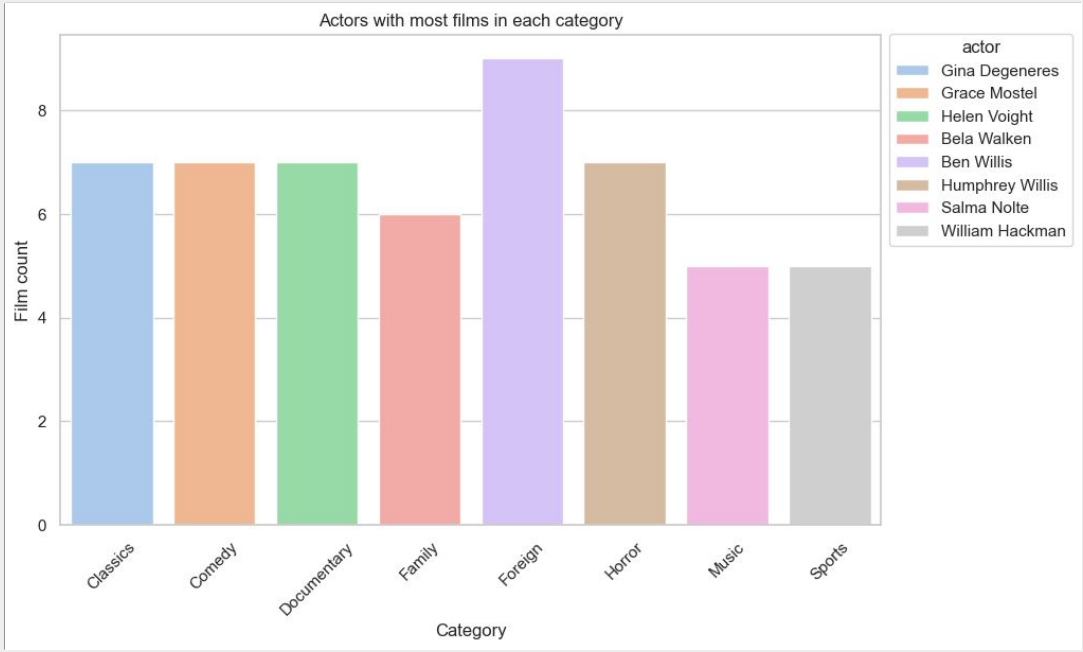
Which actors have appeared in the most films in each category, and how many films did they appear in?



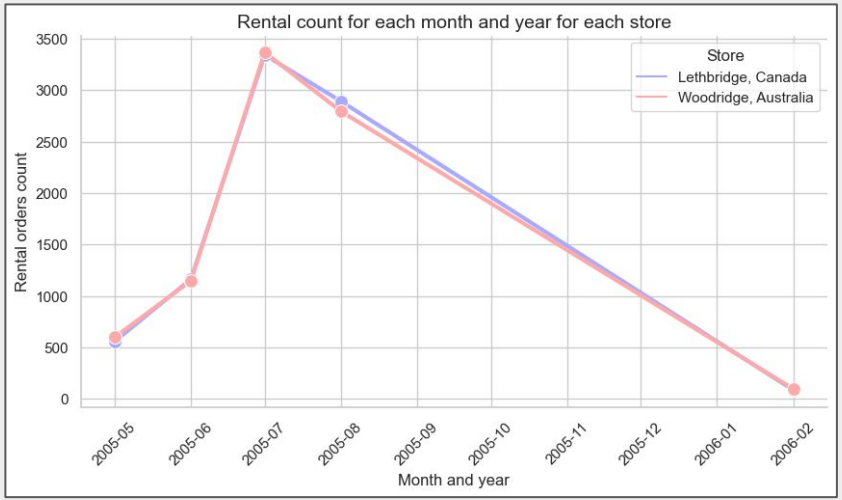
This query ranks actors by the number of films they appeared in each film category and returns the top actor for each category.

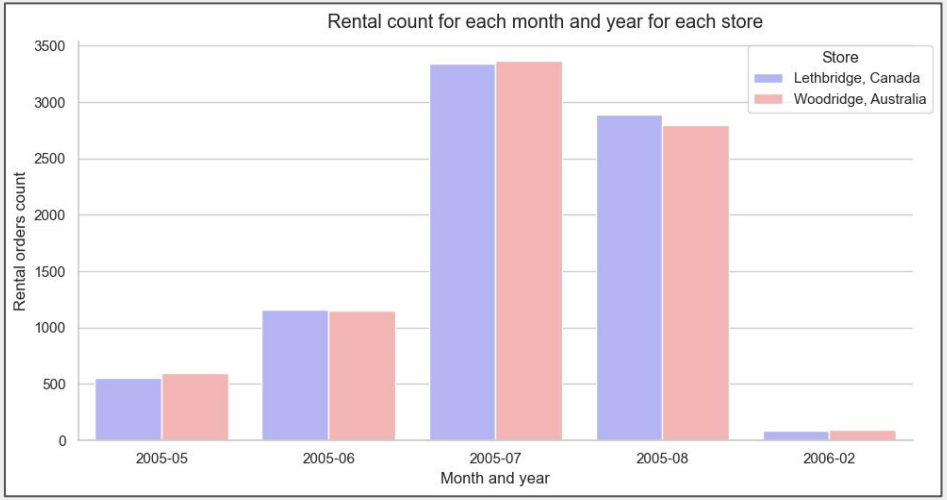
This can provide valuable information about which actors are most popular in each category and can help inform casting decisions for future films.

Kindly refer to Query 1 in the .txt file

To answer this question, I joined the `actor`, `film_actor`, `film`, `category` and `film_category`tables together. I then grouped the result by category and actor, and used the `COUNT` aggregation function to count the number of films each actor appeared in within each category. Finally, I used a window function to rank the actors within each category by the number of films they appeared in. For visualisation I took only the top actor in each category and created a bar chart with a bar per each category showing the film count and the colour representing the actor.

How do the two stores compare in the count of rentals by month?



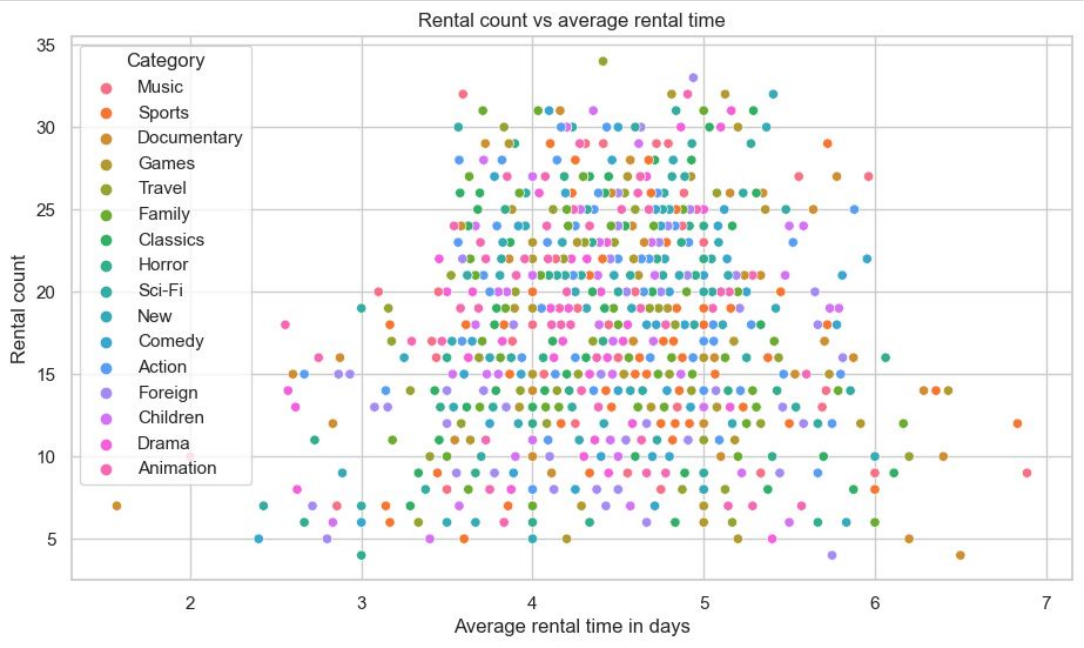


This graph shows the rental count for each month and year for each store. We can see that both stores function almost equally in terms of rental counts throughout the year. However, an interesting observation is that the peak season for rental activity (with the top month being August) in Canada is summer while in Australia it's winter.

Kindly refer to Query 2 in the .txt file

I chose this question provided by Udacity because I thought it would be interesting to observe changes over time. So for visualisation, I first tried to create a time series graph but there is not enough data to make a continuous timeline, so a grouped bar chart is a better option. I added the store location for easier identification (instead of referring to them by the id number).

Which movie categories have the highest average rental duration and rental count?

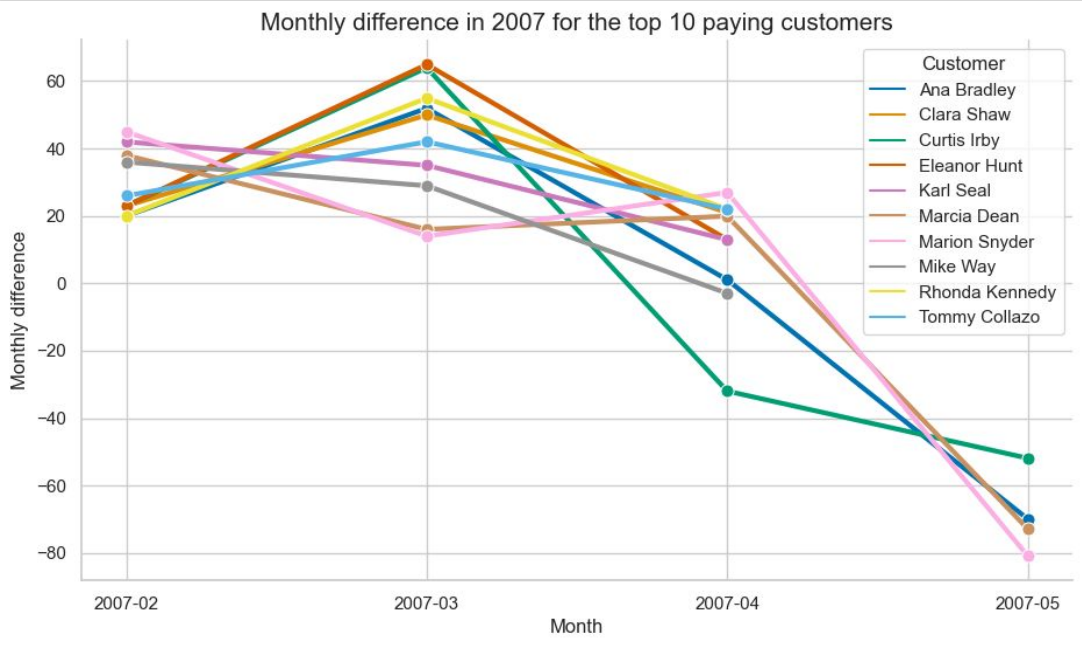


Based on this scatter plot, we can gain insights into the popularity of movies based on rental count and average rental rank. We can also determine if there are any correlations between the rental count and average rental days. Additionally, we can analyse which movie categories are more or less popular based on the same factors.

Kindly refer to Query 3 in the .txt file

To answer this question, I joined the `film`, `rental`, and `inventory` tables together and calculate the rental duration of each filmin days by subtracting the rental date from the return date, and summing up the duration for all the rentals of that film. Next, I used a CTE to calculate the average rental duration for each film. Finally, I used a window function to rank the films by their average rental duration. I also added the `film_category` and `category` tables to represent the category of each film with colour on the scatter plot.

What is the difference across monthly payments during 2007 for each of the top 10 paying customers?



Looking at the results, we can see that some customers have a consistent increase in monthly payments, while others have inconsistent patterns. The information can be used to identify patterns in customer payment behavior and to better understand customer payment trends.

Kindly refer to Query 4 in the .txt file

This is Question 3 from the Question Set #2 from the course. This query result provides information on the monthly payment difference for the top 10 customers who have made the highest payments overall in descending order. The table shows the customer's full name, payment month, and monthly payment difference. The payment month is truncated to show only the month and year, and the monthly difference is the difference between the current month's payment and the previous month's payment.