MAJOR PROJECT 4 - SANJAY ANAND V

SWIGGY EDA AND VISUALISATION

Importing Libraries and dataset

In [1]:

```
import pandas as pd
import numpy as np
import json as js
```

In [2]:

```
from matplotlib import pyplot as plt
import seaborn as sns
```

In [3]:

```
df=pd.read_csv("swiggy.csv")
df.head()
```

Out[3]:

	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

1)The top 10 Most common restaurants in India

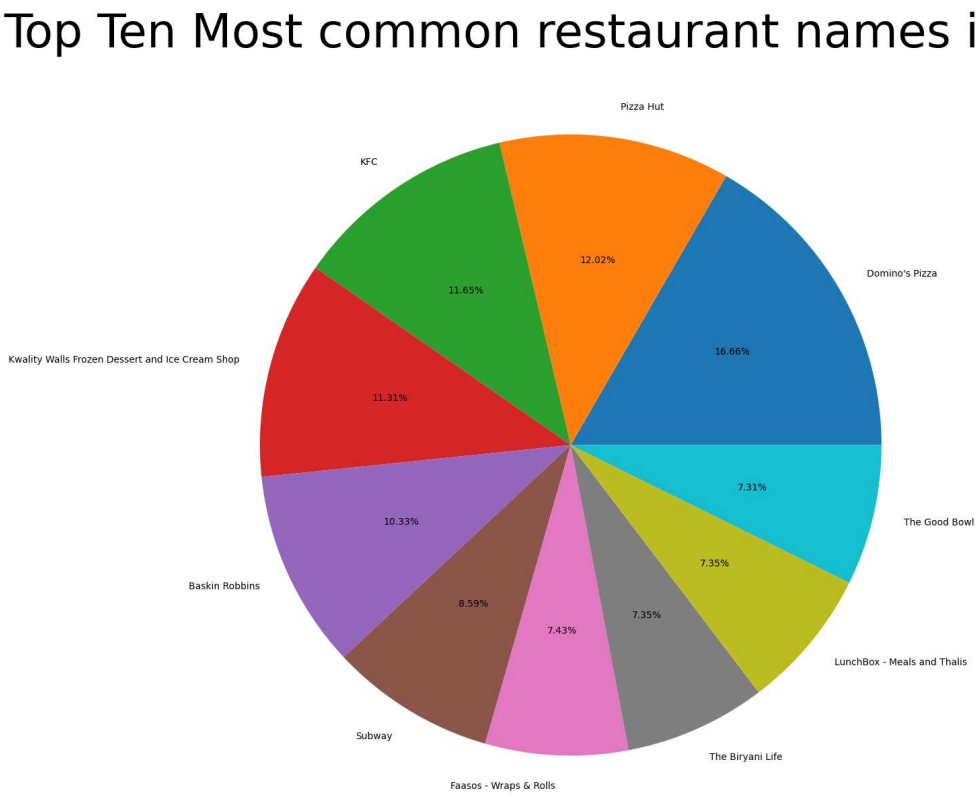
In [4]:

```
data = df.groupby('name').size().sort_values(ascending = False).head(10)
print("The top 10 most common restaurant names in India are")
print("-------------------------------------------------------")
print(data)
print("-------------------------------------------------------")
```

```
The top 10 most common restaurant names in India are
-------------------------------------------------------
name
Domino's Pizza                                  442
Pizza Hut                                       319
KFC                                             309
Kwality Walls Frozen Dessert and Ice Cream Shop 300
Baskin Robbins                                  274
Subway                                          228
Faasos - Wraps & Rolls                          197
The Biryani Life                                195
LunchBox - Meals and Thalis                     195
The Good Bowl                                   194
dtype: int64
-------------------------------------------------------
```

In [5]:

```
percent = data.values
names  = data.index
fig = plt.figure(figsize =(15, 15))
plt.pie(data, labels = names,autopct = '%1.2f%%')
plt.title('Top Ten Most common restaurant names in India', fontsize = 50 )
plt.show()
```



2) The top 10 most common restaurants in Bangalore

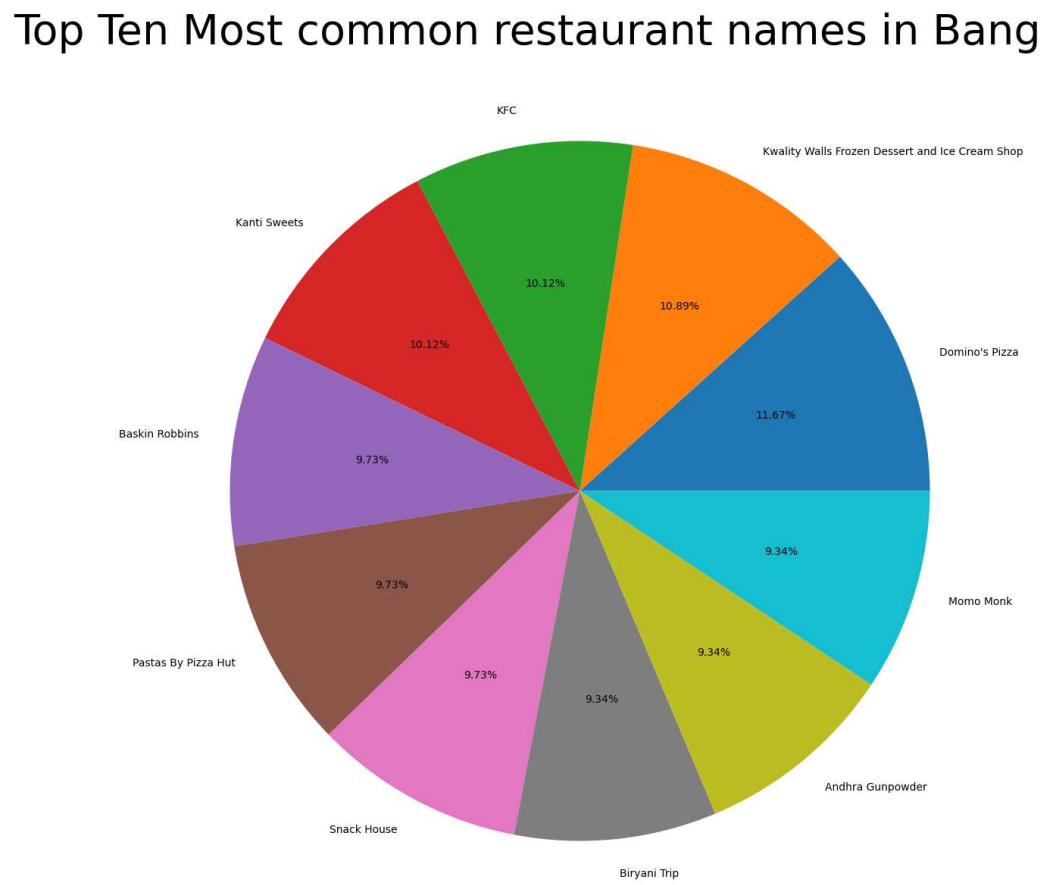
In [6]:

```
cities = [ i for i in df['city'].unique() if 'Bangalore' in i ]

data   = df[['name','city']]
data['include'] = data['city'].apply( lambda city : 1 if city in cities else 0)

data = data[data['include'] == 1]

data = data.groupby('name').size().sort_values(ascending = False).head(10)

print("The top 10 most common restaurant names in Bangalore are")
print("-------------------------------------------------------")
print(data)
print("-------------------------------------------------------")
```

```
The top 10 most common restaurant names in Bangalore are
-------------------------------------------------------
name
Domino's Pizza                                  30
Kwality Walls Frozen Dessert and Ice Cream Shop 28
KFC                                             26
Kanti Sweets                                    26
Baskin Robbins                                  25
Pastas By Pizza Hut                             25
Snack House                                     25
Biryani Trip                                    24
Andhra Gunpowder                                24
Momo Monk                                       24
dtype: int64
-------------------------------------------------------

C:\Users\anand\AppData\Local\Temp\ipykernel_5668\2245140167.py:4: SettingW
ithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-doc
s/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://
pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
view-versus-a-copy)
  data['include'] = data['city'].apply( lambda city : 1 if city in cities
else 0)
```

In [7]:

```
count = data.values
names  = data.index
fig = plt.figure(figsize =(15, 15))
plt.pie(data, labels = names,autopct = '%1.2f%%')
plt.title('Top Ten Most common restaurant names in Bangalore', fontsize = 40 )
plt.show()
```



3) Number of Branches of Dominos

In [8]:

```
l=len(df['name'])
c=0
for i in range(l):
    if(df['name'][i]=="Domino's Pizza"):
        c+=1
print("The number of branches of Dominos are",c)
```

```
The number of branches of Dominos are 442
```

4)Comparison of number of branches on KFC,MacD,Dominos and Subway

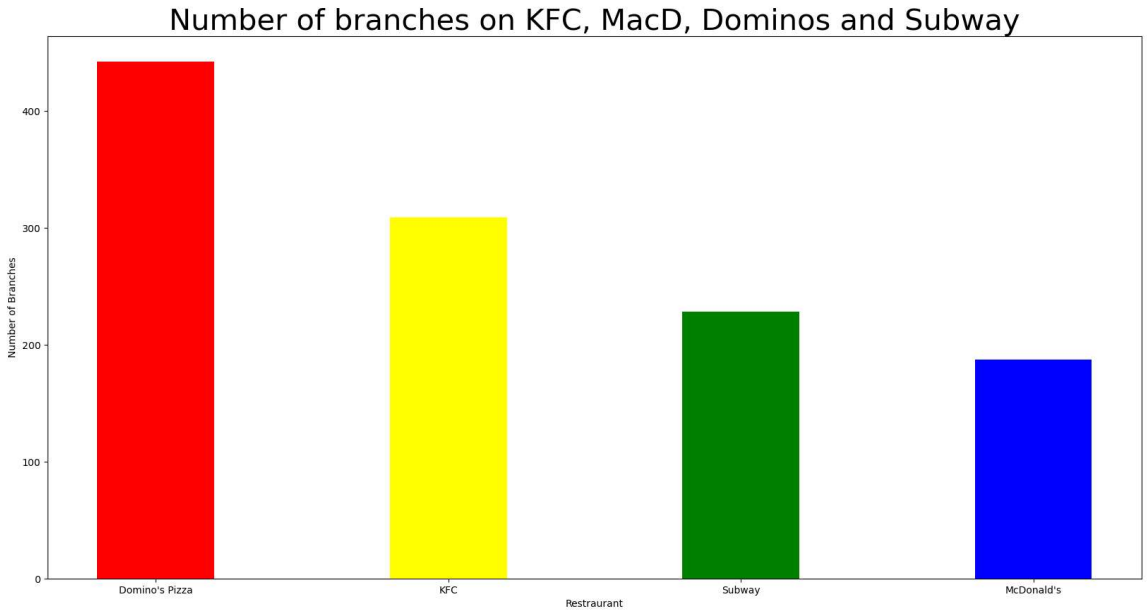
In [9]:

```
data = df.groupby('name').size()[["Domino's Pizza", "KFC", "McDonald's",'Subway']].sort_
print("The number of branches of KFC,MacD,Dominos and Subway are")
print("-------------------------------------------------------")
print(data)
print("-------------------------------------------------------")
```

```
The number of branches of KFC,MacD,Dominos and Subway are
-------------------------------------------------------
name
Domino's Pizza    442
KFC               309
Subway            228
McDonald's        187
dtype: int64
-------------------------------------------------------
```

In [10]:

```
x    = data.index
y    = data.values


plt.figure(figsize = (20,10))

color = ['red','yellow','green','blue']

ax = plt.bar(x,y,width = 0.4,color = color)
plt.xlabel('Restraurant')
plt.ylabel('Number of Branches')
plt.title('Number of branches on KFC, MacD, Dominos and Subway', fontsize=30)
plt.show()
```



5)Top 10 most popular cuisines

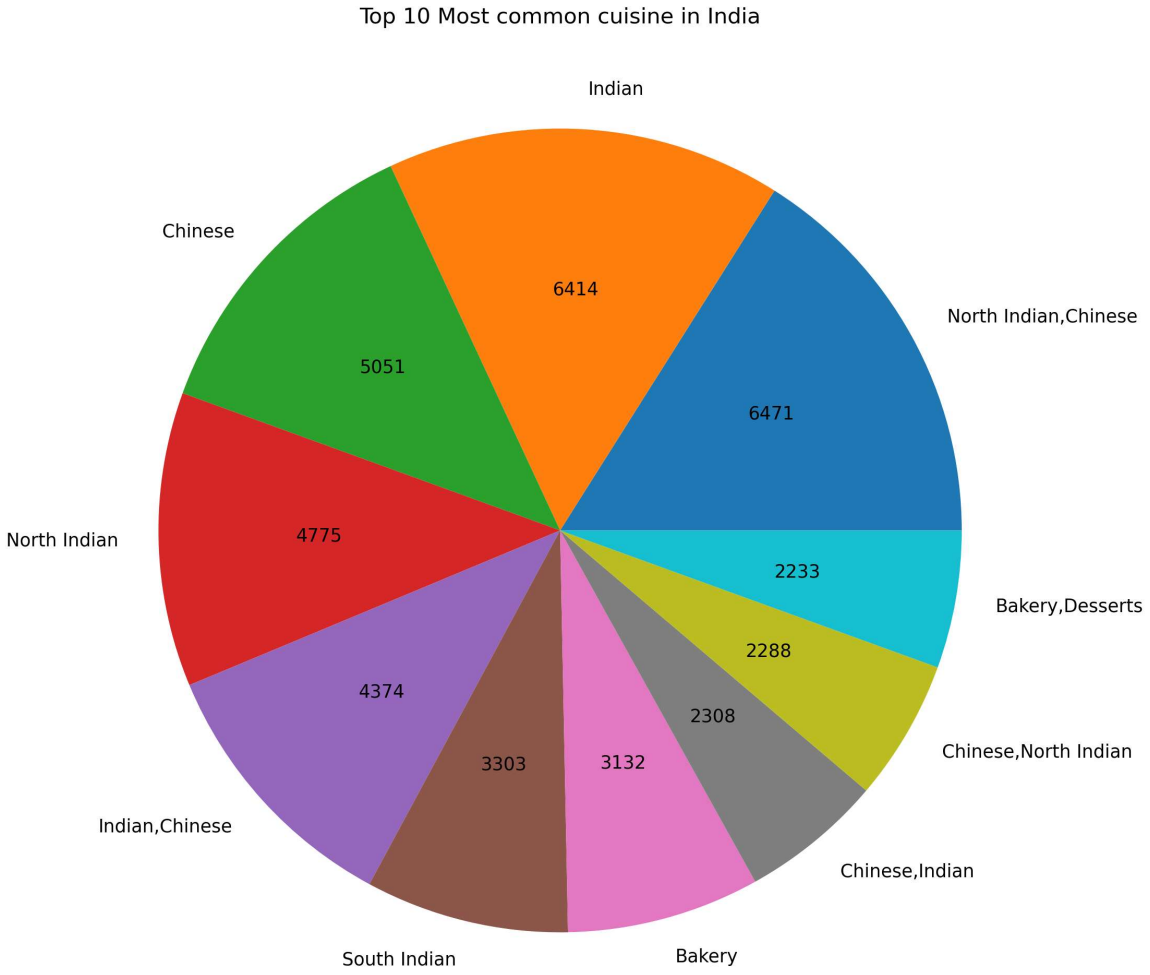
In [11]:

```
dt = df.groupby('cuisine').size().sort_values(ascending = False).head(10)
print("The top 10 most popular cuisines are")
print("-----------------------------------------------------------")
print(dt)
print("-----------------------------------------------------------")
```

```
The top 10 most popular cuisines are
-----------------------------------------------------------
cuisine
North Indian,Chinese    6471
Indian                  6414
Chinese                 5051
North Indian            4775
Indian,Chinese          4374
South Indian            3303
Bakery                  3132
Chinese,Indian          2308
Chinese,North Indian    2288
Bakery,Desserts         2233
dtype: int64
-----------------------------------------------------------
```

In [12]:

```
data = df.groupby('cuisine').size().sort_values(ascending = False).head(10)

count = data.values
names  = data.index

def format(p,count):
    value = p/100 * sum(count)

    value += 1 if float('.'+str(value).split('.')[-1]) > 0.5 else 0
    return f'{int(value)}'

plt.rc('font', **{'size':25})
plt.figure(figsize=(25,25))
plt.pie(count,labels = names,autopct = lambda p : format(p,count))

plt.title('Top 10 Most common cuisine in India', fontsize = 30 )
plt.show()
```



Loading JSON Dataset

In [13]:

```
fd   = open('data.json','r')
dct = js.loads(fd.read())
fd.close()
dct.keys()
```

Out[13]:

```
dict_keys(['Abohar', 'Adilabad', 'Adityapur', 'Adoni', 'Agartala', 'Agra',
'Ahmedabad', 'Ahmednagar', 'Aizawl', 'Ajmer', 'Akola', 'Alappuzha', 'Aliga
rh', 'Alipurduar', 'Allahabad', 'Almora', 'Alwar', 'Amalapuram', 'Ambala',
'Ambikapur', 'Ambur', 'Amravati', 'Amreli', 'Amritsar', 'Anand', 'Anantapu
r', 'Angul', 'Ankleshwar', 'Arakkonam', 'Arambagh', 'Arrah', 'Aruppukotta
i', 'Asansol', 'Aurangabad', 'Aurangabad_bihar', 'Azamgarh', 'Baddi', 'Bag
alkot', 'Bagdogra', 'Bagula', 'Bahadurgarh', 'Bahraich', 'Balaghat', 'Bala
ngir', 'Balasore', 'Ballari', 'Balrampur', 'Balurghat', 'Banda', 'Bangalor
e', 'Bankabihar', 'Bankura', 'Bantwal', 'Bapatlachirala', 'Baramati', 'Bar
an', 'Baraut', 'Bardhaman', 'Bardoli', 'Bareilly', 'Baripada', 'Barmer',
'Barnala', 'Barshi', 'Barwani', 'Basirhat', 'Basti', 'Batala', 'Bathinda',
'Beawar', 'Beed', 'Begusarai', 'Bela-pratapgarh', 'Belgaum', 'Berhampore',
'Berhampur', 'Bettiah', 'Betul', 'Bhadohi', 'Bhadrachalam', 'Bhadrak', 'Bh
adravati', 'Bhagalpur', 'Bhandara', 'Bharabanki', 'Bharatpur', 'Bharuch',
'Bhatkal', 'Bhavnagar', 'Bhawanipatna', 'Bhilai', 'Bhilwara', 'Bhimavara
m', 'Bhind', 'Bhiwadi', 'Bhiwani', 'Bhopal', 'Bhubaneswar', 'Bhuj', 'Bhusa
wal', 'Bidar', 'Biharsharif', 'Bijapur', 'Bijnor', 'Bikaner', 'Bilaspur',
'Bilaspur-hp', 'Bilimora', 'Biswanath-chariali', 'Bodhan-rural', 'Bodinaya
kanur', 'Boisar', 'Bokaro', 'Bolpur', 'Bongaigaon', 'Bongaon', 'Budaun',
'Budhwal', 'Bulandshahr', 'Buldana', 'Bundi', 'Burhanpur', 'Buxar', 'Centr
al-goa', 'Chakdaha', 'Chalakkudy', 'Chalisgaon', 'Chandausi', 'Chandigar
h', 'Chandrapur', 'Changanassery', 'Chengannur', 'Chennai', 'Cherthala',
'Chhapra', 'Chhatarpur', 'Chhindwara', 'Chidambaram', 'Chikhli', 'Chikkaba
llapur', 'Chikmagalur', 'Chiplun', 'Chitradurga', 'Chittoor', 'Chittorgar
h', 'Chopda', 'Churu', 'Coimbatore', 'Cooch-behar', 'Cuddalore', 'Cuttac
k', 'Dahanu', 'Dahod', 'Daltonganj', 'Daman', 'Darbhanga', 'Darjeeling',
'Daund', 'Dausa', 'Davanagere', 'Dehradun', 'Dehri', 'Delhi', 'Deoghar',
'Dewas', 'Dhanbad', 'Dhar', 'Dharamshala', 'Dharapuram', 'Dharmapuri', 'Dh
arwad', 'Dhoraji', 'Dhule', 'Dibrugarh', 'Digboi', 'Dimapur', 'Dindigul',
'Diu', 'Doddaballapura', 'Duliajan', 'Dumka', 'Durgapur', 'Eluru', 'Erod
e', 'Etah', 'Etawah', 'Faizabad', 'Faridabad', 'Faridkot', 'Farrukhabad',
'Fatehabad', 'Fatehgarh-sahib', 'Fatehpur', 'Fazilka', 'Firozabad', 'Firoz
pur', 'Freelancer', 'Gadag-betigeri', 'Gadwal', 'Gandhidham', 'Gangapur-ci
ty', 'Gangarampur', 'Gangtok', 'Gauriganj', 'Gaya', 'Giridih', 'Godhra',
'Gokak', 'Golaghat', 'Gonda', 'Gondal', 'Gondia', 'Gopalganj', 'Gorakhpu
r', 'Gudivada', 'Guna', 'Guntakal', 'Guntur', 'Gurdaspur', 'Gurgaon', 'Guw
ahati', 'Gwalior', 'Habra', 'Hajipur', 'Haldia', 'Haldwani', 'Halol', 'Ham
pi', 'Hansi', 'Hanumangarh', 'Hapur', 'Hardoi', 'Haridwar', 'Hassan', 'Hat
hras', 'Himmatnagar', 'Hindaun', 'Hinganghat', 'Hisar', 'Hoshangabad', 'Ho
shiarpur', 'Hospet', 'Hosur', 'Hubli', 'Hyderabad', 'Ichalkaranji', 'Idukk
i', 'Imphal', 'Indore', 'Irinjalakuda', 'Itanagar', 'Itarsi', 'Jabalpur',
'Jagdalpur', 'Jagraon', 'Jagtial', 'Jahanabad', 'Jaigaon', 'Jaipur', 'Jala
ndhar', 'Jalaun', 'Jalgaon', 'Jalpaiguri', 'Jammu', 'Jamnagar', 'Jamshedpu
r', 'Jaunpur', 'Jhalawar', 'Jhansi', 'Jhargram', 'Jhunjhunu', 'Jind', 'Jod
hpur', 'Jorhat', 'Junagadh', 'Kadapa', 'Kadayanallur', 'Kadiri', 'Kaitha
l', 'Kakinada', 'Kalaburagi', 'Kalady', 'Kalna', 'Kamareddy', 'Kanchrapar
a', 'Kannauj', 'Kannur', 'Kanpur', 'Kanyakumari', 'Kapurthala', 'Karad',
'Karaikkudi', 'Karimnagar', 'Karnal', 'Karunagappaly', 'Karur', 'Karwar',
'Kasaragod', 'Kashipur', 'Katihar', 'Katni', 'Kavali', 'Kayamkulam', 'Kend
rapada', 'Kendujhar', 'Khalilabad', 'Khamgaon', 'Khammam', 'Khandwa', 'Kha
nna', 'Kharagpur', 'Khopoli', 'Kishanganj', 'Kishangarh', 'Kochi', 'Kodaik
anal', 'Kohima', 'Kolar', 'Kolhapur', 'Kolkata', 'Kollam', 'Kopargaon', 'K
oppal', 'Korba', 'Kota', 'Kotdwar', 'Kothagudem', 'Kothamanagalam', 'Kotta
kkal', 'Kottarakkara', 'Kottayam', 'Kovilpatti', 'Kozhikode', 'Krishnagir
i', 'Krishnanagar', 'Kumarakom', 'Kumbakonam', 'Kumta', 'Kundapura', 'Kunn
amkullam', 'Kurnool', 'Kurukshetra', 'Lakhimpur', 'Lalitpur', 'Latur', 'Lo
navala', 'Lonavla', 'Lucknow', 'Ludhiana', 'Machilipatnam', 'Madanapalle',
'Madhubani', 'Madikeri', 'Madurai', 'Mahabaleshwar', 'Mahbubnagar', 'Mahob
a', 'Malappuram', 'Malda', 'Malegaon', 'Malout', 'Manali', 'Mancherial',
'Mandi-dabwali', 'Mandi-gobindgarh', 'Mandi-hp', 'Mandsaur', 'Mandya', 'Ma
ngaluru', 'Manipal', 'Manjeri', 'Mannargudi', 'Mansa', 'Markapur', 'Mathur
```

6)Ratio of veg and non veg restaurants in India

In [14]:

```
category={}
for city in dict.values():
    if len(city)==2:
        for restaurant in city['restaurants'].values():
            if len(restaurant)!=5:
                for menus in restaurant['menu'].values():
                    for menu in menus.values():
                        val=menu['veg or non veg']
                        category[val]=category.get(val,0) +1
            else:
                for subcity in city.values():
                    for restaurant in subcity['restaurants'].values():
                        if len(restaurant)!=5:
                            for menus in restaurant['menu'].values():
                                for menu in menus.values():
                                    val=menu['veg or non veg']
                                    category[val]=category.get(val,0) +1
#NonVeg means even if one of the item in menu is NonVeg and Veg means pure
answer=int(category['Veg']/category['Non veg'])
print(answer,'is the ratio of veg and nonveg restaurants in India.')
```

```
2 is the ratio of veg and nonveg restaurants in India.
```