

# Análisis de Clustering y Métodos de Pronóstico para Series Temporales Geográficas en el Dataset Global Superstore 2018

Saul Martínez<sup>a</sup>

<sup>a</sup>*Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Nuevo León, Monterrey, 66451, México*

---

## Abstract

Este estudio presenta un análisis de clustering y métodos de pronóstico aplicados al dataset de Global Superstore 2018. Se realizó un clustering con K-means para identificar grupos de productos o clientes con características similares, determinando el número óptimo de clusters mediante el método del codo y el índice de Davies-Bouldin. Adicionalmente, se aplicó el algoritmo DBSCAN y se evaluó su rendimiento utilizando varias métricas. Los resultados de los diferentes métodos de clustering fueron comparados y discutidos.

En cuanto a los métodos de pronóstico para series temporales geográficas, se resumió el estudio de Huddleston, Porter y Brown sobre la previsión de eventos delictivos en Pittsburgh. Se discutió el uso de modelos ARIMA y métodos de pronóstico de arriba hacia abajo, destacando la importancia de seleccionar el método de pronóstico adecuado según las necesidades específicas del análisis y las capacidades del analista.

*Keywords:* Clustering, K-means, DBSCAN, Series Temporales, ARIMA, Métodos de Pronóstico, Global Superstore 2018

---

## 1. Introducción

Este documento presenta un análisis de clustering y métodos de pronóstico aplicados al dataset de Global Superstore 2018. Se emplearon técnicas avanzadas para identificar patrones en los datos de ventas y para predecir eventos futuros en contextos geográficos ruidosos.

## **2. Análisis de Clustering**

Se realizó un análisis de clustering utilizando el algoritmo K-means, determinando el número óptimo de clusters mediante el método del codo y el índice de Davies-Bouldin. También se aplicó el algoritmo DBSCAN y se evaluó su rendimiento utilizando diversas métricas. Los resultados obtenidos con ambos métodos fueron comparados y discutidos para identificar la mejor estrategia de segmentación.

## **3. Métodos de Pronóstico para Series Temporales Geográficas**

Se resumió el estudio de Huddleston, Porter y Brown sobre la previsión de eventos delictivos en Pittsburgh, destacando el uso de modelos ARIMA y métodos de pronóstico de arriba hacia abajo. Se discutió la importancia de seleccionar el método de pronóstico adecuado según las necesidades específicas del análisis y las capacidades del analista.

## **4. Conclusión**

Los resultados de este estudio proporcionan insights valiosos sobre la segmentación de clientes y productos, así como sobre la previsión de eventos en series temporales geográficas. La combinación de técnicas de clustering y métodos de pronóstico robustos puede mejorar significativamente la toma de decisiones en contextos de ventas y otros campos similares.

## **Agradecimientos**

Se agradece a la Universidad Universidad Autonoma de Nuevo León por el apoyo brindado para la realización de este estudio.

## **Referencias**

### **References**

@articleHuddleston2024, title=Improving forecasts for noisy geographic time series, author=Huddleston, Samuel H and Porter, John H and Brown, Donald E, journal=Journal of Business Research, year=2024, volume=38, number=4, pages=567–584, publisher=Elsevier

@bookMakridakis1983, title=Forecasting Methods and Applications, author=Makridakis, Spyros and Wheelwright, Steven C and McGee, Victor E, year=1983, publisher=Wiley

@articleGorr2003, title=Crime forecasting using ARIMA and ARIMAX models, author=Gorr, Wilpen L and Olligschlaeger, Andreas M and Thompson, Yvonne, journal=International Journal of Forecasting, year=2003, volume=19, number=4, pages=571–594, publisher=Elsevier