# Project 7 (Sales Prediction)

**Task description**

Given a product and its corresponding details, the task is to predict the amount of sales based on the features.

**Dataset:** Train - 8k. The students are expected to split the dataset into train and validation as per the requirement.

**Training dataset link:** Click here

**Columns in the dataset**:

**Item_Identifier**: Unique identity number for a product.
**Item_Weight**: Indicates the weight of the product
**Item_Fat_Content**: Indicates the fat content- Low Fat / Regular
**Item_Visibility**: The percentage of total display area of 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 particular store. This is the outcome variable to be predicted.

**Test submission format:**

The test file will contain all the columns except the '**Item_Outlet_Sales**' as that is to be predicted. The students are expected to create a submission file in the following format, **[team_name]_SalesPredictionTask_submission.xlsx** (e.g. MLTitans_SalesPredictionTask_submission.xlsx ) with columns names **'Item_Identifier'** and **'Item_Outlet_Sales'** containing the unique ids of the test cases and the corresponding predicted sales amount, respectively.

**Evaluation**

The **Root Mean Squared Error**, will be calculated between the predicted '**Item_Outlet_Sales'** and the actual value of it.

Feel free to contact us with any inquiry.