

# McKibben\_DSC520\_Ex\_7.2

Makayla McKibben

2024-07-28

```
# Install appropriate packages
# install.packages("tidyverse")

# Importing the data set
survey <- read.csv(file = 'student-survey.csv', header = TRUE, sep = ",", stringsAsFactors = FALSE)

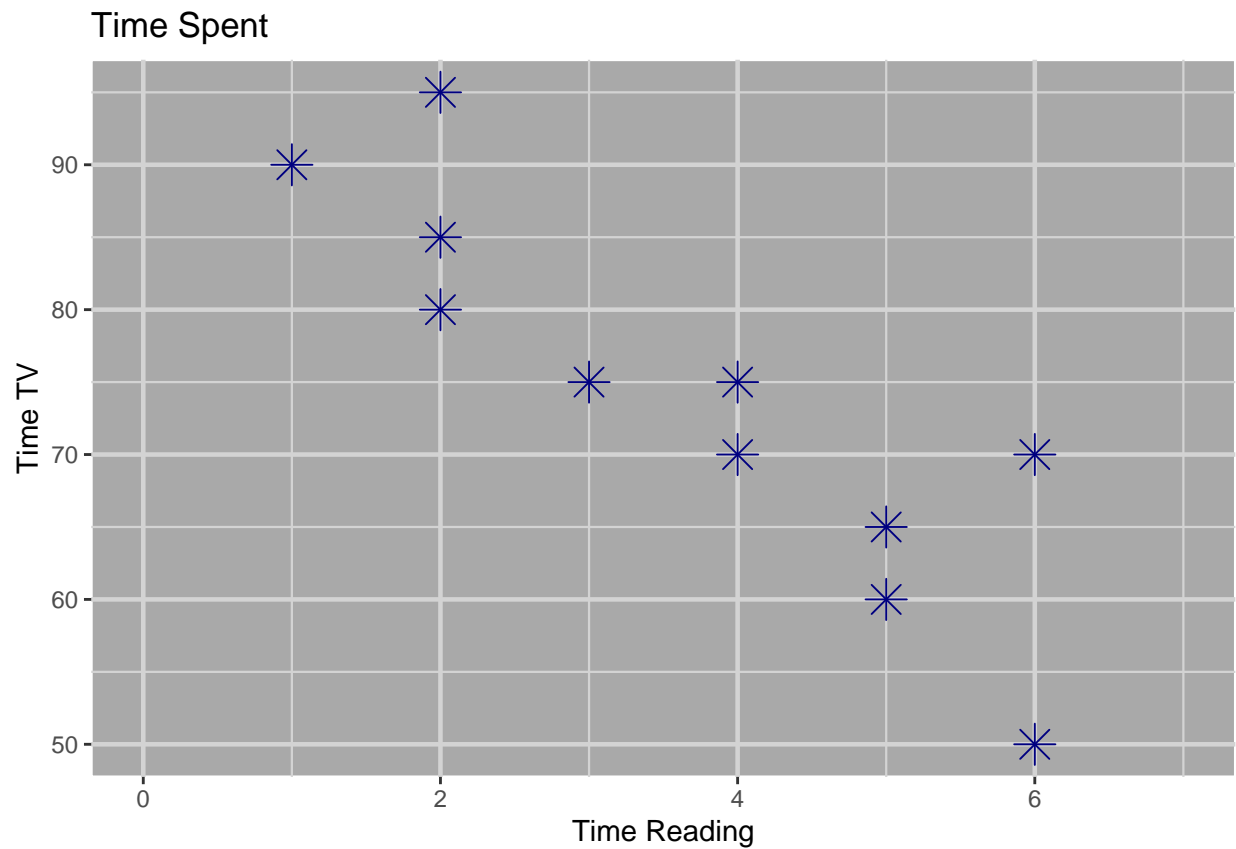
# Looking at a section of the data set to understand the structure
head(survey)
```

```
##   TimeReading TimeTV Happiness Gender
## 1           1     90      86.20      1
## 2           2     95      88.70      0
## 3           2     85      70.17      0
## 4           2     80      61.31      1
## 5           3     75      89.52      1
## 6           4     70      60.50      1
```

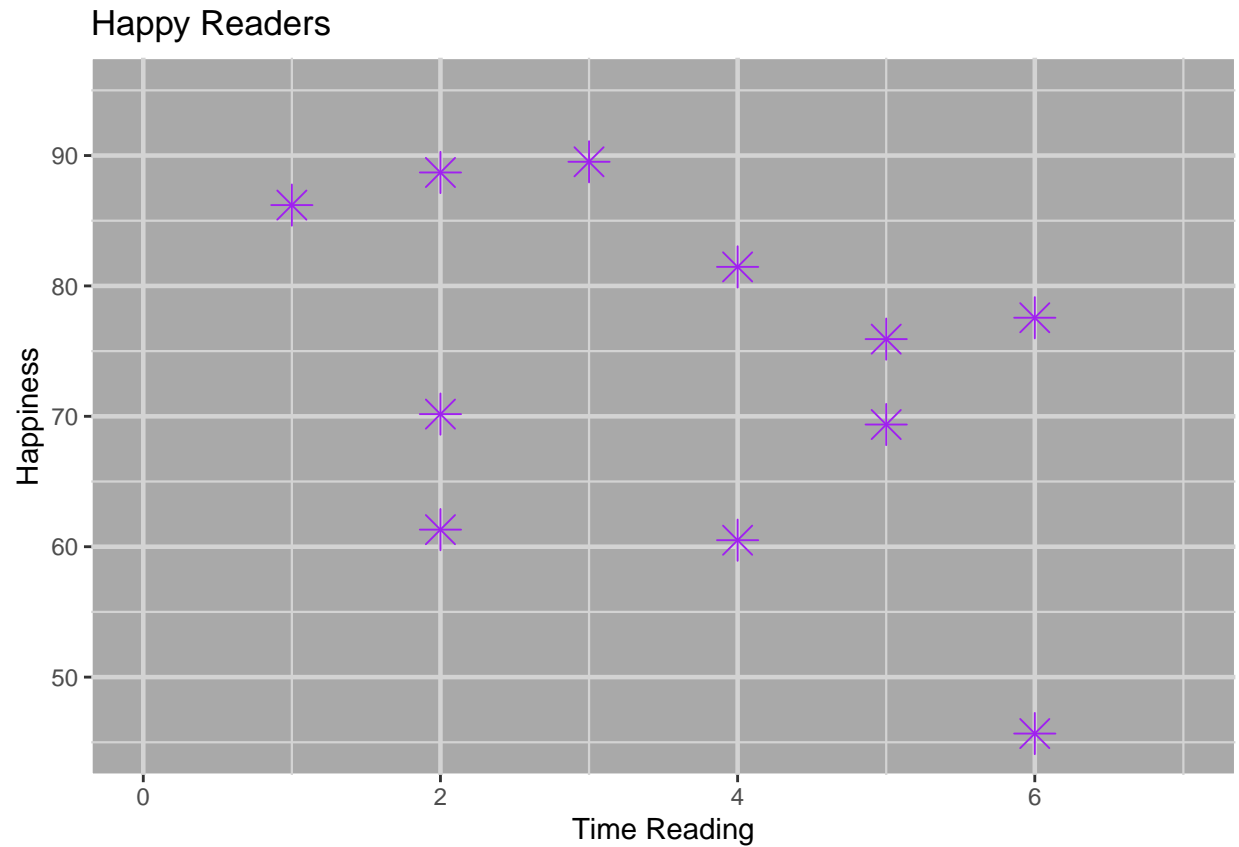
```
# Calling relevant library
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.1
```

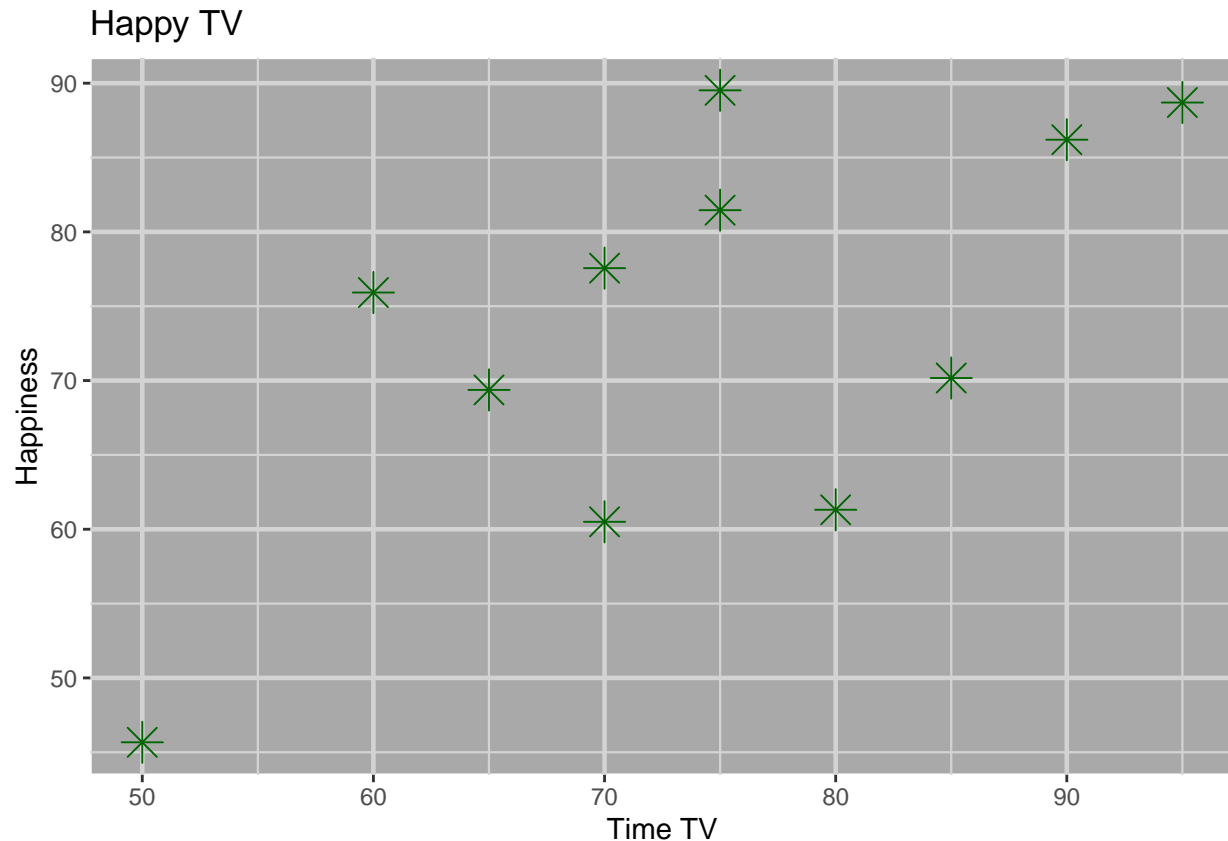
```
# Creating plots
r_v_tv <- ggplot(survey, aes(TimeReading, TimeTV))
r_v_tv + geom_point(color = "navy", shape = 8, size = 4.8) +
  theme(panel.grid = element_line(color = "lightgrey", linewidth = 0.8, linetype = 1),
        panel.background = element_rect(color = "white", fill = "darkgrey")) +
  labs(title = "Time Spent", x = "Time Reading",
        y = "Time TV") + xlim(0,7)
```



```
happy_reader <-ggplot(survey, aes(TimeReading, Happiness))
happy_reader + geom_point(color = "purple", shape = 8, size = 4.8) +
  theme(panel.grid = element_line(color = "lightgrey", linewidth = 0.8, linetype = 1),
        panel.background = element_rect(color = "white", fill = "darkgrey")) +
  labs(title = "Happy Readers", x = "Time Reading",
        y = "Happiness") + xlim(0,7) + ylim(45,95)
```



```
happy_tv <- ggplot(survey, aes(TimeTV, Happiness))
happy_tv + geom_point(color = "darkgreen", shape = 8, size = 4.8) +
  theme(panel.grid = element_line(color = "lightgrey", linewidth = 0.8, linetype = 1),
        panel.background = element_rect(color = "white", fill = "darkgrey")) +
  labs(title = "Happy TV", x = "Time TV",
        y = "Happiness") #+ xlim(0,7) + ylim(45,95)
```



Exercise 7.2 Q3 TimeReading vs. TimeTV shows a strong negative correlation TimeReading vs. Happiness shows a less strong, less negative correlation TimeTV vs. Happiness shows a weak positive correlation

```
# Find the covariance matrix
data_group <- cbind(survey$TimeReading, survey$TimeTV, survey$Happiness)
cov_data <- cov(data_group)
cov_data
```

```
##           [,1]      [,2]      [,3]
## [1,]    3.054545 -20.36364 -10.35009
## [2,]   -20.363636 174.09091 114.37727
## [3,]   -10.350091 114.37727 185.45142
```

Exercise 7.2 Q4 # TimeReading has a negative covariance with both TimeTV and Happiness. Happiness decreases half as fast as TimeTV as TimeReading increases

**As TimeTV increases Happiness increases rapidly as they have a relatively large positive covariance**

```
cor_data <- cor(data_group)
cor_data
```

```
##           [,1]      [,2]      [,3]
```

```
## [1,] 1.0000000 -0.8830677 -0.4348663
## [2,] -0.8830677 1.0000000 0.6365560
## [3,] -0.4348663 0.6365560 1.0000000
```

## Exercise 7.2 Q5

TimeReading has a negative correlation with both TimeTV and Happiness. Happiness decreases half as fast as TimeTV as TimeReading increases

As TimeTV increases Happiness increases rapidly as they have a relatively large positive correlation

I think that correlation is better and easier to interpret for one primary reason. I believe it's better because there's a reference number other than zero i.e. it's bounded by 1 and -1

```
# Find the correlation between TimeReading and TimeTV
cor_r_tv <- cor(survey$TimeReading, survey$TimeTV)
cor_r_tv
```

```
## [1] -0.8830677
```

## Exercise 7.2 Q6

TimeReading has a strong negative correlation with TimeTV.

**We cannot assume causation from correlation. We cannot assume causation from correlation.** In all seriousness, we should not imply causation based solely on the correlation coefficient. I would imagine we could still say that if you spend more time reading there's less time to spend watching TV, so it probably does have an effect but that's an observation separate from the correlation coefficient.