

Analyse des Correspondances Multiples

Serigne Fallou Mbacke NGOM

2023-09-22

L'objectif de l'Analyse des Correspondances Multiples (ACM) est de résumer un tableau de données où les individus sont décrits par des variables qualitatives. L'ACM permet d'étudier les ressemblances entre individus du point de vue de l'ensemble des variables et de dégager des profils d'individus. Elle permet également de faire un bilan des liaisons entre variables et d'étudier les associations de modalités.

Importer les donnees:

```
## Warning: le package 'readxl' a été compilé avec la version R 4.2.3
```

```
for (i in c("SEXE", "TOUX", "DOUL_THOR", "TACHYPNEE", "GRIPPE")){
  Data_CI[, i] = factor(Data_CI[, i])
}

for (i in c("TOUX", "DOUL_THOR", "TACHYPNEE", "GRIPPE")){
  levels(Data_CI[,i]) = c(paste(i, "Oui", sep = "_"), paste(i, "Non", sep = "_"))
}

levels(Data_CI[, 'SEXE']) = c("Homme", "Femme", "Nsp")
head(Data_CI)
```

```
## SEXE TOUX DOUL_THOR TACHYPNEE GRIPPE
## 1 Homme TOUX_Oui DOUL_THOR_Oui TACHYPNEE_Oui GRIPPE_Oui
## 2 Femme TOUX_Oui DOUL_THOR_Oui TACHYPNEE_Oui GRIPPE_Oui
## 3 Homme TOUX_Oui DOUL_THOR_Oui TACHYPNEE_Oui GRIPPE_Oui
## 4 Femme TOUX_Non DOUL_THOR_Oui TACHYPNEE_Oui GRIPPE_Non
## 5 Femme TOUX_Non DOUL_THOR_Oui TACHYPNEE_Oui GRIPPE_Non
## 6 Femme TOUX_Oui DOUL_THOR_Oui TACHYPNEE_Oui GRIPPE_Non
```

```
summary(Data_CI)
```

```
## SEXE TOUX DOUL_THOR TACHYPNEE
## Homme:1747 TOUX_Oui:2798 DOUL_THOR_Oui:3708 TACHYPNEE_Oui:3708
## Femme:1879 TOUX_Non: 925 DOUL_THOR_Non: 0 TACHYPNEE_Non: 0
## Nsp : 97 NA's : 15 NA's : 15
## GRIPPE
## GRIPPE_Oui:2834
## GRIPPE_Non: 889
##
```

Analyse des Correspondances Multiples (ACM2):

```
library(FactoMineR)
```

```
## Warning: le package 'FactoMineR' a été compilé avec la version R 4.2.3
```

```
library(factoextra)
```

```
## Warning: le package 'factoextra' a été compilé avec la version R 4.2.3
```

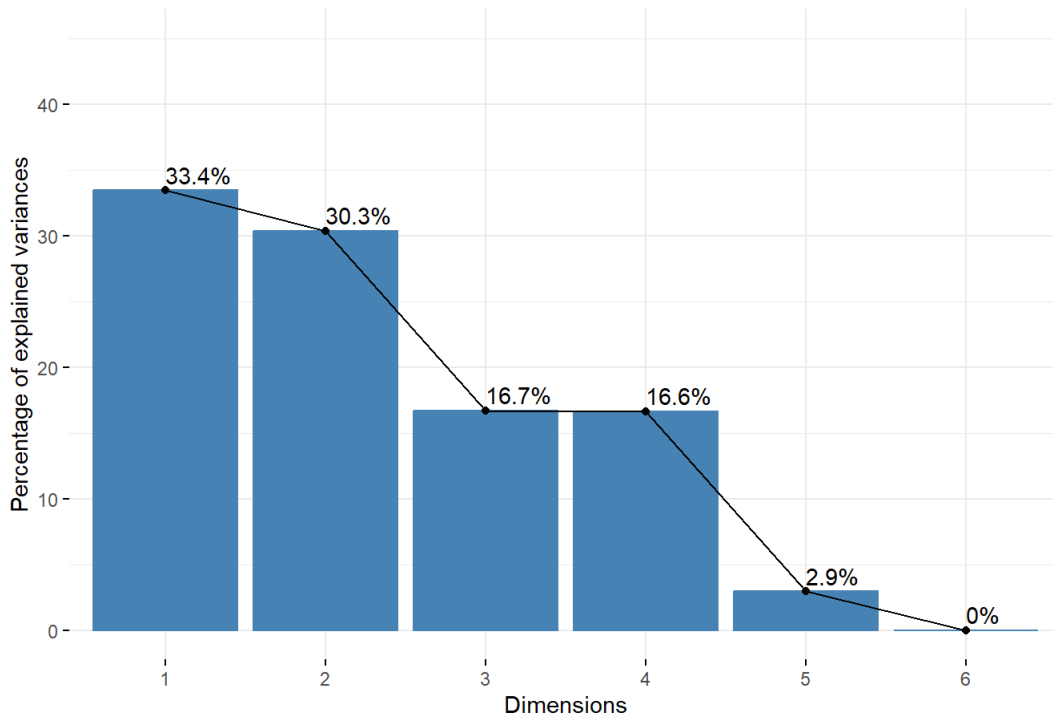
```
## Le chargement a nécessité le package : ggplot2
```

```
## Warning: le package 'ggplot2' a été compilé avec la version R 4.2.3
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
mca <- MCA(Data_CI, graph = FALSE)
fviz_screplot(mca, addlabels = TRUE, ylim = c(0, 45))
```

Scree plot

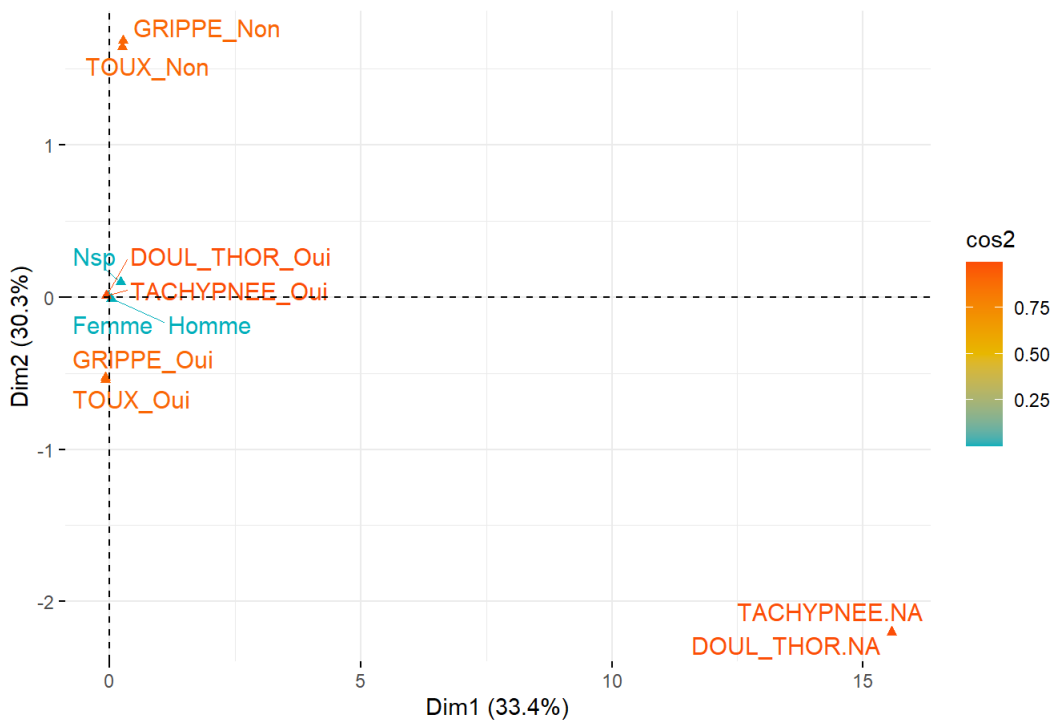


```
fviz_mca_biplot(mca,axes = c(1, 2), col.var='cos2', gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),repel = TRUE, invisible="ind", ggtheme=theme_minimal())
```

Scale for colour is already present.

Adding another scale for colour, which will replace the existing scale.

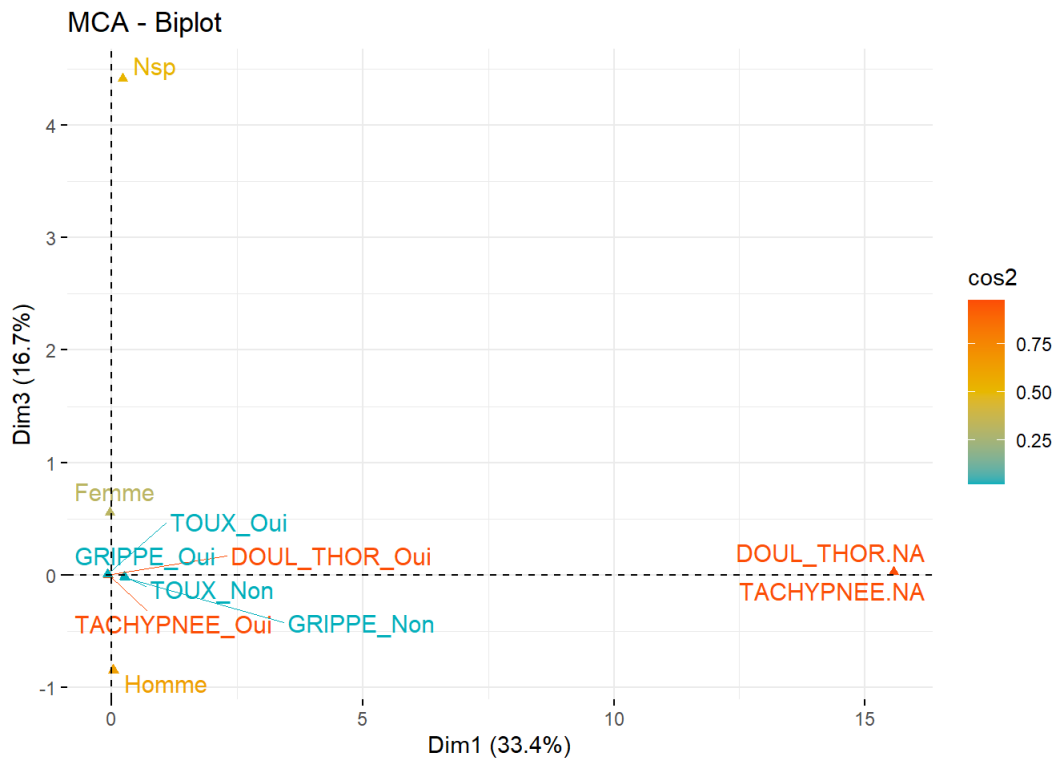
MCA - Biplot



```
fviz_mca_biplot(mca,axes = c(1, 3), col.var='cos2', gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),repel = TRUE, invisible="ind", ggtheme=theme_minimal())
```

Scale for colour is already present.

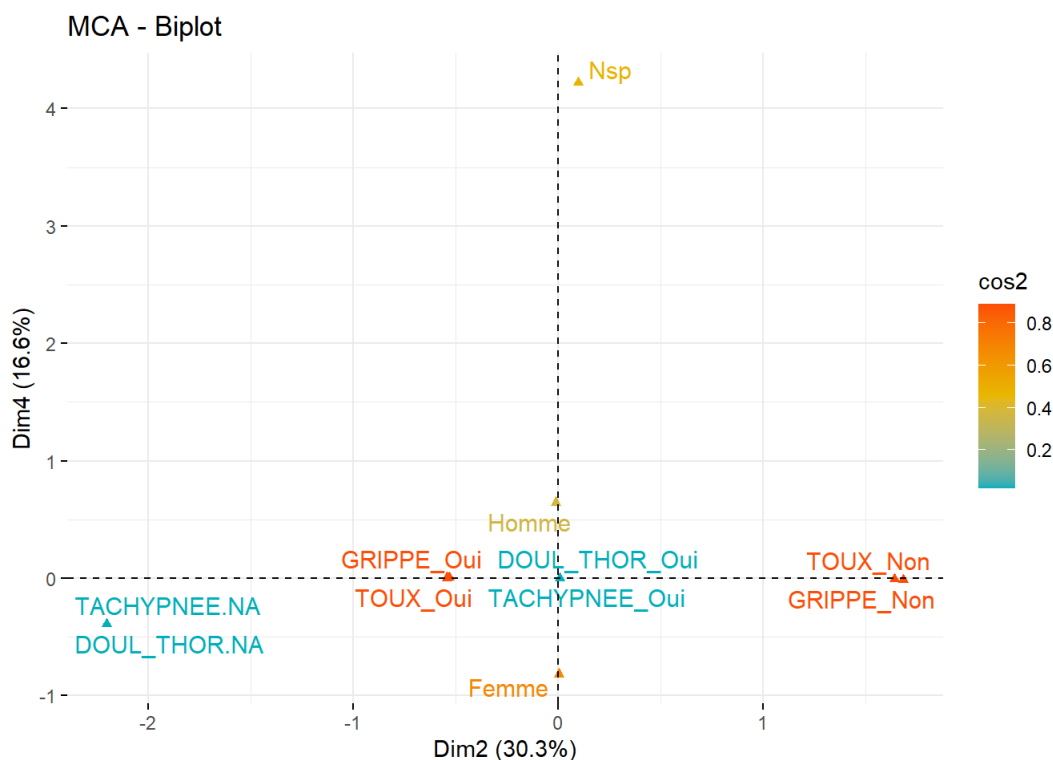
Adding another scale for colour, which will replace the existing scale.



```
fviz_mca_biplot(mca, axes = c(2, 4), col.var = 'cos2', gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE, invisible = "ind", ggtheme = theme_minimal())
```

Scale for colour is already present.

Adding another scale for colour, which will replace the existing scale.



Nous constatons que les quatres(4) premiers composantes expliquent 97% de l'information. L'analyse des composantes nous permet de déduire qu'en général:

- Les patients grippés (GRIPPE_Oui), présentent des toux (TOUX_Oui) et vice versa. Et aussi ceux non grippés (GRIPPE_Non), ne présentent pas de toux (TOUX_Non) et vice versa. Donc on pourra regrouper les deux variables en une seule : TOUX + GRIPPE = GRIP_T ou TOUX_G ...
- Les personnes qui présentent des douleurs thoraciques (DOUL_THOR_Oui), présentent également de la tachynée (TACHYPNEE_Oui).
- Il existe un faible rapprochement entre les deux(2) groupes de patients [avec toux (TOUX_Oui), ceux grippés (GRIPPE_Oui)] et [ceux présentant des tachynées (TACHYPNEE_Oui) et douleurs thoraciques (DOUL_THOR_Oui)].

Arbres:

Les arbres sont des outils d'exploration des données et d'aide à la décision qui permettent d'expliquer et de prédire une variable quantitative (arbre de régression) ou qualitative (arbre de classification) à partir de variables explicatives quantitatives et/ou qualitatives

```
library("rpart")
```

```
## Warning: le package 'rpart' a été compilé avec la version R 4.2.3
```

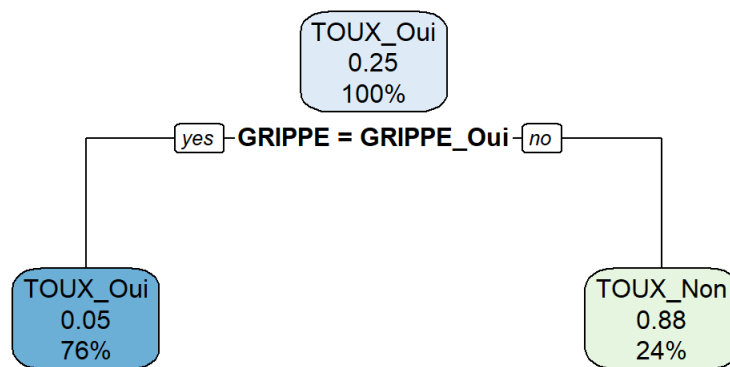
```
library("rpart.plot")
```

```
## Warning: le package 'rpart.plot' a été compilé avec la version R 4.2.3
```

```
library(tree)
```

```
## Warning: le package 'tree' a été compilé avec la version R 4.2.3
```

```
tree.mode2 <- rpart(TOUX ~ GRIPPE + DOUL_THOR + TACHYPNEE, method="class", data=Data_CI)  
rpart.plot(tree.mode2)
```



L'arbre de décision rejoint notre

hypothèse selon laquelle dès l'instant que le patient est grippé, donc il présente des toux.