

# Analysis of Ebola data from HealthMap

*true*

*06 December, 2017*

## Introduction

## Parameters for Ebola and Reproduction Number Estimation

Culled from literature

## Gravity model parameters

## Data clean-up

Visualising the raw data.

The data clean-up consists of the following steps: 1. Extract the cumulative case count as a sum of suspected and confirmed cases. 2. Merge duplicate alerts. 3. Remove outliers and interpolate missing data. 4. Determine the incidence count from the cumulative case count.

We now use the incidence count to estimate reproduction number.

```
## [[1]]
## NULL
##
## [[2]]
## NULL
##
## [[3]]
## NULL
```

We assume that the reproduction number remains unchanged for the time period over which we wish to project. For each location, distribution of  $r_t$  at  $t_{proj}$  is  $r_t$  over the next  $n_{days}$ .sim.

Determine the flow matrix for the countries of interest only.

At this point, all the pieces are in place. `by.location_incid` contains the incidence count `r.j.t` contains the estimates of reproduction numbers. `p.movement` contains the probabilities. `SI_Distr` is the serial interval distribution. The model is:  $\lambda_{j.t} = p.movement * (incidence * r_t) * serial\_interval$  taking care of the dimensions of course. Now divide the dataset into training and validation sets.

We will now split our data into training and validation sets.