**Earthquakes and Bodies of Water Significantly Increase Spread of Zika**

Liz DiCesare, Elyssa Kiva, Emet Marwell, Emily Serleth

STAT-241, A. Foulkes

Fall 2016

**Abstract**

The spread of Zika Virus, a mosquito-born disease, has notably increased in Ecuador and Mexico since June 2016. Both countries experienced large earthquakes in April 2016. Regressions of Zika cases versus distance to the earthquake’s epicenter for each country, respectively, shows distance is not a strong predictor of Zika spread for either state. In Ecuador, declaration of State of Emergency is a stronger predictor of increased Zika cases, whereas in Mexico, distance from large bodies of water is a stronger predictor. Due to time constraints, we did not analyze social, economic, or political factors. Future research and integration of such factors to our geological data could yield much stronger predictors of Zika spread in Ecuador and Mexico.

**Introduction**

Zika Virus is a disease spread by mosquitoes1. It can cause birth defects when transmitted from mother to fetus, and may be a trigger of Guillain-Barré syndrome which can cause paralysis and death2. Transmission is possible between people through the exchange of bodily fluids such as blood, saliva, semen, breast milk, and from the placenta in mother-to-fetus transmission3. Because Zika is carried by mosquitoes, prevention of the virus hinges on elimination of mosquito breeding sites. In settled areas, breeding sites are any bodies of stagnant water such as bird baths, large puddles, and trash containers4. Additionally, Zika can be curbed through the prevention of mosquito bites in general via such methods as insect repellant and mosquito nets5.

On April 16, 2016, Ecuador experienced a massively destructive Magnitude 7.8 earthquake6. The earthquake originated just off the western coast of Ecuador, about 27 kilometers south-southeast of Muisne7, and was only 20 kilometers in depth8. UNICEF reported a twelvefold increase in reported Zika cases in the three months after the earthquake9, with 80% of the cases occurring in Manabí Province, which experienced the most damage from the quake10. This is spike in Zika cases is likely due to the destruction and subsequent accumulation of stagnant waters which served as mass breeding sites for Zika-carrying mosquitoes. Decreased protection against bodily fluids such as blood, saliva, and semen due to the chaos of the earthquake damage is another probable cause of the increase of Zika.

On April 25th, 2016, Mexico experienced two earthquakes of magnitudes 5.6 and 6.0 off the coast of Southern Mexico near the Chiapas state11. These earthquakes were only about 16 kilometers in depth12. For the same reasons as outlined with Ecuador, this earthquake could have caused a rise in Zika case in Mexico.

**Approach**

At first glance of the data, each member of our group looked at a different set of countries and decided to focus our attention on countries with weekly data sets. This list included countries such as Panama, the Dominican Republic, Ecuador, México, Puerto Rico, and Columbia. We researched data for these different countries for variables such as location and distance from other countries or bodies of water, number of cases reported versus expected, age, and pregnancy rate.

While looking at this data, we noticed a large spike in Zika in Ecuador in June of 2016. In order to find some explanation for this, we researched possible events that occurred in or prior to June of 2016 in Ecuador. According to UNICEF, an earthquake in Ecuador in April of 2016 caused a large spike in Zika. We found this interesting and decided to investigate it further.

Once we researched more about the earthquake, we were able to add in other factors to our model to test for a correlation with the number of Zika cases. The first thing we looked at was the distance of each smaller province in Ecuador from the epicenter of the earthquake. We compared that to the amount of cases in each province and to its population. After this, we then looked at each province to see if there was a state of emergency warning, and if so, at what level.

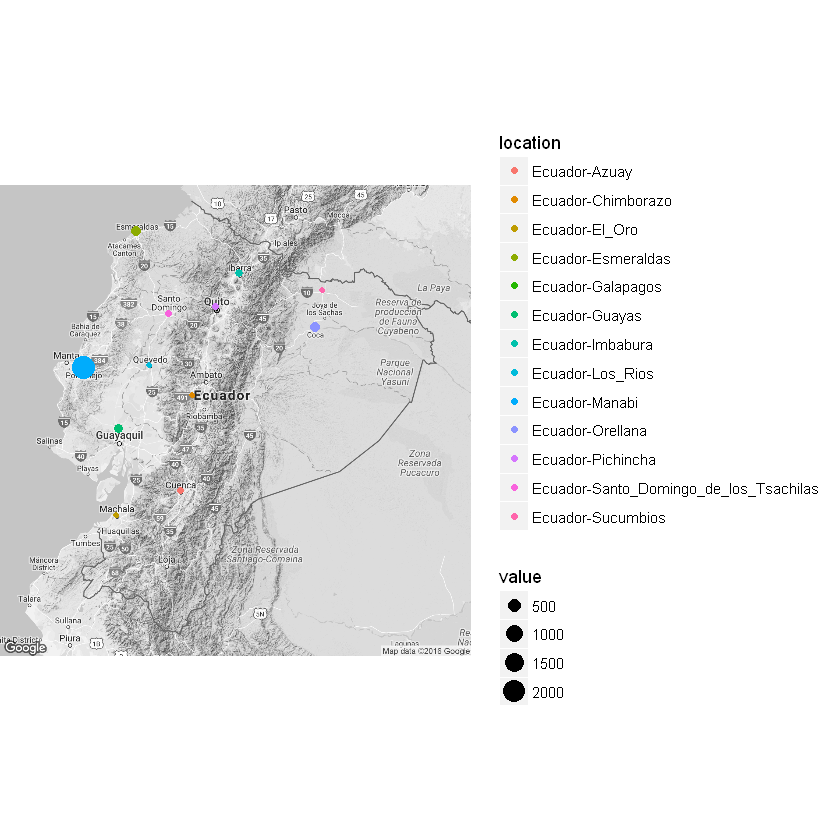
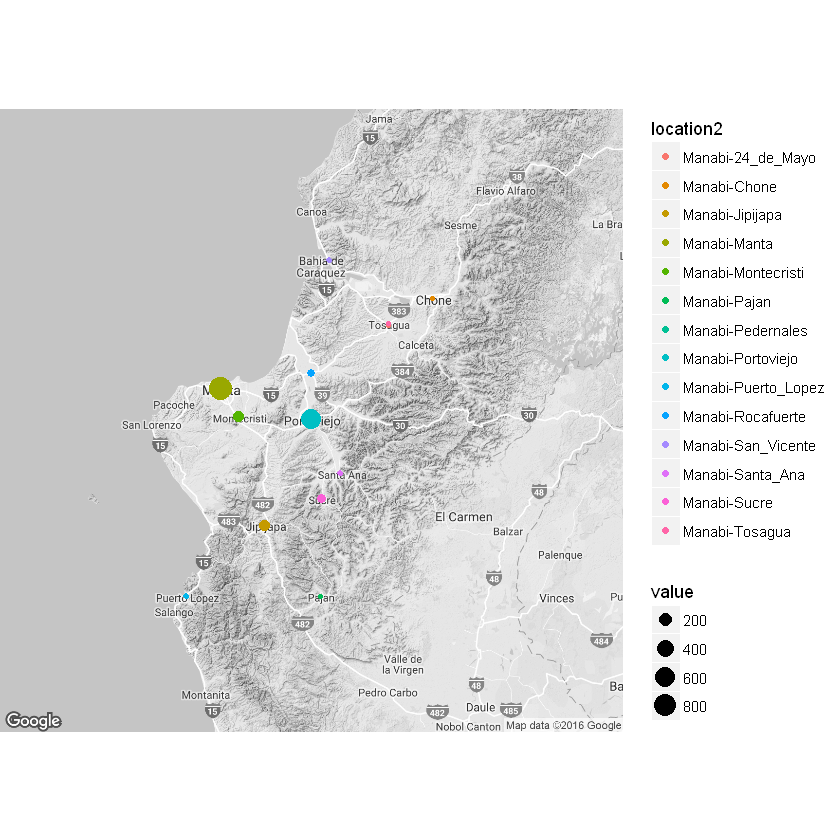
After discovering this spike in Zika cases in June in Ecuador, we looked at other countries with weekly data to see if any of them reflected a similar spike. We discovered that Mexico had a similar spike in Zika cases in June of 2016. We ran a similar search of events that happened during or prior to June of 2016 in Mexico and found another earthquake off the coast of Mexico in April of 2016. However, after testing for a correlation between the cases in Mexico and the earthquake, there did not seem to be one.

In order to find out what could have caused the sudden rise in Zika in June 2016, we controlled for population and tested for variables such as distance from water and distance from the border and found that those were somewhat more significant. In order to do this, we pinpointed the nearest large body of water to each state by finding longitude and latitude points. In the data of cases we received, some states did not have any cases of Zika. For this reason, we ran regressions with all states included and regressions with just the states that have Zika included. We have since done more research on what could have caused the rise in Zika cases, such as immigration of Haitians through the southern borders of Mexico13.

**Results**

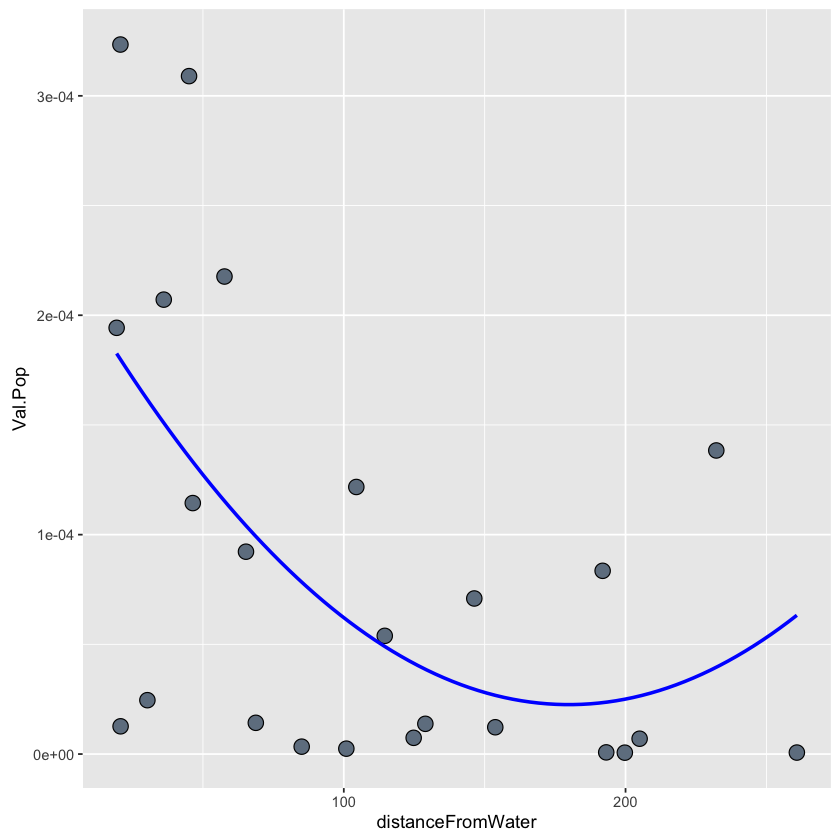
*Ecuador*

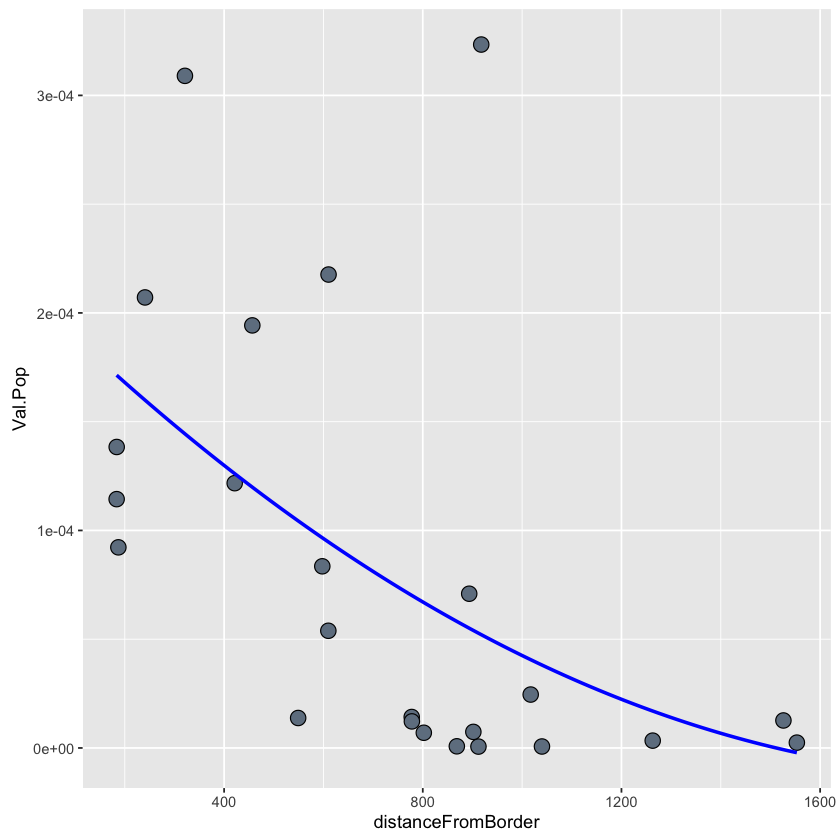
Figure 1 shows the clustering of Zika cases around the west coast of Ecuador, near the earthquake’s epicenter. Figure 2 shows Zika cases throughout Manabi Province, the province most afflicted by Zika following the earthquake.

We ran a total of 6 regressions for Ecuador, called Lm1 through Lm6. Lm1 of the number of Zika cases versus distance from the earthquake plus population had a p-value of 2.5e-11 with an adjusted R-squared value of 0.2455. Lm2 of the number of Zika cases adjusted for population versus distance from the earthquake had a p-value of 0.002283 and an adjusted R-squared value of 0.0487. Lm3, which was of cases versus distance plus population for all provinces of Ecuador, returned a p-value of 0.2716 and an adjusted R-squared value of 0.01025. Lm4 of cases adjusted for population versus distance from the quake gave a p-value of 0.5032 with an adjusted R-squared value of -0.008614. Lm5 looked at cases versus declaration of state of emergency and Lm6 did the same but adjusted cases for population. Lm5 had a p-value of 0.0008221 with an adjusted R-squared of 0.1506 and Lm6 had a p-value of 0.004319 with an adjusted R-squared value of 0.1082.

*Mexico*

We ran a total of 10 regressions for Mexico, testing for the distance from water and the distance from the border. First, we ran a regression on the number of cases by population versus the distance from the earthquake in Mexico which proved to be insignificant with a p-value of 0.1279 and an adjusted R Squared of 0.04472.

The first regression we ran was of the total number of cases versus the distance from water plus population which had a p-value of 0.006391 and an adjusted R squared value of 0.2456 with all states included. Without all the states the p-value was 0.008434 and the adjusted R Squared was 0.305. After this we ran a regression of the value of cases per person versus the distance from water. With all states the p-value was 0.002224 and the adjusted R squared was 0.2473 whereas without all states, the p-value was 0.01605 and the adjusted R squared was 0.2015. We then ran a regression for this with an exponential value for distance from water which gave a p-value of 0 .005955 and adjusted R squared of 0.2492 with all states and a p-value of 0.01619 and adjusted R squared of 0.2605 without all the states. Figure 3 shows the graph of the value of the number of cases by population versus the exponential value of distance from water, including only the states with reported Zika cases. 

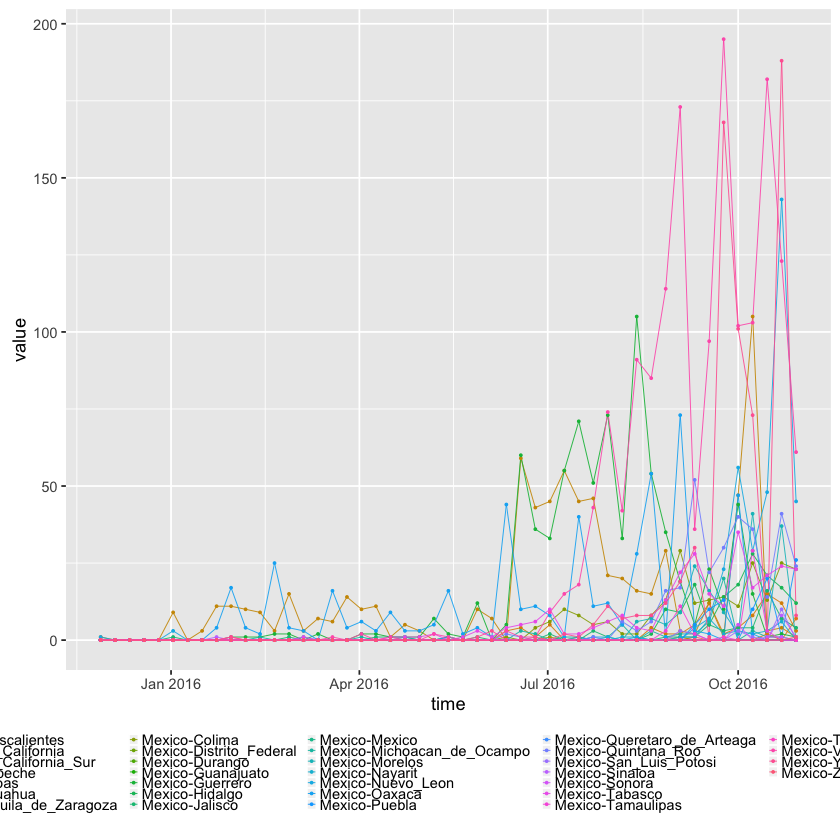
Next, we ran was the total number of cases versus the distance from the border plus population which yielded a p-value of 0.01688 with an adjusted R squared of 0.1933 with all states and a p-value of 0.004013 and adjusted R squared of 0.3525 without all the states. Next, calculating the value of cases per person versus the distance from the border, we got a p-value of 0.003174 and an adjusted R squared of 0.2306 with all states and a p-value of 0.009212 and adjusted R squared of 0.2371 without all the states. Finally, we tested for the value of cases per person by the exponential value of distance from the border which gave a p-value of 0.005955 and adjusted R squared of 0.2492 with all the states and a p-value of 0.03151 and adjusted R squared of 0.212 without all the states. Figure 4shows the graph of the value of the number of cases by population versus the exponential value of the distance from the border. 

We then ran a regression of the total number of cases versus the distance from water plus the distance from the border plus the population. With all of the states we got a p-value of 0.00207 and an adjusted R squared of 0.3399 and without all of the states we got a p-value of 0.001878 and an adjusted R squared of 0.4456. Next, we ran a regression of the number of cases per person versus the distance from water plus the distance from the border which produced a p-value of 0.0002978 and an adjusted R squared of 0.3894 with all the states and a p-value of 0.0009737 and adjusted R squared of 0.4342 without all the states. Lastly, we ran a regression of the number of cases per person versus the distance from the water plus the distance from the border plus the exponential value of the distance from the water which provided us with a p-value of 8.554e-05 and an adjusted R squared of 0.4786 with all the states included and a p-value of 0.0006256 and adjusted R square of 0.5054. 

**Discussion**

*Ecuador*

The first statistical tests of Zika in Ecuador showed distance from the earthquake to be a weak predictor of Zika’s spread, though it was very significant in Manabi Province. A seemingly better predictor was the declaration of a state of emergency. Assuming state of emergency is associated with the location’s ability to respond to extreme events, then the correlation of Zika spread to places that declared a state of emergency could be due to the region’s lack of protections against Zika spread. Alternatively, the declaration of state of emergency could be a direct response to Zika spread rather than a factor that was indirectly related.

*Mexico*

We began work on the Mexico data by looking at the number of Zika cases over time by week. From Figure 5, we noticed a large spike in Zika in June of 2016. We started by testing the number of cases in Mexico by their distance from the earthquake, and it was clear the earthquake was insignificant to the number of cases. In order to find other explanations for the rise in Zika cases, we looked at the distance from water and from the border. In order to do this, we ran regressions on all of the states and just the states with Zika cases. The regressions on just the states with Zika proved to be more significant than the regressions that included every state overall.

We then controlled for population to get a per person analysis of the number cases per state compared to their distance from the border and to water. This analysis showed us that the distance from the water was more significant to the number of Zika cases than the distance from the border. By looking at our graphs of this data, the points seemed to have a curvature to them rather than a linear pattern. For this reason, we tested each distance variable exponentially to see which fit would work best for our data. The exponential distance from the water of each state proved to be the most significant factor in number of cases. For this reason, we then found that a regression with both distance from the border and distance from water with the exponential distance from water to show the most significance to number of cases. 

From maps of where Zika is most prominent in Mexico, we can see that the most cases are in southern Mexico. Because of this, we started to research why that may be. If we had more time, we would have tested other factors related to the distance from water such as precipitation and climate factors. We also would have tested things such as the immigrant population in southern Mexico to see if immigrants were bringing in Zika. Currently, we have found some sources that say many haitians were immigrating through southern Mexico and the CDC data shows high numbers of cases of Zika starting in January of 2016 in Haiti. However, we do not have data for Haiti past January of 2016 so it is unclear what is causing the spike in Zika in Mexico in June of 2016. All we know as a group is that it may have to do with the state’s distance from the nearest large body of water.

*Challenges and Future Directions*

We faced many hurdles throughout this project, largely availability (or lack of) of consistent data. Additionally, it was difficult to isolate variables to find the true relationships amongst them. Because Zika is such a complex and current issue involving geological, biological, social, economic, and political factors spread over the globe, it is incredible challenging to find a reliable predictor(s) of Zika spread. Geological factors seem to be a good start, but more research must be conducted to integrate social, economic, and political variables. In order to conduct in-depth statistical analysis of these factors, consistent and reliable data collection is necessary.

**Notes**

1. “Zika Virus”; Moghadam et al., 2016
2. “Zika Virus and Complications: Questions and Answers”; Moghadam et al., 2016
3. Moghadam et al., 2016
4. Goddard, 2016
5. Goddard, 2016; Moghadam et al., 2016; “Zika Virus”
6. “M 7.8 - 27km SSE of Muisne, Ecuador: Overview”
7. “M 7.8 - 27km SSE of Muisne, Ecuador: Regional Information”
8. “M 7.8 - 27km SSE of Muisne, Ecuador: Overview”
9. “Twelvefold Increase in Zika Cases Since Ecuador Earthquake”
10. Ibid.
11. “Strong Earthquake - Puerto Madero, Mexico on April 25, 2016”
12. “M 6.0 - 83km WSW of Puerto Madero, Mexico: Overview”
13. “Mexican Officials Quietly Help Thousands of Haitian Illegal Immigrants Reach U.S.”

**Acknowledgements**

We would like to acknowledge the online GitHub community for sharing data of Zika from all over the world, allowing us to find the data that worked best for our project. Additionally, the United States Center for Disease Control and the World Health Organization provided us with invaluable information on Zika and its spread. Lastly, thank you to Professor Andrea Foulkes for her unwavering support of our research and help with R coding.

**Bibliography**

Earthquake Report. “Strong Earthquake - Puerto Madero, Mexico on April 25, 2016.” Last modified April 25, 2016. http://earthquake-report.com/2016/04/25/strong-earthquake-puerto-madero-mexico-on-april-25-2016/.

Goddard, Jerome. 2016. “Essential Facts About Mosquito Control and Zika Virus.” *The American Journal of Medicine,* 129: 1149-1152.

Moghadam, Saeed R.J., Samaneh Bayrami, Sepideh J. Moghadam, Raheleh Golrokhi, Fatemeh G. Pahlaviarri, Seyed A.S. Alinaghi. 2016. “Zika virus: A Review of Literature.” *Asian Pacific Journal of Tropical Biomedicine* 6 (12): 989-994.

United Nations International Children’s Emergency Fund. “Twelvefold Increase in Zika Cases Since Ecuador Earthquake.” Last modified July 18, 2016. https://www.unicef.org/media/media\_91912.html.

United States Geological Survey. “M 6.0 - 83km WSW of Puerto Madero, Mexico: Overview.” http://earthquake.usgs.gov/earthquakes/eventpage/us20005li9#executive.

United States Geological Survey. “M 7.8 - 27km SSE of Muisne, Ecuador: Overview.” <http://earthquake.usgs.gov/earthquakes/eventpage/us20005j32#origin>.

United States Geological Survey. “M 7.8 - 27km SSE of Muisne, Ecuador: Regional Information.” http://earthquake.usgs.gov/earthquakes/eventpage/us20005j32#region-info.

Washington Times. “Mexican Officials Quietly Help Thousands of Haitian Illegal Immigrants Reach U.S.” Last modified October 10, 2016. <http://www.washingtontimes.com/news/2016/oct/10/mexican-officials-helping-haitian-illegal-immigran/>.

World Health Organization. “Zika Virus.” Last modified September 6, 2016. http://www.who.int/mediacentre/factsheets/zika/en/.

World Health Organization. “Zika Virus and Complications: Questions and Answers.” Last modified November 15, 2016. http://www.who.int/features/qa/zika/en/.

**Appendix I. R Code for Ecuador Case Study Analysis**

library(dplyr)

library(ggplot2)

library(ggmap)

library(geosphere)

library(Imap)

files <- list.files("Ecuador/GACETA-ZIKA/data")

all <- rep(list(NA),length(files))

for(i in 1:length(files)){

f <- read.csv(paste("Ecuador/GACETA-ZIKA/data/",files[i],sep=""),stringsAsFactors = F)

all[[i]] <- f

}

Ecuador\_All<-as.data.frame(do.call(rbind,all))

Ecuador\_province<-filter(Ecuador\_All,location\_type=="province")

Ecuador\_manabi<-Ecuador\_All[substr(Ecuador\_All$location,1, 15)=="Ecuador-Manabi-",]

geo1 <- geocode(Ecuador\_province$location)

Ecuador\_manabi$location2 <- sub("Ecuador-","",Ecuador\_manabi$location)

geo2 <- geocode(Ecuador\_manabi$location2)

Ecuador\_manabi$long <- geo2[,1]

Ecuador\_manabi$lat <- geo2[,2]

Ecuador\_manabi$Earth\_long<-rep(-79.94,222)

Ecuador\_manabi$Earth\_lat<-rep(.371,222)

Ecuador\_manabi$distance<-gdist(lon.1 =Ecuador\_manabi$long , lat.1 = Ecuador\_manabi$lat,

lon.2 = Ecuador\_manabi$Earth\_long, lat.2 = Ecuador\_manabi$Earth\_lat, units="miles")

Ecuador\_province$long <- geo1[,1]

Ecuador\_province$lat <- geo1[,2]

Ecuador\_province$Earth\_long<-rep(-79.94,379)

Ecuador\_province$Earth\_lat<-rep(.371,379)

Ecuador\_province$distance<-gdist(lon.1 =Ecuador\_province$long ,lat.1 = Ecuador\_province$lat,

lon.2 = Ecuador\_province$Earth\_long, lat.2 = Ecuador\_province$Earth\_lat,

units="miles")

Ecuador\_manabi <- data.frame(Ecuador\_manabi,population=rep(0,222))

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Manta",]$population<-192322

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Pedernales",]$population<-303682

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Rocafuerte",]$population<-29321

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Montecristi",]$population<-43400

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Sucre",]$population<-52158

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Jipijapa",]$population<-65796

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Santa\_Ana",]$population<-45287

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-San\_Vicente",]$population<-19116

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-24\_de\_Mayo",]$population<-28294

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Puerto\_Lopez",]$population<-16626

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Pajan",]$population<-35952

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Tosagua",]$population<-33922

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Chone",]$population<-117634

Ecuador\_manabi[Ecuador\_manabi$location2=="Manabi-Portoviejo",]$population<-303682

Ecuador\_province$location2 <- sub("Ecuador-","",Ecuador\_province$location)

Ecuador\_province <- data.frame(Ecuador\_province,population=rep(0,379))

Ecuador\_province[Ecuador\_province$location2=="Sucumbios",]$population<-176472

Ecuador\_province[Ecuador\_province$location2=="Santo\_Domingo\_de\_los\_Tsachilas",]$population<-368013

Ecuador\_province[Ecuador\_province$location2=="Santa\_Elena",]$population<-308693

Ecuador\_province[Ecuador\_province$location2=="Pichincha",]$population<-2576287

Ecuador\_province[Ecuador\_province$location2=="Orellana",]$population<-136396

Ecuador\_province[Ecuador\_province$location2=="Manabi",]$population<-1369780

Ecuador\_province[Ecuador\_province$location2=="Los\_Rios",]$population<-778115

Ecuador\_province[Ecuador\_province$location2=="Guayas",]$population<-3645483

Ecuador\_province[Ecuador\_province$location2=="Galapagos",]$population<-25124

Ecuador\_province[Ecuador\_province$location2=="Esmeraldas",]$population<-534092

Ecuador\_province[Ecuador\_province$location2=="El\_Oro",]$population<-600659

Ecuador\_province[Ecuador\_province$location2=="Cotopaxi",]$population<-409205

Ecuador\_province[Ecuador\_province$location2=="Bolivar",]$population<-183641

Ecuador\_province[Ecuador\_province$location2=="Canar",]$population<-225184

Ecuador\_province[Ecuador\_province$location2=="Loja",]$population<-448966

Ecuador\_province[Ecuador\_province$location2=="Napo",]$population<-103697

Ecuador\_province[Ecuador\_province$location2=="Carchi",]$population<-164524

Ecuador\_province[Ecuador\_province$location2=="Imbabura",]$population<-398244

Ecuador\_province[Ecuador\_province$location2=="Chimborazo",]$population<-458581

Ecuador\_province[Ecuador\_province$location2=="Tungurahua",]$population<-504583

Ecuador\_province[Ecuador\_province$location2=="Azuay",]$population<-712127

Ecuador\_province <- data.frame(Ecuador\_province,State\_of\_emergency=rep(0,379))

Ecuador\_province[Ecuador\_province$location2=="Sucumbios",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Santo\_Domingo\_de\_los\_Tsachilas",]$State\_of\_emergency<-1

Ecuador\_province[Ecuador\_province$location2=="Santa\_Elena",]$State\_of\_emergency<-1

Ecuador\_province[Ecuador\_province$location2=="Pichincha",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Orellana",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Manabi",]$State\_of\_emergency<-1

Ecuador\_province[Ecuador\_province$location2=="Los\_Rios",]$State\_of\_emergency<-1

Ecuador\_province[Ecuador\_province$location2=="Guayas",]$State\_of\_emergency<-1

Ecuador\_province[Ecuador\_province$location2=="Galapagos",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Esmeraldas",]$State\_of\_emergency<-1

Ecuador\_province[Ecuador\_province$location2=="El\_Oro",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Cotopaxi",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Bolivar",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Canar",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Loja",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Napo",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Carchi",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Imbabura",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Chimborazo",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Tungurahua",]$State\_of\_emergency<-0

Ecuador\_province[Ecuador\_province$location2=="Azuay",]$State\_of\_emergency<-0

head(Ecuador\_province)

head(Ecuador\_manabi)

*#analysis of manabi province*

lm1<-lm(value~distance+population,data=subset(Ecuador\_manabi,data\_field\_code %in% c("EC0003")))

summary(lm1)

lm2<-lm(value/population~distance,data=subset(Ecuador\_manabi,data\_field\_code %in% c("EC0003")))

summary(lm2)

*#distance is positive and not negative probably due to missing county data (look at plot)*

*#analysis of all provinces*

lm3<-lm(value~distance+population,data=subset(Ecuador\_province,data\_field\_code %in% c("EC0003")))

summary(lm3)

lm4<-lm(value/population~distance,data=subset(Ecuador\_province,data\_field\_code %in% c("EC0003")))

summary(lm4)

lm5<-lm(value~State\_of\_emergency,data=subset(Ecuador\_province,data\_field\_code %in% c("EC0003")))

summary(lm5)

lm6<-lm(value/population~State\_of\_emergency,data=subset(Ecuador\_province,data\_field\_code %in% c("EC0003")))

summary(lm6)

Ecuador\_manabi$time <- as.Date(Ecuador\_manabi$"report\_date")

ggplot(data=subset(Ecuador\_manabi,data\_field\_code %in% c("EC0003"))

,aes(x=time,y=value)) +

geom\_line(aes(colour=location),size=.25) +

geom\_point(aes(colour=location),size=.4) + theme(legend.key.size = unit(.2, "cm"),legend.position = 'bottom')

Ecuador\_province$time <- as.Date(Ecuador\_province$"report\_date")

ggplot(data=subset(Ecuador\_province,data\_field\_code %in% c("EC0003"))

,aes(x=time,y=value)) +

geom\_line(aes(colour=location),size=.25) +

geom\_point(aes(colour=location),size=.4) +

theme(legend.key.size = unit(.2, "cm"),legend.position = 'bottom')

*#Manabi Plots*

ggplot(data=subset(Ecuador\_manabi,data\_field\_code %in% c("EC0003")),aes(x=(distance),y=(value),color=report\_date))+geom\_point()

ggplot(data=subset(Ecuador\_manabi,data\_field\_code %in% c("EC0003")),aes(x=population,y=value,color=location))+geom\_point()

*#Provicnes plots*

ggplot(data = subset(Ecuador\_province,data\_field\_code %in% c("EC0003")),aes(x=as.factor(State\_of\_emergency),y=value))+geom\_boxplot()+

geom\_point(aes(colour=location),size=1)

EcuadorMap <- qmap("Ecuador",zoom=7, color = "bw")

EcuadorMap+geom\_point(aes(x = long, y = lat, colour = location, size=value),

data = subset(Ecuador\_province,data\_field\_code %in% c("EC0003")))

EcuadorMap <- qmap("Ecuador\_manabi",zoom=9, color = "bw")

EcuadorMap+geom\_point(aes(x = long, y = lat, colour = location2, size=value),

data =subset(Ecuador\_manabi,data\_field\_code %in% c("EC0003")))

**Appendix II. R Code for Mexico Case Study Analysis - All States Included**

library(dplyr)

library(ggplot2)

library(ggmap)

files <- list.files("Mexico/DGE\_Zika/data")

files

all <- rep(list(NA),length(files))

for(i in 1:length(files)){

f <- read.csv(paste("Mexico/DGE\_Zika/data/",files[i],sep=""),stringsAsFactors = F)

all[[i]] <- f

}

Mexico <- as.data.frame(do.call(rbind,all))

Mexico <- arrange(Mexico,report\_date)

Mexico <- subset(Mexico, data\_field == "weekly\_zika\_confirmed")

*#Mexico*

Mexico$time <- as.Date(Mexico$"report\_date")

ggplot(data=Mexico,aes(x=time,y=value)) +

geom\_line(aes(colour=location),size=.25) +

geom\_point(aes(colour=location),size=.4) +

theme(legend.key.size = unit(.2, "cm"),legend.position = 'bottom')

num <- aggregate(value~location,Mexico,length)

names(num)[2] <- 'num'

totalValue <- aggregate(value~location,Mexico,sum)

names(totalValue)[2] <- 'totalValue'

Mexico\_Totals<-merge(num,totalValue)

*#Mexico\_Totals*

Mexico\_Totals <- arrange(Mexico\_Totals,totalValue)

*#Mexico\_Totals*

geo <- geocode(Mexico\_Totals$location)

Mexico\_Totals$long <- geo[,1]

Mexico\_Totals$lat <- geo[,2]

*#dim(Mexico\_Totals)*

Mexico\_Totals$Earth\_long<-rep(-105.173,32)

Mexico\_Totals$Earth\_lat<-rep(18.364,32)

*#install.packages("geosphere")*

library(geosphere)

*#install.packages("Imap")*

library(Imap)

Mexico\_Totals$distance<-gdist(lon.1 =Mexico\_Totals$long , lat.1 = Mexico\_Totals$lat, lon.2 = Mexico\_Totals$Earth\_long, lat.2 = Mexico\_Totals$Earth\_lat, units="miles")

Mexico\_Totals$nearestWaterLat <- rep(0,32)

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chiapas",]$nearestWaterLat<-16.969457

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Oaxaca",]$nearestWaterLat<-16.000276

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guerrero",]$nearestWaterLat<-17.032495

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tabasco",]$nearestWaterLat<-18.497192

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Morelos",]$nearestWaterLat<-19.432277

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Puebla",]$nearestWaterLat<-19.432277

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Quintana\_Roo",]$nearestWaterLat<-18.827666

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Colima",]$nearestWaterLat<-19.063335

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Distrito\_Federal",]$nearestWaterLat<-20.093982

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Veracruz\_de\_Ignacio\_de\_la\_Llave",]$nearestWaterLat<-19.348035

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tlaxcala",]$nearestWaterLat<-19.534514

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Mexico",]$nearestWaterLat<-20.516429

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Michoacan\_de\_Ocampo",]$nearestWaterLat<-18.813937

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Yucatan",]$nearestWaterLat<-21.350142

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guanajuato",]$nearestWaterLat<-21.500991

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Aguascalientes",]$nearestWaterLat<-21.474716

Mexico\_Totals[Mexico\_Totals$location == "Mexico-San\_Luis\_Potosi",]$nearestWaterLat<-22.357061

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Zacatecas",]$nearestWaterLat<-22.351267

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Durango",]$nearestWaterLat<-23.66583

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tamaulipas",]$nearestWaterLat<-24.213778

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sinaloa",]$nearestWaterLat<-25.464277

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nuevo\_Leon",]$nearestWaterLat<-25.271399

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California\_Sur",]$nearestWaterLat<-26.025859

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Coahuila\_de\_Zaragoza",]$nearestWaterLat<-26.870018

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chihuahua",]$nearestWaterLat<-27.750316

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sonora",]$nearestWaterLat<-28.731388

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Campeche",]$nearestWaterLat<-19.840503

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Hidalgo",]$nearestWaterLat<-20.545044

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Queretaro\_de\_Arteaga",]$nearestWaterLat<-20.832823

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Jalisco",]$nearestWaterLat<-20.635194

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nayarit",]$nearestWaterLat<-21.49005

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California",]$nearestWaterLat<-30.98758

*#Mexico\_Totals*

Mexico\_Totals$nearestWaterLong <- rep(0,32)

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chiapas",]$nearestWaterLong<--89.629211

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Oaxaca",]$nearestWaterLong<--97.846985

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guerrero",]$nearestWaterLong<--100.307922

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tabasco",]$nearestWaterLong<--92.782288

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Morelos",]$nearestWaterLong<--96.27594

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Puebla",]$nearestWaterLong<--96.27594

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Quintana\_Roo",]$nearestWaterLong<--88.071899

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Colima",]$nearestWaterLong<--104.317932

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Distrito\_Federal",]$nearestWaterLong<--96.715393

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Veracruz\_de\_Ignacio\_de\_la\_Llave",]$nearestWaterLong<--96.296539

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tlaxcala",]$nearestWaterLong<--96.362457

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Mexico",]$nearestWaterLong<--97.001038

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Michoacan\_de\_Ocampo",]$nearestWaterLong<--103.922424

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Yucatan",]$nearestWaterLong<--89.241943

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guanajuato",]$nearestWaterLong<--97.506409

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Aguascalientes",]$nearestWaterLong<--105.262756

Mexico\_Totals[Mexico\_Totals$location == "Mexico-San\_Luis\_Potosi",]$nearestWaterLong<--97.792053

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Zacatecas",]$nearestWaterLong<--105.680237

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Durango",]$nearestWaterLong<--106.844788

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tamaulipas",]$nearestWaterLong<--97.748108

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sinaloa",]$nearestWaterLong<--108.800354

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nuevo\_Leon",]$nearestWaterLong<--97.68219

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California\_Sur",]$nearestWaterLong<--111.332703

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Coahuila\_de\_Zaragoza",]$nearestWaterLong<--97.484436

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chihuahua",]$nearestWaterLong<--110.805359

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sonora",]$nearestWaterLong<--111.868286

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Campeche",]$nearestWaterLong<--90.548286

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Hidalgo",]$nearestWaterLong<--97.064209

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Queretaro\_de\_Arteaga",]$nearestWaterLong<--97.261963

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Jalisco",]$nearestWaterLong<--105.276489

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nayarit",]$nearestWaterLong<--105.224304

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California",]$nearestWaterLong<--114.813995

*#Mexico\_Totals*

Mexico\_Totals$distanceFromWater<-gdist(lon.1 =Mexico\_Totals$long , lat.1 = Mexico\_Totals$lat, lon.2 = Mexico\_Totals$nearestWaterLong, lat.2 = Mexico\_Totals$nearestWaterLat, units="miles")

Mexico\_Totals$border\_long<-rep(-90.41748,32)

Mexico\_Totals$border\_lat<-rep(16.214675,32)

Mexico\_Totals$distanceFromBorder<-gdist(lon.1 =Mexico\_Totals$long , lat.1 = Mexico\_Totals$lat, lon.2 = Mexico\_Totals$border\_long, lat.2 = Mexico\_Totals$border\_lat, units="miles")

Mexico\_Totals$population <- rep(0,32)

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chiapas",]$population<-5217908

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Oaxaca",]$population<-3967889

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guerrero",]$population<-3533251

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tabasco",]$population<-2395272

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Morelos",]$population<-1903811

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Puebla",]$population<-6168883

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Quintana\_Roo",]$population<-1501562

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Colima",]$population<-711235

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Distrito\_Federal",]$population<-8918653

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Veracruz\_de\_Ignacio\_de\_la\_Llave",]$population<-8112505

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tlaxcala",]$population<-1272847

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Mexico",]$population<-16187608

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Michoacan\_de\_Ocampo",]$population<-4584471

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Yucatan",]$population<-2097175

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guanajuato",]$population<-5853677

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Aguascalientes",]$population<-1312544

Mexico\_Totals[Mexico\_Totals$location == "Mexico-San\_Luis\_Potosi",]$population<-2717820

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Zacatecas",]$population<-1579209

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Durango",]$population<-1754754

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tamaulipas",]$population<-3441698

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sinaloa",]$population<-2966321

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nuevo\_Leon",]$population<-5119504

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California\_Sur",]$population<-712029

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Coahuila\_de\_Zaragoza",]$population<-2954915

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chihuahua",]$population<-3556574

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sonora",]$population<-2850330

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Campeche",]$population<-899931

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Hidalgo",]$population<-2858359

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Queretaro\_de\_Arteaga",]$population<-2038372

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Jalisco",]$population<-7844830

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nayarit",]$population<-1181050

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California",]$population<-3315766

*#Mexico\_Totals*

Mexico\_Totals$Val.Pop <- rep(0,32)

Mexico\_Totals$Val.Pop <- Mexico\_Totals$totalValue/Mexico\_Totals$population

*#Mexico\_Totals <- subset(Mexico\_Totals, totalValue != 0)*

Mexico\_Totals

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=totalValue), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=Val.Pop), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, colour=distanceFromWater), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, colour=distanceFromBorder), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=Val.Pop, colour = distanceFromWater), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=Val.Pop, colour=distanceFromBorder), data = Mexico\_Totals)

total.water.population<-lm(totalValue~distanceFromWater+population,data=Mexico\_Totals)

summary(total.water.population)

total.border.population<-lm(totalValue~distanceFromBorder+population,data=Mexico\_Totals)

summary(total.border.population)

total.water.border.population<-lm(totalValue~distanceFromWater+distanceFromBorder+population,data=Mexico\_Totals)

summary(total.water.border.population)

perPerson.water<-lm(Val.Pop~distanceFromWater,data=Mexico\_Totals)

summary(perPerson.water)

perPerson.border<-lm(Val.Pop~distanceFromBorder,data=Mexico\_Totals)

summary(perPerson.border)

perPerson.water.border<-lm(Val.Pop~distanceFromWater+distanceFromBorder,data=Mexico\_Totals)

summary(perPerson.water.border)

ggplot(data=Mexico\_Totals,

aes(x=distanceFromBorder,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

geom\_smooth(method=lm, *# Add linear regression line*

se=FALSE, *# Don't add shaded confidence region*

colour="darkblue")

ggplot(data=Mexico\_Totals,

aes(x=distanceFromWater,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

geom\_smooth(method=lm, *# Add linear regression line*

se=FALSE, *# Don't add shaded confidence region*

colour="darkblue")

ggplot(data=Mexico\_Totals,

aes(x=distanceFromBorder,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

stat\_smooth(method = "lm", formula = y ~ x + I(x^2),

size = 1, se = FALSE, colour = "blue")

exponential.perPerson.border<-lm(Val.Pop~distanceFromBorder + I(distanceFromBorder^2),data=Mexico\_Totals)

summary(exponential.perPerson.border)

ggplot(data=Mexico\_Totals,

aes(x=distanceFromWater,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

stat\_smooth(method = "lm", formula = y ~ x + I(x^2),

size = 1, se = FALSE, colour = "blue")

exponential.perPerson.water<-lm(Val.Pop~distanceFromWater + I(distanceFromWater^2),data=Mexico\_Totals)

summary(exponential.perPerson.water)

exponential.perPerson.water.border<-lm(Val.Pop~distanceFromBorder +

distanceFromWater +

I(distanceFromWater^2),

data=Mexico\_Totals)

summary(exponential.perPerson.water.border)

**Appendix III. R Code for Mexico Case Study Analysis - Excluding States with No Zika Cases**

library(dplyr)

library(ggplot2)

library(ggmap)

files <- list.files("Mexico/DGE\_Zika/data")

files

all <- rep(list(NA),length(files))

for(i in 1:length(files)){

f <- read.csv(paste("Mexico/DGE\_Zika/data/",files[i],sep=""),stringsAsFactors = F)

all[[i]] <- f

}

Mexico <- as.data.frame(do.call(rbind,all))

Mexico <- arrange(Mexico,report\_date)

Mexico <- subset(Mexico, data\_field == "weekly\_zika\_confirmed")

*#Mexico*

Mexico$time <- as.Date(Mexico$"report\_date")

ggplot(data=Mexico,aes(x=time,y=value)) +

geom\_line(aes(colour=location),size=.25) +

geom\_point(aes(colour=location),size=.4) +

theme(legend.key.size = unit(.2, "cm"),legend.position = 'bottom')

num <- aggregate(value~location,Mexico,length)

names(num)[2] <- 'num'

totalValue <- aggregate(value~location,Mexico,sum)

names(totalValue)[2] <- 'totalValue'

Mexico\_Totals<-merge(num,totalValue)

*#Mexico\_Totals*

Mexico\_Totals <- arrange(Mexico\_Totals,totalValue)

*#Mexico\_Totals*

geo <- geocode(Mexico\_Totals$location)

Mexico\_Totals$long <- geo[,1]

Mexico\_Totals$lat <- geo[,2]

*#dim(Mexico\_Totals)*

Mexico\_Totals$Earth\_long<-rep(-105.173,32)

Mexico\_Totals$Earth\_lat<-rep(18.364,32)

*#install.packages("geosphere")*

library(geosphere)

*#install.packages("Imap")*

library(Imap)

Mexico\_Totals$distance<-gdist(lon.1 =Mexico\_Totals$long , lat.1 = Mexico\_Totals$lat, lon.2 = Mexico\_Totals$Earth\_long, lat.2 = Mexico\_Totals$Earth\_lat, units="miles")

Mexico\_Totals$nearestWaterLat <- rep(0,32)

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chiapas",]$nearestWaterLat<-16.969457

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Oaxaca",]$nearestWaterLat<-16.000276

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guerrero",]$nearestWaterLat<-17.032495

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tabasco",]$nearestWaterLat<-18.497192

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Morelos",]$nearestWaterLat<-19.432277

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Puebla",]$nearestWaterLat<-19.432277

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Quintana\_Roo",]$nearestWaterLat<-18.827666

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Colima",]$nearestWaterLat<-19.063335

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Distrito\_Federal",]$nearestWaterLat<-20.093982

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Veracruz\_de\_Ignacio\_de\_la\_Llave",]$nearestWaterLat<-19.348035

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tlaxcala",]$nearestWaterLat<-19.534514

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Mexico",]$nearestWaterLat<-20.516429

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Michoacan\_de\_Ocampo",]$nearestWaterLat<-18.813937

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Yucatan",]$nearestWaterLat<-21.350142

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guanajuato",]$nearestWaterLat<-21.500991

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Aguascalientes",]$nearestWaterLat<-21.474716

Mexico\_Totals[Mexico\_Totals$location == "Mexico-San\_Luis\_Potosi",]$nearestWaterLat<-22.357061

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Zacatecas",]$nearestWaterLat<-22.351267

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Durango",]$nearestWaterLat<-23.66583

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tamaulipas",]$nearestWaterLat<-24.213778

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sinaloa",]$nearestWaterLat<-25.464277

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nuevo\_Leon",]$nearestWaterLat<-25.271399

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California\_Sur",]$nearestWaterLat<-26.025859

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Coahuila\_de\_Zaragoza",]$nearestWaterLat<-26.870018

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chihuahua",]$nearestWaterLat<-27.750316

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sonora",]$nearestWaterLat<-28.731388

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Campeche",]$nearestWaterLat<-19.840503

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Hidalgo",]$nearestWaterLat<-20.545044

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Queretaro\_de\_Arteaga",]$nearestWaterLat<-20.832823

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Jalisco",]$nearestWaterLat<-20.635194

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nayarit",]$nearestWaterLat<-21.49005

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California",]$nearestWaterLat<-30.98758

*#Mexico\_Totals*

Mexico\_Totals$nearestWaterLong <- rep(0,32)

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chiapas",]$nearestWaterLong<--89.629211

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Oaxaca",]$nearestWaterLong<--97.846985

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guerrero",]$nearestWaterLong<--100.307922

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tabasco",]$nearestWaterLong<--92.782288

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Morelos",]$nearestWaterLong<--96.27594

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Puebla",]$nearestWaterLong<--96.27594

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Quintana\_Roo",]$nearestWaterLong<--88.071899

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Colima",]$nearestWaterLong<--104.317932

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Distrito\_Federal",]$nearestWaterLong<--96.715393

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Veracruz\_de\_Ignacio\_de\_la\_Llave",]$nearestWaterLong<--96.296539

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tlaxcala",]$nearestWaterLong<--96.362457

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Mexico",]$nearestWaterLong<--97.001038

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Michoacan\_de\_Ocampo",]$nearestWaterLong<--103.922424

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Yucatan",]$nearestWaterLong<--89.241943

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guanajuato",]$nearestWaterLong<--97.506409

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Aguascalientes",]$nearestWaterLong<--105.262756

Mexico\_Totals[Mexico\_Totals$location == "Mexico-San\_Luis\_Potosi",]$nearestWaterLong<--97.792053

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Zacatecas",]$nearestWaterLong<--105.680237

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Durango",]$nearestWaterLong<--106.844788

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tamaulipas",]$nearestWaterLong<--97.748108

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sinaloa",]$nearestWaterLong<--108.800354

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nuevo\_Leon",]$nearestWaterLong<--97.68219

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California\_Sur",]$nearestWaterLong<--111.332703

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Coahuila\_de\_Zaragoza",]$nearestWaterLong<--97.484436

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chihuahua",]$nearestWaterLong<--110.805359

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sonora",]$nearestWaterLong<--111.868286

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Campeche",]$nearestWaterLong<--90.548286

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Hidalgo",]$nearestWaterLong<--97.064209

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Queretaro\_de\_Arteaga",]$nearestWaterLong<--97.261963

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Jalisco",]$nearestWaterLong<--105.276489

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nayarit",]$nearestWaterLong<--105.224304

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California",]$nearestWaterLong<--114.813995

*#Mexico\_Totals*

Mexico\_Totals$distanceFromWater<-gdist(lon.1 =Mexico\_Totals$long , lat.1 = Mexico\_Totals$lat, lon.2 = Mexico\_Totals$nearestWaterLong, lat.2 = Mexico\_Totals$nearestWaterLat, units="miles")

Mexico\_Totals$border\_long<-rep(-90.41748,32)

Mexico\_Totals$border\_lat<-rep(16.214675,32)

Mexico\_Totals$distanceFromBorder<-gdist(lon.1 =Mexico\_Totals$long , lat.1 = Mexico\_Totals$lat, lon.2 = Mexico\_Totals$border\_long, lat.2 = Mexico\_Totals$border\_lat, units="miles")

Mexico\_Totals$population <- rep(0,32)

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chiapas",]$population<-5217908

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Oaxaca",]$population<-3967889

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guerrero",]$population<-3533251

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tabasco",]$population<-2395272

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Morelos",]$population<-1903811

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Puebla",]$population<-6168883

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Quintana\_Roo",]$population<-1501562

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Colima",]$population<-711235

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Distrito\_Federal",]$population<-8918653

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Veracruz\_de\_Ignacio\_de\_la\_Llave",]$population<-8112505

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tlaxcala",]$population<-1272847

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Mexico",]$population<-16187608

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Michoacan\_de\_Ocampo",]$population<-4584471

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Yucatan",]$population<-2097175

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Guanajuato",]$population<-5853677

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Aguascalientes",]$population<-1312544

Mexico\_Totals[Mexico\_Totals$location == "Mexico-San\_Luis\_Potosi",]$population<-2717820

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Zacatecas",]$population<-1579209

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Durango",]$population<-1754754

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Tamaulipas",]$population<-3441698

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sinaloa",]$population<-2966321

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nuevo\_Leon",]$population<-5119504

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California\_Sur",]$population<-712029

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Coahuila\_de\_Zaragoza",]$population<-2954915

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Chihuahua",]$population<-3556574

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Sonora",]$population<-2850330

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Campeche",]$population<-899931

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Hidalgo",]$population<-2858359

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Queretaro\_de\_Arteaga",]$population<-2038372

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Jalisco",]$population<-7844830

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Nayarit",]$population<-1181050

Mexico\_Totals[Mexico\_Totals$location == "Mexico-Baja\_California",]$population<-3315766

*#Mexico\_Totals*

Mexico\_Totals$Val.Pop <- rep(0,32)

Mexico\_Totals$Val.Pop <- Mexico\_Totals$totalValue/Mexico\_Totals$population

Mexico\_Totals <- subset(Mexico\_Totals, totalValue != 0)

Mexico\_Totals

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=totalValue), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=Val.Pop), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, colour=distanceFromWater), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, colour=distanceFromBorder), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=Val.Pop, colour = distanceFromWater), data = Mexico\_Totals)

MexicoMap <- qmap("Mexico",zoom=4, color = "bw")

MexicoMap + geom\_point(aes(x = Mexico\_Totals$long, y = Mexico\_Totals$lat, size=Val.Pop, colour=distanceFromBorder), data = Mexico\_Totals)

total.water.population<-lm(totalValue~distanceFromWater+population,data=Mexico\_Totals)

summary(total.water.population)

total.border.population<-lm(totalValue~distanceFromBorder+population,data=Mexico\_Totals)

summary(total.border.population)

total.water.border.population<-lm(totalValue~distanceFromWater+distanceFromBorder+population,data=Mexico\_Totals)

summary(total.water.border.population)

perPerson.water<-lm(Val.Pop~distanceFromWater,data=Mexico\_Totals)

summary(perPerson.water)

perPerson.border<-lm(Val.Pop~distanceFromBorder,data=Mexico\_Totals)

summary(perPerson.border)

perPerson.water.border<-lm(Val.Pop~distanceFromWater+distanceFromBorder,data=Mexico\_Totals)

summary(perPerson.water.border)

ggplot(data=Mexico\_Totals,

aes(x=distanceFromBorder,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

geom\_smooth(method=lm, *# Add linear regression line*

se=FALSE, *# Don't add shaded confidence region*

colour="darkblue")

ggplot(data=Mexico\_Totals,

aes(x=distanceFromWater,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

geom\_smooth(method=lm, *# Add linear regression line*

se=FALSE, *# Don't add shaded confidence region*

colour="darkblue")

ggplot(data=Mexico\_Totals,

aes(x=distanceFromBorder,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

stat\_smooth(method = "lm", formula = y ~ x + I(x^2),

size = 1, se = FALSE, colour = "blue")

exponential.perPerson.border<-lm(Val.Pop~distanceFromBorder + I(distanceFromBorder^2),data=Mexico\_Totals)

summary(exponential.perPerson.border)

ggplot(data=Mexico\_Totals,

aes(x=distanceFromWater,y=Val.Pop)) +

geom\_point(alpha=1, shape=21, fill="slategrey", colour="black", size=4) +

stat\_smooth(method = "lm", formula = y ~ x + I(x^2),

size = 1, se = FALSE, colour = "blue")

exponential.perPerson.water<-lm(Val.Pop~distanceFromWater + I(distanceFromWater^2),data=Mexico\_Totals)

summary(exponential.perPerson.water)

exponential.perPerson.water.border<-lm(Val.Pop~distanceFromBorder +

distanceFromWater +

I(distanceFromWater^2),

data=Mexico\_Totals)

summary(exponential.perPerson.water.border)