

# SY09 - TP1

## Statistique descriptive, Analyse en composantes principales

Antoine Jeannot

31 mars 2016

### Résumé

Le premier TP de SY09 a pour objectifs d'apprendre à manipuler les outils fondamentaux de statistiques descriptives du logiciel R, afin de comprendre, d'analyser et de tirer ou non des conclusions à propos de jeux de données complexes issus de paris sportifs ou concernant deux espèces de crabes. Il a également pour but d'illustrer et de nous faire utiliser l'analyse en composantes principales, nous expliquant les besoins et les avantages de l'utilisation d'une telle méthode.

## 1 Statistique descriptive

La première partie du TP concerne les statistiques descriptives. La première sous-partie traite de l'analyse de paris sportifs, la seconde de deux espèces de crabes.

### 1.1 Le racket du tennis

Le jeu de données fourni contient au total quelques 129 271 prises de positions décrites par 16 variables, tel l'identifiant du match de tennis, le bookmaker, l'année du match, le gagnant, le perdant ainsi que la probabilité de gain du dit match. Dans un premier temps, et suite au script de pré-traitement, nous avons réalisé une analyse descriptive de ces données afin de mieux comprendre leur organisation, leur structure et de devenir à l'aise avec leur manipulation.

Ainsi, 25993 matches ont été trouvés dans ces données, ayant lieu de 2009 à 2015. 1523 joueurs différents sont également présents, ainsi que sept matchbookers (notés de A à G).

Une fois cette analyse descriptive générale effectuée, nous avons calculé le nombre de matchs gagnés et perdus par joueur avant de classer les joueurs selon leur niveau. Pour cela, et étant donné qu'un même match possède plusieurs entrées dans le jeu de données (à cause des différents matchbookers), il fal-

lait tout d'abord projeter nos données sur les colonnes `match_uid`, `winner` et `looser` pour obtenir des triplets (`match_uid`, `winner`, `looser`) uniques. Nous avons ensuite calculer deux tables de contingence : une sur la colonne `winner` et une sur la colonne `looser` afin d'obtenir, pour un joueur, son nombre de victoires et de défaites. Enfin, nous avons calculer le niveau de chaque joueur grâce à une fonction `ratio` :  $\text{Victoires} / (\text{Victoires} + \text{Défaites})$ . Un graphique de dispersion des joueurs selon leur niveau a pu être réalisé. (sachant qu'un niveau/ratio de 0.5 correspond à un niveau neutre, autant de victoires que de défaites).

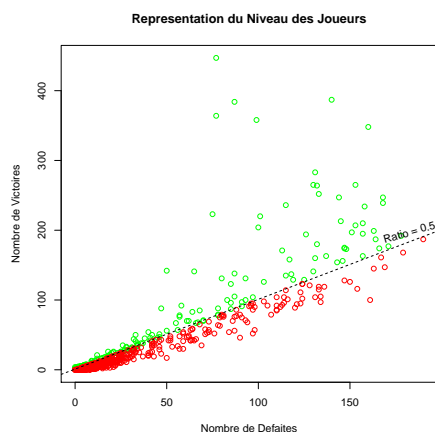


FIGURE 1 – Graphique de dispersion des 1523 joueurs.

Nous avons ensuite étudié les matches suspects et la manière de les distinguer. Pour cela, nous avons calculé l'évolution de la probabilité de gain du gagnant, du perdant ainsi que l'évolution de la probabilité en valeur absolue. Puis nous avons appliqué une restriction sur cette dernière pour ne sélectionner que les matchs qui présentent une évolution de probabilité de gain supérieure à 0.1. Ainsi 2798 matches suspects ont été authentifiés, menant à 4298 paris suspects. Nous avons ainsi généré un data.frame composé uniquement de match suspects. De plus, nous avons recherché les bookmakers impliqués dans ces matches suspects en calculant la table de contingence pour la colonne book sur l'ensemble de ces matches suspects. Tous ont participé au moins 121 fois dans des matches suspects. Nous avons également effectué un ratio nombre de matches suspects / nombre de matches par bookmaker pour avoir la proportion de potentiels litiges par nombre de matches joués et par conséquent pouvoir comparer les "petits" bookmakers aux "grands".

TABLE 1 – Nombre de matches joués par bookmaker.

A	B	C	D	E	F	G
22681	22285	21482	16688	16685	14161	12479

TABLE 2 – Nombre de matches suspects par bookmaker.

B	C	A	D	E	F	G
1363	1117	1007	243	229	218	121

TABLE 3 – Ratio nombre de matches suspects / nombre de matchs par bookmaker. (en %)

B	C	A	D	E	F	G
6.12	5.20	4.44	1.46	1.37	1.54	0.97

Ainsi, on se rend compte que les bookmakers B, C et A sont les plus impliqués dans des matches suspects avec un ratio allant de 4.44% à 6.12%. Ceux-ci sont également les plus gros bookmakers (plus grand nombre de matches joués).

Ensuite, nous avons cherché les joueurs suspectés d'être associés aux matches jugés litigieux. Pour cela, nous avons regroupé les gagnants dont la probabilité de gain du perdant a augmenté de manière significative (supérieure strictement à 0.1). De même pour les perdants dont l'évolution de la probabilité de gain du gagnant a diminué de manière significative (inférieure strictement à -0.1). Nous avons ensuite appliqué une union sur ces deux ensembles de joueurs afin de regrouper les gagnants et perdants suspects au sein d'un même tableau "joueurs suspects". (la fonction union de R enlevant les doublons). Ainsi, nous avons relevés 333 gagnants suspects, 365 perdants suspects, pour un total de 462 joueurs suspects, ce qui nous confirme le fait que de nombreux joueurs sont suspectés à la fois de gagner et de perdre des matchs de manière suspicieuse. Enfin, nous avons calculé le nombre de joueurs impliqués dans un grand nombre de défaites suspectes à l'aide d'une table de contingence sur l'ensemble des perdants précédemment généré et d'une restriction qui nous a permis de garder les joueurs impliqués dans plus de 10 (strictement) défaites suspectes. 37 joueurs sont donc impliqués dans plus de 10 défaites suspectes.

## 1.2 Données crabs

Le jeu de donnée crabs est un jeu de donnée dit "classique" et est constitué de 200 observations décrites par 8 variables, dont 3 qualitatives et 5 quantitatives. Tout comme pour les paris sportifs, nous avons dans un premier temps commencé notre travail par une analyse descriptive des données. Pour mieux cerner les différences entre espèces et sexes, nous avons réalisé de nombreuses visualisation sous la forme de boxplots (Fig. 12 à Fig. 17 en annexe), de paires de variables (Fig. 3) et de tables sommaires (Tables 7 à 12) réalisées par la fonction R summary.

Les boxplots, paires de variables, sommaires décrivent les caractéristiques à propos des :

- Crabs de sexe M
- Crabs de sexe F
- Crabs de sexe M et d'espèce O
- Crabs de sexe M et d'espèce B
- Crabs de sexe F et d'espèce O
- Crabs de sexe F et d'espèce B

Après avoir réaliser ces représentations et sommaires, nous nous sommes rendus comptes qu'il existait des différences physiques entre les deux espèces, et entre sexes d'une même espèce. En effet, à première vue en comparant la taille des boxplots, l'espèce O apparait plus grande que l'espèce B. De plus, et pour toute espèce, les crabs de sexe M apparaissent également quelque peu plus grand que les crabs de sexe F. On notera toutefois que la différence de taille entre l'espèce O et B est plus marquée pour les crabs de sexe F, par rapport aux crabs de sexe M. Les paires de variables confirment également nos différentes constatations à propos de l'existence de différences physiques entre espèces et sexes.

Malgré cette étude descriptive et la présence de différences physiques, l'overlapping (recouvrement) général entre sexes et espèces rend difficile l'identification précision du sexe ou de l'espèce d'un crabe à partir de mesures de ses caractéristiques. On pourrait toutefois envisager de déterminer le sexe d'un "grand" crabe grâce aux variables RW et CL en calculant les coefficients directeurs des droites correspondant aux crabs de sexe M et F, bien qu'il y ait recouvrement pour les "petits" crabs ayant un RW

inférieur à 13 et un CL inférieur à 30. On pourrait également envisager de construire un arbre de décision statistique basé sur les coefficients directeurs des droites des paires de variables.

Dans un second temps, nous avons étudié la corrélation entre les différentes variables en réalisant un graphique de paires de variables, et en calculant les matrices de corrélation. Les valeurs présentes au sein de la matrices de corrélation (Table 4) nous informent que toutes les variables (FL,RW,CL,CW) sont extrêmement dépendantes, qu'une forte dépendance linéaire existe entre elles. (la corrélation minimale entre deux variables étant de 0.889, entre BD et RW) Les graphiques de paires de variables (Fig.2 et Fig.3) viennent nous confirmer la forte dépendance linéaire entre chacune des caractéristiques.

TABLE 4 – Matrix de corrélation entre les caractéristiques des crabs.

	FL	RW	CL	CW	BD
FL	1	0.907	0.979	0.965	0.988
RW	0.907	1	0.893	0.900	0.889
CL	0.979	0.893	1	0.995	0.983
CW	0.965	0.900	0.995	1	0.968
BD	0.988	0.889	0.983	0.968	1

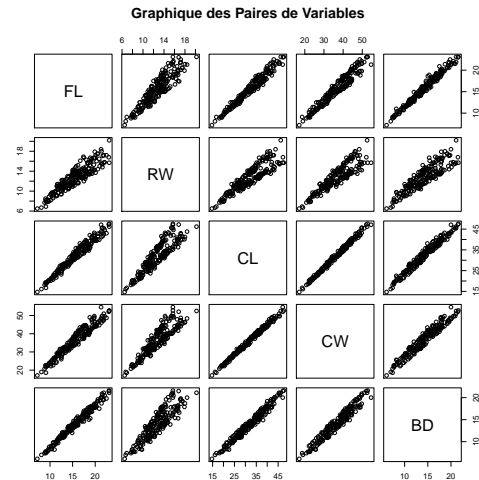


FIGURE 2 – Représentation des paires de variables.

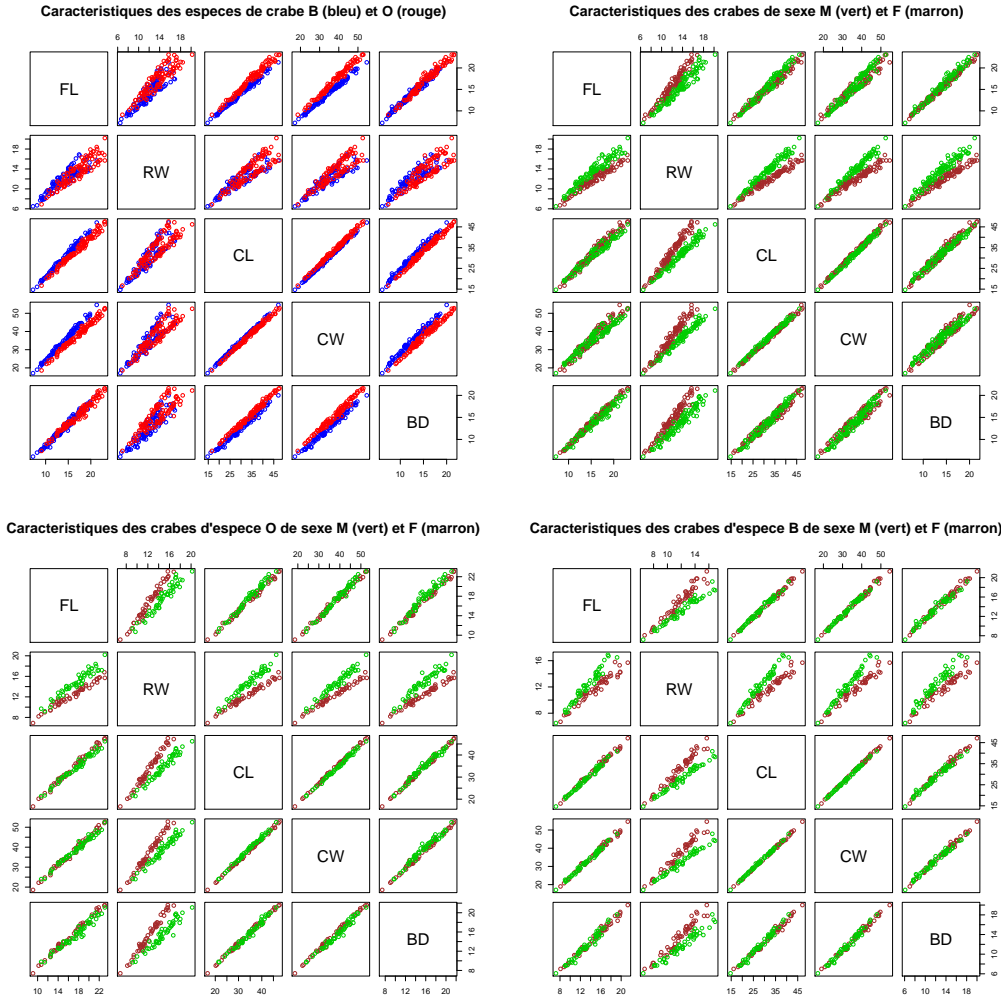


FIGURE 3 – Représentations des paires de variables entre différentes espèces et sexes.

La cause de la forte dépendance des variables peut être expliquée par le fait que ces dernières correspondent à des mensurations et que, comme pour la plupart des espèces, ces dernières sont proportionnelles entre elles. En effet, chez l'homme par exemple, la taille des jambes est généralement proportionnelle à la taille des bras, tout comme l'est la taille des pieds. Ici, la longueur de la carapace d'un crabe est bien proportionnelle à sa largeur (CL pour Carapace Length, CW pour Carapace Width), tout comme la taille du lobe frontal (FL pour Frontal Lobe), la taille de son postérieur (RW pour Rear Width) et la profondeur de son corps

(BD pour Body Depth).

Toutefois, pour s'affranchir de ce phénomène nous pourrions utiliser l'Analyse en Composantes Principales qui consiste à transformer des variables liées entre elles, corrélées, en nouvelles variables décorrélées les unes des autres, : les composantes principales.

## 2 Analyse en composantes principales

La deuxième partie de ce TP a pour objectifs de nous faire découvrir et appliquer l'Analyse en Composantes Principales, en l'implémentant à la fois manuellement et par le biais de fonctions pré-conçues R.

### 2.1 Exercice Théorique

Cet exercice théorique a pour but d'illustrer le fonctionnement de l'A.C.P. sur un exemple relativement simple et compréhensible avec une matrice initiale de taille réduite :

$$X = \begin{pmatrix} 3 & 4 & 3 \\ 1 & 4 & 3 \\ 2 & 3 & 6 \\ 4 & 1 & 2 \end{pmatrix}$$

Dans cet exercice, nous prendrons :

$$D_p = \frac{1}{n}I_n \quad \text{et} \quad M = I_p$$

Pour utiliser l'ACP, il nous faut tout d'abord centrer en colonne X en soustrayant la moyenne correspondante à chaque colonne. Les moyennes des trois variables sont respectivement 2.5, 3.0 et 3.5.

La matrice X devient alors :

$$X = \begin{pmatrix} 0.50 & 1.0 & -0.50 \\ -1.50 & 1.0 & -0.50 \\ -0.50 & 0.0 & 2.50 \\ 1.50 & -2.0 & -1.50 \end{pmatrix}$$

Puis, on calcule la matrice de variance :

$$V = X^T D_p X = \frac{1}{n} X^T X = \frac{1}{4} X^T X$$

$$V = \begin{pmatrix} 1.25 & -1.0 & -0.75 \\ -1.0 & 1.50 & 0.50 \\ -0.75 & 0.50 & 2.25 \end{pmatrix}$$

Enfin, on diagonalise la matrice de variance V afin d'obtenir les valeurs propres et vecteurs propres :

	$\lambda_1$	$\lambda_2$	$\lambda_3$
Valeur	3.199	1.469	0.333
% Inertie Expliquée	63.98	29.37	6.65
% I. E. Cumulée	63.98	93.35	100

$$u_1 = \begin{pmatrix} 0.524 \\ -0.509 \\ -0.683 \end{pmatrix}; u_2 = \begin{pmatrix} 0.339 \\ -0.611 \\ 0.716 \end{pmatrix};$$

$$u_3 = \begin{pmatrix} 0.781 \\ 0.606 \\ 0.148 \end{pmatrix}$$

On remarque ainsi que l'essentielle de l'inertie expliquée est contenue dans les deux premiers axes factoriels. (93.35%)

Ensuite, il nous fallait calculer les composantes principales avant de représenter les quatre individus dans le premier plan factoriel.

$$C = X M U = X U$$

$$C = \begin{pmatrix} 0.0940 & -0.799 & 0.923 \\ -0.954 & -1.477 & -0.640 \\ -1.969 & 1.620 & -0.022 \\ 2.829 & 0.656 & -0.262 \end{pmatrix}$$

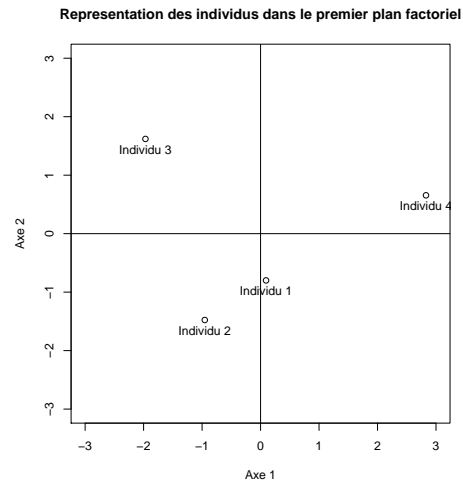


FIGURE 4 – Représentation des individus dans le premier plan factoriel.

Puis, nous avons tracé la représentation des trois variables dans le premier plan factoriel suite au calcul de la matrice de corrélation entre

la matrice des composantes principales C et la matrice initiale centrée X.

$$\text{cor}(\alpha, j) = \begin{pmatrix} 0.838 & 0.367 & 0.403 \\ -0.744 & -0.604 & 0.285 \\ -0.814 & 0.578 & 0.057 \end{pmatrix}$$

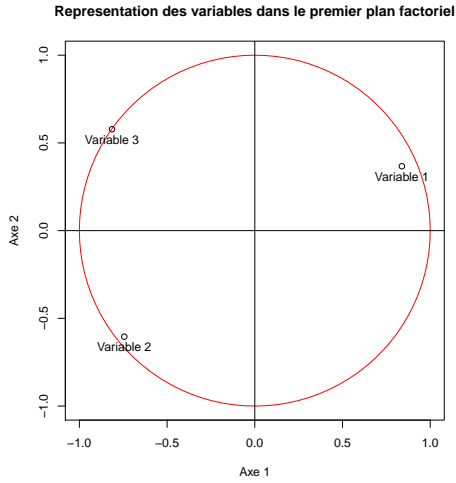


FIGURE 5 – Représentation des variables dans le premier plan factoriel.

Nous avons ensuite calculer l'expression  $\sum_{\alpha=1}^k c_{\alpha} u_{\alpha}^T$  pour tous les k :

$$k = 1 : \begin{pmatrix} 0.049 & -0.048 & -0.064 \\ -0.500 & 0.486 & 0.651 \\ -1.032 & 1.003 & 1.344 \\ 1.482 & -1.441 & -1.931 \end{pmatrix}$$

$$k = 2 : \begin{pmatrix} -0.221 & 0.440 & -0.636 \\ -1.000 & 1.388 & -0.406 \\ -0.483 & 0.013 & 2.503 \\ 1.704 & -1.841 & -1.461 \end{pmatrix}$$

$$k = 3 : \begin{pmatrix} 0.50 & 1.0 & -0.50 \\ -1.50 & 1.0 & -0.50 \\ -0.50 & 0.0 & 2.50 \\ 1.50 & -2.0 & -1.50 \end{pmatrix} = X$$

Lorsque k=3, cette somme correspond à  $X = CU^T$  et permet de reconstituer la matrice initiale centrée X.

## 2.2 Utilisation des outils R

L'objectif de cet exercice est de se familiariser avec le logiciel R et les fonctions pré-faites concernant l'Analyse en Composantes Principales. Pour tester ces fonctions, nous avons chargé la matrice de notes puis tester la fonction `princomp`. `Princomp` permet de calculer automatiquement les composantes principales d'une matrice passée en argument et retourne un objet R constitué de plusieurs variables membres :

- `sdev` : écarts types des composantes principales. ( $\text{sdev}^2$  correspondant aux valeurs propres)
- `loadings` : matrice des vecteurs propres.
- `scores` : matrice des composantes principales.

Ainsi, on calcule la matrice des composantes principales de notes et on l'affiche à l'aide de `C$scores` :

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
jean	8.671	1.419	-2.810	-0.233	0.088
alin	3.916	0.519	-1.919	-0.054	0.113
anni	3.195	-3.484	-0.0212	0.131	-0.149
moni	-10.049	-0.817	-3.205	0.129	0.164
didi	-6.377	2.099	-1.069	-0.199	-0.139
andr	2.772	4.540	2.113	0.456	0.020
pier	1.011	-6.362	-1.112	0.089	-0.084
brig	-1.701	4.518	1.061	-0.136	-0.135
evel	-1.437	-2.431	6.962	-0.183	0.122

Les valeurs propres sont ensuite calculées grâce au carré des écarts-types donnés par `C$sdev`, on peut également calculer l'inertie expliquée par chaque valeur propre ainsi que l'inertie expliquée cumulée :

TABLE 5 – Valeurs propres, % Inertie Expliquée et % Inertie Expliquée Cumulée

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
Valeur	28.450	11.878	8.699	0.044	0.014
% Inertie Expliquée	57.960	24.198	17.723	0.090	0.029
% Inertie Expliquée Cumulée	57.960	82.158	99.881	99.971	100

La matrice des vecteurs propres (ou axes factoriels) est donnée par `C$loadings` :

$u_1$	$u_2$	$u_3$	$u_4$	$u_5$
-0.513	0.571	0	0.344	0.540
-0.505	0.375	0	-0.197	-0.752
-0.498	-0.653	0	0.551	-0.134
-0.482	-0.326	0	-0.732	0.353
0	0	0.997	0	0

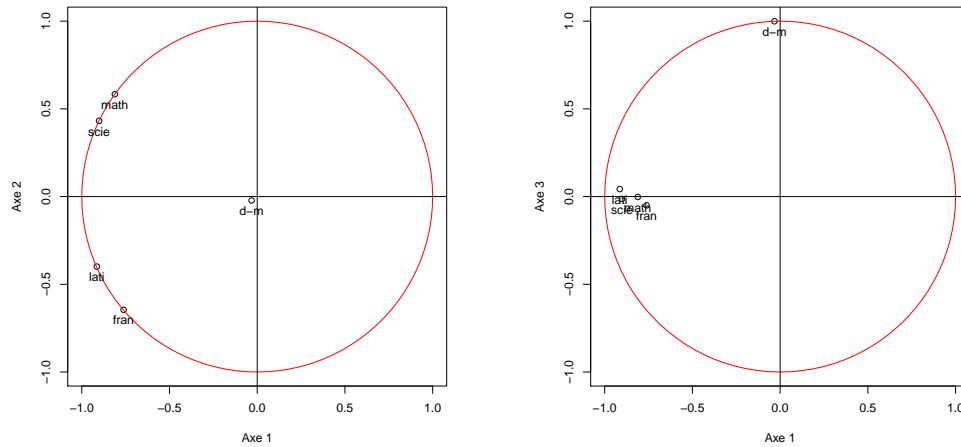


FIGURE 6 – Représentations des variables dans le premier et deuxième plan factoriel.

La table des pourcentages d'inertie expliquée cumulée (Table 5) nous informe que le nuage de points est pratiquement dans un espace de dimension 3, l'inertie expliquée cumulée correspondante étant de 99.88%.

La fonction R `plot` sur le résultat de la fonction R `princomp` génère l'historgramme des variances des composantes principales, nous permettant de visualiser l'importance de la première composante principale et de constater que les trois premières composantes représentent, "contiennent" pratiquement tout l'information ce qui confirme les constatations précédentes liées à la table des pourcentages d'inertie cumulée.

La fonction `biplot` quant à elle affiche les individus et les variables dans le premier plan factoriel. La fonction `biplot.princomp` bénéficie de paramètres supplémentaires par rapport à la fonction `biplot` : "choices" pour choisir les composantes principales à représenter, "scale" pour obtenir une représentation standard des données.

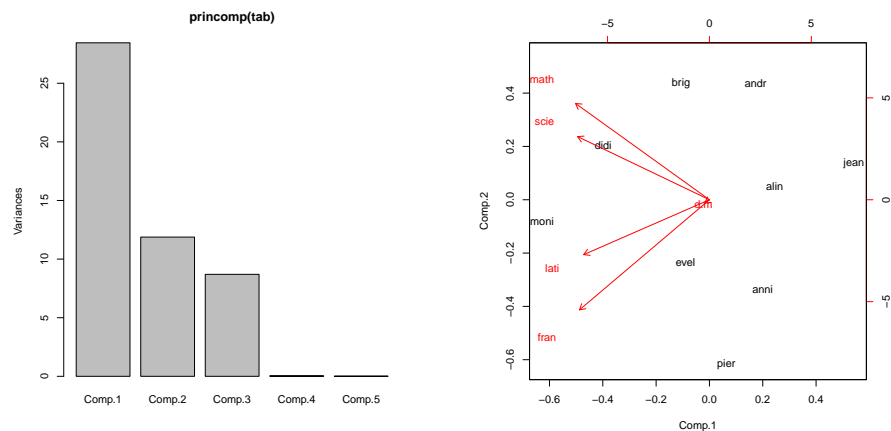


FIGURE 7 – Représentations obtenues à l'aide des fonctions `plot` et `biplot`.



## 2.3 Traitement des données Crabs

Dans cette dernière partie, nous avons réutilisé les données Crabs pour leur appliquer une A.C.P. et ainsi essayer de distinguer visuellement les espèces et sexes de crabes. Après avoir centré et appliqué une A.C.P. sur les cinq variables quantitatives des données crabs, nous avons obtenu la matrice de vecteurs propres suivantes :

$u_1$	$u_2$	$u_3$	$u_4$	$u_5$
-0.289	-0.323	0.507	0.734	0.125
-0.197	-0.865	-0.414	-0.148	-0.141
-0.599	0.198	0.175	-0.144	-0.742
-0.662	0.288	-0.491	0.126	0.471
-0.284	-0.160	0.547	-0.634	0.439

Ainsi que les valeurs propres suivantes :

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
Valeur	140.00	1.29	1.00	0.13	0.078
% Inertie Expliquée	98.25	0.91	0.70	0.094	0.054
% Inertie Expliquée Cumulée	98.25	99.15	99.85	99.95	100

Nous pouvons déduire, grâce à la table des pourcentages d'inertie expliquée cumulée, que le nuage de points est presque entièrement dans un espace de dimension 2, voire 1 tant la première valeur propre est importante comparée aux autres. On peut alors se demander si celle-ci risque de nuire au lieu d'aider à la distinction des classes. Puis en effectuant la matrice de corrélations entre les différentes composantes principales, nous remarquons qu'il s'agit de la matrice identité.

TABLE 6 – Matrice de corrélation des composantes principales  $cor(\alpha, j)$ .

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Comp.1	1	0	0	0	0
Comp.2	0	1	0	0	0
Comp.3	0	0	1	0	0
Comp.4	0	0	0	1	0
Comp.5	0	0	0	0	1

Nous pouvons ainsi en déduire que les composantes principales sont totalement dé-corrélées grâce à l'A.C.P. Pour visualiser les différentes espèces et sexes de crabes, nous avons créé une nouvelle catégories à quatre modalités à laquelle nous avons fais correspondre une couleur :

- 1 : crabes d'espèce O de sexe M (Bleu)
- 2 : crabes d'espèce O de sexe F (Violet)
- 3 : crabes d'espèce B de sexe M (Cyan)
- 4 : crabes d'espèce B de sexe F. (Rouge)

Nous avons ensuite représenté les individus et axes principaux dans le premier plan factoriel ainsi que les paires de composantes principales, nous permettant de distinguer le sexe et l'espèce d'un crabe en fonction de ses caractéristiques, grâce aux composantes 2 et 3 notamment. Ainsi, malgré l'importante valeur de la première valeur propre, on constate que celle-ci n'apporte que très peu d'information pour la distinction des groupes de crabes tant il y a du recouvrement (overlapping) entre la première composante principale et les suivantes.(voir la première colonne de la Fig. 9)

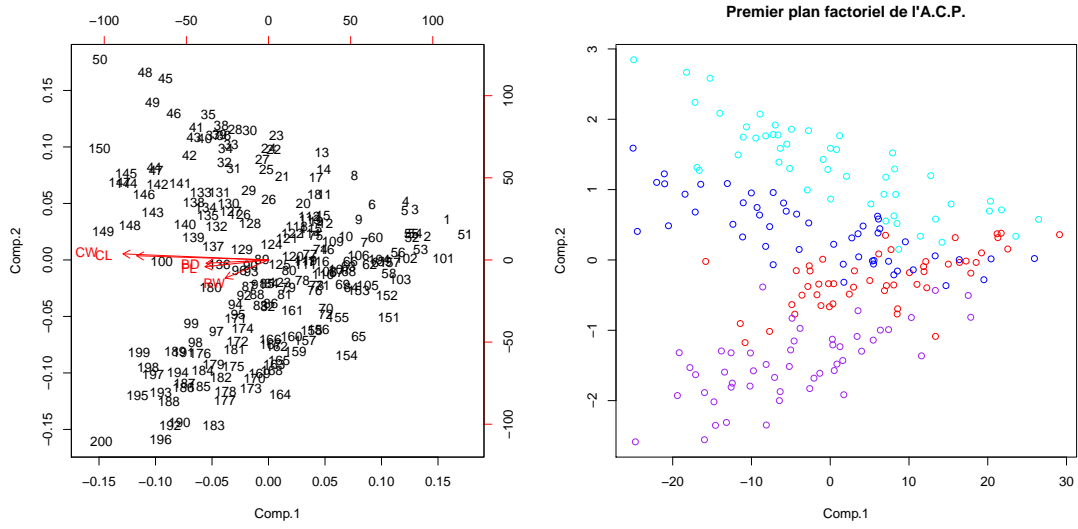


FIGURE 8 – Représentations des axes principaux et des individus dans le premier plan factoriel.

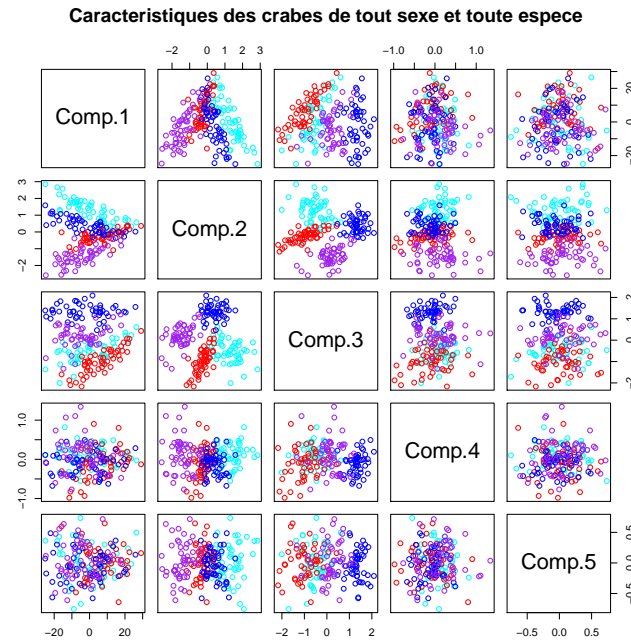


FIGURE 9 – Représentation des paires de composantes principales.

Pour améliorer la qualité de cette représentation, nous avons donc supprimé la première composante principale en divisant chaque composante principale par cette première avant d'effectuer les mêmes représentations. Sur ces dernières, les composantes principales 1 et 2 nous permettent de

distinguer plus facilement et directement le sexe et l'espèce d'un crabe, ces différents groupes étant plus distants et séparables linéairement les uns des autres.

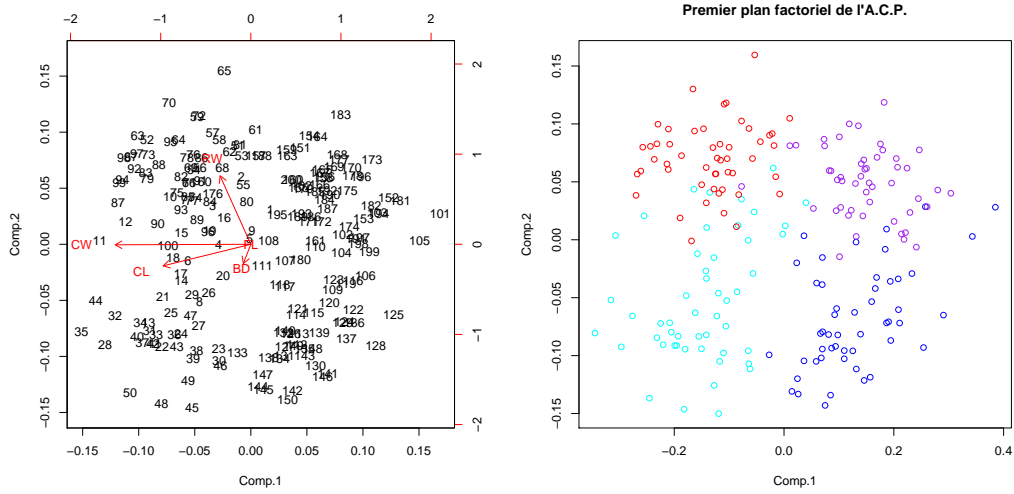


FIGURE 10 – Nouvelles représentations des individus et axes principaux dans le premier plan factoriel.

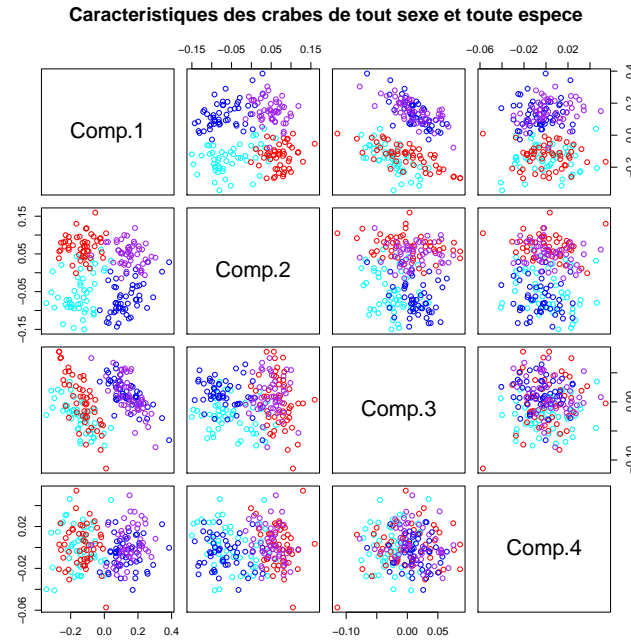


FIGURE 11 – Nouvelles représentations des paires de composantes principales.

### 3 Conclusion

Ce premier TP de SY09 nous aura permis de prendre en main le logiciel R, sa syntaxe et ses différentes fonctions. Il nous aura également appris à effectuer des analyses descriptives sur des jeux de données complexes et variés tirés de la vie courante, de nous confronter aux problèmes de corrélations de variables et de découvrir les avantages de l'A.C.P. pour les dé-corréler et mieux distinguer différents groupes (comme les espèces de crabes) à travers des visualisations, mais aussi par la suite lors de classifications automatiques.

## A Annexes

TABLE 7 – Description des crabes de sexe M

FL	RW	CL	CW	BD
Min. : 8.10	Min. : 6.70	Min. :16.10	Min. :18.60	Min. : 7.00
1st Qu. :13.18	1st Qu. :10.70	1st Qu. :27.65	1st Qu. :31.60	1st Qu. :11.85
Median :15.70	Median :11.85	Median :32.45	Median :37.10	Median :14.20
Mean :15.73	Mean :11.99	Mean :32.85	Mean :37.00	Mean :14.34
3rd Qu. :18.05	3rd Qu. :13.50	3rd Qu. :37.95	3rd Qu. :42.83	3rd Qu. :16.85
Max. :23.10	Max. :16.80	Max. :47.60	Max. :54.60	Max. :21.60

TABLE 8 – Description des crabes de sexe F

FL	RW	CL	CW	BD
Min. : 7.20	Min. : 6.50	Min. :14.70	Min. :17.10	Min. : 6.10
1st Qu. :12.80	1st Qu. :11.50	1st Qu. :26.55	1st Qu. :31.18	1st Qu. :11.28
Median :15.45	Median :13.80	Median :31.70	Median :36.45	Median :13.80
Mean :15.43	Mean :13.49	Mean :31.36	Mean :35.83	Mean :13.72
3rd Qu. :18.07	3rd Qu. :15.55	3rd Qu. :35.92	3rd Qu. :40.77	3rd Qu. :16.02
Max. :23.10	Max. :20.20	Max. :46.20	Max. :52.50	Max. :21.10

TABLE 9 – Description des crabes de sexe M et d'espece B

FL	RW	CL	CW	BD
Min. : 8.10	Min. : 6.70	Min. :16.10	Min. :19.00	Min. : 7.00
1st Qu. :12.65	1st Qu. :10.65	1st Qu. :27.23	1st Qu. :31.52	1st Qu. :11.00
Median :15.10	Median :11.70	Median :32.45	Median :37.10	Median :13.60
Mean :14.84	Mean :11.72	Mean :32.01	Mean :36.81	Mean :13.35
3rd Qu. :17.05	3rd Qu. :13.28	3rd Qu. :37.15	3rd Qu. :42.52	3rd Qu. :15.60
Max. :21.30	Max. :15.80	Max. :47.10	Max. :54.60	Max. :20.00

TABLE 10 – Description des crabes de sexe M et d'espece O

FL	RW	CL	CW	BD
Min. : 9.10	Min. : 6.90	Min. :16.70	Min. :18.60	Min. : 7.40
1st Qu. :14.10	1st Qu. :10.78	1st Qu. :28.75	1st Qu. :31.73	1st Qu. :12.95
Median :16.70	Median :12.10	Median :33.35	Median :36.30	Median :15.00
Mean :16.63	Mean :12.26	Mean :33.69	Mean :37.19	Mean :15.32
3rd Qu. :18.80	3rd Qu. :13.70	3rd Qu. :39.02	3rd Qu. :43.08	3rd Qu. :17.77
Max. :23.10	Max. :16.80	Max. :47.60	Max. :52.80	Max. :21.60

TABLE 11 – Description des crabes de sexe F et d'espece B

FL	RW	CL	CW	BD
Min. : 7.20	Min. : 6.50	Min. :14.70	Min. :17.10	Min. : 6.10
1st Qu. :11.53	1st Qu. :10.62	1st Qu. :23.93	1st Qu. :27.90	1st Qu. :10.03
Median :13.15	Median :12.20	Median :27.90	Median :32.35	Median :11.60
Mean :13.27	Mean :12.14	Mean :28.10	Mean :32.62	Mean :11.82
3rd Qu. :15.28	3rd Qu. :13.88	3rd Qu. :32.75	3rd Qu. :37.83	3rd Qu. :13.88
Max. :19.20	Max. :16.90	Max. :40.90	Max. :47.90	Max. :18.10

TABLE 12 – Description des crabes de sexe F et d'espece O

FL	RW	CL	CW	BD
Min. :10.70	Min. : 9.20	Min. :21.40	Min. :24.00	Min. : 9.70
1st Qu. :15.60	1st Qu. :13.60	1st Qu. :31.05	1st Qu. :34.88	1st Qu. :13.80
Median :18.00	Median :14.65	Median :34.70	Median :39.55	Median :15.65
Mean :17.59	Mean :14.84	Mean :34.62	Mean :39.04	Mean :15.63
3rd Qu. :19.90	3rd Qu. :16.68	3rd Qu. :39.70	3rd Qu. :44.05	3rd Qu. :17.85
Max. :23.10	Max. :20.20	Max. :46.20	Max. :52.50	Max. :21.10

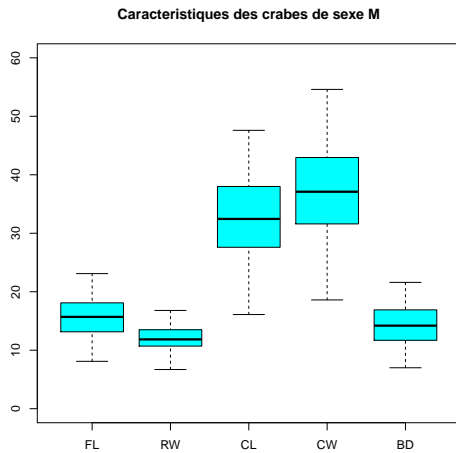


FIGURE 12 – Caractéristiques des crabes de sexe M.

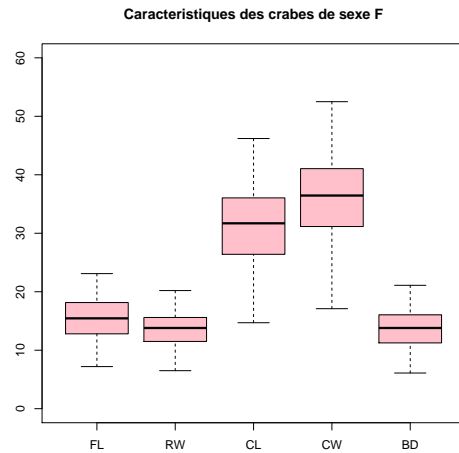


FIGURE 13 – Caractéristiques des crabes de sexe F.

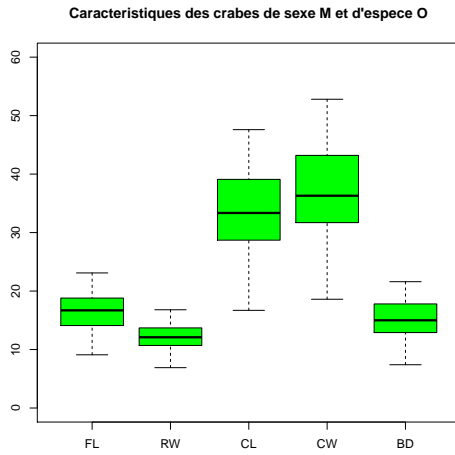


FIGURE 14 – Caractéristiques des crabes de sexe M et d'espèce O.

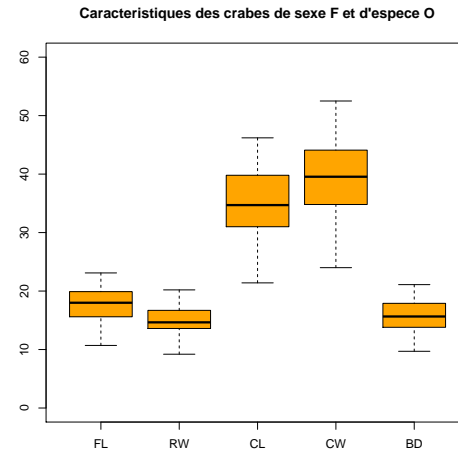


FIGURE 16 – Caractéristiques des crabes de sexe F et d'espèce O.

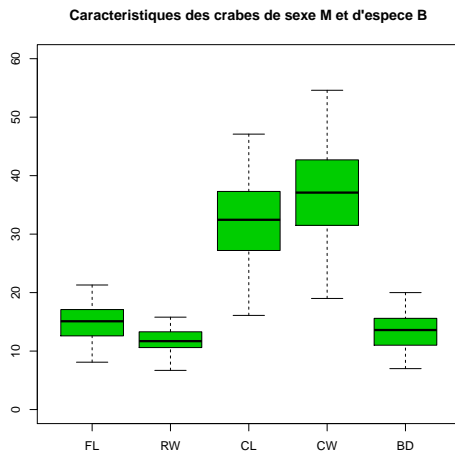


FIGURE 15 – Caractéristiques des crabes de sexe M et d'espèce B.

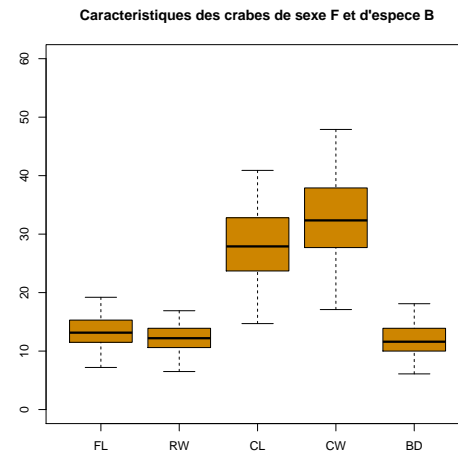


FIGURE 17 – Caractéristiques des crabes de sexe F et d'espèce B.