ARIMA

**Descripción**: Es un modelo estadístico que utiliza datos de series de tiempo para comprender mejor el conjunto de datos o para predecir tendencias futuras. Es una forma de análisis de regresión que mide la fuerza de una variable dependiente en relación a otra variable cambiante.

**Siglas**: Autoregresivo(AR), Integrado(I), Media móvil(MA)

* **AR**: **AutoRegressive** (Autoregresivo)  
  Indica que el modelo utiliza las observaciones pasadas de la serie temporal para predecir valores futuros. Específicamente, el componente autoregresivo (AR) se refiere al uso de una relación lineal entre el valor actual de la serie y sus valores anteriores.
* **I**: **Integrated** (Integrado)  
  Se refiere al grado de diferenciación necesaria para hacer que la serie de datos temporales sea estacionaria, es decir, una serie cuya media y varianza sean constantes a lo largo del tiempo. También se puede decir que ndica cuántas veces los datos han sido diferenciados para eliminar la tendencia.
* **MA**: **Moving Average** (Media Móvil)  
  Este componente se basa en la relación entre el valor actual de la serie y los errores de predicción pasados. El término "media móvil" (MA) indica que el modelo utiliza el promedio de los errores pasados para hacer predicciones.

**Serie de tiempo**: Es una secuencia en la que se registra una métrica en intervalos de tiempos regulares. En nuestro caso, el gasto total de la compra del mes es la métrica mensual.

Para aplicar el modelo ARIMA, es fundamental verificar si la serie temporal es estacionaria. Una serie temporal es estacionaria si sus propiedades estadísticas, como la media y la varianza, son constantes a lo largo del tiempo.

**Prueba de Dickey-Fuller Aumentada** (ADF): Es una prueba estadística utilizada para determinar si una serie temporal es estacionaria. Evalúa la hipótesis nula de que existe una raíz unitaria en la serie (es decir, la serie no es estacionaria). Si el valor p es menor que un umbral crítico (normalmente 0.05), se rechaza la hipótesis nula y se concluye que la serie es estacionaria.

**Diferenciacion**: La diferenciación es una técnica utilizada para hacer que una serie temporal no estacionaria se vuelva estacionaria. Esto se logra restando el valor anterior de la serie del valor actual:

* Primera Diferenciación: Se calcula restando cada valor de la serie por su valor anterior. Si la serie aún no es estacionaria, se pueden aplicar diferenciaciones adicionales.
* Segunda Diferenciación: Se calcula tomando la diferencia de las diferencias (aplicando la primera diferenciación nuevamente a la serie ya diferenciada).

El parámetro 𝑑 en el modelo ARIMA indica el número de veces que la serie ha sido diferenciada.

**Fórmula**: ARIMA (p, d, q)

- p: es el orden del término AR: modelo de regresión lineal que usa sus propios rezagos como predictores.

- d: es el número de diferenciaciones necesarias para que la serie de tiempo sea estacionaria, si la serie de tiempo ya es estacionaria, entonces d = 0.

- q: es el orden del término MA: Se refiere al número de errores de pronóstico retrasados que deben incluirse en el modelo ARIMA.

**Características**:

- Modelo aplicable a series de tiempo no estacionarias.

- Puede predecir una serie de tiempo utilizando los valores pasados de la serie.

- Para construir un modelo ARIMA es necesario hacer que la serie temporal sea estacionaria.