Assignment 11: Surviving the Titanic

MinJae Jo

2025-11-05

Exercise 1

```r
library(readr)
library(dplyr)

train_df <- read_csv(
  file = "train.csv",
  col_types = cols(
    Pclass = col_character(),
    SibSp = col_character(),
    Parch = col_character()
  )
)

train_df <- train_df %>%
  mutate(did_survive = as.logical(Survived))
```
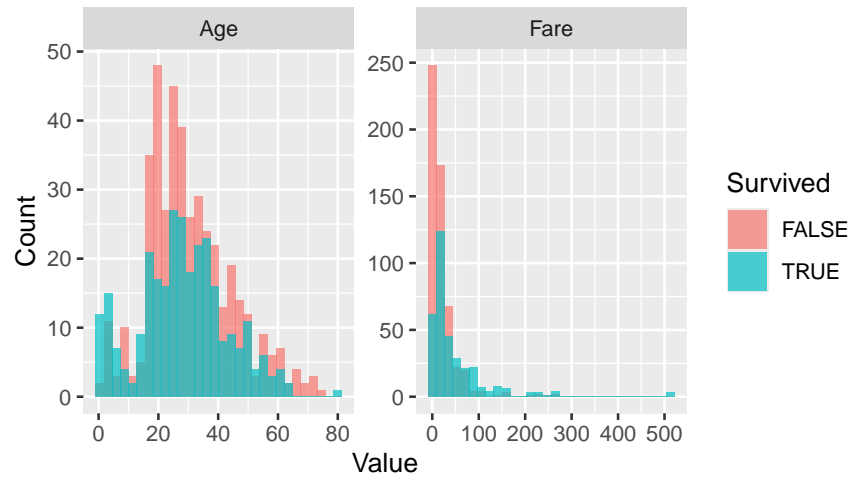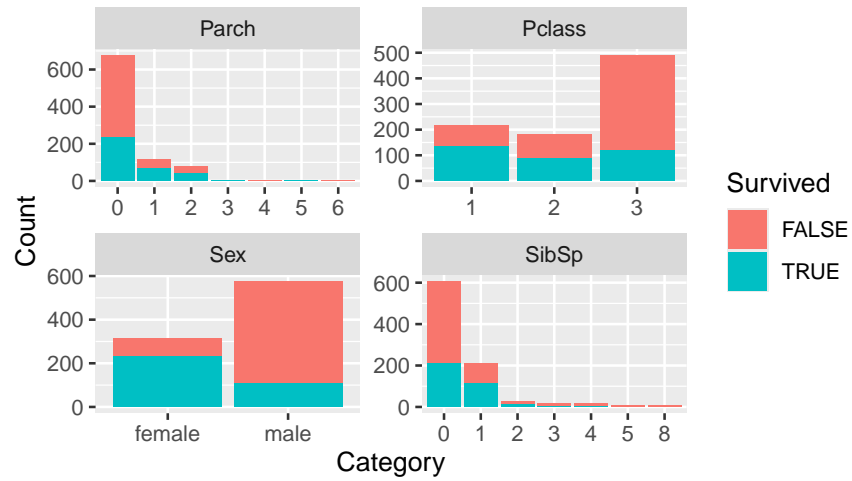
Exercise 2

```
## Warning: Removed 177 rows containing non-finite outside the scale range
## ('stat_bin()').
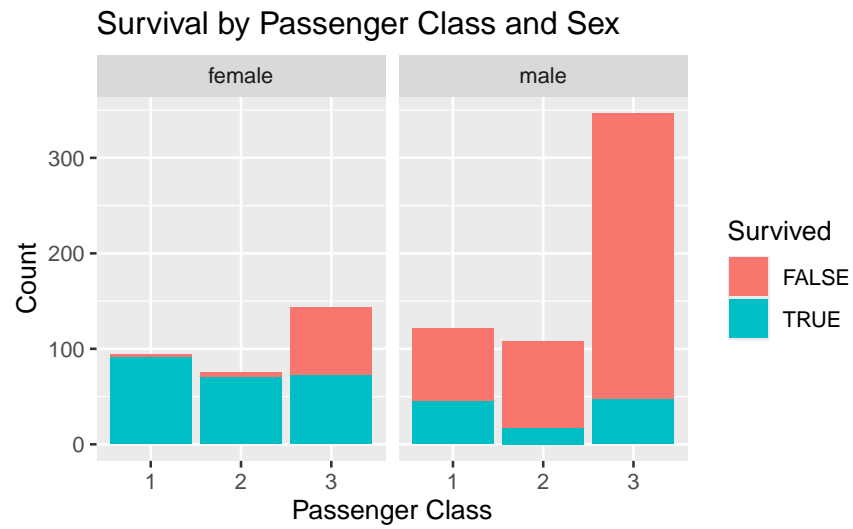```

## Distributions of Age and Fare by Survival Status



**Exercise 3**

## Categorical Variables and Survival on the Titanic

## Exercise 4

**Survival by Passenger Class and Sex**



## Exercise 5

| count | missing | fraction_missing |
|-------|---------|------------------|
| 891 | 177 | 0.1986532 |

| missing_age_imputed |
|---------------------|
| 0 |

## Exercise 6

| total_correct | accuracy |
|---------------|----------|
| 549 | 0.6161616 |

## Exercise 7

```
## [1] 0.3838384 0.3838384
```

Exercise 8

```
## [1] 0.1997755 0.2080828
```

```
## [1] 0.1874299 0.1854308
```

Bonus Exercise

If you choose to do the bonus exercise, write your code here. Otherwise you can delete this section.

Academic Integrity statement