

Measuring Anti-Immigrant Sentiment Using Big Data Techniques



CENTER FOR STUDIES IN
DEMOGRAPHY & ECOLOGY

Michelle L. O'Brien – Department of Sociology
University of Washington
shanml@uw.edu

Background


After the collapse of the Soviet Union, a growing body of research has examined the evolution of anti-immigrant sentiment in the Russian Federation. Fears of nationalist mobilization and ethnic conflict materialized in many cases during the Glasnost' period, and that reality generated public anxiety about migration policy in the former Soviet bloc. Claims that President Vladimir Putin had co-opted a neo-nationalist platform gained momentum during the crisis in Crimea circa 2014.

Yet, anti-immigrant sentiment is notoriously difficult to measure (Butz and Kehrberg 2016), especially in diplomatic speeches in which language is often coded. Nevertheless, anti-immigrant sentiment, mobilization, and policies are typically linked to outcomes like broader ethnic conflict (Flores 2014), ethnic segregation (Della Posta 2013), family policy (Kingsbury 2015), welfare reform (2013) and even trends in gun sales (Flores 2015).



Data and Methods

This study utilizes the English-language archival material of online Russian newspaper *Pravda.ru*. This media source has been accused of distributing pro-government propaganda. Archives begin in January 2000 and end in November 2016. Not every day contains an article, but every month contains at least one day of archival material. Articles collected did not exclude opinion pieces. The text was collected and parsed with Python (visit <https://github.com/OBMich/RuMediaAnalysis> for details). The resulting corpus contained a maximum of 500 lines of text for over 6000 dates. For each date, I consider the text of all articles for that day. This means that each sentiment score is calculated based on an entire issue of *Pravda.ru* instead of simply one author's writing style or opinion piece. While more recent news consumption has been transformed by the prevalence of single-article sharing on social media, most of the archival years included occur before internet penetration reached 50% in Russia (starting from 2% in 2000).



Father of all Turkmen, President Saparmurat Niyazov, dies of sudden cardiac arrest
World » Former USSR

Turkmenistan's authoritarian President Saparmurat Niyazov, who has ruled the gas-rich Central Asian country for 20 years, died early Thursday. The 66-year-old leader had died of a sudden cardiac arrest. The TV ran still images of a national flag in a black-bordered frame.

```

150 <div class="date"><div>
151 <a href="/world/">World</a> &raquo; <a href="/world/ussr/">Former
152 USSR</a>
153 <a href="/world/ussr/21-12-2006/86035-niyazov-0/" class="articleAn
154 ons">
155 Turkmenistan's authoritarian President Saparmurat Niyazov, who
has ruled the gas-rich Central Asian country for 20 years, died early Thursday.
The 66-year-old leader had died of a sudden cardiac arrest. The TV ran still
images of a national flag in a black-bordered frame. </a>

```

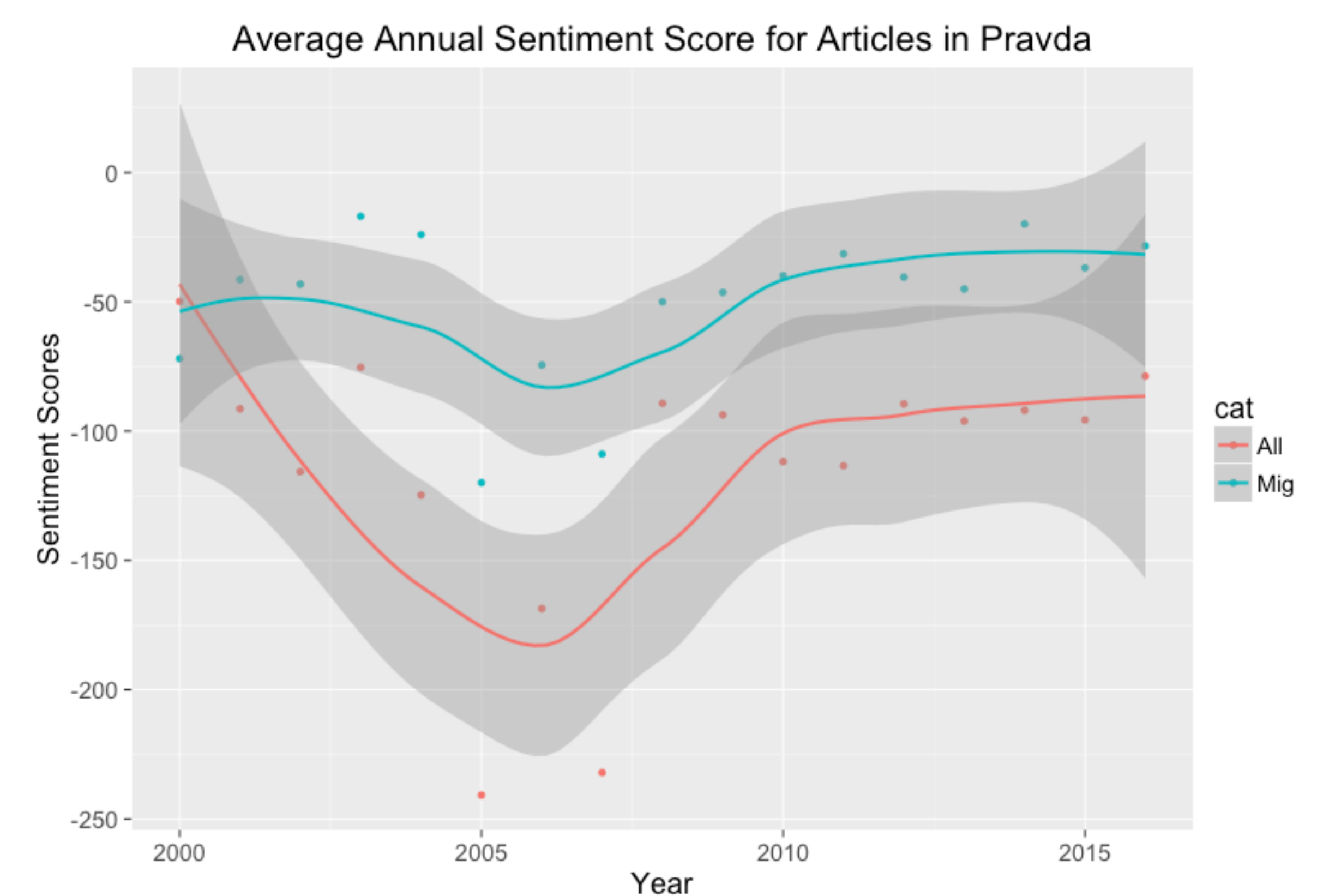
I used sentiment analysis to score articles. I used a list of over 2000 positive-scoring unigrams, such as "accomplish", "aspire", "advocate", "exonerate", "improve", "inspire", "success", "uphold", and "virtuous". Negative-scoring words, of which I use nearly 4800 unigrams, include "abolish", "abuse", "antagonize", "backwardness", "chide", "fanatic", "irreparable", "resent", "upheaval", and "zealot". Full lists are included in the documentation on the aforementioned github repository. Each date was given a sentiment score, and these scores were averaged (mean) over the calendar year.

Approximately 5% (338 days) of the corpus contains one or more migration-related search terms (listed below). I compare the baseline (all articles) to the migration-related corpus in order to measure anti-immigrant sentiment *relative to* an overall media sentiment.

Migration-Related Search Terms						
migration	migrant	migrated	diaspora	refugee	asylum seeker	asylum seeking
seeking asylum	Emigration	immigration	emigrated	immigrated	emigrant	immigrant

Findings

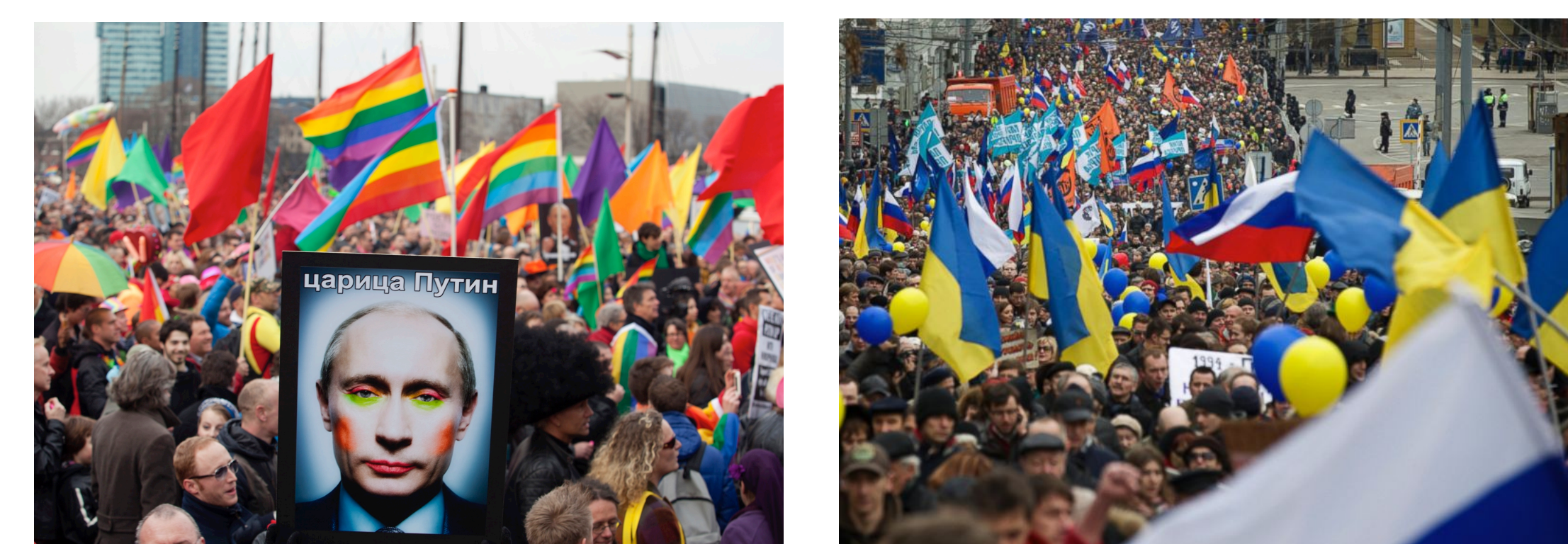
The findings are contrary to popular belief that anti-immigrant sentiment has been rising during the Putin administration. In most years, the average annual sentiment scores for migration-related topics were *less negative* compared to the baseline scores for all articles. It is worth noting here that all average annual scores are negative. This means that while we do measure anti-immigrant sentiment in Russian media, it is no more negative than sentiment around non-migration related topics.



Including this baseline is an important step for social scientists analyzing text data, particularly in the age of big data. Tracking the trends in anti-immigrant sentiment in Russia without a baseline measure of overall sentiment may result in misleading findings. Social scientists should consider a random sampling of the population in which their data live in order to create a baseline for comparison.

Discussion

One potential implication of this study is the impetus of contextualizing sentiment analysis in macro political and economic events. A clear decline in sentiment occurs in Pravda around 2006, but recovers to previous levels by 2010. A closer examination of global politics and a sampling of the articles during this time frame may reveal factors in negative news reporting. For example, in the 18 months leading up to U.S. President Obama's first term, perhaps *Pravda* ramped up negative reporting.



When we contextualize anti-immigrant sentiment in this case, we must consider the political climate in Russia. Whereas migrants certainly exist in a vulnerable social position, the range of groups targeted by negative media reporting in contemporary Russia is extensive. From the so-called 'Propaganda' Laws against public discussion of homosexuality, to the arrest and imprisonment of Pussy Riot members for 'hooliganism', and the annexation of Crimea, several issues that do not fall under migration-related terms have been gaining media attention during Putin's administration. Researchers not considering a baseline in this political climate may overestimate the relative anti-immigrant sentiment in the media.

References

Brown, Hana E. 2013. "Race, Legality, and the Social Policy Consequences of Anti-Immigration Mobilization." *American Sociological Review* . 78:2, p. 290-314.

Butz, Adam M. and Jason E. Kehrberg. 2016. "Estimating anti-immigrant sentiment for the American states using multi-level modeling and post-stratification, 2004-2008." *Research and Politics* . 3:2, p. 1-6.

Della Posta, Daniel J. 2013. "Competitive Threat, Intergroup Contact, or Both? Immigration and the Dynamics of Front National Voting in France." *Social Forces* . 92:1, p. 249-273.

Flores, Rene D. 2014. "Living in the Eye of the Storm: How Did Hazleton's Restrictive Immigration Ordinance Affect Local Interethnic Relations?" *American Behavioral Scientist* . 58, p. 1743-1763.

Kingsbury, Marina A. 2015. "Let's Have More Russian Babies: How Anti-Immigrant Sentiment Shapes Family Policy in Russia." *Council for European Studies Conference Proceedings*.