



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**DESARROLLO DE UN MODELO DE ESTIMACIÓN DE DEMANDA
DE MEDICAMENTOS PARA UNA EMPRESA FARMACÉUTICA**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL
INDUSTRIAL**

DOMINIQUE GABRIELA LECLERC CORREA

**PROFESOR GUÍA:
ALEJANDRA PUENTE CHANDÍA**

**MIEMBROS DE LA COMISIÓN:
PABLO MARÍN VICUÑA
BLAS DUARTE ALLEUY**

**SANTIAGO DE CHILE
2020**

**RESUMEN DE LA MEMORIA PARA
OPTAR AL TÍTULO DE:** Ingeniera Civil
Industrial

POR: Dominique Gabriela Leclerc Correa

FECHA: 06/04/2020

PROFESOR GUÍA: Alejandra Puente Chandía

**DESARROLLO DE UN MODELO DE ESTIMACIÓN DE DEMANDA
DE MEDICAMENTOS PARA UNA EMPRESA FARMACÉUTICA**

Uno de los principales desafíos dentro de la logística de una empresa farmacéutica, y en general de cualquier empresa que lidia con la comercialización de productos de cualquier tipo, es la estimación de demanda. Hacer un buen pronóstico de lo que se venderá en períodos futuros afecta directamente en los ingresos de la empresa, ya que: si se sobreestima la demanda, se tiene sobre stock de productos que se deberá vender a precios más bajos; y si se subestima, se tienen ventas perdidas. En una empresa farmacéutica estos riesgos son aún mayores porque se trata de medicamentos con fechas de expiración estrictas y con la salud de las personas. La empresa farmacéutica en estudio tiene un nivel de cumplimiento de pronóstico del 56% en promedio, es decir, el 56% de sus productos cumple con lo predicho, con un rango de error del 20%.

El objetivo de este trabajo es desarrollar un modelo de estimación de demanda mensual de los medicamentos de la empresa en estudio, a nivel de SKU y para los diferentes canales de venta. Esto para poder robustecer el proceso de estimación de demanda, para dar satisfacción a sus clientes (instituciones que venden medicamentos e instituciones de salud) y controlar grandes variaciones de stock. Para este objetivo se emplea la metodología de CRISP-DM, donde se aplican modelos estadísticos para hacer las estimaciones de demanda para cada una de las series de tiempo de producto-canal de venta. Se aplican modelos que van desde lo más simple (Promedio móvil simple) a lo más complejo (modelos ARIMA).

Se selecciona un medicamento por cada principio activo para aplicar los modelos, considerándolos como medicamentos representativos de todo el catálogo de la empresa. Posteriormente se estima y se selecciona el “mejor modelo” para cada una de las series SKU-canal, considerando el que tenga menor RMSE y menor desvío entre la demanda real y pronosticada. Como principal resultado se obtiene un 78% de cumplimiento de pronóstico en los medicamentos con los “mejores modelos” seleccionados, frente a un promedio histórico de cumplimiento de pronóstico del 62%. Se propone para trabajos futuros capturar la demanda desde los pedidos que se hacen a la empresa para tener pronósticos más certeros.

TABLA DE CONTENIDO

I. ANTECEDENTES GENERALES	1
II. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN	4
III. OBJETIVOS	8
a. Objetivo General	8
b. Objetivos Específicos	8
IV. ALCANCES Y RESULTADOS ESPERADOS	9
V. MARCO CONCEPTUAL	10
a. Cross Industry Standard Process for Data Mining (CRISP-DM)	10
b. Series de tiempo	11
c. Modelos de estimación de demanda	12
b.1 Promedio móvil simple	12
b.2 Suavización exponencial simple	14
b.3 Método Holt-Winters	15
b.4 Modelos ARIMA (Autoregressive Integrated Moving Average)	16
d. Métricas de error	17
VI. METODOLOGÍA	18
VII. DESARROLLO DE LA METODOLOGÍA Y ANÁLISIS DE RESULTADOS	20
a. Estudio de la situación actual	20
b. Procesamiento y limpieza de datos	26
b.1 Datos disponibles	26
b.2 Limpieza de datos	28
b.2 Aproximación a una caracterización de la demanda	29
c. Selección de medicamentos a estudiar y nivel de agregación de la estimación	36
c.1 Demanda de medicamentos por principio activo y canal de venta	36

c.2 Características de series de tiempo a estimar	40
d. Elección de modelos a utilizar y su estimación	43
d.1 Estimación de demanda del medicamento OL116 – canal pseudo-privado.....	44
d.2 Resumen de modelos para todos los productos por canal de venta	52
e. Evaluación de modelos.....	53
f. Propuesta de desarrollo y aplicación.....	59
VIII. CONCLUSIONES Y RECOMENDACIONES PARA FUTUROS	
TRABAJOS.....	61
a. Conclusiones.....	61
b. Limitaciones y trabajo futuros	63
IX. BIBLIOGRAFÍA.....	65
X. ANEXOS	67
Anexo 1: Evolución histórica del indicador <i>forecast accuracy</i>	67
Anexo 2: Porcentaje de cumplimiento por principio activo.....	68
Anexo 3: Demanda mensual por modalidad de venta y línea de negocio	70
Anexo 4: Comparación de demandas por unidad de negocio.....	72
Anexo 5: Cantidad de medicamentos por principio activo (familia).....	77
Anexo 6: Modelos estimados por serie de tiempo producto-canal.....	79
Anexo 7: Mejor método por serie de tiempo producto-canal.....	86
Anexo 8: Forecast accuracy de mejor modelo comparado con promedio de forecast accuracy del 2019	90

INDICE DE TABLAS

Tabla 1: Descripción de unidades de negocio y sus líneas asociadas.....	21
Tabla 2: Forecast accuracy promedio por línea y unidad de negocios en años 2018 y 2019.....	24
Tabla 3: Cumplimiento de forecast por línea y unidad de negocio	25
Tabla 4: Detalle de variables de archivo de ventas históricas.....	27
Tabla 5: Detalle de variables de archivo de backorders.....	27
Tabla 6: Métricas y coeficientes de estimación de promedio móvil simple de la serie OL116 - canal pseudo-privado	47
Tabla 7: Forecast accuracy del producto OL116.....	47
Tabla 8: Métricas y coeficientes de estimación de suavización exponencial simple de la serie OL116 - canal pseudo-privado	49
Tabla 9: Métricas y coeficientes de promedio móvil doble de la serie OL116 - canal pseudo-privado.....	50
Tabla 10: Métricas y coeficientes de suavización exponencial doble de la serie OL116 - canal pseudo-privado	51
Tabla 11: Métricas y coeficientes de ARIMA(1,0,1) de la serie OL116 - canal pseudo-privado	52
Tabla 12: Modelos estimados para una selección de SKU-canal.....	53
Tabla 13: Comparación de forecast accuracy con mejor modelo y su promedio del 2019	56
Tabla 14: Comparación de cumplimiento de forecast real con lo estimado	58

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Sector farmacéutico en Chile (Vasallo, 2010)	1
Ilustración 2: Porcentaje de cumplimiento 2018 y 2019	6
Ilustración 3: Forecast accuracy desde el año 2018 a junio de 2019 por línea de negocio	23
Ilustración 4: Cumplimiento de forecast desde el año 2018 al 2019 por línea de negocio	25
Ilustración 5: Demanda mensual por línea de negocio.....	30
Ilustración 6: Demanda en unidades por línea de negocio, versión 2.....	30
Ilustración 7: Utilidades reportadas semanalmente por línea de negocio.....	31
Ilustración 8: Demanda semanal por unidad de la línea de negocio FARMA, unidades Cardio metabólica y Neurociencias	32
Ilustración 9: Demanda semanal por unidad de la línea de negocio FARMA, unidades Primary Care, Dolor y UGI	33
Ilustración 10: Demanda semanal por unidad de la línea de negocio OLE.....	33
Ilustración 11: Demanda mensual por modalidad de venta	34
Ilustración 12: Utilidades reportadas semanalmente por modalidad de venta.....	35
Ilustración 13: Demanda semanal para 4 productos de la línea FARMA, unidad DOLOR.....	37
Ilustración 14: Demanda semanal de medicamentos de principio activo Drospirenona	39
Ilustración 15: Demanda por canal de venta del medicamento FA47	40
Ilustración 16: Detalle de estacionalidad de las series de tiempo a estudiar	42
Ilustración 17: Series de tiempo con tendencia al alza	43
Ilustración 18: Diagrama de grupos de entrenamiento y validación.....	44
Ilustración 19: Demanda de medicamento OL116 por canal pseudo-privado	45
Ilustración 20: Gráfico de autocorrelación para la serie OL116 - canal pseudo privado	45
Ilustración 21: Estimación de promedio móvil simple de la serie OL116 - canal pseudo-privado	46
Ilustración 22: Residuales de la estimación por promedio móvil simple de medicamento OL116 - pseudo privado.....	48
Ilustración 23: Estimación de suavización exponencial simple de la serie OL116 - canal pseudo-privado.....	48
Ilustración 24: Estimación de promedio móvil doble de la serie OL116 - canal pseudo-privado	49

Ilustración 25: Estimación de suavización exponencial doble de la serie OL116 - canal pseudo-privado.....	50
Ilustración 26: Estimación de modelo ARIMA(1,0,2) de la serie OL116 - canal pseudo-privado	51
Ilustración 27: Mejores modelos agrupados por unidad de negocio	55
Ilustración 28: Mejores modelos agrupados por canal de venta	55

I. ANTECEDENTES GENERALES

La industria farmacéutica en el mundo es un sector empresarial que se dedica a la investigación, desarrollo, fabricación y comercialización de productos químicos medicinales que alivian dolencias y buscan la prevención y tratamiento de enfermedades (Bajaña & Moncayo, 2012). En Chile esta industria representa cercano al 1,2% del PIB en importaciones, moviendo alrededor de 270.000.000 unidades de productos al año 2017 (Oficina Económica y Comercial de España en Santiago de Chile, 2018).

El informe “*El mercado de medicamentos en Chile: caracterización y recomendaciones para la regulación económica*” (Vasallo, 2010) emitido a petición del Ministerio de Salud de Chile, detalla la morfología del mercado de medicamentos. Los principales actores que están involucrados en la industria son: laboratorios farmacéuticos, droguerías, cadenas farmacéuticas y farmacias. Una ilustración del sector farmacéutico del país se puede ver en la siguiente figura.

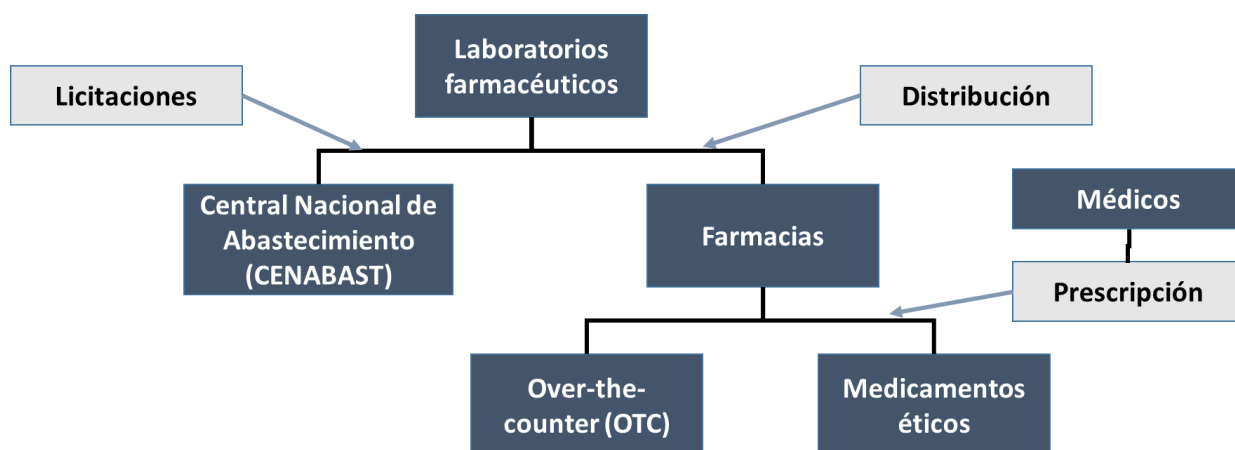


Ilustración 1: Sector farmacéutico en Chile (Vasallo, 2010)

Partiendo por los laboratorios, en Chile existen alrededor de 180 laboratorios farmacéuticos que importan, producen y distribuyen medicamentos. En el Libro VI del Código Sanitario se explicita que “los laboratorios farmacéuticos autorizados por el

Instituto de Salud Pública (ISP) son los únicos actores a cargo de la producción de medicamentos” y que “los laboratorios farmacéuticos sólo pueden distribuir medicamentos a droguerías autorizadas por el ISP” (Ministerio de Salud, 2006). Es importante destacar que el Instituto de Salud Pública puede otorgar registro sanitario a cualquier empresa que pueda demostrar que cuenta con infraestructura que permita fabricar, almacenar o distribuir medicamentos. Esto crea una dinámica especial en la industria: muchos laboratorios farmacéuticos tienen una tercera empresa, con registro sanitario, encargada de la distribución de sus medicamentos. Así los laboratorios pueden ocuparse sólo de la importación, producción y/o comercialización de medicamentos.

Los siguientes actores de la industria son las droguerías. Una droguería es una institución que centraliza las compras de medicamentos. Según el informe mencionado anteriormente “son el instrumento logístico de compra, que permite negociar y establecer las condiciones con los laboratorios proveedores” (Vasallo, 2010). Además el Decreto 466 del Código Sanitario de Chile establece que las droguerías son establecimientos que “funcionarán bajo la dirección técnica de un profesional químico farmacéutico” (Ministerio de Salud, 2013).

Siguen las farmacias, cadenas de farmacias y otras instituciones de salud que requieren de medicamentos. Cada uno de estos establecimientos, que son los que venden medicamentos a consumidores finales (persona individual), deben estar calificados como droguería. Estas droguerías pueden ser el intermediario de una o muchas farmacias. Por ejemplo, las cadenas de farmacias tienen una droguería propia que abastece a todos sus establecimientos y también hay droguerías que luego venden medicamentos a otras droguerías, es decir, a otras farmacias de diferentes nombres. Un laboratorio farmacéutico le venderá medicamentos a la droguería, no necesariamente a cada farmacia individualmente.

Otro actor presente en Chile es la Central Nacional de Abastecimiento. Esta es una institución pública que intercede en todas las licitaciones de productos para instituciones de salud públicas (Ministerio de Salud, 2006).

Es desde las farmacias y otras instituciones de salud que los productos farmacéuticos llegan a manos del consumidor final, el paciente.

La industria farmacéutica está compuesta por un conjunto de sub-mercados que varían en sus características. En primer lugar, se encuentra el mercado de los medicamentos de venta libre (OTC: *Over-the-counter*, es decir, sobre el mostrador). Son aquellos medicamentos comprados directamente por el paciente, sin necesidad de poseer una receta médica (Vasallo, 2010). Este sub-mercado se puede considerar similar a un mercado competitivo ya que consumidores tienen total control en la decisión de compra y pueden utilizar información que reciben a través de publicidad.

Por otro lado, está el mercado de los medicamentos éticos o de venta bajo receta. Aquí existen principalmente dos tipos: el producto farmacéutico innovador y el ético similar. El innovador es aquel que innova el mercado farmacéutico mundial y el que podrá ser vendido únicamente por el laboratorio que impulsa su desarrollo, por un período determinado de tiempo regido por una patente. Mientras que el medicamento ético similar es aquel que tiene igual dosis de principio(s) activo(s) e igual forma farmacéutica que el producto farmacéutico innovador (Farma ERP, 2010). Este último se comienza a comercializar cuando la patente del medicamento innovador ha terminado. En este mercado participan médicos encargados de la prescripción de medicamentos.

El trabajo de título se enmarca dentro de un laboratorio farmacéutico que participa en el mercado de medicamentos éticos similares sujetos a receta médica. El laboratorio en estudio importa medicamentos terminados, es decir, medicamentos que no se producen en Chile. Estos bienes son en su totalidad medicamentos que ya han estado en el mercado mundial por tiempo considerable, es decir, que la patente del medicamento innovador ya ha terminado hace más de 1 año.

II. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN

Uno de los principales desafíos dentro de la cadena de suministro de una empresa farmacéutica, y en general de cualquier empresa que lidia con inventario, es la estimación de demanda.

La organización donde se realiza el proyecto de título es un laboratorio farmacéutico radicado en Chile. Como se menciona en el apartado anterior, un laboratorio farmacéutico es una institución que importa, produce y distribuye medicamentos. En el caso del laboratorio en estudio, este sólo importa bienes terminados que distribuye a instituciones certificadas. Los productos que importa y distribuye son medicamentos éticos similares. Es decir, medicamentos no innovadores vendidos bajo receta médica.

Los productos comercializados se agrupan en dos familias: FARMA, donde se concentran productos de mayor volumen de ventas y OLE (Oncología, Licencia y Especialidades), que une los medicamentos para enfermedades o condiciones específicas con precios por unidad más elevados y con menor volumen de ventas. Dentro de estas dos unidades de negocio existen más clasificaciones dependiendo de las enfermedades o condiciones que tratan los medicamentos.

El laboratorio farmacéutico tiene 3 canales de venta: institucional, pseudo-privado y privado. A continuación se detalla cada uno:

- El canal institucional considera la venta de medicamentos a instituciones de salud pública como hospitales, Fuerzas Armadas, consultorios, entre otros. Se da la modalidad de venta a través de licitaciones, a través de la Central de Abastecimiento del Ministerio de Salud (CENABAST).
- El canal privado se enfoca en las instituciones de salud privadas como lo son clínicas y consultorios privados.
- El canal pseudo-privado a su vez concentra las cadenas de farmacias y pequeñas farmacias. En este canal la empresa farmacéutica se relaciona con las droguerías de cada establecimiento farmacéutico.

El inventario de los productos se hace a través de un tercero, llamado “operador logístico”. Esta empresa externa es la que se encarga del almacenamiento y transporte de los medicamentos (acorde a lo que la farmacéutica en estudio le indica). Este inventario es centralizado en la Región Metropolitana.

En este trabajo no se indaga en cómo llegan los productos a cada uno de sus clientes ni cómo se almacenan antes de entregarse. Es otra área de la farmacéutica la que asegura que los productos se encuentren en las bodegas del operador logístico. El presente trabajo se concentra en la necesidad de determinar los productos que deben estar disponibles en el almacenamiento del operador logístico para satisfacer la demanda mensual del laboratorio.

Cada mes, la empresa farmacéutica tiene la tarea de estimar la demanda (hacer un *forecast*) de sus productos. Se debe responder a la pregunta de cuánto se vende mensualmente cada producto (SKU). Las estimaciones se hacen para los 6 meses próximos y se revisan mes a mes, con la posibilidad de modificar los valores de los meses siguientes. La revisión de los pronósticos se hace en la tercera semana de cada mes, es decir, cuando ya queda sólo 1 semana para que el mes pronosticado comience.

El método de estimación de demanda utilizado actualmente es un juicio de expertos. Actores que diariamente están en contacto con distintas partes del negocio se reúnen una vez al mes a definir el pronóstico mensual tomando en cuenta: ventas pasadas, entradas o salidas de producto de la competencia, comportamiento de los clientes, expectativas de la empresa, expectativa de quienes venden, intuiciones personales, entre otras variables. Los expertos son principalmente los representantes de grupos de medicamentos que son los encargados de venderlos y un analista encargado de supervisar estas reuniones y dar información más concreta sobre la logística del negocio.

El año 2019 se han implementado cambios en el proceso de estimación de demanda, convocando a personas de diferentes áreas del negocio para dar sus opiniones e interpretaciones. Esta medida busca quitar sesgos en la estimación de demanda, al antes estar sólo formada por la fuerza de ventas de la empresa. Entre todos los expertos se determina el *forecast* de cada producto, pero quien tiene la decisión final es el representante de los medicamentos.

Tener un mal pronóstico de demanda es particularmente riesgoso y dañino para una empresa farmacéutica porque se trata de productos con fecha de vencimiento, con composiciones químicas y otras particularidades que inciden directamente en su venta. El sobreestimar la demanda significa tener un sobre stock de productos con fecha de expiración inamovible, donde clientes no aceptarán productos que estén muy cercanos a vencer. Para lograr vender estos productos se debe recurrir a precios con descuentos perdiendo rentabilidad, o simplemente se deberán eliminar, perdiendo potenciales ingresos. Por otro lado, subestimar la demanda significa tener ventas perdidas por quiebres de stock resultando en pérdida de clientes al irse a la competencia, o en el comprometer la salud de un paciente al no tener disponible el medicamento que necesita.

Actualmente, la empresa mide la certeza de sus pronósticos considerando un margen de error del $\pm 20\%$. Si se encuentran dentro del rango de error, se considera que el pronóstico fue acertado y se cumplió lo predicho. A continuación se muestra el porcentaje de cumplimiento a lo largo de los años 2018 y 2019.

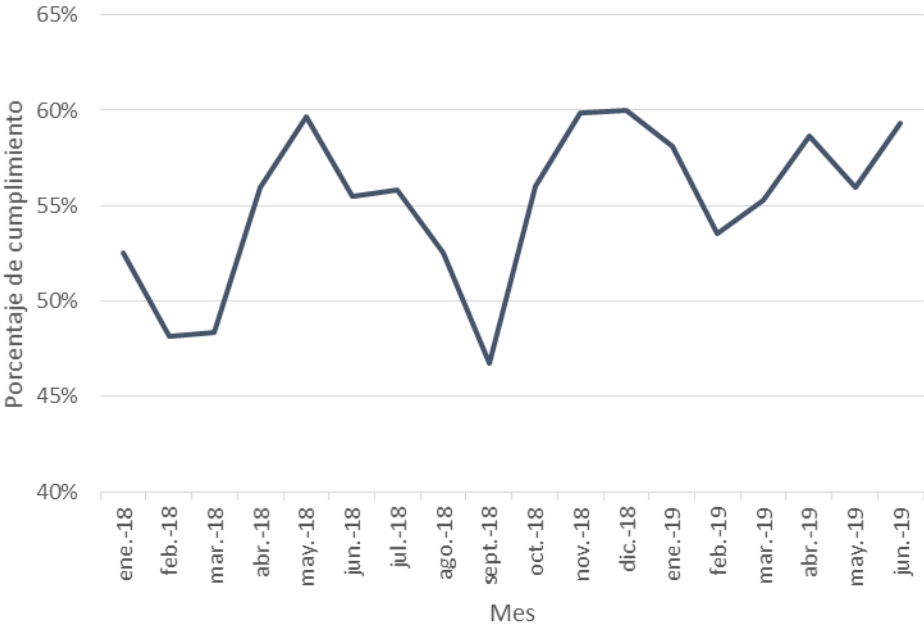


Ilustración 2: Porcentaje de cumplimiento 2018 y 2019

Ya se han hecho mejoras en el proceso de estimación de demanda. Desde el 2019 (cuando se implementaron los cambios) el porcentaje de cumplimiento no ha bajado la barrera

del 50%, como se puede ver en la Ilustración 2. En el mes de septiembre de 2018 hubo una baja considerable del cumplimiento de la empresa, que muestra la variabilidad de este indicador antes de que se implementaran los nuevos cambios.

Aun existiendo un aumento del cumplimiento desde el año 2019, se consideran porcentajes bajos. Es en este espacio que resulta necesario implementar nuevos *inputs* de información para lograr tener una predicción de demanda más certera. La empresa farmacéutica está intentando tener un pronóstico de demanda más certero, una de las metas es llegar al 70% a fines del año 2020.

El proyecto implica la implementación de herramientas y modelos estadísticos en el proceso de estimación de demanda dentro de la empresa farmacéutica, para reducir sesgos que pueda tener el juicio de expertos y que la información que otorguen las herramientas implementadas sea un input más en el proceso de estimación de demanda.

III. OBJETIVOS

a. Objetivo General

Desarrollar un modelo de estimación de demanda mensual para los medicamentos que distribuye una empresa farmacéutica, a nivel de SKU y para los diferentes canales de venta, de forma de robustecer el proceso de estimación de demanda y que sea un sistema de apoyo a la toma de decisiones, para así dar satisfacción a sus clientes y controlar variaciones de stock.

b. Objetivos Específicos

1. Determinar métricas de desempeño del proceso de estimación de demanda actual, de forma de definir una línea base actual
2. Caracterizar los medicamentos en base a sus características y demanda histórica, de forma de elegir un conjunto de productos representativos del catálogo total de medicamentos de la empresa
3. Estimar la demanda de los medicamentos representativos escogidos, probando diferentes modelos y evaluando su ajuste y capacidad de pronóstico
4. Generar una propuesta de desarrollo y aplicación de los modelos para hacerlos utilizables en el futuro, y replicables a los medicamentos no seleccionados

IV. ALCANCES Y RESULTADOS ESPERADOS

Los alcances del trabajo a realizar están dados por:

1. Se utilizan datos de ventas y *backorders* desde el año 2018 hasta agosto del año 2019 y sólo se consideran los productos que actualmente están activos.
2. Se dejan fuera del análisis productos de reciente lanzamiento y productos que se manejan con modelo *order-up-to*.
3. Se dejan fuera del estudio las ventas adjudicadas en licitaciones en el Mercado Público, a través de CENABAST.
4. No se indaga en cómo se desarrolla el almacenamiento ni traslado de los medicamentos a los clientes ya que esto es manejado por un tercero.
5. Los modelos resultantes son un apoyo al proceso actual de estimación de demanda, es decir, sirve de complemento al método de juicio de expertos usado.

Mientras que se espera tener como resultado final lo siguiente:

1. Una caracterización y clasificación de los medicamentos de la empresa farmacéutica.
2. Modelos que permitan estimar la demanda de productos característicos que tome como principal insumo las ventas históricas.
3. Una propuesta de desarrollo y aplicación de los modelos escogidos para hacerlos utilizables en el futuro.

V. MARCO CONCEPTUAL

A continuación se exponen los conceptos que se necesitan para el desarrollo del trabajo de título. Se detallan desde la metodología de minería de datos, métodos de estimación de demanda y métricas para la evaluación de estos.

a. Cross Industry Standard Process for Data Mining (CRISP-DM)

Para abordar la minería de datos se define utilizar la metodología CRISP-DM (Gallardo, 2009), la técnica más utilizada en proyectos de esta índole. Esta herramienta tiene 6 pasos:

1. **Comprensión del negocio o problema:** Se busca entender los objetivos y requerimientos del proyecto desde la perspectiva de la empresa.
2. **Comprensión de los datos:** Se refiere a la etapa de recopilación de datos, identificando la calidad de ellos.
3. **Preparación de los datos:** Se hace una limpieza de datos ocupando técnicas de visualización y de relación entre variables para adaptar los datos a las diversas técnicas que se podrían ocupar.
4. **Modelado:** Se seleccionan las técnicas de modelado más pertinentes tomando en cuenta: si responden a los objetivos y requerimientos definidos, si están los datos disponibles y si el tiempo es adecuado para resolver el problema. También, aquí se determinan métricas de evaluación de los modelos elegidos y posteriormente, los modelos son estimados.
5. **Evaluación:** Se evalúa cada modelo con respecto a los objetivos del proyecto y las métricas definidas. En esta etapa es necesario revisar con detención cada modelo para identificar posibles errores que requieran repetir la fase 4.

6. Implementación: Con el conocimiento entregado por la etapa anterior, se implementa la solución en el negocio ya sea recomendando acciones o aplicando el modelo a otros datos dentro del proceso.

b. Series de tiempo

En el trabajo de memoria se trabaja con datos de ventas para diferentes medicamentos. Estos datos están agrupados en lo que se llama “series de tiempo” que se puede definir como

Para el análisis de series de tiempo es útil descomponerlas en elementos básicos como: tendencia, estacionalidad y errores (Schroeder, Meyer Goldstein, & Rungtusanatham, 2011). Estos elementos se detallan a continuación (Hyndman & Athanasopoulos, 2018):

- Tendencia: Una tendencia existe cuando hay un aumento o disminución a largo plazo en los datos. Puede ser lineal o no.
- Estacionalidad: Una serie de tiempo tiene estacionalidad cuando se ve afectada por factores estacionales como un día de la semana o un momento determinado del año. La estacionalidad es siempre fija y de frecuencia conocida.

Para determinar si una serie es estacional se debe ver si sus valores están autocorrelacionados. La autocorrelación mide la relación lineal entre los valores rezagados de una serie de tiempo, donde se tiene que en el periodo donde hay mayor autocorrelación es la frecuencia de la estacionalidad.

Las series de tiempo que no presentan estacionalidad se llaman: ruido blanco (*white noise*). Es en estas series que se espera que su autocorrelación sea cercana a 0. Para probar si una serie es realmente ruido blanco se puede aplicar la prueba estadística de Ljung-Box, el cual tiene como hipótesis nula: “la autocorrelación de los datos es 0, es decir, los datos se distribuyen de forma independiente” (Aguirre, 1994).

- Errores: Una serie de tiempo tiene un término de error aleatorio, que es conocido como “ruido”. La autocorrelación de los errores siempre debe ser cercana a 0, ya que eso representa su aleatoriedad.

c. Modelos de estimación de demanda

Esta sección se basa en el libro “*Administración de Operaciones: Conceptos y casos contemporáneos*” de Roger G. Schroeder, Susan Meyer Goldstein y M. Johnny Rungtusanatham de la Universidad de Minnesota y en el trabajo de Rob J. Hyndman y George Athanasopoulos de la Universidad de Monash (Australia), en su libro “*Forecasting: Principles and Practice*”.

Los modelos de estimación de demanda son variados para las diferentes aplicaciones que puedan tener y también dependiendo de los datos disponibles a usar.

En el trabajo de título se buscará ir de modelos más simples a más complejos, para también poder determinar si la complejidad se traduce en mejores pronósticos. A continuación se detallan los modelos que pueden ser utilizados en el desarrollo de la metodología (Hyndman & Athanasopoulos, 2018).

b.1 Promedio móvil simple

El método del promedio móvil simple es uno de los más sencillos para estimar demanda y puede resultar efectivo en series de tiempo que tienen un comportamiento estable. Este método considera que no hay presencia de estacionalidad, tendencia ni componentes cíclicos en los datos de demanda. (Schroeder, Meyer Goldstein, & Rungtusanatham, 2011).

Al estimar demandas futuras con promedio móvil simple, se considera que todos los valores futuros son iguales al promedio de los datos históricos. A continuación, se calcula la demanda promedio (A_t) para los períodos 1 a t , que finalmente será el pronóstico para el siguiente periodo ($t+1$):

$$F_{t+1} = A_t = \frac{D_1 + D_2 + \dots + D_t}{t}$$

Es uno de los métodos que pueden servir como *benchmark* para comparar modelos más complejos.

Existe una variación de este modelo llamada promedio móvil doble, que trata de pronosticar series de tiempo que tienen una tendencia lineal. Se calcula un conjunto de promedios móviles y después se calcula un segundo conjunto como promedio móvil del primero (Pineda, 1999).

En primer lugar se tiene el promedio móvil simple:

$$S_{t+1} = \frac{D_1 + D_2 + \dots + D_t}{t}$$

Se introduce el concepto de promedio móvil doble:

$$S'_{t+1} = \frac{S_1 + S_2 + \dots + S_t}{t}$$

Luego, se definen los parámetros a y b :

$$a = 2S_{t+1} - S'_{t+1}$$

$$b = \frac{2(S_{t+1} - S'_{t+1})}{t - 1}$$

El pronóstico final será la ecuación de la recta:

$$F_{t+1} = a + b * m$$

Donde m representa el número de periodos futuros a proyectar.

b.2 Suavización exponencial simple

El método de suavización exponencial simple sirve para predecir datos que no tienen tendencia o patrones estacionales. Se añade más complejidad con respecto al modelo anterior ya que se le otorga cierta relevancia (peso) a los datos, dependiendo si son recientes o no. Las predicciones son calculadas usando promedios ponderados, donde los pesos pueden disminuir o aumentar exponencialmente mientras las observaciones se alejan del periodo actual.

El promedio ponderado (A_t) que calcula este método es entre la observación de demanda más reciente (D_t) y el promedio de demandas observadas antes del periodo t (A_{t-1}). Esto queda como sigue:

$$\begin{aligned} A_t &= \alpha D_t + (1 - \alpha)A_{t-1} \\ 0 &< \alpha < 1 \end{aligned}$$

Donde α es el parámetro de suavización. Si este parámetro es pequeño (cercano a 0), entonces más peso es dado a observaciones en el pasado más distante. Cuando α es igual a 1, entonces las predicciones serán exactamente al dato del período anterior. El pronóstico para el siguiente periodo es este promedio ponderado:

$$F_{t+1} = A_t$$

También existe el modelo de suavización exponencial doble, que intenta agregar tendencia al modelo de suavización exponencial (Pineda, 1999). Se agrega una ecuación para la tendencia y una constante de suavización para la tendencia (β). Las ecuaciones para este método quedan como sigue:

$$\begin{aligned} A_t &= \alpha D_t + (1 - \alpha)(A_{t-1} + T_{t-1}) \\ T_t &= \beta(A_t - A_{t-1}) + (1 - \beta)T_{t-1} \end{aligned}$$

$$F_{t+1} = A_t + T_t$$

b.3 Método Holt-Winters

El método de Holt-Winters de tendencia lineal extiende la suavización exponencial simple para permitir estimar datos que tienen tendencia e incluye 3 ecuaciones: una para la predicción, una para el nivel de los datos y otra para la tendencia.

$$\begin{aligned}F_{t+1} &= l_t + hb_t \\l_t &= \alpha A_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \\b_t &= \beta^*(l_t + l_{t-1}) + (1 - \beta^*)b_{t-1}\end{aligned}$$

Donde:

- A_t es el promedio de las observaciones de demanda hasta el periodo t y F_{t+1} es el pronóstico del periodo $t+1$,
- l_t se refiere al nivel de los datos al tiempo t ,
- b_t es la estimación de la tendencia (pendiente) de la serie al tiempo t ,
- α y β^* son los parámetros de suavización para el nivel y tendencia, respectivamente.

También existe el método de Holt-Winters que incluye estacionalidad. Este tiene 3 ecuaciones de suavización: una para el nivel, una para la tendencia, y otra para la estacionalidad. Sus ecuaciones son:

$$\begin{aligned}F_{t+1} &= l_t + hb_t + s_{t+h-m(k+1)} \\l_t &= \alpha(A_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \\b_t &= \beta^*(l_t + l_{t-1}) + (1 - \beta^*)b_{t-1} \\s_t &= \gamma^*(y_t - l_t) + (1 - \gamma^*)s_{t-m}\end{aligned}$$

Este componente de estacionalidad podrá ser aditiva o multiplicativa.

b.4 Modelos ARIMA (Autoregressive Integrated Moving Average)

Los modelos ARIMA tienen un enfoque diferente a los pronósticos de series de tiempo. Mientras que los modelos de suavización exponencial y promedios móviles analizan principalmente en las tendencias y estacionalidades de los datos, los modelos ARIMA se basan en las autocorrelaciones en los datos para determinar sus pronósticos. Se pueden aplicar en series de tiempo que son “ruido blanco”, es decir, que no presentan tendencia ni estacionalidad y que tienen media y varianza constante.

Los modelos ARIMA sin estacionalidad combinan los modelos de autorregresivos (AR) y de media móvil (MA, *moving average*). A continuación se detallan estos modelos antes de especificar los modelos ARIMA a utilizar en este trabajo:

- Modelos autorregresivos (AR): Se hacen predicciones de una variable de interés (\mathbf{y}_t) usando una combinación lineal de los valores pasados de esa variable. Es decir, se hace una regresión de la variable consigo misma. Esto se puede escribir como sigue:

$$\mathbf{y}_t = \mathbf{c} + \phi_1 \mathbf{y}_{t-1} + \phi_2 \mathbf{y}_{t-2} + \dots + \phi_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

Donde $\boldsymbol{\varepsilon}_t$ es ruido blanco. La ecuación especificada es un modelo AR(p), un modelo autorregresivo de orden p.

- Modelos de media móvil (MA, *moving average*): Este modelo ocupa errores de pronóstico pasados para realizar predicciones, en un símil a una regresión.

$$\mathbf{y}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t + \theta_1 \boldsymbol{\varepsilon}_{t-1} + \theta_2 \boldsymbol{\varepsilon}_{t-2} + \dots + \theta_q \boldsymbol{\varepsilon}_{t-q}$$

Donde $\boldsymbol{\varepsilon}_t$ es ruido blanco. La ecuación especificada es un modelo MA(q), un modelo de media móvil de orden q.

Combinando estos dos modelos con la diferenciación, que es transformar una serie de tiempo no-estacionaria a una estacionaria, se obtiene un modelo ARIMA.

El modelo ARIMA(p,d,q) se puede escribir como:

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Donde y'_t es la serie estacionaria (ruido blanco), y los “regresores” ($y'_{t-1}, \dots, y'_{t-p}, \varepsilon_{t-1}, \dots, \varepsilon_t$) son valores rezagados de y_t y de los errores. En la notación ARIMA(p,d,q) se tiene que p es el orden de la parte autorregresiva, d es el grado de la diferenciación implicada, y q es el orden de la parte de media móvil. Hay diversos casos especiales de los modelos ARIMA dependiendo de estos parámetros.

d. Métricas de error

Algunas métricas de error a considerar cuando se desarrollen los modelos de estimación de demanda son (Hyndman R. , 2014):

1. Errores dependientes de la escala de datos: estos errores están en la misma escala de los datos. No sirve para hacer comparaciones entre series que no tienen la misma escala. Algunos de los más usados son: (Considerando el error como $e_t := Demanda_t - Forecast_t$)

Error promedio absoluto (*Mean Absolute Error*): $MAE = \sum_{t=1}^n \frac{|e_t|}{n}$

Error promedio cuadrado (*Root Mean Squared Error*): $RMSE = \sqrt{\sum_{t=1}^n \frac{|e_t|^2}{n}}$

2. Errores porcentuales: Los errores son presentados en porcentajes y tienen la ventaja que son independientes de cualquier escala, por lo que sirven en comparaciones. El más usado es el error absoluto porcentual promedio (Mean absolute percentage error, $MAPE = promedio(\frac{100 * errores}{y})$).

Porcentaje de error absoluto (*Mean Absolute Error*): $MAPE = \sum_{t=1}^n \frac{|e_t/D_t|}{n} * 100\%$

VI. METODOLOGÍA

La metodología del trabajo se basa principalmente en la metodología de minería de datos CRISP-DM, con algunos detalles específicos al presente trabajo. Se detalla a continuación.

a. Estudio de la situación actual

En esta etapa se analiza el contexto actual de la empresa farmacéutica a estudiar. Se indaga en sus productos, su terminología y sus objetivos.

Además, se detalla en profundidad su proceso de estimación de demanda. Se determina la situación actual de la empresa farmacéutica estudiada, de manera que queda definida por métricas que sirvan como “línea base” para comparar los posteriores resultados del trabajo de título.

b. Limpieza y procesamiento de datos

En esta etapa de la metodología se realiza la recopilación, limpieza y procesamiento de los datos disponibles. Esto implica: crear y consolidar bases de datos provenientes de diferentes fuentes dentro de la empresa, identificar valores atípicos, identificar valores no realistas que son errores, entre otras tareas. Se determina el procedimiento con que tratar estos valores atípicos, es decir, si eliminarlos del análisis o hacerles alguna transformación.

Además, se busca la manera de que los datos de ventas históricas se aproximen lo más posible a la demanda a la que se pudo haber visto afectada la empresa farmacéutica.

c. Selección de medicamentos a estudiar y nivel de agregación de la estimación

Se busca caracterizar la demanda de los distintos medicamentos para así identificar patrones dentro de sus ventas históricas y según sus atributos, para lograr algún tipo de agrupación. Es aquí donde se determinan qué series de tiempo presentan estacionalidad.

Se busca la agrupación de medicamentos para seleccionar un set de productos. Es a este set a los cuales aplicar los modelos de estimación de demanda, de tal manera que estos sean productos representativos de todo el portafolio de la empresa.

d. Elección de modelos a utilizar y su estimación

En esta etapa se elegirán los modelos a utilizar, para así construirlos y ejecutarlos en algún software estadístico. Estos modelos son escogidos dependiendo de los medicamentos, de forma que cada medicamento tenga modelos específicos para sus atributos.

Se buscan modelos simples y complejos, tal de ver si la complejidad aporta en al estimar demanda en esta empresa farmacéutica.

e. Evaluación de modelos

Se comparan todos los modelos elegidos en base a su ajuste (métricas de error) y su capacidad de pronóstico en el periodo de evaluación.

Se estima cuánta ganancia (o pérdida) significaría implementar cada modelo. Luego se escoge el modelo que más se ajusta a lo que busca la empresa, para cada uno de sus productos. Además, se hace una validación de los modelos en meses pasados de las ventas históricas.

f. Propuesta de desarrollo y aplicación

Se determina cómo el modelo seleccionado apoya a la gestión de la demanda. Aquí se define en qué parte del proceso se inserta, quiénes la usan y con qué grado de confiabilidad se implementa la herramienta.

VII. DESARROLLO DE LA METODOLOGÍA Y ANÁLISIS DE RESULTADOS

a. Estudio de la situación actual

La empresa farmacéutica en estudio sólo comercializa y distribuye medicamentos que deben ser vendidos a pacientes que tengan una receta médica. Este tipo de productos no pueden ser promocionados ni desplegados en el mostrador de farmacias.

Como se mencionaba en apartados anteriores, los medicamentos se separan en dos líneas de negocio: FARMA y OLE. La línea FARMA agrupa 96 medicamentos que tratan condiciones comunes en la población, es en esta línea de negocio que se pueden encontrar productos como: paracetamol, anticonceptivos, antibióticos, entre otros. Mientras que la línea OLE (Oncología, Licencia y Especialidades) concentra 35 productos que van enfocados a enfermedades y condiciones específicas. Dentro de las líneas de negocio expuestas, existen otras divisiones que se muestran en la Tabla 1 en conjunto con su descripción.

Línea	Unidad	Descripción	Cantidad	
FARMA	CARDIO METABÓLICA	Área de cardiología y metabolismo. Concentra medicamentos que tratan la obesidad, hipertensión, entre otros.	21	96
	DOLOR	Trata el dolor en sus distintas formas. Incluye antiinflamatorios y analgésicos, generalmente recetados para tratamientos de artrosis y artritis.	11	
	NEUROCIENCIAS	Área de la neurociencia. Agrupa medicamentos antidepresivos, antiepilépticos, hipnóticos y otros que tratan la enfermedad de Alzheimer.	24	

	PRIMARY CARE	Asociados a condiciones comunes. Contiene analgésicos, antiinflamatorios y antibióticos. Generalmente asociados a afecciones agudas de tipo inflamatorio o doloroso.	20	
	UGI	Área de urología y ginecología. Concentra anovulatorios, anticonceptivos, antibióticos y otros que tratan condiciones de esta área.	20	
OLE	ESPECIALIDADES	Asociados a enfermedades muy particulares.	2	35
	ONCO HEMATO	Área onco-hematológicas. Agrupa medicamentos inmunoestimulantes, inmunosupresores y otros que tratan principalmente mieloma múltiple.	12	
	ONCOLOGIA	Área oncológica. Une medicamentos indicados para el tratamiento de múltiples cánceres.	11	
	URO-ONCO	Área uro-oncológica. En su mayoría concentra medicamentos para el tratamiento del cáncer de próstata.	10	

Tabla 1: Descripción de unidades de negocio y sus líneas asociadas

De los 131 productos activos de la empresa, hay 9 productos para los cuales la empresa no realiza predicciones. Esto se debe a que son productos de venta esporádica, donde se tiene un método de *order-up-to* para gestionar el stock. Corresponden a los 2 medicamentos de la unidad de Especialidades, 6 de Oncología y 1 de Onco-Hemato. Además, hay 4 medicamentos nuevos en el catálogo de la empresa los cuales aún no son agregados al proceso de estimación de demanda. Estos productos no son analizados de aquí en adelante.

Actualmente se utilizan datos de venta para hacer pronósticos del resto de los 118 medicamentos. La empresa cuenta con la información de *backorders* desde enero del año 2018. Los *backorders* son los pedidos pendientes que la empresa no pudo satisfacer por no tener suficientes unidades en un determinado período, y pueden ocurrir errores en el pronóstico de unidades o factores externos. En el proceso actual no se incluye información de *backorders*.

Para tener una línea base de la situación actual, y para poder comparar posteriormente los resultados de este trabajo, se pueden tomar las dos métricas que la empresa utiliza para evaluar el proceso de estimación de demanda: *forecast accuracy* y el cumplimiento del *forecast*.

El indicador de *forecast accuracy* sigue la siguiente fórmula:

$$\textit{Forecast accuracy} = 1 - \frac{\textit{Ventas Reales}}{\textit{Forecast}}$$

Es decir, mide el desvío que tiene la predicción con respecto a las ventas reales que se perciben en ese mes, para cada producto de la compañía. Este indicador puede tener:

- Valores negativos, lo que significa que la predicción queda bajo las ventas reales (subestimación de demanda);
- Valores positivos, que es cuando se predice más de lo que realmente se vende (sobreestimación de demanda).

En el siguiente gráfico se puede ver cómo se ha comportado este indicador a lo largo de los años. Los valores mostrados son promedios mensuales del *forecast accuracy* de los productos. Es decir, se calcula el promedio mensual de este indicador para cada producto, y luego estos valores son promediados en cada mes. La evolución histórica del *forecast accuracy* se puede ver con más detalle en el Anexo 1.

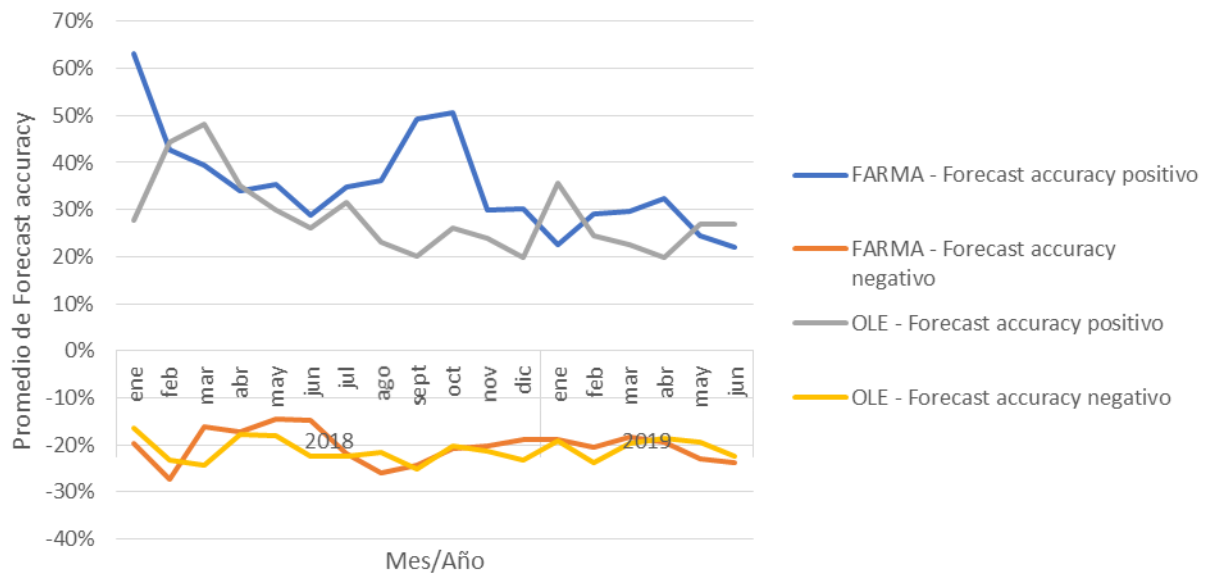


Ilustración 3: Forecast accuracy desde el año 2018 a junio de 2019 por línea de negocio

Ver los valores del *forecast accuracy* positivo y negativo, y no en valor absoluto, ayuda a evidenciar cómo se comportan los desvíos en los pronósticos. En general, el desvío negativo es menor en valor absoluto que el desvío positivo. Esto significa que la empresa tiende a sobreestimar la demanda en más unidades que cuando subestima la demanda. Los medicamentos inciden directamente en la vida de las personas, por lo que el tener que vender un producto con descuento o llegar a destruirlo se preferirá antes que quedarse sin stock. En los últimos meses se ha logrado disminuir el porcentaje de desvíos positivos.

Existe una diferencia entre los valores de *forecast accuracy* de las diferentes líneas y dentro de sus unidades también. Desagregando por las unidades de negocio dentro de las líneas se ve lo siguiente:

<i>Forecast accuracy</i>	positivo		negativo	
	2018	2019	2018	2019
FARMA	43 %	27 %	-19 %	-21 %
CARDIO METABOLICA	44 %	21 %	-21 %	-18 %
DOLOR	64 %	27 %	-16 %	-18 %
NEUROCIENCIAS	36 %	32 %	-20 %	-22 %
PRIMARY CARE	51 %	39 %	-24 %	-30 %
UGI	23 %	15%	-16 %	-16 %

OLE	29 %	23 %	- 20 %	-19 %
ONCO HEMATO	31 %	18 %	-24 %	-20 %
ONCOLOGIA	21 %	10 %	-16 %	-13 %
URO-ONCO	36 %	39 %	- 22 %	-25 %

Tabla 2: Forecast accuracy promedio por línea y unidad de negocios en años 2018 y 2019

La línea FARMA tiene porcentajes de *forecast accuracy* positivos más altos que la línea OLE, donde las unidades que tendrían menos certeza en sus pronósticos son Primary Care, Neurociencias y Dolor. Mientras que la con peor certeza en la línea OLE es Uro-Onco y Onco-Hemato. Estos resultados se podrían deber a algún medicamento en particular dentro de las unidades. Se analiza este punto más adelante.

El segundo indicador utilizado en el proceso de estimación de demanda es el de cumplimiento del *forecast*. Este mide cuántos productos “cumplen” el pronóstico estimado y tiene la siguiente forma:

$$Cumplimiento_i = \begin{cases} 1 & \text{si forecast accuracy del producto } i \in [-20\%, 20\%] \\ 0 & \text{si no} \end{cases}$$

$$Cumplimiento \text{ de forecast} = \frac{\sum_{i=1}^n Cumplimiento_i}{n}$$

El indicador de cumplimiento se calcula para todos los productos de la compañía cada mes. Así, la empresa puede ver el porcentaje de cumplimiento mes a mes. La evolución de este último indicador se puede ver en la siguiente figura, por línea de negocio.

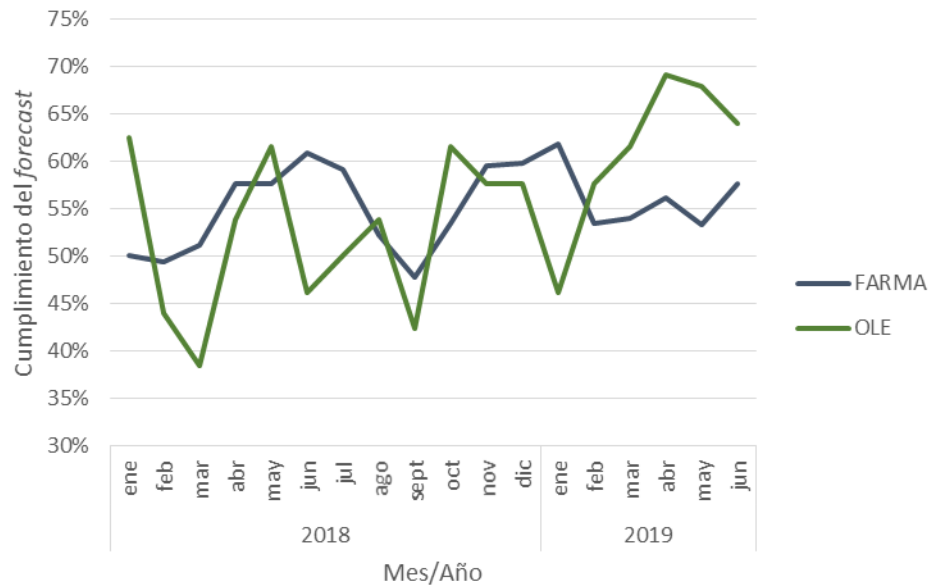


Ilustración 4: Cumplimiento de forecast desde el año 2018 al 2019 por línea de negocio

Desagregando por líneas y unidad de negocio se ve que el cumplimiento de *forecast* varía dependiendo de las características de los productos, como se ve en la siguiente tabla:

Cumplimiento de <i>forecast</i>		
Unidad	2018	2019
FARMA	55 %	56 %
CARDIO METABOLICA	52 %	63 %
DOLOR	54 %	52 %
NEUROCIENCIAS	54 %	50 %
PRIMARY CARE	46 %	42 %
UGI	68 %	73 %
OLE	52 %	61 %
ONCO HEMATO	46 %	70 %
ONCOLOGIA	65 %	93 %
URO-ONCO	53 %	34 %

Tabla 3: Cumplimiento de forecast por línea y unidad de negocio

La empresa ha tenido un aumento de su cumplimiento en el último año, sobre todo en la línea OLE. El principal cambio entre diciembre 2018 y enero 2019 es la implementación de una nueva etapa en el proceso de estimación de demanda: la reunión con más expertos del laboratorio para determinar los pronósticos. Se puede inferir que al tener mayor cantidad de perspectivas del negocio los sesgos de la fuerza de ventas se han logrado disminuir y así mejorar los indicadores del proceso.

Aun así, unidades como Primary Care, Neurociencias, Dolor y Uro-Onco siguen teniendo porcentajes de cumplimiento muy. Analizar por unidad no ayuda a esclarecer la razón de esta disminución, ya que dentro de cada unidad hay gran variación en los porcentajes de cumplimientos de los principios activos, y por ende, de los medicamentos. Por ejemplo, la unidad de Neurociencias agrupa 13 principios activos, donde 4 de ellos tienen porcentajes cercanos al 100%, mientras que el resto promedia un 36% de cumplimiento. En el Anexo 2 se muestra en detalle el porcentaje de cumplimiento para cada uno de los principios activos.

Con este detalle se puede empezar a notar que cada medicamento dentro de las unidades y líneas de negocio presenta una situación particular, donde pareciera que la empresa logra captar muy bien el comportamiento de algunos productos, como los de Oncología, pero no tanto con otros como los de Primary Care.

b. Procesamiento y limpieza de datos

b.1 Datos disponibles

Para hacer la predicción de demanda de los diferentes productos de la compañía se dispone de los siguientes datos:

- Ventas históricas de los 118 productos a analizar, desde el año 2018 hasta agosto del 2019. Este set de datos contiene:

Variables	Detalle
Fecha de facturación	Fecha en que se concreta la facturación y así la fecha en que se despacha el producto. En el formato día/mes/año.
Cliente que hace la compra	Cliente a quien se le facturan los productos. Este cliente puede ser un hospital, municipalidad, clínica, farmacia, droguería, entre otros. Siempre es una institución certificada, nunca una persona natural. Se cuenta con un código único para cada cliente y su alias.

Canal	Canal al que pertenece la venta: institucional, privado o pseudo-privado. Esta información también ayuda a describir al cliente que hace la compra.
Producto (SKU)	Nombre y código del producto facturado. Hay información para los 118 productos, es decir, 118 SKUs.
Línea de negocio	Línea de negocio a la que pertenece el producto. Existen 2 líneas de negocio: FARMA y OLE.
Unidad de negocio	Unidad de negocio a la que pertenece el producto. Existen 8 líneas de negocio.
Familia	Familia a la que pertenece el producto, se refiere al principio activo. Existen 58 familias.
Cantidad Facturada	Cantidad facturada, es decir, la cantidad de cada medicamento (SKU) que el cliente compra.
Precio del producto	Precio del producto facturado.
Porcentaje de descuento	Porcentaje de descuento que se le hace a un producto. En caso de no existir es 0%.

Tabla 4: Detalle de variables de archivo de ventas históricas

- *Backorders* de productos con algún evento de esta índole, desde el año 2018 hasta agosto del 2019.

Variables	Detalle
Fecha de <i>backorder</i>	Fecha en que se registra el evento de <i>backorder</i> . En el formato día/mes/año.
Cliente que pide <i>backorder</i>	Cliente que registra el evento de <i>backorder</i> .
Producto	Nombre y código del producto con evento de <i>backorder</i> .
Cantidad de <i>backorder</i>	Cantidad que falta para cumplir el pedido del cliente.

Tabla 5: Detalle de variables de archivo de *backorders*

Como el objetivo de este trabajo es estimar la demanda de los productos de la empresa farmacéutica, se quiere obtener una aproximación lo más cercana posible a datos de demanda real. Es por eso que se suman los eventos de *backorder* a las ventas, considerando que se venden al precio lista del producto sin porcentaje de descuento. Desde ahora en adelante a esta variable se le llama “demanda”. La base de datos consolidada queda con una estructura igual a la de ventas históricas original, es decir,

se tiene la fecha de compra, cliente que hace la compra, el canal al que pertenece la compra, cantidad facturada, precio y descuento si corresponde.

La estimación de demanda de la empresa en estudio se hace mensualmente porque así es la política de pedidos, pero se decide agrupar los datos por semana. Esto es principalmente por 2 razones. En primer lugar, porque consolidarla semanalmente permite tener más datos con los cuales estimar posteriores periodos. Y en segundo lugar, muchos de los pedidos se hacen en semanas en particular de algún mes por lo que se podría capturar algún tipo de estacionalidad semanal de esta manera. Luego de estimar las demandas semanales, se pueden sumar para obtener la demanda mensual para, más adelante, evaluar los modelos.

b.2 Limpieza de datos

Antes de comenzar a indagar en los datos y posibles comportamientos de los productos, se sigue a la limpieza de datos. En primer lugar, se filtran todas las ventas que corresponden a licitaciones, es decir, al canal institucional. Estas no entran dentro del estudio ya que siguen otro proceso de estimación de demanda asociado a licitaciones, y corresponden al 6% de las ventas.

En la historia de ventas de la empresa han ocurrido quiebres de stock, ventas *spot* de cantidades muy altas o simplemente situaciones puntuales que han hecho tener valores atípicos en la demanda de algunos productos. El hecho de consolidar ventas históricas con información de *backorders* logra solucionar los eventos de quiebres de stock, ya que agrega las unidades demandadas pero que la empresa no tenía disponible.

Luego, exhaustivamente para cada uno de los 118 productos, se identifican los *outliers* que aún puedan existir ocupando el método no paramétrico *Interquantile Range* (IQR). Este método identifica valores que están 1,5 veces por sobre o bajo la distancia desde el primer y tercer cuartil de los datos.

Los valores atípicos de 106 productos son interpolados de tal manera que son el promedio de 2 períodos anteriores, y de los 2 siguientes. Se cree que de esta manera se pueden suavizar ventas *spot* y eventos que no se volverán a repetir.

El resto de los productos están en un período de alza o baja en los últimos meses. Es por eso que la herramienta seleccionada para identificar *outliers* podría no ser útil en estos casos. Para estas series de tiempo no se interpolan los valores atípicos ya que una tendencia al alza o a la baja es información valiosa para un modelo de estimación de demanda. Conservar estas características permite capturar la naturaleza de estos productos.

b.2 Aproximación a una caracterización de la demanda

Con la base de datos post-limpieza, se procede a analizar la demanda de los productos. Se puede ver cómo se comporta la demanda para las dos líneas de negocio definidas por la empresa.

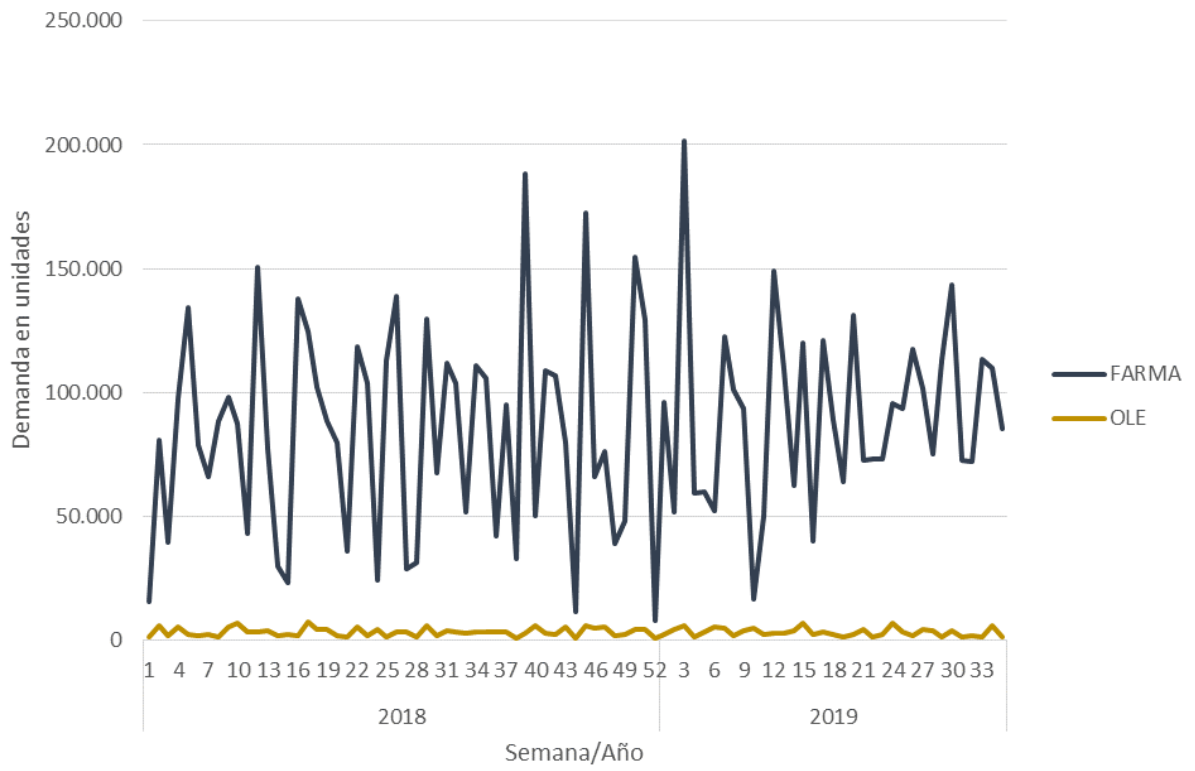


Ilustración 5: Demanda mensual por línea de negocio

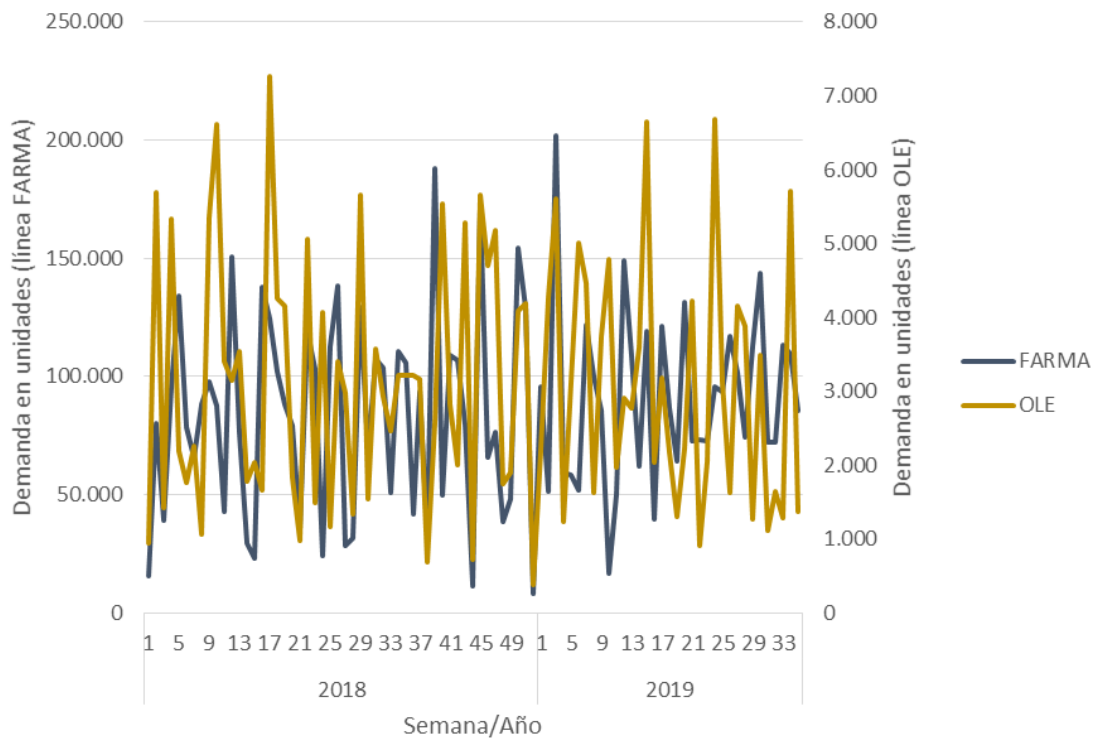


Ilustración 6: Demanda en unidades por línea de negocio, versión 2

Lo primero que se puede ver en la Ilustración 5 es la naturaleza de ambas líneas de negocio, descrita en la sección de Antecedentes Generales. La línea de FARMA agrupa medicamentos de uso común que concentran la mayor parte de las unidades vendidas, oscilando entre las 8.000 unidades aproximadamente a 200.000 unidades semanales. Mientras que la línea OLE, tiene un nivel de ventas mucho menor por tratar condiciones más específicas, moviéndose entre las 300 unidades a 7.000 unidades semanales.

Con respecto a las variaciones semanales de cada una de las líneas hay *peaks* y valles. En la Ilustración 5 y 6 se ve que las grandes cantidades de unidades demandadas se concentran en las semanas pares. Esto se puede deber a los calendarios de pedidos de los clientes.

A continuación, se puede ver la comparación de utilidades entregadas por cada línea en el año 2018 y 2019.

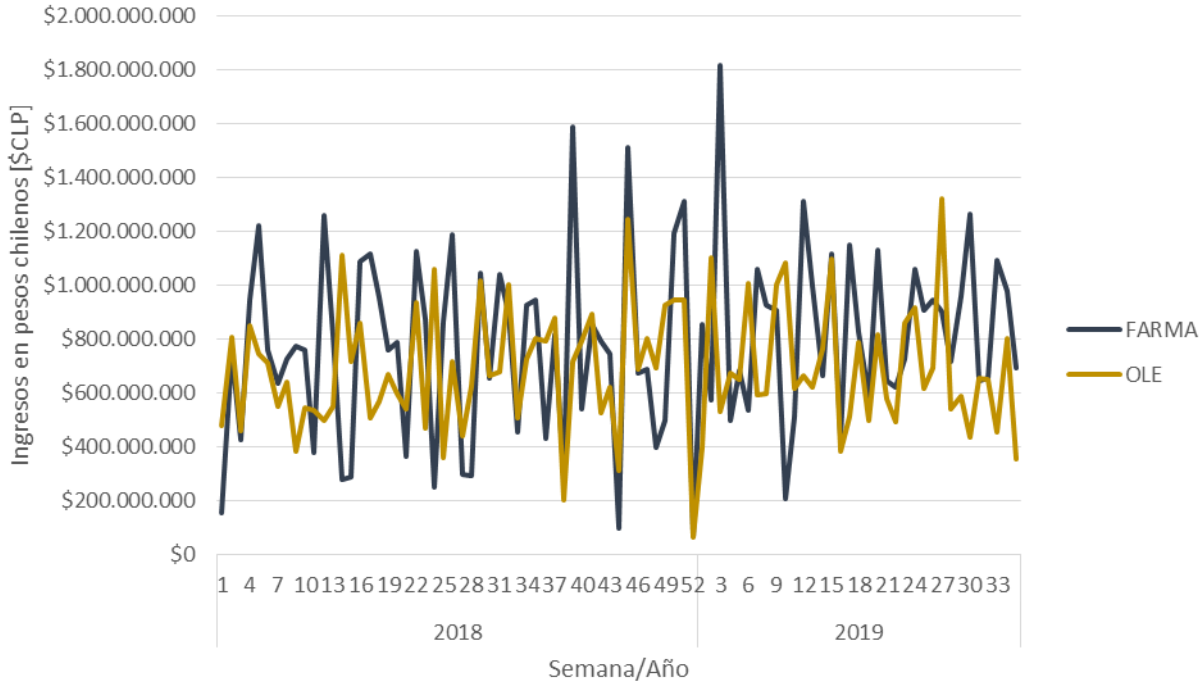


Ilustración 7: Utilidades reportadas semanalmente por línea de negocio

De la Ilustración 7 se puede ver que los ingresos reportados por cada línea son similares. Esto se debe a que los medicamentos de la línea OLE, al tratar condiciones específicas y particulares, tienen valores más elevados. Es más, hay un medicamento dentro de esta línea que ronda los \$8.000.000 por unidad. En promedio el valor de un medicamento vendido de la línea OLE en el año 2019 es de \$1.360.000 aproximadamente, mientras que el promedio de un producto vendido perteneciente a la línea FARMA es de \$12.730 aproximadamente.

La clara diferencia de ingresos por unidad de producto, dependiendo de qué medicamento se trata hace ver algo importante para la posterior estimación de demanda. Subestimar una unidad de la línea OLE es más perjudicial que subestimar una unidad de la línea FARMA. Esto se debe tener presente al elegir los modelos para cada producto.

La demanda se puede desagregar además en las unidades de negocio. En particular, para las unidades dentro de las líneas de negocio de la empresa se muestran a continuación.

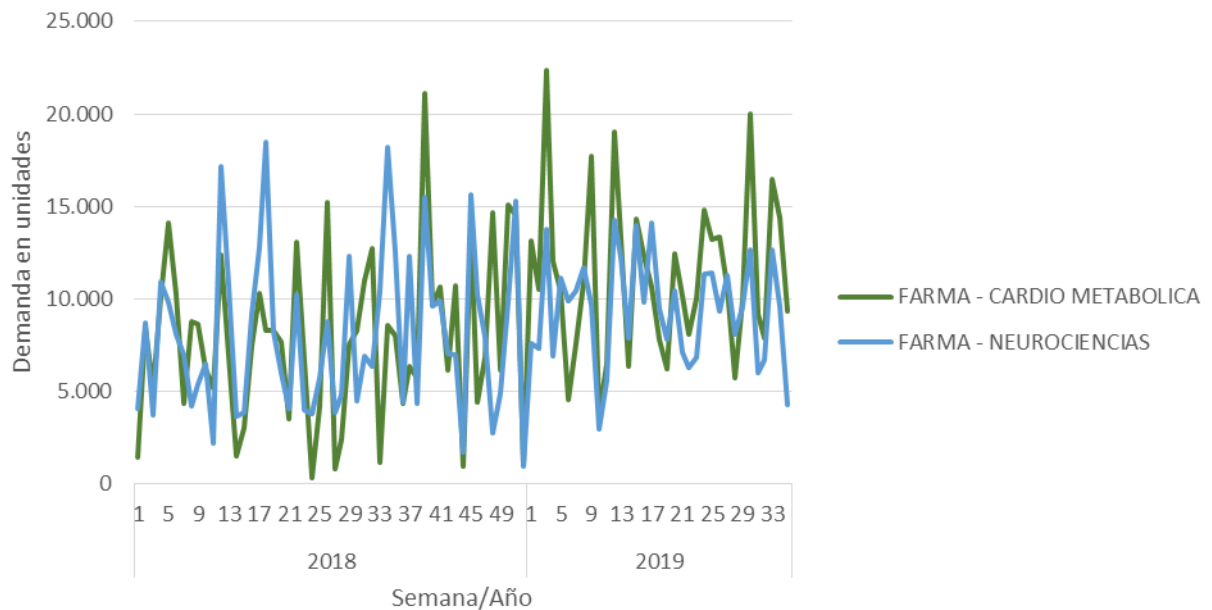


Ilustración 8: Demanda semanal por unidad de la línea de negocio FARMA, unidades Cardio metabólica y Neurociencias

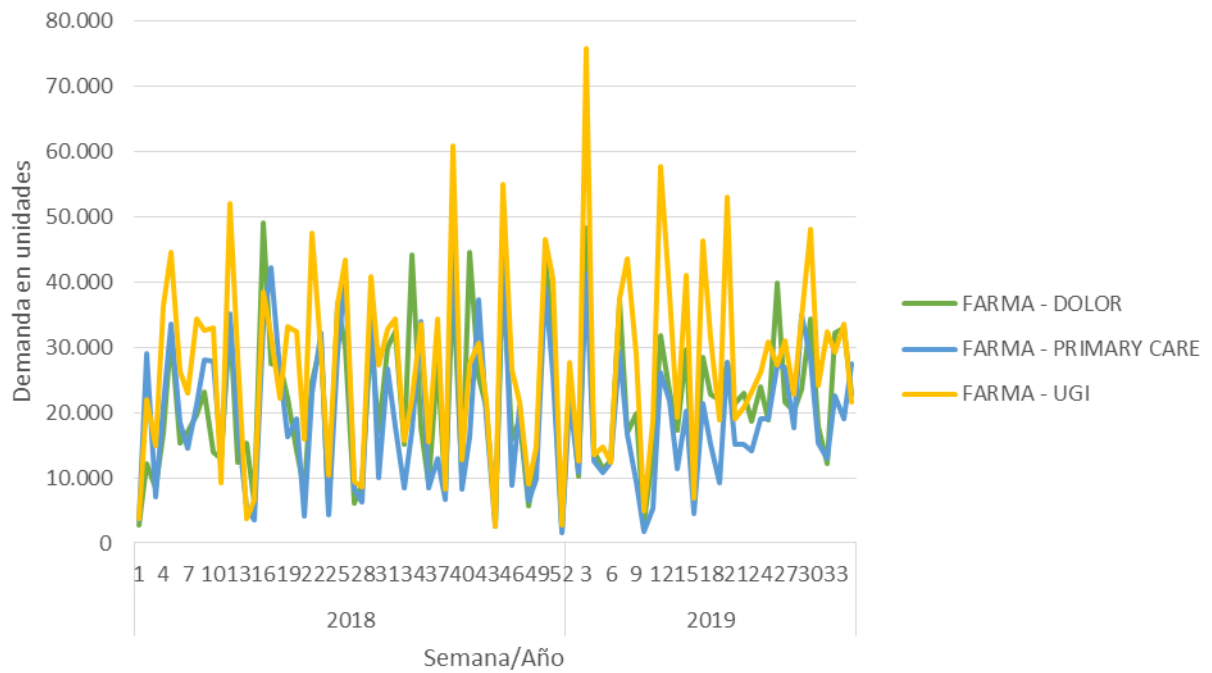


Ilustración 9: Demanda semanal por unidad de la línea de negocio FARMA, unidades Primary Care, Dolor y UGI

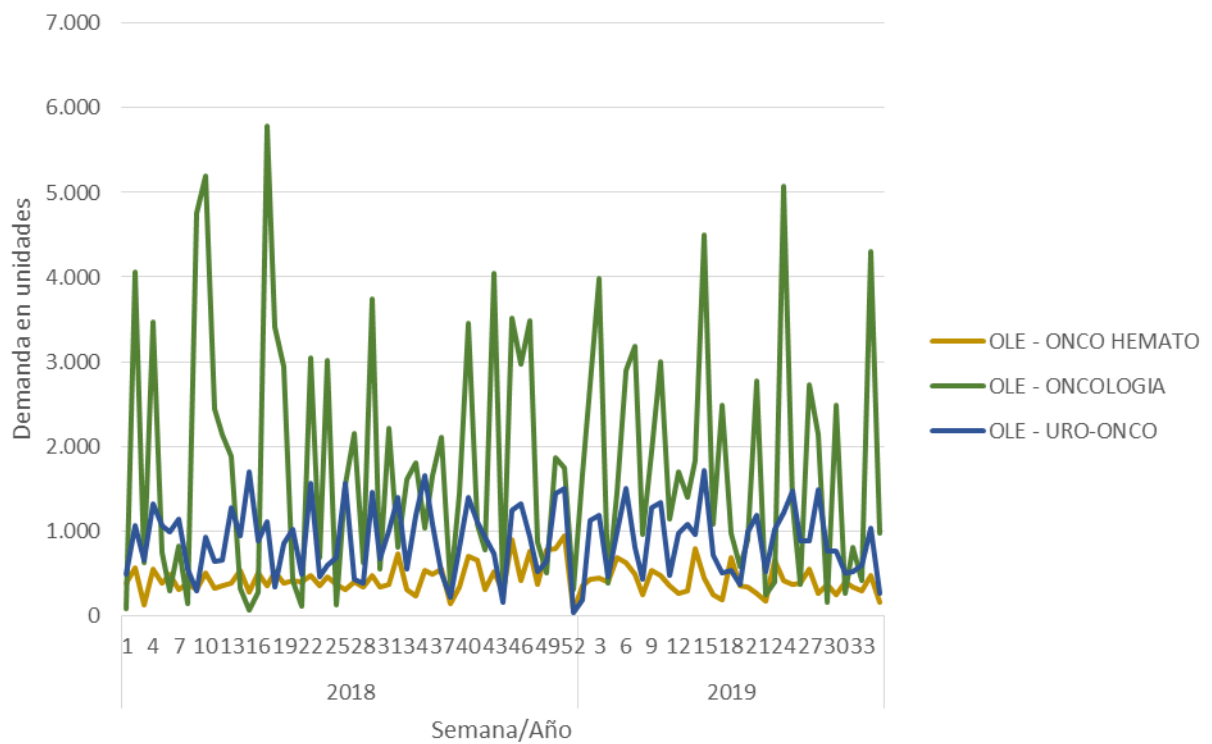


Ilustración 10: Demanda semanal por unidad de la línea de negocio OLE

Cada una de las unidades de negocio se comporta de manera diferente. No se indaga más en estas características porque cada una de ellas agrupa medicamentos que atienden a enfermedades o condiciones similares, pero no por eso significa que dentro de ellas tengan un comportamiento similar. Por ejemplo, en la línea Neurociencias existen medicamentos antipsicóticos para el tratamiento de la esquizofrenia, y medicamentos ansiolíticos para el tratamiento de la ansiedad. Ambos tratan dolencias del área de la Neurociencia, pero se recetan de manera distinta para enfermedades diferentes por lo que habría de esperar que se comporten diferente. Se indaga más en este punto y en la estacionalidad de las series en la siguiente sección.

Para entender y lograr caracterizar la demanda de cada producto también interesa desglosarla en los canales de venta: privado y pseudo-privado, ya que la estimación también es a nivel de tipo de cliente. A continuación se puede ver cómo se comportan las demandas de ambos canales.

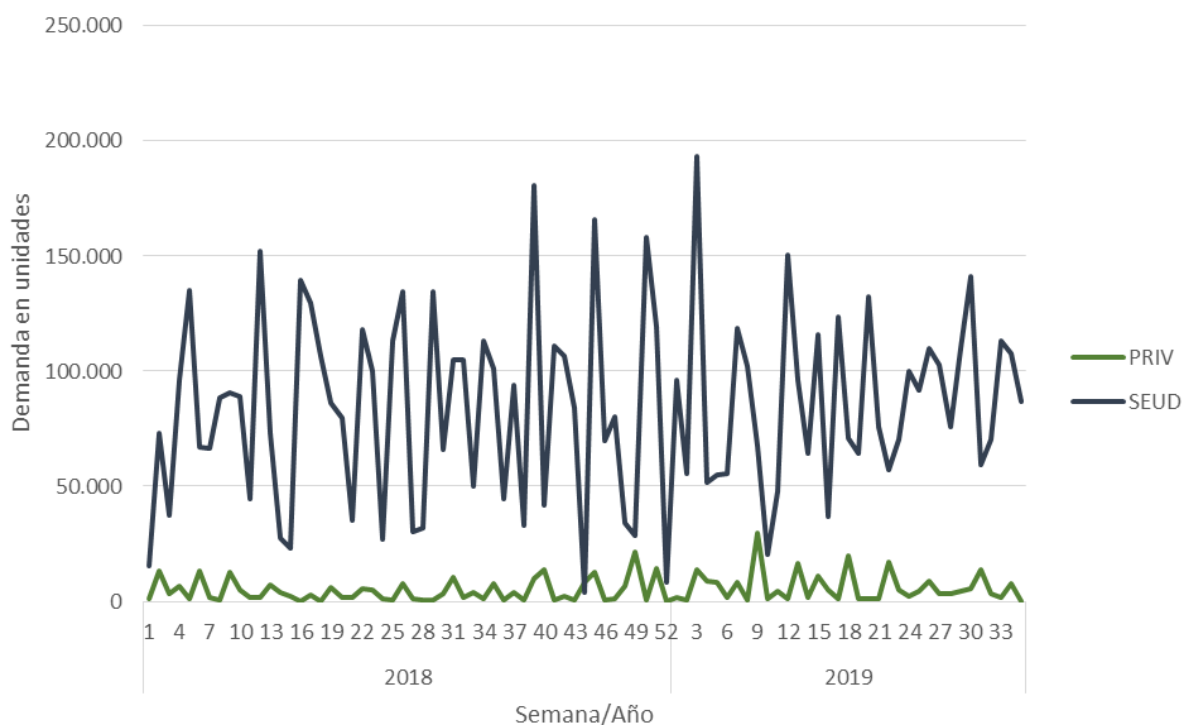


Ilustración 11: Demanda mensual por modalidad de venta

Se ve una clara diferencia entre el canal pseudo-privado y privado, donde es el primero el que concentra más unidades vendidas a lo largo de las semanas. Este comportamiento tiene sentido en la realidad, ya que es donde regularmente el cliente final adquiere los medicamentos. Mientras que es menos común que se vendan o dispensen medicamentos dentro de alguna institución privada de la salud y se hace en casos particulares.

También, se puede ver cómo se comportan las utilidades reportadas semanalmente.

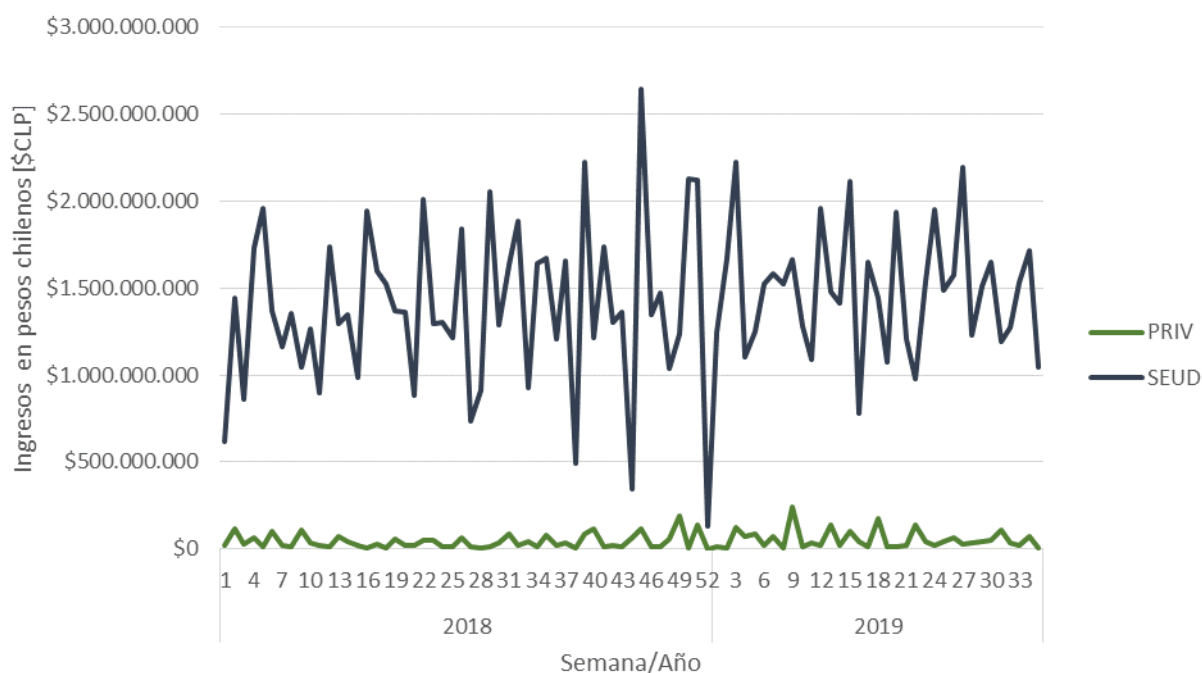


Ilustración 12: Utilidades reportadas semanalmente por modalidad de venta

Las utilidades calculadas son simplemente el precio por la cantidad vendida en la semana. Esto hace reforzar el hecho que el canal pseudo-privado que más demanda medicamentos a la empresa.

El desglosar la demanda por canal de venta, línea y unidad de negocio ayuda a entender la realidad de la empresa, saber cómo se distribuyen los clientes y qué es lo que más se demanda. Esta información sirve para ver cierta información que es útil para posteriormente seleccionar los medicamentos a utilizar. En el Anexo 3 se puede ver más en detalle la desagregación de demanda por unidad de venta y línea de negocio.

c. Selección de medicamentos a estudiar y nivel de agregación de la estimación

Antes de seleccionar los medicamentos a estudiar se determina tener un grupo de entrenamiento y validación, donde todo el análisis que se haga de este punto en adelante sea el grupo de testeo.

Esto se hace para no contaminar las predicciones con información que se predecirá en el futuro. Como testeo se toma todo el año 2018 y las 30 primeras semanas del 2019 (semana del 1° de enero del 2018 a la semana del 22 de julio del 2019). Como set de datos de calibración quedan las siguientes 5 semanas del 2019 (semana del 5 de agosto del 2019 al 26 de agosto del 2019).

c.1 Demanda de medicamentos por principio activo y canal de venta

Para seleccionar medicamentos a estudiar resulta intuitivo pensar que medicamentos dentro de cada línea y unidad de negocio se comportan de manera similar, ya que tratan condiciones en áreas de salud similares. Pero como se ha mencionado anteriormente esto no sucede en esta empresa. Las unidades de negocio se hacen por área de la salud en que se enfocan y no por su función. Cada medicamento trata alguna particularidad específica de una enfermedad lo que no los hace ni sustitutos ni competidores entre sí.

De manera de ejemplificar esa realidad, a continuación se muestran en un mismo gráfico la demanda de algunos productos de la línea FARMA de la unidad Dolor. Una comparación de productos por las otras unidades de negocio se puede ver en el Anexo 4.

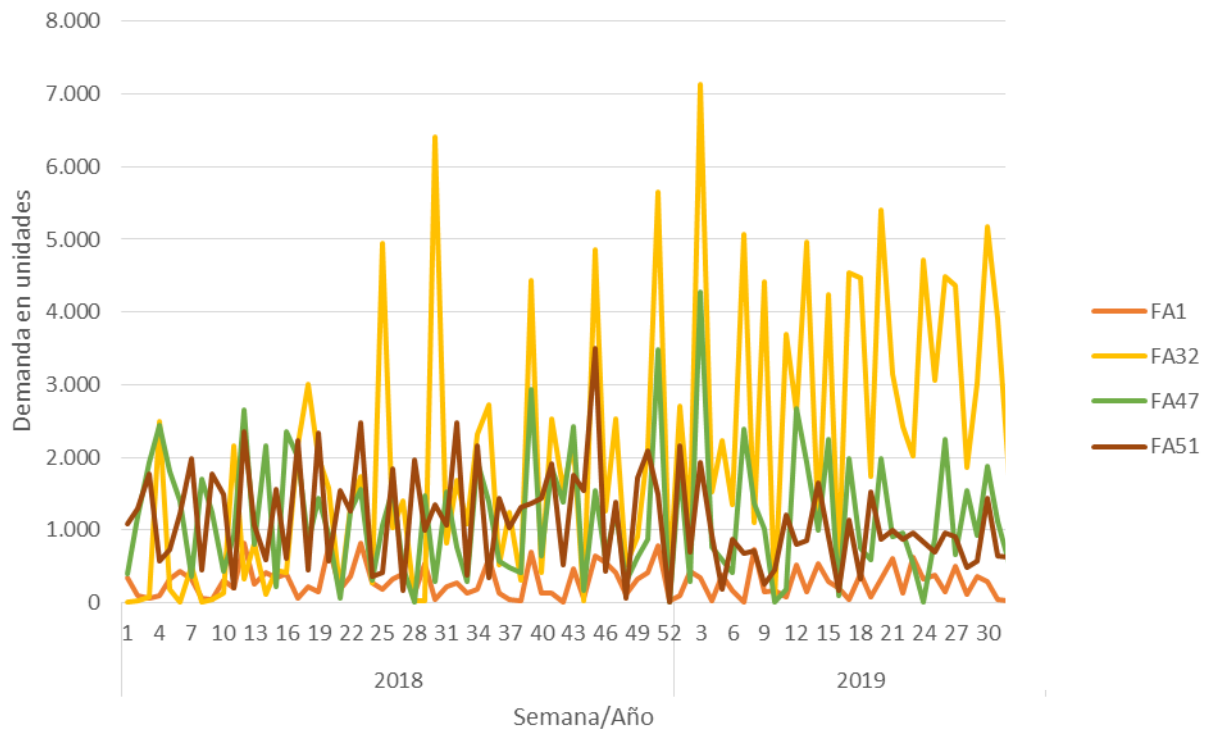


Ilustración 13: Demanda semanal para 4 productos de la línea FARMA, unidad DOLOR

Aunque todos estos medicamentos tratan el dolor, en particular de la artrosis o artritis, tienen un comportamiento diverso. En primer lugar, el producto FA32 muestra una tendencia al alza en las últimas semanas mientras que los demás no. Esta alza se puede deber a que es el medicamento que a más tratamientos apunta, sirve para múltiples formas de artrosis y artritis y para el tratamiento a corto plazo del dolor asociado a cirugía dental. El resto de los productos sólo se indican como medicamentos suplementarios en el tratamiento de la artrosis o artritis. Por otro lado, se espera que el comportamiento del producto F1 sea radicalmente diferente a los demás ya que la presentación del producto es diferente. Mientras los medicamentos FA32, FA47, y FA51 vienen en *blisters*¹ con cápsulas, el medicamento F1 se administra mediante una inyección entonces incluye una jeringa.

La forma en que el consumidor final compra el medicamento en farmacias o cómo es recetado en instituciones privadas de la salud determina la manera en que estas mismas instituciones hacen sus pedidos a la empresa farmacéutica. Es por eso que ocupar un

¹ Blisters: Envase para manufacturados pequeños que consiste en un soporte de cartón o cartulina sobre el que va pegada una lámina de plástico transparente con cavidades en las que se alojan los distintos artículos. (Real Academia Española, 2019)

mismo modelo de estimación de demanda para los productos en conjunto es un error y pasa a llevar características esenciales de cada producto.

Se podría hacer el ejercicio de hacer *clusters* con el método *k-means*, o ver correlaciones entre las demandas históricas de los productos para así poder hacer grupos. Así poder elegir un producto representativo y hacer el pronóstico de esos medicamentos y luego extrapolarlos a los demás. Pero hacer eso podría hacer agrupar productos que por coincidencia son parecidos y no por razones del negocio o de sus características. Los modelos estimados necesitan ser lo más certeros posibles en cada medicamento y agrupar equivocadamente es peligroso para la posterior estimación de los modelos.

Es por esto que se decide agrupar los medicamentos por familia, es decir, por el principio activo que tienen. Como se detallaba anteriormente, los 118 productos en estudio se agrupan en 58 familias/principios activos. Existe un principio activo presente en un producto de FARMA y otro de OLE que se contabiliza por separado. En el Anexo 5, se detallan cuántos medicamentos están en cada una de las familias.

Los medicamentos dentro de cada familia son químicamente iguales y sólo se diferencian en la cantidad de cápsulas o gramos de medicamento que vienen en su presentación. Esto quiere decir que al agrupar por familia también se está agrupando por presentación: medicamentos vendidos en cápsulas o en formato de jeringa. A continuación se puede ver una ejemplificación de esta situación en el principio activo de Drospirenona, familia que agrupa los anticonceptivos vendidos por la empresa.

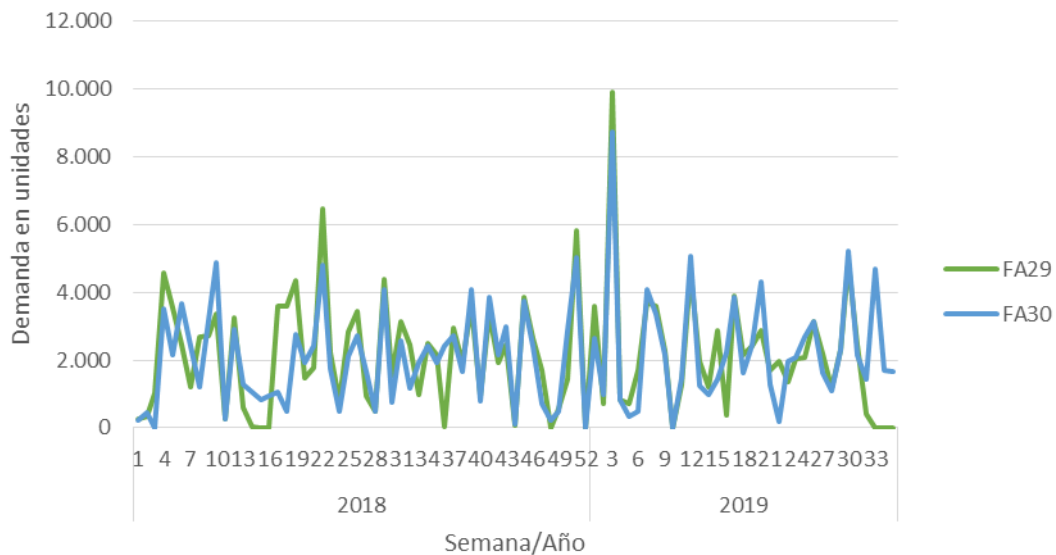


Ilustración 14: Demanda semanal de medicamentos de principio activo Drospirenona

Para hacer la estimación de demanda se selecciona 1 medicamento por familia. Para seleccionar este medicamento representativo dentro de cada grupo, se elige el que más ingresos promedio reporta en el año 2018 y 2019. Se hace la elección por ingresos y no por demanda ya que, aunque todos los productos tienen la misma ponderación dentro del cálculo del *forecast accuracy*, un medicamento que reporta mayor utilidad es considerado más importante. Estos son medicamentos que en el día a día se les presta más atención.

Esta selección se hace para todas las familias, excepto para la familia de *Diclofenac*. La acción terapéutica de este principio activo se enfoca en aliviar el dolor leve a moderado y agrupa 13 medicamentos. Los productos dentro de esta familia son los únicos que varían en presentación y dosis entre sí, por lo que dentro de este se pueden diferenciar en 6 marcas diferentes.

Finalmente se tienen 64 productos de cada una de las familias de principios activos de la empresa. Con estos medicamentos se está contabilizando el 70% de las utilidades al año de la empresa, y tienen un promedio de cumplimiento del 62% el año 2019. Representan aproximadamente el 82% de la demanda total mensual.

Además de ver comportamientos por principio activo, también interesa indagar en los clientes que compran cada medicamento. Los comportamientos del canal privado y pseudo-privado pueden variar porque compran estos medicamentos para diferentes razones. Una farmacia perteneciente al canal pseudo-privado debe siempre mantener en stock disponible de productos para sus pacientes, mientras que una clínica o consultorio podría seguir esa misma modalidad o podría pedir medicamentos sólo cuando los requiere o cuando un paciente necesita cierto medicamento. Pueden existir ciertos comportamientos propios de cada uno de los canales que se pierden si se estima la demanda total de los productos y no por canal. Esto se puede ejemplificar con la demanda del medicamento por canal a continuación.

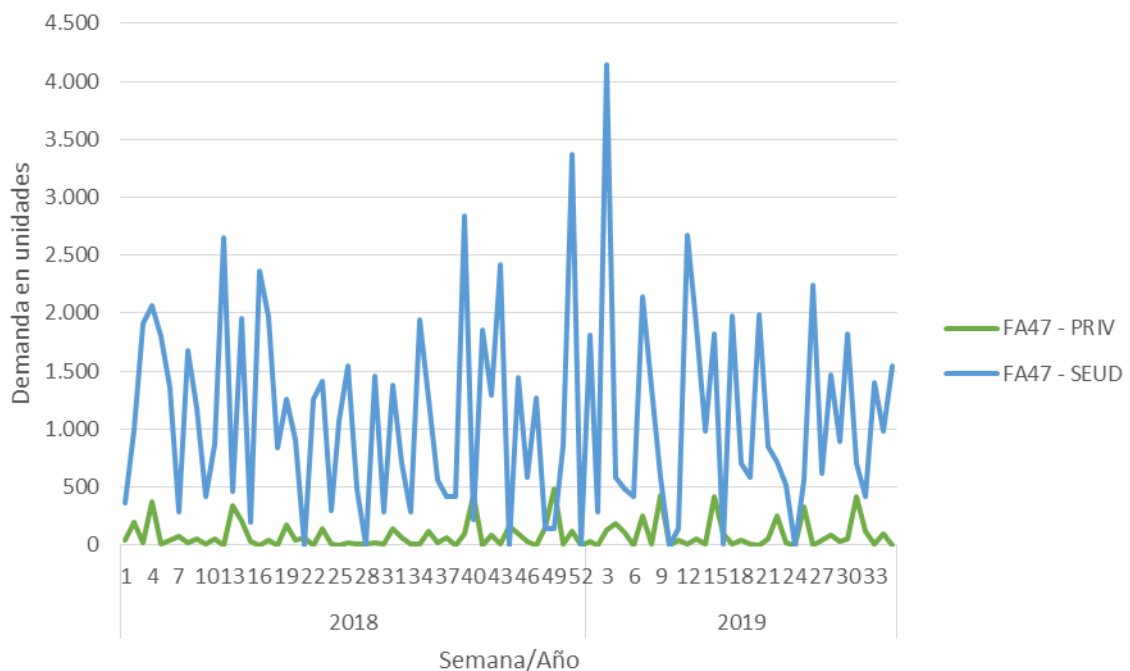


Ilustración 15: Demanda por canal de venta del medicamento FA47

c.2 Características de series de tiempo a estimar

Con la separación de la demanda en principios activos y canales de venta, sigue a ver las características de cada serie de tiempo. En este punto interesa ver qué series tienen estacionalidades y cuáles son estacionarias o ruido blanco (*white noise*). En total se tienen 117 series de tiempo y para cada una de ella se realiza la prueba de *Ljung-Box* para ver si las series son significativamente distintas a *white noise*. Para ver los ciclos

de esta estacionalidad se ve el gráfico de autocorrelación, donde se ve el retraso significativo que será el largo de la estacionalidad detectada.

La siguiente tabla detalla las series que presentan estacionalidad.

Producto	Unidad	Familia	Canal	Ciclo en semanas	Ciclo en meses
FA27	FARMA	CARDIO MET	Pseudo	4	1
FA56	FARMA	CARDIO MET	Pseudo	4	1
FA69	FARMA	CARDIO MET	Pseudo	3	1
FA70	FARMA	CARDIO MET	Pseudo	14	4
FA72	FARMA	CARDIO MET	Privado	12	3
FA72	FARMA	CARDIO MET	Privado	10	3
FA33	FARMA	DOLOR	Pseudo	4	1
FA34	FARMA	DOLOR	Privado	4	1
FA34	FARMA	DOLOR	Pseudo	16	4
FA51	FARMA	DOLOR	Privado	4	1
FA51	FARMA	DOLOR	Pseudo	2	1
FA52	FARMA	DOLOR	Privado	4	1
FA78	FARMA	DOLOR	Pseudo	9	2
FA12	FARMA	NEUROCIENCIAS	Privado	31	8
FA41	FARMA	NEUROCIENCIAS	Privado	2	1
FA60	FARMA	NEUROCIENCIAS	Privado	8	2
FA62	FARMA	NEUROCIENCIAS	Pseudo	16	4
FA66	FARMA	NEUROCIENCIAS	Pseudo	9	2
FA67	FARMA	NEUROCIENCIAS	Pseudo	12	3
FA84	FARMA	NEUROCIENCIAS	Pseudo	5	1
FA86	FARMA	NEUROCIENCIAS	Privado	4	1
FA86	FARMA	NEUROCIENCIAS	Pseudo	11	3
FA92	FARMA	NEUROCIENCIAS	Pseudo	8	2
F15	FARMA	PRIMARY CARE	Privado	25	6
FA16	FARMA	PRIMARY CARE	Pseudo	19	5
FA17	FARMA	PRIMARY CARE	Pseudo	3	1
FA21	FARMA	PRIMARY CARE	Privado	13	3
FA23	FARMA	PRIMARY CARE	Privado	36	9
FA53	FARMA	PRIMARY CARE	Privado	9	2
FA82	FARMA	PRIMARY CARE	Privado	8	2
FA82	FARMA	PRIMARY CARE	Pseudo	5	1

FA10	FARMA	UGI	Pseudo	10	3
FA48	FARMA	UGI	Privado	6	2
FA74	FARMA	UGI	Privado	26	7
FA87	FARMA	UGI	Pseudo	5	1
OL100	OLE	ONCO HEMATO	Pseudo	4	1
OL101	OLE	ONCO HEMATO	Pseudo	4	1
OL107	OLE	ONCO HEMATO	Pseudo	25	6
OL95	OLE	URO-ONCO	Pseudo	18	5
OL103	OLE	URO-ONCO	Privado	16	4
OL117	OLE	URO-ONCO	Pseudo	7	2

Ilustración 16: Detalle de estacionalidad de las series de tiempo a estudiar

De los medicamentos que presentan estacionalidad, el 41% pertenecen a series de tiempo de demanda del canal privado y 59% restante al canal pseudo-privado.

Los medicamentos que presentan una estacionalidad de mayor a 2 meses son de uso crónico, con indicaciones terapéuticas que explican su estacionalidad. Mientras que la mayoría de los medicamentos con estacionalidad de 1 mes son aquellos que son indicados para tratamientos de corto plazo o enfermedades que necesitan ser tratadas constantemente. Aquí se concentran antiinflamatorios y analgésicos recetados para dolores moderados a severos, antihipertensivos, antidepresivos y antiepilépticos. Instituciones de canales privados o pseudo-privados, dependiendo del producto, deben estar constantemente reabasteciendo su stock de estos productos porque son constantemente demandados por personas nuevas a los tratamientos. Esto explica su estacionalidad.

Existen también series de productos-canal que tienen una clara tendencia al alza como lo son: FA32-pseudo privado, FA52-pseudo privado y FA71-pseudo privado. La tendencia sólo se confirma visualmente. Se comprueba si realmente esta tendencia es real si los modelos posteriormente estimados contienen esta componente de tendencia.

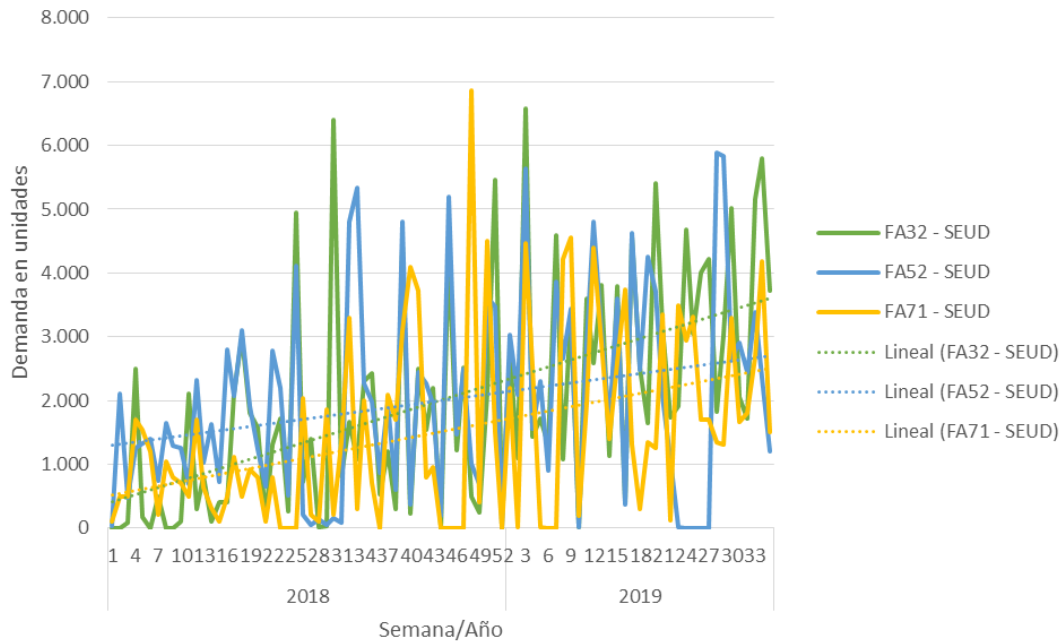


Ilustración 17: Series de tiempo con tendencia al alza

Con las estacionalidades y tendencias especificadas, se ha llegado a una caracterización de la demanda de los productos. Más allá de que esta caracterización de demanda por principio activo y canal sirva para posteriores modelos de estimación de demanda, es útil porque es más información que se puede usar para entender el contexto de la empresa.

d. Elección de modelos a utilizar y su estimación

Para cada una de las series de tiempo se sigue el siguiente método: en primer lugar se parte por probar modelos no estacionarios y sin tendencia (promedio móvil simple y suavización exponencial simple), luego se sigue por modelos no estacionarios con tendencia (promedio móvil doble y suavizado exponencial doble) y por último los modelos auto-regresivos integrados de media móvil (ARIMA). Para las series estacionales se agregan los modelos estacionales (aditivo de Holt-Winters y multiplicativo de Holt-Winters). Para realizar las estimaciones se ocupa el paquete estadístico *fpp2* y *forecast* en el software R, que incluye todos los modelos explicitados.

Como se menciona en apartados anteriores, se determina un grupo de datos de entrenamiento y validación de manera que se tiene lo siguiente:



Ilustración 18: Diagrama de grupos de entrenamiento y validación

Entre cada una de las estimaciones de los modelos se analizan sus métricas de ajuste y pronóstico así como sus residuos. Las métricas de ajuste se analizan con el set de datos de entrenamiento y las de pronóstico con el set de datos de validación.

Se escoge un producto para ilustrar el desarrollo de la aplicación de los modelos a estimar, considerando como set de datos de entrenamiento y validación los de la Ilustración 18.

d.1 Estimación de demanda del medicamento OL116 – canal pseudo-privado

El producto OL116 pertenece a la línea de OLE, unidad Uro-Onco y a la familia de Triptorelina. Este medicamento se indica a pacientes en tratamiento de cáncer de próstata avanzado, y su dosis recomendada es una inyección intramuscular cada tres meses. Esto significa que es un medicamento de uso crónico. Para este ejemplo se analiza la demanda en el canal pseudo-privado. A continuación se muestra esta serie de tiempo y su gráfico de autocorrelación.

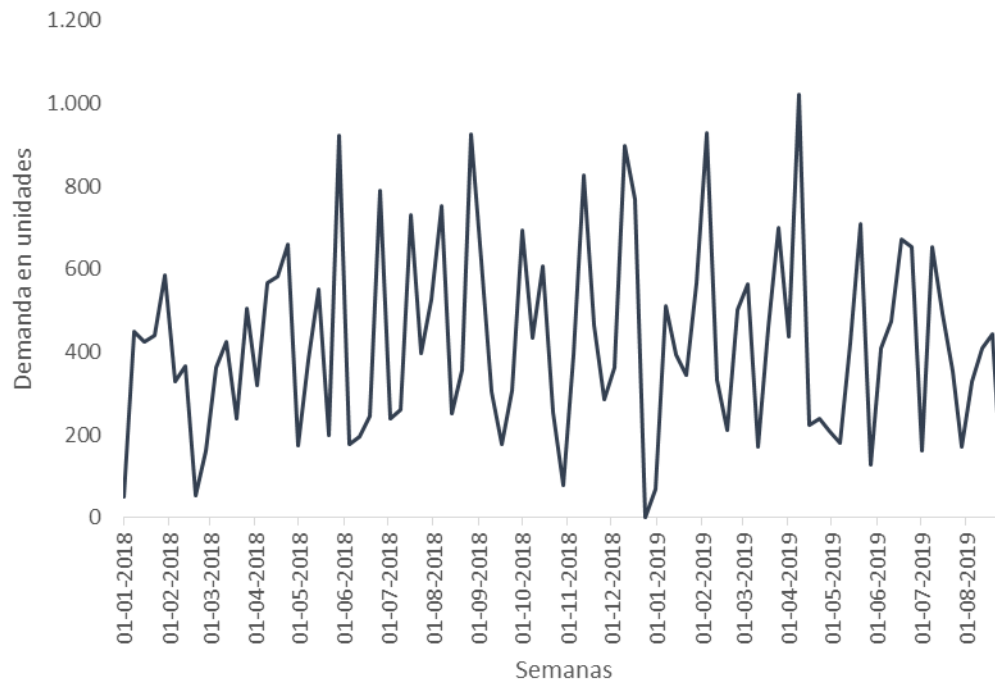


Ilustración 19: Demanda de medicamento OL116 por canal pseudo-privado

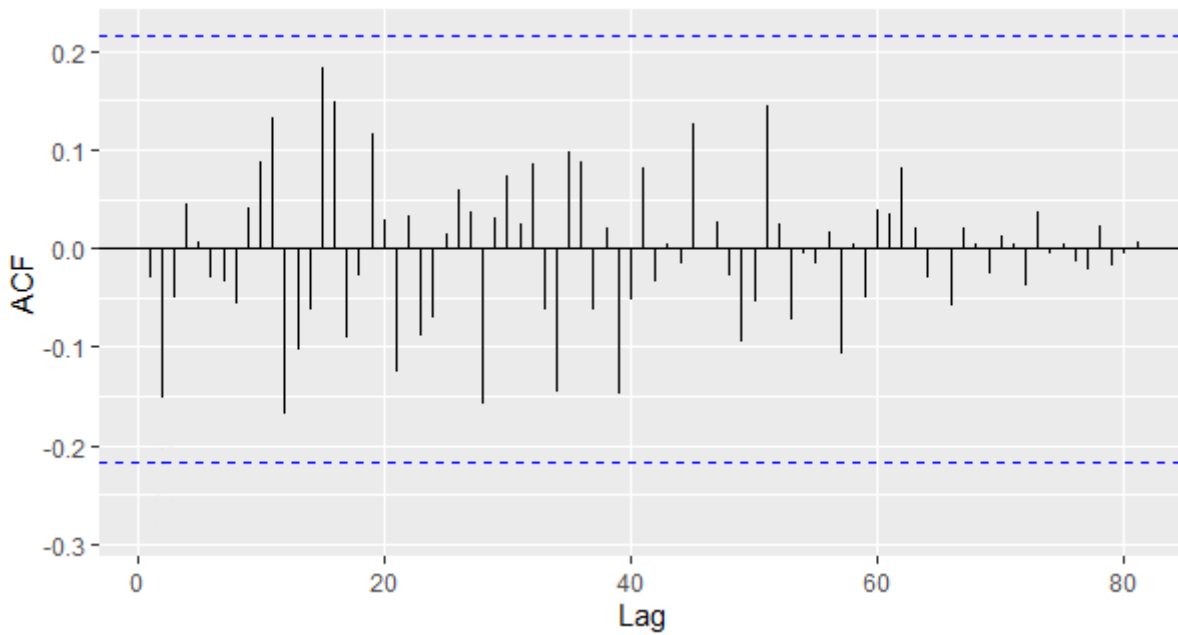


Ilustración 20: Gráfico de autocorrelación para la serie OL116 - canal pseudo privado

Se observa que la media y varianza son relativamente constantes a través del tiempo, lo que da la intuición de que la serie puede ser estacionaria (*white noise*). Para comprobar esto, se realiza la prueba de Ljung-Box.

Se debe recordar que la hipótesis nula de Ljung-Box es que los datos tienen autocorrelación igual a 0. El p -valor de esta prueba da un 0,77. Al ser mayor que 0,05 se puede afirmar que la serie no es significativamente diferente a una serie de ruido blanco. El gráfico de autocorrelación tampoco muestra ningún retraso significativo para esta serie.

Se sigue a estimar la demanda futura. En primer lugar, se aplica el modelo de estimación de demanda de promedio móvil simple. Se recalca que este modelo es el que más se asimila a las estimaciones que se hacen actualmente en el proceso de estimación. La estimación se puede ver en el siguiente gráfico:

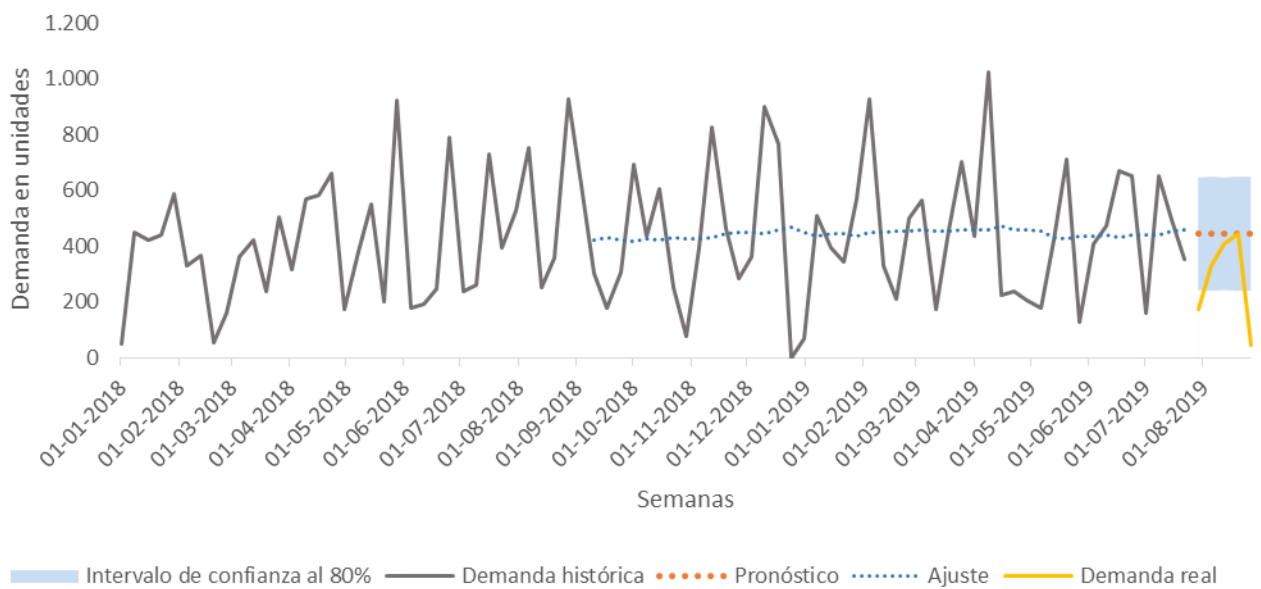


Ilustración 21: Estimación de promedio móvil simple de la serie OL116 - canal pseudo-privado

A continuación se muestran las métricas de ajuste y pronóstico de esta estimación.

Modelo	Métricas de ajuste			Métricas de pronóstico	
	MAD	MAPE	RMSE	MAD	<i>Accuracy</i>
Promedio móvil simple	195	42,61%	241	164	37%

Tabla 6: Métricas y coeficientes de estimación de promedio móvil simple de la serie OL116 - canal pseudo-privado

Las métricas de este pronóstico no explican bien el comportamiento del producto ni tampoco logran predecir el futuro. Un *accuracy* de sólo 37% está por sobre lo que la empresa determina como aceptable. El pronóstico sobreestima la demanda, lo cual es similar a la realidad ya que todo el año 2017 este producto no ha cumplido con su pronóstico por sobreestimarse la cantidad que se vende. Esto último se puede ver en la siguiente tabla.

OL116	enero 2019	febrero 2019	marzo 2019	abril 2019	mayo 2019	junio 2019
<i>Forecast accuracy</i>	45,03%	40,59%	52,38%	52,16%	39,98%	39,28%

Tabla 7: Forecast accuracy del producto OL116

Además de observar las métricas de ajuste y pronóstico de cada modelo, se debe verificar que la distribución de los errores de ajuste, llamados residuales, se asimilen a una serie de tiempo de ruido blanco. Esto es porque si llegan a presentar alguna tendencia o estacionalidad, significa que el modelo de estimación de demanda no captura toda la información de los datos históricos. A continuación se presenta el gráfico de los residuales de este modelo:

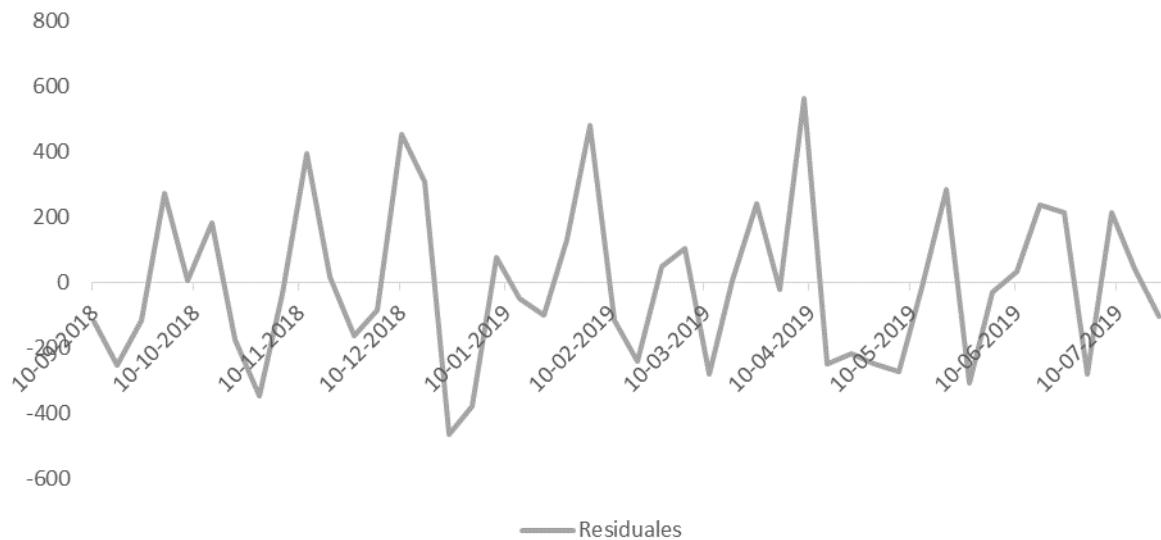


Ilustración 22: Residuales de la estimación por promedio móvil simple de medicamento OL116 - pseudo privado

Si se hace el test de Ljung-Box para los residuales se tiene un p -valor de 0,273. Esta información muestra que los residuales no son significativamente distintos a ruido blanco.

Se sigue a aplicar el modelo de suavización exponencial simple a la misma serie de tiempo.

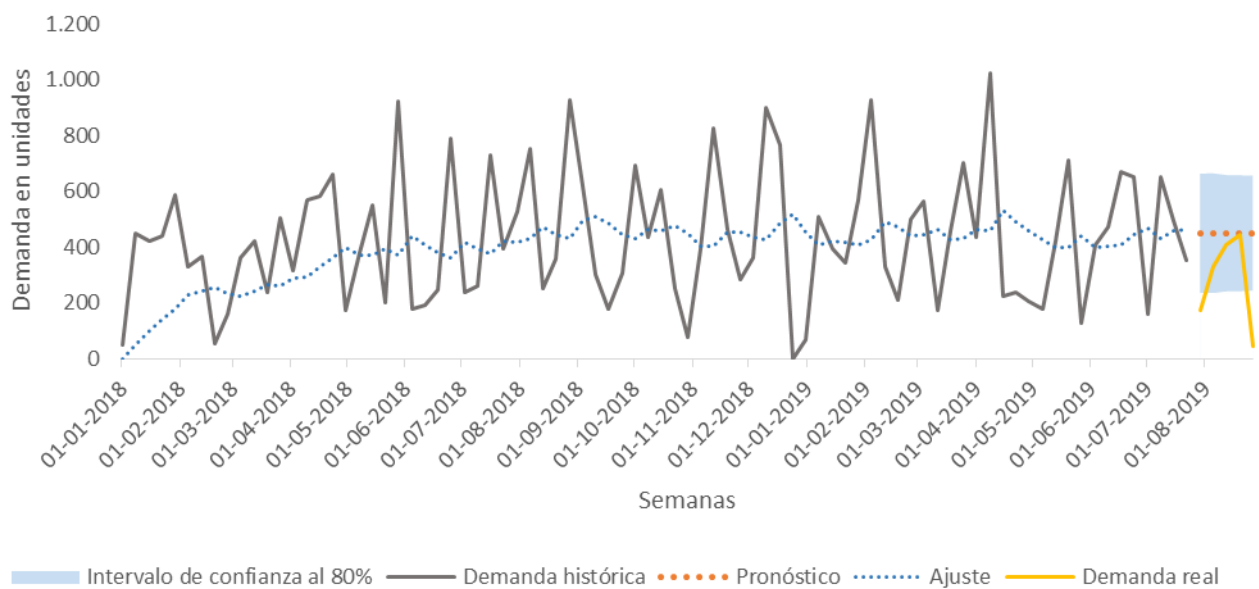


Ilustración 23: Estimación de suavización exponencial simple de la serie OL116 - canal pseudo-privado.

Modelo	Coefficientes	Métricas de ajuste			Métricas de pronóstico		Residuales
	α	MAD	MAPE	RMSE	MAD	Accuracy	<i>p</i> -valor <i>Ljung-Box</i>
Suavización exponencial simple	0,13	213	50,06%	255	169	38%	0,3833

Tabla 8: Métricas y coeficientes de estimación de suavización exponencial simple de la serie OL116 - canal pseudo-privado

El modelo de suavización exponencial simple logra ajustarse mejor a la demanda histórica, otorgando un coeficiente α del 0,13. Esto quiere decir que el 13% del valor predicho está dado por la observación más reciente, mientras que el 87% se basa en los valores pasados. En términos de certeza, este modelo no mejora con respecto al promedio móvil simple. Verificando los residuales, se puede decir que no son significativamente distintos a *white noise*.

Sigue estimar el modelo de promedio móvil doble, que incorpora la posible tendencia que puede tener el producto en su demanda.

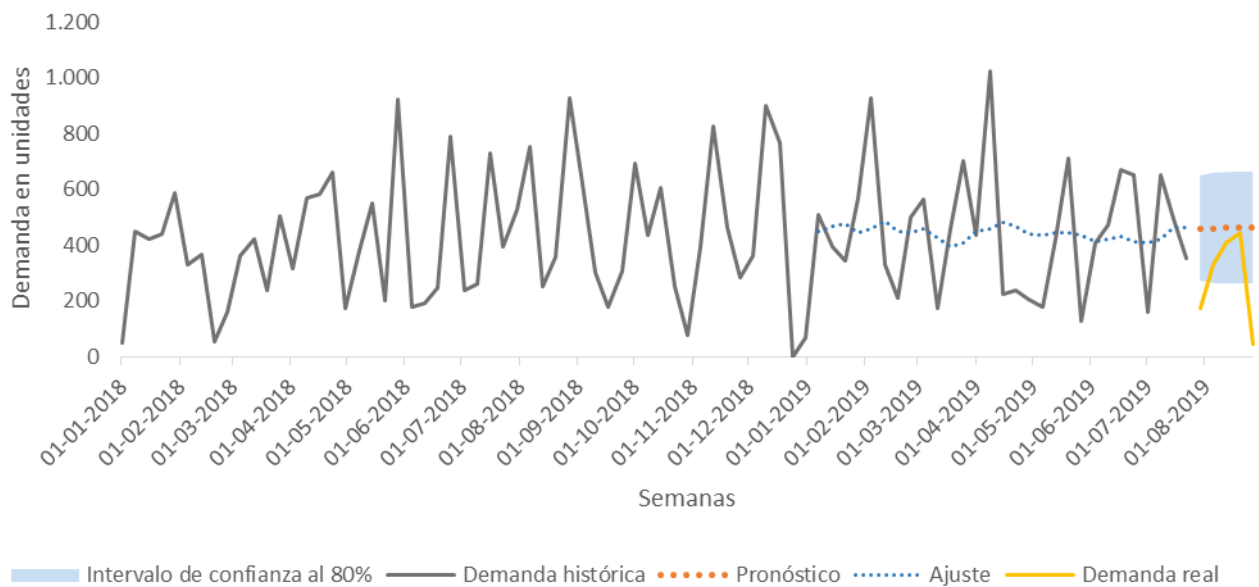


Ilustración 24: Estimación de promedio móvil doble de la serie OL116 - canal pseudo-privado

Modelo	Métricas de ajuste			Métricas de pronóstico		Residuales
	MAD	MAPE	RMSE	MAD	Accuracy	<i>p</i> -valor <i>Ljung-Box</i>
Promedio móvil doble	184	41,03%	226	181	39%	0,9752

Tabla 9: Métricas y coeficientes de promedio móvil doble de la serie OL116 - canal pseudo-privado

Aunque las métricas de ajuste mejoran en comparación a los 2 modelos anteriores, el pronóstico no logra seguir esta misma tendencia, y empeora la métrica de MAD.

Sigue el modelo de suavización exponencial doble.

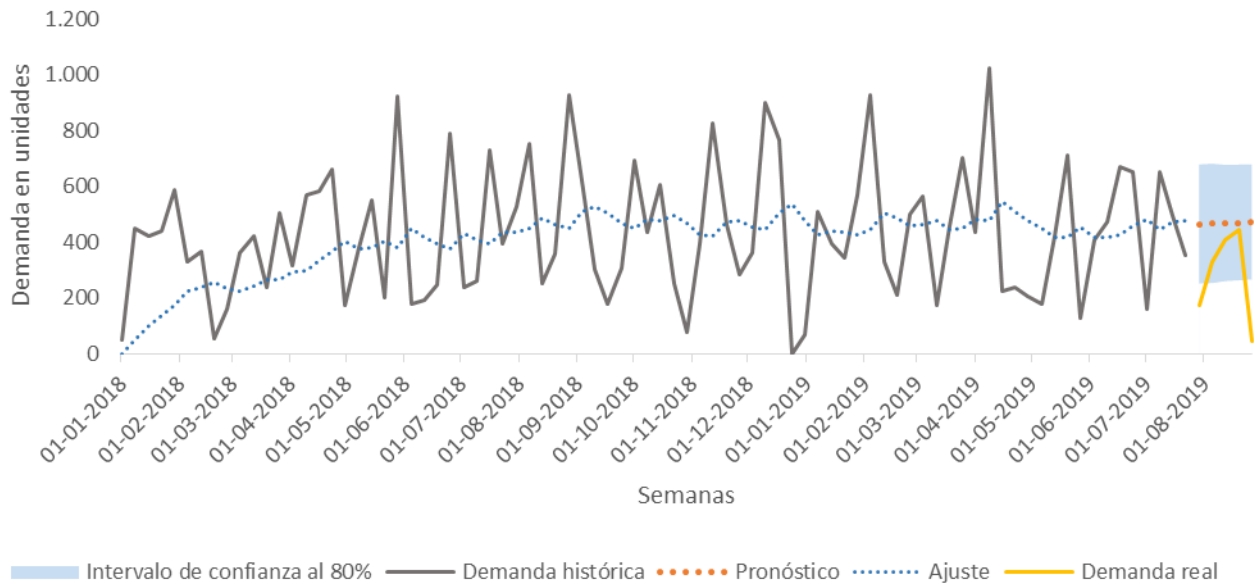


Ilustración 25: Estimación de suavización exponencial doble de la serie OL116 - canal pseudo-privado

Se muestran los coeficientes y métricas de ajuste y pronóstico.

Modelo	Coeficientes		Métricas de ajuste			Métricas de pronóstico		Residuales
	α	β	MAD	MAPE	RMSE	MAD	Accuracy	<i>p-valor</i> <i>Ljung-Box</i>
Suavización exponencial doble	0,12	0,007	213	49,37%	255	188	40%	0,4039

Tabla 10: Métricas y coeficientes de suavización exponencial doble de la serie OL116 - canal pseudo-privado

Al igual que el modelo anterior, el modelo de suavización exponencial doble hace ver una leve tendencia al alza de este producto que en la realidad no es verdadera. Ocupar este modelo podría sesgar los pronósticos al pensar que este producto está siendo cada vez más vendido al canal pseudo-privado cuando esto no es verdad.

Por último, sigue el modelo ARIMA. Este modelo es estimado en el software estadístico *R*, el cual tiene una función que selecciona los mejores parámetros de múltiples modelos autorregresivos de media móvil.

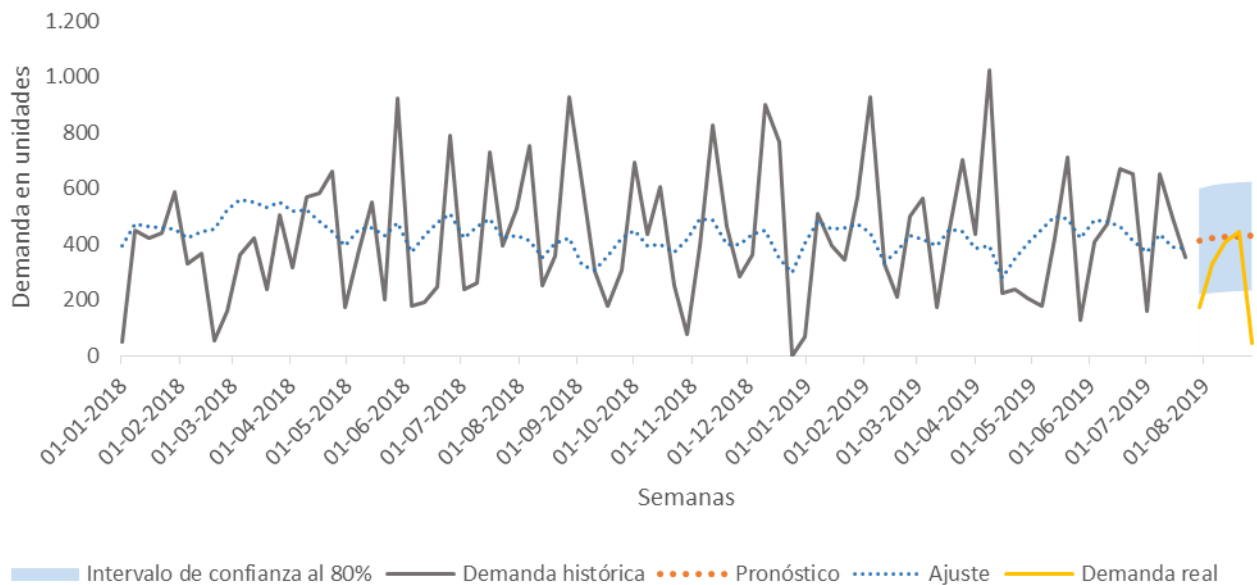


Ilustración 26: Estimación de modelo ARIMA(1,0,2) de la serie OL116 - canal pseudo-privado

Se muestran los coeficientes y métricas de ajuste y pronóstico.

Modelo	Métricas de ajuste			Métricas de pronóstico					Res.
	MAD	MAPE	RMSE	MAD	<i>Accuracy</i>	BIC	AIC	AICc	p-valor Ljung-Box
ARIMA (1,0,2)	174	40,04%	220	146	33%	11	10,89	10,89	0,773

Tabla 11: Métricas y coeficientes de ARIMA(1,0,1) de la serie OL116 - canal pseudo-privado

Este modelo muestra que el orden de la parte autorregresiva de la ecuación es 1, que no existe diferenciación (ya que la serie de tiempo ya es estacionaria, por lo tanto no necesita diferenciación) y el orden del promedio móvil es 1.

Este procedimiento es realizado para cada una de las series de tiempo representativas de su familia de principios activos. Para las series que presentan estacionalidad, se agregan aquellos modelos que la incluyen.

d.2 Resumen de modelos para todos los productos por canal de venta

Es importante recalcar que, aunque hasta ahora se ha estado estimando la demanda semanal, el objetivo principal es estimar la demanda mensual de cada SKU por canal. Una métrica relevante para evaluar es cómo se traducen las estimaciones de los modelos semanales a certeza mensual. Para esto solo basta agregar las demandas semanales en mensuales y hacer el cálculo de *accuracy* para las demandas y pronósticos mensuales.

Por razones de espacio, no se detallan todos los modelos estimados para cada serie. Para un detalle de los 2 modelos con menor *accuracy* mensual y menor RMSE de ajuste de cada serie se puede ver el Anexo 6. A continuación a modo de ilustrar los modelos estimados se muestran algunas series de SKU-canal con sus 2 modelos.

Producto	Unidad	Canal	Modelos	RMSE	Accuracy mensual
FA65	CARDIO MET	Privado	ARIMA(1,0,1)	42	47%
FA65	CARDIO MET	Privado	Promedio móvil doble	41	48%
FA65	CARDIO MET	Pseudo	Suavizado exponencial doble	431	0%
FA65	CARDIO MET	Pseudo	Promedio móvil doble	446	3%
FA1	DOLOR	Privado	ARIMA(1,0,0)	13	27%
FA1	DOLOR	Privado	Suavizado exponencial doble	12	75%
FA1	DOLOR	Pseudo	ARIMA(0,0,1)	210	11%
FA1	DOLOR	Pseudo	Suavizado exponencial doble	216	18%
OL116	URO-ONCO	Privado	Promedio móvil simple	2	16%
OL116	URO-ONCO	Privado	Suavizado exponencial simple	2	19%
OL116	URO-ONCO	Pseudo	ARIMA(1,0,1)	225	33%
OL116	URO-ONCO	Pseudo	Promedio móvil simple	241	37%
FA80	NEUROCIENCIAS	Privado	Promedio móvil simple	6	-27%
FA80	NEUROCIENCIAS	Privado	Promedio móvil doble	7	-38%
FA80	NEUROCIENCIAS	Pseudo	ARIMA(2,0,2)	183	24%
FA80	NEUROCIENCIAS	Pseudo	Suavizado exponencial doble	202	24%

Tabla 12: Modelos estimados para una selección de SKU-canal

e. Evaluación de modelos

Para evaluar los modelos se ven principalmente 2 métricas: RMSE de ajuste y el *accuracy* mensual (desvío entre el pronóstico y la demanda real). En primer lugar, al comparar los métodos por ajuste, se elige el RMSE de ajuste ya que es el criterio más importante de ajuste si el objetivo final es la predicción de valores. Ocupar otro tipo de error, por ejemplo porcentual no es muy útil en este caso porque se indetermina cuando los valores son 0, algo que pasa en algunas semanas que no hay demanda. El *accuracy* es una medida para luego poder obtener el *forecast accuracy* de los modelos estimados.

Pero se prioriza la segunda métrica mencionada: el *accuracy* mensual, medida de la cual luego se puede obtener el *forecast accuracy* de los modelos estimados. Así, el mejor modelo para cada serie de demanda SKU-canal es aquel que tiene menor porcentaje de

accuracy (es decir, el desvío entre lo predicho y lo real es menor), y posteriormente menor RMSE de ajuste.

Además, para elegir el mejor modelo se priorizan aquellos modelos que sobreestiman la demanda antes que subestimarla. Esto es porque en el rubro de los medicamentos es preferible vender un producto a un precio más bajo o perder una venta, que arriesgar que el cliente se vaya a la competencia o peor, comprometer la salud de un paciente al no tener disponible un medicamento.

Con esta información en mente, se seleccionan los mejores modelos para cada uno de los productos y canal de venta. En el Anexo 7 se detalla cada uno de estos modelos. A continuación se muestran los modelos seleccionados y la cantidad de series que se tienen con ese mejor modelo, por unidad de negocio y por canal.

Mejores modelos	FARMA					OLE		
	Cardio Met	Dolor	Neurociencias	Primary Care	UGI	Onco Hemato	Onco- logía	Uro Onco
Promedio móvil simple	6	3	5	2	3	1	0	4
Promedio móvil doble	1	1	7	3	3	0	4	0
Suavizado exponencial simple	2	1	5	5	3	2	0	2
Suavizado exponencial doble	3	3	4	3	5	0	0	0
Aditivo de Holt-Winters	1	0	0	1	0	0	0	0
ARIMA(1,0,1)	2	0	0	0	0	0	0	1
ARIMA(0,0,1)	1	1	2	4	4	3	0	2
ARIMA(1,0,0)	0	1	2	0	3	0	2	0
ARIMA(2,0,2)	0	0	1	0	1	0	1	0

ARIMA(0,1,1)	0	0	0	2	0	1	0	0
--------------	---	---	---	---	---	---	---	---

Ilustración 27: Mejores modelos agrupados por unidad de negocio

Modelos	Privado	Pseudo
Promedio móvil simple	15	9
Promedio móvil doble	8	11
ARIMA(1,0,1)	1	2
Suavizado exponencial simple	9	11
Aditivo de Holt-Winters	1	1
Suavizado exponencial doble	9	9
ARIMA(0,0,1)	3	14
ARIMA(1,0,0)	5	3
ARIMA(2,0,2)	0	3
ARIMA(0,1,1)	2	1

Ilustración 28: Mejores modelos agrupados por canal de venta

De las 117 series de tiempo estimadas con diversos modelos estimados en cada una, estas resultan en sólo 10 modelos considerados como los mejores modelos, cada uno calibrado con la serie de datos correspondiente.

En primer lugar, se destaca que en ninguna de las series se tiene como igual mejor modelo para el canal privado y pseudo privado. Esto hace reafirmar los hechos de que se comportan distinto, que cada canal tiene diferentes necesidades para abastecerse de productos. Por lo que estudiar estos canales por separado es efectivo.

Para algunas series el mejor modelo de estimación de demanda sigue siendo el que se usa hoy en la compañía: el promedio móvil simple. Esto se debe a que ciertas series de tiempo son erráticas, haciendo difícil la tarea de encontrar algún modelo que se ajuste a los datos, para posteriormente pronosticar.

Los modelos que consideran estacionalidad sólo resultan ser el mejor para 2 de las 42 series que presentaban estacionalidad. Estas series son la del producto FA56 – canal pseudo-privado y FA23 – canal privado. El primero es medicamento para el tratamiento de la obesidad, que se receta en pacientes adultos con un índice de masa corporal inicial

alto para el manejo crónico del peso. Tiene sentido que este medicamento tenga estacionalidad ya que es de uso crónico en pacientes que cumplirán un ciclo de uso y lo dejarán de usar. Así, nuevos pacientes entrarán a un nuevo ciclo de la estacionalidad. El segundo es un antiinflamatorio para el tratamiento de inflamaciones dolorosas osteoarticulares, que es de uso crónico para pacientes con artritis.

Como se explicita en el Objetivo General, el nivel de agregación de la estimación que se busca en este trabajo es de SKU-canal de venta. Lo que hace natural desagregar la demanda de los 64 medicamentos representativos en demanda privada y pseudo-privada, y estimar los modelos como ya se realizó. Aun así, hay que recordar que el desempeño de los pronósticos se mide a nivel de SKU y no SKU-canal.

Los medicamentos estimados para un canal, no están comprometidos necesariamente para ese canal. Ejemplificando en la práctica, si la demanda del canal pseudo-privado se subestima por 2 unidades y la del canal privado se sobreestima por esas 2 unidades, las unidades proyectadas para el canal privado pueden destinarse al pseudo-privado. El *accuracy* del pronóstico SKU-canal podrá estar errado, pero en conjunto puede ser bueno.

Así, para ver efectivamente el poder de los modelos escogidos frente a la situación actual de la empresa, se calcula el *forecast accuracy* de cada SKU, dado por los modelos estimados considerados como “mejores modelos”, comparado con el promedio de *forecast accuracy* del año 2019. Los resultados de esta comparación de algunos medicamentos se pueden ver en el Anexo 8.

Producto	Unidad	Familia	Forecast accuracy de mejor modelo	Promedio de forecast accuracy en 2019
FA1	DOLOR	ACIDO HIALURONICO	12%	-12%
FA65	CARDIO METABOLICA	NEBIVOLOL	2%	-7%
FA80	NEUROCIENCIAS	PRAMIPEXOL	22%	-11%
OL116	URO-ONCO	TRIPTORELINA	33%	44%

Tabla 13: Comparación de *forecast accuracy* con mejor modelo y su promedio del 2019

Se seleccionan estos 4 productos para ejemplificar lo que ocurre con la comparación de *forecast accuracy*. Los valores de RMSE y *accuracy* mensual de cada producto se puede ver en la Tabla 12 en la sección de Estimación de modelos.

En primer lugar, se tienen productos como el FA1, que es un medicamento para mejorar la función articular aliviando dolor. Este medicamento estaba siendo constantemente subestimado, resultando en ventas pérdidas, pero estaba dentro del rango de error por lo que se consideraba como una buena predicción. Los mejores modelos seleccionados para este producto son ARIMA(1,0,0) para el canal privado y ARIMA(0,0,1) para el canal pseudo-privado, que tienen *accuracy* mensual de 27% y 11% respectivamente. Esto se traduce a un 12% de *forecast accuracy* para el producto. Este modelo más complejo finalmente otorga mejor pronóstico. Además, se pasa de subestimar la demanda a sobrestimarla en la misma cantidad, lo que finalmente se prefiere. El valor predicho está dentro del rango de error aceptable.

El segundo producto FA65 se receta para pacientes mayores de 70 años que presentan insuficiencia cardiaca crónica estable, ligera o moderada. Este medicamento tiene una evolución positiva, pasando de un *forecast accuracy* de -7% a 2% con el mejor modelo estimado. Lo particular de este medicamento es que aunque su *forecast accuracy* es bastante cercano a 0% (lo cual es deseable), esto sólo se debe al buen pronóstico de la serie FA65-canal pseudo privado. Esta última serie es estimada usando el modelo de suavización exponencial doble, mientras que la serie de canal privado es con ARIMA(1,0,1). Lo principal que ocurre aquí es que la cantidad que se demanda en el canal privado es significativamente menor al canal pseudo-privado, por lo que aunque el pronóstico de la serie de canal privado es no es preciso, no afecta tanto en el *forecast accuracy* general. Las unidades destinadas para un canal logran suplir las del otro.

El producto FA80 es un producto que trata los síntomas y signos de la enfermedad de Parkinson y presenta la única situación desfavorable en el desarrollo de este trabajo. Los valores antes del proyecto tienen mejor *forecast accuracy* que las dadas por el mejor modelo estimado. Otros 6 productos (de los 64 estudiados) presentan esta situación. Aunque sea una situación desfavorable sí hay que destacar que cada modelo viene asociado con un intervalo de confianza. Es este intervalo el que permite a los encargados de definir el *forecast* de ver si se quiere tomar un pronóstico optimista o pesimista, dada

las condiciones del mercado en ese momento. Aunque el *forecast accuracy* no sea el deseado, estos modelos aún pueden ser útiles en la práctica.

Por último se encuentra el producto OL116, indicado para el tratamiento del cáncer de próstata avanzado, que presenta un mejor pronóstico con el modelo seleccionado. Aunque sigue estando fuera del rango de error aceptable sí se contabiliza como un avance.

Con la métrica de *forecast accuracy* calculada se puede seguir a ver el porcentaje de cumplimiento que se tiene con la utilización de los mejores modelos para los medicamentos representativos seleccionados.

Cumplimiento de <i>forecast</i>	Mejores modelos	Promedio 2019	Ene-2019	Feb-2019	Mar-2019	Abr-2019	May-2019	Jun-2019	Jul-2019
	78%	62%	62%	60%	60%	65%	63%	66%	61%

Tabla 14: Comparación de cumplimiento de *forecast* real con lo estimado

Utilizar los modelos seleccionados como “mejores modelos” hace que inmediatamente el indicador de cumplimiento de *forecast* alcance el 78%. Se puede decir entonces que se cumple el objetivo principal, y que ocupar modelos estadísticos sirve para tener pronósticos más certeros.

Para analizar si los modelos determinados como “mejores modelos” son consistentemente los mejores, se hace una validación con meses anteriores. Es decir, se hace el mismo análisis de mejores modelos tomando diferentes grupos de entrenamiento y calibración de la siguiente manera:

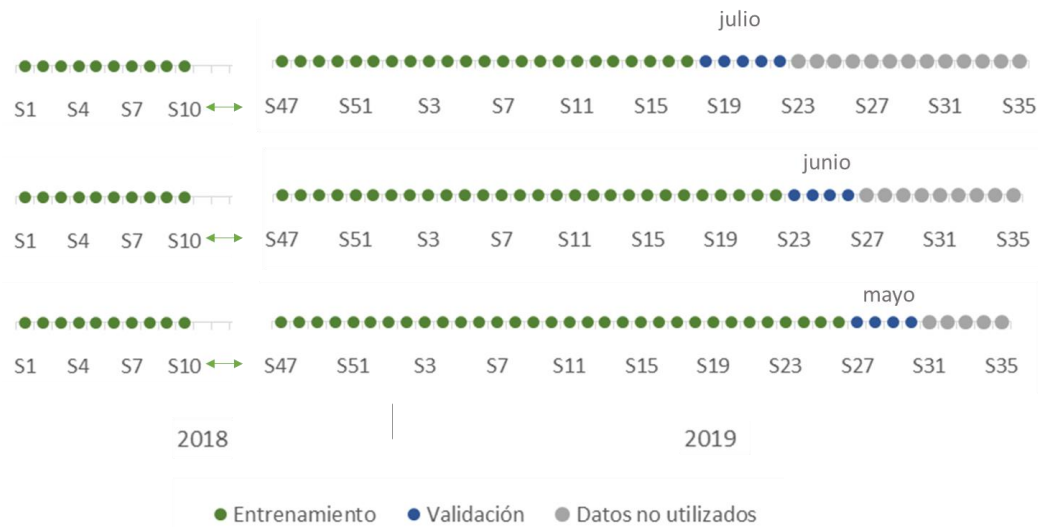


Ilustración 29: Grupos de entrenamiento y calibración para validación de "mejores modelos"

De las 117 series de tiempo estimadas, se tiene que 74 de ellas presenta el mismo modelo como "mejor modelo" en los últimos 3 meses. Para el resto de las 43 series de SKU-canal el "mejor modelo" determinado para las semanas del mes de agosto, aparece como el segundo o tercer modelo con mejor *accuracy* para el resto de los meses. Este resultado da luces de cómo mientras más datos se agregan, más información se tiene para predecir el futuro y cómo se debe iterar en la aplicación de modelos mientras pasan los meses.

f. Propuesta de desarrollo y aplicación

Para implementar este proyecto, es crucial tener una base de datos actualizada todos los meses. En esta se deben seguir los pasos detallados en este trabajo, indagar cada mes en demandas irregulares y *outliers*. Se requiere una revisión constante de los datos, tal que la aproximación de demanda sea lo más cercana a la realidad posible.

Los modelos detallados en este trabajo quedan programados en el software estadístico R para todos los productos y sus respectivos canales. Así, se pueden estimar los modelos seleccionados. Se exportan las predicciones a una planilla de *Excel* para que se integre a los variados análisis que se hacen en toda la empresa y para que se encuentre en una plataforma amigable para todos.

Para aplicar de manera satisfactoria este trabajo, es necesario iterar la metodología. Cada semana que pasa, se tiene un nuevo dato en la serie de tiempo que puede entregar información valiosa para predecir el futuro. El análisis debe ser constante, ya que cada vez se tiene más información y cada vez la empresa añade más medicamentos a su catálogo. Estos modelos deben ir actualizándose a medida que crezca la cantidad de datos históricos, ya que así se puede ver cómo la demanda va cambiando a lo largo de los meses, lo que serviría para hacer modelos cada vez más certeros y reales.

Se recalca que los pronósticos entregados por los modelos estadísticos son un *input* más en el proceso de estimación de demanda, y que por ningún motivo se toma como verdad absoluta. Una gran ventaja de aplicar estos modelos, por más simples que sean, es que otorgan más que un solo número de predicción sino que también entregan los intervalos de confianza que tienen asociados. Estos intervalos de confianza permiten a los tomadores de decisión tener un rango en donde basar sus estimaciones, con la información cualitativa que poseen, pueden ubicarse en diferentes escenarios con la ayuda de estos rangos.

Se espera que a largo plazo, mientras más datos limpios y actualizados se tengan, se pueda prescindir de las reuniones mensuales y que esta sea la única herramienta para predecir la demanda.

VIII. CONCLUSIONES Y RECOMENDACIONES PARA FUTUROS TRABAJOS

a. Conclusiones

Actualmente la predicción de demanda en la empresa farmacéutica recae más en juicios de expertos e intuiciones que en análisis de datos. De esta manera se tiene conocimiento práctico de cómo se comportan las ventas de los productos, pero no se indaga en patrones de comportamiento o en cómo se comportan productos similares. Se está perdiendo información contenida en los datos que no es evidente sin hacer un análisis como el presentado en este trabajo.

Las herramientas desarrolladas a través de este trabajo permiten llegar a un pronóstico de demanda más preciso, de manera que le dan a la empresa otro *input* de información para tomar decisiones más informadas. El trabajo logra entender comportamientos de ciertos datos que con intuiciones e información cualitativa no se pueden saber.

Comenzando con la etapa de limpieza y procesamiento de datos, se muestra lo importante que es trabajar con una base de datos que realmente represente la realidad. Hoy en día en la empresa no existe un protocolo de cómo tratar datos atípicos, y muchas veces estos logran sesgar la predicción actual. En el desarrollo de este trabajo se identificaron dichos valores atípicos con personas de la empresa que estuvieron presentes en dichos períodos, corroborando la existencia de un valor atípico o no.

Se logra hacer una caracterización de los datos, agrupándolos por principio activo y canal de venta. Esta caracterización es importante ya que da nueva información para los tomadores de decisión. Comparar productos dentro de un principio activo podría ayudar también a reducir sesgos, muchas veces pronósticos de productos del mismo principio activo difieren por intuiciones de los expertos, cuando en la práctica debieron ser parecidos.

Al estimar la demanda de los medicamentos usando diferentes tipos de modelo, se ve que hay ciertos medicamentos que toman como “mejor modelo” modelos simples como

promedio simple móvil o suavización exponencial, mientras otras siguen modelos más complejos como ARIMA. Es importante esta distinción porque generalmente las series más erráticas son aquellas que su mejor modelo son modelos simples. Estas probablemente necesitan una mayor limpieza de datos para comenzar a estimar futuros valores, o simplemente tratar de capturar mejor los datos de demanda.

La mayoría de los modelos logran hacer un buen pronóstico de la demanda futura al agregar las demandas de los canales privados y pseudo privados. Esto finalmente es lo que busca la empresa, ya que las unidades no están destinadas exclusivamente para un canal en especial.

Se destaca el hecho de que los modelos de pronóstico tienen asociado un intervalo de confianza, que en este trabajo se considera al 80%. Tener este intervalo es esencial, ya que otorga un rango de valores en los cuales puede estar la demanda, haciendo que en la reunión de *forecast* tengan un rango en el cual moverse dada la información cualitativa que puedan tener.

Se logra verificar que aplicar métodos estadísticos y modelos de pronóstico de demanda mejora de sobremanera el pronóstico actual, donde pasa de un 61% de cumplimiento a un 78%.

Es posible que los “mejores modelos” determinados en este trabajo cambien, a medida que la naturaleza del mercado farmacéutico cambie. Lo importante de este trabajo es el *insight* que en la actualidad la empresa está perdiendo información, al no analizar a fondo sus datos de venta. Ocupando modelos estadísticos que logran captar características que a simple vista no son observables, hace que el proceso de estimación de demanda se acerque cada vez más a la realidad.

b. Limitaciones y trabajo futuros

Aunque los resultados de este trabajo fueron mayoritariamente favorables, no hay que olvidar que se tienen ciertas limitaciones que dan espacios de oportunidad para trabajos futuros.

En primer lugar, está la clara limitación de los datos con los que se desarrolla este trabajo. Las series de tiempo vienen principalmente de ventas y no de demanda real. Aunque se trató lo más posible de aproximar los datos de venta a demanda real ocupando los datos de *backorders*, aún puede existir demanda no capturada. Si un cliente hace un pedido de medicamentos a la empresa y esta última no tiene disponible, el cliente puede tener 2 actitudes: hacer la compra de igual manera como *backorder* y así esperar hasta que el producto esté disponible; o irse a la competencia y/o simplemente no comprar. Los datos de esta última situación no son capturados hoy en día, y son valores que ayudarían a estimar con mayor certeza la demanda futura.

Como solución, se podría digitalizar toda la información de los pedidos que entran a la empresa, con tal de capturar la demanda real aunque no se concrete una venta o *backorder*. Sólo de esta manera se podrán hacer modelos que realmente estén estimando una demanda lo más aproximada a lo real posible.

Existe otra limitación en consideración a los datos ocupados. Para estimar la demanda futura sólo se ocupan datos de demanda histórica, pero claramente la situación futura de la empresa no sólo está determinada por su pasado. Gracias a los resultados obtenidos, tan sólo aplicando modelos simples de estimación, sí se pudo mejorar la situación actual. Pero hay otras variables que inciden en la demanda que no se están contabilizando al ocupar modelos de promedio móvil, suavización exponencial o modelos ARIMA.

Una de estas variables es la participación de mercado que tiene cada uno de los medicamentos dentro de su sub-mercado. Con esta información se podría saber qué rol juega cada producto dentro de su entorno, saber si es un medicamento fácilmente de sustituir o si es único y poder entender realmente a la competencia. La empresa farmacéutica tiene esta información de manera cualitativa, son intuiciones que entran

por medio del juicio de expertos al *forecast* de cada mes. Al no tener esta información de manera detallada se pierde la oportunidad de añadirla cuantitativamente a los modelos y así tener más caracterizada la demanda de cada producto.

Para trabajos futuros sería interesante poder contar con información más detallada sobre el comportamiento de pedido de los clientes. Los clientes del laboratorio son empresas que a su vez deben satisfacer a un cliente final, por lo que el calendario de sus pedidos podría estar escondiendo estacionalidades o tendencias. Los datos de demanda ocupados en este trabajo reflejan cuando el pedido es procesado, es decir, la demanda se registra cuando se le asigna un lote de medicamentos al cliente. Este procesamiento de pedidos no está estandarizado, puede hacerse a los pocos minutos de recibirlo o a la semana. La variabilidad que esta situación presenta hace que, nuevamente, no se esté capturando la demanda en su período real y que pueda estar contaminando los modelos estimados.

Como última variable que incide en la demanda de medicamentos que no es considerada en este trabajo es la fuerza de ventas. La empresa farmacéutica cuenta con representantes para cada línea de negocio, que son los encargados de hacer vínculos comerciales con droguerías e instituciones de salud para que compren los productos. En trabajos futuros puede ser útil considerar cuán grande es el equipo que tiene a su cargo cada representante y cuántas visitas hacen para visibilizar los medicamentos en los distintos establecimientos de salud.

Finalmente, este trabajo se evalúa con el *forecast accuracy* y cumplimiento del *forecast* de cada uno de los productos. Agregar de esta manera la precisión de los pronósticos hace que se oculten comportamientos que sería bueno analizar mes a mes. Por ejemplo, si se desagregara el *forecast accuracy* por producto-canal se podría identificar qué canal es el que presenta mayores o menores desvíos o porcentajes de cumplimiento y la razón. Esto es importante ya que como se veía anteriormente, cada canal pesa distinto en términos de ingresos para la empresa y ayudaría ver, por ejemplo, a qué clientes se les está cumpliendo y a cuáles no.

IX. BIBLIOGRAFÍA

- Aguirre, A. (1994). *Introducción al tratamiento de series temporales*. Madrid: Diaz de Santos.
- Bajaña, W., & Moncayo, C. (octubre de 2012). *La visita médica como alternativa en la promoción de los productos farmacéuticos*. Obtenido de <http://repositorio.ucsg.edu.ec/bitstream/3317/419/3/T-UCSG-POS-MGSS-4.pdf>
- Farma ERP. (2010). *Estudio de Oferta y Demanda del Sector Farmacéutico - Chile*. Obtenido de http://www.farma-erp.cl/biblioteca/estudioOD_sectorfarmaceutico_2011.pdf
- Gallardo, J. (2009). *Metodología para el Desarrollo de Proyectos en Minería de Datos CRISP-DM*. Obtenido de <http://oa.upm.es/1946/>
- Hyndman, R. (marzo de 2014). *Measuring forecast accuracy*. Obtenido de <https://pdfs.semanticscholar.org/af71/3d815a7caba8dff7248ecea05a5956b2a487.pdf>
- Hyndman, R., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. Obtenido de <https://otexts.com/fpp2/>
- Ministerio de Salud. (abril de 2006). *Código Sanitario*. Obtenido de <https://www.leychile.cl/Navegar?idNorma=5595>
- Ministerio de Salud. (16 de Octubre de 2013). *Decreto 466: Reglamento de farmacias, droguerías, almacenes farmacéuticos, botiquines y depósitos autorizados*. Obtenido de http://www.supersalud.gob.cl/observatorio/671/articles-8907_recurso_1.pdf
- Oficina Económica y Comercial de España en Santiago de Chile. (10 de 5 de 2018). *Sector farmacéutico en Chile*. Obtenido de <https://www.icex.es/icex/es/navegacion-principal/todos-nuestros-servicios/informacion-de-mercados/paises/navegacion-principal/el-mercado/estudios-informes/DOC2018789896.html?idPais=CL>
- Pineda, L. (1999). *Métodos y modelos econométricos: una introducción*. México D.F.: Editorial Limusa.

Real Academia Española. (2019). *Diccionario de la lengua española*. Obtenido de <https://dle.rae.es/bl%2525C3%2525ADster>

Schroeder, R. G., Meyer Goldstein, S., & Rungtusanatham, M. J. (2011). *Administración de operaciones: Conceptos y casos contemporáneos*. México D.F.: McGraw Hill.

Vasallo, C. (junio de 2010). *El mercado de medicamentos en Chile: caracterización y recomendaciones para la regulación económica*. Obtenido de <https://www.minsal.cl/wp-content/uploads/2015/09/Informe-Vassallo-2010.pdf>

X. ANEXOS

Anexo 1: Evolución histórica del indicador *forecast accuracy*

Se presenta el detalle de la evolución histórica del indicador *forecast accuracy*, desagregado por valores positivos y negativos.

También se ve la cantidad mensual en que este indicador fue positivo o negativo, es decir, en cuántos productos se sobrestima (*forecast accuracy* positivo) o subestima (*forecast accuracy* negativo) la demanda.

Año	Mes	<i>Forecast accuracy</i> positivo		<i>Forecast accuracy</i> negativo	
		Cantidad	Promedio	Cantidad	Promedio
2018	enero	74	56 %	40	-19 %
	febrero	57	43 %	45	-26 %
	marzo	69	41 %	45	-19 %
	abril	56	34 %	55	-18 %
	mayo	54	33 %	57	-16 %
	junio	50	28 %	63	-17 %
	julio	56	34 %	58	-22 %
	agosto	60	33 %	54	-25 %
	septiembre	46	43 %	70	-25 %
	octubre	48	45 %	66	-20 %
	noviembre	44	28 %	71	-19 %
	diciembre	46	28 %	67	-19 %
2019	enero	53	25 %	62	-19 %
	febrero	66	33 %	46	-21 %
	marzo	75	28 %	38	-19 %
	abril	60	29 %	55	-22 %
	mayo	57	25 %	60	-28 %
	junio	58	23 %	59	-23 %

Anexo 2: Porcentaje de cumplimiento por principio activo

Cada unidad de negocio está conformada por una cantidad de medicamentos de distintos principios activos. A continuación se muestra el porcentaje de cumplimiento promedio de cada principio activo en el año 2019.

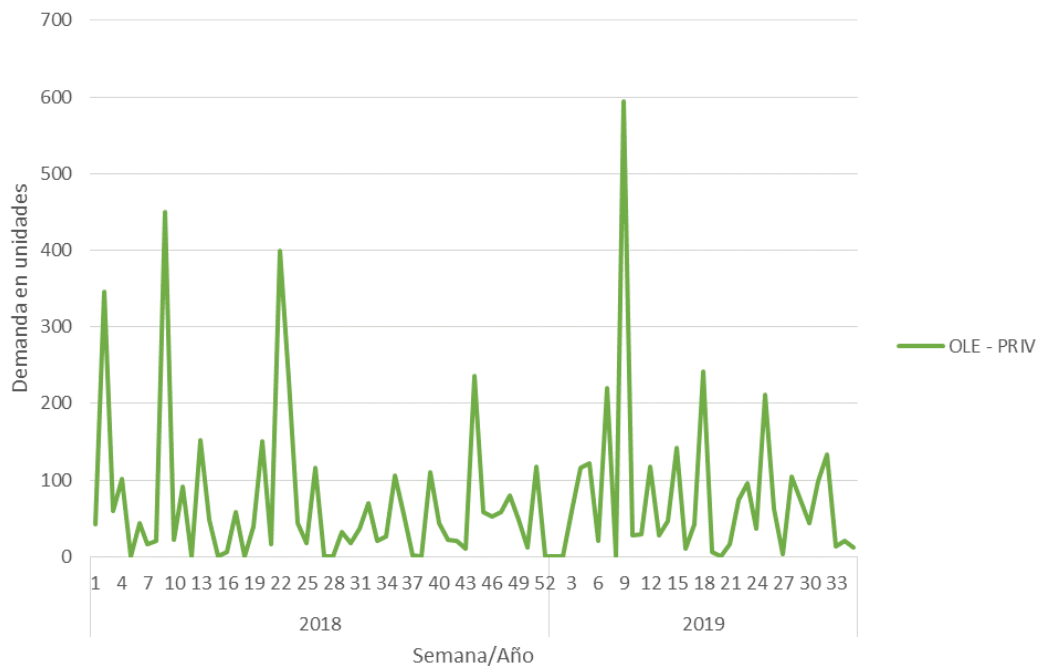
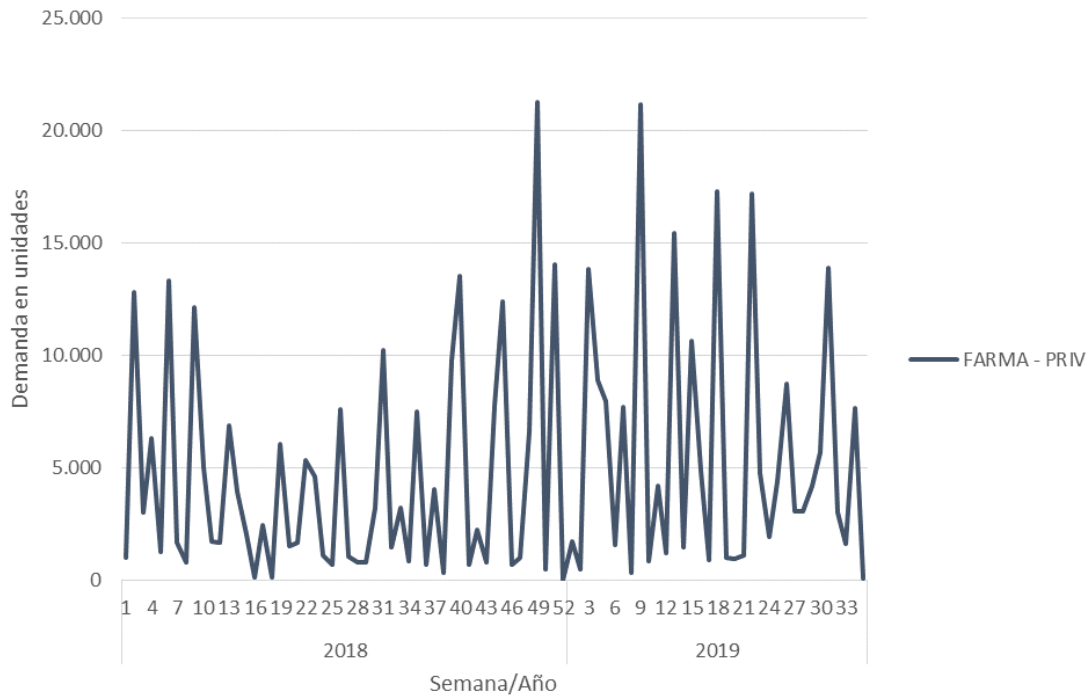
Línea	Unidad	Cumplimiento promedio en 2019
FARMA	CARDIO METABOLICA	
	BEZAFIBRATO	83%
	CILOSTAZOL	17%
	DIOSMINA	50%
	LORCASERINA	100%
	NEBIVOLOL	67%
	OLMESARTAN	100%
	PANTOPRAZOL	100%
	POLIETILENGLICOL	83%
	DOLOR	
	ACIDO HIALURONICO	67%
	ETORICOXIB	0%
	GLUCOSAMINA	100%
	KETOROLAC	100%
	PARACETAMOL	17%
	NEUROCIENCIAS	
	BENZERAZIDA	25%
	DESVENLAFAXINA	0%
	FLUNARIZINA	50%
	FLUOXETINA	83%
	KETAZOLAM	100%
	MEMANTINE	33%
	MIRTAZAPINA	100%
	NIMODIPINA	100%
	PRAMIPEXOL	67%
	QUETIAPINA	25%
	SERTRALINA HCL	33%
	TOPIRAMATO	20%
	ZOLPIDEM	67%
	PRIMARY CARE	
	DICLOFENAC	44%
	GABAPENTIN	67%

	LEVOFLOXACINA	50%
	MAGNESIO	67%
	PREGABALINA	0%
	UGI	
	CABERGOLINA	0%
	CIPROTERONA+ETINILESTRADIOL	83%
	CLINDAMICINA	67%
	DENOSUMAB-FA	83%
	DROSPIRENONA	100%
	FENAZOPIRIDINA	83%
	FLAVOXATO	100%
	FLUCONAZOL	83%
	IBANDRONATO	100%
	OXIBUTININA	100%
	TAMSULOSINA	50%
OLE	ONCO HEMATO	
	AZACITIDINA	67%
	FILGRASTIM	67%
	LENALIDOMIDA	100%
	POMALIDOMIDA	83%
	ROMIPLOSTIN	17%
	TALIDOMIDA	83%
	ONCOLOGIA	
	ANASTRAZOLE	100%
	DOXORRUBICINA LIPOSOMAL	100%
	NAB-PACLITAXEL	100%
	PANITUMUMAB	67%
	TAMOXIFENO	100%
	URO-ONCO	
	BICALUTAMIDA	17%
	DENOSUMAB-OL	33%
	ENZALUTAMIDA	100%
	FLUTAMIDA	0%
	PAZOPANIB	67%
	TRIPTORELINA	0%

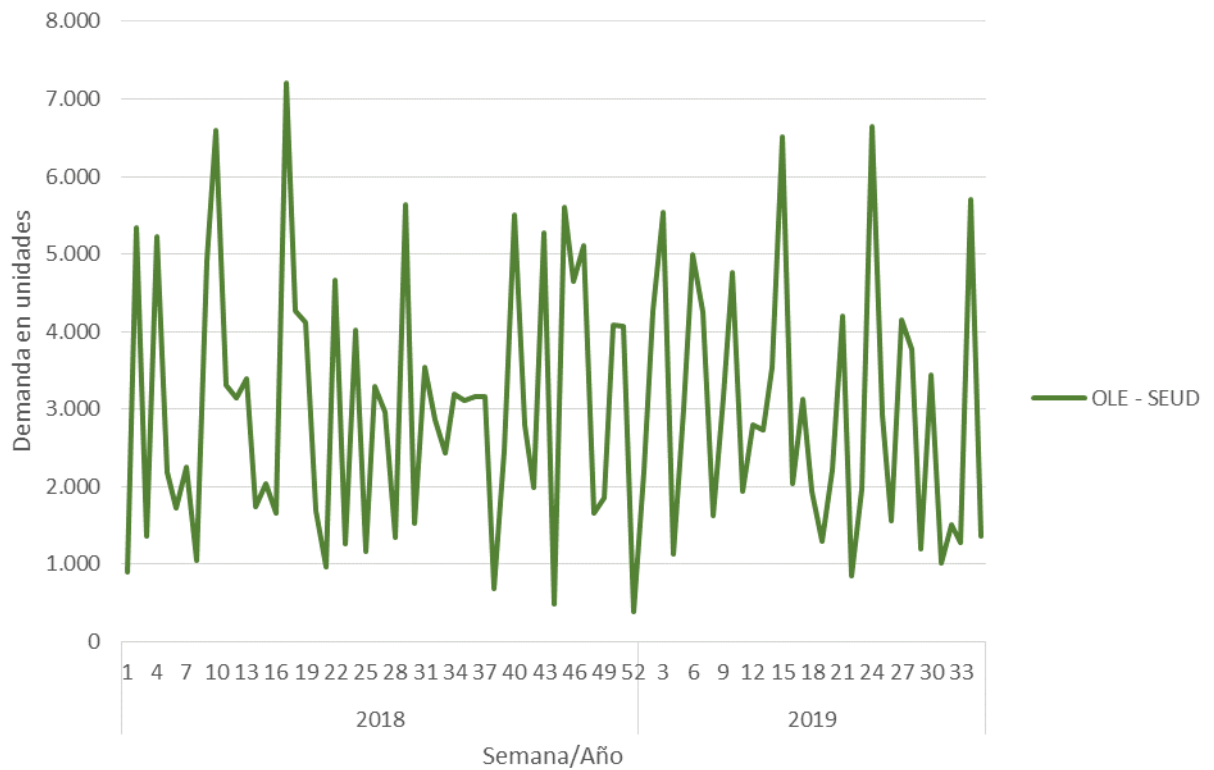
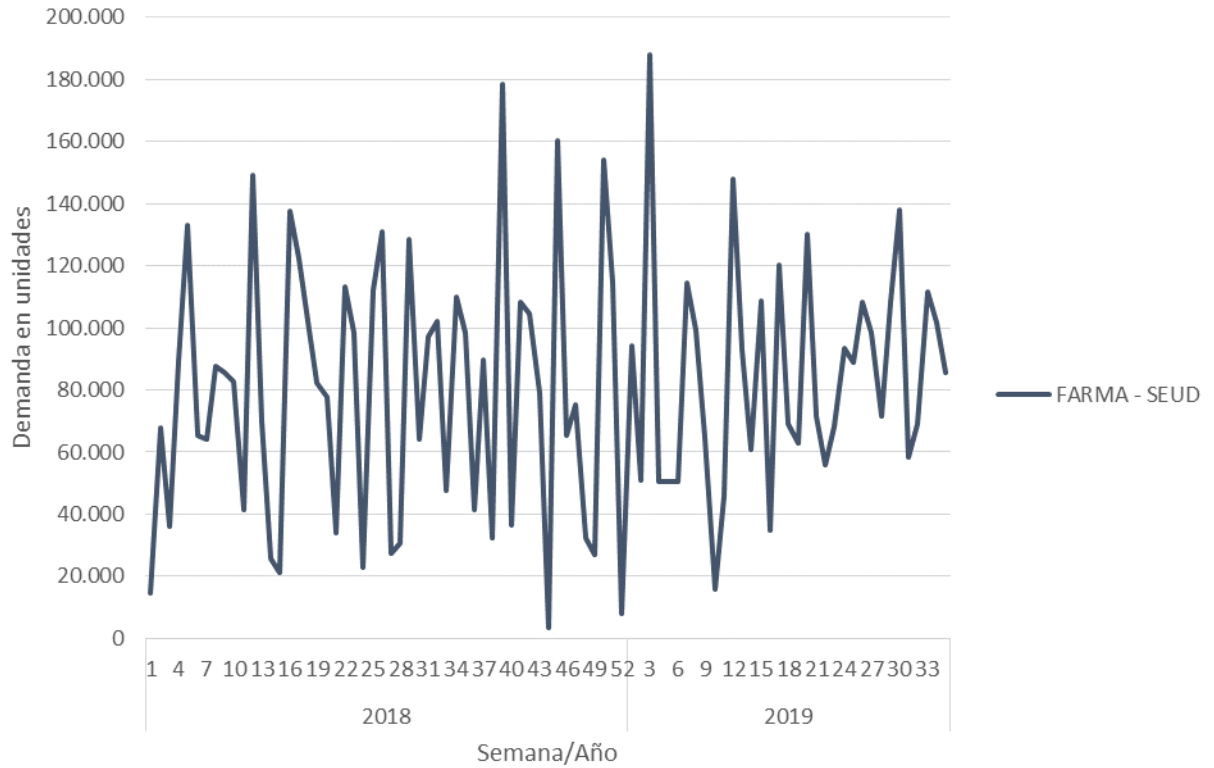
Anexo 3: Demanda mensual por modalidad de venta y línea de negocio

El cómo se comporta la demanda depende del canal de venta y línea de negocio.

A continuación se muestra la demanda del canal privado, para las líneas de FARMA y OLE.



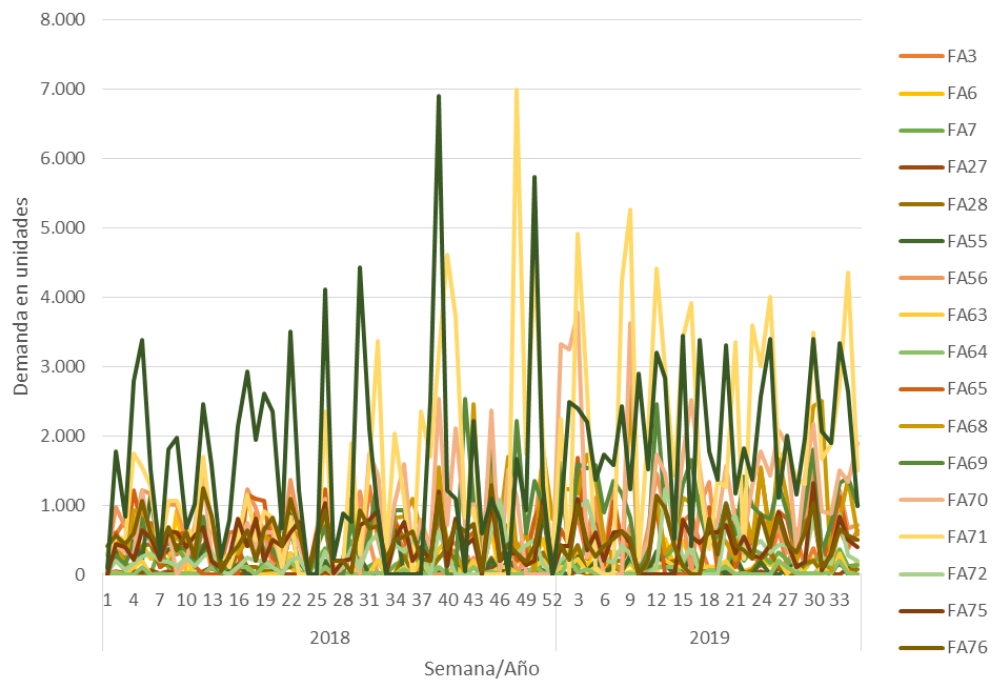
A continuación se muestra la demanda del canal pseudo-privado, para las líneas de FARMA y OLE.



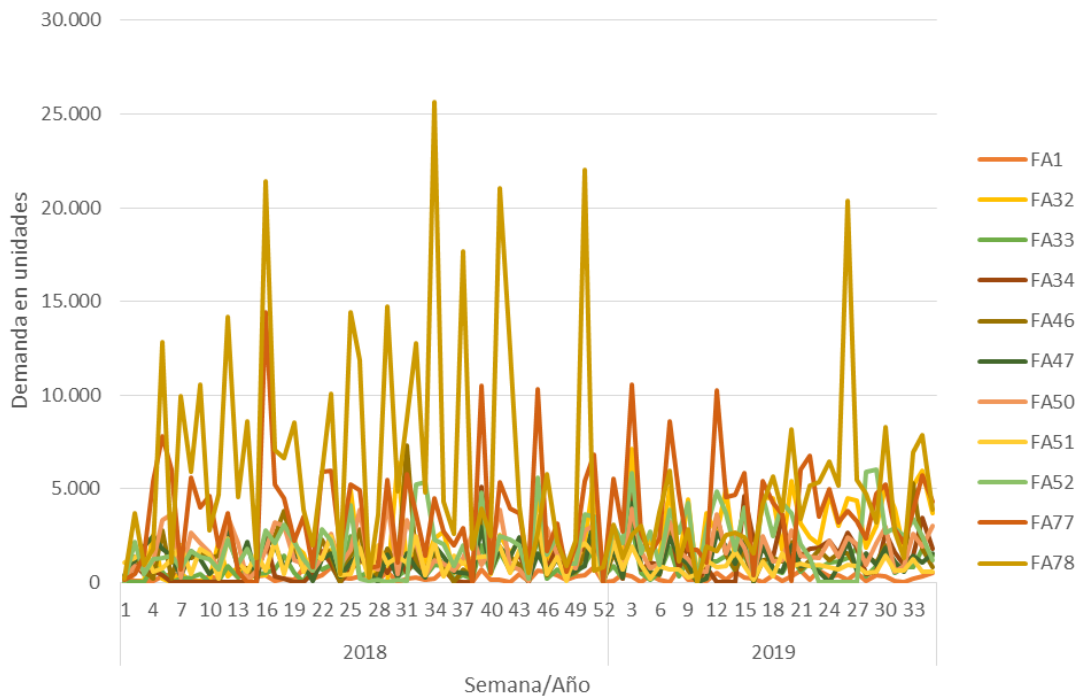
Anexo 4: Comparación de demandas por unidad de negocio

Cada línea de negocio tiene unidades que agrupan medicamentos que tratan condiciones en áreas de la salud similares. La demanda por cada medicamento se puede ver a continuación.

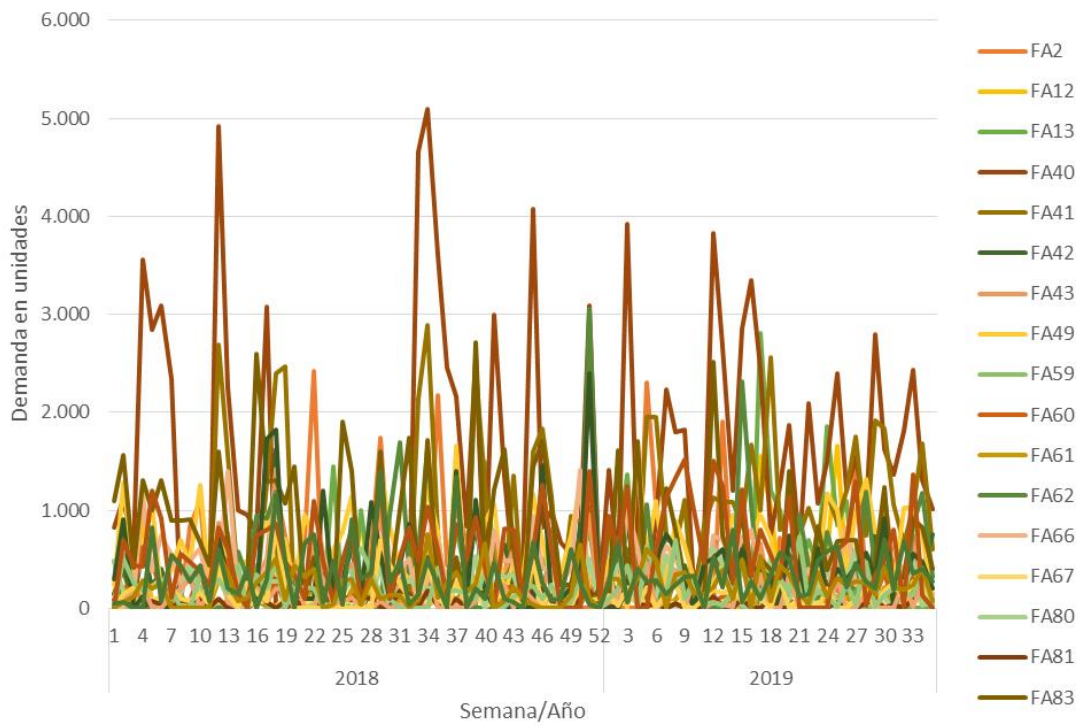
- Línea FARMA- Unidad Cardio Metabólica



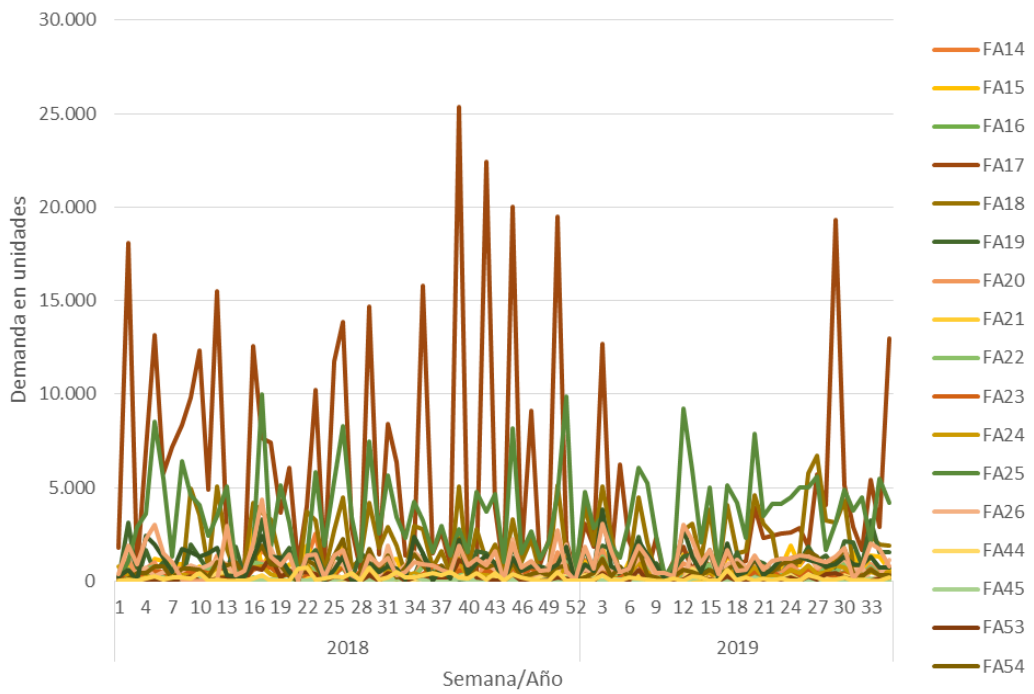
- Línea FARMA- Unidad Dolor



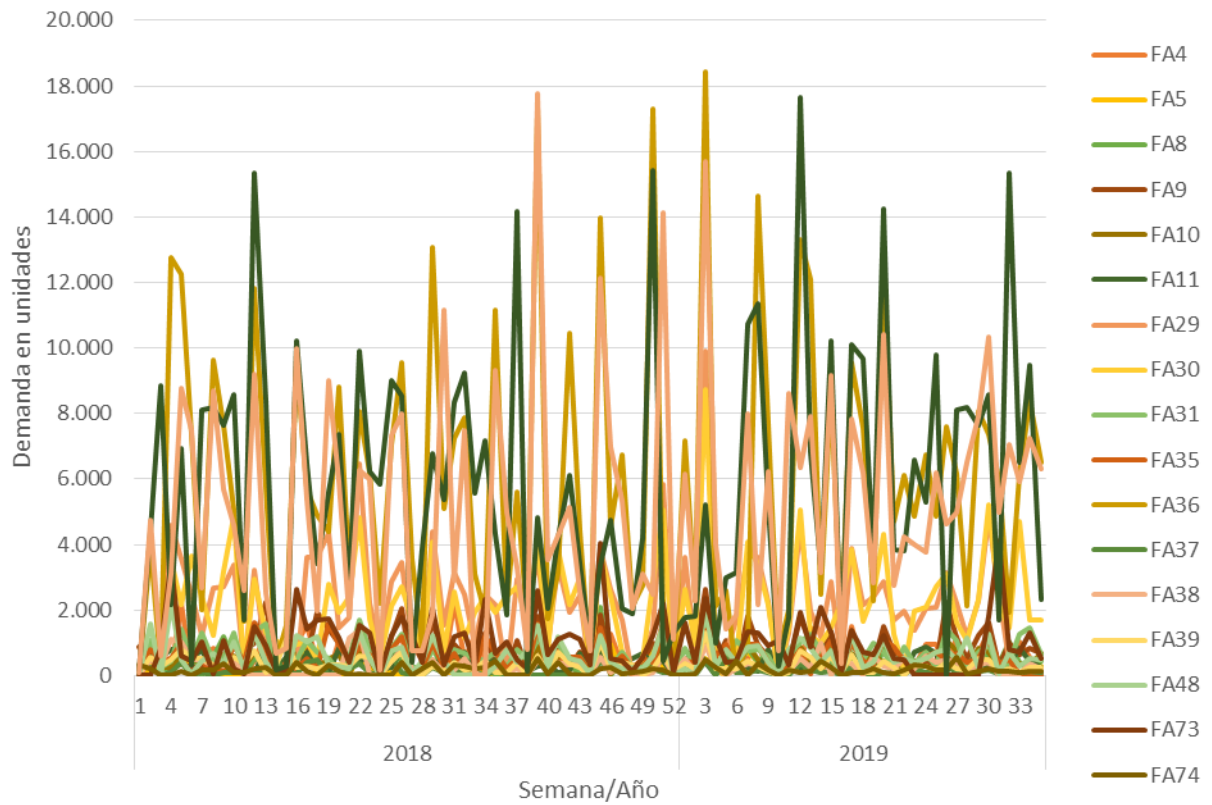
- Línea FARMA- Unidad Neurociencias



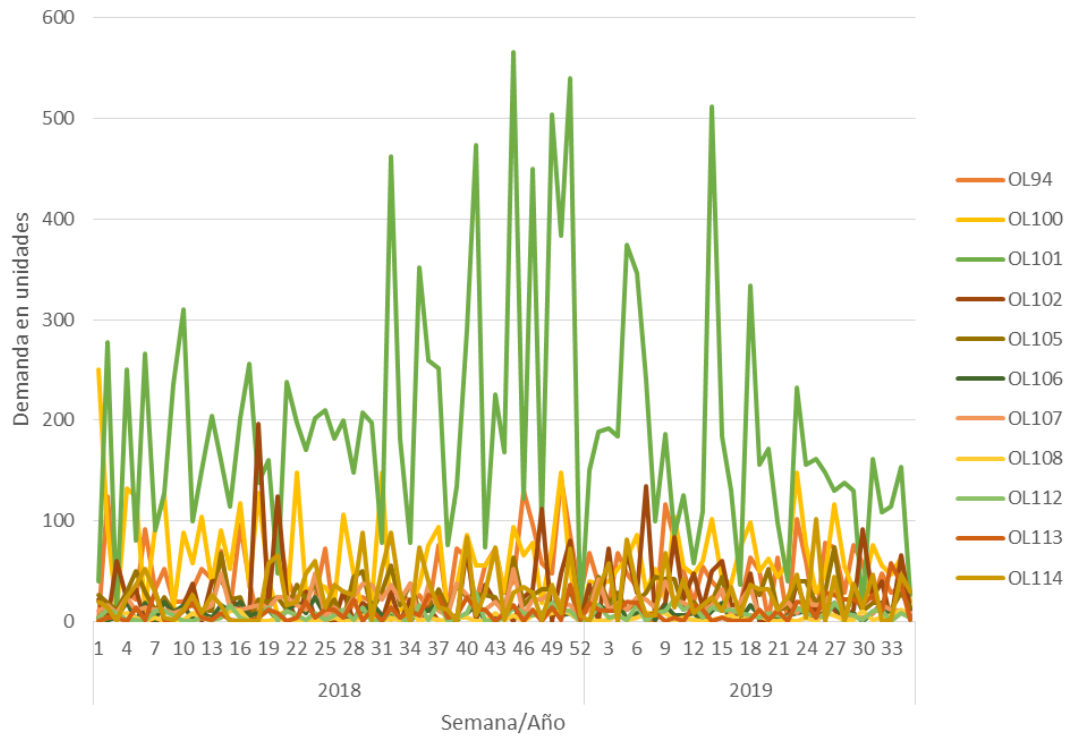
- Línea FARMA- Unidad Primary Care



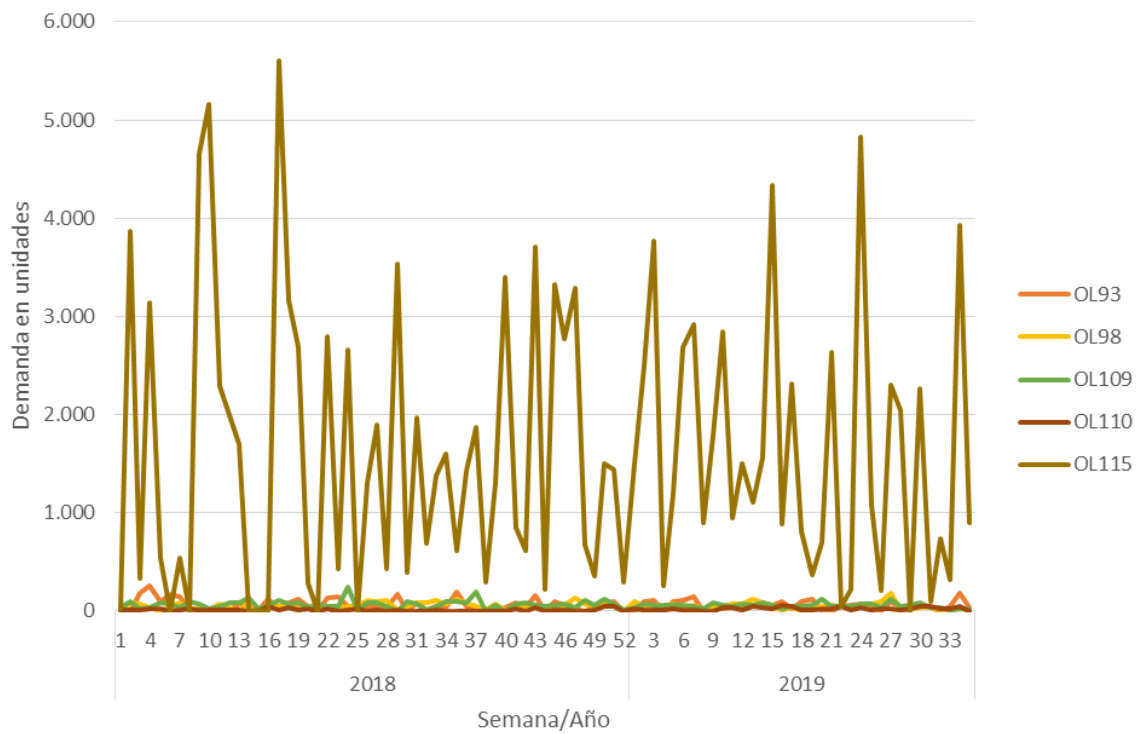
- Línea FARMA- Unidad UGI



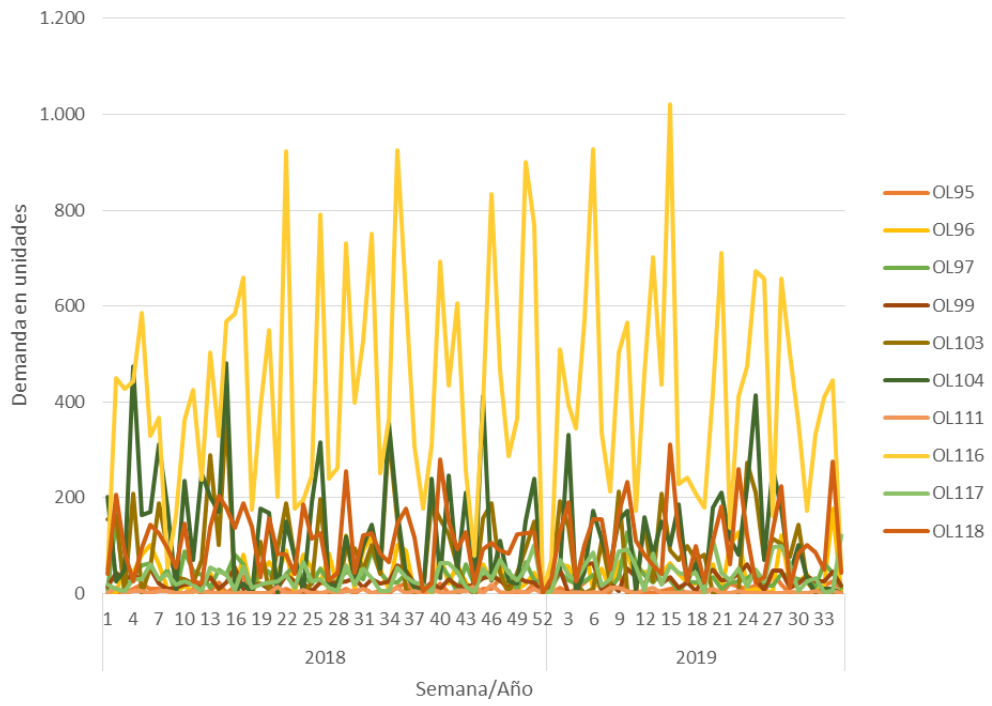
- Línea OLE - Unidad Onco Hemato



- Línea OLE - Unidad Oncología



- Línea OLE- Uro-Onco



Anexo 5: Cantidad de medicamentos por principio activo (familia)

Los medicamentos comercializados por el laboratorio farmacéutico, además de estar agrupados en líneas y unidades de negocio, se pueden desagregar en principios activos. A estos principios activos se les llama familia y dentro de cada una se encuentran medicamentos de la misma composición química, utilizados para el mismo propósito médico. La única diferencia entre ellos es la cantidad (cápsulas, gramos) en que viene su presentación.

Línea	Unidad	Familia	Cantidad
FARMA	CARDIOMETABOLICA	BEZAFIBRATO	1
FARMA	CARDIOMETABOLICA	CILOSTAZOL	2
FARMA	CARDIOMETABOLICA	DIOSMINA	2
FARMA	CARDIOMETABOLICA	LORCASERINA	2
FARMA	CARDIOMETABOLICA	NEBIVOLOL	3
FARMA	CARDIOMETABOLICA	OLMESARTAN	5
FARMA	CARDIOMETABOLICA	PANTOPRAZOL	2
FARMA	CARDIOMETABOLICA	POLIETILENGLICOL	1
FARMA	DOLOR	ACIDO HIALURONICO	1
FARMA	DOLOR	ETORICOXIB	3
FARMA	DOLOR	GLUCOSAMINA	2
FARMA	DOLOR	KETOROLAC	3
FARMA	DOLOR	PARACETAMOL	2
FARMA	NEUROCIENCIAS	BENZERAZIDA	1
FARMA	NEUROCIENCIAS	DESVENLAFAXINA	2
FARMA	NEUROCIENCIAS	FLUNARIZINA	2
FARMA	NEUROCIENCIAS	FLUOXETINA	2
FARMA	NEUROCIENCIAS	KETAZOLAM	1
FARMA	NEUROCIENCIAS	MEMANTINE	3
FARMA	NEUROCIENCIAS	MIRTAZAPINA	1
FARMA	NEUROCIENCIAS	NIMODIPINA	2
FARMA	NEUROCIENCIAS	PRAMIPEXOL	2
FARMA	NEUROCIENCIAS	QUETIAPINA	2
FARMA	NEUROCIENCIAS	SERTRALINA HCL	2
FARMA	NEUROCIENCIAS	TOPIRAMATO	3
FARMA	NEUROCIENCIAS	ZOLPIDEM	1
FARMA	PRIMARY CARE	DICLOFENAC	13
FARMA	PRIMARY CARE	GABAPENTIN	2
FARMA	PRIMARY CARE	LEVOFLOXACINA	2

FARMA	PRIMARY CARE	MAGNESIO	2
FARMA	PRIMARY CARE	PREGABALINA	1
FARMA	UGI	CABERGOLINA	2
FARMA	UGI	CIPROTERONA+ETINILESTRADIO L	1
FARMA	UGI	CLINDAMICINA	2
FARMA	UGI	DENOSUMAB-FA	1
FARMA	UGI	DROSPIRENONA	3
FARMA	UGI	FENAZOPIRIDINA	1
FARMA	UGI	FLAVOXATO	1
FARMA	UGI	FLUCONAZOL	3
FARMA	UGI	IBANDRONATO	1
FARMA	UGI	OXIBUTININA	2
FARMA	UGI	TAMSULOSINA	2
OLE	ONCO HEMATO	AZACITIDINA	1
OLE	ONCO HEMATO	FILGRASTIM	3
OLE	ONCO HEMATO	LENALIDOMIDA	4
OLE	ONCO HEMATO	POMALIDOMIDA	1
OLE	ONCO HEMATO	ROMIPLOSTIN	1
OLE	ONCO HEMATO	TALIDOMIDA	1
OLE	ONCOLOGIA	ANASTRAZOLE	1
OLE	ONCOLOGIA	DOXORRUBICINA LIPOSOMAL	1
OLE	ONCOLOGIA	NAB-PACLITAXEL	1
OLE	ONCOLOGIA	PANITUMUMAB	1
OLE	ONCOLOGIA	TAMOXIFENO	1
OLE	URO-ONCO	BICALUTAMIDA	2
OLE	URO-ONCO	DENOSUMAB-OL	1
OLE	URO-ONCO	ENZALUTAMIDA	1
OLE	URO-ONCO	FLUTAMIDA	2
OLE	URO-ONCO	PAZOPANIB	1
OLE	URO-ONCO	TRIPTORELINA	3

Anexo 6: Modelos estimados por serie de tiempo producto-canal

La tabla que sigue muestra las métricas de ajuste y pronóstico (RMSE y Accuracy mensual) para los modelos estimados para cada medicamento.

SKU	Unidad	Canal	Modelos	RMSE	Accuracy mensual
FA28	CARDIO MET	Privado	Promedio móvil doble	31	84%
FA28	CARDIO MET	Privado	Promedio móvil simple	28	90%
FA28	CARDIO MET	Pseudo	Promedio móvil simple	116	14%
FA28	CARDIO MET	Pseudo	Promedio móvil doble	123	24%
FA3	CARDIO MET	Privado	Promedio móvil simple	41	76%
FA3	CARDIO MET	Privado	Suavizado exponencial simple	41	76%
FA3	CARDIO MET	Pseudo	ARIMA(1,0,1)	312	-4%
FA3	CARDIO MET	Pseudo	Promedio móvil simple	338	-21%
FA56	CARDIO MET	Privado	Suavizado exponencial simple	136	83%
FA56	CARDIO MET	Privado	Promedio móvil doble	131	83%
FA56	CARDIO MET	Pseudo	Aditivo de Holt-Winters	439	-3%
FA56	CARDIO MET	Pseudo	Multiplicativo de Holt-Winters	443	-4%
FA6	CARDIO MET	Privado	Promedio móvil simple	70	95%
FA6	CARDIO MET	Privado	Suavizado exponencial simple	114	99%
FA6	CARDIO MET	Pseudo	Promedio móvil simple	112	-4%
FA6	CARDIO MET	Pseudo	Suavizado exponencial doble	122	-9%
FA65	CARDIO MET	Privado	ARIMA(1,0,1)	42	47%
FA65	CARDIO MET	Privado	Promedio móvil doble	41	48%
FA65	CARDIO MET	Pseudo	Suavizado exponencial doble	431	0%
FA65	CARDIO MET	Pseudo	Promedio móvil doble	446	3%
FA71	CARDIO MET	Privado	Suavizado exponencial doble	220	30%
FA71	CARDIO MET	Privado	Promedio móvil doble	242	47%
FA71	CARDIO MET	Pseudo	Suavizado exponencial doble	1.410	-1%
FA71	CARDIO MET	Pseudo	Promedio móvil doble	1.534	-6%
FA76	CARDIO MET	Privado	Promedio móvil simple	71	-28%
FA76	CARDIO MET	Privado	ARIMA(2,0,2)	62	58%
FA76	CARDIO MET	Pseudo	ARIMA(0,0,1)	314	3%
FA76	CARDIO MET	Pseudo	Suavizado exponencial simple	321	10%
FA79	CARDIO MET	Privado	Promedio móvil simple	248	-15%

FA79	CARDIO MET	Privado	ARIMA(0,0,1)	228	-25%
FA79	CARDIO MET	Pseudo	Suavizado exponencial simple	1.370	-3%
FA79	CARDIO MET	Pseudo	Suavizado exponencial doble	1.366	4%
FA1	DOLOR	Privado	ARIMA(1,0,0)	13	27%
FA1	DOLOR	Privado	Suavizado exponencial doble	12	75%
FA1	DOLOR	Pseudo	ARIMA(0,0,1)	210	11%
FA1	DOLOR	Pseudo	Suavizado exponencial doble	216	18%
FA32	DOLOR	Privado	Suavizado exponencial simple	298	-79%
FA32	DOLOR	Privado	Promedio móvil simple	310	-199%
FA32	DOLOR	Pseudo	Suavizado exponencial doble	1.545	-5%
FA32	DOLOR	Pseudo	Promedio móvil doble	1.597	-7%
FA46	DOLOR	Privado	Promedio móvil simple	105	-42%
FA46	DOLOR	Privado	Promedio móvil doble	106	-55%
FA46	DOLOR	Pseudo	Suavizado exponencial doble	1.230	-43%
FA46	DOLOR	Pseudo	ARIMA(0,0,1)	1.127	-44%
FA50	DOLOR	Privado	Suavizado exponencial doble	86	0%
FA50	DOLOR	Privado	Suavizado exponencial simple	86	0%
FA50	DOLOR	Pseudo	Promedio móvil doble	950	-4%
FA50	DOLOR	Pseudo	Suavizado exponencial simple	1.181	4%
FA78	DOLOR	Privado	Promedio móvil simple	237	7%
FA78	DOLOR	Privado	Suavizado exponencial simple	294	-8%
FA78	DOLOR	Pseudo	Promedio móvil simple	5.199	-3%
FA78	DOLOR	Pseudo	Promedio móvil doble	4.643	8%
FA13	NEUROCIENCIAS	Privado	Suavizado exponencial doble	4	100%
FA13	NEUROCIENCIAS	Privado	ARIMA(0,1,1)	4	100%
FA13	NEUROCIENCIAS	Pseudo	Suavizado exponencial doble	469	100%
FA13	NEUROCIENCIAS	Pseudo	ARIMA(0,1,1)	473	100%
FA2	NEUROCIENCIAS	Privado	ARIMA(1,0,0)	150	93%
FA2	NEUROCIENCIAS	Privado	Promedio móvil simple	161	94%
FA2	NEUROCIENCIAS	Pseudo	Promedio móvil simple	496	19%
FA2	NEUROCIENCIAS	Pseudo	ARIMA(0,0,1)	538	29%
FA40	NEUROCIENCIAS	Privado	Promedio móvil doble	309	28%
FA40	NEUROCIENCIAS	Privado	Suavizado exponencial simple	312	46%
FA40	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple	1.266	-3%
FA40	NEUROCIENCIAS	Pseudo	Promedio móvil doble	1.066	6%
FA43	NEUROCIENCIAS	Privado	Promedio móvil doble	38	-37%

FA43	NEUROCIENCIAS	Privado	Promedio móvil simple	38	-50%
FA43	NEUROCIENCIAS	Pseudo	Promedio móvil doble	228	9%
FA43	NEUROCIENCIAS	Pseudo	Promedio móvil simple	222	25%
FA49	NEUROCIENCIAS	Privado	Suavizado exponencial simple	64	3%
FA49	NEUROCIENCIAS	Privado	Promedio móvil doble	32	19%
FA49	NEUROCIENCIAS	Pseudo	Promedio móvil simple	390	2%
FA49	NEUROCIENCIAS	Pseudo	ARIMA(0,0,1)	390	-3%
FA60	NEUROCIENCIAS	Privado	Promedio móvil simple	4	-7%
FA60	NEUROCIENCIAS	Privado	SARIMA(0,0,1)(0,0,1)	3	24%
FA60	NEUROCIENCIAS	Pseudo	ARIMA(1,0,0)	173	6%
FA60	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple	171	42%
FA62	NEUROCIENCIAS	Privado	Promedio móvil doble	15	16%
FA62	NEUROCIENCIAS	Privado	Promedio móvil simple	14	-53%
FA62	NEUROCIENCIAS	Pseudo	Promedio móvil doble	508	-1%
FA62	NEUROCIENCIAS	Pseudo	Promedio móvil simple	591	4%
FA66	NEUROCIENCIAS	Privado	Suavizado exponencial simple	30	18%
FA66	NEUROCIENCIAS	Privado	Suavizado exponencial doble	30	19%
FA66	NEUROCIENCIAS	Pseudo	Promedio móvil doble	270	8%
FA66	NEUROCIENCIAS	Pseudo	SARIMA(1,0,1)(1,0,1)	286	12%
FA80	NEUROCIENCIAS	Privado	Promedio móvil simple	6	-27%
FA80	NEUROCIENCIAS	Privado	Promedio móvil doble	7	-38%
FA80	NEUROCIENCIAS	Pseudo	ARIMA(2,0,2)	183	24%
FA80	NEUROCIENCIAS	Pseudo	Suavizado exponencial doble	202	24%
FA84	NEUROCIENCIAS	Privado	Suavizado exponencial doble	2	-567%
FA84	NEUROCIENCIAS	Privado	Promedio móvil simple	2	-650%
FA84	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple	475	20%
FA84	NEUROCIENCIAS	Pseudo	Aditivo estacional	510	27%
FA86	NEUROCIENCIAS	Privado	Suavizado exponencial doble	6	-6%
FA86	NEUROCIENCIAS	Privado	Suavizado exponencial simple	6	-101%
FA86	NEUROCIENCIAS	Pseudo	Promedio móvil simple	115	-50%
FA86	NEUROCIENCIAS	Pseudo	SARIMA(1,0,0)(1,0,0)	150	-55%
FA90	NEUROCIENCIAS	Privado	ARIMA(0,0,1)	36	33%
FA90	NEUROCIENCIAS	Privado	Suavizado exponencial simple	38	41%
FA90	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple	454	5%
FA90	NEUROCIENCIAS	Pseudo	Suavizado exponencial doble	454	6%
FA92	NEUROCIENCIAS	Privado	ARIMA(0,0,1)	49	1%

FA92	NEUROCIENCIAS	Privado	Suavizado exponencial simple	51	26%
FA92	NEUROCIENCIAS	Pseudo	Promedio móvil doble	217	1%
FA92	NEUROCIENCIAS	Pseudo	SARIMA(0,0,1)(0,0,1)	267	1%
FA17	PRIMARY CARE	Privado	Suavizado exponencial simple	338	23%
FA17	PRIMARY CARE	Privado	Promedio móvil simple	287	29%
FA17	PRIMARY CARE	Pseudo	Suavizado exponencial doble	6.170	-16%
FA17	PRIMARY CARE	Pseudo	Suavizado exponencial simple	6.168	-16%
FA18	PRIMARY CARE	Privado	ARIMA(0,1,1)	219	32%
FA18	PRIMARY CARE	Privado	Suavizado exponencial simple	220	36%
FA18	PRIMARY CARE	Pseudo	ARIMA(0,0,1)	1.662	26%
FA18	PRIMARY CARE	Pseudo	Promedio móvil simple	1.621	43%
FA20	PRIMARY CARE	Privado	Suavizado exponencial doble	88	20%
FA20	PRIMARY CARE	Privado	Suavizado exponencial simple	88	-52%
FA20	PRIMARY CARE	Pseudo	Promedio móvil doble	404	2%
FA20	PRIMARY CARE	Pseudo	Suavizado exponencial simple	453	9%
FA23	PRIMARY CARE	Privado	Aditivo de Holt-Winters	41	-2%
FA23	PRIMARY CARE	Privado	Suavizado exponencial doble	41	2%
FA23	PRIMARY CARE	Pseudo	Promedio móvil simple	238	3%
FA23	PRIMARY CARE	Pseudo	Suavizado exponencial simple	261	7%
FA24	PRIMARY CARE	Privado	Suavizado exponencial doble	60	0%
FA24	PRIMARY CARE	Privado	Promedio móvil doble	66	-8%
FA24	PRIMARY CARE	Pseudo	ARIMA(0,0,1)	289	-7%
FA24	PRIMARY CARE	Pseudo	Promedio móvil simple	302	13%
FA25	PRIMARY CARE	Privado	Suavizado exponencial simple	460	9%
FA25	PRIMARY CARE	Privado	ARIMA(0,0,1)	439	-12%
FA25	PRIMARY CARE	Pseudo	Suavizado exponencial simple	2.547	1%
FA25	PRIMARY CARE	Pseudo	Suavizado exponencial doble	2.546	2%
FA44	PRIMARY CARE	Privado	ARIMA(0,1,1)	14	76%
FA44	PRIMARY CARE	Privado	Suavizado exponencial simple	14	76%
FA44	PRIMARY CARE	Pseudo	ARIMA(0,0,1)	181	-10%
FA44	PRIMARY CARE	Pseudo	Promedio móvil doble	148	71%
FA54	PRIMARY CARE	Privado	Suavizado exponencial simple	52	-13%
FA54	PRIMARY CARE	Privado	Suavizado exponencial doble	53	17%
FA54	PRIMARY CARE	Pseudo	ARIMA(0,0,1)	560	24%
FA54	PRIMARY CARE	Pseudo	Promedio móvil simple	473	45%
FA58	PRIMARY CARE	Privado	Suavizado exponencial simple	105	69%

FA58	PRIMARY CARE	Privado	Suavizado exponencial doble	105	72%
FA58	PRIMARY CARE	Pseudo	Promedio móvil doble	718	5%
FA58	PRIMARY CARE	Pseudo	Suavizado exponencial simple	864	5%
FA82	PRIMARY CARE	Privado	Promedio móvil simple	18	13%
FA82	PRIMARY CARE	Privado	SARIMA(1,0,0)(1,0,0)	19	15%
FA82	PRIMARY CARE	Pseudo	Promedio móvil doble	158	21%
FA82	PRIMARY CARE	Pseudo	SARIMA(0,0,1)(1,0,0)	197	24%
FA11	UGI	Privado	Promedio móvil simple	4	-16%
FA11	UGI	Privado	Promedio móvil doble	4	-26%
FA11	UGI	Pseudo	Suavizado exponencial doble	295	-5%
FA11	UGI	Pseudo	Suavizado exponencial simple	295	-6%
FA30	UGI	Privado	Suavizado exponencial doble	73	-12%
FA30	UGI	Privado	Promedio móvil simple	79	-54%
FA30	UGI	Pseudo	Suavizado exponencial simple	1.669	3%
FA30	UGI	Pseudo	Promedio móvil simple	1.600	4%
FA35	UGI	Privado	ARIMA(1,0,0)	52	91%
FA35	UGI	Privado	Promedio móvil doble	57	92%
FA35	UGI	Pseudo	ARIMA(0,0,1)	407	54%
FA35	UGI	Pseudo	Promedio móvil simple	361	67%
FA36	UGI	Privado	Promedio móvil simple	540	22%
FA36	UGI	Privado	Promedio móvil doble	562	30%
FA36	UGI	Pseudo	ARIMA(0,0,1)	3.802	-4%
FA36	UGI	Pseudo	Promedio móvil doble	4.544	5%
FA39	UGI	Privado	Promedio móvil simple	6	-35%
FA39	UGI	Privado	ARIMA(1,0,0)	8	-67%
FA39	UGI	Pseudo	ARIMA(2,0,2)	254	18%
FA39	UGI	Pseudo	Promedio móvil doble	269	33%
FA4	UGI	Privado	Promedio móvil doble	17	4%
FA4	UGI	Privado	Suavizado exponencial simple	16	50%
FA4	UGI	Pseudo	Suavizado exponencial doble	546	-15%
FA4	UGI	Pseudo	Promedio móvil doble	464	16%
FA48	UGI	Privado	Promedio móvil doble	66	30%
FA48	UGI	Privado	Promedio móvil simple	63	51%
FA48	UGI	Pseudo	ARIMA(0,0,1)	473	24%
FA48	UGI	Pseudo	Promedio móvil simple	442	29%
FA73	UGI	Privado	Suavizado exponencial doble	171	-4%

FA73	UGI	Privado	ARIMA(1,0,0)	155	-51%
FA73	UGI	Pseudo	ARIMA(1,0,0)	792	-61%
FA73	UGI	Pseudo	Promedio móvil simple	765	-117%
FA8	UGI	Privado	Promedio móvil doble	91	-1%
FA8	UGI	Privado	ARIMA(0,0,1)	91	-19%
FA8	UGI	Pseudo	Suavizado exponencial doble	384	-8%
FA8	UGI	Pseudo	Suavizado exponencial simple	385	-16%
FA88	UGI	Privado	ARIMA(1,0,0)	520	-19%
FA88	UGI	Privado	Suavizado exponencial doble	565	23%
FA88	UGI	Pseudo	Suavizado exponencial simple	3.867	-2%
FA88	UGI	Pseudo	Suavizado exponencial doble	3.853	3%
FA9	UGI	Privado	Suavizado exponencial simple	71	50%
FA9	UGI	Privado	Suavizado exponencial doble	71	56%
FA9	UGI	Pseudo	ARIMA(0,0,1)	574	0%
FA9	UGI	Pseudo	Suavizado exponencial doble	595	12%
OL100	ONCO HEMATO	Pseudo	Suavizado exponencial simple	50	-8%
OL100	ONCO HEMATO	Pseudo	Promedio móvil simple	32	13%
OL105	ONCO HEMATO	Pseudo	ARIMA(0,0,1)	15	30%
OL105	ONCO HEMATO	Pseudo	Suavizado exponencial doble	16	31%
OL112	ONCO HEMATO	Pseudo	ARIMA(0,1,1)	6	19%
OL112	ONCO HEMATO	Pseudo	Suavizado exponencial simple	6	21%
OL113	ONCO HEMATO	Pseudo	Promedio móvil simple	8	-38%
OL113	ONCO HEMATO	Pseudo	Suavizado exponencial doble	9	-61%
OL114	ONCO HEMATO	Privado	ARIMA(0,0,1)	1	44%
OL114	ONCO HEMATO	Privado	Promedio móvil doble	1	-110%
OL114	ONCO HEMATO	Pseudo	Suavizado exponencial simple	26	11%
OL114	ONCO HEMATO	Pseudo	Promedio móvil simple	25	11%
OL94	ONCO HEMATO	Pseudo	ARIMA(0,0,1)	35	7%
OL94	ONCO HEMATO	Pseudo	Suavizado exponencial simple	36	38%
OL109	ONCOLOGIA	Pseudo	Promedio móvil doble	24	56%
OL109	ONCOLOGIA	Pseudo	Promedio móvil simple	25	56%
OL110	ONCOLOGIA	Pseudo	Promedio móvil doble	16	1%
OL110	ONCOLOGIA	Pseudo	Promedio móvil simple	16	-7%
OL115	ONCOLOGIA	Privado	ARIMA(1,0,0)	91	3%
OL115	ONCOLOGIA	Privado	Suavizado exponencial simple	95	-8%
OL115	ONCOLOGIA	Pseudo	Promedio móvil doble	1.248	18%

OL115	ONCOLOGIA	Pseudo	Suavizado exponencial simple	1.503	24%
OL93	ONCOLOGIA	Pseudo	ARIMA(2,0,2)	49	6%
OL93	ONCOLOGIA	Pseudo	Promedio móvil doble	43	-9%
OL98	ONCOLOGIA	Privado	Promedio móvil doble	9	100%
OL98	ONCOLOGIA	Privado	Promedio móvil simple	9	100%
OL98	ONCOLOGIA	Pseudo	ARIMA(1,0,0)	33	65%
OL98	ONCOLOGIA	Pseudo	Promedio móvil simple	35	66%
OL104	URO-ONCO	Privado	Promedio móvil simple	7	100%
OL104	URO-ONCO	Privado	Promedio móvil doble	7	100%
OL104	URO-ONCO	Pseudo	ARIMA(0,0,1)	126	85%
OL104	URO-ONCO	Pseudo	Promedio móvil simple	107	91%
OL111	URO-ONCO	Pseudo	Suavizado exponencial simple	5	10%
OL111	URO-ONCO	Pseudo	Promedio móvil doble	4	-30%
OL116	URO-ONCO	Privado	Promedio móvil simple	2	16%
OL116	URO-ONCO	Privado	Suavizado exponencial simple	2	19%
OL116	URO-ONCO	Pseudo	ARIMA(1,0,1)	225	33%
OL116	URO-ONCO	Pseudo	Promedio móvil simple	241	37%
OL96	URO-ONCO	Privado	Promedio móvil simple	1	100%
OL96	URO-ONCO	Privado	ARIMA(0,0,1)	1	100%
OL96	URO-ONCO	Pseudo	ARIMA(0,0,1)	36	-20%
OL96	URO-ONCO	Pseudo	Suavizado exponencial doble	38	-25%
OL97	URO-ONCO	Pseudo	Promedio móvil simple	21	-4%
OL97	URO-ONCO	Pseudo	Suavizado exponencial simple	28	6%
OL99	URO-ONCO	Pseudo	Suavizado exponencial simple	16	-2%
OL99	URO-ONCO	Pseudo	Promedio móvil simple	16	4%

Anexo 7: Mejor método por serie de tiempo producto-canal

La siguiente tabla detalla los considerados “mejores métodos” para cada serie de tiempo.

SKU	Unidad	Canal	Mejor modelo
FA28	CARDIO MET	Privado	Promedio móvil doble
FA28	CARDIO MET	Pseudo	Promedio móvil simple
FA3	CARDIO MET	Privado	Promedio móvil simple
FA3	CARDIO MET	Pseudo	ARIMA(1,0,1)
FA56	CARDIO MET	Privado	Suavizado exponencial simple
FA56	CARDIO MET	Pseudo	Aditivo de Holt-Winters
FA6	CARDIO MET	Privado	Promedio móvil simple
FA6	CARDIO MET	Pseudo	Promedio móvil simple
FA65	CARDIO MET	Privado	ARIMA(1,0,1)
FA65	CARDIO MET	Pseudo	Suavizado exponencial doble
FA71	CARDIO MET	Privado	Suavizado exponencial doble
FA71	CARDIO MET	Pseudo	Suavizado exponencial doble
FA76	CARDIO MET	Privado	Promedio móvil simple
FA76	CARDIO MET	Pseudo	ARIMA(0,0,1)
FA79	CARDIO MET	Privado	Promedio móvil simple
FA79	CARDIO MET	Pseudo	Suavizado exponencial simple
FA1	DOLOR	Privado	ARIMA(1,0,0)
FA1	DOLOR	Pseudo	ARIMA(0,0,1)
FA32	DOLOR	Privado	Suavizado exponencial simple
FA32	DOLOR	Pseudo	Suavizado exponencial doble
FA46	DOLOR	Privado	Promedio móvil simple
FA46	DOLOR	Pseudo	Suavizado exponencial doble
FA50	DOLOR	Privado	Suavizado exponencial doble
FA50	DOLOR	Pseudo	Promedio móvil doble
FA78	DOLOR	Privado	Promedio móvil simple
FA78	DOLOR	Pseudo	Promedio móvil simple
FA13	NEUROCIENCIAS	Privado	Suavizado exponencial doble
FA13	NEUROCIENCIAS	Pseudo	Suavizado exponencial doble
FA2	NEUROCIENCIAS	Privado	ARIMA(1,0,0)
FA2	NEUROCIENCIAS	Pseudo	Promedio móvil simple

FA40	NEUROCIENCIAS	Privado	Promedio móvil doble
FA40	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple
FA43	NEUROCIENCIAS	Privado	Promedio móvil doble
FA43	NEUROCIENCIAS	Pseudo	Promedio móvil doble
FA49	NEUROCIENCIAS	Privado	Suavizado exponencial simple
FA49	NEUROCIENCIAS	Pseudo	Promedio móvil simple
FA60	NEUROCIENCIAS	Privado	Promedio móvil simple
FA60	NEUROCIENCIAS	Pseudo	ARIMA(1,0,0)
FA62	NEUROCIENCIAS	Privado	Promedio móvil doble
FA62	NEUROCIENCIAS	Pseudo	Promedio móvil doble
FA66	NEUROCIENCIAS	Privado	Suavizado exponencial simple
FA66	NEUROCIENCIAS	Pseudo	Promedio móvil doble
FA80	NEUROCIENCIAS	Privado	Promedio móvil simple
FA80	NEUROCIENCIAS	Pseudo	ARIMA(2,0,2)
FA84	NEUROCIENCIAS	Privado	Suavizado exponencial doble
FA84	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple
FA86	NEUROCIENCIAS	Privado	Suavizado exponencial doble
FA86	NEUROCIENCIAS	Pseudo	Promedio móvil simple
FA90	NEUROCIENCIAS	Privado	ARIMA(0,0,1)
FA90	NEUROCIENCIAS	Pseudo	Suavizado exponencial simple
FA92	NEUROCIENCIAS	Privado	ARIMA(0,0,1)
FA92	NEUROCIENCIAS	Pseudo	Promedio móvil doble
FA17	PRIMARY CARE	Privado	Suavizado exponencial simple
FA17	PRIMARY CARE	Pseudo	Suavizado exponencial doble
FA18	PRIMARY CARE	Privado	ARIMA(0,1,1)
FA18	PRIMARY CARE	Pseudo	ARIMA(0,0,1)
FA20	PRIMARY CARE	Privado	Suavizado exponencial doble
FA20	PRIMARY CARE	Pseudo	Promedio móvil doble
FA23	PRIMARY CARE	Privado	Aditivo de Holt-Winters
FA23	PRIMARY CARE	Pseudo	Promedio móvil simple
FA24	PRIMARY CARE	Privado	Suavizado exponencial doble
FA24	PRIMARY CARE	Pseudo	ARIMA(0,0,1)
FA25	PRIMARY CARE	Privado	Suavizado exponencial simple
FA25	PRIMARY CARE	Pseudo	Suavizado exponencial simple
FA44	PRIMARY CARE	Privado	ARIMA(0,1,1)
FA44	PRIMARY CARE	Pseudo	ARIMA(0,0,1)

FA54	PRIMARY CARE	Privado	Suavizado exponencial simple
FA54	PRIMARY CARE	Pseudo	ARIMA(0,0,1)
FA58	PRIMARY CARE	Privado	Suavizado exponencial simple
FA58	PRIMARY CARE	Pseudo	Promedio móvil doble
FA82	PRIMARY CARE	Privado	Promedio móvil simple
FA82	PRIMARY CARE	Pseudo	Promedio móvil doble
FA11	UGI	Privado	Promedio móvil simple
FA11	UGI	Pseudo	Suavizado exponencial doble
FA30	UGI	Privado	Suavizado exponencial doble
FA30	UGI	Pseudo	Suavizado exponencial simple
FA35	UGI	Privado	ARIMA(1,0,0)
FA35	UGI	Pseudo	ARIMA(0,0,1)
FA36	UGI	Privado	Promedio móvil simple
FA36	UGI	Pseudo	ARIMA(0,0,1)
FA39	UGI	Privado	Promedio móvil simple
FA39	UGI	Pseudo	ARIMA(2,0,2)
FA4	UGI	Privado	Promedio móvil doble
FA4	UGI	Pseudo	Suavizado exponencial doble
FA48	UGI	Privado	Promedio móvil doble
FA48	UGI	Pseudo	ARIMA(0,0,1)
FA73	UGI	Privado	Suavizado exponencial doble
FA73	UGI	Pseudo	ARIMA(1,0,0)
FA8	UGI	Privado	Promedio móvil doble
FA8	UGI	Pseudo	Suavizado exponencial doble
FA88	UGI	Privado	ARIMA(1,0,0)
FA88	UGI	Pseudo	Suavizado exponencial simple
FA9	UGI	Privado	Suavizado exponencial simple
FA9	UGI	Pseudo	ARIMA(0,0,1)
OL100	ONCO HEMATO	Pseudo	Suavizado exponencial simple
OL105	ONCO HEMATO	Pseudo	ARIMA(0,0,1)
OL112	ONCO HEMATO	Pseudo	ARIMA(0,1,1)
OL113	ONCO HEMATO	Pseudo	Promedio móvil simple
OL114	ONCO HEMATO	Privado	ARIMA(0,0,1)
OL114	ONCO HEMATO	Pseudo	Suavizado exponencial simple
OL94	ONCO HEMATO	Pseudo	ARIMA(0,0,1)
OL109	ONCOLOGIA	Pseudo	Promedio móvil doble

OL110	ONCOLOGIA	Pseudo	Promedio móvil doble
OL115	ONCOLOGIA	Privado	ARIMA(1,0,0)
OL115	ONCOLOGIA	Pseudo	Promedio móvil doble
OL93	ONCOLOGIA	Pseudo	ARIMA(2,0,2)
OL98	ONCOLOGIA	Privado	Promedio móvil doble
OL98	ONCOLOGIA	Pseudo	ARIMA(1,0,0)
OL104	URO-ONCO	Privado	Promedio móvil simple
OL104	URO-ONCO	Pseudo	ARIMA(0,0,1)
OL111	URO-ONCO	Pseudo	Suavizado exponencial simple
OL116	URO-ONCO	Privado	Promedio móvil simple
OL116	URO-ONCO	Pseudo	ARIMA(1,0,1)
OL96	URO-ONCO	Privado	Promedio móvil simple
OL96	URO-ONCO	Pseudo	ARIMA(0,0,1)
OL97	URO-ONCO	Pseudo	Promedio móvil simple
OL99	URO-ONCO	Pseudo	Suavizado exponencial simple

Anexo 8: Forecast accuracy de mejor modelo comparado con promedio de forecast accuracy del 2019

Se compara el *forecast accuracy* dado por los considerados “mejores modelos” con el promedio de este indicador a largo del año 2019.

Producto	Forecast accuracy con mejores modelos	Promedio de forecast accuracy en 2019
FA1	12%	-12%
FA11	-5%	7%
FA13	-8%	52%
FA17	-14%	-56%
FA18	26%	58%
FA2	26%	40%
FA20	5%	38%
FA23	2%	25%
FA24	-6%	19%
FA25	2%	12%
FA28	17%	-7%
FA3	1%	7%
FA30	3%	10%
FA32	-9%	46%
FA35	57%	14%
FA36	-3%	0%
FA39	17%	-12%
FA4	-15%	-16%
FA40	-1%	-1%
FA43	9%	-1%
FA44	14%	5%
FA46	-43%	3%
FA48	24%	1%
FA49	2%	-7%
FA50	-3%	6%
FA54	20%	-13%
FA56	3%	-2%
FA58	8%	16%
FA6	2%	13%
FA60	6%	-21%
FA62	1%	-3%

FA65	2%	-7%
FA66	8%	-13%
FA71	0%	-2%
FA73	-4%	-4%
FA76	2%	-1%
FA78	-3%	-44%
FA79	-4%	0%
FA8	-8%	5%
FA80	22%	-11%
FA82	20%	59%
FA84	19%	-20%
FA86	-5%	-8%
FA88	-3%	16%
FA9	2%	14%
FA90	6%	38%
FA92	1%	-15%
OL100	-8%	-10%
OL104	15%	-33%
OL105	30%	5%
OL109	16%	-13%
OL110	1%	9%
OL111	9%	-3%
OL112	18%	9%
OL113	-38%	-5%
OL114	12%	-8%
OL115	18%	2%
OL116	33%	44%
OL93	6%	1%
OL94	7%	11%
OL96	-17%	20%
OL97	-4%	-28%
OL98	68%	-9%