

|  |
| --- |
| Capstone Project |
| Sakila DVD Rentals |
| October 1’ 23  *Rushikesh S Sathe* |

# Overview

|  |
| --- |
| Store Database is designed with a relational structure comprising tables for Customers, Films, Rentals, Staff, Stores, and Categories. These tables collectively organize and manage customer data, film inventory, rental transactions, staff information, store details, and film categories, facilitating efficient data retrieval and analysis for rental store operations. |
|  |

|  |
| --- |
| Process1. Data Acquisition from GitHub:Obtain the requisite dataset from a designated GitHub repository, containing essential information on Rental Store, encompassing various countries and their performance across distinct rental system.2. Data Transformation and Enhancement:If necessary, execute data transformation procedures to ensure data quality and consistency. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential.3. Connecting with Tools:Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, & Excel. Workbench, facilitating seamless data integration and processing.4. Problem Statement Solution in Power BI:Utilize Power BI to delve into the specified problem statements employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions. |
| 6. Creation of Visual and Insightful PowerPoint:Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights.7. Detailed Documentation:Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions, EDA insights, and PowerPoint visualizations. |

# Objective

# The primary objective for analyzing a DVD rental shop’s operation is to gain actionable insights that enable informed decision-making and business improvement. This analysis aims to:

# 1. Optimize Customer Experience: - Understand customer preferences, behaviors, and rental patterns to enhance customer satisfaction and retention.

# 2. Maximize Revenue: - Identify high-value films and customer segments, optimize pricing strategies, and improve inventory management to increase revenue streams.

# 3. Enhance Operational Efficiency: - Evaluate staff performance, assess store operations, and allocate resources effectively to streamline operations.

# 4. Improve Marketing Strategies: - Utilize demographic and category data to tailor marketing efforts, offer personalized recommendations, and drive customer engagement.

# 5. Sustain Competitiveness: - Stay competitive in the DVD rental market by adapting strategies based on data-driven insights and market trends.

# 6. Ensure Inventory Diversity: - Maintain a diverse and appealing film collection to meet the preferences of a wide customer base.

# Overall, the analysis aims to make the DVD rental shop more efficient, customer-centric, and profitable in a competitive market landscape.

# Significance

# DVD rental store analysis plays a significant role in the entertainment industry landscape, providing valuable insights for customers, businesses, and policymakers. Let's explore the significance of DVD rental store analysis in more detail.

# Customers:

# Selection and Variety: DVD rental store analysis helps customers make informed decisions about where to rent DVDs. By understanding the store's catalog, customers can assess if it aligns with their movie preferences, including new releases, classics, or niche genres.

# Availability and Convenience: Customers can use analysis to determine which stores offer convenient rental options, such as online reservations, multiple locations, or extended hours. This information allows them to choose a store that best suits their schedule and location.

# Pricing and Membership: Analyzing rental store data can help customers find the best deals, discounts, or membership benefits. This knowledge empowers them to make cost-effective choices and enjoy their movie-watching experience without breaking the bank.

# Recommendations and Reviews: Accessing user reviews and recommendations from DVD rental store analysis can guide customers in selecting movies that align with their tastes and interests, enhancing their overall viewing satisfaction.

# Policymakers:

# Consumer Protection: Policymakers can use DVD rental store analysis to ensure fair pricing and transparent rental terms for customers, protecting them from deceptive practices in the industry.

# Market Regulation: Analysis can assist policymakers in monitoring the health and competitiveness of the DVD rental market. They can intervene if necessary to prevent monopolistic behavior or anticompetitive practices.

# Copyright and Licensing: Policymakers can assess the impact of DVD rental store analysis on copyright and licensing agreements, helping to strike a balance between the interests of content creators, rental stores, and consumers.

# Digital Transition: As the industry transitions from physical DVDs to digital formats, policymakers can use analysis to understand the implications on DVD rental stores, potentially offering incentives or support during this transition.

# In summary, DVD rental store analysis is crucial for customers, businesses, and policymakers alike, providing valuable insights for decision-making, business optimization, and regulatory oversight within the entertainment sector.

# Data Dictionary

# City: Represents different cities where the rental stores are located. Each city is identified by a unique city ID and contains attributes such as city name and postal code.

# Address: Stores address details for various entities like customers and stores. Each address is identified by a unique address ID and includes attributes like street address, district, and postal code.

# Country: Contains information about different countries. Each country is identified by a unique country ID and may have attributes like country name and region.

# Language: Represents different languages in which films are available. Each language is identified by a unique language ID and may have attributes like language name.

# Customer List: Serves as a junction table between customers and rental stores. It stores information about the customers associated with specific rental stores.

# Customer: Contains details about individual customers who rent films. Each customer is identified by a unique customer ID and includes attributes such as customer name, email, and membership details.

# Category: Represents film categories or genres. Each category is identified by a unique category ID and includes attributes like category name.

# Store: Represents individual rental stores. Each store is identified by a unique store ID and contains attributes like store name, manager staff ID (foreign key), and address ID (foreign key).

# Inventory: Stores film inventory information for each rental store. It includes film ID (foreign key) and store ID (foreign key).

# Rental: Records film rental transactions. It includes rental ID, rental date, return date, customer ID (foreign key), staff ID (foreign key), and inventory ID (foreign key).

# Sales by Store: Contains sales data grouped by store. It includes attributes such as store ID (foreign key) and total sales.

# Sales by Film Category: Contains sales data grouped by film category. It includes attributes such as category ID (foreign key) and total sales.

# Actor Info: Contains actor information, including actor ID (foreign key), first name, and last name.

# Film Text: Stores film description and additional textual data for films. It includes film ID (foreign key) and film description.

# Staff: Represents staff members working at rental stores. Each staff member is identified by a unique staff ID and contains attributes such as staff name, email, and store ID (foreign key).

# Staff list: Serves as a junction table between staff and rental stores, storing information about staff members associated with specific rental stores.

# Payment: Records payment transactions for film rentals. It includes payment ID, payment date, rental ID (foreign key), and payment amount.

# Film: Contains details about individual films available for rent. Each film is identified by a unique film ID and includes attributes such as film title, release year, and rental duration.

# Actor: Represents individual actors. Each actor is identified by a unique actor ID and includes attributes like actor name and gender.

# Nicer But Slower Film: Contains data about films considered "nicer but slower" versions of original films. It includes film ID (foreign key) and additional film details.

# Film List: Serves as a junction table between films and categories, linking films to their corresponding categories.

# Film Category: Represents the relationship between films and categories. Each film can belong to multiple categories, and each category can have multiple films.

# Film Actor: Represents the relationship between films and actors. Each film can have multiple actors, and each actor can appear in multiple films.

# Relevant Data Missing and Errors Identified:

# 1. Customer Satisfaction: No customer satisfaction ratings available.

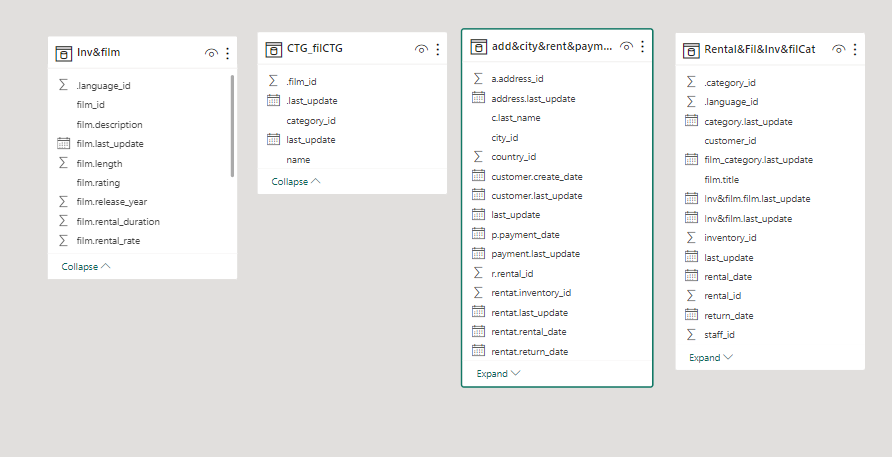
# 2. Language Options: Only one language film (ID-1 English) is available.

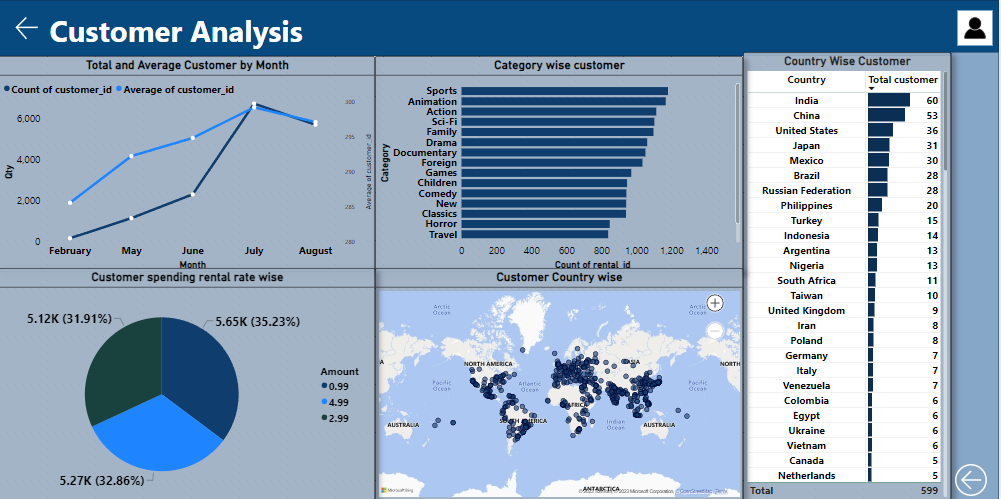
# 3. Limited Customer Age Data: Incomplete information makes it impossible to define customer age accurately.

# 4. Payment Method Data: Payment method information is not provided.

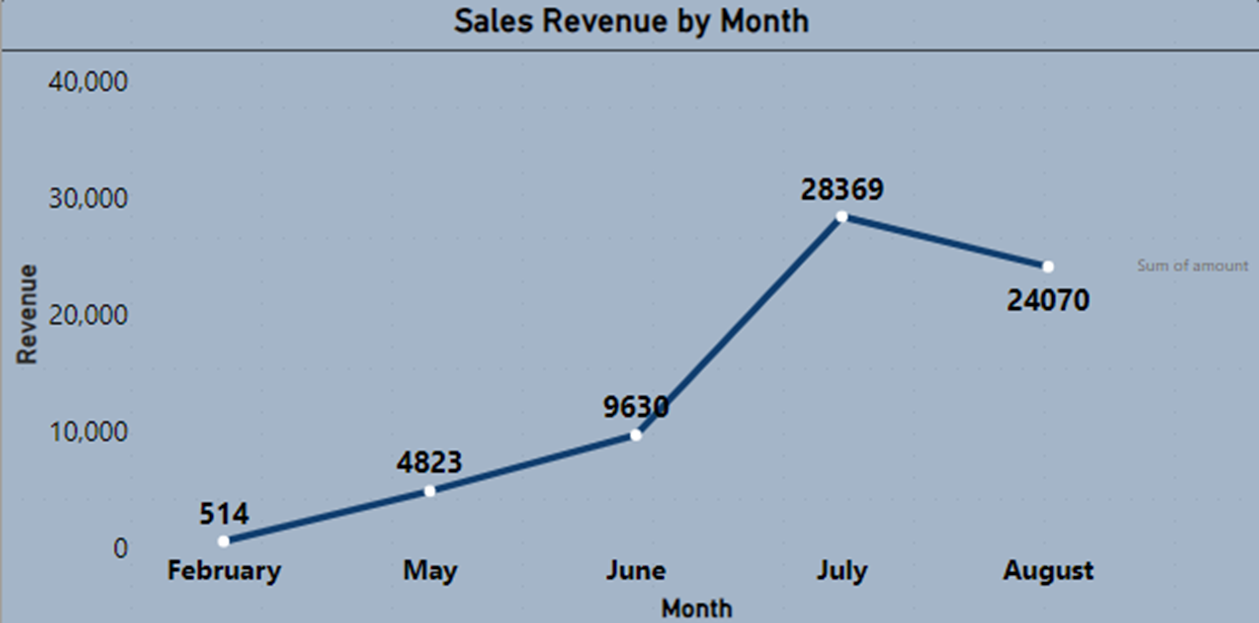
# 5. Purchase History Date Conflict: Customer purchase history dates are earlier than the customer ID creation dates, indicating a data error.

# ER Diagram



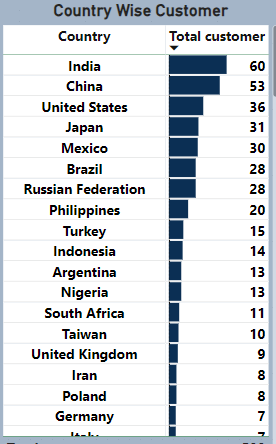


How does the sales revenue vary by month?

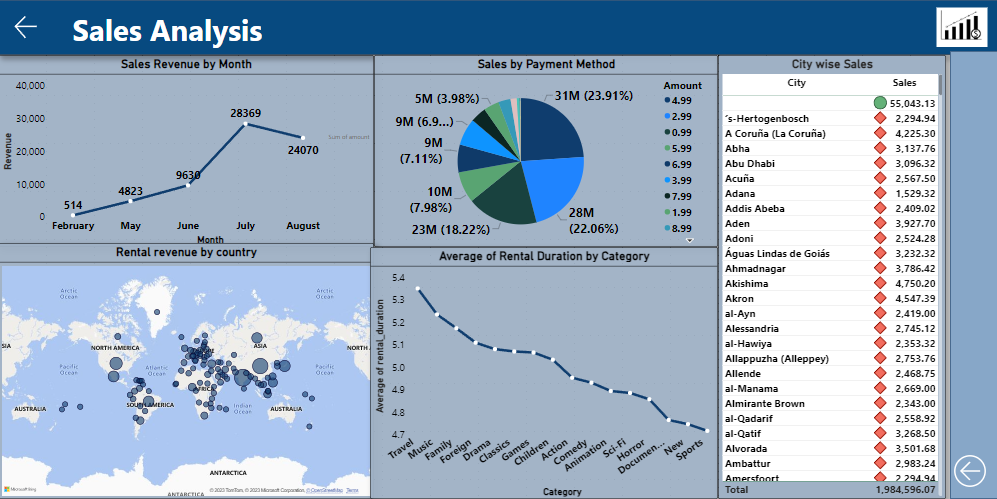


Based on the data provided, we observe that sales revenue experiences a gradual increase starting in June, reaches its zenith in July, and then undergoes a gradual decline of approximately 15%.

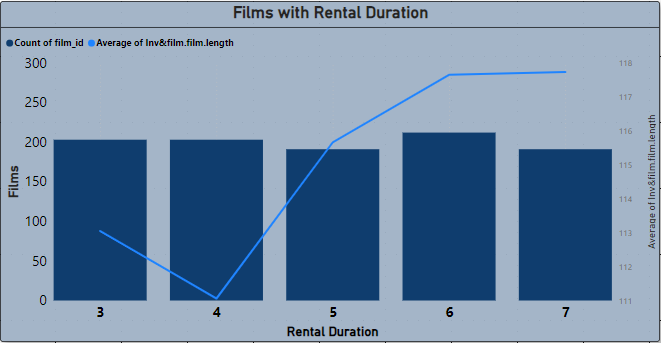
Which customer segments generate the highest sales?



Based on the available data, it is evident that India, China, the United States, and Japan have the highest number of customers. When considering sales by category, Sports and Animation Action movies exhibit the highest sales figures.

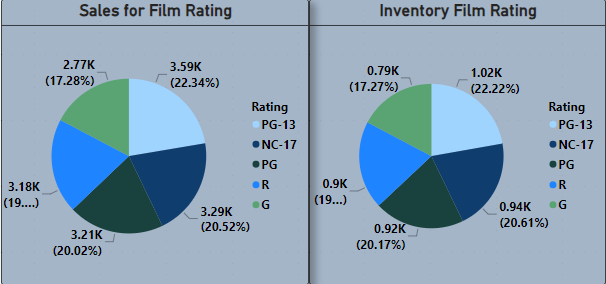


What is the distribution of films by rental duration?

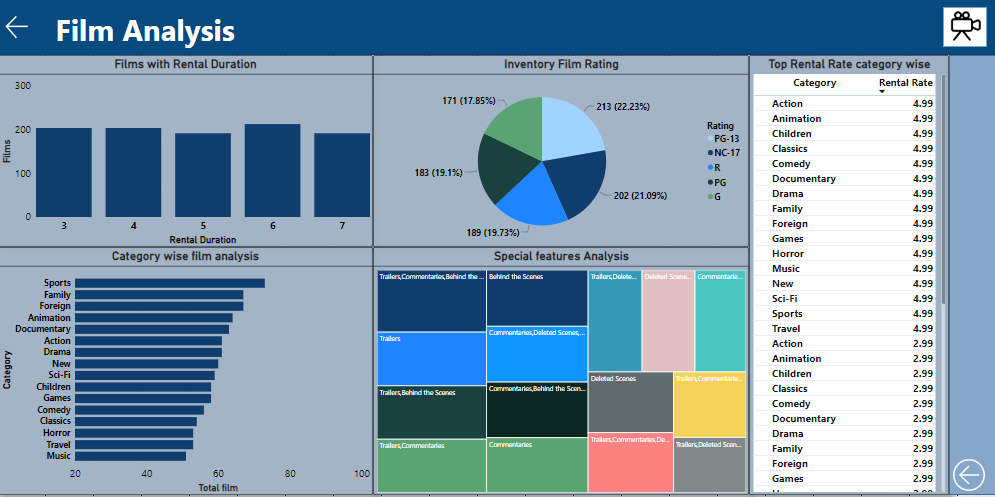


In the purchase data, it appears that rental durations do not show significant differences, but there is variation in the length of film, with the highest frequency occurring for a duration of 7.

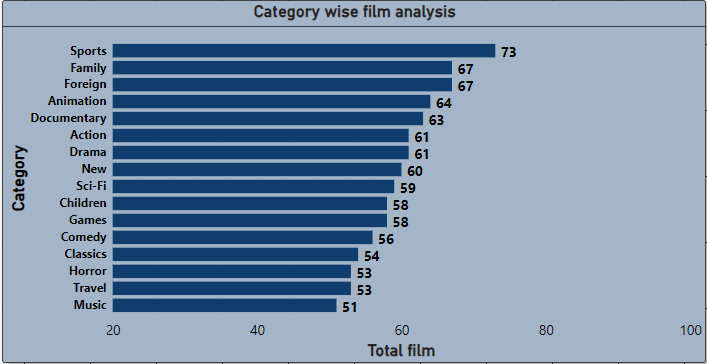
How does the inventory vary by film rating?



Based on the analysis above, it is evident that the PG-13 rating holds the highest stock at 22.23%, while the G rating has the lowest stock. This trend also extends to sales figures, where PG-13 movies outperform those with a G rating.

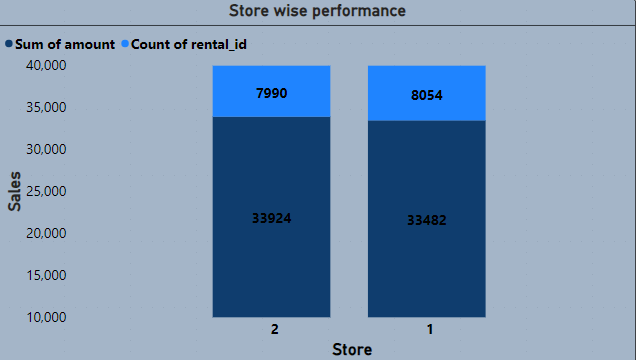


What is the breakdown of film categories in the inventory?



It can be observed that the inventory stock for stores is highest in the categories of Sports, Family, Foreign, and Animation, followed by other categories.

How does the store performance vary by location?

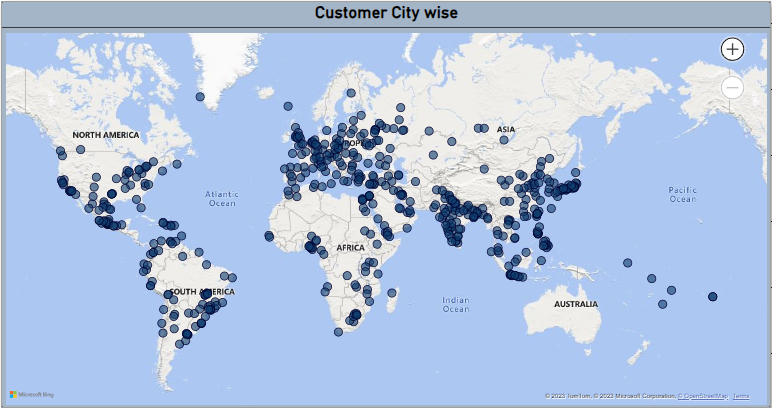




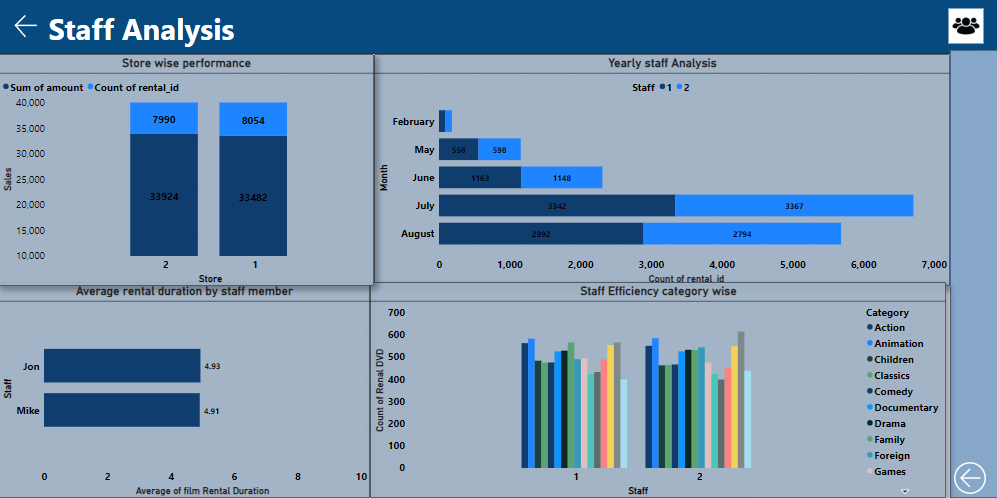
When evaluating store performance, it becomes evident that Store 1 boasts a larger customer base compared to Store 2, yet it generates lower revenue than Store 2. Furthermore, considering market penetration in City, Store 1 has a more significant share of the city's customer base compared to Store 2.

From these insights, it can be deduced that Store 1 may require improvements in operational efficiency to convert its larger customer base into higher revenue. Conversely, Store 2 could consider expanding its customer base within the city through strategic marketing initiatives to enhance its market presence.

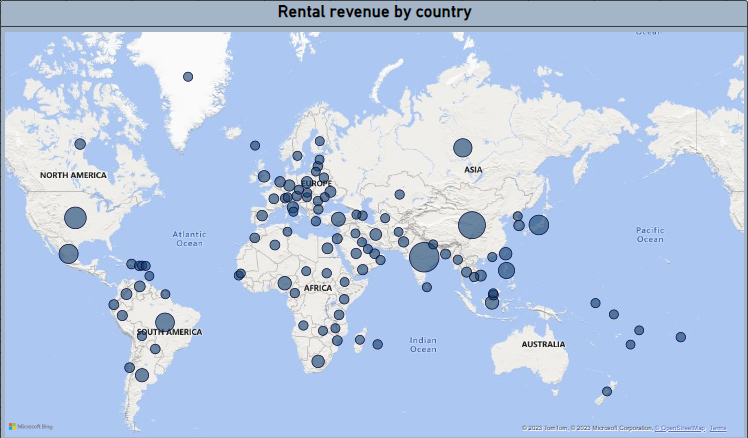
What is the distribution of customers across different cities?



Based on the map data, it's apparent that we have a substantial market presence in Southeast Asia and Europe. However, there's room for growth and expansion in countries like Australia, Africa, and South America. Therefore, it would be advisable to focus on these regions to tap into untapped market segments and potentially increase our market share and global reach.

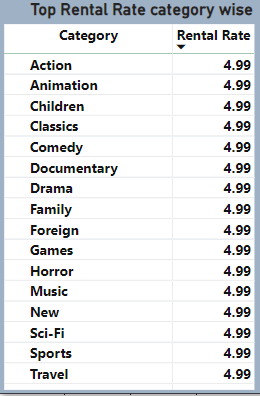


How does the rental revenue vary by country?



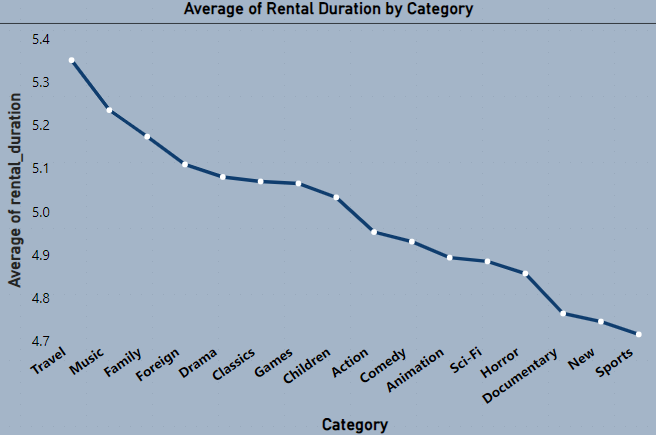
Based on the map data, it's evident that we generate substantial revenue in Southeast Asia, Europe, and certain parts of Africa. Nevertheless, there is an opportunity for growth and expansion in countries such as Australia, Africa, South America, and the former USSR region. Therefore, it is recommended that we concentrate our efforts in these areas to target untapped market segments and potentially enhance our market share and global presence.

Which film categories have the highest rental rates?

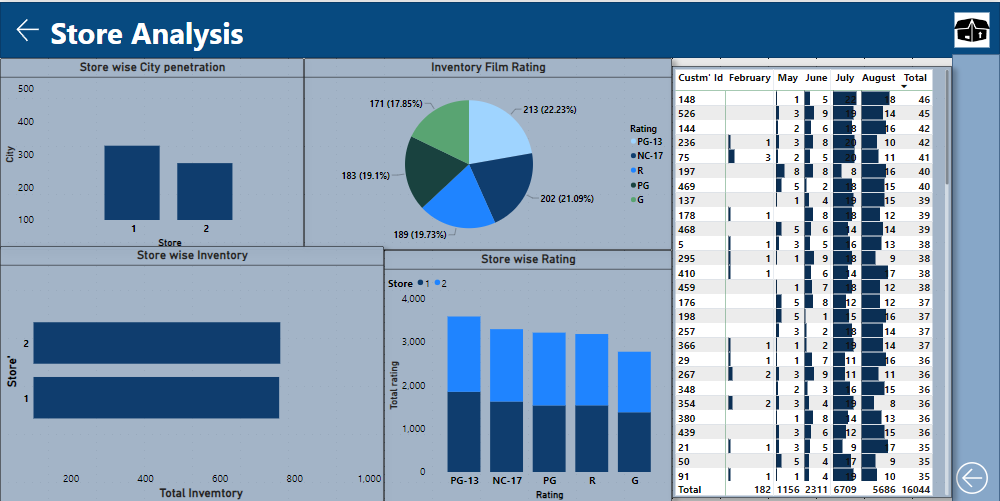


From the data, it is discernible that the genres with the highest film rental rates are Action, Animation, Children, and Classics, in that order.

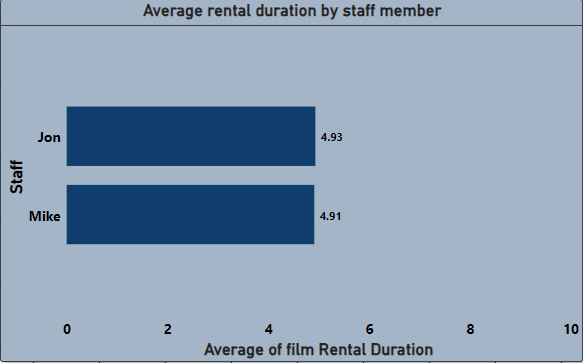
How does the average rental duration vary by film category?



It can be observed that the average rental duration for Sports, New, Documentary, Horror, and Sci-Fi genres is relatively low, while genres such as Travel, Music, Family, and Foreign have the highest rental duration. This suggests that people are more inclined towards exploring and enjoying outdoor experiences, as reflected in their longer rental durations for genres associated with travel and family-oriented entertainment.



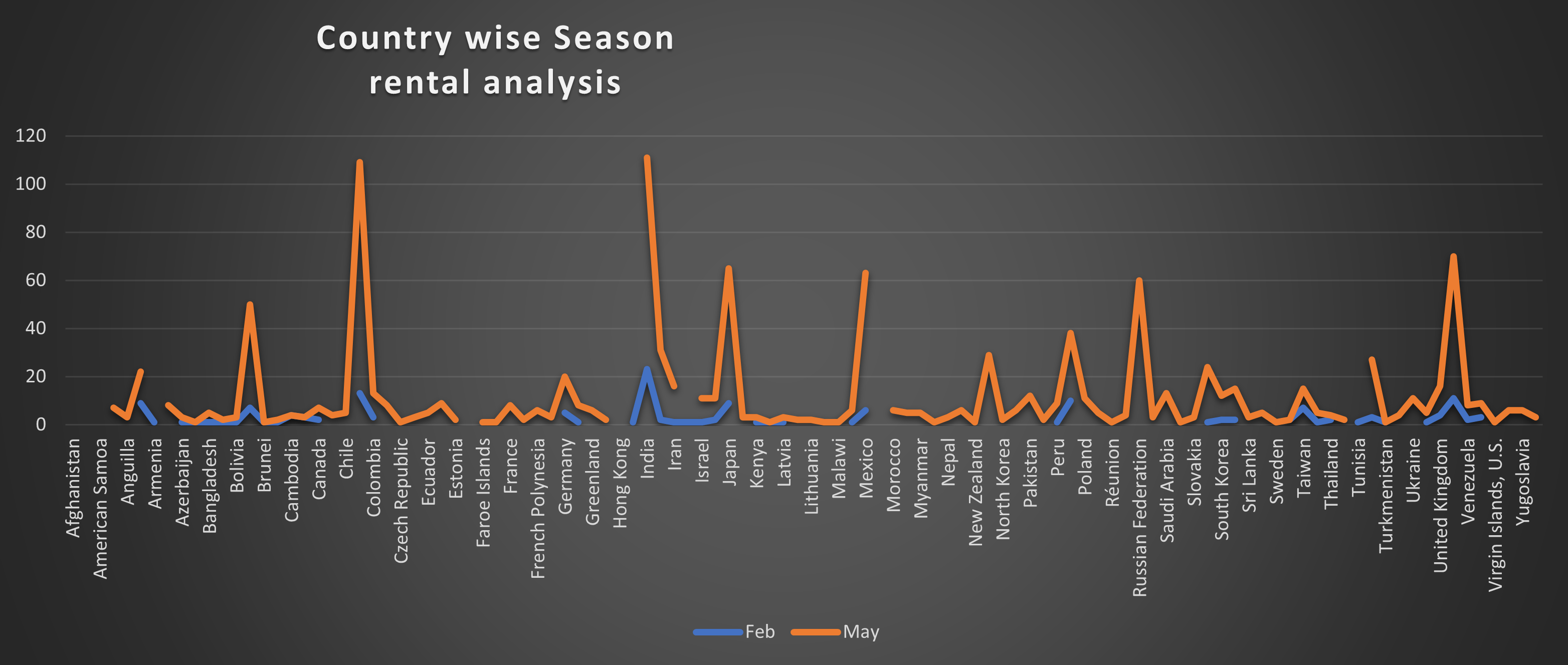
What is the average rental duration by staff member?

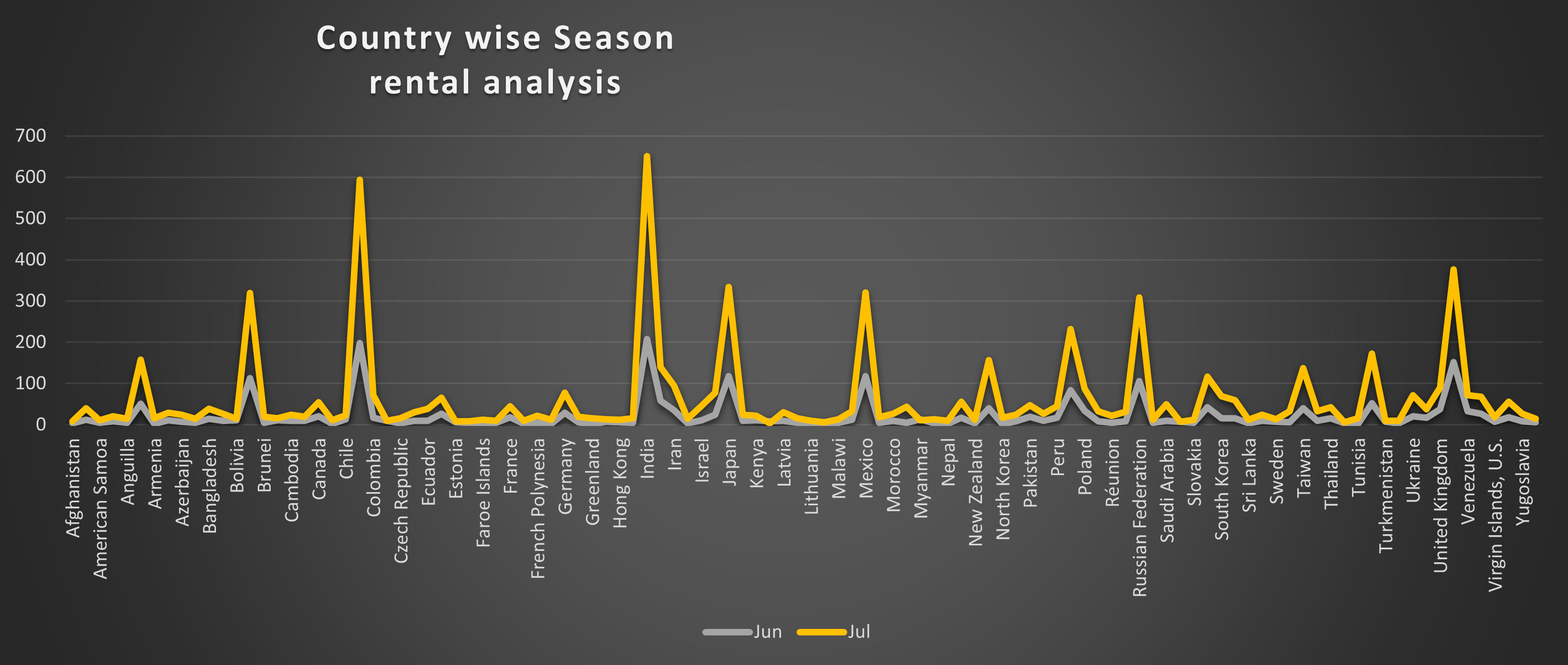


The average rental duration for staff doesn't exhibit significant differences. Therefore, it would be prudent to concentrate efforts on enhancing market penetration for Store 2. Additionally, optimizing efficiency during non-peak seasons through targeted offers and sales strategies tailored to specific city and country dynamics can maximize overall store performance.

EDA Statement

Are there seasonal trends in customer behavior across different locations?



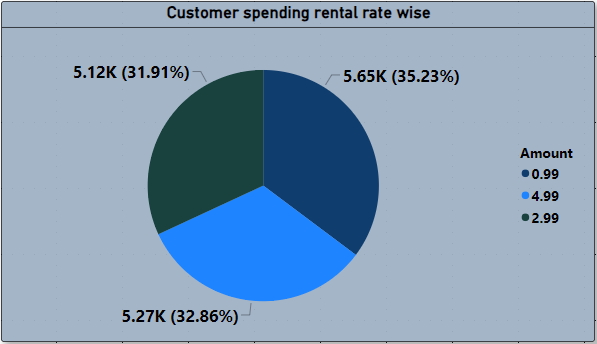


From the data, it's evident that in February, there isn't a substantial response from any country, with a slight increase in sales observed in India, Peru, and the UK.

On the other hand, during the month of May, there is an uptick in sales, particularly from countries such as Bolivia, Chile, Iran, Japan, the Russian Federation, and the UK. This suggests a seasonal trend or specific factors driving increased sales during that period in these countries.

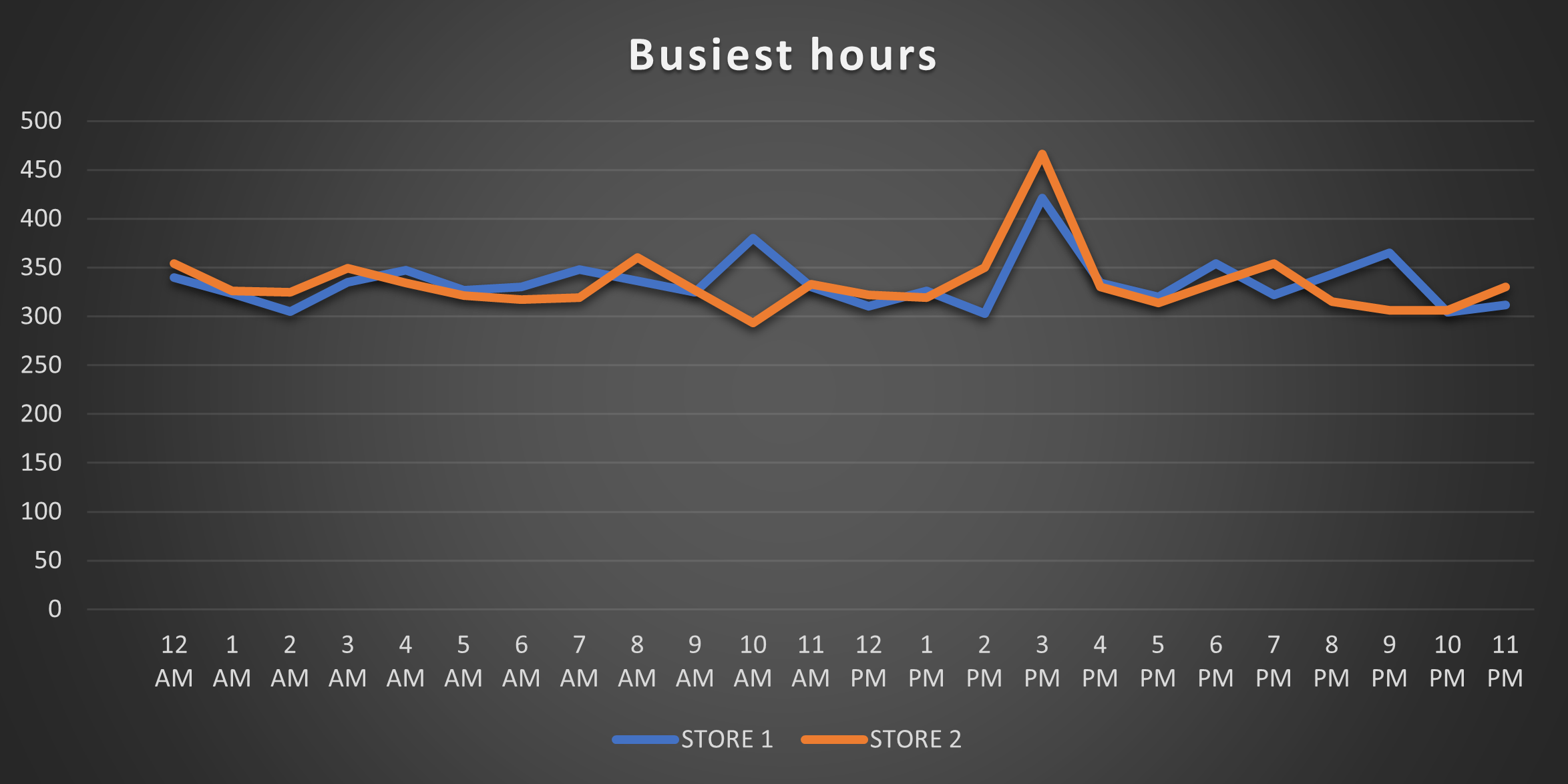
It appears that June has lower sales compared to July. In July, there is a more favorable response from all countries, with notable increases in sales from Chile, India, Bolivia, Japan, Mexico, the Russian Federation, and the UK. This trend continues into August, where the response remains similar to that of July.

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



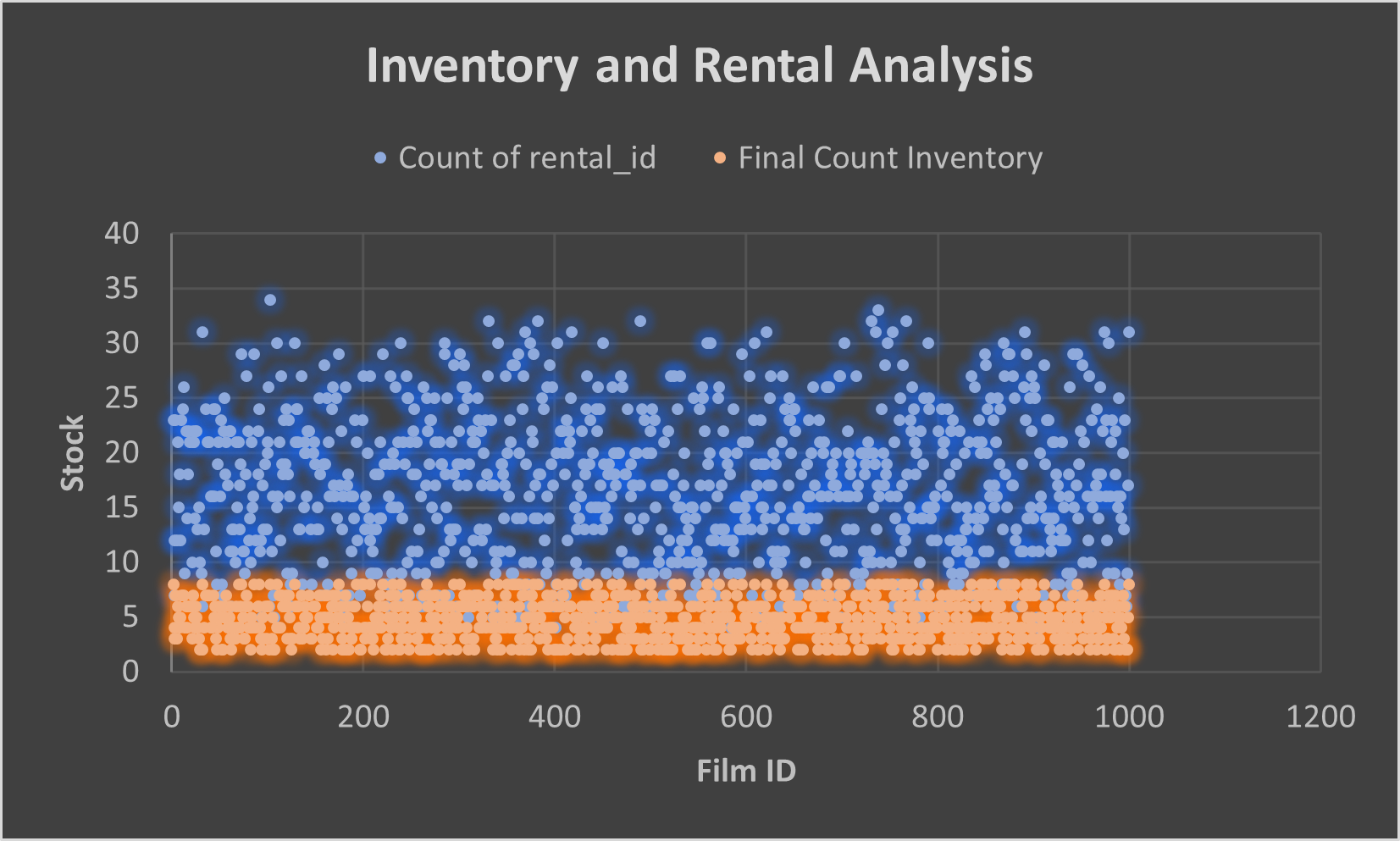
From the analysis, it is evident that cities such as Aurora, London, Cape Coral, Tanza, and Shimonoseki have the highest customer spending. Additionally, it appears that expensive rental rates, contributing approximately 32.86% of the total rental rate shares, play a significant role in revenue generation.

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



The busiest hours for Store 2 are between 7 AM and 11 AM, while for Store 1, they are from 9 AM to 11 AM. Peak hours for both stores occur from 1 PM to 5 PM simultaneously. To optimize staffing and resource allocation, it's advisable to consider reducing staff during non-peak hours and potentially reallocating them for marketing or other productive tasks. This can help improve efficiency and cost-effectiveness during different parts of the day.

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



When comparing inventory and film sales data, it's apparent that there is a notable gap between stock levels and sales trends. To address this issue and optimize revenue generation, a thorough analysis of stock replication at a granular level is necessary. Streamlining the inventory management process can help ensure that available films align more closely with customer demand, ultimately leading to improved revenue outcomes.

Recommendations

Sales:

Target non-peak seasons like February, May, and June for promotions and discounts to boost revenue. Create festival-related marketing campaigns to attract more in-store customers.

Customer Segment:

Focus on expanding in countries like the UK, Iran, and Poland, where there's potential for higher purchases despite a smaller customer base.

Store Efficiency:

Store 1 has a larger customer base but lower revenue than Store 2. Enhance Store 1's operational efficiency to convert more customers into higher revenue. For Store 2, expand its customer base through strategic marketing efforts.

Market Expansion:

Explore growth opportunities in Australia, Africa, and South America to tap into untapped markets and increase market share.

Country Revenue:

Concentrate efforts in Australia, Africa, South America, and the former USSR region to target new market segments and enhance global presence.

Store Operations:

Optimize staffing by reducing during non-peak hours and reallocating resources for marketing or other productive tasks.

Inventory Management:

Analyze and streamline inventory management to align stock with customer demand, improving revenue outcomes.