

```
import os
import pandas as pd
os.chdir('./My Documents/MIT')
df = pd.read_csv('Fast_Food_Restaurants_US.csv')

df.shape

(10000, 11)

df.index

RangeIndex(start=0, stop=10000, step=1)

df.columns

Index(['Unnamed: 0', 'address', 'categories', 'city', 'country',
      'latitude',
      'longitude', 'name', 'postalCode', 'province', 'websites'],
      dtype='object')

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   Unnamed: 0            10000 non-null  int64   
1   address               10000 non-null  object  
2   categories            10000 non-null  object  
3   city                  10000 non-null  object  
4   country               10000 non-null  object  
5   latitude              10000 non-null  float64 
6   longitude             10000 non-null  float64 
7   name                  10000 non-null  object  
8   postalCode            10000 non-null  object  
9   province              10000 non-null  object  
10  websites              10000 non-null  object  
dtypes: float64(2), int64(1), object(8)
memory usage: 859.5+ KB

df.head()
```

| | Unnamed: 0 | address | categories | city | country |
|---|------------|----------------------|------------|------|---------|
| 0 | 0 | 800 N Canal Blvd | | | |
| 1 | 1 | 800 N Canal Blvd | | | |
| 2 | 2 | 206 Wears Valley Rd | | | |
| 3 | 3 | 3652 Parkway | | | |
| 4 | 4 | 2118 Mt Zion Parkway | | | |

| | | | |
|---|----------------------------------------------|--------------|----|
| 0 | American Restaurant and Fast Food Restaurant | Thibodaux | US |
| 1 | Fast Food Restaurants | Thibodaux | US |
| 2 | Fast Food Restaurant | Pigeon Forge | US |
| 3 | Fast Food | Pigeon Forge | US |
| 4 | Fast Food Restaurant | Morrow | US |

| | latitude | longitude | name | postalCode | province | \ |
|---|-----------|------------|----------------|------------|----------|---|
| 0 | 29.814697 | -90.814742 | SONIC Drive In | 70301 | LA | |
| 1 | 29.814697 | -90.814742 | SONIC Drive In | 70301 | LA | |
| 2 | 35.803788 | -83.580553 | Taco Bell | 37863 | TN | |
| 3 | 35.782339 | -83.551408 | Arby's | 37863 | TN | |
| 4 | 33.562738 | -84.321143 | Steak 'n Shake | 30260 | GA | |

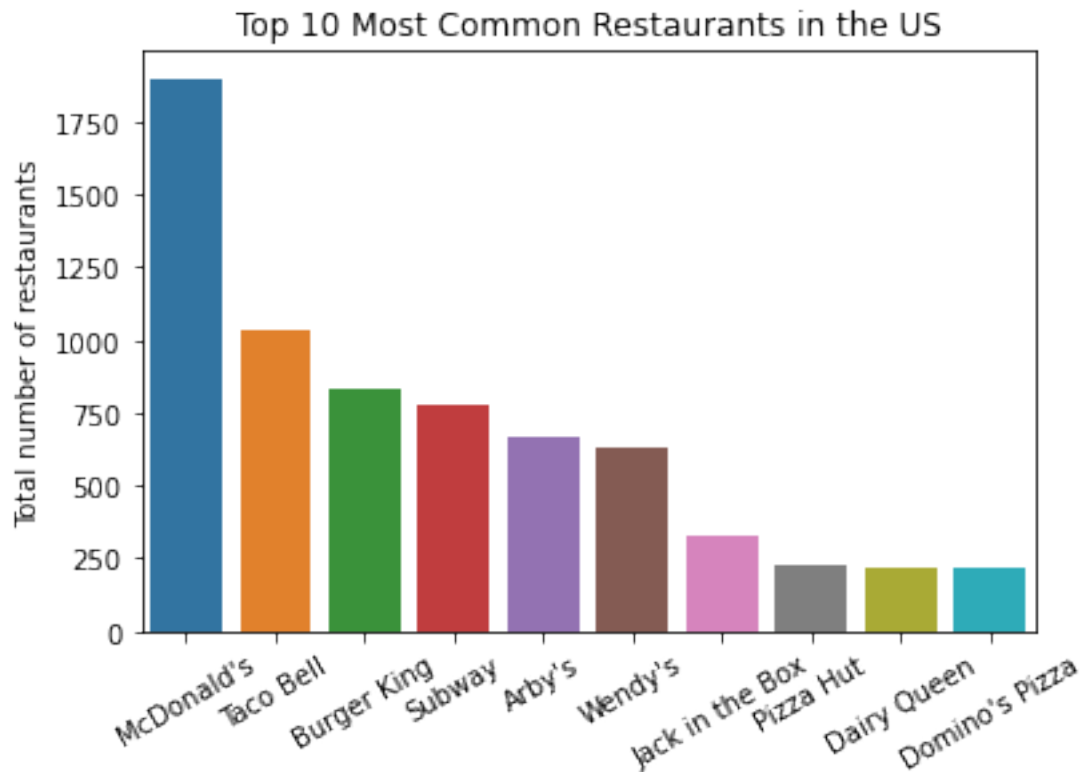
| | websites |
|---|-----------------------------------------------------------------------------------------------------------------------------------|
| 0 | https://locations.sonicdrivein.com/la/thibodau... |
| 1 | https://locations.sonicdrivein.com/la/thibodau... |
| 2 | http://www.tacobell.com , https://locations.taco... |
| 3 | http://www.arbys.com , https://locations.arbys.c... |
| 4 | http://www.steaknshake.com/locations/23851-ste... |

```
data2 = df.groupby('name')['name'].count()
```

```
top10restaurants = data2.sort_values(ascending=False).head(10)
top10restaurants
```

```
name
McDonald's      1898
Taco Bell       1032
Burger King      833
Subway           776
Arby's           663
Wendy's          628
Jack in the Box  330
Pizza Hut        230
Dairy Queen      218
Domino's Pizza   215
Name: name, dtype: int64
```

```
import matplotlib.pyplot as plt
import seaborn as sns
g = sns.barplot(y=top10restaurants.values, x=top10restaurants.index)
g.set_xticklabels(labels=top10restaurants.index, rotation=30)
g.set(title='Top 10 Most Common Restaurants in the US', ylabel="Total
number of restaurants", xlabel="")
plt.show()
```



```
data3 = df.groupby('city')['city'].count()
top10cities = data3.sort_values(ascending=False).head(10)
top10cities
```

```
city
Houston      107
Las Vegas     82
Phoenix       78
Columbus      72
Dallas        66
Orlando       65
Los Angeles   64
Miami         62
Chicago       57
San Antonio   54
Name: city, dtype: int64
```

```
g2 = sns.barplot(y=top10cities.values, x=top10cities.index)
g2.set_xticklabels(labels=top10cities.index, rotation=30)
g2.set(title='Top 10 Cities for Fast Food Accessibility',
        ylabel='Total number of restaurants', xlabel='')
plt.show()
```

Top 10 Cities for Fast Food Accessibility

