

## **Proyecto #3: Exploración y Creación de Características**

Para completar el análisis y creación de pipeline utilice toda la información que recabó y desarrolló en la fase anteriores del proyecto, para implementar las siguientes. Esta fase servirá de base para definir el pipeline de pro-procesamiento del sistema de machine learning. Antes de iniciar a desarrollar esta fase, le recomendamos fuertemente investigar como funcionan los modelos de forecasting que utilizan variables externas o de segregación para manejar valores de tienda y producto como variables exógenas y parte de la estructura del modelo.

1. Cree el notebook para análisis y selección del proceso de ingeniería de características llamado 02\_feature\_exploration.ipynb, en este notebook deberá realizar los siguientes pasos:
  - a. Cargar el dataset definido en los notebooks anteriores y utilizando la investigación que desarrolló sobre las relaciones entre las ventas y las demás variables para determinar que lags son los más adecuados para desarrollar modelos de forecasting.
  - b. Deberá aplicar al menos una operación de cada uno de los procesos de ingeniería de características. Indicando el criterio de selección de la operación. Si su dataset está completo o no es necesario aplicar alguna de esos proceso deberá forzar a que sea posible aplicarlos, es decir eliminar algunos datos de forma aleatoria y demás, lo importante es aplicar al menos una transformación de estas. No es necesario crear ningún operador ya que eso lo definiremos en la siguiente fase.  
Recuerde que los procesos de ingeniería de características son:
    - i. Imputación de variables numéricas,
    - ii. Imputación de variables categóricas,
    - iii. Codificación de variables categóricas,
    - iv. Tratamiento de outliers,
    - v. Transformación de variables numéricas.
    - vi. Escalado de características.
  - c. Al finalizar su proceso deberá listar con palabras el proceso deberá definir la lista de variables a las que se le aplicará dichas transformaciones así como los mecanismos de mapeo que considere necesarios (similar a los vistos en clase).