Study of Proportion of Students Utilizing Campus Fitness Center

Haobo Wang, Xiaoyu Wang, Jiamin Zheng, Cami Cai, Yi Shan

2024 Dec 2

1 Abstract

In general, many academic institutions offer gym services, which are more accessible to students. However, not all students choose to frequently use the campus fitness center. Many factors might influence their frequency of utilizing the campus fitness. The factors could probably be related to parking, commuting time, gym facilities, complementary services, and so on. Without focusing on how good the campus fitness is, the demographic differences (age, gender, major) may also determine their frequency of use.

This study examines the essential factors that influence the utilization rate of the school gym and whether different demographics affect the use of the campus fitness center. In a third-year statistics survey and sampling course, we asked students for their thoughts on the campus fitness center through an online Google form during the lectures' breaks. The result indicates no significant difference in the use of the fitness center between different types of students. Also, it suggests that facility variety is the crucial factor associated with gym frequency. In response, we plan to call on schools to provide more gym services to encourage more students to use the campus gym.

2 Introduction

Nowadays many people choose to exercise in their free time. Staying healthy and having a fit body has become their goals. Rather than exercising alone,

many people prefer going to the gym, using more professional sports equipment and making friends at the same time. Our campus, UTM, has its gym, which is open to students. Many students think it is a convenient place that provides most of the services offered in gyms, but some students have never used it for various reasons.

In this study, we examine the proportion of students utilizing the campus fitness center. We distributed questionnaires both in person and in Piazza to collect data on the use of the campus gym among STA304H5 students enrolled in Fall 2024. The questionnaire also includes influencing factors and demographic differences. We aim to study the following research questions:

- **(RQ0)** What are the differences in the percentage of different types of students using the school fitness center?
- **(RQ1)** What factors influence the utilization rate of the school gym?
  - We examined 4 factors and separated RQ1 into 4 smaller questions.
  - **(RQ1.1)** Does commuting time affect the utilization rate of the school gym?
    - H_0: commuting time has no connection to using the school fitness center.
    - H_a: commuting time is associated with using the school fitness center.
  - **(RQ1.2)** Does using an off-campus gym affect the utilization rate of the school gym?
    - H_0: using an off-campus gym has no connection to using the school fitness center.
    - H_a: using an off-campus gym is associated with using the school fitness center.
  - **(RQ1.3)** Does the variety of equipment and facilities in the campus fitness center affect the utilization rate of the school gym?
    - H_0: the variety of equipment and facilities in the campus fitness center has no connection to using the school fitness center.
    - H_a: the variety of equipment and facilities in the campus fitness center is associated with using the school fitness center.

  - **(RQ1.4)** Does the number of courses taken by students affect the utilization rate of the school gym?
    * H_0: The number of courses taken by students has no connection to using the school fitness center.
    * H_a: The number of courses taken by students is associated with using the school fitness center.

- **(RQ2)** What are the demographic differences among students who frequently utilize the campus fitness center?
  - Similar to RQ1, we separated it into 3 questions.
  - **(RQ2.1)** Does gender affect the utilization rate of the school gym?
    * H_0: Gender has no connection to using the school fitness center.
    * H_a: Gender is associated with using the school fitness center.
  - **(RQ2.2)**Does major affect the utilization rate of the school gym?
    * H_0: Major has no connection to using the school fitness center.
    * H_a: Major is associated with using the school fitness center.
  - **(RQ2.3)** Does status affect the utilization rate of the school gym?
    * H_0: Status has no connection to using the school fitness center.
    * H_a: Status is associated with using the school fitness center.

The structure of the paper is as follows: Section 1 is the abstract of the content of the entire study. Section 2 introduces the topic and the research questions of our study. Section 3 shows how data is collected. Section 4 detailed analyzes the result and answers the research questions. Section 5 interprets and discusses the result. In section 6, the limitations are argued. Section 7 concludes the result of this study. Section 8 covers the R code we used to examine the data.

3 Methodology

During October 2024, We collected information on Fall 2024 STA304 students' use of the campus Fitness Center using a questionnaire posted on

Piazza. From the 111 valid responses collected ($N = 200$), we randomly selected 60 students ($n = 60$) from R using simple random sampling (SRS). Compared to stratified and convenience sampling, simple random sampling (SRS) ensures that each individual in the sample has an equal chance of being selected, which reduces selection bias and makes the sample more representative, thus enhancing the reliability of the results. The SRS was used to ensure randomness and fairness in the sampling process, and the consistency of results was ensured by setting a random seed. The survey inquired about the average weekly frequency of student's use of the campus gym, students' commute time to campus, and information on demographic factors (e.g., age, gender, major).

4 Analysis

We ensured the randomness of the sampling by using a random seed and each participant received a separate questionnaire, and each person answering the questionnaire was required to fill in a specific ID to take responsibility for the authenticity of the questionnaire they answered, thus ensuring the independence of the responses. In addition, the sample size n = 60 ($N = 200$) was large enough. Here it is assumed that the proportion of people who use the gym is as large as the proportion of people who do not use the gym. We set a small bound of error $B = 0.1083$, and calculated the sample size,

$$n = \frac{Npq}{(N-1)D + pq}$$

with $q = 1 - p$ and $D = \frac{B^2}{4}$, we have:

$$n = \frac{200 \cdot 0.5 \cdot 0.5}{199 \cdot \frac{0.1083^2}{4} + 0.5 \cdot 0.5} = 59.9871 \approx 60$$

So, after receiving 111 valid responses, we randomly selected data from 60 students for analysis by using R. This method is intended to simulate simple random sampling. We found that approximately 38% ($n = 23$) of participants were female and approximately 62% ($n = 37$) were male. Coincidentally, there are exactly 50% ($n = 30$) of domestic students and 50% ($n = 30$) of international students. 35 of these 60 participants had not been to the school fitness center once, only 25 of them had been there once a week and more.

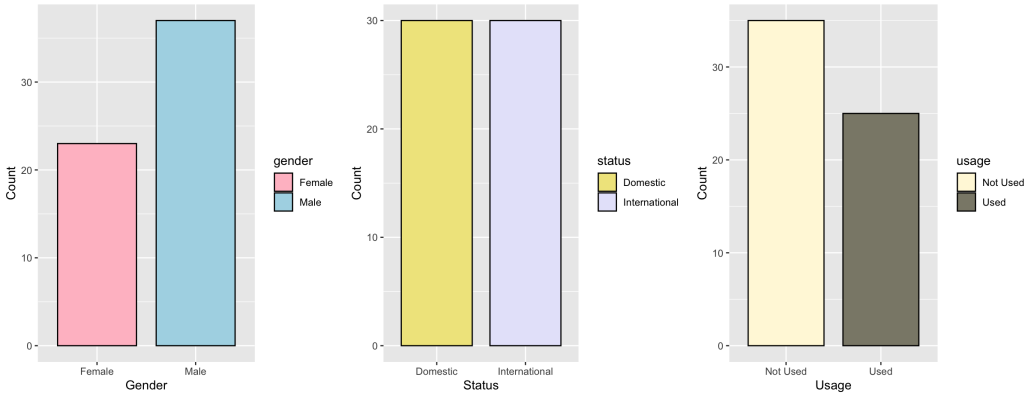

Figure 1: Bar Plot for Gender, Status and Usage Distribution in Sample

The above can answer **Research Question 0**.

RQ1.1 Does commuting time affect the utilization rate of the school gym?

The samples are independent from each other and are obtained randomly by SRS and random Seeds. The samples are not normally distributed since the normality assumption was violated as shown by Shapiro-Wilk tests ($W = 0.684, p < 0.001$ for "Used" group and $W = 0.814, p < 0.001$ for "Not Used" group). The variances for the two independent groups are equal.(Bartlett's test: $p = 0.2256$; Levene's test: $p = 0.1401$)

Because the normality assumption was not satisfied, a Wilcoxon Rank Sum Test was conducted as a non-parametric alternative for the t-test, which indicated that commute time did not have a significant effect on the utilization rate of the school gym ($W = 506$, $p = 0.261$).

RQ1.2 Does using an off-campus gym affect the utilization rate of the school gym?

The assumptions for the chi-square test were verified before conducting the analysis. The data were randomly sampled and responses were independent, ensuring the independence assumption was met. Additionally, all expected

frequencies in the contingency table were greater than 5, satisfying the assumption of sufficient cell counts. The results of the chi-square test indicated that there was no significant association between the use of off-campus gyms and the use of campus fitness centers ($\chi^2 = 0.048$, $p = 0.827$).

RQ1.3 Does the variety of equipment and facilities in the campus fitness center affect the utilization rate of the school gym?

The assumptions for the chi-square test were partially violated as one cell in the contingency table had an expected frequency less than 5 (4.17). To address this, Fisher's Exact Test was conducted as an alternative. The results indicated a significant association between the perception of facility variety and the utilization of the campus fitness center ($p = 0.0147$). This indicating that students who perceive facility variety as adequate are more likely to use the fitness center.

RQ1.4 Does the number of courses taken by students affect the the utilization rate of the school gym?

The samples are independent from each other and are obtained randomly by SRS and random Seeds. The samples are not normally distributed since the normality assumption was violated as shown by Shapiro-Wilk tests ($W = 0.684, p < 0.001$ for "Used" group and $W = 0.814, p < 0.001$ for "Not Used" group).The variances for the two independent groups are equal(Bartlett's test: $p = 0.199$; Levene's test: $p = 0.4089$).

Because the normality assumption for the t-test was not satisfied, a Wilcoxon Rank Sum Test was conducted as a non-parametric alternative, which indicated that number of courses taken by students did not have a significant effect on the utilization rate of the school gym ($W = 410.5$, $p = 0.5216$).

RQ2.1 Does gender affect the utilization rate of the school gym?

The assumptions for the chi-square test were verified before conducting the analysis. The data were randomly sampled and responses were independent,

ensuring the independence assumption was met. Additionally, all expected frequencies in the contingency table were greater than 5, satisfying the assumption of sufficient cell counts. The results of the chi-square test showed no significant association between gender and whether or not they used the campus fitness center ($\chi^2 = 1.259$, $p = 0.2618$), indicating that there was no significant difference between males and females in whether or not they used the fitness center.

RQ2.2 Does major affect the utilization rate of the school gym?

We performed a box plot to visualize the distribution of the gym utilization rate of each major:

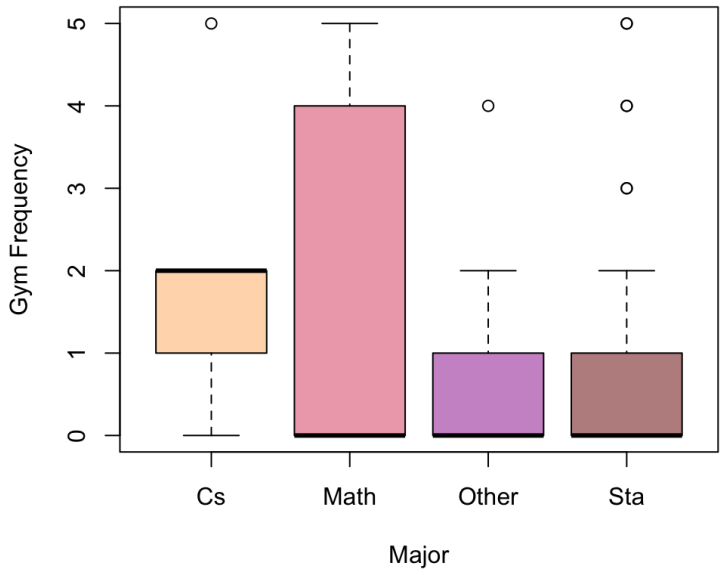

Figure 2: Boxplot of Gym Frequency by Major

The data were randomly sampled and responses were independent, ensuring the independence assumption was met. The samples are not normally

distributed since the normality assumption was violated as shown by Shapiro-Wilk tests ($W = 0.7416, p > 0.05$ for the "Math" group, $W = 0.82162, p > 0.05$ for the "Cs" group, $W = 0.6395, p < 0.05$ for the "Sta" group and $0.69383, p < 0.05$ for the "Other" group,). To verify the common variance assumption, we performed Levene's Test for Homogeneity of Variance,

$$\text{Levene's Test Results: } F = 2.2377,\ p = 0.1401$$

The p-value is greater than the significance level ($\alpha = 0.05$), indicating that we failed to reject the null hypothesis of equal variances. Thus, the assumption of common variance holds for our analysis.

However, as the assumptions for the ANOVA test were violated, we did the one-way ANOVA, Kruskal–Wallis test instead. This test shows there's no significant difference between different majors (Kruskal-Wallis $\chi^2 = 3.7674, p = 0.2877$) and here are the means of the utilization rate of gym use per week by students of each major:

Major	**Cs**	**Math**	**Other**	**Sta**
Mean	1.8571429	1.8000000	0.8000000	0.8684211

Table 1: Average Weekly Gym Utilization Rates by Major

The assumptions for the chi-square test were partially violated as many cells in the contingency table had an expected frequency less than 5. To address this, Fisher's Exact Test was conducted as an alternative. The results indicated no significant association between the major and the utilization of the campus fitness center ($p = 0.4338$).

RQ2.3 Does status affect the utilization rate of the school gym?

The assumptions for the chi-square test were verified before conducting the analysis. The data were randomly sampled and responses were independent, ensuring the independence assumption was met. Additionally, all expected frequencies in the contingency table were greater than 5, satisfying the assumption of sufficient cell counts.

The results of the chi-square test showed no significant association between status (international or domestic student) and whether or not they

used the campus fitness center ($\chi^2 = 0$, $p = 1$), indicating that there was not any significant difference between the two groups of students in terms of their use of the campus fitness center.

5 Results

In this study we used the chi-square test; the Wilcoxon Rank Sum Test, Fisher's test, and the Kruskal-Wallis test were also used when the assumptions for the t-test, the chi-square test, and the ANOVA test were violated. With such statistical tests, we explore the effects of different factors on campus fitness center usage and frequency. The results indicated that most factors (e.g., commute time to school, use of off-campus gyms, major, and status) did not have a significant influence on fitness center frequency and use. However, subjective evaluation of the variety of fitness center equipment influenced the usage of fitness centers. Specifically, students who think there is enough variety are more likely to use the fitness center.

For RQ1, factors influence the utilization rate of the school gym, we examined commute time to school, use of off-campus gyms, equipment variety, and courses taken per semester. Within these four factors, we find that there is only an association between equipment variety and utilization rate of the school gym, but no connection with the rest three factors. This result shows participants may focus on the experience when working out, but not additional factors.

Some past studies have shown that commute time will affect the attendance of students and workers. Longer commuting times are associated with lower attendance rates [1–3]. From these, we assumed that commuting time may also associated with the use rate of the campus fitness center. However, we didn't find any association between commuting time and the utilization rate of the campus fitness center.

All the 3 demographic factors we focused on do not have an association with the use of campus gym. The public generally believes that males are more motivated to exercise [4–6], so we assumed that there would be a correlation between gender and the use of the campus gym. However, in our study, the sample shows gender is not associated with the use of the campus gym. Still, many studies have found that gender is not related to motivation, participation, or frequency of exercise in adolescents or adults [7–9]. These inconsistent findings show that the association between gender and exercise

may be more complicated. For our study, further research would also better demonstrate the correlation between gender and the use of campus gyms.

6 Limitation

In the questionnaire, we raised some leading questions. For example, in RQ1, we considered parking availability and fees as an external factor of the utilization proportion. To examine this, we set this question:

> *Do you think the availability and price of parking will affect your visit to the fitness center on campus?*

When we do the analysis, we look back at this question and find it was too leading. This question directly correlates parking availability and fees with the use of the campus gym, which may lead the participants to choose "Yes" or "Maybe". Considering this, we ignored this factor in our research. In the future, we need to better revise the questions in surveys, in order to avoid similar errors.

During the research process, we discovered that the study population was not representative enough to adequately reflect the broader student population. The questionnaire was targeted at students in STA 304 and thus lacked coverage of other majors and grades, which made it difficult to fully reflect overall student fitness behaviors and usage habits. Our interest group in this study primarily consists of statistics students, who may be less likely to frequent the gym compared to students majoring in fields such as nutrition, kinesiology, or health sciences. Students in these disciplines above are often more health-conscious and motivated to improve their cardiovascular fitness through aerobic exercises (e.g., treadmill, elliptical) and build strength through gym-based training. If our sample included more students from these majors, the results might vary significantly. To address these limitations, future studies should focus on ensuring question neutrality and increasing the diversity of the sample to enhance representativeness and minimize potential response bias.[10]

7 Conclusion

In this study, we evaluated the effect of different factors on students' use of campus fitness centers. The analysis revealed that most of the external

factors (e.g. commute time, off-campus gym access, etc.) and demographic factors (e.g. major and status) did not significantly affect the use of the campus fitness center. However, significant associations were found regarding the variety of equipment types. The findings of this study on campus fitness show that increasing the variety of equipment may help attract more students to come to the campus fitness center. Additionally, for gender differences, further research is needed to provide a better experience for students of different genders. Including more students from different years of study and majors should also develop interventions to stimulate their interest in fitness. This will support the development of more scientific and accurate campus wellness policies. Multidimensional variables and more comprehensive measures could also be included.

8 Appendix

8.1 Codes[11]

```
# Loading libraries
library(car)
library(ggplot2)
library(gridExtra)


# Loading data
STA304_Group9_Data_Cleaned
<- read.csv("C:/Users/Haobo Wang/Desktop/STA304_Group9_Data_Cleaned.csv")
View(STA304_Group9_Data_Cleaned)


# Random sampling
set.seed(304)
sample_data <- STA304_Group9_Data_Cleaned[sample(nrow(STA304_Group9_Data_Cleaned), 60), ]
View(sample_data)

# Gender analysis
gender_counts <- table(sample_data$gender)
print(gender_counts)

# Creating a barplot
gender_plot <- ggplot(sample_data, aes(x = gender, fill = gender)) +
  geom_bar(color = "black", width = 0.8) +
  scale_fill_manual(values = c("Male" = "lightblue", "Female" = "pink")) +
  labs(x = "Gender", y = "Count")


# Status analysis
status_counts <- table(sample_data$status)
print(status_counts)

# Creating a barplot
status_plot <- ggplot(sample_data, aes(x = status, fill = status)) +
```

```
  geom_bar(color = "black", width = 0.8) +
  scale_fill_manual(values = c("D" = "khaki", "I" = "lavender"),
                    labels = c("D" = "Domestic", "I" = "International")) +
  labs(x = "Status", y = "Count") +
  scale_x_discrete(labels = c("D" = "Domestic", "I" = "International"))


# Frequency analysis
frequency_counts <- table(sample_data$gym_frequency)
print(frequency_counts)
sample_data$campus_gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

# Creating a barplot
usage_plot <- ggplot(sample_data, aes(x = campus_gym_use, fill = campus_gym_use)) +
  geom_bar(color = "black", width = 0.8) +
  scale_fill_manual(values = c("Not Used" = "cornsilk", "Used" = "cornsilk4")) +
  labs(x = "Usage", y = "Count", fill = "usage")

# Combine barplots
grid.arrange(gender_plot, status_plot, usage_plot, nrow = 1)


# Common Var analysis
sample_data$gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

levene_test_result <- leveneTest(commute_time ~ gym_use, data = sample_data)
print(levene_test_result)


# Commute time vs gym usage analysis:
sample_data$gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")
group1 <- sample_data$commute_time[sample_data$gym_use == "Used"]
group2 <- sample_data$commute_time[sample_data$gym_use == "Not Used"]

shapiro_test_group1 <- shapiro.test(group1)
shapiro_test_group2 <- shapiro.test(group2)
print(shapiro_test_group1)
print(shapiro_test_group2)
```

```
bartlett_test <- bartlett.test(commute_time ~ gym_use, data = sample_data)
print(bartlett_test)

levene_test <- leveneTest(commute_time ~ gym_use, data = sample_data)
print(levene_test)

wilcox_test_result <- wilcox.test(commute_time ~ gym_use, data = sample_data)
print(wilcox_test_result)


# Off-campus gym usage vs gym frequency analysis
sample_data$campus_gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

table_off_campus <- table(sample_data$off_campus_gym, sample_data$campus_gym_use)
chi_off_campus <- chisq.test(table_off_campus)

print(chi_off_campus$expected)
print(chi_off_campus)


# Equipment variety vs gym frequency analysis
sample_data$campus_gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

table_variety <- table(sample_data$enough_variety, sample_data$campus_gym_use)
chi_variety <- chisq.test(table_variety)

print(chi_variety$expected)
print(chi_variety)

fisher_variety <- fisher.test(table_variety)
print(fisher_variety)


# Courses taking vs gym frequency t-test analysis
sample_data$gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

bartlett_test <- bartlett.test(courses_per_semester ~ gym_use, data = sample_data)
```

```
print(bartlett_test)

levene_test <- leveneTest(courses_per_semester ~ gym_use, data = sample_data)
print(levene_test)

wilcox_test_result <- wilcox.test(courses_per_semester ~ gym_use, data = sample_data)
print(wilcox_test_result)


# Gender vs gym frequency chi-square analysis
sample_data$campus_gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

table_gender <- table(sample_data$gender, sample_data$campus_gym_use)
chi_gender <- chisq.test(table_gender)
print(chi_gender$expected)
print(chi_gender)


# Major vs gym frequency chi-square analysis
sample_data$campus_gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")

table_major <- table(sample_data$major, sample_data$campus_gym_use)
chi_major <- chisq.test(table_major)
print(chi_major$expected)
print(chi_major)

fisher_major <- fisher.test(table_major)
print(fisher_major)

math_group <- sample_data$gym_frequency[sample_data$major == "Math"]
cs_group <- sample_data$gym_frequency[sample_data$major == "Cs"]
sta_group <- sample_data$gym_frequency[sample_data$major == "Sta"]
other_group <- sample_data$gym_frequency[sample_data$major == "Other"]

shapiro_test_math <- shapiro.test(math_group)
shapiro_test_cs <- shapiro.test(cs_group)
shapiro_test_sta <- shapiro.test(sta_group)
shapiro_test_other <- shapiro.test(other_group)
```

```
print(shapiro_test_math)
print(shapiro_test_cs)
print(shapiro_test_sta)
print(shapiro_test_other)

kruskal.test(gym_frequency ~ major, data = sample_data)
tapply(sample_data$gym_frequency, sample_data$major, mean)


#Creating a boxplot
boxplot(gym_frequency ~ major, data = sample_data, xlab = "Major", ylab = "Gym Frequency", col = c


# Status vs gym frequency chi-square analysis
sample_data$campus_gym_use <- ifelse(sample_data$gym_frequency > 0, "Used", "Not Used")
table_status <- table(sample_data$status, sample_data$campus_gym_use)
chi_status <- chisq.test(table_status)

print(chi_status$expected)
print(chi_status)

qqnorm(sample_data$gym_frequency, main = "QQ Plot of Gym Frequency")
qqline(sample_data$gym_frequency, col = "blue")
```

8.2 Reference

1. van Ommeren, J. N. & Gutiérrez-i-Puigarnau, E. Are workers with a long commute less productive? An empirical analysis of absenteeism. Reg. Sci. Urban Econ. 41, 1–8 (2011). https://www.sciencedirect.com/science/article/pii/S0166046210000633

2. Lopes, R. & Silveira-Neto, R. Commuting and School Absenteeism: Evidence from Brazil. https://api.semanticscholar.org/CorpusID:264446992

3. Coutts, S., Aird, B., Mitra, R. & Siemiatycki, M. Does commute influence post-secondary Students' social capital? A study of campus participa-

tion at four universities in Toronto, Canada. J. Transp. Geogr. 70, 172–181 (2018). https://www.sciencedirect.com/science/article/pii/S0966692318300656

4. Garcia, A. W. et al. Gender and Developmental Differences in Exercise Beliefs Among Youth and Prediction of Their Exercise Behavior. J. Sch. Health 65, 213–219 (1995). https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1746-1561.1995.tb03365.x

5. Portela-Pino, I., López-Castedo, A., Martínez-Patiño, M. J., Valverde-Esteve, T. & Domínguez-Alonso, J. Gender Differences in Motivation and Barriers for The Practice of Physical Exercise in Adolescence. Int. J. Environ. Res. Public. Health 17, 168 (2020). https://www.mdpi.com/1660-4601/17/1/168

6. Lauderdale, M. E., Yli-Piipari, S., Irwin, C. C. & Layne, T. E. Gender Differences Regarding Motivation for Physical Activity Among College Students: A Self-Determination Approach. Phys. Educ. 72, (2015). https://js.sagamorepub.com/index.php/pe/article/view/4682

7. Davis, C., Fox, J., Brewer, H. & Ratusny, D. Motivations to exercise as a function of personality characteristics, age, and gender. Personal. Individ. Differ. 19, 165–174 (1995). https://www.sciencedirect.com/science/article/pii/019188699500030A

8. Mao, H.-Y., Hsu, H.-C. & Lee, S.-D. Gender differences in related influential factors of regular exercise behavior among people in Taiwan in 2007: A cross-sectional study. PLOS ONE 15, e0228191 (2020). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0228191

9. Jakobsen, A. & Evjen, E. Gender differences in motives for participation in sports and exercise among Norwegian adolescents. Balt. J. Health Phys. Act. 10, (2022). https://www.balticsportscience.com/journal/vol10/iss2/10

10. Sample bias and Diversity. ChatGPT (2024). Available at: https://chatgpt.com/share/674e8591-44c8-8004-a0bc-0572d52e3aec. (Accessed: 2nd December 2024)

11. Box plot in R. ChatGPT (2024). Available at: https://chatgpt.com/share/674e8551-a5fc-800c-b1fe-8f6a8b94abab. (Accessed: 2nd December 2024)