

Titanic_question

Bonny Runameso

2022-11-19

```
library(readxl)
library(tidyverse)

## -- Attaching packages ----- tidyverse
1.3.1 --

## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.5      v dplyr  1.0.9
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   2.0.1      v forcats 0.5.1

## -- Conflicts -----
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(extrafont)

## Registering fonts with R

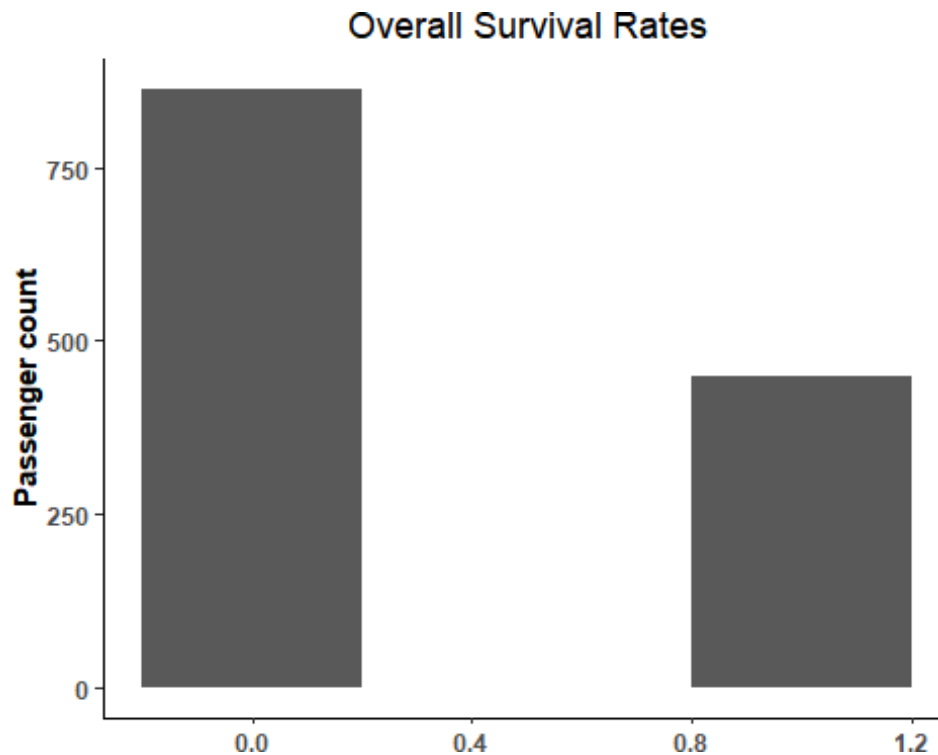
library(extrafontdb)
titanic_data <- read_excel("C:/Users/bonny/Desktop/Data
Visualisation/Titanic.xlsx")
#View(Titanic)
train <- titanic_data
summry <- summary(Titanic)

# 1a)
table(titanic_data$Survived)

##
##    0    1
## 863 450

# Graph
titanic_data %>%
  ggplot(aes(x = Survived)) +
  geom_bar(width = 0.4) +
  theme_classic() +
  theme(
    plot.title = element_text(family = "Times New Roman", hjust = 0.5),
    axis.text = element_text(family = "Times New Roman", face = "bold"),
    axis.title = element_text(family = "Times New Roman", face = "bold")
  )
```

```
) +  
labs(title = "Overall Survival Rates", x = NULL, y = "Passenger count")
```



```
#1b)  
Ccounts <- table(train$Survived, train$PClass)  
#confusion Matrix  
Ccounts  
  
##  
##      1st 2nd 3rd  
##  0 129 160 573  
##  1 193 119 138  
  
counts <- table(train$Survived, train$Sex)  
#confusion Matrix  
counts  
  
##  
##      female male  
##  0      154   709  
##  1      308   142  
  
# survival for male and female  
female <- nrow(train[train$Sex=='female',])  
male <- nrow(train[train$Sex=='male',])  
slice <- c(female, male)  
divi <- c("Female", "Male")
```

```

survival_percent <- c(counts[2]/(counts[1]+counts[2]),
counts[4]/(counts[3]+counts[4]))
# Survival Percent for Female, Male
survival_percent

## [1] 0.6666667 0.1668625

#1c) survival by class of ticket
Ccounts <- table(train$Survived, train$PClass)
#confusion Matrix
Ccounts

##
##      1st 2nd 3rd
##    0 129 160 573
##    1 193 119 138

Class_survival_percent <- c(Ccounts[2]/(Ccounts[1]+Ccounts[2]),
Ccounts[4]/(Ccounts[3]+Ccounts[4]), Ccounts[6]/(Ccounts[5]+Ccounts[6]))
Class_survival_percent

## [1] 0.5993789 0.4265233 0.1940928

#1d)Pclass and gender
Ccounts <- table(train$Survived, train$PClass,train$Sex)
Ccounts

## , , = female
##
##
##      1st 2nd 3rd
##    0   9  13 132
##    1 134  94  80
##
## , , = male
##
##
##      1st 2nd 3rd
##    0 120 147 441
##    1  59  25  58

Class_gendersurvival_percent <- c(Ccounts[2]/(Ccounts[1]+Ccounts[2]),
Ccounts[4]/(Ccounts[3]+Ccounts[4]),
Ccounts[6]/(Ccounts[5]+Ccounts[6]),
Ccounts[8]/(Ccounts[7]+Ccounts[8]),
Ccounts[10]/(Ccounts[9]+Ccounts[10]),
Ccounts[12]/(Ccounts[11]+Ccounts[12])

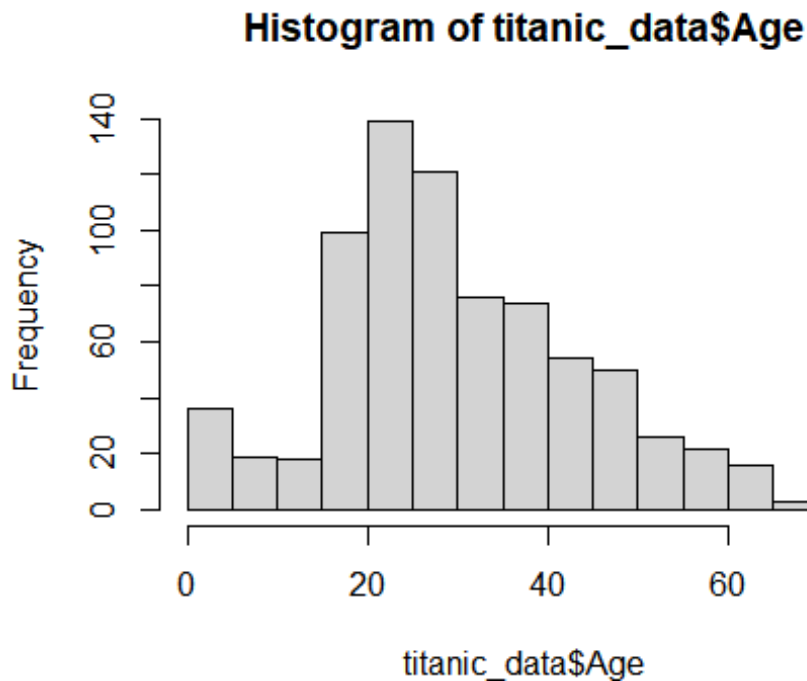
)

Class_gendersurvival_percent

## [1] 0.9370629 0.8785047 0.3773585 0.3296089 0.1453488 0.1162325

```

```
#1e)
hist(titanic_data$Age)
```



#the age histogram graph is relatively normally distributed and skewed to the right.

```
#1f
Ccounts <- table(train$Survived,train$Age)
Ccounts
```

```
##
##      0.17 0.33 0.8 0.83 0.92  1 1.5  2  3  4  5  6  7  8  9 10 11 12 13 14
15 16
##    0    0    1    0    0    0  0    2  3  0  1  1  3  0  0  5  2  2  0  2  3
1  5
##    1    1    0    1    2    1  5    0  4  6  6  2  2  1  4  2  0  1  2  2  1
4  5
##
##      17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
##    0  7 19 11 17 21 22 14 16 16 19 15 17 11 23  7 11  8  8  8 13  5  7  8
##    1  6 11 12  6 10 13  9 11  7  8  9  7  4  8  7 11  6  4  6 16  2  4  6
5  2
##
##      42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64
65 67
```

```
## 0 8 5 5 9 7 6 5 4 5 3 3 0 4 3 2 4 2 2 1 3 1 1 5
2 1
## 1 3 2 3 11 1 3 7 6 6 1 3 3 3 3 2 0 4 1 4 0 1 2 1
0 0
##
## 69 70 71
## 0 0 1 3
## 1 1 0 0
```

code to be done

Class 1 is Classy and others NonClassy

```
CustomerType <- factor(train$PClass,
                        levels = c("1st", "2nd", "3rd"),
                        labels = c("Classy", "Non Classy", "Non Classy"))
```

CustomerType

```
## [1] Classy Classy Classy Classy Classy Classy
## [7] Classy Classy Classy Classy Classy Classy
## [13] Classy Classy Classy Classy Classy Classy
## [19] Classy Classy Classy Classy Classy Classy
## [25] Classy Classy Classy Classy Classy Classy
## [31] Classy Classy Classy Classy Classy Classy
## [37] Classy Classy Classy Classy Classy Classy
## [43] Classy Classy Classy Classy Classy Classy
## [49] Classy Classy Classy Classy Classy Classy
## [55] Classy Classy Classy Classy Classy Classy
## [61] Classy Classy Classy Classy Classy Classy
## [67] Classy Classy Classy Classy Classy Classy
## [73] Classy Classy Classy Classy Classy Classy
## [79] Classy Classy Classy Classy Classy Classy
## [85] Classy Classy Classy Classy Classy Classy
## [91] Classy Classy Classy Classy Classy Classy
## [97] Classy Classy Classy Classy Classy Classy
## [103] Classy Classy Classy Classy Classy Classy
## [109] Classy Classy Classy Classy Classy Classy
## [115] Classy Classy Classy Classy Classy Classy
## [121] Classy Classy Classy Classy Classy Classy
## [127] Classy Classy Classy Classy Classy Classy
## [133] Classy Classy Classy Classy Classy Classy
## [139] Classy Classy Classy Classy Classy Classy
## [145] Classy Classy Classy Classy Classy Classy
## [151] Classy Classy Classy Classy Classy Classy
## [157] Classy Classy Classy Classy Classy Classy
## [163] Classy Classy Classy Classy Classy Classy
## [169] Classy Classy Classy Classy Classy Classy
## [175] Classy Classy Classy Classy Classy Classy
## [181] Classy Classy Classy Classy Classy Classy
## [187] Classy Classy Classy Classy Classy Classy
## [193] Classy Classy Classy Classy Classy Classy
## [199] Classy Classy Classy Classy Classy Classy
```

[illegible]

[illegible]

[illegible]

[illegible]