Visualization Lab5 Group2

Assignment 1

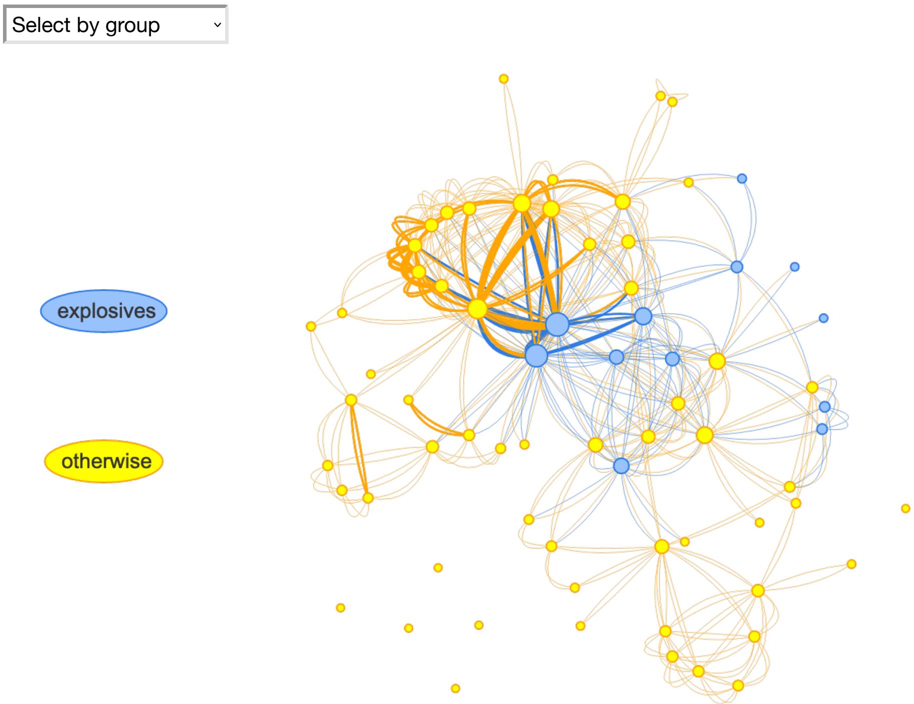
Task 1

```
train_data=read.table("trainData.dat")
train_meta=read.table("trainMeta.dat")

train_meta$id=1:70
train_meta=train_meta[,c(3,1,2)]
colnames(train_data)=c("from","to","connect")
#net <- graph_from_data_frame(d=train_data, vertices=train_meta$id, directed=T)
#visIgraph(net)
#visNetwork(train_meta, train_data)

#use strength of links variable
train_data$value=train_data$connect
train_meta$label=train_meta$V1
#nodes are colored by Bombing Group.
train_meta$group=train_meta$V2
train_meta$group[train_meta$group==1]=" explosives"
train_meta$group[train_meta$group==0]=" otherwise"
#size of nodes is proportional to the number of connections
data_value=data.frame(rle(sort(train_data$from))[2],rle(sort(train_data$from))[1])
train_meta$value=0
train_meta$value[data_value$values]=data_value$lengths

#use a layout that optimizes repulsion forces
#all nodes that are connected to a currently selected node by a path of length one are highlighted
visNetwork(nodes=train_meta, edges=train_data)%>%
  visPhysics(solver="repulsion")%>%
  visOptions(highlightNearest=list(enabled=TRUE,degree=1),
             selectedBy = "group")%>%
               visLegend()
```

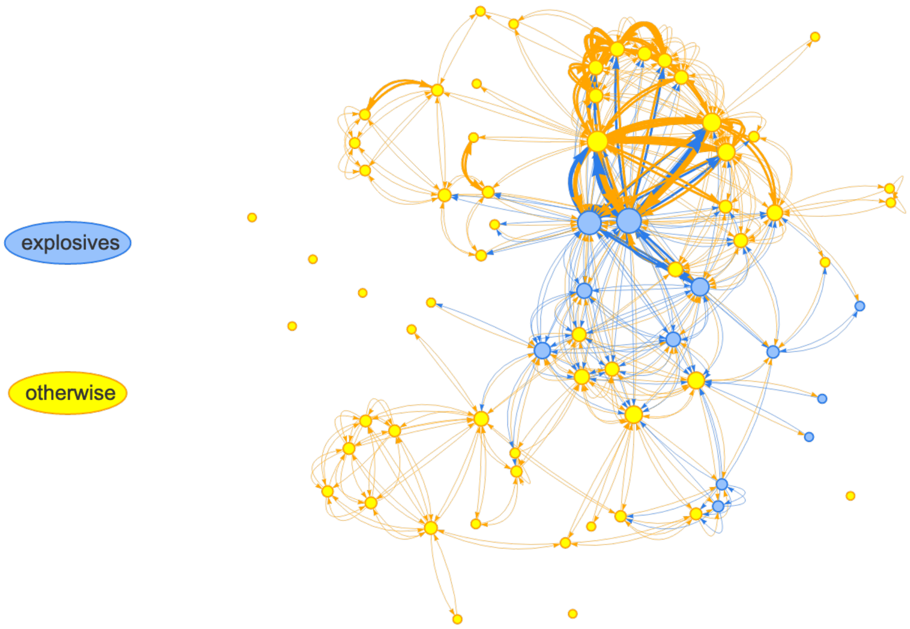


The blue nodes represent people who participated in placing the explosives,yellow represent otherwise.we could easily recognize the largest cluster,which is related to Jamal Zougam and Mohamed Chaoui,they are the people involved in bombing.We can also find other 4 small cluster:the cluster around Semaan Gaby Eid,Abderrahim Zbakh,Imad Eddin Barakat and Abdelkarim el Mejjati respectively.

Task 2

```
visNetwork(nodes=train_meta, edges=train_data)%>%
  visPhysics(solver="repulsion")%>%
  visEdges(arrows = "to")%>%
  visOptions(highlightNearest=list(enabled=TRUE,algorithm="hierarchical",degree=list(from=1,to=2)),
             selectedBy = 'group')%>%
  visLegend()
```

Select by group ⌄



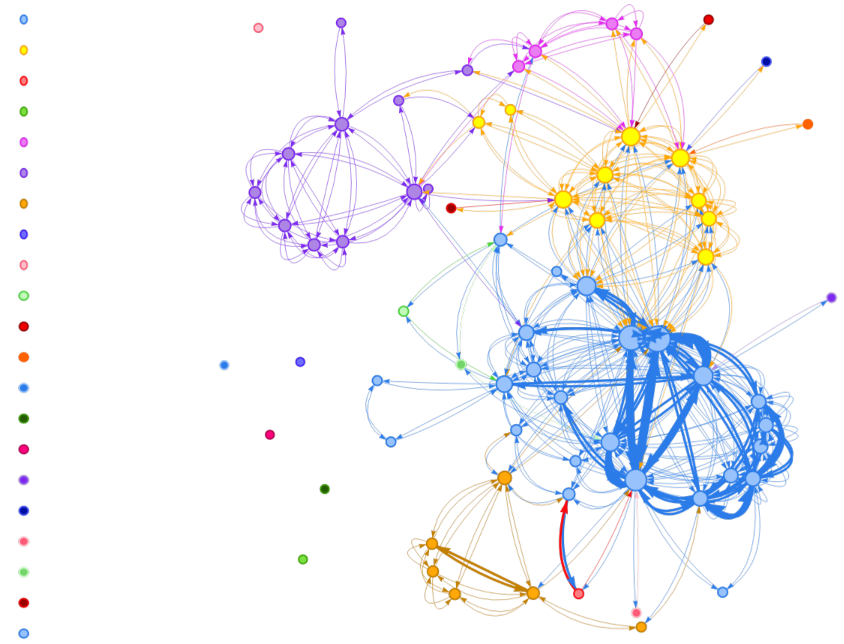
Jamal Zougam and Abdeluahid Berrak has the best opportunity to spread the information in the network.Jamal Zougam is believed to be the person who sold telephones which were used to detonate the bombs in the attack,maybe this why he had so many chance to spread the information in the network.

Task 3

```
#Community identification
nodes1<-train_meta
graph=graph_from_data_frame(d=train_data,vertices = train_meta,directed = F)
cluster=cluster_edge_betweenness(graph,directed = T)
nodes1$group=cluster$membership

visNetwork(nodes=nodes1, edges=train_data)%>%
  visPhysics(solver="repulsion")%>%
  visEdges(arrows = "to")%>%
  visOptions(highlightNearest=list(enabled=TRUE,degree=1),
             selectedBy = 'group')%>%
  visLegend()
```

Select by group ⌄



The largest cluster i.e. cluster relate to Jamal Zougam and Mohamed Chaoui could still be found,and the cluster around Semaan Gaby Eid,Imad Eddin Barakat and Abdelkarim el Mejjati respectively could be clearly recognized with the color purple,yellow and orange,but the cluster with Abderrahim Zbakh in it is divided into the cluster of Jamal Zougam.otherwise,some nodes which have no relationship with the other are divide into single clusters.

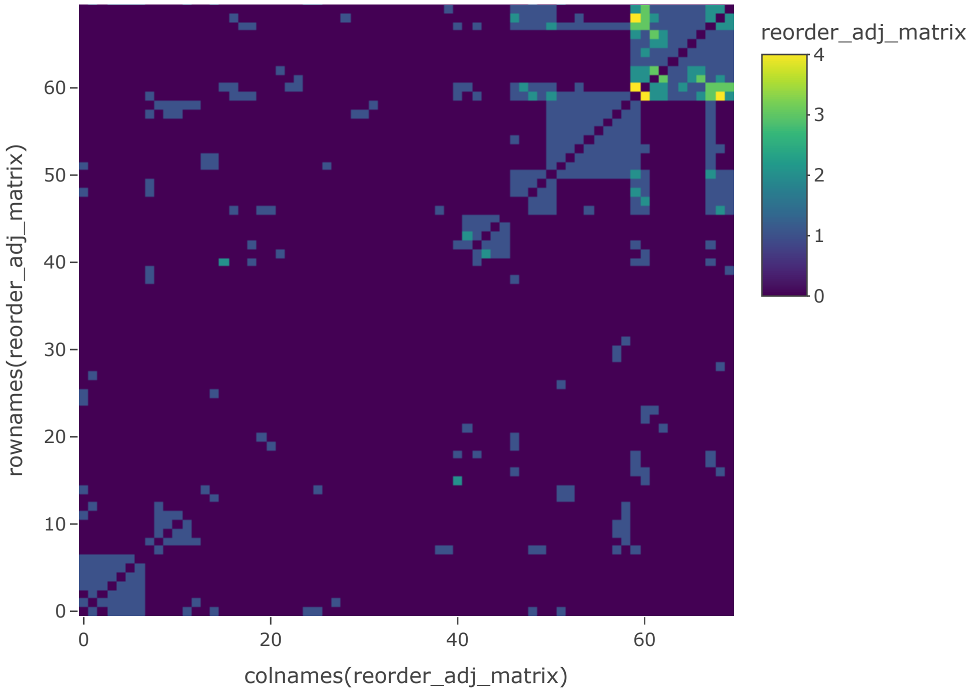
Task 4

```
#Adjacency representation
adj_matrix<-get.adjacency(graph,attr="value",sparse=F)
colnames(adj_matrix)<-V(graph)$media
rownames(adj_matrix)<-V(graph)$media

rowdist<-dist(adj_matrix)

order1<-seriate(rowdist,"HC")
ord1<-get_order(order1)

reorder_adj_matrix<-adj_matrix[ord1,ord1]
plot_ly(z=~reorder_adj_matrix, x=~colnames(reorder_adj_matrix),
        y=~rownames(reorder_adj_matrix), type="heatmap")
```



The most pronounced cluster is on the top right area,which relate to Jamal Zougam and Mohamed Chaoui.We have already discover this cluster in step 1 and 3.

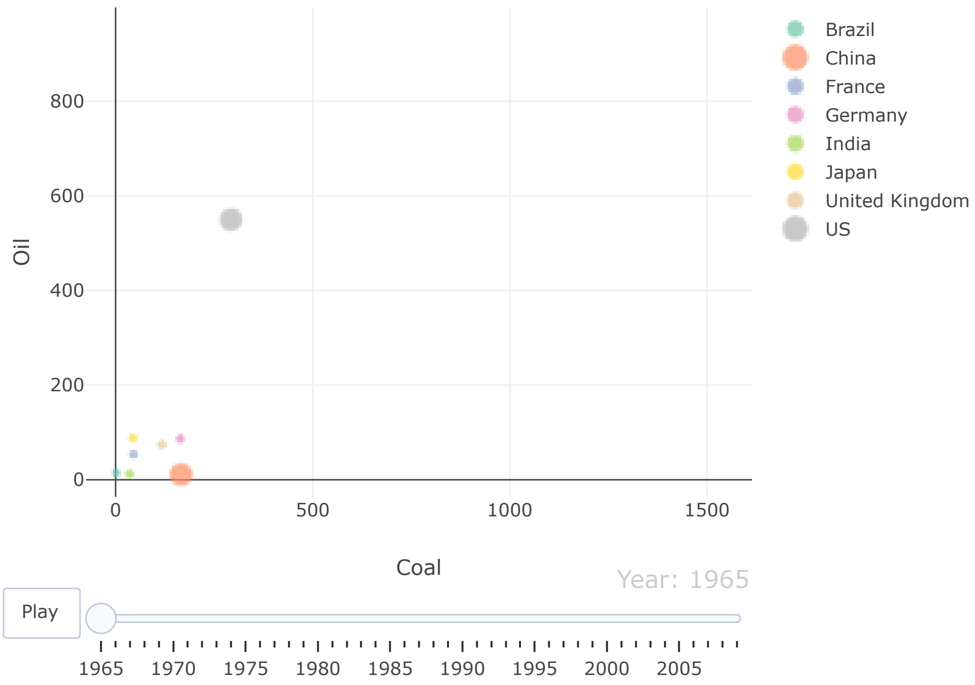
Assignment 2

1. Animated bubble chart

```
oilcoal <- read.csv2("Oilcoal.csv")

p2_1 <- oilcoal %>%
  plot_ly(
    x = ~Coal,
    y = ~Oil,
    size = ~Marker.size,
    color = ~Country,
    frame = ~Year,
    text = ~Country,
    hoverinfo = "text",
    type = 'scatter',
    mode = 'markers'
  )

p2_1
```

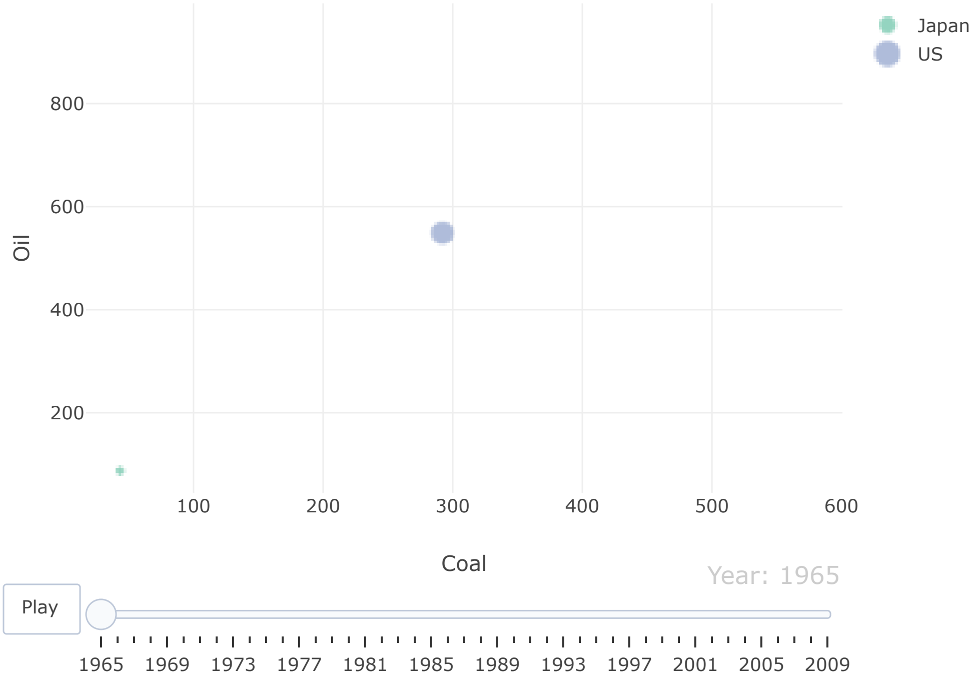


Q1: List several noteworthy features of the investigated animation.

- From 1965 to 1979, the oil consumption in US increased rapidly.
- There is a short decrease on oil consumption in most countries (except China) from 1973 and return to growth after three years. Since 1979, the oil consumption in most countries (except China) also decreased until 1983. And there is also a decrease in 1990 and last one year.
- Both oil and coal consumption descended in US during 2008 ~ 2009.
- Coal consumption increased rapidly in China after 2002.

2. Two countries that had similar motion patterns

```
p2_2 <- oilcoal %>%
  filter(Country == "US" | Country == "Japan") %>%
  plot_ly(
    x = ~Coal,
    y = ~Oil,
    size = ~Marker.size,
    color = ~Country,
    frame = ~Year,
    text = ~Country,
    hoverinfo = "text",
    type = 'scatter',
    mode = 'markers'
  )
```



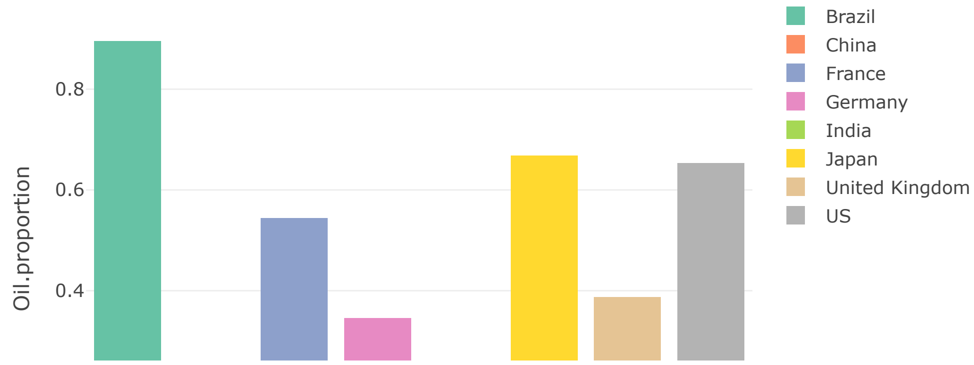
Q1: Try to find historical facts that could explain some of the sudden changes in the animation behavior.

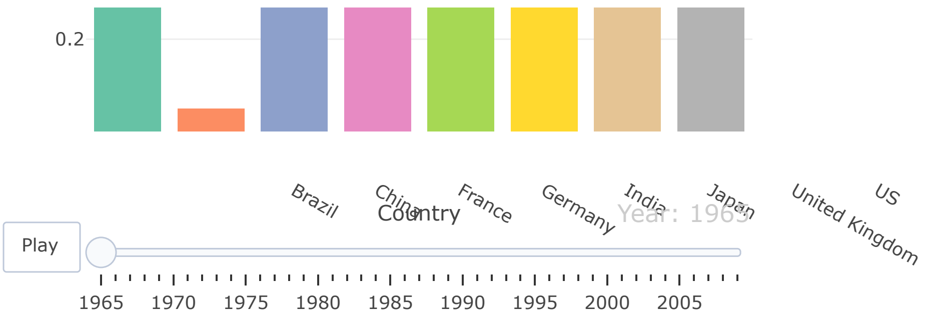
US and Japan seems to be the two countries had similar motion patterns. We can see they both had a rapidly increase in oil consumption from 1965 to 1973. For Japan, which had just hosted the Olympic Games in 1964 and was in a period of rapid economic growth, after World War II. During this period Japan as a bridgehead for the United States in Asia and US also gave Japan a lot of economic aid. So, maybe Japan and the United States are more economically linked. And as we mentioned in task 1, they both had three periods of decline in oil consumption. After searching, these three periods correspond to the three oil crises.

3. Bar chart of the proportion of oil consumption

```
oilcoal$Oil.proportion <- oilcoal$Oil / (oilcoal$Oil + oilcoal$Coal)

p2_3 <- oilcoal %>%
  plot_ly(x = ~Country, y = ~Oil.proportion, frame = ~Year, color = ~Country) %>%
  add_bars() %>%
  animation_opts(1000)
```





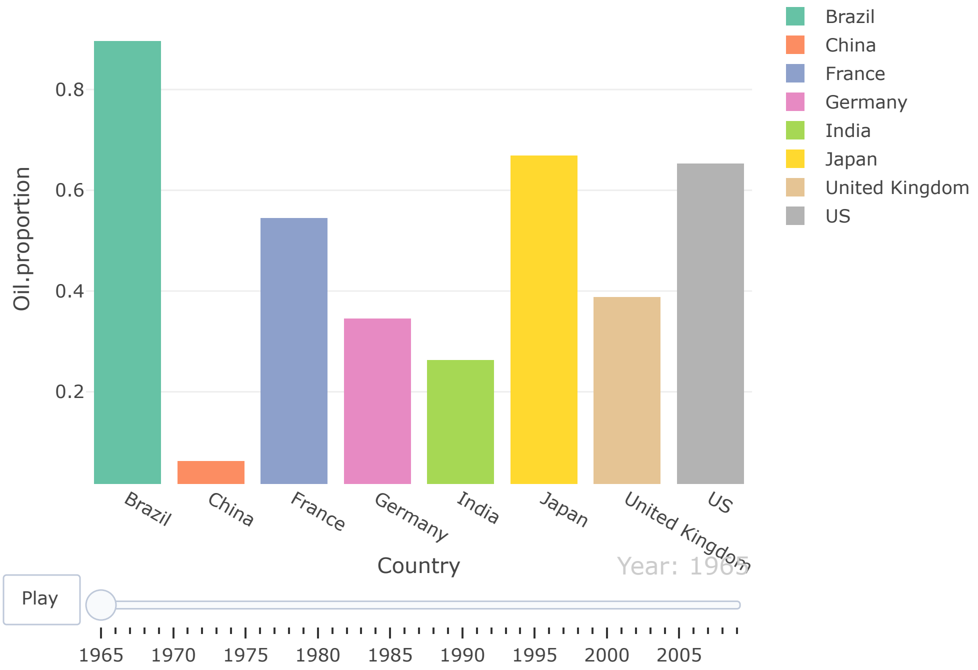
Q1: What are the advantages of visualizing data in this way compared to the animated bubble chart? What are the disadvantages?

Advantages: The all animations occur in one direction - the Y-axis direction in the bar chart. So we can focus and compare the bars with each other. In the bubble chart, bubble move among X-axis and Y-axis, we can follow one or two points but can not follow all in the same time.

Disadvantages: The bar chart has less information than bubble chart, because there is no coal data in bar chart. In the bubble chart, we can see China and US are moving in different directions in terms of energy consumption because of the different energy structures. But in the bar chart we can only see that the proportion of oil consumption is decreasing, but we don’t know whether it is less oil consumption or more coal consumption.

4. Animation easing

```
p2_4 <- oilcoal %>%
  plot_ly(x = ~Country, y = ~Oil.proportion, frame = ~Year, color = ~Country) %>%
  add_bars() %>%
  animation_opts(1000, easing = "elastic")
```

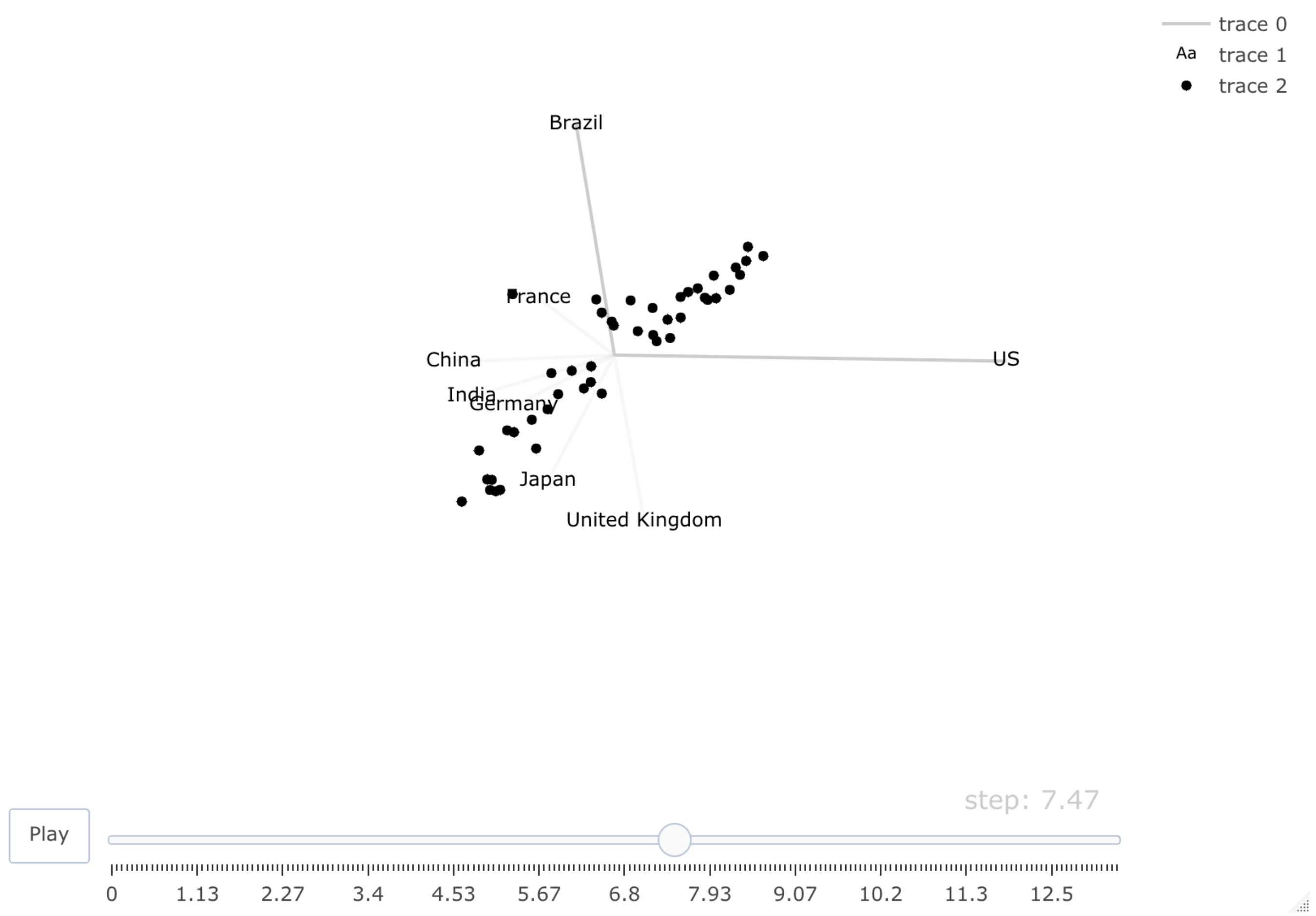


Q1: Which advantages and disadvantages can you see with this animation?

Advantages: From easings.net (https://easings.net), the elastic animation has two spring effects before and after the data change. With this effect, we can clearly see which is the real data and which is the data we generated by interpolation. We can not see this in the default linear animation, where all values change evenly.

Disadvantages: We did not found the disadvantages of elastic animations.

5. Guided 2D-tour



Q1: Do clusters correspond to different Year ranges?

The clusters are correspond to different Year ranges, the left bottom cluster is before 1983 (contain 1983), the right top one is after 1983 and the outlier (or call it one member cluster) on the left top is in 2009.

Q2: Which variable has the largest contribution to this projection? How can this be interpreted?

From the path graph, US has the largest contribution to this projection. We do not know the exact connection between these variables but we found some information as below:

- US has a strong economy and the US dollar has a strong connection to oil settlements.
- US Strategic Petroleum Reserve has been on an upward trend since 1983.

Appendix

Codes For Assignment 1

```
library(plotly)
library(ggraph)
library(igraph)
library(visNetwork)
library(seriation)

train_data=read.table("trainData.dat")
train_meta=read.table("trainMeta.dat")

#1
train_meta$id=1:70
train_meta=train_meta[,c(3,1,2)]
colnames(train_data)=c("from","to","connect")
#net <- graph_from_data_frame(d=train_data, vertices=train_meta$id, directed=T)
#visIgraph(net)
visNetwork(train_meta, train_data)

#use strength of links variable
train_data$value=train_data$connect
train_meta$label=train_meta$V1
#nodes are colored by Bombing Group.
train_meta$group=train_meta$V2
train_meta$group[train_meta$group==1]=" explosives"
train_meta$group[train_meta$group==0]=" otherwise"
#size of nodes is proportional to the number of connections
data_value=data.frame(rle(sort(train_data$from))[2],rle(sort(train_data$from))[1])
train_meta$value=0
train_meta$value[data_value$values]=data_value$lengths

#use a layout that optimizes repulsion forces
#all nodes that are connected to a currently selected node by a path of length one are highlighted
visNetwork(nodes=train_meta, edges=train_data)%>%
  visPhysics(solver="repulsion")%>%
  visOptions(highlightNearest=list(enabled=TRUE,degree=1),
             selectedBy = "group")%>%
               visLegend()

#2
visNetwork(nodes=train_meta, edges=train_data)%>%
  visPhysics(solver="repulsion")%>%
  visEdges(arrows = "to")%>%
  visOptions(highlightNearest=list(enabled=TRUE,algorithm="hierarchical",degree=list(from=1,to=2)),
             selectedBy = 'group')%>%
  visLegend()

#3
#Community identification
nodes1<-train_meta
graph=graph_from_data_frame(d=train_data,vertices = train_meta,directed = F)
cluster=cluster_edge_betweenness(graph,directed = T)
nodes1$group=cluster$membership

visNetwork(nodes=nodes1, edges=train_data)%>%
  visPhysics(solver="repulsion")%>%
  visEdges(arrows = "to")%>%
  visOptions(highlightNearest=list(enabled=TRUE,degree=1),
             selectedBy = 'group')%>%
  visLegend()

#4
#Adjacency representation
adj_matrix<-get.adjacency(graph,attr="value",sparse=F)
colnames(adj_matrix)<-V(graph)$media
rownames(adj_matrix)<-V(graph)$media

rowdist<-dist(adj_matrix)

order1<-seriate(rowdist,"HC")
ord1<-get_order(order1)

reorder_adj_matrix<-adj_matrix[ord1,ord1]
plot_ly(z=~reorder_adj_matrix, x=~colnames(reorder_adj_matrix),
        y=~rownames(reorder_adj_matrix), type="heatmap")
```

Codes For Assignment 2

```r
# Task 1
oilcoal <- read.csv2("Oilcoal.csv")

p2_1 <- oilcoal %>%
  plot_ly(
    x = ~Coal,
    y = ~Oil,
    size = ~Marker.size,
    color = ~Country,
    frame = ~Year,
    text = ~Country,
    hoverinfo = "text",
    type = 'scatter',
    mode = 'markers'
  )

p2_1

# Task 2
p2_2 <- oilcoal %>%
  filter(Country == "US" | Country == "Japan") %>%
  plot_ly(
    x = ~Coal,
    y = ~Oil,
    size = ~Marker.size,
    color = ~Country,
    frame = ~Year,
    text = ~Country,
    hoverinfo = "text",
    type = 'scatter',
    mode = 'markers'
  )

p2_2

# Task 3
p2_3 <- oilcoal %>%
  plot_ly(x = ~Country, y = ~Oil.proportion, frame = ~Year, color = ~Country) %>%
  add_bars() %>%
  animation_opts(1000, easing = "elastic")

p2_3

# Task 4
p2_4 <- oilcoal %>%
  plot_ly(x = ~Country, y = ~Oil.proportion, frame = ~Year, color = ~Country) %>%
  add_bars() %>%
  animation_opts(1000, easing = "elastic")

p2_4

# Task 5
years <- oilcoal$Year
years <- years[!duplicated(years)]

countries <- oilcoal$Country
countries <- countries[!duplicated(countries)]

df_tour <- data.frame(row.names = years)
for (country in countries) {
  c <- oilcoal[which(oilcoal$Country == country),]
  df <- as.data.frame(c$Coal)
  df_tour <- cbind(df_tour, df)
}
colnames(df_tour) <- countries

mat <- rescale(df_tour)
set.seed(12345)
tour <- new_tour(mat, guided_tour(cmass), NULL)

steps <- c(0, rep(1/15, 200))
Projs <- lapply(steps, function(step_size) {
  step <- tour(step_size)
  if(is.null(step)) {
    .GlobalEnv$tour <- new_tour(mat, guided_tour(cmass), NULL) # err
    step <- tour(step_size)
  }
  step
})

# projection of each observation
tour_dat <- function(i) {
  step <- Projs[[i]]
  proj <- center(mat %*% step$proj)
  data.frame(x = proj[,1], y = proj[,2], state = rownames(mat))
}

# projection of each variable's axis
proj_dat <- function(i) {
  step <- Projs[[i]]
  data.frame(
    x = step$proj[,1], y = step$proj[,2], variable = colnames(mat)
  )
}

stepz <- cumsum(steps)

# tidy version of tour data

tour_dats <- lapply(1:length(steps), tour_dat)
tour_datz <- Map(function(x, y) cbind(x, step = y), tour_dats, stepz)
tour_dat <- dplyr::bind_rows(tour_datz)

# tidy version of tour projection data
proj_dats <- lapply(1:length(steps), proj_dat)
proj_datz <- Map(function(x, y) cbind(x, step = y), proj_dats, stepz)
proj_dat <- dplyr::bind_rows(proj_datz)

ax <- list(
  title = "", showticklabels = FALSE,
  zeroline = FALSE, showgrid = FALSE,
  range = c(-1.1, 1.1)
)

# for nicely formatted slider labels
options(digits = 3)
tour_dat <- highlight_key(tour_dat, ~state, group = "A")
tour <- proj_dat %>%
  plot_ly(x = ~x, y = ~y, frame = ~step, color = I("black")) %>%
  add_segments(xend = 0, yend = 0, color = I("gray80")) %>%
  add_text(text = ~variable) %>%
  add_markers(data = tour_dat, text = ~state, ids = ~state, hoverinfo = "text") %>%
  layout(xaxis = ax, yaxis = ax)
tour
```