

H1N1 Exploration

Jose Figueroa, Liam Donohoe

April 6, 2020

Contents

1	Introduction	1
2	Dataset Exploration	1

1 Introduction

The goal of this project is to create a visualization of the progression of past epidemic or pandemic viruses throughout their lifetimes. In doing this, we hope to be able to provide a better understanding of not only the rate at which these types of viruses can spread throughout the population, but also the time it takes for scientists to develop and distribute a vaccine or cure for the spreading virus.

2 Dataset Exploration

Dimensions:

```
## [1] 1822    4
```

Dataframe Summary:

```
str(H1N1.df)
```

```
## 'data.frame':    1822 obs. of  4 variables:
## $ Country      : Factor w/ 152 levels "\xa0\xa0\xa0\xa0French Polynesia, FOC",...: 19 20 21 22 23 ...
## $ Cases        : int   5 2 2485 5298 19 7 15 18 12 54 ...
## $ Deaths       : int   0 0 60 10 0 0 0 0 0 0 ...
## $ Update.Time: Factor w/ 22 levels "5/23/2009 8:00",...: 22 22 22 22 22 22 22 22 22 ...
```

```
summary(H1N1.df)
```

##	Country	Cases	Deaths	Update.Time
##	Argentina: 22	Min. : 1.00	Min. : 0.000	7/6/2009 9:00 : 136
##	Australia: 22	1st Qu.: 3.00	1st Qu.: 0.000	7/3/2009 9:00 : 126
##	Austria : 22	Median : 13.50	Median : 0.000	7/1/2009 9:00 : 121
##	Belgium : 22	Mean : 901.06	Mean : 4.316	6/29/2009 9:00 : 116
##	Brazil : 22	3rd Qu.: 88.75	3rd Qu.: 0.000	6/26/2009 7:00 : 113
##	Canada : 22	Max. : 94512.00	Max. : 429.000	6/24/2009 7:00 : 109
##	(Other) :1690		NA's :1	(Other) :1101

Head:

```
head(H1N1.df)
```

Country	Cases	Deaths	Update.Time
Algeria	5	0	7/6/2009 9:00
Antigua and Barbuda	2	0	7/6/2009 9:00
Argentina	2485	60	7/6/2009 9:00
Australia	5298	10	7/6/2009 9:00
Austria	19	0	7/6/2009 9:00
Bahamas	7	0	7/6/2009 9:00

Tail:

```
tail(H1N1.df)
```

	Country	Cases	Deaths	Update.Time
1817	Switzerland	2	0	5/23/2009 8:00
1818	Thailand	2	0	5/23/2009 8:00
1819	Turkey	2	0	5/23/2009 8:00
1820	United Kingdom	117	0	5/23/2009 8:00
1821	United States of America	6552	9	5/23/2009 8:00
1822	Grand Total	12022	86	5/23/2009 8:00

Unique Dates:

```
unique.dates <- levels(H1N1.df$Update.Time)
length(unique.dates)
```

```
## [1] 22
```

```
unique.dates
```

```
## [1] "5/23/2009 8:00" "5/25/2009 8:00" "5/26/2009 6:00" "5/27/2009 8:00"
## [5] "5/29/2009 6:00" "6/1/2009 6:00" "6/10/2009 6:00" "6/11/2009 14:00"
## [9] "6/12/2009 7:00" "6/15/2009 17:00" "6/17/2009 12:00" "6/19/2009 7:00"
## [13] "6/22/2009 7:00" "6/24/2009 7:00" "6/26/2009 7:00" "6/29/2009 9:00"
## [17] "6/3/2009 6:00" "6/5/2009 6:00" "6/8/2009 6:00" "7/1/2009 9:00"
## [21] "7/3/2009 9:00" "7/6/2009 9:00"
```

There are 22 unique weeks to account for. We should look into how to plot on a world map style heatmap (I have an idea on this), and then allowing for the option to traverse time on said heatmap. For this week, we can probably focus on having just 1 week plotted (5/23/2009). We can also easily transform this notebook into a webapp with RShiny