Lecture Assignment 4

Viraj Vijaywargiya

2022-04-11

```
library(tidyverse)
```

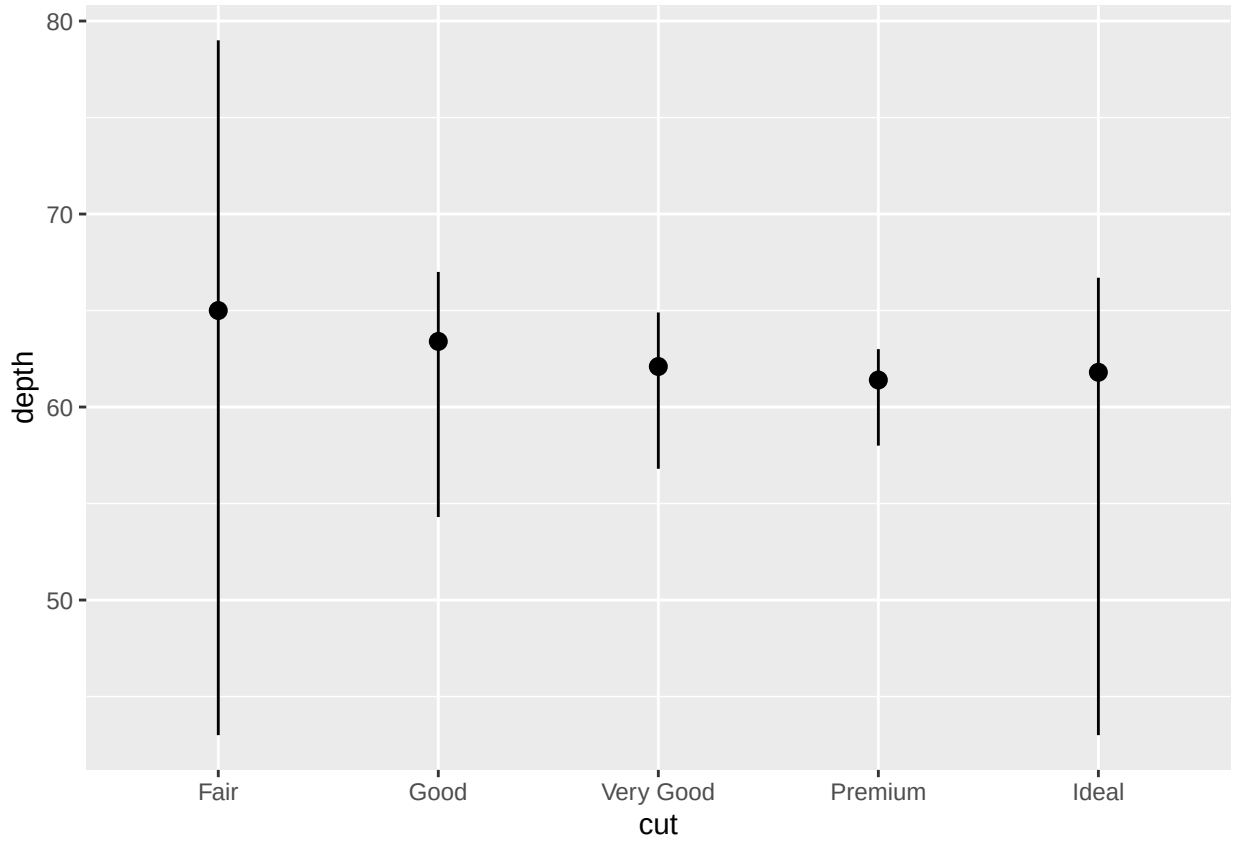
```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

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

Part 3.7.1

Question 1

The default geom associated with stat_summary() is geom_pointrange(). Below is how you could rewrite the previous plot to use geom_pointrange() instead of the stat function,

```
ggplot(data = diamonds) +
  geom_pointrange(mapping = aes(x = cut, y = depth), stat = "summary",
                  fun.min = min, fun.max = max, fun = median)
```



Question 2

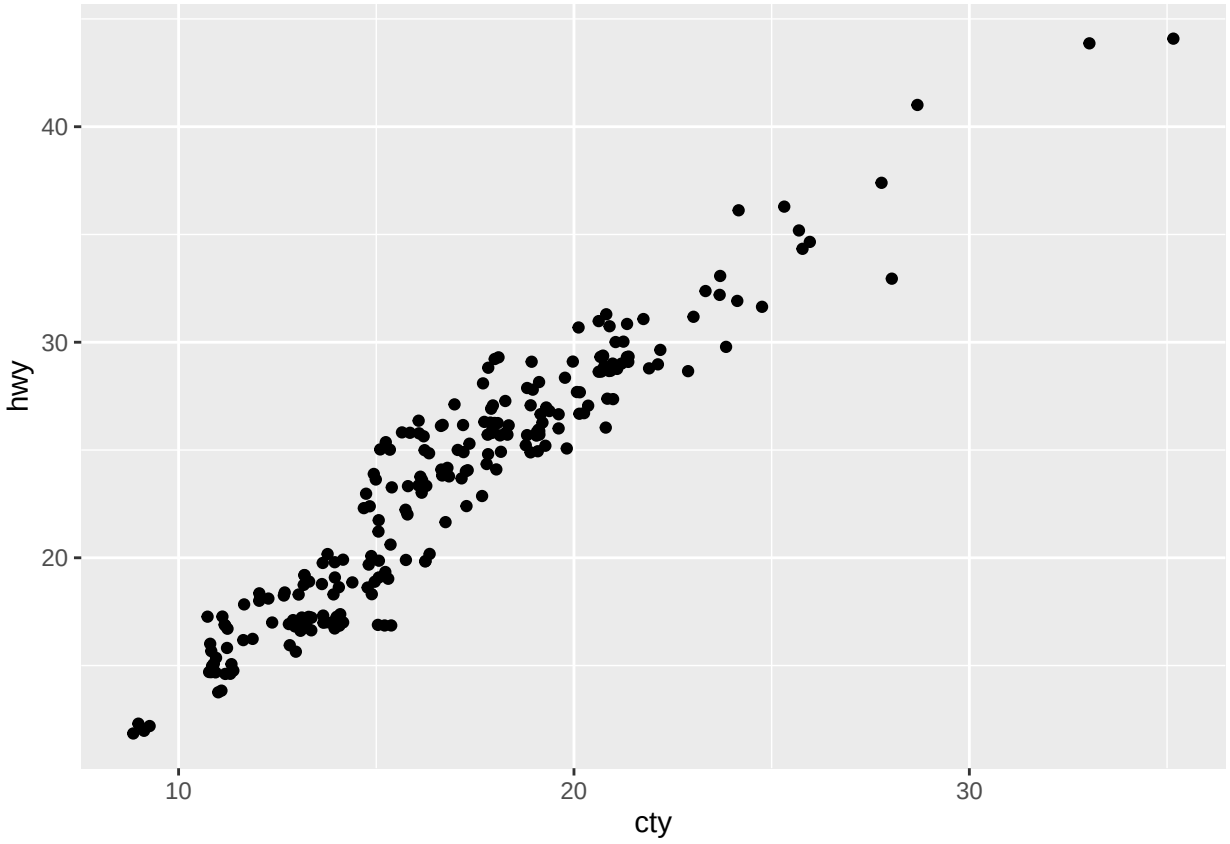
geom_col() is a wrapper over geom_bar() and has stat_identity() as the default stat, which is different than the default stat of geom_bar(). The geom_col() needs both x and y values from the data, representing the bar height. Whereas, geom_bar() only needs an x variable as it uses stat_count() as a default stat, which pre-processes input data by counting the number of observations for each value of x and so, the values of these counts are used for the y variable.

Part 3.8.1

Question 1

Because there are multiple observations for each combination of cty and hwy values, there is overplotting. We could improve it by using a jitter position adjustment as it shows the area with more observations.

```
ggplot(data = mpg, mapping = aes(x = cty, y = hwy)) +
  geom_point(position = "jitter")
```



Question 2

```
?geom_jitter()
```

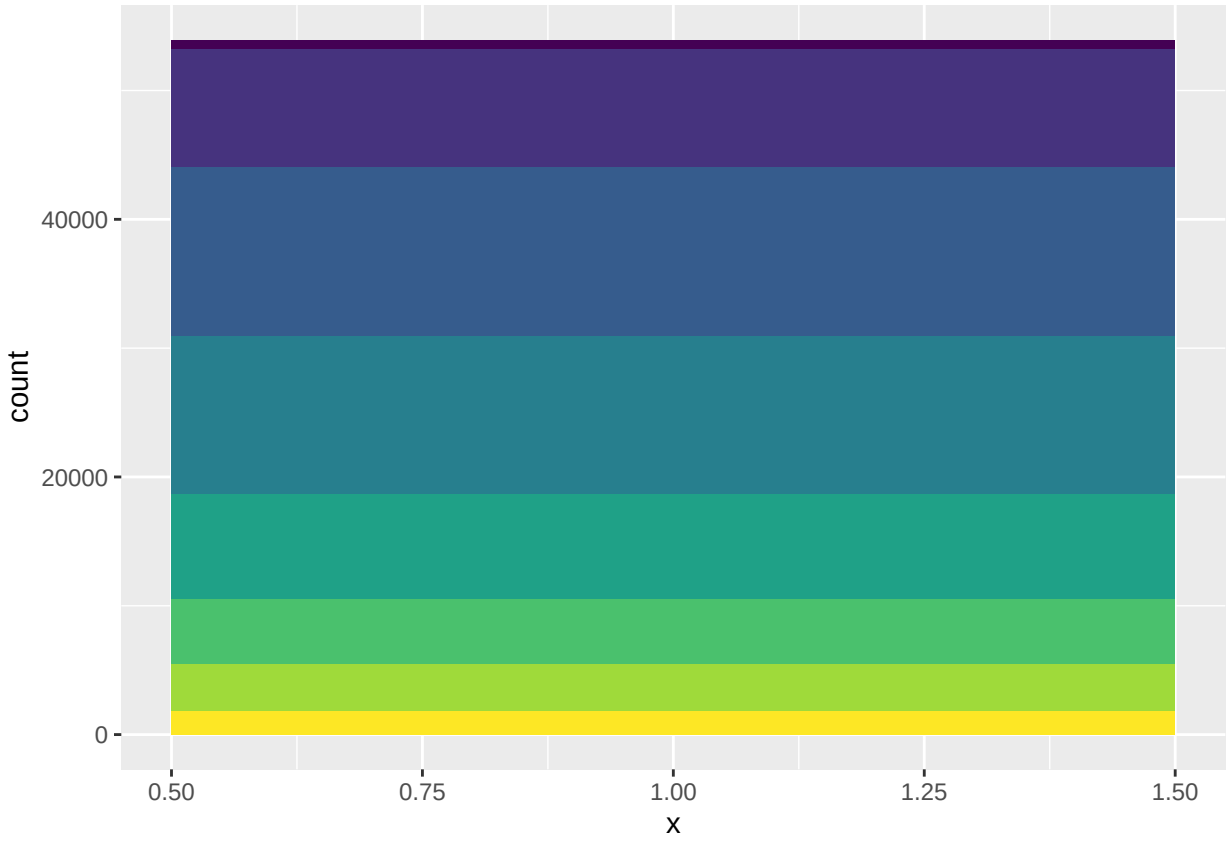
The "width" and "height" parameters to geom_jitter() control the amount of jittering. Width controls the amount of horizontal displacement, and height controls the amount of vertical displacement.

Part 3.9.1

Question 1

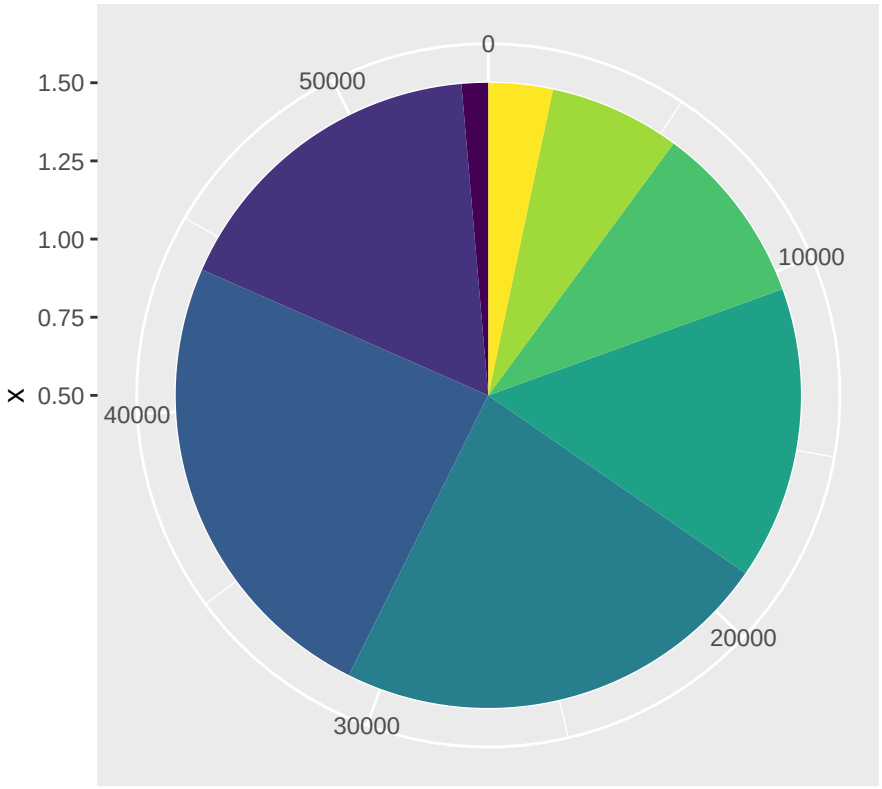
A stacked bar chart.

```
ggplot(data = diamonds) +
  geom_bar(mapping = aes(x = 1, fill = clarity), show.legend = FALSE, width = 1)
```



A stacked bar chart with polar coordinates is a pie chart. We also need to add the argument theta = "y" to coord_polar() to map "y" to the angle of each section.

```
ggplot(data = diamonds) +
  geom_bar(mapping = aes(x = 1, fill = clarity), show.legend = FALSE, width = 1) +
  coord_polar(theta = "y")
```



Question 2

```
?labs()
```

labs() is used for modifying axis, legend, and plot labels. It adds plot title, plot subtitle, plot caption, and axis titles.