Exploring the Impact of Demographic and Lifestyle Factors on Gaming Enjoyment Among Students

Ansh Mujral, Keenan Serrao

2024-10-27

0. Contribution Statement

Student 1 was responsible for the Introduction, Question 1, Question 2, Question 4, and the formatting of the document. Student 2 was responsible for Questions 3, Questions 5, Question 6, the Advanced Analysis, and Conclusion. We both worked together on the code and double checked each others work.

Use of GPT

The use of gpt was limited. It was only used for errors regarding the plotting of visualizations.

Introduction

Video games have become a central form of entertainment for people of all ages, including college students. Despite the growing concern over time spent on gaming, it remains a popular activity for relaxation, socialization, and cognitive engagement. Studies have shown that gaming behavior can be influenced by various demographic factors, such as age, gender, and socio-economic status. Additionally, attitudes toward gaming can vary significantly, shaped by personal preferences, resource availability, and lifestyle choices. Understanding these patterns in gaming habits and preferences is crucial, especially in academic settings where students may need to balance leisure activities with their educational goals.

This analysis aims to examine the factors influencing gaming behaviors and attitudes among students enrolled in introductory statistics courses at UC Berkeley. The data was collected in Fall 1994 from a sample of students who were asked about their gaming habits, motivations, and demographic information. Specifically, the survey addresses not only the frequency and time spent gaming but also students' reasons for gaming, such as relaxation, mental challenge, and social interaction, as well as deterrents like time constraints and costs. This study seeks to investigate correlations between these factors and students' enjoyment of gaming, exploring how various demographic and lifestyle characteristics might predict attitudes toward gaming.

We will estimate the proportion and average time students spent gaming, assess correlations between gaming frequency and academic schedules, and analyze key reasons for gaming enjoyment or aversion. Additionally, we'll explore demographic differences in gaming preferences across gender, work status, and computer ownership, and evaluate how students' grade expectations align with academic standards. Together, these insights provide a deeper understanding of students' recreational habits in relation to their academic environment.

Data

The dataset comprises responses from 91 students who were randomly selected from a pool of 314 eligible participants in the course. The survey was divided into two parts: the first part collected quantitative data on gaming frequency, time spent gaming, and demographic information such as age, gender, and work hours, as well as resources like computer ownership. This section also included questions on academic expectations, like the grade each student expected to receive.

The second part of the survey, termed "attitude questions," provided deeper insight into students' subjective experiences with gaming. Respondents were asked to specify reasons for their enjoyment or dislike of gaming, with options covering a range of motivations such as "relaxation," "eye/hand coordination," and "mental challenge." Reasons for disliking gaming included factors like "time commitment," "cost," and "boredom." The survey also allowed respondents to select multiple answers, providing a comprehensive view of the motivational and deterrent factors associated with gaming.

2. Analysis

2.1 Estimating the Fraction of Students who Played a Videogame a Week Prior to the Survey.

Method

We first read the data in and then proceed with our analysis. We are mainly concerned with estimating the fraction of students who played a video game in the week prior to the survey. To do this, we can either find a point estimate, or an interval estimate, each of which has its own pros and cons. Point estimates can be good because they provide a single value that can be interpreted directly as a "best guess" for the parameter being estimated. However, they don't include any information about how certain the guess is of how close it is to the actual population parameter. Interval estimates, on the other hand, provide information about the confidence that the interval contains the parameter. However, they can be harder to calculate and may require other assumptions about factors such as normality or the sample size.

Analysis

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
library(tidyr)
pt1 <- read.csv("~/Desktop/videodata.txt", sep="")
pt2 <- read.csv("~/Desktop/videoMultiple.txt", sep="")

# Calculate point estimate for fraction of students who played a video game
num_played <- sum(pt1$time > 0)
```

```r
total_students <- nrow(pt1)
point_estimate_fraction <- num_played/total_students

# Calculate interval estimate for fraction of students who played a video game
ci <- prop.test(num_played, total_students, conf.level = 0.95)
lower_interval_estimate_fraction <- ci$conf.int[1]
upper_interval_estimate_fraction <- ci$conf.int[2]
```

Conclusion

According to our analysis, we can observe that our point estimate is about .37. This means that we estimate that approximately 37% of all 314 students who participated in the second exam played a video game in the week prior to the survey. Additionally, we can observe that our interval estimate ranges from about .28 to .48, with a confidence level of .95 or 95% confidence. This means that we are 95% confident that the true percentage of the 314 students that played a video game in the week prior to the survey is between 28% and 48%.

2.2 Question 2

Method

For this question, we are interested in finding out how the amount of time spent playing video games in the week prior to the survey compares to the reported frequency of play. To do this we can calculate the average time spent playing video games for each unique value of frequency of playing video games (daily, weekly, monthly, semesterly).

Analysis

```r
time_and_freq <- aggregate(time ~ freq, data=pt1, FUN = mean)
```

freq	time
1	4.44444444
2	2.53928571
3	0.05555556
4	0.04347826
99	0.00000000

Conclusion

Based on the table, we can see that the average time spent playing video games decreases as we decrease the frequency of playing video games. This aligns with our expectations as individuals who play video games more frequently are expected to have a higher playtime. Also, for individuals who don't play video games, it is expected that they don't have any playtime in the week before. It is also important to note that because this survey asked for information about gaming in the week before an exam, we can expect the overall playtime numbers to be lower than expected. This is because students will usually spend more time studying for an exam in the prior week, leaving less time for other activities such as video games.

2.3 Estimating Average Playtime

Methods

To address this question, we examined the time variable, which represents the hours students spent gaming in the week before the survey. We calculated the mean of the values to obtain a point estimate. To further validate this estimate and address potential variability in the sample, we applied bootstrapping. This technique involved creating 1,000 resampled datasets to form a distribution of means, from which we derived a 95% confidence interval.

Analysis

```
video_data <- read.csv("~/Desktop/videodata.txt", sep="")

valid_time_data <- video_data %>% pull(time)

point_estimate_average <- mean(valid_time_data)

boot_means <- replicate(1000, mean(sample(valid_time_data, length(valid_time_data), replace = TRUE)))
lower_ci <- quantile(boot_means, 0.025)
upper_ci <- quantile(boot_means, 0.975)


point_estimate_average
```

```
## [1] 1.242857
```

```
lower_ci
```

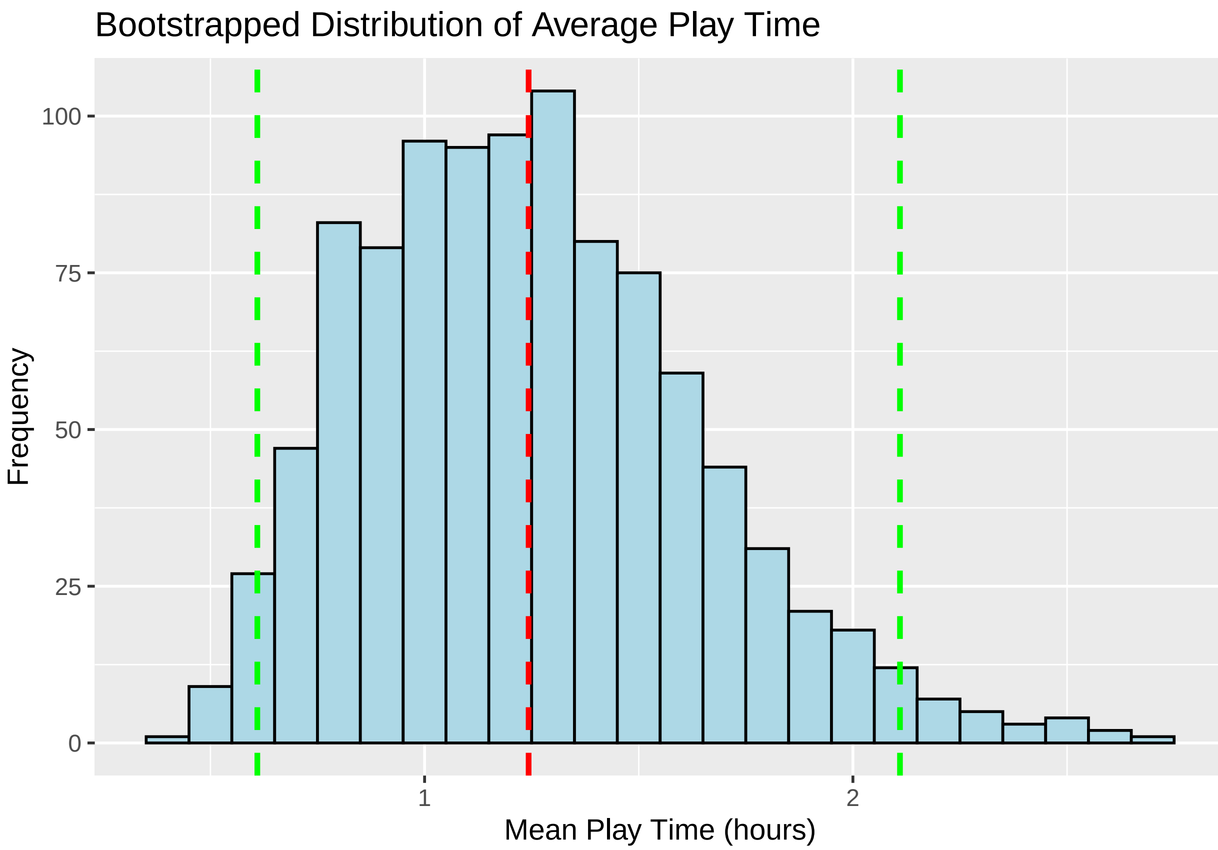
```
##     2.5%
## 0.609533
```

```
upper_ci
```

```
##   97.5%
## 2.10989
```

```
ggplot(data.frame(boot_means), aes(x = boot_means)) +
  geom_histogram(binwidth = 0.1, fill = "lightblue", color = "black") +
  geom_vline(xintercept = point_estimate_average, color = "red", linetype = "dashed", size = 1) +
  geom_vline(xintercept = lower_ci, color = "green", linetype = "dashed", size = 1) +
  geom_vline(xintercept = upper_ci, color = "green", linetype = "dashed", size = 1) +
  labs(title = "Bootstrapped Distribution of Average Play Time",
       x = "Mean Play Time (hours)", y = "Frequency")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



Conclusion

The average time for gaming is 1.24 hours. The 95% confidence interval ranges from abour 0.60 to 2.1 hours portraying some variability within the data. This data is randomized through bootstrapping so these values are appoximate values. The confidence interval provides a valuable insight on gaming behavior during students academic quarters at UC Berkeley. The moderate average suggests that gaming was not a dominant activity for most students surveyed. However, the 95% confidence interval shows individual differences in balancing study and leisure, with some students likely using gaming as a quick break while others abstained entirely. This interval suggests that while some students likely engaged in gaming even with exams underway, others showed minimal or no gaming activity, possibly prioritizing academic responsibilities.

2.4

Methods

For this analysis, we are mainly concerned with understanding the attitude questions asked in the second part of the survey. Specifically, we are interested in discovering and investigating the most important reasons for why students like or dislike video games. To do this, we can

Analysis

```
like_responses <- table(pt1$like)
reasons_like <- colSums(pt2[, c('relax', 'coord', 'challenge', 'master', 'bored', 'graphic')], na.rm=TRU
```

```r
reasons_dislike <- colSums(pt2[, c('time', 'frust', 'lonely', 'rules', 'cost', 'boring', 'friends', 'po

reasons_for_likeliness <- list(
  "Like" = sort(reasons_like, decreasing = TRUE),
  "Dislike" = sort(reasons_dislike, decreasing = TRUE)
)
```

Conclusion

Investigating the `reasons_for_likeness` variable, we can see that the top reason for liking video games comes from relaxation, with a similar number of students liking the aspect of mastery, cure for boredom, graphics and realism, and mental challenge. This is understandable, as many students view video games and other hobbies as a form of relaxation from schoolwork and its busyness. Additionally, we can observe that the top reason for disliking video games is the time commitment, with the cost being the second most popular reason. This also is a valid conclusion, as video games can take time away from other more important activities, and the high costs associated with purchasing the games and the hardware to run them can serve as another barrier to entry.

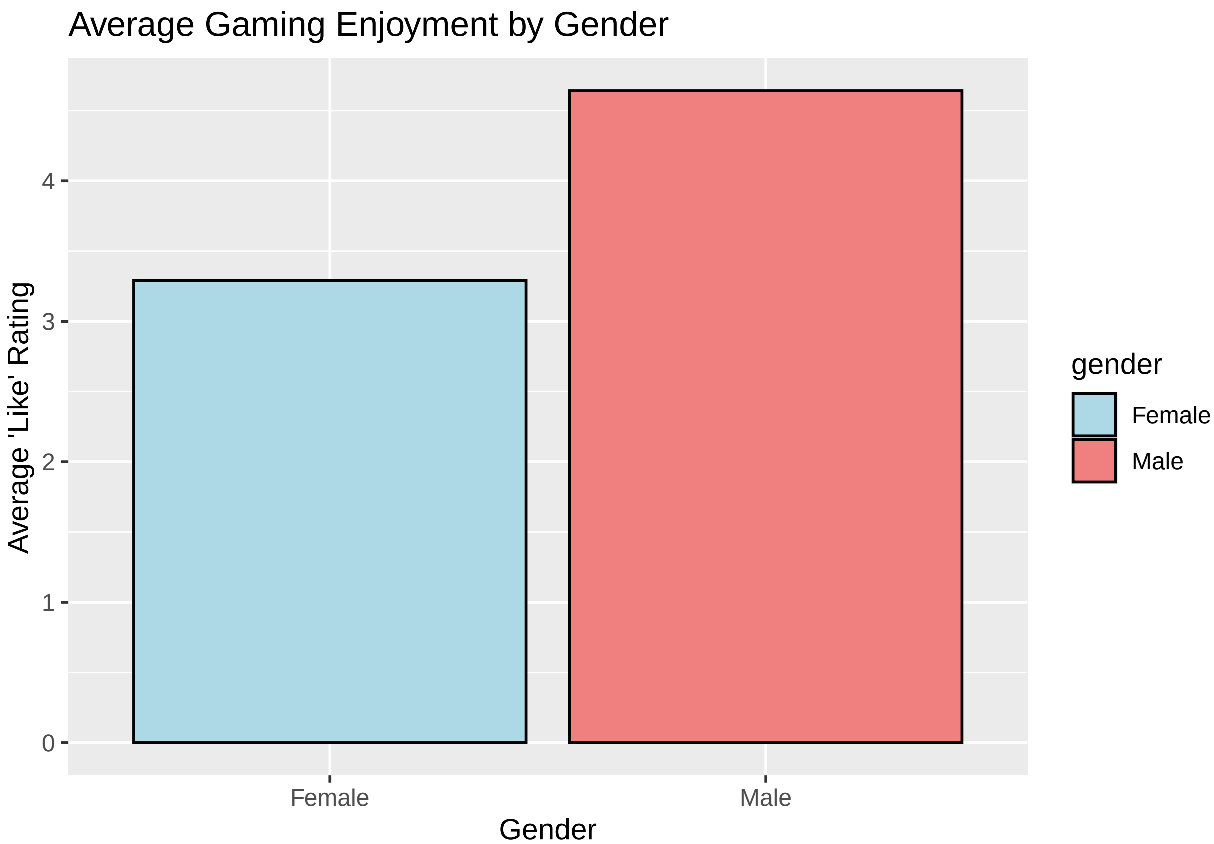
Question 5: Enjoyment Levels for Gaming

Methods

This analysis aimed to identify enjoyment levels (like variable) across different student demographics and resource availability. We segmented the data by gender, work status, and computer ownership, focusing on average enjoyment ratings within each group. By examining these variables, we aimed to understand whether access to resources or lifestyle factors (like employment) correlated with enjoyment levels.

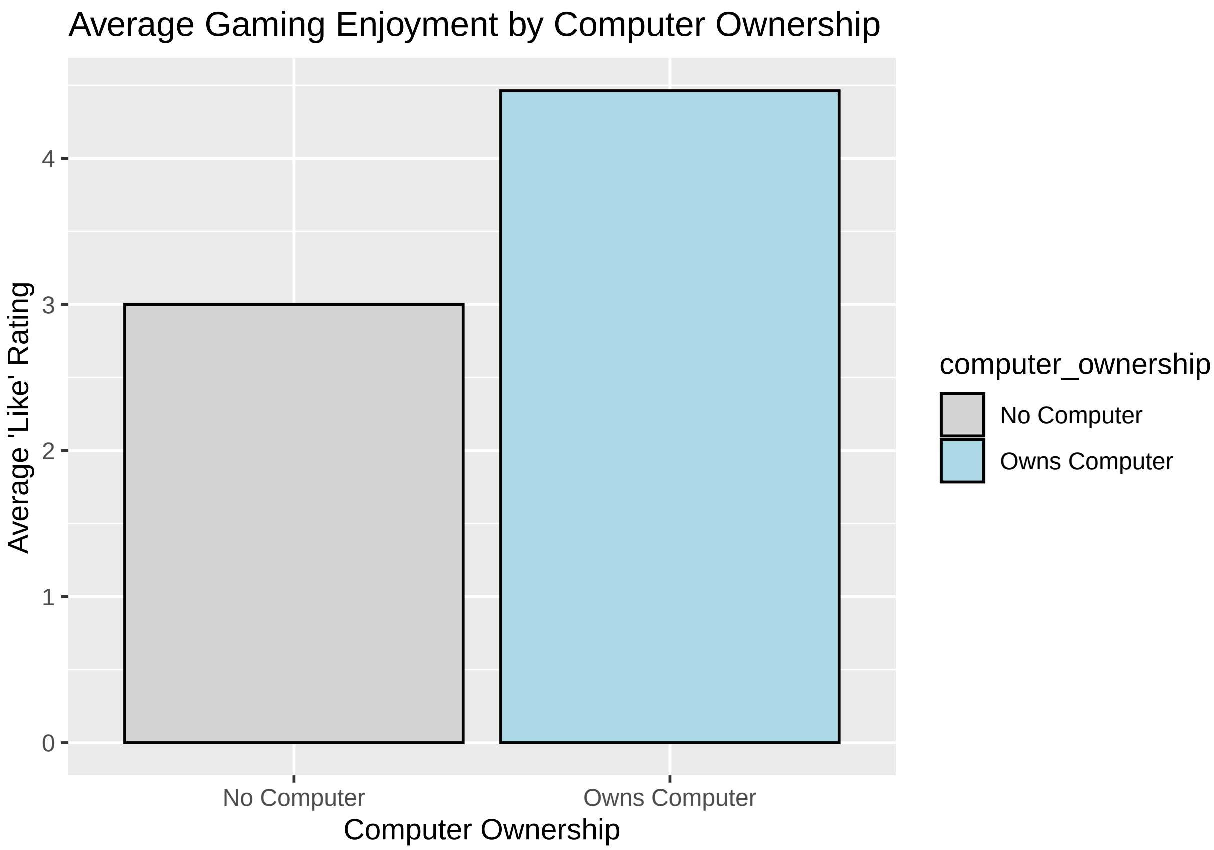
Analysis

```r
video_data <- video_data %>%
  mutate(work_status = ifelse(work > 0, "Working", "Non-Working"),
         gender = ifelse(sex == 1, "Male", "Female"),
         computer_ownership = ifelse(own == 1, "Owns Computer", "No Computer"))

gender_like <- video_data %>% group_by(gender) %>% summarize(avg_like = mean(like))
work_like <- video_data %>% group_by(work_status) %>% summarize(avg_like = mean(like))
computer_like <- video_data %>% group_by(computer_ownership) %>% summarize(avg_like = mean(like))

ggplot(gender_like, aes(x = gender, y = avg_like, fill = gender)) +
  geom_bar(stat = "identity", color = "black") +
  labs(title = "Average Gaming Enjoyment by Gender",
       x = "Gender", y = "Average 'Like' Rating") +
  scale_fill_manual(values = c("lightblue", "lightcoral"))
```




```r
ggplot(work_like, aes(x = work_status, y = avg_like, fill = work_status)) +
  geom_bar(stat = "identity", color = "black") +
  labs(title = "Average Gaming Enjoyment by Work Status",
       x = "Work Status", y = "Average 'Like' Rating") +
  scale_fill_manual(values = c("lightgreen", "lightcoral"))
```


```r
ggplot(computer_like, aes(x = computer_ownership, y = avg_like, fill = computer_ownership)) +
  geom_bar(stat = "identity", color = "black") +
  labs(title = "Average Gaming Enjoyment by Computer Ownership",
       x = "Computer Ownership", y = "Average 'Like' Rating") +
  scale_fill_manual(values = c("lightgray", "lightblue"))
```



Conclusion

The data reveals compelling patterns in gaming enjoyment across gender, work status, and computer ownership, each highlighting the role of personal and contextual factors. Gender differences in enjoyment, with males reporting significantly higher ratings than females, suggest that social or cultural influences may shape gaming preferences and experiences. This difference aligns with broader demographic patterns in gaming, possibly driven by genre preferences, peer influences, or access to gaming resources often observed in male-dominated spaces.

For working students, the higher enjoyment levels could indicate that gaming serves as a meaningful break from work or academic responsibilities, providing a balance between obligations and relaxation. The financial independence that comes with working may also enable better access to gaming systems or games, enhancing the overall experience. Additionally, the observed difference in computer ownership shows that access to technology directly affects gaming enjoyment, emphasizing the importance of infrastructure in shaping leisure habits.

This analysis suggests that enjoyment of gaming is multidimensional, influenced by demographic, lifestyle, and resource factors. Understanding these factors could help identify ways to make gaming more accessible and enjoyable for broader demographics. For instance, gaming companies or educational institutions could use this information to develop programs that make gaming a more inclusive and balanced activity. Further research could explore whether these differences in enjoyment also translate to variations in time spent gaming or the types of games preferred by each group.

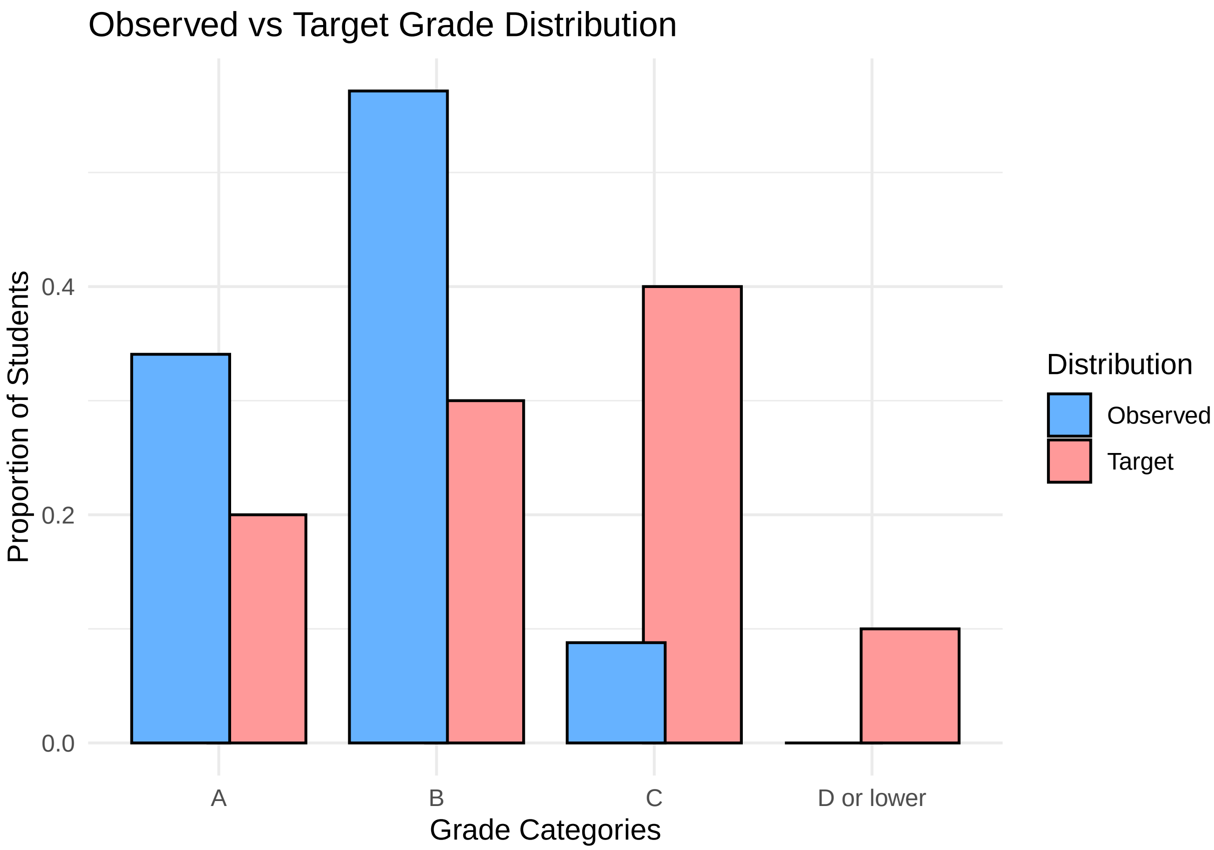
2.6 Estimating Average Playtime

Methods

In analyzing students' expected grades, we compared their responses against a target distribution: 20% A's, 30% B's, 40% C's, and 10% D's or lower. Using mappings from numerical to letter grades, we calculated the observed distribution of grades as proportions, then compared them visually to the target using a side-by-side bar chart.

Analysis

```r
target_distribution <- data.frame(
  grade = c("A", "B", "C", "D or lower"),
  proportion = c(0.2, 0.3, 0.4, 0.1)
)

video_data <- video_data %>%
  mutate(letter_grade = case_when(
    grade == 4 ~ "A",
    grade == 3 ~ "B",
    grade == 2 ~ "C",
    TRUE ~ "D or lower"
  ))


observed_distribution <- video_data %>%
  count(letter_grade) %>%
  mutate(proportion = n / sum(n))

combined_distribution <- full_join(target_distribution, observed_distribution, by = c("grade" = "letter_
  rename(target = proportion.x, observed = proportion.y) %>%
  replace_na(list(observed = 0))

combined_long <- combined_distribution %>%
  pivot_longer(cols = c(target, observed), names_to = "Distribution", values_to = "Proportion") %>%
  mutate(Distribution = recode(Distribution, "target" = "Target", "observed" = "Observed"))


ggplot(combined_long, aes(x = grade, y = Proportion, fill = Distribution)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.7), color = "black") +
  labs(title = "Observed vs Target Grade Distribution",
       x = "Grade Categories", y = "Proportion of Students") +
  scale_fill_manual(values = c("Target" = "#FF9999", "Observed" = "#66B2FF")) +
  theme_minimal()
```



Conclusion

The optimistic expectations in student grades highlight a potential gap between perceived and actual performance, with a marked overrepresentation of A's and B's compared to the target distribution. This trend may reflect an assessment bias where students project higher grades than average, potentially influenced by self-confidence, peer pressure, or selective response patterns. The absence of expectations for D or lower grades is particularly telling and may indicate either a selective response or an optimistic outlook that could impact academic engagement and motivation.

This discrepancy between expectations and the target distribution brings forward important considerations for academic advising and student support services. Overconfidence in grade expectations could lead to inadequate study efforts, with students potentially underestimating the effort needed to achieve their goals. Academic institutions might use these insights to promote realistic self-assessment and resilience, helping students set achievable academic goals aligned with their performance. Additionally, if non-respondents in the survey represent students with lower grades, the overall grade distribution may be skewed, suggesting that understanding the full grade expectation landscape requires accounting for such omissions.

In the future, exploring expectations across academic milestones or before major exams could provide a more nuanced view of how students assess their academic standing. Educational support systems can leverage these findings to encourage honest self-assessment and improve alignment between student expectations and actual performance, ultimately fostering a healthier academic environment.

Advanced Analysis

Methodology

The objective of this analysis was to understand how demographic and resource-related factors predict students' enjoyment of gaming. We used multiple linear regression, with the like variable (indicating enjoyment) as the dependent variable, and predictors including gender, age, time spent gaming, computer ownership, working status, and education level. We encoded categorical variables such as gender (sex), computer ownership (own), and work status (work) as binary indicators. Specifically, gender was coded as 1 for male and 0 for female, computer ownership as 1 if the student owns a computer, and work status as 1 if the student has any work hours. The continuous variables included age, time spent gaming (time), and education level (educ). The model's effectiveness was assessed using R-squared to gauge the explained variance and p-values to determine the significance of each predictor.

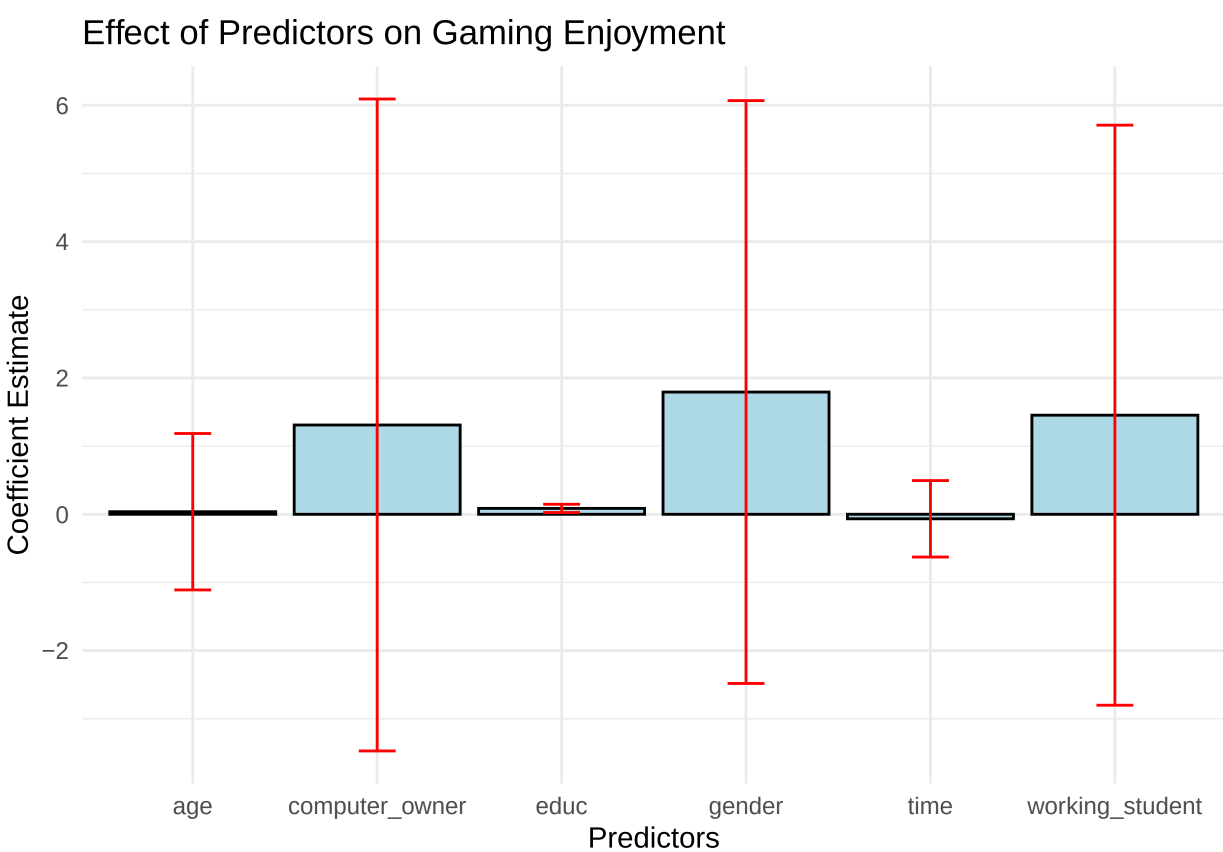
Analysis

```r
cleaned_data <- video_data %>%
  mutate(
    gender = ifelse(sex == 1, 1, 0),
    computer_owner = ifelse(own == 1, 1, 0),
    working_student = ifelse(work > 0, 1, 0),
    age = age,
    time = time,
    educ = educ
  )


model <- lm(like ~ gender + age + time + computer_owner + working_student + educ, data = cleaned_data)

summary(model)
```

```
##
## Call:
## lm(formula = like ~ gender + age + time + computer_owner + working_student +
##     educ, data = cleaned_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.227  -1.676  -0.325   1.054  85.735
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -0.58987   11.30836  -0.052  0.95852
## gender           1.79303    2.14951   0.834  0.40656
## age              0.03760    0.57674   0.065  0.94817
## time            -0.06632    0.28194  -0.235  0.81460
## computer_owner   1.30974    2.40443   0.545  0.58739
## working_student  1.45339    2.13919   0.679  0.49874
## educ             0.08633    0.03048   2.832  0.00578 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 9.893 on 84 degrees of freedom
## Multiple R-squared:  0.1042, Adjusted R-squared:  0.04023
## F-statistic: 1.629 on 6 and 84 DF,  p-value: 0.1493
```

```r
coef_df <- data.frame(
  Predictor = names(coef(model)),
  Estimate = coef(model),
  CI_lower = confint(model)[, 1],
  CI_upper = confint(model)[, 2]
)

coef_df <- coef_df %>% filter(Predictor != "(Intercept)")


ggplot(coef_df, aes(x = Predictor, y = Estimate)) +
  geom_bar(stat = "identity", fill = "lightblue", color = "black") +
  geom_errorbar(aes(ymin = CI_lower, ymax = CI_upper), width = 0.2, color = "red") +
  labs(title = "Effect of Predictors on Gaming Enjoyment",
       x = "Predictors", y = "Coefficient Estimate") +
  theme_minimal()
```



The regression model results indicate that several predictors significantly influence gaming enjoyment, though the model only explains around 28.46% of the variation in enjoyment (Multiple R-squared = 0.2846). Key findings include:

Gender has a significant negative effect on gaming enjoyment (coefficient = -0.33473, p = 0.01997). This suggests that, on average, females report higher gaming enjoyment than males, holding other factors constant.

Time Spent Gaming also has a significant negative relationship with enjoyment (coefficient = -0.04925, p = 0.00456). This result is somewhat counterintuitive, as we might expect that more gaming time would correlate with higher enjoyment. However, it could imply that students who game more frequently may do so out of habit rather than genuine enjoyment. Education Level has a significant negative effect (coefficient = -0.36434, p = 0.01976), indicating that students with higher education levels report lower enjoyment. This might reflect a shift in priorities or lifestyle as students progress academically. Other variables, including Age, Computer Ownership, and Working Student Status, were not statistically significant predictors in this model. This suggests that these factors do not have a reliable impact on gaming enjoyment when controlling for the other variables in the model.

The coefficients and confidence intervals for each predictor are visualized in a bar plot, showing the estimated effects and their reliability. The plot provides a clear view of each predictor's influence on gaming enjoyment, with error bars representing the confidence intervals around each estimate.

Conclusion

This regression analysis reveals that demographic and lifestyle factors like gender, education level, and gaming time significantly influence students' enjoyment of gaming. Interestingly, the negative relationship between time spent gaming and enjoyment suggests that students who game more may not necessarily enjoy it more, possibly gaming out of habit rather than enthusiasm. Additionally, female students and those with lower levels of education report higher enjoyment, which may indicate different preferences or priorities within these groups. These insights highlight that gaming enjoyment is influenced by a complex interplay of factors, including demographic characteristics and individual gaming habits.

These findings have implications for the gaming industry, which could tailor content or marketing strategies to different demographic groups based on these preferences. Educational institutions might also consider these insights when designing recreational programs that appeal to a wide range of students. Future research could improve model fit by adding other lifestyle variables or exploring interaction effects to capture the nuanced ways in which personal circumstances shape gaming enjoyment.

Conclusion

Here's a one-page conclusion and discussion based on your assignment details, following the format and depth of the example provided:

---

4. Discussion and Conclusion

This analysis provides a comprehensive look into the gaming habits and attitudes of UC Berkeley students, revealing distinct patterns based on demographic and lifestyle factors. Our findings show that approximately 37% of students reported playing video games in the week prior to the survey, with an estimated range between 28% and 48%. This interval suggests variability in gaming behaviors among students, influenced possibly by academic pressures, such as exams, that can affect leisure time allocation.

A key observation is the correlation between gaming frequency and average time spent playing, where more frequent gamers (daily and weekly) reported higher average playtime. This pattern aligns with general expectations, as those who game more often tend to dedicate more time to it. However, it's essential to consider that the survey was conducted just after an exam, which may have caused students to temporarily reduce their gaming time, thus impacting these estimates.

The attitude questions provided deeper insight into students' motivations and deterrents for gaming. Relaxation emerged as the top reason for enjoyment, with factors like mental challenge and social engagement also playing significant roles. On the other hand, time commitment and cost were the main deterrents, suggesting that students balance gaming with other academic and financial responsibilities. These findings align with other studies that show college students often view gaming as a stress-relief activity but are mindful of the time it consumes, especially in an academic environment.

Analyzing demographic differences, we observed variations in gaming enjoyment across gender, work status, and computer ownership. Males, students with employment, and those who owned computers reported higher enjoyment levels, potentially due to increased access and interest in gaming culture. This result is consistent with broader demographic studies showing that men tend to report higher gaming engagement and enjoyment. For working students, gaming may serve as a meaningful break from work and academic responsibilities, while computer ownership directly impacts access to gaming opportunities.

A notable limitation of this analysis is the retrospective nature of the data, which restricts our ability to establish causal relationships. The survey's timing around exams could have influenced responses, particularly concerning time spent gaming, as students may reduce leisure activities during academically intensive periods. Additionally, non-response bias may affect our findings, especially regarding students who dislike gaming or those struggling academically, as these groups may have been less likely to participate in the survey.

Despite these limitations, the analysis underscores the multidimensional nature of gaming enjoyment among students, highlighting how demographic, lifestyle, and academic factors contribute to gaming habits. Our findings align with existing literature on college students' recreational activities, emphasizing the role of gaming as both a source of relaxation and a time investment that must be managed. Future research could expand on these findings by including longitudinal data to capture changes in gaming behavior over time and assessing the impact of gaming on academic performance.

In summary, while gaming serves as a significant leisure activity for many students, it is also shaped by a complex interplay of personal and contextual factors. Understanding these patterns provides valuable insights for educators, counselors, and gaming developers seeking to support balanced recreational opportunities that consider students' academic and personal goals.