

# ACP\_Carctéristiques\_Entreprises\_Non\_Agricoles

Michel

2025-05-15

## Contents

<b>1</b>	<b>1. Installation et chargement des packages</b>	<b>1</b>
<b>2</b>	<b>2. Définition du chemin vers le dossier de la base</b>	<b>3</b>
<b>3</b>	<b>3. Importation et nettoyage des données</b>	<b>3</b>
<b>4</b>	<b>4. Aperçu des variables, types et labels</b>	<b>3</b>
<b>5</b>	<b>5. Sélection et préparation des variables quantitatives</b>	<b>4</b>
<b>6</b>	<b>6. Standardisation des données</b>	<b>4</b>
<b>7</b>	<b>7. Analyse en Composantes Principales (ACP)</b>	<b>4</b>
<b>8</b>	<b>8. Visualisations</b>	<b>4</b>
8.1	8.1. Variance expliquée par les composantes principales . . . . .	4
8.2	8.2. Cercle des corrélations des variables . . . . .	5
8.3	8.3. Représentation des individus . . . . .	6
<b>9</b>	<b>9. Résultats détaillés</b>	<b>7</b>
9.1	9.1. Coordonnées des individus . . . . .	7
9.2	9.2. Contributions des variables . . . . .	8
<b>1</b>	<b>1. Installation et chargement des packages</b>	

```
options(repos = c(CRAN = "https://cloud.r-project.org"))
```

```
packages <- c("haven", "dplyr", "FactoMineR", "factoextra", "stringr", "knitr", "kableExtra",
new.packages <- packages[!(packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages)
lapply(packages, library, character.only = TRUE)
```

```
## [[1]]
## [1] "haven"      "stats"      "graphics"   "grDevices" "utils"      "datasets"
## [7] "methods"    "base"
##
## [[2]]
## [1] "dplyr"      "haven"      "stats"      "graphics"   "grDevices" "utils"
## [7] "datasets"   "methods"    "base"
##
## [[3]]
## [1] "FactoMineR" "dplyr"      "haven"      "stats"      "graphics"
## [6] "grDevices"  "utils"      "datasets"   "methods"    "base"
##
## [[4]]
## [1] "factoextra" "ggplot2"    "FactoMineR" "dplyr"      "haven"
## [6] "stats"      "graphics"   "grDevices"  "utils"      "datasets"
## [11] "methods"    "base"
##
## [[5]]
## [1] "stringr"    "factoextra" "ggplot2"    "FactoMineR" "dplyr"
## [6] "haven"      "stats"      "graphics"   "grDevices"  "utils"
## [11] "datasets"   "methods"    "base"
##
## [[6]]
## [1] "knitr"      "stringr"    "factoextra" "ggplot2"    "FactoMineR"
## [6] "dplyr"      "haven"      "stats"      "graphics"   "grDevices"
## [11] "utils"      "datasets"   "methods"    "base"
##
## [[7]]
## [1] "kableExtra" "knitr"      "stringr"    "factoextra" "ggplot2"
## [6] "FactoMineR" "dplyr"      "haven"      "stats"      "graphics"
## [11] "grDevices"  "utils"      "datasets"   "methods"    "base"
##
## [[8]]
## [1] "rstudioapi" "kableExtra" "knitr"      "stringr"    "factoextra"
## [6] "ggplot2"     "FactoMineR" "dplyr"      "haven"      "stats"
## [11] "graphics"    "grDevices"  "utils"      "datasets"   "methods"
## [16] "base"
```

## 2 2. Définition du chemin vers le dossier de la base

```
if(requireNamespace("rstudioapi", quietly = TRUE) && rstudioapi::isAvailable()) {
  senegaldata <- dirname(rstudioapi::getSourceEditorContext()$path)
} else {
  senegaldata <- "."
}
cat("Dossier des données :", senegaldata, "\n")
```

```
## Dossier des données : .
```

## 3 3. Importation et nettoyage des données

```
ma_base <- read_dta(file.path(senegaldata, "s10_2_me_SEN2018.dta"))
ma_base <- ma_base %>% filter(!(s10q17b %in% 1:5))
```

## 4 4. Aperçu des variables, types et labels

```
var_types_labels <- data.frame(
  Variable = names(ma_base),
  Type = sapply(ma_base, function(x) class(x)[1]),
  Label = sapply(ma_base, function(x) attr(x, "label"))
)
knitr::kable(var_types_labels[1:20, ], caption = "Variables, Types et Labels (extrait)") %>%
  kableExtra::kable_styling(full_width = FALSE, position = "left")
```

Table 1: Variables, Types et Labels (extrait)

	Variable	Type	Label
vague	vague	numeric	Vague
grappe	grappe	numeric	grappe
menage	menage	numeric	Identifiant du ménage
s10q12a__1	s10q12a__1	numeric	10.12a.Lister entreprises possédées par le MEN: NUMERO ORDR
s10q12a__2	s10q12a__2	character	10.12a.Lister entreprises possédées par le MEN: NOM ENTREPRI
s10q13	s10q13	numeric	10.13.Quel est le répondant principal, pour cette entreprise?
s10q15__0	s10q15__0	numeric	10.15.Code ID du propriétaire1 de cette entreprise
s10q15__1	s10q15__1	numeric	10.15.Code ID du propriétaire2 de cette entreprise
s10q17a	s10q17a	haven_labelled	10.17a Code de la branche d'activité
s10q17b	s10q17b	haven_labelled	10.17b CODE DE BRANCHE D'ACTIVITÉ

s10q17c	s10q17c	haven_labelled	10.17c CODE D'ACTIVITÉ
s10q18a	s10q18a	haven_labelled	10.18a Y a-t-il des membres de votre ménage qui gèrent cette entreprise?
s10q18__0	s10q18__0	numeric	10.18_0. Code ID de la personne1 qui gère cette entreprise?
s10q18__1	s10q18__1	numeric	10.18_1. Code ID de la personne2 qui gère cette entreprise?
s10q18__2	s10q18__2	numeric	10.18_2. Code ID de la personne3 qui gère cette entreprise?
s10q19a	s10q19a	haven_labelled	10.19a Y a-t-il des mbres de votre mén qui gèrent les revenus de ce
s10q19__0	s10q19__0	numeric	10.19_0. Code ID du membre1 qui gère les revenus de cette entrep
s10q19__1	s10q19__1	numeric	10.19_1. Code ID du membre2 qui gère les revenus de cette entrep
s10q20	s10q20	numeric	10.20 Depuis quand (date ) cette entreprise fonctionne-t-elle?
s10q21	s10q21	numeric	10.21 Comb d pers non mbres du mén st co-propriétaires de cette c

## 5 5. Sélection et préparation des variables quantitatives

```
vars_quant <- ma_base %>%
  select(where(is.numeric)) %>%
  select(where(~ mean(is.na(.)) < 0.3)) %>%
  select(where(~ sd(., na.rm = TRUE) > 0)) %>%
  select(-any_of(c("vague", "grappe", "menage")))

vars_imputed <- vars_quant %>%
  mutate(across(everything(), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))

stopifnot(!any(is.na(vars_imputed)))
stopifnot(all(sapply(vars_imputed, sd) > 0))
```

## 6 6. Standardisation des données

```
vars_scaled <- scale(vars_imputed)
```

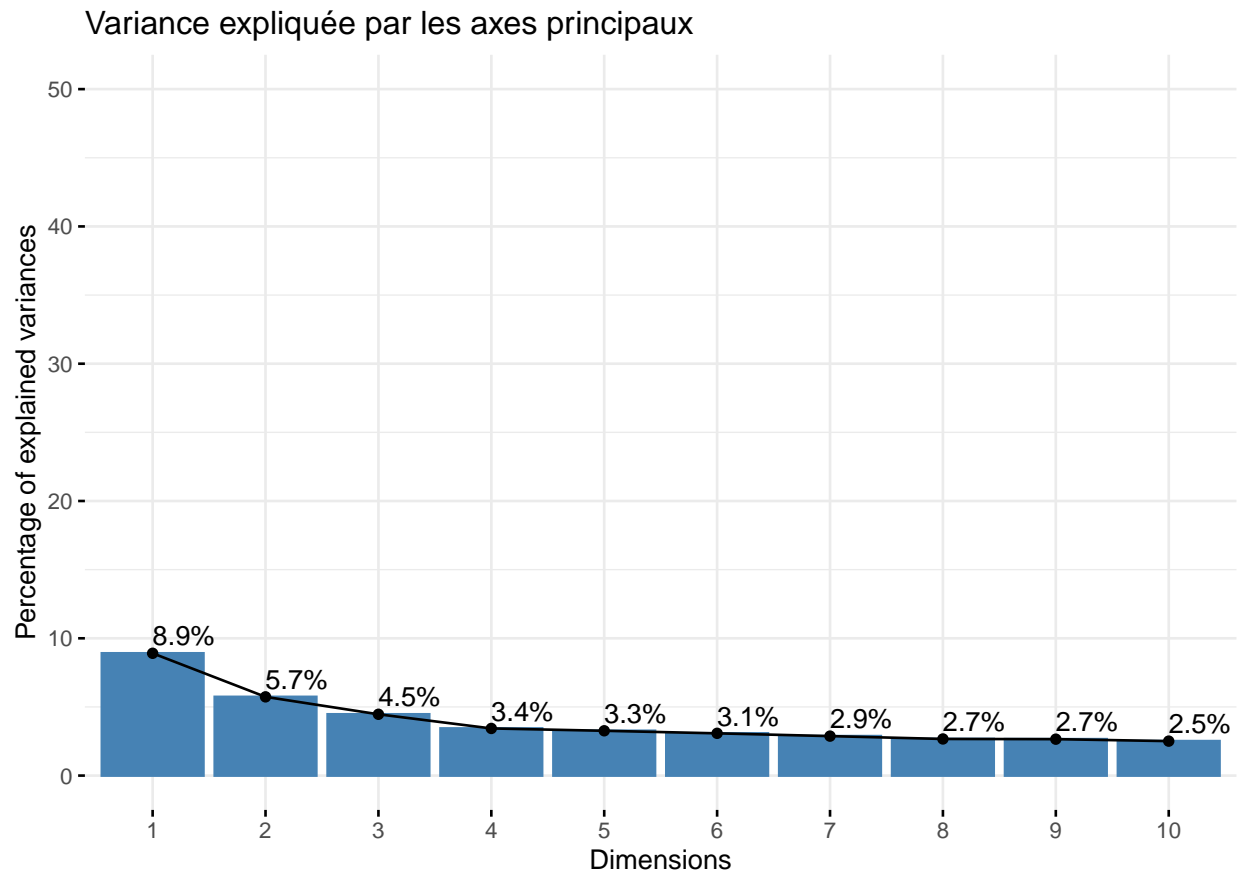
## 7 7. Analyse en Composantes Principales (ACP)

```
res_acp <- PCA(vars_scaled, graph = FALSE)
```

## 8 8. Visualisations

### 8.1 8.1. Variance expliquée par les composantes principales

```
fviz_eig(res_acp, addlabels = TRUE, ylim = c(0, 50)) +  
  ggtitle("Variance expliquée par les axes principaux")
```

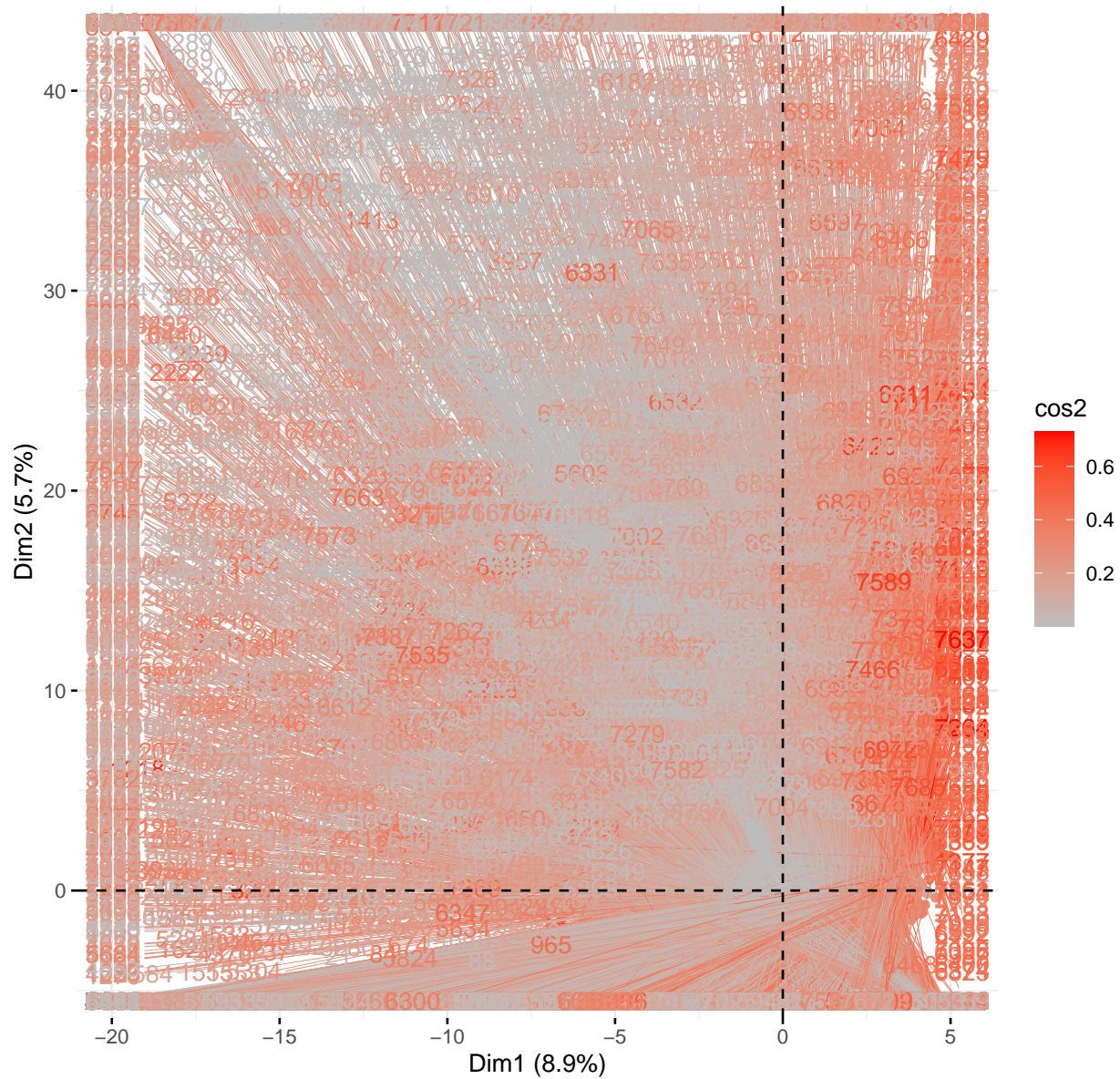


## 8.2 Cercle des corrélations des variables

```
fviz_pca_var(res_acp, col.var = "cos2", gradient.cols = c("grey", "blue"), repel = TRUE) +  
  ggtitle("Cercle des corrélations")
```



## Projection des individus



## 9 9. Résultats détaillés

### 9.1 9.1. Coordonnées des individus

```
coord_individus <- as.data.frame(res_acp$ind$coord)
coord_individus <- cbind(id = rownames(coord_individus), coord_individus)
knitr::kable(head(coord_individus), caption = "Coordonnées des individus (extrait)") %>%
  kableExtra::kable_styling(full_width = FALSE)
```



Table 2: Coordonnées des individus (extrait)

id	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
1	-0.4639700	-0.5195670	0.5101594	-0.7105956	-1.0087238
2	-0.4657207	-0.4884308	0.9440010	0.6633286	-1.4757352
3	0.6198622	-0.8730473	1.0302181	0.6610029	-1.3749322
4	1.0130290	-0.6209831	1.3811877	-0.1206627	-0.4089295
5	0.3081650	0.1360724	1.1604901	-0.0358335	-1.3660937
6	-2.6941145	2.6273728	1.8128393	2.5736232	-1.2238856

## 9.2 9.2. Contributions des variables

```

contrib_variables <- as.data.frame(res_acp$var$contrib)
contrib_variables <- cbind(variable = rownames(contrib_variables), contrib_variables)
knitr::kable(head(contrib_variables), caption = "Contributions des variables (extrait)") %>%
  kableExtra::kable_styling(full_width = FALSE)

```

Table 3: Contributions des variables (extrait)

	variable	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
s10q12a_1	s10q12a_1	0.0132642	0.1170097	0.0867358	3.8614429	13.6459347
s10q13	s10q13	0.0000001	0.0000002	0.0000015	0.0003175	0.0023852
s10q15__0	s10q15__0	0.0000002	0.0000012	0.0000041	0.0006005	0.0039062
s10q15__1	s10q15__1	0.1602752	0.0057632	0.0288940	0.0656282	1.4586342
s10q17a	s10q17a	1.0807747	9.1149327	14.7957609	1.2685970	0.0111496
s10q17b	s10q17b	1.4976836	10.2552546	13.5722228	1.2635125	0.0000575