Cordero_week7.2

2024-07-17

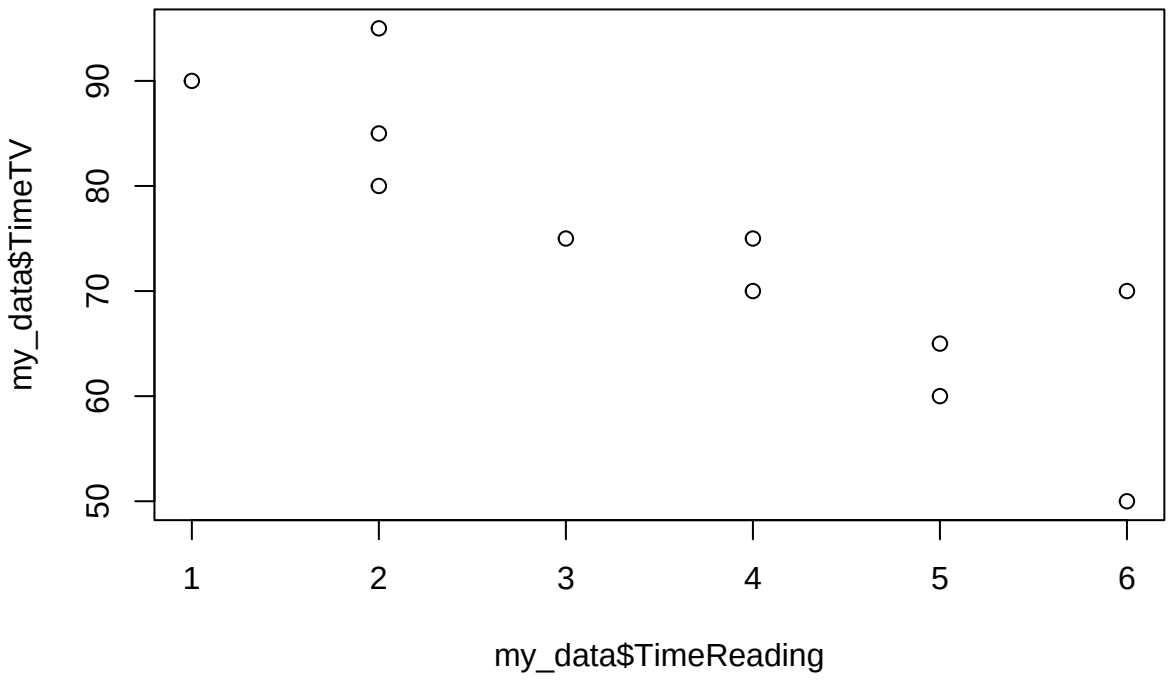
```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Loading required package: ggplot2

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

Survey Variable Plots

```
my_data <- read.csv('student-survey.csv')
```

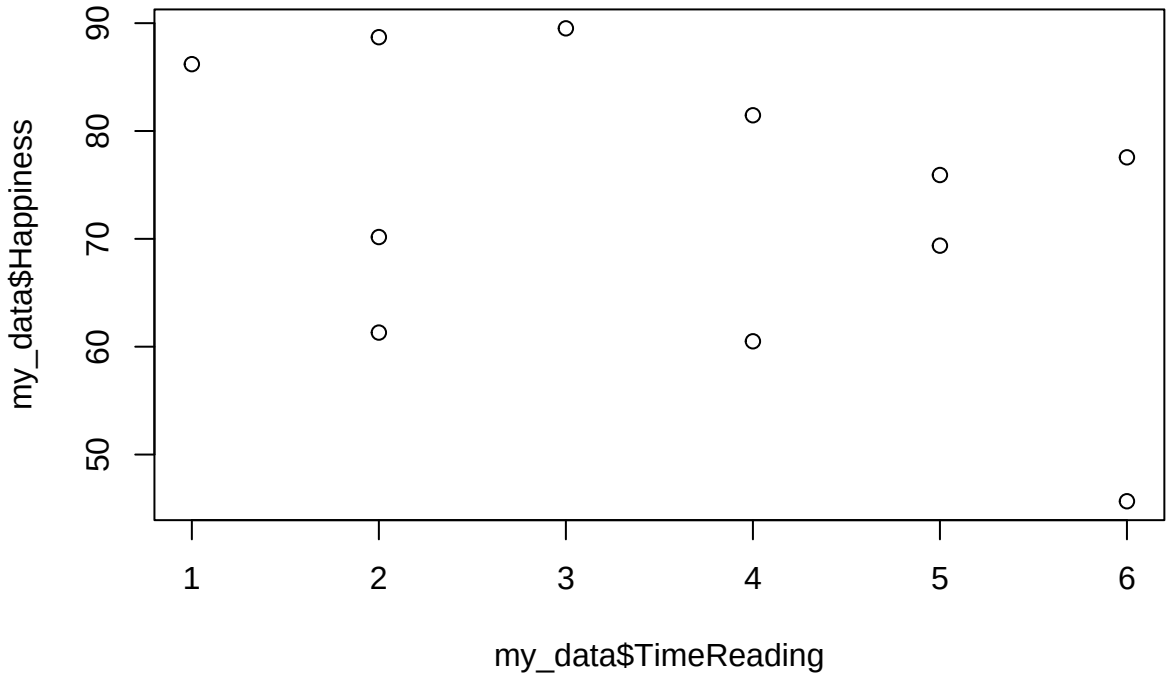
2a.

```
reading_tv <- plot(x = my_data$TimeReading, y = my_data$TimeTV)
```



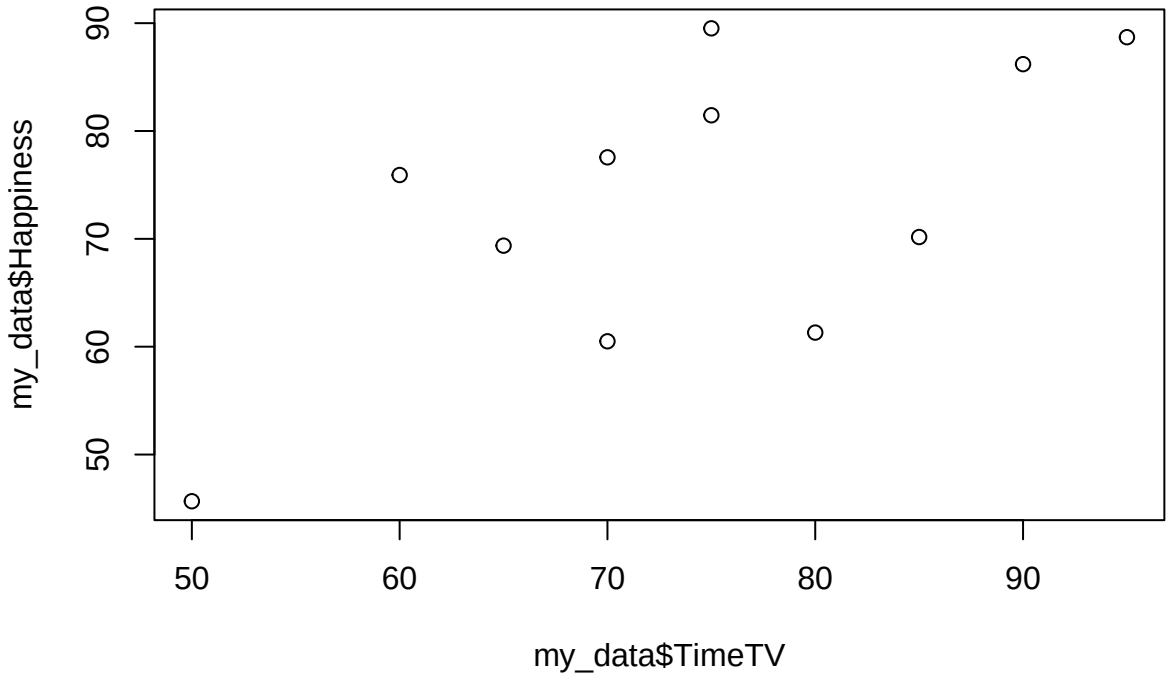
2b.

```
reading_happiness <- plot(x = my_data$TimeReading, y = my_data$Happiness)
```



2c.

```
tv_happiness <- plot(x = my_data$TimeTV, y = my_data$Happiness)
```



3a. The variables TimeReading and TimeTV slope indicate a negative relationship

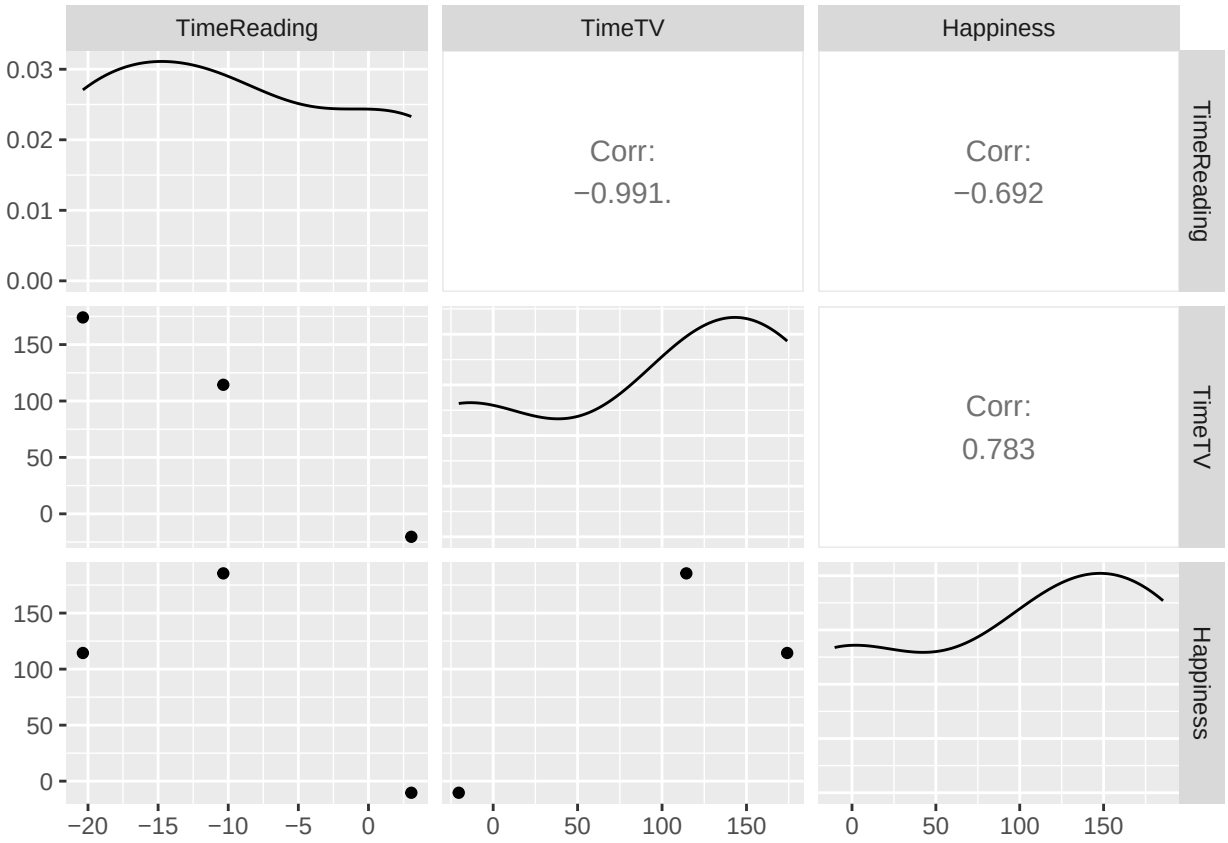
3b. The variables TimeReading and Happiness slope somewhat indicate a negative relationship

3c. The variables TimeTV and Happiness slop indicate a positive relationship

```
reading_tv_happiness <- my_data %>% select(TimeReading, TimeTV, Happiness)
```

4.

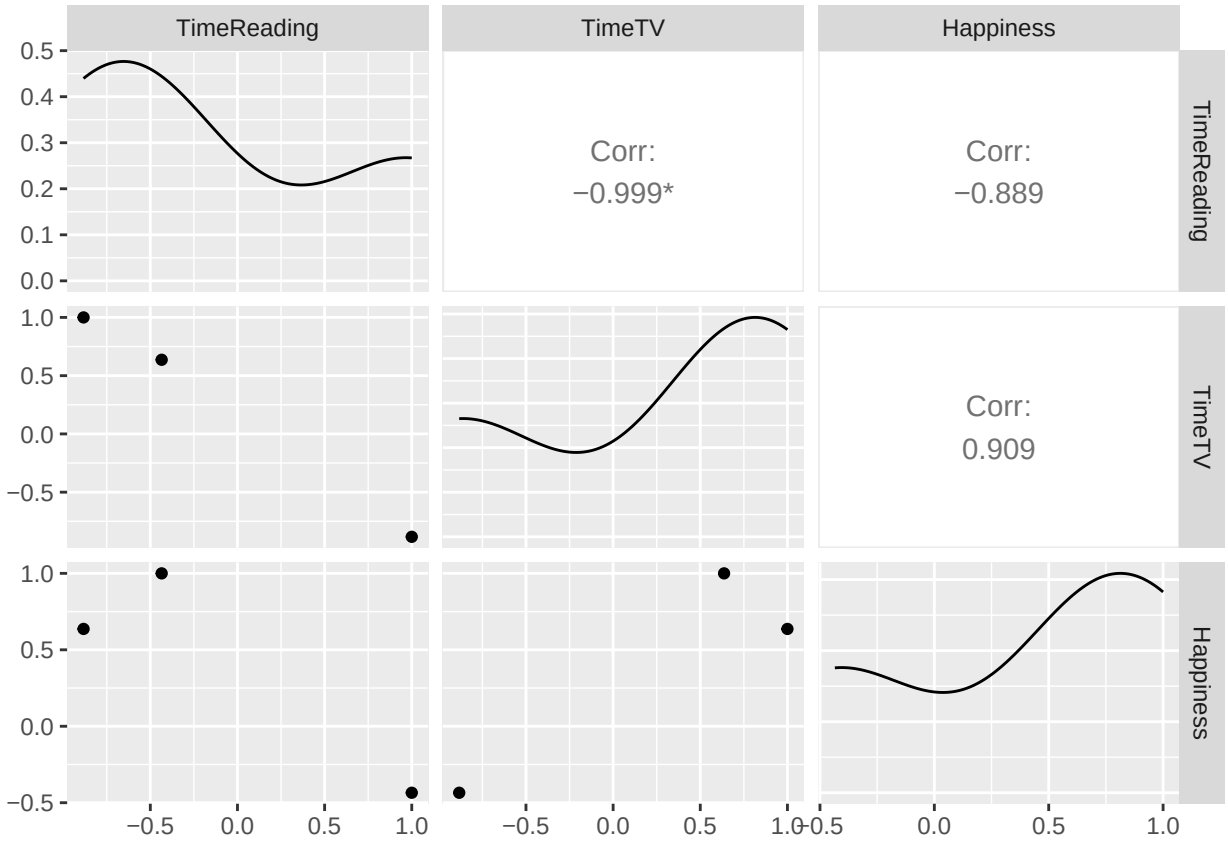
```
cov_matrix <- cov(reading_tv_happiness)

ggpairs(cov_matrix[])
```



TimeReading and Happiness have a negative correlation, TimeReading and TimeTV also has a negative correlation. Lastly, TimeTV and Happiness has a positive correlation.

5.

```
cor_matrix <- cor(reading_tv_happiness)

ggpairs(cor_matrix[])
```



TimeReading and Happiness still have a negative correlation, as well as TimeTV and TimeReading. Happiness and TimeTV still has a postive correlation. I think its easier to determine relationship using correlation matrix as its closer to -1 and 1, as well as the direction of the lines are more clear.

6.

```
cor(my_data$TimeReading, my_data$TimeTV)
```

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

The correlation between TimeReading and TimeTV is negative correlation. TimeTV has an effect on TimeReading, probably more time spent on TimeTV means less time spent on TimeReading and vice versa.