

**Homework 1**  
**Problem 3**  
**Ayad Masud**  
**733009045**

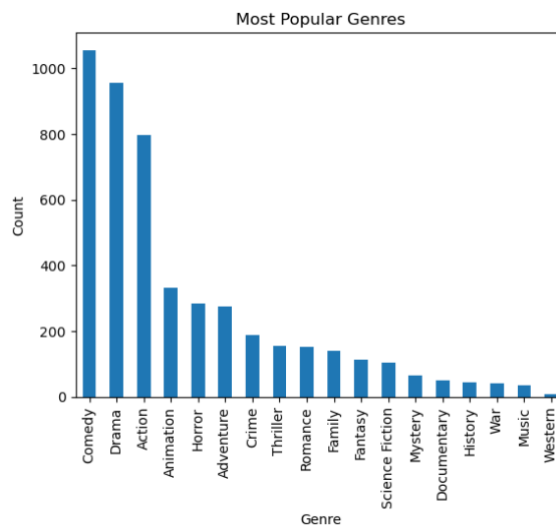
*I certify that I have personally done the coding, generated the figures and written the report without aid from anybody else, and that I have not plagiarized, self-plagiarized, or used AI-generated text. I certify that I have acknowledged any sources I used to complete this assignment. ARM.*

## 1 Part 1: Starting and Refining the Question

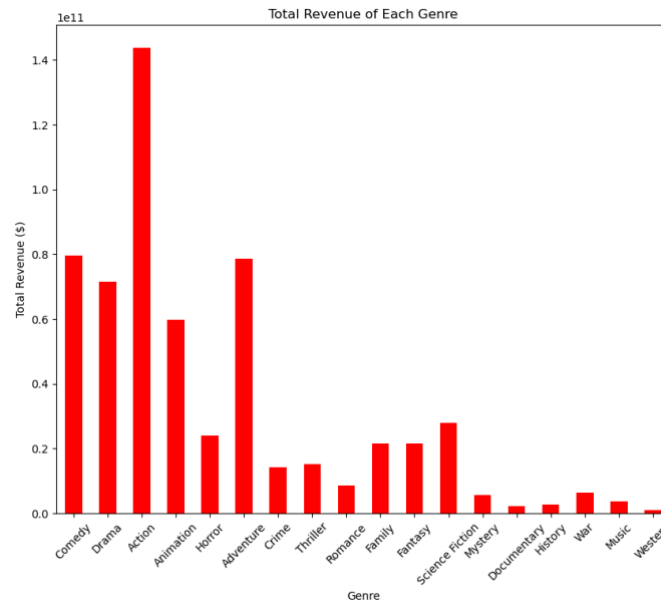
The filmmaking industry a multi-billion-dollar industry where the revenue generated is dependent on many of different factors. One of which is the genre of movies produced. Some genres may perform better while others perform subpar. Understanding the relationship between revenue and genre can be incredibly valuable to producers and investors who plan the making of these films. We can understand the relationship by asking and analyzing the question: “How does the genre of a movie impact its revenue?” Certain genres may be more appealing to a wider audience such as adventure or action themed movies. If there are more people watching movies that generates more revenue. Therefore, our working hypothesis will be that genre does significantly impact the total revenue of movies produced in the film industry.

## 2 Part 2: Exploring the Data

In **Figure 1** we can see the total number of movies produced per genre in this dataset. Comedy, drama, and action have a significant lead over other genres. The western genre does not have many movies produced anymore. This can be due to the fact that the peak of the “wild west” was many years ago now and is not as appealing to the newer generations that are watching the films currently being produced. Looking at **Figure 2** we can see the total revenue generated per genre. Matching the top 3 genres from **Figure 1** we would expect them to be in the top three revenue generating genres in **Figure 2**. However, this is not the case. We can see that comedy, action, and adventure are the top three. An explanation for this can be franchise movies and a strong, loyal fan base. Movie franchises such as Marvel, Lord of the Rings, Star Wars, and DC are all labeled as action or adventure movies. Due to the large fan base of these movies, people come back to the theater over and over to keep watching the movies produced for these series. The producers know that if they make a movie for a certain franchise, it will be sure to generate revenue and the genre for the most popular of these franchises just happens to be action, which explains why action is the largest bar in **Figure 2**.



**Figure 1.** A bar plot showing the total number of movies per genre.



**Figure 2.** A bar plot showing the total revenue per genre.

### 3 Part 3: Building Statistical Model

We will be using the ANOVA test to test our working hypothesis. The independent variable will be defined as the genre of the films and the dependent variable is the total revenue generated. Our **null** hypothesis or the default statement will be that there is no significant difference in revenue among genres. The **alternative** hypothesis or the claim that we are testing will be that there is a significant difference in revenue between genres. We will test using a significance level of 0.05 the “f\_oneway” function from the SciPy library to conduct the test.

#### Results of the test:

**F-stat:** 32.32

**P-score:**  $7.33 \times 10^{-10}$

**Action (Highest Revenue):** \$143 billion

**Comedy (2<sup>nd</sup> Revenue):** \$79 billion

**Adventure (3<sup>rd</sup> Revenue):** \$78 billion

**Western (Lowest Revenue):** \$944 million

### 4 Part 4: Interpreting Results

The statement  $p < 0.05$  is true after running the ANOVA test. This means that we reject the null hypothesis, or we reject that the genre does not have a significant impact on revenue. From the f-stat we can see that it is quite high indicating that there is a significant difference between the revenue generated by each genre. This confirms our hypothesis that genre affects the revenue generated. We can clearly see this by comparing how much the action genre made, \$143 billion, and how much the western genre made, \$944 million. The difference is massive and underlines how much genre effects the revenue generated by a film.

## 5 Part 5: Communicating Results

Our analysis has led us to the conclusion that genre does in fact have a significant impact on revenue. We can see that certain genres like action movies perform better and generate more revenue, while other more niche genres such as western movies don't do as well. Action movies generate the more revenue likely due to large amounts of funding, ad campaigns, and loyal fans. Understanding the relationship between genre and revenue can help people working in the industry plan and make decisions accordingly to maximize the revenue generated from the movies they make.

## 6 Resources used to achieve this goal

**Python Libraries:** NumPy, matplotlib, pandas, SciPy

**Canvas Home Page:** Homework template and Presentation on Epicycles of Analysis

## 7 References

- ADITYA JILLA. (2024). Movies Box Office Dataset (2000-2024). *Kaggle.com*.  
<https://doi.org/10347532/118ff04a5fb9db66a72e1e4cf797a13f>
- teagrackind. (2025). *movies-and-genre-analysis/movie-genres-analysis.ipynb at main* · *teagrackind/movies-and-genre-analysis*. GitHub. <https://github.com/teagrackind/movies-and-genre-analysis/blob/main/movie-genres-analysis.ipynb>