PSC 103B - Lab 1 Assignment Answer Key

WRITE YOUR NAME HERE

Invalid Date

Question 1

Video Game (hours)	Aggression Score	Visuospatial Cognition
4	58	79
3	52	20
7	63	82
6	54	81
3	59	79
6	55	79
6	61	81
7	58	83
6	60	80
5	67	83

Enter the dataset into R, so that you have a dataframe object with the variables "VideoGames", "Aggression", and "Cognition". Show your code. (1 pt)

```
df <- data.frame("VideoGames" = c(4, 3, 7, 6, 3, 6, 6, 7, 6, 5),
"Aggression" = c(58, 52, 63, 54, 59, 55, 61, 58, 60, 67),
"Cognition" = c(79, 20, 82, 81, 79, 79, 81, 83, 80, 83))
```

Question 2

Find the mean and standard deviation of each variable. Show your code. (2 pts)

```
lapply(df, mean)
```

```
$VideoGames
[1] 5.3

$Aggression
[1] 58.7

$Cognition
[1] 74.7
```

```
lapply(df, sd)
```

```
$VideoGames
[1] 1.494434

$Aggression
[1] 4.423423

$Cognition
[1] 19.28183
```

Question 3

What do you think is the most appropriate measure of central tendency for the variable "Cognition"? Why? (1 pt)

`Median, due to presence of an outlier value (20)`

Question 4

How are variance and standard deviation related? Why might we prefer to use the standard deviation? (1 pt)

`Standard deviation is the square root of variance. We might prefer to use standard deviation because its units match the units of our variable, making it easier to interpret (variance has units that are the squared units of our variable)`

Plot VideoGames and Aggression as a scatterplot where the hours of each participant spending on VideoGames display on x-axis and their Aggression scores displays on y-axis. Make sure

your axes are appropriately labeled and include a title (put whatever you think make sense). Show your code. (1 pt)

```
plot(x = df$VideoGames, y = df$Aggression,
     xlab = "Hours Spent Playing Video Games",
     ylab = "Aggression Score",
     main = "Relation Between Time Spent Playing \nVideo Games and Aggression")
```