**Big Sales Mart Problem**

**Dataset Description**

We have train (8523) and test (5681) data set, train data set has both input and output variable(s). We need to predict the sales for test data set.

* 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 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 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
* Item\_Outlet\_Sales: Sales of the product in the particulat store. This is the outcome variable to be predicted.