

Quantifying Voter Suppression

Technology Review



Andres De La Fuente

Anmol Srivastava

Juan Solorio

Background & Needs

Visualize voter suppression across the US:

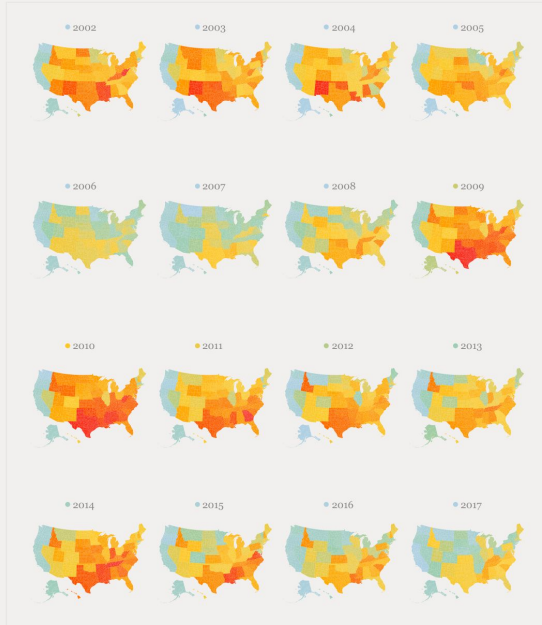
- analyses that group regions by degree of inhibited turnout
- interactive map, with proper labels and contributing factors

Specifically, we need Python tools that can:

- perform cluster analysis, regression, or group comparisons
- generate an interactive browser-friendly map, based on our findings

At Minimum.

% of all U.S. hourly workers earning minimum wage or less



Source: U.S. Bureau of Labor Statistics
Inspired by FiveThirtyEight's Ella Kocze

Tricky additions:

- user-adjusted time attribute
- on-click information (pop-ups)

Use Case: Analysis

Objective

- user wants to compare voting patterns in one year, across states
- group states based on available attributes
- give map information about each state's grouping and statistics

Interaction

- year is fixed
- user hovers over and selects different states
- map highlights information about states' categorization

Technology Options: Analysis

scipy

- library oriented around mathematics, engineering and computation

scikit-learn

- third-party extension of scipy designed for machine learning

Technology Comparisons: Analysis

scipy

- **more robust support for statistical tests**
 - `scipy.stats.ttest_ind`
 - `scipy.stats.f_oneway`
- **includes ML features, to a lesser capacity**
 - `scipy.cluster.vq`

scikit-learn

- **state-of-the-art ML methods, with interpretability issues**
 - `sklearn.cluster`
- **limited support for statistical tests**

Technology Comparisons: Analysis

scipy

- more robust support for statistical tests
 - `scipy.stats.ttest_ind`
 - `scipy.stats.f_oneway`
- includes ML features, to a lesser capacity
 - `scipy.cluster.vq`

scikit-learn

- **state-of-the-art ML methods, with interpretability issues**
 - `sklearn.cluster`
- **limited support for statistical tests**

Use Case: Map Interactivity

Objective

- user compares changes in voting patterns in the US, as time progresses
- map allows the user to manipulate the 'year' variable and updates itself

Interaction

- user adjusts a slider for the 'year' variable
- map updates itself based on analysis for that year
- user hover and click on specific states
- map lists selected states' details

Technology Options: Map Interactivity

bokeh

- library meant for interactive web-based visualizations

folium

- library for constructing leaflet.js visualizations through Python

Technology Comparisons: Map Interactivity

bokeh

- convenient solutions to interactivity requirements
 - `bokeh.models.Slider`
 - `bokeh.models.HoverTool`
- cleaner, more versatile, and high-performance heat maps and graphs

folium

- restrictive interactivity features
 - `TimeStampedGeoJson(..., time_slider_drag_update=True, ...)`
 - JavaScript-editing for hover functionality
- noisier and rigid choropleth method for map construction

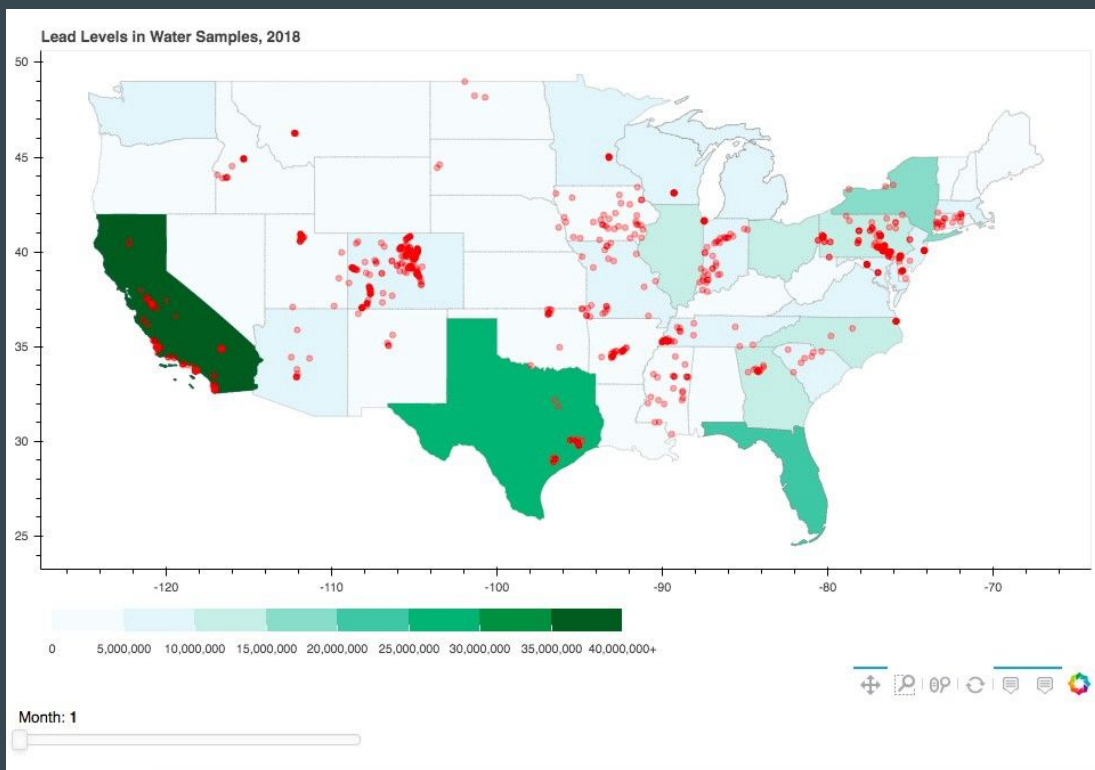
Technology Comparisons: Map Interactivity

bokeh

- convenient solutions to interactivity requirements
 - `bokeh.models.Slider`
 - `bokeh.models.HoverTool`
- cleaner, more versatile, and high-performance heat maps and graphs

folium

- restrictive interactivity features
 - `TimeStampedGeoJson(..., time_slider_drag_update=True, ...)`
 - JavaScript-editing for hover functionality
- noisier and rigid choropleth method for map construction



towardsdatascience.com

Closing Remarks

- `scipy` over `scikit-learn`
 - necessarily involves `pandas` and `numpy`
- `bokeh` over `folium`
 - potential issues in support and documentation
 - may require switch to friendlier packages
 - `Plotly`
 - `Dash` (`React` + `Plotly`)
 - `Tableau` (easy visuals and publication)