Lecture 2 in class

Jung-Jin Lee

1/14/2020

```
first_name <- c("Lisa", "John", "Chuck", "Matt")
last_name <- c("Simpson", "Smith", "Williams", "June")
age_yrs <- c(8, 42, 81, 23)

length(first_name)
```

```
## [1] 4
```

```
last_name[3]
```

```
## [1] "Williams"
```

```
last_name[c(3, 4)]
```

```
## [1] "Williams" "June"
```

```
book <- data.frame(first = first_name,
                   last = last_name,
                   age = age_yrs)

dim(book)
```

```
## [1] 4 3
```

```
nrow(book)
```

```
## [1] 4
```

```
ncol(book)
```

```
## [1] 3
```

```
dim(book)
```

```
## [1] 4 3
```

```
book[2, 2]
```

```
## [1] Smith
## Levels: June Simpson Smith Williams
```

```
book$age
```

```
## [1]  8 42 81 23
```

```
book[, 3]
```

```
## [1]  8 42 81 23
```

```r
book[1:3, 2:3]
```

```
##       last age
## 1  Simpson   8
## 2    Smith  42
## 3 Williams  81
```

```r
gender <- c("Female", "Male", "Male", "Unknown")
book$sex <- gender

book$remark <- "friend"

book$extra <- c("A", "B")

df <- data.frame(a = 1:6, b = "some random")

df$third <- c("aa", "bb", "cc")

df$fourth <- c("xxx", "zzz")
```

```r
d <- read.table(file = "heights.txt", header = TRUE, sep = " ")
```

```r
dim(d)
```

```
## [1] 1375    2
```

```r
names(d)
```

```
## [1] "Mheight" "Dheight"
```

```r
head(d)
```

```
##   Mheight Dheight
## 1    59.7    55.1
## 2    58.2    56.5
## 3    60.6    56.0
## 4    60.7    56.8
## 5    61.8    56.0
## 6    55.5    57.9
```

```r
tail(d)
```

```
##      Mheight Dheight
## 1370    69.5    70.4
## 1371    69.1    70.1
## 1372    65.0    71.6
## 1373    66.3    71.4
## 1374    70.8    71.0
## 1375    63.0    73.1
```

```r
vec <- 1:100
head(vec)
```

```
## [1] 1 2 3 4 5 6
```

```r
tail(vec)
```

```
## [1]  95  96  97  98  99 100
```

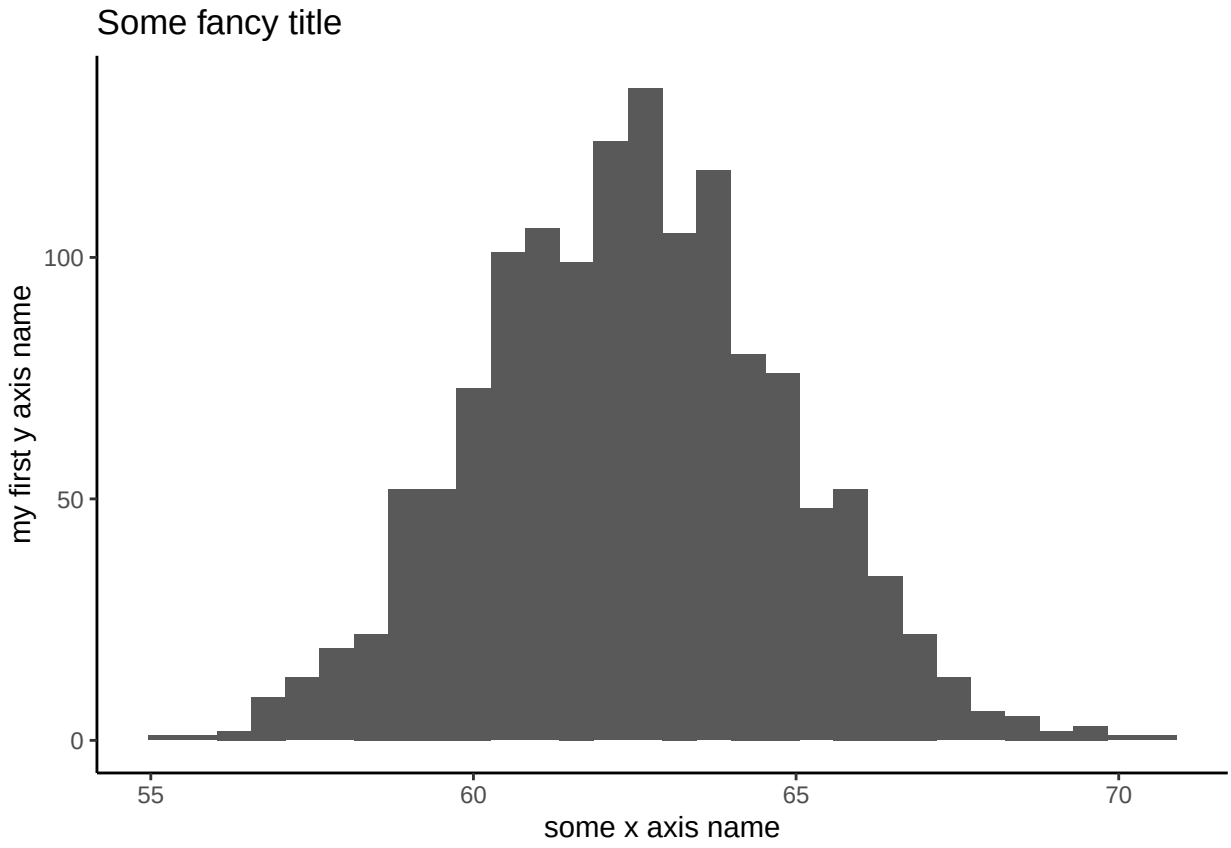
Some sentence

Another sentence

```
library(tidyverse)

g1 <- ggplot(d, aes(x = Mheight)) +
  geom_histogram() +
  theme_classic() +
  ggtitle("Some fancy title") +
  xlab("some x axis name") +
  ylab("my first y axis name")

print(g1)
```



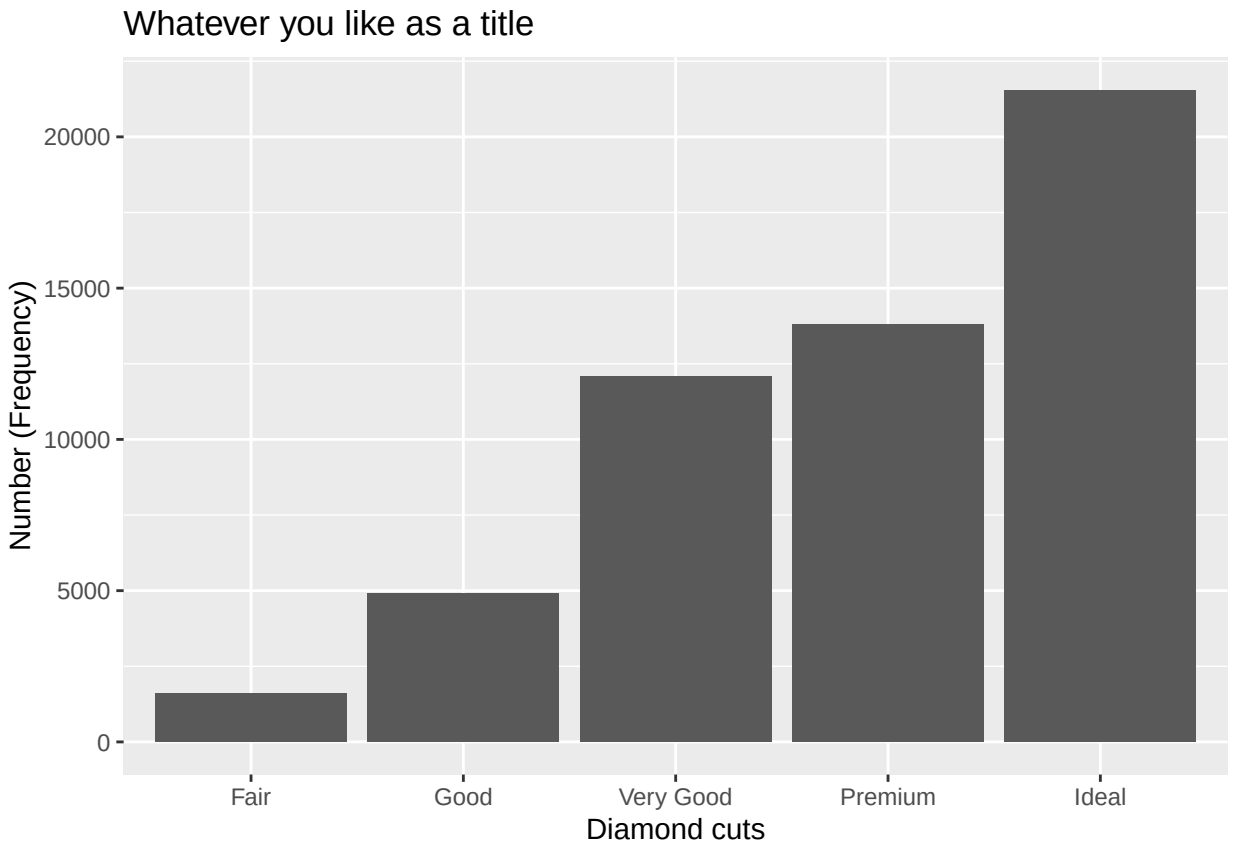

```
library(tidyverse)

df <- diamonds ## diamonds comes with ggplot2
dim(df)
```

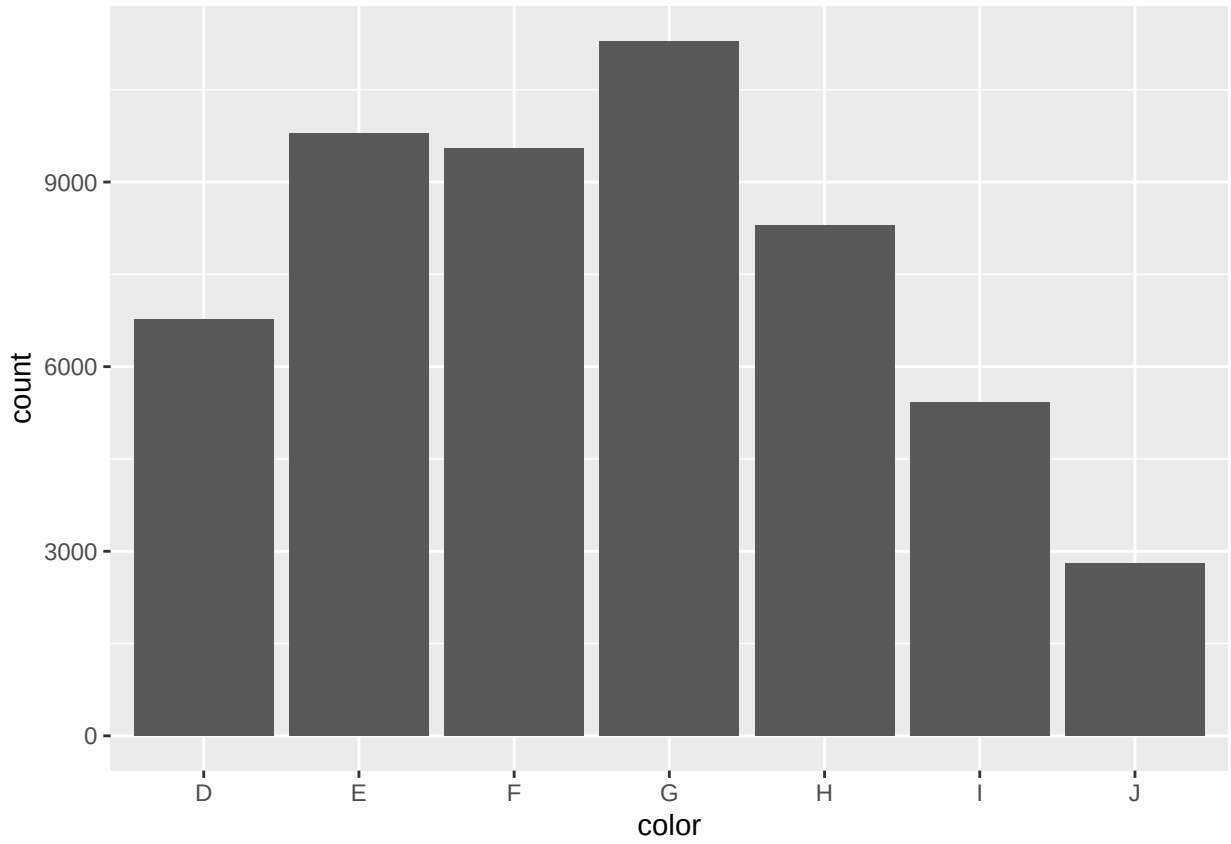
```
## [1] 53940    10
```

```
g <- ggplot(data = df, aes(x = cut)) +
  geom_bar() +
  ggtitle("Whatever you like as a title") +
  xlab("Diamond cuts") +
  ylab("Number (Frequency)")
```

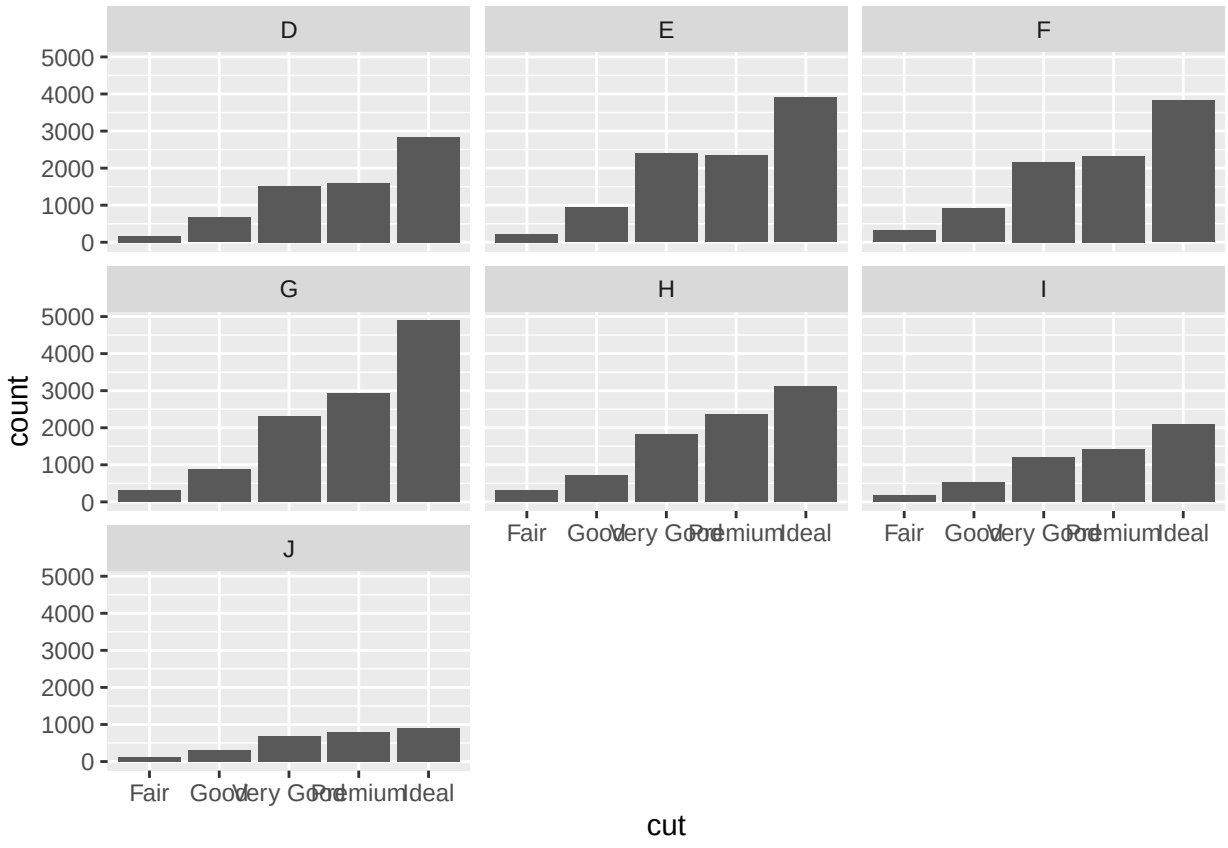
```r
print(g)
```



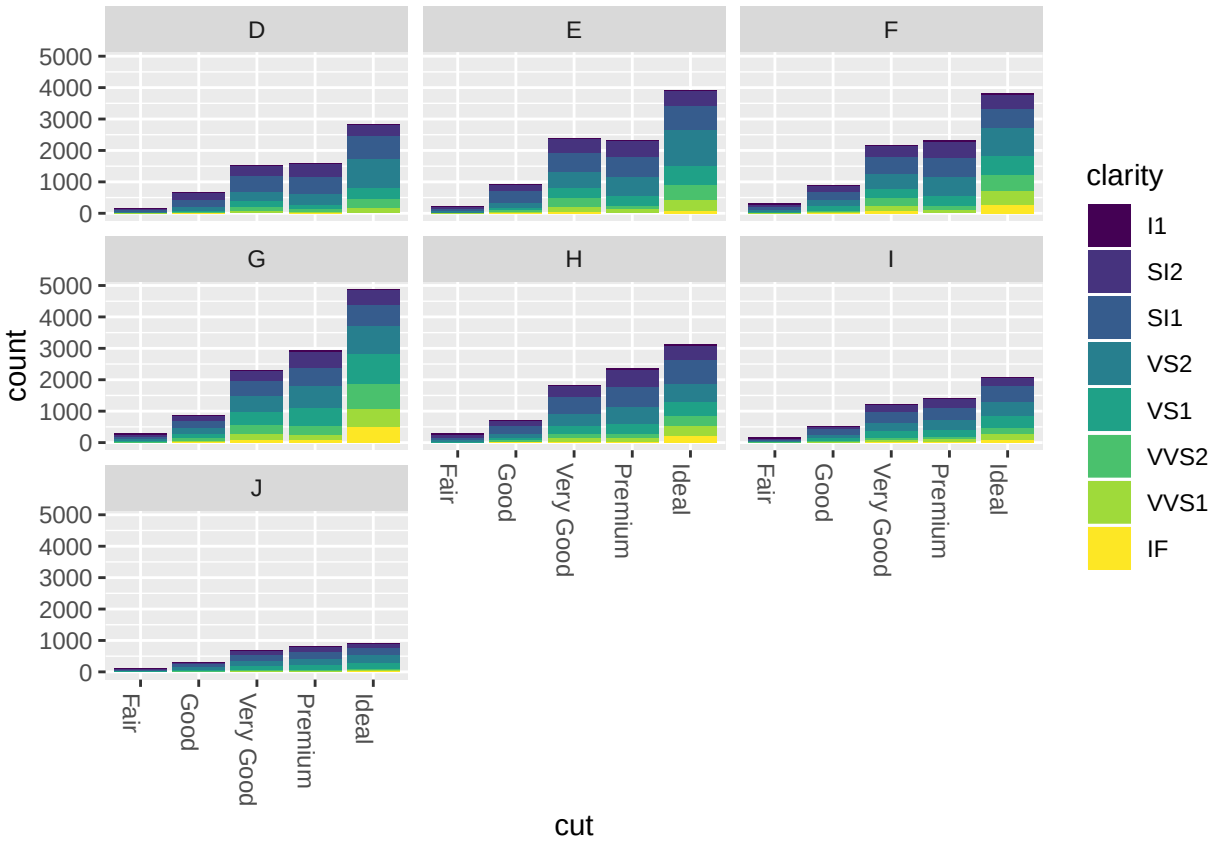

```r
ggplot(data = df, aes(x = color)) + geom_bar()
```



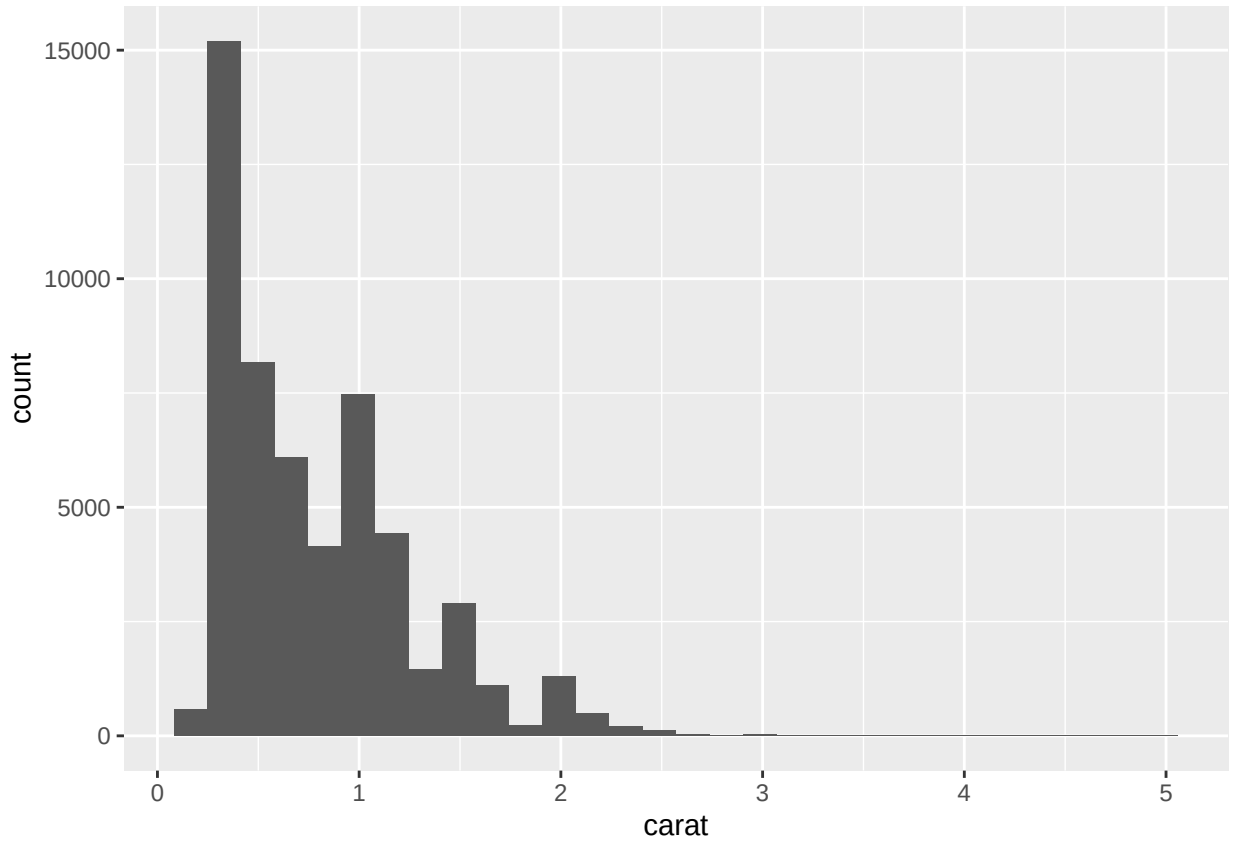

```
ggplot(df, aes(cut)) + ## can omit 'data = ' and 'x = '
  geom_bar() +
  facet_wrap(~color)
```




```
ggplot(df, aes(cut, fill = clarity)) + ## can omit 'data = ' and 'x = '
  geom_bar() +
  facet_wrap(~color) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0))
```




```
ggplot(df, aes(carat)) +
  geom_histogram(bins = 30)
```




```
?quantile

sqrt(5.546511)
```

```
## [1] 2.355103
```