Introduction

With COVID-19 diminishing and universities reopening to offer lectures in person, courses might choose to stop certain practices such as lecture recordings. However, such practices can be an important part of the toolkit for students' learning. This study will examine the impact of using online resources during the duration of a course, and a student's self-reported academic satisfaction. Additionally, we aim to answer whether utilizing more online resources correlates with better academic satisfaction. In general, online resources pertain to a specific course providing additional resources to supplement learning from traditional lectures that are accessed over the web. These resources include recordings of lectures and/or tutorials, having a Piazza forum, annotated slides, and past tests or exams. Understanding how online resources can impact learning will be important as universities return to normalcy.

Hypotheses

To determine if using online resources impact a student's view of their own academic performance, we define the hypothesis as follows for a Mann-Whitney U Test:

H_0: There is no difference between the ranks of academic satisfaction between Groups 1 and 2

H_A: There is difference between the ranks of academic satisfaction between Groups 1 and 2

Where Group 1 is Engagement-study with Online Learning Resources and Group 2 is Engagement-Study only.

Furthermore, we want to investigate the correlation between using professor edited materials (Annotated Lecture Slides/Tutorial Notes, Lecture/Zoom Recordings, and Piazza/Discussion Forums) and reported academic satisfaction. We define the hypothesis as follows:

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$

H_a: at least one $\beta_i \neq 0$

Let x_i be independent variables with:

- x_1 is the impact of **Annotated Lecture Slides/Tutorial Notes**
- x_2 is the impact of **Lecture/Zoom Recordings**
- x_3 is the impact of **Piazza/Discussion Forums**

Let β_i being the corresponding slope for each x_i in a multiple linear regression model $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$, where Y is the reported academic satisfaction.

Methodology

Between October 2023 and November 2023, The data was collected between October 2023 and November 2023 using Piazza as the platform for distributing the questionnaire to STA304H5 F 2023 students. The sampling method used was Simple Random Sampling (SRS), achieved by posting the questionnaire on Piazza. By posting our questionnaire to Piazza, each student in the population had an equal chance of being selected for the survey if they decided to participate. Out of N = 200 students, 30 students decided to fill out our questionnaire, therefore we are left with n = 30 responses. The survey asked for demographic information (gender, program of study, etc.), rating effectiveness of a couple online learning resources (zoom/lecture recordings, piazza, etc.), the participants preferred mode of learning (self-study, engagement-study, engagement-study with online learning resources), what online resources they use before lecture and after lecture, along with their academic satisfaction.

Analysis

To calculate our bound on the mean of academic satisfaction for students who use OLS and engagement learning given N = 200, n = 21, and $s_2^2 = 0.3619$, we use the simple random sampling sample size for 95% CI formula and isolate for B. Then,

$$n = 21 = \frac{Ns_2^2}{(N-1)D + s_2^2} = \frac{200(0.3619)}{(200-1)D + 0.3619} = \frac{72.38}{199D + 0.3619}$$

Isolating for D, we get

$$3(199D + 0.3619) = 72.38$$
$$597D + 1.0857 = 72.38$$
$$\frac{B^2}{4} = D = 0.1194$$
$$B = 0.6911$$

Per sub-category: Male = 21, Female = 9, International = 15 , Domestic = 15

Hypothesis 1 Analysis/Computations:

We want to investigate the association between learning modes and academic satisfaction. Where Group 1 is Engagement-study with Online Learning Resources and Group 2 is Engagement-Study only. We will be doing a Mann-Whitney U Test since our sample is not normally distributed.

Restating our hypothesis:

H_0: There is no difference between the ranks of academic satisfaction between Groups 1 and 2

There is difference between the ranks of academic satisfaction between Groups 1 and 2

We choose a significance level of $\alpha = 0.05$, and assumptions/conditions are met for the Mann-Whitney U Test because

1. The sample drawn from the population is random.
2. Each response is independent of each other
3. Sample independence and mutual independence is assumed.

Dataset (Academic Satisfaction):

Engagement-study	Engagement-study with Online Learning Resources
2	3
4	4
4	2
	3
	4
	4
	3
	4
	3
	4
	4
	3
	3
	4
	4
	3
	4
	4
	3
	3
	4

Ranking of the Data (w/ Ties):

Rank	Data
2	2
2	2
2	2
8	3
8	3
8	3
8	3
8	3
8	3
8	3
8	3
8	3
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4
18.5	4

Rank Sum for Group 1 (Engagement-Study w/ OLR) = 2 + 9(8) + 11(18.5) = 277.5

Rank Sum for Group 2 (Engagement-Study) = 2 + 2(4) = 10

U-statistics:

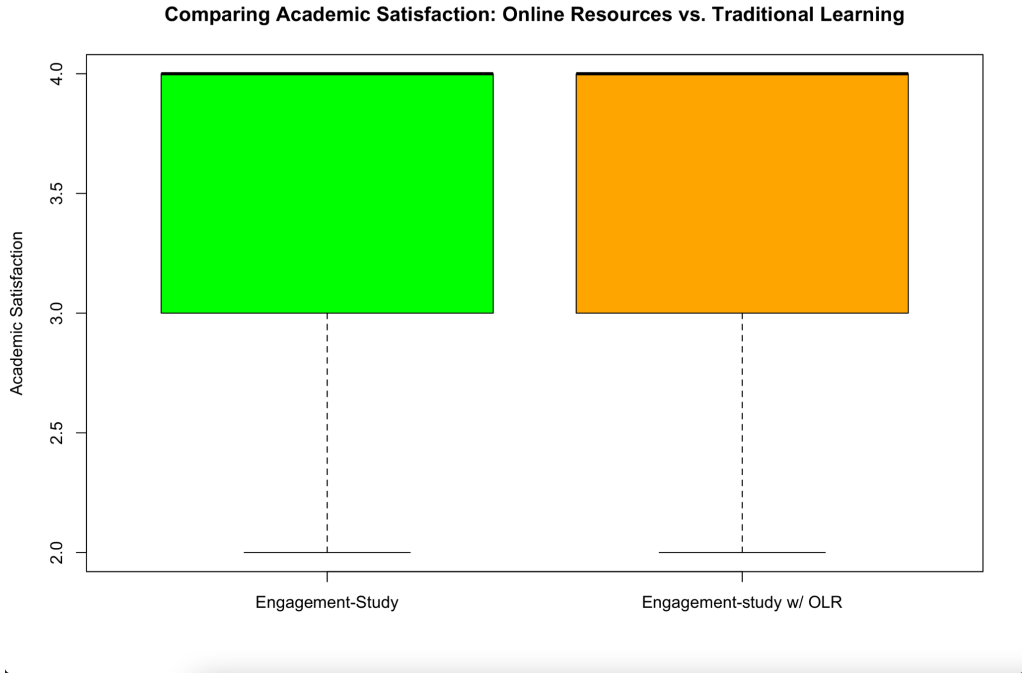
We use $n_1 = 20$ since 20 is the max value on the Mann-Whitney table is n=20

$$U_1 = \sum Rank_1 - \frac{n_1(n_1+1)}{2} = 277.5 - \frac{20(20+1)}{2} = 227.5 - 210 = 17.5$$

$$U_2 = \sum Rank_2 - \frac{n_2(n_2+1)}{2} = 10 - \frac{3(3+1)}{2} = 10 - 6 = 4$$

Let $U_{STAT} = 17.5$

Using the critical values table for Mann-Whitney U statistic, if $n_1 = 20$, $n_2 = 3$, then at the $\alpha = 0.05$ critical level, $U_{CRIT} = 8$. Since $U_{STAT} > U_{CRIT}$, we fail to reject the null hypothesis H_0 meaning that the two groups are not different in terms of academic satisfaction.

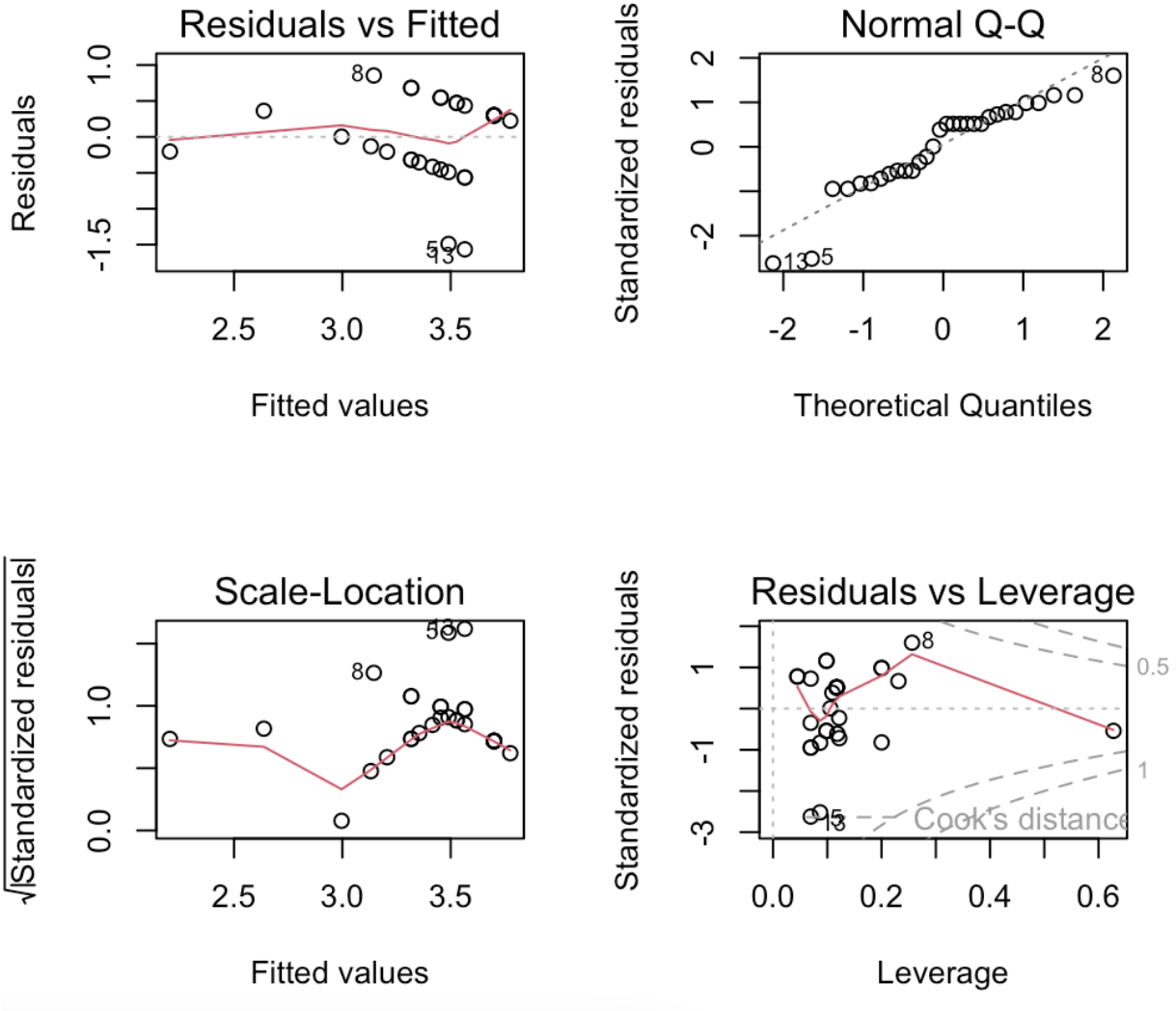


Hypothesis 2 Analysis/Computations (Advanced Methodology):

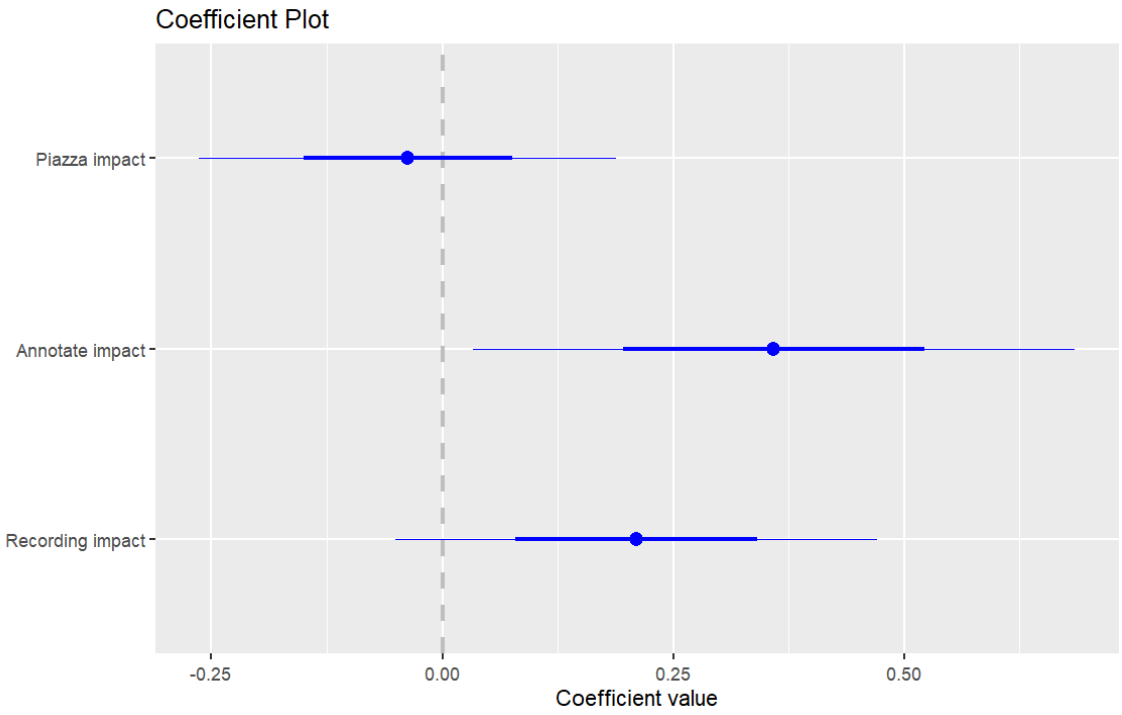
We want to investigate the correlation between professor edited materials (Annotated Lecture Slides/Tutorial Notes, Lecture/Zoom Recordings, and Piazza/Discussion Forums) on academic satisfaction as defined in the Hypotheses section above.

All Conditions/Assumptions for Linear Regression suffice such as:

- Linearity (using Residuals vs. Fitted Plot)
- Independence (using QQ Plot)
- Homoscedasticity (using Scale-Location Plot)
- Normality (using Residuals vs. Leverage Plot)



```
(Intercept)  recording_impact   annotate_impact     piazza_impact
    4.23154           0.20998           0.35878          -0.03749
```



Coefficient values for β_i using MLR

Using ordinary linear regression, we have the coefficients of:

β_0 = 4.23154

β_1 = 0.20998 p-value: 0.07041 (recording_impact)

β_2 = 0.35878 p-value: 0.03869 (annotate_impact)

β_3 = -0.03749 p-value: 0.74251 (piazza_impact)

The linear model has p-value > 0.05 for two β_i, which implies that recordings and having Piazza forum has no significant impact on reported academic satisfaction, with only annotated slides having a significant positive impact on improving course satisfaction. Since β_2 is statistically significant at p-value < 0.05, we reject null hypothesis.

Discussion/Results

Based on the calculations we have confirmed, we can conclude that for our first hypothesis, $U_{STAT} > U_{CRIT}$, we fail to reject the null hypothesis. This means engagement groups with online learning resources and engagement-study-only groups have no difference in terms of their academic satisfaction. From the second hypothesis results, we observe that certain online resources are more beneficial to perceived academic satisfaction than others. Some online resources have no significant impact, while others like annotated slides do. From both of the hypotheses, we can conclude that the impact of online resources is minimal to students' belief of their academic performance matching their expectations. We might interpret the result of both groups being statistically the same as a result of people from online learning groups not effectively using the online resources. The engagement-study group has the same academic satisfaction as the OLS study group because the engagement-study group has a more in-person focus on the material than the OLS, where students retain more by actively listening and writing notes.

Limitations

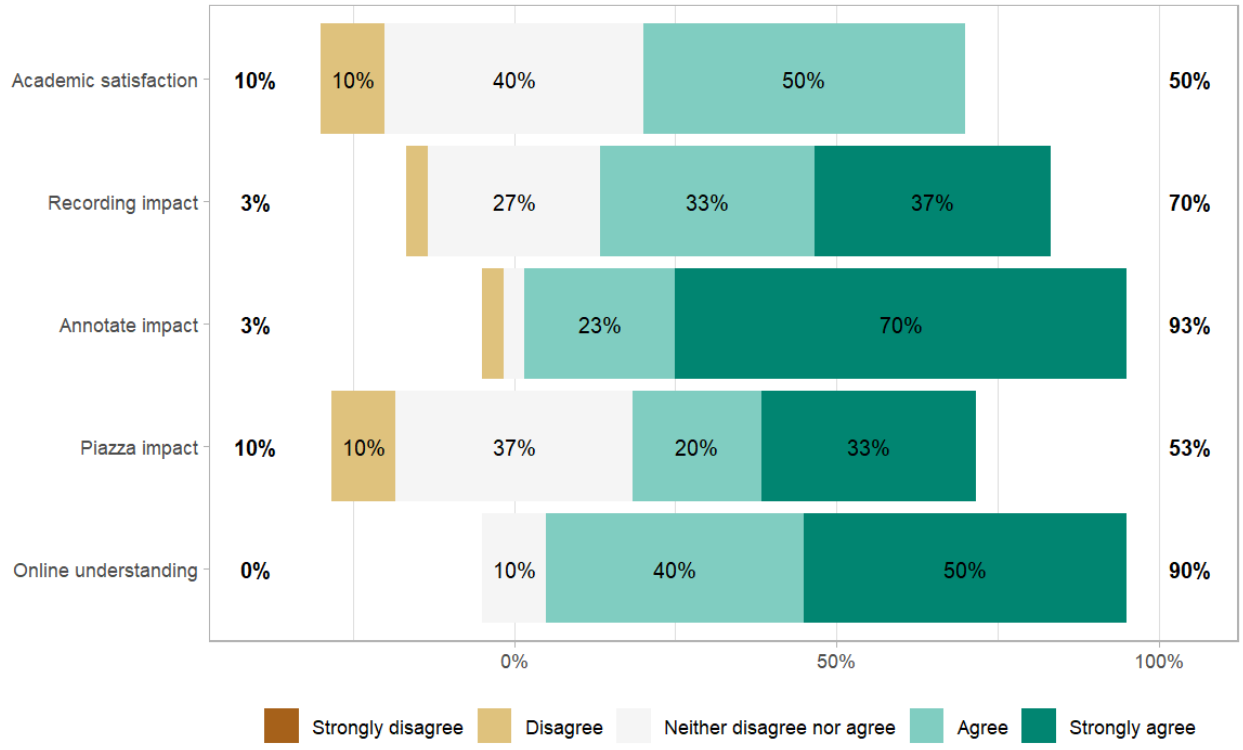
During our analysis we realized we had several data points that had insufficient data that we needed for our analysis. For example, we could have separated online understanding into individual sections for each type of online resource, in our current survey it is not clear how much influence each individual type of online resource had. Also, our mode of learning question didn't have a middle ground option, which results in our data being very skewed towards one option. As a result our analysis of the data was biased. For post-study and pre-study questions. We provided too many options, some were not relevant to our hypothesis and clogged up our responses. Lowering our

options to only ones that are relevant to our study would have helped us get more concise data.

Conclusion

Our study attempted to find out if providing online resources improved a student's self reported academic satisfaction. We reflect on the results of the study and understand that it was not performed under optimal conditions. Due to the limitation on sample sizes, we could not perform any tests that require variables to be normally distributed. Additionally, our questionnaire can be improved upon to prevent bias by adding more options for certain questions and removing unnecessary questions. We conclude that there is no difference between academic satisfaction among those who use online resources and those who do not use online resources, provided both groups attend in-person lectures and tutorials. Additionally, the more online resources a student engages with does not necessarily imply an improved belief in academic performance. More importantly, it is the kind of resource that can enhance a student’s learning, specifically having and using annotated slides as supplementary material.

Appendix



R Script:

```
library(coefplot)
library(ggstats)
library(ggplot2)
library(dplyr)
library(readxl)

#import sample data change work space
setwd("C:/Users/10713/Desktop/304GW R")
Dataset <- read_excel("Dataset.xlsx")
Tidy_data <- attach(Dataset)

############################ Hypothesis 1 Mann-Whitney Test
df <- data.frame(
  as = Tidy_data$academic_satisfaction,
  mode = Tidy_data$mode_of_learning
)

withOls_df <- filter(df, mode == 1)
withoutOls = c(2,4,4)
```

```
data <- data.frame(
  group = rep(c("withoutOls", "withOls"),
           times=c(length(withOls_df$as), length(withoutOls))),
  value = c(withOls_df$as, withoutOls)
)

wilcox.test(value ~ group, data=data)

#define groups for boxplot
engage <- Dataset$academic_satisfaction[Dataset$mode_of_learning ==
2]
engage_OLR <- Dataset$academic_satisfaction[Dataset$mode_of_learning
== 1]

# Create a boxplot
boxplot(engage, engage_OLR, names = c("Engagement-Study",
"Engagement-study w/ OLR"),
        main = "Comparing Academic Satisfaction: Online Resources vs.
Traditional Learning",
        ylab = "Academic Satisfaction",
        col = c("green", "orange"))

##########################

#set Likert level
likert_level <- c(
  "1",
  "2",
  "3",
  "4",
  "5"
)


#copy to new variable with level
df_all <- tibble(academic_satisfaction, recording_impact,
annotate_impact, piazza_impact, online_understanding) %>%
  mutate(across(everything(), ~ factor(.x, levels = likert_level ,
                                       labels = c("Strongly disagree",
                                                  "Disagree",
                                         "Neither disagree nor agree",
                                                  "Agree",
                                                 "Strongly agree"))))
```

```
#generate centered bar plot with all data
gglikert(df_all, variable_labels = c(
  academic_satisfaction = "Academic satisfaction",
  recording_impact = "Recording impact",
  annotate_impact = "Annotate impact",
  piazza_impact = "Piazza impact",
  online_understanding = "Online understanding"
))

#multiple linear regression model Hypothesis 2
model <- lm(academic_satisfaction ~ recording_impact +
annotate_impact + piazza_impact, data = Dataset)
summary(model)

# Diagnostic plots, set to 2x2 grid
par(mfrow = c(2, 2))

# 1. Residuals vs. Fitted
plot(model, which = 1)

# 2. Q-Q Plot
plot(model, which = 2)

# 3. Scale-Location Plot
plot(model, which = 3)

# 4. Residuals vs. Leverage Plot
plot(model, which = 5)

# set back to 1x1 grid
par(mfrow = c(1, 1))

#use coefficient plot to visualize multiple linear regression slope
coefplot(model, intercept = FALSE, title = "Coefficient Plot",
         xlab = "Coefficient value",
         ylab = "",
         newNames=c(academic_satisfaction = "Academic satisfaction",
                    recording_impact = "Recording impact",
                    annotate_impact = "Annotate impact",
                    piazza_impact = "Piazza impact"
                    ))
```