In [1]:

```
import numpy as np
import pandas as pd
```

In [2]:

```
from datetime import datetime

##datetime_object = datetime.strptime('2018-10-15 20:59:29', '%Y-%m-%d %H:%M:%S')
```

In [3]:

```
import glob
import pandas as pd

# get data file names
path =r'C:\Users\LEELA SURYA TEJA\Desktop\syncs2'
filenames = glob.glob(path + "/*.csv")

dfs = []
for filename in filenames:
    dfs.append(pd.read_csv(filename))

# Concatenate all data into one DataFrame
big_frame = pd.concat(dfs, ignore_index=True)
```

ALL NAMES OF SYNCHRONY IS STORED IN THE filename with indices

In [4]:

```
filenames[2].split('#')[1]
```

Out[4]:

```
'bestfanarmy.csv'
```

In [5]:

```
for i in range(len(dfs)):
    dfs[i]['date'] =  pd.to_datetime(dfs[i]['date'], format='%Y-%m-%d %H:%M:%S')
```

In [6]:

```
 dfs[1]['date'].dtype
```

Out[6]:

```
dtype('<M8[ns]')
```

In [75]:

```
y=[]
for j in range(len(dfs)):
    y.append('#'+filenames[j].split('#')[1].split('.')[0])

synchrony_name=pd.DataFrame(y,columns=['name of synchrony'])
```

In [8]:

```
no_tweet=dfs[1].groupby('date',as_index=False).count()
```

In [9]:

```
no_tweet.head(2)
```

Out[9]:

	date	favorites	retweet	tweet	tweet_id	user_id
0	2018-02-15 01:08:36	1	1	1	1	1
1	2018-02-15 04:06:43	1	1	1	1	1

In [10]:

```
for i in range(len(dfs)):
    dfs[i]['Dates'] = pd.to_datetime(dfs[i]['date']).dt.date
    dfs[i]['Time'] = pd.to_datetime(dfs[i]['date']).dt.time
```

In [11]:

```
dfs[2].head(1)
```

Out[11]:

	date	favorites	retweet	tweet	tweet_id	user_id	Dates	Time
0	2018-02-21 01:19:19	0	0	Prince Myeon . #iHeartAwards #BestFanArmy #EXO...	966037066726375424	926792375522832384	2018-02-21	01:19:19

In [12]:

```
no_tweet=[]

for i in range(len(dfs)):
    temp=dfs[i].groupby('Dates',as_index=False).count()
    no_tweet.append(temp)
```

In [13]:

```
#no_tweet is a dataframe grouped by date

a=[]
for i in range(len(no_tweet)):
    a.append(no_tweet[i][['Dates','tweet']])
```

In [14]:

```
len(a)
#len(no_tweet)
a[1]['Dates'].count()
```

Out[14]:

```
4
```

In [15]:

```
a[200].head()
```

Out[15]:

	Dates	tweet
0	2018-02-14	1
1	2018-02-15	4
2	2018-02-17	3

	Dates	tweet

In [16]:

```
import seaborn as sns
```

In [17]:

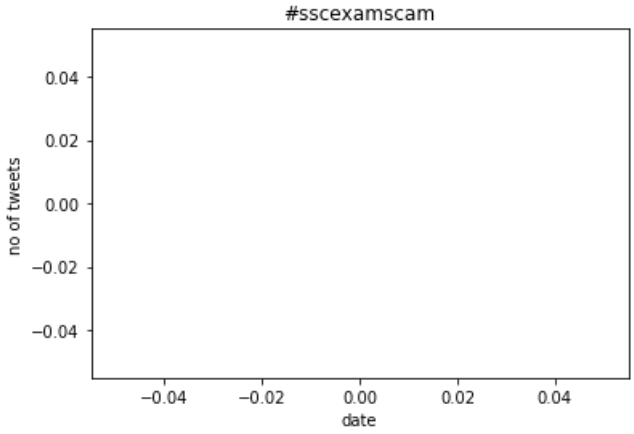
```
import matplotlib.dates as date
import matplotlib.pyplot as plt
```

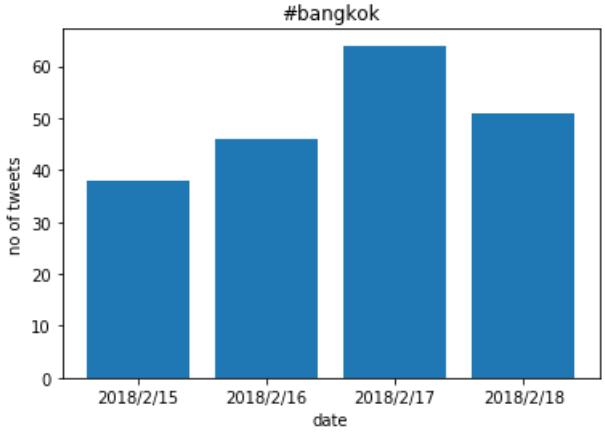
analysing no of unique tweets in each synchrony

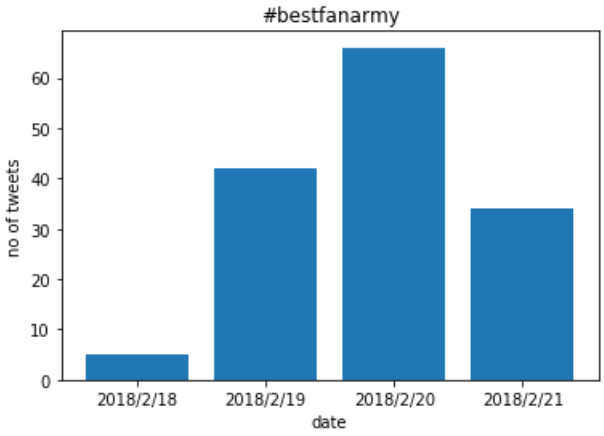
In [18]:

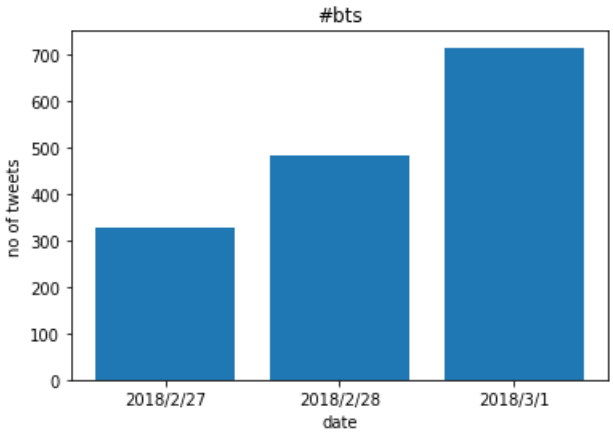
```
for j in range(25):
    v=a[j]['tweet']
    d=[]
    for i in range(a[j]['Dates'].count()):
        d.append(str(a[j]['Dates'][i].year) +'/' + str(a[j]['Dates'][i].month) +'/' + str(a[j]
['Dates'][i].day))
    plt.figure()
    plt.bar(d,v)
    plt.xlabel('date')
    plt.ylabel('no of tweets')
    plt.title('#'+filenames[j].split('#')[1].split('.')[0])
```

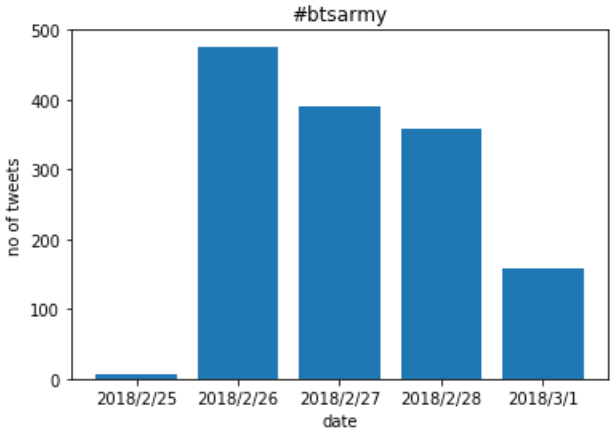
```
L:\ANACONDA\lib\site-packages\matplotlib\pyplot.py:537: RuntimeWarning: More than 20 figures have
been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are
retained until explicitly closed and may consume too much memory. (To control this warning, see th
e rcParam `figure.max_open_warning`).
  max_open_warning, RuntimeWarning)
```

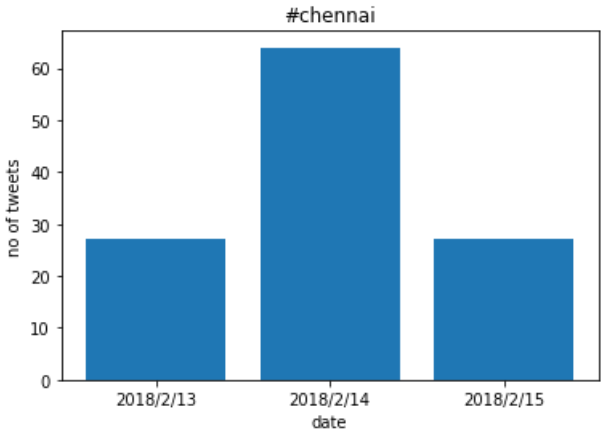


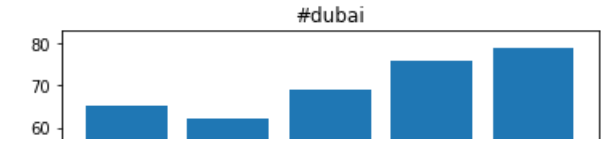


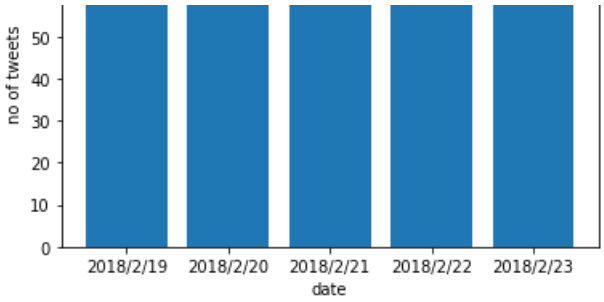

#bestfanarmy
no of tweets
0
10
20
30
40
50
60
2018/2/18
2018/2/19
2018/2/20
2018/2/21
date

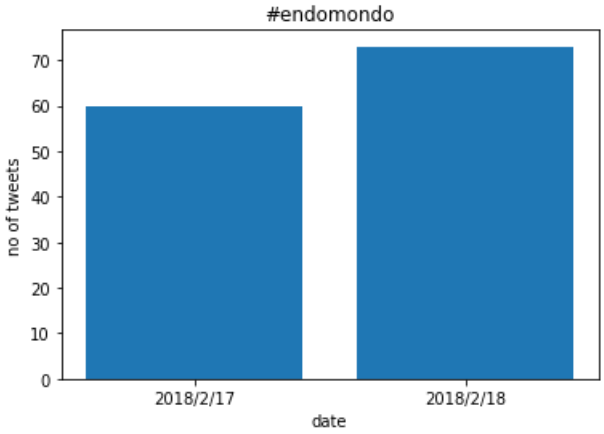

#bts
no of tweets
0
100
200
300
400
500
600
700
2018/2/27
2018/2/28
2018/3/1
date

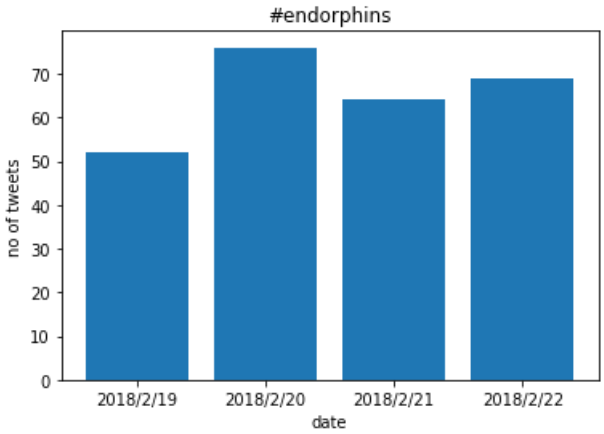

#btsarmy
no of tweets
0
100
200
300
400
500
2018/2/25
2018/2/26
2018/2/27
2018/2/28
2018/3/1
date

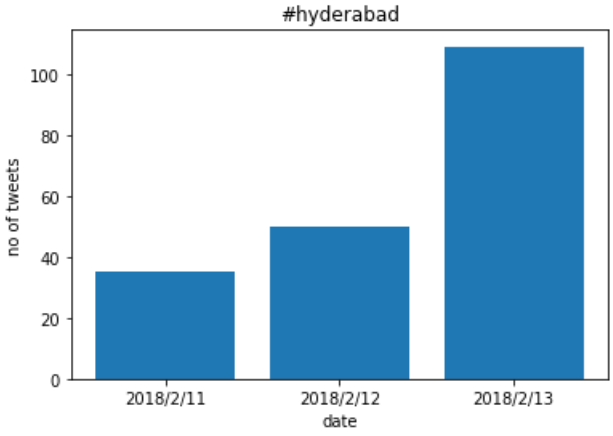

#chennai
no of tweets
0
10
20
30
40
50
60
2018/2/13
2018/2/14
2018/2/15
date

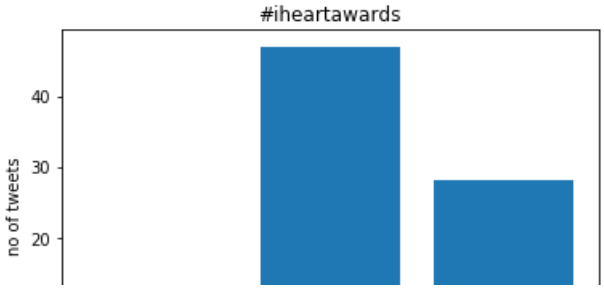

#dubai
60
70
80

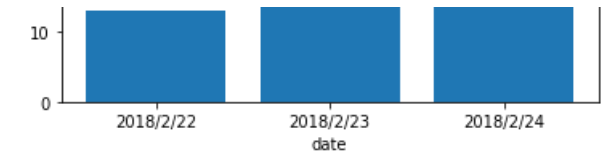

no of tweets
0
10
20
30
40
50
2018/2/19
2018/2/20
2018/2/21
2018/2/22
2018/2/23
date

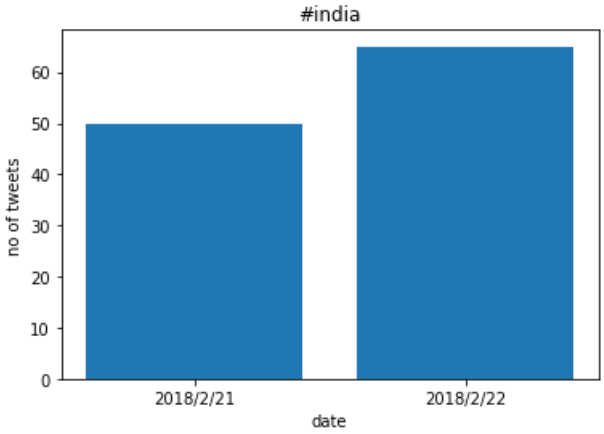

#endomondo
no of tweets
0
10
20
30
40
50
60
70
2018/2/17
2018/2/18
date

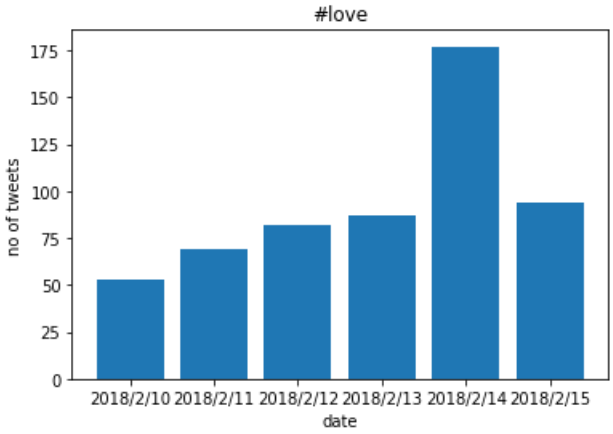

#endorphins
no of tweets
0
10
20
30
40
50
60
70
2018/2/19
2018/2/20
2018/2/21
2018/2/22
date

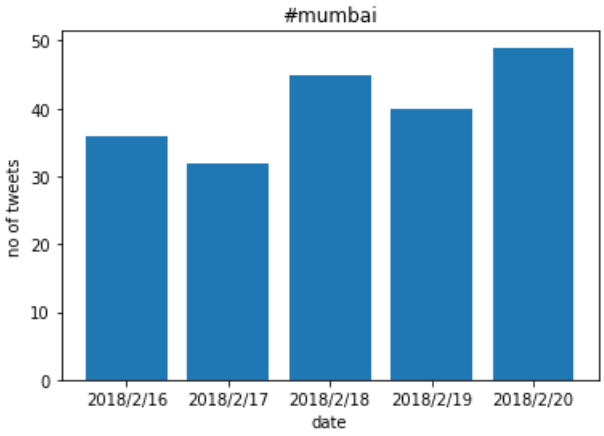

#hyderabad
no of tweets
0
20
40
60
80
100
2018/2/11
2018/2/12
2018/2/13
date

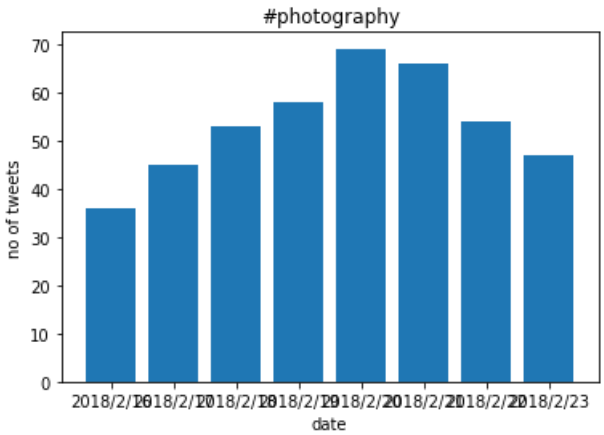

#iheartawards
no of tweets
20
30
40

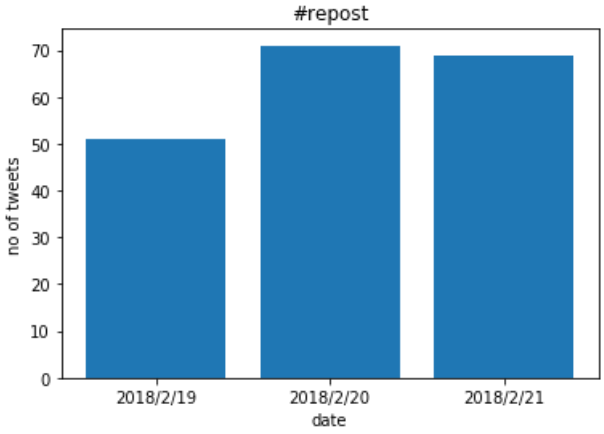

10
0
2018/2/22
2018/2/23
2018/2/24
date

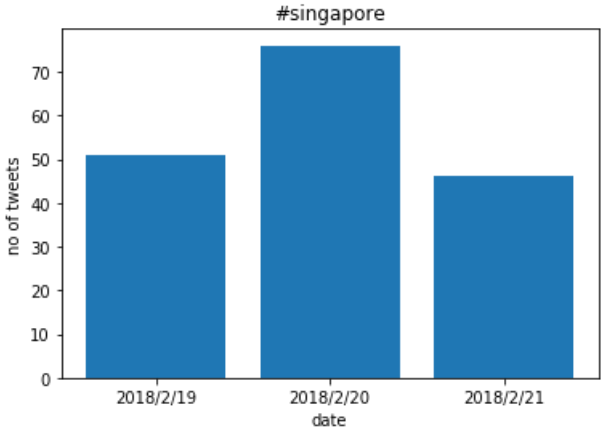

#india
no of tweets
60
50
40
30
20
10
0
2018/2/21
2018/2/22
date

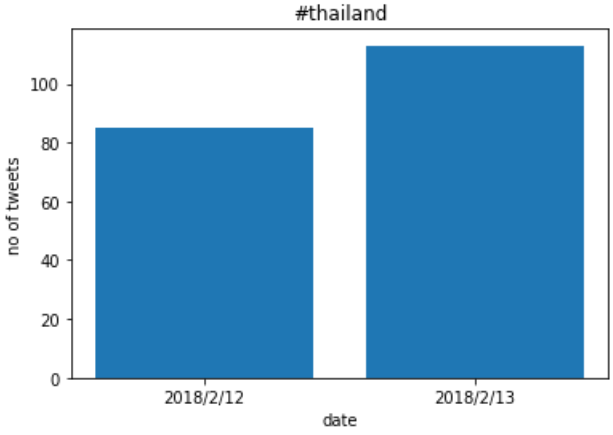

#love
no of tweets
175
150
125
100
75
50
25
0
2018/2/10 2018/2/11 2018/2/12 2018/2/13 2018/2/14 2018/2/15
date

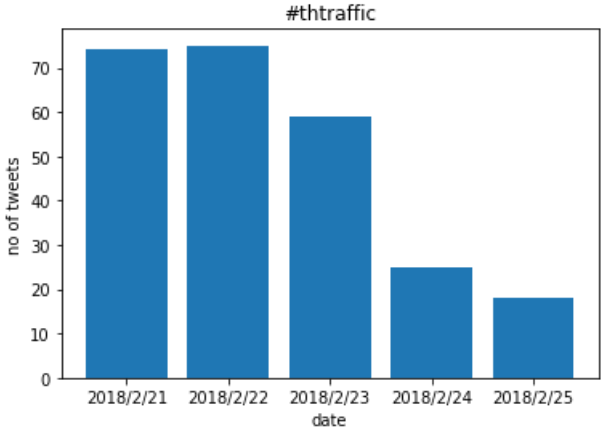

#mumbai
no of tweets
50
40
30
20
10
0
2018/2/16
2018/2/17
2018/2/18
2018/2/19
2018/2/20
date

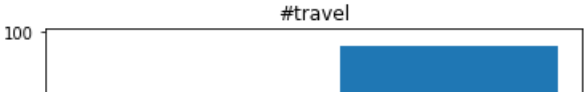

#photography
no of tweets
70
60
50
40
30
20
10
0
date

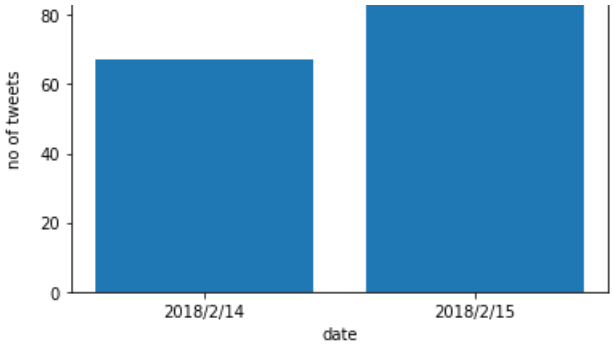

#repost
no of tweets
0
10
20
30
40
50
60
70
2018/2/19
2018/2/20
2018/2/21
date

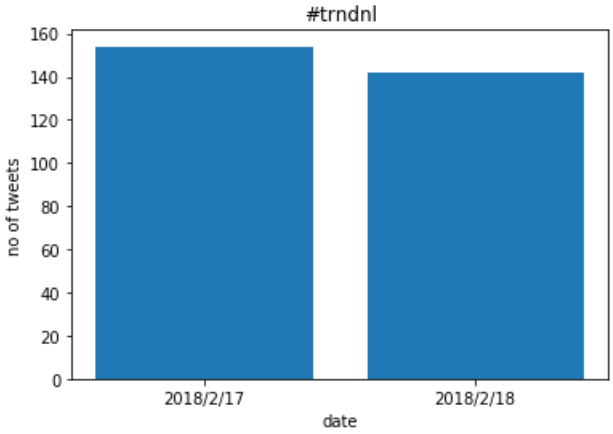

#singapore
no of tweets
0
10
20
30
40
50
60
70
2018/2/19
2018/2/20
2018/2/21
date

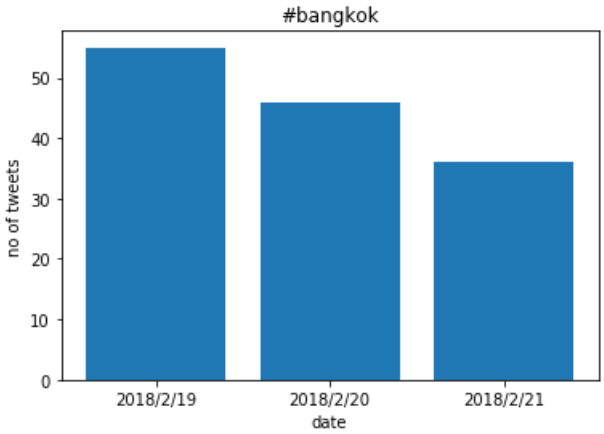

#thailand
no of tweets
0
20
40
60
80
100
2018/2/12
2018/2/13
date

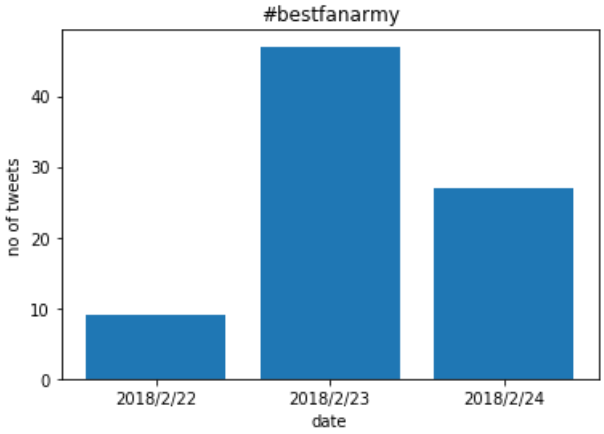

#thtraffic
no of tweets
0
10
20
30
40
50
60
70
2018/2/21
2018/2/22
2018/2/23
2018/2/24
2018/2/25
date

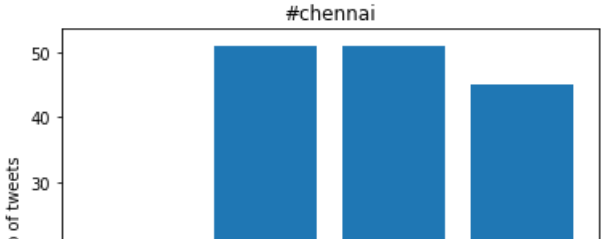

#travel
100

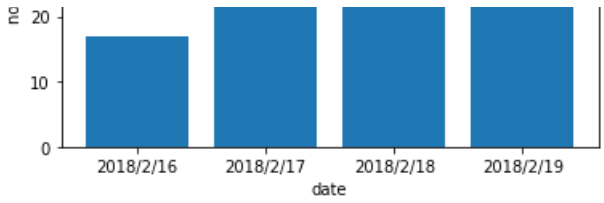

80
60
40
20
0
no of tweets
2018/2/14
2018/2/15
date

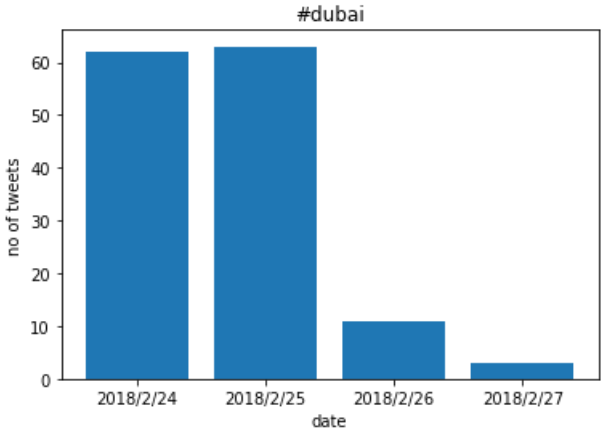

#trndnl
160
140
120
100
80
60
40
20
0
no of tweets
2018/2/17
2018/2/18
date


#bangkok
50
40
30
20
10
0
no of tweets
2018/2/19
2018/2/20
2018/2/21
date


#bestfanarmy
40
30
20
10
0
no of tweets
2018/2/22
2018/2/23
2018/2/24
date


#chennai
50
40
30
of tweets



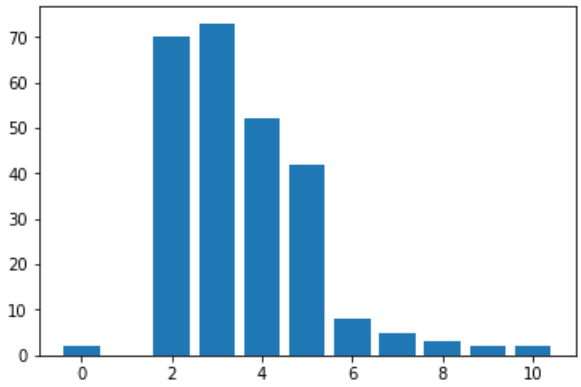


In [19]:

```
X=[]
for i in range(len(a)):
    X.append(a[i]['Dates'].count())
dframe=pd.DataFrame(X,columns=['number_of_days'])
#X=dframe.groupby('abc',as_index=False).count()
#X
x=dframe['number_of_days'].value_counts().keys().tolist()
y=dframe['number_of_days'].value_counts().tolist()
plt.bar(x,y)
```

Out[19]:

```
<BarContainer object of 10 artists>
```



In [20]:

```
dframe['number_of_syncs']=filename
x=dframe.groupby('number_of_days',as_index=False).count()
```

In [21]:

```
dframe.head()
```

Out[21]:

	number_of_days	number_of_syncs
0	0	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...

	number_of_days	number_of_syncs
1	1	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync8...
2	4	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
3	3	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
4	5	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...

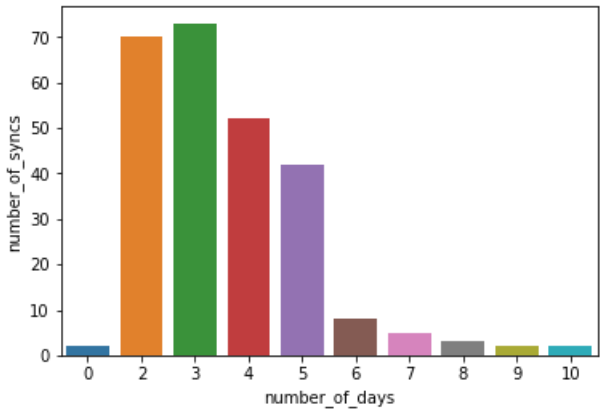
below graph represents the frequency of no days synchronies occured

In [22]:

```
sns.barplot('number_of_days','number_of_syncs',data=x,label='length of synchronies in terms of days')
```

Out[22]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x2497f86a0b8>
```



In [23]:

```
big_frame.head(5)
```

Out[23]:

	date	favorites	retweet	tweet	tweet_id	user_id
0	2018-02-15 06:54:15	0	0	Thailand. Next up Cambodia. #chaingmai #krabi ...	9.639470e+17	23811362.0
1	2018-02-15 22:02:53	0	0	For the last evening in #vibrant and #bubling ...	9.641757e+17	112967218.0
2	2018-02-15 04:20:42	0	0	Rooftops #Bangkok @Octave Rooftop Lounge & Bar...	9.639084e+17	759256634.0
3	2018-02-15 23:05:56	0	0	Red dragons #dragons #dragones #chinesenewyear...	9.641916e+17	940683308.0
4	2018-02-15 20:51:41	1	0	Dinner à¹à¸ nahm #à¸à¸²à¸¡à¸à¸´à¸£à¹à¸...	9.641578e+17	90110020.0

In [24]:

```
big_frame['Dates'] = pd.to_datetime(big_frame['date']).dt.date
big_frame['Time'] = pd.to_datetime(big_frame['date']).dt.time
```

In [25]:

```
big_frame.head(4)
```

Out[25]:

	date	favorites	retweet	tweet	tweet_id	user_id	Dates	Time
0	2018-02-15 06:54:15	0	0	Thailand. Next up Cambodia. #chaingmai #krabi ...	9.639470e+17	23811362.0	2018-02-15	06:54:15

	date	favorites	retweet	tweet	tweet_id	user_id	Dates	Time
1	2018-02-15 22:02:53	0	0	For the last evening in #vibrant and #bubling ...	9.641757e+17	112967218.0	2018-02-15	22:02:53
2	2018-02-15 04:20:42	0	0	Rooftops #Bangkok @Octave Rooftop Lounge & Bar...	9.639084e+17	759256634.0	2018-02-15	04:20:42
3	2018-02-15 23:05:56	0	0	Red dragons #dragons #dragones #chinesenewyear...	9.641916e+17	940683308.0	2018-02-15	23:05:56

In [26]:

```
big_frame['date'] =  pd.to_datetime(big_frame['date'], format='%Y-%m-%d %H:%M:%S')
```

In [27]:

```
big_frame['hour']=big_frame['date'].dt.hour
```

In [28]:

```
big_frame.head(1)
```

Out[28]:

	date	favorites	retweet	tweet	tweet_id	user_id	Dates	Time	hour
0	2018-02-15 06:54:15	0	0	Thailand. Next up Cambodia. #chaingmai #krabi ...	9.639470e+17	23811362.0	2018-02-15	06:54:15	6

In [29]:

```
t=big_frame.groupby('hour',as_index=False).count()
```

In [30]:

```
t.head()
```

Out[30]:

	hour	date	favorites	retweet	tweet	tweet_id	user_id	Dates	Time
0	0	2812	2812	2812	2812	2812	2812	2812	2812
1	1	2213	2213	2213	2213	2213	2213	2213	2213
2	2	2280	2280	2280	2280	2280	2280	2280	2280
3	3	1945	1945	1945	1945	1945	1945	1945	1945
4	4	2017	2017	2017	2017	2017	2017	2017	2017

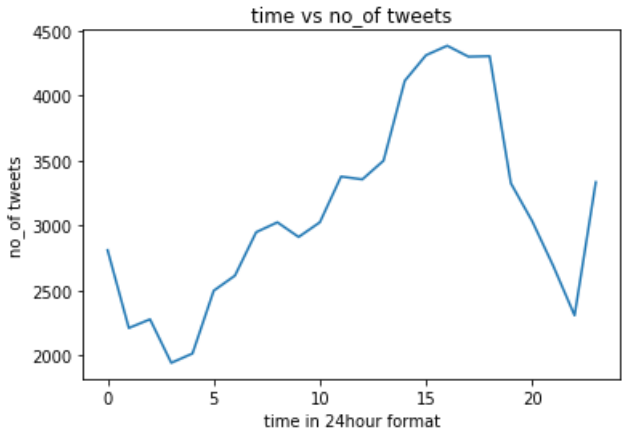
Below graph represents the total no of tweets occuring at a particular time considering all synchronies

In [31]:

```
x=t['hour']
y=t['tweet']
plt.plot(x,y)
plt.title('time vs no_of tweets')
plt.xlabel('time in 24hour format')
plt.ylabel('no_of tweets')
```

Out[31]:

```
Text(0,0.5,'no_of tweets')
```



In [138]:

```
a=[]
for i in range(len(dfs)):
    temp=dfs[i]['user_id'].count()
    a.append(temp)
```

In [139]:

```
b=range(1,260)
```

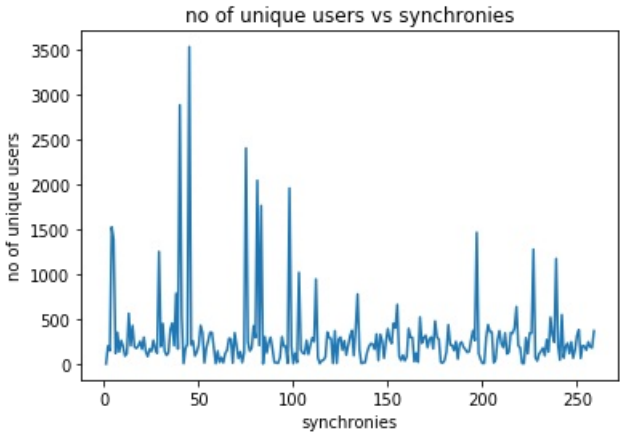
below graph represents total number of unique users in each synchrony

In [140]:

```
plt.plot(b,a)
plt.xlabel('synchronies')
plt.ylabel('no of unique users')
plt.title('no of unique users vs synchronies')
```

Out[140]:

```
Text(0.5,1,'no of unique users vs synchronies')
```



In [141]:

```
a=np.array(a)
```

In [145]:

```
from scipy import stats
import numpy as np
z = np.abs(stats.zscore(a))

threshold = 2
```

```
ar=(np.where(z > 2))

j=[]
for i in ar:
    j=i
```

below hashtags represents the outliers in synchronies in terms of total number of unique users

In [146]:

```
for i in j:
    print(synchrony_name['name of synchrony'][i])
```

```
#bts
#btsarmy
#iheartawards
#bestfanarmy
#iheartawards
#kca
#savesyrianchildren
#spiritualleadersaintrampalji
#favpinoynewbieinigo
#trndnl
#exol
#trndnl
```

In [35]:

```
big_frame.head()
```

Out[35]:

	date	favorites	retweet	tweet	tweet_id	user_id	Dates	Time	hour
0	2018-02-15 06:54:15	0	0	Thailand. Next up Cambodia. #chaingmai #krabi ...	9.639470e+17	23811362.0	2018-02-15	06:54:15	6
1	2018-02-15 22:02:53	0	0	For the last evening in #vibrant and #bubling ...	9.641757e+17	112967218.0	2018-02-15	22:02:53	22
2	2018-02-15 04:20:42	0	0	Rooftops #Bangkok @Octave Rooftop Lounge & Bar...	9.639084e+17	759256634.0	2018-02-15	04:20:42	4
3	2018-02-15 23:05:56	0	0	Red dragons #dragons #dragones #chinesenewyear...	9.641916e+17	940683308.0	2018-02-15	23:05:56	23
4	2018-02-15 20:51:41	1	0	Dinner à¹à¸ nahm #à¸à¸²à¸¡à¸ ´à¸¡à¸£à¸²à¸...	9.641578e+17	90110020.0	2018-02-15	20:51:41	20

In [36]:

```
c=[]
for i in range(len(dfs)):
    c.append(dfs[i]['retweet'].sum())
```

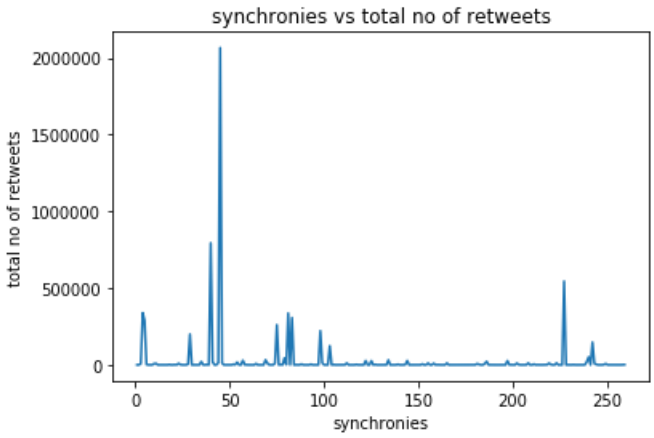
below graph represents total no of retweets in each synchrony

In [37]:

```
plt.plot(b,c)
plt.xlabel('synchronies')
plt.ylabel('total no of retweets')
plt.title("synchronies vs total no of retweets")
```

Out[37]:

```
Text(0.5,1,'synchronies vs total no of retweets')
```



In [89]:

```
from scipy import stats
import numpy as np
z = np.abs(stats.zscore(c))

threshold = 2
ar=(np.where(z > 2))

j=[]
for i in ar:
    j=i
```

below hashtags represents the outliers in synchronies in total number of retweets

In [91]:

```
for i in j:
    print(synchrony_name['name of synchrony'][i])
```

```
#bts
#bestfanarmy
#iheartawards
#savesyrianchildren
#exol
```

In [38]:

```
d=[]
for i in range(len(dfs)):
    d.append(sum(dfs[i]['retweet']!=0))
```

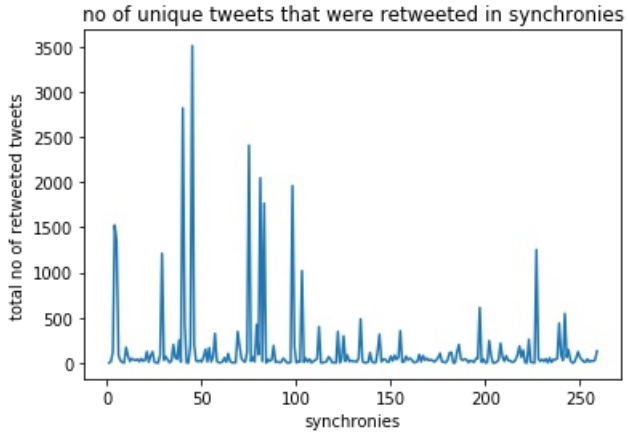
below graph represents total no of unique tweets that were retweeted in each synchrony

In [39]:

```
plt.plot(b,d)
plt.xlabel('synchronies')
plt.ylabel('total no of retweeted tweets')
plt.title("no of unique tweets that were retweeted in synchronies")
```

Out[39]:

```
Text(0.5,1,'no of unique tweets that were retweeted in synchronies')
```



In [92]:

```
from scipy import stats
import numpy as np
z = np.abs(stats.zscore(d))

threshold = 2
ar=(np.where(z > 2))

j=[]
for i in ar:
    j=i
```

below hashtags represents the outliers in synchronies in number of unique tweets that were retweeted

In [93]:

```
for i in j:
    print(synchrony_name['name of synchrony'][i])
```

```
#bts
#btsarmy
#iheartawards
#bestfanarmy
#iheartawards
#kca
#savesyrianchildren
#spiritualleadersaintrampalji
#favpinoynewbieinigo
#kaalateaser
#exol
```

In [40]:

```
r=[]
for i in range(len(dfs)):
    r.append(dfs[i]['tweet'].count())
```

In [41]:

```
c=pd.Series(c)
r=pd.Series(r)
```

In [42]:

```
c=np.array(c)
r=np.array(r)
```

In [43]:

```
ratio=c/r
```

```
L:\ANACONDA\lib\site-packages\ipykernel_launcher.py:1: RuntimeWarning: invalid value encountered i
n true_divide
  """Entry point for launching an IPython kernel.
```

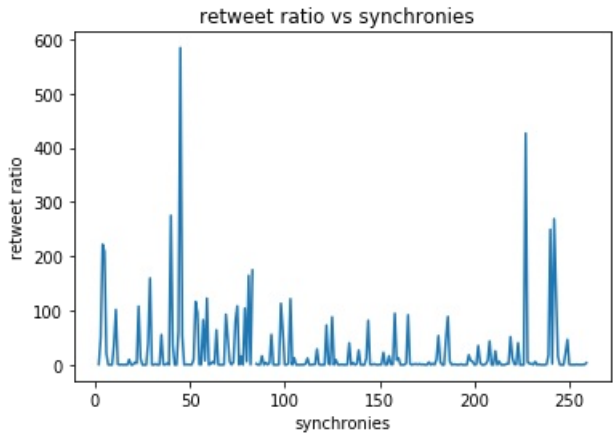
below graph represents the retweet ratio in each synchrony.

In [44]:

```
plt.plot(b,ratio)
plt.xlabel('synchronies')
plt.ylabel('retweet ratio')
plt.title('retweet ratio vs synchronies  ')
```

Out[44]:

```
Text(0.5,1,'retweet ratio vs synchronies  ')
```



In [125]:

```
mean=np.mean(ratio)
sd =np.std(ratio)

rt=np.nan_to_num(rt, copy=True)
```

In [127]:

```
from scipy import stats
import numpy as np
z = np.abs(stats.zscore(rt))

threshold = 2
ar=(np.where(z > 2))

j=[]
for i in ar:
    j=i
```

below hashtags represents the outliers in synchronies in retweet ratio

In [128]:

```
for i in j:
    print(synchrony_name['name of synchrony'][i])
```

```
#bts
#btsarmy
```

```
#iheartawards
#bestfanarmy
#iheartawards
#savesyrianchildren
#spiritualleadersaintrampalji
#exol
#bangkok
#bts
```

In [45]:

```
X=[]
for i in range(len(dfs)):
    X.append(dfs[i]['user_id'].count())
dframe=pd.DataFrame(X,columns=['number_of_days'])
#X=dframe.groupby('abc',as_index=False).count()
#X
#x=dframe['number_of_days'].value_counts().keys().tolist()
#y=dframe['number_of_days'].value_counts().tolist()
#plt.bar(x,y)
```

In [46]:

```
dframe.describe()
```

Out[46]:

	number_of_days
count	259.000000
mean	288.189189
std	415.142955
min	0.000000
25%	113.000000
50%	198.000000
75%	300.500000
max	3538.000000

In [47]:

```
dframe['synchrony']=filename
```

In [48]:

```
dframe.head()
```

Out[48]:

	number_of_days	synchrony
0	0	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
1	199	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
2	147	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
3	1526	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
4	1389	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...

In [49]:

```
dframe=dframe.sort_values('number_of_days',axis=0, ascending=True, inplace=False, kind='quicksort',
na_position='last')
```

In [50]:

```
dframe.head()
```

Out[50]:

	number_of_days	synchrony
0	0	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
83	0	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
58	4	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
99	4	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
113	4	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...

In [51]:

```
#(dframe['number_of_days']>0) & (dframe['number_of_days']<500)

filter=(dframe['number_of_days']>0) & (dframe['number_of_days']<=500)
dframe[filter]=500

filter=(dframe['number_of_days']>500) & (dframe['number_of_days']<=1000)
dframe[filter]=1000

filter=(dframe['number_of_days']>1000) & (dframe['number_of_days']<=1500)
dframe[filter]=1500

filter=(dframe['number_of_days']>1500) & (dframe['number_of_days']<=2000)
dframe[filter]=2000

filter=(dframe['number_of_days']>2000) & (dframe['number_of_days']<=2500)
dframe[filter]=2500


filter=(dframe['number_of_days']>2500) & (dframe['number_of_days']<=3000)
dframe[filter]=3000

filter=(dframe['number_of_days']>3000) & (dframe['number_of_days']<=3500)
dframe[filter]=3500

filter=(dframe['number_of_days']>3500) & (dframe['number_of_days']<=4000)
dframe[filter]=4000



frame=dframe.groupby('number_of_days',as_index=False).count()
```

In [52]:

```
frame
```

Out[52]:

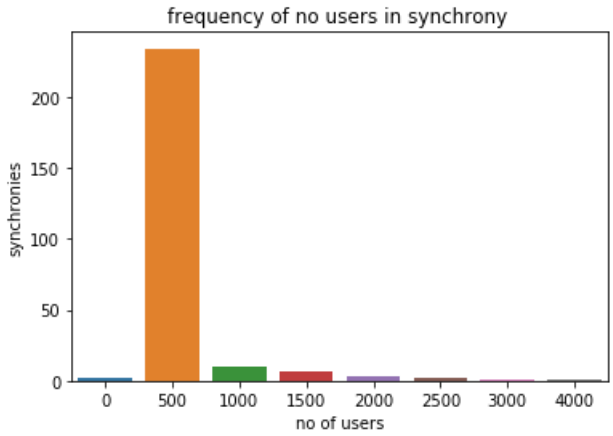
	number_of_days	synchrony
0	0	2
1	500	234
2	1000	10
3	1500	6
4	2000	3
5	2500	2
6	3000	1

	number_of_days	synchrony
7	4000	1

Below graph represents frequency of no of users in terms of no of synchronies

In [53]:

```
fig=sns.barplot('number_of_days','synchrony',data=frame)
fig.set(xlabel='no of users',ylabel='synchronies')
plt.title('frequency of no users in synchrony ')
plt.show()
```



In [54]:

```
X=[]
for i in range(len(dfs)):
    X.append(dfs[i]['retweet'].sum())
frame=pd.DataFrame(X,columns=['retweet'])
```

In [55]:

```
frame['synchrony']=filename
```

In [56]:

```
frame.head()
```

Out[56]:

	retweet	synchrony
0	False	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
1	199	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
2	7297	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
3	338525	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
4	289857	C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...

In [57]:

```
frame.std()
```

Out[57]:

```
retweet    148103.673486
dtype: float64
```

In [58]:

```
frame.min()
```

Out[58]:

```
retweet                                                    False
synchrony    C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
dtype: object
```

In [59]:

```
frame=frame.sort_values('retweet',axis=0, ascending=True, inplace=False, kind='quicksort', na_posit
ion='last')
```

In [60]:

```
frame.iloc[258]
```

Out[60]:

```
retweet                                                  2066476
synchrony    C:\Users\LEELA SURYA TEJA\Desktop\syncs2\Sync9...
Name: 44, dtype: object
```

In [61]:

```
num=frame['retweet']
```

In [62]:

```
filter=(frame['retweet']>=0) & (frame['retweet']<=1000)
frame[filter]=1000

filter=(frame['retweet']>1000) & (frame['retweet']<=2500)
frame[filter]=2500


filter=(frame['retweet']>2500) & (frame['retweet']<=5000)
frame[filter]=5000

filter=(frame['retweet']>5000) & (frame['retweet']<=10000)
frame[filter]=10000


filter=(frame['retweet']>10000) & (frame['retweet']<=25000)
frame[filter]=25000


filter=(frame['retweet']>25000) & (frame['retweet']<=50000)
frame[filter]=50000

filter=(frame['retweet']>50000) & (frame['retweet']<=75000)
frame[filter]=75000

filter=(frame['retweet']>75000) & (frame['retweet']<=100000)
frame[filter]=100000



filter=(frame['retweet']>100000) & (frame['retweet']<=250000)
frame[filter]=250000


filter=(frame['retweet']>25000) & (frame['retweet']<=500000)
frame[filter]=500000

filter=(frame['retweet']>500000) & (frame['retweet']<=1000000)
frame[filter]=1000000

filter=(frame['retweet']>1000000) & (frame['retweet']<=2000000)
frame[filter]=2000000
```

```
filter=(frame['retweet']>2000000) & (frame['retweet']<=5000000)
frame[filter]=5000000


f=frame.groupby('retweet',as_index=False).count()
```

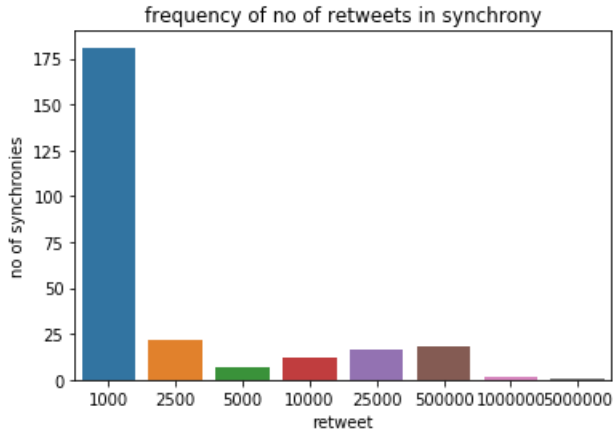
Below graph represents frequency of no of retweets in terms of synchronies

In [63]:

```
fig=sns.barplot('retweet','synchrony',data=f)
plt.ylabel('no of synchronies')
plt.title('frequency of no of retweets in synchrony')
```

Out[63]:

```
Text(0.5,1,'frequency of no of retweets in synchrony')
```



In [64]:

```
y=[]
for j in range(len(dfs)):
    y.append('#'+filenames[j].split('#')[1].split('.')[0])

synchrony_name=pd.DataFrame(y,columns=['name of synchrony'])
```

In [65]:

```
synchrony_name.head(20)
```

Out[65]:

	name of synchrony
0	#sscexamscam
1	#bangkok
2	#bestfanarmy
3	#bts
4	#btsarmy
5	#chennai
6	#dubai
7	#endomondo
8	#endorphins
9	#hyderabad
10	#ibeartawards

	name of synchrony
11	#india
12	#love
13	#mumbai
14	#photography
15	#repost
16	#singapore
17	#thailand
18	#thtraffic
19	#travel

synchronies with hashtag stored in the csv file named as "synchrony_name"

In [66]:

```
synchrony_name.to_csv('synchrony_name.csv',index=True)
```