DATA ANALYST PROJECT CASE STUDY

by: Padmasini Venugopal

INTRODUCTION

Rockbuster Stealth is a movie rental company that has a worldwide market. Facing stiff competition from streaming services such as Netflix and Amazon Prime, the management team plans to use its existing movie licenses to launch an online video rental service to stay competitive."

PROJECT OVERVIEW

DATA

The data set contains information about Rockbuster's film inventory, customers, and payments, among other things

- DATA: DATASET

DURATION

It took almost 3 weeks for me to complete the entire analysis.

TOOLS USED

- Microsoft Excel
- SQL (Structured Query Language)
- Tableau
- Microsoft Powerpoint

TECHNIQUES APPLIED

- Database querying
- Data filtering
- Data cleaning
- Summarizing
- Joining tables
- Subqueries
- Common Table Expressions


```
1   SELECT A.customer_id AS number_of_customers,
2   A.first_name AS "customer first name",
3   A.last_name AS "customer last name",
4   D.country, C.city,
5   SUM(E.amount) AS "total amount paid"
6   FROM customer A
7   INNER JOIN address B ON A.address_id = B.addr
8   INNER JOIN city C on B.city_id = C.city_id
9   INNER JOIN country D on C.country_id = D.coun
10  INNER JOIN payment E ON A.customer_id = E.cus
11  WHERE city IN('Aurora','Acua','Citrus Heights
12  Leopoldo','Tianjin','Hami','Cianjur')
13  GROUP BY A.customer_id, first_name, last_name
14  ORDER BY "total amount paid" DESC
15  LIMIT 5
```



```
Query        Query History
1    SELECT
2    MIN(rental_duration) AS minimum_rental_duration,
3    MAX(rental_duration) AS maximum_rental_duration,
4    AVG(rental_duration) AS average_rental_duration,
5    COUNT (rental_duration) AS count_rental_duration,
6    MIN(length) AS minimum_length,
7    MAX(length) AS maximum_length,
8    AVG(length) AS average_length,
9    COUNT (length) AS count_length,
10   MIN(rental_rate) AS minimum_rental_rate,
11   MAX(rental_rate) AS maximum_rental_rate,
12   AVG(rental_rate) AS average_rental_Rate,
13   COUNT (rental_rate) AS count_rental_rate,
14   MIN(replacement_cost) AS minimum_replacement_cost,
15   MAX(replacement_cost) AS maximum_replacement_cost,
16   AVG(replacement_cost) AS average_replacement_cost,
17   COUNT (replacement_cost) AS count_replacement_cost,
18   COUNT (*) AS count_rows,
19   MODE() WITHIN GROUP (ORDER BY rating)
20   AS modal_value_rating,
21   MODE() WITHIN GROUP (ORDER BY language_id)
22   AS modal_value_language_id,
23   MODE() WITHIN GROUP (ORDER BY release_year)
24   AS modal_value_release_year,
25   MODE() WITHIN GROUP (ORDER BY last_update)
26   AS modal_value_last_update,
27   MODE() WITHIN GROUP (ORDER BY title)
28   AS modal_value_title,
29   MODE() WITHIN GROUP (ORDER BY special_features)
30   AS modal value special features
Total rows: 1 of 1        Query complete 00:00:00.104
```

THE PROCESS

1. THE DATA PREPARATION PHASE

To streamline the Rockbuster Stealth data, I conducted a thorough scan for duplicates, missing numbers, and other anomalies. I subsequently eliminated extraneous columns and standardized data types.

2. THE ANALYSIS PHASE

Proceeding with my additional investigation was easier as the management team had already answered all my queries. I created separate SQL queries for each instance and saved the file as a CSV. I meticulously went through every question and gained a better understanding of the information I analyzed.

3. THE RESULTS

Finally, after all the analysis the final presentation was created using PowerPoint, and the visualizations were created using Tableau. The stakeholders were advised based on the results of the analysis with a clear disclaimer regarding the limitations of the data.
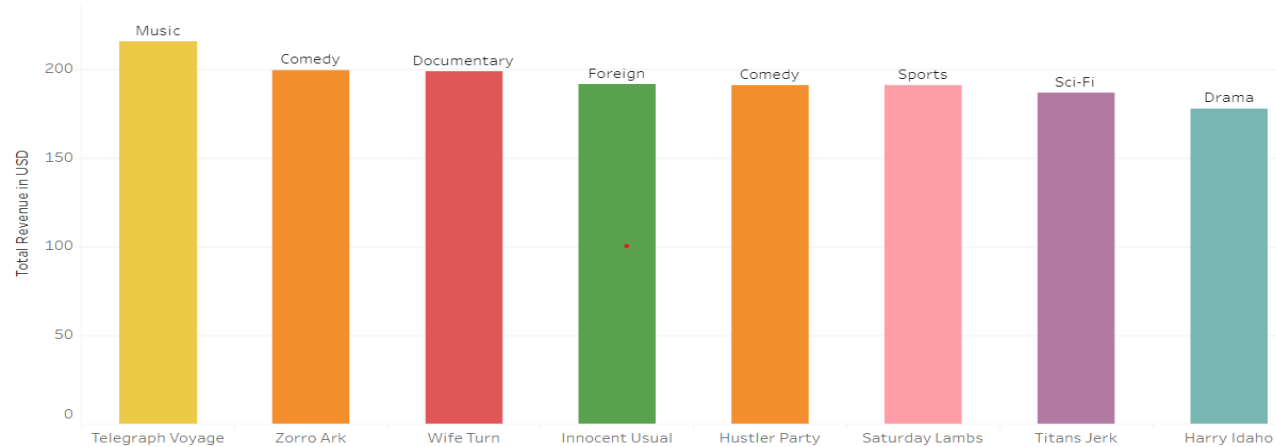
PERFORMANCE OF ROCK BUSTER STEALTH

- Average rental rate - $0.99- $4.99

- Total no of movies - 1000

- Average rental duration - 3-7 days

- No of active users – 584

- Average film length - 115 minutes

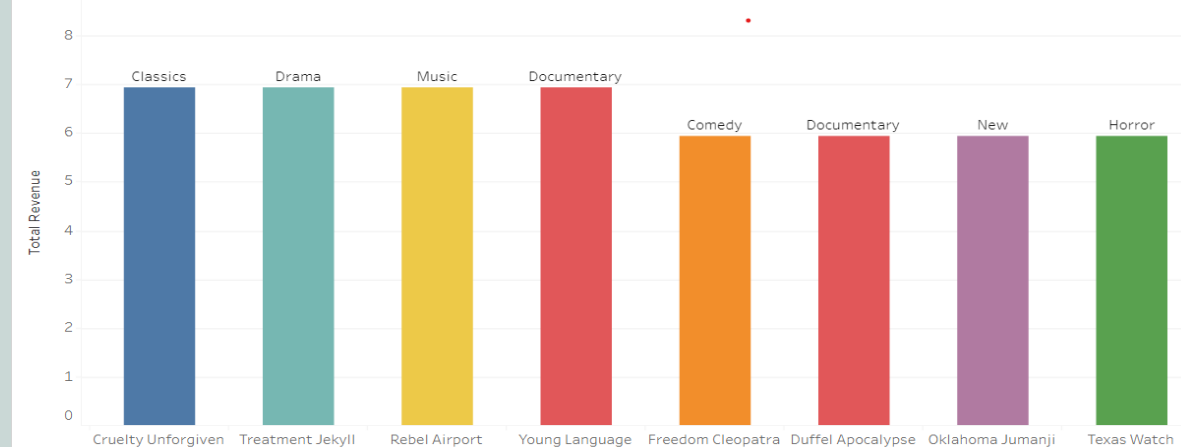- Movie ratings - G (178), PG (194), PG - 13 (223), R (195), NC-17 (210)

MOVIES AND REVENUE

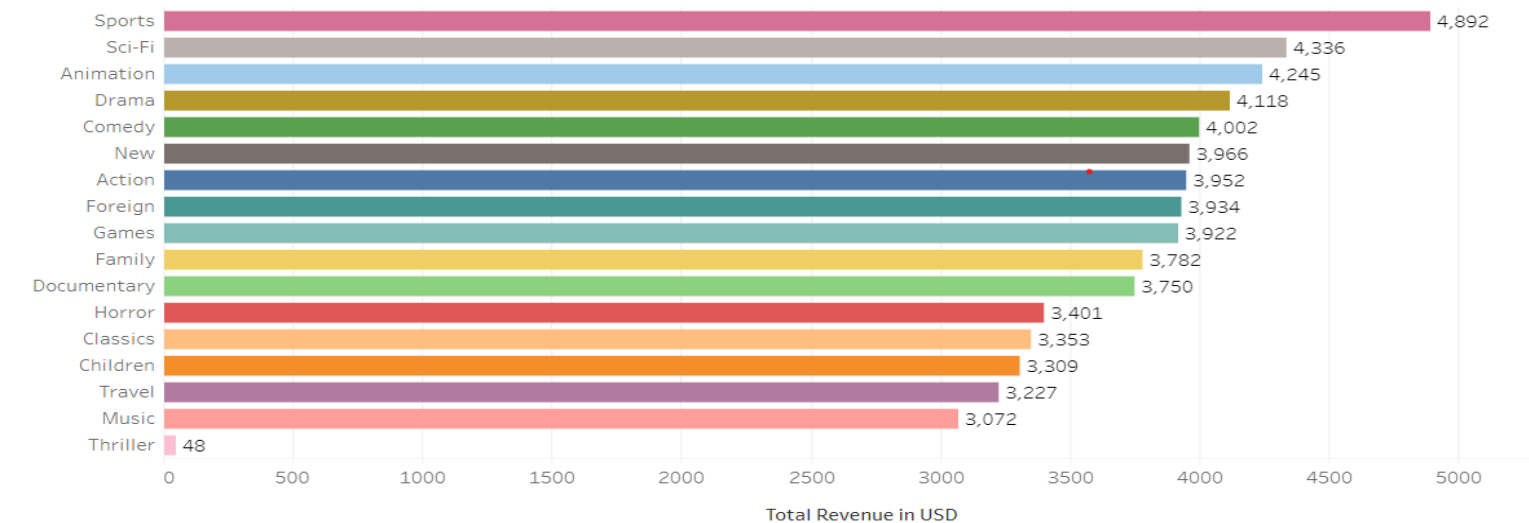## Top 10 movies contributing higher for revenue based on revenue

To maximize revenue, it's important to focus on movies in the music, comedy, and documentary genres. Telegraphic Voyage, Zorro Ark, and Wife Turn are all top performers in these categories.

The movies having lowest Revenue

The new and horror genres have yielded the lowest revenue among movies, with Texas Watch and Oklahoma Jumanji being the specific titles.

The total revenue by genres

Genre	Total Revenue in USD
Sports	4,892
Sci-Fi	4,336
Animation	4,245
Drama	4,118
Comedy	4,002
New	3,966
Action	3,952
Foreign	3,934
Games	3,922
Family	3,782
Documentary	3,750
Horror	3,401
Classics	3,353
Children	3,309
Travel	3,227
Music	3,072
Thriller	48

The analysis reveals that the sports genre movie generates the highest revenue, followed by science fiction, animation, drama, and comedy.

CUSTOMER ANALYSIS

THE TOP 10 CITIES WITH THE HIGHEST NUMBER OF CUSTOMERS

Countries	
India	60
China	53
United States	36
Japan	31
Mexico	30
Brazil	28
Russian Federation	28
Philippines	20
Turkey	15
Indonesia	14

THE TOP 10 CITIES WITH THE HIGHEST NUMBER OF CUSTOMERS WITHIN THE TOP 10 COUNTRIES

Coun..	City	
United States	Aurora	2
	Citrus Heights	1
Mexico	Acua	1
Japan	Iwaki	1
Indonesia	Cianjur	1
India	Ambattur	1
China	Shanwei	1
	Tianjin	1
	Hami	1
Brazil	So Leopoldo	1

THE TOP 5 CUSTOMERS FROM THOSE TOP CITIES, WHO HAVE PAID THE HIGHEST AMOUNT

Country	City	
India	Ambattur	111.76
China	Shanwei	109.71
Japan	Iwaki	106.77
Mexico	Acua	100.77
United States	Aurora	98.76

RETROSPECTIVE

WHAT WENT WELL?

I really enjoyed learning a new tool and writing queries. Additionally learned to create ERD (Entity relationship diagram) which helped me in understanding the large dataset with multiple columns better.

CHALLENGES FACED?

At first, learning a new tool was quite challenging for me as I come from a different background. Writing and recalling lengthy queries for even minor adjustments, tasks that can easily be accomplished using Excel posed significant difficulties. However, as I observed its positive impact on larger datasets and persevered through consistent practice, I am now able to use the tool effectively.

RECOMMENDATIONS

✓ To increase revenue, it is recommended to focus on certain high-revenue-generating movies rather than specific genres. Consider removing movies with the least revenue. Additionally, offering special discounts to customers who rent movies for more than five days can significantly increase rental duration.

✓ To expand sales, it is suggested to increase marketing strategies in major markets such as Asia & Pacific, Europe, and South/Latin America. Marketing promotions should also be implemented in other regions to increase sales.

✓ To cater to specific markets, it is advised to introduce more movies in the top five countries and in their preferred language. This can create more interest among customers and improve sales.

✓ To retain loyal customers and attract new ones, it is recommended to reward high-paying customers for their loyalty and implement a referral code system to increase customer numbers.

THANK YOU!

For further queries kindly contact

drpadmasini.v@gmail.com