VISITORS TO KENYA

# PROBLEM STATEMENT

Owing to the fact that there are many people who benefit from visitors and visits in the country. It would be prudent to understand the trends related to these visitors coming in and out of the country as it would be beneficial to many.

# OBJECTIVES AND GOALS

The objective of this project is to get a summary of the trend related to the visits in Kenya and have data visualizations of the same.

Some of the questions to be answered are:

1. What is the trend of visitors over the years?
2. Which purpose is most dominantly used by visitors? And what has its trend been?
3. Do the quarters affect the visitors, and if yes how?
4. What is the trend of the quarters over the years?

# PROJECT PLAN

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| PLAN | DURATION | RESOURCES | INPUT | OUTPUT | DEPENDENCIES |
| Problem statement | 15 min | N/A | Local Observations | Problem statement | N/A |
| Objectives and goal setting | 15 min | N/A | Problem statement | Project goals and objectives | Problem statement |
| Data collection | 5 min | N/A | N/A | Working data | Problem statement |
| Data preparation and quality | 15 min | Python libraries | Working data | Data quality report | Collected data |
| Analysis |  | Python libraries and Vscode | Prepared data | Analysis report and data visualization | Tools and libraries used |
| Conclusion | 15 min | Word document | Project summary | Project findings | Project summary |
| Recommendation and next step | 15 min | Word document | conclusions | Recommendations from the project | Analytical findings |

# DATA COLLECTION

The data was collected from a csv file known as visitors\_to\_kenya.csv

# DATA PREPARATION AND QUALITY

The initial steps to preparing the data were to open a jupyter note book and to import python libraries then load the csv file in our working environment.

The data set had 832 entries and 8 columns. The columns of the data set were as follows:

|  |  |  |
| --- | --- | --- |
| COLUMN INDEX | COLUMN NAME | DATA TYPE |
| 0 | Arrival\_or\_Depature | Object |
| 1 | Date\_End\_of\_Quater | Object |
| 2 | Year\_Text | Object |
| 3 | No\_of\_Visitors\_to\_Kenya | Int64 |
| 4 | Visitors\_by\_Purpose | object |
| 5 | Period\_Quater | Object |
| 6 | Result\_Status | object |
| 7 | OBECTID | Int64 |

There were no recorded null values as the information showed 832 non-null values in every column.

# DATA CLEANING

First the columns were renamed for easy manipulation. The new columns were as follows: move, year, num, purpose, Q, status, id. From there I only picked the column that I’d need for the analysis.

The purpose column had long strings that would make manipulation more tedious than if they are one word. The values in this column were changed as follows; ‘Visitors on Holiday’ to ‘holiday’, ‘Visitors on Business’ to ‘business’ and ‘Visitors in Transit’ to ‘transit’.

# ANALYSIS

# CONCLUSION

# RECCOMMENDATION AND NEXT STEP