**NAME:**

**ALI RAZA**

**ROLL NO:**

**103**

**AI PROJECT REPORT**

## ****Sales Forecasting Report****

### ****Objective****

The primary goal of this notebook is to analyze and forecast sales data using machine learning techniques. It involves:

* Understanding the sales dataset.
* Cleaning and preparing the data.
* Building predictive models to forecast sales.
* Evaluating model performance.

### ****Methods****

The following methodologies were applied in the analysis:

**1.Data Cleaning**:

* + Missing values in columns such as Item\_Weight and Outlet\_Size were filled using appropriate strategies (mean or mode).
  + The dataset was preprocessed for better compatibility with machine learning models.

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

* + Insights into the distribution of features such as Item\_Type, Outlet\_Type, and Item\_Outlet\_Sales were explored.
  + Relationships between variables like Item\_MRP and sales were analyzed using statistical and visual methods.

**3.Machine Learning Models**:

* + Regression models were developed to predict Item\_Outlet\_Sales.
  + The dataset was split into training and testing subsets to evaluate model performance.

**4.Performance Metrics**:

* + Models were evaluated using the **R-squared (R²)** metric to assess their predictive accuracy.

### ****Results****

**1.Dataset Overview**:

* + The dataset contains **8523 rows** and **12 columns**.
  + Missing values in columns such as Item\_Weight and Outlet\_Size were successfully handled.
  + Key features include Item\_Type, Outlet\_Type, and Item\_Outlet\_Sales.

**Key Insights**:

* + The Item\_MRP feature shows a strong correlation with Item\_Outlet\_Sales.
  + Outlets categorized by type and size exhibit varying sales patterns.

**Model Performance**:

* + A regression model achieved a high **R² value of ~0.876**, indicating strong predictive accuracy.
  + Another model's **R² value was ~0.502**, showing moderate performance.

**Predictions**:

* + The final predictions highlight varying sales trends across different outlet types and item categories.

### ****Conclusion****

The notebook effectively forecasts sales using regression models. The results demonstrate the importance of preprocessing and the significant impact of features like Item\_MRP and Outlet\_Type on sales. The high-performing model can be utilized for practical sales prediction, while further refinement is needed for the lower-performing model.