Report on Tourist Attractions Dataset

Student name: Talamanchi shrijith

Student I'd: 23017322

Git hub link: https://github.com/shrijith2002/shrijith

Introduction:

The tourism sector is a vital contributor to regional economies, drawing visitors and supporting local businesses. Analyzing data on tourist attractions, including visit duration, Google review ratings, entrance fees, and other details, is crucial for strategic planning and development. This report delves into such data to glean insights into the characteristics and factors that shape tourist attractions.

Overview:

The dataset contains information on 325 tourist attractions spanning various zones, states, and cities. It encompasses details like establishment year, visit duration, Google review ratings, entrance fees, and more. Our objective through exploratory data analysis is to reveal insights into these attractions' attributes and their interrelations.

Statistical Summary:

A statistical analysis of the data reveals the following insights:

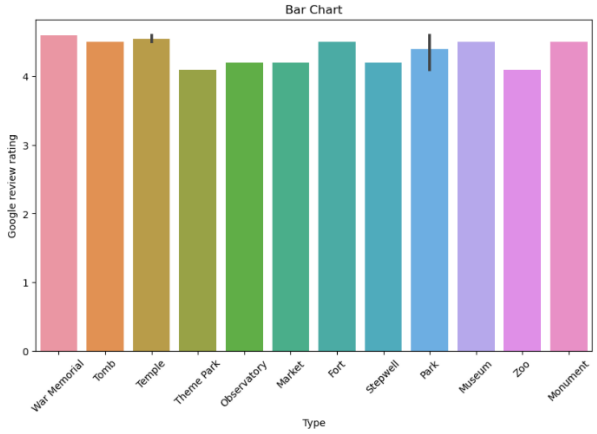
Time needed to visit in hrs: The mean time required for a visit is 1.81 hours, with a standard deviation of 0.97 hours, indicating moderate variability. The duration ranges from a minimum of 0.5 hours to a maximum of 7 hours, with the 25th, 50th, and 75th percentiles falling at 1 hour, 1.5 hours, and 2 hours respectively.

Google review ratings exhibit a mean of 4.49 with a standard deviation of 0.27, showcasing generally high ratings. Ratings range from a minimum of 1.4 to a maximum of 4.9, with the median rating at 4.5 and the 25th and 75th percentiles at 4.4 and 4.6 respectively. Lastly, the number of Google reviews has a mean of 0.41 lakhs and a standard deviation of 0.65 lakhs, indicating a substantial range in review counts. Reviews span from a minimum of 0.01 lakhs to a maximum of 7.4 lakhs, with the median review count at 0.17 lakhs and the 25th and 75th percentiles at 0.059 lakhs and 0.5 lakhs respectively.

Exploratory Data Analysis Report:

Bar Chart: Type vs. Google Review Rating

This report seeks to examine how different types of tourist attractions correlate with their respective Google review ratings. The aim is to ascertain whether there are disparities in visitor satisfaction levels among various types of attractions. By utilizing a bar chart, we visualize the average Google review ratings across different types of tourist attractions. Through this analysis, we gain valuable insights into the levels of visitor satisfaction across diverse types of tourist attractions.



Observations:

Temple: Temples have the highest average Google review rating among all types of attractions, indicating a high level of visitor satisfaction with temple visits.

Observatory: Attractions categorized as observatories also exhibit a high average rating, suggesting positive visitor experiences.

Theme Park: Theme parks receive relatively high ratings, indicating satisfactory experiences for visitors.

Tomb: Tombs also have a decent average rating, reflecting positive visitor feedback.

Park: Parks are generally well received, with a moderate to high average rating.

Museum: Museums receive favourable ratings, indicating positive visitor experiences with cultural and historical attractions.

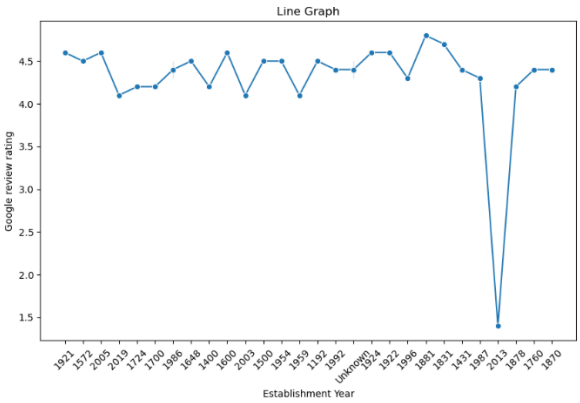
Monument: Monuments have a moderate average rating, suggesting varying levels of satisfaction among visitors.

Market, Fort, Stepwell, Zoo: These attractions exhibit lower average ratings compared to others, indicating potential areas for improvement in visitor experiences or attraction management.

Line Graph: Establishment Year vs. Number of Google Reviews

The goal of this report is to examine how the establishment year of tourist attractions relates to the number of Google reviews they receive, aiming to identify any trends or patterns in attraction popularity over time. Through a line graph, we depict the trend in the number of Google reviews over the years since

the establishment of the attractions. This analysis provides valuable insights into the popularity and engagement levels of tourist attractions over time, shedding light on their long-term appeal to visitors.



Observations:

Overall Trend: The line graph shows fluctuations in the number of Google reviews over the years, indicating varying levels of popularity for attractions established at different times.

Early Years: Attractions established in the early years may exhibit a gradual increase in the number of Google reviews over time, reflecting sustained interest and visitor engagement.

Recent Years: Attractions established in more recent years may show a steep increase or fluctuating trend in the number of Google reviews, indicating changing visitor preferences or marketing efforts.

Outliers: Certain attractions may stand out as outliers, experiencing exceptionally high or low numbers of Google reviews compared to others established in the same period.

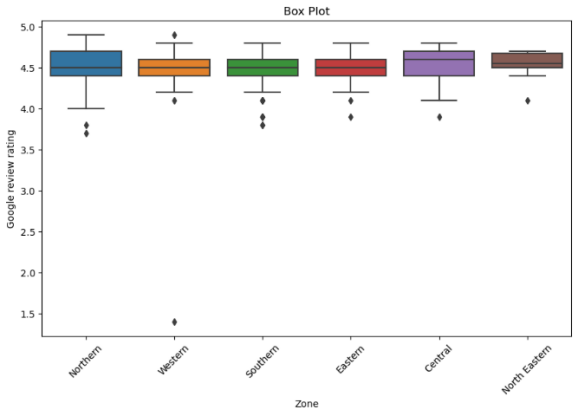
Box plot: Zone vs. Google Review Rating

The objective of this report is to examine the connection between various zones and the corresponding Google review ratings of tourist attractions, with the goal of determining if there are discrepancies in visitor satisfaction levels across different zones. Using a box plot, we represent the distribution of Google review ratings across different zones.

Observations:

Northern Zone: Attractions in the Northern zone tend to have relatively high Google review ratings, indicating high visitor satisfaction levels.

Southern Zone: The Southern zone also exhibits favourable Google review ratings, suggesting positive visitor experiences.



Eastern Zone: Attractions in the Eastern zone show moderate to high average ratings, reflecting satisfactory visitor experiences.

Western Zone: The Western zone demonstrates varying levels of Google review ratings, with some attractions receiving high ratings while others have lower ratings.

Central Zone: The Central zone exhibits a wide range of Google review ratings, possibly indicating diverse visitor experiences across attractions in this zone.

Correlation Matrix:

A correlation matrix reveals the relationships between different variables in the dataset. For example, we can examine if there's any correlation between visit duration, Google review ratings, entrance fees, and the number of Google reviews.

Insights and Observations:

The Google review ratings generally reflect high levels of visitor satisfaction across the attractions.

There is a wide variation in entrance fees, implying diverse pricing strategies among different attractions.

The visit durations at attractions vary considerably, indicating a range of visitor experiences and levels of engagement Certain correlations may exist between attributes; for example, attractions with longer visit durations may tend to have higher entrance fees.

Conclusion:

Analyzing the dataset on tourist attractions provides valuable insights into their features and influencing factors. By understanding visit durations, Google review ratings, entrance fees, and other attributes, tourism officials and businesses can make informed decisions, refine marketing strategies, and enhance visitor experiences. Further exploration and analysis could offer deeper insights into the dynamics of tourist attractions and their contributions to local economies and tourism industries.