UBER RIDE BOOKING ANALYSIS (EDA SUMMARY)

1. Ride Volume & Patterns

- **Peak ride hours:** 5 PM–7 PM (office return traffic).
- **Lowest ride hours:** Midnight to 5 AM.
- **Rides per weekday:** Almost evenly distributed; demand is stable across the week.
- **Rides per month:** Uniform throughout the year; no major seasonality.

2. Vehicle Type Usage

- **Auto, Go Mini, Go Sedan** dominate total rides.
- **EBike & Uber XL** have the lowest usage.
- Customers prefer small and affordable vehicles.

3. Payment Behaviour

- **UPI** is the most used method (~45k rides).
- **Cash** is second-highest.
- All payment methods show **~91–92% completion rate**, meaning cancellations don't depend on payment type.

4. Booking Status / Ride Outcomes

- **Completed rides:** 93,000 (majority).
- **Cancelled by Driver:** 27,000.
- **No Driver Found:** 10,500.
- **Cancelled by Customer:** 10,500.
- **Incomplete:** 9,000.

Driver-side issues (cancellation + no driver) make up **~25% of all failures**, dominating non-completions.

5. Fare & Distance Insights

- **Fare distribution:** Right-skewed; most rides are ₹200–₹700.
- **Distance distribution:** Mostly 5–40 km.
- **Fare vs Distance correlation:** Extremely low (~0.005).
    - Pricing appears fixed-band or has caps → not proportional to distance.
- **Fare per km:** Highest for **Go Sedan**, lowest for **EBike**.

6. Ratings Patterns

- Driver Rating → Avg **4.23**
- Customer Rating → Avg **4.40**

Both ratings are high with very low variation.

- No relationship between **ratings & distance**.
- No relationship between **driver and customer ratings**.

Ratings mostly independent of ride characteristics.

7. Cancellations & Incomplete Rides

Driver Cancellations

- Peak between **5 PM–7 PM** across all vehicles.
- Highest for **Auto and Go Mini** (supply issues during demand spikes).

No Driver Found

- Also peaks in evening → strong demand–supply gap.

Incomplete Rides

- Main reasons:
  - Customer Demand
  - Other Issue
  - Vehicle Breakdown
- Auto has the highest incomplete count.

8. Multivariate Heatmaps

- **Ride hour × Weekday heatmap** shows:
  - Consistent patterns across all days.
  - Evening demand spike is universal.
- **Ride hour × Vehicle type cancellations**:
  - Autos show the largest spikes → operational bottleneck.

FINAL SUMMARY

Demand is stable year-round with strong evening peaks, autos and small cars dominate trips, cancellations are primarily driver-related during peak hours, pricing doesn't scale strongly with distance, and ratings remain high regardless of ride characteristics.