

## Analytics with R

1. A physician is evaluating a new diet for her patients with a family history of heart disease. To test the effectiveness of this diet, 16 patients are placed on the diet for 6 months. Their weights and triglyceride levels are measured before and after the study, and the physician wants to know if either set of measurements has changed. ([Data set: dietstudy.csv](#))
  
2. An analyst at a department store wants to evaluate a recent credit card promotion. To this end, 500 cardholders were randomly selected. Half received an ad promoting a reduced interest rate on purchases made over the next three months, and half received a standard seasonal ad. Is the promotion effective to increase sales? ([Data set: creditpromo.csv](#))
  
3. An experiment is conducted to study the hybrid seed production of bottle gourd under open field conditions. The main aim of the investigation is to compare natural pollination and hand pollination. The data are collected on 10 randomly selected plants from each of natural pollination and hand pollination. The data are collected on fruit weight (kg), seed yield/plant (g) and seedling length (cm). ([Data set: pollination.csv](#))
  - a. Is the overall population of Seed yield/plant (g) equals to 200?
  - b. Test whether the natural pollination and hand pollination under open field conditions are equally effective or are significantly different.
  - c. Test whether hand pollination is better alternative in comparison to natural pollination.

4. An electronics firm is developing a new DVD player in response to customer requests. Using a prototype, the marketing team has collected focus data for different age groups viz. Under 25; 25-34; 35-44; 45-54; 55-64; 65 and above. Do you think that consumers of various ages rated the design differently? (**Data set:** [dvdplayer.csv](#)).
  
5. A trial was designed to evaluate 15 rice varieties grown in soil with a toxic level of iron. Guard rows of a susceptible check variety were planted on two sides of each experimental plot. Scores for tolerance for iron toxicity were collected from each experimental plot as well as from guard rows. Data on the tolerance score of each variety and on the score of the corresponding susceptible check are captured. (**Data set:** [tolerance.csv](#)).
  - a. Are all the varieties having same tolerance?
  - b. What difference in result you are observing after using susceptible check as covariates in the analysis?
  
6. A survey was conducted among 2800 customers on several demographic characteristics. Working status, sex, age, age-group, race, happiness, no. of child, marital status, educational qualifications, income group etc. had been captured for that purpose. (**Data set:** [sample\\_survey.csv](#)).
  - a. Is there any relationship in between labour force status with marital status?
  - b. Do you think educational qualification is somehow controlling the marital status?
  - c. Is happiness is driven by earnings or marital status?