



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

MODELO DE ESTIMACIÓN DE VENTA DE POLLO EN GRANDES CADENAS DE SUPERMERCADOS EN CHILE

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL
INDUSTRIAL

JAVIERA LISETTE BELÉN LANDEROS MATELUNA

PROFESORA GUÍA:
ALEJANDRA PUENTE CHANDIA

PROFESOR CO-GUÍA:
OMAR CERDA INOSTROZA

COMISIÓN:
RONALD FISCHER BARKAN

SANTIAGO DE CHILE
2023

**RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE:** Ingeniera Civil Industrial
ESTUDIANTE: Javiera Lisette Belén Landeros Mateluna
FECHA: 2023
PROFESOR GUÍA: Alejandra Puente Chandia.

MODELO DE ESTIMACIÓN DE VENTA DE POLLO EN GRANDES CADENAS DE SUPERMERCADOS EN CHILE

La empresa donde se desarrolla el proyecto de título es un conglomerado de marcas que se dedica a la comercialización, producción y distribución de productos cárnicos, específicamente de pollo, cerdo, pavo y algunos peces como el salmón.

La principal problemática que se quiere abordar con el proyecto es el no cumplimiento de metas de participación de mercado, ya que para poder seguir siendo líder de este mercado es importante preocuparse por cómo se comporta la demanda y que medidas se pueden tomar para anteponerse a cambios en esta, con el fin de mantener o mejorar el posicionamiento de la empresa en el mercado.

Para resolver esta problemática es que se define el objetivo principal del proyecto como: Construir un modelo de estimación de demanda de mercado para pollo en supermercados, con el fin de definir metas de participación de mercados ambiciosas pero alcanzables. Este proyecto se aborda con la metodología CRISP-DM, lo que permite tener un proyecto iterativo que consulta con expertos de la empresa.

Se aborda el objetivo con 3 tipos de modelos diferentes, alisamiento exponencial, regresiones lineales y ARIMA, estos modelos se evalúan con las métricas estadísticas MSE, MAPE y, además, se utiliza el AIC y BIC para algunos modelos. Se obtienen resultados con menores errores y con mejores métricas para los modelos realizados con ARIMA, tanto para cada cadena individualmente como para el total nacional que las incluye a todas. En particular el modelo con mejores métricas y resultados más cercanos a los reales es el modelo total nacional.

Se concluye que, en general el modelo ARIMA logra estimar de mejor manera la estacionalidad y tendencia de las series, esto se debe a la cantidad de variantes que puede tener el modelo, lo que lo hace más flexible y con mayor cantidad de ajustes. Además, que el modelo "total nacional" al incluir todas las cadenas tiene una tendencia y estacionalidad más marcada, por lo que todos los modelos utilizados, fueron capaces de ajustar de mejor manera el modelo.

Se recomienda como trabajo futuro, por un lado, incluir más historia en los modelos y volver a ajustarlos, ya que se incluyen solo datos desde enero 2020 a agosto 2022, mayormente tiempo de pandemia, lo que pudo afectar el consumo de pollo. Por otro lado, obtener detalles de venta de cadenas importantes como Walmart y Tottus, ya que estas dos se encuentran en el grupo "otras cadenas", donde también están incluidas cadenas de menores ventas. "Otras cadenas" es el modelo con peores resultados, ya que agrupa supermercados con distintos públicos objetivos, lo que genera que no exista una tendencia marcada en el grupo.

Tabla de Contenido

1. Introducción	1
1.1. Antecedentes generales	1
1.2 Descripción del proyecto y justificación	5
2. Objetivos	7
2.1 Objetivo General	7
2.2 Objetivo Específicos	7
3. Marco Teórico	8
3.1 Conceptos claves	8
3.2 Modelos	8
3.3 Métricas de evaluación.....	9
4. Metodología	12
4.1 Comprensión del negocio:	12
4.2 Comprensión de los datos:.....	13
4.3 Análisis y preparación de datos	13
4.4 Modelamiento	14
4.5 Evaluación	14
4.6 Plan de Implementación piloto	14
5. Alcances	16
6. Resultados.....	17
6.1 Entendimiento del negocio	17
6.2 Comprensión y preparación de los datos	17
6.3 Análisis de datos.....	18
6.3.1 Análisis Univariado.....	18
6.3.2 Análisis Multivariado.....	25
6.4 Modelamiento	26
6.4.1 Modelo ARIMA	26
6.4.1.1 Modelo Total Nacional	28
6.4.1.2 Modelo Jumbo	30
6.4.1.3 Modelo Santa Isabel.....	31
6.4.1.4 Modelo Unimarc.....	33
6.4.1.5 Modelo otras cadenas.....	35

6.4.2	Modelo Regresión Lineal	37
6.4.2.1	Modelo Total Nacional	37
6.4.2.2	Modelo Jumbo	39
6.4.2.3	Modelo Santa Isabel.....	40
6.4.2.4	Modelo Unimarc	41
6.4.2.5	Modelo Otras cadenas	42
6.4.3	Modelo Alisamiento Exponencial.....	44
6.4.3.1	Modelo Total Nacional	44
6.4.3.2	Modelo Jumbo	45
6.4.3.3	Modelo Santa Isabel.....	46
6.4.3.4	Modelo Unimarc	48
6.4.3.5	Modelo Otras cadenas	49
6.5	Evaluación de Modelos	50
6.5.1	Modelos Total Nacional	51
6.5.2	Modelos Jumbo.....	52
6.5.3	Modelos Santa Isabel	52
6.5.4	Modelos Unimarc.....	53
6.5.5	Modelos de otras cadenas	54
6.6	Plan de Implementación	55
7.	Discusión	56
8.	Conclusiones	57
9.	Recomendaciones de trabajo futuro	58
	Bibliografía.....	59
	Anexos	62

1. Introducción

1.1. Antecedentes generales

Para entender la importancia del análisis de la venta en el mercado de pollo en los supermercados en Chile, es relevante entender el comportamiento del consumo mensual de alimentos en el país. En el año 2021, el 21% de la compra de alimentos en el país, corresponde, según Deloitte Chile, al consumo de carnes, ya sea rojas (bovino, cerdo u otras) o blancas (pollo, pavo u otras) siendo el segundo mayor gasto, después de la categoría otros alimentos, que incluye productos como, platos preparados, bebidas, aguas, entre otros. Este consumo de carnes según un análisis de la Asociación Gremial de Productores de Cerdo de Chile (ASPROCER), el 39% corresponde a pollo, siendo esta la carne de mayor consumo en el país, seguido por la carne de bovino con un 33% y luego por la carne de cerdo con un 23%, lo que se puede apreciar de mejor manera en el gráfico a continuación:

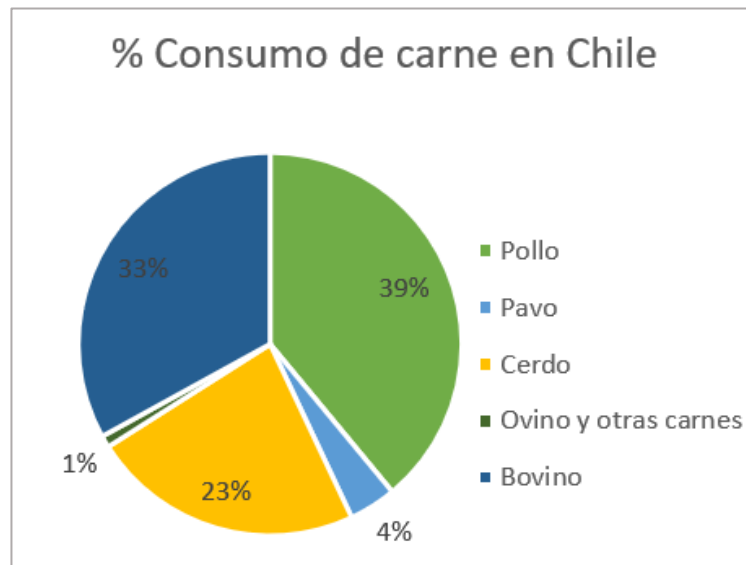


Gráfico 1: Porcentaje de Consumo de carne en Chile por tipo en el año 2021.
Fuente: Adaptación de análisis sectorial de ASPROCER año 2021.

El consumo per cápita de carne de pollo ha ido disminuyendo con los años, sin tener realmente clara la razón de esto, lo que se puede ver en el gráfico a continuación:

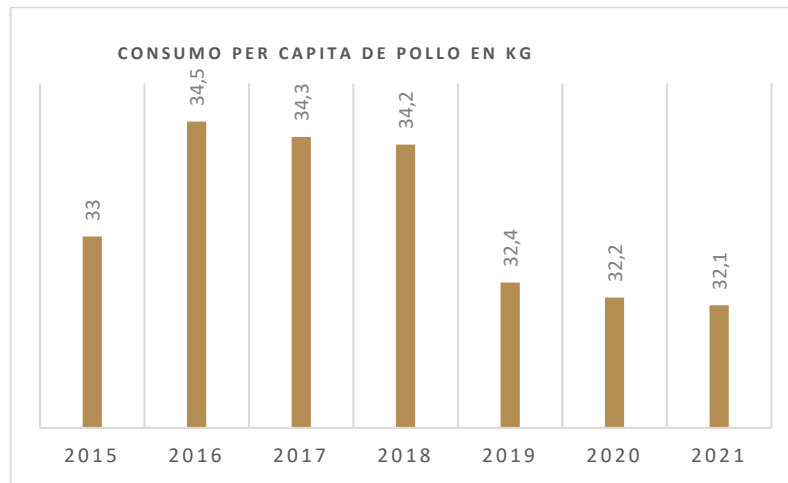


Gráfico 2: Consumo per cápita de carne de pollo en Chile
 Fuente: Adaptación de reporte Integrado de Agrosuper 2021.

En el gráfico 2 es posible apreciar en detalle, como ha disminuido la venta per cápita de carne de pollo en el mercado nacional, mostrando que el 2021 el consumo fue más bajo que años anteriores, esta diferencia no es tan grande comparada con la que hubo entre los años 2018-2019.

Dentro de las empresas que se dedican a la venta de pollo, está Agrosuper S.A., la cual es un conglomerado de marcas que nace en el año 1955 y es la empresa con mayores ventas de esta carne en Chile. El propósito principal de Agrosuper es “Alimentar lo bueno de la vida todos los días” (Matriz Agrosuper, 2021) para esto cuenta con 29 oficinas comerciales nacionales y 11 oficinas internacionales, en los países Corea del Sur, China, Japón, Estados Unidos, México e Italia.

La empresa cuenta con más de 19 mil trabajadores, separados en tres segmentos importantes, por un lado, se tiene el de carnes que incluye pollo, pavo y Cerdo, luego, el corporativo, el cual, se encarga de áreas más transversales para la compañía y, por último, el segmento acuícola, este se encarga de todo el proceso de producción, comercialización y distribución del salmón y otros peces. En la imagen a continuación se puede ver más en detalle el organigrama de Agrosuper.

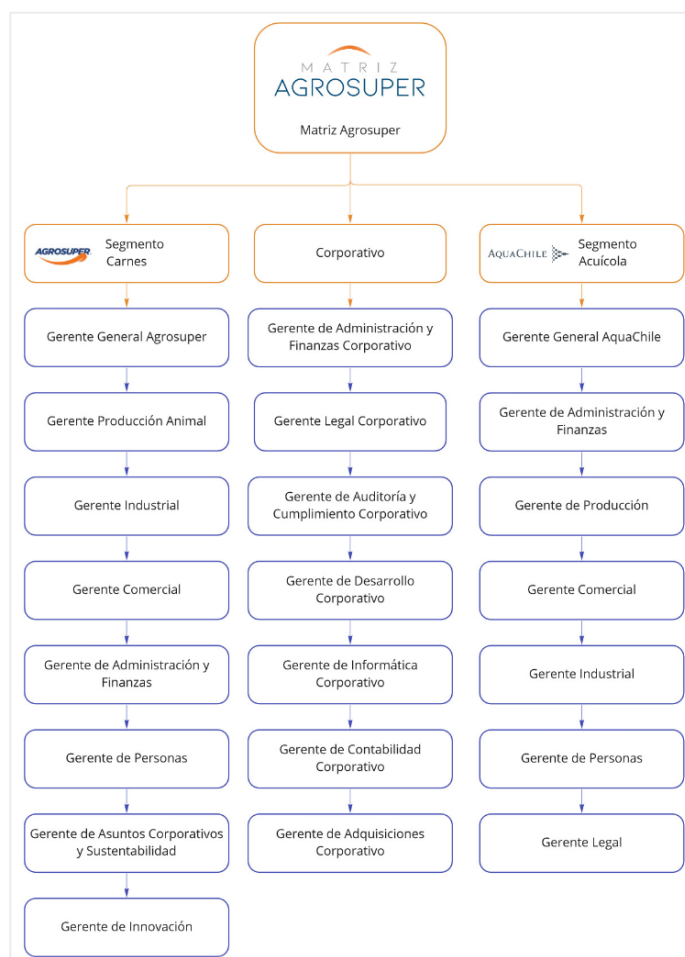


Imagen 1: Organigrama Agrosuper.

Fuente: Adaptación de organigrama de Reporte Integrado Matriz Agrosuper 2021.

El segmento de carnes de Agrosuper tiene distintos canales de venta a nivel nacional, los cuales se explican a continuación:

1. Canal Tradicional: corresponde a los negocios de barrio, minimarkets, carnicerías, entre otros, los cuales venden los productos de Agrosuper a un consumidor final.
2. Canal Supermercados: corresponde a los supermercados, ya sean de regiones o grandes cadenas donde las más reconocidas en Chile son Jumbo, Walmart, etc. Esta agrupación vende a un consumidor final de forma más masiva que el canal Tradicional.
3. Canal Industrial: corresponde a las empresas u organizaciones que utilizan los productos Agrosuper como materias primas y con estas crean otros productos para la venta.
4. Canal Food Service: corresponde a los locales de venta de comida preparada, como restaurantes, comida rápida, entre otros, los cuales ofrecen un servicio de comida preparada.

Cada uno de estos canales de ventas, generan un nivel distinto de ingresos para la empresa. En particular, el canal supermercados es el que genera mayor cantidad de ventas. Como se puede apreciar en el gráfico siguiente:

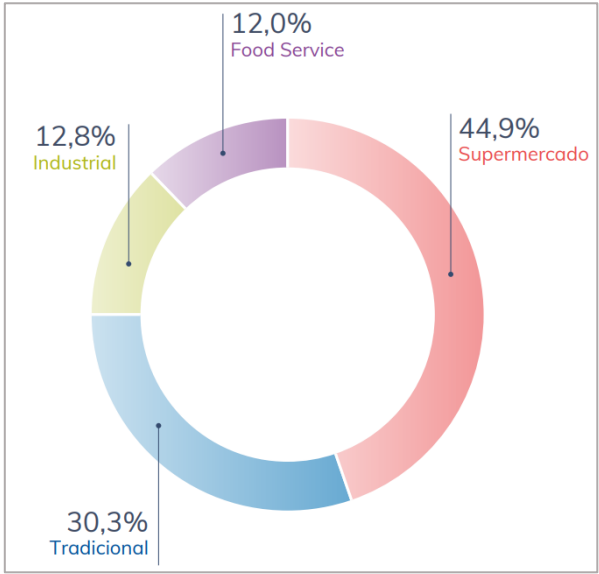


Gráfico 3: Ventas Nacionales por canal de Distribución.
Fuente: Reporte Integrado Matriz Agrosuper 2021.

Del gráfico 3 se desprende que el canal de ventas Supermercado alcanza el 44,9% de las ventas totales de Agrosuper, seguido del canal Tradicional con un 30,3% en el año 2021, esto muestra que el canal supermercado es el más relevante para Agrosuper por su alto nivel de ventas. Estas ventas son analizadas también, por sector, los cuales son pollo, cerdo, pavo, salmón, entre otros. En el gráfico 4 se puede ver como se distribuyen las ventas.

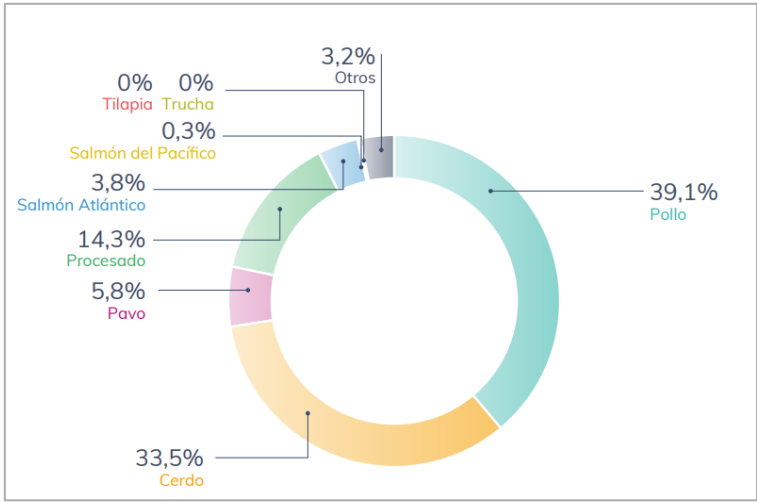


Gráfico 4: Ventas nacionales por negocio.
Fuente: Agrosuper.

Como se poder ver en el gráfico 4 el pollo es el sector de mayor venta en el año 2021 con un 39,9% de estas, seguido del cerdo con un 33,5% en el país.

1.2 Descripción del proyecto y justificación

Agrosuper está continuamente buscando maneras de poder mejorar y lograr cumplir su propósito. Para esto define metas de la participación que quiere tener en el mercado, esta tarea es realizada el área de Inteligencia del Consumidor (IC).

IC define las metas de participación considerando el comportamiento del mercado en los años anteriores. El cumplimiento o no cumplimiento de estas metas están asociadas a las distintas áreas de Agrosuper donde cada trabajador tiene como objetivo llegar a estas. Por lo que es muy importante poder definir metas posibles de cumplir, pero sin dejar de hacer todos los esfuerzos por aumentar la participación de mercado.

En el año 2021 se fijaron metas para pollo de supermercados según las distintas cadenas y total nacional que incluye todos los grupos de cadenas. Como ejemplo se muestran la imagen 2, donde es posible ver los resultados obtenidos de participación por cadena y el total nacional, en la segunda columna se puede ver las metas fijadas para cada grupo y finalmente el cumplimiento de estas.

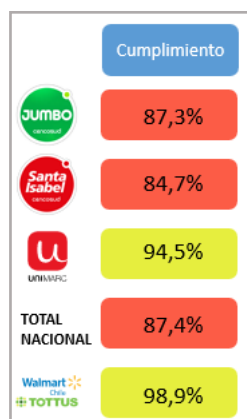


Imagen 1: Metas y resultados de la Participación de Mercado de Agrosuper en Pollo de Supermercados.
Fuente: Elaboración propia.

En la tercera columna de la imagen 2 se puede ver como fue el cumplimiento y que se considera malo (rojo), regular (amarillo) o bueno (verde). En detalle, se considera malo un cumplimiento menor al 90%, regular un cumplimiento entre 90% y 100% y bueno mayor o igual al 100%. Por ejemplo, se puede ver que se obtuvo un cumplimiento del 87,4% de la meta, este está en rojo, ya que es considerado un mal resultado.

Además, no es solo el hecho de que las metas no se cumplan, sino que también es necesario considerar como ha ido la participación de mercado de Agrosuper a lo largo del tiempo.

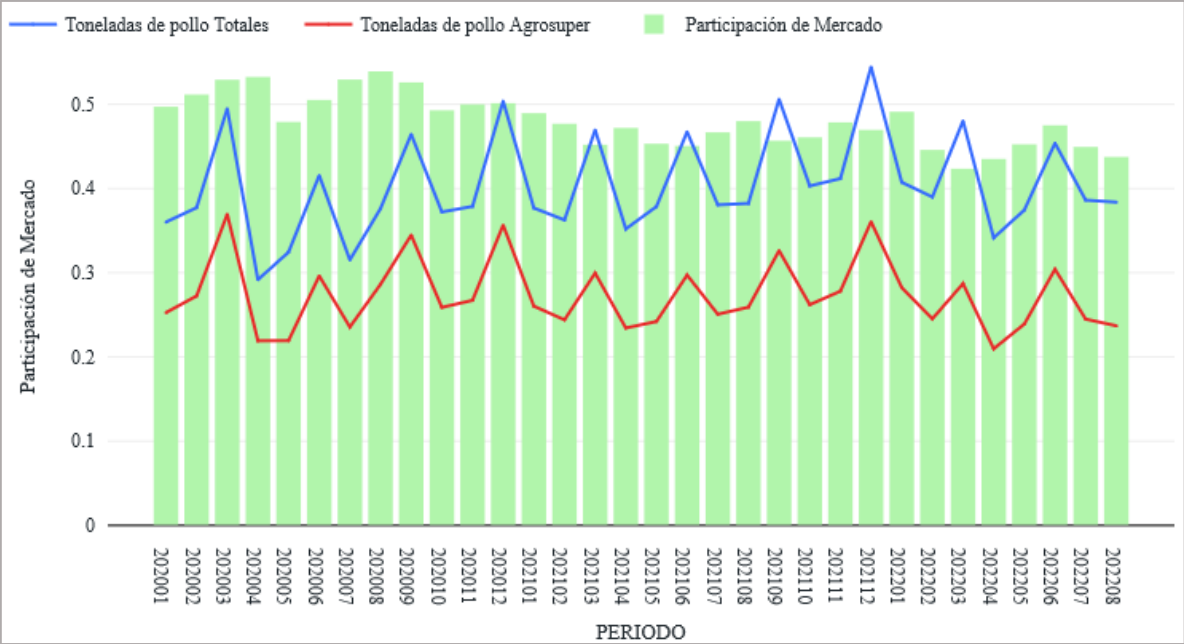


Gráfico 5: Participación de mercado de pollo de supermercados Agrosuper.

Las barras verdes del gráfico representan el comportamiento de la participación de mercado de Agrosuper, se puede ver que en general tiene alta variabilidad, pero siempre se mantiene sobre el 40% (0,4), también, se puede apreciar claramente la tendencia a la baja desde enero 2020 a agosto 2022. En anexo 1, se puede obtener el detalle del comportamiento de la participación de mercado por cadena.

Es importante considerar que hasta el momento el área de inteligencia del consumidor no tiene una forma estándar de generar la metas, si no que va cambiando, dependiendo de la persona de las herramientas y conocimientos que tengan en el área. Esto implica que pueden cambiar las predicciones dependiendo de las personas y solo realizan un modelo general, es decir, del total pollo de todas las cadenas en Chile.

2. Objetivos

2.1 Objetivo General

Construir un modelo de estimación de demanda de mercado para pollo en supermercados, con el fin de definir metas de participación de mercados ambiciosas pero alcanzables.

2.2 Objetivo Específicos

1. Caracterizar la demanda actual y las tendencias de ventas tanto de Agrosuper como de la competencia, con el fin de incluir estos aprendizajes en el modelo de estimación y también, entender el comportamiento del consumidor.
2. Identificar y analizar las variables relacionadas con la estimación de demanda para generar un tablón con las variables útiles para el modelo, ya sean internas o externas.
3. Desarrollar distintos modelos estadísticos para estimar demanda y evaluar el resultado según distintas métricas, con la finalidad de definir qué modelo responde de mejor manera a la problemática.
4. Aportar información estadística que permita generar estrategias para la aumentar o mantener la participación de mercado de Agrosuper

3. Marco Teórico

3.1 Conceptos claves

Una serie temporal es un conjunto de observaciones ordenadas y equidistantes cronológicamente sobre alguna característica o sobre varias de algún dato observable en distintos momentos del tiempo. Esta puede tener diferentes características, como la estacionalidad la cual refiere al movimiento periódico que se produce durante un tiempo corto y conocido. Además, pueden tener una tendencia que es el comportamiento que predomina en la serie, la cual son ciclos dentro que marcan oscilaciones dentro de esta y puede tener un comportamiento difícil de explicar y que sea duradero en el tiempo. Como última característica una serie temporal puede tener aleatoriedades las cuales son erráticas, obedecen a diversos motivos y son poco duraderas.

3.2 Modelos

Se buscan modelos aplicables al problema y que sean capaces de dar una respuesta a este sin importar lo complejo del modelo. Se eligen 3 tipos de modelos, ARIMA, regresiones lineales y alisamiento exponencial.

3.2.1 Modelo ARIMA (AutoRegressive Integrated Moving Average)

El Modelo de ARIMA, viene de la combinación del modelo autorregresivo (AR) y el de medias móviles (MA). Para analizar y predecir series de tiempo se puede utilizar los modelos autorregresivos, como AR.

El modelo autorregresivo AR(p) de orden p, se define de la siguiente forma:

Ecuación 1: Modelo Autorregresivo.

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \epsilon_t$$

Donde los parámetros ϕ_i ($i=1,2,3, \dots, p$) y δ son constantes y ϵ_t están no correlacionada con los X_{t-i} .

El modelo de medias móviles de orden q, MA(q), se define de la siguiente manera:

Ecuación 2: Modelo medias móviles.

$$X_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}$$

Donde ϵ_t tiene análogo significado que en el modelo AR, y μ y θ_i ($i=1,2,3, \dots, q$) son constantes. El modelo ARMA es una suma de los modelos AR y MA.

Para obedecer a un modelo ARIMA la serie original X_t es homogénea, de orden d y no estacionaria, entonces debe cumplir que:

Ecuación 3: Condición de modelo ARIMA. Fuente: Libro econometría y series de predicción.

$$\Delta^d X_t = Z_t, t = 1, 2, \dots, T$$

Ahora si admite que Z_t obedece a un proceso ARMA(p,q), entonces se puede decir que es un proceso ARIMA(p,d,q) ya que se convierte en un proceso integrado.

3.2.2 Modelo de Regresión Lineal

El modelo de regresión lineal simple es un proceso experimental donde intervienen dos variables, por un lado, la variable dependiente Y, la cual no es controlada por el experimento y depende de la variable X, la cual es controlada por el experimento, por lo tanto, no tiene la característica de ser aleatoria. Si se asume que todas las medias de las variables dependientes Y_i caen en una recta, se puede representar de la siguiente manera:

Ecuación 4: Regresión Lineal. Fuente: Elaboración propia.

$$Y_i = \alpha + \beta X_i + E_i$$

Donde E_i es el error de la regresión lineal, el cual debe tener una media cero.

3.2.3 Modelo Alisamiento exponencial

Este modelo se puede caracterizar entre modelos estacionales o no estacionales, dependiendo si la se definió una periodicidad en el conjunto de datos. Existen varios tipos de alisamiento exponencial:

1. **Simple:** se utiliza más para ventas más estables en las cuales no hay tendencia ni elasticidad.
2. **Tendencia lineal de Holt:** es más apropiado cuando se tiene alguna tendencia lineal, pero no hay estacionalidad en la serie de tiempo.
3. **Tendencia lineal de Brown:** el cual es un caso más específico de la tendencia lineal de Holt, variando principalmente en forma del alisamiento exponencial. Este modelo se asimila más a un modelo ARIMA.
4. **Simple estacional:** la cual es más apropiada en una serie de tiempo sin tendencia, pero con una clara estacionalidad, que es constante en el tiempo.
5. Existen algunos tipos más de alisamiento exponencial, los cuales tienen una variación en los niveles de alisamiento y si presentan o no estacionalidad y tendencia, lo que va variando como se abordan las series de tiempo.

3.3 Métricas de evaluación

Se establecen métricas aplicables a los modelos anteriores con la finalidad de poder comparar los resultados obtenidos en cada caso.

3.3.1 MAPE

La métrica MAPE o error porcentual absoluto medio, es una métrica para calcular el error de un modelo de pronóstico. Se calcula de la siguiente manera:

Ecuación 5: Cálculo de MAPE. Fuente: Recuperado de Internet.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Siendo y_i el resultado real esperado e \hat{y}_i el resultado obtenido con el modelo aplicado. Se calcula obteniendo la diferencia entre estos dos valores y dividiéndose por el primero, a esto se le calcula el valor absoluto y se va sumando con cada una de las observaciones obtenidas. Al calcular este error de esta forma, genera que no impacten demasiado los valores atípicos a diferencia del error MSE, pero esto no quiere decir que esta métrica sea mejor, si no que depende de lo que se esté buscando con el modelo. Hay veces que se necesita que el error castigue más por estos valores atípicos y otras veces es preferible que este no lo haga. Dependerá del tipo de modelo y lo que se esté buscando.

3.3.2 MSE

La métrica MSE o error cuadrático medio es una métrica que evalúa los errores del modelo y se calcula de la siguiente manera.

Ecuación 6: Cálculo de MSE.

$$MSE = \frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Siendo y_i el resultado real esperado e \hat{y}_i el resultado obtenido con el modelo aplicado. Se calcula la diferencia entre los dos la cual se potencia al cuadrado obteniendo siempre valores positivos, los cuales se van sumando, considerando todos los resultados del modelo. Es menos el error cuando este indicador es más pequeño, es decir, un modelo perfecto tendrá un MSE = 0. Es útil cuando existen valores extraños, que están muy alejados de los esperado. Pero en el caso de tener muchos errores mayores a 1 esta métrica podría sobre estimar el error final, y en el caso contrario subestimarlos.

3.3.3 AIC y BIC

AIC es el criterio de información de Akaike y BIC es el criterio de información bayesiano, son métricas para seleccionar el modelo de mejor comportamiento. El AIC se calcula de esta forma:

Ecuación 7: Formula de AIC.

$$AIC = 2k - 2\ln(\widehat{L})$$

Donde k es el número de parámetros y \widehat{L} es el máximo de la función de verosimilitud del modelo. La métrica BIC, se calcula de la siguiente forma:

Ecuación 8: Formula de BIC.

$$BIC = \ln(n)k - 2\ln(\widehat{L})$$

Donde k y \widehat{L} es lo mismo que en la métrica AIC y n en este caso es número de la muestra. El BIC a diferencia del AIC, penaliza por cantidad y parámetros utilizados.

Cabe mencionar que la métrica AIC y BIC, solo se puede calcular en modelos que utilicen variables extras a la variable independiente. De los tres tipos de modelos, se puede calcular para el ARIMA y las regresiones lineales.

4. Metodología

El proyecto se realizará utilizando la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining) la cual considera 6 etapas en la construcción del proyecto, partiendo por el entendimiento del negocio, luego la compresión de los datos, preparación de los datos, modelamiento e implementación.

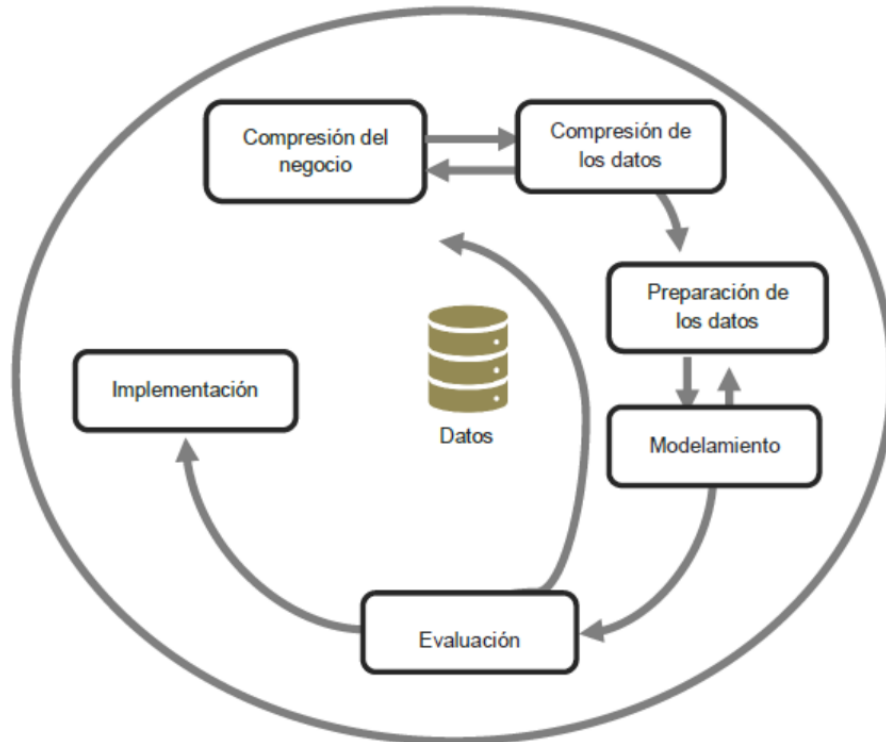


Imagen 2: Metodología CRISP-DM.
Fuente: IBM.

Se utilizará esta metodología debido a que a diferencia de otras como KDD, CRISPM-DM contempla esta primera parte de entendimiento del negocio y además permite ir obteniendo retroalimentación en cada paso, por lo que se puede volver a los pasos anteriores e ir iterando la solución del problema.

4.1 Comprensión del negocio:

Esta es la primera fase del proyecto, consiste en estudiar y analizar este. Es la fase más importante ya que si se realiza mal alguno de los procesos de esta parte como la fijación de objetivos, comprensión del problema o planificación del proyecto, el algoritmo utilizado para resolver la problemática entregará resultados erróneos, teniendo que ser modificada o fracasando el proyecto.

Las principales tareas por definir en esta fase son:

1. Determinar objetivos del negocio: consiste en estudiar la problemática, plantear los objetivos y los criterios de éxito de cada uno de estos.
2. Verificar la situación actual: es realizar la comprobación del estado actual, es decir, antes de ejecutar el proyecto. Es importante tener claridad respecto de tres ítems:
 1. Conocimiento del problema
 2. Datos requeridos para el análisis
 3. Relación entre costo/beneficio de la realización del proyecto
3. Producir plan de proyecto: por último, se realiza un plan que describa metodología de desarrollo, técnicas y plazos para cada una de las etapas.

4.2 Comprensión de los datos:

Corresponde a la etapa de entender los datos disponibles con que se cuenta y cuáles son las limitaciones de estos, con el fin de familiarizarse y elaborar las directrices necesarias para el procesamiento de estos en los modelos a realizar. Para dar por completa esta es necesario cumplir con algunas actividades:

1. Recopilación de datos: consiste en buscar todas las variables disponibles para el modelo. Entender con que se cuenta y que no es posible de tener para el desarrollo del proyecto.
2. Descripción de los datos: consiste en definir que es cada variable, como está calculada y cuál es la relación con el proyecto.
3. Exploración de los datos: consiste en realizar un análisis de las variables, si es útil como está o es necesario modificarla para que sea más útil para el modelo.
4. Calidad de datos: analizar la calidad de los datos, fijándose en si hay elementos fuera de rango o elementos que no estén dentro de lo esperado.

4.3 Análisis y preparación de datos

Una vez terminada la comprensión de los datos, comienza la etapa de preparación de estos, la cual puede ir desde la limpieza de la data hasta la definición de nuevas variables. Esta fase es importante realizarla bien, debido a que es la base para realizar un buen modelamiento. Considerando esto, las actividades a realizar en esta etapa son:

1. Selección de datos: definir qué datos se utilizarán para el modelo, ya sean externos o internos de la organización, desde el conjunto de datos recolectados en la parte anterior.
2. Limpieza de datos: realizar la limpieza de estos datos, eliminando variables correlacionadas, reestructurarlas, editar el largo o que no sean explicativas del modelo y también, viendo elementos extraños que no tengan sentido lógico con las información vista.
3. Estructuración de los datos: consiste en realizar la creación de nuevas variables en caso de ser necesario, con los datos que se tienen, generar datos que se relacionen con otros, pero aporten información más explícita sobre una característica en particular.
4. Integración de los datos: consiste en unir los datos obtenidos para dejar solo un registro por una categoría definida.

5. Formateo de datos: Consiste en cambiar los datos, eliminando caracteres complicados u otros problemas, o puede ser reestructurar completamente la información en beneficio del modelo.

4.4 Modelamiento

En esta etapa se realiza la elección y ejecución de la técnica de modelamiento elegida, para esto es necesario tener en cuenta lo hecho anteriormente en la etapa de preparación de datos también el conocimiento del negocio para así aplicar la técnica más adecuada, además, de manejar la técnica de modelamiento a aplicar.

Esta etapa incluye los siguientes pasos:

1. Selección de técnica de modelado: este paso apunta a la elección de una técnica de modelamiento. Cabe destacar que, si el modelo se basa en suposiciones, estas deben ser registradas.
2. Generación plan de prueba: Después de construir un modelo, este debe ser testeado para comprobar la calidad y validez de este. Es común separar la data en conjuntos de entrenamiento y prueba, en el conjunto de entrenamiento se construye el modelo y en el de pruebas se estima su calidad. Con esto se mide que tan bien el modelo puede predecir la historia antes de decidir utilizar para predecir el futuro.
3. Construcción del modelo: Tras la etapa de pruebas, se ejecuta la herramienta de modelamiento en el conjunto de datos correspondiente para crear uno o varios modelos.
4. Evaluación el modelo: Finalmente, se interpretan los modelos y se trabaja con otros analistas para evaluar los resultados en el contexto empresarial, en algunos casos estos analistas son incluidos desde el principio con el fin de detectar posibles errores

4.5 Evaluación

Es importan verificar que el modelo sea capaz de responder a las preguntas que tenía la organización al respecto. Por lo que es necesarios, evaluar la situación y si es capaz de realizar eso o se deben generar cambios para que el modelo sea capaz de dar las respuestas necesarias. Esta etapa debe seguir los siguientes pasos:

1. Evaluación de resultados: Esta etapa evalúa que el modelo cumpla con los objetivos empresariales, otra opción posible en este paso es probar el modelo en ambientes reales siempre que el tiempo o el presupuesto lo permitan. En este paso también se evalúa si los resultados cumplen con los criterios de éxito empresarial.
2. Retroalimentación y siguientes pasos: En esta última etapa se evalúa si fueron considerados todos los criterios relevantes para el proyecto o si alguno se pasó por alto, a su vez se decide si se da fin al proyecto o se inician nuevas iteraciones de este.

4.6 Plan de Implementación piloto

Una vez elegido el modelo final, se realizará una documentación donde se establecerá como implementar el modelo y dejando las condiciones necesarias para realizar un piloto y evaluar el funcionamiento de este.

Solo se realizará un plan de implementación debido a los tiempos del trabajo de título, en otros casos esta etapa debería ser la implementación del proyecto con sus etapas respectivas.

5. Alcances

El modelo de demanda de mercado incluirá solo productos de pollo de supermercados, ya que es este el mercado que se requiere analizar, dada la gran cantidad de ventas de esta carne en los supermercados de Chile.

El modelo se realizará con una apertura por cadena de supermercado, en específico, se analizará la cadena Jumbo, Santa Isabel, Unimarc y Mayorista 10. También, se incluirá la categoría "otros" en la cual están incluidas principalmente Walmart y Tottus, entre otros supermercados más pequeños, esto debido a que la información con la que se trabaja no tiene desagregada los datos de las otras cadenas.

No se realizará la implementación del modelo como tal, pero se dejará un plan para la implementación de un piloto del modelo, con la finalidad de evaluar su funcionamiento.

Para la construcción y ejecución del modelo se utilizará principalmente los recursos proporcionados por el lenguaje de programación R y SQL en la plataforma Databricks que es lo que utiliza Agrosuper donde se está desarrollando el proyecto.

6. Resultados

6.1 Entendimiento del negocio

Uno de los principales resultados de esta parte de la metodología es el proceso de definición de metas de participación de mercado. El cual, parte en el área de inteligencia del consumidor, donde realizan un proceso de estimación que maneje el principal encargado de este, luego estas son aprobadas por el subgerente y gerente de inteligencia del consumidor, para entregárselas a la unidad de comercial correspondiente que las analiza y ajusta según distintas variables que ellos tengan, como producción, ventas, precios, etc.



Imagen 4: Proceso de metas de definición de metas en Agrosuper

Este proceso mientras que si no hay un buen trabajo desde el inicio, donde se estima la demanda, el resultado final que son las metas no serán realistas y acordes con el mercado.

6.2 Comprensión y preparación de los datos

Primero en la recopilación de datos se obtuvieron distintas tablas, las cuales se pueden separar en 2 grandes grupos, internos y externos. De los datos internos se recopilan 4 tablas:

1. Clientes de Agrosuper: se filtra para obtener los datos necesarios, como el código interno de local, a que canal de distribución pertenece, es decir, si es de supermercado, canal tradicional, etc. Además, se recopila la cadena a la que pertenece cada local y cuál es su ubicación geográfica, estos pasos se realizan con la intención de entender que datos están disponibles para utilizar en el modelamiento.
2. Materiales: tiene la información de todos los productos con los que trabaja Agrosuper. Se recopiló principalmente a que sector pertenece cada producto, teniendo los sectores 01, 02 y 03 los cuales son pollo, cerdo y pavo respectivamente, existen otros sectores, pero principalmente se trabajará con estos tres. Ya que, se tiene la hipótesis que estos bienes pueden ser considerados como complementarios al producto pollo.
3. Pedidos: tiene información de los pedidos realizados por los distintos clientes de Agrosuper, mostrando la cantidad de kilos pedidos, el valor unitario de los productos, el valor total y el tipo de producto solicitado. A esta tabla se le realizó un filtro para extraer solo los clientes considerados como supermercados en la tabla de clientes y los productos de la tabla de materiales, además se le realizó una agrupación por mes y por cadena, obteniendo finalmente una tabla con las toneladas de pollo pedidas por cada cadena y por sector (Pollo, Cerdo, Pavo) y los precios de venta de Agrosuper a sus clientes ponderados por toneladas de los productos.
4. Facturación: es una tabla con columnas como la fecha de facturación, los kilos de cada producto por cliente, además incluye las devoluciones realizadas por los supermercados a la empresa. Estas devoluciones pueden ser por distintos motivos, entre ellos, productos en mal estado, productos cerca de caducar, por no venta de

estos, etc. Finalmente se realiza la misma agrupación que para la tabla de pedidos quedando organizada de igual forma.

En cuanto a los datos externos se recopilaron de dos tipos:

1. Datos Nielsen: son proporcionados por Nielsen, la cual es una empresa internacional que se dedica a entregar información, datos, realizar estudios de mercado, entre otras actividades. En esta información viene los kilos vendidos en cada cadena por producto y semanal, teniendo la marca de estos que puede pertenecer a Agrosuper o a la competencia, el precio ponderado de estos productos y el valor total por la cantidad de producto. Se realizaron algunos ajustes para obtener los datos mencionados anteriormente de forma mensual y a nivel total nacional y por cadena.
2. IPC: tabla con los valores mensuales desde el 2019 a la fecha.
3. IMACEC: Índice Mensual de Actividad Económica, se recopiló la información de este índice desde enero del 2019 hasta agosto 2022.
4. UF: Unidad de Fomento, es una medida monetaria en Chile. Se recopila el valor en pesos chilenos de esta, desde el 2019 a agosto 2022
5. M1: Agregado monetario, que representa el valor más líquido del dinero, billetes y monedas. Además, se incluyen la cantidad de
6. Feriados: cantidad de días feriados por mes en Chile y si estos son o no irrenunciables.

Finalmente, se realiza un cruce final de las tablas, para tener todos los datos unificados y proceder con la siguiente etapa del proyecto.

6.3 Análisis de datos

6.3.1 Análisis Univariado

Para iniciar esta etapa se realizaron distintos gráficos para poder entender el comportamiento de las variables. El primer gráfico realizado es de las ventas por cadena, el cual se puede ver a continuación.

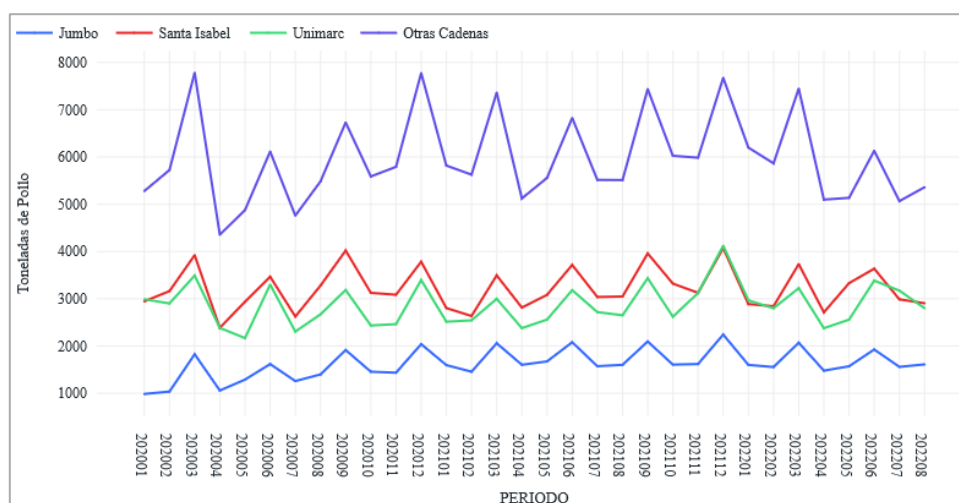


Gráfico 6: Ventas por cadena desde enero 2020 a agosto 2022 por toneladas de pollo.

De este gráfico se puede ver que existen distintos peaks los cuales tienen una periodicidad de 3 meses, marzo, junio, septiembre y diciembre. Esto es debido a la consideración de meses definidos por Nielsen (proveedor de estos datos), ya que los meses mencionados anteriormente tienen 5 semanas a diferencia del resto que solo se consideran 4 semanas para su construcción.

También, se puede destacar que la cadena con menor cantidad de ventas es Jumbo, que se puede apreciar en color azul y en la parte más baja del gráfico, esto no era un resultado esperado, principalmente debido a que Jumbo es uno de los clientes más importantes de Agrosuper. Además, el peak más grande para la mayoría de las cadenas ocurre en el periodo 202003, es decir en marzo del 2020, esto puede relacionarse con el inicio de la pandemia en Chile, ya que en esos momentos estuvieron los supermercados con un gran flujo de Clientes.

La venta puede guardar una relación con la variable de precios, es por esto que es importante analizar cómo se han comportado estos durante el periodo de datos (enero 2020 a mayo 2022).

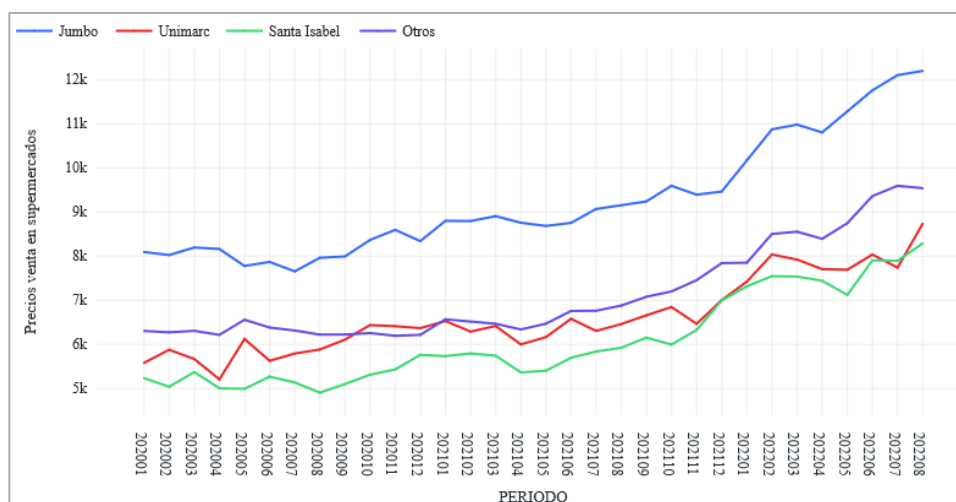


Gráfico 7: Precios del mercado por cadena desde 2020 al 2022 por kilogramo de pollo.

Se destaca del gráfico 7 que Jumbo es la cadena que tiene el valor de los productos de pollo más alto a nivel mercado, lo que se condice con que sea el de menor cantidad de ventas, luego viene Unimarc con los precios más altos el cual también es lo que se esperaba debido a la menor cantidad de ventas y luego vienen Santa Isabel y otros con los precios más bajos en estos productos.

Otro análisis importante es el comportamiento de la facturación y pedidos de Agrosuper, ya que se espera que estos guarden relación con el comportamiento de las ventas en los supermercados, a continuación, se puede visualizar un gráfico con las toneladas de pollo que factura Agrosuper en la fecha correspondiente.

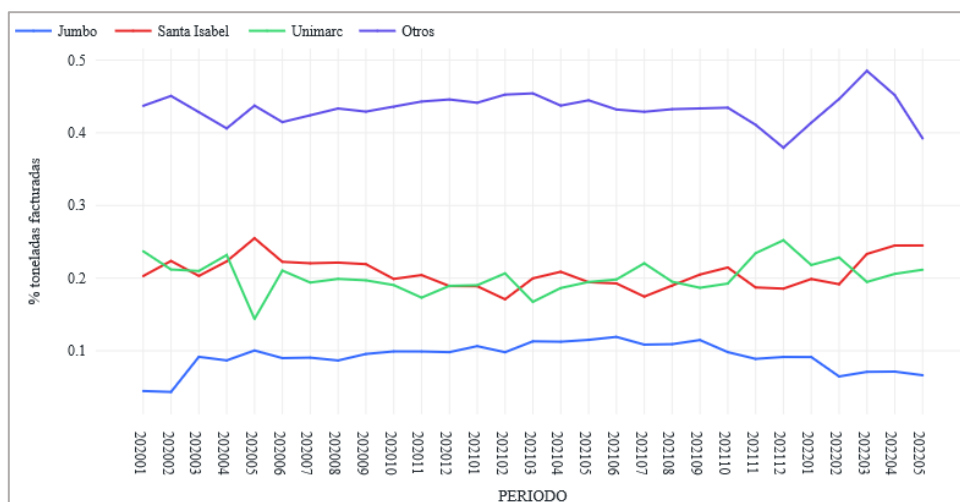


Gráfico 8: Facturación por cadena desde 2020 al 2022 por toneladas de pollo.

Acá no se ve la misma periodicidad que en el gráfico de ventas de los supermercados, pero si se puede apreciar los picos en septiembre y en diciembre. Lo cual es muy parecido a lo que pasa con la información de pedidos que se puede ver a continuación.

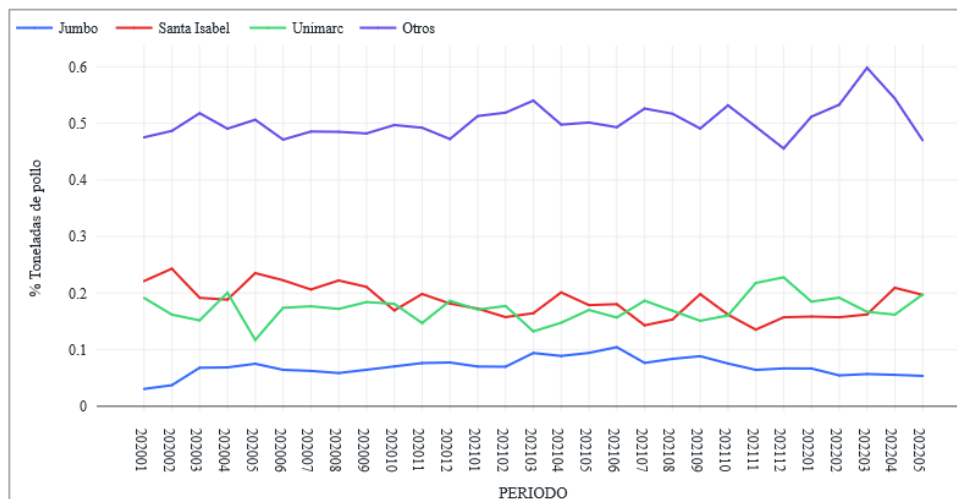


Gráfico 9: Pedidos por cadena desde 2020 al 2022 por toneladas de pollo.

Que no se vea la misma distribución depende por sobre todo de la definición de meses, ya que en la información de Nielsen se definen los meses con semanas cerradas, es decir, la semana 1 a la 4 del año pertenecen a enero, de la 5 a la 8 a febrero, etc. Y la información interna define los meses por calendario.

Además, se realiza un análisis de las variables externas (M1, IMACEC, IPC) para entender su comportamiento y cómo este puede posteriormente afectar la variable principal, toneladas de pollo por cadena.

Primeramente, se muestra la variación en el tiempo de los valores del M1, lo cual se aprecia en el siguiente gráfico.

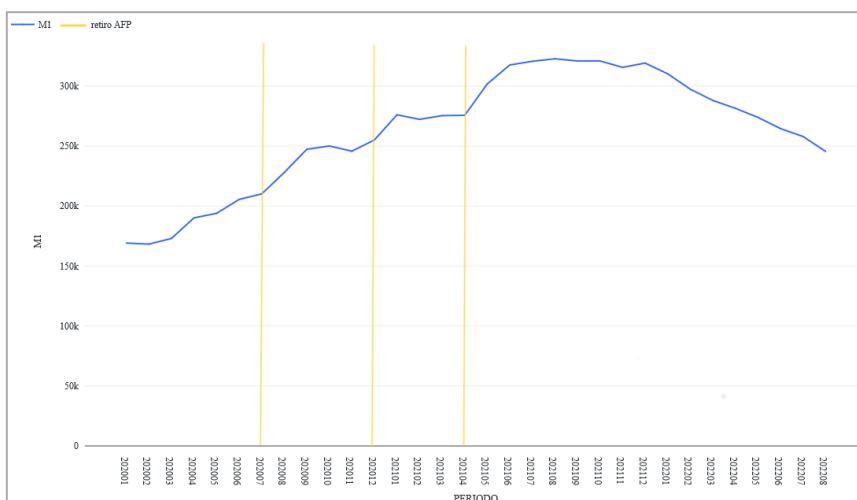


Gráfico 10: Variación del M1 (agregado monetario) en el tiempo y retiros de AFP.

Se puede apreciar la tendencia de crecimiento desde enero del 2020 hasta más o menos agosto del 2021, luego se ve un descenso de la curva. Las líneas de color amarillo representan la fecha donde se aprobaron retiros de fondo desde las AFP (Administradoras de Fondos de Pensiones), lo que generó después de cada uno un aumento en la curva que luego disminuía sin perder la tendencia creciente que tenía a la fecha.

Luego se analizó el comportamiento de la variable IMACEC, lo que se puede ver en el gráfico a continuación:

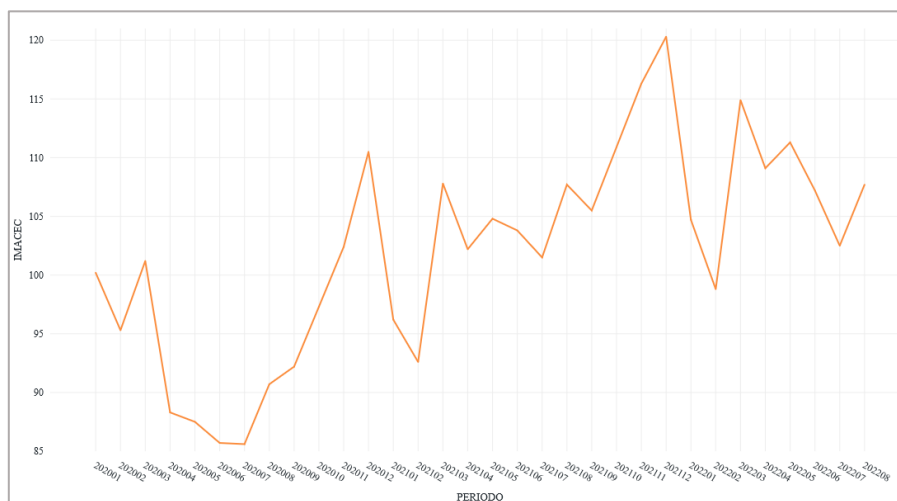


Gráfico 11: Variación del IMACEC en el tiempo.

Se puede interpretar del gráfico, que en los primeros meses del 2020 el IMACEC tiene una tendencia a la baja, llegando a su punto más bajo, entre junio y julio del 2020, luego de esto, la tendencia cambia a un aumento, pero con bastantes variaciones entremedios, generando peaks más altos en los meses de diciembre de cada año.

Además, se realiza un análisis del comportamiento del Índice de Precios al Consumidor (IPC) en el tiempo. Lo que se puede ver en el siguiente gráfico:

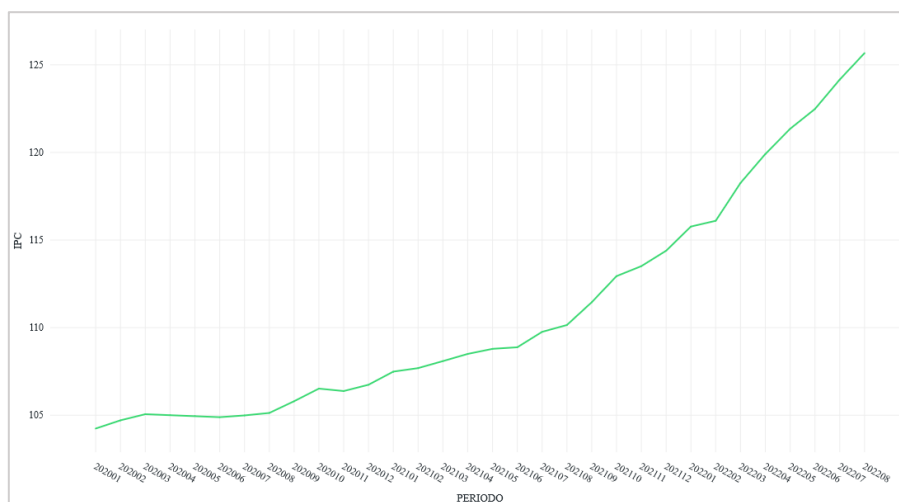


Gráfico 12: Comportamiento del IPC en el tiempo.

Acá se puede notar la tendencia de alza que tiene el IPC, teniendo un comportamiento parecido a la curva de la exponencial, es decir, su nivel de aumento es mayor mientras pasa el tiempo.

Finalmente, se realiza un entendimiento de los días feriados irrenunciables en Chile, notando que los meses con mayor cantidad de feriados irrenunciables, son abril, septiembre y diciembre, destacando por sobre todos diciembre del 2021, donde además de los feriados de todos los años, fue la elección presidencial, por lo que hay un feriado irrenunciable extra. Lo que se puede ver en el siguiente gráfico:

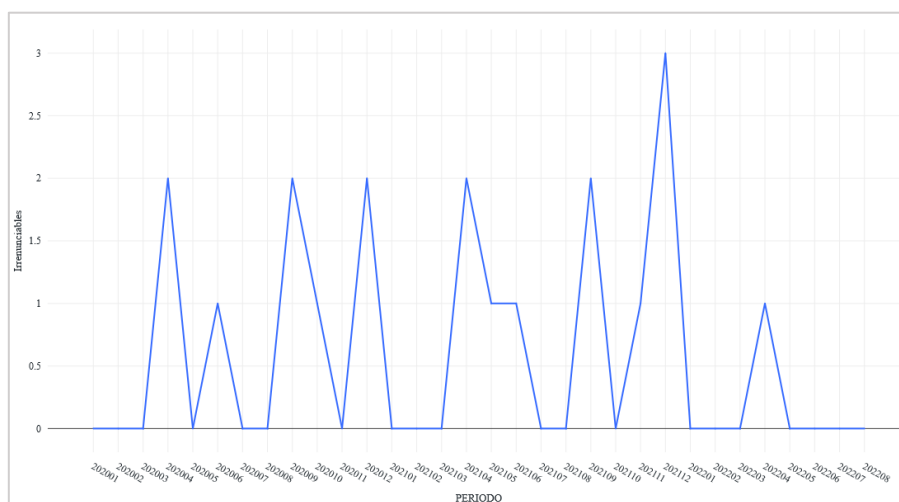


Gráfico 13: Cantidad de feriados en el tiempo.

También, se realiza el análisis de la demanda de pollo por cadena y total nacional, para ver si esta tiene alguna tendencia o periodicidad. En particular se realiza una descomposición de la serie de forma aditiva y multiplicativa, no encontrando diferencias significativas entre una y otra, los resultados de la descomposición se muestran a continuación:

1. Total nacional: la demanda tiene una tendencia de crecimiento y una componente estacional, que se puede ver en los gráficos siguiente donde se descompone la serie de forma multiplicativa y aditiva respectivamente.

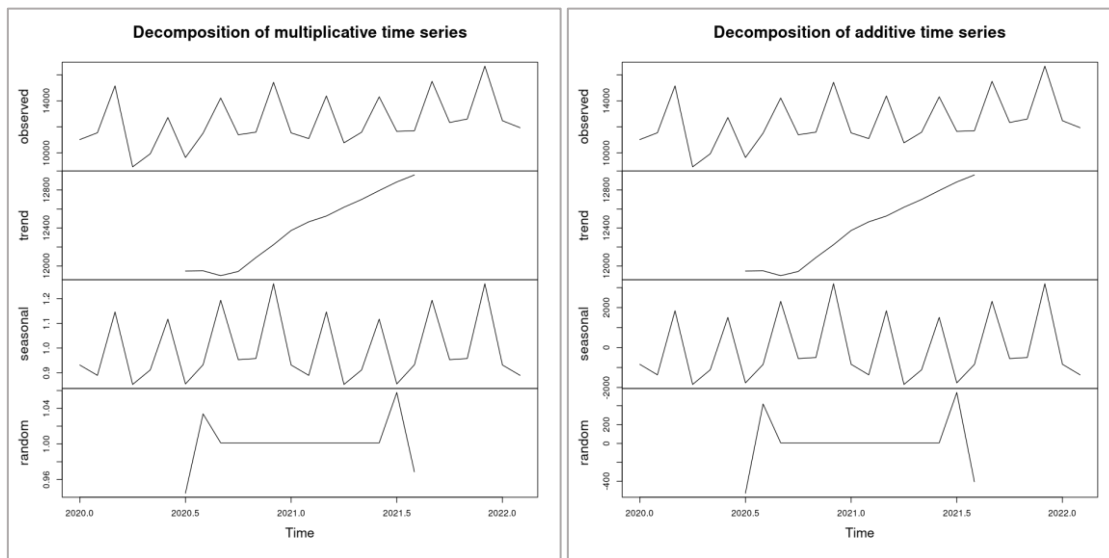


Gráfico 14: Descomposición multiplicativa y aditiva de la serie ventas del total nacional

2. Jumbo: Esta cadena también representa una tendencia positiva y la componente de estacionalidad

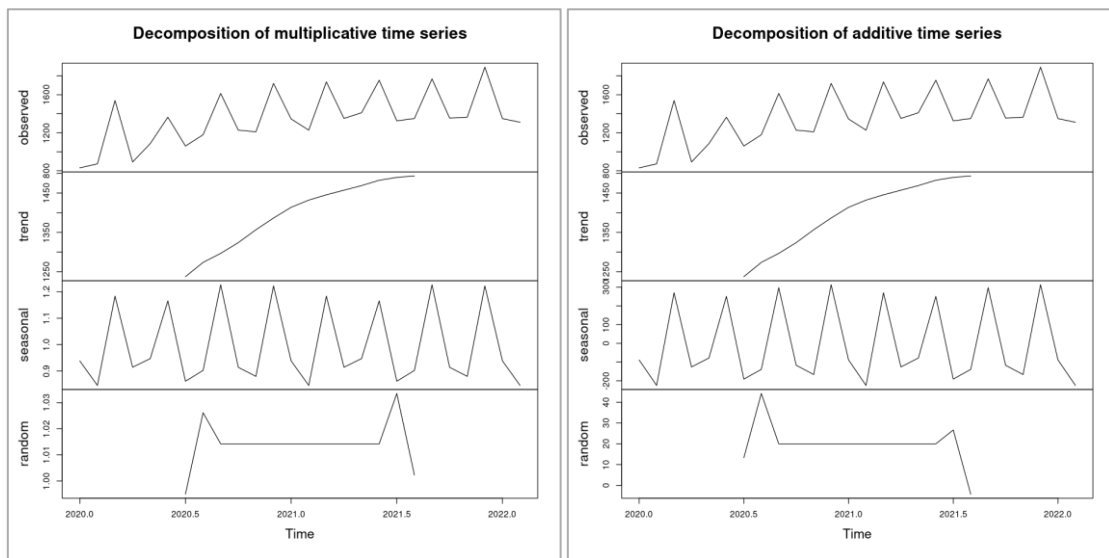


Gráfico 15: Descomposición multiplicativa y aditiva de la serie ventas de Jumbo.

3. Santa Isabel: Esta serie no tiene un comportamiento de crecimiento tan marcado como las anteriores, pero de igual manera se obtiene una tenencia marcada, lo mismo que en la estacionalidad.

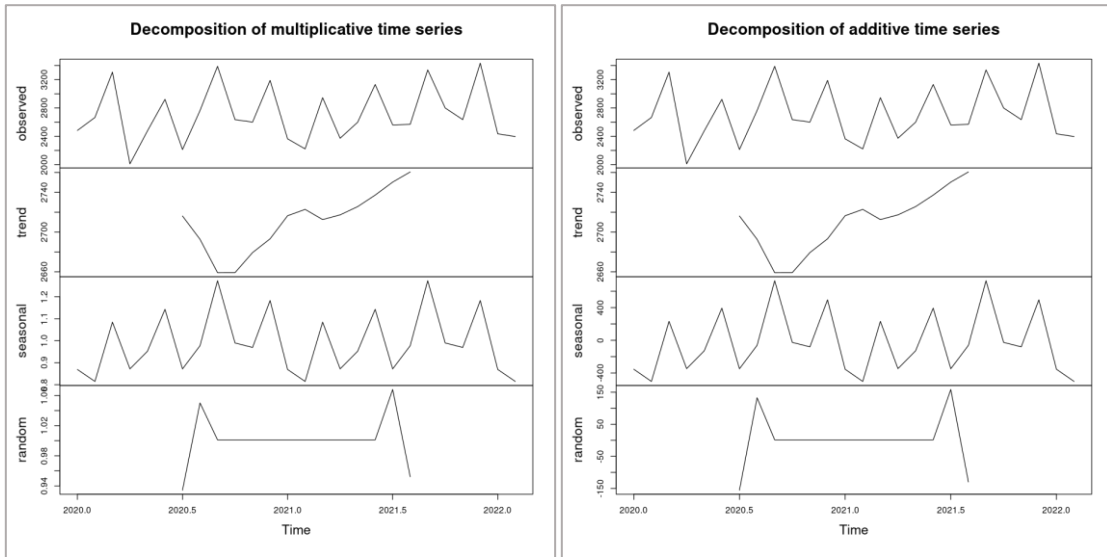


Gráfico 16: Descomposición multiplicativa y aditiva de la serie ventas de Santa Isabel.

4. Unimarc: Al igual que las series anteriores, esta tiene una tendencia y estacionalidad marcada, lo que se puede apreciar en los gráficos de más abajo.

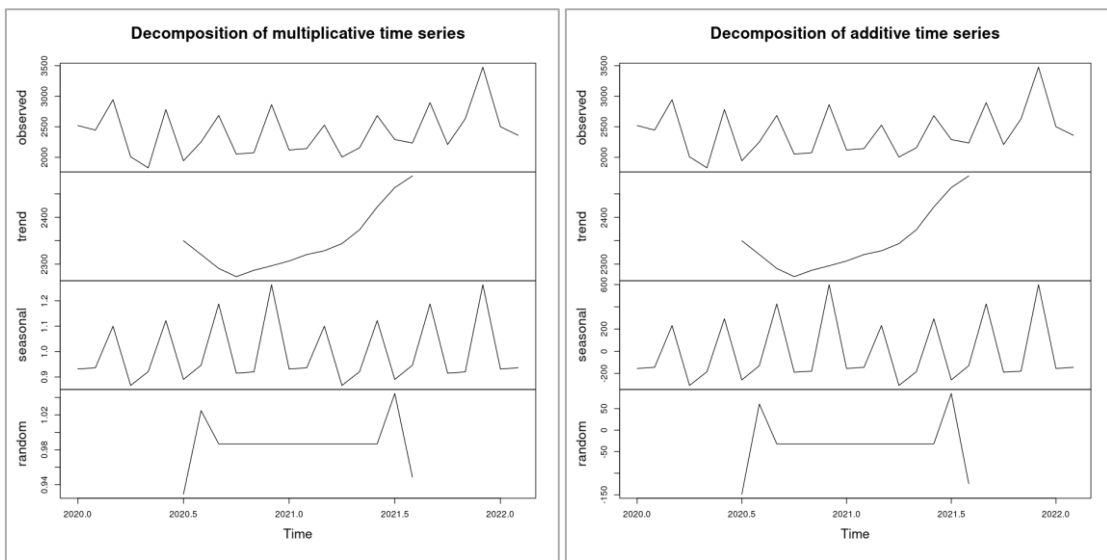


Gráfico 17: Descomposición multiplicativa y aditiva de la serie ventas de Unimarc

5. Otras Cadenas: Se puede ver una tendencia clara de alza, más marcada durante el año 2021 y una estacionalidad donde destacan los peaks que existen cada 3 meses.

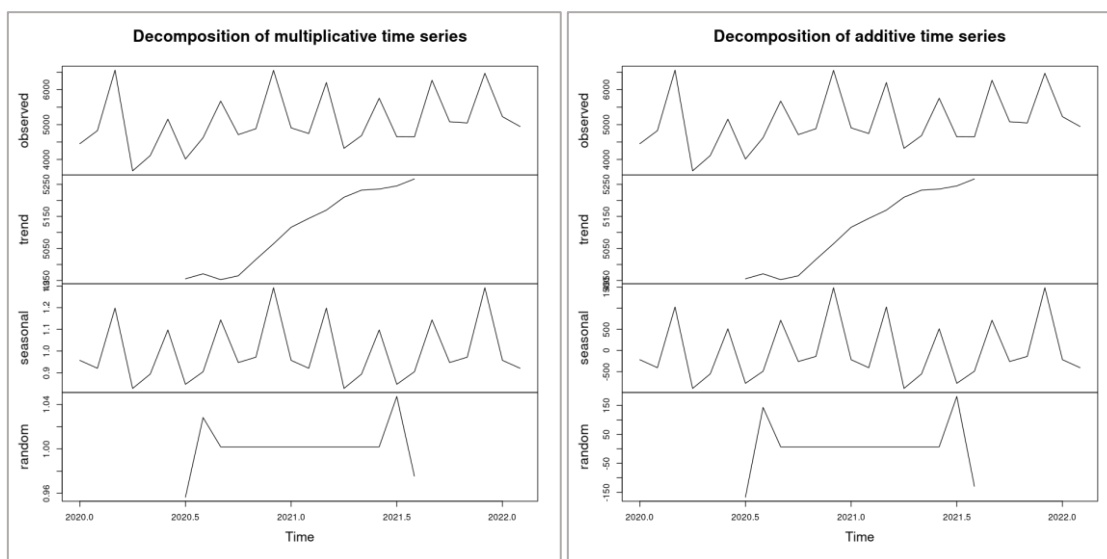


Gráfico 18: Descomposición multiplicativa y aditiva de la serie ventas de otras cadenas

6.3.2 Análisis Multivariado

Para el análisis multivariado se realizaron correlaciones entre las distintas variables mencionadas anteriormente, principalmente para entender como estas se relacionan con la variables objetivo y entre ellas. Ya que, si existe una correlación entre los datos y la variable respuesta, estos datos podrían ser útiles para predecir la demanda.

Se considera como una correlación alta valores menor o iguales a $-0,6$ o mayores e iguales a $0,6$.

Este análisis se separa por cadena y total nacional. A continuación, se muestra un gráfico con las correlaciones para el total nacional:

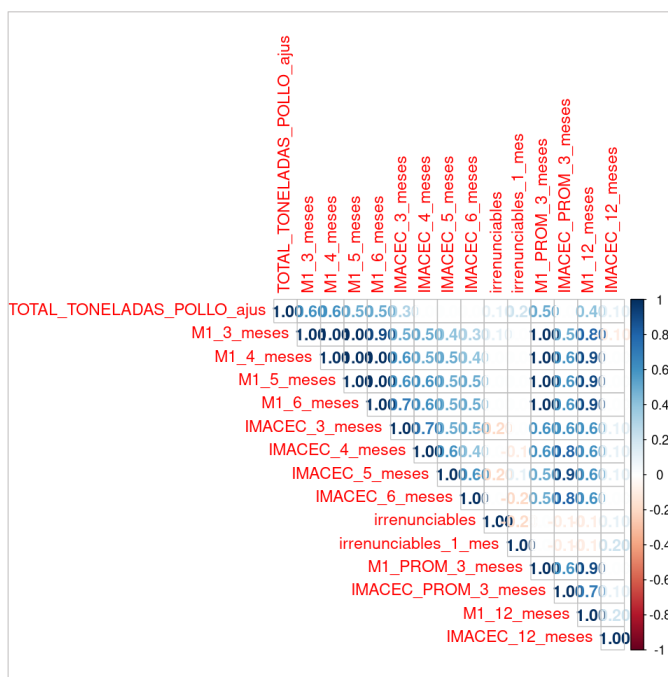


Gráfico 19: Correlaciones entre variables con total toneladas de pollo nivel nacional

En el gráfico 19 se puede ver las correlaciones de la variable objetivo con las otras variables, cabe mencionar que, los nombres de estas hacen referencia a la variable utilizada y la cantidad de meses de diferencia entre la dependiente y la independiente, por ejemplo, la variable M1_3_meses son los datos del M1 3 meses atrás y como estos se relacionan con las otras variables.

Las variables más correlacionadas con la variable dependiente son el M1 de 3 meses atrás y el M1 de 4 meses atrás. Además, estas variables están altamente correlacionadas entre ellas por lo que no se pueden utilizar juntas en un modelo, ya que esto produciría errores de colinealidad.

Se realiza este mismo análisis para cadenas Jumbo, Santa Isabel, Unimarc y Otros. Obteniendo como variables más correlacionadas las siguientes:

1. Jumbo: M1 de tres meses con un coeficiente de correlación de 0,6
2. Santa Isabel: feriados irrenunciables con 0,4 como coeficiente de correlación
3. Unimarc: feriados irrenunciables e IMACEC de tres meses atrás, con coeficiente de correlación de 0,4
4. Otros: feriados irrenunciables e IMACEC de tres meses atrás con 0,3 como coeficiente de correlación.

Se utilizan estas variables en particular para construir distintos modelos, pero de igual manera se prueban otras. Ya que los coeficientes de correlación no son muy altos. Los gráficos con los resultados para todas las variables se pueden ver en el anexo 2.

6.4 Modelamiento

Las técnicas de modelamiento a utilizar se condicen con lo que se expresó en el apartado de marco conceptual para el proyecto los cuales son los siguientes modelos: ARIMA, regresiones lineales y suavizamiento exponencial

6.4.1 Modelo ARIMA

Para trabajar con este modelo primero se definió la cantidad de meses que se tenían disponibles, estos son 32 meses que van desde enero del 2020 a agosto del 2022, se decidió utilizar 26 meses de entrenamiento y 6 meses de predicción.

Se realizaron aproximadamente 36 modelos y alrededor de 100 variaciones de cada uno. Se utilizaron distintas variables y la combinación de estas, ya sea el M1, IMACEC, UF, IPC, entre otras para armar los modelos.

Luego de tener todos los modelos realizados se elige el de mejor comportamiento, según las métricas definidas en el marco conceptual y las sugerencias de los involucrados, ya que estos esperaban una variación de menos del 10% en todos los meses de predicción. Además, se evalúa el comportamiento del modelo en los meses de entrenamiento, con el fin de elegir el modelo más cercano a la realidad.

Se realiza el análisis de la función de autocorrelación (ACF), se utiliza para determinar la parte de media móvil del modelo y la función de autocorrelación parcial (PACF), se utiliza para determinar el proceso autorregresivo. Para cada modelo se obtuvieron los siguientes resultados:

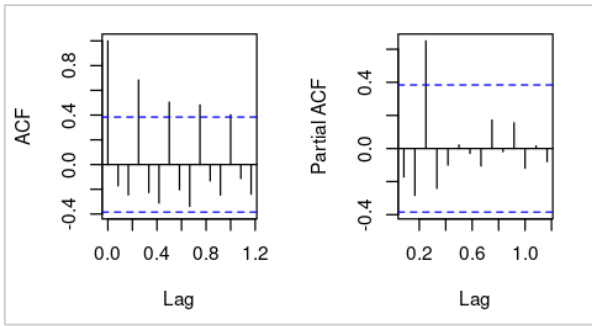


Gráfico 20: ACF y PACF del modelo total nacional.

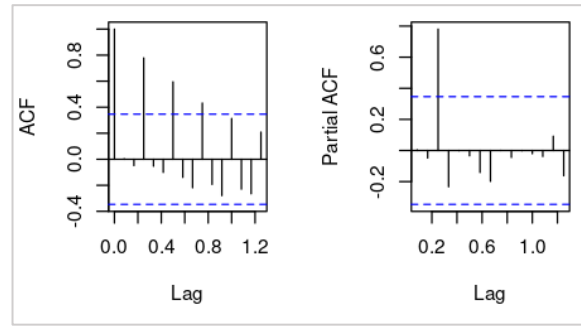


Gráfico 21: ACF y PACF del modelo Jumbo.

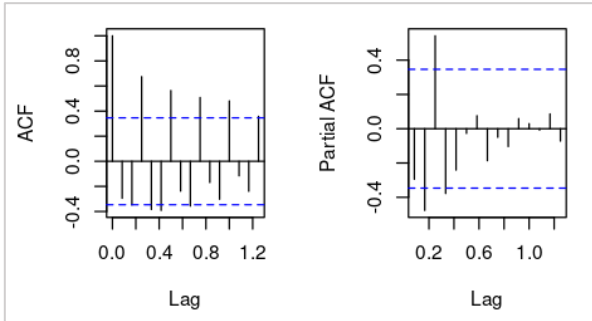


Gráfico 22: ACF y PACF del modelo Santa Isabel.

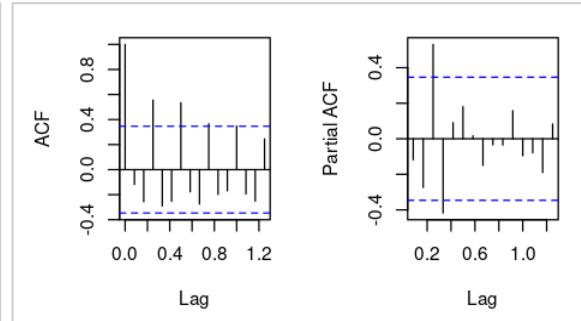


Gráfico 23: ACF y PACF del modelo Unimarc.

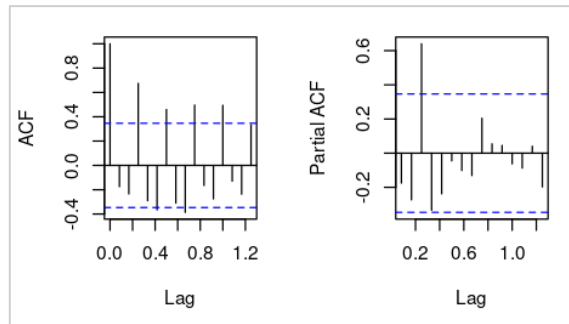


Gráfico 24: ACF y PACF del modelo otras cadenas.

Se analiza como en cada gráfico se ve sobrepasada la línea en color azul, lo que infiere que los modelos muestran una componente autorregresiva y una correspondiente medias móviles, las diferencias en cada modelo muestran que orden llevará cada uno. Por ejemplo en el gráfico 22, correspondiente al análisis de la cadena Santa Isabel, ACF, la componente debería tener al menos un orden 3. Se realizó el mismo análisis para cada uno obtenido los modelos que se muestran en la siguiente tabla.

Los modelos elegidos para cada cadena son modelos ARIMAX con la siguiente estructura.

Modelo	ARIMAX (p,d,q) (P,D,Q)
Total Nacional	(3,0,3) (0,1,0)
Jumbo	(1,1,2) (0,1,2)
Santa Isabel	(0,1,3) (1,0,2)
Unimarc	(2,0,2) (0,1,0)
Otras Cadenas	(1,1,3) (0,1,0)

Tabla 1: Resultados predicción modelo general.

6.4.1.1 Modelo Total Nacional

El primer modelo definido es el total nacional, es decir, el que incluye todas las cadenas de supermercados. Este modelo se realizó con la variable M1, pero de 3 meses antes al mes de predicción y los feriados irrenunciables del mes siguiente a la predicción. Se utilizan variables de meses antes, ya que permite tener información para predecir los meses futuros. En la imagen X se puede ver la significancia de los elementos usados en el modelo, en particular se tiene que el valor p del ARIMA que representa la componente autorregresiva del modelo y la variable del M1 de tres meses antes es significativa para el modelo.

```

z test of coefficients:

              Estimate  Std. Error  z value  Pr(>|z|)
ar1           8.1166e-01  3.7007e-01  2.1933  0.028287 *
ar2          -5.4734e-01  4.1482e-01 -1.3195  0.187013
ar3          -2.5548e-01  3.3549e-01 -0.7615  0.446347
ma1          -1.1458e+00  4.3455e-01 -2.6368  0.008369 **
ma2           1.1458e+00  7.6091e-01  1.5058  0.132109
ma3          -9.9994e-01  6.5395e-01 -1.5291  0.126243
M1_3_meses    1.0885e-02  6.4059e-04 16.9915 < 2.2e-16 ***
irrenunciables_1_mes 3.5695e+01 2.4987e+02 0.1429 0.886403
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
  
```

Imagen 5: Prueba Z de coeficientes modelo ARIMA total nacional

Para este modelo se obtuvieron los resultados mostrados en la tabla siguiente donde se puede ver específicamente el porcentaje de diferencia entre el mes real y el mes de predicción.

Mes del 2022	Diferencia %
Marzo	1,6%
Abril	3,3%
Mayo	1,1%
Junio	4,0%
Julio	0,9%
Agosto	-1,0%

Tabla 2: Resultados predicción modelo general.

Aquí podemos notar que el porcentaje máximo es de un 4% y el menor error es de un 0,9%. Además, también es importante analizar que de entre los modelos tomados este tenía las siguientes métricas.

AIC	BIC	MSE	MAPE
235,06	240,81	87092,64	1,98%

Tabla 3: métricas de evaluación del modelo.

Otros resultados mencionados anteriormente, que es necesario analizar, son las diferencias en los datos de entrenamiento del modelo, lo cual se puede ver en la siguiente tabla.

MES	2020	2021	2022
ENE	-0,1%	2,1%	-5,5%
FEB	-0,1%	5,3%	-2,0%
MAR	-0,1%	4,7%	
ABR	-0,1%	-4,6%	
MAY	-0,1%	1,9%	
JUN	-0,1%	6,7%	
JUL	-0,1%	-3,7%	
AGO	-0,1%	0,9%	
SEP	-0,1%	-1,4%	
OCT	-0,1%	-0,7%	
NOV	-0,1%	1,8%	
DIC	-0,1%	-1,0%	

Tabla 4: diferencia % meses de entrenamiento.

En esta tabla 4 se puede ver que el error máximo es de -5,5% y el mínimo de 0,1%.

Para este modelo de total supermercado se muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

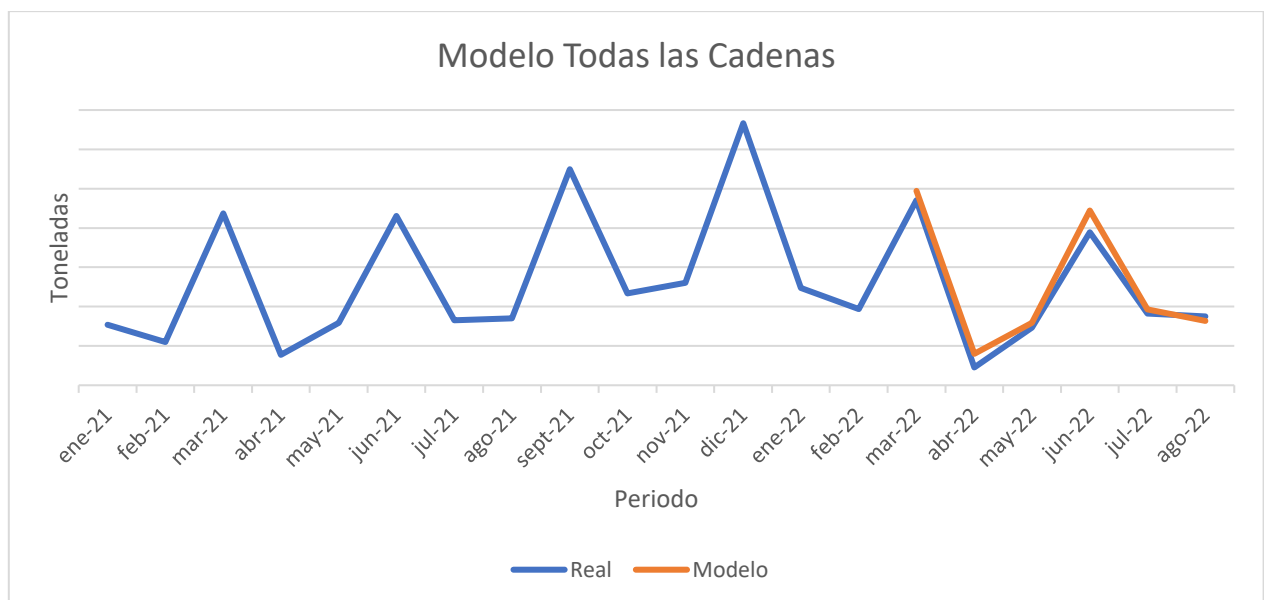


Gráfico 26: Modelo ARIMA total nacional.

6.4.1.2 Modelo Jumbo

Segundo modelo realizado es el modelo de la cadena Jumbo, el modelo seleccionado se realizó con la variable M1 con un desfase de 3 meses, es decir que para marzo del 2022 se utilizó el valor de M1 de diciembre. La componente autorregresiva y las componentes de medias móviles son relevantes en el modelo, no así las componentes de medias móviles, pero de la parte estacional (sma1 y sma2), se puede ver en detalle en la imagen 6.

```

z test of coefficients:

              Estimate  Std. Error   z value  Pr(>|z|)
ar1              1.0000e+00  1.1437e-05  87433.3751 < 2.2e-16 ***
ma1             -1.9389e+00  2.6394e-01  -7.3460  2.043e-13 ***
ma2              9.9926e-01  2.6431e-01   3.7806  0.0001564 ***
sma1            -1.7247e+00  1.2677e+00  -1.3605  0.1736841
sma2              7.4353e-01  1.2681e+00   0.5863  0.5576613
M1_5_meses     -8.9934e-04  1.4613e-03  -0.6154  0.5382588
irrenunciabes_1_mes  4.3525e+01  4.0155e+01   1.0839  0.2783965
irrenunciabes    4.1512e+01  3.8243e+01   1.0855  0.2777124
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Imagen 6: Prueba Z de coeficientes modelo ARIMA Jumbo.

Los resultados de la comparación entre los datos reales y el resultado del modelo se muestran a continuación:

Mes del 2022	Diferencia %
Marzo	5,5%
Abril	-1,5%
Mayo	0,4%
Junio	-0,8%
Julio	-3,5%
Agosto	-2,9%

Tabla 5: Resultados predicción modelo Jumbo.

De la tabla 5 se desprende que el error absoluto máximo entre la predicción y el valor real es de un 5,5% y mínimo de 0,4%. Del modelo utilizado y resultados obtenidos se pueden calcular las siguientes métricas:

AIC	BIC	MSE	MAPE
171,07	176,16	2266,91	2,45%

Tabla 6: Métricas de evaluación modelo Jumbo.

Se obtiene un valor de MAPE de 0,03% lo cual es menor que el modelo general. A continuación, se muestran los resultados de la diferencia porcentual entre los meses de entrenamiento del modelo.

Mes	2020	2021	2022
ENE	-0,07%	0,2%	-0,1%
FEB	-0,03%	4,0%	-2,7%
MAR	-0,04%	4,6%	
ABR	0,01%	-3,2%	

MAY	-0,02%	1,0%
JUN	-0,03%	-2,7%
JUL	-0,01%	1,0%
AGO	-0,01%	3,0%
SEP	-0,03%	0,9%
OCT	-0,01%	1,7%
NOV	-0,01%	2,1%
DIC	0,00%	-0,9%

Tabla 7: Resultado de meses de entrenamiento modelo Jumbo.

La máxima diferencia absoluta en este caso es de un 4,6% y el mínimo de 0,00%. Esto muestra que no hay un sobre ajuste del modelo, lo que apunta decir que como máximo el valor real es un 0,046 veces más que el mes de entrenamiento del modelo.

Para este modelo de Jumbo se muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

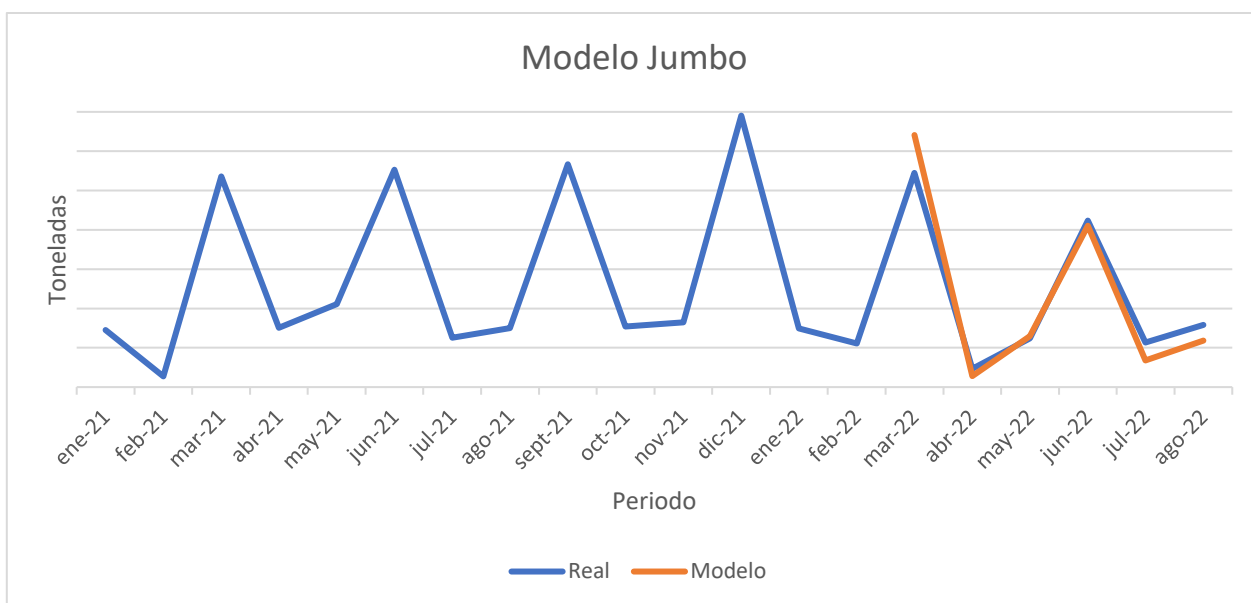


Gráfico 27: Modelo ARIMA cadena Jumbo.

6.4.1.3 Modelo Santa Isabel

Las variables utilizadas para este modelo es el IMACEC con un desfase de cuatro meses y los feriados irrenunciables del mes. De este modelo es significativa la componente autorregresiva 1 (ar1), la de medias móviles (ma1) y la variable IMACEC de 4 meses antes, como se puede apreciar en la imagen 7.

```

z test of coefficients:

              Estimate Std. Error z value Pr(>|z|)
ar1              0.72069   0.26385   2.7314 0.0063062 **
ar2             -0.39147   0.32774  -1.1945 0.2322913
ar3             -0.26688   0.27018  -0.9878 0.3232496
ma1             -0.99992   0.28879  -3.4624 0.0005353 ***
sma1            -0.37895   1.35089  -0.2805 0.7790781
sma2             1.00000   4.40602   0.2270 0.8204541
IMACEC_4_meses  9.23548   1.95354   4.7276 2.272e-06 ***
irrenunciables -74.37577 101.77885 -0.7308 0.4649266
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Imagen 7: Prueba Z de coeficientes modelo ARIMA Santa Isabel.

Se realiza la comparación entre la predicción de este modelo y los valores reales de marzo a agosto del 2022. Se exponen los resultados en la siguiente tabla:

Mes del 2022	Diferencia %
Marzo	-0,2%
Abril	-2,3%
Mayo	-3,2%
Junio	5,0%
Julio	-7,8%
Agosto	4,7%

Tabla 8: Resultados predicción modelo ARIMA Santa Isabel.

En el modelo ARIMA para Santa Isabel se obtienen resultados con mayor diferencia que los dos anteriores (7,8%), del cual se obtienen las siguientes métricas:

AIC	BIC	MSE	MAPE
365,61	376,58	14426,23	6,06%

Tabla 9: Métricas de Evaluación del modelo de Santa Isabel.

Obteniendo un valor de MAPE y MSE menor que el modelo total nacional. Por último, también se muestran las diferencias porcentuales en los meses de entrenamiento y los reales.

Mes	2020	2021	2022
ENE	-0,2%	-0,9%	-3,7%
FEB	-1,7%	4,0%	-3,6%
MAR	-5,4%	-4,0%	
ABR	6,8%	-8,4%	
MAY	-0,8%	-3,4%	
JUN	1,7%	-5,4%	
JUL	4,3%	-4,4%	
AGO	1,4%	2,0%	
SEP	-5,0%	-3,2%	
OCT	4,4%	2,8%	

NOV	-2,5%	-3,3%
DIC	-0,8%	-3,0%

Tabla 10: resultado de meses de entrenamiento modelo Santa Isabel.

Para este modelo de Santa Isabel se muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

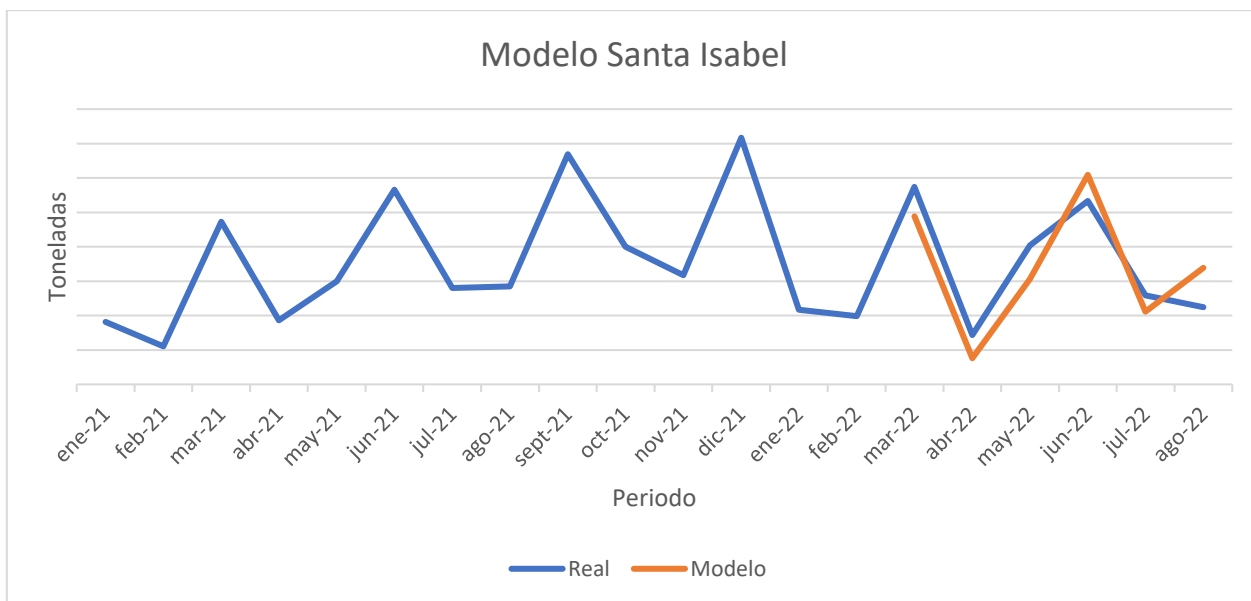


Gráfico 28: Modelo ARIMA cadena Santa Isabel.

6.4.1.4 Modelo Unimarc

El cuarto modelo realizado es de la cadena Unimarc, con las variables: promedio de los 3 meses anteriores del IMACEC de 3 meses antes y los feriados irrenunciables del mes. Se obtienen las siguientes significancias para los parámetros del modelo, se destaca principalmente la componente de medias móviles 2 (ma2) y las variables utilizadas para el modelo.

```

z test of coefficients:

              Estimate Std. Error z value Pr(>|z|)
ar1              0.135581   0.419747   0.3230 0.7466898
ar2             -0.224679   0.296832  -0.7569 0.4490958
ma1             -0.078324   0.308896  -0.2536 0.7998344
ma2              0.999974   0.282013   3.5458 0.0003914 ***
IMACEC_PROM_3_meses 21.094523   5.257806   4.0120 6.020e-05 ***
irrenunciables    238.029735  56.494023   4.2134 2.516e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Imagen 8: Prueba Z de coeficientes modelo ARIMA Unimarc.

Se realiza la comparación de la predicción del modelo para los meses marzo del 2022 a agosto del mismo año, se obtienen los resultados mostrados en la tabla 11.

Mes del 2022	Diferencia %
Marzo	2,9%

Abril	-0,6%
Mayo	-0,7%
Junio	-8,1%
Julio	-8,6%
agosto	0,6%

Tabla 11: Resultados predicción modelo Unimarc.

En tabla anterior se muestra que la mayor diferencia absoluta entre los meses es en el mes Julio, donde se ve una diferencia del 8,6%, es decir, el valor real es mayor que el resultado del modelo. Se calculan las distintas métricas antes expuestas y se obtienen los siguientes valores:

AIC	BIC	MSE	MAPE
192,52	196,99	18851,55	3,58%

Tabla 13: Métricas de Evaluación del modelo de Unimarc.

Finalmente, se muestra la tabla de la diferencia de los meses de entrenamiento:

Mes	2020	2021	2022
ENE	-0,0%	7,7%	-8,7%
FEB	-0,0%	5,8%	-3,9%
MAR	-0,0%	9,7%	
ABR	0,0%	-3,4%	
MAY	0,0%	-9,6%	
JUN	-0,0%	3,8%	
JUL	0,0%	-7,4%	
AGO	-0,0%	3,7%	
SEP	-0,0%	3,8%	
OCT	0,0%	-4,0%	
NOV	-0,0%	-1,5%	
DIC	-0,0%	1,2%	

Tabla 13: Resultados entrenamiento modelo Unimarc.

En esta última tabla de resultados se puede ver que hay una mayor diferencia que en los modelos anteriores, teniendo 3 meses con valores mayores al 8%.

Para este modelo de Unimarc muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

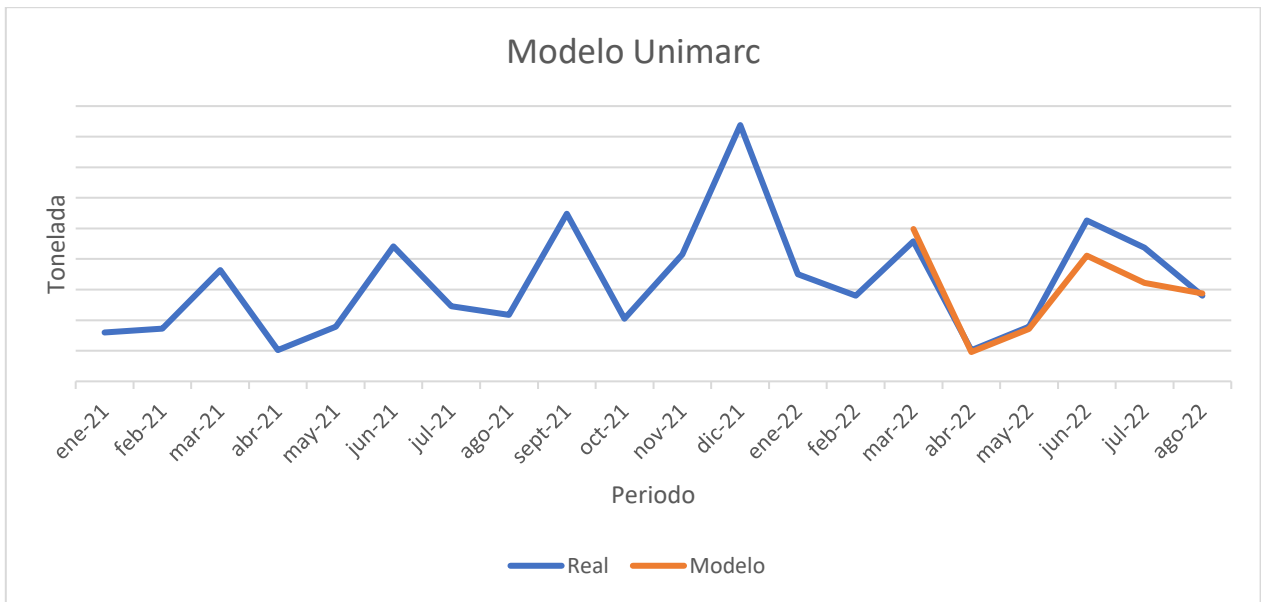


Gráfico 29: Modelo ARIMA cadena Unimarc.

6.4.1.5 Modelo otras cadenas

El modelo realizado para otras cadenas se utiliza las variables IMACEC de 12 meses antes, los feriados irrenunciables del mes de predicción y los del mes siguiente. Se obtienen las siguientes significancias para los parámetros del modelo, se destaca únicamente la componente de autorregresiva 2 (ar2), esto quiere decir que las otras componentes no se saben si son realmente útiles para el modelo.

```

z test of coefficients:

                Estimate Std. Error z value Pr(>|z|)
ar1              0.304476   0.310105   0.9818   0.32618
ar2              0.694945   0.310107   2.2410   0.02503 *
ma1              0.018237   0.989526   0.0184   0.98530
ma2             -0.979495   0.973888  -1.0058   0.31453
IMACEC_3_meses   1.522167  13.148469   0.1158   0.90784
irrenunciables_1_mes -68.552615 175.475939 -0.3907   0.69604
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Imagen 9: Prueba Z de coeficientes modelo ARIMA otras cadenas.

Para el modelo realizado se calculó la predicción de los meses siguientes y la diferencia entre estos y los valores reales, se obtuvieron los siguientes resultados:

Mes del 2022	Diferencia %
Marzo	-1,5%
Abril	0,8%
Mayo	2,2%
Junio	5,4%

Julio	2,8%
agosto	-3,1%

Tabla 14: Resultados predicción modelo ARIMAX otras cadenas.

Tiene mejores resultados que los otros modelos, con un máximo de 5,4% de diferencia absoluta. Para la evaluación general del modelo, se calcularon las métricas:

AIC	BIC	MSE	MAPE
199,01	203,53	22069,98	2,65%

Tabla 15: Métricas de Evaluación del modelo ARIMAX de otras cadenas.

El valor del AIC y BIC es mucho mayor que el de los modelos anteriores, pero el MAPE y MSE están dentro del rango de los otros. A continuación, se muestra las diferencias entre los meses de entrenamiento del modelo y los reales:

Mes	2020	2021	2022
ENE	-0,1%	0,5%	-5,7%
FEB	0,0%	6,2%	1,9%
MAR	0,0%	6,6%	
ABR	0,0%	-4,7%	
MAY	0,0%	0,4%	
JUN	0,0%	-4,7%	
JUL	0,0%	-3,5%	
AGO	0,0%	2,8%	
SEP	0,0%	-3,6%	
OCT	0,0%	-5,2%	
NOV	0,0%	-4,8%	
DIC	0,0%	2,7%	

Tabla 16: Resultados entrenamiento modelo ARIMAX otras cadenas.

Acá se puede ver una menor variación comprando con el modelo Unimarc, teniendo un máximo absoluto de 6,6%.

Para este modelo de otras cadenas se muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

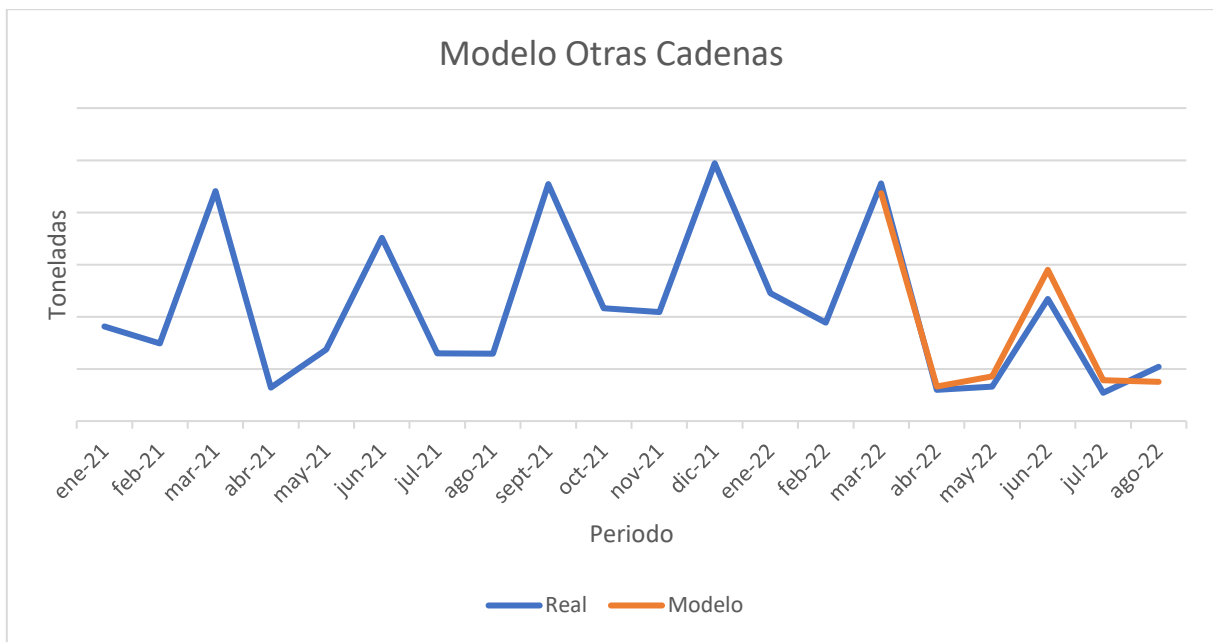


Gráfico 30: Modelo ARIMA Otras cadenas.

6.4.2 Modelo Regresión Lineal

Se realizaron distintas pruebas de modelo obteniendo el de mejor comportamiento según el valor del R^2 y del R^2 buscando que estos valores se acerquen lo más posible a 1 ya que esto muestra qué tanto las variables independientes, es decir, las externas que se están usando para predecir sirven para predecir el comportamiento de la variable independiente.

Cabe mencionar que se probaron al menos 10 combinaciones distintas por cadena, considerando la correlación de las variables y valores del R^2 .

6.4.2.1 Modelo Total Nacional

Se realizó un modelo de utilizando un ajuste de los datos de demanda, con el fin de aplanar los peaks que se producen cada 3 meses.

Para este modelo se utiliza la variable M1 de 4 meses antes, obteniendo un R^2 de 0,45, un R^2 ajustado de 0,43 y p-valor = 0,0001753 lo que permite rechazar la hipótesis nula, dando a entender que estas variables son representativas en el modelo. A continuación, se muestran los coeficientes asociados al modelo:

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.861e+03	6.038e+02	14.676	1.74e-13 ***
M1_4_meses	1.094e-02	2.466e-03	4.434	0.000175 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Imagen 10: Coeficientes regresión lineal total nacional.

Realizando un reverso al ajuste antes mencionado, se obtienen las diferencias porcentuales por mes mostradas en la tabla 17.

Mes del 2022	Diferencia %
Marzo	4,7%
Abril	18,2%
Mayo	6,9%
Junio	9,0%
Julio	1,6%
agosto	1,6%

Tabla 17: Resultados predicción modelo regresión lineal Total nacional.

Se obtiene un máximo error absoluto de 18,2% en el mes de abril, este modelo sobreestima la demanda en todos los meses, ya que se obtienen valores positivos para todos ellos. Este modelo tiene las siguientes métricas:

AIC	BIC	MSE	MAPE
420,85	424,63	1056430	7,00%

Tabla 18: Métricas de medición modelo regresión lineal Total Nacional.

Las métricas mostradas en la tabla 18, se usarán posteriormente para hacer la comparación entre modelos.

Para este modelo de total supermercado se muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

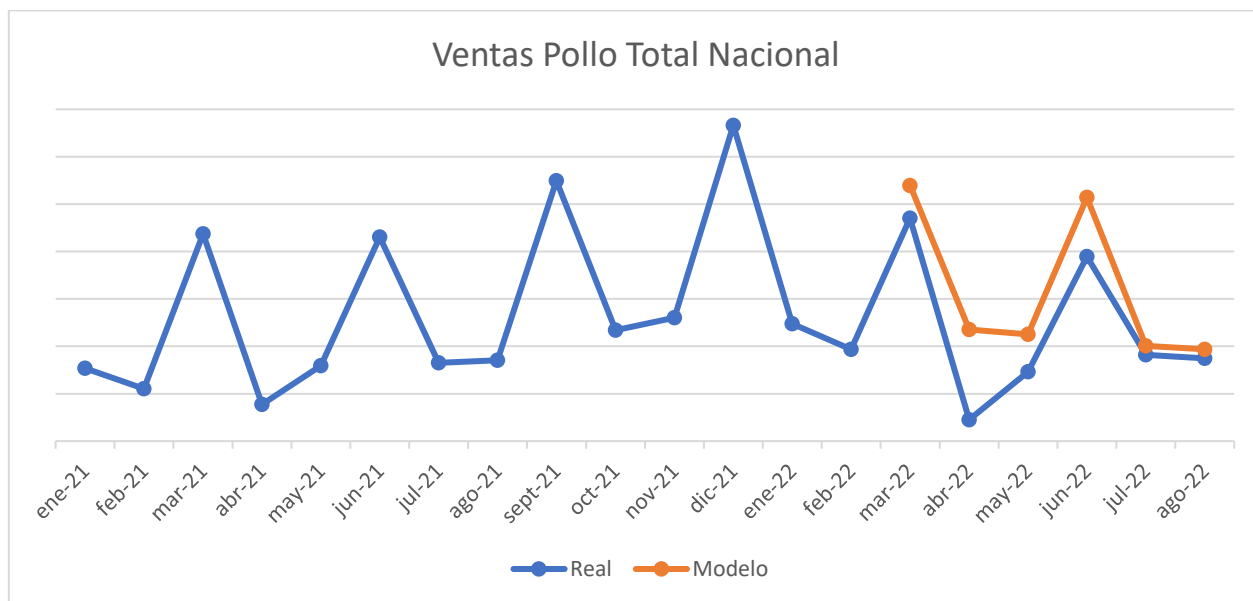


Gráfico 31: Modelo Regresión lineal Total Nacional.

6.4.2.2 Modelo Jumbo

Para este modelo se utilizó el mismo ajuste que para el modelo total nacional y se utilizaron las variables M1 de 3 meses antes y el IMACEC con 5 meses de anticipación. Resultando un R^2 de 0,73 y un R^2 ajustado de 0,71 aproximadamente y un p-valor = 2.841e-07. El resultado del p-valor ($<0,05$) indica que es un modelo estadísticamente significativo, es decir, que las variables son capaces de predecir significativamente la variable dependiente (la venta de pollo en supermercados Jumbo). A continuación, se muestran los coeficientes asociados al modelo:

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.104e+03  2.622e+02  4.210 0.000334 ***
M1_3_meses   2.690e-03  3.410e-04  7.889 5.44e-08 ***
IMACEC_5_meses -5.187e+00  2.750e+00  -1.886 0.072000 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Imagen 11: Coeficientes regresión lineal Jumbo.

Se calcula la a diferencia porcentual por mes, obteniendo los resultados de la tabla 19.

Mes del 2022	Diferencia %
Marzo	-1,3%
Abril	9,1%
Mayo	-0,7%
Junio	4,8%
Julio	4,0%
agosto	-6,8%

Tabla 19: Resultados predicción modelo regresión lineal Jumbo.

Este modelo tiene las métricas mostradas en la tabla 20, las cuales son usadas para realizar la comparación con los otros tipos de modelo.

AIC	BIC	MSE	MAPE
316,27	321,30	5127,421	4,45%

Tabla 20: Métricas de medición modelo regresión lineal Jumbo.

Para este modelo de Jumbo se muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

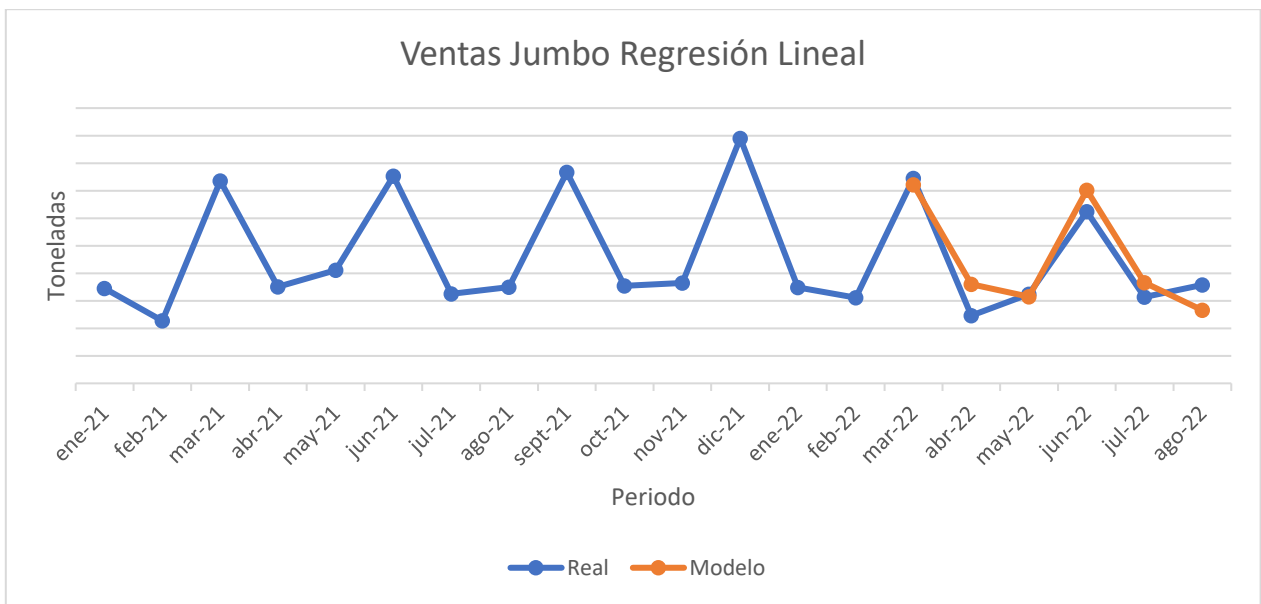


Gráfico 32: Modelo Regresión lineal cadena Jumbo.

6.4.2.3 Modelo Santa Isabel

El modelo de regresión lineal para la cadena Santa Isabel, se armó con la variable feriados irrenunciables del mes siguiente. Resultando un R^2 de 0,12, un R^2 ajustado de 0,08 y un p-valor de 0,0867. El resultado del p-valor indica que este resultado no es capaz de rechazar la hipótesis nula, lo que muestra que el modelo no es estadísticamente significativo. A continuación, se muestran los coeficientes asociados a las variables del modelo:

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2463.73	44.96	54.798	<2e-16 ***
irrenunciables_1_mes	70.22	39.32	1.786	0.0868 .

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Imagen 12: Coeficientes regresión lineal total Santa Isabel.

A pesar de que el modelo no sea estadísticamente significativo se mostrarán los resultados del modelo y sus respectivas métricas en las siguientes tablas.

Mes del 2022	Diferencia %
Marzo	0,6%
Abril	7,7%
Mayo	-12,2%
Junio	0,5%
Julio	-2,2%
agosto	6,3%

Tabla 21: Resultados predicción modelo regresión lineal Santa Isabel.

Para este modelo también se realizó el ajuste para aplanar las curvas, buscando obtener un mejor modelo. Obteniendo las siguientes métricas de medición:

AIC	BIC	MSE	MAPE
348,45	352,22	29541,17	4,95%

Tabla 22: Métricas de medición modelo regresión lineal Santa Isabel.

Para este modelo de Santa Isabel muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

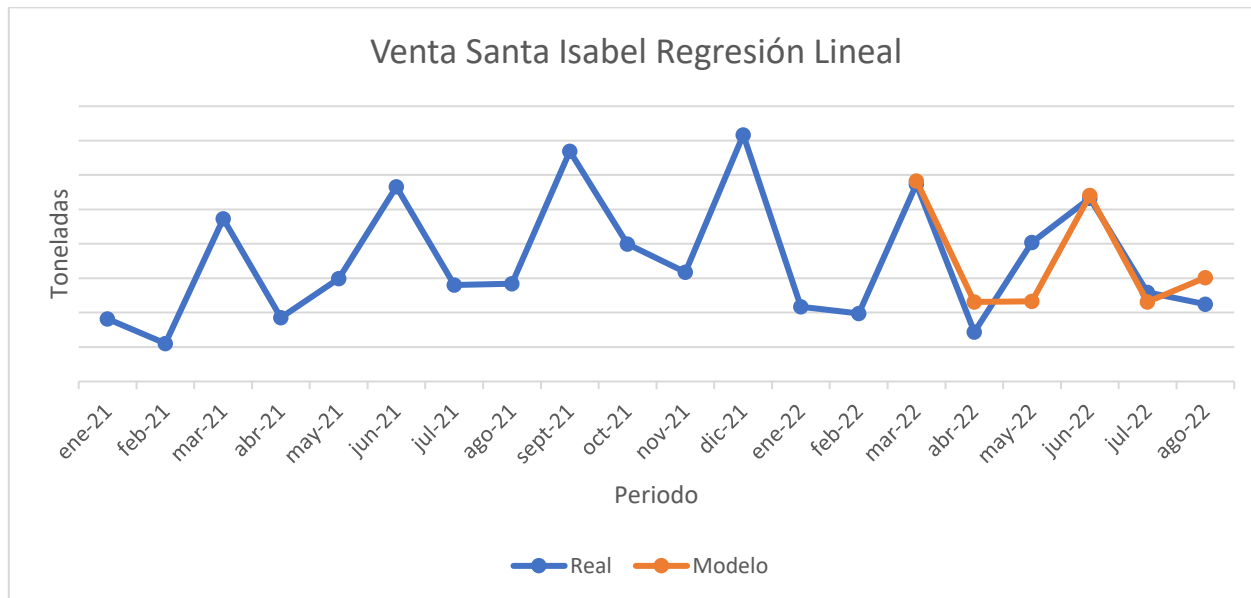


Gráfico 33: Modelo Regresión lineal cadena Santa Isabel.

6.4.2.4 Modelo Unimarc

Para el modelo de regresión lineal de la cadena Unimarc se utilizaron las variables IMACEC de meses antes y los feriados irrenunciables de cada mes. Resultando un modelo con un R^2 0,45, un R^2 ajustado 0,40 y un p-valor = 0,0009846. Ya que el resultado del p-valor es menor a 0,5 se puede concluir que se rechaza la hipótesis nula, por lo que es un modelo estadísticamente significativo. En la imagen 13 se puede ver el valor de los coeficientes asociados a las variables del modelo.

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	69.578	726.319	0.096	0.92451
IMACEC_3_meses	21.850	7.172	3.046	0.00573 **
irrenunciables	228.583	64.592	3.539	0.00175 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Imagen 13: Coeficientes regresión lineal Unimarc.

Se realiza la comparación de la predicción del modelo con los valores reales, obteniendo los siguientes resultados:

Mes del 2022	Diferencia %
Marzo	-0,7%
Abril	29,1%
Mayo	3,3%
Junio	-9,6%
Julio	-8,3%
agosto	6,0%

Tabla 23: Resultados predicción modelo regresión lineal Unimarc.

El mes de abril es el con mayor diferencia con 29,1%, esto quiere decir que el resultado del modelo es 0,29 veces mayor que el mes original. Con estos resultados se procede a calcular las métricas de medición que permitirán realizar la comparación entre los modelos. Las métricas de este modelo se muestran en la siguiente tabla:

AIC	BIC	MSE	MAPE
374,56	379,59	81284,32	9,5%

Tabla 24: Métricas de medición modelo regresión lineal Unimarc.

Para este modelo de la cadena Unimarc muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

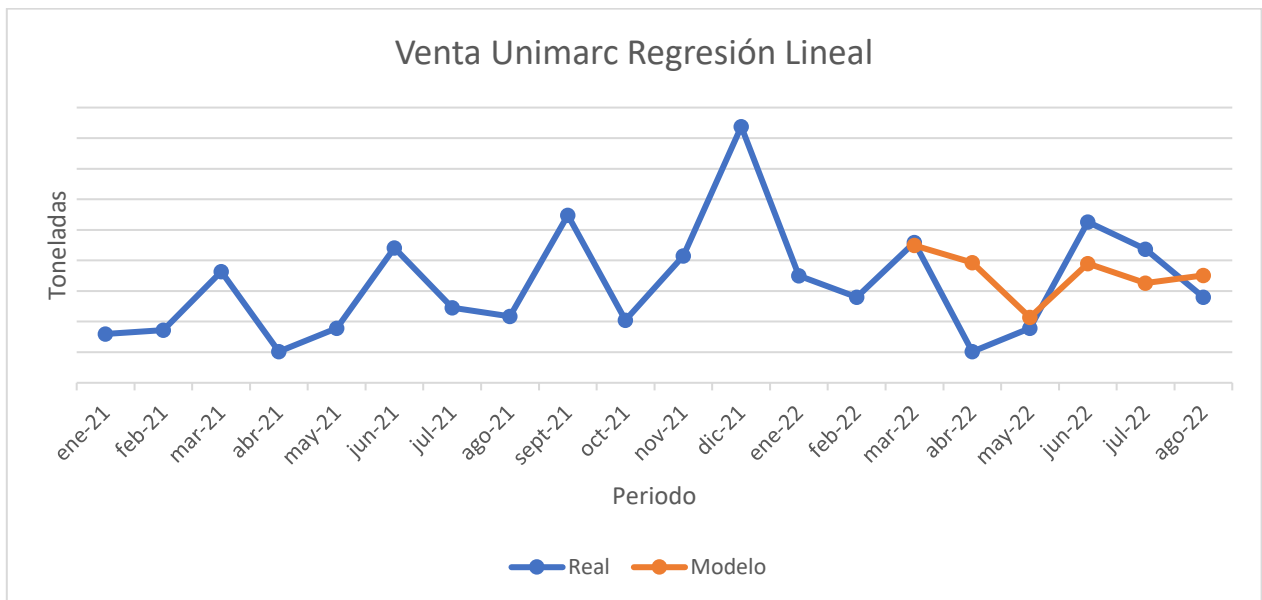


Gráfico 34: Modelo Regresión lineal cadena Unimarc.

6.4.2.5 Modelo Otras cadenas

El modelo para las otras cadenas se realiza con la variable M1 de 3 meses antes. Con esta variable se obtiene un modelo con un R^2 igual a 0,32, un R^2 ajustado de 0,29 y un p-valor = 0.002757, ya que este valor es menor a 0,05 se puede decir que es un modelo estadísticamente significativo. Se muestran los detalles de los coeficientes a continuación:

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.760e+03  2.927e+02  12.847 3.01e-12 ***
M1_3_meses  3.894e-03  1.167e-03   3.336 0.00276 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Imagen 14: Coeficientes regresión lineal otras cadenas.

Además, se obtiene la siguiente diferencia porcentual entre los valores de predicción y del modelo de:

Mes del 2022	Diferencia %
Marzo	-3,3%
Abril	7,9%
Mayo	4,2%
Junio	14,6%
Julio	12,7%
agosto	3,4%

Tabla 25: Resultados predicción modelo regresión lineal para otras cadenas.

Para este modelo se calculan las siguientes métricas, las cuales permitirán realizar la comparación de los modelos:

AIC	BIC	MSE	MAPE
380,91	387,20	180078,7	7,68%

Tabla 26: Métricas de medición modelo regresión lineal para otras cadenas.

Para este modelo de Otras Cadenas muestra el siguiente gráfico, que muestra la diferencia entre el modelo y los valores reales para los meses de testeo.

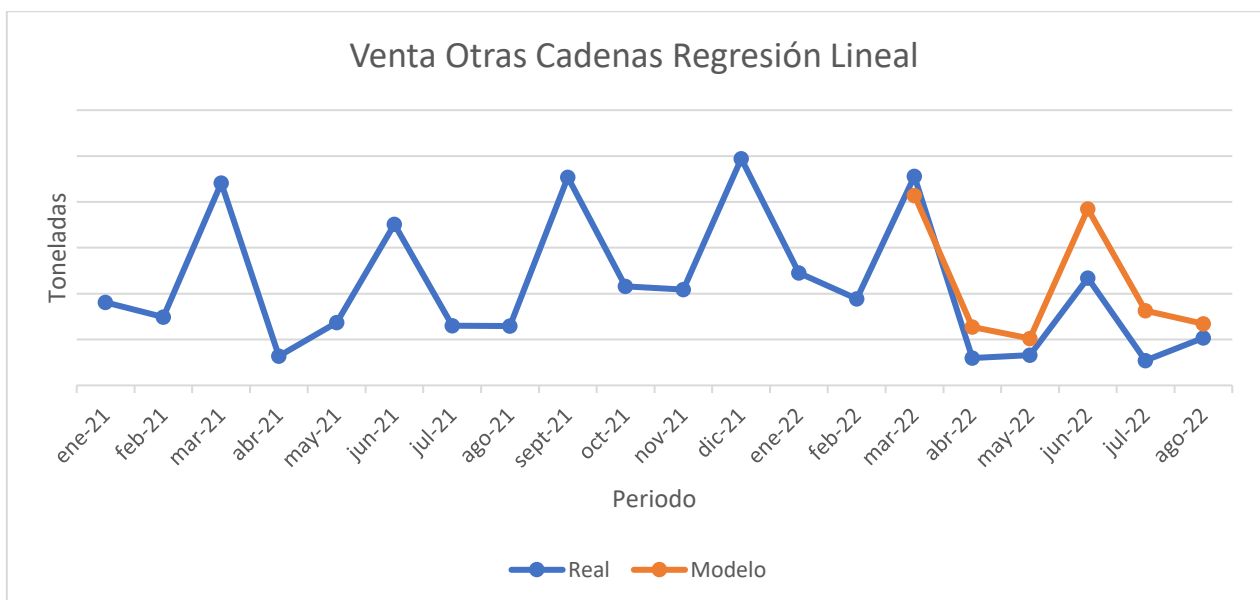


Gráfico 32: Modelo Regresión lineal Otras Cadenas.

6.4.3 Modelo Alisamiento Exponencial

Se realizaron 4 modelos distintos de Alisamiento exponencial para cada cadena y el total nacional, evaluando cada uno de estos según sus métricas y la diferencia porcentual de cada mes de predicción. Se eligió uno como el modelo de mejor comportamiento el cual será mostrado en cada apartado con sus respectivas métricas de medición.

6.4.3.1 Modelo Total Nacional

El modelo para la serie total nacional se realizó con un alisamiento exponencial “AAA”, es decir, utilizando un error, una tendencia y una estacionalidad aditiva respectivamente. El hecho de que sea AAA y no tenga N en el tipo de modelo, quiere decir que captura la componente estacional y la tendencia de la serie.

Este modelo utiliza unos valores alfa y beta cercanos a cero, esto quiere decir que se utilizan en mayor manera las tendencias y estacionalidad de la serie para la predicción. A continuación, se muestran los parámetros del modelo:

```
Smoothing parameters:  
alpha = 1e-04  
beta = 1e-04  
gamma = 0.5687  
phi = 0.98
```

Imagen 15: Parámetros modelo suavizamiento exponencial para total nacional.

Como se mencionó anteriormente se logra capturar la tendencia y la estacionalidad de la serie, lo que se puede ver el gráfico siguiente:



Gráfico 26: Modelo alisamiento exponencial modelo total nacional.

En el gráfico 26 la línea azul representa la predicción del modelo con su respectivo intervalo de confianza. Con este modelo, se realizó la comparación de la predicción con los valores reales para los meses de marzo a agosto del 2022 y se obtuvieron los siguientes resultados:

Mes del 2022	Diferencia %
Marzo	6,7%
Abril	4,6%
Mayo	3,1%
Junio	4,7%
Julio	-1,1%
agosto	6,5%

Tabla 27: Resultados predicción modelo alisamiento exponencial para total nacional.

Para este modelo se calculan las siguientes métricas:

MSE	MAPE
390190,7	4,45%

Tabla 28: Métricas de medición modelo alisamiento exponencial para total nacional.

No se calculan las métricas AIC y BIC debido a que este modelo no utiliza otras variables, solo utiliza la venta de total de pollo en Supermercados en Chile.

6.4.3.2 Modelo Jumbo

El modelo de alisamiento exponencial para Jumbo se realiza con un modelo “AAA”, es decir, el error, la tendencia y la estacionalidad se utilizan de forma aditiva. Para este modelo se utilizan los siguientes parámetros:

```
Smoothing parameters:
alpha = 0.0095
beta  = 1e-04
gamma = 4e-04
phi   = 0.9482
```

Imagen 16: Parámetros modelo suavizamiento exponencial para Jumbo.

En el gráfico 27 se puede notar como el modelo logra capturar la tendencia y la estacionalidad de la serie, ya que los meses de predicción que están representados por la línea azul del gráfico tienen peaks cada 3 meses igual que la serie original.

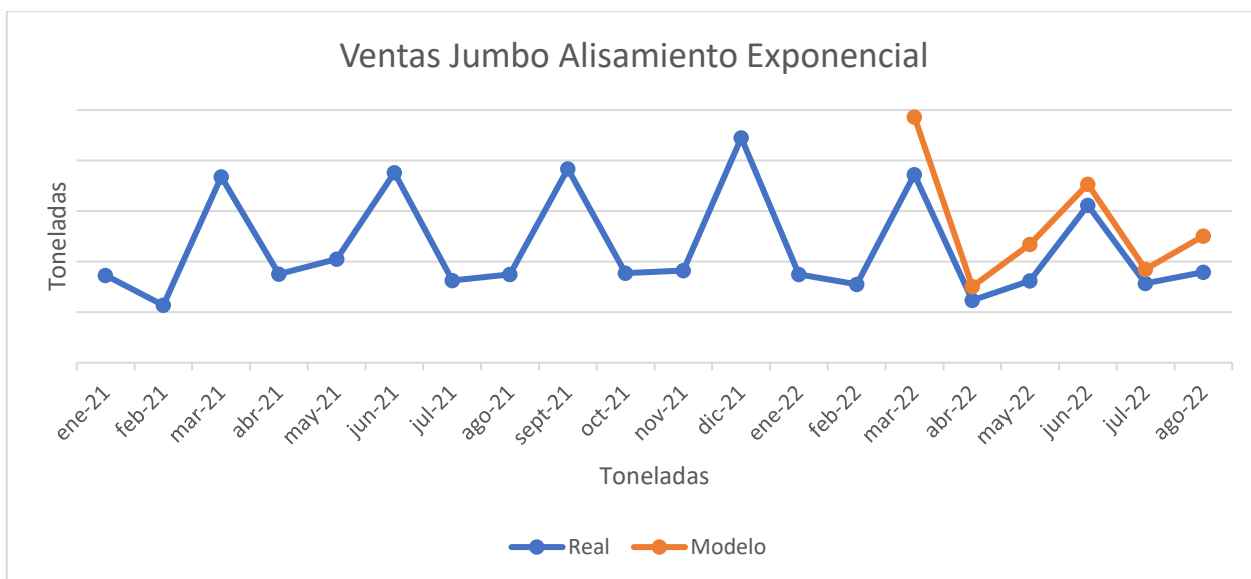


Gráfico 27: Modelo alisamiento exponencial modelo Jumbo.

Con este modelo se realiza una predicción de 6 meses, desde marzo a agosto del 2022 y estos se comparan con los valores reales de la serie en los meses mencionados. La comparación por mes se muestra en la tabla 29.

Mes del 2022	Diferencia %
Marzo	13,1%
Abril	4,4%
Mayo	10,9%
Junio	5,1%
Julio	4,3%
agosto	10,6%

Tabla 29: Resultados predicción modelo alisamiento exponencial para Jumbo.

Se calculan las métricas que indican como se está comportando el modelo, los valores resultantes se muestran en la tabla 30.

MSE	MAPE
17762,8	8,07%

Tabla 30: Métricas de medición modelo alisamiento exponencial para Jumbo.

6.4.3.3 Modelo Santa Isabel

El modelo de alisamiento exponencial para Santa Isabel se realiza con un modelo “MAM”, es decir, el error y la estacionalidad se utilizan de forma multiplicativa, a diferencia de la tendencia que se utiliza de manera aditiva.

Para este modelo se utilizan los parámetros de la imagen 17, la cual muestra que los parámetros alfa y beta son números cercanos a cero, lo que implica, que no se le está dando

importancia a la historia de los datos, si no que la serie utiliza la estacionalidad y la tendencia para ajustar el modelo.

Smoothing parameters:
 alpha = 0.0294
 beta = 2e-04
 gamma = 0.2106
 phi = 0.98

Imagen 17: Parámetros modelo suavizamiento exponencial para Santa Isabel.

El gráfico 28 se puede ver en la línea de color azul, como está realizando la predicción el modelo, se logra visualizar como rescata los peaks que ocurren cada 3 meses.

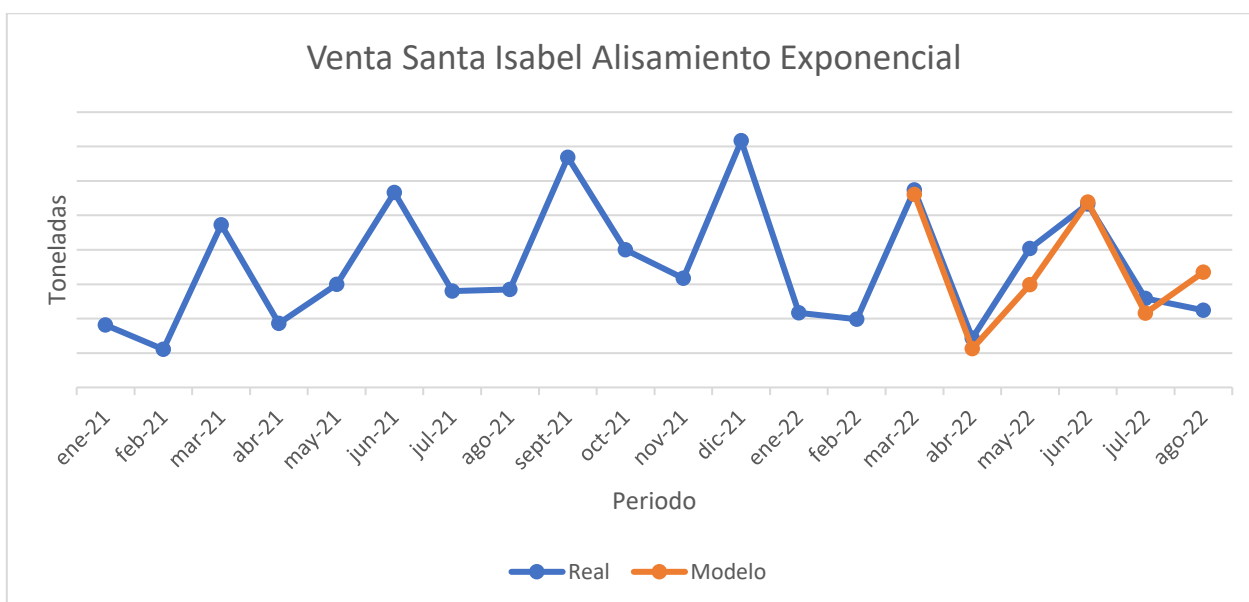


Gráfico 28: Modelo alisamiento exponencial modelo Santa Isabel.

Se realiza la comparación de los 6 meses de predicción con el valor de real de cada mes, obteniendo los resultados mostrados en la tabla 31.

Mes del 2022	Modelo 1
Marzo	-0,8%
Abril	-2,7%
Mayo	-7,5%
Junio	0,4%
Julio	-3,4%
agosto	9,0%

Tabla 31: Resultados predicción modelo alisamiento exponencial para Santa Isabel.

Con los meses de predicción del modelo se realiza el cálculo de las métricas. Se obtienen los resultados mostrados en la tabla 32.

MSE	MAPE
17516,3	3,97%

Tabla 32: Métricas de medición modelo alisamiento exponencial para Santa Isabel.

6.4.3.4 Modelo Unimarc

El modelo de alisamiento exponencial para Unimarc se realiza con un modelo “MAM”, es decir, el error y la estacionalidad es multiplicativa, a diferencia de la tendencia la usa de forma aditiva.

Para este modelo se utilizan los parámetros de la imagen 18, donde se aprecia que los parámetros alfa y beta y gamma son números cercanos a cero, lo que implica, que no se le está dando importancia a la historia de los datos ni a la tendencia, si no que la serie utiliza en mayor medida la estacionalidad para ajustar el modelo.

Smoothing parameters:

alpha = 0.0364

beta = 1e-04

gamma = 1e-04

phi = 0.96

Imagen 18: Parámetros modelo suavizamiento exponencial para Unimarc.

El gráfico 29 se puede ver en la línea de color azul, como está realizando la predicción el modelo, se logra visualizar como rescata los peaks que ocurren cada 3 meses.

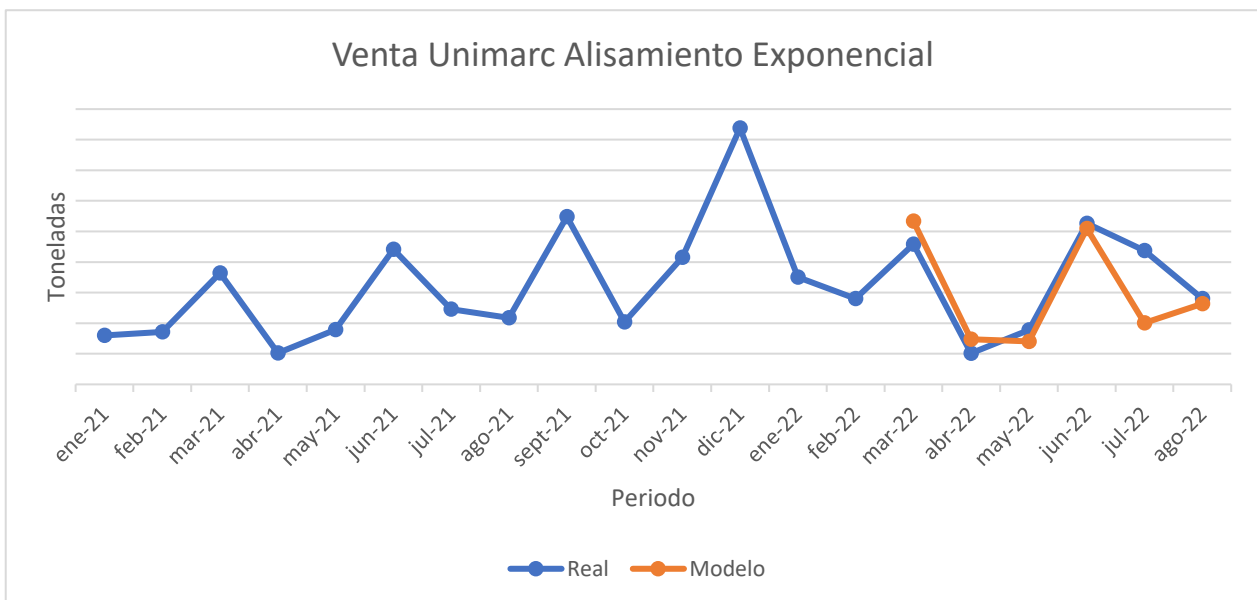


Gráfico 29: Modelo alisamiento exponencial modelo Unimarc.

Se realiza la comparación de los 6 meses de predicción con el valor de real de cada mes, obteniendo los resultados mostrados en la tabla 33.

Mes del 2022	Modelo 1
Marzo	5,5%
Abril	4,6%
Mayo	-3,5%
Junio	-1,2%
Julio	-17,7%
agosto	-1,4%

Tabla 33: Resultados predicción modelo alisamiento exponencial para Unimarc.

Con los meses de predicción del modelo se realiza el cálculo de las métricas. Se obtienen los resultados mostrados en la tabla 34.

MSE	MAPE
43883,0	5,65%

Tabla 34: Métricas de medición modelo alisamiento exponencial para Unimarc.

6.4.3.5 Modelo Otras cadenas

El modelo de alisamiento exponencial para otras cadenas se realiza con un modelo “MAM”, es decir, el error y la estacionalidad es multiplicativa, a diferencia de la tendencia la usa de forma aditiva.

Para este modelo se utilizan los parámetros de la imagen 19, donde se aprecia que los parámetros alfa y beta y gamma son números cercanos a cero, lo que implica, que no se le está dando importancia a la historia de los datos ni a la tendencia, si no que la serie utiliza en mayor medida la estacionalidad para ajustar el modelo.

```
Smoothing parameters:
alpha = 0.0565
beta  = 2e-04
gamma = 0.0042
phi   = 0.978
```

Imagen 19: Parámetros modelo suavizamiento exponencial para otras cadenas.

El gráfico 30 se puede ver en la línea de color azul, como está realizando la predicción el modelo, se logra visualizar como rescata los peaks que ocurren cada 3 meses.

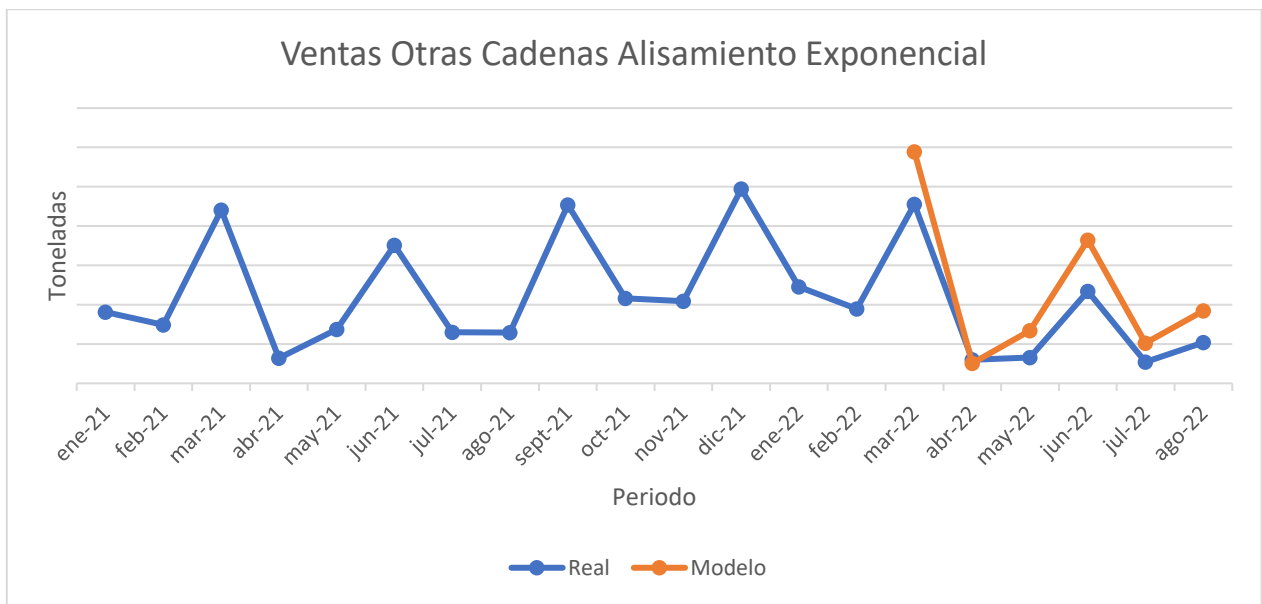


Gráfico 30: Modelo alisamiento exponencial modelo otras cadenas.

Se realiza la comparación de los 6 meses de predicción con el valor de real de cada mes, obteniendo los resultados mostrados en la tabla 35.

Mes del 2022	Modelo 1
Marzo	10,6%
Abril	-1,0%
Mayo	7,9%
Junio	12,6%
Julio	5,6%
agosto	8,9%

Tabla 35: Resultados predicción modelo alisamiento exponencial para otras cadenas.

Con los meses de predicción del modelo se realiza el cálculo de las métricas. Se obtienen los resultados mostrados en la tabla 36.

MSE	MAPE
201997,4	7,77%

Tabla 36: Métricas de medición modelo alisamiento exponencial para Otras Cadenas.

6.5 Evaluación de Modelos

Se realiza la comparación entre modelos ARIMA, regresión lineal y alisamiento exponencial, se considera la diferencia porcentual entre los meses de predicción y los meses reales y las métricas calculadas para cada modelo.

6.5.1 Modelos Total Nacional

En la tabla 37 se muestra la diferencia porcentual entre los meses reales y la predicción de los distintos tipos de modelos para la serie total nacional.

Mes del 2022	ARIMA	Regresión lineal	Alisamiento exponencial
Marzo	1,6%	4,7%	6,7%
Abril	3,3%	18,2%	4,6%
Mayo	1,1%	6,9%	3,1%
Junio	4,0%	9,0%	4,7%
Julio	0,9%	1,6%	-1,1%
Agosto	-1,0%	1,6%	6,5%

Tabla 37: Resumen de diferencia porcentual por modelo para total nacional.

En particular en la tabla anterior, se observa que todos los modelos en general sobreestiman la demanda, solo el modelo ARIMA en agosto y el modelo alisamiento exponencial en Julio subestiman la demanda, obteniendo valores negativos en la diferencia. Además, se observa que el modelo ARIMA tiene menores diferencias porcentuales en cada mes.

El mes con mayor error del modelo ARIMA es el mes Junio con un 4%, este valor es menos de la mitad del error del mismo mes de la regresión lineal y el modelo de alisamiento exponencial tiene un error 0,7% mayor. Considerando esto ARIMA, es el modelo que logra acercarse más a los valores reales. Para analizar el comportamiento el modelo en cuanto a las métricas de cada modelo, se muestra la tabla 38 con la comparación de cada tipo de modelo.

Modelos	AIC	BIC	MSE	MAPE
ARIMA	235,06	240,81	87.092,6	1,98%
Regresión Lineal	420,85	424,63	1.056.430	7%
Alisamiento Exponencial	-	-	390.190,7	4,45%

Tabla 38: Resumen de métricas por modelo para total nacional.

Al comparar los valores AIC y BIC, se aprecia que los valores son menores para el modelo ARIMA, por lo que se puede concluir que el modelo ARIMA es mejor modelo que la regresión lineal en cuanto a los estimadores usados. Además, si se comparan los valores del MSE, se busca el valor menor para cada modelo, lo que ocurre para el Modelo ARIMA siendo 4,48 veces menor que el modelo de alisamiento exponencial y 12,12 veces menor que el MSE de la regresión lineal, lo que apoya la conclusión de que el modelo ARIMA es el mejor modelo para la serie total nacional. Por último se compara la métrica MAPE, que al igual que MSE se busca que sea cercano a cero, el modelo que logra acercarse más a cero es el modelo ARIMA con un 1,98% siendo menor que la regresión lineal con un 7% y que el modelo alisamiento exponencial con un 4,45% absoluto.

Con este análisis se concluye que el modelo que logra representar mejor a la serie total nacional es el modelo ARIMA.

6.5.2 Modelos Jumbo

Para el modelo Jumbo se analiza el error porcentual por mes entre los distintos tipos de modelos, considerando que el negocio espera diferencias de menos del 10% para cada mes.

Mes del 2022	ARIMA	Regresión lineal	Alisamiento exponencial
Marzo	5,5%	-1,3%	13,1%
Abril	-1,5%	9,1%	4,4%
Mayo	0,4%	-0,7%	10,9%
Junio	-0,8%	4,8%	5,1%
Julio	-3,5%	4,0%	4,3%
Agosto	-2,9%	-6,8%	10,6%

Tabla 39: Resumen de diferencia porcentual por modelo para Jumbo.

En la tabla 39 se muestran los errores porcentuales para cada mes, se puede notar que el modelo de alisamiento exponencial tiene errores mayores al 10%, lo que no cumple con la condición definida por el negocio. Se analizan los modelos ARIMA y regresión lineal, obteniendo que el primer modelo tiene errores absolutos menores que la regresión lineal en todos los meses excepto en marzo, que el modelo ARIMA tiene una diferencia de 4,2% más. Además, regresión lineal tiene un máximo error absoluto porcentual de 9,1% a diferencia del modelo ARIMA que presenta un 5,5% de máximo error porcentual. Para lograr concluir que modelo es mejor, se analizan las métricas de cada modelo.

Modelos	AIC	BIC	MSE	MAPE
ARIMA	171,07	176,16	2.266,9	2,45%
Regresión Lineal	316,27	321,30	5.127,4	4,55%

Tabla 40: Resumen de métricas por modelo para Jumbo.

En la tabla 40 se muestran la comparación de métricas de cada modelo, en particular el modelo ARIMA presenta valores menores en todas las métricas, lo que permite concluir que es modelo es más representativo para esta serie.

6.5.3 Modelos Santa Isabel

Se realiza la comparación de los modelos evaluados para la cadena Santa Isabel, se puede observar en la tabla 41, la diferencia porcentual entre los meses reales y la predicción de cada modelo.

Mes del 2022	ARIMA	Regresión lineal	Alisamiento exponencial
Marzo	-0,2%	0,6%	-0,8%
Abril	-2,3%	7,7%	-2,7%
Mayo	-3,2%	-12,2%	-7,5%
Junio	5,0%	0,5%	0,4%
Julio	-7,8%	-2,2%	-3,4%

Agosto	4,7%	6,3%	9,0%
--------	------	------	------

Tabla 41: Resumen de diferencia porcentual por modelo para Santa Isabel.

En este caso, el modelo regresión lineal no cumple la condición de tener errores de menor del 10%, por lo que solo se analizan los resultados de los modelos ARIMA y alisamiento exponencial.

El error absoluto máximo del modelo ARIMA es un 7,8% el cual es menor que el del alisamiento exponencial que tiene un 9,0%, esto indica que el primer modelo podría ser mejor que el otro. Al comparar la diferencia entre los meses, se desprende que, de seis meses, en cuatro el modelo ARIMA tiene menor error, y en 2 tiene menor error el modelo alisamiento exponencial. Lo que aun no permite concluir que modelo logra realizar una mejor predicción, para poder analizar las métricas se muestra la tabla 42 con la comparación de estas para cada modelo.

Modelos	MSE	MAPE
ARIMA	14.426,2	6,06%
Alisamiento Exponencial	17.516,3	3,97%

Tabla 42: Resumen de métricas por modelo para Santa Isabel.

No se muestran los valores de AIC y BIC debido a que este valor no se puede calcular para el modelo alisamiento exponencial. Considerando las métricas MSE y MAPE, donde cada modelo tiene una métrica menor al otro. El modelo ARIMA tiene absolutos menores en general, con una diferencia del 1,2 puntos menos que el modelo Alisamiento exponencial, lo que permite concluir que el mejor modelo en este caso es el ARIMA.

6.5.4 Modelos Unimarc

Se realiza la comparación de los distintos tipos de modelos para la cadena Unimarc, primeramente, se nota en la tabla 43 que el único modelo que cumple con tener errores de menos del 10% es el modelo ARIMA. Por lo que, es el único modelo que es capaz de cumplir con las peticiones del negocio.

Mes del 2022	ARIMA	Regresión lineal	Alisamiento exponencial
Marzo	2,9%	-0,7%	5,5%
Abril	-0,6%	29,1%	4,6%
Mayo	-0,7%	3,3%	-3,5%
Junio	-8,1%	-9,6%	-1,2%
Julio	-8,6%	-8,3%	-17,7%
agosto	0,6%	6,0%	-1,4%

Tabla 43: Resumen de diferencia porcentual por modelo para Unimarc.

A pesar de que el modelo ARIMA es el único con errores menores al 10%, de igual manera se realiza un análisis estadístico, para entender si a pesar de tener errores menores igual es mejor modelo en cuanto a métricas, lo que se puede ver en la tabla 44.

Modelos	AIC	BIC	MSE	MAPE
ARIMA	192,52	196,99	18.851,55	3,58%
Regresión Lineal	374,56	379,59	81.284,32	9,5%
Alisamiento Exponencial	-	-	43.883,0	5,65%

Tabla 44: Resumen de métricas por modelo para Unimarc.

Se observan los valores de AIC y BIC se desprende que el modelo ARIMA tiene mejores estimadores, debido a que tiene un menor valor de estas métricas. En cuanto al MSE y MAPE, el modelo ARIMA también es el modelo con menor valor, por lo que se asume que tiene errores más bajos en general.

Finalmente, se concluye que el mejor modelo es el ARIMA, ya que logra cumplir con los pedidos del negocio y además es el con menor error según las métricas.

6.5.5 Modelos de otras cadenas

Al igual que los modelos para Unimarc, se aprecia en la tabla 45 que el único modelo para las otras cadenas que cumple con tener errores de menos del 10% es el modelo ARIMA. Por lo que, es el único modelo que es capaz de cumplir con las peticiones del negocio.

Mes del 2022	ARIMA	Regresión lineal	Alisamiento exponencial
Marzo	-1,5%	-3,3%	10,6%
Abril	0,8%	7,9%	-1,0%
Mayo	2,2%	4,2%	7,9%
Junio	5,4%	14,6%	12,6%
Julio	2,8%	12,7%	5,6%
agosto	-3,1%	3,4%	8,9%

Tabla 45: Resumen de diferencia porcentual por modelo para otras cadenas.

A pesar de que el modelo ARIMA es el único con errores menores al 10%, de igual manera se realiza un análisis estadístico, para entender si a pesar de tener errores menores igual es mejor modelo en cuanto a métricas, lo que se puede ver en la tabla 46.

Modelos	AIC	BIC	MSE	MAPE
ARIMA	199,01	203,53	22.069,98	2,65%
Regresión Lineal	380,91	387,20	180.078,7	7,68%
Alisamiento Exponencial	-	-	201.997,4	7,77%

Tabla 46: Resumen de métricas por modelo para otras cadenas.

Se observan los valores de AIC y BIC se desprende que el modelo ARIMA tiene mejores estimadores, debido a que tiene un menor valor de estas métricas. En cuanto al MSE y MAPE, el modelo ARIMA también es el modelo con menores valores, por lo que se asume que tiene errores más bajos en general.

Finalmente, se concluye que el mejor modelo es el ARIMA, ya que logra cumplir con los pedidos del negocio y además es el con menor error según las métricas.

6.6 Plan de Implementación

El plan de implementación consiste en generar un proceso automatizado que pueda obtener de forma mensual las variables externas del modelo, en particular el valor del M1, IMACEC y feriados irrenunciables para que no se necesiten personas para actualizar las tablas. Esto se puede realizar con Web Scraping, que permite extraer los valores de una tabla en una página web en este caso del banco central que es de donde se obtiene los valores del IMACEC y M1. En el caso de los feriados irrenunciables se pueden obtener con una api del gobierno que guarda los feriados desde el 2013 hasta el 2023 en formato json. Esto va a permitir generar la actualización de las tablas mensualmente y de esta manera el área de inteligencia del consumidor obtendrá la predicción de 3 meses todos los meses, esto permitirá que fijen metas trimestrales que podrán ser ajustadas mes a mes.

Una vez funcionando el proceso automatizado, se generará un correo que enviará la predicción mes a mes para los próximos tres meses a todas las áreas interesadas.

El área de Inteligencia del Consumidor realizará la estimación de metas, utilizando los datos del área de producción que tiene una estimación de cual será la cantidad de toneladas de carne de pollo que se tendrá cada mes, con lo que se ajustarán las metas.

En el caso del área de Supermercados una vez recibido el correo utilizaran la información para definir las campañas de marketing y las ofertas que utilizarán cada mes para las cadenas, por ejemplo, si se viene un mes con menor cantidad de demanda, pueden ajustar su campaña para aumentar esta venta, ya sea con una baja de precio o mayor visualización de los productos en góndola, etc.

Cabe destacar que este modelo debe ajustarse con el tiempo, por lo que es importante estar continuamente evaluando su desempeño, por lo que se implementará una automatización que calcule la métrica MAPE, MSE y que guarde la diferencia entre mayor entre los meses de predicción y los meses reales una vez que se tenga este dato. Lo que se mostrará en una plataforma de visualización como PowerBI, que permitirá ver la variación de estas métricas en el tiempo, cosa de detectar a tiempo cuando sea necesarios realizar un ajuste al modelo.

7. Discusión

La estimación de la demanda de carne de pollo en cadenas de supermercados es un aspecto crucial para la toma de decisiones estratégicas en la industria alimentaria.

Se destaca la necesidad de tomar en cuenta los precios de los sustitutos al realizar la estimación de la demanda de carne de pollo. Los consumidores tienen la capacidad de optar por otros productos similares, como carne de res, pescado u otras proteínas animales, en función de su precio relativo. Por lo tanto, los precios de los sustitutos pueden influir directa o indirectamente en la demanda de carne de pollo en supermercados. Al incluir los precios de los sustitutos como variables explicativas, se logra una estimación más precisa y completa de la demanda de carne de pollo, si bien los precios no son un aspecto abordado en este trabajo, si debiesen ser considerados para dar mayor profundidad al estudio y generar mayor valor para la organización.

Los modelos utilizados en este estudio, como ARIMA, regresiones lineales y alisamiento exponencial, pudieron ofrecer herramientas eficaces para abordar la relación entre los precios y la demanda de carne de pollo. El modelo ARIMA permitió capturar patrones estacionales y tendencias en los datos históricos de demanda y utilizar los precios en estos modelos podría ser una variable explicativa clave para obtener aún mejores resultados en esta herramienta. Las regresiones lineales podrían proporcionar una forma de cuantificar el impacto del precio y los sustitutos en la demanda de carne de pollo, controlando otros factores relevantes como los utilizados en este estudio.

Si bien estos modelos presentan ventajas en la estimación de la demanda, es importante tener en cuenta sus limitaciones. Estas incluyen la necesidad de contar con datos históricos confiables, la suposición de que las relaciones pasadas se mantendrán en el futuro y la incapacidad para prever factores externos impredecibles que puedan afectar la demanda, como cambios económicos o eventos inesperados.

8. Conclusiones

Actualmente el área inteligencia del consumidor de Agrosuper no cuenta con las herramientas para tener un proceso de estimación de demanda, lo realizan más por juicio de expertos o las herramientas y conocimiento que cuenten en el momento. De esta forma más la estimaciones que tienen de producción y ventas desde el negocio logran la definición de metas.

Las herramientas desarrolladas durante este proyecto permiten tener una predicción más precisa, lo que logrará tener un input nuevo para la toma de decisiones que nos solo será útil para la definición de metas, si no que puede aportar información para el área encargada de supermercados.

Se destaca la importancia de tener clara la información con la que se cuenta y los inputs que puede aportar esta en la organización. En este caso se compran datos a Nielsen que solo están siendo usados por un área en Agrosuper, siendo que podrían ser de aporte para el área de supermercados, el área que fija los precios (Revenue Management), podría generar conocimiento para los otros canales de ventas, como canal tradicional, etc.

Se logra realizar una caracterización de los datos y obtener la agrupación por cadena y el total nacional. Además de incluir variables externas como el IMACEC, M1 y feriados irrenunciables. Con esto se logran generar modelos con buenos resultados en la mayoría de los casos, con métricas de buenos resultados.

En general, el modelo ARIMA logra predecir mejor la demanda para cada agrupación, esto se puede deber a dos motivos. Por un lado, el modelo ARIMA tiene mayor cantidad de ajustes, lo que permite generar una mayor cantidad de modelos distintos para cada cadena con las mismas variables. Por otro lado, a que las series de tiempo no son muy complejas, por lo que para un modelo complejo como lo es el modelo ARIMA, logra ajustar con mayor precisión los modelos.

Se logra disminuir en un 60% el error del modelo total nacional comparado con que fue realizado para el año 2020. No se realiza la comparación con los otros modelos, debido a que periodos anteriores no fueron realizados modelos por cadena.

Finalmente, se destaca de este trabajo que Agrosuper está continuamente intentando encontrar maneras de generar información útil para la toma de decisiones dentro de la organización. Pero aún quedan bastantes oportunidades de proyectos de datos que sean útiles dentro de la organización.

9.Recomendaciones de trabajo futuro

Primero, se recomienda volver a realizar este tipo de modelos cuando se obtenga más información, esto debido a que solo se están utilizando la información mensual desde enero del 2020, esto significa que todo los datos utilizados en entrenamiento corresponden a tiempos de pandemia, donde hubo un gran cambio en el comportamiento de compra de las personas, por lo que sería útil generar un análisis más amplio.

Segundo, se podría unificar más las áreas para lograr que las metas fijadas tengan relación con todas las otras partes de Agrosuper, la producción, la exportación de productos, los precios fijados, entre otras.

Tercero, se recomienda incluir los otros canales de venta de Agrosuper, debido a que son competencia entre ellos, ya que si una persona va a comprar pollo puede hacerlo en un Supermercado o en un negocio de barrio, no comprará en los dos establecimiento al mismo tiempo.

Cuarto, se recomienda estudiar el consumo de carnes per cápita, debido a que cada vez existen más personas veganas y vegetarianas, lo que afecta directamente el mercado. Este es un ámbito que también podría incluirse en futuros estudios.

Quinto, se debe utilizar de mejor manera los datos obtenidos por Agrosuper ya que estos pueden servir para la toma de decisiones no solo del área de inteligencia del consumidor, si no que igual podrían ser útiles para áreas como Revenue Management y el canal super-mercados por completo. Además, con estos datos se pueden realizar estudios por zona geográfica, ya que el comportamiento de compra de los clientes puede cambiar según en que parte del país se estudie.

Sexto, buscar una forma por medio de los otros datos externos que tienen dentro de la organización de separar la venta de Walmart y Tottus, para generar un modelo específico para cada cadena que logre ser representativo del comportamiento de compra de cada tipo de cliente. Ya que el modelo otras cadena es el que tiene peores resultados y puede deberse a que cada cadena tiene públicos objetivos distintos, que son incentivados por diversos factores.

Finalmente, se recomienda general un análisis de la elasticidad de los compradores de pollo en supermercados, ya que esto explicaría más del comportamiento de compra de los consumidores, lo que se podría incluir en los distintos estudios.

Bibliografía

1. 4 Metodologías para proyectos de Data Science – INVESTIGACIÓN DATLAS. (2021, 21 octubre). Blog Datlas. Recuperado 9 de septiembre de 2022, de <https://blogdatlas.wordpress.com/2020/02/16/4-metodologias-para-proyectos-de-data-science-datlas-research/>
2. Agrosuper. (2020, 7 octubre). Nuestra Empresa. Recuperado 15 de agosto de 2022, de <https://www.agrosuper.cl/gobierno-corporativo/nuestra-empresa/>
3. Análisis sectorial. (2022, 26 mayo). ASPROCER. <http://www.asprocer.cl/industria/analisis-sectorial/>
4. Brockwell, P. J. & Davis, R. A. (2016, 31 agosto). Introduction to Time Series and Forecasting (Springer Texts in Statistics) (3rd ed. 2016). Springer.
5. Caridad, J. C. & Ocerin, J. O. (s. f.). *ECONOMETRÍA: MODELOS ECONOMÉTRICOS Y SERIES TEMPORALES*.
6. Castro, B. E. (2020, 20 octubre). Rediseño de un proceso de estimación de la demanda para Guallarauco. Recuperado 9 de septiembre de 2022, de <https://repositorio.uchile.cl/handle/2250/177241>
7. CLAVERO BARRANQUERO, A. C. B. (2005). Una revisión de modelos econométricos aplicados al análisis de demanda y utilización de servicios sanitarios. Hacienda Pública Española.
8. colaboradores de Wikipedia. (2022, 26 marzo). Regresión lineal. Wikipedia, la enciclopedia libre. Recuperado 9 de septiembre de 2022, de https://es.wikipedia.org/wiki/Regresi%C3%B3n_lineal
9. Data, S. B. (2020, 20 enero). Aprendizaje automatico y las Metricas de regresión. sitio-bigdata.com. Recuperado 9 de septiembre de 2022, de <https://sitiobigdata.com/2018/08/27/machine-learning-metricas-regresion-mse/>
10. Factores asociados al éxito de los estudiantes en modalidad de aprendizaje en línea: un análisis en minería de datos. (s. f.). Recuperado 9 de septiembre de 2022, de https://www.scielo.cl/scielo.php?pid=S0718-50062020000600023&script=sci_arttext_plus&tlng=es
11. González, M. G. (s. f.). Análisis de series temporales: Modelos ARIMA.
12. *Home - Banco Central de Chile*. (s. f.). <https://www.bcentral.cl/en/web/banco-central>
13. IBM. (s. f.). IBM Analytics Metodología Fundamental para la Ciencia de Datos [Diapositivas].

14. IBM. (s. f.-b). Modelos personalizados de suavizado exponencial. Recuperado 9 de septiembre de 2022, de <https://www.ibm.com/docs/es/spss-statistics/saas?topic=modeler-custom-exponential-smoothing-models>
15. Jaime, A. A. (1994, 15 junio). Introducción al tratamiento de series temporales: Aplicación a las ciencias de la salud (Spanish Edition) (1.a ed.). Ediciones Díaz de Santos, S.A.
16. John Villavicencio. (s. f.). Introducción a Series de Tiempo.
17. José Alberto Mauricio, J. A. M. (s. f.). Introducción al Análisis de Series Temporales.
18. Kido-Cruz, A. K. (2005). Estimación de un Modelo de Demanda Casi Ideal para Determinar Cambios en la Estructura de Consumo de Carnes en los Estados Unidos de América.
19. Lin, T. Y., Xie, Y., Wasilewska, A. & Liau, C. (2010, 23 diciembre). Data Mining: Foundations and Practice (Studies in Computational Intelligence, 118) (Softcover reprint of hardcover 1st ed. 2008). Springer.
20. Lilien, Gary & Rangaswamy, Arvind. (2008). Marketing Engineering: Models that Connect with Practice. 10.1007/978-0-387-78213-3_16.
21. Linear regression. (s. f.). Recuperado 9 de septiembre de 2022, de https://web.archive.org/web/20080222195200/http://www.curvefit.com/linear_regression.htm
22. Matriz Agrosuper. (2021). Reporte Integrado Matriz agrosuper.
23. MORALES ROJAS, D. M. R. (s. f.). PRONOSTICO DE LA DEMANDA POR MÉTODOS MULTIVARIANTES [Tesis]. UNIVERSIDAD NACIONAL DEL CENTRO DEL PERÚ.
24. Nielsen, A. (2019, 5 noviembre). Practical Time Series Analysis: Prediction with Statistics and Machine Learning (1.a ed.). O'Reilly Media.
25. Otero, J. O. (s. f.). Econometría series temporales y predicción.
26. Pascual, C. R. (1990, 25 octubre). Previsión Tecnológica y de la Demanda. Productica 43 (Spanish Edition) (1.a ed.). Marcombo.
27. Pepió, M. P. (s. f.). Series temporales.
28. Pereira, A. P. G. (s. f.). ANALISIS PREDICTIVO DE DATOS MEDIANTE TECNICAS DE REGRESION ESTADISTICA [Tesis]. Univesidad Complutense de Madrid.
29. Preferencias y Tendencias del Consumo de Alimentos en Chile. (2021, noviembre). Deloitte. Recuperado 4 de marzo de 2023, de <https://www2.deloitte.com/content/dam/Deloitte/cl/Documents/customer-marketing/cl-preferencias-y-tendencias-de-los-alimentos-en-chile.pdf>
30. RAMÍREZ RODRÍGUEZ, I. R. R. (s. f.). MEJORA DEL PROCESO DE PRONÓSTICO DE DEMANDA EN PRODUCTOS DE CONSUMO MASIVO EN EL MERCADO NACIONAL DE AGROSUPER, MEDIANTE UN ENFOQUE DE REDISEÑO DE PROCESOS Y LA

APLICACIÓN DE TÉCNICAS DE DEEP LEARNING [Memoria de Título]. Universidad de Chile.

31. Rectas de regresión. (s. f.). Recuperado 9 de septiembre de 2022, de http://descartes.cnice.mec.es/materiales_didacticos/bidimensional_lbarrios/regresion_est.htm

30. Rodriguez-Ravines, R. (2021, 5 febrero). ¿Qué Calidad Predictiva tiene el Modelo? Romy R. Ravines. Recuperado 9 de septiembre de 2022, de https://ravinesromy.org/post/post_201709_diagnosisrespuestacontinua/

31. S, P. (2010, 3 octubre). AIC and BIC. Difference Between. Recuperado 9 de septiembre de 2022, de <http://www.differencebetween.net/miscellaneous/difference-between-aic-and-bic/>

32. Silva, A. R. S. (2000). Estimación de la Función de Demanda de Pollo a Nivel Minorista para la República Argentina . Recuperado 9 de septiembre de 2022, de https://www.researchgate.net/publication/265686323_Estimacion_de_la_Funcion_de_Demanda_de_Pollo_a_Nivel_Minorista_para_la_Republica_Argentina

33. Susan, M. (2007, 20 febrero). ESTADISTICA PARA BIOLOGIA Y CIENCIAS DE LA SALUD. (EDICION REVISADA. ACTUALIZADA Y AMPLIADA) (Spanish Edition) (3.a ed.). McGraw-Hill Interamericana de España S.L.

34. THE DATA WAREHOUSING INSTITUTE. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. Journal of data warehousing, 5(4). <https://mineracaodados.files.wordpress.com/2012/04/the-crisp-dm-model-the-new-blueprint-for-data-mining-shearer-colin.pdf>

35. Tutoriales, G. (2016, 5 febrero). Error Porcentual Absoluto Medio (MAPE) en un Pronóstico de Demanda. Gestión de Operaciones. Recuperado 9 de septiembre de 2022, de <https://www.gestiondeoperaciones.net/proyeccion-de-demanda/error-porcentual-absoluto-medio-mape-en-un-pronostico-de-demanda/>

36. Vargas, A. V. (s. f.). Estadística descriptiva e inferencial.

37. VILLARROEL PONCE, I. V. P. (s. f.). REDISEÑO DE PROCESOS DEL ÁREA DE ADMINISTRACIÓN Y SOPORTE DE PRECIOS EN EMPRESA AGROSUPER [Memoria de Título]. Universidad de Chile.

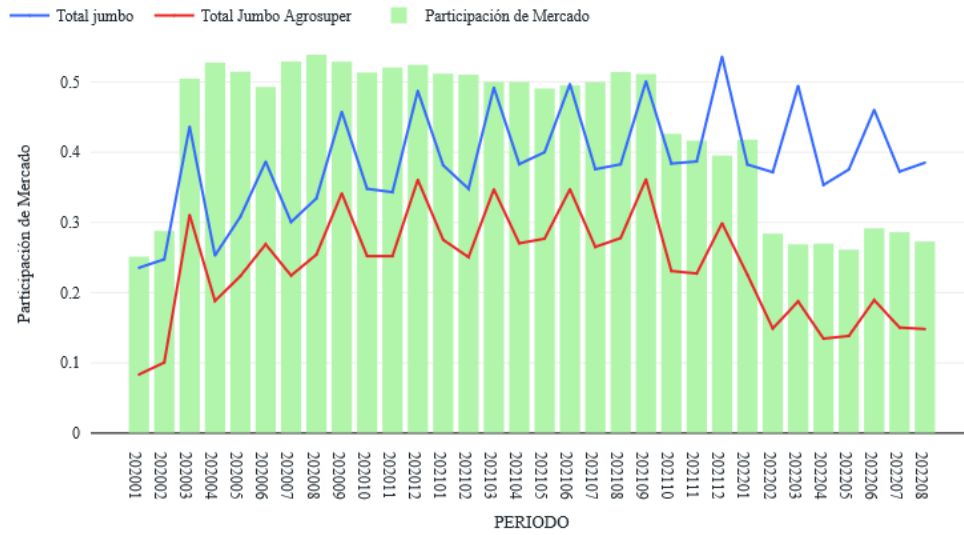
38. Zhang, Peter & Patuwo, Eddy & Hu, Michael. (1998). Forecasting With Artificial Neural Networks: The State of the Art. International Journal of Forecasting. 14. 35-62. 10.1016/S0169-2070(97)00044-7.

Anexos

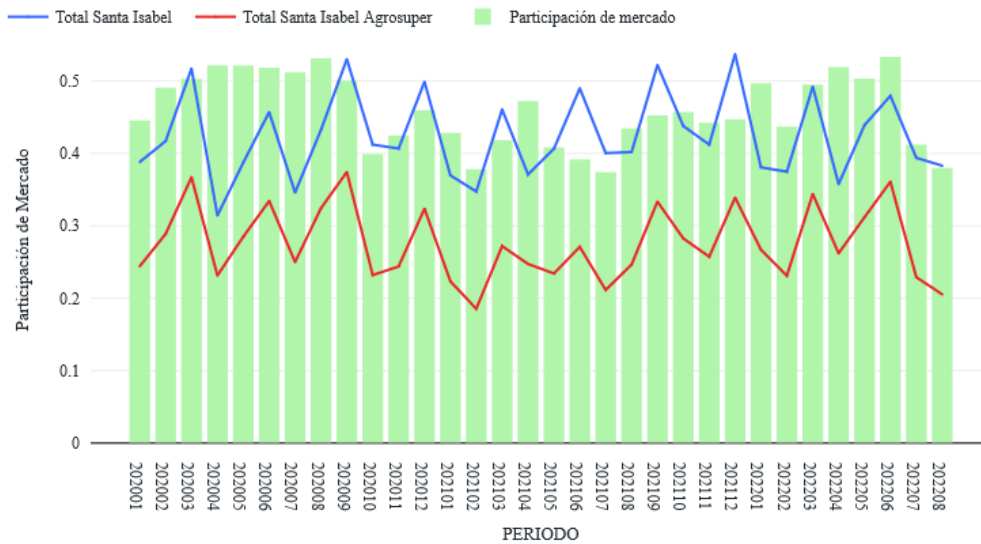
Anexo A

Participación de mercado y venta de Agrosuper por cadena:

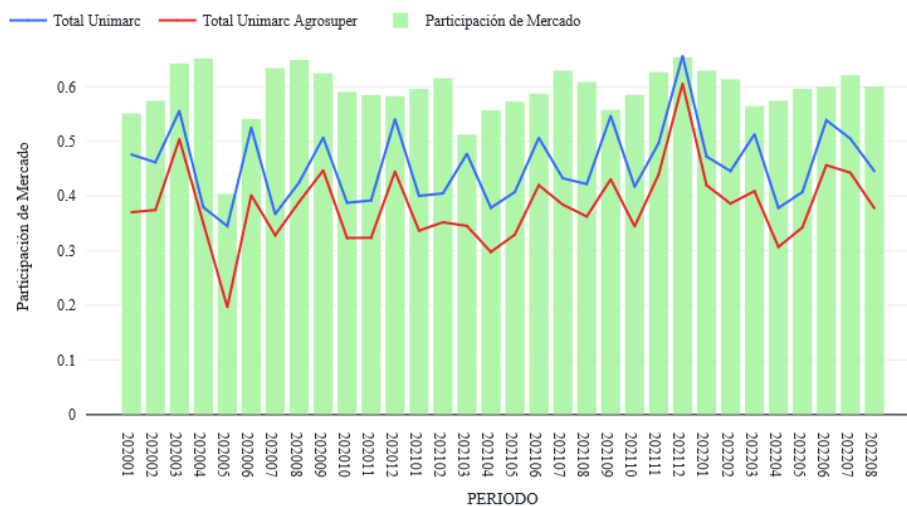
Anexo A.1. Jumbo



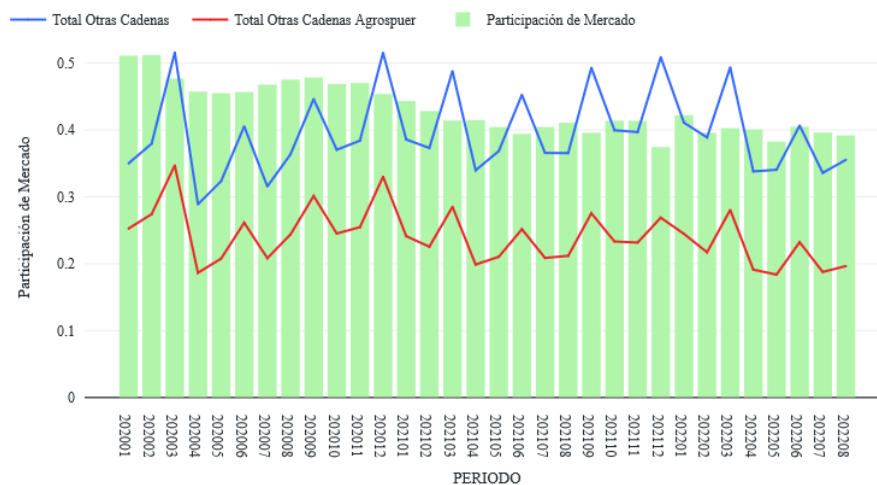
Anexo A.2. Santa Isabel



Anexo A.3. Unimarc



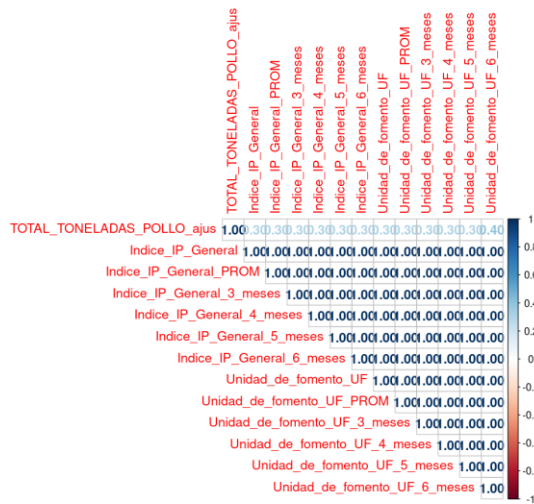
Anexo A.4. Otras Cadenas



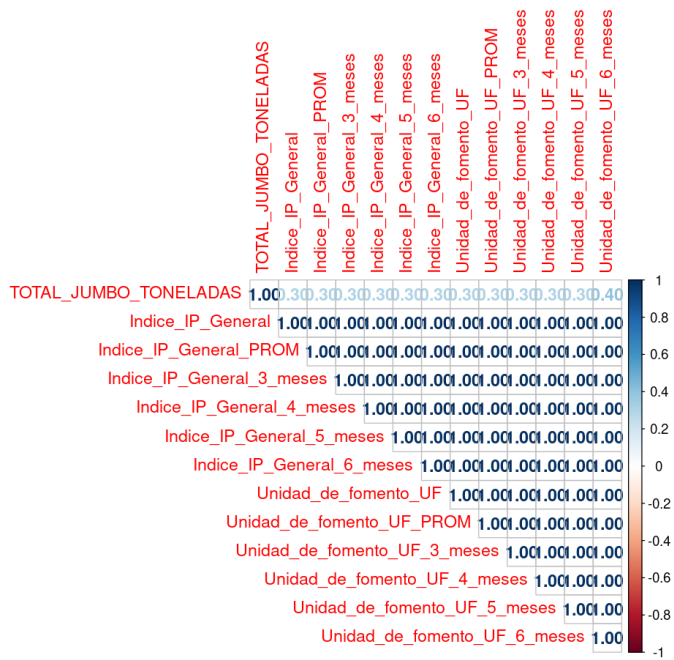
Anexo B

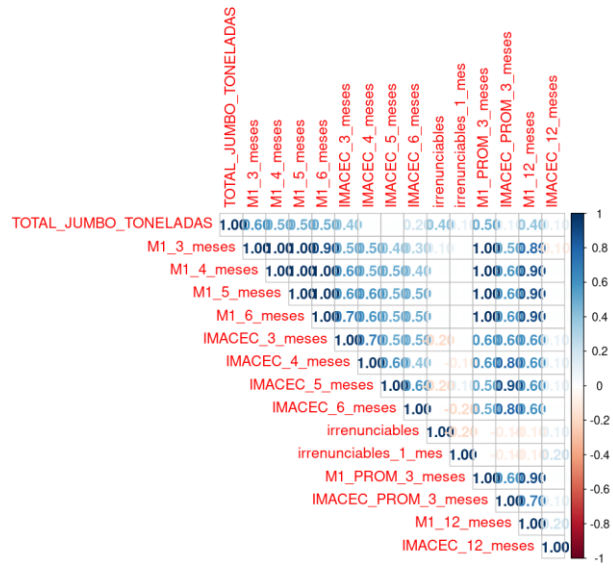
Correlaciones entre las variables utilizadas para la predicción de los modelos, separadas por cadenas.

Anexo B.1. Total nacional

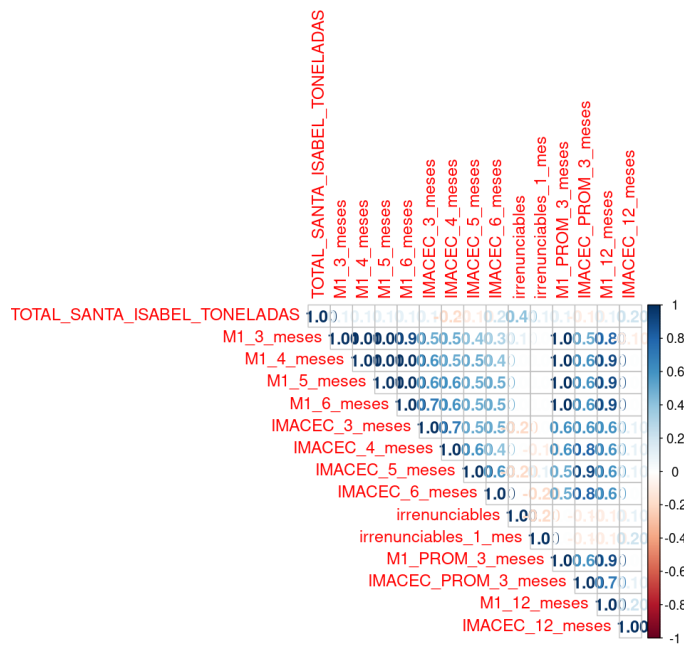


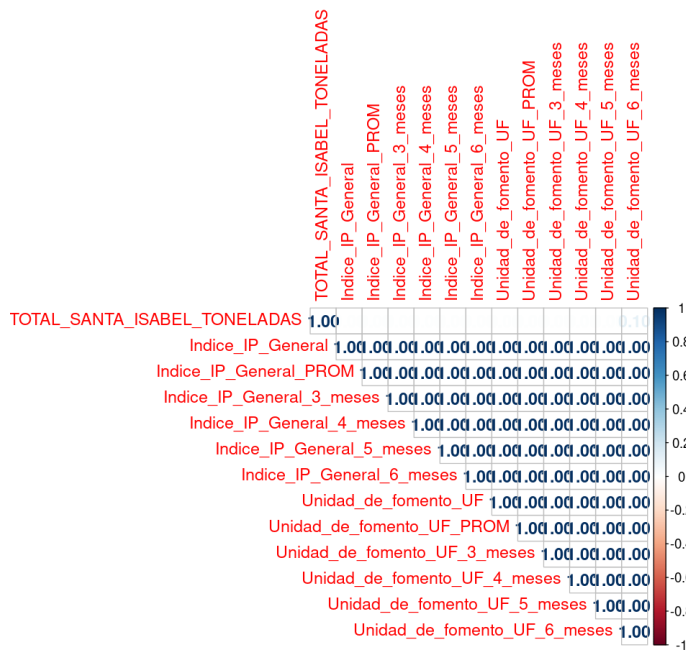
Anexo B.2. Jumbo





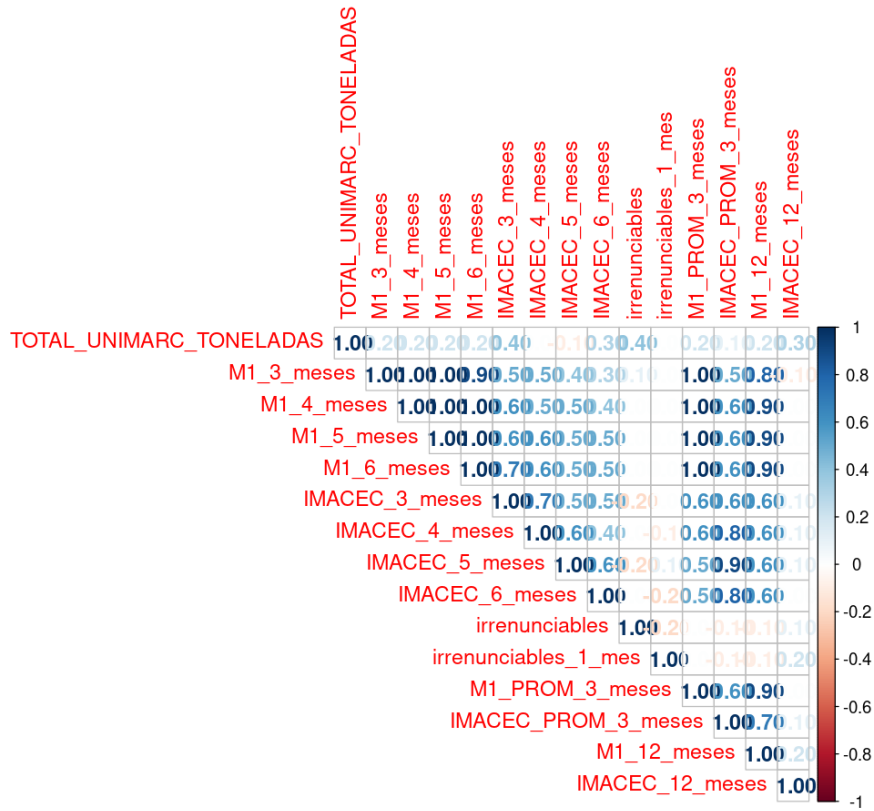
Anexo B.3. Santa Isabel





Anexo B.4. Unimarc





Anexo B.5. Otras Cadenas

