Codebook
Aditi Malik. 2023. Kenya Violent Elections Dataset (KVED), 1991-2015.

Missing values throughout the dataset are represented by .

Violent events that occurred or spread across multiple places are coded as such using a slash for the descriptive geographic variables in the dataset. These descriptive variables are: province, district, division, county, city, village, area, place, old_constituency, and new constituency. As such a violent event that affected both the Rift Valley and Nyanza is coded as Rift Valley/Nyanza for the province variable

year = year of violent event

month = month of violent event

month_num = month of violent event (by month number)
Note: If a violent event occurred across multiple months, the first month is reported in month_num

day = date of violent event
Note: If a violent event occurred across multiple days, the first day is reported in day

elec_num = broad election period in which violent event occurred (1992 = 1; 1997 = 2; 2002 = 3; 2007 = 4; 2013 = 5)

country = Kenya (for all observations)

province = province where violent event occurred

province_id = numerical province identifier (Coast = 1; Northeastern = 2; Eastern = 3; Central = 4; Rift Valley = 5; Western = 6; Nyanza = 7; Nairobi = 8)

district = district where violent event occurred

division = division where violent event occurred

county = county where violent event occurred
Note: county information is provided for all violent events – including those before 2013 – where counties could be accurately determined

county_id = numerical county identifier (as per the Kenya Election Database)

city = city where violent event occurred

village = village where violent event occurred

area = (broad) area where violent event occurred
Note: The reports used to develop KVED sometimes referred to broad areas of violence rather than specific cities or villages.

place = (specific) place where violent event occurred
Note: The reports used to develop KVED sometimes referred to specific sites of violence (e.g., particular farms in the Rift Valley)

old_constituency = constituency in which violent event occurred at the time that it was recorded (string variable)

old_constituency_id = numerical constituency identifier at the time of the occurrence of violence (as per the Kenya Election Database – numeric variable)

new_constituency = constituency in which violent event occurred as per the 2010 constitution (string variable)

new_constituency_id = numerical constituency identifier as per the 2010 constituency (and as listed in the Kenyan Election Database – numeric variable)

groups = groups involved in violence. Multiple sets of antagonists are separated by semicolons

deaths = number of people killed
Note: If fatalities figures are provided in a range (e.g., 3-5 people or more than 10 people, then the lowest number is reported for deaths).

deaths_bin = binary variable for deaths (0 = no deaths)

injuries = number of people injured
Note: If the number of individuals injured is reported in a range (e.g., 3-5 people or more than 10 people, then the lowest number is reported for injuries).

displacements = number of people displaced
Note: If the number of individuals displaced is reported in a range (e.g., 10-20 people or more than 20 people, then the lowest number is reported for injuries).

arrests = number of people arrested
Note: If the number of individuals arrested is reported in a range (e.g., 10-20 people or more than 20 people, then the lowest number is reported for arrests).

duration = duration of violence (in days, months, or years – string variable)

duration_days = duration of violence in days (numeric variable)

rural = binary variable for rural location (0 = urban; 1 = rural)

source = name of report from which event is coded

source_date = date of source

timing = timing of event (pre-election; inter-election; post-election)
Note: If an event spread across multiple periods, these periods are separated by semicolons

precipitant = reported trigger for violent event
Note: Multiple reported precipitants for a single event are separated by semicolons

elite_involvement = binary variable for direct or indirect involvement in a violent event (0 = no elite involvement)
Note: Elite involvement was inferred through a close reading of qualitative information in source materials about each violent event