

Assignment 2.2 - Week 3&4 in R

Aarti Ramani

2023-06-27

```
library(readxl)
library(ggplot2)
library(RColorBrewer)
library(plotly)
```

Load required libraries

```
##
## 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(treemapify)
```

```
## Warning: package 'treemapify' was built under R version 4.2.3
```

```
library(treemap)
```

```
## Warning: package 'treemap' was built under R version 4.2.3
```

```
unemployment_rate_df <- read.csv("C:/Masters/GitHub/Summer2023/DSC640-Data Presentation & Visualization")
nrow(unemployment_rate_df)
```

Read xls into a dataframe

```
## [1] 746
```

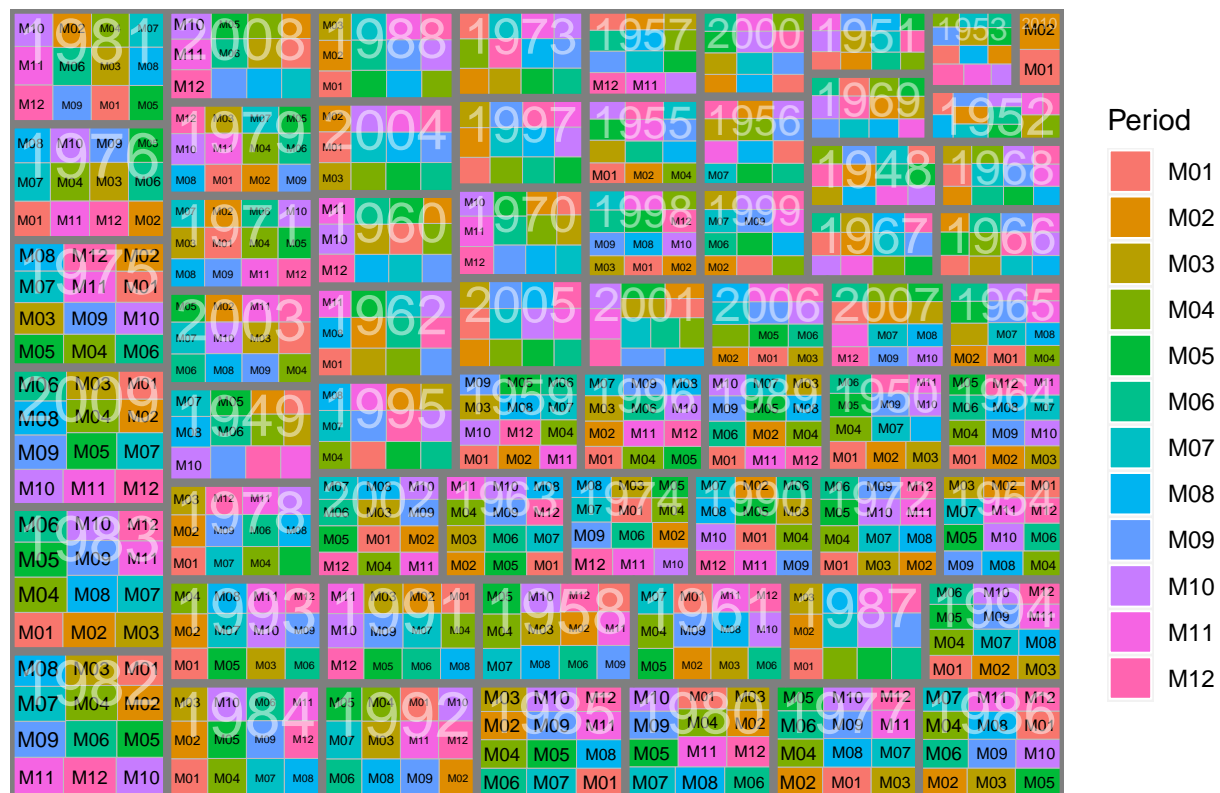
```
head(unemployment_rate_df,5)
```

```
##      Series.id Year Period Value
## 1 LNS14000000 1948   M01   3.4
## 2 LNS14000000 1948   M02   3.8
## 3 LNS14000000 1948   M03   4.0
## 4 LNS14000000 1948   M04   3.9
## 5 LNS14000000 1948   M05   3.5
```

TREE MAP

```
p <- ggplot(unemployment_rate_df, aes(area = Value, fill = Period, label = Period, subgroup = Year)) +
  geom_treemap() + geom_treemap_subgroup_border() +
  geom_treemap_text(place = "middle", grow = F, size = 10) +
  geom_treemap_subgroup_text(place = "top", grow = F, alpha = 0.5, color = "white") +
  labs(title = "R - TreeMap for Unemployment Rate by Year and Month")
p
```

R – TreeMap for Unemployment Rate by Year and Month



The multidimensional chart above appears cluttered and difficult to read because of overlapping. Modifying the same for readability.

```
exp_agg_df <- aggregate(unemployment_rate_df$Value, by = list(Year = unemployment_rate_df$Year), FUN = sum)
group <- exp_agg_df$Year
```

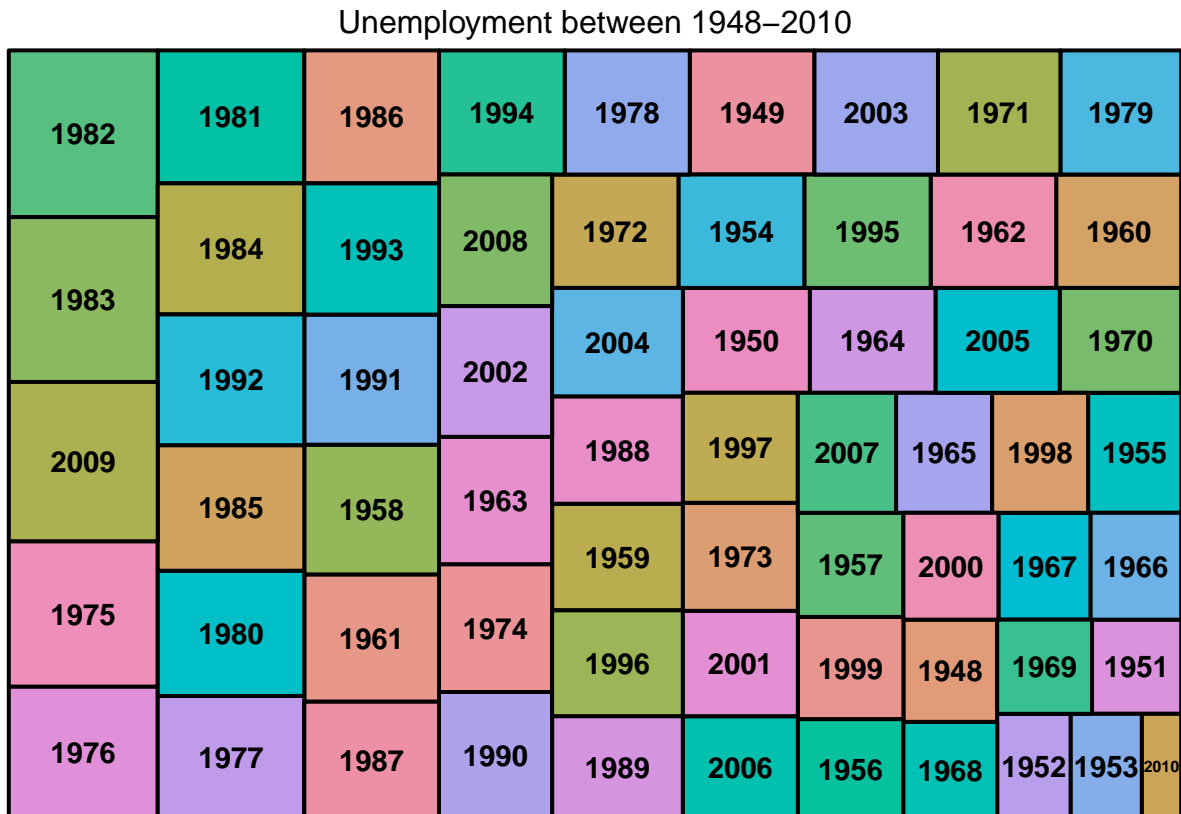
```

value<-exp_agg_df$x

df_exp_agg<-data.frame(group,value)

treemap(df_exp_agg,index="group",vSize="value",type="index",
        title="Unemployment between 1948-2010",fontsize.title = 12)

```



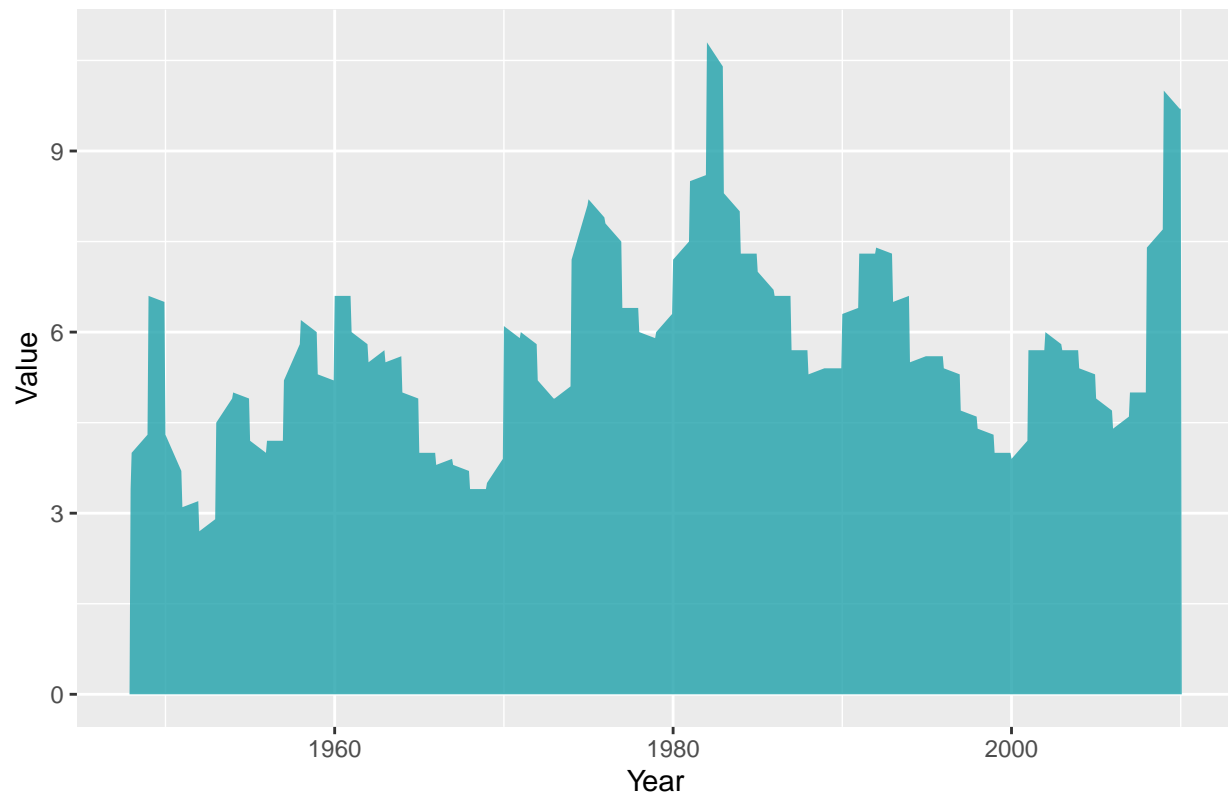
AREA CHART

```

ggplot(unemployment_rate_df,aes(x=Year,y=Value)) +
  geom_area(fill = "#20a2ab", size=0.2,alpha=0.8) +
  ggtitle("R - Unemployment Per Year - Area Chart")

```

R – Unemployment Per Year – Area Chart

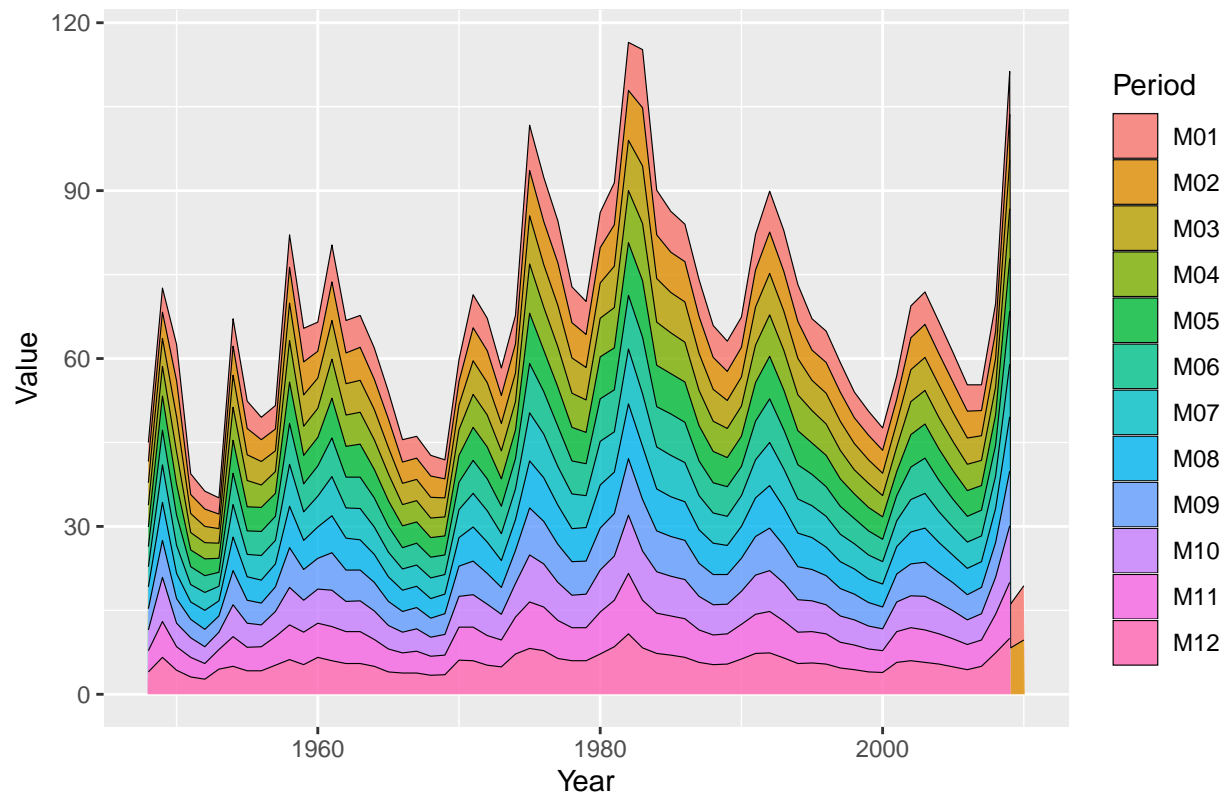


```
#fig <- plot_ly(data = unemployment_rate_df, x = ~Year, y = ~Value, type = 'scatter', mode = 'lines',
#fig <- fig %>% layout(xaxis = list(title = 'Year'),
#      yaxis = list(title = 'Unemployment Rate'),
#      title = "R - Stacked Area Chart for Unemployment Rate by Year")
#fig
```

STACKED AREA CHART

```
ggplot(unemployment_rate_df,aes(x=Year,y=Value,fill=Period)) +
  geom_area(color="black",size=0.2,alpha=0.8) +
  ggtitle("R - Unemployment Per Year - Stacked Area Chart")
```

R – Unemployment Per Year – Stacked Area Chart



STEP CHART

```
p <- ggplot(unemployment_rate_df, aes(Year, Value)) + geom_step()+
  labs(title="R - Step Chart for Unemployment Rate by Year" ) + xlab("Year") + ylab("Unemployment Rate".
p
```

R – Step Chart for Unemployment Rate by Year

