Data Visualization Examples

Edmund Neil Quiros

1 de marzo de 2018

## References

library(tidyverse)

## Loading tidyverse: ggplot2  
## Loading tidyverse: tibble  
## Loading tidyverse: tidyr  
## Loading tidyverse: readr  
## Loading tidyverse: purrr  
## Loading tidyverse: dplyr

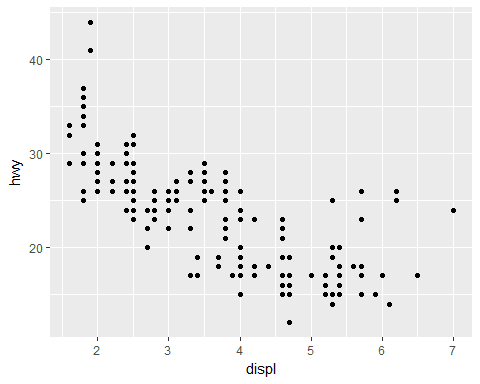
## Conflicts with tidy packages ----------------------------------------------

## filter(): dplyr, stats  
## lag(): dplyr, stats

#MPG Data Frame  
mpg

## # A tibble: 234 x 11  
## manufacturer model displ year cyl trans drv cty hwy  
## <chr> <chr> <dbl> <int> <int> <chr> <chr> <int> <int>  
## 1 audi a4 1.8 1999 4 auto(l5) f 18 29  
## 2 audi a4 1.8 1999 4 manual(m5) f 21 29  
## 3 audi a4 2.0 2008 4 manual(m6) f 20 31  
## 4 audi a4 2.0 2008 4 auto(av) f 21 30  
## 5 audi a4 2.8 1999 6 auto(l5) f 16 26  
## 6 audi a4 2.8 1999 6 manual(m5) f 18 26  
## 7 audi a4 3.1 2008 6 auto(av) f 18 27  
## 8 audi a4 quattro 1.8 1999 4 manual(m5) 4 18 26  
## 9 audi a4 quattro 1.8 1999 4 auto(l5) 4 16 25  
## 10 audi a4 quattro 2.0 2008 4 manual(m6) 4 20 28  
## # ... with 224 more rows, and 2 more variables: fl <chr>, class <chr>

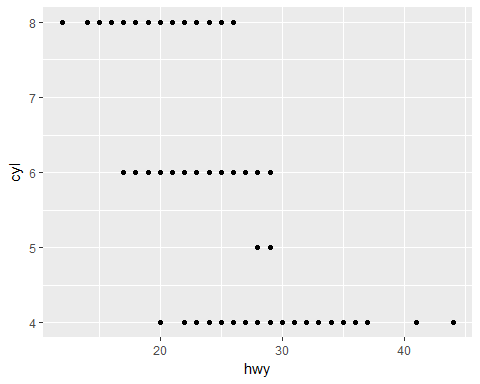
#Creating a ggplot  
  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y= hwy))



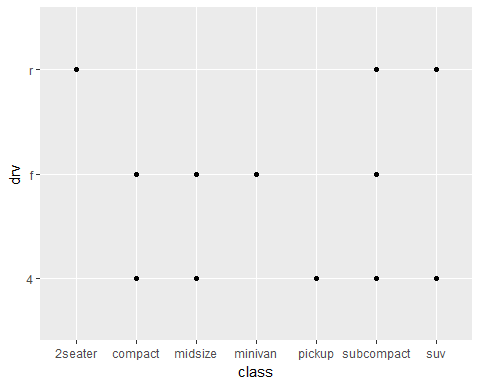
#A graphing template  
#ggplot(data= <DATA>) +  
 #<GEOM\_FUNCTION>(mapping= aes(x=DATA.VAR, y=DATA.VAR))  
  
  
#Exercises  
  
ggplot(data = mpg)



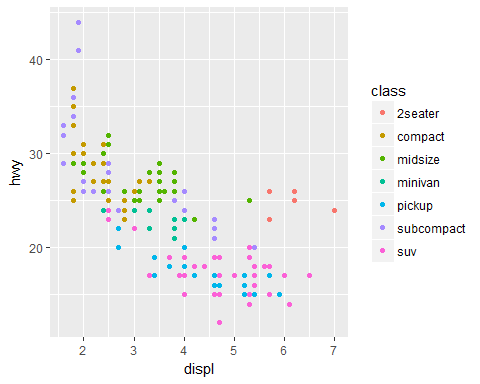
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=hwy, y=cyl))



ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=class, y=drv))

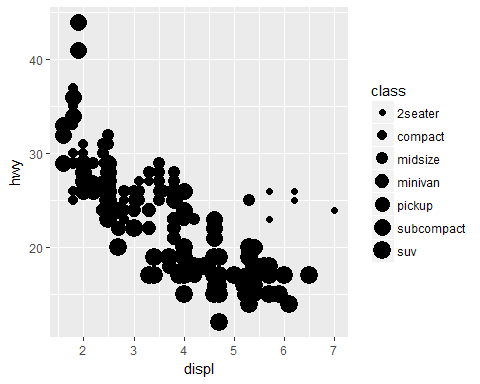


#Aesthetic mappings  
  
#Color  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y=hwy, color = class))

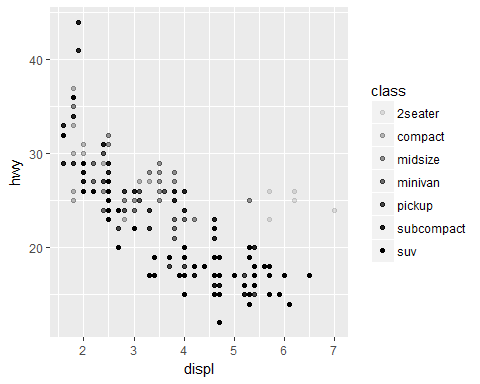


#Size  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y=hwy, size = class))

## Warning: Using size for a discrete variable is not advised.



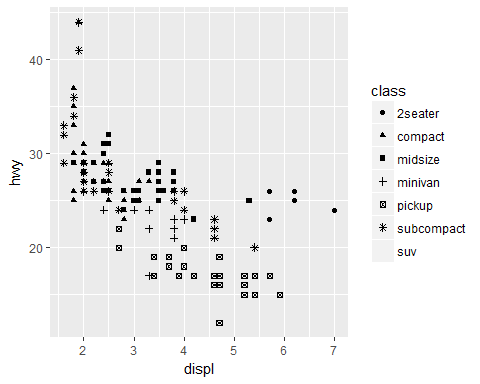
#Transparency  
ggplot(data=mpg) +  
 geom\_point(mapping = aes(x=displ, y=hwy, alpha = class))



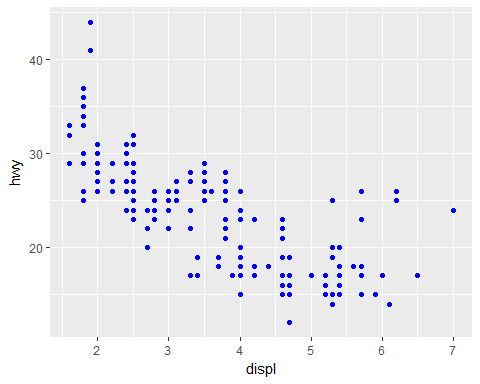
#Shape  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y=hwy, shape=class))

## Warning: The shape palette can deal with a maximum of 6 discrete values  
## because more than 6 becomes difficult to discriminate; you have 7.  
## Consider specifying shapes manually if you must have them.

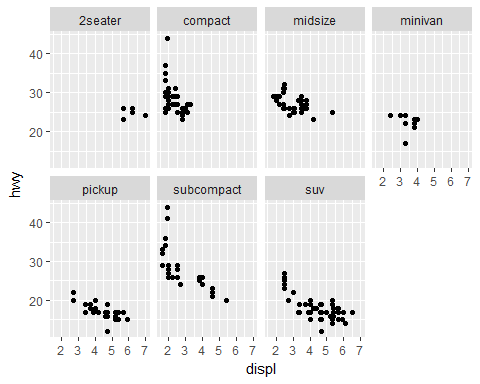
## Warning: Removed 62 rows containing missing values (geom\_point).



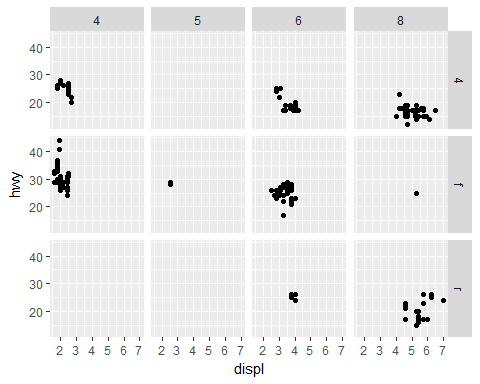
#All aesthetic points in blue  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y=hwy),color = "blue")



#Facets  
  
#Facet wrap  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y=hwy)) +  
 facet\_wrap(~ class, nrow = 2)

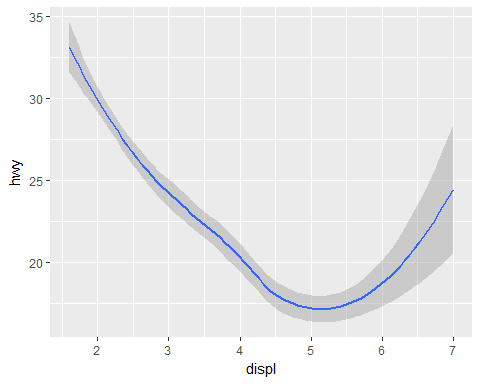


#Facet grid  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy)) +  
 facet\_grid(drv~cyl)



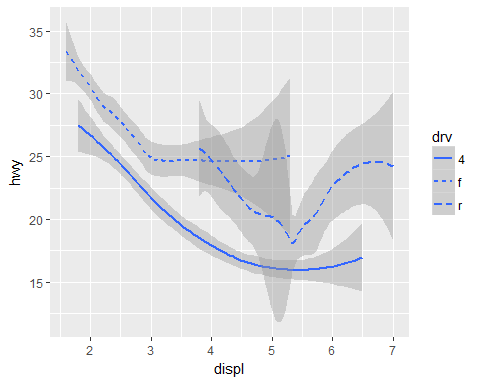
#Geometric objects  
  
ggplot(data = mpg) +  
 geom\_smooth(mapping = aes(x = displ, y = hwy))

## `geom\_smooth()` using method = 'loess'



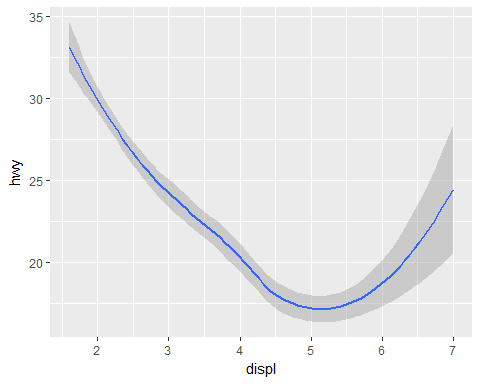
ggplot(data = mpg) +  
 geom\_smooth(mapping = aes(x = displ, y = hwy, linetype = drv))

## `geom\_smooth()` using method = 'loess'



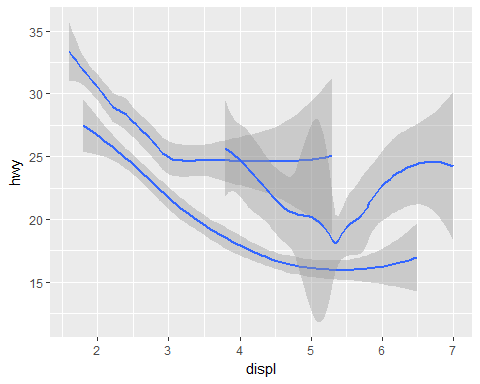
ggplot(data = mpg) +  
 geom\_smooth(mapping = aes(x = displ, y = hwy))

## `geom\_smooth()` using method = 'loess'



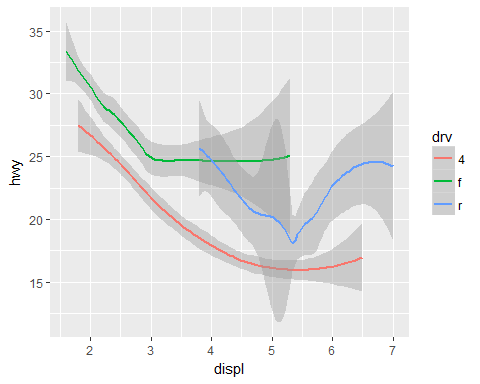
ggplot(data = mpg) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy, group = drv))

## `geom\_smooth()` using method = 'loess'



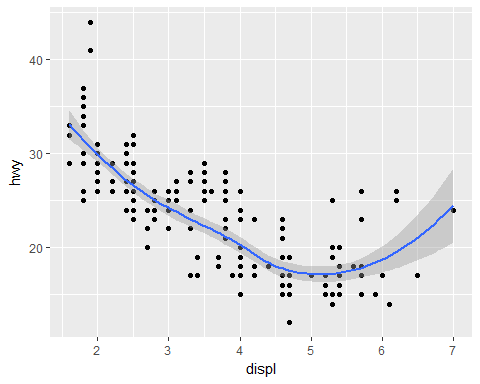
ggplot(data = mpg) +  
 geom\_smooth(mapping = aes(x = displ, y = hwy, color = drv), show.legend = TRUE)

## `geom\_smooth()` using method = 'loess'



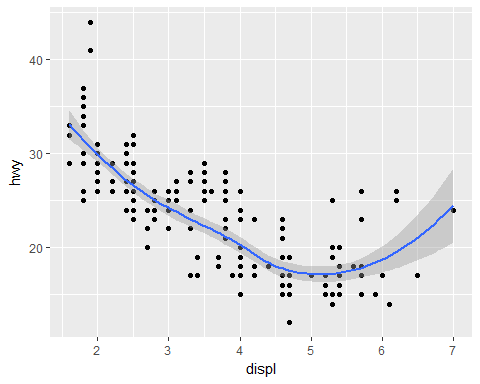
#Multiple geoms in the same plot  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x = displ, y =hwy)) +  
 geom\_smooth(mapping = aes(x = displ, y=hwy))

## `geom\_smooth()` using method = 'loess'



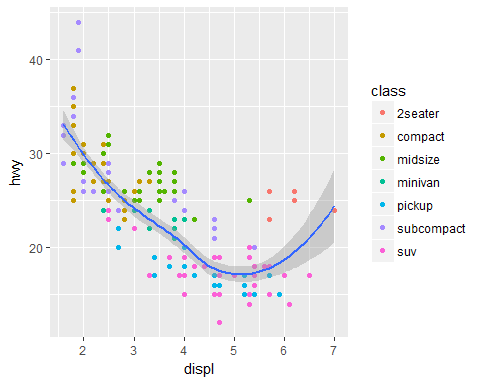
#Multiple geoms in the same plot - less lines  
ggplot(data = mpg, mapping = aes(x= displ, y = hwy)) +  
 geom\_point() +  
 geom\_smooth()

## `geom\_smooth()` using method = 'loess'



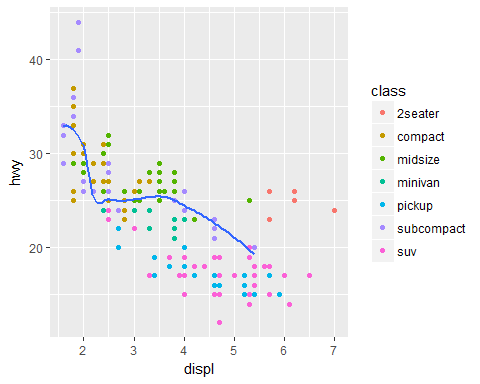
ggplot(data = mpg, mapping = aes(x = displ, y = hwy)) +  
 geom\_point(mapping = aes(color = class)) +  
 geom\_smooth()

## `geom\_smooth()` using method = 'loess'



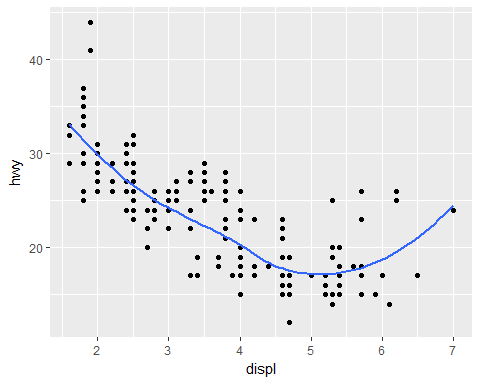
ggplot(data = mpg, mapping = aes(x=displ, y= hwy)) +  
 geom\_point(mapping = aes(color = class)) +  
 geom\_smooth(data = filter(mpg, class == "subcompact"), se = FALSE)

## `geom\_smooth()` using method = 'loess'



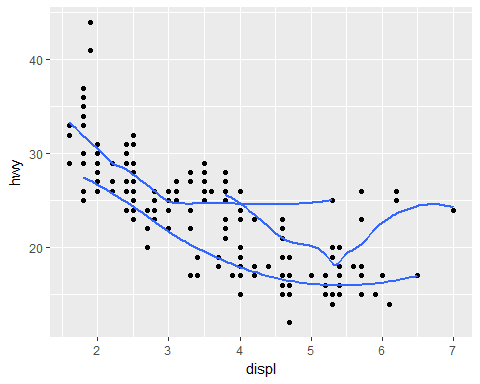
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy)) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy), se = FALSE)

## `geom\_smooth()` using method = 'loess'



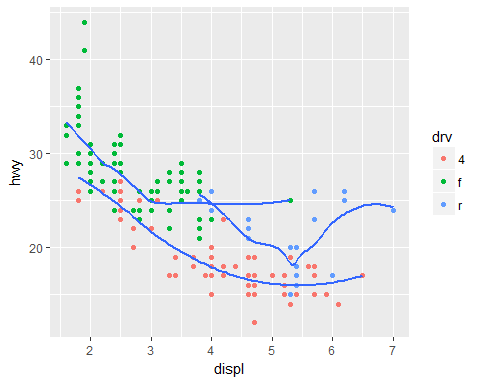
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy)) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy, group = drv), se = FALSE)

## `geom\_smooth()` using method = 'loess'



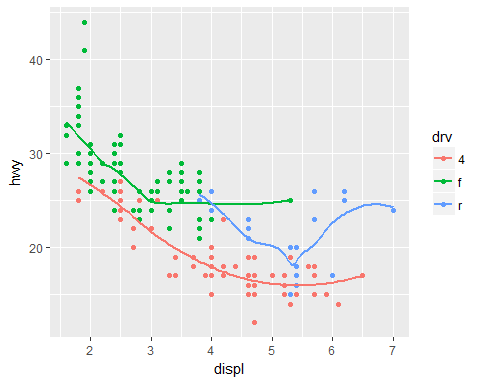
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy, color = drv)) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy, group = drv), se = FALSE)

## `geom\_smooth()` using method = 'loess'



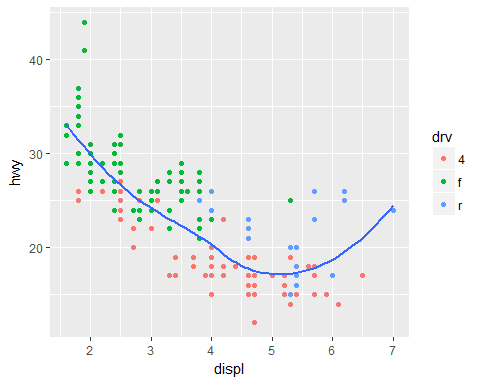
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy, color = drv)) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy, color = drv, group = drv), se = FALSE)

## `geom\_smooth()` using method = 'loess'



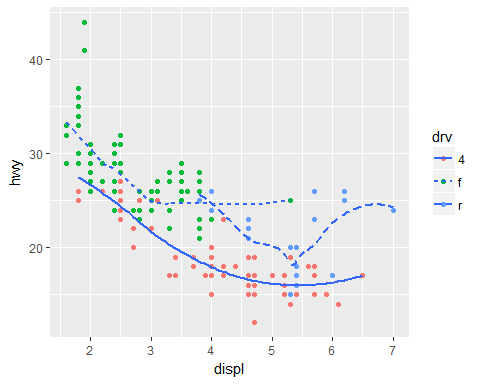
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy, color = drv)) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy), se = FALSE)

## `geom\_smooth()` using method = 'loess'

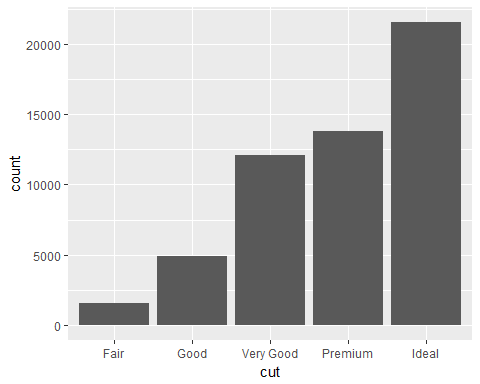


ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y = hwy, color = drv)) +  
 geom\_smooth(mapping = aes(x=displ, y = hwy, linetype = drv), se = FALSE)

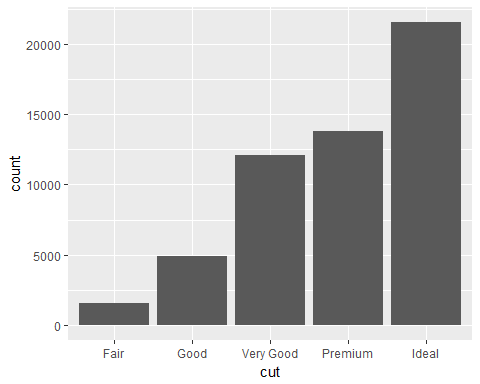
## `geom\_smooth()` using method = 'loess'



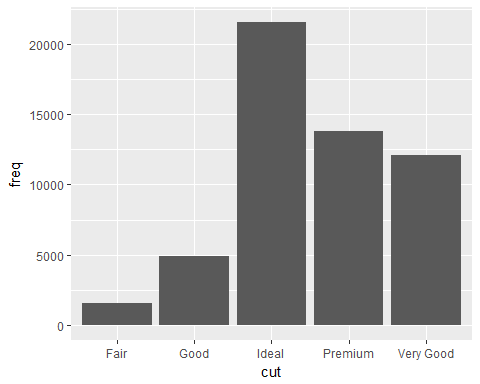
#Statistical transformations  
#Bar plots  
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x=cut))



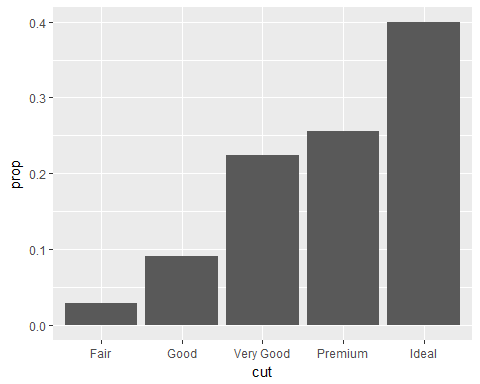
#Bar plot using stat\_count() instead of geom\_bar()  
ggplot(data = diamonds) +  
 stat\_count(mapping = aes(x = cut))



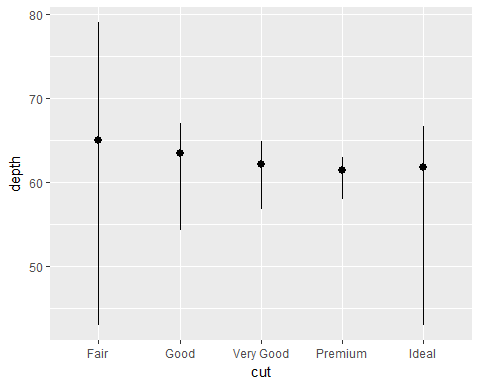
#Bar plot overwriting stat\_count with stat\_identity  
  
demo <- tribble(  
 ~cut, ~freq,  
 "Fair", 1610,  
 "Good", 4906,  
 "Very Good", 12082,  
 "Premium", 13791,  
 "Ideal", 21551  
)  
  
ggplot(data = demo) +  
 geom\_bar(mapping = aes(x=cut,y=freq), stat = "identity")



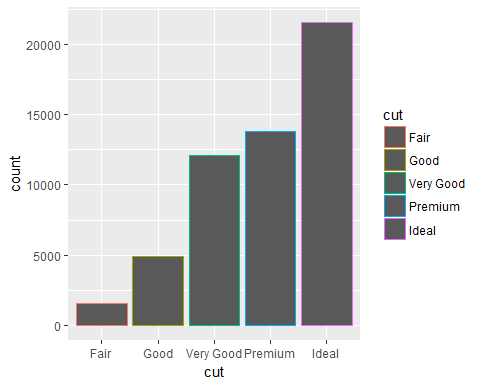
#display a bar chart of proportion, rather than count  
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x = cut, y= ..prop.., group = 1))



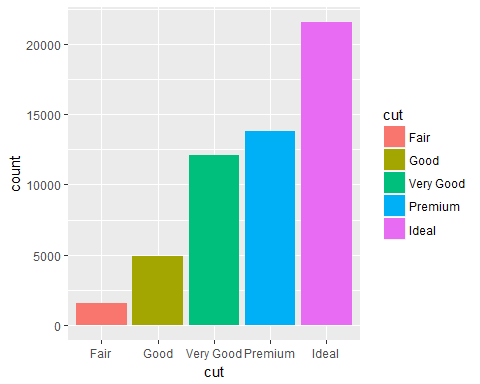
ggplot(data = diamonds) +  
 stat\_summary(  
 mapping = aes(x=cut, y = depth),  
 fun.ymin = min,  
 fun.ymax = max,  
 fun.y = median  
 )



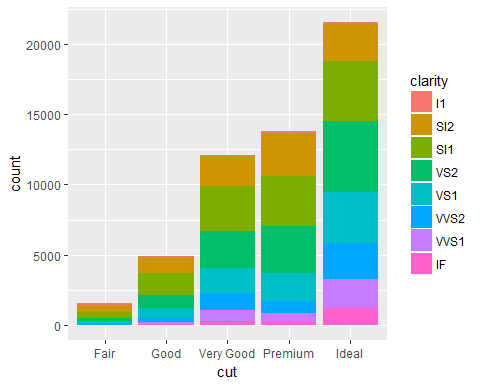
#Position adjustments  
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x=cut, color = cut))



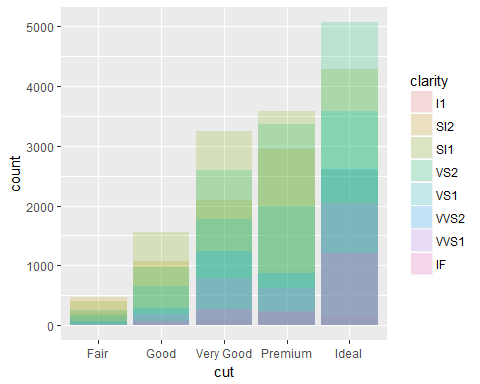
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x=cut, fill = cut))



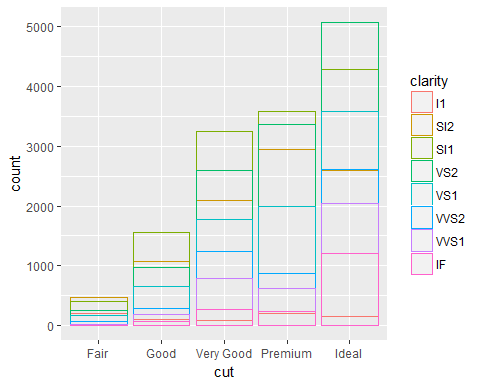
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x=cut, fill = clarity))



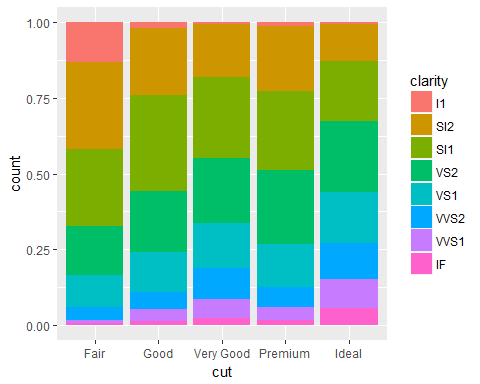
#POSITION = IDENTITY  
  
ggplot(data = diamonds, mapping = aes(x= cut, fill = clarity)) +  
 geom\_bar(alpha = 1/5, position = "identity")



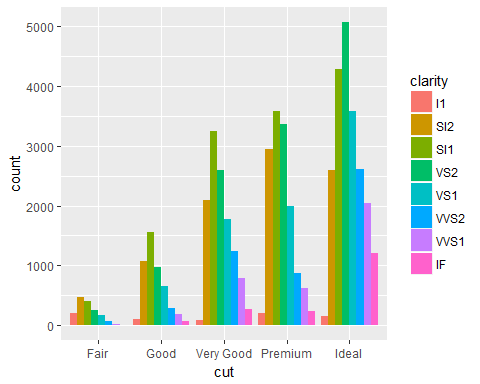
ggplot(data = diamonds, mapping = aes(x=cut, color = clarity)) +  
 geom\_bar(fill = NA, position = "identity")



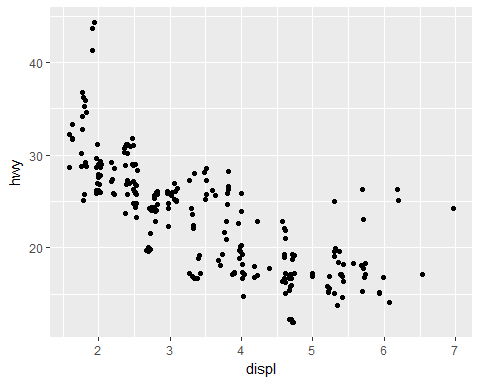
#Position = "fill"  
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x = cut, fill = clarity), position = "fill")



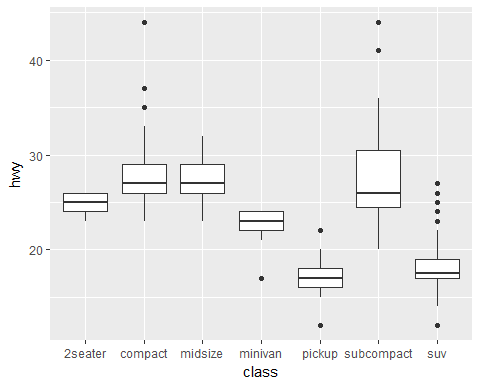
#Position = "dodge"  
  
ggplot(data = diamonds) +  
 geom\_bar(mapping = aes(x=cut, fill = clarity), position = "dodge")



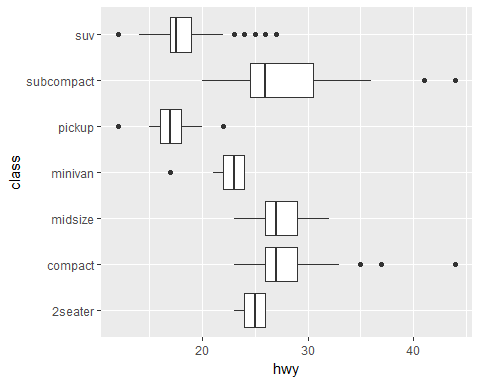
# Fix overplotting  
ggplot(data = mpg) +  
 geom\_point(mapping = aes(x=displ, y= hwy), position = "jitter")



#Coordinate Systems  
ggplot(data = mpg, mapping = aes(x=class, y = hwy)) +  
 geom\_boxplot()



ggplot(data = mpg, mapping = aes(x=class, y = hwy)) +  
 geom\_boxplot() +  
 coord\_flip()

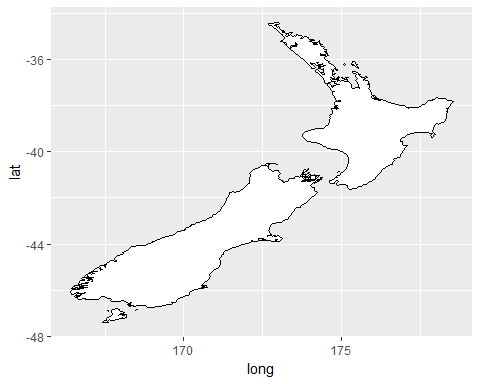


#coord\_quickmap()  
nz <- map\_data("nz")

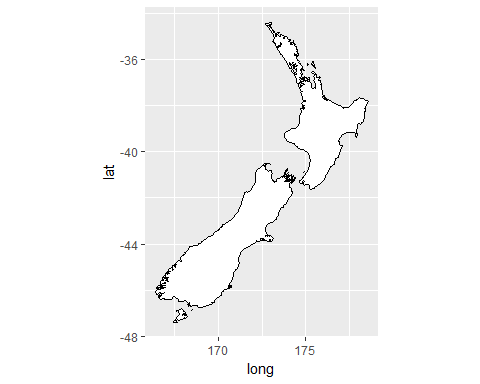
##   
## Attaching package: 'maps'

## The following object is masked from 'package:purrr':  
##   
## map

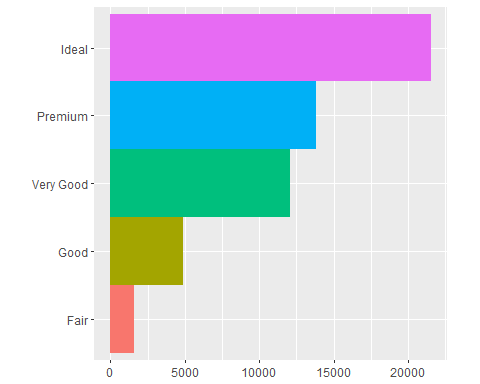
ggplot(data = nz, aes(long, lat, group = group)) +  
 geom\_polygon(fill = "white", color = "black")



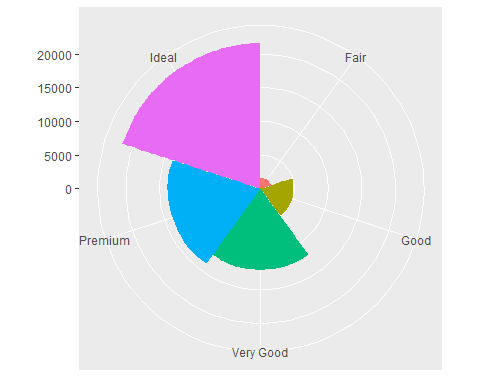
ggplot(data = nz, aes(long,lat,group = group)) +  
 geom\_polygon(fill = "white", color = "black") +  
 coord\_quickmap()



#coord\_polar()  
  
bar <- ggplot(data = diamonds) +  
 geom\_bar(  
 mapping = aes(x=cut, fill=cut),  
 show.legend = FALSE,  
 width = 1  
 ) +  
 theme(aspect.ratio = 1) +  
 labs(x=NULL, y = NULL)  
  
bar + coord\_flip()



bar + coord\_polar()



#NEW FULL GGPLOT TEMPLATE  
  
# ggplot(data = <DATA>) +   
# <GEOM\_FUNCTION>(  
# mapping = aes(<MAPPINGS>),  
# stat = <STAT>,   
# position = <POSITION>  
# ) +  
# <COORDINATE\_FUNCTION> +  
# <FACET\_FUNCTION>