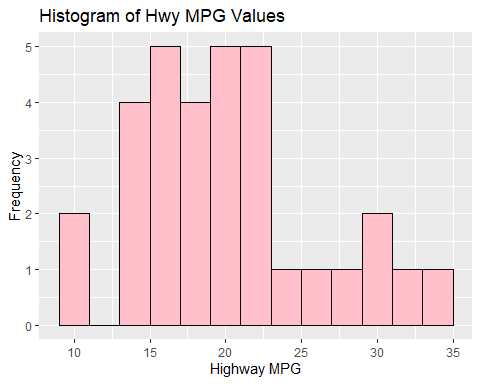
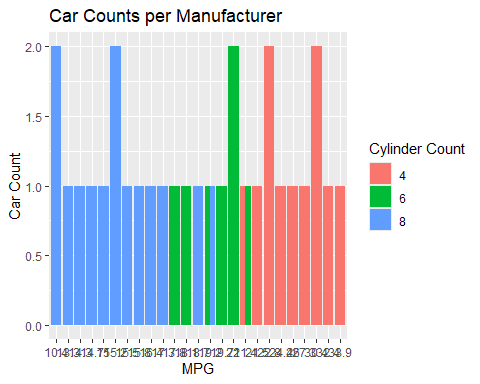
ggplot\_question

2024-04-17

### 1 ->> Use "mtcars" and "txhousing" dataset.  
library(ggplot2)  
data(mtcars)  
# 1. Histogram of hwy mpg values  
ggplot(mtcars, aes(x = mpg)) +  
 geom\_histogram(binwidth = 2, fill = "pink", color = "black") +  
 labs(title = "Histogram of Hwy MPG Values", x = "Highway MPG", y = "Frequency")

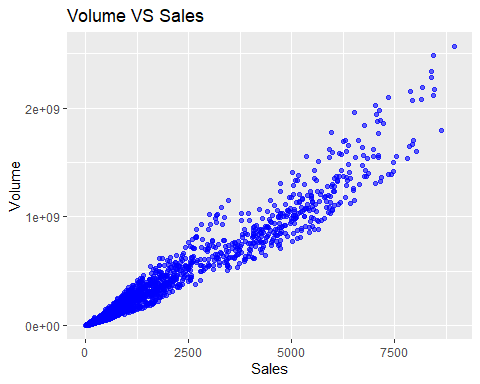


# 2. Barplot of car counts per MPG with color fill defined by cyl count  
ggplot(mtcars, aes(x = factor(mpg))) +  
 geom\_bar(aes(fill = factor(cyl)), position = "dodge") +  
 labs(title = "Car Counts per Manufacturer", x = "MPG", y = "Car Count", fill = "Cylinder Count")



data(txhousing)  
  
# 3. Create a scatterplot of volume versus sales  
ggplot(txhousing, aes(x = sales, y = volume)) +  
 geom\_point(alpha = 0.6, color = "blue") +  
 labs(title = "Volume VS Sales", x = "Sales", y = "Volume")

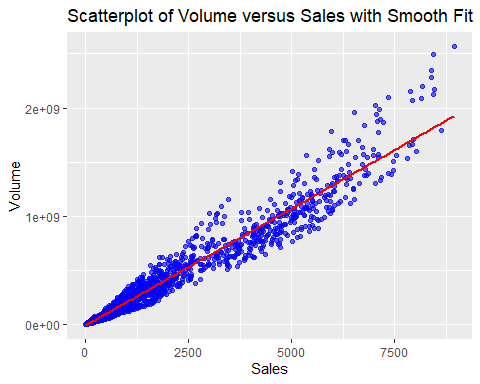
## Warning: Removed 568 rows containing missing values or values outside the scale range  
## (`geom\_point()`).



# 4. Add a smooth fit line to the scatterplot  
ggplot(txhousing, aes(x = sales, y = volume)) +  
 geom\_point(alpha = 0.6, color = "blue") +  
 geom\_smooth(method = "lm", se = FALSE, color = "red") +  
 labs(title = "Scatterplot of Volume versus Sales with Smooth Fit", x = "Sales", y = "Volume")

## `geom\_smooth()` using formula = 'y ~ x'

## Warning: Removed 568 rows containing non-finite outside the scale range  
## (`stat\_smooth()`).  
## Removed 568 rows containing missing values or values outside the scale range  
## (`geom\_point()`).

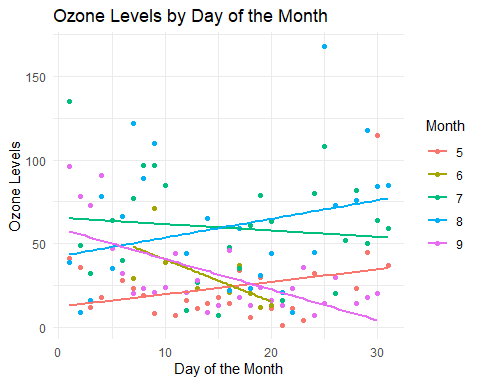


### 2 -> Use "airquality" dataset.   
 # 1. The x-axis represents the Day of the month.  
 # 2. The y-axis represents the Ozone levels.  
 # 3. Color the points by Month.  
 # 4. Add a smooth trend line to the plot.  
library(ggplot2)  
data(airquality)  
ggplot(airquality, aes(x = Day, y = Ozone, color = factor(Month))) +  
 geom\_point() + # Scatterplot  
 geom\_smooth(method = "lm", se = FALSE) + # Add a linear trend line  
 labs(title = "Ozone Levels by Day of the Month", x = "Day of the Month", y = "Ozone Levels", color = "Month") +  
 theme\_minimal() # Use minimal theme

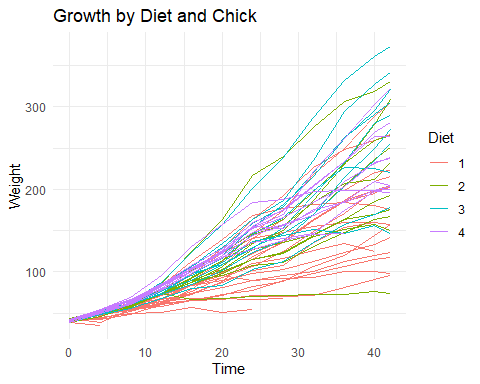
## `geom\_smooth()` using formula = 'y ~ x'

## Warning: Removed 37 rows containing non-finite outside the scale range  
## (`stat\_smooth()`).

## Warning: Removed 37 rows containing missing values or values outside the scale range  
## (`geom\_point()`).



### 3 ->> USe ChickWeight dataset.   
 # 1. Plots Time on the x-axis and weight on the y-axis.  
 # 2. Uses different colors to represent each Diet.  
 # 3. Groups the lines by Chick to show individual growth trajectories within each diet.  
library(ggplot2)  
ChickWeight <- read.csv("C:/Users/ChickWeight.csv") #Add your file path of csv file.  
ggplot(ChickWeight, aes(x = Time, y = weight, color = factor(Diet), group = Chick)) +  
 geom\_line() + # Line plot  
 labs(title = "Growth by Diet and Chick", x = "Time", y = "Weight", color = "Diet") +  
 theme\_minimal() # Use minimal theme



library(ggplot2)  
ChickWeight <- read.csv("C:/Users/ChickWeight.csv") #Add your file path of csv file.  
  
### 4 ->> Create a histogram to show the distribution of carat sizes  
ggplot(ChickWeight, aes(x = weight)) +  
 geom\_histogram(binwidth = 2, fill = "lightgreen", color = "black") +  
 labs(title = "Distribution of Carat Sizes", x = "Carat Size", y = "Frequency") +  
 theme\_minimal() +  
 xlim(0,75) # Limit x-axis to diamonds up to 3 carats for better detail

## Warning: Removed 376 rows containing non-finite outside the scale range  
## (`stat\_bin()`).

## Warning: Removed 1 row containing missing values or values outside the scale range  
## (`geom\_bar()`).

