Final Project

Diti Shashidhar

2025-03-18

Introduction

Housing availability is a critical issue for university students, particularly at institutions like the University of California, Santa Barbara (UCSB), where high demand and limited supply create challenges in securing stable accommodations. As enrollment grows and local housing markets fluctuate, discussions about whether UCSB should guarantee on-campus housing for all students have become increasingly relevant.

The struggle for adequate student housing is not unique to UCSB. Across the country, universities are grappling with how to balance expansion, affordability, and their responsibility to both students and surrounding communities. Nishimoto (2022), in an article by Harvard university, highlights the increasing tensions between universities and local housing markets, noting that inadequate student housing can force students into long commutes, contribute to homelessness, and inflate rent prices surrounding neighborhoods. These challenges underscores the importance of examining student perspectives on housing policy, as such opinions may inform future university decisions.

This study explores how different framings of UCSB's housing policy influence student opinions. Specifically, it examines whether emphasizing student benefit, resource constraints, or presenting a neutral statement impacts agreement levels among students. Given that housing security directly affects students' well-being, academic performance, and financial stability, understanding these perspectives is crucial for policymakers. By analyzing responses across different class standings, this study also aims to determine whether student perspectives shift based on their time at UCSB.

Through a Generalized Randomized Block Design (GRBD) approach, this project investigates the extent to which wording influences perceptions of housing policy, providing insight into how public messaging can shape opinions on university policies.

Methods

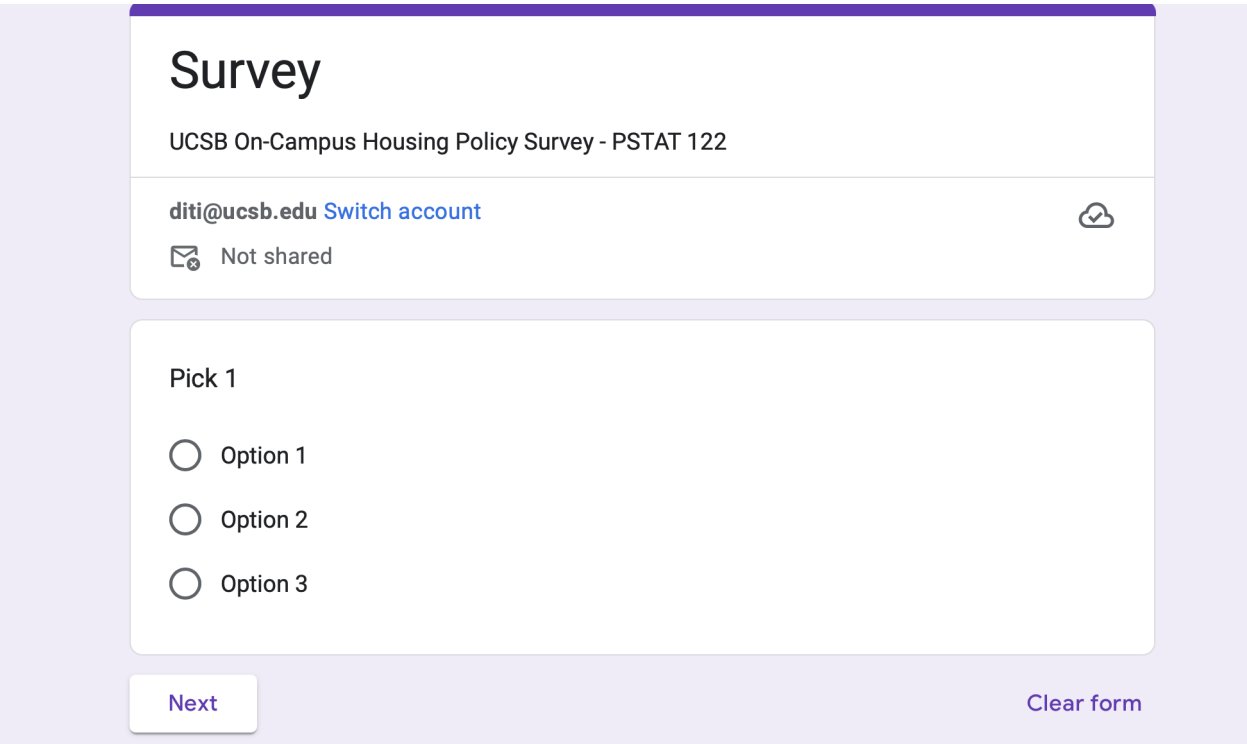
Data Collection

The data was collected through a survey. I achieved that using google forms, where I sent a link to all my friends and asked them to circulate is as well. I managed to get 27 responses. The survey contained 4 sections, the first section was where each person had to pick a version, this, although was random (since each applicant picked one), it wasn't truly randomizes (as it would've been had i used the sample function). Each version would take the person to a different section corresponding to the version they'd pick. Suppose A person picked 'Option 1', on clicking this option, the person would be taken to a unique section that had two contained two questions, one of the questions was to select their class standing (block) and the other question was specific to the option. Option 1's statement was: "Neutral - UCSB should guarantee on-campus housing for all students who request it", Option 2's statement was: "Emphasizing Student Benefit - To reduce housing insecurity, UCSB should guarantee on-campus housing for all students who want it", and Option 3's statement was: "Emphasizing resource Constraints - Due to limited space and resources, UCSB should

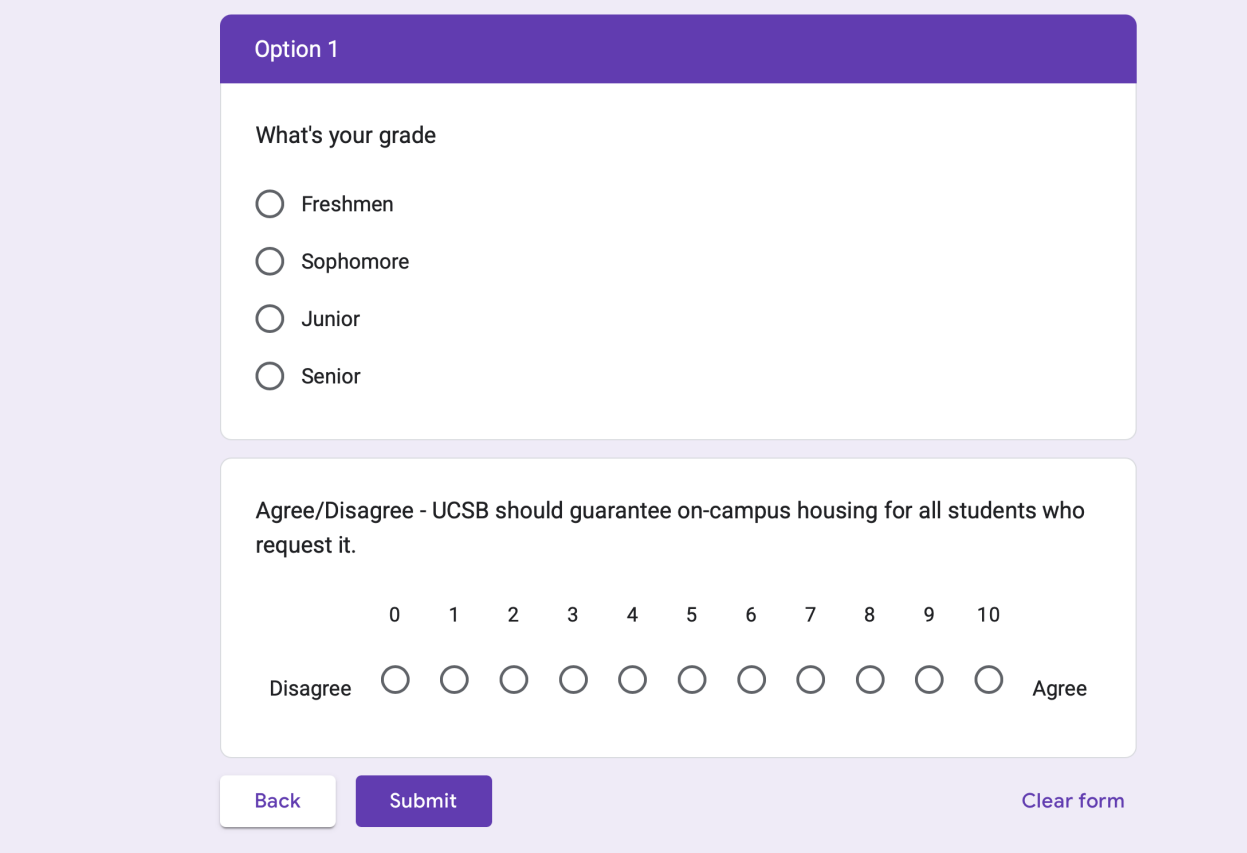
not be expected to guarantee on-campus housing for students. For each of the three statements, the student had to choose a number between 0 and 1, 0 meaning that they disagree, and 1 meaning that they agree. To summarize, the Primary Factor is the Wording of the statement - Option 1, Option 2, and Option 3 (Neutral, Student Benefit, Resource Constraints), the Blocking Variable is the Class Standing (Freshmen, Sophomore, Junior, Senior), the Outcome Variable is the Agreement Score (0 - 10). The survey was distributed via personal networks and social media.

Images of data collection process (Online Survey):

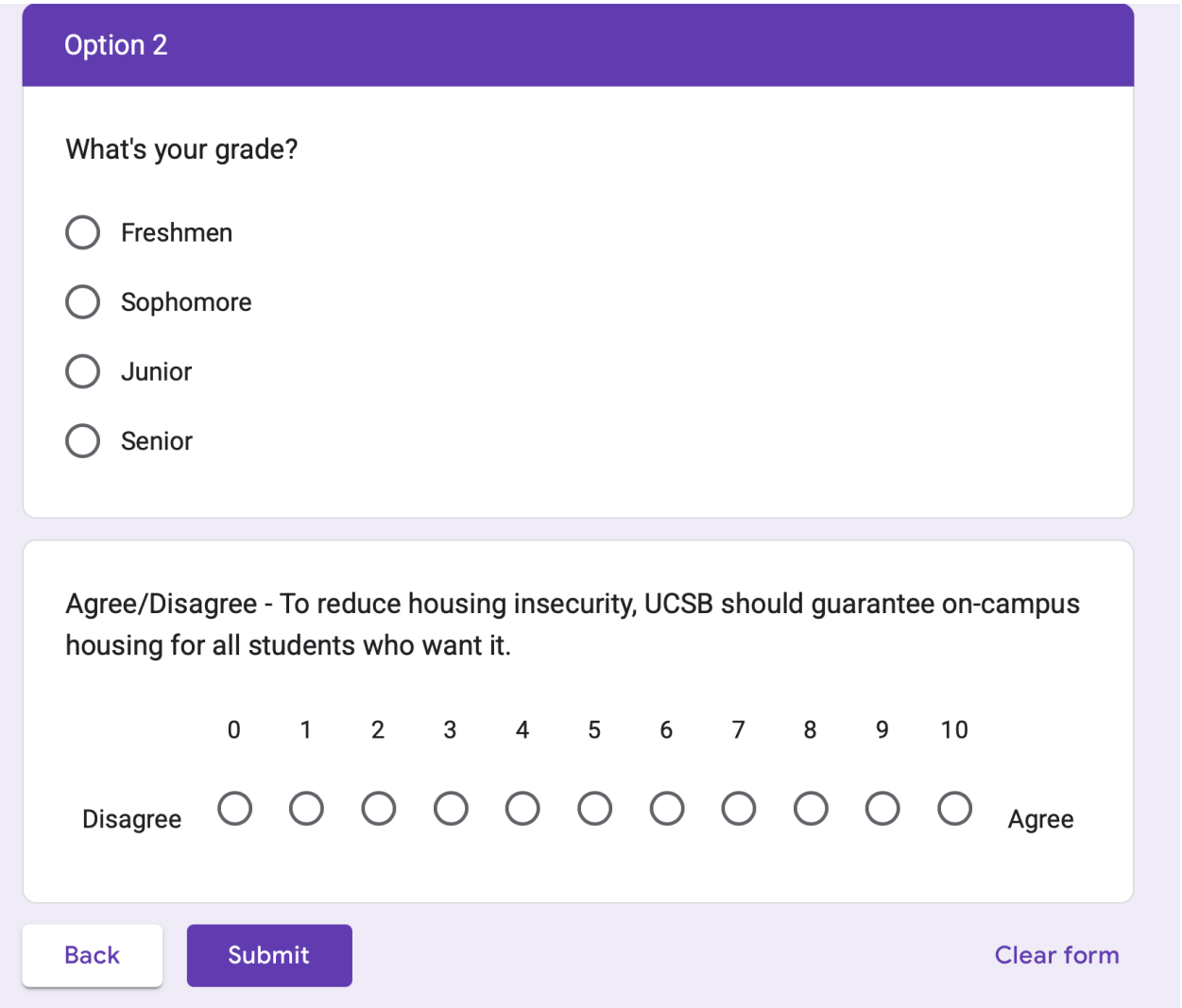
```
knitr::include_graphics("ss1.png")
```



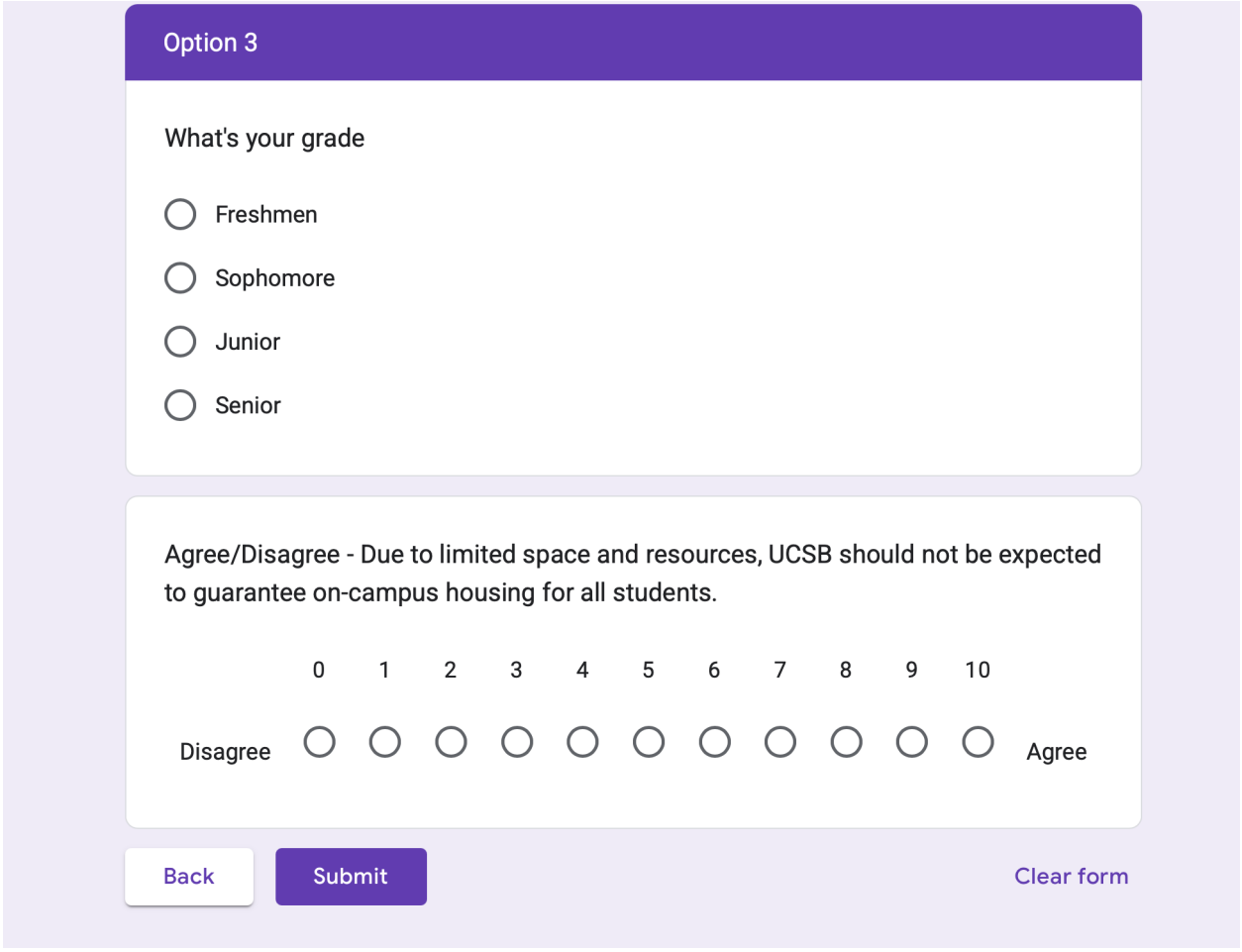

```
knitr::include_graphics("ss2.png")
```



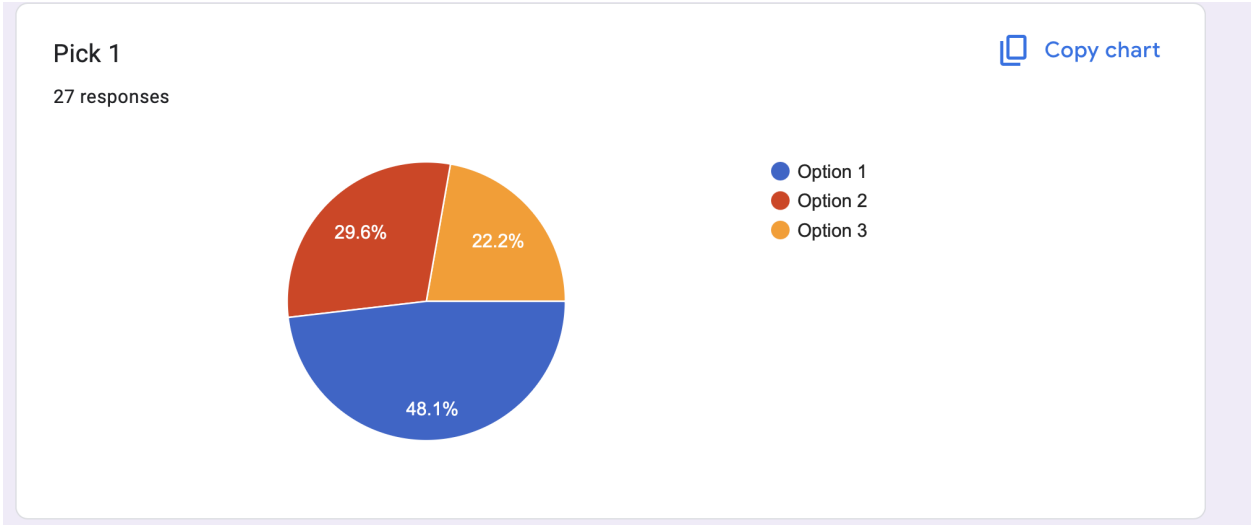

```
knitr::include_graphics("ss3.png")
```



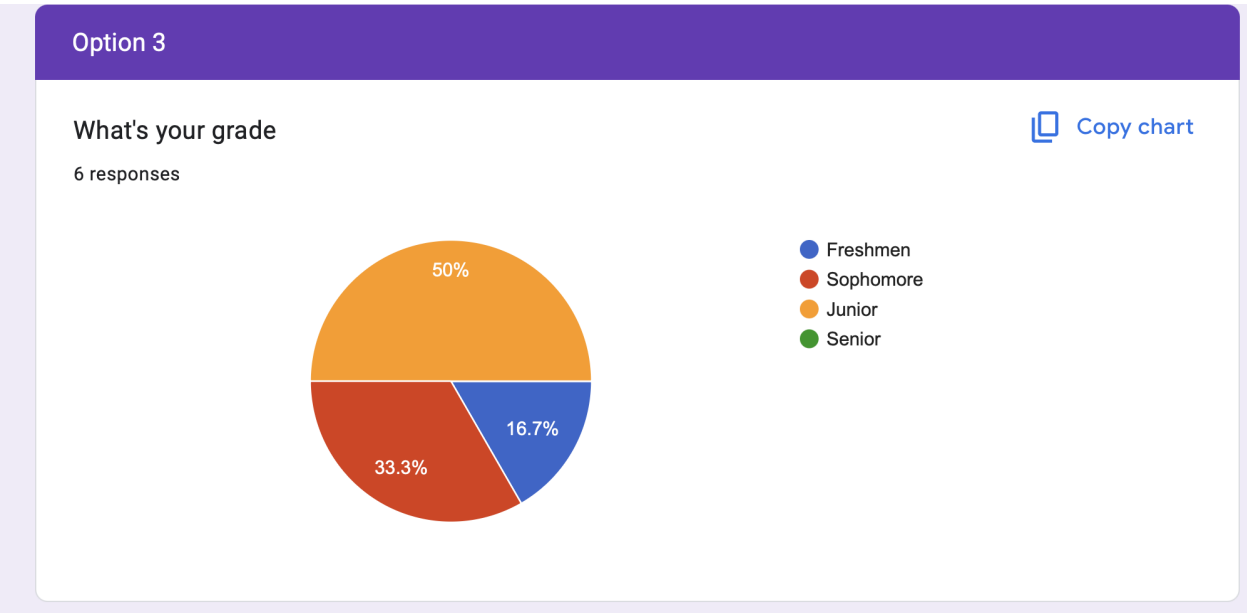

```
knitr::include_graphics("ss4.png")
```



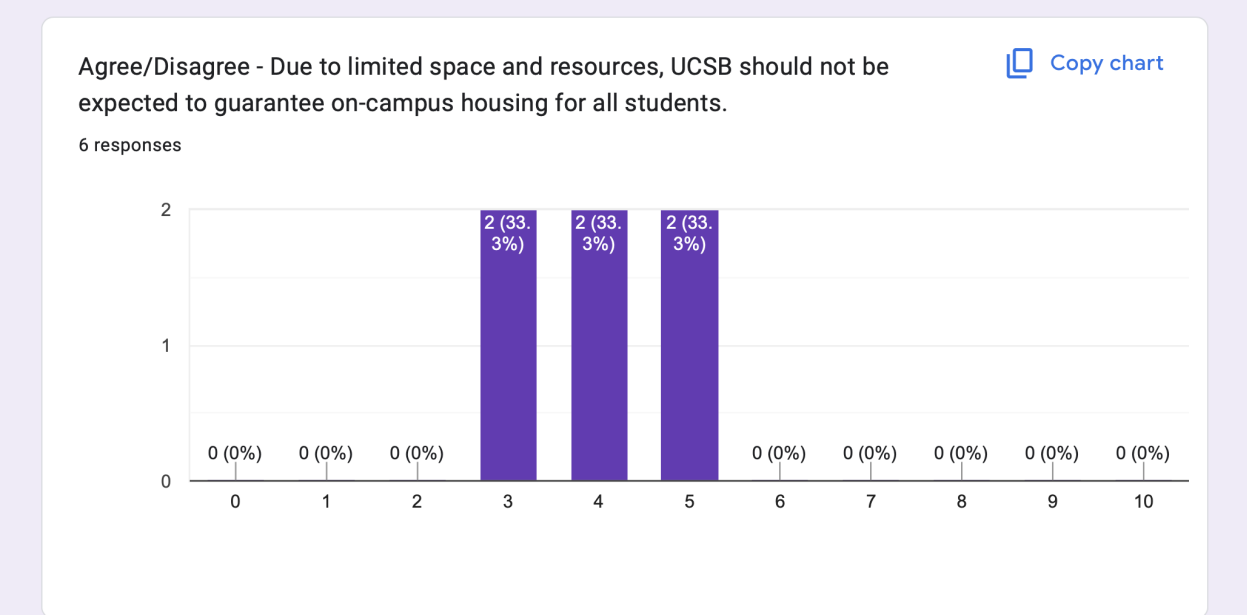

```
# responses

knitr::include_graphics("ss5.png")
```



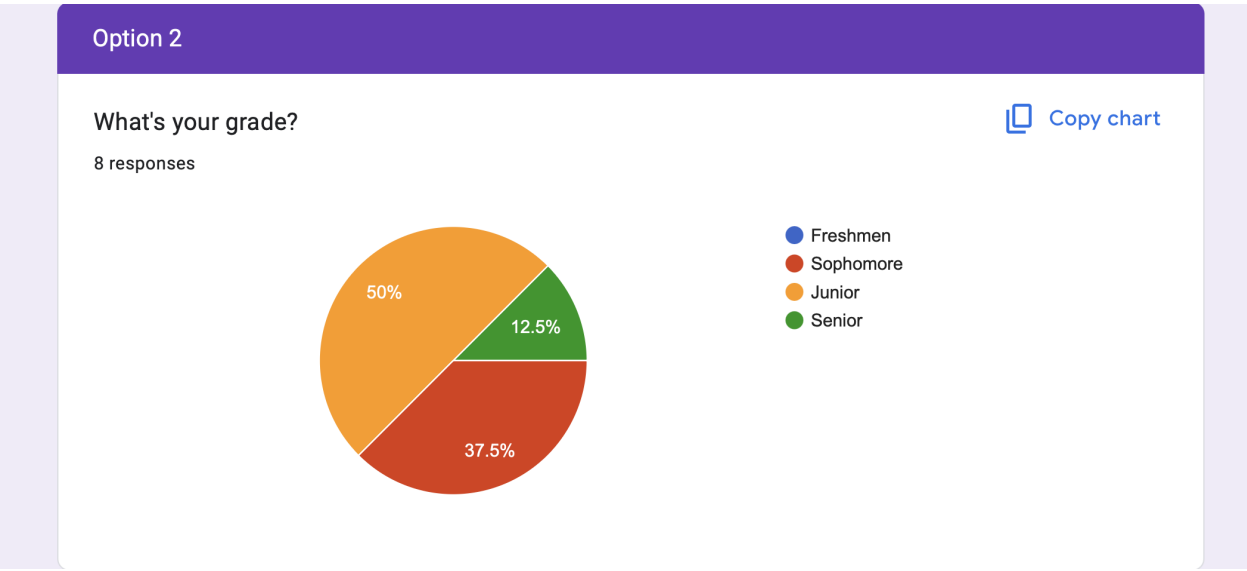
```
knitr::include_graphics("ss6.png")
```



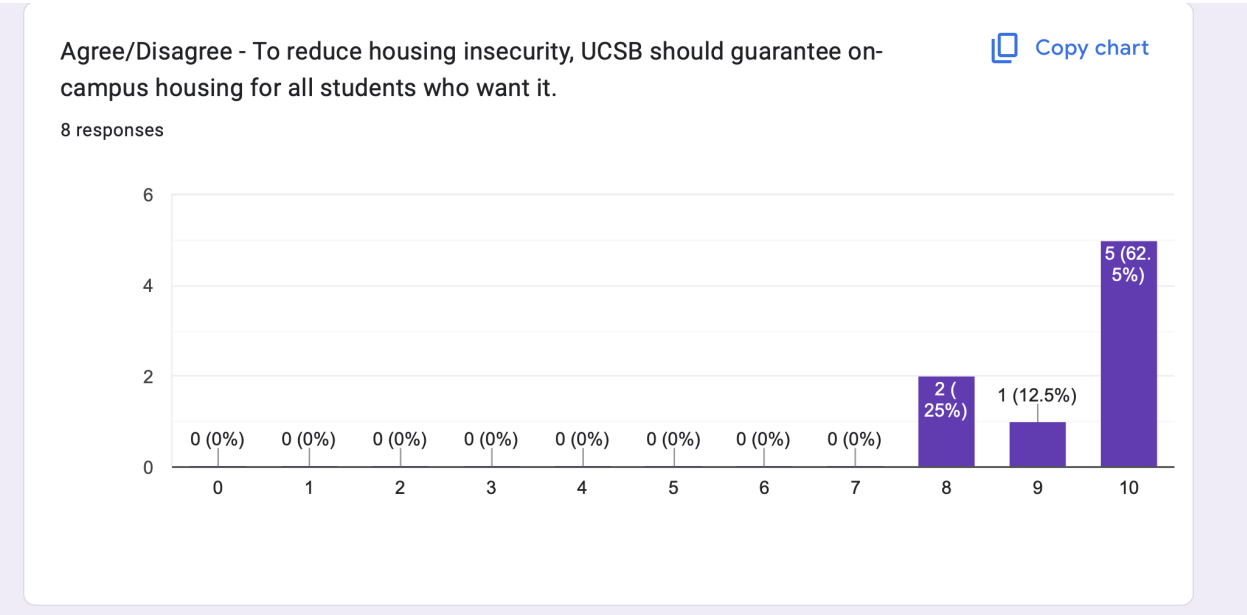

```
knitr::include_graphics("ss7.png") # response for Option 3
```



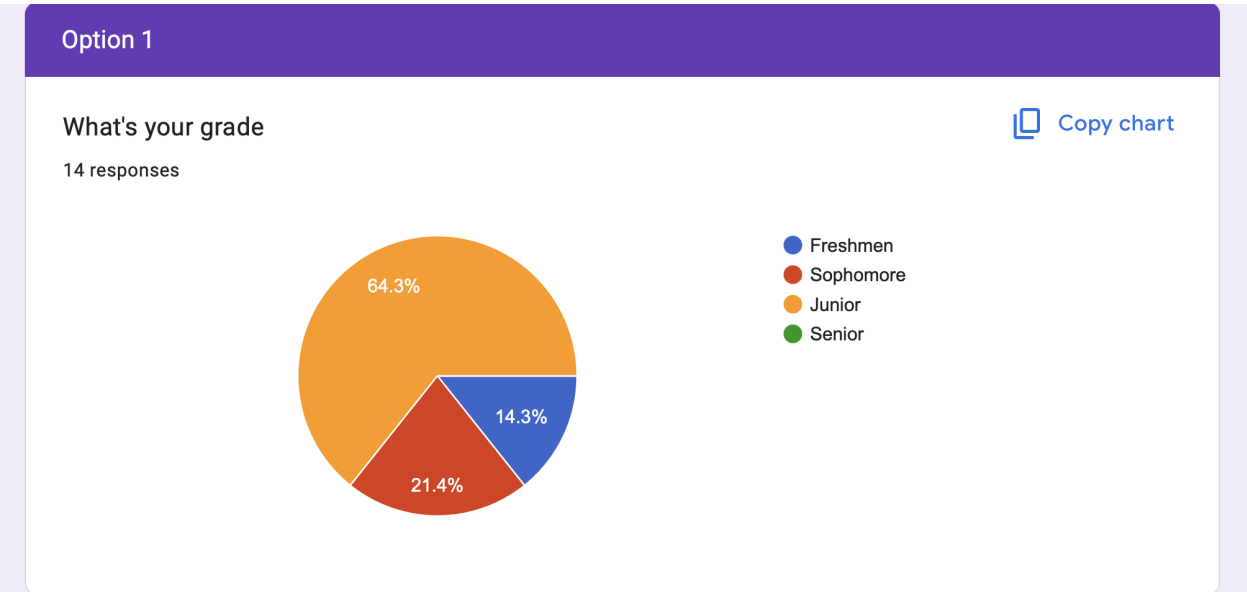

```
knitr::include_graphics("ss8.png")
```



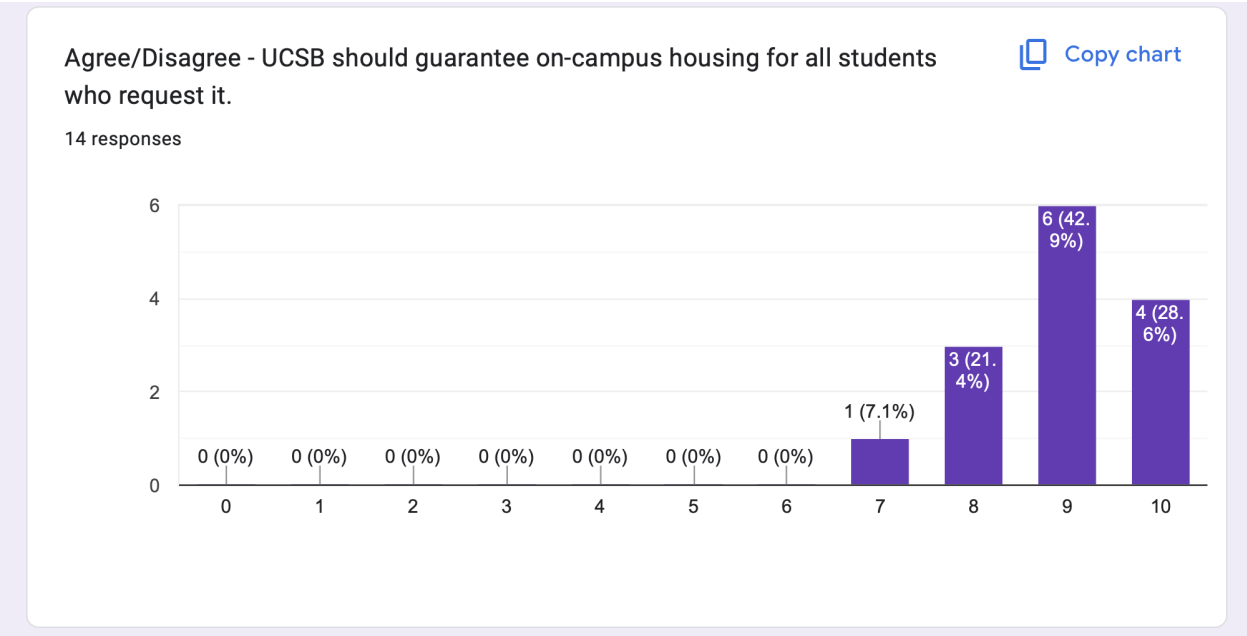

```
knitr::include_graphics("ss9.png") # reponse for Option 2
```




```
knitr::include_graphics("ss10.png")
```




```
knitr::include_graphics("ss11.png") # respose for Option 1
```



Statistical Methods used in Project

For the statistical analysis of the collected survey data, a Generalized Randomized Block Design (GRBD) approach was implemented. The primary factor under consideration was the wording of the housing policy statement, categorized into three levels: Neutral (Option 1), Student Benefit (Option 2), and Resource Constraints (Option 3). To account for potential variability in responses due to differences in student experience, class standing (Freshmen, Sophomore, Junior, and Senior) was used as a blocking variable. The outcome variable was the Agreement Score, measured on a scale from 0 to 10.

A one-way Analysis of Variance (ANOVA) was conducted to determine if the wording of the housing statement significantly affected student agreement levels. The ANOVA model included both the primary factor

(Option) and the blocking variable (Class Standing). The significance of each factor was assessed using p-values, with a threshold of 0.05 for statistical significance.

Technical Issues

I didn't run into amy technical issues on the course of this project, although the only issue I faced was the randomization of data in that it may not be truly randomized.

Results

Data Preparation

```
housing_data <- read.csv("pstat_survey_data.csv", stringsAsFactors = TRUE)

option1_data <- housing_data[housing_data$Versions == "Option 1", c("block1", "FOI1")]
colnames(option1_data) <- c("Block", "Rating")
option1_data$Option <- "Option 1"


option2_data <- housing_data[housing_data$Versions == "Option 2", c("block2", "FOI2")]
colnames(option2_data) <- c("Block", "Rating")
option2_data$Option <- "Option 2"


option3_data <- housing_data[housing_data$Versions == "Option 3", c("block3", "FOI3")]
colnames(option3_data) <- c("Block", "Rating")
option3_data$Option <- "Option 3"


stacked_data <- rbind(option1_data, option2_data, option3_data)


cat("Freshmen Rows:\n")
```

```
## Freshmen Rows:
```

```
print(stacked_data[stacked_data$Block == "Freshmen", ])
```

```
##       Block Rating   Option
## 24 Freshmen      7 Option 1
## 26 Freshmen      9 Option 1
## 21 Freshmen      5 Option 3
```

```
cat("\nSophomore Rows:\n")
```

```
##
## Sophomore Rows:
```

```r
print(stacked_data[stacked_data$Block == "Sophomore", ])
```

```
##        Block Rating   Option
## 20 Sophomore     10 Option 1
## 23 Sophomore      9 Option 1
## 25 Sophomore     10 Option 1
## 13 Sophomore      9 Option 2
## 22 Sophomore     10 Option 2
## 27 Sophomore     10 Option 2
## 15 Sophomore      4 Option 3
## 17 Sophomore      4 Option 3
```

```r
cat("\nJunior Rows:\n")
```

```
##
## Junior Rows:
```

```r
print(stacked_data[stacked_data$Block == "Junior", ])
```

```
##     Block Rating   Option
## 2  Junior      9 Option 1
## 4  Junior      9 Option 1
## 6  Junior      8 Option 1
## 7  Junior      9 Option 1
## 9  Junior      8 Option 1
## 10 Junior     10 Option 1
## 11 Junior      8 Option 1
## 18 Junior      9 Option 1
## 1  Junior     10 Option 2
## 5  Junior     10 Option 2
## 8  Junior     10 Option 2
## 12 Junior      8 Option 2
## 3  Junior      5 Option 3
## 16 Junior      3 Option 3
## 19 Junior      3 Option 3
```

```r
cat("\nSenior Rows:\n")
```

```
##
## Senior Rows:
```

```r
print(stacked_data[stacked_data$Block == "Senior", ])
```

```
##     Block Rating   Option
## 14 Senior      8 Option 2
```

Summary Statistics

```r
sum_stats <- stacked_data %>%
  group_by(Block) %>%
  summarize(mean=mean(Rating), var=var(Rating))

knitr::kable(sum_stats)
```

Block	mean	var
Freshmen	7.000000	4.000000
Junior	7.933333	5.638095
Sophomore	8.250000	7.071429
Senior	8.000000	NA

Sample Size Calculation

```r
# "Best case"

groupmeans <- c(7,7.93, 8.25, 8)
power.anova.test(groups=length(groupmeans),
between.var=var(groupmeans),
within.var=4,
power=0.8, sig.level=0.05, n=NULL)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##           groups = 4
##                n = 49.48136
##     between.var = 0.2997667
##      within.var = 4
##       sig.level = 0.05
##           power = 0.8
##
## NOTE: n is number in each group
```

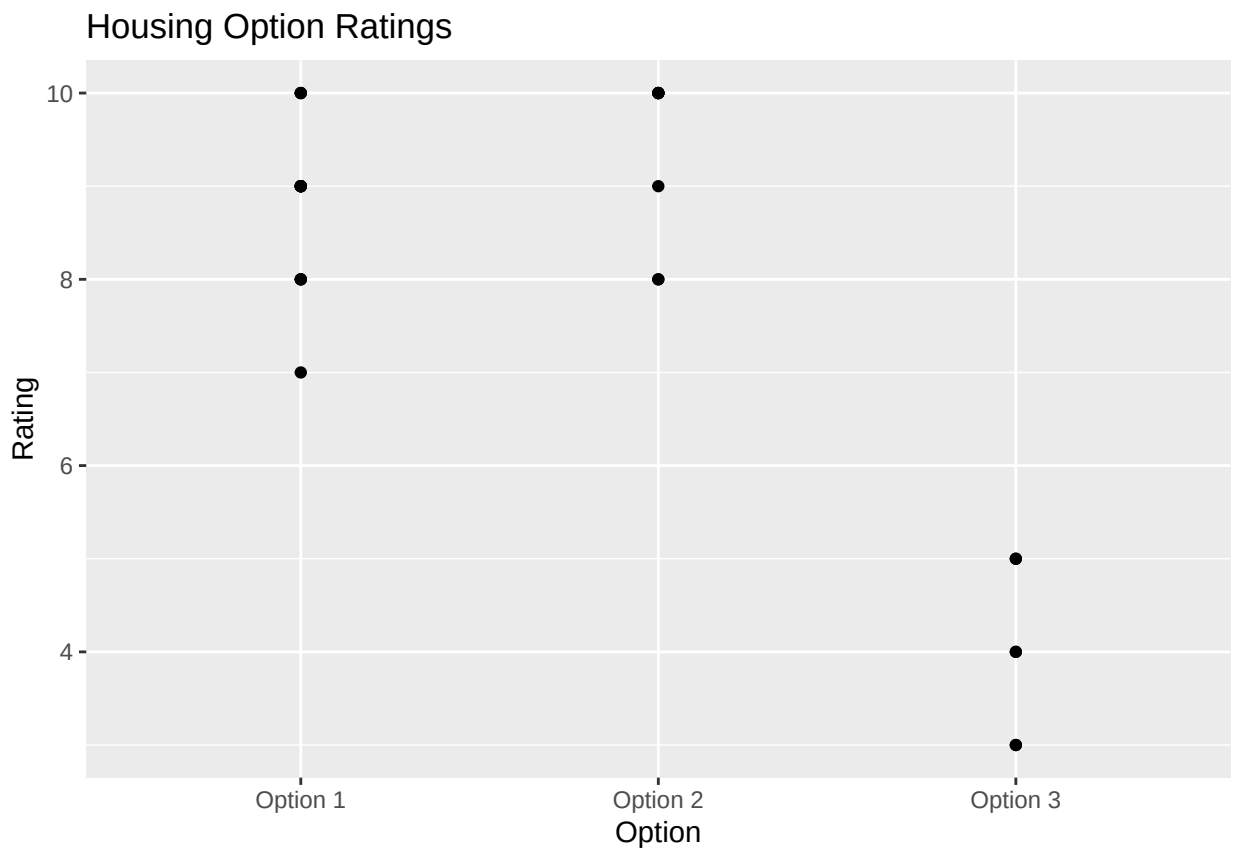
```r
# "Worst case"

groupmeans <- c(7,7.93, 8.25, 8)
power.anova.test(groups=length(groupmeans),
between.var=var(groupmeans),
within.var=8,
power=0.8, sig.level=0.05, n=NULL)
```

```
##
##      Balanced one-way analysis of variance power calculation
##
##           groups = 4
##                n = 97.96947
##     between.var = 0.2997667
```

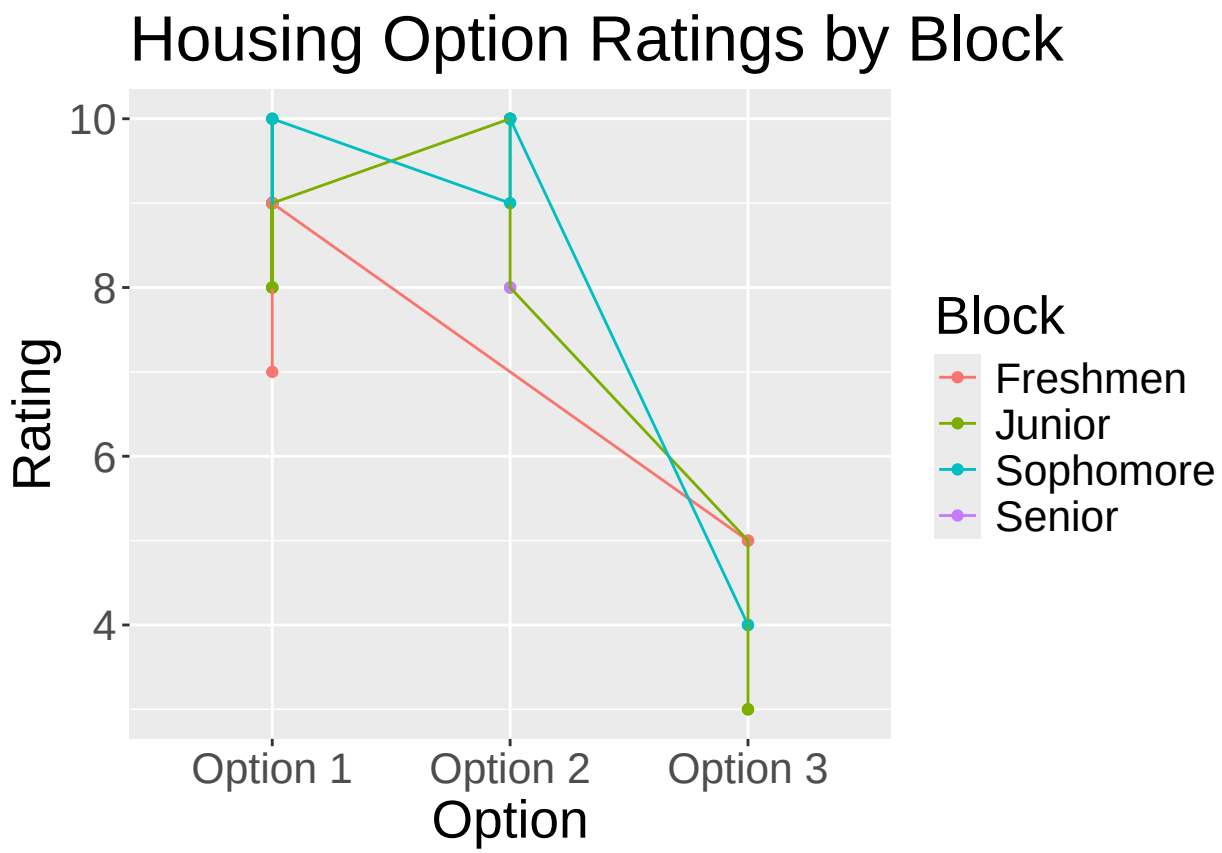
```
##      within.var = 8
##       sig.level = 0.05
##           power = 0.8
##
## NOTE: n is number in each group
```

Graphical Representation

```
ggplot(stacked_data, aes(y=Rating, x=Option)) +
geom_point() + ggtitle("Housing Option Ratings")
```




```
theme_update(text = element_text(size=20))

ggplot(data=stacked_data, aes(y=Rating, x=Option, group = Block, color = Block)) + geom_point() + geom_l
```



From the data, we observe that Option 1 and 2 have the highest agreement values whereas Option three has the lowest agreement values (more students disagree than agree)

Statistical Analysis & Model Fitting

```
fit <- aov(Rating ~ Option + Block, data=stacked_data)
sum_fit <- summary(fit)
knitr::kable(sum_fit[[1]])
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Option	2	120.284544	60.1422721	79.116542	0.0000000
Block	3	3.603671	1.2012238	1.580198	0.2239431
Residuals	21	15.963636	0.7601732	NA	NA

The ANOVA results indicate whether the different wording options (Neutral, Student Benefit, and Resource Constraints) significantly impact agreement scores, while accounting for variation in class standing. We observe a p-value of 0. This might indicate that at least one wording style leads to a statistically different response.

Additionally, the blocking factor `Block` helps determine whether agreement levels significantly vary across different class standings. If significant, this means that students at different levels (Freshmen, Sophomore, Junior, Senior) perceive housing policies differently. We observe that the p-value for Block is greater than the significance level (0.05), this indicates that there is no statistically significant effect of the wording class standing on agreement levels.

Discussion

The results of the ANOVA indicate that the wording of the housing policy statement had a significant effect on student agreement levels (p-value < 0.05). Specifically, the options emphasizing student benefit and neutrality garnered higher agreement scores compared to the statement emphasizing resource constraints. This suggests that students are more likely to support housing policies when framed positively rather than when limitations are highlighted.

On the other hand, the blocking variable (Class Standing) did not exhibit a statistically significant effect, implying that student perspectives on housing policies do not vary considerably across different academic levels. This finding suggests that concerns regarding housing availability are relatively uniform across UCSB students, regardless of their academic seniority.

A major limitation of this study is the lack of true randomization in data collection. Since participants self-selected their survey option, potential biases may exist in the distribution of responses. Additionally, the small sample size (n=27) limits the generalizability of the findings to the entire UCSB student population. Future research should aim for a larger, more representative sample and employ a fully randomized experimental design to improve validity.

Furthermore, external factors such as students' personal housing experiences or external economic conditions were not controlled for, which could influence agreement levels. Incorporating qualitative data or additional demographic variables could provide a more nuanced understanding of student opinions on housing policies.

Despite these limitations, the findings contribute to the broader discussion on university housing policies and the impact of framing effects on public opinion. Universities considering similar policy changes may benefit from carefully crafting their messaging to emphasize student benefits, thereby increasing policy acceptance among their student body.

References

Nishimoto, M., The Duty of Universities and the Right to the City: Balancing Campus Expansion with Community Impacts.