INDIAN RESTAURANTS 2023 COMPREHENSIVE EDA - PYTHON

By: Zahida Kauser

PROJECT BACKGROUND:

Dineout is a table booking platform helping customers to do table booking in their favorite restaurants for free and help them get great discounts.

The website is well known for displaying user ratings for restaurants, hotels, b&b, touristic attractions, and other places, with a total word count of all reviews is more than 10 million.

The project surveys and analyzes Dineout restaurants in Indian market. To determine regional performance and customer behaviour; region-based analysis is performed.

PROJECT GOALS:

This project surveys and analyzes Dineout restaurants in Indian market. To determine regional performance and customer behaviour; region-based analysis is performed.

To provide further information in regard to the restaurants details that make them successful and appreciated by the users, with the possibility to compare the common features of different European countries regarding the average ratings, awards, open hours, reviews count, etc.

1. Analyzing Dataframe
2. How are restaurants distributed across India?
3. How are average ratings distributed across India?
4. How is cost distributed across India?
5. How are votes distributed across India?
6. How is the overall performance of restaurants across different states?
7. What are top cuisines in India?
8. How are the cuisines distributed among states?
9. What are top restaurant locations in Maharashtra, Delhi and Karnataka?
10. References and etc.

ABOUT THE DATASET:

Data has been retrieved from the publicly available website https://www.dineout.co.in/.

All the restaurants from the main Indian cities have been scraped in early March 2023.

The Dineout dataset includes thousands of restaurants with attributes such as location data, average rating, number of reviews, cuisine types, etc. The dataset combines the restaurants from the main Indian cities.

IMPLEMENTATION:

Phase I: Data Processing:

strong positive correlation with our label or gives strong negative correlation. From pandas library use scatter_matrix for attribute combination.

Phase II: To fill missing attributes:

There are three ways to set a missing values in data as:

1) get rid of the messing data point.

2) Get rid of the whole attribute.

3) set the value to some value (0, mean or median). Here, can't use the first option because we cannot drop the data point from the data. Option second is not valid. We have to use option no three forset missing attributes.

Phase III: Python Packages:

import pandas as
pd import
numpy as np

import plotly.figure_factory
as ff from plotly.offline import
iplot

import plotly.express as px
from plotly.subplots import make_subplots

import plotly.graph_objects as go import
warnings warnings.filterwarnings('ignore')

REFERENCES:

1. Salim Lahmiri and Stelios Bekiros. Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos, Solitons and Fractals, 118:35–40, jan 2019. ISSN 09600779. doi:10.1016/j.chaos.2018.11.014.

2. Suhwan Ji, Jongmin Kim, and Hyeonseung Im. A comparative study of bitcoin price prediction using deep learning. Mathematics, 2019. ISSN 22277390. doi: 10.3390/math7100898.

3. Byeonghwa Park and Jae Kwon Bae. Using machine learning algorithms for housing price prediction: The case of fairfax county, virginia housing data. Expert Systems with Applications, 42(6):2928–2934,2015.

4. K. T. Chandrashekhara, M. Thungamani, C. N. Gireesh Babu, and T. N. Manjunath. Smartphone price prediction in retail industry using machine learning techniques. In Lecture Notes in Electrical Engineering, 2019. ISBN 9789811358012. doi: 10.1007/978-981-13-5802-9 34.

5. Muhammad Asim and Zafar Khan. Mobile price class prediction using machine learning techniques. In- ternational Journal of Computer Applications, 179(29):6–11, 2018. doi: 10.5120/ijca2018916555.

6. Meysam Asgari-Chenaghlu, M. Reza Feizi-Derakhshi, Leili Farzinvash, M. A. Balafar, and Cina Motamed. A multimodal deep learning approach for named entity recognition from social media,2020.

7. N. Nikzad-Khasmakhi, M. A. Balafar, M. Reza Feizi-Derakhshi, and Cina Motamed. Berters: Multimodal representation learning for expert recommendation system with transformer, 2020.

8. Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. Deep learning-based text classification: A comprehensive review, 2020.

****