



Data Visualization & Preprocessing Report

Student Name:

1. Ganaparthi Nageswari
2. Manasa Chandragiri
3. Kaduru Pavithra
4. Polisetti Obulesu
5. Balija Veerendra

Project Title: UNESCO World Heritage Sites Analysis System

Technology Used: Tableau Public, Python, Flask

Dataset: UNESCO World Heritage Sites (2019)

1. Dataset Description

The dataset contains detailed information about UNESCO World Heritage Sites across different countries and regions.

- ID No
- Name_en
- Region
- Country
- Location
- Date_inscribed
- Danger (In Danger / Not in Danger)
- Endangered Year
- Category (Cultural, Natural, Mixed)
- Criterion 1–10

2. Preprocessing Steps

1. **Data Cleaning** – Checked for missing values, removed duplicates, handled null entries.
2. **Data Formatting** – Standardized date formats for Date_inscribed column.
3. **Data Type Conversion** – Converted categorical fields like Danger and Category into structured format for visualization.
4. **Handling Incomplete Data** – Managed empty Endangered Year values appropriately.
5. **Calculated Fields Creation** – Created fields such as Total Sites, Risk Percentage, and Year-wise Site Count.

6. **Data Filtering Preparation** – Enabled filters for Country, Region, Year, and Category in Tableau.

3. Business Questions with Visualization

- Which countries have the highest number of heritage sites? (Treemap)
- What percentage of sites are listed as “In Danger”? (Pie Chart)
- How have heritage site inscriptions increased over the years? (Line Chart)
- Which region has the most heritage sites? (Bar Chart)
- What is the distribution of Cultural, Natural, and Mixed sites? (Stacked Bar Chart)

Conclusion

The analysis shows that certain countries and regions dominate in heritage site counts. A small but significant percentage of sites are categorized as “In Danger,” highlighting the need for preservation efforts. Regional inscription trends indicate increasing global recognition of heritage sites over time. Visual analytics helps stakeholders make informed decisions for conservation and tourism development.