MATH 189 Case Study 2

Berkeley Video Game Survey Analysis

Yuxin Zou, A13996888
Yifei Li, A92082060
Bo Hu, A13805750
Bo Zhang, A13616488
Xuanyu Wu, A13569778

Literature Review

Nowadays, video games have become the dominant culture in the world for teenagers. A research noticed that 97% of the children in the United States play video games for at least 1 hour per day (Granic). While many psychologists have found out the negative effect of playing video games on children, we should realize that video games have changed overtime since they were first being developed. More and more researchers are trying to discover the benefits from playing video games.

The benefits of playing video games are not limited to cognitive development, emotional health, and potential social learning experiences. Cognitively, playing video games would not only improve eye-hand coordination or hasten reaction time, but also has a widespread benefits on vision and cognition like peripheral localization or the capacity of visual attention (Green). Emotionally, during physical or psychological therapies, people use video games to distract patients from pain in order to smooth the therapy process (Primack). More interestingly, Video games could also be brought to the stage of educational purpose, since they are really motivating for students. Other interesting traits of educational video games is that they provide immersive in-game worlds that can be explored freely by the students, promoting self-directed learning, their short feedback cycles with perception of progress, or their relation to constructivist theories and support of scaffolded learning (Torrente). Now we would like to see the example from UC Berkeley on the research of video games connecting to a new designed computer lab.

Introduction and Data Description

Every year, three to four thousand students enroll in statistics courses at UC Berkeley and half of them take introductory statistics courses to satisfy their quantitative reasoning requirement. In order to facilitate those students' studies in these courses, a committee of faculty and students has designed a series of computer labs, aiming at extending the traditional syllabus by providing an interactive learning environment. As a good lab has parallel characteristics with an attractive video game, a survey of undergraduate students who were enrolled in a lower-division statistics course was conducted, studying which aspects of video games those students find most and least fun. The discovery from this survey will be used to help design a good computer lab of this statistics course. The survey was designed and conducted by students enrolled in another advanced statistics course and most of those students surveyed are business major students, meaning that we cannot extend the result of this paper to students of other majors confidently.

In this study, not all students enrolled in the class were surveyed. Only 95 out of 314 students were sampled to answer the questionnaire and only 91 finished. Since this study is based upon data collected from a survey, survey methodology and sampling strategy play an important role in determining the degree of reliability of the discovery from this study. A survey follows scientific methodologies to collect data from individuals, usually samples from a large population, for the purpose of describing, exploring and explaining the general characteristics and patterns of the whole population. A good survey research should be quantitative, self-monitoring, replicable, impartial, theory-based, and representative of the entire population (in this survey the entire population is 314 students enrolled in this lower division statistics course). And the sampling technique is only used when appropriate. If all members of the population are identical, sampling would be redundant. In general, when sampling, every individual must have equal chance of being selected, in order to represent the interests of the total population.

The following table is the data dictionary for the encoded data set:

Column Name	Data Description and Dictionary	Type
time	Number of hours played in the week prior to the survey	Numerical
like	The degree to which the student likes to play video games. 1=never played; 2=very much; 3=somewhat; 4=not really; 5=not at all.	A combination of ordinal and nominal
where	Where the student primarily plays video games. 1=arcade; 2=home system; 3=home computer; 4=arcade and either home computer or system; 5=home computer and system; 6=all three.	Nominal
freq	How often does the student play video games. 1=daily; 2=weekly; 3=monthly; 4=semesterly.	Ordinal
busy	Whether the student tends to play video games when they are busy. 1=yes; 0=no.	Categorical
educ	Whether the student thinks playing video games can be	Categorical

	educational 1=yes; 0=no.	
sex	The biological sex of the student 1=male; 0=female.	Categorical
age	Age of the surveyed student.	Numerical
home	Whether the student has a computer at home 1=yes; 0=no.	Categorical
math	Whether the student likes math 1=yes; 0=no.	Categorical
work	Number of hours the student has worked the week prior to the survey	Numerical
own	Whether the student owns a personal computer 1=yes; 0=no.	Categorical
cdrom	1=yes; 0=no.	Categorical
email	Whether the student has an email account 1=yes; 0=no.	Categorical
grade	The grade type the student expects to receive for this course 4=A; 3=B; 2=C; 1=D; 0=F.	Nominal

Investigation:

Scenario 1:

To provide an estimate of the proportion of students who played video games prior to the survey, we start with a point estimate. Since sample proportion is an unbiased estimate of the population proportion, a reasonable point estimate is the fraction of students who played video games in the sample, which is 0.374.

Then, instead of a single estimate, we want to find an interval that the true proportion is likely to fall in. The first procedure is using Central Limit Theorem because the sample size is 91, which is greater than 30. By Central Limit Theorem, the sample proportion follows normal distribution with mean p and the standard deviation $\sqrt{\frac{p(1-p)}{n}}$, where p is the true population proportion and n is the sample size. Since we do not know the population proportion, we use our point estimate $\widehat{p}$ instead. Therefore, the 95% confidence interval is

$$(\widehat{p} - 1.96 * \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}},\ \widehat{p} + 1.96 * \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}) = (0.274,\ 0.473)\ .$$

However, the size of the population in this case is 314, which is not large enough for the population to be considered infinite. Therefore, it makes sense to include finite sample correction factor. The adjusted confidence interval is now $(\widehat{p} - 1.96 * \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n-1}}\sqrt{\frac{N-n}{N}},\ \widehat{p} + 1.96 * \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n-1}}\sqrt{\frac{N-n}{N}}) = (0.288,\ 0.459)$, which, as expected, has a smaller range due to the corrector.

Lastly, we can also get a confidence interval by the empirical distribution of sample proportions. Generating 400 bootstrap samples, we calculate the mean of each sample and take 2.5% and 97.5% percentiles of bootstrap means to be the endpoints of the interval. Here, the point estimate would be the mean of the bootstrap means, which is 0.373, and the corresponding 95% bootstrap confidence interval is (0.264, 0.462).

We can see from three intervals above that the one using finite sample correction factor has the smallest range, which means more informative. Its range is smaller than bootstrap interval because bootstrap is a nonparametric method, which uses less information and is therefore more conservative. Also, because it takes out the infinite population assumption in our standard Central Limit Theorem approach, it is therefore more precise with a smaller range for finite population.

Scenario 2:

In the questionnaire, sampled students in the class were asked to report the amount of time they spent on video games in the previous week and their usual frequency of playing video games. There are four categories of frequencies: “daily”, “weekly”, “monthly” and “semesterly”. In order to explore the relationship between their time spent on video game last week and their self-reported frequency of playing, we draw the boxplot of the amount of time playing video games for each frequency in Figure 2.1 and display the statistics of each frequency in Table 2.1.

Figure 2.1: Boxplot: Playing Frequency VS Playing Time Last Week

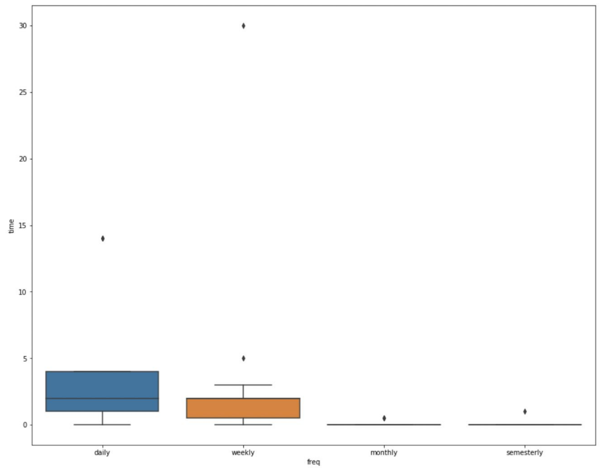


Table 2.1: Statistics of time spent playing video games grouped by reported frequency

frequency	count	mean	std	min	25%	50%	75%	max
daily	9.0	4.444444	5.570258	0.0	1.0	2.0	4.0	14.0
weekly	18.0	0.055556	0.161690	0.0	0.0	0.0	0.0	0.5
monthly	23.0	0.043478	0.208514	0.0	0.0	0.0	0.0	1.0
semesterly	28.0	2.539286	5.499046	0.0	0.5	2.0	2.0	30.0

Based upon Figure 2.1, when the frequency of playing decreases from "daily" to "semesterly", the reported time spent on video games in the previous week also decreases. Thus, there should be a positive correlation between time spent on playing video games and self-reported frequency of playing, which confirms our intuition that those who played more frequently tend to spend more time playing during the last week. From a closer look at boxplots of frequency "daily" and "weekly" in Figure 2.1, we notice that the third quantile overlaps with the "maximum" in the boxplot of frequency "daily" and the median overlaps with the 3rd quantile in the boxplot of frequency "weekly". Thus, based on statistics in Table 2.1, we claim that for those students who played daily, the upper quarter of them (there is only one outlier) all reported 4 hours spent on video games last week and for students who played weekly, everyone within median and 3rd quartiles reported 2 hours spent on video games.

Figure 2.2: Boxplot: Playing Time VS Playing Frequency (grouped by playing if busy)

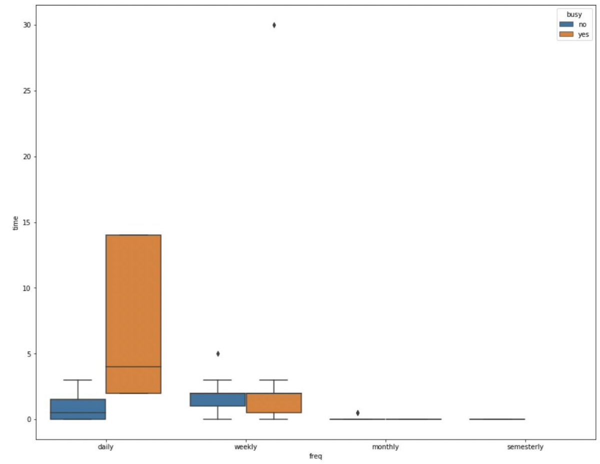


Table 2.2: Statistics of time spent playing video games grouped by play when busy

	count	mean	std	min	25%	50%	75%	max
busy								
no	63.0	0.509524	0.997208	0.0	0.0	0.0	0.5	5.0
yes	17.0	4.705882	7.782309	0.0	0.5	2.0	3.0	30.0

Figure 2.2 draw the same boxplot as Figure 2.1 but with students divided into two categories: whether the student play when busy. In order to investigate the effect of a coming exam on the previous estimates, we could simulate the “existence of an exam” by the variable “play if busy.” When the answer to “play if busy” is “yes”, we consider the student will play even if there is an exam coming. When the answer to “play if busy” is “no”, we consider the student will not play if there is an exam coming. From Table 2.2, we can see that students answer “yes” under the variable “busy” have both higher mean and median time of playing during last week than those answering “no”. This confirms our intuition that when they are busying preparing the exam, those answering “no” would play shorter time than those answering “yes”. However, we noticed that the standard deviation of time spent playing video game of those answering “no” is 0.997208, which is much smaller than that of answering “yes”, which is 7.782309. This means that among those answering “yes”, there are large variations of their time spent on video game before the exam. In Figure 2.2, the boxplot demonstrates the discrepancy of time spent on video game grouped by self-reported frequency between students answering “yes” and “no”. In the figure, the frequency “daily” has the greatest discrepancy between students answering “yes” and “no”, that is, students with self-reported frequency of “daily” tend to be affected the most by their state of being busy or not. Thus, students who play video games daily tend to be affected the most by the coming exam, those play weekly are affected less and those who play monthly and semesterly are affected the least. Therefore, we claim that there is a positive correlation between the frequency of playing and the potential of being affected by the coming exam, that is, the more frequent a student plays video games, the more his time spent on

video games tend to be affected by the coming exam. Thus, if there is not going to be an exam in the coming weeks, those who play daily and answers “no” when busy will spend more time on video games, enlarging the overall number of hours spent on video games. We claim that the presence of an exam in the week prior to the survey will skew the data.

Scenario 3:

In this part, we are trying to do an interval estimation from the survey data in order to generalize an inference on the mean playing time for those students. There are two ways to approach the solution. One is bootstrapping, and the other one is calculating directly from the confidence interval formula. In order to get the confidence interval from bootstrapping, we would like to generate a bootstrap population first, which mimic the population that we are desired for the undergraduate students at UC Berkeley, and then keep drawing samples from this population in order to get those samples' means. After that, we could rank the data and get the 2.5 and 97.5 quantiles to construct the endpoints of the confidence interval.

The other way of doing the confidence interval would be calculating directly from the confidence interval formula. The formula first calculate the point estimator, which would be the sample mean from the data to estimate the true population mean. Then it calculates the standard error and then times the t statistics for 2.5 and 97.5 quantile of the t distribution to get the interval. $(\widehat{x} - 1.97 * \frac{se}{\sqrt{n}}\sqrt{\frac{N-n}{N}}, \widehat{x} + 1.97 * \frac{se}{\sqrt{n}}\sqrt{\frac{N-n}{N}})$, where $\sqrt{\frac{N-n}{N}}$ is the finite population correction factor, which takes into account that we have a finite population of size 314. When we are doing surveys, the random selection of people to answer the survey would make the chance of the next person increase under finite population. As a result, the drawing of people to answer the survey questions would become dependent from each other under this assumption. We will calculate both with the correction factor and without the correction factor.

After bootstrapping and collecting the means, we got a confidence interval, which indicates that we have a 95 percent chance that if we keep drawing samples from the bootstrap population, the mean time of students playing games would fall in the interval from (0.593, 1.965). By calculating from the formula, we get the interval (0.456, 2.029) without the finite population correction, and (0.580,1.906) with finite population correction. Comparing those three confidence intervals, we could find out that the directed calculation from the formula with finite population correction gives us the narrowest interval, which might mean more useful or informative. Since the bootstrapping is non-parametric, it gives a wider interval considering that it has no assumption of the distribution. On the other hand, using the information that we expect and will get a normal like distribution for samples of many means under the central limit theorem, we could generate a shorter confidence interval by the formula.

Scenario 4:

By calculation, 76.92% of students reported that they at least somewhat like the game(like=2 and like=3), so students like to play video games in general. Figure 4-1 shows the popularity of each game type among surveyed students. Table 4-1 shows that the Strategy and Action games are the most popular ones among students.

Figure 4-1 Histogram of the Types of Games Students Play

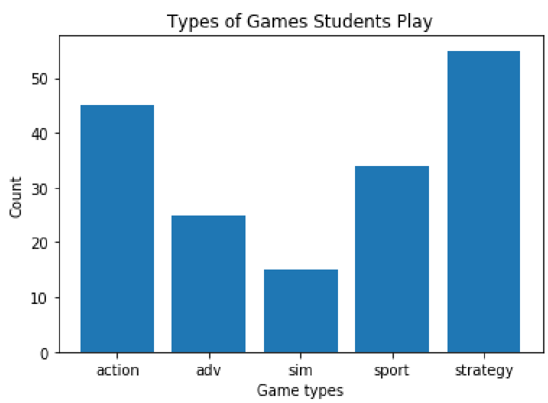


Table 4-1: Count of the Type of the Games that Students Play

Type	Count	Percentage
Action	**45**	**50%**
Adventure	**25**	**28%**
Simulation	**15**	**17%**
Sports	**34**	**39%**
Strategy	**55**	**63%**

Figure 4-2: Histogram of the Reasons Why Students Like Games

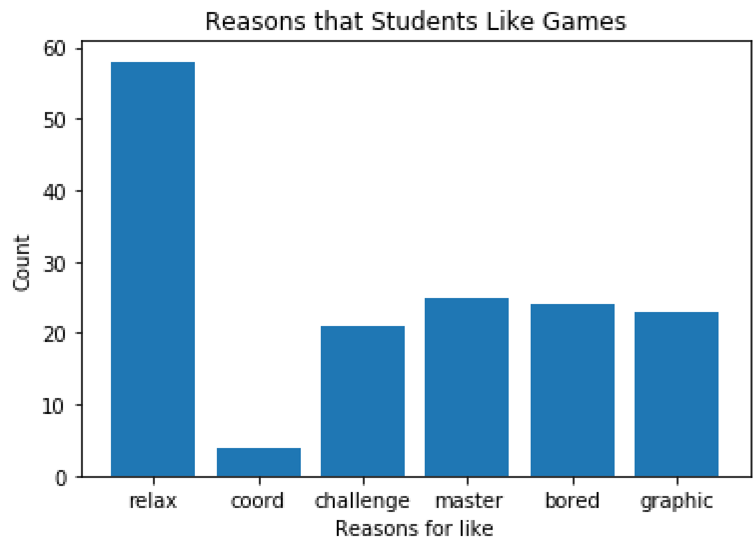


Table 4-2: Count of the Reasons Why Students Like to Play Games

Reason for like	**Count**	**Percentage**
Graphics/Realism	**23**	**26%**
Relaxation	**58**	**66%**
Eye/hand coordination	**4**	**5%**
Mental challenge	**21**	**24%**
Feeling of mastery	**25**	**28%**
Bored	**24**	**27%**

Figure 4-2 shows the reasons that students like to play games. In the survey, students were asked to choose at most 3 main reasons that they play games. According to the table 4-2, the three biggest reasons that students play the games are relaxation, feeling of mastery and bored, led by 66%, 28%, 27%. Note that relaxation became the biggest reason that students play games and it is the only reason that surpass 50%. It can be further inferred that, students mostly play games to relax from the heavy workload of school.

Three reasons that students play games:

1. Students play games to relax from the heavy workload
2. Students play games to feel of mastery
3. Students play games to fill up their free time

Figure 4-3: Histogram of the Reasons Why Students Dislike to Play Games

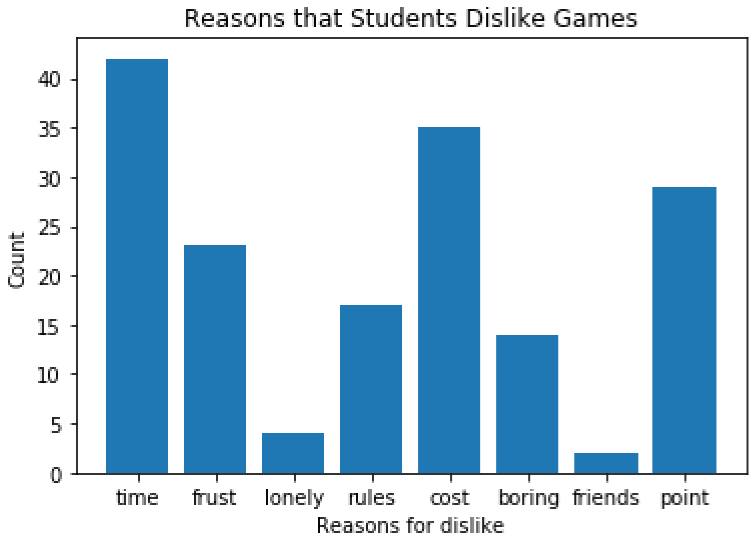


Table 4-3: Count of the Reasons Why Students Dislike Games

Reason for dislike	Count	Percentage
Too much time	**42**	**48%**
Frustrating	**23**	**26%**
Lonely	**4**	**6%**
Too many rules	**17**	**19%**
Cost too much	**35**	**40%**
Boring	**14**	**17%**
Friend's don't play	**2**	**17%**
It is pointless	**29**	**33%**

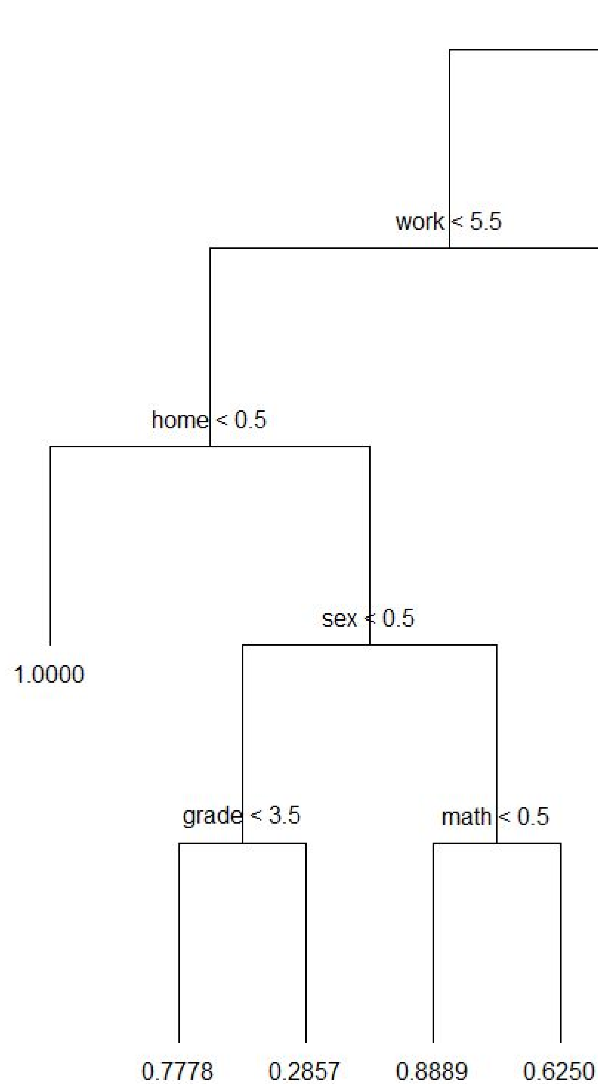
Figure 4-3 shows the reasons why students dislike to play games. In the survey, students were asked to choose at most 3 main reasons that they do not play games. According to the table 4-2, among the students who do not like to play games, three biggest reasons are too much time, costs too much and pointless. It can be inferred that such students believe that the main duty of the students is to learn from the schools, and since playing games won't bring them the knowledge they want, they could have spent the time on studying course materials.

Three reasons that students do not play games:

1. Students do not play games due to waste of time
2. Students do not play games due to limit of their budget

3. Students do not play games because they do not think it's educational

Figure 4-4: Regression Tree of Videodata



In a regression tree the idea is: since the target variable does not have classes, we fit a regression model to the target variable using each of the independent variables. Then for each independent variable, the data is split at several split points. At each split point, the "error" between the predicted value and the actual values is squared to get a "Sum of Squared Errors (SSE)". The split point errors across the variables are compared and the variable/point yielding the lowest SSE is chosen as the root node/split point. This process is recursively continued. The figure above provides the prediction from video data.

From the tree, 97.06% of students who worked more than 5.5 hours prior to the survey like video games, which might be explained by the fact that, after tiring work, they want to find relaxation though games. Interestingly, all of those who worked less than 5.5 hours and do not have a computer at home like playing games. Moreover, for male students who worked less than

5.5 hours, do not have a computer at home and do not expect to get A, only 28.57% of them like playing video games.

Scenario 5:

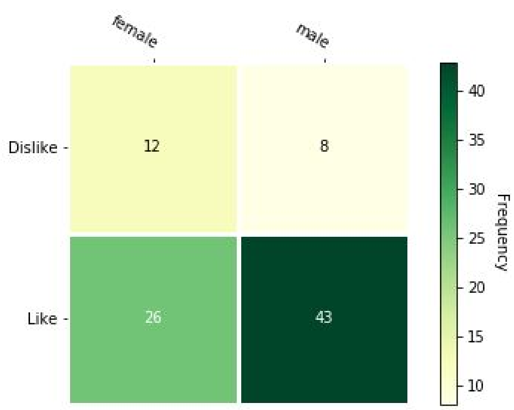
For classifying students into two groups of people who like to play video games and those who don't, we define that students who responded "very much" and "somewhat" to the like question are those who like to play video games, and those who responded "not really" and "not at all" are those who do not like to play video games. The remaining category of students who responded "never played" will be purged for this particular analysis. This category only contains 1 student, so it does not matter much to this grouping method.

Category	Like to play video games	Like to play video games
Responses	2: "Very much", 3: "somewhat"	4: "Not really", 5: "not at all"

Sex

Cross-tabulation for sex

Sentiment\Sex	female	male
Dislike	12	8
Like	26	43



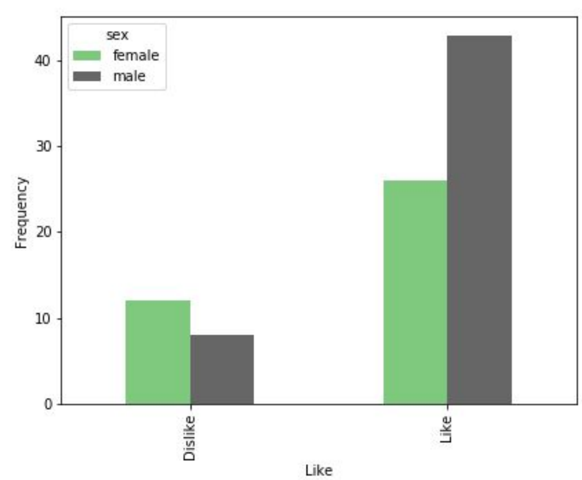


Figure 5.1: Heatmap for contingency table for likeness by sex (left)
Figure 5.2: Bar chart for likeness by sex (right)

From these graphs, we see that both males and females are more likely to have a positive sentiment to video games while those who don't like video games are more likely to be a female. To test if there is a statistical significance between these two categories, we use Pearson's Chi-square test with a degree of freedom of 1. The null hypothesis is if sex is independent of likeness of video games in this survey. We get:

$$\mathscr{X}^2 = 2.311,\ \text{p-value} = 0.128$$

Since p-value 0.128 is larger than 0.05, we fail to reject the null hypothesis under the 0.05 significance level, meaning a person's likeness to video games is independent of sex.

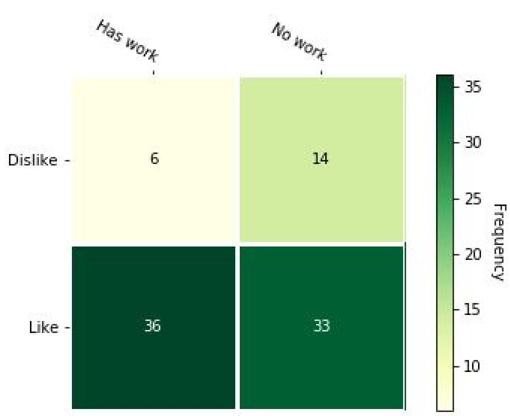
Work

We classify those who have a response time of more than 0 hour of work in the previous week as people with work, and those who have 0 hour of work in the previous work don't have a work.

Category	Work	No work
Responses	X > 0	X = 0

Cross-tabulation for work:

Sentiment\Work	Has work	No work
Dislike	6	14
Like	36	33



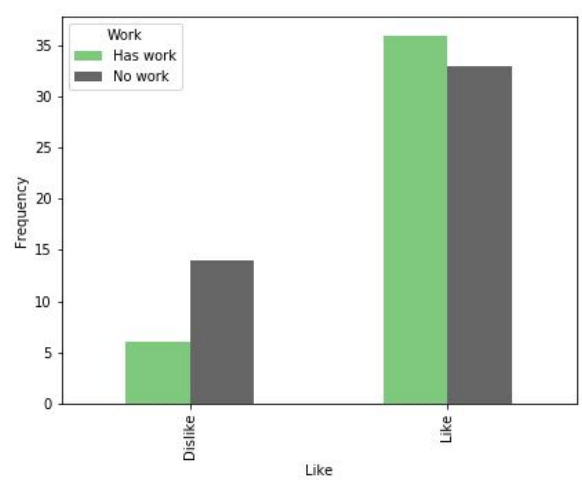


Figure 5.3: Heatmap for contingency table for likeness by work (left)
Figure 5.4: Bar chart for likeness by work (right)

From these graphs, we see that both students who have work and those who don't work are more likely to have a positive sentiment to video games while those who don't like video games are more likely to not have work. To test if there is a statistical significance between these two categories, we use Pearson's Chi-square test with a degree of freedom of 1. The null hypothesis is that having work is independent of liking playing video games, and the alternative hypothesis is that having work is dependent of liking playing video games. We get:

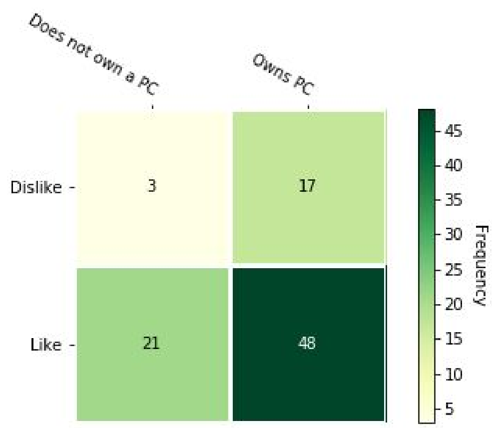
$\mathcal{X}^2 = 2.234$, p-value = 0.135

Since p-value 0.128 is larger than 0.05, we fail to reject the null hypothesis under the 0.05 significance level, meaning a person's likeness to video games is independent of having to work.

Owns PC

Cross-tabulation for owning a personal computer

Sentiment\Sex	Does not own PC	Owns PC
Dislike	3	17
Like	21	48



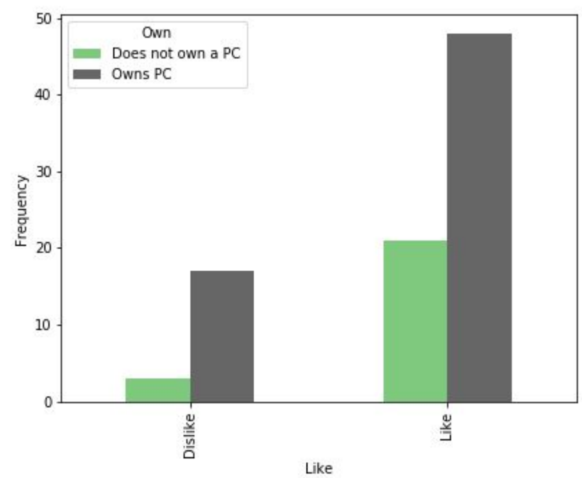


Figure 5.5: Heatmap for contingency table for likeness by own (left)
Figure 5.6: Bar chart for likeness by own (right)

From these graphs, we see that both students who own a personal computer and those who don't have one are more likely to have a positive sentiment to video games. To test if there is a statistical significance between these two categories, we use Pearson's Chi-square test with a degree of freedom of 1. The null hypothesis is that owning a PC is independent of liking playing video games, and the alternative hypothesis is that owning a PC is dependent of liking playing video games. We get:

$\mathscr{X}^2 = 1.174$, p-value = 0.279

Since p-value 0.128 is larger than 0.05, we fail to reject the null hypothesis under the 0.05 significance level, meaning a person's likeness to video games is independent of owning a personal computer.

Advanced Analysis (Scenario 6):

To further investigate the grade that students expect, we want to compare the distribution of expected grades by students and the target distribution in grade assignment. Although it is intuitive to do Chi-Square Goodness of Fit Test, we need to make sure that the assumptions are all met. First, Chi-Square Goodness of Fit Test is appropriate when the sampling method is simple random sampling. In this case, each subset of the population has the same chance of being selected, so these 91 students are a simple random sample. Second, the expected value of number of observations in each category must be at least 5. In a sample of size 91, if the distribution of grades expected by students matches grade assignment, which is 20% A's, 30% B's, 40% C's, and 10% D's or lower, the expected number of students who want to get each grade are summarized in table 1. Therefore, we can see that each number is greater than 5. Thus, assumptions are all met, and it is safe to do Chi-Square Goodness of Fit Test.

Table 6.1: Expected value of number of students expecting each grade

Grade	A	B	C	D&F
Number	18.2	27.3	36.4	9.1

Here, the hypotheses are that H_0 : the observed grades expectations are consistent with the grade assignment, and H_1 : the observed data are not consistent with the grade assignment. We set the significance level to be 0.05. Table 2 summarizes the observed number of students who expect each grade. Performing Chi-Square Goodness of Fit Test, with degree of freedom 3, we get a p-value of 1.6288*10^(-13), which is extremely smaller than 0.05. Therefore, we reject the null hypothesis and conclude that the grades that students expect in this course does not match the target distribution in the grade assignment.

Table 6.2: Observed value of number of students expecting each grade

Grade	A	B	C	D&F
Number	31	52	8	0

The result matches our intuition because students tend to expect better grades than what they are capable of. So there tend to be more students expecting A and B and less students expecting C or even D, causing a significant difference between the two distributions.

Furthermore, we want to know the effect of nonrespondents on the result. If most of them are failing students who no longer bothered to the discussion section, they are likely to expect to pass the class with a final grade of C or higher. Assuming most of nonrespondents expect to get C, the real distribution of grades expected by students will change. The percentage of C will be higher while that of A and B will be lower. The observed distribution is therefore closer to the expected one, leading to a smaller statistic and larger p-value. We may or may not reject the null hypothesis. Thus, to make sure that the sample is representative, it is important to minimize the number of nonrespondents who might change the whole picture dramatically.

Further Discussion and Conclusion

From scenario 1, we found that 37.4% of students played video games a week before the survey when there was an exam. Therefore, we might expect 37.4% of students to go to the lab when there are exams. Since it is a relatively small size of lab, the content of the lab might be designed to be more interactive.

According to the findings from scenario 2, there is a strong association between the number of hours students spend on video games and whether the students play if busy. Typically, if the student is busy, he tends to spend less time playing video games. Thus, we infer from such findings that the existence of an exam will diminish the chance that students come to the lab. Thus, it is recommended that the lab hours should be arranged to avoid exam periods.

In scenario 3, the 95% confidence interval of the mean time of playing video games from 0.58hr to 1.9hr. We may infer from this finding that students immerse themselves into video games usually less than 2 hours. Thus, in order to attract students' attention most concentratedly, the lab should not be designed over 1.9 hours.

Furthermore, scenario 4 tells us that 76.92% of the students at least somewhat like playing video games. Therefore, the faculty should expect 76.92% of the students to go to the lab. As action and strategy games are most popular among students, the faculty might consider incorporating these elements into the lab. For example, the contexts of statistics problems may be related to designing the best strategies for a captain. Since 66% of students have listed "relaxation" as one of the reasons why they like playing video games, it is better that the lab is designed to have a relaxed instead of intensive atmosphere. Also, costing too much time and money and being pointless are the top 3 reasons why students dislike video games. In order for the lab to attract more students, its duration should not be too long and each lab's purpose should be straightforward.

Scenario 5 reveal the fact that in the survey, the sampled student's favor over video games is independent of sex, if the student have to work and if the student is owning a personal computer.

From scenario 6, the distribution of grades expected by students and that of grade assignment are significantly different. More students expect themselves to get A or B and less students expect to get C or lower. If nonrespondents' data are included, the distribution will be closer to what we expect.

Theory

Regression tree: A regression tree is built through a process known as binary recursive partitioning, which is an iterative process that splits the data into partitions or branches, and then continues splitting each partition into smaller groups as the method moves up each branch. It is used to predict the values of a continuous variable from one or more continuous and/or categorical predictor variables.

Chi-square test of independence: The chi-square test of independence is a test to check if there is a significant association between categorical variables. In scenario 5, we check if the preference of video games are related to other variables such as whether the student owns a pc, etc. To reduce the degree of freedom and ensuring we have an expected frequency count to be more than 5 for each cell of the table, we reduced the number of choice in a categorical variable to be 2, thus making the overall degree of freedom (2-1)*(2-1) = 1. First, we specify a null hypothesis that assumes there is an independent relationship between two variables, and we verify if the test statistic can reject the null hypothesis. We then calculate the expected frequencies of each cell. This is done by multiplying the column frequency and row frequency and divided by the total sample size. The test statistic is then calculated by the sum of all squared deviations from the expected cell size and the observed cell size over the expected cell size. Finally, we find the p-value for this test statistic given the degree of freedom. If the p-value is less than a set significance level, say 0.05, the null hypothesis is rejected, and we can conclude that there is a relationship between two variables.

Chi-square Goodness of Fit Test: Chi-squared goodness of fit test is used to determine whether the sample data matches the hypothesized distribution. This is a perfect match to our purpose in scenario 6 because we want to investigate if the grades students expect match the grade distribution according to grade assignment. In this case, we have 4 categories for grades and we do not need to estimate any parameters. Therefore, the degree of freedom is 4-1=3. To get the test statistics, given sample size 91, we calculate the sum of squared distances between expected and observed number of students expecting each grade divided by the expected number. Lastly, we get the p-value from chi-square table and decide whether to reject the null hypothesis, which is that grades expected by students match the grade assignment distribution.

Confidence Interval:

Confidence interval, in statistics, is one kind of interval estimation. It uses the observed data to compute an interval for the true population parameter which might contained in that interval calculated from the data. Related with confidence interval, there is confidence level. As the confidence level increases, we would have a wider interval for the true parameter. In a rigorous way, confidence level is the frequency (i.e. the proportion) of possible confidence intervals that contain the true value of the unknown population parameter. Confidence intervals consist of a range of potential values of the unknown population parameter. However, the interval computed from a particular sample does not necessarily include the true value of the parameter. In that case, we would replicate the experiment and calculate multiple confidence interval in order to capture the true parameter.

Bootstrap:

Bootstrap, in statistics, is the practice of estimating properties of an estimator such as variance when random sampling from a finite population with replacement. Since the sample from the original distribution with simple random sample method shares the similar distribution with the original one, the bootstrap population from this random sampling can be considered as a new population generated with the same process as the original distribution. By calculating sample statistic from many bootstrap distribution, an empirical confidence interval of such statistic can be computed by taking the corresponding percentile values from the distribution of these resampled statistic by bootstrapping.

Work cited:

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist, 69*(1), 66-78.

Green, C. Shawn, and Daphne Bavelier. “The Cognitive Neuroscience of Video Games.” *Learningtransferlab.wiscweb.wisc.edu*.

J. Torrente, Á. del Blanco, E. J. Marchiori, P. Moreno-Ger and B. Fernández-Manjón, "<e-Adventure>: Introducing educational games in the learning process," *IEEE EDUCON 2010 Conference*, Madrid, 2010, pp. 1121-1126.

Primack, Brian A., et al. “Role of Video Games in Improving Health-Related Outcomes: A Systematic Review.” *NeuroImage*, Academic Press, 15 May 2012, www.sciencedirect.com/science/article/pii/S0749379712001729.