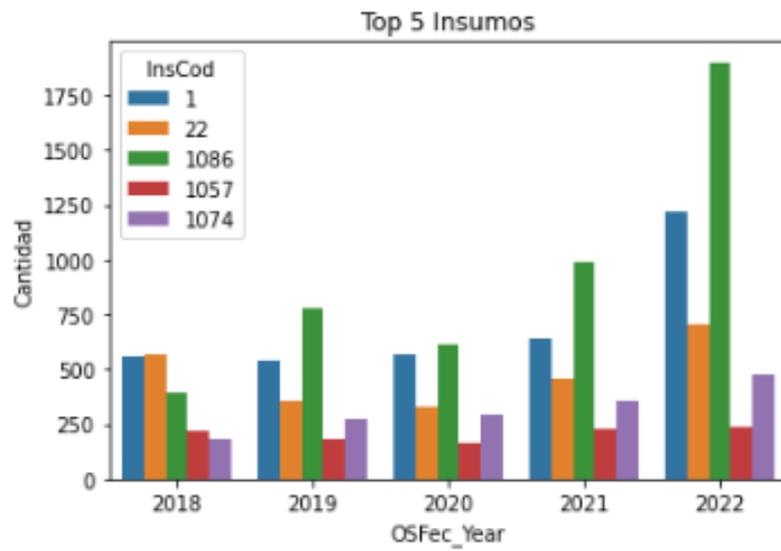
Del Gráfico 6 resulta relevante observar la distribución de la utilización de insumos a lo largo del año. Los tres insumos más requeridos, correspondientes al pico, al acople y al filtro, presentan una distribución similar entre sí y parecida en su forma a una distribución binomial. Con este resultado, se decide hacer foco en los últimos cinco años.

Gráfico 7. Top Five Insumos más utilizados. Distribución anual periodo 2018-2022



En el Gráfico 7 se observa con mayor claridad la distribución de la demanda de los insumos a lo largo del año. Los filtros 1" (1086) y los acoples (22) presentan una demanda más alta en los meses de mayo, junio y julio, mientras que los picos (1) presentan cierta estacionalidad con pequeñas variaciones a lo largo de los meses.

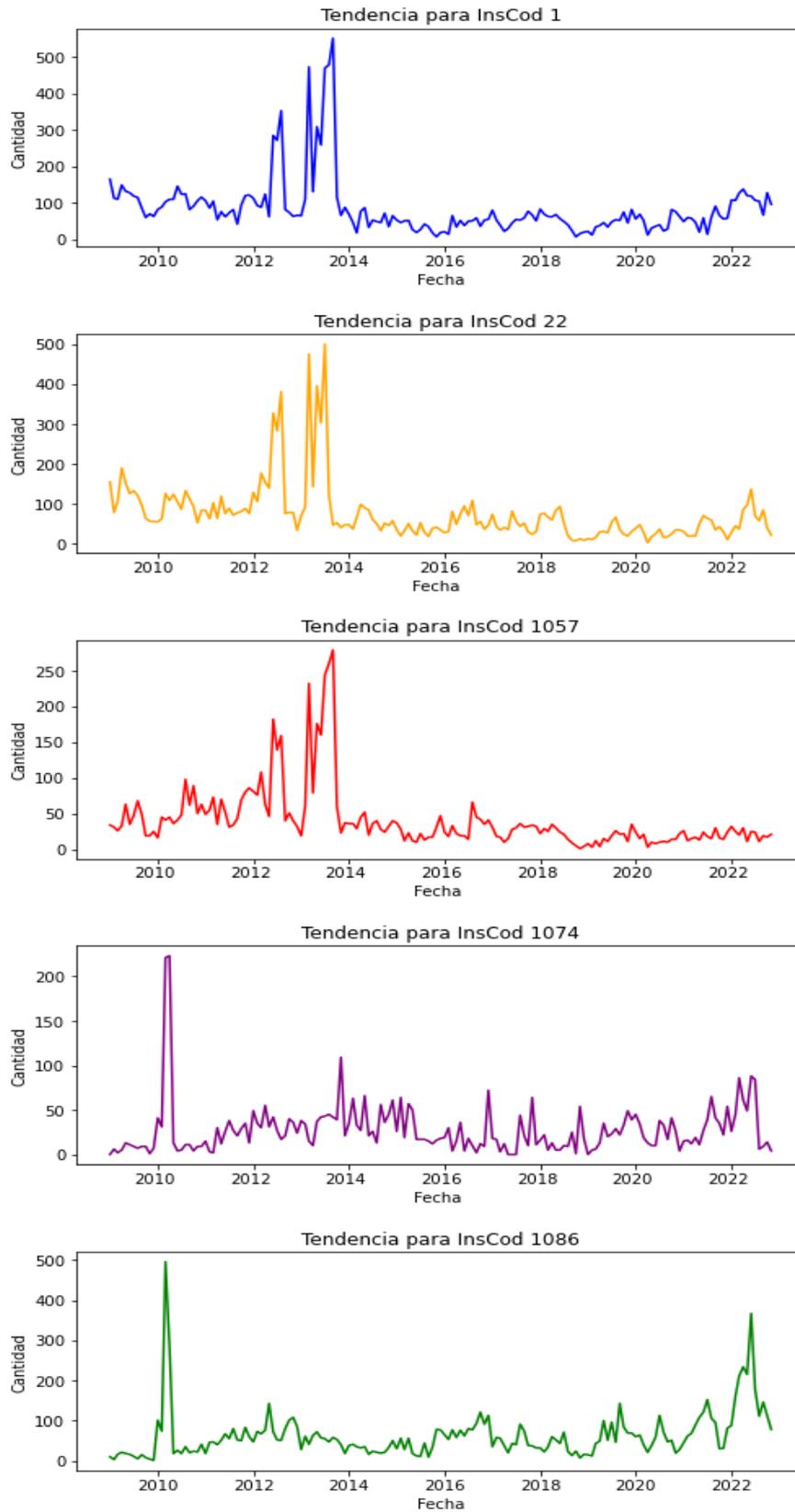
**Gráfico 8.** Top Five Insumos más utilizados. Distribución por año periodo 2018-2022



Del Gráfico 8 se puede observar el crecimiento en los últimos cinco años de los filtros (1086 y 1074) y picos (1) mientras que la manguera (1057) muestra cantidades estables para el mismo periodo.

Consumibles: Tendencia

Gráfico 9. Tendencia de los consumibles para el periodo 2009-2022



En el Gráfico 9 se puede observar a simple vista, dos grandes grupos donde los “InsCod” 1, 22 y 1086 muestran tendencias similares, mientras que los “InsCod” 1074 y 1086 también presentan patrones similares entre sí.

*Consumibles: Correlación*

Luego de haber revisado las tendencias de cada uno de los insumos se procedió a confeccionar una matriz de correlación para comprobar estadísticamente si se presentaba una relación entre los insumos seleccionados.

La Matriz de Correlación mide el grado de relación lineal entre cada par de variables (consumibles). Un valor de 1 muestra una correlación positiva perfecta, indicando que las variables aumentan al mismo tiempo. Por otro lado, un valor de -1 indica una correlación negativa perfecta, comprobando que las variables disminuyen en su tendencia al mismo tiempo. Finalmente, un valor de 0 indica una ausencia de correlación entre el par de variables. Se debe tener en cuenta que la correlación no necesariamente implica causalidad.

*Tabla 13. Matriz de Correlación*

Matriz de Correlación					
	1	22	1057	1074	1086
1	1,00	0,75	0,92	0,11	0,06
22	0,75	1,00	0,76	0,06	0,06
1057	0,92	0,76	1,00	0,07	-0,03
1074	0,11	0,06	0,07	1,00	0,68
1086	0,06	0,06	-0,03	0,68	1,00

La matriz de correlación detallada en la Tabla 13 muestra que hay una fuerte correlación positiva (cerca de 1) entre las variables 1 y 1057, y también entre las variables 22 y 1057. Además, se puede visualizar una correlación moderada positiva (alrededor de 0,75) entre las variables 1 y 22, y también entre las variables 1 y 1057.

Por otro lado, se puede identificar una correlación positiva entre las variables 1074 y 1086 mientras que sus correlaciones con las otras variables son muy bajas, sugiriendo que las mismas no están relacionadas con las otras variables en el conjunto de datos.

El análisis de las correlaciones entre los pares de variables mejora el análisis presentado en el Gráfico 9. Podemos afirmar que tanto los picos (1) como las mangueras (1057) se demandan casi de igual forma a lo largo del tiempo. A este par de variables se puede adicionar el consumible acople (22) que presenta una correlación fuerte y similar a ambos consumibles. En otro grupo encontramos el filtro de una pulgada y el filtro de 3/4 que correlacionan positivamente.

## 2.4.2 Zoom in Estaciones de Servicio

“M Surtidores” tiene 2669 estaciones de servicio dadas de alta distribuidas en las 23 provincias que componen el territorio argentino, sin embargo, en el periodo enero 2009 – noviembre 2022 solo fueron atendidas 720. El 92% de dichas estaciones pertenecen a la empresa YPF mientras que el resto corresponden a estaciones particulares. El negocio utiliza la abreviatura “EESS” para referirse a las estaciones de servicio, en el presente trabajo también se hace uso de dicha abreviatura.

YPF es el principal cliente de “M Surtidores” y es por este motivo que la mayoría de las EESS que presentan incidencias corresponden a esa empresa. Además, debido al contrato negociado con YPF, “M Surtidores” debe brindar tanto un servicio preventivo como correctivo haciendo que las incidencias aumenten si se trata de una EESS de YPF.

Luego de haber explorado la tabla “df\_incidencias” se procedió a identificar cuáles eran las estaciones de servicio con mayor cantidad de insumos solicitados durante el periodo comprendido entre 2009 y 2022.

**Tabla 14.** Ranking EESS

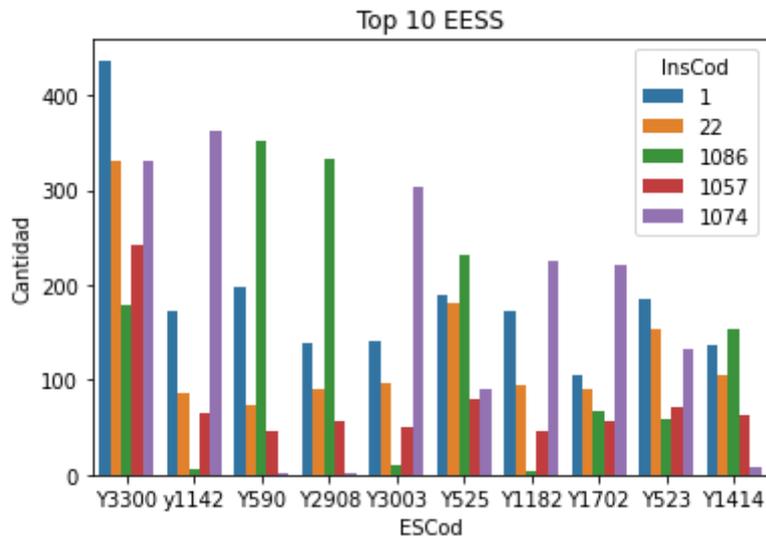
ESCod	Cantidad Incidencias (% total EESS)	ESNom	ESLoc	ESCantSurt
Y3300	3%	OPESSA (DOLORES)	DOLORES	64
Y590	2%	OPESSA (CAÑUELAS)	CAÑUELAS	20
Y525	2%	OPESSA (BERISSO)	BERISSO	28
y1142	1%	OPESSA (CAMPANA ESTE)	PACHECO	34
Y523	1%	OPESSA (BERISSO PTE ROMA)	BERISSO	24
Y2908	1%	OPESSA (SALADILLO)	SALADILLO	30
Y3003	1%	OPESSA (AUSOL PILAR SUR)	TORTUGUITAS	34
Y2904	1%	OPESSA (BR ZARATE) (ENC)	ZARATE	32
Y2991	1%	OP (ACCESO OESTE LUJAN)	GRAL RODRIGUEZ	28
y3253	1%	OP (CAMPANA AUSOL ASC)	CAMPANA	28
Y1702	1%	OPESSA (LAS FLORES)	LAS FLORES	29
Y504	1%	OPESSA (LA PLATA)	LA PLATA	24
Y1182	1%	OPESSA(AUSOL PILAR NORTE)(ENC)	TORTUGUITAS	34
Y421	1%	OPESSA (ACA ZARATE)	ZARATE	42
Y335	1%	OPESSA (MORENO)	MORENO	32
Y3329	1%	OPESSA (CAMPANA OESTE)	PACHECO	34
Y1414	1%	OPESSA (BELLA VISTA)	BELLA VISTA	24
Y3150	1%	OPESSA (FCIO.VARELA)	FCIO.VARELA	34
y2916	1%	OPESSA (JUMBO PILAR) (ENC)	PILAR	40
Y2992	1%	OPESSA (ACCESO OESTE C)	GRAL RODRIGUEZ	24

Como se puede apreciar en la Tabla 14, las 20 EESS con mayor demanda de insumos corresponden a la empresa YPF y todas se encuentran en la provincia de Buenos Aires. Las tres

primeras contienen el 7% de las incidencias totales mientras que el resto de las EESS tienen un 1% del total general.

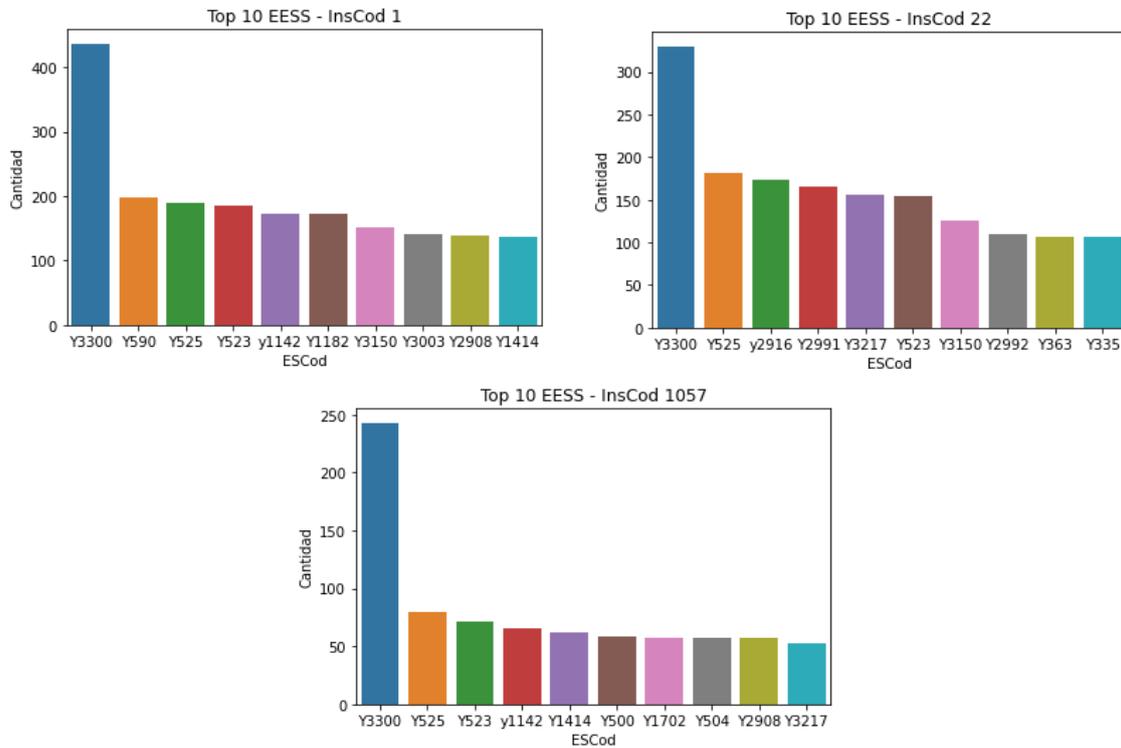
El Gráfico 10 se centra exclusivamente en los consumibles, mostrando las 10 estaciones de servicio que registraron la mayor demanda de este tipo de insumos durante el periodo 2009 - 2022.

**Gráfico 10.** Top Five Insumos más utilizados. Principales EESS



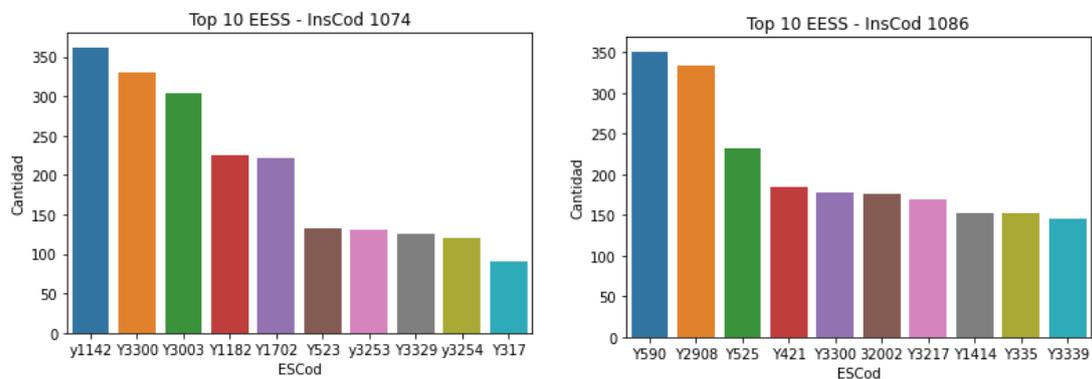
Profundizando el análisis sobre las EESS, se decidió confeccionar un ranking de estaciones de servicio para cada uno de los consumibles.

**Gráfico 11.** Top 10 EESS para los consumibles 1,22 y 1057



En los primeros tres gráficos se puede ver que la EESS Y3300 ubicada en Dolores, Buenos Aires concentra la mayor demanda de picos, mangueras y acoples.

**Gráfico 12.** Top 10 EESS para los consumibles 1074 y 1086



A diferencia de lo que se ha observado previamente en el Gráfico 11 relacionado con los consumibles 1, 22 y 1057, se ha constatado que durante el periodo comprendido entre 2009 y 2022, la EESS Y3300 no ha registrado la mayor demanda de filtros de 1" y 3/4. Al analizar el Gráfico 12, que se refiere al "InsCod" 1074, se visualiza que la estación Pacheco ha sido la que ha consumido la mayor cantidad de filtros 3/4, mientras que la EESS de Dolores se ha ubicado en segundo lugar. Si se considera únicamente dentro del Gráfico 12 el correspondiente al "InsCod" 1086, se puede afirmar que la EESS de Cañuelas ha solicitado la mayor cantidad de filtros de 1", seguida, en segundo lugar, por la EESS de Saladillo. Es importante destacar que la EESS de Dolores se ha posicionado en el quinto lugar del ranking de consumo de este insumo, ubicándose detrás de la estación de Zarate.

## 2.5 Análisis Exploratorio de Datos – Conclusión final

Finalizado el análisis realizado en la primera parte del presente trabajo se concluye que los insumos a predecir serán los detallados debajo:

- InsCod 1: Pico OPW 11 A
- InsCod 22: Acople Articulado OPW 3 / 4
- InsCod 1086: Filtro 1"
- InsCod 1057: Manguera 5/8"
- InsCod 1074: Filtro 3 / 4

Se ha comprobado que estos cinco insumos ("Consumibles") son los más demandados por las EESS que "M Surtidores" atiende. En consecuencia, el modelo de machine learning se enfocará en predecir la demanda del periodo julio 2023 – junio 2024 para cada uno de ellos.

Se presume que un Modelo Arima podría ser adecuado para realizar la predicción. En la segunda parte del trabajo, se realiza un análisis estadístico para verificar la presencia de estacionalidad en los datos y determinar si dicho modelo es adecuado.

## 3. Capítulo 3. Forecasting

En esta parte del trabajo se procede a crear distintas variantes de los modelos de machine learning de tipo ARIMA para predecir la cantidad de consumibles que el negocio debería utilizar en el periodo julio 2023 – junio 2024. Los datos fueron divididos en datos de entrenamiento para darlos como input a los modelos que luego fueron evaluados con la otra parte de los datos (los datos de evaluación). Para dicha evaluación se tuvo en cuenta la métrica de performance MSE y AIC.

### 3.1 Modelo ARIMA

#### 3.1.1 Modelo ARIMA: Introducción

El modelo ARIMA (Autoregressive Integrated Moving Average) es un modelo estadístico univariado<sup>1</sup> utilizado para analizar series temporales (Box and Jenkins, 1970). Este modelo es de tipo dinámico ya que busca hacer predicciones con los mismos datos de la serie en cuestión pero del pasado, busca entender si los datos actuales de la serie se explican con sus datos predecesores.

El modelo ARIMA se compone de tres partes: la parte autoregresiva (AR), la parte integrada (I) y la parte de media móvil (MA). El parámetro AR describe cómo los valores previos de la serie temporal afectan a los valores actuales; la parte I describe cómo la serie temporal debe ser diferenciada para hacerla estacionaria<sup>2</sup>; y la parte MA describe cómo los errores de la serie temporal afectan a los valores actuales.

El modelo ARIMA se ajusta a los datos mediante el uso de técnicas de estimación de máxima verosimilitud, y generalmente se utiliza para predecir valores futuros en series temporales.

Usualmente, estos modelos se suelen expresar como ARIMA(p,d,q) donde los parámetros p, d y q son números enteros no negativos que indican el orden de las distintas partes del modelo. Respectivamente estos parámetros refieren a la parte “AR”, la parte “I” y la parte “MA”.

#### 3.1.2 Modelo ARIMA: Ejemplos

Para poder entender un poco más sobre estos modelos se procede a plantear tres ejemplos distintos con diferentes valores en sus parámetros p, d y q.

##### ARIMA (1,0,0)

Un modelo ARIMA(1,0,0)<sup>3</sup> será un modelo que solo “va hacia atrás” 1 periodo ya que p=1 indica que se utiliza un proceso autorregresivo de orden 1, d=0 indica que no se requiere una diferencia para hacer estacionaria la serie de tiempo y q=0 indica que no se utiliza un proceso de media móvil. Este modelo entonces muestra que la predicción de la serie de tiempo en un momento t depende únicamente del valor de la serie en el momento t-1. Matemáticamente se puede escribir como:

$$x(t) = \alpha + \phi * x(t-1) + w(t)$$

- x(t) es el valor de la serie en el momento t,
- $\alpha$  es una constante,

---

<sup>1</sup> Los métodos univariados se enfocan en una sola variable a lo largo del tiempo.

<sup>2</sup> Las series de tiempo son estacionarias cuando su media y varianza son constantes a lo largo del tiempo.

<sup>3</sup> Abreviado como AR(1)

- $\phi$  es el coeficiente autorregresivo,
- $w(t)$  es un término de error aleatorio

Matemáticamente se puede ver que el valor actual de la serie ( $x_t$ ) se puede explicar como función de los “p” valores pasados:  $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ , donde p determina el número de pasos hacia atrás que se debe dar para poder predecir el valor actual ( $x_t$ ).

#### ARIMA (1,1,0)

Un modelo ARIMA(1,1,0) es un modelo que además de utilizar un parámetro autorregresivo de orden 1 (AR(1)) y de no utilizar un proceso de media móvil (MA) ya que  $q=0$ , como vimos en el ejemplo anterior, utiliza un proceso de diferencia de orden 1 para hacer estacionaria la serie de tiempo.

En este modelo, la serie de tiempo original se diferencia una vez para hacerla estacionaria. Esto significa que se calcula la diferencia entre cada valor de la serie de tiempo y su valor anterior. Posteriormente, se utiliza un proceso autorregresivo de orden 1 en la serie de tiempo diferenciada para modelar la relación entre las observaciones consecutivas. Matemáticamente se puede representar como:

$$\Delta x(t) = \alpha + \phi * (\Delta x(t-1)) + w(t)$$

- $\Delta x(t) = x(t) - x(t-1)$  es la serie de tiempo diferenciada,
- $\alpha$  es una constante,
- $\phi$  es el coeficiente autorregresivo,
- $w(t)$  es un término de error aleatorio.

El modelo predice el cambio en el valor de la serie de tiempo en un momento t utilizando el cambio en el valor de la serie en el momento anterior t-1.

#### ARIMA (1,1,1)

Un modelo ARIMA(1,1,1), además de utilizar un proceso autorregresivo de orden 1 (AR(1)) y un proceso de diferencia de orden 1 para hacer estacionaria la serie de tiempo, utiliza un proceso de media móvil de orden 1 (MA(1)).

En este ejemplo, al igual que en el caso anterior, la serie de tiempo original se diferencia una vez para hacerla estacionaria y luego se utiliza un proceso autorregresivo de orden 1 en la serie diferenciada para modelar la relación entre las observaciones consecutivas. La diferencia con ARIMA (1,1,0) es que en este caso se incluye un proceso de media móvil de orden 1 para capturar cualquier patrón en los errores de predicción. Este modelo utiliza tanto el pasado de la serie de tiempo diferenciada como los errores de predicción anteriores para predecir el cambio en el valor de la serie de tiempo en el siguiente período.

Matemáticamente, se puede representar como:

$$\Delta x(t) = c + \phi * (\Delta x(t-1)) + \theta * (w(t-1)) + w(t)$$

- $\Delta y(t) = y(t) - y(t-1)$  es la serie de tiempo diferenciada,
- $c$  es una constante,
- $\phi$  es el coeficiente autorregresivo,
- $\theta$  es el coeficiente de media móvil,
- $w(t-1)$  es el error en el periodo anterior,
- $w(t)$  es el error en el periodo actual.

### 3.1.3 Modelos ARIMA: Aplicación

A continuación se procede a utilizar el lenguaje de programación Python para construir un modelo ARIMA teniendo en cuenta las demandas mensuales de los consumibles en el periodo enero 2009 – noviembre 2022 para cada uno de los cinco insumos seleccionados anteriormente. Para ello se utiliza la clase “statsmodels.tsa.arima.model.ARIMA” de la librería “statsmodels”<sup>4</sup>, un módulo de Python especializado en la confección y análisis de modelos estadísticos.

Antes de avanzar con el modelo se procede a analizar si cada una de las series temporales es estacionaria o si se debe realizar una transformación previa para convertirlas. Posteriormente se lleva a cabo la búsqueda de la mejor combinación de parámetros p, d y q que minimicen el error cuadrático medio (MSE) para seleccionar el mejor modelo ARIMA.

### 3.2 Análisis de estacionalidad

Se procedió a realizar un análisis de la estacionalidad de los datos para cada consumible utilizando la prueba de Dickey-Fuller aumentada, ADF por sus siglas en ingles. Dicho test es una prueba estadística que busca determinar si una serie de tiempo tiene una raíz unitaria lo que significa que la serie no es estacionaria y puede ser afectada por tendencias a largo plazo.

De esta manera, el test plantea la siguiente hipótesis nula y alternativa:

Hipótesis nula (H0): La serie de tiempo tiene una raíz unitaria, lo que significa que no es estacionaria y puede ser afectada por tendencias a largo plazo.

Hipótesis alternativa (Ha): La serie de tiempo no tiene una raíz unitaria, lo que significa que es estacionaria y no es afectada por tendencias a largo plazo.

Para rechazar la hipótesis nula el p valor brindado por el test (la probabilidad de obtener un resultado igual o más extremo que el observado si la hipótesis nula fuera verdadera) debe ser menor que el nivel de significancia elegido (el riesgo de cometer un error de tipo I al rechazar la hipótesis nula cuando en realidad era verdadera). El nivel de significancia puede establecerse en 1%, 5% o 10%. En este caso particular, para analizar los resultados obtenidos en el test utilizamos un nivel de significancia del 5%.

Para realizar el test se utilizó la herramienta “adfuller” disponible en la librería “statsmodels”.

#### 3.2.1 Análisis de estacionalidad: Resultados

Los resultados arrojados por el test para cada uno de los consumibles se visualizan en la Tabla 15.

Tabla 15. Test ADF

	InsCod				
	1	22	1057	1074	1086
ADF Statistic	-3,632	-1,798	-1,599	-5,998	-6,392
P Valor	0,005	0,382	0,484	0,000	0,000
Valor Crítico 1%	-3,471	-3,473	-3,473	-3,471	-3,470
Valor Crítico 5%	-2,879	-2,880	-2,880	-2,879	-2,879
Valor Crítico 10%	-2,576	-2,577	-2,577	-2,576	-2,576

<sup>4</sup> <https://www.statsmodels.org/>

Los consumibles 1, 1074 y 1086 presentaron un p valor menor que el nivel de significancia establecido del 5%. Incluso, en los tres casos, el p valor es menor que 1%. De esta manera se rechaza la H0 pudiendo concluir que los datos tienen una raíz unitaria y que por ende son estacionarios. Esta conclusión también se puede obtener comparando el estadístico ADF con los valores críticos. Los valores críticos son los valores límite para las estadísticas de prueba para diferentes niveles de significancia. Si el estadístico de prueba es menor que el valor crítico correspondiente, se puede rechazar la hipótesis nula de que la serie de tiempo tiene una raíz unitaria. En los tres casos el estadístico ADF obtenido es menor al valor crítico correspondiente para el nivel de significancia 1%.

Por otro lado, podemos ver que los consumibles 22 y 1057 tienen un p valor mayor al nivel de significancia establecido del 5%. Incluso ambos consumibles tienen un p valor mayor al 10%. De esta manera, no podemos rechazar la hipótesis nula de que la serie tiene raíces unitarias y, por lo tanto, no es estacionaria. Esta conclusión también se puede ver en la comparación del estadístico ADF que es mayor a los valores críticos correspondientes a 1%, 5% y 10%.

El test ADF nos indica que cada una de las series de tiempo de los consumibles 1, 1074 y 1086 es estacionaria y no requiere ser transformada para ser procesada por el modelo ARIMA. Por el contrario, las series de tiempo de los “InsCod” 22 y 1057 no son estacionarias, por ende, deben ser diferenciadas para que el modelo ARIMA pueda procesarlos y realizar una predicción. De esta manera, se aplica una diferenciación de un periodo tanto para la serie del consumible 22 como para la del consumible 1057 y se vuelve a repetir el test ADF.

**Tabla 16.** Test ADF V2

	InsCod	
	22	1057
ADF Statistic	-8,172	-6,869
P Valor	0,000	0,000
Valor Crítico 1%	-3,473	-3,473
Valor Crítico 5%	-2,880	-2,880
Valor Crítico 10%	-2,577	-2,577

Se puede ver que en ambos casos el p valor es menor que el nivel de significancia establecido del 5% (incluso es menor que 1%). De esta manera se avanza con el rechazo de la hipótesis nula pudiendo concluir que con la diferenciación de un periodo las series de tiempo de ambos consumibles se vuelven estacionarias.

Para continuar con la selección del mejor modelo a utilizar, las series de tiempo del “InsCod” 22 y 1057 se diferencian un periodo antes de ser procesadas por el modelo.

### 3.3 Selección Mejor Modelo ARIMA

Para poder determinar el mejor modelo ARIMA para cada uno de los consumibles, se procedió a realizar una búsqueda de los parámetros p, d y q a través del método Grid Search. El método seleccionará al mejor modelo en base a la performance de la métrica MSE.

#### 3.3.1 Método Grid Search

En primer lugar, cada una de las series temporales fue dividida en “datos de entrenamiento” y “datos de evaluación”. El primer conjunto de datos abarcaba la demanda mensual desde enero

2009 hasta diciembre 2020 mientras que los datos de test se componían de las demandas mensuales del periodo enero 2021 - noviembre 2022. Para realizar dicha división se utilizó la función “iloc” de la librería “Pandas”<sup>5</sup>, la cual se especializa en brindar herramientas para facilitar la manipulación de los data frames.

En segundo lugar, se establecieron los rangos de valores que podían tomar los parámetros p, d y q. Para p y q se generaron rangos de cero a tres en ambos casos, mientras que para d se estableció un rango de cero a dos. Se utilizó la función “itertools”<sup>6</sup> de Python para crear 18 posibles combinaciones de parámetros a probar.

Una vez separado los datos y establecidos los rangos de parámetros, el código itera en cada posible combinación ajustando el modelo sobre los datos de entrenamiento para luego hacer las predicciones para el periodo de los datos de test y, posteriormente, comparar lo predicho con los datos reales de ese conjunto.

El código realiza esta evaluación utilizando la métrica de performance MSE (error cuadrático medio), que calcula la diferencia al cuadrado entre cada valor observado (valor real de los datos de test) y su correspondiente valor predicho para tomar el promedio de esos valores cuadrados. De esta manera, el método Grid Search logra indicar cuál es el conjunto de parámetros que indica un menor MSE.

Luego de aplicar este método para cada una de las series temporales, se identificaron los mejores parámetros con su correspondiente MSE.

Tabla 17. Modelo Arima

ARIMA Model				
	p	d	q	MSE
1	2	0	1	1034,82
22	2	1	0	908,80
1057	2	0	1	37,87
1074	2	1	2	755,86
1086	2	0	1	10756,51

### 3.4 Modelos ARIMA: Extensiones y variantes

En este punto donde ya contamos con un mejor modelo “ARIMA base” para cada consumible, procederemos a profundizar en distintas variantes del modelo para identificar cual será el mejor modelo a utilizar en cada caso y así emplearlo en la realización de la predicción de demanda requerida.

#### 3.4.1 Modelo ARIMAX

El Modelo ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) es una extensión del modelo ARIMA en el que se incluyen variables exógenas para ayudar a explicar la variabilidad en la variable de interés. De esta manera, un modelo ARIMAX(1,0,0) se puede escribir como:

<sup>5</sup> [https://pandas.pydata.org/docs/user\\_guide/index.html#user-guide](https://pandas.pydata.org/docs/user_guide/index.html#user-guide)

<sup>6</sup> <https://docs.python.org/3/library/itertools.html>

$$x(t) = \alpha + \phi x(t-1) + \beta z(t) + w(t)$$

- $x(t)$  es el valor de la serie en el momento  $t$ ,
- $\alpha$  es una constante,
- $\phi$  es el coeficiente autorregresivo,
- $\beta$  es el coeficiente de la variable exógena  $z(t)$
- $w(t)$  es un término de error aleatorio.

### 3.4.2 Modelo ARIMAX: Aplicación

A cada uno de los “modelos ARIMA base” se les agregaron distintas variables exógenas para buscar aquel modelo ARIMAX que minimice el MSE. Nuevamente se llevó a cabo la búsqueda del mejor modelo a través del método Grid Search.

Las variables exógenas tenidas en cuenta para cada uno de los modelos fueron:

- Días laborables en el mes.
- Precio del USD oficial en ARS para cada mes.
- EESS atendidas: Esta variable comprende la cantidad de EESS que “M Surtidores” atiende, independientemente si estas necesitan de insumos o de servicios de reparación de surtidor o purga.

En el *apéndice B* se muestran los valores de las tres variables para el periodo 2009-2022.

Dichas variables se agregaron a cada uno de los modelos “ARIMA base” para repetir el proceso de búsqueda de parámetros que minimicen el MSE. Los modelos obtenidos se pueden visualizar en la Tabla 18.

*Tabla 18. Modelo Arimax*

ARIMAX Model				
	p	d	q	MSE
1	2	1	2	445,21
22	0	0	0	721,94
1057	1	1	2	217,67
1074	0	0	2	641,43
1086	0	0	0	5429,05

### 3.4.3 Modelos Arima Estacionales (SARIMA)

Un modelo SARIMA (Seasonal Autoregressive Integrated Moving Average) es una extensión del modelo ARIMA que, además de considerar los componentes AR, MA e I, tiene en cuenta la estacionalidad en los datos de la serie de tiempo.

Este modelo se expresa como SARIMA ( $p, d, q$ )  $\times$  ( $P, D, S$ ):

- P: es el orden del término de autorregresión estacional.
- D: es el orden de diferenciación estacional.
- Q: es el orden del término de media móvil estacional.
- s: es el número de periodos de una temporada estacional.

### 3.4.4 Modelos SARIMA: Aplicación

Para ajustar y estimar el modelo SARIMA de cada consumible se utilizó la función “statsmodels.tsa.statespace.SARIMAX” de la librería “statsmodels”.

Para identificar el mejor modelo en cada caso se procedió a replicar el método Grid Search creado anteriormente pero, en este caso, se adicionaron parámetros P, D, Q y s correspondientes al orden estacional. Se estableció el parámetro “s” en el valor 12 ya que los valores de cada serie son mensuales. Para el caso de P, D y Q se crearon rangos de cero a tres, de cero a dos y de cero a tres respectivamente. Se obtuvo así una combinación que 324 parámetros para ser aplicados al modelo. En esta oportunidad, para la serie temporal de “InsCod” 22 e “InsCod” 1057, los datos no fueron diferenciados previamente, sino que, se limitó al parámetro d a tomar un rango entre uno y dos.

Lo mejores modelos SARIMA, teniendo en cuenta el menor MSE, se detallan en la Tabla 19.

*Tabla 19. Modelo Sarima*

SARIMA Model							
	p	d	q	P	D	Q	MSE
1	0	1	2	2	0	2	1072,25
22	0	2	2	1	0	2	696,97
1057	2	1	2	1	0	0	37,42
1074	1	1	1	0	0	0	755,21
1086	1	2	0	0	0	1	10310,27

### 3.4.5 Modelos SARIMAX

Al igual que en el modelo ARIMA, es posible agregar variables exógenas al modelo SARIMA. Esta variante del modelo se conoce como SARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables for Seasonality).

### 3.4.6 Modelos SARIMAX: Aplicación

Para aplicar este modelo a los datos se tuvieron en cuenta las mismas variables exógenas aplicadas al modelo ARIMAX. A la función “statsmodels.tsa.statespace.SARIMAX” se le agregó el parámetro “exog” como un data frame que incluía las tres variables exógenas seleccionadas. Luego de aplicó el método Grid Search para encontrar la combinación de parámetros con mejor performance MSE.

*Tabla 20. Modelo Sarimax*

SARIMAX Model							
	p	d	q	P	D	Q	MSE
1	0	0	2	1	0	1	361,56
22	0	1	0	2	1	0	661,49
1057	1	1	2	2	0	2	112,98
1074	2	1	2	0	0	0	616,36
1086	0	0	0	1	1	0	4899,71

### 3.5 Mejor Modelo MSE

Para cada uno de los consumibles se eligió el mejor modelo considerando la performance del error cuadrático medio (MSE). En la tabla 21 se resaltan los mejores modelos en color verde.

*Tabla 21. Mejores modelos por performance MSE*

	MSE			
	ARIMA	ARIMAX	SARIMA	SARIMAX
1	1034,82	445,21	1072,25	361,56
22	908,80	721,94	696,97	661,49
1057	37,87	217,67	37,42	112,98
1074	755,86	641,43	755,21	616,36
1086	10756,51	5429,05	10310,27	4899,71

### 3.6 Desempeño in sample vs. out of sample

Una vez elegidos los modelos que mejor desempeño tuvieron según los criterios mencionados en el punto anterior, se procede a realizar la predicción para cada uno de los consumibles.

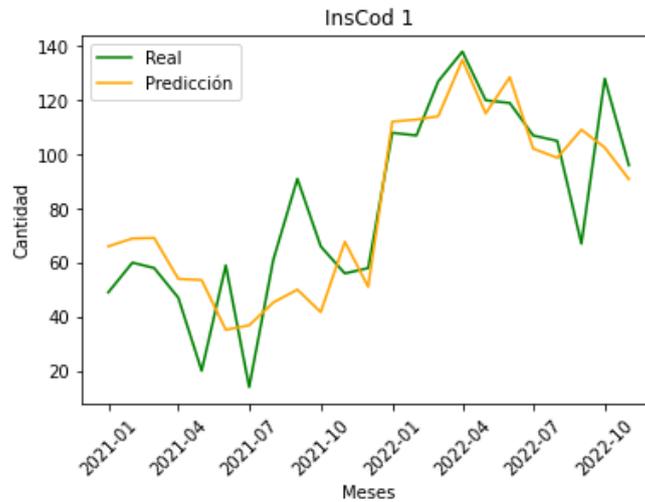
Para calcular los valores de las variables exógenas “USD Oficial” y “EESS atendidas” en el periodo de predicción (julio 2023 – junio 2024) se tuvieron en cuenta las siguientes fuentes de información:

- Precio del USD oficial en ARS para cada mes: Se consideró el precio del dólar futuro ROFEX al día 08/05/2023 para el periodo julio 2023 – abril 2024. El precio futuro para los meses mayo y junio 2024 no estaban disponibles por ende se decidió estimarlos asumiendo que aumentará un 6% cada mes como lo hacen en el mes de marzo y abril 2024.
- EESS atendidas: En este caso se tomó el promedio de EESS atendidas por mes de los últimos 5 años.

#### *InsCod 1*

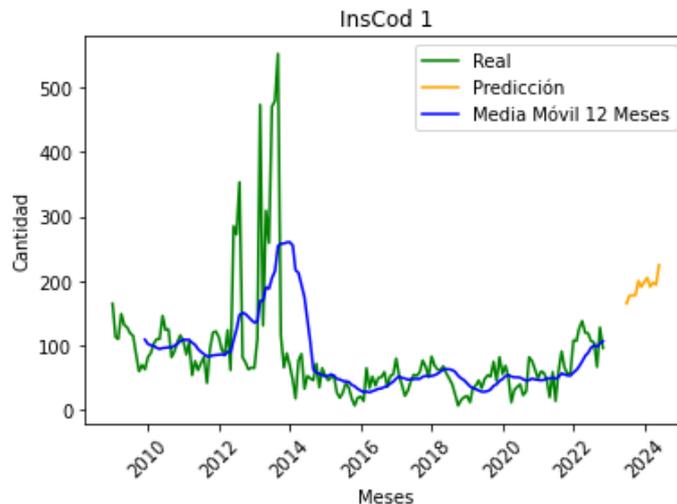
El modelo SARIMAX (0, 0, 2)(1, 0, 1) fue el elegido para predecir la demanda del Pico OPW 11A. Este consumible muestra un buen desempeño dentro de la muestra que se ve reflejado en el Gráfico 13.

Gráfico 13. Predicción InsCod 1 in sample



Sin embargo, al momento de generar las predicciones y mostrar un desempeño “out of sample”, el modelo predice una demanda muy por encima de la media.

Gráfico 14. Predicción InsCod 1 out of sample



Se presume que este resultado puede estar influenciado por dos motivos:

1. Un *overfitting* de los datos de entrenamiento dado por la elección del MSE como métrica de performance. El MSE elige el mejor modelo teniendo en cuenta el promedio de errores entre la predicción y el dato real pero no aporta ningún tipo de penalización al modelo por su complejidad. Este comportamiento puede llevar a un sobreajuste en los datos de entrenamiento haciendo que el modelo elegido no generalice correctamente en datos nuevos.
2. Las variables exógenas elegidas no explican la variable incógnita. De hecho solo una de ellas es de significancia para el modelo y esta es la variable “EESS atendidas por mes”. Sin embargo, esta variable no es 100% conocida por nosotros al momento de utilizarla para predecir la demanda futura, por lo cual, el modelo está utilizando datos que no son ciertos para hacer predicciones futuras.

Se procede entonces a utilizar la métrica AIC (Akaike Information Criterion) para buscar un modelo que equilibre la precisión en la predicción y la complejidad del modelo creado. Esta métrica indica que cuanto menor es el AIC, mejor desempeño tendrá el modelo.

Debido al análisis realizado sobre el desempeño "out of sample", solo se tendrán en cuenta los modelos ARIMA y SARIMA, ya que no se incorporarán las variables exógenas anteriormente seleccionadas.

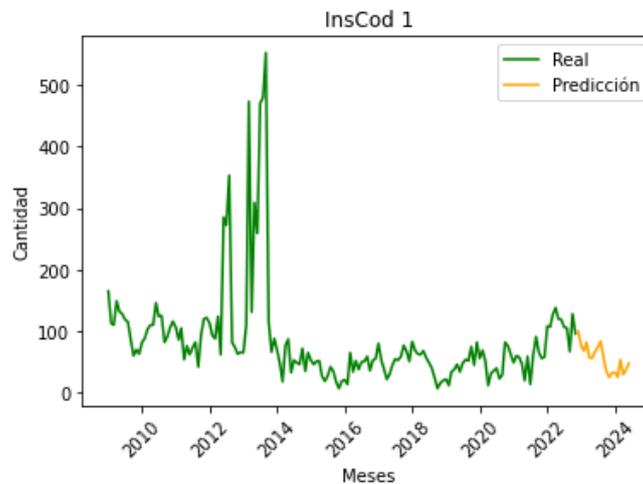
Para encontrar el modelo que minimice el AIC, se utiliza nuevamente el método de Grid Search. En este caso, se solicita al código que itere en busca del mejor AIC, utilizando la función "model.fit().aic" de la librería "statsmodels". De esta manera, se obtienen los mejores modelos para el consumible 1 como se visualizan en la Tabla 22.

**Tabla 22.** Mejores modelos AIC - InsCod 1.

INS COD 1							
	p	d	q	P	D	Q	AIC
ARIMA	2	1	2	0	0	0	1597,00
SARIMA	2	1	2	0	1	1	1497,00

En el Grafico 15 se muestran las predicciones out of sample creadas por el modelo SARIMA(2,1,2)(0,1,1). Dicho modelo presentó mejor desempeño según performance AIC.

**Gráfico 15.** Predicción InsCod 1 out of sample



Las predicciones del modelo se detallan en la Tabla 23.

**Tabla 23.** Predicciones InsCod 1

	Ins Cod 1 (unidades)
jul-23	74
ago-23	84
sep-23	57
oct-23	36
nov-23	26
dic-23	33
ene-24	34
feb-24	26
mar-24	54
abr-24	30
may-24	38
jun-24	49

*InsCod 22*

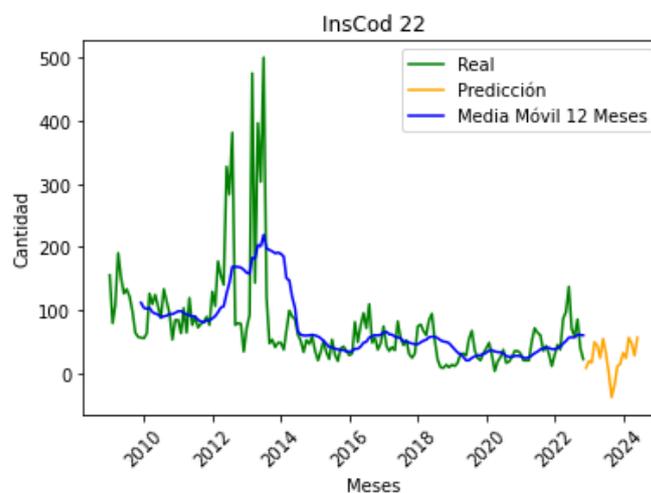
Tomando como ejemplo lo sucedido con el consumible 1, se procede a buscar el modelo con el mejor rendimiento según el criterio AIC para evitar el sobreajuste. En este caso, se realiza un paso previo de diferenciación en la serie para lograr su estacionalidad. Los resultados obtenidos se muestran en la Tabla 24.

**Tabla 24.** Mejores modelos AIC - InsCod 22

INS COD 22							
	p	d	q	P	D	Q	AIC
ARIMA	2	0	2	0	0	0	1591,59
SARIMA	2	0	2	0	1	1	1483,02

Nuevamente se visualiza que el modelo SARIMA presenta mejor desempeño que el modelo ARIMA.

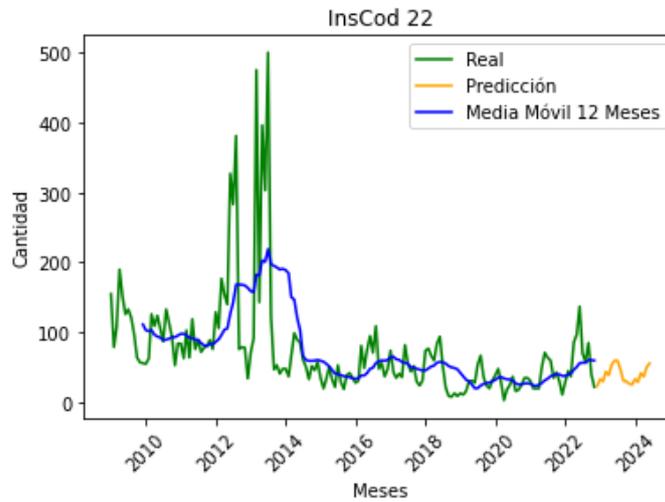
**Gráfico 16.** Predicción InsCod 22 out of sample



Al ajustar el modelo SARIMA para realizar las predicciones de demanda, el modelo arroja cantidades negativas que pueden apreciarse en el Gráfico 16.

Para abordar esta problemática, se opta por aplicar una transformación diferente a los datos de entrada. En lugar de diferenciar directamente la serie de tiempo, se realiza una transformación logarítmica a los datos y luego se aplica la diferenciación especificando  $d=1$  en los parámetros del modelo SARIMA. El valor del AIC obtenido para este modelo es de 247.56.

**Gráfico 17.** Predicción *InsCod 22 out of sample*



Las predicciones generadas se detallan en la Tabla 25.

**Tabla 25.** Predicciones *InsCod 22*

	<b>Ins Cod 22 (unidades)</b>
jul-23	60
ago-23	48
sep-23	31
oct-23	31
nov-23	27
dic-23	26
ene-24	34
feb-24	30
mar-24	42
abr-24	37
may-24	51
jun-24	56

### *InsCod 1057*

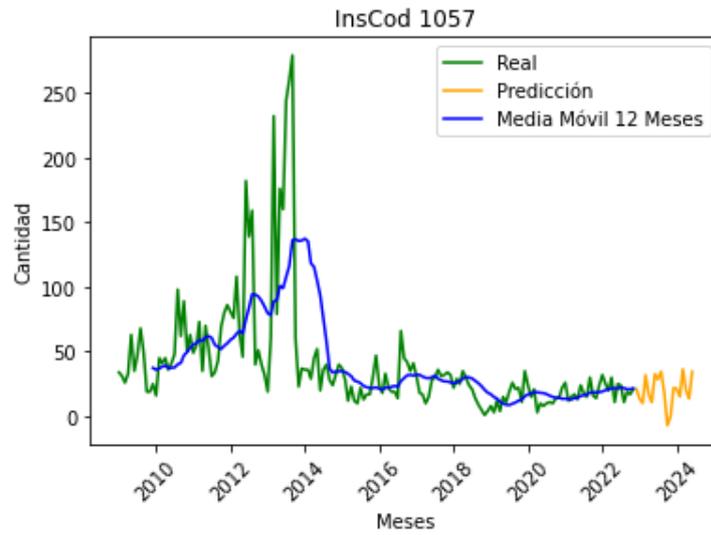
Para el consumible 1057 se procede a diferenciar la serie antes de entrenar el modelo en búsqueda del mejor AIC. En la Tabla 26 se identifica al modelo SARIMA como el de menor AIC.

**Tabla 26.** Mejores modelos AIC - *InsCod 1057*

<b>INS COD 1057</b>							
	<b>p</b>	<b>d</b>	<b>q</b>	<b>P</b>	<b>D</b>	<b>Q</b>	<b>AIC</b>
<b>ARIMA</b>	2	0	2	0	0	0	1423,18
<b>SARIMA</b>	2	0	2	0	1	1	1332,86

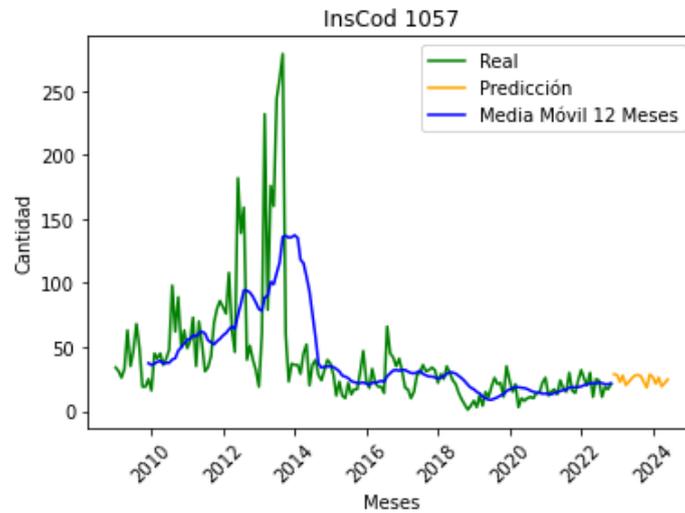
Así como sucedió con el consumible 22, el modelo SARIMA seleccionado también arroja predicciones negativas que se visualizan en el Gráfico 18.

Gráfico 18. Predicción InsCod 1057 out of sample



Se procede a realizar una transformación logarítmica y a confeccionar un test ADF para comprobar la estacionalidad de los datos. Con el test ADF confirmando la estacionalidad, se avanza en el entrenamiento del modelo SARIMA (2,0,2)(0,1,1). Este modelo arroja un AIC de 277,40.

Gráfico 19. Predicción InsCod 1057 out of sample



Las predicciones obtenidas por el modelo se muestran en la Tabla 27.

**Tabla 27.** Predicciones InsCod 1057

	Ins Cod 1057 (unidades)
jul-23	28
ago-23	29
sep-23	28
oct-23	23
nov-23	19
dic-23	29
ene-24	27
feb-24	22
mar-24	27
abr-24	20
may-24	22
jun-24	25

*InsCod 1074*

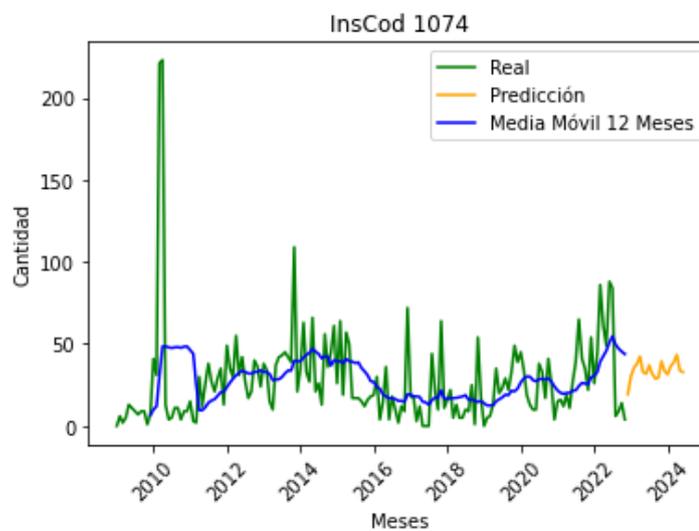
Al realizar la búsqueda del modelo con mejor desempeño por medio del método Grid Search, se concluye que el modelo SARIMA(0,1,2)(0,1,1) es el que minimiza el AIC para el consumible 1074.

**Tabla 28.** Mejores modelos AIC - InsCod 1074

INS COD 1074							
	p	d	q	P	D	Q	AIC
ARIMA	0	1	2	0	0	0	1356,00
SARIMA	0	1	2	0	1	1	1279,00

En el Gráfico 20 se pueden visualizar las predicciones out of sample del modelo seleccionado.

**Gráfico 20.** Predicción InsCod 1074 out of sample



Las predicciones generadas por el modelo de mejor desempeño AIC se muestran en la Tabla 29.

**Tabla 29.** Predicciones InsCod 1074

	<b>Ins Cod 1074 (unidades)</b>
jul-23	37
ago-23	32
sep-23	29
oct-23	30
nov-23	40
dic-23	34
ene-24	32
feb-24	37
mar-24	39
abr-24	44
may-24	34
jun-24	33

#### InsCod 1086

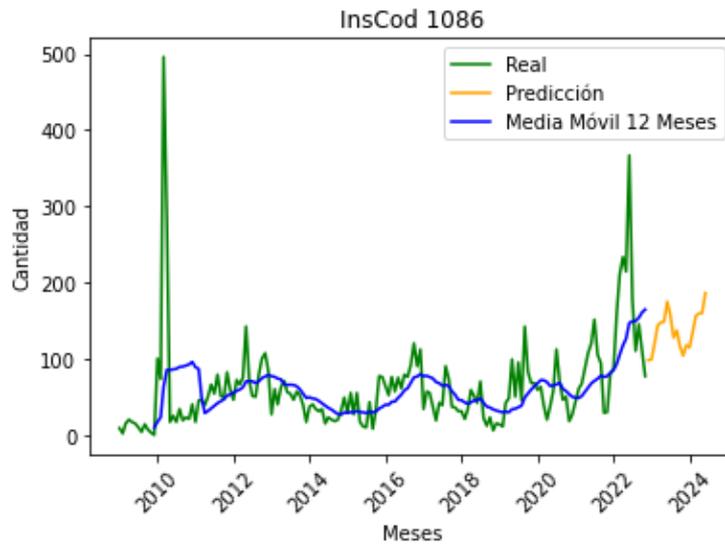
Por último, se utiliza nuevamente el método Grid Search para encontrar el modelo que mejor desempeño tendrá al momento de predecir la demanda del consumible 1086. Se obtuvieron los modelos detallados en la Tabla 30.

**Tabla 30.** Mejores modelos AIC - InsCod 1086

<b>INS COD 1086</b>							
	<b>p</b>	<b>d</b>	<b>q</b>	<b>P</b>	<b>D</b>	<b>Q</b>	<b>AIC</b>
<b>ARIMA</b>	1	1	1	0	0	0	1504,00
<b>SARIMA</b>	1	1	1	0	1	1	1417,00

Las predicciones out of sample se visualizan en el Gráfico 21.

**Gráfico 21.** Predicción InsCod 1086 out of sample



Las predicciones generadas por el modelo de mejor desempeño AIC se muestran en la Tabla 31.

**Tabla 31.** Predicciones InsCod 1086

	<b>Ins Cod 1086 (unidades)</b>
jul-23	158
ago-23	129
sep-23	138
oct-23	118
nov-23	105
dic-23	119
ene-24	116
feb-24	137
mar-24	157
abr-24	161
may-24	161
jun-24	187

## 4. Capítulo 4. Optimización de compra de insumos

El output de los modelos utilizados en la sección anterior se toma como input de un modelo de optimización que utiliza programación lineal para poder generar una recomendación de compra óptima de consumibles a “M Surtidores” para el periodo julio 2023 – junio 2024.

La programación lineal es una técnica de optimización que se emplea para resolver problemas de maximización o minimización sujetos a restricciones lineales. El objetivo de la programación lineal es calcular el valor de las variables de una función que la maximizan o minimizan dependiendo de cómo se haya planteado el problema. Para ello se brindan restricciones en forma de ecuaciones o inecuaciones que limitan los valores que puede tener cada variable.

### 4.1 Describiendo el problema de optimización

Para llevar a cabo la optimización se toma como referencia un modelo de optimización de tipo “Lot Sizing” donde, dada una demanda estimada de un tipo de producto (o varios) para un periodo futuro específico, se determina qué cantidad de productos se deben comprar (o fabricar) para satisfacer dicha demanda.

En el problema planteado para “M Surtidores” los productos serían los cinco consumibles demandados por las EESS que atienden.

#### 4.1.1 Función objetivo

La función objetivo busca minimizar la sumatoria de los costos de cada insumo “i” en cada periodo “t”.

$$\min \sum_{i \in I} \sum_{t \in T} c_{it} x_{it}$$

#### 4.1.2 Componentes de la función objetivo

- El conjunto “I” se compone de los cinco consumibles “i”: 1, 22, 1057, 1074 y 1086.
  - $I = \{1, \dots, 5\}$
- El conjunto “T” se compone de los 12 periodos “t”. Abarca desde el mes de julio 2023 hasta el mes de junio 2024.
  - $T = \{1, \dots, 12\}$
- $c_{it}$  es el costo de compra de “i” en el periodo “t”. Es un parámetro que se brinda al modelo.
- $x_{it}$  es la cantidad de insumo “i” a comprar en el periodo “t”. Esta variable será uno de los outputs del modelo de optimización.

#### 4.1.3 Restricciones

**Cumplimiento de la demanda:** Para todos los insumos “i” en todos los periodos “t”, la demanda del insumo “i” debe ser cumplida. Esta restricción es brindada al modelo a través de la ecuación e inecuación:

$$s_{it} = s_{i,t-1} + x_{it} - d_{it} \quad \forall i \in I, t \in T$$

$$s_{it} \geq 0 \quad \forall i \in I, t \in T$$

- La variable  $s_{it}$  representa la cantidad de stock del insumo “i” al final del periodo “t”. Esta variable debe ser no negativa para poder cumplir con la demanda.
- El parámetro  $d_{it}$  representa la demanda del insumo “i” en el periodo “t”. Este parámetro es brindado por los modelos entrenados en la sección anterior.
- El stock del insumo “i” en el periodo “t-1” sumado a lo que debe comprarse de dicho insumo en el periodo “t” debe ser mayor o igual a la demanda de ese insumo en el periodo “t”.

Output no negativo: El modelo restringe a la variable  $x_{it}$  con la siguiente inecuación para que solo pueda tomar valores positivos.

$$x_{it} \geq 0 \quad \forall i \in I, t \in T$$

Este primer modelo de optimización es “naive” ya que supone que “M Surtidores” no tiene restricción de dinero para comprar cualquiera de sus consumibles y a su vez tampoco considera ningún tipo de restricción de ocupación en su depósito. Además, presume que el modelo de machine learning creado en la sección anterior predijo la demanda para cada consumible y cada periodo con una extrema exactitud. De esta manera se procede a tomar este modelo como base y a adaptarlo con restricciones adicionales para crear una correcta recomendación al negocio.

## 4.2 Profundizando el modelo de optimización

Costos: Cada insumo “i” tiene un costo de compra asociado a cada periodo “t”. “M Surtidores” compartió el costo actualizado a marzo 2023 de los cinco insumos.

**Tabla 32.** Costo (USD) adquisición de insumos

Insumos	Costo Marzo 2023 (USD)
1	128,80
22	56,60
1057	139,85
1074	12,13
1086	11,70

Si bien todos los consumibles son insumos importados, “M Surtidores” no los importa directamente, sino que, los compra localmente al distribuidor autorizado.

Los consumibles “Pico”, “Acople” y “Manguera” (1, 22 y 1057 respectivamente) son insumos provenientes de Brasil por ende su costo se ve afectado por la inflación del país vecino. Por otro lado, los consumibles “Filtros” (tanto el “InsCod” 1074 como el 1086) son insumos importados de Estados Unidos y su costo se ve afectado por la inflación de ese país. De esta manera, para poder calcular un costo estimado para el periodo julio 2023 – junio 2024 y así poder crear el conjunto de parámetros  $C$ , se tuvieron en cuenta las expectativas inflacionarias de cada país.

En el caso de Brasil, se consultó la página del Banco Central de dicho país donde se informan las expectativas de inflación del mercado<sup>7</sup>. En el *apéndice C* se detalla la expectativa de inflación

<sup>7</sup> <https://www.bcb.gov.br/estatisticas/detalhamentoGrafico/graficosestatisticas/precos>

para el periodo abril 2023 – julio 2024 y se puede ver como evolucionarían los costos de los “InsCod” 1, 22 y 1057 durante ese periodo de tiempo.

En el caso de EE.UU., se tuvo en consideración la inflación real del mes de abril 2023 para luego proyectar la variación del índice teniendo en cuenta la expectativa de inflación anual para los años 2023 y 2024. Las expectativas de inflación fueron obtenidas de la página web oficial del Fondo Monetario Internacional<sup>8</sup>. En el *apéndice D* se detalla la evolución de los costos de compra de los insumos 1074 y 1086 basados en la evolución mensual de la inflación proyectada.

**Restricción de capital disponible:** En este punto se debe aclarar de qué manera “M Surtidores” cobra sus servicios brindados ya que de estos ingresos depende el presupuesto disponible con el que contará de manera mensual para poder comprar los insumos demandados.

Por un lado, la empresa obtiene **ingresos por venta de insumos** por cada insumo requerido en cada una de las incidencias atendidas. El beneficio que la empresa obtiene de estas ventas es de un 10% sobre el costo del insumo. De esta manera, el beneficio por insumo para el mes de marzo 2023 se detalla en la Tabla 33.

*Tabla 33. Beneficio (USD) por venta de insumos*

Insumos	mar-23		
	Costo (USD)	Precio de Venta (USD)	Beneficio (USD)
1	128,80	141,68	12,88
22	56,60	62,26	5,66
1057	139,85	153,84	13,99
1074	12,13	13,34	1,21
1086	11,70	12,87	1,17

El plazo de cobro de estos ingresos por ventas es de 60 días, esto implica que los insumos demandados por las EESS en el mes “t” serán cobrados por “M Surtidores” en el mes “t+2”.

Por otro lado, “M Surtidores” **cobra un abono mensual** a cada una de las EESS que solicitan sus servicios, independientemente de si durante ese mes en cuestión la estación informó algún tipo de incidencia o no. La empresa indicó que este abono es la principal fuente de ingreso del negocio y proporcionó información sobre el último monto actualizado en ARS correspondiente a marzo 2023.

El abono mensual se utiliza para pagar todos los costos asociados al negocio, como sueldos y gastos fijos, y la empresa estima que puede llegar a destinar hasta un 10% de ese monto en la compra de consumibles. Sin embargo, durante el mes de junio y de diciembre este porcentaje se reduciría a un 5% debido al gasto extra que deben hacer para pagar los aguinaldos de los empleados.

A los fines de poder llevar a cabo el análisis en una sola moneda, se toma el abono en USD para poder agregar este parámetro al modelo. La empresa comenta que el abono debe seguir a la inflación por ende se asumirá que el tipo de cambio oficial va a seguir el ritmo de la inflación. De esta manera, se toma un tipo de cambio oficial de 227,5<sup>9</sup> ARS por USD para poder transformar el abono en moneda extranjera.

<sup>8</sup> <https://www.imf.org/>

<sup>9</sup> Dólar Oficial para la venta al 26/04/2023 según página web del BCRA.

Teniendo en cuenta estos dos ingresos, se restringe al modelo con la siguiente inecuación para que solo compre insumos en el mes “t” considerando el capital disponible durante dicho mes:

$$e_t + v \sum_{i \in I} c_{i,t-2} x_{i,t-2} - \sum_{i \in I} c_{it} x_{it} \geq 0 \quad \forall i \in I, t \in T$$

- El parámetro  $e_t$  hace referencia al **ingreso por abono** disponible en el periodo “t”.
- El término  $v \sum_{i \in I} c_{i,t-2} x_{i,t-2}$  indica el ingreso por ventas del mes “t-2” multiplicando al término  $v$  que será igual a 1,1 por el costo.
- El término  $\sum_{i \in I} c_{it} x_{it}$  representa el costo por compra de insumo en el mes “t”.

Al indicar que dicha inecuación debe ser mayor o igual a 0 se restringe al modelo para que no gaste más que el ingreso disponible para insumos durante el mes “t”.

**Compra mínima:** Para que el modelo no recomiende compras mensuales de montos pequeños se agrega una restricción de compra mínima de consumible por mes. “M Surtidores” informa que está dispuesto a colocar una orden de compra mensual a los proveedores si las mismas igualan o superan las cantidades informadas en la Tabla 34.

**Tabla 34.** Compra mínima

Insumos	Compra Mínima (unidades)
1	20
22	30
1057	20
1074	200
1086	200

Esta restricción se representa matemáticamente a través de dos inecuaciones:

$$x_{it} \leq \text{BIGNUMBER} * y_{it} \quad \forall i \in I, t \in T$$

$$x_{it} \geq m_i * y_{it} \quad \forall i \in I, t \in T$$

- La variable  $y_{it}$  es una variable binaria que indica 1 si en el mes “t” se debe comprar el insumo “i”, caso contrario indica 0.
- El parámetro  $m_i$  indica la cantidad mínima mensual a comprar del insumo “i”.
- Si se debe hacer una compra en el mes “t” la variable  $y_{it}$  será igual a 1 restringiendo al modelo para que  $x_{it}$  sea mayor o igual a la compra mínima mensual.
- La inecuación que multiplica a  $y_{it}$  por un número grande (parámetro BIG NUMBER) actúa como complemento de la segunda inecuación: Si  $y_{it}$  es 0 (no hay compras de “i” en el mes “t”) entonces  $x_{it}$  debe ser 0. Si el modelo necesita hacer una compra en el mes “t”, necesariamente deberá hacer que  $y_{it}$  sea igual a 1 haciendo que  $x_{it}$  iguale o supere la compra mínima ( $m_i$ )

**Stock de seguridad:** La empresa necesita contar con un stock adicional para hacer frente a variaciones en la obtención de los insumos y/o fluctuaciones en la demanda. De esta manera, la inecuación de la variable de stock se restringe con un nuevo parámetro:

$$s_{it} \geq w_{it} \quad \forall i \in I, t \in T$$

- El parámetro  $w_{it}$  representa el stock de seguridad del insumo “i” en el periodo “t”.
- Esta inecuación afecta a la ecuación de stock  $s_{it} = s_{i,t-1} + x_{it} - d_{it} \quad \forall i \in I, t \in T$  ya que restringe a  $s_{it}$  para que cumpla con la demanda y adicionalmente con el stock de seguridad.

Para calcular el stock de seguridad de cada insumo, se realizó la fórmula:

$$(\text{Plazo máximo de entrega} - \text{Plazo de entrega normal}) \times \text{Demanda Media} = \text{Stock de Seguridad}$$

- *Plazo máximo de entrega:* es el periodo máximo que puede tardar el proveedor en entregar los insumos una vez realizado el pedido.
- *Plazo de entrega normal:* es el tiempo en el que el proveedor entrega los insumos generalmente.
- *Demanda media:* es la demanda promedio mensual.

El negocio proporcionó información sobre los plazos de entrega de sus proveedores para poder calcular el stock de seguridad en cada caso.

**Tabla 35.** Stock de Seguridad

Insumos	Plazo Max Entrega	Plazo Entrega Normal	Demanda Media <sup>10</sup> (UN)	Stock de Seguridad (UN)
1	5	3	81	162
22	5	3	50	100
1057	5	3	20	40
1074	30	20	36	360
1086	30	20	125	1250

**Restricción de espacio:** La empresa informa que no cuenta con restricciones de espacio ya que tienen un amplio depósito y están dispuestos a extenderlo si así su demanda lo necesitara.

<sup>10</sup> Como demanda mensual se tomó el promedio mensual de cada insumo para el periodo 2021-2022.

#### 4.2.1 Modelo Final

Luego de haber adaptado el modelo de “lot sizing” al problema particular de la empresa, se concluye con la obtención de las siguientes constantes y variables:

##### Constantes

- $q = 5 \rightarrow$  Cantidad de consumibles. Define la cantidad de consumibles en el conjunto  $I = \{1 \dots q\}$
- $n = 12 \rightarrow$  Número de meses del año. Define la cantidad de periodos en el conjunto  $T = \{1 \dots n\}$
- $v = 1.1 \rightarrow$  Valor utilizado para calcular el beneficio de la venta del insumo.

##### Variables

- $x_{it} \rightarrow$  Variables de compra. Representa la cantidad de insumo “i” a comprar en el periodo “t”. Esta variable solo puede tomar valores enteros.
- $s_{it} \rightarrow$  Variables de stock. Representa la cantidad de stock del insumo “i” al final del periodo “t”. Esta variable solo puede tomar valores enteros.
- $y_{it} \rightarrow$  Variable binaria que indica 1 si en el mes “t” se debe comprar el insumo “i”, caso contrario indica 0.

Además, se detalla la función objetivo con todas sus restricciones:

$$\begin{aligned} \min \quad & \sum_{i \in I} \sum_{t \in T} c_{it} x_{it} \\ s_{it} = & s_{i,t-1} + x_{it} - d_{it} \quad \forall i \in I, t \in T \\ e_t + v \sum_{i \in I} c_{i,t-2} x_{i,t-2} - \sum_{i \in I} c_{it} x_{it} \geq & 0 \quad \forall i \in I, t \in T \\ x_{it} \leq & \text{BIGNUMBER} * y_{it} \quad \forall i \in I, t \in T \\ x_{it} \geq & m_i * y_{it} \quad \forall i \in I, t \in T \\ s_{it} \geq & w_{it} \quad \forall i \in I, t \in T \\ x_{it} \geq & 0 \quad \forall i \in I, t \in T \end{aligned}$$

Para llevar a cabo la optimización se utiliza el programa de solver SCIP<sup>11</sup>(Solving Constraint Integer Programs), un software de código abierto que se emplea para resolver problemas de programación entera mixta (Gamrath, G., Anderson, D., Bestuzheva, K., Chen, W. K., Eifler, L., Gasse, M., ... & Witzig, J, 2020). Para que el software pueda leer el modelo confeccionado y proporcionar una solución, se debe traducir el modelo en un lenguaje que el programa pueda leer. En este caso se utilizó el lenguaje Zimpl (Koch, 2004) para describir el problema de optimización al solver.

---

<sup>11</sup> scipopt.org

El modelo en lenguaje ZIMPL se puede visualizar en el [apéndice E](#).

#### 4.2.2 Presunciones

El modelo creado tiene en consideración las siguientes presunciones:

- El costo de los insumos fluctúa únicamente a causa de la inflación del país exportador.
- La empresa paga al proveedor en el mismo mes en el que el proveedor le envía la mercadería.
- La inflación sigue al tipo de cambio oficial.
- En el mes “t0” la empresa ya cuenta con stock de seguridad, de esta manera no es necesario comprar stock de seguridad en el mes “t1”.
- En el mes “t0” el stock de seguridad es el único stock que la empresa posee en dicho mes.

#### 4.3 Escenarios

Para poder dar una recomendación al negocio sobre la compra de los insumos para el periodo julio 2023- junio 2024, se plasman a continuación distintos escenarios teniendo en cuenta posibles fluctuaciones en la demanda calculada en el capítulo anterior.

Escenario 1: La demanda indicada por el modelo de ML es exacta ( $d_{it} * 1$ ).

Escenario 2: La demanda indicada por el modelo de ML aumenta un 10% ( $d_{it} * 1,1$ ).

Escenario 3: La demanda indicada por el modelo de ML aumenta un 15% ( $d_{it} * 1,15$ ).

##### 4.3.1 Recomendación

Escenario 1:  $d_{it} * 1$

En este escenario “M Surtidores” debería comprar los insumos para el periodo julio 2023 – junio 2024 como se detallan en la Tabla 36.

**Tabla 36.** Escenario 1

	ESCENARIO 1					
	Picos (Unidades)	Acoples (Unidades)	Mangueras (Unidades)	Filtro 3/4 (Unidades)	Filtro 1" (Unidades)	Total a invertir (USD)
jul-23	74	60	28	200	201	\$ 21.972,78
ago-23	85	59	33	-	224	\$ 21.977,88
sep-23	165	354	26	-	-	\$ 46.140,51
oct-23	125	-	188	-	200	\$ 46.151,68
nov-23	92	-	24	221	1061	\$ 31.243,38
dic-23	-	-	-	-	-	-
ene-24	-	-	-	-	-	-
feb-24	-	-	-	-	-	-
mar-24	-	-	-	-	-	-
abr-24	-	-	-	-	-	-
may-24	-	-	-	-	-	-
jun-24	-	-	-	-	-	-

La inversión total a afrontar en este escenario sería de 167.486,23 USD para poder comprar los consumibles que las estaciones de servicio demandarán. El modelo indica que se debe comprar toda la demanda en los primeros cinco meses para minimizar los costos de compra de estos insumos.

Escenario 2:  $d_{it} * 1.10$

Este escenario hace que modelo no pueda brindar una solución factible ya que la demanda de los primeros meses supera el monto disponible para comprar dichos insumos. En esta situación, se recomienda modificar el parámetro “Stock de Seguridad” o aumentar el porcentaje de abono destinado a la compra de insumos.

*Opción 1:* Se modifica el Stock de Seguridad disminuyéndolo en un 10%.

$$\text{subto ss: forall } \langle i, t \rangle \text{ in } I * T: \\ s[i, t] \geq 0.9 * w[i, t] ;$$

**Tabla 37.** Escenario 2 – V1

ESCENARIO 2 - V1						
	Picos (Unidades)	Acoples (Unidades)	Mangueras (Unidades)	Filtro 3/4 (Unidades)	Filtro 1" (Unidades)	Total a invertir (USD)
jul-23	75	58	28	200	200	\$ 21.976,78
ago-23	86	83	41	-	-	\$ 21.977,25
sep-23	128	375	29	-	273	\$ 46.151,59
oct-23	118	-	211	-	-	\$ 46.149,24
nov-23	177	-	20	232	1261	\$ 44.555,49
dic-23	-	-	-	-	-	-
ene-24	-	-	-	-	-	-
feb-24	-	-	-	-	-	-
mar-24	-	-	-	-	-	-
abr-24	-	-	-	-	-	-
may-24	-	-	-	-	-	-
jun-24	-	-	-	-	-	-

En la Tabla 37 se visualizan los resultados de esta optimización. El negocio tendrá que invertir 180.810,35 USD para cubrir el aumento de la demanda en un 10% para el periodo estudiado.

Opción 2: Aumentando el abono 10% en jul-23 y 5% en ago-23.

Tabla 38. Escenario 2 – V2

ESCENARIO 2 - V2						
	Picos (Unidades)	Acoples (Unidades)	Mangueras (Unidades)	Filtro 3/4 (Unidades)	Filtro 1" (Unidades)	Total a invertir (USD)
jul-23	85	66	31	200	200	\$ 24.174,06
ago-23	92	67	33	-	200	\$ 23.076,43
sep-23	133	393	40	-	200	\$ 48.569,26
oct-23	139	-	200	-	-	\$ 47.353,44
nov-23	151	-	29	268	1259	\$ 42.810,59
dic-23	-	-	-	-	-	-
ene-24	-	-	-	-	-	-
feb-24	-	-	-	-	-	-
mar-24	-	-	-	-	-	-
abr-24	-	-	-	-	-	-
may-24	-	-	-	-	-	-
jun-24	-	-	-	-	-	-

La opción 2 indica que se debe invertir 185.983,78 USD para cubrir un aumento en la demanda de un 10% si el negocio no desea flexibilizar su política de stocks.

Escenario 3:  $d_{it} * 1.15$

Al igual que el escenario anterior, en esta situación el modelo arroja una solución no factible ya que la demanda supera al monto disponible para adquirir los insumos. Se plantean así dos opciones para solucionar este problema.

Opción 1: Se modifica el Stock de Seguridad disminuyéndolo en un 10%.

Tabla 39. Escenario 3 – V1

ESCENARIO 3 - V1						
	Picos (Unidades)	Acoples (Unidades)	Mangueras (Unidades)	Filtro 3/4 (Unidades)	Filtro 1" (Unidades)	Total a invertir (USD)
jul-23	70	63	30	200	207	\$ 21.977,13
ago-23	98	73	34	-	-	\$ 21.977,97
sep-23	109	404	33	-	294	\$ 46.148,73
oct-23	105	-	223	-	-	\$ 46.153,08
nov-23	230	-	27	253	1319	\$ 53.609,36
dic-23	-	-	-	-	-	-
ene-24	-	-	-	-	-	-
feb-24	-	-	-	-	-	-
mar-24	-	-	-	-	-	-
abr-24	-	-	-	-	-	-
may-24	-	-	-	-	-	-
jun-24	-	-	-	-	-	-

En este escenario “M Surtidores” invertiría 189.866,27 USD los primeros cinco meses del periodo estudiado para cubrir el aumento de la demanda en un 15%.

*Opción 2:* Aumentando el abono 15% en jul-23 y 5% en ago-23.

**Tabla 40.** Escenario 3 – V2

ESCENARIO 3 - V2						
	Picos (Unidades)	Acoples (Unidades)	Mangueras (Unidades)	Filtro 3/4 (Unidades)	Filtro 1" (Unidades)	Total a invertir (USD)
jul-23	89	71	33	200	200	\$ 25.270,39
ago-23	94	55	36	-	200	\$ 23.074,39
sep-23	125	424	41	-	226	\$ 49.766,53
oct-23	126	-	212	-	-	\$ 47.357,28
nov-23	194	-	29	289	1319	\$ 49.536,68
dic-23	-	-	-	-	-	-
ene-24	-	-	-	-	-	-
feb-24	-	-	-	-	-	-
mar-24	-	-	-	-	-	-
abr-24	-	-	-	-	-	-
may-24	-	-	-	-	-	-
jun-24	-	-	-	-	-	-

En este escenario, el negocio debe invertir 195.005,27 USD en la compra de consumibles para poder abastecer la demanda. Al igual que en los escenarios anteriores, el modelo recomienda que “M Surtidores” compre todo el stock a utilizar para el periodo julio 2023 – junio 2024 en los primeros cinco meses de dicho periodo.

## 5. Capítulo 5. Conclusión

### 5.1 Recomendación final

El presente trabajo final de maestría ha logrado proporcionar herramientas a “M Surtidores” para poder tomar decisiones fundamentadas en la valiosa información con la que cuentan. La base de datos brindada por la empresa permitió la creación de los modelos de machine learning de tipo ARIMA y del modelo de optimización de tipo lot sizing que fueron confeccionados y adaptados especialmente a la naturaleza de los datos del negocio y al problema que se quería resolver. Los modelos de machine learning lograron predecir la demanda de los cinco insumos más solicitados para el periodo comprendido entre julio 2023 y junio 2024. Con esta información, el modelo de optimización pudo brindar recomendaciones de compra para minimizar el costo de adquisición de los insumos. Luego de analizar los tres escenarios planteados en el capítulo anterior, se recomienda a la empresa comprar todos los insumos demandados durante los primeros cinco meses del periodo de estudio.

### 5.2 Síntesis final: Conocimientos adquiridos, limitaciones y futuras mejoras

Analizar, asociar y limpiar las diversas tablas compartidas por la empresa representó un desafío por la gran cantidad de información que cada una de ellas contenía. Además, por sí solas no proporcionaban información completa. En este sentido, los conocimientos aportados por el negocio fueron de gran ayuda para comprender qué información brindaba cada atributo en cada tabla. A medida que nos adentrábamos en el mundo de los surtidores, la colaboración del negocio resultó crucial para dar sentido a los primeros *insights* obtenidos.

La investigación exhaustiva de la base de datos relacional nos llevó a ajustar el objetivo de predicción y optimización para centrarnos en los “consumibles”, los insumos más demandados por los clientes de “M Surtidores”. Una vez identificados los insumos a predecir, nos aventuramos en la elección del modelo de machine learning. Se analizó cada una de las series de tiempo para determinar si los modelos ARIMA, inicialmente considerados, eran adecuados para abordar el problema. Comprendimos que “el mejor modelo”, debía necesariamente ser aquel que pudiera equilibrar la precisión en la predicción y la complejidad del algoritmo para ser capaz de adaptarse a datos desconocidos y dar una buena performance “out of sample”.

En esta etapa de forecasting también se prestó especial atención a la correcta división de los datos en “datos de entrenamiento” y “datos de test” para evitar que el modelo utilice información del futuro para predecir datos pasados. Dimos cuenta que las variables exógenas agregadas a los modelos no respondían como esperábamos, ya que no estaban relacionadas con las variables a predecir. A este acontecimiento se adicionaba el hecho de que los datos no eran del todo certeros al momento de ser brindados al modelo para hacer predicciones futuras. En próximas investigaciones, se podría explorar la creación de modelos de forecasting con otros algoritmos que permitan utilizar los atributos de la base de datos brindada por el negocio para poder explicar la demanda de los insumos.

En cuanto a la optimización, inicialmente se consideró un modelo de optimización de portfolio, pero se comprendió que el modelo de lot sizing se adaptaba mejor a la necesidad del negocio de determinar cuándo sería conveniente comprar los insumos para abastecer a la demanda predicha. Fue acertado partir de un modelo base y adaptarlo a la situación particular de “M Surtidores”. Sin embargo, debemos hacer mención a una limitante de esta optimización: un parámetro tan sensible e importante como el aumento del costo de adquisición de cada insumo, no se puede conocer con precisión y debe ser estimado. Para mejorar la optimización, en futuras investigaciones, se podrían utilizar técnicas de modelado de datos para construir un modelo que

prediga el aumento de los costos de compra de los insumos, utilizando información pasada y datos macroeconómicos.

Para poder implementar el modelo matemático en el solver, se utilizó el lenguaje ZIMPL y se tuvo que ser muy cuidadoso al escribirlo. Aprendimos que, para evitar errores en la escritura, debíamos crear variables y varias inecuaciones simples para llegar a una sola restricción en vez de crear una sola inecuación compleja.

Aunque para brindar las recomendaciones se plantearon tres escenarios, es importante destacar que las posibilidades son ilimitadas. En nuestro caso, se decidió mostrar que pasaría si la demanda presentara fluctuaciones. Solo se permitió relajar los supuestos de stock de seguridad y/o la posibilidad de aumentar el capital disponible para invertir en la compra de insumos durante los primeros meses.

Con el resultado del modelo para los tres escenarios, se pudo observar que la recomendación es adquirir todo el stock necesario en los primeros cinco meses del periodo estudiado. Otras extensiones podrían incluir la implementación de restricciones para evitar que el stock del mes ( $S(i,t)$ ) supere un límite establecido por la empresa. Esta medida no estaría relacionada a restricciones de espacio físico, sino que buscaría evitar que "M Surtidores" acumule excesivo stock de una demanda con fluctuaciones. Además, para ampliar el análisis y enfocarse en el beneficio de la empresa, se podría agregar un modelo de optimización de portfolio que tome en cuenta el capital disponible después de haber realizado la compra recomendada por el modelo de lot sizing. De esta manera, se podría optimizar el rendimiento del capital disponible después de la compra de insumos.

En conclusión, las herramientas de análisis, predicción y optimización presentadas en este trabajo se combinaron con los conocimientos aportados por el negocio para mejorar la eficiencia de su proceso de compra de insumos. Esta sinergia representa una oportunidad para que "M Surtidores" capitalice su amplia experiencia en la industria y la combine con las avanzadas herramientas de análisis de datos disponibles en la actualidad para, por ejemplo, minimizar sus costos y, en consecuencia, obtener mayores beneficios en el futuro.

## Bibliografía

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). Time series analysis: Forecasting and control (4th ed.). Wiley.

Brockwell, P. J., & Davis, R. A. (2016). Introduction to time series and forecasting (3rd ed.). Springer.

Greene, W. H. (2011). Econometric analysis (7th ed.). Pearson.

G. Gamrath, D. Anderson, K. Bestuzheva, W.-K. Chen, L. Eifler, M. Gasse, P. Gemander, A. Gleixner, L. Gottwald, K. Halbig, G. Hendel, C. Hojny, T. Koch, P. Le Bodic, S. J. Maher, F. Matter, M. Miltenberger, E. Mühmer, B. Müller, M. Pfetsch, F. Schlösser, F. Serrano, Y. Shinano, C. Tawfik, S. Vigerske, F. Wegscheider, D. Weninger, J. Witzig (2020). The SCIP Optimization Suite 8.0. ZIB-Report 20-10.

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.).

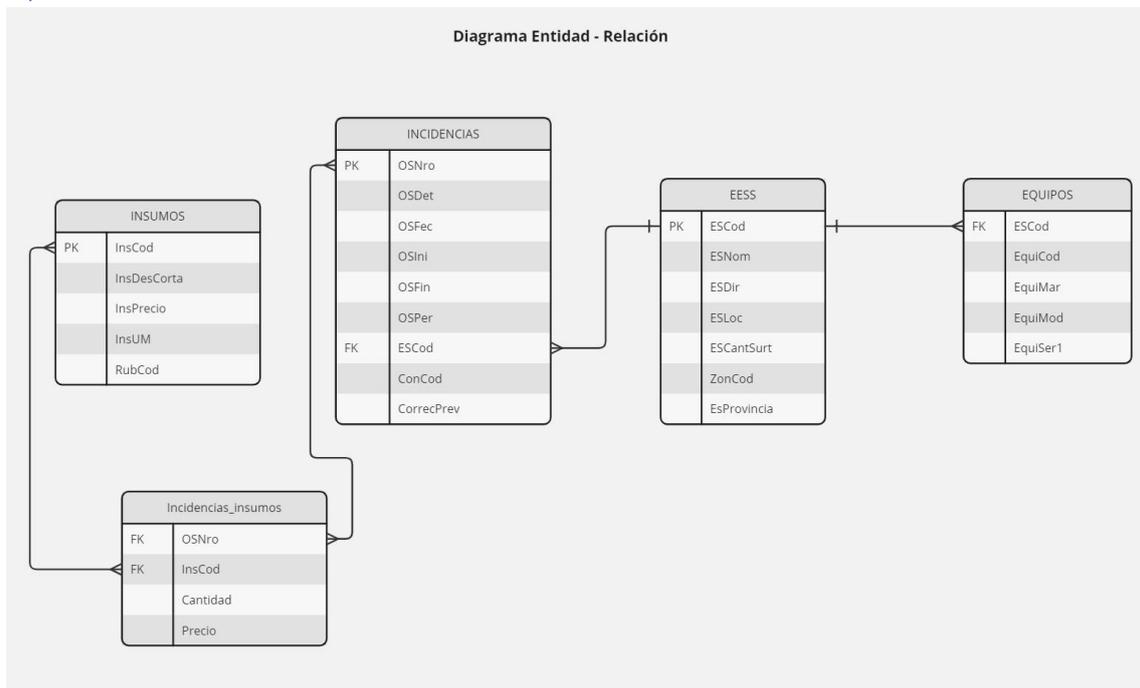
Koch, T. (2004). Rapid mathematical programming (Doctoral dissertation, Ph. D. dissertation, Technische Universität Berlin, 2004, ZIB-Report 04-58. {Online}. Available: <http://www.zib.de/Publications/abstracts/ZR-04-58>).

Konishi, S., & Kitagawa, G. (2008). Information criteria and statistical modeling.

Shumway, R. H., & Stoffer, D. S. (2017). Time series analysis and its applications: With R examples (4th ed.). Springer.

# Apéndices

## Apéndice A



## Apéndice B

Fecha	Laborable day	USD oficial	EESS Atendidas	Fecha	Laborable day	USD oficial	EESS Atendidas	Fecha	Laborable day	USD oficial	EESS Atendidas
1/1/2009	22	3,49	364	1/9/2013	20	5,79	342	1/5/2018	22	24,96	228
1/2/2009	19	3,57	344	1/10/2013	22	5,91	334	1/6/2018	20	28,94	236
1/3/2009	22	3,72	412	1/11/2013	20	6,14	331	1/7/2018	22	27,42	207
1/4/2009	21	3,71	395	1/12/2013	20	6,52	338	1/8/2018	22	36,90	195
1/5/2009	20	3,75	347	1/1/2014	22	8,02	367	1/9/2018	20	41,29	127
1/6/2009	22	3,80	392	1/2/2014	18	7,89	345	1/10/2018	23	35,90	115
1/7/2009	23	3,83	366	1/3/2014	22	8,00	331	1/11/2018	20	37,73	127
1/8/2009	20	3,85	388	1/4/2014	22	8,00	351	1/12/2018	21	37,65	123
1/9/2009	22	3,84	370	1/5/2014	21	8,08	301	1/1/2019	23	37,30	126
1/10/2009	21	3,82	300	1/6/2014	20	8,13	240	1/2/2019	20	39,13	123
1/11/2009	20	3,81	301	1/7/2014	23	8,21	293	1/3/2019	22	43,36	156
1/12/2009	23	3,80	283	1/8/2014	21	8,40	262	1/4/2019	19	44,30	160
1/1/2010	21	3,84	321	1/9/2014	20	8,43	228	1/5/2019	22	44,77	166
1/2/2010	19	3,86	351	1/10/2014	22	8,50	209	1/6/2019	20	42,48	165
1/3/2010	22	3,88	400	1/11/2014	19	8,53	182	1/7/2019	23	43,82	182
1/4/2010	22	3,89	396	1/12/2014	22	8,55	208	1/8/2019	21	59,50	213
1/5/2010	18	3,93	362	1/1/2015	22	8,63	182	1/9/2019	20	57,59	217
1/6/2010	21	3,93	374	1/2/2015	16	8,72	155	1/10/2019	23	59,57	200
1/7/2010	23	3,94	387	1/3/2015	21	8,82	187	1/11/2019	21	59,93	196
1/8/2010	21	3,95	359	1/4/2015	22	8,92	170	1/12/2019	20	59,87	203
1/9/2010	20	3,96	350	1/5/2015	21	9,00	188	1/1/2020	23	60,28	219
1/10/2010	22	3,96	355	1/6/2015	20	9,10	187	1/2/2020	20	62,13	203
1/11/2010	22	3,99	340	1/7/2015	23	9,20	194	1/3/2020	22	64,31	152
1/12/2010	21	3,98	378	1/8/2015	21	9,30	173	1/4/2020	21	66,78	94
1/1/2011	21	4,01	385	1/9/2015	20	9,43	194	1/5/2020	21	68,47	117
1/2/2011	19	4,03	326	1/10/2015	22	9,52	230	1/6/2020	22	70,39	135
1/3/2011	23	4,05	343	1/11/2015	19	9,66	235	1/7/2020	22	72,31	158
1/4/2011	20	4,08	312	1/12/2015	22	12,94	260	1/8/2020	19	74,02	175
1/5/2011	20	4,09	357	1/1/2016	22	13,89	223	1/9/2020	22	76,17	184
1/6/2011	22	4,11	308	1/2/2016	19	15,83	257	1/10/2020	20	78,32	177
1/7/2011	21	4,15	325	1/3/2016	22	14,70	301	1/11/2020	23	81,30	181
1/8/2011	22	4,20	326	1/4/2016	21	14,29	283	1/12/2020	21	84,08	197
1/9/2011	21	4,21	313	1/5/2016	21	13,98	246	1/1/2021	22	87,30	227
1/10/2011	20	4,24	287	1/6/2016	20	15,04	258	1/2/2021	20	89,82	212
1/11/2011	22	4,28	321	1/7/2016	22	15,00	288	1/3/2021	22	91,98	226
1/12/2011	20	4,30	296	1/8/2016	22	14,93	257	1/4/2021	20	93,36	172
1/1/2012	22	4,33	308	1/9/2016	20	15,36	261	1/5/2021	21	94,49	185
1/2/2012	19	4,35	318	1/10/2016	21	15,17	246	1/6/2021	22	95,70	158
1/3/2012	21	4,38	345	1/11/2016	20	15,87	256	1/7/2021	21	96,68	167
1/4/2012	20	4,41	327	1/12/2016	22	15,87	256	1/8/2021	23	97,74	196
1/5/2012	22	4,47	390	1/1/2017	22	15,90	242	1/9/2021	20	98,73	178
1/6/2012	21	4,53	365	1/2/2017	19	15,48	254	1/10/2021	22	99,62	164
1/7/2012	23	4,59	360	1/3/2017	22	15,38	247	1/11/2021	20	100,93	196
1/8/2012	21	4,64	360	1/4/2017	21	15,39	237	1/12/2021	21	102,68	173
1/9/2012	20	4,70	342	1/5/2017	22	16,09	176	1/1/2022	20	105,00	285
1/10/2012	22	4,76	364	1/6/2017	20	16,62	234	1/2/2022	20	107,39	276
1/11/2012	21	4,84	350	1/7/2017	22	17,65	263	1/3/2022	22	111,00	306
1/12/2012	20	4,92	360	1/8/2017	21	17,34	235	1/4/2022	21	115,31	326
1/1/2013	23	4,98	344	1/9/2017	20	17,31	267	1/5/2022	20	120,19	290
1/2/2013	18	5,05	358	1/10/2017	22	17,64	213	1/6/2022	20	125,21	317
1/3/2013	22	5,12	397	1/11/2017	20	17,31	230	1/7/2022	22	131,26	286
1/4/2013	22	5,19	379	1/12/2017	22	18,61	203	1/8/2022	22	138,72	280
1/5/2013	20	5,28	418	1/1/2018	22	19,64	185	1/9/2022	21	147,31	282
1/6/2013	20	5,39	390	1/2/2018	20	20,12	214	1/10/2022	21	156,89	259
1/7/2013	23	5,51	390	1/3/2018	22	20,12	199	1/11/2022	18	167,28	212
1/8/2013	22	5,67	393	1/4/2018	19	20,53	191				

## Apéndice C

<b>DateTime</b>	<b>Expectativa Mensual</b>	<b>Costo proyectado (InsCod 1)</b>	<b>Costo proyectado (InsCod 22)</b>	<b>Costo proyectado (InsCod 1057)</b>
<b>1/3/2023</b>	0,71%	128,8	56,6	139,85
<b>1/4/2023</b>	0,61%	129,59	56,95	140,70
<b>1/5/2023</b>	0,34%	130,03	57,14	141,18
<b>1/6/2023</b>	0,33%	130,45	57,33	141,65
<b>1/7/2023</b>	0,42%	131,00	57,57	142,24
<b>1/8/2023</b>	0,47%	131,61	57,83	142,90
<b>1/9/2023</b>	0,52%	132,29	58,13	143,64
<b>1/10/2023</b>	0,50%	132,96	58,43	144,36
<b>1/11/2023</b>	0,50%	133,62	58,72	145,08
<b>1/12/2023</b>	0,49%	134,27	59,01	145,79
<b>1/1/2024</b>	0,48%	134,92	59,29	146,50
<b>1/2/2024</b>	0,46%	135,54	59,56	147,17
<b>1/3/2024</b>	0,43%	136,12	59,82	147,80
<b>1/4/2024</b>	0,41%	136,67	60,06	148,40
<b>1/5/2024</b>	0,39%	137,21	60,30	148,98
<b>1/6/2024</b>	0,37%	137,72	60,52	149,53

## Apéndice D

<b>DateTime</b>	<b>Expectativa Mensual</b>	<b>Costo proyectado (InsCod 1074)</b>	<b>Costo proyectado (InsCod 1086)</b>
<b>1/3/2023</b>	0,33%	12,13	11,7
<b>1/4/2023</b>	0,51%	12,19	11,76
<b>1/5/2023</b>	0,28%	12,23	11,79
<b>1/6/2023</b>	0,28%	12,26	11,82
<b>1/7/2023</b>	0,28%	12,29	11,86
<b>1/8/2023</b>	0,28%	12,33	11,89
<b>1/9/2023</b>	0,28%	12,36	11,92
<b>1/10/2023</b>	0,28%	12,40	11,96
<b>1/11/2023</b>	0,28%	12,43	11,99
<b>1/12/2023</b>	0,28%	12,46	12,02
<b>1/1/2024</b>	0,19%	12,49	12,05
<b>1/2/2024</b>	0,19%	12,51	12,07
<b>1/3/2024</b>	0,19%	12,54	12,09
<b>1/4/2024</b>	0,19%	12,56	12,11
<b>1/5/2024</b>	0,19%	12,58	12,14
<b>1/6/2024</b>	0,19%	12,61	12,16

## Apéndice E

```
param n := 12; #meses del año
param q := 5; #consumibles
param v := 1.1; # para calcular el beneficio de las compras.

#Conjunto de periodos
set T := {1 .. n}; #meses
set T0 := { 0 .. n };
set T1 := { -1 .. n };
set I := {1 .. q}; #cantidad consumibles
set Z := {1 .. 14}; #costo diferido
# Matriz de demanda (insumos * periodos)
param d[I*T] := read "demandas.txt" as "n+";
# Matriz de SS (insumos * periodos)
param w[I*T] := read "ss.txt" as "n+";
# Matriz de costos (insumos * periodos)
param c[I*T] := read "costo_compra.txt" as "n+";
# Matriz de costos (insumos * periodos)
param f[I*Z] := read "costo_compra_diferido.txt" as "n+";
# Abono en cada periodo
param e[T] := read "abono.txt" as "n+";
# Compra mínima
param m[I] := read "min_compra.txt" as "n+";

#Variables

# Variables compra, cantidad de insumo i a comprar en el periodo t
var x[I*T1] integer >= 0;

# Variables de stock, especifican la cantidad que queda en stock luego de cada periodo
```

```
var s[I*T0] >= 0;
```

#Variable binaria: indica 1 si se debe comprar insumo "i" en el mes "t", caso contrario indica 0.

```
var y[I*T1] binary;
```

#Variable h, indica la cantidad de dinero destinado a comprar insumos en el mes t.

```
var h[T] >= 0;
```

# Función objetivo

minimize obj:

```
sum <i,t> in I*T: c[i,t] * x[i,t];
```

# Ligando las variables de stock

subto ligado: forall <i,t> in I\*T:

```
s[i,t] == s[i,t-1] + x[i,t] - d[i,t];
```

# Stock inicial: Se presume que en el periodo 0 la empresa ya contaba con stock de seguridad.

subto inicial: forall <i,t> in I\*T0:

```
s[i,0] == w[i,1];
```

#Stock seguridad

subto ss: forall <i,t> in I\*T:

```
s[i,t] >= w[i,t] ;
```

# Cantidad comprada del insumo i en el periodo t-2

subto iniciali: forall <i,t> in I\*T1:

```
x[i,-1] == 0;
```

# Cantidad comprada del insumo i en el periodo t-1

subto iniciali0: forall <i,t> in I\*T0:

```
x[i,0] == 0;
```

# Ligando las variables de restricción de capital.

subto ligadob: forall <t> in T: sum <i> in I:

```
c[i,t]*x[i,t] == h[t];
```

subto ligadoc: forall <t> in T: sum <i> in I:

$$f[i,t]*x[i,t-2]*v - h[t] + e[t] \geq 0;$$

#Si hay compra en el mes, y[i,t] es igual a 1. Le indico que x[i,t] tiene que ser menor a un maximo de compra(número grande)

subto maxcompra: forall <i,t> in I\*T:

$$x[i,t] \leq 1000000 * y[i,t];$$

#Si hay compra en el mes, y[i,t] es igual a 1. Le indico que x[i,t] tiene que ser mayor o igual al mínimo de compra.

subto mincompra: forall <i,t> in I\*T:

$$x[i,t] \geq m[i] * y[i,t];$$