



**Kharchaf Idris**

**Coran**

**&**

**Data Mining**

**RABAT 5-7-2020**

# **Coran & Data Mining**

**D<sup>r</sup> KHARCHAF IDRIS**

Professeur de l'enseignement supérieur en Analyse des données (Data Mining)

Département de Mathématiques Informatique, Faculté des sciences, Université M<sup>d</sup> V, Rabat

Maroc

# Table des matières

**Partie I** : Avant-propos

**Partie II** : Introduction aux fouilles de données

**PII :1** : Présentation

**PII :2** : Objectifs

**PII :3** : Plan d'action

**PII :4** : Data Mining

**PII :5** : A quoi sert le D.M.

**PII :7** : D.M. & résolution de lacunes.

**PII :8** : Espace stratégique

**PII :9** : Nature de l'information

**PII :10** : Critères

**PII :11** : les difficultés rencontrées

**PII :12** : Algorithme de D.M

**Partie III** :

**Partie : III:1**: Analyse et fouille de données

**Partie : III:2**: Analyse de données coraniques : Théorique & Pratique

**Partie : III:3**: Représentation des facteurs, Discussions

**Partie IV** : Annexe

**PIV : A**/ Coran.

**PIV : B**/ Mathématiques et Analyse factorielle de correspondance.



## AVANT-PROPOS

La recherche dont s'inspire le présent travail a pour objectif principal, la rationalisation de la raison, et la compréhension objective du Coran, de manière nouvelle et à l'aide de moyens scientifiques, afin que le lecteur non averti puisse comprendre la structure du Coran et son contenu, quelles que soient sa religion et ses affiliations idéologiques.

Avec l'expansion du monde numérique et l'intelligence artificielle et sa connaissance ininterrompue, cibler le Coran et l'école prophétique devient une nécessité civilisationnelle.

La saisie des données du Coran sur un support technique en vue de leur analyse et traitement automatique, ne peut pas être le fruit du hasard pour suivre un mode, et elle ne peut être décidée qu'en vertu d'une réflexion professionnelle, et d'une stratégie bien définie des recherches dans lesquelles elle s'intègre ou qu'elle va déterminer.

Pour cette raison, on a fait recours aux techniques de Data Mining (Analyse factorielle des correspondances, Classification hiérarchique, Analyse en composantes principales), qui doit précisément être le facteur principal et scientifique qui clarifie le sens objectif du Coran, afin de ne pas projeter sur le Coran la subjectivité du lecteur.

Ceux qui sont étrangères à l'esprit mathématiques s'étonneront d'une telle approche originale et féconde, à cause du langage mathématique avec lequel ils ne sont pas familiers. Ainsi, ils peuvent prendre l'attitude dogmatique du refus arbitraire et gratuit.

La matrice utilisée comme base de données (input) est constituée de 60 Hizbs (240 Rubu's), pouvant être décrits à partir d'un nombre fini de mots, cette matrice est représentée par le tableau  $K(M, H)$  où  $M = \{m_{i=1,340}\}$ , et  $H = \{h_{(j,t)/j=1,60 ; t=1,4}\}$ .

La difficulté essentielle de l'analyse réside dans le fait que, le choix des mots à chaque essai est la résultante d'un mécanisme de recherche et de traitement de l'information (mots qui contribuent à la formation du tableau des données).

Pour résoudre cette difficulté, on doit traiter la matrice de plusieurs façons (voir plus de quatre analyses).

Les résultats de recherche montrent ce qui suit :

- 1- La foi islamique ne s'oppose pas à la science, bien au contraire, elle a consolidé le lien entre ce qui est matériel à ce qui est moral, en prescrivant à chaque musulman et musulmane le devoir de la science.
- 2- La recherche illustre à tort la notion que le Coran est écrit par cinquante écrivains ou plus.
- 3- Le Coran demande à chaque chercheur de rationaliser sa raison, d'utiliser les dimensions civilisationnelles et humaines (Perception de Soi, Valeurs, Relations humaines, Expression individuelle, Prise de décision, Créativité positive)
- 4- Les résultats de l'exploration scientifique des données nous montrent que le Coran est une unité cohérente et compacte.
- 5- Cet article montre que le Coran est basé sur trois composantes :
  5. a : le côté rationnel (la contemplation de l'univers de Big Bang au monde numérique et au-delà de numérique, le voyage dans le monde infiniment petit et dans le monde infiniment grand),
  5. b : le côté Étique (représenté par les enseignements des prophètes).
  5. c : le côté législatif (engendré par la prière, le pèlerinage, les croyants, etc. structure soutenue par la plupart des spécialistes de la charia)
- 6- L'importance de cette recherche peut être résumée en quelques points:
  - 6.1. Authenticité (Il s'agit d'un travail original, lancé en 1978 à Paris au centre

de calcul Électronique d'Orsay, CNRS, Paris).

6.2. Travail rationnel et transparent.

6.3. Découverte de la science prioritaire en Islam (فقه الأولويات في الإسلام)

6.4. Preuve de la sainteté du Coran.

6.5. Moyen efficace pour faire comprendre le Coran aux non-musulmans.

## **PII :1 :Présentation**

Cette recherche rationnelle espère interroger la place du numérique dans l'analyse du Coran, S'appuyant tant sur des contextes réels directes que sur un ensemble d'études approfondies en Analyse des données.

Il convient alors que la recherche scientifique rationnelle, ouvre le débat sur la statistique appliquée comme instrument nécessaire à l'élaboration de la science, et spécialement son application à des disciplines aussi variées que la sociologie, l'analyse du Coran, l'analyse de contenu du discours prophétique à ARAFAT, la théorie de l'information, les élections législatives, la reconnaissance des formes, l'analyse des préférences, l'analyse des données biologiques, géologiques linguistiques médicales ou autres, fait ressortir son intérêt.

L'analyse des données a permis de tirer parti de cette information pour la synthétiser, ou pour appréhender la nature des phénomènes sous-jacents aux données.

## **PII :2 : Objectifs**

Les objectifs à atteindre par cette recherche sont :

- L'objectif principal de ce travail est l'étude statistique des faits humains, mais le cerveau humain ne peut accomplir une synthèse multidimensionnelle, sans faire de nombreux choix arbitraires, qui ôtent souvent toute signification au résultat, il faut donc l'aide de l'ordinateur et les techniques du Data Mining (dans le monde numérique), qui permettrons le traitement jusqu'au bout de volumineux calculs et données (Big Data) telles qu'on puisse lire avec sûreté à la sortie (output) qui, à l'entrée (input) étaient indéchiffrables.
- La réalisation de cette recherche devra apporter à l'homme universel compréhension, souplesse, explication des sigles, rapidité et la créativité.

- Porter assistance aux chercheurs dans leur quête d'informations dispersées sur les deux livres (coranique et prophétique).
- Amener au chercheur une information conforme à son centre d'intérêt.
- Finalement, les recherches sont par la suite orientées vers l'identification et la définition de problématiques spécifiques dans un cadre multidisciplinaire.
- Pour comprendre la relation entre le Coran et Data Mining, il faut tout d'abord connaître le Coran et son histoire dans l'espace-temps humain.

### **PII :3: Introduction**

Le travail présenté ici, est basé sur l'étude de l'analyse du Coran (la technique de Fouille de Données : Data Mining). L'originalité de l'approche mathématique réside dans la nature de raisonnement, Raisonner pour comprendre, comprendre pour savoir, (سأعرف في الأرض) savoir pour croire et croire pour être lieutenant sur terre (Khalifa

La saisie des données sur un support technique en vue de leur analyse et traitement automatique, ne peut pas être le fruit du hasard pour suivre une mode, et elle ne peut être décidée qu'en vertu d'une réflexion rationnelle, et d'une stratégie bien définie des recherches dans lesquelles elle s'intègre ou qu'elle va déterminer.

### **PII: 3 : Plan d'action : QQQQCP**

L'une des questions importantes qui doivent être soulevées, à chaque recherche scientifique expérimentale, est la création d'une plate-forme professionnelle permettant de définir le chemin de but désiré.

Donc, il s'agit de mettre en place un plan d'action pour maîtriser toutes les tâches, afin d'aboutir à une explication transparente concise et rationnelle du Coran.

Ce plan d'action (Tab.1), est un outil d'aide à la résolution de problèmes, et il permet de décrire une base de données, en répondant aux questions suivantes : Quoi, Qui, où, Quand, comment, Pourquoi.

| <b>Questions</b>  | <b>Forme de la question</b>         | <b>Cible</b>   |
|-------------------|-------------------------------------|--|
| Quoi?             | C'est quoi le problème ?            | Il s'agit de la relecture logique et rationnelle du Coran afin de comprendre son contenu, en faisant recours aux techniques de l'intelligence artificielle (Ex : Data Mining ).  |
| Qui ?             | <b>Qui</b> est concerné ?           | Concerne les gens, les chercheurs et les scientifiques qui recherchent la vérité universelle, quelles que soient leurs croyances religieuses, idéologiques ou géographiques  |
| <b>où</b>         | <b>où</b> apparaît le problème ?    | Quand on fait de la recherche scientifique dans le domaine du Coran, que ce soit dans des laboratoires universitaires, ou dans des centres de recherches scientifiques ou sur le terrain.  |
| <b>Quand ?</b>    | <b>Quand</b> apparaît le problème ? | Lorsque l'homme « intellectuel » ignore les connaissances techniques, éthiques et scientifiques du Coran.  |
| <b>Pourquoi ?</b> | <b>Pourquoi</b> repenser le Coran   | Une lecture rapide de ce qui se passe dans l'esprit de l'homme ordinaire montre, qu'il existe un fossé artificiel pour empêcher l'être humain de lire le Coran. Ainsi, il pense que le Coran n'a pas de zone géographique dans la marche civilisationnelle de l'être humain moderne. |

|                  |   |   |
|------------------|---|---|
| <b>Comment ?</b> | <b>Comment</b> mesurer le problème et ses solutions ? | La construction d'une base de données, d'une façon rationnelle et technique en vue de leur analyse et traitement automatique, ne peut pas être le fruit du hasard, et elle ne peut être décidée qu'en vertu d'une réflexion responsable et transparente. Les méthodes quantitatives, lorsqu'elles sont bien conduites, assurent aux chercheurs une méthodologie objective logique et transparente de l'analyse, dans la phase de la recherche professionnelle |
|------------------|---|---|

[Tab.1]

#### **PII :4 : Data Mining (Définition)**

Le Data Mining (de point de vue mathématique : AFC) a pour but d'approcher, par des variétés géométriques (droites, plans, corps, etc.), les nuages des profils d'un tableau de données brutes (linguistique, Biologie, géologie, accidents de la route, médecine, anthropologie, archéologie, Enseignement, Ethique, etc.), afin d'en révéler la structure globale de base des données.

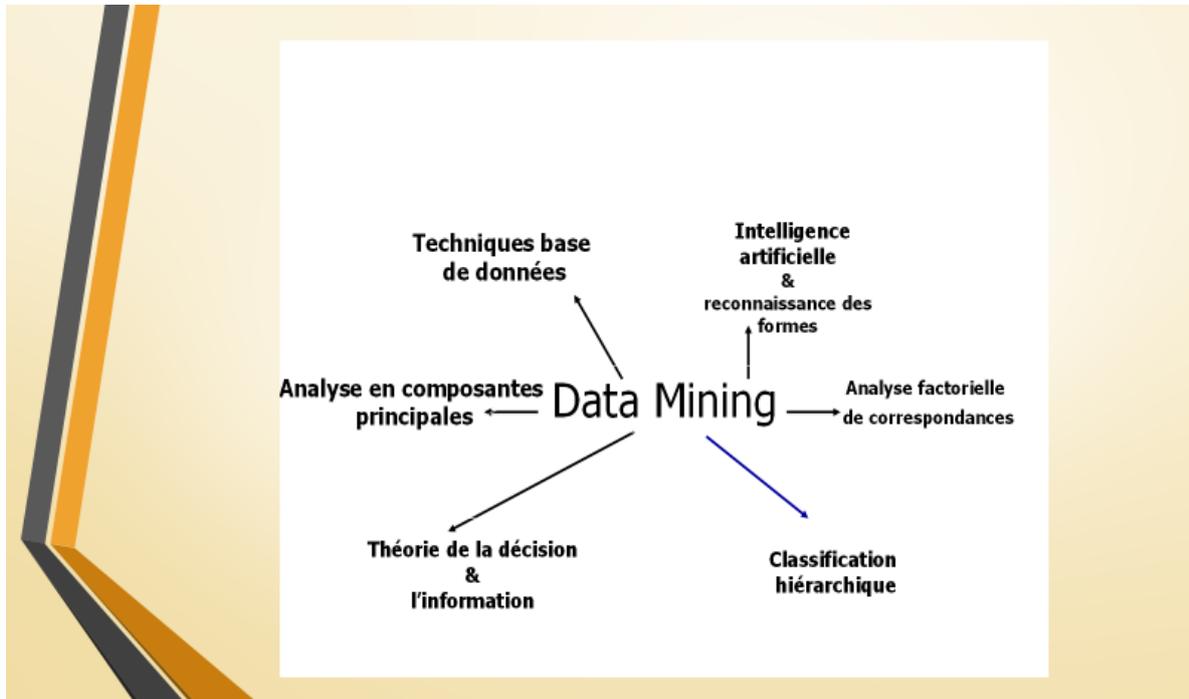
#### **PII :5 : A quoi sert le Data Mining (Définition)**

C'est un ingrédient crucial pour les chercheurs, et un outil d'exploration des données concrètes et efficaces.

Le **Data Mining** est en fait un terme récent (1990), englobant toute une famille d'outils, facilitant l'exploration et l'analyse des données contenues dans différentes perspectives, et le fait de transformer ces données en informations utiles, en établissant des relations entre les données.

## PII :6 : DIFFERENTES TECHNIQUES DE DATA MINING

Le coût d'une telle opération et les contraintes qu'elle implique, sont si importants que seule l'analyse minutieuse de la démarche, et la claire conscience des conséquences, qui vont en découler doivent être maîtrisées par avance, pour que cette action soit entreprise avec un minimum de risque et un maximum de succès.



[Fig.1] Techniques de D.M.

## PII :7 : Data Mining & résolution de lacunes

De même, les mathématiciens-statisticiens notamment les utilisateurs de **Data Mining** (étape clé de l'extraction des connaissances) sont fréquemment confrontés au problème suivant :

Le tableau de leurs données est, pour une raison ou pour une autre soit très grand (Big Data). De telles données traitées sans précautions ne conduiront probablement pas à des résultats satisfaisants.

Lorsqu'il est possible de réduire les ensembles de données sans pour autant affaiblir la compréhension de la correspondance, (Exemple : Analyse factorielle des correspondances, méthode de classification hiérarchique : construction des classes de Base de données), la difficulté précédemment signalée peut-être surmontée, en mettant les colonnes ou les lignes d'un tableau de données en élément supplémentaires, que l'on sait ou que l'on ne connaît pas.

Une telle démarche revient cependant à abandonner une partie de l'information ou plutôt certains aspects des problèmes étudiés.

La solution est donc de lever cette difficulté, tentant de mesurer l'angle intercepté par les plans engendrés par les facteurs trouvés (vecteurs propres) à chaque analyse, et de chercher l'angle de déviation(D), lorsque D est minimum.

### **PII :8 : l'Espace stratégique de data mining [E.S.D.M.]**

Pour réaliser une recherche positive, rationnelle, étiq, et responsable dans le domaine de data mining, le chercheur doit manipuler ses techniques dans un espace ambiant stratégique (appelé : espace stratégique de data mining) engendré par trois composantes :

8 :1. La connaissance rationnelle universelle (l'algèbre linéaire, désignées sous le nom de méthodes factorielles, et les techniques de classification automatique), la capacité de raisonner et de résoudre des problèmes dans des situations nouvelles (Ex. l'intelligence fluide, l'intelligence émotionnelle).

8 :2. L'Ethiques de la science.

8 :3. L'Innovation positive.

## PII :9 : Nature de l'information

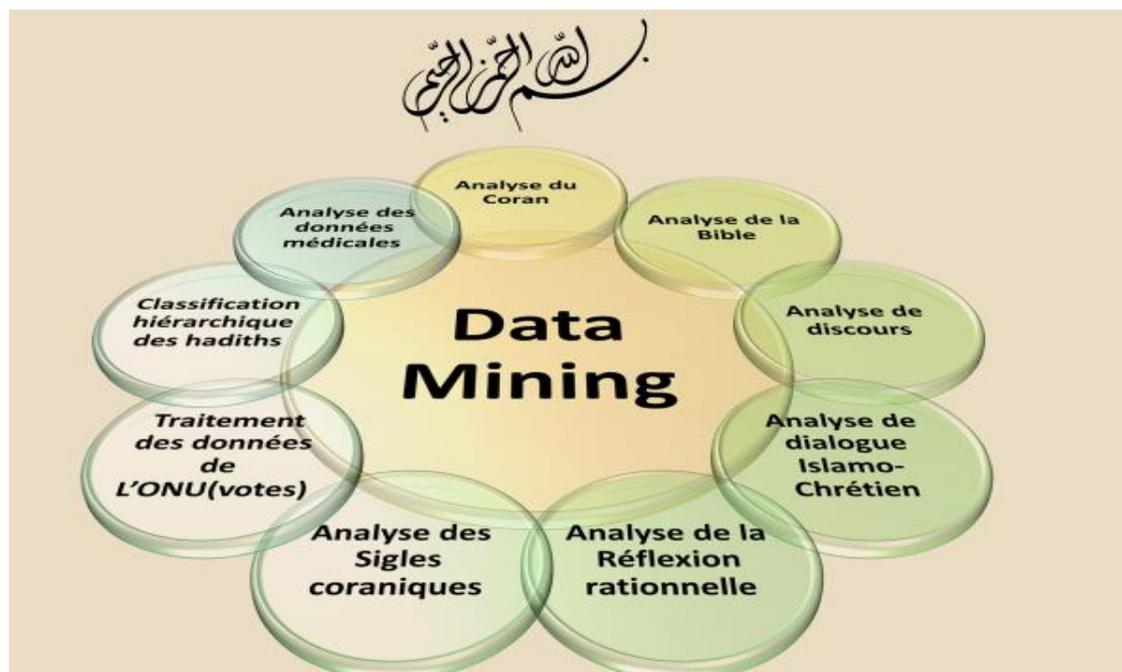
En Data Mining, l'information est une chose capitale, elle se présente généralement sous la forme suivante :

Un certain nombre d'informations sont décrites par un nombre de variables (ou caractères), les variables peuvent être quantitatives ou qualitatives (Fig :2) on procède alors à un codage.

Avec l'ensemble d'informations, on constitue le tableau de données qui doit répondre à deux conditions :

9. a/ **homogénéité** : toutes les grandeurs recensées dans le tableau ont la même unité de mesure (sinon transformer les données hétérogènes en données homogènes) .

9.b/ **exhaustivité** (précise et complète) : les ensembles M (Ensemble des individus ou des mots) et V (ensemble des variables ou des versets), doivent refléter exactement les données à analyser.



(Fig.2) quelques exemples de réussite en matière de Data Mining.

Exemple : Prenons une technique de Data Mining (l'Analyse Factorielle des correspondances d'une matrice appelée matrice de fréquence), qui permet d'étudier

l'association entre deux variables qualitatives (M : l'ensemble de mots coraniques en lignes) et (S : l'ensemble de sourates coraniques en colonnes), étant deux ensembles finis, alors la matrice construite est une matrice rectangulaire de nombres positifs (ou à (m) lignes et (s) colonnes).

Donc le tableau de données obtenu répond aux deux conditions citées (ci-dessus).

A l'intersection de la ligne  $m_i$  et de la colonne  $v_j$ , on trouve  $k(m_i, v_j)$  qui représente la fréquence absolue du mot coranique  $m_i$  dans le verset  $v_j$ .

Il s'agit d'une méthode qui permet d'étudier l'association entre deux variables (quantitatives ou qualitatives) (voir Tab.2).

|          |       |       |               |  |            |       |           |
|----------|-------|-------|---------------|--|------------|-------|-----------|
| Sourates | $s_1$ | $s_2$ | $s_3$         |  | $s_\alpha$ | ..... | $s_{114}$ |
| Mots     |       |       |               |  |            |       |           |
| $m_1$    |       |       | :             |  |            |       |           |
| $m_2$    |       |       | $k(m_2, s_3)$ |  |            |       |           |
| :        |       |       | :             |  |            |       |           |
| $m_t$    |       |       | :             |  |            |       |           |
|          |       |       | :             |  |            |       |           |

(Tab.2)  $k(m_2, s_3)$  représente à la croisée de la ligne  $m_2$  et de la colonne  $s_3$ , alors  $k(m_2, s_3)$  représente la fréquence du mot  $m_2$  dans la sourate( $s_3$ )

### PII :10 : Critères

Parmi les critères qui conduisent au choix des moyens informatiques deux noms ont semblé essentiels :

#### a/ La potentialité :

La saisie des données, suivie par une base de données (nous désignons par l'expression base de données, les corpus textuels saisis exhaustivement sous leur forme séquentielle, sans aucune perte d'information paléographique).

#### b/ L'objectivité :

L'objectivité est assurée par la neutralité du système de codification qui, doit être conçu de façon large pour conserver toute l'information, dont le corpus textuel est porteur au niveau du signifiant, sous lequel il se présente dans le manuscrit, de façon à pouvoir répondre à toutes les interrogations auxquelles le texte pourra être soumis dans l'avenir.

Les méthodes quantitatives, lorsqu'elles sont bien conduites, assurent aux chercheurs une méthodologie objective de l'analyse dans la phase de la recherche.

### **PII :11 : Les Difficultés rencontrées en Data Mining :**

En analyse textuelle, on se trouve placé face à deux situations distinctes :

A -Dans le premier cas, on a affaire à une information faite d'un très grand nombre d'unités discrètes (Big Data), telles les occurrences des mots, et, dans ce cas, on peut utiliser l'outil statistique (Histogramme, corrélation, théorie de l'information, etc.) qui est destiné à l'étude des relations entre des phénomènes mesurés sur des vastes ensembles.

B -Dans le second cas, on se trouve devant des petits nombres, et la question se pose de savoir si l'outil statistique peut être employé en critique textuelle, sans risquer de fausser l'analyse. Et le problème sera de définir le niveau d'intersection de l'outil statistique dans l'analyse textuelle, car la dichotomie qui oppose, en analyse textuelle, les données existant en grand nombre (Big Data), et les données qui n'existent qu'en petit nombre, pose les limites des méthodes d'analyses statistiques, dites « paramétriques » ; Et pour résoudre les problèmes textuels, des méthodes d'analyses et de classifications mises au point par le Professeur Jean Paul Benzecri (Directeur de laboratoire de l'analyse des données,(ISUP) Institut de Statistique de l'Université de Paris, Paris VI, Jussieu, FR).

### **PII :12 : Algorithme de Data Mining et la Connaissance rationnelle du Coran.**

Partant d'un model réel de données coraniques, on introduit les étapes principales de Data mining (Analyse factorielle des correspondances).

1. Définir les objectifs de l'analyse du Coran
2. Comprendre les données coraniques
  - 2 :1 : recueillir l'espace ambiant et sélectionner les données (Sourates, Hizbs, Versets, sens de versets, etc.)

2 :2 : Déterminer des renseignements valables

3 : détecter la nature de l'algorithme

4 : Préparer les données de base : codage et recodage des Données (Découpage en classes, dédoublement d'un tableau, transposition, disjonctive complète...)

4 :1 : homogénéiser les données (transformer les données de bases en formes appropriées : voir les tableaux de données : logique, contingence, dédoublées, Burt, etc.)

4 :2 : préparer à partir des données de base, celles qui seront utilisés.

5 : Choisir les techniques info-mathématique à appliquer.

5 :1 : Calculer les profils marginaux de la matrice des données.

5 :2 : diagonaliser la matrice des données (Calculer les valeurs propres, les vecteurs propres, et tirer les facteurs des nuages :  $\mathcal{N}(f^m_V)$  et  $\mathcal{N}(f^V_M)$ ).

5 :3 : choisir la métrique (distance au sens de  $\chi^2$  entre les différents éléments des nuages  $\mathcal{N}(f^m_V)$  et  $\mathcal{N}(f^V_M)$ , indice de similitude, information, etc).

6 : Programmer la structuration

On peut choisir les méthodes suivantes :

6 :1 : Analyse factorielle des correspondances.

6 :2 : classification hiérarchique (Ascendante, descendante).

6 :3 : Analyse en composantes principales.

7 : Représenter graphiquement le nuage de M :  $\mathcal{N}(M) = \{ (f^m_s, f_m) / m \in M \}$ , et le nuage de V :  $\mathcal{N}(V) = \{ (f^V_m, f_v) / v \in V \}$ .

8 : Analyser les résultats

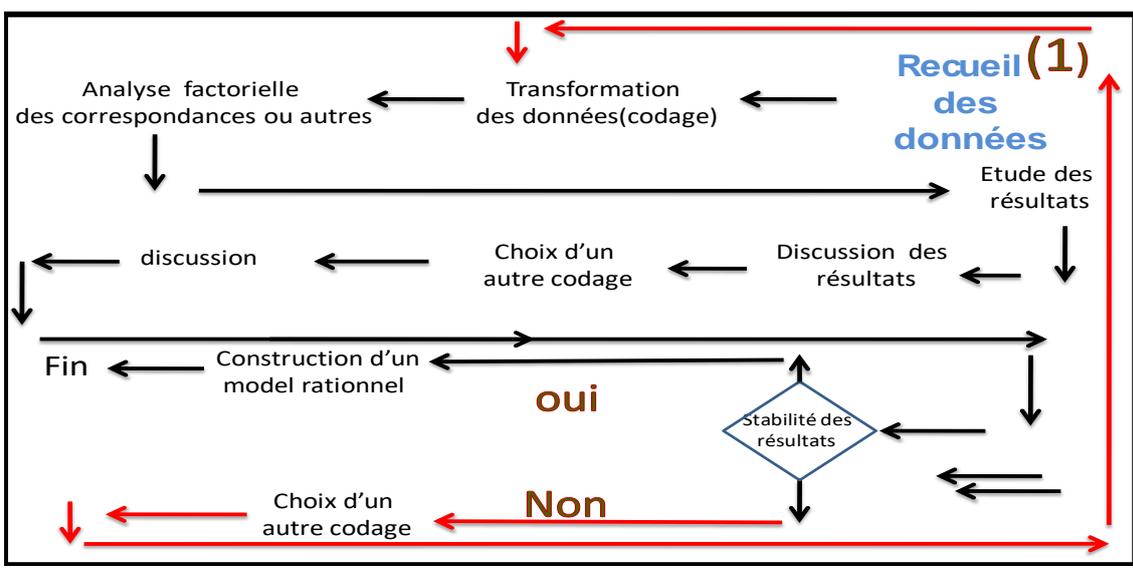
L'analyse des données n'est pas seulement la description d'un tableau numérique (matrice numérique), c'est une manière d'extraire une connaissance, et de produire le savoir.

9 : Evaluer la qualité et la crédibilité des résultats.

9:1 : retour au codage et changement de celui-ci pour des causes de précision ou d'aberration.

- 9 :2 : retour aux données brutes et éventuellement changement de codage dans le cas de non précision.
- 10 : Refaire d'autres analyses pour pouvoir déterminer des méthodes et des résultats valables en fonction de l'objectif défini.
- 11 : Modéliser le résultat [Fig.3].

*Algorithme*



[Fig.3]

**Partie III : Analyse et fouille de données coraniques (f.d.c)**

Le Coran a été divisé en trente parties (جزء) en une époque ancienne, plus récemment, le texte coranique a été partagé en soixante Hizb (حزب), chacun de ces Hizb est divisé sur quatre quart.

Le Hizb contient un ou plusieurs sourates, et chaque sourate est composée de plusieurs versets (Ayats : آيات) (voir l'explication plus loin).

A l'heure actuelle, « Ayat » signifie strictement « verset » dans la vulgate cairote, ou l'on compte près de 6236 versets (version Madani), et 6208 versets (version Koffi).

### **Longueur des versets coraniques**

La longueur de ces versets est fort variable :

a/Dans les révélations médinoises, elle atteint plusieurs lignes (parfois plus de 5 lignes)

b/Dans les révélations mecquoises, au contraire, il est fréquent d'avoir des versets de deux mots (voir un mot).

Les sourates sont au nombre de cent quatorze, elles sont d'inégale étendue, la plus longue contient plus de cinquante pages (cas de la deuxième sourate : la Vache), et la plus courte contient deux lignes seulement (exemple : l'instant, abondance, assistance), et ce qui est étrange ici, c'est que le rangement par ordre de longueur décroissante n'a pas été nettement observé, c'est-à-dire le rangement ne tient absolument pas compte de leur ordre chronologique de révélation.

### **PIII :1 : Traitement des données coraniques**

L'étude qui a été faite porte sur la fréquence des concepts (mots) qui existent dans le Coran. Cette fois-ci, nous avons découpé le Coran selon les Hizbs (60 Hizbs) qui sont découpé à leur tour, chacun en quatre Rubu' (paragraphe), nous avons pris les mots qui ont une bonne représentation théologique, ou selon leur degré de fréquence.

NB. Le coran se constitue de 30 Jouze(جزء), et chaque Jouze se consitue de deux Hizb(حزب), donc le nombre totale de **hizb** est de **60** .

Exemple : le cinquante quatrième Hizb (H (54)) est découpé en 4 Rubu's sous forme d'intervalles :

Hizb n°54: de la Sourate Arrahman, numéro 55, verset 1, à la Sourate Al Hadid, numéro 57, verset 29.

Notation :  $H(54) = [1, 77] S_{55} \cup [1, 77] S_{56} \cup [78, 14] S_{56} \cup S_{57} \cup [15, 29] S_{57}$

$$H(54) = H(54)_1 \cup H(54)_2 \cup H(54)_3 \cup H(54)_4$$

Où  $[a, b]_{S\alpha}$  { a, b : représentent les bornes inférieure (a), et supérieure (b) de la sourate S  
 $\alpha$  : numéro de la sourate S  
 et  
 $H(n^\circ) \beta$  { H : indique le hizb  
 $(n^\circ)$ : représente le numéro du Hizb H  
 $\beta$  : indique le n° du RUBU' (n°1, 2, 3, 4)

Application :  $H(54)_2$  est le deuxième Rubu' du Hizb n°54, qui commence par le verset n°1 de la sourate n°56, et se termine par le verset n° 77 de la même sourate (voir E1).

Les données que nous avons recueillies constituent un tableau ternaire :

$$K(R, M, H) = \{k(r, m, h) / r \in R ; m \in M ; h \in H \}$$

Où R : Rubu' (240 Rubu')  
 M : mots (340 mots)  
 H : Hizbs (60 Hizbs).

PIII :2 : **Liste de quelques mots utilisés** (voir l'ensemble M ci-dessous)

M={ mille(ألف), barrière(برزخ), museau(الخرطوم), Moucheron(بعوضة); menstruation (المحيض), planètes (الخنس), Sujet(الحديث), ceux qui observent (حافظون), étoile(النجم), brillante(الثاقب), Pain(الخبز), bête étouffée (المختنقة), Péché(ذنب), rêve(حلم), lorsque la terre sera violemment secouée (إذا رجت الأرض رجا), Six (سنة), année (سنة), Récits(القصص) mois(شهر), le grand bruit(الراجفة), livre(كتاب), Dou Alqarnaïn (ذو القرنين) le grand bruit

retentira(الراجفة) , Etoiles (نجوم) ,fécondés (لواقح) ,Ramadan(رمضان), Cent(مائة) , viande (لحم), individus(أناس) , Sommeil (النوم) , mâchée de chaire(مضغة), Lampe ,(مصباح),etc.}(Tab.3)

| <b>Mots</b>   |  | $m_1$ | $m_2$ | $m_3$ | $m_4$ | ..... $m_\beta$   | ..... | ..... | $m_{340}$ |
|---------------|--|-------|-------|-------|-------|-------------------|-------|-------|-----------|
| <b>Rubu's</b> |  |       |       |       |       |                   |       |       |           |
| $r_1$         |  |       |       |       |       |                   |       |       |           |
| $r_2$         |  |       |       |       |       |                   |       |       |           |
| $r_3$         |  |       |       |       |       | $k(r_3, m_\beta)$ |       |       |           |
| :             |  |       |       |       |       |                   |       |       |           |
| :             |  |       |       |       |       |                   |       |       |           |
| :             |  |       |       |       |       |                   |       |       |           |
| $r_{239}$     |  |       |       |       |       |                   |       |       |           |
| $r_{240}$     |  |       |       |       |       |                   |       |       |           |

$k(r_3, m_\beta)$  représente la croisée de la ligne  $r_3$ , et de la Colonne  $m_\beta$ , d'une manière plus particulière si on désigne par  $r_3$  le 3<sup>ème</sup> Rubu', et le mot d'ordre  $\beta$ , alors  $k(r_3, m_\beta)$  sera donc le nombre de fois que se trouve le mot  $m$  d'ordre  $\beta$  dans le Rubu'3 d'ordre  $\beta$  [**Tab.3**]

Ce tableau (Tab.3) tri dimensionnel ne se prête pas tel quel, à l'analyse factorielle des correspondances qui traite habituellement des binaires.

Mais on peut construire à partir du tableau binaire cumulé en juxtaposant les Hizbs :  $\{H_1, H_2, H_3, \dots, H_{60}\}$

Le tableau obtenu n'est autre que le tableau  $K((R^*H) * M)$  lui-même, mais considéré comme produit de deux ensembles dont l'un est l'ensemble  $R^*M$  et le second représenté par l'ensemble H.

**PIII :3 : Liste des Hizbs (Tab.4)**

| <b>Hizb</b> | <b>Intervalle1</b>                              | <b>Intervalle 2</b>                      | <b>Intervalle 3</b>                      | <b>Intervalle4</b>                      |
|-------------|---|--|--|---|
| 1           | [1,24] <sub>s<sub>1</sub> ∪ s<sub>2</sub></sub> | [25,40] s <sub>2</sub>                   | [41,58] s <sub>2</sub>                   | [59,74] s <sub>2</sub>                  |
| 2           | [75,90] s <sub>2</sub>                          | [91,104]s <sub>2</sub>                   | [105,122]s <sub>2</sub>                  | [123 ,140]s <sub>2</sub>                |
| 3           | [141,156]s <sub>2</sub>                         | [157,175]s <sub>2</sub>                  | [176,187]s <sub>2</sub>                  | [188,200]s <sub>2</sub>                 |
| 4           | [201,216]s <sub>2</sub>                         | [217,230]s <sub>2</sub>                  | [231,240]s <sub>2</sub>                  | [241,250]s <sub>2</sub>                 |
| 5           | [251,261]s <sub>2</sub>                         | [262,270]s <sub>2</sub>                  | [271,281]s <sub>2</sub>                  | [282,14]s <sub>2</sub> ∪ s <sub>3</sub> |
| 6           | [15,32]s <sub>3</sub>                           | [33,50]s <sub>3</sub>                    | [51,73]s <sub>3</sub>                    | [74,90]s <sub>3</sub>                   |
| 7           | [91,112]s <sub>3</sub>                          | [113,132]s <sub>3</sub>                  | [133,152]s <sub>3</sub>                  | [153,170]s <sub>3</sub>                 |
| 8           | [171,185]s <sub>3</sub>                         | [186,5]s <sub>3</sub> ∪ s <sub>4</sub>   | [6,12] s <sub>4</sub>                    | [13,23] s <sub>4</sub>                  |
| 9           | [24, 35]s <sub>4</sub>                          | [36,56]s <sub>4</sub>                    | [57,72]s <sub>4</sub>                    | [73,85]s <sub>4</sub>                   |
| 10          | [86,98]s <sub>4</sub>                           | [99,112]s <sub>4</sub>                   | [113,128]s <sub>4</sub>                  | [129,146]s <sub>4</sub>                 |
| 11          | [147,164]s <sub>4</sub>                         | [165,3]s <sub>4</sub> ∪ s <sub>5</sub>   | [4,12]s <sub>5</sub>                     | [13,24]s <sub>5</sub>                   |
| 12          | [25,42]s <sub>5</sub>                           | [43,50]s <sub>5</sub>                    | [51,68]s <sub>5</sub>                    | [69,83]s <sub>5</sub>                   |
| 13          | [84,98]s <sub>5</sub>                           | [99,112]s <sub>5</sub>                   | [113,13]s <sub>5</sub> ∪ s <sub>6</sub>  | [14,36]s <sub>6</sub>                   |
| 14          | [37,59]s <sub>6</sub>                           | [60,80]s <sub>6</sub>                    | [81,95]s <sub>6</sub>                    | [96,111]s <sub>6</sub>                  |
| 15          | [112,127]s <sub>6</sub>                         | [128 ,140]s <sub>6</sub>                 | [141,151]s <sub>6</sub>                  | [152,3]s <sub>6</sub> ∪ s <sub>7</sub>  |
| 16          | [4,29]s <sub>7</sub>                            | [30,45]s <sub>7</sub>                    | [46,67]s <sub>7</sub>                    | [68,86]s <sub>7</sub>                   |
| 17          | [87,115] s <sub>7</sub>                         | [116,141]s <sub>7</sub>                  | [142,155]s <sub>7</sub>                  | [156,170]s <sub>7</sub>                 |
| 18          | [171,187]s <sub>7</sub>                         | [188,4]s <sub>7</sub> ∪ s <sub>8</sub>   | [5,21]s <sub>8</sub>                     | [22,40]s <sub>8</sub>                   |
| 19          | [41,60]s <sub>8</sub>                           | [61,4]s <sub>8</sub> ∪ s <sub>9</sub>    | [5,18]s <sub>9</sub>                     | [19,33]s <sub>9</sub>                   |
| 20          | [34,45]s <sub>9</sub>                           | [46,60]s <sub>9</sub>                    | [61,75]s <sub>9</sub>                    | [76,93]s <sub>9</sub>                   |
| 21          | [94,111]s <sub>9</sub>                          | [112,123]s <sub>9</sub>                  | [124,10]s <sub>9</sub> ∪ s <sub>10</sub> | [11,25]s <sub>10</sub>                  |
| 22          | [26,48]s <sub>10</sub>                          | [49,70]s <sub>10</sub>                   | [71,92]s <sub>10</sub>                   | [93,5]s <sub>10</sub> ∪ s <sub>11</sub> |
| 23          | [6,23]s <sub>11</sub>                           | [24,40]s <sub>11</sub>                   | [41,60]s <sub>11</sub>                   | [61,82]s <sub>11</sub>                  |
| 24          | [83,104]s <sub>11</sub>                         | [105,9]s <sub>11</sub> ∪ s <sub>12</sub> | [10,32]s <sub>12</sub>                   | [33,52]s <sub>12</sub>                  |
| 25          | [53,76]s <sub>12</sub>                          | [77,100]s <sub>12</sub>                  | [101,4]s <sub>12</sub> ∪ s <sub>13</sub> | [5,20]s <sub>13</sub>                   |
| 26          | [21,35]s <sub>13</sub>                          | [36,12]s <sub>13</sub> ∪ s <sub>14</sub> | [13,29]s <sub>14</sub>                   | [30,54]s <sub>14</sub>                  |
| 27          | [1,48]s <sub>15</sub>                           | [49,99]s <sub>15</sub>                   | [1,29]s <sub>16</sub>                    | [30,50]s <sub>16</sub>                  |
| 28          | [51,70]s <sub>16</sub>                          | [71,89]s <sub>16</sub>                   | [90,110]s <sub>16</sub>                  | [111,128]s <sub>16</sub>                |
| 29          | [1,22]s <sub>17</sub>                           | [23,49]s <sub>17</sub>                   | [50,69]s <sub>17</sub>                   | [70,98]s <sub>17</sub>                  |

|    |                 |                |                |                 |
|----|-----------------|----------------|----------------|-----------------|
| 30 | [99,16]s17∪s18  | [17,29]s18     | [30,48]s18     | [49,73]s18      |
| 31 | [74,97]s18      | [98,20]s18∪s19 | [21,58]s19     | [59,99]s19      |
| 32 | [1,53]s20       | [54,80]s20     | [81,107]s20    | [108,134]s20    |
| 33 | [1,29]s21       | [30,50]s21     | [51,85]s21     | [86,111]s21     |
| 34 | [1,18]s22       | [19,35]s22     | [36,57]s22     | [58,76]s22      |
| 35 | [1,38]s23       | [39,75]s23     | [76,115]s23    | [116,20]s23∪s24 |
| 36 | [21,34]s24      | [35,50]s24     | [51,60]s24     | [61,20]s24∪s25  |
| 37 | [21,44]s25      | [45,77]s25     | [1,49]s26      | [50,110]s26     |
| 38 | [111,180]s26    | [181,5]s26∪s27 | [6,26]s27      | [27,57]s27      |
| 39 | [58,83]s27      | [84,10]s27∪s28 | [11,28]s28     | [29,50]s28      |
| 40 | [51,75]s28      | [76,6]s28∪s29  | [7,24]s29      | [25,45]s29      |
| 41 | [46,7]s29∪s30   | [8,28]s30      | [29,52]s30     | [53,20]s31      |
| 42 | [21,10]s31∪s32  | [11,30]s32     | [1,17]s33      | [18,30]s33      |
| 43 | [31,48]s33      | [49,59]s33     | [60,9]s33∪s34  | [10,23]s34      |
| 44 | [24,45]s34      | [46,14]s34∪s35 | [15,40]s35     | [41,26]s35∪s36  |
| 45 | [27,58]s36      | [59,21]s36∪s37 | [22,82]s37     | [83,144]s37     |
| 46 | [145,19]s37∪s38 | [20,50]s38     | [51,8]s38∪s39  | [9,30]s39       |
| 47 | [31,50]s39      | [51,72]s39     | [1,20]s40      | [21,40]s40      |
| 48 | [41,65]s40      | [66u,7]s40∪s41 | [8,23]s41      | [24,45]s41      |
| 49 | [46,10]s41∪s42  | [11,24]s42     | [25,47]s42     | [48,22]s42∪s43  |
| 50 | [23,62]s43      | [63,23]s43∪s44 | [24,12]s44∪s45 | [13,36]s45      |
| 51 | [1,19]s46       | [20,10]s46∪s47 | [11,33]s47     | [34,17]s47∪s48  |
| 52 | [18,29]s48      | [1,13]s49      | [14,26]s49∪s50 | [27,30]s50∪s51  |
| 53 | [31,21]s51∪s52  | [22,25]s52∪s53 | [26,9]s53∪s54  | [10,55]s54      |
| 54 | [1,77]s55       | [1,77]s56      | [78,14]s56∪s57 | [15,28]s57      |
| 55 | [1,13]s58       | [14,10]s58∪s59 | [11,6]s60      | [7,14]s60∪s61   |
| 56 | [1,8]s62∪s63    | [9,18]s63∪s64  | [1,12]s65      | [1,12]s66       |
| 57 | [1,18]s67∪s68   | [19,17]s68∪s69 | [18,39]s69∪s70 | [40,30]s70∪s71  |
| 58 | [1,17]s72∪s73   | [18,55]s73∪s74 | [1,18]s75∪s76  | [19,50]s76∪s77  |
| 59 | [1,40]s78∪s79   | [41,29]s79∪s81 | [1,15]s82∪s84  | [16,17]s84-s86  |
| 60 | [1,32]s87-s89   | [1,11]s90-s93  | [1,8]s94-s100  | [1,6]s101-s114  |

[Tab.4]

Explication :

a/  $H_{54}=[1,77]_{s_{55}}$ : le Hizb n°54, commence par le verset n°1 au verset n°77, dans la sourate n°55

b/  $H_{59}=[16,17]_{s_{84-s_{86}}}$  : le Hizb n°59 commence par le verset n°16 de la sourate n°84, au n° 17 de la sourate n°86

### **PIII :4 :Réduction des données et choix des mots**

#### **PIII : 4 :1/ Partie théorique**

L'analyse factorielle des correspondances a pour but d'approcher de représenter les nuages d'une matrice des nombres positifs afin d'en révéler la structure.

Dans cette partie nous présentons, l'ensemble et le choix des mots, ce travail représente la première phase de la recherche.

Donc, soucieux d'élaborer les méthodes les plus efficaces pour comprendre le Coran ou la Bible, les linguistes expérimentalistes sont conduits à apprendre le vocabulaire fondamental du texte.

#### **1 :1 :But**

Il est intéressant de trier judicieusement le lexique élémentaire, c'est-à-dire de pouvoir connaître la structure essentielle dont le lecteur (débutant) aura le plus besoin pour effectuer des tâches simples :

- Lire facilement le texte coranique.
- Comparer deux sourates (ou Hizbs ou versets, etc.) textes ou plusieurs.
- Comprendre le contenu d'une sourate, ou texte coranique.

- Etudier la formation d'un texte sur le plan de l'unité, ou de l'hétérogénéité des sources à partir de la distribution des éléments : mots, noms propres, verbes choisis, qui composent le discours du texte.

### **1 :2 : L'exploitation de l'information :**

Pour exploiter au mieux toute l'information contenue dans un texte, nous avons examiné les classes des mots dont la fréquence est au voisinage de 50 sinon plus (comme le cas de certains mots : création, terre, etc.).

### **PIII :4 :2/ Partie pratique :**

La matrice utilisée donc est constituée des Rubu's (Hizbs, Sourates ou Chapitres) pouvant être décrits à partir d'un nombre fini de mots.

La difficulté essentielle de l'analyse réside dans le fait que le choix du mot à chaque essai est la résultante d'un mécanisme de recherche et de traitement de l'information (mots qui contribuent à la formation du tableau des données).

Pour résoudre cette difficulté, nous serons conduits à définir quelques étapes dans notre analyse. Donc les règles méthodologiques retenues étaient les suivantes :

**Première étape :** consiste à traiter la matrice entière à partir des données brutes, ces mots ont été dénombrés à la main en se basant sur le dictionnaire du Coran (Tableau de contingence initial 340 mots et 60 Hizbs).

**Deuxième étape :** cherchons ensuite les mots qui ont les mêmes coordonnées (c'est-à-dire mêmes abscisses et mêmes ordonnées), ou plutôt les mêmes représentativités), l'examen des corrélations et des qualités permet de savoir, pour chaque mot, s'il a ou non une représentativité identique aux autres mots.

**Troisième étape :** consiste à mettre en exécution les tableaux choisis.

**Quatrième étape :** Après avoir traité le tableau de tous les mots et les hizbs, nous avons obtenu un nuage (résultat), dont certains mots ont les mêmes représentations sur 7 facteurs (coordonnées).

Nous dressons ci-dessous (Tab.5, 6,...11) la réduction faite après avoir analysé un tableau constitué de 340 mots \* 60 Hizbs, dont les mots ont les mêmes coordonnées (F1, F2,...F7).

**1<sup>ère</sup> réduction :**

| Mots  | Qualité de représentation | F1   | F2   | F3    | F4     | F5  | F6    | F7  |
|---|---------------------------|------|------|-------|--------|-----|-------|-----|
| Nuit de décret  | 950                       | 2831 | 6240 | - 951 | - 3887 | 575 | - 885 | 218 |
| Sinine  |                           |      |      |       |        |     |       |     |
| Dispersé  |                           |      |      |       |        |     |       |     |
| Houtamah  |                           |      |      |       |        |     |       |     |
| Impénétrable  |                           |      |      |       |        |     |       |     |
| Coursiers   |                           |      |      |       |        |     |       |     |
| Rapides<br>Quraïch<br>Ville à la<br>colonne<br>Misérable<br>Deux voies<br>tentateur |                           |      |      |       |        |     |       |     |

[Tab.5]

**2<sup>ème</sup> réduction :**

| Mots      | Qualité | F1   | F2   | F3 | F4   | F5   | F6   | F7   |
|-----------|---------|------|------|----|------|------|------|------|
| Salsabil  | 440     | 1795 | 1692 | 40 | 2000 | -327 | 1633 | 3956 |
| Sombre    |         |      |      |    |      |      |      |      |
| Planètes  |         |      |      |    |      |      |      |      |
| Enveloppé |         |      |      |    |      |      |      |      |

|  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|
| Froid<br>glacial<br>Gingembre<br>Camphre |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|

[Tab.6]

**3<sup>ème</sup> réduction**

| Mots | Qualité | F1  | F2   | F3  | F4   | F5  | F6   | F7 |
|------|---------|-----|------|-----|------|-----|------|----|
| TSM  | 70      | 146 | -894 | 510 | -943 | 883 | -    | 7  |
| TS   |         |     |      |     |      |     | 1358 |    |
| TH   |         |     |      |     |      |     |      |    |

[Tab.7]

**4<sup>ème</sup> réduction**

|   |     |      |     |     |      |          |           |     |
|---|-----|------|-----|-----|------|----------|-----------|-----|
| Tasnim<br>Premier<br>état<br>Débordants<br>Vin rare<br>Bruit<br>retentira<br>Le fracas<br>Cataclysme<br>Ceux qui<br>arrachent | 850 | 2512 | 844 | -40 | 5173 | -<br>687 | -<br>5389 | 914 |
|---|-----|------|-----|-----|------|----------|-----------|-----|

[Tab.8]

**5<sup>ème</sup> réduction**

|  |     |      |     |      |     |      |     |           |
|--|-----|------|-----|------|-----|------|-----|-----------|
| Nun<br>Celle<br>qui<br>doit<br>venir<br>Baltard<br>Par<br>surcroît | 563 | 1574 | 450 | -426 | 559 | -417 | 689 | -<br>5866 |
|--|-----|------|-----|------|-----|------|-----|-----------|

[Tab.9]

### 6<sup>ème</sup> réduction

|            |     |      |       |      |       |       |      |     |
|------------|-----|------|-------|------|-------|-------|------|-----|
| K.HY.AIN   | 637 | 1530 | -1348 | 1756 | -1714 | -5360 | -989 | -   |
| Sad        |     |      |       |      |       |       |      | 232 |
| Dou al     |     |      |       |      |       |       |      |     |
| qarnain    |     |      |       |      |       |       |      |     |
| Bouillante |     |      |       |      |       |       |      |     |
| Deux       |     |      |       |      |       |       |      |     |
| monts      |     |      |       |      |       |       |      |     |

[Tab.10]

### 7<sup>ème</sup> réduction

|          |     |     |     |      |    |     |     |     |
|----------|-----|-----|-----|------|----|-----|-----|-----|
| Poterie  | 378 | 826 | 316 | -166 | 78 | 408 | 544 | -   |
| Montagne |     |     |     |      |    |     |     | 231 |
| Tamoud   |     |     |     |      |    |     |     |     |
| Mise en  |     |     |     |      |    |     |     |     |
| marche   |     |     |     |      |    |     |     |     |

[Tab.11]

### PIII :4 :2 :1 : Résultat

Les tableaux ont été constitués des mots qu'on juge importants et qui ont été choisis selon leur degré de fréquence (voir ci-dessus).

En premier temps, nous allons traiter un tableau de contingence dont les lignes représentent les Hizbs éclatés en 4 parties (Rubu's), et les colonnes représentent les mots (118 mots choisis).

### PIII :4 :2 :2 : Tableau des valeurs propres (Tab.12) :

Pour l'analyse factorielle des correspondances, on a analysé un tableau de 240 rubu's et 118 mots (après réduction des données).

Dans cette analyse, on a les résultats d'un tableau de contribution, l'histogramme des valeurs propres et des graphes.

Le tableau indiqué ci-dessous, montre la variation des valeurs propres et leur pourcentage, dans l'inertie totale pour 7 facteurs extraits de l'analyse.

| <b>Facteur <math>F_\alpha</math></b> | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> | <b>5</b> | <b>6</b> | <b>7</b> |
|--------------------------------------|----------|----------|----------|----------|----------|----------|----------|
| Valeur propre $\lambda_\alpha$       | 0.28     | 0.20     | 0.17     | 0.16     | 0.16     | 0.14     | 0.12     |
| Pourcentage d'inertie $\zeta_\alpha$ | 5.997    | 4.33     | 3.75     | 3.56     | 3.47     | 3.17     | 2.62     |

[Tab.12] Tableau de facteurs  $F_\alpha$

### **PIII :4 :2 :3: Interprétation**

#### **A/ L'histogramme des valeurs propres :**

On remarque la décroissance dans l'explication des axes factoriels successifs. Le premier axe explique 5.997 de l'inertie totale, c'est-à-dire l'équivalence de  $5.997/100 /118 \approx 17$  variables (mots) et le deuxième axe explique 4.33 soit 5 variables.

#### **B/ Contribution relative (CTR) :**

Exprime l'inertie de l'élément parallèle à l'axe factoriel en tant que proportion de l'inertie du nuage parallèle à cet axe (multipliée par un facteur 1000).

Si  $f_m$  est la masse de l'élément (m),  $F_\alpha(m)$  est l'abscisse de l'élément (m) sur l'axe  $\alpha$ , l'inertie de (m) parallèle à  $\alpha$  est alors :

$$\lambda_\alpha = \sum \{ f_r * F_\alpha(r) / r \in R \}$$

Alors sa contribution relative est :  $CTR_\alpha(r) = (f_r * F_\alpha(r)^2 / \lambda_\alpha) * 1000$

**C/Corrélation** : indique la contribution relative de facteur  $\alpha$  à la distance au carré, du mot (m) au centre de gravité de nuage.

$$Corr_\alpha(m) = (F_\alpha^2(m) / \sum \{F_\alpha^2(m) ; \alpha=1, 118\})$$

#### **D/Représentation des facteurs :**

L'examen des graphiques plans, suffit à l'interprétation des indications numériques.

#### **D :1 :Facteurs**

**D :1 :1 :Premier facteur : Facteur de croyance ( $\lambda_1=0.28 ; \zeta_1 =5.977\%$ )**

Il s'agit d'un facteur général sur lequel, se distinguent nettement trois blocs fondamentaux :

1 : Les rubu's mecquois (Tab.13), qui se trouvent à l'extrême gauche du 1<sup>er</sup> axe  $\alpha$  (côté négatif), forment ce qu'on appelle : côté rationnel (voir : Explication Générale plus loin).

2 : Les rubu's médinois (Tab.14) qui se trouvent à l'extrême droite du 1<sup>er</sup> axe  $\alpha$  (Côté positif), forment la partie législative.

3 : Un bloc intermédiaire dont le rubu' est mecquois ou bien la révélation était à la médina ou le contraire (la position des rubu's est près du centre), ou bien ils n'ont pas une forte corrélation et contribution avec le 1<sup>er</sup>axe  $\alpha$  (fig.4).

Aux non spécialistes, il suffira de dire que le Coran est divisé en 114 sourates de très inégales longueurs.

Le texte du Coran comprend une partie mecquoise et une partie médinoise : il y a des sourates de la Mecque, des sourates de Médine et quelques sourates dont les versets sont rapportés aux deux villes (résultat trouvé ci-dessus).

Alors, sur l'extrême gauche de l'axe, il y a un regroupement de rubu's mecquois dont la corrélation avec le 1<sup>er</sup> axe est très forte, et qui sont (Tab.13) :

| <b>Hizbs<br/>(Mecquois)</b> | <b>Corrélation</b> | <b>Contribution relative</b> | <b>Poids</b> | <b>Nombre de versets</b> |
|-----------------------------|--------------------|------------------------------|--------------|--------------------------|
| H(17)2=<br>[113,132]s3      | 163                | 14                           | 5            | 20                       |
| H(17)3= [133,152]s3         | 145                | 13                           | 4            | 20                       |
| H(27)1=[1,48]s15            | 72                 | 77                           | 4            | 48                       |
| H(30)1=<br>[99,16]s17∪s18   | 53                 | 12                           | 5            | 28                       |
| H(31)2=<br>[98,20]s18∪s19   | 74                 | 13                           | 4            | 28                       |
| H(31)4=[59,99]s19           | 92                 | 15                           | 5            | 41                       |
| H(32)2= [1,53]s20           | 306                | 17                           | 4            | 53                       |
| H(32)3= (81,107) s20        | 112                | 11                           | 2            | 27                       |
| H(32)4= (108,134)s20        | 97                 | 12                           | 4            | 27                       |
| H(33)2=(30,50)s21           | 171                | 8                            | 3            | 21                       |
| H(37)2=[45,77]s25           | 110                | 7                            | 4            | 33                       |
| H(39)4= [29,50]s28          | 141                | 13                           | 4            | 22                       |
| H(54)1<br>=(1,77)s55        | 354                | 31                           | 6            | 77                       |
| H(58)3=[1,18]s7<br>5∪s76    | 62                 | 12                           | 3            | 57                       |
| H(58)4=[19,50]s76∪s77       | 27                 | 11                           | 3            | 63                       |
| H(59)1=(1,40)s78 ∪<br>s79   | 235                | 15                           | 4            | 80                       |
| H(59)2= [41,29]s79∪s81      | 136                | 8                            | 2            | 76                       |

|  |     |   |   |    |
|--|-----|---|---|----|
| H(59) <sub>3</sub> =[1,15] <sub>s82</sub><br>∪ s84 | 144 | 8 | 3 | 70 |
| H(60) <sub>1</sub> =[s87 ∪ s89]                    | 216 | 7 | 3 | 77 |

[Tab.13]

Sur l'extrême droite du 1° axe, on trouve aussi une concentration des Rubu's (أرباع القرآن الكريم) médinois dont la contribution et corrélation de ces derniers sont très fortes, ces Rubu's sont (Tab.14)

| Hizbs médinois  | Corrélation | Contributi<br>on relative | Poids | Nombre de<br>versets |
|---|-------------|---------------------------|-------|----------------------|
| H(3) <sub>4</sub> = [188,200] <sub>s2</sub>                   | 43          | 24                        | 3     | 13                   |
| H(4) <sub>2</sub> = [217,230] <sub>s2</sub>                   | 188         | 15                        | 6     | 14                   |
| H(4) <sub>3</sub> = [231,240] <sub>s2</sub>                   | 161         | 9                         | 3     | 10                   |
| H(7) <sub>1</sub> = [91,112] <sub>s3</sub>                    | 159         | 6                         | 5     | 22                   |
| H(7) <sub>4</sub> = [153,170] <sub>s3</sub>                   | 291         | 8                         | 5     | 18                   |
| H(8) <sub>3</sub> = [6,12] <sub>s4</sub>                      | 56          | 7                         | 2     | 7                    |
| H(8) <sub>4</sub> = [13,23] <sub>s4</sub>                     | 84          | 10                        | 3     | 11                   |
| H(9) <sub>1</sub> = (24,35) s4                                | 264         | 11                        | 4     | 12                   |
| H(10) <sub>1</sub> = (86,98) <sub>s4</sub>                    | 190         | 10                        | 5     | 13                   |
| H(10) <sub>4</sub> = (129,146) <sub>s4</sub>                  | 240         | 11                        | 7     | 18                   |
| H(13) <sub>1</sub> = [84,98] <sub>s5</sub>                    | 38          | 9                         | 6     | 15                   |
| H(19) <sub>2</sub> = [61,4] <sub>s8</sub> U <sub>s</sub><br>9 | 366         | 18                        | 7     | 20                   |
| H(19) <sub>4</sub> = [19,33] <sub>s9</sub>                    | 240         | 7                         | 5     | 15                   |
| H(20) <sub>3</sub> = [61,75] <sub>s9</sub>                    | 154         | 10                        | 6     | 15                   |
| H(20) <sub>4</sub> = [76,93] <sub>s9</sub>                    | 152         | 8                         | 4     | 18                   |

|                                 |     |    |    |    |
|---------------------------------|-----|----|----|----|
| H(21)4=(94,111)s9               | 189 | 11 | 6  | 18 |
| H(21)2=[112,123)s9              | 168 | 7  | 4  | 12 |
| H(26)1= [21,34]s24              | 254 | 9  | 4  | 14 |
| H(42)4=[18,30]s33               | 311 | 10 | 4  | 13 |
| H(43)1= [31,48]s33              | 294 | 16 | 5  | 18 |
| H(43)2=[49,59]s33               | 165 | 23 | 4  | 11 |
| H(52) <sub>2</sub> =[1 , 13]s49 | 342 | 10 | 4  | 13 |
| H(52)1=[18,29]s48               | 298 | 8  | 4  | 12 |
| H(55)1= [1,13] s58              | 286 | 11 | 10 | 13 |
| H(55)2= [14,10]s58 U<br>s59     | 215 | 8  | 6  | 18 |
| H(55)4= [7,14]s60 U<br>s61      | 311 | 14 | 8  | 21 |
| H(56)1= [1, 8]s62 Us63          | 102 | 7  | 5  | 19 |
| H(56)3 = [1, 12]s56             | 286 | 6  | 4  | 12 |

[Tab.14]

Le centre est représenté par tous les Rubu's non cités ci-dessus et qui ont une faible contribution et corrélation avec le 1<sup>er</sup> axe, c'est-à-dire ils ne sont pas représentés sur cet axe.

| <b>Hizb</b>        | <b>Corrélation</b> | <b>Contribution relative</b> | <b>Poids</b> | <b>Nombre de versets</b> |
|--------------------|--------------------|------------------------------|--------------|--------------------------|
| H(3)4= [188,200]s2 | 43                 | 24                           | 3            | 13                       |
| H(4)2= [217,230]s2 | 188                | 15                           | 6            | 14                       |
| H(4)3=[213,240]s2  | 161                | 9                            | 3            | 10                       |

|                          |     |    |   |    |
|--------------------------|-----|----|---|----|
| H(7)1= [91,112]s3        | 159 | 6  | 5 | 22 |
| H(7)4=[153, 170]s3       | 291 | 8  | 5 | 18 |
| H(8)3=[6,12]s4           | 56  | 7  | 2 | 7  |
| H(8)4=[13,23]s4          | 84  | 10 | 3 | 11 |
| H(9)1=[24,35]s4          | 264 | 11 | 4 | 12 |
| H(10)1=[86,98]s4         | 190 | 10 | 5 | 13 |
| H(10)4=(129,146)s4       | 240 | 11 | 7 | 18 |
| H(13)1=[84,98]s5         | 38  | 9  | 6 | 15 |
| H(19)2=[61,4]s8 U s 9    | 366 | 18 | 7 | 20 |
| H(19)4 = [19,33]s9       | 240 | 7  | 5 | 15 |
| H(20)3= [61,75]s9        | 154 | 14 | 6 | 15 |
| H(20)4=[76,93]s9         | 152 | 8  | 4 | 18 |
| H(21)4=(94,111)s9        | 189 | 11 | 6 | 18 |
| H(21)2=[112,123]s9       | 168 | 7  | 4 | 12 |
| H(26)1= [21,34]s24       | 254 | 9  | 4 | 14 |
| H(42)4=[18,30]s33        | 311 | 10 | 4 | 13 |
| H(43)1= [31,48]s33       | 294 | 16 | 5 | 18 |
| H(43)2=[49,59]s33        | 165 | 23 | 4 | 11 |
| H(51)4= [34,17]s47Us48   | 322 | 7  | 6 | 23 |
| H(52)1=[18,29]s48        | 298 | 8  | 4 | 12 |
| H(52)2=[1, 13]s49        | 342 | 10 | 4 | 13 |
| H(55)2= [14,10]s58 U s59 | 215 | 8  | 6 | 18 |
| H(55)4= [7,14]s60 U s61  | 311 | 14 | 8 | 21 |
| H(56)1=[1, 8] s62 U S63  | 102 | 7  | 5 | 19 |
| H(56)3 = [1, 12] s65     | 286 | 6  | 4 | 12 |

(Tab :15)

On peut résumer les deux tableaux comme suit (Tableaux :13,14) :

Le premier facteur, nous montre que le Coran est composé de deux parties :

A/ Une partie mecquoise qui représente la première étape de la révélation, c'est-à-dire la période avant l'émigration du prophète (Que Dieu Lui accorde la bénédiction et la paix).

Dans cette partie, le Coran nous donne une idée complète sur les récits des prophètes et des nations, sur la création de l'univers, d'Adam, l'enfer, le Paradis, etc.

Les versets de cette partie sont courts, on peut dire (que cette partie représente la situation avant l'émigration).

B/ Une autre partie qui représente la deuxième étape après l'émigration.

Durant la prédication médinoise se trouve un bon nombre des thèmes déjà développés à la Mecque, les thèmes en rapport avec la vie politique individuelle sont évoqués souvent.

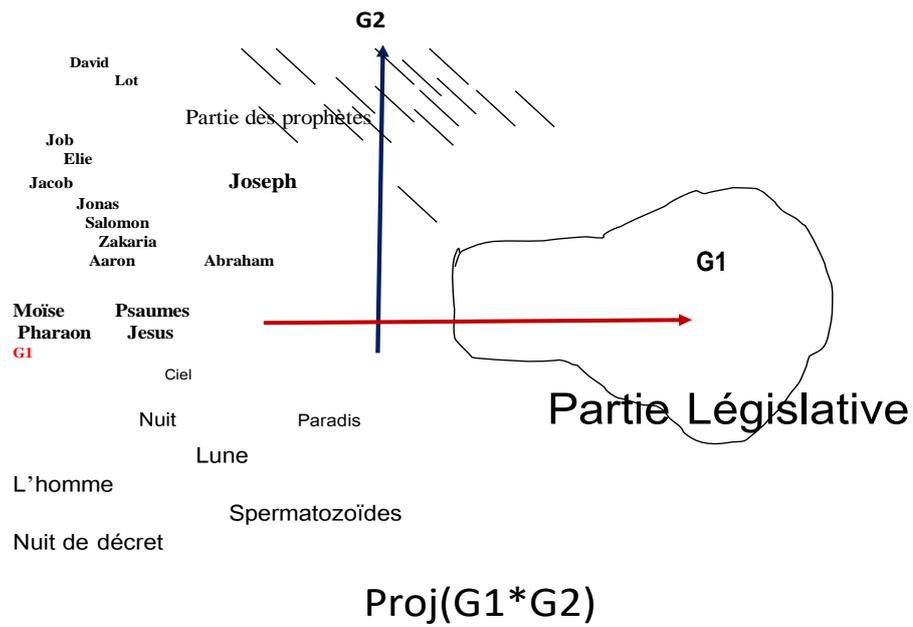
**D :1 :2 : Deuxième facteur : Facteur de Prophètes (  $\lambda^2=0.20$  ;  $\zeta^2= 4.33$  %)**

**Par rapport aux mots  $\mathcal{N}(M)$  [Fig.4]**

Malgré sa faible contribution de 4.33% à l'inertie du nuage, cet axe mérite une interprétation par rapport à l'axe 1. La projection sur ce facteur, nous donne toute une série de prophètes qui sont passés dans l'histoire, cette série se trouve bien représenté et, elle a une forte corrélation et contribution avec l'axe  $G2 > 0$  (partie supérieure du plan) (figure :4), et surtout :  $\Lambda = \{\text{Abraham, Joseph, Isaak}\}$ .

La contribution élevée de Joseph est expliquée par le fait que l'histoire de ce dernier est complète et occupe la totalité de la sourate  $S_{12}$ (Fig.4).

D'autre part, sur  $G2 < 0$  (partie inférieure du plan :  $1*2$  ; Fig.4), nous constatons l'existence des mots :  $\eta = \{\text{Ciel, Homme, Lune, réer, spermatozoïdes}\}$ , qui forment un bloc homogène (Fig. 2, bloc gauche-bas, et Tableau : 4).



(fig:4)

La partie positive (Fig.4) :

| <b>G2 &gt;0</b> | <b>Poids</b> | <b>Corrélation</b> | <b>Contribution relative</b> |
|-----------------|--------------|--------------------|------------------------------|
| Abraham         | 7            | 334                | 98                           |
| Joseph          | 3            | 132                | 99                           |
| Moïse           | 13           | 157                | 79                           |
| Ismail          | 1            | 307                | 62                           |
| Abraham         | 7            | 334                | 98                           |
| Joseph          | 3            | 132                | 99                           |
| Moïse           | 13           | 157                | 79                           |
| Ismail          | 1            | 307                | 62                           |
| Job             | 1            | 204                | 22                           |
| Jonas           | 1            | 130                | 13                           |

|                          |   |     |    |
|--------------------------|---|-----|----|
| Jacob                    | 2 | 389 | 59 |
| David                    | 2 | 71  | 19 |
| Zakaria                  | 1 | 69  | 13 |
| Pharaon                  | 7 | 53  | 19 |
| Lot                      | 3 | 90  | 25 |
| Prophètes                | 4 | 81  | 19 |
| Christ                   | 2 | 77  | 10 |
| Aaron                    | 2 | 153 | 26 |
| Ceux qui font<br>Le bien | 3 | 112 | 26 |
| Elie                     | 1 | 131 | 16 |

. (Tab:16)

La partie négative (Fig.4) :

| <b>G2&lt;0</b> | <b>Poids</b> | <b>Corrélation</b> | <b>contribution</b> |
|----------------|--------------|--------------------|---------------------|
| Ciel           | 30           | 157                | 24                  |
| Homme          | 6            | 103                | 36                  |
| Paradis        | 14           | 70                 | 11                  |
| Lune           | 3            | 75                 | 10                  |
| Nuit           | 8            | 74                 | 12                  |
| Spermatozoïdes | 1            | 53                 | 10                  |
| Créer          | 13           | 104                | 18                  |

[Tab.17]

### Explication Générale :

Nul doute que la lecture transparente responsable et lisible du Coran, montre cette homogénéité, il suffit de lire - en utilisant le 1er verset coranique révélé à son prophète (Que Dieu Lui accorde la Bénédiction et la paix)

*Lis au nom de ton Seigneur qui a créé\**

*Qui a créé l'homme\* de quelque chose qui s'accroche\**

*Lis, car ton Seigneur le Très Noble, c'est Lui qui enseigné*

*Par la plume, Il a enseigné à l'homme ce qu'il ne savait pas] (LVI-1,...5)*

Une lecture lisible montre clairement que le coran a invité l'homme à acquérir le savoir et la connaissance qui lui sont utiles dans les divers domaines de l'univers, et autres sciences.

Les mots de ces versets ont constitué la toute première révélation à Mohammed, ils ont constitué la lumière par ordre du Seigneur afin qu'il soit la délivrance de toutes les ténèbres de l'ignorance, de l'illusion, et les amène vers la pleine lumière, la lumière de vérité qui illumine les cœurs et guide les clairvoyants vers la voie du salut.

Encore, l'Islam a encouragé les croyants à chercher, à acquérir des connaissances dans toutes les spécialités et à accroître leur savoir dans tous les domaines.

Il a appelé à une foi fondée sur l'observation de la création et, la méditation métaphysique et qu'elle soit le fruit d'une quête personnelle de l'absolu.

En plus, l'Islam va même plus loin, il considère que pour être complet, la foi de l'homme en Le Tout Puissant, doit s'accompagner de connaissance et du savoir.

Le Seigneur dit :

[*Où il y a bien des signes vraiment, pour les gens qui comprennent*] (XIII-4)

[*Où il y a un des signes de Dieu, peut-être s'en souviendront-ils !*] (VII-26)

[*Dans la création des cieux et de la terre ; Dans l'alternance de la nuit et du jour ; Il y a des signes Pour ceux qui sont doués d'intelligence*] (III-190)

Et c'est ainsi, on constate que le Coran a bien démontré à travers les nombreux versets, que la foi islamique ne s'oppose pas à la science, bien au contraire, elle a consolidé le lien entre ce qui est matériel à ce qui est moral, en prescrivant à chaque musulman et musulmane le devoir de la science, et, en leur recommandant d'en puiser tout au long de la vie, en faisant de la science un critère de préférence entre les hommes au même titre que la piété.

Dans ce cadre, on conclut l'idée fondamentale à savoir que :

a/  $\eta = \{\text{Ciel, Homme, Lune, créer, spermatozoïdes}\}$ , qui forment un bloc homogène (Fig. 2, bloc gauche-bas). cité au- paravent trouve sa place convenable.

b/ l'Islam est une religion rationnelle (qui donne la priorité à la raison), et ceci on le remarque au 1<sup>er</sup> verset du Coran, où le Seigneur a demandé à son prophète Mohammed (Que Dieu Lui accorde la bénédiction et la paix) de lire cet univers avant d'exécuter les piliers de l'Islam.

Donc Il a fait de la recherche scientifique un devoir sacré et noble, qu'il soit :

### A/ Dans un monde infiniment petit

[Le poids d'un atome n'échappe à ton Seigneur ni sur terre, ni dans les cieux, Il n'y a rien de plus petit ou de plus grand que cela qui ne soit inscrit dans un livre explicite] (X-61)

### B/ Dans un monde infiniment grand

[Ceux qui mécroient n'ont-ils pas vu que les cieux et la terre étaient bel et bien cousus ? Ensuite Nous les avons dégagés tous deux] (XXI-30)

### C/ Dans nous-mêmes dans notre corps

[Nous avons créé l'homme d'un choix d'argile  
Puis Nous l'avons conigné, goutte de sperme dans un reposoir sûr.  
Puis Nous avons fait du sperme quelque chose qui s'accroche  
Puis de quelque chose qui s'accroche Nous avons créé un morceau de chair.  
Puis du morceau de la chair Nous avons créé des os  
Nous en avons produit une toute autre créature] (XXIII- 12,13,14)

## La partie médinoise (Partie positive)

En procédant à une étude minutieuse de la partie médinoises (bloc de droite), nous remarquons que le coran invite les musulmans a adhérer à une foi saine, et leur montre que la révélation médinoise a des normes rigoureuses et invariables, elle présente de façon détaillée, des éléments de la législation islamique, les œuvres de culte, les jugements légaux relatifs au jeûne, pèlerinage, aux orphelins, aux menstrues de la femme, la dépense l'usure.

Elle montre les catégories des gens (croyants, mécréants, hypocrites...) et présente la création d'Adam, paix sur lui.

### Lecture générale (Résumé) :

On peut résumer les deux tableaux comme suit (Tableaux :1,2)

Le premier facteur nous montre que le Coran est composé de deux parties :

- a- Une partie mecquoise qui représente la première étape de la révélation, c'est-à-dire la période avant l'émigration du prophète (Que Dieu Lui accorde la bénédiction et la paix) de la Mecque à la Médina.

b- Une autre partie qui représente la deuxième étape après l'émigration.

Dans la partie mecquoise, le Coran nous donne une idée complète sur les récits des prophètes et des nations, sur la création de l'univers, d'Adam et qu'il y a bien une résurrection, un paradis, un enfer .....

Les versets de cette partie sont courts, et on peut dire (si nous prenons l'unité du temps) que cette partie représente ce qui est passé avant l'émigration) (Tab.13).

Les textes coraniques ont été distribués en trois groupes pour tous ceux qui appartiennent à la période mecquoise, les critères pour l'élaboration de ce groupement sont trois sortes :

Les uns sont prédominance stylistique et tiennent compte avant tout de l'allure brève de versets, les autres concernent les circonstances qui ont provoqué des attitudes de Mohammed (محمد صلى الله عليه وسلم), devant ses opposants ; d'autres font surtout état de textes organiques destinés à structurer une communauté réclamant des interdictions alimentaires, du droit privé, etc.

Durant la prédication médinoise se trouve un bon nombre des thèmes déjà développés à la Mecque, les thèmes en rapport avec la vie politique individuelle sont évoqués souvent.

Dans la partie droite (voir Fig. :3), on constate l'existence des mots suivants : {Pèlerinage, Prière, jeûne, Allah et ses attributs}, cette distribution montre bien la partie législative de l'Islam, ou plutôt les préceptes fondamentaux obligatoires pour les musulmans.

**1/Profession de foi**, elle est la plus importante croyance islamique, implique que l'univers est créé par le seul Allah Unique et créateur des cieux et de la terre, Signifie que nous sommes convaincus que cet univers n'aurait pas existé sans la puissance d'Allah

**2/ La prière**, où chaque musulman est tenu d'effectuer cinq prières quotidiennes obligatoires (faridat : (فريضة) Elle s'effectue tournée vers la Kaaba (située à La Mecque).

**3/Zakat** (الزكاة), C'est un impôt légal purificateur.

**4/ Ramadan** (jeûne), C'est le jeûne effectué durant le mois lunaire du Ramadan

:

يا أيها الذين آمنوا كتب عليكم الصيام كما كتب على الذين من قبلكم لعلكم تتقون  
وكلوا واشربوا حتى يتبين لكم الخيط الأبيض من الخيط الأسود من الفجر

Mangez et buvez jusqu'à ce que l'on puisse distinguer

A l'aube un fil blanc d'un noir. Jeûnez, ensuite, jusqu'à la nuit. » (II; 187)

### 5/ pèlerinage à la Mecque

|                                      |       |      |      |      |      |      |      |
|--------------------------------------|-------|------|------|------|------|------|------|
| Valeur propre $\lambda_\alpha$       | 0.28  | 0.20 | 0.17 | 0.16 | 0.16 | 0.14 | 0.12 |
| Pourcentage d'inertie $\zeta_\alpha$ | 5.997 | 4.33 | 3.75 | 3.56 | 3.47 | 3.17 | 2.62 |

### D :1 :3 : Troisième facteur : Facteur de Prophètes ( $\lambda_3=0.17$ ; $\zeta_3= 3.75$ %)

Sur ce facteur, nous avons l'histoire de Moïse avec Pharaon, cette histoire est bien représentée par  $H(22)3 = [71,92]_{s10}$ ,

dont  $CTR(H(22)3, F3) = 6.7\%$  ;  $\text{Cos}^2(H(22)3, F3) = 0.461$ .

C'est-à-dire que le 3<sup>o</sup> Rubu' du 22<sup>ième</sup> Hizb exploite 17% de la contribution totale, aussi le Rubu' H(17)2, ces deux Rubu's expliquent l'histoire de Moïse avec Pharaon (Tab.16), où la contribution de Moïse à cet axe est de 24.5% (fig.3).

**Tableau de Rubu's :**

1/ Côté négatif  $G_3 < 0$  (Tab.18)

| $G_3 < 0$ | Poids | Corrélation | Contribution |
|-----------|-------|-------------|--------------|
| H(2)4     | 7     | 71          | 30           |
| H(14)3    | 4     | 333         | 77           |
| H(24) 2   | 5     | 268         | 31           |
| H(24)3    | 3     | 206         | 54           |
| H(24)4    | 3     | 191         | 21           |
| H(25)1    | 3     | 226         | 36           |
| H(25)2    | 4     | 224         | 109          |
| H(33)3    | 4     | 49          | 12           |
| H(46)2    | 4     | 48          | 23           |

(Tab.18)

2/côté positif :  $F_3 > 0$  (Tab.19)

| $F_3 > 0$ | Poids | Corrélation | Contribution |
|-----------|-------|-------------|--------------|
| H(1)3     | 2     | 212         | 22           |
| H(17)1    | 4     | 158         | 20           |
| H(17) 2   | 5     | 427         | 59           |
| H(22)3    | 5     | 461         | 67           |
| H(17)3    | 4     | 224         | 32           |

|        |   |     |    |
|--------|---|-----|----|
| H(32)2 | 3 | 369 | 43 |
| H(32)1 | 4 | 301 | 26 |
| H(32)3 | 2 | 114 | 19 |
| H(47)4 | 5 | 224 | 24 |
| H(39)4 | 4 | 328 | 50 |

[Tab.19]

✚ **Tableau de mots** (côté négatif ; fig.5)

| <b>G3&lt;0</b>        | <b>Poids</b> | <b>Corrélation</b> | <b>Contribution</b> |
|-----------------------|--------------|--------------------|---------------------|
| Abraham               | 7            | 69                 | 23                  |
| Joseph                | 3            | 243                | 212                 |
| Ismaël                | 1            | 61                 | 14                  |
| Homme                 | 6            | 68                 | 28                  |
| Jacob                 | 2            | 127                | 22                  |
| Nuit                  | 8            | 51                 | 10                  |
| Malédiction           | 3            | 29                 | 23                  |
| Ceux qui font le bien | 3            | 63                 | 13                  |

[Tab.20]

✚ **Tableau de mots** (côté positif ; fig.5)

| <b>G3&gt;0</b> | <b>Poids</b> | <b>Corrélation</b> | <b>Contribution relative</b> |
|----------------|--------------|--------------------|------------------------------|
| Moïse          | 13           | 421                | 245                          |
| Fils d'Israël  | 4            | 179                | 30                           |
| Pharaon        | 7            | 360                | 152                          |
| Corrupteurs    | 2            | 67                 | 10                           |
| Aaron          | 2            | 82                 | 16                           |

[Tab.21]

Sur le côté négatif, on trouve la contribution de Joseph à cet axe est 21.2%, (Tab.20) on ajoute cette contribution à celle de Moïse (24.5%) cela dépassera 50% (Tab.21).

Les tableaux : (Tab.20), (Tab.21) montrent une correspondance parfaite entre les rubu's qui engendrent l'axe 3 mentionnés dans les tableaux (Tab.18) ;(Tab.19) et les mots cités dans les tableaux (Tab.20); (Tab.21).

La sourate de Joseph est emplie d'événements qui ne s'accordent pas avec notre logique :

i/ l'amour du père provoque la haine des frères. Joseph fût jeté dans le puits. Cet événement, bien qu'il soit un mal à nos yeux, lui a permis d'arriver en Egypte.

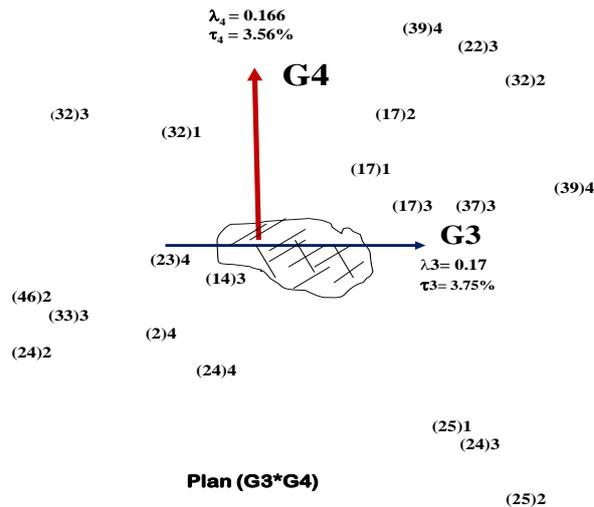
ii/ Son emprisonnement, qui ne semble qu'une rude épreuve, a fait de lui le « Aziz d'Egypte ».

iii/ La confiance placée par le roi en Joseph, due à sa haute moralité dans la maison d'Aziz, et à la sincérité de son interprétation des rêves des prisonniers.

iv/ Sa pensée stratégique était son ingénierie rationnelle alors qu'il courait des années de sécheresse et de prospérité dans le royaume.

Nous concluons de tout ce qui précède que nous devons nous en remettre à Allah, et c'est cette idée que nous allons renforcer tout au long de l'explication de la sourate [voir livre intitulé : « Explication scientifique du Sigle ALR ». [ kharchaf Idriss, Edition Okad1990.Rabat, Maroc]

La sourate de Joseph déborde également d'épreuves et de tentations, nous citons entre autres: En comparant la vie de deux prophètes Joseph et Moïse,



[Fig.5]

N.B : La correspondance entre les deux figures (Fig 4 ; Fig. 5) montre bien la stratégie des mots (Fig 4) dans les rubu's indiqués dans la figure 4 ci-dessus.

**D :1 :4 : Quatrième facteur : Facteur de Joseph ( $\lambda_4 = 0.166$  ;  $\zeta_4 = 3.56\%$ )**

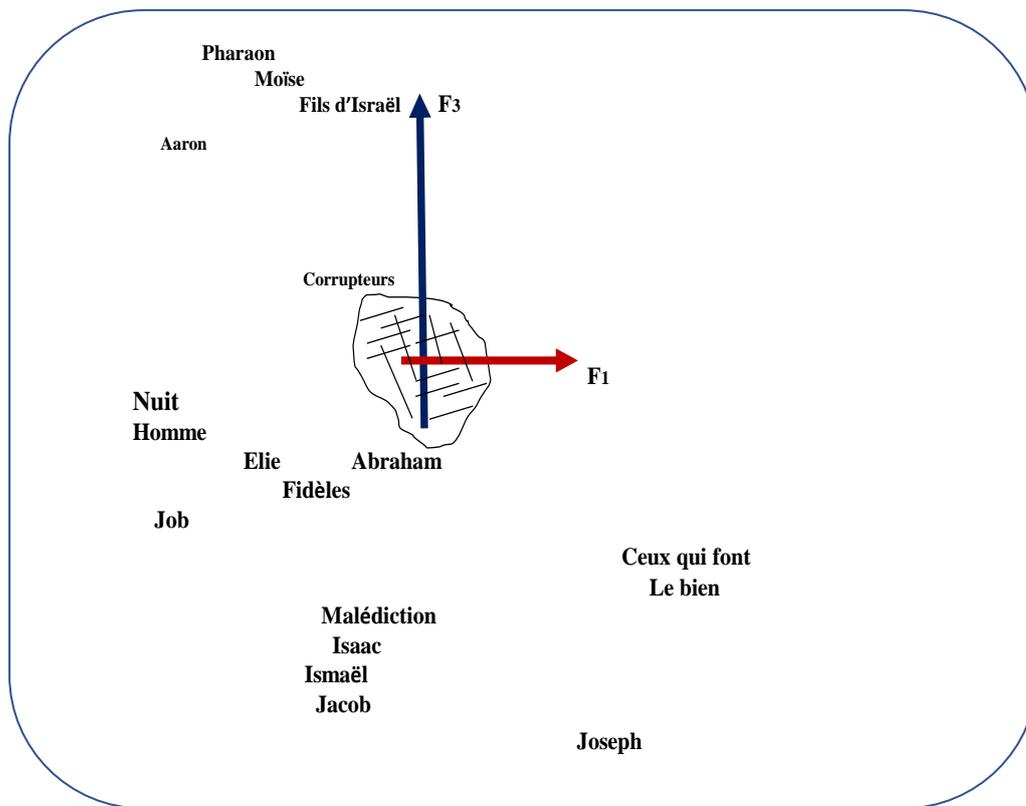
On remarque sur ce facteur que le rubu' H (25)2 est le plus représenté (Tab.22)

$$CTR (H (77, 100), F4) = 22.2\% \quad \text{Cos}^2 (H (77,100), F4) = 0.432$$

| Rubu'  | Corrélation | Contribution relative |
|--------|-------------|-----------------------|
| H(3)4  | 212         | 198                   |
| H(24)3 | 380         | 104                   |
| H(25)1 | 353         | 54                    |
| H(25)2 | 432         | 222                   |

[Tab.22]

La sourate de H (25)<sub>2</sub> est la sourate de Joseph, particulièrement appelée ainsi dans le Coran, car elle est un témoignage de toute époque (fig.4 ; 5)



[Fig. 6]

:

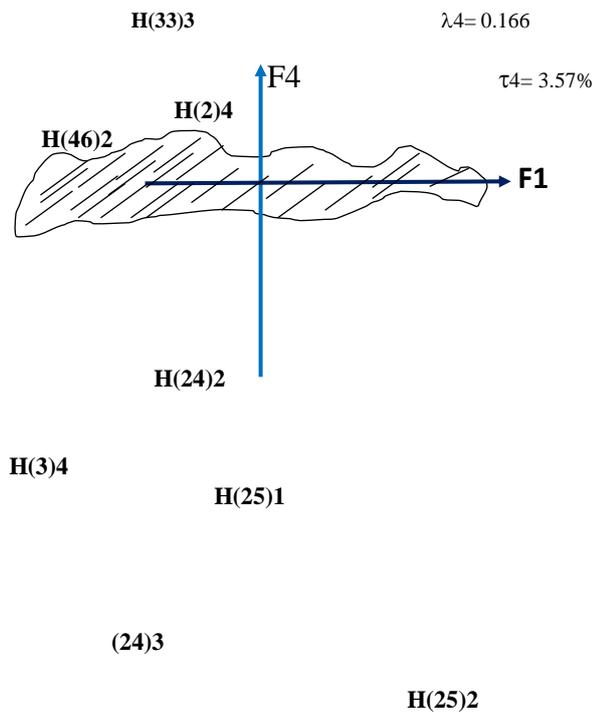
A/ Elle évoque l'histoire de Joseph avec ses frères

B/ Surtout elle révèle la grande foi de Joseph en DIEU, cette foi l'a protégé de tout malheur ordonné ainsi :(fig.5 , 6)

1. La haine qu'éprouvent ses frères à son égard.
2. Le fait d'être jeté dans le puits et l'éloignement de son père.
4. L'épreuve de l'esclavage et de la servitude.
5. Le désir et la séduction des femmes de la ville.

6. L'épreuve de la prison.
7. L'épreuve d'être attaqué dans l'honneur.
8. L'épreuve du pardon.

Huit points ordonnés selon l'histoire de Joseph sur Lui la prière et le salut.



[Fig. 7]

Typologie et forme du nuage (fig.8)

Dans le plan  $1*2$ , la projection nous permet de détailler géométriquement l'interprétation des nuages.

On remarque que la distribution des points par rapport à la parabole est bien interprétable.

En effet, d'après le principe du centre de gravité, un mot (لفظ) Ou Rubu' est d'autant plus à l'intérieure de la parabole, que les caractéristique qui leur ont été attribués sont dispersées sur l'échelle des valeurs.

Maintenant que nous connaissons géométriquement la formule du nuage, nous pouvons chercher la meilleure courbe d'ajustement du nuage.

Pour cela, on doit chercher un polynôme de second degré pour la branche parabolique :  $F_2 = \alpha (F_1)^2 + \beta F_1 + \gamma$

On remarque que  $F_2$  est un facteur de moyenne nulle et de variance  $\lambda\alpha$  ( $\alpha \in A$  : axes factoriels).

Si on prend la somme de chaque côté :

$$\sum\{F_2(r) * f_r / r \in \mathbb{R}\} = \alpha \sum\{(F_1(r))^2 * f_r / r \in \mathbb{R}\} + \beta \sum\{(F_1(r)) * f_r / r \in \mathbb{R}\} + \gamma \sum\{f_r / r \in \mathbb{R}\}$$

$$\text{On aura : } 0 = \alpha \lambda\alpha + \gamma \Rightarrow \gamma = -\alpha \lambda\alpha$$

Si de plus, on suppose que la forme est symétrique, c'est-à-dire que :

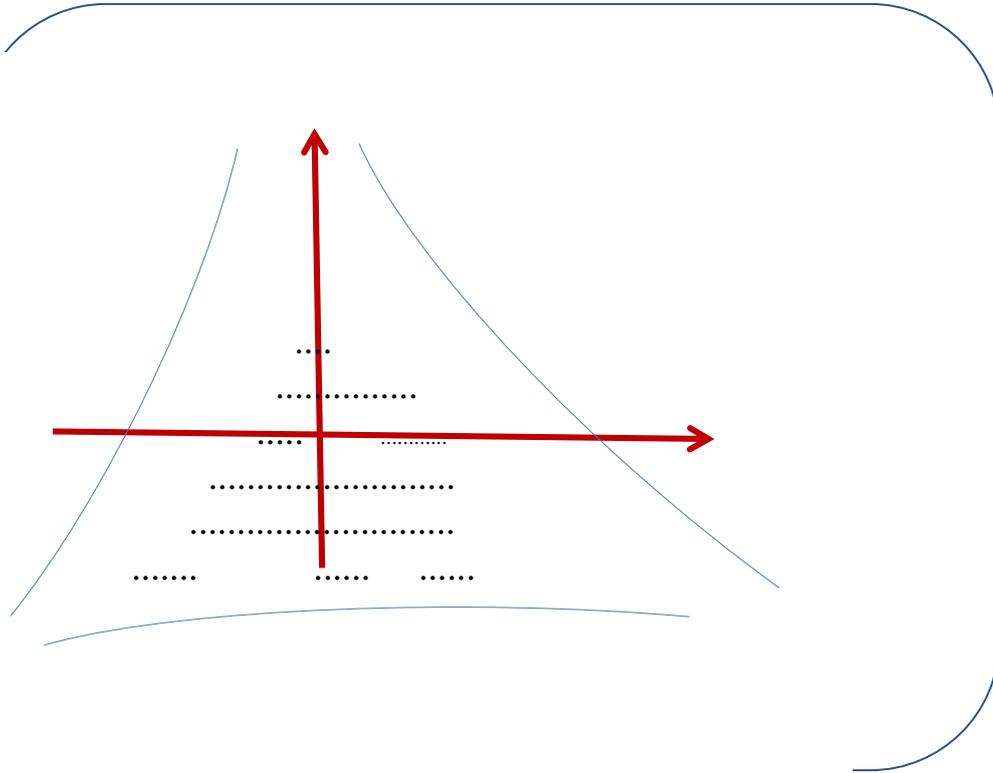
$$F(r) = F(-r) ; \text{ on obtient pour } \beta = 0 ; F_2 = \alpha (F_1)^2 - \alpha \lambda\alpha = \alpha (F_1^2 - \lambda\alpha)$$

En normalisant  $F_2$  on aura :

$$\begin{aligned} \text{var } F_2 = \lambda_2 &= \sum\{(F_2^2)(r) * f(r) / r \in \mathbb{R}_1\} = \sum\{\alpha^2((F_1)^2 - (\lambda\alpha))^2 * f(r) / r \in \mathbb{R}_1\} \\ &= \alpha^2 * 2 [\sum\{F_1^4 * f_r / r \in \mathbb{R}_1\} + \sum\{\lambda\alpha^2 * f_r / r \in \mathbb{R}_1\} - 2\sum\{\lambda\alpha (F_1)^2 * f_r / \\ &r \in \mathbb{R}_1\}] \end{aligned}$$

$$\text{Donc } \lambda_2 = \alpha^2 [\sum\{F_1^4 * f_r / r \in \mathbb{R}_1\} + \lambda_1^2 - 2\lambda_1^2] = \alpha^2 [M(F_1(r))^4 - \lambda_1^2]$$

$$\text{où } \alpha^2 = \lambda_2 / [M(F_1(r))^4 - \lambda_1^2]$$



(Fig.8)

## **Partie IV : Annexes**

### **Annexe 1 :**

#### **Le Coran, message universelle**

On sait (Selon les historiens) que les Mecquois étaient idolâtres, les idoles leur servaient d'intercesseurs auprès de Dieu. Ce Dieu avait un temple, une maison, qui est un cube de pierre, construit sur trois pierres sacrées. Les besoins de Dieu étaient administrés par une famille de prêtres dont l'office se transmettait de génération en génération.

En effet, si nous retournions au VII<sup>ème</sup> siècle, nous constaterions que la somme des connaissances humaines était peu importante.

L'homme d'Arabie avait une conception étroite, ses idées sur la création se bornaient à son horizon limité, et le milieu était plongé dans les ténèbres de l'ignorance, il ne pouvait pas acquérir le savoir des pays voisins (la Perse, la Grèce, l'Égypte) ; même pendant la période de commerce, il vivait dans un État sans loi, la propriété et l'honneur étaient menacés, il avait le droit de tuer même ses filles, il a oublié qu'il était le descendant d'Adam et Ève .Sa haine l'avait ramené à un niveau plus bas que celui d'une bête. Une minorité de personnes se retirait du monde et s'adonnait aux pratiques spirituelles, mais la société humaine n'en tirait aucun profit, elle avait besoin de la direction d'une religion.

C'est à cette époque et dans ce contexte géographique et social que va se produire ce grand événement qu'est la naissance de l'envoyé de l'Islam Mohammed (Dieu Lui accorde la bénédiction et la paix), Dieu le très haut dit: « Mohammed, envoyé de Dieu »[ محمد ]

رسول الله (XLVIII ; 29) qui suscita à la Mecque un réveil de la vie spirituelle, plein de vigueur, mais modeste et charitable, illettré mais intelligent et honnête, Mohammed est appelé Al Amin (homme de confiance) «الأمين» par les quraïchis.

A l'âge de 40 ans, il quitta sa vie privée pour se retirer dans une grotte et passa les jours dans la plus complète solitude, il est à la recherche d'une lumière qui puisse dissiper les ténèbres environnantes.

Soudain, son cœur est illuminé par la lumière divine, il écoute la première révélation du Seigneur le tout-puissant, lorsque le Seigneur des cieux et de la terre, dit :

*« Lis au nom de ton Seigneur qui a créé, qui a créé l'homme d'une adhérence, Lis car ton Seigneur le Noble, qui a enseigné par la plume (Al Qalam), Il a enseigné à l'homme ce qu'il ne savait pas »*  
[اقرأ باسم ربك الذي خلق، خلق الإنسان من علق وربيك الأكرم الذي علم بالقلم علم الإنسان ما لم يعلم]  
(XCVI ; 1-5).

Ce message universel va faire sortir l'homme des ténèbres de l'ignorance à la lumière de la Science. Il va, dès lors, transformer un agglomérat de tribus hostiles les unes aux autres, en une nation qui deviendra le centre de la culture et de la civilisation, grâce au message qui porte le symbole de « la lecture scientifique et la lecture spirituelle ».

Donc le coran procure aux musulmans l'incitation nécessaire pour partir dans la voie des conquêtes intellectuelles pour acquérir des connaissances dans toutes les spécialités et à accroître leur savoir dans tous les domaines en étudiant les lois de la nature et en tirant profit de la création de Dieu.

Aussi, l'univers a-t-il été créé pour le bénéfice de l'homme :

*« Pour vous, Il assujettit la nuit et le jour ; le soleil et la lune. Et à son ordre sont assujetties les étoiles. Voilà bien là des preuves pour des gens qui raisonnent »*. [وسخر لكم الليل والنهار والشمس والقمر والنجوم]. (XVI ; 12)

« en vérité, dans la création des cieux et de la terre, dans l'alternance de la nuit et du jour, il y a certes des signes pour les douer d'intelligence » [إن في خلق السماوات والأرض واختلاف الليل والنهار لآيات لأولي الأبواب]. (III ; 190)

Ainsi, le Livre sacré, le Coran, fût révélé dans une époque qui s'est distinguée des précédentes par le grand nombre d'hommes éloquents et habiles dans l'art de la parole.

Son apparition était le plus grand des miracles et la preuve tangible que le Coran est l'œuvre de Dieu). Il a bien démontré à travers ses nombreux versets, que la foi islamique ne s'oppose pas à la science qui peut être définie, comme une tentative pour comprendre le monde en retraçant les rapports de divers éléments les uns avec les autres pour en connaître les causes. « Dis : « Sont-ils égaux, ceux qui savent, et ceux qui ne savent pas ? » [قُلْ هَلْ يَسْتَوِي الَّذِينَ يَعْلَمُونَ وَالَّذِينَ لَا يَعْلَمُونَ] (XXXIX ; 9)

« Et afin que ceux à qui le savoir a été donné sachent que (le Coran) est en effet, la Vérité venant de ton Seigneur, qu'ils croient alors, et que leurs cœurs s'y soumettent en toute humilité Allah guide certes vers le droit chemin ceux qui croient ». [وَلِيَعْلَمَ الَّذِينَ أُوتُوا الْعِلْمَ أَنَّهُ الْحَقُّ مِنْ رَبِّكَ فَيُؤْمِنُوا بِهِ فَتُخْبِتَ لَهُ قُلُوبُهُمْ وَإِنَّ اللَّهَ لَهَادِ الَّذِينَ آمَنُوا] (XXII ; 54) [إلى صِرَاطٍ مُسْتَقِيمٍ]

Bien au contraire, elle a consolidé le lien entre ce qui est matériel et ce qui est moral, en prescrivant à chaque musulman et musulmane le devoir de la science, et en leur recommandant d'en puiser tout au long de la vie.

« Ô, vous qui avez cru ! Quand on vous dit : « Faites place (aux autres) dans les assemblées », alors faites place : Allah vous ménagera une place. Et quand on vous dit : « de vous lever » levez-vous. Allah élèvera en degrés ceux d'entre vous qui aurez cru et ceux qui auront reçu le savoir. Allah est parfaitement connaisseur de ce que vous faites ». [ يَا ]

أيها الذين آمنوا إذا قيل لكم تفسحوا في المجالس فافسحوا يفسح الله لكم وإذا قيل لكم انشزوا فانشزوا يرفع الله الذين آمنوا منكم والذين أوتوا العلم درجات والله بما تعملون خبير] (LVIII ; 11)

« Nous leur avons certes, apporté un livre que Nous avons détaillé, en toute connaissance, à titre de guide et de miséricorde pour les gens qui croient » [VII ; 52] *ولقد جنناهم بكتاب فصلناه على علم هدى ورحمة لقوم يؤمنون*

« Il consiste plutôt en des versets évidents (préservés), dans les poitrines de ceux à qui le savoir a été donné. Et seuls les injustes renient nos versets » [XXIX ; 49] *بل هو آيات بينات في صدور الذين أوتوا العلم وما يجحد بآياتنا إلا الظالمون*

« Telles sont les paraboles que Nous citons aux gens, cependant, seuls les savants les comprennent; *الأمثال نضربها للناس وما يعقلها إلا العالمون* (XXIX :43) [ *وتلك* ] « Nous avons (effectivement) bien détaillé les signes à des gens qui se rappellent » [VI ; 126] *قد فصلنا الآيات لقوم يذكرون*.

Donc l’Islam fait de la recherche scientifique et rationnelle, un devoir sacré et noble, « La recherche de l’acquisition de la science est une obligation pour tout musulman » (hadith cité par Ahmed).

L’accomplissement de cette obligation religieuse doit donc être réalisé avec piété et dévotion, sincérité et loyauté intellectuelle. Ace propos, les livres d’histoire nous rapportent que le prophète exigeait de chaque prisonnier de BADR qui savait écrire, d’enseigner l’écriture à dix enfants de médina pour recouvrer sa liberté tandis que les autres étaient libérés parce qu’ils étaient pauvres.

En plus, la science et la technique islamiques (mécanique, astronomie, optique...) étaient étroitement liées, il en était de même pour les sciences fondamentales et les sciences appliquées.

De même, l’Islam n’a pas rejeté la révolution technologique occidentale, bien au

contraire, toute la philosophie de développement dans les pays musulmans repose sur le recours aux acquisitions techniques et aux transferts technologiques.

En vue d'atteindre la vérité scientifique et de comprendre la réalité des phénomènes (l'existence de l'être humain, La vie dans le cosmos) Sommes-nous seuls dans l'Univers ?), les lois naturelles, etc.), le Coran a incité l'homme à exercer ses fonctions de méditation, d'analyse et d'expérience en se basant sur les règles et méthodes scientifiques. Dieu le très haut dit :

« Par la sagesse et la bonne exhortation appelle (les gens) au sentier de ton Seigneur. Et discute avec eux de la meilleure façon » [XVI ; 125] ادع إلى سبيل ربك بالحكمة والموعظة الحسنة وجادلهم بالتتي هي أحسن]. « meilleure façon »

Dans nos relations et entretiens avec les gens, le Coran nous ordonne à les appeler au chemin de Dieu, en se basant sur les méthodes scientifiques (recherche, analyse, fouille des données, dialogue, preuves, argumentation, etc.) Ce verset coranique enracine les principales méthodes scientifiques de déduction, d'induction et de discussion. Les méthodes (déduction et induction) sont exprimées par deux termes universels : « al Hikmat » qui englobe la sagesse, la raison, l'intelligence, la pensée, la connaissance des choses, le savoir, la science et la prudence. Et la bonne exhortation « al Mawada Al Hassanat والمودة والحسنات » c'est-à-dire la bonne exhortation qui signifie la meilleure et excellente leçon, le bon exemple, le raisonnement logique et la preuve, alors que dans la méthode de controverse ou de discussion, le Coran exige une discussion meilleure, satisfaisante et qui aboutit à un bon résultat.

Parmi les versets scientifiques, il y a ceux qui enracinent la méthode de déduction par la méditation naturelle et instinctive par la spéculation des finesses, de l'existence en commençant par la vérité de soi puis par celle des phénomènes de l'univers. Dieu, le Très Haut dit :

« Nous leur montrerons Nos signes dans l'univers et en eux –mêmes, jusqu'à ce qu'il leur devienne évident que c'est cela (le Coran) la Vérité ». [سنريهم آياتنا في الآفاق وفي أنفسهم حتى يتبين لهم أنه الحق] (XLI ;53) cela (le Coran) la Vérité ».

« Dis : Louange à Allah : Il vous fera voir Ses preuves, et vous les reconnaîtrez, Ton Seigneur, n'est pas inattentif à ce que vous faites ». [وقل الحمد لله سيريكم آياته فتعرفونها وما ربك بغافل عما تعملون]

(XXVII :93)

Le Coran, par les signes scientifiques, les concepts et les instructions, éveille dans l'esprit humain le sens de la recherche scientifique, afin de percevoir la vérité (que tout être humain cherche à connaître.).

Celui qui lit le Coran remarquera que plusieurs versets incitent l'homme à l'observation, à la réflexion à l'analyse, à la recherche et à la discussion, éléments de la méthode d'induction. Parmi ces versets, nous citons :

« Dis regardez ce qui est dans les cieux et sur la terre » [قل انظروا ماذا في السماوات والأرض] (X ;110)

« Dis : Voyez-vous ? Si Allah prenait votre ouïe et votre vue, et scellait vos cœurs, quelle divinité autre qu'Allah vous les rendrait ? Regarde comment, à leur intention, Nous clarifions les preuves ! Pourtant ils s'en détournent ». "

قل أرأيتم إن أخذ الله سمعكم وأبصاركم وختم على قلوبكم من إله غير الله يأتيكم به انظر كيف نصرّف الآيات ثم

"هم يصدفون" [VI ;46]

« Dis : « Il est capable, Lui, de susciter contre vous, d'en haut ou de dessous vos pieds, un châtimeur, ou de vous confondre dans le sectarisme. Et Il vous fait goûter l'ardeur (au combat) les uns aux autres, Regarde comment Nous exposons Nos versets ! Peut-être comprendront-ils ? »

"قل هو القادر على أن يبعث عليكم عذابا من فوقكم أو من تحت أرجلكم أو يلبسكم شيئا ويذيق بعضكم

[بأس بعض انظر كيف نصرّف الآيات لعلهم يفقهون" (VI : 65)]

Au cours d'une recherche scientifique, le Coran affirme la méthode qui s'appuie sur la preuve, l'argument et la bonne discussion afin d'atteindre les bons résultats, écoutons quelques versets :

« Ô gens ! Certes une preuve évidente vous est venue de la part de votre Seigneur. Et Nous avons fait descendre vers vous une lumière éclatante. »

[يا أيها الناس قد جاءكم برهان من ربكم وأنزلنا إليكم نورا مبينا] (IV ; 174)

« Nous avons fait du ciel un toit protégé. Et cependant ils se détournent de ses signes »

[وجعلنا السماء سقفا محفوظا وهم عن آياتها معرضون] (XXI ; 32)

Toute réalité naturelle est régie par des lois et des règles que Dieu a posées, avec lesquelles il accomplit sa volonté et son acte (Exemple : l'unification du modèle standard de la physique des particules et de la cosmologie qui permet la proposition des scénarios de différenciation des quatre interactions fondamentales c'est: gravitationnelle, nucléaire-forte, nucléaire- faible, et électromagnétique). Ces lois permettent de comprendre la signification des phénomènes complexes, et de la réalité, le Coran les a exprimés par les causes et nous conseille de les découvrir par le biais de la recherche scientifique,

« Il en fut ainsi et Nous embrassons de Notre Science ce qu'il détenait. Puis il suivit (une autre voie)».

[كذلك وقد أحطنا بما لديه خبرا ثم اتبع سببا] (XVIII ; 91,92)

## ANNEXE 2

### TECHNIQUES DE MODELISATION

Le datamining met en œuvre un ensemble de technique issues des méthodes suivantes :

1 :Méthode statistique : dans une première étape on prend connaissance avec ses données, en étudiant les statistiques et indicateurs élémentaires a partir du tableau de base  $k(m,v)$ .

2 :Choix de codage : choix qui doit être cohérent avec le choix de la métrique (voir n°3).

Parmi les situations de codage les plus utilisées, citons la construction des tableaux suivants :

2 : a : le tableau logique (oui, non).

2 : b : le tableau disjonctif complet (Votes au sein des Nations unies : oui, non, abstention, absent)

2 : c : le tableau de Burt associé à un tableau logique, ou mis sous forme disjonctive complète.

2 : d : le découpage en classes d'éléments mesurés dans des unités différentes (découpage en classes d'effectifs égaux, selon la forme des histogrammes, selon les bornes fixées à l'avance par l'analyse.

### **3 : Choix de la métrique**

3 : a :la métrique de  $\chi^2$  (voir partie mathématique).

3 : b : information

3 : c : indices de similitude (Indice de Jaccard, coefficient de corrélation. etc.)\*.

|                    |         |        |
|--------------------|---------|--------|
| Verset V<br>Mots M | Présent | Absent |
| Présent            | a       | b      |
| Absent             | c       | D      |

[Tab.A1] Exemple : Indice de similitude

(\*) Indice de similitude de Jaccard =  $a / [a+b+c]$

Indice : Coefficient de corrélation =  $(ad - bc) / [(a+b)(c+d)(a+c)(b+d)]$

### ANNEXE :3

#### A3.1. L'Analyse factorielle des correspondances (Généralités)

Cette partie comporte deux étapes :

La première étape rassemblera un formulaire de l'analyse des correspondances, tout en essayant de reprendre tous les développements mathématiques qui y sont liés.

La deuxième étape parcourt le domaine de l'analyse du Coran, en faisant recours à l'analyse factorielle des correspondances.

#### A3.2. L'objectif principal de l'analyse des données

L'objectif principal de l'analyse des données est, d'élaborer et de représenter géométriquement dans un espace euclidien de dimension  $p < n$ , les informations les plus diverses, autrement dit .....

Son application est pratiquement générale, il s'agit d'assurer une meilleure adéquation formation-insertion dans la vie quotidienne (Médecine, Astronomie, Biologie, Linguistique, Hadiths, Coran, Data Mining, Classification automatique,

Géologie, Paléontologie, Economie, Ecologie, Transport, traitement d'image, Politique Internationale, chimie, Biométrie, etc.

### **A3.3. Outils utilisés**

L'objet principal de ce programme est, l'étude analytique du Coran, technique de Fouille de Données –Data Mining-).

La saisie des données sur un support technique en vue de leur analyse et traitement automatique ne peut pas être le fruit du hasard pour suivre un modèle, et elle ne peut être décidée qu'en vertu d'une réflexion rationnelle, et d'une stratégie bien définie des recherches dans lesquelles elle s'intègre ou qu'elle va déterminer.

Il faut donc l'aide d'un ordinateur qui permettra le traitement jusqu'au bout de volumineux calculs et données (Big Data) telles qu'on puisse lire avec sûreté à la sortie ce qui, à l'entrée, étaient indéchiffrables.

### **A3.4. Caractéristiques de l'Analyse des données**

L'analyse des données est caractérisée par un fond commun d'idées de base :

\*Utiliser de façon efficace l'informatique du plus puissant des ordinateurs au micro-ordinateur.

\*Réduire au maximum la subjectivité et les a priori, rester proche des données en les déformant le moins possible.

\*Mettre au point des approches linéaires ou non linéaires pour la recherche de directions révélatrices utilisant les nouvelles possibilités graphiques et des critères de type nouveau issus par exemple de techniques dites de « projection poursuite ».

### **A3.5. Base des données**

L'étude qui été faite porte sur le traitement du Coran, en faisant recours aux techniques d'extraction de connaissance à partir des données coraniques (EDC).

### A3.6. Pourquoi choisir les données coraniques ?

Le Coran est une révélation qu'Allah a fait descendre sur Mohammed (Que Dieu Lui accorde la bénédiction et la paix) vers l'âge de 40 ans.

Il comporte 114 sourates dont le classement a été fait par le prophète Mohammed (Que Dieu Lui accorde la bénédiction et la paix).

Coran signifie lecture, Il est la source primordiale des enseignements de l'Islam (Il vérifie scientifiquement les deux conditions citées ci-dessus), que ce soit par le biais de l'élucidation de ses versets, ou à travers la somme infinie de valeurs, concepts et principes moraux qui s'en dégagent.

De plus, le Coran cherche à guider l'homme universel (sans distinction de races, de régions, ni même d'époques) dans tous les domaines de sa vie (spirituel, individuel, scientifique, social, etc.), en lui demandant de contempler l'univers de l'infiniment petit à l'infiniment grand.

A3.7. Le coût de l'analyse (voir article KHARCHAF IDRIS « Sur la recherche des plus proches voisins suivant une décomposition cellulaire de l'espace en classification », cahiers de l'analyse des données CAD: Vol. XII, n° 2 - 1987. Centre National de la Recherche Scientifique (CNRS) - Paris - France.

Le coût d'une telle opération et les contraintes qu'elle implique, sont si importants que seule l'analyse minutieuse de la démarche et la claire conscience des conséquences qui vont en découler doivent être maîtrisées par avance, pour que cette action soit entreprise avec un minimum de risque et un maximum de succès.

De même, les mathématiciens-statisticiens notamment les utilisateurs de *Data Mining* (étape clé dans l'extraction de connaissances) sont fréquemment confrontés au problème suivant :

### A3.8. Réduction de Base de données (voir l'article)

Lorsqu'il est possible de réduire les ensembles de données sans pour autant affaiblir la compréhension de la correspondance, (exemple : Analyse factorielle des correspondances, la classification hiérarchique : construction des classes de Base de données), la difficulté précédemment signalée peut-être surmontée, en mettant les colonnes ou les lignes d'un tableau de données en élément supplémentaires, que l'on sait ou que l'on ne connaît pas.

Une telle démarche revient cependant à abandonner une partie de l'information ou plutôt certains aspects des problèmes étudiés.

La solution est donc de lever cette difficulté, tentant de mesurer l'angle intercepté par les plans engendrés par les facteurs trouvés (vecteurs propres) à chaque analyse, et chercher l'angle de déviation (D), lorsque D est minimum.

## ANNEXE4 : METHODE MATHEMATIQUE DE L'ANALYSE

### FACTORIELLE DES CORRESPONDANCES (VOIR ANNEXE 2)

L'analyse factorielle des correspondances a pour but d'approcher, par des variétés linéaires (droites, plans, tétraèdre, etc.), et de représenter les nuages dans un espace de dimension  $t_{1, 2, 3, \dots} \leq p$  ; ce qui revient à rechercher le meilleur espace de dimension  $t$ , tel que le nuage soit représenté par sa projection.

Dans cette partie nous présentons, la méthode mathématique, ensuite l'ensemble et le choix des mots, ce dernier représente la première phase essentielle de la recherche.

Supposons qu'on ait  $n$  individus (mots, mesures, températures.), et que sur chaque individu, on fait  $p$  mesures (fréquences, degré de liberté, taux d'apparition) chacun de ces  $n$  individus est représentable par un point de  $i^p$  ; les coordonnées d'un individu  $n$ 'étant autres que les résultats des  $p$  mesures effectuées sur cet individu.

On a donc un nuage de  $n$  points dans  $i^p$ , nuage non visualisable, dès que  $p > 3$ . le but de l'analyse factorielle est de trouver la représentation la meilleure du nuage.

#### A4.1. Mesures et fréquences relatives (Rappels)

##### A4.1.1/ $\sigma$ - anneau (ou $\sigma$ - clan)

On appelle  $\sigma$ -anneau (de parties de  $\Omega$ ), une famille non vide où  $\Omega$  un ensemble non vide  $\mathcal{B}$  de parties de  $\Omega$ (un ensemble non vide) telle que :

**A4.1.1. i/**  $B_1, B_2 \in \mathcal{B} \Rightarrow B_1 \cap B_2 \in \mathcal{B}$

**A4.1.1. ii/**  $\forall j : B_j \in \mathcal{B} \Rightarrow \bigcup_{j=1, \infty} (B_j) \in \mathcal{B}$

**A4.2.  $\sigma$ -algèbre (Tribu),** si  $\Omega \in \sigma$ -anneau

**A4.3. Tribus d'événements**

Soit  $\Omega$  un ensemble non vide. On appelle « tribu » de parties de  $\Omega$ , tout ensemble  $\mathcal{F}$  de parties de  $\Omega$  vérifiant les axiomes suivants :

a.  $\Omega \in \mathcal{F}$

b. si  $A \in \mathcal{F}$  alors  $(\bar{A}) \in \mathcal{F}$  ( $\bar{A}$  est le complémentaire de  $A$ )

c. si  $\{A_i\}_{i \in \mathbb{N}} \in \mathcal{F}$  alors  $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$

Le couple  $(\Omega, \mathcal{F})$  est appelé espace mesurable

**A4.4. Fréquence relative**

Soit un espace mesurable  $(\Omega, \mathcal{F})$ , on appelle fréquence relative sur la tribu  $\mathcal{F}$ , toute fonction  $f : \mathcal{F} \rightarrow [0, 1]$  tels que :

a-  $f(\Omega) = 1$

b-  $f(\bigcup_{i=1, n} A_i) = \sum_{i=1, n} f(A_i)$  ;  $\forall \{A_i\}_{i \in \mathbb{N}} : A_i \cap A_j = \emptyset$  pour  $i \neq j$

**A4.5. Espace de fréquence relative**

Tout triplet  $(\Omega, \mathcal{F}, f)$  formé de :

a.  $\Omega \neq \emptyset$

b.  $\mathcal{F} = \mathcal{P}(\Omega)$ , où  $\mathcal{P}(\Omega)$  représente les parties de  $\Omega$ .

c.  $f$  est une fréquence relative sur  $\mathcal{F}$ .

**A4.6. Fréquence conditionnelle**

Soit  $(\Omega, \mathcal{F}, f)$  l'espace de fréquence relative, soit  $B \in \mathcal{F}$  où  $f(B) > 0$ .

On définit la fréquence conditionnelle :  $f(A/B) = f(B \cap A) / f(B)$ , notée  $f_{B/A}$

**A4.7. Simplexe : Dans  $\mathbb{R}_I$**

L'ensemble  $\phi_I$  des mesures (ou lois de fréquences relatives sur I, mesure de masse totale égale 1),  $\phi_I = \{ m_i / m_i \in \mathbb{R}_I; \sum \{ m_i / i \in I \} = 1, \forall i \in I \}$ .

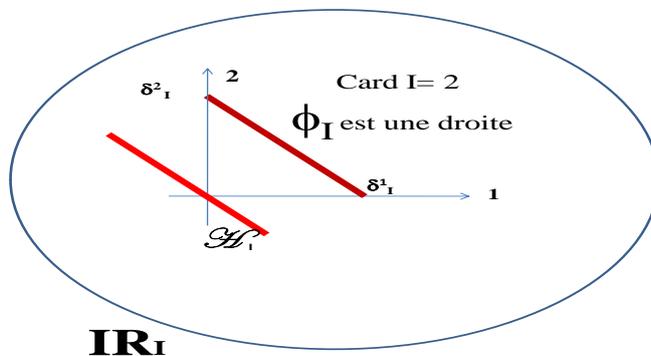
si card I=2 alors  $\phi_I$  est alors un segment (voir Fig :1)

card I=3 alors  $\phi_I$  est alors un triangle (voir Fig. :1)

[Fig A :1]

**A4.8. Hyperplan :  $\mathcal{H}_I$**

$\mathcal{H}_I = \{ m_i / m_i \in \mathbb{R}_I; \sum \{ m_i / i \in I \} = 0 \}$  (sous -espace vectoriel de Co-dimension 1) est un hyperplan



**[Fig A :1]**

**A4.9 : Moyenne et covariance d'une fonction :**

Soit une fonction  $h^I$  : est une variable aléatoire qui prend la valeur  $h^i$ , avec la fréquence relative  $f_i$ .

On définit  $moy(h) = h^I \circ f_I = \sum \{ h^i f_i / i \in I \}$ .

Et la covariance  $(u_I, v_I)$  : notée par commodité d'écriture :  $cov(u,v)$

$Cov (u,v) = moy (u - moy(u), v - moy (v)) = Esp ([ u - moy (u)], [v - moy(v)]) =$

$$= \sum \{u^i v^i f_i / i \in I\} - \sum \{u^i f_i / i \in I\} * \sum \{v^i f_i / i \in I\}$$

Etant sur l'espace vectoriel :  $H^I(f_I)$  (des fonctions de moyenne nulle). Pour  $u=v$  (forme bilinéaire), alors la covariance définit une forme quadratique définie par :  
 $\text{Cov}(h,h) = \text{Esp}(h^2) = \sum \{(h^i)^2 f_i / i \in I\}$ .

#### A4.10. Relations entre les deux espaces vectoriels $H^I$ et $H^I(p_I)$

A4.10.1/ Ces espaces et leurs duals sont isomorphes entre eux.

A4.10.2/ ces espaces ont pour dimension = Card (I)-1

A4.10.3/  $m_I \rightarrow (m/f)^I ; \forall i \in I : (m/f)^i = m_i / f_i$

Et  $h^I \rightarrow (h f)_I ; \forall i \in I : (h f)_i = h^i f_i$

L'Isomorphisme réciproque associant à la fonction  $h^I$  de moyenne nulle, la mesure dont elle est la densité par rapport à  $f_I$  :

$$h^I \rightarrow (hf)_I = (h^I \times \delta_I^I) \circ f_I ; \forall i \in I : (h f)_i = h^i f_i$$

C'est-à-dire que  $(h f)_I$  est l'image de  $f_I$  par la transition (associée à  $h^I$ ) qui à la masse ponctuelle unité  $\delta_I^I$ , fait correspondre la masse ponctuelle  $h^i$  placée au même point  $i$ .

Pour une mesure  $\mu_I$ , on calcule une fonction appelée densité relativement à  $f_I$ , c'est-à-dire  $\mu^I = (\mu/f)^I : \mu^j = (\mu/f)^j$ , c'est-à-dire  $\mu^j = \mu_j / f_j$ .

#### A4.11. Relations entre les deux espaces vectoriels :

$\mathcal{H}^I$  (sous-espace de  $\mathbb{R}^I$ ) &  $\mathcal{H}^I(f_I)$  (sous-espace de  $\mathbb{R}^I$ , et  $f_I$  fréquence relative).

$\mathcal{H}^I(f_I)$  : un sous-espace des fonctions de moyenne nulle, c'est un hyper plan de  $\mathbb{R}^I$ , autrement dit :  $\mathcal{H}^I(f_I) = \{ \mathcal{H}^I \in \mathbb{R}^I : \sum \{h^i f_i / i \in I\} = 0$

entre  $\mathcal{H}_I$  et  $\mathcal{H}^1(f_I)$ , il y a un isomorphisme qui à la mesure  $m_I$  (de masse totale zéro) associe sa densité par rapport à la mesure  $f_I$  (densité : dont la moyenne  $\text{moy}(f_I) = 0$ ).

L'isomorphisme réciproque associant à la fonction  $h^I$  ( $\text{moy}(h^I) = 0$ ), la mesure  $h^I \rightarrow (h^I f)_I = (h^I \times \delta^I_I) \circ f_I$  ;  $\forall i \in I : (h^I f)_i = h^I_i f_i$ .

Pour une mesure  $\mu^J = (\mu / f)^J$  :  $\mu^j = (\mu / f)^j = (\mu_j / f_j)$

$$\text{Alors } \mu^j = \mu_j / f_j$$

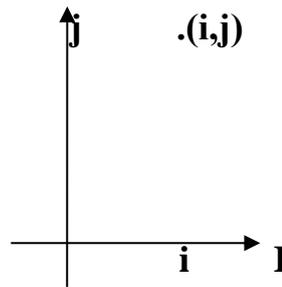
#### A4.12. Projection de I (Fig.A.4.2)

Soit  $I \times J = \{(i, j) / i \in I ; j \in J\}$

La projection sur I de  $I \times J$  : notée  $\text{proj}_I(I \times J) = \delta^{I \times J}_I$

La projection sur J de  $I \times J$  : notée  $\text{proj}_J(I \times J) = \delta^{I \times J}_J$

$$\delta^{i \times j}_i = \begin{cases} 1 & \text{si } i = i' \\ 0 & \text{sinon} \end{cases}$$



[Fig.A.4.2]

#### A4.13. Identité

C'est une application  $I \rightarrow I : i \rightarrow i$ , A cette application est associée la transition  $\delta^I_I$  :

$$\forall (i, i') \in I \times I \quad ; \quad \delta^i_{i'} = \begin{cases} 1 & \text{si } i = i' \\ 0 & \text{sinon} \end{cases}$$

#### A4.14. Application diagonale

**Soit**  $I \rightarrow I \times I : i \rightarrow i \times i$ , A cette application est associée la transition  $\delta^I_{I \times I}$  telle que :

$$\forall i, i', i'' \in I \quad : \delta^i_{i' i''} = \begin{cases} 1 & \text{si } i = i' = i'' \\ 0 & \text{sinon} \end{cases}$$

**A4.15. Application constante :**  $I \rightarrow \{ 1 \}$

## **Annexe5 : Statistique & fouille des données (D.M)**

Définitions

### **A5.1. Population**

Définition : Une population (finie) est un ensemble  $\Omega$  d'éléments  $\omega$  muni d'une pondération  $\alpha$ , autrement dit, elle comprend tous les êtres ou objets ou objets d'un groupe bien défini sur lequel porte notre étude.

### **A5.2. Echantillon**

Définition : l'échantillon est un groupe relativement petit et choisi scientifiquement de manière à représenter le plan fidèlement possible une population.

### **A5.3. Exemples :**

- l'ensemble des versets coraniques de la sourate « Al Baqara ».
- l'ensemble de tous les versets de la sourate «Arrahmane»(Miséricordieux).
- les mots cités dans la lettre prophétique Mohammed (Que Dieu Lui accorde la bénédiction et la paix) destinée au roi de l'Ethiopie

### **A5.4. Variables quantitatives & qualitatives**

En statistique, l'information est une chose capitale, elle se présente généralement sous la forme suivante : Un certain nombre d'informations sont décrites par un certain nombre de caractères.

Les caractères peuvent être quantitatifs ou qualitatifs, Avec l'ensemble d'informations on constitue des tableaux qui sont à la base de l'analyse.

### **A5.5. Notations**

Un certain nombre de symboles est essentiel pour la description des méthodes statistiques. Malheureusement, il y a désaccord entre les auteurs, ce qui peut créer des confusions pour le débutant.

Pour cette raison, un essai de méthode conforme de notation est fait dans ce chapitre.

### **A5.6. Note aux étudiants**

L'étudiant qui travaille consciencieusement la statistique, apprend comment utiliser les méthodes connues de statistiques.

Il ne devra pas oublier cependant que les méthodes statistiques sont essentiellement des instruments pour un travail de recherche, elles le rendent capable de fixer la valeur de ses observateurs.

L'étudiant averti doit d'abord, essayer d'acquérir et de comprendre le but spécifique de chaque méthode statistique puis, apprendre à utiliser les formules et les relations scientifiques.

### **A5.7. Collecte de données**

Les collectes basées sur un dénombrement complet présupposent :

**A5.7.a/** l'existence d'un certain minimum de moyens.

**A5.7.b/** ressources financières statisticiens pour la planification

**A5.7.c/** des méthodes de collecte.

**A5.7.d/** la surveillance générale des opérations

**A5.7.e/** agents recenseurs suffisamment qualifiés.

**A5.7.f/** documents cartographiques, matériels informatiques de  
Classement des informations, etc.

Or ces moyens n'existent pas toujours, de sorte qu'il est impossible d'effectuer un recensement.

Alors lorsque n'existent pas les moyens d'effectuer la collecte par dénombrement complets, la mise au point d'une théorie et d'une pratique des enquêtes par sondage, offre actuellement une deuxième formule commode, qui est celle des collectes par sondage.

La collecte par sondage consiste à réunir les informations voulues pour les besoins d'une collecte usuelle, en s'adressant seulement à un échantillon d'unités appartenant à l'univers que l'on veut recenser.

Par conséquent, les collectes des données par sondage sont essentiellement des enquêtes sur échantillon.

Toutefois seules les enquêtes sur échantillons, soit le but est de fournir des informations d'ordre général sur un grand nombre de caractéristiques inscrites d'ordinaire dans les programmes de recensement, seront désignées par le terme de collectes par sondage.

### **A5.8. Comment choisir adéquatement un échantillon ?**

Supposons que l'on désire connaître la grandeur des habitants d'une petite ville, on désire un échantillon de  $n$  personnes, et on désire que l'échantillon représente le

plus possible, la population de la petite ville, ce qui permettra de généraliser les résultats.

Comment choisir un échantillon pour qu'il soit représentatif ?

**A5.8.1/** Un échantillon peut être choisi automatiquement :

**A5.8.1: a /** les unités qui appartiennent à l'échantillon sont connues progressivement au fur et à mesure que se déroule le processus de dénombrement.

**A5.8.1: b /** lorsqu'aucune décision particulière n'est prise d'établir une base de sondage avant le dénombrement lui-même.

**N.B.** Il faut savoir que l'avantage du choix automatique s'accompagne d'un inconvénient très sérieux, à savoir que l'échantillon est exposé à des distorsions de sélection dont on ignore l'ordre de grandeur.

**A5.8.2/** Un échantillon ne peut être choisi par commodité.

**A5.8.3/** Un échantillon peut théoriquement être choisi avec ou sans remise. Dans la plupart des cas (en pratique), on le choisit sans remise.

### **A5.9. Notations et définitions :**

Nous décrivons un ensemble de résultats, que l'on peut obtenir à partir d'un tableau des nombres positifs, ou tableau de correspondance ; que l'on note  $K_{MS}$ , où  $M$  représente l'ensemble des individus  $m$  (ex. mots en lignes), et  $S$  représente l'ensemble des variables  $s$  (ex. Les sourates en colonnes), le cardinal (effectif total) de l'ensemble  $M$  est noté  $\text{card } M$  ou  $|M|$  (respectivement  $\text{card } S$  ou  $|S|$ ).

|                |                |                |                                    |  |                |       |                  |
|----------------|----------------|----------------|------------------------------------|--|----------------|-------|------------------|
| S<br>M         | s <sub>1</sub> | s <sub>2</sub> | s <sub>3</sub>                     |  | s <sub>α</sub> | ..... | s <sub>114</sub> |
| m <sub>1</sub> |                |                | :                                  |  |                |       |                  |
| m <sub>2</sub> |                |                | k(m <sub>2</sub> ,s <sub>3</sub> ) |  |                |       |                  |
| :              |                |                | :                                  |  |                |       |                  |
| m <sub>t</sub> |                |                | :                                  |  |                |       |                  |
|                |                |                | :                                  |  |                |       |                  |

[Tab A2] K(m<sub>2</sub>,s<sub>3</sub>) représente la croisée de la ligne m<sub>2</sub> et de la colonne s<sub>3</sub>, d'une manière plus particulière, si on désigne par m<sub>2</sub> le mot choisi dans la sourate s<sub>3</sub>, alors k(m<sub>2</sub>,s<sub>3</sub>) représente le deuxième mot (m<sub>2</sub> choisi) de la troisième sourate s<sub>3</sub>.

### A5.9.1. Les marges :

**A5.9.1.a/** la somme des éléments de la ligne m ;  $K(M)=\{k(m) / m \in M\}$  ; où

$$k(m) = \sum\{ k(m, s) / s \in S \}.$$

**A5.9.1.b/** la somme des éléments de la colonne S ;  $K(S)=\{k(s)/s \in S\}$  ;

Où  $k(s) = \sum\{ k(m, s) / m \in M \}.$

**A5.9.1.c/** On note pour  $K = \sum\{ k(m, s) / s \in S, m \in M \}$  le total.

Du tableau de contingence  $K_{SM}$ , on obtient un tableau de fréquence, où

$$f_{MS} = \{f_{ms} / m \in M ; s \in S\}, \text{ avec } f_{ms} = k(m, s) / K$$

$$f_M = \delta^{MS}_M \text{ o } f_{MS} = \{f_m / m \in M\} ; \text{ où } f_m = \sum\{ f(m, s) / s \in S \},$$

$f_m$  est la somme d'une sourate m (fréquence relative marginale de m), avec

$$f_s = \sum\{ f(m, s) / m \in M \}, \text{ est la somme d'une sourate s (fréquence relative}$$

marginale de s), avec  $f_s = k(s) / K ; s \in S$ .

### A5.9.2. Notation de profil :

**A5.9.2 :1 : Profil conditionnel de la sourate S sachant m :**

On note pour  $f^m_s$  le rapport  $[k(m, s) / k(m)]$ , appelé : profil conditionnel de la sourate S sachant m, on écrit :

$f^m_s = \{f^m_s / s \in S\} = k(m, s) / k(m)$ , la fréquence conditionnelle de m quand s parcourt S.

### **A5.9.2 :2 : Profil de l'élément m (mot) sur la sourate S :**

Dans l'espace  $\mathbf{IR}_S$  (espace des mesures sur S), alors on peut écrire :

$f^m_s = \{f^m_s / s \in S\} = \Sigma \{ f(m, s) / f_m / s \in S \} \in \mathcal{P}_S \subset \mathbf{IR}_S$  ; où  $\mathcal{P}_S$  est le simplexe de  $\mathbf{IR}_S$ , correspondre aux lois de probabilités sur S (card S est fini), avec  $\mathbf{IR}_S$  l'espace vectoriel des mesures sur l'ensemble de S .

### **A5.9.2.3: Profil de l'élément m(mot) sur la sourate M:**

Dans l'espace  $\mathbf{IR}_M$  (espace des mesures sur M), alors on peut écrire :

$f^s_M = \{f^s_M / m \in M\} = \Sigma \{ f(m, s) / f_s / m \in M \} \in \mathcal{P}_M \subset \mathbf{IR}_M$  ; où  $\mathcal{P}_M$  est le simplexe de  $\mathbf{IR}_M$  correspondre aux lois de probabilités sur M (card M est fini), avec  $\mathbf{IR}_M$  l'espace vectoriel des mesures sur l'ensemble de M .

Définition : On peut dire pour  $\mathbf{IR}^M$  (respectivement sur  $\mathbf{IR}^S$ ) l'espace des fonctions sur M (respectivement sur  $\mathbf{IR}^S$ ).

### **A5.9.3: Le nuage $\mathcal{N}(M)$ :**

Nous entendons par nuage de M, un ensemble de points de masse, alors à chaque élément  $m \in M$ , on fait correspondre un point de profil  $f^m_s$  et une masse de fréquence  $f_m$ , tel que :  $\mathcal{N}(M) = \{ (f^m_s, f_m) / m \in M \} \subset \mathbf{IR}_S$  .

### **A5.9.4: Le nuage $\mathcal{N}(S)$ :**

Nous entendons par nuage de S, un ensemble de points de masse, alors à chaque élément  $s \in S$ , on fait correspondre un point son profil  $f^s_M$  et une masse sa fréquence est  $f_s$ , tel que :  $\mathcal{N}(S) = \{ (f^s_M, f_s) / s \in S \} \subset \mathbf{IR}_M$  .

### **A 6 : Le centre de gravité de M**

Nous noterons  $g_s$  le centre de gravité des points  $f^m_s$ , affectés chacun de sa masse  $f_m$ , par la moyenne pondérée des  $S^{\text{ième}}$  coordonnées des points du nuage  $\mathcal{N}(M)$  .

$G_s = \Sigma \{ f^m_s * f_m / m \in M \} / \Sigma \{ f_m / m \in M \}$ , et puisque  $\Sigma \{ f_m / m \in M \} = 1$

ce qui fait que  $g_s = \sum \{ f_m^s \cdot f_m / m \in M \}$ .

### A7: Le centre de gravité de S:

Nous noterons  $G_M$  le centre de gravité des points  $f_M^s$ , affectés chacun de sa masse  $f_s$ , par la moyenne pondérée des  $M^{\text{ième}}$  coordonnées des points du nuage  $\mathcal{N}(S)$ .

$$G_m = \sum \{ f_m^s \cdot f_s / s \in S \} / \sum \{ f_s / s \in S \}, \text{ et puisque } \sum \{ f_s / s \in S \} = 1$$

ce qui fait que  $G_m = \sum \{ f_m^s \cdot f_s / s \in S \}$ .

### A8 : Détermination des axes factoriels et les facteurs :

Dans  $IR_S$  (l'espace des mesures), l'axe factoriel de rang  $\alpha$ , de vecteur unitaire  $U_{\alpha S}$ , c'est-à-dire que  $\sum \{ (U_{\alpha S}^2 / f_s) / s \in S \} = 1$ .

On note la fonction de cette mesure  $\varphi_\alpha^S = U_{\alpha S} / f_s$ , cette fonction  $\varphi_\alpha^S$  a pour variance 1 et de moyenne égale à zéro, c'est-à-dire :

$$\sum \{ (\varphi_\alpha^S)^2 \cdot f_s / s \in S \} = 1 \quad ; \quad \sum \{ (\varphi_\alpha^S) \cdot f_s / s \in S \} = 0$$

L'analyse factorielle a mis en place des droites  $D_\alpha$  dans l'espace euclidien où est situé le nuage  $\mathcal{N}(M)$ , ces droites passent par  $f_s$ , profil moyen sur  $S$ , centre de gravité du nuage  $\mathcal{N}(M)$ , alors  $\mathcal{D}_1$  sera l'axe principal d'allongement du nuage,  $\mathcal{D}_2$  est la droite perpendiculaire à  $\mathcal{D}_1$ , celle sur lequel le nuage se projette avec la plus grande dispersion (c.à.d celle le long de laquelle l'inertie du nuage est la plus grande).

D'une manière générale, les axes factorielles successifs, sont orthogonaux deux à deux, qui pour les fonctions  $(\varphi_\alpha^S)$  s'écrit :

$$\delta_{\alpha\beta}^S = \sum \{ (\varphi_\alpha^S \varphi_\beta^S f_s) / s \in S \} = \begin{cases} 0 & \text{si } \alpha \neq \beta \\ 1 & \text{si } \alpha = \beta \end{cases}$$

De cette formule, on dit que les fonctions  $\varphi_\alpha^S$  sont deux à deux non corrélées sur  $S$ , muni de systèmes de masse  $f_s$ .

### A9. Mesures sur les mots (M) et les sourates (S)

#### A9.a/ mesure sur S

Une mesure sur  $S$  (sourates) est donnée par un système de nombres réels qui sont les masses des points de  $S$ , on a :  $\mu_S = \{ \mu_s / s \in S \}$ .

L'ensemble de toutes les mesures sur  $\mathbf{S}$  est un espace vectoriel de dimension  $\text{Card } \mathbf{S}$ , cet espace est noté  $\text{IR}_{\mathbf{S}}$ .

### A9.b/ mesure sur $\mathbf{M}$

Une mesure sur  $\mathbf{M}$  (mots utilisés) est donnée par un système de nombres réels qui sont les masses des points de  $\mathbf{M}$ , on a :  $\mu_{\mathbf{M}} = \{ \mu_m / m \in \mathbf{M} \}$ .

L'ensemble de toutes les mesures sur  $\mathbf{M}$  est un espace vectoriel de dimension  $\text{Card } \mathbf{M}$ , cet espace est noté  $\text{IR}_{\mathbf{M}}$ .

### A10. Les valeurs propres

Si on prend le moment d'inertie de nuage  $\mathcal{N}(\mathbf{S})$  dans la direction de l'axe factoriel de rang  $\alpha$  de la source  $s$  noté  $U_{\alpha s}$ , nous aurons la définition suivante :  $\lambda_{\alpha} = \Sigma \{ F(m)^2 f_m / m \in \mathbf{M} \}$ .

**A10.1** : L'inertie ici au sens de statisticien et non de physicien,

Alors on peut dire (par définition) que  $F(m)$  est une fonction :

**A10.1.1**/ de moyenne nulle, c'est-à-dire :

$$\text{Moy}(F(m)) = \Sigma \{ F(m) * f_m / m \in \mathbf{M} \} = 0.$$

**A10.1.2**/ de variance  $\lambda_{\alpha}$ , C.à.d.:  $\text{var}(F(m)) = \Sigma \{ F(m)^2 f_m / m \in \mathbf{M} \} = \lambda_{\alpha}$

**A10.2**. L'inertie appartient à  $[0, 1]$

On prend la corrélation entre deux fonctions :  $\varphi_{\alpha}^s$ ,  $\varphi_{\alpha}^M$  notée

$$\begin{aligned} \text{Corr}(\varphi_{\alpha}^M, \varphi_{\alpha}^s) &= \Sigma \{ (\varphi_{\alpha}^m - \text{moy}(\varphi_{\alpha}^M)) * ((\varphi_{\alpha}^s - \text{moy}(\varphi_{\alpha}^s)) * f_{ms} / m \in \mathbf{M}, s \in \mathbf{S}) \} / \\ &= \Sigma \{ \varphi_{\alpha}^m (\varphi_{\alpha}^s * f_{ms}) f_m / s \in \mathbf{S}, m \in \mathbf{M}, \} = \Sigma \{ \varphi_{\alpha}^m \sqrt{(\lambda_{\alpha})} \varphi_{\alpha}^m * f_m / m \in \mathbf{M} \} = \\ &= \sqrt{(\lambda_{\alpha})} \Sigma \{ (\varphi_{\alpha}^m)^2 * f_m / m \in \mathbf{M} \} = \sqrt{(\lambda_{\alpha})} \end{aligned}$$

$(\sqrt{(\lambda_{\alpha})})$  : la racine carrée des valeurs propres  $\lambda_{\alpha}$ .

### A11. Formule de transition

On a considéré  $\Sigma \{ \varphi_{\alpha}^s \circ f_{ms} / s \in \mathbf{S} \}$  = formule de transition d'un ensemble à l'autre ensemble, c'est-à-dire les facteurs étant connus sur l'un des ensembles soit  $s$ ).

On peut par des simples calculs linéaires, déterminer leur valeur sur l'autre ensemble  $m$  (on l'appelle formule de transition, parce qu'elle permet le passage de  $S$  à  $M$ ).

$$\text{Alors nous avons } F_\alpha(m) = \sqrt{(\lambda_\alpha)} \varphi_\alpha^m = \varphi_\alpha^s \circ f_s^m.$$

On définit de la manière analogue, un autre facteur sur l'ensemble  $S$ ,  $G(s)$  :

$$\sum \{ G_\alpha(s) f_s / s \in S \} = 0 \quad \& \quad \sum \{ (G_\alpha(s))^2 \cdot f_s / s \in S \} = \lambda_\alpha ;$$

$$\text{Avec } G_\alpha(s) = \sqrt{(\lambda_\alpha)} \varphi_\alpha^s = \varphi_\alpha^m \circ f_m^s.$$

Ces facteurs jouent le rôle le plus important dans la représentation de nuage sur un plan, or ils représentent la projection du nuage  $N(M)$  (respectivement  $N(S)$  sur les axes factoriels).

Alors on dit que  $\varphi_\alpha^M, \varphi_\alpha^S$  sont des facteurs normalisés de variance 1 issus du  $f_m$  tableau des  $K(m,s)$ , les  $F_\alpha, G_\alpha$  sont appelés tout simplement facteurs, et les  $\lambda_\alpha$  sont dits valeurs propres, leur somme qui donne à la fois d'inertie totale du nuage  $\mathcal{N}(M)$  et du nuage  $\mathcal{N}(S)$  est appelé trace.

### A12. Métrique du $\chi^2$ de centre $f_s$ :

On prend une métrique,  $m^{SS}$  métrique sur  $IR_S * IR_S$ , et on définit

$$m^{ss'} = \delta_{s'}^s / f_s \quad \text{où} \quad \delta_{s'}^s = \begin{cases} 0 & \text{si } s \neq s' \\ 1 & \text{si } s = s' \end{cases}$$

On définit le produit scalaire  $V_{\alpha S} \& V_{\beta S}$ , de la manière suivante :

$$\langle V_{\alpha S}, V_{\beta S} \rangle = \sum \{ V_{\alpha S} \cdot V_{\beta S} / f_s ; s \in S \}$$

Donc le carré de la distance

$$d^2(f_s^1, f_s^2) = \langle f_s^1 - f_s^2, f_s^1 - f_s^2 \rangle = \sum \{ (f_s^1 - f_s^2)^2 / f_s ; s \in S \}$$

$$\text{Alors } d^2(f_s^1, f_s^2) = \sum \{ (f_s^1 - f_s^2)^2 / f_s ; s \in S \} = \| f_s^1 - f_s^2 \|_{f_s}^2 = \\ = \sum \{ (f_s^1 - f_s^2)^2 * 1/f_s / s \in S \}$$

De la même façon, on aura pour l'élément  $s$ :

$$d^2(f_M^1, f_M^2) = \| f_M^1 - f_M^2 \|_{f_M}^2 = \sum \{ (f_M^1 - f_M^2)^2 * 1/f_m / m \in M \}$$

**Remarque :**

L'intérêt de cette métrique vérifie une équivalence distributionnelle, c'est-à-dire si deux lignes sont proportionnelles, alors on ne change rien si on remplace deux lignes par une autre ligne, i.e.

$$\forall s \in S : f_{m_1, m_2} = \lambda f_{m_2 s} ; f_{m_1 s} + f_{m_2 s} = f_{m_{tot} s} \quad \forall s \text{ où } m_{tot} = m_1 + m_2$$

et on trouve :  $f_{m_1 s} = f_{m_2 s}$

**A13.les vecteurs propres de la matrice des données**

Soit  $U_{\alpha s}$  le vecteur unitaire correspondant au  $\alpha^{i\grave{e}me}$  axe factoriel, alors

$$\| | U_{\alpha s} | |^2_{fs} = \sum \{ U_{\alpha s} / f_s ; s \in S \} = 1.$$

Il est possible de définir une fonction  $\varphi^S_{\alpha} : S \rightarrow \mathbf{IR}$ , qui donne une valeur réelle à chaque élément de l'ensemble  $S : \varphi^S_{\alpha} = \{ \varphi^S_{\alpha} / s \in S \} = \{ U_{\alpha s} / f_s ; s \in S \}$ .

Nous dirons que  $\varphi^S_{\alpha}$  est le  $\alpha^{i\grave{e}me}$  facteur normalisé issu du tableau  $K_{MS}$

**A14.Formules pour le calcul des contributions**

**A14.a : contribution absolue du mot m à la trace**

Soit un nuage  $\mathcal{N}(M) = \{ (f^m_s, f_m) / m \in M \}$  de centre  $f_s$ , on note

$$d(f^m_s, f_s) = \rho(m) \text{ [rayon polaire ]}, \text{ où } \rho^2(m) = \| | f^m_s - f_s | |^2_{fs}$$

$$\text{et } F(m) = \text{proj}_{U_{\alpha s}} (f^m_s - f_s) = \sum_{\alpha} \langle f^m_s - f_s ; U_{\alpha s} \rangle U_{\alpha s}$$

Alors l'inertie totale du nuage des mots M, ou trace est :

$$\sum_{\alpha} \{ \lambda_{\alpha} / \alpha \in A \} = \sum \{ \rho^2(m) f_m / m \in M \} = \sum \{ \rho^2(s) f_s / s \in S \} =$$

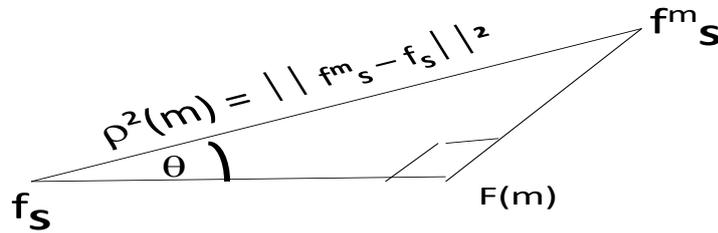
$$= \text{tra} (f^S_M \circ f^M_S) = \text{tra}(f^M_S \circ f^S_M), \text{ où } \mathcal{A} \text{ représente l'ensembles des axes factoriels.}$$

$$\text{Et } \text{tra} (f^M_S \circ f^S_M) = \sum \{ \rho^2_m f_m / m \in M \} = \sum \{ \rho^2_m f_m / m \in M \} =$$

$$= \sum \{ \| | f^m_s - f_s | |^2_{fs} * f_m / m \in M \}$$

Alors  $\rho^2_m f_m$  : appelée contribution absolue du mot m à la trace (Fig :1)

## Contribution absolue



[Fig A :3] [Fig.14.a]

### A14.b : contribution absolue dde la variable S(sourate) à la trace

De la même façon, on représente la contribution de la sourate s à la trace, notée :  $\rho^2(s) f_s$ : appelée absolue de la sourate s à la trace(Fig A3)

### A14.c: contribution relative du facteur F(m) au mot m

Soit un nuage  $N (M) = \{ (f_s^m, f_m) / m \in M \}$  de centre  $f_s$  , on note

$$d(f_s^m, f_s) = \rho(m) \text{ [rayon polaire] , où } \rho^2(m) = || f_s^m - f_s ||^2_{f_s}$$

$$\text{et } F(m) = \text{proj}_{U_{\alpha S}}(f_s^m - f_s) = \sum_{\alpha} \langle f_s^m - f_s ; U_{\alpha S} \rangle U_{\alpha S}$$

le rapport  $F(m)^2 / \rho(m)^2 \dots \dots (1)$  , est appelé contribution relative du facteur  $F_{\alpha}(m)$  au mot : m

et de la même façon par rapport au facteur sur les sourates s :

$$G(s)^2 / \rho(s)^2 \dots \dots (2) , \text{ est appelé contribution relative du facteur } G_{\alpha}(s) \text{ à la sourate s.}$$

### Démonstration

La contribution relative est un  $\cos^2\theta$  Ces relations relatives ne sont rien, Car  $\cos( U_{\alpha S} , (f_s f_s^m)) = F_{\alpha} (m) / \rho(m)$

La contribution relative est un coefficient de corrélation :

$$\text{Corr}( G_{\alpha} , f_s^S ) = \text{Cov} (G_{\alpha} , f_s^S) / ( \sqrt{\text{var } G_{\alpha}} \cdot \sqrt{\text{var } f_s^S} )$$

$$\begin{aligned} \text{Cov}((G_\alpha, f_m^S)) &= \sum \{ f_m [G_\alpha(S) - \text{moy}(G_\alpha(S))] [f_m^S - \text{moy } f_m^S] / s \in \mathbf{S} \} \\ &= f_m \cdot \sqrt{\lambda_\alpha} \sum \{ f_m^s \varphi_\alpha^s / s \in \mathbf{S} \} = f_m \cdot \sqrt{\lambda_\alpha} F_\alpha^m \end{aligned}$$

Mais  $\text{Var } G_\alpha = \sum \{ f_s [G_\alpha(S) - \text{moy}(G_\alpha(S))]^2 / s \in \mathbf{S} \} = \sum \{ f_s [G_\alpha^S]^2 / s \in \mathbf{S} \} = \lambda_\alpha$

Et  $\text{Var } f_m^S = \sum \{ f_s (f_m^S - \text{moy}(f_m^S))^2 / s \in \mathbf{S} \} = \sum \{ f_s (f_m^S - f_m)^2 / s \in \mathbf{S} \} = f_m^2 \rho_m^2$

D'où finalement

$$\text{Corr}^2(G_\alpha, f_m^S) = \text{cov}(G_\alpha, f_m^S) / (\sqrt{\text{var } G_\alpha} * \sqrt{\text{var } f_m^S}) = F_\alpha^2(m) / \rho^2(m)$$

### A15. Les axes factoriels

Soit dans  $\mathbb{R}^M$ , une base n-orthonormée constituée de vecteurs propres  $\varphi^M$  de la matrice (m o s).

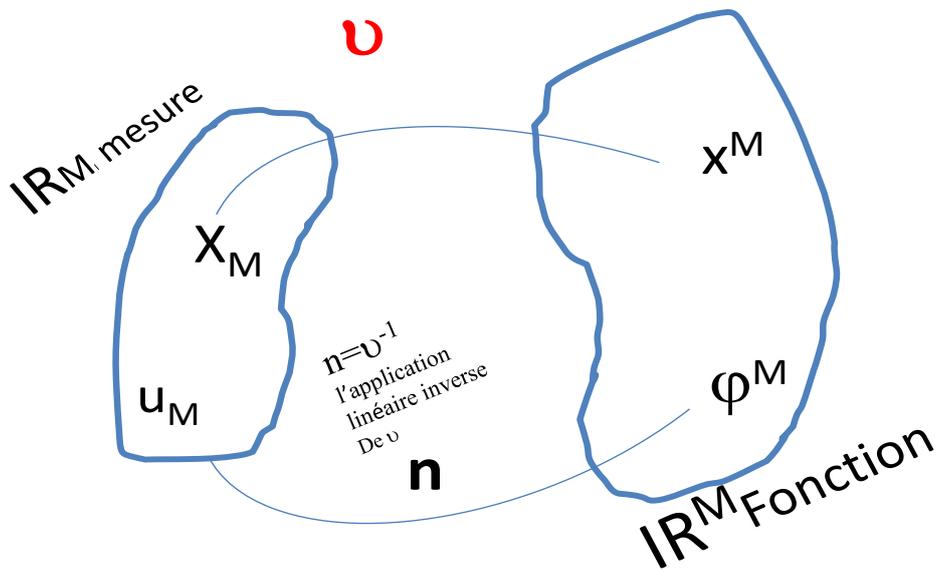
On choisit dans  $\mathbb{R}_M$  une base n-orthonormée, constituée des images de  $\varphi^M$ , notée  $n(\varphi^M)$ .

Alors  $n(\varphi^M)$  est un système d'axes factoriels de nuage  $\mathbf{N}(S)$ , où :

$$n : \varphi^M \in \mathbb{R}^M \xrightarrow{\quad} u_M \in \mathbb{R}_M \quad (u_M : \text{est une mesure})$$

$$\text{i.e.} \quad \varphi_\alpha^M = U_{\alpha M} / f_M$$

$\varphi^M$  est une fonction de densité pour la mesure  $u_M$  par rapport à  $f_M$



[Fig A.15]

On dit que  $\varphi^M \in IR^M$  est un axe factoriel relatif à la matrice  $(m_{os})$  sur  $M$ , si :  
 $(v \circ s) \varphi^M = \lambda \varphi^M$  où  $n = v^{-1}$  l'application linéaire inverse de  $v$ .

$\varphi^M$  est appelé vecteur propre de la matrice  $v \circ s$ , par rapport à la valeur propre  $\lambda$ , où  $v : IR_M \rightarrow IR^M$  où :  $v^{mm'} = \delta_m^{m'} / f_m$

$$\text{i.e. } \sum \{ v^{mm'} s_{mm'} \varphi^{m'} / m \in M, m' \in M \} = \lambda \varphi^m$$

$$\sum \{ \delta_m^{m'} / f_m * (f_m^s - f_m) * (f_{m'}^s - f_{m'}) f_s \varphi^{m'} / m' \in M, s \in S \} = \lambda \varphi^m$$

$$\sum \{ (f_m^s - f_m) * (f_{m'}^s - f_{m'}) f^s \varphi^{m'} / f_m / m' \in M, s \in S \} = \lambda \varphi^m$$

$$\sum \{ (\varphi_{m'} / f_m) * f_s (f_m^s - f_{m'}^s) / m' \in M, s \in S \} - \sum \{ (\varphi_{m'} / f_m) * f_s (f_m^s - f_{m'}^s) / m' \in M, s \in S \} -$$

$$- \sum \{ \varphi^{m'} / f_m \} f_s f_m f_{m'}^s / m' \in M, s \in S \} + \sum \{ \varphi^{m'} / f_m \} f_s f_m f_{m'} / m' \in M, s \in S \} = \lambda \varphi^m$$

$$\text{Finalement : } \sum \{ \varphi^{m'} f_{m'}^s f_{m'}^s / m' \in M, s \in S \} = \lambda \varphi^m$$

$$\text{d'où : } \varphi^M \circ f_{M_s}^S \circ f_{M_s}^M = \lambda \varphi^M$$

De la même façon, on construit l'équation générale dans  $S$  :

$$\varphi^S \circ f_{S_M}^M \circ f_{S_M}^S = \lambda \varphi^S$$

### A16. La métrique du $\chi^2$

Considérons le nuage  $\mathcal{N}(S)$ , la métrique  $\nu$  définie sur  $\mathbb{R}_M$  a l'expression suivante :

$$\nu(f_M^s, f_M^{s'}) = \sum \{ f_m^s f_{m'}^{s'} / f_m ; m \in M \} = \sum \{ \delta_{m'}^m f_{m'}^{s'} f_m^s / f_m ; m, m' \in M \}$$

où  $\delta_M^m$  est une base canonique de  $\mathbb{R}_M$ , i.e.:  $(m \in M) \wedge (m' \in M) : \delta_{m'}^m = \begin{cases} 0 & \text{si } m \neq m' \\ 1 & \text{sinon} \end{cases}$

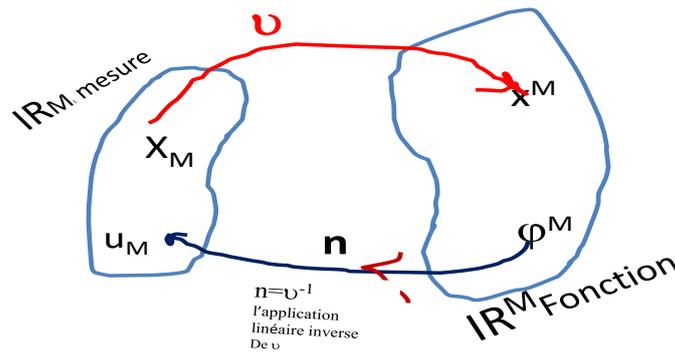
d'où  $\nu(f_M^s, f_M^{s'}) = \sum \{ f_m^s f_{m'}^{s'} \nu_{m m'} ; m \in M, m' \in M \}$

$$\nu_{m m'} = \delta_{m'}^m / f_m$$

Alors :  $\nu_{m m'} = \begin{pmatrix} 1/f_1 & & & & \\ & 1/f_2 & & & \\ & & 1/f_3 & & \\ & & & \vdots & \\ & & & & \vdots \\ & & & & & 1/f_m \end{pmatrix}$

où  $m, m'$  sont deux mots de la matrice de correspondance  $M^*S$  (mots\*sourates) de  $1/f_m$ .

Cette métrique est appelée : la métrique du  $\chi^2$  de centre  $f_M$ , et la distance correspondante est égale à la distance distributionnelle.



[Fig. A 16]

Reste à montrer que  $(n=v^{-1})$  est l'application linéaire inverse de  $v$ .

En effet, soit  $x_I \in \mathbb{R}^I$ , alors  $v(x_I) = x^I \in \mathbb{R}^I \iff x_I = v^{-1}(x^I)$

$$n(x^I, x^I) = (x_I, x_I) = (v^{-1}(x^I), v^{-1}(x^I)) = v^{-1}(x^I, x^I)$$

Donc  $n_{i,i'} = v^{-1}_{i,i'} = \delta^i_{i'} f_i$ .

Alors si  $\varphi^I \in \mathbb{R}^I$  et  $n(\varphi^I) \in \mathbb{R}^I$

$$\varphi_i = \sum \{ n_{i,i'} \varphi^{i'} / i' \in I \} = \sum \{ \delta^i_{i'} f_i \varphi^{i'} / i' \in I \} = \sum \{ \delta^i_{i'} f_i \varphi^{i'} / i' \in I \} = \sum \{ \varphi^{i'} f_i / i' \in I \}.$$

D'où  $\varphi_I = \varphi^I \cdot f_I$

### A17. Détermination des axes factoriels d'inertie et les facteurs

Soit  $\mathcal{N}(M) = \{(f^m_s, f_m) / m \in M\}$  du centre  $f_s$ , il est toujours possible de chercher les axes principaux d'inertie du nuage  $\mathcal{N}(M)$ , la métrique de  $\chi^2$  de centre  $f_s$ , nous appelons axes factoriels des axes principaux d'inertie :  $\cos(U_{\alpha M}, (f^s_M - f_M)) = G(s)/\rho(s)$

#### Propriétés des facteurs $F_\alpha(m)$ :

On peut dire que, La projection n-orthonormé du profil  $f^m_s$  sur l'axe engendré par le vecteur unitaire  $U_{\alpha S}$  est :

$$F_\alpha(m) = \text{proj}_{U_{\alpha S}} f^m_s = \sum \{ \varphi_\alpha^s f^m_s / s \in S \}, \text{ Alors : } F_\alpha^M = \varphi_\alpha^s f^m_s$$

et de la même façon, on aura :  $F_\alpha(s) = \text{proj}_{U_{\alpha M}} f^s_M = \sum \{ \varphi_\alpha^m f^s_m / m \in M \},$

Donc :  $F_\alpha^S = \varphi_\alpha^M \circ f^s_M$ ;  $\text{moy}(F_\alpha^M) = 0$  et de variance  $\text{var}(F_\alpha^M) = \lambda_\alpha$

### A17.a. contribution relative du mot (m) à l'axe factoriel de rang $\alpha$

(CTR), est Représentée par le rapport  $f_m \cdot F(m)^2 / \lambda_\alpha$ .

### A17.b. Covariance de deux vecteurs propres

Soient  $\varphi_\alpha^M$  et  $\varphi_\alpha^S$ , sont deux fonctions centrées et de variance 1

$$\text{Cov}(\varphi_\alpha^M, \varphi_\alpha^S) = \sum \{ f_{ms} * (\varphi_\alpha^M - \text{moy}(\varphi_\alpha^M)) \cdot (\varphi_\alpha^S - \text{moy}(\varphi_\alpha^S)) / m \in M, s \in S \}$$

Etant donnée que les deux fonctions  $\varphi_\alpha^M$  et  $\varphi_\alpha^S$  sont de moyenne nulle, d'où :

$$\begin{aligned} \text{Cov}(\varphi_\alpha^M, \varphi_\alpha^S) &= \sum \{ f_{ms} * (\varphi_\alpha^M - \text{moy}(\varphi_\alpha^M)) \cdot (\varphi_\alpha^S - \text{moy}(\varphi_\alpha^S)) / m \in M, s \in S \} = \\ &= \sum \{ f_{ms} * (\varphi_\alpha^M) \cdot (\varphi_\alpha^S) / m \in M, s \in S \} = \sum \{ (\varphi_\alpha^S f_s \sum \{ f_{ms} \varphi_\alpha^M / m \in M \}, s \in S) \} = \\ &= \sqrt{\lambda_\alpha} \sum \{ f_s (\varphi_\alpha^S)^2 / s \in S \} = \sqrt{\lambda_\alpha}. \end{aligned}$$

C'est-à-dire que les valeurs propres sont comprises entre 0 et 1.

## **Bibliographie**

- 1- Amari S. (1990) - Mathematical foundation of neurocomputing. Proc of the IEEE, 78, n°9
- 2- Art D., Gnanadesikan R, Kettenring J.R. (1982), Data based metrics for cluster analysis. Utilitas Mathematica, 21 A, P 75-99
- 3- ASU, Lebart L., éd.(1992), La qualité de l'information dans les enquêtes. Dunod, Paris.
- 4- Benzécri J.-P. (1969b) - Approximation stochastique dans une algèbre normée non commutative. Bull. Soc. Math. France, 97, p 225-241.
- 5- Benzécri J.-P. (1973), L'Analyse des Données. Tome 1: La Taxinomie. Tome 2: L'Analyse des Correspondances (2de. éd. 1976). Dunod, Paris.
- 6- Benzécri J.-P. (1974) - La place de l'a priori. In: Organum- Encyclopaedia Universalis. Paris.
- 7- Benzécri J.-P. (1977 a) - Analyse discriminante et analyse factorielle. Les Cahiers de l'Analyse des Données, 4, p 369-406.
- 8- Benzécri J.-P. 0977 b) - Choix des unités et des poids dans un tableau en vue d'une analyse des correspondances. Cahiers de l'Analyse des Données, 2, p 333-352.
- 9- Benzécri J.-P. (1979) - Sur le calcul des taux d'inertie dans l'analyse d'un questionnaire. Cahiers de l'Analyse des Données, 4, p 377-378 .
- 10- Benzécri J.-P. (1982 a) - Histoire et préhistoire de l'analyse des données. Dunod, Paris.
- 11- Benzécri J.-P. (1982 b) - Sur la généralisation du tableau de Burt et son analyse par bandes. Cahiers de l'Analyse des Données, 7, p 33-43.
- 12- Benzécri, J.-P. (1982 c) - Construction d'une classification ascendante hiérarchique par la recherche en chaîne de voisins réciproques. Cahiers d'Analyse des Données, 7, p 209-218.

- 13- Benzécri, J.-P. (1983) - Analyse d'inertie intraclasse par l'analyse d'un tableau de correspondance. Les Cahiers d'Analyse des Données, 8, p 351-358.
- 14- Benzécri J.-P. (1992) - Correspondence Analysis Handbook. Marcel Dekker, New York.
- 15- Benzécri J.-P., Cazes P. (1978) - Problème sur la classification. Les Cahiers de l'Analyse des Données, 3,1, P 95-101.
- 16- Benzécri J.-P., Jambu M. (1976) - Agrégation suivant le saut minimum et arbre de longueur minimum. Les Cahiers de l'Analyse des Données, 1, p 441-452.
- 17- Benzécri, J.-P., Lebeaux M.-O., and Jambu M. (1980) - Aides a l'interpretation en classification automatique. Les Cahiers de l'Analyse des Données, 5, p 101-123.
- 18- Bouroche J.-M., Saporta G. (1980) - L'analyse des données. coll."Que sais-je", n01854, PUF, Paris.
- 19- Caillez F., Pagès J.P. (1976) - Introduction à l'Analyse des Données. S.M.A.S.H., Paris.
- 20- Cazes P. (1977) - Etude des propriétés extrêmes des sous-facteurs issus d'un sous-tableau d'un tableau de Burt. Les Cahiers de l'Analyse des Données, 2, p 143-160.
- 21- Cazes P. (1980) - Analyse de certains tableaux rectangulaires décomposés en blocs. Les Cahiers de l'Analyse des Données, 5, p 145-161, et p 387-403.
- 22- Cazes P. (1981) - Note sur les éléments supplémentaires en analyse des correspondances. Les Cahiers de l'Analyse des Données, 1, p 9-23; 2, P 133-154.
- 23- Cazes P. (1982) - Analyse de certains tableaux rectangulaires décomposés en blocs: Codage simultané de variables qualitatives et quantitatives. Les Cahiers de l'Analyse des Données, 6, p 9-18.
- 24- Cazes P. (1984) - Correspondance hiérarchiques et ensembles associés. Cahiers du B.U.R.O., n° 43-44, Université Pierre et Marie Curie, p 43-142.

- 25- Cazes P. (1986 a) - Une généralisation des correspondances multiples et des correspondances hiérarchiques. Cahiers du B.U.R.D., 46-47, Université Pierre et Marie Curie, p 37-64.
- 26- Cazes P. (1986 b) - Correspondance entre deux ensembles et partition de ces deux ensembles. Les Cahiers de l'Analyse des Données, 11, p 335-340.
- 27- Cazes P. (1990) - Codage d'une variable continue en vue de l'analyse des correspondances. Re1JUE Stat/. App/., 38,3, P 35-51. Cazes P., Chessel D., Doledec S. (1988) - L'analyse des correspondances interne d'un tableau partitionné : son usage en hydrobiologie. Revue SIalis/. Appl. 36, (1), p 39- 54
- 28- Diday E. , Lemaire J.L., Pouget J., Testu F. (1982) - Eléments d'Analyse des Données. Dunod, Paris.
- 29- Escofier B. (1979 b) - Traitement simultané de variables qualitatives et quantitatives. Les Cahiers de l'Analyse des Données, 4, (2), P 137-146.
- 30- Escofier B. (1979 c) - Une représentation des variables dans l'analyse des correspondances multiples. Revue de Statis/. Appl. ,27, P 37-47.
- 31- Jambu M. (1991) - Exploration statistique et informatique des données. Dunod, Paris.
- 32- Jambu M., Lebeaux M.O. (1978) - Classification automatique pour l'analyse des données. Dunod, Paris.
- 33- Kharchaf Idris, Rousseau R. 1988, (1989) Reconnaissance de la structure de blocs d'un tableau de correspondance par la classification ascendante hiérarchique: parties 1 et 2, Les Cahiers de l'Analyse des Données, 13, p 439-443; et: 14, p 257-266.
- 34- Kharchaf Idris, « Analyse des votes à l'ONU,1967-1976 », Doctorat, Institut supérieur des universités de Paris VI (ISUP)\*, Laboratoire de l'analyse des données, Paris VI, Juillet 1978 - France.
- 35- Kharchaf Idris, “ Analyse factorielle des correspondances, Analyse de la table pastorale de la Bible”, Colloque international Belgique et Nice 1985, Les Cahiers de l'Analyse des Données, CAH.Vol. XI, 1986, n° 1, p : 23 - France.

- 36- Benzecri,J.P. & KHARCHAF IDRIS : Distribution des noms dans les sourates du Coran »publié dans CAH (Les Cahiers de l'Analyse des données)Vol XI, 1986, n° 1, p : 19 - France.
- 37- KHARCHAF IDRIS, « Coût de la recherche du plus proche voisin », Comptes rendus de l'Académie des sciences, Paris, France (réf. I-432, 9/12/1987).
- 38- KHARCHAF IDRIS « Sur la recherche des plus proches voisins suivant une décomposition cellulaire de l'espace en classification », cahiers de l'analyse des données CAD: Vol. XII, n° 2, 1987.centre national de la recherche scientifique, Paris , France.
- 39-KHARCHAF IDRIS , « Sur la complexité des algorithmes de la classification hiérarchique » CAD. Vol XLI,1987, Paris, France.
- 40-KHARCHAF IDRIS, Etude statistique d'un problème en micropaléontologie » CAD, X<sup>ème</sup> colloque africain de micropaléontologie, Faculté des Sciences 29-30 Septembre 1987, Rabat - Maroc.
- 41- KHARCHAF IDRIS, « Reconnaissance de la structure des blocs d'un tableau de correspondance par la classification hiérarchique », CAD. Vol XIII, 1988, Paris, France.
- 42- KHARCHAF IDRIS, « Stylométrie et théorie des sources ». CAD Vol XII, 1989, Paris - France
- 43- KHARCHAF IDRIS, « Traitement informatique des verbes trilitères hébreux ». Vol. I n° 7, revue RMAITS - 1989, Rabat - Maroc.
- 44- KHARCHAF IDRIS « Etude de la pollution métallique (cu, PI, Zn, Cu) : Témara-Kénitra » (revue chemosphere-Oxford-OX38JD-England, 1989. U.K.
- 45- KHARCHAF IDRIS, « Etude informatique de la stratégie de votes », colloque international de mathématiques appliquées. Ecole Mohammedia des ingénieurs, Rabat 1992 - Maroc.

- 46- KHARCHAF IDRIS : “Pollution and data analysis “The first conference of environmental problems - 14-16 April 1998 - Faculty of Agriculture Mutah University - karak - Jordan, p5.
- 47- K.Lerhain, J.Auray, B. Khbaya, **Idris Kharchaf**, Filali Maltouf, “**phenotypic Characteristics of rhizobia isolates nodulating acacia species in the arid and Saharian regions of Morocco**” [letters in Applied Microbiology, Revue 2000, 351-357] Nederland.
- 48-S.H. Mohamed, A. Smouni, M. Neyra, **I. Kharchaf**, A. Filali Maltouf : “Phenotypic characteristics of root modulating bacteria Acacia Spp. Grown in Libya” Nederland (2000).
- 49-KHARCHAF IDRIS : » l’analyse informatique des sigles, Alif, Lam, Ra “ 3° journées d’optique et du traitement de l’information Cryptologie, optique 2002, Accepté, Ref : TI/17 ,18 Octobre 2002, Facultés des sciences de Kenitra. Maroc
- 50-KHARCHAF IDRIS :« Classification automatique d’une matrice en bloc », Quatrième conférence internationale sur les Mathématiques appliquées et les Sciences de l’Ingénieur, CIMASI 2002 ;Accepté, Réf : MN16 , Casablanca 23-25/10/2002 , Maroc
- 51-KHARCHAF IDRIS, « contribution de l’analyse des données dans le cryptage », article n°65, congrès « The second international conférence of Mathematics », Aleppo University , Faculty of science ,Alep, Syria,26-30/10/2008.
- 52-Lebart L. (1969 b) - Introduction à l'analyse des données: Analyse des correspondances et validité des résultats. Consommation, Dunod. 4, p 65-87.
- 53-Lebart L. (1974) - On the Benzécri's method for finding eigenvectors by stochastic approximation. Proceedings in Comp. Statist., COMPSTAT, Physica verlag, Vienna, p 202-211.

54-Lebart L. (1975 a) - L'orientation du dépouillement de certaines enquêtes par l'analyse des correspondances multiples. *Consommation*, 2, p 73-96. Dunod.

55-Lebart L. (1975 b) - Validité des résultats en analyse des données. Rapport Credoc Cordes. Credoc, Paris.

56-Ludovic Lebart (C.N.R.S. France), Alain Morineau (Centre international de statistique et d'informatique appliquées CISIA), Marie Piron (Institut français de recherche scientifique pour le développement en coopération), « Statistique exploratoire multidimensionnelle », Dunod, Paris, 1995.

57-Nakache J.P. (1973) - Influence du codage des données en analyse factorielle des correspondances. Etude d'un exemple pratique médical. *Revue Sialisi. Appl.*, 21, (2).

58- Sabatier R (1984) - Quelques généralisations de l'analyse en composantes principales de variables instrumentales. *Statist. et Anal. des Données*, 9, (3), P 75-103.

59-Sabatier R (1987) - Analyse factorielle de données structurées et métriques. *Statist. et Anal. des Données*, 12, (3), P 75-96.

60-Saporta G. (1975 a) - Liaisons entre plusieurs ensembles de variables et codages de données qualitatives. Thèse 3<sup>o</sup>C., Université Paris V.

---

\*) ISUP : L'Institut de Statistique de l'Université de Paris (ISUP) est une formation pionnière dans le domaine de la statistique. Depuis sa création en 1922 par le mathématicien Emile Borel, l'institut a su s'adapter aux nombreuses évolutions de cette discipline. Aujourd'hui, les formations de l'ISUP, tournées vers l'analyse et l'exploitation des données, en font un acteur majeur de la formation aux technologies de la science des données et du Big Data.

Les liens étroits entre l'ISUP et les entreprises construisent un modèle pédagogique axé sur les innovations et en phase avec les besoins du secteur professionnel. Le Diplôme de Statisticien de l'ISUP, avec ses deux filières (actuariat, ingénierie statistique et data science), couvre un spectre varié d'applications et favorise l'interdisciplinarité des parcours, permettant à nos étudiants de s'insérer au cœur des problématiques contemporaines. La filière actuariat de l'ISUP est à ce titre reconnue par l'Institut des Actuaire comme l'une des huit formations initiales permettant à ses diplômés d'accéder au statut d'Actuaire Associé.

En tant que département de formation de Sorbonne Université (5<sup>ème</sup> université en mathématiques au niveau mondial, classement de Shanghai 2015), l'ISUP bénéficie d'un corps enseignant d'une qualité exceptionnelle, auquel s'adjoint un nombre important d'intervenants issus de grandes entreprises du secteur privé. Les liens étroits qui le

relient notamment au Laboratoire de Probabilités Statistique et Modélisation apportent à nos étudiants un lien constant avec une recherche scientifique de haut niveau.



La recherche dont s'inspire le présent travail a pour objectif principal, la rationalisation de la raison, et la compréhension objective du Coran, de manière nouvelle et à l'aide de moyens scientifiques, afin que le lecteur non averti puisse comprendre la structure du Coran et son contenu, quelles que soient sa religion et

ses affiliations idéologiques .

Pour cette raison, on a fait recours aux techniques de Data Mining (Analyse factorielle des correspondances, Classification hiérarchique, Analyse en composantes principales), qui doit précisément être le facteur principal et scientifique qui clarifie le sens objectif du Coran, afin de ne pas projeter sur le Coran la subjectivité du lecteur.

Ceux qui sont étrangères à l'esprit mathématiques s'étonneront d'une telle approche originale et féconde, à cause du langage mathématique avec lequel ils ne sont pas familiers. Ainsi, ils peuvent prendre l'attitude dogmatique du refus arbitraire et gratuit.

Prof. Kharchaf Idris