In [1]:
```
import numpy as np
import pandas as pd

import seaborn as sns
%matplotlib inline
```

In [4]:
```
df=pd.read_csv('zomato.csv',encoding='ISO-8859-1')
df.head()
#df.columns
#df.info()
# df.describe()
# df.isnull().sum()
```

Out[4]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Cer Mal Pob Mal
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Litt Leg Villa Mal
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Eds La, Mar City
3	6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, O...	SM Meg Orti Mar City
4	6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas...	SM Meg Orti Mar City

5 rows × 21 columns

In [15]:
```
df_con=pd.read_excel('Country-Code.xlsx')
df_con.head()
```

Out[15]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Cent Mall, Pobl Maka
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi...	Little Lega Villag Maka
2	6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandal...	Edsa La, O Manc City

3 rows × 22 columns

In [17]:
```
#check data type
Final_df.dtypes
```

Out[17]:
```
Restaurant ID            int64
Restaurant Name         object
Country Code             int64
City                    object
Address                 object
Locality                object
Locality Verbose        object
Longitude              float64
Latitude               float64
Cuisines                object
Average Cost for two     int64
Currency                object
Has Table booking       object
Has Online delivery     object
Is delivering now       object
Switch to order menu    object
Price range              int64
Aggregate rating       float64
Rating color            object
Rating text             object
Votes                    int64
Country                 object
dtype: object
```

In [20]:
```
Final_df.columns
Final_df.Country.value_counts()
```

Out[20]:
```
Country
India             8652
United States      434
United Kingdom      80
Brazil              60
South Africa        60
UAE                 60
New Zealand         40
Turkey              34
Australia           24
Phillipines         22
Indonesia           21
Qatar               20
Singapore           20
Sri Lanka           20
Canada               4
Name: count, dtype: int64
```

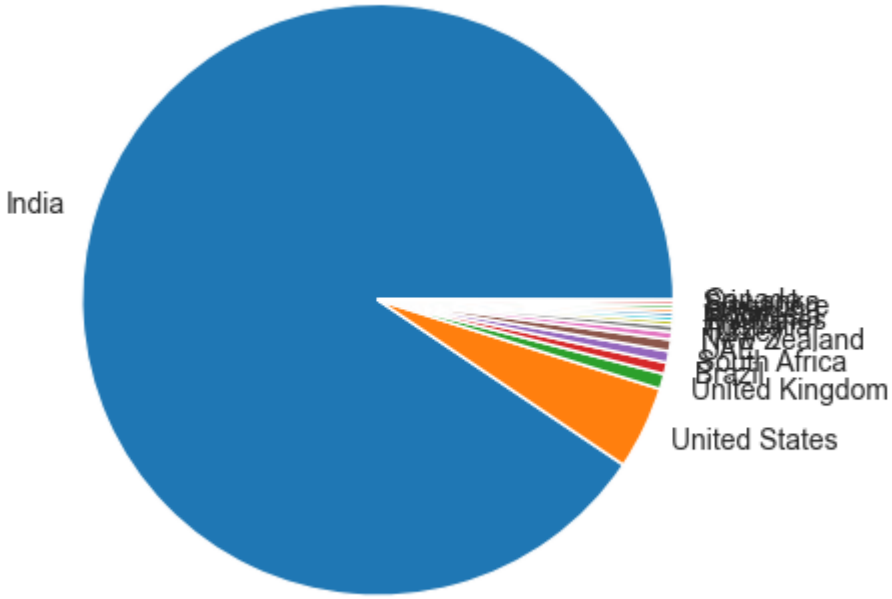
In [ ]:
```

```

In [33]:
```
C_names=Final_df.Country.value_counts().index
C_Values=Final_df.Country.value_counts().values
```

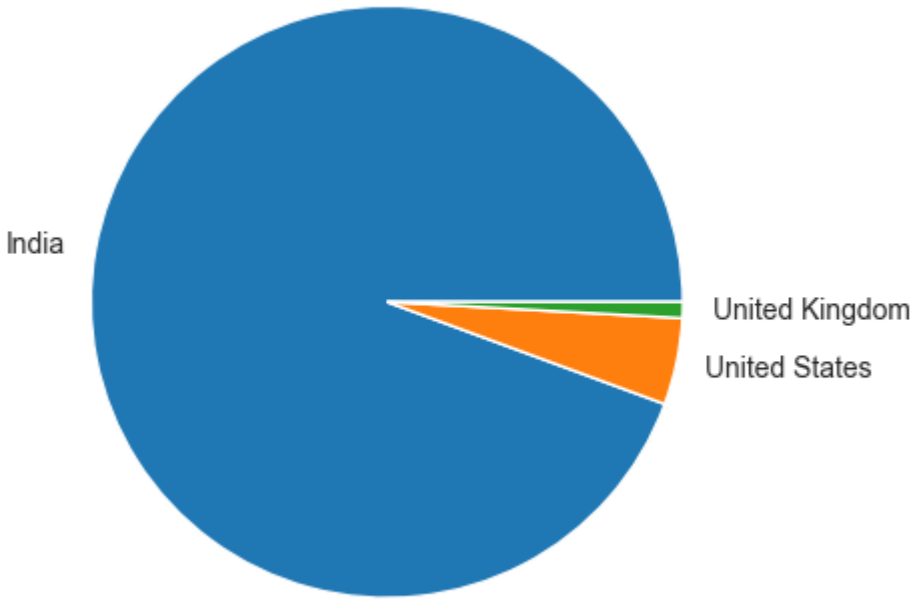
In [35]:
```
#pie chart
import matplotlib.pyplot as plt
plt.pie(C_Values,labels=C_names)
```

Out[35]:
```
([<matplotlib.patches.Wedge at 0x15ec701f450>,
  <matplotlib.patches.Wedge at 0x15ecaa91ad0>,
  <matplotlib.patches.Wedge at 0x15ecacf1e90>,
  <matplotlib.patches.Wedge at 0x15ecacf0e90>,
  <matplotlib.patches.Wedge at 0x15ecacf3190>,
  <matplotlib.patches.Wedge at 0x15ecad303d0>,
  <matplotlib.patches.Wedge at 0x15ecad311d0>,
  <matplotlib.patches.Wedge at 0x15ecad323d0>,
  <matplotlib.patches.Wedge at 0x15ecad33450>,
  <matplotlib.patches.Wedge at 0x15ecad301d0>,
  <matplotlib.patches.Wedge at 0x15eccd41810>,
  <matplotlib.patches.Wedge at 0x15eccd42550>,
  <matplotlib.patches.Wedge at 0x15eccd439d0>,
  <matplotlib.patches.Wedge at 0x15eccd48b90>,
  <matplotlib.patches.Wedge at 0x15eccd49c90>],
 [Text(-1.0522561700723039, 0.32055725314640227, 'India'),
  Text(0.9911331050301793, -0.4771322333622366, 'United States'),
  Text(1.057285832834238, -0.30355669600259505, 'United Kingdom'),
  Text(1.0701388534934553, -0.2545640081467003, 'Brazil'),
  Text(1.0793507650271223, -0.21213657401628266, 'South Africa'),
  Text(1.086881175607599, -0.16937919030932916, 'UAE'),
  Text(1.0918636626261928, -0.13354303514790955, 'New Zealand'),
  Text(1.094790433001231, -0.10692945248610065, 'Turkey'),
  Text(1.0966310607987864, -0.08602509221924036, 'Australia'),
  Text(1.097807123683944, -0.06942275699499531, 'Phillipines'),
  Text(1.098679169865659, -0.053889532409427066, 'Indonesia'),
  Text(1.0993059904136664, -0.03906839439531931, 'Qatar'),
  Text(1.0997248602793597, -0.02460145693943169, 'Singapore'),
  Text(1.0999533477346455, -0.010130785129799997, 'Sri Lanka'),
  Text(1.0999990483203896, -0.001446960344028858, 'Canada')])
```



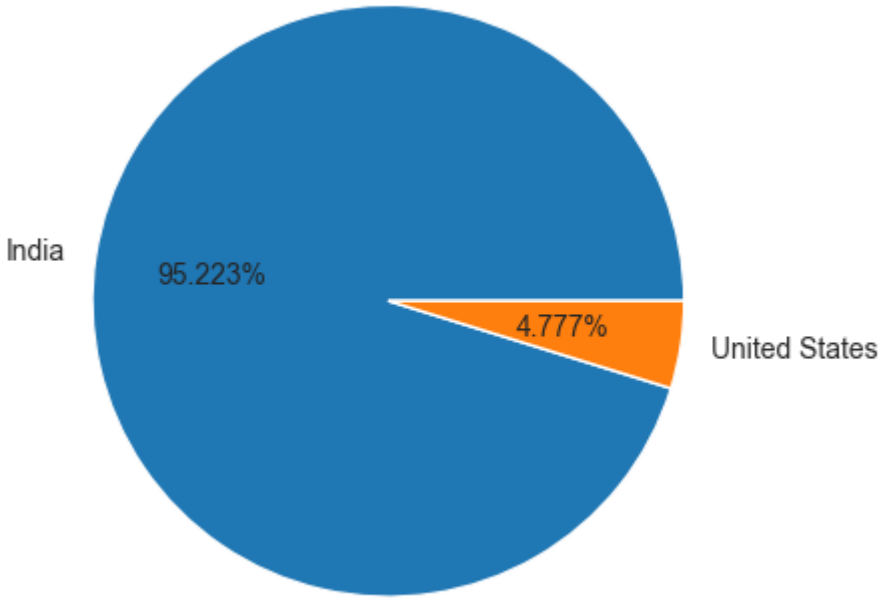
In [36]:
```
plt.pie(C_Values[:3],labels=C_names[:3])
```

Out[36]:
```
([<matplotlib.patches.Wedge at 0x15eccdafad0>,
  <matplotlib.patches.Wedge at 0x15ec8ede9d0>,
  <matplotlib.patches.Wedge at 0x15eca9cb390>],
 [Text(-1.082974277862112, 0.1927867046480056, 'India'),
  Text(1.0772816964394372, -0.22240536530526556, 'United States'),
  Text(1.0995865232164619, -0.030157552300104404, 'United
Kingdom')])
```



In [45]:
```
#percentage
plt.pie(C_Values[:2],labels=C_names[:2],autopct='%1.3f%%')
```

Out[45]:
```
([<matplotlib.patches.Wedge at 0x15ecf3a3150>,
  <matplotlib.patches.Wedge at 0x15ecf3d7010>],
 [Text(-1.0876382277344787, 0.16444782020629697, 'India'),
  Text(1.08763824211096, -0.16444772512193884, 'United States')],
 [Text(-0.5932572151278974, 0.08969881102161653, '95.223%'),
  Text(0.5932572229696146, -0.08969875915742118, '4.777%')])
```



In [47]:
```
Final_df.columns
```

Out[47]:
```
Index(['Restaurant ID', 'Restaurant Name', 'Country Code', 'City',
'Address',
       'Locality', 'Locality Verbose', 'Longitude', 'Latitude',
'Cuisines',
       'Average Cost for two', 'Currency', 'Has Table booking',
       'Has Online delivery', 'Is delivering now', 'Switch to
order menu',
       'Price range', 'Aggregate rating', 'Rating color', 'Rating
text',
       'Votes', 'Country'],
      dtype='object')
```

In [52]:
```
Final_df.groupby(['Aggregate rating','Rating color','Rating
text']).size().reset_index()
```

Out[52]:

	Aggregate rating	Rating color	Rating text	0
0	0.0	White	Not rated	2148
1	1.8	Red	Poor	1
2	1.9	Red	Poor	2
3	2.0	Red	Poor	7
4	2.1	Red	Poor	15
5	2.2	Red	Poor	27
6	2.3	Red	Poor	47
7	2.4	Red	Poor	87
8	2.5	Orange	Average	110
9	2.6	Orange	Average	191
10	2.7	Orange	Average	250
11	2.8	Orange	Average	315
12	2.9	Orange	Average	381
13	3.0	Orange	Average	468
14	3.1	Orange	Average	519
15	3.2	Orange	Average	522
16	3.3	Orange	Average	483
17	3.4	Orange	Average	498
18	3.5	Yellow	Good	480
19	3.6	Yellow	Good	458
20	3.7	Yellow	Good	427
21	3.8	Yellow	Good	400
22	3.9	Yellow	Good	335
23	4.0	Green	Very Good	266
24	4.1	Green	Very Good	274
25	4.2	Green	Very Good	221
26	4.3	Green	Very Good	174
27	4.4	Green	Very Good	144
28	4.5	Dark Green	Excellent	95

	Aggregate rating	Rating color	Rating text	0
29	4.6	Dark Green	Excellent	78
30	4.7	Dark Green	Excellent	42
31	4.8	Dark Green	Excellent	25
32	4.9	Dark Green	Excellent	61

In [ ]:

In [2]:
```
import sklearn
import nltk
import seaborn

print("All libraries installed successfully")
```

```
All libraries installed successfully
```