



## Fun Exercise 7

### Applied Longitudinal Data Analysis

- 1a. How did you spend yesterday? Please insert your activities by hour of the day into the table below. You may distinguish states, such as SLEEP, EAST, WORK (including education), LEISURE (sports, friends etc.), MEDIA, OTHER.

ID	TIME	Wednesday
1	8	
1	9	
1	10	
1	11	
1	12	
1	13	
1	14	
1	15	
1	16	
1	17	
1	18	
1	19	
1	20	

**1b. Insert the data into R and translate the data into wide format.**

- Declare the data as a sequence file (`seqdef`)
- Plot the sequence index plots (`seqIplot`)
- Make the output “nicer”

```
install.packages("TraMineR")  
library(TraMineR)
```

**#Insert data (example)**

```
ID <-as.numeric(c(1,1,1,1,1,1,1,1,1,1,1,1,1))  
TIME <-as.numeric(c(8,9,10,11,12,13,14,15,16,17,18,19,20))  
ACTIVITY<-  
as.factor(c("SLEEP","SLEEP","SLEEP","WORK","WORK","WORK","WORK","MED  
IA","EAT", "MEDIA", "WORK","MEDIA","SLEEP"))  
DATA01 = data.frame(ID,TIME,ACTIVITY)  
DATA02<-reshape(DATA01,idvar="ID",timevar="TIME",direction="wide")
```

**#Define sequence object**

```
DATA03 <- seqdef(DATA02[,2:13])
```

**#Draw sequence index plot**

```
seqIplot(DATA03)
```

**# Nicer colors**

```
seqstat1(DATA03[, 1:12])  
cpal(DATA03) <- c("blue","black", "red", "red4")  
seqIplot(DATA03)
```

**# Nicer x-axis**

```
seqIplot(DATA03, xaxis=FALSE, with.legend=FALSE, )  
axis(1, at=c(0,12), labels=c("8 a.m.", "8 p.m."))
```

**# Legend**

```
seqlegend(DATA03, bty="n")
```

**#Everything together**

```
par(mfrow=c(2,1))  
cpal(DATA03) <- c("yellow","green", "red", "blue")  
seqIplot(DATA03, xaxis=FALSE,with.legend=FALSE, )  
axis(1, at=c(0,12), labels=c("8 a.m.", "8 p.m."))  
seqlegend(DATA03, bty="n")
```

**2a. Insert the data below into R. Change the party names and the distribution so that it fits to a country and preference of your choice.**

```
library(TraMineR)
library(reshape2)
library(cluster)
library(WeightedCluster)

#Insert Data
ID <-as.numeric(c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4,4,5,5,5,5,5,6,6,6,6,6))
TIME01 <-as.numeric(c(1,2,3,4,5,1,2,3,4,5,1,2,3,4,5,1,2,3,4,5,1,2,3,4,5,1,2,3,4,5))
TIME<-as.factor(c("CDU","SPD","GREEN","SPD","SPD","CDU","SPD","GREEN","GREEN","GREEN",
"CDU","SPD","GREEN","SPD","GREEN","CDU","SPD","SPD","SPD","SPD","SPD","SPD","SPD","GREEN","GREEN",
"SPD","CDU","SPD","GREEN","SPD","SPD"))
DATA01 = data.frame(ID,TIME01,TIME)
DATA02<-reshape(DATA01,idvar="ID",timevar="TIME01",direction="wide")
```

**2b. Conduct a sequence analysis with the data.**

- Declare the data as a sequence file (seqdef)
- Plot the sequence index plots (seqIplot)
- Plot the state distribution (seqdplot)
- Generate and display (print) the distance matrix (seqdist)
- Cluster the data (hclust)
- Group the data into 2 clusters (cutree)
- Display the sequence index plots of the two clusters
- Label the two clusters in an appropriate manner

```
# Declare the data as a sequence file
DATA03 <- seqdef(DATA02[,2:6])

#Plot the sequence index
cpal(DATA03) <- c("black","green", "red")
seqIplot(DATA03, border=T, with.legend="right")

#Plot the state distribution
seqdplot(DATA03, border=T, with.legend="right")

# Generate and display (print) the distance matrix
DISTANCE <- seqdist(DATA03, method = "LCS")
print(DISTANCE)

# Cluster the data (hclust)
CLUSTER <-hclust(as.dist(DISTANCE), method="ward.D")
plot(CLUSTER)

# Group the data into 2 clusters
TREE <- cutree(CLUSTER, k = 2)

# Display the sequence index plots of the two clusters
seqIplot(DATA03, border=T, with.legend="right", group = TREE)

# Label the two clusters in an appropriate manner
LABELS <- c("SPD changers", "Green Transitioners")
TREE01<- factor(TREE, levels = 1:2, labels = LABELS)
seqIplot(DATA03, border=T, with.legend="right", group = TREE01)
```