Flexible, Beautiful, Customizable Graphs

Understanding ggplot2's Grammar

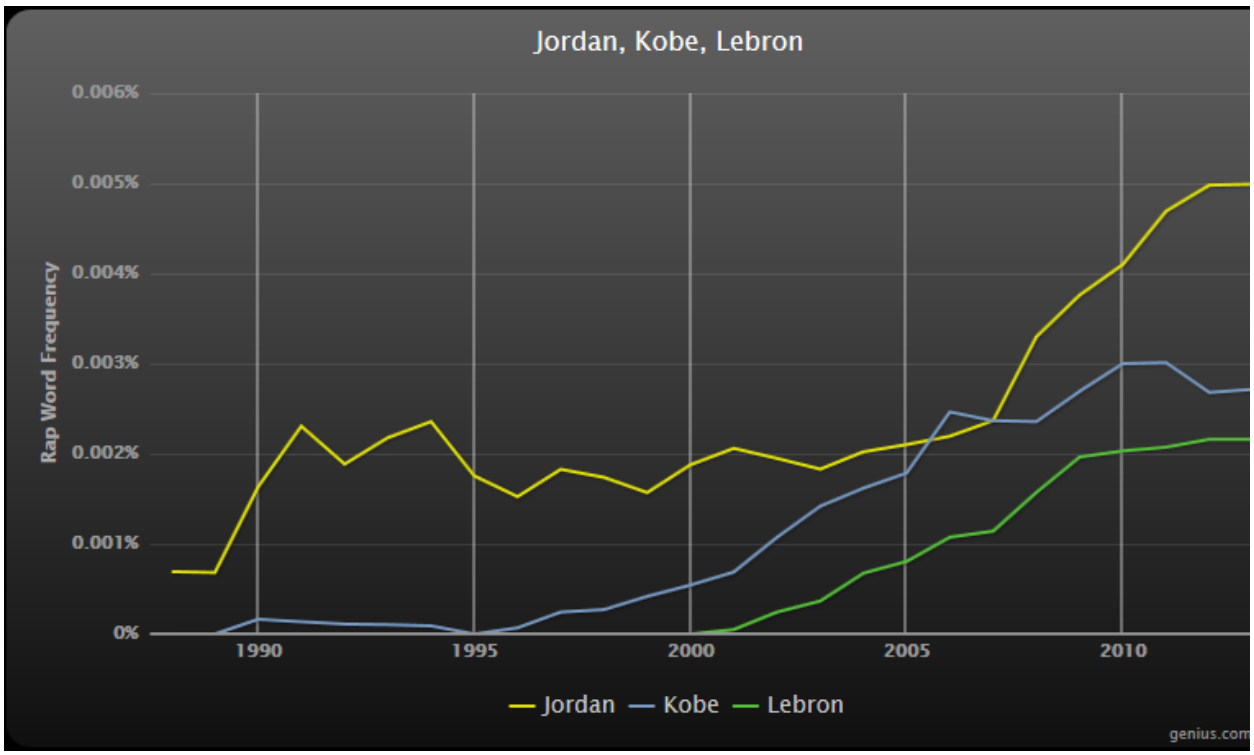
Kyle Nickodem

Monday, February 2nd, 2015

Take Home Points

1. Become aware of the various components that comprise a graph
2. Understand how ggplot2 uses these components to construct a plot

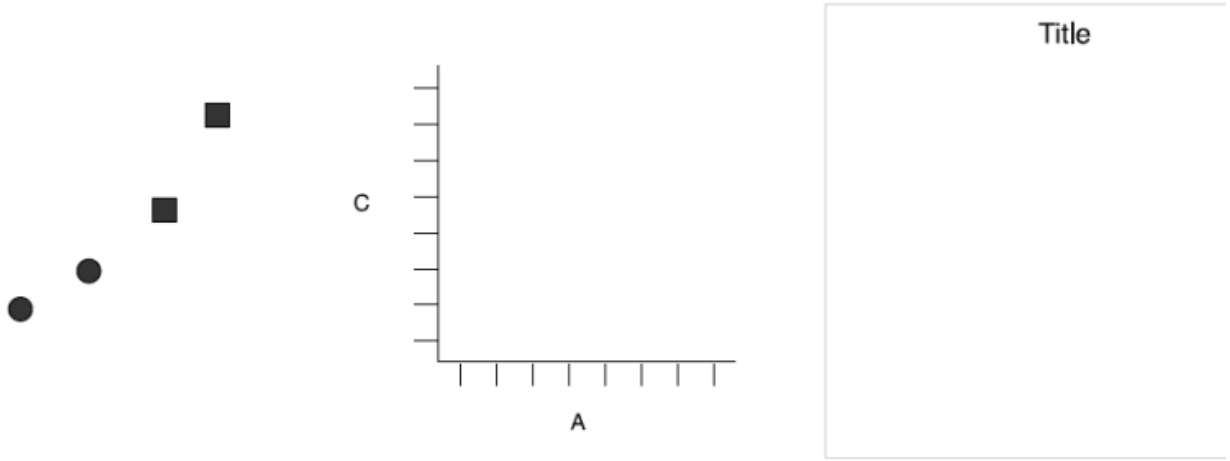
What is this graphic trying to tell us?



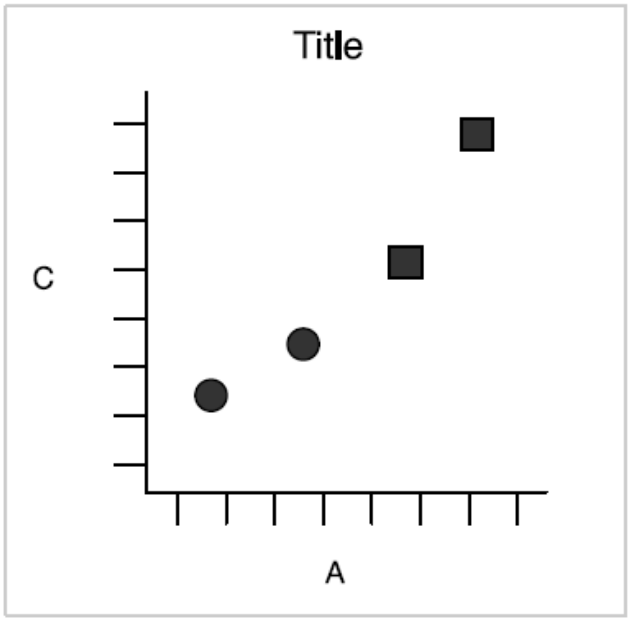
The Grammar of Graphics

Why is it necessary to understand the grammar?

1. ggplot2 operates using this grammar
2. It provides us with a process to think about the structure that underlies statistical grapics



- Data and asthetic mapping
- Geometric objects
- Scales and coordinate system
- Plot annotations and themes

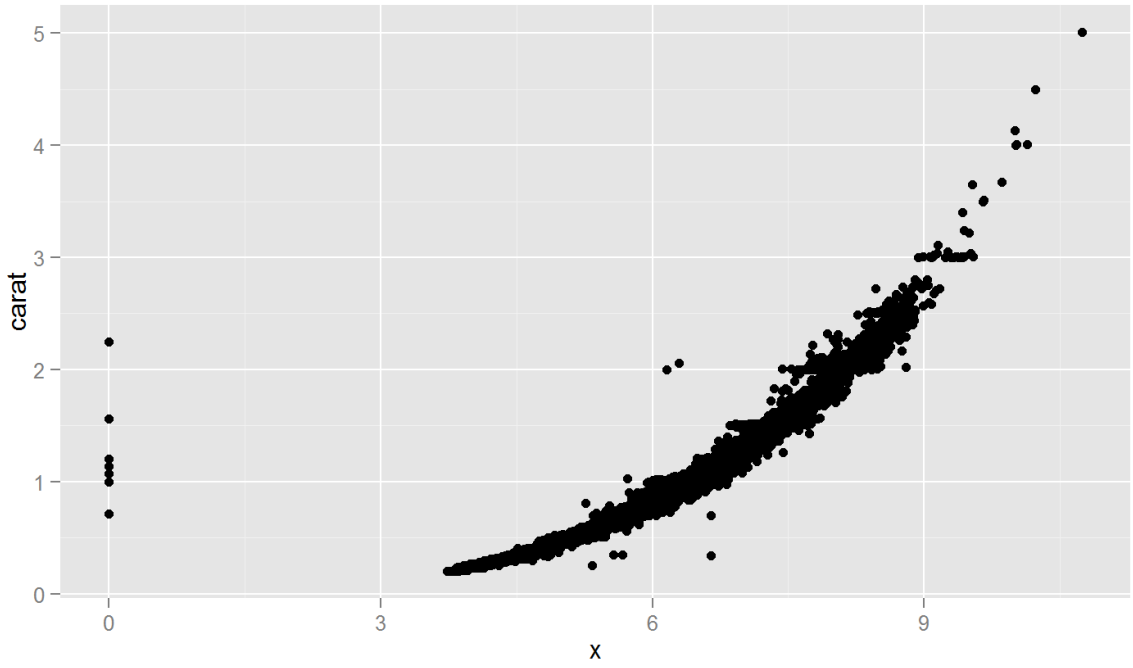

Title
C
A

How does this work in ggplot2?

```
library(ggplot2)
head(diamonds)
```

```
##   carat       cut color clarity depth table price    x    y    z
## 1  0.23     Ideal     E     SI2  61.5    55   326 3.95 3.98 2.43
## 2  0.21   Premium     E     SI1  59.8    61   326 3.89 3.84 2.31
## 3  0.23      Good     E     VS1  56.9    65   327 4.05 4.07 2.31
## 4  0.29   Premium     I     VS2  62.4    58   334 4.20 4.23 2.63
## 5  0.31      Good     J     SI2  63.3    58   335 4.34 4.35 2.75
## 6  0.24 Very Good     J    VVS2  62.8    57   336 3.94 3.96 2.48
```

```
ggplot(data=diamonds, aes(x = x, y = carat)) +
  geom_point()
```



8/22

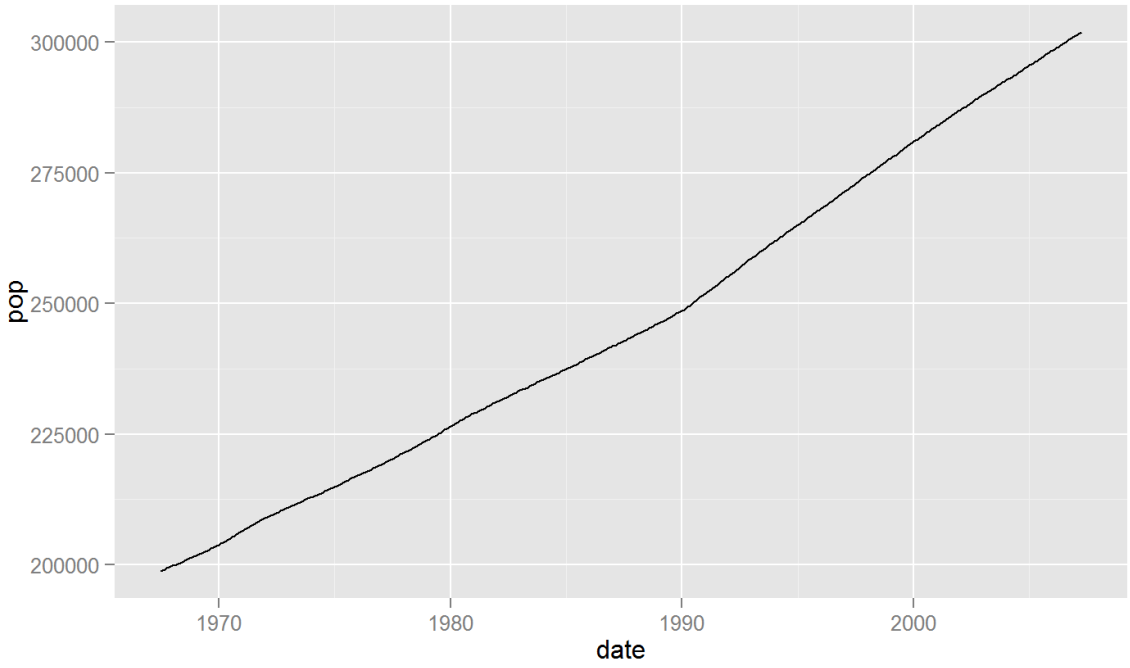
```
ggplot(data=diamonds, aes(x = x, y = carat)) +
  geom_point()


ggplot() +
  layer(data = diamonds, mapping = aes(x = x, y = carat),
                geom = "point", stat = "identity", pos = "identity")
  scale_x_continuous() +
  scale_y_continuous() +
  coord_cartesian() +
  theme()


ggplot(diamonds, aes(x,carat)) +
  geom_point()
```

Quiz Time

What will this plot look like?

```
ggplot(data = economics, aes(x = date, y = pop)) +
  geom_line()
```

```
ggplot(data = economics, aes(x = date, y = pop)) +
  geom_line()
```

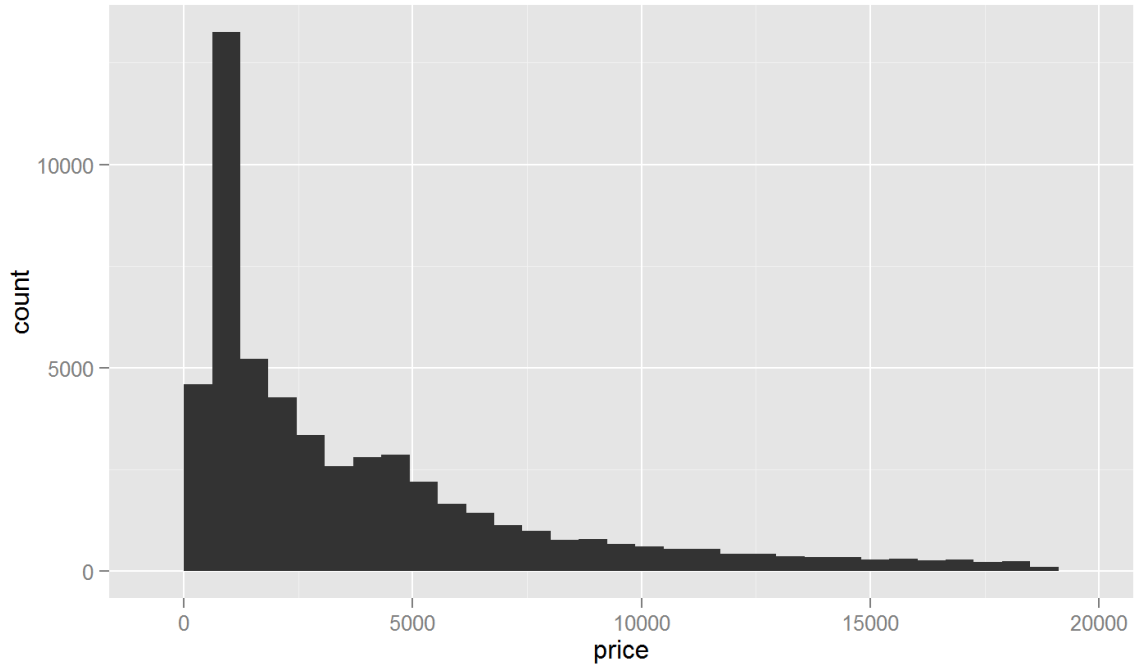


11/22

Question #2

```
ggplot(data = diamonds, aes(x = price)) +
  geom_histogram()
```

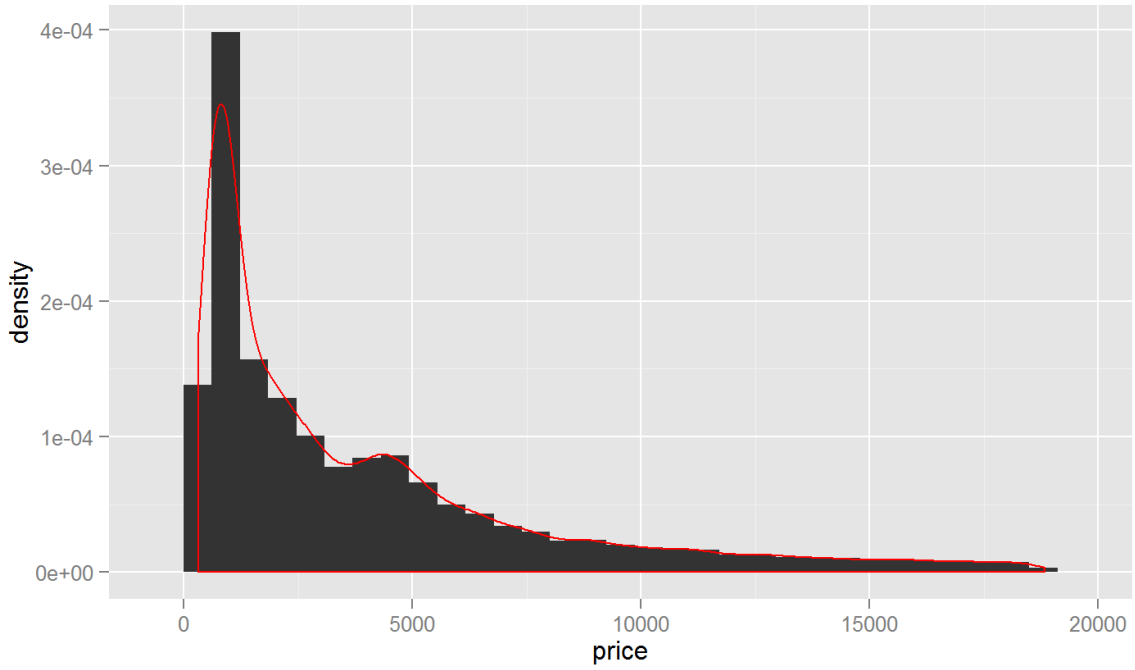
```
ggplot(data = diamonds, aes(x = price)) +
  geom_histogram()
```



13/22

Add multiple geometric objects

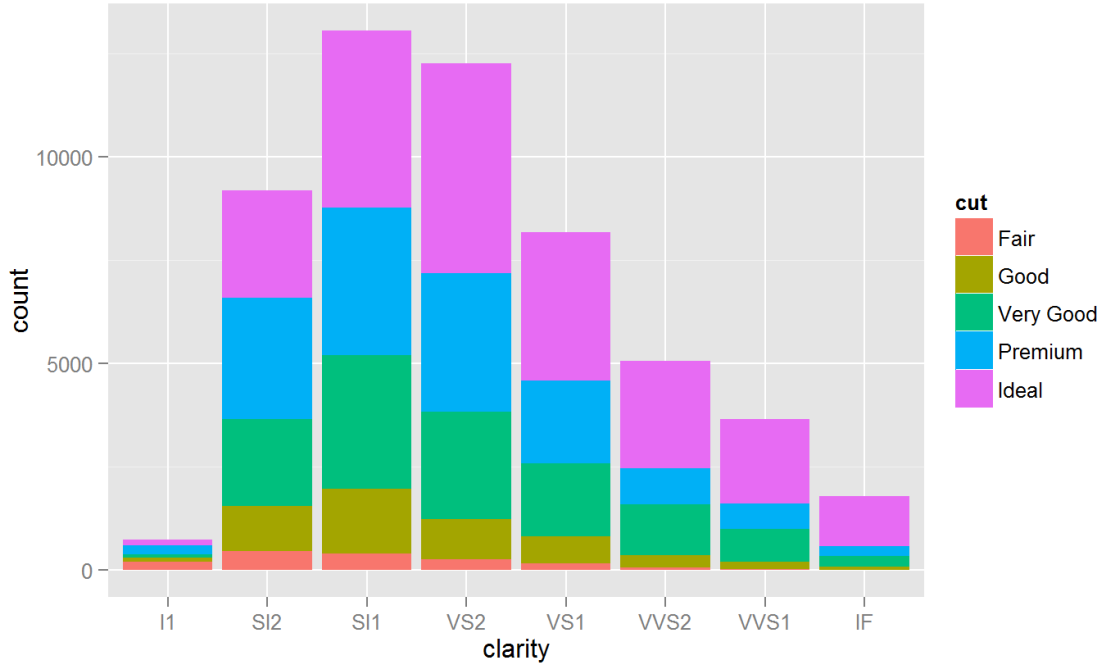
```
ggplot(data = diamonds, aes(x = price)) +
  geom_histogram(aes(y = ..density..)) + geom_density(color = "red")
```



14/22

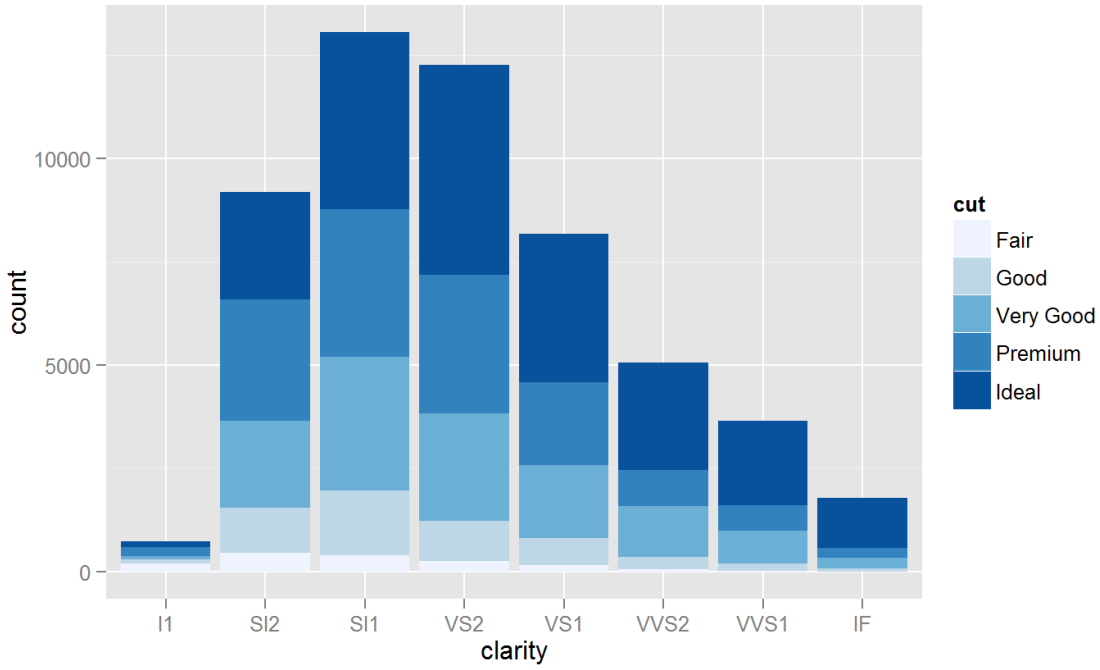
Frequency of diamond clarity by cut?

```
ggplot(data=diamonds, aes(x = clarity, fill = cut)) +
  geom_bar()
```



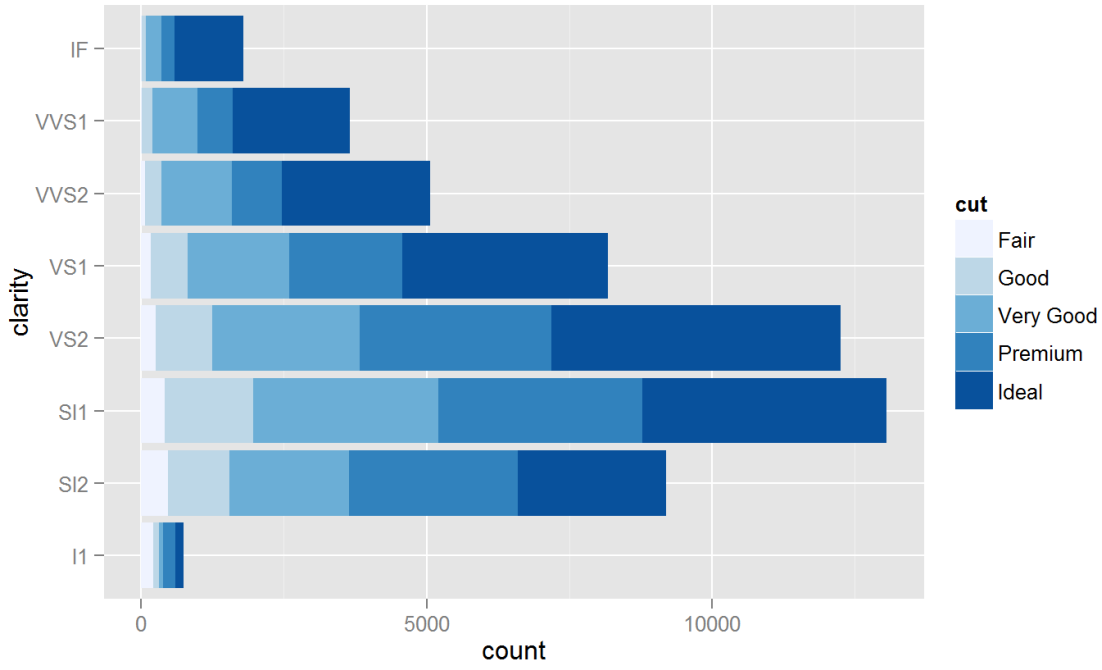
Change color scheme

```
library(RColorBrewer)
ggplot(data=diamonds, aes(x = clarity, fill = cut)) +
  geom_bar() + scale_fill_brewer()
```



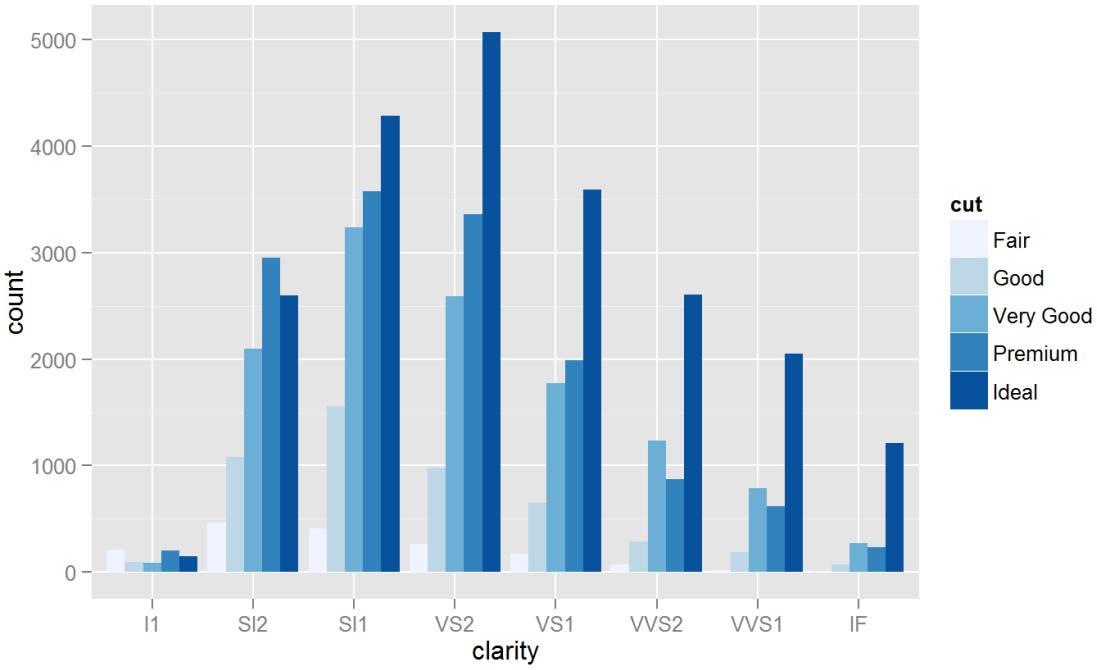
Flip the coordinate grid

```
ggplot(data=diamonds, aes(x = clarity, fill = cut)) +
  geom_bar() + scale_fill_brewer() +
  coord_flip()
```



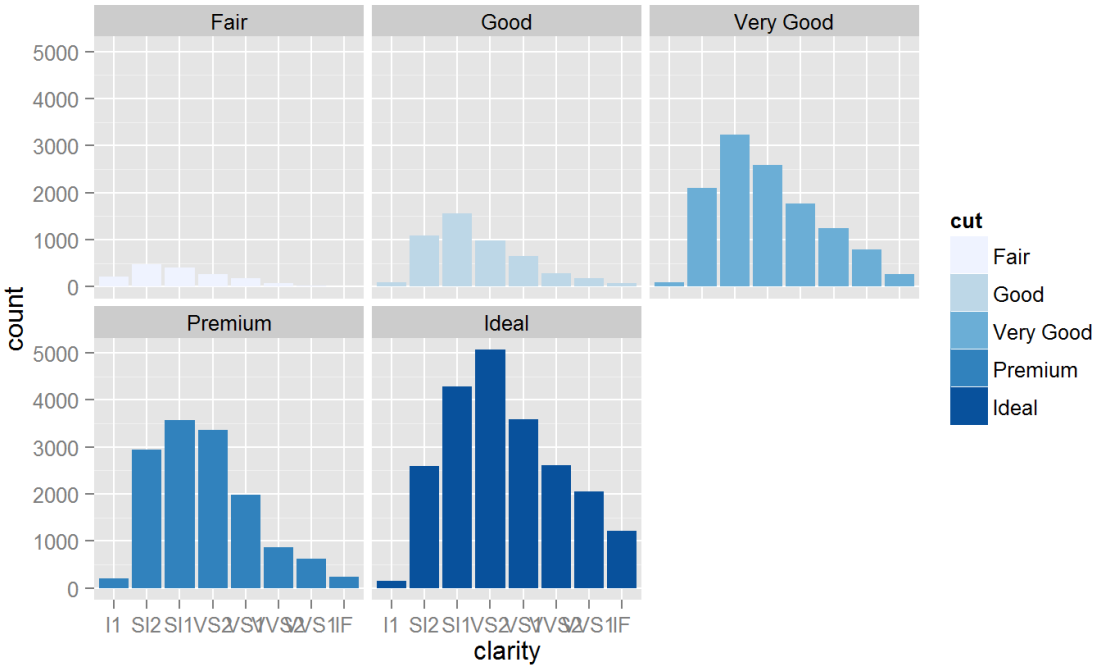
Dodged bar chart

```
ggplot(data=diamonds, aes(x = clarity, fill = cut)) +
  geom_bar(position = "dodge") + scale_fill_brewer()
```



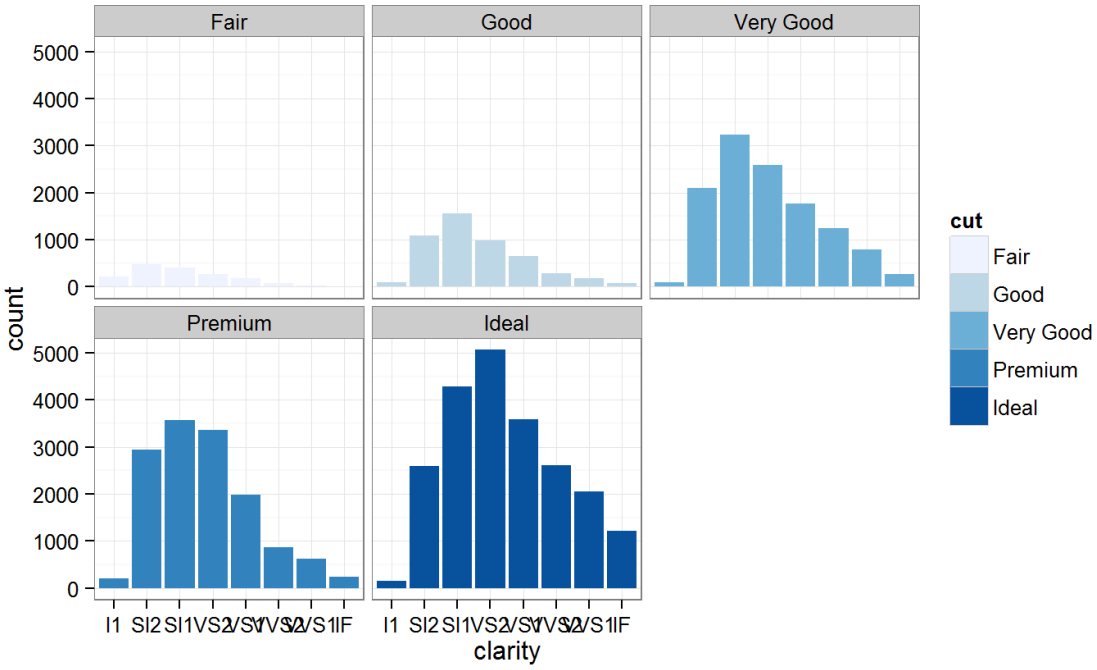
Facet

```
ggplot(data=diamonds, aes(x = clarity, fill = cut)) +
  geom_bar() + scale_fill_brewer() + facet_wrap(~cut)
```



Alter the theme

```
ggplot(data=diamonds, aes(x = clarity, fill = cut)) +
  geom_bar() + scale_fill_brewer() +  facet_wrap(~cut) + theme_bw()
```

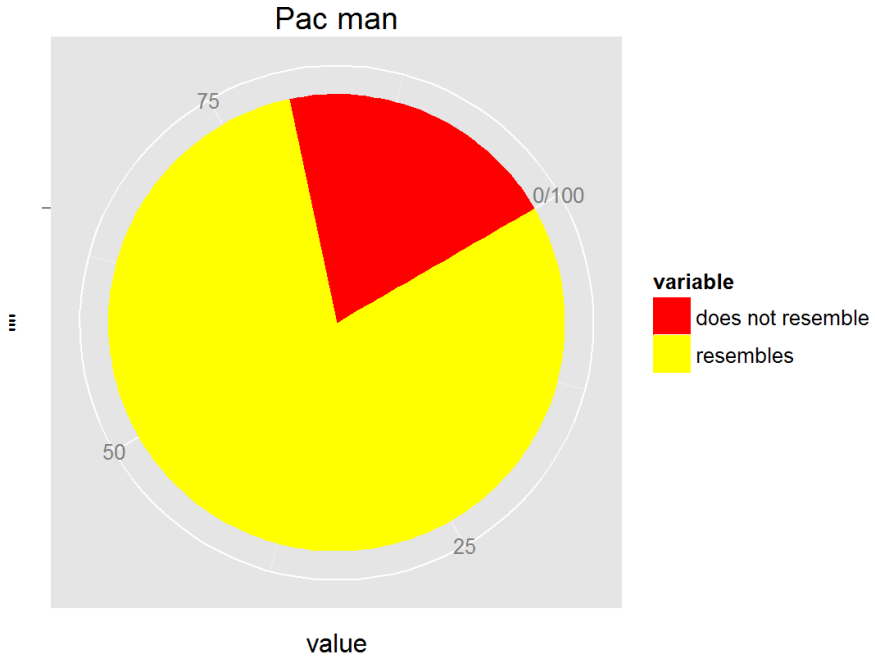


Resources

- Hadley's ggplot2 documentation - docs.ggplot2.org
- ZevRoss ggplot2 cheatsheet
- R Graphics Cookbook
- R Color Brewer
- Wilkinson, L. (2006). *The grammar of graphics*. Springer - Available for free from through the UM Library portal.

Hadley's favorite pie chart

```
ggplot(df, aes(x = "", y = value, fill = variable)) + geom_bar(width
stat = "identity") + scale_fill_manual(values = c("red", "yellow")) +
  coord_polar("y", start = pi / 3) + labs(title = "Pac man")
```



22/22