# Je’Kolby Worthy

## BAN 502 Predictive Analytics

### R and RStudio Refresher Assignment

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.0 ──

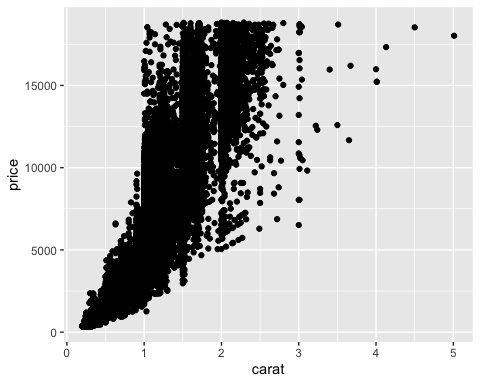
## ✓ ggplot2 3.3.3 ✓ purrr 0.3.4  
## ✓ tibble 3.0.5 ✓ dplyr 1.0.3  
## ✓ tidyr 1.1.2 ✓ stringr 1.4.0  
## ✓ readr 1.4.0 ✓ forcats 0.5.0

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

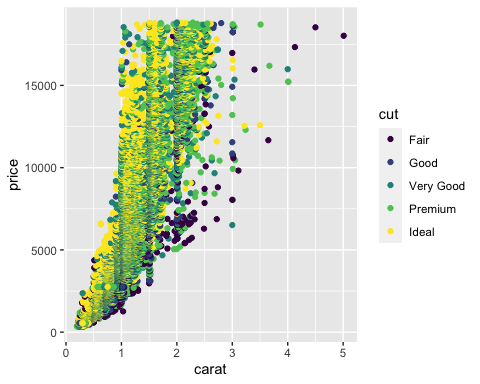
diamonddata = diamonds

There are 10 columns and 53,940 rows in the “diamonddata” data set.

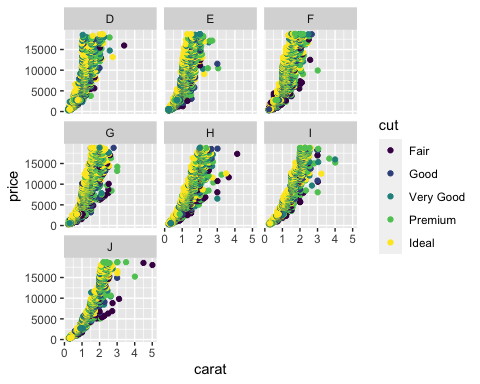
ggplot(diamonddata, aes(x= carat, y=price)) + geom\_point()

 The relationship between the two variables show that the bigger the carat size, the more expensive the diamond is.

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

 The relationship between the four variables show that you can purchase a fair piece of diamond that is way bigger in carat size for the same price you can purchase an decent size cut of 2 carat diamonds. The better the cut is in the diamond, the smaller the size will be due to the expensive prices. Terrible cut diamonds can be purchased at a much larger carat size and usually would be a lower price compared to the more better looking cuts.

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

 While looking at this graph of data, the nicer the color make the diamond more expensive. Also, the better the cut, the more expensive the diamond will be. You can purchase a larger carat that is not much of a nice color for a cheaper price than buying a normal and nice looking colored diamond.

library(readr)  
inventory <- read\_csv("InventoryData.csv")

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

inventoryA <- inventory %>% filter(Supplier == "SupplierA")

0 rows are in this data frame.

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

This line of code shows a new column titled "On Hand Ratio and is formulated by dividing the On Hand variable over the Annual Demand variable.

avg\_cost <- inventory %>% group\_by('Item SKU') %>% summarise("SKUAvgCost" = mean("Cost per Unit ($)"), .groups ='drop')

## Warning in mean.default("Cost per Unit ($)"): argument is not numeric or  
## logical: returning NA

Given my previous course experience with R/RStudio, I have not found any topic/concept to be challenging. So far, everything has been very straight forward with the lecture videos that were given in previous course and also reading through the book how to navigate through R and R Studio.