STK-INF4000 Group Project

Datascience Warriors 2017

February 19, 2017

Abstract

A short summary of our goal and how well we achieved it.

1 Introduction and Motivation

2 About ACLED

ACLED (Armed Conflict Location and Event Data Project) is designed for disaggregated conflict analysis and crisis mapping. This dataset codes the dates and locations of all reported political violence and protest events in over 60 developing countries. Political violence includes events that occur within civil wars and periods of instability.

ACLED collects real-time data on political violence and protest in both African and Asian states. Real-time data is available on this website and at the Climate Change and African Political Stability Project (CCAPS) website. All ACLED data can be downloaded on this site by country in excel and GIS formats. While the information is designed for disaggregated conflict analysis and crisis mapping, these data can be used in any GIS, mapping program, or statistical package.

3 The Africa Real-time Dataset

Real-time ACLED data is available as .csv and .xlsx files from their webpage http://www.acleddata.com/data/realtime-data. For the Africa real-time dataset, each data row contains the following information (tags in order):

GWNO, EVENT_ID_CNTY, EVENT_ID_NO_CNTY, EVENT_DATE, YEAR, TIME_PRECISION, EVENT_TYPE, ACTOR1, ALLY_ACTOR_1, INTER1, ACTOR2, ALLY_ACTOR_2, INTER2, INTERACTION, COUNTRY, ADMIN1, ADMIN2, ADMIN3, LOCATION, LATITUDE, LONGITUDE, GEO_PRECISION, SOURCE, NOTES, FATALITIES

In particular, $YEAR$ ranges from 2001 to 2017, and $EVENT\_ID\_NO\_CNTY$ appears to be an empty column (possibly since this is all taken from Africa). The data is real-time in the sense that it is collected and published weekly.

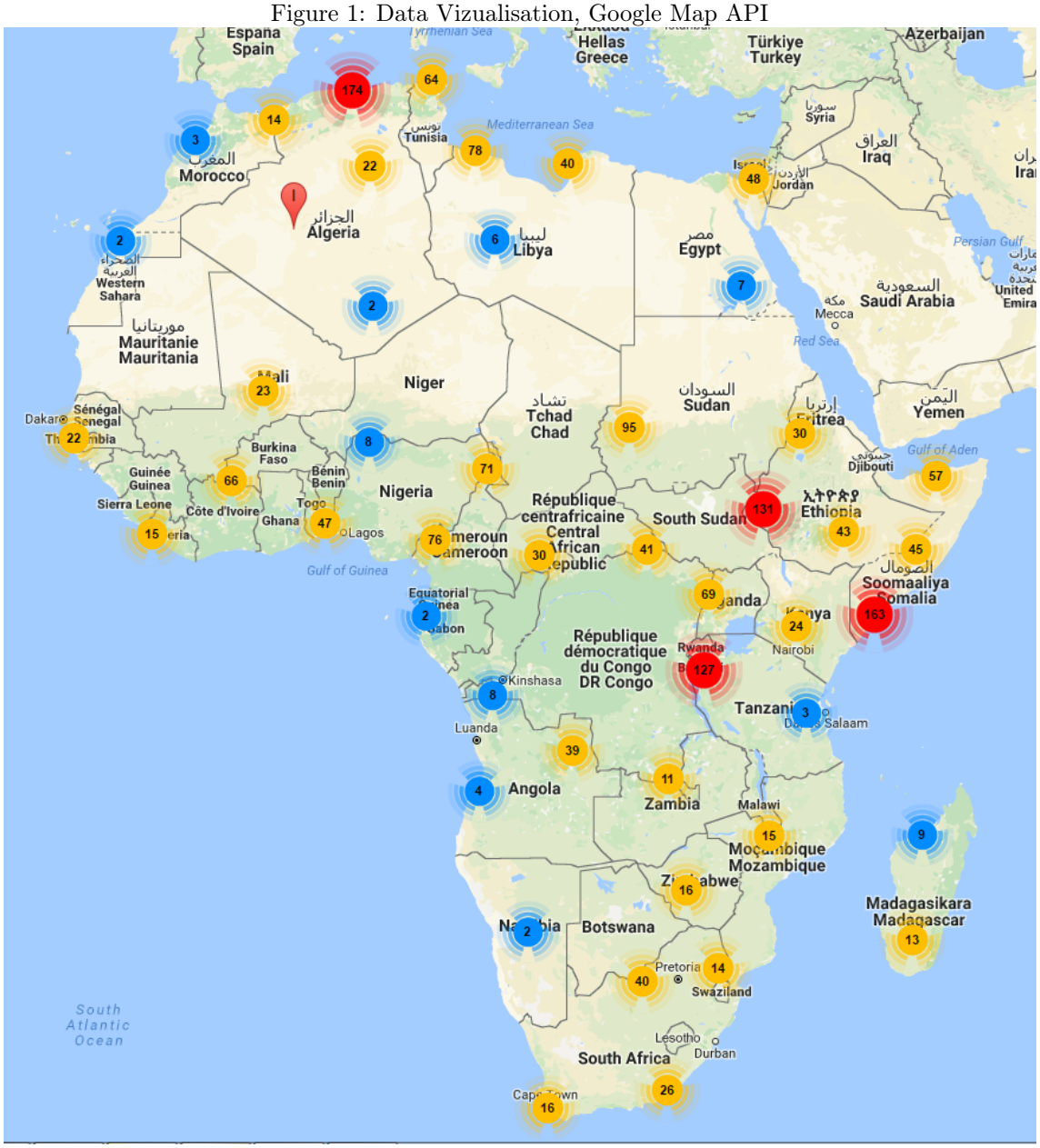
3.1 Preliminary Analysis

3.1.1 Visualizing location data using Google Map API

API Documentation webpage: https://developers.google.com/maps/documentation/javascript/. In figure 1 is an image generated using the Google Map API from the geo-location data $LATITUDE$, $LONGTITUDE$ from the edited latlongdata.csv file (code: webapp2.html). As is clear from the image, there are a few areas nearly devoid of conflict data, namely in the Democractic Republic of Congo, and "desert" terrain areas. The location data is clustered (and zoomable in the app).

We can use the location data to train a Neural Network (of at least three linear neurons) to predict whether or not a given latitude, longitude is likely to be an area of conflict. We present

Figure 1: Data Vizualisation, Google Map API



some very basic statistics of the geo-location data:

$$\text{mean(latitude)} \approx 10.2, \tag{1}$$
$$\text{mean(longitude)} \approx 22.6, \tag{2}$$
$$\text{std(latitude)} \approx 16.9, \tag{3}$$
$$\text{std(longtitude)} \approx 15.9. \tag{4}$$

References