BISC-305 Discussion #2

09/09/2019

Exercise 1: Trojan's Scores Boxplot

Q: Trojans' score for the last season are as follows: 43, 3, 14, 39, 24, 31, 28, 35, 38, 14, 27, 17. We want to make a boxplot and see if there are any outliers. Repeat this with Example 2.4.2 in your textbook about pulses of 12 college patients: 43, 3, 14, 39, 24, 31, 28, 35, 38, 14, 27, 17

```
trojans.score <- c(43, 3, 14, 39, 24, 31, 28, 35, 38, 14, 27, 17)
# c(43, 3, 14, 39, 24, 31, 28, 35, 38, 14, 27, 17)
```

Let's look at the sorted scores:

```
sort(trojans.score)
```

Let's look at all the quantiles:

```
quantile(trojans.score)
```

Does it match with your calculations? R uses different algorithms to calculate quantiles. These can be specified using `type` argument. You can look up the various algorithms by typing `?quantile` in the console or searching for `qunatile` in the help section. Let's try a different algorithm for calculating quantile:

```
quantile(trojans.score, type=2)
```

```
IQR(trojans.score, type=2)
```

Boxplot:

```
boxplot(trojans.score , main="Boxplot of Trojans scores" , col = "#b2df8a")
```

Excercise 2: Scatter plot

For this and the next excercice we will be using the `mtcars` dataset that comes with `R`. It comprises of fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models). You can get alist of all datasets available inside `R` using `data()` command.

Let's first look at how the data looks like. The `head` command shows the first five rows along with all columns in the dataset.

Q: Can you determine which variables are categorical and which are numerical (continous/discontinous)?

```
head(mtcars)
```

Variables and their description:

```
[, 1]   mpg Miles/(US) gallon
[, 2]   cyl Number of cylinders
[, 3]   disp    Displacement (cu.in.)
[, 4]   hp  Gross horsepower
```

```
[, 5]   drat    Rear axle ratio
[, 6]   wt  Weight (1000 lbs)
[, 7]   qsec    1/4 mile time
[, 8]   vs  Engine (0 = V-shaped, 1 = straight)
[, 9]   am  Transmission (0 = automatic, 1 = manual)
[,10]   gear    Number of forward gears
[,11]   carb    Number of carburetors
```

Q: What is the relationship between the weight of the car and the mileage?

```
plot(mtcars$wt, mtcars$mpg,
     main="Scatterplot of wt vs mpg",
     xlab="Car Weight (lbs)",
     ylab="Miles Per Gallon")
```

Q: What relationship did you observe between the car weight and its mileage?

`ablines` command in `R` draws a straight line. `lines` command is a generic function to draw a curve passing it the coordinates. `lty` parameter determines the line type (1-continuous, 2-dashed).

```
plot(mtcars$wt, mtcars$mpg,
     main="Scatterplot of wt vs mpg",
     xlab="Car Weight (lbs)",
     ylab="Miles Per Gallon")
abline(lm(mpg~wt,data=mtcars), col="#e34a33", lty=1) # regression line
lines(lowess(mtcars$wt, mtcars$mpg), col="#2b8cbe", lty=2) # lowess line
legend(4.5, 32, legend=c("Regression line", "Lowess line"),
       col=c("#e34a33", "#2b8cbe"), lty=1:2, cex=0.8)
```

Q: How does the shape of the engine (vs) vary across the number of gears (gear)?

```
counts <- table(mtcars$vs, mtcars$gear)
barplot(counts,
        main="Car Distribution by Gears and shape of the engine",
        xlab="Number of Gears", col=c("#fc8d62","#8da0cb"),
        legend = rownames(counts))
```

Q: How does the 1/4 mile time (qsec) vary across Transmission (am) types?

```
boxplot(qsec ~ am,
        data = mtcars,
        main="Distribution of qsec by the type of transimission",
        xlab="Transimission type",
        ylab="1/4 mile time (sec)",
        col=c("#fc8d62","#8da0cb"))
```

Bonus Q: Is their a difference between the gross horsepower of the cars with manual vs those with automatic transmission?