



Pontificia Universidad Católica Argentina

TRABAJO FINAL: Ingeniería Industrial

***Analítica Aplicada dentro de una Red
Logística: Nuevo Enfoque para la
Optimización de Inventarios***

Tutor	Alumno	Carrera
Ing. EMBA Carlos Raúl Arredondo, PhD.	Simón Pedro García Morillo 15-162144-6	Ingeniería Industrial

Índice

Resumen Ejecutivo	3
Diccionario	4
1 Planteo del Problema	5
2 Objetivos	7
3 Metodología	8
4 Marco Teórico	9
5 Desarrollo	15
5.1 Introducción	15
5.2 Supuestos	16
5.3 Análisis de Datos y Visualizaciones	19
5.4 De SKUs a GEPs	25
5.5 Parámetros Generales	29
5.6 Diseño de Red: Ubicación de los Centros de Distribución	33
5.6.1 Introducción	33
5.6.2 Modelo Matemático	36
5.7 Política de Inventarios Base	40
5.7.1 Introducción	40
5.7.2 Modelo Matemático	43
5.8 Metodología Alternativa: Nueva Política de Inventarios	47
5.8.1 Introducción	47
5.8.2 Modelo Matemático	54
6 Simulación	61
7 Análisis de Resultados	62
7.1 Planteo Pseudo Determinístico	62
7.2 Planteo Estocástico	64
8 Conclusión	67
9 Apéndice	69

Resumen Ejecutivo

En la actualidad, los mercados suelen caracterizarse por su competitividad. Cada vez, los clientes demandan un servicio más rápido, de mayor calidad y a un mejor precio; lo cual implica que las empresas deben comprometerse a mantenerse actualizadas y replantear sus estrategias de negocio, para permanecer relevantes.

En este contexto, el uso de herramientas de analítica avanzada es cada vez más relevante en la optimización de los procesos del negocio. Lo cual permite incrementar la rentabilidad de la empresa y agregarle valor al cliente; generando así, una ventaja competitiva que le permita destacarse.

Dada esta necesidad, se decidió analizar el caso de una empresa autopartista ficticia llamada “Accesorios Deportivos”. Esta compañía, distribuye una enorme variedad de productos a distintos clientes ubicados en todo el país. En el último año, a pesar de incrementar el volumen de ventas, la empresa reportó un aumento desmedido en el costo por transporte de inventario. Agregado a esto, la extensa variedad de productos ofrecidos y la demanda estocástica que percibe la empresa, la obligan a mantener un nivel de inventario desmedido, incurriendo así en un alto costo de almacenamiento.

Por esta razón, se propuso rediseñar la red logística e implementar una nueva política que permita disminuir los costos de transporte y almacenamiento de inventario; manteniendo el nivel de servicio requerido, con un tiempo de abastecimiento competitivo.

Para encontrar una solución óptima al problema, se entrenaron modelos predictivos para analizar y proyectar las ventas futuras, se aplicaron algoritmos de *clusterización* para agrupar los productos finales y se utilizó programación lineal para resolver los modelos matemáticos planteados para replicar el entorno de la red logística.

De esta manera, se logró disminuir en un 15.9%¹ el costo total anual incurrido por transporte y almacenamiento de inventario dentro de la nueva red logística. Lo cual refuerza la intuición de que el uso de modelos más sofisticados y el aprovechamiento de la información es necesario para encontrar soluciones más flexibles, con mejor rendimiento.

Por otra parte, dada la naturaleza pseudo determinística del planteo, se realizaron distintas simulaciones para validar los resultados obtenidos en los modelos matemáticos. Se buscó simular el costo “real” que se incurriría al aplicar dichas políticas de inventario, afrontando una demanda con cierto grado de aleatoriedad; simulando suficientes veces el problema para obtener resultados estadísticamente significativos.

No solo se buscó exponer los resultados de aplicar la metodología en la red planteada, sino que se hizo énfasis en introducir a las etapas necesarias para resolver un problema complejo y desestructurado; desglosándolo en distintas secciones relativamente simples, que puedan ser resueltas con nivel de abstracción razonable, utilizando herramientas de analítica avanzadas.

¹ Calculado según el promedio de los costos simulados en el capítulo 6.

Diccionario

A continuación, se realizará un pequeño listado de palabras y conceptos; cuyas abreviaciones serán usadas reiteradas veces, a lo largo del trabajo:

- Red logística: conjunto interconectado de fábricas y almacenes distribuidos en el país; que reciben pedidos de proveedores, almacenan inventario y satisfacen la demanda percibida de los clientes finales a los que abastecen.
- Nodo: puede ser cualquier centro de distribución (fábrica, almacén o proveedor), dentro de la red logística.
- Política de inventarios: se refiere a la definición tanto del nivel de inventario en cada nodo, cómo la frecuencia y traslado de inventario entre distintos nodos de la red logística.
- Nivel de servicio: proporción de pedidos de clientes finales que pueden ser satisfechos en tiempo y forma.
- CTE: Costo Total Esperado. Referido al costo anual esperado, tanto por transporte cómo almacenamiento de inventario.
- CTI: Costo de Transporte de Inventario. En donde se considera, tanto el costo de traslado de inventarios de proveedores, cómo traslado interno en la red.
- CAI: Costo de Almacenamiento de Inventario. Se incluye tanto el almacenamiento en fábricas, cómo en almacenes.
- SS: Inventario (o “Stock”) de Seguridad. Se refiere a la cantidad de inventario adicional que se guarda en los nodos para satisfacer distintos picos de demanda inesperada.
 - Naturalmente, el SS está íntegramente relacionado con el nivel de servicio deseado.
- SO: Inventario (o “Stock”) Objetivo: Se refiere a la cantidad de inventario máximo que se buscará tener, luego de realizar un pedido Q .
- LT: Tiempo de Transporte (o “Lead Time”). Referido al tiempo necesario para transportar inventario desde un nodo a otro.
- SKU: “Stock Keeping Unit”. El mismo es un código único que se usa para representar y distinguir cada producto final vendido.

1 Planteo del Problema

A continuación, se describirá la empresa ficticia sobre la cual se realizará el trabajo. Es decir, se detallarán los problemas y el comportamiento de la red logística que se usará cómo “instancia” para aplicar la metodología descrita en el *Resumen Ejecutivo*.

La empresa “Accesorios Deportivos” es una empresa autopartista que distribuye y comercializa accesorios para automóviles en todo el país. Actualmente, cuenta con un único “nodo” de distribución: una fábrica ubicada en San Justo, Buenos Aires, en las inmediaciones del aeropuerto internacional Ezeiza.

La totalidad de los productos importados son almacenados en un depósito en las inmediaciones de la fábrica; desde donde salen los camiones que entregan el producto a los consumidores finales (minoristas y particulares), en todo el país.

Naturalmente, las largas distancias llevan consigo un considerable tiempo de respuesta; razón lo la cual, han recibido continuos reclamos de clientes que amenazan con cambiar de proveedor. Además, a pesar de vender un producto altamente competitivo con demanda creciente, los elevados costos de transporte impactan negativamente en los márgenes de venta, lo cual se traduce en un declive en la renta de la empresa.

Agregado a esto, en función a un análisis de competencia, el área de marketing ha informado que es indispensable satisfacer al 80% de la demanda con tiempos de entrega de 24 hs., consolidando así una ventaja competitiva que les permita posicionarse debidamente en el mercado y crecer conforme a las siguientes proyecciones:

Estimación del crecimiento del mercado					
Año 1	Año 2	Año 3	Año 4	Año 5	Año 6 en adelante
3%	2%	2%	3.5%	3.50%	2%

Incremento del Market Share estimado					
Año 1	Año 2	Año 3	Año 4	Año 5	Año 6 en adelante
6%	5%	2%	0%	0%	0%

Nota: no se especificarán las fuentes de las imágenes, diagramas y/o gráficas que hayan sido realizadas por el autor del trabajo.

Por otra parte, la empresa ofrece una enorme variedad de SKUs distintos, lo cual la obliga a tener una cantidad excesiva de SS almacenado en la fábrica. En consecuencia, esto implica un alto el costo de almacenamiento y un considerable costo financiero de oportunidad, que se incurre al tener semejante dinero inmovilizado.

Dado el problema en cuestión, es evidente que la empresa necesita implementar cambios drásticos, tanto en el diseño de su red logística, cómo en la política de inventarios dentro de la misma. Esto será necesario para poder reducir los elevados CTI y CAI; manteniendo el nivel de servicio requerido (en este caso, 95%), y un tiempo de abastecimiento competitivo.

2 Objetivos

Cómo fue previamente explicado, la empresa *Accesorios Deportivos* sufre de altos costos que están directamente relacionados con una clara ineficiencia en el diseño y política de inventarios dentro de la red.

A su vez, esto es una problemática que toma particular significancia hoy en día, ya que el mercado tiende a ser cada vez más competitivo y los clientes tienden a demandar cada vez un servicio más rápido, de mejor calidad y a un mejor precio.

Es por esto, que el objetivo *general* del trabajo será:

- Diseñar una metodología robusta, basada en analítica avanzada, que permita reducir el costo total esperado por transporte y almacenamiento de inventario, abasteciendo una demanda estocástica y cumpliendo con los requerimientos mínimos de nivel de servicio y tiempo de abastecimiento.
- Implementar dicha metodología dentro de la empresa en cuestión, eligiendo herramientas analíticas gratuitas que cualquier persona pueda utilizar.

Luego, para lograr este objetivo general, se definen los siguientes objetivos *particulares*; los cuales serán necesarios para cumplir con el objetivo fundamental:

- Realizar un análisis preliminar de datos, para comprender mejor la situación de la empresa
- Diseñar una nueva configuración, agregando nodos dentro de la red logística, que permita cumplir con las necesidades de la empresa; logrando un tiempo de abastecimiento competitivo
- Determinar una política de inventario “base” sobre la cual se pueda calcular el CTE; para luego usarlo como un “benchmark” para comparar los resultados finales.
- Diseñar una nueva política en donde se defina el nivel de inventario y el itinerario de transporte óptimo que minimice el CTE de la empresa; logrando el nivel de servicio requerido
- Validar resultados con simulaciones

3 Metodología

En cuanto a la metodología, se puede distinguir claramente un tinte analítico a lo largo del trabajo. Esto se debe a que se buscará que cada método, cada modelo y cada conclusión que se obtenga de los mismos esté siempre fundamentado en con su respectivo análisis.

Agregado a esto, se utilizarán con frecuencia modelos matemáticos con distinto nivel de abstracción para “esculpir” el entorno sobre el cual se encuentra la red logística en la “realidad”. En consecuencia, se empleará frecuentemente una metodología conocida cómo *Programación Lineal* para resolver los modelos en donde el objetivo sea *optimizar* (maximizar, o reducir) cierta función objetivo que se haya planteado.

A su vez, se hará uso de modelos predictivos para analizar, explicar y proyectar la demanda a la que se enfrenta la red; y se aplicarán algoritmos de *clusterización* para agrupar los productos vendidos, en lo que se definirán cómo: *Grupos Equivalentes de Productos*. Además, una vez encontradas las políticas de inventario, se realizarán múltiples simulaciones para medir el rendimiento “real” de cada metodología.

Dada la complejidad del trabajo, se buscará acompañar las distintas secciones del desarrollo con sus correspondientes visualizaciones, para contribuir a un entendimiento más profundo en cada etapa de este.

Por otra parte, el objetivo del trabajo no es simplemente exponer los resultados de implementar distintos modelos matemáticos en la red logística elegida, sino que se busca que el lector logre comprender e incorporar distintas herramientas analíticas que puedan ayudarle en su carrera profesional. Razón por la cual, en la siguiente sección del trabajo se incorporará un *Marco Teórico*, en donde se introducirán, brevemente, los distintos conceptos teóricos clave, que serán de gran utilidad para desentrañar los próximos capítulos.

4 Marco Teórico

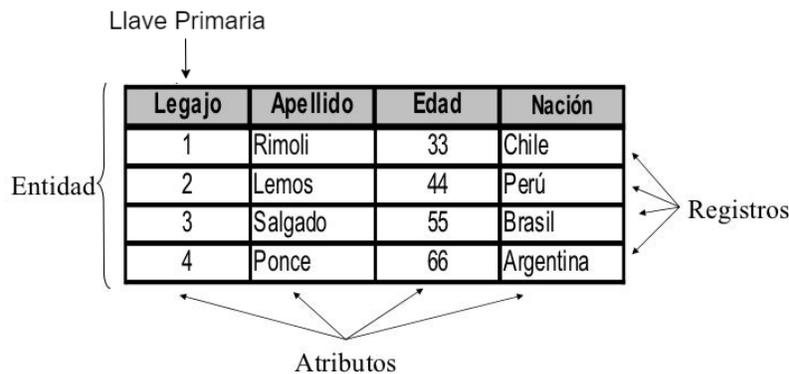
Dada la naturaleza analítica del trabajo, se dedicará esta sección a introducir algunos conceptos teóricos clave, sobre los cuales se construirá la mayor parte del desarrollo y conclusiones del trabajo.

Cabe resaltar que esta sección no será un conjunto de tutoriales profundos sobre los distintos temas, sino que serán más bien breves introducciones que permitan comprender la intuición y la utilidad de cada herramienta.

Diagrama Entidad Relación:

Un diagrama entidad-relación (o *DER*), es un tipo de diagrama de flujo que ilustra cómo distintas “entidades”, cómo personas, objetos o conceptos, se relacionen entre sí dentro de un sistema.

Más concretamente, los DER se suelen usar para diseñar, ilustrar, comprender y depurar las interrelaciones entre las tablas que conviven dentro de una base de datos relacional.



Fuente: cyta.com.ar

Nota: a lo largo del trabajo, se usarán las palabras atributo, columna y campo de manera indistinta; así como registro, fila y observación.

De esta forma, los DER suelen emplear un conjunto definido de símbolos, tales como rectángulos, diamantes, óvalos y líneas de conexión para representar la interconexión de entidades, relaciones y sus atributos.

A su vez, estas líneas de conexión permiten realizar uniones entre tablas, conectando las llaves primarias y foráneas de las mismas; lo cual permite crear y extraer tablas con mayor información.

Web Scrapping:

Web scrapping es una técnica utilizada para automatizar la extracción de datos de una, o muchas, páginas web.

La misma es particularmente conveniente para extraer datos cuando la página web no cuenta con una *api* (application programming interface) que facilite la extracción de los mismos, o cuando se quiere extraer información “desestructurada” que no se encuentre explícitamente en formato tabla.

Clusterización:

La clusterización es una técnica analítica basada en algoritmos de *Machine Learning* (aprendizaje automatizado) no supervisado; usado para encontrar patrones en los datos que permitan agrupar los mismo en grupos denominados *clusters*.

En otras palabras, es una técnica que permite encontrar regularidades en los datos, que permite asociar los mismos en grupos que naturalmente presentan características similares o que tienen un comportamiento similar.

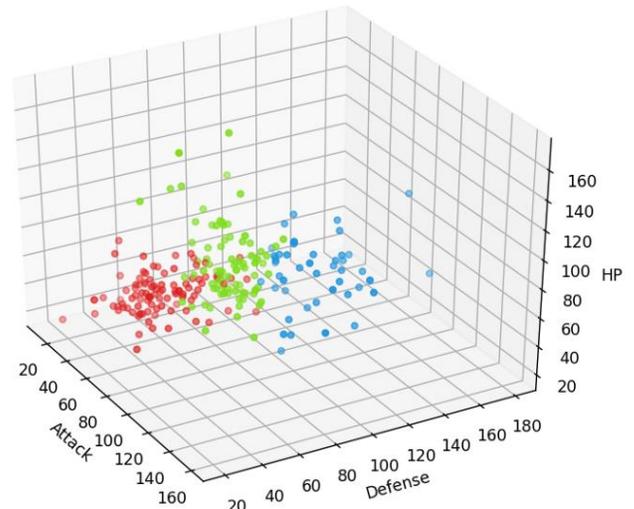
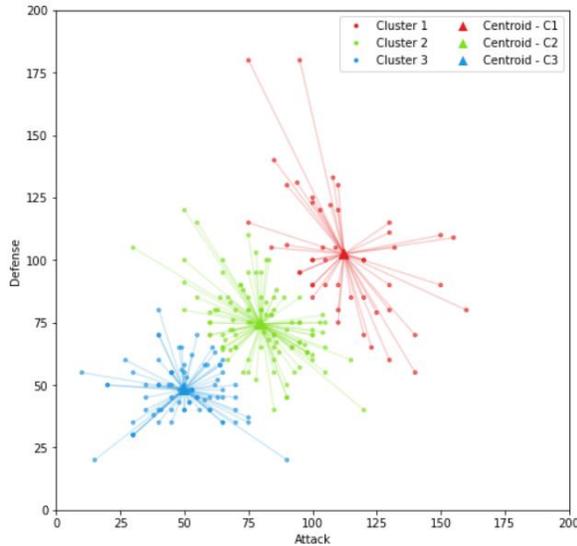
De esta manera, se busca asignar cada observación dentro de los datos ingresados a alguno de los k clusters creados; en donde el criterio para agrupar los datos es por su cercanía a ciertos puntos en el espacio llamados *centroides*.

Luego, el algoritmo busca minimizar una función de costo (J) que es la suma de las distancias al cuadrado de los puntos al centroide de su cluster; es decir, la varianza dentro de cada cluster:

$$J = \sum_{j=1}^k \sum_{\vec{x} \in S_j} \|\vec{x} - \vec{c}_j\|^2$$

en donde \vec{c}_j es el centroide del cluster S_j .

A continuación, se ilustra un ejemplo en donde se utiliza un algoritmo de clusterización para agrupar distintos *Pokemones* (creaturas ficticias creadas por la empresa Game Freak), según su poder de ataque, defensa y puntos de vida (HP):



Fuente: towardsdatascience.com (Autor: Thiago Carvalho)

Nota: se puede ver cómo a medida que incrementa la dimensionalidad del problema, los clusters se hacen cada vez más difíciles de visualizar.

Por otra parte, para poder elegir el número óptimo de clusters, es necesario puntuar la “validez” de los grupos creados, usando el *Silhouette Score*. El mismo mide cuán cercano es un punto al resto de los que están en su mismo cluster, en relación a cuán cercano es a los puntos del cluster más próximo. Es decir que para cada punto i , se tiene un coeficiente S_i :

$$S_i = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

en donde $a(i)$ es la distancia promedio a todos los otros puntos dentro del mismo cluster y $b(i)$ es la distancia promedio a todos los puntos del cluster vecino más cercano.

Finalmente, para obtener un score del clustering podemos promediar los coeficientes:

$$S(C) = \frac{1}{k} * \sum_{j=1}^k \bar{S}_j$$

en donde \bar{S}_j es el coeficiente promedio de todos los puntos dentro del cluster j -ésimo.

De esta forma, al considerar tanto las distancias entre los puntos internos de un mismo cluster, cómo las distancias entre los puntos con el resto de los clusters, el *silhouette score* ayuda a tener una mejor intuición sobre qué tan “densos” son cada uno de los clusters; y ayuda a ver qué tan bien separados están los clusters entre sí.

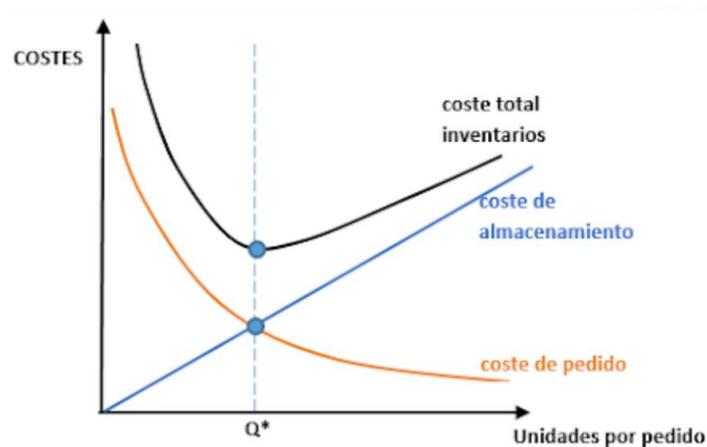
Además, el valor $S(C)$ está matemáticamente comprendido entre -1 y 1; donde un valor cercano a 1 nos dice que todos los clusters son muy densos y están bien separados;

mientras que un valor de 0 indica que puede haber un problema de superposición de clusters; y un valor de -1 indica que deben haber puntos asignados a clusters incorrectos.

Modelo de Wilson:

El modelo de Wilson es un modelo matemático relativamente sencillo, que permite calcular la cantidad o la frecuencia de pedido “óptima” que minimizan el *Costo Total Esperado*, conformado tanto por el costo de almacenamiento cómo el costo de pedido (o costo de transporte).

A continuación, se muestra un diagrama cualitativo que permite ilustrar el objetivo básico del modelo:



Fuente: econosublime.com

Nótese que existe un “trade-off” entre el **costo de pedido** (o transporte) y el **costo de almacenamiento**, en donde a mayor frecuencia de pedidos, se reduce la cantidad de inventario promedio en almacenamiento (y por lo tanto, su costo asociado de almacenamiento) pero, naturalmente, incrementa sustancialmente el costo de pedido.

Prophet:

Prophet es un modelo de pronóstico de series de tiempo, basado en un modelo aditivo en el que las tendencias no lineales se ajustan a la estacionalidad anual, semanal y diaria presentes en los datos. El mismo es particularmente útil para pronosticar series temporales con fuertes efectos estacionales y varias temporadas de datos históricos.

Nota: esta definición fue traducida de la descripción del modelo, realizada en facebook.github.io.

En su forma más simple, el modelo de Prophet puede ser desglosado de la siguiente manera:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

En donde:

- $g(t)$: es la función de tendencia
- $s(t)$: es la función de estacionalidad (anual, mensual, semanal, diaria)
- $h(t)$: es la función de efectos por feriados, días festivos y otros regresores
- ε_t : es el término del error donde, usualmente, se asume $\varepsilon_t \sim Normal$

A su vez, la *tendencia* suele ser de la forma:

$$g(t) = \frac{C}{1 + e^{-k*(t-m)}}$$

En donde:

- C: representa la capacidad
- k: representa la tasa de crecimiento
- m: representa la parámetro de compensación

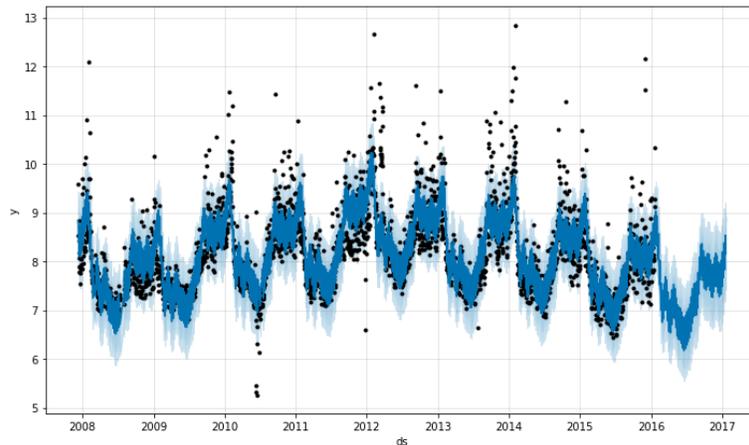
Y la *estacionalidad* (y feriados) suelen estar modelados aproximando la siguiente serie de *Fourier*:

$$s(t) = \sum_{n=1}^N \left(a_n * \cos\left(\frac{2\pi nt}{P}\right) + b_n * \sin\left(\frac{2\pi nt}{P}\right) \right)$$

En donde:

- P: representa el período de estacionalidad

En definitiva, el modelo no solo permite capturar en gran medida el comportamiento *explicativo* de los datos, sino que permite también realizar un pronóstico probabilístico de la serie de tiempo. Esta cualidad puede ser visualizada en el siguiente gráfico:



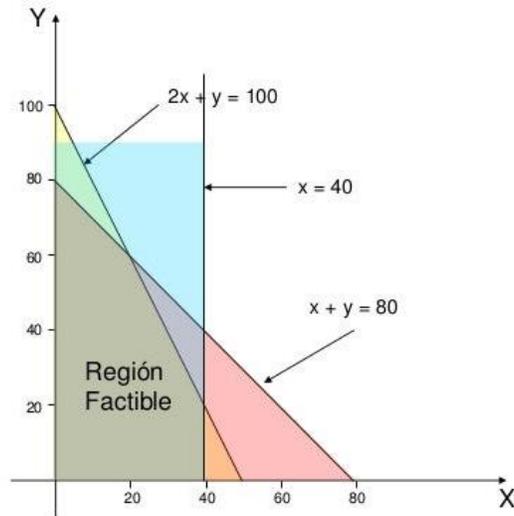
Fuente: facebook.github.io

Programación Lineal:

La programación lineal es una metodología de optimización dentro del campo de la programación matemática, en donde se busca maximizar o minimizar una *función objetivo*; la cual debe ser una función *lineal*. Es decir, debe ser una función donde todos los exponentes de las *variables decisión* sean estrictamente igual a 1.

Agregado a esto, la función objetivo suele estar acompañada de un conjunto de inecuaciones, también lineales, que representan las restricciones que deben cumplir las variables decisión dentro del problema.

A continuación, se muestra un diagrama explicativo en donde se puede ver la *región factible*, delimitada por ciertas inecuaciones (restricciones), sobre la cual se debe elegir un punto (generalmente un vértice de la región factible), en donde se maximice la función objetivo que pueda ser descripta:



Fuente: sites.google.com/site/elpetroleousil/geometria-analitica-1/programacion-lineal

Nótese que este ejemplo puede ser fácilmente diagramado, dado que solamente se cuenta con dos variables de decisión (x e y). No obstante, se debe recordar que para ejercicios con mayor complejidad donde incrementa la dimensionalidad del problema, este tipo de visualizaciones se vuelven prácticamente imposibles de realizar en un plano.

5 Desarrollo

5.1 Introducción

Luego de haber detallado el problema en cuestión, los objetivos a realizar, la metodología que se seguirá y el marco teórico que sustenta las herramientas analíticas a utilizar; se proseguirá a realizar el desarrollo del trabajo, en donde se describirá la evolución de los distintos pasos; desde la definición de los modelos hasta las soluciones de los mismos.

En concreto, el desarrollo estará subdividido en los siguientes capítulos:

- *Supuestos*: sección dedicada a detallar los múltiples supuestos que se asumirán para dar sustento a los modelos. Es importante notar que los mismos son necesarios para delimitar el problema en cuestión y, en cierto sentido, reducir la complejidad del mismo.
- *Análisis de Datos*: sección dedicada a definir, visualizar y comprender el comportamiento de ciertos componentes claves dentro del “sistema” (la red logística); cómo los proveedores, los productos vendidos y la demanda geográfica dentro del país. Agregado a esto, se aprovechará para verificar la “bondad” de los datos y realizar la correspondiente “limpieza” de los mismos.
- *De SKUs a GEPs*: sección dedicada a realizar la agrupación de productos en *clusters* de productos similares; y a justificar la intuición y el razonamiento detrás de la implementación de esta metodología.
- *Parámetros Generales*: sección dedicada a estimar ciertos parámetros que serán utilizados en los modelos matemáticos dentro de los siguientes capítulos.
- *Diseño de la Red Logística: Ubicación de los Centros de Distribución*: sección dedicada a la descripción de la eurística y el planteo del modelo matemático que se utilizará para encontrar la red logística sobre la cual se aplicarán las distintas políticas de inventarios.
- *Política de Inventarios Base*: sección dedicada al planteo y a la resolución del modelo matemático que será usado como línea “base” de comparación.
- *Metodología Alternativa: Nueva Política de Inventarios*: sección dedicada al diseño, planteo y resolución del modelo matemático alternativo que se propondrá; para encontrar una nueva política de inventarios que minimice el CTE dentro de la red logística. La misma estará subdividida en:
 - *Introducción*
 - *Traslado de Inventario entre Almacenes*
 - *Planeamiento de Pedidos Anticipados*
 - *Proyecciones de ventas*
 - *Modelo Matemático*

5.2 Supuestos

Antes de describir los supuestos, resulta pertinente resaltar las siguientes cuestiones:

- Por un lado, es necesario que los supuestos sean lo suficientemente robustos cómo para “retratar” correctamente la realidad en la que “vive” el sistema (la red logística)
- Por el otro lado, los supuestos sirven también para delimitar la “realidad” sobre la cual se construirán los modelos; recordando que nunca va a ser posible crear un entorno “hiper realista” y que siempre se deberá trabajar con algún nivel de abstracción, o simplificación de la realidad.

Dicho esto, se procederá a realizar un listado exhaustivo de los supuestos *generales*:

Red Logística:

- La misma estará compuesta por proveedores, una fábrica, almacenes y clientes finales ubicados en distintas ciudades del país.
- Cada semana, la fábrica debe realizar un pedido a los proveedores; quienes luego transportan las cajas pedidas.
 - Nota: no es necesario que se realicen pedidos todas las semanas
- Cada semana, los almacenes pueden realizar un pedido a la fábrica (o a otros almacenes); los cuales luego transportan las cajas pedidas
- Cada nodo de la red recibe pedidos diarios de los clientes finales.
- Dado que el tiempo de entrega al cliente es de gran interés para la empresa, cada *cliente* deberá ser abastecido por su nodo más cercano

Proveedores:

- Cada proveedor cuenta con una capacidad máxima de suministro anual, para cada GEP, igual al 50% de la demanda anual total de dicho GEP.
- La calidad es la misma para cada proveedor. Es decir, al cliente le es indistinto el producto de cualquier proveedor.

Fábricas:

- Se considerará una única fábrica ubicada en San Justo, Buenos Aires (pero los modelos matemáticos se plantearán para que puedan ser resueltos con f fábricas)
- Todos los artículos importados se almacenan inicialmente en un centro de distribución ubicado en la misma fábrica.
- Las fábricas tienen una capacidad de almacenamiento ilimitado, para cada GEP.

- Se ignora la capacidad productiva de la fábrica.

Almacenes:

- Si bien inicialmente no se encuentran almacenes dentro de la red, se asume que se podrán construir los suficientes almacenes para satisfacer los requisitos de tiempo de respuesta detallados en el capítulo 1 *Planteo del Problema*.
 - Se debe definir un criterio para decidir sobre la cantidad y ubicaciones idóneas para los mismos.
- Todos los almacenes construidos cuentan con una capacidad de almacenamiento ilimitado, para cada GEP.

Consumidor Final:

- Los consumidores finales son minoristas y/o particulares distribuidos en distintas ciudades de Argentina.
- Los mismos pueden realizar pedidos diarios a su nodo más cercano.
- La red llega hasta el cliente final, quien “consume” el stock.

Producto:

- Dado que solamente se tiene información sobre el precio de venta de cada producto, se asume que el precio de “compra” de cada producto es igual al 40% del precio de venta; es decir, se asume un margen constante de 60% para cada producto.
- Se asume que todos los productos dentro de un GEP tienen un peso igual al peso promedio de los productos de este.

Demanda:

- Dado que las decisiones sobre pedidos y transporte de inventario se realizan semanalmente, se estudiará la demanda semanal agregada, de cada GEP.
- La demanda semanal podrá tener tanto estacionalidad (semanas y mensual), cómo tendencia. Ambos parámetros deberán ser consideradas dentro del análisis.
- La demanda media crece según las estimaciones del mercado e incremento del market share; ambos especificados en el capítulo 1 *Planteo del Problema*
- Las demandas en las diferentes regiones de Argentina no tienen correlación entre sí.
- Las demandas de los distintos SKUs no tienen correlación entre sí.

Inventarios:

- Los inventarios pueden ser almacenados tanto en las fábricas como en los almacenes.
- Se debe idear la política de inventarios para reabastecer los mismos.
- Todas las semanas se realizarán las compras a proveedores y/o pedidos de transporte de inventarios dentro de la red
- Cada nodo debe tener, al menos, suficiente inventario almacenado para abastecer la demanda semanal proyectada que percibe (de los clientes finales)
- Dado que la demanda no es determinística, cada nodo deberá tener una cantidad de SS necesaria para lograr el nivel de servicio requerido. Si bien este podría ser un parámetro variable, en este caso se buscará un nivel de servicio del 95%.

Transporte:

- La fábrica puede realizar pedidos a los proveedores, pero posee un costo de envío fijo; e independiente de la cantidad ordenada.
- El transporte desde los proveedores hacia las fábricas se realiza en avión
- El transporte interno (dentro de la red) se realiza en camiones
- Se asume que todos los camiones se transportan a una velocidad promedio de 50 km/h
- El tiempo de transporte entre cada par de nodos se calcula como la velocidad promedio asumida, por la distancia entre ambos nodos.
- Se debe estimar el CTI entre cada par de nodos dentro de la red.
 - Este cálculo se detallará en el capítulo 5.5 *Parámetros Generales*
- Los conductores de camión pueden manejar hasta 10hs por día.
- Ya se cuenta con todos los camiones (y conductores) para realizar los viajes necesarios

Otras consideraciones:

- Se ignoran los costos fijos.
 - Es decir, solamente se trabajará con los CTI y CAI.
- Los costos serán considerados en dólares
- No se considerarán los efectos de inflación
- No se considerarán los impuestos
- Se busca generar una estrategia que minimice el CTE de manera sostenible (a largo plazo); no obstante, se asume repetitividad anual.

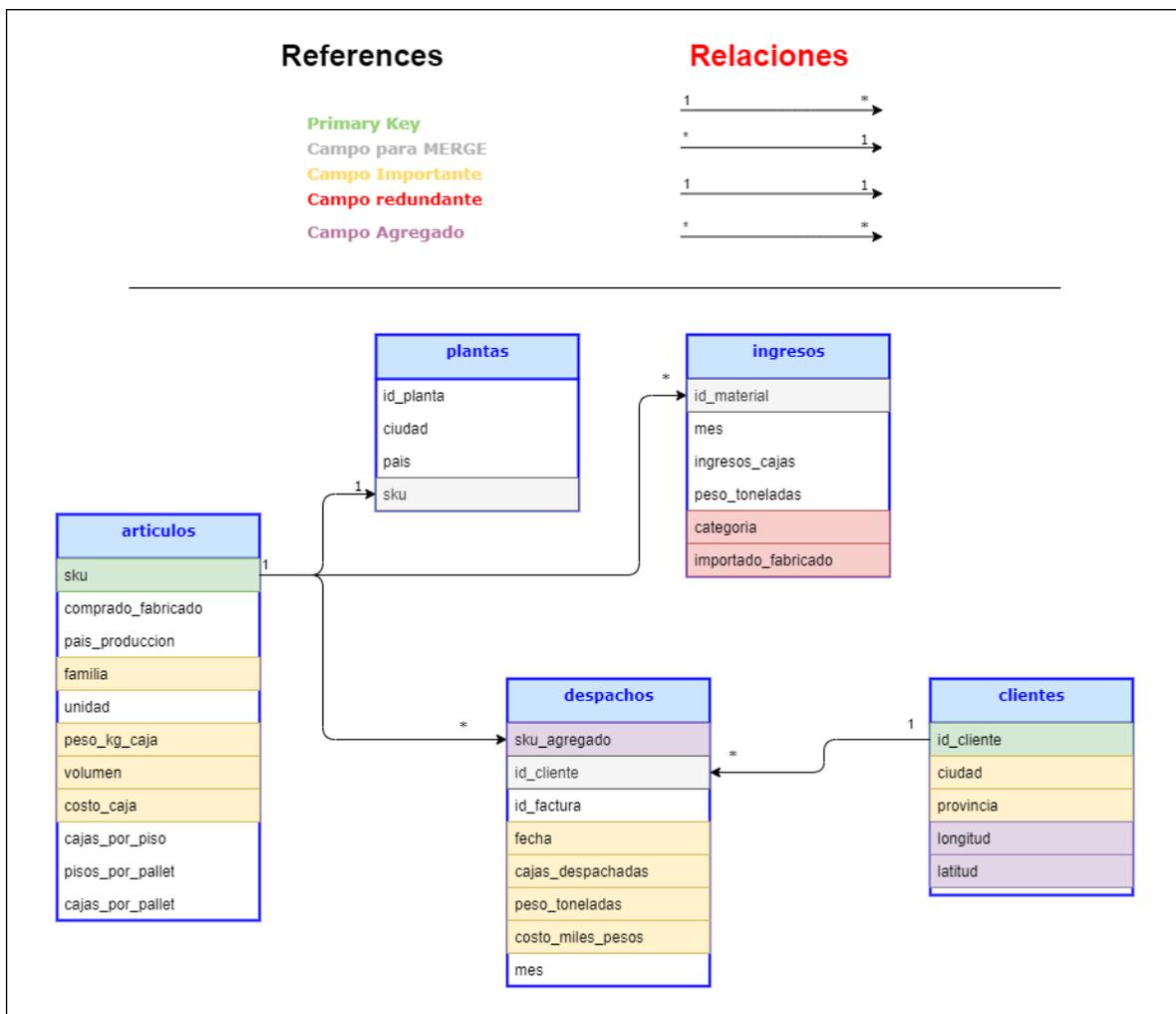
5.3 Análisis de Datos y Visualizaciones

Para iniciar el análisis, se construirá un *Diagrama de Entidad Relación (DER)*, para lograr visualizar y comprender las relaciones que poseen las tablas con las que estaremos trabajando.

Cómo primera observación, se puede notar que la tabla **despachos** no posee ningún atributo que se pueda relacionar con el SKU despachado. Esto resulta particularmente alarmante ya que, sin tener esta relación, solamente se podría analizar la demanda agregada de todos los SKUs juntos.

Es por esta razón que se buscará agregar el campo *sku_agregado*, usando la tabla **ingresos** para tener una noción de la distribución de SKUs.

*Nota: se asumirá que existe una correspondencia entre los **ingresos** y la demanda (contenida en **despachos**), tal que la distribución de la demanda por SKU sea la misma para ambas tablas*

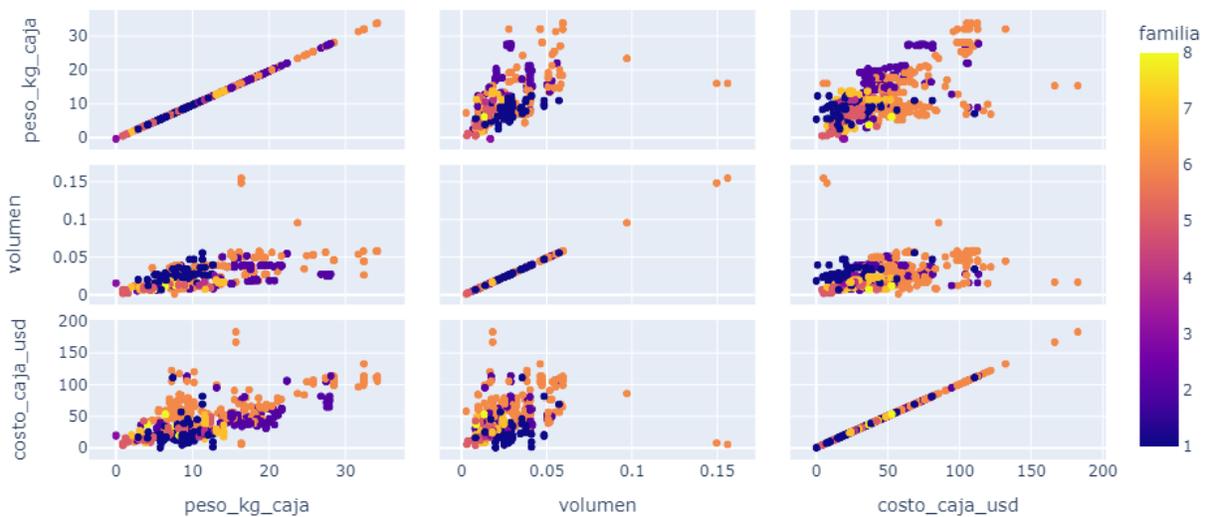


Para poblar el campo *sku_agregado*, será necesario mantener la misma distribución de ingresos que egresos, para cada SKU. Es decir que, a modo de ejemplo, si el 5% de los ingresos fueron del SKU 0001, entonces el 5% de los egresos también deberán ser del mismo SKU; para mantener la correspondencia.

Luego de haber agregado este atributo, se comenzará realizando el análisis de los productos vendidos. A continuación, se observa una matriz de diagramas de dispersión; la cual puede ser útil para visualizar posibles patrones o comportamientos particulares dentro de los productos vendidos.

Nota: inicialmente se cuenta con 730 SKUs distintos.

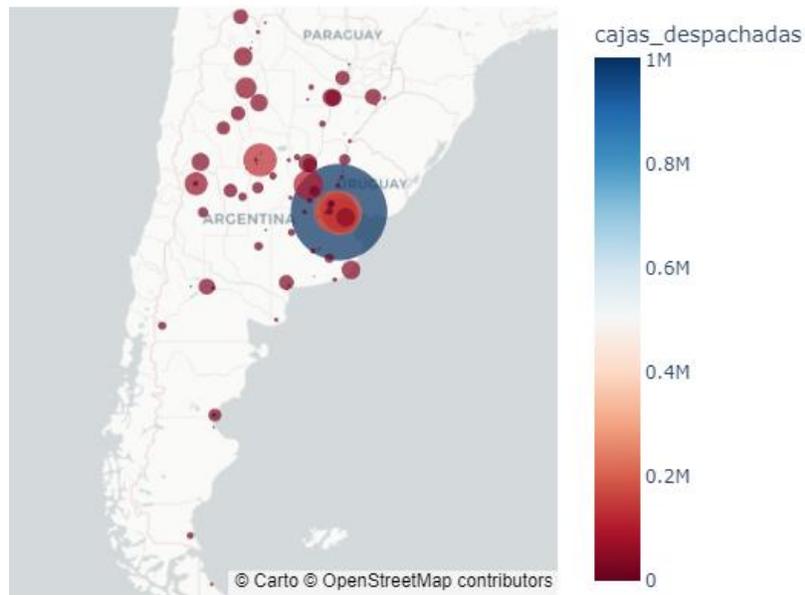
Matriz de Diagramas de Dispersión: Peso | Volumen | Costo



Si bien la figura no presenta patrones muy evidentes, sí nos da la intuición de que los SKUs, dentro de una misma familia de producto, parecen tener características similares entre sí.

Por otra parte, se realizará un análisis geográfico de la demanda; por lo cual, se hará uso de la librería **geopy** para extraer la *latitud* y la *longitud* de las ciudades donde residen los clientes.

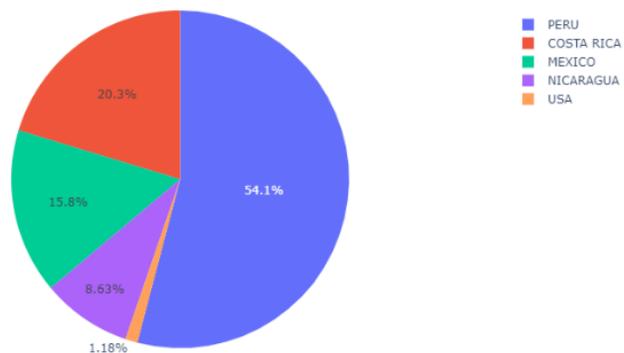
A continuación, se realiza un mapa de dispersión geográfico para visualizar tanto la ubicación como la cantidad de cajas demandadas a lo largo del país.



En donde se nota cómo a pesar de que la gran mayoría de las cajas son demandadas por clientes ubicados en Buenos Aires, sí será necesario construir otros centros de distribución (almacenes) en donde almacenar inventario y transportar cajas a los clientes finales del resto de las provincias.

En cuanto a los ingresos que recibe la empresa, se puede distinguir los ingresos y ubicación de los distintos proveedores; en donde cada sección representa la proporción de ingresos de cada proveedor, agrupados por país.

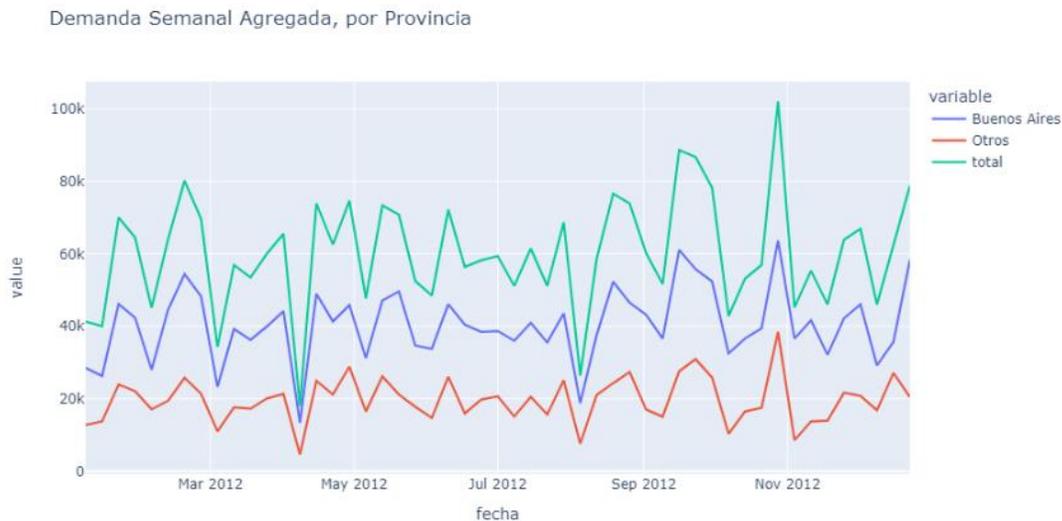
Ingresos por proveedor (externos)



Nota: se excluyen las cajas producidas en la planta

Por otra parte, será necesario llevar a cabo un análisis sobre la demanda percibida por los consumidores finales (contenida en la tabla de despachos),.

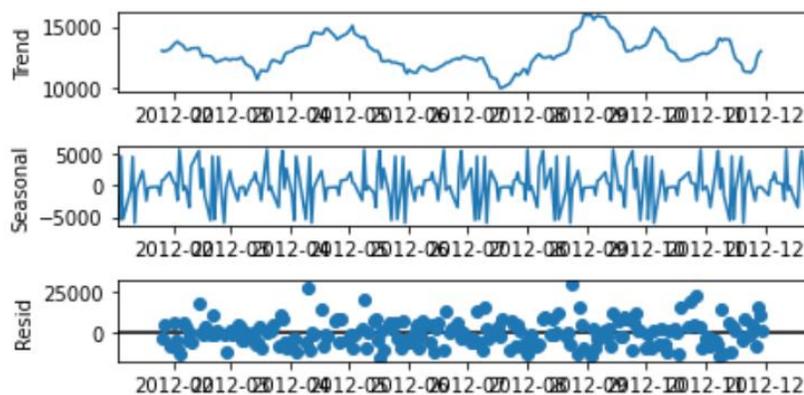
A continuación, se realizará un gráfico de línea para visualizar la demanda agregada “final” que deberá ser satisfecha por los nodos de la red. A pesar de contener los datos registrados con granularidad diaria, se visualizará la demanda de manera semanal, ya que el análisis final también será semanal.



La figura distingue la demanda de los clientes ubicados en la provincia de Buenos Aires, frente a la demanda de las demás provincias, dado que la enorme mayoría de las cajas son despachadas a clientes que residen en Buenos Aires.

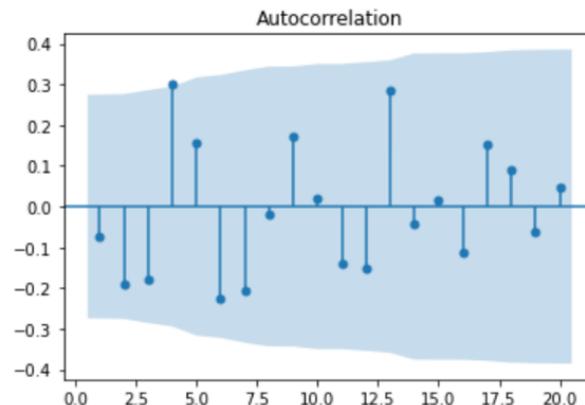
Es destacable cómo tanto la demanda total, cómo la demanda en Buenos Aires y la demanda en las distintas provincias presentan todas una tendencia y estacionalidad excepcionalmente similares; sugiriendo que se podría usar un modelo como *Prophet* para explicar y proyectar la demanda.

Luego, para ahondar sobre esta hipótesis, se realizará una descomposición estacional aditiva para distinguir los distintos componentes que presenta la serie de tiempo:



En definitiva, parece que la demanda agregada podría ser descompuesta en una componente de tendencia y una componente estacional que expliquen, con cierto grado de confianza, el comportamiento de la serie.

Por otro lado, se graficarán las autocorrelaciones de los egresos semanales, para confirmar la viabilidad de utilizar un modelo autorregresivo (en vez de un modelo estacional).



En definitiva, se puede ver que la demanda no presenta autocorrelaciones significativas en los distintos *lags*; con lo cual se descartará el uso de este tipo de modelos.

Finalmente, es necesario aclarar que se podría ignorar la tendencia y estacionalidad de la serie de tiempo; y continuar el análisis asumiendo una demanda con comportamiento normal (con observaciones independientes e idénticamente distribuidas).

No obstante, se argumenta que el uso de modelos predictivos tiene ventajas que justifican su uso:

1. Permite predecir la demanda futura, con cierto grado de certeza.
2. Contiene un componente explicativo de la demanda pasada, que puede ser usado para inferir ciertos descubrimientos que tengan valor de negocio, por fuera del proyecto en cuestión.

En definitiva, ayudan a enriquecer el análisis y, posiblemente, lograr resultados más precisos.

5.4 De SKUs a GEPs

Cómo fue brevemente mencionado en el capítulo 5.2 *Supuestos*, se asumirá que los SKUs podrán ser agrupados en diversos grupos, o *clusters* de productos similares; los cuales se denominarán *Grupos Equivalentes de Productos* (desde ahora en más, referidos como *GEP*). Luego, este capítulo estará destinado a ahondar sobre la necesidad de este supuesto; justificando la intuición y el razonamiento detrás del mismo.

Se comenzará entonces, desarrollando acerca de los beneficios de agrupar los SKUs:

1. *Reducción de la dimensionalidad del problema:*

La *Programación Lineal* es una metodología frecuentemente usada para resolver problemas de optimización, por su relativa facilidad de aplicación y por su efectividad para encontrar soluciones en un tiempo factible. Sin embargo, esta metodología cuenta con el factor de que a medida que se incrementan las *variables de decisión*, el tiempo de resolución puede incrementar desmedidamente.

Agregado a esto, resulta inviable intentar resolver el problema con 730 SKUs distintos, sin utilizar un software pago o un servidor con mayor capacidad computacional (también pago).

Luego, recordando que parte del objetivo *general* es resolver el problema usando herramientas de analítica gratuitas, resulta necesario agrupar los SKUs en grupos, reduciendo así sustancialmente las variables decisión y las restricciones necesarias.

2. *Mayor poder predictivo de demanda:*

Al realizar un análisis descriptivo de la demanda, es frecuente notar cómo a mayor granularidad y a menor nivel de agregación, el comportamiento de la variable se torna cada vez más errático e indistinguible de un proceso puramente aleatorio.

Por otra parte, cuando se incrementa el nivel de agregación es cuando ciertos patrones, cómo la tendencia y estacionalidad, se tornan cada vez más evidentes y, por ende, más fáciles de incorporar en un modelo.

Es por esta razón, que es sumamente más probable que un modelo a nivel GEP tenga un poder predictivo notablemente mayor que un conjunto de modelos predictivos a nivel SKU.

Luego de haber ahondado sobre la necesidad de este supuesto, parece razonable describir un ejemplo simple (hipotético) de cómo podría ser llevado a cabo en la práctica:

Asúmase el caso en que se agrupan tres tipos de alerones, sutilmente distintos (quizás de distintas gamas), en un mismo GEP; donde las ventas históricas de dichos alerones fueron 10.000 unidades del alerón A, 10.000 unidades del alerón B y 5.000 unidades del alerón C (es decir, 40% de las ventas del GEP fueron de A, 40% fueron de B y 20% fueron de C).

Luego, si se proyectaran unas ventas semanales (agregadas) de 100 unidades, entonces se realizaría un pedido de 40 unidades de A, 40 unidades de B y 20 unidades de C (es decir, se realizarían pedidos proporcionales a las ventas históricas relativas de los productos del GEP).

No obstante, cabe resaltar que el SS será calculado a nivel producto, asumiendo una demanda normal y aplicando la ecuación correspondiente:

$$SS_{as} = Z_{95\%} * \sigma_{as(T+L)}$$

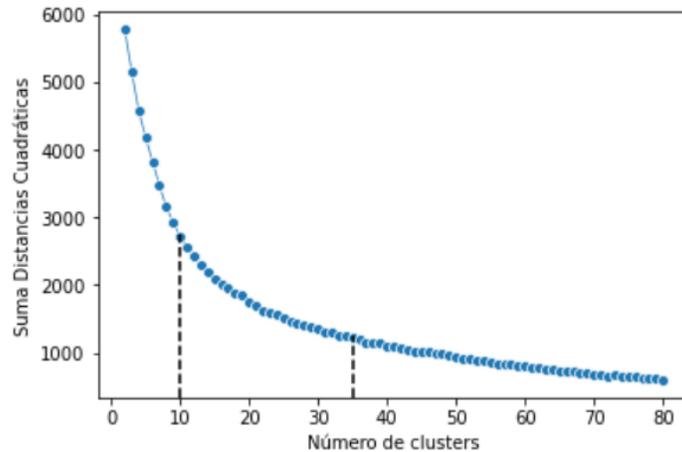
Nota: también podría asumir “intercambiabilidad” de los productos dentro de un mismo GEP (a los ojos del consumidor); justificando que los mismos tendrían un precio similar, características físicas similares y una demanda similar; tanto en magnitud cómo en estacionalidad.

En este caso, se podría calcular el SS a nivel GEP, lo cual reduciría la cantidad de SS agregada (dado que la varianza de la demanda agregada es menor a la suma de las varianzas de la demanda de cada producto)

A continuación, se especificará las características, o *features*, numéricas que se utilizarán para agrupar los productos:

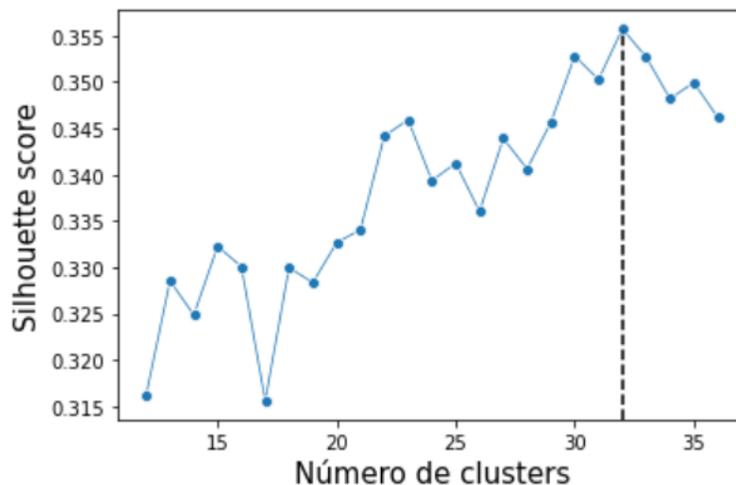
- Familia: familia de producto
- peso_kg_caja: peso de una “caja” de SKUs, medida en kg
- volumen: volumen de una “caja” de SKUs, medida en m³
- precio_caja_usd: precio al cual se vende una “caja” de SKUs
- ventas_trimestre_1_perc: el porcentaje de ventas que se realizan en el primer trimestre (respecto a las ventas totales)
- ventas_trimestre_2_perc: el porcentaje de ventas que se realizan en el segundo trimestre
- ventas_trimestre_3_perc: el porcentaje de ventas que se realizan en el tercer trimestre
- ventas_trimestre_4_perc: el porcentaje de ventas que se realizan en el cuarto trimestre
- ventas_totales: ventas totales (anuales) que se realizan; medida en “cajas”
- std_ventas_totales: desviación estándar de las ventas diarias

A continuación, se muestra un análisis del cambio en la suma de las distancias cuadráticas J , a medida que se aumenta el número de *clusters*; recordando que el número de clusters es un hiperparámetro que se debe fijar previo a “correr” el algoritmo:



Naturalmente, a medida que se incrementa el número de clusters n , la suma de distancias baja (siempre el mínimo estará dado cuando n sea igual a la cantidad de observaciones). De todas formas, se utilizará la “regla del codo”, la cual permite distinguir una secuencia de puntos en donde la disminución en las distancias cuadráticas comienza a disminuir; es decir, el punto en donde agregar un cluster más no se condice con un incremento sustancial en la varianza explicada.

Luego, una vez elegido el rango de clusters a evaluar, se calcula el *silhouette score* para determinar la “bondad” de cada cluster y lograr determinar el número “natural” de clusters que se debería usar:



Finalmente, obtenemos el número óptimo de clusters: 31. Es decir, los SKUs serán agrupados en 31 GEPs; en donde las características promedio de cada GEP están dadas por las siguientes tablas.

cluster	0	1	2	3	4	5	6	7	8	9	10
familia	7.000	6.000	2.000	1.000	2.000	6.000	2.000	1.000	2.000	2.000	5.000
peso_kg_caja	10.854	16.564	18.287	9.545	14.330	28.711	15.674	8.004	26.609	22.217	5.246
volumen	0.021	0.043	0.034	0.022	0.031	0.053	0.032	0.026	0.026	0.030	0.010
costo_caja_usd	29.531	51.174	48.874	20.733	42.844	102.158	40.941	20.926	62.416	52.925	19.811
ventas_trimestre_1_perc	0.000	0.221	0.000	0.126	0.000	0.236	0.000	0.000	1.000	0.532	0.087
ventas_trimestre_2_perc	0.000	0.258	0.000	0.242	0.000	0.266	0.000	0.000	0.000	0.389	0.178
ventas_trimestre_3_perc	0.000	0.301	0.000	0.297	0.000	0.270	1.000	0.000	0.000	0.059	0.477
ventas_trimestre_4_perc	0.000	0.220	0.000	0.335	1.000	0.125	0.000	0.000	0.000	0.020	0.259
ventas_totales	0.000	5196.973	0.000	25740.294	1054.480	3475.690	1939.706	0.000	1453.167	2366.429	5141.167
coeficiente_variacion	0.000	0.352	0.000	0.366	0.000	0.375	0.000	0.000	0.000	0.547	0.633

cluster	11	12	13	14	15	16	17	18	19	20	21
familia	2.000	6.000	1.000	6.000	6.000	6.000	6.000	2.000	6.000	5.000	2.000
peso_kg_caja	16.410	9.413	9.912	13.598	9.417	11.186	19.613	15.864	16.399	4.134	11.021
volumen	0.040	0.017	0.027	0.038	0.019	0.026	0.052	0.021	0.153	0.011	0.025
costo_caja_usd	38.588	26.440	12.281	52.705	41.843	44.388	57.115	45.888	6.623	15.934	25.810
ventas_trimestre_1_perc	0.248	0.288	0.293	0.037	0.000	0.000	0.000	0.216	0.000	0.000	1.000
ventas_trimestre_2_perc	0.277	0.206	0.358	0.074	0.000	1.000	0.000	0.280	0.000	0.000	0.000
ventas_trimestre_3_perc	0.170	0.262	0.189	0.710	0.000	0.000	0.079	0.309	0.000	0.000	0.000
ventas_trimestre_4_perc	0.304	0.245	0.160	0.180	1.000	0.000	0.921	0.195	0.000	0.000	0.000
ventas_totales	9532.611	4453.200	109726.500	4369.250	1022.130	2386.143	1075.545	5039.308	0.000	0.000	4271.333
coeficiente_variacion	0.326	0.344	0.606	0.955	0.000	0.000	0.035	0.336	0.000	0.000	0.000

cluster	22	23	24	25	26	27	28	29	30	31
familia	6.000	6.000	6.000	6.000	6.000	6.000	6.000	1.000	6.000	1.000
peso_kg_caja	15.700	14.183	8.466	7.780	9.717	15.760	8.690	7.239	25.501	8.500
volumen	0.019	0.042	0.026	0.014	0.019	0.046	0.018	0.028	0.051	0.025
costo_caja_usd	168.772	54.837	51.352	51.949	48.296	42.008	80.572	31.743	92.983	16.603
ventas_trimestre_1_perc	0.000	0.000	0.244	0.000	0.000	0.252	0.295	0.000	0.000	0.200
ventas_trimestre_2_perc	0.000	0.000	0.453	0.000	0.000	0.259	0.239	0.847	0.000	0.189
ventas_trimestre_3_perc	0.000	0.000	0.107	0.000	0.483	0.250	0.242	0.011	0.190	0.280
ventas_trimestre_4_perc	0.000	0.000	0.196	0.000	0.517	0.240	0.224	0.142	0.810	0.331
ventas_totales	0.000	0.000	7820.043	0.000	3377.077	36009.846	3166.368	7422.000	6772.600	73174.400
coeficiente_variacion	0.000	0.000	0.507	0.000	0.243	0.233	0.381	1.150	0.876	0.460

5.5 Parámetros Generales

Una vez encontrados los GEPs, se proseguirá a estimar los parámetros generales que serán compartidos en los modelos matemáticos que se plantearán a continuación. Es necesario aclarar, que algunos de los parámetros serán posibles de calcular con los datos contenidos en las tablas, mientras que otros deberán ser estimados usando información de distintas fuentes, o asumidos con algún criterio.

Cualquiera sea la metodología de cálculo, extracción o suposición de parámetros; se debe recordar que los mismos son simplemente “inputs” del modelo. Es decir que, si bien el valor de estos tendrá un claro impacto en el resultado final de los modelos, el verdadero valor del trabajo está en comprender y lograr desmenuzar la metodología y las herramientas analíticas usadas.

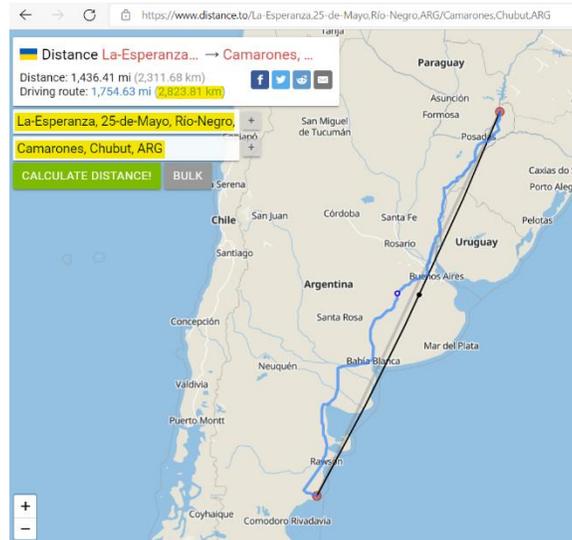
A continuación, se realiza el listado de los parámetros, con sus metodologías, cálculo y/o suposiciones:

- d_c : demanda anual agregada de los clientes en la ciudad c [cajas]
 - El mismo será una tabla calculada agrupando la demanda anual, por ciudad:

ubicacion	cajas_despachadas
Ciudad Autonoma de Buenos Aires, Buenos Aires,...	1007213
Avellaneda, Buenos Aires, Argentina	215940
3 de Febrero, Buenos Aires, Argentina	193587
La Matanza, Buenos Aires, Argentina	180132
Vicente Lopez, Buenos Aires, Argentina	149808
Florencio Varela, Buenos Aires, Argentina	140708
Córdoba, Córdoba, Argentina	126538
Rosario, Santa Fe, Argentina	101544
Mendoza, Mendoza, Argentina	62386
San Miguel de Tucumán, Tucumán, Argentina	52764

- dis_{fc} : distancia entre la fábrica f y la ciudad c
- dis_{fa} : distancia entre la fábrica f y el almacén a
- dis_{ac} : distancia entre el almacén a y la ciudad c
- $dis_{aa'}$: distancia entre el almacén a y el almacén a'
 - Las mismas serán obtenidas diseñando un *web-scrapper* que automatice el siguiente proceso:
 1. Entrar a la página: <https://www.distance.to/>
 2. Buscan la ciudad inicial (salida)
 3. Buscar la ciudad final (llegada)
 4. Extraer la distancia calculada en la página

5. Agregar la distancia a la matriz de distancias.



- PV_j : Precio de venta (a los consumidores finales)
 - Dado por el precio de venta promedio de cada GEP

- CV_j : Costo de venta: solo para la planta
 - Independiente de la cantidad (no hay descuento por cantidad)
 - Es igual a un 40% del precio de venta final

- H_j : Costo de mantenimiento de inventario

$$H_j = p_j * h = p_j * (c_{op} + h_c)$$
 - C_{op} : Costo financiero de oportunidad
 - Se tomará cómo el rendimiento actual anual de los bonos del tesoro americano: 2.15%
 - h_c : Holding Cost
 - a. El mismo depende de una pluralidad de factores, cómo:
 - i. Costos de almacén, costos de mano de obra, impuestos y otros
 - b. Dado que no se cuenta con ninguno de estos datos, se asume 15%

- w_j : peso promedio por caja de la familia de productos j [kg]
 - Dado por el peso promedio de cada cluster
- v : velocidad promedio de los camiones
 - Se asume 50 km/h

- It_{fa} : tiempo de recorrido (lead time) desde la fábrica f hasta el almacén a [días]
 - Dada la velocidad promedio de 50 km/h y el tiempo máximo de recorrido de 10 hs/día , se puede recorrer un máximo de 500 km/día
 - Luego, se calculan los días necesarios para transportar las cajas como el cociente entre la distancia entre la fábrica f hasta el almacén a (dis_{fa}) y 500 km/día (redondeado al mayor número entero)

- ct_{km} : costo de transporte por kilómetro

$$ct_{km} = p_{comb} * c_{comb} + c_{cond} + c_{amort}$$

- p_{comb} : precio de combustible.
 - Se asume $109 \text{ \$/l} \approx 1 \text{ usd/l}$ (ver: capítulo 9 *Apéndice*)
- c_{comb} : consumo de combustible.
 - Se asume $30 \text{ l/h} = 0,6 \text{ l/km}$ (ver: capítulo 9 *Apéndice*)
- c_{cond} : costo de conductor
 - Se asume un sueldo promedio de $102.638 \text{ \$/mes}$ (ver: capítulo 9 *Apéndice*)
 - Que conduce un total de 11.000 km/mes , en 22 días hábiles)
 - Tasa de cambio (actual): $109,20 \text{ usd/\$}$
 - $c_{cond} = 102.638 \frac{\$}{\text{mes}} * \frac{1}{109,20 \frac{\text{usd}}{\$} * 11.000 \frac{\text{km}}{\text{mes}}} = 0,0855 \frac{\text{usd}}{\text{km}}$
- c_{amort} : costo de amortización
 - Costo de camión: $6.480.000 \text{ \$}$ (ver: capítulo 9 *Apéndice*)
 - VU: 10 años
 - $11.000 \text{ km/mes} \rightarrow 1.320.000 \text{ km}$
 - $c_{amort} = 0,044 \text{ usd/km}$
- $c_{cond} = 6.480.000 \text{ \$} * \frac{1}{109,20 \frac{\text{usd}}{\$} * 1.320.000 \text{ km}} = 0,044 \frac{\text{usd}}{\text{km}}$

$$ct_{km} = 0,7295 \text{ usd/km}$$

- w_{mc} : peso máximo que podrá transportar el camión c [kg]
 - Se asume 52.000 kg (ver: capítulo 9 *Apéndice*)
- ct_{fa} : costo de transportar un camión (cargado con productos) desde la fábrica f , hasta el almacén a [usd]
 - *Surge de multiplicar el costo por kilómetro por la cantidad de kilómetros entre la fábrica y el almacén.*

- ct_{aa} : costo de transportar un camión (cargado con productos) desde el almacén a , hasta el almacén a' [usd]
 - Surge de multiplicar el costo por kilómetro por la cantidad de kilómetros entre ambos almacenes.
- s_{pj} : capacidad de suministro del proveedor p (anual) [cajas]
 - Se asume que la capacidad máxima de cada proveedor será igual al 50% de la demanda total (para cada GEP)
- k_{pfj} : costo de realizar un pedido al proveedor p , para la fábrica f , del GEP j [usd]
 - Parámetro que depende de los proveedores y del poder de negociación de la empresa.
 - Para este caso, se usará UPS para estimar los valores (ver: capítulo 9 Apéndice):

San-Justo, La-Matanza, Buenos-Aires, ARG

k [usd/pedido]	
peru	8311.30
costa_rica	10002.85
mexico	11056.35
nicaragua	10651.65
usa	13279.55

- lt_{pf} : tiempo de recorrido (lead time) desde el proveedor p , hasta la fábrica f [días]
 - Se asumen los siguientes tiempos:

San-Justo, La-Matanza, Buenos-Aires, ARG

lt [días]	
peru	2
costa_rica	3
mexico	3
nicaragua	3
usa	4

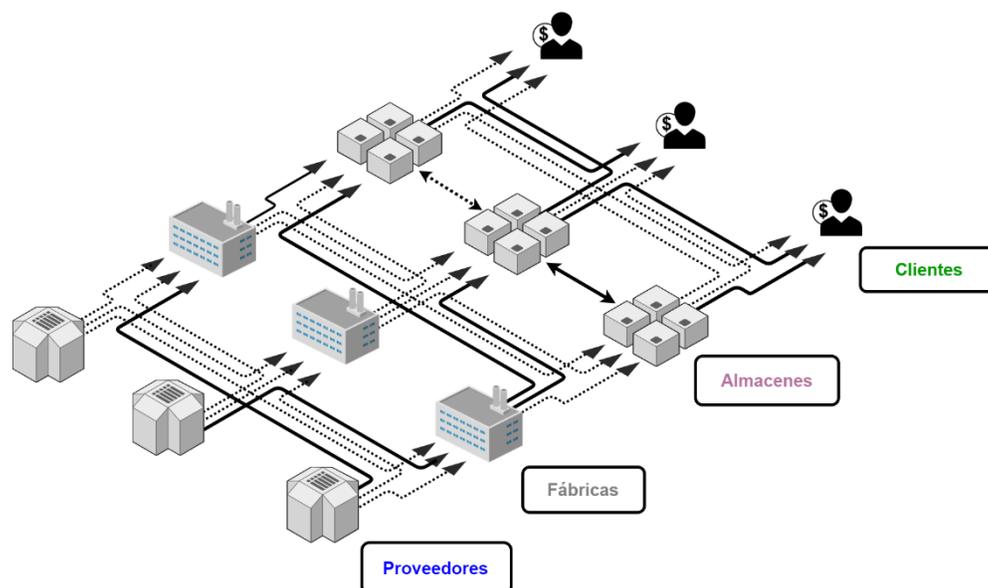
5.6 Diseño de Red: Ubicación de los Centros de Distribución

5.6.1 Introducción

Cómo fue previamente explicado en el capítulo 1 *Planteo del Problema*, actualmente la empresa cuenta con un único nodo (una fábrica ubicada en San Justo), que usa para recibir, almacenar y transportar inventario a toda la red. Esto, naturalmente, conlleva un elevado costo de transporte y un tiempo de respuesta sub-óptimo, que genera una notable pérdida en la rentabilidad de la empresa.

Por esta razón, se deberá conformar un nuevo diseño de la red logística, en donde se incorporen nuevos centros de almacenamiento y distribución de inventario: almacenes.

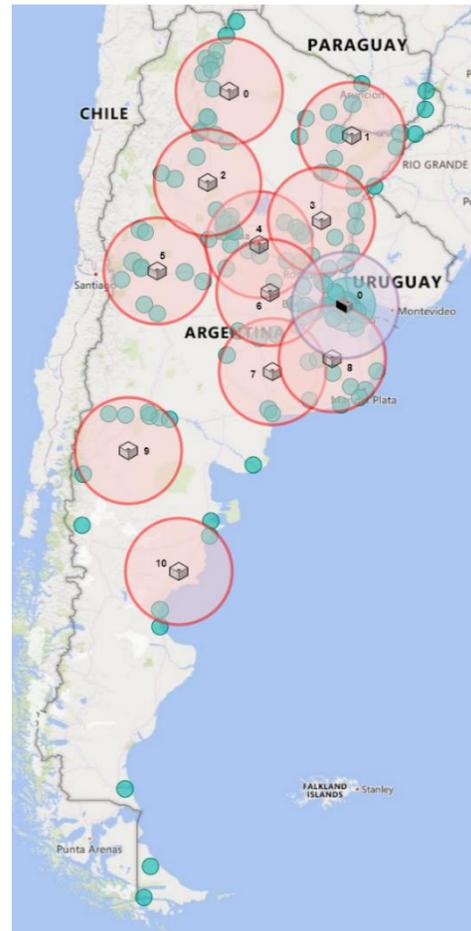
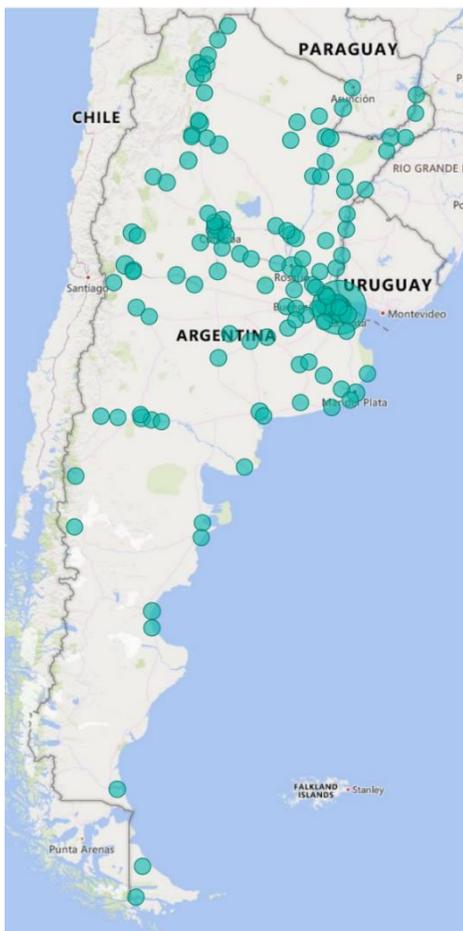
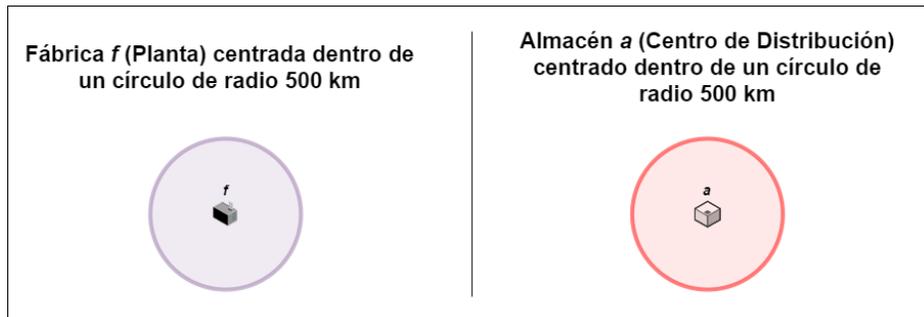
A continuación, se muestra un esquema cualitativo del problema, en donde se visualizan los distintos nodos de la red: proveedores, fábricas, almacenes y clientes finales; en donde las flechas punteadas representan los transportes posibles de inventario y las flechas rellenas representan los transportes reales:



Nota: se recuerda que el modelo matemático se planteará para f fábricas; a pesar de que ya se sabe que se trabajará con tan solo una fábrica, ubicada en San Justo. Esto se realiza para facilitar la replicabilidad del modelo para otras redes con mayor cantidad de fábricas.

Luego, se planteará la siguiente eurística; diseñada para construir la menor cantidad de almacenes posibles, en donde se logre satisfacer (al menos) al 80% de la demanda con tiempos de entrega de 24 horas o menos:

1. Realizar un mapa de dispersión geográfica de la demanda
2. Proponer distintas ubicaciones tentativas para los almacenes, intentando de que cada ubicación logre abastecer la mayor cantidad de demanda
3. Se realizará un modelo matemático que elija cuales de estos almacenes construir, considerando que se debe cubrir el 80% de la demanda



Nota: no se considerarán agregar Centros de Distribución que contengan a la demanda de las ciudades: Tartagal, San Ramón de la Nueva Orán, Puerto Iguazú, El Dorado, Oberá, Viedma, Esquel, Puerto Madryn, Caleta Olivia, Río Gallegos, Río Grande e Ushuaia. Esto se debe a que la demanda total de estas ciudades

tan solo representa al 0.9% del total; lo cual no justifica considerar otros Centros de Distribución extra.

Ubicaciones propuestas:

- 0: Argentina, Salta, Ntra Sra de Talavera
- 1: Argentina, Corrientes, Itaí
- 2: Argentina, Santiago del Estero, Guanaco Sombriana
- 3: Argentina, Entre Ríos, La Paz
- 4: Argentina, Córdoba, El Fortín
- 5: Argentina, Mendoza, La Dormida
- 6: Argentina, Santa Fe, Venado Tuerto
- 7: Argentina, Buenos Aires, La Colina
- 8: Argentina, Buenos Aires, Rauch
- 9: Argentina, Río Negro, La Esperanza
- 10: Argentina, Chubut, Camarones

5.6.2 Modelo Matemático

A Conjuntos y Parámetros

F: Conjunto de fábricas.

- $F = (0)$
- f : fábrica / $f \in F$

A: Conjunto de almacenes posibles.

- $A = (0, 1, 2 \dots 10)$
- a : almacén / $a \in A$

C: Conjunto de ciudades (donde se encuentran los clientes).

- $C = (0, 1, 2 \dots 129)$
- c : ciudad / $c \in M$
- dc : demanda anual agregada de los clientes en la ciudad c [cajas]
- $disfc$: distancia entre la fábrica f y la ciudad c
- $disac$: distancia entre el almacén a y la ciudad c

$$bfc = \begin{cases} 1, & \text{si la fábrica } f \text{ se encuentra a una distancia } d \leq d_{max} \text{ de la ciudad } c \\ 0, & \text{si la fábrica } f \text{ se encuentra a una distancia } d > d_{max} \text{ de la ciudad } c \end{cases}$$
$$bac = \begin{cases} 1, & \text{si el almacén } a \text{ se encuentra a una distancia } d \leq d_{max} \text{ de la ciudad } c \\ 0, & \text{si el almacén } a \text{ se encuentra a una distancia } d > d_{max} \text{ de la ciudad } c \end{cases}$$

- p : porcentaje de la demanda que debe ser cubierta en menos de 24 hs [%]

B Variables Decisión

Para cada $a \in A$, $c \in C$:

$$va = \begin{cases} 1, & \text{si se construirá el almacén } a \\ 0, & \text{si no se construirá el almacén } a \end{cases}$$

- $va \in$ variable binaria

xac = cantidad de cajas enviadas por el almacén a , hacia la ciudad c

- $xac \in$ variable entera

yfc = cantidad de cajas enviadas por la fábrica f , hacia la ciudad c

- $yfc \in$ variable entera

C Función Objetivo

Se diseñará con las siguientes prioridades:

1. Construir la menor cantidad de fábricas (para este caso, ya se cuenta con una fábrica construida y no se podrán construir más)
2. Construir la menor cantidad de Almacenes
3. Reducir el transporte de cajas

De manera tal que:

$$w1 \gg w2 \gg w3$$

$$Z = w1 * \sum_{f \in F} vf + w2 * \sum_{a \in A} va + w3 * \left(\sum_{f \in F} \sum_{c \in C} yfc * disfc + \sum_{a \in A} \sum_{c \in C} xac * disac \right)$$

*Nota: el factor $(\sum_{f \in F} \sum_{c \in C} yfc * disfc + \sum_{a \in A} \sum_{c \in C} xac * disac)$ no representa el costo de transporte, sino que es un factor diseñado para "incentivar" al modelo matemático a elegir una red que (en lo posible) tenga nodos cerca de las ciudades con mayor demanda.*

Nota: Se recuerda que este modelo matemático no está diseñado para construir una red óptima que minimice los costos totales, sino que es simplemente un método para obtener una red con nodos distribuidos de manera coherente, que sea lo suficientemente compleja como para justificar la implementación de políticas de inventario alternativas a la política de inventarios tradicional.

D Restricciones

- 1 La demanda anual de los clientes en cada ciudad debe ser satisfecha:

$$dc = \sum_{a \in A} xac + \sum_{f \in F} yfc \quad (\forall c \in C)$$

- 2 Cada almacén tiene que estar construido para poder abastecer algún mercado.

$$\sum_{c \in C} xac \leq va * 1000000 \quad (\forall a \in A)$$

Nota: el número 1000000 se elige cómo un número arbitrariamente grande

- 3 Cada fábrica tiene que estar construida para poder abastecer algún mercado.

$$\sum_{c \in C} y_{fc} \leq v_f * 100000 \quad (\forall f \in F)$$

Nota: para esta red en particular, se conoce que $v_f = 1$, ya que la fábrica ya se encuentra construida.

- 4 Se debe cubrir al menos un porcentaje p de la demanda en menos de un día.

Dicho de otra forma: debe haber al menos un porcentaje p de clientes que sean abastecidos por un almacén, o fábrica que se encuentre a una distancia menor o igual que la distancia máxima.

$$\sum_{a \in A} \sum_{c \in C} x_{ac} * b_{ac} + \sum_{a \in A} \sum_{c \in C} y_{fc} * b_{fc} \geq p * \sum_{c \in C} d_c$$

- 5 Se construirá, al menos, un total de 5 almacenes:

$$\sum_{a \in A} v_a \geq 5$$

Nota: esta restricción se agrega con el objetivo de trabajar con una red logística que sea lo suficientemente “compleja” y poder extraerle mayor provecho a todo el análisis realizado.

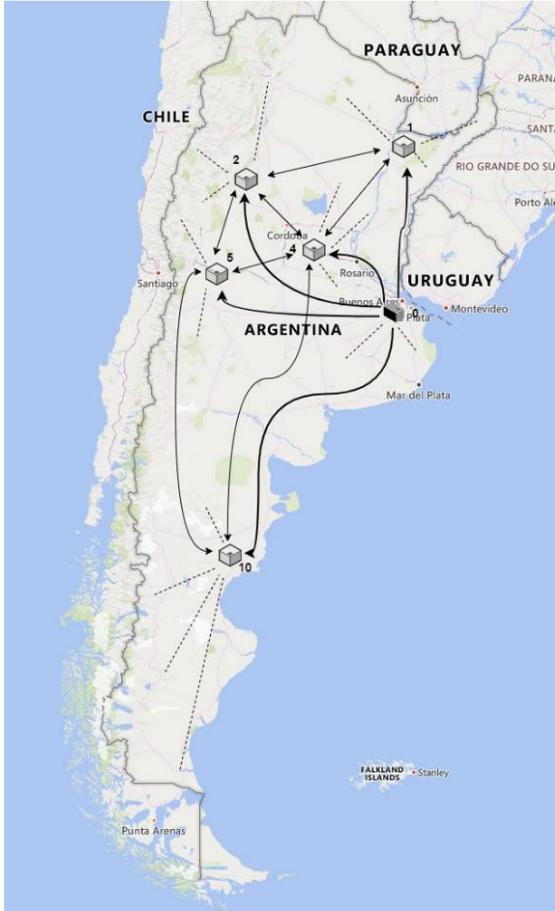
Solución:

Luego de realizar el planteo matemático del problema, se expondrá la solución encontrada, usando la metodología de *Programación Lineal*:

- Se construirá el almacén 1, 2, 4, 5 y 10
- Demanda total = 3066502
- Demanda satisfecha en $t \leq 24$ hs = 2887033
- Se satisfecerá un 94.15 % de la demanda en $t \leq 24$ hs

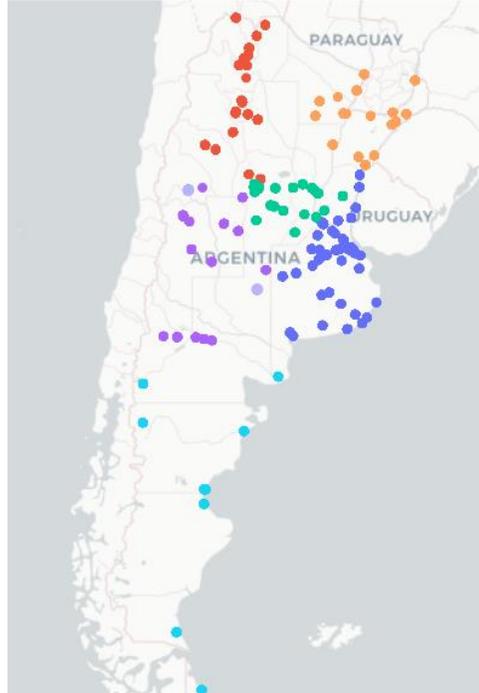
A continuación, se adhieren las siguientes figuras, en donde se visualizarán:

- Figura 1: diagrama cualitativo de la solución; en donde se observa la ubicación de los almacenes construídos
- Figura 2: dado que cada consumidor será satisfecho por su nodo más cercano, se presenta un mapa de dispersión geográfica, donde el color de cada punto representa el nodo de la red que satisfecerá a los clientes de dicha ciudad.



almacen

- San Justo, La Matanza, Buenos Aires, Argentina
- Guanaco Sombriana, Atamisqui, Santiago del Estero, Argentina
- El Fortín, San Justo, Córdoba, Argentina
- La Dormida, Santa Rosa, Mendoza, Argentina
- Itá Ibaté, General Paz, Corrientes, Argentina
- Camarones, Chubut, Argentina



5.7 Política de Inventarios Base

5.7.1 Introducción

Cómo primera aproximación, se deberá definir una política de inventario “base”, que sirva para determinar un nivel de inventario y frecuencia de transporte razonable dentro de la red. Luego, se calculará el CTE que se incurriría tomando estas decisiones, para obtener así un “benchmark”, que pueda ser utilizado posteriormente para determinar el ahorro potencial de la metodología alternativa.

En particular, se evaluarán distintos tiempos entre pedidos T (medida en semanas), para encontrar así la frecuencia global de la red que minimice el CTE, según las condiciones planteadas en el modelo.

Se trabajará con los siguientes supuestos “clásicos”:

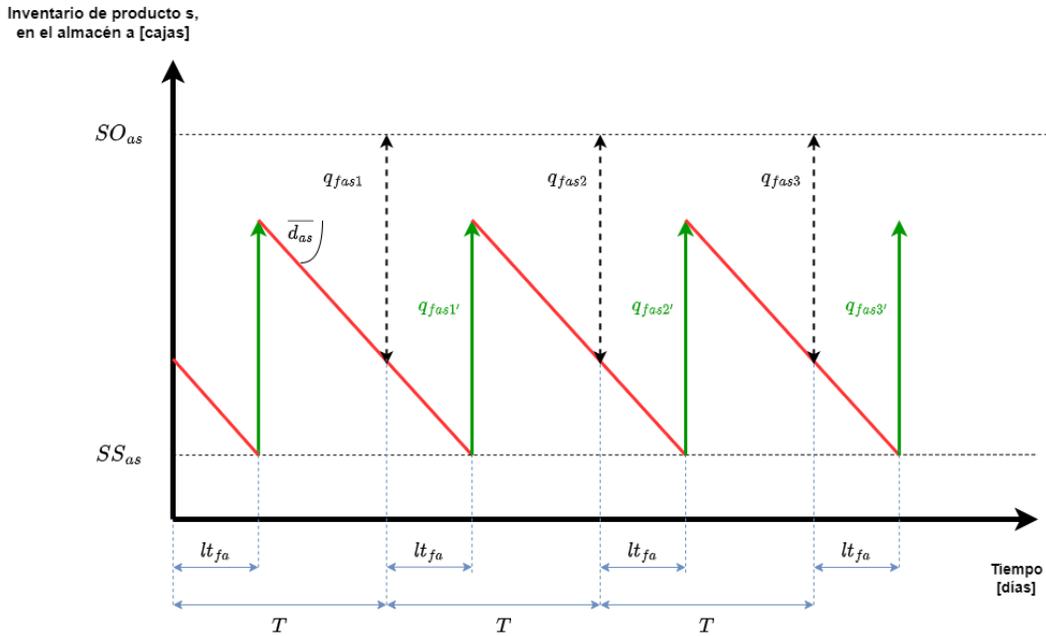
- SO y SS constante, para cada producto s y cada semana n :

$$SO_{sn} = SO, \quad \forall n \in N, \forall s \in S$$

$$SS_{sn} = SS, \quad \forall n \in N, \forall s \in S$$

- *Nota: tanto el SO como el SS se plantean a nivel producto, ya que la agrupación a nivel GEP (descrita en el capítulo 5.4) se reserva para el planteo de la metodología alternativa.*
- La demanda semanal (d_{sn}) tiene un comportamiento normal:
$$d_{sn} = d_s \sim N(\mu = \bar{d}_s, \sigma = \sigma_{ds}), \quad \forall s \in S$$
 - *Nota: inicialmente se trabajará asumiendo una demanda semanal normal e independiente entre sí (es decir, se ignorarán los efectos de tendencia y estacionalidad).*
- Cada nodo puede ser abastecido por un único nodo
- Todos los nodos de la red tienen tiempo entre pedidos (T).

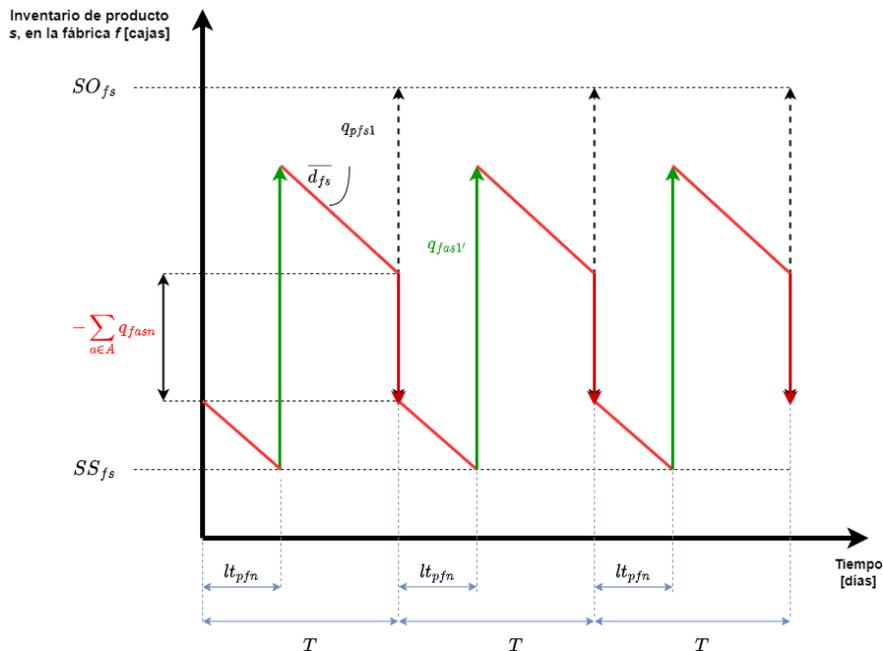
Luego, la situación de cada almacén (abastecido semanalmente por una fábrica), se puede representar usando el siguiente diagrama (cualitativo), en donde se ilustra la evolución de inventarios dentro de un almacén a , que recibe pedidos q_{fasn} de la fábrica f y percibe una demanda media \bar{d}_{as} , del producto s :



Nota: el parámetro L_{fa} representa el tiempo de transporte desde la fábrica f al almacén a

Por otra parte, se debe evaluar el nivel de inventario dentro de cada fábrica f , la cual debe abastecer a los consumidores finales cercanos; y debe también enviar los pedidos necesarios a los almacenes distribuidos por el país.

Para ilustrar dicho comportamiento, se muestra el siguiente diagrama en donde se representa la evolución de inventarios dentro de una fábrica f , que recibe pedidos q_{pfsn} del proveedor p y percibe una demanda media \bar{d}_{fs} , del producto s :



Nota: además de percibir la demanda media $\overline{d_{fs}}$, también debe satisfacer los pedidos de los almacenes; razón por la cual hay una bajada abrupta de inventario cada T días.

Es necesario aclarar, que este tipo de modelo relativamente sencillo, puede ser de gran utilidad para resolver problemas de políticas de inventario en situaciones simplificadas, cómo un único nodo, que vende un único producto y se enfrenta a una demanda con comportamiento constante a lo largo del tiempo.

Sin embargo, cuando se intenta aplicar el modelo a un conjunto de nodos interrelacionados, que se enfrentan a una demanda con comportamiento estacional, se pone en evidencia algunos sesgos y sobre simplificaciones que pueden disminuir notablemente el rendimiento de éste.

En particular, se ponen en evidencia las siguientes *limitaciones*:

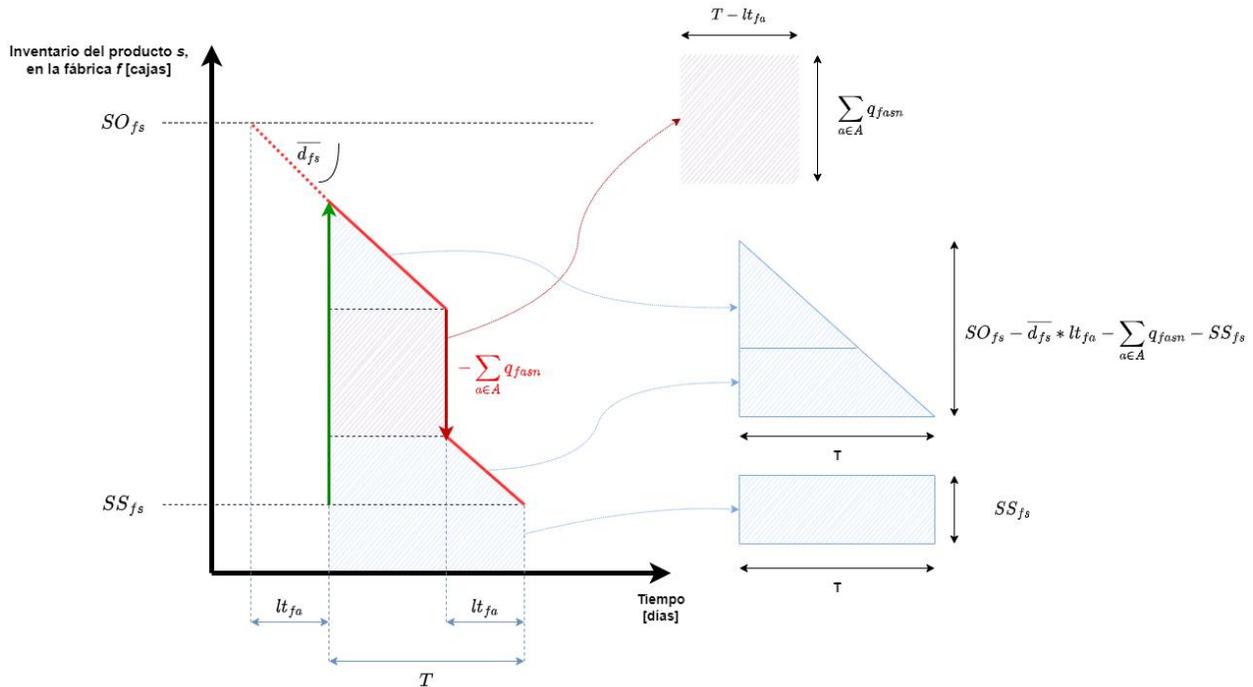
1. *Cada nodo de la red debe ser abastecido por un “proveedor” previamente fijado:*
Esto, a su vez, obliga a forzar la siguiente eurística:
 - a. Las fábricas priorizarán recibir pedidos de su proveedor más cercano, cambiando de proveedor cuando se supere la capacidad de suministro del mismo.
 - b. Los almacenes realizarán los pedidos a su fábrica más cercana
2. No se permite el planeamiento anticipado de pedidos
3. No se permite trabajar con proyecciones de la demanda, en donde se pueda considerar el efecto de estacionalidades y tendencias.
4. La frecuencia de pedidos tiene que ser la misma para todos los nodos de la red.

Agregado a esto, por naturaleza, el modelo intentará encontrar el mínimo *local* para cada nodo; es decir, se intenta minimizar el CTE de cada nodo por separado, sin considerar cómo podría afectar en el CTE de la red en conjunto.

Finalmente, es por esta razón, que no se pretendrá obtener un resultado óptimo para el problema, sino que se usará el resultado cómo el “benchmark” (base de comparación), sobre la cual se compararán los resultados de la nueva metodología que se detallará en el siguiente capítulo.

5.7.2 Modelo Matemático

Previo a realizar un listado de los *parámetros* y *fórmulas* que serán utilizadas, se ilustra el siguiente diagrama, que permitirá ayudar a justificar el cálculo del CAI para una fábrica f .



Parámetros:

- H_s : costo de mantenimiento de inventario, para el producto s
- ct_{fa} : costo de transporte desde la fábrica f , hacia el almacén a
- $\overline{ct_{pfs}}$: promedio ponderado del costo de transporte desde los distintos proveedores hasta la fábrica f , para el producto s
- σ_{das} : desviación standard de la demanda diaria, para el almacén a y el producto s
- σ_{dfs} : desviación standard de la demanda diaria, para la fábrica f y el producto s
- $\overline{d_{as}}$: demanda diaria promedio, para el almacén a y el producto s
- $\overline{d_{fs}}$: demanda diaria promedio, para la fábrica f y el producto s
- I_{asn} : inventario inicial en la semana n , para el almacén a y el producto s
- I_{fsn} : inventario inicial en la semana n , para la fábrica f y el producto s

- w_m : peso máximo que puede transportar el camión de carga
- T : tiempo entre pedidos
- L_a : tiempo de transporte que percibe el almacén a
 - Se asume como el tiempo de transporte entre cada almacén y su fábrica más cercana
- L_f : tiempo de transporte que percibe la fábrica f
 - Se asume como el promedio ponderado del tiempo de transporte para los distintos proveedores hasta la fábrica f

Fórmulas:

Para cada almacén a y para cada producto s , se calcula:

- Stock de Seguridad (SS_{as}):

$$SS_{as} = Z_{95\%} * \sigma_{as(T+L)}$$

$$\sigma_{as(T+L)} = \sqrt{(T + L_a)} * \sigma_{das}$$

- Stock Objetivo (SO_{as}):

$$SO_{as} = \bar{d}_{as} * (T + L_a) + SS_{as}$$

- Pedido n (q_n):

$$q_{asn} = SO_{as} - I_{asn}$$

- Pedido n (q): caso determinístico

$$q_{as} = SO_{as} - \bar{d}_{as} * L_a - SS_{as}$$

- Cantidad de camiones usados para transportar las cajas pedidas hacia el almacén: nc_a
 - El mismo será obtenido cómo el número entero mínimo que satisfaga:

$$nc_a * w_m \geq \sum_{s \in S} (q_{as} * w_s)$$

Finalmente, se calcula el *Costo Total Esperado* (CTE_A) anual:

$$CTE_A = \sum_{a \in A} \sum_{s \in S} \left(\frac{q_{as}}{2} * H_s + SS_{as} * H_s \right) + 54 * nc_a * ct_{fa}$$

*Nota, se ignora el término: $CV_s * \bar{d}_{as} * 365$, porque el mismo resulta (casi) independiente de la política de inventario dentro de la red*

Luego, para cada fábrica f (en este caso, solamente una) y para cada producto s , se calcula:

- Stock de Seguridad (SS_{fs}):

$$SS_{fs} = Z_{95\%} * \sigma_{fs(T+L)}$$

$$\sigma_{fs(T+L)} = \sqrt{(T_f + L_f) * \sigma_{dfs}}$$

- Stock Objetivo (SO_{fs}):

$$SO_{fs} = \overline{d_{fs}} * (T_f + L_f) + \sum_{a \in A} q_{fas} + SS_{fs}$$

- Pedido n (q_n):

$$q_{fsn} = SO_{fs} - I_{fsn}$$

- Pedido n (q): caso determinístico

$$q_{fs} = SO_{fs} - \overline{d_{fs}} * L_f - SS_{fs}$$

- Costo Total Esperado (fábricas):

$$CTE_f = \sum_{j \in J} 54 * \overline{ct_{pfs}} + \sum_{f \in F} \sum_{s \in S} H_s * \left(\frac{q_{fs} - \sum_{a \in A} q_{fas}}{2} + \sum_{a \in A} q_{fas} * \frac{T - L}{T} + SS_{fs} \right)$$

- Costo Total:

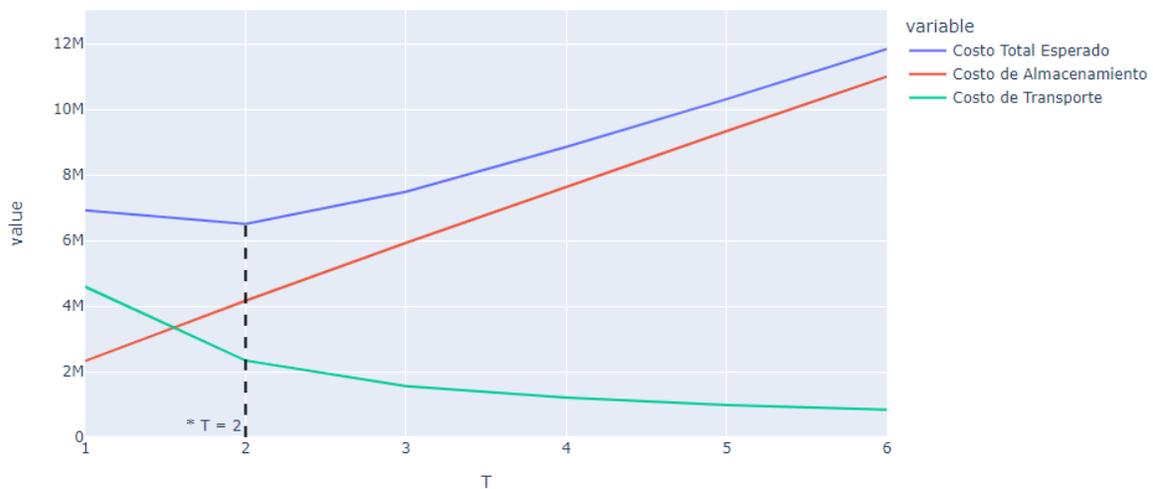
$$CTE = CTE_F + CTE_A$$

Solución:

Cómo fue previamente mencionado, se evaluarán distintos tiempos entre pedidos, T , cómo alternativas para encontrar aquel que minimice el CTE, calculado con el modelo matemático descrito.

A continuación, se muestra la evolución del CTE, respecto del T (donde, a su vez, se desglosa tanto el Costo de Almacenamiento cómo el Costo de Transporte por separado):

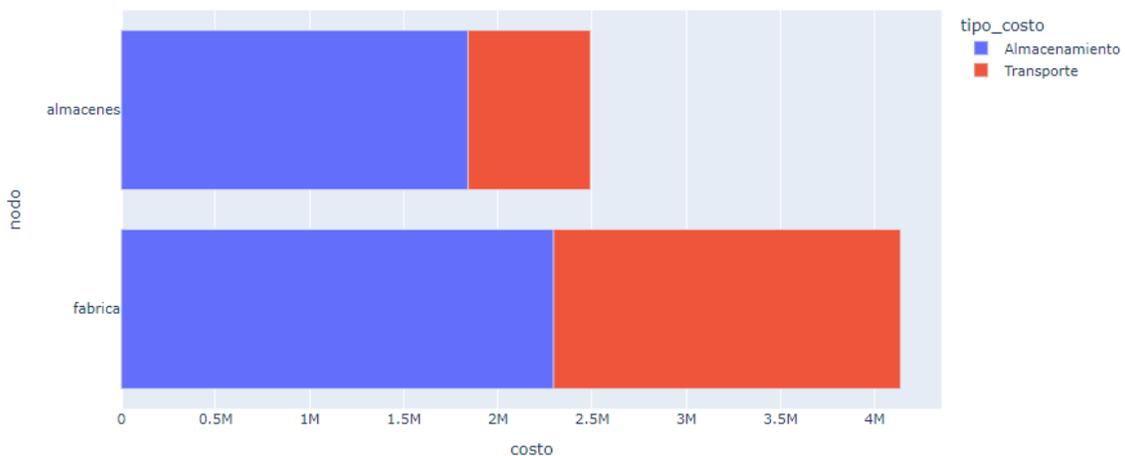
Costo Total Esperado, por Tiempo entre Pedidos



Aquí se puede observar cómo el tiempo entre pedidos “óptimo” para el modelo en cuestión es $T = 2$; es decir, realizando pedidos cada dos semanas (en todos los nodos) se logra un CTE = 6.629.212.1 usd/año; el cual será usado cómo *costo base*.

Luego, se realiza un desglose de los costos encontrados:

Costo Total Esperado, por tipo de costo (T = 14 días)



5.8 Metodología Alternativa: Nueva Política de Inventarios

5.8.1 Introducción

En el capítulo 5.7 *Política de Inventarios Base*, se mencionó cómo el modelo usado resulta ser un modelo relativamente sencillo, que puede ser utilizado para encontrar una política de inventario aceptable para un nodo dentro de la red. Sin embargo, el mismo tiene la limitación de estar sesgado a buscar mínimos *locales* del problema; lo cual no necesariamente implica encontrar el mínimo *global* para la red.

Agregado a esto, se hizo mención de las suposiciones propias del modelo; y de cómo estas pueden implicar una sobre simplificación de la realidad, limitando así la efectividad del modelo.

Es por esta razón, que se propondrá una nueva metodología que pueda ser utilizada para encontrar una política de inventarios más robusta; en donde se logre eludir las limitaciones previamente mencionadas, encontrando así soluciones más “flexibles” y coherentes con el entorno en el que “vive” la red.

A continuación, se detallarán las distintas optimizaciones que se implementarán en el modelo matemático final:

1) Traslado de Inventario entre Almacenes:

En el modelo base, cada nodo de la red debe ser abastecido por un “proveedor” previamente fijado. Es decir, el modelo no cuenta con la *flexibilidad* necesaria para “elegir” de dónde se transportarán las cajas, semana a semana.

Luego, al plantear el modelo matemático, se agregará una “funcionalidad” en donde, de ser necesario, se permita realizar transportes de inventario entre almacenes.

Cabe aclarar que, si bien la solución óptima del problema no tiene por qué, necesariamente, utilizar este transporte “interno” en la red, sí se puede garantizar que agregar esta funcionalidad permitirá encontrar una CTE menor o igual al CTE que se obtendría sin esta permisión.

A continuación, se agregan dos diagramas que ilustran el transporte permitido, previo y luego de esta implementación:



2) Agrupación de productos en GEPs:

Cómo fue previamente discutido en el capítulo 5.4 *De SKUs a GEPs*, una de las implementaciones principales de la metodología en cuestión, será la agrupación de productos en *clusters* de productos con características similares. Es decir, el planteo de la nueva metodología se realizará a nivel de agrupación GEP, en vez de SKU.

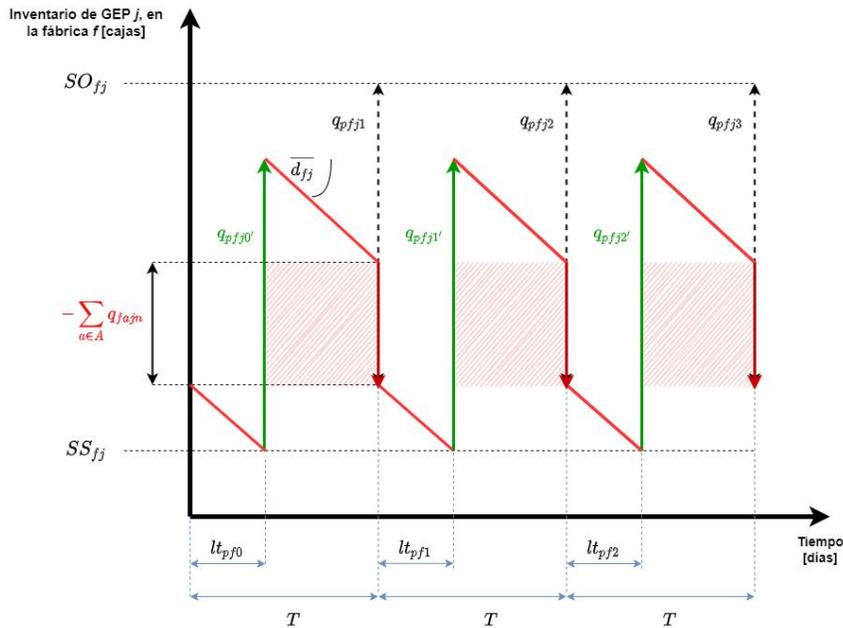
3) Planeamiento de Pedidos Anticipados

Uno de los beneficios de realizar un modelo matemático de pedidos, es que, una vez obtenida la solución, se podrá saber con anticipación la cantidad pedida de cajas que se realizará en cada nodo, cada semana, y para cada GEP.

Luego, es posible anticipar o retrasar los pedidos acordemente, para evitar mantener una cantidad de inventario innecesaria, que vaya a ser redistribuida posteriormente a otro nodo (en la misma semana).

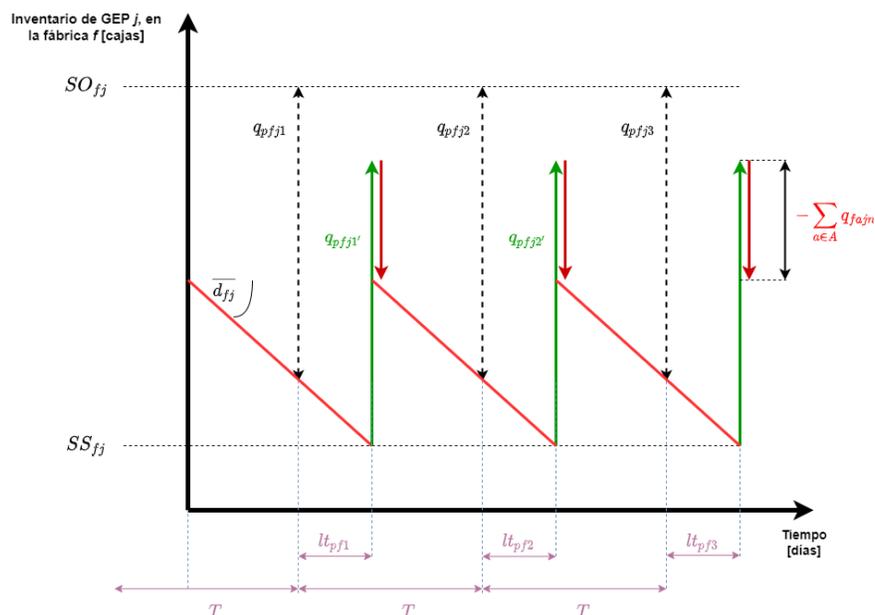
Para ilustrar esta idea, se comenzará por realizar un diagrama explicativo, en donde se observa el nivel de inventario en el tiempo, dentro de una fábrica f . En el mismo, se plantea el caso en donde cada T días, la fábrica realiza un pedido q_{pfjn} al proveedor p , del GEP j , en la semana n .

Además, en el mismo día, la fábrica debe enviar los pedidos $\sum_{a \in A} q_{fajn}$, a los almacenes que abastece cada semana, para cada GEP.



En la figura, se puede notar el área $(T - lt_{pfjn}) * (\sum_{a \in A} q_{fajn})$ rayada en colorado; la cual representa el inventario “extra” que debe contener la fábrica f para satisfacer el pedido próximo de los almacenes a los que satisface.

No obstante, dado que la fábrica conoce con anticipación los pedidos que realizarán los almacenes, ésta podría realizar el pedido q_{pdjn} , con lt_{pfjn} días de anticipación, de manera tal que reciba el pedido de los proveedores el mismo día que despacha las cajas hacia los almacenes.



Proyecciones de Ventas (Forecasting)

Cómo fue previamente detallado, el modelo de Wilson asume que la semanal (d_{jn}) que percibe cada nodo, tiene un comportamiento normal:

$$d_{jn} = d_j \sim N(\mu = \bar{d}_j, \sigma = \sigma_{d_j}), \quad \forall j \in J$$

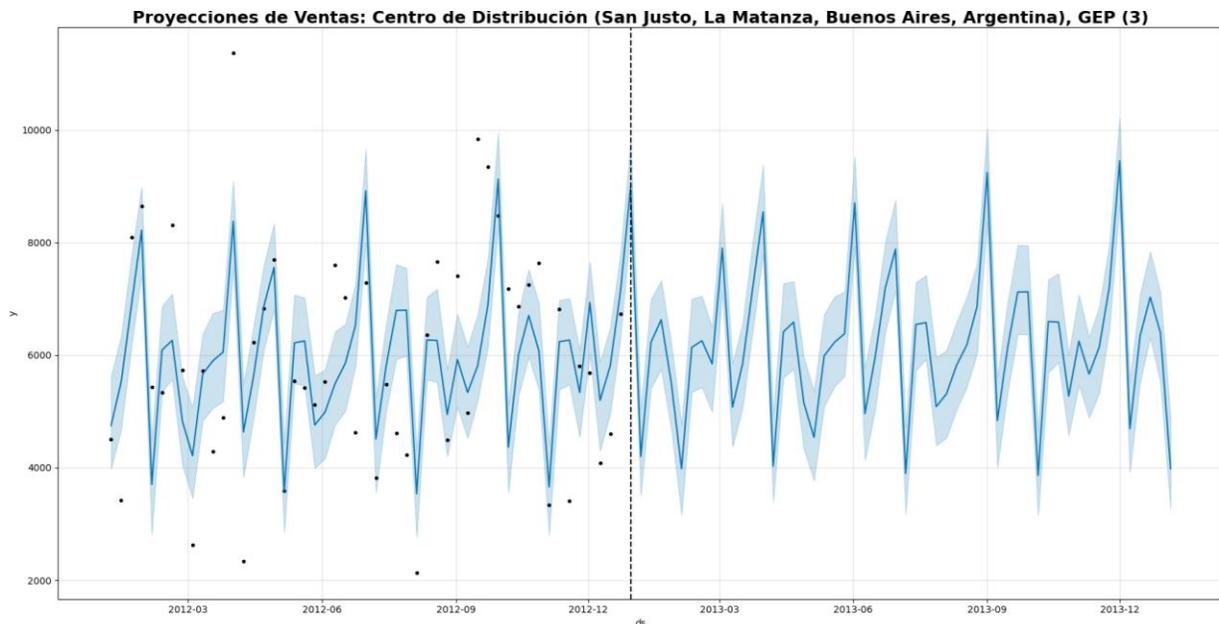
No obstante, en el capítulo 5.3 *Análisis de Datos y Visualizaciones*, se pudo observar cómo dicha demanda semanal parece tener un comportamiento con cierto componente estacional, acompañado de una tendencia sutil.

Esto quiere decir, que se podría utilizar un modelo predictivo como *Prophet*, que capture dicho comportamiento y logre para proyectar la demanda semanal de cada GEP, para cada nodo de la red.

A su vez, esta nueva metodología tiene tres claros beneficios:

1. No se sobreestima la cantidad de inventario necesario en las semanas donde el componente estacional es menor que la demanda promedio (bajas),
2. No se subestima la cantidad de inventario necesario en las semanas donde el componente estacional es mayor que la demanda promedio (picos).
3. La variabilidad en la demanda pasa a estar dado por el componente residual (ε_t) dentro del modelo. Es decir, si se logra componer un modelo con suficiente poder predictivo, se necesitará una cantidad de SS menor para satisfacer la demanda con el nivel de servicio requerido (95%).

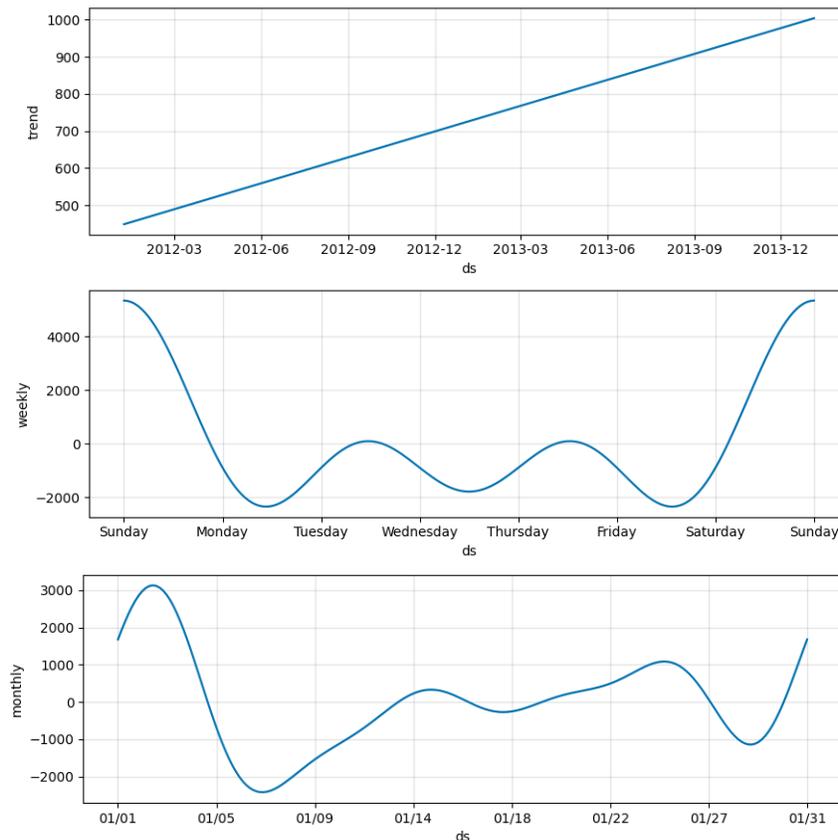
A modo de ejemplo, se realizará una visualización de la proyección de cajas demandadas semanalmente, que se obtiene entrenando a un modelo *Prophet*, para la fábrica ubicada en San Justo, del GEP 3:



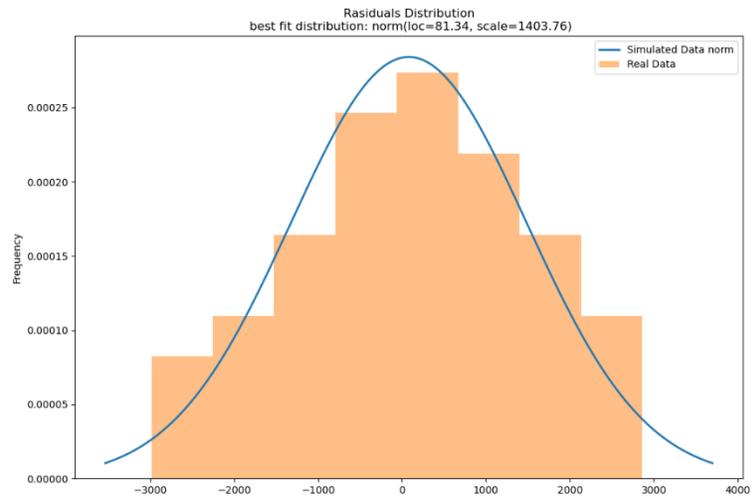
En donde se observan los siguientes componentes:

- En azul, se ven las predicciones realizadas para el año 2012 (del set de entrenamiento) y las proyecciones realizadas para el año 2013.
- En celeste, se denota el intervalo de confianza del 60%, tanto para las predicciones del set de entrenamiento, cómo para las proyecciones a futuro.
- En negro, se puede ver los *outliers* del modelo

Por otra parte, se ilustran los componentes estacionales semanales y mensuales; agregado al componente de tendencia capturado por el modelo:



Finalmente, se observa un histograma de los residuos del modelo (en naranja); sobre el cual se agregó la función de densidad de probabilidad de la distribución normal que mejor describe este histograma:



En donde se puede verificar cómo, para este modelo, la suposición de normalidad de los residuos parece estar correctamente justificada.

5.8.2 Modelo Matemático

Finalmente, se buscará idear un modelo matemático en el cual se minimice la esperanza del costo total anual; implementando las optimizaciones previamente mencionadas.

Nota: se habla de minimizar la esperanza del costo total anual, dado que el mismo depende de la demanda (variable aleatoria). Es decir, se minimizará el costo total que incurriría la empresa, si la demanda real fuera igual a la demanda proyectada.

Para encontrar la solución óptima del modelo, el mismo se planteará cómo un ejercicio de programación lineal entera. No obstante, es importante resaltar que el planteo cuenta con la limitación de no poder incorporar directamente el costo por almacenamiento de SS, ya que esto implicaría que el problema pierda su linealidad.

Esto se debe a que, dicho costo dependerá de la frecuencia de pedidos que perciba cada nodo; en donde, si un nodo recibe pedidos todas las semanas (de un cierto GEP), el SS necesario para lograr un nivel de servicio de 95% estaría dado por:

$$SS = Z_{95\%} * \sqrt{T} * \sigma e_{fj}$$

No obstante, si la frecuencia promedio cambiara, y se empezaran a recibir pedidos cada dos semanas, entonces el SS debería ser calculado cómo:

$$SS = Z_{95\%} * \sqrt{2 * T} * \sigma e_{fj}$$

En donde, la cantidad de stock de seguridad necesario para satisfacer la demanda, dependerá de $\sqrt{\frac{1}{freq}}$, en donde *freq* representa la frecuencia de recibimiento de pedidos (medido en pedidos recibidos por semana).

Luego, calcular este parámetro implicaría elevar una variable decisión con un exponente distinto a 1, ergo, rompiendo con el concepto de linealidad (necesario para satisfacer el problema con la metodología elegida).

Es por esta razón, que se incluirá un componente en la función objetivo, que aproxime el costo lineal que incurriría el modelo al disminuir la frecuencia de entrega de pedidos a un cierto nodo:

$$\sum_{f \in F} \sum_{j \in J} \left(h_j * Z_{95\%} * \frac{1}{2} * \left(1 - \sum_{n \in N} \frac{b_{fjn}}{54} \right) * \sigma e_{fj} \right)$$

A Conjuntos y Parámetros

J: Conjunto de GEPs. $J = (0, 1, 2 \dots j)$

- j : GEP; $j \in J$
- p_{vj} : precio de venta para la familia de productos j [usd]
- cv_j : costo de venta para la familia de productos j [usd]
- h_j : holding cost para una caja de la familia de productos j [usd/caja.año]
- w_j : peso promedio por caja de la familia de productos j [kg]

N: Conjunto de semanas. $N = (0, 1, 2 \dots 53)$

- n : semana; $n \in N$
 - *Nota: cómo fue previamente explicitado, sólo se considerará el próximo año.*

P: Conjunto de Proveedores. $P = (0, 1, 2, 3, 4)$

- p : proveedor; $p \in P$
 - $\{0: peru, 1: costa_rica, 2: mexico, 3: nicaragua, 4: usa\}$
- s_{pj} : capacidad de suministro del proveedor p (anual) [cajas]
- k_{pf} : costo de realizar un pedido al proveedor p , para la fábrica f [usd]
- lt_{pf} : tiempo de suministro (lead time) del proveedor p hasta la fábrica f [días]

F: Conjunto de Fábricas (Plantas). $F = (0)$

- f : fábrica; $f \in F$
 - $\{0: San Justo\}$
- dp_{fjn} : demanda proyectada que percibe la fábrica f , de la familia de productos j , en la semana n [cajas]
- σe_{fj} : desviación estándar del error esperado en la demanda proyectada que percibe la fábrica f , de la familia de productos j [cajas]

A: Conjunto de Almacenes (Centros de Distribución). $A = (0, 1, 2 \dots 10)$

- a : almacén; $a \in A$
 - $\{0: Itá Ibaté, 1: Guanaco Sombriana, 2: El Fortín, 3: La Dormida, 4: Camarones\}$
- dp_{ajn} : demanda por día que percibe el almacén a , para la familia de productos j [cajas]
- σe_{aj} : desviación estándar del error esperado en la demanda proyectada que percibe el almacén a , de la familia de productos j [cajas]

C: Conjunto de camiones

- ct_{fa} : costo de transportar un camión (cargado con productos) desde la fábrica f , hasta el almacén a [usd]
- ct_{aa} : costo de transportar un camión (cargado con productos) desde el almacén a , hasta el almacén a' [usd]
- wm_c : peso máximo que podrá transportar el camión c [kg]
 - *Nota: se asume que el factor limitante es el peso*
- lt_{fa} : tiempo de recorrido desde la fábrica f hasta el almacén a [días]
- lt_{aa} : tiempo de recorrido desde el almacén a hasta el almacén a' [días]
- T : tiempo entre cada pedido
 - $T = 7$ días

B Variables Decisión

Para cada $p \in P$, $f \in F$, $a \in A$:

- q_{pfjn} : cantidad total de cajas trasladadas desde el proveedor p , a la fábrica f , del GEP j , en la semana n [cajas]
- q_{fajn} : cantidad total de cajas trasladadas desde la fábrica f , al almacén a , del GEP j , en la semana n [cajas]
- $q_{aa'jn}$: cantidad total de cajas transportadas desde el almacén a , al almacén a' , del GEP j , en la semana n [cajas]
 - $q_{pfjn}, q_{fajn}, q_{aa'jn} \in$ variable entera ≥ 0
- ii_{fjn} : inventario inicial esperado en la fábrica f , del GEP j , en la semana n [cajas]
- if_{fjn} : inventario final esperado en la fábrica f , del GEP j , en la semana n [cajas]
- ii_{ajn} : inventario inicial esperado en el almacén a , del GEP j , en la semana n [cajas]
- if_{ajn} : inventario final esperado en el almacén a , del GEP j , en la semana n [cajas]
 - $ii_{fjn}, if_{fjn}, ii_{ajn}, if_{ajn}, ss_{fjn}, ss_{ajn} \in$ variable entera ≥ 0

$$b_{pfjn} = \begin{cases} 1, & \text{proveedor } p \text{ envía un pedido a la fábrica } f, \text{ del GEP } j, \text{ en la semana } n \\ 0, & \text{proveedor } p \text{ NO envía un pedido a la fábrica } f, \text{ del GEP } j, \text{ en la semana } n \end{cases}$$

- $b_{pfjn} \in$ variable binaria ≥ 0
- nc_{fan} : cantidad de camiones usados para transportar productos desde la fábrica f , al almacén a , en la semana n [camiones]

- $nC_{aa'n}$: cantidad de camiones usados para transportar productos desde el almacén a , al almacén a , en la semana n [camiones]
 - $nC_{fan}, nC_{aa'n} \in$ variable entera ≥ 0

$$b_{fjn} = \begin{cases} 1, & \text{fábrica } f \text{ recibe algún pedido, del GEP } j, \text{ en la semana } n \\ 0, & \text{fábrica } f \text{ NO recibe algún pedido, del GEP } j, \text{ en la semana } n \end{cases}$$

$$b_{ajn} = \begin{cases} 1, & \text{almacén } a \text{ recibe algún pedido, del GEP } j, \text{ en la semana } n \\ 0, & \text{almacén } a \text{ NO recibe algún pedido, del GEP } j, \text{ en la semana } n \end{cases}$$

- $b_{fjn}, b_{ajn} \in$ variable binaria ≥ 0

C Función Objetivo

- Costo de Mantenimiento de Inventarios:
 - Costo de Mantenimiento de inventarios en la(s) fábricas:

$$CI_F = \sum_{f \in F} \sum_{j \in J} \sum_{n \in N} \left(\frac{ii_{fjn} + if_{fjn}}{2} * \frac{h_j}{54} \right) + \sum_{f \in F} \sum_{j \in J} \left(h_j * Z_{95\%} * \frac{1}{2} * \left(1 - \sum_{n \in N} \frac{b_{fjn}}{54} \right) * \sigma_{e_{fj}} \right)$$

Nota: el segundo término del sumando es un componente pensado para “desmotivar” al modelo a disminuir la frecuencia de recibimiento de pedidos para cada nodo.

Una vez encontrado la solución óptima, se deberá calcular el costo por almacenamiento de stock de seguridad, debidamente.

- Costo de Mantenimiento de inventarios en lo(s) almacenes:

$$CI_A = \sum_{a \in A} \sum_{j \in J} \sum_{n \in N} \left(\frac{ii_{ajn} + if_{ajn}}{2} * \frac{h_j}{54} \right) + \sum_{a \in A} \sum_{j \in J} \left(h_j * Z_{95\%} * \frac{1}{2} * \left(1 - \sum_{n \in N} \frac{b_{ajn}}{54} \right) * \sigma_{e_{aj}} \right)$$

- Costo de Transporte de Inventarios:
 - Proveedores hacia fábricas:

$$CT_{PF} = \sum_{p \in P} \sum_{f \in F} \sum_{n \in N} b_{pfn} * k_{pf}$$

- *Fábricas hacia Almacenes:*

$$CT_{FA} = \sum_{f \in F} \sum_{a \in A} \sum_{n \in N} nc_{fan} * ct_{fa}$$

- *Almacenes hacia Almacenes:*

$$CT_{AA} = \sum_{a \in A} \sum_{a' \in A} \sum_{n \in N} nc_{a'an} * ct_{aa'}$$

- *Costo total a minimizar:*

$$Z = CMI_F + CMI_A + CT_{PF} + CT_{FA} + CT_{AA}$$

D Restricciones

- D.1 Cada proveedor no podrá enviar más cajas que su capacidad de suministro:

$$s_{pj} \leq \sum_{f \in F} \sum_{j \in J} \sum_{n \in N} q_{pfjn} \quad (\forall p \in P, j \in J)$$

- D.2 Restricción que relaciona las cajas pedidas, con su correspondiente variable binaria, para cada nodo, cada GEP y cada semana:

$$q_{pfjn} \leq b_{pfjn} * 100000 \quad (\forall p \in P, f \in F, j \in J, n \in N)$$

$$b_{fjn} \leq \sum_{p \in P} q_{pfjn} \quad (\forall f \in F, a \in A, j \in J, n \in N)$$

$$b_{ajn} \leq \sum_{f \in F} q_{fajn} + \sum_{a' \in A} q_{a'ajn} \quad (\forall a' \in A, a \in A, j \in J, n \in N)$$

- D.3 Restricción que relaciona el stock inicial con la demanda proyectada y el stock final, para cada nodo, cada GEP y cada semana:

$$ii_{fjn} - dp_{fjn} = if_{fjn} \quad (\forall f \in F; j \in J; n \in N)$$

$$ii_{ajn} - dp_{ajn} = if_{ajn} \quad (\forall a \in A; j \in J; n \in N)$$

D.4 Restricción que relaciona el stock inicial, con el stock final de la semana previa y las cajas trasladadas (entrantes y salientes), para cada nodo, cada GEP y cada semana ($n \neq 0$):

$$if_{fj(n-1)} + \sum_{p \in P} q_{pfjn} - \sum_{a \in A} q_{fajn} = ii_{fjn} \quad (\forall f \in F; j \in J; n \neq 0 \in \mathbb{N})$$

$$if_{aj(n-1)} + \sum_{f \in F} q_{fajn} + \sum_{a' \in A} q_{a'ajn} - \sum_{a' \in A} q_{aa'jn} = ii_{ajn} \quad (\forall a \in A; j \in J; n \neq 0 \in \mathbb{N})$$

D.5 Se asume que, en la primera semana ($n=0$), se tiene un stock igual a la demanda proyectada para la semana siguiente

$$dp_{fj1} + \sum_{p \in P} q_{pfj0} - \sum_{a \in A} q_{faj0} = ii_{fj0} \quad (\forall f \in F; j \in J)$$

$$dp_{aj1} + \sum_{f \in F} q_{faj0} + \sum_{a' \in A} q_{a'aj0} - \sum_{a' \in A} q_{aa'j0} = ii_{aj0} \quad (\forall a \in A; j \in J)$$

D.6 Los camiones no podrán transportar una cantidad mayor a su capacidad máxima:

$$nc_{fan} * wm \geq \sum_{j \in J} (q_{fajn} * w_j) \quad (\forall f \in F; a \in A; n \in \mathbb{N})$$

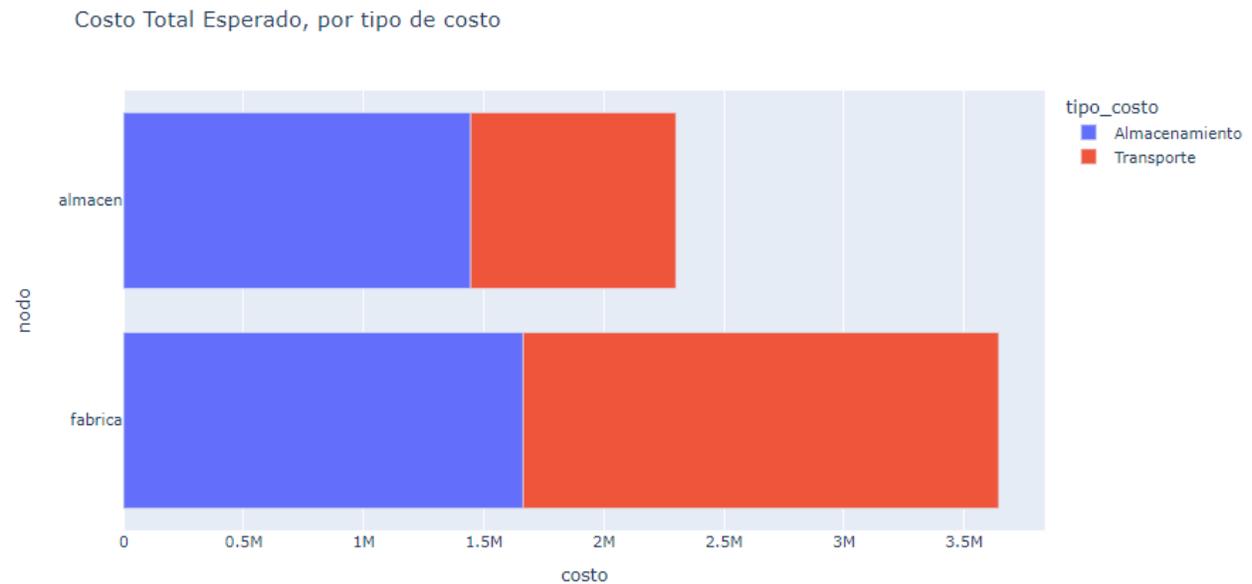
$$nc_{aa'n} * wm \geq \sum_{j \in J} (q_{aa'jn} * w_j) \quad (\forall a \in A; a' \in A; n \in \mathbb{N})$$

E Solución:

CTE = 5.984.087,8 usd/año

Cabe mencionar que las ecuaciones previamente expuestas sirven para obtener las decisiones de almacenamiento y traslado de inventarios dentro de la red. No obstante, al CTE calculado con la solución encontrada se le agregó el costo de almacenamiento de SS a nivel producto (cómo fue previamente explicado en el capítulo 5.4).

Similarmente, se realizará un desglose de los costos divididos por tipo.



Nota: para que la solución converja, fue necesario subdividir el problema en tres “secciones”; en donde cada sección se utilizó para hallar la solución óptima de 10, 10 y 11 GEPs. Luego de encontrar los tres mínimos locales del problema, se debió adicionar los costos incurridos en cada uno.

6 Simulación

Si bien los resultados de los CTE para ambas metodologías parecen ser, intuitivamente, convincentes; es necesario recordar que los mismos son simplemente aproximaciones que resultan de resolver una instancia simplificada del problema “real”.

En particular, ambas metodologías intentan resolver el caso “pseudo determinístico”, en donde se asume, entre otras cosas, que se conoce el valor de ciertos parámetros que en la realidad son valores aleatorios que varían en el tiempo.

A continuación, se realiza un listado, no exhaustivo, de las variables principales a las que se está aludiendo:

- *Demanda*: se trabaja con la demanda media (modelo base) o con la demanda proyectada (modelo alternativo), a pesar de que la misma es una variable aleatoria en donde, incluso los modelos predictivos más sofisticados contienen un cierto error.
- *Tiempo de Suministro*: se asume constante y conocido; a pesar de que en la realidad puede haber variaciones considerables en el mismo.
- *Costo de Transporte*: nuevamente, se asumen constantes y conocidos; a pesar de ser variables.
- *Costo de Almacenamiento*: a pesar de ser asumidos como conocidos y constantes, también pueden variar.

Simplemente, es innegable que en la realidad pueden existir situaciones como: picos de demanda imprevistos, retrasos en los transportes, incumplimiento en el tiempo de suministro de proveedores y/o incrementos en el sueldo de los choferes de camión que aumenten el costo de transporte.

Luego, la mejor manera de conseguir una aproximación más acertada y confiable es simulando el entorno en que se implementarían las distintas políticas de inventario; repitiendo dicho experimento suficientes veces, para lograr obtener un intervalo de confianza del CTE de cada política que permita validar los resultados estimados previamente.

No obstante, dado que la simulación será más bien una metodología de validación de resultados (y no el objetivo central del trabajo), solamente se simulará el caso donde la demanda semanal de cada GEP sea igual a la suma algebraica entre la demanda proyectada por el modelo entrenado y un componente aleatorio ε_t , que se obtiene extrayendo una muestra de la distribución normal que mejor explique a los residuos.

Es decir, para cada una de las 54 semanas del año, la demanda simulada para cada nodo (de cada GEP) de la red, estará dado por:

$$ds_{fjn} = dp_{fjn} + \varepsilon_{fjn}$$

- ds_{fjn} : demanda simulada en el nodo f (fábrica o almacén), del GEP j , en la semana n .
- dp_{fjn} : demanda proyectada (por el modelo predictivo) en el nodo f (fábrica o almacén), del GEP j , en la semana n .
- ε_{fjn} : error aleatorio (positivo o negativo) en el nodo f (fábrica o almacén), de la predicción de la demanda del GEP j en la semana n , que surge de tomar una muestra de la distribución normal $N(\mu = \overline{\varepsilon_{jn}}, \sigma = \sigma_{\varepsilon_{jn}})$
 - Nota: el error de cada semana es independiente al error del resto de las semanas

Luego, una vez simulada la demanda de cada semana del año próximo, se realiza el cálculo del CTE que se incurriría (con la metodología base y la metodología alternativa) dada la demanda simulada y finalmente se repite todo el proceso (con distintas demandas simuladas) para cada una de las 300 veces que se realiza la simulación.

A continuación, se expondrán los resultados de las simulaciones en el capítulo 7.2 *Planteo Estocástico*.

7 Análisis de Resultados

7.1 Planteo Pseudo Determinístico

En los capítulos 5.7 *Política de Inventarios Base* y 5.8 *Metodología Alternativa: Nueva Política de Inventarios*, se buscó implementar distintos modelos matemáticos, basados en un planteo pseudo determinístico del problema. En donde la variabilidad de la demanda se veía reflejada en la necesidad de un cierto nivel de SS (necesario para mantener el nivel de servicio deseado); y luego, se realizaba el cálculo del costo esperado, asumiendo cierta demanda predeterminada.

Una vez “corridos” los modelos, se encontraron las políticas de inventario, con sus respectivos CTE; en donde se pueden distinguir algunas claras diferencias:

- *Costo Total Esperado:*
Utilizando la metodología propuesta, el CTE fue reducido un **9.7%**, en relación con el CTE calculado con el modelo Base; lo cual implicaría un ahorro anual de **645.124,27 usd**.
- *Costo de Transporte:*
 - El costo de transporte incurrido por la fábrica se *incrementó* un **9.3%**
 - El costo de transporte incurrido por los almacenes *incrementó* un **36.2%**
- *Costo de Almacenamiento:*
 - El costo de Almacenamiento incurrido por la fábrica *disminuyó* un **26.8%**
 - El costo de Almacenamiento incurrido por los almacenes *disminuyó* un **22.9%**

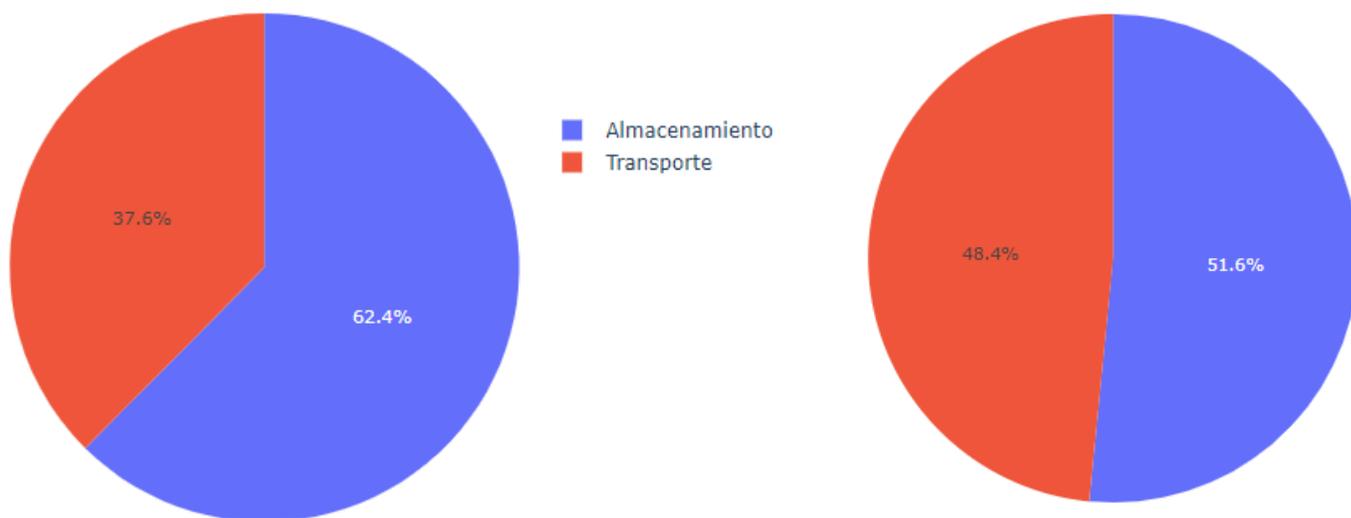
En concreto, se puede notar cómo en la propuesta alternativa, se logró reducir considerablemente el costo de almacenamiento; mientras que el costo de transporte hacia las fábricas sufrió pequeñas modificaciones y se incrementó el costo de transporte hacia los almacenes.

Esto sugiere, que el optimizador logró “aprovechar” las permisiones adicionales en cuanto al traslado de inventarios para encontrar una solución más flexible y *balanceada*, en donde se distribuye el inventario de manera más eficiente (reconociendo el costo relativamente “barato” de traslado de inventario hacia los almacenes).

De esta forma, se logra el siguiente balance de costos:

Costo Total Esperado, por Tipo de Costo: Modelo de Base

Costo Total Esperado, por Tipo de Costo: Modelo Alternativo



En cuanto al costo de almacenamiento, se puede ver el planeamiento anticipado de pedidos tiene como consecuencia un mejor aprovechamiento del inventario dentro de la fábrica, lo cual se traduce en un menor costo de almacenamiento

Además, se debe tener presente que la utilización de modelos predictivos de demanda supone (entre otros factores), una reducción en la variabilidad de la demanda percibida; que implica, a su vez, un nivel de SS menor, en cada nodo.

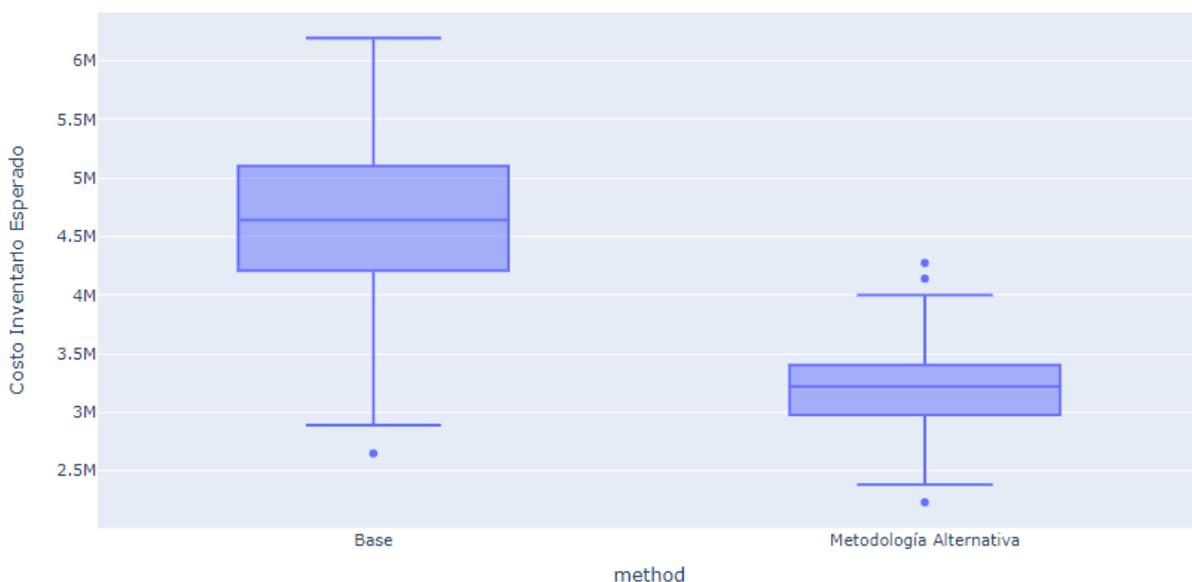
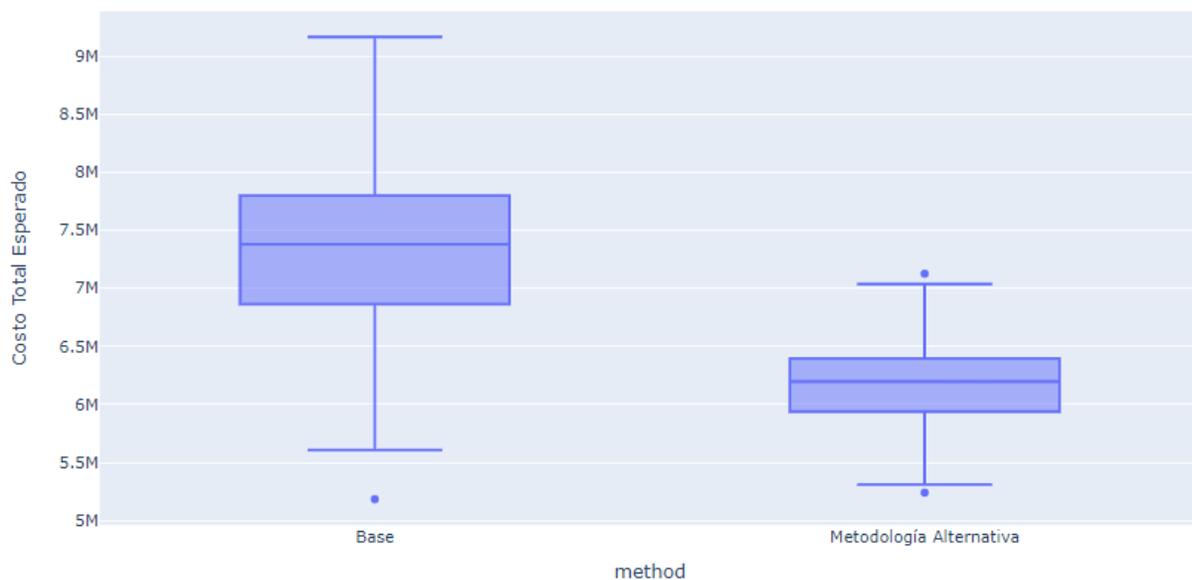
En definitiva, lo que se pone en evidencia es que un modelo más sofisticado, con menores restricciones, puede lograr resultados significativamente mejores. Es decir, que se justifica la implementación de distintas optimizaciones que ayuden a *flexibilizar* al modelo.

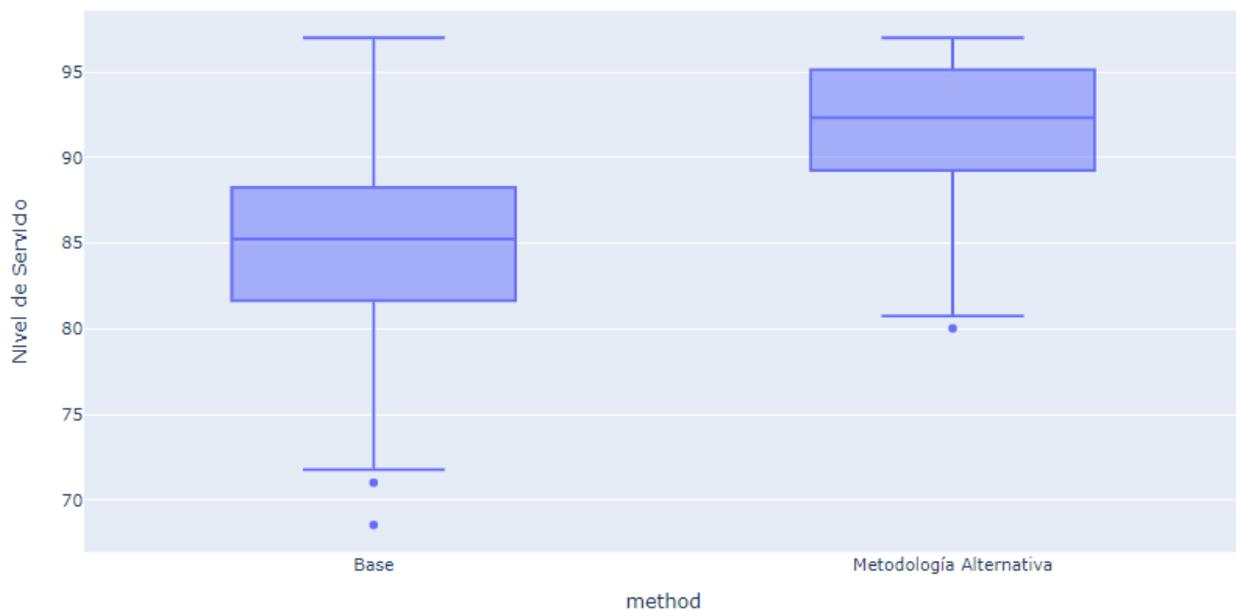
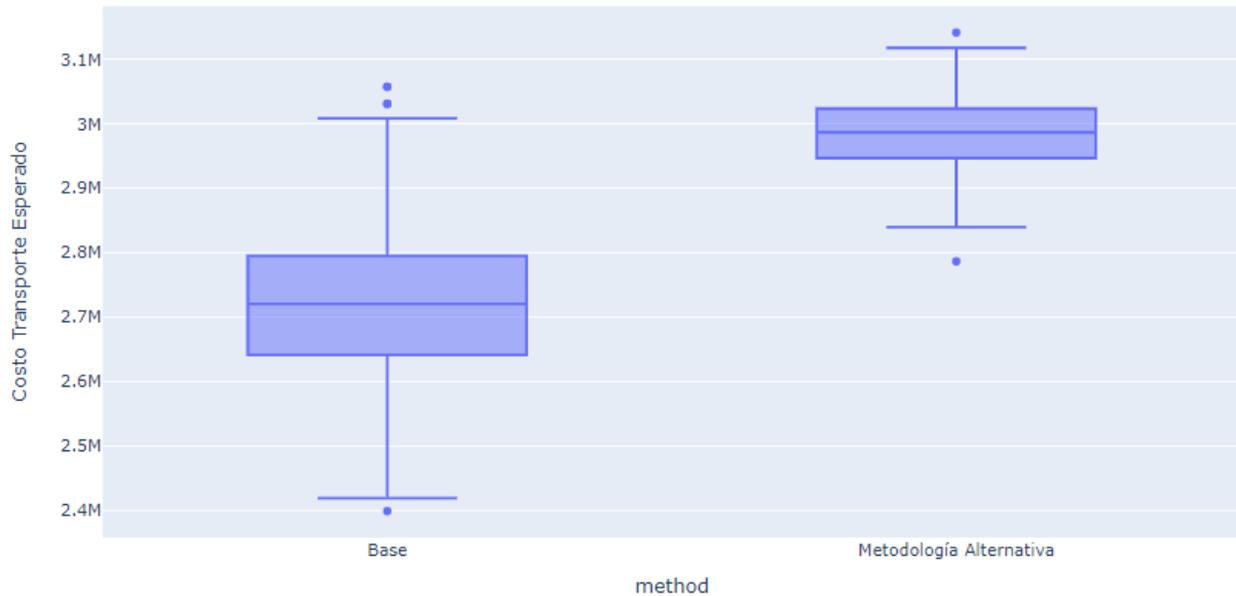
7.2 Planteo Estocástico

Cómo fue mencionado en el capítulo 6 *Simulación*, el planteo pseudo determinístico es simplemente una aproximación obtenida de un planteo sobre simplificado de la realidad en la que actúa la red.

Es por esta razón, que se simulará el costo “real” que se incurriría al aplicar dichas políticas de inventario, afrontando una demanda con cierto grado de aleatoriedad. Luego, se repetirá dicha simulación un total de 300 veces, para lograr obtener resultados estadísticamente significativos.

A continuación, se expondrán los resultados de dichas simulaciones:





- La diferencia entre los CTE promedio simulados fue incrementada, desde un 9.7% (planteo pseudo determinístico) hasta un **15.9%**.
 - Lo cual representa un ahorro anual de **1.172.961 usd**.
- Si bien el CTE simulado con la metodología alternativa parece ser notoriamente similar al CTE estimado en el planteo pseudo determinístico (tan solo un 3.41% de diferencia); el CTE simulado con la metodología basada en el modelo base es significativamente mayor al calculado previamente (8.54% de diferencia)

- Esto sugiere que la metodología alternativa no solamente es más eficiente, sino que es también una manera considerablemente más apropiada de estimar el costo real que se incurriría.
- El costo incurrido por almacenamiento de inventario no solo es considerablemente menor en la metodología alternativa, sino que además tiene una menor variabilidad.
- El costo incurrido por transporte de inventario sutilmente mayor en la metodología alternativa.
- Para ambos casos, el costo incurrido por transporte de inventario parece tener una variabilidad reducida (en comparación con la variabilidad en el costo incurrido por almacenamiento de inventario)
- Se logró conseguir un mejor nivel de servicio usando la metodología alternativa; lo cual es razonable, dado que la misma logra predecir y anticiparse los picos de demanda heterogéneos que se incurrirán, por la estacionalidad de cada producto.

Si bien los resultados parecen indicar, cómodamente, que el CTE de la alternativa es significativamente menor al CTE obtenido con el modelo base, se procederá a calcular un intervalo de confianza (de 95% de certeza) para el CTE medio de ambas metodologías; el cual se calcula de la siguiente forma:

$$IC_{CTE} = \left(\overline{CTE} - Z_{\frac{\alpha}{2}} * \frac{S_{CTE}}{\sqrt{N}} ; \overline{CTE} + Z_{\frac{\alpha}{2}} * \frac{S_{CTE}}{\sqrt{N}} \right)$$

En donde:

\overline{CTE} : Promedio del Costo Total Esperado

S_{CTE} : desviación estándar del Costo Total Esperado

α : error esperado (5%)

N : Cantidad de observaciones (coincide con el número de simulaciones)

Luego:

$$IC_{CTE \text{ base}} = (7.274686,8 ; 7.431.150,7)$$

$$IC_{CTE \text{ alternativo}} = (6.141.196,7 ; 6.218.718,3)$$

Es decir, que se puede afirmar, casi con total seguridad, que el CTE de la Metodología Alternativa es significativamente menor al CTE base.

8 Conclusión

En el capítulo 7 *Análisis de Resultados*, se describió el impacto real de aplicar la nueva política de inventarios dentro de la red logística en cuestión. Allí se logró demostrar, con suficiente nivel de certeza, que la nueva política es considerablemente superadora; tanto en ahorro del CTE, cómo en el nivel de servicio logrado.

De esta manera, se puede verificar cómo el diseño de modelos matemáticos y la implementación de herramientas de analítica avanzadas pueden ser sumamente útiles para incrementar la rentabilidad de la empresa y satisfacer los objetivos de negocio.

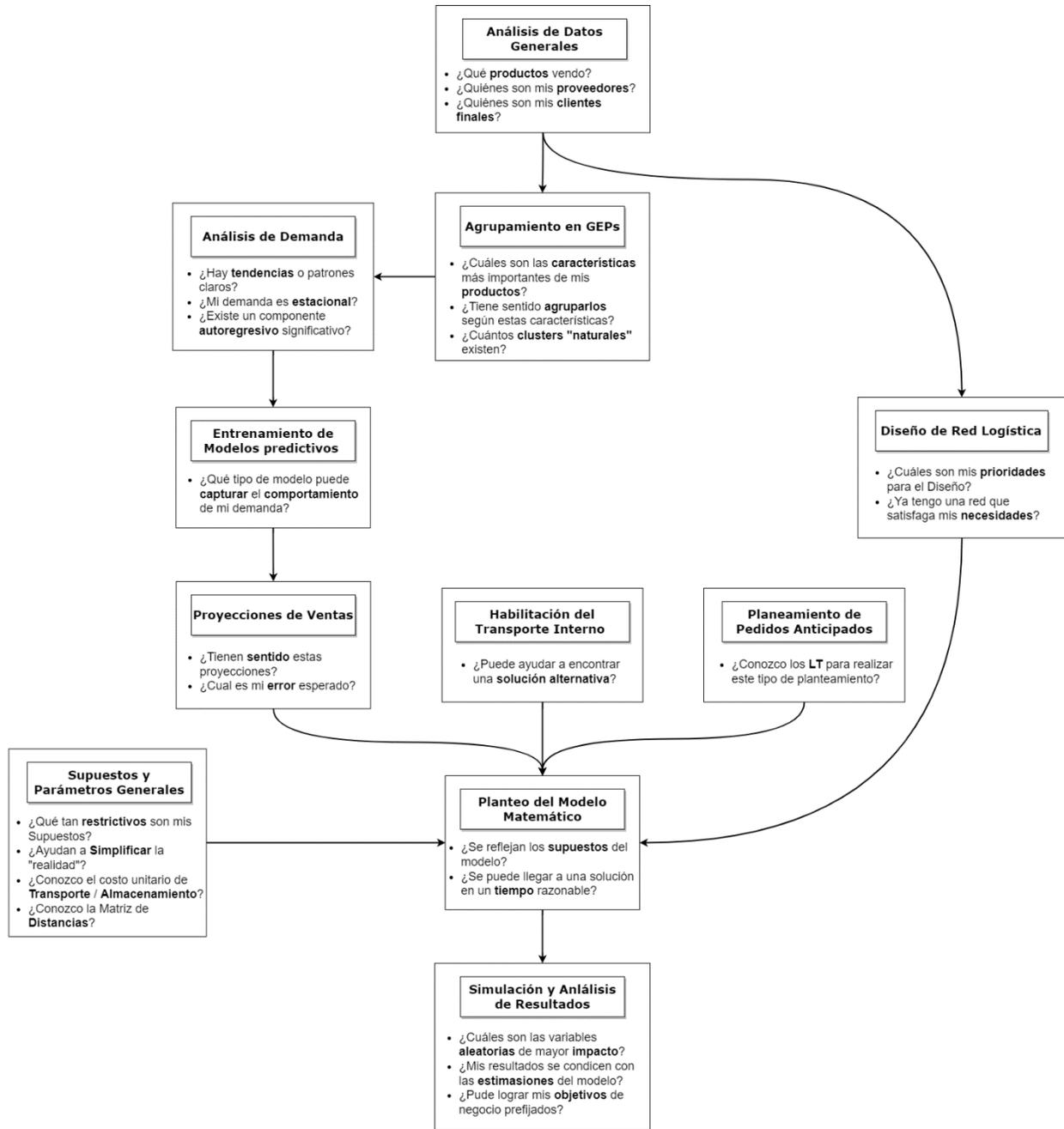
Se destacó también, cómo este tipo de metodología cuenta con la gran ventaja de poder tomar “perspectiva”, flexibilizando el problema y midiendo el impacto de determinadas decisiones sobre la red logística completa. Es decir, dada la naturaleza abarcativa de la metodología; ésta es capaz de capturar el mínimo *global* del problema, sin estar sesgada por los mínimos locales de cada nodo.

No obstante, cabe señalar que los resultados expuestos son el resultado de implementar la nueva política de inventarios en la red logística trabajada. Es decir, que a pesar de contar con una metodología notablemente robusta, no es posible garantizar este tipo de ahorros y nivel de servicio para cualquier otra red planteada.

Dicho esto, se argumenta que el objetivo del trabajo no es encontrar una “fórmula” que solucione todo tipo de problemas de inventario, sino que el valor está en introducir al lector a las etapas necesarias para resolver un problema complejo y desestructurado; desglosándolo en distintas secciones relativamente simples, que puedan ser resueltas con nivel de abstracción razonable, utilizando herramientas de analítica avanzadas.

Es por esta razón, que se buscó desarrollar cada etapa dentro de la metodología, justificando la intuición detrás de la aplicación de cada herramienta, para que el lector pueda adquirir (o al menos introducirse en) distintos conceptos e instrumentos que pueda usar para inspirarse y replicar este tipo de trabajo en situaciones comparables.

Finalmente, a modo de resumen, se agregará un diagrama explicativo completo, en donde se ilustra las etapas más significativas realizadas en el trabajo; acompañado de sus respectivas preguntas “guía” más pertinentes:



9 Apéndice

Recursos

- Estimador de precios de transporte para proveedores:
 - https://wwwapps.ups.com/ctc/request?loc=en_AR
- Consumo de combustible:
 - <https://www.lavoz.com.ar/agro/actualidad/estiman-que-los-bitrenes-reducen-30-el-gasto-en-gasoil/#:~:text=En%20principio%2C%20suponen%20un%20consumo,entr e%2028%20y%2035%20litros.>
- Sueldo promedio de un chofer de camión de carga:
 - https://www.glassdoor.com.ar/Sueldos/buenos-aires-chofer-de-camiones-sueldo-SRCH_IL.0,12_IM963_KO13,31.htm
- Costo de un camión de carga:
 - <https://listado.mercadolibre.com.ar/camiones-para-carga-pesada>
- Peso máximo de transporte:
 - <https://galeon.com.ar/ya-rige-el-aumento-del-peso-en-los-camiones-en-argentina/#:~:text=Ya%20est%C3%A1%20vigente%20en%20Argentina,55%2C5%20toneladas%20de%20PBT.>