Sales Predictions

Item Sales Prediction Analysis

Author: George Ajayi

Business problem:

We need to be able to predict the Item Outlet Sales value in a target
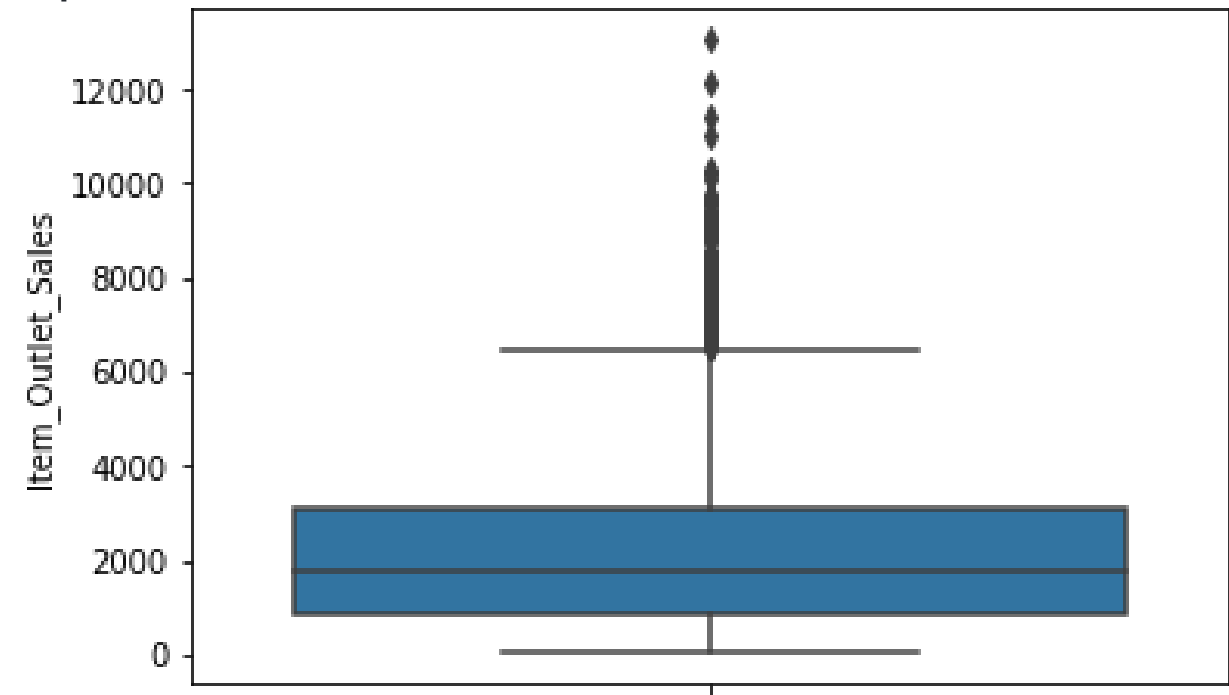
Data:

The Data source is a csv flat file Data named sales prediction which consist of 12 columns and 8524 rows. The dataset consist of 3 datatypes which are objects, floats and an integer. Both Numeric and Categorical data.

Methods

- Import Packages
- Load Data
- Explore Data
- Perform a Validation Split
- Preprocess the data
- Create the Model
- Fit the Model Pipeline on the Training Data and Make Predictions
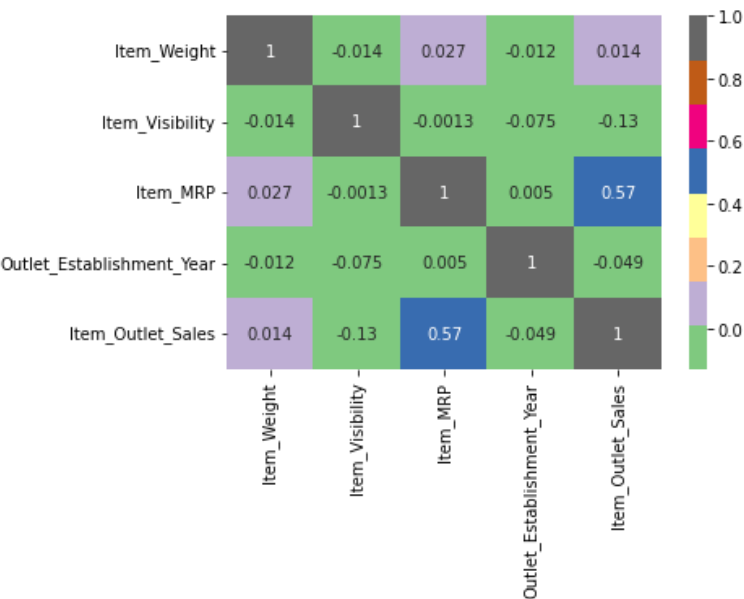- Evaluate the Model

Results

Boxplot Visualization

Above is a boxplot visualization of the (y) target which signifies the mean, median, mode of the item Outlet sales as we can see some outliers are also present in sales

Visual 2 Title

By my right is a heatmap

Visualization that shows a positive

Coefficient correlation between

2 columns in our data;

Item_MRP (the feature) and

Item_Outlet_Sales which is the target.

Summary of the Model and the Metrics

In this project, we see 2 different models and its evaluations. The linear regression model and the decision tree. In this project our aim is to predict the Item Outlet Sales which always comes with a continuous numeric values which also makes it a linear regression problem. We used the RMSE to evaluate the accuracy of the model while we used the R2 to evaluate the percentage of variation in the target that our model can explain.