



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

TRABAJO DE FIN DE GRADO

FACULTAD DE EMPRESARIALES

**FACTORES QUE IMPACTAN EL PRONÓSTICO DE VENTAS EN EL SECTOR
MINORISTA: UNA APLICACIÓN DE MODELOS EXPLICATIVOS AL CASO
DE WALMART**

Autor: Jose María de Almansa Garrido

Curso: 5º E3-A

Director: Alex Escolá Gascón

Madrid

2024/2025

Resumen

El presente trabajo examina la capacidad de los factores macroeconómicos para mejorar el pronóstico de ventas semanales en Walmart, integrando registros históricos de tiendas con indicadores de inflación, desempleo, confianza del consumidor, carga financiera de los hogares y pendiente de la curva de tipos. A través de varios modelos de regresión lineal múltiple se demuestra que los indicadores macroeconómicos, considerados en conjunto, incrementan sustancialmente la explicación de la demanda frente a una especificación basada solo en variables internas. El desempleo se confirma como el elemento contractivo más consistente; la confianza del consumidor actúa como estímulo de la demanda; la inflación pierde protagonismo cuando se controlan los demás factores; y tanto la carga de deuda familiar como la pendiente de la curva de tipos resultan significativas, aunque con una dirección opuesta a la prevista, lo que apunta a un posible desplazamiento del gasto hacia distribuidores de bajo precio en contextos macroeconómicos menos favorables. Las pruebas diagnósticas acreditan la solidez estadística del modelo y respaldan la conclusión de que el entorno macroeconómico, lejos de ser un mero telón de fondo, condiciona de forma decisiva el comportamiento de la demanda minorista.

Palabras clave: Pronóstico de ventas minoristas; indicadores macroeconómicos; regresión lineal múltiple; Walmart; modelos explicativos.

Abstract

The present study examines the ability of macro-economic factors to improve weekly-sales forecasting at Walmart by integrating historical store records with indicators of inflation, unemployment, consumer confidence, household debt ratio, and the yield-curve slope. Across several multiple-linear-regression models it is shown that external determinants, taken together, substantially raise the model's explanatory power compared with a specification that relies only on internal variables. Unemployment proves to be the most consistently contractionary factor; consumer confidence stimulates demand; inflation loses prominence once the other factors are controlled for; and both the household debt burden and the yield-curve slope are significant yet work in the opposite direction to that originally hypothesized, pointing to a possible shift in spending toward low-price retailers in financially stressed contexts. Diagnostic tests confirm the statistical significance of the model and underscore the conclusion that the macro-economic environment, far from being a mere backdrop, decisively conditions retail-demand behavior.

Keywords: Retail sales forecasting; macro-economic indicators; multiple linear regression; Walmart; demand modelling.

TABLA DE CONTENIDOS

1. INTRODUCCIÓN AL TEMA	7
1.1. Marco teórico.....	7
1.2. Análisis conceptual de investigaciones previas.....	7
1.2.1. Enfoques estadísticos tradicionales: alcance y limitaciones.....	8
1.2.2. Integración de variables externas: una tendencia consolidada	9
1.2.3. Aplicación de modelos predictivos al caso de Walmart.....	9
1.2.4. Convergencia de la literatura en las hipótesis del trabajo	10
2. OBJETIVOS E HIPÓTESIS	12
2.1. Objetivos de la investigación.....	12
2.2. Modelo inicial e hipótesis planteadas	12
3. METODOLOGÍA.....	14
3.1. Base de datos	14
3.1.1. Fuente de la base de datos	14
3.1.2. Limpieza y preprocesamiento de datos	14
3.2. Selección de variables	15
3.2.1. Descripción de variables.....	15
3.2.2. Transformaciones para captar relaciones no lineales	17
3.3. Análisis descriptivo de las variables.....	17
3.3.1. Tabla de estadísticos resumen	17
3.3.2. Matriz de correlación.....	19
4. DESARROLLO Y RESULTADOS	21
4.1. Modelos entrenados	21
4.2. Análisis sobre el modelo	25
4.2.1. Multicolinealidad.....	25
4.2.2. Heterocedasticidad.....	27
4.3. Modelo de elección.....	27
4.4. Predicciones del modelo.....	28
4.5. Verificación de las hipótesis de investigación.....	29
5. DISCUSIÓN.....	32
5.1. Interpretación de estimaciones de los parámetros	32
5.2. Relación del análisis con el marco teórico	33
5.3. Implicaciones metodológicas y sociales: proyección más allá del caso Walmart.	34
6. DECLARACIÓN DE USO DE INTELIGENCIA ARTIFICIAL GENERATIVA	35
7. REFERENCIAS	36

LISTADO DE ABREVIATURAS

- ARMAV: Autoregressive Moving Average Vector (Modelo Vectorial Autorregresivo de Medias Móviles)
- BIC: Bayesian Information Criterion (Criterio de Schwarz)
- CCI: Consumer Confidence Index (Índice de Confianza del Consumidor)
- CPI / IPC: Consumer Price Index (Índice de Precios al Consumidor)
- CSI: Consumer Sentiment Index (Índice de Sentimiento del Consumidor)
- DSR: Debt-Service Ratio (Ratio de Servicio de la Deuda de los Hogares)
- EMD: Empirical Mode Decomposition (Descomposición en Modo Empírico)
- GDP / PIB: Gross Domestic Product (Producto Interior Bruto)
- LSTM Long Short-Term Memory (tipo de red neuronal recurrente)
- MCO / OLS: Mínimos Cuadrados Ordinarios / Ordinary Least Squares
- RF: Random Forest (Bosque Aleatorio)
- RLSE: Recursive Least-Squares Estimation (Estimación Recursiva de Mínimos Cuadrados)
- SARIMA: Seasonal Autoregressive Integrated Moving Average (Modelo Autorregresivo-Media Móvil Integrado Estacional)
- TDSP Household Debt Service Payments as a Percent of Disposable Personal Income (Pagos del Servicio de la Deuda de los Hogares como % de la Renta Disponible)
- VIF: Variance Inflation Factor (Factor de Inflación de la Varianza)
- XGBoost: Extreme Gradient Boosting (algoritmo de gradiente extremo)
- 10Y-3M (Spread) : Diferencial entre la rentabilidad del bono del Tesoro de EE. UU. a 10 años y el de 3 meses

1. INTRODUCCIÓN AL TEMA

1.1. Marco teórico

En un entorno minorista caracterizado por su dinamismo e intensa competencia, la precisión en el pronóstico de la demanda se ha convertido en un elemento estratégico clave. Una estimación adecuada permite optimizar la gestión del inventario, reducir costes operativos y mejorar la capacidad de respuesta ante fluctuaciones de consumo evitando tanto exceso de inventario como roturas de stock. No obstante, alcanzar un nivel adecuado de precisión, sigue siendo un reto, particularmente en mercados no consolidados o ante productos innovadores, donde el comportamiento del consumidor resulta menos predecible (Chern et al., 2010).

En este contexto, han proliferado el uso de modelos avanzados de regresión con el objetivo de mejorar la capacidad predictiva. Pese a sus avances, estos enfoques suelen prescindir de variables macroeconómicas, a pesar de la evidencia que vincula estos factores a una mayor predictibilidad de los modelos (Tseng & Turkmen, 2024).

El presente trabajo se iniciará con una revisión de la literatura, la cual permitirá identificar y delimitar las principales lagunas en la aplicación de modelos de predicción de la demanda. Estas carencias se trasladarán al estudio del caso de Walmart, con el objetivo de evaluar empíricamente si la combinación de datos históricos con variables macroeconómicas aporta un valor añadido significativo al análisis de la demanda en el sector minorista.

Más allá de su contribución técnica, este estudio busca reforzar el vínculo analítico entre las condiciones económicas generales y el comportamiento del consumidor, avanzando así hacia modelos de predicción más integradores y representativos.

1.2. Análisis conceptual de investigaciones previas

La literatura académica ha documentado una evolución metodológica clara en este ámbito, desde enfoques estadísticos tradicionales hacia modelos avanzados basados en algoritmos de aprendizaje automático (Tseng & Turkmen, 2024). De forma paralela, se ha consolidado una línea de investigación que promueve la incorporación de variables

macroeconómicas capaces de enriquecer los modelos predictivos y mejorar su precisión en contextos como el comercio minorista (Lin, 2024).

1.2.1. Enfoques estadísticos tradicionales: alcance y limitaciones

Los métodos tradicionales de predicción han sido ampliamente utilizados por su simplicidad, solidez teórica y capacidad para captar patrones estacionales. Estos se pueden clasificar en subjetivos y objetivos. Los modelos subjetivos surgen ante la inexistencia de datos históricos y se clasifican en críticos, basados en el juicio de un experto, o en experimentales, donde se encuestan a clientes para hacer predicciones. Por su parte, los modelos objetivos tradicionales, son aquellos enfocados en series temporales cuando existen datos históricos (Tseng & Turkmen, 2024). Modelos como Holt, Holt-Winters o SARIMA han sido particularmente relevantes, por su capacidad de capturar componentes de tendencia y estacionalidad. Este último, por ejemplo, ha mostrado buenos resultados en sectores como el tecnológico (Chuang, 1991) o el de bebidas alcohólicas (Velos et al., 2020), gracias a su capacidad para modelar series temporales con componentes estacionales. Asimismo, los enfoques causales, como la regresión lineal, han permitido explorar relaciones explicativas entre variables internas del sistema (Seyedan & Mafakheri, 2020).

A pesar de funcionar estos métodos tradicionales de predicción para productos con largos históricos y bajas complejidades técnicas, estos presentan dificultades para adaptarse a nuevos productos con falta de datos de demanda, como el lanzamiento de nuevos productos (Goodwin et al., 2014; Lynn et al., 1999). En estos casos, se han utilizado técnicas como las sesiones Delphi (Caplice & Ponce, 2023), donde se consideran opiniones de experto para futuras predicciones, o los modelos análisis por analogía, que utilizan datos de lanzamientos de productos similares (Kahn, 2002).

En respuesta a las limitaciones de los enfoques tradicionales, la literatura ha evidenciado una mejora en el rendimiento por los modelos de previsión basados en técnicas de aprendizaje automático por su capacidad para identificar patrones complejos, no lineales y de alta dimensionalidad (véase, e.g., Smirnov & Sudakov, 2020; Hamoudia & Vanston, 2023; Amar et al., 2022).

1.2.2. Integración de variables externas: una tendencia consolidada

Más allá del tipo de algoritmo utilizado, la literatura reciente subraya cómo la incorporación de variables exógenas emerge como una línea de trabajo clave para optimizar estos modelos predictivos.

Grzywińska-Rapca y Ptak-Chmielewska (2023) examinan la utilidad del CCI como predictor del gasto de los hogares, mientras que Ye y Le (2023) desarrollan un modelo de previsión de ventas de automóviles que incorpora renta disponible entre otros indicadores macroeconómicos, logrando una mejora significativa en la precisión del pronóstico.

Por último y más concretamente para mejorar la previsión de demanda minorista, Tseng y Turkmen (2024) destacan la relevancia de incorporar variables macroeconómicas en los modelos de predicción de demanda, especialmente en mercados no maduros. Frente a las limitaciones de los enfoques univariantes tradicionales, los autores demuestran que factores como el PIB, los salarios medios o la fiscalidad local mejoran significativamente la precisión y explicabilidad de los modelos.

1.2.3. Aplicación de modelos predictivos al caso de Walmart

La investigación sobre la predicción de ventas en Walmart ha evolucionado desde aproximaciones con regresión lineal clásica hasta marcos de aprendizaje automático que explotan grandes volúmenes de datos no estructurados.

En primer lugar, Wu et al. (2012) demostraron que un modelo ARMAV con tendencia lineal explicaba adecuadamente la demanda de productos nuevos, sentando así las bases para el análisis de series temporales en el retail.

A continuación, Singh et al. (2017)) aplicaron herramientas de big data (Spark y MapReduce) para procesar los históricos de Walmart, enfatizando la visualización interactiva como apoyo a la toma de decisiones. Su trabajo subraya que los precios del combustible y los periodos vacacionales concentran picos de venta siempre que los primeros se mantengan en rangos “razonables”.

Ese mismo año, Xu (2017) propuso un esquema híbrido (EMD-XGBoost-RLSE) que, tras descomponer la serie temporal, redujo la varianza de pronóstico, lo que reforzó la

tesis de que los modelos no lineales capturan patrones complejos ignorados por la regresión convencional.

Posteriormente, Jeswani (2021) mostró la superioridad del Gradient Boosting para predecir las ventas semanales y halló correlaciones ligeramente positivas entre tamaño de tienda, periodos vacacionales y ventas, así como una débil asociación con el desempleo.

Más tarde, Yao (2023) obtuvo mejoras sustanciales en la capacidad predictiva con los modelos de Random Forest y K-Nearest Neighbors tras un fine-tuning de las variables, subrayando la importancia de la selección de características.

Ya en 2024, Lin (2024), mediante regresión lineal múltiple, verificó los días festivos, el índice de precios al consumidor (IPC) y la tasa de desempleo presentan un efecto estadísticamente significativo sobre la facturación, confirmando la pertinencia de incorporar factores macroeconómicos en la modelización de la demanda.

Por último, Harsoor et al. (2024) evidenciaron la heterogeneidad espacial del consumo al segmentar la red de tiendas por ubicación geográfica y resaltaron la necesidad de ajustar los parámetros del modelo a escala local.

1.2.4. Convergencia de la literatura en las hipótesis del trabajo

Con todo, la literatura parece converger en favor de lo siguiente: (i) los indicadores macroeconómicos mejoran la capacidad explicativa; y (ii) la combinación de aprendizaje automático, técnicas de descomposición y *big data* amplía la precisión predictiva frente a los enfoques lineales puros. No obstante, la mayoría de trabajos se circunscribe a un conjunto limitado de variables exógenas, lo que deja margen para integrar factores adicionales y desarrollar modelos multivariantes más robustos, tal como se propone en la presente tesis.

En el presente trabajo se analizarán los factores que inciden en la demanda en el caso de Walmart. El caso de Walmart, otorga una base de datos con 3 principales virtudes: primero, la empresa genera series históricas extensas y desagregadas en múltiples formatos de tienda, lo que permite contrastar la respuesta de la demanda a distintos ciclos macroeconómicos. Segundo, su combinación de bienes de primera necesidad y consumo discrecional amplifica la sensibilidad a indicadores macroeconómicos, facilitando la

identificación de elasticidades diferenciadas. Tercero, la política de precios bajos permanentes reduce el ruido promocional y aísla el impacto de las variables exógenas.

Con base en las lagunas detectadas, el presente estudio se propone evaluar el valor incremental que aporta la incorporación de un espectro ampliado de indicadores macroeconómicos (inflación, desempleo, confianza del consumidor, endeudamiento de los hogares y pendiente de la curva de tipos) al pronóstico de ventas semanales de Walmart. El objetivo no es únicamente aumentar la precisión predictiva, sino dilucidar los canales causales a través de los cuales las fluctuaciones del ciclo económico se traducen en variaciones de la demanda minorista, aportando evidencia empírica que permita:

1. Cuantificar la contribución marginal de cada factor macroeconómico a la estimación de ventas semanales frente a un modelo basado exclusivamente en variables internas.
2. Comprobar las hipótesis básicas del modelo para garantizar la validez econométrica de las estimaciones de los valores asignados a los parámetros.
3. Ofrecer lineamientos prácticos para la toma de decisiones estratégicas para los agentes de mercado, tanto comerciantes, como responsables de política económica interesados en el consumo retail.

2. OBJETIVOS E HIPÓTESIS

2.1. Objetivos de la investigación

En coherencia con el marco teórico expuesto y la revisión de literatura realizada, el presente estudio se propone:

- Determinar el impacto individual de variables macroeconómicas en las predicciones de ventas semanales.
- Comparar la eficacia de predictiva de modelos que integran variables macroeconómicas aplicados al pronóstico de ventas minoristas con la de un modelo sin ellas.

2.2. Modelo inicial e hipótesis planteadas

El presente estudio aplicará la regresión lineal múltiple como método de análisis sobre las tiendas de Walmart como unidad observacional. Este modelo busca describir la relación entre las variables independientes y la variable dependiente ventas semanales. La ecuación del modelo inicial se expresa como sigue:

$$\text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{non_holiday} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + \beta_4 \cdot \text{cpi} + \beta_5 \cdot \text{unemployment} + \beta_6 \cdot \text{monthly_csi} + \beta_7 \cdot \text{quarterly_dsr_pct_di} + \beta_8 \cdot \text{spread_10a-3m} + u$$

Con el fin de mantener el foco analítico en la influencia de los factores macroeconómicos, no se formulan hipótesis específicas para las variables operativas internas: `holiday_flag`, `temperature` y `fuel_price`. Estas se incorporan exclusivamente como controles, destinados a aislar los efectos ajenos al objetivo principal del estudio.

Bajo este planteamiento, el modelo inicial conduce a las siguientes cinco hipótesis relativas a las variables macroeconómicas seleccionadas:

Hipótesis 1 – El Índice de Precios al Consumidor (CPI) tiene efecto significativo en la predicción de ventas semanales

$$H_0: \beta_4 = 0 ; H_1: \beta_4 \neq 0$$

Hipótesis 2 – La tasa de desempleo afecta negativamente a las ventas semanales.

$$H_0: \beta_5 \geq 0 ; H_1: \beta_5 < 0$$

Hipótesis 3 – El Índice de Sentimiento del Consumidor (monthly_csi) ejerce un efecto positivo y significativo sobre las ventas semanales de Walmart.

H₀: $\beta_6 \leq 0$; H₁: $\beta_6 > 0$

Hipótesis 4 – El ratio de deuda del hogar como porcentaje de la renta disponible (quarterly_dsr_pct_di) tiene un efecto negativo en las ventas.

H₀: $\beta_7 \geq 0$; H₁: $\beta_7 < 0$

Hipótesis 5 – La pendiente de la curva de tipos (diferencial entre bono a 10 años y a 3 meses) influye significativamente en las ventas semanales.

H₀: $\beta_8 = 0$; H₁: $\beta_8 \neq 0$

3. METODOLOGÍA

3.1. Base de datos

3.1.1. Fuente de la base de datos

La base de datos utilizada combina información interna de ventas con indicadores macroeconómicos externos. El núcleo procede del conjunto Walmart Store Sales disponible en Kaggle (M. Yasser H, 2022), que recoge 6435 observaciones de 45 establecimientos de Walmart en Estados Unidos durante 3 años. El archivo original (.csv) se transformó a formato Excel (.xlsx) para facilitar su manipulación y posterior análisis.

El conjunto de datos incluye ocho variables: Store (identificador numérico de la tienda), Date (fecha de la observación), Weekly_Sales (ventas semanales en dólares), Holiday_Flag (indicador binario de semana festiva), Temperature (temperatura media semanal en °F), Fuel_Price (precio medio semanal del combustible), CPI (Índice de Precios al Consumidor) y Unemployment (tasa de desempleo).

A este núcleo se añadieron tres series mensuales provenientes de la web oficial de la Reserva Federal (Federal Reserve Bank of St. Louis, 2025a, 2025b; University of Michigan, 2025). En primer lugar, el Índice de Sentimiento del Consumidor (CSI) refleja, mediante encuestas de percepción, las expectativas de gasto de los hogares y actúa como indicador previo de demanda. En segundo lugar, el Ratio de Servicio de la Deuda de los Hogares (DSR) cuantifica la porción de renta destinada a amortizar intereses y principal, revelando la holgura financiera real que condiciona el consumo. En tercer lugar, la pendiente de la curva de tipos (spread 10 años-3 meses) sintetiza las expectativas sobre política monetaria, anticipando inflexiones de la demanda.

La integración temporal se realizó asignando a cada observación semanal el valor mensual o trimestral más próximo de estas variables macroeconómicas, lo que permite evaluar el efecto combinado de la dinámica interna de ventas y las condiciones económicas agregadas.

3.1.2. Limpieza y preprocesamiento de datos

Para la manipulación y procesamiento del conjunto de datos se ha utilizado Python, concretamente las librerías pandas y sklearn.

En primer lugar, se creó una copia de seguridad para preservar la trazabilidad del conjunto de datos. Una vez garantizada, se suprimió la variable *Date* por carecer de valor analítico en el modelo propuesto. A continuación, se llevó a cabo una normalización léxica de los encabezados: todos los nombres de columna se convirtieron a “snake_case”, se eliminaron espacios superfluos y se homogenizaron caracteres especiales, con lo que se facilita tanto la legibilidad como el procesamiento posterior. Posteriormente, se comprobó que no había muestras nulas o incompatibles.

Por último, se aplicó codificación one-hot sobre la variable categórica *store*, descartando la categoría de referencia para evitar colinealidad perfecta e integrar dicha dimensión espacial en los algoritmos de aprendizaje supervisado. El resultado se almacenó sobrescribiendo la copia de respaldo, sobre la que se realizará el análisis estadístico subsecuente.

3.2. Selección de variables

3.2.1. Descripción de variables

Weekly_Sales – Variable endógena, numérica continua, medida en dólares. Representa el total de ventas semanales realizadas por cada tienda Walmart durante el período analizado. Esta variable es el foco del análisis y se busca explicar su comportamiento.

Holiday_Flag – Variable explicativa, dicotómica. Toma el valor 1 si la semana analizada contiene un día festivo significativo en Estados Unidos (*holiday* = 1), y 0 en caso contrario (*non-holiday* = 0). Se espera que esta variable se relacione de manera positiva con las ventas semanales, debido al impacto del consumo estacional asociado a festividades.

Temperature – Variable explicativa, numérica continua, medida en grados Fahrenheit. Indica la temperatura media registrada durante la semana correspondiente a cada observación. Puede afectar de forma directa o indirecta el comportamiento de los consumidores, en función de la estación del año y de los productos predominantes.

Fuel_Price – Variable explicativa, numérica continua, medida en dólares por galón. Representa el precio promedio semanal del combustible en la zona correspondiente. Se

espera que influya negativamente en las ventas semanales, dado que un aumento en el precio del combustible puede reducir el poder adquisitivo de los hogares o disuadir desplazamientos a tiendas físicas.

CPI – Variable explicativa, numérica continua. Este indicador mide la evolución del coste de una cesta representativa de bienes y servicios. Se espera que se relacione de forma negativa con las ventas semanales.

Unemployment – Variable explicativa, numérica continua, expresada como porcentaje de la población activa. Indica la tasa de desempleo semanal en la región correspondiente. Se anticipa una relación negativa con las ventas, ya que mayores tasas de desempleo suelen disminuir la renta disponible y, en consecuencia, el gasto en bienes.

Monthly_CSI – Variable explicativa, numérica continua. Esta variable, obtenida a partir de encuestas mensuales del University of Michigan, mide las expectativas de los consumidores respecto a la situación económica futura. Se espera que se relacione positivamente con las ventas, ya que una mayor confianza suele incentivar el consumo.

Quarterly_dsr_pct_di (Deuda como % de Renta Disponible) – Variable explicativa, numérica continua, expresada como porcentaje. Representa el porcentaje de los ingresos de los hogares que se destina al pago de deuda (intereses y principal). Se espera que se relacione negativamente con las ventas semanales, dado que una mayor carga de deuda reduce la capacidad de gasto.

Spread_10a_3m (Pendiente de la curva de tipos) – Variable explicativa, numérica continua. Calculada como la diferencia entre el rendimiento de los bonos del Tesoro a 10 años y a 3 meses, esta variable refleja las expectativas del mercado sobre la evolución económica. Una pendiente invertida (spread negativo) suele anticipar recesiones, por lo que se espera que la variable tenga un efecto positivo sobre las ventas minoristas.

Store – Variable categórica nominal entre 1 y 45, codificada como dicotómica. Representa el identificador numérico de cada tienda Walmart. Aunque no se incluye como variable explicativa directa, se ha codificado mediante one-hot encoding para capturar posibles efectos espaciales o estructurales entre establecimientos. La categoría de referencia se ha eliminado para evitar colinealidad.

3.2.2. Transformaciones para captar relaciones no lineales

Con el objetivo de capturar posibles relaciones no lineales entre determinadas variables explicativas y la variable dependiente, se han incorporado transformaciones funcionales específicas:

Temperature²: Dado que los efectos del clima sobre el comportamiento del consumidor pueden no ser lineales, se introduce el cuadrado de la temperatura para permitir que el modelo capte un efecto en forma de U invertida. Se espera así, que las ventas sean más altas en temperaturas moderadas y disminuyan cuando el clima es extremo.

Interacción Holiday_Flag × Fuel_Price: Se considera que el precio del combustible podría tener un efecto diferenciado en semanas festivas, donde el transporte del consumidor es mayor. Esta interacción permitirá evaluar si el coste del desplazamiento adquiere mayor relevancia en esos periodos.

3.3. Análisis descriptivo de las variables

3.3.1. Tabla de estadísticos resumen

Tabla 1

Tabla de Estadísticos Resumen

Variable	Media	Mediana	Mín	Máx	Desv. Típica.	C.V.	Asimetría	Exc. de curtosis
weekly_sales	1,046,964.88	960,746.04	209,986.25	3,818,686.45	564,366.62	54%	0.67	0.05
holiday_flag	0.07	0.00	0.00	1.00	0.26	365%	3.37	9.38
temperature	60.66	62.67	-2.06	100.14	18.44	30%	-0.34	-0.61
fuel_price	3.36	3.45	2.47	4.47	0.46	14%	-0.10	-1.18
cpi	171.58	182.62	126.06	227.23	39.36	23%	0.06	-1.84
unemployment	8.00	7.87	3.88	14.31	1.88	23%	1.19	2.64
monthly_csi	71.36	73.20	55.80	82.60	5.73	8%	-0.76	0.40
quarterly_dsr_pct_di	13.05	13.06	11.75	14.51	0.78	6%	0.12	-1.10
spread_10a-3m	2.46	2.47	0.00	3.80	0.78	32%	-0.43	0.16

La Tabla 1 describe el conjunto de datos con estadísticos. Cada fila corresponde a una variable del análisis, mientras que las columnas recogen sus medidas de tendencia central (media, mediana), dispersión (mínimo, máximo, desviación típica, coeficiente de variación) y forma de la distribución (asimetría, exceso de curtosis). Esta tabla busca describir el conjunto de datos antes de cualquier inferencia estadística.

En primer término, la variable dependiente weekly_sales presenta una media mayor que la mediana y un coeficiente de variación (54 %) notablemente elevado. Esta combinación, reforzada por una asimetría positiva moderada (+0,67), apunta a la existencia

de colas derechas asociadas a la diferencia de tamaño entre establecimientos y a episodios de ventas excepcionalmente altas, por ejemplo, semanas promocionales o picos estacionales. El exceso de curtosis próximo a cero (+0,05) indica, sin embargo, que estos valores extremos no se producen con una frecuencia suficiente para invalidar su tratamiento como observaciones válidas, si bien aconsejan contrastes de heterocedasticidad en la fase de estimación.

Por su parte, la *holiday_flag* cumple el patrón típico de las variables dicotómicas: una media del 0,07 se interpreta naturalmente como la proporción de semanas festivas en la muestra, mientras que la curtosis leptocúrtica (+9,38) y el coeficiente de variación ostensiblemente alto (365 %) no revelan inestabilidad, sino la concentración de la masa de la distribución en torno a cero con picos puntuales en el valor uno.

Entre las variables continuas exógenas destaca *temperature*, cuya asimetría ligeramente negativa (-0,34) y curtosis platicúrtica (-0,61) sugieren una distribución razonablemente simétrica y más plana que la normal. El rango amplio (-2 °F a 100 °F), unido a una dispersión relativa del 30 %, justifica la incorporación de un término cuadrático en el modelo para captar la posible relación cóncava entre clima y demanda detectada posteriormente. *Fuel_price* exhibe la dispersión relativa más baja (14 %) y una ligera asimetría negativa; su estabilidad temporal condiciona que su efecto se manifieste, más que de forma aislada, en interacción con variables como *holiday_flag*. El *CPI*, aunque conceptualmente central en las hipótesis, muestra una distribución prácticamente simétrica y sensiblemente platicúrtica.

La variable *unemployment* presenta una asimetría positiva destacada (+1,19) y un exceso de curtosis moderadamente alto (+2,64), reflejando un contexto en el que predominan niveles bajos de paro con puntuales episodios alcistas. A su vez, el *monthly_csi* y el *quarterly_dsr_pct_di* se caracterizan por coeficientes de variación mínimos (8 % y 6 % respectivamente) y curtosis suavemente negativas, lo que ofrece un trasfondo de gran estabilidad desde el que formular inferencias sobre su impacto.

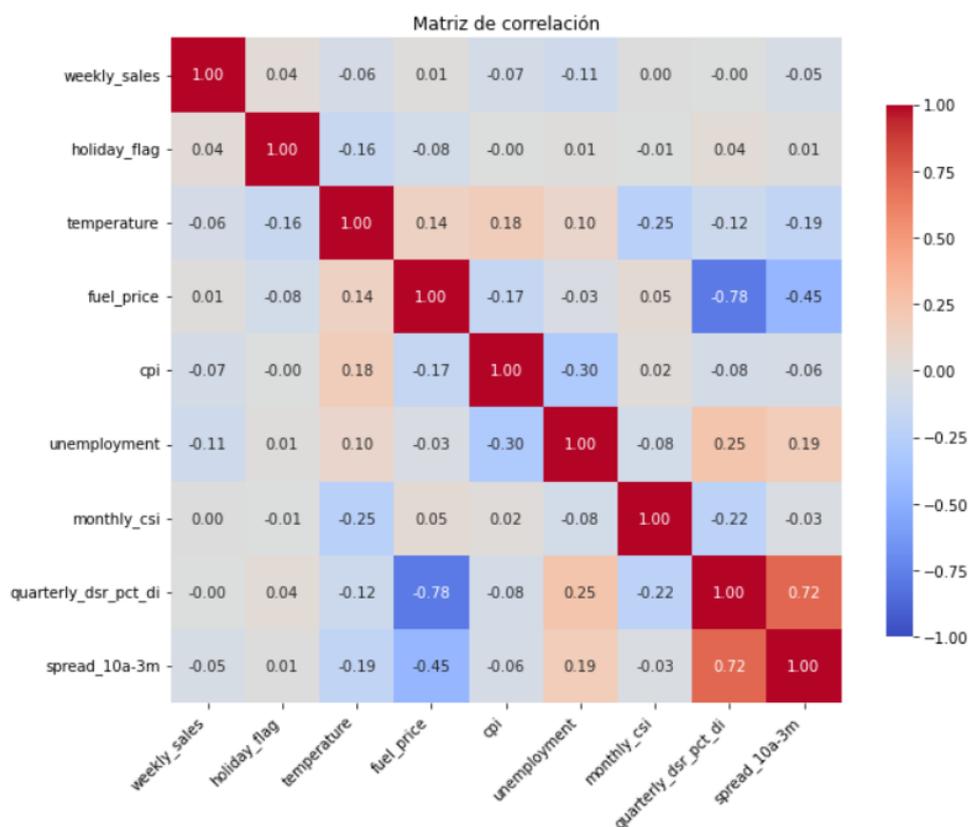
Por último, el *DSR* exhibe una distribución muy estable: media y mediana virtualmente iguales (13 %) y un coeficiente de variación del 6 % indican escasa dispersión, mientras que su asimetría levemente positiva (+0,12) y curtosis negativa (-1,10) apuntan a

una forma casi simétrica y algo aplanada; en contraste, el *spread 10a-3m* muestra mayor variabilidad (CV 32 %), con asimetría moderada a la izquierda (-0,43) y curtosis casi nula (+0,16), rasgos que describen colas algo más largas en valores bajos pero, en conjunto, una distribución próxima a la normal.

3.3.2. Matriz de correlación

Figura 1

Matriz de correlación



La Figura 1 representa la matriz de correlación entre las variables. Esta muestra el coeficiente de correlación de Pearson, una medida de dependencia lineal entre dos variables numéricas, para todas las variables del modelo en una matriz cuadrada, escalar y simétrica. De la matriz se desprende que la mayoría de las variables presentan correlaciones débiles entre sí. No obstante, se identifican dos excepciones notables: la pendiente de la curva de tipos (*spread_10a_3m*) muestra una correlación moderada con el precio del combustible (*fuel_price*), mientras que el ratio de servicio de la deuda (*quarterly_dsr_pct_di*) presenta una correlación fuerte tanto con *spread_10a_3m* como con *fuel_price*.

Esta matriz, resulta útil para dos fines principales:

En primer lugar, contextualiza las hipótesis al confirmar la dirección teórica esperada de las variables independientes respecto a la dependiente: unemployment y fuel_price muestran correlaciones negativas débiles (-0,11 y -0,17) con las ventas, coherentes con la idea de que menos renta disponible o mayores costes de desplazamiento merman la demanda.

En segundo lugar, advierte de posibles colinealidades que podrían inflar la varianza de los estimadores (lo que hace que sea difícil en el modelo distinguir cuál de ellas explica qué parte de la variación en la variable dependiente): destacan la relación inversa fuerte entre fuel_price y quarterly_dsr_pct_di (-0,78) y la positiva entre este último y spread_10a-3m (+0,72); ambos valores, cercanos en magnitud al umbral de alerta ($|\rho| \approx 0,70$).

4. DESARROLLO Y RESULTADOS

4.1. Modelos entrenados

Recordemos el modelo inicial propuesto para estudiar la capacidad explicativa de las variables macroeconómicas en el pronóstico de ventas semanales para el caso de Walmart:

$$\text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{non_holiday} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + \beta_4 \cdot \text{cpi} + \beta_5 \cdot \text{unemployment} + \beta_6 \cdot \text{monthly_csi} + \beta_7 \cdot \text{quarterly_dsr_pct_di} + \beta_8 \cdot \text{spread_10a-3m} + \beta_9 \cdot \text{temperature_spr} + \beta_{10} \cdot \text{holiday_fuel} + u$$

Dado que la variable dependiente (*weekly_sales*) es de naturaleza numérica continua, se optó por el método de estimación por Mínimos Cuadrados Ordinarios (MCO), que permite minimizar la suma de los errores cuadráticos y obtener estimadores lineales, insesgados, óptimos y consistentes, siempre que se cumplan las hipótesis clásicas del modelo. Se destinó el 90 % de la muestra para entrenamiento (Training Set), reservando el 10 % restante para validación (Test Set).

Con el objetivo de identificar el modelo más robusto y parsimonioso, se ha desarrollado una secuencia de estimaciones progresivas en la que se ha mantenido constante la variable endógena (*weekly_sales*) y se han ido incorporando nuevas variables explicativas. Este procedimiento permite evaluar tanto la contribución individual de cada variable como su efecto sobre la especificación global del modelo.

Se consignan a continuación las ecuaciones estimadas, manteniendo la notación convencional para los parámetros (β_i) y el término de perturbación (u):

1. Modelo I: $\text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{holiday_flag} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + u$
2. Modelo II: $\text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{holiday_flag} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + \beta_4 \cdot \text{cpi} + \beta_5 \cdot \text{unemployment} + u$
3. Modelo III: $\text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{holiday_flag} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + \beta_4 \cdot \text{cpi} + \beta_5 \cdot \text{unemployment} + \beta_6 \cdot \text{monthly_csi} + \beta_7 \cdot \text{quarterly_dsr_pct_di} + \beta_8 \cdot \text{spread_10a-3m} + u$

$$4. \text{ Modelo IV: } \text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{holiday_flag} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + \beta_4 \cdot \text{cpi} + \beta_5 \cdot \text{unemployment} + \beta_6 \cdot \text{monthly_csi} + \beta_7 \cdot \text{quarterly_dsr_pct_di} + \beta_8 \cdot \text{spread_10a-3m} + \beta_9 \cdot \text{temperature}^2 + \beta_{10} \cdot \text{holiday_flag} \cdot \text{fuel_price} + u$$

El primer modelo, usa las variables del conjunto de datos original excluyendo totalmente las variables macroeconómicas, el segundo añade las variables macroeconómicas del conjunto inicial, el tercero añade las tres variables adicionales del estudio y el último las variables transformadas para la captura de relaciones no lineales.

Los cuatro modelos obtenidos (I-IV) se exponen en la siguiente tabla:

Tabla 2

Tabla de secuencia de modelos obtenidos

	Modelo I				Modelo II				Modelo III				Modelo IV			
	Coef.	Est. t	Valor p	Sign.	Coef.	Est. t	Valor p	Sign.	Coef.	Est. t	Valor p	Sign.	Coef.	Est. t	Valor p	Sign.
const	1626809.68	77.17	0	0	1277907.25	5.61	2.09E-08	***	53469.66	0.13	0.8927		111416.98	0.28	0.7789	
holiday_flag	71487.09	8.93	5.55E-19	***	69239.28	8.67	5.23E-18	***	62947.29	8.13	5.04E-16	***	0.08	0.00	1	
temperature	-825.84	-6.24	4.63E-10	***	-833.71	-6.33	2.65E-10	***	-1457.70	-10.66	2.67E-26	***	141.35	0.23	0.8193	
fuel_price	-6253.35	-1.33	0.1848		-41354.15	-5.77	8.09E-09	***	38104.76	4.29	0.0000183	***	34218.81	3.81	0.0001	***
cpi					2933.35	2.84	0.0045	***	2406.92	1.64	0.1002		2190.26	1.49	0.135	
unemployment					-22467.01	-5.24	0.000000169	***	-24680.27	-5.32	0.00000011	***	-25418.15	-5.47	4.78E-08	***
monthly_csi									873.52	2.24	0.025	**	771.91	1.97	0.0489	**
quarterly_dsr_pct_di									98722.14	11.28	3.07E-29	***	96457.96	10.98	8.73E-28	***
spread_10a3m									-84567.36	-20.91	5.9E-94	***	-83557.18	-20.59	2.95E-91	***
temperature_sq													-13.94	-2.66	0.0078	***
holiday_fuel													19879.94	1.07	0.2833	
Store Effects ⁽¹⁾		Yes				Yes				Yes				Yes		
R ²		0.9194				0.9201				0.9254				0.9255		
R ² corregido		0.9188				0.9195				0.9248				0.9249		
V-p de F		0.0000				0.0000				0.0000				0.0000		
Criterio de Schwarz		172924.01				172881.73				172466.89				172476.89		
Log-verosimilitud		-86251.54				-86221.63				-86001.05				-85996.89		

(1) Para información adicional sobre store_*, mirar Apéndice 1.

(2) * p < 0.10, ** p < 0.05, *** p < 0.01.

La Tabla 2 resume los coeficientes estimados, niveles de significación y métricas de ajuste de los cuatro modelos de regresión lineal múltiple desarrollados en este estudio. Esta secuencia permite evaluar cómo evoluciona la capacidad explicativa del modelo a medida que se incorporan variables macroeconómicas adicionales y transformaciones funcionales. El análisis comparado de los coeficientes y sus significados estadísticos permite extraer conclusiones sobre la evidencia de las hipótesis planteadas.

De nuestra secuencia de modelos observamos:

En primer lugar, la variable *unemployment* o destaca por su estabilidad y consistencia: es significativa al 1 % en todos los modelos en los que está presente (Modelos II, III y IV), y su coeficiente estimado se mantiene negativo en todas las especificaciones (mínimo de -25.418 en el Modelo VI y máximo de -22.467 en el Modelo II). Este patrón apoya con claridad la hipótesis 2 del trabajo, según la cual mayores niveles de desempleo contraen la demanda, lo que se traduce en menores ventas semanales. El efecto es además cuantitativamente relevante, con una reducción estimada de entre 22.000 y 25.000 dólares por punto porcentual adicional de desempleo.

En segundo lugar, *CPI*, muestra un comportamiento menos robusto. Aunque resulta estadísticamente significativo en los Modelos II, al 1 %, pierde significación cuando se incorporan otras variables macroeconómicas y las variables funcionales en el Modelo III y IV. En términos de dirección del efecto, el signo estimado es positivo, en contra de lo esperado, lo que puede interpretarse como una indicación de que Walmart capta parte del consumo en contextos inflacionarios, presumiblemente por su posicionamiento como minorista de bajo precio. No obstante, dado que esta significación se diluye en la especificación más completa, la evidencia empírica en favor de la H_1 resulta débil.

En tercer lugar, *monthly_csi* se detecta un efecto positivo y estadísticamente significativo al 5 % en los Modelos III y IV. Este resultado sugiere que las expectativas favorables de los hogares se traducen en un mayor gasto en el canal minorista, lo que confirma su papel como variable relevante en la modelización de la demanda.

Por último, los resultados relativos al *quarterly_dsr_pct_di*, que mide el peso de la deuda del hogar sobre la renta disponible, contradicen la hipótesis formulada: el coeficiente estimado es positivo y significativo al 1 % tanto en el Modelo III como en el IV. Lejos de reducir el consumo, una mayor carga financiera parece asociarse a un incremento de las ventas. Lo mismo sucede con el *spread 10a-3m*, cuya influencia sobre las ventas resulta también significativa y negativa, contraviniendo lo anticipado en la hipótesis 4. En ambos casos, estos signos opuestos a los esperados invitan a reflexionar sobre la relación entre las condiciones macroeconómicas adversas y el comportamiento de los consumidores.

Respecto a las variables de control, cabe destacar el comportamiento de *temperature*, cuya influencia es significativa y negativa en los tres primeros modelos. No

obstante, al introducir el término cuadrático *temperature_sq* en el Modelo IV, la significación de la variable lineal desaparece, mientras que el parámetro cuadrático mantiene un signo negativo y resulta significativo al 1 %. Esto indica que la relación entre temperatura y ventas presenta una forma cóncava, es decir, que las ventas son más elevadas en climas moderados y se reducen cuando la temperatura se aleja de estos valores óptimos, confirmando así la pertinencia de introducir transformaciones no lineales. Por último, la variable *fuel_price* aunque no es significativa en el Modelo I, adquiere significación estadística y signo negativo en el Modelo II; sin embargo, en los modelos siguientes el signo se invierte y el efecto estimado se vuelve positivo y significativo al 1 %. Esta inversión sugiere que el precio del combustible, una vez controladas las condiciones macroeconómicas agregadas, puede estar asociado a una mayor afluencia a tiendas con una oferta más completa, como Walmart. En cuanto al término de interacción *holiday × fuel*, incorporado únicamente en el Modelo VI, este no resulta significativo, lo que sugiere que el efecto del coste de transporte no varía de manera diferenciada durante las semanas festivas.

En relación con la validez global del modelo, el contraste F es altamente significativo en todas las especificaciones consideradas ($p\text{-valor} < 0,01$), lo que indica que, en conjunto, las variables explicativas introducidas mejoran de forma sustancial la capacidad predictiva sobre las ventas semanales. Esta validez estadística apoya la inclusión progresiva de variables macroeconómicas y transformaciones funcionales en el modelo de análisis.

Por último, en lo que respecta a la bondad del ajuste, se observa una mejora progresiva del coeficiente de determinación (R^2) a medida que se amplía el conjunto de variables, pasando de 0,919 en el Modelo I hasta 0,926 en el Modelo IV. No obstante, para evitar sobreajustes, resulta más pertinente analizar el R^2 ajustado, que también se incrementa hasta alcanzar su valor máximo en el Modelo IV (0,9249). Sin embargo, al considerar conjuntamente este indicador con el Criterio de Schwarz (BIC), el Modelo III presenta el mejor equilibrio entre ajuste y parsimonia, al registrar el valor de BIC más bajo de todos los modelos estimados. En consecuencia, se adopta este modelo como la especificación preferente del estudio, dado que logra capturar adecuadamente la influencia de los factores macroeconómicos sin incurrir en una complejidad excesiva.

4.2. Análisis sobre el modelo

A continuación verificaremos las hipótesis básicas del modelo para entender la idoneidad y validez de los métodos empleados y por tanto de los resultados – las estimaciones de los valores asignados a los parámetros por el método MCO.

4.2.1. Multicolinealidad

La multicolinealidad se clasifica en dos categorías: perfecta e imperfecta. En ambas, dos o más variables explicativas exhiben una relación lineal, total en la primera, parcial en la segunda. Cuando la multicolinealidad es perfecta, la estimación del modelo es inviable; si es imperfecta, los coeficientes obtenidos resultan imprecisos, inestables y carentes de interpretabilidad.

Para comprobar que el nuevo modelo no adolece de este problema, se ha aplicado el cálculo del Factor de Inflación de la Varianza (VIF, por sus siglas en inglés). Valores de VIF superiores a 10 se consideran indicativos de multicolinealidad severa. Los resultados de esta prueba se muestran en la siguiente figura.

Figura 2

Factores de Inflación de Varianza

```
Factores de inflación de varianza (VIF)
Mínimo valor posible = 1.0
Valores mayores que 10.0 pueden indicar un problema de colinealidad

holiday_flag  1.047
temperature   1.709
fuel_price    4.467
store_2       1.957
store_3       2.129
store_4       96.998
store_5       2.180
store_6       2.119
store_7       8.428
store_8       2.427
store_9       2.432
store_10      97.995
store_11      2.129
store_12     103.495
store_13      97.098
store_14     13.304
store_15     84.173
store_16      8.279
store_17     97.004
store_18     84.507
store_19     84.174
store_20      2.608
store_21      1.957
store_22     76.345
store_23     83.955
store_24     84.399
store_25      2.618
store_26     83.976
store_27     76.444
store_28     103.495
store_29     85.256
store_30      1.957
store_31      1.957
store_32      8.398
store_33     98.134
store_34     98.793
store_35     76.714
store_36      1.987
store_37      1.987
store_38     103.495
store_39      1.987
store_40     83.954
store_41      8.199
store_42     97.995
store_43      3.573
store_44     97.050
store_45     13.304
cpi          891.503
unemployment  20.372
monthly_csi   1.337
quarterly_dsr_pct_di 12.577
spread_10a3m  2.684

VIF(j) = 1/(1 - R(j)^2), donde R(j) es el coeficiente de correlación múltiple
entre la variable j y las demás variables independientes
```

Los resultados de esta prueba se presentan en la Figura 2. La mayoría de los regresores principales arrojan valores claramente inferiores al umbral ($VIF < 5$), lo que indica una correlación limitada con el resto de variables. Dos excepciones merecen atención: *unemployment* ($VIF \approx 20$) y *quarterly_dsr_pct_di* ($VIF \approx 13$) rebasan el umbral convencional de 10. Aunque estos valores sugieren cierta superposición de información con otros indicadores macroeconómicos, la estabilidad de sus coeficientes y su relevancia teórica justifican mantenerlos en la especificación principal.

Por otra parte, varias *dummies* de tienda (*store_**) exhiben VIF muy elevados (> 80), fenómeno previsto en la codificación one-hot cuando se controlan efectos fijos.

4.2.2. Heterocedasticidad

Para comprobar si se verifica la hipótesis de la homocedasticidad, se analizará la varianza de las perturbaciones aleatorias de los residuos. Así se aplicará la prueba de White; un contraste no paramétrico que evalúa la constancia de la varianza de los errores.

Figura 3

Prueba de White

```
Contraste de heterocedasticidad de White -  
Hipótesis nula: [No hay heterocedasticidad]  
Estadístico de contraste: LM = 2270.7  
con valor p = P(Chi-cuadrado(439) > 2270.7) = 5.19996e-244
```

Observamos en la Figura 3, un p-valor superior a un nivel de significación $\alpha = 10\%$, por lo que las diferencias entre los datos y la hipótesis no son significativas al $\alpha = 10\%$, no habiendo suficiente evidencia para descartar la H_0 (homocedasticidad). De ello se desprende que la varianza de los errores puede considerarse constante e igual para todas las observaciones, lo que respalda la validez de los estimadores MCO y la fiabilidad de las inferencias basadas en ellos.

4.3. Modelo de elección

Como se indicó en el apartado anterior, la selección del modelo final se sustenta en un análisis comparativo, con el objetivo de identificar aquel modelo que ofrezca el mejor equilibrio entre capacidad explicativa y simplicidad estructural. Para ello, se han considerado de forma conjunta varios indicadores de ajuste (R^2 y R^2 ajustado), penalización por complejidad (Criterio de Schwarz, BIC) y verosimilitud del modelo.

En términos concretos, el Modelo III alcanza un R^2 ajustado de 0,9248, superior al de las especificaciones previas, y un BIC de 1.724.687, inferior incluso al del Modelo IV, que añade variables funcionales adicionales sin obtener mejoras sustanciales en los indicadores de ajuste. La log-verosimilitud también mejora sensiblemente, lo que refuerza la consistencia interna del modelo seleccionado. Además, el contraste F de significación global muestra niveles altamente significativos en todos los modelos, validando que las variables incluidas aportan valor explicativo en conjunto.

A partir de este análisis comparado, se concluyó que el Modelo III es el más adecuado para responder a la pregunta de investigación. Su estructura permite integrar un conjunto representativo de variables macroeconómicas, sin introducir redundancias ni comprometer la robustez estadística.

El modelo estimado se expresa de la siguiente forma:

$$\text{weekly_sales} = \beta_0 + \beta_1 \cdot \text{non_holiday} + \beta_2 \cdot \text{temperature} + \beta_3 \cdot \text{fuel_price} + \beta_4 \cdot \text{cpi} + \beta_5 \cdot \text{unemployment} + \beta_6 \cdot \text{monthly_csi} + \beta_7 \cdot \text{quarterly_dsr_pct_di} + \beta_8 \cdot \text{spread_10a-3m} + \beta_9 \cdot \text{temperature_spr} + \beta_{10} \cdot \text{holiday_fuel} + u$$

Por último, la elección del modelo, se ve reforzada por la posterior verificación de las hipótesis clásicas del modelo lineal. En efecto, los diagnósticos realizados confirman la ausencia de multicolinealidad severa entre las variables principales ($VIF < 10$ en la mayoría de los regresores), así como la homocedasticidad de los residuos (según la prueba de White), lo que garantiza la validez de las estimaciones obtenidas por Mínimos Cuadrados Ordinarios.

4.4. Predicciones del modelo

Figura 4

Predicciones del Modelo III



La Figura 4 presenta visualmente la capacidad predictiva del Modelo III de regresión lineal múltiple, que ha sido seleccionado como la especificación preferente del

estudio por su equilibrio entre ajuste y parsimonia. En dicha gráfica se muestran: (i) la serie real de ventas semanales (*weekly_sales*, línea verde), (ii) las predicciones generadas por el modelo (predicción, línea naranja), y (iii) el intervalo de confianza al 95 % (área sombreada en gris).

En esta, se observa cómo las dos series mantienen una trayectoria paralela y convergente en los tramos de ventas estables; las discrepancias se concentran en los picos aislados (sobre-estimaciones puntuales en torno a la semana 6330 y sub-estimaciones en 6150), lo que sugiere que el modelo captura bien la tendencia central pero tiende a suavizar extremos.

A pesar de esto, el objeto de estudio de este análisis no es predecir las ventas semanales a través de una serie de variables, sino analizar cuáles de ellas son las que afectan de manera significativa a las mismas.

4.5.Verificación de las hipótesis de investigación

Hipótesis 1 – El Índice de Precios al Consumidor (CPI) tiene un efecto significativo sobre las ventas semanales.

$$H_0: \beta_4 = 0; H_1: \beta_4 \neq 0$$

Resultados. En el Modelo III el coeficiente asociado al *CPI* no es significativo al $\alpha = 0,10$ ($p = 0,1002$). Por tanto, no se rechaza H_0 ; la evidencia resulta insuficiente para afirmar que el *CPI* explique variaciones en las ventas. El efecto observado en el Modelo II desaparece al introducir controles adicionales, lo que sugiere que el resultado inicial podía deberse a omisión de variables relevantes.

Hipótesis 2 – La tasa de desempleo afecta negativamente a las ventas semanales.

$$H_0: \beta_5 \geq 0; H_1: \beta_5 < 0$$

Resultados. El coeficiente de *unemployment* es negativo y altamente significativo en todos los modelos ($p < 0,001$). Se rechaza H_0 con fuerte evidencia a favor de H_1 : un incremento de un punto porcentual en la tasa de paro reduce las ventas semanales de una tienda en ~24 500 dólares.

Hipótesis 3 – El Índice de Sentimiento del Consumidor (*monthly_csi*) ejerce un efecto positivo y significativo sobre las ventas semanales de Walmart.

$$H_0: \beta_6 \leq 0 ; H_1: \beta_6 > 0$$

Resultados. El parámetro asociado es significativo al 5 % ($p = 0,025$). Se rechaza H_0 con evidencia moderada en favor de H_1 y se confirma el efecto positivo previsto. Además, cada punto extra en el *monthly_csi* suma de media y ceteris paribus 8 900 USD a las ventas semanales.

Hipótesis 4 – El ratio de servicio de la deuda de los hogares (DSR) reduce las ventas semanales.

$$H_0: \beta_7 \geq 0; H_1: \beta_7 < 0$$

Resultados. El parámetro asociado a *quarterly_dsr_pct_di* es altamente significativo ($p < 0,001$) pero exhibe signo positivo, contrario a lo previsto. Por tanto, no se acepta H_1 ; la relación detectada sugiere que, en el periodo analizado, un mayor apalancamiento coincide con mayor gasto minorista.

Hipótesis 5 – La pendiente de la curva de tipos (*spread 10a-3m*) influye significativamente en las ventas.

$$H_0: \beta_8 = 0; H_1: \beta_8 \neq 0$$

Resultados. El coeficiente de *spread_10a3m* es significativo al $\alpha = 0,01$ en los modelos, pero con signo negativo, opuesto a la hipótesis inicial. Se rechaza H_0 ; existe efecto estadísticamente robusto, aunque la dirección sugiere que una pendiente más pronunciada (expectativas de expansión) coincide con moderación del gasto en Walmart.

Así podemos concluir:

- No hay suficiente evidencia en favor de la hipótesis 1.
- Hay evidencia fuerte en favor de la hipótesis 2; el desempleo es un predictor robusto y negativo.
- La hipótesis 3 y la hipótesis 4 muestran evidencia significativa pero de signo inverso al inicialmente hipotetizado, lo que invita a replantear la

interpretación económica de ambas variables en el contexto minorista analizado.

5. DISCUSIÓN

5.1. Interpretación de estimaciones de los parámetros

De los resultados comentados en el apartado anterior, podemos concluir que los resultados avalan la relevancia de los factores macroeconómicos, si bien subrayan que su impacto puede diferir de la intuición teórica.

La serie de modelos estimados confirma que los indicadores macroeconómicos mejoran de forma sustancial la capacidad explicativa: el modelo finalmente elegido explica más del 92 % de la variabilidad de las ventas y el BIC desciende de manera apreciable. En otras palabras, el contexto económico agrega un componente explicativo significativo.

Ahora bien, la dirección y la intensidad de algunos coeficientes difieren de la intuición teórica. El desempleo, por ejemplo, se comporta conforme a lo previsto, su incremento erosiona las ventas, pero la pendiente de la curva de tipos y el ratio de servicio de la deuda muestran signos opuestos a los anticipados. Este fenómeno podría leerse bajo dos claves:

- Estrategias de consumo defensivo. Cuando los hogares perciben riesgo macroeconómico, tipos de interés al alza o mayor apalancamiento, pueden re-optimizar la cesta y desplazarse hacia distribuidores de precio bajo. Walmart, con su política “everyday low prices”, se beneficia de ese trasvase aunque el entorno sea, en abstracto, adverso para el gasto agregado.
- Financiación de corto plazo y crédito al consumo. En periodos de tensión financiera, la compra minorista puede mantenerse mediante facilidades de crédito revolvente, programas “buy now, pay later” o tarjetas de fidelización. El aumento del DSR no implicaría, por tanto, una contracción inmediata de la demanda; más bien reflejaría la utilización de deuda como mecanismo de suavización del consumo.

Por su parte, la pérdida de significación del IPC al introducir controles adicionales sugiere que la inflación incide sobre las ventas a través de canales más complejos (por ejemplo, modificando la mezcla de productos o fomentando la compra de marcas propias) y no como un efecto directo y homogéneo.

5.2. Relación del análisis con el marco teórico

Los hallazgos empíricos confirman la tesis de Tseng & Turkmen (2024) y Lin (2024): incluir indicadores macroeconómicos refuerza la capacidad explicativa de los modelos de demanda. Ahora bien, nuestros coeficientes muestran que el efecto práctico de esos indicadores no es uniforme, sino que está condicionado por la estrategia competitiva del minorista.

En un entorno como el de Walmart, caracterizado por precios sistemáticamente bajos, marcas propias asequibles y líneas de crédito al consumo, los choques macroeconómicos tradicionales (p. ej., un repunte de la carga de deuda de los hogares o una curva de tipos más inclinada) no se traducen necesariamente en una contracción del gasto agregado dentro de la cadena. De hecho, pueden estimular un “efecto trasvase”: consumidores que, ante la presión sobre su renta disponible, sustituyen compras en establecimientos minoristas más caros por adquisiciones en establecimientos de precio agresivo. El resultado es una ganancia de cuota para el distribuidor de bajo costo que amortigua, o incluso revierte, la caída de volumen relativa al conjunto del mercado.

En términos macroeconómicos, esto matiza la lógica keynesiana clásica, según la cual un empeoramiento de las condiciones financieras reduce indiscriminadamente el consumo. Nuestros datos sugieren que el ajuste no es homogéneo: el gasto se redistribuye internamente entre formatos comerciales en lugar de disiparse por completo. Así, la sensibilidad de la demanda al ciclo económico depende tanto de la salud macro como del posicionamiento micro de cada retailer, un matiz que la literatura empieza a documentar pero que aún no se refleja de forma sistemática en los modelos de previsión tradicionales.

Como línea especulativa, y abierta a contrastación futura, cabe plantear que la elasticidad de la demanda minorista frente a ciertos choques macro varía no solo por la naturaleza de los bienes, sino también por la reputación de precio y la oferta de servicios financieros del retailer. En el caso de Walmart, ambos activos podrían amortiguar o incluso revertir el signo de variables que, en un distribuidor menos competitivo en precio, se manifestarían como claramente contractivas.

5.3. Implicaciones metodológicas y sociales: proyección más allá del caso

Walmart

En primer lugar, resulta pertinente subrayar la repercusión metodológica del estudio. Al constatar empíricamente que la incorporación sistemática de indicadores macroeconómicos incrementa el poder explicativo del modelo, se refuerza la idea, ya adelantada por la literatura, de que los análisis centrados exclusivamente en variables internas incurrir en sesgo por omisión. Asimismo, el procedimiento propuesto aporta una hoja de ruta replicable en sectores donde el consumo es sensible al ciclo (por ejemplo, hostelería, transporte o energía), dado que permite homogenizar la comparación entre empresas sometidas a marcos regulatorios heterogéneos y, así, construir escenarios contrafactuales más informativos para la planificación estratégica.

Ahora bien, el valor del enfoque no se limita a grandes corporaciones. De hecho, para cadenas regionales o plataformas de comercio electrónico con menor densidad estadística, la metodología sugiere priorizar indicadores macro de amplia disponibilidad y emplear proxies cuando el histórico sea escasa. En consecuencia, cualquier organización que disponga de series históricas mínimas puede beneficiarse de esta ampliación del marco analítico.

Desde un punto de vista social y de política pública, los reguladores y autoridades competentes podrían apoyarse en esta evidencia para decidir cómo afectar en los indicadores macroeconómicos.

Finalmente, la investigación abre varias líneas futuras. En primer término, comprobar si este “efecto trasvase” hacia cadenas de bajo precio se repite en otros formatos retail; En segundo término, añadir información sobre crédito y programas de fidelidad de los hogares para medir hasta qué punto la financiación minorista suaviza las caídas del consumo; y, por último, emplear modelos dinámicos más flexibles (p. ej., redes LSTM con variables exógenas) que capten retrasos y efectos no lineales y verifiquen si estos resultados persisten en distintos contextos.

6. DECLARACIÓN DE USO DE INTELIGENCIA ARTIFICIAL GENERATIVA

Por la presente, yo, Jose María de Almansa Garrido, estudiante de Derecho y Administración de Empresas de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "Factores que impactan el pronóstico de ventas en el sector minorista: una aplicación de modelos explicativos al caso de Walmart", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. Brainstorming de ideas de investigación: Utilizado para idear y esbozar posibles áreas de investigación.
2. Metodólogo: Para descubrir métodos aplicables a problemas específicos de investigación.
3. Interpretador de código: Para realizar análisis de datos preliminares.
4. Estudios multidisciplinares: Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
5. Corrector de estilo literario y de lenguaje: Para mejorar la calidad lingüística y estilística del texto.
6. Sintetizador y divulgador de libros complicados: Para resumir y comprender literatura compleja.
7. Traductor: Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 28/05/2025

Firma: Jose María de Almansa Garrido

7. REFERENCIAS

- Amar, J., Rahimi, S., Surak, Z., & von Bismarck, N. (2022, 15 de febrero). *AI-driven operations forecasting in data-light environments*. McKinsey & Company. Obtenida el 05/04/2025 de <https://www.mckinsey.com/capabilities/operations/ourinsights/ai-driven-operations-forecasting-in-data-light-environments>
- Caplice, C., & Ponce, E. (2023). *MITx MicroMasters Program in SCM Key Concepts*. MITx MicroMasters® Program in Supply Chain Management.
- Chern, C.-C., Ao Jeong, K., Wu, L.-L., & Kung, L.-C. (2010). Designing a decision-support system for new product sales forecasting. *Expert Systems with Applications*, 37(2), 1654-1665. <https://doi.org/10.1016/j.eswa.2009.06.042>
- Chuang, A. Z. (1991). Time series analysis: Univariate and multivariate methods. *Technometrics*, 33(1), 108-109. <https://doi.org/10.1080/00401706.1991.10484802>
- Federal Reserve Bank of St. Louis. (2025a). *10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity*. FRED. Obtenida el 05/04/2025 de <https://fred.stlouisfed.org/series/T10Y3M>
- Federal Reserve Bank of St. Louis. (2025b). *Household Debt Service Payments as a Percent of Disposable Personal Income*. FRED. Obtenida el 05/04/2025 de <https://fred.stlouisfed.org/series/TDSP>
- Grzywińska-Rapca, M., & Ptak-Chmielewska, A. (2023). Backward assessments or expectations: What determines the consumer confidence index more strongly? Panel model based on the CCI of European countries. *Wiadomości Statystyczne*, 68(2), 1-15. Obtenida el 05/04/2025 de <https://doi.org/10.5604/01.3001.0055.0575>
- Hamoudia, M., & Vanston, L. (2023). Machine learning for new product forecasting. En M. Hamoudia, S. Makridakis, & E. Spil (Eds.), *Forecasting with artificial intelligence* (pp. 77-106). Palgrave Macmillan.

- Harsoor, A. S., & Patil, A. (2024). *Forecast of sales of Walmart Store using big data applications*.
- Jeswani, R. (2021). *Predicting Walmart sales, exploratory data analysis, and Walmart sales dashboard*. Rochester Institute of Technology.
- Kahn, K. B. (2002). An exploratory investigation of new product forecasting practices. *Journal of Product Innovation Management*, 19(2), 133-143. <https://doi.org/10.1111/1540-5885.1920133>
- Lin, A. (2024). Walmart sales prediction using multiple linear regression. En *Highlights in Science, Engineering and Technology: AMMSAC 2024* (Vol. 107, pp. 124-129). Science Publications.
- Seyedan, M., & Mafakheri, F. (2020, julio). Predictive big data analytics for supply chain demand forecasting: Methods, applications, and research opportunities. *Journal of Big Data*, 7, 1-24. <https://doi.org/10.1186/s40537-020-00329-2>
- Singh, M., Ghutla, B., Jnr, R. L., et al. (2017). Walmart's sales data analysis — A big-data analytics perspective. En *Proceedings of the 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*. IEEE.
- Smirnov, P. S., & Sudakov, V. A. (2020). Forecasting new product demand using machine learning. *Journal of Physics: Conference Series*, 1925(1), 012042. <https://doi.org/10.1088/1742-6596/1925/1/012042>
- Tseng, C. S., & Turkmen, T. (2024). *Demand forecasting with machine learning* (Proyecto de fin de máster, Massachusetts Institute of Technology). Massachusetts Institute of Technology.
- University of Michigan. (2025). *University of Michigan: Consumer Sentiment*. FRED, Federal Reserve Bank of St. Louis. Obtenida el 05/04/2025 de <https://fred.stlouisfed.org/series/UMCSENT>
- Velos, S. P., Marivel, G. B., Bate, G. P., et al. (2020). A seasonal autoregressive integrated moving average (SARIMA) model to forecasting tourist arrival in the Philippines:

A case study in Moalboal, Cebu (Philippines). *Recoletos Multidisciplinary Research Journal*, 8(1), 20-30. <https://doi.org/10.32871/rmrj2008.01.03>

Wu, L., Yan, J. Y., & Fan, Y. J. (2012). Data mining algorithms and statistical analysis for sales data forecast. En *Proceedings of the Fifth International Joint Conference on Computational Sciences & Optimization* (pp. 577-581). IEEE.

Xu, Z. G. (2017). Complex production process prediction model based on EMD-XGBoost-RLSE. En *Proceedings of the 9th International Conference on Modelling Identification and Control*.

Yao, B. (2023). Walmart sales prediction based on decision tree, random forest, and K-neighbors regressor. *Highlights in Business, Economics and Management*, 5, 330-335.

Yasser, H. M. (2022). *Walmart store sales prediction – regression problem*. Kaggle. Obtenida el 05/04/2025 de <https://www.kaggle.com/datasets/yasserh/walmart-dataset>

Ye, X., & Le, W. (2023). A new energy vehicle sales forecasting model based on high-dimensional tensor CNN. En *Third International Conference on Artificial Intelligence and Computer Engineering* (Proc. SPIE 12610). <https://doi.org/10.1117/12.2671046>

APÉNDICE 1

	Modelo I				Modelo II				Modelo III				Modelo IV			
	Coef.	Est. t	Valor p	Sign.	Coef.	Est. t	Valor p	Sign.	Coef.	Est. t	Valor p	Sign.	Coef.	Est. t	Valor p	Sign.
const	1626809.68	77.17	0	0	1277907.25	5.61	2.09E-08	***	53469.66	0.13	0.8927		111416.98	0.28	0.7789	
holiday_flag	71487.09	8.93	5.55E-19	***	69239.28	8.67	5.23E-18	***	62947.29	8.13	5.04E-16	***	0.08	0.00	1	
temperature	-825.84	-6.24	4.63E-10	***	-833.71	-6.33	2.65E-10	***	-1457.70	-10.66	2.67E-26	***	141.35	0.23	0.8193	
fuel_price	-6253.35	-1.33	0.1848		-41354.15	-5.77	8.09E-09	***	38104.76	4.29	0.0000183	***	34218.81	3.81	0.0001	***
store_2	370412.27	19.47	4.49E-82	***	371741.59	19.63	2.57E-83	***	371530.33	20.29	9.73E-89	***	371948.14	20.32	5.29E-89	***
store_3	-1149977.21	-60.43	0	***	-1169648.22	-60.89	0	***	-1166869.01	-61.09	0	***	-1165961.38	-61.07	0	***
store_4	534432.35	28.07	1.47E-163	***	753446.46	7.90	3.34E-15	***	700277.18	5.43	5.78E-08	***	679607.09	5.27	0.000000144	***
store_5	-1236341.39	-64.99	0	***	-1267545.31	-64.35	0	***	-1269467.91	-65.69	0	***	-1269971.78	-65.74	0	***
store_6	10614.36	0.56	0.5769		-16485.90	-0.85	0.3937		-17018.30	-0.89	0.3718		-17010.79	-0.89	0.3717	
store_7	-1008118.26	-51.96	0	***	-920154.07	-31.65	2.23E-204	***	-949326.69	-24.98	1.67E-131	***	-949682.93	-25.00	1.08E-131	***
store_8	-651298.80	-34.21	1.2E-235	***	-695559.20	-35.03	5.12E-246	***	-700722.80	-34.36	1.4E-237	***	-701651.63	-34.42	2.77E-238	***
store_9	-1011722.87	-53.18	0	***	-1056311.81	-53.24	0	***	-1058075.90	-51.83	0	***	-1057998.52	-51.86	0	***
store_10	349636.96	18.31	5.05E-73	***	635211.17	6.87	6.99E-12	***	565060.34	4.36	0.0000132	***	548935.02	4.23	0.0000232	***
store_11	-195434.23	-10.27	1.51E-24	***	-215097.00	-11.20	7.89E-29	***	-211664.73	-11.08	2.79E-28	***	-210346.20	-11.01	5.8E-28	***
store_12	-542229.13	-28.38	6.25E-167	***	-148802.71	-1.71	0.0866		-212090.51	-1.59	0.1113		-224280.84	-1.68	0.0925	
store_13	436706.20	22.83	6.66E-111	***	681382.32	7.28	3.87E-13	***	619672.30	4.80	0.0000159	***	600573.65	4.65	0.00000336	***
store_14	458267.23	23.98	9.49E-122	***	575612.78	16.20	7.08E-55	***	539982.63	11.31	2.23E-29	***	533960.35	11.18	9.62E-29	***
store_15	-943313.51	-48.98	0	***	-684867.07	-7.90	3.2E-15	***	-765375.94	-6.37	1.98E-10	***	-781856.12	-6.51	3.2E-10	***
store_16	-1055102.80	-54.74	0	***	-1014451.47	-31.64	3.6E-204	***	-1044975.79	-27.75	4.52E-160	***	-1049045.18	-27.85	3.63E-161	***
store_17	-679369.55	-35.28	2.71E-249	***	-444915.48	-4.70	0.0000026	***	-512187.53	-3.97	0.0000718	***	-530378.65	-4.11	0.0000399	***
store_18	-481491.82	-25.11	8.96E-133	***	-208908.06	-2.47	0.0136	**	-275743.61	-2.29	0.022	**	-292237.93	-2.43	0.0153	**
store_19	-121246.08	-6.30	3.21E-10	***	137203.99	1.58	0.1134		56982.90	0.47	0.6352		40287.04	0.34	0.7375	
store_20	543033.58	28.37	6.65E-167	***	564921.87	27.06	8.7E-153	***	536982.31	25.40	1.05E-135	***	535119.73	25.32	8.61E-135	***
store_21	-798748.67	-41.99	0	***	-797414.38	-42.10	0	***	-797231.74	-43.53	0	***	-796541.43	-43.51	0	***
store_22	-536449.07	-28.01	6.54E-163	***	-292623.71	-3.57	0.0004	***	-358146.07	-3.13	0.0017	***	-374959.33	-3.28	0.0011	***
store_23	-180116.69	-9.36	1.13E-20	***	1613.03	0.02	0.9861		-77018.38	-0.64	0.5208		-95272.59	-0.79	0.4257	
store_24	-208056.92	-10.82	4.59E-27	***	60534.08	0.70	0.4811		-17606.75	-0.15	0.8836		-34271.79	-0.28	0.7758	
store_25	-860657.77	-44.86	0	***	-838795.56	-40.09	0	***	-868802.40	-41.02	0	***	-870274.77	-41.10	0	***
store_26	-571319.78	-29.51	1.73E-179	***	-320520.79	-3.71	0.0002	***	-395555.69	-3.30	0.001	***	-410339.99	-3.42	0.0006	***
store_27	213113.57	11.11	1.99E-28	***	460231.76	5.57	2.72E-08	***	385215.73	3.37	0.0008	***	368746.64	3.22	0.0013	***
store_28	-227708.50	-11.92	2.13E-32	***	165717.92	1.91	0.0553		102430.12	0.77	0.4418		90239.80	0.68	0.4983	
store_29	-1025498.68	-53.54	0	***	-731152.96	-8.78	0	***	-794893.77	-6.58	5.19E-11	***	-811274.31	-6.71	2.14E-11	***
store_30	-1116238.13	-58.67	0	***	-1114903.84	-58.86	0	***	-1114721.21	-60.87	0	***	-1114030.89	-60.86	0	***
store_31	-158916.31	-8.35	6.86E-17	***	-157582.02	-8.32	7.87E-08	***	-157399.39	-8.60	1.03E-17	***	-156709.07	-8.56	1.39E-17	***
store_32	-401408.98	-20.98	1.61E-94	***	-313342.24	-10.85	3.55E-27	***	-334385.96	-8.82	1.52E-18	***	-338648.06	-8.93	5.74E-19	***
store_33	-1286220.39	-67.30	0	***	-996787.21	-10.81	5.02E-27	***	-1063761.53	-8.20	2.8E-16	***	-1077073.93	-8.30	1.21E-16	***
store_34	-596602.13	-31.29	4.58E-200	***	-288421.05	-3.26	0.0011	***	-335148.09	-2.58	0.01	**	-353254.60	-2.71	0.0067	***
store_35	-643382.48	-33.66	9.11E-229	***	-384226.86	-4.76	0.00000196	***	-444733.57	-3.88	0.0001	***	-461009.73	-4.02	0.0000542	***
store_36	-1179492.72	-61.99	0	***	-1170484.60	-61.66	0	***	-1167567.67	-63.27	0	***	-1167493.09	-63.30	0	***
store_37	-1034007.53	-54.34	0	***	-1024455.47	-53.96	0	***	-1022769.87	-55.43	0	***	-1022656.90	-55.45	0	***
store_38	-1165499.09	-60.99	0	***	-772072.66	-8.89	7.62E-19	***	-835360.46	-6.27	3.77E-10	***	-847550.79	-6.36	2.09E-10	***
store_39	-102704.63	-5.40	6.98E-08	***	-93157.01	-4.91	0.000000949	***	-91822.72	-4.98	0.00000663	***	-91819.55	-4.98	0.00000656	***
store_40	-606786.77	-31.48	2.46E-202	***	-425065.95	-4.60	0.00000434	***	-504402.82	-4.21	0.0000264	***	-522335.02	-4.35	0.0000138	***
store_41	-303433.55	-15.80	3.58E-55	***	-251625.79	-8.09	6.81E-16	***	-278944.44	-7.44	1.12E-13	***	-283851.40	-7.57	4.29E-14	***
store_42	-993383.75	-52.02	0	***	-707809.53	-7.66	2.2E-14	***	-777960.37	-6.00	2.13E-09	***	-794085.69	-6.13	9.59E-10	***
store_43	-921468.19	-48.44	0	***	-845007.17	-39.52	8.02E-306	***	-843855.69	-34.10	2.49E-234	***	-843561.95	-34.11	2.34E-234	***
store_44	-1264165.24	-66.08	0	***	-1025469.90	-10.89	2.25E-27	***	-1087768.47	-8.44	4.04E-17	***	-1107134.33	-8.58	1.2E-17	***
store_45	-776729.76	-40.64	0	***	-659384.21	-18.56	6.39E-75	***	-695014.36	-14.56	3E-47	***	-701036.65	-14.68	5.53E-48	***
cpi					2933.35	2.84	0.0045	***	2406.92	1.64	0.1002		2190.26	1.49	0.135	
unemployment					-22467.01	-5.24	0.000000169	***	-24680.27	-5.32	0.00000011	***	-25418.15	-5.47	4.78E-08	***
monthly_csi									873.52	2.24	0.025	**	771.91	1.97	0.0489	**
quarterly_dsr_pct_di									98722.14	11.28	3.07E-29	***	96457.96	10.98	8.73E-28	***
spread_10a3m									-84567.36	-20.91	5.9E-94	***	-83557.18	-20.59	2.95E-91	***
temperature_sq													-13.94	-2.66	0.0078	***
holiday_fuel													19879.94	1.07	0.2833	
R-cuadrado		0.9194				0.9201				0.9254				0.9255		
R-cuadrado corregido		0.9188				0.9195				0.9248				0.9249		
V-p de F		0.0000				0.0000				0.0000				0.0000		
Criterio de Schwarz		172924.01				172881.73				172466.8873				172476.89		
Log-verosimilitud		-86251.54				-86221.63				-86001.0517				-85996.89		