7COM1079-0901-2024 - Team Research and Development Project

Final report title: "Does the average length of hospital stay differ between patients admitted for acute care and those admitted for childbirth care?"

Group ID: A283

Dataset number:

Prepared by: Piyush Kumar Mallick 23079409

Ahmad Mujtaba Khan 23111308

Ankit Joshi 22032491

Haider Abid 23081929

Jeswanth Kumar Botcha

***Please make sure*** *the document spelled correctly (including image labels, section headings, and table of contents). Please use correct punctuation.*

*Make sure your report is grammatically correct.*

University of Hertfordshire

Hatfield, 2024

Table of Contents

1. Introduction
   1. Problem statement and research motivation
   2. The data set
   3. Research question
   4. Null hypothesis and alternative hypothesis (H0/H1)

1. Background research
   1. Research papers (at least 3 relevant to your topic / DS)
   2. Why RQ is of interest (research gap and future directions according to the literature)

1. Visualisation
   1. Appropriate plot for the RQ *output of an R script (NOT a screenshot)*
   2. Additional information relating to understanding the data (optional)
   3. Useful information for the data understanding

1. Analysis
   1. Statistical test used to test the hypotheses and output
   2. The null hypothesis is rejected /not rejected based on the p-value

1. Evaluation – group’s experience at 7COM1079
   1. What went well
   2. Points for improvement
   3. Group’s time management
   4. Project’s overall judgement
   5. Comment on GitHub log output

1. Conclusions
   1. Results explained.
   2. Interpretation of the results
   3. Reasons and/or implications for future work, limitations of your stud

1. Reference list

Harvard (author, date) format.

1. Appendices
2. R code used for analysis and visualisation.
3. GitHub log output.

*The list below outlines the chapter/subchapter numbers, names, word count limits, and explanations of what to write in each section.*

1. Introduction

* 1. Problem statement and research motivation **(100 words)**
* *What is the problem in the area we want to learn more about (motivation for study).*
* *Use at least one citation from related literature for top marks.*
  1. The data set **(75 words)**
  2. Research question **(50 words).** *Explain how you are going to answer your RQ.* **(50 words)**
  3. Null hypothesis and alternative hypothesis (H0/H1) **(100 words)**

1. Background research
   1. Research papers (at least 3 relevant to your topic / DS) **(200 words)**

* *Was the data set used for some research papers?* *Reference at least 3 relevant research papers to your topic / DS.*
  1. Why RQ is of interest (research gap and future directions according to the literature) **(100 word**s)

1. Visualisation
   1. Appropriate plot for the RQ *output of an R script (NOT a screenshot)* (**50 words)**

* *Explain the choice of the plot.*
* *Anything on the plot from R is not counted towards word count limit*
* *Make sure that the plot is from output of an R script (NOT a screenshot)*
* *Make sure that the plot has a caption or title, X and Y-axis labels, legend if appropriate and units.*
* *Make sure the title or caption and axis labels are informative.*
  1. Additional information relating to understanding the data (optional) (**50 words)**
* *Per plot: explain the purpose and insights.*
  1. Useful information for the data understanding (**50 words)**
* *Summarise key observations from the plot.*

1. Analysis
   1. Statistical test used to test the hypotheses and output (**75 words)**

* *Explain the choice of the test.*
* *Make sure the test is appropriate for the RQ and data.*
  1. The null hypothesis is rejected /not rejected based on the p-value (**100 words)** *(interpret the results)*

1. Evaluation – group’s experience at 7COM1079
   1. What went well **(75 words)**
   2. Points for improvement **(75 words)**
   3. Group’s time management (**50 words)**
   4. Project’s overall judgement (**50 words)**
   5. Note any changes to group since submission of Assignment 1. Add new or amended GitHub Ids for new members **(75 words, write only if applies to your group arrangements)**
   6. Comment on the GitHub log output **(50 words)**

*Please comment on the GitHub log output, and refer to it as being placed into**Appendix B.*

*From your Git log, select the three most significant commits during this project and include the following for each:*

1. ***Commit Message:*** *[Insert Commit Message] Brief explanation of the broader impact of the change*
2. ***Commit Message:*** *[Insert Commit Message] Brief explanation of the broader impact of the change*
3. ***Commit Message:*** *[Insert Commit Message] Brief explanation of the broader impact of the change*

1. Conclusions
   1. Results explained (**75 words)**
   2. Interpretation of the results (**75 words)**

* *Interpretation of what the results mean in terms of your RQ and the effect this may have on your population and the wider context of your topic.*
  1. Reasons and/or implications for future work, limitations of your study (**50 words)**

1. Reference list ***(not included in the work count)***

Harvard (author, date) format.

1. Appendices
2. R code used for analysis and visualisation ***(not included in the word count)***

Analysis.R code with the appropriate statistics to test the hypotheses.

* ***No word count****, but ensure the code is without redundant lines, well-commented and produces the correct output.*
* *Make sure it runs (look in Rscript.log for output from a statistical test)*
* *It should compute appropriate statistics to test the hypotheses*

1. GitHub log output.

1.Introduction

1.1 Problem Statement and research motivation

This dataset is interesting to us because it shows how, over the years, technological advancements in medical care have affected the average length of stay at a hospital.

1.2 The Data Set

This dataset is interesting to us because it shows how, over the years, technological advancements in medical care have affected the average length of stay at a hospital.  
  
  
Our Independent variable is: Subject   
 This Independent variable datatype is: Nominal.  
  
Our Dependent variable is: Value  
 This Dependent variable datatype is: Interval

1.3 Research Question

Research Question :- **How does the average length of hospital stay differ between patients admitted for an acute care and childbirth care?”**

1.4 Null Hypothesis And Alternative Hypothesis

Null hypothesis (H0): There is no difference in the mean of the hospital stay between acute care and childbirth care.  
  
Alt hypothesis (H1): There is a difference in the mean of the hospital stay between acute care and childbirth care.

2.Background Research

2.1 Research Papers

**1."Length of Stay After Childbirth in 92 Countries and Associated Factors in 30 Low- and Middle-Income Countries: Compilation of Reported Data and a Cross-sectional Analysis from Nationally Representative Surveys"**  
*Authors*: Campbell, O. M. R., Cegolon, L., Macleod, D., & Benova, L.  
*Published in*: PLOS Medicine, 2016.  
*Summary*: This study analyzed postpartum hospital stay durations across 92 countries, revealing that the mean length of stay ranged from 1.3 to 6.6 days. Specifically, for singleton vaginal deliveries, stays ranged from 0.5 to 6.2 days, and for cesarean-section deliveries, from 2.5 to 9.3 days. The research highlights significant global variations and discusses factors influencing these differences.

2. study related to length of stay in acute conditions in countries

3.” Postpartum Length of Hospital Stay Among Obstetric Patients in a Tertiary Care Centre in Eastern Nepal: A Cross-Sectional Study"

2.2 Why RQ is of interest

This Research Question interests us because in this we get to know how advancements in medical care ,technologies and practices has affected the average length of stay of patients in acute conditions and in child birth/ post partum stays.

3.Visualization

3.1 Appropriate Plot for RQ

3.2 Additional information relating to understand the data.

3.3 Useful information for the data understanding

4.Analysis

4.1 Statistical test used to test the hypothesis and output

For this Experiment we have used the Welch two sample t-test because our Research Question is about comparison of means of two types of stays in hospitals , The dependent variable the number of days of stay or the length of stay looks normally distributed and there’s exactly two values of independent variable.

4.2 The null hypothesis is rejected / not rejected based on the p-value

Test results :- t = 40.84 , df = 1530.5 , p-value < 2.2e-16

Alternative hypothesis is true and the Null hypothesis is rejected/false because the p-value of the test is too small so it almost negligible differences in means between group Acute and group CHILDBIRTH is not equal to 0 95 percent confidence interval :

4.760922 5.241327

Mean in group acute :- 8.550328 , Mean in group CHILDBIRTH :- 3.549204

5.Evaluation

5.1 What went well

5.2 Points for improvement

5.3 Group’s time management

5.4 Project’s overall judgement

5.5 Note any changes to group since submission of Assignment

1.Add new or amended GitHub Ids for new members

5.6 Comment on the Github log output

6.Conclusions

1.Results explained

2.Interpretation of the results

3.Reasons and/or implications for future work,limitations of your study

7.Reference List

8. Appendices

A. R code used for analysis and visualization Analysis R code with the appropriate statistics to test the hypothesis.

B. Github log output