# Module 3 - Assignment 2

## Agyabeng Prince

### Exploratory Data Analysis

library("tidyverse")

## Warning: package 'tidyverse' was built under R version 3.4.4

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

## v ggplot2 2.2.1 v purrr 0.2.4  
## v tibble 1.4.1 v dplyr 0.7.5  
## v tidyr 0.8.0 v stringr 1.3.1  
## v readr 1.1.1 v forcats 0.2.0

## Warning: package 'dplyr' was built under R version 3.4.4

## Warning: package 'stringr' was built under R version 3.4.4

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

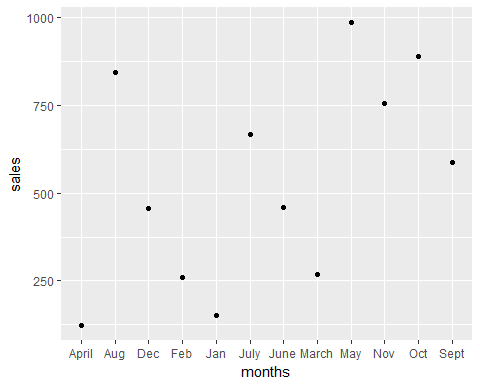
## Diamond Color and Price

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



## What do you notice from the scatterplot as the carat size increases? As carat size increases, prices of diamond increases too.  
## From the scatterplot, what carats are most represented within the diamonds dataset? 0.30

Yearly\_sales <- data.frame(months = c("Jan", "Feb", "March", "April", "May", "June", "July", "Aug", "Sept", "Oct", "Nov", "Dec"), sales = c(150.25, 258.54, 268.55, 122.52, 987, 458.82, 667.23, 845.54, 586.78, 888.58, 756.12, 456.84))  
  
ggplot(Yearly\_sales,aes(months,sales,colors)) +  
 geom\_point()



ggplot(diamonds,aes(carat,price,colors()))+  
 geom\_point()



dsample <- diamonds[sample(nrow(diamonds), 100), ]  
ggplot(diamonds,aes(carat,price,colors()))+  
 geom\_point()+  
 geom\_smooth()

## `geom\_smooth()` using method = 'gam'

