Learning Outcomes:

- Quick recap of the linear regression model and its assumptions
- Building linear regression models and interpret its output
- A recap of logistic regression, interpretation and ROC curves to arrive at the best probability threshold
- Time series decomposition, studying various components, adjusting seasonality, and stationarity in data
- Applying ARIMA, and Auto ARIMA models for forecasting time series data

---

1. Linear Regression
   a) Read the Housing data.
   b) Visualize the relationship between the median value of house with various attributes
   c) Split the data into train and test sets
   d) Apply simple and multiple linear regression
   e) Observe the summary, and residuals and compare them with the assumptions
   f) Observe the evaluation metrics
   g) Apply VIF or other techniques to improve the model

2. Logistic Regression
   a) Read the German Credit Ratings dataset
   b) Apply necessary preprocessing steps
   c) Split the data into train and test sets
   d) Apply logistic regression and observe the output
   e) Apply dimension reduction techniques and rebuild the model
   f) Study the error metrics before and after
   g) Apply ROC curve to arrive at the best probability threshold value for classification

3. Time series

   a) Read the HDFC data
   b) Consider the Close Price as Stock Price and aggregate the data at month level
   c) Adjust trend, seasonality and non-stationary behavior of data if necessary
   d) Apply ARIMA, and Auto ARIMA models to forecast and observe the errors