## Task 1

#install.packages("tidyverse")   
library(tidyverse)

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

## ✓ ggplot2 3.3.2 ✓ purrr 0.3.4  
## ✓ tibble 3.0.4 ✓ dplyr 1.0.2  
## ✓ tidyr 1.1.2 ✓ stringr 1.4.0  
## ✓ readr 1.4.0 ✓ forcats 0.5.0

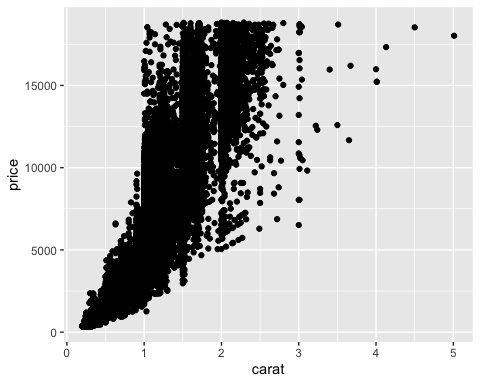
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

## Task 2

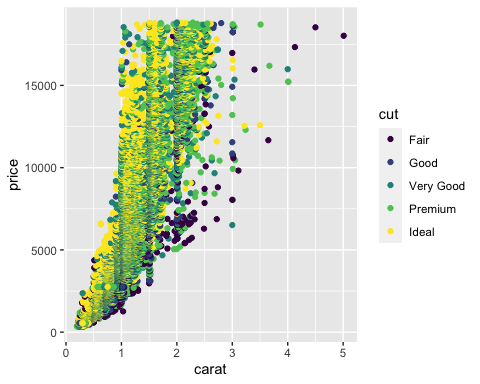
diamonddata = diamonds

## Task 3

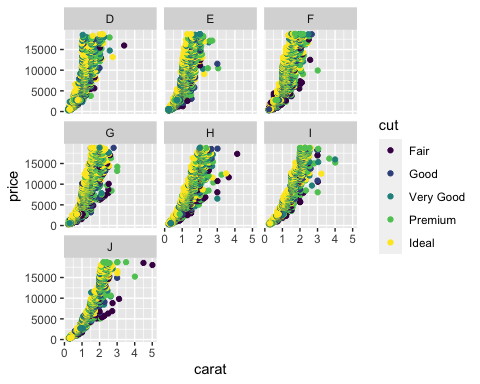
ggplot(diamonddata, aes(x=carat, y=price)) +  
 geom\_point()

 Price and Carat are positively correlated. As carat increases, price increases.  
## Task 4

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

 Carat has a greater impact on price than cut does. There is an obvious positive correlation between carat and price, but no correlation between cut, price, and carat.  
## Task 5

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

 ## Task 6

inventory<- read\_csv("~/Documents/MBA\_Predictive/Week 1/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()  
## )

## Task 7

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

There are 3695 rows in the data frame inventory A.

## Task 8

inventoryA = mutate(inventoryA, OnHandRatio = On Hand / Annual Demand) The code above creates a new column called OnHandRatio. The column contains the product of “On Hand” divided by “Annual Demand.”  
## Task 9

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

## `summarise()` regrouping output by 'Item SKU' (override with `.groups` argument)

## Task 10

I thought this exercise was a great refresher to R and Rstudio. Focusing on Data visualization in Tidyverse is a great way to start the course. Tidyverse is very straight forward and easy to use. Remembering the syntax is the most challenging part for me. The little things like double spacing after a heading trip me up from time to time, but those are easy to fix.