

Examen d'Analyse de Données de Janvier 2019

Masters 1 IEF et FE - Université de Rennes 1

Cours de Sylvain Barthélémy

Prénom	
Nom	
Master	

INSTRUCTIONS IMPORTANTES

L'épreuve dure 1h30.

Vous devrez utiliser le logiciel R (ou RStudio) pour répondre à un certain nombre de questions. N'oubliez pas de copier le code, ainsi que vos résultats, dans les encadrés prévus à cet effet.

A l'issue de l'épreuve, vous devrez rendre ce document, complété avec vos réponses (et le code R), en envoyant un email à :

sylvain.barthelemy@taceconomics.com

Première Partie : Programmation (12pts)

Téléchargez le fichier Excel « usstates.xlsx » se trouvant sur le github du cours (<https://github.com/sylbarth/m1mbfa>).

a) Importez l'onglet « data » du fichier Excel « usstates.xlsx »

b) Renommez les individus (lignes du tableau) par le code des Etats (colonne « iso »), puis supprimer la colonne « iso » de la table.

c) Ecrivez une fonction qui effectue automatiquement les étapes (a) et (b), ne prenant en argument que le nom du fichier et le nom de l'onglet, et qui renvoie automatiquement les données.

d) Produisez des statistiques descriptives simples sur les données des Etats (moyennes, médianes, minimum, maximum).

e) Produisez les graphiques des distributions de probabilités sur les variables « WTOT » et « GDPPC ».

f) Quelles conclusions en tirez-vous sur la dispersion des variables ?

Examen d'Analyse de Données de Janvier 2019

Masters 1 IEF et FE - Université de Rennes 1

Cours de Sylvain Barthélémy

g) Calculez et affichez les corrélations entre les variables, et interprétez les résultats obtenus.

h) Faites un graphique croisé de la variable « WTOT » contre la variable « RGSP »

i) Faites une Analyse en Composante Principale sur l'ensemble des données

j) Déterminez le nombre de composantes à analyser à l'aide de deux critères différents. Justifiez votre choix.

k) Affichez et interprétez les axes sélectionnés

l) Affichez et interprétez les graphiques des individus

Examen d'Analyse de Données de Janvier 2019

Masters 1 IEF et FE - Université de Rennes 1

Cours de Sylvain Barthélémy

m) Refaire la même analyse, mais en mettant en variables supplémentaires les variables « WTOT » et « RGSP » ainsi que la variable « REGION ».

--

n) Interprétez les résultats obtenus

--

o) Déterminez les Etats à exclure (si nécessaire en cas de valeurs extrêmes) du calcul de l'ACP (mettre ces états en individus supplémentaires pour le reste de l'analyse)

--

Deuxième Partie : Cours (8pts)

1. A l'aide des graphiques de l'Annexe 1, faites une analyse des résultats de cette ACP

a) Que représente le chiffre 71.892 dans le tableau ?

b) Que vaut l'inertie totale du nuage après passage sur les 4 facteurs ?

c) Interprétez les graphs des individus en fonction de l'interprétation des axes

2. Questions de cours

a) Expliquez quels sont les objectifs d'une ACP

b) Quelles sont les différentes étapes d'une ACP

c) Donner deux exemples de questions auxquelles une ACP peut répondre

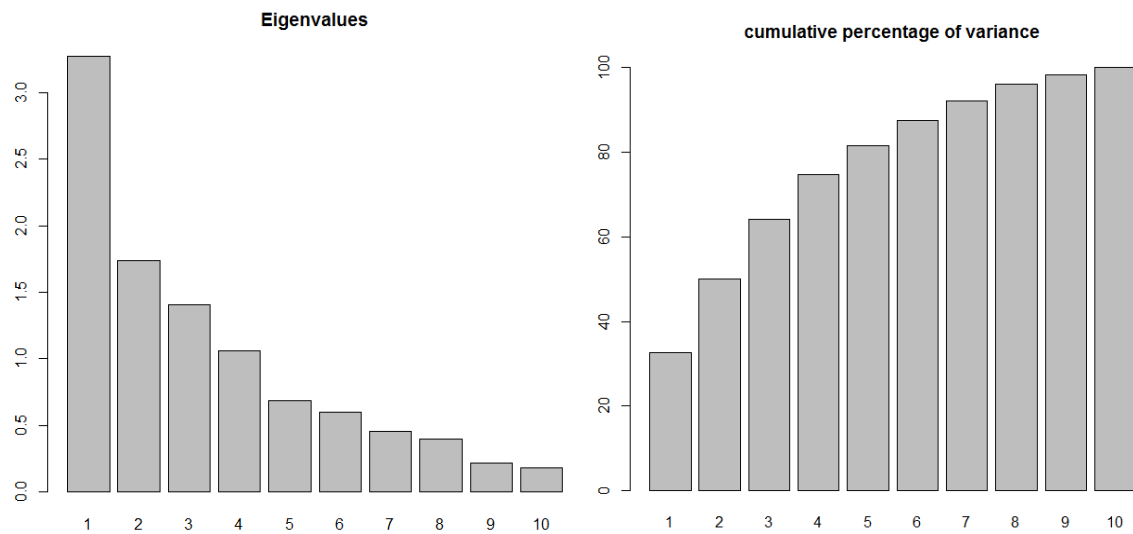
d) Expliquer les différences entre une ACP et une Classification Ascendante Hiérarchique

Examen d'Analyse de Données de Janvier 2019

Masters 1 IEF et FE - Université de Rennes 1

Cours de Sylvain Barthélémy

e) D'après les deux graphiques suivants, combien pouvons-nous retenir de composantes ? (justifier ce choix)



Examen d'Analyse de Données de Janvier 2019

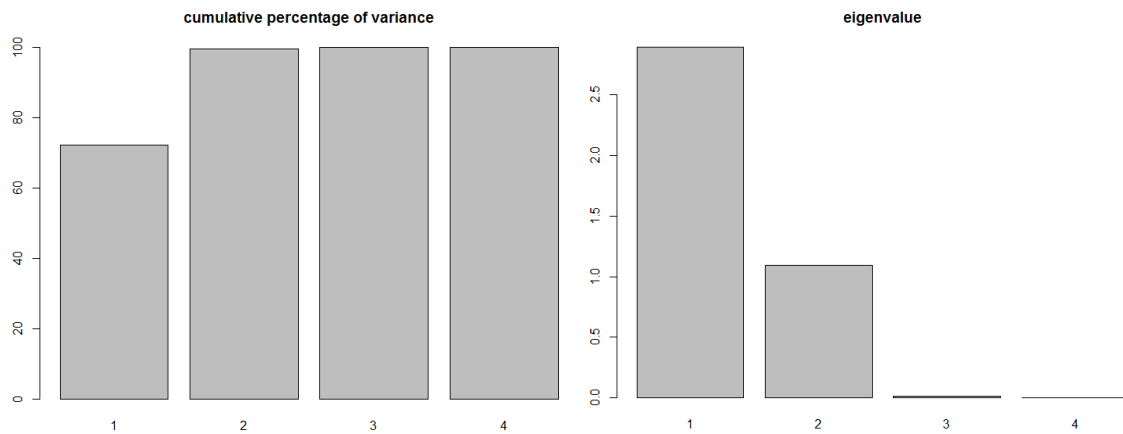
Masters 1 IEF et FE - Université de Rennes 1

Cours de Sylvain Barthélémy

Annexe 1

Nous étudions par une ACP les notes d'élèves dans les matières suivantes :
Mathématiques (maths), Physique (phys), Français (Fran) et Anglais (Angl)

	Dim 1	Dim 2	Dim 3	Dim 4
Variance	2.876	1.120	0.004	0.001
% of var.	71.892	27.992	0.089	0.026
Cumulative % of var.	71.892	99.884	99.974	100.00

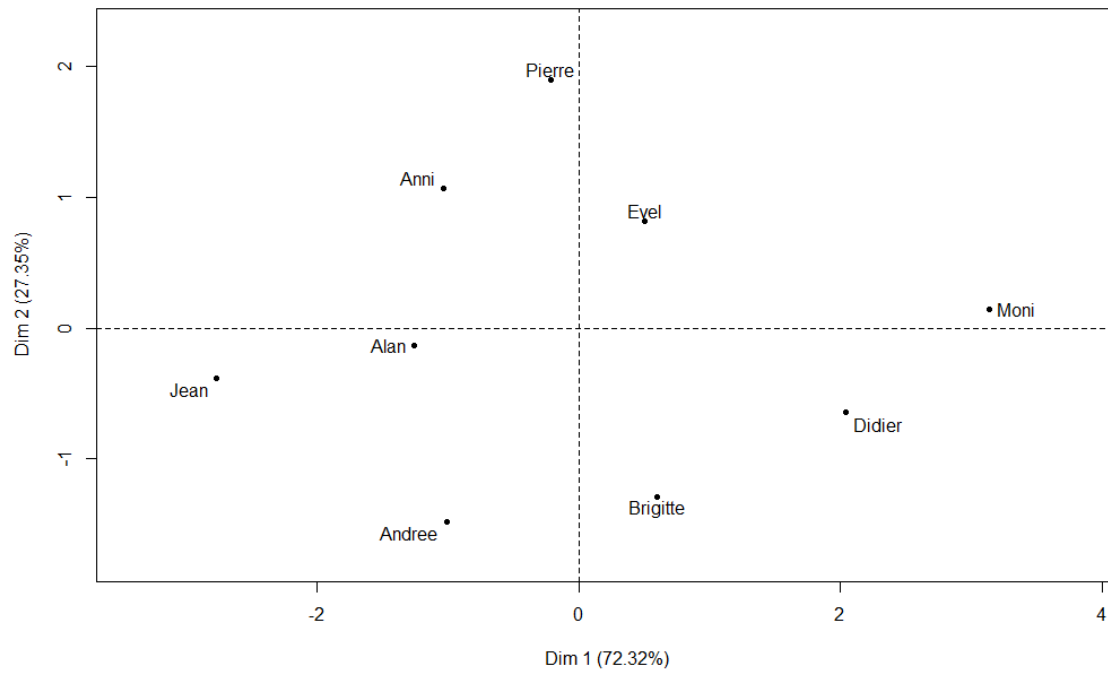


Examen d'Analyse de Données de Janvier 2019

Masters 1 IEF et FE - Université de Rennes 1

Cours de Sylvain Barthélémy

Individuals factor map (PCA)



Variables factor map (PCA)

