

# **OLA RIDE DATA ANALYSIS - PUNE**

**SQL-Based Data Exploration**

**Presented By  
Sanika Mestry**

# Project Overview

- \* Objective: Analyze Ola rides in Pune using SQL
- \* Goal: Explore ride patterns, user behavior, and driver performance to generate actionable insights

# Dataset Description

1. Ride Details: Pickup and drop locations, ride type, distance
2. User Information: User ID, Name
3. Driver Information: Driver ID, Rating
4. Ride Status: Completed, Cancelled
5. Additional Fields: Booking Time, Payment Mode, Fare

# Key SQL Queries

## 1. Completed Bookings

```
CREATE VIEW complete_booking AS  
SELECT * FROM booking WHERE  
Ride_Status = 'completed';
```

## 2. Average Distance by Vehicle Type:

```
CREATE VIEW ride_distance_for_each_vehicle AS  
SELECT Ride_Type, AVG(Distance) AS avg_Distance  
FROM booking GROUP BY Ride_Type;
```

### 3. Cancelled Rides Count:

```
SELECT COUNT(*) FROM booking WHERE  
    Ride_Status = 'cancelled';
```

# User and Driver Insights

## 1. Top Users by Rides:

```
SELECT User_Name,  
       COUNT(User_Name) AS total_rides  
FROM booking GROUP BY User_Name  
ORDER BY total_rides DESC LIMIT 5;
```

## 2. Driver Ratings for Prime Rides:

```
SELECT MAX(Driver_Rating) AS  
       max_rating, MIN(Driver_Rating) AS  
       min_rating  
FROM booking WHERE Ride_Type =  
       'Prime';
```

### 3. Rides Using UPI Payment:

```
SELECT * FROM booking WHERE Payment_Mode = 'UPI';
```

# Ride Demand Analysis

## 1. Ride Demand by Time of Day

```
SELECT CASE
WHEN HOUR(Booking_Time) BETWEEN 0 AND 5 THEN 'Night (12am-5am)'
WHEN HOUR(Booking_Time) BETWEEN 6 AND 11 THEN 'Morning (6am-11am)'
WHEN HOUR(Booking_Time) BETWEEN 12 AND 16 THEN 'Afternoon (12pm-4pm)'
WHEN HOUR(Booking_Time) BETWEEN 17 AND 20 THEN 'Evening (5pm-8pm)'
ELSE 'Late Night (9pm-11pm)'
END AS time_of_day, COUNT(*) AS ride_count
FROM booking GROUP BY time_of_day ORDER BY ride_count DESC;
```

## 2. Ride Type Distribution

```
SELECT Ride_Type, COUNT(*) AS ride_count  
FROM booking GROUP BY Ride_Type ORDER  
BY ride_count DESC;
```

# Location and Time Trends

1. Ride Demand by Location:

```
SELECT Pickup_Location, COUNT(*) AS  
ride_count FROM booking GROUP BY  
Pickup_Location ORDER BY ride_count DESC;
```

2. Completed Rides per User:

```
SELECT User_Name, COUNT(*) FROM booking  
WHERE Ride_Status='completed' GROUP BY  
User_Name;
```

### 3. Total Rides Per Day:

```
SELECT DATE(Booking_Time) AS ride_date,  
       COUNT(*) AS total_rides FROM booking  
GROUP BY ride_date ORDER BY ride_date;
```

# Fare and Route Analysis

1. Average Fare by Ride Type:

```
SELECT Ride_Type, ROUND(AVG(Fare),2) AS  
avg_fare FROM booking GROUP BY  
Ride_Type;
```

2. Rides by Hour of Day:

```
SELECT HOUR(Booking_Time) AS hour, COUNT(*) AS total_rides  
FROM booking GROUP BY hour ORDER BY hour;
```

### 3. Most Frequent Routes:

```
SELECT Pickup_Location, Drop_Location,  
COUNT(*) AS ride_count FROM booking GROUP  
BY Pickup_Location, Drop_Location ORDER BY  
ride_count DESC LIMIT 5;
```

# Insights & Recommendations

1. Peak Demand: Focus driver availability during peak hours
2. Popular Routes\*: Optimize services along frequently traveled routes
3. Payment Preferences\*: UPI is widely used; offer related incentives
4. Driver Performance\*: High ratings on Prime rides; replicate training model

# Conclusion

- 1.SQL-based analysis helped reveal crucial patterns in Ola's operations in Pune
- 2.These insights can improve user experience, driver management, and strategic planning.
- 3.Demonstrates the value of data-driven decision making in ride-hailing services



# **THANK YOU**