innomatics-research-lab

April 19, 2025

[ ]:
```
!pip install pandas numpy matplotlib seaborn geopy folium
```

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from geopy.geocoders import Nominatim
from geopy.exc import GeocoderTimedOut
from IPython.display import display, HTML
import time
import folium
```

[2]:
```
data1 = pd.read_csv("C:\\Users\\venka\\Downloads\\zomato_data.csv")

data2 = pd.read_csv("C:\\Users\\venka\\Downloads\\Geographical Coordinates.csv")

data1.head()
data2.head()
```

[2]:

	listed_incity	Latitude	Longitude
0	Banashankari	12.939333	77.553982
1	Bannerghatta Road	12.952660	77.605048
2	Basavanagudi	12.941726	77.575502
3	Bellandur	12.925352	77.675941
4	Brigade Road	12.967358	77.606435

[66]:
```
data1['rate'] = data1['rate'].replace(['-', 'NEW'], np.nan)
data1['rate'] = data1['rate'].astype(str).str.replace('/5', '').str.strip()
data1['rate'] = pd.to_numeric(data1['rate'], errors='coerce')
data1['rate'] = data1['rate'].fillna(data1['rate'].median())
```

[67]:
```
data1['approx_costfor_two_people'] = data1['approx_costfor_two_people'].
 ↪astype(str).str.replace(',', '')
data1['approx_costfor_two_people'] = pd.
 ↪to_numeric(data1['approx_costfor_two_people'], errors='coerce')
data1['approx_costfor_two_people'] = data1['approx_costfor_two_people'].
 ↪fillna(data1['approx_costfor_two_people'].median())
```

[68]:
```
data1['dish_liked'] = data1['dish_liked'].fillna('Not Available')
data1['cuisines'] = data1['cuisines'].fillna('Other')
data1['rest_type'] = data1['rest_type'].fillna('Unknown')
```

[69]:
```
data1['votes'] = data1['votes'].fillna(data1['votes'].median())
```

[70]:
```
data1['online_order'] = data1['online_order'].map({'Yes': 1, 'No': 0})
data1['book_table'] = data1['book_table'].map({'Yes': 1, 'No': 0})
```

[71]:
```
data1['votes'] = data1['votes'].astype(int)
data1['approx_costfor_two_people'] = data1['approx_costfor_two_people'].
 ↪astype(int)
```

[72]:
```
merged_df = pd.merge(data1, data2, on='listed_incity', how='left')
merged_df.head()
```

[72]:

	online_order	book_table	rate	votes	rest_type	\
0	1	1	4.1	775	Casual Dining	
1	1	0	4.1	787	Casual Dining	
2	1	0	3.8	918	Cafe, Casual Dining	
3	0	0	3.7	88	Quick Bites	
4	0	0	3.8	166	Casual Dining	

	dish_liked	\
0	Pasta, Lunch Buffet, Masala Papad, Paneer Laja…	
1	Momos, Lunch Buffet, Chocolate Nirvana, Thai G…	
2	Churros, Cannelloni, Minestrone Soup, Hot Choc…	
3	Masala Dosa	
4	Panipuri, Gol Gappe	

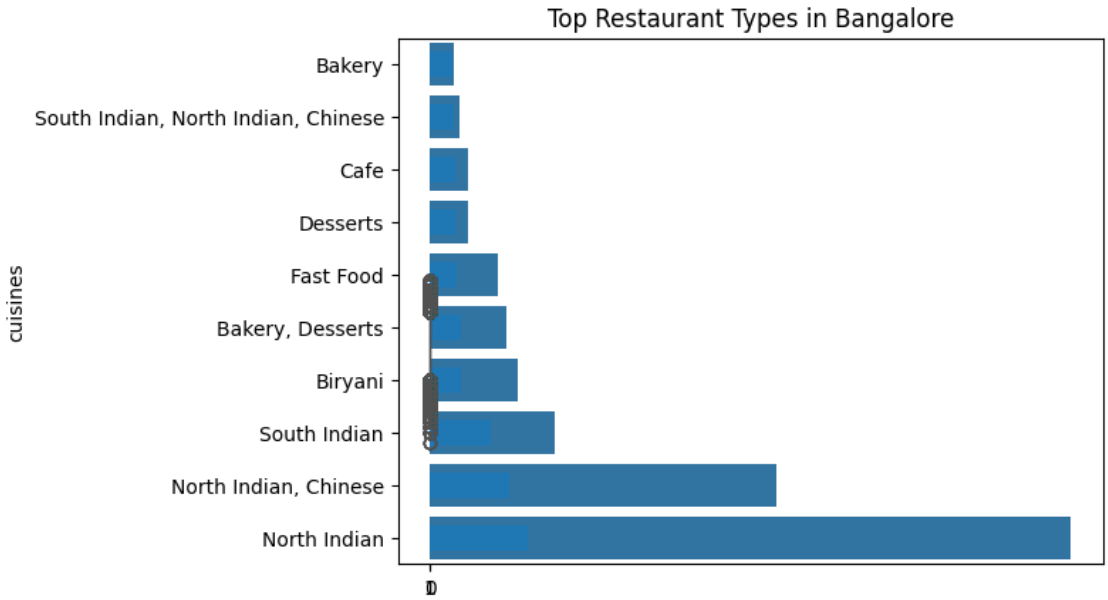
	cuisines	approx_costfor_two_people	listed_intype	\
0	North Indian, Mughlai, Chinese	800	Buffet	
1	Chinese, North Indian, Thai	800	Buffet	
2	Cafe, Mexican, Italian	800	Buffet	
3	South Indian, North Indian	300	Buffet	
4	North Indian, Rajasthani	600	Buffet	

	listed_incity	Latitude	Longitude
0	Banashankari	12.939333	77.553982
1	Banashankari	12.939333	77.553982
2	Banashankari	12.939333	77.553982
3	Banashankari	12.939333	77.553982
4	Banashankari	12.939333	77.553982

[73]:
```
sns.countplot(data=merged_df, y='rest_type', order=merged_df['rest_type'].
 ↪value_counts().iloc[:10].index)
plt.title("Top Restaurant Types in Bangalore")
```

```
sns.boxplot(x='online_order', y='rate', data=merged_df)
merged_df['cuisines'].value_counts().head(10).plot(kind='barh')
```

[73]: <Axes: title={'center': 'Top Restaurant Types in Bangalore'}, ylabel='cuisines'>



[74]:
```
merged_df = merged_df.dropna(subset=['Latitude', 'Longitude'])

bangalore_map = folium.Map(location=[12.9716, 77.5946], zoom_start=11)

for i, row in merged_df.iterrows():
    folium.CircleMarker(
        location=[row['Latitude'], row['Longitude']],
        radius=1,
        color='blue',
        fill=True,
        fill_color='blue'
    ).add_to(bangalore_map)

bangalore_map
```

[74]: <folium.folium.Map at 0x22745d44210>

[79]:
```
italian_map = folium.Map(location=[12.9716, 77.5946], zoom_start=11)
for i, row in italian_df.iterrows():
    folium.Marker(
        location=[row['Latitude'], row['Longitude']],
        popup=row['rest_type']
```

```
    ).add_to(italian_map)

italian_map
```

[79]:
```
<folium.folium.Map at 0x22759915710>
```

[ ]:
```

```