

Project Report

Analysis & Segmentation of Top Indian Places to Visit

Focus: Geographic, Significance, and Google Review variables

Prepared By:

Mrinmoy Mondal

Mayank Kumar

Machine Learning Intern, Fynn Lab

14th August, 2025

1. Dataset Overview

This dataset contains curated information about 325 tourist attractions across India, including Zone, State, City, Name, Type, Establishment Year, visit time, entrance fee, weekly off, DSLR policy, *Significance* (Low/Medium/High), Google review rating, and the number of Google reviews (in lakhs). The analysis emphasizes the geographic spread, the distribution of significance levels, and review-driven popularity.

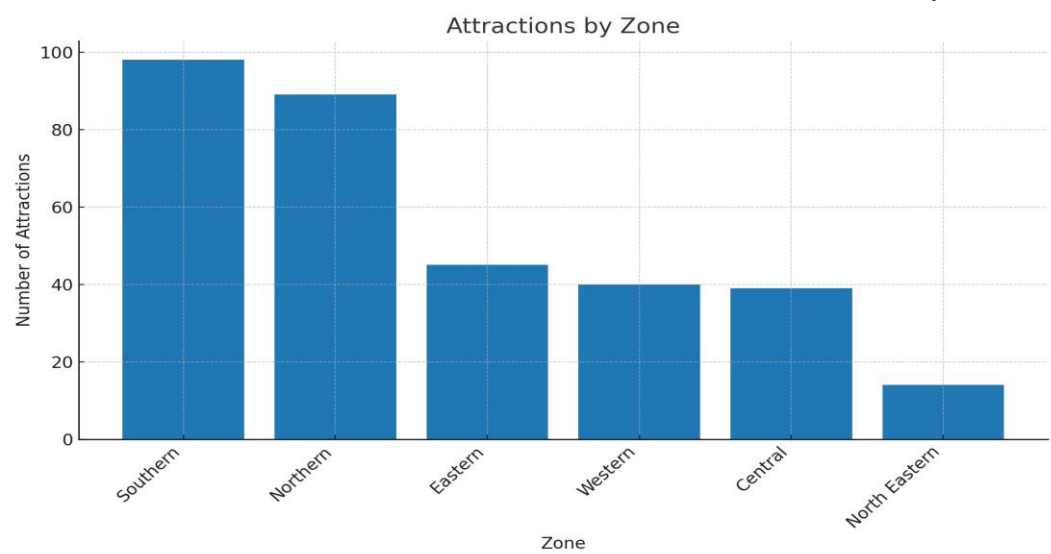
Dataset size: 325 attractions | 33 States | 214 Cities

Methodology. We performed descriptive analytics over geographic variables (Zone/State/City), significance, and Google review metrics (rating and number of reviews). Visualizations summarize distributions and relationships. A pragmatic, rule-based segmentation groups attractions into actionable tiers by combining rating, review volume, and significance. The segments are designed for planning, marketing, and itinerary design.

Metric	Value
Attractions	325
Zones	Central, Eastern, North Eastern, Northern, Southe
States (unique)	33
Cities (unique)	214
Avg Google Rating	4.49
Median Google Rating	4.50
Median Reviews (lakhs)	0.17

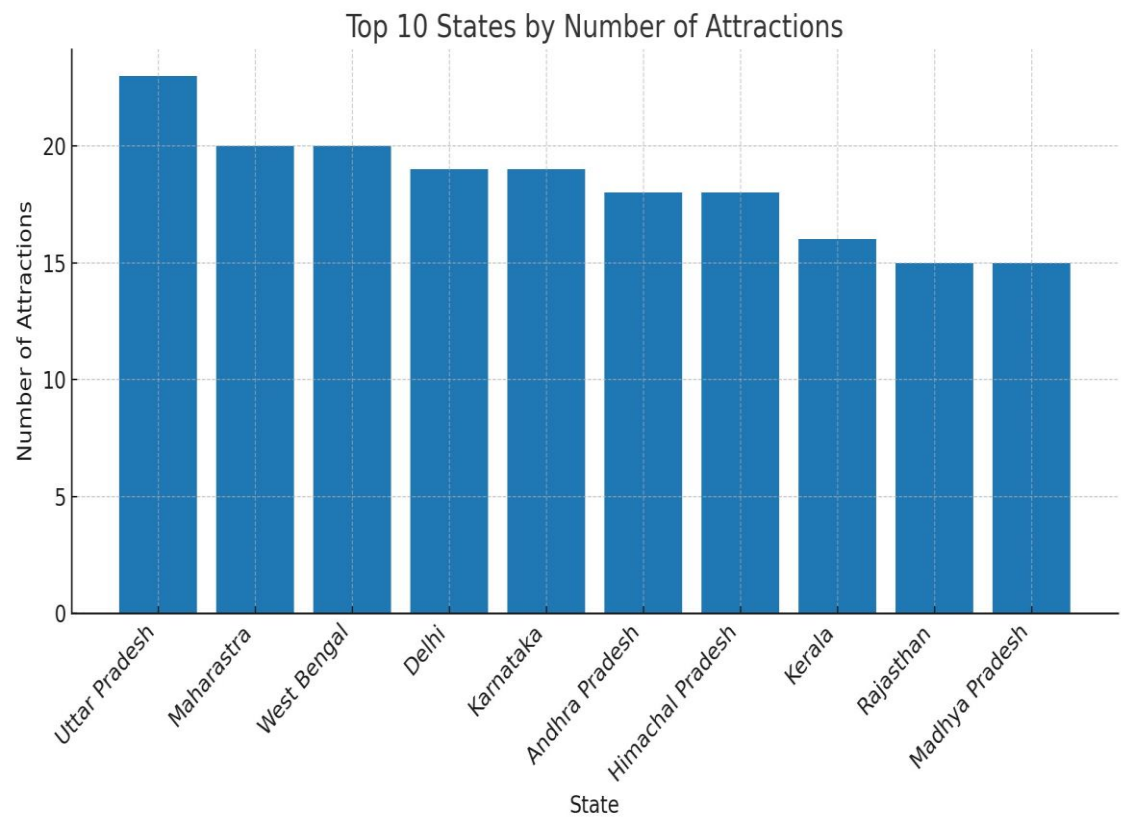
2. Geographic Analysis

Attractions are distributed across zones and states. The bar chart shows the count by Zone.



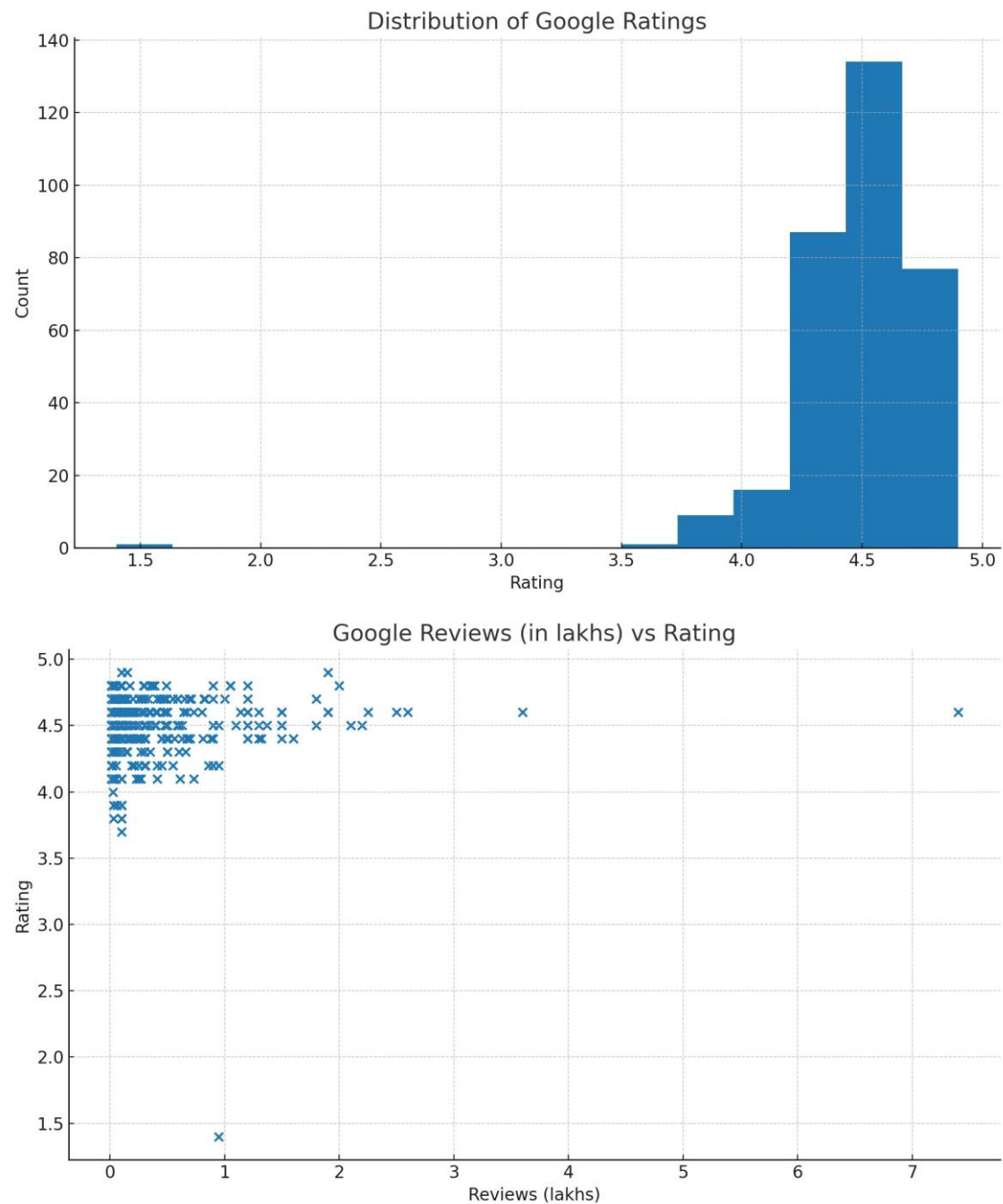
Top 10 states by number of attractions:

State	Attractions
Uttar Pradesh	23
Maharastra	20
West Bengal	20
Delhi	19
Karnataka	19
Andhra Pradesh	18
Himachal Pradesh	18
Kerala	16
Rajasthan	15
Madhya Pradesh	15

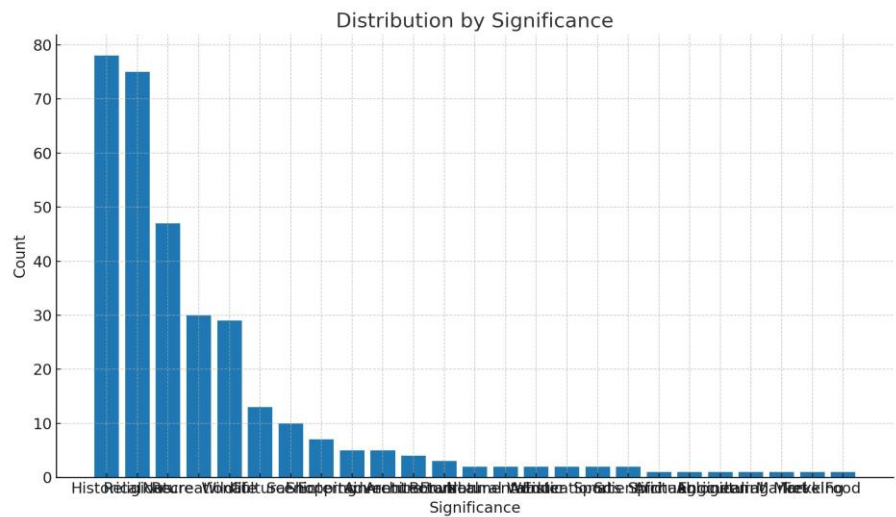


3. Google Review Analysis

The rating distribution skews high, while review volumes vary widely by site.

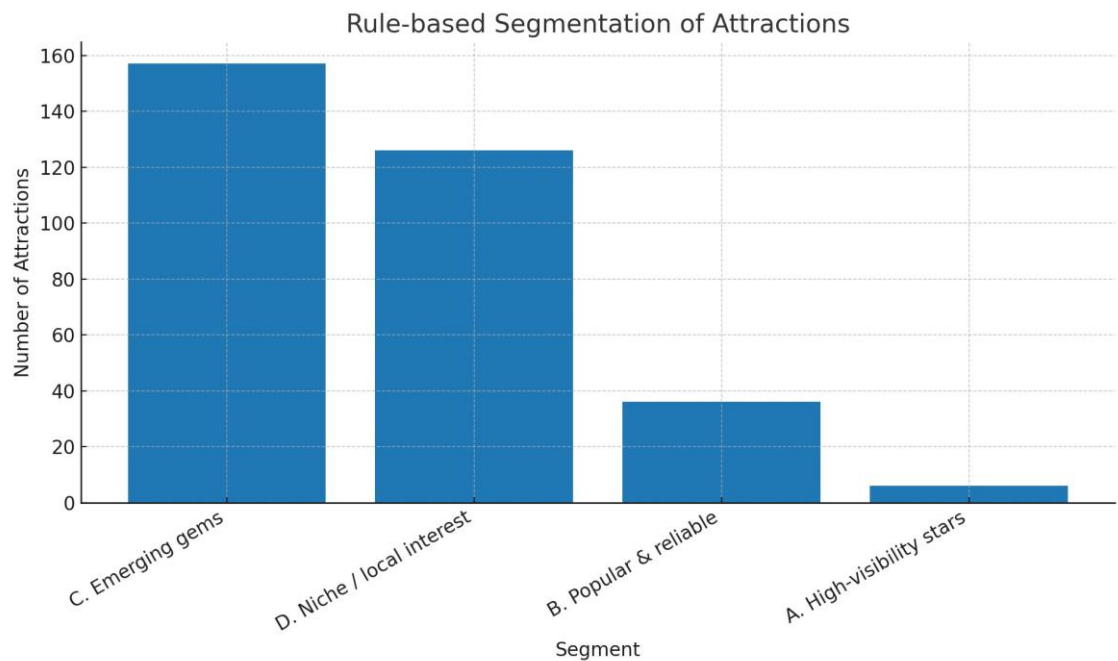


Wildlife	29
Cultural	13
Scenic	10
Shopping	7
Entertainment	5
Adventure	5
Architectural	4
Botanical	3
Environmental	2
Natural Wonder	2
Artistic	2
Educational	2
Sports	2
Scientific	2
Spiritual	1
Archaeological	1
Agricultural	1
Engineering Marvel	1
Market	1
Trekking	1
Food	1



Rule-based Segments

- A. **High-visibility stars**: rating ≥ 4.6 and reviews ≥ 2.0 lakh.
- B. **Popular & reliable**: $4.3 \leq \text{rating} < 4.6$ and 0.5–2.0 lakh reviews.
- C. **Emerging gems**: rating ≥ 4.5 with < 0.5 lakh reviews, or High significance with low reviews.
- D. **Niche / local interest**: remaining sites (use for themed itineraries and nearby add-ons).



5. Insights & Recommendations

- **Portfolio design**: Use Segment A as anchors for itineraries and promotions; pair with Segment C to showcase lesser-known gems.
- **Regional strategy**: Prioritize states with dense attractions for multi-stop circuits; in low-density states, build thematic trails.
- **Experience timing**: Combine average visit time and weekly-offs to optimize crowd management and scheduling.
- **Quality control**: For sites with high reviews but lower ratings, investigate pain points (access, amenities, fees).
- **Content marketing**: For Segment C, invest in storytelling, photography (DSLR policy), and seasonal campaigns (Best Time to visit).
- **Data improvements**: Track monthly review velocity and sentiment to detect rising stars early and adjust segments dynamically.