**SQL capstone project**

**Description:**

**Using:**

Sakila database schema is used for this analysis

**Case study:**

Help Movie on rent company to analysis the demand and stock up the inventory accordingly.

**Analysis:**

**Task 1: Display the full names of actors available in database.**

I used the Concat() function to join the first\_name and last\_name

**Task 2: Management wants to know if there are ant names of the actors appearing frequently**

1. **Display the number of time each first name appears in the data base**
2. **What is the count of actors that have unique first names in the database? Display name of all these actors.**

So I selected first\_name and count() function to count the first name of the actors and group by the first\_name having count(first\_name)=1

**Task 3: The management is interested to analyze the similarity in the last names of the actors.**

1. **Display the number of times each last name appears in the database.**
2. **Display all unique last names in the database.**

So I selected last\_name and count() function to count the last name of the actors and group by the last\_name having count(last\_name)=1

**Task 4: The management wants to analyze the movies based on their ratings to determine if they are suitable for kids or some parental assistance is required. Perform the following tasks to perform the required analysis.**

1. **Display the list of records for the movies with the rating "R". (The movies with the rating "R" are not suitable for audience under 17 years of age).**
2. **Display the list of records for the movies that are not rated "R"**
3. **Display the list of records for the movies that are suitable for audience below 13 years of age.**

Selecting title and rating from film and used where to see rating=r

Selecting title and rating from film and used where to see rating!=r

Selecting title and rating from film and used where to see rating and used or to see rating=g or pg

**Task 5: The board members want to understand the replacement cost of a movie copy(disc - DVD/Blue Ray). The replacement cost refers to the amount charged to the customer if the movie disc is not returned or is returned in a damaged state.**

1. **Display the list of records for the movies where the replacement cost is up to $11.**
2. **Display the list of records for the movies where the replacement cost is between $11 and $20.**
3. **Display the list of records for the all movies in descending order of their replacement costs.**

Selected title and replacement cost from film where replacement cost is lesser then equal to 11 using comparison operator

Selected title and replacement cost from film where replacement cost is between 11 and 20 used between function

Selected title and replacement cost from film used order by replacement cost in descending order

**Task 6: Display the names of the top 3 movies with the greatest number of actors.**

Joined 2 table film table and film actor on film id table using inner join in count() counted the actor id group by film title order by actor count as descending and limited to 5

**Task 7: 'Music of Queen' and 'Kris Kristofferson have seen an unlikely resurgence. As an unintended consequence, films starting with the letters K and Q have also soared in popularity. Display the titles of the movies starting with the letters K and o**

Used like function to title starts with k and % sign says any number of characters used or logical operator for another title that starts with q and % sign for any number of characters

**Task 8: The film 'Agent Truman has been a great success. Display the names of all actors who appeared in this film**

Joined 3 table actor table, film actor table on actor id and film table on film id using inner join in concat() first and last name as full name where title is “Agent Truman”

**Task 9. Sales have been lagging among young families, so the management wants to promote family movies. Identify all the movies categorized as family films.**

Selected film title and category name from film table and film category on inner join film id and inner join category table on category id where category id is 8

**Task 10: The management wants to observe the rental rates and rental frequencies(Number of time the movie disc is rented).**

1. **Display the maximum, minimum, and average rental rates of movies based on their ratings. The output must be sorted in descending order of the average rental rates.**
2. **Display the movies in descending order of their rental frequencies, so the management can maintain more copies of those movies.**

Selected ranking and used aggregate function of avg, min and max of rental rate column group by rating order by avg rental rate in descending

Selected everything from film order by rental duration descending

**Task 11: In how many film categories, the difference between the average film replacement cost ((disc - DVD/Blue Ray) and the average film rental rate is greater than $15?**

1. **Display the list of all film categories identified above, along with the corresponding average film replacement cost and average film rental rate.**

Selected category name used avg function for film replacement cost and film rental rate after that avg replacement cost – avg rental rate as avg diff used inner join on 3 tables film table, film category table and category table on film id and category id respectively group by category name and having avg diff greater than 15

**Task 12: Display the film categories in which the number of movies is greater than 70.**

Selected category name and count of film title as film count from joining 3 tables film table, film category table and category on film id and category id respectively grouped by category name having film count grater then 70