STA304 - Fall 2022

Assignment 1

Tshegofatso Mable Kelesitse - 1006724298

6 October 2022

Part 1

Goal

The goal of my survey is to determine the relationship between different levels of physical activity and happiness in university students. According to Human Kinetics (2020), physical fitness refers to the ability of our body systems to work together efficiently to allow you to be healthy and perform activities of daily living.Studies have shown that physical activity plays an important role in preventing and treating certain diseases such as depression and anxiety disorders (Peluso and Guerra de Andrade, 2005).

This topic is relevant because physical activity is linked to enhancing brain function and improving on-task classroom behaviors (Communications, n.d.). Exercise improves mood, boosts energy and promotes better sleep (Mayo Clinic, 2021). It is especially important for university students who face a lot of academic pressures. The topic explores possible strategies that students may apply.

The survey will contribute to the overall goal by gathering data through asking questions related to physical activity and happiness while taking demographic variables into account. .

Procedure

Data is collected through a survey link that will be posted on Instagram and Facebook where University of Toronto students will be reached. Additionally, the survey link will be posted on course discussion boards and discord groups. The link will be accompanied with a strict instruction that only U of T students above 18 years are requested to submit responses.

The target population is University of Toronto students above 18 years of age. The frame population is everyone who fills in the survey and successfully submits their responses. The sample population included 42 students.

Participants are chosen based on their willingness to take part in the survey. Some of the drawbacks in the volunteer-based sampling procedure are, the sample population may not be a representative of the entire target population. Additionally, there is a possibility of volunteer bias, whereby there is a difference between participants who responded to the survey and those who did not (Shantikumar, 2018). Some respondents may not respond truthfully to questions they find uncomfortable. Another concern is students may avoid taking the survey because of fears of being hacked through social media. However, there are strengths to this procedure. It is convenient as responses are automatically recorded and saved, cost-effective and time efficient.

Showcasing the survey.

https://forms.gle/N83oRJsJAPk4cNFP7 [1]

Question 1

Which of the following currently describes your current relationship status? - Married - Widowed - Divorced - Separated - Cohabiting with a significant other or domestic partnership - Single, never married - Prefer not to answer

Social relationships play an important role in the lives of humans. Our thoughts feelings and behaviors are influenced by the presence of the people around us. I selected this question because social relationships are a good predictor for a happy life (Oppong, 2019).

Some respondents may feel uncomfortable answering questions relating to their marital status and opt not to answer. It may lead to fewer responses regarding this demographic variable, which studies have shown is a good predictor for a happy life.

Question 2 In general I consider myself:

not a very happy person 1 2 3 4 5 6 7 8 9 10 a very happy person

This question relies on self-report answers to determine if the student is happy. I selected this question because it is important to this study since happiness is the response variable we are interested in.

There are different definitions of happiness, and this question only serves to measure subjective happiness and doesn't contribute to answering other forms of the definition. Such as that proposed by Aristotle and other Greek philosophers (David et al., 2014).

Question 3 What is your current age?

Age may have an influence on peoples subjective well being. Different age groups experience different pressures and priorities. Studies have shown well-being tends to be U-shaped over a life-cycle (Blanchflower & Oswald, 2007)

Certain individuals may feel uncomfortable providing an exact answer about their age, therefore they may choose to not answer the question truthfully.

Part 2

Data

Data was collected by requesting University of Toronto students above 18 years to volunteer to fill out a survey with questions relating to physical activity, demographic variables and answer the Subjective Happiness Scale. The demographic variables assessed in the study are age, gender and marital status.The survey was created using Microsoft forms, therefore responses were automatically recorded when participants submitted their results.

The survey link was posted on Instagram, WhatsApp, course Discussion boards and school related Dischord groups. A drawback faced during data collection was some students declined to answer the survey due to fear of the possibility of their social media accounts being hacked, if they pressed on unfamiliar links. Since the survey was posted on personal social media feeds, it may be difficult to reach students who are not on my followers list.

An excel spreadsheet was created for all the survey responses. For question four namely,'Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. To what extend does this characterization describe you?' Question four was was reversed scored in the following way, 7 changes to a 1, 6 to a 2, 5 to a 3, 4 remains as 4, 3 to a 5, 2 to a 6 and 1 to a 7.A variable for the overall subjective happiness was created by summing up the scores continuously.

Additionally, the data frame columns were formatted to be more readable.

Variables Gender This describes the gender the individual most identifies with

Age The current age of the individual

Relationship Status Current relationship or marital status of the individual

Fitness_Level The current overall fitness level of the individual ranging from Perfect, Good, Average, Poor and Unfit

Alcohol_Drugs This describes whether the individual consumes alcohol, partakes in recreational drugs or smokes. Response is a binary, Yes or No

Physical Activity per week The number of days per week an individual engages in physical activity for more than 30 minutes which resulted in increased breathing rate.

Overall_SH The overall subjective happiness of an individual. Happiness was measured using the Subjective Happiness Scale (Lyubomirsky & Lepper) This is a four question scale. To determine the overall subjective happiness of the participants, the scores from the questions are summed up continuously, keeping in mind question 4 is reversed scored. Larger score indicate a higher overall subjective happiness.

Table 1: Figure1: Statistics about Overall happiness for each fitness level

Fitness Level	Mean Overall Happiness	Standard Deviation of Overall Fitness	Minimum Overall Happiness	Maximum Overall Happiness
Average	18.90909	4.104691	10	28
Good	21.58333	4.122187	11	26
Perfect	23.66667	2.081666	22	26
Poor	20.40000	4.159327	16	25

The sample population is 42 students with 22 women and 20 men. The ages are between 19 and 26 inclusive with an average age of 22.36 years.The mean overall subjective happiness is 20.19 units with a standard deviation of 4.17 units.

Figure 2: Histogram of the Number of days per week an individual engages

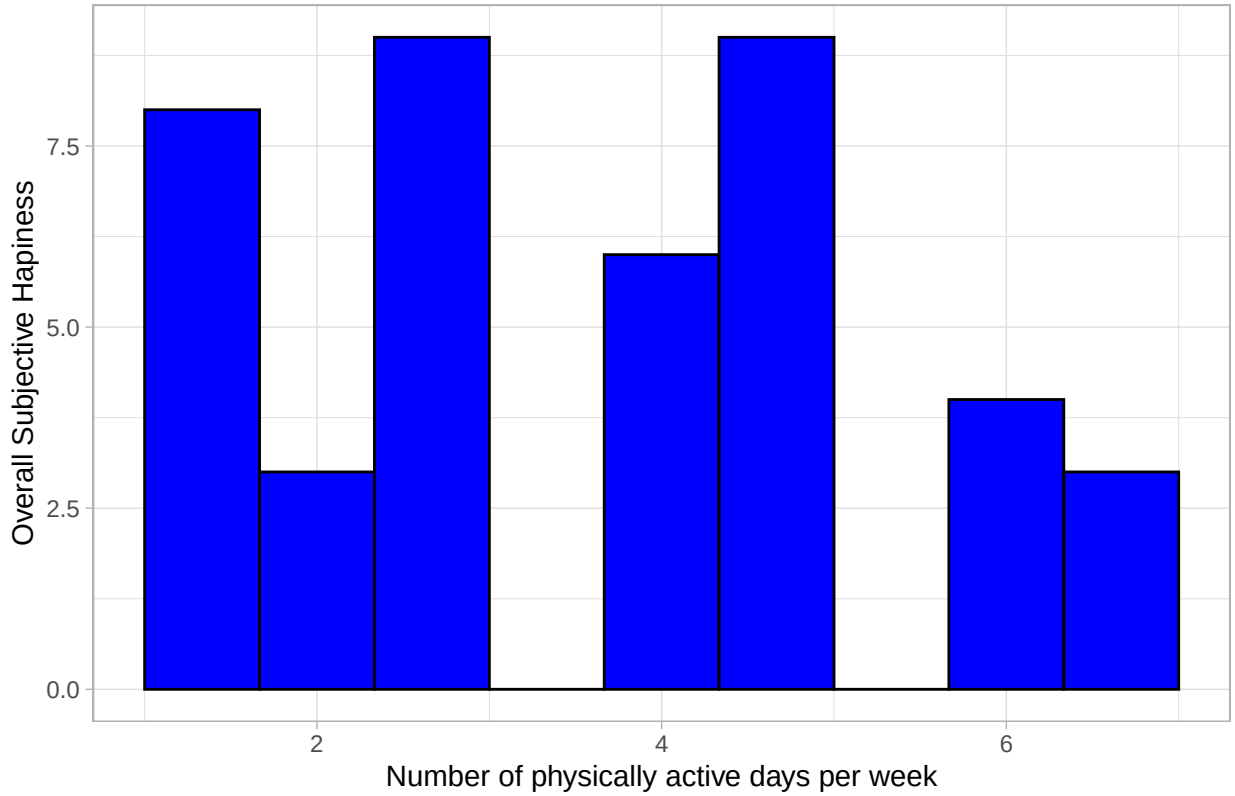
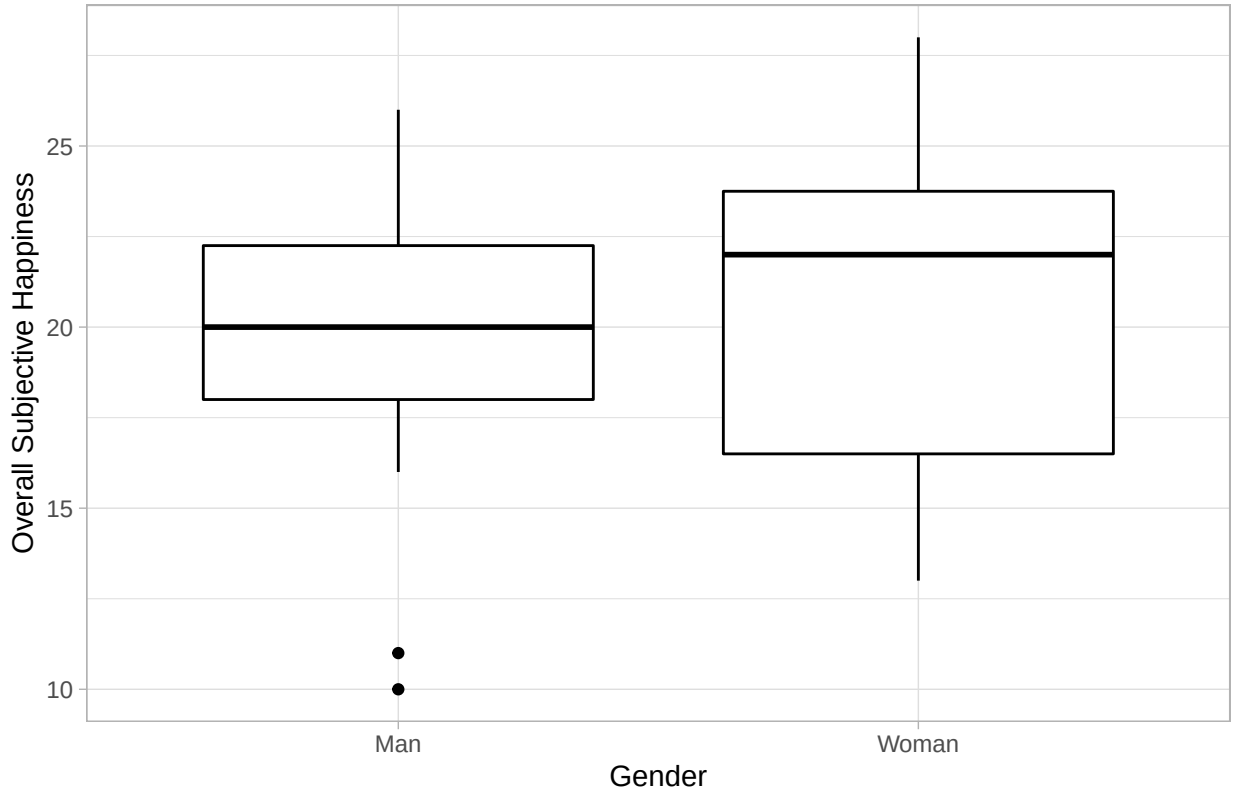


Figure 3: Boxplot of the Overall Subjective happiness for each Gender



The mean overall happiness for fitness levels perfect, good, average and poor are 23.67, 21.58, 18.91 and 20.40 respectively based on Figure 1. The Perfect fitness level has the lowest standard deviation at 2.08 out of all the four groups. The Average, Good and Poor fitness levels have relatively similar standard deviations at 4.10, 4.12 and 4.16 respectively.

Based on Figure 2, the distribution of the Number of days per week an individual engages in physical activity and their Overall Subjective Happiness is bimodal. With an average of 3.69 days where individuals are physically active.

Based on Figure 3, Women generally have a higher overall subjective happiness than men, 20.82 units compared to 19.50. With woman recording a minimum of 13 and a maximum of 28 compared to men with 10 and 26 for minimum and maximum respectively. The interquartile range for overall happiness in women is 7.25 units which is higher than that of men at 4.25 units.

The hypothesis test and confidence intervals will be conducted on the fitness level and Overall Subjective happiness variables.

All analysis for this report was programmed using `R version 4.0.2`.

Methods

The fitness level variable is dichotomized, and cases are grouped into High and low. The Perfect and Good fitness levels are classified as high. The Average and Poor fitness levels are classified as Low. Using the data with the new grouped cases of fitness level, a hypothesis test conducted to see if there is a difference in the median overall happiness between the high and low levels of fitness. The test is carried out at the 5% significance level.

The following assumptions are made, independence between the two groups because individuals from high fitness group vs low fitness group are not related. Additionally, the data form the two groups do not follow a

normal distribution. We also assume equal variances between the two groups.

The Null hypothesis H_0 is that Fitness levels do not have an effect on the average overall subjective happiness of students. The alternative hypothesis H_1 is that Fitness levels have an effect on the average overall subjective happiness of students

$$H_0 : \mu_0 = \mu_1$$
$$H_1 : \mu_0 \neq \mu_1$$

An F-test was conducted to check the assumption for equal variances between the two groups. The F-test results in a p-value of 0.84 which is greater than the significance level of 0.05. Therefore we can conclude that there is no statistically significant difference between the variances of the high level vs low level fitness groups. Hence use the equal variance assumption.

A Shapiro-Wilk normality test is conducted to verify the normality assumption. The p value for the high fitness group is $0.014 < 0.05$ significance level. Therefore the distribution of the high fitness data has a statistically significant difference from the normal distribution. Hence the normality assumption is not maintained.

Based on the assumptions we conduct a two-samples Wilcoxon rank test.(Unpaired Two-Samples T-Test in R - Easy Guides - Wiki - STHDA, n.d.)

Results

The p-value from the test is 0.015 , which is smaller than `alpha = 0.05`. Therefore we can reject the null hypothesis hypothesis and conclude the difference in the overall subjective happiness between high level and low level fitness groups is statistically significant. Based on the data, we some some relationship between overall subjective happiness and physical activity.

The 95% confidence interval for the mean overall subjective happiness is (18.89, 21.49).

It is important to note that a limitation to this study is that, it only explores the subjective well-being and happiness of individuals and doesn't take other definitions into consideration. Additionally, to obtain more reliable measure for physical fitness rather than relying on self-reports, participants can be asked to perform fitness tests. A physical fitness assessment includes measures of body composition, cardiorespiratory endurance, muscular fitness, and musculoskeletal flexibility (Wilder et al., 2006)

Bibliography

1. Grolemund, G. (2014, July 16) *Introduction to R Markdown.* RStudio. https://rmarkdown.rstudio.com/articles_intro.html. (Last Accessed: May 5, 2021)
2. Dekking, F. M., et al. (2005) *A Modern Introduction to Probability and Statistics: Understanding why and how.* Springer Science & Business Media.
3. Allaire, J.J., et. el. *References: Introduction to R Markdown.* RStudio. https://rmarkdown.rstudio.com/docs/. (Last Accessed: May 5, 2021)
4. *What is physical fitness? Human Kinetics.* (n.d.). https://us.humankinetics.com/blogs/excerpt/what-is-physical-fitness. (Last Accessed September 29, 2022,)
5. Peluso, M. A., & Guerra de Andrade, L. H. (2005). *Physical activity and mental health: the association between exercise and mood.* Clinics (Sao Paulo, Brazil), 60(1), 61–70. https://doi.org/10.1590/s1807-59322005000100012
6. Communications, N. W. (n.d.). *Physical Activity.* Www.nyu.edu. Retrieved September 29, 2022, from https://www.nyu.edu/life/safety-health-wellness/live-well-nyu/priority-areas/physical-activity.html#:~:text=Physical%20activity%20has%20been%20shown

7. Mayo Clinic. (2021, October 8). *7 Great Reasons Why Exercise Matters.* Mayo Clinic; Mayo Clinic. https://www.mayoclinic.org/healthy-lifestyle/fitness/in-depth/exercise/art-20048389

8. Shantikumar, S. (2018, August 20). *Methods of Sampling from a Population. Health Knowledge.* https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/methods-of-sampling-population

9. Wronski, L. (2020). *Why (and how!) to ask survey questions on gender identity and sexual orientation* | SurveyMonkey. SurveyMonkey. https://www.surveymonkey.com/curiosity/ask-survey-questions-sexual-orientation-gender-identity/

10. Oppong, T. (2019, October 18). *Good Social Relationships Are the Most Consistent Predictor of a Happy Life.* The Center for Compassion and Altruism Research and Education.http://ccare.stanford.edu/press_posts/good-social-relationships-are-the-most-consistent-predictor-of-a-happy-life/

11. David, S. A., Boniwell, I., & Amanda Conley Ayers. (2014). The Oxford handbook of happiness. Oxford University Press.

12. Blanchflower, D., & Oswald, A. (2007). Is Well-being U-Shaped over the Life Cycle? https://doi.org/10.3386/w12935

13. Galbicsek, C. (2019). *College Alcoholism and Binge Drinking - Alcohol Rehab Guide. Alcohol Rehab Guide.* https://www.alcoholrehabguide.org/resources/college-alcohol-abuse/

14. Physical Activity Survey. (n.d.). Www.surveymonkey.co.uk. Retrieved October 1, 2022, from https://www.surveymonkey.co.uk/r/STeesPA

15. SUBJECTIVE HAPPINESS SCALE (also known as General Happiness Scale). (n.d.). https://fetzer.org/sites/default/files/images/stories/pdf/selfmeasures/SATISFACTION-SubjectiveHappiness.pdf

16. Wilder, R. P., Greene, J. A., Winters, K. L., Long, W. B., Gubler, K., & Edlich, R. F. (2006). Physical fitness assessment: an update. Journal of Long-Term Effects of Medical Implants, 16(2), 193–204. https://doi.org/10.1615/jlongtermeffmedimplants.v16.i2.90

17. LYUBOMIRSKY, S. (2019). *Subjective Happiness Scale* | Positive Psychology Center. Upenn.edu. https://ppc.sas.upenn.edu/resources/questionnaires-researchers/subjective-happiness-scale

18. Physical fitness survey. (n.d.).* Physical Fitness Survey.* https://www.survio.com/survey-template/physical-fitness-survey

19. Physical Activity Survey. (n.d.). Www.surveymonkey.co.uk. https://www.surveymonkey.co.uk/r/STeesPA

20. Tables. (n.d.). Rmarkdown.rstudio.com.https://rmarkdown.rstudio.com/lesson-7.html

21. Unpaired Two-Samples T-test in R - Easy Guides - Wiki - STHDA. (n.d.). Www.sthda.com. http://www.sthda.com/english/wiki/unpaired-two-samples-t-test-in-r

22. Ph.D, L. K. (n.d.). Practice 9 Calculating Confidence Intervals in R | R Practices for Learning Statistics. In bookdown.org. https://bookdown.org/logan_kelly/r_practice/p09.html

Appendix

Here is a glimpse of the data set simulated/surveyed:

```
## Rows: 42
## Columns: 12
## $ Gender              <chr> "Man", "Man", "Man", "Woman", "Woman", "Man", "Man~
## $ Age                 <int> 22, 24, 23, 22, 24, 22, 25, 23, 20, 22, 25, 24, 23~
## $ Relationship_Status <chr> "Cohabitating with a significant other or in a dom~
## $ Fitness_Level       <chr> "Average", "Good", "Perfect", "Average", "Average"~
## $ Alcohol_drugs       <chr> "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", "No", "Y~
## $ phys_active_perweek <int> 1, 5, 6, 4, 3, 3, 3, 3, 1, 1, 5, 7, 4, 2, 5, 5, 3,~
## $ h_self              <int> 7, 4, 7, 5, 7, 6, 4, 4, 5, 7, 5, 4, 3, 5, 4, 5, 5,~
## $ h_comp_peers        <int> 7, 3, 6, 6, 7, 6, 4, 1, 4, 5, 5, 6, 3, 6, 4, 5, 5,~
## $ h_generally         <int> 7, 2, 6, 6, 7, 6, 6, 5, 4, 2, 5, 2, 1, 7, 4, 5, 4,~
## $ unh_generally       <int> 4, 6, 5, 2, 1, 6, 3, 2, 5, 6, 5, 6, 5, 3, 4, 5, 2,~
## $ reverse_score_Q4    <int> 4, 2, 3, 6, 7, 2, 5, 6, 3, 2, 3, 2, 3, 5, 4, 3, 6,~
## $ overall_SH          <int> 25, 11, 22, 23, 28, 20, 19, 16, 16, 16, 18, 14, 10~
```