Importing libraries:

```
library("readxl")
```

```
## Warning: package 'readxl' was built under R version 4.1.3
```

```
library(ggplot2)
```

```
raw_data <- read_excel("C:/Users/dhire/Downloads/STA304 group 20 dataset.xlsx")
```

Cleaning Data and for now removing an entry that doesn't make sense

```
data <- raw_data[-c(31:33),]
data <- subset(data, Nickname != "crest")
```

Extracting variables for analysis

```
happiness <- data$Happiness
hobbytime <- data$time_hobby
studytime <- data$time_studies
```

Data Exploration

The number in column 1 is the number of people who said the time they spend on hobbies contributes more to their happiness, in column two, the number who said the time they spend on studies contributes more.

```
table(data$contribute_happiness)
```

```
##
##  1  2
## 25  4
```

Calculates the averages of the time participants spent on hobbies and studies, and their happiness level.

```
colMeans(data[, 4:6])
```

```
##   time_hobby time_studies    Happiness
##     2.810345     4.672414     6.689655
```

Counts how many times students rated their happiness an 8 out of 10, a statistic that was mentioned in the methodology section.

```
sum(data$Happiness == 8)
```

```
## [1] 11
```

This code will be used to add coloring to our plots, based of the students happiness score.

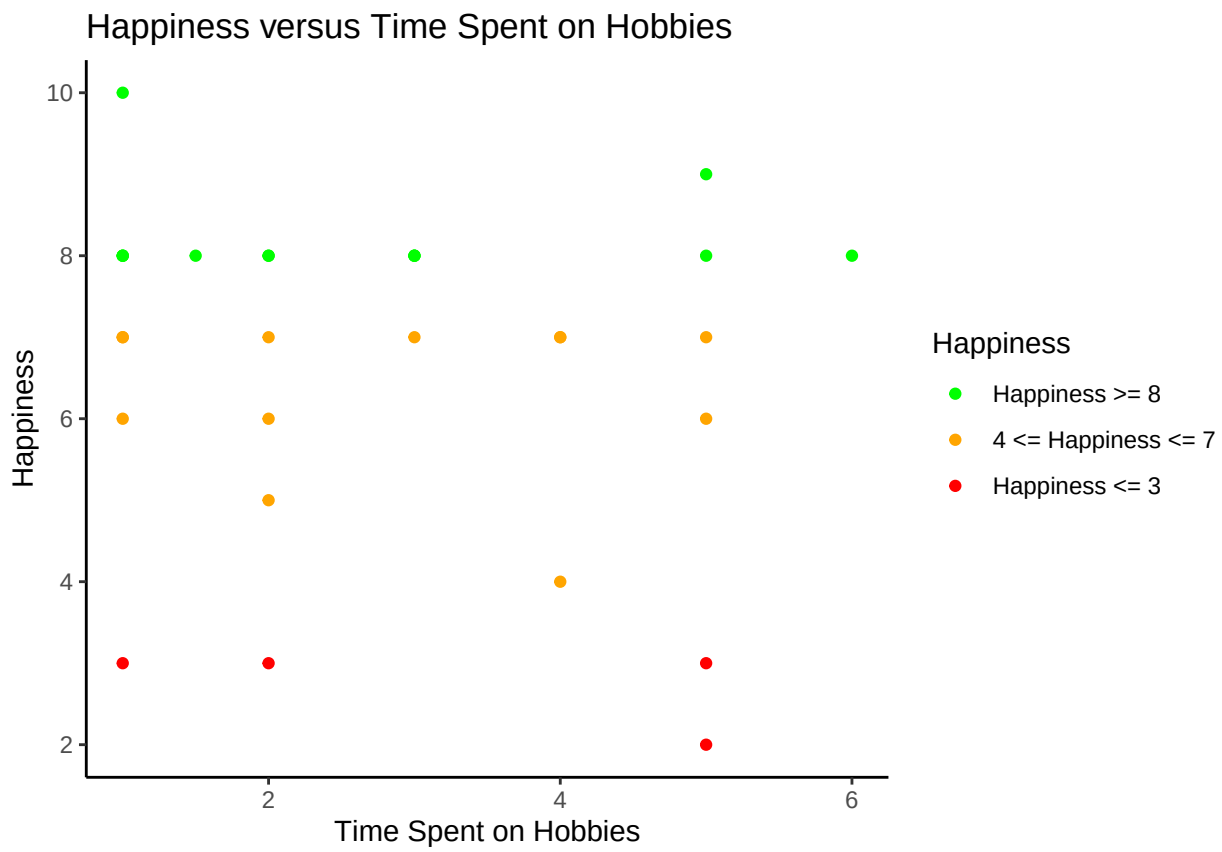
```
colors_happiness = happiness
colors_happiness[happiness >= 8] <- "green"
colors_happiness[happiness <= 7 & happiness >= 4] <- "orange"
colors_happiness[happiness < 4] <- "red"
colors_happiness
```

```
##  [1] "orange" "green"  "orange" "orange" "green"  "green"  "green"  "orange"
##  [9] "orange" "green"  "orange" "green"  "orange" "green"  "green"  "green"
## [17] "orange" "green"  "orange" "orange" "red"    "red"    "orange" "green"
## [25] "red"    "green"  "red"    "orange" "green"
```

We used the ggplot2 package to make a scatter plot for Happiness vs Time Spent on Hobbies.

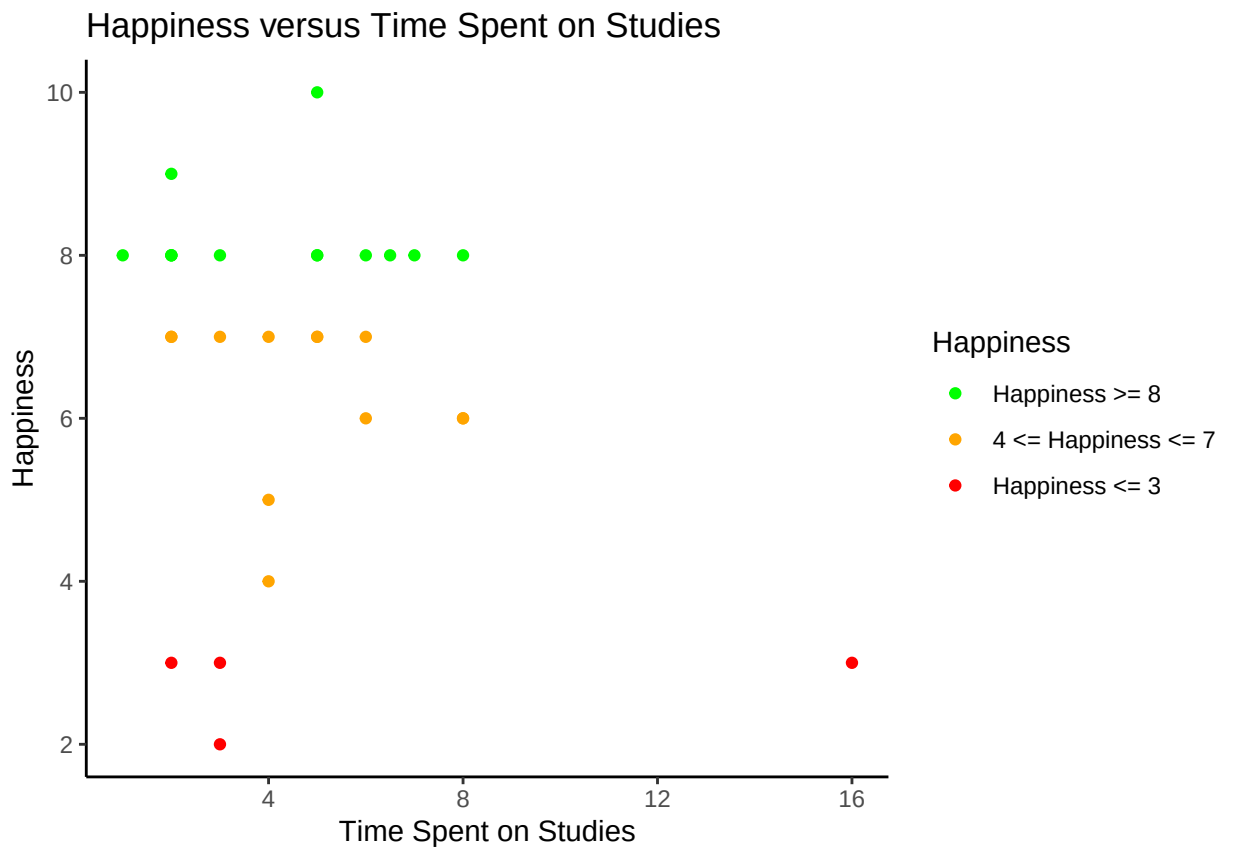
```
ggplot(mapping = aes(x=hobbytime, y=happiness, color = colors_happiness)) +
  geom_point() +
```

```
  scale_color_manual(values = c("green" = "green", "orange" = "orange", "red" = "red"), labels=c("Happi
  theme_classic() + labs(title = "Happiness versus Time Spent on Hobbies", x = "Time Spent on Hobbies",
```



We used the ggplot2 package to make a scatter plot for Happiness vs Time Spent on Studies.

```
ggplot(mapping = aes(x=studytime, y=happiness, color = colors_happiness)) +
  geom_point() +
  scale_color_manual(values = c("green" = "green", "orange" = "orange", "red" = "red"), labels=c("Happi
  theme_classic() + labs(title = "Happiness versus Time Spent on Studies", x = "Time Spent on Studies",
```



Testing Hypotheses

We run Kendall Tau's ranked correlation test on the time a student spends on their studies and their happiness level.

```
cor.test(studytime, happiness, method = "kendall", exact = FALSE)
```

```
##
##  Kendall's rank correlation tau
##
## data:  studytime and happiness
## z = -0.57363, p-value = 0.5662
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##         tau
## -0.08491301
```

We run Kendall Tau's ranked correlation test on the time a student spends on their hobbies and their happiness level.

```
cor.test(hobbytime, happiness, method = "kendall", exact = FALSE)
```

```
##
##  Kendall's rank correlation tau
##
## data:  hobbytime and happiness
## z = -0.67888, p-value = 0.4972
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
```

```
##        tau
## -0.1025904
```

We run a binomial test to test our third hypothesis. Recall from our code earlier, 25 is the number of students who believe that the time they spend on their hobbies contributes more to their happiness, 29 is the total number of valid participants.

```
binom.test(25, 29, p = 0.5, alternative = "two.sided")
```

```
##
##  Exact binomial test
##
## data:  25 and 29
## number of successes = 25, number of trials = 29, p-value = 0.0001037
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.6833594 0.9611052
## sample estimates:
## probability of success
##               0.862069
```