tmp

May 18, 2021

R Project

Loading the data

[1]:
```
path = "/home/otiose/repos/epita/rlang/Enquete_Budget_Temps.xlsx"
library("readxl")
df01 <- read_excel(path)
```

[59]:
```
summary(df01)
```

```
      ID                 PROF            TRAN             MENA      
 Length:28          Min.   : 10.0   Min.   :  0.00   Min.   : 50.0  
 Class :character   1st Qu.:386.8   1st Qu.: 47.50   1st Qu.: 96.5  
 Mode  :character   Median :535.5   Median : 95.50   Median :256.0  
                    Mean   :450.5   Mean   : 86.11   Mean   :277.1  
                    3rd Qu.:631.0   3rd Qu.:127.00   3rd Qu.:424.0  
                    Max.   :656.0   Max.   :148.00   Max.   :710.0  
      ENFA             COUR            TOIL             REPA      
 Min.   :  0.00   Min.   : 52.0   Min.   : 77.00   Min.   : 85.0  
 1st Qu.: 10.00   1st Qu.: 85.0   1st Qu.: 90.00   1st Qu.:100.0  
 Median : 22.00   Median :112.0   Median : 92.00   Median :111.0  
 Mean   : 33.32   Mean   :108.8   Mean   : 94.93   Mean   :118.1  
 3rd Qu.: 56.00   3rd Qu.:131.8   3rd Qu.: 96.25   3rd Qu.:132.5  
 Max.   :110.00   Max.   :170.0   Max.   :130.00   Max.   :180.0  
      SOMM            TELE             LOIS            SEX       
 Min.   :745.0   Min.   : 40.00   Min.   :228.0   Min.   :1.000  
 1st Qu.:762.2   1st Qu.: 64.75   1st Qu.:308.8   1st Qu.:1.000  
 Median :775.0   Median : 91.50   Median :347.0   Median :2.000  
 Mean   :785.9   Mean   : 99.43   Mean   :345.8   Mean   :1.571  
 3rd Qu.:809.2   3rd Qu.:122.75   3rd Qu.:385.8   3rd Qu.:2.000  
 Max.   :849.0   Max.   :180.00   Max.   :475.0   Max.   :2.000  
      ACT             CIV             PAY       
 Min.   :1.000   Min.   :1.000   Min.   :1.00  
 1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.75  
 Median :9.000   Median :2.000   Median :2.50  
 Mean   :5.714   Mean   :4.714   Mean   :2.50  
 3rd Qu.:9.000   3rd Qu.:9.000   3rd Qu.:3.25  
 Max.   :9.000   Max.   :9.000   Max.   :4.00  
```

Normalizing the distributions

[2]:
```
df02 <- df01
df02[, 2:11] <- df01[, 2:11] / 2400
```

[3]:
```
df02
```

A tibble: 28 × 15

ID <chr>	PROF <dbl>	TRAN <dbl>	MENA <dbl>	ENFA <dbl>	COUR <dbl>	TOIL <dbl>	RE <dl
HAU	0.254166667	0.058333333	0.02500000	0.004166667	0.05000000	0.03958333	0.0
FAU	0.197916667	0.037500000	0.10416667	0.012500000	0.05833333	0.05000000	0.0
FNU	0.004166667	0.000000000	0.20625000	0.045833333	0.07083333	0.04583333	0.0
HMU	0.256250000	0.058333333	0.02708333	0.004166667	0.04791667	0.03750000	0.0
FMU	0.074583333	0.012083333	0.17541667	0.036250000	0.06708333	0.04666667	0.0
HCU	0.243750000	0.047916667	0.02083333	0.000000000	0.06250000	0.04375000	0.0
FCU	0.200833333	0.039166667	0.08166667	0.007500000	0.05875000	0.05416667	0.0
HAW	0.272083333	0.041666667	0.03958333	0.002916667	0.02375000	0.03541667	0.0
FAW	0.212916667	0.029166667	0.12791667	0.012500000	0.03333333	0.03958333	0.0
FNW	0.008333333	0.002916667	0.23666667	0.036250000	0.04666667	0.03750000	0.0
HMW	0.273333333	0.040416667	0.04041667	0.004166667	0.02166667	0.03541667	0.0
FMW	0.070000000	0.009166667	0.22000000	0.028750000	0.04250000	0.03458333	0.0
HCW	0.267916667	0.043750000	0.03000000	0.000000000	0.02583333	0.03208333	0.0
FCW	0.178750000	0.014166667	0.10916667	0.005833333	0.03833333	0.04041667	0.0
HAY	0.270833333	0.058333333	0.05000000	0.006250000	0.03541667	0.03750000	0.0
FAY	0.233333333	0.043750000	0.15625000	0.018750000	0.03750000	0.03750000	0.0
FNY	0.004166667	0.004166667	0.29583333	0.022916667	0.06041667	0.03541667	0.0
HMY	0.270833333	0.060416667	0.04666667	0.006250000	0.03541667	0.03750000	0.0
FMY	0.108333333	0.021666667	0.24000000	0.024583333	0.04833333	0.03541667	0.0
HCY	0.256250000	0.052083333	0.03958333	0.000000000	0.04791667	0.03750000	0.0
FCY	0.180416667	0.037083333	0.13250000	0.009583333	0.04666667	0.04000000	0.0
HAE	0.270833333	0.059166667	0.05083333	0.009166667	0.03166667	0.03916667	0.0
FAE	0.240833333	0.044166667	0.14083333	0.017500000	0.04416667	0.03916667	0.0
FNE	0.010000000	0.003333333	0.24750000	0.030000000	0.06583333	0.03833333	0.0
HME	0.271666667	0.055416667	0.05583333	0.009166667	0.02833333	0.03916667	0.0
FME	0.181666667	0.032916667	0.18041667	0.025000000	0.04958333	0.03750000	0.0
HCE	0.261250000	0.061666667	0.02833333	0.000000000	0.03666667	0.03833333	0.0
FCE	0.180833333	0.035833333	0.12375000	0.008750000	0.05375000	0.04250000	0.0

Boxplot Analytics

[4]:
```
df03 <- df02[2:11]
```

[68]:
```
pdf("plot/box-plot.pdf")
boxplot(df03, use.cols=FALSE, las=2, ylab="Time percentage", main="Analytics of␣
 ↪time spent per-activity")
dev.off()
```

png: 2

Correlation Matrix

[62]:
```
pdf("plot/correlation-matirx.pdf")
library("PerformanceAnalytics")
chart.Correlation(df03, histogram=TRUE, title="hello")

mtext("Correlation Matrix", side=3, line=3)
dev.off()
```

png: 2

Radar Charts

[87]:
```
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()
}
```

png: 2

Bar Plots

[88]:
```
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()
}
```

PCA

[73]:
```
pca = prcomp(df03)
```

[70]:
```
summary(pca)$importance[2,]
```

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

[72]:
```
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()
```

png: 2

[79]:
```
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()
```

png: 2

[80]:
```
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()
```

png: 2

Clustering

[81]:
```
pdf("plot/kmeans-inertia.pdf")
fviz_nbclust(df03, kmeans, method="wss")
dev.off()
pdf("plot/kmeans-silhouette.pdf")
fviz_nbclust(df03, kmeans, method="silhouette")
dev.off()
pdf("plot/kmeans-gap-stat.pdf")
fviz_nbclust(df03, kmeans, method="gap_stat")
dev.off()
```

png: 2

png: 2

png: 2

[82]:
```
pdf("plot/clusters.pdf")
km = kmeans(df03, centers=2, nstart=25)
fviz_cluster(km, data=df03)
dev.off()
```

png: 2

In-Cluster Instance Comparison

[84]:
```
pdf("plot/fne-hme-comparison.pdf")
instanceIndices = c(24,25)
df04 <- rbind(rep(max(df03),28), rep(min(df03),28), df03[instanceIndices,])
title = paste(df01$ID[instanceIndices[1]],"-",df01$ID[instanceIndices[2]],␣
 ↪"Comparison", sep=" ")
radarchart(df04, title=title)
dev.off()
```

png: 2

[85]:
```
pdf("plot/hme-fay-comparison.pdf")
instanceIndices = c(25,16)
df04 <- rbind(rep(max(df03),28), rep(min(df03),28), df03[instanceIndices,])
title = paste(df01$ID[instanceIndices[1]],"-",df01$ID[instanceIndices[2]],␣
 ↪"Comparison", sep=" ")
radarchart(df04, title=title)
dev.off()
```

png: 2