MINI PROJECT

AIM:

To perform data visualization and analysis on the World Tourism Economy Dataset using Python, Power BI, and Tableau, in order to explore global tourism patterns, understand economic contributions, and uncover insights on tourism's impact on GDP and employment across regions and years.

ALGORITHM:

A. Python Visualization (Matplotlib & Seaborn)

1. Import Libraries
   Import required libraries — pandas, matplotlib.pyplot, and seaborn.
2. Load Dataset
   Load tourism_economy_dataset.csv using pandas.read_csv().
3. Data Preprocessing
   - Check for null or missing values.
   - Convert data types if needed (e.g., Year to integer).
   - Clean or filter data for relevant analysis.
4. Exploratory Data Analysis (EDA)
   - Display info and statistics using .info() and .describe().
   - Identify relationships between tourism receipts, GDP contribution, and employment.
5. Visualization using Matplotlib & Seaborn
   - Line Chart → Tourism Receipts over Years.
   - Bar Chart → Top 10 Countries by Receipts.
   - Scatter Plot → Receipts vs GDP Contribution.
   - Heatmap → Correlation between numerical variables.

Analyze charts to understand tourism growth trends, economic impact, and regional differences.

B. Power BI Visualization (Power BI Online)

Algorithm:

1. **Import Dataset**
   Upload tourism_economy_dataset.csv into Power BI Service (Online Workspace).
2. **Data Preparation**
   - Review column data types (Year, Region, GDP, etc.).
   - Rename fields and format numeric columns for readability.
3. **Report Editor Setup**
   Open **Report Editor View** to create visuals.
4. **Create Visualizations**
   - **Line Chart:** Tourism Receipts vs Year (Region as Legend).
   - **Bar Chart:** Top 10 Countries by Receipts.
   - **Scatter Chart:** GDP Contribution vs Tourism Receipts.
   - **Pie Chart:** Regional share of GDP Contribution.
5. **Publish / Save**
   Save and publish the report in Power BI Online Workspace for cloud access.

C. Tableau Visualization (Tableau Public / Online)

Algorithm:

1. **Load Dataset**
   Open Tableau Public → Upload tourism_economy_dataset.csv.
2. **Data Connection**
   Verify field types (Country, Region = Dimension; GDP, Receipts = Measure).
3. **Create Individual Sheets for Visuals**
   - **Line Chart:** Receipts over Years (by Region).
   - **Bar Chart:** Top 10 Countries by Receipts.
   - **Scatter Plot:** Receipts vs GDP Contribution (Bubble size = Employment).
   - **Donut Chart:** GDP Contribution by Region.
4. **Design Dashboard Layout**
   - Place charts and filters in a balanced layout.
   - Apply consistent color palette by region.
5. **Publish**

   Save and publish to Tableau Public.

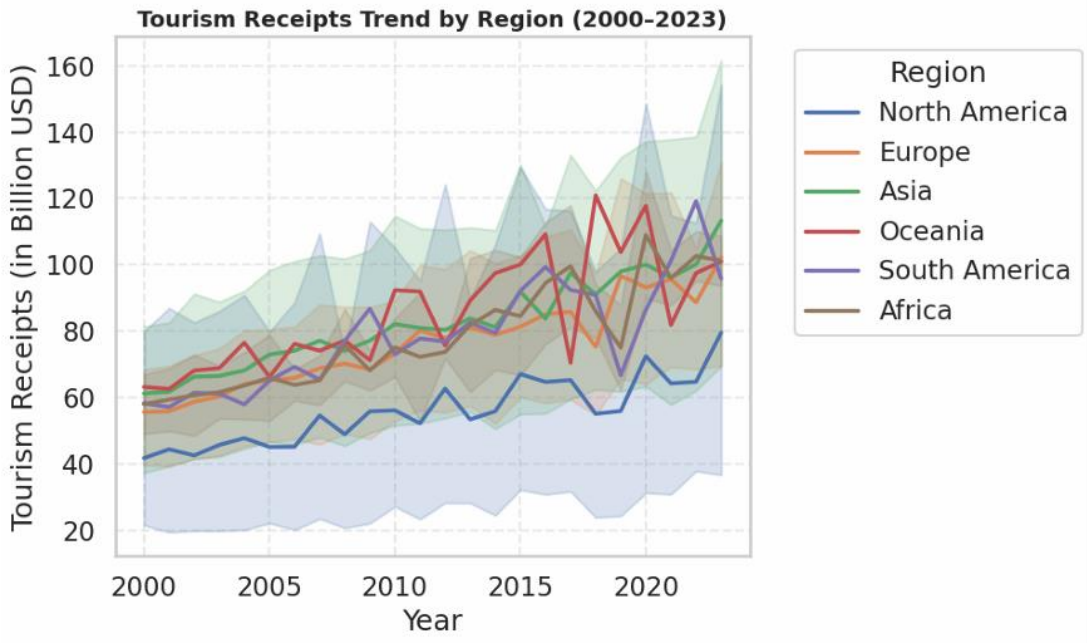
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

sns.set_theme(style="whitegrid", context="talk")

# 2️⃣ Load Dataset
df = pd.read_csv("/content/tourism_economy_dataset.csv")   # Replace with your file
print("✅ Dataset Loaded Successfully!")
```

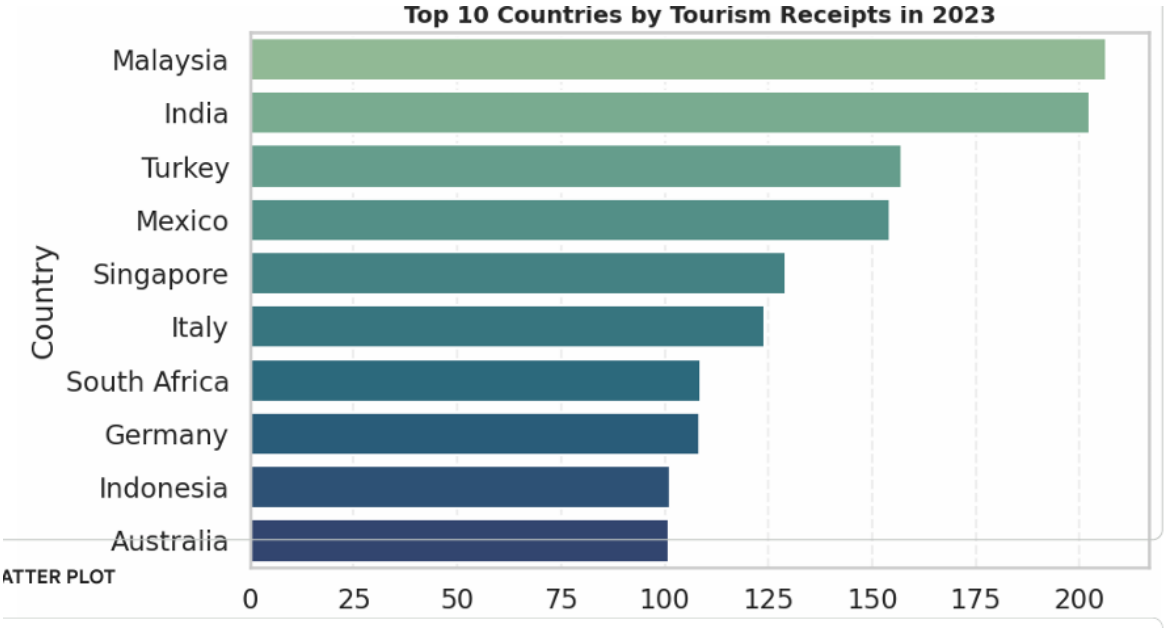
```
✅ Dataset Loaded Successfully!
```

```
plt.figure(figsize=(10,6))
sns.lineplot(data=df, x="Year", y="Tourism_Receipts", hue="Region", linewidth=2.5)
plt.title("Tourism Receipts Trend by Region (2000–2023)", fontsize=14, fontweight="bold")
plt.xlabel("Year")
plt.ylabel("Tourism Receipts (in Billion USD)")
plt.legend(title="Region", bbox_to_anchor=(1.05, 1), loc="upper left")
plt.grid(True, linestyle="--", alpha=0.4)
plt.tight_layout()
plt.show()
```



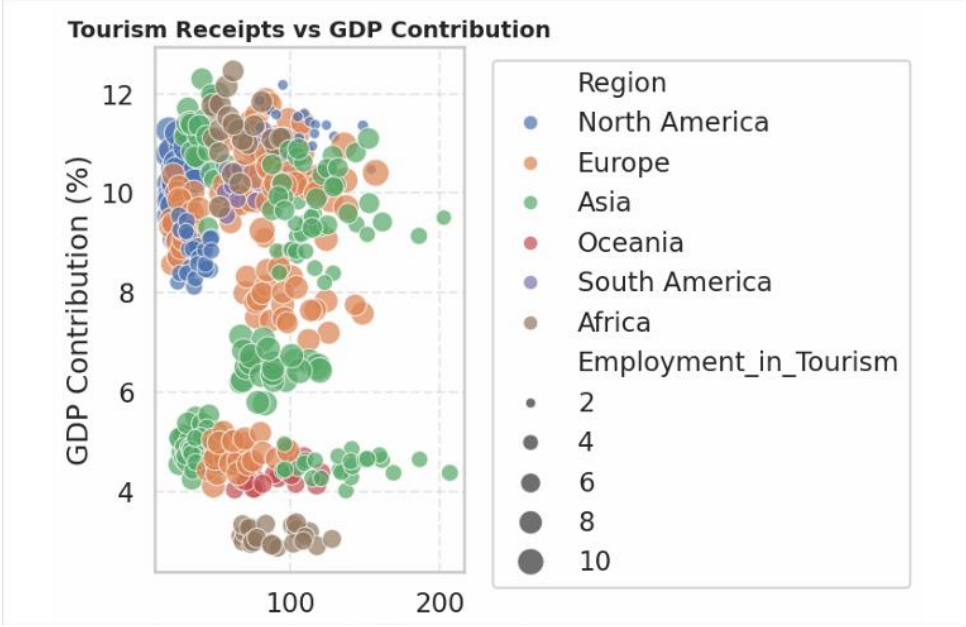
```
latest_year = df['Year'].max()
top10 = df[df['Year'] == latest_year].nlargest(10, 'Tourism_Receipts')

plt.figure(figsize=(10,6))
sns.barplot(data=top10, x="Tourism_Receipts", y="Country", palette="crest")
plt.title(f"Top 10 Countries by Tourism Receipts in {latest_year}", fontsize=14, fontweight="bold")
plt.xlabel("Tourism Receipts (in Billion USD)")
plt.ylabel("Country")
plt.grid(axis='x', linestyle='--', alpha=0.3)
plt.tight_layout()
plt.show()
```



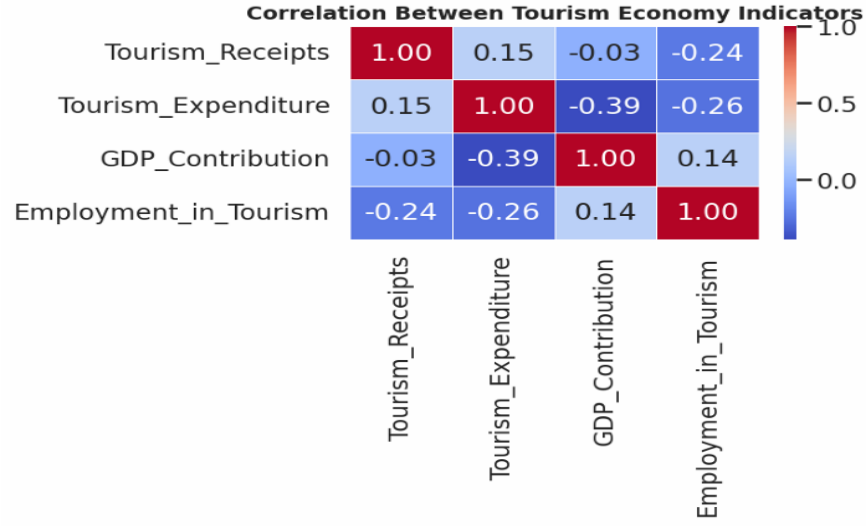
ATTER PLOT

```
plt.figure(figsize=(8,6))
sns.scatterplot(
    data=df,
    x="Tourism_Receipts",
    y="GDP_Contribution",
    hue="Region",
    size="Employment_in_Tourism",
    sizes=(40,300),
    alpha=0.7
)
plt.title("Tourism Receipts vs GDP Contribution", fontsize=14, fontweight="bold")
plt.xlabel("Tourism Receipts (Billion USD)")
plt.ylabel("GDP Contribution (%)")
plt.legend(bbox_to_anchor=(1.05, 1), loc="upper left")
plt.grid(True, linestyle="--", alpha=0.4)
plt.tight_layout()
plt.show()
```



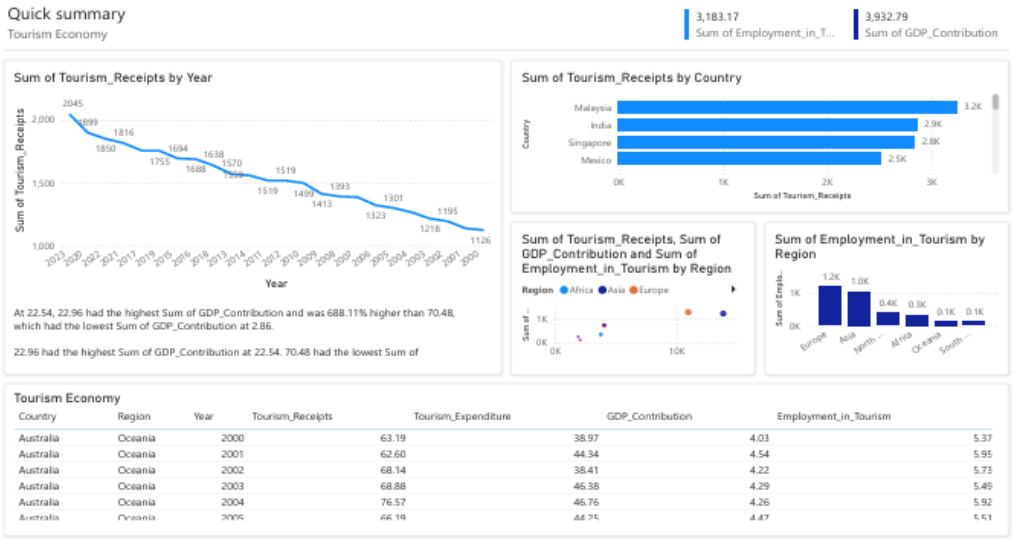
HEATMAP

```
plt.figure(figsize=(8,6))
sns.heatmap(
    df[['Tourism_Receipts','Tourism_Expenditure','GDP_Contribution','Employment_in_Tourism']].corr(),
    annot=True, cmap="coolwarm", fmt=".2f", linewidths=0.5
)
plt.title("Correlation Between Tourism Economy Indicators", fontsize=14, fontweight="bold")
plt.tight_layout()
plt.show()
```

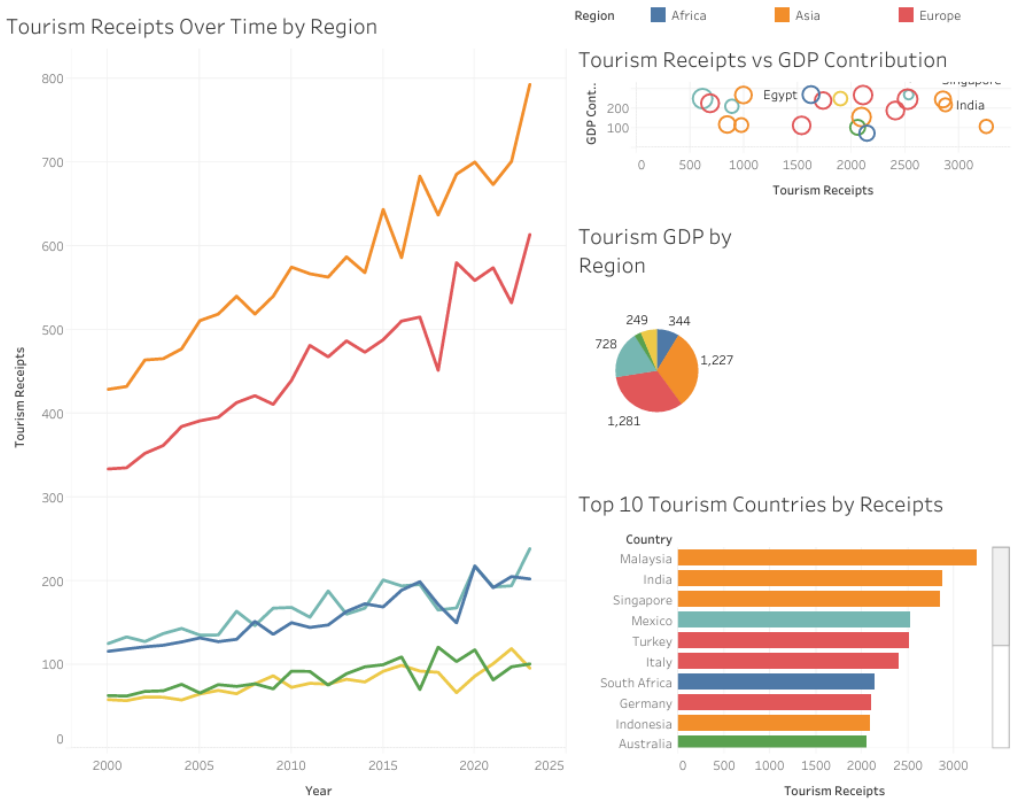


DASHBOARD

POWER BI



TABLEAU



RESULT:

Python, PowerBi ,Tableau visualizations has been executed successfully.