

ANALYSE :



**OUTIL D'ANALYSE DE DONNEES POUR LES SCIENCES
HUAMINES**

MANUEL DE L'UTILISATEUR : PRISE EN MAIN

Statistics for Social Sciences and Humanities

Your stat file

Select the type of your file

Select the Operating System of your computer

A. PREMIER PAS

1. INTEGRATION DU TABLEAU DE DONNEES

❖ Envoyer le tableau de données au serveur

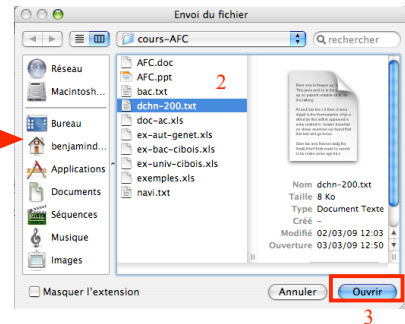
Statistics for Social Sciences and Humanities

Your stat file 1 Parcourir...

Select the type of your file

Select the Operating System of your computer

Send Clear the values and reset



1. Cliquer sur le bouton « Parcourir »
2. Sélectionner le fichier contenant votre tableau de données
3. Cliquer sur le bouton « Ouvrir »

❖ Sélection du type de fichier

Statistics for Social Sciences and Humanities

Your stat file Parcourir...

Select the type of your file 2

Select the Operating System of your computer

Send Clear the values and reset

À cette étape, *Analyse* vous demande de lui indiquer le type de fichier dans lequel est contenu votre tableau de données.

Analyse est capable d'intégrer trois type de fichiers :

- les fichiers textes (extension : « .txt ») dans lesquels les colonnes du tableau de données sont séparées par un caractère de tabulation (« \t ») ;
- les fichiers textes (extension : « .txt ») dans lesquels les colonnes du tableau de données sont séparées par un caractère point-virgule (« ; ») ;
- les fichiers textes produits à l'aide d'open office.

❖ Sélection du système d'exploitation et des paramètres d'encodage

Statistics for Social Sciences and Humanities

Your stat file
/Users/benjaminderuelle/univ-paris1/tdhist/cours/AFC/cours-AFC/dchn-200 (Parcourir...)

Select the type of your file

Select the Operating System of your computer

- Standard Unix, Linux or Windows (ISO 8859-1)
- Macintosh
- Macintosh with Unix line breaks
- DOS
- DOS with Unix line breaks
- NextSTEP
- UTF-8

Il s'agit ici d'indiquer à Analyse le système d'exploitation qui vous a servi à produire votre tableau de données. Cela permet à Analyse de reconnaître l'encodage des caractères utilisés dans votre fichier, ainsi que les données.

❖ Envoi du fichier au serveur

Statistics for Social Sciences and Humanities

Your stat file
/Users/benjaminderuelle/univ-paris1/tdhist/cours/AFC/cours-AFC/dchn-200 (Parcourir...)

Select the type of your file

Select the Operating System of your computer

Cliquer sur le bouton « *Send* » (« envoyer »)

2. CONFIGURATION DU TABLEAU DE DONNEES

Choisissez le type de tableau :

Vous souhaitez faire :

❖ Choisir le type de tableau

Choisissez le type de tableau :

- 1- Tableau de comptage variables par variables (tableau de contingence)
- 2- Tableau avec individus en lignes et variables en colonne (1er colonne est une variable)
- 3- Tableau avec individus en lignes et variables en colonne (1er colonne indique les individus)

Vous souhaitez faire :

Analyse est capable d'intégrer trois types de tableau de données :

- 1. Les tableaux de contingence ou tableaux croisés présentant en ligne et en colonnes une variable et les effectifs d'individus porteurs de ces deux variables sont comptés dans chacune des cellules ;
- 2. Les tableaux de données présentant en colonnes les différentes variables et en lignes les individus ;
- 3. Les tableaux de données présentant en colonnes les différentes variables, en lignes les individus, et dont la première colonne porte un identifiant d'individus type numéro ou dénomination.

❖ Sélectionner le type d'analyse

Choisissez le type de tableau : -1- Tableau de comptage variables par variables (tableau de contingence)

Vous souhaitez faire :
Description des variables
Description des variables
Analyse factorielle des correspondances
Analyse en correspondances multiples
Analyse en correspondances principales
Classements hiérarchiques
utiliser directement R

Send Clear the values and reset

À ce stade d'avancement, vous devez indiquer à *Analyse* ce que vous désirez faire de votre tableau :

1. « Description de variables » : *Analyse* repère les variables de votre tableau de données et vous en propose une définition (variable quantitative : discrète ou continue ; qualitative) afin de déterminer vos besoins en termes d'analyse statistique.
2. « Analyse factorielle des correspondances » : vous ne devez sélectionner ce type d'analyse qu'à la condition d'avoir envoyé au serveur un tableau de contingence (cf. « choisissez un type de tableau »).
3. « Analyse des correspondances multiples » : vous ne devez sélectionner ce type d'analyse qu'à la condition d'avoir envoyé au serveur un tableau de données (cf. « choisissez un type de tableau »).
4. « Analyse des correspondances principales » :
5. « Classements hiérarchiques » :
6. « Utiliser R directement » : cette option vous permet d'utiliser R directement en ligne de commande sur le tableau que vous avez envoyé à *Analyse*.

❖ Valider votre requête

Choisissez le type de tableau : -1- Tableau de comptage variables par variables (tableau de contingence)

Vous souhaitez faire : Analyse factorielle des correspondances

Send Clear the values and reset

Il s'agit d'envoyer votre requête une fois celle-ci paramétrée.

II. LES ANALYSES

A. LA DESCRIPTION DE VARIABLES

1. CARACTERES ET CHOIX DES VARIABLES POUR L'ANALYSE

Les variables

Nouvelle analyseDescriptionACMR directement

AideRapportsGlossaire

Définition de la nature des variables à étudier (518 enregistrements)

Nom de la variable	Type	NA	NA en % des enregistrements	Choix de variable
LANCEMENT	Discrete	0	0	☐
PORT	Qualitative	0	0	☐
LAN	Qualitative	0	0	☐
TYPE	Qualitative	0	0	☒
CANONS	Discrete	0	0	☐

soumettre

Une fois le tableau intégré et le choix de l'analyse paramétré sur l'outil de description de variable, *Analyse* repère les variables de votre tableau de données et vous en propose une définition.

Le menu qui vous accompagne est contextuel. Il évolue en fonction du nombre et du type de variables sélectionnées.

À ce stade, vous pouvez :

❖1. Changer la qualité d'une variable quantitative

Les variables

Nouvelle analyseDescriptionACMR directement

AideRapportsGlossaire

Définition de la nature des variables à étudier (518 enregistrements)

Nom de la variable	Type	NA	NA en % des enregistrements	Choix de variable
LANCEMENT	Discrete	0	0	☐
PORT	Qualitative	0	0	☐
LAN	Qualitative	0	0	☐
TYPE	Qualitative	0	0	☒
CANONS	Discrete	0	0	☐

soumettre

Analyse vous offre la possibilité de changer la qualité des variables quantitatives. Vous pouvez ainsi transformer une quantitative discrète en quantitative continue, et inversement.

Cette fonctionnalité n'est pas permise pour les variables qui ont été reconnues comme qualitatives et dont les modalités d'analyse descriptives sont différentes.

Attention, les variables quantitatives discrètes doivent avoir comme séparateur de décimal le point (« 1.2 ») et non la virgule (« 1,2 »). La présence d'une virgule dans une cellule engendre la reconnaissance de la variable comme qualitative et non comme quantitative.

❖2. Les « na » (valeurs non attribuées)

Les variables

Nouvelle analyse	Description	ACM	R directement
------------------	-------------	-----	---------------

Aide
Rapports
Glossaire

Définition de la nature des variables à étudier (518 enregistrements)

Nom de la variable	Type	NA	NA en % des enregistrements	Choix de variable
LANCEMENT	Discrete	0	0	☐
PORT	Qualitative	0	0	☐
LAN	Qualitative	0	0	☐
TYPE	Qualitative	0	0	☒
CANONS	Discrete	0	0	☐

soumettre

Analyse vous permet également de connaître le nombre de valeurs non renseignées, en valeur absolue et en valeur relative, pour chacune variable de votre tableau de données.

❖3. Choisir les variables à décrire

Les variables

Nouvelle analyse	Description	ACM	R directement
------------------	-------------	-----	---------------

Aide
Rapports
Glossaire

Définition de la nature des variables à étudier (518 enregistrements)

Nom de la variable	Type	NA	NA en % des enregistrements	Choix de variable
LANCEMENT	Discrete	0	0	☒
PORT	Qualitative	0	0	☒
LAN	Qualitative	0	0	☐
TYPE	Qualitative	0	0	☒
CANONS	Discrete	0	0	☐

soumettre

Sélectionner les variables que vous désirez utiliser pour la suite de l'analyse descriptive.

❖4. Passer à l'analyse

Les variables

Nouvelle analyse	Description	ACM	R directement
------------------	-------------	-----	---------------

Aide
Rapports
Glossaire

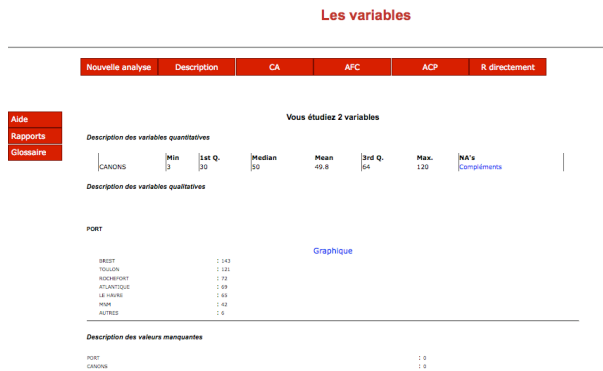
Définition de la nature des variables à étudier (518 enregistrements)

Nom de la variable	Type	NA	NA en % des enregistrements	Choix de variable
LANCEMENT	Discrete	0	0	☒
PORT	Qualitative	0	0	☒
LAN	Qualitative	0	0	☐
TYPE	Qualitative	0	0	☒
CANONS	Discrete	0	0	☐

soumettre

Il vous faut ensuite soumettre votre requête.

2. LA DESCRIPTION



❖ Pour les variables quantitatives

Analyse propose un résumé statistique de la variable comprenant :

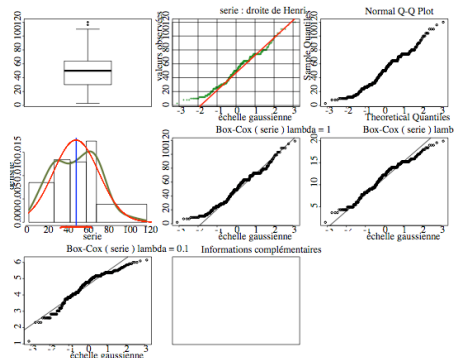
- Le minimum (Min)
- Les 1^{er} et 3^e quartiles (1st Q., 3rd Q.)
- La médiane (Median)
- La moyenne (Mean)
- Le maximum (Max.)
- Le nombre de na (NA's)

Description des variables quantitatives

	Min	1st Q.	Median	Mean	3rd Q.	Max.	NA's
CANONS	3	30	50	49.8	64	120	Compléments

Il vous également en « compléments » une série de graphiques :

- box plot
- droite d'Henri
- Q-Q plot
- Densité
- Box Cox



❖ Les variables qualitatives

PORT

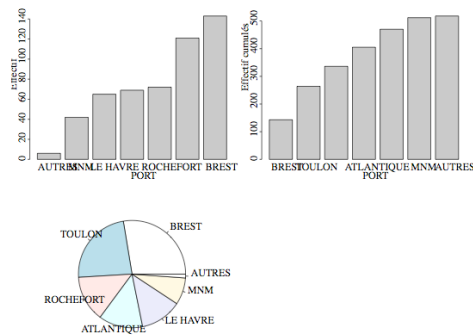
BREST	: 143
TOULON	: 121
ROCHEFORT	: 72
ATLANTIQUE	: 69
LE HAVRE	: 65
MNM	: 42
AUTRES	: 6

Graphique

La description des variables qualitatives vous renvoie la liste des différentes valeurs de la variable accompagné du décompte des effectifs d'individus correspondant à chacune de ces valeurs.

Cette description est, comme dans le cas des variables quantitatives, accompagnée de graphiques rendant visible les phénomènes statistiques :

- un diagramme en barre représentant les effectifs des différentes valeurs de la variable ordonnées du plus faible au plus important ;
- un diagramme en barre représentant les effectifs des différentes valeurs de la variable en effectifs cumulés ;
- un diagramme circulaire représentant les effectifs des différentes valeurs de la variable.



3. CONTINUER L'ANALYSE

Nouvelle analyse	Description	CA	AFC	ACP	R directement
------------------	-------------	----	-----	-----	---------------

ou

Nouvelle analyse	Description	ACM	R directement
------------------	-------------	-----	---------------

La suite de l'analyse se gère grâce à la barre supérieure de navigation :

1. « Nouvelle analyse » : vous renvoie à la page d'accueil d'*Analyse* afin que vous puissiez intégrer un nouveau tableau de données pour une nouvelle analyse.
2. « Description » : vous renvoie à la page « caractère et choix des variables pour analyse » afin de choisir de nouvelles variables dans votre tableau de données.

3. « CA » : Classement hiérarchique.
4. « AFC » : vous permet de demander à *Analyse* de faire une AFC sur les variables sélectionnées à l'étape précédente et qui font l'objet de la description. Attention, vous ne pouvez demander une AFC que si vous ne disposez, à cette étape, que de deux variables. R recompose alors un tableau de contingence avant de lancer une AFC.
4. « ACM » : vous permet de demander à *Analyse* de faire une ACM sur les variables sélectionnées à l'étape précédente et qui font l'objet de la description. Attention, vous ne pouvez demander une ACM que si vous disposez, à cette étape, d'au moins trois variables. R recompose alors un tableau de Burt avant de lancer une ACM.
5. « ACP » : Analyse des composantes principales. Cette option n'est disponible que si vous ne disposez que de deux variables.
6. « R directement » : cette option vous permet d'utiliser R directement en ligne de commande sur le tableau que vous avez envoyé à *Analyse*.

B. ANALYSE FACTORIELLE DES CORRESPONDANCES ET ANNALYSE DES CORRESPONDANCES MUTILPLES

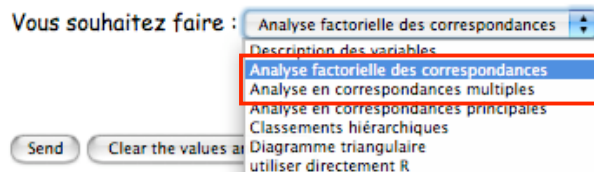
1. LANCER UNE ANALYSE FACTORIELLE

L'analyse factorielle des correspondances est disponible à toutes les étapes du processus de travail.

1. L'interface de configuration du tableau de données :

- dans ce cas l'analyse factorielle se fera automatiquement sur votre tableau de contingence ;
- l'analyse des correspondances multiples sur l'ensemble des variables de votre tableau de données.

Choisissez le type de tableau : -1- Tableau de comptage variables par variables (tableau de contingence)



2. L'interface de caractérisation et de choix des variables pour l'analyse ;

- vous pouvez ici lancer l'ACM directement sur l'ensemble des variables de votre tableau de données en cliquant simplement sur l'icône « ACM » ;
- vous pouvez également sélectionner les variables sur lesquelles vous voulez pratiquer l'ACM en taguant les cellules correspondantes.

Les variables

Nouvelle analyseDescriptionACMR directement

Aide
Rapports
Glossaire

Définition de la nature des variables à étudier (518 enregistrements)

Nom de la variable	Type	NA	NA en % des enregistrements	Choix de variable
LANCEMENT	Discrete	0	0	☐
PORT	Qualitative	0	0	☒
LAN	Qualitative	0	0	☒
TYPE	Qualitative	0	0	☒
CANONS	Discrete	0	0	☐

soumettre

3. et de l'interface de description des données

- l'ACM est lancée sur les variables qui ont été préalablement sélectionnées et qui font l'objet de l'analyse descriptive.

Les variables

Nouvelle analyse	Description	ACM	R directement
------------------	-------------	------------	---------------

Aide	Vous étudiez 3 variables Description des variables qualitatives
Rapports	
Glossaire	

2. L'INTERFACE D'ANALYSE FACTORIELLE

❖ Paramètre du graphique

La première partie de l'interface d'analyse factorielle se présente sous la forme d'un tableau contenant l'ensemble des éléments paramétrables pour modifier le graphique factoriel :

Nouvelle analyse	CA	AFC	ACP	R directement
------------------	----	------------	-----	---------------

Tableau de paramétrage de l'analyse : AFC					
Aide Rapports Glossaire	1	Titre du graphique			Il y a 0 fichiers
	2	Nombre de facteurs (max.8)	2		
	3	Choix des facteurs	1	2	
	4	Seuil moyen en colonne (val. moyenne 12.5)	12.5		
	5	Seuil moyen en ligne (val. moyenne 0.5)	0.5		
	6	Choix des variables supplémentaires en ligne	il nous est tous ils ne foutre vous je on a ont *sans sont faire culottes elle tout République tre		
	7	Choix des variables supplémentaires en colonne	Formes juillet aout septembre octobre novembre décembre janvier février		
	8	Modification			

Souhaitez-vous le code R

1. « Titre du graphique » : cette cellule vous permet d'introduire un titre qui apparaît ensuite directement en dessous du graphique factorielle
2. « Nombre de facteurs » : vous permet de déterminer le nombre de facteurs dont vous souhaitez voir la description dans les tableaux se trouvant en dessous du graphique factorielle (ces tableaux décrivent le facteur et vous donne les coordonnées ainsi que les contributions de chaque valeur des variables prises en compte pour l'analyse)
3. « Choix des facteurs » : il s'agit de sélectionner les facteurs que vous désirez observer sur le graphique factoriel. Par défaut le facteur 1 apparaît à l'horizontal, et le facteur 2 à la vertical.
4. « Seuil moyen en colonne (valeur moyenne X) » : vous pouvez déterminer la valeur contributive minimum requise pour l'affichage d'un élément en colonne sur le graphique factoriel

(on parle de valeur seuil, ou plus simplement de seuil). Par défaut, *Analyse* choisit comme seuil la valeur moyenne de contribution en %.

5. « Seuil moyen en ligne (valeur moyenne Y) » : idem mais avec les éléments en ligne.

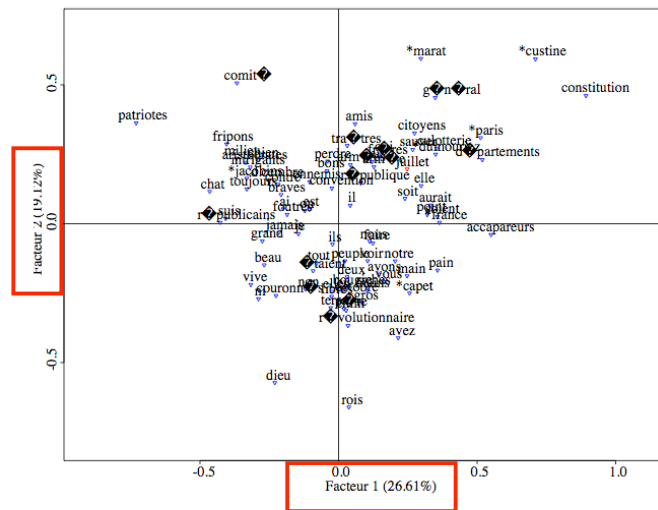
6. « Choix des variables supplémentaires en ligne » : Lorsque la contribution d'un élément en ligne est écrasante, *Analyse* vous propose d'en faire une « variable supplémentaire ». Le calcul de l'analyse est recommencé sans prendre en compte cette valeur, qui est ensuite replacée sur le graphique à la place qu'elle aurait du y occuper.

7. « Choix des variables supplémentaires en colonne » : idem mais avec les valeurs en lignes.

8. « Modifier » : bouton d'application des modifications.

❖ Le graphique factoriel

Le graphique factoriel vous indique le numéro des facteurs et leur valeur propre en pourcentage.



❖ Les informations sur l'analyse factorielle

Analyse met à votre disposition toutes les informations nécessaires à l'interprétation de votre analyse factorielle.

1	X-squared :			
	3391.219			
2	df :			
	1386			
3	p.value :			
	7.69592e-169			
	test :			
	Pearson's Chi-squared test			
	data :			
	AFC			
4		Valeur propre	% exp	% cumulé
	facteur 1	0.03	26.61	26.61
	facteur 2	0.02	19.12	45.74
	facteur 3	0.02	16.13	61.87
	facteur 4	0.01	13.76	75.63
	facteur 5	0.01	10.49	86.12
	facteur 6	0.01	7.47	93.59
	facteur 7	0.01	6.41	100

1. « X-squared » : Le χ^2 de votre tableau de contingence, ou du tableau recomposé à partir des variables sélectionnées dans votre tableau de données.
2. « df » : Le degré de liberté (degree of freedom) du tableau de contingence ou du tableau de données.
3. p.value : valeur probable de l'hypothèse nulle, elle permet de déterminer l'intervalle de confiance dans lequel se place votre analyse factorielle.
4. La description des facteurs : en fonction de leur valeur propre, de la valeur propre en % exprimé du χ^2 , et en % cumulé du χ^2 .

3. LES TABLEAUX DE COORDONNEES ET DE CONTRIBUTION

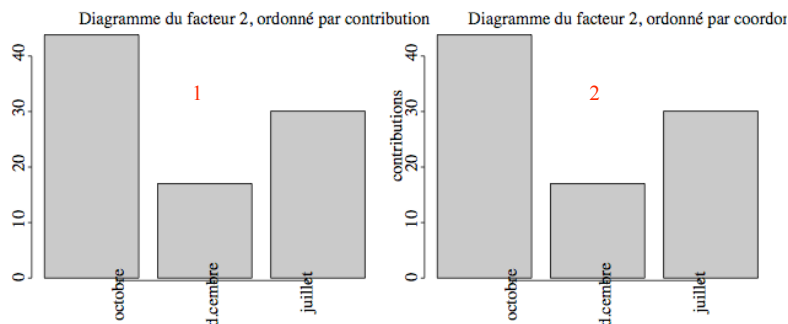
Analyse vous fournit également les tableaux contenant les coordonnées et les contributions des différents éléments en ligne et en colonne pour chaque facteur.

modalité en colonne				
Sauvegarder	1		2	
	coord	contrib.	coord	contrib.
juillet	0.247	34.139	0.197	30.067
août	0.129	9.235	-0.052	2.084
septembre	0.089	4.233	0.002	0.002
octobre	0.044	0.929	-0.255	43.813
novembre	-0.157	12.276	0.012	0.101
d.cembre	-0.249	30.528	0.157	17.002
janvier	-0.06	1.645	-0.104	6.835
f.vrier	-0.139	7.015	0.014	0.097

Pour vous aider dans l'interprétation, *Analyse* fait apparaître pour chaque facteur :

- en rouge les contributions supérieures à la moyenne des contributions dont la coordonnée positive ;
- en bleu les contributions supérieures à la moyenne des contributions dont la coordonnée est négative.

Vous avez également accès, en cliquant sur le numéro du facteur, aux diagrammes de contribution.



1. Le premier diagramme est ordonné en fonction de la valeur des contributions et permet de mettre en évidence les grandes oppositions.
2. Le second diagramme est ordonné en fonction des coordonnées et permet de visualiser des regroupements de chaque côté de l'origine du facteur.