

Lab 04 - The RMS Titanic

Team gitData - Nathan Kim, Avanti Shah, Blossom Mojekwu

Due Thursday, Feb 13 at 11:59p

```
library(tidyverse)
titanic <- read_csv("data/titanicsurvival.csv")
typeof("Class")
```

```
## [1] "character"
```

```
typeof("Age")
```

```
## [1] "character"
```

```
typeof("Sex")
```

```
## [1] "character"
```

```
typeof("Survive")
```

```
## [1] "character"
```

Exercise 1

The titanic survival file includes 4 observations, which are all of type “character”. These include class, age, sex, and whether they survived or not. Class indicates which type of passenger they are, being a first to third-class passenger or crewmember. Age indicates if the passenger is an adult or child. Sex indicates the sex of the passenger. Survive is either a yes or no, depending on if they survived the cruise crash.

Exercise 2

```
titanic %>%
  group_by(Class) %>%
  count(Survive) %>%
  mutate(total_count = sum(n)) %>%
  mutate(prop_survived = n / total_count) %>%
  filter(Survive == "Yes") %>%
  select(Class, Survive, prop_survived)
```

```
## # A tibble: 4 x 3
## # Groups:   Class [4]
##   Class Survive prop_survived
##   <chr> <chr>          <dbl>
## 1 Crew  Yes             0.240
## 2 First Yes             0.625
```

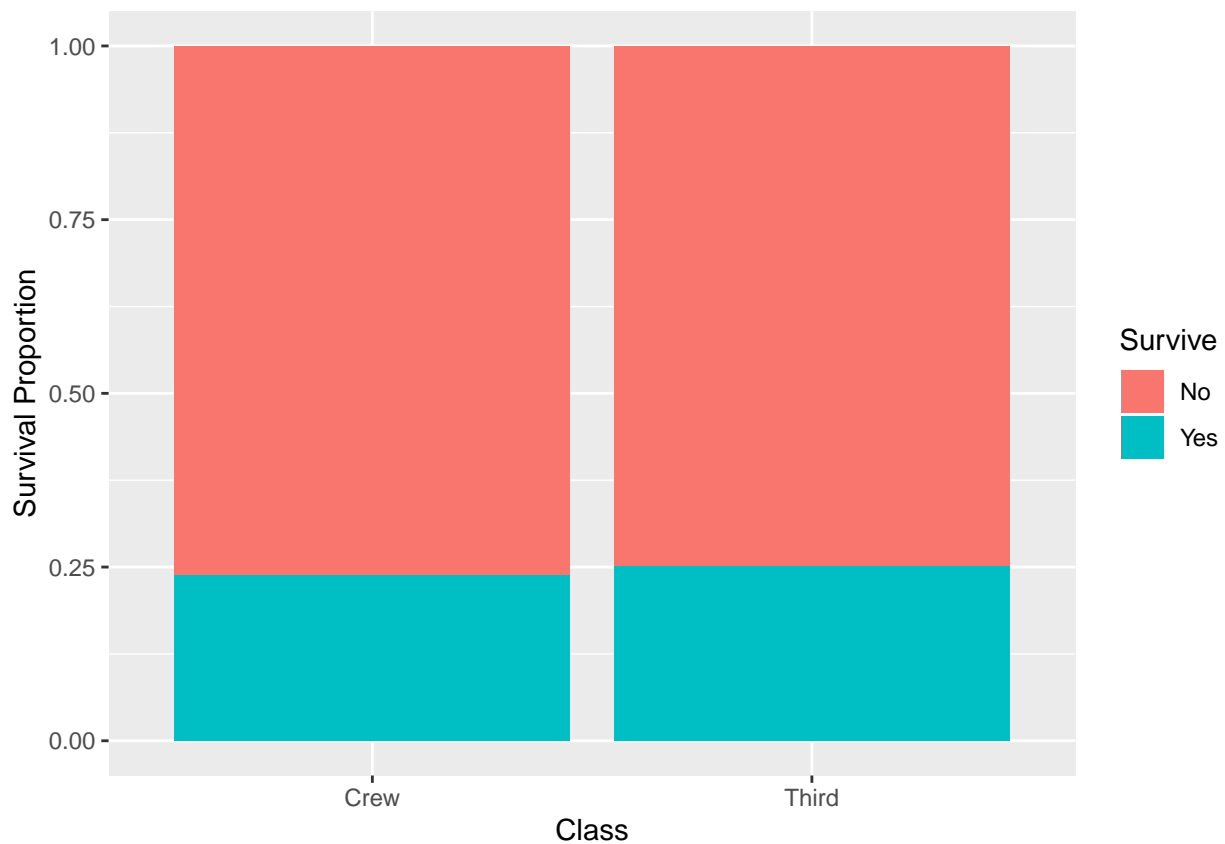
```
## 3 Second Yes      0.414
## 4 Third  Yes      0.252
```

Crew members had the lowest conditional probability of survival at around 23.95%. First class members had a 62.46% rate of survival on the Titanic, Second class had a 41.40% rate of survival and Third class had 25.21%.

Exercise 3

```
third_crew <- titanic %>%
  filter(Class == "Third" | Class == "Crew")

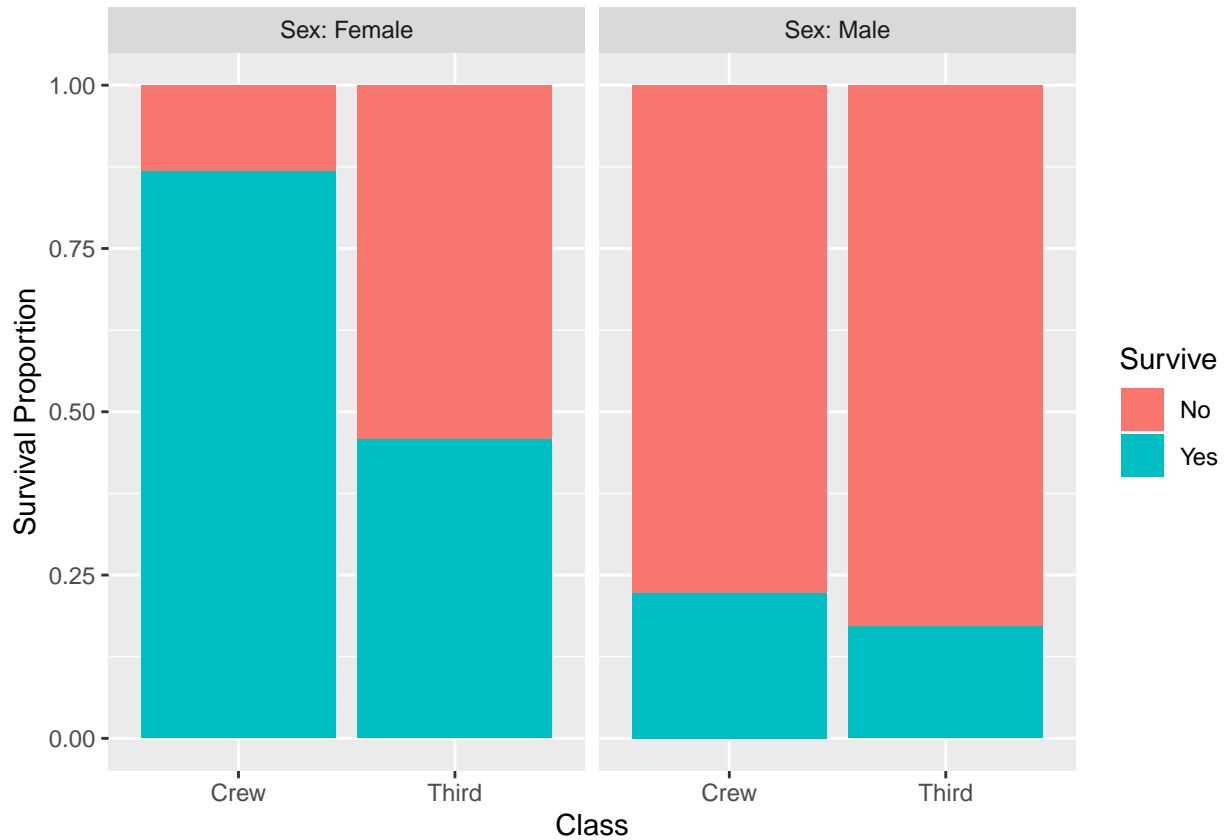
ggplot(data=third_crew, aes(x = Class, fill = Survive)) +
  geom_bar(position = "fill") +
  labs(title = "Survival Rate of Third-Class Passengers Compared to
  Crew Members", x = "Class", y = "Survival Proportion")
```



Exercise 4

```
ggplot(data = third_crew, aes(x = Class, fill = Survive)) +
  geom_bar(position = "fill") +
  facet_grid(~ Sex, labeller = label_both) +
```

```
labs(Title = "Survival Probability of Third-Class Passengers Compared to  
Crew Members", x = "Class", y = "Survival Proportion")
```



Exercise 5

When the bar graph is faceted with consideration for Sex, it is seen that crew members had a higher survival rate than third-class passengers for both genders. Female passengers had a higher proportion of survival as both crew members and third-class passengers. You can conclude that female crew members had the highest survival rate than any other group of passengers.