**Music Preferences among University Students**(Ivan Izhbirdeev, Anqi Lin, Ranjana Reddy, Yifan Zhang)

**Abstract:**

For our project, we are trying to find the correlation between music preferences and five other variables, such as GPA, Major, activity, region and age. The method that we use to collect data is survey mainly and direct observation. To test our hypothesis of each variable, we show that distribution of GPA and categorize qualitative variables. According to our analysis, there is not much correlation between the time spent on listening to music and GPA. There is a correlation between music genre and major. The most common activity associated with music is leisure.

**Introduction:**

In our study we analyzed music preferences among university students. Mainly we were interested in any correlation that arises between music genres, hours spent listening and students’ GPA scores, or other variables that are described in this paper. The main questions that we tried to answer were:

1. How does the number of hours students spend listening to music affect their GPA?

2. Correlation between music genre preference and major

3. Is listening to music associated with some activity (e.g. studying, resting)?

4. Correlation between music preferences and region of origin

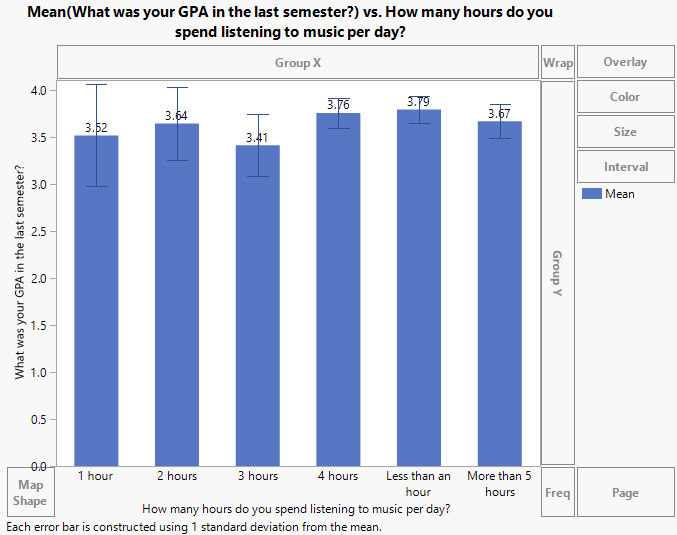
5. How age influences music preferences?

For each of the questions we tried to prove or disprove that there exists a described correlation.

**Methods:**

* Data collection method: survey (main one), direct observation (alternative)
* Send to all classmates in MA 213
* Send to the group chat in other classes in BU
* Population: all BU students
* Sample: Students in our class

**Data visualizations and inference results:**

****

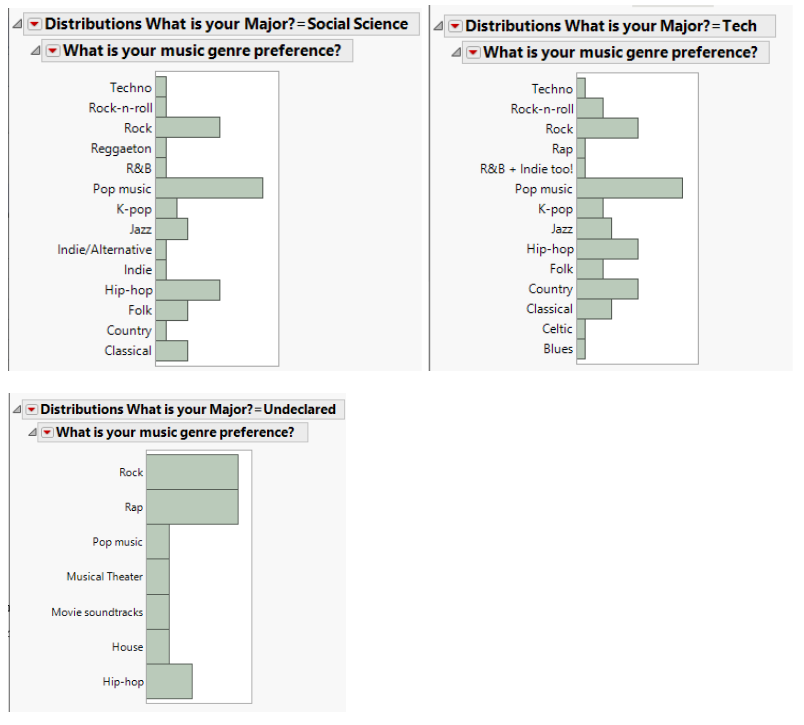
**How does the number of hours you spend listening to music affect your GPA?**

Students spend between 1 to 5 or more hours listening to music everyday.

As it can be observed from the graph

above, time spent on music **does not impact** the students’ GPA in a major way.

Students score approximately the same GPAs without high disparities. The outliers that can be found are from students who are attending their first semester at BU and do not have a GPA to report yet.

****

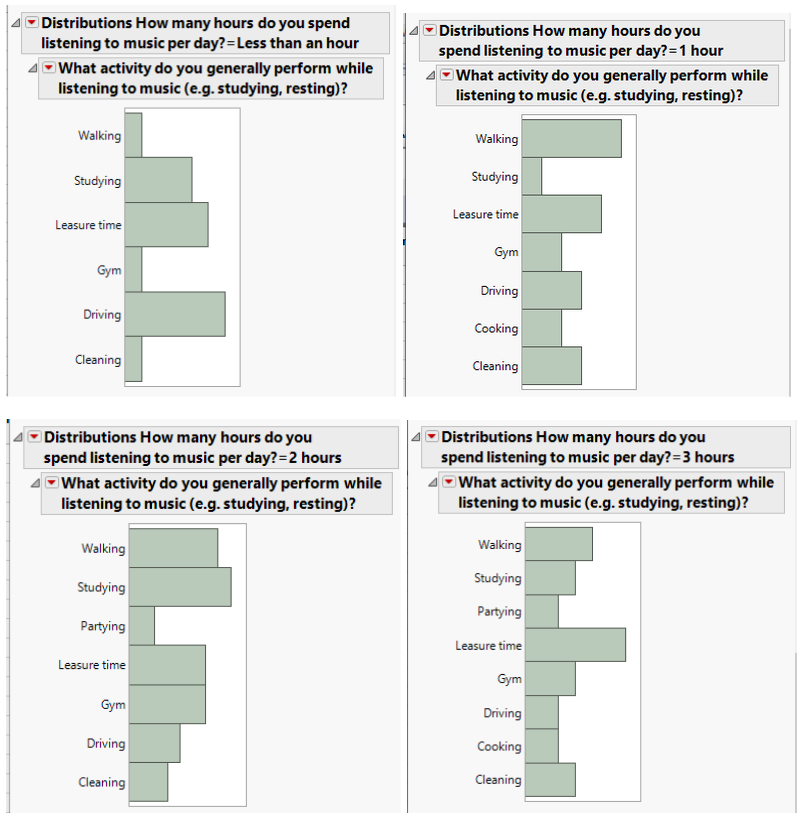
**What’s your music genre preference and your major’s category?**

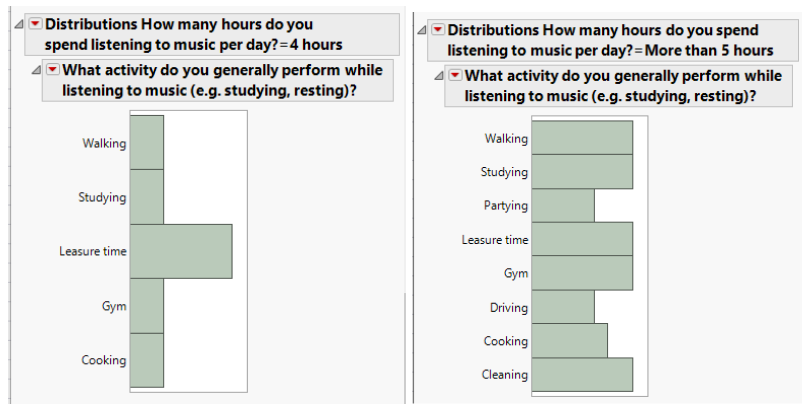
The above graphs show the correlation between the students music genre preference and their major’s category. The students of MA 213 class could be classified into 3 categories; **Social Science, Tech, and Undeclared.**

We can observe that most students in the **Tech and Social Science category** have a high preference towards pop music, rock, and hip-hop genres.

It can also be observed that a considerable amount of students in the **Social Science major** category prefer older music like folk, classical, reggaeton, and jazz. We can also notice that the **Tech major category** students are the only ones that have a preference towards celtic, blues, and R&B + Indie genres.

The **Undeclared major category** students have an equal preference towards rock and rap, followed by hip-hop. This group of students have the least diversity in music genre preference among the three categories mentioned above.

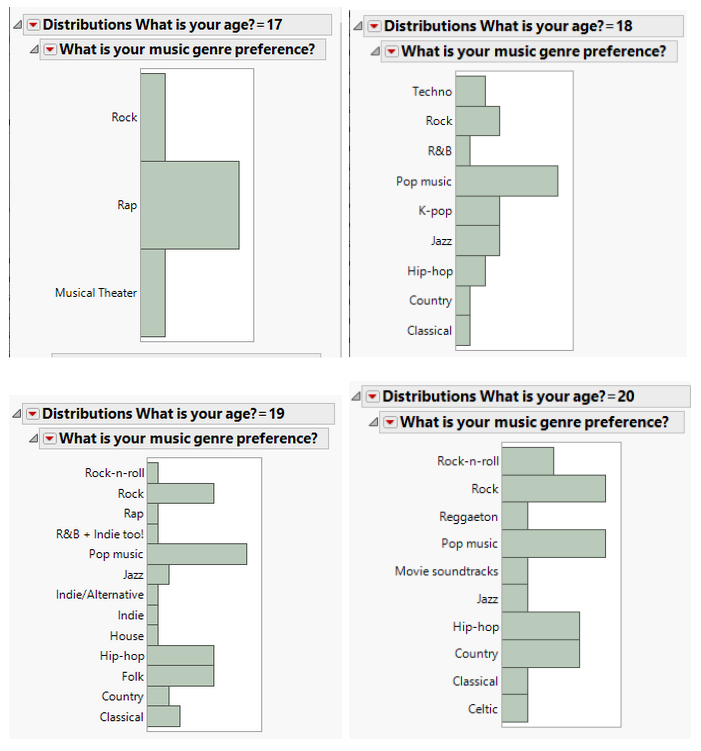
****

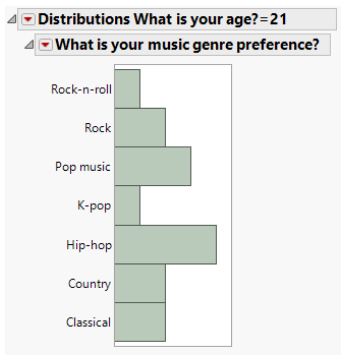
****

**Most popular activities associated with listening to music:**

The most common activity associated with music is leisure time. There is a pattern between the number of hours spending on listening to music and activities.

When people listen to music for less than 3 hours, a specific activity is outstanding, such as walking, leisure time and driving. When people spend around 4 hours, except one outstanding activity, others start to become average. When people spend more than 5 hours on music, almost all activities tend to be average.

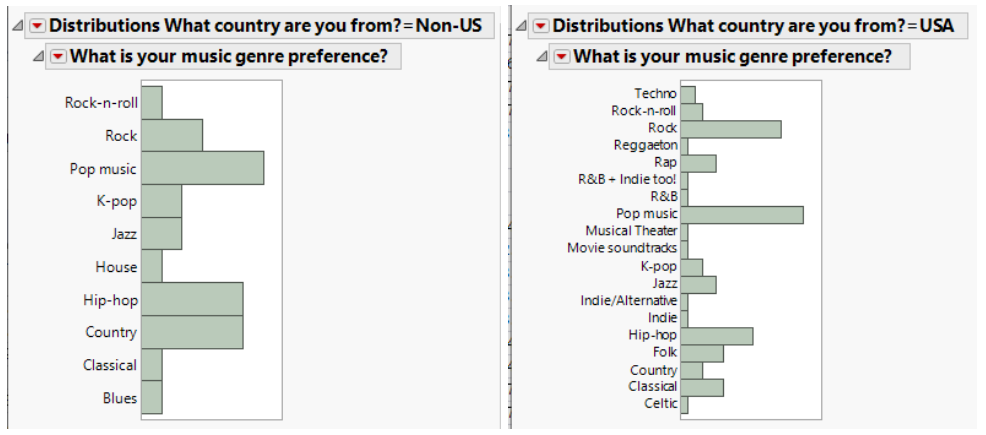




**Correlation between music genre and age:**

In the analysis of the data for age distribution, we found one outlier - age of 22 years old. Only two people were of this age, so we disregarded it in our final analysis. We can see that people of smaller age have less diversity in their music preferences compared to older students (e.g. for people of age 17, there are only 3 prominent music genres, with rap dominating the graph). Again, the overall most popular music genres were pop music and rock, however, genres such as hip-hop, country, and folk were

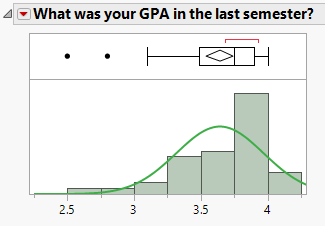
prominent for some age groups.



**Correlation between music genre and region of origin**

For the analysis of the correlation between music genre and region of origin, we divide people into two categories:

Non-US and USA. The reason that we categorize people like this is because either dividing USA people by states or dividing other people by countries, there is not enough data to support. From the above graph, we can see that pop music is popular in both categories, and for non-US, mainly Asian in our data, hip-hop and country music are in higher preference than for people from the USA. Alse, for people from the USA, the most popular music genre other than pop music is Rock.



**Normally Distributed Variable:**

The p-value is less than 0.001, which means we need to reject the null hypothesis - it’s normally distributed.

However, since we only have a sample size around 40, which is not a large sample size. Also, the survey is generally taken by students from introductory courses like MA 213. Generally we believe that GPA should be normally distributed.

**Random Sampling:**

For a sample to be truly random, every subject has to have an equal chance of being selected. Our survey was handed out to the students and we got 43 responses. Since our pool of responses was low, we were able to pick all of them for our data analysis. Another important aspect of a random sample is that every member of the population has equal opportunity to respond to it. The survey was sent to all the students via email so all of them got an equal opportunity. For certain responses like Majors and region of origin we had a lot of varied answers, so we narrowed the responses down to simplified categories. Our dataset had a considerable amount of diversity, which is reflected in our analysis. Hence, we have a random sample here.

**Confidence Intervals:**

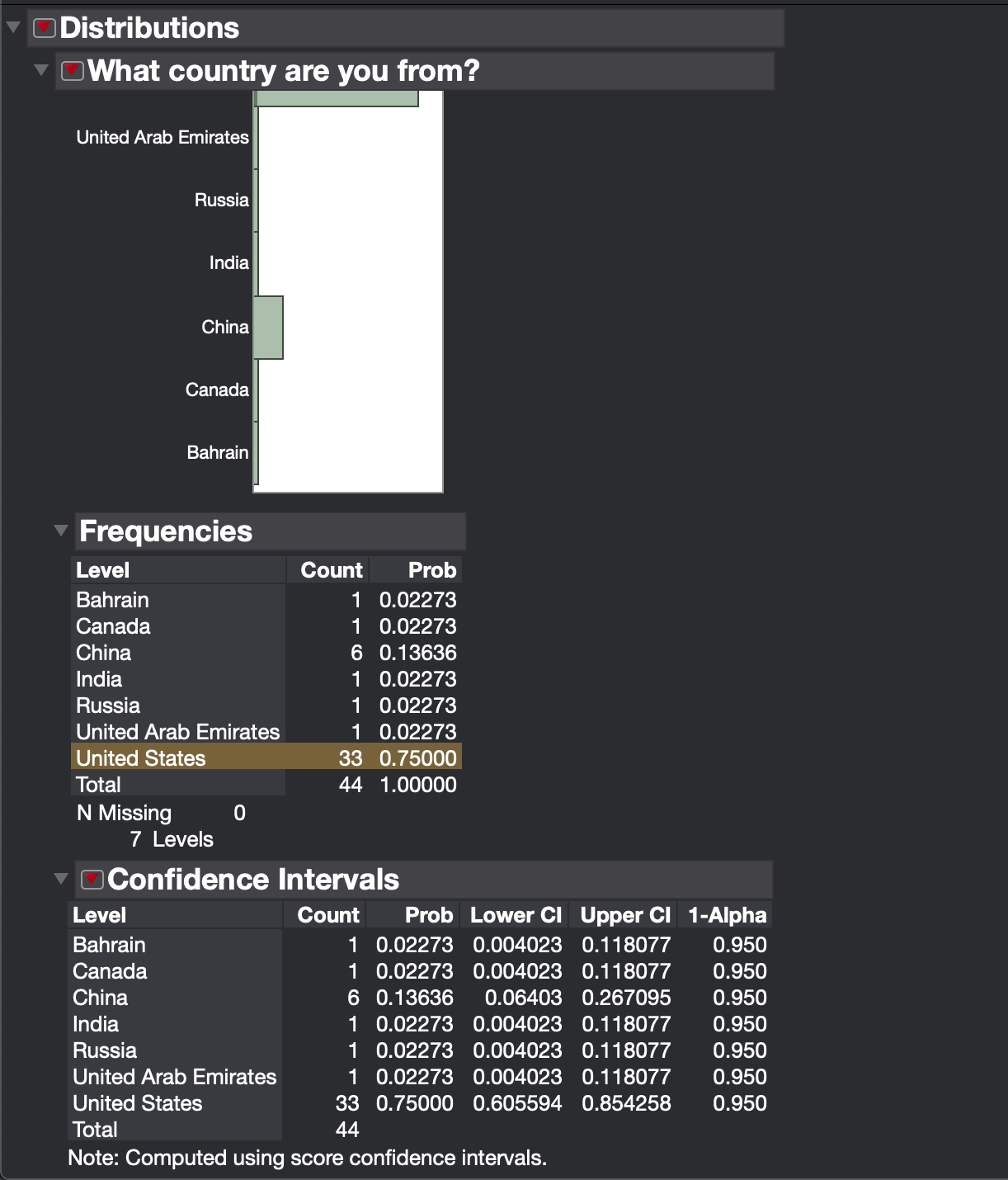
**Qualitative variable: Country**

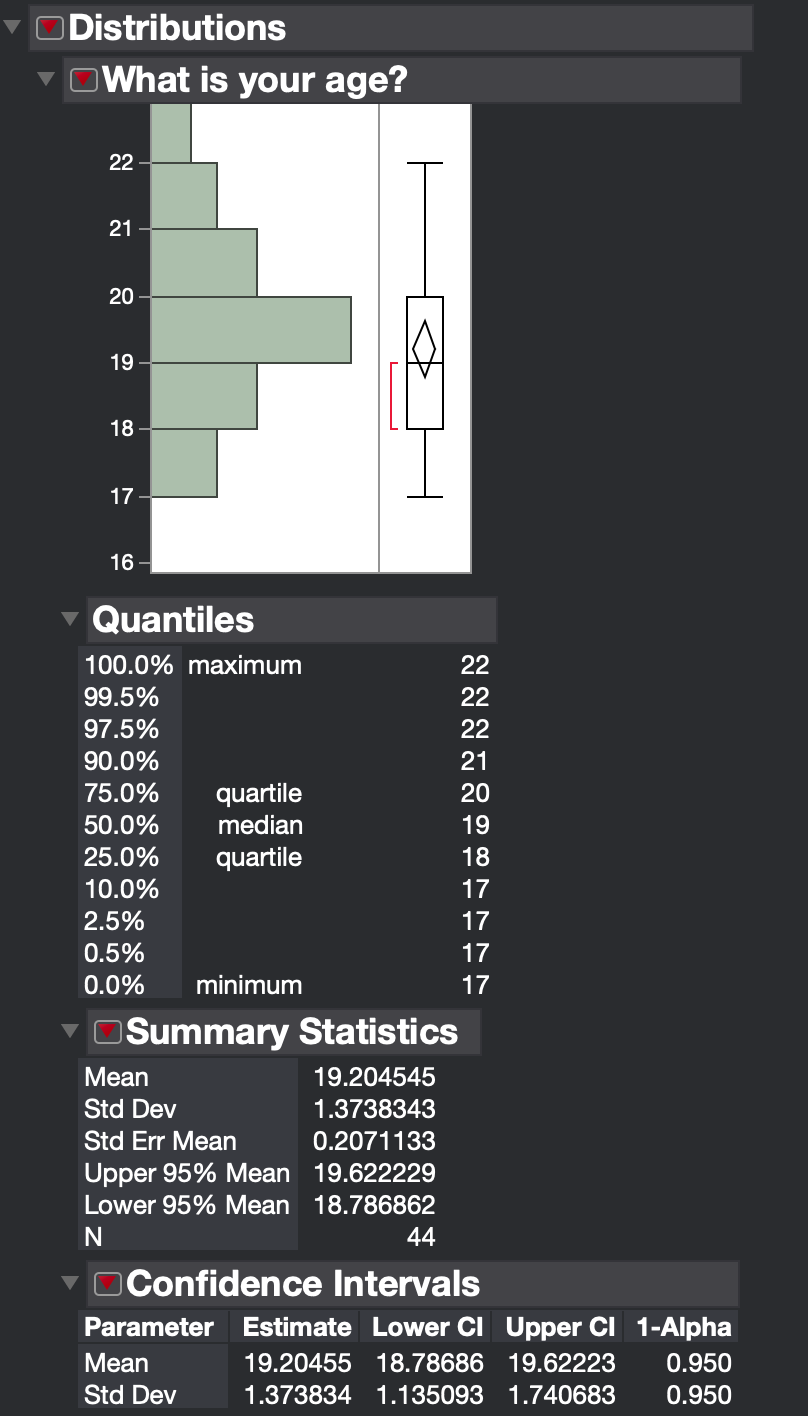
We could only find the confidence interval for United States as we did not have

sufficient data on other countries as their sample size was small (𝑛𝑝≥15 and 𝑛𝑞≥15)

For the region parameter, we are 95% confidence that the population proportion

mean of region lies between 0.0606 and 0.854





**Quantitative variable: Age**

Our sample size is > 30, so we can assume the distribution of the sample mean.

For the Age parameter, we are 95% confidence that the population proportion mean of the values lies between 18.79 and 19.63

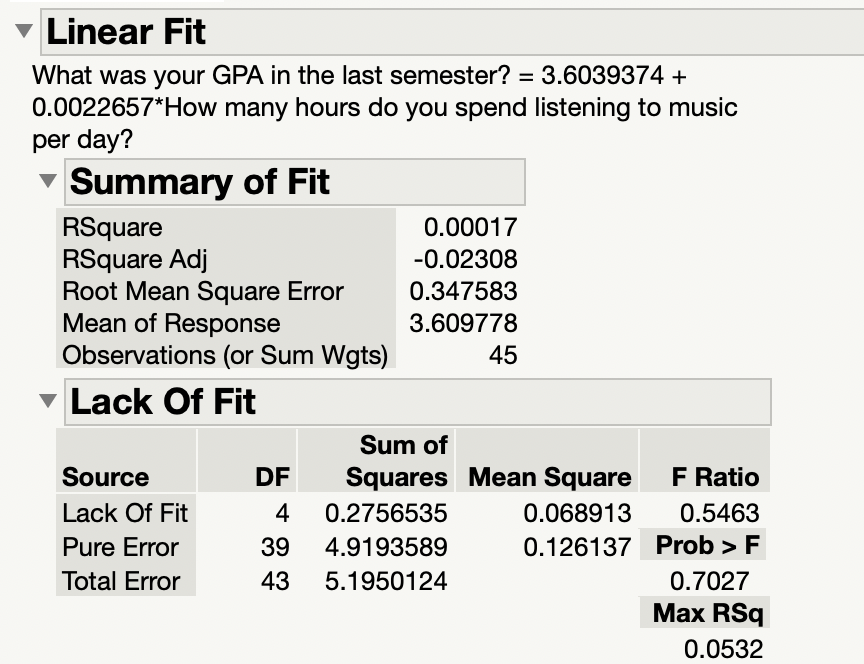
**Hypothesis Testing:**

**\*Due to the special cases of our survey result, we can only do one hypothesis testing for our five research question. The rest of the questions can be answered by the data analysis part above.**

1. How does the number of hours students spend listening to music affect their GPA?

Null hypothesis: The correlation coefficient IS NOT significantly different from 0. There IS NOT a significant linear relationship (correlation) between hours spent listening to music and GPA in the population.

Alternative hypothesis: The correlation coefficient IS significantly different from 0. There IS a significant linear relationship (correlation) between hours spent listening to music and GPA in the population.



Significance level: 5%

As we can see, there’s only 0.00017 Rsquare, which means only 0.017% of the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model.

Because p-value is greater than 0.05, we do not reject the null hypothesis.

Conclusion: There IS NOT a significant linear relationship (correlation) between hours spent listening to music and GPA in the population.

2. Correlation between music genre preference and major

Since we have two qualitative variables here, we cannot conduct a hypothesis testing.

3. Is listening to music associated with some activity (e.g. studying, resting)?

Same as above, since we have two qualitative variables here, we cannot conduct a hypothesis testing.

4. Correlation between music preferences and region of origin.

Same as above, since we have two qualitative variables here, we cannot conduct a hypothesis testing.

5. How age influences music preferences?

Same as above, since we have one qualitative variables here, we cannot conduct a hypothesis testing.

**Conclusions and Discussions:**

* According to our findings, there is not much correlation between the time spent on listening to music and GPA as students score approximately the same GPAs without high disparities.
* We can observe that there is a correlation between music genre and major, as different majors prefer different genres of music.
* We can also observe that the most common activity associated with music is leisure.
* For the Age parameter, we are 95% confidence that the population proportion mean of the values lies between 18.79 and 19.63
* - For the region parameter, we are 95% confidence that the population proportion mean of region lies between 0.0606 and 0.854
* For the age group of 17 years, the overall most popular music genres were pop music and rock
* In the analysis of the data for age distribution, we found one outlier - age of 22 years old. Only two people were of this age, so we disregarded it in our final analysis.
* We can observe that pop music is popular in both categories, and for non-US, mainly Asian in our data, hip-hop and country music are in higher preference than for people from the USA.
* We hope that our findings will help researchers understand the various preferences/trends that are present among college students.

**Appendix:** Survey questionnaire can be found at<https://forms.gle/Zut2zXUQktaVhAgh6>

**References:**  McClave and Sincich. *“Statistics” 13th Ed*

**Participation:** Ivan Izhbirdeev, Anqi Lin, Ranjana Reddy, Yifan Zhang. All work was split evenly among four of us. We met up together for every deliverable and worked as a collective group.