

You should have:

- A Python file called shareTopicModeling.py
- A file called lowercase_words.pkl
- A file called vader_lexicon.txt
- A sample CSV called Sample_CSV.csv

SET-UP (IF YOU DO NOT HAVE AN EDITOR):

- Download PyCharm Edu
- Download Anaconda 2.7 (not 3.0)
 - Anaconda is great because it makes sure all Python packages are synchronized. Sometimes one package is updated but not another, and they no longer function together. By updating your packages with Anaconda (below), it's hard to screw things up.
- Make sure Anaconda is your interpreter:
 - Pycharm Edu → Preferences → Project Interpreter → 2.7.11 (~/anaconda/bin/python.app)
- To update Anaconda:
 - Go to Terminal
 - Enter "conda update" (no quotes)
 - Enter Y to accept updates

*As long as you're using Anaconda 2.7 and "conda" when you're in Terminal, nothing should seriously break (I've had issues with "pip" in Terminal).

*I highly recommend trying the PyCharm tutorials. They're a great starter to Python if you're not already familiar with it. Note that the "tutorial" version of PyCharm looks different than its "regular" version.

TO RUN CODE:

- File → New Project
- In your New Project:
 - Drag shareTopicModeling.py file into left section (with all files)
 - Drag lowercase_words.pkl to same section
 - Drag vader_lexicon.txt to same section
 - Drag your csv file (the one you're analyzing) to the same section
- In shareTopicModeling.py, adjust the sections I've noted. You'll need to enter the file path for your csv, column headings for columns that you're analyzing, variables for lists of data you want to analyze, and similarity words.
- Press the green play button at the top to run.

*Take a look at the Sample_CSV.csv to get an idea of how to set up data. Helpful tips are included.

*Errors are inevitable, especially errors related to unicode. Copy and paste the last line of error messages into Google - you'll find results quickly. StackOverflow is a great resource.

*Don't be afraid of playing with your code. If you're not sure about a line, duplicate it and put a # in front of one of them - this will save the line for reference but "inactivate" it.

*DO be careful if you decide to write to a csv (instructions not included here). I almost

lost all my Master's data that way.