Case Study

Context:

Assume that you are working as a Data Analyst at a consumer retail company that sells a wide range of products across multiple cities in Texas. Each customer transaction includes information about the product category, sales revenue, discount offered, and the resulting profit.

Question:

Can you reliably estimate the profit of a retail order using only its sales value?

Objective:

The finance and business strategy teams want a simple way to predict profit based on the sales value of an order. They are particularly interested in whether sales alone (without other variables like discount or product category) can be a strong predictor of profit.

Expected Approach:

1. Perform Exploratory Data Analysis on the data provided to uncover meaningful insights enabling the company to understand its consumer base better, other insights are welcome.
2. Fit a simple linear regression model after testing the necessary assumptions.
3. Tests preferred: Goldfeld Quandt Test, Bartlett's test, Durbin-Watson Test, Cook's Distances.
4. Model Evaluation should be based on MAPE and R-Squared value.

Expected Deliverable:

1. Interpretation of results from the model and the tests performed.

Timeline:

Candidates get 1-week time starting from the time they receive the case study to prepare a presentation.