

## Documentation for News Sentiment Scores

Latest revision: 09/17/2020

This code was developed in the study “Measuring News Sentiment” by Adam Shapiro, Moritz Sudhof, and Daniel Wilson. Forthcoming in *Journal of Econometrics*, <https://doi.org/10.1016/j.jeconom.2020.07.053>. Please cite accordingly.

### Copyright

Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson. 2020. “Measuring News Sentiment.” Federal Reserve Bank of San Francisco Working Paper 2017-01. <https://doi.org/10.24148/wp2017-01>

**Copyright ©2020, Adam Shapiro, Moritz Sudhof, and Daniel Wilson and the Federal Reserve Bank of San Francisco. All rights reserved.**

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Redistributions [and use] in source and binary forms must be for Noncommercial purposes only. “Noncommercial” means not primarily intended for or directed towards commercial advantage or monetary compensation.
- Neither the name “Federal Reserve Bank of San Francisco” nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

### Folder overview

This folder contains materials to replicate the News Sentiment score described in Federal Reserve Bank of San Francisco Working Paper 2017-01, “Measuring News Sentiment” by Shapiro, Sudhof, and Wilson, as listed below.

|                            |  |
|----------------------------|--|
| main_NS_score.py           | The main python script developed to calculate the News PMI model sentiment score of .txt files. Before running this file, set the directories in lines 10-12 and the names of the folders with the .txt files in line 13. You will also need to set the directory for the lexicons in line 10 of NS_functions.py<br><br>This file outputs a .csv file with the file name and corresponding sentiment score |
| NS_functions.py            | An auxiliary python script that contains functions to load the lexicons used by main_NS_score.py   |
| LoughranMcDonald_2016.csv  | Loughran-McDonald 2016 lexicon. Loaded in main_NS_score.py<br><br>Dictionary source: Loughran, T., and B. McDonald. 2011. “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks.” <i>The Journal of Finance</i> 66: 35-65.<br><a href="https://sraf.nd.edu/textual-analysis/">https://sraf.nd.edu/textual-analysis/</a>  |
| ns.vader.sentences.20k.csv | Lexicon created using a corpus of economic news articles and scored with Vader   |