Big Mart Sales Prediction

The dataset is downloaded from:
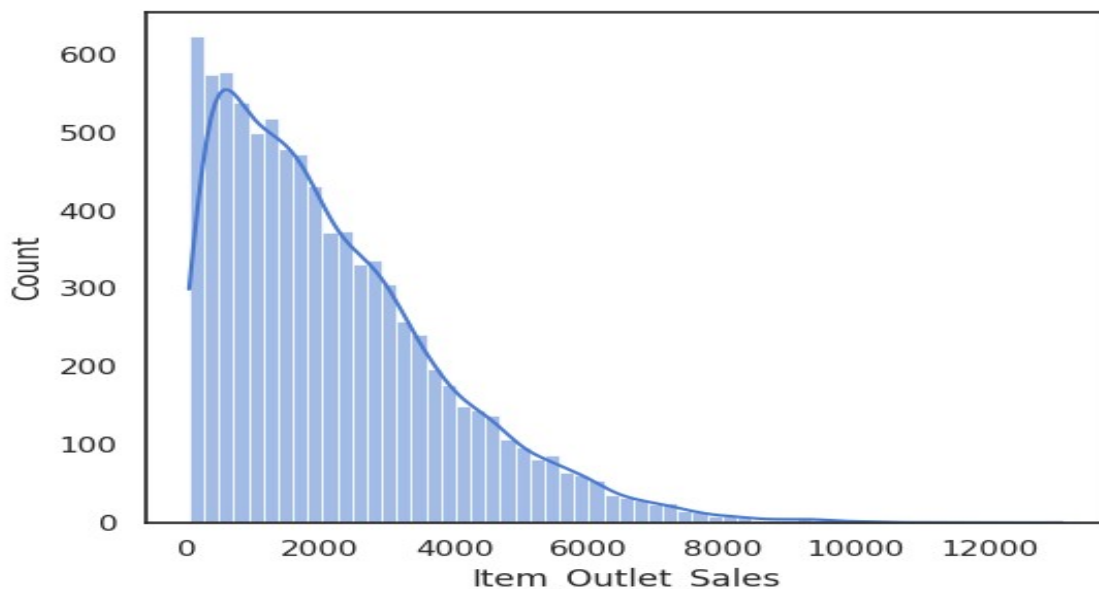
Variable Description

Independent Features:

- Item_Identifier ---- Unique product ID

- Item_Weight ---- Weight of product

- Item_Fat_Content ---- Whether the product is low fat or not

- Item_Visibility ---- The % of the total display area of all products in a store allocated to the particular product

- Item_Type ---- The category to which the product belongs

- Item_MRP ---- Maximum Retail Price (list price) of the product

- Outlet_Identifier ---- Unique store ID

- Outlet_Establishment_Year ---- The year in which the store was established

- Outlet_Size ---- The size of the store in terms of ground area covered

- Outlet_Location_Type ---- The type of city in which the store is located

- Outlet_Type ---- Whether the outlet is just a grocery store or some sort of supermarket
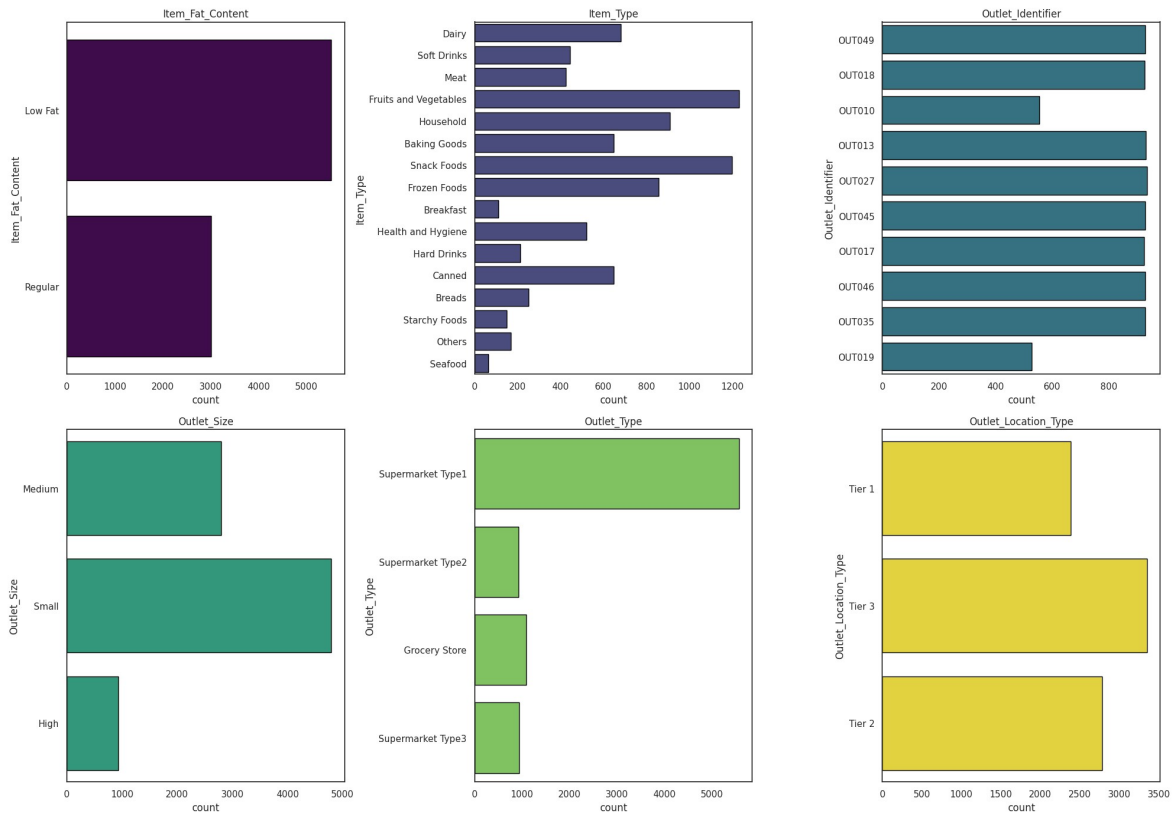
Dependent Feature:
Item_Outlet_Sales ---- sales of the product in t particular store. This is the outcome variable to be predicted.
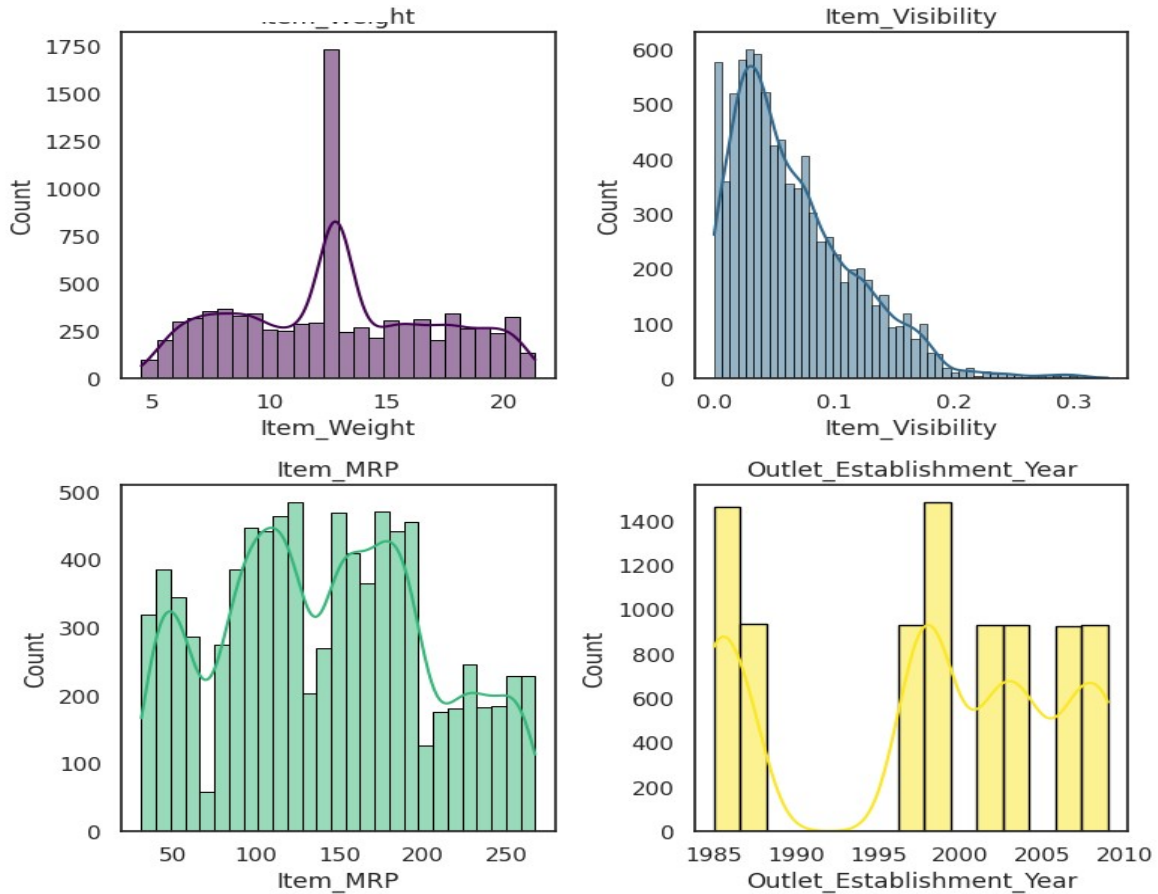
Descriptive Statistics

Dependent Feature

Categorical Features

Numerical Features

Machine Learning

Table 1: Metrics of different classifiers

Classifiers	Best CV	MSE	RMSE	MAE	R²
XGB	1206267.708	124134.819	1060.252	749.347	0.599
LR	-	1209515.003	1099.780	810.075	0.568
KNN	1528070.907	1445075.050	1202.113	864.266	0.484