**Data Analysis Project (Big\_Mart\_Sale\_Analaysis)**

This project focuses on forecasting the sales of Big Mart stores by leveraging historical sales data. The primary tool employed for this analysis is PowerBI.

**Project Overview**

* **Data Manipulation:**

Cleaned and transformed the dataset to ensure accuracy and consistency.

* **Data Visualization:**

Utilized PowerBI to create insightful visualizations of the data.

**Key Steps**

* **Data Cleaning:**

Tackled missing values and outliers to ensure data integrity.

* **Exploratory Data Analysis (EDA):**

Created various plots to visualize data distribution and uncover relationships.

**Dataset Overview**

* **Total Records**: 8,523 rows
* **Columns**: 12

**Key Columns and Observations:**

1. **Item Identifier**: A unique code for each item.
2. **Item Weight**: Numerical data representing the item's weight in kilograms. Some values are missing (about 17%).
3. **Item Fat Content**: Categorical data with values like "Low Fat" and "Regular," indicating the fat content of the items.
4. **Item Visibility**: Numerical data showing the item's visibility (display proportion) in stores. Contains values ranging from 0 to around 0.3.
5. **Item Type**: Categorical data that categorizes items into types like Dairy, Soft Drinks, Meat, etc.
6. **Item MRP**: The Maximum Retail Price (MRP) of each item, with a wide range of values, indicating varied product pricing.
7. **Outlet Identifier**: Unique identifier for each outlet/store.
8. **Outlet Establishment Year**: Year each outlet was established, indicating outlet age.
9. **Outlet Size**: Categorical data representing outlet sizes (e.g., Small, Medium, Large). Missing values in about 28% of records.
10. **Outlet Location Type**: Describes the type of city where the outlet is located, categorized as Tier 1, Tier 2, or Tier 3.
11. **Outlet Type**: Categorizes outlets into types like Supermarket Type1, Supermarket Type2, Supermarket Type3, and Grocery Store.
12. **Item Outlet Sales**: Numeric data on sales of each item in each outlet, which serves as the target variable for analysis.

**Key Insights & Summary Statistics:**

1. **Sales Data**:
   * **Total Sales** (sum of Item Outlet Sales): Can be calculated as a quick snapshot of overall revenue.
   * **Average Sales per Item**: Shows how much, on average, each item contributes to sales.
2. **Item Pricing and Type**:
   * **Average MRP by Item Type**: Reveals average pricing across item categories, with possible insights into premium versus budget items.
3. **Outlets and Growth**:
   * **Sales Growth** over the years can be explored to show outlet performance and sales trends based on establishment years.
   * **Sales by Outlet Type and Location**: Highlights the contribution of different store types and locations to total sales.

**Missing Data:**

* **Item Weight** and **Outlet Size** columns have missing values that might require handling for certain analyses, like filling with zero (0) or examining them separately.