In [48]:
```
######################## CITY VERSUS SUBSCRIPTION RATE ####################

import pandas as pd
import numpy as num
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]:
```
path = "/Users/yarkingazi/Desktop/dognition_data_no_aggregation_with_zip_co
data = pd.read_excel(path)
print(data)
```

```
                Created_at          Updated_at  \
0      2013-05-13 16:17:00 2013-05-13 16:17:00
1      2013-05-22 18:54:00 2013-05-22 18:54:00
2      2013-05-22 18:55:00 2013-05-22 18:55:00
3      2013-05-22 18:56:00 2013-05-22 18:56:00
4      2013-05-22 18:58:00 2013-05-22 18:58:00
...                    ...                 ...
177476 2013-05-14 08:45:00 2013-05-14 08:45:00
177477 2013-05-14 08:50:00 2013-05-14 08:50:00
177478 2013-05-14 08:54:00 2013-05-14 08:54:00
177479 2013-05-19 04:45:00 2013-05-19 04:45:00
177480 2013-05-19 05:05:00 2013-05-19 05:05:00

                                      User_ID  \
0       ce134a78-7144-11e5-ba71-058fbc01cf0b
1       ce134a78-7144-11e5-ba71-058fbc01cf0b
2       ce134a78-7144-11e5-ba71-058fbc01cf0b
3       ce134a78-7144-11e5-ba71-058fbc01cf0b
4       ce134a78-7144-11e5-ba71-058fbc01cf0b
...                                      ...
177476  ce2777e6-7144-11e5-ba71-058fbc01cf0b
177477  ce2777e6-7144-11e5-ba71-058fbc01cf0b
177478  ce2777e6-7144-11e5-ba71-058fbc01cf0b
177479  ce2777e6-7144-11e5-ba71-058fbc01cf0b
177480  ce2777e6-7144-11e5-ba71-058fbc01cf0b

                                       Dog_ID               Test_name  \
0       fd3d1b44-7144-11e5-ba71-058fbc01cf0b           Yawn Warm-up
1       fd3d1b44-7144-11e5-ba71-058fbc01cf0b              Yawn Game
2       fd3d1b44-7144-11e5-ba71-058fbc01cf0b    Eye Contact Warm-up
3       fd3d1b44-7144-11e5-ba71-058fbc01cf0b       Eye Contact Game
4       fd3d1b44-7144-11e5-ba71-058fbc01cf0b          Treat Warm-up
...                                      ...                    ...
177476  fd444202-7144-11e5-ba71-058fbc01cf0b        One Cup Warm-up
177477  fd444202-7144-11e5-ba71-058fbc01cf0b        Two Cup Warm-up
177478  fd444202-7144-11e5-ba71-058fbc01cf0b  Memory versus Pointing
177479  fd444202-7144-11e5-ba71-058fbc01cf0b    Memory versus Smell
177480  fd444202-7144-11e5-ba71-058fbc01cf0b       Delayed Cup Game

       Subcategory_name Gender  Birthday     Breed  \
0               Empathy   male      2009  Shih Tzu
1               Empathy   male      2009  Shih Tzu
2               Empathy   male      2009  Shih Tzu
3               Empathy   male      2009  Shih Tzu
4         Communication   male      2009  Shih Tzu
...                 ...    ...       ...       ...
177476           Memory   male      2012     Mixed
177477           Memory   male      2012     Mixed
177478           Memory   male      2012     Mixed
177479           Memory   male      2012     Mixed
177480           Memory   male      2012     Mixed

                                  Breed_Type  ... State     Zip  Country  Exc
```

```
lude  \
0                              Pure Breed  ...    NC   27701       US
1
1                              Pure Breed  ...    NC   27701       US
1
2                              Pure Breed  ...    NC   27701       US
1
3                              Pure Breed  ...    NC   27701       US
1
4                              Pure Breed  ...    NC   27701       US
1
...                                   ...  ...   ...     ...      ...
...
177476  Mixed Breed/ Other/ I Don't Know  ...   NaN  466311       SG
NaN
177477  Mixed Breed/ Other/ I Don't Know  ...   NaN  466311       SG
NaN
177478  Mixed Breed/ Other/ I Don't Know  ...   NaN  466311       SG
NaN
177479  Mixed Breed/ Other/ I Don't Know  ...   NaN  466311       SG
NaN
177480  Mixed Breed/ Other/ I Don't Know  ...   NaN  466311       SG
NaN

       Free_Start_User       Last_Active_At  Membership_Type  Rating  \
0                    0  2015-02-23 13:39:00                2     NaN
1                    0  2015-02-23 13:39:00                2     NaN
2                    0  2015-02-23 13:39:00                2     NaN
3                    0  2015-02-23 13:39:00                2     NaN
4                    0  2015-02-23 13:39:00                2     NaN
...                ...                  ...              ...     ...
177476             NaN  2014-04-28 01:18:00                1     NaN
177477             NaN  2014-04-28 01:18:00                1     NaN
177478             NaN  2014-04-28 01:18:00                1     NaN
177479             NaN  2014-04-28 01:18:00                1     NaN
177480             NaN  2014-04-28 01:18:00                1     NaN

        Rank_by_UserID Rank_by_DogID
0                  1.0           1.0
1                  2.0           2.0
2                  3.0           3.0
3                  4.0           4.0
4                  5.0           5.0
...                ...           ...
177476            12.0          12.0
177477            13.0          13.0
177478            14.0          14.0
177479            15.0          15.0
177480            16.0          16.0

[177481 rows x 30 columns]
```

In [3]:
```
uniqueValues = data['City'].nunique()
print('Number of unique values in column "City" of the dataframe : ')
print(uniqueValues)
##I wanted to have a look at the number of the cities among all.
```

```
Number of unique values in column "City" of the dataframe :
3398
```

In [4]:
```
dataframe0 = pd.read_excel("/Users/yarkingazi/Desktop/dognition_data_no_agg
dataframe1 = pd.DataFrame(dataframe0)
print(dataframe0['City'])
```

```
0             Durham
1             Durham
2             Durham
3             Durham
4             Durham
              ...
177476     Singapore
177477     Singapore
177478     Singapore
177479     Singapore
177480     Singapore
Name: City, Length: 177481, dtype: object
```

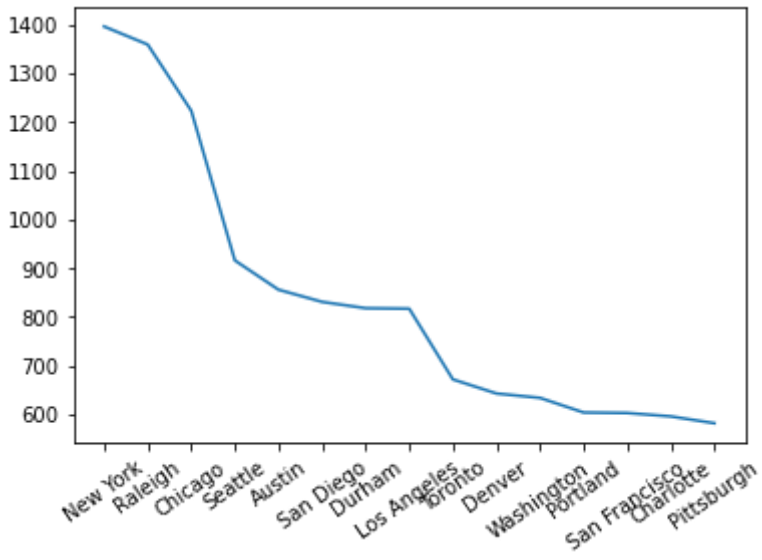
In [43]:
```
dataframe2 = dataframe1.groupby('City')['Subscribed'].sum().nlargest(15,'fi

print(dataframe2)
```

```
City
New York         1396
Raleigh          1359
Chicago          1223
Seattle           916
Austin            856
San Diego         831
Durham            818
Los Angeles       817
Toronto           672
Denver            643
Washington        634
Portland          604
San Francisco     603
Charlotte         596
Pittsburgh        582
Name: Subscribed, dtype: int64
```

In [52]:
```
plt.plot(dataframe2)
plt.xticks(rotation=35)
plt.show()

###Below figure says that most subscribers are from New York. And the most
###are the cities can be seen in the x-axis.
```




In [ ]: