Lab Name

Code

Author Name

2021-04-11

Hide

```
library(tidyverse)
```

```
## Warning: package 'tibble' was built under R version 4.0.4
```

```
## Warning: package 'tidyr' was built under R version 4.0.4
```

```
## Warning: package 'dplyr' was built under R version 4.0.4
```

```
## Warning: package 'forcats' was built under R version 4.0.4
```

Hide

```
library(titanic)
```

```
## Warning: package 'titanic' was built under R version 4.0.4
```

Part 1

Getting the titanic Data

Hide

```
titanic <- titanic_train %>%
  select(Survived, Pclass, Sex, Age, SibSp, Parch, Fare, Embarked,
         Cabin, Ticket, Name) %>%
  mutate(Survived = factor(Survived),
         Pclass = factor(Pclass),
         Sex = factor(Sex))
str(titanic)
```

```
## 'data.frame':    891 obs. of  11 variables:
##  $ Survived: Factor w/ 2 levels "0","1": 1 2 2 2 1 1 1 1 2 2 ...
##  $ Pclass  : Factor w/ 3 levels "1","2","3": 3 1 3 1 3 3 1 3 3 2 ...
##  $ Sex     : Factor w/ 2 levels "female","male": 2 1 1 1 2 2 2 2 1 1 ...
##  $ Age     : num  22 38 26 35 35 NA 54 2 27 14 ...
##  $ SibSp   : int  1 1 0 1 0 0 0 3 0 1 ...
##  $ Parch   : int  0 0 0 0 0 0 0 1 2 0 ...
##  $ Fare    : num  7.25 71.28 7.92 53.1 8.05 ...
##  $ Embarked: chr  "S" "C" "S" "S" ...
##  $ Cabin   : chr  "" "C85" "" "C123" ...
##  $ Ticket  : chr  "A/5 21171" "PC 17599" "STON/O2. 3101282" "113803" ...
##  $ Name    : chr  "Braund, Mr. Owen Harris" "Cumings, Mrs. John Bradley (Florence Brigg
s Thayer)" "Heikkinen, Miss. Laina" "Futrelle, Mrs. Jacques Heath (Lily May Peel)" ...
```

Hide

```
head(titanic)
```

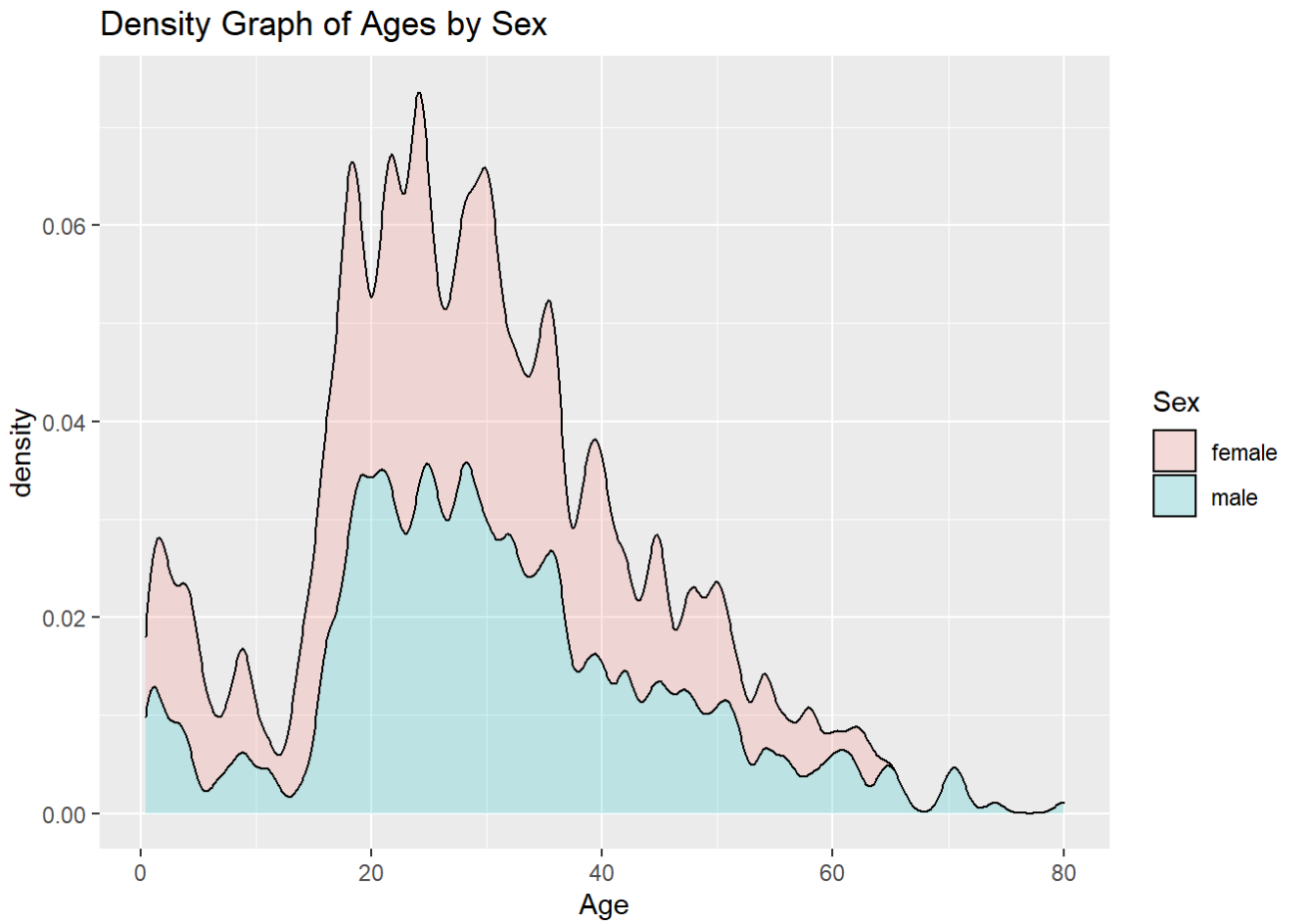
```
##   Survived Pclass    Sex Age SibSp Parch    Fare Embarked Cabin
## 1        0      3   male  22     1     0  7.2500        S
## 2        1      1 female  38     1     0 71.2833        C   C85
## 3        1      3 female  26     0     0  7.9250        S
## 4        1      1 female  35     1     0 53.1000        S  C123
## 5        0      3   male  35     0     0  8.0500        S
## 6        0      3   male  NA     0     0  8.4583        Q
##             Ticket                                               Name
## 1        A/5 21171                             Braund, Mr. Owen Harris
## 2         PC 17599 Cumings, Mrs. John Bradley (Florence Briggs Thayer)
## 3 STON/O2. 3101282                              Heikkinen, Miss. Laina
## 4           113803        Futrelle, Mrs. Jacques Heath (Lily May Peel)
## 5           373450                            Allen, Mr. William Henry
## 6           330877                                    Moran, Mr. James
```

Density Graphs

Hide

```
#density graph of ages by sex
titanic %>%
  group_by(Sex) %>%
  ggplot(aes(Age, fill=Sex)) +
  geom_density(alpha=.2, bw=.8, position="stack") +
  labs(title = 'Density Graph of Ages by Sex') +
  theme_get()
```

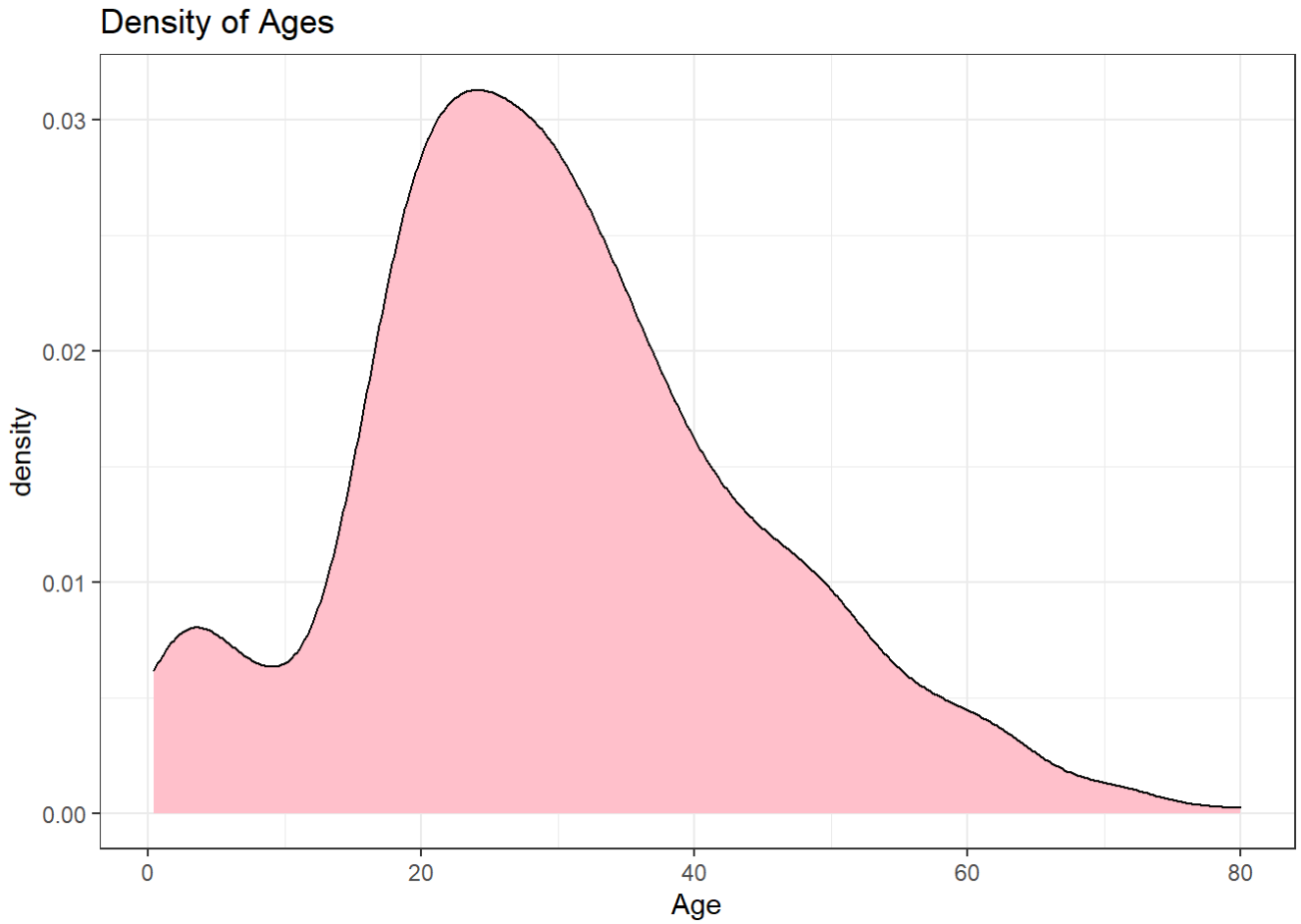
```
## Warning: Removed 177 rows containing non-finite values (stat_density).
```



Hide

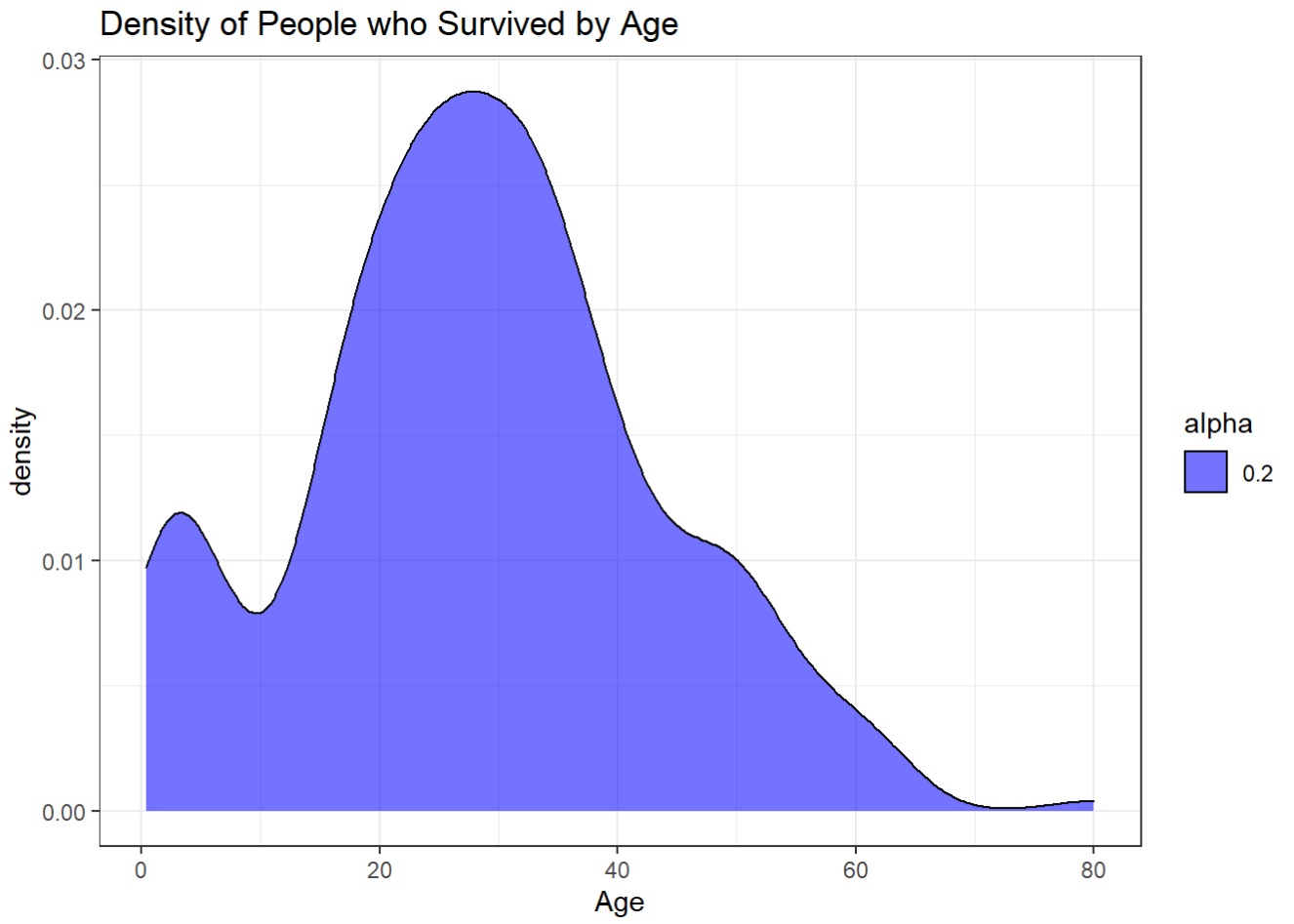
```
#density of ages
titanic %>%
  ggplot(aes(Age)) +
  geom_density(fill = 'pink') +
  labs(title = "Density of Ages") +
  theme_bw()
```

```
## Warning: Removed 177 rows containing non-finite values (stat_density).
```



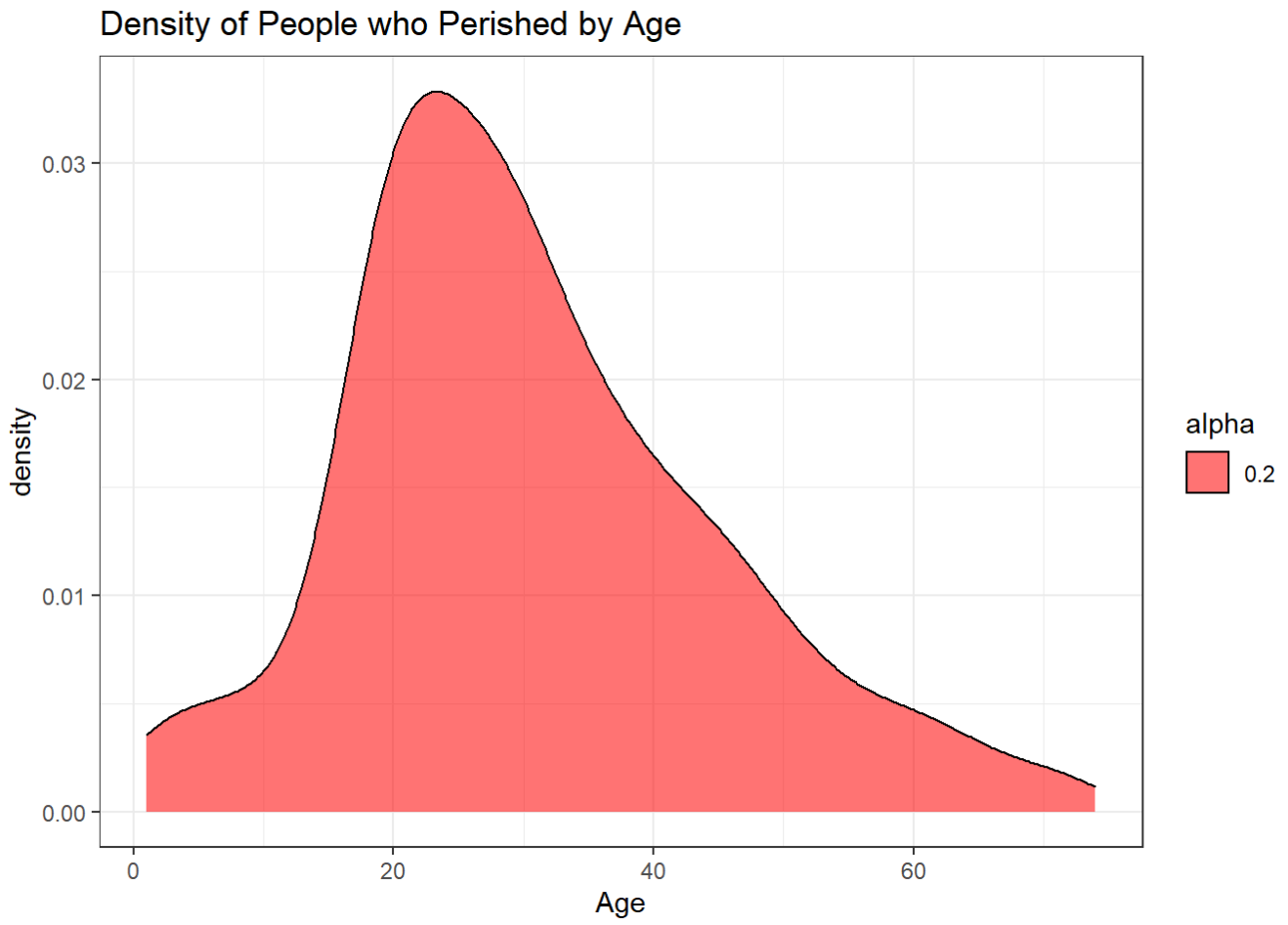
Hide

```
#density plot of people who survived/perished by age
titanic %>%
  filter(Survived==1, !is.na(Age)) %>%
  ggplot(aes(Age, alpha=.2)) +
  geom_density(fill="blue") +
  labs(title = 'Density of People who Survived by Age') +
  theme_bw()
```



Hide

```
titanic %>%
  filter(Survived==0, !is.na(Age)) %>%
  ggplot(aes(Age, alpha=.2)) +
  geom_density(fill="red") +
  labs(title = 'Density of People who Perished by Age') +
  theme_bw()
```



Hide

```
#density of people who perished based on age sex and class type
titanic %>%
  filter(Survived==0) %>%
  ggplot(aes(Age)) +
  geom_density(alpha=.2, legend=TRUE, fill = 'red') +
  facet_grid(Sex~Pclass) +
  labs(title = 'Density of People Who Perished Based on Age, Sex, and Class Type') +
  theme_bw()
```

```
## Warning: Ignoring unknown parameters: legend
```

```
## Warning: Removed 125 rows containing non-finite values (stat_density).
```

Hide

```
#Density plot of people who survived based on age & sex in each class
titanic %>%
 filter(Survived==1) %>%
  ggplot(aes(Age)) +
  geom_density(alpha=.2, legend=TRUE, fill = 'blue') +
  facet_grid(Sex~Pclass) +
  labs(title = 'Density of People Who Survived Based on Age, Sex, and Class Type') +
  theme_bw()
```

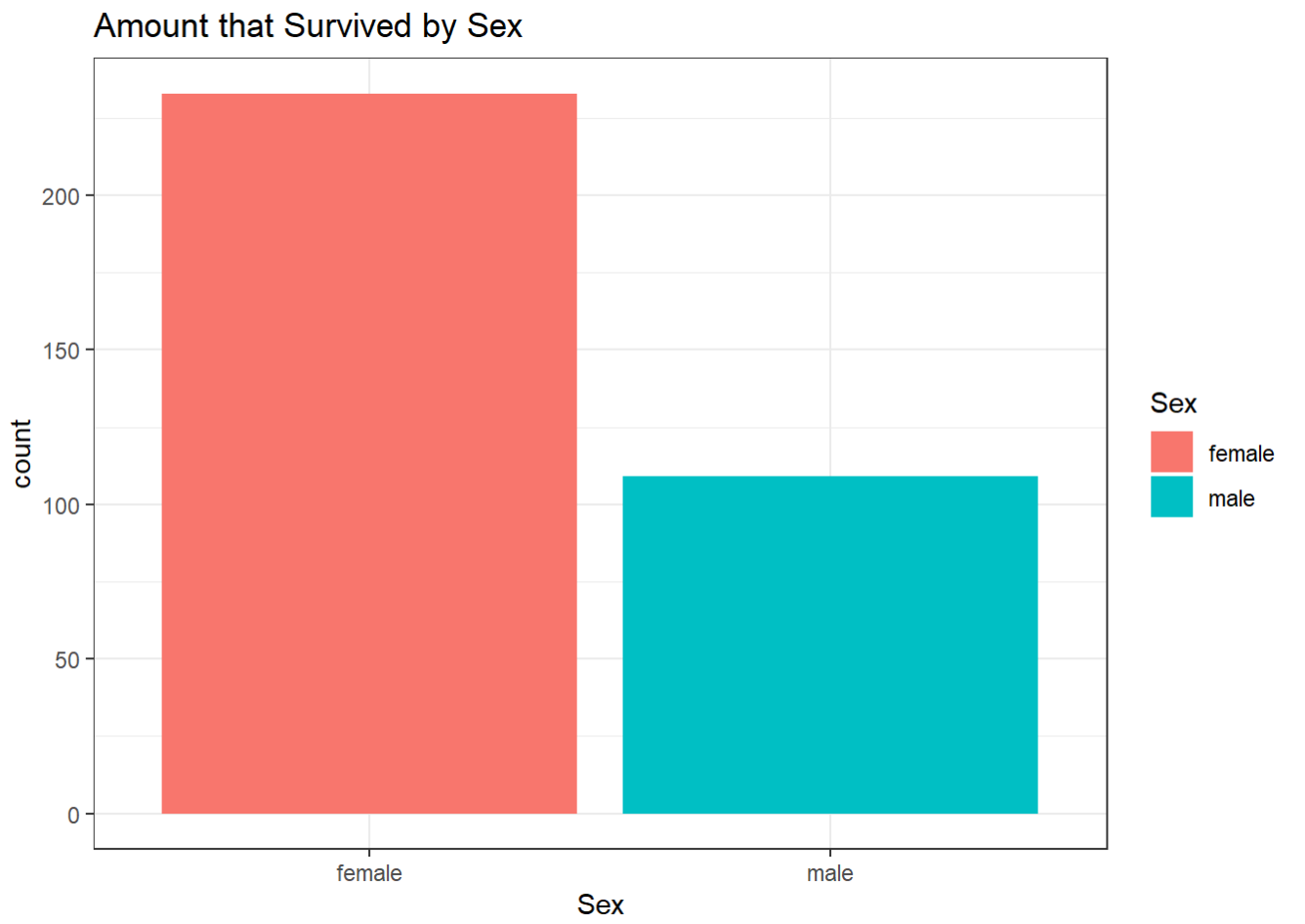
```
## Warning: Ignoring unknown parameters: legend
```

```
## Warning: Removed 52 rows containing non-finite values (stat_density).
```

Barplots

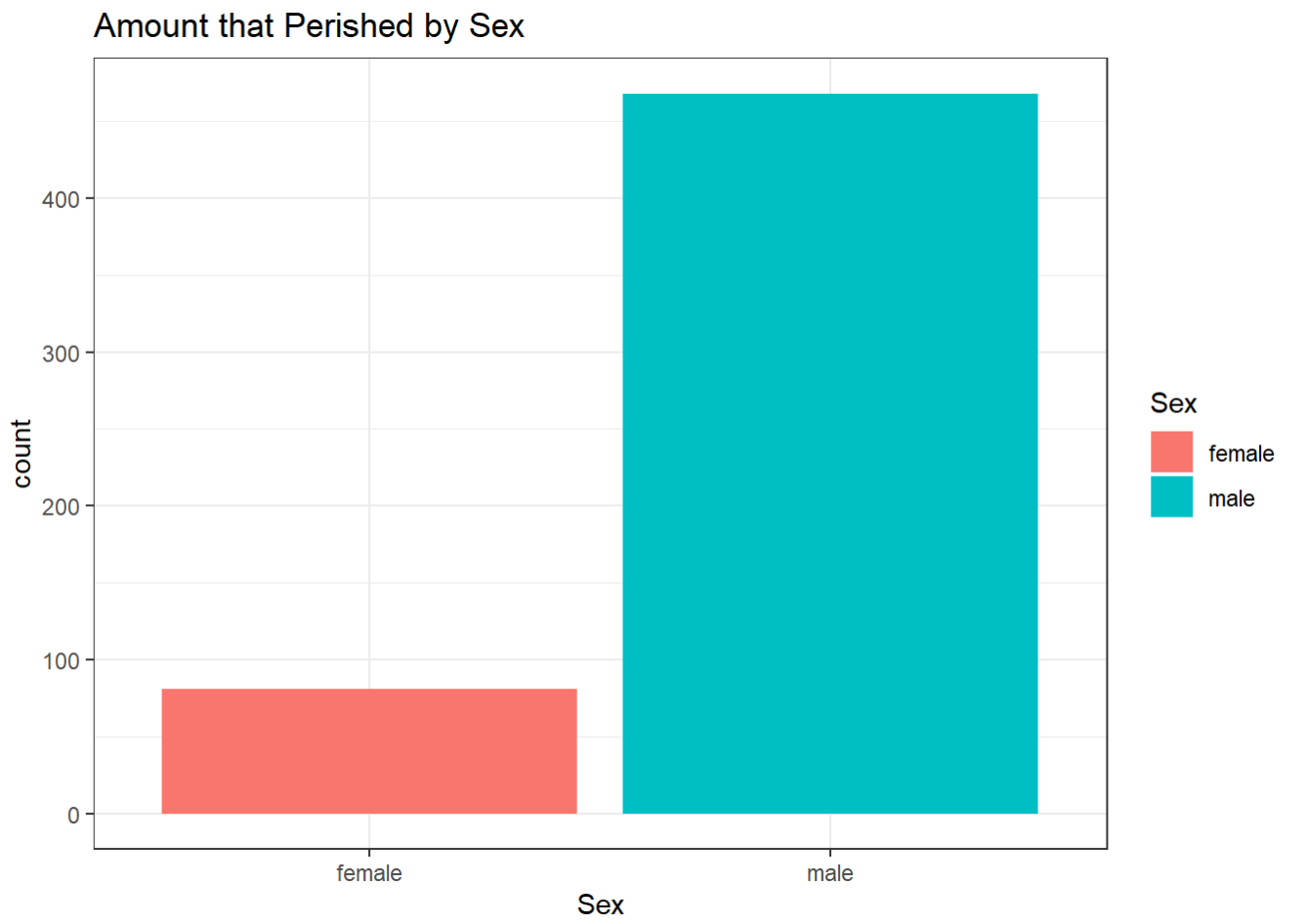
Hide

```
#barplot of amount of peolpe that survived/perished by sex
titanic %>%
  filter(Survived==1)%>%
  ggplot(aes(Sex)) +
  geom_bar(aes(fill= Sex)) +
  labs(title = "Amount that Survived by Sex") +
  theme_bw()
```



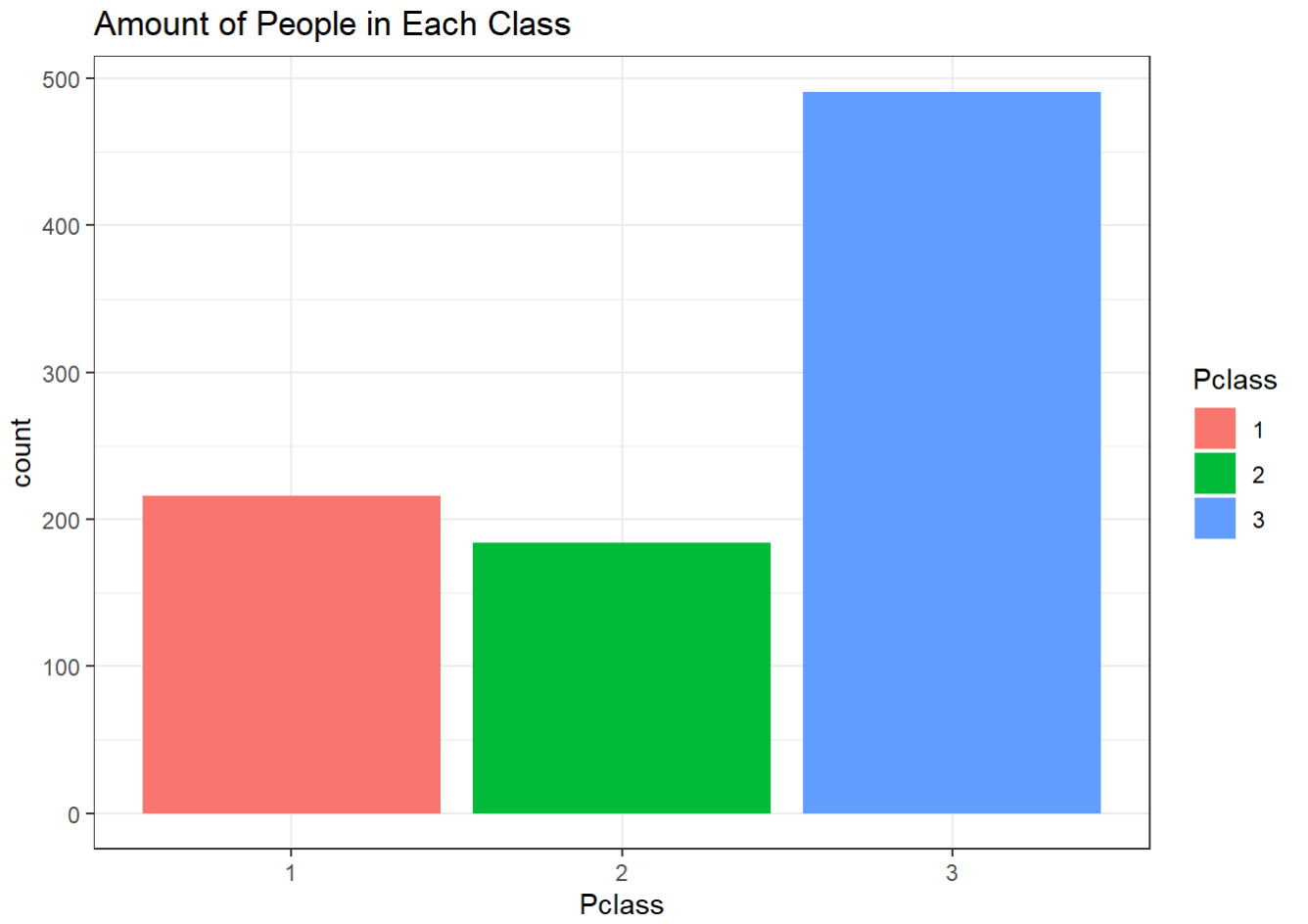
Hide

```
titanic %>%
  filter(Survived==0) %>%
  ggplot(aes(Sex)) +
  geom_bar(aes(fill = Sex)) +
  labs(title = 'Amount that Perished by Sex') +
  theme_bw()
```



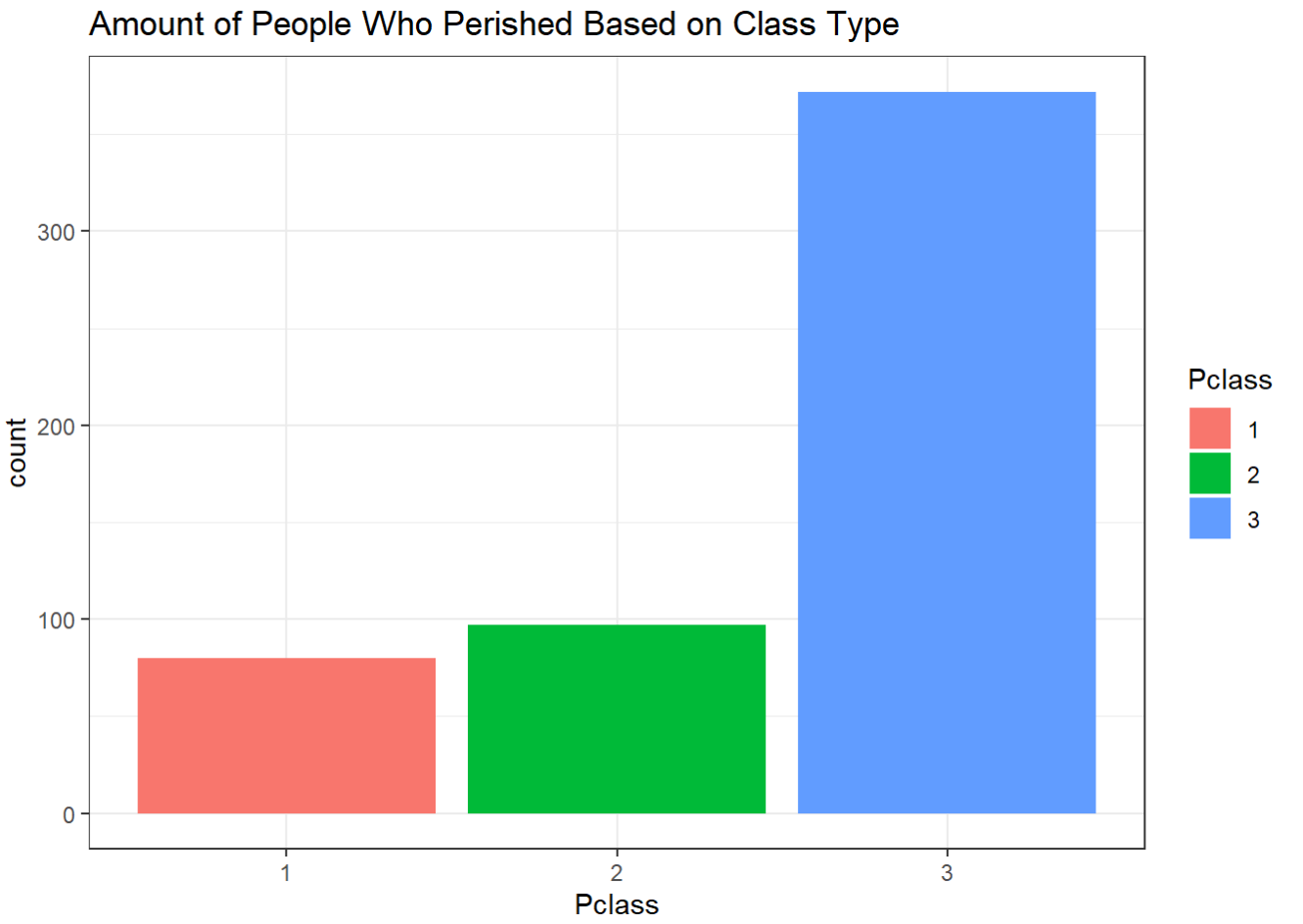
Hide

```
#barplot of amount of people in each class
titanic %>%
  ggplot(aes(Pclass)) +
  geom_bar(aes(fill = Pclass)) +
  labs(title = 'Amount of People in Each Class') +
  theme_bw()
```



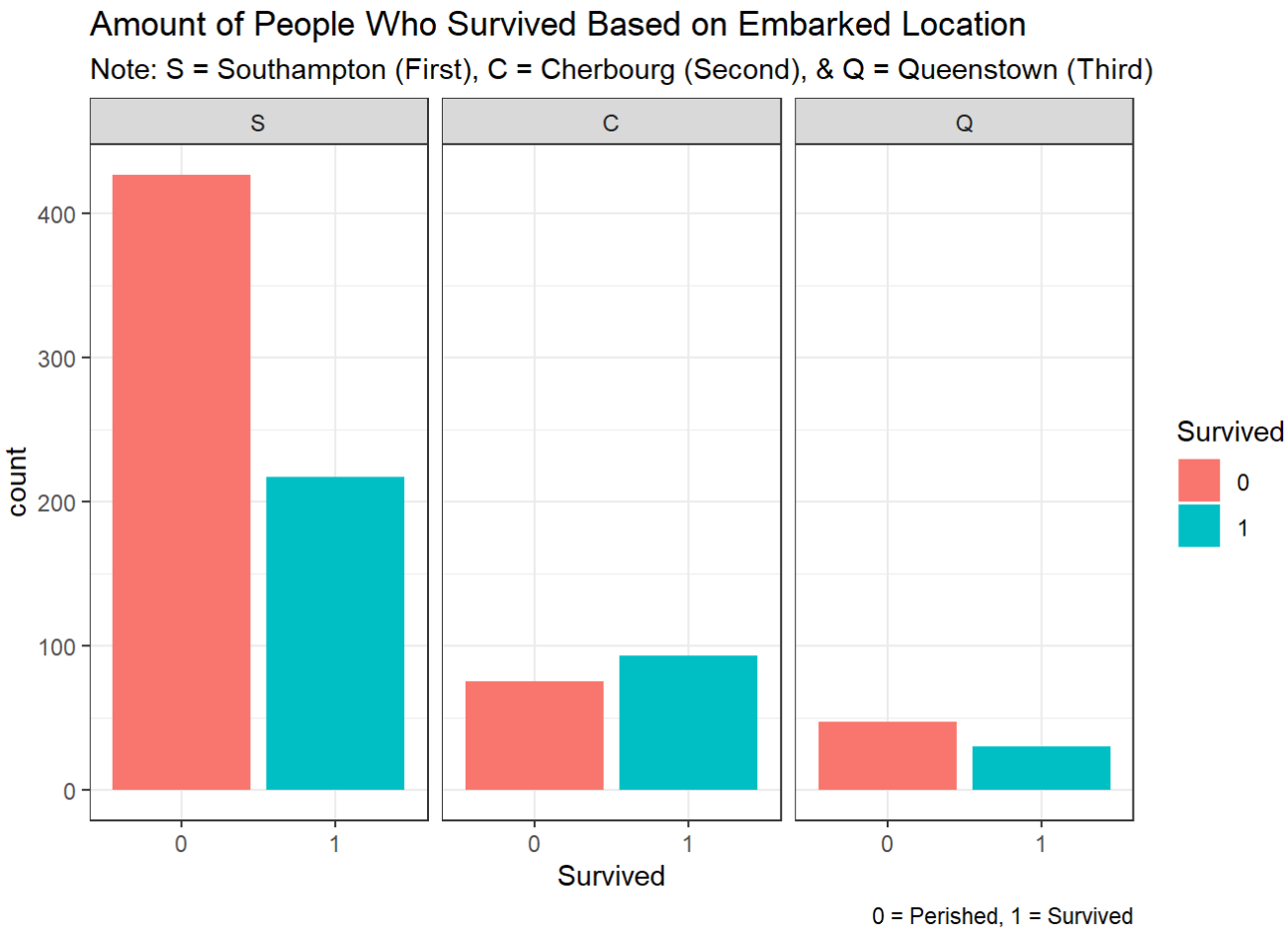
Hide

```
#barplot of amount of people who perished based on class
titanic %>%
  filter(Survived==0) %>%
  ggplot(aes(Pclass)) +
  geom_bar(aes(fill = Pclass)) +
  labs(title = 'Amount of People Who Perished Based on Class Type') +
  theme_bw()
```



Hide

```
#barplot of survival rate based on embarked
titanic %>%
  filter(Embarked != "") %>%
  ggplot(aes(Survived)) +
  geom_bar(aes(fill = Survived)) +
  facet_grid(.~ factor(Embarked, levels = c('S', 'C', 'Q'))) +
  labs( title = 'Amount of People Who Survived Based on Embarked Location',
        subtitle = 'Note: S = Southampton (First), C = Cherbourg (Second), & Q = Queenstow
         n (Third)',
        caption = '0 = Perished, 1 = Survived') +
  theme_bw()
```



Boxplots

Hide

```
#boxplot of amount of people who survived base on their fare
titanic %>%
  filter(Fare != 0) %>%
  ggplot(aes(Survived, Fare, color= Survived))+
  geom_boxplot() +
  geom_jitter(width = .2,alpha=.2) +
  scale_y_continuous(trans = "log2") +
  labs(title = 'Amount of People that Survived based on Their Fare Price',
       caption = '0 = Perished and 1 = Survived') +
  theme_bw()
```

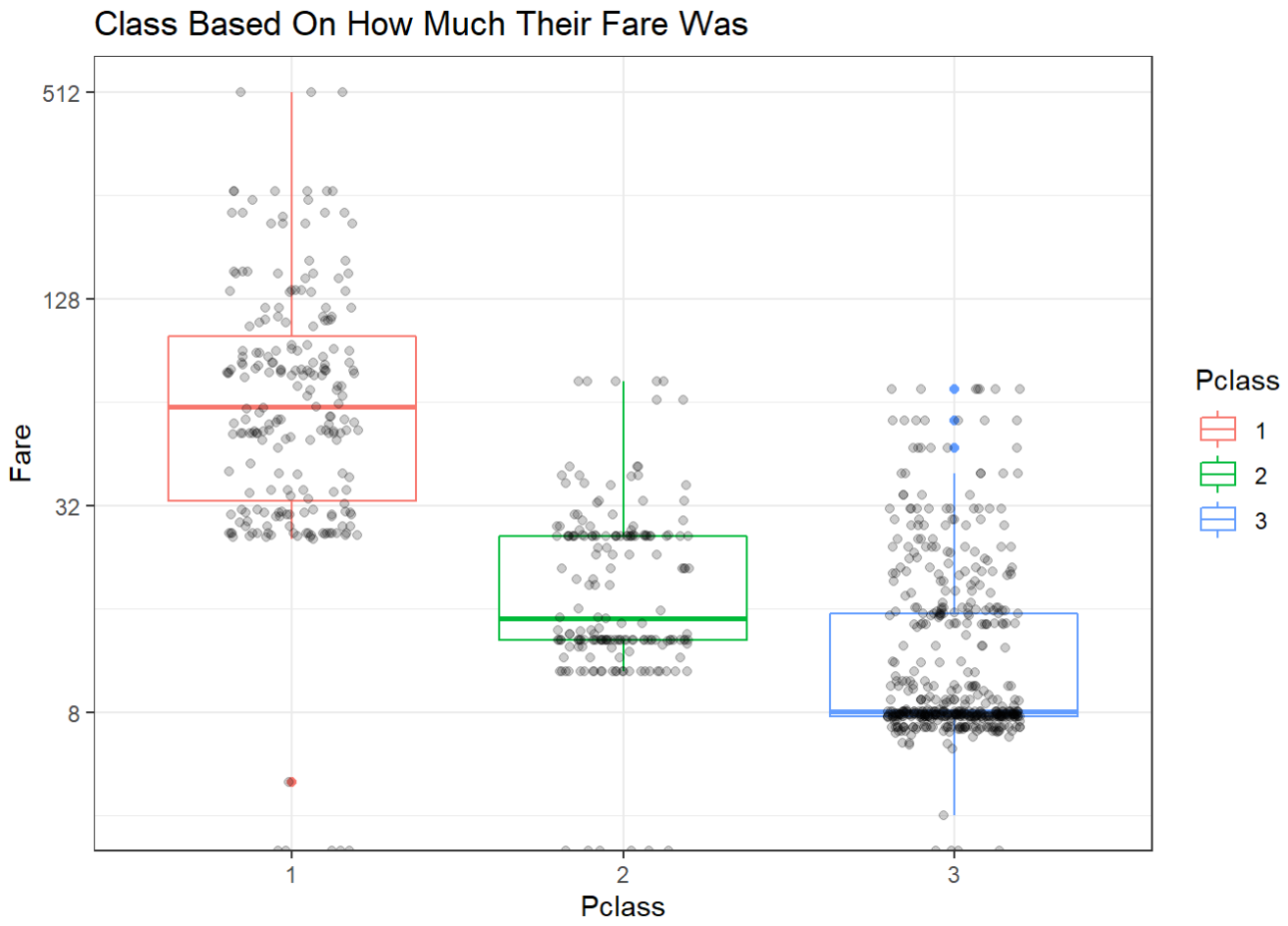
Hide

```
#boxplot of each class based on fare
titanic %>%
  ggplot(aes(x= Pclass, y= Fare)) +
  geom_boxplot(aes(color= Pclass)) +
  geom_jitter(width = .2, alpha = .2) +
  scale_y_continuous(trans = 'log2') +
  labs(title = "Class Based On How Much Their Fare Was") +
  theme_bw()
```

```
## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous y-axis
```

```
## Warning: Removed 15 rows containing non-finite values (stat_boxplot).
```


Class Based On How Much Their Fare Was
512
128
32
8
Fare
1
2
3
Pclass
Pclass
1
2
3