Disney+ Database Analytics

BY SAIFUDDIN SANUSI

Agenda

- Problem Statement
- Data Overview
- Data Cleaning
- Data Analysis (conclusion & recommendations)

Problem Statement

Future content strategies are unsure on how and what they need to consider to improve Disney+ platform's offering.

What factors have more influence on view, and how do the relationships vary across different settings?

Data Overview

- **Data:** Disney+ movies and TV shows
- **Source:** Kaggle / JustWatch
- **Date:** March 2023
- **Size:** titles.csv (1k titles, 15 columns) ; credits.csv (30k credits of actors and directors, with 5 columns)
- **Fields:** titles, show type, description, release year, age certification, runtime, genre, production countries, number of seasons, imdb score, imdb votes, tmdb popularity, tmdb score

Data Cleaning

```sql
-- Part 1 - DATA CLEANING
-- by running this code, the entire sql script will be executed.

-- Step 1 - Creating a new database 'disney'
DROP DATABASE disney;
CREATE DATABASE disney;



-- Step 2 - Creating new tables in the database
-- dropping existing tables in the database
DROP TABLE IF EXISTS titles CASCADE;
DROP TABLE IF EXISTS credits CASCADE;

-- creating new tables to hold the raw data
CREATE TABLE titles(
id VARCHAR(15),
t_title TEXT,
t_type VARCHAR(5),
description TEXT,
release_year INTEGER,
age_certification VARCHAR(10),
runtime INTEGER,
genres VARCHAR(100),
production_countries VARCHAR(50),
seasons FLOAT,
imdb_id VARCHAR(15),
imdb_score FLOAT,
imdb_votes FLOAT,
tmdb_popularity FLOAT,
tmdb_score FLOAT
);
```

```sql
CREATE TABLE credits(
person_id INTEGER,
id VARCHAR(15),
name TEXT,
characters TEXT,
role VARCHAR(8)
);



-- Step 3 - loading the data into the titles and credits table
-- showing the first 10 rows from the tables
SELECT *
FROM titles
LIMIT 10;

SELECT *
FROM credits
LIMIT 10;



-- Step 4 - Checking the size of the dataset
SELECT COUNT(*) FROM titles;
SELECT COUNT(*) FROM credits;



-- Step 5 - Removing the redundant columns
ALTER TABLE titles DROP COLUMN imdb_id;
```

```sql
-- Step 6 - Looking for null values
-- from titles table
SELECT COUNT(*) FROM titles WHERE id IS NULL;
SELECT COUNT(*) FROM titles WHERE t_title IS NULL;
SELECT COUNT(*) FROM titles WHERE t_type IS NULL;
SELECT COUNT(*) FROM titles WHERE release_year IS NULL;
SELECT COUNT(*) FROM titles WHERE age_certification IS NULL;
SELECT COUNT(*) FROM titles WHERE age_certification= '[]';
SELECT COUNT(*) FROM titles WHERE runtime = 0;
SELECT COUNT(*) FROM titles WHERE genres IS NULL;
SELECT COUNT(*) FROM titles WHERE genres= '[]';
SELECT COUNT(*) FROM titles WHERE production_countries IS NULL;
SELECT COUNT(*) FROM titles WHERE production_countries= '[]';
SELECT COUNT(*) FROM titles WHERE imdb_id IS NULL;
SELECT COUNT(*) FROM titles WHERE imdb_score IS NULL;
SELECT COUNT(*) FROM titles WHERE tmdb_popularity IS NULL;
SELECT COUNT(*) FROM titles WHERE tmdb_score IS NULL;

-- from credits table
SELECT COUNT(*) FROM credits WHERE person_id IS NULL;
SELECT COUNT(*) FROM credits WHERE id IS NULL;
SELECT COUNT(*) FROM credits WHERE name IS NULL;
SELECT COUNT(*) FROM credits WHERE characters IS NULL;
SELECT COUNT(*) FROM credits WHERE role IS NULL;



-- Step 7 - Deleting the null values
DELETE FROM titles
WHERE tmdb_popularity IS NULL
      OR tmdb_score IS NULL
      OR imdb_score IS NULL
      OR imdb_votes IS NULL;
```

Data Cleaning

```sql
-- Step 8 - Replacing the Null Values
-- Setting the null values in character column to 'No information'
UPDATE credits
SET characters = 'N/A'
WHERE characters IS NULL;

-- Setting the null values in seasons to 0 which corresponds to Movies
UPDATE titles
SET seasons = 0
WHERE seasons IS NULL;

-- Setting the null values in age_certification to 'Others'
UPDATE titles
SET age_certification = 'Others'
WHERE age_certification IS NULL;

-- Setting the values with [] in genres column with 'N/A'
UPDATE titles
SET genres = 'N/A'
WHERE genres = '[]';

-- Setting the values with [] in production_countries column with 'N/A'
UPDATE titles
SET production_countries = 'N/A'
WHERE production_countries = '[]';
```

```sql
-- Step 9 - Looking for duplicates
-- in titles table
SELECT id, COUNT(id) as count
FROM titles
GROUP BY id
HAVING COUNT(id) >1 ;
```

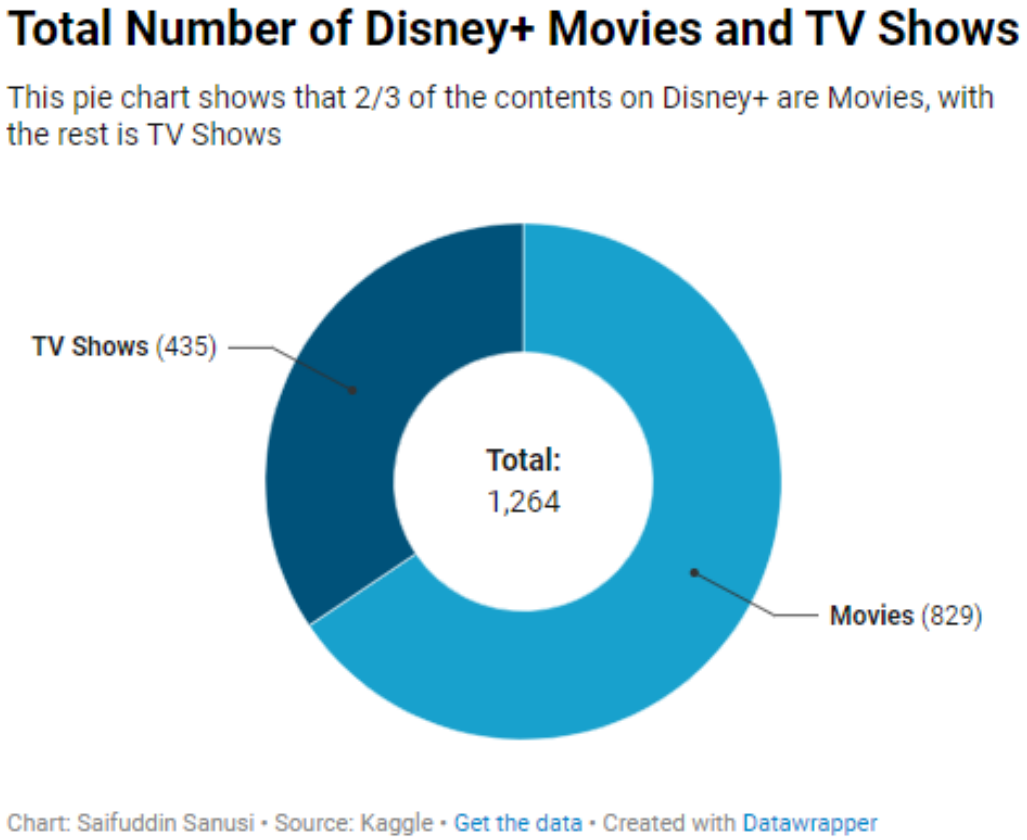
```sql
-- In credits table
-- This SQL code is selecting id, person_id, and the count of each group as count
-- from the raw_credits table. The GROUP BY clause groups the rows by id, person_id,
-- and role. The HAVING clause filters the groups with a count greater than 1.

SELECT id, person_id, COUNT(*) as count
FROM credits
GROUP BY id, person_id, role
HAVING COUNT(*) > 1
ORDER BY count desc limit 10;
```

```sql
-- Step 10- Some more unusual titles starting with '#' and '@'
-- The SQL query selects the title column from the titles table.
-- The WHERE clause filters out the rows where the titles column starts
-- with a digit or an uppercase letter.

SELECT t_title
FROM titles
WHERE t_title ~ '^#'
ORDER BY t_title;

SELECT t_title
FROM titles
WHERE t_title ~ '^@'
ORDER BY t_title;
```

```sql
-- trimming the # from the Left hand side of the titles
UPDATE title
SET t_title = LTRIM(t_title, '#')
WHERE t_title ~ '^#';

-- trimming the @ from the Left hand side of the titles
UPDATE title
SET t_title = LTRIM(t_title, '@')
WHERE t_title ~ '^@';
```

Data Analysis – Total number movies and TV show available on Disney+

```sql
-- 1. What is the total number of movies and TV shows available on Disney+?
SELECT
  t_type, COUNT(*) as total_count
FROM titles
GROUP BY t_type;
```



Conclusion:

- 2/3 of Disney+ dominated by movies selection

Data Analysis – Top 10 movies and TV show on Disney+ (IMDB Score)

```sql
-- 2. What were the top 10 movies according to IMDB score?
SELECT t_title,
t_type,
imdb_score
FROM titles
WHERE imdb_score >= 8.0
AND t_type = 'MOVIE'
ORDER BY imdb_score DESC
LIMIT 10;
```

```sql
-- 3. What were the top 10 shows according to IMDB score?
SELECT t_title,
t_type,
imdb_score
FROM titles
WHERE imdb_score >= 8.0
AND t_type = 'SHOW'
ORDER BY imdb_score DESC
LIMIT 10;
```

Top 10 Disney + TV Shows Based on IMDB Score

Title	Type	IMDB Score
Bluey	SHOW	9.50
The Beatles: Get Back	SHOW	9.00
Critter Fixers: Country Vets	SHOW	8.80
Heartland Docs, DVM	SHOW	8.80
Alaska Animal Rescue	SHOW	8.80
Light & Magic	SHOW	8.80
The Simpsons	SHOW	8.70
One Strange Rock	SHOW	8.70
The Mandalorian	SHOW	8.70
Super/Natural	SHOW	8.70

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Top 10 Disney + Movie Based on IMDB Score

Title	Type	IMDB Score
The Empire Strikes Back	MOVIE	8.70
Star Wars	MOVIE	8.60
BTS: Permission to Dance on Stage - LA	MOVIE	8.60
The Lion King	MOVIE	8.50
WALLÂ·E	MOVIE	8.40
Avengers: Infinity War	MOVIE	8.40
Avengers: Endgame	MOVIE	8.40
Hamilton	MOVIE	8.40
Folklore: The Long Pond Studio Sessions	MOVIE	8.40
The Rescue	MOVIE	8.40

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Data Analysis –Bottom 10 Movies & TV Show on Disney+ (IMDB Score)

```sql
-- 4. What were the bottom 10 movies according to IMDB score?
SELECT t_title,
t_type,
imdb_score
FROM titles
WHERE t_type = 'MOVIE'
ORDER BY imdb_score ASC
LIMIT 10;
```

```sql
-- 5. What were the bottom 10 shows according to IMDB score?
SELECT t_title,
t_type,
imdb_score
FROM titles
WHERE t_type = 'SHOW'
ORDER BY imdb_score ASC
LIMIT 10;
```

Bottom 10 Disney + TV Shows Based on IMDB Score

Title	Type	IMDB Score
Storage Wars: Miami	SHOW	4.40
Paradise Islands	SHOW	4.40
Pickle & Peanut	SHOW	4.20
Walk the Prank	SHOW	4.10
The Montaners	SHOW	4.00
The Quest	SHOW	3.90
Bizaardvark	SHOW	3.80
Dance Moms: Miami	SHOW	3.40
Billy Dilleyâ™s Super-Duper Subterranean Summer	SHOW	3.20
The Proud Family: Louder and Prouder	SHOW	2.80

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

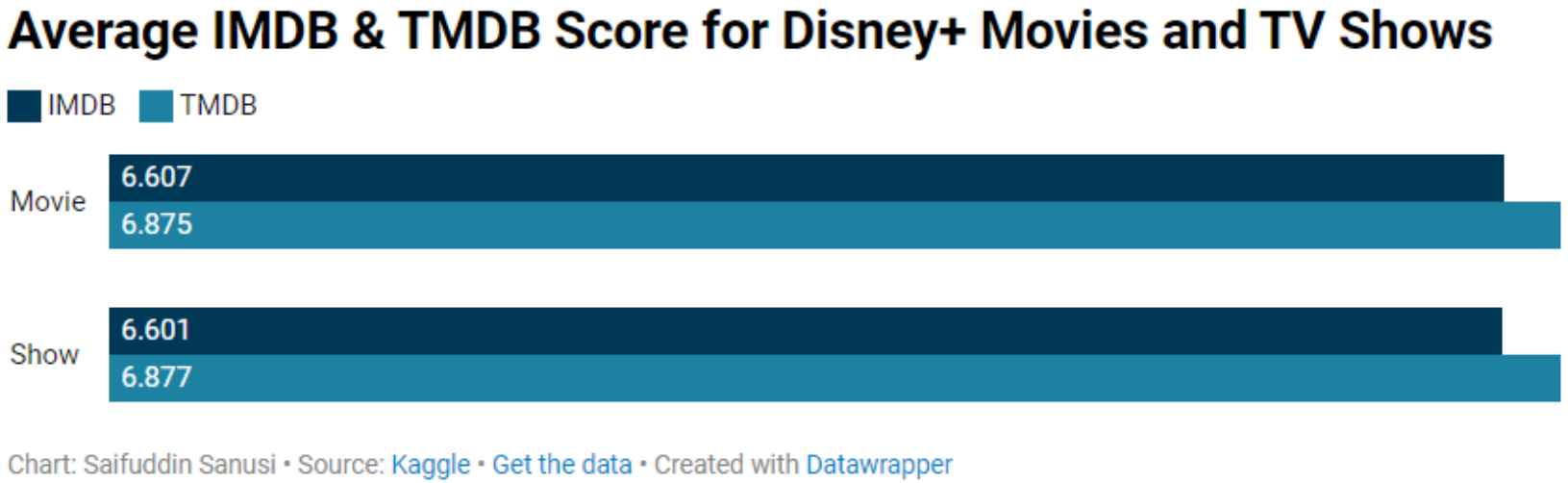
Bottom 10 Disney + Movies Based on IMDB Score

Title	Type	IMDB Score
Beverly Hills Chihuahua	MOVIE	3.90
Freaky Friday	MOVIE	3.80
Kim Possible	MOVIE	3.70
Dory's Reef Cam	MOVIE	3.60
Home Sweet Home Alone	MOVIE	3.60
Home Alone: The Holiday Heist	MOVIE	3.50
Kazaam	MOVIE	3.10
Home Alone 4	MOVIE	2.60
Hannah Montana & Miley Cyrus: Best of Both Worlds Concert	MOVIE	2.50
Jonas Brothers: The Concert Experience	MOVIE	1.60

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Data Analysis – Average IMBD & TMDB Score

```sql
-- 6. What were the average IMDB and TMDB scores for shows and movies?
SELECT DISTINCT type,
ROUND(AVG(imdb_score),3) AS avg_imdb_score,
ROUND(AVG(tmdb_score),3) as avg_tmdb_score
FROM titles
GROUP BY t_type ;
```



Conclusion:

1.Minimal Difference Between Movies and Shows:

The difference of IMDB and TMDB score for movies and shows is negligible. This indicate that Disney+ content, whether movie or show, is equally appreciated by viewers across these platforms.

2.Slight Preference for TMDB Scores:

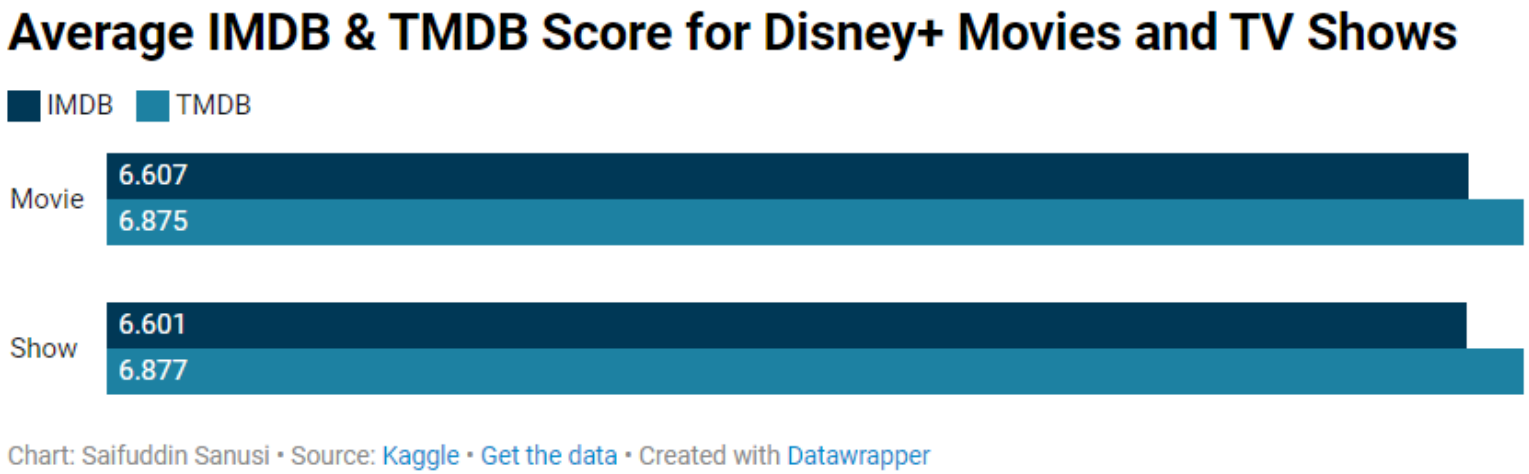
Both movies and shows receive slightly higher scores on TMDB than on IMDB. This could indicate that TMDB users might be slightly more favorable in their ratings for Disney+ content, or it could reflect differences in user demographics between the two platforms.

3.Above-Average Ratings:

With both scores being above 6.6, Disney+ content is generally well-received by viewers. While these scores do not represent top-tier content, they suggest that Disney+ offers a consistent range of good, but not exceptional content, which meets the expectations of its target audience.

Data Analysis – Average IMBD & TMDB Score

```sql
-- 6. What were the average IMDB and TMDB scores for shows and movies?
SELECT DISTINCT type,
ROUND(AVG(imdb_score),3) AS avg_imdb_score,
ROUND(AVG(tmdb_score),3) as avg_tmdb_score
FROM titles
GROUP BY t_type ;
```



Recommendations:

1.Targeted Improvement for Higher Ratings:

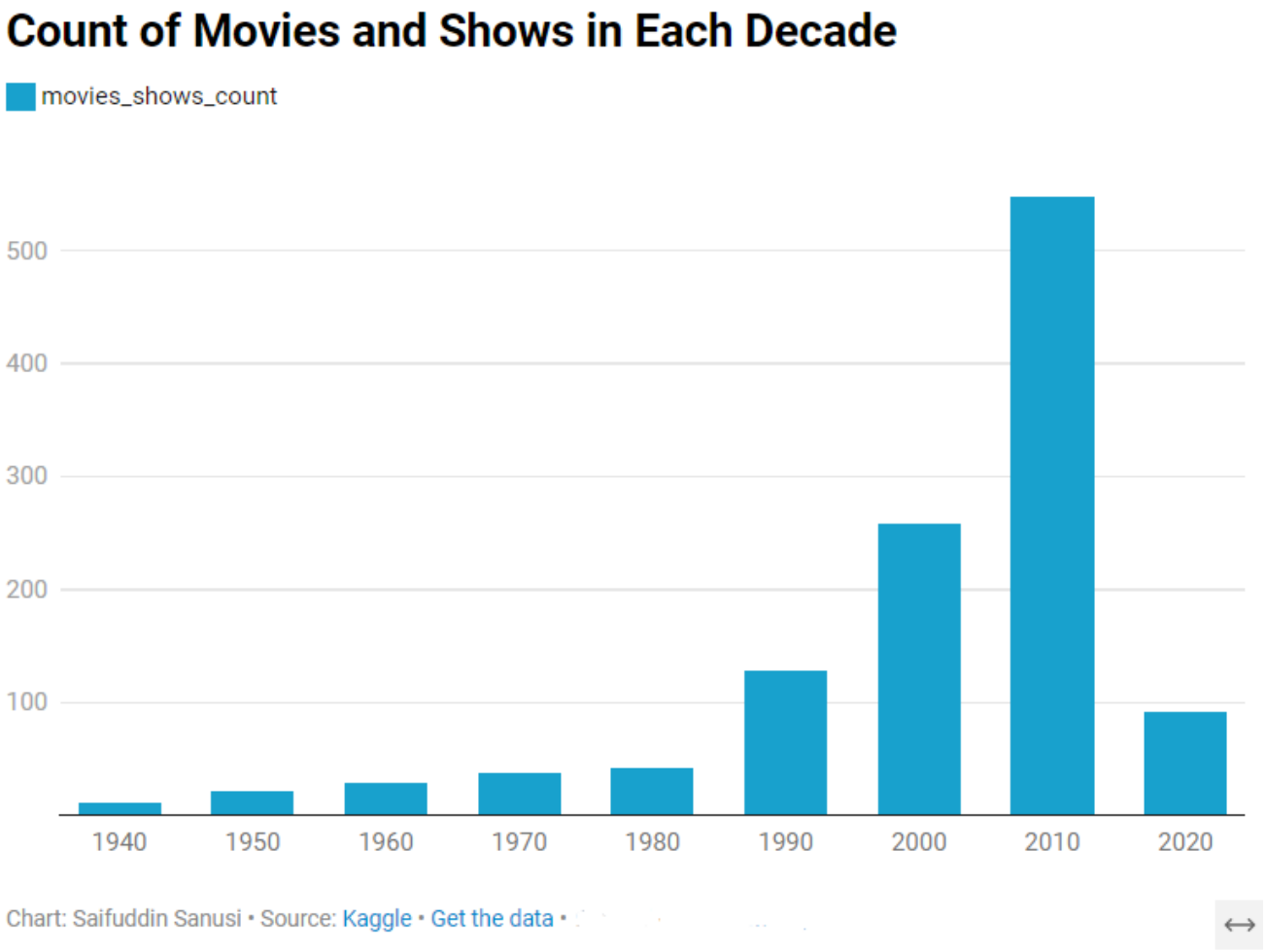
Although the average ratings are good, there is room for improvement to bring scores closer to or above 7.0. Disney+ could focus on increasing the quality of storytelling, character development, and production value in both movies and TV shows. This could involve increasing investment in original content and big-budget productions that can attract higher ratings and critical acclaim.

2.Encourage User Engagement on Review Platforms:

Since TMDB shows a slightly higher rating, Disney+ could partner with or promote user engagement on platforms like TMDB to ensure that its content receives a fair representation. Encouraging users to rate content more frequently across platforms may help boost visibility and credibility. Disney+ could also work on strategies to improve ratings through audience interaction (e.g., providing review prompts after viewing).

Data Analysis – Number of Movies & TV Shows in Each Decades

```sql
-- 7. Count of movies and shows in each decade
SELECT CONCAT(FLOOR(release_year / 10) * 10, 's') AS decade,
       COUNT(*) AS movies_shows_count
FROM titles
WHERE release_year >= 1940
GROUP BY CONCAT(FLOOR(release_year / 10) * 10, 's')
ORDER BY decade;
```



Analysis

- **Increasing Trend Over Time:** There is a clear increasing trend in the number of movies and shows from the 1940s through the 2010s

- **Decline in 2020s:** There is a noticeable decline in the number of movies and shows in the 2020s with only 92, which could be attributed to the decade not being complete at the time of data collection

Data Analysis – Average IMDB & TMDB (Production Countries)

```sql
-- 8. What were the average IMDB and TMDB scores for each production country?
SELECT DISTINCT production_countries,
ROUND(AVG(imdb_score),2) AS avg_imdb_score,
ROUND(AVG(tmdb_score),2) AS avg_tmdb_score
FROM titles
GROUP BY production_countries
ORDER BY avg_imdb_score DESC;
```

Average IMDB & TMDB Score by Production Countries

Page 1 of 3

	Countries	Average of imdb_score	Average of tmdb_score
1	['NZ', 'US', 'GB']	9	8
2	['KR', 'US']	9	9
3	['US', 'MX']	8	8
4	['PR']	8	7
5	['KR']	8	9
6	['AU']	8	7
7	['ES', 'CO']	8	7
8	['DE', 'US']	8	8
9	['ES', 'CH', 'US', 'FR', 'MC']	8	7
10	['TW']	8	5
11	['DE']	8	8
12	['CL']	8	8
13	['AE']	8	8
14	['BR', 'CA', 'DK', 'US']	8	7
15	['GB', 'US', 'AU']	8	7
16	['ES', 'US']	7	9
17	['AT']	7	8
18	['CN', 'GB', 'US']	7	7
19	['US', 'FR']	7	7
20	['US', 'KI']	7	7
21	['GB']	7	7
22	['CI', 'UG', 'US']	7	7
23	['NL']	7	7
24	['GB', 'FR']	7	6
25	['US', 'JP']	7	7

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Average IMDB & TMDB Score by Production Countries

	Countries	Average of imdb_score	Average of tmdb_score
26	['FR']	7	7
27	['AU', 'CA', 'FR', 'US']	7	7
28	['CO']	7	9
29	['AR', 'BR']	7	8
30	['GB', 'BE', 'FR']	7	6
31	['PH']	7	7
32	['GB', 'US']	7	7
33	['CN']	7	8
34	['US', 'BE', 'GB']	7	7
35	['JP']	7	7
36	['IN', 'US']	7	7
37	['US']	7	7
38	['US', 'GB']	7	7
39	['US', 'GT']	7	9
40	['IN']	7	8
41	['PL', 'SI', 'US', 'CZ']	7	7
42	['SE', 'US', 'NO']	7	7
43	['BR', 'US']	6	6
44	['CA', 'DK', 'US']	6	7
45	['GB', 'ZA', 'DE', 'US', 'FR']	6	8
46	['JP', 'US']	6	7
47	['CA', 'FR', 'US', 'GB']	6	7
48	['KR', 'US', 'FR']	6	7
49	['AR']	6	8
50	['FR', 'US']	6	6

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Average IMDB & TMDB Score by Production Countries

	Countries	Average of imdb_score	Average of tmdb_score
51	['GB', 'US', 'DE']	6	6
52	['MX', 'US', 'JP']	6	7
53	['CA', 'US']	6	7
54	['ES']	6	6
55	['NZ', 'US']	6	6
56	['BR']	6	6
57	['US', 'CA']	6	6
58	['AU', 'US']	6	6
59	['DK', 'US', 'CA']	6	6
60	['US', 'NZ']	6	6
61	['CA', 'US', 'AU']	6	6
62	['GB', 'US', 'DE', 'IE']	6	6
63	['MX', 'AR']	6	8
64	['TR']	6	6
65	['IT']	6	8
66	['US', 'CN']	6	7
67	['SK', 'US']	6	6
68	['CA']	6	6
69	['US', 'AU']	5	6
70	['ES', 'AR', 'IT']	5	7
71	['US', 'DE']	5	6
72	['CN', 'HK', 'US']	5	7
73	['US', 'GB', 'DE']	4	4

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Average IMDB & TMDB Scores for Disney+ Each Production Country

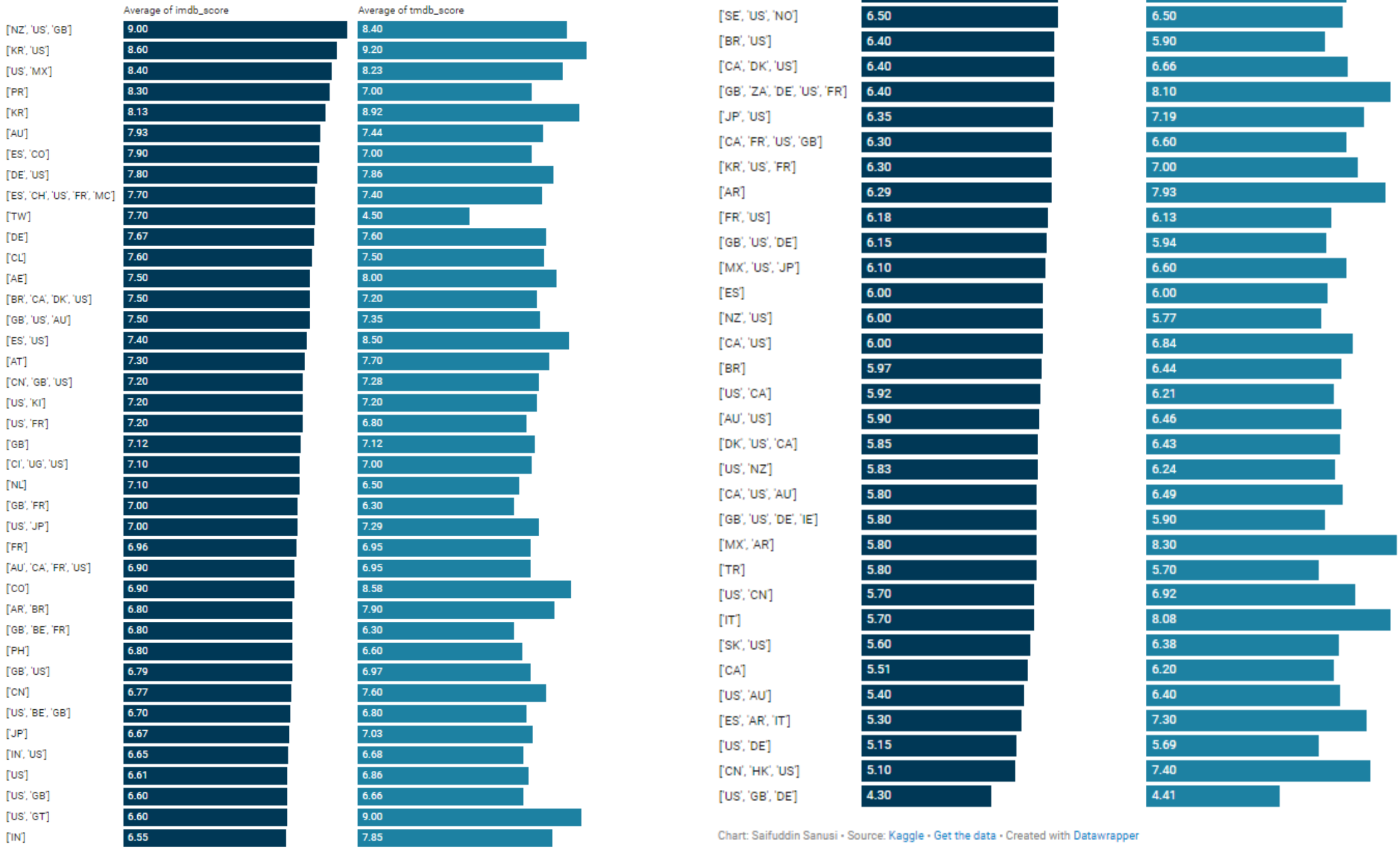


Chart: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Data Analysis – Average IMDB & TMDB (Production Countries)

Analysis:

Top-Performing Countries:

- The highest average IMDB score is 9.00 for the combination of countries ['NZ', 'US', 'GB'], with a TMDB score of 8.40.
- Korea ('KR') and the US ('US') have the highest TMDB score at 9.20, with a strong IMDB score of 8.60. Other notable high-scoring combinations include ['US', 'MX'] (IMDB: 8.40, TMDB: 8.23), and Puerto Rico ('PR') (IMDB: 8.30, TMDB: 7.00).

Strong Regional Variations:

There is a significant range of ratings across different country combinations, indicating that regional preferences vary greatly. For instance, ['NZ', 'US', 'GB'] and ['KR', 'US'] rate content much higher than countries like ['US', 'GB', 'DE'] (IMDB: 4.30, TMDB: 4.41).

Underperforming Regions:

Some countries, particularly combinations like ['US', 'GB', 'DE'] (IMDB: 4.30, TMDB: 4.41) and ['US', 'DE'] (IMDB: 5.15, TMDB: 5.69), have relatively low scores. This could indicate less favorable reception for Disney+ content or lower engagement with rating platforms in these regions.

Recommendations:

Target High-Scoring Regions for New Releases:

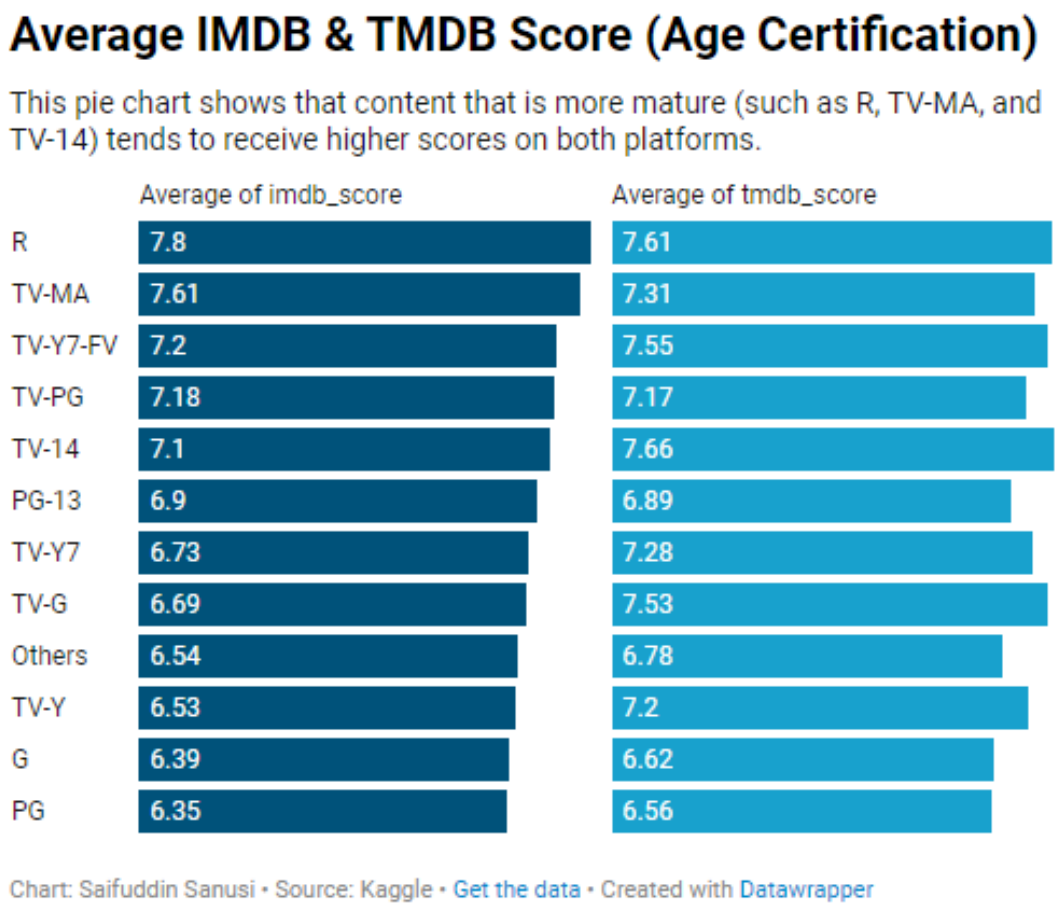
Disney+ should leverage regions that show consistently high ratings for content distribution, especially for high-budget and original content. Countries like ['NZ', 'US', 'GB'], ['KR', 'US'], and ['US', 'MX'] show high appreciation of content, making them ideal targets for new exclusive releases and marketing campaigns.

Address Low-Scoring Markets:

For countries with lower scores, such as ['US', 'GB', 'DE'] and ['US', 'DE'], Disney+ should consider conducting regional market research to identify viewer preferences, expectations, and feedback on why content is underperforming. Localized content or adjustments to marketing strategies could help improve reception

Data Analysis –Average IMBD & TMDB Score(Age Certification)

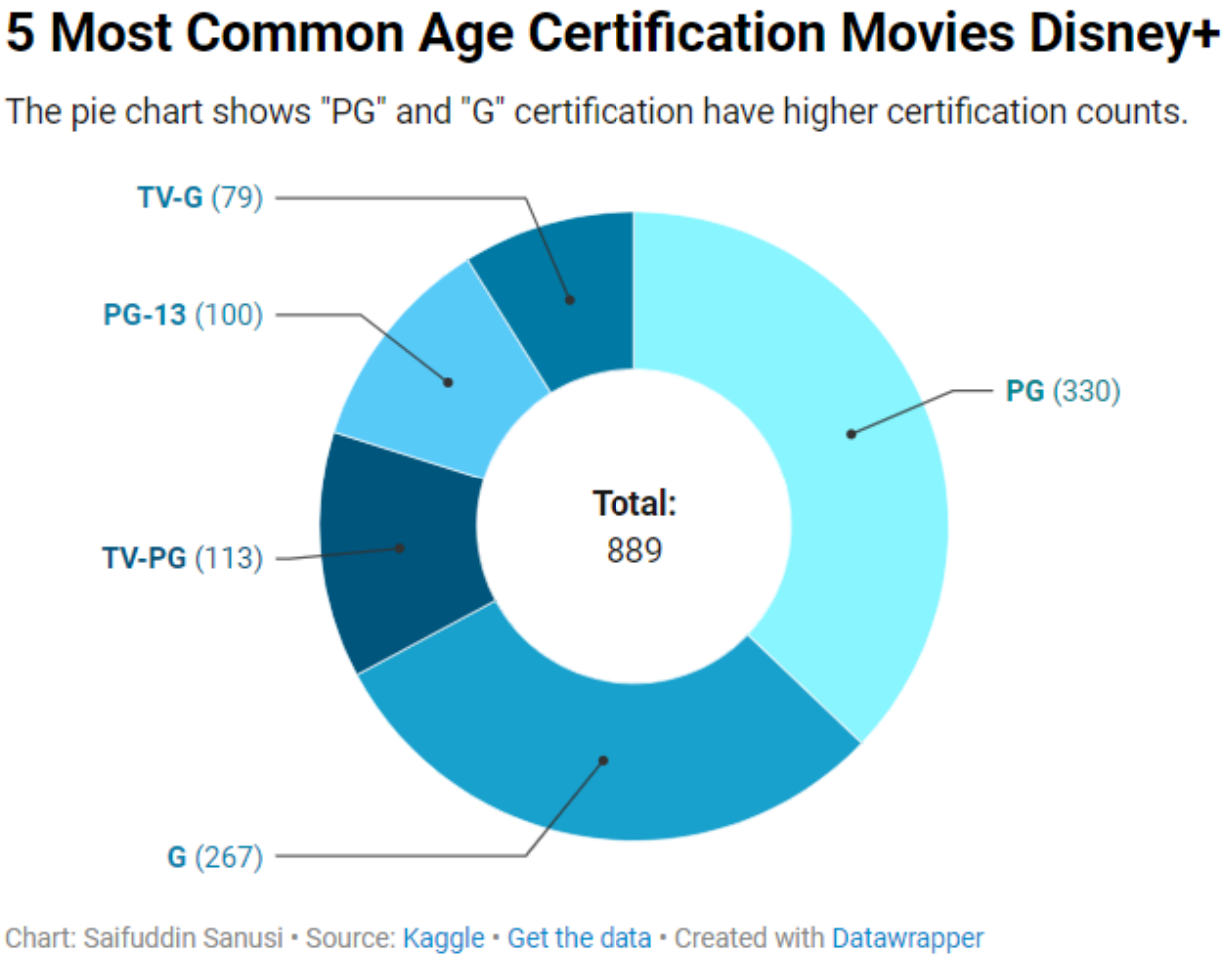
```sql
-- 9. What were the average IMDB and TMDB scores for each age certification for shows and movies?
SELECT DISTINCT age_certification,
ROUND(AVG(imdb_score),2) AS avg_imdb_score,
ROUND(AVG(tmdb_score),2) AS avg_tmdb_score
FROM titles
GROUP BY age_certification
ORDER BY avg_imdb_score DESC;
```



* Number 9 will be analyse together with number 10

Data Analysis – Top 5 Common Age Certification in Disney+

```sql
-- 10. What were the 5 most common age certifications for movies?
SELECT age_certification,
COUNT(*) AS certification_count
FROM titles
WHERE t_type = 'Movie'
AND age_certification != 'N/A'
GROUP BY age_certification
ORDER BY certification_count DESC
LIMIT 5;
```



Overall Analysis- Number 9 & 10

Conclusion:

1.Family-Oriented Content Has Lower Average Scores:

- PG-rated and G-rated content, despite having a high certification count (330 for PG and 267 for G), have lower average ratings on both IMDB (6.35 for PG, 6.39 for G) and TMDB (6.56 for PG, 6.62 for G).
- This suggests that while this type of content is widely produced, it is generally less favorably rated compared to content with more mature themes.

2.Mature Content Scores Higher, Despite Lower Certification Counts:

- PG-13 and R-rated content has higher ratings on both IMDB (6.9 for PG-13, 7.8 for R) and TMDB (6.89 for PG-13, 7.61 for R), but the certification count is lower (100 for PG-13).
- Similarly, TV-14 and TV-MA content also receive high ratings (TV-14: 7.1 on IMDB, 7.66 on TMDB; TV-MA: 7.61 on IMDB, 7.31 on TMDB) but their certification counts are comparatively lower.

3.Children's Programming Is Rated Better on TMDB Than on IMDB:

- TV-Y and TV-Y7 content shows higher scores on TMDB (7.20 for TV-Y and 7.28 for TV-Y7) compared to IMDB (6.53 for TV-Y and 6.73 for TV-Y7), indicating that the TMDB user base may rate children's content more favorably than the IMDB audience.

Analysis:

1.Higher Volume Doesn't Always Translate to Higher Ratings:

- PG and G content has a high certification count, but the average ratings are lower. This might be due to the simple or less engaging nature of content targeted at younger audiences, which tends to be rated lower compared to content with more mature or complex themes.

2.Mature Content Is Valued More by Audiences:

- Even though R, PG-13, TV-MA, and TV-14 have lower certification counts, they generally receive higher ratings. This indicates that audiences tend to value more complex, mature, or adult-themed content, as it may offer deeper narratives, more sophisticated characters, or higher production quality.

3.TV Ratings Show More Favorable Trends on TMDB:

- TV content, especially for younger viewers (TV-Y, TV-G, and TV-Y7), tends to be rated higher on TMDB than on IMDB. This suggests that different user bases might have slightly different preferences or scoring behaviors. TMDB users might be more lenient or appreciate family-oriented or children's content more than IMDB users.

4.Underperformance of Family-Oriented Content:

- The PG and G categories, despite having the highest certification count (597 combined), show consistently lower ratings. This implies that while family-friendly content is widely produced, it may not be meeting audience expectations or generating as much enthusiasm as content aimed at older demographics.

Overall Analysis- Number 9 & 10

Recommendations:

1.Improve the Quality of Family-Oriented Content:

- Given the high volume of PG and G-rated content, there is a significant opportunity to improve the quality of family-friendly movies and shows. Conducting audience surveys to understand the reasons and references.

2.Invest More in Mature-Themed Content:

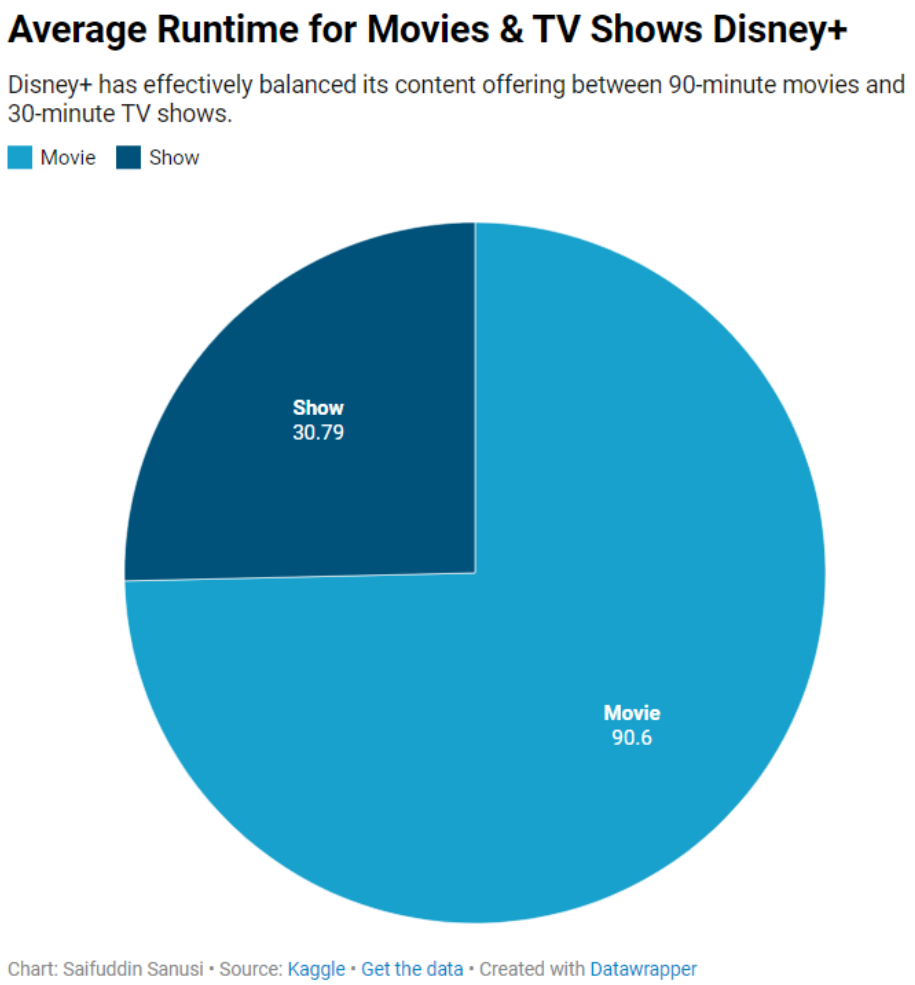
- Since PG-13, R, and TV-MA content receive higher ratings, it might be beneficial for content producers to invest in creating more mature-themed content, especially considering the positive audience reception.

3.Leverage TMDB's Favorable Audience for Children's Programming:

- The higher ratings for children's content (e.g., TV-Y, TV-Y7, TV-G) on TMDB suggest that there is a more appreciative audience for this type of content on that platform. Consider targeting promotions, marketing, or content launches for children's programming to platforms like TMDB where such content might be better received.

Data Analysis – Average Runtime of Movies & TV Shows

```sql
-- 11. Calculating the average runtime of movies and TV shows separately
SELECT
'Movies' AS content_type,
ROUND(AVG(runtime),2) AS avg_runtime_min
FROM titles
WHERE t_type = 'Movie'
UNION ALL
SELECT
'Show' AS content_type,
ROUND(AVG(runtime),2) AS avg_runtime_min
FROM titles
WHERE t_type = 'Show';
```



Conclusion:

• The average runtime of movies on Disney+ is 90.60 minutes, while the average runtime of TV shows is 30.79 minutes.

• There is a clear distinction in runtimes between movies and TV shows, with movies having a longer average duration, and TV shows having a more typical episodic length.

Data Analysis – Average Runtime of Movies & TV Shows

Analysis:

1.Standard Movie Runtime:

- The average runtime of 90.60 minutes for Disney+ movies falls within the typical range for feature films, which usually run between 90 to 120 minutes. This indicates that Disney+ offers a wide selection of full-length feature films, aligning with industry standards for movie duration.

2.Target Audience Consideration:

- The shorter episode length of TV shows suggests that Disney+ continues to cater to its core audience, including families and younger viewers, who may prefer shorter, easily digestible content.
- On the other hand, the 90-minute movie average caters well to family movie nights and offers a range of content from animated classics to live-action films, maintaining the platform's appeal to a broad audience and that allows for binge-watching or quick viewing, which has become popular among streaming audiences.

Recommendations:

1.Expand Variety in TV Show Length:

- While the 30-minute format works well for younger audiences and family viewing, Disney+ could consider offering more variety in show runtimes by introducing some longer-form series (45-60 minute episodes) aimed at older viewers or teens. This would allow the platform to cater to a wider demographic, especially for those interested in more in-depth, serialized storytelling.

2.Leverage the Binge-Watching Trend:

- With the 30-minute average runtime for shows, Disney+ is well-positioned to capitalize on the binge-watching trend. By creating and promoting more serialized content with multiple shorter episodes, the platform could further encourage back-to-back viewing habits. Shows with ongoing story arcs that encourage viewers to keep watching would align well with this model.

3.Continue Producing Family-Friendly Movies:

- The 90-minute average runtime is perfect for family movie nights, which is a key market for Disney+. The platform should continue producing and curating family-friendly films that fit into this 90-minute format while also offering some longer, more epic films for audiences who enjoy extended movie experiences.

Data Analysis – Titles & Directors of Movies and TV Shows (after 2023)

```sql
-- 12. Finding the titles and  directors of movies released on or after 2023
SELECT DISTINCT t.t_title, c.name AS director
FROM titles AS t
JOIN credits AS c
ON t.id = c.id
WHERE t.t_type = 'Movie'
AND t.release_year >= 2023
AND c.role = 'director'
ORDER BY release_year DESC;
```

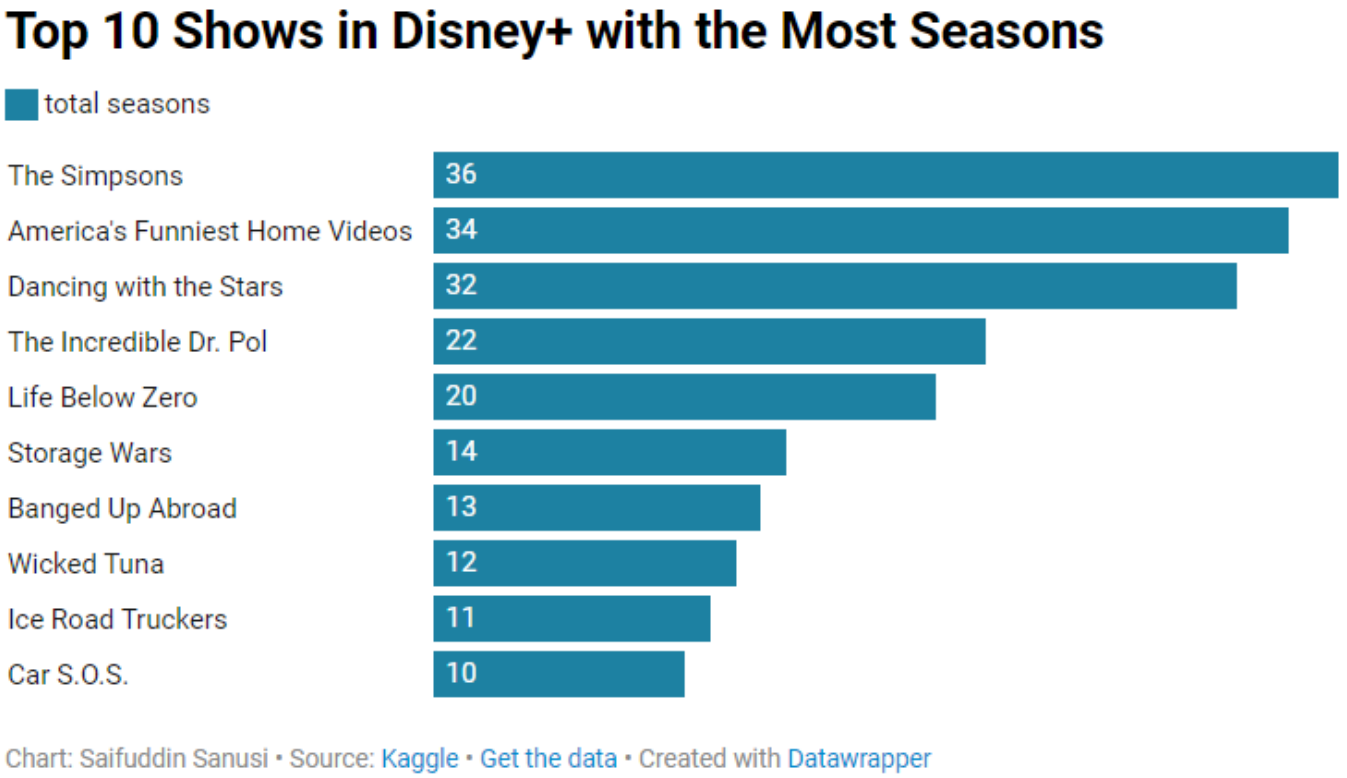
Titles & Directors of Movies and TV Shows (after 2023)

Directors	Titles
Suzy Milstein	Schoolhouse Rock! 50th Anniversary Singalong
Luiza Trigo	Whispers: An Elephant's Tale
Juliana Vonlanten	Baby Sharks

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

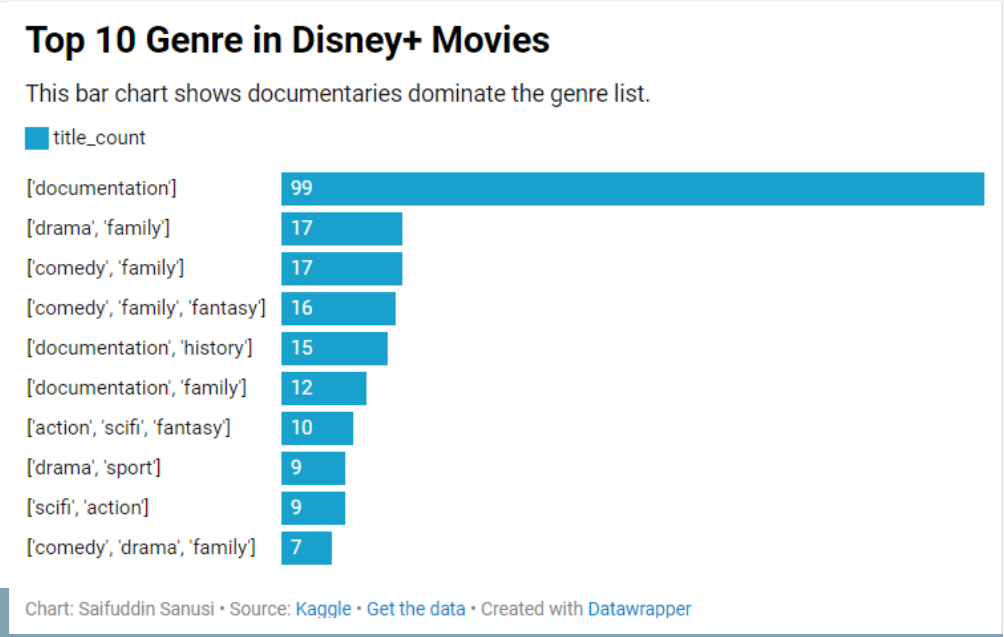
Data Analysis – Top 10 Most Seasons Disney+ TV Shows

```sql
-- 13. Which shows on Disney+ have the most seasons?
SELECT t_title,
SUM(seasons) AS total_seasons
FROM titles
WHERE t_type = 'Show'
GROUP BY title
ORDER BY total_seasons DESC
LIMIT 10;
```



Data Analysis – Top 10 Genre in Disney+ Movies

```sql
-- 14. Which genres had the most movies?
SELECT genres,
COUNT(*) AS title_count
FROM titles
WHERE t_type = 'Movie'
GROUP BY genres
ORDER BY title_count DESC
LIMIT 10;
```

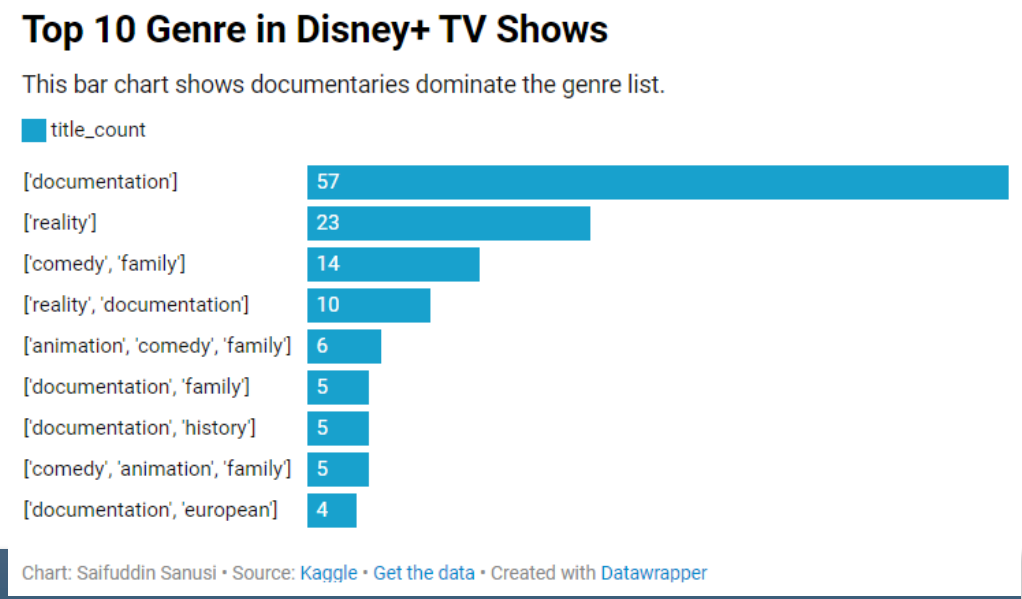


Conclusion

- **Documentation Dominance**: "Documentation" genres dominate the movie category with 99 titles, which is far ahead of the next most popular genres. This suggests a strong emphasis on educational, factual content.
- **Family-Oriented Content**: Genres like "drama", "comedy", and "family" frequently appear together, reflecting Disney's reputation for family-friendly content. Titles such as "drama, family" and "comedy, family" rank high with 17 titles each.
- **Fantasy and Sci-Fi**: There's notable representation of genres blending "fantasy" and "sci-fi" with action and family genres, aligning with Disney's focus on franchises like Marvel, Star Wars, and other fantasy properties.

Data Analysis – Top 10 Genre in Disney+ TV Shows

```sql
-- 15. Which genres had the most shows?
SELECT genres,
COUNT(*) AS title_count
FROM titles
WHERE t_type = 'Show'
GROUP BY genres
ORDER BY title_count DESC
LIMIT 10;
```



Conclusion:

1.Documentation Leads: Similar to movies, "documentation" is the leading genre for TV shows, with 57 titles.

2.Reality Shows Growing: "Reality" TV is the second most prominent genre (23 titles), suggesting that Disney+ is tapping into the demand for unscripted content that appeals to viewers' tastes for real-life storytelling.

3.Family Content: "Comedy, family" is a significant genre for TV shows with 14 titles, mirroring the focus on family-friendly entertainment in the movie segment.

4.Animation as a Key TV Genre: Animated shows, especially those combined with comedy and family themes (e.g., "animation, comedy, family"), are highly represented, with Disney leveraging its long-standing expertise in animation.

Data Analysis – Top 10 Genre in Disney+ Movies and TV Shows

Recommendations

1.Leverage Documentation Strengths: Disney+ should continue investing in its strong documentation catalog. To maintain and grow this audience, the platform could diversify within the genre by producing or acquiring documentaries in trending or emerging topics (e.g., environmental issues, modern technology, space exploration).

2.Expand Family-Oriented Content: Given the high number of family-oriented titles across both movies and TV shows, Disney+ should continue to produce content that appeals to children, parents, and families alike. Expanding narratives with contemporary themes while retaining family-friendly values will ensure Disney+ continues to be the go-to platform for this demographic.

3.Enhance Reality TV Offerings: Since reality TV is gaining traction (particularly in TV shows with 23 titles), Disney+ should consider further investments in high-quality reality shows. Focusing on niche areas like travel, cooking, or competition-based formats could attract broader audiences.

4.Capitalize on Sci-Fi and Fantasy: With genres like "scifi, action, fantasy" being popular in both movies and TV shows, Disney+ could benefit from investing more in speculative fiction, expanding franchises like Star Wars, Marvel, and other original properties that fall within these genres.

5.Continue to Innovate with Hybrid Genres: Disney+ is doing well by blending genres (e.g., "comedy, family, fantasy" or "animation, comedy, family"). Further exploration of hybrid genres and unique combinations could create standout content that appeals to multiple audience segments simultaneously.

6.Invest in Animated TV Shows: Since animation combined with family and comedy themes is popular, Disney+ should continue investing in animated series that appeal to both younger audiences and nostalgic adults, perhaps reviving older properties or creating new ones.

Data Analysis –Movies with high IMDB &TMDB popularity scores

```sql
-- 16. Titles and Directors of movies with high IMDB scores (>7.5) and high TMDB popularity scores (>80)
SELECT t.t_title,
c.name AS director
FROM titles AS t
JOIN credits AS c
ON t.id = c.id
WHERE t.t_type = 'Movie'
AND t.imdb_score > 7.5
AND t.tmdb_popularity > 80
AND c.role = 'director';
```

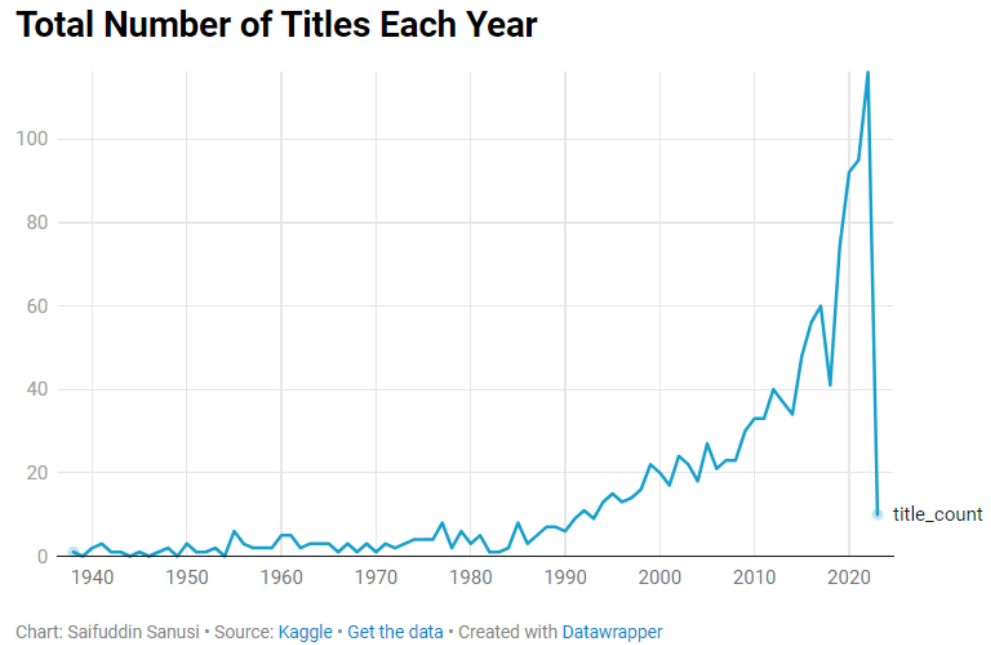
Titles and Directors of movies with high IMDB scores (>7.5) and high TMDB popularity scores (>80)

Titles	Directors
The Empire Strikes Back	Broni Likomanov
Star Wars	Nikos Dayandas
The Lion King	Dean DeBlois
WALLÂ·E	Chris Sanders
Avengers: Infinity War	John Schultz
Newsies	John Pasquin
Avengers: Endgame	Luke Cormican
Hamilton	Don MacKinnon
Folklore: The Long Pond Studio Sessions	Victor Cook
Coco	Lisa Schaffer
BTS: Permission to Dance on Stage - LA	Steve Loter
Hubble's Cosmic Journey	Lisa Feit
Just Love and a Thousand Songs	Renan Ozturk

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Data Analysis - Number Movies and Shows by Years

```sql
-- 17. What were the total number of titles for each year?
SELECT release_year,
COUNT(*) AS title_count
FROM titles
GROUP BY release_year
ORDER BY release_year DESC;
```



Conclusion :

1.Steady Growth in Title Count:

- There is a steady increase in the number of titles being released from the late 1990s to 2022, with the most notable surge starting around 2009.
- The biggest jump is from 2019 (74 titles) to 2020 (92 titles), then increasing further to 116 titles in 2022.

2.Consistent Growth (2000-2022):

- From 1999 onward, the number of titles steadily increases. Particularly, the 2010s (2010-2019) show a consistent upward trend, with no drop in the number of releases year-to-year, culminating in the peak in 2022.

3.Notable Decline in 2023:

- There is a sharp drop in releases in 2023 (10 titles), which might indicate incomplete data or early in the year, or it could reflect a major event affecting content production, such as post-pandemic disruptions.

Data Analysis – Actors Who Have Starred in The Most Highly Rated Movies or Shows

```sql
-- 18. Actors who have starred in the most highly rated movies or shows
SELECT c.name AS actor,
COUNT(*) AS num_highly_rated_titles
FROM credits AS c
JOIN titles AS t
ON c.id = t.id
WHERE c.role = 'actor'
AND (t.t_type = 'Movie' OR t.t_type = 'Show')
AND t.imdb_score > 8.0
AND t.tmdb_score > 8.0
GROUP BY c.name
ORDER BY num_highly_rated_titles DESC;
```

ctors Starred in the Most Highly Rated Disney+ Movies or TV Shows

oth IMBD Score and TMDB Score are above 8.

Page 1 of 3 ›

Name (actor) ▲	Number (num_highly_rated_titles)
Alec Guinness	1
Anthony Daniels	1
Billy Dee Williams	1
Bob Anderson	1
Brigitte Kahn	1
Bruce Boa	1
Burnell Tucker	1
Carrie Fisher	1
Christopher Malcolm	1
Clive Revill	1
David Prowse	1
Denis Lawson	1
Des Webb	1
Frank Oz	1
Harrison Ford	1
Ian Liston	1
Jack McKenzie	1

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Actors Starred in the Most Highly Rated Disney+ Movies or TV Shows

Both IMBD Score and TMDB Score are above 8.

Page 2 of 3

Name (actor)	Number (num_highly_rated_titles)
Jack Purvis	1
Jason Wingreen	1
Jeremy Bulloch	1
Jerry Harte	1
John Cannon	1
John Dicks	1
John Hollis	1
John Morton	1
John Ratzenberger	1
Julian Glover	1
Kathryn Mullen	1
Kenneth Colley	1
Kenny Baker	1
Lightning Bear	1
Marjorie Eaton	1
Mark Capri	1
Mark Hamill	1

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper

Actors Starred in the Most Highly Rated Disney+ Movies or TV Shows

Both IMBD Score and TMDB Score are above 8.

Page 3 of 3

Name (actor)	Number (num_highly_rated_titles)
Mark Jones	1
Martin Dew	1
Michael Culver	1
Michael Sheard	1
Milton Johns	1
Norman Chancer	1
Norwich Duff	1
Oliver Maguire	1
Peter Mayhew	1
Ray Hassett	1
Richard Bonehill	1
Richard Oldfield	1
Robin Scobey	1

Table: Saifuddin Sanusi • Source: Kaggle • Get the data • Created with Datawrapper