Maha Kumbh Mela 2025 - Full Data Journey & Analysis

About the Event

The **Maha Kumbh Mela** is the **world's largest public gathering**, rooted in Hindu mythology and attracting ascetics, saints, sadhus, sadhvis, kalpvasis, and millions of pilgrims from across the world. It rotates between **Haridwar, Ujjain, Nashik, and Prayagraj**, based on unique astrological alignments of the Sun, Moon, and Jupiter.

In **2025**, the Maha Kumbh Mela is held in **Prayagraj** from **January 13 to February 26**, at the confluence of the Ganga, Yamuna, and the mythical Sarasvati.

The Challenge

When I began my analysis, there was **no ready-made dataset** anywhere online.
The required numbers existed only in **scattered news reports**, primarily from Times of India and other sources.

Steps I Took:

1. **Data Discovery** – Reading dozens of articles for fragmented statistics on footfall, meal distribution, weather, and incidents.

2. **Data Structuring** – Turning text snippets into a consistent Excel format.

3. **Prompt Engineering** – Using ChatGPT to produce realistic projections after multiple refinements.

4. **Data Cleaning** – Filling missing values using **mean**, **median**, and **standard deviation**.

5. **Validation** – Checking numbers for logical consistency before moving forward.

Data Transformation

Before: Disconnected facts, missing values, inconsistent formats.

After: A clean Excel dataset with:

- Date & Day

- Total Footfall

- Age & gender split

- Foreign tourist %

- Peak hours

- Processions count

Before – Scattered Raw Data	After – Clean Structured Data

Snippet
Footfall: around 2 million visitors reported on Jan 13
Friday saw the largest crowd
Medical emergencies recorded in news articles
Meal distribution happened daily
Weather: Rainy days had high crowd turnout

Date	Total Footfall	Male %	Female %	Foreign Tourists	Security Incidents
10-01-2025	1390652	50	50	5	10
11-01-2025	1150221	51	49	8	0
12-01-2025	2810217	59	41	4	1

Power BI Dashboard

Key Insights from Analysis

- **Dashashwamedh Ghat** had the highest footfall and meals served.

- **Rainy & sunny days** brought the largest crowds.

- **Fridays** were the busiest for food distribution.

- Medical & security teams were most active in high-footfall Ghats.

Conclusion

This wasn't just a visualization task — it was **data creation from scratch**.
It required **research, structuring, cleaning, and validation** before visualization in Power BI.
The outcome is not only a dashboard that tells a compelling story but also a **blueprint for future large-scale event planning**.