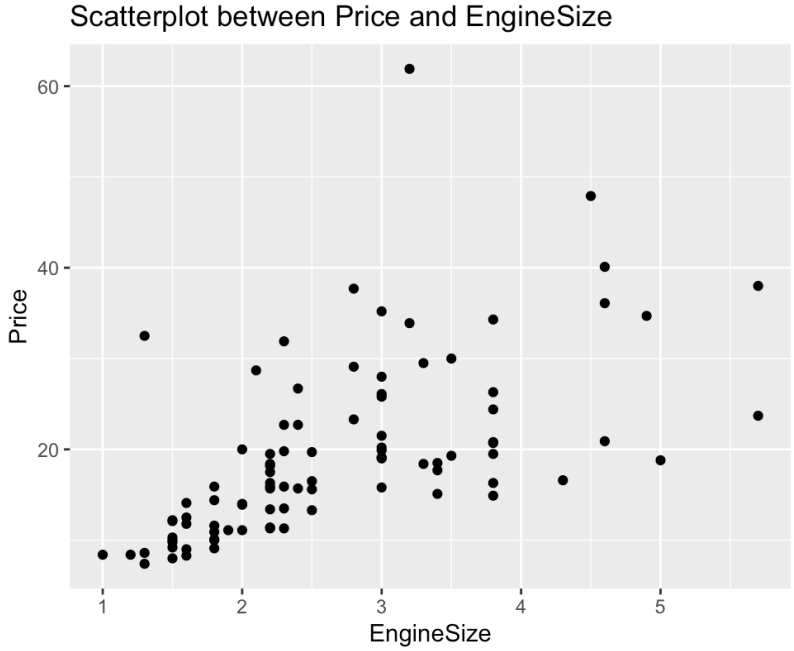
Name: Komal Patil
Course: Foundation of AI

ASSIGNMENT 3

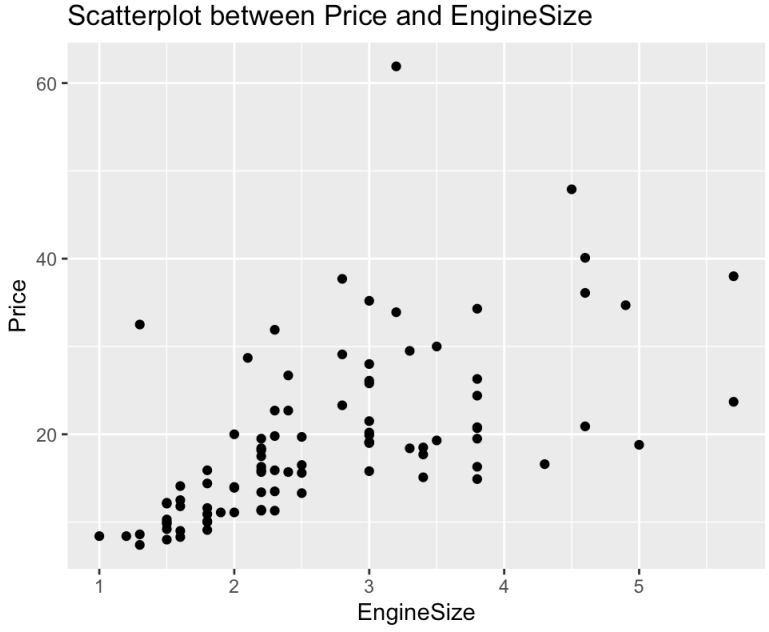
TASK 01

a) Use q plot to create a scatter plot with Price on the y-axis and Engine Size on the x-axis.

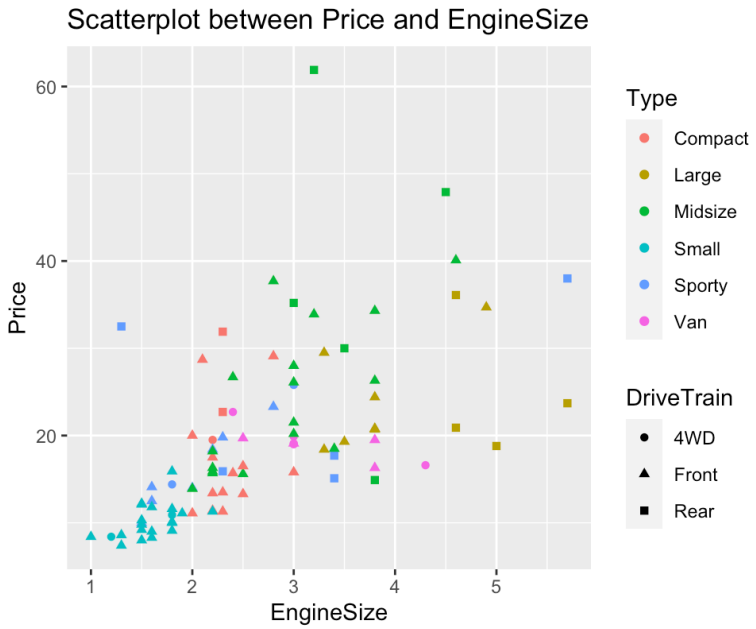


a)After conducting the analysis, it is evident that there exists a moderately positive linear relationship between Engine Size and Price. The correlation between these variables suggests that, on average, as Engine Size increases, there is a corresponding increase in Price.

b) Repeat part a) using the ggplot function and geom_point() layer.



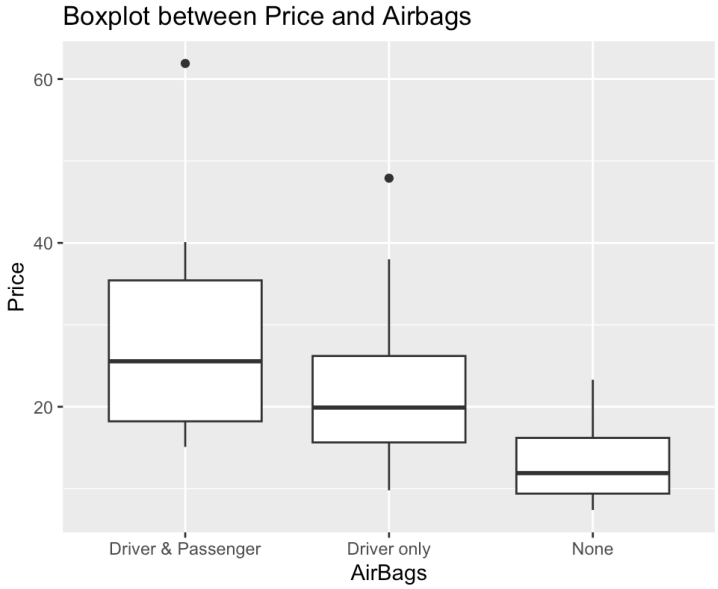
c) Repeat part b), but this time specifying that the color mapping should depend on Type and the shape mapping should depend on DriveTrain



a) Many small-type vehicles do tend to fall into the lower price and engine size category (between 1 and 2). Even though they are a little more expensive than small-type vehicles, compact-type vehicles also fall into the lower price and engine size range. It's interesting to note that compact-type automobiles have higher price ranges and larger engines than small-type vehicles.

b) In comparison to other drive train types, a sizable fraction of front-wheel-drive vehicles are seen in the lower price and engine size range.

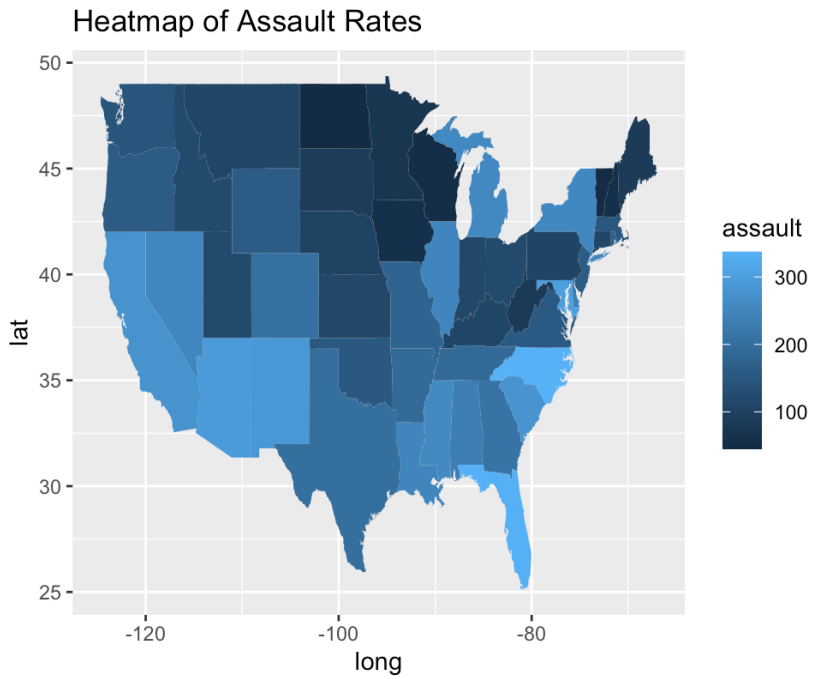
d) Construct box plots showing Price on the y-axis and Airbags on the x-axis. (Hint: boxplot is a valid ggplot2 geometry) a. Do you observe any association between Airbag type and Price? Explain.



a) Yes, vehicles with driver and passenger airbags cost more on average than those with a driver-only airbag. Airbags are available at the lowest median price for non-vehicles.

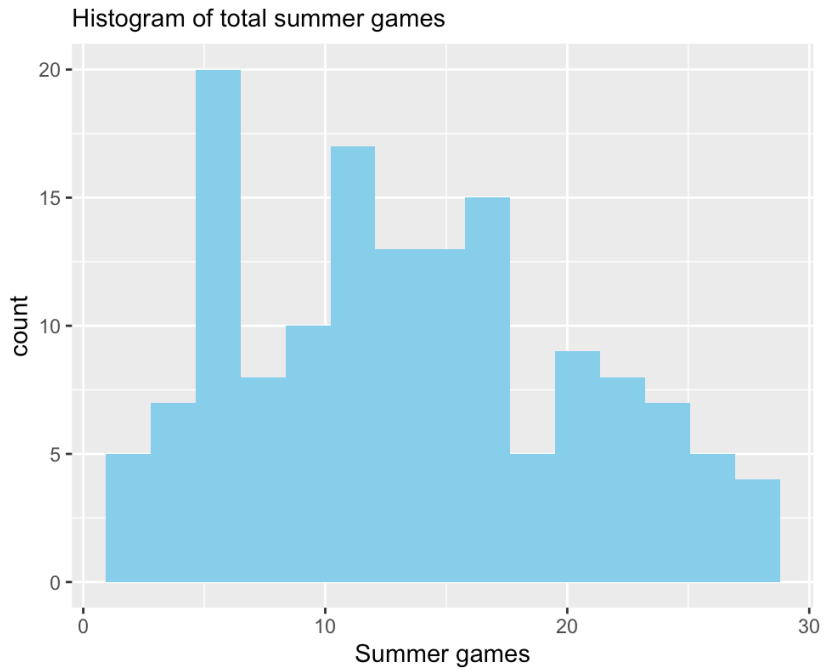
TASK 02:

Plotting a map
By using USArrest data from maps
package library(maps)
Produce a heatmap of assault rates.

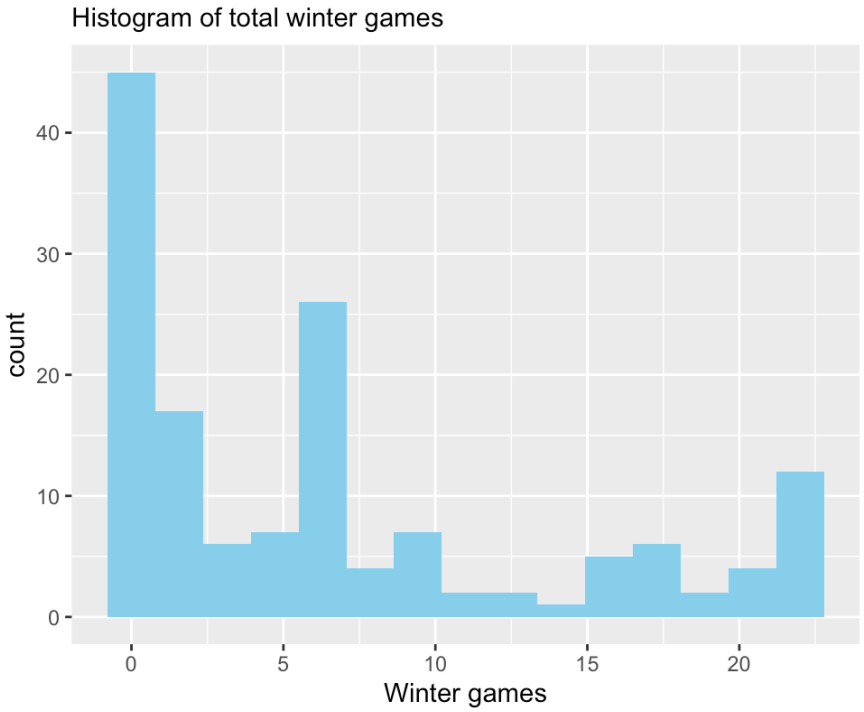


TASK 03

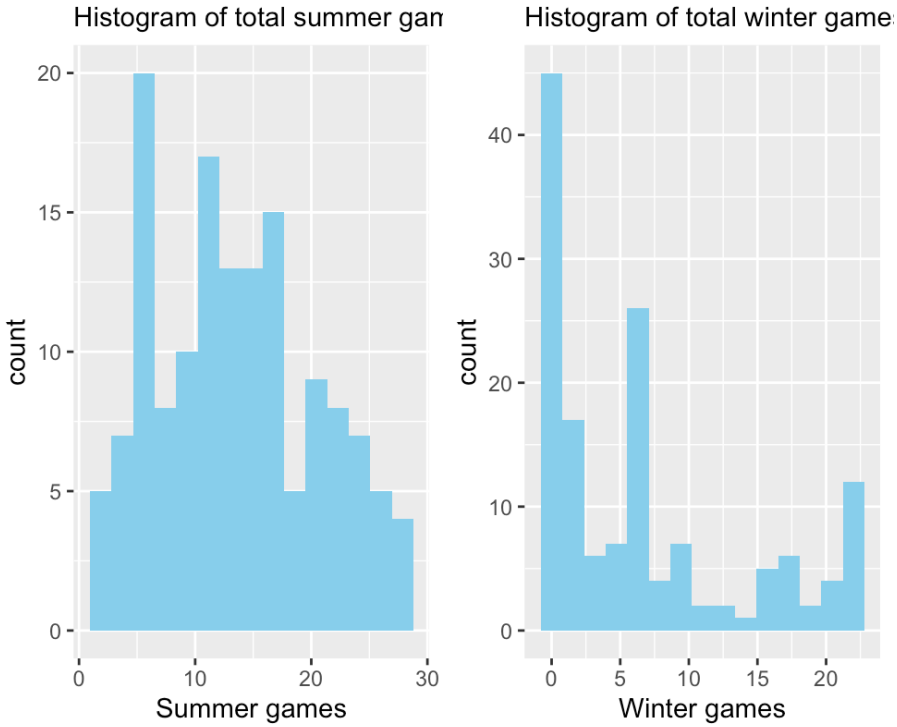
a) do histogram of summer games(total)



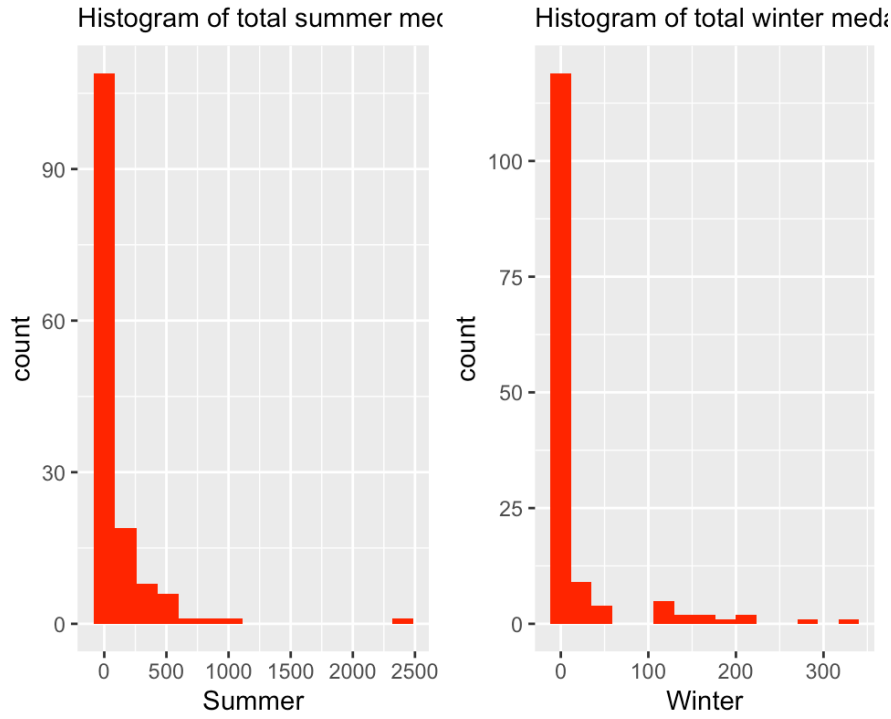
b) do histogram of winter games (total)



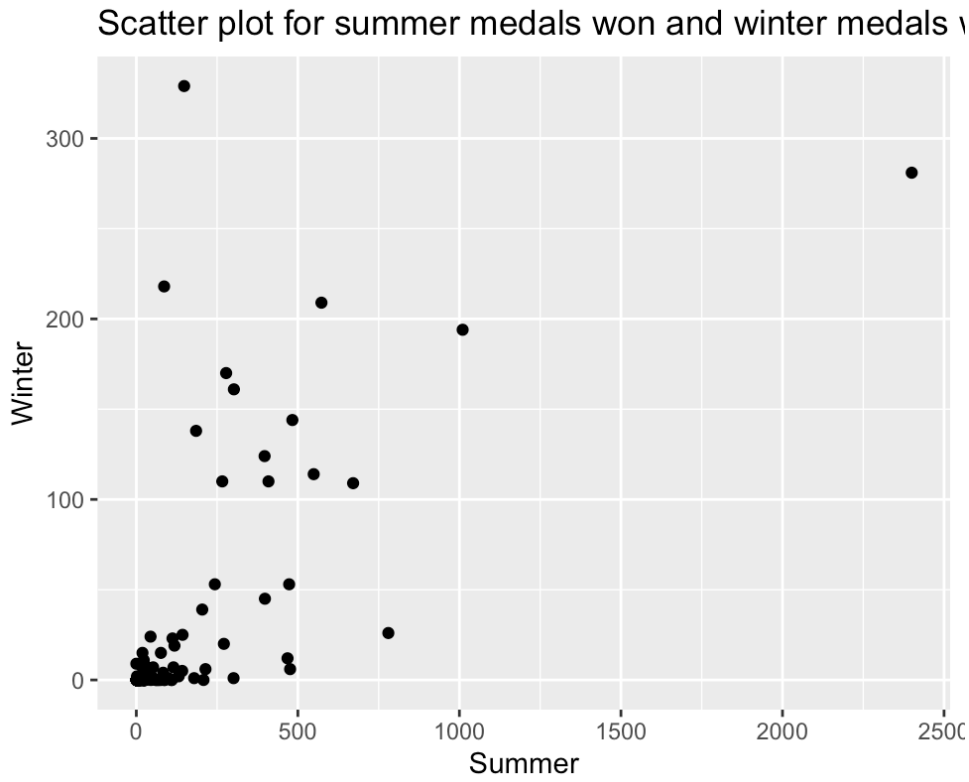
c) put above two histograms on one page



d) do two histograms on one page: total summer, total winter medals won

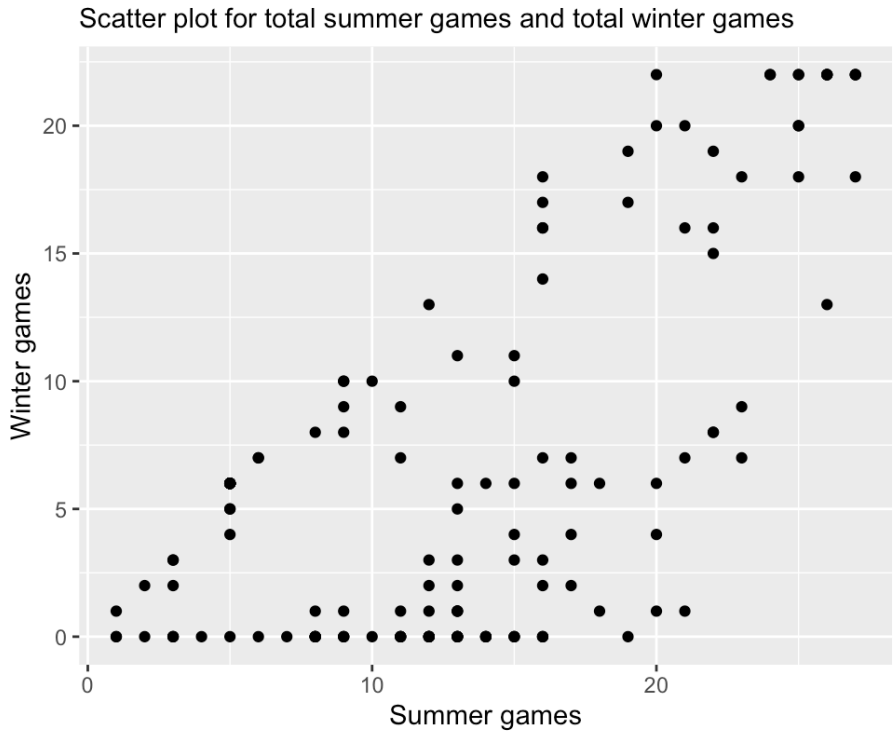


e) is there a correlation between the number of medals given out in winter and summer? Use ggplot2, scatter plot, and add additional dimension into this plot.

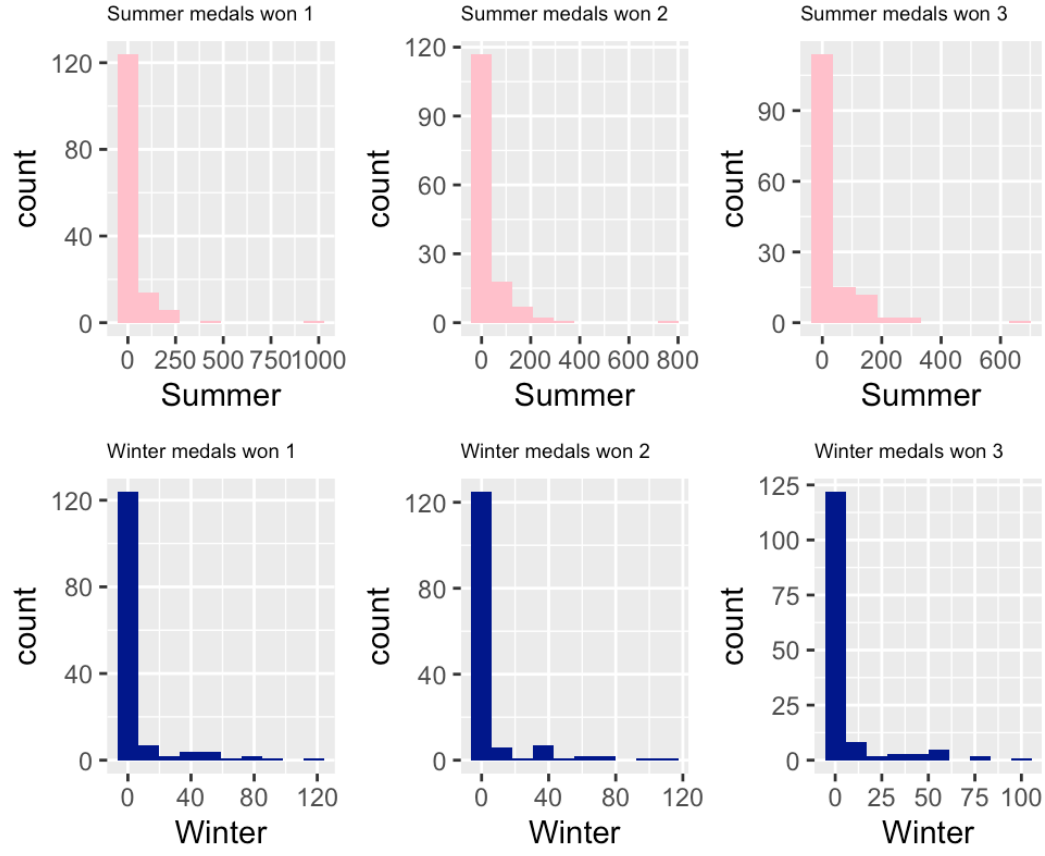


a) There is a moderate positive linear relationship between summer medals won and winter medals won.

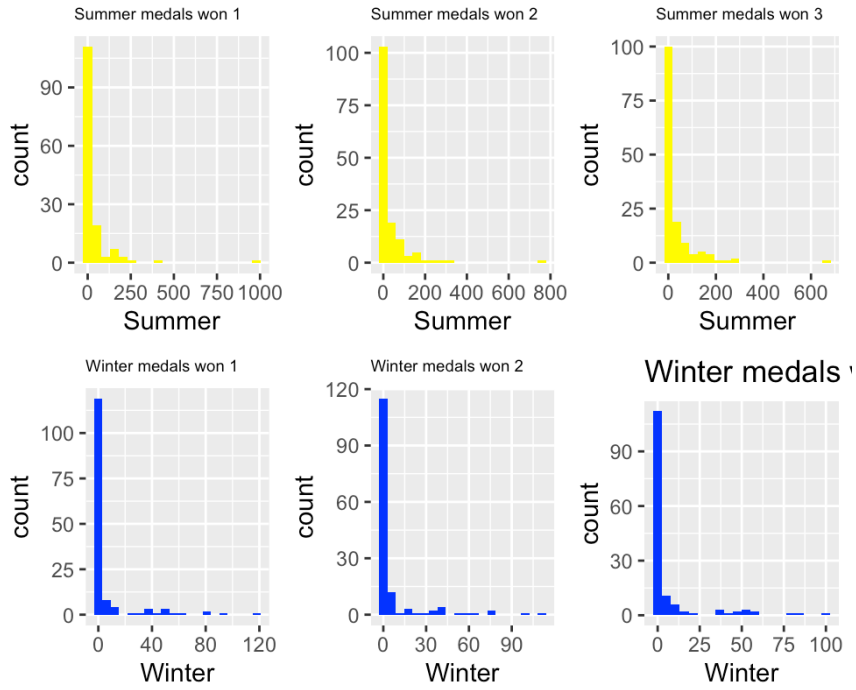
f) How about the number of games each country competes in. Is there a correlation between winter and summer? Use ggplot 2 scatter plot and add additional dimension into this plot.



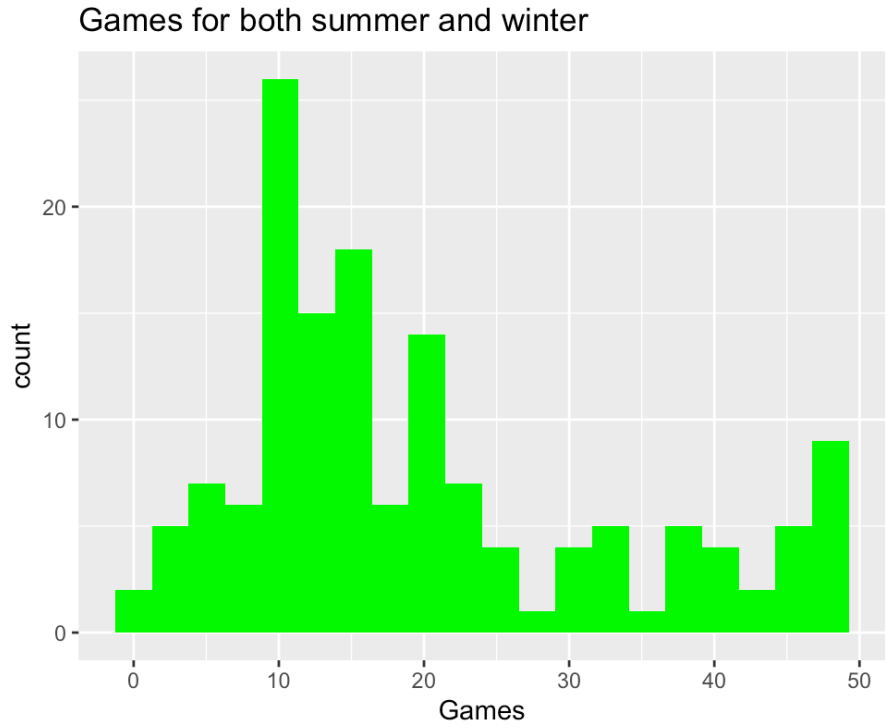
g) look at distribution of each of the types of medals, by season (6 histograms on one page)



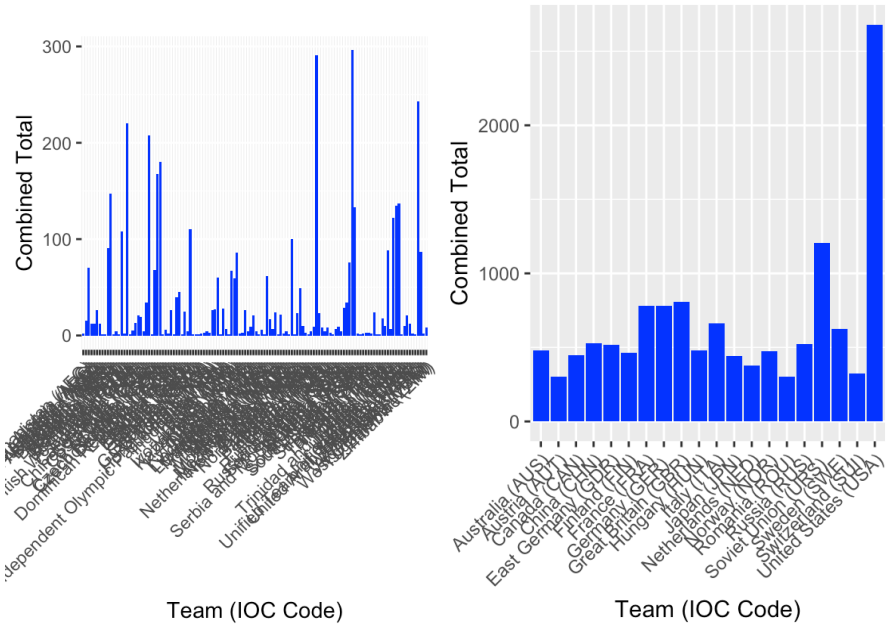
h) redo g) with different number of bins (10 instead of 20)



i) (extra) explore data on your own



i)b)



#From the above observation we can see that most of the countries have earned the medals less than 300.