Visualization

Alex Lee

5/17/2017

```
## Set Workspace
setwd("~/R/Classes/ECOL 8540 Intro/ECOL8540")

## Load Packages
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(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
## Load Data
mers <- read.csv('cases.csv')

# Correcting errors in mers data
mers$hospitalized[890] <- c('2015-02-20')
mers <- mers[-471,]

mers$onset2 <- ymd(mers$onset)
mers$hospitalized2 <- ymd(mers$hospitalized)
```

```
## Warning: 5 failed to parse.
```

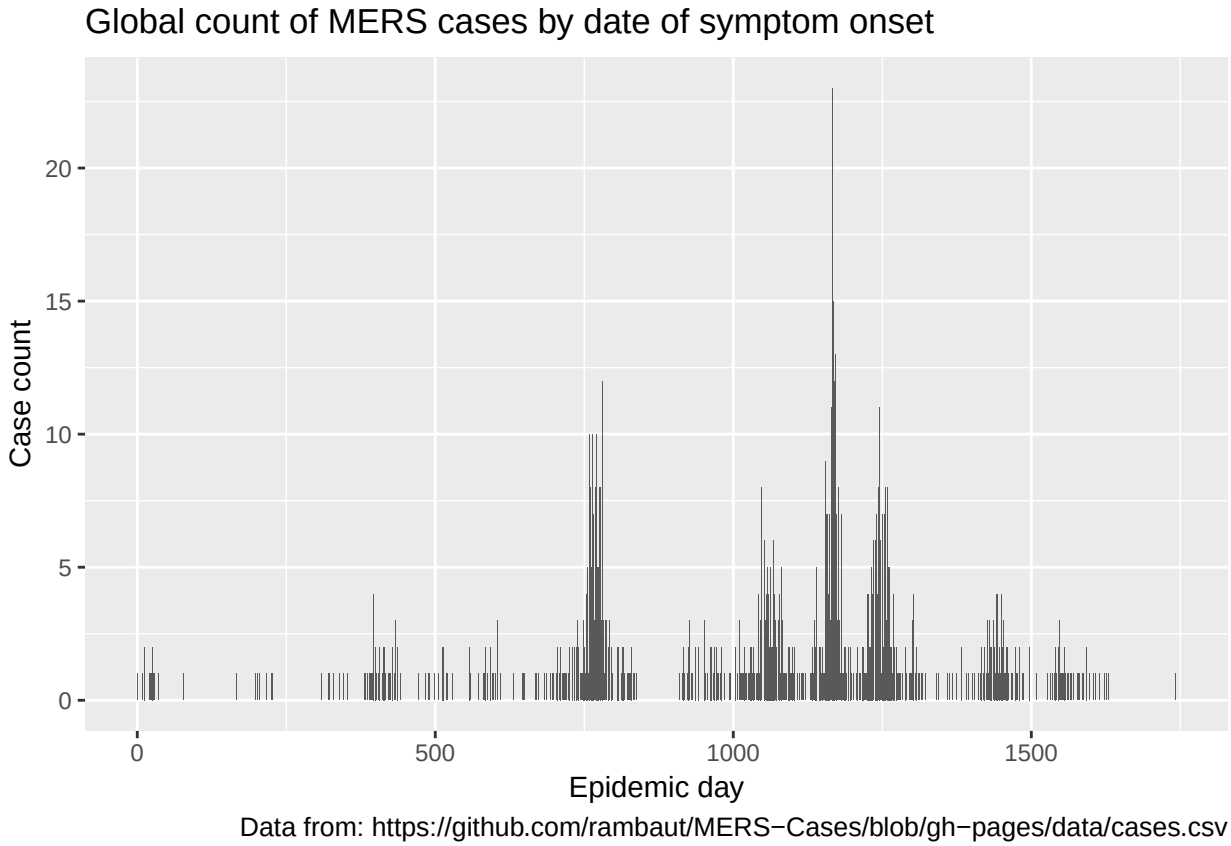
```
day0 <- min(na.omit(mers$onset2))

mers$epi.day <- as.numeric(mers$onset2 - day0)
```

Introduction

Making a Plot

```
# Making a Plot
ggplot(data = mers) +
  geom_bar(mapping = aes(x = epi.day)) +
  labs(x     = 'Epidemic day',
       y     = 'Case count',
       title = 'Global count of MERS cases by date of symptom onset', caption = "Data from: https://gitl
```

```
## Warning: Removed 535 rows containing non-finite values (stat_count).
```

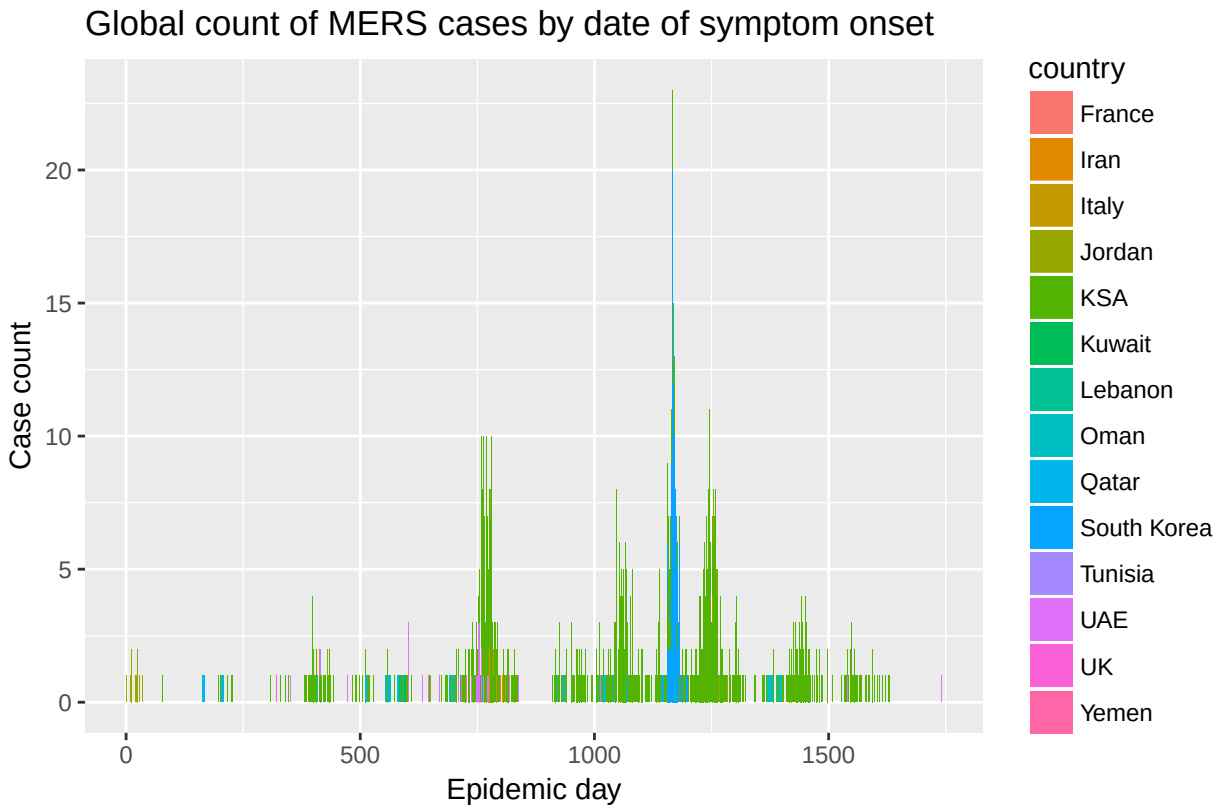


By Country

```
# fill = country
ggplot(data = mers) +
  geom_bar(mapping = aes(x    = epi.day,
                         fill = country)) +
  labs(x     = 'Epidemic day',
       y     = 'Case count',
       title = 'Global count of MERS cases by date of symptom onset', caption = "Data from: https://gitl
```

```
## Warning: Removed 535 rows containing non-finite values (stat_count).
```

```
## Warning: position_stack requires non-overlapping x intervals
```



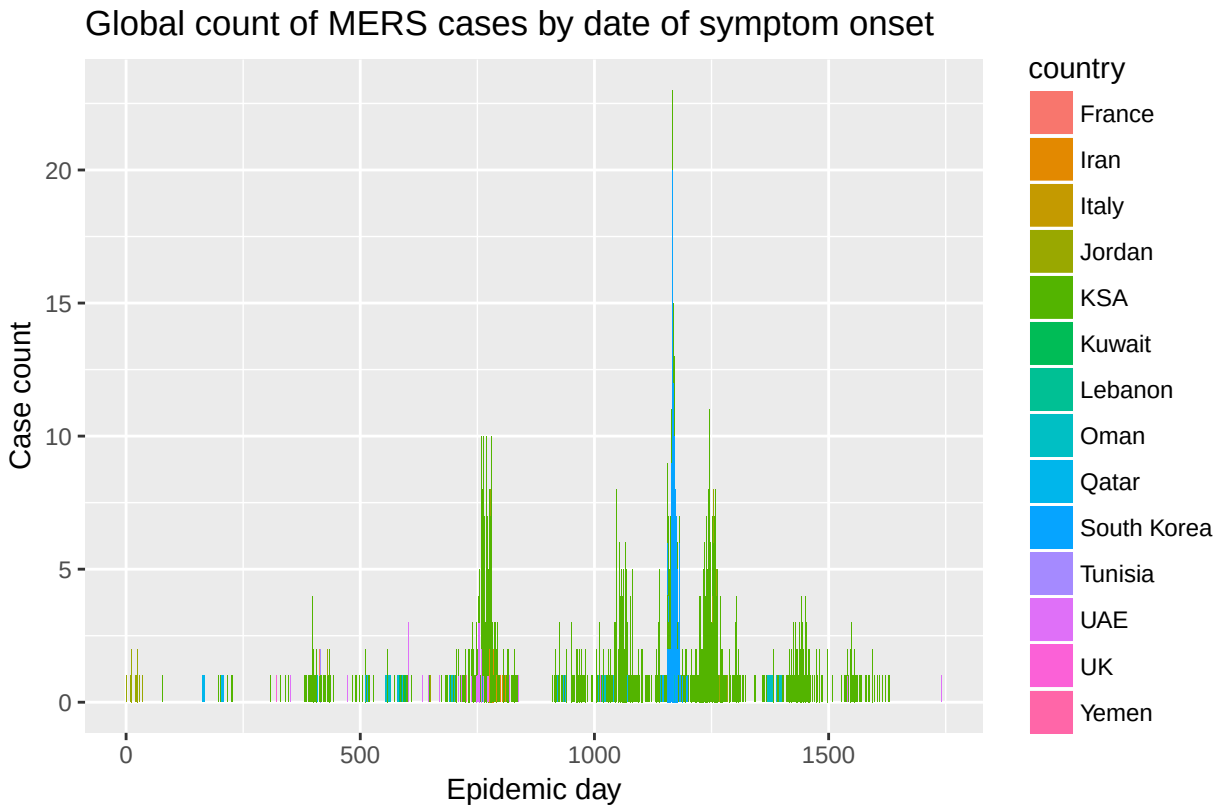
Data from: https://github.com/rambaut/MERS−Cases/blob/gh−pages/data/cases.csv

Position = Fill

```
# position = fill
ggplot(data = mers, position = fill) +
  geom_bar(mapping = aes(x    = epi.day,
                         fill = country)) +
  labs(x     = 'Epidemic day',
       y     = 'Case count',
       title = 'Global count of MERS cases by date of symptom onset', caption = "Data from: https://gitl
```

```
## Warning: Removed 535 rows containing non-finite values (stat_count).
```

```
## Warning: position_stack requires non-overlapping x intervals
```



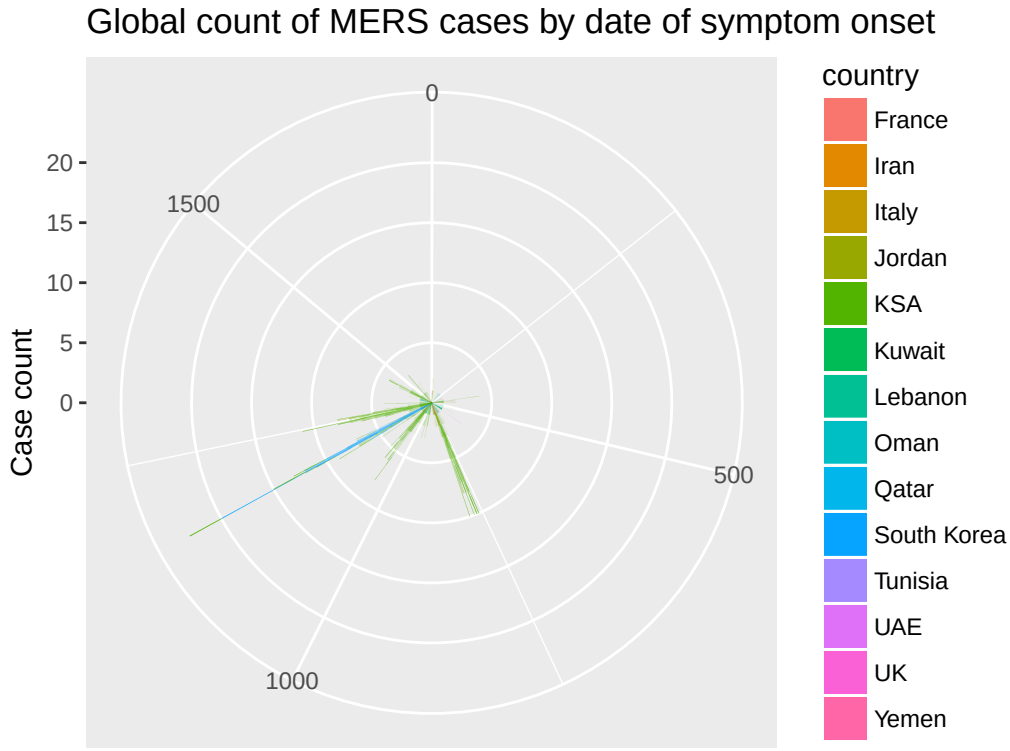
Data from: https://github.com/rambaut/MERS−Cases/blob/gh−pages/data/cases.csv

Polar Chart

```
# coord_flip and coord_polar
ggplot(data = mers, position = fill) +
  geom_bar(mapping = aes(x    = epi.day,
                         fill = country)) +
  labs(x     = 'Epidemic day',
       y     = 'Case count',
       title = 'Global count of MERS cases by date of symptom onset', caption = "Data from: https://gitl
  coord_flip() +
  coord_polar()
```

```
## Warning: Removed 535 rows containing non-finite values (stat_count).
```

```
## Warning: position_stack requires non-overlapping x intervals
```



n: https://github.com/rambaut/MERS−Cases/blob/gh−pages/data/cases.csv

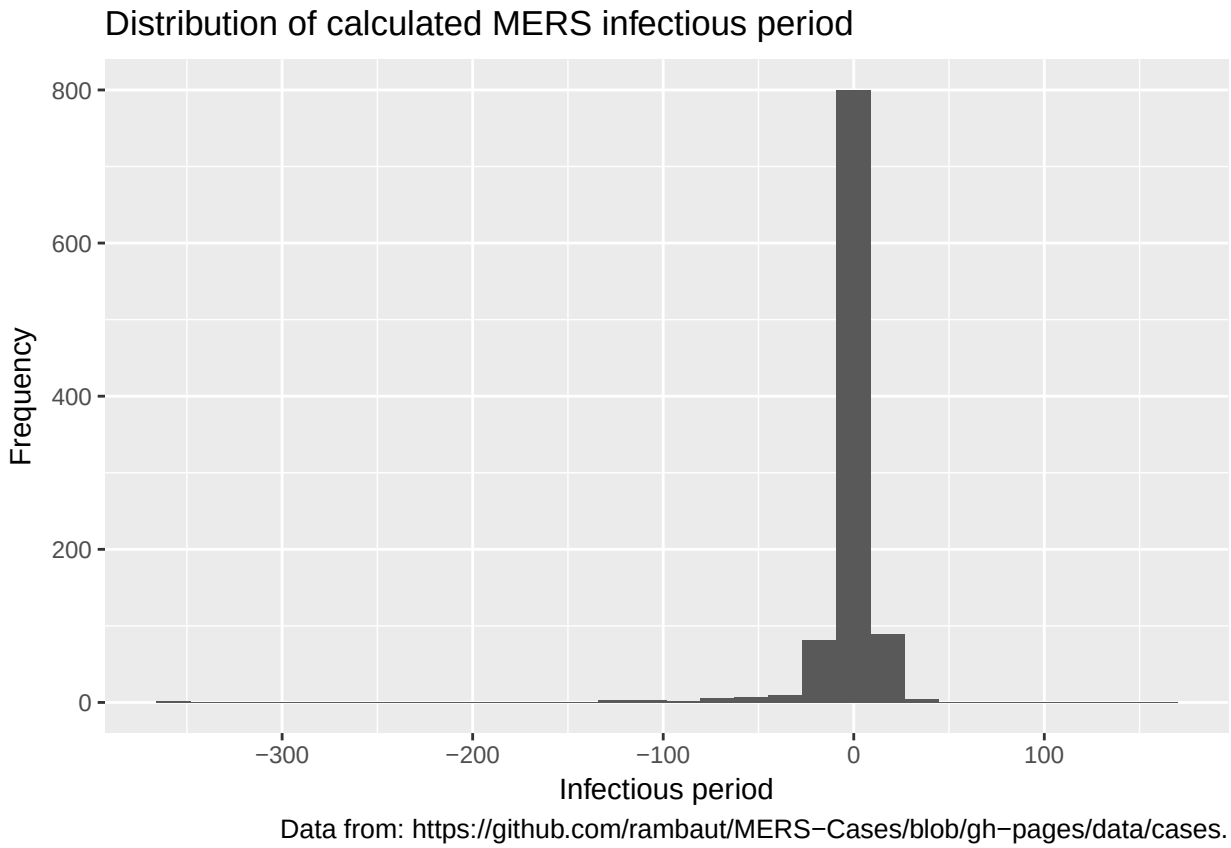
Univariate Plots

Distribution of of infectious period

```
## UNIVARIATE PLOTS ##
mers$infectious.period <- mers$hospitalized2 - mers$onset2
mers$infectious.period <- as.numeric(mers$infectious.period,
                                     units = "days")
ggplot(data = mers) +
  geom_histogram(aes(x = infectious.period)) +
  labs(x = 'Infectious period',
       y = 'Frequency',
       title = 'Distribution of calculated MERS infectious period',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 727 rows containing non-finite values (stat_bin).
```



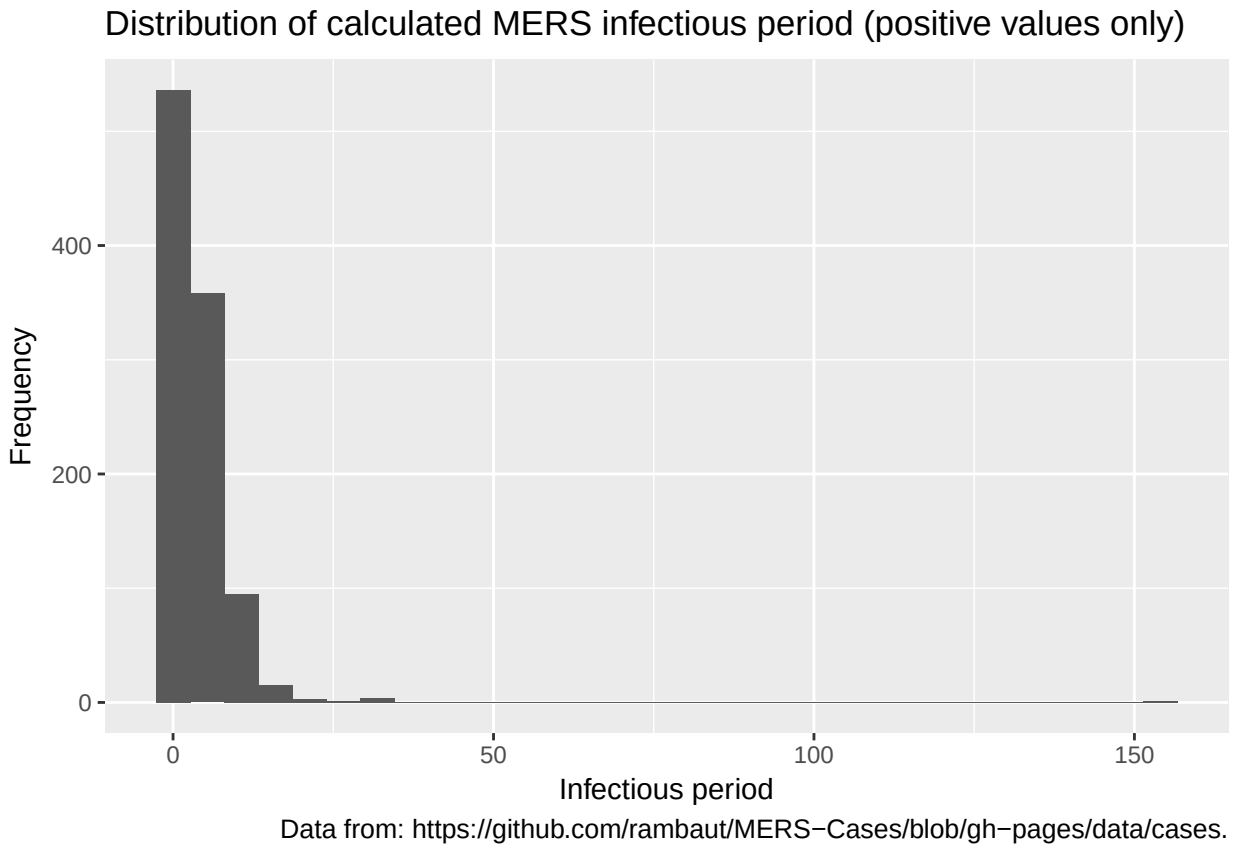
Distribution of of infectious period

Positive values only.

```
# Positive infectious periods only
mers$infectious.period2 <- ifelse(mers$infectious.period < 0,
                                  0,
                                  mers$infectious.period)
ggplot(data = mers) +
  geom_histogram(aes(x = infectious.period2)) +
  labs(x = 'Infectious period',
       y = 'Frequency',
       title = 'Distribution of calculated MERS infectious period (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

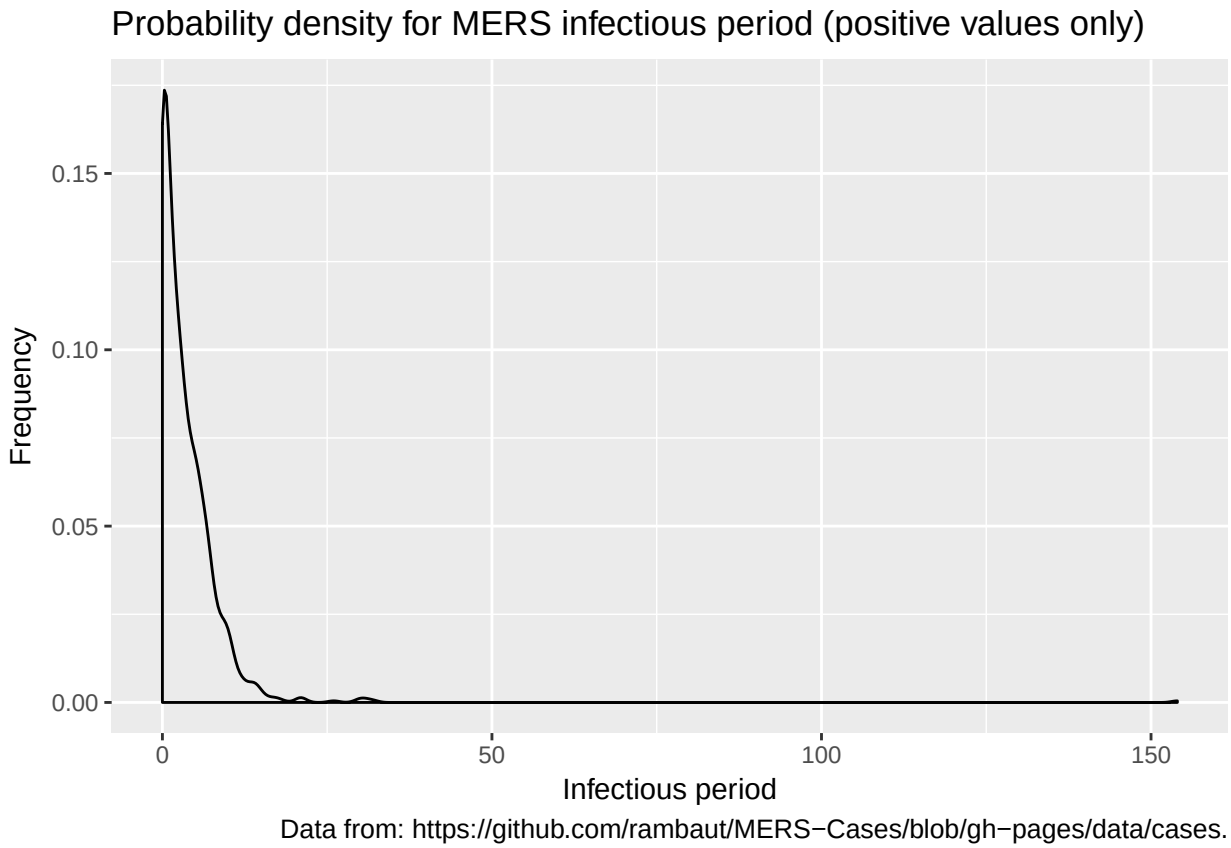
```
## Warning: Removed 727 rows containing non-finite values (stat_bin).
```



Density Plot

```
# Density Plot
ggplot(data = mers) +
  geom_density(mapping = aes(x = infectious.period2)) +
  labs(x = 'Infectious period',
       y = 'Frequency',
       title = 'Probability density for MERS infectious period (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 727 rows containing non-finite values (stat_density).
```

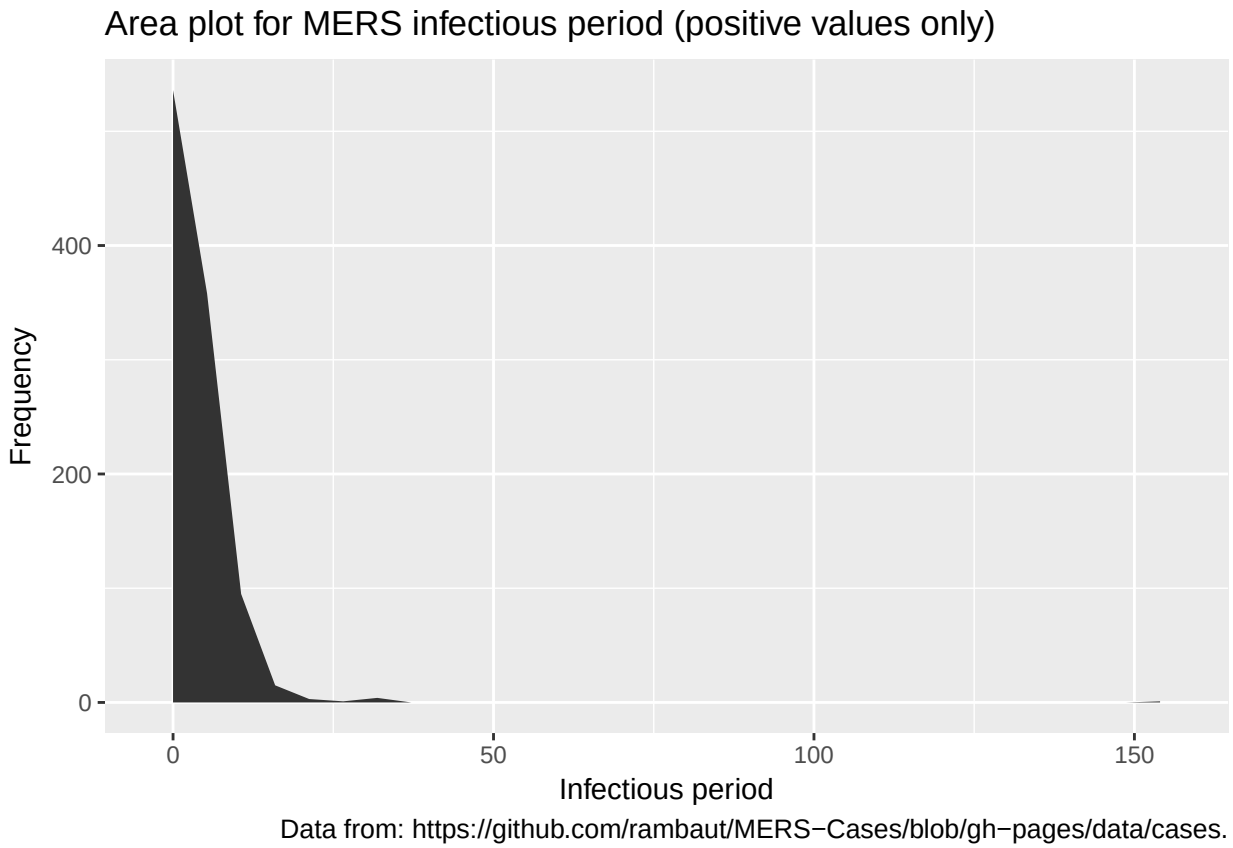


Area Plot

```
# Area Plot
ggplot(data = mers) +
  geom_area(stat = 'bin',
            mapping = aes(x = infectious.period2)) +
  labs(x = 'Infectious period',
       y = 'Frequency',
       title = 'Area plot for MERS infectious period (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 727 rows containing non-finite values (stat_bin).
```

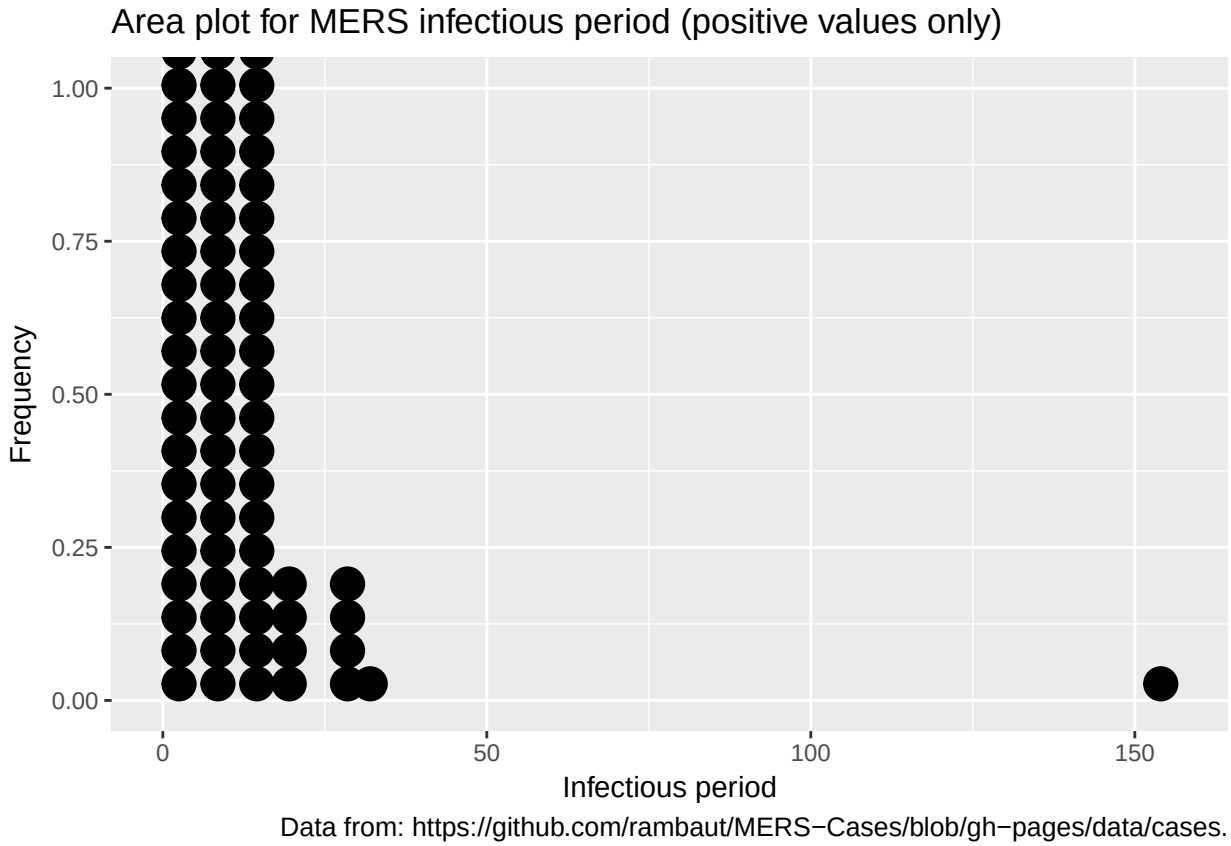


Dot Plot

```
# Dot Plot
ggplot(data = mers) +
  geom_dotplot(mapping = aes(x = infectious.period2)) +
  labs(x = 'Infectious period',
       y = 'Frequency',
       title = 'Area plot for MERS infectious period (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## `stat_bindot()` using `bins = 30`. Pick better value with `binwidth`.
```

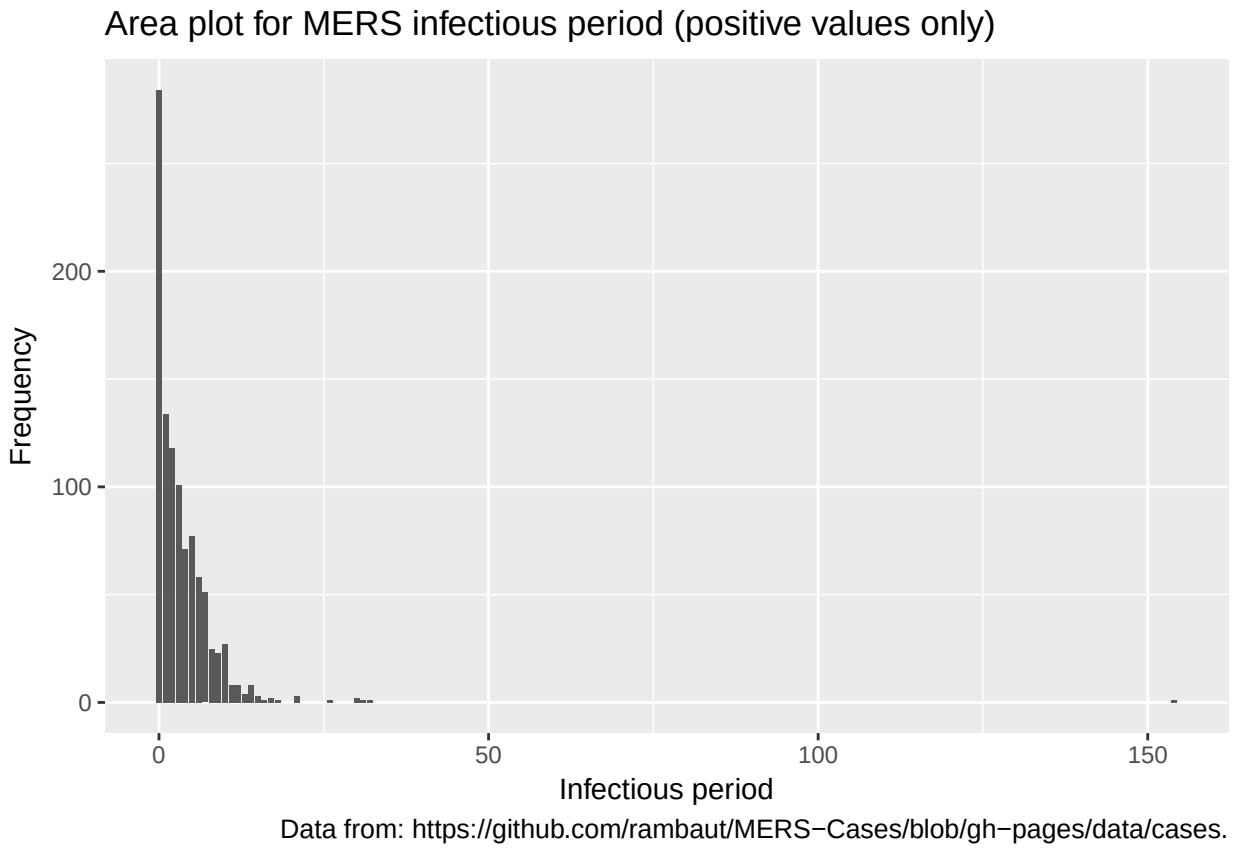
```
## Warning: Removed 727 rows containing non-finite values (stat_bindot).
```



Bar Plot

```
# Bar Plot
ggplot(data = mers) +
  geom_bar(mapping = aes(x = infectious.period2)) +
  labs(x = 'Infectious period',
       y = 'Frequency',
       title = 'Area plot for MERS infectious period (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 727 rows containing non-finite values (stat_count).
```

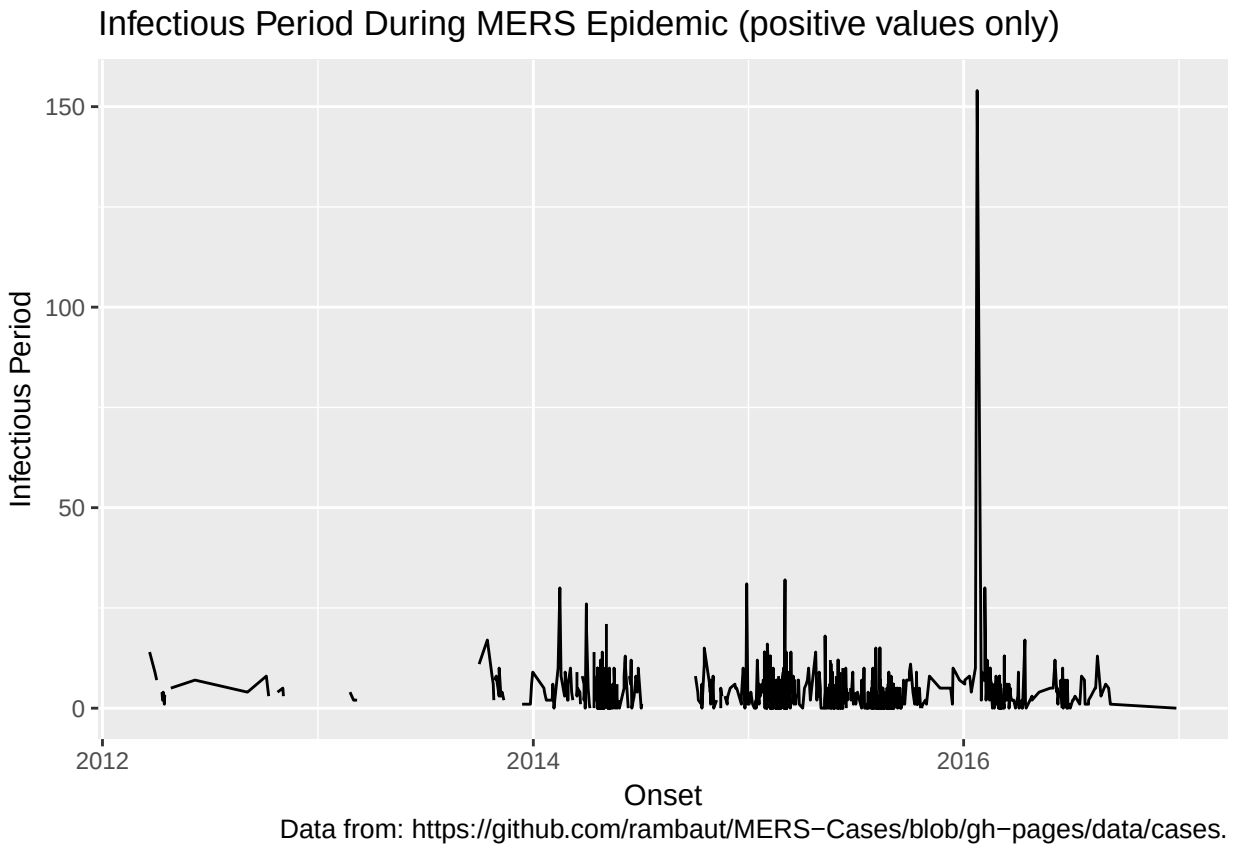


Bivariate Plots

Infectious Period Over Time

```
## BIVARIATE PLOTS ##

# Infectious period over the course of the MERS epidemic
ggplot(data = mers) +
  geom_line(mapping = aes(y = infectious.period2,
                          x = onset2))  +
  labs(x = 'Onset',
       y = 'Infectious Period',
       title = 'Infectious Period During MERS Epidemic (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 535 rows containing missing values (geom_path).
```

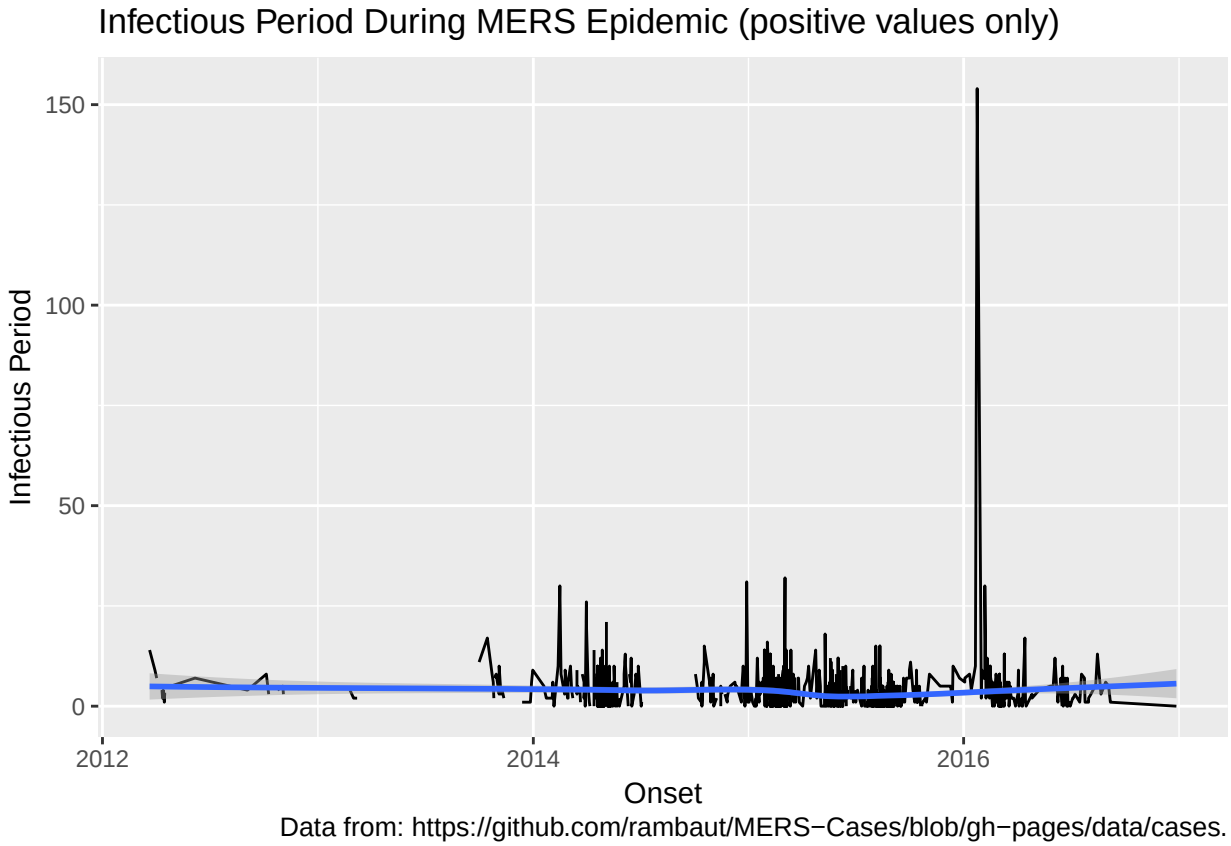


Societal Learning?

```
# Societal Learning
ggplot(data = mers) +
  geom_line(mapping = aes(y = infectious.period2,
                          x = onset2))  +
  geom_smooth(mapping = aes(y = infectious.period2,
                            x = onset2),
              method = "loess")  +
  labs(x = 'Onset',
       y = 'Infectious Period',
       title = 'Infectious Period During MERS Epidemic (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 727 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 535 rows containing missing values (geom_path).
```



By Country

```
# By Country
ggplot(data = mers) +
  geom_line(mapping = aes(y = infectious.period2,
                          x = onset2))  +
  geom_smooth(mapping = aes(y = infectious.period2,
                            x = onset2,
                            fill = country),
              method = "loess")  +
  labs(x = 'Onset',
       y = 'Infectious Period',
       title = 'Infectious Period During MERS Epidemic (positive values only)',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 727 rows containing non-finite values (stat_smooth).
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 16199
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 27.555
```

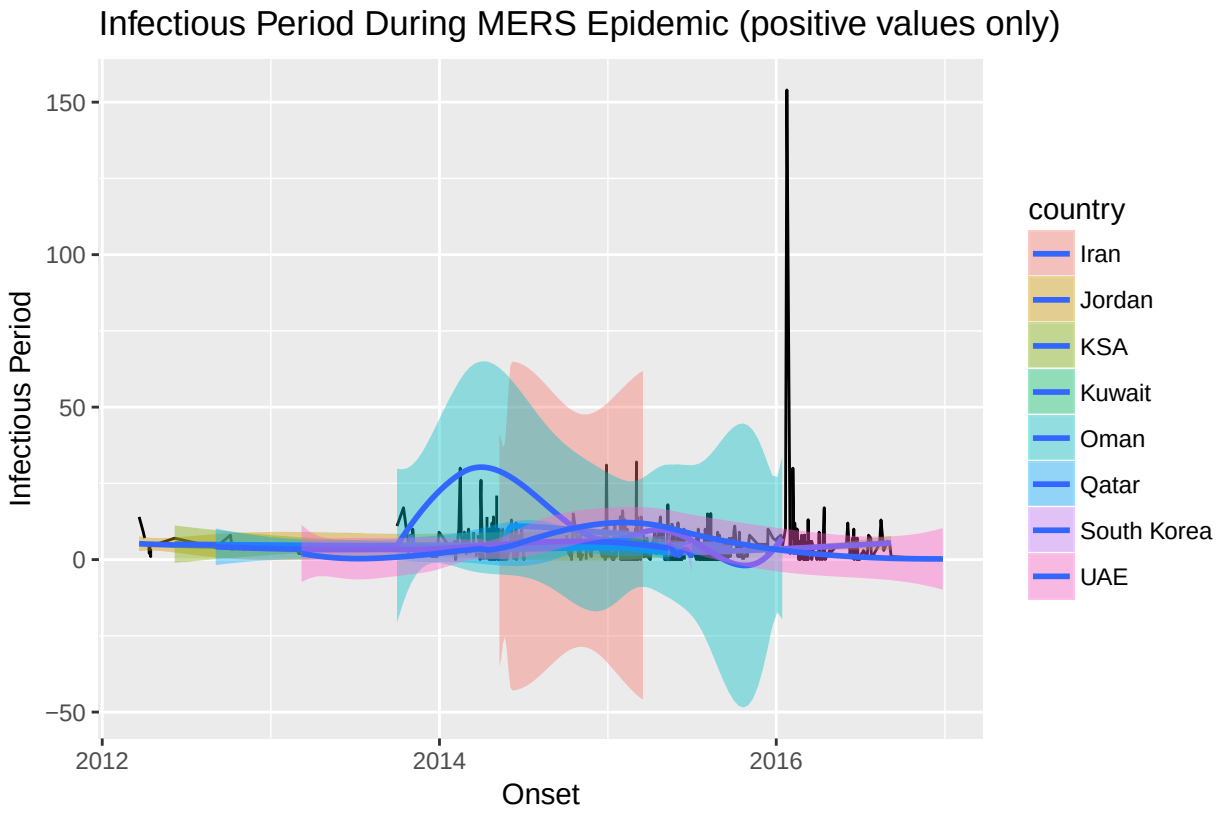
```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 97691

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 16199

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 27.555

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 97691

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 16005

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 11.39

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.5345e+05

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 16005

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 11.39

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
```

```
## condition number 0
```

```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.5345e+05
```

```
## Warning: Removed 535 rows containing missing values (geom_path).
```



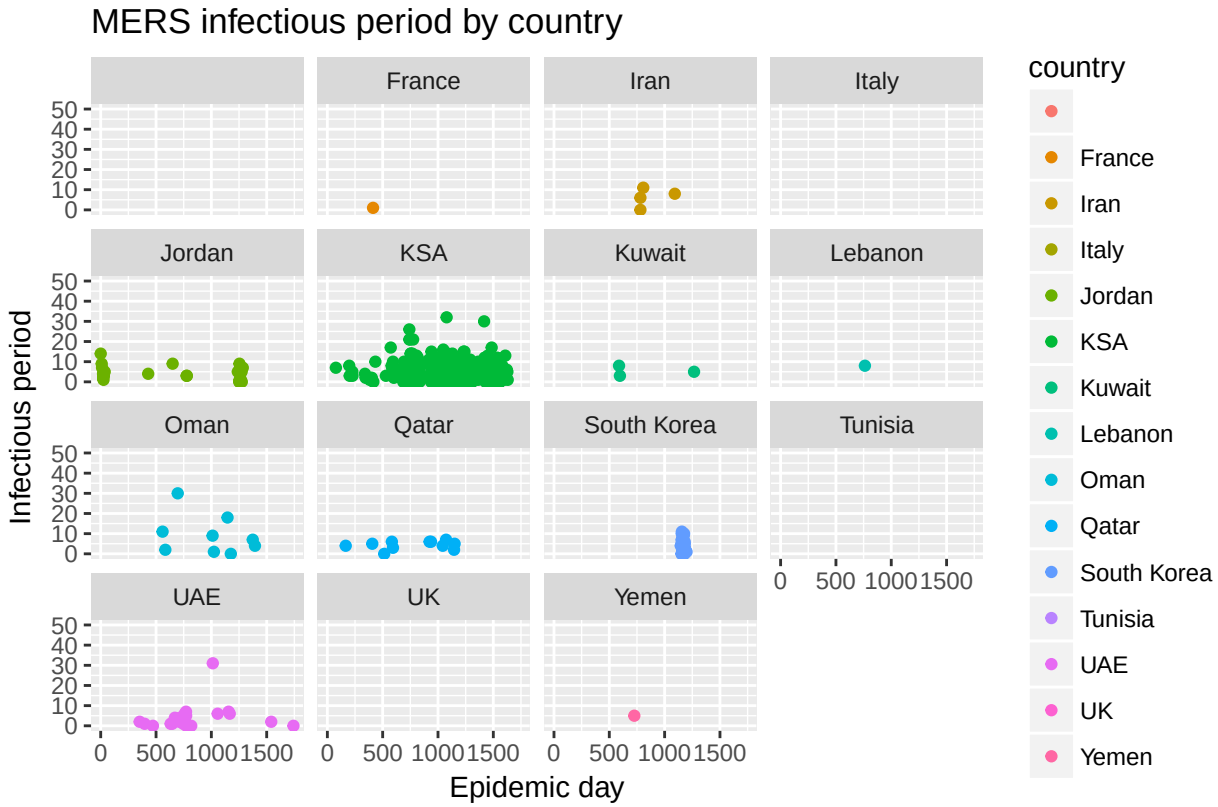
Data from: https://github.com/rambaut/MERS−Cases/blob/gh−pages/data/cases.

Faceting

Country

```
## FACETING ##
# By Country
ggplot(data = mers,
       mapping = aes(x=epi.day, y=infectious.period2)) +
  geom_point(mapping = aes(color=country)) +
  facet_wrap(~ country) +
  scale_y_continuous(limits = c(0, 50)) +
  labs(x='Epidemic day',
       y = 'Infectious period',
       title='MERS infectious period by country',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 728 rows containing missing values (geom_point).
```

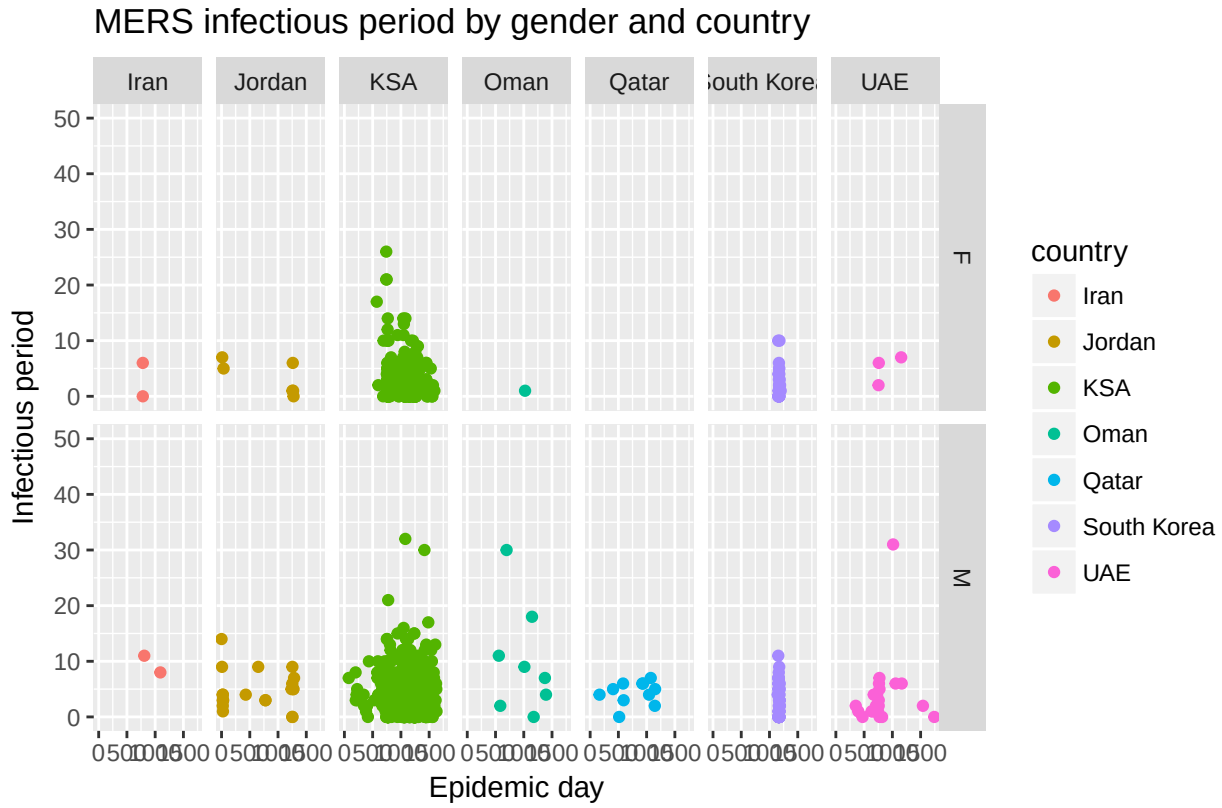


Data from: https://github.com/rambaut/MERS−Cases/blob/gh−pages/data/cases.

Country & Gender

```
# By Country & Gender
ggplot(data = subset(mers, gender %in% c('M', 'F') & country %in% c('KSA', 'Oman', 'Iran', 'Jordan', 'Q
       mapping = aes(x=epi.day, y=infectious.period2 )) +
  geom_point(mapping = aes(color=country)) +
  facet_grid(gender ~ country) +
  scale_y_continuous(limits = c(0, 50)) +
  labs(x = 'Epidemic day',
       y = 'Infectious period',
       title = 'MERS infectious period by gender and country',
       caption = "Data from: https://github.com/rambaut/MERS-Cases/blob/gh-pages/data/cases.")
```

```
## Warning: Removed 692 rows containing missing values (geom_point).
```



Data from: https://github.com/rambaut/MERS−Cases/blob/gh−pages/data/cases.

Case Fatality Rate

Work in progress, code does not work.

Interactive

For HTML output only.