1. ggplot

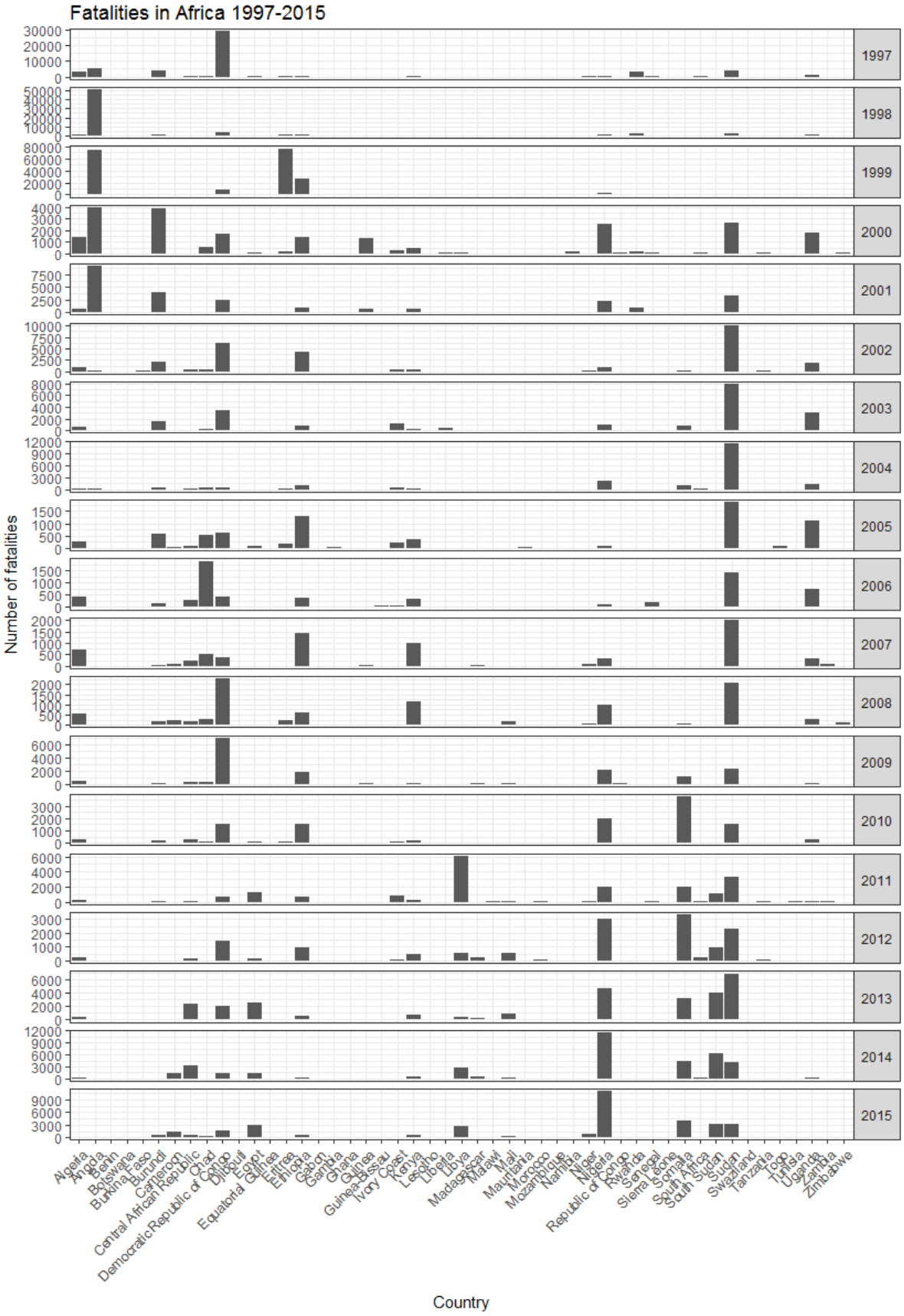
a. **Create a faceted plot that shows the temporal distribution (e.g. bar chart) of the number of fatalities (not events) per country.**

There were struggles while creating the bar chart that counts the fatalities, not event count. After some research, I found that there are two types of bar charts: geom_bar, which makes the height of the bar proportional to the number of cases in each group (it uses stat_count by default), and geom_col, which let the heights of the bars to represent values in the data.

Here's the code for geom_col:

```
ggplot(acled, aes(x=COUNTRY,y=FATALITIES)) + geom_col()+    facet_grid(YEAR ~ ., scales = "free_y") + theme_bw() + theme(strip.text.y = element_text(angle = 0), axis.text.x = element_text(angle = 45, hjust = 1)) + xlab("Country") + ylab("Number of fatalities") + ggtitle("Fatalities in Africa 1997-2015")
```

Result as follows:


Fatalities in Africa 1997-2015
Number of fatalities
Country
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Algeria
Angola
Benin
Botswana
Burkina Faso
Burundi
Cameroon
Central African Republic
Democratic Republic of Congo
Chad
Djibouti
Egypt
Equatorial Guinea
Eritrea
Ethiopia
Gabon
Gambia
Ghana
Guinea
Guinea-Bissau
Ivory Coast
Kenya
Lesotho
Liberia
Libya
Madagascar
Malawi
Mali
Mauritania
Morocco
Mozambique
Namibia
Niger
Nigeria
Republic of Congo
Rwanda
Senegal
Sierra Leone
Somalia
South Africa
South Sudan
Sudan
Swaziland
Tanzania
Togo
Tunisia
Uganda
Zambia
Zimbabwe

It may be a good idea to add colors to the plot (colored by event type):

Here’s the code for the coloured version:

```
ggplot(acled, aes(x=COUNTRY,y=FATALITIES,fill = factor(EVENT_TYPE))) + geom_col()+
facet_grid(YEAR ~ ., scales = "free_y") + theme_bw() + theme(strip.text.y = element_text(angle =
0), axis.text.x = element_text(angle = 45, hjust = 1)) + xlab("Country") + ylab("Number of
fatalities") + ggtitle("Fatalities in Africa 1997-2015")
```

Coloured result as follows:


Fatalities in Africa 1997-2015
Number of fatalities
Country
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
factor(EVENT_TYPE)
Battle-Government regains territory
Battle-No change of territory
Battle-Non-state actor overtakes territory
Headquarters or base established
Non-violent transfer of territory
Remote violence
Riots/Protests
Strategic development
Violence against civilians

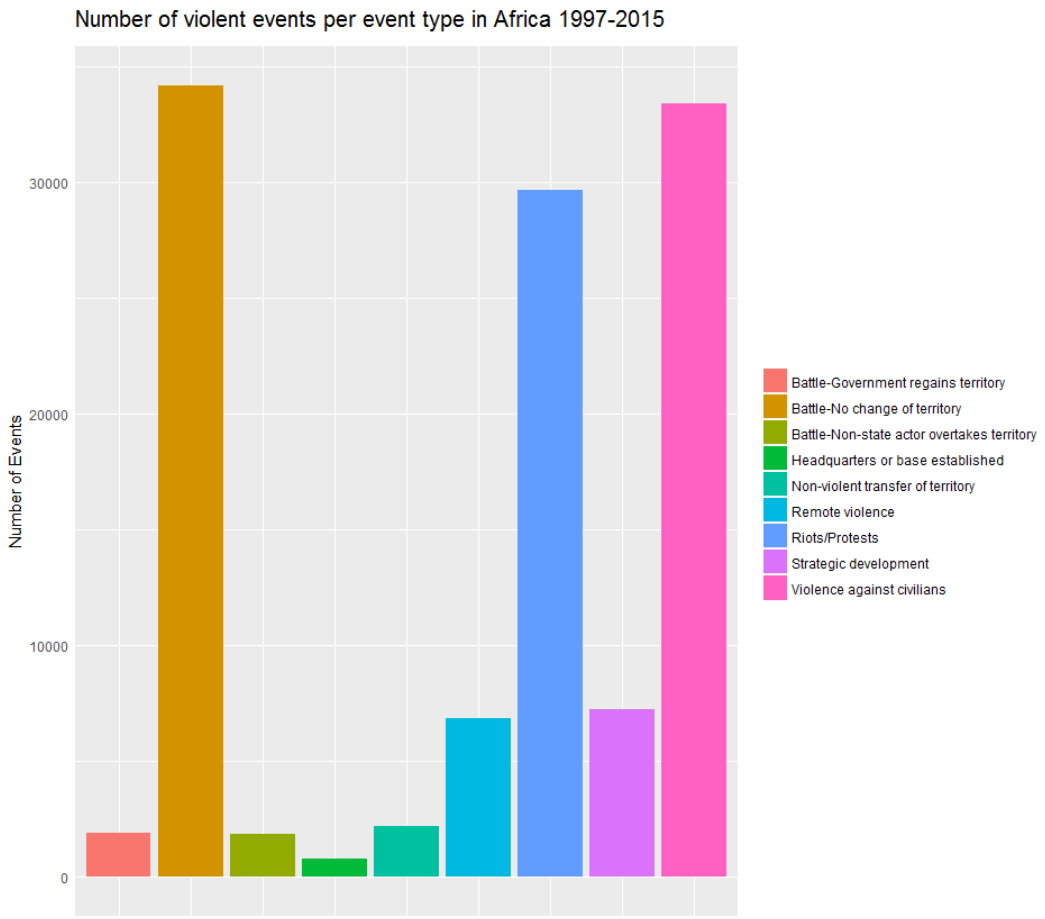
b. Discuss the advantages and disadvantages of such an approach (faceted plot) vis-à-vis a set of small multiple maps where each map represents a single year.

	Faceted Plot	Set of Maps
Pros	• Convenient for side-by-side comparison, as we could align all facets by axis • Easier to generate then generating multiple plots • More organized plot arrangements	• Provide more freedom for creative plot arrangements • Do not need to deal with complex syntax for plotting faceted plot
Cons	• The size of individual plot is compromised especially when many individual plots are created • There are two scaling options for the y-axis, but neither is ideal: when the data range is diverse, standardized y-axis will make the data difference hard to detect, but non-standardized y-axis may results in biased comparison (when axis-scaling factor is overlooked)	• It requires writing multiple code, a largely repetitive task with minor difference, which is prone to human error • It may be difficult to resize and align the individual plot efficiently

c. Create a bar chart, pie chart, coxcomb and bullseye chart of the number of events per event type. Discuss the similarities and differences between each chart type in a short paragraph.

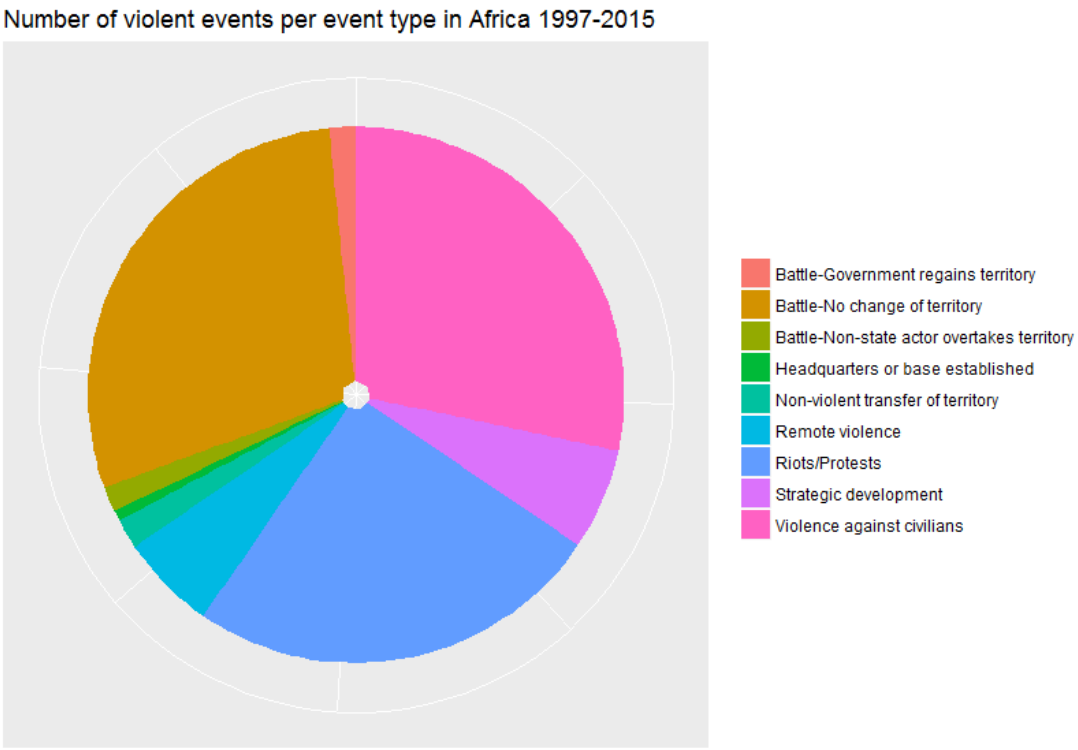
Bar Chart

ggplot(acled, aes(EVENT_TYPE, fill = factor(EVENT_TYPE))) + geom_bar() + xlab("") + ylab("Number of Events") + ggtitle("Number of violent events per event type in Africa 1997-2015") + theme(strip.text.y = element_text(angle = 0), axis.text.x = element_blank(),axis.ticks = element_blank()) + guides(fill=guide_legend(title=NULL)) + theme(plot.margin = unit(c(1,1,1,1), "cm"))+theme(plot.title=element_text(size=15, vjust=3))



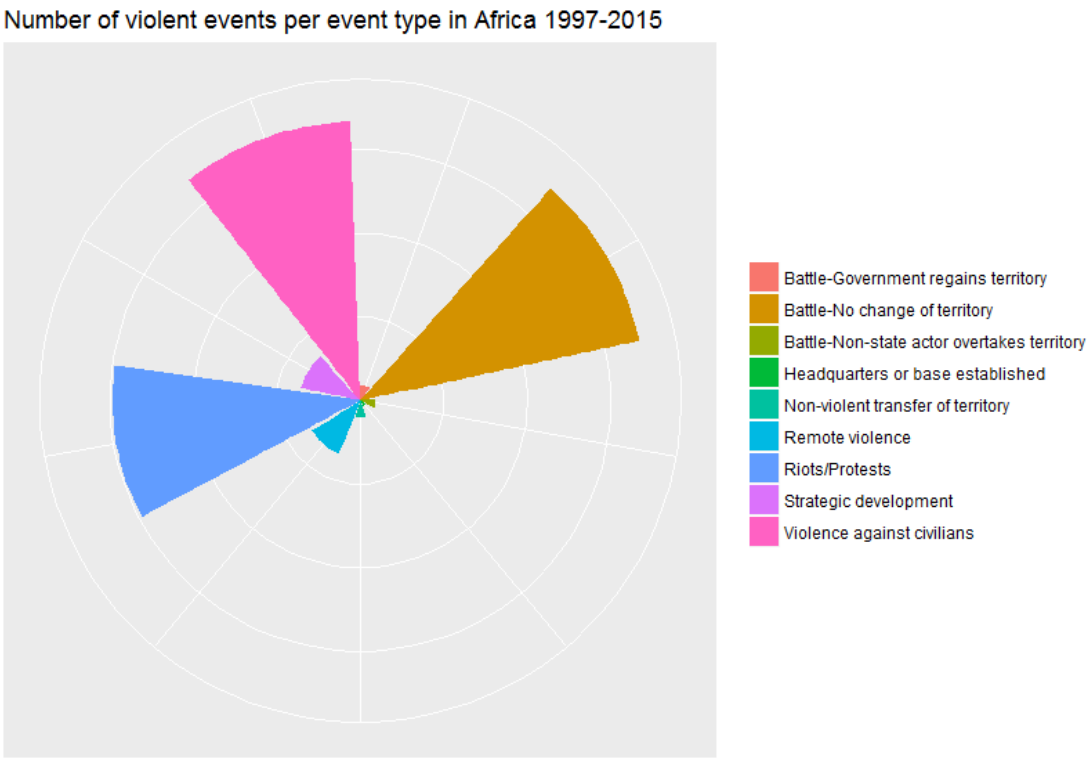
Pie Chart

```
ggplot(acled, aes(x = factor(1), fill=factor(EVENT_TYPE))) + geom_bar()+ coord_polar("y") +
ggtitle("Number of violent events per event type in Africa 1997-2015") + theme(axis.text.x =
element_blank(),axis.text.y = element_blank(),axis.ticks = element_blank()) + xlab("") + ylab("")
+ guides(fill=guide_legend(title=NULL))
```



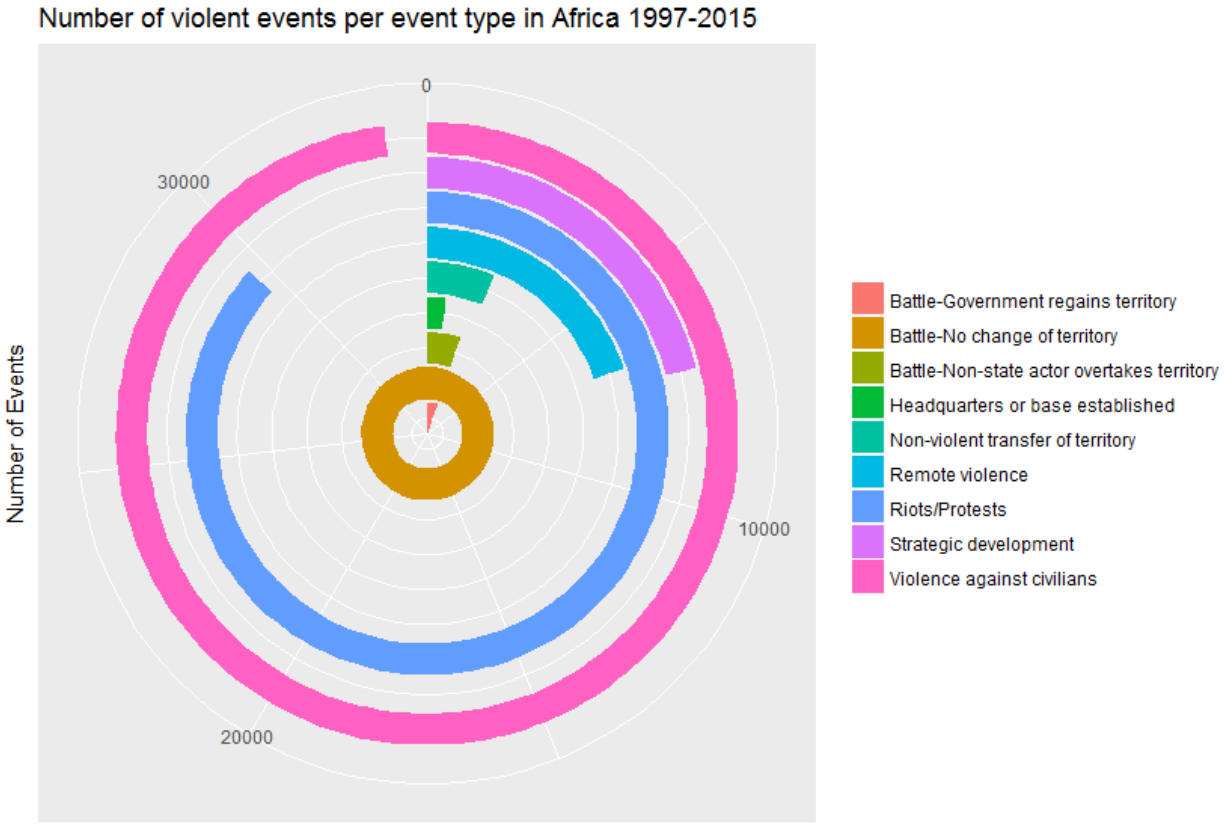
Coxcomb Chart

```
ggplot(acled, aes(x = factor(EVENT_TYPE), fill=factor(EVENT_TYPE))) + geom_bar()+
coord_polar() + ggtitle("Number of violent events per event type in Africa 1997-2015") +
theme(axis.text.x = element_blank(),axis.text.y = element_blank(),axis.ticks = element_blank()) +
xlab("") + ylab("") + guides(fill=guide_legend(title=NULL))
```



Bullseye Chart

```
ggplot(acled, aes(x = factor(EVENT_TYPE), fill=factor(EVENT_TYPE))) + geom_bar()+
coord_polar("y") + ggtitle("Number of violent events per event type in Africa 1997-2015") +
theme(axis.text.y = element_blank(),axis.ticks = element_blank()) + xlab("Number of Events") +
ylab("") + guides(fill=guide_legend(title=NULL))
```



Similarities

All chart types shows proportional relationship between data size and the area of chart element, be it the bar or the pie. Also, they all allow direct comparison between each data set in the same scale and coordinate system.

Differences

Different charts are designed for different purposes, that's why each put emphasis in a unique area of the data, and each may be biased in a special way. For example, a pie chart shows relative ratios better than a bar chart, but a bar chart allows reading off of data values easier.

2. plotly

Publish your two interactive charts to rpubs.com. Again, make sure that your charts have appropriate labels, titles, scales and colors. In your homework, include a short paragraph on why you chose the two graphs types and, of course, include a link to both graphs.

I chose two unconventional graph types: the 2D histogram and the 3D scatter plot, as it is more exciting to create graphs that you cannot easy reproduce in by grammar of graphics plot. Also, these two graph types allow me to present three data fields without using faceted plots.

2D Histogram

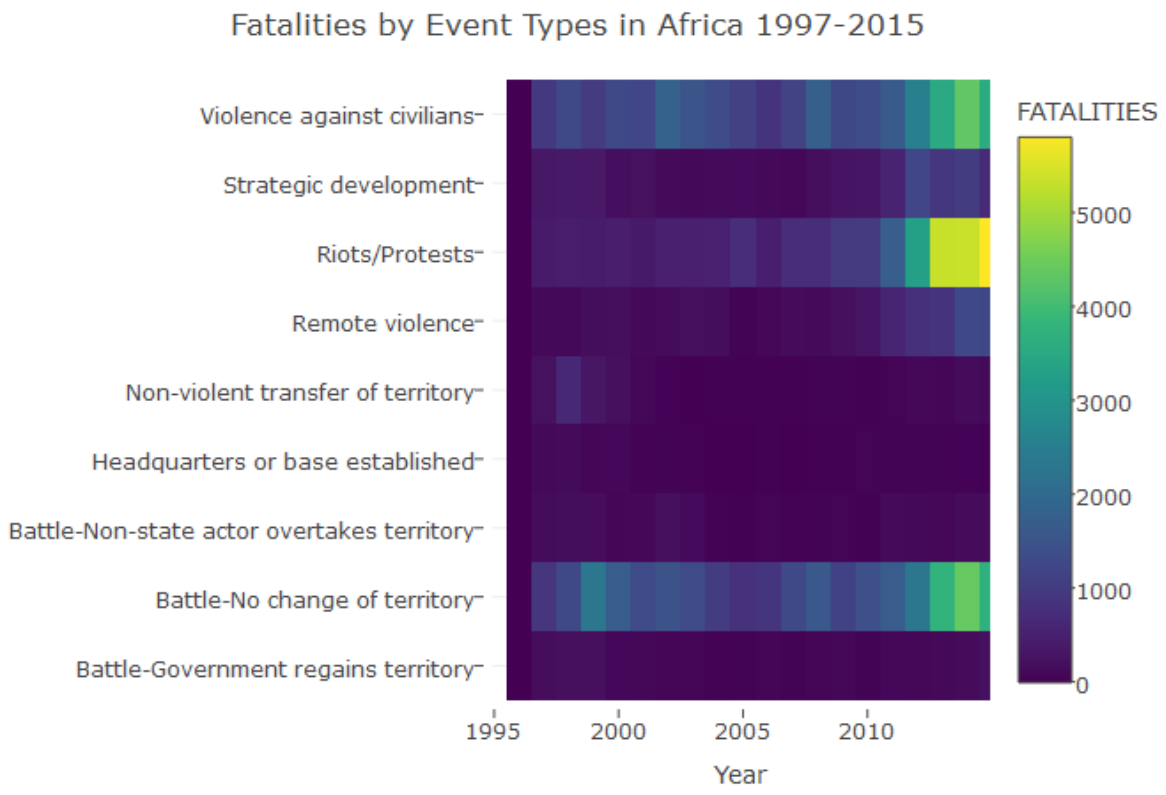
```
x <- list( title = "Year" )

y <- list( title = "Event Type" )

m <- list( l = 300,  r = 80,  b = 50,  t = 50,  pad = 4)

p <- plot_ly(acled, x = ~YEAR, y = ~EVENT_TYPE, z = ~FATALITIES) %>% layout(title = "Fatalities by Event Types in Africa 1997-2015", xaxis = x, yaxis = y, autosize = F, margin = m)

p %>% add_histogram2d()
```

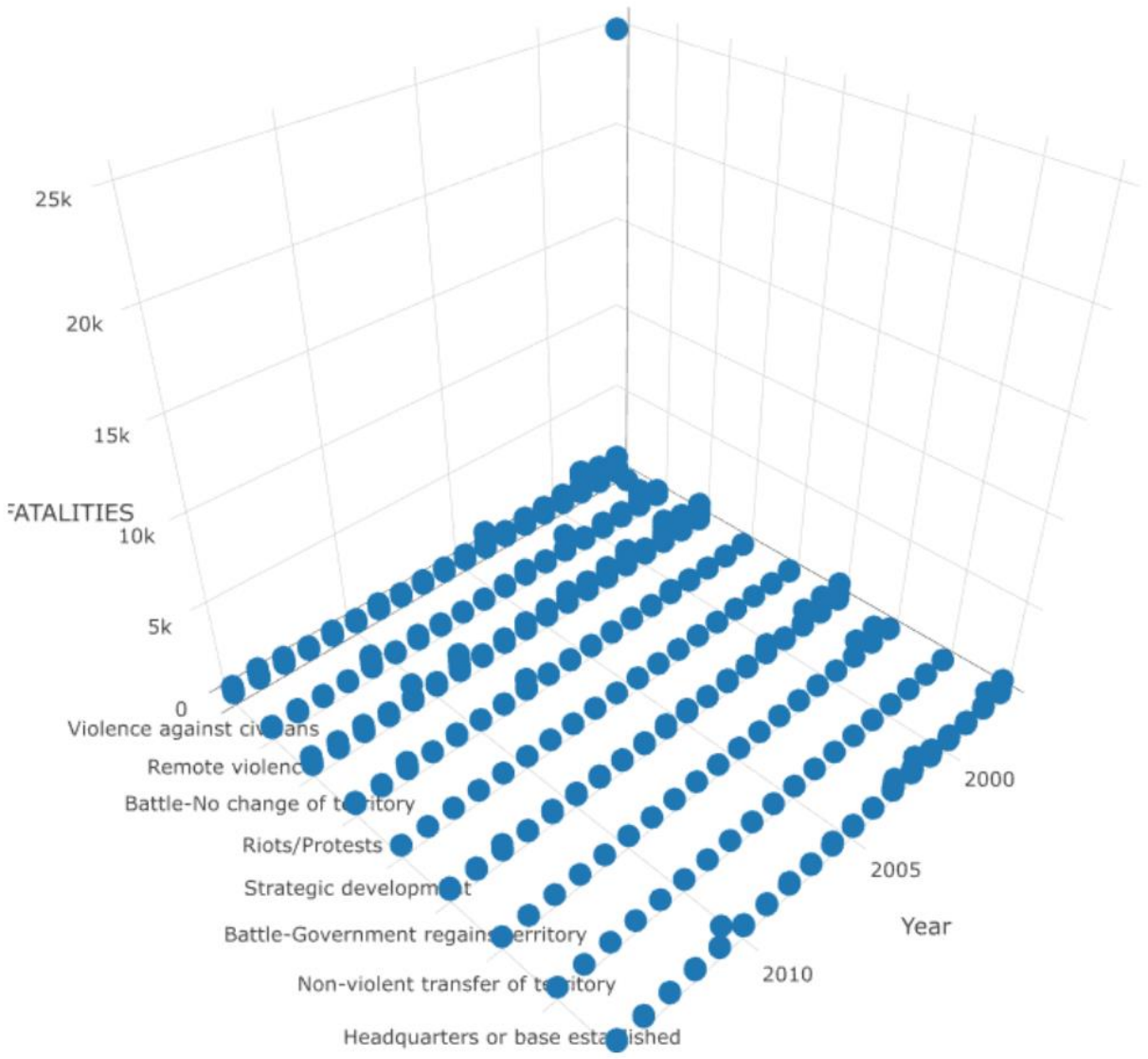


Link: http://rpubs.com/zhexxian/263289

3D Scatter Plot

```
p <- plot_ly(acled, x = ~ YEAR, y = ~ EVENT_TYPE, z = ~ FATALITIES) %>%  layout(title = "Fatalities by Event Types in Africa 1997-2015", scene = list(xaxis = list(title = 'Year'),  yaxis = list(title = ''), autosize = F , margin = list( l = 50,  r = 50,  b = 50, t = 50, pad = 10)))

p %>%  add_markers()
```



Link: http://rpubs.com/zhexxian/263298