Jason Larosiliere

Professor Nguyen

Applied Mathematics & Statistics Capstone Seminar

10 May 2023

Capstone Final Project

Intro:

It is shown that youth culture and people’s perceptions constantly change over time. At one point something is not acceptable and then out of nowhere that same thing is cool and its all the rage. Japanese anime has experienced something like this with its fans. The anime industry has seen a massive boom here in America, due to the advancement of social media and networks broadcasting these shows. Although, I just recently started watching anime in highschool, it was something that my brother was very passionate about. Eventually, he persuaded me to give it a try and I found that anime is deeply different from other genres. My goal for this project is to analyze popularity (members count) of animes using data from MAL. I think gaining further insight into the factors that cause high popularity can help people distinguish anime from each other. Essentially, it would also prove useful for production teams and up incoming manga artists when trying to make an anime that people will enjoy. At the end of the day, anime is something I have fell deeply in love with in the past few years, so this is more than good enough a reason to analyze this topic.

Literature Research:

Anime is a form of cartoons created in Japan that follow a certain design type. Although it has always been popular in Japan, it started gaining popularity in America in the 1960s, and has increased in traction ever since. This paper will look at and review different sources on the popularity of anime in America, as well as some of the most popular animes to date. It is interesting to see culture from another country make such an influence on Americans and the impact that anime has had on American media, and review will attempt to look into the reasons why it has become popular as well as subcultures within the anime fandom.

The first study is “Study on Anime and Its Impacts Among University Students.” The authors are Nor Hassan, and Iza Sallehuddin from University Tunku Abdul Rahman (UTAR). A survey of 247 students in Malaysia was conudcted to survey to see how the perception and attitudes towards anime, affects society’s behavior. The study included the initial attitude towards anime, acceptance of anime and the post acceptance behavior. An independent t-test to compare male and female students and a pilot test to check reliability of survey. As a result, it showed how youth culture’s perception on anime has changed tremendously over time. It’s important to understand people’s attitude towards anime to analyze why it has so much of an impact today.

The second study that we will look at was published in Bad Subjects: Political Education for Everyday Life in 1994, and is titled “Anime Otaku: Japanese Animation Fans Outside Japan”. This essay starts by discussing the Japanese term “otaku”, which is translated in English to mean “a person who is so involved with a particular type of fan subculture that he or she becomes obsessed, even insane” (Newitz, 1994). While it is seen as an insult in Japanese culture, Americans use the term with adoration, to show appreciation for their love of anime. This author did research on what the target demographic for anime watchers in America is. According to the text, “the target audience for anime culture in America: racially diverse men between the ages 18-25.” (Newitz, 1994). This study also discusses some of the reasons as to why so many Americans enjoy the anime stories. In many animes, there are predictable character stereotypes and plotlines that follow a romantic comedy theme. While this is not always the case, as there are plenty of popular animes that are action packed and filled with “magical powers”, there are a majority that are seen as realistic yet showing fantasy stories at the same time. This study ends with the author stating that American anime fans believe that Japanese culture is superior to American culture, and how that can be a “danger” if taken too far.

The third study to be discussed was written by Susan Napier in 2004, and is titled “Why Anime?”. This author conducted interviews and passed out questionnaires in order to answer the question as to whether or not there is a “typical anime fan” in America. What she found was that “anime fans cannot be typecast either politically or socially” (Napier, 2004). She also attempted to determine what attracts American fans to watching anime. Many of the answers showed that many adults enjoyed these shows because they were cartoons with much more complex stories than those they watched as children, such as Disney. They also said that anime was more “realistic” than Hollywood movies, because often times the main characters die or do not get a happy ending such as mainstream American media portrays, or that they find the characters relatable. An interesting point that the author makes is that “what seems evident in these replies is that these young Americans do not see the protagonists of anime as strongly Japanese. Instead, they seem to identify with them as fellow humans” (Napier, 2004). This reference gives us a good idea as to what different reasons Americans enjoy animes and that there are no barriers on who enjoys this type of media.

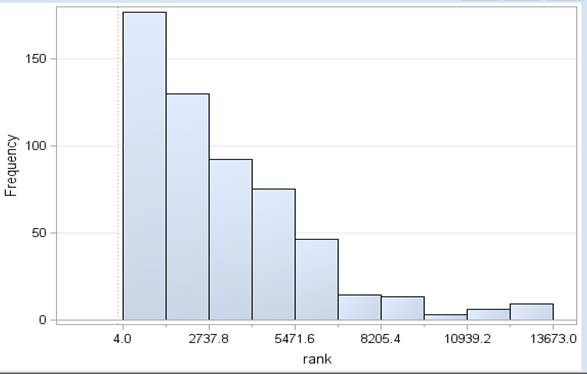
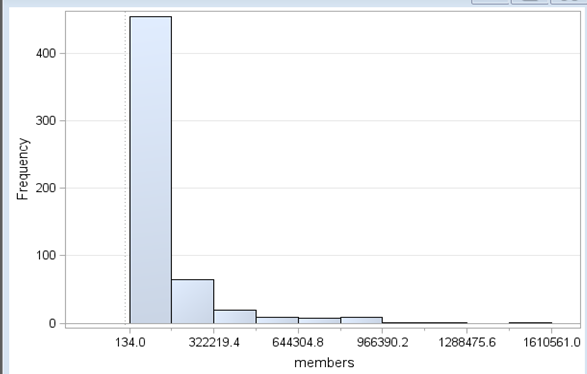
The last study that we will look at will shed light on a different popular area of anime, which is Manga Cosplay. The article is titled “Costuming the Imagination: Origins of Anime and Manga Cosplay”, written by Theresa Winge in 2006. Cosplaying is when people get together and dress up as their favorite characters from the anime that they watch. Around the world, including in America, there are large conventions held for people to show up in costume and meet creators or voice actors from the shows. The origins of cosplaying manga are described to be debated, but that the idea most likely to have started when Nov Takahashi, founder of Studio Hard, went to a sci fi convention in L.A. and wrote about the experience for his Japanese fans. Shortly after, in the 1980s, many North American anime fans started attending sci fi conventions dressed in manga cosplay, and eventually created conventions of their own. According to the article, “Certain anime and manga characters are more popular than others, which results in trends within cosplay. The popularity of anime and manga characters is most evident by the numerous observations of cosplayers dressed as the same character at a convention” (Winge, 2006). These conventions are a great way for people who share a similar interest in manga and anime to meet each other and discuss what they enjoy about their hobby and passion, which is why it has grown in popularity over the years.

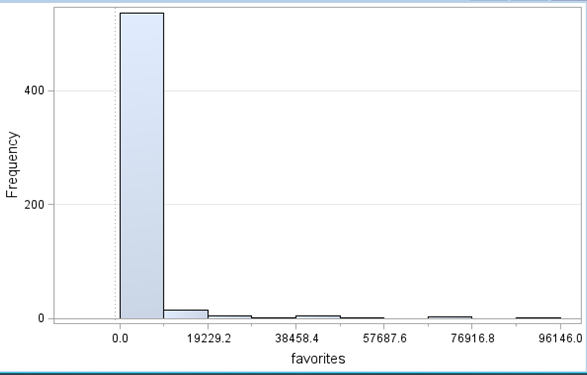
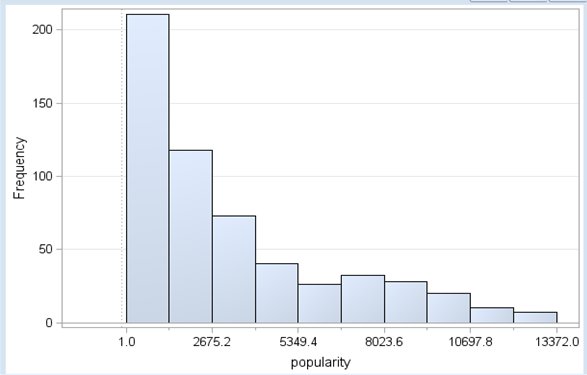
In conclusion, this literature review looks at the different reasons why anime has become popular in the United States, what types of people enjoy anime, and different ways these people meet and interact with one another. Anime has become popular in the United States because Americans enjoy the different types of storylines that they do not experience in their own media. These articles taught us that there is no certain criteria that people must meet to be anime fans, and that all different types of people enjoy the culture. Also, anime has a strong effect on on college students and society’s behavior. Lastly, many of the people who enjoy anime also enjoy attending conventions and cosplaying as their favorite characters, which is another subculture within the anime culture. Overall, anime is a form of media that many Americans enjoy watching and continues to gain popularity as the internet progresses and we are more able to share media resources with one another, therefore we should expect it to continue to grow in popularity in the future.

Data:

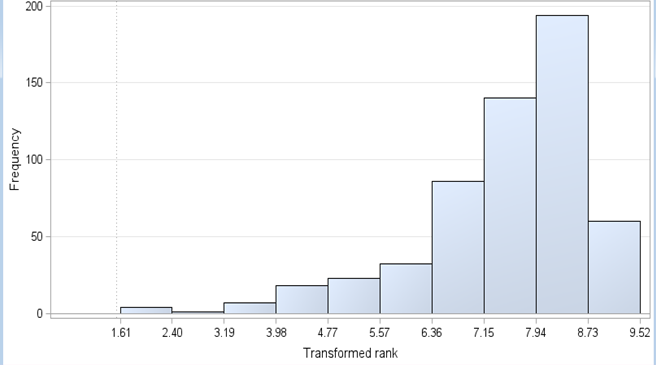
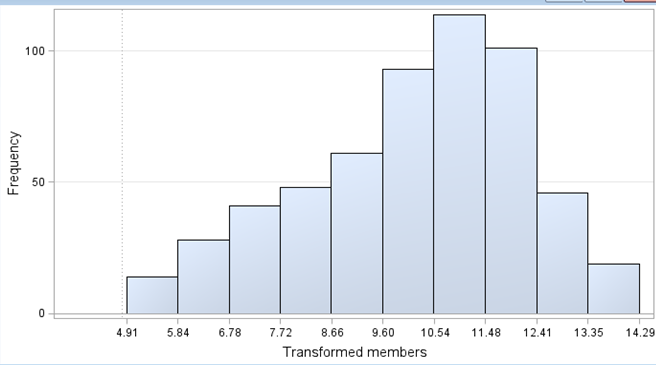
The dataset for this project contains 19,000 observations and 26 variables. Before modeling, missing data and distributions of the variables were checked. Since the starting dataset contained more than enough observations, the ones with any missing values under any of the indicators, are deleted. Along with that were handful of duplicates and useless variables. These variables were “Japanese Title”, “Title Synopsis”, ”Source”,”Status”,”Start\_date”, “End\_Date”, “Premiered”, and “Broadcast.” As a result, the dataset has 1129 observations and the variables are name, “Type”, “Source”, “Producers”, “Genre”, “Studio”, “Episodes”, “Rating”, “Score”, “Rank”, “Popularity”, and “Members.” Type represents whether the anime falls under TV, Movie, Manga or OVA. For this study, all the animes that are not TV type are filtered out. Producers and Studio represents who is responsible for animated the specific anime. Score represents that Anime’s given ImDB Score. Interent Movie Database is an online database of information related to films, television series, podcasts, home videos, and more. It is the most popular and authorative source for tv ratings and reviews. Rank rrepresents the anime’s placement upon a hierarchy of animes in MyAnimeList. The MyAnimeList website has every anime/manga in existence and has an account-based style where people can sign up and add shows to their list and debate amongst each other while rating anime. Favorites is also a predictor variable that is recorded through MyAnimeList which represents the number of members that add the anime to their favorites list. The target variable is set as members. This represents the sum of users currently watching a show, completed the show or plan on watching it. Now that the dataset is cleaned up, the distributions had to be checked for non-normality. From here on, the program being used to perform this analysis is SAS Enterprise Minor. The SAS Institute created the data mining software tool known as SAS Enterprise Miner. Predictive modeling, machine learning, and data analysis all use it. Users of SAS Enterprise Miner can explore and analyze data sets using a graphical user interface, build predictive models, and use those models to score new data. In SAS Enterprise Miner before cleaning the data, the Data Partition Node is used to split the data 50% Training and 50% Validation. Then the Stat Explore is used to bring up a histogram of every variable’s distribution. The variables that appear to have a skew distribution are episodes, members, popularity, rank, and score. As a result, the Transform Variables node from SAS Enterprise is used to transform members, popularity, and rank using log transformation. As for the variable score, the square transformation is used.

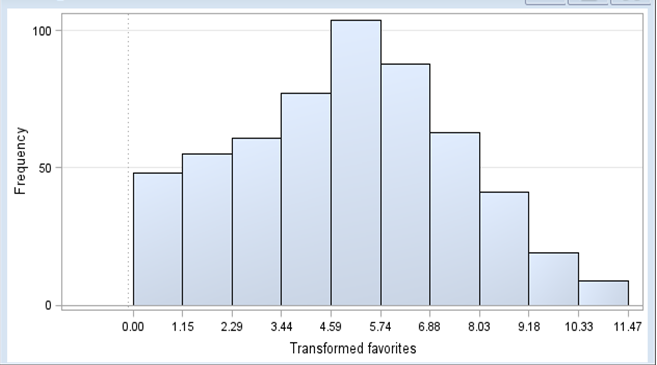
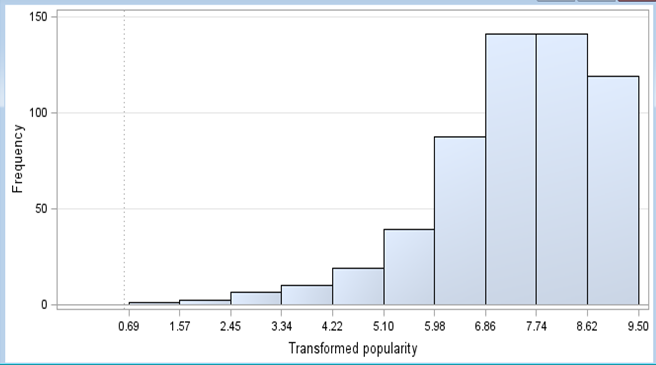
Before Transformations:





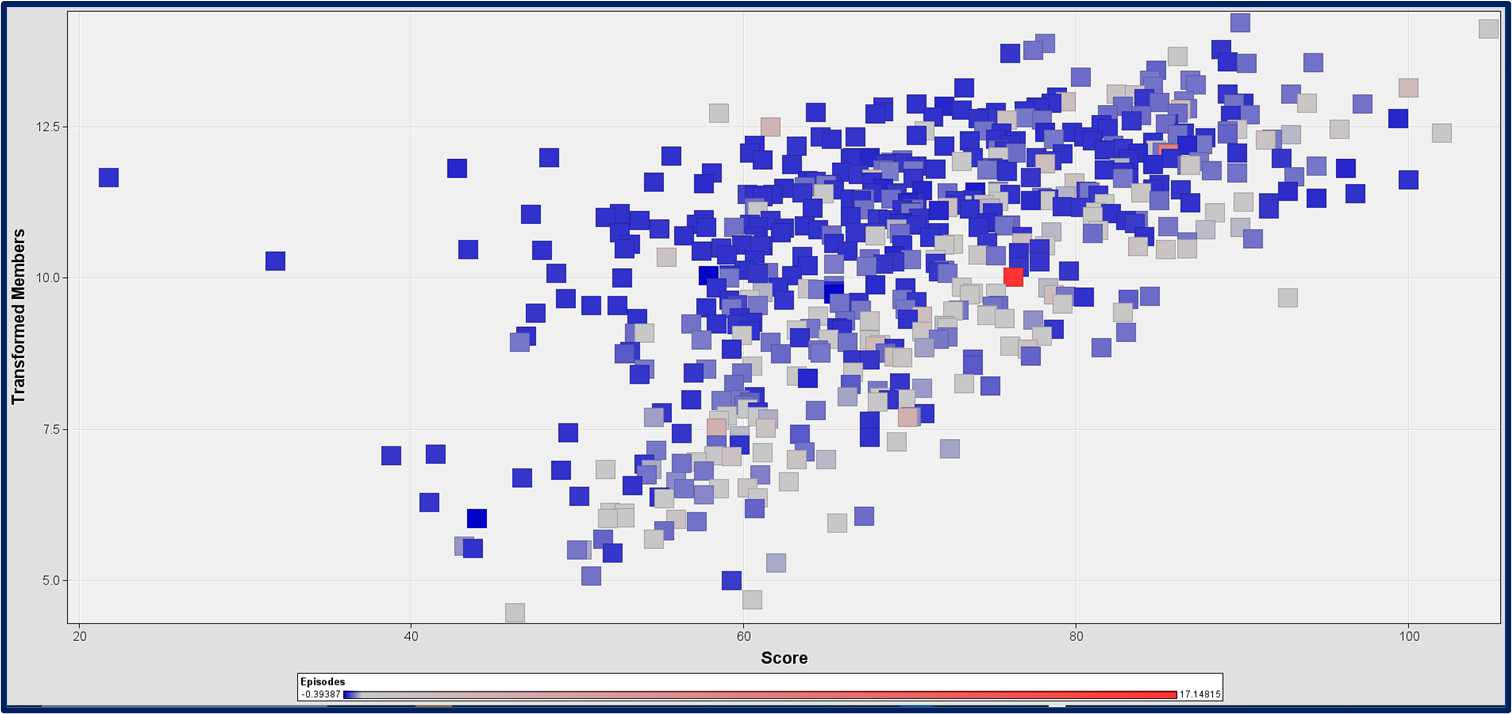
After Transformation:



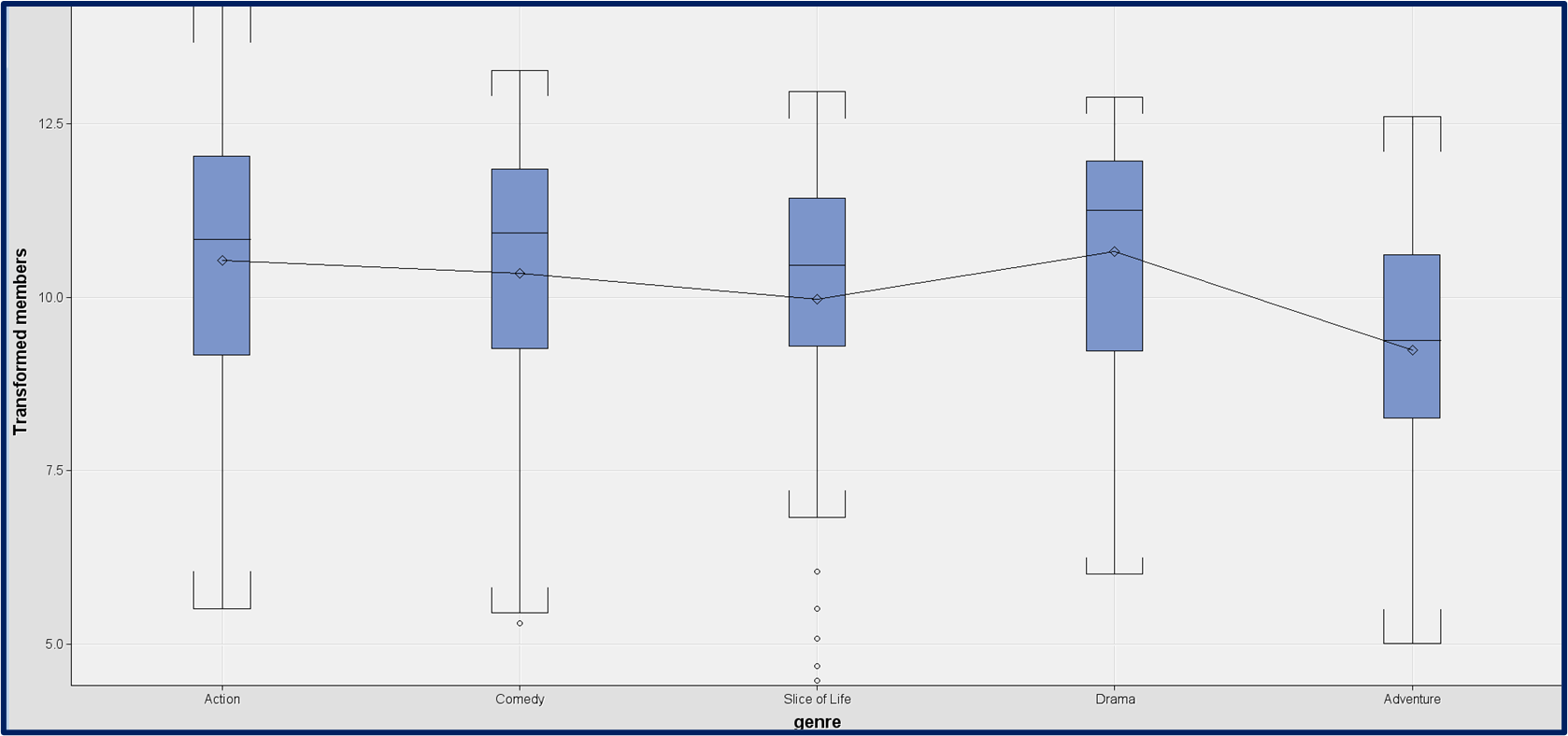


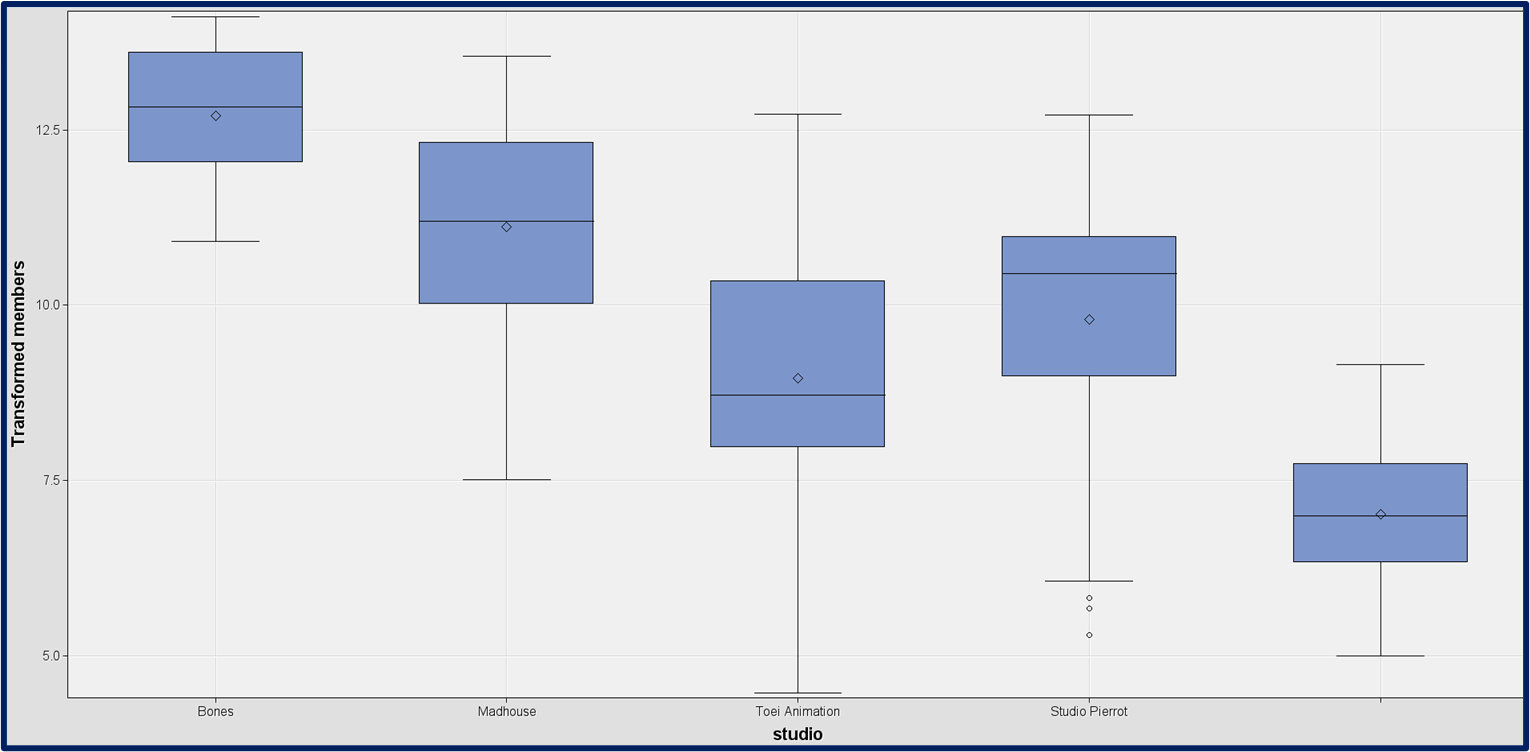
Exploratory analysis:

After running the decision tree model and regression model, a number of variables came up as impactful. Further exploration was completed on variables and their separate relationship with the target variable members.

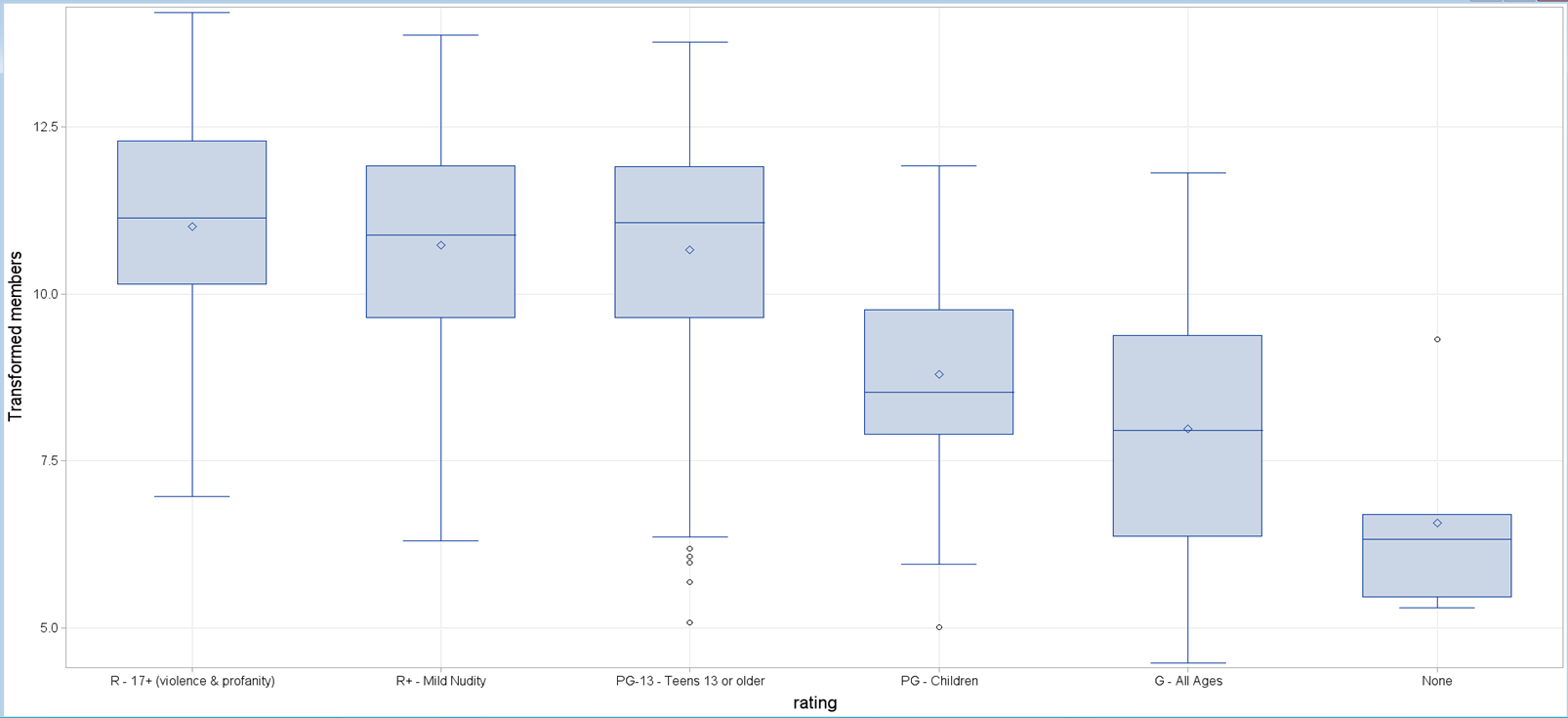


The first graph is a scatterplot showcasing IMDB score vs members. The observations were also color coded by # of episodes. The higher the IMDB score the higher the # of members. A majority of the lower rated shows have lesser # of episodes, so as the number of episodes increases, so does the number of members, for the most part.





The second and third graph showcase a boxplot of the top 5 most frequent genres and anime production studios. The genre, action was heavily favored by fans followed by comedy and drama. Iconic shows like Attack on Titan, Fullmetal Alchemist: Brotherhood, Gintama and My Hero Academia are all action shows. Among the dataset the clear winner in terms of production studio is Bones and Madhouse. After taking a look at the top 10 popular anime produced by the 4 widely known production studios, none had member counts that came close to Madhouse and Bones’ top projects.



The last graph showcases a boxplot of rating vs members. It is clear that the ratings R-17, PG and G avg more members than the rating PG 13 and none

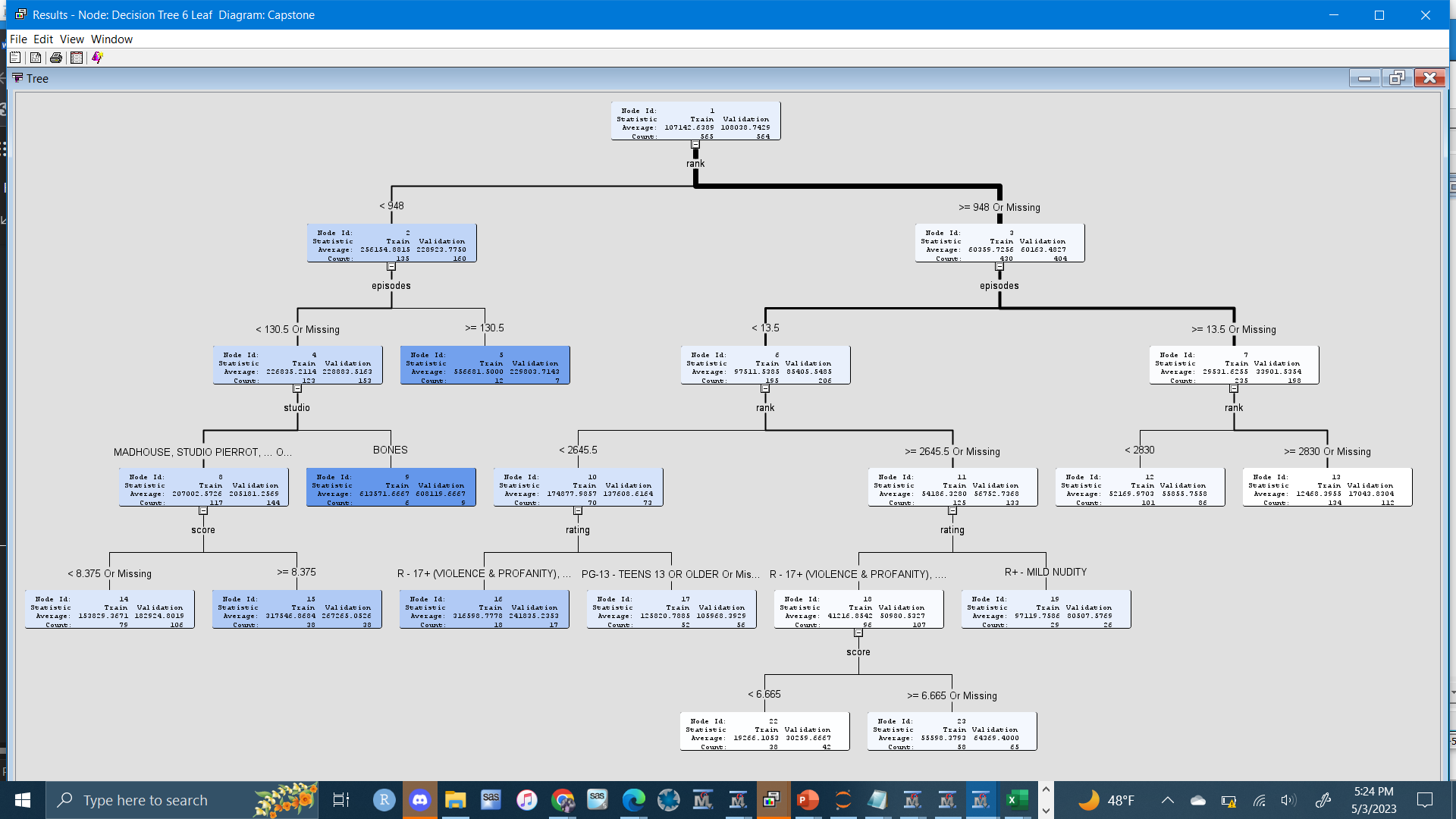
Models:

Decision tree is a flowchart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label. A decision tree will serve as a method to predict the # members a certain anime will have based on their features. During this analysis, a bit of trial and error was the plan. A total of 6 Random Forests and 3 Decision Trees were ran.

Model Comparison

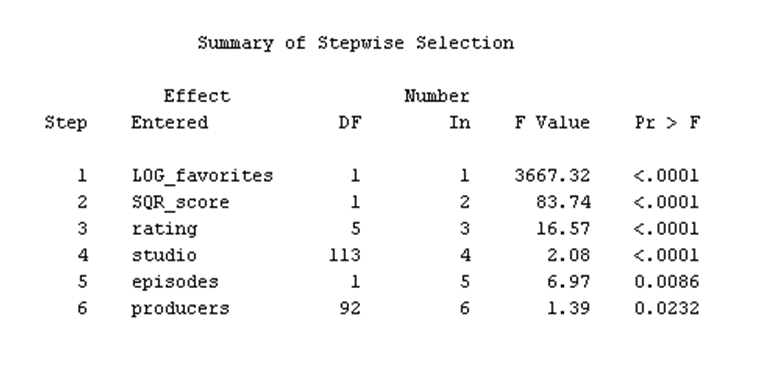
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Maximum Number of Trees | Maximum Depth | Signifcance Level | Selection Criterion: AVG Squared Error | Maxium Absolute Error |
| Forest 7 | 100 | 10 | 0.05 | 2.074E10 | 121959 |
| Forest 8 | 50 | 10 | 0.05 | 2.07E10 | 121526 |
| Forest 9 | 50 | 20 | 0.05 | 2.07E10 | 121526 |
| Forest 10 | 5 | 20 | 0.05 | 2.151E10 | 122587 |
| Forest 11 | 50 | 50 | 0.05 | 2.07E10 | 121527 |
| Forest 12 | 10 | 50 | 0.05 | 2.093E10 | 123345 |
| Model | Interval Target Criterion | Leaf Size | Signifance Level | Selection Criterion: AVG Squared Error | Maxium Absolute Error |
| Tree 3 | ProbF | 4 | 0.2 | 2.075E10 | 118341 |
| Tree 6 | ProbF | 5 | 0.2 | 2.075E10 | 118341 |
| **Tree 7** | **ProbF** | **6** | **0.2** | **2.068E10** | **118341** |

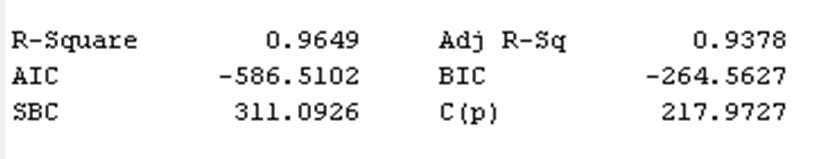
Decision Tree 7



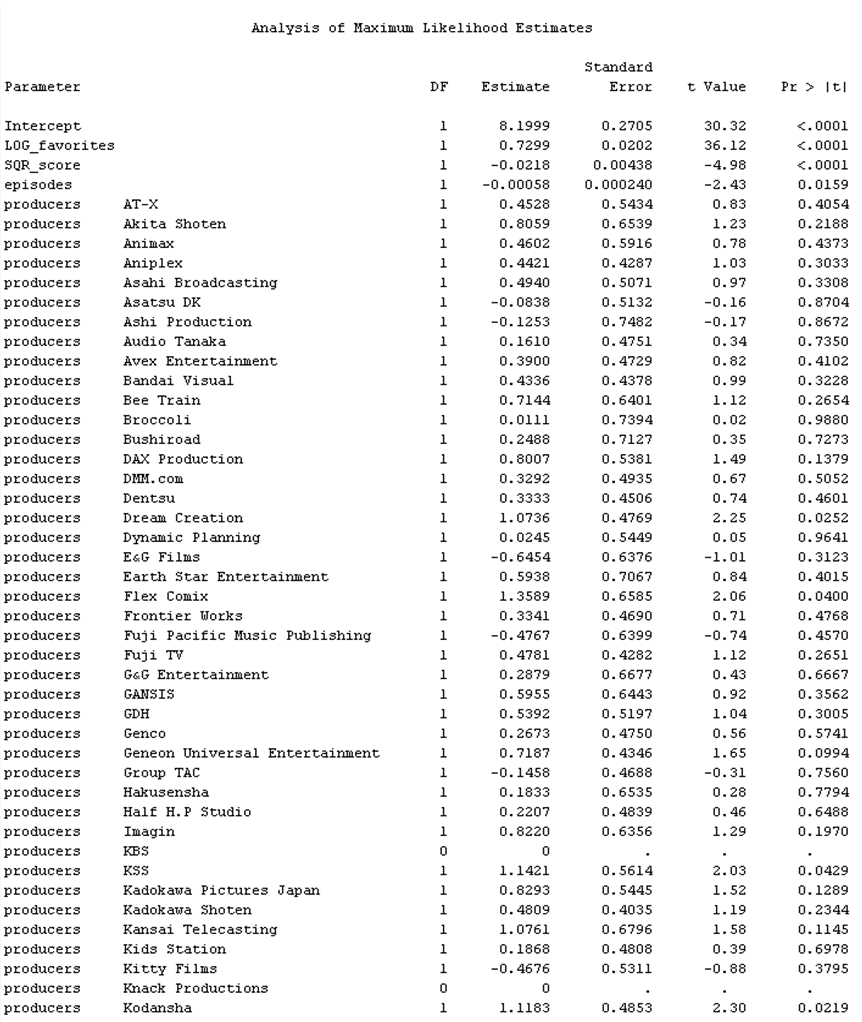
The model comparison went with the decision tree #7. The assessment measure used is decision, and is a 6 leaf tree. The nominal target criterion is ProbChisq. Variables that came up as most important are rank with 1.000 importance and 5 splitting rules, studio with .656 importance and 1 splitting rule, then episodes, score and rating. The validation MAE is 1183411 and the validation ASE is 2.075E10. From the tree its apparent that 130 episodes will positively affect your popularity since longevity seems most likely to correlate with popularity. However, if your anime is ranked lower having less episodes will lead to more members. As for ratings, R-17+ and R+ will lead to higher success as well.

A regression analysis is a set of statistical methods used for estimating the relationship between a dependent variable and one or more independent variables. It's basically a way of sorting out which variables have an impact and which ones to ignore and finding trends in data.





For this analysis the dependent variable is once again members. The regression type ran is linear and the selection model is stepwise. After the model was ran it received an adjusted R-Sq of 93.8%. This represents how accurate the chosen variables are in predicting the target’s variability, in this case members.



Conclusion:

In conclusion, there are many observations that can be taken away from this analysis. 15 out of 110 producers have a negative effect on popularity. ImDB Score is strongly correlated with popularity. Generally, animes with more epsiodes are less popular however more epsiodes will the boost the popularity of high ranked shows. Also, the genres’s action, harem, josei, military and mystery have a positive impact on popularity. The ratings R-17+(violence & profanity), PG and G have positive impacts on popularity whil PG-13 has a negative impact.

There are many changes that could be done to this study to further improve for the future. One potential change is incorporating a way of controlling the effect of sequels. There are many instances when a sequel does not live up to expectations but have a high number of members due to the anime’s first season. Another change is implementing a sentiment analysis on comments on Instagram, YouTube and Twitter tweets to see how people feel about certain animes that are trending now and get a better sense of the emotional tone. This analysis also contributes a great deal to society and the anime community. It can help people distinguish anime from each other and gain a better undertsnading of what elements of anime do people enjoy the most. It also proves very useful for anime production teams and up incoming manga artists when trying make an anime that will grab the anime community’s attention.

References:

*Study on Anime and Its Impacts among University Students - Researchgate*, www.researchgate.net/publication/319494624\_Study\_on\_Anime\_and\_Its\_Impacts\_Among\_University\_Students. Accessed 8 May 2023.

Napier, Susan J. "Why Anime?." Anime from Akira to Princess Mononoke: Experiencing Contemporary Japanese Animation. New York: Palgrave Macmillan US, 2001. 3-14.

Newitz, Annalee. "Anime otaku: Japanese animation fans outside Japan." Bad Subjects 13.11 (1994): 1-14.

Winge, Theresa. "Costuming the imagination: Origins of anime and manga cosplay." Mechademia 1.1 (2006): 65-76.