**PROJECT REPORT**

**Movie Rental Analysis**

By Rishabh Tyagi

OVERVIEW

In this project, we conducted queries on the Sakila DVD Rental database, which contains information related to a company engaged in renting movie DVDs. Our primary objective was to analyse and comprehend the characteristics of the customer base. This involved investigating patterns in movie-watching behaviours across diverse customer groups, examining how these groups differ in terms of payment earnings, and assessing the comparative performance of different stores within the company.

**Steps of Project Lifecycle**

1. **Data Acquisition from GitHub:** The project commenced with the acquisition of the dataset from a dedicated GitHub repository. This dataset encapsulates vital information pertaining to the operations of a movie rental store, encompassing details about customers, staff productivity, and store performance across diverse locations.

**2. Data Transformation and Enhancement:** To ensure data quality and consistency, rigorous data transformation procedures were implemented. Additionally, the dataset was augmented with a problem statement, enhancing its analytical potential and providing a clear direction for subsequent analyses.

**3. Connecting with Tools:** Establishing seamless connections between the dataset and various analytical tools was a critical step. This involved interfacing the dataset with Power BI, Excel, and MYSQL Workbench to facilitate comprehensive data integration and processing.

**4. Problem Statement Solution in Power BI:** Leveraging the robust features of Power BI, we delved into the specified problem statement. Power BI's capabilities in data visualization, exploration, and analysis were harnessed to derive meaningful insights and formulate effective solutions.

**5. Exploratory Data Analysis (EDA):** Exploratory Data Analysis (EDA) was conducted using either Excel or MYSQL, depending on the complexity of the analysis. This phase involved extracting meaningful patterns, relationships, and trends from the data to inform subsequent decision-making processes.

**6. Creation of Visuals and Insightful PowerPoint:** A comprehensive PowerPoint presentation was crafted to encapsulate the project's objectives, methodologies, solutions to problem statements, and key visualizations. Each problem statement was accompanied by a dedicated section presenting pertinent conclusions and insights.

**7. Detailed Documentation:** A meticulous record documenting the entire project lifecycle was compiled. This comprehensive document includes sections on data collection, transformation, problem statement formulations, tools integrations, Power BI solutions, EDA insights, and PowerPoint visualizations. The documentation is presented in a structured format with clear and concise sentences, providing a valuable reference for future analyses and collaborations.

**OBJECTIVE**

The primary objective of this project is to leverage the Sakila DVD Rental Store Database to create a robust Power BI dashboard. This dashboard aims to provide a holistic understanding of the rental store business, focusing on key aspects such as customer behaviour, film inventory management, staff performance, and overall store operations.

**Data-Driven Decision-Making:** Enable data-driven decision-making processes for rental store owners. Utilize insights derived from the dashboard to inform strategic business decisions.

**Customer Segmentation:** Analyse customer behaviour and preferences to create effective customer segmentation. Provide insights into different customer segments for targeted marketing campaigns.

**Sales Trends and Film Performance:** Explore and visualize sales trends over time, particularly focusing on monthly variations.Evaluate film performance metrics to optimize the film inventory for higher returns.

**Staff Productivity and Training:** Assess staff performance metrics to identify areas for improvement.Provide recommendations for staff training initiatives to enhance overall productivity.

**Store Revenue Optimization:** Understand store performance across different locations.Offer actionable insights to optimize store revenue and improve customer satisfaction.

**Actionable Recommendations:** Provide actionable recommendations for targeted marketing campaigns. Suggest enhancements to the film collection based on customer preferences.

**Final Deliverables:** Generate a comprehensive report detailing the findings and recommendations derived from the Power BI dashboard. Create an engaging presentation to effectively communicate the insights to stakeholders.

**Business Success in Competitive Market:** Empower rental store owners with the tools needed to succeed in the competitive DVD rental market.Facilitate informed decision-making for long-term business success.

In summary, this project aims to harness the potential of the Sakila DVD Rental Store Database to create a powerful Power BI dashboard. The insights derived will not only optimize current business operations but also provide a roadmap for sustained success in the dynamic and competitive DVD rental market.

**Significance of the Project**

The creation of this project holds significant value for various stakeholders involved in the DVD rental business, offering a range of benefits and opportunities:

**Informed Decision-Making:**  Understanding customer behaviour, sales trends and operational metrics facilitates informed decision making aligned with market dynamics.

**Optimized Inventory Management:** Insights into film performance, rental rates and customer preferences optimise inventory management, ensuring a diverse and in-demand collection.

**Enhanced Customer Satisfaction**: Customer segmentation analysis and targeted marketing recommendations improve customer satisfaction through tailored film collections and experiences.

**Maximized Revenue Streams:** Optimization of sales trends, store performance, and staff productivity directly impacts revenue streams, fostering business success.

**Strategic Staff Training:** Insights into staff metrics guide strategic training initiatives, enhancing employee capabilities and overall store efficiency.

**Competitive Edge in the Market**: Equipping rental store owners with tools to stay ahead of market trends and outperform competitors ensures a competitive edge and adaptability to changing preferences.

**Actionable Recommendations:** Delivering specific and actionable insights enables owners to implement changes directly impacting business performance.

**Long-Term Business Success:** Providing a roadmap for strategic decision making and continuous improvement supports long term business success and resilience in a competitive landscape.

**Summary:** The project's significance lies in transforming raw data into actionable insights, contributing to the success and longevity of DVD rental businesses. It aligns with the industry shift toward data-driven decision-making, ensuring rental store owners are well-equipped to navigate market complexities.

**DATA DICTIONARY**

A data dictionary is a structured collection of metadata or information about the data used in a database, including definitions, relationships, and constraints. Here's a simplified data dictionary for the Sakila DVD Rental database schema provided:

**Actor Table (actor):** The actor table lists information for all the actors, including first name and last name of actors.

actor\_id (INT): Unique identifier for actors.

first\_name (VARCHAR(45)): First name of the actor.

last\_name (VARCHAR(45)): Last name of the actor.

last\_update (TIMESTAMP): Timestamp of the last update.

**Address Table (address):** The address table contains address information for customers, staff, and stores.

address\_id (INT): Unique identifier for addresses.

address (VARCHAR(50)): Street address.

address2 (VARCHAR(50)): Additional address information.

district (VARCHAR(20)): District or region.

city\_id (INT): Foreign key referencing the city table.

postal\_code (VARCHAR(10)): Postal code.

phone (VARCHAR(20)): Phone number.

last\_update (TIMESTAMP): Timestamp of the last update.

**Category Table (category):** The category table lists the categories that can be assigned to films.

category\_id (INT): Unique identifier for categories.

name (VARCHAR(25)): Category name.

last\_update (TIMESTAMP): Timestamp of the last update.

**City Table (city):** city table contains a list of cities.

city\_id (INT): Unique identifier for cities.

city (VARCHAR(50)): City name.

country\_id (INT): Foreign key referencing the country table.

last\_update (TIMESTAMP): Timestamp of the last update.

**Country Table (country):** The country table contains a list of countries or regions.

country\_id (INT): Unique identifier for countries.

country (VARCHAR(50)): Country name.

last\_update (TIMESTAMP): Timestamp of the last update.

**Customer Table (customer):** The customer table contains a list of all customers.

customer\_id (INT): Unique identifier for customers.

store\_id (INT): Foreign key referencing the store table.

first\_name (VARCHAR(45)): First name of the customer.

last\_name (VARCHAR(45)): Last name of the customer.

email (VARCHAR(50)): Email address of the customer.

address\_id (INT): Foreign key referencing the address table.

active (BOOLEAN): Customer's active status.

create\_date (DATETIME): Date of customer creation.

last\_update (TIMESTAMP): Timestamp of the last update.

**Film Table (film):** The film table lists all the films that may be in stock in the store.

film\_id (INT): Unique identifier for films.

title (VARCHAR(255)): Title of the film.

description (TEXT): Description of the film.

release\_year (YEAR): Release year of the film.

language\_id (INT): Foreign key referencing the language table.

original\_language\_ id (INT): Foreign key referencing the language table.

rental\_duration (TINYINT UNSIGNED): Rental duration in days.

rental\_rate (DECIMAL(4,2)): Rental rate of the film.

length (SMALLINT UNSIGNED): Duration of the film in minutes.

replacement\_cost (DECIMAL(5,2)): Replacement cost of the film.

rating (ENUM('G','PG','PG-13','R','NC-17')): Film rating.

special\_features (SET): Special features of the film.

last\_update (TIMESTAMP): Timestamp of the last update.

**Film Actor Table (film\_ actor):** The film\_actor table is used to support many-to-many relationships between films and actors.

actor\_id (INT): Foreign key referencing the actor table.

film\_id (INT): Foreign key referencing the film table.

last\_update (TIMESTAMP): Timestamp of the last update.

**Film Category Table (film\_ category):** The film category table is used to support many-to-many relationships between films and categories.

film\_id (INT): Foreign key referencing the film table.

category\_id (INT): Foreign key referencing the category table.

last\_update (TIMESTAMP): Timestamp of the last update.

**Film Text Table (film\_ text):** The content of the film\_text table is kept in synchrony with the film table by means of triggers on the film table INSERT, UPDATE, and DELETE operations.

film\_id (INT): Foreign key referencing the film table.

title (VARCHAR(255)): Title of the film.

description (TEXT): Description of the film.

**Film Inventory Table (inventory):** A row in the inventory table represents a copy of a given film in a given store.

inventory\_id (INT): Unique identifier for inventory items.

film\_id (INT): Foreign key referencing the film table.

store\_id (INT): Foreign key referencing the store table.

last\_update (TIMESTAMP): Timestamp of the last update.

**Language Table (language):** The language table lists all possible values for the film language and original language.

language\_id (INT): Unique identifier for languages.

name (CHAR(20)): Language name.

last\_update (TIMESTAMP): Timestamp of the last update.

**Payment Table (payment):** The payment table records every payment made by the customer, including information such as the amount and rent paid.

payment\_id (INT): Unique identifier for payments.

customer\_id (INT): Foreign key referencing the customer table.

staff\_id (INT): Foreign key referencing the staff table.

rental\_id (INT): Foreign key referencing the rental table.

amount (DECIMAL(5,2)): Payment amount.

payment\_date (DATETIME): Date of payment.

last\_update (TIMESTAMP): Timestamp of the last update.

**Rental Table (rental):** The rental table contains a row for each rental of each inventory item, which contains information about who rented what, when it rented it, and when it was returned.

rental\_id (INT): Unique identifier for rentals.

rental\_date (DATETIME): Date of rental.

Inventory\_id (INT): Foreign key referencing the inventory table.

customer\_id (INT): Foreign key referencing the customer table.

return\_date (DATETIME): Date of return.

staff\_id (INT): Foreign key referencing the staff table.

last\_update (TIMESTAMP): Timestamp of the last update.

**Staff Table (staff):** The staff table lists all staff information, including email addresses, login information, and pictures.

staff\_id (INT): Unique identifier for staff members.

first\_name (VARCHAR(45)): First name of the staff member.

last\_name (VARCHAR(45)): Last name of the staff member.

address\_id (INT): Foreign key referencing the address table.

picture (MEDIUMBLOB): Staff member's picture.

email (VARCHAR(50)): Email address of the staff member.

store\_id (INT): Foreign key referencing the store table.

active (BOOLEAN): Staff member's active status.

username (VARCHAR(16)): Staff member's username.

password (VARCHAR(40) BINARY): Staff member's password.

last\_update (TIMESTAMP): Timestamp of the last update.

**Store Table (store):** The store table lists all stores in the system.

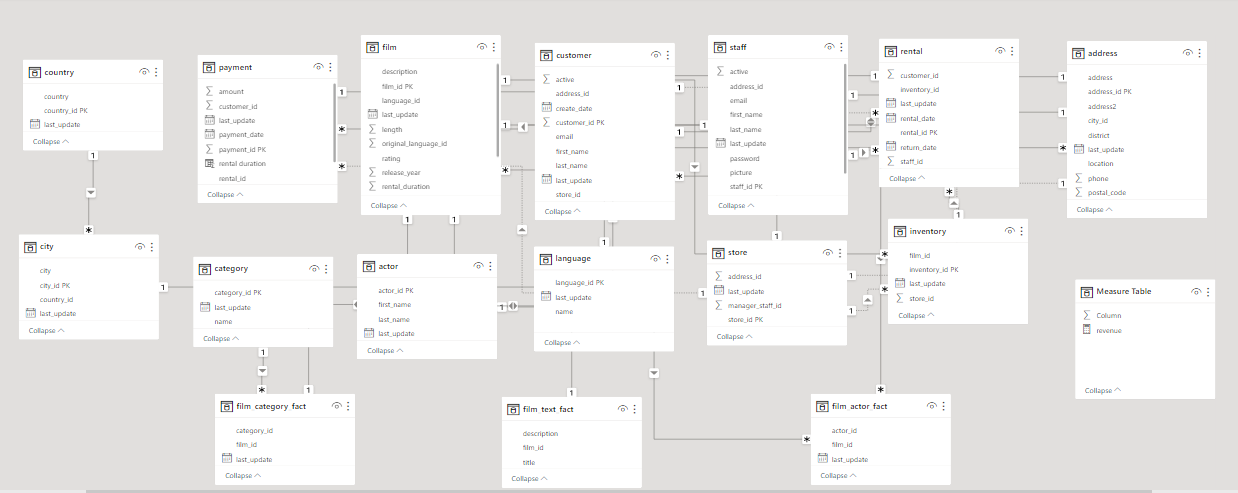
store\_id (INT): Unique identifier for stores.

manager\_staff\_id (INT): Foreign key referencing the staff table.

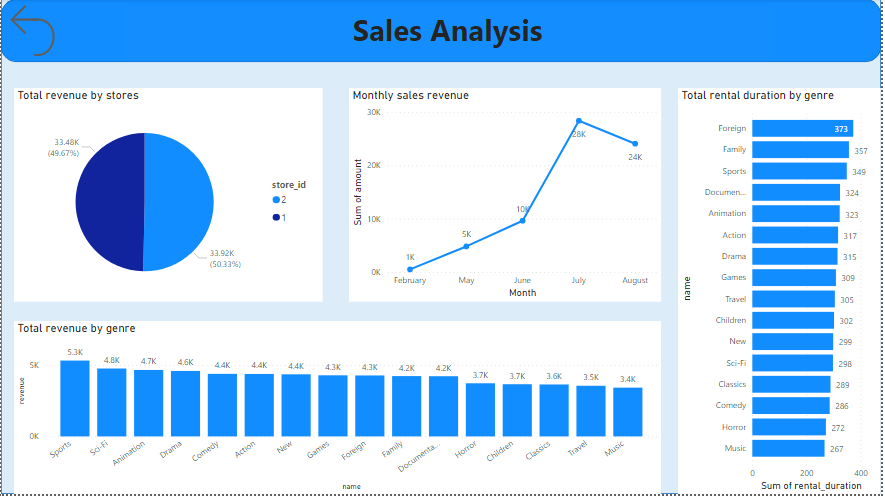
address\_id (INT): Foreign key referencing the address table.

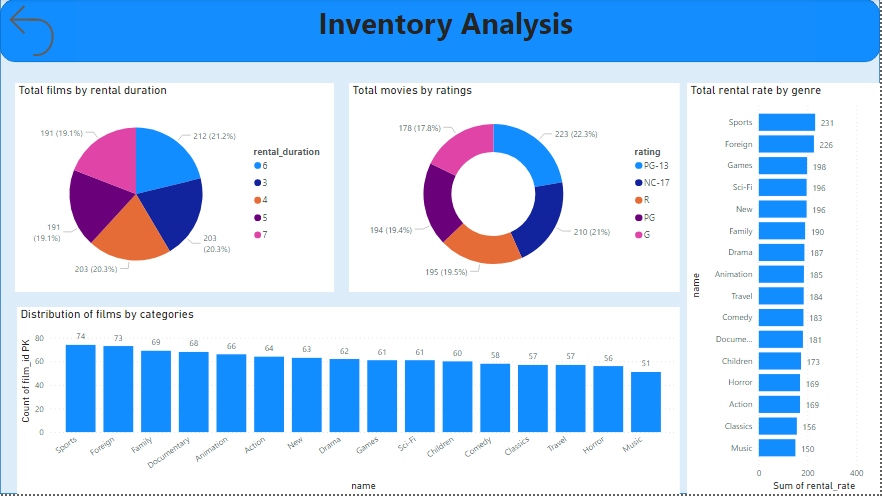
last\_update (TIMESTAMP): Timestamp of the last update.

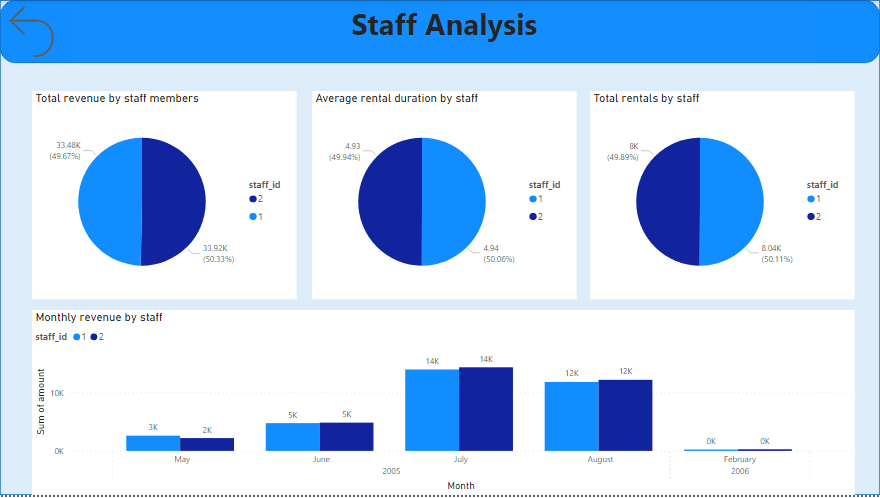
**Entity Relationship Diagram**

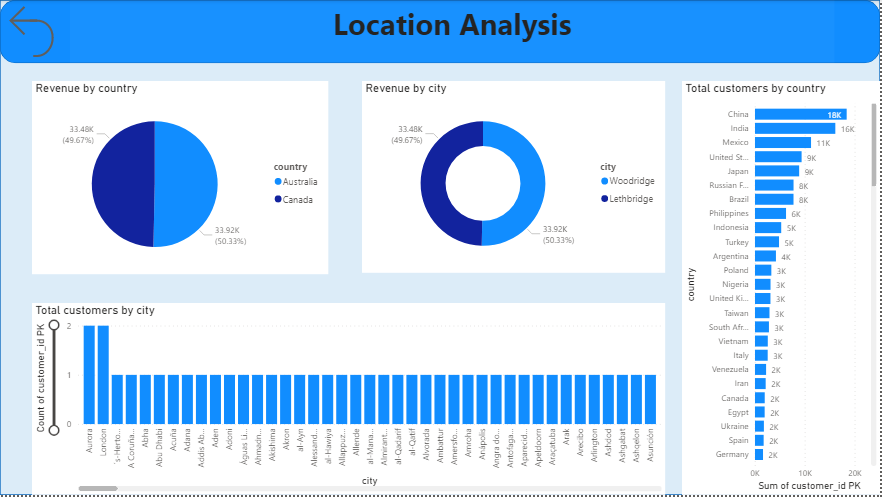


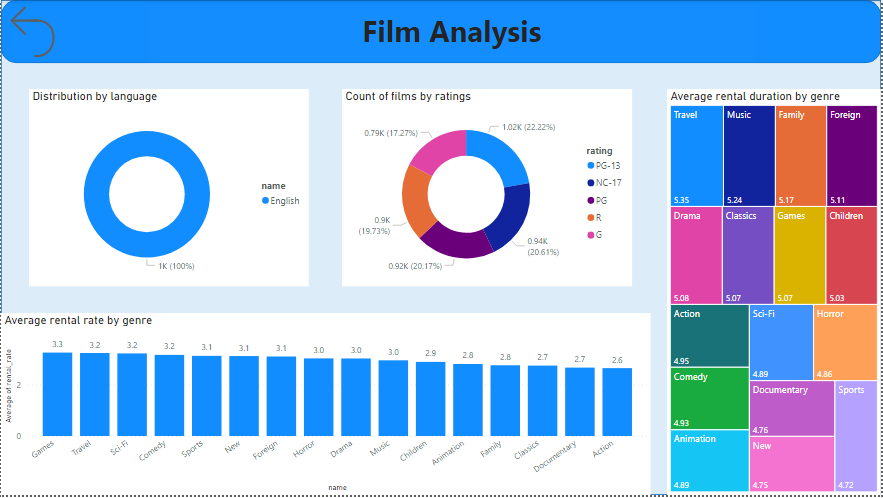
**POWER BI PROBLEM STATEMENT**











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

The analysis reveals distinctive patterns in customer segments based on rental duration, shedding light on their respective contributions to overall sales within the DVD rental store. Notably, customers opting for a rental duration of 3 emerge as the primary contributors to the store's highest sales. This suggests that individuals who prefer shorter rental periods, perhaps indicative of more frequent movie-watching habits, play a significant role in driving revenue through increased transaction frequency. On the contrary, customers selecting a rental duration of 7 exhibit the lowest contribution to sales among the segmented groups. This implies that customers with a preference for longer rental durations, potentially indicative of less frequent transactions or an inclination to keep movies for an extended period, may generate comparatively lower revenue for the rental store.

This information allows the rental store to optimize its strategies for different customer segments. For instance, understanding the preferences of customers with different rental durations can help in tailoring promotions, recommending movies, and adjusting inventory levels to meet demand effectively.

In summary, the analysis highlights the importance of considering customer behaviour, specifically their rental duration preferences, in driving sales strategies. By recognizing the varying contribution of different customer segments to sales, the DVD rental store can tailor its services to enhance customer satisfaction and overall business performance.

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

The analysis of customer distribution across different cities provides insights into the geographic spread of the customer base within the dataset. The dataset encompasses a total of 597 unique cities, indicating a diverse representation of customers across a wide range of locations. This diversity is crucial for understanding the reach and popularity of the DVD rental store's services.

A closer look at the specific cities with the highest number of customers reveals interesting patterns. Aurora stands out as the city with the most significant customer presence, boasting 50 customers. Following closely are London and Saint-Denis, with 48 and 46 customers, respectively. These cities with high customer counts may represent areas of strong market interest or high population density, making them key locations for the store's business.

Conversely, cities like Bydgoszcz, Basel, and Enshi have the least number of customers, each having 12, 14, and 14 customers, respectively. The lower customer counts in these cities might indicate areas with lower demand for DVD rental services, potentially influenced by factors such as local preferences, competition, or demographic characteristics.

Suggested actions based on this analysis:

Target Marketing Efforts: Focus marketing efforts on cities with higher customer counts, such as Aurora, London, and Saint-Denis, to maximize outreach and capitalize on existing popularity.

Explore Untapped Markets: For cities with lower customer counts, like Bydgoszcz, Basel, and Enshi, consider strategies to explore and tap into potential untapped markets. This might involve targeted promotions, partnerships, or understanding local preferences to make the services more appealing.

Customer Engagement: Engage with customers in cities with higher counts to understand their preferences better and tailor services to meet local demands. This could involve customer surveys, feedback mechanisms, or loyalty programs.

Evaluate Operational Efficiency: Assess the operational efficiency and resource allocation in cities with varying customer counts. Optimize inventory management, staffing, and promotional activities based on the unique characteristics of each city.

Diversification: Consider diversification or expansion strategies in cities with high demand to further strengthen the market presence. Conversely, for cities with lower demand, evaluate the feasibility of service adjustments or collaborations to enhance appeal.

**Which film category enjoys the highest popularity among customers?**

The analysis indicates that out of the 17 film categories, the sports category has garnered the highest popularity, attracting preferences from 1,179 customers. Following closely, animation and action genres also show substantial popularity with 1,166 and 1,112 customer preferences, respectively. On the other hand, music, horror, and travel categories are relatively less popular, each having preferences from 830, 846, and 837 customers, respectively.

Recommendations:

Content Emphasis: Given the high popularity of sports, animation, and action categories, it is advisable to emphasize and expand the film collection in these genres to cater to the strong customer demand.

Marketing Strategies: Incorporate targeted marketing strategies for popular categories, such as promotions, discounts, or themed events, to further engage customers and boost rentals.

Diversification: While maintaining a focus on popular categories, consider diversifying the film collection by introducing new titles and genres to attract a broader audience.

Customer Feedback: Collect and analyse customer feedback to understand preferences better and fine-tune the film collection strategy accordingly.

Inventory Management: Optimize inventory management by adjusting the stock levels based on the popularity of different categories, ensuring a balance between customer demand and available resources.

**Find top 5 customer according to sales?**

The top 5 customers based on sales are as follows:

Karl Seal with a total sales amount of $221.55.

Eleanor Hunt with a total sales amount of $216.54.

Clara Shaw with a total sales amount of $195.58.

Marion Synder with a total sales amount of $194.61.

Rhonda Kennedy with a total sales amount of $194.61.

These customers have demonstrated a high level of engagement and contribution to the store's revenue. Recognizing and appreciating their patronage could be beneficial, such as offering loyalty rewards, exclusive promotions, or personalized recommendations to further enhance their experience and encourage continued business. Additionally, analysing the preferences and behaviours of these top customers can provide insights into the factors that contribute to their higher spending, informing targeted marketing strategies to attract and retain similar high-value customers.

**Find customers who had made maximum orders?**

The list of the top 5 customers who made the maximum number of orders is as follows:

Eleanor Hunt with a total of 46 orders.

Karl Seal with a total of 45 orders.

Clara Shaw with a total of 42 orders.

Marcia Dean with a total of 42 orders.

Tammy Sanders with a total of 41 orders.

These customers stand out for their frequent engagement and transactions with the rental store. Recognizing their loyalty and consistent patronage could be valuable for building strong customer relationships. Consider implementing loyalty programs, personalized recommendations, or exclusive offers to further enhance their experience and encourage continued loyalty. Understanding the preferences and rental habits of these top customers can also guide inventory management and marketing strategies to cater to the preferences of the most active customer segments.

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

The analysis of inventory distribution by film rating provides insights into the composition of the film collection. The highest percentage of films falls into the PG-13 category, indicating a substantial presence of movies with a more mature target audience. The NC-17 and PG categories closely follow, suggesting a diverse selection catering to different age groups. The distribution reflects a balanced inventory with a variety of film ratings, contributing to a broad audience appeal.

Recommendations:

Diversification: While the inventory seems balanced, continuous monitoring and periodic updates are essential to ensure a diverse and up-to-date film collection that aligns with changing audience preferences.

Audience Segmentation: Consider conducting audience surveys or analysing rental patterns to better understand the preferences of different customer segments. This can aid in refining the inventory to cater to specific demographics.

Promotional Strategies: Leverage the popularity of PG-13 films by implementing targeted promotional strategies or themed events to attract a larger audience.

Refresh G and PG Categories: If the G and PG categories have lower representation, evaluate opportunities to introduce popular and high-quality films in these categories to cater to families and younger audiences.

Monitor Trends: Stay attuned to industry trends and emerging film ratings to ensure the inventory remains relevant and competitive in the ever-evolving entertainment landscape.

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

The inventory breakdown highlights a diverse collection of films, with Sports leading at 344 films, closely followed by Animation at 335. The inventory reflects a well-rounded offering, with genres such as Family, Drama & Foreign, Documentary, and Games contributing significantly. The presence of New releases and Classics suggests a balance between fresh content and enduring favourites, ensuring a varied and appealing selection for customers

Analysis:

Audience Variety: The presence of diverse categories suggests an effort to cater to various audience tastes, promoting inclusivity and broadening the customer base.

Popular Genres: The prominence of Sports and Animation suggests a high demand for these genres. Consider expanding or highlighting these categories in marketing efforts to capitalize on their popularity.

Timeless Appeal: The inclusion of both New Releases and Classics signifies a balance between staying current with the latest content and providing timeless favourites. This approach helps maintain customer interest over the long term.

Recommendations:

Regular Updates: Continuously assess and update the inventory to keep up with changing trends and customer preferences. Regularly refreshing the selection ensures a dynamic and engaging film catalogue.

Promotional Strategies: Implement targeted promotions or themed events around popular genres like Sports and Animation to attract and retain customers. Consider bundling related films for special offers.

Customer Feedback: Encourage customer feedback on film preferences to gain insights into evolving tastes. Use this information to fine-tune the inventory and enhance customer satisfaction.

Collaborations and Special Events: Explore collaborations or special events with local communities or organizations to showcase films from genres with cultural or seasonal relevance, fostering community engagement.

Monitor Performance: Regularly analyse the performance of each film category in terms of rentals and popularity. Use this data to make informed decisions about adjusting the inventory mix for optimal business outcomes.

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

The distribution of films by rental duration provides insight into the variety of film options available to customers based on different timeframes. In this analysis:

Rental Durations:

3 and 4: A total of 203 films fall under the rental durations of 3 and 4 days. This suggests a significant portion of the film catalogue is optimized for a relatively short viewing period, possibly appealing to customers seeking quick entertainment options or those with limited time commitments.

5 and 7: There are 191 films with rental durations of 5 and 7 days. This category may attract customers looking for a more extended viewing window, allowing them more flexibility in watching films at their own pace. It could cater to individuals who prefer a more leisurely viewing experience.

6: The largest group consists of 212 films with a rental duration of 6 days. This duration sits between the shorter and longer options, providing a balanced choice for customers who want a moderate timeframe for film enjoyment.

Analysis:

Customer Preferences: The distribution reflects an effort to cater to diverse customer preferences regarding the time they want to keep a rented film. Some may prefer quick, one-time entertainment, while others may want a more extended period for leisurely viewing.

Flexibility: Offering a range of rental durations ensures flexibility in accommodating various customer schedules and preferences. This flexibility can contribute to a positive customer experience and satisfaction.

Optimized Catalogue: By strategically allocating films across different durations, the inventory is optimized to meet the needs of a broader customer base, potentially leading to increased rentals and customer loyalty.

Recommendations:

Promotional Strategies: Implement targeted promotions or discounts based on rental duration to encourage customers to explore films with different viewing timeframes.

Customer Education: Provide clear information to customers about the rental duration options and benefits of each. This can help them make informed choices based on their preferences and availability.

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

The examination of film categories based on rental rates uncovers valuable insights into the pricing dynamics of various genres. Notably, the Games category stands out with the highest rental rate of 3.3, indicating a premium associated with films in this niche. Possible explanations for this could include a specialized and unique content offering within the Games genre or a perceived higher value by consumers. Following closely with a rental rate of 3.2 are the Travel, Sci-Fi, and Comedy categories, suggesting a consistent pricing strategy, possibly reflective of similar production costs or consistent viewer demand. The Sports, New, and Foreign categories share a rental rate of 3.1, indicating a moderate pricing approach, striking a balance between affordability and perceived value for films in these genres. This analysis underscores the genre-specific nature of pricing strategies, catering to the distinctive characteristics and market dynamics of each film category. To optimize revenue, recommendations include exploring dynamic pricing strategies, considering promotional packages, soliciting customer feedback for value perceptions, and staying vigilant on competitive pricing in the film rental market.

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

The analysis of the distribution of films by language indicates a peculiar scenario within the dataset. It appears that there is only one language represented, denoted by language\_id -1, and this singular language is identified as English. However, it's essential to note that this finding could be attributed to specific data handling practices or limitations within the dataset. The presence of only one language might be a result of data encoding choices, placeholder values, or potential data cleaning procedures. In a more typical scenario, one would expect to encounter a diverse array of languages associated with films to cater to global audiences. Given this anomaly, it is advisable to review the data collection and processing steps, ensuring that language information is accurately captured and represented. Additionally, exploring external sources or consulting the dataset documentation could provide insights into any constraints or peculiarities associated with language representation in this specific dataset.

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

The analysis of sales revenue variation by month reveals interesting trends in the dataset. From February to June, there is a consistent upward trajectory in sales revenue, suggesting a gradual increase in business during this period. This could be attributed to various factors such as seasonal demand, marketing strategies, or changes in customer behavior. The notable observation, however, is the sharp and significant spike in revenue between June and July. This abrupt increase points to a specific event, promotion, or external factor that significantly boosted sales during July.

Subsequent to the peak in July, there is a decline in sales, indicating a decrease in revenue in the months following the peak. This decline could be influenced by factors such as the conclusion of a particular promotion, a decrease in customer demand, or a shift in market dynamics. Analyzing the contributing factors to the peak in July and the subsequent decline can provide valuable insights for future business strategies. It's recommended to investigate the specific circumstances surrounding the notable increase in July to identify successful practices that can be replicated and understand the factors contributing to the subsequent decline to inform strategic adjustments and maintain a steady revenue stream.

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

The analysis of store performance by location indicates variations in revenue generation among different stores. Specifically, the Lethbridge store is reported to have a revenue collection of 33.48K, reflecting its financial performance within the dataset. In contrast, the Woodridge store outperforms Lethbridge with a higher revenue of 33.92K. This observation highlights a nuanced performance distinction between the two stores, with Woodridge leading in terms of revenue generation.

The factors contributing to the performance differences could be multifaceted, encompassing aspects such as local market dynamics, customer preferences, promotional strategies, and operational efficiency. Understanding the specific elements that contribute to the success of the Woodridge store can provide valuable insights for optimizing operations and strategies across other locations. Additionally, identifying areas of improvement for the Lethbridge store can help enhance its performance and contribute to overall business growth. This analysis underscores the importance of location-specific strategies and adaptability in maximizing store performance within a diverse market landscape.

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

The analysis of rental revenue by country in the Sakila movie rental dataset reveals notable variations in financial performance across different nations. India emerges as the leading contributor to rental revenue, boasting the highest income of 6.6K. Following closely, China secures the second position with 5.8K in rental revenue, and the United States follows suit with 4.1K.

This revenue distribution could be attributed to several factors, including the popularity of movie-watching culture, economic factors, and the size of the customer base in each country. The high revenue from India and China may be indicative of a robust movie rental market and a substantial audience engagement in these regions.

Conversely, countries such as Afghanistan and American Samoa register the least rental revenue, both at 0.1K. This could be influenced by factors like limited access to the movie rental service, lower customer participation, or potentially a smaller population contributing to the overall revenue.

To enhance rental revenue globally, it's essential to consider tailored strategies for each country. This may involve targeted marketing campaigns, understanding local preferences, and ensuring a diverse and appealing film selection. Additionally, exploring opportunities to expand services or partnerships in countries with lower revenue can contribute to overall growth.

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

The analysis of average rental duration by film category provides insights into the viewing habits and preferences of customers across different genres. The findings reveal notable variations in the average duration that customers keep films based on their categories.

Films in the Travel category stand out with the highest average rental duration, indicating that customers tend to retain movies from this genre for a longer period, with an average duration of 5.35. This could be attributed to the nature of travel-related content, often encompassing documentaries, scenic views, and cultural exploration, which might encourage viewers to savor the content over an extended period.

Following closely, the Music category boasts the second-highest average rental duration at 5.24, suggesting that customers engaging with musical content may find value in prolonged enjoyment of performances, concerts, or music-related documentaries.

Family films, with an average rental duration of 5.17, also show a tendency for longer viewer retention. This could be indicative of families taking their time to enjoy films together or the enduring appeal of family-friendly content.

On the contrary, the Sports category has the lowest average rental duration at 4.72, implying that customers may consume sports-related content more quickly, possibly driven by the immediate and time-sensitive nature of sporting events.

The New category, with an average rental duration of 4.75, suggests that recently released films may also have a shorter viewing span, possibly due to higher demand and quicker turnover in this category.

Understanding these variations in rental duration by film category is crucial for inventory management and customer satisfaction. Tailoring marketing strategies, promotions, and inventory decisions based on these insights can contribute to a more customer-centric approach, enhancing the overall effectiveness of the DVD rental business within the Sakila dataset.

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

The analysis of average rental duration by staff member in the Sakila movie rental database sheds light on the performance and efficiency of individual staff members in handling customer transactions. The findings indicate that Staff 1 and Staff 2 have relatively similar average rental durations, with Staff 2 holding a slightly higher average.

The average rental duration is a key metric as it reflects the amount of time customers retain DVDs before returning them. A lower average rental duration may suggest that a staff member is effective in processing returns

promptly and ensuring a quicker turnover of DVDs, potentially leading to increased customer satisfaction and improved availability of movies for other customers.

In this case, Staff 1's average rental duration of 4.93 indicates that, on average, DVDs rented by customers under the supervision of Staff 1 are returned within this timeframe. Similarly, Staff 2's slightly higher average of 4.94 suggests a comparable efficiency in managing rentals.

While the difference between the two staff members is marginal, it could be due to various factors, such as customer interactions, workflow management, or individual work styles. It's essential for the rental business to monitor these metrics regularly to identify any patterns or areas for improvement.

**EDA PROBLEM STATEMENT**

**What are the purchasing patterns of new customers versus repeat customers?**

The analysis of purchasing patterns in the Sakila movie rental dataset reveals interesting trends in customer behaviour, distinguishing between new customers and repeat customers over several months. The findings provide valuable insights into customer acquisition, retention, and potential areas for business strategy adjustments.

In May, there is a notable surge in new customer activity, with 520 new customers making purchases. This suggests a strong marketing impact or external factors driving new customers to the rental service. However, the subsequent months show a drastic decline in the number of new customers, dropping to just 60 in June, and no new customers added in the following months. This decline could be attributed to various factors such as the effectiveness of marketing campaigns, seasonality, or external market conditions.

Contrastingly, the count of repeat customers exhibits a different trend. From May to June, there is a clear increase in the number of repeat customers, indicating a positive shift in customer loyalty and satisfaction. The stability of repeat customer counts in the subsequent months suggests a baseline level of customer retention. However, there is a slight decline in August and a more significant drop in February, indicating potential challenges in retaining customer interest during these periods.

The analysis underscores the importance of balancing customer acquisition and retention strategies. While the initial focus may be on attracting new customers, maintaining and nurturing relationships with repeat customers is equally crucial for sustained business success. Strategies to enhance customer engagement, loyalty programs, and personalized recommendations could be implemented to encourage repeat business.

To address fluctuations in repeat customer counts, the business could implement customer feedback mechanisms, analyse customer satisfaction trends, and introduce incentives to encourage consistent patronage. A proactive approach to customer relationship management can help mitigate declines in customer retention and foster a loyal customer base.

Overall, this analysis provides a comprehensive view of customer behaviour, guiding the Sakila movie rental business to optimize its approach to customer acquisition and retention for sustained growth and success.

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

The analysis of film rental rates and demand in the dataset highlights key insights into the popularity and performance of specific films within the inventory. The film "Bucket Brotherhood" emerges as the most in-demand title, with the highest rental rate of $4.99 and an impressive 34 rental occurrences. This combination of a high rental rate and a substantial number of rentals underscores its popularity among customers.

Two other films, "Scalawag Duck" and "Goodfellas Salute," share the same rental rate as "Bucket Brotherhood" but exhibit variations in rental counts. While they have the same rental rate, their differing rental frequencies provide nuanced information about customer preferences. "Scalawag Duck" and "Goodfellas Salute" may appeal to specific segments of the customer base, and understanding these variations can aid in tailoring inventory management and marketing strategies.

The identification of these top-performing films is crucial for strategic decision-making within the movie rental business. It enables the business to focus on promoting and stocking films that have demonstrated high rental rates and demand. This, in turn, can lead to optimized revenue generation and customer satisfaction.

For inventory management, the business can prioritize stocking more copies of films like "Bucket Brotherhood" to meet the demand and capitalize on their popularity. Additionally, understanding the factors contributing to the success of these films, whether it be genre, cast, or storyline, can guide the acquisition of similar titles to diversify the inventory and attract a broader audience.

The analysis also emphasizes the importance of monitoring rental rates and demand trends over time. Regularly updating the inventory based on customer preferences and emerging film trends ensures that the movie rental business remains competitive and aligned with changing consumer interests.

In conclusion, the analysis of film rental rates and demand provides actionable insights for the Sakila movie rental business. It guides decision-making related to inventory stocking, marketing strategies, and customer engagement, ultimately contributing to the business's overall success in a dynamic and competitive market.

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

The analysis of customer behavior across different store locations in the dataset reveals distinct seasonal trends. The rental counts exhibit a noticeable increase during the summer months (June to August) and a corresponding decrease during the winter months (February). This observed pattern implies a seasonal influence on customer rental behavior, with potential factors such as holidays, weather conditions, or vacation periods contributing to the fluctuations.

During the summer months, customers tend to be more active in renting movies, leading to a surge in rental counts. This could be attributed to various factors associated with the summer season, such as school vacations, holidays, and generally more leisure time for individuals and families. The warmer weather and extended daylight hours may encourage people to engage in indoor activities like movie watching, contributing to the observed uptick in rental activity.

Conversely, during the winter months, particularly in February, there is a decline in rental counts. This decline could be influenced by factors such as colder weather, potential travel or holiday-related activities, and a general tendency for people to spend more time indoors. The reduced rental activity during these colder periods aligns with the expectation that customers may be less inclined to engage in movie-watching during the winter season.

Understanding these seasonal trends in customer behavior is valuable for the movie rental business. It allows the business to anticipate and plan for fluctuations in demand, adjusting inventory levels and marketing strategies accordingly. For instance, during peak summer months, the business might consider promoting specific genres or launching targeted marketing campaigns to capitalize on increased customer engagement. Conversely, during slower winter months, promotions or incentives could be introduced to stimulate rental activity.

In conclusion, the identification of seasonal trends in customer behavior provides actionable insights for optimizing business strategies. By aligning operations with these trends, the Sakila movie rental business can enhance customer satisfaction, improve resource allocation, and remain responsive to the dynamic patterns of customer demand throughout the year.

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

The analysis of language films within the movie rental dataset reveals a limitation in diversity, as all films in the dataset belong to a single language, which is English. Consequently, drawing conclusions about the popularity of films in specific languages among different customer segments becomes challenging.

In a more comprehensive dataset featuring a variety of language options, it would be possible to analyze whether certain language films resonate more with specific customer segments. Preferences for films in particular languages can be influenced by cultural factors, demographic characteristics, and individual preferences. For example, customers from different cultural backgrounds or regions may show distinct preferences for films in their native languages.

However, the current dataset's restriction to only English-language films limits the ability to explore such nuances in customer behavior. To gain meaningful insights into the popularity of films among specific customer segments based on language preferences, a dataset with a broader representation of languages would be required.

In practical terms, expanding the dataset to include films in multiple languages could provide a more accurate and nuanced understanding of customer preferences. This expanded dataset would enable the identification of patterns or trends related to language preferences among diverse customer segments. For a movie rental business aiming to cater to a diverse customer base, this information could be valuable for curating a well-rounded inventory that appeals to a broad range of tastes and preferences

In conclusion, while the current analysis highlights the presence of only English-language films in the Sakila dataset, the exploration of language preferences among specific customer segments necessitates a more diverse dataset with a broader array of language options. This would allow for a more comprehensive understanding of how language influences film preferences across different customer groups.

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

The analysis reveals a clear and impactful relationship between customer loyalty and sales revenue over time. Customer loyalty, measured by the repeat purchase rate, serves as a key driver of sales performance. The data illustrates that as the repeat purchase rate increases, there is a corresponding rise in sales revenue.

In July 2005, a month with a 100% repeat purchase rate, the sales revenue peaked at $28,373.89. This signifies that a high level of customer loyalty, where all purchases were made by repeat customers, had a direct and positive impact on the overall sales performance. The substantial revenue during this period underscores the significance of retaining and engaging existing customers, as they contribute significantly to the business's financial success.

Conversely, in February 2006, a month with a lower repeat purchase rate of 15%, the sales revenue experienced a notable decline, plummeting to $514.18. This stark contrast emphasizes the vulnerability of sales when customer loyalty diminishes. A lower repeat purchase rate indicates a reduced proportion of sales from repeat customers, which, in turn, affects the overall revenue negatively. This decline underscores the importance of consistent efforts to retain customers and foster loyalty over time.

The insights drawn from this analysis suggest that businesses should prioritize strategies and initiatives that enhance customer loyalty. This could involve implementing loyalty programs, personalized marketing efforts, and exceptional customer service to encourage repeat business. Building and maintaining strong relationships with customers can lead to sustained revenue growth and insulate the business from the fluctuations associated with a less loyal customer base.

In conclusion, the analysis demonstrates the dynamic interplay between customer loyalty and sales revenue, emphasizing the pivotal role of customer retention in achieving and sustaining business success over time. Businesses that invest in cultivating and nurturing customer loyalty are likely to experience more stable and prosperous revenue streams.

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

The analysis of film category preferences across different locations, specifically in Woodridge and Lethbridge, highlights variations in customer tastes and demands based on geographic factors. In Woodridge, it is observed that customers exhibit a notable preference for films in category 15, which corresponds to the sports genre. This suggests a specific interest or demand for sports-related content in that particular store. On the other hand, in Lethbridge, film category 1, corresponding to the action genre, emerges as the most popular among customers, indicating distinct preferences in this location.

This information is valuable for guiding inventory management and marketing strategies tailored to each store's unique audience. For instance, in Woodridge, the store could focus on expanding its sports-related movie offerings, ensuring a diverse selection of sports-themed films to cater to the apparent demand. Marketing initiatives in Woodridge could also highlight and promote sports films to capitalize on the customer interest in this genre.

In Lethbridge, the store could prioritize stocking a variety of action films, including new releases and popular classics. Marketing campaigns in this location could emphasize the extensive collection of action movies available for rent, potentially attracting more customers who favor this genre.

By aligning the movie inventory with the specific genre preferences of customers in each location, the movie rental business can enhance customer satisfaction and drive rentals. Furthermore, targeted marketing efforts that resonate with the unique tastes of each local audience can contribute to increased customer engagement and loyalty.

In summary, the analysis provides actionable insights for tailoring movie selections and marketing strategies to better meet the preferences of customers in Woodridge and Lethbridge, ultimately optimizing the performance of each store.

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

The analysis of the highest-spending customers, exemplified by individuals like Karl Seal and Eleanor Hunt, reveals a diversity in their film category preferences. Despite this diversity, category 2 emerges as a common choice among these high-spending customers. This suggests that these customers share an interest in films belonging to this particular category.

Understanding the film category preferences of these top-spending customers is valuable for developing targeted marketing strategies that can enhance customer engagement and satisfaction. By recognizing the commonality in category 2 preferences, the movie rental business can curate promotions, discounts, or exclusive offers specifically for films in this category. This targeted approach aims to capture the attention of high-spending customers, providing them with incentives to explore and rent more movies from category 2.

Moreover, the identification of shared preferences among top spenders allows for the optimization of the movie inventory. The business can prioritize stocking a diverse and appealing selection of films within category 2, ensuring that high-spending customers find a wide array of options to choose from during their visits.

In summary, the analysis not only highlights the diversity in film category preferences among top-spending customers but also identifies a common interest in category 2. Leveraging this insight, the movie rental business can implement targeted marketing strategies and optimize its inventory to cater to the preferences of these valuable customers, fostering increased satisfaction and loyalty.

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

The analysis of store activity reveals that the busiest hours for both Store 1 and Store 2 occur at 3 pm in the afternoon. This information holds significance in determining staffing requirements, as it identifies the peak operational hours during which customer activity, transactions, and overall store demand are at their highest.

Understanding the busiest hours is crucial for optimizing staffing levels and ensuring efficient store operations. During peak times, there is a greater need for customer service, assistance with transactions, restocking inventory, and overall operational support. By aligning staffing levels with the busiest hours, the stores can enhance customer service, reduce wait times, and maintain a smooth flow of operations.

In practical terms, this insight allows store managers to schedule additional staff during the identified peak hours, ensuring that there are an adequate number of employees available to handle the increased workload. Conversely, during less busy hours, staffing levels can be adjusted to reflect the lower demand, helping to manage operational costs effectively.

Efficient staffing aligned with the busiest hours not only improves customer satisfaction by providing prompt service but also contributes to the overall productivity and effectiveness of store operations. The ability to adapt staffing levels based on the observed patterns in store activity enables a more strategic and responsive approach to managing human resources within each store location.

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

The analysis of the data suggests a positive correlation between the availability of inventory and customer satisfaction, emphasizing the role of a well-stocked inventory in influencing customer behavior. This correlation is evident from the unique film rate of 1 for all customers, indicating that customers consistently find the films they are looking for in stock.

When customers can easily access the films they desire without facing stockouts or unavailability issues, it contributes to a positive experience and satisfaction. A film rate of 1 suggests that customers are successful in renting the specific movies they want, enhancing their overall satisfaction with the rental service.

The return rate of 1 further reinforces the idea that the availability of inventory is linked to repeat business. A return rate of 1 implies that customers are highly likely to return for additional rentals, indicating a level of loyalty and trust in the store's ability to maintain a diverse and well-stocked collection of films.

In summary, maintaining a sufficient and diverse inventory positively impacts customer satisfaction by meeting their movie preferences consistently. This satisfaction, in turn, fosters repeat business, with customers being more inclined to return to a store where they consistently find the films they enjoy. As a result, strategic inventory management plays a crucial role in not only satisfying current customers but also in building long-term customer loyalty and encouraging repeat transactions.

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

The analysis of the dataset indicates a limitation in language diversity, as all films are exclusively in English. Given this lack of variation in language options, it may not have a direct impact on customer satisfaction or rental frequency based on language preferences. However, the absence of multiple language options suggests a potential area for improvement to better cater to a diverse customer base.

While the current dataset does not provide insights into the impact of language diversity, it highlights an opportunity for the rental store to enhance its offerings by expanding the film collection to include movies in different languages. Offering films in multiple languages has the potential to positively influence customer satisfaction and rental frequency among a broader audience with diverse language preferences.

Customers from various linguistic backgrounds may have distinct preferences for films in their preferred languages. By incorporating a more diverse range of languages in the inventory, the rental store can better meet the entertainment needs of a broader customer base, potentially attracting new customers and encouraging existing ones to rent more frequently.

In summary, while the current dataset does not reveal specific impacts on customer satisfaction or rental frequency based on language diversity, it underscores the importance of considering language preferences in the overall business strategy. Expanding the film collection to include movies in different languages can be a strategic move to enhance customer satisfaction and attract a more varied audience, ultimately contributing to increased rental frequency and business growth.

**Conclusion**

To capitalize on customer segmentation and sales patterns, a strategic approach involves delving deeper into customer behavior. By identifying specific preferences, the business can tailor targeted promotions to high-sales segments, fostering increased engagement and loyalty. In terms of the distribution of customers across cities, implementing location-specific marketing campaigns and events can help strengthen the customer base in areas with lower presence, while also catering to the preferences of high-traffic locations.

Film category popularity can be further leveraged through strategic promotions. By promoting less popular categories through themed offers or bundled discounts, the business can increase visibility and attract a more diverse audience. Recognizing and rewarding top customers, such as Karl Seal and Eleanor Hunt, can be achieved through loyalty programs, exclusive offers, and personalized recommendations.

Inventory management plays a crucial role, and a proactive strategy involves regularly updating the inventory with new releases, classics, and popular categories. Monitoring customer feedback provides valuable insights for adjusting the inventory mix to align with evolving preferences. For rental duration and rates, introducing flexible pricing strategies or discounts for longer durations can encourage customers to opt for extended rental periods, contributing to increased revenue.

Diversifying the film collection by adding titles in various languages is a key strategy for enhancing inclusivity. This approach aligns with the goal of appealing to a broader customer base. Understanding sales revenue trends enables the business to plan promotions and exclusive releases during peak months, ensuring preparedness for high-demand periods.

Tailoring strategies to each store's unique characteristics is crucial for store performance and rental revenue. This involves conducting location-specific marketing campaigns, optimizing store layouts, and adjusting inventory based on customer preferences in each location. Implementing a robust loyalty program, complete with rewards and personalized recommendations, fosters repeat business and positively impacts sales.

Analyzing film category preferences by location allows for regular updates to inventory based on local preferences. Targeted marketing efforts can then be employed to promote films that align with each store's specific audience. Understanding the busiest hours and staffing accordingly, through flexible schedules and cross-training staff for efficiency during peak periods, optimizes operations.

Ensuring inventory availability is a critical factor in customer satisfaction. Implementing real-time inventory tracking systems and addressing any issues through customer feedback contribute to a seamless rental experience. Finally, recognizing the impact of language on customer satisfaction involves collaborating with distributors to source films in different languages and creating language-specific sections to cater to diverse language preferences.

Implementing these comprehensive strategies collectively enhances customer satisfaction, drives sales, and positions the movie rental business for sustained growth in a competitive market. Regular reassessment and adaptation of these strategies based on evolving customer trends and industry dynamics are essential for continued success.