"An Introduction to Statistical Learning" (ISLR) Chapter 1 & 2 and finish the following questions.
Page 63, Chapter 2.4 Exercises

Question 2

Explain whether each scenario is a classification or regression prob- lem, and indicate whether we are most interested in inference or pre- diction. Finally, provide n and p.

(a) We collect a set of data on the top 500 firms in the US. For each firm we record profit, number of employees, industry and the CEO salary. We are interested in understanding which factors affect CEO salary.

Understanding CEO Salary:

- **Type:** This is a regression problem because the outcome variable (CEO salary) is a continuous variable.
- **Interest:** The primary interest is in inference, understanding which factors (profit, number of employees, industry) affect CEO salary.
- **n (Sample Size):** The number of firms in the dataset (top 500 firms).
- **P(Number of Predictors):**The number of predictors includes profit,number of employees and industry
-

(b) We are considering launching a new product and wish to know whether it will be a success or a failure. We collect data on 20 similar products that were previously launched. For each product we have recorded whether it was a success or failure, price charged for the product, marketing budget, competition price, and ten other variables.

Product Success Prediction:

- **Type:** This is a classification problem because the outcome variable (success or failure) is categorical.
- **Interest:** The primary interest is in prediction, determining whether a new product will be a success or a failure based on similar products' data.
- **n (Sample Size):** Data on 20 similar products that were previously launched.
- **P(Number of Predictors):**The number of predictors includes price,marketing budget,competition price,and ten other variables

(c) We are interested in predicting the % change in the USD/Euro exchange rate in relation to the weekly changes in the world stock markets. Hence we collect weekly data for all of 2012. For each week we record the % change in the USD/Euro, the % change in the US market, the % change in the British market, and the % change in the German market.

USD/Euro Exchange Rate Prediction:

- **Type:** This is a regression problem because the outcome variable (% change in the USD/Euro exchange rate) is continuous.
- **Interest:** The primary interest is in prediction, forecasting the % change in the exchange rate based on weekly changes in world stock markets.
- **n (Sample Size):** 52-Data collected weekly for the entire year of 2012.

p (Number of Predictors): The number of predictors includes % change in the US market, % change in the British market, and % change in the German market

Question 4

You will now think of some real-life applications for statistical learning.

(a) Describe one real-life applications in which classification might be useful. Describe the response, as well as the predictors. Is the goal of each application inference or prediction? Explain your answer.

Credit Approval:

- **Response Variable:** Binary outcome (approved or rejected).
- **Predictors:** Income, credit score, debt-to-income ratio, employment status.

  Goal: Prediction. The objective is to predict whether an individual's credit application should be approved or rejected based on their financial information

(b) Describe one real-life applications in which regression might be useful. Describe the response, as well as the predictors. Is the goal of each application inference or prediction? Explain your answer.

Educational Performance Prediction:

- **Response Variable:** Continuous variable (student's academic performance).
- **Predictors:** Study hours, attendance, socioeconomic factors.
- **Goal:** Prediction. The objective is to predict a student's academic performance based on various factors, assisting in identifying potential areas for improvement.

(C) Describe one real-life applications in which cluster analysis might be useful.

Customer Segmentation:

- **Variables:** Various customer characteristics, purchase history.
- **Goal:** Inference. The goal is to identify groups of customers with similar characteristics and behaviors for targeted marketing strategies.