In [37]:

```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib notebook
```

In [2]:

```
df_day = pd.read_csv('../datasets/bike_sharing/day.csv')
df_hour = pd.read_csv('../datasets/bike_sharing/hour.csv')
```

In [3]:

```
df_day.head()
```

Out[3]:

| | instant | dteday | season | yr | mnth | holiday | weekday | workingday | weathersit | temp | a |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 2011-01-01 | 1 | 0 | 1 | 0 | 6 | 0 | 2 | 0.344167 | 0.3 |
| **1** | 2 | 2011-01-02 | 1 | 0 | 1 | 0 | 0 | 0 | 2 | 0.363478 | 0.3 |
| **2** | 3 | 2011-01-03 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0.196364 | 0.1 |
| **3** | 4 | 2011-01-04 | 1 | 0 | 1 | 0 | 2 | 1 | 1 | 0.200000 | 0.2 |
| **4** | 5 | 2011-01-05 | 1 | 0 | 1 | 0 | 3 | 1 | 1 | 0.226957 | 0.2 |

In [5]:

```
df_hour.sample(5)
```

Out[5]:

| | instant | dteday | season | yr | mnth | hr | holiday | weekday | workingday | weathersit | temp |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **13102** | 13103 | 2012-07-05 | 3 | 1 | 7 | 3 | 0 | 4 | 1 | 1 | 0.72 |
| **16666** | 16667 | 2012-12-02 | 4 | 1 | 12 | 5 | 0 | 0 | 0 | 2 | 0.26 |
| **1914** | 1915 | 2011-03-25 | 2 | 0 | 3 | 13 | 0 | 5 | 1 | 1 | 0.32 |
| **3778** | 3779 | 2011-06-11 | 2 | 0 | 6 | 8 | 0 | 6 | 0 | 1 | 0.74 |
| **11977** | 11978 | 2012-05-19 | 2 | 1 | 5 | 6 | 0 | 6 | 0 | 1 | 0.44 |

In [10]:

```
len(df_day), len(df_hour)
```

Out[10]:

```
(731, 17379)
```

In [32]:

```
for elem_cnt, elem_casual, elem_reg in zip(list(df_day.cnt),
                                           list(df_day.casual),
                                           list(df_day.registered)):
    if elem_cnt != elem_casual + elem_reg:
        print("erro")
        break
```

In [29]:

```
all(df_day.cnt == df_day.casual + df_day.registered)
```

Out[29]:

```
True
```

In [34]:

```
all(df_hour.cnt == df_hour.casual + df_hour.registered)
```

Out[34]:

```
True
```

In [36]:

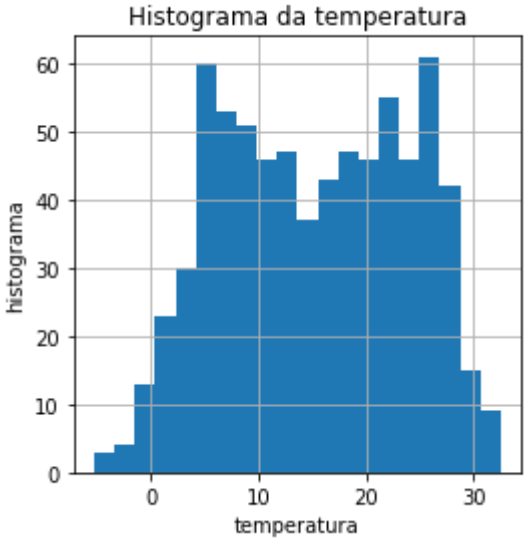
```
tmin = -8
tmax = 39

df_day['temperatura'] = df_day.temp.map(lambda t:t*(tmax-tmin)+tmin)
df_day.head()
```

Out[36]:

| | instant | dteday | season | yr | mnth | holiday | weekday | workingday | weathersit | temp | a |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 2011-01-01 | 1 | 0 | 1 | 0 | 6 | 0 | 2 | 0.344167 | 0.3 |
| **1** | 2 | 2011-01-02 | 1 | 0 | 1 | 0 | 0 | 0 | 2 | 0.363478 | 0.3 |
| **2** | 3 | 2011-01-03 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0.196364 | 0.1 |
| **3** | 4 | 2011-01-04 | 1 | 0 | 1 | 0 | 2 | 1 | 1 | 0.200000 | 0.2 |
| **4** | 5 | 2011-01-05 | 1 | 0 | 1 | 0 | 3 | 1 | 1 | 0.226957 | 0.2 |

In [46]:

```
plt.figure(figsize=(4,4))
plt.title('Histograma da temperatura')
plt.xlabel('temperatura')
plt.ylabel('histograma')
plt.grid()
plt.hist(df_day.temperatura,bins=20);
```




In [ ]: