Customer Churn and Retention Analysis Project

Project Overview

This project focuses on analyzing customer churn and retention for a subscription service, using tools like Excel, SQL, and Power BI.

Data Collection

A sample dataset was created with the following fields:

CustomerID

StartDate

EndDate

UsageFrequency

AvgSessionDuration

FeedbackScore

Age

Gender

Location

Data Cleaning and Preparation in Excel

Handled Missing Values:
Used the formulas to fill missing EndDate values.

Removed Duplicates:
Used Excel's built-in feature to remove duplicate rows based on Customer ID.

Standardized Data Formats:
Ensured date formats were consistent.

Data Storage in SQL

Loaded Data into SQL

Imported the cleaned data into a SQL database.
Created a table with the following schema:

```
CREATE TABLE CustomerData (
    CustomerID VARCHAR(10),
    StartDate DATE,
    EndDate DATE,
    UsageFrequency INT,
    AvgSessionDuration INT,
```

```
    FeedbackScore INT,
    Age INT,
    Gender VARCHAR(10),
    Location VARCHAR(50)
);
```

Imported the data using the LOAD DATA INFILE command.

Data Analysis and Manipulation in SQL

Calculated Subscription Duration:

```
ALTER TABLE CustomerData
ADD SubscriptionDuration INT;
UPDATE CustomerData
SET SubscriptionDuration = DATEDIFF(IFNULL(EndDate, CURDATE()), StartDate);
```

Categorized Usage Frequency:

Categorized Usage Frequency:

```
ALTER TABLE CustomerData
ADD UsageCategory VARCHAR(20);
UPDATE CustomerData
SET UsageCategory = CASE
    WHEN UsageFrequency >= 5 THEN 'High'
    WHEN UsageFrequency BETWEEN 3 AND 4 THEN 'Medium'
    ELSE 'Low'
END;
```

Calculated Customer Lifetime Value (CLTV):

```
ALTER TABLE CustomerData
ADD CLTV INT;
UPDATE CustomerData
SET CLTV = AvgSessionDuration * UsageFrequency * 10;
```

Identified Customer Churn:

```
ALTER TABLE CustomerData
ADD IsChurned VARCHAR(10);

UPDATE CustomerData
SET IsChurned = CASE
    WHEN EndDate IS NOT NULL AND EndDate < CURDATE() THEN 'Yes'
    WHEN EndDate IS NULL THEN 'No'
    ELSE 'No'
END;
```

Churn Rate:
SELECT COUNT(*) * 100.0 / (SELECT COUNT(*) FROM CustomerData) AS ChurnRate
FROM CustomerData
WHERE IsChurned = 'Yes';

Average Subscription Duration:
SELECT AVG(SubscriptionDuration) AS AvgSubscriptionDuration
FROM CustomerData;

Average Customer Lifetime Value:
SELECT AVG(CLTV) AS AvgCLTV
FROM CustomerData;
Data Visualization in Power BI
Imported Data:
Loaded the CSV file into Power BI.

Created Visuals

Churn Rate Analysis: Pie chart with IsChurned as the legend and count of CustomerID as values.

Usage Frequency Categories: Bar chart with UsageCategory on the X-axis and count of CustomerID on the Y-axis.

Subscription Duration: Line chart with StartDate on the X-axis and average of SubscriptionDuration on the Y-axis.

CLTV Analysis: Scatter plot with Age on the X-axis, CLTV on the Y-axis, UsageCategory as the legend, and CustomerID as values.

Conclusion

This project showcases the entire data analytics workflow, from data collection and cleaning to analysis and visualization, providing valuable insights into customer behavior and retention.