Data Science Assignment

1. Data Analysis:

Prepare a document explaining how data analysis will be done
Datasets: CocaCola_Sales_Rawdata.xlsx, Company_Data.csv

2. Feature Engineering

How feature engineering will be done

3. Feature Importance

DataSet : Company_Data.csv

About the data: Let's consider a Company dataset with around 10 variables and 400 records. The attributes are as follows:

® Sales -- Unit sales (in thousands) at each location
® Competitor Price -- Price charged by competitor at each location
® Income -- Community income level (in thousands of dollars)
® Advertising -- Local advertising budget for company at each location (in thousands of dollars)
® Population -- Population size in region (in thousands)
® Price -- Price company charges for car seats at each site
® Shelf Location at stores -- A factor with levels Bad, Good and Medium indicating the quality of the shelving location for the car seats at each site
®Age -- Average age of the local population
® Education -- Education level at each location
® Urban -- A factor with levels No and Yes to indicate whether the store is in an urban or rural location
® US -- A factor with levels No and Yes to indicate whether the store is in the US or not The company dataset looks like this:
Problem Statement: A cloth manufacturing company is interested to know about the segment or attributes causes high sale. Prepare a document explaining feature importance.

4. Forecasting:

Forecast the CocaCola prices of Airlines Passengers data set and prepare a detailed document around that.

Dataset : CocaCola_Sales_Rawdata.xlsx

4.A Model Selection:

What will be the criteria for model selection and what are the different model combination you have tried and why?

4.B. Model Evaluation:

How model evaluation process will be done and what metrics you will use?

4.C. Extra Steps: You can add extra steps as well with proper explanation around that.