swiggy_data_analysis

June 10, 2024

0.1 Swiggy data analysis

[2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.graph_objs as go
from plotly.subplots import make_subplots
```

[3]:
```
swiggy=pd.read_csv("swiggy.csv")
```

[6]:
```
swiggy
```

[6]:

	id	name	city	rating	rating_count	cost	cuisine	lic_no
0	567335	AB FOODS POINT	Abohar	--	Too Few Ratings	200	Beverages,Pizzas	22122652000138
1	531342	Janta Sweet House	Abohar	4.4	50+ ratings	200	Sweets,Bakery	12117201000112
2	158203	theka coffee desi	Abohar	3.8	100+ ratings	100	Beverages	22121652000190
3	187912	Singh Hut	Abohar	3.7	20+ ratings	250	Fast Food,Indian	22119652000167
4	543530	GRILL MASTERS	Abohar	--	Too Few Ratings	250	Italian-American,Fast Food	12122201000053
…	…	…	…	…	…	…	…	…
148536	553122	The Food Delight	Yavatmal	--	Too Few Ratings	200	Fast Food,Snacks	21522053000452
148537	562647	MAITRI FOODS & BEVERAGES	Yavatmal	--	Too Few Ratings	300	Pizzas	license
148538	559435	Cafe Bella Ciao	Yavatmal	--	Too Few Ratings	300	Fast Food,Snacks	21522251000378
148539	418989	GRILL ZILLA	Yavatmal	--	Too Few Ratings	250	Continental	21521251000241
148540	447770	Lazeez kitchen	Yavatmal	--	Too Few Ratings			

```
148540    200                        Pizzas  21521251000634

                                                      link  \
0       https://www.swiggy.com/restaurants/ab-foods-po…
1       https://www.swiggy.com/restaurants/janta-sweet…
2       https://www.swiggy.com/restaurants/theka-coffe…
3       https://www.swiggy.com/restaurants/singh-hut-n…
4       https://www.swiggy.com/restaurants/grill-maste…
…                                                     …
148536  https://www.swiggy.com/restaurants/the-food-de…
148537  https://www.swiggy.com/restaurants/maitri-food…
148538  https://www.swiggy.com/restaurants/cafe-bella-…
148539  https://www.swiggy.com/restaurants/grill-zilla…
148540  https://www.swiggy.com/restaurants/lazeez-kitc…

                                                 address              menu
0       AB FOODS POINT, NEAR RISHI NARANG DENTAL CLINI…  Menu/567335.json
1       Janta Sweet House, Bazar No.9, Circullar Road,…  Menu/531342.json
2              theka coffee desi, sahtiya sadan road city  Menu/158203.json
3         Singh Hut, CIRCULAR ROAD NEAR NEHRU PARK ABOHAR  Menu/187912.json
4       GRILL MASTERS, ADA Heights, Abohar - Hanumanga…  Menu/543530.json
…                                                     …                …
148536  The Food Delight, 94MC+X35, New Singhania Naga…  Menu/553122.json
148537  MAITRI FOODS & BEVERAGES, POLIC MITRYA SOCIETY…  Menu/562647.json
148538  Cafe Bella Ciao, SHOP NO 2 NEMANI MARKET SBI S…  Menu/559435.json
148539  GRILL ZILLA, SHO NO 2/6, POSTEL GROUND CHOWPAT…  Menu/418989.json
148540  Lazeez kitchen, 94G3+2RR, Wadgaon, Yavatmal, M…  Menu/447770.json

[148541 rows x 11 columns]
```

[8]:
```
swiggy.columns
```

[8]:
```
Index(['id', 'name', 'city', 'rating', 'rating_count', 'cost', 'cuisine',
       'lic_no', 'link', 'address', 'menu'],
      dtype='object')
```

[14]:
```
#droping not require column
col_required = ['name', 'city', 'rating', 'rating_count', 'cost', 'cuisine']
swiggy = swiggy[col_required]
```

[16]:
```
swiggy.head(10)
```

[16]:
```
                    name    city rating     rating_count   cost  \
0          AB FOODS POINT  Abohar     --  Too Few Ratings    200
1      Janta Sweet House   Abohar    4.4      50+ ratings    200
2      theka coffee desi   Abohar    3.8     100+ ratings    100
3              Singh Hut   Abohar    3.7      20+ ratings    250
```

```
4              GRILL MASTERS  Abohar     --  Too Few Ratings    250
5                  Sam Uncle  Abohar    3.6      20+ ratings    200
6           shere punjab veg  Abohar    4.0     100+ ratings    150
7  Shri Balaji Vaishno Dhaba  Abohar     --  Too Few Ratings    100
8   Hinglaj Kachori Bhandhar  Abohar    4.2      20+ ratings    100
9                  yummy hub  Abohar     --  Too Few Ratings    200

                     cuisine
0            Beverages,Pizzas
1               Sweets,Bakery
2                   Beverages
3            Fast Food,Indian
4  Italian-American,Fast Food
5                 Continental
6               North Indian
7               North Indian
8               Snacks,Chaat
9                     Indian
```

[18]:
```
swiggy.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148541 entries, 0 to 148540
Data columns (total 6 columns):
 #   Column        Non-Null Count   Dtype 
---  ------        --------------   ----- 
 0   name          148455 non-null  object
 1   city          148541 non-null  object
 2   rating        148455 non-null  object
 3   rating_count  148455 non-null  object
 4   cost          148410 non-null  object
 5   cuisine       148442 non-null  object
dtypes: object(6)
memory usage: 6.8+ MB
```

[20]:
```
swiggy.isnull().sum()
```

[20]:
```
name            86
city             0
rating          86
rating_count    86
cost           131
cuisine         99
dtype: int64
```

[30]:
```
swiggy.dropna(inplace = True)
```

[32]:
```
swiggy.isnull().sum()
```

[32]:
```
name            0
city            0
rating          0
rating_count    0
cost            0
cuisine         0
dtype: int64
```

[31]:
```
swiggy.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 148398 entries, 0 to 148540
Data columns (total 6 columns):
 #   Column        Non-Null Count   Dtype 
---  ------        --------------   ----- 
 0   name          148398 non-null  object
 1   city          148398 non-null  object
 2   rating        148398 non-null  object
 3   rating_count  148398 non-null  object
 4   cost          148398 non-null  object
 5   cuisine       148398 non-null  object
dtypes: object(6)
memory usage: 7.9+ MB
```

[34]:
```
swiggy1=swiggy.copy()
```

[36]:
```
swiggy1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 148398 entries, 0 to 148540
Data columns (total 6 columns):
 #   Column        Non-Null Count   Dtype 
---  ------        --------------   ----- 
 0   name          148398 non-null  object
 1   city          148398 non-null  object
 2   rating        148398 non-null  object
 3   rating_count  148398 non-null  object
 4   cost          148398 non-null  object
 5   cuisine       148398 non-null  object
dtypes: object(6)
memory usage: 7.9+ MB
```

[38]:
```
# Lets apply the methods to replace the values
swiggy1['rating'] = swiggy1['rating'].str.replace('--', '0')
swiggy1['cost'] = swiggy1['cost'].str.replace(' ', '')
```

```
swiggy1['rating_count'] = swiggy1['rating_count'].str.replace('+','').str.
 ↪replace('K','000').str.replace('Too Few Ratings', '0').str.
 ↪replace('ratings', '')
```

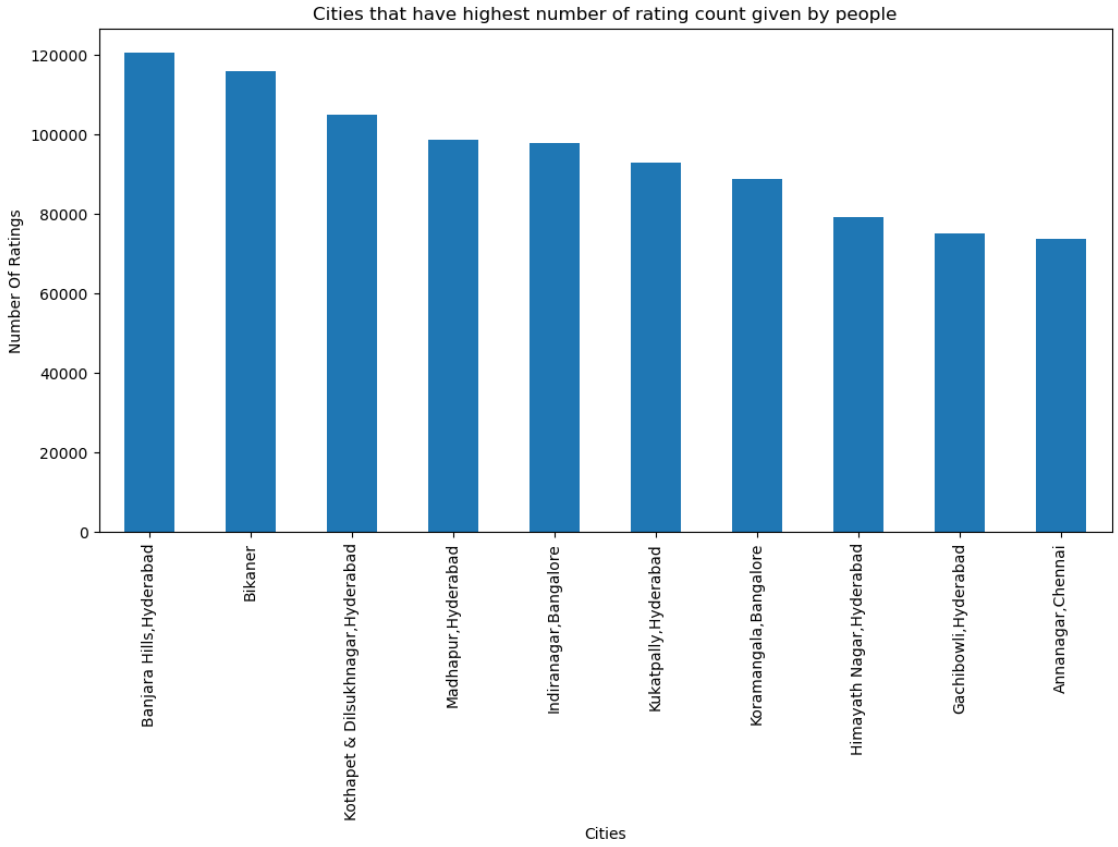
[40]:
```
swiggy1['rating'] = swiggy1['rating'].astype(float)
swiggy1['rating_count'] = swiggy1['rating_count'].astype(int)
swiggy1['cost'] = swiggy1['cost'].astype(int)
```

0.2 Data analysis

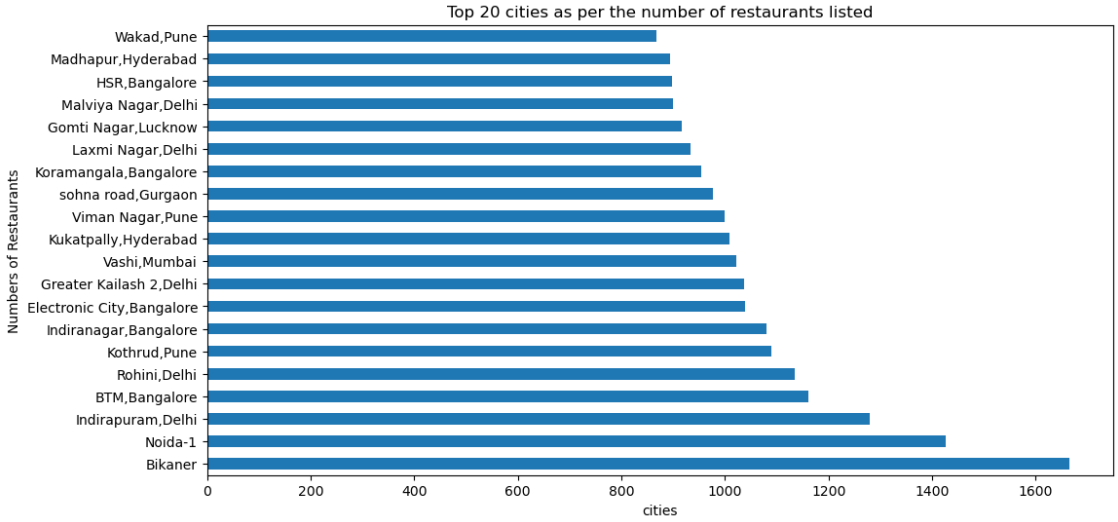
[74]:
```
plt.rcParams['font.size'] = 14
plt.rcParams['figure.figsize'] = (10, 6)
```

[46]:
```
#highest number of rating count given by people
highest_city_rating_count = swiggy1.groupby('city')['rating_count'].sum().
 ↪sort_values(ascending=False)

plt.figure(figsize=(12,6))
highest_city_rating_count.head(10).plot.bar()
plt.xlabel('Cities')
plt.ylabel('Number Of Ratings')
plt.title('Cities that have highest number of rating count given by people')
plt.show();
```

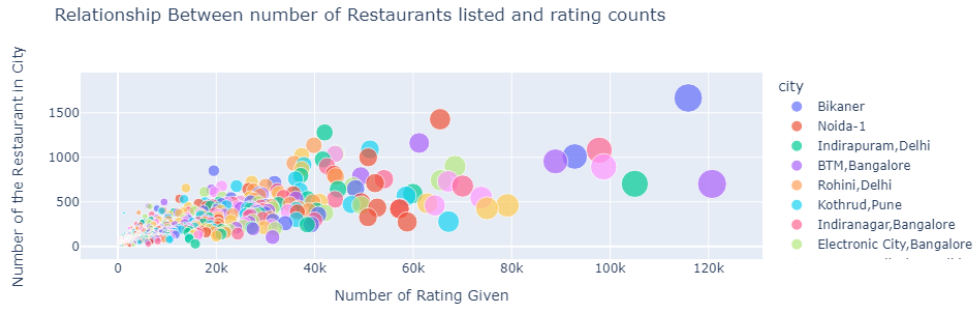


[48]:
```
#Top 20 cities as per the number of restaurants listed
top_rest=swiggy1.groupby('city')['name'].count().sort_values(ascending=False)
plt.figure(figsize=(12,6))
top_rest.head(20).plot.barh()
plt.xlabel('cities')
plt.ylabel('Numbers of Restaurants')
plt.title('Top 20 cities as per the number of restaurants listed')
plt.show()
```



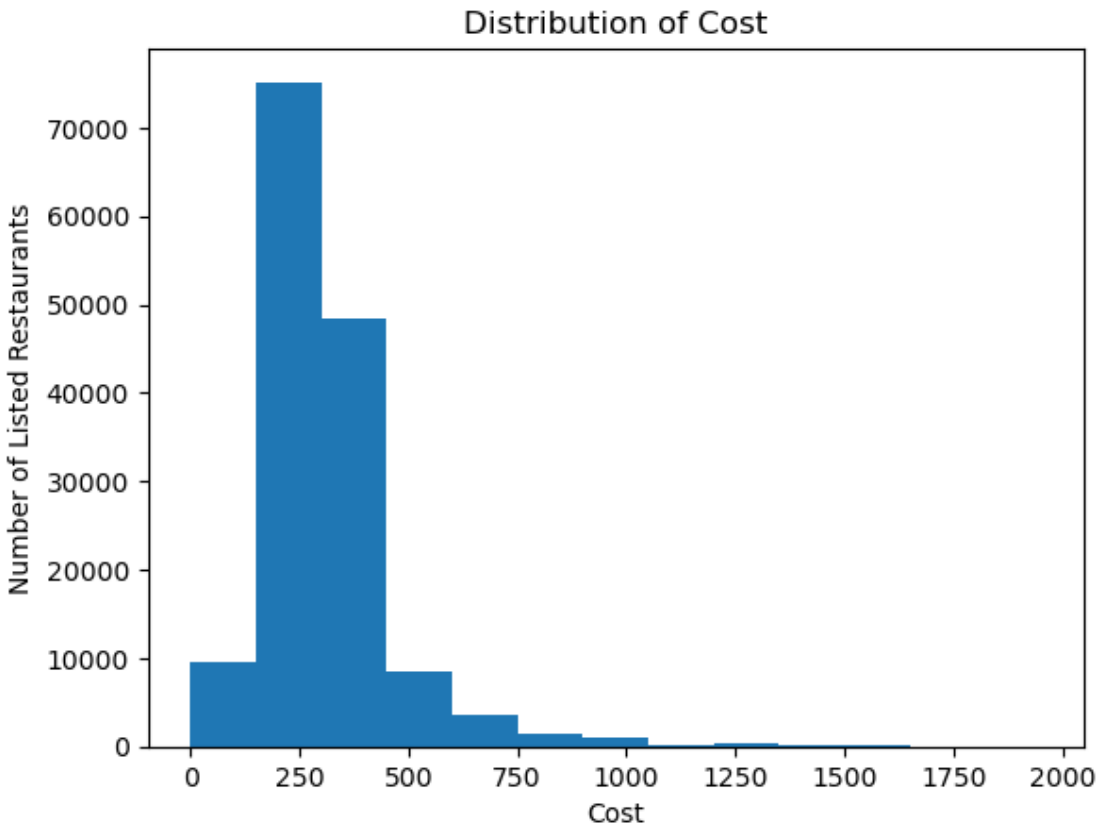
[50]:
```
combined=pd.merge(top_rest,highest_city_rating_count,on='city')
```

[52]:
```
fig=px.
 ↪scatter(combined,x='rating_count',y='name',size='rating_count',title='Relationship␣
 ↪Between number of Restaurants listed and rating counts',color=combined.index)
fig.update_yaxes(title_text='Number of the Restaurant in City')
fig.update_xaxes(title_text='Number of Rating Given')
fig.show()
```



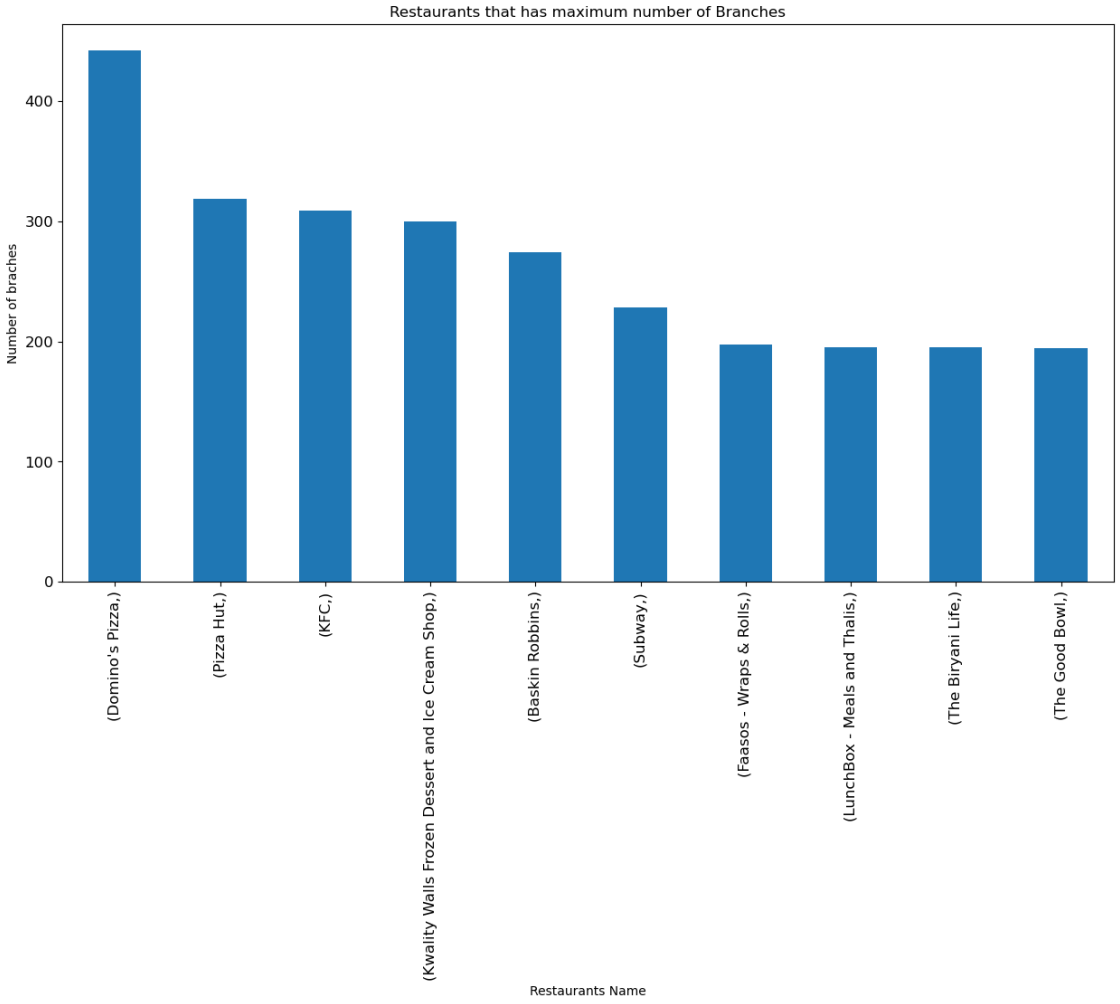
[54]:
```
# plotting graph of coast in hisgtogram
# ploting the distribution of cost
plt.hist(swiggy1['cost'],bins=np.arange(0,2000,150));
plt.title('Distribution of Cost')
plt.xlabel('Cost')
plt.ylabel('Number of Listed Restaurants')
```

[54]: Text(0, 0.5, 'Number of Listed Restaurants')



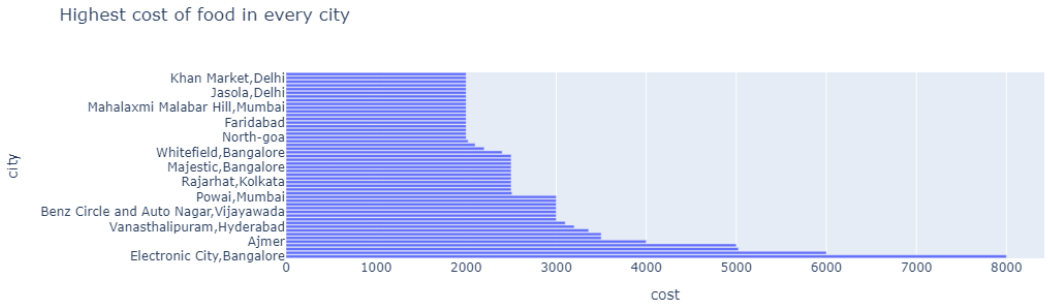
[47]:
```
#maximum number of branches
restaurant_chain=swiggy1[['name']].value_counts().sort_values(ascending=False).
 ↪head(10)
pd.options.plotting.backend='matplotlib'
restaurant_chain.plot(kind='bar',legend=False,
                      title='Restaurants that has maximum number of Branches',
                      xlabel='Restaurants Name',
                      ylabel='Number of braches',
                      figsize=(15,8),
                      fontsize=12)
```

[47]: <Axes: title={'center': 'Restaurants that has maximum number of Branches'}, xlabel='Restaurants Name', ylabel='Number of braches'>



[82]:
```
swiggy1 = swiggy1.drop(swiggy1[swiggy1['cost']>10000].index)
```

[84]:
```
# highest cost of food by city
highest_cost=swiggy1.groupby("city")["cost"].max().sort_values(ascending=False).
  ↪reset_index()
fig=px.bar(highest_cost.head(50),
          x='cost',y='city',title='Highest cost of food in every city')
fig.show()
```



[86]:
```
rating_count_100=swiggy1[swiggy1['rating_count']>=100]
```

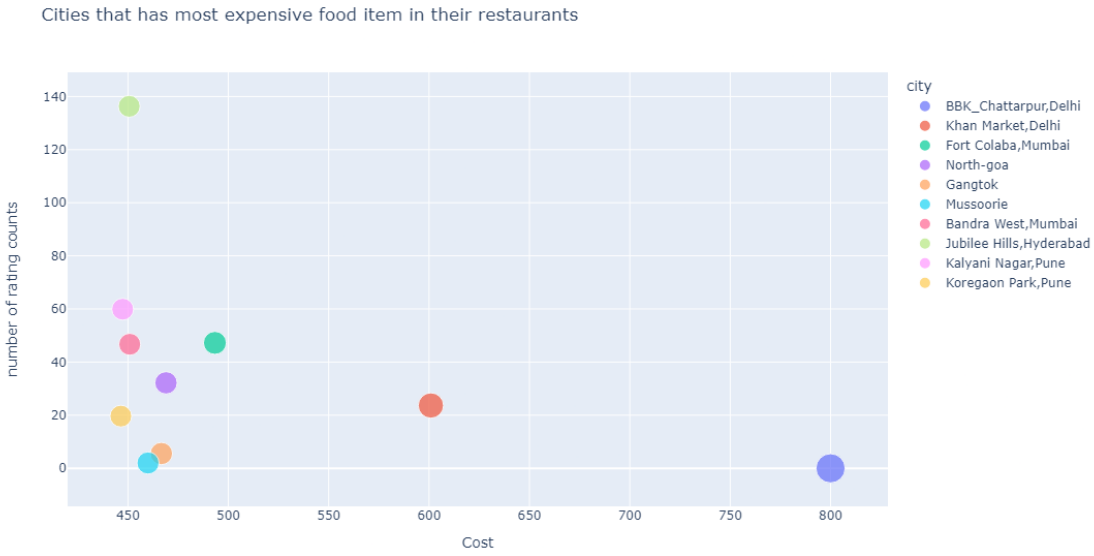
[88]:
```
top_5_popular_chain=rating_count_100.groupby('name')['rating'].mean().
 ↪sort_values(ascending=False).head(5)
top_5_popular_chain
```

[88]:
```
name
Tony Bakery                     5.0
SHRI KRISHNA FOOD ZONE          5.0
Mahi Bakery                     4.9
Aaradhya cafe and fast food     4.9
Ram Fast Food                   4.9
Name: rating, dtype: float64
```

#what are the those cities that has most expensive food in their restaurants

[90]:
```
most_exp=swiggy1.groupby('city')[['cost','rating_count','rating']].mean().
 ↪sort_values(by='cost',ascending=False).head(10)
city_cost=most_exp.reset_index()
```

[94]:
```
fig=px.scatter(city_cost,
               x='cost',
               y='rating_count',
               size='cost',
               title='Cities that has most expensive food item in their␣
 ↪restaurants',
               color='city',
               hover_data={'city' : True})
fig.update_yaxes(title_text='number of rating counts')
fig.update_xaxes(title_text='Cost')
fig.update_layout(width=800,height=600)
fig.show()
```



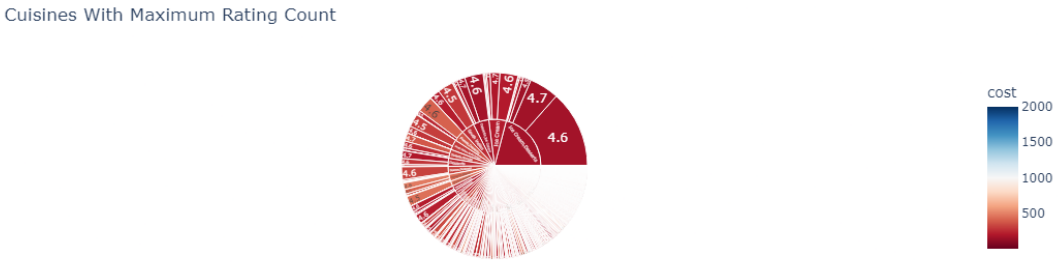
[96]:
```
highest_cost_name= swiggy1.groupby('city')['cost'].max().
  ↪sort_values(ascending=False).reset_index()
highest_cost=swiggy1.
  ↪merge(highest_cost_name,on=['city','cost'],how='inner')[['name','city','rating','rating_cou
```

[ ]:
```
```

[98]:
```
fig=px.treemap(highest_cost.head(200),
               path=['city','name','rating_count','rating'],
               color_continuous_scale='RdBu',
               values='cost',
               color='rating',
               title='cities with maximum cost')
fig.show()
```

[100]:
```
highest_rating_count= swiggy1.sort_values(by='rating',ascending =False).
↪reset_index()
```

[102]:
```
fig = px.sunburst(highest_rating_count.head(3000),
                  path = ['cuisine','rating'],
                  values = 'rating_count',
                  color = 'cost',
                  color_continuous_scale='RdBu',
                  title = 'Cuisines With Maximum Rating Count')
fig.show()
```



[ ]:
```

```