# Alex Fields

## BAN 502 Module 1 - Assignment 3

#install.packages("tidyverse") #only run this line if you have not installed the Tidyverse package previously  
library(tidyverse)

diamonddata = diamonds  
print(paste("Number of Rows:", nrow(diamonddata)))

## [1] "Number of Rows: 53940"

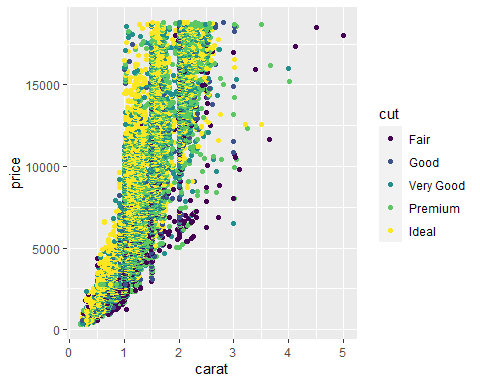
print(paste("Number of Columns:", ncol(diamonddata)))

## [1] "Number of Columns: 10"

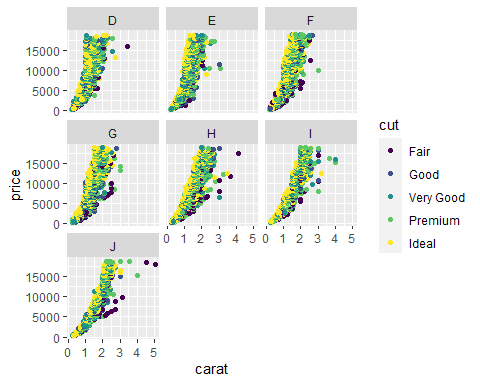
ggplot(data = diamonddata, aes(x=carat, y=price)) + geom\_point()

 The relationship between Carat and Price seem to be very correlated. One would suspect that the higher the price of the diamond, the higher quality/quantity of Carats there would be.

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

 The relationship between Carat, Price and Cut seem to be very correlated, like the before. Its easy to see that the price of a Diamond is related to the Carat and Cut of Diamond. It seems that Cut varies from range and Carat so the correlation of Cut is presumably close to 0.

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

 The relationship between Carat, Price, Cut and Color seem to be very correlated, like the before. It seems like before, with Cut, Color varies from range and Carat so the correlation of Color is presumably close to 0. All Colors seem the show the same graph (give or take a few outliers).

InventoryData <- read\_csv("InventoryData.csv")  
summary(InventoryData)

## Item SKU Store Supplier Cost per Unit ($)  
## Min. : 6 Min. : 1611 Length:13561 Min. : 0.0   
## 1st Qu.:2537 1st Qu.: 3480 Class :character 1st Qu.: 137.0   
## Median :4997 Median :20109 Mode :character Median : 377.5   
## Mean :5025 Mean :26675 Mean : 504.4   
## 3rd Qu.:7602 3rd Qu.:31779 3rd Qu.: 775.5   
## Max. :9998 Max. :80212 Max. :1982.3   
## On Hand Annual Demand   
## Min. : 0.0 Min. : 0.0   
## 1st Qu.: 50.0 1st Qu.: 483.0   
## Median :101.0 Median : 965.0   
## Mean :100.5 Mean : 966.2   
## 3rd Qu.:151.0 3rd Qu.:1448.0   
## Max. :200.0 Max. :2150.0

inventoryA =   
 InventoryData %>% filter(Supplier == 'A')  
  
print(paste("Number of Rows in 'InventoryA' Dataframe:", nrow(inventoryA)))

## [1] "Number of Rows in 'InventoryA' Dataframe: 3695"

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

The code above creates a new variable called ‘OnHandRatio’ and this calculates the ratio of SKU’s on hand vs the demand of said SKU.

avg\_cost =   
 inventoryA %>%  
 group\_by(`Item SKU`) %>%  
 summarise(SKUAvgCost = median(`Cost per Unit ($)`))

print(paste("Q: Given your previous course experience with R/RStudio, what topics/concepts did you find to be  
most challenging?  
"))

## [1] "Q: Given your previous course experience with R/RStudio, what topics/concepts did you find to be\nmost challenging?\n"

print(paste("A: Task 9 seemed to be a little bit of a challenge but overall it was not too difficult."))

## [1] "A: Task 9 seemed to be a little bit of a challenge but overall it was not too difficult."