


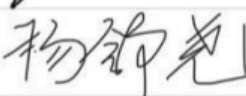

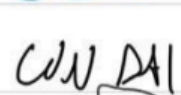
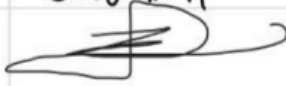


**University of Toronto at Mississauga**  
**Fall 2023**  
**SURVEYS, SAMPLING AND OBSERVATIONAL DATA**  
**STA304H5F LEC0101**

**COURSE PROJECT**

**Final Due Date: 23-11-2023**

*Late submissions will not be accepted.*

All undersigned students have made a significant contribution to this project:

| Name (Print)            | Student Number | Signature   |
|-------------------------|----------------|---|
| Jaewook Park            | 1005265197     |   |
| SHUYAO YANG             | 1007283645     |  |
| Steven Wong             | 1007196277     |  |
| LIN PAI                 | 1007226994     |  |
| Hyunwoo Jung            | 100860349      |  |
| Jessica Batta           | 1006250174     |  |
| Dhiren Kishore Tulsiani | 1005703240     |  |
|                         |                |   |

**Group Members who do not sign this sheet receive a mark of zero.**

# Written Report

Jessica Batta, Lin Dai, Hyunwoo Jung, Jaewook Park, Dhiren Tulsiani, Steven Wang,

Shuyao Yang

University of Toronto, Mississauga

STA304H5 FALL 2023 LEC0101: Surveys, Sampling and Observational Data

Professor: Dr. Luai Al Labadi

2023-11-27

## **Contents**

|                                       |          |
|---------------------------------------|----------|
| <b>1 Introduction</b>                 | <b>4</b> |
| <b>2 Elaboration of the Variables</b> | <b>4</b> |
| <b>3 Methodology</b>                  | <b>5</b> |
| <b>4 Data Overview and Insights</b>   | <b>5</b> |
| <b>5 Results</b>                      | <b>7</b> |
| <b>6 Conclusion</b>                   | <b>7</b> |
| <b>A R codes and Questionnaire</b>    | <b>8</b> |

## Introduction

As university students, most of our main goals are to succeed academically. However, before we are students, we are also humans; we strive for happiness. The goal in this analysis is to find out whether spending time on hobbies and/or studying contributes to how happy a given student is. We also understand that each and every student's self-measured happiness levels depend on what priorities they have. We expect the results of this study to help students efficiently allocate time to studying and hobbies to achieve a higher level of happiness.

To achieve our goal, we plan to consider 3 hypotheses. We will then use our survey statistics to test these hypotheses. We first hypothesize that there exists a relationship between time spent on hobbies and student happiness. Secondly, we hypothesize that there exists a relationship between time spent on studies and student happiness. Lastly, to see if students truly have different views on how to increase their happiness, we hypothesized the proportion of students who believe that spending time on hobbies leads to higher levels of happiness to be around half, or  $\frac{1}{2}$  of the entire STA304H5F students of the year 2023, while the other half believe spending time on studies leads to higher levels of happiness.

## Elaboration of the Variables

- **Lecture Section:** The participants' lecture section from STA304H5F: either LEC0101 or LEC0102.
- **Nickname:** The name that participants used to distinguish themselves in the survey.
- **Category:** Category of hobbies the participant engages with.
- **Time spent on hobbies:** The number of hours each participant spends on their hobbies per day on average.
- **Time spent on studies:** The number of hours each participant spends on their studies per day on average.
- **Happiness:** A 1 to 10 scale of whether the participant feels happy: 1 is very unhappy, 10 is very happy.
- **Influence of hobbies:** The participant's perception of how increasing time for hobbies might affect their happiness level.
- **Influence of studies:** The participant's perception of how increasing time for studies might affect their happiness level.
- **Academic Performance:** A measure of how close/far the participant feels towards their ideal academic performance.
- **Contribution to happiness:** Whether the participant thinks the time they spend on their hobbies contributes more to their happiness or the time they spend on their studies contributes more to their happiness.

## Methodology

From October 2023 to early November 2023, we conducted a survey which was uploaded to the course piazza. The purpose of the survey was to capture data on student happiness, time spent on study and hobbies, and the students' opinion on their relationship. Our survey was conducted anonymously, being filled by random students in the course. We received a total of 30 surveys, with one response that was physically impossible. After excluding the response, we had 29 responses to study.

To find out if there exists a relationship between the time spent on hobbies with happiness, and the time spent on studies in relation to happiness, we decided to use a statistical method called Kendall's tau rank correlation. We asked students to rate their Happiness as an integer from 1 to 10, and unfortunately most standard techniques to test the relationship between two variables don't work for data like this. Luckily for us, Kendall Tau's rank correlation does. On top of that it performs better than other alternatives when there are a lot of people who answer the same number, which is notable given that 11 people rated their happiness an 8 out of 10.

To test our third hypothesis, we need to look at our survey first. The survey asked the participants if they believe allocating more time to hobbies would lead to higher happiness level or allocating more time to studies would lead to higher happiness level. We treated the response as a binary response, which is simply a "yes" if they believe that hobbies lead to higher happiness levels. and a "no" if they believe that studying leads to higher happiness levels. Treating the responses as binary, we decided to use the binomial test to deal with the binary responses.

## Data Overview and Insights

To dive deeper into our data, we will look into the true proportion of students who believe that allocating more time on hobbies lead to higher levels of happiness as  $\hat{p}$ . The true proportion of students who think time spent on hobbies contributes more to their happiness is calculated as follows:

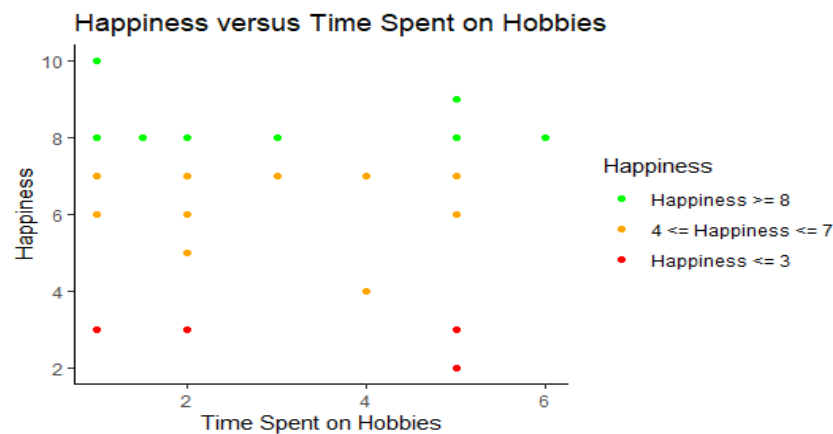
$$\hat{p} = \frac{\text{the number of students who think time spent on hobbies contribute to higher happiness}}{\text{the total number of valid responses}} = \frac{25}{29} = 0.86201$$

Here is a table containing the averages of some of the other factors in our data:

| Variable              | Average  |
|-----------------------|----------|
| Happiness             | 6.689655 |
| Time spent on hobbies | 2.810345 |
| Time spent on studies | 4.672414 |

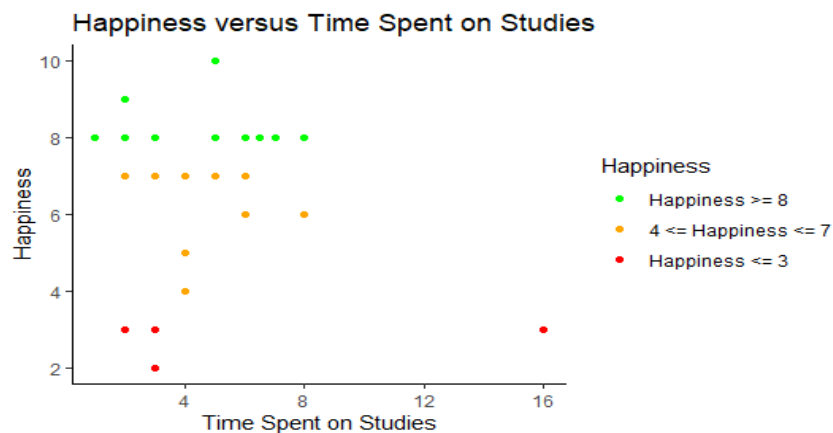
Before diving into any statistical testing, reviewing some graphs and visual representations may help us discern a noticeable connection between Happiness and the two other variables.

Let us visualize the relationship between time spent on hobbies and happiness.



Visually, there does not seem to be any linear relationship. There are students who spend the same amount of time on their hobbies, yet their happiness levels completely differ.

Next, let us visualize the relationship between time spent on studies and happiness.



Again, there does not seem to be any obvious relationship between students' time spent on studies and happiness. We should look into the results of the statistical test to confirm if there exist no relationship between the pairs.

## Results

Computations by hand using the method mentioned is very difficult, so we used RStudio, a programming language for statistical computing to execute our computations. After testing our first hypothesis in RStudio, we were 95% certain that there exists no relationship between time spent on hobbies and happiness.

Testing our second hypothesis in RStudio returned similar results: with 95% certainty, there existed no relationship between time spent on studies and happiness.

After testing our last hypothesis, where we assumed the proportion of students who believe that allocating more time to hobbies would lead to higher happiness level, we were 95% sure that the proportion of students who believe that allocating more time on hobbies would lead to happiness lies between 0.6834 and 0.9611 out of 1, or 68.34% to 96.11% of the students.

Comparing the visuals and our first two results as evidence, we can now understand why it seemed like there existed no relationship between the pairs.

## Conclusion

In conclusion, our main goal of finding a correlation between time spent on hobbies, time spent on studies and happiness was unsuccessful. From the collected data, there is no evidence to suggest that relationships exist between happiness and time spent on either hobbies or studies. In fact, as we can see in our graphical displays, student happiness varies at all levels of hobby or study time commitments.

From our last hypothesis we found out that the true proportion of students who think spending time on their hobbies contributes more to their happiness is not 0.5, but rather interestingly much higher, at approximately 70 - 95%.

This means that the majority of the students believe that the time they spend on hobbies contributes more to their happiness. Yet our other results show that the time a student spends on hobbies does not correlate with happiness. This may indicate that students overestimate the effects of hobby time on happiness.

Upon conducting this research, the result of no correlation between time spent on either hobbies or studies to happiness made us question if our hypothesis was too subjective or if our method of

research was not optimal; we believed that there were a few limitations to our research that returned the above results:

1. Small Sample Size:

We were only able to receive 30 responses from 2023 STA304H5F students. If we had more samples, we believe that the result could be different.

2. Limited options:

When asking students whether the time they spend on hobbies or the time they spend on studies contributes more to happiness, we didn't give them an option to say both equally contribute.

3. Existence of "Ties" when using Kendall's tau rank coefficient:

Though it is a minor effect, having ties in our observations can impact the precision and validity of the tests.

For future research, aside from collecting more responses from students, modification to the survey to obtain data usable for statistical tests with higher accuracy. Also, the limited number of options when asking which factor contributes to higher happiness level likely made us come to the conclusion that students may overestimate the effects of hobbies on happiness. Allowing students to freely respond to such questions can help model a more effective study of this topic.

## **Appendix: R Codes and Questionnaire**



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
```

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"
```

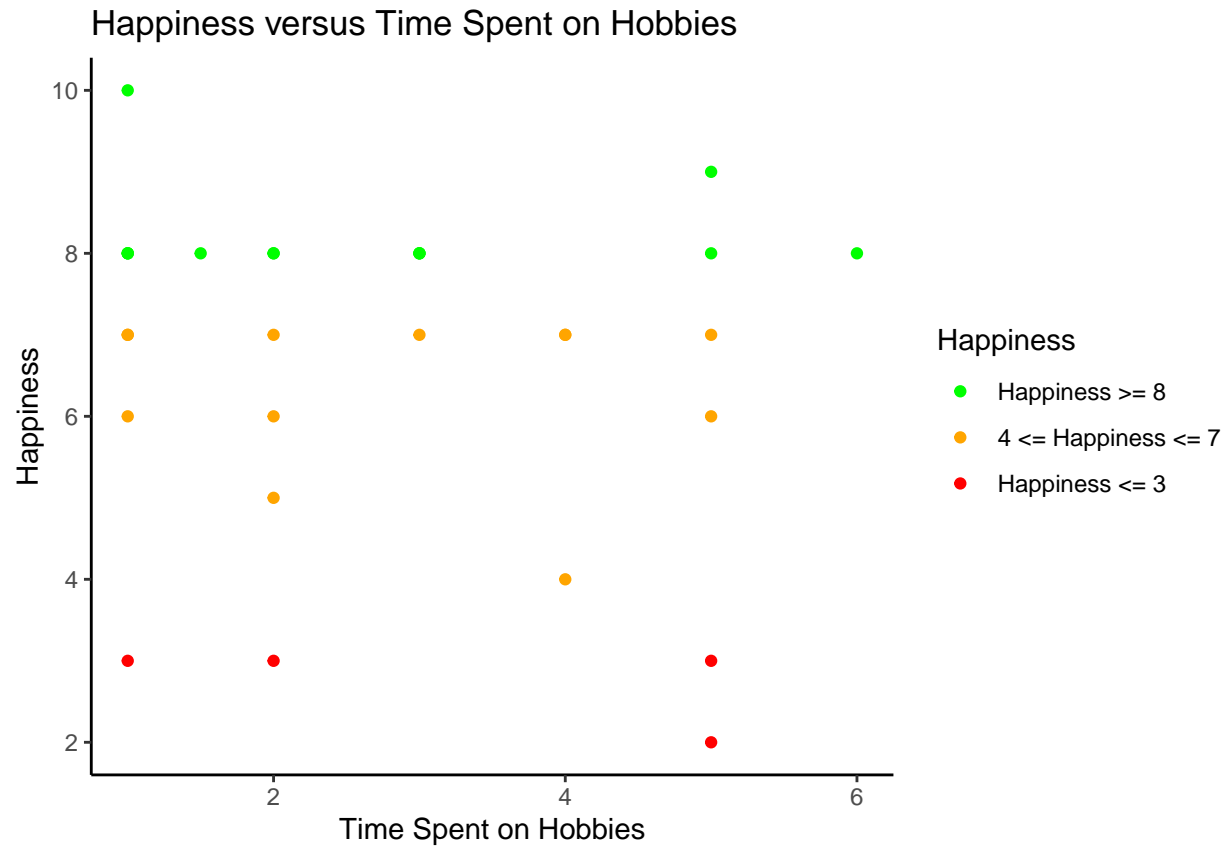
We used the ggplot2 package to make a scatter plot for Happiness vs Time Spent on Hobbies.

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

```
geom_point() +
```

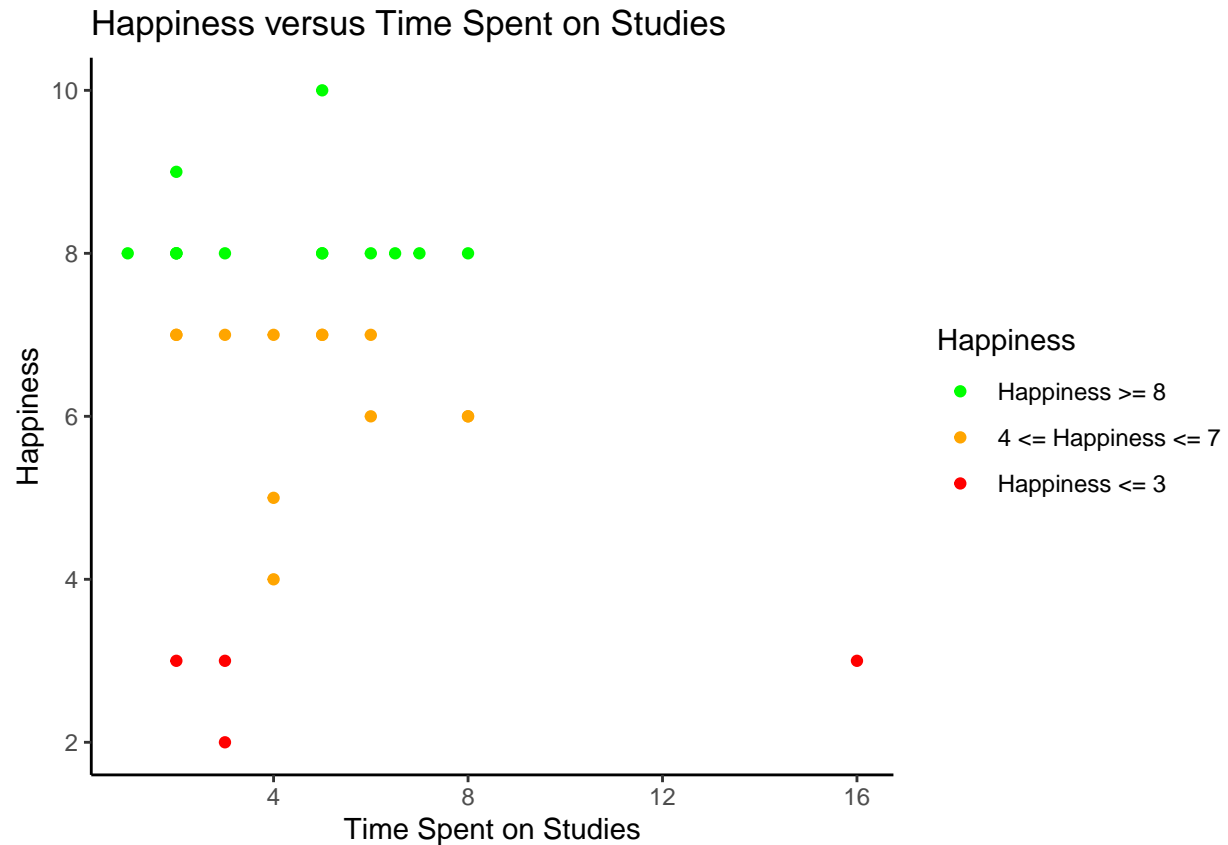
```
scale_color_manual(values = c("green" = "green", "orange" = "orange", "red" = "red"), labels=c("Happi
```

```
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("Happiness >= 8", "4 <= Happiness <= 7", "Happiness <= 3")) +
  theme_classic() + labs(title = "Happiness versus Time Spent on Studies", x = "Time Spent on Studies", y = "Happiness")
```



#### 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
```

# STA304H5F Relationship between Happiness and Time Spent on Hobbies/Studies Questionnaire

Dear Participant,

All of your data will be kept anonymous and confidential within our group. You are not forced to fill out this survey; you may exit any time if you feel there is intrusion in your personal information. Thank you.

---

\* Indicates required question

1. Which lecture section are you in? \*

*Mark only one oval.*

☐ LEC0101

☐ LEC0102

2. Please share your nickname to distinguish you! \*

---

3. Which categories do your hobbies fall under? \*

*Check all that apply.*

- ☐ None
- ☐ Exercise (Working out, Sports, etc)
- ☐ Seeing friends
- ☐ Video games
- ☐ Social Networking Services (Instagram, Facebook, etc)
- ☐ Television or Streaming (YouTube, Netflix, Disney+, Apple TV, etc)
- ☐ Creative or Performative Arts (Writing, Instrument, etc.)
- ☐ Culinary Arts
- ☐ Other: \_\_\_\_\_

4. If you have hobbies, how many hours do you spend on hobbies on a daily basis? \*

\_\_\_\_\_

5. How many hours do you spend on your studies on a daily basis? \*

\_\_\_\_\_

6. How would you rate your happiness from 1 to 10? 1 - very unhappy, 10 - very happy \*

*Mark only one oval.*

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ 6

☐ 7

☐ 8

☐ 9

☐ 10

7. How do you believe allocating more time to your hobbies might influence your happiness? \*

*Mark only one oval.*

- ☐ Would significantly increase my happiness
- ☐ Would somewhat increase my happiness
- ☐ No significant impact on my happiness
- ☐ Would somewhat decrease my happiness
- ☐ Would significantly decrease my happiness
- ☐ I'm unsure/not certain

8. How do you believe allocating more time to your studies might influence your happiness? \*

*Mark only one oval.*

- ☐ Would significantly increase my happiness
- ☐ Would somewhat increase my happiness
- ☐ No significant impact on my happiness
- ☐ Would somewhat decrease my happiness
- ☐ Would significantly decrease my happiness
- ☐ I'm unsure/not certain



9. How close do you feel you are to your ideal academic performance?

*Mark only one oval.*

- ☐ Very close to my ideal
- ☐ Somewhat close to my ideal
- ☐ Neutral - neither close nor far
- ☐ Somewhat far from my ideal
- ☐ Very far from my ideal
- ☐ I haven't defined an ideal academic performance for myself

10. What do you think contributes more to your happiness? \*

*Mark only one oval.*

- ☐ Time spent on hobbies
- ☐ Time time spent on studies

---

This content is neither created nor endorsed by Google.

Google Forms

