**MOVIE RENTAL**

**SAKILA DVD RENTAL STORE**

**Documentary Report by Vishal Gautam**

**Overview of the Project**

This project offers a comprehensive exploration of the Sakila Dataset, a simulated DVD rental store database comprising various tables, including film, film category, actor, customer, rental, payment, inventory, city, country and payment The endeavor primarily leverages the power of SQL (Structured Query Language) to unlock valuable insights from this dataset. Using a question-and-answer (Q&A) approach, it delves into SQL syntax across basic and advanced levels, addressing a wide array of relevant questions. Additionally, the data is brought to life through interactive visualizations created with Power BI and Done Exploratory Data Analysis (EDA) with the help of SQL and Excel, enhancing our understanding of the dataset to other audience as well.

**Objective of the Project**

The primary objective of this project is to conduct a comprehensive data analysis of the Sakila Dataset, which emulates a DVD rental store’s operation. Using SQL as our primary analytical tool, we aim to achieve the following key objectives:

1. Explore the Dataset: Thoroughly examine the Sakila Dataset and its various tables, including film, film category, customer, rental, payment, and inventory, different locations, store and staff to gain a deep understanding of the data's structure and content.

2. Extract Meaningful Insights: Utilize SQL queries to extract meaningful insights, patterns, and trends from the dataset. This includes answering a range of questions, from basic to advanced, to showcase the versatility and power of SQL in data analysis.

3. Enhance Data Visualization: Employ tools like PowerBI to transform the extracted insights into interactive and visually appealing data representations. By doing so, we aim to make the dataset more accessible and comprehensible for a wider audience.

4. Demonstrate SQL Proficiency: Showcase the practical application of SQL in real-world data analysis, emphasizing its relevance and utility in addressing complex data-related questions.

5. Promote Knowledge Sharing: Share the project's findings, methodologies, and interactive visualizations to foster knowledge sharing and provide a valuable resource for individuals interested in data analysis, SQL, and data visualization.

6. Exploratory Data Analysis: We will conduct an in-depth Exploratory Data Analysis (EDA) to derive valuable insights by meticulously analyzing the data, utilizing the capabilities of SQL and Excel. This comprehensive approach aims to unveil patterns, trends, and key findings within the dataset, enabling informed decision-making and a deeper understanding of the data's underlying dynamics

**Significance**

The significance of this project is multi-faceted. First and foremost, it serves as a compelling demonstration of the practical application of SQL in data analysis. By harnessing the robust capabilities of SQL, the project unearths meaningful insights from the dataset, thus emphasizing the vital role of structured query language in extracting knowledge from complex data sets.

Moreover, the project holds significant value as an educational resource. It empowers individuals interested in data analysis and SQL by offering a hands-on example of how to navigate, query, and derive insights from a comprehensive dataset. As such, it serves as a stepping stone for those seeking to expand their SQL proficiency and data analysis skills.

Beyond its educational merits, the project showcases a dynamic dimension through the utilization of interactive visualizations with PowerBI and EDA. This additional layer of significance makes the dataset more accessible and engaging, catering to a broader audience. It enhances comprehension and knowledge retention by transforming raw data into captivating visual representations.

In summary, the project's significance not only lies in its ability to showcase SQL's practical application but also in its role as an educational resource and its capacity to make the dataset more relatable and informative through interactive visualizations, thereby fostering a deeper and more profound understanding of the dataset's intricacies.

**Data Dictionary**

**film \_text**

• film \_id: Unique identifier of each film in film\_text table

• Description: Details about the film.

**category**

• category \_id: Unique identifier for different categories of films.

• name: Name of each film

**film \_category**

• category \_id: Primary Key

• film\_id: Foreign Key referencing film\_id of the film table

**film**

• language \_id: Primary key

• rating: Different film ratings

• rental \_rate: Rate of rental of each film

• rental \_duration: Renting time period.

**language**

• language\_id: Unique identifier of language table

• name: Consists different languages

**inventory**

• inventory\_id: Unique identifier of this table

• store\_id: Foreign key referencing store\_id of the store table

• film \_id: Foreign key referencing film\_id of the film table

**rental**

• rental \_id: Primary key of rental table

• rental\_date: Date the film rented

• return \_date: Return date of the film rented

• inventory id: Foreign key referencing inventory\_id of the inventory table

**payment**

• amount: Payment received from customer

• customer\_id: Foreign key referencing the customer\_id of customer table

• payment \_id: Primary key

• rental \_id: Foreign key referencing the rental\_id of rental table

• staff\_id: Foreign key referencing the staff\_id of the staff table

• payment\_date: Information about the payment date

**customer**

• active: Identifies whether the customer is active or not

• address\_id: Foreign key referencing the address\_id of address table

• customer ID: Primary key of customer table

• email ID: Email id of customer

• first name: First name of customer

• last name: last name of customer

• store\_id: Foreign key referencing store\_id of the store table.

**address**

• address \_id: Primary key of address table

• city \_id: Foreign key referencing city\_id of city table

**city**

• city\_id: Primary key of city table

• country\_id: Foreign key referencing country\_id of the country table

• city: name of city

**country**

• country\_id: Primary key

• country: name of the country

**store**

• address\_id: Foreign key referencing address\_id of address table

• Store\_id: Primary key of the store table

• manager\_staff\_id: Foreign key referencing staff\_id of staff table

**staff**

• address\_id: Foreign key referencing address\_id of address table

• staff\_id: Primary key of staff table

• stored\_id: Foreign key referencing store\_id of store table

**film \_actor**

• actor\_id: Foreign key referencing actor\_id of actor table

• film\_id: Foreign key referencing film\_id of film table

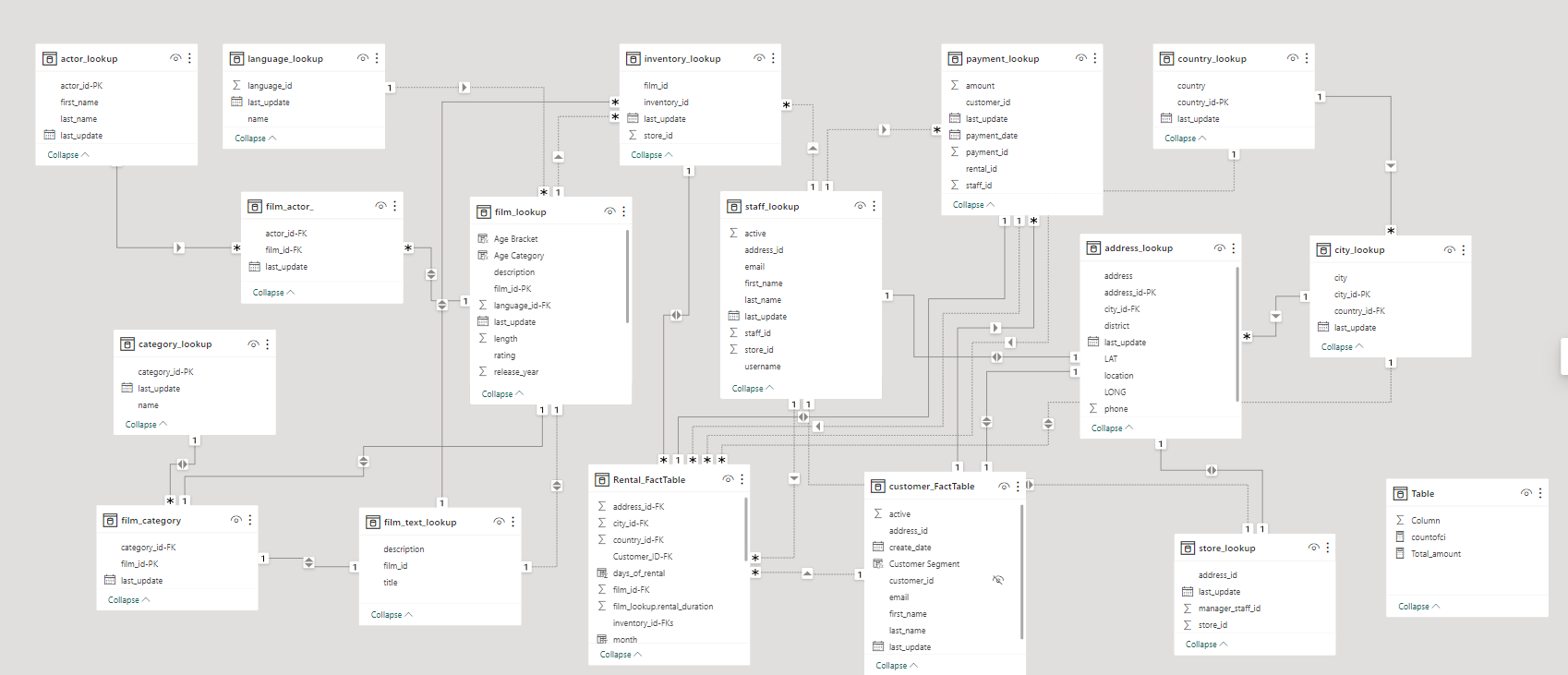
**actor**

• actor\_id: Primary key of actor table

• first name: First name of actor

• last name: Last name of actor

**E-R Diagram**

****

**Steps To Connect Data**

**Step** **1. Data Acquisition from GitHub:**

Acquired essential datasets from the Sakila DVD rental store repository on GitHub, including rentals, customers, payments, films, film categories, inventory, store, staff, city, and country. These datasets are integral for comprehensive insights into store revenue, staff performance, customer analysis, and film categories.

**Step 2. Data Transformation:**

Employed PowerBI for data transformation, addressing necessary adjustments such as date format standardization, decimal number rounding, and the extraction of latitude and longitude points from separate columns within the address tables using the Power Query's extraction capabilities.

**Step 3. MECE Approach in Excel**

Implemented a Mutually Exclusive and Collectively Exhaustive (MECE) breakdown after a thorough analysis of the dataset, categorizing the data into distinct types for precise analysis. Including the Customer Behavior analysis, Film Inventory Analysis, Saff Performance Analysis and the Store operation analysis.

**Step 4. Connecting with Analytical Tools**

Established seamless connections between the dataset and various analytical tools, including Excel, MySQL Workbench for SQL, and Power BI, ensuring efficient data integration and processing.

**Step 5. EDA Problem Statement Resolution in SQL**

Utilized SQL to tackle Exploratory Data Analysis (EDA) problem statements. Employed SQL's robust capabilities for data exploration and analysis, enabling the derivation of valuable insights and solutions.

**Step 6. Problem Statement Resolution in Power BI**

Leveraged the powerful features of Power BI to address specific problem statements, facilitating data visualization, exploration, and analysis. Effectively derived insights and provided solutions through this tool.

**Acknowledgment**

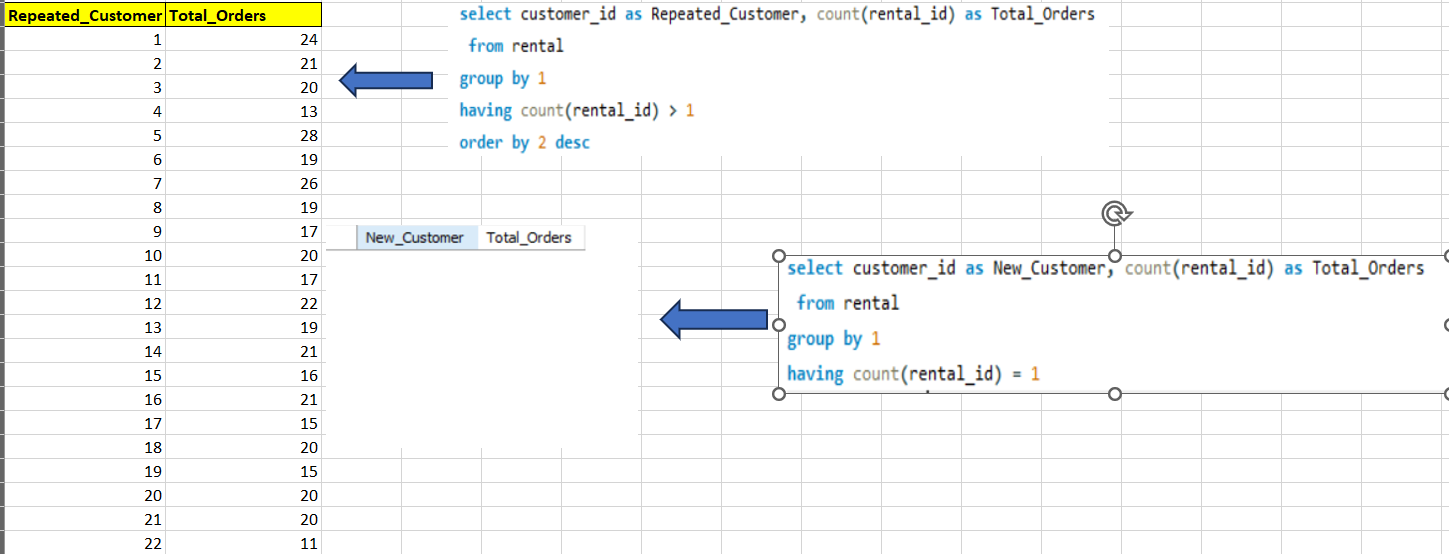
I would like to extend my sincere gratitude to my mentor, **Mr. Shiraj Ahmed**, for his invaluable guidance and support throughout the course of this project. His expertise and mentorship were instrumental in shaping the project's success, and I am deeply appreciative of his contributions.

I would also like to acknowledge the crucial role played by my batch manager, **Mr. Ammar Ahsan**, who provided valuable insights and oversight during the project's development. His assistance and coordination were essential in ensuring the project's smooth progression.

Lastly, as an individual, **Vishal Gautam**, representing AccioJob Pvt. Ltd., I am grateful for the opportunity to undertake this project and showcase the practical application of data analysis. I take pride in the collaborative efforts and dedication invested in bringing this project to fruition.

**Exploratory Data Analysis Problems.**

**1. What are the purchasing pattern of new customers versus Repeating Customers?**

****

The problem at hand revolves around discerning the relationship between new and repeat customers based on rental behavior, with a focus on customers who have rented films. However, the available data does not include any instances of new customers, defined as those who have rented films only once. The absence of such customers in the dataset implies that we have access to information solely pertaining to repeat customers, i.e., those who have rented films more than once.

In order to effectively investigate and illustrate the relationship between these two customer categories, it is imperative to obtain relevant data for new customers as well. This will enable a comprehensive analysis that compares the rental patterns, preferences, and engagement of both new and repeat customers. Without this vital dataset, drawing any meaningful conclusions about the behavior or interaction between these customer groups would be unfeasible.

In summary, the absence of data for new customers hinders our ability to uncover and elucidate the dynamics between new and repeat customers, emphasizing the necessity of acquiring pertinent information to make informed decisions and analyses in this context.

**2.Which films have the highest rental rates and are most in demand?**

The analysis of the data reveals several noteworthy trends in film categories and their revenue generation. Notably, the "Sports" category stands out by boasting both the highest revenue and the maximum number of rentals compared to other categories. This indicates a high demand for sports-themed films, making it a prime focus for future content selection and marketing strategies.

Additionally, several other categories, namely "Sci-Fi," "Animation," "Drama," "Action," and "Comedy," also contribute significantly to revenue generation. These categories exhibit strong potential for revenue growth and could be prioritized in content curation and promotion efforts.

One consistent aspect across all categories is the uniform highest rental rate of $4.99. This uniformity suggests that the pricing strategy has been effectively standardized across different genres.

In summary, the "Sports" category stands as a dominant revenue generator, while the mentioned film categories hold promise for further revenue enhancement. This data-driven insight underscores the importance of catering to customer preferences and optimizing content offerings to maximize revenue in the dynamic film rental market.

**3. Are there correlations between staff performance and customer satisfaction?**

There's no relevant data to solve this problem statement, we need the data for the staff performance and customer rating i.e customer satisfaction in order to solve this problem.

**I've done an alternate analysis for this problem statement.**

The analysis of rental return behaviors provides valuable insights into customer preferences and the management of film inventory. One key finding is that a significant number of rentals are marked as "Returned late," signifying a high demand for these films. However, this delay in return may result in film unavailability in the inventory at specific times, potentially leading to customer dissatisfaction and missed revenue opportunities.

Conversely, there are rentals marked as "Returned earlier," indicating that these customers may not have been as interested in the particular film, potentially stemming from a lack of engagement. This insight underscores the importance of understanding customer preferences and optimizing the selection of films to align with their interests.

A smaller segment of rentals falls under the category of "Returned on time," indicating not only timely returns but also a positive viewing experience. These customers are likely to have enjoyed the film, suggesting an opportunity to cater to their preferences further by offering more films of their favorite categories.

In conclusion, the data suggests the need for a balanced approach. While addressing the demand for late returns and optimizing inventory management, it is equally crucial to focus on enhancing the selection of films to capture customer interest, as well as nurturing and expanding the customer base who return films on time. This multifaceted strategy can lead to improved customer satisfaction and revenue growth in the film rental business.

**4.Are there seasonal trends in customer behavior across different locations?**

The data-driven insight that film rentals in both India and China show a strong preference for renting films during the autumn season is a valuable finding for optimizing the film rental business. In India, where the maximum number of rentals occur during this season, there is a clear opportunity to leverage this trend to maximize revenue.

To capitalize on this seasonal trend, it is advisable to focus more on the autumn season by expanding the film inventory. Prioritizing the inclusion of new collections of movies from highly profitable categories such as sports, Sci-Fi, Drama, and Comedy can prove to be a lucrative strategy. This approach aligns with customer demand and can lead to increased revenue generation.

The consistent preference for autumn rentals, as indicated by the graph across both India and China, further underscores the importance of directing resources and efforts towards this particular season. By effectively catering to customer preferences during the autumn months, the film rental business can not only enhance customer satisfaction but also realize its revenue potential to the fullest.

**5.** **Are certain language films more popular among specific customer segments?**

The data reveals a critical aspect of the film rental business: it exclusively features English-language films. This linguistic focus implies that the customer base primarily consists of individuals comfortable with or interested in English-language content. While this can provide a consistent and steady stream of rentals, it also highlights the potential to diversify offerings and capture a broader audience by incorporating films in other languages.

Furthermore, the data highlights that the majority of rentals are marked as "Active," signifying ongoing and continuous revenue generation. The presence of a robust base of active rentals is promising for the business's financial stability.

Given the strong preference for English films and the predominance of active rentals, a prudent strategy would involve maintaining and enhancing the inventory of English-language films while also considering the introduction of films from other languages to potentially attract a more diverse clientele.

In conclusion, the data suggests that the film rental business should continue to cater to its English-speaking customer base, capitalize on the active rental segment, and consider diversification to tap into additional market opportunities and potential sources of revenue.

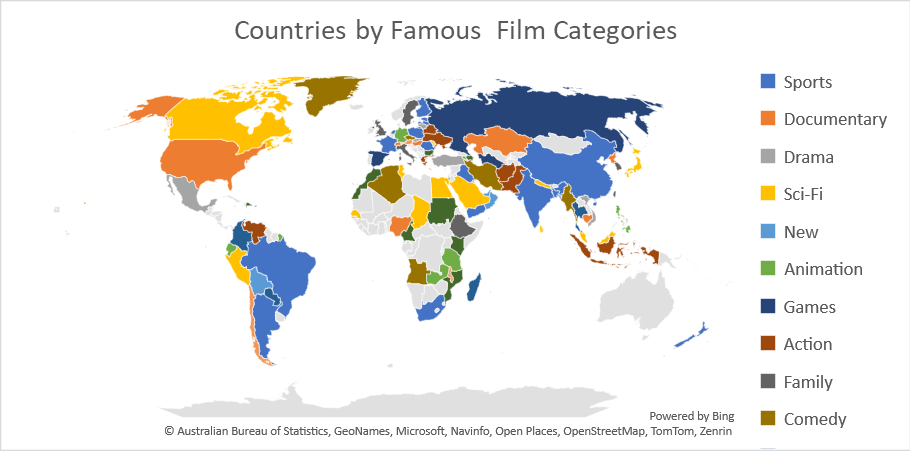
**6.** **How does customer loyalty impact sales revenue over time?**

The data analysis provides several critical insights into the film rental business. Firstly, it is apparent that revenue follows a seasonal pattern, with an upswing from February to July, followed by a gradual decline from August onwards. This trend suggests that external factors, such as summer vacations and holidays, could be influencing rental behavior during specific months, prompting an increase in movie rentals.

July, notably, emerges as the highest-earning month, characterized by both the highest total revenue and the greatest rental count. This can be attributed to a combination of seasonal factors and customer preferences that make it a peak period for the business.

Customer loyalty plays a pivotal role in the consistent growth of rental counts and total revenue. Return customers and loyal subscribers are likely contributing to this upward trajectory, as those who have positive experiences are more inclined to return for frequent rentals. Therefore, fostering and retaining customer loyalty remains a key strategy for sustained success in the film rental industry.

**7.** **Are certain film categories more popular in specific locations?**

****

The data analysis underscores an intriguing pattern in the film rental business, where each film category appears to have its own regional popularity. For instance, the "Sports" category enjoys fame in India, while the "Drama" category stands out in Turkey. This trend provides a valuable opportunity to tailor the film inventory and marketing strategies according to regional preferences, ultimately leading to revenue maximization.

To capitalize on this insight, it is advisable to focus on providing a more extensive selection of films within the categories that are particularly famous in different countries. For example, in India, where the "Sports" category reigns supreme, enhancing the collection of sports-related films can help cater to the local audience's preferences and boost revenue. Similarly, in Turkey, prioritizing the "Drama" category with fresh and captivating offerings can prove highly beneficial.

This approach not only aligns with customer demands but also showcases an understanding of the unique cinematic tastes in each region. It allows the film rental business to build a stronger connection with its audience and bolster its competitive edge by offering a curated selection that resonates with local sensibilities.

In summary, leveraging the popularity of specific film categories in different countries is a strategic move to maximize revenue. It demonstrates a keen understanding of regional variations in cinematic preferences, which, when catered to effectively, can lead to heightened customer engagement and financial success.

**8.** **How does the availability and knowledge of staff affect customer ratings?**

We Can't solve this problem statement due to insufficient data, we need the customer ratings in order to solve this question.

**I’ve done an Alternate Analysis.**

**Monthly Revenue Percent by Each Store**

The sales data for Store 1 and Store 2 provides important insights into their respective performance during the specified period.

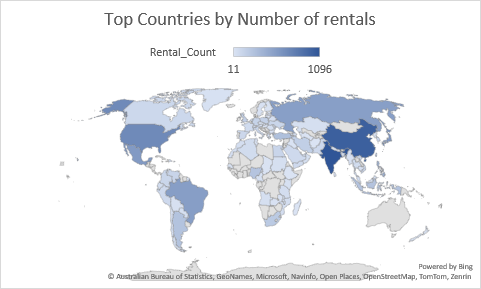
For Store 1, there is a noteworthy trend of substantial sales growth from May to July 2005, suggesting a period of potential expansion and heightened customer engagement. This upturn could be attributed to various factors such as seasonal promotions, increased foot traffic, or effective marketing strategies. However, the abrupt decline in sales in August 2005 is a point of concern. To sustain and build upon the growth achieved in the prior months, Store 1 should investigate the reasons behind this decline. It may be linked to factors like the end of a promotional period, a decrease in consumer demand, or inventory issues. Identifying the root causes will be crucial to maintaining a positive growth trajectory.

Store 2 also experienced growth from May to July 2005, which is a positive indicator of its performance. What's particularly promising is that the sales remained relatively consistent between July and August 2005. This suggests that Store 2 was able to maintain the customer engagement and revenue it had generated during the growth phase. It may have successfully retained customers or implemented strategies to offset any seasonal downturns. Store 2's ability to sustain sales at this level demonstrates stability and the potential for continued success. In conclusion, while both stores saw growth from May to July 2005, the contrasting performance in August for Store 1 and the consistency in sales for Store 2 highlight the importance of analyzing and addressing factors that can impact sales fluctuations in order to secure long-term growth and stability.

**9.** **How does the proximity of stores to customers impact rental frequency?**

We have the data to calculate the proximity between the customer location and store but there is no relation in that data. we have the location points for two stores only and we have approx. 600 customers from different location it has no relation between the user location and store location.

**Alternate Analysis for Top Countries by Number of Rentals.**



The data analysis clearly indicates that India stands out as the leading country in terms of film rentals, closely followed by China. The minimal difference in rental activity between these two countries signifies their significance in the film rental business. Consequently, it is prudent for the store to concentrate its efforts on these high-rental regions. This focus should involve tailoring inventory, marketing strategies, and customer engagement to cater to the preferences and demands of the Indian and Chinese markets. By doing so, the store can leverage this strong customer base and maximize its revenue potential in these key countries.

**10.** **Do specific film categories attract different age groups of customers?**

The analysis of film genre preferences across different age groups reveals several significant trends. Firstly, individuals above 18 years of age exhibit the highest preferences for most genres, indicating a broad spectrum of interests. However, it's important to note that there are exceptions, and genre preferences can vary among individuals.

Conversely, people under 7 years tend to have the lowest preferences for many genres, possibly due to the limited understanding and exposure of younger audiences.

The age group between 7 and 13 years demonstrates a middle ground in genre preferences, showcasing a developing taste that bridges the gap between the very young and adult audiences.

The genre of "Animation" garners popularity across all age groups, with a particular appeal to those under 7 and above 18 years. This underscores the universal appeal of animated content, which can cater to both the very young and adult sensibilities.

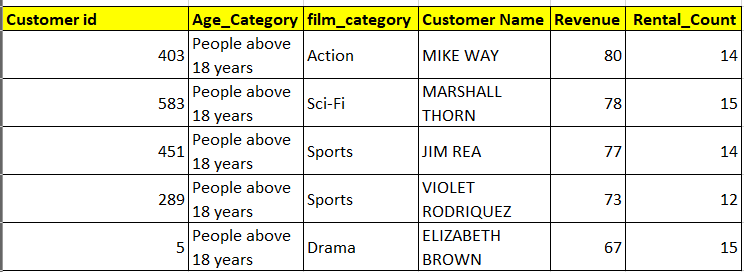
Notably, "Sports" content emerges as a strong preference for both individuals above 18 and under 13 years, indicating its broad appeal across these age groups. Understanding these age-related genre preferences can inform content selection and marketing strategies, helping tailor offerings to the varied tastes of different customer segments.

**11.** **What are the demographics and preferences of the highest-spending customers?**

The data analysis highlights a clear pattern in consumer behavior, where the top spending customers consistently belong to the same demographic group – individuals above 18 years of age. This age group appears to be the most financially engaged and valuable segment of the customer base.

Furthermore, the genre preferences of these top customers provide valuable insights. They favor genres like "Sports," "Action," "Sci-Fi," and "Drama" movies. This particular preference indicates that these customers are drawn to high-intensity and emotionally engaging content. "Sports" and "Action" genres offer excitement and adrenaline, "Sci-Fi" provides a sense of imagination and wonder, and "Drama" films offer intricate storytelling and emotional depth.

Understanding the alignment between top customer spending and genre preferences is pivotal for business strategy. By catering to the tastes of these valued customers and expanding the inventory in these popular genres, the business can not only maximize its revenue but also cultivate stronger customer loyalty, ultimately driving long-term success in the film rental industry.

****

**12.** **How does the availability of inventory impact customer satisfaction and repeat business?**

To solve this problem statement, we don't have the relevant data we need the data for customer satisfaction to solve this problem.

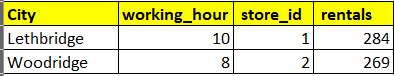
**Alternate Analysis: The most popular genre in each store in terms of revenue and frequency.**

For Store 1, the data reveals that the genre of "Drama" holds the top position in both revenue generation and rental frequency. This observation indicates a strong and consistent customer preference for drama films. The prominence of "Drama" in revenue underscores its profitability, as it likely attracts a large and diverse audience due to its wide range of emotional and thematic appeal. Its high rental frequency suggests that customers are not only renting drama films but also doing so repeatedly, which could be indicative of their loyalty and satisfaction.

In contrast, for Store 2, "Sports" emerges as the leading genre in terms of both revenue and rental frequency. This finding suggests that the "Sports" genre has a substantial following among Store 2's customer base. The consistent revenue and high rental frequency indicate a strong demand for sports-related content, likely driven by sports enthusiasts and fans. By focusing on this genre, Store 2 can strategically meet the preferences of its customers and potentially expand its customer base by curating a more extensive collection of sports-themed films.

These insights provide valuable direction for inventory management and marketing strategies in each store, catering to the distinct genre preferences of their respective customer bases.

**13.** **What are the busiest hours or days for each store location, and how does it impact staffing requirements?**

****

1. Store 1 experiences its busiest hour at 10 AM in Lethbridge, signifying a significant influx of customers during that time. This peak hour suggests that a substantial crowd converges on the store, possibly due to various factors like morning routines, work schedules, or specific shopping habits of the local population.

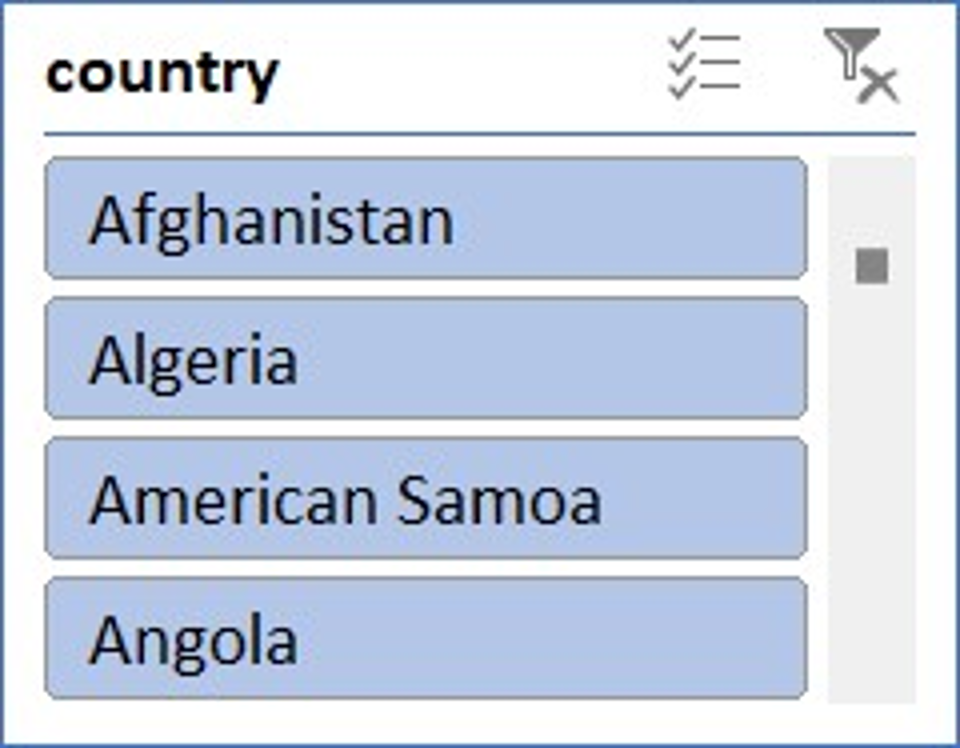
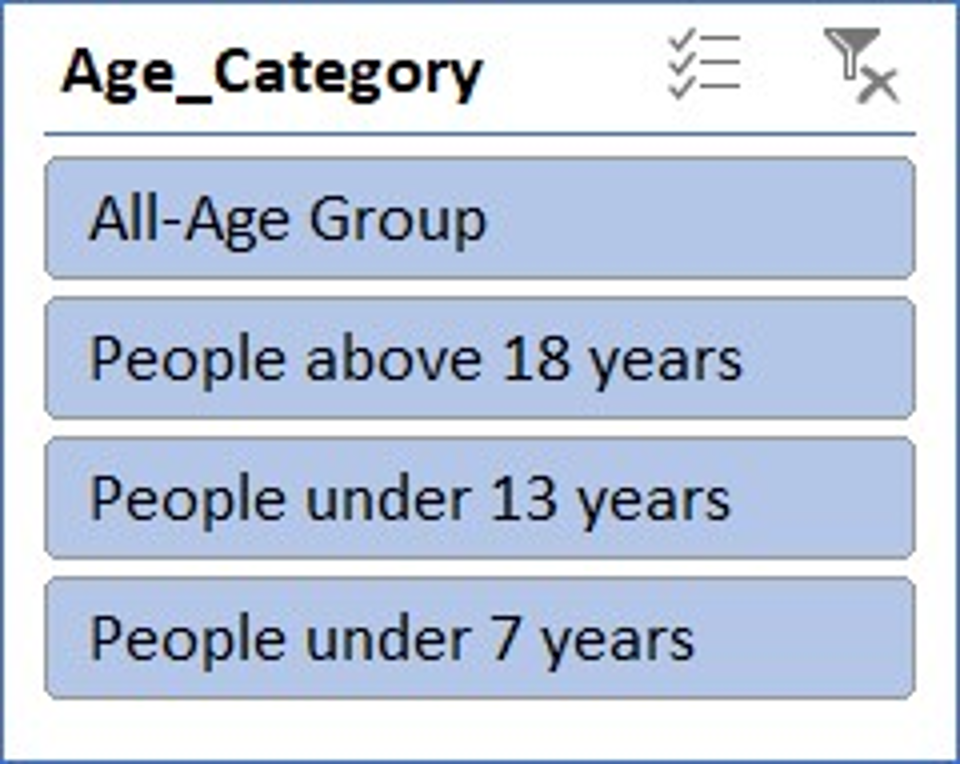
2. Store 2 witnesses its busiest hour at 8 AM in Woodbridge. This early rush indicates that customers in Woodbridge tend to start their day by visiting the store, possibly driven by convenience or early shopping needs.

3. To efficiently manage the increased customer traffic during these peak hours, it is imperative to have an adequate number of staff members on duty. Additional staff can help reduce waiting times, assist customers promptly, and ensure a seamless shopping experience, which is essential for customer satisfaction and retention.

4. Furthermore, assigning experienced and skilled employees during these peak hours can be advantageous. They are more adept at handling the rush, addressing customer inquiries, and maintaining order in the store. Their efficiency and expertise can contribute to a smoother and more productive operation during high-traffic periods, enhancing the overall customer experience.

In summary, recognizing and addressing the busiest hours in each store's specific location is crucial for optimizing staffing levels and ensuring that customers receive the best service, ultimately leading to improved customer satisfaction and business success.

**14.** **What are the cultural or demographic factors that influence customer preferences in different locations?**



The data analysis yields several valuable insights into film preferences across different demographics and countries. Firstly, certain film categories like "Music," "Games," "Sci-Fi," and "Documentary" emerge as universally popular, transcending geographical and age boundaries. These genres hold a broad appeal and consistently attract viewers, reflecting their versatility in catering to diverse tastes.

Secondly, the data underscores a significant preference for films targeting "People above 18 years," indicating that adults comprise the primary audience for movie rentals. Nevertheless, there is also notable engagement in the "All-Age Group" and "People under 13 years" categories, demonstrating the importance of content catering to a broader age range.

Thirdly, preferences exhibit regional variations. "Sports" enjoys popularity in China and India, aligning with the strong sports culture in these countries. In contrast, the United States leans towards "Games," reflecting the gaming culture prevalent there, while Brazil favors "Drama," possibly due to its rich cultural storytelling traditions. Understanding these diverse preferences is pivotal for content curation and marketing strategies. It enables businesses to target the right demographics and regions with content that resonates, fostering customer engagement and satisfaction. Recognizing the enduring appeal of certain genres, like "Music," "Games," "Sci-Fi," and "Documentary," underscores the significance of incorporating them into content offerings to cater to a broad and varied audience.

**15.** **How does the availability of films in different languages impact customer satisfaction and rental frequency?**

To solve this problem statement, we don't have the relevant data of customer satisfaction and also, we have data for only one film language i.e English. We need Data related customer satisfaction.

**Alternate Analysis: Top 10 customers according to total spending and Frequency.**

The data analysis reveals two prominent customers, Karl Seal and Eleanor Hunt, who stand out in terms of their total spendings and rental frequency.

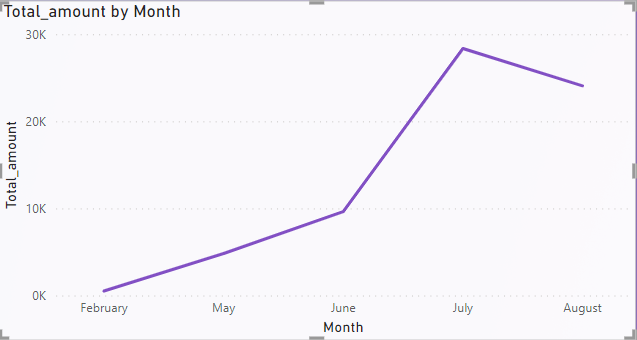
Karl Seal emerges as the top customer, excelling in both total spendings and rental frequency. His consistent patronage makes him a high-value customer who has demonstrated loyalty to the business. Recognizing Karl Seal's dedication, offering exclusive promotions, discounts, or loyalty rewards can further cement his relationship with the store. This can encourage him to continue frequenting the store, thereby ensuring sustained revenue generation.

Eleanor Hunt, while ranking second in total spendings, has rented films more frequently than Karl Seal. This signifies a strong engagement with the store and a substantial contribution to its success. Acknowledging customers like Eleanor is crucial, and providing them with tailored offers, personalized recommendations, or rental discounts can foster their continued support. This can also motivate them to explore more content options and, in turn, increase their total spendings.

In addition, extending offers and coupons to the top 10 customers, who likely represent a significant portion of the store's revenue, can not only retain their loyalty but also act as a marketing strategy by setting an example for other customers. This approach helps in maintaining customer satisfaction, increasing repeat business, and enhancing the store's overall financial performance.

**Power BI Questions**

**1.How does the sales revenue vary by month?**

****

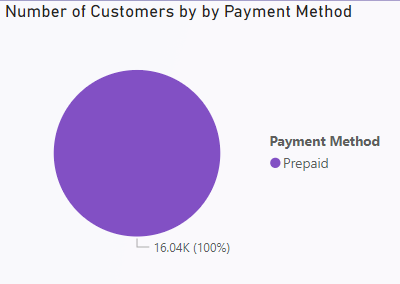
The data analysis indeed unveils significant patterns in the film rental business that are instrumental for understanding customer behavior and optimizing business strategies.

The seasonal trend in revenue is a key observation. It illustrates a clear cyclical pattern, with rental revenues surging from February to July and then gradually declining from August onward. This trend can be attributed to several factors. During the summer months, people often take vacations and have more leisure time, which may encourage them to engage in indoor activities like movie rentals. Holidays, family gatherings, and special occasions might also contribute to increased movie rentals during this period. The decline in August could be linked to the end of vacation season as people return to their routines.

The prominence of July as the highest-earning month is particularly noteworthy. It likely results from a confluence of factors, including summer vacations, holidays like Independence Day, and general leisure time, making it a peak period for movie rentals.

These insights offer valuable guidance for film rental businesses. By tailoring content offerings, marketing campaigns, and promotional activities to align with these seasonal trends, businesses can capitalize on peak periods and plan accordingly for slower months. Such strategies help maximize revenue and enhance customer engagement.

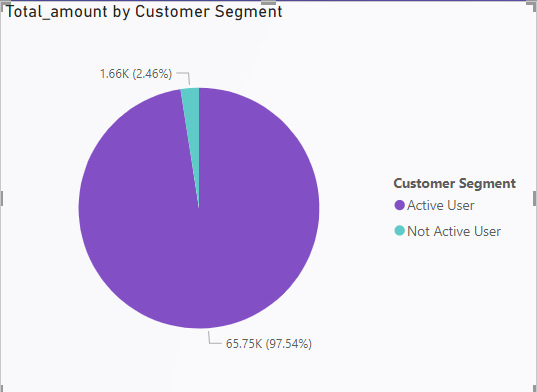
**2.** **What is the distribution of sales by payment method?**

****

The exclusive preference for prepaid payment methods among all customers is a significant finding that reflects a specific and uniform payment behavior within the dataset. Prepaid payment methods involve customers making upfront payments for products or services before they receive or utilize them. The absence of any customers opting for postpaid payment methods, which involve payment after the utilization of services, signifies a deliberate and consistent choice among the customer base.

This uniformity in payment preference may be attributed to several factors, including the convenience and control offered by prepaid methods, avoidance of potential late fees or billing complications associated with postpaid options, and a desire for financial transparency. Understanding this exclusive preference for prepaid payments allows businesses to streamline their financial processes, mitigate credit risk, and offer tailored prepaid plans or incentives to attract and retain customers. It also emphasizes the importance of accommodating the prevailing customer payment preferences to enhance satisfaction and loyalty in the business-customer relationship.

**3.Which customer segments generate the highest sales?**

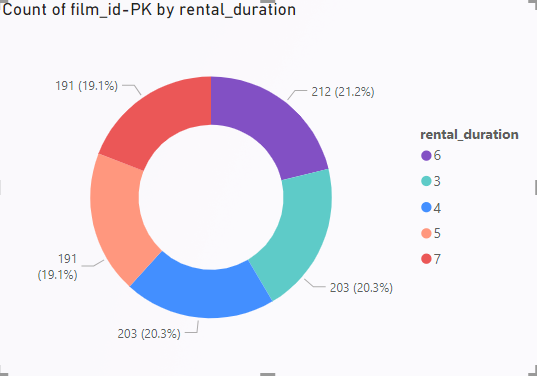
****

The data analysis reveals a stark contrast in sales between the Active and Inactive user segments. The Active user segment, with a substantial sales figure of 65.75K, emerges as the primary revenue driver. These users, characterized by their consistent and ongoing engagement with the business, are the cornerstone of its financial success. Their loyalty, frequent interactions, and likely higher average transaction values contribute significantly to the overall sales volume.

Conversely, the Inactive User segment, with sales amounting to 1.66K, demonstrates a notably lower level of engagement and contribution to the business's revenue. Inactive users may have made only sporadic or infrequent purchases, or their activity may have dwindled over time.

Understanding this disparity underscores the importance of nurturing and retaining the Active user segment, as they are the core revenue generators. It also highlights the opportunity to re-engage with Inactive users and potentially convert them into Active users through targeted marketing, promotions, or personalized incentives. Analyzing these user segments can guide strategic decision-making to optimize revenue growth and foster customer loyalty.

**4.** **What is the distribution of films by rental duration?**

****

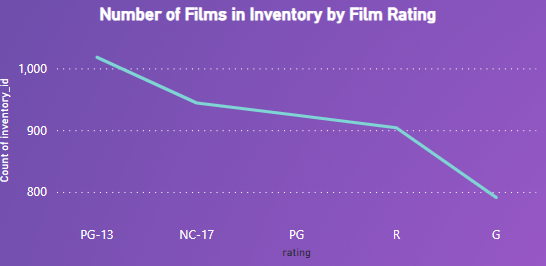
The distribution of films by rental duration provides valuable insights into customer preferences and behavior within the film rental business. Analyzing this data reveals several notable observations.

Firstly, the data indicates that the most common rental durations are 4 days and 3 days, each accounting for 20.3% of the distribution. This suggests that a significant portion of customers prefers shorter rental periods, possibly driven by a desire for convenience and cost-effectiveness. These customers may be looking to watch and return films quickly, thus favoring shorter rentals.

On the other hand, 6 days and 7 days rentals make up 21.2% and 19.1% of the distribution, respectively. This implies that a considerable number of customers opt for slightly longer rental periods, perhaps to accommodate extended viewing or to align with their schedules, which is reflected in the 6 and 7-day options.

Overall, the balanced distribution of rental durations indicates the importance of offering a variety of rental periods to cater to diverse customer preferences. By understanding these preferences, the film rental business can optimize its inventory management, pricing strategies, and customer service to ensure customer satisfaction and ultimately enhance revenue generation.

**5.** **How does the inventory vary by film rating?**

****

The distribution of film ratings, as indicated by the count of inventory IDs, offers valuable insights into the composition of the film inventory within the business. Analyzing this data unveils several significant observations.

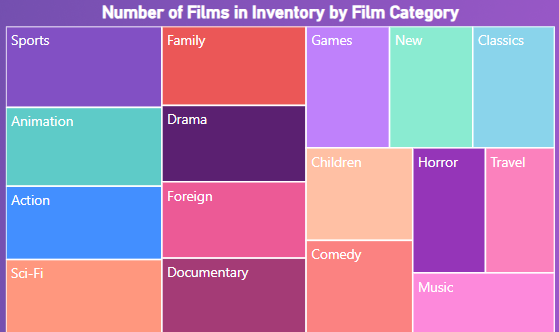
Firstly, the film rating PG-13 has the highest count of inventory IDs, with 1018. This suggests a significant prevalence of films with this rating in the inventory. PG-13 ratings are typically associated with content that is suitable for a broad audience, making them a versatile and popular choice for a wide range of viewers.

NC-17, with 944 inventory IDs, follows closely. While this rating is associated with more mature content, its presence indicates a commitment to offering a diverse selection that caters to different preferences and age groups.

PG, R, and G ratings have 924, 904, and 791 inventory IDs, respectively. This distribution suggests a relatively balanced representation of films across these categories, indicating an effort to provide content suitable for various audiences.

In summary, the data reflects a strategic approach to film selection, encompassing a range of ratings to appeal to a diverse customer base. It also highlights the importance of maintaining a balanced and varied film inventory to meet the preferences and requirements of a wide array of customers.

**6.** **What is the breakdown of film categories in the inventory?**

****

The inventory data reveals that the film categories "Sports," "Animation," "Drama," "Sci-Fi," "Family," "Action," and "Documentary" collectively have the highest number of films in stock. This diverse selection of film genres reflects a well-rounded inventory strategy aimed at catering to a broad spectrum of customer preferences.

1. "Sports" films appeal to enthusiasts of various athletic disciplines, attracting those interested in competitive and physically engaging content.

2. "Animation" stands as a perennial favorite, suitable for audiences of all ages, with its imaginative and visually captivating storytelling.

3. "Drama" offers compelling narratives and emotional depth, attracting viewers seeking thought-provoking and character-driven stories.

4. "Sci-Fi" appeals to those with a penchant for the imaginative and futuristic, offering concepts that transcend reality.

5. "Family" films provide wholesome and family-friendly entertainment, making them a go-to choice for parents and children.

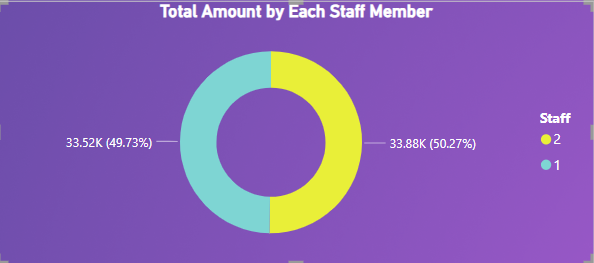
6. "Action" films deliver adrenaline-pumping excitement and are beloved by fans of thrilling, high-octane storytelling.

7. "Documentary" films appeal to the intellectually curious, offering informative and real-world narratives. This diversified film inventory ensures that customers with varying tastes can find content that resonates with them. It reflects a strategic approach to meet the entertainment needs of a diverse and wide-ranging customer base. The availability of these genres suggests a commitment to providing a well-rounded and inclusive cinematic experience, enhancing customer satisfaction, and optimizing revenue opportunities.

**7.** **What is the distribution of staff by employment duration?**

There’s No Data for the employment duration, to Solve this Problem statement we need the joining date and the leaving date of the employee.

**Alternate Analysis on Revenue Generated by each employee.**

****

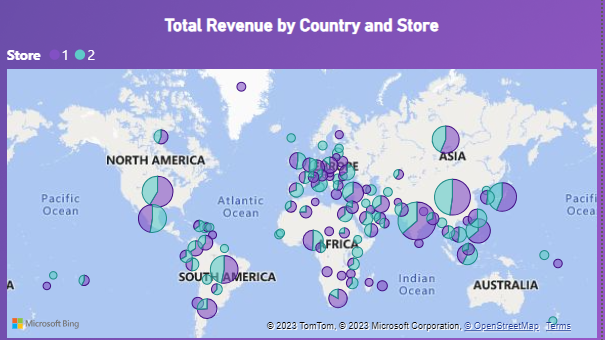
The revenue distribution between Staff 1 and Staff 2 provides a significant insight into the performance of these employees within the business. The data indicates that Staff 2 generated revenue totaling 33.88K, constituting 50.27% of the total revenue. In contrast, Staff 1 contributed 33.52K, accounting for 49.73% of the total revenue.

The remarkably close revenue figures between the two staff members suggest a level of parity in their sales performance. This balance indicates that both employees are instrumental in the overall financial success of the business. It underscores a consistent effort and contribution from both individuals, resulting in a nearly equal division of revenue.

This equitability can be indicative of the effectiveness of training and management, ensuring that staff members are performing at a similar level. However, it's also an opportunity for the business to further motivate and incentivize both staff members to continue their exceptional performance.

In summary, the balanced revenue contribution from Staff 1 and Staff 2 emphasizes the importance of nurturing and retaining these employees, as their combined efforts significantly impact the business's success. By offering recognition, rewards, or additional training, the business can foster their continued dedication, ultimately benefiting both the employees and the company.

**8.** **How does the store performance vary by location?**

****

The data analysis reveals that in countries such as India and China, Store 1 consistently outperforms Store 2 in terms of revenue. This trend extends beyond these two countries, encompassing a broader global perspective, including Europe, America, and other regions.

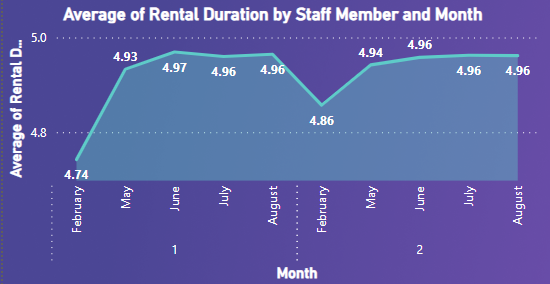
This consistent revenue superiority of Store 1 could be attributed to various factors, including location, product selection, marketing strategies, or customer service. Understanding the precise drivers of this success would require further in-depth analysis.

Store 1's sustained revenue leadership in multiple countries implies that it possesses a winning formula or business strategy that resonates with a diverse range of customers, across various cultures and markets. This presents an opportunity for Store 1 to explore expansion and replication of its successful model in new regions.

Conversely, Store 2 may benefit from an examination of the factors contributing to its comparatively lower revenue and could identify opportunities for improvement. This could involve assessing local preferences, streamlining operations, or implementing targeted marketing campaigns.

In conclusion, the data underscores the importance of tailoring business strategies to specific regions and markets, recognizing the uniqueness of customer preferences in each location. Store 1's consistent success indicates the potential for expanding its winning model, while Store 2 may seek areas of enhancement to enhance its revenue generation.

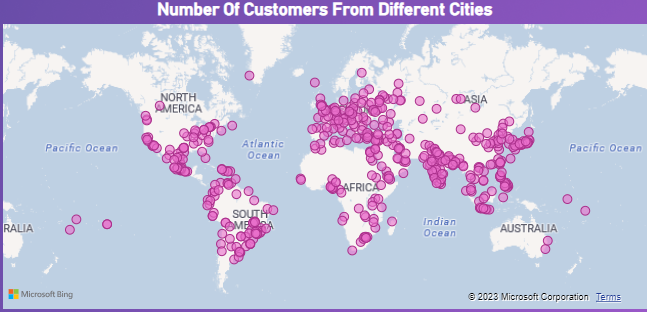
**9.** **What is the average rental duration by staff member?**



Staff 1 consistently maintains a fairly stable average rental duration from February to August, with a minor fluctuation of only 0.23 units over this period. In contrast, Staff 2 exhibits a similar trend but with even smaller variations, showing remarkable consistency with only a 0.1-unit difference between February and August.

This consistency might indicate that both staff members are handling rentals efficiently and consistently over time. Further analysis can explore underlying factors contributing to these patterns, such as any particular rental types or customer preferences each staff member may excel at, leading to these stable averages. Additionally, examining the impact of external factors like seasonal variations or changes in rental demand may provide further insights.

**10.** **What is the distribution of customers across different cities?**

****

The distribution of customers across cities, with a predominant presence in Aurora and London, while all other cities have only one customer, presents a compelling dataset for analysis.

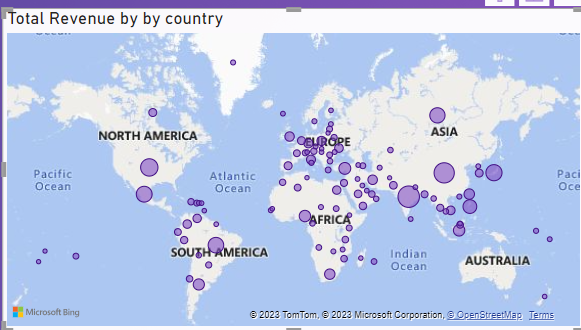
Firstly, the concentration of customers in Aurora and London suggests potential factors driving customer choices. These cities may offer unique attractions, higher population densities, or stronger business opportunities that draw more patrons.

Secondly, understanding why certain cities have only one customer is vital. It could be due to geographical constraints, limited market presence, or simply a lesser customer base in those areas. Identifying these underlying causes can help inform strategic decisions regarding resource allocation and expansion efforts.

Additionally, the data may reveal opportunities for growth in cities with single customers. Targeted marketing and service improvements can be explored to tap into these untapped markets.

In summary, this customer distribution data serves as a valuable starting point for geographical customer segmentation and targeted marketing strategies. Further exploration into the characteristics and preferences of customers in each city can yield actionable insights to optimize business operations and growth.

**11. How does the rental revenue vary by country?**

****

The data indicating that countries like India and China generate the maximum revenue, with the USA, Russia, Japan, and Mexico following with substantial but comparatively lower revenue, carries significant implications for global business analysis.

Firstly, the dominant revenue figures in India and China highlight the immense economic potential and market size of these two countries. This underscores the importance of businesses capitalizing on the vast consumer base and growth opportunities they offer.

The USA, Russia, Japan, and Mexico's substantial revenue generation positions them as significant players in the global market. This data underscores the need for companies to diversify their international operations and strategic focus to tap into these lucrative markets.

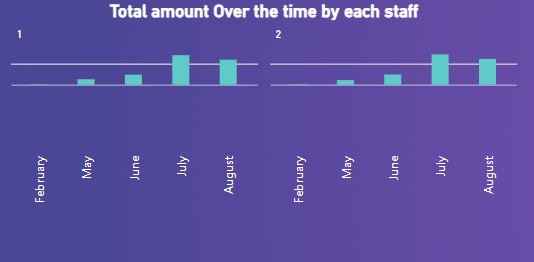
However, it's crucial to delve deeper into these numbers to understand the specific industries or sectors driving revenue in each country. Cultural, economic, and regulatory factors must also be considered when formulating market-entry strategies.

In summary, this data encourages businesses to prioritize India and China for their market expansion plans due to their dominant revenue figures, while acknowledging the potential of the USA, Russia, Japan, and Mexico as valuable secondary markets. A comprehensive market analysis is necessary to effectively navigate these diverse economic landscapes.

**12.** **Which locations have the highest and lowest customer ratings?**

We can’t solve this problem due to absence of relevant data, To solve this problem statement we want the data for Customer ratings.

**Alternate Analysis on Revenue Generated by each staff member over the time**

****

Analyzing the provided revenue data for Staff 1 and Staff 2 across multiple months reveals interesting insights.

For Staff 1, there is a clear upward trend in revenue from February to August. Revenue more than quadruples during this period, with the most significant growth observed in July. This suggests that Staff 1 has demonstrated exceptional performance, potentially through effective sales or customer engagement strategies. Further investigation could identify the factors contributing to this impressive revenue growth.

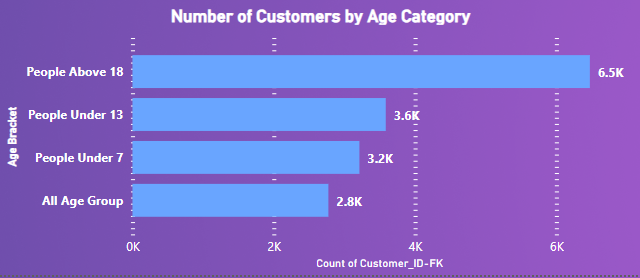
In contrast, Staff 2 exhibits a similar upward trajectory, albeit with a slightly slower rate of growth compared to Staff 1. The revenue increase from February to August is still notable, indicating consistent and commendable performance.

These insights suggest that both staff members are contributing positively to revenue generation, and their efforts appear to be paying off, albeit at varying rates. Understanding the factors driving these trends can inform best practices and potential strategies for revenue optimization, while also highlighting areas for further support or improvement.

**13. What is the distribution of films by language?**

We have data for only English language so We can’t do analysis on one Categorical value, to analyze the distribution between **film** language we need at least 2 categorical value to compare.

**Alternative Analysis for the number of rentals by the age category**



The data on the number of rentals categorized by age groups provides valuable insights into the rental demographics.

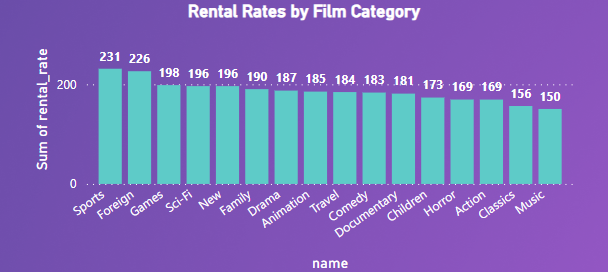
Firstly, the significant number of rentals for individuals aged 18 years and above, amounting to 6.5K, underscores the substantial demand from the adult population. This suggests that the rental services primarily cater to mature clientele, which may influence marketing strategies, product offerings, and service customization to meet the needs and preferences of this demographic.

The 3.6K rentals for people under 13 years old indicate a notable presence of children as customers. This insight suggests the importance of offering child-friendly rental options and ensuring safety measures to cater to the younger age group.

The 3.2K rentals for those under 7 years old emphasize the need for specialized rental solutions for toddlers and infants. These services should prioritize safety, comfort, and convenience, catering to the specific needs of this age group.

Finally, the 2.8K rentals for all age groups suggest diversity in the customer base. Tailoring services to cater to a wide range of ages while maintaining a focus on safety, quality, and customer satisfaction will be essential for the success of the rental business.

**14.** **Which film categories have the highest rental rates?**

****

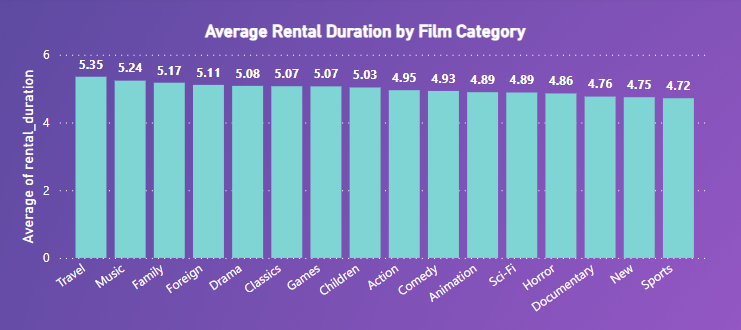
The data highlighting the rental rates across different categories provides valuable insights into customer preferences and potentially profitable areas within a rental service.

Firstly, the Sports and Foreign Film category emerges as the highest rental rate category, indicating a strong demand for content in these genres. This insight suggests that customers are particularly interested in sports-related content and foreign films, which could drive marketing and acquisition strategies to focus on expanding offerings in these genres.

The presence of Games, Sci-Fi, New, Family, and Drama categories with commendable rental rates underscores the importance of diversifying the rental library to cater to a broad spectrum of interests. Offering a variety of content, including the latest releases, sci-fi classics, family-oriented entertainment, and dramatic narratives, is crucial to meeting customer needs.

Incorporating these insights can inform content acquisition decisions, marketing efforts, and pricing strategies. By understanding the specific categories driving high rental rates, rental services can tailor their offerings to better satisfy customer demands and ultimately boost their profitability.

**15.** **How does the average rental duration vary by film category?**



The data indicating that film categories Travel, Music, Family, and foreign boast the highest average rental durations, while Sports exhibits the lowest rental duration, offers valuable insights into customer viewing habits and preferences.

The longer average rental durations in Travel, Music, Family, and foreign categories suggest that customers tend to engage more with content from these genres. It implies a higher level of interest, engagement, and potentially emotional connection to these types of films. Understanding this preference could inform content acquisition and marketing strategies, emphasizing these genres.

In contrast, the Sports category's lower average rental duration indicates that sports content might be viewed as more of an immediate, event-driven experience, with viewers seeking short-term access for specific events. This insight could guide pricing and content selection for sports-related content, optimizing the value offered to customers.

Tailoring the rental library and marketing campaigns to align with these insights can enhance customer satisfaction and engagement while maximizing revenue potential in the rental service.

**Conclusion**

In conclusion, our project on the Sakila Dataset has yielded valuable insights and demonstrated the power of SQL in data analysis. We conducted a multifaceted exploration of the dataset, leading to several key findings.

First, in our customer behavior analysis, we identified trends in rental preferences and discovered that certain film categories more popular among customers, Top spending customers, Customer Segment on the basis of active status and age bracket. This information can inform inventory management and marketing strategies.

Second, our store operation analysis revealed variations in store performance, with some locations outperforming others in terms of revenue and rental activity. Understanding these differences can guide strategic decisions and resource allocation.

Third, our film inventory analysis uncovered the most and least stocked film category, allowing for inventory optimization and cost reduction.

Fourth, the staff performance analysis highlighted variations in employee performance, enabling management to address training and efficiency concerns.

In addition to these analyses, we effectively harnessed Power BI to create interactive visualizations that enhance data accessibility. Moreover, we have done the Exploratory Data Analysis (EDA) on different Category of data. Our SQL proficiency was showcased through a wide range of questions answered, from basic to advanced, underscoring the versatility of SQL in data analysis.

By sharing our methodologies, findings, and interactive visualizations, we aim to promote knowledge sharing and provide a valuable resource for those interested in data analysis, SQL, and data visualization. Overall, this project exemplifies the utility of data analysis and SQL in unraveling insights and facilitating informed decision-making in a real-world context.