Popular article on research project national GI-minor
Marieke Statema- February 2016

Mapping immigration views in California (USA)

Every day, people from Central America try to reach the USA via Mexico. Often this is a very dangerous trip. They take the risk to travel on top of a train or are at risk of getting kidnapped by gangs. To limit illegal immigration several initiatives arose to complete the fence along the entire Mexican border. But not all inhabitants of the USA agree with this. The assumption arises that location may matter in explaining these opinions. So, it is interesting to have a look at where people agree and where they disagree. We will have a look at the state of California and answer the question: *'How do the inhabitants of the different counties in California think about completing the fence along the entire Mexican border to prevent illegal immigration in the USA?'*

To answer this question, data of Kieskompas is used. This Dutch company develops online and offline voting advice tools. These tools present lists of statements about political issues that respondents are asked to review. After indicating whether or not they agree with each individual statement, the tool indicates the respondent's position in the political landscape. The vote advice application for the US presidential elections of 2008 included a statement about the completion of the fence along the Mexican border. Besides the political statements, Kieskompas also included some extra questions. With the answers to the question in which city the respondent lives it is possible to link their opinion to a location on a map.

In addition to asking where the respondent lives, there were also questions about their age, gender and education. To improve the representativeness of the obtained responses, the answers are weighted to better match the population of a city. The weighted data of the different cities of a county is merged to get a score per county. Weighting data is a good method when the respondents are not proportionally distributed over the whole population of a city. To do this, basic statistical data on the inhabitants of cities in California is needed. This data was obtained from the '*United States Census Bureau'*. When the data of Kieskompas is compared with the data of the population, each class in the Kieskompas data gets a value. This is the weight of the class and it is used to calculate the score of this class for the whole population. For calculating the scores of the different counties in California the data of the different cities is used. The values of the different cities are weighted towards the average composition of those cities. This gives a weighted score for the cities in California where respondents from the Kieskompas data live. A simple example of weighting is the distribution of males and females in a sample. If the percentage male respondents in a city is 25%, all its males receive a weight of 2 (calculated by dividing the actual share in population - 50% - by the observed share in the sample).

After doing this, the data is aggregated from city-level to county-level. This is done with a list of cities per county. The result of weighting the data and aggregating it to county-level is shown in Figure 1. As a result of weighting on age, gender and education, around 90% of the respondents lost, because only a small fraction of respondents filled out all extra questions. This results in counties without weighted data. For these counties, the unweighted data is used. In the map, stars mark the difference between weighted and unweighted counties. If counties have less than 10 respondents, weighting may not be very reliable because the opinions of specific subgroups will be based on very small samples. Also in these counties the unweighted score is used.

Mapping the results shows that there are differences between counties. Many urban counties are less positive about completing the fence along the Mexican border than the more rural counties. Especially around the notoriously liberal city of San Francisco. However, most rural counties have a limited number of respondents. So the data of these counties may be not very representative for their whole population. Moreover, the percentages between the most positive and the most negative counties do not differ very and fall within the range of 0% to 7.5%. With a maximum score of 7.5% most of the inhabitants of California are negative about completing the fence along the Mexican border.

If we have a look at the research there are some aspects that can be improved in the future. Firstly, it would be interesting to have data for all counties. This could be achieved by further promoting the use of Kieskompas to increase the chances of having respondents in all counties. This could be done via the media partners (newspapers, television stations) Kieskompas cooperates with. Secondly, it may be better to weight data at county level rather than city level. This implies that weighting should be done after aggregating the data from city to county level. In the weighting process other personal characteristics such as religion or ethnicity could also be included. Finally, an attempt could be made to explain regional variation in opinions in a more formal, statistical analysis that includes both personal and spatial characteristics. The presented analysis highlights only one of the many opinions that can be mapped based on the extensive election database of Kieskompas.

Popular article on research project national GI-minor
Marieke Statema- February 2016

Figure 1: Spatial distribution of agreement with proposition *'To prevent illegal immigration, the US should complete the fence along the entire Mexican border'* per county in California based on 2008 presidential election questionnaire of Kieskompas