PROBLEM-1

BACKGROUND: The Lending Club is a peer-to-peer lending site where members make loans to each other. The site makes anonymized data on loans and borrowers publicly available.

BUSINESS PROBLEM:

Using lending club loans data, the team would like to test below hypothesis on how different factors affecting each other (Hint: You may leverage hypothesis testing using statistical tests)

    a. Interest rate is varied for different loan amounts (Less interest charged for high loan amounts)
    b. Loan length is directly affecting interest rate.
    c. Interest rate varies for different purpose of loans
    d. There is a relationship between FICO scores and Home Ownership. It means that people owning a home will have high FICO scores.

DATA AVAILABLE:

     LoansData.csv

The data have the following variables (with data type and explanation of meaning) ·
    Amount.Requested - *numeric*. The amount (in dollars) requested in the loan application. ·
    Amount.Funded.By.Investors - *numeric*. The amount (in dollars) loaned to the individual. ·
    Interest.rate – character. The lending interest rate charged to the borrower. · **Loan.length** - character. The length of time (in months) of the loan.
    · **Loan.Purpose –** categorical variable. The purpose of the loan as stated by the applicant. ·
    Debt.to.Income.Ratio – character. The % of the consumer's gross income going toward paying debts.
    · **State** - character. The abbreviation for the U.S. state of residence of the loan applicant. ·
    Home.ownership - character. Indicates whether the applicant owns, rents, or has a mortgage.
    · **Monthly.income** - categorical. The monthly income of the applicant (in dollars). · **FICO.range –** categorical (expressed as a string label e.g. "650-655"). A range indicating the applicants FICO score.
    · **Open.CREDIT.Lines** - numeric. The number of open lines of credit at the time of application.
    · **Revolving.CREDIT.Balance** - numeric. The total amount outstanding all lines of credit. · Inquiries.in.the.Last.6.Months - numeric. Number of credit inquiries in the previous 6 months.
    · **Employment.Length** - character. Length of time employed at current job.

PROBLEM - 2

BACKGROUND: When an order is placed by a customer of a small manufacturing company, a price quote must be developed for that order. Because each order is unique, quotes must be established on an order-by-order basis by a pricing expert. The price quote process is labor intensive, as prices depend on many factors such as the part number, customer, geographic location, market, and order volume. The sales department manager is concerned that the pricing process is too complex, and that there might be too much variability in the quoted prices. An improvement team is tasked with studying and improving the pricing process.

After interviewing experts to develop a better understanding of the current process, the team designed a study to determine if there is variability between pricing experts. That is, do different pricing experts provide different price quotes? Two randomly selected pricing experts, Mary and Barry, were asked to independently provide prices for twelve randomly selected orders. Each expert provided one price for each of the twelve orders.

BUSINESS PROBLEM: We would like to assess if there is any difference in the average price quotes provided by Mary and Barry.

DATA AVAILABLE:

  Price_Quotes.csv

The data set contains the order number, 1 through 12, and the price quotes by Mary and Barry for each order. Each row in the data set is the same order. Thus, Mary and Barry produced quotes for the same orders.

PROBLEM-3:

BACKGROUND: The New Life Residential Treatment Facility is a NGO that treats teenagers who have shown signs of mental illness. It provides housing and supervision of teenagers who are making the transition from psychiatric hospitals back into the community. Because many of the teenagers were severely abused as children and have been involved with the juvenile justice system, behavioral problems are common at New Life. Employee pay is low and staff turnover (attrition) is high.

A reengineering program was instituted at New Life with the goals of lowering behavioral problems of the kids and decreasing employee turnover rates. As a part of this effort, the following changes were made:

- Employee shifts were shortened from 10 hours to 8 hours each day. · Employees were motivated to become more involved in patient treatments. This included encouraging staff to run various therapeutic treatment sessions and allowing staff to have more say in program changes.
- The activities budget wasincreased.
- A facility-wide performance evaluation system was putinto place that rewarded staff participation andinnovation.
- Management and staff instituted a program designed to raise expectations about appropriate behavior from the kids. This included strict compliance with reporting of behavioral violations, insistence on participation in therapeutic sessions, and a lowered tolerance for even moderate behavioral infractions.

To determine the effectiveness of the reengineering effort, a data set comprising pre- and post-reengineering periods was compiled. The information contains two measures of behavioral problems. A *critical incident* occurs when a resident goes AWOL (leaves the premises without permission), destroys property (e.g., punching a hole in a wall or throwing furniture through windows), is caught in possession of street drugs, or engages in assault against other residents or staff members. A teenager is temporarily *removed from the facility* when s/he is sent to jail or back to a psychiatric hospital

BUSINESS PROBLEM: Determine what effect, if any, the reengineering effort had on the incidence behavioral problems and staff turnover. i.e To determine if the reengineering effort changed the critical incidence rate. Is there evidence that the critical incidence rate improved?

DATA AVAILABLE: Price_Quotes.csv

The data set contains 20 months of data; the first 13 months were prior to reengineering. The variables in the data include:

- **Reengineer:** Whether the month was before (Prior) or after (Post) reengineering   **EmployeeTurnover:** The percentage of employees who quit in a given month, out of the total number of employees

- **TRFF(%):** The percentage of residents who were temporarily removed from the facility, out of the total number of residents
- **CI (%):** The percentage of critical incident reports written that month, out of the total number of residents

PROBLEM - 4

BACKGROUND: Software development projects typically follow six basic phases: Requirements, design, implementation (and integration), testing (validation), deployment (installation) and maintenance. First, general requirements are gathered, and the scope of the functionality is defined. Then, alternative scenarios for the required functionality are developed and evaluated. Implementation, usually 50% or more of the development time, is the phase in which the design is translated into programs and integrated with other parts of the software – this is when software engineers actually develop the code. During the final phases, programs are tested, software is put into use, and faults or performance issues are addressed.

ApDudes, a developer of applications for tablet computers, was having difficulty meeting project deadlines; only 10% of their projects had been completed within budget and on time last year and that was starting to hurt business. The group's project manager wastasked with studying problems within the implementation phase. He found that software engineers were having difficulty prioritizing their work, and that they often became overwhelmed by the magnitude of the projects.

As a result, two changes were made. Each project was broken down into smaller, distinct tasks, or jobs, and each job was assigned a priority. The project manager believes that this classification and prioritization system would speed the completion of high priority jobs, and thus lower overall project completion time

BUSINESS PROBLEM: We will focus on the prioritization system. If the system is working, then high priority jobs, on average, should be completed more quickly than medium priority jobs, and medium priority jobs should be completed more quickly than low priority jobs. Use the data provided to determine whether this is, in fact, occurring.

DATA AVAILABLE:

  Priority_Assessment.csv

The data set contains a random sample of 642 jobs completed over the lastsix months. The variables in the data set are:
  Days The number of days it took to complete the job
  Priority The priority level assigned to that job

PROBLEM - 5

BackGround: Film on the Rocks is a summer movie series held at the world-renowned Red Rocks Amphitheatre, which is situated on a hillside of the Rocky Mountains in Morrison, Colorado fifteen miles west of Denver. The film series features classic films, and pre-show entertainment including bands and comedians. Among the features that have made Red Rocks an internationally-famous concert stage are sweeping views of Denver, outstanding acoustics, and cool, dry Colorado summers.

The series is jointly promoted by the Denver Film Society (DFS) and the City and County of Denver's Division of Theatres and Arenas(DT&A). It is marketed through various outlets including newspaper, radio, and the Red Rocks and Denver Film Society websites. Film on the Rocks patrons also benefit from corporate sponsorship. In return for on-site posters and banners at Red Rocks, and recognition in pre- show marketing materials, corporations donate funds that keep ticket prices low.

Although the Red Rocks Amphitheatre provides a cinematic experience unlike any other venue, there *are* tradeoffs. Red Rocks is a farther commute for most people than the local movie theater or movie rental store. Given the uphill walk to the amphitheatre from the parking lot, getting there can be challenging. And, as an outdoor venue, the viewing experience is dependent on the weather.

Patron satisfaction with Red Rocks as the venue for the film series is critical to its success. But, the series promoters would also like to increase attendance at the film series, and are unsure how to do this. Promoters recognize that they need a better understanding of the customer base, and of the current level of satisfaction. Knowing the demographics of those who attend the film series will help attract and expand corporate sponsorship. In addition, knowing which media outlets are most effective will provide information about how best to target future marketing campaigns.

To this end, the promoters conducted surveys during a recent Film on the Rocks season. Questionnaires were handed out at the entrance. Volunteers walked through the crowd to remind people about the free soft drink given to those who returned the survey.

BUSINESS PROBLEM: Use the survey results to address the following questions ·
    What is the overall level of customer satisfaction?
    · What factors are linked to satisfaction?
    · What is the demographic profile of Film on the Rocks patrons?
    · In what media outlet(s) should the film series be advertised?

DATA AVAILABLE: Data provided by the Denver Film Society

 Films.csv

The data set contains 330 surveys collected during three Film on the Rocks movies:
Ferris Bueller's Day Off, Old School, and Willy Wonka and the Chocolate Factory. The variables are:

Gender The patron's gender: 1 = male; 2 = female
Marital Status The patron's marital status: 1 = married; 2 = single **Age** The patron's age in years:
  1= 1-12; 2 = 13-30; 3 = 31-60; 4 = 60+ **Income** The patron's annual household income: 1 = Less
              than $50,000; 2 = $50- $100,000; 3 = $100,000+

Hear About The patron's response to this question: "How did you hear about Film on the Rocks?" Respondents could check any of the following that applied: 1 = television; 2 = newspaper; 3 = radio; 4 = website; 5 = word of mouth

The survey also contained four Likert-scaled questions about satisfaction; each of the following questions is coded: 1 = Excellent; 2 = Good; 3 = Average/Fair; 4 = Poor; 5 = Very Poor.

Signage "How was the signage directing you to Red Rocks?" **Parking** "How was the venue's parking?"
Clean "How was the cleanliness of the venue?"
Overall "How was your overall customer service experience?"