# Enquette Budget Temps

## Langage R et Analyse de donnée

**Etienne Sharpin**
**Jose A. Henriquez Roa**

May 19, 2021

# Contents

# 1 Carte Radar

For comparion of the time allocated to different tasks between individuals of different groups we have choosen the method of radar chart. The following code snippet shows how we plotted these charts for all 28 groups, shown in the annex. Along with a radar chart containing all groups containing all groups shown at the end of this section.

```
library(fmsb)

pdf("plot/radar-chart-all.pdf")
df04 <- rbind(rep(max(df03),28), rep(min(df03),28), df03)
radarchart(df04, title="Per-Instance comparison")
dev.off()

for (i in 1:28) {
    pdf(paste("plot/radar-chart/radar-chart-", tolower(df01$ID[i]), ".pdf", sep=""))
    df05 <- rbind(rep(max(df03),28), rep(min(df03),28), df03[i,])
    radarchart(df05, title=df01$ID[i])
    dev.off()
}
```

For this we have used the library *fmsb* which includes the function *radarchart(...)* from which the plots were drawn. Figure 1 shows a comparison of time distributions between all groups present in the dataset. The individual radar charts can be found in the annex **??**

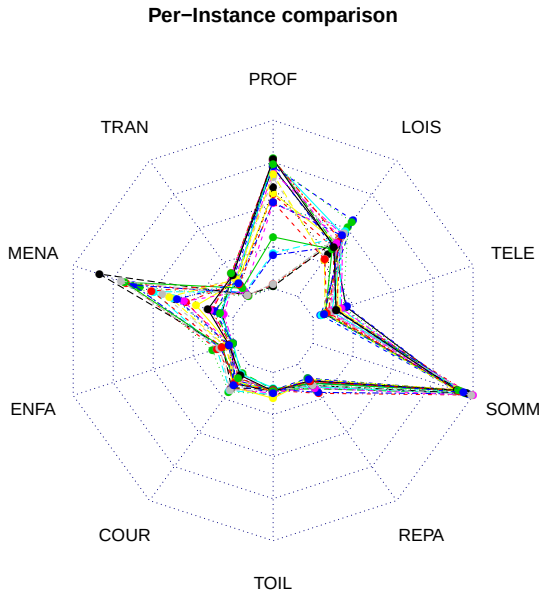


Figure 1: Comparison in between all profile groups

## 2 Graphique à barres

Since, radar charts that incorporate reference values tend to look rather clutered we have also chosen to visualize time distribution with the use of bar plots, in which we have included the time percentages in the y axis.

```
for (i in 1:28) {
    pdf(paste("plot/bar-plot/bar-plot-", tolower(df01$ID[i]), ".pdf", sep=""))
    barplot(t(as.matrix(df03[i,])), beside=TRUE, main=df01$ID[i],
            ylab="Time percentage", names.arg=colnames(df03), las=2)
    dev.off()
}
```

The generated plots are in the annex **??**

## 3 ACP

The following line performs the PCA which we use in the following clustering section for visualization of instance clustering results.

```
  pca = prcomp(df03)
```

The percentages of the explained variance for each principal component is gotten through the following command.

```
  summary(pca)$importance[2,]
```

which produces the following output:

```
PC1: 0.8804 PC2: 0.07166 PC3: 0.02654 PC4: 0.01521 PC5: 0.0032 PC6: 0.00158
PC7: 0.00072 PC8: 0.00043 PC9: 0.00026 PC10: 0
```

As we can see a little more that 95% of the variance is explained by the two first principal components. Which means that the two dimensional projection of the data will be of fairly good quality.

The followng plots a graphical representation of the above output

```
pdf("plot/principal-component-explained-variance.pdf")
barplot(summary(pca)$importance[2,],ylab="Explained Variance Proportion",
        ylim=c(0,1), main="Principal Component Explained Variance")
dev.off()
```

The resulting plot can be seen in figure 2.

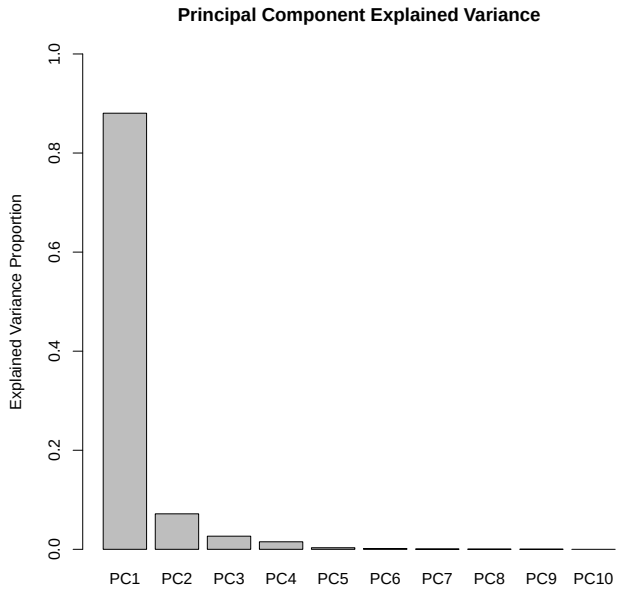


Figure 2: Principal component explained variance

The following code snippet shows the projection of all attributes into the two first principal components.

```
pdf("plot/pca-attribute-projection.pdf")
library("factoextra")
fviz_pca_var(pca, col.var = "cos2", col.ind = "cos2",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"))
dev.off()
```

As show in figure 3. In this one we see that the two best projected attributes are by far the PROF and MENA attributes, corresponding respectively to the attributes designating the time spent working and cleaning.

Next, the visualization of the instance projection was done through the following code.

```
pdf("plot/pca-instance-projection.pdf")
library("factoextra")
fviz_pca_ind(pca, col.var = "cos2", col.ind = "cos2",
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"))
dev.off()
```

The resulting plot is in figure 4. As according to the percentages of explained variance we see that most instance have a good projection quality.

# 4 Clustering

To analyse similarities in between the group profiles themselves we performed clustered the instances throught the use of the KMeans Machine Learning algorithm.

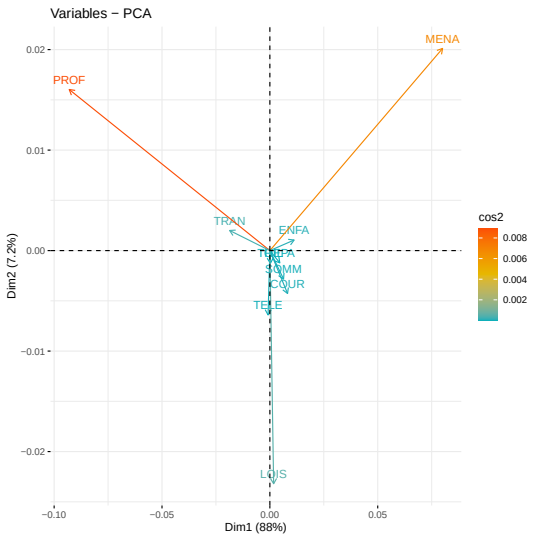


Figure 3: PCA attribute projection

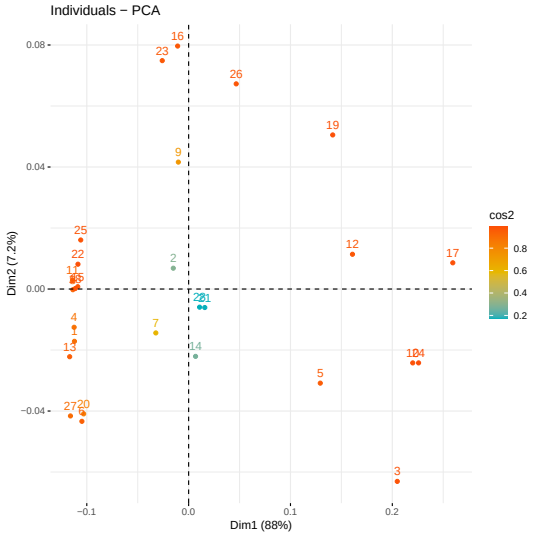


Figure 4: PCA instance projection