Mira Wadehra

Framing Real World Questions

- Hypothesis Testing:
  - a) Question 1: Does a new drug treatment reduce the average blood pressure of patients compared to placebo?
    - Null Hypothesis: The new drug treatment does not reduce the average blood pressure of patients compared to a placebo.
    - Alternative Hypothesis: The new drug treatment reduces the average blood pressure of patients compared to a placebo.
    - Data: Blood pressure measurements of patients in a controlled clinical trial, with one group receiving the new drug treatment and another group receiving a placebo.
  - b) Question 2: Is there a significant difference in customer satisfaction scores between two different customer service approaches?
    - Null Hypothesis: The two customer service approaches have no significant difference in customer satisfaction scores.
    - Alternative Hypothesis: There is a significant difference in customer satisfaction scores between the two customer service approaches.
    - Data: Customer satisfaction survey responses from customers who experienced both customer service approaches.
- Regression:
  - a) Question 1: Can we predict the sales revenue of a product based on its price, marketing budget, and competitor prices?
    - Data: Historical sales data, product price, marketing budget, and competitor prices.
  - b) Question 2: What is the relationship between the number of hours studied and students' exam scores?
    - Data: Exam scores and corresponding number of hours studied for a group of students.
- Classification:
  - a) Question 1: Can we classify customer churn based on their demographic information, usage patterns, and customer service interactions?
    - Data: Customer information (age, gender, etc.), usage data (usage duration, frequency, etc.), and customer service interaction records.
  - b) Question 2: Can we classify emails as spam or non-spam based on their content and email metadata?
    - Data: Emails labeled as spam or non-spam, along with their content and metadata (sender, subject, etc.).
- Clustering:

- a) Question 1: Can we identify distinct customer segments based on their purchasing behavior and demographic information?
    - Data: Customer purchase history, demographic data, and other relevant customer information.
- b) Question 2: Can we cluster news articles based on their content to identify different topics or categories?
    - Data: News articles, along with their content, titles, and other relevant metadata.

Please note that the specific data needed for each question depends on the chosen industry or domain and the availability of relevant data sources.