## Refresher Assignment

### Adam Samstag

#### Task 1

#the install.packages is commented out as to not affect the knit  
#install.packages("tidyverse")  
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 3.5.2

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

## v ggplot2 3.1.0 v purrr 0.2.5  
## v tibble 1.4.2 v dplyr 0.7.7  
## v tidyr 0.8.2 v stringr 1.3.1  
## v readr 1.1.1 v forcats 0.3.0

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

#### Task 2

diamonddata = diamonds  
nrow(diamonds)

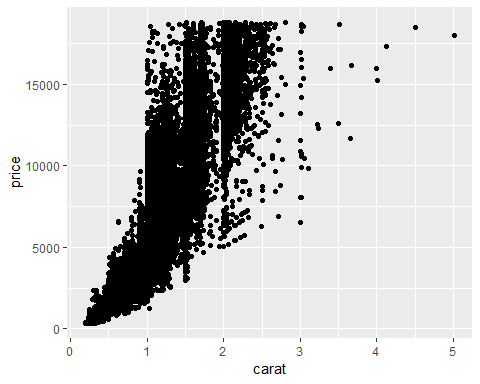
## [1] 53940

ncol(diamonds)

## [1] 10

#### Task 3

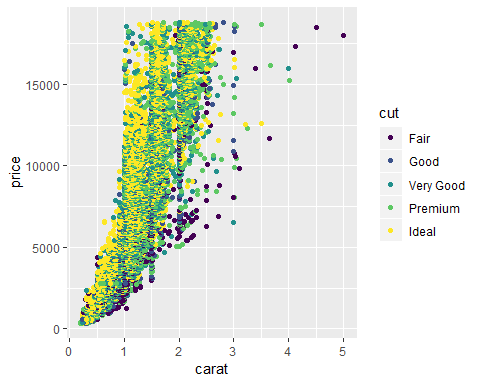
ggplot(diamonds, aes(x=carat, y=price))+  
geom\_point()



#You can see that a majority of the points show that as the carat size increases so does the price.

#### Task 4

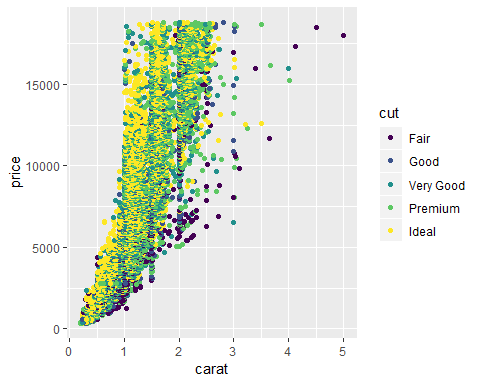
ggplot(diamonds, aes(x=carat, y=price, color=cut))+  
 geom\_point()



#In this plot you can see that the highest priced diamonds have an ideal or premium cut and you can see that a majority of the ideal and premium cuts diamonds are between 1 and 2 carats.

#### Task 5

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



#By adding facet to this plot you can see that many more ideal cut diamonds show up than there was before.

#### Task 6

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

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

#### Task 7

InventoryA = filter(InventoryData, Supplier == "A")  
#There are 3695 rows

#### Task 8

InventoryA = mutate(InventoryA, OnHandRatio = `On Hand` / `Annual Demand`)  
  
#This code adds a new variable or column to the data set.

#### Task 9

Avg\_Cost = InventoryA %>% group\_by(`Item SKU`) %>% summarise(SKUAvgCost = mean(`Cost per Unit ($)`))

#### Task 10

#Some areas where I start to have issues is when you start adding a bunch of different features to a plot, most things in plots are simple but some can be confusing if the question is not descriptive enough.