Offleaks

This project for Social and Economic Networks course uses Offshore Leaks Database found here https://offshoreleaks.icij.org/.

Yearly counts development

One would like to know how many accounts have been opened throughout years.

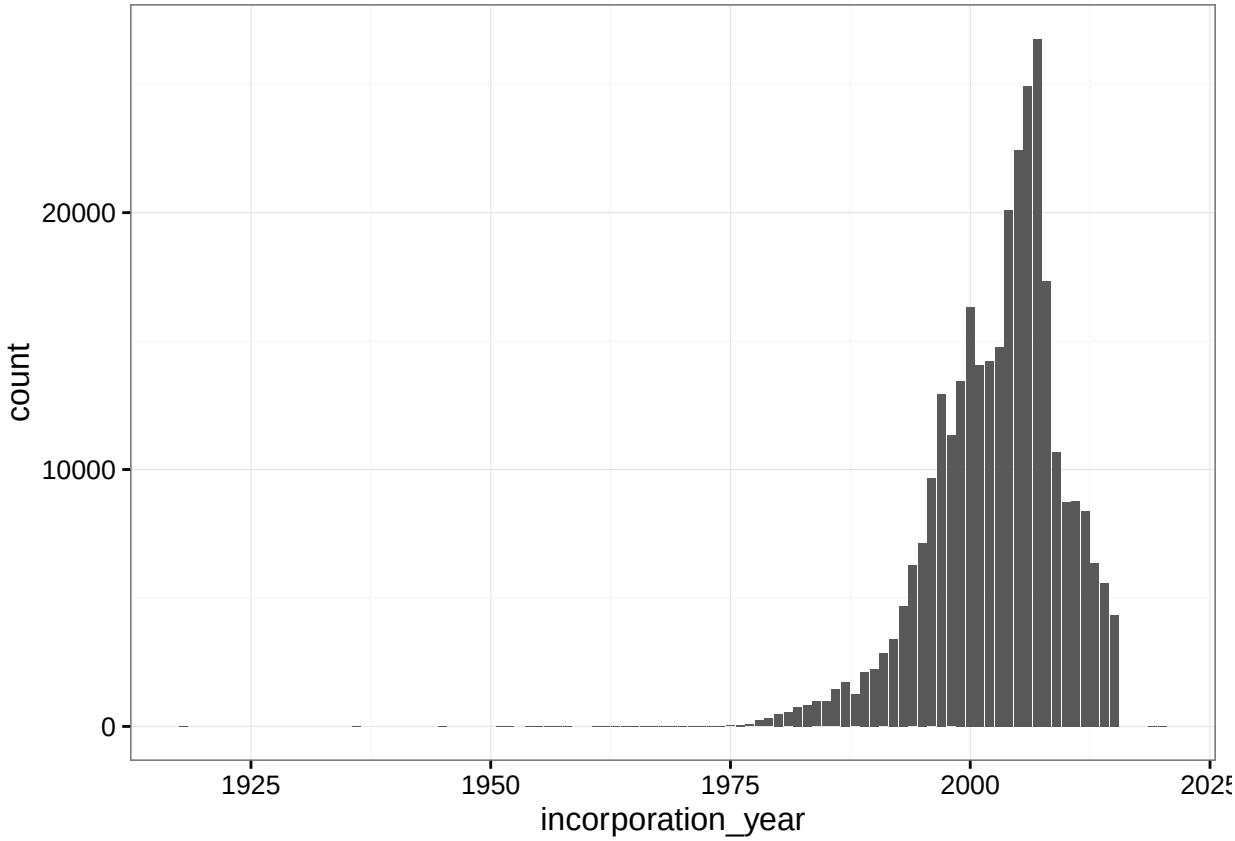
The sample size for answering this question is 309736.

The earliest opened account belongs to 1918 and the latest comes from the future 2020.

The most fruitful years for the offshore accounts are the following:

```
## Source: local data frame [6 x 2]
##
##   incorporation_year     n
##                (dbl) (int)
## 1               2007 26774
## 2               2006 24924
## 3               2005 22454
## 4               2004 20109
## 5               2008 17342
## 6               2000 16327
```

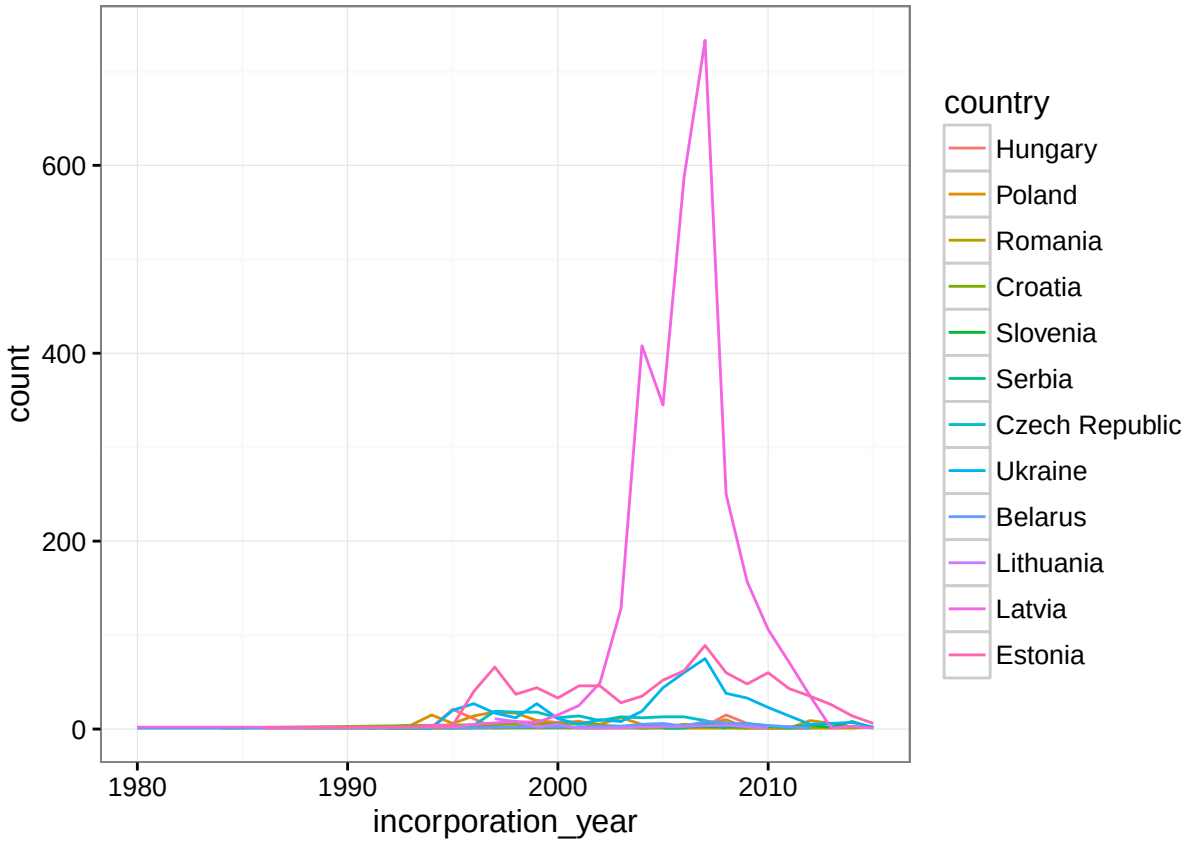
Overall, that is how the process looks like:



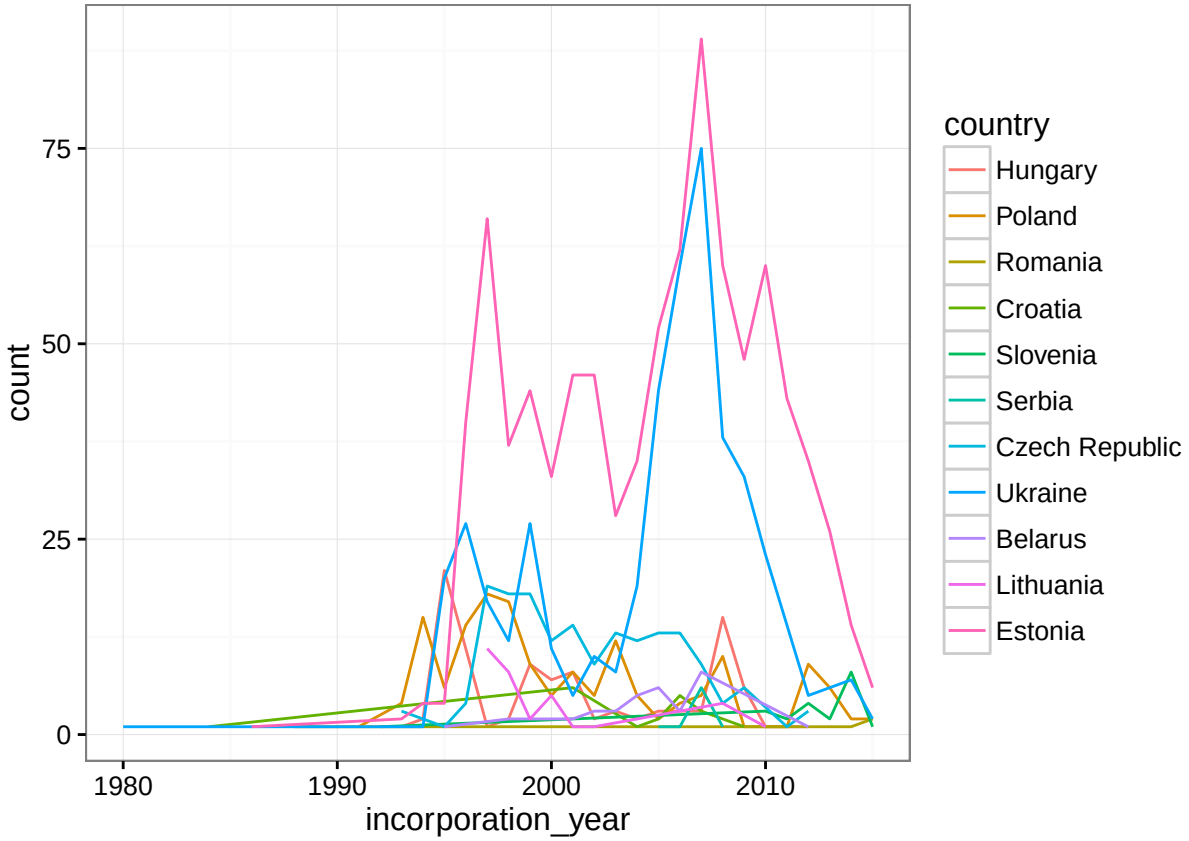
CEE countries counts

Number of firms related to different Central-Eastern European countries.

```
##
##        Hungary         Poland        Romania        Croatia       Slovenia
##             90            161              8             19             21
##         Serbia Czech Republic        Ukraine        Belarus      Lithuania
##              9            172            465             34             33
##         Latvia        Estonia
##           2938            881
```



Without Latvia:



CEE countries relationships

It is interesting to see how are the relationships built around the entities from these countries.

There are 13188 known links related to CEE countries entities.

The count allocation of the nodes from origin to end is the following:

```
##                 end
## origin           address cee-entity
##   cee-entity        1948          9
##   intermediary         0       4831
##   non-cee-entity       0          6
##   officer              0       6394
```

The relationship types involved in the entities from CEE countries from origins and ends respectivelly are:

```
##                                    origin
## relationship                        cee-entity intermediary
##   beneficial owner of                        0            0
##   beneficiary of                             0            0
##   director of                                0            0
##   intermediary of                            0         4831
##   registered address                      1948            0
##   same name and registration date as         9            0
##   secretary of                               0            0
```

```
##   shareholder of                                0           0
##                                     origin
## relationship                         non-cee-entity officer
##   beneficial owner of                             0       7
##   beneficiary of                                  0     199
##   director of                                     0    2502
##   intermediary of                                 0       0
##   registered address                              0       0
##   same name and registration date as              6       0
##   secretary of                                    0      19
##   shareholder of                                  0    3667

##                                     end
## relationship                         address cee-entity
##   beneficial owner of                      0          7
##   beneficiary of                           0        199
##   director of                              0       2502
##   intermediary of                          0       4831
##   registered address                    1948          0
##   same name and registration date as       0         15
##   secretary of                             0         19
##   shareholder of                           0       3667
```

Nodes decomposed

The main connections file provides 1269796 links/edges. At both ends one can find different players.

Origin nodes

There are 462468 unique origin nodes/vertices.

They are distributed among the types in the following manner:

Entities	Addresses	Intermediaries	Officers
93572	960	23600	345475
20.1834744%	0.2070719%	5.0905185%	74.5189352%

A note. The decomposition does not sum up to the number of unique entries. 463607 = 462468 is FALSE.

Receiving nodes

Looking at the unique receiving nodes/vertices, one can find that there are 492143 of them.

They are distributed among the types in the following manner:

Entities	Addresses	Intermediaries	Officers
319122	151043	355	21664
64.8379468%	30.6883198%	0.0721275%	4.4016059%

In the same manner 492184 = 492143 is FALSE.

Origins and ends overall

The cross tabulated origins and ends from all the edges available are presented in the table below.

```
##
##                address entity intermediary officer
##   address          960      0            0       0
##   entity         93623   3768            0       0
##   intermediary    9358 362425           21      56
##   officer       213491 538166          349   47579
```