



Máster Universitario en Técnicas Estadísticas | Curso 2025/2026 | Resumen extendido

Análisis de procesos en retail y diagnóstico de necesidades de previsión coordinada

Iván Lado Canosa

Director: Manuel Febrero Bande (USC) Tutores: Natalia Vallejo Paredes y David Polo Tascón (Inditex)

Este documento constituye un resumen extendido del Trabajo Fin de Máster titulado “*Análisis de procesos en retail y diagnóstico de necesidades de previsión coordinada*”, realizado por Iván Lado Canosa bajo la dirección de Manuel Febrero Bande (catedrático de universidad de la Universidade de Santiago de Compostela) y la tutorización de Natalia Vallejo Paredes y David Polo Tascón, del equipo de *Business Analytics* de Industria de Diseño Textil, S.A. (en adelante, Inditex). Esta versión ha sido adaptada específicamente para su difusión en el repositorio de acceso público del Máster en Técnicas Estadísticas, omitiendo información de carácter confidencial en cumplimiento de los compromisos suscritos con Inditex.

Introducción y motivación

Las tiendas físicas de Zara, la cadena más representativa del grupo Inditex, ejecutan a diario numerosas tareas operativas cuya correcta planificación de recursos humanos resulta clave para su funcionamiento. Estas tareas se agrupan en procesos (recepción de camión, venta en caja, devoluciones, gestión comercial, reposición, etc.), y la compañía está trabajando en cuantificar la carga operativa que cada uno genera, midiendo su productividad en unidades manipuladas por hora. Conocida esa productividad, disponer de una previsión de las unidades que manipulará cada proceso permite anticipar las horas de trabajo necesarias y planificar la plantilla de forma eficiente.

La **reposición** —la preparación en el almacén de la mercancía que repondrá las unidades vendidas en la superficie de venta— es uno de los procesos que más carga operativa supone a las tiendas. Sin embargo, a diferencia de otros procesos, actualmente no se dispone de una predicción del número de unidades que maneja, lo que dificulta la planificación de los recursos asociados. El presente trabajo aborda directamente esta carencia.

Objetivos

El trabajo se plantea un objetivo doble:

- (1) Proponer un **modelo explicativo** del número de unidades manipuladas por el proceso de reposición, que identifique los factores operativos que más influyen en dicha carga.
- (2) A partir del modelo explicativo, construir un **modelo predictivo** capaz de anticipar la carga futura de reposición empleando únicamente otras predicciones ya disponibles en la compañía, lo que garantiza coherencia entre las distintas estimaciones internas.

Datos y metodología

Los datos proceden de los registros RFID de la compañía, que permiten rastrear cada unidad de producto a lo largo de su ciclo de vida en la tienda. Tras una fase de limpieza, depuración de valores atípicos y construcción de variables con sentido operativo, el análisis se desarrolla sobre el dato semanal y se centra en los periodos de actividad regular, en los que la previsión resulta operativamente relevante. Para evitar que las diferencias de tamaño entre tiendas distorsionen las conclusiones, las correlaciones y los modelos se calculan a nivel de tienda individual y se promedian después. El análisis exploratorio confirmó la intuición de negocio: la reposición está estrechamente correlacionada con la venta.

Modelos y resultados

El desarrollo del modelo explicativo se abordó de forma progresiva, motivando cada paso a partir de las limitaciones del anterior. Como referencia se emplearon varios modelos naif, siendo el mejor aquel que toma como predicción el valor observado la semana anterior.

- Un **único modelo de regresión lineal múltiple** para todas las tiendas ofreció un ajuste razonable, pero su diagnosis reveló un sesgo sistemático: no lograba ajustar el nivel absoluto de reposición de cada tienda.
- Ajustar **un modelo por clasificación de tienda** apenas mejoró los resultados, lo que evidenció que la heterogeneidad relevante no se captura agrupando tiendas por volumen de ventas.
- Los **modelos de efectos mixtos**, que reconocen la estructura jerárquica de los datos (mediciones agrupadas por tienda), resolvieron en gran medida el problema. El modelo con intercepto aleatorio y pendiente fija mejoró el error de nivel manteniendo la capacidad explicativa, confirmando que las diferencias entre tiendas son un factor determinante. El modelo con **intercepto y pendiente aleatorias** mejoró todas las métricas y resultó el más adecuado.

Se concluye que la reposición se explica adecuadamente mediante un modelo lineal jerárquico de efectos mixtos, en el que la venta resulta el factor de mayor peso.

Al evaluar los modelos se detectó además un fenómeno de *concept drift*: un cambio en la dinámica operativa de la reposición a lo largo del periodo analizado que introducía un sesgo sistemático. La solución adoptada fue un **reentrenamiento deslizante**, en el que para cada semana el modelo se ajusta únicamente con las semanas inmediatamente anteriores, lo que mejoró de forma apreciable el error.

Modelo predictivo

Para el modelo predictivo se reajustaron los modelos explicativos sustituyendo las variables reales por las predicciones ya disponibles en la compañía, corrigiendo así su posible sesgo. Como era esperable, el rendimiento empeoró respecto al modelo explicativo al incorporarse el error propio de las variables de entrada. Aun así, el mejor modelo predictivo mejora al modelo naif de referencia y consigue anticipar la carga de reposición con un error relativo contenido, para un proceso que hasta ahora carecía de cualquier previsión.

Conclusiones y trabajo futuro

El trabajo cubre la necesidad que lo motivó: proporciona una estimación de la carga futura de la reposición que, convertida en horas mediante las productividades de la compañía, sirve de apoyo a la planificación de los recursos humanos en las tiendas. Metodológicamente, muestra cómo un enfoque progresivo —desde la regresión lineal hasta los modelos de efectos mixtos con reentrenamiento deslizante— permite modelar de forma consistente un proceso con estructura jerárquica y dinámica cambiante en el tiempo.

Como principales líneas de trabajo futuro se identifican dos: ofrecer predicciones a nivel diario —ajustando un modelo distinto para cada día de la semana, dado que la carga de reposición varía según el día para unos mismos valores de las variables explicativas— y extender las predicciones a horizontes temporales mayores, por necesidad de negocio.