

An explanatory memo for implementing the rescaling method described in the manuscript, “Building Longitudinal *Google Trends* to Measure Dynamic Local-Level Issue Attention”

This data archive includes several files used to implement the rescaling method described in the manuscript. Analysis 1 is the first example in the manuscript which is about the jobs index. Analysis 2 is the second example which is about the cdc-cnn-fox news comparison index. In each of the two analyses, first, run the CrossSection and TimeSeries files to collect Google Trends data.¹ Once new folders are created and the Google Trends data are saved in the folders, run the Rescale file in Jupyter Notebook to rescale the Google Trends cross-sectional indices.

Analysis 1: The “jobs” index

- `CrossSection_jobs.py`: This code is to collect DMA-level cross-sectional Google Trends indices for the jobs search term for the period between December 31, 2006 and December 31, 2016. DMA refers to the Designated Market Area, a geographic area where people receive the same television station offerings in the United States.
- `TimeSeries_jobs.py`: This code is to collect the DMA-level time-series Google Trends indices for the jobs search term for the period between December 31, 2006 and December 31, 2016. Though this code generates 210 DMAs’ jobs time-series, the manuscript only uses the Honolulu time-series as the reference for the rescaling method.
- `Rescale_jobs.ipynb`: This Jupyter Notebook code is to rescale the cross-sectional jobs indices using the reference Honolulu time-series. This code corresponds to the second and third steps in the manuscript.
- `dma_googleid.tsv`: This text data file includes individual DMAs’ Google ID. This file is called in the Python code where needed.

¹Note that it may take long.

Analysis 2: The “cdc-cnn-fox news” index

- `CrossSection_comparison.py`: This code is to collect DMA-level cross-sectional Google Trends indices for the CDC-CNN-Fox News comparison-search term for the period between February 1, 2020 and May 10, 2020.
- `TimeSeries_comparison.py`: This code is to collect DMA-level time-series Google Trends indices for each of the CDC, CNN, and Fox News single-search term for the period between February 1, 2020 and May 10, 2020. Though this code generates 210 DMAs’ CDC, CNN, and Fox News time-series, the manuscript only uses the CNN time-series for Atlanta as the reference for the rescaling method.
- `Rescale_comparison.ipynb`: This Jupyter Notebook code is to rescale the cross-sectional CDC-CNN-Fox News comparison-search indices using the reference Atlanta time-series. This code corresponds to the second and third steps in the manuscript.
- `dma_googleid.tsv`: This text data file includes individual DMAs’ Google ID. This file is called in the Python code where needed.