Lab 12 - Statistics, Coordinates, Facets, and Themes

Ignacio Cabezudo

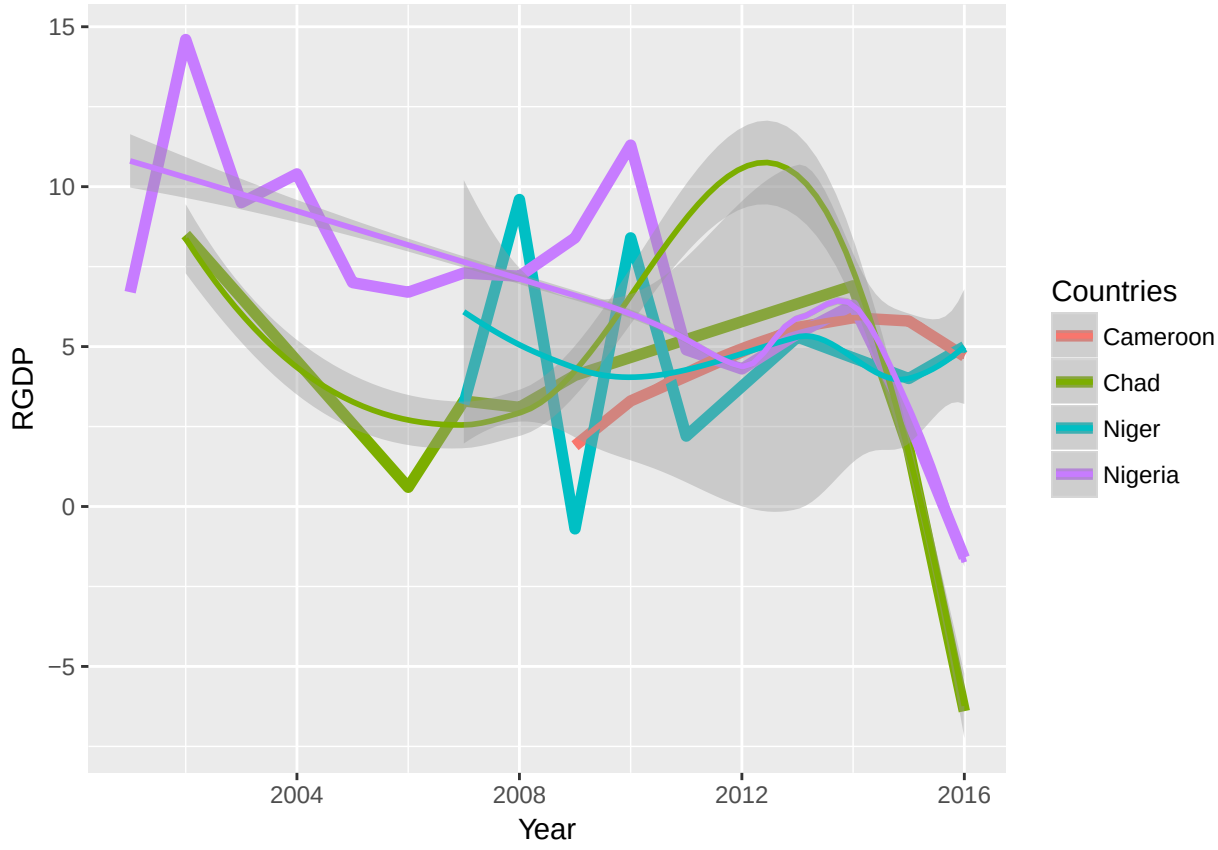
November 14, 2017

Complete the following exercises below. Knit together the PDF document and commit both the Lab 12 RMDfile and the PDF document to Git. Push the changes to GitHub so both documents are visible in your public GitHub repository.

1. Choose one or more graphics you created for Lab 11 and either experiment with the underlying statistical layer if it already has one (i.e. if you made a histogram experiement with different bin widths) or add a separate statistical layer to your plot (i.e. a smooothing curve). Choose something you think will offer meaningful insight and describe why you made the choice you did. What additional information does this provide viewers of your graphic?

```
library(ggplot2)
library(tidyverse)
library(hexbin)
library(readxl)

gtd_filter <- c("country_txt", "nkill", "iyear", "imonth", "nwound")
gtd_full <- read_csv("GTD FULL DB.csv")
IMF_Real_GDP_Growth <- read_excel("IMF Real GDP Growth.xls")


#Set the ranges for the data using filter
plotTerror <- filter(gtd_full, country_txt == c("Nigeria", "Niger", "Chad", "Cameroon") & iyear > 2000)
plotTerror <- select(plotTerror,gtd_filter)

#Fix the GDP growth data
temp <- gather(IMF_Real_GDP_Growth, key = Year, value = RGDP, ...=-'X__1')

#Rename columns to setup keys for join
names(temp)[names(temp) == 'X__1'] <- 'Countries'
names(plotTerror)[names(plotTerror) == 'country_txt'] <- 'Countries'
names(plotTerror)[names(plotTerror) == 'iyear'] <- "Year"

#Join the two datasets

joinT <- merge(plotTerror, temp, by = c("Countries", "Year"))

#Make some plots

ggplot(joinT, aes(x = Year, y = log(nkill), col = Countries )) +
  geom_jitter()+
  geom_smooth()
```

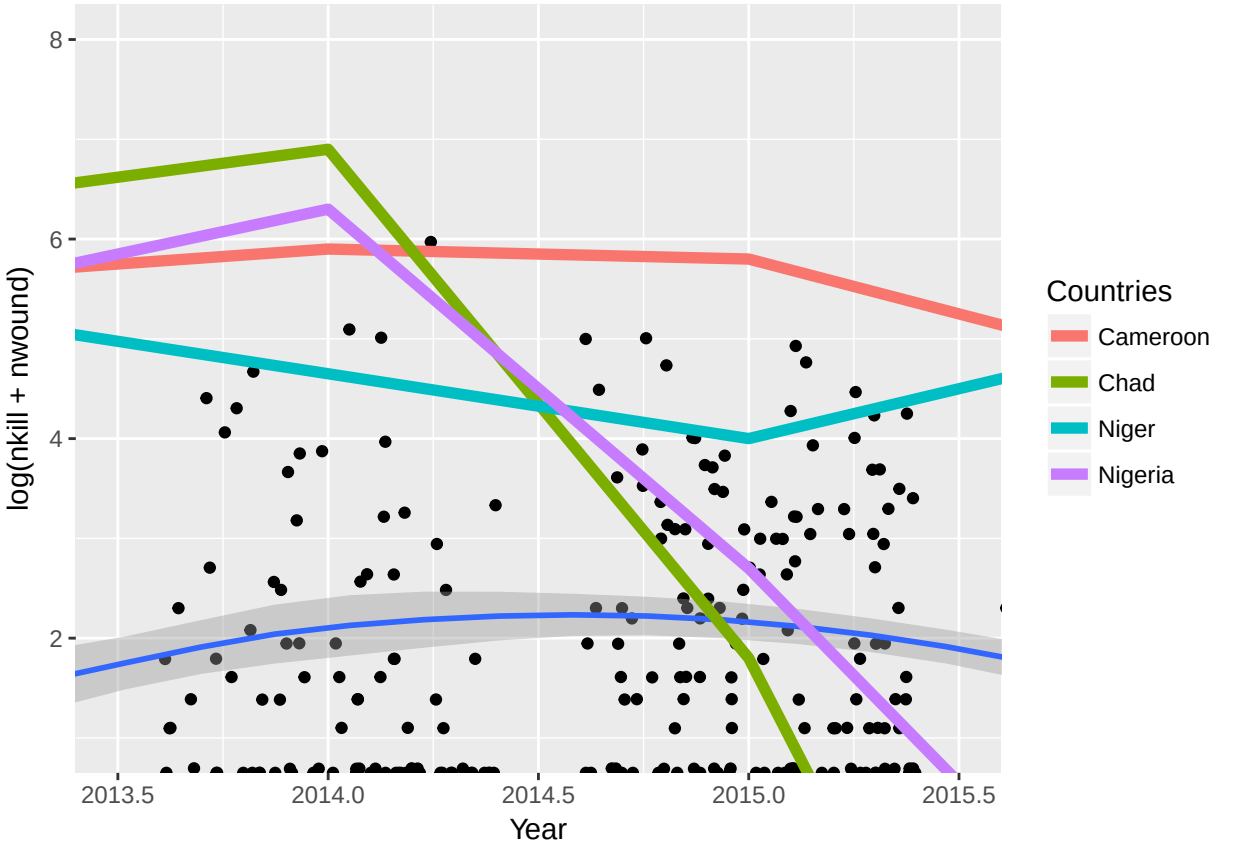

```
ggplot(joinT, aes(y = RGDP, x = Year, col = Countries)) +
  geom_line(size = 2)+
  geom_smooth()
```



I'm honestly at a loss with what to do with the data as far as statistical manipulation goes.

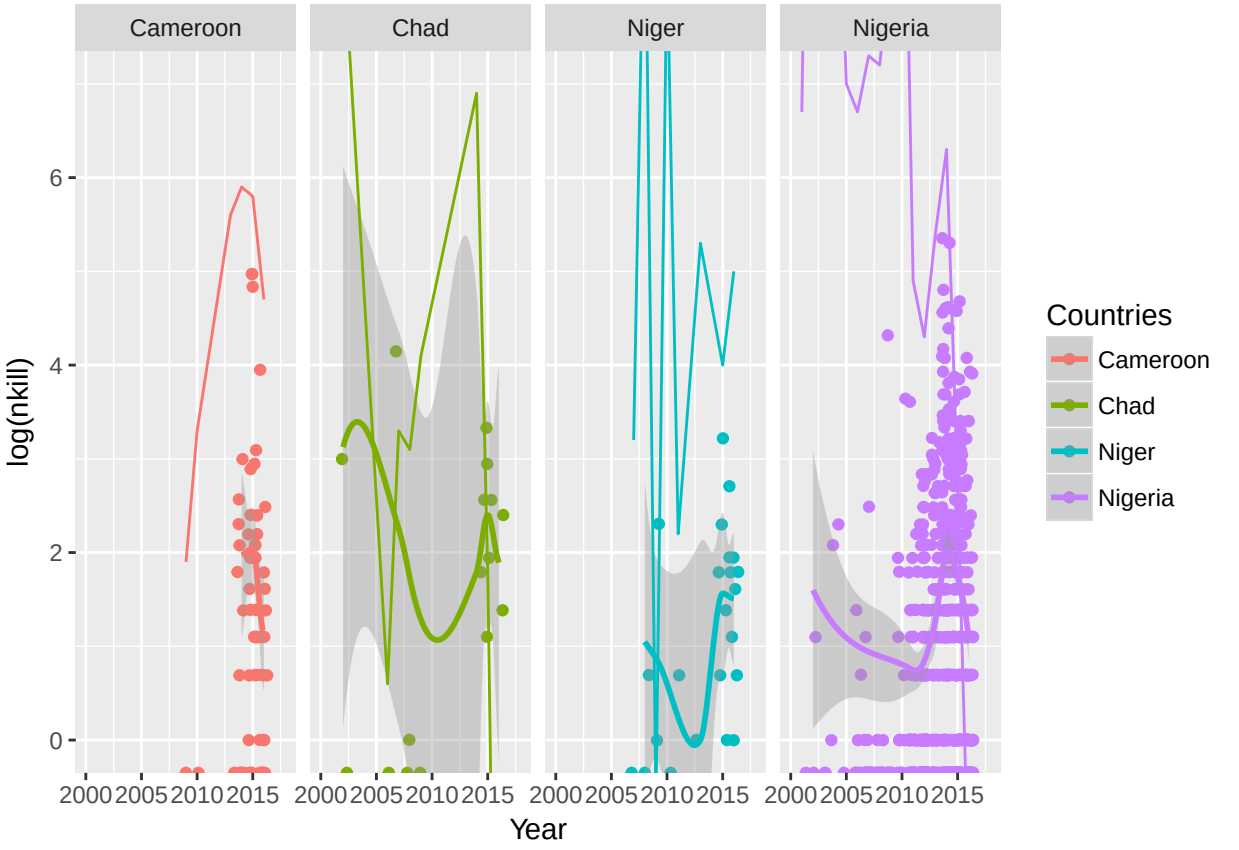
2. With the same or a different plot created in Lab 11, experiment with zooming in on specific areas of your graphic and changing the aspect ratio. Are their any benefits/drawbacks with either or both of these approaches for the visualizations you've created? What are they?

```
#Zooming in on most active portion of data


ggplot(joinT, aes(x = Year, y = log(nkill + nwound))) +
  geom_jitter() +
  coord_cartesian(xlim = c(2013.5, 2015.5), ylim = c(1, 8)) +
  geom_smooth() +
  geom_line(size = 2, data = joinT, aes( x = Year, y = RGDP, col = Countries))
```



3. Try facetting a plot you have made by another categorical variable in your data (this can even be as simple as Male/Female). What is the difference between `facet_wrap()` and `facet_grid()`? How might facetting be useful in data visualization?

```
ggplot(joinT, aes(x = Year, y = log(nkill), col = Countries)) +
  geom_jitter() +
  coord_cartesian(xlim = c(2000, 2018), ylim = c(0, 7)) +
  geom_smooth() +
  geom_line(data = joinT, aes( x = Year, y = RGDP))+
  facet_grid(. ~ Countries, scales = "free", space = "free")
```



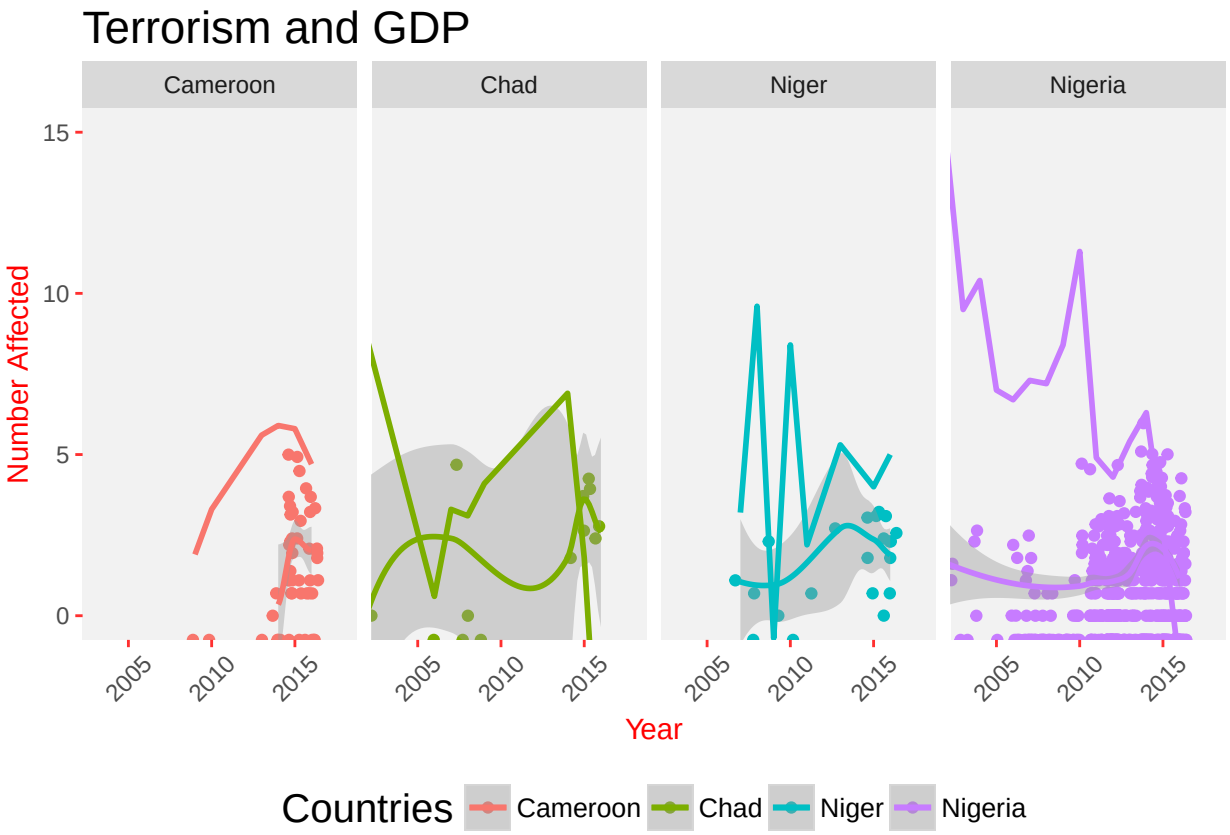
Faceting seems perfect for me to show a visual comparrison between the RGDP of each country and the number killed.

4. Use the `theme()` layer to change the appearance of a plot of your choice including the

- plot, axes, and legend titles
- axes tick marks
- text size
- legend position

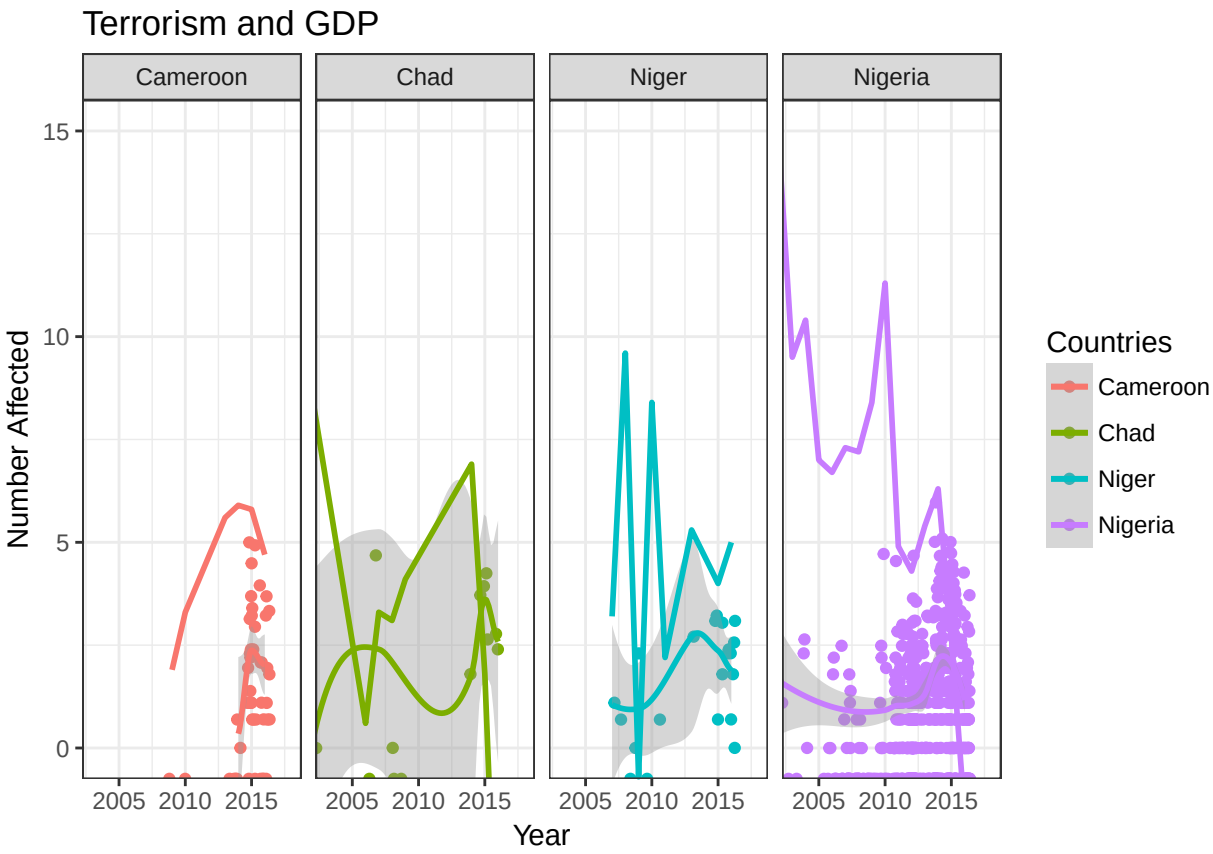
```
ggplot(joinT, aes(x = Year, y = log(nkill + nwound), col = Countries)) +
  geom_jitter() +
  coord_cartesian(xlim = c(2003, 2018), ylim = c(0, 15)) +
  geom_smooth() +
  geom_line(size = 1, data = joinT, aes( x = Year, y = RGDP))+
  facet_grid(. ~ Countries, scales = "free", space = "free") +
  labs( x = "Year",
        y = "Number Affected") +
  ggtitle("Terrorism and GDP") +
  theme(
    legend.position = "bottom",
    legend.title = element_text(size = 15),
    legend.text = element_text(size = 10),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.background = element_rect(fill = 'grey95'),
    axis.text.x = element_text(angle = 45, vjust=0.5),
    axis.title.x = element_text(color="red", vjust=-0.5),
```

```
    axis.title.y = element_text(color="red" , vjust=0.5),
    axis.ticks = element_line(colour = "firebrick1"),
    plot.title = element_text(size = rel(1.5)))
```

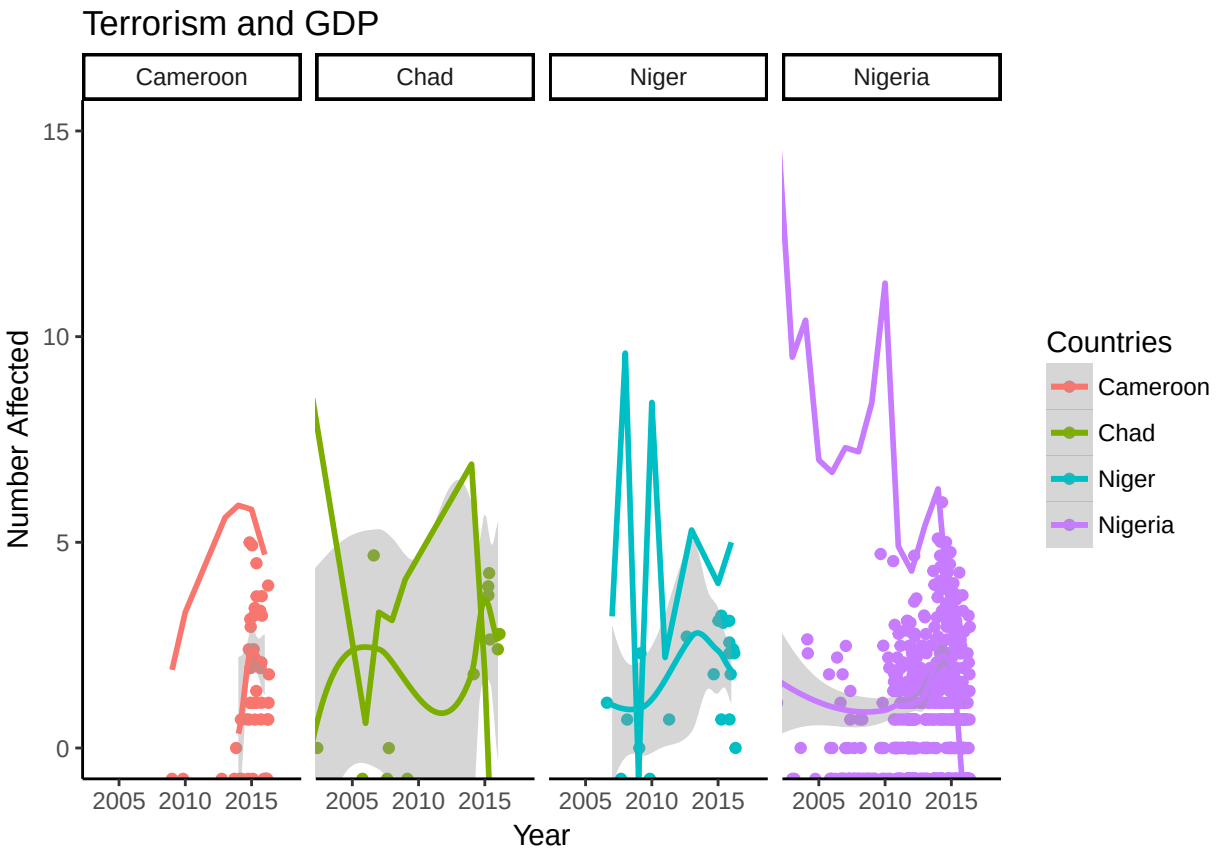


5. Create three versions of a graphic of your choice using different built-in themes or a theme created from `ggthemes`. Which ones do you think are best for presenting in an academic journal? A poster session? What are the qualities of the themes that you choice that you think make them more appropriate for presentation?

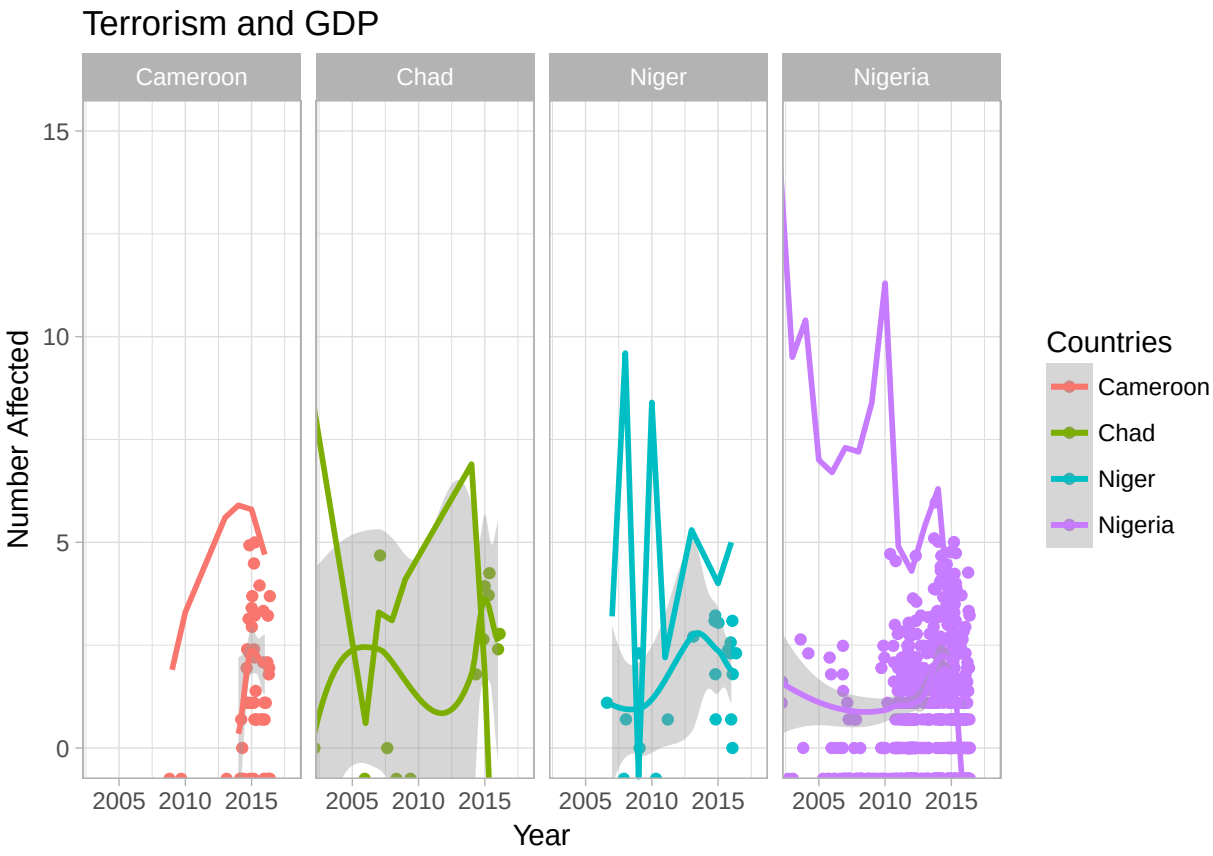
```
ggplot(joinT, aes(x = Year, y = log(nkill + nwound), col = Countries)) +
  geom_jitter() +
  coord_cartesian(xlim = c(2003, 2018), ylim = c(0, 15)) +
  geom_smooth() +
  geom_line(size = 1, data = joinT, aes( x = Year, y = RGDP))+
  facet_grid(. ~ Countries, scales = "free", space = "free") +
  labs( x = "Year",
        y = "Number Affected") +
  ggtitle("Terrorism and GDP") +
  theme_bw()
```




```
ggplot(joinT, aes(x = Year, y = log(nkill + nwound), col = Countries)) +
  geom_jitter() +
  coord_cartesian(xlim = c(2003, 2018), ylim = c(0, 15)) +
  geom_smooth() +
  geom_line(size = 1, data = joinT, aes( x = Year, y = RGDP))+
  facet_grid(. ~ Countries, scales = "free", space = "free") +
  labs( x = "Year",
        y = "Number Affected") +
  ggtitle("Terrorism and GDP")+
  theme_classic()
```




```
ggplot(joinT, aes(x = Year, y = log(nkill + nwound), col = Countries)) +
  geom_jitter() +
  coord_cartesian(xlim = c(2003, 2018), ylim = c(0, 15)) +
  geom_smooth() +
  geom_line(size = 1, data = joinT, aes( x = Year, y = RGDP))+
  facet_grid(. ~ Countries, scales = "free", space = "free") +
  labs( x = "Year",
        y = "Number Affected") +
  ggtitle("Terrorism and GDP")+
  theme_light()
```


Terrorism and GDP
Cameroon
Chad
Niger
Nigeria
Number Affected
0
5
10
15
2005
2010
2015
Year
Countries
Cameroon
Chad
Niger
Nigeria