

In [1]:

```
import matplotlib
import pandas as pd
%matplotlib inline
```

In [3]:

```
registrations_dataset = './events_recommender/resources/user_registrations.csv'
```

In [6]:

```
regs_df = pd.read_csv(registrations_dataset)
```

Registrations Information

In [7]:

```
regs_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1999 entries, 0 to 1998
Data columns (total 3 columns):
event_edition_id    1999 non-null int64
user_id             1999 non-null int64
registered          1999 non-null int64
dtypes: int64(3)
memory usage: 46.9 KB
```

Total Number of Null and Non Null Counts

In [8]:

```
regs_df.count()
```

Out[8]:

```
event_edition_id    1999
user_id             1999
registered          1999
dtype: int64
```

```
regs_df.groupby('user_id')['event_edition_id'].count().plot.bar(width=5,log=True)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f4dfb027550>
```

