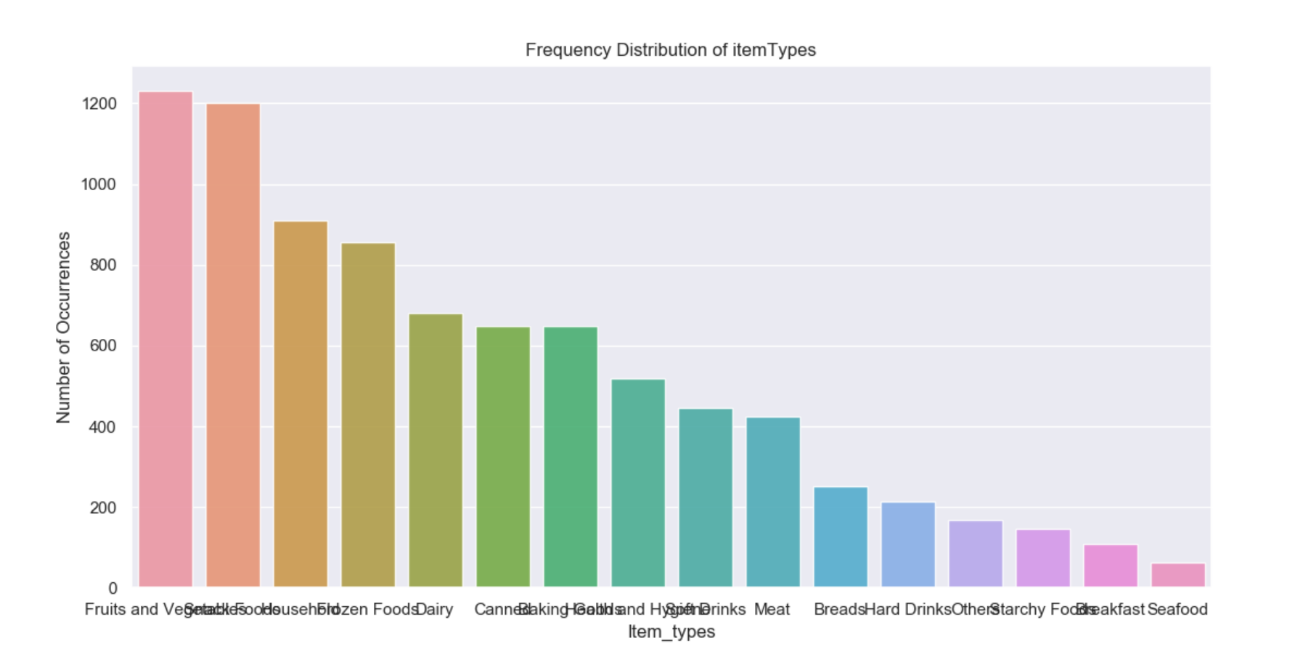
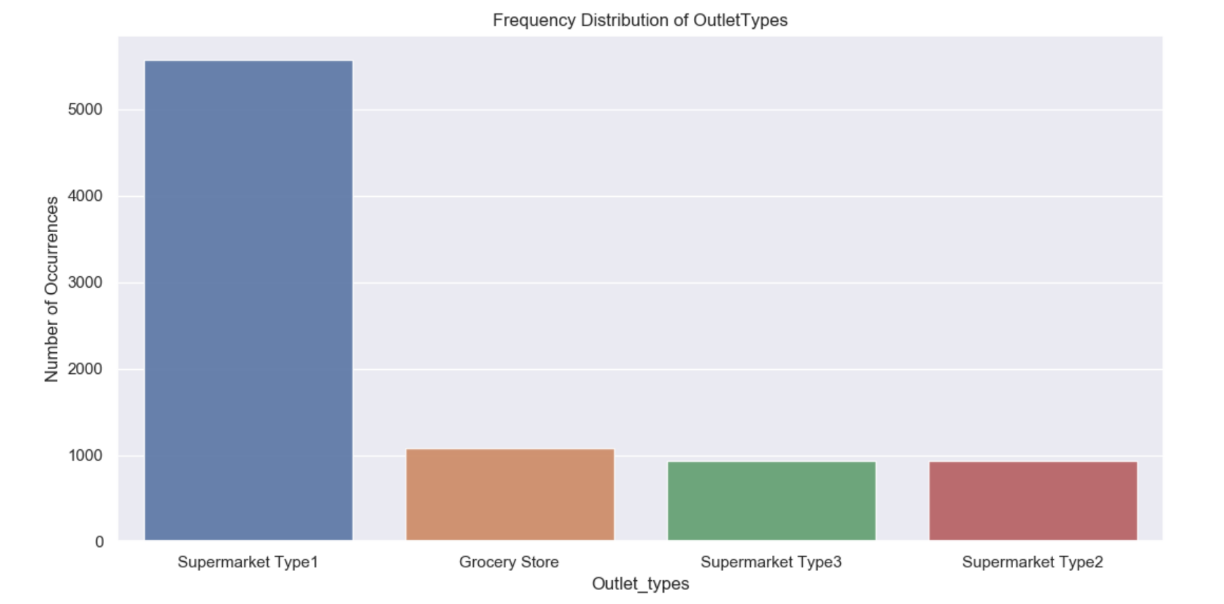
**Big Mart Sales Analysis**

Problem Statement - The data scientists at BigMart have collected 2013 sales data for 1559 products across 10 stores in different cities. Also, certain attributes of each product and store have been defined. The aim is to build a predictive model and find out the sales of each product at a store. Using this model, BigMart will try to understand the properties of products and stores which play a key role in increasing sales.

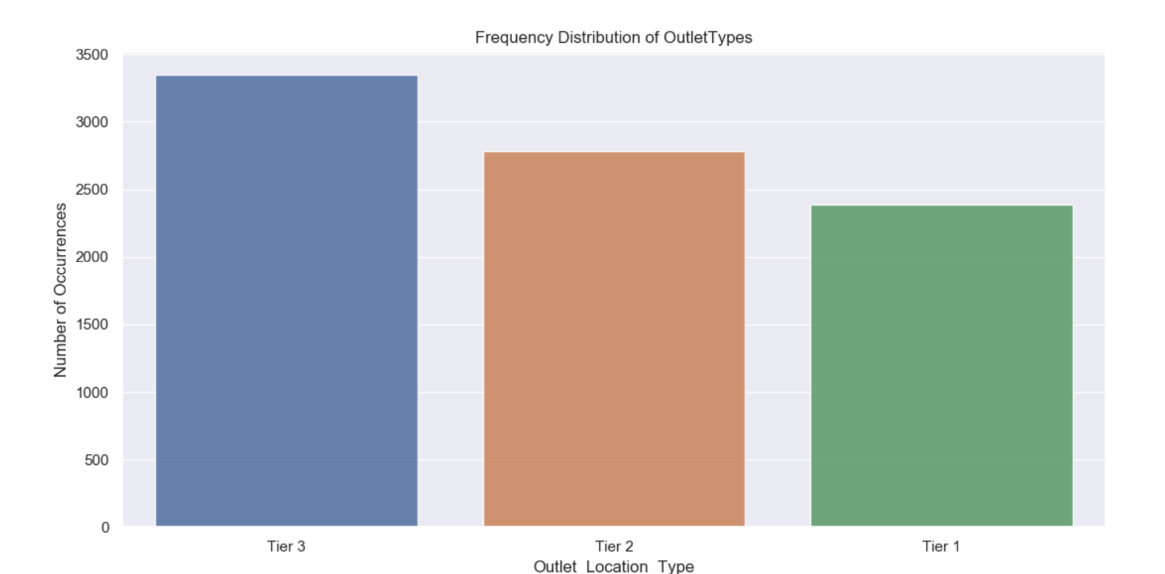
* Performed data wrangling as there were many missing values in the dataset.
* Performed some exploratory analysis on the relevant variables and derived the following conclusions:
  + - We see that the maximum frequency is for the item type Fruits and Vegetables.



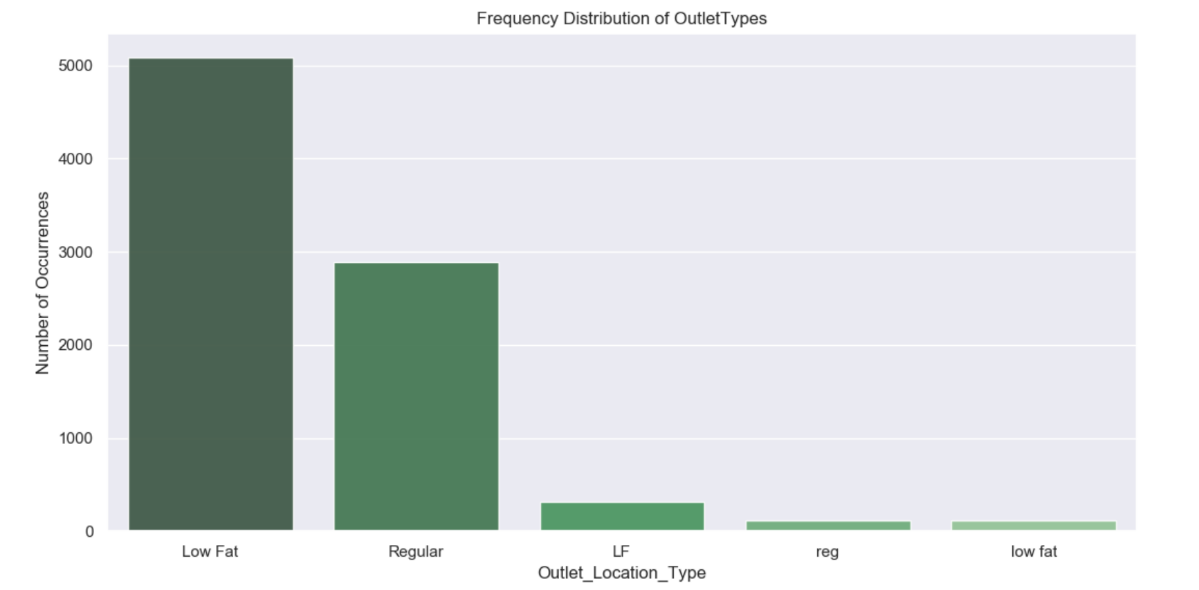
* + - We see that the maximum number of transactions were made from the Supermarket Type1 outlet type.



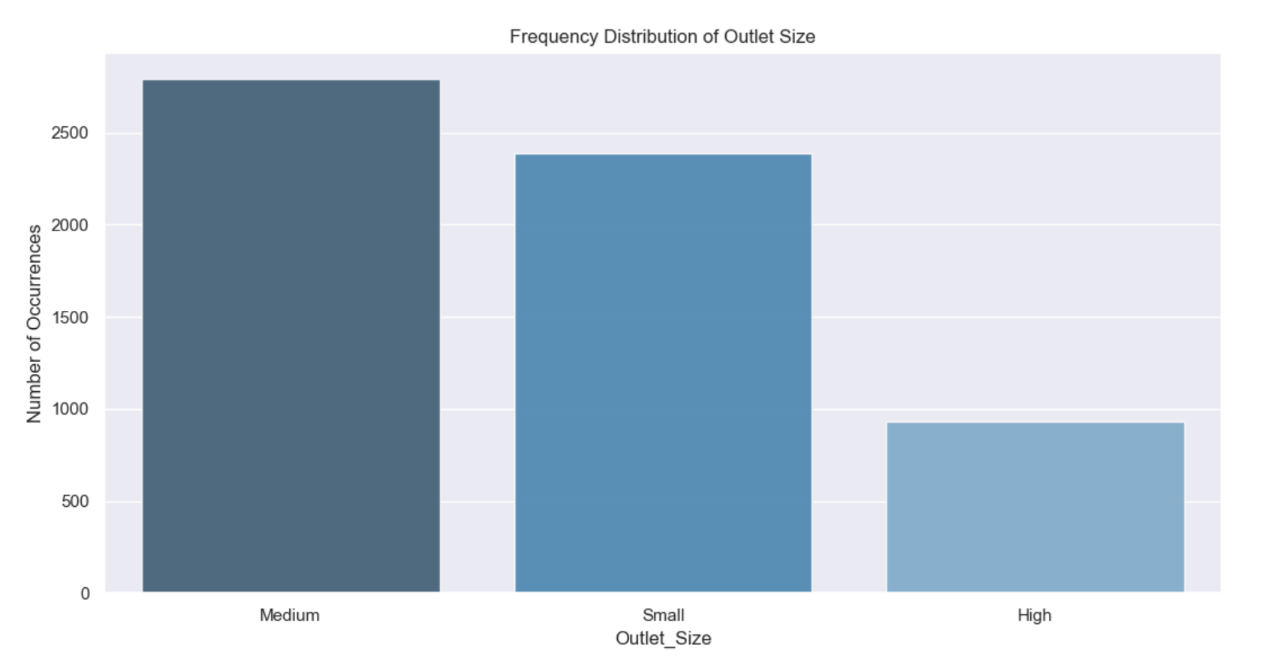
* We see that the maximum number of transactions were made from the Tier 3outlet location type.



* We see that the maximum frequency of the item sold are the item having low fat content.



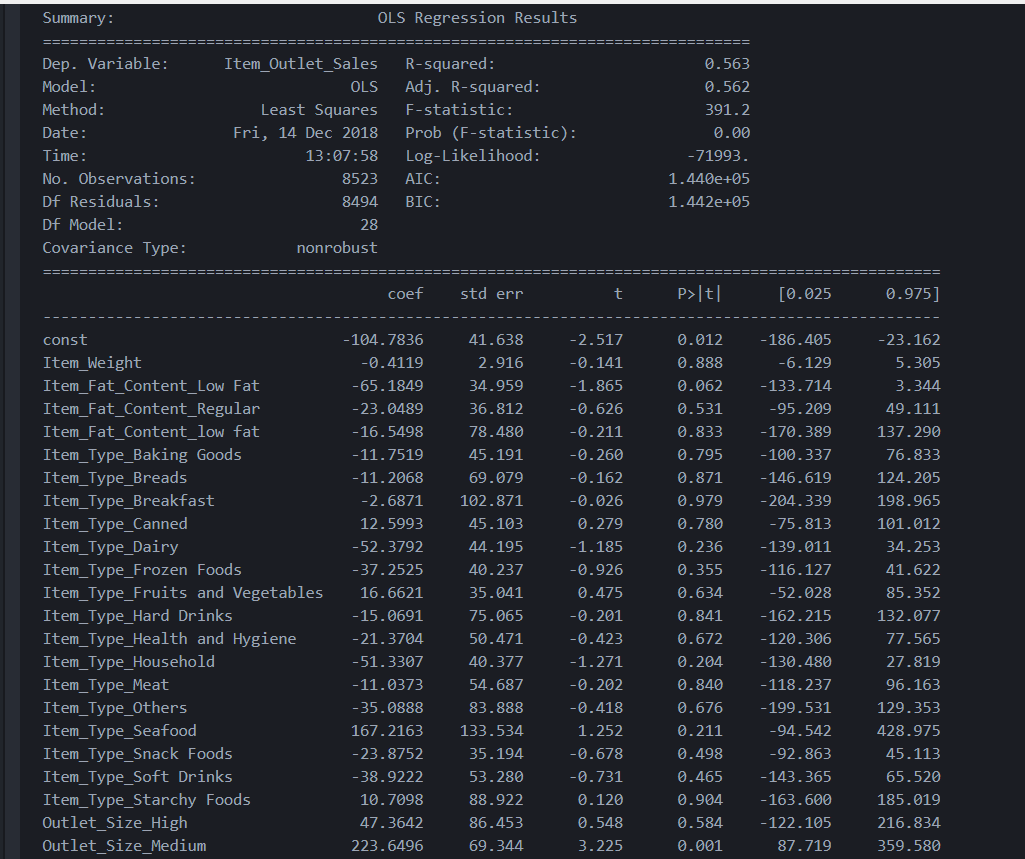
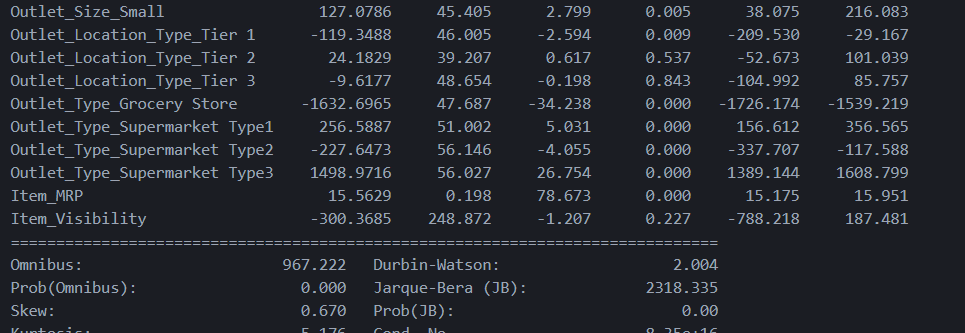
* We observe that the maximum number of items are sold from medium sized outlets.



* Performed One hot encoding to convert categorical variables to binary form so that they can be used in the model.
* Performed valid imputations for the missing values.

Modelling:

Below is the summary of the regression results:

However this model can be improved further .