Codebook

main_data.RDS

Election results by electoral precinct and each election. Each observation is electoral precinct (UIK) by each election. Unless stated otherwise, the source of each variable is Ukraine's Central Election Commission (UCEC).

- *name* is name of electoral precinct according to the nomenclature used starting with 2012 elections. The results of earlier elections (2004 to 2010) were merged by spatial proximity.
- *uik* is the name of the precinct according to its original nomenclature; starting with 2012, *uik* takes the same value as *name*.
- *election* is the date of election.
- *size* is the classification of electoral precinct according to its size: small (mala), intermediate (serednya), large (velika).
- *address* is address of the voting station associated with the precinct.
- *lon* and *lat* are the longitude and latitude of the voting station associated with the precinct. These correspond to the locations in 2014. We did not geo-code this information, but rather took the coordinates from the website of the UCEC.
- *district* is the electoral district.
- *oblast* is the oblast in which the precinct is located.
- *ADM1_UA*, *ADM2_UA*, *ADM2_PCODE*, *ADM3_UA*, *ADM3_PCODE* are codes or names of level 1 (oblast), level 2 (rayon), and level 3 (rada) administrative units. Source: https://data.amerigeoss.org (accessed on 2017/12/21). These were merged with election data by geography. *ADM1_UA* in some cases does not coincide with *oblast*, because of errors in geographic matching (*oblast* variable is more precise).
- *voters* is the number of registered voters in the precinct (in some elections this is actually the number of ballots issued in the precinct).
- *votes* is the number of votes cast (including invalid ballots).
- *pro_russian* is the number of votes for what the paper refers as 'pro-Soviet' candidates/parties. See Appendix of the paper for which parties and candidates are included in this set.
- *nationalist* is the number of votes for the "nationalist" candidates (available for 2012 and 2014 presidential election only); the the paper and the appendix for further details.
- *ru_votes* is *pro_russian/votes* in percentages.
- *ru_turnout* is *pro_russian/voters* in percentages.
- *nationalist_turnout* is *nationalist/voters* in percentages.
- *n* is the number of Lenin's monuments in the precinct at the time of election. These were merged geographically: the lon-lat of the statue is inside the polygon of the precinct. Source: see the paper on the sources of data for statues.
- *dist_to_kiev* is the kilometers to kiev from the center of the precinct.

- *km_road* is the length of the road that falls within the precinct created by overlaying road network data from OpenStreetMap (dated March 8, 2016) with precinct polygons.
- *time* and *cycle* are convenience variables that simply rank each election by date.

protest.RDS

Protest events which happened in Ukraine during Euromaidan (21 November 2013 – 23 February 2014). Each observation is a protest event in specific location. The data was gathered from mass media archives and supplemented by mass media and social media posts.

-*location* is the name of an administrative unit (city, village), where the protest event took place.

-*participants* is the number of participants of the protest event in specific location.

-*source* is the link to the mass media article describing the protest event.

-*start_date* is the date of the beginning of the protest event.

-*end_date* is the date of the end of the protest event.

-*add_source* is the link to the mass media or social media post which validates information about the protest event.

-*add_source_2* is the link to the mass media or social media post which adds additional information about the protest event.

-*geometry* is the longitude and latitude of protest activities in specific location.

statues.RDS

All Lenin statues which were built at the territory of Ukraine after the October Revolution 1917. The data was collected from user-generated list at leninstatues.ru. Each observation is an individual Lenin statue.

-*id* is the id number of each statue generated for analysis purposes.

-*url* is the link to the individual webpage of the statue at leninstatues.ru.

-*removedby* is the classification of the entity which demolished the statue. “Sov” denotes Soviet authorities, “activists” denotes individuals who removed the statue without official authorization, “gov” denotes Ukrainian government.

-*source_of_update* is the link to the mass media or social media post which updates information from leninstatues.ru.

-*source* is the original source of information at leninstatues.ru.

-*img_url* is the link to the image of the statue.

-*sculptor* is the artist who created the statue.

-*tags* is the additional information about the statue, such as standing position, extended hand, removal status, condition of the pedestal, etc. This variable collects tags about the statue as seen at leninstatues.ru.

-*demolished* is the demolition status of the statue. TRUE denotes that the statue was removed from its location. FALSE denotes that the statue was not removed.

-*region* is the oblast where the statue is located.

-*raion* is the district where the statue is located.

-*location* is the administative unit (city, village) where the statue is located.

-*date_built* is the year when the statue was built.

-*date_demolished* is the date when the statue was demolished.

-*description* is the additional description of the statue removal as seen at leninstatues.ru.

-*year* is the year of statue removal.

-*month* is the month of statue removal.

-*day* is the day of statue removal.

-*date_ymd* is the date of statue removal in yyyy-mm-dd format.

-*cycle* is the convenience variable that shows in which electoral cycle the statue was removed.

-*lon* and *lat* are the geocoordinates of the statue.

-*Problematic* is the variable which indicates whether the date of statue removal is known precisely.

-*yandex_address* is the address of the statue, as seen at Yandex.Maps (these are the most precise digital maps in Ukraine).

-*yandex_lon* and *yandex_lat* are the geocoordinates of the statue, as seen at Yandex.Maps.

-*geo_flag* is the variable which indicates whether the statue location is known precisely.

stm_text.RDS

New articles used in structural topic analysis. These articles were scraped from eight news websites (see variable *site* in the dataset) and this dataset contains only articles that had word "Lenin".

-*site* the website from which articles were scraped.

-*date* is the date of the article.

-*text* stemmed text of the article with stop words removed.

-*day* is the day counter

-*period* is one of the three periods (the independent variable in our analysis)

twitter.RDS

Geo-locations of tweets from December 2013 to the end of February 2014 obtained from Twitter API. Due to the terms of use we did not have access to full text of the tweets.

-*lon* and *lat* are the geocoordinates of the tweet.

-*date* is the date.

-*followers_count* is the number of followers that the tweeting account has.

-*include_maidan* is an indicator whether the tweet contains the word "Maidan" or its derivatives.

uik_polygons.RDS

This is the shapefile of electoral precincts obtained from the central election commission (year 2019).

UKR_adm1.RDS

This is the shapefile of Ukraine's administrative oblasts. Source: https://data.amerigeoss.org (accessed on 2017/12/21).