Assignment Four

Gerardo Palacios[1,2]

[1] DSC 430 - Data Visualization

[2] DePaul University

Assignment Four

Covid-19

Pre-processing was necessary in order to tune the plots appropriately. The field names
and data types are standardized. In addition, the stringency variable index was binned to
represent 4 different stringency levels as opposed to the index value. Finally, the data was
summarized and averaged on a weekly basis in order to portray the data in smoothed equal,
time intervals. The following visualizations are centered around the relationships between
the number of new cases or new deaths and a continents stringency response over the course
of the pandemic.

```r
# Import data
data <- read.table("cleandata_rm.csv", header=TRUE, sep=",")


# Convert to date types
data$date <- as.Date(as_date(data$date,format = '%Y-%m-%d'))


# Remove special characters
data$continent <- gsub("\n"," " , data$continent ,ignore.case = TRUE)


data <- data %>% mutate_if(is.character,as.factor)


# Bin the stringency index into 4 parts
data$stry_index <- as.factor(ifelse(
  data$stry_index <= 25, 'No Response' , ifelse(
   data$stry_index > 25 & data$stry_index <= 50 , 'Least Stringent', ifelse(
    data$stry_index > 50 & data$stry_index <= 75 ,'Moderately Stringent',ifelse(
     data$stry_index > 75 ,  'Most Stringent','None')))))
```

```r
# Define mode function
getmode <- function(v) {

   uniqv <- unique(v)

   uniqv[which.max(tabulate(match(v, uniqv)))]

}


# Copy standardized raw data set
data_raw <- data


# Pivot and summarize data by 7 days
data <- data %>% group_by(
  continent=continent,

  date=floor_date(date, "7 days"),


  ) %>%
   summarize(

     location = getmode(location),

     new_cases=mean(new_cases),

     new_deaths=mean(new_deaths),

     stry_index=getmode(stry_index),

     total_deaths=mean(total_deaths),

     total_cases = mean(total_cases),

     month = mean(month),

     population=mean(population))


# Add new date format
```

```r
data$date2 <- as.POSIXct(as.yearmon(data$date,

                   format = "%Y-%m-%d"),

                   format = "%m-%d-%Y")


# Create new variable

data$deaths_per_cases <- (data$total_deaths/data$total_cases) * 100

data_raw$deaths_per_cases <- (data_raw$total_deaths/data_raw$total_cases) * 100


# Manually apply log base 10 scale

data$new_cases <- log(data$new_cases+1,10)

data$new_deaths <- log(data$new_deaths+1,10)


data_raw$new_cases <- log(data_raw$new_cases+1,10)

data_raw$new_deaths <- log(data_raw$new_deaths+1,10)
```

Visualization One

In order to achieve the visualization below, the stringency index had to be binned in to 4 levels. Figure 1 is a mosaic plot showing the total proportion of stringency responses by continent. Each continent had varying response levels. Note that Asia and South America had the most stringent response while North America and Europe responded less urgently.
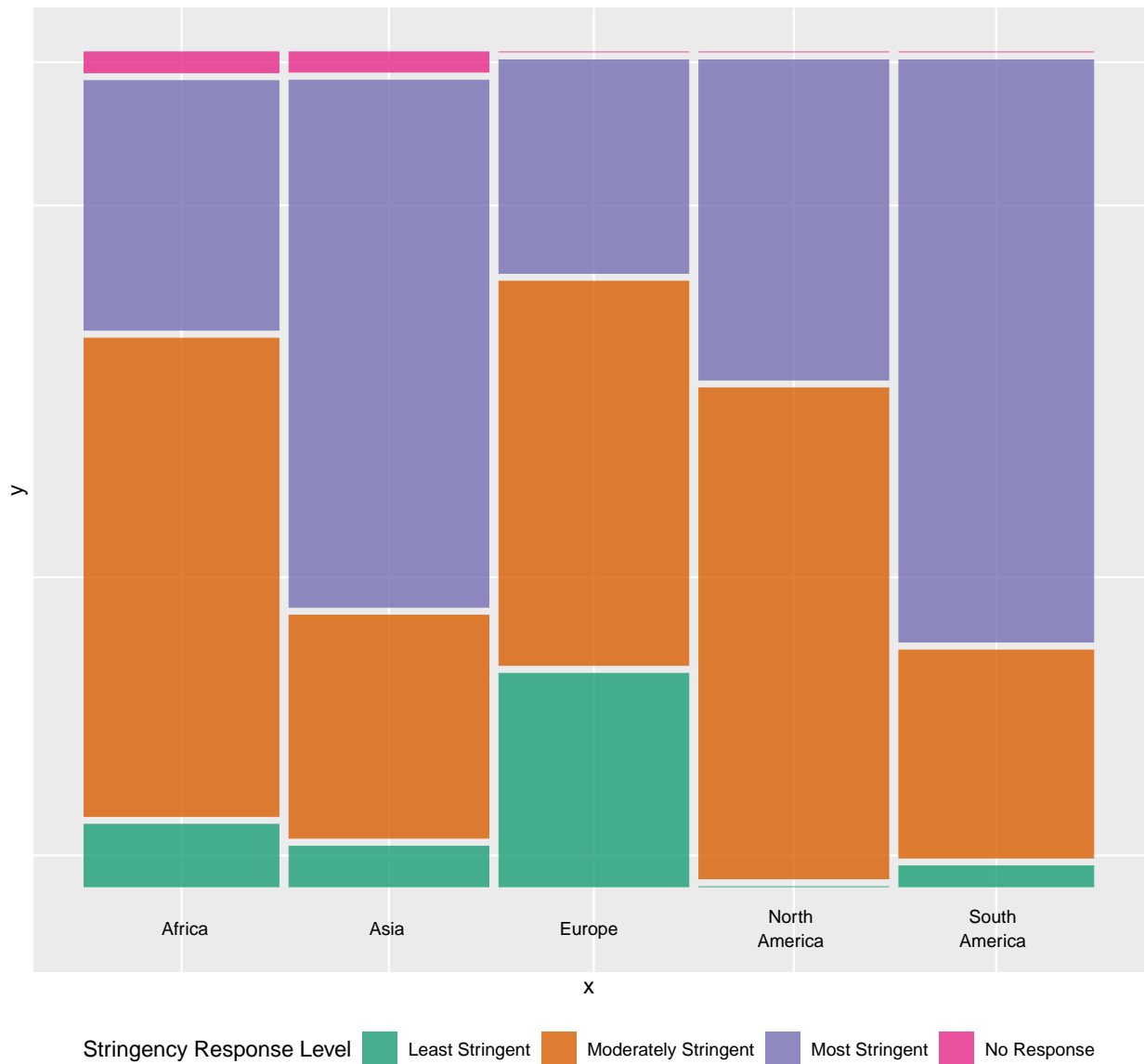

Figure 1. Covid 19 - Stringency Level vs Continent

Visualization Two

Figure 2 are the summarized weekly distributions for number of new cases by Continent and Stringency level. Continuing with the logic from the previous plot, the stringency index is able to show drastic differences in the number of cases. It can be easily seen in Figure 2, that continents that applied that most stringent responses to the pandemic had exponentially lower number of new weekly cases and much wider standard deviations. Conversely, continents with less stringent protocols had higher numbers of cases and a much more narrow standard deviation.
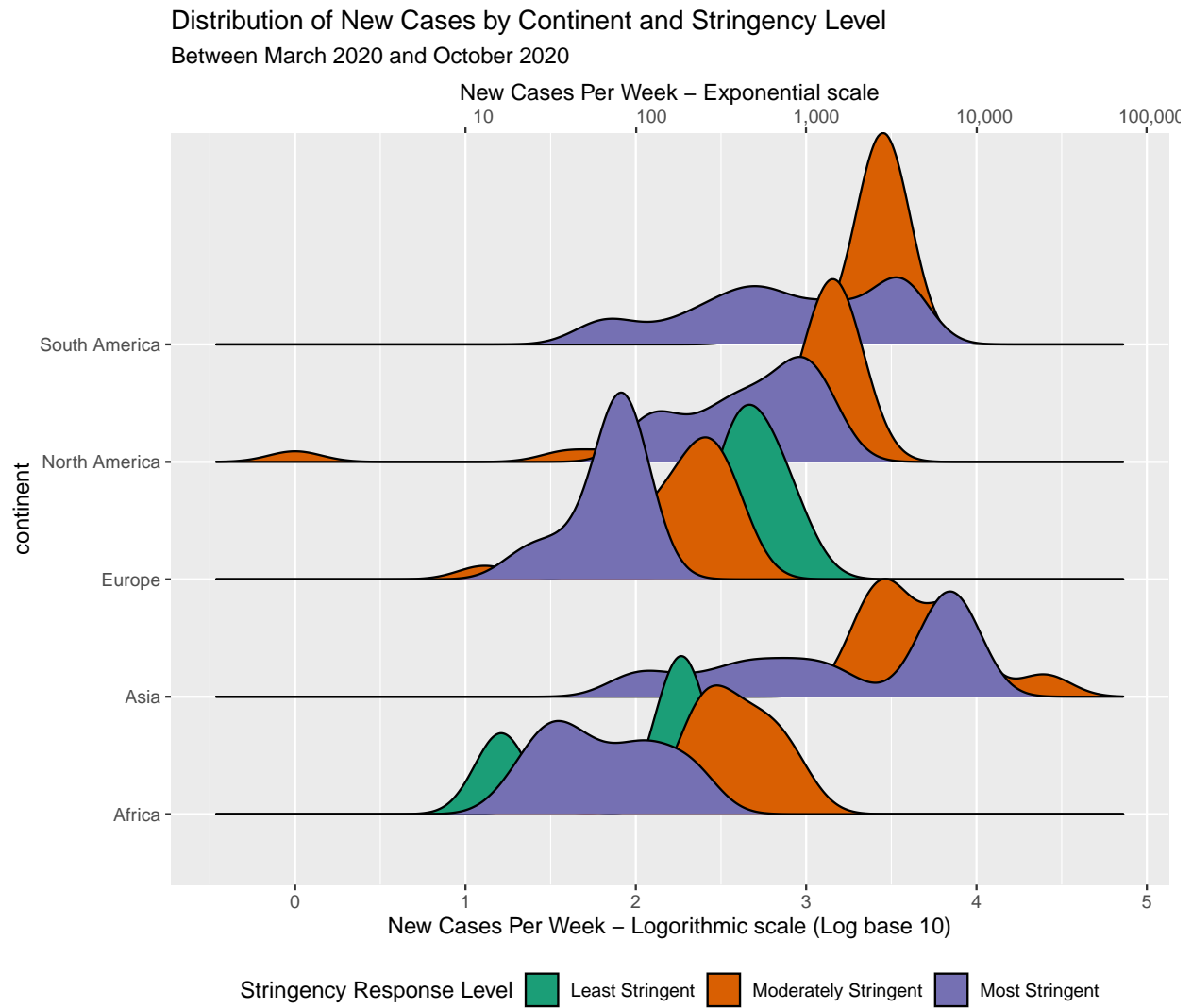
Figure 2. Distributions by Stringency Level and Continent

Visualization Three and Four

Figure 3 has a different point of view. Instead of looking at the weekly cases, it looks at the weekly death rate by total number of weekly cases. Top portion represent the ratio by continent and by stringency index. Figure 3 reflects the steady but controlled increase in the pandemic with Asia and South American continents whereas Africa, Europe and and North America experienced a sudden wave of deaths between April and June. The bottom portion reinforces the importance of continent response level and its impact on the weekly death rates. Those continents that responded poorly saw an enormous exponential impact of death cases in the same time frame. Figure 4 is shown to have wild upward curves for Least stringent and some periods of no responses while those that imposed moderate and strict protocols saw a steady decline.
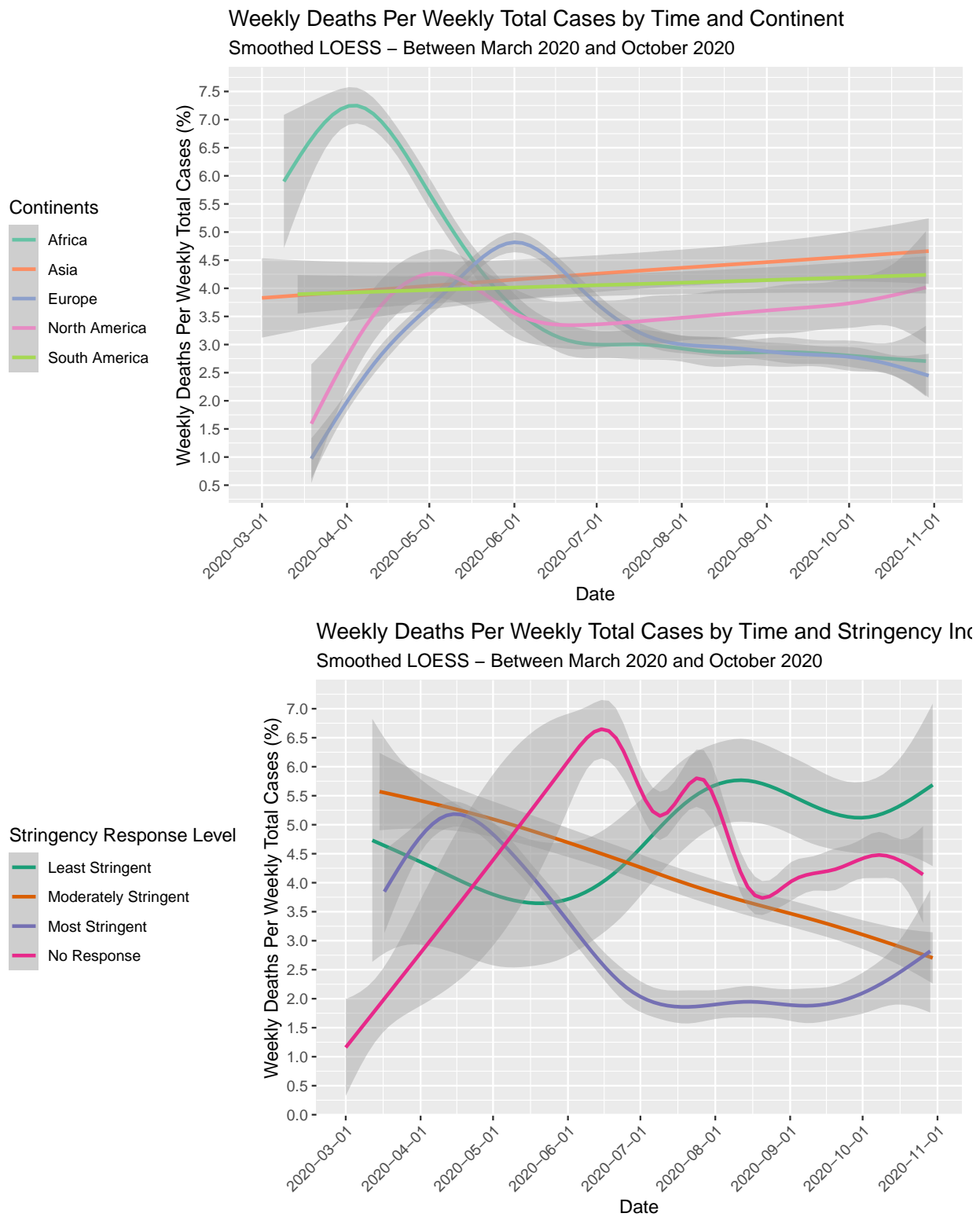
Figure 3. Smoothed LOESS by Continent and Stringency Level

Visualization Five

Figure 4 is a choropleth. depicting the total deaths caused by Covid 19 between March 2020 and October 2020. The animation is a better representation of the time lapse while the static version aggregates the entire time frame, Although the static may not the perfectly representative the animation provides fuller context of the extent the virus to hold.
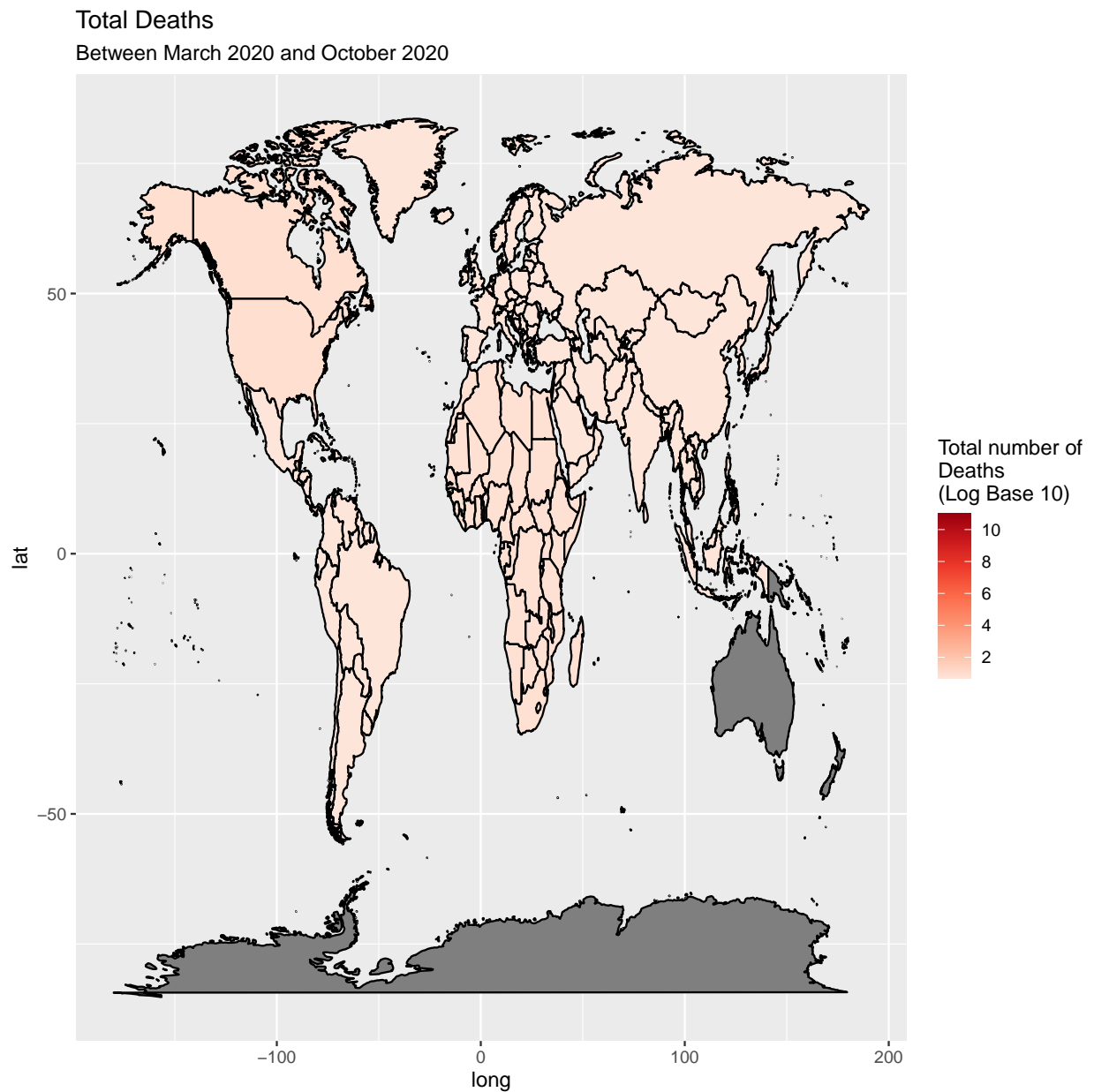

Figure 4. Choropleth. of Total Deaths caused by Covid 19