

# NLP For Economists

## Topic Modeling

Sowmya Vajjala

Munich Graduate School of Economics - LMU Munich

Guest Course, October 2020

# Goals

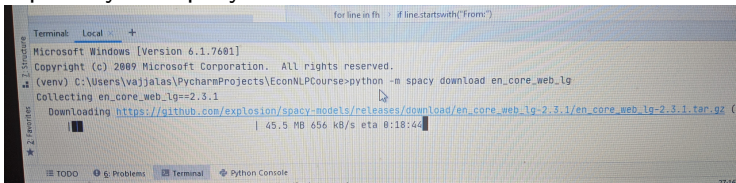
- ▶ Just a few general instructions on installing python libraries
- ▶ note: I am taking pycharm as the example IDE.

# Libraries needed

- ▶ nltk - for various NLP functions
- ▶ spacy - for various NLP functions
- ▶ textacy - for key phrase extraction examples
- ▶ gensim - for text representation example and others I don't mention here
- ▶ sumy - for text summarization algorithms
- ▶ sklearn - for training machine learning models and some basic vectorization of text
- ▶ lime - for visualizing text classification output
- ▶ pandas - for csv file processing
- ▶ snorkel -for corpus collection example

# How to install

- ▶ <https://www.jetbrains.com/help/pycharm/installing-uninstalling-and-upgrading-packages.html>
- ▶ Once installed, you may need to download additional models, especially for Spacy.



```
Terminal: Local x +
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.
(venv) C:\Users\vajjalas\PycharmProjects\EconNLPCourse>python -m spacy download en_core_web_lg
Collecting en_core_web_lg==2.3.1
  Downloading https://github.com/explosion/spacy-models/releases/download/en_core_web_lg-2.3.1/en_core_web_lg-2.3.1.tar.gz (7
    | 45.5 MB 656 kB/s eta 0:18:44
```

- ▶ NLTK also requires it, and I mentioned in an earlier video about installations.