

Analyzing Netflix User Data with SQL and Custom Functions

1. Company

We explored Netflix user data using SQL, aiming to unearth meaningful patterns, trends, and correlations that can significantly impact strategic decision-making and amplify the overall user experience. As a seminal force in the realm of entertainment, Netflix, an American video-on-demand over-the-top streaming service, has not only revolutionized the way audiences consume content but has also emerged as a global cultural phenomenon [1]. At the core of Netflix's success lies its unparalleled ability to deliver a diverse array of original and acquired movies, TV shows, and exclusive content to an expansive subscriber base, spanning diverse demographics and geographical locations. This analysis is positioned at the intersection of technology, entertainment, and data, seeking to provide a nuanced understanding of user engagement, content preferences, and the evolving dynamics of the streaming landscape [2].

Netflix's strategic prowess is underscored by its data-driven approach, leveraging insights gleaned from user behavior to inform content creation, curation, and platform optimization. By harnessing the power of data analytics, Netflix continues to refine its recommendation algorithms, tailor content libraries, and introduce innovative features that resonate with its global audience. Through this project, we endeavor to contribute to this ongoing narrative by peeling back layers of data to reveal actionable intelligence that aligns with Netflix's commitment to pushing the boundaries of storytelling and providing an immersive streaming experience for its users [3]. Our analysis not only seeks to unravel the intricate tapestry of user preferences but also serves as a testament to the evolving landscape of on-demand entertainment in the digital age, where data insights become the compass guiding strategic decisions for industry leaders like Netflix.

2. Tables

Our dataset, sourced from Kaggle (<https://www.kaggle.com/datasets/arnavsmayan/netflix-userbase-dataset>), constitutes a synthetic yet comprehensive representation of a simulated Netflix user base, encompassing 2500 rows, each uniquely identified by a User ID. Within this dataset, a rich tapestry of user-centric information unfolds, featuring key attributes such as subscription type (Basic, Standard, or Premium), monthly revenue, join date, last payment date, country of residence, and detailed demographics including age and gender. This multifaceted dataset extends its scope to include additional dimensions such as Device Type (e.g., Smart TV, Mobile, Desktop, Tablet) and Account Status, contributing nuanced insights into user behaviors and preferences.

To facilitate a nuanced and efficient analysis, we have meticulously organized the data into three distinct tables, each with a specific focus: "Subscription_detail"(Fig2.1) provides a condensed snapshot with UserID, Subscription Type, and Country, offering an initial lens into user subscriptions. "Subscription_revenue" (Fig2.2) delves into the financial aspects, including UserID, Monthly Revenue,

Join Date, and Last Payment Date, enabling a detailed exploration of revenue-related metrics. Meanwhile, "Subscription_user"(Fig2.3) broadens the perspective, incorporating UserID, Age, Gender, Device, and Plan Duration, thus affording a comprehensive view of user demographics and behavior patterns.

It is imperative to clarify that our dataset, while robust and insightful, is a synthetic representation, deliberately designed to not disclose any actual Netflix user data. The carefully structured tables not only lay the foundation for our exploration but also serve as integral tools for unraveling user trends, preferences, and the intricacies of revenue generation within this hypothetical Netflix user base.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	UserID	Subscription_Type	Country												
2	1	Basic	United States												
3	2	Premium	Canada												
4	3	Standard	United Kingdom												
5	4	Standard	Australia												
6	5	Basic	Germany												
7	6	Premium	France												
8	7	Standard	Brazil												
9	8	Basic	Mexico												
10	9	Standard	Spain												
11	10	Premium	Italy												
12	11	Basic	United States												
13	12	Premium	Canada												
14	13	Standard	United Kingdom												
15	14	Basic	Australia												
16	15	Standard	Germany												
17	16	Premium	France												
18	17	Basic	Brazil												
19	18	Standard	Mexico												
20	19	Premium	Spain												
21	20	Basic	Italy												
22	21	Premium	United States												
23	22	Basic	Canada												
24	23	Standard	United Kingdom												

Fig2.1 “Subscription_detail” Table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	UserID	Monthly_Revenue	Join_Date	Last_Payment_Date														
2	1	15	06/05/21	22/06/23														
3	38	12	11/05/21	24/06/23														
4	23	12	12/05/21	27/06/23														
5	58	12	10/08/21	24/06/23														
6	68	12	11/09/21	24/06/23														
7	7	12	12/09/21	25/06/23														
8	186	12	03/01/22	29/06/23														
9	36	15	03/01/22	27/06/23														
10	483	11	06/01/22	26/06/23														
11	383	10	06/01/22	27/06/23														
12	112	10	06/01/22	28/06/23														
13	819	14	07/01/22	07/07/23														
14	1286	15	07/01/22	07/07/23														
15	1010	13	07/01/22	07/07/23														
16	1150	15	07/01/22	07/07/23														
17	1200	13	07/01/22	07/07/23														
18	1305	11	07/01/22	07/07/23														
19	1293	11	07/01/22	07/07/23														
20	1364	14	07/01/22	07/07/23														
21	1434	13	07/01/22	07/07/23														
22	1363	11	07/01/22	07/07/23														
23	1653	13	07/01/22	07/07/23														
24	1791	14	07/01/22	07/07/23														
25	1514	12	07/01/22	07/07/23														
26	1722	13	07/01/22	07/07/23														
27	442	12	07/01/22	27/06/23														
28	443	10	07/01/22	27/06/23														
29	515	15	07/01/22	29/06/23														
30	806	15	07/01/22	30/06/23														

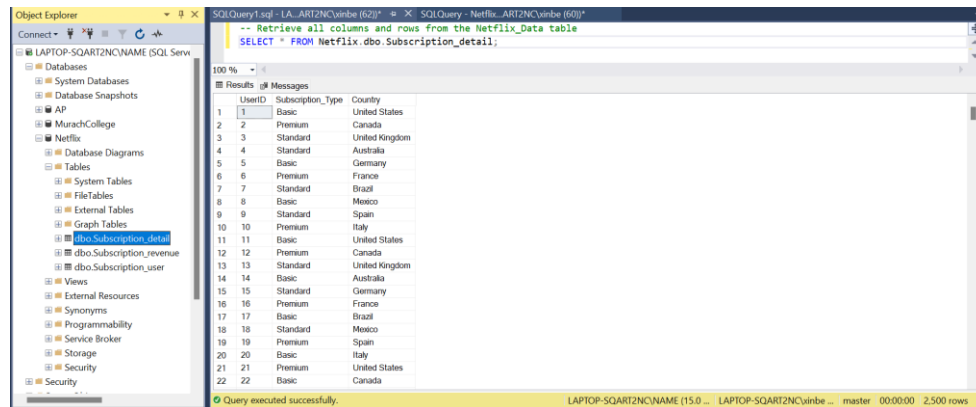
Fig2.2 “Subscription_revenue” Table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	UserID	Age	Gender	Device	Plan_Duration									
2	1	28	Male	Smartphon	1 Month									
3	2	35	Female	Tablet	1 Month									
4	3	42	Male	Smart TV	1 Month									
5	4	51	Female	Laptop	1 Month									
6	5	33	Male	Smartphon	1 Month									
7	6	29	Female	Smart TV	1 Month									
8	7	46	Male	Tablet	1 Month									
9	8	39	Female	Laptop	1 Month									
10	9	37	Male	Smartphon	1 Month									
11	10	44	Female	Smart TV	1 Month									
12	11	31	Female	Smartphon	1 Month									
13	12	45	Male	Tablet	1 Month									
14	13	48	Female	Laptop	1 Month									
15	14	27	Male	Smartphon	1 Month									
16	15	38	Female	Smart TV	1 Month									
17	16	36	Male	Tablet	1 Month									
18	17	30	Female	Laptop	1 Month									
19	18	43	Male	Smartphon	1 Month									
20	19	32	Female	Smart TV	1 Month									
21	20	41	Male	Tablet	1 Month									
22	21	26	Female	Laptop	1 Month									

Fig2.3 “Subscription_user” Table

3. Views

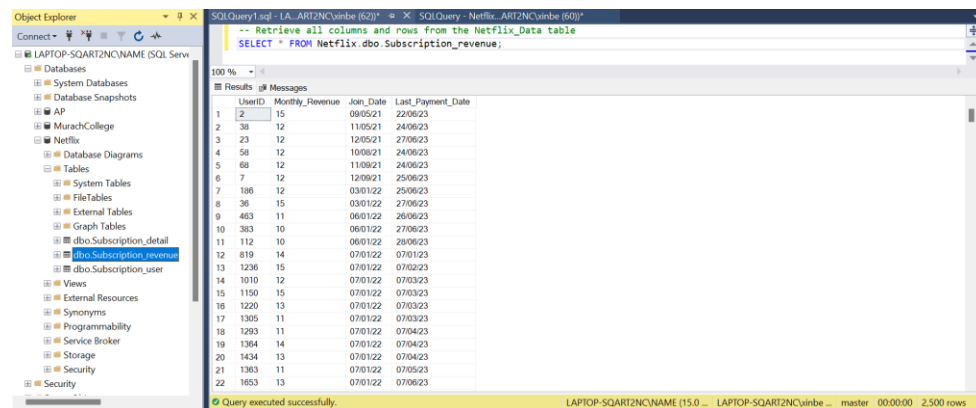
We first retrieve all columns and rows from three tables in the "Netflix" database: "Subscription_detail," (Fig3.1), "Subscription_revenue," (Fig3.2), and "Subscription_user." (Fig3.3). These queries offer a comprehensive snapshot of user subscription details, and demographic information, providing essential data for analysis and exploration.



```
-- Retrieve all columns and rows from the Netflix_Data table
SELECT * FROM Netflix.dbo.Subscription_detail;
```

	UserID	Subscription_Type	Country
1	1	Basic	United States
2	2	Premium	Canada
3	3	Standard	United Kingdom
4	4	Standard	Australia
5	5	Basic	Germany
6	6	Premium	France
7	7	Standard	Brazil
8	8	Basic	Mexico
9	9	Standard	Spain
10	10	Premium	Italy
11	11	Basic	United States
12	12	Premium	Canada
13	13	Standard	United Kingdom
14	14	Basic	Australia
15	15	Standard	Germany
16	16	Premium	France
17	17	Basic	Brazil
18	18	Standard	Mexico
19	19	Premium	Spain
20	20	Basic	Italy
21	21	Premium	United States
22	22	Basic	Canada

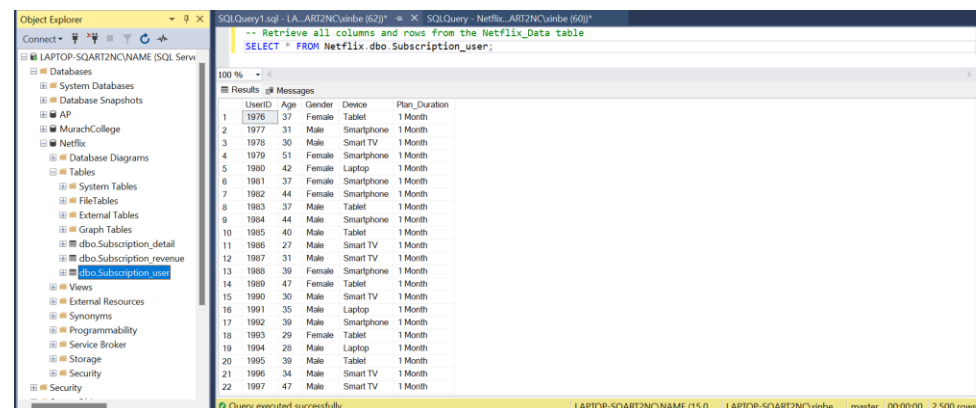
Fig3.1 "Subscription_detail" Table Retrieving



```
-- Retrieve all columns and rows from the Netflix_Data table
SELECT * FROM Netflix.dbo.Subscription_revenue;
```

	UserID	Monthly_Revenue	Join_Date	Last_Payment_Date
1	2	15	09/05/21	22/06/23
2	38	12	11/05/21	24/06/23
3	23	12	12/05/21	27/06/23
4	58	12	10/06/21	24/06/23
5	68	12	11/06/21	24/06/23
6	7	12	12/06/21	25/06/23
7	198	12	03/01/22	25/06/23
8	38	15	03/01/22	27/06/23
9	403	11	06/01/22	26/06/23
10	383	10	06/01/22	27/06/23
11	112	10	06/01/22	28/06/23
12	819	14	07/01/22	07/01/23
13	1236	15	07/01/22	07/02/23
14	1010	12	07/01/22	07/03/23
15	1150	15	07/01/22	07/03/23
16	1220	13	07/01/22	07/03/23
17	1305	11	07/01/22	07/03/23
18	1203	11	07/01/22	07/04/23
19	1304	14	07/01/22	07/04/23
20	1434	13	07/01/22	07/04/23
21	1363	11	07/01/22	07/05/23
22	1853	13	07/01/22	07/06/23

Fig3.2 "Subscription_revenue" Table Retrieving



```
-- Retrieve all columns and rows from the Netflix_Data table
SELECT * FROM Netflix.dbo.Subscription_user;
```

	UserID	Age	Gender	Device	Plan_Duration
1	1976	37	Female	Tablet	1 Month
2	1977	31	Male	Smartphone	1 Month
3	1978	30	Male	Smart TV	1 Month
4	1979	51	Female	Smartphone	1 Month
5	1980	42	Female	Laptop	1 Month
6	1981	37	Female	Smartphone	1 Month
7	1982	44	Female	Smartphone	1 Month
8	1983	37	Male	Tablet	1 Month
9	1984	44	Male	Smartphone	1 Month
10	1985	40	Male	Tablet	1 Month
11	1986	27	Male	Smart TV	1 Month
12	1987	31	Male	Smart TV	1 Month
13	1988	30	Female	Smartphone	1 Month
14	1989	47	Female	Tablet	1 Month
15	1990	30	Male	Smart TV	1 Month
16	1991	35	Male	Laptop	1 Month
17	1992	30	Male	Smartphone	1 Month
18	1993	29	Female	Tablet	1 Month
19	1994	28	Male	Laptop	1 Month
20	1995	39	Male	Tablet	1 Month
21	1996	34	Male	Smart TV	1 Month
22	1997	47	Male	Smart TV	1 Month

Fig3.3 "Subscription_user" Table Retrieving

We calculate the total revenue for each country in the Netflix dataset by summing the monthly revenue, presenting the results in descending order, as Fig3.4 shows. It achieves this by joining the "Subscription_detail" and "Subscription_revenue" tables based on the common "UserID" column. The SELECT statement retrieves the "country" column and computes the sum of monthly revenue for each country, aliasing it as "total_revenue." The GROUP BY clause organizes the results by country, allowing for distinct aggregations. The final ORDER BY clause arranges the outcomes in descending order of total revenue, providing a ranked overview of countries contributing the most to the Netflix revenue stream. The result shows that United States owns the highest total revenue while Mexico owns the lowest.

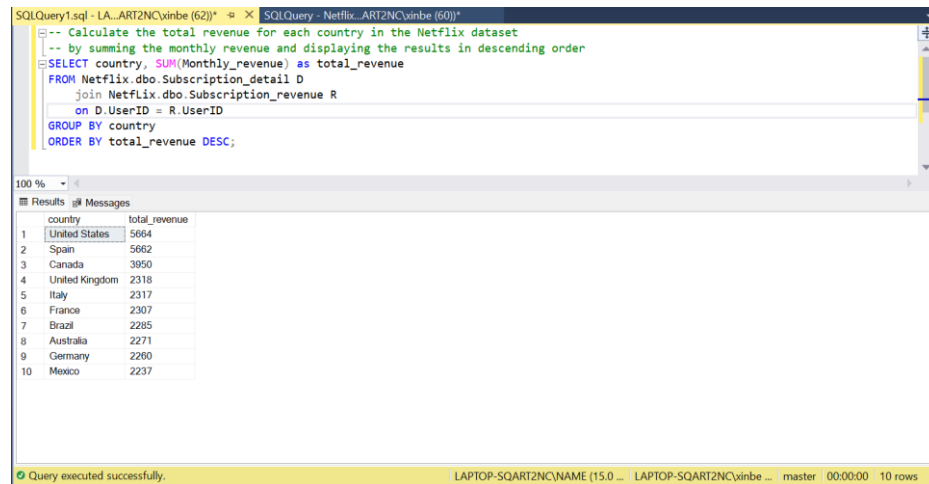


Fig3.4 Calculating Total Revenue for Each Country

We extract information on gender and count the number of premium subscriptions for each gender in the Netflix dataset as Fig3.5 shows. Utilizing a join operation between the "Subscription_user" and "Subscription_detail" tables based on the common "UserID" column, the query filters for premium subscriptions and groups the results by gender. The result shows that the number of premium subscriptions for males and females are very close while males have a little larger number of premium subscriptions.

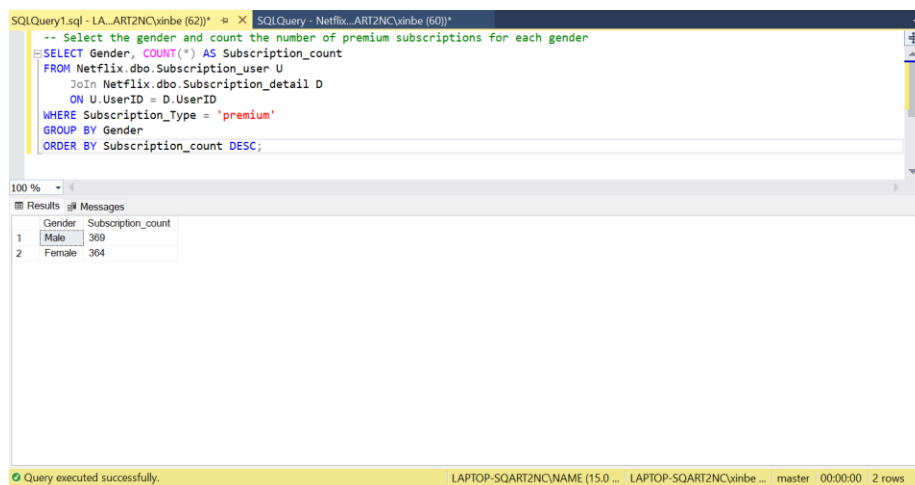


Fig3.5 Counting Premium Subscriptions for Different Genders

We extract data on device types and calculate the subscribers for each device type in the Netflix dataset as Fig3.6 shows. The results show that most of the subscribers use laptops while the least of subscribers use smart TVs.

The screenshot shows a SQL query window with the following code:

```
-- Select the device_type and count the number of subscribers for each device_type
SELECT Device, COUNT(Subscription_Type) AS Subscriber_count
FROM Netflix.dbo.Subscription_user U
JOIN Netflix.dbo.Subscription_detail D
ON U.UserID = D.UserID
GROUP BY Device
ORDER BY Subscriber_count DESC;
```

The results pane displays a table with 4 rows:

Device	Subscriber_count
Laptop	636
Tablet	633
Smartphone	621
Smart TV	610

The status bar at the bottom indicates: "Query executed successfully. LAPTOP-SQART2NC\NAME (15.0 ... LAPTOP-SQART2NC\name ... master 00:00:00 4 rows".

Fig3.6 Counting Subscribers for Different Device Types

We select the gender and count the number of subscribers for each gender in the Netflix dataset as Fig3.7 shows. Results show that Females have larger subscriptions than males.

The screenshot shows a SQL query window with the following code:

```
-- Select the gender and count the number of subscribers for each gender
SELECT Gender, COUNT(*) AS Subscriber_count
FROM Netflix.dbo.Subscription_user
GROUP BY Gender;
```

The results pane displays a table with 2 rows:

Gender	Subscriber_count
Male	1243
Female	1257

The status bar at the bottom indicates: "Query executed successfully. LAPTOP-SQART2NC\NAME (15.0 ... LAPTOP-SQART2NC\name ... master 00:00:00 2 rows".

Fig3.7 Counting Subscribers for Each Gender

We check for subscribers who joined in 2021 and are still active in the Netflix dataset as Fig3.8 shows.

The screenshot shows a SQL query window with the following code:

```
-- Check for subscribers who joined in 2021 and are still active
SELECT *
FROM Netflix.dbo.Subscription_revenue
WHERE YEAR(CONVERT(date, join_date, 3)) >= 2021;
```

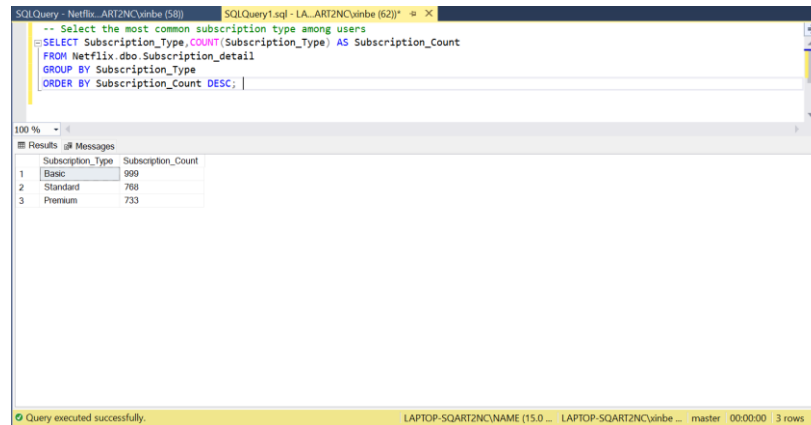
The results pane displays a table with 18 rows of subscriber data:

UserID	Monthly_Revenue	Join_Date	Last_Payment_Date
2	15	09/05/21	22/06/23
38	12	11/05/21	24/06/23
23	12	12/05/21	27/06/23
58	12	10/06/21	24/06/23
68	12	11/09/21	24/06/23
7	12	12/09/21	25/06/23
186	12	03/01/22	25/06/23
36	15	03/01/22	27/06/23
483	11	06/01/22	28/06/23
383	10	06/01/22	27/06/23
112	10	06/01/22	28/06/23
819	14	07/01/22	07/01/23
1236	15	07/01/22	07/02/23
1010	12	07/01/22	07/03/23
1150	15	07/01/22	07/03/23
1220	13	07/01/22	07/03/23
1305	11	07/01/22	07/03/23
1293	11	07/01/22	07/04/23

The status bar at the bottom indicates: "Query executed successfully. LAPTOP-SQART2NC\NAME (15.0 ... LAPTOP-SQART2NC\name ... master 00:00:00 2,500 rows".

Fig3.8 Checking Subscribers Joining in 2021

We identify the most common subscription type among users in the Netflix dataset. The query provides a clear view of the distribution of subscription types, highlighting the most common type at the top of the list based on user counts. As Fig3.9 shows that most people choose the basic subscription type while fewer people choose the premium subscription type.



The screenshot shows a SQL query window with the following query:

```
-- Select the most common subscription type among users
SELECT Subscription_Type, COUNT(Subscription_Type) AS Subscription_Count
FROM Netflix.dbo.Subscription_detail
GROUP BY Subscription_Type
ORDER BY Subscription_Count DESC;
```

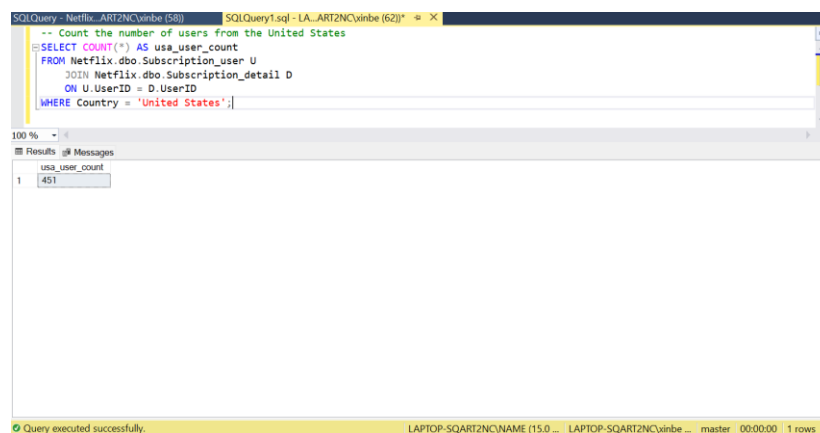
The results table shows the following data:

	Subscription_Type	Subscription_Count
1	Basic	999
2	Standard	768
3	Premium	733

The status bar at the bottom indicates: "Query executed successfully. LAPTOP-SQART2NC\NAME (15.0 ... LAPTOP-SQART2NC\vinbe ... master 00:00:00 3 rows".

Fig3.9 Most Common Subscription Type

We count the number of users from the United States in the Netflix dataset. As Fig3.10 shows We find that there are a total of 451 users from the United States.



The screenshot shows a SQL query window with the following query:

```
-- Count the number of users from the United States
SELECT COUNT(*) AS usa_user_count
FROM Netflix.dbo.Subscription_user U
JOIN Netflix.dbo.Subscription_detail D
ON U.UserID = D.UserID
WHERE Country = 'United States';
```

The results table shows the following data:

	usa_user_count
1	451

The status bar at the bottom indicates: "Query executed successfully. LAPTOP-SQART2NC\NAME (15.0 ... LAPTOP-SQART2NC\vinbe ... master 00:00:00 1 rows".

Fig3.10 Counting Users from United States

We calculate the average age of users as Fig3.11 shows. This query offers a concise metric for understanding the central tendency of user ages in the analyzed subscription data. The result shows that the average age is 38.8.

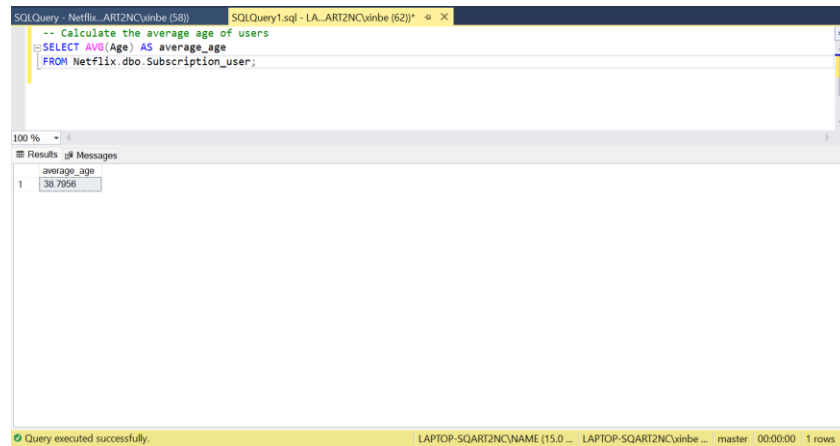


Fig3.11 Calculating Average Age of Users

We identify the country with the highest monthly revenue as Fig3.12 shows. As a result, the query provides a clear view that the United States contributes the highest monthly revenue to Netflix based on the aggregated data while Mexico contributes the lowest.

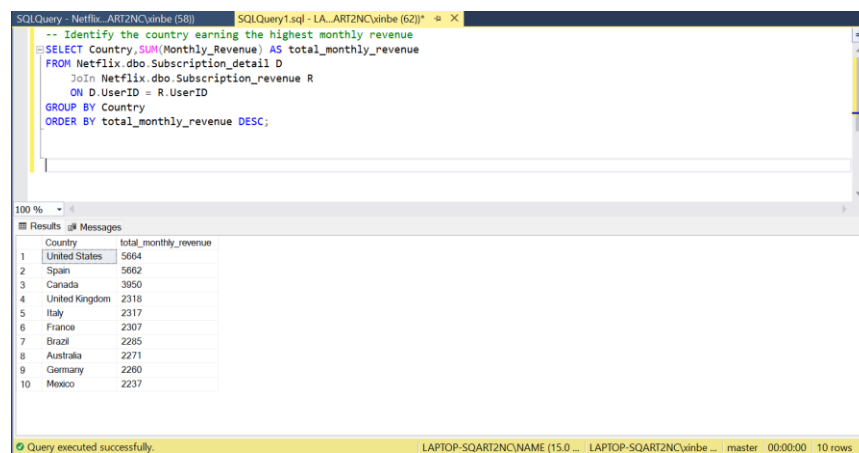


Fig3.12 Identifying Country Earning Highest Monthly Revenue

Moreover, we use three user-defined functions for analysis within the context of Netflix user data. The first function(Fig 3.13), CalculateAge, determines a user's age based on their join date and last payment date, employing intricate date calculations to yield accurate results. The second function(Fig 3.14) calculates the total revenue for a specified subscription type, enhancing the ability to analyze revenue metrics dynamically. The third function(Fig 3.15) computes the average age of users from a specified country, showcasing the versatility of scalar functions in extracting nuanced demographic insights. These functions are then applied by providing calculated ages, GetMonthly_Revenue offering a monthly revenue metric for the 'Basic' subscription type, and GetAverageAgeByCountryV2 yielding the average age for users from the United States. The functions contribute to a more detailed and dynamic exploration of user demographics and subscription-based financial metrics within the dataset.

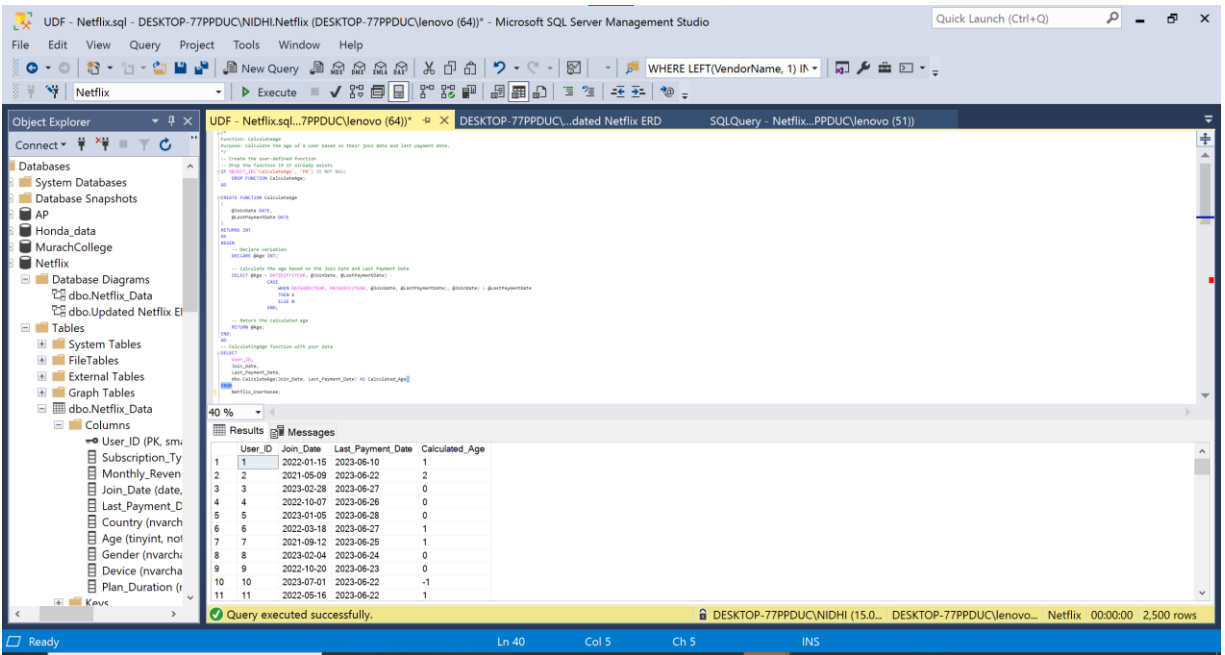


Fig3.13 CalculateAge Function

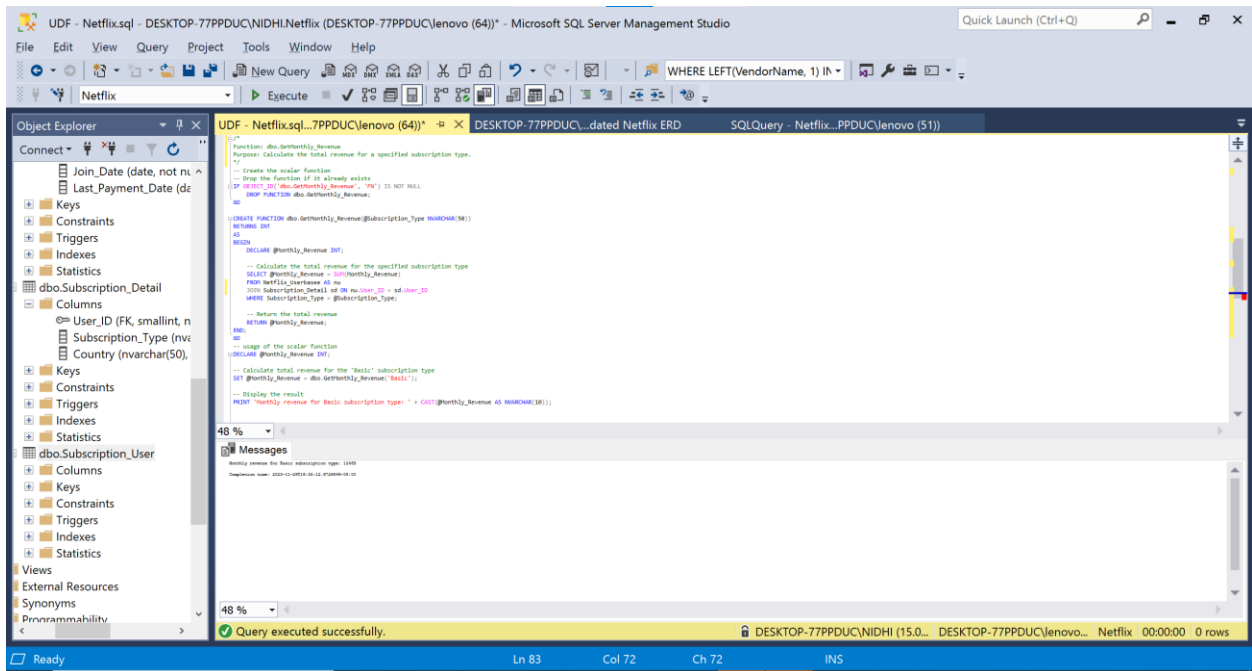


Fig3.14 Calculating Total Revenue Function

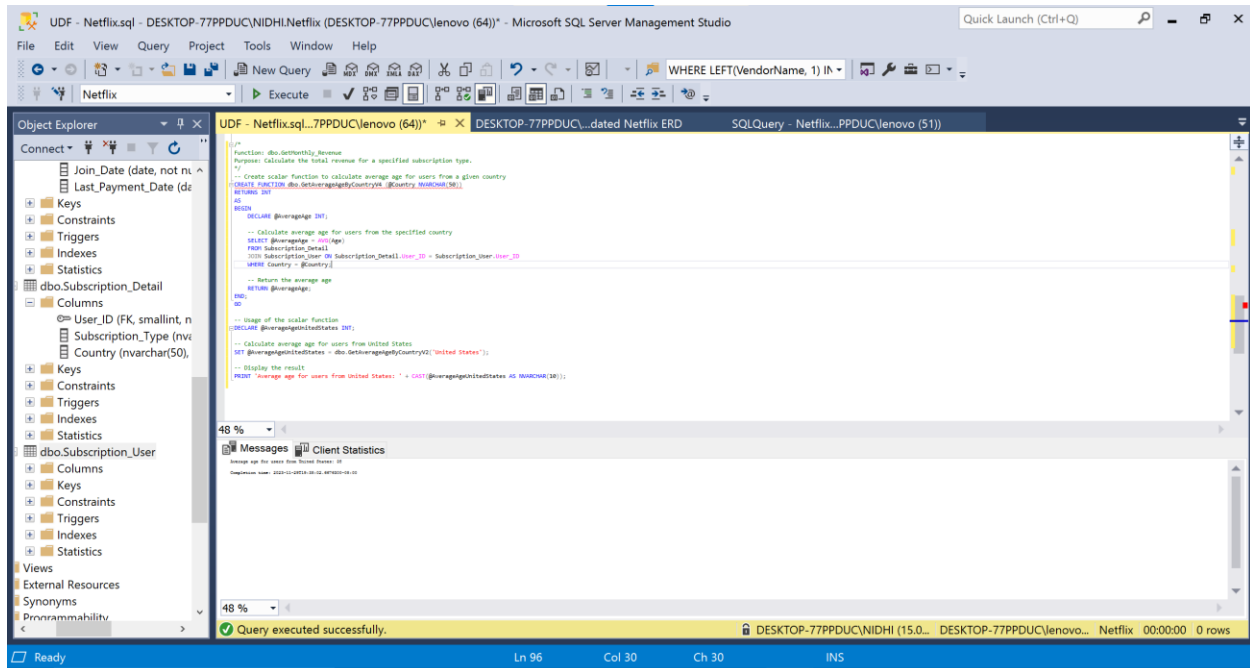


Fig3.15 Computing Average Age Function

We also use stored procedures presented to serve specific analytical purposes within the Netflix user dataset. The first procedure(Fig3.16) retrieves user information based on specified country and subscription types. The second(Fig 3.17) identifies users with premium subscriptions before a specified date, enabling temporal analyses. The third(Fig3.18) allows for targeted exploration by extracting user data with a specified subscription type and age range. The fourth(Fig3.19) aggregates revenue metrics based on user age. The fifth(Fig3.20) provides a multidimensional approach by extracting country and gender-specific data within a given age range. These procedures offer a tailored suite of functionalities for focused and meaningful analyses aligned with specific research objectives within the Netflix user dataset.

Stored procedure for Netflix.sql - DESKTOP-77PPDUC\NIDH1\Netflix (DESKTOP-77PPDUC\lenovo (53)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Netlix

Object Explorer

Connect

Monthly_Revenue (tinyint, null)

Join_Date (date, not null)

Last_Payment_Date (datetime, null)

Keys

Constraints

Triggers

Indexes

Statistics

dbo.Subscription_Detail

Columns

User_ID (FK, smallint, not null)

Subscription_Type (nvarchar(50), not null)

Country (nvarchar(50), not null)

Keys

Constraints

Triggers

Indexes

Statistics

dbo.Subscription_User

Columns

User_ID (PK, smallint, not null)

Age (tinyint, null)

Gender (nvarchar(50), not null)

Device (nvarchar(50), not null)

Plan_Duration (nvarchar(50), not null)

Keys

Constraints

Triggers

Stored procedure f...PPDUC\lenovo (53)*

UDF - Netflix.sql...7PPDUC\lenovo (64)

DESKTOP-77PPDUC...dated Netflix ERD

```

-- Create the stored procedure if it already exists
IF OBJECT_ID('dbo.GetPremiumUsers') IS NOT NULL
    DROP PROCEDURE dbo.GetPremiumUsers;
GO

-- Create the stored procedure
CREATE PROCEDURE dbo.GetPremiumUsers
AS
BEGIN
    -- Select users with Premium subscription before the specified join date
    SELECT *
    FROM Netflix_Subscription AS su
    JOIN Netflix_Subscription_Detail AS sd ON su.User_ID = sd.User_ID
    WHERE Subscription_Type = 'Premium' AND Join_Date < @Join_Date;
END;
GO

-- Execute the stored procedure with a specific join date
EXEC GetPremiumUsers @Join_Date = '2023-01-01';

```

38 %

Results Messages

User_ID	Subscription_Type	Monthly_Revenue	Join_Date	Last_Payment_Date	Country	Age	Gender	Device	Plan_Duration
21	Premium	15	2023-10-06	2023-06-22	United States	28	Female	Laptop	1 Month
31	Premium	15	2023-03-02	2023-06-22	United States	28	Female	Smart TV	1 Month
41	Premium	15	2023-01-20	2023-06-22	United States	31	Female	Laptop	1 Month
51	Premium	15	2023-02-27	2023-06-22	United States	30	Female	Smart TV	1 Month
61	Premium	15	2023-08-02	2023-06-22	United States	28	Female	Laptop	1 Month
81	Premium	10	2023-11-01	2023-06-24	United States	37	Male	Tablet	1 Month
91	Premium	11	2022-03-10	2023-06-23	United States	39	Male	Tablet	1 Month
101	Premium	11	2022-10-17	2023-06-24	United States	46	Female	Tablet	1 Month
111	Premium	14	2022-07-14	2023-06-27	United States	40	Male	Tablet	1 Month
121	Premium	15	2022-03-10	2023-06-24	United States	27	Male	Smart TV	1 Month
131	Premium	15	2022-06-10	2023-06-23	United States	47	Male	Laptop	1 Month

Query executed successfully.

DESKTOP-77PPDUC\NIDH1 (15.0... DESKTOP-77PPDUC\lenovo... Netflix 00:00:00 145 rows

Ready Ln 12 Col 7 Ch 7 INS

Fig3.16 Retrieving User Information Based on Specific Country and Subscription Type

Stored procedure for Netflix.sql - DESKTOP-77PPDUC\NIDH1\Netflix (DESKTOP-77PPDUC\lenovo (53)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Netlix

Object Explorer

Connect

Monthly_Revenue (tinyint, null)

Join_Date (date, not null)

Last_Payment_Date (datetime, null)

Keys

Constraints

Triggers

Indexes

Statistics

dbo.Subscription_Detail

Columns

User_ID (FK, smallint, not null)

Subscription_Type (nvarchar(50), not null)

Country (nvarchar(50), not null)

Keys

Constraints

Triggers

Indexes

Statistics

dbo.Subscription_User

Columns

User_ID (PK, smallint, not null)

Age (tinyint, null)

Gender (nvarchar(50), not null)

Device (nvarchar(50), not null)

Plan_Duration (nvarchar(50), not null)

Keys

Constraints

Triggers

Stored procedure f...PPDUC\lenovo (53)*

UDF - Netflix.sql...7PPDUC\lenovo (64)

DESKTOP-77PPDUC...dated Netflix ERD

```

-- Create a stored procedure to retrieve users with Premium subscription before a specific date
IF OBJECT_ID('dbo.GetPremiumUsers_IDBeforeDate') IS NOT NULL
    DROP PROCEDURE dbo.GetPremiumUsers_IDBeforeDate;
GO

-- Delimiter for batch execution
GO

-- Create the stored procedure
CREATE PROCEDURE dbo.GetPremiumUsers_IDBeforeDate
    @Join_Date DATE
AS
BEGIN
    -- Select users with Premium subscription before the specified join date
    SELECT *
    FROM Netflix_Subscription AS su
    JOIN Netflix_Subscription_Detail AS sd ON su.User_ID = sd.User_ID
    WHERE Subscription_Type = 'Premium' AND Join_Date < @Join_Date;
END;
GO

-- Execute the stored procedure with a specific join date
EXEC GetPremiumUsers_IDBeforeDate '2023-01-01';

```

59 %

Results Messages Client Statistics

User_ID	Monthly_Revenue	Join_Date	Last_Payment_Date	User_ID	Subscription_Type	Country	User_ID	Age	Gender	Device	Plan_Duration
2	15	2021-08-09	2023-06-22	2	Premium	Canada	2	35	Female	Tablet	1 Month
6	15	2022-03-18	2023-06-27	6	Premium	France	6	29	Female	Smart TV	1 Month
16	15	2022-07-04	2023-06-27	16	Premium	France	16	36	Male	Tablet	1 Month
26	15	2022-12-01	2023-06-27	26	Premium	France	26	29	Male	Smartphone	1 Month
29	15	2022-12-19	2023-06-23	29	Premium	Spain	29	36	Female	Laptop	1 Month
36	15	2022-01-03	2023-06-23	36	Premium	France	36	35	Male	Tablet	1 Month
39	15	2022-12-28	2023-06-23	39	Premium	Spain	39	27	Female	Smart TV	1 Month
46	15	2022-02-23	2023-06-27	46	Premium	France	46	34	Male	Smartphone	1 Month
49	15	2022-10-11	2023-06-23	49	Premium	Spain	49	28	Female	Laptop	1 Month
56	15	2022-03-15	2023-06-27	56	Premium	France	56	36	Male	Tablet	1 Month
59	15	2022-03-12	2023-06-23	59	Premium	Spain	59	27	Female	Smart TV	1 Month

Query executed successfully.

DESKTOP-77PPDUC\NIDH1 (15.0... DESKTOP-77PPDUC\lenovo... Netflix 00:00:00 722 rows

Ready Ln 60 Col 77 Ch 77 INS

Fig3.17 Identifying Users with Premium Subscriptions Before a Specified Date

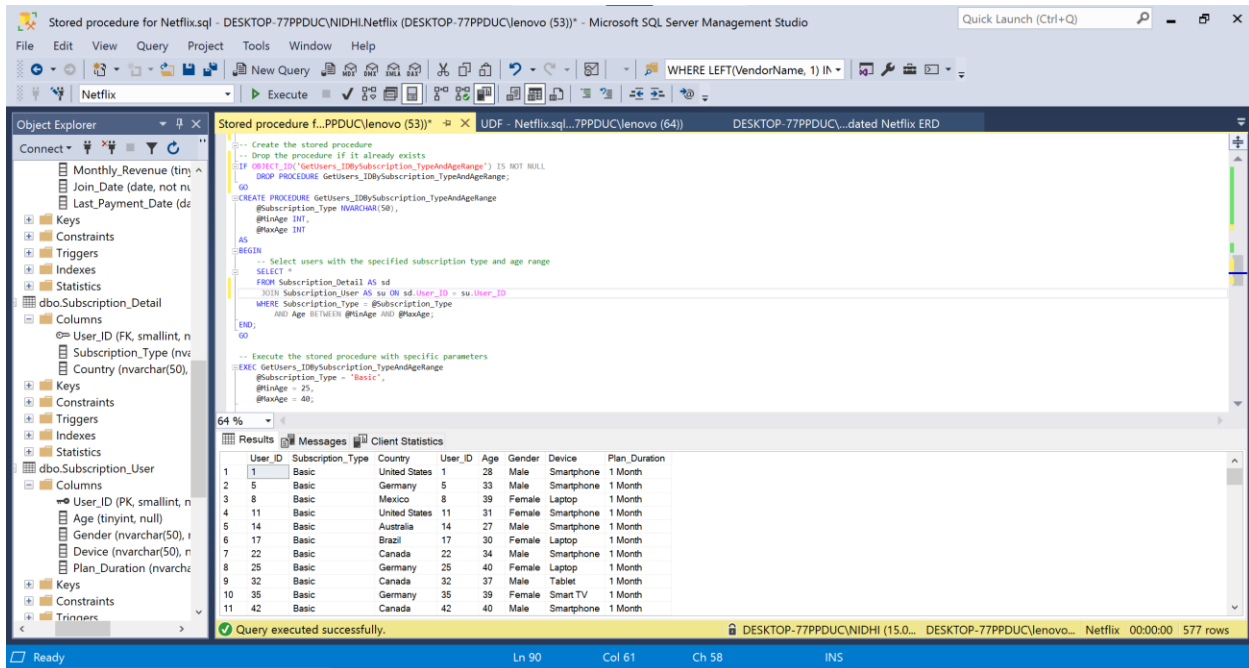


Fig3.18 Extracting User Data with a Specified Subscription Type and Age Range

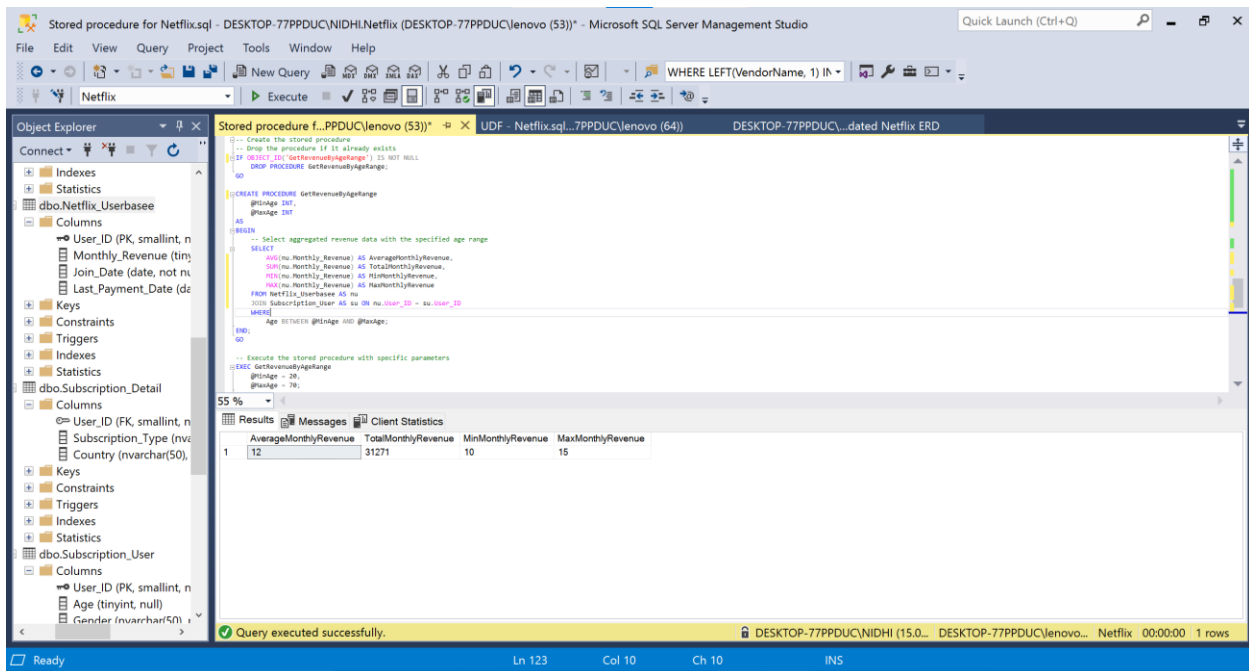


Fig3.19Aggregating Revenue Metrics based on User Age

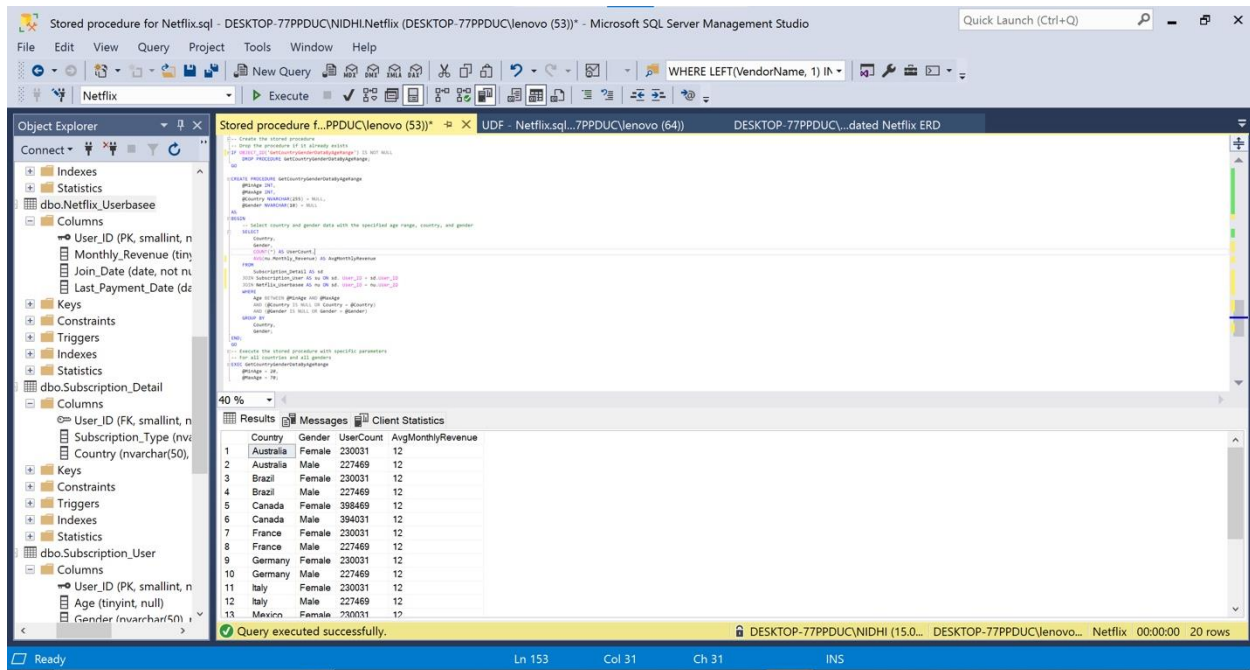


Fig3.20 Extracting Country and Gender-specific Data

4. Summaries

In this comprehensive exploration of Netflix user data using SQL, we initiated our analysis by retrieving all columns and rows from tables, offering a holistic snapshot of user subscription details and demographic information. Subsequent queries delved into revenue analysis, demographic insights on gender and premium subscriptions, device usage patterns, and user subscriptions based on gender, providing nuanced perspectives on user behavior. Examining subscribers who joined in 2021 and their activity status added a temporal dimension to our analysis. Identifying the most common subscription type sheds light on user preferences, revealing a prevalent inclination towards the basic subscription type. Furthermore, we quantified the number of U.S. users, calculated the average user age, and pinpointed the country contributing the highest monthly revenue. These findings collectively form a comprehensive narrative, offering insights into user preferences, revenue dynamics, and regional contributions, contributing to a nuanced understanding of Netflix's diverse user base.

5. Assumptions

In our analysis, several underlying assumptions have been considered to contextualize and interpret the results effectively. We assume that the dataset maintains a high level of consistency and integrity, free from significant anomalies or errors and that foreign key relationships between tables are appropriately established. Additionally, our analysis presumes that the "Subscription_Type" column primarily consists of categories such as 'Basic,' 'Standard,' and 'Premium,' excluding other potential subscription types. The definition of 'active' users in the query assessing subscribers who joined in 2021 is assumed to be contingent upon the absence of specific deactivation or termination status. Moreover, the "Device" column in queries

related to device analysis is assumed to represent consistent and distinct device types without variations. Our revenue calculations rely on the accuracy of the "Monthly_Revenue" column, while queries involving country-specific analysis assume that the "Country" column accurately reflects user geography. Other assumptions encompass gender classification, join date format, data representativeness as a synthetic sample, and accurate user age data. It is crucial to acknowledge and document these assumptions, as they shape the context in which our analysis unfolds and influence the interpretations derived from the dataset.

6. References

- [1] Netflix. "About Netflix." (<https://about.netflix.com/en>)
- [2] Dredge, S. (2017, August 24). "How Netflix's Recommendation System Works." The Guardian.
- [3] Holloway, D. (2019, January 2). "How Netflix Is Using Big Data." Forbes.