Tower of Babble? Predicting UN (in)action

DAT7

Alex Kapitanskaya

In situations of crisis…

Iraq, Syria, Afghanistan, Yemen, Congo, Darfur, South Sudan, Chad, Mali, Nigeria, Somalia, Ukraine, Caucasus, Libya…

Can we predict what the UN Security Council will do?

Why does this matter?
A real-life example:

- Haiti: elections scheduled for 2015
- MINUSTAH, the UN peacekeeping mission in Haiti, was deployed in 2004
- MINUSTAH may leave in 2015
- Concerns of possible chaos in its aftermath
- Operational decisions on investment, security, etc. depend on whether MINUSTAH will stay or go
  - MNCs, NGOs, World Bank, IDB must decide how to structure operations

Building a predictive model

- Inputs: records of UN Security Council meetings
- Outcomes: UN Security Council Resolutions
- Timeframe: 1994 – 2014
- Total observations: 1,250
- Only considering matters of peace and security, not admission of new members, appointment of judges to international courts, etc.

Possible outcomes

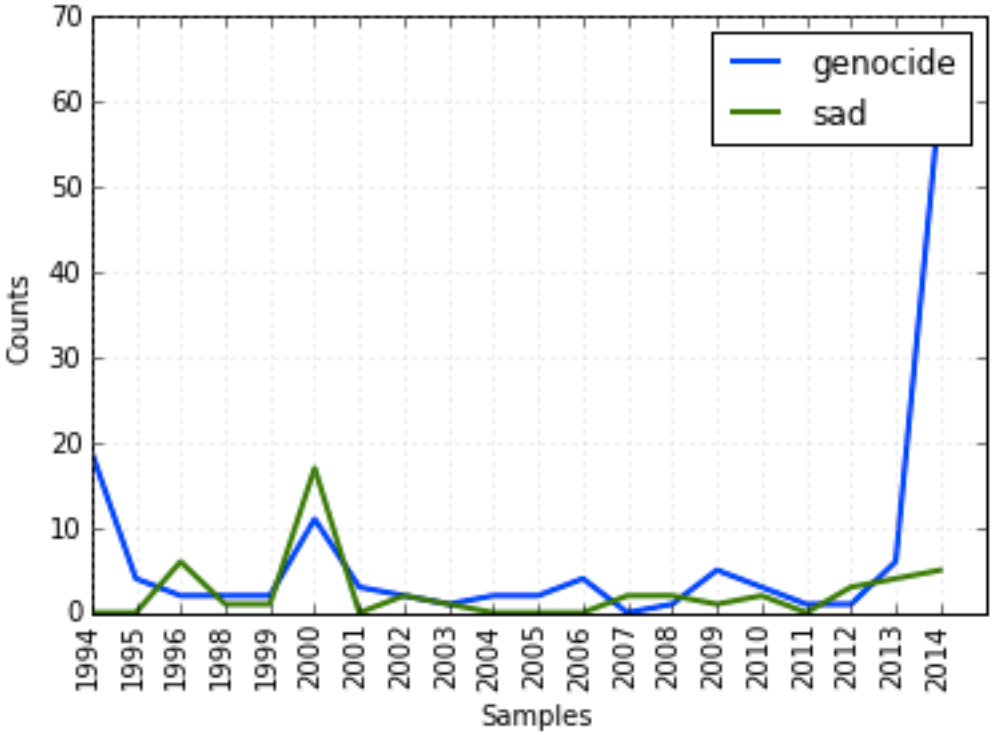
- 0 – no action
- 1 – soft measures (observers)
- 2 – prohibitive measures (embargoes)
- 3 – hard measures (peacekeeping missions)
- 4 – reversal of soft measures
- 5 – reversal of prohibitive measures
- 6 – reversal of hard measures

What I've done

- Scraped tables of meeting records and corresponding resolutions for all years into a single CSV file
- Cleaned data (removed meetings that didn't lead to resolutions, added a year column)
- Basic data exploration by topic in pandas
- Downloaded all meeting record files & resolutions (manually ☹)

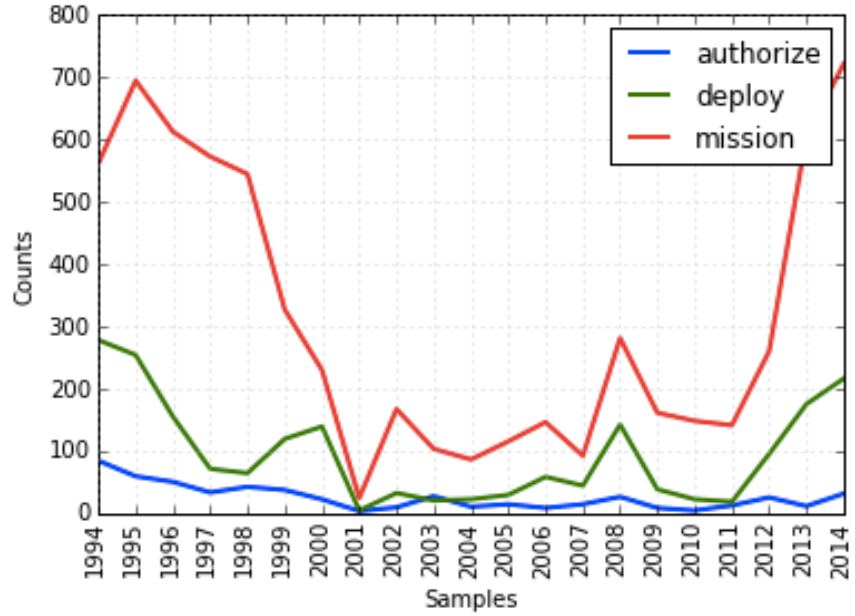
What I've done

- Lots of human learning
- Built two corpora (meeting records (PV) & resolutions (SCRs))
- Explored frequency distributions with NLTK



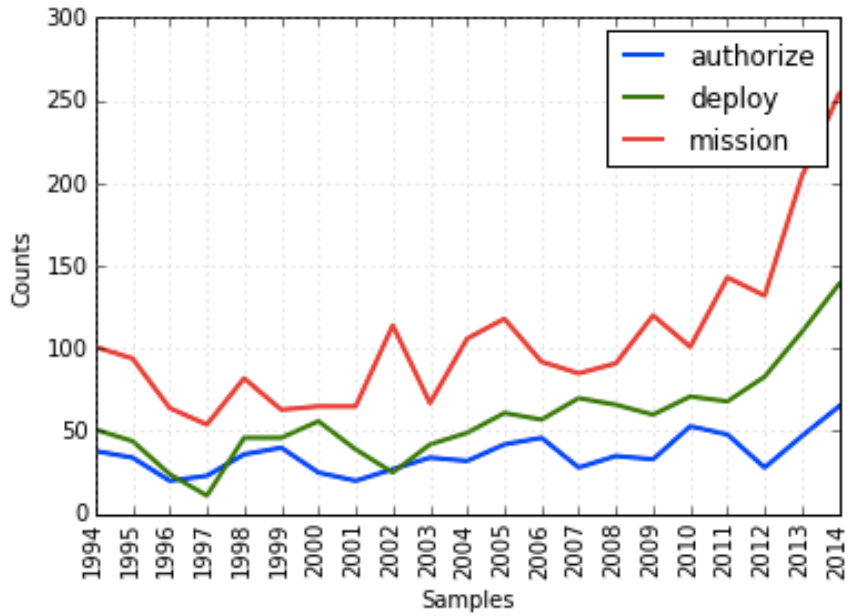
Example: peacekeeping missions mentioned in **meeting records**

```
In [47]: cfd = nltk.ConditionalFreqDist(
    ...:     (target, fileid[:4]) # separates the years out of the filenames for the X axis
    ...:     for fileid in pvs.fileids()
    ...:     for w in pvs.words(fileid)
    ...:     for target in ['authorize', 'mission', 'deploy']
    ...:     if w.lower().startswith(target))
    ...:
    ...: cfd.plot()
```

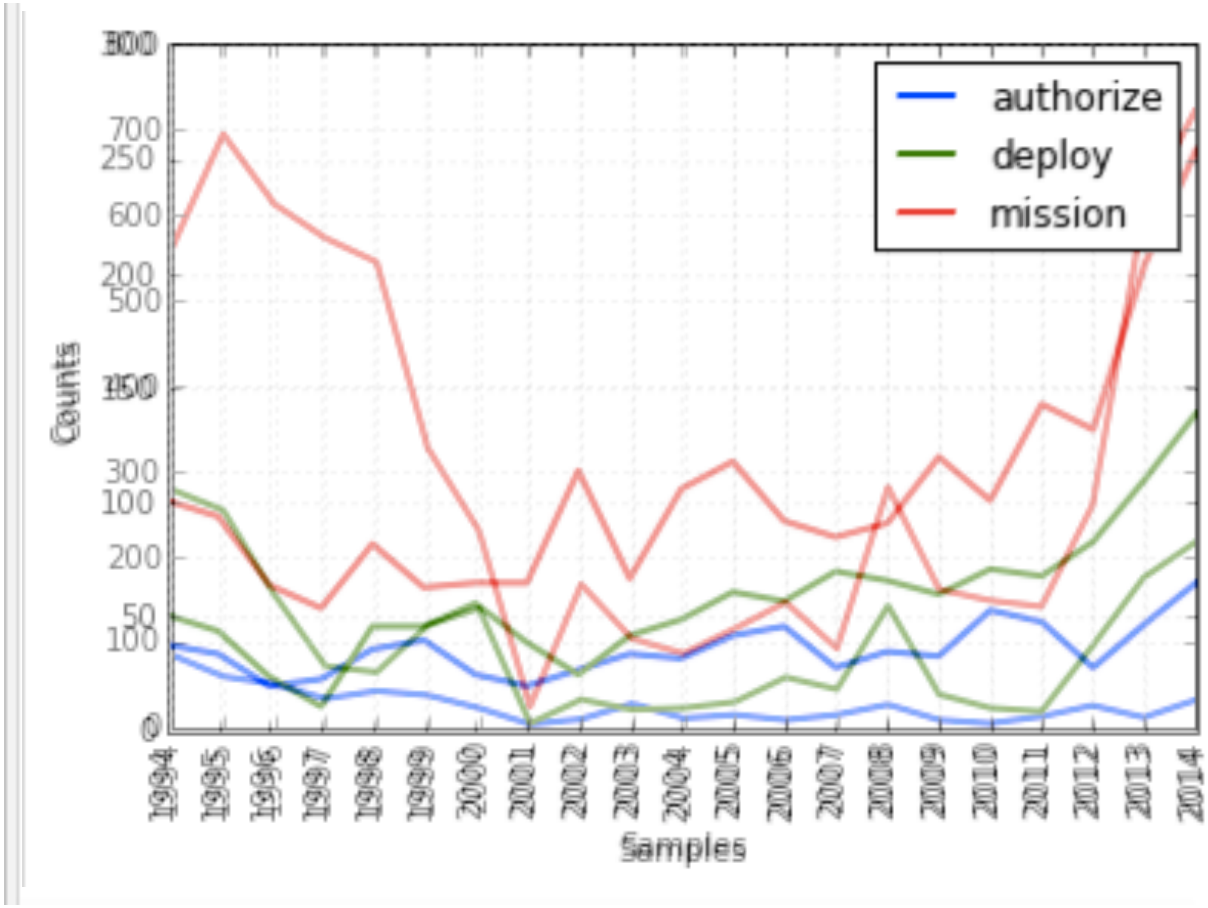


Example: peacekeeping missions mentioned in **resolutions**

```
In [46]: cfd = nltk.ConditionalFreqDist(
    ...:     (target, fileid[:4]) # separates the years out of the filenames for the X axis
    ...:     for fileid in scrs.fileids()
    ...:     for w in scrs.words(fileid)
    ...:     for target in ['authorize', 'mission', 'deploy']
    ...:     if w.lower().startswith(target))
    ...:
    ...: cfd.plot()
```



Overlay – possible relationship?



Tools for Mac OS X

- PDF Converter with OCR ($12.99)
  - For PDF files with text as images
- TEConverter ($4.99)
  - Batch converts .txt files to UTF-8
  - Tried bash scripts, accidentally overwrote all files with blank ones…