Module 2 Project

Performance Analysis, probability and probability trees

This project is due on Sunday at 11:59 pm

Individual work

Overview and Rationale

This assignment is designed to provide you with hands-on experience in applying experimental and theoretical probability theories. The data set is provided in an Excel workbook and contains a data you need as well as space to complete the tasks outlined in the assignment.

Course Outcomes

This assignment is directly linked to the following key learning outcomes from the course syllabus:

CO3: Apply the theory of probability to calculate events' likelihoods, understanding the differences between experimental and theoretical probabilities (the Law of Large Numbers), and calculate posterior probabilities by using the Bayes' Law with emphasis on applications

CO7: Interpret meaningful relationships and patterns in the data in relation to a given business question

Assignment Summary

Read the scenarios described below and follow the instructions in the project description below to analyze the data presented in the Excel workbook (**Module 2 Project data.xlsx**). Then complete a report summarizing the results of your calculations. Submit both the report and the Excel workbook (every person in each team should submit both files for grading). The Excel workbook contains all statistical work. The report should include all your findings along with important analysis and relevant figures and tables.

For this project, make a team of maximum 2 (**two**) students to solve the Excel file and to prepare the report. Yes, the report can be presented together, make sure your two names are listed in the title page. **IMPORTANT**: each one of you must submit the two files to your corresponding blackboards.

Format & Guidelines

Project Instructions

For this project you must submit:

1. An R File containing your code-based calculations.
2. A DOC or PDF report explaining your observations, analysis of the data, and conclusions.
   a. Report should include an introduction to explain the topics, for example, why probability is important, what do you know about the topics and its applications in real life, etc.
   b. The report should include the following sections:
      i. The title Page
      ii. The main Body, which must include the following sections:
         Introduction
         Analysis
         Conclusions
         References.
   c. Report must be presented as indicated by the instructor in class, and in **Maximum 5** pages.
   d. Your introduction and conclusions must be supported by at least two (2) academic references.

Create a good introduction:

On your notes, make a list of the topics you must cover in your report (this is a good practice for any report).

Make sure that all the topics are addressed in your introduction, use academic references when needed. Be brief.

Part 1: Using the data Excel in Tab Data of the workbook provided:

The quality control manager of a manufacturing company has randomly selected 50 of the company's recent projects. Each project has four attributes:

1.   A **PIP** number that identifies each project uniquely, a **Quality** score, a **Speed** value, which is the time duration in days during which the project was completed, and the **Cost** of the project.

2.   In order to evaluate the efficiency of each project, the manager assigns values to the Quality, Speed, and Cost attributes.

•   (Event **Q**) The Quality criterion is satisfied if Quality higher than 550.

•   (Event **S**) The Speed criterion is satisfied if project completed in less than 11 days.

•   (Event **C**) The Cost criterion is satisfied if cost less than $180,000.

3.   The manager then assigns an overall score (from zero to 7) to each project depending on how many of the three criteria have been satisfied by that project.

- P(score = **0**) if none of the three criteria is satisfied.
- P(score = **1**) if only the Quality criterion is satisfied but none of the other two are satisfied
- P(score = **2**) if only the Speed criterion is satisfied but none of the other two are satisfied.
- P(score = **3**) if only the Cost criterion is satisfied but none of the other two are satisfied.
- P(score = **4**) if Quality and Speed are satisfied but the Cost is not satisfied
- P(score = **5**) if Quality and Cost are satisfied but the Speed is not satisfied.
- P(score = **6**) if Speed and Cost are satisfied but the Quality is not satisfied.
- P(score = **7**) if all three criteria are satisfied.

Tasks:

Q.1. Complete **Tables 1** and **2** in the designated cells in the Excel worksheet provided. Note that cells should contains formulas that clearly show the details of in-cell calculations. Copy and paste the tables in the tab "Your tables" and follow the recommendations indicated in that tab before presenting the tables on your report. In your report, remember to indicate what do the results mean, what conclusions you can draw from them.

The conclusions should (1) be based on the main question for the particular problem, (2) summarize the major observations, and (3) be used to clarify and support your arguments.

Q.2. Using information from tables 1 and 2, create a **Venn diagram** to display the relationships among Quality, Speed and Cost (observe their counts from table 1 and the 8 regions from table 2). Present the diagram in your report and draw meaningful conclusions.

Q.3. Fill **table** 3 to complete the following steps about specific probability of events.

    a) What percentage satisfied exactly two of the three criteria?
    b) Of those who satisfied Quality, what percentage also satisfied Cost?
    c) Of those who satisfied Cost, what percentage also satisfied Speed?
    d) Of those who did not satisfy Speed, what percentage satisfied Quality and Cost?
    e) Of those who satisfied Quality, what percentage also satisfied Speed but did not satisfy the Cost?
    f) Of those who did not satisfy Cost, what percentage satisfied the Speed criterion?
    g) Of those who satisfied at least one of the three criteria, what percentage satisfied exactly one criterion?
    h) Of those who satisfied Cost, what percentage also satisfied Speed but did not satisfy the Quality?

In your report, present your results and make meaningful observations about the application of probability in the context of the data you just analyzed. Present a practical example of a real-life scenario.

Part 2: Probability trees

Imagine you are running a customer satisfaction survey regarding insurance coverage and MD office visits.

Q.4. From a sample of 250 people, 143 men and 107 women, you found that 92 men and 83 women have health insurance, the rest not. Forty (40) of the men with health insurance and 18 of the men without insurance, visited the MD office more than 5 times last year. In the other hand, 45 of the women with insurance and 9 of the women without insurance, visited the MD office more than 5 times last year.

Giving this information, construct a probability tree. You have to add the missing data to complete the tree, use your imagination.

Using this tree, ask yourself 2 probability-based questions. These two questions are your choice, look at your tree and see which information is interesting to you. The task is to demonstrate that you can build probability trees and to show how they can be used to answer practical questions.

Provide the tree with a good presentation format and include it on your report. Make sure all the labels are easy to read.

Q.5. Now imagine you survey 120 young people under 15 (57 boys and 63 girls) and ask the same questions about health insurance and number of visits to the MD office. How many boys and girls have insurance, and from those groups, how many visited the MD office more or less than 5 times last year? Again, you put those numbers (use your imagination). Then create a probability tree with this information and ask yourself 2 practical probability-based questions about the data that seem interesting to you.

Remember: the probability trees must contain branches for all the groups, they should display all data (N values and probabilities at each branch), and each level must add to 100% of its previous branch.

Also present this tree in your report and explain what are the advantages or disadvantages of their application.

Conclusions section

Make general and meaningful conclusions about the whole project, the techniques you used (probability, Venn Diagrams, probability trees), what are their advantages and what did you learn from their application.

References

Do not forget to add at least 2 academic references, books are the best. No web links are accepted unless they are to a trustful 100% academic source.

Please remember: your reports are very important, in class or at work. Make them look professional, make them as short as possible but containing all the relevant information. Let me know you have learnt and understood the statistical tools you just used, and for extra points, use deep critical thinking to provide examples of practical applications in real life scenarios.

Rubric

Category	Above Standard	Meets Standards	Approaching Standards	Below Standards	Not Evident
Excel (or R): Problem Modeling & Set-up	Thoroughly and concisely modeled the problem in Excel (or R) for each method	Accurately modeled the problem in Excel (or R) for each method	Satisfactorily modeled the problem in Excel (or R) for each method.	Partially modeled the problem in Excel (or R) for each method, but there are some gaps in the problem modeling and setup	Did not submit or incompletely modeled the problem in Excel (or R)
Excel (or R): Problem Solution & Accuracy	Thoroughly and efficiently obtained correct and accurate solutions in Excel (or R) by using the appropriate analytic tools of the software	Thoroughly obtained accurate solutions in Excel (or R) by using the appropriate analytic tools of the software	Satisfactorily obtained correct solutions in Excel (or R) by using the appropriate analytic tools of the software	Partially obtained accurate solutions in Excel (or R) by using the appropriate analytic tools of the software	Did not submit or did not obtain accurate solutions in Excel (or R) using the appropriate analytic tools of the software
Word/Report: Problem Description & Introduction	Thoroughly provided a summary of the problem descriptions and introduced the problem using rich and significant ideas	Thoroughly provided a summary of the problem descriptions and problem introduction	Satisfactorily provided a summary of the problem descriptions and problem introduction	Partially provided a summary of the problem descriptions and problem introduction	Did not submit or did not provide a summary of the problem descriptions and problem introduction
Word/Report: Description of Problem Analysis	Thoroughly and accurately described the analytic concepts and theories used in analyzing the problem	Accurately described the analytic concepts and theories used in analyzing the problem	Satisfactorily described the analytic concepts and theories used in analyzing the problem	Partially described the analytic concepts and theories used in analyzing the problem	Did not submit or did not provide a summary of the problem descriptions and problem introduction
Word/Report: Description of Conclusions	Thoroughly described the conclusions and results obtained in the project using a high level of critical thinking and reasoning	Thoroughly described the conclusions and results obtained in the project	Satisfactorily described the conclusions and results obtained in the project	Partially described the conclusions and results obtained in the project	Did not submit or did not describe the conclusions and results obtained in the project
Word/Report: Writing Mechanics, Title Page, & References	Completely free of errors in grammar, spelling, and punctuation; and completely correct usage of title page, citations, and references. The report contains a minimum of 1000 words	There are no noticeable errors in grammar, spelling, and punctuation; and completely correct usage of title page, citations, and references. The report contains a minimum of 1000 words	There are very few errors in grammar, spelling, and punctuation; and completely correct usage of title page, citations, and references. The report contains a minimum of 1000 words	There are more than five errors in grammar, spelling, and punctuation; or the usage of title page, citations, and references are incomplete; or the report contains less than 1000 words	Did not submit; or there are many errors in grammar, spelling, and punctuation; or the usage of title page, citations, and references are totally incomplete; or the report contains very few words