



UAI
UNIVERSIDAD ADOLFO IBÁÑEZ



MEJORA EN LA PRODUCTIVIDAD DEL PROCESO DE OUTBOUND DE CARGA DEL CD 6011 QUILICURA DE WALMART CHILE

30 DE NOVIEMBRE, 2023

Alumna: Micaela Campusano Padilla

Carrera: Ingeniería Civil Industrial

Contenido

1. Resumen ejecutivo	3
2. Abstract	4
3. Contexto	5
4. Problema	9
5. Objetivos y medidas de desempeño	13
6. Análisis Causas del problema	15
7. Estado del arte	16
8. Solución	21
9. Metodología	23
10. Implementación y desarrollo solución	26
11. Resultados	30
12. Conclusiones y discusión	33
13. Referencias	35
14. Anexos	36

1. Resumen ejecutivo

Dentro de la empresa Walmart Chile, uno de los principales retailer del país en cuanto a alimentos y mercancías, se presenta el área de reabastecimiento, perteneciente a la gerencia de Supply Chain, dentro del área encontramos distintos segmentos siendo uno de estos los alimentos perecederos.

Un proceso relevante del segmento corresponde al de preparar las cajas en el centro de distribución para despachar a locales, esto es realizado por los llamados “surtidores”. Es acá donde se ha visto que el personal en el proceso de surtido de cajas ha presentado altos y bajos durante las semanas del segundo semestre del año, con diferencias del objetivo de la empresa que corresponde a surtir 21¹ cajas/hr por surtidor, mientras que los surtidores en promedio surten 13 cajas/hr cada uno.

Esta diferencia principalmente más baja del objetivo de la empresa nos lleva a buscar como objetivo de este presente proyecto aumentar la productividad de los surtidores del proceso mencionado a 19 cajas/hr por cada uno.

Se indaga entonces en la principal causa que nos lleva a entender que el área de reabastecimiento declara constantemente una cantidad de cajas por surtir en el CD desviado al que efectivamente se tiene la necesidad de surtir por necesidad del negocio, ya que la declaración se hace de manera subjetiva sin ni una estandarización o cálculo, entonces el CD suele prepararse con una capacidad de personal mayor a la verdadera necesidad del negocio y de la real cantidad de productos que deben despacharse. Al entender esto, se propone realizar un método de pronóstico objetivo para declarar el volumen de cajas que se surtirá en CD.

A través de la prueba de algunos métodos de pronósticos de series de tiempo con diferentes cualidades, se obtienen los que poseen menor error con respecto al resultado real y se emplea uno de ellos para calcular un inventario de seguridad que se añade al cálculo de la declaración de cajas que se deben surtir al CD.

Con este nuevo cálculo de pronósticos tenemos una proyección objetiva con la cual declarar con menor grado de incertidumbre la cantidad de cajas que se espera surtir semanalmente en el centro de distribución, la cual este utilizará como input para dotar su capacidad de personal de la manera correcta generando mayor eficiencias en productividad y a su vez ahorrando el costo de tener personal extra que termina siendo innecesario si la necesidad real de cajas por surtir no equipara a la capacidad de personal.

Se reduce la brecha entre capacidad y necesidad en un 87% y la productividad por surtidor sobrepasa las 19 cajas/hr incluso en algunos casos llegando al objetivo de la empresa que contempla 21 cajas/hr.

Se sugieren próximos pasos tales como añadir variables exógenas a métodos de regresión lineal más complejos que permitan pronosticar con aún mayor precisión contemplando una mayor gama de variables que afectan la cantidad de cajas por despachar del centro de distribución de Walmart Chile.

¹ Las cifras de productividad tales como cuántas cajas/hora se surten, han sido ponderadas por un factor por temas de confidencialidad de Walmart Chile.

2. Abstract

Within Walmart Chile, one of the leading retailers in the country for food and goods, the replenishment department operates under the Supply Chain management. Within this department, various segments exist, with one of them focusing on perishable goods.

A crucial process within this segment involves preparing boxes at the distribution center for dispatch to local stores, a task carried out by the so-called "surtidores" (pickers). It has been observed that the personnel involved in the box-picking process has shown fluctuations in performance during the second half of the year, deviating from the company's goal of picking 21 boxes per hour per picker. On average, the pickers currently achieve 16 boxes per hour each.

The primary objective of this project is to increase the productivity of the pickers in the process to 19 boxes per hour per picker, aiming to narrow the gap between the current performance and the company's target.

Investigation into the main cause reveals that the replenishment department consistently declares a quantity of boxes to be picked at the distribution center that deviates from the actual business need. This discrepancy arises due to subjective declarations without standardization or calculation. As a result, the distribution center is often staffed with more personnel than necessary, leading to an excess of resources.

To address this issue, a method for an objective forecasting of the volume of boxes to be picked at the distribution center is proposed. After testing various time series forecasting methods, the ones with the lowest errors compared to actual results are identified. One of these methods is then employed to calculate a safety inventory, which is added to the box declaration calculation for the distribution center.

With this new forecasting calculation, there is a more accurate projection of the weekly quantity of boxes expected to be picked at the distribution center. This projection serves as input for determining the correct staffing levels, thereby increasing productivity and saving costs associated with unnecessary personnel.

The gap between capacity and actual need is reduced by 34%, and picker productivity surpasses 19 boxes per hour, even reaching the company's target of 21 boxes per hour in some cases.

Future steps are suggested, including incorporating exogenous variables into more complex linear regression methods to enhance forecasting precision by considering a broader range of factors influencing the quantity of boxes to be dispatched from Walmart Chile's distribution center.

3. Contexto

3.1 Contexto de la empresa:

Walmart Chile es una filial de Wal-Mart Stores Inc., una cadena de supermercados y retail con origen en Bentonville, Arkansas, Estados Unidos en el año 1969. Es además una de las principales empresas en la industria del retail en el país, su actividad principal se caracteriza por la operación de venta de alimentos y mercaderías generales en distintos formatos de supermercados a lo largo del territorio nacional, dentro de ellos encontramos:

- **Líder:** Posee una gran variedad de productos a la venta, desde alimentos hasta mercaderías de tipo hogar, juguetería, vestuario, electrónica, entre otros. Cuenta con productos de marcas propias, importados y nacionales.
- **Express de Líder:** Formato que busca ofrecer compras rápida y fácil para entregar comodidad a clientes cerca de sus hogares, se enfoca principalmente en alimentos perecederos y comidas preparadas.
- **Líder.cl:** Sistema de compra online del formato líder, se caracteriza por ofrecer productos de forma rápida, cómoda y segura, trabajando por el cumplimiento de entregas según lo prometido.
- **SuperBodega aCuenta:** Se caracteriza por ofrecer precios aún más bajos que los demás formatos de líder a través de operaciones y ambientes más simples, también conteniendo variedad de surtido de productos.
- **Central Mayorista:** Este formato es exclusivo para comerciantes, elaborado para abastecer sus negocios a los mejores precios, encontrando todo en un solo lugar, de manera cómoda, sencilla y rápida.

Walmart Chile y todos sus formatos se caracterizan por su misión que consiste en: Ahorrarle dinero a sus clientes para que puedan vivir mejor, buscando constantemente diferenciarse con esta estrategia de sus competidores.

3.2 Contexto del problema:

Walmart se divide en 15 gerencias, una de ellas corresponde a la gerencia Omnichannel Supply Chain, la cual cuenta con el área de replenishment & E2E Flow omnicanal o también llamada reabastecimiento, esta área se encarga de a través de distintos flujos de abastecimiento, efectuar la llegada de mercadería a los locales, definiendo el cómo, cuándo y cuánto de esta. Dentro del área existen 3 segmentos, ACP (alimentos no perecederos o de larga vida útil), GM (mercancías generales) y PPS (alimentos perecederos). Durante este informe nos enfocaremos en el segmento de PPS o perecederos, ya que es el cual en que me encuentro desempeñando.

Para más claridad del segmento se presenta un organigrama de las personas que lo componen y sus cargos.

PPS se divide en distintas categorías de alimentos, tales como proteínas, lácteos, huevos, panadería, platos preparados, pastelería, verduras, frutas y congelados. Además, se divide en equipo de fulfillment y demanda, los cuales contienen distintos analistas encargados del abastecimiento de sus categorías en los distintos locales a lo largo del país.

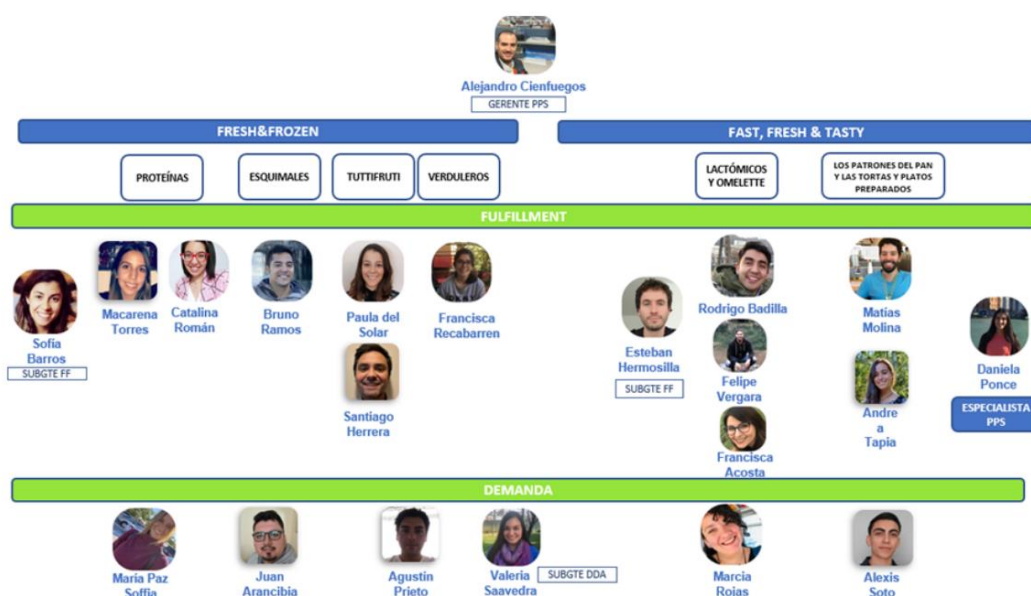


Ilustración 1: Organigrama segmento PPS del área de reabastecimiento en Walmart Chile. (Elaboración y propiedad de Walmart Chile)

Para entender de mejor manera el problema encontrado se deben explicar los distintos tipos de flujos de abastecimiento que posee Walmart para sus locales:

- **Flujo Directo:** Los proveedores distribuyen la mercadería directamente en los locales, encargándose de esta gestión.
- **Flujos Centralizados:** Son aquellos en que la mercadería de los proveedores tiene un paso por el centro de distribución de Walmart, existen 2 principales tipos de este flujo:

- Flujo Inventario: Los proveedores distribuyen la mercadería en el centro de distribución de Walmart, donde será almacenado y gestionado por el área de reabastecimiento cuándo se despachará esta mercadería, a qué locales y en qué cantidad.
- Flujo Continuo: Los proveedores distribuyen la mercadería en el centro de distribución de Walmart, pero esta viene predestinada por parte de los proveedores para ser despachada a ciertos locales, por lo que su paso por el centro de distribución se considera más breve.

En este informe nos enfocaremos en el flujo centralizado de tipo inventario, ya que es el flujo en que se encontró el problema, específicamente en la etapa del flujo en que los productos se encuentran almacenados en el CD² y deben ser gestionados para ser despachados a locales.

Esta etapa del flujo la definimos como proceso de outbound de carga del CD y a continuación se presenta un diagrama de este.

El proceso se efectúa por dos actores, el área de reabastecimiento y el centro de distribución, reabastecimiento declara semanalmente al CD la cantidad de cajas de mercadería que deben ser surtidas en este, esto se hace en base a una estimación subjetiva de los especialistas de cada categoría revisando indicadores y datos históricos, el término surtir refiere a preparar las cajas para ser despachadas a locales, de esta manera el CD se dota en recursos humanos para la semana, luego reabastecimiento realiza diariamente durante la semana la carga de pedidos, la cual refiere a la cantidad de cajas que efectivamente se deben despachar a locales según la necesidad, a partir de esto el personal contratado por el CD, llamados surtidores, se dedica a preparar las cajas que se deben despachar a locales.

Si las cajas de la carga de pedido logran ser surtidas en el mismo día, entonces el proceso se da por finalizado, si este no se logra, entonces quedan cajas pendientes para el siguiente día hábil por surtir.

² Centro de distribución.

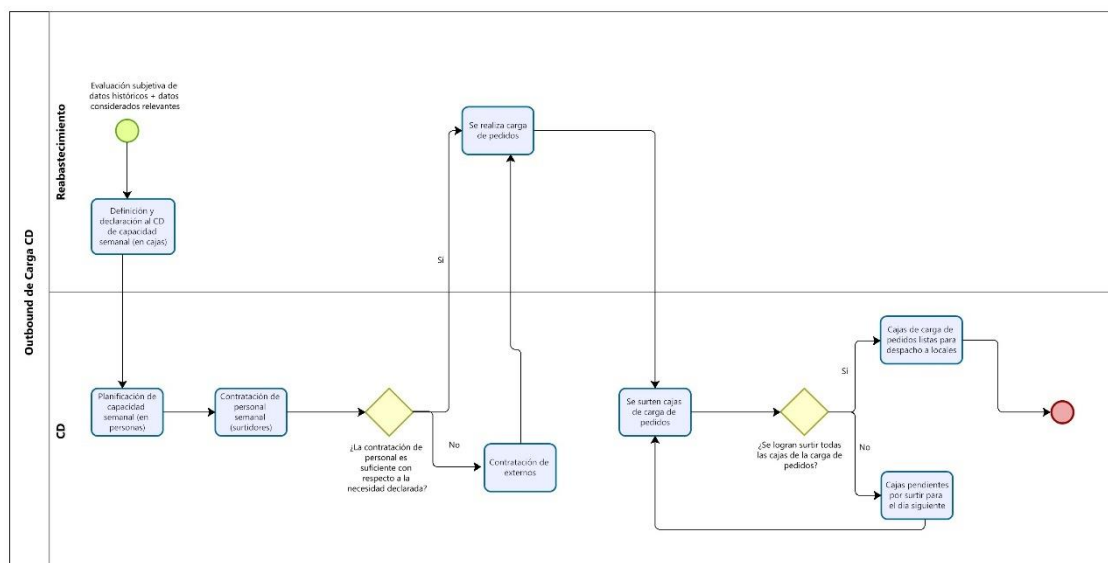


Ilustración 2: Diagrama de flujo del proceso de outbound de carga del CD 6011 Walmart Chile. (Elaboración propia con datos obtenidos de Walmart Chile)

Para efectos de este proyecto se trabajará con las categorías de fiambrería y lácteos, ya que representan el mayor porcentaje de la capacidad otorgada en el proceso, con el 37%, debido a su peso de venta y volumen dentro de la compañía.

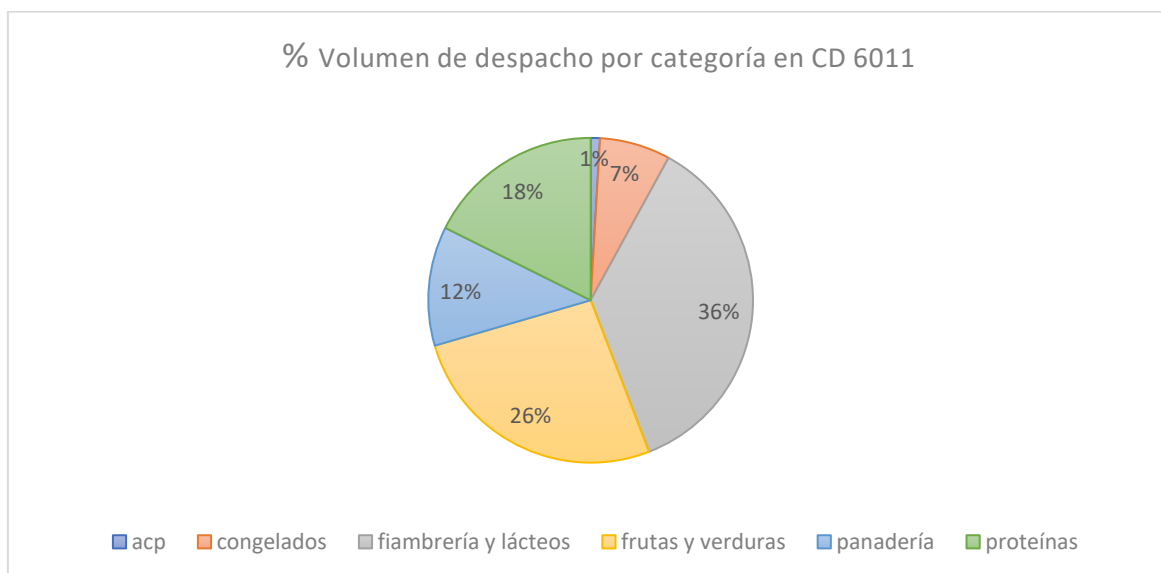


Gráfico 1: % de volumen de carga de despacho del CD 6011 Walmart Chile. (Elaboración propia con datos obtenidos de Walmart Chile)

4. Problema

4.1 Problema:

Durante un período de 17 semanas en el segundo semestre del presente año (semanas 27 a 44) la productividad del personal en el proceso de surtido de cajas en el centro de distribución 6011 de Walmart, ha presentado diferencias del objetivo de 21 cajas/hr de la empresa, representando un promedio de 13 cajas/hr por surtidor.³

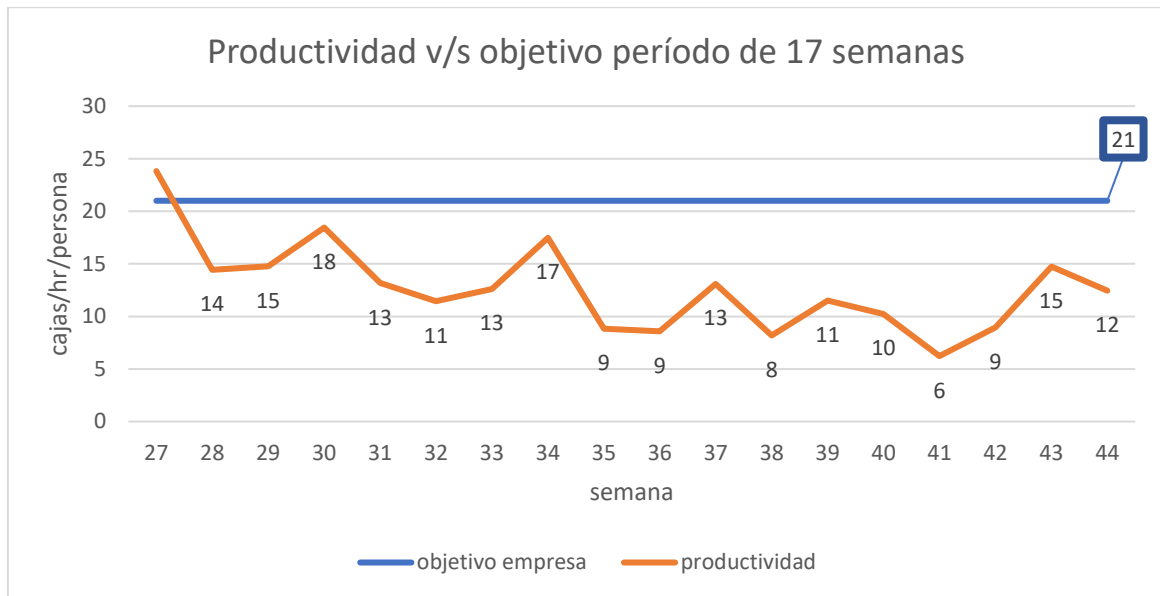


Gráfico 2: Productividad v/s objetivo empresarial en el proceso de surtido de cajas despacho en CD 6011 Walmart Chile. (Elaboración propia con datos obtenidos de Walmart Chile)

Además, la productividad diaria del personal varía semanalmente de manera contraria a la necesidad, generando falta de capacidad los últimos días de la semana y excesos de capacidad los primeros días de la semana como se aprecia en el siguiente gráfico.

³ Todos los datos de cifras de productividad han sido ponderados por un factor por aspectos de confidencialidad de Walmart Chile, por lo que no representan la realidad, pero sí una similitud directamente proporcional.

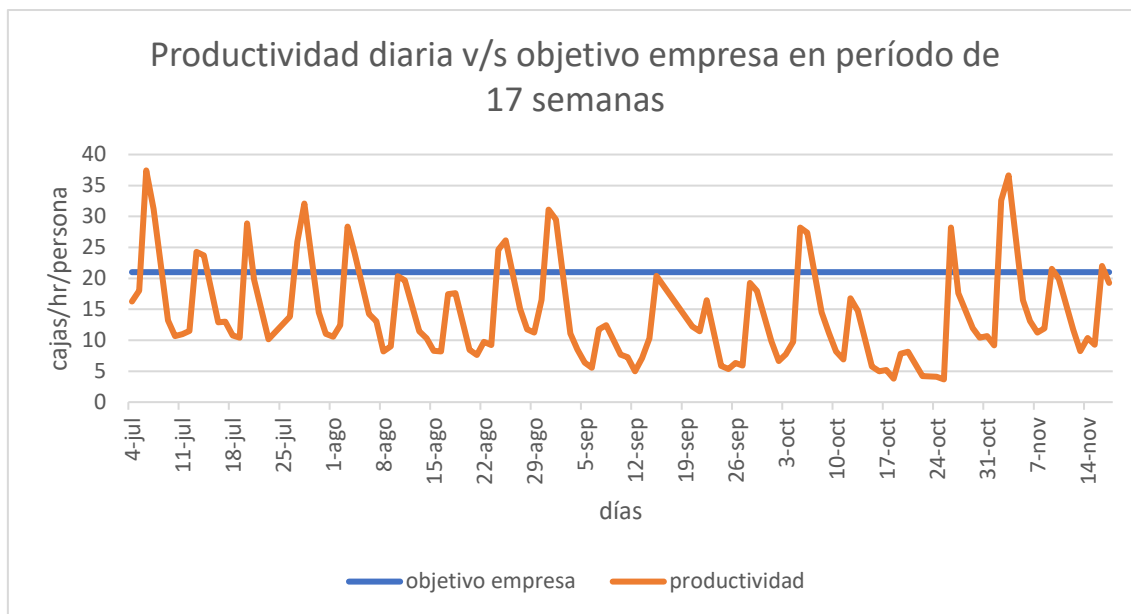


Gráfico 3: Productividad v/s objetivo empresarial en el proceso de surtido de cajas despacho en CD 6011 Walmart Chile a nivel diario. (Elaboración propia con datos obtenidos de Walmart Chile)

Se puede evidenciar que, si bien a nivel semanal el promedio de productividad es bajo, existen días donde el promedio supera la productividad objetivo, específicamente en el 15% de los días estudiado cuando sobrepasa la productividad máxima objetivo, el promedio es de 28 cajas por hora por surtidor.

Sin embargo, esto no quiere decir que sea bueno o entregue mayor eficiencia al proceso, ya que el CD se prepara para surtir 21 cajas/hr por surtidor, por lo que este número refleja el caso de quedar con cajas pendientes para el día siguiente, atrasando finalmente las entregas a locales, surgiendo entonces un problema secundario al estudiado que se refleja en las cajas pendientes por surtir que amenazan la disponibilidad en locales, en el gráfico se ve como 20 días el CD quedó con cajas pendientes por surtir en el período estudiado.

Es por esto por lo que se estudia si afecta en la disponibilidad, pero el indicador Instock⁴ (que mide la disponibilidad) no parece verse mayormente afectado, ya que ha mantenido en promedio un 95% de este durante las semanas estudiadas, lo que se encuentra dentro del objetivo de la compañía, por lo que no ahondaremos en esa arista del problema, que bien si es relevante de haber mencionado.

⁴ Fórmula Instock en Anexo A.

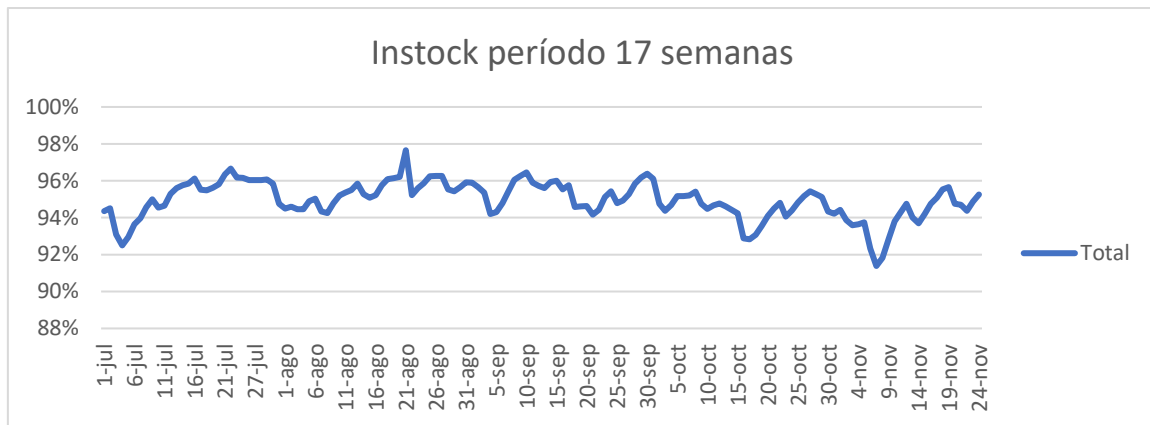


Gráfico 4: Instock evolutivo entre 17 semanas del año 2023. (Elaboración propia con datos obtenidos de Walmart Chile)

Se entiende del problema entonces que, la capacidad supera a la necesidad, dado que la productividad queda semanalmente en el 94% de los casos a la baja (solo 1 de 17 semanas estudiadas quedó por sobre cómo se ve en el gráfico), la capacidad de personal para la que se planifica el CD se vuelve mayor a la necesidad real de cajas por despachar, siendo la diferencia porcentual promedio entre estas dos curvas de un 39%.

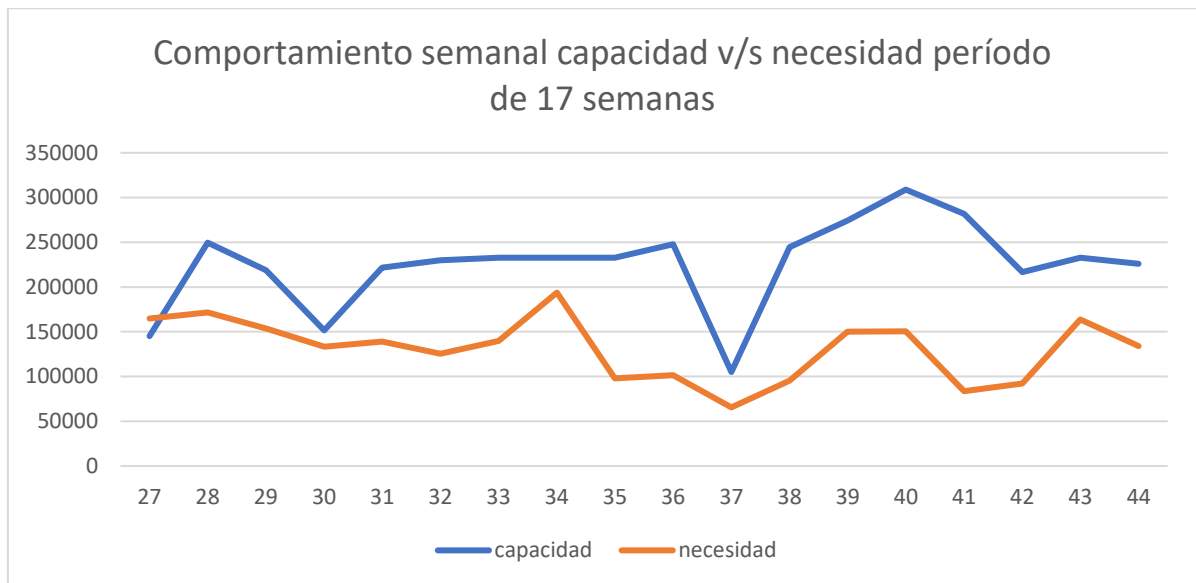


Gráfico 5: Comportamiento capacidad planificada v/s necesidad real en el proceso de surtido de cajas para despacho a locales en CD 6011 Walmart Chile a nivel semanal. (Elaboración propia con datos obtenidos de Walmart Chile)

En relación con todo lo anterior y a la productividad baja, vemos como se vuelve innecesaria la sobre contratación de personal, donde en promedio se contrata a 24 surtidores (personal) más de lo necesario, en el siguiente gráfico se aprecia como la curva de personal contratado se mantiene mayormente alejada del personal realmente necesario.

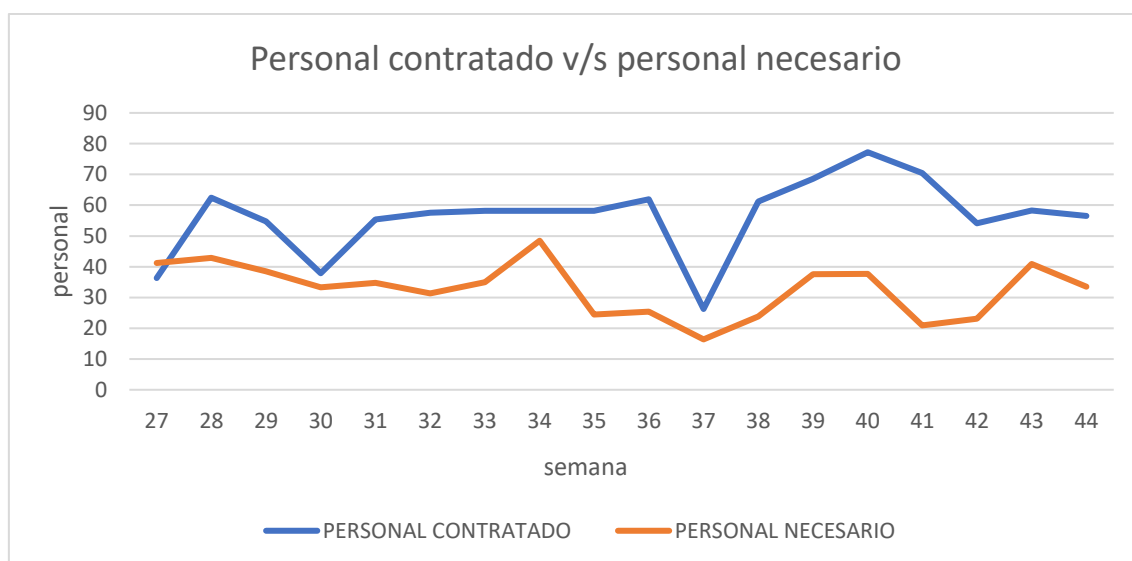


Gráfico 5: Personal contratado v/s necesario en un período de 17 semanas en el proceso de outbound de carga en CD 6011 Walmart Chile a nivel semanal. (Elaboración propia con datos obtenidos de Walmart Chile)

Para confirmar la utilización del promedio como medida de tendencia central para el análisis de estos datos, se ha elaborado una prueba estadística de normalidad que comprueba que los datos que representan la productividad se comportan de forma de distribución normal, el valor p que se obtiene es de 0.3691⁵, esto quiere decir que es mayor a 0.05 y no hay evidencia suficiente para poder rechazar la hipótesis nula de normalidad.

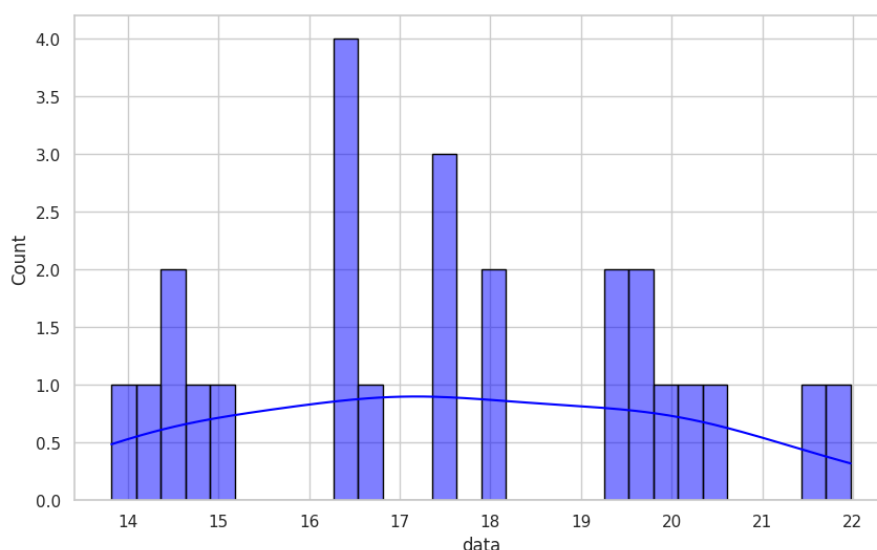


Gráfico 6: Comportamiento distribución normal datos de productividad. (Elaboración propia con datos obtenidos de Walmart Chile)

⁵ Test de normalidad en Anexo B.

4.2 Impacto económico del problema

La baja de productividad en el CD se traduce en pérdidas monetarias debido a que se contrata más personal del necesario para la semana, por términos de confidencialidad de la empresa no se puede reflejar los montos de pérdida precisos, por lo que se presentan en formato de % de dinero del total de los sueldos de la semana.

Semana	% De pérdida
28	-31%
29	-30%
30	-12%
31	-33%
33	-47%
34	-40%
35	-17%

Tabla 1: % Pérdida monetaria por baja productividad

En la tabla se aprecia la pérdida monetaria por semana, esta fue calculada en base al % de sueldos del total de estos que no resultaron necesarios de pagar debido a la sobrecontratación de personal.

El promedio de las últimas semanas analizadas corresponde a un 30% de pérdida del total de sueldos, es decir el 30% del dinero del total semanal de remuneraciones no debía ser necesariamente gastado.

El otro impacto secundario, son los quiebres que se pueden llegar a producir por los días en que el personal de surtidores es menor a la necesidad de cajas, pero al cuantificarlo descubrimos que no representa una mayor amenaza en el indicador de disponibilidad, por lo que nos centraremos en el principal impacto de baja de productividad.

5. Objetivos y medidas de desempeño

5.1 Objetivo general:

El objetivo de este proyecto consistirá en aumentar la productividad del proceso de surtido de cajas para despacho a locales en el centro de distribución, desde 13 cajas/hr por persona a 19 cajas/hr, es decir en un 28% en un período de 5 meses. Cabe destacar y recordar que los números de productividad utilizados en este informe han sido ponderados por un factor por condiciones de confidencialidad de la empresa.

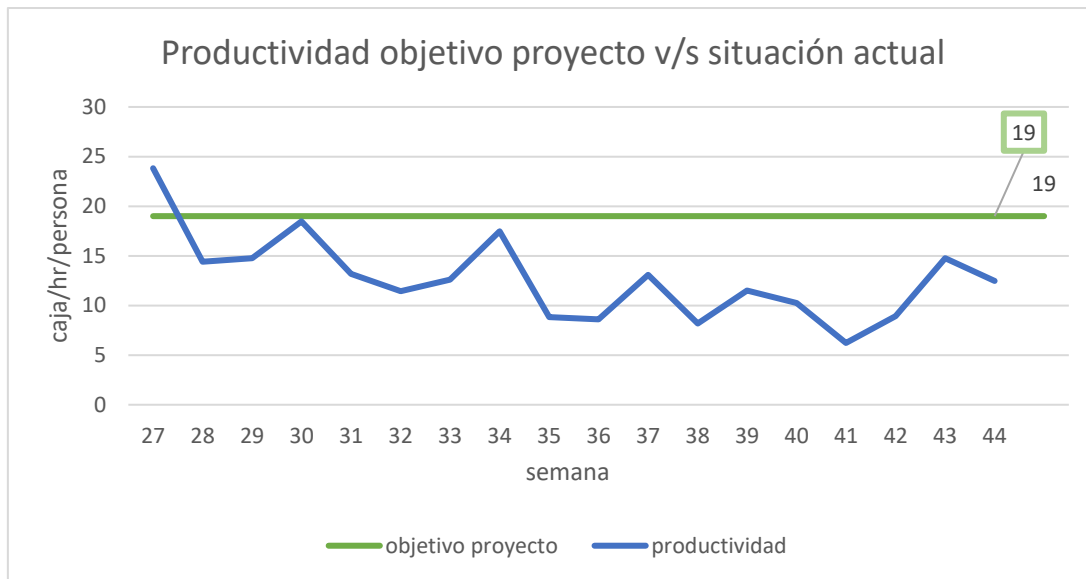


Gráfico 7: Productividad v/s objetivo proyecto en el proceso de surtido de cajas despacho en CD 6011 Walmart Chile a nivel diario. (Elaboración propia con datos obtenidos de Walmart Chile)

5.2 Objetivos específicos:

1. Disminuir la diferencia entre la capacidad planificada del CD y la necesidad de cajas por despachar en un 24% (Disminuir de 39% a 15%)
2. Disminuir la cantidad de personal innecesario por sobre contratación en un 50%.

5.3 Medidas de desempeño:

a. Productividad de un surtidor por hora

$$\frac{\text{productividad surtidor}}{\text{hora}} = \frac{\text{productividad efectiva total/hr}}{\text{cantidad de surtidores contratados}}$$

b. Diferencia % entre capacidad declarada de cajas y necesidad efectiva de cajas

$$\text{diferencia} = \frac{\text{capacidad planificada} - \text{necesidad efectiva}}{\text{necesidad efectiva}} * 100$$

c. Cantidad de personal sobre contratado

$$\text{personal contratado} = \text{personal contratado} - \text{personal necesario}$$

$$\text{personal necesario} = \text{necesidad real}/4000^6$$

⁶ Se calcula que un surtidor prepara una cantidad de 4000 cajas semanales.

6. Análisis Causas del problema

Con el fin de identificar las causas raíz de este problema y así abordar una posible solución, se realiza un diagrama con referencias del diagrama de “los 5 por qué”. (Academy, 2016)

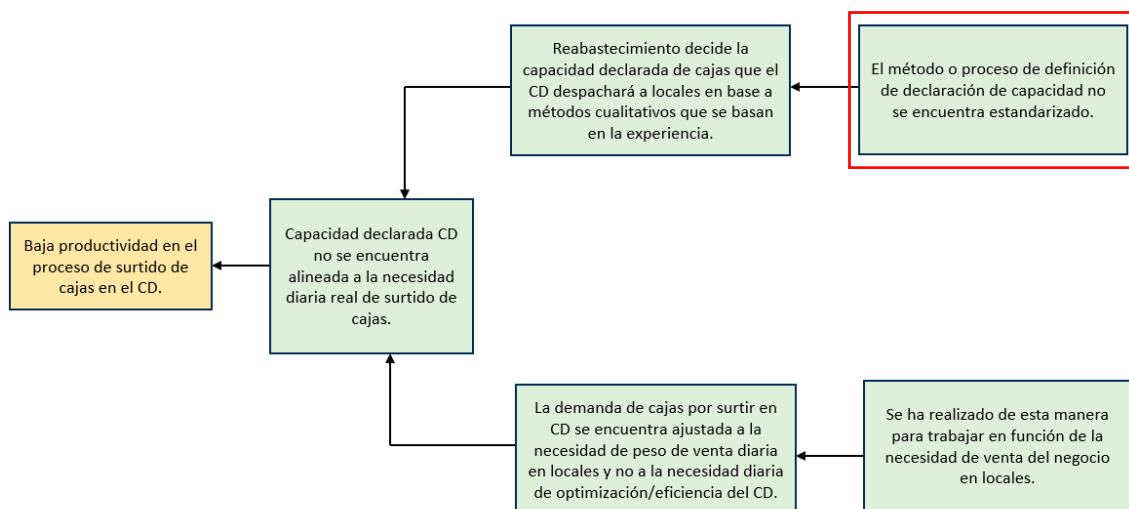


Ilustración 3: Diagrama de los 5 por qué sobre causas del problema. (Elaboración propia con datos obtenidos de Walmart Chile)

En este proyecto atacaremos principalmente la causa de la capacidad que se encuentra desalineada a la necesidad, esto quiere decir que el área de reabastecimiento entrega una declaración de capacidad de cajas al CD desviada de la necesidad efectiva real del día a día de cajas que deben ser surtidas para despacho, investigando en estas causas se encuentra que la decisión se elabora de manera subjetiva como de hecho se puede evidenciar en el proceso mencionado anteriormente. La determinación subjetiva cualitativa de este dato de declaración de cajas a despachar en una semana al CD se da porque nunca se ha definido un proceso o método estandarizado de cómo tomar la decisión, por lo que cada persona que se ve enfrentada a la situación utiliza sus propias maneras, siendo la principal de ellas la experiencia o “intuición en base a proyecciones”, lo que nos da a entender que no se sustenta en un dato firme objetivo.

Validando lo anterior, se puede ver en el gráfico como la capacidad de cajas con la que cuenta el CD no conversa con la necesidad real de cajas que deben despacharse día a día, representando en promedio una diferencia del 39%.

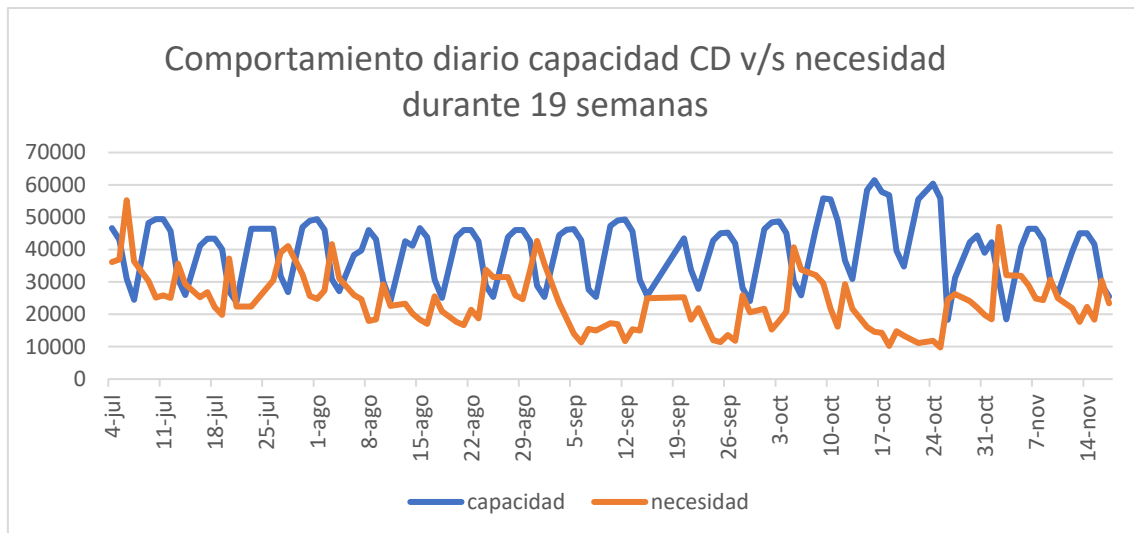


Gráfico 7: Capacidad en cajas v/s necesidad real de despacho diaria en cajas el proceso de surtido de cajas despacho en CD 6011 Walmart Chile en un período de 19 semanas. (Elaboración propia con datos obtenidos de Walmart Chile)

Además, no se selecciona la causa acerca de la demanda porque esta hace referencia a que la demanda no se mueve de forma lineal durante la semana, como si lo hace la capacidad, sino que se mueve en base a la venta. Esto no lo buscaremos modificar, ya que la venta es el CORE del negocio, lo que lo guía y en base a lo que se toman decisiones, por lo que nos adaptaremos a la venta en vez de buscar cambiarla o adaptarla a la curva de capacidad, así preservando los valores de la compañía.

7. Estado del arte

Recordando la situación actual mencionada en la explicación del proceso y en las causas, actualmente en la empresa la medida que se toma para apalancar la situación, es brindarle foco a la declaración de la capacidad, es decir cuantas cajas esperan despachar a través del CD para enviar a locales por parte de reabastecimiento al CD, sin embargo, esta declaración se realiza de manera subjetiva según cada subgerente especialista analizando volúmenes de cajas despachados semanas anteriores cualitativamente en base a intuición de experiencia en el negocio.

7.1 Revisión literatura:

Para adaptarnos a la declaración de capacidad u “oferta” incierta existen métodos subjetivos de pronósticos como es el caso de Walmart en el área de reabastecimiento a la hora de declarar la suma de cajas que esperan enviar del centro de distribución a los locales, de hecho, esto ha sido estudiado en la literatura y cabe dentro de la categoría del método de “Delphi”, el cual hace referencia a que decidir, “Se basa en solicitar las opiniones de los expertos. La diferencia radica en la forma en la que se combinan las opiniones individuales.” (Nahmias, 2007)

Sin embargo, en la literatura surgen los métodos objetivos de pronóstico en búsqueda de mayor precisión y asertividad de la demanda, a través de análisis de datos. En el libro Análisis de la producción y las operaciones (Nahmias, 2007), encontramos métodos de series de tiempo, donde se utiliza valores pasados en cuanto al evento que deseamos predecir y existen

otros métodos de tipo modelo causal, donde se usan datos de distintas fuentes que podrían encontrarse vinculados al evento por pronosticar.

Dentro de los métodos de series de tiempo encontramos varios tipos como:

i. Promedios móviles:

$$F_t = \left(\frac{1}{N}\right) \sum_{i=t-N}^{t-1} D_i = \left(\frac{1}{N}\right)(D_{t-1} + D_{t-2} + \dots + D_{t-N})$$

F_t : Pronóstico de demanda para el período t

D_{t-N} : Demanda para el período t-N

(Nahmias, 2007)

ii. Suavizamiento exponencial:

$$F_t = \alpha D_{t-1} + (1 - \alpha)F_{t-1}$$

F_t : Pronóstico de demanda para el período t

D_{t-N} : Demanda para el período t-N

α : Constante de suavizamiento [0,1]

(Nahmias, 2007)

Encontramos ventajas y desventajas en cada uno, por ejemplo, en los promedios móviles un valor atípico se eliminará luego de los N periodos, sin embargo, permanecerá en suavizamiento exponencial, por otro lado, para usar los promedios móviles se deben guardar todos los N datos anteriores mientras que para el suavizamiento exponencial solo necesitamos mantener el último pronóstico.

Ahora, hay métodos que consideran tendencia en los datos, permitiéndose adaptar a series con tendencias definidas, tal como:

iii. Suavizamiento exponencial doble o Método de Holt:

$$S_t = \alpha D_t + (1 - \alpha)(S_t + G_{t-1})$$

$$G_t = \beta(S_t - S_{t-1}) + (1 - \beta)G_{t-1}$$

$$F_{t,t+\tau} = S_t + \tau G_t$$

D_t : Demanda para el período t

S_t : Valor de la intercepción en el tiempo t

G_t : Valor de la pendiente en el tiempo t

α : Constante de suavizamiento [0,1]

β : Constante de suavizamiento para la tendencia [0,1]

(Nahmias, 2007)

Existen también métodos como los estacionales que resultan convenientes para series con patrones que se repiten cada N períodos, tales como:

iv. Holt-Winters aditivo (para patrón estacional constante):

$$S_t = \alpha \left(\frac{D_t}{C_{1-N}} \right) + (1 - \alpha)(S_{t-1} + G_{t-1})$$

$$G_t = \beta[S_t - S_{t-1}] + (1 - \beta)G_{t-1}$$

$$C_t = \gamma \left(\frac{D_t}{S_t} \right) + (1 - \gamma)C_{t-N}$$

$$F_{t,t+\tau} = (S_t + \tau G_t)_{C_{t+\tau-N}}$$

D_t : Demanda para el período t

S_t : Valor de la serie

G_t : Valor de la tendencia

C_t : Valor de estacionalidad

α : Constante de suavizamiento $[0,1]$

β : Constante de suavizamiento para la tendencia $[0,1]$

(Nahmias, 2007)

También encontramos métodos más complejos como ARIMA que dados su naturaleza, se puede adaptar a datos más complejos otorgando mayor flexibilidad.

v. ARIMA

Permite describir un valor como una función lineal de datos anteriores y errores debidos al azar, también adiciona el componente cíclico. El modelo ARIMA (p, d, q) posee tres componentes AR (autorregresivo), I (integrado) y MA (medias móviles)

La parte izquierda de la ecuación representa el componente autorregresivo y el componente de diferenciación, mientras que la parte derecha representa el componente de media móvil. La idea central de ARIMA es ajustar un modelo autorregresivo y de media móvil a la serie temporal después de aplicar la diferenciación para hacerla estacionaria

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d \gamma_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t$$

γ_t : serie temporal en el tiempo t .

B : operador de diferencia (reduciendo la serie temporal a $\gamma_t - \gamma_{t-1}$)

P : orden del componente autorregresivo (AR)

d : orden de diferenciación

q : orden del componente de media móvil (MA)

$\phi_1, \phi_2, \dots, \phi_p$ son los coeficientes autorregresivos.

$\theta_1, \theta_2, \dots, \theta_q$ son los coeficientes de media móvil

ε_t : termino de error en el tiempo t

Existen además métodos de carácter más complejo que incluyen variables exógenas, que pueden ser ventas, promociones, entre otros elementos, se conocen como modelos dinámicos de series temporales que logran describir un valor como una función de línea, algunos de estos métodos de Box-Jenkins son conocidos como ARIMAX, SARIMAX.

vi. ARIMAX

ARIMAX, además de las variables endógenas (las propias series temporales), se incorporan variables exógenas que afectan la serie temporal de interés, es decir se podría utilizar una variable como promociones de precio, feriados, entre otros.

Supongamos que tenemos una serie temporal y_t y k variables exógenas $X_{1t}, X_{2t}, \dots, X_{kt}$,

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_k X_{kt} + \varepsilon_t$$

y_t : serie temporal en el tiempo t .

c : constante

$\phi_1, \phi_2, \dots, \phi_p$ son los coeficientes autorregresivos.

$\theta_1, \theta_2, \dots, \theta_q$ son los coeficientes de media móvil

ε_t : termino de error en el tiempo t

$X_{1t}, X_{2t}, \dots, X_{kt}$: variables exógenas

$\beta_1, \beta_2, \beta_k$: coeficientes asociados a las variables exógenas

En Nahmias se explica que la determinación del modelo adecuado depende de las características de las observaciones como el contexto en el que se requieren los pronósticos. (Nahmias, 2007). También se aclara que no necesariamente un pronóstico debe ser extremadamente sofisticado para cumplir su propósito.

“Los resultados provenientes de la literatura sugieren que los métodos más sencillos son con frecuencia tan exactos como los más sofisticados. Armstrong (1984) revisó 25 años de casos de estudios de pronósticos para averiguar si los métodos sofisticados funcionan mejor. Al comparar los resultados de 39 casos de estudio, descubrió que en 20 casos los métodos sofisticados funcionaron prácticamente de la misma manera que los sencillos, en 11 casos superaron a los sencillos y en 7 resultaron significativamente peores” (Nahmias, 2007)

Además, en la literatura estos métodos de pronósticos han medido su desempeño y vinculado a inventarios a través de la distribución del error de los pronósticos, según esta distribución obtenemos una media y una varianza y desviación estándar.

Se puede calcular el stock de seguridad realizando la multiplicación de un intervalo de confianza por la desviación estándar del error del pronóstico, según el libro mencionado durante este informe, la desviación estándar la obtenemos de la raíz cuadrada de la varianza.

Varianza del error de pronóstico es:

$$\sigma_e^2 = \sigma^2 \left(\frac{2}{2 - \alpha} \right)$$

(Nahmias, 2007)

7.2 Revisión de la industria

En la industria encontramos diversas formas de abordar la capacidad limitada en los centros de distribución y aun así seguir cumpliendo con la demanda incierta del mercado. Por un lado, encontramos a Amazon que destaca globalmente por su gama de robots que participan en el proceso de picking aumentando la capacidad de almacenamiento de los centros y a su vez el de picking, “Los robots mejoran la eficiencia y la seguridad en los centros logísticos. Permiten almacenar un 40% más de inventario, lo que a su vez hace más fácil entregar los pedidos de Amazon Prime y otros pedidos a tiempo, ya que es menos probable que un artículo se agote” (Amazon, 2019). En otras empresas como Falabella, el gerente de la cadena de suministro comenta: “esta ampliación nos ha permitido aumentar un 50% la capacidad instalada del Centro de Distribución” siendo la principal inversión el camino a un almacén tecnológico, añadiendo a la operación 132 robots “Shuttle” (Logistec, 2022)). Por último, el mismo Walmart Chile pero en otros centros de distribución (distintos al de este informe) para otras divisiones de productos, ha instalado robots autónomos donde la empresa afirma que se cuadruplicará de operación logística, sobre todo de cara al foco e-commerce, Walmart afirma: “Utilizará 16 robots autónomos para eliminar el desplazamiento redundante de los trabajadores, mejorar la precisión del picking y reducir la intensidad del trabajo, todo ello impulsado por inteligencia artificial, y que representará un aumento de productividad del 400% solo en su primera etapa.” (Walmart, 2022)

Por otro lado, en la industria se ve como empresas de retail reconocidas como Zara adoptan sistemas de inventario que se adaptan ágilmente a la demanda implementando extra esfuerzos en capacidad, de esta forma se adaptan a la demanda incierta a través de mayores refuerzos en recursos, con el ideal de enfocar sus esfuerzos y presupuesto en el aceleramiento de la cadena de suministro, el sistema de Zara de producción rápida y amplio gasto en capacidad les ha permitido adaptarse a las fluctuantes demandas del mercado, entendiendo que más capacidad y más esfuerzos y presupuesto en esta llevarán a una mejor adaptabilidad a la demanda variante a la que se exponen los retailers cada año. (Harvard, 2004)

También en la industria encontramos empresas que se han caracterizado por adaptarse exitosamente a la demanda sin excesiva variación de capacidad implementando modelos de machine learning y similares, Unilever afirma escalar en eficiencia gracias a estas prácticas (Unilever, 2023). Por último, Amazon logra su reconocida reputación a través de una gestión de inventario basado en machine learning, lo que lo ha llevado incluso a ofrecer este servicio. (Amazon, 2023). El Machine learning aprende de la historia y sus eventos, transformándolo así en una poderosa herramienta para abordar problemas complejos desarrollando respuestas eficaces y eficientes a problemáticas.

Por último, en estudios sobre industria de retail se ve el uso de APIs y softwares como Prophet, que pueden realizar funciones de pronósticos desde series sencillas hasta complejas, usando modelos tradicionales a algunos más complejos que consideran más variables como ha sido

mencionado en la literatura. (Comparing Prophet and Deep Learning to ARIMA in forecasting Wholesale Food prices)

8. Solución

8.1 Alternativas de solución:

- I. Implementar un modelo avanzado de machine learning que prediga correctamente la cantidad de cajas que deben ser despachadas para enviar a locales considerando las restricciones de capacidad y la restricción de satisfacer correctamente la demanda semanal, además este modelo debería considerar las sensibilidades de vida útil de las distintas categorías, ya que el machine learning puede aprender de la historia y sus eventos, esto le generaría insights para sí mismo que le permitan perfeccionarse, como se comentó en el estado del arte, existen diversas empresas que utilizan estos modelos y de hecho el mismo Walmart lo utiliza para otras tareas y procesos, como por ejemplo “SMART” que fue elaborado por la empresa como software que utiliza machine learning para predecir correctamente la demanda en locales, analizando eventos pasados y diversos inputs que permiten obtener resultados precisos que son de relevante importancia para la correcta gestión de la cadena de suministro.
- II. Se propone utilizar robots en el sistema de picking como los mencionados que utilizan distintas empresas de la industria y rubro, específicamente se proponen los robots del sistema “Goods to person”, corresponde a una metodología en la que los robots llevan productos a las personas como bien dice su nombre en inglés, estos AMRs (autonomous mobile robots) permiten que las personas puedan preparar pedidos/cajas sin la necesidad del desgaste del desplazamiento, incrementando de esta manera la eficiencia del CD y así la capacidad, lo que permitiría aumentar la capacidad los días que se necesita mayor productividad en el proceso de surtido de cajas y los días que se requiere menor capacidad no se incurriría en mayores gastos, ya que la inversión se realizaría una vez y se retribuiría en el tiempo, representando solo beneficios a largo plazo resolviendo el problema de la capacidad limitada distribuida en categorías.
- III. Por último, se propone utilizar un método de pronóstico objetivo como son los métodos de series de tiempo y causales vistos en el estado del arte, para declarar las cajas que deben ser despachadas semanal y detalle diario al CD y a su vez entregando un stock de seguridad en base a los pronósticos, de esta manera estandarizando esta parte proceso y quitándole la subjetividad de la decisión que actualmente lleva a niveles de error de en promedio un 39% como se mencionó en el contexto del problema. Se utilizarían los métodos vistos en el estado del arte y se compararían entre ellos, evaluando cual representa un menor error para la solución. Para realizar todo esto se empleará lo aprendido en el curso de gestión de operaciones en cuanto a métodos de pronósticos, además de herramientas computacionales como Python y R aprendidas en la universidad.

8.2 Matriz de decisión para la solución

Solución / Criterio	Costo	Complejidad implementación	Relación con la carrera y área de trabajo	Rapidez resultados	¿Ajusta la planificación de la declaración de cajas al CD?
Modelo de machine learning que planifique y pronostique las cajas a despachar del CD a locales.	Medio (2)	Alto (1)	Medio (2)	Medio (2)	Alto (3)
AMRs GTP (Goods to person)	Alto (1)	Alto (1)	Bajo (1)	Bajo (1)	Bajo (1)
Método objetivo de pronóstico de series de tiempo o causalidad.	Bajo (3)	Medio (2)	Alto (3)	Alto (3)	Medio (2)

Tabla 2: Matriz de decisión de la solución. (Elaboración propia)

Se definen los criterios presentados en la matriz para decidir que solución utilizar: Costo de la solución, complejidad de la implementación, relación con la carrera y área de trabajo, rapidez en ver resultados y si ataca la causa principal del problema.

Costo de la solución haciendo referencia al precio en términos de implementación del proyecto, complejidad de implementación refiriéndose a que tan posible de realizar por el estudiante se aprecia el problema, relación con la carrera y área de trabajo, rapidez de resultados y finalmente si ajusta la precisión de la declaración de cajas por despachar al CD, que viene siendo la causa raíz atacada.

Se definen puntos para cada criterio: 1 el peor escenario, 2 el escenario medio y 3 el mejor escenario en base a la necesidad del problema.

Se distribuye entonces, por ejemplo, un 3 a costo bajo pero un 1 a rapidez de los resultados, porque los números se trabajan en pos de buscar la mayor suma para las soluciones que representen el mejor escenario para cada criterio.

Así se obtienen los siguientes puntajes:

- AMRs goods to person: 5 pts
- Modelo de machine learning: 10 pts
- Método objetivo de pronósticos: 13 pts

Finalmente, siendo la mejor opción la solución del método objetivo de pronósticos, ya que representa altos puntajes en cuanto a costos, relación con la carrera y área de trabajo, rapidez de resultados y puntajes medios en cuanto a atacar la causa elegida y complejidad de implementación.

9. Metodología

9.1 Plan de acción:

Con el fin de realizar este proyecto, se utilizará la metodología PDCA (Planificar, Hacer, Consultar y Actuar), la elección se basa en que es un reconocido modelo de mejora continua que otorga gran flexibilidad y un enfoque práctico que conecta con el modelo Capstone de esta vía de titulación. Se siguen las siguientes etapas:

1. Planificar (Plan): Se establece lo que se busca mejorar y se establecen objetivos y fechas propuestas de cuándo lograrlos, esto ya se declara en los objetivos del presente informe.
2. Hacer (Do): Se sigue el plan de la primera etapa, se recopilan datos se analizan y ponen a pruebas diversos modelos de pronóstico que fueron estudiados en el estado del arte.
3. Verificar (Check): Evaluar los resultados obtenidos a través de los distintos modelos en base a los objetivos esperados.
4. Actuar (Act): Decidir que modelo implementar permanentemente en base a su relación con los objetivos y evaluar su desempeño.

9.2 Evaluación económica:

Para evaluar económicamente el problema se realiza primero la identificación de inversión, costos e ingresos, como inversión solo se considera el notebook comprado por la empresa para

la realización de la práctica ya que no existen inversiones mayores o de otro tipo, a este se le considera la depreciación que representa su inversión. Por el lado de los costos, se considera el sueldo de un ingeniero y la licencia de Microsoft Office 365 en la que se está incurriendo para desarrollar las diversas labores de la práctica.

Por último como ingresos se está considerando el ahorro en productividad que ahora será utilizada y anteriormente reflejaba un costo de oportunidad, para calcularla se utilizó una estimación de proyección de las nuevas capacidades declaradas en base a los objetivos esperados, donde se obtiene la nueva productividad y sus costos asociados a la sobre contratación de personal, restando los costos de oportunidad de tener personal que no está trabajando según el objetivo de productividad que representa antes y después de la solución, obtenemos el ingreso por ahorros en costo de oportunidad de productividad.⁷

	MES 0	MES 1	MES 2	MES 3	MES 4
Ingresos					
Ahorro por mejora en productividad		\$ 10,583,499	\$ 14,579,618	\$ 17,593,726	\$ 18,059,396
Costos					
Sueldo ingeniero		\$ (1,400,000)	\$ (1,400,000)	\$ (1,400,000)	\$ (1,400,000)
Microsoft Office 360		\$ (6,990)	\$ (6,990)	\$ (6,990)	\$ (6,990)
Depreciación computador		\$ (10,000)	\$ (10,000)	\$ (10,000)	\$ (10,000)
Total Costos		\$ (1,416,990)	\$ (1,416,990)	\$ (1,416,990)	\$ (1,416,990)
Resultados antes de impuestos		\$ 12,000,489	\$ 15,996,607	\$ 19,010,715	\$ 19,476,386
Impuestos		\$ (3,240,132)	\$ (4,319,084)	\$ (5,132,893)	\$ (5,258,624)
Resultado después de impuestos		\$ 8,760,357	\$ 11,677,523	\$ 13,877,822	\$ 14,217,762
Inversión	\$ (599,990)				
Flujo de caja	\$ (599,990)	\$ 8,760,357	\$ 11,677,523	\$ 13,877,822	\$ 14,217,762

Flujo de caja que representaría la solución. (Elaboración propia)

Es así como a través del flujo de caja se obtienen los siguientes indicadores de VAN y TIR a modo de evaluación de rentabilidad del flujo de caja de este proyecto:

VAN	\$40,019,273.80
TIR	1492%

Se deduce entonces que el proyecto resulta rentable, la inversión se recupera al ver el VAN positivo y la alta TIR, se entiende también que el valor tan alto del % de la TIR se presentó en base a que no existen costos ni inversión de mayor envergadura, por lo que finalmente la mayor parte del proyecto resulta en beneficios y el riesgo de pérdidas es muy bajo con retornos muy altos.

Ahora, comparando la situación se puede ver que el flujo de caja con respecto a los costos sin el proyecto es el siguiente:

	COSTO ANTES	COSTO AHORA	AHORRO	% DISMINUCIÓN
MES 1	\$ (14,454,868.75)	\$ (3,871,369.69)	\$ (10,583,499)	73%
MES 2	\$ (18,078,375.00)	\$ (3,498,757.50)	\$ (14,579,618)	81%
MES 3	\$ (20,623,356.25)	\$ (3,029,630.63)	\$ (17,593,726)	85%
MES 4	\$ (21,166,556.25)	\$ (3,107,160.00)	\$ (18,059,396)	85%

⁷ Detalle en Anexo C.

Tabla 3: Disminución de costos con proyecto. (Elaboración propia)

Vemos que en promedio los costos de oportunidad por sobrecontratación con la proyección realizada disminuyen en un 81%.

9.3 Otras evaluaciones de impactos:

- **Ético:** Se realiza un análisis ético y se concluye que este proyecto contribuye a la estandarización objetiva del proceso de declaración de cajas a despachar en el período de 1 semana al CD, por lo que al estandarizar disminuye el riesgo de los juicios personales de cada trabajador, disminuyendo así el riesgo de que posibles intereses personales o influencias puedan verse reflejadas en la honestidad de la declaración.
- **Medio ambiente:** Con respecto al impacto medio ambiental, este proyecto al calcular con mayor precisión la cantidad de cajas a ser surtidas en CD 6011 Walmart Chile, permite la mejor planificación del personal, y donde al existir productos perecibles dada la naturaleza de las categorías estudiadas, permitiría siempre tener el personal necesario para surtir y no llegar a la instancia de dejar cajas pendientes que puedan terminar mermando y generando no solo pérdidas económicas sino que una pérdida para el medio ambiente desde el punto de vista del deshecho.

9.4 Matriz de riesgo de implementación:

En esta matriz, se evalúan diferentes eventos en términos de su probabilidad de ocurrencia y el impacto que tendrían en el entorno en el que se espera aplicar la implementación de la solución, para esto se utiliza una matriz típica de evaluación de riesgos.⁸

Riesgo	Probabilidad	Impacto	Nivel de riesgo	Mitigación
Falla en SQL	Moderado (3)	Significativo (3)	Medio (9)	Utilizar respaldos y buscar fuentes de información de respaldo.
Falla en Dataiku	Moderado (3)	Significativo (3)	Medio (9)	Utilizar respaldos y buscar fuentes de información de respaldo.
Escalabilidad del problema	Poco probable (2)	Mayor (4)	Medio (8)	Dividir etapas claras de plan de acción
Problemas con fuentes de información	Poco probable (2)	Mayor (4)	Medio (8)	Analizar siempre el sentido general de los datos. Enseñar de donde se obtienen los datos y cómo se obtiene el respaldo de estos.

Tabla 4: Matriz de riesgos del proyecto. (Elaboración propia)

En base a este análisis se ve que, si bien los riesgos no representan valores graves, sí se debe tener en cuenta las mitigaciones, ya que un nivel de riesgo medio puede influir en la correcta implementación del proyecto, en conclusión, se deben mantener respaldos y analizar el sentido general de los datos.

⁸ Matriz de riesgo en Anexo D

9.5 El nuevo proceso:

Al realizar el diagrama del flujo del proceso con la implementación de la solución se puede obtener que, si bien el proceso cambia levemente, este ahorra un paso/tarea, que precisamente es la relacionada a la declaración de la capacidad de cajas a despachar a nivel semanal al CD. Esto permitiría una mejor precisión de la declaración al poseer el dato estandarizado, agilizando así la completitud del proceso.

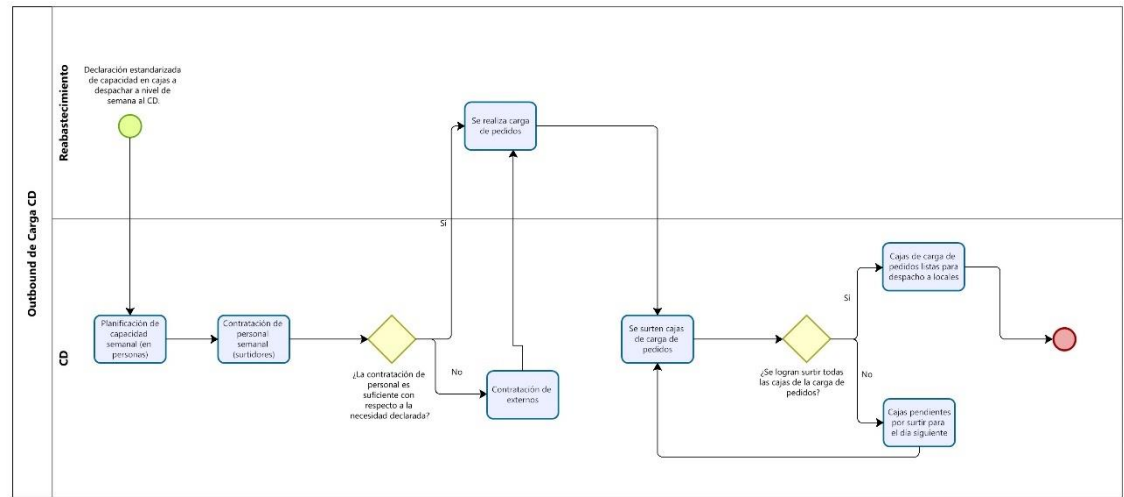


Ilustración 4: Diagrama del proceso de surtido de cajas en CD 6011 Walmart Chile con la solución. (Elaboración propia)

10. Implementación y desarrollo solución

10.1 Carta Gantt:

Según la metodología anteriormente presentada, se crea un plan para el desarrollo e implementación de la solución, se pueden ver las tareas que conlleva cada una de las 4 etapas de la metodología:

	10-sep	1-oct	15-oct	29-oct	12-nov	26-nov	10-dic
1. Planificar							
Investigar modelos de pronósticos							
Se definen objetivos y métricas							
2. Hacer							
Análisis exploratorio serie de tiempo							
Se desarrolla en Python y R studio los modelos							
3. Verificar							
Se evalúa el error de los pronósticos							
Comparación resultados entre modelos							
4. Actuar							
Decisión de modelo más acertado							
Cálculo de stock de seguridad para la declaración al CD							
Elaboración reporte automático para la entrega periódica de pronósticos							

Ilustración 5: Carta Gantt para elaboración e implementación de la solución. (Elaboración propia)

10.2 Paso a paso implementación:

1. Planificar:

Se utilizará la plataforma Jira que cuenta con diversas opciones de planificación de proyectos, se insertan ítems que corresponden a las tareas a realizar con su fecha de entrega y la aplicación lo visualiza de distintas formas.⁹

En esta etapa se estudian los métodos de pronósticos que dan resultado y se adaptan a distintas necesidades, se decide realizar un análisis exploratorio de los datos de la serie de tiempo y analizar que método de pronóstico de los investigados en el estado del arte pueden resultar convenientes para el caso.

2. Hacer:

Se decide qué métodos de series de tiempo y causales incluir para su ejecución, para esto primero se realiza un análisis exploratorio de la serie de tiempo a utilizar:

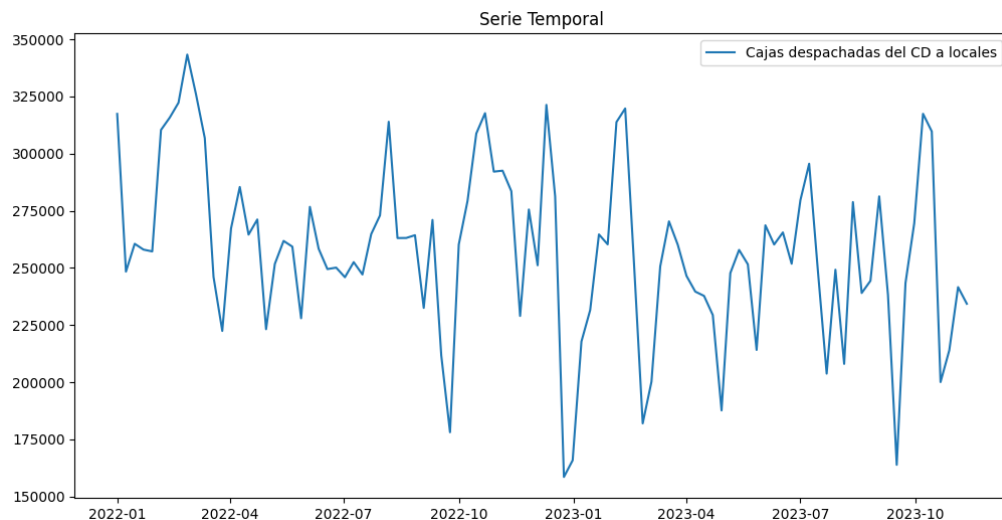


Gráfico 8: Serie temporal de dos años que representa las cajas históricamente despachadas semanalmente por el CD 6011 de Walmart Chile. (Elaboración propia)

Del gráfico se puede ver que no existe una tendencia muy clara en los datos, ya que los datos no muestran una tendencia al alza, por ejemplo, lo cual se vería representado por curvas sostenidas en una dirección, sin embargo, sí hay una moderada leve estacionalidad, se puede ver que cada un cierto período de tiempo las cajas despachadas del CD a locales disminuyen y luego vuelven a aumentar y así sucesivamente demostrando el concepto de estacionalidad, estos factores se deben apreciar en un gráfico y analizar patrones.

Se decide entonces en esta etapa que se utilizarán métodos de pronóstico de serie de datos que consideren la estacionalidad, en base a los dos más populares y estudiados en la literatura que prometen resultados, se eligen los siguientes modelos:

i. Holt-Winters aditivo (para patrón estacional constante):

⁹ Anexo E Imágenes Jira.

$$S_t = \alpha \left(\frac{D_t}{C_{1-N}} + (1 - \alpha)(S_{t-1} + G_{t-1}) \right)$$

$$G_t = \beta[S_t - S_{t-1}] + (1 - \beta)G_{t-1}$$

$$C_t = \gamma \left(\frac{D_t}{S_t} \right) + (1 - \gamma)C_{t-N}$$

$$F_{t,t+\tau} = (S_t + \tau G_t)_{C_{t+\tau-N}}$$

D_t : Demanda para el período t

S_t : Valor de la serie

G_t : Valor de la tendencia

C_t : Valor de estacionalidad

α : Constante de suavizamiento [0,1]

β : Constante de suavizamiento para la tendencia [0,1]

(Nahmias, 2007)

También encontramos métodos más complejos como ARIMA que dados su naturaleza, se puede adaptar a datos más complejos otorgando mayor flexibilidad.

ii. ARIMA

Permite describir un valor como una función lineal de datos anteriores y errores debidos al azar, también adiciona el componente cíclico. El modelo ARIMA (p, d, q) posee tres componentes AR (autorregresivo), I (integrado) y MA (medias móviles)

La parte izquierda de la ecuación representa el componente autorregresivo y el componente de diferenciación, mientras que la parte derecha representa el componente de media móvil. La idea central de ARIMA es ajustar un modelo autorregresivo y de media móvil a la serie temporal después de aplicar la diferenciación para hacerla estacionaria

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d \gamma_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t$$

γ_t : serie temporal en el tiempo t.

B : operador de diferencia (reduciendo la serie temporal a $\gamma_t - \gamma_{t-1}$)

P : orden del componente autorregresivo (AR)

d : orden de diferenciación

q : orden del componente de media móvil (MA)

$\phi_1, \phi_2, \dots, \phi_p$ son los coeficientes autorregresivos.

$\theta_1, \theta_2, \dots, \theta_q$ son los coeficientes de media móvil

ε_t : termino de error en el tiempo t

Ahora, se procede a utilizar Python y R Studio para el desarrollo de los modelos, así se consigue programar los métodos de pronósticos escogidos en Python a través de librerías y APIs¹⁰ que permitan su ejecución y además la simplifiquen.

Se utilizan principalmente 3 herramientas en Python:

- *Pandas*: para manipular y analizar datos
- *Matplotlib.pyplot*: visualizar los datos a través de gráficos
- *Statsmodels.tsa.api*: modela y analiza las series temporales para realizar los diferentes métodos de pronóstico

Así, a través de estas herramientas se ejecutan códigos¹¹ para cada método de pronósticos, los cuales se adjuntan en el anexo de este informe, así se obtienen pronósticos para las futuras 6 semanas y sus gráficos respectivos de cómo se comportaría la curva.

Se calcula el stock de seguridad en base al error del pronóstico, la media y la varianza de este, así no solo entregando un pronóstico, sino el inventario recomendado asociado a este, que viene siendo para cuánta capacidad de cajas debemos prepararnos en el centro de distribución en base a las que se planifica despachar a locales a través de este, de esta manera aseguramos entregar un valor declarado que nos entregue seguridad de no declarar una menor cantidad de cajas de la necesidad.

Inventario de seguridad = $z\sigma$

Cantidad de cajas semanales a declarar al centro de distribución = $z\sigma + \mu$

σ : desviación estándar error absoluto

μ : media (pronóstico)

Z: se utilizará un nivel de confianza de 95% en base a políticas de Walmart Chile

(Nahmias, 2007)

3. Verificar:

Con los resultados de pronósticos listos, se evalúa el error de cada uno para así definir cuál tiene menor error y definirlo como más confiable para este proyecto.

Con los resultados de los pronósticos se comienza a probar el error de cada método en comparación a la cantidad de cajas efectivamente despachadas a locales por parte del CD, esto se hizo con las siguientes formulas:

i. Error absoluto

Medida de cuanto se subestima o sobreestima el pronóstico, permite observar sencillamente qué tan preciso fue el pronóstico a través de magnitudes.

¹⁰ Pieza de código con la cual se puede obtener comunicación de diferentes aplicaciones y permite compartir funcionalidades e información.

¹¹ Anexo F códigos Python.

$$Error = |A_t - F_t|$$

F_t : Pronóstico período t

A_t : Cantidad de cajas real despachadas de CD a locales en período t

ii. MAPE

Error porcentual absoluto medio o MAPE (Mean Absolute Percentage Error), medida para calcular el tamaño del error porcentualmente, al hablar de % y no volumen facilita la interpretación.

$$MAPE = \frac{\sum |A_t - F_t|}{|A_t|} * \frac{100}{n}$$

F_t : Pronóstico período t

A_t : Cantidad de cajas real despachadas de CD a locales en período t

n : número total períodos

(Nahmias, 2007)

4. Actuar:

En base a lo anterior se decide oficialmente y en conjunto a otros factores el pronóstico a usar oficialmente.

En base al menor error entregado se evalúa ARIMA como el método con mejor performance de pronósticos.

Finalmente, se utiliza el pronóstico y su inventario de seguridad para planificar la declaración de capacidad al centro de distribución semanal de cajas que se esperan despachar a locales y se mide cómo afecta a la productividad de este, se propone realizar una visualización de este a través de reportes automáticos a través de la herramienta Dataiku¹², la cual permitiría obtener el pronóstico de las próximas seis semanas periódicamente en el correo de los analistas encargados de la declaración.

11. Resultados

11.1 Proyecciones:

Al realizar proyecciones¹³ para seis períodos, calcular el inventario de seguridad para cada método, entregar una declaración de cajas que despachar en base a la suma del pronóstico y el inventario de seguridad y luego calcular el error de cada una a través del MAPE y error absoluto indicados anteriormente, se obtuvo que con un promedio de 5% de MAPE y 17417 cajas de error absoluto el método con las menores cifras de error es ARIMA.

¹² Anexo G.

semana año	semana modelo	pronóstico ARIMA	demanda real	Error Abs	desv estandar	inv seguridad	declaración cajas CD	MAPE	Error abs
45	98	242296.79	241585	711.8	9147.0	15045.4	257342.2	-7%	15757.2
46	99	246557.95	234306	12251.9	9147.0	15045.4	261603.4	-12%	27297.4
47	100	248830.25	272211	23380.8	9147.0	15045.4	263875.7	3%	8335.3
48	101	250041.98	236565	13477.0	9147.0	15045.4	265087.4	-12%	28522.4
49	102	250688.14	272910	22221.9	9147.0	15045.4	265733.6	3%	7176.4

Tabla 5: Consolidado de cálculos para determinar la declaración de cajas al CD en base a modelo ARIMA y sus respectivas medidas de error. (Elaboración propia)

semana año	semana modelo	pronóstico HOLT WINT	demanda real	Error Abs	desv estandar	inv seguridad	declaración cajas CD	MAPE	Error abs
45	98	233204.42	241585	8380.6	17015.6	27988.2	261192.6	-8%	19607.6
46	99	220383.86	234306	13922.1	17015.6	27988.2	248372.1	-6%	14066.1
47	100	227151.26	272211	45059.7	17015.6	27988.2	255139.5	6%	17071.5
48	101	239609.39	236565	3044.4	17015.6	27988.2	267597.6	-13%	31032.6
49	102	266130.67	272910	6779.3	17015.6	27988.2	294118.9	-8%	21208.9

Tabla 6: Consolidado de cálculos para determinar la declaración de cajas al CD en base a modelo HOLT WINTERS y sus respectivas medidas de error. (Elaboración propia)

Por otro lado, se puede ver que el modelo Holt-Winters presentó en promedio un MAPE de 6% y 20597 cajas de error absoluto, si bien no se encuentra tan alejado de ARIMA el criterio de decisión para el modelo se basa en elegir al con menor error.

11.2 Resultados medida de desempeño primaria

Una vez entonces definido el modelo a utilizar (ARIMA), se empieza a declarar al CD desde la semana 45 la cantidad de cajas que se espera despachar a locales para las categorías estudiadas, se pueden ver los siguientes resultados en la productividad para las semanas entre la 45 a la 49:

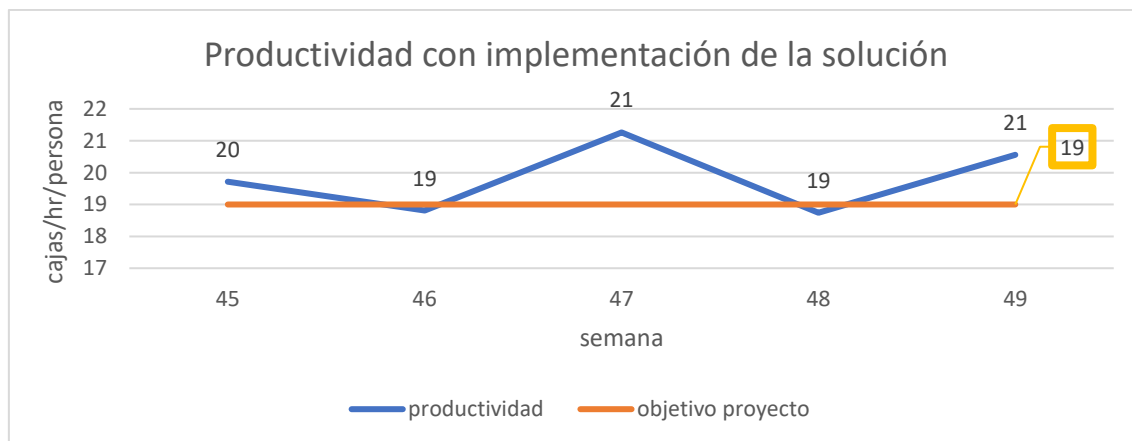


Gráfico 9: Productividad obtenida en el proceso de surtido de cajas en el CD 6011 de Walmart Chile luego de la implementación de la solución. (Elaboración propia)

Se puede ver que la productividad alcanza el objetivo propuesto para este proyecto, con un promedio de 20 cajas/hora por surtidor, donde incluso en 2 semanas alcanzó el objetivo de la empresa, si ahora vemos el gráfico que representa todas las semanas desde la primera estudiada hasta en las que se implementó el proyecto, vemos como la curva sube al objetivo en las últimas 5 semanas:

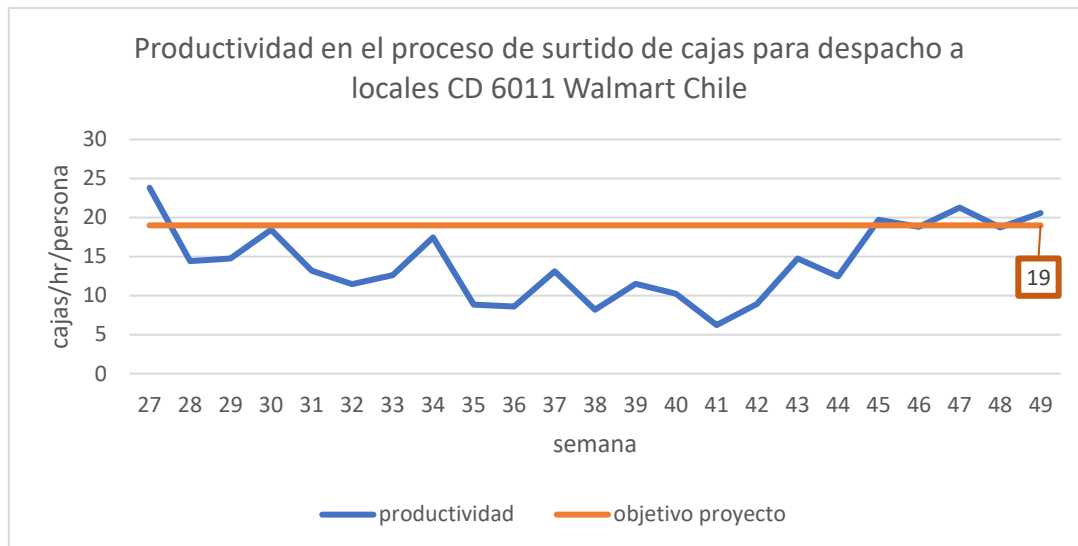


Gráfico 10: Productividad obtenida en el proceso de surtido de cajas en el CD 6011 de Walmart Chile durante la semana 27 a 49 del año 2023. (Elaboración propia)

Con el fin de comprobar que los resultados resultan significantes y demuestran un cambio en el proceso, se lleva a cabo la prueba estadística “Two sample t test”, donde se rechaza la hipótesis nula, ya que el valor p es 0.0224, es decir menor que 0.05, esto quiere decir que los resultados muestran una significancia de cambio con respecto al estado inicial.¹⁴

11.3 Resultados medidas de desempeño secundarias

Para la métrica secundaria de desempeño que mide la diferencia porcentual entre capacidad y necesidad, se puede ver que esta es lo mismo (en fórmula) que el MAPE, y anteriormente pudimos ver que para el modelo ARIMA el MAPE fue en promedio del 5%, lo que cumple con el objetivo de disminuir esta diferencia porcentual de 39% a 15%.

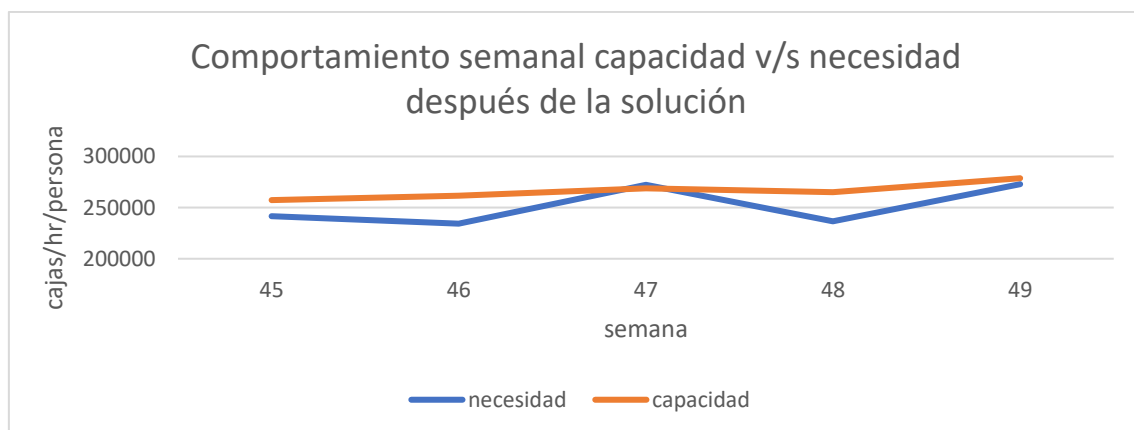


Gráfico 11: Comportamiento de la capacidad v/s necesidad en el proceso de surtido de cajas en el CD 6011 de Walmart Chile durante la semana 45 a 49 del año 2023. (Elaboración propia)

¹⁴ Prueba estadística en Anexo H.

Por último, para la métrica secundaria que mide la cantidad de personas sobre contratadas, es decir que no era necesario contratar en vista de la necesidad real, disminuyó en un 83% ya que bajó de 24 promedio a 4 personas promedio, sobre pasando el objetivo en un 33%.

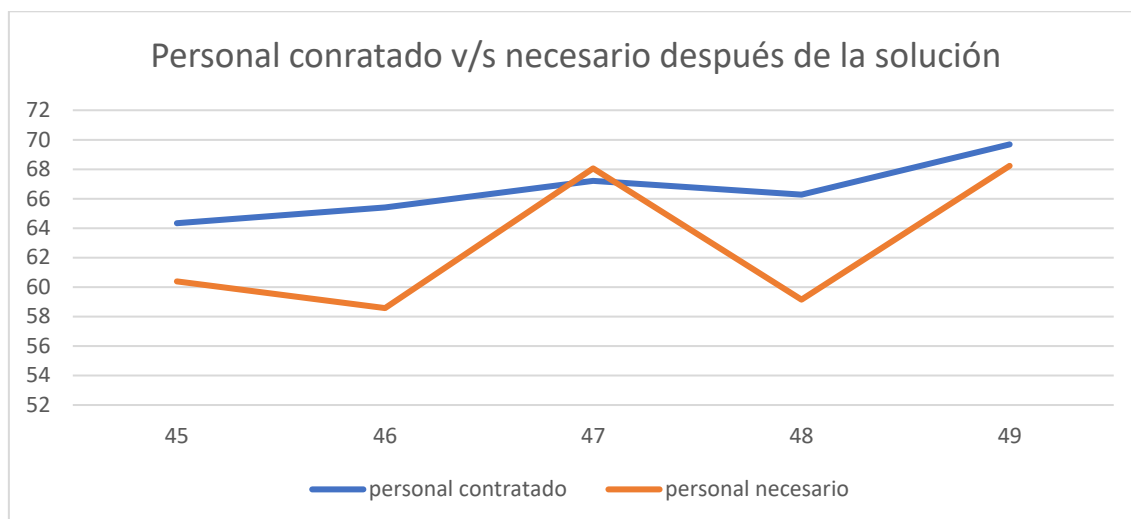


Gráfico 12: Personal contratado v/s necesario en el proceso de surtido de cajas en el CD 6011 de Walmart Chile durante la semana 45 a 49 del año 2023. (Elaboración propia)

12. Conclusiones y discusión

Como conclusiones de este proyecto se puede apreciar que el uso de métodos objetivos de pronósticos de la demanda para la planificación de la operación en la cadena de suministros representa mayor precisión en los resultados que los métodos subjetivos, involucrando así a la productividad y eficiencia de áreas como el centro de distribución, donde estandarizando un input de esta se puede aumentar toda la productividad. Específicamente vimos como la productividad aumentó de 13 a 20 cajas/hora por surtidor, superando el objetivo del proyecto y casi alcanzando el objetivo de la empresa.

Al aumentar la productividad vimos cómo se proyectan ahorros que anteriormente representaban un costo de oportunidad, disminuyendo los costos en un 81%.

Bien cabe destacar en este proyecto que no se consideran algunas variables que pueden ser muy relevantes para la declaración de volumen de cajas por despachar del centro de distribución a locales, tales como promociones activas, ventas, feriados, entre otros. Se propone agregar estas variables exógenas a la regresión de las series de tiempo, de hecho en el estado del arte se estudia el método ARIMAX que propone la realización de esto, existiendo

softwares de no mayor completitud para usarlo, por efectos de tiempo y datos en este proyecto no se realizó, pero la invitación y la propuesta es a mejorar las regresiones agregándole la opción de la variable exógena que pueda relacionarse temporalmente con la serie de tiempo entregando valor en el pronóstico, sobre todo de un gran retailer como Walmart Chile donde los factores a medir para tomar una decisión son abundantes.

También se puede concluir que en problemas que surgen en la operación muchas veces hay decisiones por detrás provenientes de otras áreas que les afectan directamente, las otras áreas al no verse involucradas en su mayoría en el día a día de la operación pueden afectar indicadores de productividad de estas, incluso llega a ser difícil encontrar las causas de estos problemas cuando en las empresas tan grandes pareciera encontrarse todo más fragmentado que conectado, se entiende luego de este proyecto que ese es uno de los principales desafíos en una compañía de la envergadura de Walmart Chile.

Para la empresa este proyecto entrega un valor sustancial en cuanto a estandarización de los inputs para la toma de decisiones de operación, si bien, al presentar y generar pronósticos y stock de seguridad se invita a la gerencia que toma la decisión de la declaración de cajas a guiarse, estos pueden decidir no hacerlo y tener en cuenta una opinión subjetiva, pero el input de pronósticos y declaración estandarizada de cajas se encontrará libre para todos los usuarios ahorrándoles tiempo e incertidumbre donde además como se pudo apreciar en los resultados, entrega pronósticos que entregan mayor precisión para el cálculo de capacidad de personal, ahorrando y aumentando la productividad del proceso de surtido de cajas en el CD 6011 Quilicura de Walmart Chile.

13. Referencias

Academy, G. (2016). *Análisis de los 5 ¿Por qué?*

Amazon. (2023).

(s.f.). *Comparing Prophet and Deep Learning to ARIMA in forecasting Wholesale Food prices* .

Culture, S. (2023). Obtenido de safetyculture.com.

Harvard. (2004). *Harvard Business Review*.

Logistec. (2022).

Nahmias. (2007). *Análisis de la producción y las operaciones*.

Unilever. (2023).

Walmart. (2022).

14. Anexos

A. Instock:

Principal indicador de reabastecimiento y en base al cual se toman la mayor parte de las decisiones, por eso resulta siempre sensato analizarlo y ver si tiene relación con distintos fenómenos.

$$Instock = \frac{\sum \text{Combinaciones ítem local sin quiebres}}{\text{Total de combinaciones ítem local}} * 100$$

B. Test de normalidad Shapiro Wilk:

Programado en Python (Collab)

```
import pandas as pd

excel_file = '/content/prod-norm-5.xlsx'
df = pd.read_excel(excel_file)

data = df['data'].dropna()

stat, p_value = shapiro(data)

print(f'Estadístico de prueba: {stat:.4f}')
print(f'Valor p: {p_value:.4f}')
```

Estadístico de prueba: 0.9576

Valor p: 0.3691

No podemos rechazar la hipótesis nula (la muestra parece provenir de una distribución normal)

C. Detalle evaluación económica:

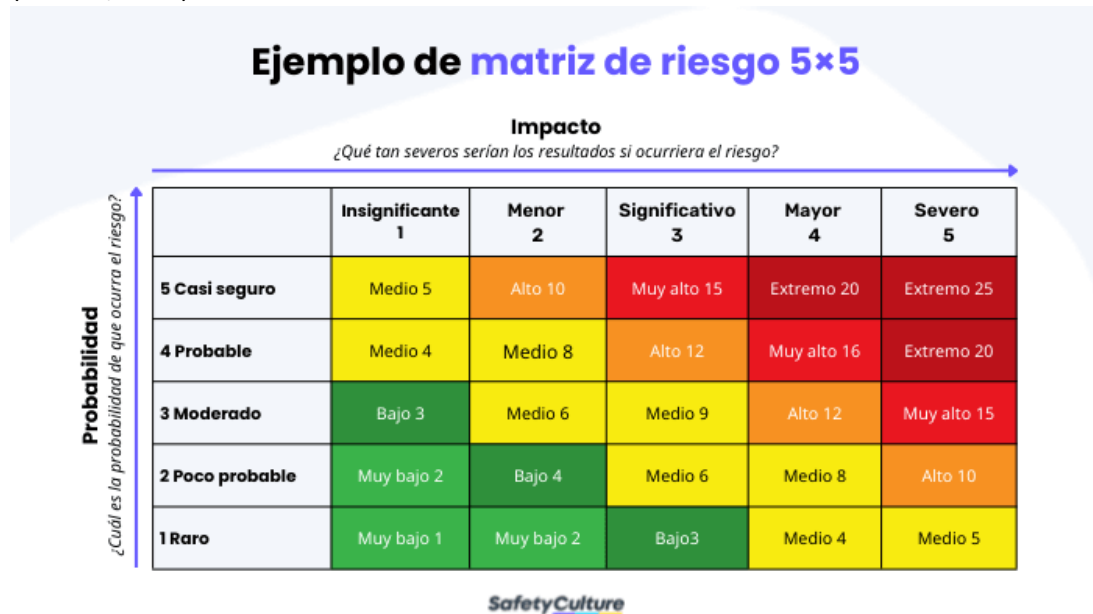
Capacidad y proyección se cuantifican en cajas.

NUEVA CAPACIDAD	PROYECCIÓN NECESIDAD	DESV OBJETIVO	PERSONAL CONTRATADO	PERSONAL NECESARIO	NUEVO COSTO
189450	164739	15%	47	41	\$ (1,081,099.69)
197211	171488	15%	49	43	\$ (1,125,390.00)
176961	153879	15%	44	38	\$ (1,009,830.94)
153171	133192	15%	38	33	\$ (874,072.50)
159944	139082	15%	40	35	\$ (912,725.63)
144295	125474	15%	36	31	\$ (823,423.13)
160836	139857	15%	40	35	\$ (917,811.56)
222878	193807	15%	56	48	\$ (1,271,858.44)
112746	98040	15%	28	25	\$ (643,387.50)
116656	101440	15%	29	25	\$ (665,700.00)
75315	65491	15%	19	16	\$ (429,784.69)
109654	95351	15%	27	24	\$ (625,740.94)
172794	150256	15%	43	38	\$ (986,055.00)
173144	150560	15%	43	38	\$ (988,050.00)
96042	83515	15%	24	21	\$ (548,067.19)
106076	92240	15%	27	23	\$ (605,325.00)
188124	163586	15%	47	41	\$ (1,073,533.13)
154251	134131	15%	39	34	\$ (880,234.69)

D. Matriz de riesgo

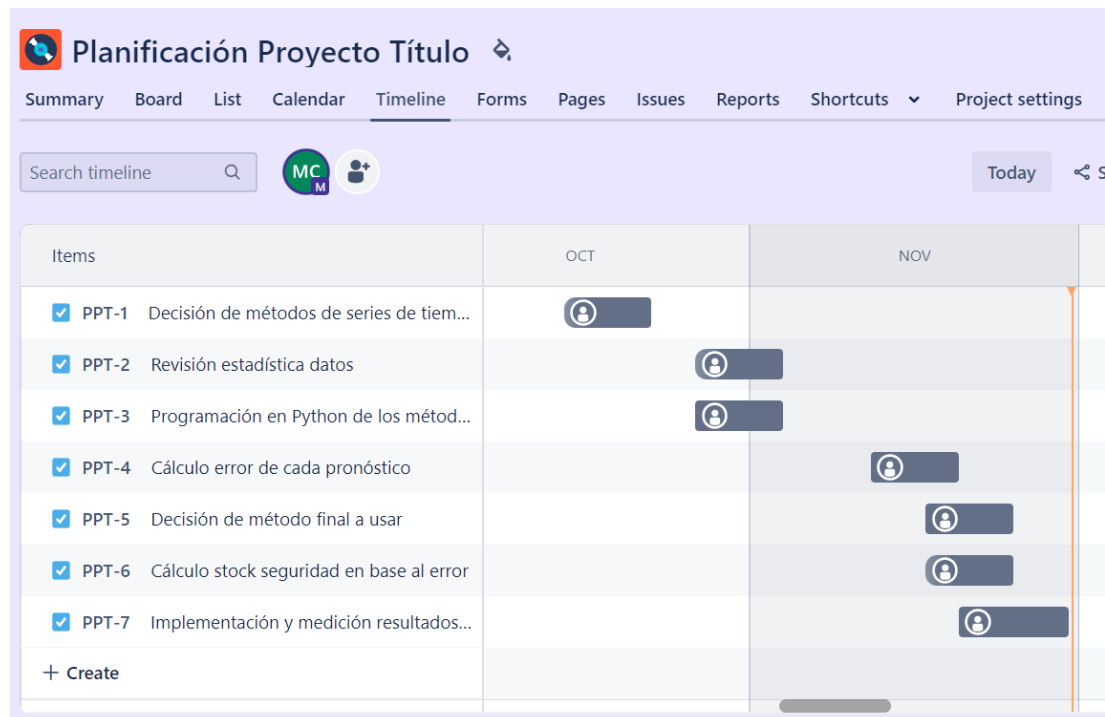
Matriz que se utilizó de guía para realizar la del presente proyecto.

(Culture, 2023)



E. Plataforma Jira

Visión timeline de la plataforma Jira para planificación personal con fechas y objetivos.



F. Códigos Python modelos de pronósticos

Algoritmo Python para el método ARIMA en base a statsmodels

```
!pip install pandas numpy matplotlib statsmodels xlrd

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

archivo_excel = '/content/semana_rolling_sstk.xlsx'
datos = pd.read_excel(archivo_excel)

print(datos.head())

serie_temporal = pd.Series(datos['VALOR'].values,
index=pd.to_datetime(datos['SEMANA']))

# Serie temporal
plt.plot(serie_temporal)
plt.title('Serie Temporal')
plt.show()

# Modelo ARIMA
p, d, q = 1, 1, 1 # Defecto
modelo_arima = ARIMA(serie_temporal, order=(p, d, q))
resultado_arima = modelo_arima.fit()

# Pronósticos
numero_de_semanas_a_predecir = 6
pronostico =
resultado_arima.get_forecast(steps=numero_de_semanas_a_predecir)
pronostico_mean = pronostico.predicted_mean
pronostico_ci = pronostico.conf_int()

# Resultados
plt.plot(serie_temporal, label='Histórico')
plt.plot(pronostico_mean.index, pronostico_mean.values,
color='red', label='Pronóstico')
plt.fill_between(pronostico_ci.index, pronostico_ci.iloc[:, 0],
pronostico_ci.iloc[:, 1], color='pink', alpha=0.2)
plt.title('Pronóstico ARIMA')
plt.legend()
plt.show()

# Imprimir
print("Pronósticos:")
print(pronostico_mean)
```

Algoritmo Python para el método Holt Winters en base a statsmodels

```
import pandas as pd

import matplotlib.pyplot as plt
import statsmodels.api as sm

file_path = '/content/semana_rolling_sstk.xlsx'
data = pd.read_excel(file_path, parse_dates=['SEMANA'])

plt.figure(figsize=(10, 6))
plt.plot(data['SEMANA'], data['VALOR'], label='Datos originales')
plt.title('Datos Originales')
plt.legend()
plt.show()

# Holt-Winters
model = sm.tsa.ExponentialSmoothing(data['VALOR'], trend='add',
seasonal='add', seasonal_periods=12)
model_fit = model.fit()

# Realizar pronósticos
forecast_steps = 6
forecast = model_fit.predict(start=len(data), end=len(data) +
forecast_steps - 1)

# Resultados
plt.figure(figsize=(12, 8))
plt.plot(data['SEMANA'], data['VALOR'], label='Datos originales')
plt.plot(pd.date_range(start=data['SEMANA'].max() +
pd.Timedelta(days=1), periods=forecast_steps, freq='W'), forecast,
label='Pronóstico', linestyle='--')
plt.title('Suavizamiento Exponencial Triple (Holt-Winters) con
Pronóstico')
plt.legend()
plt.show()

print("Pronósticos para los próximos pasos:")
print(forecast)
```

G. Dataiku

Plataforma de Data Science unificada, funcionalidades para preparación de datos, análisis y creación de reportes automáticos (uso del presente). (Dataiku)

H. Test estadístico Two Sample T test

Programado en Python Collab:

```
from scipy.stats import shapiro
import pandas as pd

excel_file = '/content/prod nueva tettt.xlsx'
df = pd.read_excel(excel_file)

column_sample1 = 'prod nueva'
column_sample2 = 'prod antes'

sample1 = df[column_sample1]
sample2 = df[column_sample2]

from scipy.stats import ttest_ind
import numpy as np

datos_antes_limpio = sample1[~np.isnan(sample1)]
datos_despues_limpio = sample2[~np.isnan(sample2)]

resultado = ttest_ind(datos_antes_limpio, datos_despues_limpio)

print(f"Valor p: {resultado.pvalue}")

Estadística de prueba (t): 4.360959664763869
Valor p: 0.0024095393929412377
```