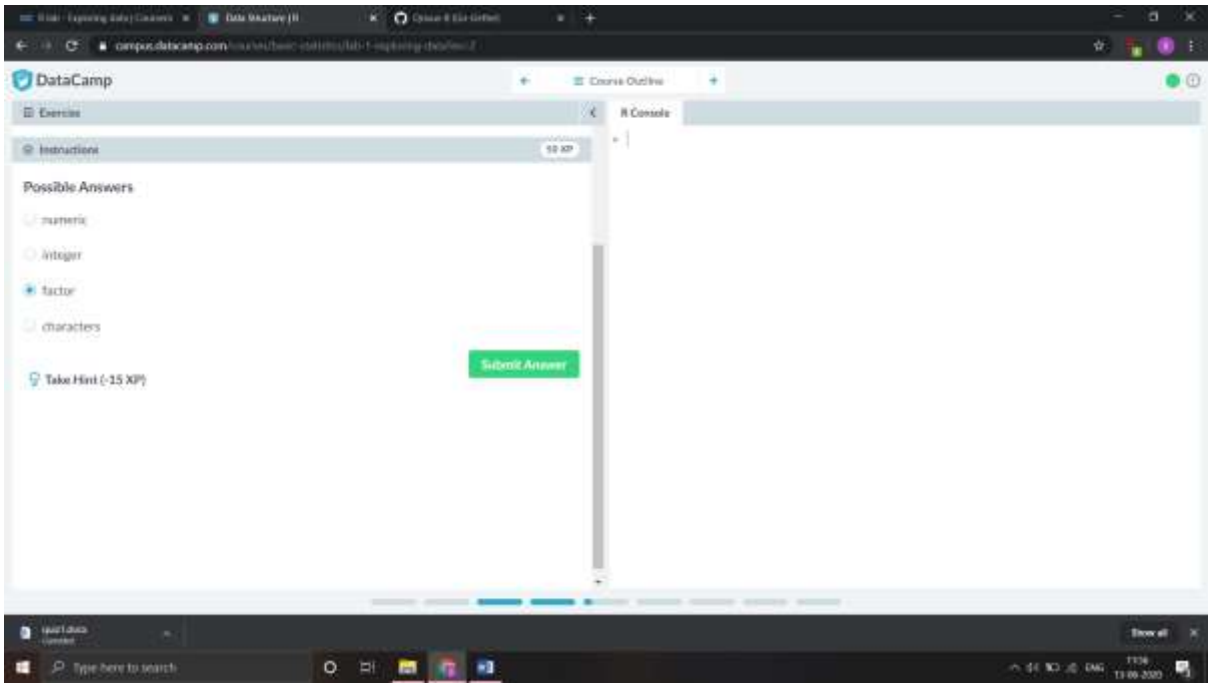
#Use the dim() function on mtcars

dim(mtcars)



Look at the levels of the variable am

levels(mtcars$am)

#Assign the value of mtcars to the new variable mtcars2

mtcars2 <- mtcars

mtcars2

#Assign the label "high" to mpgcategory where mpg is greater than or equal to 20

mtcars2$mpgcategory[mtcars2$mpg >= 20] <- "high"

#Assign the label "low" to mpgcategory where mpg is less than 20

mtcars2$mpgcategory[mtcars2$mpg < 20] <- "low"

#Assign mpgcategory as factor to mpgfactor

mtcars2$mpgfactor <- as.factor(mtcars2$mpgcategory)

#How many of the cars have a manual transmission?

table(mtcars$am)

mtcars

13

What percentage of cars have 3 or 5 gears?

table(mtcars$gear)

62.5

#Assign the frequency of the mtcars variable "am" to a variable called "height"

```
 height <- table(mtcars$am)
 height
```

#Create a barplot of "height"

```
barplot(height)
```

vector of bar heights

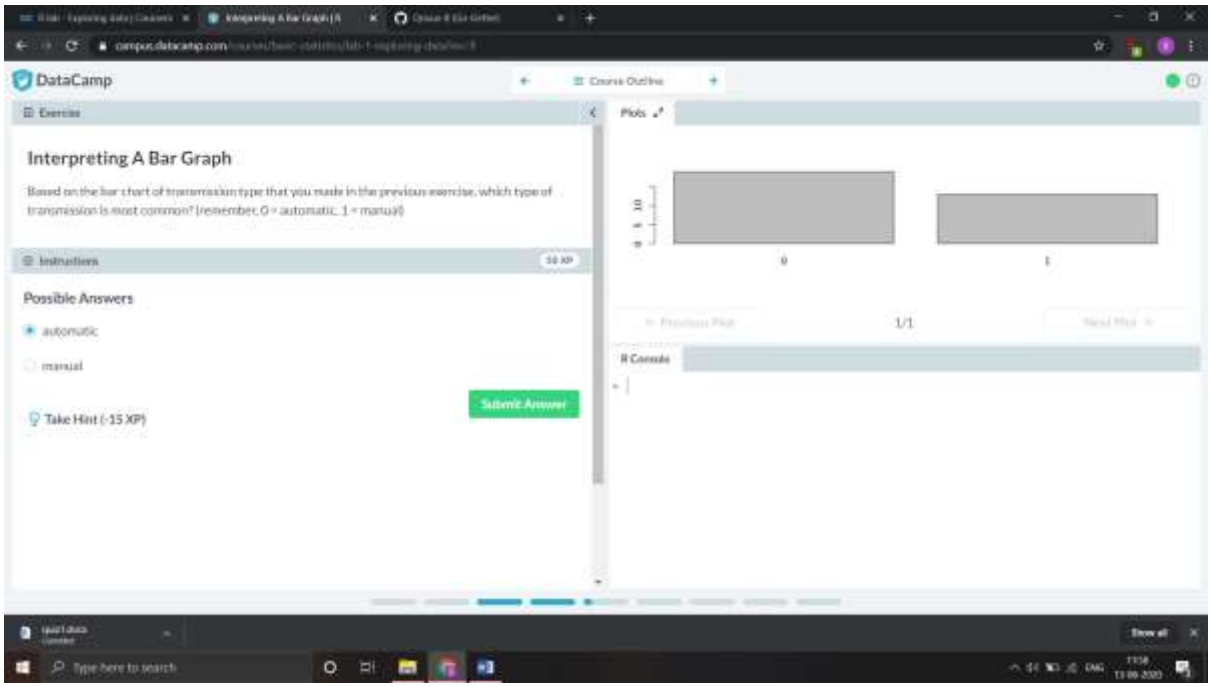
```
height <- table(mtcars$am)
```

Make a vector of the names of the bars called "barnames"

```
barnames <- c("automatic","manual")
```

Label the y axis "number of cars" and label the bars using barnames

```
barplot(height,ylab="number of cars", names.arg=barnames)
```

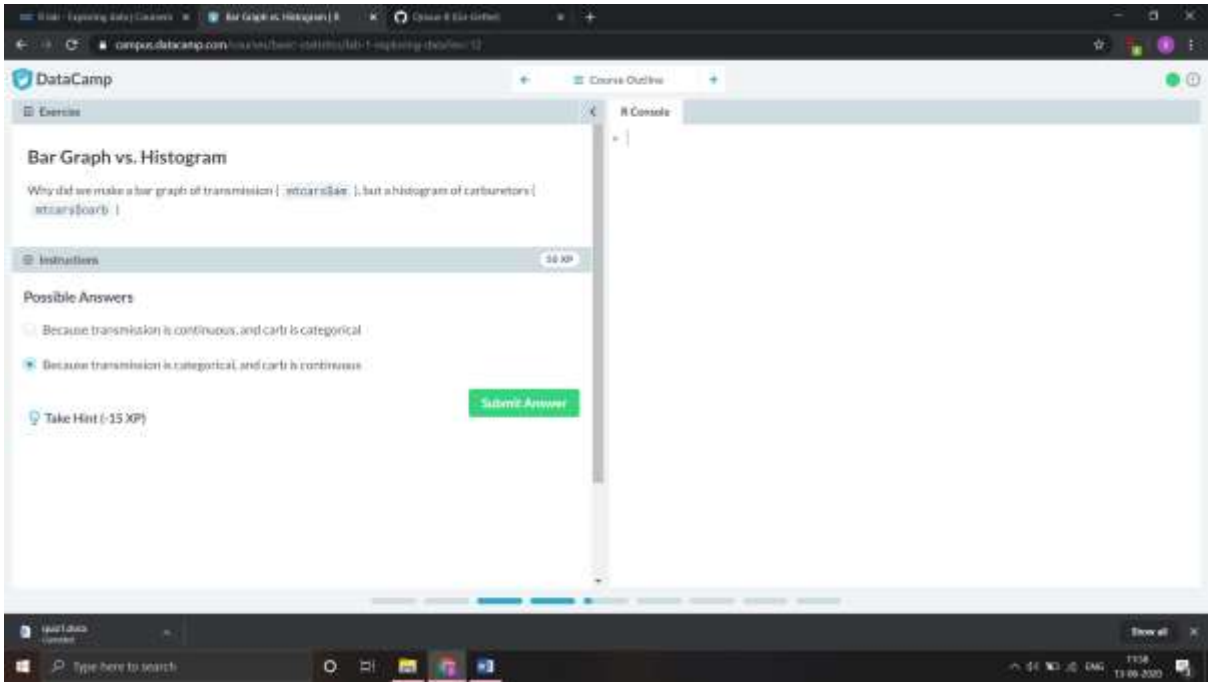


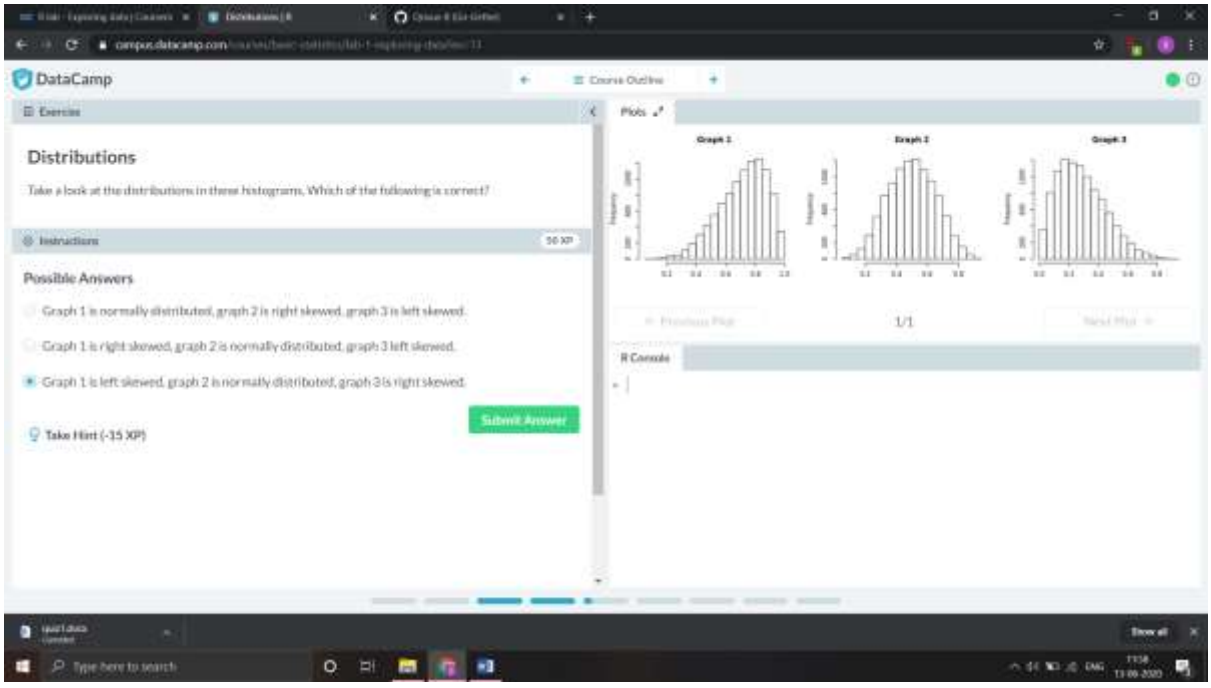
Make a histogram of the carb variable from the mtcars data set. Set the title to "Carburetors"

```
hist(mtcars$carb, main="Carburetors")
```

arguments to change the y-axis scale to 0 - 20, label the x-axis and colour the bars red

```
hist(mtcars$carb, main = "Carburetors", ylim=c(0,20), col="red", xlab="Number of Carburetors")
```

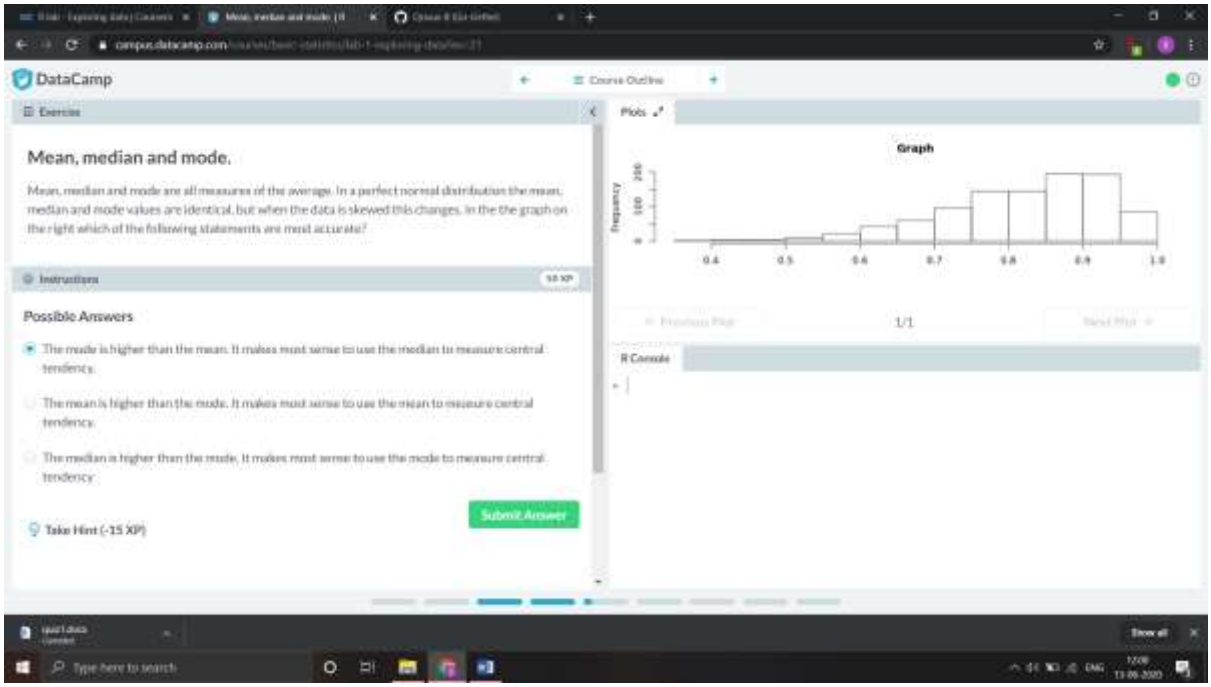





```
# Calculate the mean miles per gallon

mean(mtcars$mpg)

# Calculate the median miles per gallon

median(mtcars$mpg)

# Produce a sorted frequency table of `carb` from `mtcars`

sort(table(mtcars$carb), decreasing = TRUE)

# Minimum value

x <-min(mtcars$mpg)
```

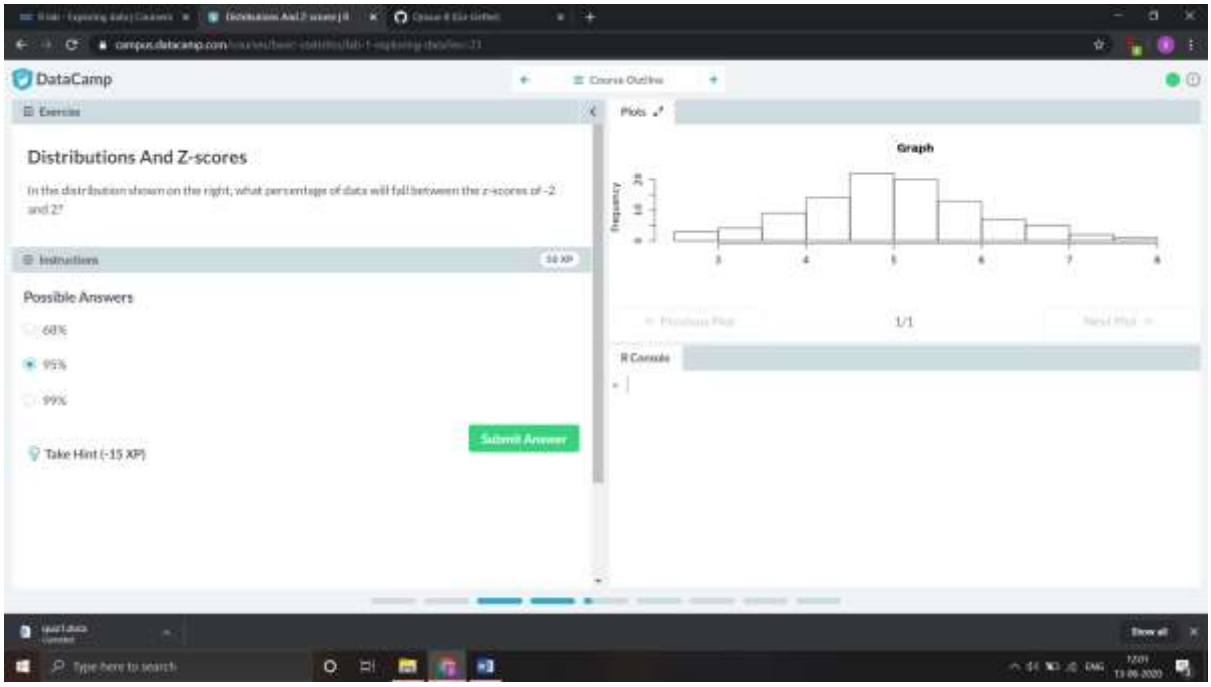
```
x
# Maximum value
y <-max(mtcars$mpg)
y
# Calculate the range of mpg using x and y
range <- (y - x)
range
# What is the value of the second quartile?
17.7100
# What is the value of the first quartile?
16.8925
quantile(mtcars$qsec)
# Make a boxplot of qsec
boxplot(mtcars$qsec)
# Calculate the interquartile range of qsec
IQR(mtcars$qsec)
IQR(mtcars$qsec)
quantile(mtcars$qsec)
# What is the threshold value for an outlier below the first quartile?
 lower_threshold <- (16.8925 - (1.5 * 2.0075))
lower_threshold
# What is the threshold value for an outlier above the third quartile?
higher_threshold <- (18.9000 + (1.5 * 2.0075))
higher_threshold
# Find the IQR of horsepower
IQR(mtcars$hp)
# Find the standard deviation of horsepower
sd(mtcars$hp)
# Find the IQR of miles per gallon
IQR(mtcars$mpg)
# Find the standard deviation of miles per gallon
```

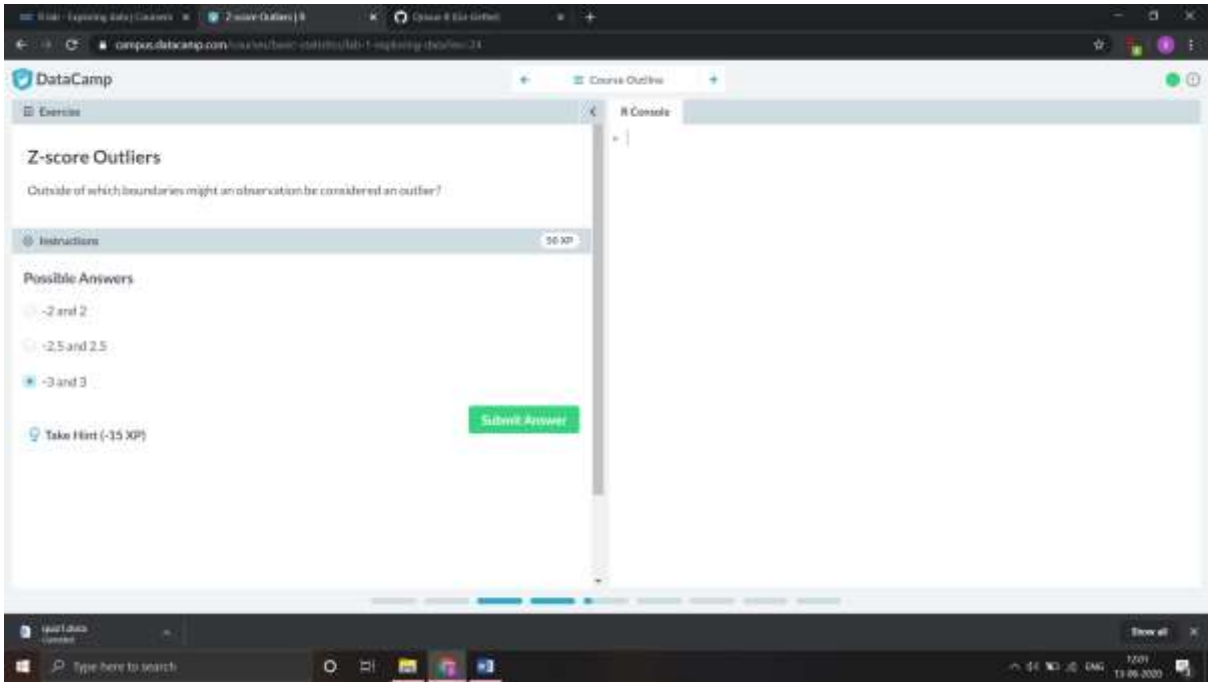
sd(mtcars$mpg)



Calculate the z-scores of mpg

(mtcars$mpg - mean(mtcars$mpg)) /sd(mtcars$mpg)




DataCamp
Course Outline
Exercise
Z-score Outliers
Outside of which boundaries might an observation be considered an outlier?
Instructions
50 XP
Possible Answers
-2 and 2
-2.5 and 2.5
-3 and 3
Submit Answer
Take Hint (-15 XP)
R Console
Type here to search