Visualizing Data

75694189 16257626 13341225

Chen Wang, Junke Wang, Zhuozhi Xiong

due:Oct 9, 2019 at the end of discussion

```
library(tidyverse)
library(fivethirtyeight)
library(titanic)
```

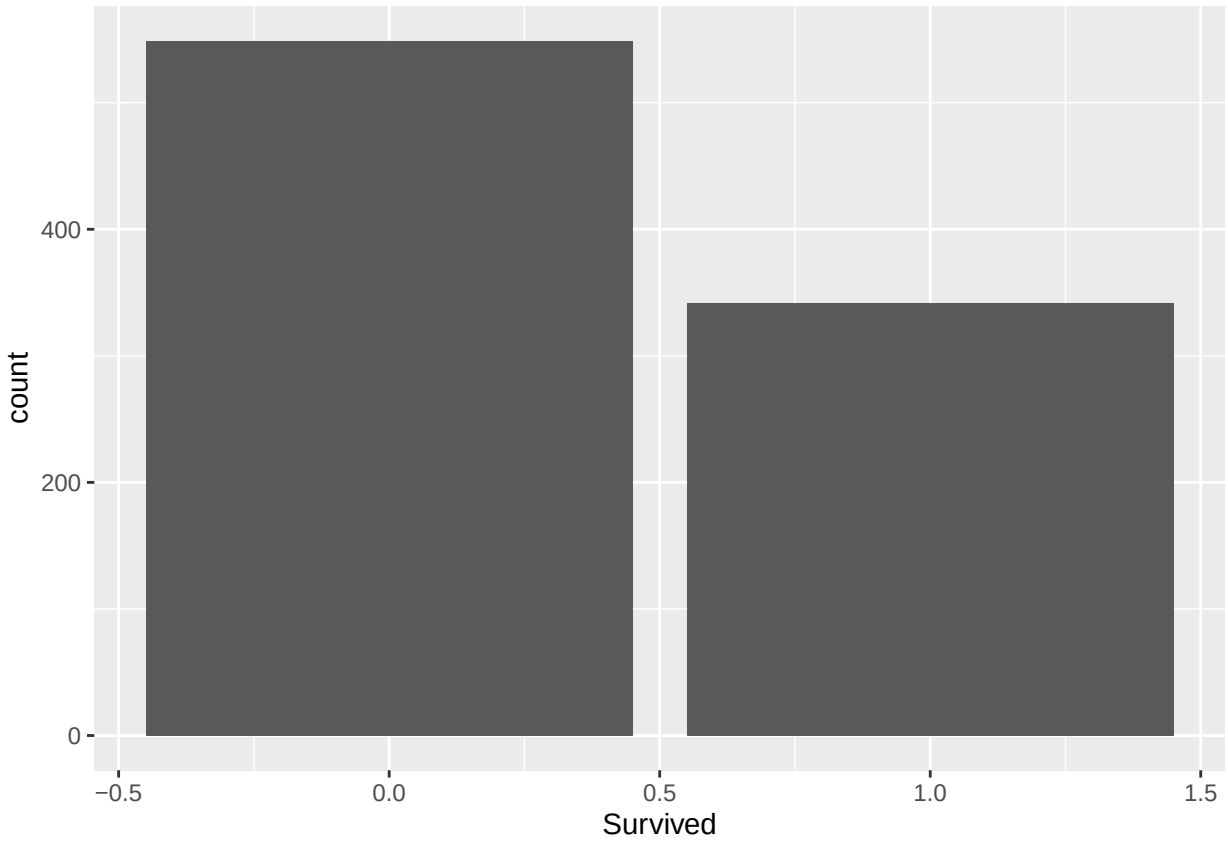
Slide 9 Plot

Text:

1. Survived
2. Survived is on the x-axis
3. Bar plot

Code:

```
titanic_train %>%
  ggplot(aes(x = Survived)) +
  geom_bar()
```



Slide 10 Plot

Text:

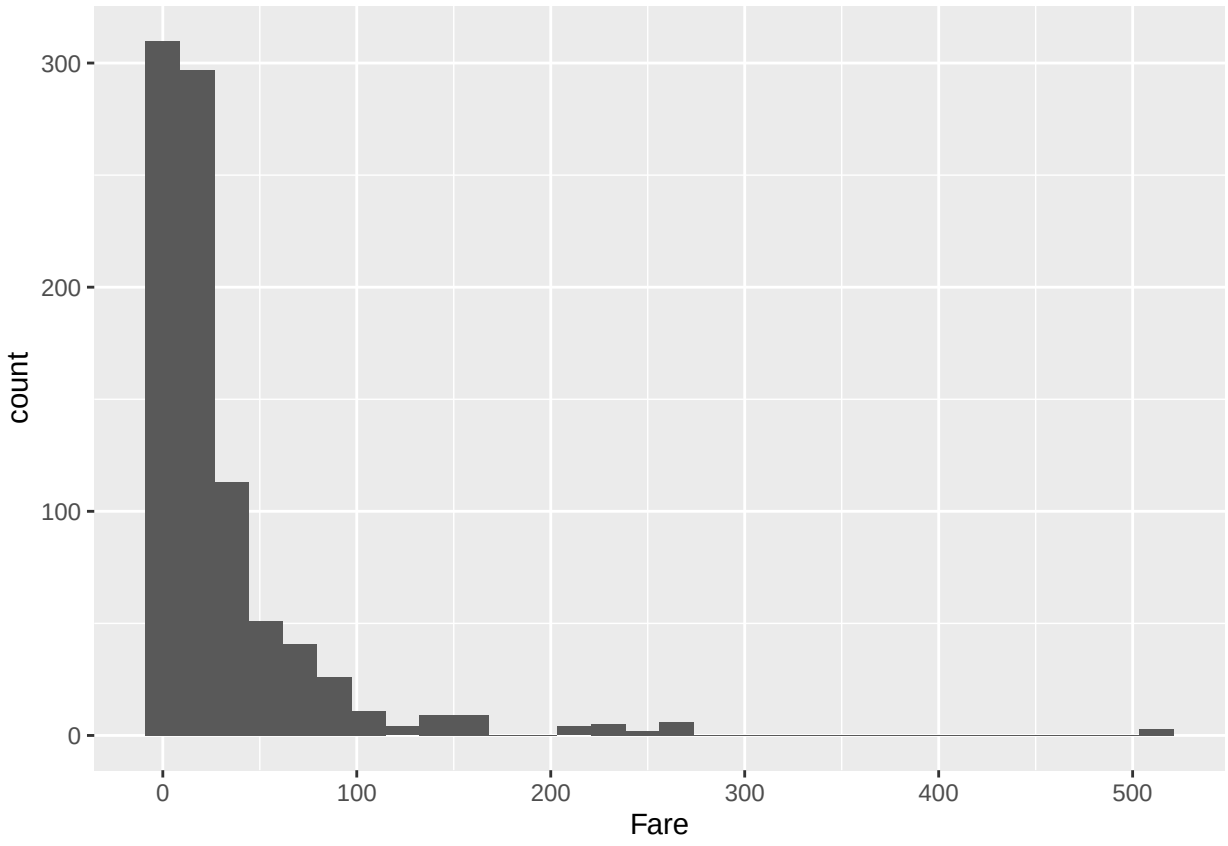
1. Fare
2. Fare is on the x-axis
3. Histogram

Code:

```
titanic_train %>%

  ggplot(aes(x=Fare)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Slide 11 plot

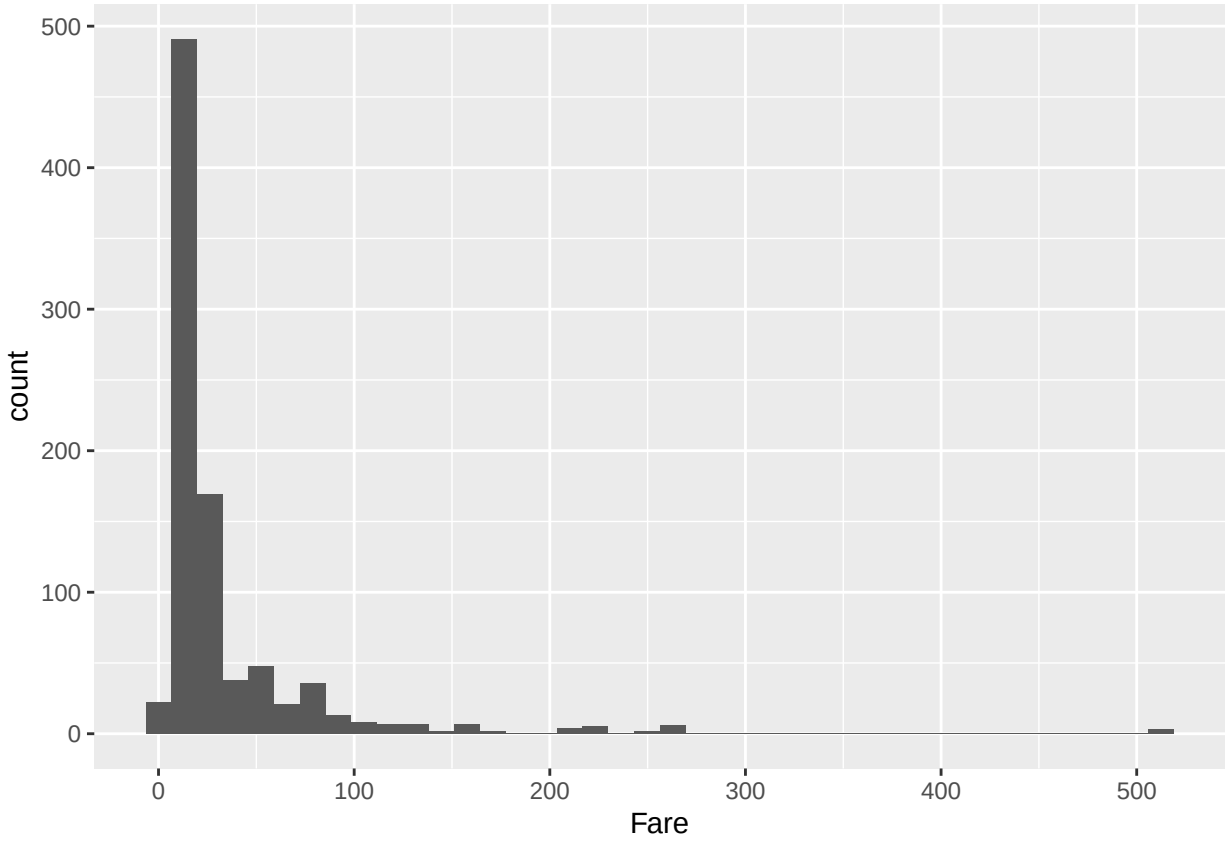
Text:

1. Fare
2. Fare is on the x-axis
3. Histogram

Code:

```
titanic_train %>%

  ggplot(aes(x=Fare)) +
  geom_histogram(bins = 40)
```

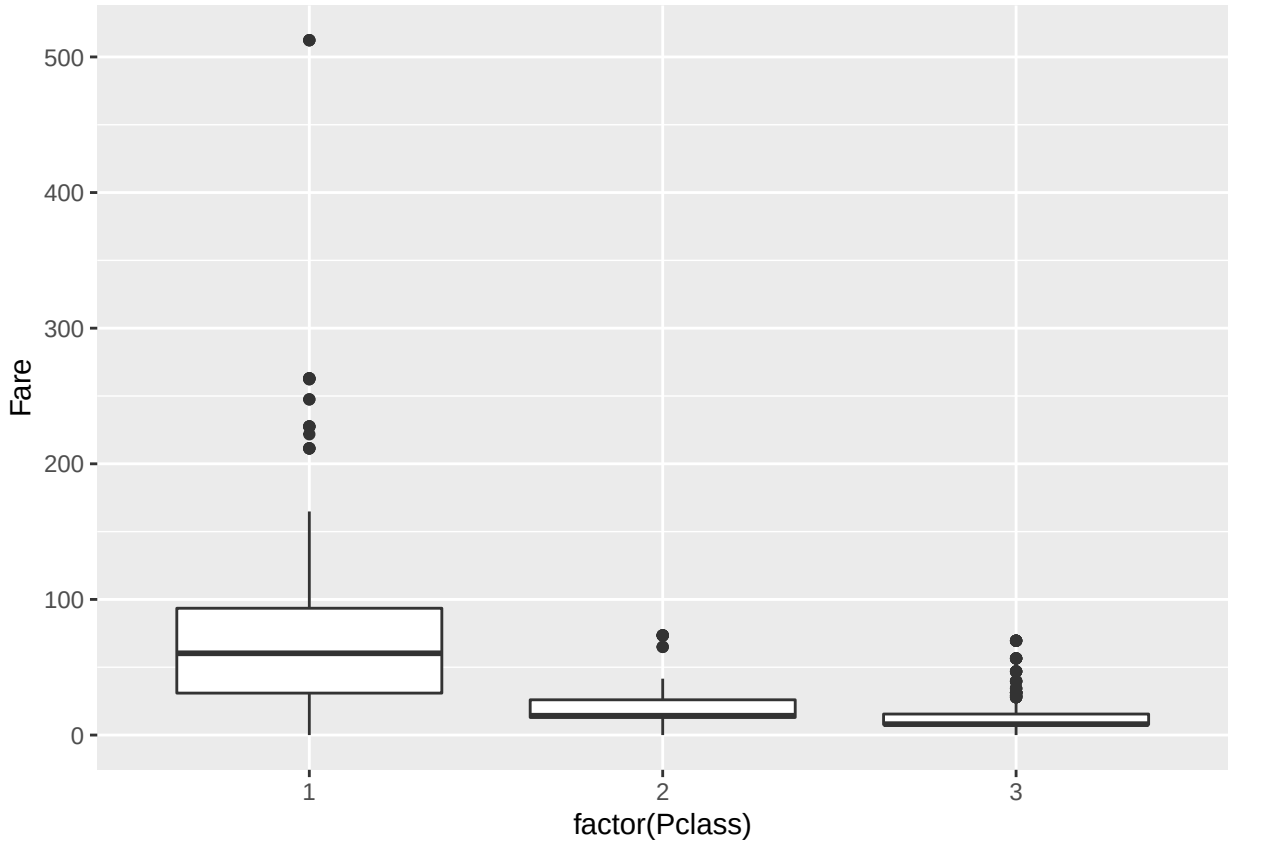


Slide 13 plot

Text:

1. Pclass and Fare
2. Pclass on x-axis, Fare on y-axis
3. Boxplot

```
titanic_train %>%
  ggplot(aes(x=factor(Pclass), y= Fare)) +
  geom_boxplot()
```



Slide 15 plot

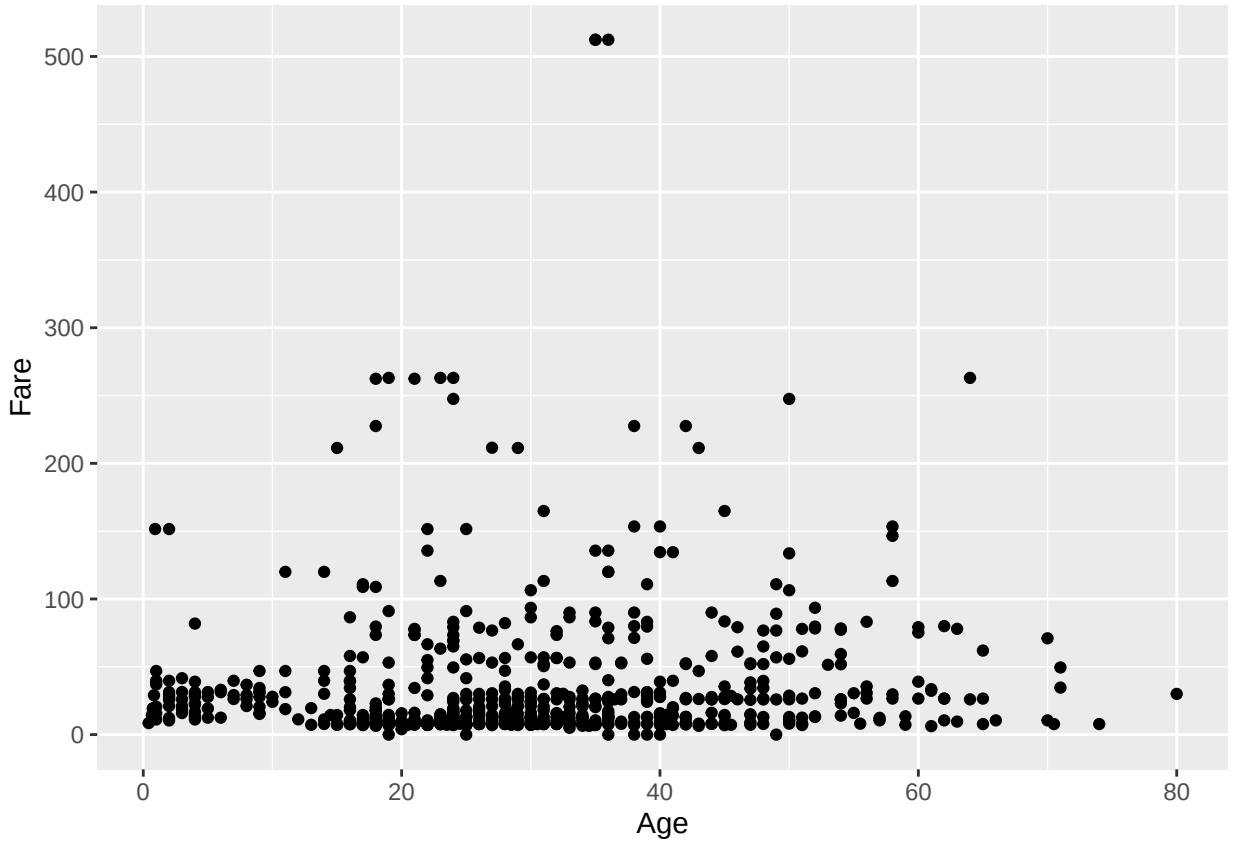
Text:

1. Age and Fare
2. Age is on x-axis and Fare is on y-axis
3. Scatter-plot

Code:

```
titanic_train %>%
  ggplot(aes(x=Age, y= Fare)) +
  geom_point()
```

```
## Warning: Removed 177 rows containing missing values (geom_point).
```



No real relationship between Age and Fare

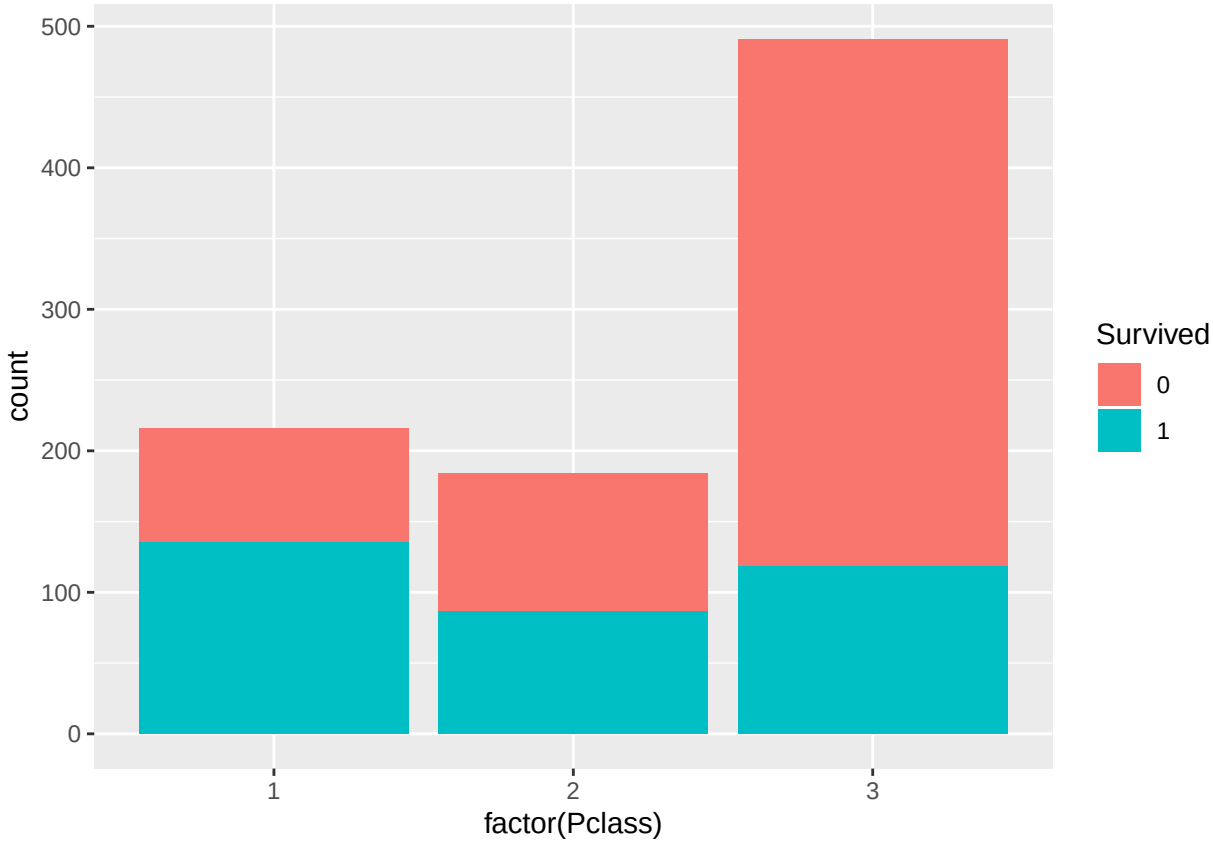
Slide 21 plot

Text:

1. Pclass, Survived
2. Pclass on the x-axis, Survived stacked
3. Stacked bar plot

Code:

```
titanic_train %>%
  ggplot(aes(x=factor(Pclass), fill=Survived)) +
  geom_bar(aes(fill=factor(Survived)))
```



Slide 22 plot

Text:

1. domgross, binary(FAIL/PASS)
2. domgross on the x-axis, binary stacked
3. Stacked histogram

Code:

```
bechdel %>%
  ggplot(aes(x=domgross, fill = binary))+
  geom_histogram(aes(fill=binary),bins = 80) +
  scale_fill_manual(values = c("blue", "magenta"))
```

```
## Warning: Removed 17 rows containing non-finite values (stat_bin).
```