Geocoding The EMDAT Dataset

Ram Mukund Kripa

4/15/2021

## Abstract

## Keywords

## Introduction

## Structure of The Package

The structure of the Wrangler for Emergency Events Database is best explained through a sample workflow.

### Loading

First, one may download the subset of EM-DAT required from the public EM-DAT query tool located at [this link](https://public.emdat.be/). From here, researchers can exploit the loading functionality of WEED. EM-DAT files tend to be excel sheets with a few lines of metadata, followed by the subsection of the dataset downloaded. The read\_emdat function allows researchers to load the dataset as a List in one of two ways: the dataset, as well as its metadata, or just the dataset.

sample\_data <- read\_emdat(here("data", "India\_SriLanka.xlsx"))

## New names:  
## \* `` -> ...1  
## \* `` -> ...2

summary(sample\_data)

## Length Class Mode  
## file\_data 2 tbl\_df list  
## disaster\_data 43 tbl\_df list

Sample Data

Dis No

Year

Seq

Disaster Group

Disaster Subgroup

Disaster Type

Disaster Subtype

Disaster Subsubtype

Event Name

Entry Criteria

Country

ISO

Region

Continent

Location

Origin

Associated Dis

Associated Dis2

OFDA Response

Appeal

Declaration

Aid Contribution

Dis Mag Value

Dis Mag Scale

Latitude

Longitude

Local Time

River Basin

Start Year

Start Month

Start Day

End Year

End Month

End Day

Total Deaths

No Injured

No Affected

No Homeless

Total Affected

Reconstruction Costs (’000 US$)</th> <th class="gt\_col\_heading gt\_columns\_bottom\_border gt\_right" rowspan="1" colspan="1">Insured Damages ('000 US$)

Total Damages (’000 US$)

CPI

1900-9001-IND

1900

9001

Natural

Climatological

Drought

Drought

NA

NA

NA

India

IND

Southern Asia

Asia

Bengal

NA

NA

NA

NA

No

No

NA

NA

Km2

NA

NA

NA

NA

1900

NA

NA

1900

NA

NA

1250000

NA

NA

NA

NA

NA

NA

NA

3.261389

1905-0003-IND

1905

0003

Natural

Geophysical

Earthquake

Ground movement

NA

NA

Kill

India

IND

Southern Asia

Asia

Kangra

NA

NA

NA

NA

NA

NA

NA

8

Richter

32.04

76.16

06:20

NA

1905

4

4

1905

4

4

20000

NA

NA

NA

NA

NA

NA

25000

3.522300

1907-0001-IND

1907

0001

Natural

Biological

Epidemic

Bacterial disease

NA

Bubonic

Kill

India

IND

Southern Asia

Asia

NA

NA

NA

NA

NA

NA

NA

NA

NA

Vaccinated

NA

NA

NA

NA

1907

NA

NA

1907

NA

NA

1300000

NA

NA

NA

NA

NA

NA

NA

3.652756

1916-0004-IND

1916

0004

Natural

Meteorological

Storm

Tropical cyclone

NA

NA

Kill

India

IND

Southern Asia

Asia

Cuddalore, Pondicherry

NA

NA

NA

NA

NA

NA

NA

NA

Kph

NA

NA

NA

NA

1916

11

21

1916

11

21

300

NA

NA

NA

NA

NA

NA

NA

4.265897

1920-0001-IND

1920

0001

Natural

Biological

Epidemic

Bacterial disease

NA

Bubonic

Kill

India

IND

Southern Asia

Asia

NA

NA

NA

NA

NA

NA

NA

NA

NA

Vaccinated

NA

NA

NA

NA

1920

NA

NA

1920

NA

NA

2000000

NA

NA

NA

NA

NA

NA

NA

7.853425

1920-0002-IND

1920

0002

Natural

Biological

Epidemic

Bacterial disease

NA

Cholera

Kill

India

IND

Southern Asia

Asia

NA

NA

NA

NA

NA

NA

NA

NA

NA

Vaccinated

NA

NA

NA

NA

1920

NA

NA

1920

NA

NA

500000

NA

NA

NA

NA

NA

NA

NA

7.853425

### Exploration

The next step might be to explore the given data. One of the most pressing issues with EM-DAT at the moment is its geocoding data. The fact that a very small fraction of disasters have usable geocoding data, in terms of Latitude and Longitude, severely hampers location analysis. Another issue is the presence of multiple locations per disaster.

sample\_df <- sample\_data[['disaster\_data']] %>%  
 sample\_n(10) %>%  
 select(`Dis No`, Year, Country, `Disaster Type`, Location, Latitude, Longitude) %>%  
 filter(!is.na(`Dis No`))

## Warning: The `.dots` argument of `group\_by()` is deprecated as of dplyr 1.0.0.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_warnings()` to see where this warning was generated.

Sample Data

The problems

Dis No

Year

Country

Disaster Type

Location

Latitude

Longitude

2001-0337-IND

2001

India

Storm

Uttar Pradesh province

NA

NA

1974-0122-IND

1974

India

Storm

Midnapore

NA

NA

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

1985-0380-IND

1985

India

Storm

NA

NA

NA

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

Having a single latitude and longitude refer to multiple locations also makes analysis significantly more challenging, and nigh on impossible. To counteract these problems, the recommended Weed workflow is to change the unit of analysis of the data frame from “one row per disaster” to “one row per disaster-location pair”. This process will henceforth be referred to as “locationizing”.

locationized\_df <- sample\_df %>%  
 split\_locations(column\_name = "Location")

Locationized Sample Data

Note the change of unit of analysis

Dis No

Year

Country

Disaster Type

Location

Latitude

Longitude

location\_word

2001-0337-IND

2001

India

Storm

Uttar Pradesh province

NA

NA

uttar pradesh

1974-0122-IND

1974

India

Storm

Midnapore

NA

NA

midnapore

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

bombay

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

assam

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

mekhliganj

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

haldibari

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

fukaldabri

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

nijtaraf

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

kasiabari

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

bholarhat

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

beltali areas

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

kochbihar

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

west bengal

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

new delhi

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

rajasthan

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

madhya pradesh

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

mannar

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

northern

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

ratnapura

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

sabaragamuwa

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

gampaha

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

western

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

puttalam

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

north western

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

uva

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

eastern

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

bihar

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

uttar pradesh

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

jharkhand

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

madhya pradesh

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

punjab

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

himachal pradesh

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

haryana

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

jamuu

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

kashmir

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

delhi

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

galle

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

matara

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

southern

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

badulla

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

uva

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

western

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

north western

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

sabaragamuwa

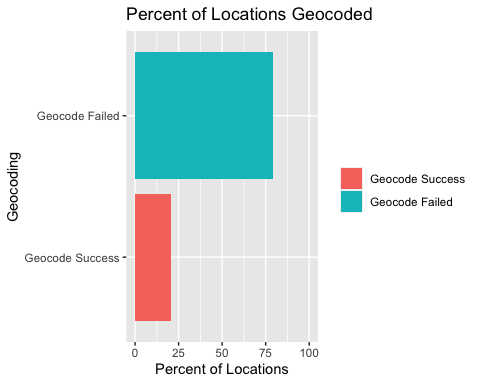
The split\_locations function allows users to execute the process of “locationizing”. It comes with a default method of splitting, defined by its parameters ‘dummy\_words’, which indicate which words to altogether remove from the location strings, and ‘joiner\_regex’, which indicates how the locations have been concatenated to form the location strings.

locationized\_sample\_data <- sample\_data[['disaster\_data']] %>%  
 split\_locations(column\_name = "Location")

The locationized Data frame is compatible with exploratory functions like percent\_located\_locations and percent\_located\_disasters which allow for easy visualization of the coverage provided in the data, with respect to latitudes and longitudes.

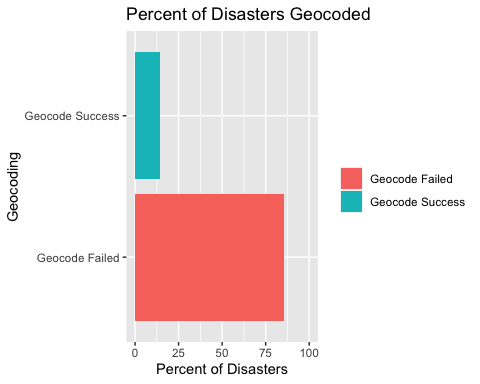
Percent Located Location-Disaster Pairings

locationized\_sample\_data %>%  
 percent\_located\_locations(lat\_column = "Latitude",  
 lng\_column = "Longitude")



Percent Located Disasters

locationized\_sample\_data %>%  
 percent\_located\_disasters(lat\_column = "Latitude",  
 lng\_column = "Longitude")



As we can see, th ecoverage is very sparse. Certainly not enough for proper analysis.

Once the data has been locationized, it is ready to be geocoded.

### Geocoding

Weed uses the free [geonames API](https://www.geonames.org/) to geocode each location. To use this functionality, one must first create a free account and then supply their username to the geocode function in Weed. This function comes with a few options, depending on the kind of analysis that is being performed. Researchers can utilize the n\_results parameter to get the n closest matches to the input location and decide which one to use. The unwrap parameter also allows researchers to keep the geocoded data in a nested Data frame structure, possibly good for exporting, or in unwrapped from, where each lat and long gets a separate column (lat1, lng1, lat2, lng2, etc.)

geocoded\_df <- locationized\_df %>%  
 geocode(unwrap = FALSE, geonames\_username = sample\_username)

## <error/rlang\_error>  
## Can't select within an unnamed vector.  
## Backtrace:  
## 1. rmarkdown::render(...)  
## 43. dplyr:::select.data.frame(., toponymName, lat, lng)  
## 44. tidyselect::eval\_select(expr(c(...)), .data)  
## 45. tidyselect:::eval\_select\_impl(...)  
## <error/rlang\_error>  
## Can't select within an unnamed vector.  
## Backtrace:  
## 1. rmarkdown::render(...)  
## 43. dplyr:::select.data.frame(., toponymName, lat, lng)  
## 44. tidyselect::eval\_select(expr(c(...)), .data)  
## 45. tidyselect:::eval\_select\_impl(...)  
## <error/rlang\_error>  
## Can't select within an unnamed vector.  
## Backtrace:  
## 1. rmarkdown::render(...)  
## 43. dplyr:::select.data.frame(., toponymName, lat, lng)  
## 44. tidyselect::eval\_select(expr(c(...)), .data)  
## 45. tidyselect:::eval\_select\_impl(...)  
## <error/rlang\_error>  
## Can't select within an unnamed vector.  
## Backtrace:  
## 1. rmarkdown::render(...)  
## 43. dplyr:::select.data.frame(., toponymName, lat, lng)  
## 44. tidyselect::eval\_select(expr(c(...)), .data)  
## 45. tidyselect:::eval\_select\_impl(...)  
## <error/rlang\_error>  
## Can't select within an unnamed vector.  
## Backtrace:  
## 1. rmarkdown::render(...)  
## 43. dplyr:::select.data.frame(., toponymName, lat, lng)  
## 44. tidyselect::eval\_select(expr(c(...)), .data)  
## 45. tidyselect:::eval\_select\_impl(...)  
## <error/rlang\_error>  
## Can't select within an unnamed vector.  
## Backtrace:  
## 1. rmarkdown::render(...)  
## 43. dplyr:::select.data.frame(., toponymName, lat, lng)  
## 44. tidyselect::eval\_select(expr(c(...)), .data)  
## 45. tidyselect:::eval\_select\_impl(...)

Geocoded Data

Dis No

Year

Country

Disaster Type

Location

Latitude

Longitude

location\_word

lat

lng

2001-0337-IND

2001

India

Storm

Uttar Pradesh province

NA

NA

uttar pradesh

26.46523

80.34975

1974-0122-IND

1974

India

Storm

Midnapore

NA

NA

midnapore

22.33971

87.32501

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

bombay

19.07283

72.88261

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

assam

26.00000

93.00000

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

mekhliganj

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

haldibari

26.33887

88.78220

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

fukaldabri

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

nijtaraf

26.31761

88.87669

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

kasiabari

26.24912

89.31505

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

bholarhat

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

beltali areas

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

kochbihar

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

west bengal

22.56263

88.36304

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

new delhi

28.65195

77.23149

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

rajasthan

26.91962

75.78781

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

madhya pradesh

23.25469

77.40289

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

mannar

8.86667

80.08333

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

northern

9.66845

80.00742

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

ratnapura

6.68580

80.40360

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

sabaragamuwa

6.75000

80.50000

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

gampaha

7.20830

79.83580

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

western

6.93548

79.84868

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

puttalam

8.03620

79.82830

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

north western

7.75000

80.16667

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

uva

6.98472

81.05639

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

eastern

8.57780

81.22890

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

bihar

25.75000

85.75000

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

uttar pradesh

26.46523

80.34975

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

jharkhand

23.34316

85.30940

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

madhya pradesh

23.25469

77.40289

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

punjab

31.62234

74.87534

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

himachal pradesh

31.91667

77.25000

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

haryana

29.25000

76.33333

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

jamuu

NA

NA

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

kashmir

34.08565

74.80555

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

delhi

28.65195

77.23149

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

galle

6.04610

80.21030

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

matara

5.94851

80.53528

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

southern

6.93548

79.84868

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

badulla

6.98020

81.05770

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

uva

6.98472

81.05639

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

western

6.93548

79.84868

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

north western

7.75000

80.16667

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

sabaragamuwa

6.75000

80.50000

Further exploration with the percent\_located\_locations and percent\_located\_disasters functions is advisable, to visualize the success of our geocoding. As the data is indeed locationized, a choice must be made as to how to decide if a disaster has been “located”. Two popular choices are any and all. “Any” considers a disaster located if any one of its constituent locations has valid lat-long data, while “all” requires every constituent location to be geocoded. These can be set by the how parameter of percent\_located\_disasters, which also allows user defined functions!

Percent Located Locations

geocoded\_df %>%  
 percent\_located\_locations(plot\_result = FALSE) %>%  
 gt() %>%  
 tab\_header(title = "Geocoding success rate") %>%  
 data\_color(columns = c("percent"), colors = "orange")

Geocoding success rate

coords\_nonexistent

count

percent

Geocode Success

38

86.36364

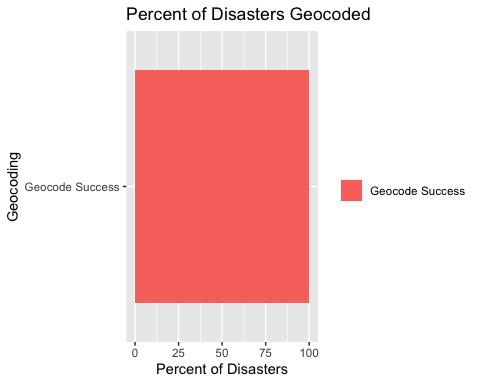
Geocode Failed

6

13.63636

Percent Located Disasters

geocoded\_df %>%  
 percent\_located\_disasters(how = "any")



The goal with the geocoding and subsequent exploration was to provide as much flexibility and modularity to this step of the workflow as possible, to allow for diverse analyses and use cases.

### Elementary Analysis

One of the most common uses of lat-long data is checking whether a point lies in some defined region. Weed allows for regions to be defined either as a lat-long box, or as a shapefile. For increased modularity, the shapefile may be defined as either a shape object or even as a string containing the file name.

Assume the required box is Lat(8 to 23) and Lng(80 to 90)

inbox\_df <- geocoded\_df %>%  
 located\_in\_box(top\_left\_lat = 23, top\_left\_lng = 80, bottom\_right\_lat = 8, bottom\_right\_lng = 90)

Lat Long Box Data

Dis No

Year

Country

Disaster Type

Location

Latitude

Longitude

location\_word

lat

lng

in\_box

2001-0337-IND

2001

India

Storm

Uttar Pradesh province

NA

NA

uttar pradesh

26.46523

80.34975

FALSE

1974-0122-IND

1974

India

Storm

Midnapore

NA

NA

midnapore

22.33971

87.32501

TRUE

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

bombay

19.07283

72.88261

FALSE

1990-0714-IND

1990

India

Flood

Bombay (Assam state)

NA

NA

assam

26.00000

93.00000

FALSE

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

mekhliganj

NA

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

haldibari

26.33887

88.78220

FALSE

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

fukaldabri

NA

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

nijtaraf

26.31761

88.87669

FALSE

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

kasiabari

26.24912

89.31505

FALSE

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

bholarhat

NA

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

beltali areas

NA

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

kochbihar

NA

NA

NA

2003-0234-IND

2003

India

Storm

Mekhliganj, Haldibari, Fukaldabri, Nijtaraf, Kasiabari, Bholarhat, Beltali areas (Kochbihar district, West Bengal province)

NA

NA

west bengal

22.56263

88.36304

TRUE

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

new delhi

28.65195

77.23149

FALSE

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

rajasthan

26.91962

75.78781

FALSE

1994-0624-IND

1994

India

Extreme temperature

New Delhi, Rajasthan, Madhya Pradesh

NA

NA

madhya pradesh

23.25469

77.40289

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

mannar

8.86667

80.08333

TRUE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

northern

9.66845

80.00742

TRUE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

ratnapura

6.68580

80.40360

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

sabaragamuwa

6.75000

80.50000

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

gampaha

7.20830

79.83580

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

western

6.93548

79.84868

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

puttalam

8.03620

79.82830

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

north western

7.75000

80.16667

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

uva

6.98472

81.05639

FALSE

2008-0132-LKA

2008

Sri Lanka

Flood

Mannar district (Northern province), Ratnapura district (Sabaragamuwa province), Gampaha district (Western province), Puttalam district (North Western province), Uva, Eastern provinces

7.83

80.82

eastern

8.57780

81.22890

TRUE

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

bihar

25.75000

85.75000

FALSE

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

uttar pradesh

26.46523

80.34975

FALSE

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

jharkhand

23.34316

85.30940

FALSE

2016-0237-IND

2016

India

Storm

Bihar, Uttar Pradesh, Jharkhand, Madhya Pradesh provinces

NA

NA

madhya pradesh

23.25469

77.40289

FALSE

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

punjab

31.62234

74.87534

FALSE

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

himachal pradesh

31.91667

77.25000

FALSE

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

haryana

29.25000

76.33333

FALSE

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

jamuu

NA

NA

NA

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

kashmir

34.08565

74.80555

FALSE

1988-0438-IND

1988

India

Flood

Punjab, Himachal Pradesh, Haryana, Jamuu and Kashmir, Delhi

NA

NA

delhi

28.65195

77.23149

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

galle

6.04610

80.21030

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

matara

5.94851

80.53528

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

southern

6.93548

79.84868

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

badulla

6.98020

81.05770

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

uva

6.98472

81.05639

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

western

6.93548

79.84868

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

north western

7.75000

80.16667

FALSE

2006-0602-LKA

2006

Sri Lanka

Flood

Galle, Matara districts (Southern province), Badulla district (Uva province), Western, North Western, Sabaragamuwa provinces

7.32

80.03

sabaragamuwa

6.75000

80.50000

FALSE

## Conclusion

## Acknowledgements

## References