

إستخدام التحليل العنقودي لتصنيف مجموعة السور القرآنية الكريمة المفتحة بالثناء على الله عز وجل

(PP 1 - 12)

أ.م.د. كوردستان ابراهيم مولود

قسم الاحصاء- كلية الادارة والاقتصاد جامعة صلاح الدين- أربيل

kurdistan.mawlood@su.edu.krd

الاستلام : 2016/07/18**القبول : 2016/09/08****النشر : 2018/03/26****ملخص**

إن الهدف الأساس من التحليل العنقودي هو تصنيف مجموعة من عناصر البيانات الى عناقيد ، بحيث تكون العناصر داخل العنقود أكثر تشابهاً مع بعضها البعض من العناصر في العناقيد الأخرى. تم في هذا البحث تطبيق التحليل العنقودي على مجموعة السور التي افتتحهم الله في كتابه الكريم بالثناء عليه عز وجل وعددها أربعة عشرة سورة وذلك بالاعتماد على تكرارات الحروف في كل سورة كمتغيرات الدراسة. تم استخدام طريقة الربط التام و اختيار مربع المسافة الاقليدية لإيجاد مصفوفة المسافة التي أستخدمت لإيجاد العناقيد وربط العناقيد واحدة تلو الأخرى باستخدام برنامج SPSS فضلاً عن شكل الهيكل الشجري والذي يمثل المحور السيني فيه المفردات (السور) بينما يمثل المحور الصادي المسافات بين المفردات. تم تشكيل (4) عناقيد مرتبة من حيث ترتيب السور وعدد آياتها وعلاقتهم فيما بينهم بحيث يوجد تجانس بين مفردات لكل عنقود، العنقود الأول يتضمن سورة واحدة فقط وهي: (سورة الأنعام) فقط والعنقود الثاني يتضمن سورتي (الأسراء و الكهف)، والعنقود الثالث يحتوي على خمس سور وهي (الفرقان، فاطر، سبأ، الحشر، الحديد) ، والعنقود الرابع يحتوي على ست سور المتبقية وهي (الملك، الجمعة، التغابن، الصف، الأعلى، الفاتحة).

1-1 المقدمة

اهتم المسلمون على مر العصور في علوم القرآن الكريم بكافة أشكالها، في مجال العد والأرقام والإحصاء فقد قام المسلمون بإحصاء حروف القرآن وكلماته وسوره والتوافق العددي وتناسقه في الأسلوب القرآني وقد قال الرافي رحمة الله " القرآن الكريم يعطيك معان غير محدودة في كلمات محدودة ((قُلْ لَوْ كَانَ الْبَحْرُ مِدَاداً لِكَلِمَاتِ رَبِّي لَنَفِدَ الْبَحْرُ قَبْلَ أَنْ تَنفَدَ كَلِمَاتُ رَبِّي وَلَوْ جِئْنَا بِمِثْلِهِ مَدَدًا)) (سورة الكهف الآية 110) وهذا البحر الفانض من الخيرات لا بد لاستخراج الدرر المكنونة فيه من غوص وتدبر واستخراج الحلول للمشاكل المستجدة في عصرنا وفي كل العصور وهو ما يسمى بصلاحية الإسلام لكل زمان ومكان. و يقول تعالى ((وكل شيء أحصيناه كتاباً)) (سورة النبأ الآية 29) فالله الحكيم المقدر المحصي الحسيب، أورد الحروف في القرآن بقدر، وأورد الكلمات في القرآن بقدر، وهناك تناسق و توازن في عدد مرات ورود كلمات وألفاظ القرآن الكريم. [9]

يقول تعالى: ((الله الذي انزل الكتاب بالحق والميزان)) (سورة الشورى الآية 17) والآية الأخيرة تكشف عن الدقة المطلقة للنظام القرآني الذي يضم حقائق الكون كله، لذا فإن أية دراسة جادة لآيات وكلمات وحروف وأرقام أعداد القرآن الكريم، لابد وأن تعطيناً ثماراً يانعة في مجال التناسق والتناظر والتكرار والتناغم العددي. و أن كلمات القرآن ربّتها الله بنظام رقمي معجز ليؤكد لنا أننا إذا تدبرنا هذا القرآن سوف نكتشف أنه كتاب محكم، وأننا سوف نجد البراهين الثابتة على أنه لو كان هذا القرآن قول بشرٍ لرأينا فيه التناقضات والاختلافات.

افتتح الله السور المباركة في كتابه المجيد بعشره أنواع من الكلام، وأحد أنواعها الأستفتاح بالثناء عليه عز وجل وجاء ذلك في أربعة عشر سور وهو قسمان:

1- أثبات لصفات المدح نحو ((الحمد لله)) في خمس سور هي (الفاتحة، الأنعام، الكهف، سبأ، فاطر) و ((تبارك)) في سورتين هما (الفرقان، الملك).

2- تنزيه (الله عز وجل) من صفات النقص، وجاء ذلك في سبع سور (الإسراء، الحديد، الحشر، الصف، الجمعة، التغابن، الأعلى).

والتحليل العنقودي هو تحليل احصائي و أحد طرق تحليل متعدد متغيرات لاكتشاف فيما إذا كان مفردات من البيانات تقع ضمن مجموعات معينة و متجانسة، ويعد من الأساليب الإحصائية المهمة التي يمكن استخدامها في دراسة الظواهر المعقدة بصورة أكثر



تشخيصاً وفي العلوم ومجالات الحياة المختلفة. الهدف من هذا البحث هو تطبيق التحليل العنقودي على أربعة عشر سور القرآنية الكريمة التي تبدأ بالثناء على الله عز وجل لدراسة التجانس و وأوجه الإرتباط بينهم.

1-2 التحليل العنقودي Cluster Analysis

التحليل العنقودي عبارة عن تحليل احصائي يبحث في فصل البيانات داخل مجاميع باستناد الى المتغيرات او المشاهدات وهو احد اساليب المتعدد المتغيرات وهو اداة استكشافية لحل مشاكل التصنيف مثل تصنيف المجموعة من المفردات او المتغيرات الى المجمعتين او اكثر تكون درجة التجانس قوية بين اعضاء المجموعة الواحدة في حين تكون الدرجة التجانس ضعيفة بين اعضاء المجاميع المختلفة.

ويستخدم مصطلح التصنيف classification ايضاً لغرض وصف الأساليب التي تبحث في تجميع المفردات individual أو العناصر objects تحت الدراسة في مجاميع متجانسة فيما بينهما (داخل المجموعة الواحدة) ومختلفة عن المجاميع الأخرى بالاعتماد على صفات (متغيرات) عدة، أي انه يمكن استخدام مصطلحي التحليل العنقودي و التصنيف بصورة متبادلة.[3]

2-2 مجالات تطبيق التحليل العنقودي

يطبق التحليل العنقودي على نحو واسع في شتى المجالات وفي مختلف الميادين منها ميدان علم نفس Psychology (تصنيف المفردات الى أنواع الشخصيات) والتسويق marketing (التصنيف الزبائن الى مجموعات على اساس عوامل أذواق المستهلكين والمنتج المستخدم) وعلم الكيمياء chemistry (تصنيف المكونات الكيميائية على لساس خصائصها) وعلم الأرض geology (تصنيف طبقات الأرض) وعلم الاجتماع sociology وعلم الفلك astronomy والطب medicine (التصنيف الأمراض) وعلم الحياة biology والصناعة industry والأعمال business والجغرافيا geography والزراعة agriculture وعلم الاقتصاد economics والتعليم education وعلم الوراثة genetics وعلم اللغة linguistics (تصنيفات اللغات).

3-2 أهمية التحليل العنقودي

- 1- اكتشاف علم شرح الرموز finding a true typology
- 2- مطابقة النماذج fitting models
- 3- التنبؤ المبني على أساس المجاميع prediction based on group
- 4- اختبار الفرضيات testing hypotheses
- 5- استكشاف البيانات data exploration
- 6- توليد الفرضيات hypotheses
- 7- اختزال البيانات data reduction

4-2 العنقود Cluster

يعرف العنقود بأنه مجموعة من العناصر المتجاورة في المجتمع الاحصائي، كمجموعة من الناس يعيشون في مسكن واحد، أو مجموعة من قطع الأراضي المتجاورة في الحقل. كما يعرف العنقود على انه مجموعة من الوحدات المتشابهة والتي تختلف عن الوحدات الموجودة في العناقيد الأخرى. وعليه فالعنقود يمكن ان يوصف كمناطق متلاصقة Continuous Regions من الفضاء تحوي على كثافة عالية نسبيا من المفردات وتفصل هذه المناطق عن بعضها البعض مناطق اقل كثافة، والعنقود الموصوف على هذه الشاكلة يسمى بالعنقود الطبيعي Natural Cluster [3].

5-2 التصنيف Classification

التصنيف هو عبارة عن وضع الأشياء المتشابهة في الخصائص أو في العلاقات بعضها مع بعض في فئات معينة، أي ترتيب المفردات بناءً على ما بينها من تشابه أو اختلاف، أي انه عبارة عن فصل المفردات أو تجميعها. ويعرف التصنيف العددي بأنه عبارة عن مجاميع مصنفة بطريقة عددية بحيث ان كل مجموعة تحوي على عدد من العناصر ذات الصفات المتشابهة[11].

6-2 المسافة Distance

يمكن تعريف المسافة بأنه الفضاء الذي يفصل بين أي عنصرين، وبما انه يمكن لكل من التشابه والمسافة (الاختلاف) ان يحل أحدهما محل الآخر (ولكن بصورة عكسية) فعملية يمكن الاعتماد على المسافة بدلاً من قياس التشابه في إجراءات العنقدة clustering.



ان المسافة بين عنصرين تقاس بالفترة ما بينهما , وان المسافات ليست دائما تقاس بالقوانين ,ومقاييس المسافة ذات صلة بمقاييس التشابه والاختلاف إذ ان التشابه بين عنصرين يقاس بدرجة قرب كل عنصر من آخر, والاختلاف أو المسافة هو عكس التشابه ويكمل أحدهما الآخر وكما يأتي[5]:

Dissimilarity (Distance) =1-Similarity

7-2 الهيكل الشجري (dendrogram)

يعد الهيكل الشجري أحد أكثر المظاهر المهمة للتحليل العنقودي, وهو عبارة عن شجرة بيانية تلخص عمليات عنقدة (التجميع) العناصر الى مجاميع طبيعية , يربط الحالات المتشابهة مع بعضهما البعض في الرسم البياني بتحديد التشابه بين الحالات. وتعرف الشجرة بأنها مجموعة من العناقيد المركبة كالتركيب الهرمي(المتسلسل) لعمليات العنقدة التي تنفذ في مصفوفة المسافة والتي يمكن تمثيلها بالرسم البياني ذي البعدين (Two-Dimensional) حيث يمثل محور في X فية الأهداف حين يمثل محور Y المسافات , بحيث تكون فروع الشجرة في اسفل وجذعها في أعلى (من المناسب طبع شجرة في الكومبيوتر على الجانب الأيمن والفروع على الجانب الأيسر). ان الهيكل الشجري يوضع اختلافات المجاميع للعناصر ,والتي تملك المسافات الصغيرة فيما بينها تتمثل في الفروع البعيدة من الشجرة , أما التي تملك مسافات كبيرة فتتمثل في الفروع القريبة من جذع الشجرة , وعندما تكون المسافات الفروع البعيدة كبيرة أي قريبة من الفروع القريبة من جذع الشجرة فان عملية التجميع تكون غير مؤثرة أو غير فعالة جداً. والهيكل الشجري له فائدة في استكشاف القيم شاذة (Outliers) في حالة التوزيعات المختلطة للبيانات المتعددة المتغيرات (Mixtrue Distribution of multivariate data), أي العناصر الشاذة التي تشترك أو ترتبط مع مجموعة في الفروع القريبة من جذع الشجرة (أي العناصر التي ترتبط مع المجموعة الرئيسة في آخر خطوة من عملية العنقدة). [1][11]

8-2 أساليب وطرق التحليل العنقودي Clustering Techniques & Methods

ان عملية العنقدة (Clustering) أي وضع العناصر أو المتغيرات في المجاميع (العناقيد) وفقا لدرجة التشابه بينهما، بحيث تكون درجة التجانس قوية داخل العنقود الواحد، وضعيفة بين العناقيد المختلفة، يتم وصف شكل العلاقات بين العناصر أو المتغيرات اعتمادا على مصفوفة المسافة - بأساليب عدة، منها [5] [2] :

1. أسلوب التسلسل الهرمي (Hierarchical Technique)
2. أسلوب التجزئة (Partitioning Technique)

وتتضمن أساليب التحليل العنقودي ثلاث مراحل هي :-

1. مرحلة التجزئة (The Partitioning Stage) خلال هذه المرحلة تتخذ قرارات حول كيفية قياس البيانات، وما هي افضل الخطوات (Algorithms) المناسبة لتصنيف البيانات، وعدد العناقيد المكونة.
 2. مرحلة التفسير (The Interpreting stage) في هذه المرحلة تختبر اسم الكشوفات (القوائم) المستخدمة لكشف العناقيد والتي تصف بدقة صفات (Characteristics) العنقود.
 3. مرحلة التصوير المقطعي (The Profiling Stage) يتم وصف صفات كل عنقود في هذه المرحلة، والتي توضح كيفية اختلاف كل عنقود عن الأبعاد المناسبة، وتستمر عملية التحليل باستخدام البيانات غير المصنفة (المشخصة) سابقا في العناقيد لوصف كل صفات العنقود. وفي هذه المرحلة تحصل اندماجات (Focuses) لأوصاف العنقود بعد تشخيصها، وليست لما تحدده تراكيب العناقيد. أن نقطة الاختلاف بين الطرق العنقودية هي الطريقة التي يتم فيها حساب المسافة بين العنقودين عند ربطهما، ومن أهم هذه الطرق هي:
1. طريقة الربط المفرد (المجاور الأقرب) (single linkage method(nearest neighbor))



2. طريقة الربط التام (الكامل) Complete Linkage (Furthest Neighbor) Method
 3. طريقة المعدل غير الموزون للمجموعتين Unweight pair-Groups Method Average
 4. طريقة المعدل الموزون للمجموعتين Weighted Pair-Groups Method Average
 5. الطريقة المتمركزة Centroid Method:
 6. الطريقة طريقة وارد Ward's Method
 7. الطريقة العنقودية ك من الأوساط K - Means Clustering Method
 8. طريقة لانس ووليام المرنة Lance & William's Flexible Method
 9. الطريقة الوسيطية Median Method
 10. طريقة العنقدة داخل المجاميع Within- Groups Clustering Method
- وقد تم استخدام طريقة الربط التام في هذا البحث .

9-2 طريقة الربط التام (الكامل) Complete Linkage (Furthest Neighbor) Method

وتعرف أيضا الطريقة العظمى (Maximum Method) و هذه الطريقة مشابهة لطريقة الربط المفرد (Single Linkage) باستثناء طريقة تحديد المسافة للربط بين العنقودين عند كل مرحلة من مراحل عملية العنقدة، ففي هذه الطريقة و بعد تحديد عتبة المسافة (Threshold Distance) عند بداية كل مرحلة يتم تشكيل العنقود بطريقة معاكسة للطريقة الأولى (Single Linkage) حيث انها تعتمد على ابعاد مسافة أو أكثرها (الأقل تشابها) بين كل زوج من العناصر أو أعضاء العناقيد المختلفة عند كل مرحلة، أي ان العنصر (العنقود) المرشح لتشكيل العنقود الجديد هو الذي يملك اكبر مسافة بينه وبين عنصر من عناصر العنقود الآخر، ففي هذه الحالة يتم ربط مجموعة العناقيد مع بعضها البعض اعتمادا على ابعاد مسافة بين أعضاء العناقيد (I, J) وحسب الصيغة الآتية [7] [4]:

$$D(I, J) = \max\{D_{ij} : i \in I, j \in J\} \quad (2-1)$$

$$i = 1, 2, \dots, n_I, \quad j = 1, 2, \dots, n_J$$

حيث ان:

(i, j) : يمثلان ابعاد عنصرين في العنقودين (I, J) على التوالي.

وكما في طريقة (Single Linkage) فان المسافة بين ابعاد عنصر للعنقود (T) مثلا وبين ابعاد عنصر في العنقود $(I \cup J)$ (المصاغ بوساطة العنقود (I) و العنقود (J)) باستخدام العلاقة التكرارية للمسافة تحسب بالصيغة الآتية :

$$(I \cup J, T) = \max[D_{it}, D_{jt}] \quad (2-2)$$

$$i \in I, j \in J, t \in T$$

حيث أن t : يمثل ابعاد عنصر في العنقود (T) .

وتستمر عملية ربط المجاميع.

10-2 المعايير العنقودية Clustering Criteria

تعتمد عدد إجراءات التجزئة على أساس تصغير التباين داخل المجاميع (Minimizing Within- Group Dispersion) وتعظيم التباين بين المجاميع (Maximizing Between - Groups Dispersion)، لذلك فان جميع المعايير العنقودية المستخدمة في طرق العنقدة بأسلوب التجزئة مشتقة من معادلة المصفوفة الأصلية (Fundamental Matrix Equation) الآتية [9] :-

$$T=W+B \quad \dots(3-2)$$

$$W = \sum_{i=1}^K W_i$$

حيث أن :

T : تمثل مصفوفة التباين الكلي.



W : تمثل مصفوفة التباين داخل المجاميع.

k: يمثل عدد العناقيد (المجاميع).

W_i : تمثل مصفوفة التباين للعنقود (i).

B : تمثل مصفوفة التباين بين المجاميع.

تكون مصفوفة (T) ثابتة لأية مجموعة بيانات، وفإن معيار العنقودية يكون من التحكم بالمصفوفتين (B) و (W).

ان افضل معيار للتجميع أو للتصنيف الأمثل يكون عندما تكون مصفوفة (W) اقل ما يمكن [Min(W)] وتكون مصفوفة (B) أكبر ما يمكن [Max(B)].

المعيار الأول : تقليل مجموع قطر المصفوفة (W) [min trace(W)]

ان عملية نقل الوحدات بين العناقيد تبقى مستمرة الى أن يكون جميع المفردات في المجموعة قريبة من مركز العنقود الذي تعود إليه اكثر من قربها الى المركز في أي عنقود آخر، وهذا بدوره يؤدي الى تقليل مجموع قطر المصفوفة [min trace(W)] ، وان هذا التقليل يكافؤه تكبير مجموع قطر المصفوفة (B) [max trace (B)] وذلك استنادا الى الصيغة الآتية:-

$$\min \text{trace}(W) \Leftrightarrow \max \text{trace}(B)$$

ويستنتج هذا من معادلة المصفوفة الأصلية الآتية [6] [8]:

$$\text{trace}(T) = \text{trace}(W) + \text{trace}(B)$$

المعيار الثاني : تقليل قيمة محدد المصفوفة (W) [min|W|]

اقترح هذا المعيار لأول مرة معيارا ملائم للعنقودية من قبل [Friedman & Rubin 1967]

$$\min |W| = \max [|T| / |W|]$$

المعيار الثالث : تعظيم مجموع قطر المصفوفة (B^{-1}) [max trace (B⁻¹)]

اقترح هذا المعيار أيضا من قبل [Friedman & Rubin, 1967] إن مجموعي قطري المصفوفتين كليهما (B^{-1}) و $|T| / |W|$ يعبر عنه بالقيم المميزة

(Eigenvalues) (λ_i) للمصفوفة (B^{-1}) كما يأتي :

$$\text{trace}(WB^{-1}) = \sum_{i=1}^p \lambda_i$$

$$|T| / |W| = \prod_{i=1}^p (1 + \lambda_i)$$

حيث يمثل (λ_i) الجذر المميز (القيمة المميزة Eigenvalue) للمعادلة الآتية :-

$$|B - \lambda W| = 0$$

2-3 وصف البيانات

تم استخدام برنامج (إحصاء القرآن الكريم) لاستخراج بيانات هذا البحث والتي تمثل تكرارات كل الحروف (من ألاف الى الياء) في مجموعة السور التي افتتحهم الله تعالى بالثناء عليه عز وجل في القرآن الكريم. الجدول رقم (1-3) يمثل القيمة الأدنى و الأعلى والوسط الحسابي و الانحراف المعياري لكل الحروف الهجائية العربية للسور التي افتتحهم الله تعالى بالثناء عليه عز وجل في القرآن الكريم.

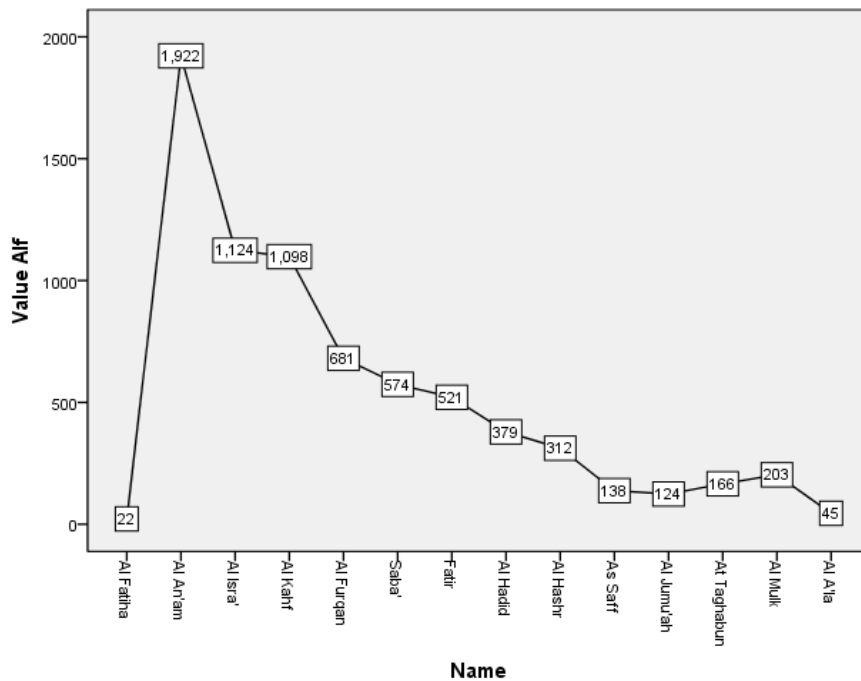
الجدول (1-3) الوسط الحسابي والانحراف المعياري لكل الحروف لمجموعة السور التي فواتحها بالثناء على الله عز وجل

| Descriptive Statistics | | | | |
|------------------------|---------|---------|--------|----------------|
| | Minimum | Maximum | Mean | Std. Deviation |
| Alf | 22 | 1922 | 522.07 | 537.702 |
| Baa | 4 | 453 | 118.50 | 128.442 |
| Taa | 3 | 404 | 98.43 | 110.570 |
| Thaa | 0 | 58 | 14.71 | 17.054 |
| Jim | 0 | 117 | 36.21 | 37.134 |



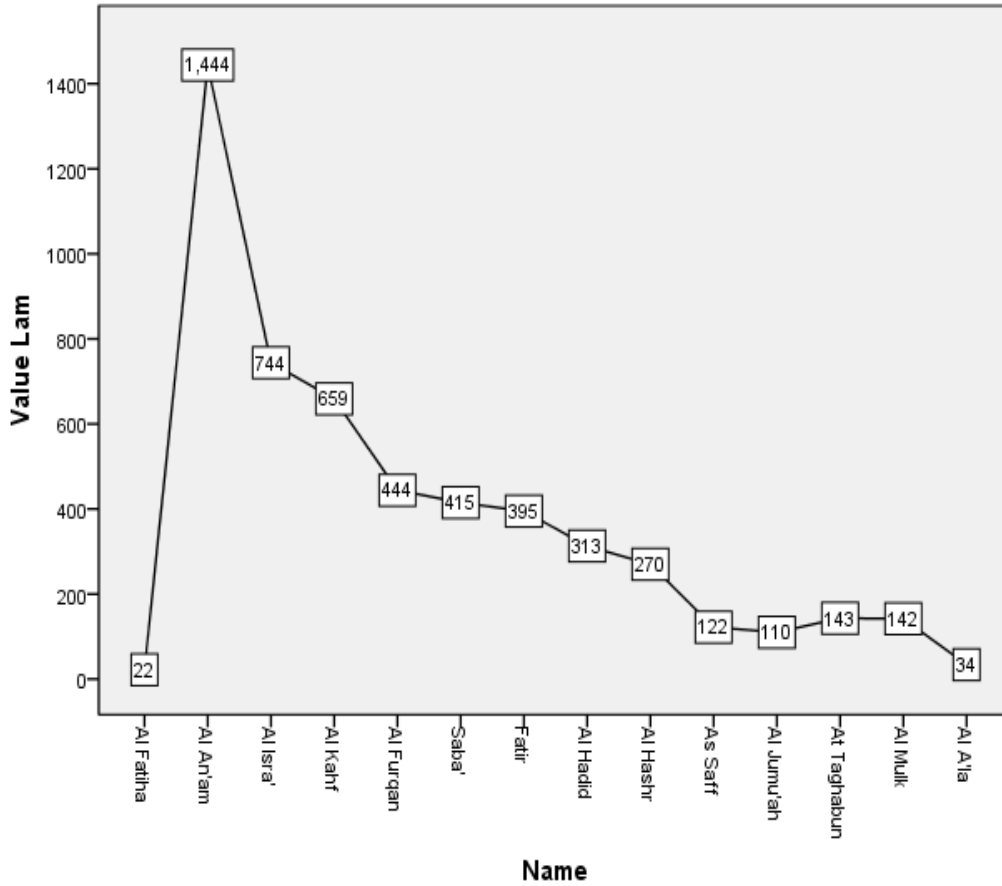
| | | | | |
|--------|----|------|--------|---------|
| Haa7 | 5 | 160 | 42.14 | 45.149 |
| Khaa | 0 | 75 | 23.36 | 23.408 |
| Dal | 4 | 199 | 61.00 | 64.497 |
| Thal | 1 | 207 | 54.29 | 55.054 |
| Raa | 8 | 504 | 140.93 | 139.848 |
| Zaa | 0 | 66 | 19.21 | 18.431 |
| Sin | 3 | 212 | 61.14 | 61.217 |
| Shin | 0 | 132 | 25.50 | 34.022 |
| Sad | 2 | 69 | 19.36 | 18.316 |
| Thad | 0 | 66 | 18.714 | 19.0038 |
| Taaa | 0 | 52 | 12.43 | 15.936 |
| Thaaaa | 0 | 48 | 9.21 | 12.367 |
| Ain | 5 | 369 | 91.14 | 102.439 |
| Ghain | 1 | 50 | 13.43 | 14.173 |
| Faa | 0 | 294 | 82.71 | 80.746 |
| Qaf | 1 | 271 | 70.36 | 76.414 |
| Kaf | 3 | 468 | 107.79 | 125.750 |
| Lam | 22 | 1444 | 375.50 | 378.706 |
| Mim | 12 | 1058 | 250.29 | 274.564 |
| Non | 11 | 1010 | 249.93 | 274.550 |
| Haa | 5 | 679 | 164.07 | 179.542 |
| Waw | 4 | 1023 | 252.29 | 266.719 |
| Yaa | 14 | 1008 | 255.64 | 263.794 |

یظهر بوضوح في الجدول أعلاه أن حرف الألف هو الأكثر تكراراً في مجموعة السور الكريمة التي فواتحها بالثناء على الله عز وجل في كتاب الله ويليه حرف اللام ويليه حرف الياء و ثم حرف الواو، بينما يعتبر حرف الظاد "ظ" هو الأقل تكراراً في كتاب الله، كما نلاحظ أن قيم الإنحراف المعياري لجميع الحروف أكبر من قيم الوسط الحسابي وذلك لوجود قيم الشاذة في البيانات أي وجود تفاوت كبير بين تكرار الحروف في هذه المجموعة من السور القرآنية الكريمة.



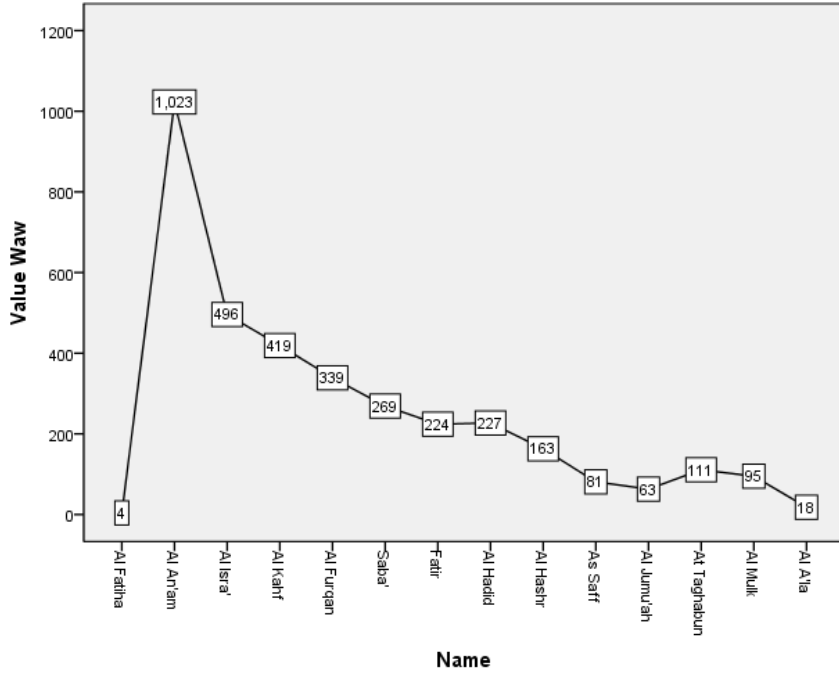
الشکل رقم (1-3)

شکل توضیحی لتکرار حرف ألاف في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم يبين الشكل أعلاه ابتداء هذه المجموعة من سور القرآن الكريم بسورة قصيرة وسرعان ما يزداد طول السور في ثاني سورة ويبدأ بالتناقص تدريجياً حتى تختتم بسورة الأعلى مع ملاحظة تذبذب المنحني في بعض الأحيان صعوداً ونزولاً. علماً أن السورة ذات أكبر تكرار لحرف (ألاف) في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم هي سورة الأنعام حيث تحوي (1922) حرفاً وتليها سورة الاسراء و ثم سورة الكهف بينما أقل تكرار حرف ألاف في سورة الفاتحة حيث تحوي (22) حرفاً فقط.



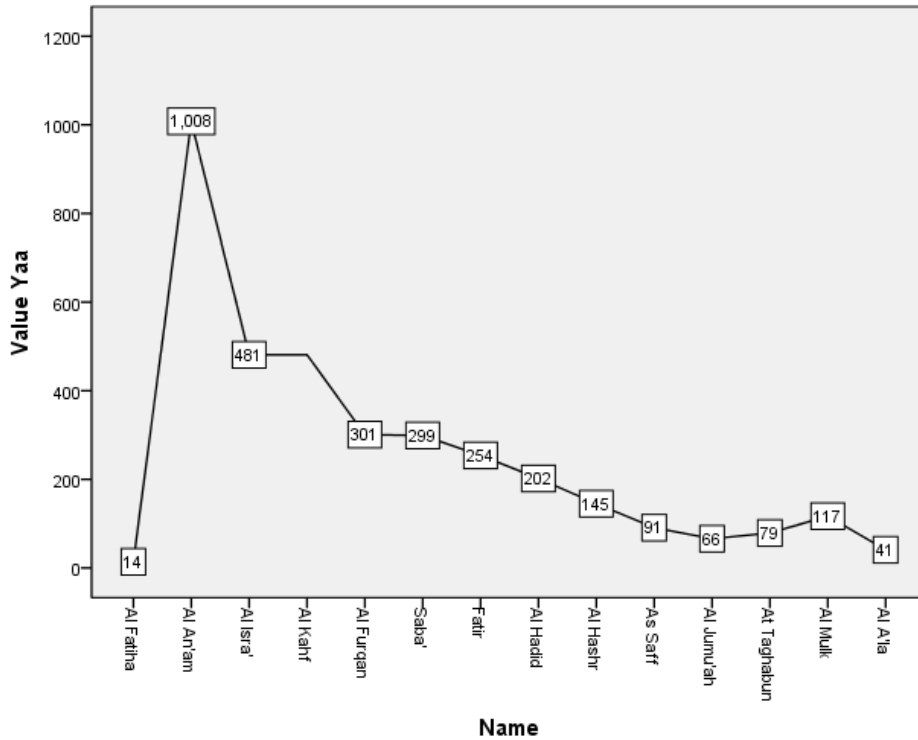
الشکل رقم (2-3)

شکل توضیحی لتکرار حرف اللام في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم يبين الشكل (2-3) أعلاه ابتداء هذه المجموعة من سور القرآن الكريم بسورة قصيرة وسرعان ما يزداد طول السور في ثاني سورة ويبدأ بالتناقص تدريجياً حتى تختتم بسورة الأعلى مع ملاحظة صعود المنحني في سورة التغابن و ثم يبدأ بالتناقص مرة أخرى. علماً أن السورة ذات أكبر تكرار لحرف (اللام) في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم هي سورة الأنعام حيث تحوي (1444) حرفاً وتليها سورة الاسراء و ثم سورة الكهف بينما أقل تكرار حرف اللام في سورة الفاتحة حيث تحوي (22) حرفاً فقط.



الشكل رقم (3-3)

شكل توضيحي لتكرار حرف الواو في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم يبين الشكل (3-3) أعلاه ابتداء هذه المجموعة من سور القرآن الكريم بسورة قصيرة وسرعان ما يزداد طول السور في ثاني سورة ويبدأ بالتناقص تدريجياً حتى تختتم بسورة الأعلى مع ملاحظة صعود المنحنى في سورة التغابن و ثم يبدأ بالتناقص مرة أخرى . علماً أن السورة ذات أكبر تكرار لحرف (الواو) في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم هي سورة الأنعام حيث تحوي (1023) حرفاً وتليها سورة الاسراء و ثم سورة الكهف بينما أقل تكرار حرف الواو في سورة الفاتحة حيث تحوي (4) حرفاً فقط.



الشكل رقم (4-3)

شكل توضيحي لتكرار حرف الياء في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم



یبن الشکل أعلاه ابتداء هذه المجموعة من سور القرآن الكريم بسورة قصيرة وسرعان ما يزداد طول السور في ثاني سورة ويبدأ بالتناقص تدريجياً حتى تختتم بسورة الأعلى مع ملاحظة تذبذب المنحني في بعض الأحيان صعوداً ونزولاً. علماً أن السورة ذات أكبر تكرار لحرف (الياء) في مجموعة السور التي فواتحها بالثناء على الله عز وجل في القرآن الكريم هي سورة الأنعام حيث تحوي (1008) حرفاً وتليها سورة الاسراء وثمر سورة الكهف بينما أقل تكرار حرف الياء في سورة الفاتحة حيث تحوي (14) حرفاً فقط.

3-3 تطبيق التحليل العنقودي

تم تطبيق التحليل العنقودي طريقة الربط التام واختيار مربع المسافة الاقليدية لإيجاد مصفوفة المسافة التي سوف تستخدم في التحليل العنقودي وربط العناقيد واحدة تلو أخرى وتم الحصول على الهيكل الشجري (Dendrogram) الموضح في الشكل رقم (3-5) من البيانات ومن مصفوفة المسافة وباستخدام مقياس مربع المسافة الاقليدية، حيث يمثل المحور العمودي (Y) في الهيكل مربع المسافة بين المفردات المدمجة (السور) والمحور الأفقي (X) المفردات المندمجة.

الجدول (2-3) جدول التقارب

| Stage | Cluster Combined | | Coefficients | Stage Cluster First Appears | | Next Stage |
|-------|------------------|-----------|--------------|-----------------------------|-----------|------------|
| | Cluster 1 | Cluster 2 | | Cluster 1 | Cluster 2 | |
| 1 | 1 | 14 | 2238 | 0 | 0 | 9 |
| 2 | 10 | 12 | 3314 | 0 | 0 | 3 |
| 3 | 10 | 11 | 5714 | 2 | 0 | 4 |
| 4 | 10 | 13 | 12969 | 3 | 0 | 9 |
| 5 | 6 | 7 | 18619 | 0 | 0 | 8 |
| 6 | 8 | 9 | 25232 | 0 | 0 | 10 |
| 7 | 3 | 4 | 36075 | 0 | 0 | 12 |
| 8 | 5 | 6 | 39186 | 0 | 5 | 10 |
| 9 | 1 | 10 | 55245 | 1 | 4 | 11 |
| 10 | 5 | 8 | 150982 | 8 | 6 | 11 |
| 11 | 1 | 5 | 448038 | 9 | 10 | 12 |
| 12 | 1 | 3 | 2E+06 | 11 | 7 | 13 |
| 13 | 1 | 2 | 8E+06 | 12 | 0 | 0 |

وبملاحظة نتائج التحليل العنقودي والموضحة بالجدول (2-2) المسمى بجدول التقارب والذي يبين كيفية تكوين العناقيد في كل مرحلة من مراحل التحليل نجد أن المرحلة الأولى (Stage 1) تم فيها ربط الحالة (السورة) رقم (1) والتي تمثل سورة الفاتحة بالحالة رقم (14) والتي تمثل سورة الأعلى وان مربع المسافة الاقليدية بينهما (2238) وفي المراحل الثانية والثالثة والرابعة تم ربط الحالة (10) والتي تمثل سورة الصف بالحالات رقم (12, 11, 13) والتي تمثلان السور (التغابن، الجمعة، الملك) على التوالي، وفي المرحلة رقم (9) وتم ربط الحالة رقم (1) أو سورة الفاتحة مرة ثانية ولكن مع الحالة (10) أو مع سورة الصف ليتشكل لدينا عنقود يتكون من ست سور وهي (الفاتحة، الأعلى، الصف، التغابن، الجمعة، الملك)، وهكذا يمكن أن نفسر بقية المراحل ومتابعة كيفية تكوين بقية العناقيد.

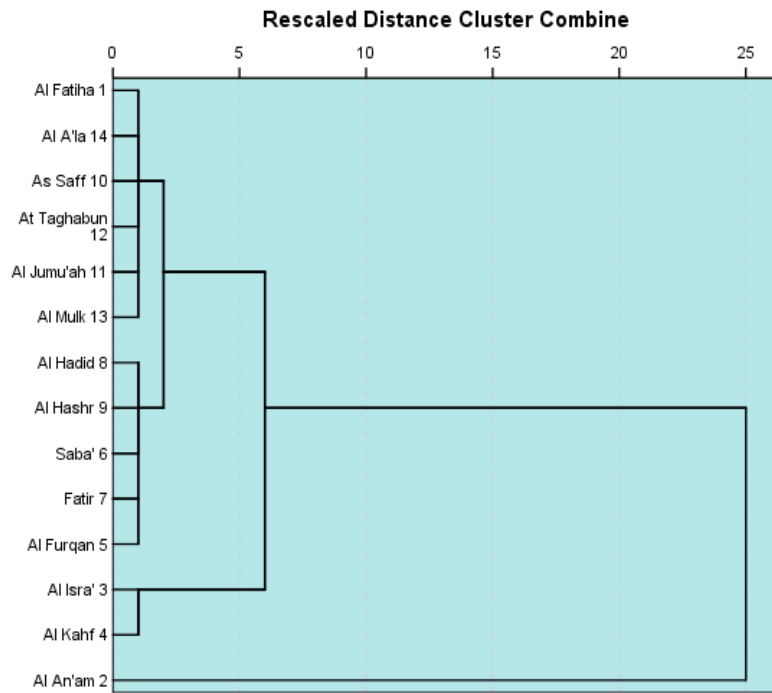
الجدول (3-3) عضوية العناقيد

| Case | 4 Clusters |
|-------------|------------|
| 1:Al Fatiha | 4 |
| 2:Al An'am | 1 |
| 3:Al Isra' | 2 |
| 4:Al Kahf | 2 |



| | |
|----------------|---|
| 5:Al Furqan | 3 |
| 6:Saba' | 3 |
| 7:Fatir | 3 |
| 8:Al Hadid | 3 |
| 9:Al Hashr | 3 |
| 10:As Saff | 4 |
| 11:Al Jumu'ah | 4 |
| 12:At Taghabun | 4 |
| 13:Al Mulk | 4 |
| 14:Al A'la | 4 |

ويلحظ من الجدول (3-3) أعلاه أن عدد العناقيد المستخدمة في التحليل هو (4) عناقيد ، فالعنقود الأول يحتوي على سورة واحدة فقط وهي: سورة الأنعام ، والعنقود الثاني يحتوي على سورتين وهما : سورة الأسراء و سورة الكهف ، والعنقود الثالث يحتوي على خمس سور وهي (الفرقان، فاطر، سبأ، الحشر، الحديد) ، والعنقود الرابع يحتوي على ست سور وهي (الملك، الجمعة، التغابن، الصف، الأعلى، الفاتحة)، وهذه العناقيد الأربعة تشكل بنسب كل منها على التوالي (% 7.14 ، % 14.29 ، % 35.71 ، % 42.86) من المفردات أو السور القرآنية الكريمة ، ومجموع السور المستخدمة في هذا التحليل يساوي (14) سورة.



الشكل (3-5)

الهيكل الشجري (Dendrogram)

نلاحظ من شكل الهيكل الشجري ان هناك اربع عناقيد مرتبة حسب تسلسل السور في المصحف الشريف وحسب (عدد) ترتيب اياتهم ترتيباً تنازلياً .

- العنقود الاول مكون من سورة واحدة فقط وهي سورة الانعام وهي أطول سورة في هذه المجموعة (المجموعة التي فواتحها بالثناء على الله عز وجل) وعدد آياتها يساوي (165) آية كريمة ، وانها أطول سورة مكيه حسب ترتيب الصحف الشريف ومايميز بها هذه السورة هي انها انزلت دفعة واحدة وكان انزالها ليلاً وحولها سبعون ألف ملك يجأرون حولها بالتسبيح (حسب قول ابن عباس) والميزة الثانية لهذه السورة عن بقية السور ان سورة الانعام من أفاضل سور القرآن لأن موضوعها هو العقيدة وفيها كل قواعد التوحيد حسب روايه عمر بن الخطاب (رضى الله عنه) انه قال سورة الانعام من نجائب القرآن وفي روايه ثانيه نواجب القرآن والنجب هو اللب. [12]



- العنقود الثاني مكون من سورتي (الاسراء و الكهف) و من أبرز أوجه التجانس بين هذين السورتين الكريمتين هي ان عدد آياتها هي قريبة جداً سورة الاسراء عدد آياتها هي (111) ايه و عدد آيات سورة الكهف (110) ايه وتسلسلها 17 و 18 على التوالي و كلاهما سور مكيه، ونزلت سورة الاسراء قبل سورة الكهف وفي تأويل انه يجوز ان يكون نزول سورة الاسراء مستمراً الى وقت نزول سورة الكهف.

- العنقود الثالث يتضمن خمس سور كريمة وهي (الفرقان، فاطر، سباء، الحشر و الحديد) والتي عدد آياتهم اقل من عدد آيات السور التي تشكل العنقودين الاول و الثاني و أيضاً ترتيبها يأتي بعد ترتيب العنقودين الاول و الثاني حسب ترتيب المصحف الشريف .

- العنقود الرابع مكون من ست سور كريمة و هي (الفاتحة، الاعلى، الصف، التغابن، الجمعة، الملك) حيث يقع جميعهم في جزئين الاخرين من القرآن الكريم عدا سورة الفاتحة وان عدد آياتهم متقاربة و اقل من عدد آيات السور لبقية العناقيد. ومن النتائج المذكورة يتبين لنا أن المفردات المتكونة لكل عنقود تكون ذات ارتباط كبير (أكثر تشابهاً) فيما بينها ، مما يدل على أن هناك تجانساً أو تقارباً كبيراً بين مفردات العنقود الواحد .

1-4 الإستنتاجات

يمكن تلخيص أهم الإستنتاجات التي توصلت اليها الدراسة فيما يلي:

1- نلاحظ وجود تفاوت كبير بين تكرار الحروف في مجموعة سور القرآن الكريم المفتحة بالثناء على الله عز وجل، وأن حرف الألف هو الأكثر تكراراً ويليه حرف اللام ويليه حرف الياء و ثم حرف الواو، بينما يعتبر حرف الظاد هو الأقل تكراراً في هذه المجموعة.
2- نلاحظ من منحي تكرارات للحروف الأربعة (الألف، اللام، الواو والياء) بأن أكبر تكرار في سورة الأنعام و تليها سورة الاسراء و ثم سورة الكهف، وأن أقل تكرار للحروف الأربعة في سورة الفاتحة و كأن موضع سجدة المنحني لجميع الحروف يكون في سورة الفاتحة والتي قراءتها ركن من أركان الصلاة بخلاف غيرها من السورة لقول النبي صلى الله عليه وسلم (لا صلاة لمن لم يقرأ بفاتحة الكتاب).

3- يتبين لنا من الشكل الهيكل الشجري ان هناك اربع عناقيد مرتبة حسب تسلسل السور في المصحف الشريف وحسب (عدد) ترتيب آياتهم ترتيباً تنازلياً . وأن المفردات المتكونة لكل عنقود تكون ذات ارتباط كبير (أكثر تشابهاً) فيما بينها ، مما يدل على أن هناك تجانساً أو تقارباً كبيراً بين مفردات العنقود الواحد . فالعنقود الأول يحتوي على سورة واحدة فقط وهي: (سورة الأنعام) ، والعنقود الثاني يحتوي على سورتين وهما : (سورة الأسراء و سورة الكهف) ، والعنقود الثالث يحتوي على خمس سور وهي (الفرقان، فاطر، سبأ، الحشر، الحديد) ، والعنقود الرابع يحتوي على ست سور وهي (الملك، الجمعة، التغابن، الصف، الأعلى، الفاتحة)، وهذه العناقيد الأربعة تشكل بنسب كل منها على التوالي (% 7.14 ، % 14.29 ، % 35.71 ، % 42.86) من المفردات أو السور القرآنية الكريمة المستخدمة في هذا التحليل و يساوي (14) سورة.

4- إن معجزة القرآن الكريم معنوية، وأن التوافق والتناسق العددي في القرآن الكريم نوع من أنواع الإعجاز البياني الذي أجمع عليه العلماء وأن التحدي بالإعجاز البياني مستمر الى يوم القيامة.



المصادر

بعد القرآن الكريم

1. Everitt, B.S., Landau, S. and Leese, M. (2001), (**Cluster Analysis, Fourth edition**), Arnold.
2. Manly, B.F.J. (2005), (**Multivariate Statistical Methods**), A primer, Third edition, Chapman and Hall.
3. Rencher, A.C. (2002), (**Methods of Multivariate Analysis**), Second edition, Wiley.
4. Mardia, K.V., Kent, T.J., Bibby, J. M (1979), (**Multivariate analysis**), Academic press, London .
5. Aziz and Zickar (2006), (**A cluster analysis investigation of work holism as a syndrome**), Journal of Occupational Health Psychology, 11, 52-62.
6. Richard A. Johnson & dean W. Wichern, (2007), (**Applied multivariate Statistics**), pearson Education, Inc.
7. Joseph F. Hair, William C. Black, Barry J. Babin and Rolph E. (2007), (**Anderson, Multivariate Data Analysis A Global Perspective**), Seventh edition, pearson.
8. John I. Marden, (2015), (**Multivariate Statistics**), Madsen and Wilson.
9. السيوطي، الإمام جلال الدين، (1996)، (الإتقان في علوم القرآن)، تعليق د. مصطفى ديب البغا، دار ابن كثير . دمشق.
10. كريم، زيان إحسان، (1992)، (استخدام التحليل العنقودي لتقييم الأراضى للأغراض الزراعية)، رسالة ماجستير، كلية الإدارة والاقتصاد، جامعة صلاح الدين.
11. عمر، جيمس أبوبكر، (2003)، (استخدام المركبات الاساسية والتحليل العنقودي لتحديد اهم العوامل المؤثرة في اصابات جيوب اللثة) (دراسة ميدانية)، رسالة ماجستير، كلية الادارة والاقتصاد، جامعة صلاح الدين.
12. أبو الفداء اسماعيل بن عمر بن كثير القريشي الدمشقي، (1991)، (تفسير القرآن العظيم)، دار الكتب العلمية بيروت.

Using cluster analysis for classification Quranic suar's which begin with Praising God Almighty

Abstract

The main objective of the cluster analysis is the classification of a set of data elements into clusters, so that elements within the cluster are more similar to each other than elements in other clusters. In this research the cluster analysis has been applied on the Quranic suar's, which begin with Praising God Almighty which are fourteen Surat depending on occurrences of letters in each Surat separately as study variables. the statistical package (SPSS) was used to find a cluster analysis; complete linkage method and choose the square Euclidean distance to find a distance matrix used in the cluster analysis and linking clusters. Four clusters was combined arranged in terms of the sequence of the fence in the Quran, according to (the number of) the order of their verses numbers in descending order. And there is a homogeneous or great convergence between the elements of one cluster. The first cluster contains only one Surat, namely: (Al An'am), and the second cluster contains two suar's namely: (Al-Isra and Al Kahf), the third cluster contains five suar's, namely, (Al-Furqan, Fater, Saba, al-Hashr, Al Hadid), and the fourth cluster contains six suar's, namely, (Al Mulk, Al Jumu'ah, At Taghabun, As Saff, Al A'la, Al Fatiha).