

CODER HOUSE

Sales Forecasting



Comisión 14075 - Proyecto Final 2021

Equipo



Gastón Saver

Business Intelligence Manager
IntegraGO | Smollan
Consultora de Trade Marketing



Germán Leloutre

Administrativo
Lions Sports & Media SA



Lucía Monchiero García

Data Science Analyst
Accenture

WHAT'S IN STORE

Índice

- Contexto
- Objetivos generales
- Análisis de Ventas en el tiempo
- Análisis de Series de tiempo
- Predicciones
- Clusterización de Series de tiempo
- Conclusiones finales



Contexto

Mercado Libre es la empresa líder y pionera en comercio electrónico y fintech en América Latina, basada en una plataforma donde los usuarios compran, venden, publicitan, entregan, financian y pagan bienes y servicios a través de Internet.

Está construyendo un ecosistema emprendedor que está democratizando el comercio, el dinero y los pagos, empoderando a millones de personas en América Latina.

Objetivos Generales



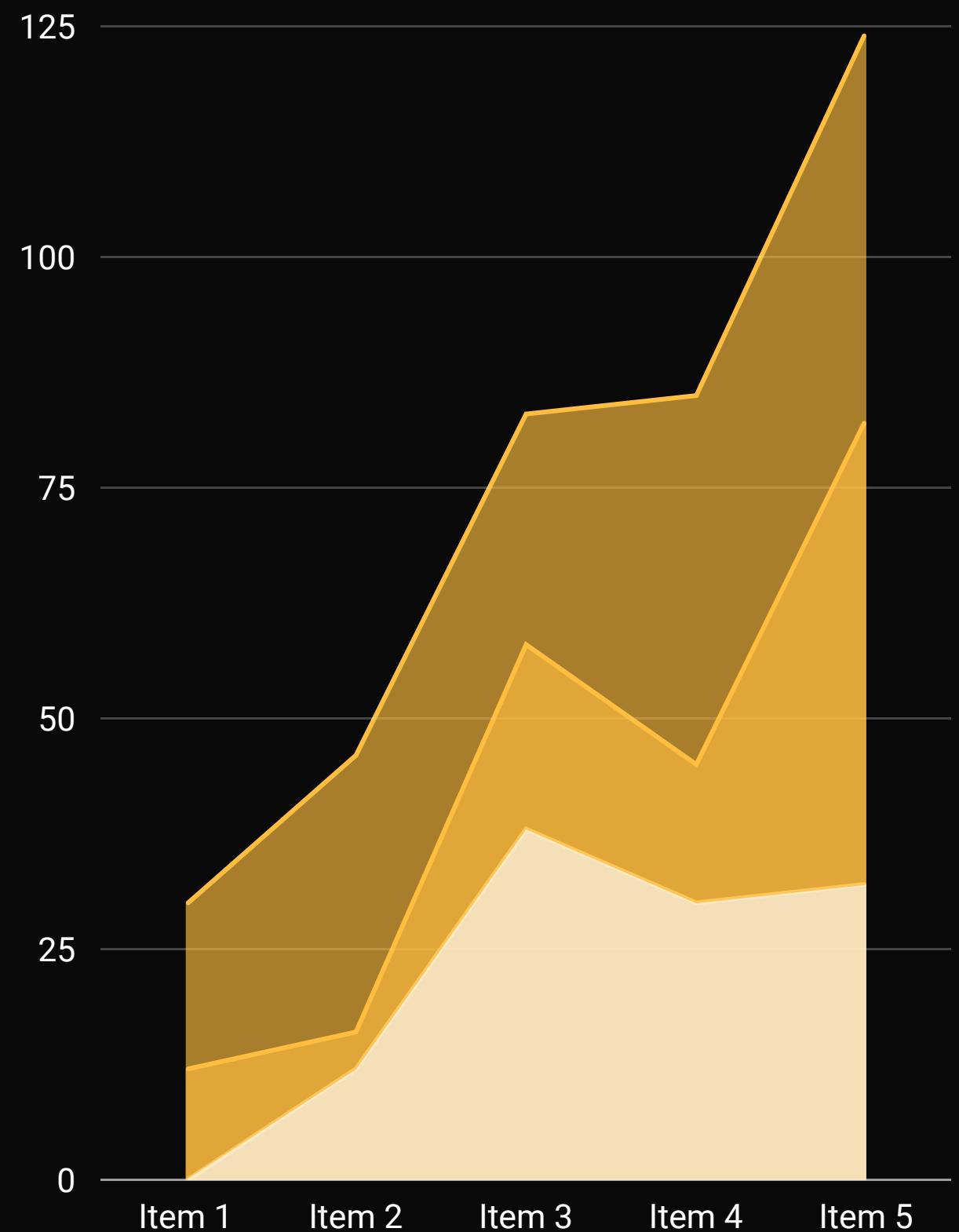
FORECASTING

Todos los días, la unidad de envío de Mercadolibre (MercadoEnvíos) entrega miles de productos en toda Latinoamérica.

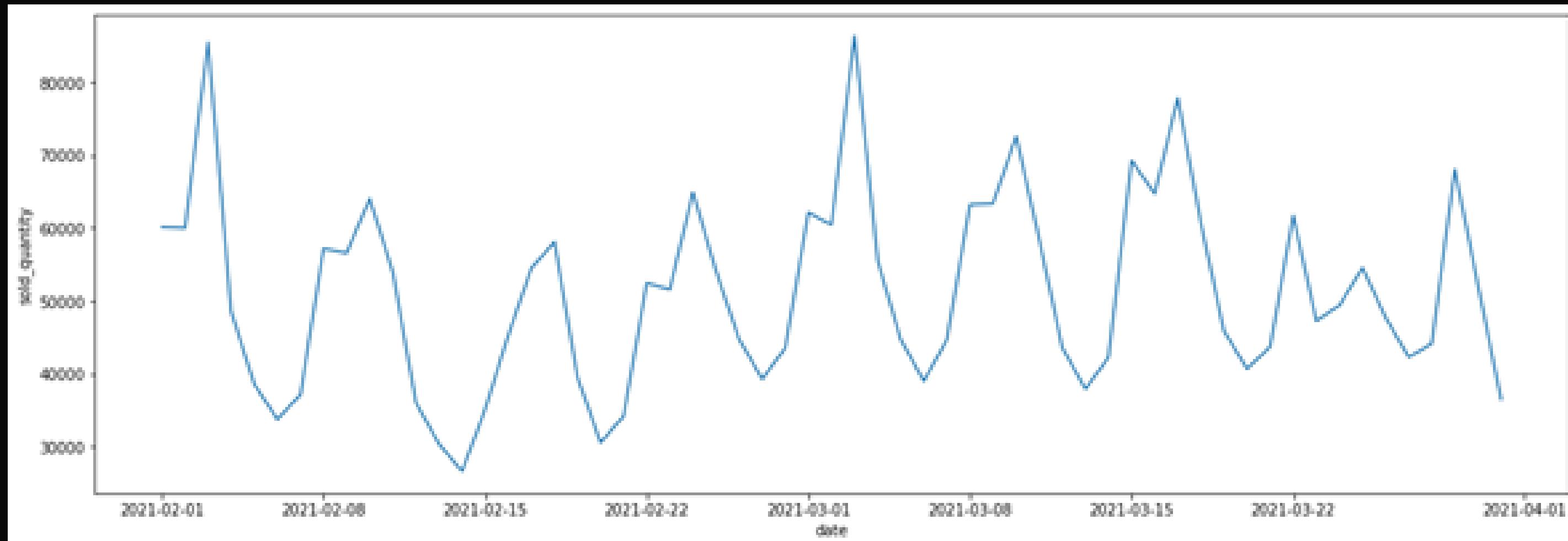
Para brindar a los clientes la mejor experiencia durante todo el proceso de envío, es fundamental contar con pronósticos de ventas precisos.

En el centro de este problema comercial, se encuentra la tarea de **Desafío de datos**: Construir un modelo para pronosticar las ventas de artículos de los 30 días siguientes en base a los datos históricos de Mercado Libre.

Análisis de las Ventas en el tiempo

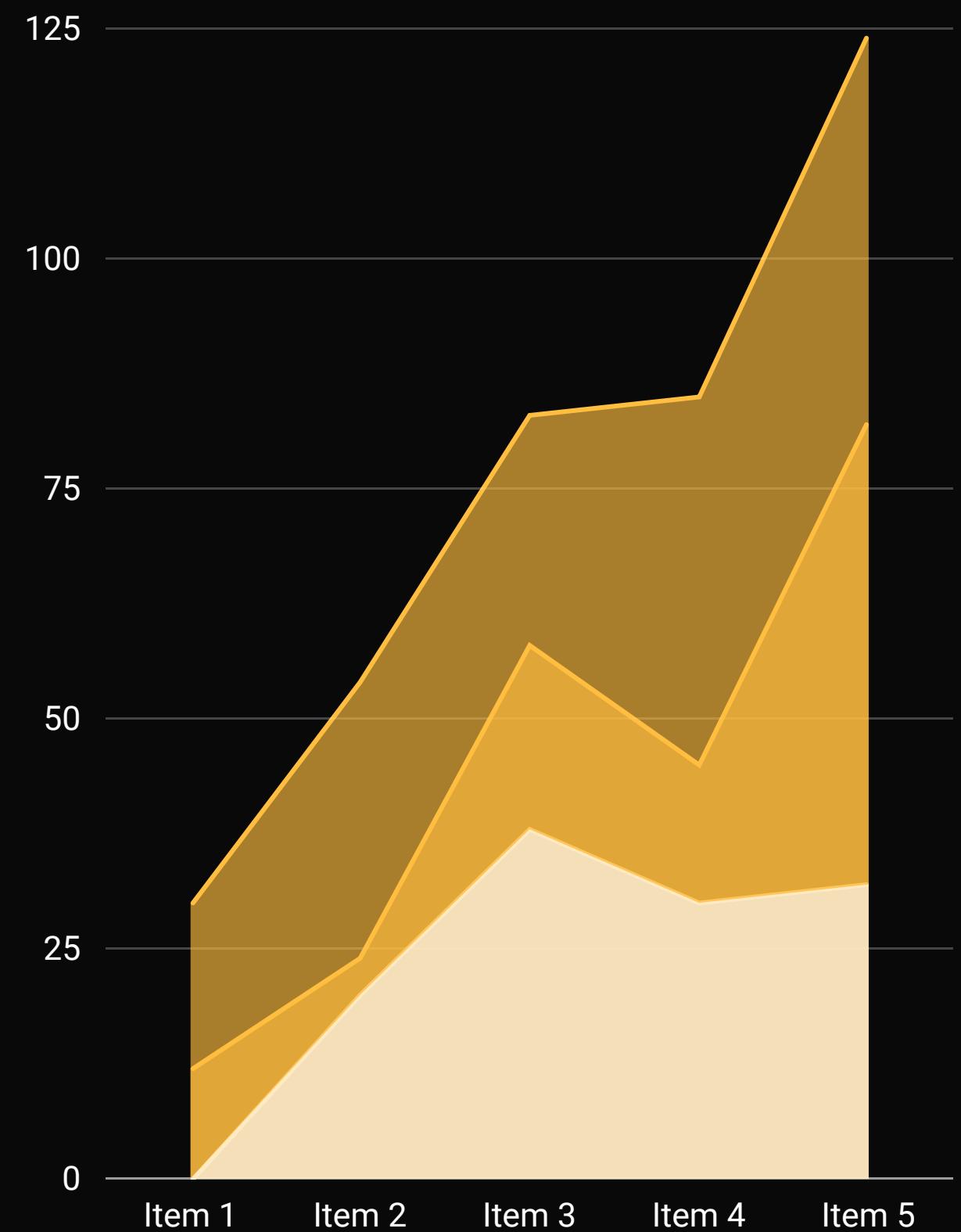


En base a los dos meses de información histórica brindados por Mercadolibre analizamos el comportamiento en un gráfico



Del mismo, podemos observar que existe una **estacionalidad** asociada a los datos. Cada semana tiene un día de ventas pico y un dia de ventas mínimas

Análisis de Series de tiempo



El análisis de series de tiempo es una de las habilidades centrales de cualquier científico de datos y cualquier persona que trabaje en el campo de la analítica a menudo se encuentra con la situación para pronosticar el futuro con los datos pasados y presentes.

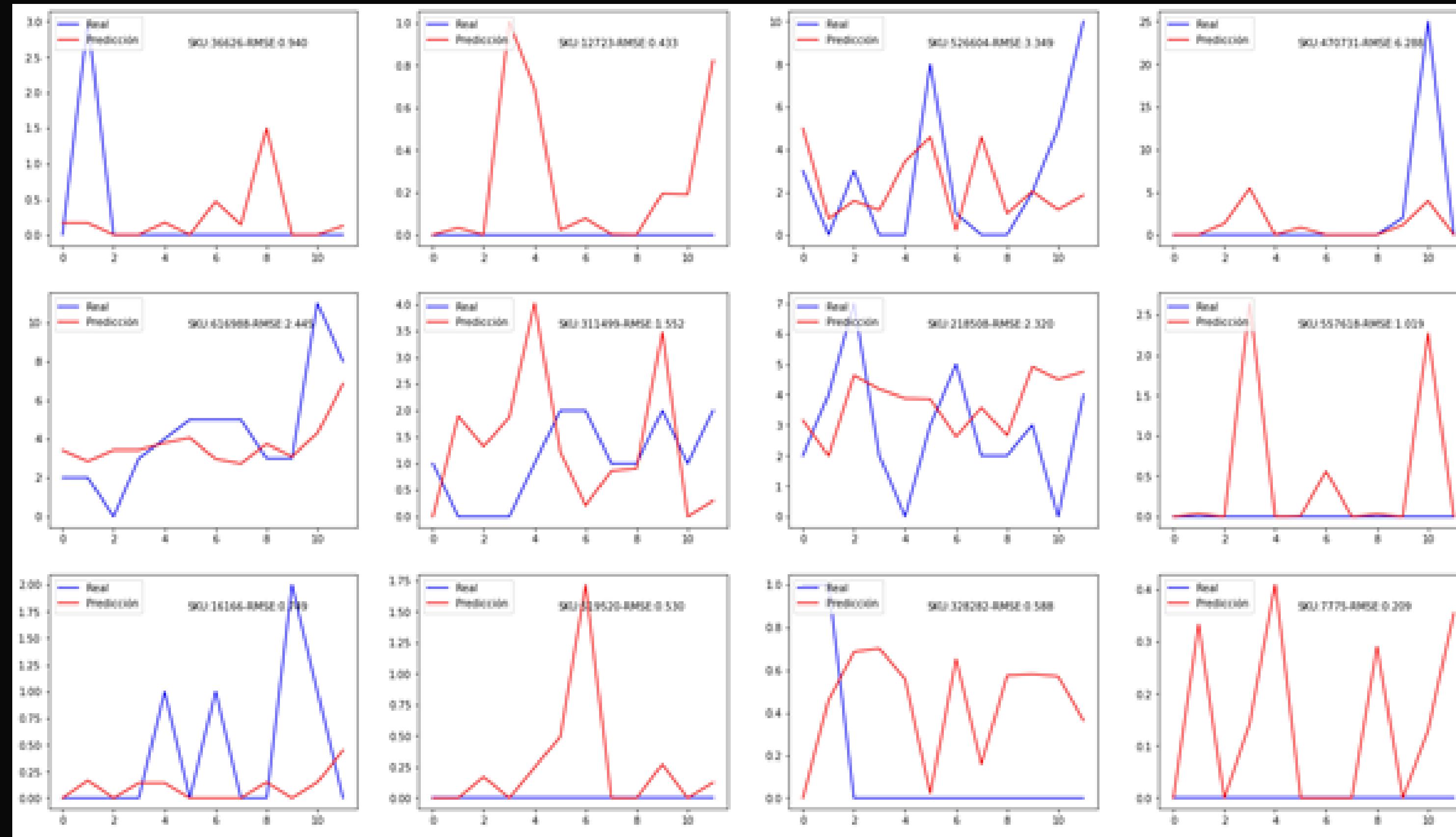
¿Qué es una serie temporal?

Es una serie de observaciones tomadas en momentos específicos básicamente a intervalos iguales. Se utiliza para predecir valores futuros basados en valores pasados observados. Los componentes que puede observar en el análisis de series de tiempo son Tendencia, Estacional, Irregular y Ciclicidad.

En el caso de estos conjuntos de datos, donde solo se observa una variable en cada momento, se denomina "Serie de tiempo univariante" y si se observan dos o más variables en cada momento, se denomina "Serie de tiempo multivariante"

Algunas Predicciones

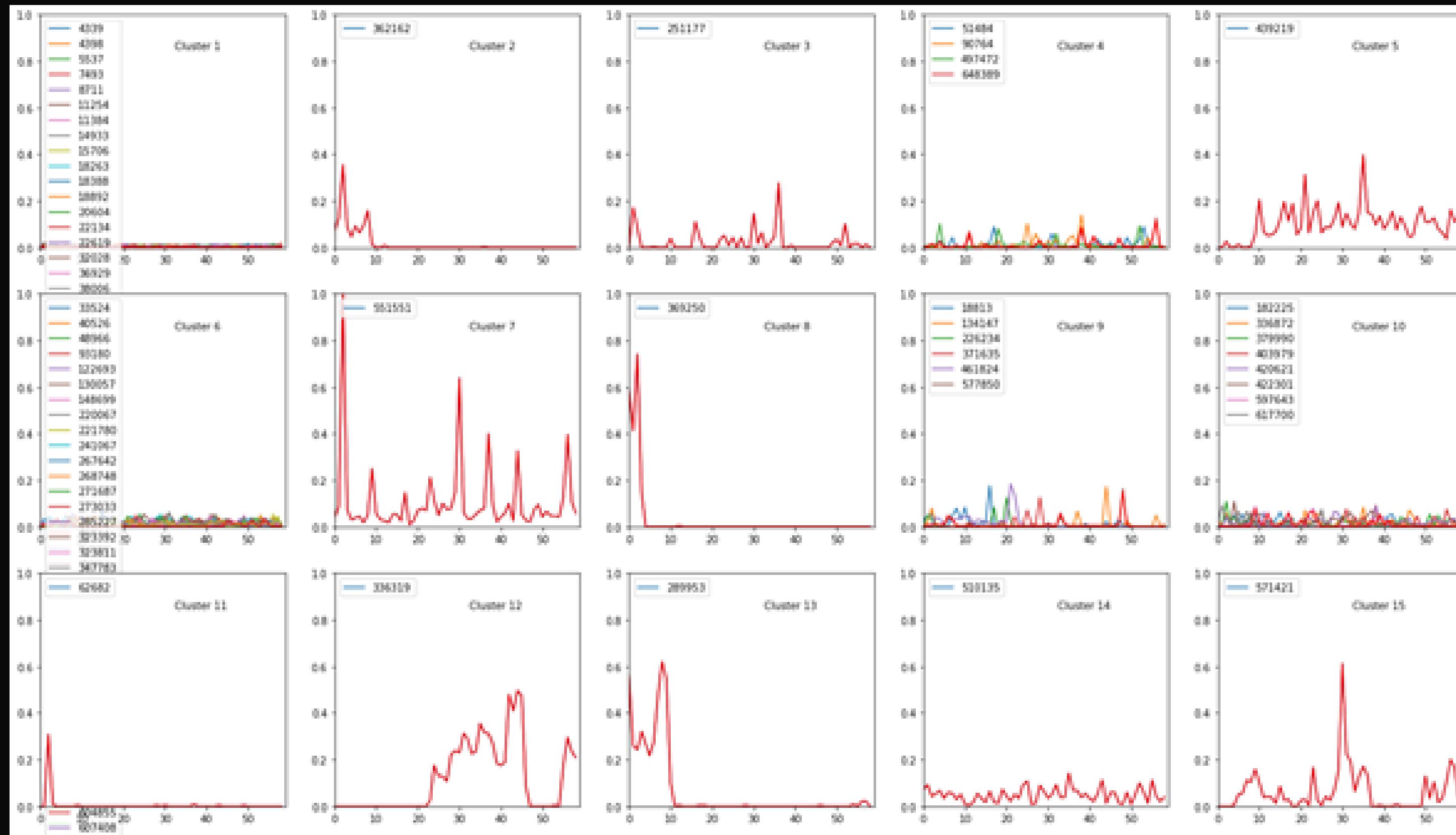




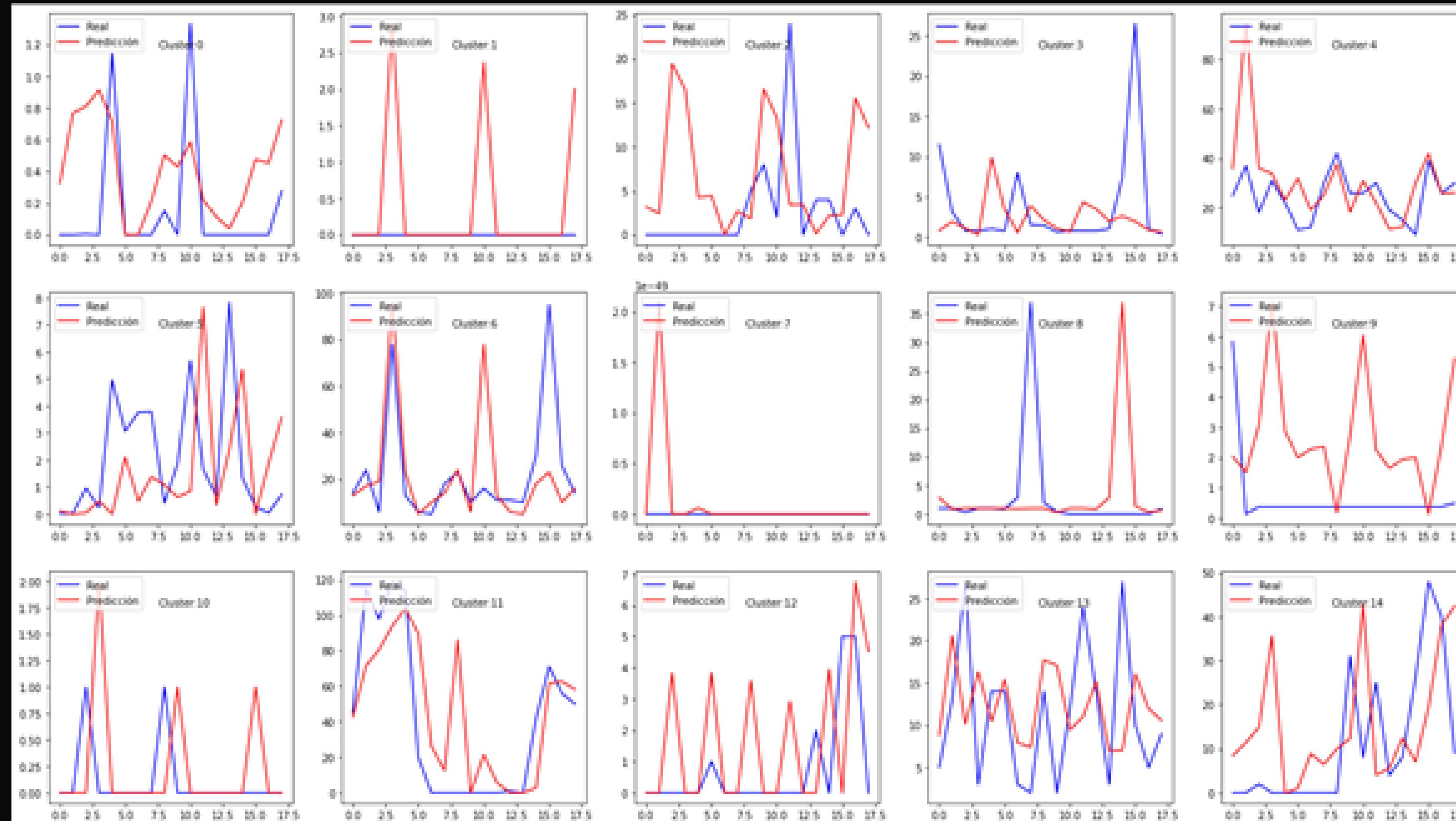
Clusterización de Series de Tiempo

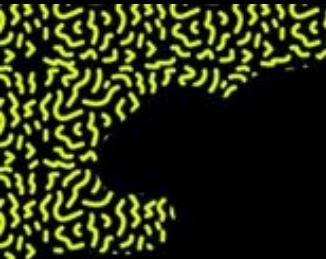
La agrupación o clustering es uno de los **métodos de extracción de datos más populares**, no sólo por su poder de exploración, sino también como paso previo al procesamiento de otras técnicas.

Dynamic time warping - Time Series Clustering



Predicciones en base a los Centroides





#NERDS4EVER

Qué logramos?

- Realizamos el análisis del dataset como una **serie de tiempo**
- Ejecutamos **predicciones** para algunos cientos de SKUs
- **Clusterizamos** el set de grupos de SKUs
- Ejecutamos **predicciones** para los distintos **centroides**
- Pusimos en **práctica** varios conceptos aprendidos durante el curso, sin embargo concluimos que la **cantidad de información suministrada no es suficiente** para realizar una predicción acertada