**1. Objective**

Goal of this project is to **predict the sales** of a Retail outlet based on the historical data provided for sales.

**2. Column Description**

**Item Identifier**: A code provided for the item of sale   
**Item Weight**: Weight of item   
**Item Fat Content**: A categorical column of how much fat is present in the item : ‘Low Fat’, ‘Regular’, ‘low fat’, ‘LF’, ‘reg’   
**Item Visibility**: Numeric value for how visible the item is   
**Item Type**: What category does the item belong to: ‘Dairy’, ‘Soft Drinks’, ‘Meat’, ‘Fruits and Vegetables’, ‘Household’, ‘Baking Goods’, ‘Snack Foods’, ‘Frozen Foods’, ‘Breakfast’, ’Health and Hygiene’, ‘Hard Drinks’, ‘Canned’, ‘Breads’, ‘Starchy Foods’, ‘Others’, ‘Seafood’.   
**Item MRP**: The MRP price of item   
**Outlet Identifier**: Which outlet was the item sold. This will be categorical column   
**Outlet Establishment Year**: Which year was the outlet established   
**Outlet Size**: A categorical column to explain size of outlet: ‘Medium’, ‘High’, ‘Small’.   
**Outlet Location Type**: A categorical column to describe the location of the outlet: ‘Tier 1’, ‘Tier 2’, ‘Tier 3’   
**Outlet Type** : Categorical column for type of outlet: ‘Supermarket Type1’, ‘Supermarket Type2’, ‘Supermarket Type3’, ‘Grocery Store’   
**Item Outlet Sales**: The amount of sales for an item.

There were two csv files provided for Training and Testing purposes namely Train\_Retail.csv and Test\_Retail.csv. Do Data imputation, EDA and feature engineering. Build a ML model and employ algorithms of your choice and use the best algorithm to predict the **Item Outlet sales** for the test data.