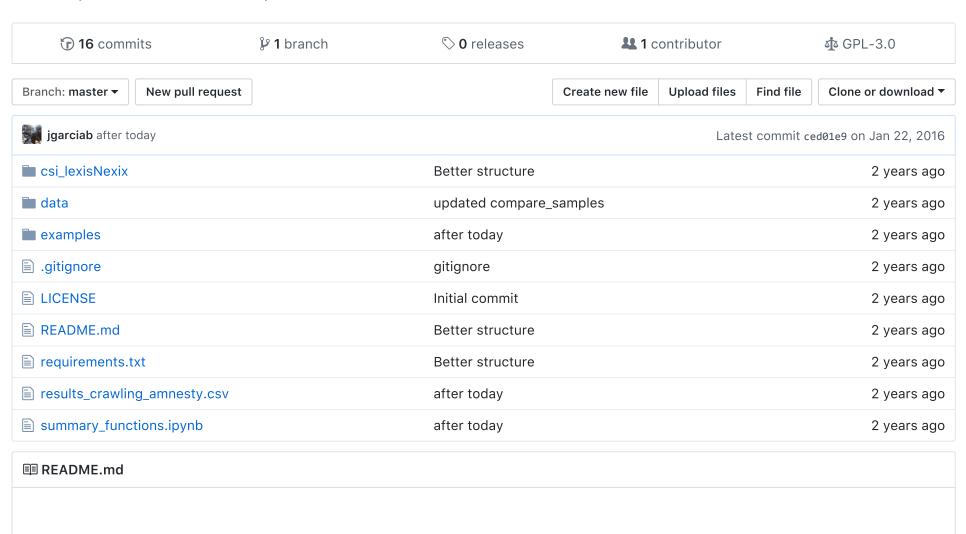
jgarciab / text_analysis_python

Text analysis in social sciences - Python



text_analysis_python

REQUIRES PYTHON 3.

This is the code for the workshop "Automated text analysis" at the University of Amsterdam (21-22nd January '15). It is tailored for beginners, so there are some bad practises that make the code easier to understand (some global variables, no Classes).

To use:

- Install Python3 with Anacondas: https://www.continuum.io/downloads
- Open an ipython notebook
- Open any of the ".ipynb" files

The notebook *summary_functions.ipynb* contains all the important stuff. The documentation is far from complete.

The file requirement.txt has the modules we need:

- Present in Anacondas: pandas, numpy, scipy, matplotlib, statsmodels, nltk, bs4, bisect
- Not present in Anacondas (Install from the terminal/anaconda prompt using pip install name_of_module): selenium, dateutil, pycountry, labMTsimple, lda, wbdata

The folder examples have the notebooks used in the workshop:

- Day 1: Introduction, Scraping_B, Visualization
- Day 2: Scraping_A, Sentiment, LDA (Semantic), Classification