**DSCI 5240 - Data Mining and Machine Learning for Business**

**Assignment 1**

**Question 2:**

* 1. **Solution**: This is a regression problem where we predict the CEO salary based on the given features, and we are interested in inference. The value of n is 500 and the value of p is 4.
  2. **Solution**: This is a classification problem where we predict the likelihood of the product to be success or failure, and we are interested in prediction. The value of n is 20 and the value of p is 14.
  3. **Solution**: This is a regression time series problem where we predict the percentage change in the USD/Euro exchange rate based on the weekly observations. We are interested in prediction. The value of n is 52 and p is 4

**Question 4:**

1. **Solution**: A real-life application in which classification is used - Customers churn prediction in the Insurance industry.

The goal of the classification model is to predict/classify customers who are more likely to cancel their insurance policy.

The relevant features that can be used in the model are:

**Response**: Status of the policy (0 – closed, 1 – Active)

**Predictors**: Age of the customer, Gender of the customer, Insurance Premium of the customer, claim history, Change in the Premium amounts in last six months, state of residence, FICO score of the customer.

1. **Solution**: A real-life application in which regression is used - Predicting the price of the Insurance premium.

The goal of the regression model is to predict the ideal premium rates for the individuals based on the demographics and driving history.

The relevant features that can be used in the model are:

**Response**: Premium of the policy.

**Predictors**: Age, Gender, Years of driving experience, Number of claims in the past, Number of Driving tickets, Vehicle type, Ownership of the vehicle, state of residence.

1. **Solution**: A real-life application in which cluster analysis is used - Insurance industry. A cluster model is developed to group the customers based on the demographics, product ownerships, education, number of years associated with the company, Income. The Cluster assignment for the members helps companies in targeted marketing of their products.