

Rapport de projet Analyse de données & Eléments de modélisation statistique

Xiaoya Wang, Mickael Song, Yessine Jmal, Florian Grivet

2022-11-18

Analyse du jeu de données

Statistiques descriptives et préparation du jeu de données

```
# Chargement des données
data = read.delim("Data_Etudiants_2023.txt",header=TRUE, sep=";")

datapy = pd.read_csv("Data_Etudiants_2023.txt", sep=";")

# Affichage des premières lignes des données
datapy.head(n=6)

##      T1_1h_R1  T1_2h_R1  T1_3h_R1  ...  T3_4h_R2  T3_5h_R2  T3_6h_R2
## G1  0.167009  0.682881 -0.180178  ...  2.525922  3.268041  3.089573
## G2  0.187621  0.815365 -0.045735  ...  2.867855  2.622922  2.051829
## G3 -0.052076 -0.029437  0.261409  ... -3.886458 -3.129329 -3.110750
## G4 -0.230785 -0.753947 -0.244923  ...  1.155446  1.890564  2.442747
## G5 -0.205473 -0.688821 -0.181084  ... -2.849342 -2.938259 -3.390693
## G6 -0.619510 -0.856138 -0.021084  ... -2.483709 -2.461097 -2.970514
##
## [6 rows x 36 columns]
```

Le jeu de données contient 1615 individus et 36 variables, toutes quantitatives.

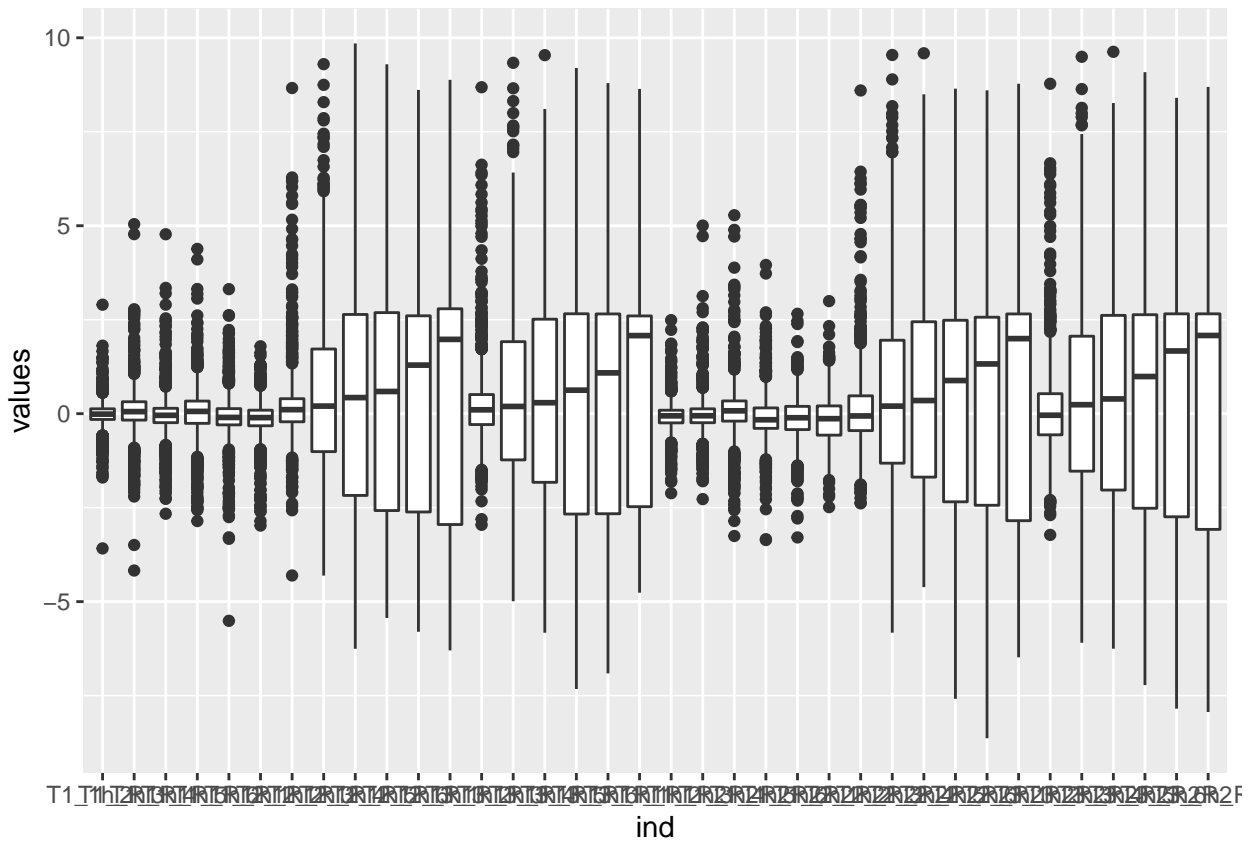
Les attributs du jeu de données sont : T1_1h_R1, T1_2h_R1, T1_3h_R1, T1_4h_R1, T1_5h_R1, T1_6h_R1, T2_1h_R1, T2_2h_R1, T2_3h_R1, T2_4h_R1, T2_5h_R1, T2_6h_R1, T3_1h_R1, T3_2h_R1, T3_3h_R1, T3_4h_R1, T3_5h_R1, T3_6h_R1, T1_1h_R2, T1_2h_R2, T1_3h_R2, T1_4h_R2, T1_5h_R2, T1_6h_R2, T2_1h_R2, T2_2h_R2, T2_3h_R2, T2_4h_R2, T2_5h_R2, T2_6h_R2, T3_1h_R2, T3_2h_R2, T3_3h_R2, T3_4h_R2, T3_5h_R2, T3_6h_R2

```
## [1] "Quelques statistiques sur les variables : "
```

```
##      T1_1h_R1      T1_2h_R1      T1_3h_R1      T1_4h_R1
## Min.      : -3.58436  Min.      : -4.17248  Min.      : -2.66211  Min.      : -2.85622
## 1st Qu.: -0.15360  1st Qu.: -0.17073  1st Qu.: -0.23907  1st Qu.: -0.25748
## Median : -0.01163  Median :  0.05521  Median : -0.04364  Median :  0.06043
## Mean    : -0.02221  Mean     :  0.06334  Mean     : -0.04633  Mean     :  0.01804
## 3rd Qu.:  0.12923  3rd Qu.:  0.31726  3rd Qu.:  0.14296  3rd Qu.:  0.33552
## Max.    :  2.89949  Max.     :  5.04213  Max.     :  4.77379  Max.     :  4.38339
##      T1_5h_R1      T1_6h_R1      T2_1h_R1      T2_2h_R1
## Min.      : -5.51058  Min.      : -2.97588  Min.      : -4.3040  Min.      : -4.3077
## 1st Qu.: -0.29766  1st Qu.: -0.32203  1st Qu.: -0.2155  1st Qu.: -1.0092
## Median : -0.09962  Median : -0.10486  Median :  0.1074  Median :  0.2015
```

## Mean : -0.08840	Mean : -0.12726	Mean : 0.1717	Mean : 0.4455
## 3rd Qu.: 0.13438	3rd Qu.: 0.09348	3rd Qu.: 0.3983	3rd Qu.: 1.7205
## Max. : 3.31269	Max. : 1.79268	Max. : 8.6635	Max. : 9.2985
## T2_3h_R1	T2_4h_R1	T2_5h_R1	T2_6h_R1
## Min. : -6.2538	Min. : -5.4355	Min. : -5.8017	Min. : -6.2970
## 1st Qu.: -2.1747	1st Qu.: -2.5772	1st Qu.: -2.6147	1st Qu.: -2.9487
## Median : 0.4284	Median : 0.5906	Median : 1.2921	Median : 1.9764
## Mean : 0.2853	Mean : 0.1813	Mean : 0.1753	Mean : 0.1602
## 3rd Qu.: 2.6392	3rd Qu.: 2.6884	3rd Qu.: 2.6041	3rd Qu.: 2.7915
## Max. : 9.8504	Max. : 9.2951	Max. : 8.6135	Max. : 8.8815
## T3_1h_R1	T3_2h_R1	T3_3h_R1	T3_4h_R1
## Min. : -2.9561	Min. : -4.9884	Min. : -5.8280	Min. : -7.32431
## 1st Qu.: -0.2891	1st Qu.: -1.2295	1st Qu.: -1.8260	1st Qu.: -2.67037
## Median : 0.1030	Median : 0.1926	Median : 0.2916	Median : 0.62522
## Mean : 0.1908	Mean : 0.4227	Mean : 0.4268	Mean : 0.07642
## 3rd Qu.: 0.5086	3rd Qu.: 1.9157	3rd Qu.: 2.5138	3rd Qu.: 2.65780
## Max. : 8.6849	Max. : 9.3343	Max. : 9.5376	Max. : 9.19767
## T3_5h_R1	T3_6h_R1	T1_1h_R2	T1_2h_R2
## Min. : -6.910	Min. : -4.7625	Min. : -2.11580	Min. : -2.27243
## 1st Qu.: -2.661	1st Qu.: -2.4740	1st Qu.: -0.24442	1st Qu.: -0.24039
## Median : 1.086	Median : 2.0769	Median : -0.05733	Median : -0.05527
## Mean : 0.128	Mean : 0.3091	Mean : -0.06376	Mean : -0.03493
## 3rd Qu.: 2.654	3rd Qu.: 2.6007	3rd Qu.: 0.09170	3rd Qu.: 0.12912
## Max. : 8.794	Max. : 8.6398	Max. : 2.48886	Max. : 5.00322
## T1_3h_R2	T1_4h_R2	T1_5h_R2	T1_6h_R2
## Min. : -3.25393	Min. : -3.3611	Min. : -3.2919	Min. : -2.4863
## 1st Qu.: -0.19771	1st Qu.: -0.3905	1st Qu.: -0.4234	1st Qu.: -0.5708
## Median : 0.07646	Median : -0.1618	Median : -0.1079	Median : -0.1329
## Mean : 0.06790	Mean : -0.1024	Mean : -0.1212	Mean : -0.1646
## 3rd Qu.: 0.33953	3rd Qu.: 0.1533	3rd Qu.: 0.1948	3rd Qu.: 0.2066
## Max. : 5.27981	Max. : 3.9568	Max. : 2.6578	Max. : 2.9948
## T2_1h_R2	T2_2h_R2	T2_3h_R2	T2_4h_R2
## Min. : -2.38587	Min. : -5.8266	Min. : -4.6135	Min. : -7.5854
## 1st Qu.: -0.44996	1st Qu.: -1.3166	1st Qu.: -1.6876	1st Qu.: -2.3444
## Median : -0.05977	Median : 0.2010	Median : 0.3500	Median : 0.8823
## Mean : 0.10076	Mean : 0.3705	Mean : 0.4685	Mean : 0.1479
## 3rd Qu.: 0.47458	3rd Qu.: 1.9544	3rd Qu.: 2.4457	3rd Qu.: 2.4853
## Max. : 8.59820	Max. : 9.5424	Max. : 9.5877	Max. : 8.6466
## T2_5h_R2	T2_6h_R2	T3_1h_R2	T3_2h_R2
## Min. : -8.6345	Min. : -6.481	Min. : -3.22436	Min. : -6.0944
## 1st Qu.: -2.4368	1st Qu.: -2.847	1st Qu.: -0.56228	1st Qu.: -1.5293
## Median : 1.3234	Median : 1.997	Median : -0.04190	Median : 0.2388
## Mean : 0.1772	Mean : 0.148	Mean : 0.07906	Mean : 0.3226
## 3rd Qu.: 2.5660	3rd Qu.: 2.654	3rd Qu.: 0.53160	3rd Qu.: 2.0612
## Max. : 8.6021	Max. : 8.778	Max. : 8.77729	Max. : 9.4943
## T3_3h_R2	T3_4h_R2	T3_5h_R2	T3_6h_R2
## Min. : -6.2521	Min. : -7.2215	Min. : -7.84771	Min. : -7.936715
## 1st Qu.: -2.0317	1st Qu.: -2.5160	1st Qu.: -2.74288	1st Qu.: -3.077880
## Median : 0.3928	Median : 0.9884	Median : 1.66822	Median : 2.080562
## Mean : 0.3293	Mean : 0.1558	Mean : 0.09246	Mean : 0.009932
## 3rd Qu.: 2.6152	3rd Qu.: 2.6327	3rd Qu.: 2.65638	3rd Qu.: 2.656590
## Max. : 9.6268	Max. : 9.0856	Max. : 8.40304	Max. : 8.692848

```
ggplot(stack(data), aes(x = ind, y = values))+
  geom_boxplot()
```

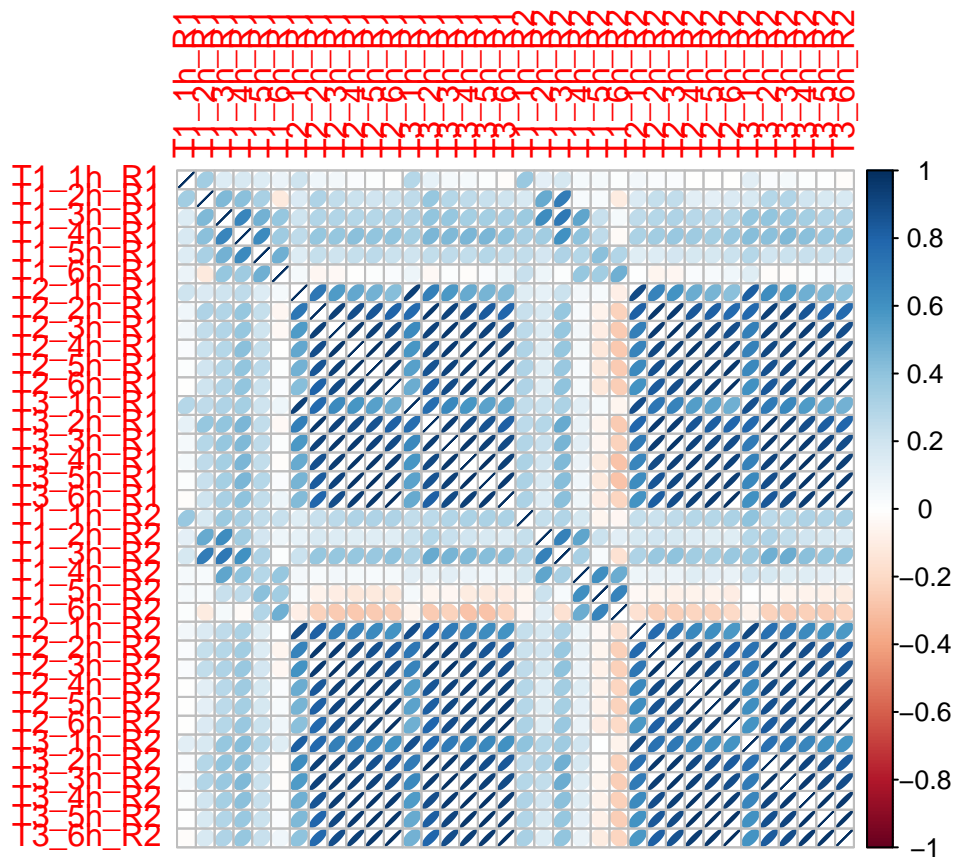


On remarque que les traitements 2 et 3 ont des boxplots similaires que ça soit pour le réplicat 1 ou pour le réplicat 2. Cela signifie que ces deux traitements ont des résultats similaires.

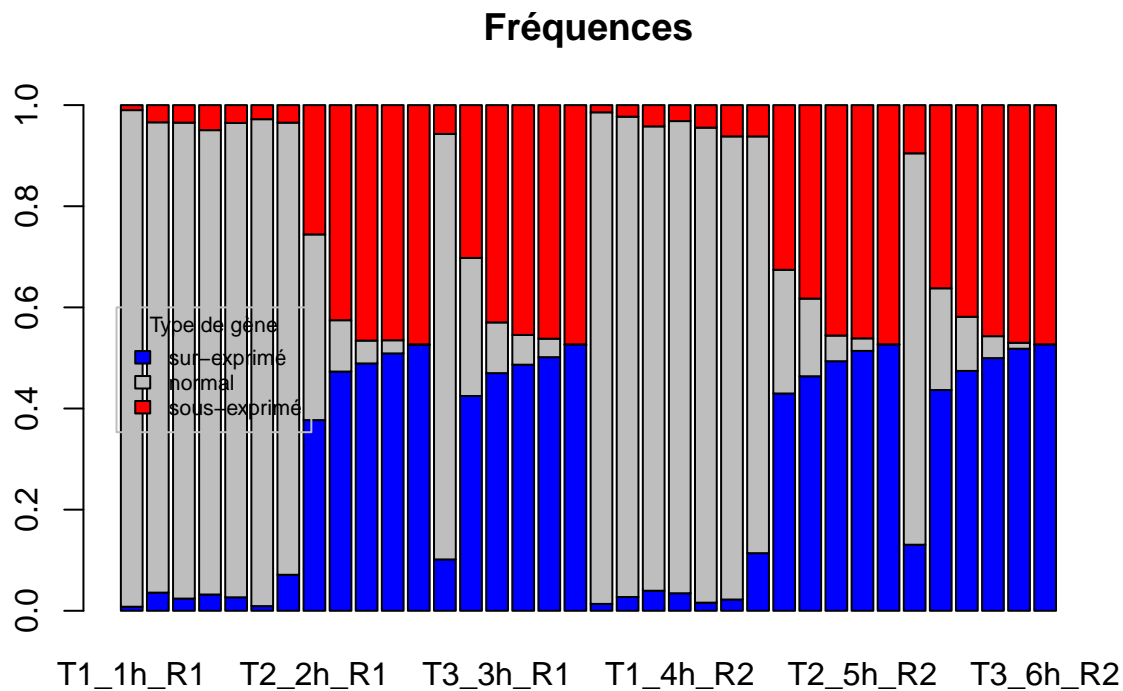
On note également une forte variabilité entre les individus.

Le traitement 1 est quant à lui beaucoup plus concentré et centré en 0. Le traitement semble donc ne pas avoir d'effet sur les gènes.

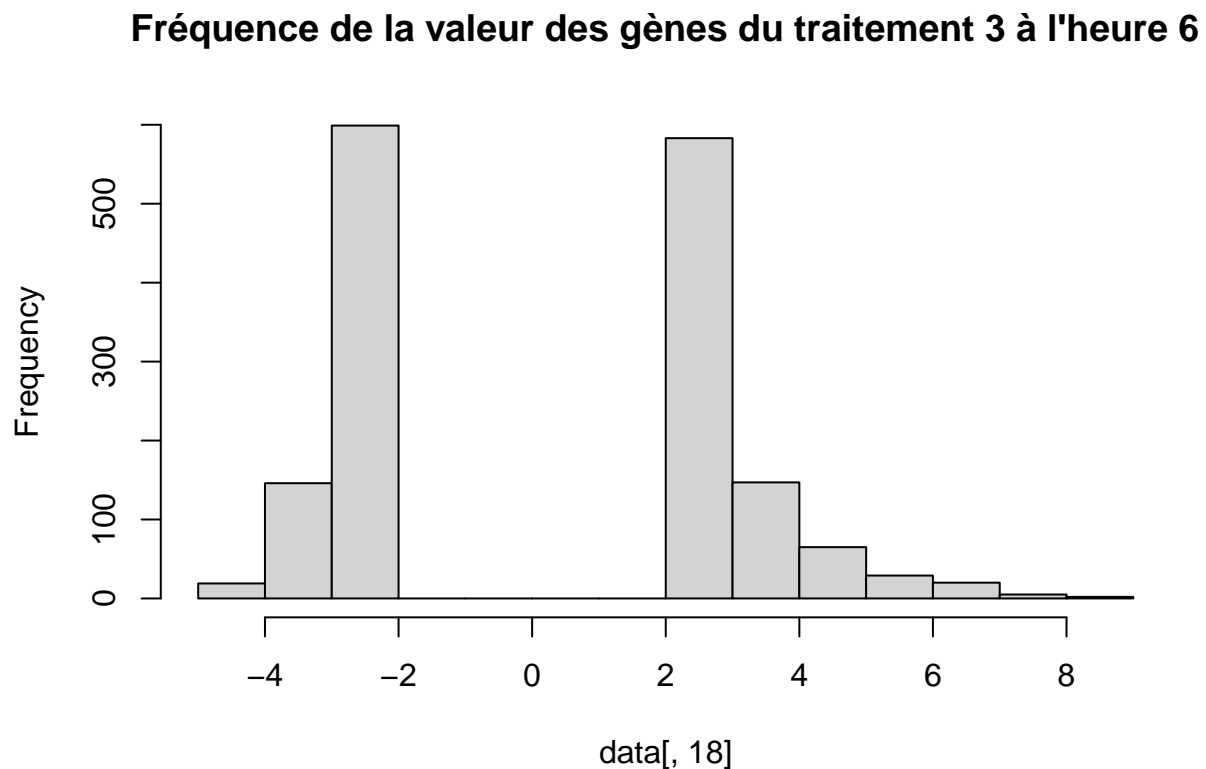
```
corrplot(cor(data), method="ellipse")
```



```
BP = array(rep(rep(0,3),36), dim=c(3,36))
BP[1,] = apply(data>1, 2, sum)
BP[2,] = apply(data<1&data>(-1), 2, sum)
BP[3,] = apply(data<(-1), 2, sum)
BP = BP/nrow(data)
barplot(BP, main="Fréquences", col=c("blue", "grey", "red"), names.arg=c(attributes(data)$names))
legend(0,0.6, legend=c("sur-exprimé", "normal", "sous-exprimé"), title="Type de gène", box.col="grey",
```



```
hist(data[,18], main="Fréquence de la valeur des gènes du traitement 3 à l'heure 6")
```



On remarque qu'à l'heure 6 (la dernière) du traitement 3, tous les gènes sont soit très sur-exprimé (valeurs ≥ 2), soit très sous-exprimé (valeurs ≤ 2). C'est comme ça que les gènes du jeu de données ont été choisis.

Analyse en composante principale

```
res.pca = PCA(data, ncp=15, graph=F)
summary(res.pca)
```

```
##
## Call:
## PCA(X = data, ncp = 15, graph = F)
##
## Eigenvalues
##
```

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7
## Variance	21.277	4.003	2.290	2.111	1.344	1.007	0.923
## % of var.	59.103	11.120	6.362	5.863	3.733	2.797	2.565
## Cumulative % of var.	59.103	70.224	76.585	82.449	86.182	88.979	91.544

```
##
```

	Dim.8	Dim.9	Dim.10	Dim.11	Dim.12	Dim.13	Dim.14
## Variance	0.526	0.488	0.418	0.279	0.242	0.213	0.157
## % of var.	1.460	1.354	1.162	0.775	0.673	0.592	0.435
## Cumulative % of var.	93.004	94.359	95.521	96.296	96.969	97.561	97.996

```
##
```

	Dim.15	Dim.16	Dim.17	Dim.18	Dim.19	Dim.20	Dim.21
## Variance	0.143	0.101	0.089	0.065	0.062	0.048	0.044
## % of var.	0.398	0.280	0.249	0.180	0.173	0.133	0.122
## Cumulative % of var.	98.394	98.674	98.922	99.103	99.276	99.409	99.531

```
##
```

	Dim.22	Dim.23	Dim.24	Dim.25	Dim.26	Dim.27	Dim.28
## Variance	0.035	0.022	0.018	0.014	0.012	0.011	0.009
## % of var.	0.096	0.062	0.051	0.039	0.033	0.030	0.025
## Cumulative % of var.	99.627	99.689	99.740	99.779	99.812	99.841	99.866

```
##
```

	Dim.29	Dim.30	Dim.31	Dim.32	Dim.33	Dim.34	Dim.35
## Variance	0.009	0.008	0.007	0.006	0.005	0.005	0.005
## % of var.	0.024	0.022	0.019	0.017	0.015	0.014	0.013
## Cumulative % of var.	99.890	99.912	99.931	99.948	99.963	99.976	99.990

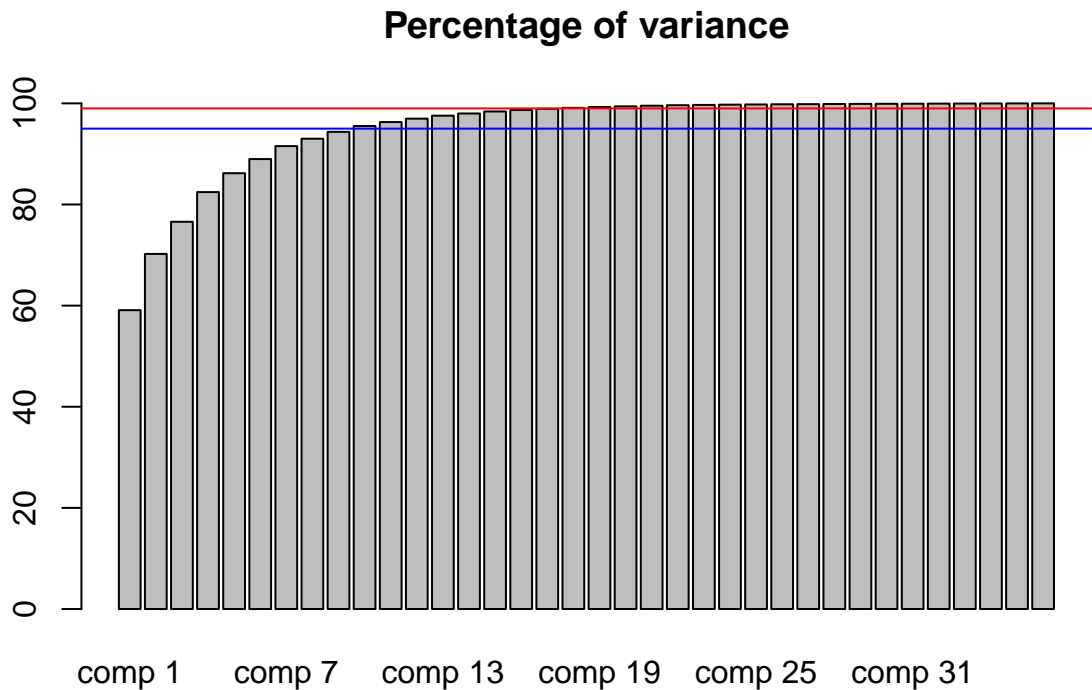
```
##
## Variance
## % of var.
## Cumulative % of var. 100.000
##
## Individuals (the 10 first)
##
```

	Dist	Dim.1	ctr	cos2	Dim.2	ctr	cos2	Dim.3	ctr
## G1	4.129	2.235	0.015	0.293	0.529	0.004	0.016	1.700	0.078
## G2	5.043	3.864	0.043	0.587	0.702	0.008	0.019	0.109	0.000
## G3	4.946	-4.406	0.056	0.793	0.690	0.007	0.019	-0.860	0.020
## G4	2.922	1.148	0.004	0.154	-1.498	0.035	0.263	0.810	0.018
## G5	6.113	-5.241	0.080	0.735	1.489	0.034	0.059	1.411	0.054
## G6	6.038	-4.864	0.069	0.649	0.881	0.012	0.021	1.882	0.096
## G7	3.737	3.095	0.028	0.686	-0.749	0.009	0.040	0.310	0.003
## G8	3.127	-1.675	0.008	0.287	0.549	0.005	0.031	-1.772	0.085
## G9	6.098	-4.035	0.047	0.438	-0.179	0.000	0.001	-1.974	0.105
## G10	3.578	2.780	0.022	0.604	-0.655	0.007	0.034	-0.227	0.001

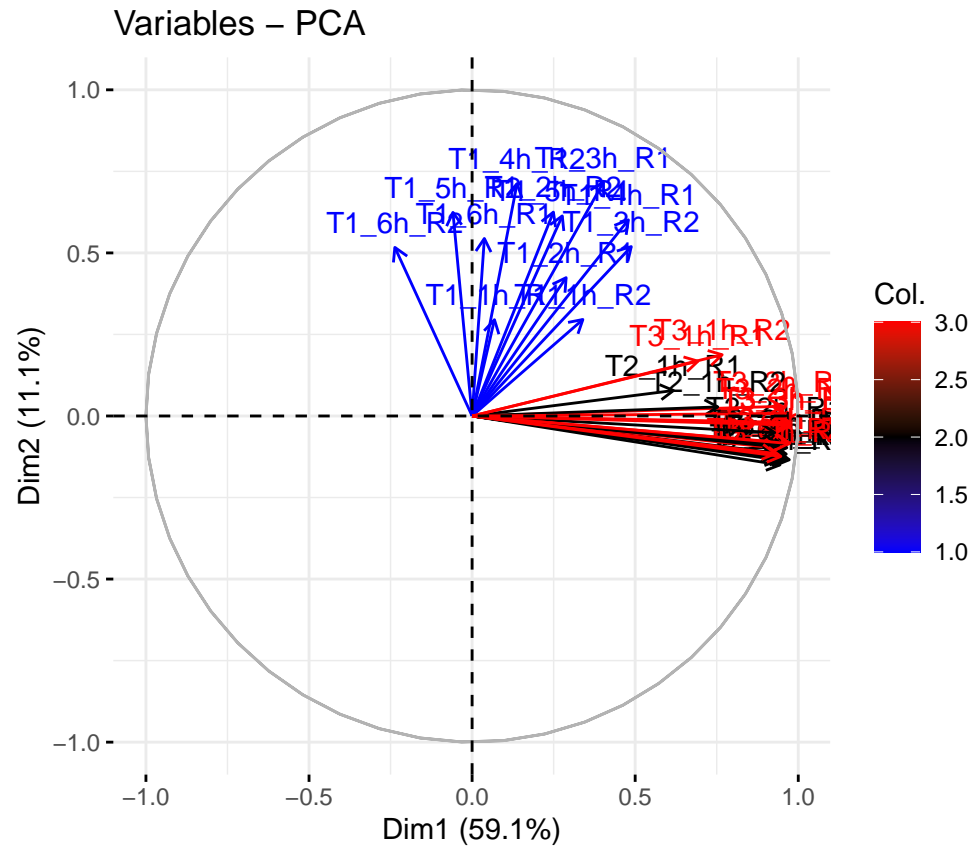
```
##
## cos2
## G1
## G2
## G3
## G4
## G5
```

```
## G6      0.097 |
## G7      0.007 |
## G8      0.321 |
## G9      0.105 |
## G10     0.004 |
##
## Variables (the 10 first)
##          Dim.1   ctr   cos2   Dim.2   ctr   cos2   Dim.3   ctr   cos2
## T1_1h_R1 | 0.069 0.022 0.005 | 0.295 2.176 0.087 | -0.326 4.631 0.106 |
## T1_2h_R1 | 0.288 0.391 0.083 | 0.424 4.494 0.180 | -0.247 2.674 0.061 |
## T1_3h_R1 | 0.403 0.762 0.162 | 0.717 12.828 0.514 | 0.049 0.105 0.002 |
## T1_4h_R1 | 0.479 1.079 0.230 | 0.601 9.030 0.362 | 0.093 0.377 0.009 |
## T1_5h_R1 | 0.277 0.359 0.076 | 0.613 9.398 0.376 | 0.206 1.849 0.042 |
## T1_6h_R1 | 0.038 0.007 0.001 | 0.544 7.397 0.296 | 0.306 4.077 0.093 |
## T2_1h_R1 | 0.618 1.792 0.381 | 0.078 0.152 0.006 | -0.640 17.872 0.409 |
## T2_2h_R1 | 0.928 4.048 0.861 | -0.051 0.065 0.003 | -0.237 2.444 0.056 |
## T2_3h_R1 | 0.963 4.360 0.928 | -0.117 0.341 0.014 | -0.001 0.000 0.000 |
## T2_4h_R1 | 0.973 4.445 0.946 | -0.134 0.449 0.018 | 0.107 0.501 0.011 |

barplot(cumsum(res.pca$eig[, "percentage of variance"]), main="Percentage of variance")
abline(h=95, col="blue")
abline(h=99, col="red")
```

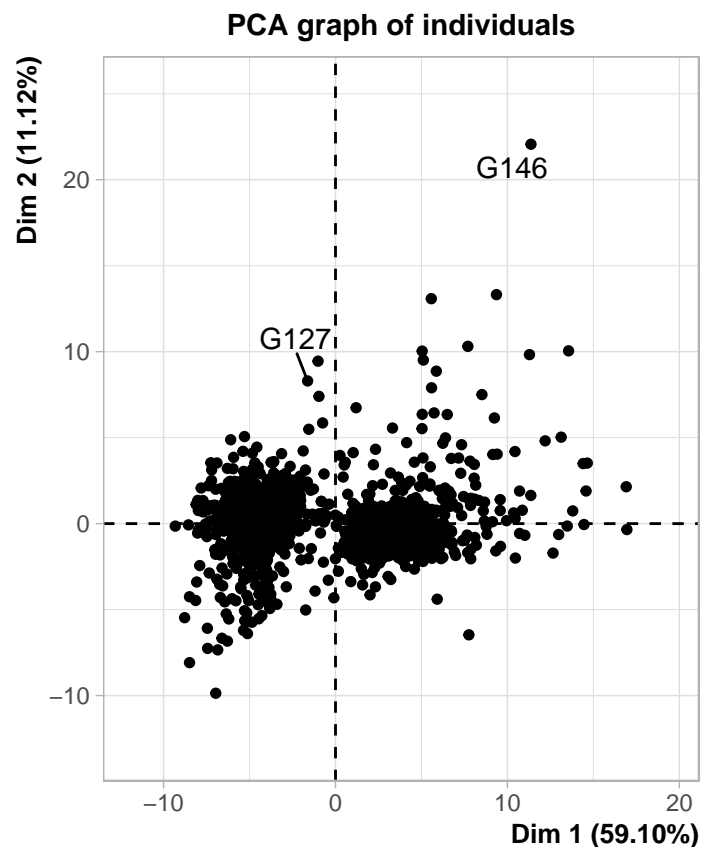


```
col_trait = rep(rep(c(1,2,3), each=6),2)
fviz_pca_var(res.pca, col.var=col_trait) + scale_color_gradient2(low="blue", mid="black", high="red", m
```



```
plot(res.pca, axes=c(1, 2), choix = "ind", autoLab="yes")
```

```
## Warning: ggrepel: 1613 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```

```
print("Pourcentage des variables dans la construction des dimensions 1 et 2 :")
```

```
## [1] "Pourcentage des variables dans la construction des dimensions 1 et 2 :"  
res.pca$var$cor[,1]/sum(res.pca$var$cor[,1])*100
```

```
##   T1_1h_R1  T1_2h_R1  T1_3h_R1  T1_4h_R1  T1_5h_R1  T1_6h_R1  T2_1h_R1  
## 0.2813629 1.1822835 1.6516723 1.9654421 1.1341337 0.1542264 2.5326636  
##   T2_2h_R1  T2_3h_R1  T2_4h_R1  T2_5h_R1  T2_6h_R1  T3_1h_R1  T3_2h_R1  
## 3.8059162 3.9502202 3.9885739 3.9446705 3.8518276 2.8403860 3.8976504  
##   T3_3h_R1  T3_4h_R1  T3_5h_R1  T3_6h_R1  T1_1h_R2  T1_2h_R2  T1_3h_R2  
## 4.0030605 3.9964707 3.9446329 3.8467296 1.3930465 1.0223162 2.0019173  
##   T1_4h_R2  T1_5h_R2  T1_6h_R2  T2_1h_R2  T2_2h_R2  T2_3h_R2  T2_4h_R2  
## 0.5649958 -0.2408849 -0.9711981 3.0894597 3.8566617 3.9582085 3.9156365  
##   T2_5h_R2  T2_6h_R2  T3_1h_R2  T3_2h_R2  T3_3h_R2  T3_4h_R2  T3_5h_R2  
## 3.8709605 3.8076235 3.1447154 3.9334063 4.0028388 3.9831928 3.8820269  
##   T3_6h_R2  
## 3.8131537
```

```
res.pca$var$cor[,2]/sum(res.pca$var$cor[,2])*100
```

```
##   T1_1h_R1  T1_2h_R1  T1_3h_R1  T1_4h_R1  T1_5h_R1  T1_6h_R1  T2_1h_R1  
## 5.7007552 8.1925907 13.8423147 11.6138771 11.8481927 10.5114045 1.5091520  
##   T2_2h_R1  T2_3h_R1  T2_4h_R1  T2_5h_R1  T2_6h_R1  T3_1h_R1  T3_2h_R1  
## -0.9857786 -2.2581646 -2.5907091 -2.5941559 -2.6034758 3.2450777 0.5755238  
##   T3_3h_R1  T3_4h_R1  T3_5h_R1  T3_6h_R1  T1_1h_R2  T1_2h_R2  T1_3h_R2  
## -0.3833647 -1.6144701 -1.8136238 -1.9069498 5.7142531 12.0738909 10.0297336  
##   T1_4h_R2  T1_5h_R2  T1_6h_R2  T2_1h_R2  T2_2h_R2  T2_3h_R2  T2_4h_R2  
## 13.7569368 12.0576906 9.9851907 0.5377747 -1.7126095 -1.8822147 -2.5712756
```

```
## T2_5h_R2 T2_6h_R2 T3_1h_R2 T3_2h_R2 T3_3h_R2 T3_4h_R2 T3_5h_R2
## -2.8978047 -2.3702032 3.6267789 0.1487847 -0.5379820 -1.6295619 -2.3941936
## T3_6h_R2
## -2.2233845
```

L'axe 1 nous dit si un gène réagit au traitement 2 ou 3. Si le gène réagit fortement à l'un de ces traitements, il se retrouve sur un côté du graphique (si le gène devient très sous-exprimé ou très sur-exprimé) et s'il ne réagit pas beaucoup à l'un de ces traitement il se trouve au milieu.

L'axe 2 nous dit si un gène réagit au traitement 1. Si le gène réagit fortement à ce traitement, il se retrouve en haut s'il devient sur-exprimé, en bas s'il devient sous-exprimé et s'il ne réagit pas beaucoup, il se trouve au milieu.

```
data_transpose = t(data)
res.pca.transpose = PCA(data_transpose, ncp=15, graph=F)
res.pca.transpose$eig
```

```
## eigenvalue percentage of variance cumulative percentage of variance
## comp 1 1302.8360199 80.67096099 80.67096
## comp 2 129.6311241 8.02669499 88.69766
## comp 3 32.7329368 2.02680723 90.72446
## comp 4 27.9050439 1.72786649 92.45233
## comp 5 22.9722824 1.42243235 93.87476
## comp 6 17.6298603 1.09163222 94.96639
## comp 7 12.0154370 0.74398991 95.71038
## comp 8 7.9108585 0.48983644 96.20022
## comp 9 6.7056534 0.41521074 96.61543
## comp 10 6.1028864 0.37788770 96.99332
## comp 11 4.6580858 0.28842637 97.28175
## comp 12 4.4147655 0.27336009 97.55511
## comp 13 4.0399164 0.25014962 97.80526
## comp 14 3.6790419 0.22780445 98.03306
## comp 15 3.0911972 0.19140540 98.22446
## comp 16 2.8740710 0.17796105 98.40243
## comp 17 2.3737595 0.14698201 98.54941
## comp 18 2.1743227 0.13463298 98.68404
## comp 19 2.0279315 0.12556851 98.80961
## comp 20 1.9353306 0.11983471 98.92944
## comp 21 1.7536513 0.10858522 99.03803
## comp 22 1.6183733 0.10020888 99.13824
## comp 23 1.4436317 0.08938896 99.22763
## comp 24 1.4219137 0.08804419 99.31567
## comp 25 1.3112718 0.08119330 99.39686
## comp 26 1.1986874 0.07422213 99.47109
## comp 27 1.1614544 0.07191668 99.54300
## comp 28 1.0939285 0.06773551 99.61074
## comp 29 1.0379893 0.06427178 99.67501
## comp 30 1.0017843 0.06202999 99.73704
## comp 31 0.9669957 0.05987590 99.79692
## comp 32 0.9006829 0.05576984 99.85269
## comp 33 0.8309085 0.05144944 99.90414
## comp 34 0.7861668 0.04867906 99.95282
## comp 35 0.7620357 0.04718487 100.00000
```

```
summary(res.pca.transpose)
```

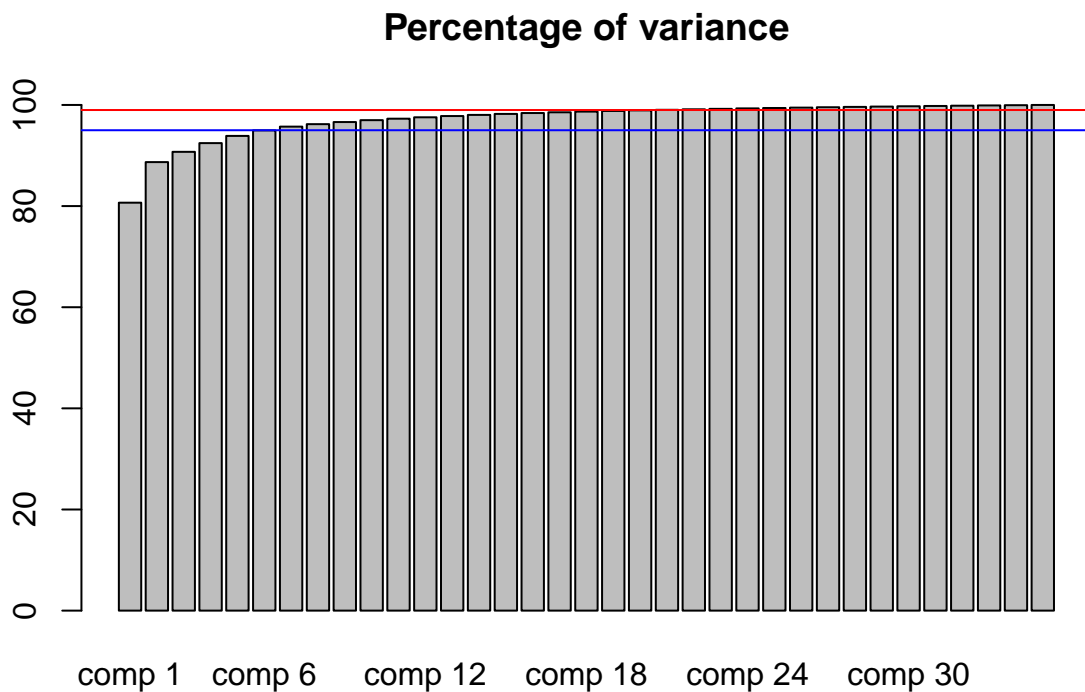
```

##
## Call:
## PCA(X = data_transpose, ncp = 15, graph = F)
##
##
## Eigenvalues
##
##          Dim.1   Dim.2   Dim.3   Dim.4   Dim.5   Dim.6
## Variance    1302.836  129.631   32.733   27.905   22.972   17.630
## % of var.      80.671    8.027    2.027    1.728    1.422    1.092
## Cumulative % of var. 80.671   88.698   90.724   92.452   93.875   94.966
##
##          Dim.7   Dim.8   Dim.9   Dim.10   Dim.11   Dim.12
## Variance     12.015    7.911    6.706    6.103    4.658    4.415
## % of var.      0.744    0.490    0.415    0.378    0.288    0.273
## Cumulative % of var. 95.710   96.200   96.615   96.993   97.282   97.555
##
##          Dim.13   Dim.14   Dim.15   Dim.16   Dim.17   Dim.18
## Variance       4.040    3.679    3.091    2.874    2.374    2.174
## % of var.      0.250    0.228    0.191    0.178    0.147    0.135
## Cumulative % of var. 97.805   98.033   98.224   98.402   98.549   98.684
##
##          Dim.19   Dim.20   Dim.21   Dim.22   Dim.23   Dim.24
## Variance       2.028    1.935    1.754    1.618    1.444    1.422
## % of var.      0.126    0.120    0.109    0.100    0.089    0.088
## Cumulative % of var. 98.810   98.929   99.038   99.138   99.228   99.316
##
##          Dim.25   Dim.26   Dim.27   Dim.28   Dim.29   Dim.30
## Variance       1.311    1.199    1.161    1.094    1.038    1.002
## % of var.      0.081    0.074    0.072    0.068    0.064    0.062
## Cumulative % of var. 99.397   99.471   99.543   99.611   99.675   99.737
##
##          Dim.31   Dim.32   Dim.33   Dim.34   Dim.35
## Variance       0.967    0.901    0.831    0.786    0.762
## % of var.      0.060    0.056    0.051    0.049    0.047
## Cumulative % of var. 99.797   99.853   99.904   99.953  100.000
##
## Individuals (the 10 first)
##
##          Dist   Dim.1   ctr   cos2   Dim.2   ctr   cos2
## T1_1h_R1 | 44.379 | -42.918  3.927  0.935 | -4.164  0.371  0.009 |
## T1_2h_R1 | 42.144 | -39.294  3.292  0.869 | -0.640  0.009  0.000 |
## T1_3h_R1 | 40.201 | -38.491  3.159  0.917 | -5.336  0.610  0.018 |
## T1_4h_R1 | 37.695 | -35.219  2.645  0.873 | -4.543  0.442  0.015 |
## T1_5h_R1 | 42.968 | -40.378  3.476  0.883 | -7.481  1.199  0.030 |
## T1_6h_R1 | 45.724 | -43.317  4.001  0.897 | -9.363  1.879  0.042 |
## T2_1h_R1 | 36.731 | -33.516  2.395  0.833 |  5.595  0.671  0.023 |
## T2_2h_R1 | 22.747 | -0.509  0.001  0.000 | 18.794  7.569  0.683 |
## T2_3h_R1 | 38.037 |  30.276  1.954  0.634 | 17.243  6.371  0.205 |
## T2_4h_R1 | 39.325 |  36.425  2.829  0.858 |  3.068  0.202  0.006 |
##
##          Dim.3   ctr   cos2
## T1_1h_R1   0.515  0.023  0.000 |
## T1_2h_R1   2.418  0.496  0.003 |
## T1_3h_R1  -2.686  0.612  0.004 |
## T1_4h_R1  -1.821  0.281  0.002 |
## T1_5h_R1  -2.866  0.697  0.004 |
## T1_6h_R1  -2.032  0.350  0.002 |
## T2_1h_R1   7.171  4.364  0.038 |
## T2_2h_R1   8.379  5.958  0.136 |
## T2_3h_R1  -1.292  0.142  0.001 |
## T2_4h_R1  -4.030  1.378  0.011 |

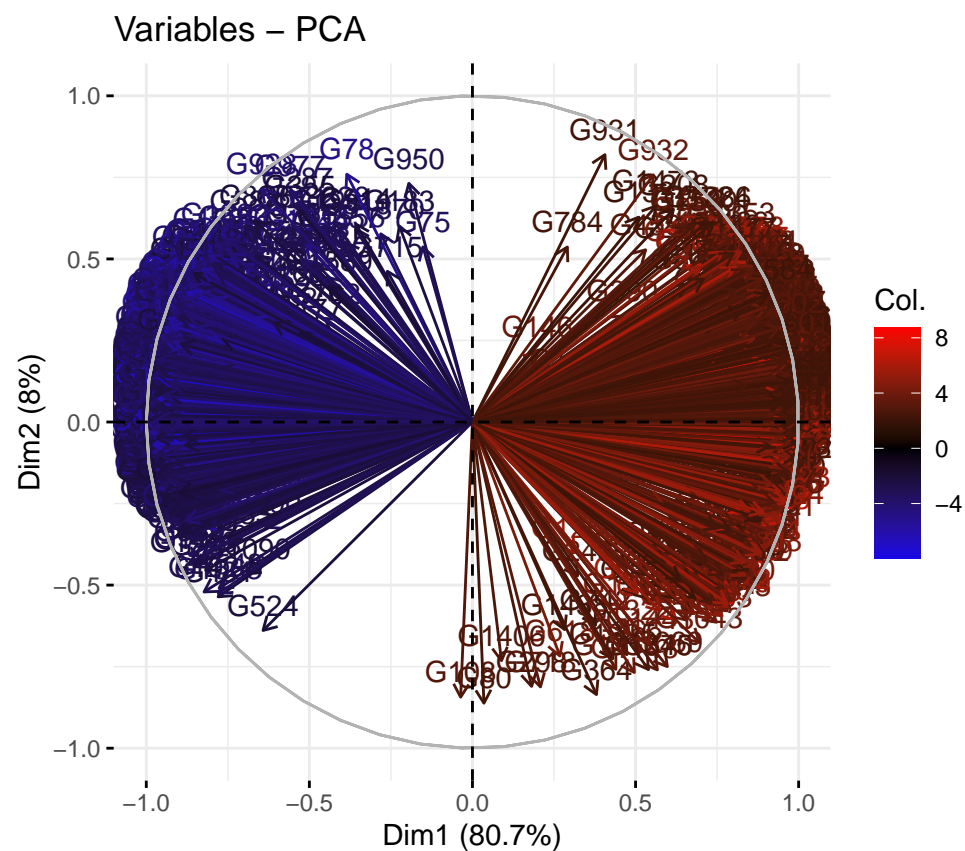
```

```
##
## Variables (the 10 first)
##          Dim.1   ctr   cos2   Dim.2   ctr   cos2   Dim.3   ctr   cos2
## G1      |  0.834  0.053  0.696 | -0.485  0.181  0.235 |  0.004  0.000  0.000 |
## G2      |  0.910  0.064  0.829 |  0.298  0.068  0.089 | -0.083  0.021  0.007 |
## G3      | -0.952  0.070  0.907 |  0.115  0.010  0.013 |  0.160  0.078  0.026 |
## G4      |  0.874  0.059  0.764 | -0.244  0.046  0.059 |  0.163  0.081  0.027 |
## G5      | -0.961  0.071  0.923 | -0.136  0.014  0.018 | -0.089  0.024  0.008 |
## G6      | -0.943  0.068  0.890 | -0.166  0.021  0.028 | -0.125  0.048  0.016 |
## G7      |  0.950  0.069  0.903 |  0.255  0.050  0.065 |  0.050  0.008  0.002 |
## G8      | -0.737  0.042  0.544 |  0.547  0.231  0.299 |  0.260  0.206  0.068 |
## G9      | -0.906  0.063  0.821 | -0.048  0.002  0.002 |  0.225  0.155  0.051 |
## G10     |  0.960  0.071  0.921 |  0.123  0.012  0.015 |  0.088  0.024  0.008 |

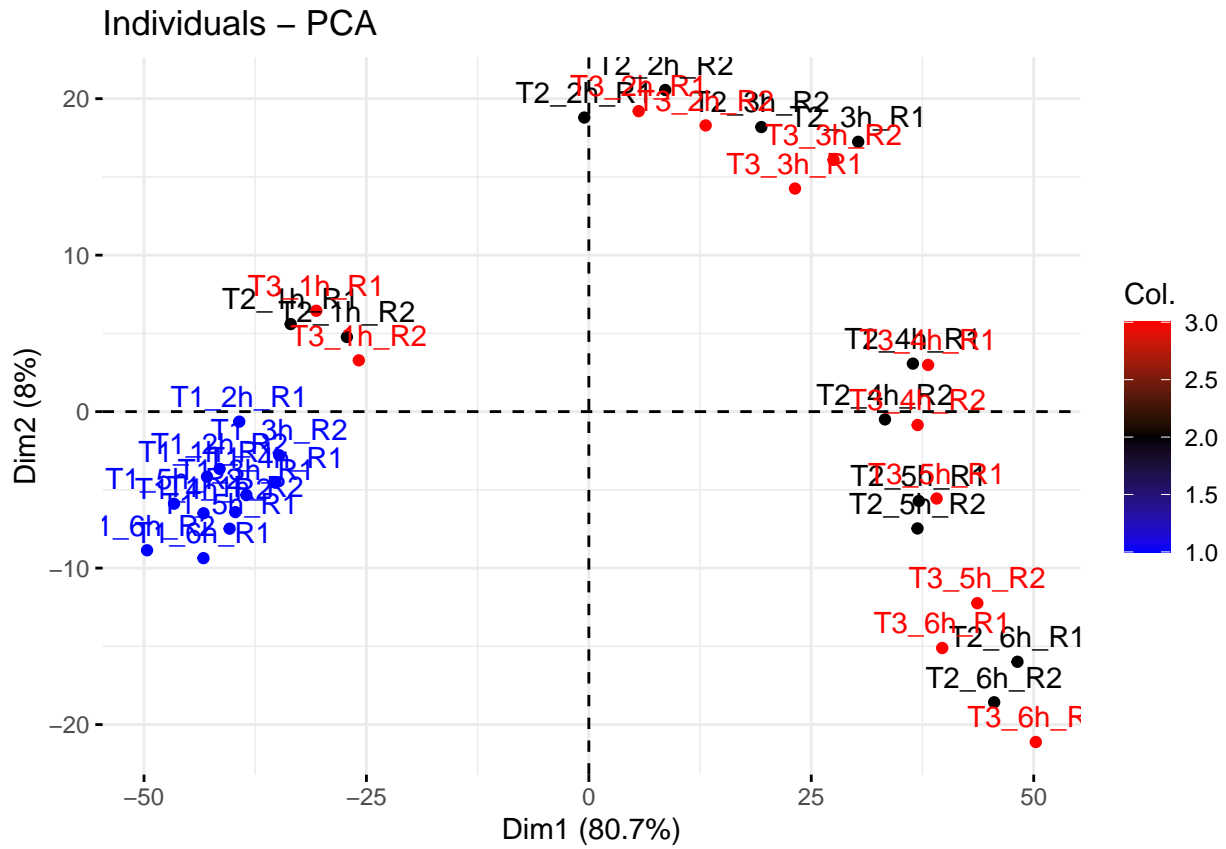
barplot(cumsum(res.pca.transpose$eig[, "percentage of variance"]), main="Percentage of variance")
abline(h=95, col="blue")
abline(h=99, col="red")
```



```
fviz_pca_var(res.pca.transpose, col.var=data$T3_6h_R2) + scale_color_gradient2(low="blue", mid="black",
```



```
fviz_pca_ind(res.pca.transpose, axes=c(1, 2), autoLab="yes", col.ind=col_trait) + scale_color_gradient2
```



Statistiques descriptives et préparation du jeu de données en Python