Analyzing the Popularity of Movie Genres in Context of Social and Economic Sentiments Within the United States Throughout the 21st Century

---

AP Research

Word Count: 4991

How does the social/economic situation of the American public from 2000 to 2020 relate to the popularity of various movie genres over the years within the domestic box office?

I. INTRODUCTION

The early 1900s marked the beginning of movies as a popular form of entertainment within the United States, and the industry has grown ever since. In 2019, the U.S. film entertainment industry alone produced an estimated revenue of $35.3 billion (Stoll, 2021). As movies became a part of mainstream entertainment, not just in the United States but on a global scale, movie genres have expanded to satisfy the diverse preferences of audiences. Although accurately quantifying such a subjective concept is difficult, there may be over 120 movie genres and subgenres (120+ Examples, 2020). Similar to literature and news, it is not unlikely that the popularity of movies may reflect the sentiments of audiences during a particular time, whether that be an era of economic recession, racial tension, or political turmoil.

This paper will explore the rise and fall in the popularity of various movie genres within the United States, how those trends are associated with indicators of social sentiments within the nation throughout the 21st century, and possible causes of any evident associations. The collection of overarching movie genres to be analyzed will be determined by IMDb, which describes itself as “the world's most popular and authoritative source for movie, TV and celebrity content”. This study aims to programmatically extract movie popularity by genre and, combined with official datasets of social sentiment indicators, identify any monotonic associations between particular genres and indicators. The results of this investigation are applicable in that they can be used to understand which movie genres achieve the most success within the domestic box office based on external factors in the nation.

II. LITERATURE REVIEW

A. Defining Genre

Over time, film industries have become advanced enough to categorize the thousands of produced movies into categories called genres. According to Merriam-Webster Dictionary, Genre is a borrowed French term defined as "a category of artistic, musical, or literary composition characterized by a particular style, form, or content". Popular overarching genres include Action, Comedy, Drama, Fantasy, Horror, Mystery, Romance, Thriller, and Western. However, there are at least a hundred more perceived genres and subcategories. As stated in the definition, films in each genre, although not identical, follow similar plot ideas, tropes, and artistic styles - hence, they are placed in the same category. According to a 2017 article by Grodal, genres are the product of "cultural constructions to adapt to the preferences of our embodied brain and its emotions and action potentials in relation to the world" (Grodal, 2017).

B. Psychology Behind Movie Appeal

It is no surprise that, as a society, we enjoy watching movies as a form of entertainment. But why do they draw us in and make us feel certain emotions - and how do these emotions add to their appeal? A 2018 paper titled "A psychology of the film" broke down our "absorption in film" into four specific aspects: narrative engagement, transportation, empathy, and flow (Tan, 2018). The qualities of these four measures create an "experience of intense attention", which we call absorption in a fictional world. In particular, the empathy and narrative engagement aspects can have lasting and beneficial effects on the viewer. There is a self-administered practice called "cinema therapy", where a patient chooses to watch movies with themes or issues that mirror their own life's problems. This can influence how they cope with the issue and make decisions

(Niemiec, 2020). Additionally, it has been proven that certain types of films can enhance positive characteristics and behaviors in viewers. Mainstream movies typically utilize a style of film technology and narrative themes that trigger strong emotional responses in the viewer and maximum concern for the characters. It also capitalizes on the viewer's sense of curiosity, as the narrative is designed to create a "characteristic systematic unfolding of interest".

C. Impact of Economic Conditions on Movie Genre Popularity

As important events unfold in the current world, the public sentiment often shifts to reflect the psychological effects of such external societal issues. One major external condition that impacts everyone in one way or another is the state of the economy. This being said, it is interesting to consider whether fluctuations in a country's economic condition result in significant changes in the popularity of certain movie genres. An article from the Journal of Cultural Economics analyzing economic and film data from 2000-2010 in three different European countries argues that the "demand for cinema entertainment as a whole is unrelated to economic indicators such as GDP, consumer confidence, and consumer prices". This study grouped genres into three umbrella categories: serious (drama, history, biography, documentary, and thriller genres), feel-good (animation, comedies, romances, and musicals), and kinetic (action, adventure, horror, sci-fi, and fantasy). A time-series analysis was conducted for the prevalence of these umbrella categories in each of the three countries, with markers for historical events. The authors of the study stated in the result that there was no significant connection between the economy and that "fluctuations in individual movie quality and the focus on film as art" superpose potential effects of the economic context on the aggregate demand and supply, respectively (von Rimscha, 2013).

However, a 2017 thesis titled What Movies Do You Watch During Bad & Good Economic Times: The Impact of Movie Genre on Attendance over Business Cycles from the University of Guelph concluded otherwise. In this study, the economic condition is measured using CSI (Consumer Sentiment Index) and the movie performance is measured as the weekly attendance. Using statistical models, the author determined that the drama genre has “the significant main effect and interaction effect with CSI [compared] with other genres for most of the models”. More specifically, consumers tend to prefer drama movies more (relative to other genres) during times of economic recession (Yi, 2017).

D. The Gap In Knowledge

Although one or two similar studies have been conducted (as mentioned in the previous subsection) to analyze the relationship between movie genre preferences and economic indicators, the overall current-day research on this specific topic is very minimal, and the results of the two studies contradict each other. This paper attempts to resolve these conflicting outcomes, and adds to the academic conversation through the use of different time frames, time scales, genre lists, economic sentiment indicators, and analysis techniques than the other studies. Additionally, it observes movie popularity in the context of racial relations within the United States - something that the other studies do not do.

III. METHOD DESIGN

This study uses content analysis to gather two different sets of data: the popularity of different movie genres over time within the United States, and the social/economic atmosphere within the United States along the corresponding periods (in this case, throughout the years 2000 to 2020). The content analysis method was chosen because the necessary data is provided by indices that are available online and professionally compiled by trusted organizations.

The first step was to find an online source that provides the movie data and satisfies the following conditions: has at least an annual breakdown, contains an indicator of the relative popularity of different movies within the United States only, and provides the genre categories of said movies. The site “Box Office Mojo” tracks domestic (US) box office revenue and offers data on a monthly and yearly basis, ranking the movies by gross box office earnings. Additionally, each movie contains a genre attribute, satisfying all the conditions for a valid source.

The next step is to look for indicators of social/political/economic atmosphere within the United States, which is more difficult since this is trying to quantify a qualitative measurement - especially on the social and political side. Because I was unable to find an all-encompassing indicator, I tried to find a couple of different indicators that could be used to evaluate social and political tension, as well as the economic situation. The sources would need to meet the following conditions: must contain stand-alone data for the U.S. (cannot be relative to other countries), provide yearly data from 2000 to 2020, and must represent some aspect of the average U.S. citizen’s views on the social/political/economic atmosphere. The indicators that were decided upon include “Race Relations” by Gallup and U.S. Consumer Confidence Index by the Conference Board. The “Race Relations” data by Gallup was gathered from a poll to quantify American public satisfaction regarding the state of racial relations then within the US. The U.S. Consumer Confidence Index (CCI) is a monthly survey to determine the “current degree of optimism among American consumers”. Compared to the U.S. Consumer Sentiment Index, a similar indicator for American consumer optimism, which surveys only 500 households, the U.S.

CCI surveys 5,000 households. It is said that the U.S. CCI does a better job of representing “job market and job security” (Morah, 2019).

After finalizing the data to be used, it was evident that the gathering of the movie data, in particular, needed to be done programmatically. While the indicator data could be easily downloaded to an Excel spreadsheet, the BoxOfficeMojo contents could not. It would be too inefficient to process twenty years’ worth of monthly U.S. box office data by hand.

IV. METHOD PROCEDURE

The process of gathering the movie data was automated with a short Python program I wrote to *web-scrape* (a term for extracting data from a website) the Box Office Mojo website. All the data was gathered, formatted, and stored in a CSV (comma-separated values) file.

Because the U.S. Consumer Confidence Index is a monthly time series, and BoxOfficeMojo offers a monthly breakdown, I decided to start by gathering the data on a monthly scale instead of yearly - as I originally intended. However, if needed, I can always go up to yearly. I needed to come up with a way to determine and store the relative popularity of different genres as a time series (Note: the genre categories were determined by IMDb), so I implemented the following processes in my program:

A. Gathering the Data

Start: Loop through the range [2000,2020] to get each year

- ➔ For each year, loop through each month
  - ◆ For each month, parse the HTML data of the Box Office Mojo list page for that corresponding year/month (ex: 2003/June) and loop through each movie in the list

- For each movie, store the ranking of the movie in a variable (ranked using gross box office earnings) and follow the hyperlink to the page with more detailed movie-specific information
  - Extract the genre categories listed that the movie falls under

B. Processing the Data

It is unnecessary to store movie-specific information such as the titles and genres since all I need is the popularity of each overall genre category for each month, so I give each genre category a "popularity score" using the following method:

1. Each movie has a rank (ex: 7th highest gross earnings out of 94 movies that month) and set of genres it falls under (ex: Action, Fantasy, and Comedy)
2. Calculate the percentile ranking of that movie with this formula: 1 - (rank/total)
3. If the movie falls under the genre categories of Action and Fantasy, then the calculated percentile ranking for that specific movie will be added to the scores set of both Action and Fantasy.
4. When the program is done looping through all the movies for the month, it will average all the values pushed into the scores list for each genre category, resulting in a single value that represents the following statement: *On average, for the month of ___, movies that fell under the ___ genre had higher U.S. gross box office earnings than ___% of other movies during the same month*. This statement is assumed to indicate the popularity of each genre category.

Figure 1: *The particular code snippet that calculates the popularity scores for each genre.*

```
    # GO THROUGH THE LIST OF GENRES THAT THIS PARTICULAR MOVIE FALLS UNDER
    for genre in genres:
        # CALCULATE THE PERCENTILE RANKING OF THE MOVIE FOR THIS MONTH/YEAR
        percentile_score = round(1-(rank/size),5)
        # PUSH THAT MOVIE PERCENTILE VALUE INTO EACH GENRE THE MOVIE FALLS UNDER
        temp_ranks[genre].append(percentile_score)

# BASICALLY, AVERAGE ALL THE PERCENTILE VALUES APPENDED UNDER EACH GENRE - IF SOME GENRE HAS NONE, JUST USE 0
temp_ranks = dict((key,round(np.mean(vals),5)) if len(vals) != 0 else (key,0) for key,vals in temp_ranks.items())
```

Figure 2: *A CSV file snippet showing the result of those calculations*:*

Year	Month	Action	Adventure	Animation	Biography	Comedy	…
2000	January	0.4265	0.53799	0.52979	0.48252	0.54099	…
2000	February	0.36067	0.57337	0.57609	0.54447	0.5123	…
2000	March	0.51782	0.68218	0.64158	0.50135	0.55504	…
2000	April	0.48438	0.55332	0.68077	0.39511	0.53526	…
2000	May	0.72959	0.65578	0.59592	0.36508	0.48792	…
…	…	…	…	…	…	…	…

****Full list of genres:*** *Action, Adventure, Animation, Biography, Comedy, Crime, Documentary, Drama Family Fantasy, History, Horror, Musical, Music, Mystery, News, Romance, Sci-Fi, Short, Sport, Thriller, War, Western.*

The indicator data did not require too much processing except for a few small modifications. The race relations dataset has the following format:

Figure 3: *Race Relations survey data as listed on the Gallup Historical Trends site.*

	Very satisfied (%)	Somewhat satisfied (%)	Somewhat dissatisfied (%)	Very dissatisfied (%)	No opinion (%)
2021	7	16	25	46	6
2020	7	29	33	25	6
…	…	…	…	…	…

I needed to combine some values to produce a single number that would indicate either the level of satisfaction or level of dissatisfaction of American adults with race relations. So, for each year, I combined the “Somewhat dissatisfied” and “Very dissatisfied” values to create a single “Dissatisfaction Percentage”. Ex: 2021 → 71% of American adults are dissatisfied with race relations, 2020 → 58% of American adults are dissatisfied with race relations.

For the U.S. Consumer Confidence Index, because the data is given on a monthly scale, to perform the statistical procedure (described in the Findings section) on the yearly movie data, I needed to make a yearly version of the U.S. Consumer Confidence Index dataset, which I did by simply averaging the monthly CCI for each year.

C. Materials

The tools used include the Visual Studio Code IDE, the Python 3.9 programming language, the BeautifulSoup library for web-scraping, and the Pandas library and NumPy library for data organization/manipulation. The statistical analysis described in the following section utilizes Jupyter Notebook, the Pandas library’s Spearman correlation method, and the Matplotlib library for data visualization.

V. FINDINGS

A. Explanation of Analysis Technique

The goal is to search for any relationship between U.S. CCI (Consumer Confidence Index) and the popularity of different movie genres, and any relationship between Race Relations Dissatisfaction and the popularity of different movie genres. Additionally, for more detailed findings, I have done an analysis for U.S. CCI vs genre popularity on both a monthly and yearly scale (the race relations dataset only has yearly data, so monthly analysis was not performed for that one).

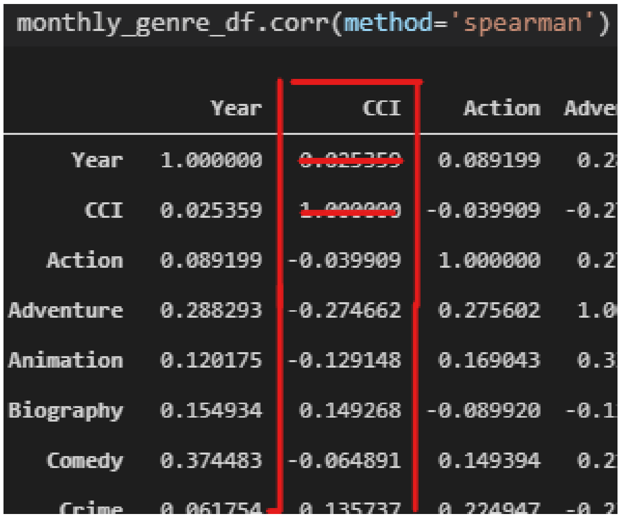
The chosen correlation method for this analysis is the Spearman Rank-Order Correlation method, rather than the most commonly used Pearson's Correlation method. This is because Pearson's Correlation coefficient calculates the strength and direction of a *linear* relationship between two datasets, while the Spearman Rank-Order Correlation coefficient calculates the strength and direction of a *monotonic* (strictly increasing/decreasing but not necessarily linear) relationship between two datasets. For the datasets in this project, it is unlikely that the x and y values will have a fairly linear relationship (if any relationship at all), so to offer more flexibility in terms of determining a relationship, a Spearman Rank-Order Correlation coefficient is used.

Additionally, with each correlation coefficient comes a P-value, which is used to determine the statistical significance of the correlation coefficient. A commonly used significance level is $\alpha = 0.05$, meaning that the correlation coefficient is statistically significant if the P-value is below 0.05 - otherwise, we *cannot* reject the null hypothesis that the correlation coefficient is 0.

B. Outcomes

After running the Spearman Correlation method on various combinations of datasets, this was the outcome*:

**The brackets indicate the [scale of the data] and the [significance level] used. The genres listed are the ones whose correlation coefficients qualify as statistically significant according to the significance level given.*



```
monthly_genre_df.corr(method='spearman')
```

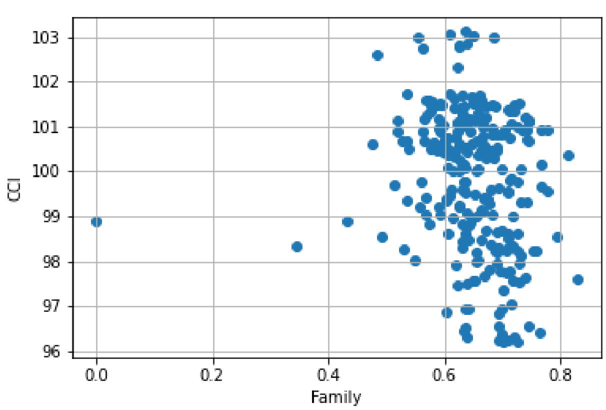
	Year	CCI	Action	Adve
Year	1.000000	~~0.025359~~	0.089199	0.2
CCI	0.025359	~~1.000000~~	-0.039909	-0.2
Action	0.089199	-0.039909	1.000000	0.2
Adventure	0.288293	-0.274662	0.275602	1.0
Animation	0.120175	-0.129148	0.169043	0.3
Biography	0.154934	0.149268	-0.089920	-0.1
Comedy	0.374483	-0.064891	0.149394	0.2
Crime	0.061754	0.135737	0.224947	-0.2

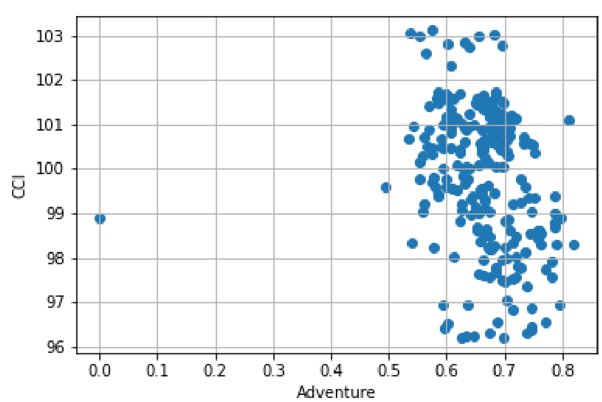
Figure 4: *Snippet of the code in Jupyter Notebook that calculates the Spearman correlation coefficients for each genre on a monthly scale. The values in the CCI column are the same correlation coefficients shown in the table immediately below.*

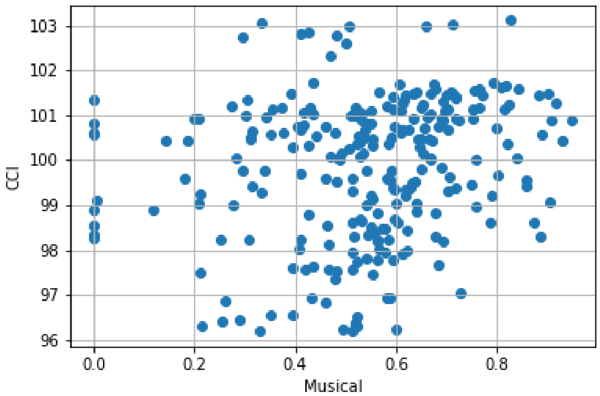
U.S. CCI vs. Genre Popularity [Monthly] [α = 0.01 → Highly Statistically Significant]

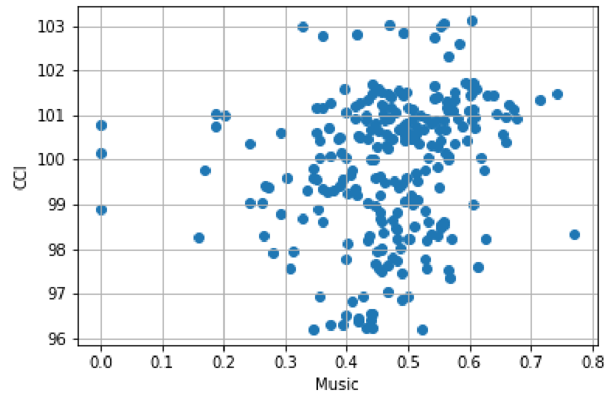
Genre	Correlation Coefficient	P-Value	Correlation Type
Music	0.280642	6.06E-06	Weak, Positive
Musical	0.241924	0.000105	Weak, Positive
Romance	-0.16654	0.00807	Very Weak, Negative
Family	-0.26444	2.11E-05	Weak, Negative
Adventure	-0.27466	9.70E-06	Weak, Negative
News	-0.34877	1.28E-08	Weak, Negative
Short	-0.3881	1.75E-10	Weak, Negative

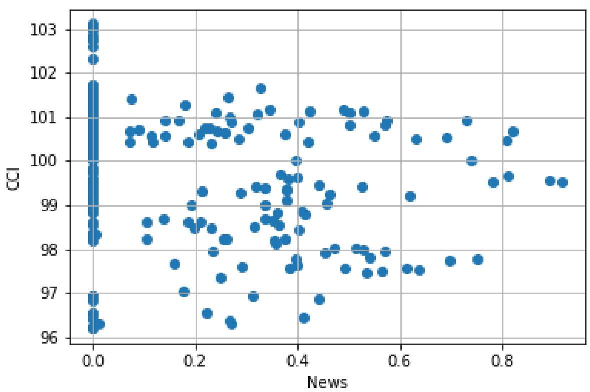
Scatterplots of the above genres:

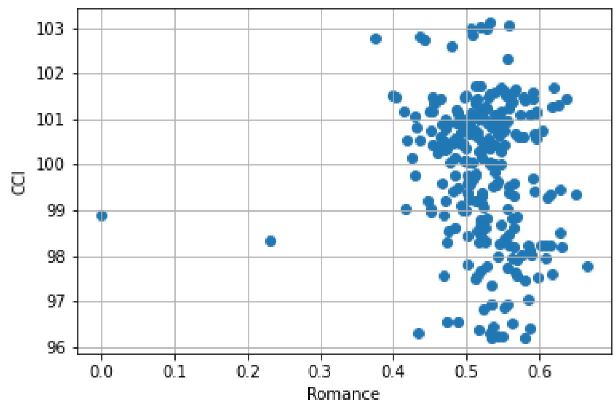


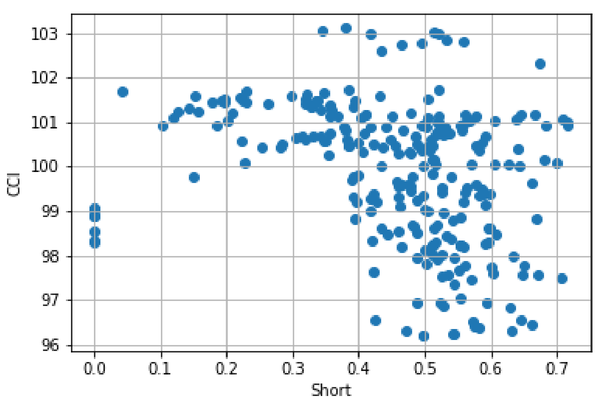












A Spearman's correlation was run to determine the relationship between 252 monthly U.S. CCI and monthly movie genre popularity values.*

- There was a **weak, positive** monotonic correlation between monthly **Music** movie genre popularity and monthly U.S. CCI ($r_s = .281$, $n = 252$, $p < .01$).

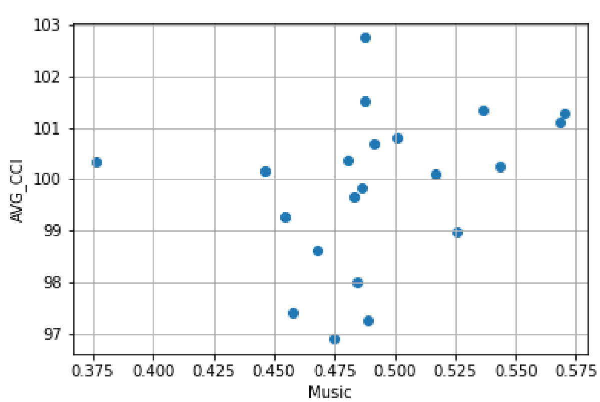
- There was a **weak, positive** monotonic correlation between monthly **Musical** movie genre popularity and monthly U.S. CCI ($r_s = 0.242$, $n = 252$, $p < .01$).
- There was a **weak, negative** monotonic correlation between monthly **Adventure** movie genre popularity and monthly U.S. CCI ($r_s = -0.275$, $n = 252$, $p < .01$).
- There was a **weak, negative** monotonic correlation between monthly **Family** movie genre popularity and monthly U.S. CCI ($r_s = 0.242$, $n = 252$, $p < .01$).
- There was a **weak, negative** monotonic correlation between **News** movie genre popularity and U.S. CCI ($r_s = -0.349$, $n = 252$, $p < .01$).
- There was a **weak, negative** monotonic correlation between **Short** movie genre popularity and U.S. CCI ($r_s = -0.388$, $n = 252$, $p < .01$).

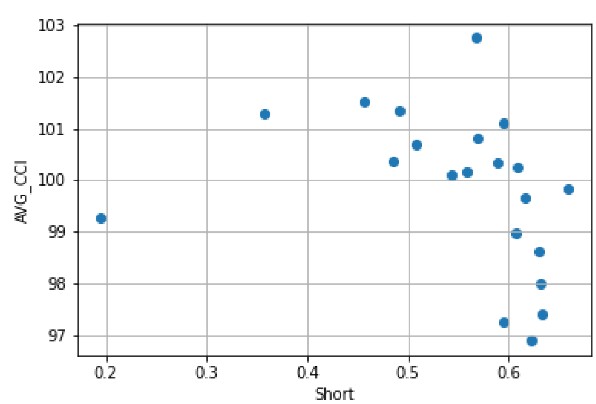
**The Romance genre was left out from the list above because the correlation is considered very weak.*

U.S. CCI vs. Genre Popularity [Yearly] [α = 0.05 → Statistically Significant]

Genre	Correlation Coefficient	P-Value	Correlation Type
Music	0.446753247	0.04232392	Moderate, Positive
Short	-0.64025974	0.001769793	Strong, Negative

Scatterplots of the above genres:





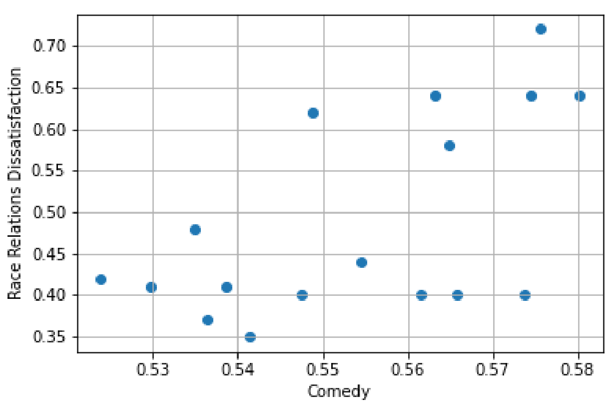
A Spearman's correlation was run to determine the relationship between 21 yearly U.S. CCI and yearly movie genre popularity values.

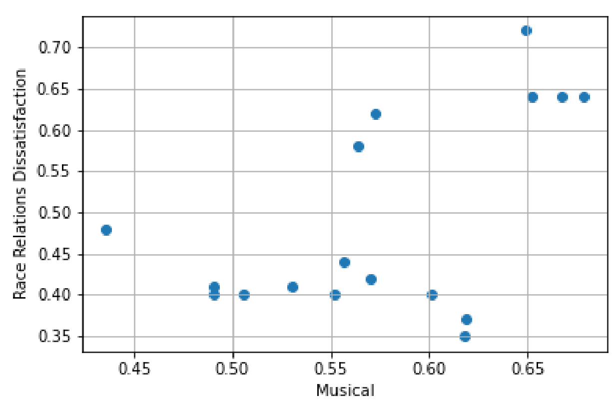
- There was a **moderate, positive** monotonic correlation between yearly **Music** movie genre popularity and U.S. CCI ($r_s = .447$, $n = 21$, $p < .05$).
- There was a **strong, negative** monotonic correlation between yearly **Short** movie genre popularity and yearly U.S. CCI ($r_s = .640$, $n = 21$, $p < .05$).

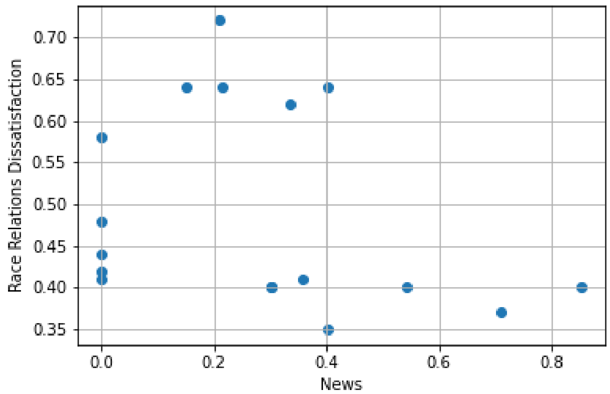
Race Relations Dissatisfaction vs. Genre Popularity [Yearly] [α = 0.10 → Weak evidence for rejecting null hypothesis]

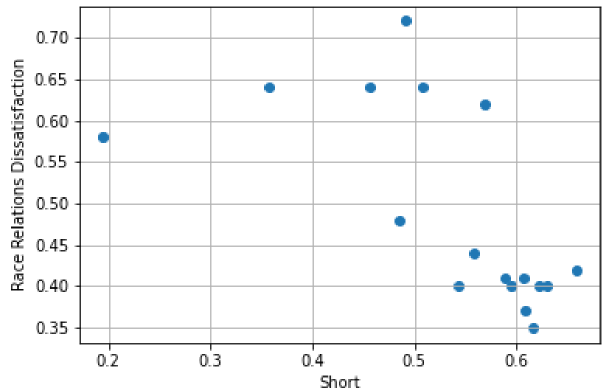
Genre	Correlation Coefficient	P-Value	Correlation Type
Comedy	0.45642	0.065536	Moderate, Positive
Musical	0.421787	0.091727	Moderate, Positive
News	-0.45586	0.065912	Moderate, Negative
Short	-0.73101	0.000857	Strong, Negative

Scatterplots of the above genres:









A Spearman's correlation was run to determine the relationship between 17 yearly movie genre popularity and yearly Race Relations Dissatisfaction values:

- There was a **moderate, positive** monotonic correlation between yearly **Comedy** movie genre popularity and yearly Race Relations Dissatisfaction (r_s = .456, n = 17, p < .10).
- There was a **moderate, positive** monotonic correlation between yearly **Musical** movie genre popularity and yearly Race Relations Dissatisfaction (r_s = .422, n = 17, p < .10).
- There was a **moderate, negative** monotonic correlation between yearly **News** movie genre popularity and yearly Race Relations Dissatisfaction (r_s = -0.455, n = 17, p < .10).
- There was a **strong, negative** monotonic correlation between yearly **Short** movie genre popularity and yearly Race Relations Dissatisfaction (r_s = -0.731, n = 17, p < .10).

VI. ANALYSIS

A. Interpret Findings

For simplicity and greater depth of analysis, I chose to look at a maximum of four genres from each table: the two genres with the strongest *positive* correlation coefficients and the two genres with the strongest *negative* correlation coefficients. The other genres will not be further analyzed throughout the rest of this paper.

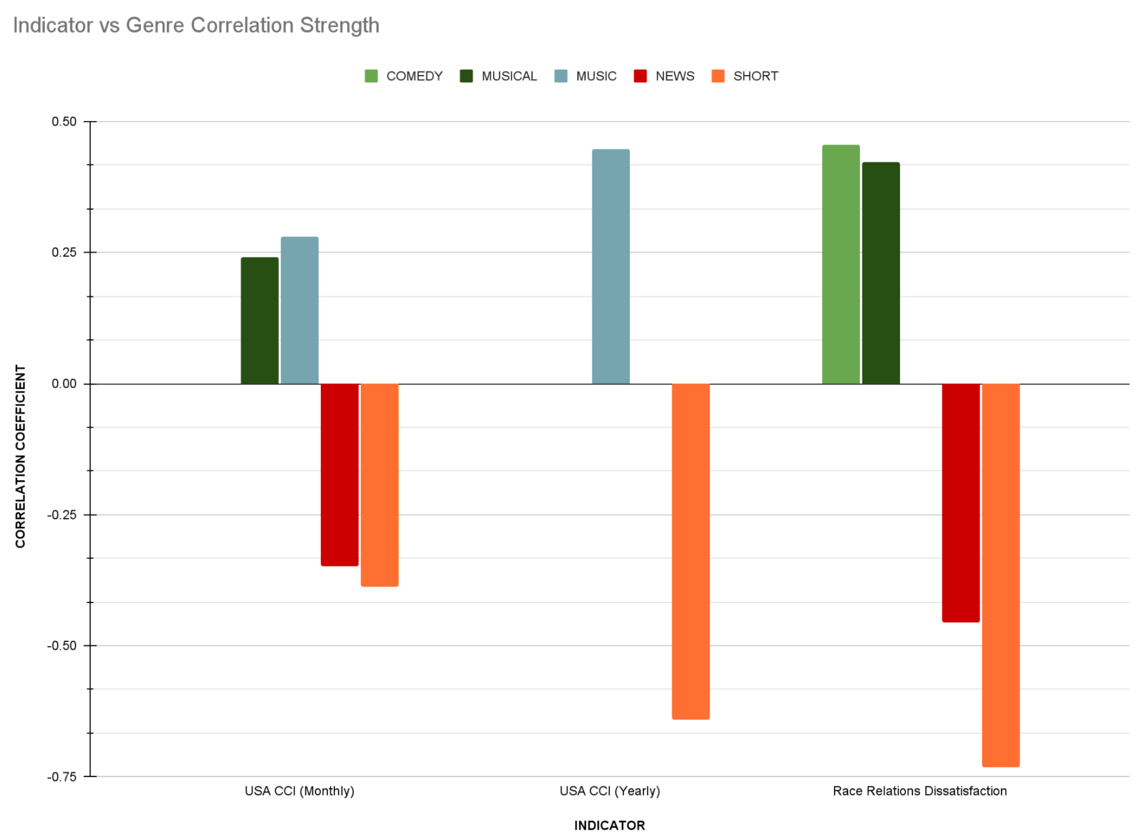
The genres that had the strongest monotonic correlations with the U.S. CCI on both a monthly and yearly scale were Music, Musicals, News, and Short. Let us recall that the U.S. Consumer Confidence Index is a measurement for “how optimistic people are about the economy and their ability to find jobs” - the higher the number, the more optimistic (Amadeo, 2021). The Music and Musical genres had positive monotonic correlations while the News and Short genres had negative monotonic correlations.

Similarly, let us recall that the Race Relations Dissatisfaction values indicate the proportion of American adults who are dissatisfied with the state of racial relations within the U.S. - the higher the number, the higher the over level of dissatisfaction. The genres that had the strongest monotonic correlations with the Race Relations Dissatisfaction values are Comedy, Musical, News, and Short. The Comedy and Musical genres had positive monotonic correlations while the News and Short genres had negative monotonic correlations.

Therefore, the answer to the research question is the following: As the U.S. Consumer Confidence Index increased, the popularity of Music and Musical genre films increased, and the popularity of News and Short genre films decreased. As Race Relations Dissatisfaction increased, the popularity of Comedy and Musical genre films increased, and the popularity of News and Short genre films decreased.

These findings differ from the two studies mentioned in the literature review: one of which found no relationship between movie genre and the economy (von Rimscha, 2013), and the other of which found that the popularity of drama movies increased during times of economic recession (Yi, 2017).

Figure 5: *A bar graph to visualize correlation strength and direction between the datasets.*



B. Potential Explanation for Findings

Before digging further into potential reasons for these findings, it is worth mentioning that the genres which demonstrated monotonic correlation with the U.S. CCI values were almost identical to the genres which demonstrated monotonic correlation with the Race Relations Dissatisfaction values (3 matches: Musical, News, Short). I chose to take a deeper look at only the U.S. CCI genre correlations because the data for that indicator is more comprehensive than that of the Race Relations Dissatisfaction indicator (further explained in the limitations section).

The IMDb site defines the Music genre as "[containing] significant music-related elements while not actually being a Musical; this may mean a concert, or a story about a band (either fictional or documentary)." The Musical genre is defined as "should contain several scenes of characters bursting into song aimed at the viewer (this excludes songs performed for the enjoyment of other characters that may be viewing) while the rest of the time, usually but not exclusively, portraying a narrative that alludes to another Genre" (IMDb Genre, 2021). Referring back to the literature review, the paper published by the Journal of Cultural Economics categorized Musicals under the "feel-good" category (von Rimscha, 2013). It was difficult to find any academic discussion surrounding the Music genre, although there was a heavy emphasis on the use of music in film. However, one could reasonably presume that Music genre movies are popular for similar "feel-good" characteristics. With this information, I could make the possible unconfirmed hypothesis that during the periods where consumers are more optimistic about the economic situation, audiences may be more inclined to enjoy spectacular and bright films such as those within the Musical and Music genres.

On the other hand, we have the News and Short genres which are negatively correlated with the U.S. CCI. The IMDb site defines the News genre as "reports and discussion of current events of public importance or interest....This generally includes newsreels, newsmagazines, daily news reports, and commentary/discussion programs that focus on news events". The Short genre is defined as "any theatrical film or made-for-video title with a running time of less than 45 minutes, i.e., 44 minutes or less, or any TV series or TV movie with a running time of less than 22 minutes, i.e. 21 minutes or less." My theory is the following: periods during which the economy is not doing well may also be characterized by rising social/political tensions, which in turn paves the way for both satirical and informative news-based films. I will refrain from forming a conjecture regarding the Short genre because although it is considered a genre according to IMDb, there are no defining qualities of Short film plots except that they are under 45 minutes long. Of course, it is important to note that everything stated in this section is simply hypothetical and not to be assumed true without further research on the matter. After all, correlation does not imply causation.

C. Noting the Effect of the COVID-19 Pandemic on the Domestic Box Office

As social distancing restrictions were widely implemented during the start of the COVID-19 pandemic in March 2020, the effect is evident on the domestic BoxOfficeMojo charts, with monthly releases dipping from over 100 to just under 10 and the gross revenue dropping significantly as well. The images below compare the domestic box office for April 2020 to that of April 2019.

Figure 6: *Domestic Box Office for April 2020 (6 Releases):*

By Month | April | 2020 | Calendar grosses

Rank	Release	Gross	Theaters	Total Gross	Release Date
1	Phoenix, Oregon Re-release	$16,846	33	$23,613	Mar 20
2	True History of the Kelly Gang	$15,030	5	$33,817	Apr 24
3	Resistance	$7,464	1	$7,464	Mar 27
4	Swallow	$7,050	39	$33,419	Mar 6
5	The Other Lamb	$4,825	1	$6,024	Apr 3
6	Strike	$800	2	$15,758	Mar 27

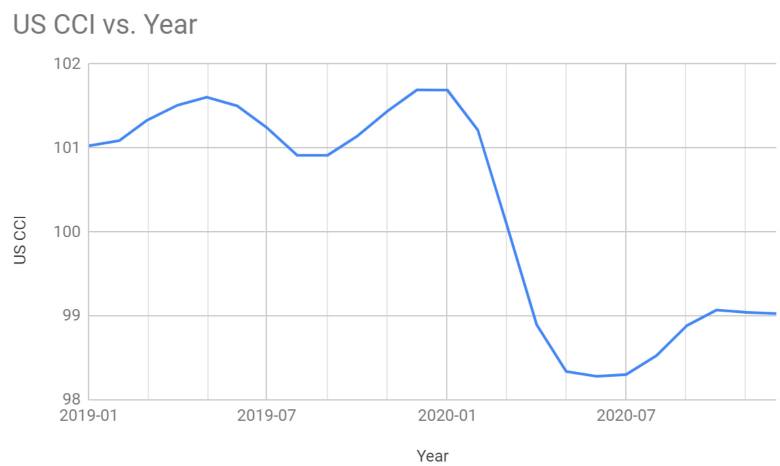
Figure 7: *Domestic Box Office for April 2019 (220 Releases):*

By Month | April | 2019 | Calendar grosses

Rank	Release	Gross	Theaters	Total Gross	Release Date
1	Avengers: Endgame	$427,099,795	4,662	$858,373,000	Apr 26
2	Shazam!	$132,092,018	4,306	$140,371,656	Apr 5
3	Dumbo	$61,843,845	4,259	$114,766,307	Mar 29
4	Captain Marvel	$61,398,320	4,310	$426,829,839	Mar 8
5	Pet Sematary	$52,972,548	3,585	$54,724,696	Apr 5
6	Us	$45,316,415	3,743	$175,084,580	Mar 22
7	The Curse of la Llorona	$43,457,094	3,372	$54,733,739	Apr 19
8	Little	$36,614,690	2,667	$40,860,481	Apr 12
9	Breakthrough	$28,192,947	2,913	$40,713,082	Apr 17
10	Hellboy	$21,643,107	3,303	$21,903,748	Apr 12
11	Missing Link	$15,702,484	3,437	$16,649,539	Apr 12
12	After	$11,748,994	2,138	$12,138,565	Apr 12
13	Unplanned	$11,420,786	1,516	$19,005,109	Mar 29
14	The Best of Enemies	$10,046,986	1,705	$10,205,616	Apr 5
15	Five Feet Apart	$9,600,631	2,866	$45,729,221	Mar 15
16	Wonder Park	$7,087,426	3,838	$45,216,793	Mar 15
17	How to Train Your Dragon: The Hidden World	$6,622,225	4,286	$160,799,505	Feb 22
18	Penguins	$6,105,864	1,815	$7,699,452	Apr 17

The U.S. Consumer Confidence Index also dipped dramatically during the first half of 2020 (see graph below). Because of the way that the popularity scores for the genres were calculated (average of movie percentile scores), the fact that there were so few total releases in April (and during the following months) likely had a large influence on the percentile scores for movie genres released during that time. Similarly, it is possible that generally popular genres suddenly dipped to a popularity score of zero because they did not fall within the 6 releases. These magnified genre popularity scores combined with the steep drop in U.S. CCI likely amplified the correlation (positive or negative) between certain genres and U.S. CCI.

Figure 8: *Graph showing a significant drop in U.S. CCI at the start of 2020.*



VII. LIMITATIONS

One assumption made was that the gross attribute of the movies in the Box Office Mojo chart is an indicator of movie popularity within the US. However, movie gross consists only of box office revenue (ticket sales at movie theatres) and does not include streaming services. With the rise of streaming services over the years, the physical ticket sales at movie theatres may not give an accurate picture of which movies are the most popular. The *total gross* of movies does include streaming services and other sources of revenue outside box office sales, but it also includes the foreign market, which would not suffice for this study since we are looking at only the United States.

As described in the methods section, I calculated a popularity score for each genre by averaging the percentile ranking of movies that fell under the said genre category. I believe this is an accurate representation of each genre's popularity for that given year/month, but there may be some inaccuracies in that process, which I missed.

Additionally, the Race Relations dataset lacked data from 2008 to 2012, bringing the number of data points down from 21 to 17. In the context of this study, 21 data points are already not much, so after losing 5 more, I was skeptical about the resulting findings. After all, the smaller the sample size, the more inaccurate the Spearman Rank-Order analysis. Additionally, the alpha level for the Race Relations Dissatisfaction vs. Movie Genres analysis was 0.10, meaning that even though some genres made the cut, there was weak evidence for rejecting the null hypothesis. It likely would have been a better idea to use another racial tensions indicator that had more data points - for instance, the number of racially motivated hate crimes in the U.S. every month.

VIII. IMPLICATIONS

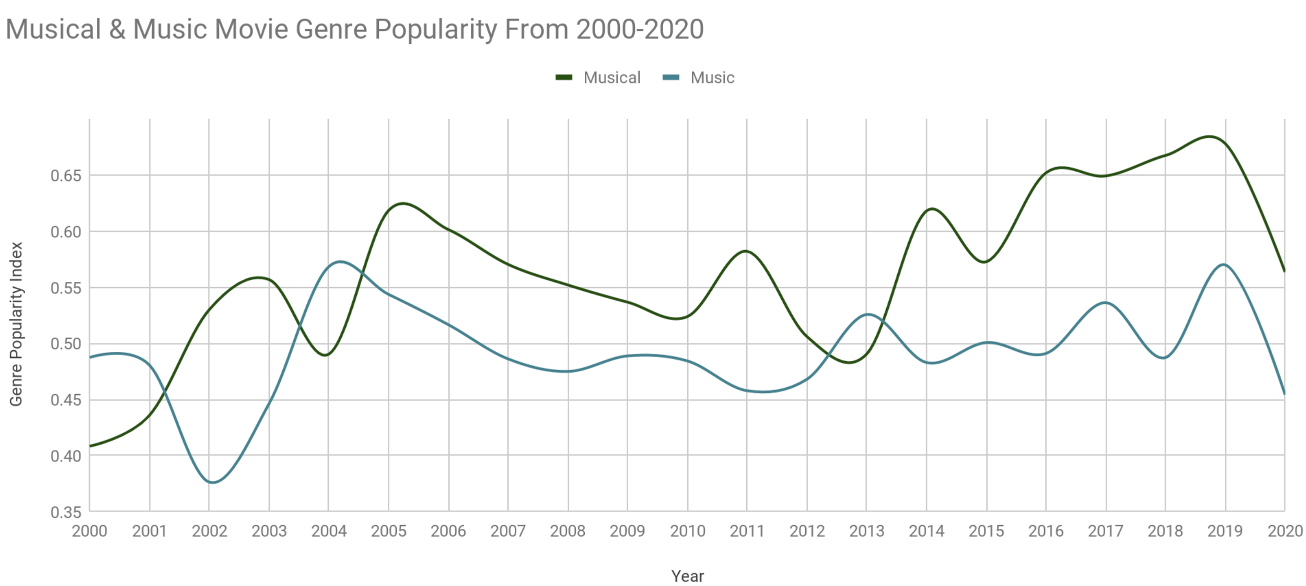
These findings can assist the entertainment industry in making more informed decisions on film production and release in the context of the economic and social sentiment within the United States at a given time. This could be in the form of choosing to release certain films during particular economic conditions that are favorable to the popularity of the film's genre or using predictive methods to determine the future state of economic conditions and based on that information, deciding the genre of film in the production stage.

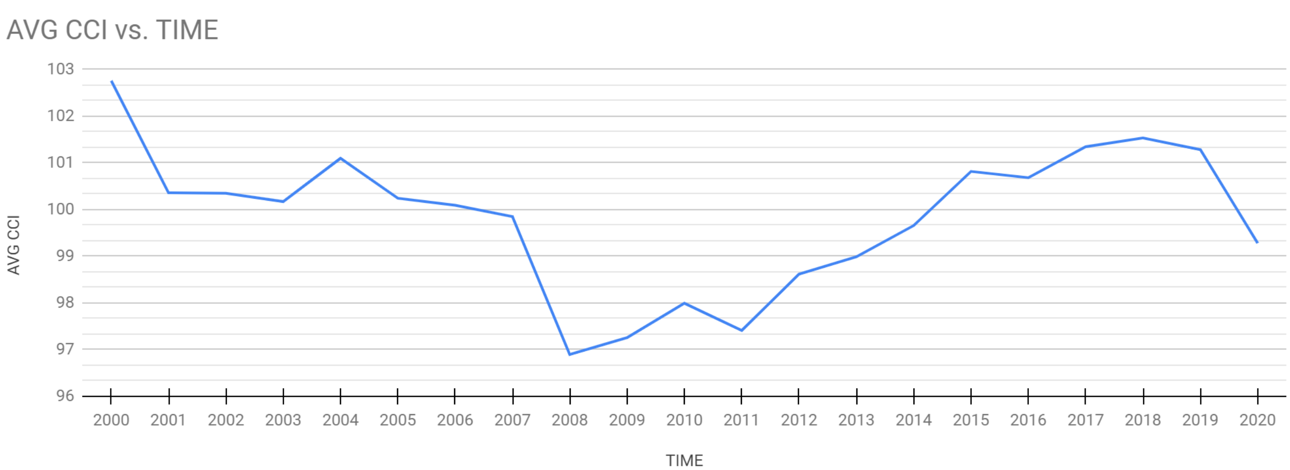
IX. CONCLUSION

In this study, a Spearman Rank-Order Correlation Coefficient was used to determine the strength of the association between the popularity of various movie genres within the U.S. and the U.S. Consumer Confidence Index as well as the association between the popularity of various movie genres within the U.S. and Race Relations Dissatisfaction. The genres which demonstrated the strongest monotonic correlations with the U.S. CCI were Music (positive), Musicals (positive), News (negative), and Short (negative). The genres which demonstrated the strongest monotonic correlations with Race Relations Dissatisfaction were Comedy (positive), Musical (positive), News (negative), and Short (negative).

X. FURTHER INVESTIGATION

Figure 9: *Comparing line graphs of U.S. CCI and the popularities of Musical and Music genres.*





After extracting these correlations and seeing interesting trends in the line graphs of movie genre popularities and the U.S. Consumer Confidence Index, further research may be conducted to determine the root cause of these relationships. Two particular relationships that warrant further investigation are those between Musical genre popularity and U.S. CCI, and Music genre popularity and U.S. CCI. Observation of the line graphs above shows strikingly similar trends from approximately 2005 to 2020. Both Music and Musical genre popularity starts a downward trend around 2004-2005 then starts increasing around 2012-2013, to sharply dip again in 2019 - just as the U.S. CCI does.

Seeing that these periods of U.S. economic growth and recession are mostly influenced by large-scale economic activity and policies, it is highly unlikely that the movie genre popularity is influencing the U.S. economy - rather, vice versa, assuming there is any causal relationship at all. Therefore, it is worth investigating in a separate study, what are some underlying social factors that may result in rising Musical and Music genre popularity during times of economic growth.

References

Amadeo, K. (2021, January 25). *What is the Consumer Confidence Index?* The Balance. Retrieved April 27, 2021, from https://www.thebalance.com/consumer-confidence-index-news-impact-3305743

Domestic box office. (2021). Retrieved April 27, 2021, from https://www.boxofficemojo.com/

Greenwood, D. (2010). Of sad men and dark comedies: Mood and gender effects on entertainment media preferences. *Mass Communication and Society*, *13*(3), 232-249. https://doi.org/10.1080/15205430903186526

Grodal, T. (2017). How film genres are a product of biology, evolution and culture—an embodied approach. *Palgrave Communications*, *3*(1). https://doi.org/10.1057/palcomms.2017.79

IMDb genre definitions. (2021). IMDb. Retrieved April 11, 2021, from https://help.imdb.com/article/contribution/titles/genres/GZDRMS6R742JRGAG#

Morah, C. (2019, November 11). *Consumer Confidence vs. Consumer Sentiment*. Investopedia. Retrieved April 27, 2021, from https://www.investopedia.com/ask/answers/09/consumer-confidence-sentiment-difference.asp

Niemiec, R. M. (2020). Character strengths cinematherapy: Using movies to inspire change, meaning, and cinematic elevation. *Journal of Clinical Psychology*, *76*(8), 1447-1462. https://doi.org/10.1002/jclp.22997

120+ examples of different movie genres. (2020, July 8). Nashville Film Institute. Retrieved April 27, 2021, from https://www.nfi.edu/movie-genres/

Race relations. (2021). *Gallup.com*. https://news.gallup.com/poll/1687/race-relations.aspx

Smithikrai, C. (2016). Effectiveness of teaching with movies to promote positive characteristics and behaviors. *Procedia - Social and Behavioral Sciences*, *217*, 522-530. https://doi.org/10.1016/j.sbspro.2016.02.033

Stoll, J. (2021, January 20). *Revenue of the U.S. motion picture and video industry 2019, by source*. Statista. Retrieved April 27, 2021, from https://www.statista.com/topics/964/film/#:~:text=Film%20entertainment%20is%20big%20business,dollars%20in%20revenue%20in%202019.

Tan, E. S. (2018). A psychology of the film. *Palgrave Communications*, *4*(1). https://doi.org/10.1057/s41599-018-0111-y

von Rimscha, M. B. (2013). It's not the economy, stupid! External effects on the supply and demand of cinema entertainment. *Journal of Cultural Economics*, *37*(4), 433-455.

Wühr, P., Lange, B. P., & Schwarz, S. (2017). Tears or fears? Comparing gender stereotypes about movie preferences to actual preferences. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00428

Yi, L. (2017). *What movies do you watch during bad & good economic times: The impact of movie genre on attendance over business cycles* [Unpublished master's thesis]. University of Guelph.