**Sales Prediction Using Linear Regression**

**Objective**

To build a statistical model that predicts company sales using multiple business drivers such as advertisement spending, store expansion, seasonal factors, and economic conditions — aiming to uncover key predictors that influence sales performance.

**Data Used**

* **Size**: 1,000 rows, simulated marketing campaign data
* **Features**:
  + Advertisement\_budget (numeric)
  + social\_media\_spend (numeric)
  + store\_openings (numeric)
  + economic\_index (numeric indicator of market conditions)
  + season (categorical: Spring, Summer, Autumn, Winter)
  + sales (target variable)
* **Preprocessing**:
  + One-hot encoding of the season variable
  + Conversion of all features to numeric format
  + Data split into independent variables (X) and target (y)

**Model Insights**

* Built an **OLS (Ordinary Least Squares)** linear regression model using statsmodels.
* Key significant predictors:
  + Advertisement\_budget: Positive and significant (p < 0.01)
  + store\_openings: Positive and significant (p < 0.01)
* Insignificant predictors:
  + social\_media\_spend, economic\_index, season dummies
* **Model performance**:
  + R² = 0.025 → Indicates the model explains only 2.5% of the variance in sales

**Limitations**

* **Low R²** suggests linear regression does not capture deeper or non-linear patterns in the data.
* Possible **multicollinearity**, as indicated by high condition number.
* Dataset may lack relevant predictors like product type, discounts, region, customer demographics.

**Business Recommendations**

* Focus marketing spend on **advertisement channels** that show measurable impact.
* Invest in **store expansion**, which has a strong correlation with increased sales.