

# Post-COVID Analysis

.

Baireddy Anjali

Email: [baireddyanjali683@gmail.com](mailto:baireddyanjali683@gmail.com)

GitHub: <https://github.com/anjali aireddy/>

LinkedIn: <https://www.linkedin.com/in/baireddy-anjali-080405314/>

# Introduction

- This presentation provides a detailed analysis of India's post-COVID trends using visitor count data collected from 2018 through Winter season.
- It highlights district-wise performance, seasonal variations, monthly trends, and data-driven recommendations.
- The project aims to assist post-COVID authorities in understanding visitor patterns, optimizing resources, and improving regional post-COVID planning.

# Dataset Description

- Source: Post-COVID\_data.csv
- Records: 8,489 | Columns: 5
- Columns Overview:
  - District: Name of the district in India
  - Month: Month of record
  - Visitors: Number of visitors in that district and month
  - Year: Year of observation
  - Season: Categorized as Summer, Monsoon, Autumn, or Winter

# Data Quality

- No Missing Values: Dataset verified to contain no null or blank entries across all fields.
- Uniform Data Types: All numeric fields (like Visitors) are stored as integers, enabling accurate aggregation and statistical operations.
- Categorical Accuracy: Seasons are mapped correctly to corresponding months (e.g., Summer → April–June).
- Ready for Visualization: The dataset is clean, structured, and directly compatible with Python and PySpark visual analytics tools.
- Validated Data Source: Data aligns with India post-COVID records, ensuring analytical reliability.

# Data Exploration & Operations

- Data Loading & Inspection:
- Imported Post-COVID\_data.csv into a Pandas DataFrame and a PySpark DataFrame for distributed processing.
- Descriptive Statistics:
- Total visitors: 715,410,282
- Average visitors per record: 84,275
- Data Aggregation:
- Grouped data by District to find top-performing post-COVID zones.
- Aggregated by Season to understand seasonal post-COVID peaks.
- Aggregated by Month to observe intra-year fluctuations.
- Visualization:
- Used Matplotlib and PySpark to create bar charts, heatmaps, and line plots for district-wise, seasonal, and monthly visitor patterns.

# Methodology / workflow

## **Data Loading:**

- Imported the raw dataset into Python and PySpark environments for inspection and processing.

## **Data Cleaning & Quality Check:**

- Ensured data consistency, verified schema, and removed any irregularities.

## **Exploratory Data Analysis (EDA):**

- Displayed initial samples of records.
- Explored patterns across districts, months, and seasons.
- Grouped data by key attributes such as district and season.

## **Visualization:**

- Created heatmaps for district-season patterns.
- Developed line and bar graphs to represent monthly and yearly trends.

## **Insight Extraction:**

- Identified top districts and months.
- Compared seasonal performance.

# Key Insights

## ❖ District-Level Trends

- Total visitors recorded: 715,410,282
- Average per observation: 84,275

### Top 5 Districts:

- Mulugu – 77,297,381
- Hyderabad, Ranga Reddy, Medchal & Vikarabad – 61,325,331
- Yadadri Bhongir – 49,289,828
- Rajanna Sircilla – 44,543,745
- Hyderabad – 43,239,548

### Interpretation:

Mulugu emerges as the most visited district, likely due to its natural attractions like waterfalls and temples. Urban centers such as Hyderabad and Ranga Reddy maintain high consistent visitor volumes due to business and leisure tourism.

Visitors Heatmap (Top 20 districts by month)

1e7

5

4

3

2

1

0

Mulugu

Hyderabad, Ranga Reddy, Medchal & Vikarabad

Yadadri Bhongir

Rajanna Sircilla

Hyderabad

Nirmal

Hyderabad, Ranga Reddy, Medchal -Malkajigiri, & Vikarabad

Jagtial

Mahbubnagar

Sangareddy

Bhadradri Kothagudem

Medak

Siddipet

YadadriBhongir

Jayashankar Bhoopalpally

Nagarkurnool

RajannaSircilla

RajanuaSircilla

Rajarina Sircilla

Jogulamba Gadwal

January

February

March

April

May

June

July

August

September

October

November

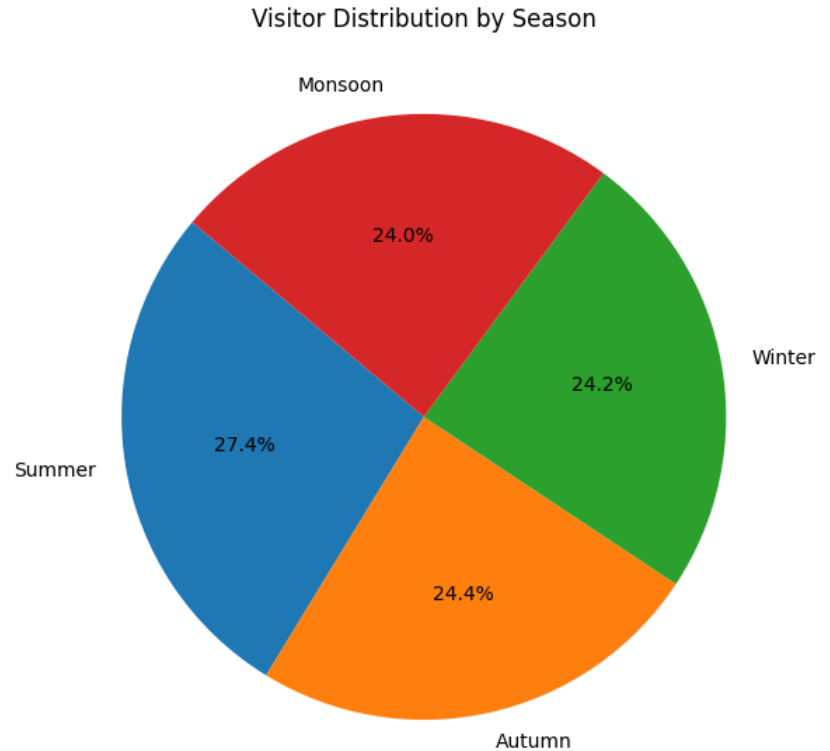
December





## ❖ Seasonal Patterns

- Summer: 196,004,551 visitors (highest)
- Autumn: 174,939,935 visitors
- Winter: 172,453,632 visitors
- Monsoon: 172,012,162 visitors

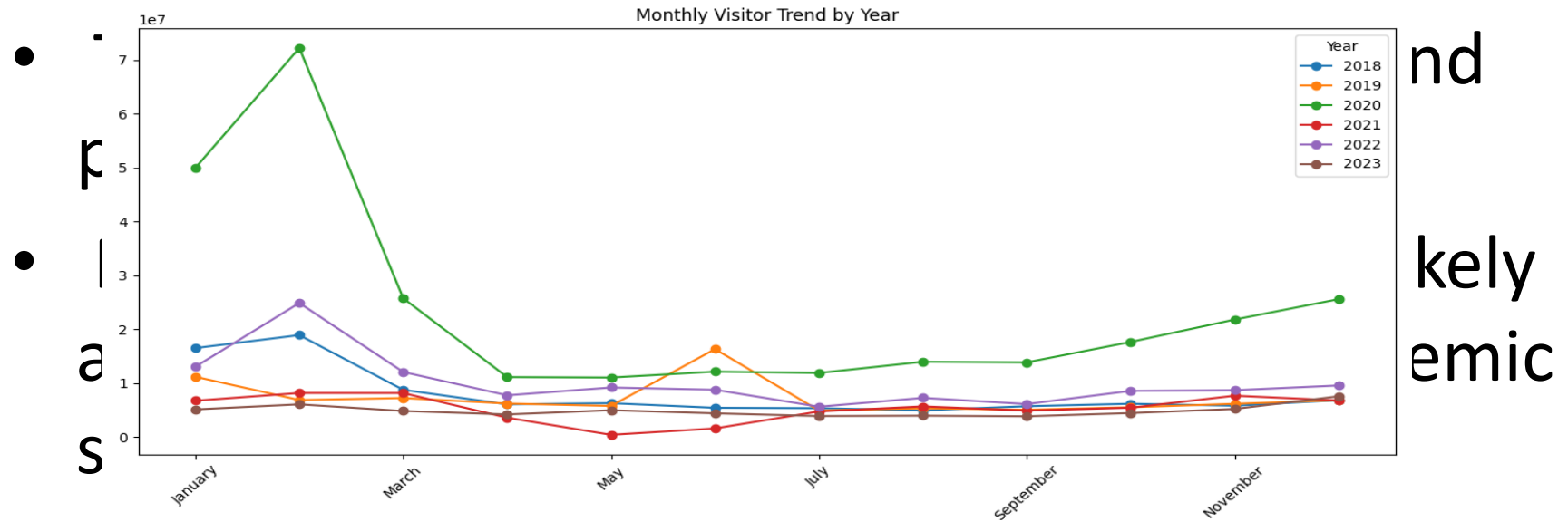


### Interpretation:

The summer season dominates tourism, possibly due to school holidays and temple festival timings. Autumn and winter show balanced inflow, indicating consistent off-peak visitation patterns suitable for year-round tourism planning.

- Monthly Trends

- Top-performing months: February, January, March, December, November



# Key Findings

- India recorded a total of 715 million visitors, with an average of 84,275 visitors.
- Top Districts:
  - Mulugu leads as the most visited district (77 million visitors).
  - Hyderabad, Ranga Reddy, Medchal, and Vikarabad follow with high urban post-COVID inflows.
- Seasonal Insights:
  - Summer had the highest visitor count (196 million).
  - Autumn and Winter followed closely with balanced post-COVID rates. Monsoon saw a slight decline, likely due to weather conditions.
- Monthly Trends:
  - Peak months: February, January, March, December, and November.
  - Indicates strong post-COVID during festive and cooler months.
- Predictable Patterns:
  - The dataset reveals clear seasonal cycles, making it highly suitable for time-series forecasting models.

# Recommendations

- Focus on top-performing districts
- Promote during low-average months
- Create dashboards for monitoring
- Build predictive models for forecasting

# Future Analytics Opportunities

- Develop time-series forecasting models
- Cluster districts by visitation patterns
- Integrate external data (weather, events, transport)

# Conclusion

- The post-COVID data analysis successfully identifies the patterns connecting district, season, and month-wise visitor trends.
- The clean dataset allowed reliable aggregation, supporting effective decision-making.
- This study lays the foundation for predictive analytics, digital dashboards, and strategic post-COVID development across India.

# Future Scope & Recommendations

- Develop machine learning models to predict post-pandemic recovery trends.
- Integrate public health and economic indicators for multi-domain insights.
- Create visual dashboards to monitor ongoing post-COVID developments.
- Encourage policy collaboration based on real-time data analytics.