Problem Statement

# Dataset:

1. [train.csv](https://drive.google.com/file/d/1GbGugeaIiTyi313AGUmgO2I1XW0PpetN/view?usp=sharing)
2. [test.csv](https://drive.google.com/file/d/1IqkufcLMPCIXWvykB4pcwPLKZxOVGQEB/view?usp=sharing)

# About the Problem Statement:

The given dataset contains sales data for different products across different stores in different cities. Also, certain attributes of each product and store have been defined.

***The aim is to build a predictive model and predict the sales of each product at a particular outlet.***

# Data Dictionary:

|  |  |
| --- | --- |
| **Variable** | **Description** |
| 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 | Whetheroutlet is just a grocery store or sort of supermarket |
| Item\_Outlet\_Sales | Sales of the product in the particular store **(Target Variable)** |