Amazon Prime SQL project Report

1.What are the top 5 most frequent genres on Amazon Prime?

```
SELECT
 genres, COUNT(*) AS total_count
FROM
 amazon_data
GROUP BY genres
ORDER BY total_count DESC
LIMIT 5;
```

genres	total_count
drama	84158
comedy	47273
thriller	40797
action	37025
romance	32989

2.What percentage of total content does each genre represent?

```
SELECT
   genres,
   COUNT(*) AS genres_count,
   ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM amazon_data), 2) AS
percentage_of_total
FROM
   amazon_data
GROUP BY
   genres
ORDER BY
   percentage_of_total DESC;
```

Result Grid | Filter Rows:

genres	genres_count	percentage_of_total
drama	84158	21.47
comedy	47273	12.06
thriller	40797	10.41
action	37025	9.45
romance	32989	8.42

Amazon Prime SQL project Report

3.How many titles are available per content type (e.g., Movie vs TV Show)?

```
SELECT
   type, COUNT(*) AS total_counts
FROM
   amazon_data
GROUP BY type;
```

Result Grid | Filter

type	total_counts
SHOW	24989
MOVIE	366939

4. Which genre has the highest average IMDb rating?

```
SELECT
   genres, AVG(imdb_score) AS highest_avg_rating
FROM
   amazon_data
GROUP BY genres
ORDER BY highest_avg_rating
LIMIT 1;
```

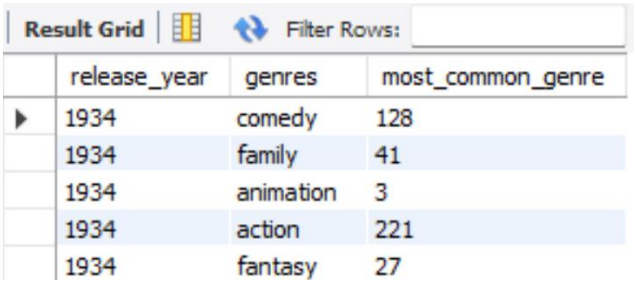
Result Grid | Filter Rows:

genres	highest_avg_rating
horror	4.997557881382789

5.Which genres are most commonly produced per year?

```
SELECT
   release_year, genres, COUNT(*) AS most_common_genre
FROM
   amazon_data
GROUP BY genres , release_year;
```

Amazon Prime SQL project Report

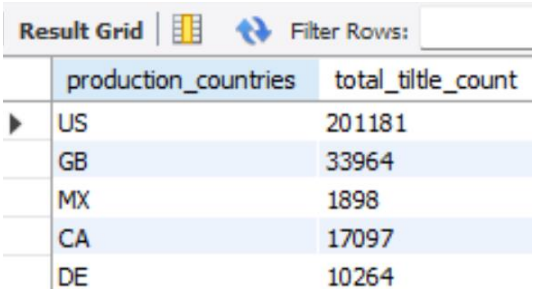


Result Grid | Filter Rows:

release_year	genres	most_common_genre
1934	comedy	128
1934	family	41
1934	animation	3
1934	action	221
1934	fantasy	27

6.How many titles are available per country or region?

```
SELECT
   production_countries, COUNT(title) AS total_tiltle_count
FROM
   amazon_data
GROUP BY production_countries;
```

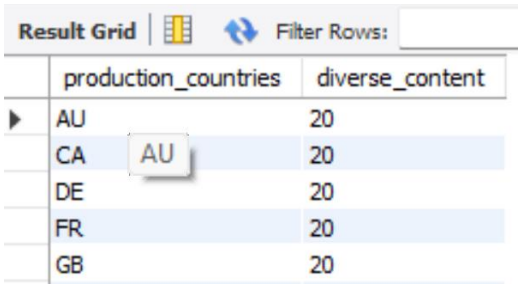


Result Grid | Filter Rows:

production_countries	total_tiltle_count
US	201181
GB	33964
MX	1898
CA	17097
DE	10264

7. Which region has the most diverse content genres?

```
SELECT
   production_countries,
   COUNT(DISTINCT genres) AS diverse_content
FROM
   amazon_data
GROUP BY production_countries
HAVING diverse_content = 20
ORDER BY diverse_content DESC
;
```



Result Grid | Filter Rows:

production_countries	diverse_content
AU	20
CA	20
DE	20
FR	20
GB	20

Amazon Prime SQL project Report

8. How many exclusive titles are available only in one region?

```
SELECT  COUNT(*)
FROM
   (SELECT
      title
   FROM
      amazon_data
   GROUP BY title
   HAVING COUNT(DISTINCT production_countries) = 1) AS exclusive_titles;
```

	COUNT(*)
▶	8800

9. Which genres are most popular per region?

```
WITH Rank AS (
   SELECT
      production_countries,
      genres,
      COUNT(genres) AS genre_count,
      RANK() OVER (
         PARTITION BY production_countries
         ORDER BY COUNT(genres) DESC
      ) AS rn
   FROM amazon_data
   GROUP BY production_countries, genres
)
SELECT
   production_countries,
   genres,
   genre_count
FROM Rank
WHERE rn = 1
ORDER BY genre_count DESC;
```

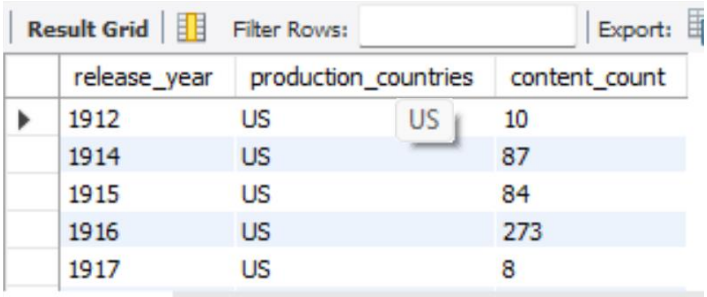
Result Grid | Filter Rows: | Exp

	production_countries	genres	genre_count
▶	US	drama	41502
	GB	drama	8312
	IN	drama	8110
	FR	IN opean	3797
	CA	drama	3582

Amazon Prime SQL project Report

10. Compare the top 3 regions by content count over the years.

```
WITH top_regions AS (
   SELECT production_countries
   FROM amazon_data
   GROUP BY production_countries
   ORDER BY COUNT(*) DESC
   LIMIT 3
)
SELECT
   release_year,
   production_countries,
   COUNT(*) AS content_count
FROM amazon_data
WHERE production_countries IN (
   SELECT production_countries FROM top_regions
)
GROUP BY release_year, production_countries
ORDER BY release_year, production_countries;
```



Result Grid | Filter Rows: | Export:

release_year	production_countries	content_count
1912	US	10
1914	US	87
1915	US	84
1916	US	273
1917	US	8

11. How has the number of new titles changed each year?

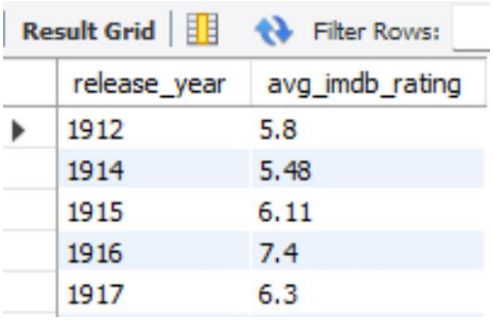
```
SELECT
   release_year,
   COUNT(*) AS total_new_titles
FROM amazon_data
GROUP BY release_year
ORDER BY release_year;
```

Amazon Prime SQL project Report

release_year	total_new_titles
1912	10
1914	87
1915	84
1916	273
1917	8

12. What is the trend of average IMDb ratings over the years?

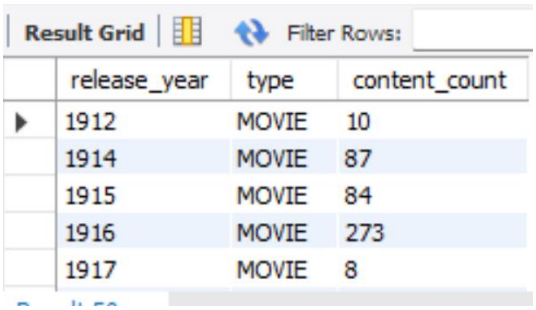
```
SELECT
   release_year,
   ROUND(AVG(imdb_score), 2) AS avg_imdb_rating
FROM amazon_data
WHERE imdb_score IS NOT NULL
GROUP BY release_year
ORDER BY release_year;
```



release_year	avg_imdb_rating
1912	5.8
1914	5.48
1915	6.11
1916	7.4
1917	6.3

13. Are more shows or movies being added recently?

```
SELECT
   release_year,
   type,
   COUNT(*) AS content_count
FROM amazon_data
WHERE release_year IS NOT NULL AND type IS NOT NULL
GROUP BY release_year, type
ORDER BY release_year, type;
```

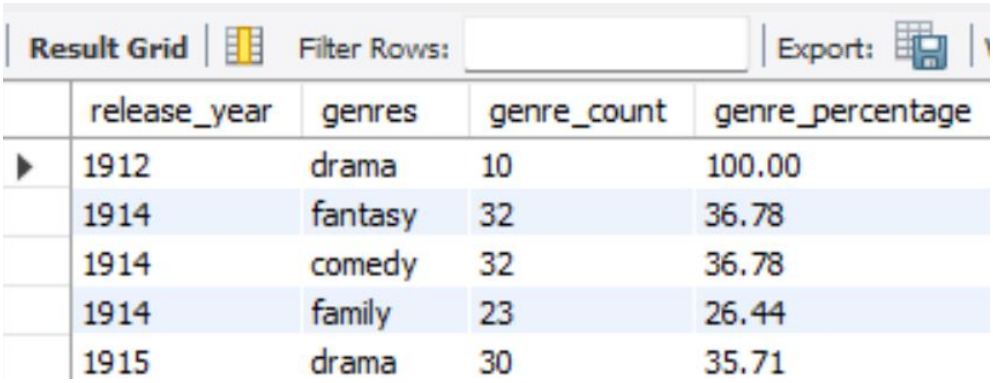


release_year	type	content_count
1912	MOVIE	10
1914	MOVIE	87
1915	MOVIE	84
1916	MOVIE	273
1917	MOVIE	8

Amazon Prime SQL project Report

14. How have genre proportions changed over time?

```
WITH genre_counts AS (
    SELECT
        release_year,
        genres,
        COUNT(*) AS genre_count
    FROM amazon_data
    WHERE release_year IS NOT NULL AND genres IS NOT NULL
    GROUP BY release_year, genres
),
total_per_year AS (
    SELECT
        release_year,
        COUNT(*) AS total_titles
    FROM amazon_data
    WHERE release_year IS NOT NULL AND genres IS NOT NULL
    GROUP BY release_year
)
SELECT
    g.release_year,
    g.genres,
    g.genre_count,
    ROUND((g.genre_count * 100.0) / t.total_titles, 2) AS genre_percentage
FROM genre_counts g
JOIN total_per_year t
ON g.release_year = t.release_year
ORDER BY g.release_year, genre_percentage DESC;
```

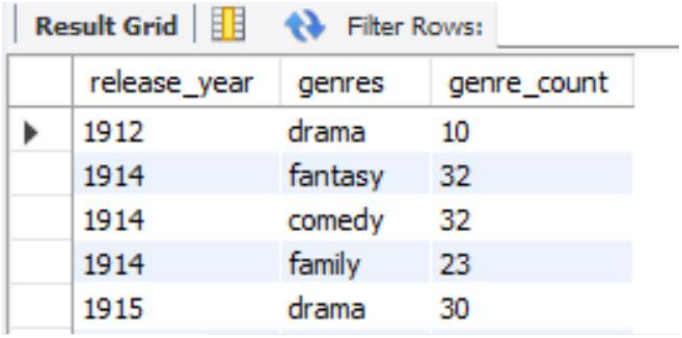

Result Grid | Filter Rows: | Export:

release_year	genres	genre_count	genre_percentage
1912	drama	10	100.00
1914	fantasy	32	36.78
1914	comedy	32	36.78
1914	family	23	26.44
1915	drama	30	35.71

Amazon Prime SQL project Report

15. What's the yearly count of content per genre?

```
SELECT
   release_year,
   genres,
   COUNT(*) AS genre_count
FROM amazon_data
WHERE release_year IS NOT NULL AND genres IS NOT NULL
GROUP BY release_year, genres
ORDER BY release_year, genre_count DESC;
```



Result Grid | Filter Rows:

release_year	genres	genre_count
1912	drama	10
1914	fantasy	32
1914	comedy	32
1914	family	23
1915	drama	30