# John R. Cobb

## MIS 502

### Module 1 Assignemnt 3

library(tidyverse)

## -- Attaching packages ----------------------------- tidyverse 1.2.1 --

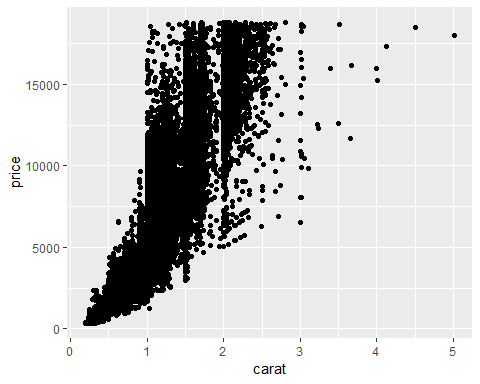
## v ggplot2 3.1.0 v purrr 0.3.2   
## v tibble 2.1.1 v dplyr 0.8.0.1  
## v tidyr 0.8.3 v stringr 1.4.0   
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts -------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

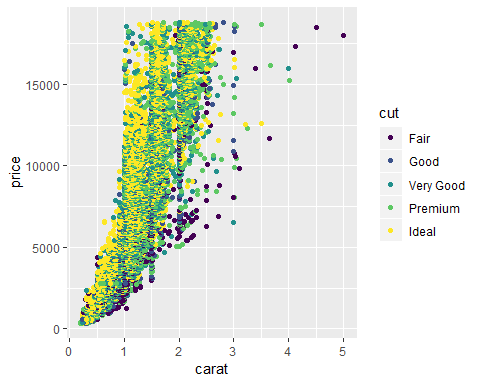
diamonds

## # A tibble: 53,940 x 10  
## carat cut color clarity depth table price x y z  
## <dbl> <ord> <ord> <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>  
## 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 0.290 Premium I VS2 62.4 58 334 4.2 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  
## 7 0.24 Very Good I VVS1 62.3 57 336 3.95 3.98 2.47  
## 8 0.26 Very Good H SI1 61.9 55 337 4.07 4.11 2.53  
## 9 0.22 Fair E VS2 65.1 61 337 3.87 3.78 2.49  
## 10 0.23 Very Good H VS1 59.4 61 338 4 4.05 2.39  
## # ... with 53,930 more rows

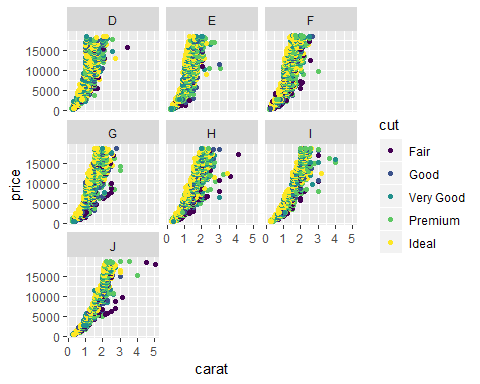
diamondsdata<-diamonds  
ggplot(diamondsdata, aes(x = carat, y = price)) +  
 geom\_point()

 #looking at the graph there is a dirrect correlation between carat size and price.

ggplot(diamondsdata, aes(x = carat, y = price, color =cut)) +  
 geom\_point()

 #looking at the grph, it can be seen that the ideal cut diamonds (yellow) are at the top of the price range for each carat size

ggplot(diamondsdata, aes(x = carat, y = price, color=cut)) +  
 geom\_point()+  
 facet\_wrap(~color)

 #for the seven different color facets, it can be seen that each has its own distribution while also keeping relatively true the the correlation between cut, carat, and price. It can, however, be seen that each color hs its own tendency, with d-g having steep exponential curves, h-j have less steep curves and seem to have a wider variety of posibilities in the cut, carat, price gambit.

library(readr)  
InventoryData <- read\_csv("~/MIS 502/Chapter 1/InventoryData.csv")

## Parsed with column specification:  
## cols(  
## `Item SKU` = col\_double(),  
## Store = col\_double(),  
## Supplier = col\_character(),  
## `Cost per Unit ($)` = col\_double(),  
## `On Hand` = col\_double(),  
## `Annual Demand` = col\_double()  
## )

Inventory<-InventoryData

InventoryA<- Inventory %>%  
 filter(Supplier == 'A')

# there are 3695 observations (rows)

InventoryA = mutate(InventoryA, OnHandRatio = `On Hand` / `Annual Demand`)

# The previous code divides ove variable by another in order to get a useful ratio, in this case the on hand ratio of products, this is just representing the percentage of product on hand currently out of the total demand for a year.

avg\_cost <- InventoryA %>%  
 group\_by(`Item SKU`) %>%  
 summarize(SKUAvgCost = mean(`Cost per Unit ($)`))

# I would say all of the concepts were just refreshers, the one that took me the most time was the very last problem because I was overthinking the programming rhetoric