CIS 467
Data Management, Warehousing, and Visualization
Spring 2025

Customer Revenue Analysis

Shan Ali Shah Sayed
Thursday, March 5, 2025

Table of Contents

Overview

This project analyzes customer behavior and revenue trends for a DVD rental company. The goal is to build a Customer Revenue Data Warehouse that consolidates data from multiple operational tables into a single, summarized view.

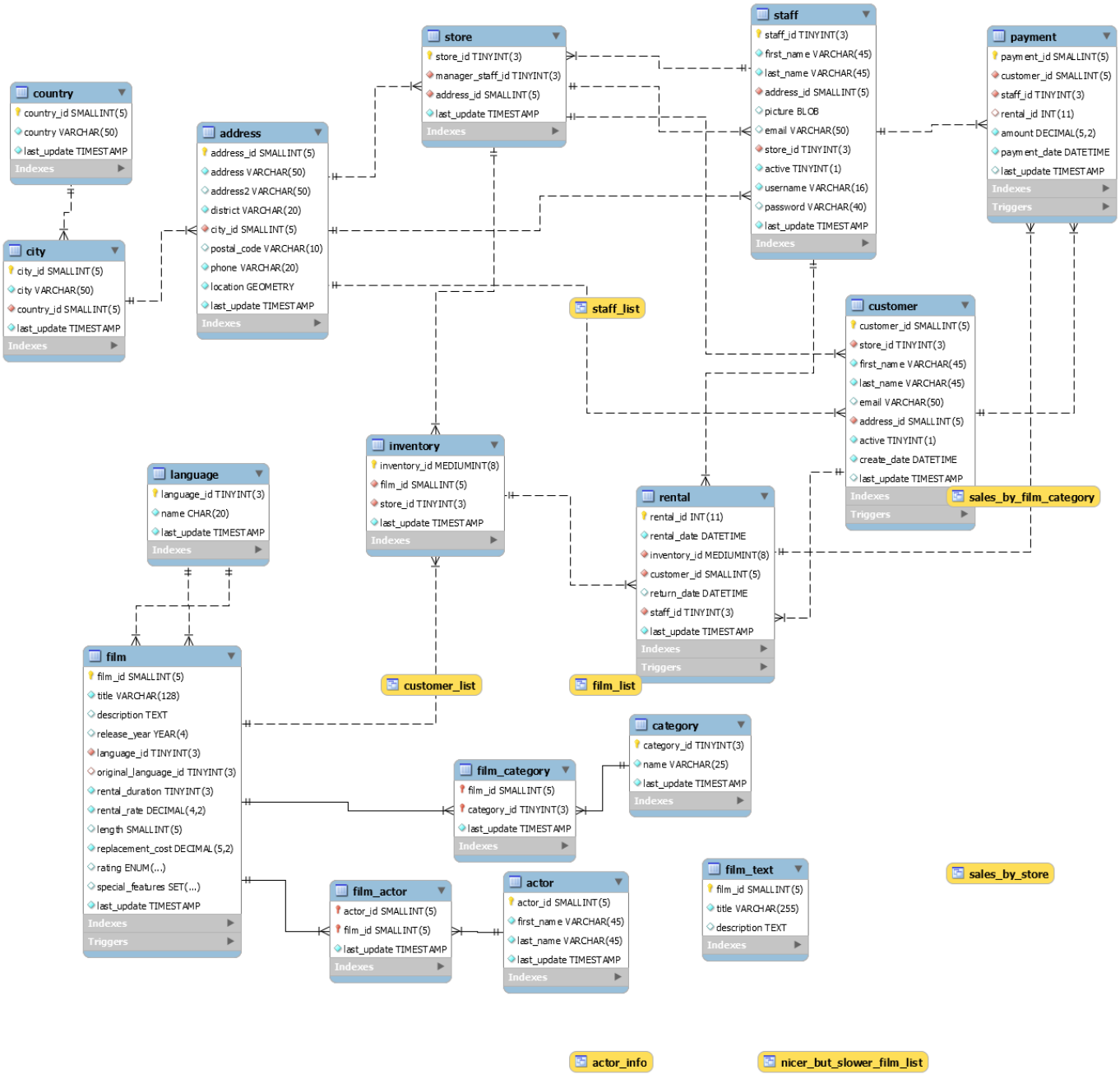
The resulting dataset (customer_revenue_summary) enables management to:

- Identify **top customers** for loyalty and retention programs
- Understand **regional spending patterns** to tailor marketing
- Analyze **genre-level performance** to guide inventory acquisition
- Detect **inactive customers** for re-engagement campaigns
- Recognize **seasonal and peak trends** for operational planning

This data-driven approach supports informed decision-making that directly impacts revenue, customer satisfaction, and long-term growth.

Data Warehouse: Customer Revenue

The below database was created using the sakila-schema.sql and sakila-data.sql scripts, which define the structure and populate it with sample transactional data related to company operations:



Customer Revenue Analytical View

The Customer Revenue Analytical View is designed to help the DVD rental company gain deeper insights into its customers' rental behaviors and revenue contributions. By aggregating transactional data from multiple operational tables in the Sakila database, this view provides a consolidated summary of customer preferences, revenue patterns, and overall rental trends. These insights enable management to make informed, data-driven decisions on marketing strategies, customer retention, and pricing optimization.

SQL:

```
-- DATA WAREHOUSE

USE sakila;

DROP VIEW IF EXISTS customer_revenue_summary;

CREATE VIEW customer_revenue_summary AS

SELECT

   c.customer_id AS Customer_ID,

   CONCAT(c.first_name, ' ', c.last_name) AS Customer_Name,

   COUNT(r.rental_id) AS Total_Rentals,

   ROUND(SUM(p.amount), 2) AS Total_Revenue,

   ROUND(SUM(p.amount) / COUNT(r.rental_id), 2) AS Avg_Spending_Per_Rental,

   COUNT(DISTINCT fc.category_id) AS Total_Categories_Rented,

   (SELECT name

    FROM category

    JOIN film_category fc ON category.category_id = fc.category_id

    JOIN inventory i ON fc.film_id = i.film_id

    JOIN rental r2 ON i.inventory_id = r2.inventory_id

    WHERE r2.customer_id = c.customer_id

    GROUP BY category.name

    ORDER BY COUNT(*) DESC
```

```
     LIMIT 1) AS Most_Frequent_Category,
    ROUND(COALESCE(SUM(CASE
        WHEN DATEDIFF(r.return_date, r.rental_date) > f.rental_duration
        THEN (DATEDIFF(r.return_date, r.rental_date) - f.rental_duration) * 0.50
        ELSE 0 END), 0), 2) AS Total_Late_Fees,
    (SELECT MIN(rental_date) FROM rental WHERE customer_id = c.customer_id) AS First_Rental_Date,
    (SELECT MAX(rental_date) FROM rental WHERE customer_id = c.customer_id) AS Last_Rental_Date,
    COUNT(p.payment_id) AS Total_Payments,
    ci.city AS Customer_City,
    co.country AS Customer_Country
FROM customer c
JOIN address a ON c.address_id = a.address_id
JOIN city ci ON a.city_id = ci.city_id
JOIN country co ON ci.country_id = co.country_id
JOIN rental r ON c.customer_id = r.customer_id
JOIN inventory i ON r.inventory_id = i.inventory_id
JOIN film f ON i.film_id = f.film_id
LEFT JOIN payment p ON r.rental_id = p.rental_id
LEFT JOIN film_category fc ON i.film_id = fc.film_id
GROUP BY c.customer_id, Customer_Name, ci.city, co.country;


SELECT * FROM customer_revenue_summary LIMIT 25;
```

Output:

```
44 •  SELECT * FROM customer_revenue_summary LIMIT 25;
45
```

100% 40:7

Result Grid Filter Rows: Search Export: Fetch rows:

Customer_ID	Customer_Name	Total_Rentals	Total_Revenue	Avg_Spending_Per_Rental	Total_Categories_Rented	Most_Frequent_Category	Total_Late_Fees	First_Rental_Date	Last_Rental_Date	Total_Payments	Customer_City	Customer_Country
1	MARY SMITH	32	118.68	3.71	14	Classics	12.50	2005-05-25 11:30:37	2005-08-22 20:03:46	32	Sasebo	Japan
2	PATRICIA JOHNSON	27	128.73	4.77	13	Sports	23.00	2005-05-27 00:09:24	2005-08-23 17:39:35	27	San Bernardino	United States
3	LINDA WILLIAMS	26	135.74	5.22	13	Action	25.00	2005-05-27 17:17:09	2005-08-23 07:10:14	26	Athenai	Greece
4	BARBARA JONES	22	81.78	3.72	14	Horror	9.00	2005-06-15 09:31:28	2005-08-23 07:43:00	22	Myingyan	Myanmar
5	ELIZABETH BROWN	38	144.62	3.81	15	Classics	22.50	2005-05-29 07:25:16	2006-02-14 15:16:03	38	Nantou	Taiwan
6	JENNIFER DAVIS	28	93.72	3.35	13	Drama	19.00	2005-05-25 08:43:32	2005-08-23 06:41:32	28	Laredo	United States
7	MARIA MILLER	33	151.67	4.60	14	Animation	26.50	2005-05-25 06:04:08	2005-08-21 04:49:48	33	Kragujevac	Yugoslavia
8	SUSAN WILSON	24	92.76	3.87	12	Drama	11.50	2005-05-30 03:43:54	2005-08-23 14:31:19	24	Hamilton	New Zealand
9	MARGARET MOORE	23	89.77	3.90	11	Foreign	10.50	2005-05-27 05:01:28	2006-02-14 15:16:03	23	Masqat	Oman
10	DOROTHY TAYLOR	25	99.75	3.99	13	Games	17.50	2005-05-31 19:36:30	2005-08-22 21:59:29	25	Esfahan	Iran
11	LISA ANDERSON	24	106.76	4.45	12	Games	19.50	2005-05-30 22:59:12	2006-02-14 15:16:03	24	Sagamihara	Japan
12	NANCY THOMAS	28	103.72	3.70	12	Drama	18.00	2005-05-30 23:08:03	2005-08-23 20:28:44	28	Yamuna Nagar	India
13	KAREN JACKSON	27	131.73	4.88	11	Classics	23.50	2005-06-17 06:54:42	2005-08-22 20:51:24	27	Osmaniye	Turkey
14	BETTY WHITE	28	117.72	4.20	13	Games	19.00	2005-05-26 00:37:28	2006-02-14 15:16:03	28	Citrus Heights	United States
15	HELEN HARRIS	32	134.68	4.21	13	Foreign	15.50	2005-06-18 21:26:56	2006-02-14 15:16:03	32	Bhopal	India
16	SANDRA MARTIN	28	118.72	4.24	13	Sports	14.50	2005-05-27 03:07:10	2005-08-21 14:45:34	28	Southend-on-...	United Kingdom
17	DONNA THOMPSON	21	98.79	4.70	9	Foreign	15.00	2005-05-26 19:44:54	2005-08-23 12:41:38	21	Elista	Russian Federation
18	CAROL GARCIA	22	91.78	4.17	14	Classics	9.00	2005-05-25 06:44:53	2005-08-20 01:29:29	22	Kaduna	Nigeria
19	RUTH MARTINEZ	24	125.76	5.24	12	Sports	22.00	2005-05-25 01:10:47	2005-08-23 03:46:47	24	Kimberley	South Africa
20	SHARON ROBINSON	30	115.70	3.86	14	New	15.00	2005-05-26 07:27:36	2005-08-23 01:10:42	30	Mardan	Pakistan
21	MICHELLE CLARK	35	155.65	4.45	15	Sci-Fi	29.50	2005-05-26 15:42:20	2006-02-14 15:16:03	35	Tangail	Bangladesh
22	LAURA RODRIGUEZ	22	113.78	5.17	12	Foreign	17.00	2005-05-27 07:49:43	2006-02-14 15:16:03	22	Salé	Morocco
23	SARAH LEWIS	30	119.70	3.99	12	Sci-Fi	18.00	2005-05-25 21:20:03	2006-02-14 15:16:03	30	Liepaja	Latvia
24	KIMBERLY LEE	25	95.75	3.83	14	Animation	12.50	2005-05-31 01:02:28	2005-08-22 21:28:59	25	Córdoba	Argentina
25	DEBORAH WALKER	29	115.71	3.99	14	Action	11.50	2005-05-25 14:31:25	2005-08-23 20:00:30	29	Shikarpur	Pakistan

Query 1 - Identifying Potential VIP Customers

Business Question: Which customers generate high revenue and rent frequently, making them ideal for premium loyalty programs?

Purpose: High-spending, frequent renters are crucial for sustained business growth. Identifying these customers enables tailored engagement—such as exclusive offers, early access to releases, and personalized recommendations—maximizing customer lifetime value.

SQL:

```
SELECT Customer_ID, Customer_Name, Total_Rentals, Total_Revenue,
    ROUND(Total_Revenue / NULLIF(Total_Rentals, 0), 2) AS Revenue_Per_Rental
FROM customer_revenue_summary
WHERE Total_Rentals > 30 AND Total_Revenue > 150
ORDER BY Total_Revenue DESC, Total_Rentals DESC
LIMIT 25;
```

Output:

Customer_ID	Customer_Name	Total_Rentals	Total_Revenue	Revenue_Per_Rental
526	KARL SEAL	45	221.55	4.92
148	ELEANOR HUNT	46	216.54	4.71
144	CLARA SHAW	42	195.58	4.66
137	RHONDA KENNEDY	39	194.61	4.99
178	MARION SNYDER	39	194.61	4.99
459	TOMMY COLLAZO	38	186.62	4.91
469	WESLEY BULL	40	177.60	4.44
468	TIM CARY	39	175.61	4.50
236	MARCIA DEAN	42	175.58	4.18
181	ANA BRADLEY	34	174.66	5.14
176	JUNE CARROLL	37	173.63	4.69
50	DIANE COLLINS	35	169.65	4.85
259	LENA JENSEN	32	168.68	5.27
522	ARNOLD HAVENS	33	167.67	5.08
410	CURTIS IRBY	38	167.62	4.41
403	MIKE WAY	35	166.65	4.76
295	DAISY BATES	38	162.62	4.28
209	TONYA CHAPMAN	32	161.68	5.05
373	LOUIS LEONE	35	161.65	4.62
470	GORDON ALLARD	32	160.68	5.02
550	GUY BROWNLEE	32	159.68	4.99
462	WARREN SHERROD	33	159.67	4.84
267	MARGIE WADE	36	159.64	4.43
293	MAE FLETCHER	31	158.69	5.12
372	STEVE MACKENZIE	34	158.66	4.67

Query 2 - Average Spending Per Rental by Country

Business Question: How much do customers from different countries spend per rental on average?

Purpose: Understanding regional spending habits allows the company to refine its pricing strategy and create country-specific promotions. High-spending regions may present opportunities for premium services, while low-spending markets may benefit from bundled offers or discounts.

SQL:

```
SELECT Customer_Country,
    ROUND(AVG(Avg_Spending_Per_Rental), 2) AS Avg_Spending
FROM customer_revenue_summary
GROUP BY Customer_Country
ORDER BY Avg_Spending DESC
LIMIT 25;
```

Output:

Customer_Count...	Avg_Spending
Nepal	5.31
Turkmenistan	5.06
Nauru	4.80
Hong Kong	4.76
Armenia	4.75
French Guiana	4.72
Réunion	4.71
Kenya	4.68
Tuvalu	4.64
Chile	4.63
Czech Republic	4.61
Ukraine	4.61
Yugoslavia	4.55
Bahrain	4.51
Holy See (Vatica…	4.49
Ecuador	4.48
Colombia	4.45
Paraguay	4.44
Malawi	4.42
Latvia	4.38
Italy	4.38
Poland	4.37
Israel	4.37
Thailand	4.35
Greece	4.34

Query 3 - Monthly Revenue Trend

Business Question: How does revenue fluctuate over time?

Purpose: Monitoring revenue trends helps identify seasonal demand patterns, enabling the company to optimize inventory levels, schedule marketing campaigns during slower months, and accurately forecast revenue.

SQL:

```
SELECT DATE_FORMAT(First_Rental_Date, '%Y-%m') AS Rental_Month,
    ROUND(SUM(Total_Revenue), 2) AS Monthly_Revenue
FROM customer_revenue_summary
GROUP BY Rental_Month
ORDER BY Rental_Month;
```

Output:

Rental_Month	Monthly_Revenue
2005-05	59083.61
2005-06	8236.14
2005-07	86.81

Query 4 - Customers Who Stopped Renting

Business Question: Which customers have not rented in the last 6 months?

Purpose: Detecting inactive customers allows for targeted re-engagement campaigns—such as promotional emails, loyalty rewards, or personalized recommendations—aimed at winning them back before they permanently churn.

SQL:

```
SELECT Customer_ID, Customer_Name, Last_Rental_Date
FROM customer_revenue_summary
WHERE Last_Rental_Date < DATE_SUB(CURDATE(), INTERVAL 6 MONTH)
LIMIT 25;
```

Output:

Customer_ID	Customer_Name	Last_Rental_Date
1	MARY SMITH	2005-08-22 20:03:46
2	PATRICIA JOHNSON	2005-08-23 17:39:35
3	LINDA WILLIAMS	2005-08-23 07:10:14
4	BARBARA JONES	2005-08-23 07:43:00
5	ELIZABETH BROWN	2006-02-14 15:16:03
6	JENNIFER DAVIS	2005-08-23 06:41:32
7	MARIA MILLER	2005-08-21 04:49:48
8	SUSAN WILSON	2005-08-23 14:31:19
9	MARGARET MOORE	2006-02-14 15:16:03
10	DOROTHY TAYLOR	2005-08-22 21:59:29
11	LISA ANDERSON	2006-02-14 15:16:03
12	NANCY THOMAS	2005-08-23 20:28:44
13	KAREN JACKSON	2005-08-22 20:51:24
14	BETTY WHITE	2006-02-14 15:16:03
15	HELEN HARRIS	2006-02-14 15:16:03
16	SANDRA MARTIN	2005-08-21 14:45:34
17	DONNA THOMPSON	2005-08-23 12:41:38
18	CAROL GARCIA	2005-08-20 01:29:29
19	RUTH MARTINEZ	2005-08-23 03:46:47
20	SHARON ROBINSON	2005-08-23 01:10:42
21	MICHELLE CLARK	2006-02-14 15:16:03
22	LAURA RODRIGUEZ	2006-02-14 15:16:03
23	SARAH LEWIS	2006-02-14 15:16:03
24	KIMBERLY LEE	2005-08-22 21:28:59
25	DEBORAH WALKER	2005-08-23 20:00:30

Query 5 - Top Movie Genre by Country (Revenue)

Business Question: Which movie genre generates the most revenue in each country?

Purpose: Understanding genre preferences by region allows the company to optimize inventory, tailor marketing messages, and acquire content that aligns with local tastes.

SQL:

```
SELECT co.country AS Customer_Country,
    fc.category_id AS Genre_ID,
    c.name AS Genre_Name,
    ROUND(SUM(p.amount), 2) AS Total_Revenue
FROM rental r
JOIN inventory i ON r.inventory_id = i.inventory_id
JOIN film_category fc ON i.film_id = fc.film_id
JOIN category c ON fc.category_id = c.category_id
JOIN payment p ON r.rental_id = p.rental_id
JOIN customer cu ON r.customer_id = cu.customer_id
JOIN address a ON cu.address_id = a.address_id
JOIN city ci ON a.city_id = ci.city_id
JOIN country co ON ci.country_id = co.country_id
GROUP BY Customer_Country, Genre_ID, Genre_Name
ORDER BY Customer_Country, Total_Revenue DESC
LIMIT 25;
```

Output:

Customer_Country	Genre_ID	Genre_Name	Total_Revenue
Afghanistan	4	Classics	12.98
Afghanistan	13	New	12.98
Afghanistan	1	Action	8.99
Afghanistan	5	Comedy	8.97
Afghanistan	7	Drama	7.98
Afghanistan	16	Travel	4.99
Afghanistan	10	Games	2.99
Afghanistan	14	Sci-Fi	2.99
Afghanistan	6	Documentary	1.98
Afghanistan	3	Children	0.99
Afghanistan	2	Animation	0.99
Afghanistan	9	Foreign	0.99
Algeria	5	Comedy	56.92
Algeria	14	Sci-Fi	46.88
Algeria	13	New	39.90
Algeria	8	Family	33.94
Algeria	2	Animation	32.93
Algeria	7	Drama	24.94
Algeria	3	Children	24.93
Algeria	15	Sports	21.94
Algeria	1	Action	17.95
Algeria	6	Documentary	14.97
Algeria	12	Music	14.96
Algeria	9	Foreign	13.96
Algeria	16	Travel	12.97

Query 6 - Most Profitable Movie Category Overall

Business Question: Which movie category generates the most revenue overall?

Purpose: Identifying top-performing genres guides inventory planning, marketing spend allocation, and potential expansion of popular categories to drive higher revenue.

SQL:

```
SELECT Most_Frequent_Category AS Movie_Category,
    ROUND(SUM(Total_Revenue), 2) AS Category_Revenue
FROM customer_revenue_summary
GROUP BY Most_Frequent_Category
ORDER BY Category_Revenue DESC
LIMIT 25;
```

Output:

Movie_Category	Category_Revenue
Sci-Fi	6711.85
Sports	6437.11
Animation	5945.00
Action	5600.84
Drama	5019.80
Documentary	4975.43
Family	4723.94
Foreign	4426.30
Games	4306.99
New	3766.18
Classics	3272.10
Comedy	2893.21
Music	2712.49
Children	2591.79
Travel	2037.24
Horror	1986.29

Query 7 - Customer Lifetime Value (CLV) Analysis

Business Question: Which customers contribute the most revenue over their lifetime?

Purpose: High-CLV customers are prime targets for retention programs. By understanding their behavior, the company can create premium offers and reward programs that maintain their loyalty.

SQL:

```
SELECT Customer_ID, Customer_Name, Total_Revenue,
    DATEDIFF(Last_Rental_Date, First_Rental_Date) AS Rental_Duration_Days,
    ROUND(Total_Revenue / NULLIF(DATEDIFF(Last_Rental_Date, First_Rental_Date), 0), 2) AS
Revenue_Per_Day
FROM customer_revenue_summary
ORDER BY Total_Revenue DESC
LIMIT 25;
```

Output:

Customer_ID	Customer_Name	Total_Revenue	Rental_Duration_Days	Revenue_Per_Day
526	KARL SEAL	221.55	87	2.55
148	ELEANOR HUNT	216.54	87	2.49
144	CLARA SHAW	195.58	88	2.22
178	MARION SNYDER	194.61	244	0.80
137	RHONDA KENNEDY	194.61	85	2.29
459	TOMMY COLLAZO	186.62	91	2.05
469	WESLEY BULL	177.60	89	2.00
468	TIM CARY	175.61	90	1.95
236	MARCIA DEAN	175.58	264	0.67
181	ANA BRADLEY	174.66	262	0.67
176	JUNE CARROLL	173.63	89	1.95
50	DIANE COLLINS	169.65	86	1.97
259	LENA JENSEN	168.68	86	1.96
522	ARNOLD HAVENS	167.67	88	1.91
410	CURTIS IRBY	167.62	244	0.69
403	MIKE WAY	166.65	87	1.92
295	DAISY BATES	162.62	263	0.62
209	TONYA CHAPMAN	161.68	88	1.84
373	LOUIS LEONE	161.65	264	0.61
470	GORDON ALLARD	160.68	88	1.83
187	BRITTANY RILEY	159.72	89	1.79
550	GUY BROWNLEE	159.68	260	0.61
462	WARREN SHERROD	159.67	89	1.79
267	MARGIE WADE	159.64	265	0.60
293	MAE FLETCHER	158.69	88	1.80

Query 8 - Peak Rental Days

Business Question: On which days of the week do rentals peak?

Purpose: Identifying high-traffic rental days allows for better staffing, marketing campaigns, and promotions that align with customer behavior patterns.

SQL:

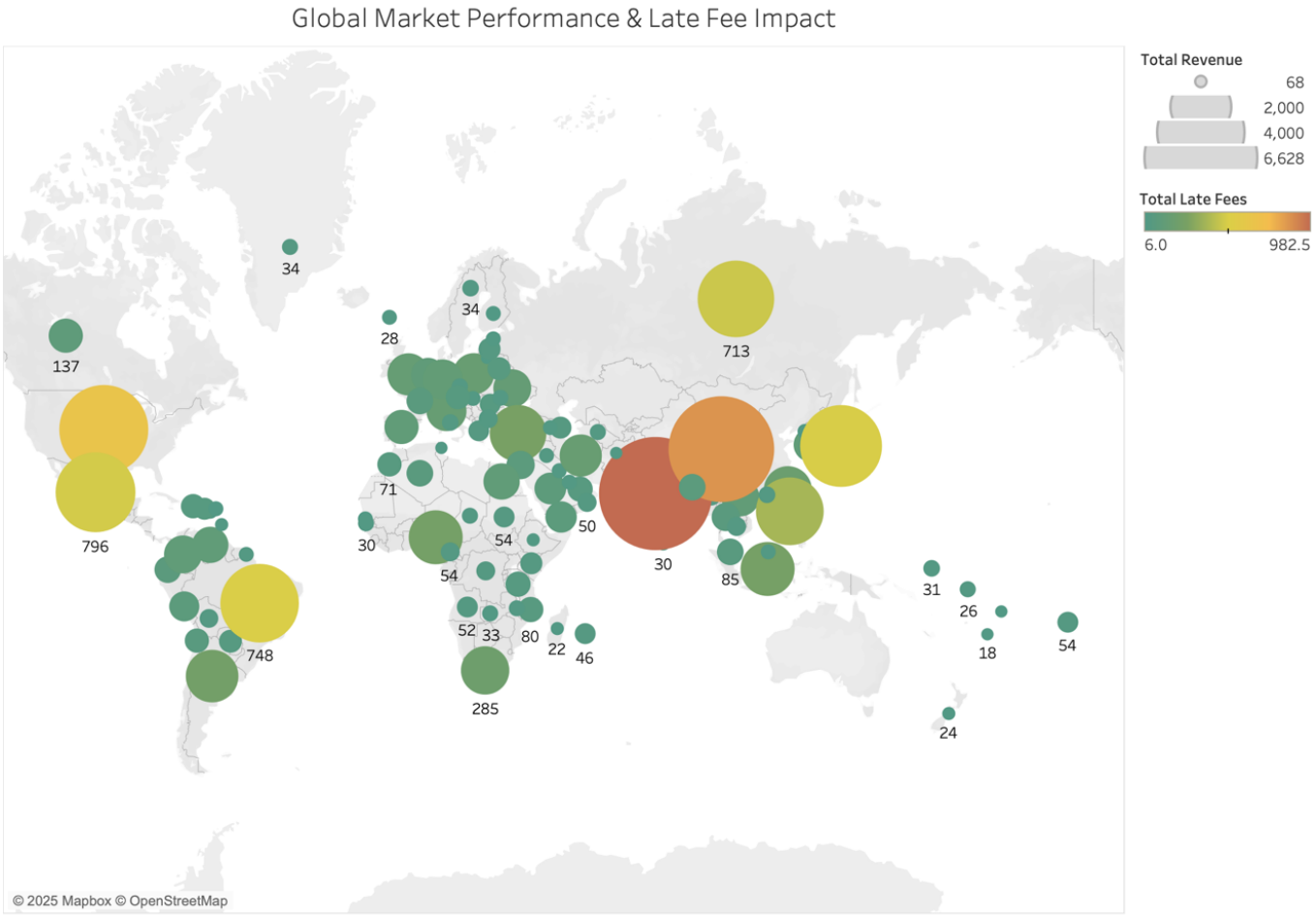
```
SELECT DAYNAME(First_Rental_Date) AS Rental_Day,
    COUNT(Customer_ID) AS Total_Rentals,
    ROUND(SUM(Total_Revenue), 2) AS Total_Revenue
FROM customer_revenue_summary
GROUP BY Rental_Day
ORDER BY Total_Rentals DESC;
```

Output:

Rental_Day	Total_Rentals	Total_Revenue
Wednesday	155	17774.52
Thursday	141	15970.63
Friday	86	9885.50
Saturday	79	8809.93
Tuesday	52	5353.16
Sunday	45	4933.13
Monday	41	4679.69

Tableau Visualizations & Dashboard

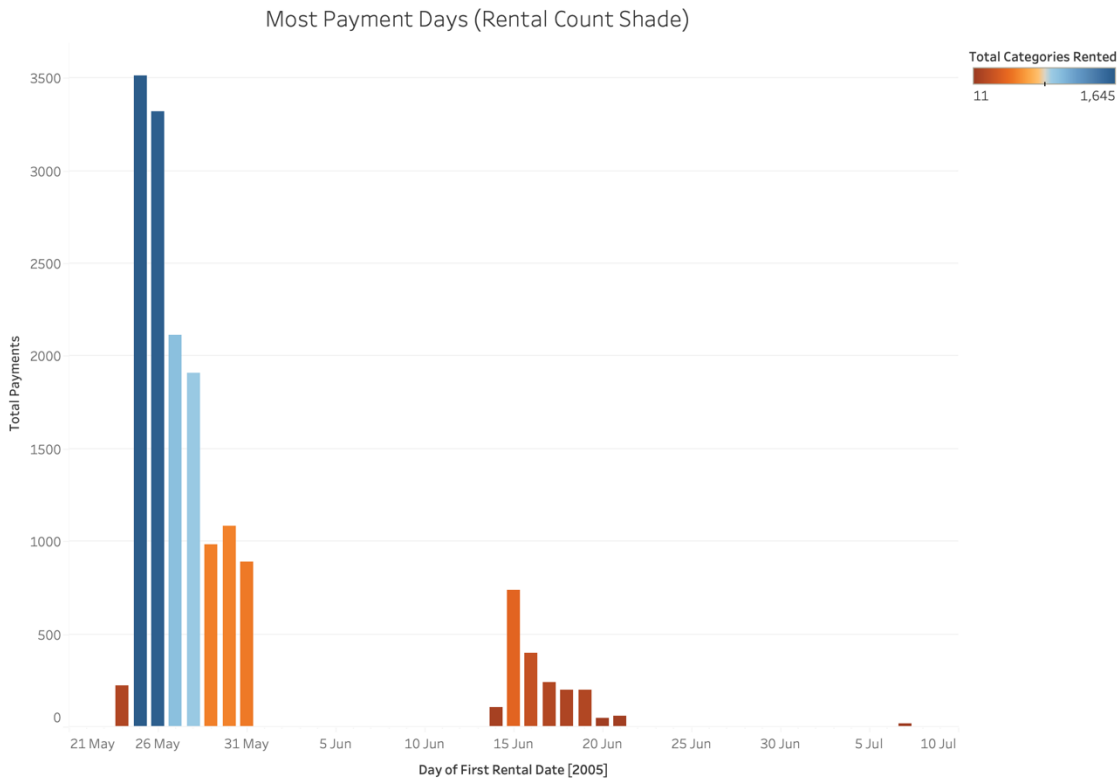
Graph 1 - Geographic Revenue & Late Fees



Map based on Longitude (generated) and Latitude (generated). Color shows sum of Total Late Fees. Size shows sum of Total Revenue. The marks are labeled by sum of Total Rentals. Details are shown for Customer Country. The view is filtered on Latitude (generated), Longitude (generated) and Customer Country. The Latitude (generated) filter keeps non-Null values only. The Longitude (generated) filter keeps non-Null values only. The Customer Country filter keeps 108 of 108 members.

A bubble map shows global revenue distribution (bubble size) and late fee impact (color intensity). Management can pinpoint profitable markets like the US and India while addressing late fee issues in other regions.

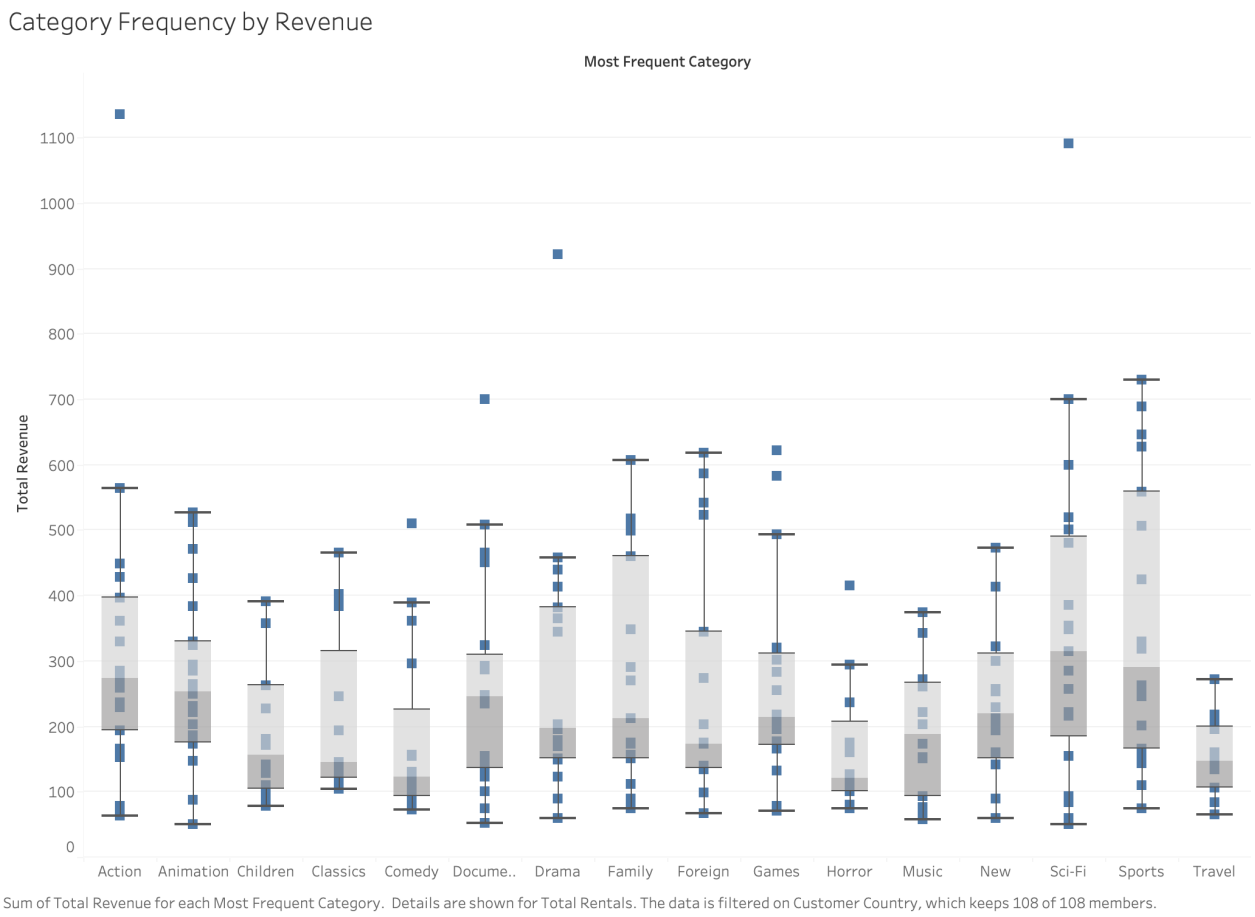
Graph 2 - Monthly Payment Trends



The plot of sum of Total Payments for First Rental Date Day. Color shows sum of Total Categories Rented. The data is filtered on Customer Country, which keeps 108 of 108 members.

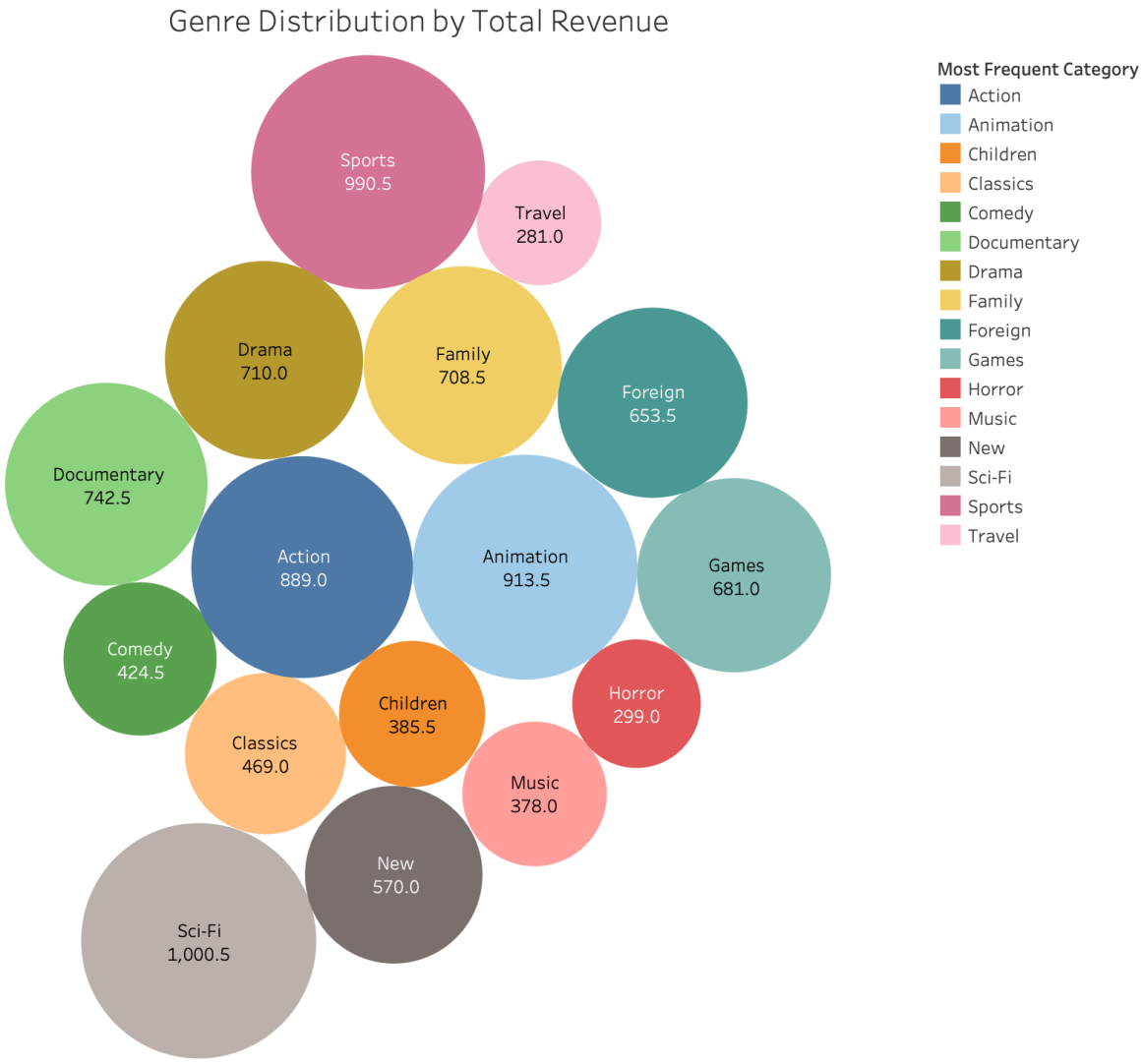
A time-series line chart displaying revenue patterns across months. This helps identify seasonal peaks and troughs for better inventory and marketing planning.

Graph 3 - Revenue by Movie Category (Box Plot)



Visualizes variability in revenue per genre, highlighting high-potential categories (e.g., Sci-Fi, Sports) for investment.

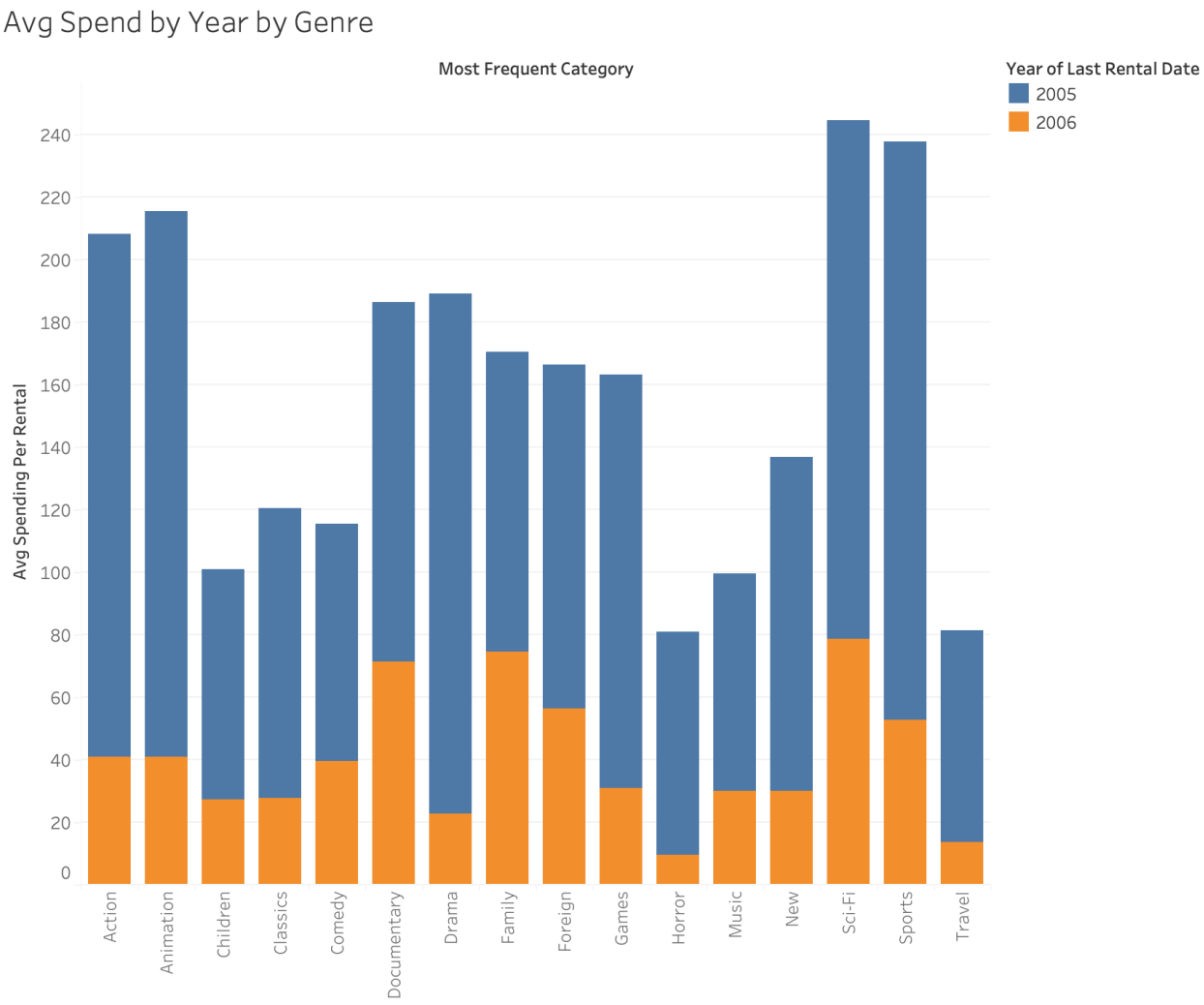
Graph 4 - Genre Market Share (Bubble Chart)



Most Frequent Category and sum of Total Late Fees. Color shows details about Most Frequent Category. Size shows sum of Total Late Fees. The marks are labeled by Most Frequent Category and sum of Total Late Fees. The data is filtered on Customer Country, which keeps 108 of 108 members.

Shows each genre's revenue share, enabling informed decisions on marketing budgets and inventory allocation.

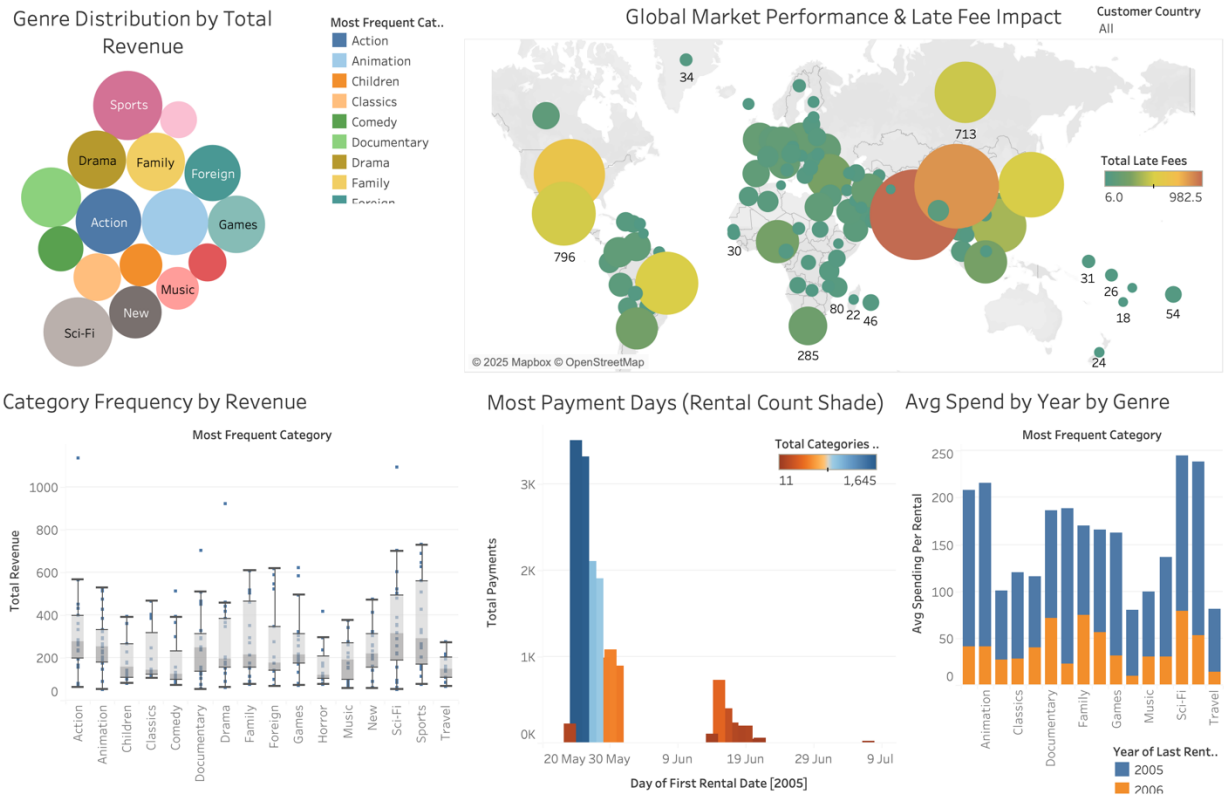
Graph 5 - Year-over-Year Genre Performance



Sum of Avg Spending Per Rental for each Most Frequent Category. Color shows details about Last Rental Date Year. The data is filtered on Customer Country, which keeps 108 of 108 members.

A bar chart comparing genre performance over two years to reveal growth or decline trends.

Dashboard - Executive Overview



A unified, interactive dashboard combining all five visualizations with a country filter, giving management instant access to market-specific insights and revenue drivers.