NAME: SHRAVAN BHAVANDLA

BRANCH: CSE

SECTION: GROUP-06

ROLL NO: 2211CS010071

BIG DATA ANALYTICS
TELANGANA TOURISM DATA ANALYSIS REPORT

1. Description Of Dataset:

- **Reference:-** The dataset utilized in this study is titled Tourism_Data.csv, comprising 10,000 records distributed across five key attributes. It encapsulates comprehensive information on the number of visitors recorded monthly across various districts of Telangana, segmented by month, year, and season.

- **Features:-**

  • **District:** Represents the geographical administrative unit of Telangana.
  • **Month:** Indicates the month of data recording.
  • **Visitors:** Numeric variable showing the total number of visitors for that period.
  • **Year:** Specifies the corresponding year of observation.
  • **Season:** Classifies the record into seasonal categories such as Winter, Summer, Monsoon, and Autumn

- **Data Accuracy:-**
  • The dataset is completely clean and free from missing values.
  • The 'Visitors' field is a numeric type, ensuring accuracy for aggregation and statistical analysis.
  • The data structure supports multi-dimensional analysis across time, location, and season.

2. Tasks Executed

- **Data Loading and Inspection:-** The dataset was loaded into a Pandas DataFrame for exploration. Schema validation was performed to confirm appropriate data types, and the numeric conversion for the Visitors column was verified. Descriptive statistics were computed to understand the distribution and variability of the data.

- **Aggregations and Visualizations:-**
  • Data was grouped by District and Season to highlight the most popular tourism regions.
  • Bar and line charts were generated to visualize visitor trends across districts, months, and seasons.
  • Additional aggregations and sorting techniques were employed to extract top-performing districts and identify months with high tourist activity.

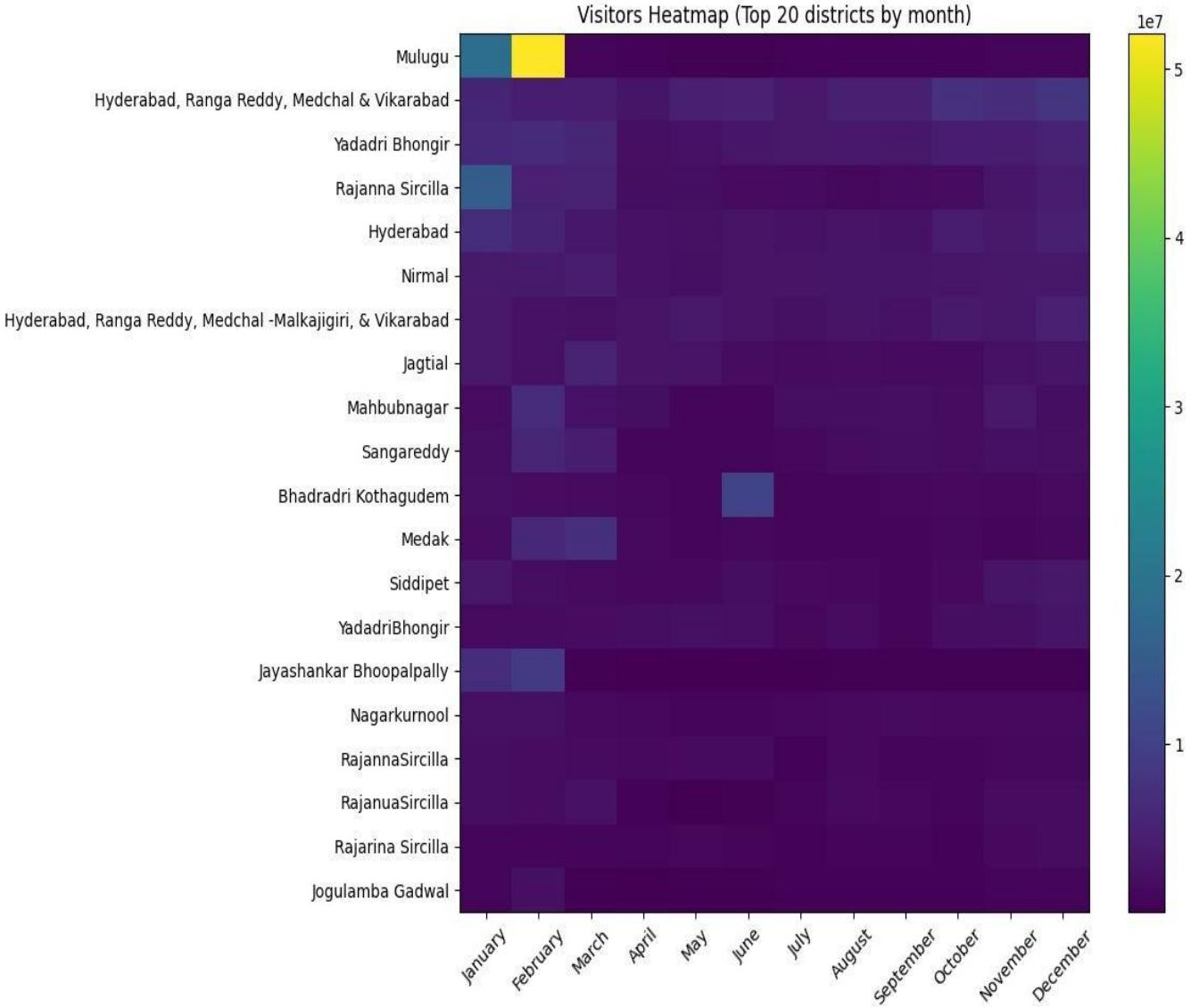
3. Principal Highlights

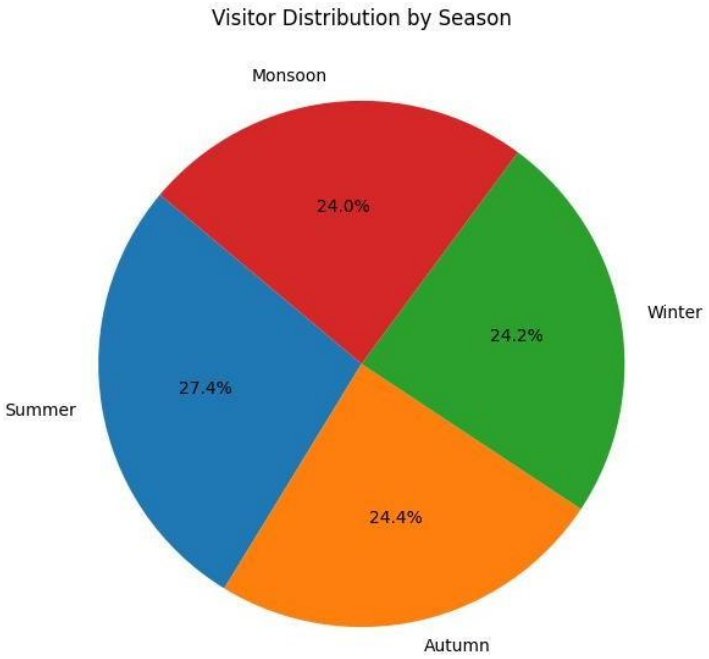
- **Overall Visitor Volume**
  - **Total Visitors:-** 843,986,772
  - **Average Visitors per Record:-** 84,398.68
  - **Minimum:-** 0 **Maximum:-** 9,61,976

- **Top Districts**
  1. **Mulugu:-** 77,297,381
  2. **Hyderabad, Ranga Reddy, Medchal & Vikarabad:-** 61,325,331
  3. **Yadadri Bhongir:-** 49,289,828
  4. **Rajanna Sircilla:-** 44,543,745
  5. **Hyderabad:-** 43,239,548
  6. **Nirmal:-** 38,298,757

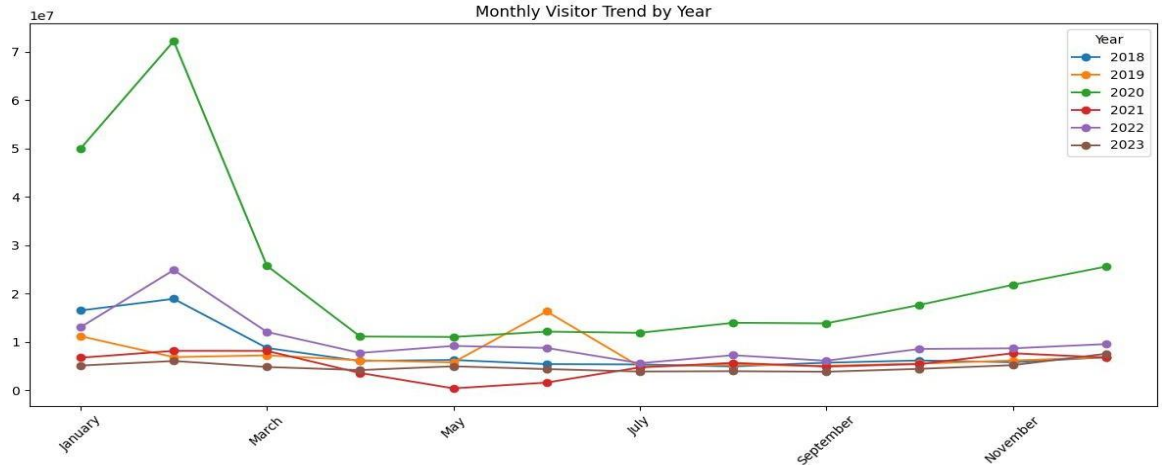
- **Seasonal Trends**

Season	Total Visitors
Summer	196,004,551
Autumn	174,939,935
Winter	172,453,632
Monsoon	172,012,162


Visitors Heatmap (Top 20 districts by month)
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
1e7
5
4
3
2
1



- **Monthly Trends:-** The top five months with the highest average visitor counts were: January, February, March, November and December.



- **Year Coverage:-** The dataset spans 2018 to Winter (latest available season), enabling both year-over-year and seasonal trend comparisons across Telangana's districts.

4. Recommendations

- **District-Level Focus:-**
  - Prioritize tourism marketing campaigns in top-performing districts.
  - Enhance infrastructure, transport, and accommodation in high-traffic zones.
- **Seasonal Readiness:-**
  - Strengthen resource allocation ahead of peak visitor seasons.
  - Introduce dynamic staffing models during high-demand months.

- **Monthly Promotions:-** Conduct targeted promotional campaigns during off-peak months to achieve balanced visitor distribution.
- **Data Monitoring**:- Establish real-time dashboards to continuously track visitor metrics at the district level.

5. Predictive Modelling and Future Opportunities

- **Predictive Analysis:-** Implementation of time-series forecasting or regression models can provide valuable insights for future visitor trends and demand forecasting.
- **Future Analytical Scope:-**
  • Cluster districts based on similar visitation patterns.
  • Integrate external data sources (weather, events, transport) to uncover deeper insights.
  • Develop interactive dashboards to enhance decision-making efficiency.

6. Conclusion:-

The Telangana Tourism Data Analysis reveals significant correlations between district, season, and visitor volume. The dataset's clean and structured nature enabled effective aggregation, visualization, and interpretation. Future work includes the integration of predictive models and interactive dashboards to support data-driven tourism planning and management in Telangana.