

A_Reproducible_Work_in_R.r

Monalisa Mishra

Sat Mar 03 19:22:46 2018

```
#####
# A reproducible work in R environment on 2017-2018 Influenza Season Week 4 ending January 27, 2018
# (presented by Influenza Deivision of CDC) based on data provided by the site and the code written by me
#
# https://www.cdc.gov/flu/weekly/weeklyarchives2017-2018/Week04.htm
#
# Synopsis:
# During week 4 (January 21-27, 2018), influenza activity increased in the United States.
# Viral Surveillance: The most frequently identified influenza virus subtype reported by public health
#                    laboratories during week 4 was influenza A(H3). The percentage of respiratory specimens
#                    testing positive for influenza in clinical laboratories remained elevated
# Pneumonia and Influenza Mortality: The proportion of deaths attributed to pneumonia and influenza (P&I) was
#                    above the system-specific epidemic threshold in the National Center for
#                    Health Statistics (NCHS) Mortality Surveillance System
# Influenza-associated Pediatric Deaths: A total of 53 influenza-associated pediatric deaths have been
#                    reported for the 2017-2018 season
# Influenza-associated Hospitalizations: A cumulative rate of 51.4 laboratory-confirmed influenza-associated
#                    hospitalizations per 100,000 population was reported
# Outpatient Illness Surveillance: The proportion of outpatient visits for influenza-like illness (ILI)
#                    was 7.1%, which is above the national baseline of 2.2%. All 10 regions
#                    reported ILI at or above region-specific baseline levels.
# ILINet State Activity Indicator Map: New York City, the District of Columbia, and 42 states experienced
#                    high ILI activity; Puerto Rico and two states experienced moderate
#                    ILI activity; three states experienced low ILI activity; and three
#                    states experienced minimal ILI activity
#####

library(ggplot2)
library(plotly)
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##   last_plot
```

```
## The following object is masked from 'package:stats':
##
##   filter
```

```
## The following object is masked from 'package:graphics':
##
##   layout
```

```
library(RColorBrewer)
library(maps)
library("ggmap")
```

```
##
## Attaching package: 'ggmap'
```

```
## The following object is masked from 'package:plotly':  
##  
##     wind
```

```
library("maptools")
```

```
## Loading required package: sp
```

```
## Checking rgeos availability: FALSE  
##     Note: when rgeos is not available, polygon geometry      computations in maptools depend on gpclib,  
##     which has a restricted licence. It is disabled by default;  
##     to enable gpclib, type gpclibPermit()
```

```
library(sqldf)
```

```
## Loading required package: gsubfn
```

```
## Loading required package: proto
```

```
## Loading required package: RSQLite
```

```
library(RSQLite)  
library(tidyr)  
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##     filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##     intersect, setdiff, setequal, union
```

```
#####
# Influenza Positive Tests Reported to CDC by U.S Clinical Laboratories, National Summary, 2017-2018 Season
#####

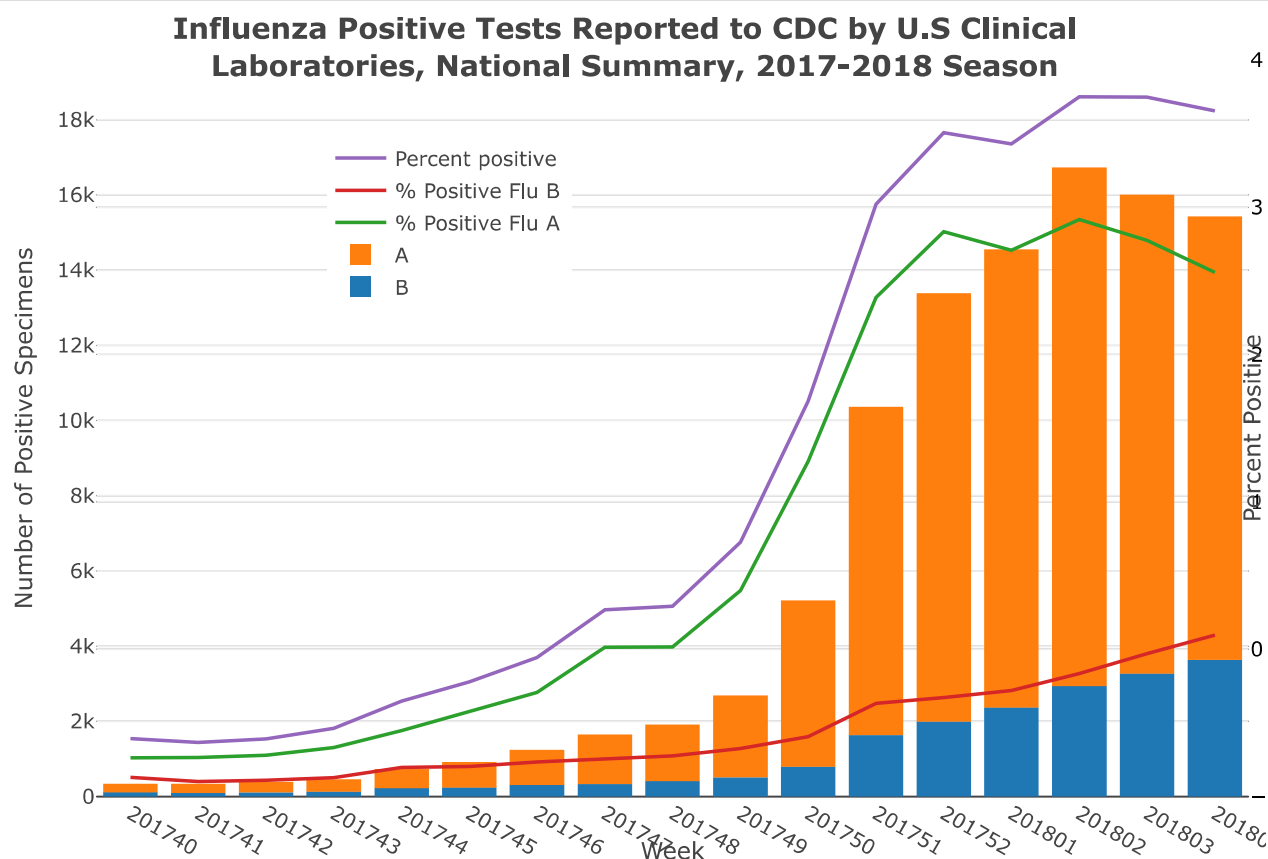
influenza_data1 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\Influenza national summary.csv")

# At the moment, ggplot is interpreting the date data as continuous date scale.
# So, turn the date data into a factor

influenza_data1$Week = as.factor(influenza_data1$Week)

ay <- list(
  tickfont = list(color = "black"),
  overlaying = "y",
  side = "right",
  title = "Percent Positive",
  zeroline = FALSE
)

plot_ly(influenza_data1, x = ~Week, y = ~Total.B, name = "B", type = 'bar', width = 750, height = 500) %>%
  add_trace(y = ~Total.A, name = "A") %>%
  add_lines(x = ~Week, y = ~700*Percent.Positive.A, name = "% Positive Flu A") %>%
  add_lines(x = ~Week, y = ~700*Percent.Positive.B, name = "% Positive Flu B") %>%
  add_lines(x = ~Week, y = ~700*Percent.Positive, name = "Percent positive") %>%
  layout(legend = list(x = 0.2, y = 0.9),
    title = "<b>Influenza Positive Tests Reported to CDC by U.S Clinical <br> Laboratories, National Summary, 2017-2018 Season</b>",
    yaxis = list(title="Number of Positive Specimens"), yaxis2 = ay,
    xaxis = list(title="Week", barmode = 'stack'))
```



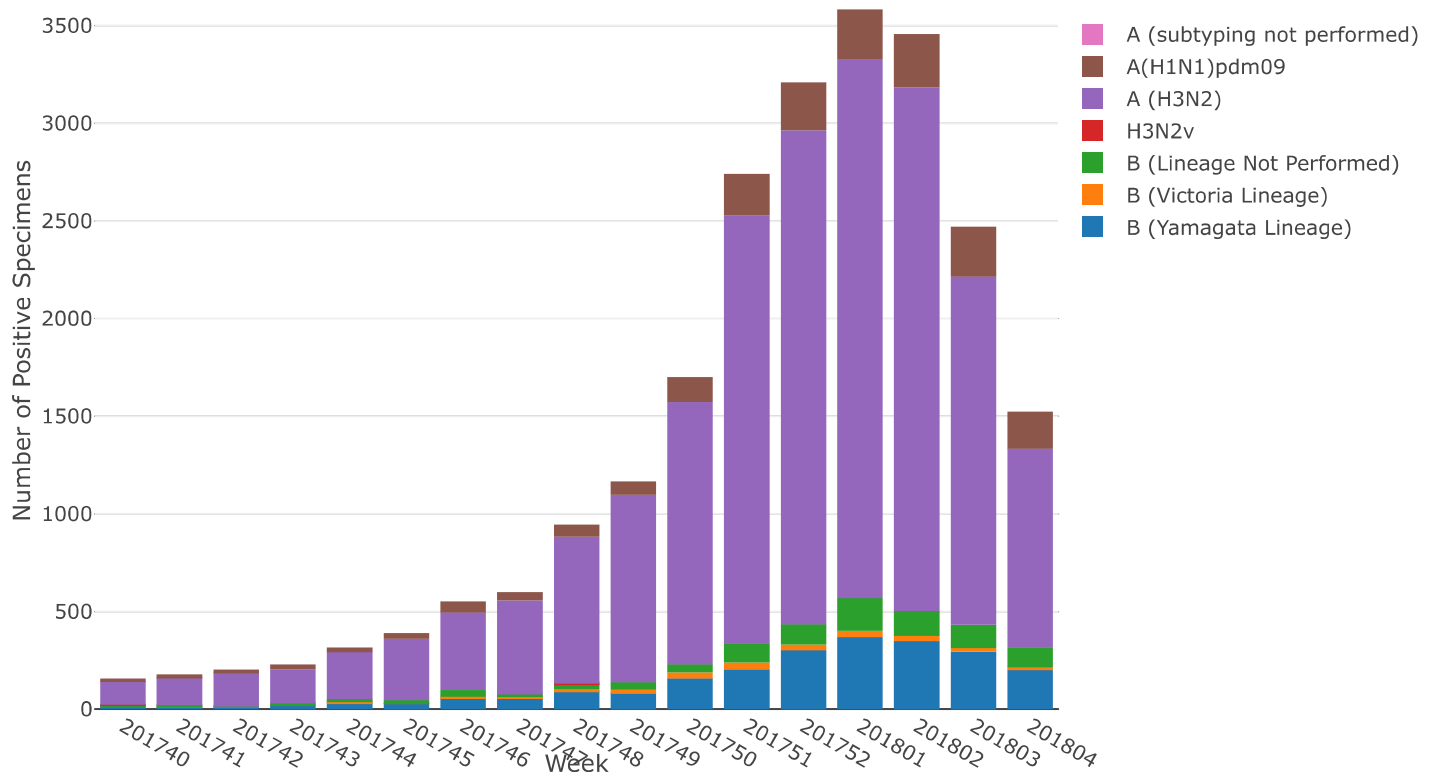
```
#####
# Influenza Positive Tests Reported to CDC by U.S Public Health Laboratories, National Summary, 2017-2018
# Season
#####

influenza_data2 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\Influenza national summary_PH.csv")

influenza_data2$Week = as.factor(influenza_data2$Week)

plot_ly(influenza_data2, x = ~Week, y = ~BYAM, name = "B (Yamagata Lineage)", type = 'bar', width = 860, height = 500)%>%
  add_trace(y = ~BVIC, name = "B (Victoria Lineage)") %>%
  add_trace(y = ~B, name = "B (Lineage Not Performed)") %>%
  add_trace(y = ~A_H3N2v, name = "H3N2v") %>%
  add_trace(y = ~A_H3, name = "A (H3N2)") %>%
  add_trace(y = ~A_H1N1_pdm09, name = "A(H1N1)pdm09") %>%
  add_trace(y = ~A_unable_to_sub_type, name = "A (subtyping not performed)") %>%
  layout(legend = list(x = 1, y = 0.95),
    title = "<b>Influenza Positive Tests Reported to CDC by U.S Public Health <br> Laboratories, National Summary, 2017-2018 Season</b>",
    yaxis = list(title="Number of Positive Specimens"),
    xaxis = list(title="Week", barmode = 'stack'))
```

Influenza Positive Tests Reported to CDC by U.S Public Health Laboratories, National Summary, 2017-2018 Season



```
#####
# Influenza Virus Characterization
#####

pie_data <- data.frame(colnames(influenza_data2[,c(-1,-2,-5,-10)]),colSums(influenza_data2[,c(-1,-2,-5,-10)]))

colnames(pie_data) = c("Categories","Sum")
pie_data$Name = c("Influenza A(H1N1)pdm09","Influenza A(H3N2)","Influenza A(subtype unknown)",
                  "Influenza B (lineage not determined)", "Influenza B Victoria", "Influenza B Yamagata")

pie_data$Percent = round((pie_data$Sum/sum(pie_data$Sum))*100,1)

main_pie <- plot_ly(pie_data, labels = ~Name, values = ~Sum, type = 'pie',
                    domain = list(
                      x = c(0, 0.85),
                      y = c(0.20, 0.91)),
                    showlegend = T, width = 350, height = 350,
                    textposition = 'inside',
                    text = ~Sum,
                    textinfo = 'text',
                    insidetextfont = list(color = 'black', size = 10),
                    marker = list(colors = brewer.pal(6, "Accent"),
                                   line = list(color = 'black', width = 0.9))) %>%
  add_pie()%>%
  layout(legend = list(x = 0.1, y = -0.35),
         title = 'Influenza Positive Specimens Reported by<br>U.S Public Health Laboratories,<br>Cumulative, 2017-2
018 Season',
         font = list(size=8.5))

# Subplots of Pie
pie_cum_data = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\Genetic04.csv")

# Delete rows where Number is 0, to avoid showing it up in pie chart
pie_cum_data = pie_cum_data[pie_cum_data$Number != 0,]

p1 = list(domain = list(
  x = c(0, 0.45),
  y = c(0.51, 0.9)),
  data = pie_cum_data[pie_cum_data$Sub_type == 'H3',],
  labels = ~Genetic_Group,
  values = ~Number,
  marker = list(colors = brewer.pal(3, "Reds"),
                line = list(color = 'black', width = 1)),
  type = "pie",
  text = ~paste(Genetic_Group,Number,Percent_of_Sub_type_Total, sep = "<br />"),
  textinfo = 'text',
  textposition = 'outside',
  hoverinfo = 'text'
)

p2 = list(domain = list(
  x = c(0.50, 0.95),
  y = c(0.51, 0.9)),
  data = pie_cum_data[pie_cum_data$Sub_type == 'H1pdm09',],
  labels = ~Genetic_Group,
  values = ~Number,
  marker = list(colors = brewer.pal(3, "PuBuGn"),
                line = list(color = 'black', width = 1)),
  type = "pie",
  text = ~paste(Genetic_Group,Number,Percent_of_Sub_type_Total, sep = "<br />"),
  textinfo = 'text',
  textposition = 'outside',
  hoverinfo = 'text')
```

```

p3 = list(domain = list(
  x = c(0, 0.45),
  y = c(0, 0.39)),
  data = pie_cum_data[pie_cum_data$Sub_type == 'B/Victoria',],
  labels = ~Genetic_Group,
  values = ~Number,
  marker = list(colors = brewer.pal(3, "YlGn"),
    line = list(color = 'black', width = 1)),
  type = "pie",
  text = ~paste(Genetic_Group,Number,Percent_of_Sub_type_Total, sep = "<br />"),
  textinfo = 'text',
  textposition = 'outside',
  hoverinfo = 'text')

p4 = list(domain = list(
  x = c(0.50, 0.95),
  y = c(0, 0.39)),
  data = pie_cum_data[pie_cum_data$Sub_type == 'B/Yamagata',],
  labels = ~Genetic_Group,
  values = ~Number,
  marker = list(colors = brewer.pal(3, "Greens"),
    line = list(color = 'black', width = 1)),
  type = "pie",
  text = ~paste(Genetic_Group,Number,Percent_of_Sub_type_Total, sep = "<br />"),
  textinfo = 'text',
  textposition = 'outside',
  hoverinfo = 'text')

layout <- list(
  showlegend = FALSE,
  title = "Sequence Results, by Genetic HA Clade/Subclade, of Specimens
    Submitted to CDC by U.S. Public Health Laboratories, Cumulative<br>2017-2018 Season",
  font = list(size=8.4))

p <- plot_ly()
p <- add_trace(p, domain=p1$domain, data = p1$data, labels = p1$labels,
  values = p1$values, marker = p1$marker,type = p1$type,
  text = p1$text, textinfo=p1$textinfo, textposition = p1$textposition,
  hoverinfo = p1$hoverinfo)

p <- add_trace(p, domain=p2$domain, data = p2$data, labels = p2$labels,
  values = p2$values, marker = p2$marker,type = p2$type,
  text = p2$text, textinfo=p2$textinfo, textposition = p2$textposition,
  hoverinfo = p2$hoverinfo)

p <- add_trace(p, domain=p3$domain, data = p3$data, labels = p3$labels,
  values = p3$values, marker = p3$marker,type = p3$type,
  text = p3$text, textinfo=p3$textinfo, textposition = p3$textposition,
  hoverinfo = p3$hoverinfo)

p <- add_trace(p, domain=p4$domain, data = p4$data, labels = p4$labels,
  values = p4$values, marker = p4$marker,type = p4$type,
  text = p4$text, textinfo=p4$textinfo, textposition = p4$textposition,
  hoverinfo = p4$hoverinfo)

p <- layout(p, showlegend=layout$showlegend, title=layout$title, font = layout$font)

# Add subtitles to subplots

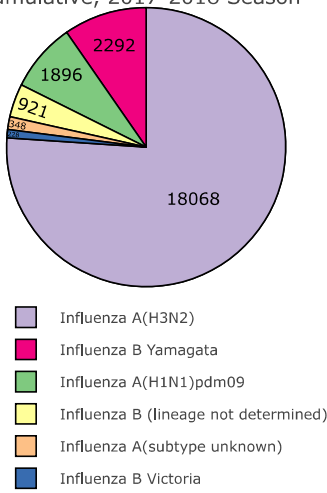
sub_pie <- p %>% layout(annotations = list(
  list(x = 0.17 , y = 0.96, text = "Influenza A(H3N2)", showarrow = F, font = list(size=10)),
  list(x = 0.80 , y = 0.96, text = "Influenza A(H1N1)pdm09", showarrow = F,font = list(size=10)),
  list(x = 0.17 , y = 0.43, text = "Influenza B Victoria", showarrow = F,font = list(size=10)),

```

```
list(x = 0.80 , y = 0.43, text = "Influenza B Yamagata", showarrow = F,font = list(size=10)))
)

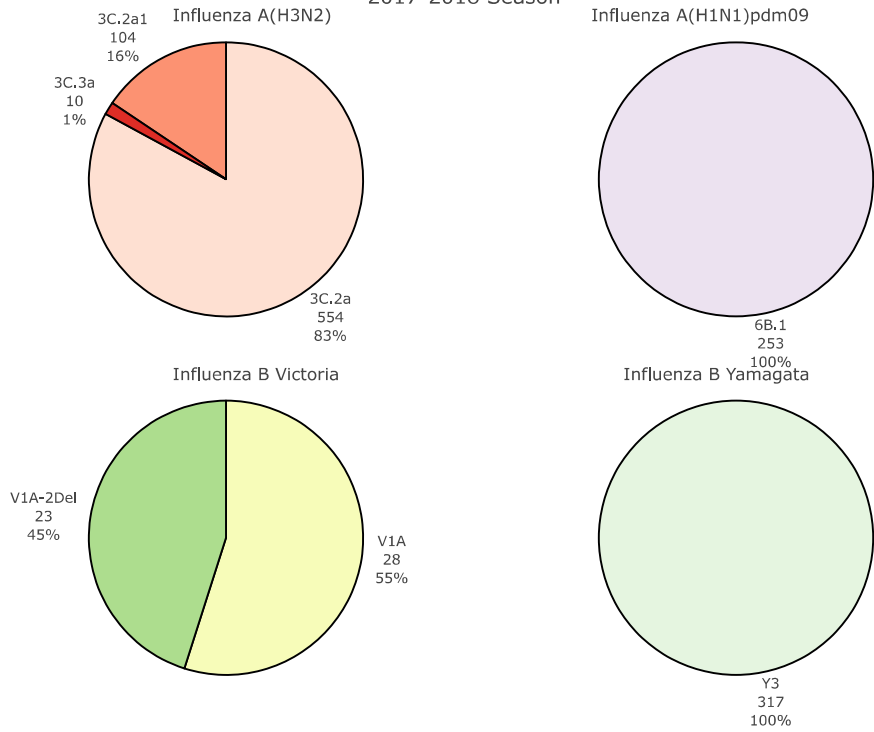
main_pie
```

Influenza Positive Specimens Reported by
U.S Public Health Laboratories,
Cumulative, 2017-2018 Season



```
sub_pie
```

Sequence Results, by Genetic HA Clade/Subclade, of Specimens
Submitted to CDC by U.S. Public Health Laboratories, Cumulative
2017-2018 Season



```
#####
# Pneumonia and Influenza (P&I) Mortality Surveillance
# National Center for Health Statistics (NCHS) mortality surveillance data as of February 1, 2018
#####

PnI_data = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\NCHSData04.csv")

PnI_data = PnI_data[(PnI_data$Year != '2009' &
                     PnI_data$Year != '2010' &
                     PnI_data$Year != '2011'),]

PnI_data = unite(PnI_data, Week, c('Year', 'Week'), sep = "-")

# To select data which are used for the graph to show, so that it is replicated
PnI_data = sqldf("SELECT * FROM PnI_data WHERE WEEK > '2012-39' AND Week <= '2018-2'")

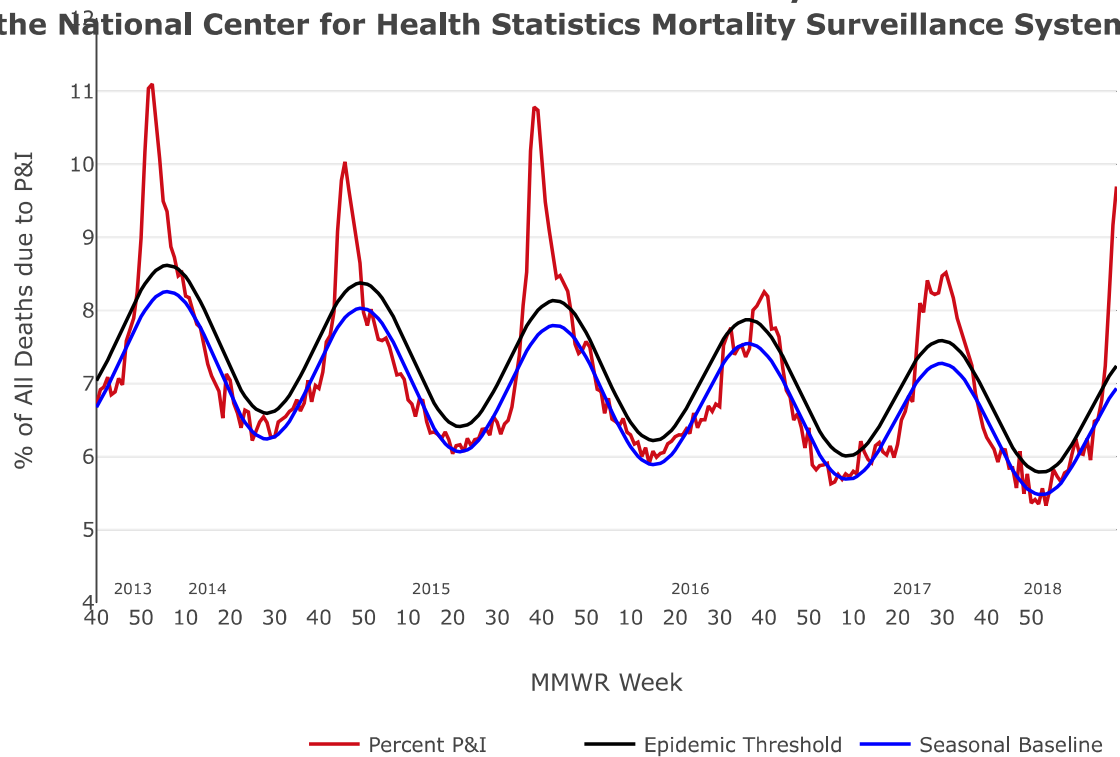
# weeks which are not relevant for this plot
PnI_data = PnI_data[(PnI_data$Week != '2012-4' & PnI_data$Week != '2012-5' &
                     PnI_data$Week != '2012-6' & PnI_data$Week != '2012-7' &
                     PnI_data$Week != '2012-8' & PnI_data$Week != '2012-9'),]

plot_ly(PnI_data, y = ~Percent.of.Deaths.Due.to.Pneumonia.and.Influenza,
        name = 'Percent P&I',
        type = 'scatter',
        mode = 'lines',
        line = list(color = 'rgb(205, 12, 24)', width = 2))%>%
add_trace(y = ~Threshold,
        name = 'Epidemic Threshold',
        type = 'scatter',
        mode = 'lines',
        line = list(color = 'black', width = 2))%>%
add_trace(y = ~Expected,
        name = 'Seasonal Baseline',
        type = 'scatter',
        mode = 'lines',
        line = list(color = 'blue', width = 2))%>%
layout(
  legend = list(orientation = 'h', x = 0.2, y = -0.2),
  title = '<b>Pneumonia and Influenza Mortality from<br>the National Center for Health Statistics Mortality Surveillance System',
  xaxis = list(
    tickvals = seq(from=0, to=260, by=12),
    ticktext = c("40", "50", "10", "20", "30", "40", "50",
                  "10", "20", "30", "40", "50", "10", "20",
                  "30", "40", "50", "10", "20", "30", "40", "50"),
    title = 'MMWR Week',
    showgrid = FALSE),
  yaxis = list(
    title = '% of All Deaths due to P&I', range = c(4,12)),
  annotations = list(list(text = "2013", showarrow = F, x = 10, y = 4.2, font = list(size = 9)),
                      list(text = "2014", showarrow = F, x = 30, y = 4.2, font = list(size = 9)),
                      list(text = "2015", showarrow = F, x = 90, y = 4.2, font = list(size = 9)),
                      list(text = "2016", showarrow = F, x = 160, y = 4.2, font = list(size = 9)),
```



```
list(text = "2017",showarrow =F, x = 220, y = 4.2, font = list(size = 9)),  
list(text = "2018",showarrow =F, x = 255, y = 4.2, font = list(size = 9)))
```

Pneumonia and Influenza Mortality from the National Center for Health Statistics Mortality Surveillance System



```
#####
# Influenza-Associated Pediatric Mortality
#####

pedriatic_mortality = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\PedFluDeath_WeeklyData.csv")

# Calculating number of deaths for each season

# 2014-2015 Season
deaths_2014_15 = sum(pedriatic_mortality[pedriatic_mortality$SEASON == '2014-15',]$PREVIOUS.WEEKS.DEATHS)

# 2015-2016 Season
deaths_2015_16 = sum(pedriatic_mortality[pedriatic_mortality$SEASON == '2015-16',]$PREVIOUS.WEEKS.DEATHS)

# 2016-2017 Season
deaths_2016_17 = sum(pedriatic_mortality[pedriatic_mortality$SEASON == '2016-17',]$PREVIOUS.WEEKS.DEATHS)

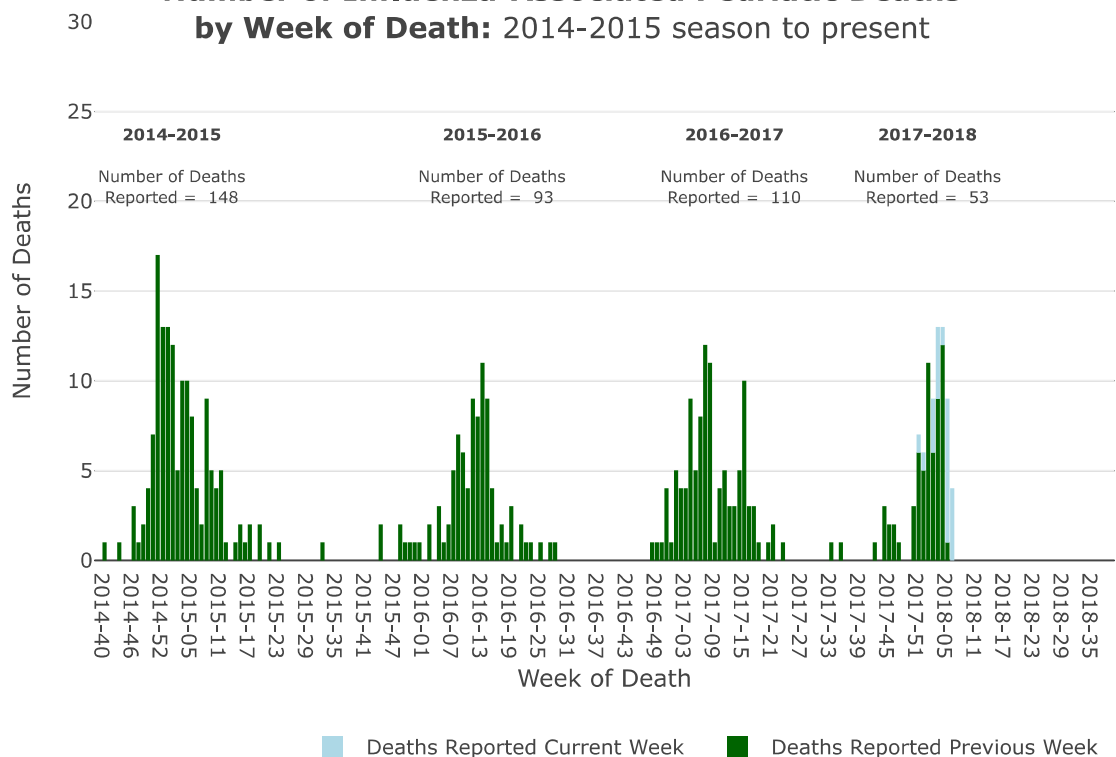
# as we need to display data upto 4th week of 2018
deaths_18 = sum(pedriatic_mortality[(pedriatic_mortality$WEEK.NUMBER == '2018-01' |
                                     pedriatic_mortality$WEEK.NUMBER == '2018-02' |
                                     pedriatic_mortality$WEEK.NUMBER == '2018-03' |
                                     pedriatic_mortality$WEEK.NUMBER == '2018-04'),]$CURRENT.WEEK.DEATHS)

deaths_17 = sum(pedriatic_mortality[pedriatic_mortality$WEEK.NUMBER %in% c('2017-40','2017-41','2017-42','2017-43',
                                     '2017-44','2017-45','2017-46','2017-47',
                                     '2017-48','2017-49','2017-50','2017-51',
                                     '2017-52','2018-01','2018-02','2018-03',
                                     '2018-04'),]$PREVIOUS.WEEKS.DEATHS)

# 2017-2018 Season
deaths_2017_18 = deaths_17 - deaths_18

plot_ly(pedriatic_mortality, x = ~WEEK.NUMBER, y = ~PREVIOUS.WEEKS.DEATHS, name = 'Deaths Reported Previous Week',
        type = 'bar',
        color = I("dark green"))%>%
  add_trace(y = ~CURRENT.WEEK.DEATHS, name = 'Deaths Reported Current Week',color = I("light blue")) %>%
  layout(legend = list(orientation = 'h', x = 0.2, y = -0.3),
        title = "<b>Number of Influenza-Associated Pedriatic Deaths<br>by Week of Death:</b> 2014-2015 season to p
resent", yaxis = list(title="Number of Deaths",range = c(0, 30)),
        xaxis = list(title="Week of Death"), barmode = 'stack',
        annotations = list(list(text = ~paste("<b>2014-2015</b><br><br>Number of Deaths<br>Reported = ",deaths_201
4_15),showarrow =F, x = 14, y=22, font = list(size = 9.5)),
                          list(text = ~paste("<b>2015-2016</b><br><br>Number of Deaths<br>Reported = ",deaths_201
5_16),showarrow =F, x = 80, y=22, font = list(size = 9.5)),
                          list(text = ~paste("<b>2016-2017</b><br><br>Number of Deaths<br>Reported = ",deaths_201
6_17),showarrow =F, x = 130, y=22, font = list(size = 9.5)),
                          list(text = ~paste("<b>2017-2018</b><br><br>Number of Deaths<br>Reported = ",deaths_201
7_18),showarrow =F, x = 170, y=22, font = list(size = 9.5))))
```

Number of Influenza-Associated Pedriatic Deaths by Week of Death: 2014-2015 season to present



```
#####
# Outpatient Illness Surveillance
#####

# 2009-2010 Season
ill_data_200910 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\FluView_LineChart_Data_2009_10.csv")
ill_data_200910 = ill_data_200910[,c(1,2,11)]
ill_data_200910 = unite(ill_data_200910, Week, c('YEAR', 'WEEK'), sep = "")
# empty data frame is created to match the rows of 2014-2015 season as it was a Leap year which has 53 weeks
df1 <- data.frame(matrix(ncol = 2, nrow = 1))
colnames(df1) = c("Week", "X..WEIGHTED.ILI")
ill_data_200910 = rbind(ill_data_200910, df1)

# 2011-2012 Season
ill_data_201112 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\FluView_LineChart_Data_2011_12.csv")
ill_data_201112 = ill_data_201112[,c(1,2,11)]
ill_data_201112 = unite(ill_data_201112, Week, c('YEAR', 'WEEK'), sep = "")
df2 <- data.frame(matrix(ncol = 2, nrow = 1))
colnames(df2) = c("Week", "X..WEIGHTED.ILI")
ill_data_201112 = rbind(ill_data_201112, df2)

# 2014-2015 Season
ill_data_201415 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\FluView_LineChart_Data_2014_15.csv")
ill_data_201415 = ill_data_201415[,c(1,2,11)]
ill_data_201415 = unite(ill_data_201415, Week, c('YEAR', 'WEEK'), sep = "")

# 2015-2016 Season
ill_data_201516 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\FluView_LineChart_Data_2015_16.csv")
ill_data_201516 = ill_data_201516[,c(1,2,11)]
ill_data_201516 = unite(ill_data_201516, Week, c('YEAR', 'WEEK'), sep = "")
df3 <- data.frame(matrix(ncol = 2, nrow = 1))
colnames(df3) = c("Week", "X..WEIGHTED.ILI")
ill_data_201516 = rbind(ill_data_201516, df3)

# 2016-2017 Season
ill_data_201617 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\FluView_LineChart_Data_2016_17.csv")
ill_data_201617 = ill_data_201617[,c(1,2,11)]
ill_data_201617 = unite(ill_data_201617, Week, c('YEAR', 'WEEK'), sep = "")
df4 <- data.frame(matrix(ncol = 2, nrow = 1))
colnames(df4) = c("Week", "X..WEIGHTED.ILI")
ill_data_201617 = rbind(ill_data_201617, df4)

# 2017-2018 Season
ill_data_201718 = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\FluView_LineChart_Data_2017_18.csv")
ill_data_201718 = ill_data_201718[,c(1,10)]
df5 <- data.frame(matrix(ncol = 2, nrow = 36))
colnames(df5) = c("Week", "Prct_Weighted.ILI")
ill_data_201718 = rbind(ill_data_201718, df5)

# Combine all data frames to one
ill_data = cbind(ill_data_200910, ill_data_201112, ill_data_201415, ill_data_201516, ill_data_201617, ill_data_201718)
colnames(ill_data) = c("Week_200910", "Percent_ILI_200910", "Week_201112", "Percent_ILI_201112",
                      "Week_201415", "Percent_ILI_201415", "Week_201516", "Percent_ILI_201516",
                      "Week_201617", "Percent_ILI_201617", "Week_201718", "Percent_ILI_201718")

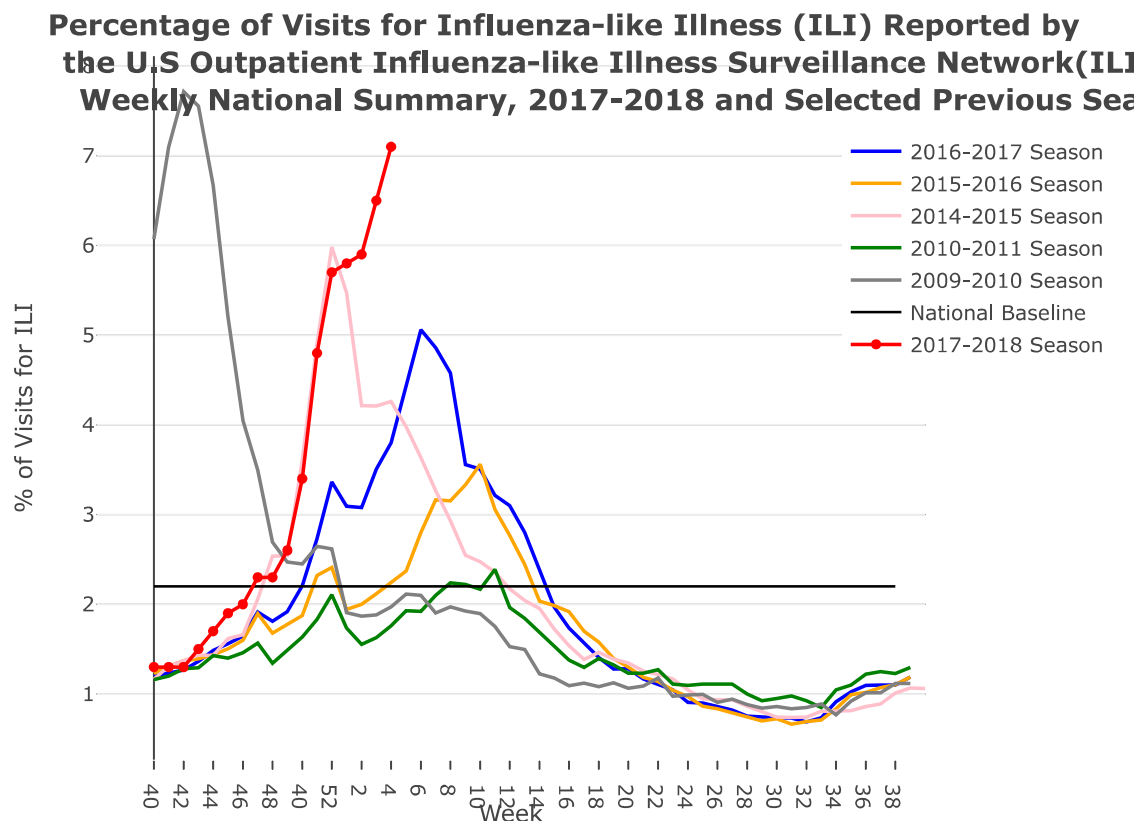
# To draw National Baseline
x1 = seq(from=0, to=51, by=2)
```

```

y1 = 2.2
data <- data.frame(x1, y1)

plot_ly(ill_data, y = ~Percent_ILI_201617, name = '2016-2017 Season',type= 'scatter',mode = 'lines',
  line = list(color = 'blue', width = 2)) %>%
  add_trace(y = ~Percent_ILI_201516, name = '2015-2016 Season',type= 'scatter',mode = 'lines',
  line = list(color = 'orange', width = 2))%>%
  add_trace(y = ~Percent_ILI_201415, name = '2014-2015 Season',type= 'scatter',mode = 'lines',
  line = list(color = 'pink', width = 2))%>%
  add_trace(y = ~Percent_ILI_201112, name = '2010-2011 Season',type= 'scatter',mode = 'lines',
  line = list(color = 'green', width = 2))%>%
  add_trace(y = ~Percent_ILI_200910, name = '2009-2010 Season', type= 'scatter',mode = 'lines',
  line = list(color = 'grey', width = 2))%>%
  add_trace(data, x = ~x1, y = ~y1, name = "National Baseline",type = 'scatter', mode = 'lines',
  line = list(color = 'black', width = 1.5))%>%
  add_trace(y = ~Percent_ILI_201718, name = '2017-2018 Season',type= 'scatter',mode = 'lines+markers',
  symbols = c("triangle-up"),line = list(color = 'red', width = 2))%>%
  layout(legend = list(x = 0.9, y = 0.9),
  title = '<b>Percentage of Visits for Influenza-like Illness (ILI) Reported by
  the U.S Outpatient Influenza-like Illness Surveillance Network(ILINet),
  Weekly National Summary, 2017-2018 and Selected Previous Seasons</b>',
  xaxis = list(
    ticks = "outside",
    tickvals = seq(from=0, to=51, by=2),
    ticktext = c("40","42","44","46","48","50","52","2","4","6","8","10","12",
      "14","16","18","20","22","24","26","28","30","32","34","36","38"),
    title = 'Week',
    showgrid = FALSE),
  yaxis = list( title = '% of Visits for ILI'))

```



```
#####
# ILINet State Activity Indicator Map
#####

influenza_div_data = read.csv("C:\\Users\\Monalisa Mishra\\Desktop\\DIC\\Lab1\\Part2\\StateDatabyWeekforMap_2017-18
week40-4.csv")

influenza_div_data = influenza_div_data %>% separate(ACTIVITY.LEVEL, into = c('ACTIVITY.LEVEL','ACTIVITY.LEVEL'), s
ep = 6)

# As we have a separate column for Level, do not require extra ACTIVITY LEVEL COLUMN
influenza_div_data = influenza_div_data[,c(-4)]

colnames(influenza_div_data) = c("STATENAME","URL","WEBSITE","ACTIVITY_LEVEL","ACTIVITY_LEVEL_LABEL","WEEKEND","WEE
K","SEASON")

# Change type from factor to num
influenza_div_data$ACTIVITY_LEVEL = as.numeric(influenza_div_data$ACTIVITY_LEVEL)

# Get the states with highest ILI Activity Level
influenza_div_data = sqldf("SELECT *, MAX(ACTIVITY_LEVEL) AS Max_level FROM influenza_div_data GROUP BY STATENAME")

influenza_div_data$region = tolower(influenza_div_data$STATENAME)

code = data.frame(setNames(state.abb, state.name))

code$STATENAME <- rownames(code)

colnames(code) = c("code","STATENAME")

map.df = merge(code,influenza_div_data, by = "STATENAME", all.x = T)

map.df$hover <- with(map.df, paste(STATENAME, '<br>', "ILI Activity Level:   ",
                                ACTIVITY_LEVEL_LABEL, '<br>',
                                "Clicking on State will take you to State of ",
                                STATENAME, "Department of Health",
                                '<br>',WEBSITE))

# give state boundaries a white border
l <- list(color = toRGB("white"), width = 2)

# specify some map projection/options
g <- list(
  scope = 'usa',
  projection = list(type = 'albers usa'),
  showlakes = TRUE,
  lakecolor = toRGB('white')
)

plot_geo(map.df, locationmode = 'USA-states') %>%
  add_trace(
    z = ~ACTIVITY_LEVEL,
    text = ~hover,
    locations = ~code,
    color = ~ACTIVITY_LEVEL,
    colors = brewer.pal(11, "RdYlGn")
  ) %>%
  colorbar(title = "ILI Activity Level") %>%
  layout(
    title = '<b>A Weekly Influenza Surveillance Report Prepared by the Influenza Division<br>Influenza-Like Illness
(ILI) Activity Level Indicator Determined by Data Reported to ILINet' ,
```

```
geo = g
)
```

Weekly Influenza Surveillance Report Prepared by the Influenza Division
Like Illness (ILI) Activity Level Indicator Determined by Data Reported

