Database Management System for JIO HOTSTAR

Name : Praveen Choudhary

Prn : 23070243065

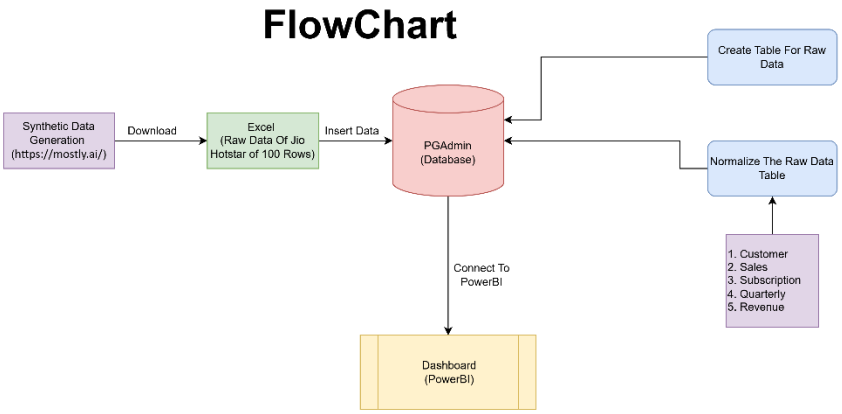
Subject : Data Engineering

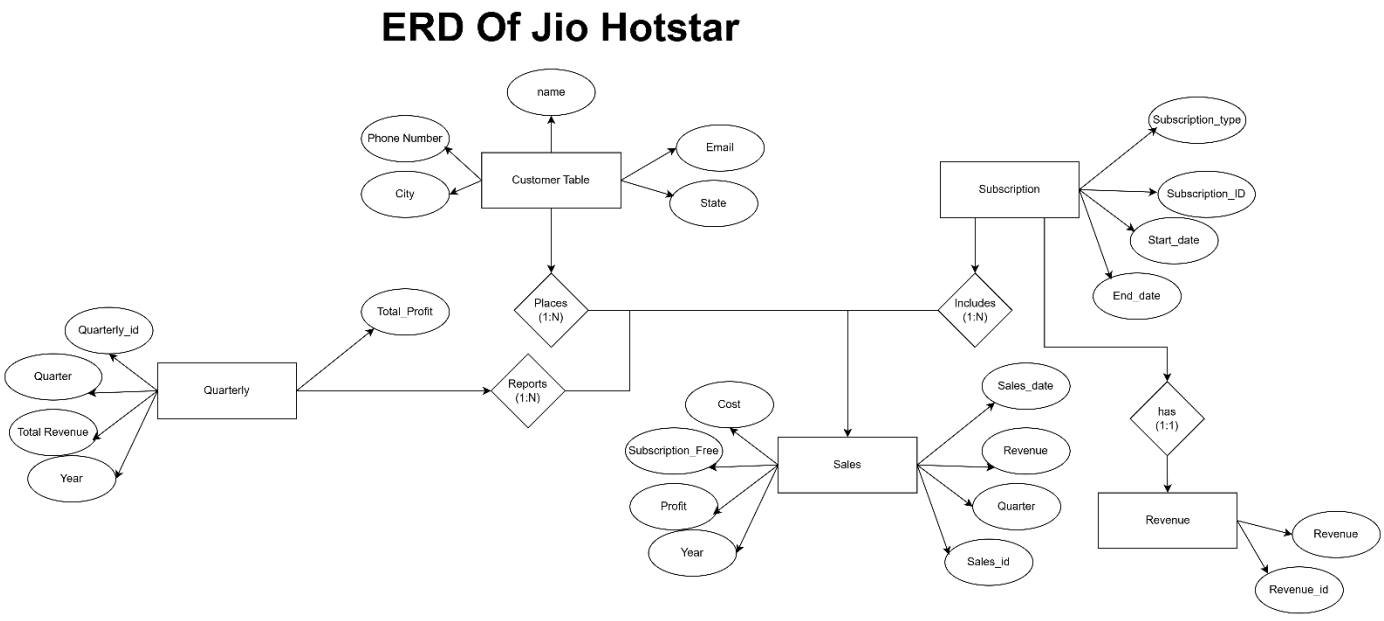
Drive Link: https://drive.google.com/drive/folders/15V4RWHrTSWL6ALuTk2vA6mVT7IRG61ld?usp=sharing

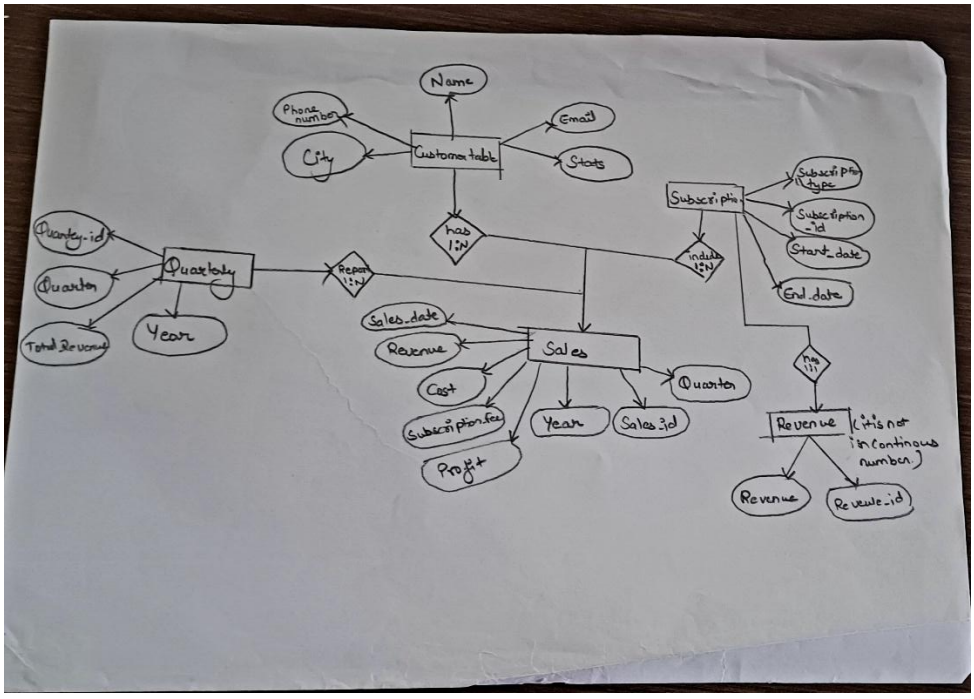
1. **Generating Synthetic Data:**

Raw Data is generate from https://mostly.ai/ .

2. **Flow Chat and ERD Diagram:**







ERD (Entity-Relationship Diagram) Description:

Entities and Attributes:

1. **Customer**
   - **Customer_ID (PK)**
   - **Customer_Name**
   - **Email**
   - **Phone_Number**
   - **City**
   - **State**
2. **Subscription**
   - **Subscription_ID (PK)**
   - **Subscription_Type**
   - **Start_Date**
   - **End_Date**
3. **Sales**
   - **Sales_ID (PK)**
   - **Customer_ID (FK to Customer)**
   - **Subscription_ID (FK to Subscription)**
   - **Sales_Date**
   - **Subscription_Fee**
   - **Revenue**
   - **Cost**
   - **Profit**
   - **Year**
   - **Quarter**
   - **(FK to Quarterly on Year and Quarter)**
4. **Revenue**
   - **Revenue_ID (PK)**
   - **Subscription_ID (FK to Subscription)**
   - **Revenue**
5. **Quarterly**
   - **Quarterly_ID (PK)**
   - **Year (Unique constraint with Quarter)**
   - **Quarter (Unique constraint with Year)**
   - **Total_Revenue**

- **Total_Profit**

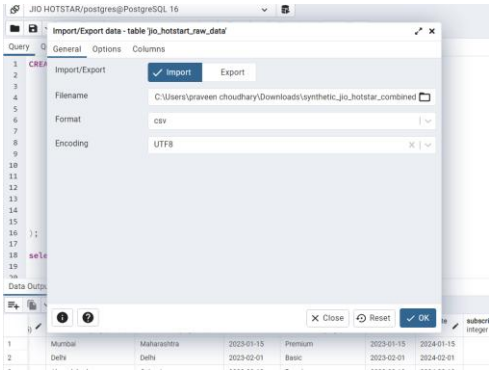
Relationships:

- **Customer to Sales: One-to-Many (A customer can have many sales, but each sale belongs to one customer).**
- **Subscription to Sales: One-to-Many (A subscription can be linked to many sales, but each sale references one subscription).**
- **Subscription to Revenue: One-to-One (Each subscription has a revenue entry).**
- **Quarterly to Sales: One-to-Many (A quarter can have many sales, but each sale belongs to one quarter based on year and quarter).**

3. Insert Data Into Data Base

Query | Query History

```
CREATE TABLE jio_hotstart_raw_data (
    Sr_No SERIAL PRIMARY KEY,
    Customer_Name VARCHAR(255),
    Email VARCHAR(255),
    Phone_Number VARCHAR(15),
    City VARCHAR(255),
    State VARCHAR(255),
    SignUp_Date DATE,
    Subscription_Type VARCHAR(50),
    Start_Date DATE,
    End_Date DATE,
    Subscription_Fee INT,
    Renewal_Status VARCHAR(50),
    Revenue INT,
    Cost INT
);
```



Show the raw data:

```
select * from jio_hotstart_raw_data
```

Data Output | Messages | Notifications

	sr_no [PK] integer	customer_name character varying (255)	email character varying (255)	phone_number character varying (15)	city character varying (255)	state character varying (255)	signup_date date	subscription_type character varying (50)	start date
1	1	John Doe	john.doe@example.com	9876543210	Mumbai	Maharashtra	2023-01-15	Premium	2023
2	2	Mary Smith	mary.smith@example.com	9123456789	Delhi	Delhi	2023-02-01	Basic	2023
3	3	Raj Patel	raj.patel@example.com	9056782345	Ahmedabad	Gujarat	2023-03-10	Premium	2023
4	4	Priya Gupta	priya.gupta@example.com	9001234567	Bangalore	Karnataka	2023-04-22	Basic	2023
5	5	Arun Kumar	arun.kumar@example.com	9532147890	Kolkata	West Bengal	2023-05-12	Premium	2023
6	6	Neha Singh	neha.singh@example.com	9965432109	Chennai	Tamil Nadu	2023-06-09	Basic	2023
7	7	Aakash Sharma	aakash.sharma@example.com	9876512345	Pune	Maharashtra	2023-07-01	Premium	2023
8	8	Meera Joshi	meera.joshi@example.com	9832105647	Jaipur	Rajasthan	2023-08-19	Basic	2023
9	9	Sanjay Verma	sanjay.verma@example.com	9021345678	Lucknow	Uttar Pradesh	2023-09-05	Premium	2023
10	10	Rita Kapoor	rita.kapoor@example.com	9445678901	Chandigarh	Chandigarh	2023-10-15	Basic	2023

Step 1: Create Tables for Normalization

1.1 Create Customer Table

```sql
CREATE TABLE Customer (
    Customer_ID SERIAL PRIMARY KEY,
    Customer_Name VARCHAR(255),
    Email VARCHAR(255),
    Phone_Number VARCHAR(15),
    City VARCHAR(255),
    State VARCHAR(255)
);
select * from Customer
```

Data Output | Messages | Notifications

customer_id [PK] integer	customer_name character varying (255)	email character varying (255)	phone_number character varying (15)	city character varying (255)	state character varying (255)

1.2 Sales Table:

```sql
CREATE TABLE Sales (
    Sales_ID SERIAL PRIMARY KEY,
    Customer_ID INT REFERENCES Customer(Customer_ID),
    Subscription_ID INT REFERENCES Subscription(Subscription_ID),
    Sales_Date DATE,
    Subscription_Fee INT,
    Revenue INT,
    Cost INT,
    Profit INT
);

select * from Sales
```

Data Output | Messages | Notifications

sales_id [PK] integer	customer_id integer	subscription_id integer	sales_date date	subscription_fee integer	revenue integer	cost integer	profit integer

1.3 Subscription Table:

```sql
CREATE TABLE Sales (
    Sales_ID SERIAL PRIMARY KEY,
    Customer_ID INT REFERENCES Customer(Customer_ID),
    Subscription_ID INT REFERENCES Subscription(Subscription_ID),
    Sales_Date DATE,
    Subscription_Fee INT,
    Revenue INT,
    Cost INT,
    Profit INT
);

select * from Sales
```

Data Output | Messages | Notifications

sales_id [PK] integer	customer_id integer	subscription_id integer	sales_date date	subscription_fee integer	revenue integer	cost integer	profit integer

1.4 Revenue Table:

```sql
CREATE TABLE Revenue (
    Revenue_ID SERIAL PRIMARY KEY,
    Subscription_ID INT REFERENCES Subscription(Subscription_ID),
    Revenue INT
);

select * from Revenue
```

Data Output | Messages | Notifications

revenue_id [PK] integer	subscription_id integer	revenue integer

1.5 Quarterly Table:

```sql
CREATE TABLE Quarterly (
    Quarterly_ID SERIAL PRIMARY KEY,
    Year INT,
    Quarter INT,
    Total_Revenue INT,
    Total_Profit INT
);
select * from Quarterly
```

Data Output | Messages | Notifications

quarterly_id [PK] integer	year integer	quarter integer	total_revenue integer	total_profit integer

Step 2: Insert Data from jio_hotstart_raw_data

2.1 Insert Customer Data:

```
INSERT INTO Customer (Customer_Name, Email, Phone_Number, City, State)
SELECT DISTINCT Customer_Name, Email, Phone_Number, City, State
FROM jio_hotstart_raw_data;

select * from Customer
```

Data Output | Messages | Notifications

customer_id [PK] integer	customer_name character varying (255)	email character varying (255)	phone_number character varying (15)	city character varying (255)	state character varying (255)
1	Sahil Mehra	sahil.mehra@example.com	9032145678	Ahmedabad	Gujarat
2	Neeraj Yadav	neeraj.yadav@example.com	9654321098	Pune	Maharashtra
3	Rani Sethi	rani.sethi@example.com	9567123456	Surat	Gujarat
4	Vivek Agarwal	vivek.agarwal@example.com	9876453120	Kolkata	West Bengal
5	Tanvi Verma	tanvi.verma@example.com	9638527410	Pune	Maharashtra
6	Priya Singh	priya.singh@example.com	9412384756	Lucknow	Uttar Pradesh
7	Rekha Sharma	rekha.sharma@example.com	9923456789	Bangalore	Karnataka

2.2 Insert Subscription Data:

```
INSERT INTO Subscription (Subscription_Type, Start_Date, End_Date)
SELECT DISTINCT Subscription_Type, Start_Date, End_Date
FROM jio_hotstart_raw_data;

select * from Subscription
```

Data Output | Messages | Notifications

subscription_id [PK] integer	subscription_type character varying (50)	start_date date	end_date date
1	Premium	2023-01-20	2024-01-20
2	Basic	2023-12-05	2024-12-05
3	Premium	2023-09-05	2024-09-05
4	Premium	2023-07-14	2024-07-14
5	Premium	2023-05-05	2024-05-05
6	Premium	2023-11-18	2024-11-18

2.3 Insert Sales Data:

```
INSERT INTO Sales (Customer_ID, Subscription_ID, Sales_Date, Subscription_Fee, Revenue, Cost, Profit)
SELECT
    c.Customer_ID,
    s.Subscription_ID,
    j.SignUp_Date AS Sales_Date,
    j.Subscription_Fee,
    j.Revenue,
    j.Cost,
    j.Revenue - j.Cost AS Profit
FROM jio_hotstart_raw_data j
JOIN Customer c ON j.Customer_Name = c.Customer_Name AND j.Email = c.Email
JOIN Subscription s ON j.Subscription_Type = s.Subscription_Type;

select * from Sales
```

Data Output | Messages | Notifications

sales_id [PK] integer	customer_id integer	subscription_id integer	sales_date date	subscription_fee integer	revenue integer	cost integer	profit integer
1	47	1	2023-01-25	499	499	280	219
2	27	1	2023-11-11	499	499	300	199
3	40	1	2023-11-11	499	499	300	199
4	38	1	2023-09-02	499	499	250	249
5	4	1	2023-07-06	499	499	270	229
6	49	1	2023-05-22	499	499	280	219

2.4 Insert Revenue Data:

```
INSERT INTO Revenue (Subscription_ID, Revenue)
SELECT s.Subscription_ID, j.Revenue
FROM jio_hotstart_raw_data j
JOIN Subscription s ON j.Subscription_Type = s.Subscription_Type;

select * from Revenue
```

Data Output | Messages | Notifications

revenue_id [PK] integer	subscription_id integer	revenue integer
1	61	499
2	60	499
3	59	499
4	58	499

2.5 Insert Quarterly Data:

```
INSERT INTO Quarterly (Year, Quarter, Total_Revenue, Total_Profit)
SELECT
    EXTRACT(YEAR FROM j.SignUp_Date) AS Year,
    EXTRACT(QUARTER FROM j.SignUp_Date) AS Quarter,
    SUM(j.Revenue) AS Total_Revenue,
    SUM(j.Revenue - j.Cost) AS Total_Profit
FROM jio_hotstart_raw_data j
GROUP BY EXTRACT(YEAR FROM j.SignUp_Date), EXTRACT(QUARTER FROM j.SignUp_Date);

select * from Quarterly
```

Data Output | Messages | Notifications

	quarterly_id [PK] integer	year integer	quarter integer	total_revenue integer	total_profit integer
1	1	2023	4	5386	1596
2	2	2023	3	5388	948
3	3	2023	1	7184	2554
4	4	2023	2	5984	2064

2.6 Make Related Between the Quarterly And Sales Table

```
ALTER TABLE Sales
ADD Year INT,
ADD Quarter INT;

UPDATE Sales
SET Year = EXTRACT(YEAR FROM Sales_Date),
    Quarter = EXTRACT(QUARTER FROM Sales_Date);

ALTER TABLE Quarterly
ADD CONSTRAINT uq_year_quarter UNIQUE (Year, Quarter);

ALTER TABLE Sales
ADD CONSTRAINT fk_sales_quarterly FOREIGN KEY (Year, Quarter)
REFERENCES Quarterly (Year, Quarter);

SELECT s.Sales_ID, s.Year, s.Quarter, q.Total_Revenue, q.Total_Profit
FROM Sales s
JOIN Quarterly q ON s.Year = q.Year AND s.Quarter = q.Quarter;
```

Data Output | Messages | Notifications

	sales_id integer	year integer	quarter integer	total_revenue integer	total_profit integer
1	1	2023	1	7184	2554
2	2	2023	4	5386	1596
3	3	2023	4	5386	1596
4	4	2023	3	5388	948
5	5	2023	3	5388	948
6	6	2023	2	5984	2064
7	7	2023	1	7184	2554

4. Drop jio_hotstart_raw_data

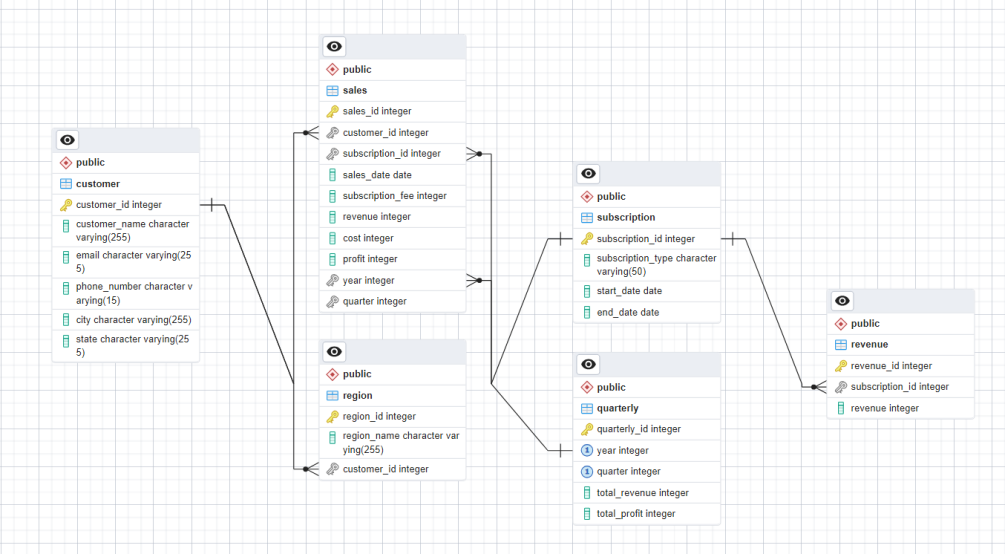
```
drop table jio_hotstart_raw_data
```

Data Output | Messages | Notifications

```
DROP TABLE

Query returned successfully in 195 msec.
```

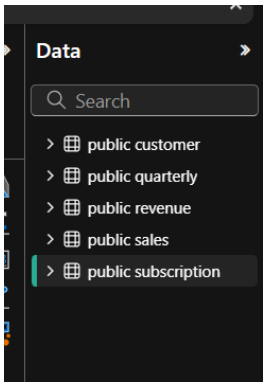
3 ERD Diagram For Jio_Hotstar DataBase:



4 Connect PGAdmin To PowerBi:

For connection use server as localhost:5433 and pgadmin database name JIO_HOTSTAR

After connection we get all the table:



Some Query

1. Query to Get Total Revenue, Profit, and Number of Customers by Subscription Type:

```sql
SELECT
    s.Subscription_Type,
    COUNT(DISTINCT c.Customer_ID) AS Total_Customers,
    SUM(sales.Revenue) AS Total_Revenue,
    SUM(sales.Profit) AS Total_Profit
FROM
    Sales sales
JOIN
    Subscription s ON sales.Subscription_ID = s.Subscription_ID
JOIN
    Customer c ON sales.Customer_ID = c.Customer_ID
GROUP BY
    s.Subscription_Type
ORDER BY
    Total_Revenue DESC;
```

Data Output | Messages | Notifications

	subscription_type character varying (50)	total_customers bigint	total_revenue bigint	total_profit bigint
1	Premium	37	541415	203515
2	Basic	36	224250	30750

Subscription Summary: Premium subscriptions generate higher revenue (₹541,415) and profit (₹203,515) compared to Basic, despite similar customer counts.

2. Query to Get Quarterly Revenue and Profit Breakdown:

Query | Query History

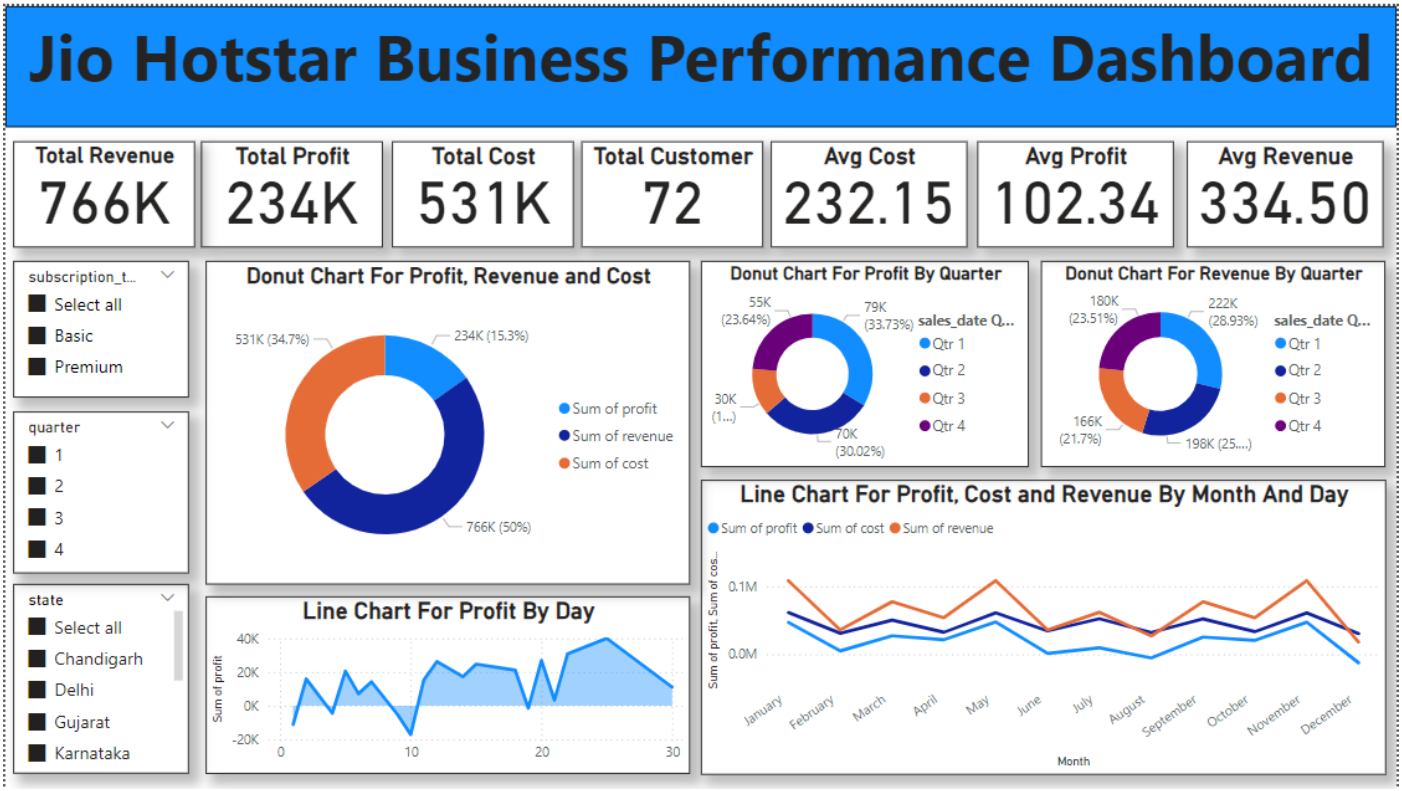
```
1  SELECT
2      q.Year,
3      q.Quarter,
4      q.Total_Revenue,
5      q.Total_Profit
6  FROM
7      Quarterly q
8  ORDER BY
9      q.Year, q.Quarter;
```

Data Output | Messages | Notifications

	year integer	quarter integer	total_revenue integer	total_profit integer
1	2023	1	7184	2554
2	2023	2	5984	2064
3	2023	3	5388	948
4	2023	4	5386	1596

Quarterly Performance: Revenue and profit peaked in Q1 2023 and steadily declined through Q4, indicating potential seasonality or operational changes.

5 Dashboard 1 For Business Performance Dashboard



Jio Hotstar Business Performance Dashboard Overview

The Jio Hotstar Business Performance Dashboard provides a comprehensive view of the key metrics related to the company's financial and customer performance. It is designed to offer actionable insights through clear visualizations and summary statistics. Below are the key components of the dashboard:

1. Key Performance Indicators (KPIs):

- Total Revenue: Displays the total income generated (766K) across all subscription types.
- Total Profit: Shows the net earnings (234K) after deducting costs.
- Total Cost: Represents the operational expenses (531K) incurred.
- Total Customers: Indicates the total number of customers (72).
- Average Cost per Customer: The mean cost (232.15K) calculated per customer.
- Average Profit per Customer: The mean profit (102.34K) calculated per customer.
- Average Revenue per Customer: The average income (334.50K) generated per customer.

2. Visualizations:

- **Donut Charts:**
    - Profit, Revenue, and Cost Distribution: Highlights the proportion of profit, revenue, and cost in total operations.
    - Profit and Revenue by Quarter: Shows the quarterly breakdown for better time-based analysis. For instance, Quarter 3 contributes significantly to profit (33.73%).
- **Line Charts:**
    - Profit by Day: Depicts daily profit trends, showing fluctuations and peak performance days.
    - Monthly Trends: Tracks monthly changes in profit, cost, and revenue to identify seasonal patterns.

3. Filters:

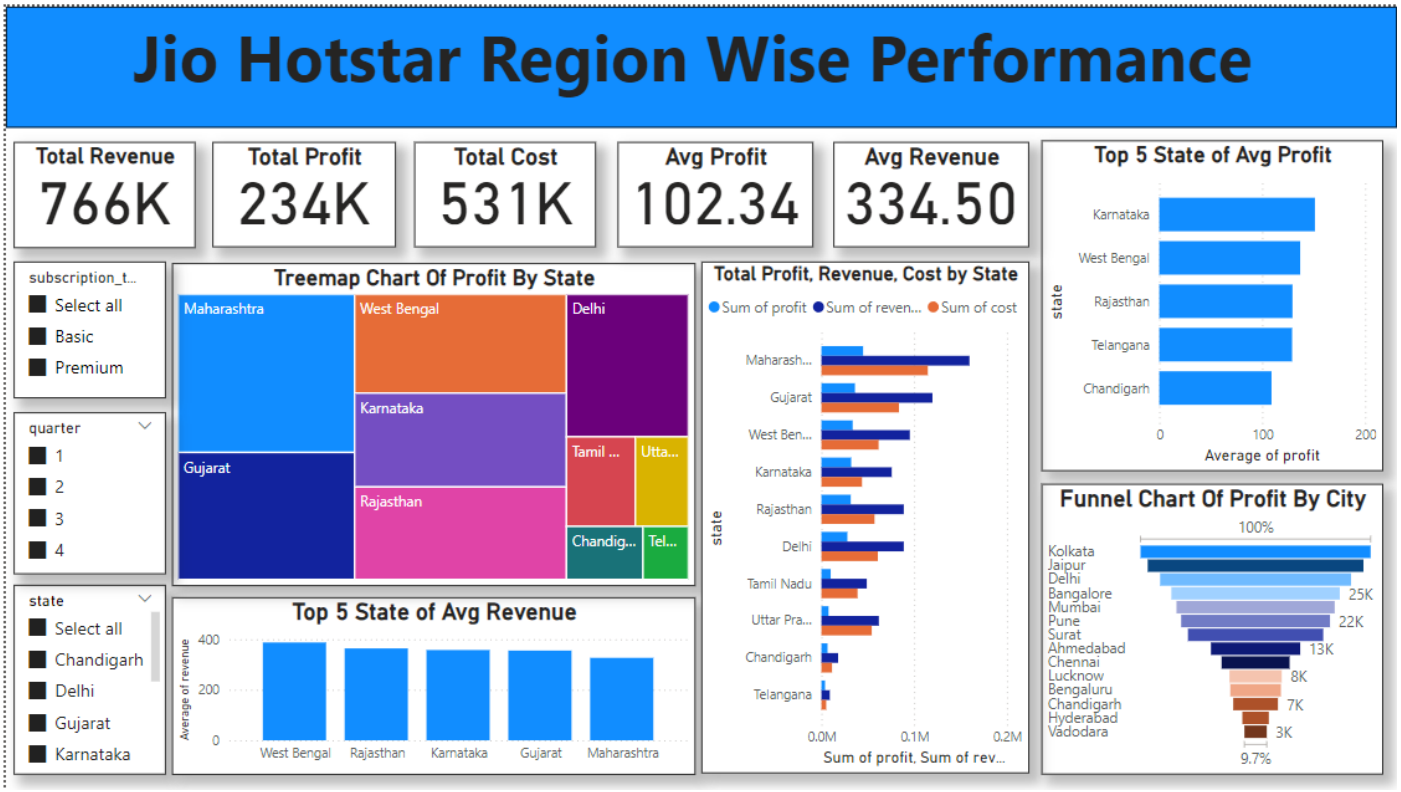
- Subscription Type Filter: Allows segmentation of data based on subscription plans (e.g., Basic, Premium).
- Quarter Filter: Enables analysis of specific quarters (1, 2, 3, or 4).
- State Filter: Provides insights by region, such as Chandigarh, Delhi, Gujarat, and Karnataka.

4. Insights from Visuals:

- The donut charts reveal a balanced revenue and cost structure, with clear profitability in Q3.
- Line charts highlight patterns in profitability over time, helping to pinpoint days or months with optimal performance or high costs.
- Regional and subscription-based filters allow focused analysis to aid in strategic decision-making.

---

6 Dashboard 2 For Region wise Performance Dashboard



Jio Hotstar Region-Wise Performance Dashboard

This dashboard provides a detailed analysis of Jio Hotstar's performance across different states and cities, highlighting key regional insights and overall business metrics.

Key Components and Visualizations

1. Key Performance Indicators (KPIs):
    - Total Revenue: 766K, showcasing the total income generated across all regions.
    - Total Profit: 234K, reflecting the net earnings after subtracting costs.
    - Total Cost: 531K, representing the operational expenses across regions.
    - Average Profit per Customer: 102.34K, indicating the mean profit per user.
    - Average Revenue per Customer: 334.50K, showing the income generated on average per customer.
2. Treemap Chart for Profit by State:
    - Displays a hierarchical view of profits by state, with larger blocks representing higher profits. Maharashtra, West Bengal, Gujarat, and Karnataka are prominent contributors.

3. Bar Chart for Total Profit, Revenue, and Cost by State:
    - Highlights the financial distribution per state. States like Maharashtra and Gujarat show higher revenue and profit margins compared to others.
4. Top 5 States by Average Profit:
    - Karnataka leads with the highest average profit, followed by West Bengal, Rajasthan, Telangana, and Chandigarh.
5. Top 5 States by Average Revenue:
    - West Bengal, Rajasthan, Karnataka, Gujarat, and Maharashtra are the leading states in terms of average revenue generation.
6. Funnel Chart for Profit by City:
    - Highlights profit contribution from major cities, with Kolkata, Jaipur, and Delhi being the top contributors.

Conclusions

From the Business Performance Dashboard:

- Jio Hotstar has a healthy revenue stream, with a total revenue of 766K and a profit margin of 234K.
- Quarter-wise performance indicates a significant profit contribution from Q3 (33.73% of total profit).
- Monthly trends reveal key periods of high profit and cost, enabling better planning for peak times.

From the Region-Wise Performance Dashboard:

- Maharashtra, Gujarat, and Karnataka stand out as key states in terms of revenue and profitability.
- Cities like Kolkata, Jaipur, and Delhi are major profit contributors, suggesting a strong customer base in urban regions.
- Karnataka leads in average profit, indicating higher profitability per customer in this state.