UNIVERSITY COLLEGE UTRECHT

Conflict in Bosnia in the Year 1993

Roos Visser

January 16, 2020

Abstract

Key words:

CONTENTS

1 *Introduction*

The research question to answer in this report is: 'What was the total number of fatalities in Bosnia in the year 1993? Interesting because this is during the Bosnian War.

2 *Methods*

Data is provided by The Uppsala Conflict Data Program. Data from this source is widely used in academic research, and has been prepared for that use, both in its definitions and in its reliability. Data set we are looking at shows one violent incident per entry with over 40 properties for that incident. Data runs till 2015. The focus of the dataset is conflict dynamics and the effects of armed violence, in the form of deaths. The dataset is constructed in such a way as maximise the comparability and consistency across time and space, and provide a globally consistent image of the phenomenon of organised violence. It is a tool for global understanding of subnational conflict patterns and trends.

Exploring the full dataset with PowerShell showed that in 1993 in Bosnia there were a total of 671 conflict events. Between the years 1989 and 2015 there were a total of 1946 conflict events which seemed like a significant enough amount to look further into the separate years.

Select Bosnia 1993 data from full dataset and create a smaller dataset with all necessary information for research question (json file).

Loaded reduced json file into python. Determine which columns are needed in order to answer the research question. I decided the columns best, high, and low will be necessary to answer the research question. These variables represent the best estimate of fatalities, the highest reliable estimate of fatalities, and the lowest reliable estimate of fatalities. The country and year do not have to be defined in the csv file since the entire data set is about Bosnia in 1993, so we already know this. created a csv file with the pre-established columns.

loaded the csv file into RStudio.

Created a new table in R that includes the sum of all fatalities for the best estimate, the highest reliable estimate, and the lowest reliable estimate. With new table created two ggplots: a scatterplot with an error bar over it. The minimum whisker of the error bar represents the total lowest reliable estimate, the maximum number the total highest reliable estimate. The

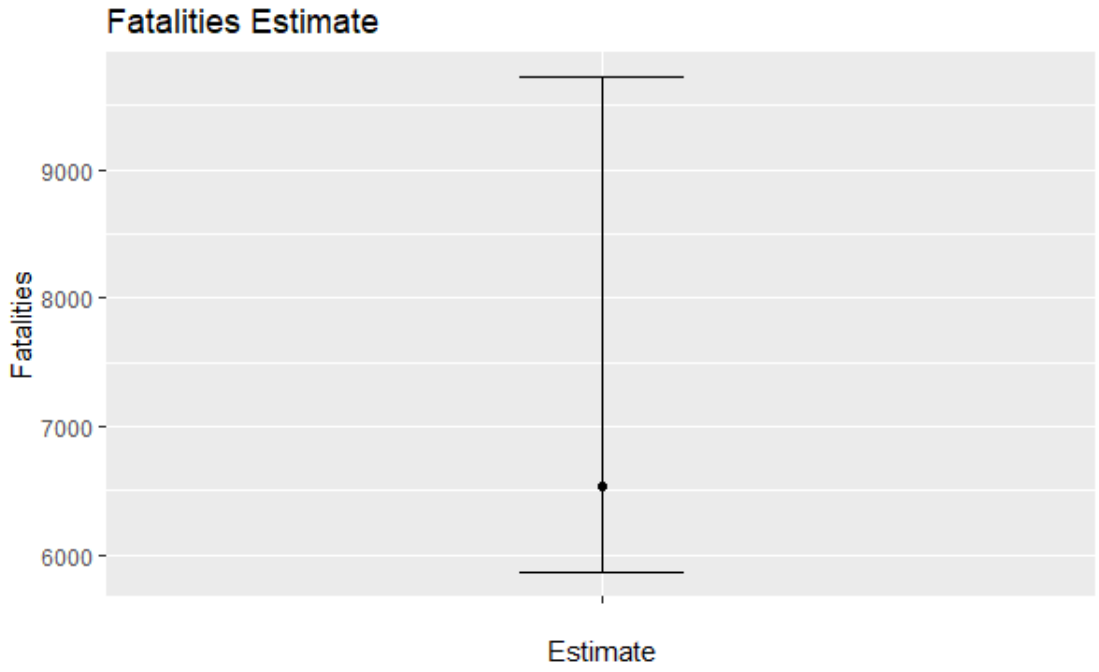
scatterplot consists of one point that represents the total best estimate.

Link to GitHub repository: here is the linky text

3 *Results*

Table 3.1: Total Fatalities

Best Estimate	Highest Reliable Estimate	Lowest Reliable Estimate
6536	9716	5866



The plot and table show that the best estimate of fatalities in Bosnia in 1993 is 6536. This is more towards the lowest reliable estimate than the highest reliable estimate.

4 *Discussion*

Reiterate pattern that found in data, and link back to research question. Discuss limitations, what can be concluded, and what cannot be concluded?

References