Mini Project 2

PSTAT100: Data Science Concepts and Analysis

Instructor: Ali Abuzaid

STUDENT NAME

- Phuc Lu (pdlu)
- Valerie De La Fuente (valeriedelafuente)

Instructions

- This mini project is designed to give you practical experience with real-world data using R and Shiny. You'll create an interactive web application that allows users to explore and visualize a dataset.
- Work in groups of **2 students** from the same discussion section.
- Individual submissions will not be accepted.
- Please use the provided `MP 2.qmd` file to type your Documentation and Presentation and submit it as a PDF file. You can utilize `RStudio`for this purpose. For guidance, refer to the Tutorial: Hello, Quarto).
- Please submit a `.zip` file that includes:

Your `app.R` file (fully working Shiny app).
A short project report (PDF).
Reminder: If your app fails to open or the .zip is incorrect, you will receive a score of **ZERO**. Test everything before submission.

Due Date

Due Date: Sunday, June 1, 2025, 11:59 PM

0.1 2 Tasks:

0.1.1 2.1.1: Data Loading:

```
library(tidyverse)
data(infert)
```

Variable	Description	Value
education	years of education	0–5 years, 6–11 years, 12+ years
age	age in years	22-24
parity	number of prior live births	1-6
induced	number of prior induced abortions	0, 1, 2 or more
case	case status	case, control
spontaneous	number of prior spontaneous abortions	0, 1, 2 or more
stratum	id for each matched set or group	1–83
pooled.stratum	represent a group or broader classification of strata (may be used for conditional logistic regression)	1 - 63

0.1.2 2.1.2: Data Preparation:

```
infert %>% head()
```

```
  education age parity induced case spontaneous stratum pooled.stratum
1    0-5yrs  26      6       1    1           2       1              3
2    0-5yrs  42      1       1    1           0       2              1
3    0-5yrs  39      6       2    1           0       3              4
4    0-5yrs  34      4       2    1           0       4              2
5   6-11yrs  35      3       1    1           1       5             32
6   6-11yrs  36      4       2    1           1       6             36
```

```
library(stringr)
# remove "yrs" from education values.
infert_ds <- infert %>% mutate(education = str_remove(education, "yrs"))
infert_ds %>% head()
```

```
  education age parity induced case spontaneous stratum pooled.stratum
1       0-5  26      6       1    1           2       1               3
2       0-5  42      1       1    1           0       2               1
3       0-5  39      6       2    1           0       3               4
4       0-5  34      4       2    1           0       4               2
5      6-11  35      3       1    1           1       5              32
6      6-11  36      4       2    1           1       6              36
```

```
summary(infert_ds)
```

```
  education               age             parity         induced      
 Length:248         Min.   :21.00   Min.   :1.000   Min.   :0.0000  
 Class :character   1st Qu.:28.00   1st Qu.:1.000   1st Qu.:0.0000  
 Mode  :character   Median :31.00   Median :2.000   Median :0.0000  
                    Mean   :31.50   Mean   :2.093   Mean   :0.5726  
                    3rd Qu.:35.25   3rd Qu.:3.000   3rd Qu.:1.0000  
                    Max.   :44.00   Max.   :6.000   Max.   :2.0000  
      case          spontaneous        stratum      pooled.stratum  
 Min.   :0.0000   Min.   :0.0000   Min.   : 1.00   Min.   : 1.00  
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:21.00   1st Qu.:19.00  
 Median :0.0000   Median :0.0000   Median :42.00   Median :36.00  
 Mean   :0.3347   Mean   :0.5766   Mean   :41.87   Mean   :33.58  
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:62.25   3rd Qu.:48.25  
 Max.   :1.0000   Max.   :2.0000   Max.   :83.00   Max.   :63.00  
```

```
str(infert_ds)
```

```
'data.frame':   248 obs. of  8 variables:
 $ education     : chr  "0-5" "0-5" "0-5" "0-5" ...
 $ age           : num  26 42 39 34 35 36 23 32 21 28 ...
 $ parity        : num  6 1 6 4 3 4 1 2 1 2 ...
 $ induced       : num  1 1 2 2 1 2 0 0 0 0 ...
 $ case          : num  1 1 1 1 1 1 1 1 1 1 ...
 $ spontaneous   : num  2 0 0 0 1 1 0 0 1 0 ...
 $ stratum       : int  1 2 3 4 5 6 7 8 9 10 ...
 $ pooled.stratum: num  3 1 4 2 32 36 6 22 5 19 ...
```

0.1.3 2.2 Shiny App Development

1 Documentation and Presentation

1.1 App Purpose

1.2 How it Works

1.3 Insights

1.4 Reflections