In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

In [2]:
```
df8 = pd.read_excel("MainJuly10.xlsx")
```

In [3]:
```
df8
```

Out[3]:

	Restaurant ID	Restaurant Name	Country	City	Cuisines	Average_cost_for_two	Has Table booking	Has Online delivery	Price range	A
0	7402935	Skye	Indonesia	Jakarta	Italian	800000	No	No	3	
1	7402935	Skye	Indonesia	Jakarta	Continental	800000	No	No	3	
2	7410290	Satoo - Hotel Shangri-La	Indonesia	Jakarta	Asian	800000	No	No	3	
3	7410290	Satoo - Hotel Shangri-La	Indonesia	Jakarta	Indonesian	800000	No	No	3	
4	7410290	Satoo - Hotel Shangri-La	Indonesia	Jakarta	Western	800000	No	No	3	
...	...	...	...	...	...	...	...	...	...	...
19696	18312106	UrbanCrave	India	Kanpur	Italian	0	No	No	1	
19697	18312106	UrbanCrave	India	Kanpur	Beverages	0	No	No	1	
19698	3900245	Deena Chat Bhandar	India	Varanasi	Street Food	0	No	No	1	
19699	18246202	VNS Live Studio	India	Varanasi	Chinese	0	No	No	1	
19700	18246202	VNS Live Studio	India	Varanasi	North Indian	0	No	No	1	

19701 rows × 13 columns

In [10]:
```
df8.columns
```

Out[10]:
```
Index(['Restaurant ID', 'Restaurant Name', 'Country', 'City', 'Cuisines',
       'Average_cost_for_two', 'Has Table booking', 'Has Online delivery',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating text',
       'Votes'],
      dtype='object')
```

In [12]:
```
df8.rename(columns={'Restaurant ID':'Restaurant_id',
                    'Restaurant Name':'Restaurant_name',
                    'Has Table booking':'Table_booking',
                    'Has Online delivery':'Online_delivery',
                    'Price range':'Price_range',
                    'Aggregate rating':'Aggregate_rating',
                    'Rating color':'Rating_color',
                    'Rating text':'Rating_text'},inplace=True)
```

In [13]:
```
df8.columns
```

Out[13]:
```
Index(['Restaurant_id', 'Restaurant_name', 'Country', 'City', 'Cuisines',
       'Average_cost_for_two', 'Table_booking', 'Online_delivery',
       'Price_range', 'Aggregate_rating', 'Rating_color', 'Rating_text',
       'Votes'],
      dtype='object')
```

In [30]:
```
df8['City'].nunique()
```

Out[30]:
```
140
```

In [174…
```
z = df8.groupby('City')
z
```

Out[174…
```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x000001CDD577AAF0>
```

In [4]:
```
for name, group in z:
    print(name)
    print(group)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-4-6d5a075b9d27> in <module>
----> 1 for name, group in z:
      2     print(name)
      3     print(group.drop_duplicates('Restaurant_id'))

NameError: name 'z' is not defined
```

In [202…
```
z.get_group('New Delhi').drop_duplicates('Restaurant_id')
```

Out[202…

	Restaurant_id	Restaurant_name	Country	City	Cuisines	Average_cost_for_two	Table_booking	Online_d
47	2701	Orient Express - Taj Palace Hotel	India	New Delhi	European	8000	Yes	
48	309548	Tian - Asian Cuisine Studio - ITC Maurya	India	New Delhi	Asian	7000	No	
53	2742	Bukhara - ITC Maurya	India	New Delhi	North Indian	6500	No	
57	301523	Nostalgia at 1911 Brasserie - The Imperial	India	New Delhi	European	6000	Yes	
59	2724	1911 - The Imperial	India	New Delhi	North Indian	6000	Yes	
...	...	...	...	...	...	...	...	...
18611	302757	Shahi Kachauri	India	New Delhi	Street Food	50	No	
18612	302835	Aggarwal Jalebi Wale	India	New Delhi	Street Food	50	No	
18613	8530	Aggarwal Confectionary	India	New Delhi	Street Food	50	No	
18615	18424202	Famous Parantha and Poori Sabzi	India	New Delhi	North Indian	50	No	
18616	18372694	Sweets n Treats	India	New Delhi	Street Food	50	No	

5473 rows × 13 columns

```
ort_values(by = 'Restaurant_id', ascending = False)
```

Out[85]:

City	Restaurant_id	Restaurant_name	Country	Cuisines	Average_cost_for_two	Table_booking	Online_del
New Delhi	5473	4295	1	81	76	2	
Gurgaon	1118	943	1	71	47	2	
Noida	1080	902	1	49	41	2	
Faridabad	251	235	1	30	25	2	
Ghaziabad	25	25	1	14	15	2	
...	...	...	...	...	...	...	
Paynesville	1	1	1	1	1	1	
Penola	1	1	1	3	1	1	
Phillip Island	1	1	1	3	1	1	
Potrero	1	1	1	3	1	1	
Yorkton	1	1	1	1	1	1	

140 rows × 12 columns

In [129…

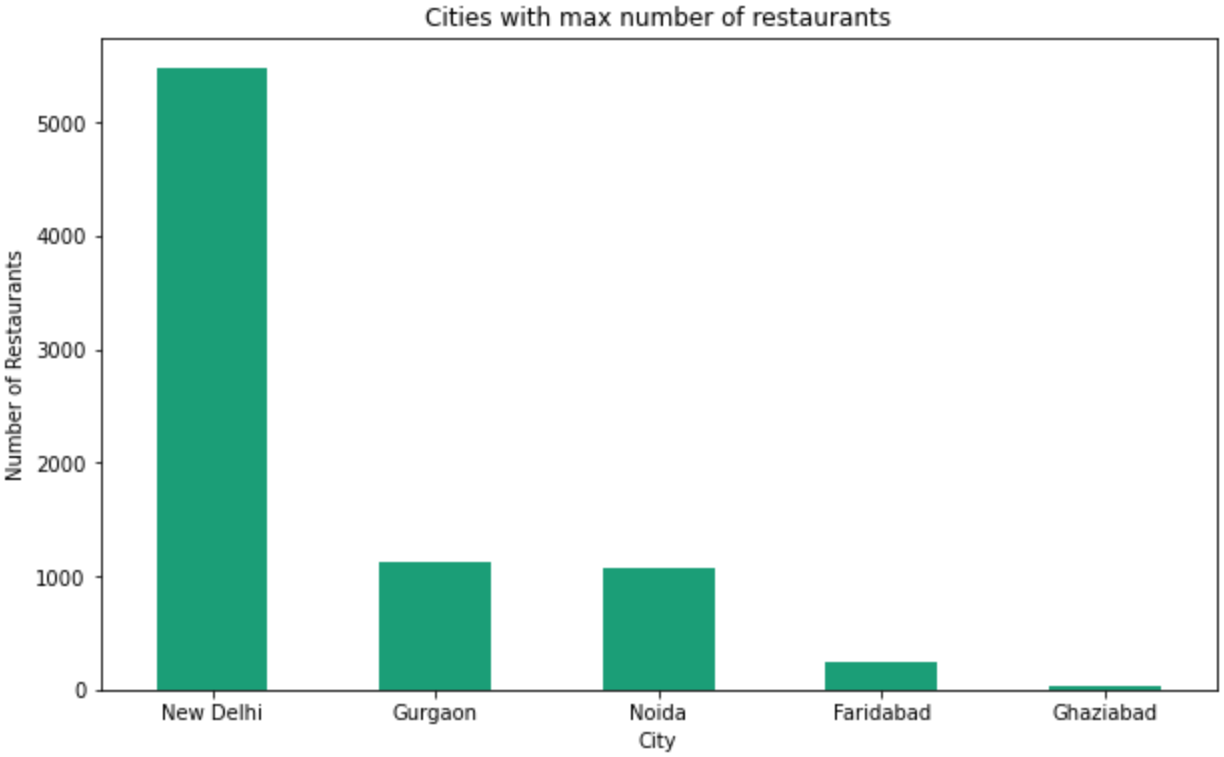
```
Cities_with_max_no_of_rest = df8.groupby('City').Restaurant_id.nunique().sort_values(ascer
Cities_with_max_no_of_rest
```

Out[129…

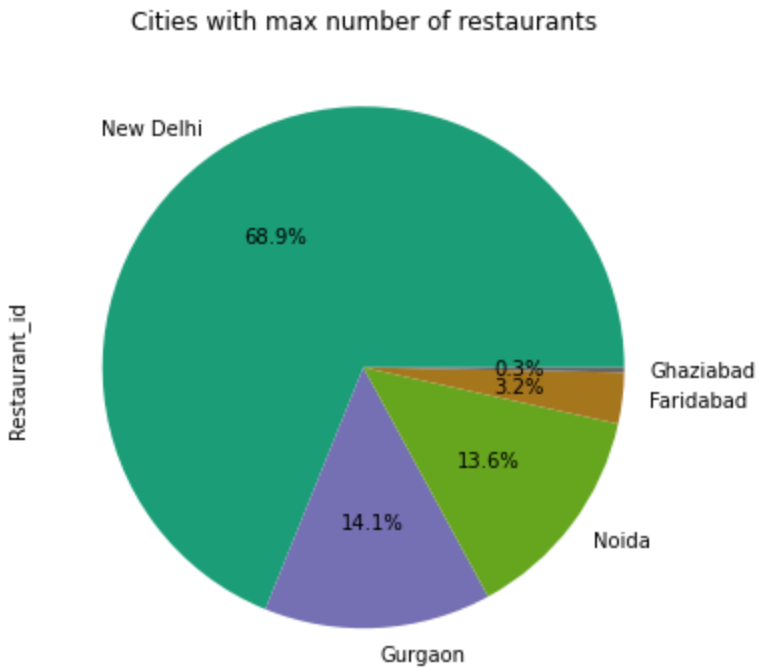
```
City
New Delhi    5473
Gurgaon      1118
Noida        1080
Faridabad     251
Ghaziabad      25
Name: Restaurant_id, dtype: int64
```

In [149…

```
Cities_with_max_no_of_rest.plot(kind='bar',
                                ylabel='Number of Restaurants',
                                title = 'Cities with max number of restaurants',
                                figsize=(10,6),
                                rot=0,
                                cmap='Dark2')
plt.show()
```



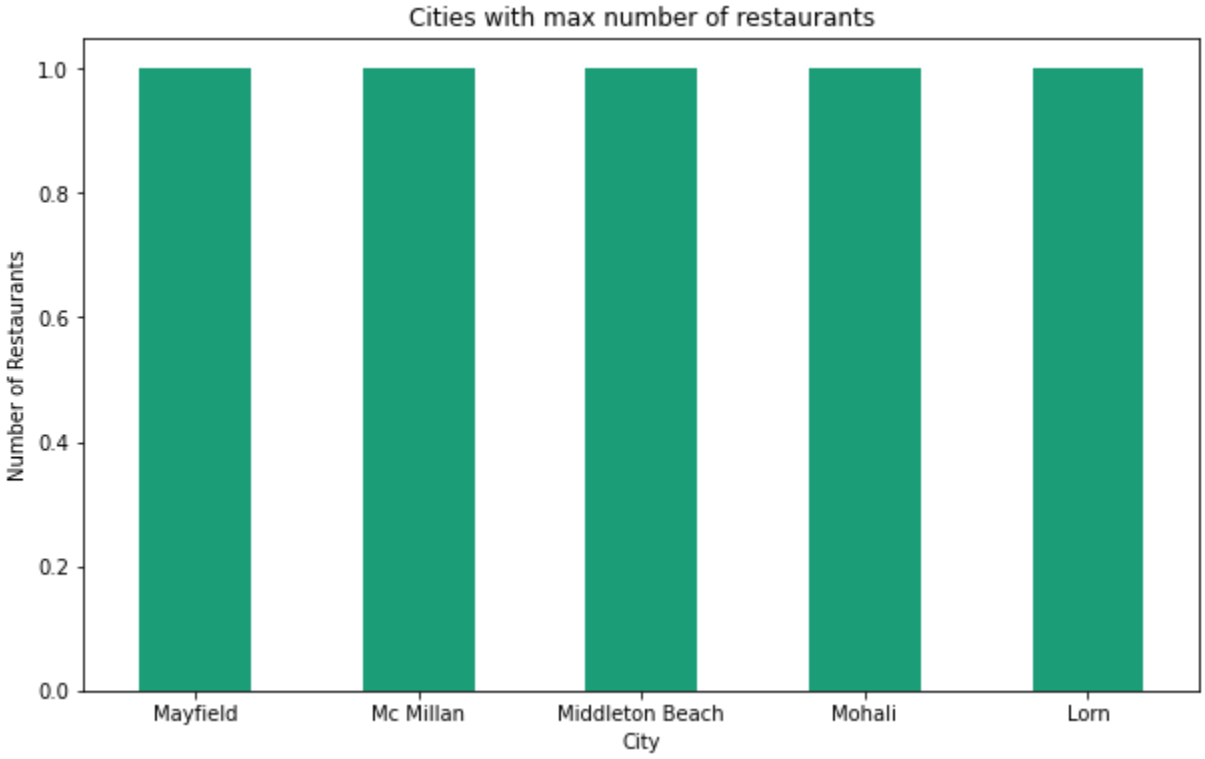
In [152…
```
Cities_with_max_no_of_rest.plot(kind='pie',
                                title = 'Cities with max number of restaurants',
                                figsize=(10,6),
                                autopct='%.1f%%',
                                cmap='Dark2')
plt.show()
```



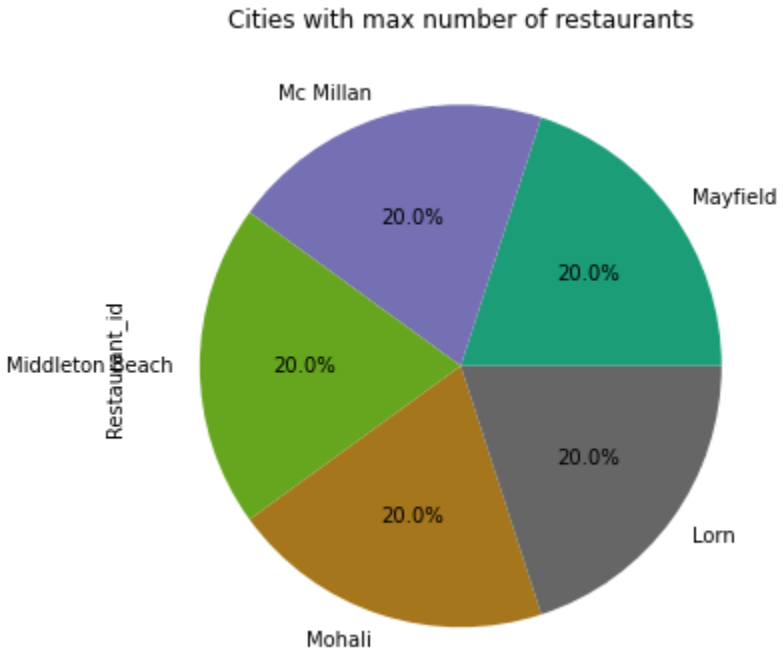
In [153…
```
Cities_with_min_no_of_rest = df8.groupby('City').Restaurant_id.nunique().sort_values(ascer
Cities_with_min_no_of_rest
```

Out[153…
```
City
Mayfield          1
Mc Millan         1
Middleton Beach   1
Mohali            1
Lorn              1
```

In [154…
```
Cities_with_min_no_of_rest.plot(kind='bar',
                                ylabel='Number of Restaurants',
                                title = 'Cities with max number of restaurants',
                                figsize=(10,6),
                                rot=0,
                                cmap='Dark2')
plt.show()
```



In [155…
```
Cities_with_min_no_of_rest.plot(kind='pie',
                                title = 'Cities with max number of restaurants',
                                figsize=(10,6),
                                autopct='%.1f%%',
                                cmap='Dark2')
plt.show()
```



In [209…
```
df8.to_excel('July 11 2022 data.xlsx',index=False)
```

In [ ]: