PrepareCattleEdgeList

January 9, 2015

1 Extract animal movement within a given time window

Thomas Selhorst, January 2015

> In order to extract the animal movements within a given time window w we will use the **pandas** and the **datetime** packages

This notebook is organized as followes: 1. read a dataset into a **pandas** dataframe, 2. extract all movement within a given timeframe.

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

```
In [2]: # this dataset consists of 38,260,192 lines
        ddata = pd.read_csv("/Users/TOSS/Documents/Projects/CattleTVG/Edgelist/elri.csv",
                            sep=',', parse_dates=['Date'], index_col='Date')
```

```
In [3]: ddata.head(10)
```

```
Out[3]:                        S                T  VOL
        Date
        2001-07-24  276010010000003  276010010000001    1
        2003-12-11  276010010000011  276010010000001   12
        2003-12-12  276010010000011  276010010000001    2
        2003-12-30  276010010000011  276010010000001    2
        2004-01-05  276010010000011  276010010000001    1
        2004-01-14  276010010000011  276010010000001    2
        2004-02-06  276010010000011  276010010000001    2
        2004-02-26  276010010000011  276010010000001    1
        2004-03-02  276010010000011  276010010000001    6
        2004-03-15  276010010000011  276010010000001    1
```

```
In [4]: ddata.shape
```

```
Out[4]: (38260192, 3)
```

```
In [5]: df = ddata.sort()
```

```
In [6]: import datetime as dt
```

```
In [7]: start = dt.datetime(2001,1,2)
```

The parameter to `timedelta()` has to be **6** because it represents a range starting at 0.

```
In [8]: end = start + dt.timedelta(6)
```

```
In [9]: start, end
```

1

```
Out[9]: (datetime.datetime(2001, 1, 2, 0, 0), datetime.datetime(2001, 1, 8, 0, 0))

In [10]: selected = df[start:end]

In [11]: selected.head()

Out[11]:                          S                  T  VOL
         Date
         2001-01-02  276091861250429    276091861250054    1
         2001-01-02  276059580402808    276059660280035    1
         2001-01-02  276094751230115    276094640000230    1
         2001-01-02  276097781610030    276097801250050    2
         2001-01-02  276072320860018    276072321010403    1

In [12]: selected.tail()

Out[12]:                          S                  T  VOL
         Date
         2001-01-08  276094731410002    276094610000387    1
         2001-01-08  276091891620172    276091831480192    1
         2001-01-08  276160610450090    276031550108533    5
         2001-01-08  276143750500065    276143750700482    1
         2001-01-08  276073352030023    276072351490813    2

In [13]: selected.shape

Out[13]: (72037, 3)

In []:
```