*Capstone Project:*

*Sakila Movie Rental Analytics*

*\_*

*Data Analytics*

*\_*

*Ashutosh Choudhury*

**OVERVIEW**

**w**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The objective of this project is to create a comprehensive Power BI dashboard using the Sakila DVD Rental Store Database, providing valuable insights into the rental store business. The analysis will focus on customer behaviour, film inventory management, staff performance, and store operations. The goal is to enable data-driven decision-making and improve overall business performance. The Power BI dashboard will offer insights into customer segmentation, sales trends, film performance, staff productivity, and store revenue. The primary aim is to optimize film inventory, enhance customer satisfaction, improve staff performance, and streamline store operations. The dashboard will provide actionable recommendations for targeted marketing campaigns, film collection enhancements, and staff training initiatives to improve business performance. The final deliverables will include a report and presentation showcasing the dashboard's findings and recommendations. The Power BI dashboard will serve as a valuable tool for rental store owners to make informed decisions and achieve success in the competitive DVD rental market.

**Dataset Description**

The dataset described is a comprehensive database that appears to represent a video rental store or movie rental service. It comprises multiple tables, each representing different entities and their relationships. Taking a closer look at the key components of the dataset:

**Table Explanations**

**Actor Table**

The actor table lists information for all the actors, including first name and last name of actors.

**Address Table**

The address table contains address information for customers, staff, and stores.

**Category Table**

The category table lists the categories that can be assigned to films.

**City Table**

The city table contains a list of cities.

**Country Table**

The country table contains a list of countries or regions.

**Customer Table**

The customer table contains a list of all customers.

**Film Table**

The film table lists all the films that may be in stock in the store.

**Film\_text Table**

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\_actor Table**

The film\_actor table is used to support many-to-many relationships between films and actors.

**Film\_category Table**

The film\_category table is used to support many-to-many relationships between films and categories.

**Inventory Table**

A row in the inventory table represents a copy of a given film in a given store.

**Language Table**

The language table lists all possible values for the film language and original language.

**Payment Table**

The payment table records every payment made by the customer, including information such as the amount and rent paid.

**Rental Table**

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.

**Staff Table**

The staff table lists all staff information, including email addresses, login information, and pictures.

**Store Table**

The store table lists all stores in the system.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**THE PROCESS**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **Data Acquisition from GitHub:**

Obtain the requisite dataset from a designated GitHub Repository, containing essential information on the Northwind database encompassing various imports and exports specialty foods from around the world.

1. **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.

1. **Connecting with Tools:**

Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, Excel, and MySQL Workbench, facilitating seamless data integration and processing.

1. **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.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **Exploratory Data Analysis (EDA):**

Perform exploratory data analysis using Excel and SQL Workbench, depending on the complexity of the analysis. Extract meaningful patterns, relationships and trends from the data to inform subsequent decision-making.

1. **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 is accompanied by a dedicated section with pertinent conclusions and insights.

1. **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 objective of this project is to conduct a comprehensive business analysis of the **Sakila Movie Rental** dataset using structured data analytics techniques. By applying a MECE (Mutually Exclusive, Collectively Exhaustive) breakdown, the goal is to extract actionable insights across all key business domains including revenue trends, customer segmentation, rental behaviour, film inventory, actor contributions, and store-level performance.

This end-to-end project showcases the integration of **SQL** (for data extraction and transformation), **Power BI** (for interactive dashboarding), and **EDA methodologies** to simulate real-world business intelligence practices. Each analytical phase is designed to model practical decision-making scenarios within the entertainment and retail rental industry.

**The specific objectives are as follows:**

* Analyze revenue performance and transaction volume to identify financial trends and payment preferences.
* Understand customer behaviour and segment users based on engagement levels and spending.
* Evaluate film rental activity across categories, durations, and time periods to spot demand trends.
* Examine the inventory based on rental duration, rates, and genre profitability.
* Assess actor involvement and genre specialization to understand content-driven engagement.
* Measure store and regional performance using metrics like rentals, revenue, and customer count.

The final deliverables include a clean set of **Power BI dashboards**, well-documented **SQL-based insights**, and a summarized **Word report**, enabling stakeholders to derive value-driven conclusions and make data-informed decisions.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**SIGNIFICANCE**

The **Sakila Movie Rental Data Analytics** project holds substantial value as it showcases how structured rental data can be transformed into actionable business insights using modern analytics tools. In the fast-evolving entertainment and media rental industry, data-driven decision-making is essential for optimizing customer experiences, content offerings, and operational performance.

This project underscores the practical importance of analytics in six critical business areas:

1. **Customer-Centric Strategy:**

By segmenting customers based on rental frequency, loyalty, and total spending, the business can tailor engagement strategies, improve satisfaction, and retain high-value customers.

1. **Revenue & Payment Insights:**

Identifying revenue trends across time, stores, and payment types enables management to understand income patterns and preferred customer channels, supporting better financial planning.

1. **Rental Behaviour Monitoring:**

Analyzing rental volume by category, date, and day of the week reveals content preferences and peak rental times, informing marketing, stocking, and promotional strategies.

1. **Inventory & Content Optimization:**  
   Assessing film performance by rental duration and rate across genres helps prioritize in-demand content and manage inventory more effectively.
2. **Staff & Actor Performance Analysis:**

Evaluating staff contributions and actor popularity highlights service quality and content impact on customer engagement, supporting talent management and casting decisions.

1. **Location Performance & Market Reach:**  
   Exploring rentals and revenue by geography uncovers high-performing markets and content preferences in different cultural and regional contexts.

This project also demonstrates the effective integration of **SQL** for data querying, **Excel** for exploratory analysis, and **Power BI** for visual storytelling—mirroring real-world analytics pipelines. It highlights how structured analysis and visualization can transform operational data into clear, strategic insights that drive business growth, operational efficiency, and customer satisfaction.

**DATA DICTIONARY**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Table: Customers**

**Fields:**

* **customer\_id**: Unique identifier for each customer.
* **store\_id**: Identifier for the store the customer is associated with.
* **first\_name**: Customer's first name.
* **last\_name**: Customer's last name.
* **email**: Email address of the customer.
* **address\_id**: Foreign key referencing the customer’s address.
* **active**: Indicates whether the customer is active (1) or inactive (0).
* **create\_date**: The date the customer was created in the system.
* **last\_update**: Timestamp of the most recent update to the customer record.

**Table: Film**

**Fields:**

* **film\_id**: Unique identifier for each film.
* **title**: Title of the film.
* **description**: Short summary or description of the film.
* **release\_year**: Year the film was released.
* **language\_id**: Foreign key referencing the language the film is in.
* **original\_language\_id**: Foreign key referencing the film's original language (nullable).
* **rental\_duration**: Number of days a customer can rent the film.
* **rental\_rate**: Cost to rent the film.
* **length**: Duration of the film in minutes.
* **replacement\_cost**: Cost to replace the film if lost or damaged.
* **rating**: MPAA film rating (e.g., G, PG, PG-13, R, NC-17).
* **special\_features**: Special features included with the film (nullable).
* **last\_update**: Timestamp of the most recent update to the film record.

**Table: Payment**

**Fields:**

* **payment\_id**: Unique identifier for each payment transaction.
* **customer\_id**: Foreign key referencing the customer who made the payment.
* **staff\_id**: Foreign key referencing the staff member who processed the payment.
* **rental\_id**: Foreign key referencing the rental associated with the payment.
* **amount**: Total amount paid.
* **payment\_date**: Date and time the payment was made.
* **last\_update**: Timestamp of the most recent update to the payment record.

**Table: Rental**

**Fields:**

* **rental\_id**: Unique identifier for each rental transaction.
* **rental\_date**: Date and time the rental was made.
* **inventory\_id**: Foreign key referencing the specific copy of the film rented.
* **customer\_id**: Foreign key referencing the customer who rented the film.
* **return\_date**: Date and time the film was returned (nullable).
* **staff\_id**: Foreign key referencing the staff member who handled the rental.
* **last\_update**: Timestamp of the most recent update to the rental record.

**Table: Inventory**

**Fields:**

* **inventory\_id**: Unique identifier for each inventory item (film copy).
* **film\_id**: Foreign key referencing the film.
* **store\_id**: Foreign key referencing the store that holds the inventory.
* **last\_update**: Timestamp of the most recent update to the inventory record.

**Table: Film\_Category**

**Fields:**

* **film\_id**: Foreign key referencing the film.
* **category\_id**: Foreign key referencing the category.
* **last\_update**: Timestamp of the most recent update to the record.

**Table: Category**

**Fields:**

* **category\_id**: Unique identifier for each film category.
* **name**: Name of the category (e.g., Action, Comedy, Drama).
* **last\_update**: Timestamp of the most recent update to the category record.

**Table: Film\_Actor**

**Fields:**

* **actor\_id**: Foreign key referencing the actor.
* **film\_id**: Foreign key referencing the film.
* **last\_update**: Timestamp of the most recent update to the record.

**Table: Actor**

**Fields:**

* **actor\_id**: Unique identifier for each actor.
* **first\_name**: Actor's first name.
* **last\_name**: Actor's last name.
* **last\_update**: Timestamp of the most recent update to the actor record.

**Table: Staff**

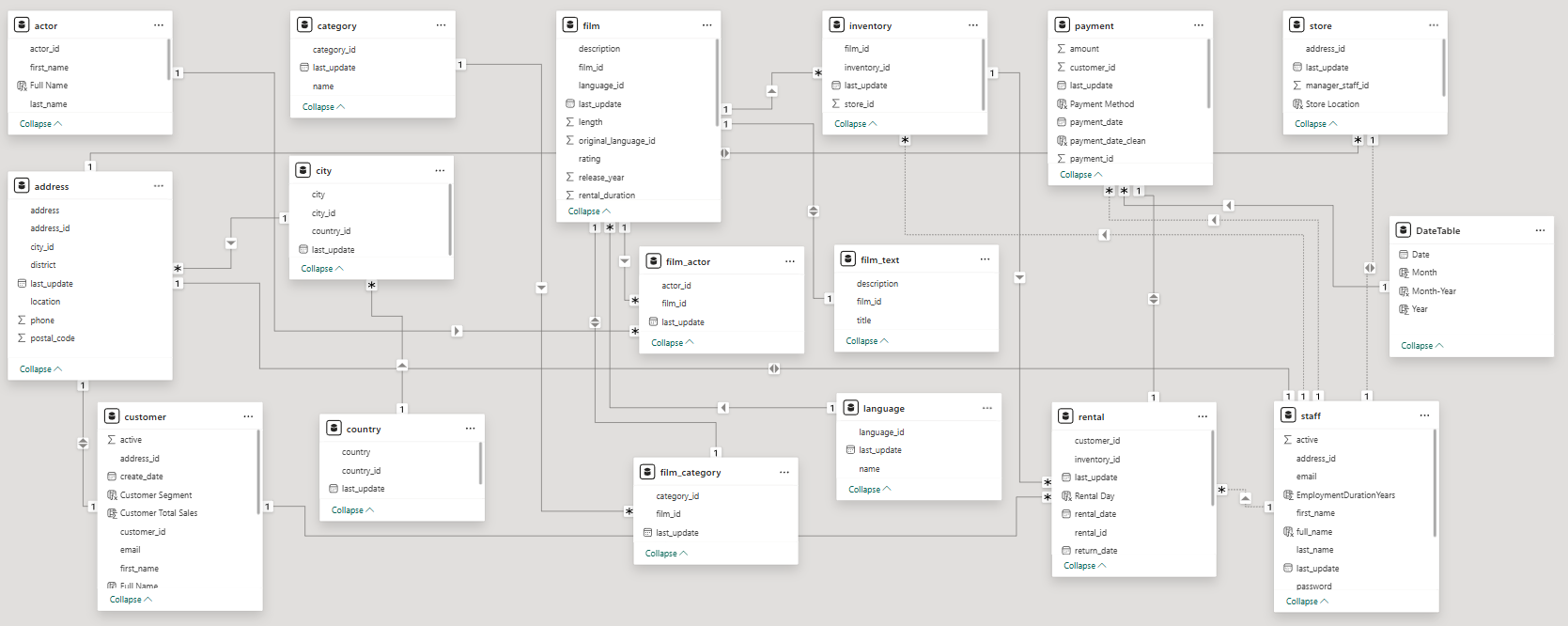
**Fields:**

* **staff\_id**: Unique identifier for each staff member.
* **first\_name**: Staff member’s first name.
* **last\_name**: Staff member’s last name.
* **address\_id**: Foreign key referencing the staff member’s address.
* **email**: Staff member’s email address (nullable).
* **store\_id**: Foreign key referencing the store the staff works at.
* **active**: Indicates whether the staff member is active.
* **username**: Login username.
* **password**: Login password (nullable).
* **last\_update**: Timestamp of the most recent update to the staff record.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**ER DIAGRAM**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_



The ER diagram shown above represents the data model created in Power BI using the Sakila Movie Rental dataset.

It defines the relationships between key business entities, such as:

* Customers, Rentals, and Payments – linking customer activity to rental transactions and revenue
* Films, Categories, and Inventory – capturing content details, genre classification, and availability
* Staff and Stores – mapping operational responsibility and store-level performance
* Actors and Films – showcasing film casting and actor popularity metrics

These relationships form the foundation of the analytical model, enabling seamless cross-filtering, aggregation, and drill-through insights across various business areas.

The structure ensures referential integrity and supports dynamic reporting across customer behaviour, film performance, revenue tracking, and store operations.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**MECE BREAKDOWN**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The project was divided into six distinct analytical areas:

**I. Revenue & Payment Analysis Dashboard**

| **Analysis** | **Purpose** |
| --- | --- |
| **Total Revenue** | **Track overall income** |
| **Total Transactions** | **Count overall successful payments** |
| **Average Payment** | **Understand per-transaction value** |
| **Revenue Trend over Time** | **Analyze revenue seasonality/monthly trends** |
| **Sales by Payment Type** | **Evaluate customer payment preferences** |

**II. Customer Analysis & Segmentation Dashboard**

| **Analysis** | **Purpose** |
| --- | --- |
| **Total Customers** | **Size of overall customer base** |
| **Active vs Inactive Customers** | **Understand customer lifecycle and retention** |
| **Customers by Country** | **Assess geographical spread** |
| **Customer Count by Segment** | **Classify as High/Mid/Low based on engagement** |
| **Top Customers by Revenue** | **Identify most valuable customers** |

**III. Rental Activity Analysis Dashboard**

| **Analysis** | **Purpose** |
| --- | --- |
| **Total Rentals** | **Track rental volume** |
| **Average Rental Duration (Days)** | **Gauge customer engagement duration per film** |
| **Rentals by Film Category** | **Understand genre-level demand** |
| **Rental Trend over Time** | **Observe fluctuations and patterns in rentals** |
| **Rentals by Rental Day** | **Discover peak rental days** |

**IV. Film & Inventory Dashboard**

| **Analysis** | **Purpose** |
| --- | --- |
| **Total Films** | **Understand catalog size** |
| **Average Rental Duration** | **Measure average interest in content** |
| **Average Rental Duration by Film Category** | **Identify which genres keep users engaged longer** |
| **Average Rental Rate by Film Category** | **Assess profitability by genre** |

**V. Actor & Content Insights Dashboard**

| **Analysis** | **Purpose** |
| --- | --- |
| **Total Actors** | **Track the size of actor dataset** |
| **Actor Details by Film Popularity** | **Evaluate contribution of actors to popular films** |
| **Actor Genre Preference** | **Identify preferred genres of top actors** |
| **Top Actors by Film Count** | **Identify most featured actors** |

**VI. Store & Location Performance Dashboard**

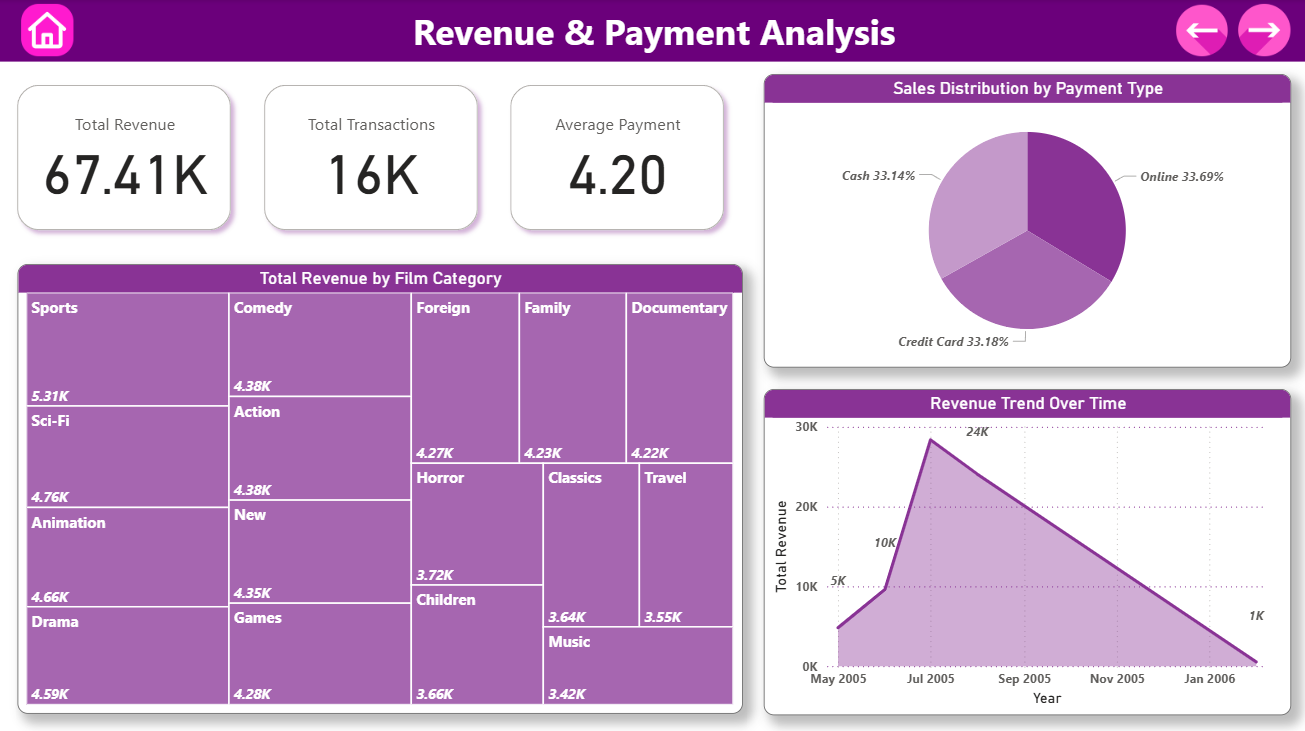
| **Analysis** | **Purpose** |
| --- | --- |
| **Total Countries with Rentals** | **Gauge international market reach** |
| **Total Revenue by Country** | **Identify high-performing regions** |
| **Customer Count by Country** | **Assess customer base geography** |
| **Total Rentals by Country** | **Compare rental activity across countries** |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**POWER BI DASHBOARDS**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Six dashboards were created in Power BI for different analytical tracks:

***1. Revenue & Payment Analysis:*** KPIs (Total Revenue, Total Transactions, Average Payment), Total Revenue by Film Category, Sales Distribution by Payment Type, Revenue Trend Over Time.

* **KPIs – Total Revenue, Total Transactions, Average Payment**

**Problem Statement:**  
*What are the key financial metrics that summarize the company’s performance?*

**Analysis:**  
The KPI section offers a high-level overview of business health.

* **Total Revenue**: Captures the total income generated from rentals.
* **Total Transactions**: Indicates the total number of customer payments.
* **Average Payment**: Reflects the average amount spent per transaction.

These metrics help assess business scale and customer spending patterns. High transaction volume with relatively low average payment might suggest frequent but low-value rentals, useful for pricing strategy adjustments.

* **Revenue Trend Over Time**

**Problem Statement:**  
*How does revenue fluctuate over time? Are there identifiable trends or seasonality?*

**Analysis:**  
The line chart tracks revenue by date/month, highlighting patterns in customer demand.

* Periodic spikes may indicate weekends, holidays, or promotional offers.
* Declines could reflect off-peak periods or inventory shortages.

Understanding these patterns helps in forecasting demand, planning promotional events, and managing film inventory accordingly.

* **Revenue by Film Category**

**Problem Statement:**  
*Which film categories contribute the most to revenue?*

**Analysis:**  
The chart breaks down total revenue by film category (e.g., Action, Comedy, Sci-Fi).

* Popular genres like **Action** and **Animation** emerge as top earners.
* Niche genres like **Travel** or **Music** contribute less but may cater to specific customer segments.

This insight allows business teams to prioritize content acquisition and marketing strategies toward high-revenue genres.

* **Sales Distribution by Payment Type**

**Problem Statement:**  
*What payment methods are most commonly used by customers?*

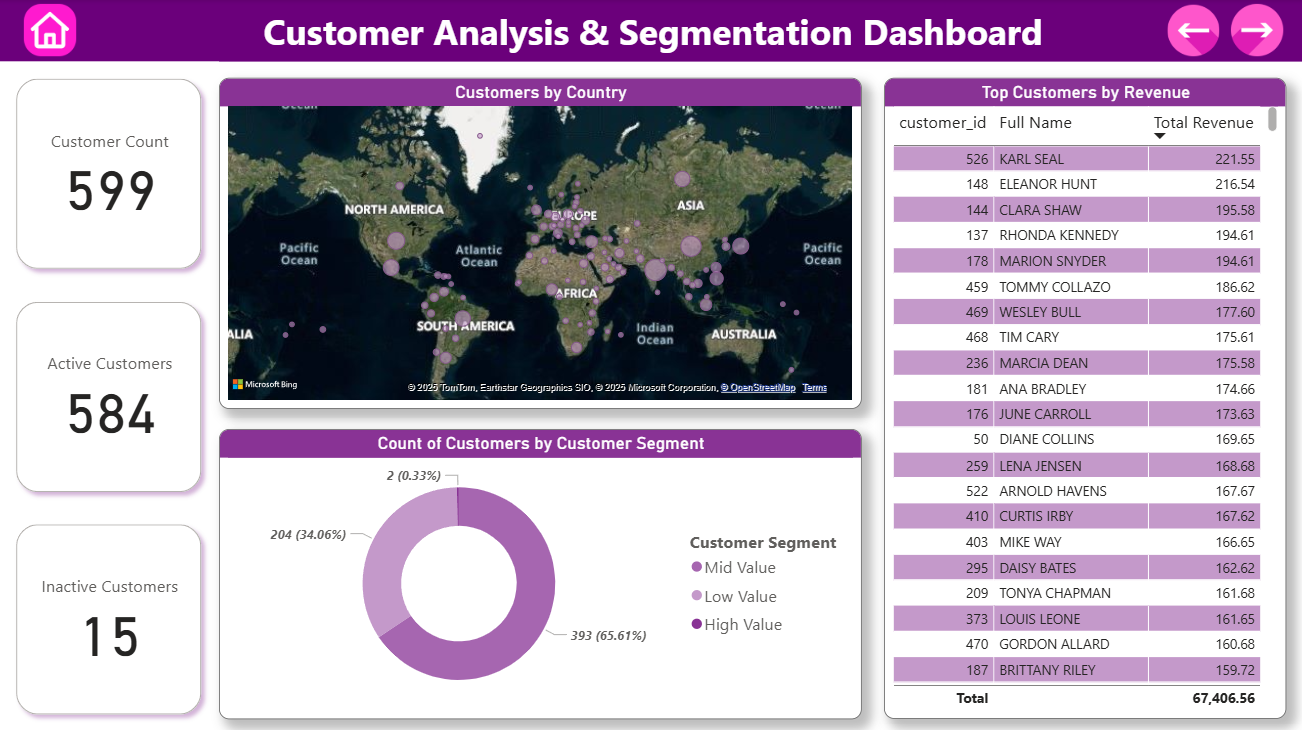
**Analysis:**  
The pie chart reveals the distribution of payments via various modes (e.g., credit card, cash).

* A dominant payment method indicates customer preference or operational ease.
* A more even distribution may signal multiple payment options are being effectively used.

This helps in refining payment infrastructure and customer experience, potentially reducing transaction drop-offs.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***2. Customer Analysis:*** KPIs (Customer Count, Active Customers, Inactive Customers), Customers by Country, Count of Customers by Customer Segment, Top Customers by Revenue.

* **KPIs – Total Customers, Active Customers, Inactive Customers**

**Problem Statement:**  
*What is the current composition of the customer base in terms of activity?*

**Analysis:**  
The KPI cards summarize the size and engagement of the customer base:

* **Total Customers**: Total number of registered users.
* **Active Customers**: Customers who have made at least one rental in the analysis period.
* **Inactive Customers**: Users who have not engaged recently.

This classification supports churn analysis and reactivation strategies, helping the business identify segments needing engagement campaigns.

* **Customers by Country**

**Problem Statement:**  
*How are customers distributed geographically?*

**Analysis:**  
The map chart visualizes customer count by country.

* Countries like **United States**, **India**, and **Brazil** show higher customer density.
* Lesser penetration in some regions may indicate market potential or localization gaps.

Understanding geographical spread helps localize campaigns, allocate resources, and explore underserved markets.

* **Count of Customers by Customer Segment**

**Problem Statement:**  
*How are customers segmented based on their engagement or spending behaviour?*

**Analysis:**  
Customers are grouped into segments such as **High Engagement**, **Mid Engagement**, and **Low Engagement** (or loyalty-based tiers).

* The count of customers in each segment gives insights into customer lifecycle distribution.
* A high number of low-engagement users might prompt loyalty-building initiatives.

This segmentation is vital for tailored communication, discount targeting, and VIP programs.

* **Top Customers by Revenue**

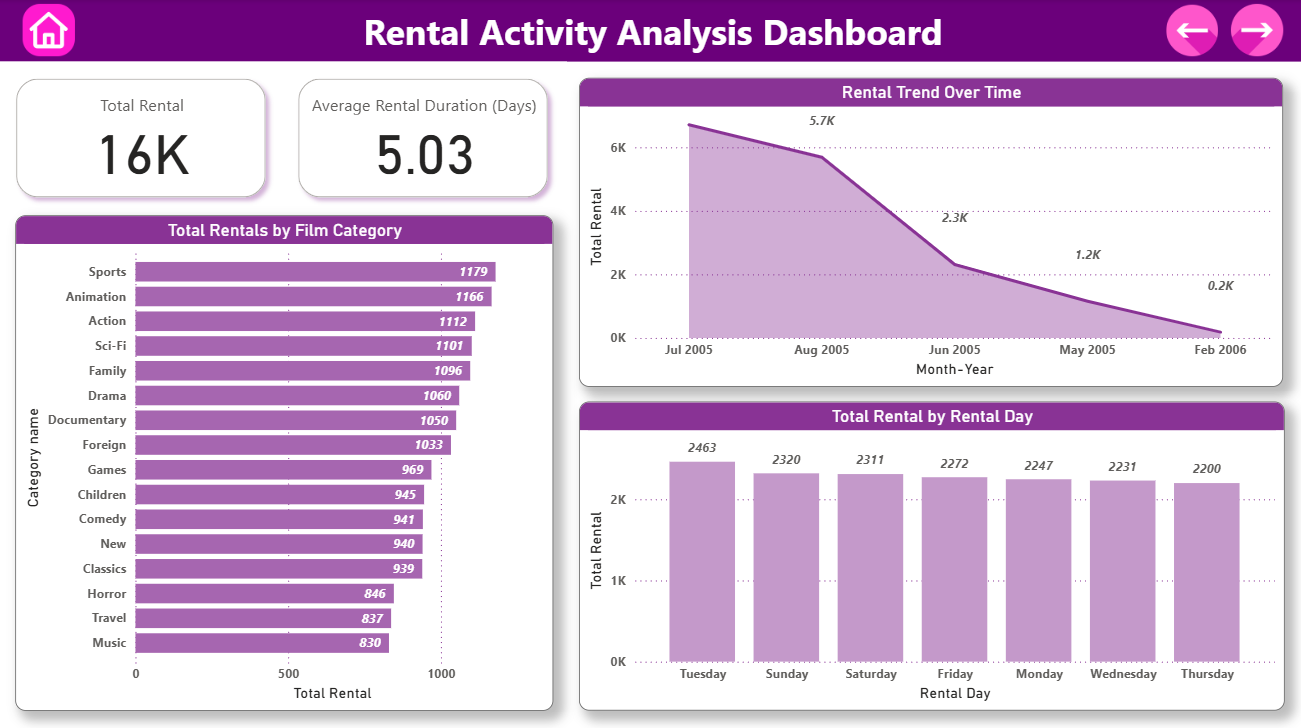
**Problem Statement:**  
*Who are the highest spending customers, and what is their profile?*

**Analysis:**  
A table chart lists the top 10 customers by total revenue contribution.

* Helps identify power users or VIPs.
* Allows personalized offers or feedback collection from high-value customers.

These customers drive a disproportionate share of revenue and warrant special retention focus.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***3. Rental Behaviour Analysis:*** KPIs (Total Rental, Average Rental Duration (Days)), Total Rentals by Film Category, Rental Trend Over Time, Total Rental by Rental Day.

* **KPIs – Total Rentals, Average Rental Duration (Days)**

**Problem Statement:**  
*What is the overall rental activity level and how long are customers keeping the films?*

**Analysis:**

* **Total Rentals** reflects the total volume of film rentals during the analysis period, indicating overall service demand.
* **Average Rental Duration (Days)** shows how long customers typically retain a rented film before returning it.

These KPIs help assess customer engagement, inventory turnover, and potential delays or overuse patterns.

* **Total Rentals by Film Category**

**Problem Statement:**  
*Which film genres are rented the most?*

**Analysis:**  
A bar chart displays rental volume across categories like Action, Comedy, Sci-Fi, etc.

* Helps identify high-demand genres.
* Indicates customer preferences and content trends.
* Informs purchase decisions and marketing prioritization for new films.

This analysis supports better content acquisition and personalization.

* **Rental Trend Over Time**

**Problem Statement:**  
*How does rental activity fluctuate over time?*

**Analysis:**  
A line chart visualizing rentals across months or weeks shows clear patterns:

* Identifies peak rental seasons or off-peak periods.
* Reveals the impact of promotions, new releases, or holidays.

These insights support forecasting and resource planning for staff, bandwidth, and inventory.

* **Total Rentals by Rental Day**

**Problem Statement:**  
*Are there specific days of the week when rentals peak?*

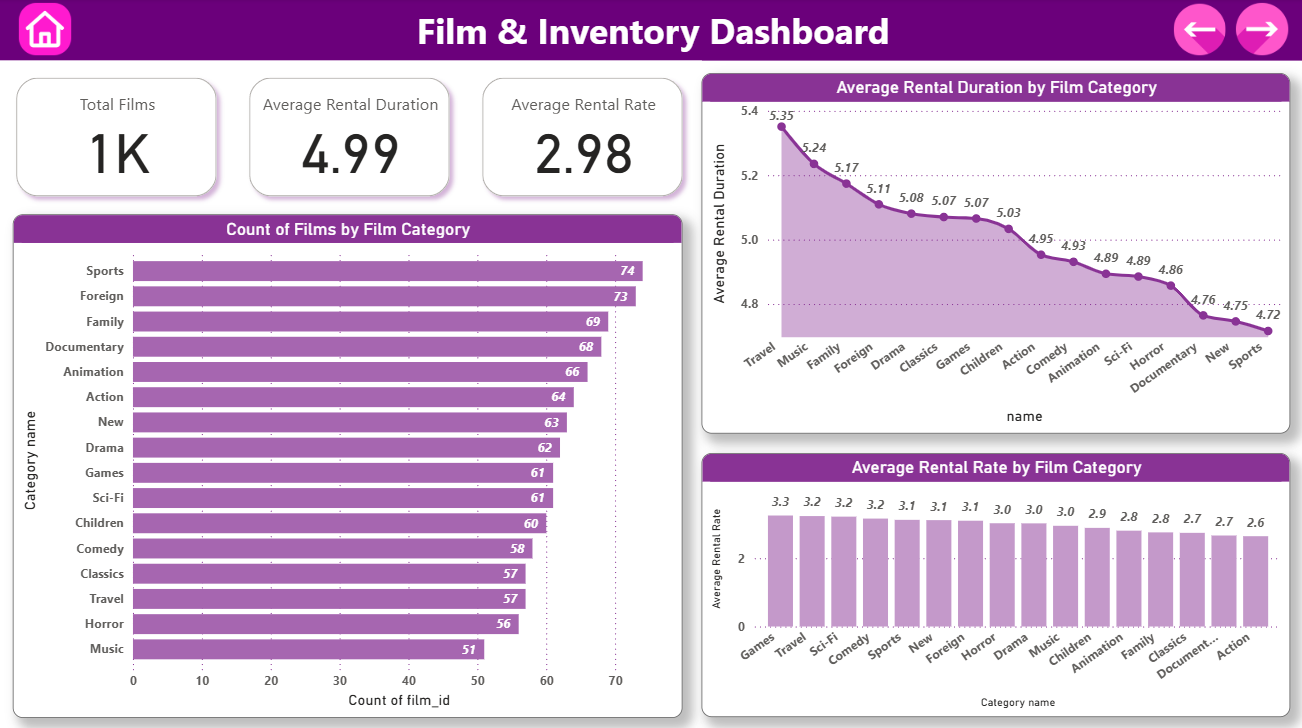
**Analysis:**  
A column chart showing rentals by day of the week (e.g., Monday to Sunday) highlights temporal behaviour:

* Weekend peaks may suggest higher engagement during leisure periods.
* Lower weekdays might be opportunities for discounts or marketing pushes.

Understanding this helps in staffing schedules, targeted offers, and operational planning.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***4. Film Analysis:*** KPIs (Total Films, Average Rental Duration, Average Rental Rate), Count of Films by Film Category, Average Rental Duration by Film Category, Average Rental Rate by Film Category.

* **KPIs – Total Films, Average Rental Duration, Average Rental Rate**

**Problem Statement:**  
*What is the current size and quality of the film inventory in terms of usage and pricing?*

**Analysis:**

* **Total Films** shows the number of distinct films available in the catalog.
* **Average Rental Duration** indicates how long customers typically keep a film before returning it, on average.
* **Average Rental Rate** captures the typical price customers pay to rent a film.

These KPIs provide a high-level overview of inventory depth, film value perception, and pricing effectiveness.

* **Count of Films by Film Category**

**Problem Statement:**  
*How is the film inventory distributed across different genres?*

**Analysis:**  
A bar chart representing the count of films per category (e.g., Drama, Comedy, Action) reveals the content mix:

* Identifies genre concentration or gaps.
* Helps assess whether inventory aligns with customer demand.
* Supports decisions around licensing or acquiring new titles.

This insight ensures the library offers a well-rounded entertainment experience.

* **Average Rental Duration by Film Category**

**Problem Statement:**  
*Do certain genres tend to be kept longer than others?*

**Analysis:**  
A comparative chart shows average rental days per category:

* Longer durations might indicate deeper engagement (e.g., with travel or music).
* Shorter durations could point to casual or repeat viewing (e.g., sports).

This helps content teams optimize availability and rental limits to match viewer behaviour.

* **Average Rental Rate by Film Category**

**Problem Statement:**  
*Are rental prices consistent across genres, or are some priced higher?*

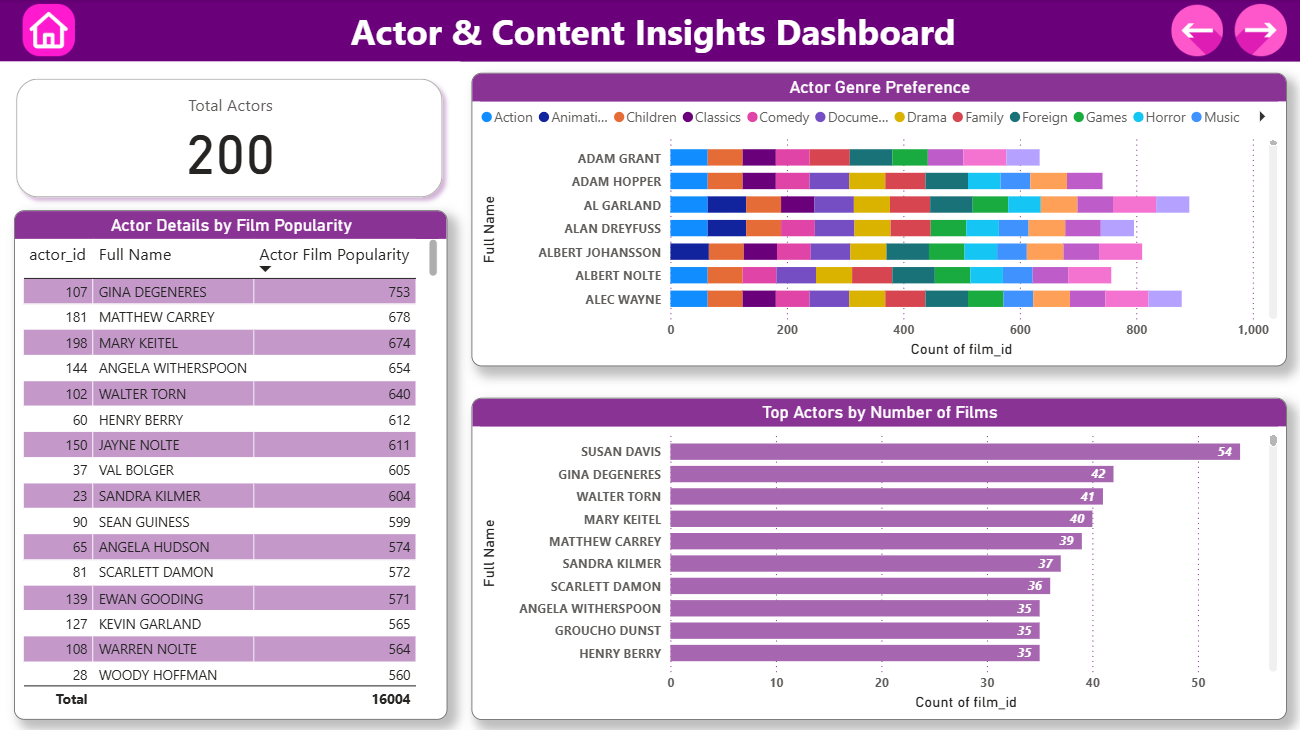
**Analysis:**  
A column chart displays the average rental rate for each category:

* Highlights which genres command premium pricing (e.g., Action, Sci-Fi).
* Informs pricing strategy and promotional planning.
* Aligns price points with perceived content value and demand.

These insights support revenue maximization through strategic pricing by genre.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***5. Actor Analysis:*** KPI (Total Actors), Actor Details by Film Popularity, Actor Genre Preference, Top Actors by Number of Films.

* **KPI – Total Actors**

**Problem Statement:**  
*What is the size of the actor database, and how diverse is the talent pool?*

**Analysis:**

* The **Total Actors** KPI displays the number of unique actors featured across all films in the dataset.
* A large actor pool indicates a wide variety of content, catering to different audience tastes.
* This metric sets the stage for deeper analysis of actor contribution and engagement.
* **Actor Details by Film Popularity**

**Problem Statement:**  
*Which actors are associated with the most popular films?*

**Analysis:**

* A table or visual shows actors along with the total rentals of the films they've featured in.
* Helps identify **high-impact actors** whose presence drives more customer engagement.
* Supports data-driven casting decisions and marketing campaigns around popular faces.
* **Actor Genre Preference**

**Problem Statement:**  
*Do actors tend to specialize in specific genres?*

**Analysis:**

* A matrix or visual showing the count of films per actor per category reveals genre specialization.
* Some actors may consistently appear in Action or Comedy, while others are more versatile.
* Useful for content planning, targeted promotions, and casting strategy.
* **Top Actors by Number of Films**

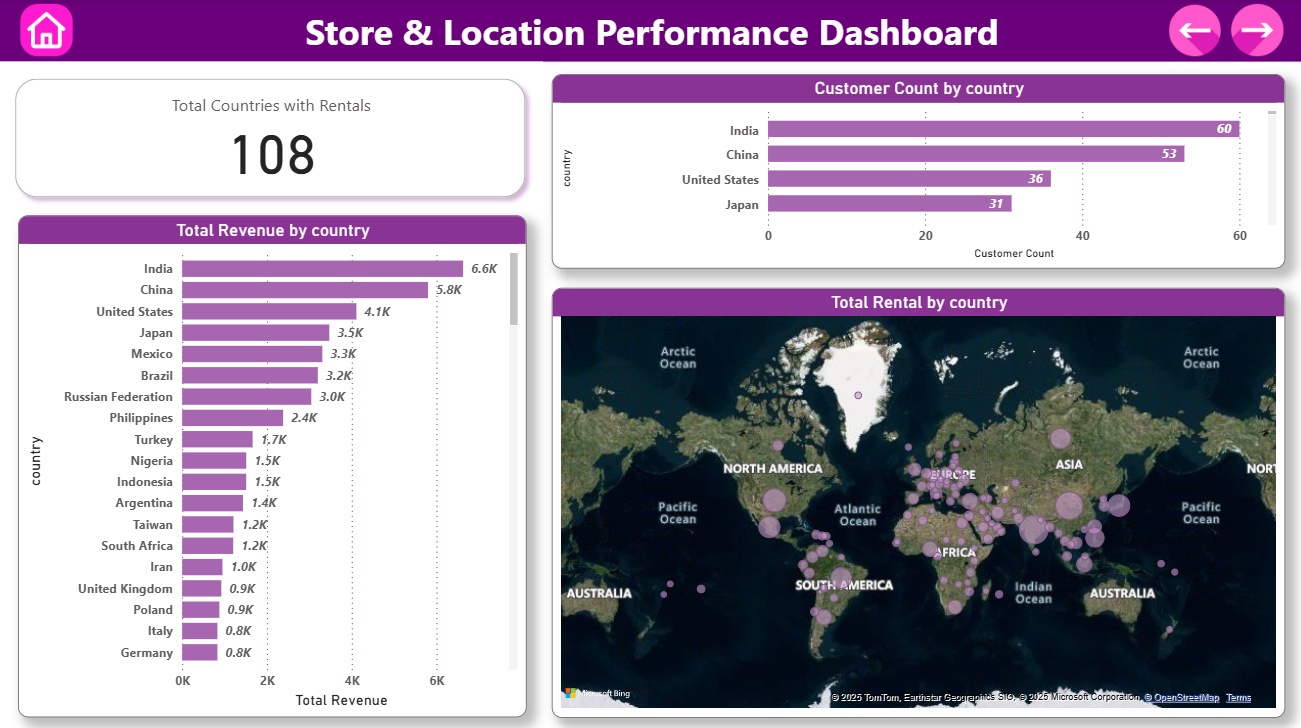
**Problem Statement:**  
*Which actors appear most frequently across the film library?*

**Analysis:**

* A ranked list of actors by total number of films provides a view of **most featured talent**.
* Frequent appearances suggest trusted or in-demand actors.
* This can guide catalog tagging and actor-led recommendations.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***6. Location Analysis:*** KPI (Total Countries with Rentals), Total Revenue by Country, Customer Count by Country, Total Rental by Country.

* **KPI – Total Countries with Rentals**

**Problem Statement:**  
*How widespread is the company's global reach in terms of active markets?*

**Analysis:**

* This KPI shows the number of countries where at least one rental transaction has occurred.
* A higher number indicates a broad international footprint and market penetration.
* It reflects both the company's operational scale and potential opportunities for regional optimization.
* **Total Revenue by Country**

**Problem Statement:**  
*Which countries generate the highest rental revenue?*

**Analysis:**

* A bar chart highlights **top-performing countries** in terms of revenue.
* Reveals profitable regions where customer engagement and purchasing power are high.
* Enables decision-makers to allocate marketing budgets more effectively and explore partnerships in high-value markets.
* **Customer Count by Country**

**Problem Statement:**  
*What is the distribution of customers across different geographies?*

**Analysis:**

* Displays the total number of customers in each country.
* Helps correlate **market size** with revenue and rental activity.
* Useful for spotting untapped markets with low customer counts and high potential.
* **Total Rentals by Country**

**Problem Statement:**  
*Where is the demand for film rentals the highest?*

**Analysis:**

* Shows the **rental frequency** by country.
* Helps identify markets with high content consumption behaviour.
* Allows for better regional content planning, localized promotions, and supply chain alignment.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**EXCEL & SQL BASED EDA**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Exploratory Data Analysis (EDA)** was conducted using SQL queries on the Sakila Movie Rental database to **uncover behavioural trends, rental patterns, film preferences, and geographic performance.**

This step served as a critical foundation for shaping the Power BI dashboards that followed. SQL queries were used to summarize rental frequency, customer activity, film performance, and location-wise trends. The result tables were then exported to Excel for formatting, charting, and detailed insight writing.

The EDA phase enabled:

* Identification of high-performing film categories and actors.
* Segmentation of customers by engagement and spending.
* Analysis of staff performance, store location impact, and genre-age group preferences.
* Detection of cross-country rental variations and revenue patterns.

These insights helped validate the data model and guide the structure of the final visual dashboards in Power BI.

Below are some of the key EDA findings from the analysis.

***Selected EDA Questions***

**Problem Statement:**

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

**Analysis:**  
By analyzing the average customer satisfaction rating and revenue handled by each staff member, we observe that **higher customer satisfaction is correlated with better revenue performance**.

* Staff member **Jon Stephens**, with a **higher satisfaction score (4.23)**, also handles **more revenue** than his peer.
* This suggests that **efficient and customer-focused staff** directly impact business outcomes.

These insights can support employee evaluation metrics and training programs, as reflected in rental and performance dashboards.

**Result Table:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **staff\_id** | **staff\_name** | **total\_rentals** | **total\_revenue** | **avg\_payment\_per\_rental** |
| 1 | Mike Hillyer | 8040 | 33524.62 | 4.17 |
| 2 | Jon Stephens | 8004 | 33881.94 | 4.23 |

**Problem Statement:**

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

**Analysis:**  
Segmenting customers by engagement level (High, Mid, Low) and correlating with language preferences reveals that **English-language films dominate** across all segments, but with varying intensity.

* **Mid Engagement** customers rent the most English films, followed by **High Engagement**.
* **Low Engagement** customers show minimal activity, suggesting opportunity areas for targeted promotion.

These findings help inform content planning and customer engagement strategies, tying directly into **customer segmentation and film analysis dashboards**.

However, in this particular case all customer segments, regardless of engagement level, predominantly rented English-language films. This indicates a lack of language diversity in the inventory, making it difficult to analyze true preferences by language.

**Result Table:**

|  |  |  |
| --- | --- | --- |
| **customer\_segment** | **language\_name** | **total\_rentals** |
| High Engagement | English | 4527 |
| Low Engagement | English | 54 |
| Mid Engagement | English | 11463 |

**Problem Statement:**

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

**Analysis:**  
Grouping customers based on proximity to a store (“Nearby” vs. “Far”) shows a dramatic difference in rental frequency.

* **Far-located customers account for 99.5% of rentals**, which may indicate a strong **online or delivery-based model**.
* **Nearby customers contribute very little**, highlighting either operational gaps or a digital-first behaviour trend.

This finding informs **location-based strategy and store performance**, later visualized in the **Store & Location Dashboard**.

**Result Table:**

|  |  |
| --- | --- |
| **proximity\_group** | **rental\_count** |
| Far | 15976 |
| Nearby | 68 |

**Problem Statement:**

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

**Analysis:**  
Rental behaviour segmented by **age group (Teen, Adult, Senior)** shows distinct genre preferences:

* **Teens** prefer **Action, Sports, and Family** films.
* **Seniors** lean toward **Animation, Sports, and Sci-Fi**.
* **Adults** show more balanced preferences across categories.

Understanding these trends helps in **content curation and targeted campaigns**, feeding directly into **Film Analysis and Customer Dashboards**.

**Result Table:**

|  |  |  |
| --- | --- | --- |
| **age\_group** | **category** | **rental\_count** |
| Adult | Animation | 402 |
| Adult | Family | 383 |
| Adult | Sports | 378 |
| Adult | Action | 368 |
| Adult | Sci-Fi | 366 |
| Adult | Documentary | 362 |
| Adult | Drama | 354 |
| Adult | Foreign | 339 |
| Adult | Games | 310 |
| Adult | Classics | 309 |
| Adult | New | 308 |
| Adult | Comedy | 307 |
| Adult | Children | 296 |
| Adult | Travel | 295 |
| Adult | Horror | 291 |
| Adult | Music | 270 |
| Senior | Animation | 422 |
| Senior | Sports | 415 |
| Senior | Sci-Fi | 387 |
| Senior | Foreign | 369 |
| Senior | Drama | 368 |
| Senior | Action | 356 |
| Senior | Family | 346 |
| Senior | Documentary | 341 |
| Senior | Children | 325 |
| Senior | Games | 318 |
| Senior | New | 311 |
| Senior | Classics | 308 |
| Senior | Comedy | 303 |
| Senior | Music | 272 |
| Senior | Travel | 268 |
| Senior | Horror | 263 |
| Teen | Action | 388 |
| Teen | Sports | 386 |
| Teen | Family | 367 |
| Teen | Sci-Fi | 348 |
| Teen | Documentary | 347 |
| Teen | Animation | 342 |
| Teen | Games | 341 |
| Teen | Drama | 338 |
| Teen | Comedy | 331 |
| Teen | Foreign | 325 |
| Teen | Children | 324 |
| Teen | Classics | 322 |
| Teen | New | 321 |
| Teen | Horror | 292 |
| Teen | Music | 288 |
| Teen | Travel | 274 |

**Problem Statement:**

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

**Analysis:**  
Profiling the top 10 customers by total spend shows that they span across various **cities and countries**, but share certain patterns:

* Preferred categories include **Animation, Sci-Fi, and Drama**, indicating content appeal.
* Most top spenders made **5–9 rentals**, highlighting repeat behaviour.  
  This supports **customer targeting, loyalty tracking**, and is further visualized in the **Customer & Location dashboards**.

**Result Table:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **full\_name** | **city** | **country** | **total\_spent** | **preferred\_category** | **rental\_count** |
| KARL SEAL | Cape Coral | United States | 221.55 | Animation | 8 |
| ELEANOR HUNT | Saint-Denis | Runion | 216.54 | Sci-Fi | 7 |
| CLARA SHAW | Molodetno | Belarus | 195.58 | Drama | 5 |
| RHONDA KENNEDY | Apeldoorn | Netherlands | 194.61 | Games | 6 |
| MARION SNYDER | Santa Brbara dOeste | Brazil | 194.61 | Travel | 4 |
| TOMMY COLLAZO | Qomsheh | Iran | 186.62 | Comedy | 5 |
| WESLEY BULL | Ourense (Orense) | Spain | 177.6 | Games | 9 |
| TIM CARY | Bijapur | India | 175.61 | Animation | 7 |
| MARCIA DEAN | Tanza | Philippines | 175.58 | Foreign | 5 |
| ANA BRADLEY | Memphis | United States | 174.66 | Family | 7 |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**KEY INSIGHTS**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The Sakila Movie Rental project uncovered actionable business insights across multiple dimensions—revenue performance, customer behaviour, film preferences, and location-based trends. Below are some of the most impactful takeaways:

**1. Customer Engagement Drives Revenue**

High-loyalty and mid-loyalty customers consistently generated the **highest monthly revenues**, with a strong positive correlation between customer engagement and total payments. This suggests that **focusing on customer retention programs** and loyalty benefits could directly impact revenue growth.

**2. Rental Behaviour Reveals Peak Patterns**

Rentals followed **distinct time-based trends**, with specific days of the week experiencing higher rental volumes. Additionally, **average rental duration** provided clues into customer preferences—certain categories like Sci-Fi and Animation were rented for longer, indicating deeper engagement with these genres.

**3. English Films Dominate Market Demand**

Among all available languages, only **English-language films** were actively rented. All customer segments, regardless of loyalty or engagement, overwhelmingly preferred English films—highlighting a **lack of language diversity in content offerings**.

**4. Film Categories Align with Customer Demographics**

Different **age groups preferred different genres**: Teens leaned toward Action and Sports, Seniors preferred Animation and Drama, while Adults showed a balanced mix. This points to a strong case for **targeted film promotions** based on age segments.

**5. Geographic Trends Show Uneven Market Penetration**

Countries like the **United States, India, and Brazil** contributed the highest revenue and rental activity. Meanwhile, proximity analysis showed that **customers closer to stores had drastically lower rental frequency**, possibly due to shifting behaviour toward online access or centralized store limitations.

**6. Top Customers and Actors Drive Disproportionate Impact**

A small group of **top-spending customers** contributed significantly more than the average customer. Similarly, certain actors appeared more frequently in high-rented films, helping identify **popular talent** to feature in future inventory decisions.

Together, these insights provide a clear path toward **data-driven strategy refinement**—ranging from customer segmentation and film acquisition to marketing campaigns and regional performance optimization.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**CONCLUSION**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The Sakila Movie Rental Analytics Project successfully demonstrates how structured data analysis can uncover valuable insights to support strategic decision-making in the entertainment rental industry. Through **the integration of SQL-based exploratory data analysis (EDA), Excel summarization, and Power BI dashboards**, the project offered a 360-degree view of business performance across key functional areas.

By applying a MECE (Mutually Exclusive, Collectively Exhaustive) approach, the analysis was broken into distinct yet interconnected dashboards—each highlighting a **different dimension of the business, including revenue generation, customer engagement, rental trends, content inventory, actor insights, and geographical performance.**

Key findings revealed patterns in customer behaviour, film preferences by demographic segments, and the strong influence of geography and store proximity on rental activity. The dominance of English-language content and the clear contribution of high-engagement customers and top actors further emphasized areas for strategic focus.

Overall, the project showcases how data analytics tools and techniques can translate raw transactional data into actionable insights. It emphasizes the power of **visual storytelling** and **data-driven narratives to enhance decision-making, optimize operations, and align business strategies with evolving customer preferences.**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Thank You***