Table of CONTENTS

06 Project Overview

KeyUserAttributes:Review,Customer_name,Rating s,Review title

O7 Libraries and Data Handling

Libraries used : Pandas, Matplotlib, Seaborn. **Data Loading and preprocessing :** Loading from CSV, data cleaning, handling dates and categorical data.

03 Data AnalysisTechnique

Descriptive statistics: Mean, median, count, standard deviation. **Visualization methods**: Bar charts, pie charts, heatmaps, count and distribution plots.

04 Key Findings

05

Rating Distribution: Most reviews are positively skewed with higher ratings.. Common Themes: Top positive words include "excellent," "friendly," and "clean," while negative words include "poor," "dirty," and "bad." Model Performance: The logistic regression model effectively predicts sentiment with an accuracy score indicating good performance..

Advance Analysis

Geographical Insights: Mapping reviews by location to identify regional satisfaction trends. **Temporal Trends**: Analyzing how review sentiments change over time.

Table of CONTENTS

Visual Insights

Gender distribution : Count plots by country. Device preference by country. **Subscription type popularity :** Visualization of plan popularity.

Conclusion

The insights derived from this analysis can help hotel management understand customer satisfaction drivers and areas needing improvement. Data-driven decision-making enables targeted strategies to enhance customer experiences and increase positive reviews. Future analysis with more comprehensive datasets can provide deeper insights into user behaviors and preferences.

Appendix

Code Snippets : Provided Python code used for loading, cleaning, transforming data, and generating visualizatins.

Datasets: Sample dataset of Hotel Reviews Sentimental Analysis

Additional References: Referenced any external datasets or tools used during the analysis process.

Github Website Link: https://nineswords.github.io/csel302/