

# Analyse en Composantes Principales

Serigne Fallou Mbacke NGOM

2023-09-18

L'objectif d'une Analyse en Composantes Principales (ACP) est de résumer et visualiser un tableau de données individus × variables. L'ACP permet d'étudier les ressemblances entre individus du point de vue de l'ensemble des variables et dégager des profils d'individus. Elle permet également de réaliser un bilan des liaisons linéaires entre variables à partir des coefficients de corrélations.

Nous avons utilisé des données fournies par notre professeur de technologie aquacole. Ces données sont issues d'une expérimentation menée dans des fermes aquacoles au Sénégal et sont réparties sur quatre fichiers Excel. Les données contiennent des informations sur l'alimentation (Ratio alimentaire, Taux de conversion alimentaire, ...) et la croissance des poissons (Biomasse, Taux de croissance, Taux de survie, ...) pendant la période d'élevage. Nous voulons dresser un bilan des relations entre les paramètres zootechniques et comparer les performances des trois aliments testés.

## 1. Importer le jeu de données

```
library(readxl)
Aqua_data = read_xlsx("C:/@Bioinformatique mini-projets/ANALYSE AQUACULTURE/JEUX DE_DONNEES_AQUA_CORR.xlsx")
head(Aqua_data)
```

```
## # A tibble: 6 × 15
##   Bacs Date      Aliment `Nombre d'individu contrôlé`
##   <chr> <dtm>      <chr>                <dbl>
## 1 BAC1 2021-09-22 00:00:00 Local_0             20
## 2 BAC4 2021-09-22 00:00:00 Local_0             20
## 3 BAC7 2021-09-22 00:00:00 Local_0             20
## 4 BAC2 2021-09-22 00:00:00 Local 10            20
## 5 BAC5 2021-09-22 00:00:00 Local 10            20
## 6 BAC8 2021-09-22 00:00:00 Local 10            20
## #> 11 more variables: `Biomasse contrôlée (g)` <dbl>,
## # `Poids moyen contrôlé (g)` <dbl>, Ration_g <dbl>,
## # `Gain Pm absolu (g)` <dbl>, `Gain Pm relatif (%)` <dbl>,
## # `Croissance journalière (g/j)` <dbl>,
## # Taux_croissance_spécifique_en_j <dbl>, Taux_de_survie <dbl>, TCA <dbl>,
## # ration <dbl>, ration_ind <dbl>
```

```
Aqua_data$Aliment = factor(Aqua_data$Aliment)
Aqua_data$Bacs = factor(Aqua_data$Bacs)
summary(Aqua_data)
```

```
##      Bacs      Date      Aliment
## BAC3   : 6   Min.   :2021-09-22 00:00:00   Local 10:15
## BAC2   : 5   1st Qu.:2021-09-22 00:00:00   Local 25:15
## BAC4   : 5   Median :2021-11-17 00:00:00   Local_0 :15
## BAC5   : 5   Mean    :2021-11-09 04:48:00
## BAC6   : 5   3rd Qu.:2021-12-16 00:00:00
## BAC7   : 5   Max.    :2021-12-31 00:00:00
## (Other):14
## Nombre d'individu contrôlé Biomasse contrôlée (g) Poids moyen contrôlé (g)
## Min.   :15.00      Min.   : 51.50      Min.   :3.450
## 1st Qu.:16.00      1st Qu.: 70.00      1st Qu.:3.750
## Median :19.00      Median : 79.50      Median :4.469
## Mean    :18.04      Mean    : 81.17      Mean    :4.539
## 3rd Qu.:20.00      3rd Qu.: 86.00      3rd Qu.:5.267
## Max.    :20.00      Max.    :115.80      Max.    :6.095
##
##      Ration_g      Gain Pm absolu (g) Gain Pm relatif (%)
## Min.   :1.360   Min.   :-0.2284   Min.   :-4.979
## 1st Qu.:2.000   1st Qu.: 0.1417   1st Qu.: 2.764
## Median :2.700   Median : 0.7053   Median :19.968
## Mean    :3.101   Mean    : 0.9509   Mean    :27.369
## 3rd Qu.:2.700   3rd Qu.: 1.6088   3rd Qu.:46.633
## Max.    :6.790   Max.    : 2.6447   Max.    :76.659
##
## Croissance journalière (g/j) Taux_croissance_spécifique_en_j Taux_de_survie
## Min.   :-0.01632   Min.   :-0.3470   Min.   : 0.00
## 1st Qu.: 0.01012   1st Qu.: 0.1948   1st Qu.: 80.00
## Median : 0.04325   Median : 1.0072   Median : 95.00
## Mean    : 0.04864   Mean    : 1.1506   Mean    : 75.94
## 3rd Qu.: 0.05620   3rd Qu.: 1.3903   3rd Qu.:100.00
## Max.    : 0.18891   Max.    : 4.0647   Max.    :100.00
##
##      TCA      ration      ration_ind
## Min.   :-7.5909   Min.   :0.000   Min.   :0.00000
## 1st Qu.: 0.6394   1st Qu.:1.326   1st Qu.:0.07524
## Median : 2.3883   Median :2.145   Median :0.12000
## Mean    : 3.0010   Mean    :2.155   Mean    :0.12368
## 3rd Qu.: 3.9194   3rd Qu.:2.691   3rd Qu.:0.13606
## Max.    :16.8600   Max.    :5.260   Max.    :0.35066
##
```

## Analyse en Composantes Principales (ACP):

```
library(FactoMineR)
res.pca <- PCA(Aqua_data[4:15], quali.sup=3, graph = FALSE)
```

- Variance expliquée:

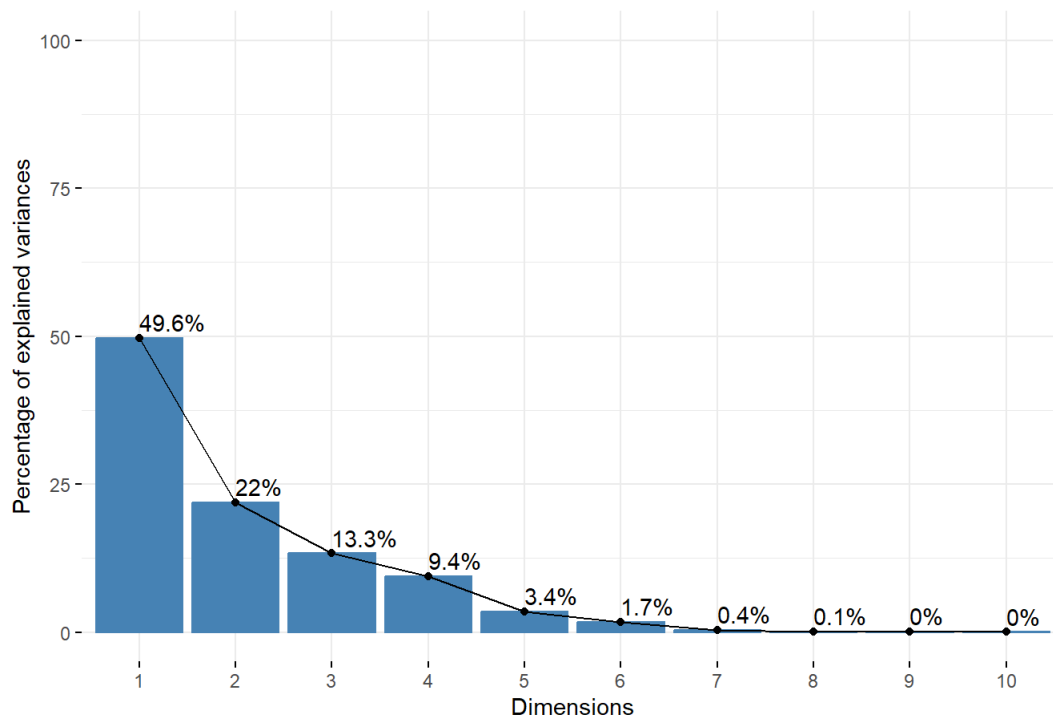
```
library("factoextra")
```

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

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

```
fviz_eig(res.pca, addlabels = TRUE, ylim = c(0, 100))
```

Scree plot



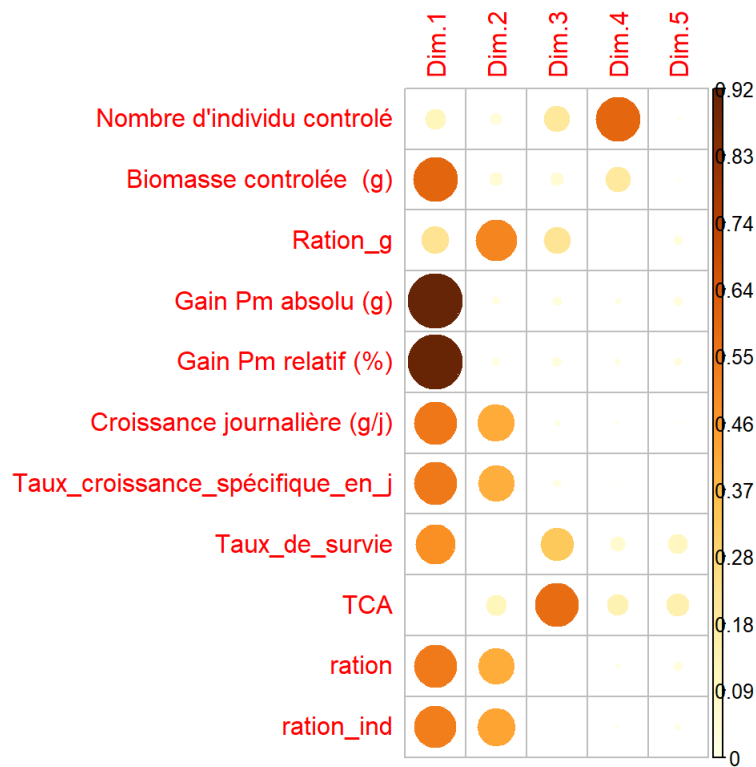
- Cercle de corrélation:

```
library("corrplot")
```

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

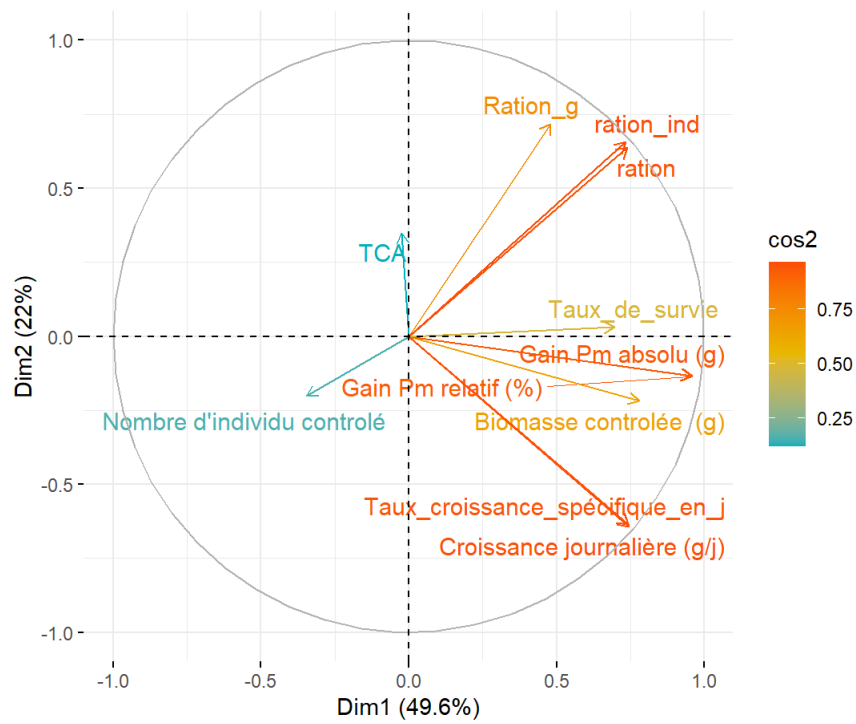
```
## corrplot 0.92 loaded
```

```
var <- get_pca_var(res.pca)
corrplot(var$cos2, is.corr=FALSE)
```



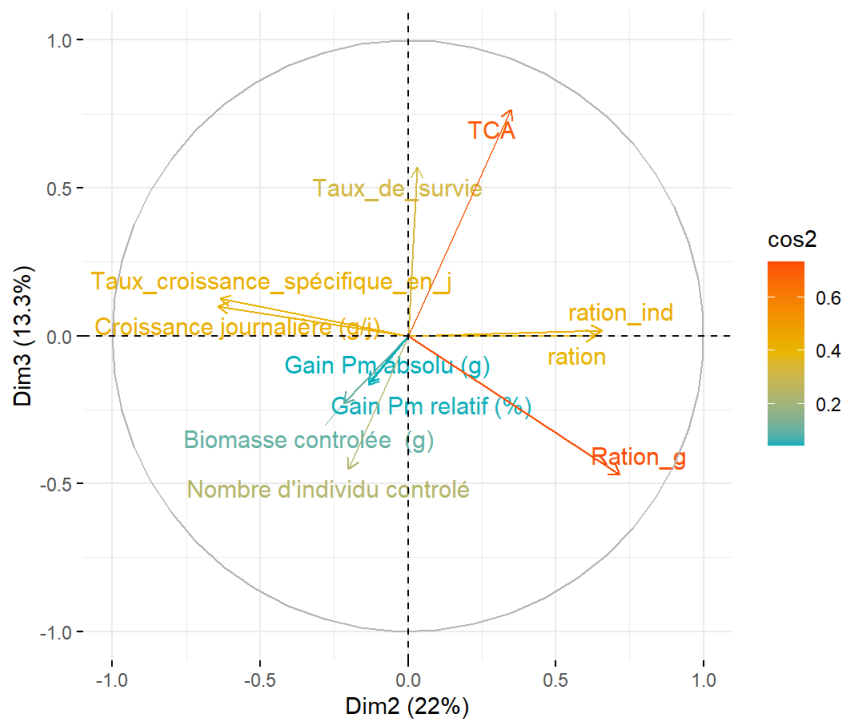
```
fviz_pca_var(res.pca, col.var = "cos2", gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE )
```

Variables - PCA



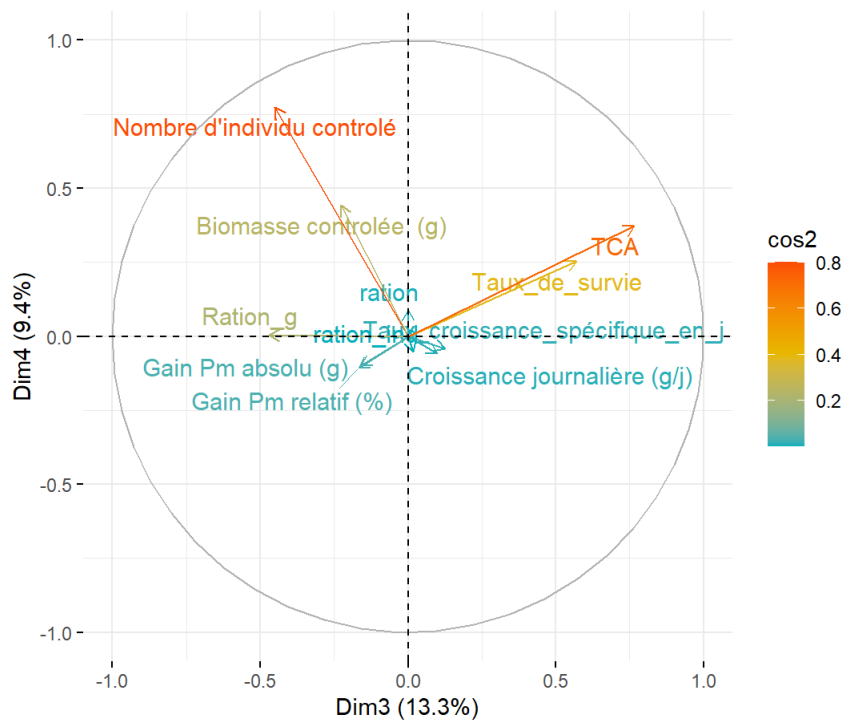
```
fviz_pca_var(res.pca, col.var = "cos2", gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE, axes = 2:3)
```

Variables - PCA



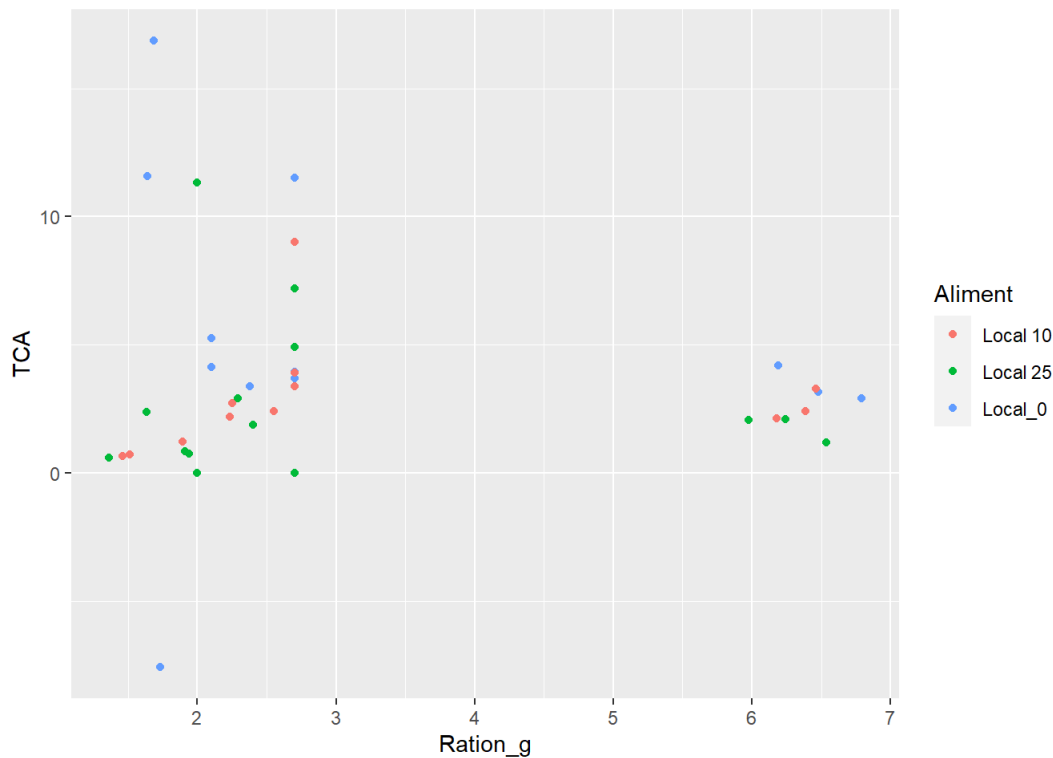
```
fviz_pca_var(res.pca, col.var = "cos2", gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE, axes = 3:4)
```

Variables - PCA

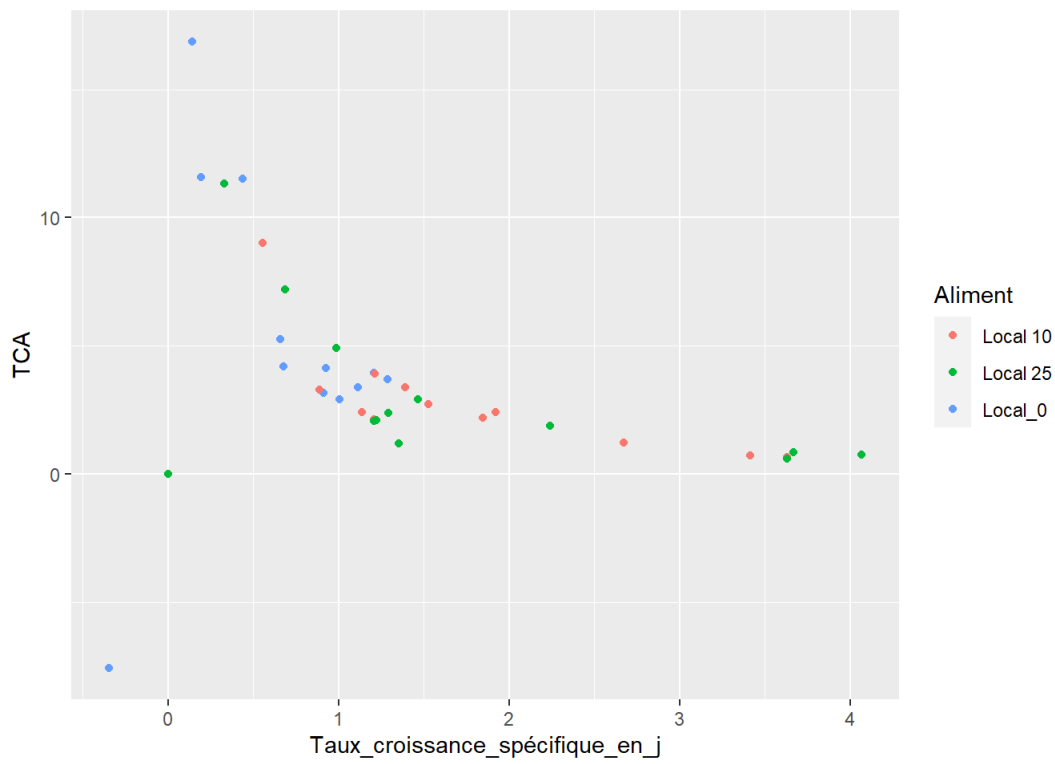


Comparaison des performances des trois (3) aliments:

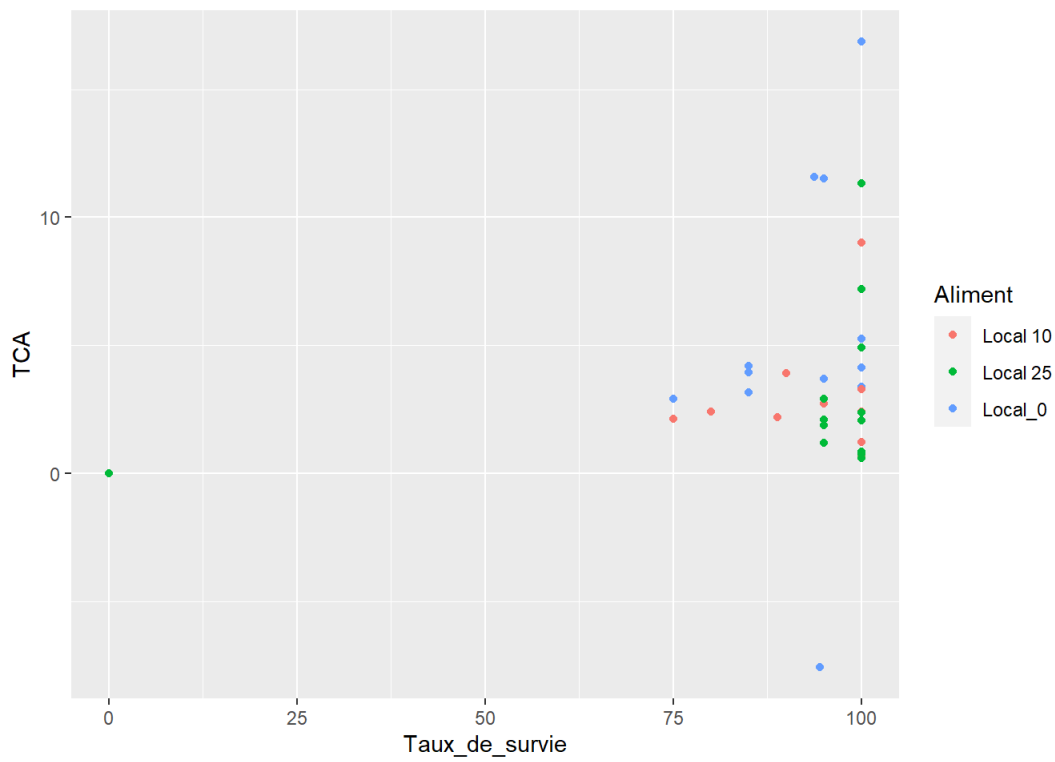
```
library(ggplot2)
ggplot(Aqua_data) + aes(x=Ration_g, y=TCA, color=Aliment) + geom_point()
```



```
ggplot(Aqua_data) + aes(x=Taux_croissance_spécifique_en_j, y=TCA, color=Aliment) + geom_point()
```



```
ggplot(Aqua_data) + aes(x=Taux_de_survie, y=TCA,color=Aliment) + geom_point()
```



```
ggplot(Aqua_data) + aes(x=Aliment, y=TCA) + geom_point()
```

