Examining Slavic Language Speaker Statistics in the United States
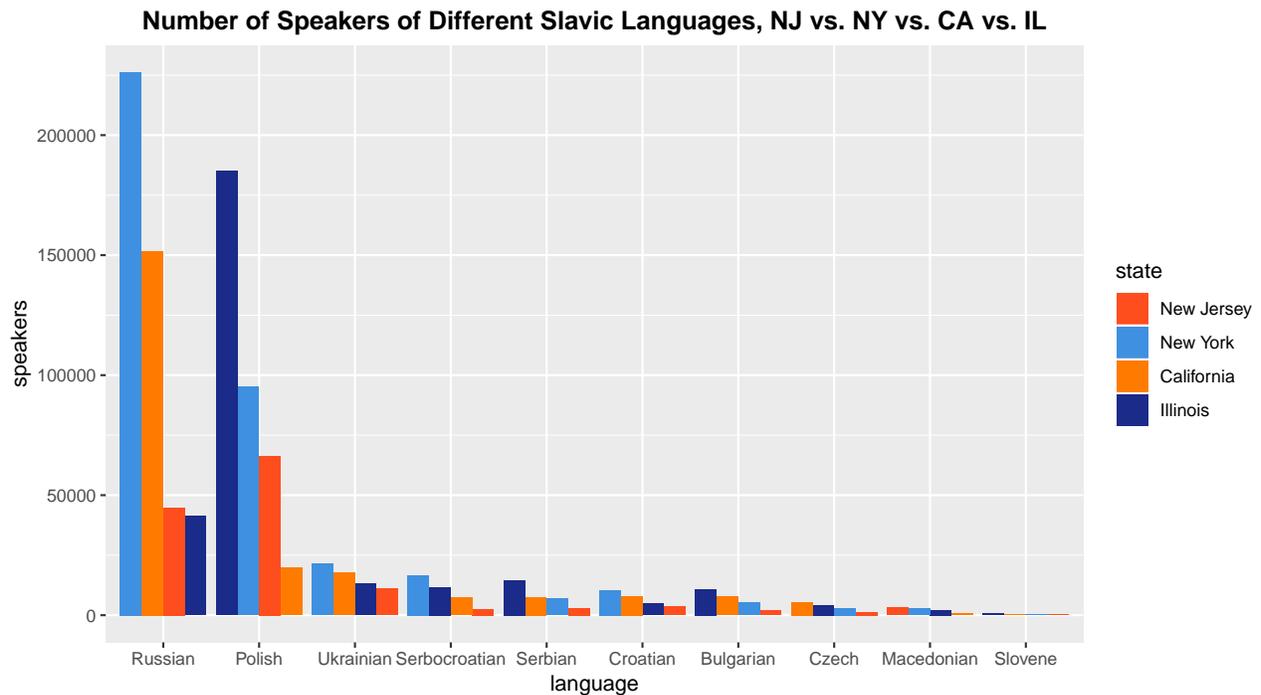
Rhea Kapur

Description

This dataset is entitled: "Detailed Languages Spoken at Home and Ability to Speak English for the Population 5 Years and Over: 2009-2013." The data is from the U.S. Census Bureau, and specifically from the American Community Survey (ACS). For many different languages, and for each state, it contains the number of speakers in the specific state who speak the language at home and the margin of error for that data collection. The dataset also holds the number of people from a specific state who, given that they speak a certain language at home, speak English less than "Very Well." The dataset can be found here: https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html. I used the one with states. I also used the built in "states" dataset in R for lists of state names, region for each state, etc.

I focused on trends in Slavic languages specifically - which ones are spoken the most in states with big cities, etc. I looked at the distribution of speakers of Slavic languages across different regions of America. I also examined how English language speaking proficiency related to number of Slavic speakers differs per state. These were pretty much my initial questions, looking at the data.
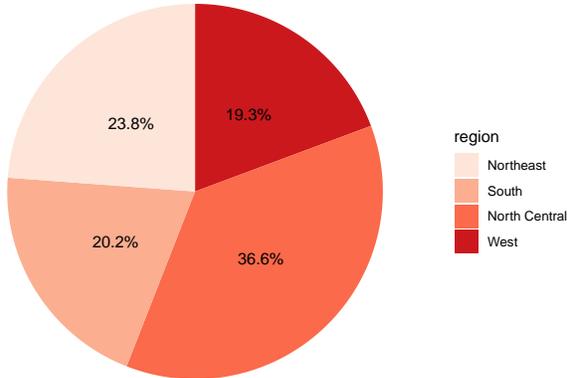
1

Bar chart of the number of speakers for each Slavic language in New Jersey, New York, California, and Illinois. Slavic language speakers tended to move to big cities or stay in the tri-state area after immigrating to the United States, so these states in particular have high populations of speakers.

Clearly, New York (NYC) has the highest population of Russian speakers, followed by California (Los Angeles, San Francisco). Illinois (Chicago) has the highest population of Polish speakers. Slavic languages other than Russian & Polish are, in most cases, the minority.
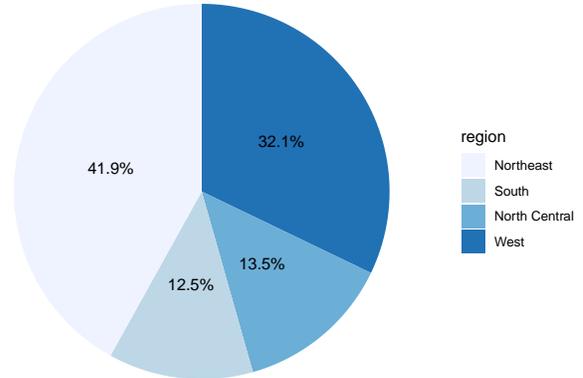
2

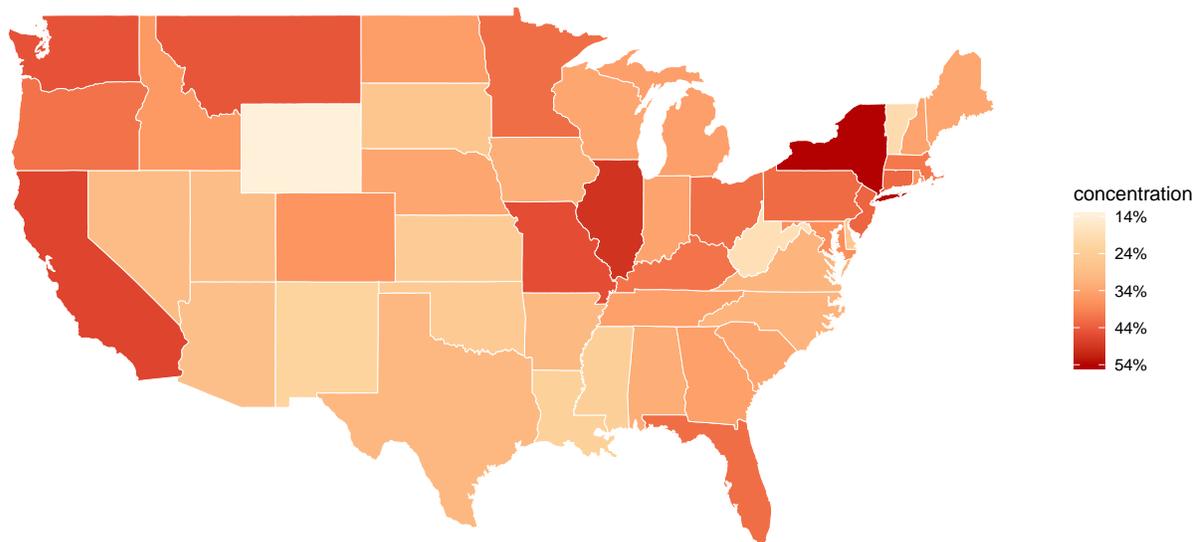**Serbo–Croatian Language Speakers by Region**   **East Slavic Language Speakers by Region**

Density of East Slavic (Russian, Belarusian, Ukrainian) vs. Serbo-Croatian (Serbocroatian, Serbian, Croatian) language speakers in the four different regions of the United States. Belarusian & Ukrainian data was not defined for all states, so I considered ones where it was.

The North Central region of the U.S. - North Dakota, South Dakota, Michigan, Illinois, etc. - has the highest population of Serbo-Croatian language speakers by 12.8%, followed by the Northeast. The Northeast took the lead for the most number of East Slavic language speakers by 9.8%, closely followed by the West.

3

Concentration of Russian or Polish Speakers with Subpar English Proficiency in Each State

Defined a ratio of number of speakers who speak English less than "Very Well" (out of total # of speakers) over the total number of speakers, for each state. Did this for Russian and Polish. Converted all ratios to percentages.

The higher the ratio, the darker the state. Therefore, darker states signify that there is a higher percentage of poor English speakers in that state out of those who speak Russian and/or Polish at home. Clearly, the darkest state is New York, followed by Illinois, then California, then Missouri, etc.

Note that in most cases, darker states are also the ones with the highest number of Russian and Polish speakers and usually have some big city in them, so the results make sense and also correspond with the high number of speakers seen in certain states with big cities in Graph 1. Because there are so many speakers in those states, there are higher populations of the native Slavic speakers who do not speak English well, or even at all.