**ANALYSIS REPORT ON AMAZON MOVIES**

**INTRODUCTION**

The dataset is for the analysis on reviewing movie data to identify trends and patterns. The dataset used in this analysis contains information about Show id, Type, Title, Director, Main cast, Date added, Release year, Rating, Duration, Main genre, Rating.1.

**ABOUT THE DATA/DATA COLLECTION**

This is a structured and secondarily collected data. The data contains a table in csv files and has 9668 rows and 11 columns. The dataset was given by Ganiyat Olajumoke Ajala.

The analysis to be carried out is to answer the following questions;

1. What is the top 6 title with the highest rating?
2. How many movies was released in 2020?
3. What is the most listed release year?
4. Who are the top 10 main cast?
5. Identify the least recorded genre.

**TOOLS**

The tools used is Microsoft Excel for analysis and Power query for data cleaning.

**DATA CLEANING AND TRANSFORMATION**

The following are the different cleaning procedures;

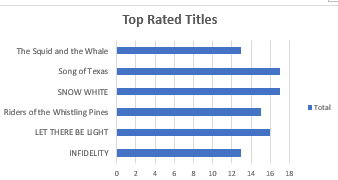
* The data contained blank, for the columns which were not needed for the analysis, the blanks were replaced with 0. For the blanks which were less than 5% of the entire data were not replaced.
* The cast column was changed to “main cast” and delimited by comma.
* The genre column was changed to “main genre” and delimited by comma.
* Checked for duplicates.

**EXPLORATORY DATA ANALYSIS AND INSIGHTS**

For the analysis to be understood and questions answered, pivot table was used to create charts. for proper analyzation and insights. Below are the questions, its analysis and insights;

1. **What is the top 6 title with the highest rating?**

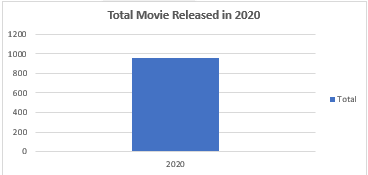
To find out the top 6 movies with the highest number of ratings, “Title” and “Rating.1” was filtered. With 17 points, “song of Texas” and “Snow white” was recorded to have the highest number of ratings; followed by “Let there be light” with 16 points in ratings.



A clustered bar chart was used to show the Top 6 rated titles.

1. **How many movies was released in 2020?**

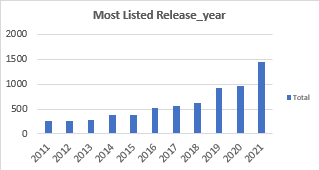
In order to find out how many movies were released in 2020, “title” and “Released Year” was filtered. Analysis showed that over 962 movies was released in the year 2020.



A clustered column chart was used to show the “total movie released in 2020”

1. What is the most listed release year?

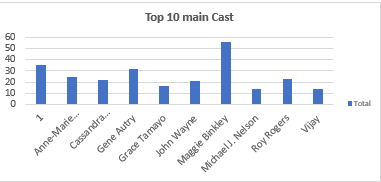
After filtering the top 10 most listed release year, 2021 was shown to have been listed multiple times. Which means 2021 is the year with highest number of released movie.



A clustered column chart was use to visualize the “most listed release year”.

1. **Who are the top 10 main cast?**

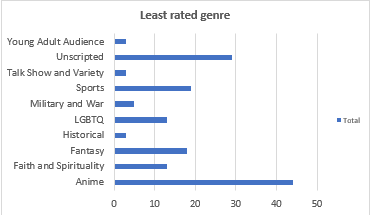
To know, the top 10 “main cast”, the top 10 main cast was filtered, and records shows that “Maggie Binkley” is the top 1 main cast with over 56 points; followed by “1” is the top 2 main cast with over 35 points.



A clustered column chart was used to show the top 10 main cast.

1. **Identify the least rated genre.**

In order to identify the least rated genre, the bottom 10 “main genre” and “Rating.1” were filtered. With only 3 points, “young adult audience”, “talk show and variety” and “historical” were the least rated genre.



A clustered bar chart was used to show the “least rated genre”.

**RECOMMENDATION**

After properly analyzing the data, I found out that the Anime and Unscripted genre have the highest number of ratings and views; whereas the young adult audience, talk show and variety, and historical genre were the least performing genre.

I recommend that the least genres should either be dropped or the creativity should be changed, and more resources should be channeled to the best performing genres. Also, the top actors should be featured more in movies to increase the ratings and viewing.

Lastly, the Anime genre should be advertised and released more as it is the best performing genre; that way it increases revenue and build the AMAZON as a brand.