

استخدام التحليل العنقودي الهرمي في تصنيف المشاهدات إلى مجاميع متجانسة مع تطبيق على دوري كرة السلة

م.م . ذنون يونس ذنون الشكرجي
المعهد التقني / نينوى

تاريخ تسليم البحث : 2007/5/29 ؛ تاريخ قبول النشر : 2008/2/14

ملخص البحث :

يقسم التحليل العنقودي إلى طريقتين هما : طريقة التكتل الهرمي (Hierarchical Agglomerative Method) , وطريقة التجزئة (Partitioning Method) وفي هذا البحث سيتم استخدام طريقة الربط التام (Complete Linkage Method) التي تُعدُّ من أهم طرق التكتل الهرمي وأكثرها شيوعاً وبالتطبيق على بيانات أخذت من كلية التربية الرياضية في جامعة الموصل , هذه البيانات مكونة من مجموعتين احتوت كل مجموعة على 48 لاعباً يمثلون أندية الدوري الممتاز لكرة السلة , أما بالنسبة لتفسير نتائج التحليل العنقودي ونظراً لكثرة المشاهدات فقد تم تفسير نتائج خمس مشاهدات من مجموع ثمانٍ وأربعين مشاهدة لكل مجموعة .

Using Hierarchical Cluster Analysis to classify the observations into homogenous groups with application on Basketball Matches

Assist . Lect
Thanoon Younis Thanoon Alshakerchy
Technical Institute /Nainavah

Abstract:

Cluster analysis is divided into two methods (Hierarchical agglomerative method) and (partitioning method) and in this research it is used (complete linkage method) which is considered as one of the important methods of hierarchical agglomerative method and most common applied to data taken from the college of sport education in Mosul university ,these data consist of two groups each group contains forty eight players representing clubs of excellent tournament The interpretation of cluster analysis results and regarding to much observations, I make an interpretation of five observations from forty eight for each group .

1- المقدمة : (كريم , 1992)

المصطلح الأكثر شيوعاً للطرق التي تبحث في فصل البيانات داخل مكونات المجاميع هو: التحليل العنقودي , وقد استعمل مصطلح: التحليل العنقودي من لدن مجموعة من الباحثين كأسلوب يبحث في تجميع المتغيرات , وقد تم اقتراح هذا الأسلوب كبديل عن التحليل العاملي (Factor Analysis) وتحليل المكونات الرئيسية (Principal Component Analysis). ولعل ابرز من استخدم هذه التسمية : (Kendall & Sturat) اللذان استخدموا أيضاً التصنيف (Classification) لغرض وصف الأساليب التي تستعمل لتجميع العناصر (objects) أو المفردات (individuals) تحت الدراسة إلى مجاميع متجانسة فيما بينها (داخل المجموعة الواحدة) ومختلفة عن المجاميع الأخرى بالاعتماد على العديد من الصفات أو (المتغيرات) . ويمكن القول انه بالامكان استخدام مصطلحي التحليل العنقودي والتصنيف بصورة متبادلة لغرض وصف الأساليب التي تبحث في تجميع المفردات المتجانسة إلى مجاميع محددة.

2- هدف البحث

- 1- تصنيف مجموعة من لاعبي كرة السلة إلى مجاميع متجانسة وباستخدام التحليل العنقودي.
- 2- مقارنة نتائج كل من المجموعتين بالاعتماد على شكل الهيكل الشجري .

3- طرق العنقدة Clustering Methods (عمر , 2003)

- تتطلب عملية العنقودية وضع العناصر (objects) في مجاميع (groups) بطريقتين :
1. بتكوين مجاميع اكبر نتيجة لدمج العناصر أو (المجاميع الصغيرة) .
 2. بتكوين مجاميع صغيرة نتيجة لتجزئة المجاميع الكبيرة ابتداءً من المجموعة المحتوية على جميع العناصر .

لذلك فان التحليل العنقودي يتكون من طريقتين هما :

1. طريقة التكتل الهرمي Hierarchical Agglomerative Method

2. طريقة التجزئة Partitioning Method

كل من الطريقتين السابقتين مكونة من عدة طرق لذلك تم استخدام طريقة الربط التام (Complete Linkage Method) في الجانب التطبيقي التي تعتبر من أهم طرق التكتل الهرمي وأكثرها استخداماً لذلك سيتم توضيح هذه الطريقة بشي من التفصيل .

4- خطوات إجراء التحليل العنقودي : (البرزنجي , 2002)

إن إجراءات الإدماج الكلي ابتداءً من دمج أول عنصرين إلى انتهاء العملية العنقودية تتضمن الخطوات الآتية :

1. حساب مصفوفة المسافة (Distance Matrix) ذات بعد $(n \times n)$:-

$$d = |d_{ij}|$$

من مجاميع العناصر الأولية فإن d_{ij} تمثل المسافة بين العنصرين (j,i) وتتم في هذه الدراسة عملية حساب المسافة بالاعتماد على مقياس المسافة المسمى بمربع المسافة الاقليدية (Square Euclidean Distance) .

2. يتم البحث عن ابسط المسافات داخل المصفوفة السابقة حيث يتم ربط العنصرين اللذين تكون المسافة بينهما اقصر المسافات ضمن المصفوفة ليكون الاختيار الأفضل أخذ أول عنصرين أو آخر عنصرين لكي يشكلوا نواة العنقود .

3. بعد أن يتم تشكيل العناقيد الأولية تأتي مرحلة حساب مصفوفة المسافة الجديدة التي تأخذ بنظر الاعتبار التغيرات التي حصلت في المرحلة الثانية .

4. الاستمرار بعملية الربط اعتماداً على اقصر المسافات والى أن يتم ربط آخر عنقودين .

طريقة الربط التام (المجاور الأبعد)

Complete Linkage Method (Furthest Neighbor)

في هذه الطريقة يشكل العنقود على أساس الاعتماد على اقل تشابه بين العناصر , ويمكن توضيح الطريقة كالآتي :

(Mardia,Kent & Biby,1979) (Sas Institute Inc.,1985)

1. يبدأ مع العناقيد (C_1, C_2, \dots, C_n) المحتوية على (X_1, X_2, \dots, X_n) على التوالي .
2. بافتراض $d_{12} = \min (d_{ij})$ لكل (j,i) وبافتراض (C_2, C_3, \dots, C_n) تكون المجموعات بعد اتحاد الزوج $(2,1)$ لتشكيل C_2 .
3. لتحديد مصفوفة المسافة الجديدة $[(n-1), (n-1)]$:

$$D^* = (d_{il}^*)$$

$$d_{2j} = \max (d_{1j}, d_{2j})$$

مع $i, j = 3, 4, \dots, n$

بعد ذلك يتبع ما في (2) أعلاه .

وتستمر العملية إلى أن تكون جميع المسافات بين العناقيد أعظم من d_0 حيث تمثل

d_0 الحد الاعتباطي .

ويمكن ملاحظة انه عندما تكون العنقودية تامة يكون الحد الأعلى
 $d_{ij} \leq d_0, \max i, j \in c$ لكل عنقود C .

وهكذا فان هذه الطريقة تتجه إلى إنتاج عناقيد مدمجة مع عدم تغير النتيجة , ويمكن
 حساب هذه الطريقة باستعمال العلاقة التكرارية للمسافة بين المجموعة k والمجموعة (j,i)
 المصاغة بوساطة المجموعة I إلى:

$$d_{k(ij)} = \max (d_{ik}, d_{jk})$$

طريقة الربط المفرد (المجاور الأقرب)

Linkage Method Single (Nearest Neighbor)

تُعدُّ هذه الطريقة من ابسط طرق العنقودية ومن أقدمها حيث يمكن استعمالها مع كل
 مقاييس التشابه ومقاييس المسافات . (Mardia, Kent & Biby, 1979)

المجموعات الأولية المؤلفة من مفردات وحيدة تدمج وفقاً للمسافة بين عناصرها الأكثر
 قرباً , فالمجموعات التي تكون مسافتها اقرب من غيرها تدمج أولاً , كل اندماج يقلل عدد
 المجموعات مجموعة واحدة . وفي هذه الطريقة ترتب أولاً $\frac{1}{2} n(n-1)$ للمسافات بين النقاط
 ترتيباً تصاعدياً .

أ. بافتراض c_1, \dots, c_n تكون بداية العناقيد التي يحتوي كل منها على نقطة واحدة أي

$$c_1 \in X_1$$

ب. مع عدم فقدان العمومية تكون $d_{r1s1} = \min_{rs}$

وهكذا تمثل X_{s1} و X_{r1} العناقيد الأقرب .

ومن ثم فان هاتين النقطتين جمعت في العنقود , كما إن هنالك العناقيد $n-1$, حيث إن
 $C_{r1} + C_{s1}$ تمثل عنقوداً جديداً .

ج. بافتراض d_{r2s2} تمثل المسافة التالية الأصغر . إذا لم تكن r_1 ولا s_1 مساوية لـ r_2 أو s_2
 فالعنقود $n-2$ الجديد يكون $C_{r1} + C_{s1}$, $C_{r2} + C_{s2}$ زائداً بقايا العناقيد القديمة , أما إذا كانت
 $r_1 = r_2$ و $s_1 \neq s_2$ فالعناقيد الجديدة $n-2$ تكون $C_{r1} + C_{s1} + C_{s2}$ زائداً بقية العناقيد القديمة .

د. وهكذا تستمر العمليات كما في (ج) أعلاه من خلال كل مسافات $\frac{1}{2} n(n-1)$. في مرحلة
 (i) بافتراض d_{r1s1} تشير إلى مسافة (i) الأصغر , ومن ثم فالعنقود الذي يحتوي r_1 يتصل
 مع العنقود الذي يحتوي على s_1 . (إذا كانت r_1 و s_1 عادة في نفس العنقود فلا يمكن تشكيل
 أية مجموعة جديدة في هذه المرحلة).

هـ. يمكن إيقاف عملية العنقدة قبل أن تتصل جميع العناقيد في مجموعة واحدة عن طريق
 التوقف عندما تكون المسافات بين العنقود كلها اكبر من d_0 حيث تكون d_0 مرتبة الحد
 الأعلى الاعباطي (Arbitrary Threshold Level) .

الهيكل الشجري Dendrogram

نعني بالهيكل الشجري رسم بياني شجري لعملية عنقدة المجاميع المعطاة حيث يمثل محور X فيه الأهداف ، بينما يمثل محور Y المسافات ، وفروع الشجرة تمثل الترتيب لـ (n-1) من الحلقات ، فالتفرع الأول يمثل الحلقة الأولى والثاني يمثل الحلقة الثانية وهكذا حتى ترتبط جميعها مع بعضها في الجذع ، وهناك طريقتان للوصول إلى ذلك: (Everitt , 1977)

1. الطريقة الأولى: تعتمد على التكتل (Agglomerative) والتي تتكون من سلسلة من الخطوات التي يتم ربط العناصر مع بعضها في أية مرحلة منها (اعتماداً على معامل التشابه أو المسافة) ومن ثم الحصول على العنقود الكبير الذي يحوي جميع العناصر .
2. الطريقة الثانية: والتي تعتمد على التقسيم (Partitioning) والذي يتم بموجبه فصل المجاميع عن بعضها على شكل مراحل إلى أن يتم فصل آخر مجموعة تتكون من عنصرين .

وفي كلتا الحالتين تكون النتيجة الحصول على هيكل شجري .

التحليل الإحصائي

استخدمت البرامج الإحصائية (Spss V11.5) و (Minitab V13.20) لإيجاد التحليل العنقودي ثم طريقة الربط التام (Complete Linkage) واختيار مربع المسافة الاقليدية (Square Euclidean Distance) لإيجاد مصفوفة المسافة التي سوف تستخدم في التحليل العنقودي وربط العناقيد واحدة تلو أخرى فضلاً عن شكل الهيكل الشجري والذي يمثل المحور X فيه المشاهدات بينما يمثل المحور Y المسافات بين المشاهدات .

جمع البيانات

تم جمع البيانات من كلية التربية الرياضية في جامعة الموصل حيث أخذت عينتان كل منهما مكون من (48) لاعباً يمثلون أندية الدوري/ الفرق الثمانية الأوائل للدوري الممتاز بكرة السلة والبالغ عددهم (12) نادياً لعام (2001-2002) حيث تمثل المجموعة الأولى الفرق الأربعة الأولى التي ترشحت إلى المربع الذهبي ، والمجموعة الثانية تمثل الفرق الأربعة خارج المربع الذهبي ، والجدول (1) يوضح الأندية وعدد لاعبيها (الطائي ، 2004)

الجدول (1)

يوضح الأندية وعدد لاعبيها

عدد اللاعبين	اسم النادي (خارج المربع الذهبي)	ت	عدد اللاعبين	اسم النادي (داخل المربع الذهبي)	ت
12	الشرطة	5	12	الجيش	1
12	الحدود	6	12	الكرخ	2
12	بابل	7	12	الدفاع الجوي	3
12	الأرمني	8	12	دهوك	4
48			48		

علماء إن الأندية التي لم تترشح لدوري الفرق الثمانية الأوائل هي : (الأعظمية , القعقاع , الناصرية , الميناء) .

وصف المتغيرات الداخلة في التحليل :

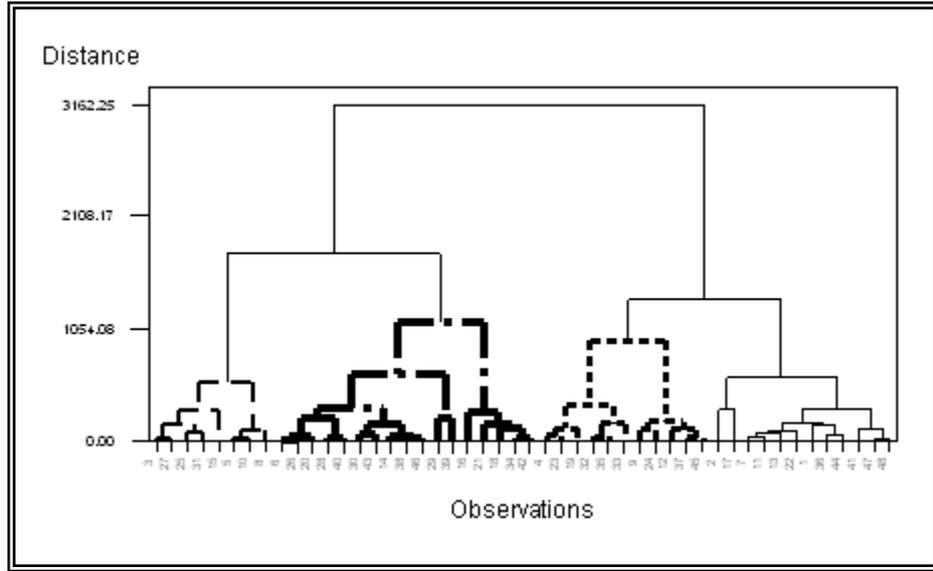
تمت عملية جمع البيانات بالاعتماد على (11) متغيراً توضيحياً منها أول ستة متغيرات تمثل القياسات الجسمية , وتليها خمسة متغيرات تمثل القياسات الوظيفية , وهذه المتغيرات هي :

- X1 : وزن الجسم (كغم) .
- X2 : طول الجسم (سم) .
- X3 : طول الجذع مع الرأس (سم) .
- X4 : طول الذراع مع الرأس (سم) .
- X5 : طول الرجل (سم) .
- X6 : طول القدم (سم) .
- X7 : السعة الحيوية (ملم / زئبق) ويقاس بجهاز خاص .
- X8 : كتف النفس (ثا) .
- X9 : النبض (ضربة / دقيقة) .
- X10 : ضغط الدم الانقباضي .
- X11 : ضغط الدم الانبساطي .

تفسير نتائج التحليل العنقودي (طريقة الربط التام) للمجموعة الأولى :
استخدم البرنامج الإحصائي (Spss V11.5) لإيجاد التحليل العنقودي الهرمي (طريقة الربط التام) ونظراً لكثرة المشاهدات تم تفسير نتائج خمس مشاهدات وهي (1,10,20,30,40) وبملاحظة نتائج التحليل العنقودي والموضحة بالجدول (1) (الملاحق) نجد أن المرحلة الأولى (Stage 1) تم فيها ربط المشاهدة رقم (47) بالمشاهدة رقم (48) وان مربع المسافة الاقليدية بينهما (20.250) وفي المرحلة رقم (21) ربطت المشاهدة رقم (47) مرة ثانية ولكن مع المشاهدة (41) ليتشكل لدينا عنقود جديد يتكون من ثلاث مشاهدات (41,47,48) . أما في المرحلة رقم (10) نجد أن المشاهدة رقم (14) ربطت مع المشاهدة رقم (38) وان مربع المسافة الاقليدية بينهما (51.500) وفي المرحلة (26) وجدنا أن المشاهدة (14) تم ربطها مرة ثانية ولكن مع المشاهدة (30) ليتشكل لدينا عنقود مكون من ثلاث مشاهدات هي (14 , 30 , 38) . أما في المرحلة رقم (20) نجد أن المشاهدة (9) ربطت مع المشاهدة (24) وان مربع المسافة الاقليدية بينهما (109.250) , وفي المرحلة (31) تم ربط المشاهدة (9) مع المشاهدة (12) ليتكون لدينا عنقود ثالث مكون من ثلاثة مشاهدات وهي (9 , 12 , 24) . أما في المرحلة (30) نجد أن المشاهدة رقم (32) قد تم ربطها مع المشاهدة رقم (33) وان مربع المسافة الاقليدية بينهما (173.250) وفي المرحلة (39) ربطت المشاهدة (32) كذلك مع المشاهدة رقم (4) ليتكون لدينا عنقود رابع مكون من ثلاث مشاهدات وهي (4 , 32 , 33) . أما في المرحلة (40) نجد أن المشاهدة رقم (3) قد تم ربطها مع المشاهدة رقم (5) وان مربع المسافة الاقليدية بينهما (549.000) وفي المرحلة (46) ربطت المشاهدة (3) كذلك مع المشاهدة رقم (6) ليتكون لدينا عنقود خامس مكون من ثلاث مشاهدات وهي (3 , 5 , 6) .

أما فيما يخص الهيكل الشجري وكما هو موضح في الشكل (1) نرى أن بيانات المجموعة الأولى قسمت إلى مجموعتين متجانستين فيما بينها وكل مجموعة مكونة من عنقودين وكل عنقود مكون من عدد من المشاهدات فمثلاً العنقود الأول في المجموعة الأولى مكون من المشاهدات (3, 27, 25, 31, 15, 5, 10, 8) بينما العنقود الثاني مكون من المشاهدات (6, 26, 20, 28, 40, 30, 43, 14, 38, 46, 29, 39, 16, 12, 18, 34, 42) .

أما المجموعة الثانية فتتكون كذلك من عنقودين وكل عنقود مؤلف من عدد من المشاهدات فمثلاً العنقود الأول مكون من المشاهدات (4, 23, 19, 32, 35, 33, 9, 24, 12, 37, 45) بينما العنقود الثاني مكون من المشاهدات (2, 17, 7, 11, 13, 22, 1, 36, 44, 41, 47, 48) .



الشكل (1)

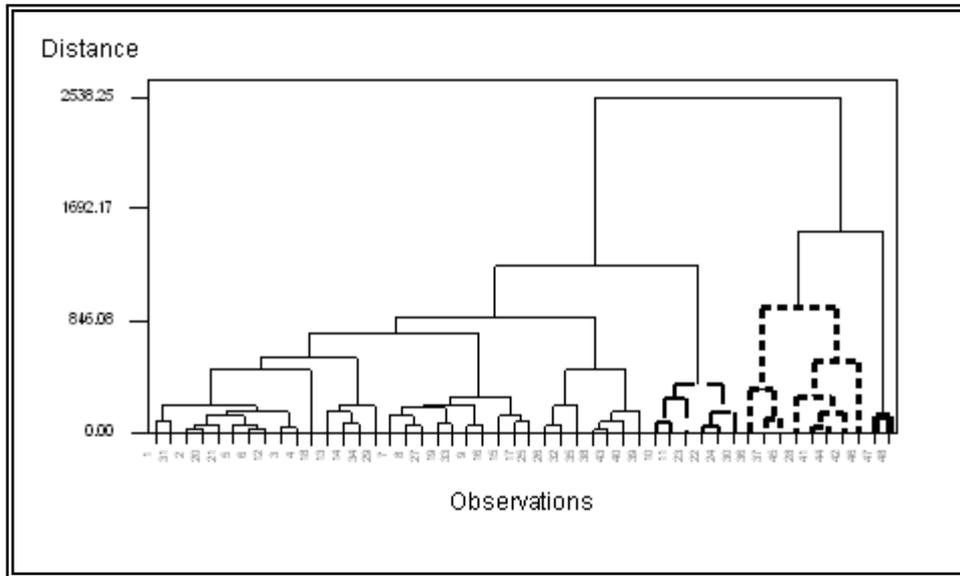
يوضح الهيكل الشجري للعناقيد المدمجة للمجموعة الأولى

تفسير نتائج التحليل العنقودي (طريقة الربط التام) للمجموعة الثانية :
استخدم البرنامج الإحصائي (Spss V11.5) لإيجاد التحليل العنقودي الهرمي (طريقة الربط التام) ونظراً لكثرة المشاهدات فقد تم تفسير نتائج خمس مشاهدات وهي (1,10,20,30,40) وبملاحظة نتائج التحليل العنقودي والموضحة في الجدول (2) (الملاحق) نجد أن المرحلة الأولى (Stage 1) تم فيها ربط مشاهدة رقم (2) بالمشاهدة رقم (20) وأن مربع المسافة الاقليدية بينهما (20.500) وفي المرحلة رقم (5) ربطت مشاهدة رقم (2) مرة ثانية ولكن مع مشاهدة (21) ليتشكل لدينا عنقود جديد يتكون من ثلاث مشاهدات (2,20,21). أما في المرحلة رقم (10) نجد أن مشاهدة رقم (14) ربطت مع مشاهدة رقم (38) وأن مربع المسافة الاقليدية بينهما (51.500) وفي المرحلة (26) وجدنا أن مشاهدة (14) قد تم ربطها مرة ثانية ولكن مع مشاهدة (30) ليتشكل لدينا عنقود مكون من ثلاث مشاهدات هي (14, 30, 38) . أما في المرحلة رقم (20) نجد أن مشاهدة (26) ربطت مع مشاهدة (32) وأن مربع المسافة الاقليدية بينهما (59.610) , وفي المرحلة (29) تم ربط مشاهدة (26) مع مشاهدة (35) ليتكون لدينا عنقود ثالث مكون من ثلاثة مشاهدات وهي (26, 32, 35) . أما في المرحلة (30) نجد أن مشاهدة رقم (13) قد تم ربطها مع مشاهدة رقم (29) وأن مربع المسافة الاقليدية بينهما (205.250) وفي المرحلة (41) ربطت مشاهدة (1) كذلك مع مشاهدة رقم (13) ليتكون لدينا عنقود رابع مكون من ثلاث مشاهدات وهي (1, 13, 29). أما في المرحلة (40) نجد أن مشاهدة رقم (28) قد تم ربطها مع مشاهدة رقم (46) وأن مربع المسافة الاقليدية بينهما (542.360) وفي المرحلة (44) ربطت مشاهدة (28)

كذلك مع المشاهدة رقم (36) ليتكون لدينا عنقود خامس مكون من ثلاث مشاهدات وهي (28) , (46, 36) .

أما فيما يخص الهيكل الشجري وكما هو موضح في الشكل (2) نرى أن بيانات المجموعة الثانية قسمت أيضاً إلى مجموعتين متجانستين فيما بينها وكل مجموعة مكونة من عنقودين وكل عنقود مكون من عدد من المشاهدات فمثلاً العنقود الأول في المجموعة الأولى مكون من المشاهدات (1, 2, 20, 21, 5, 6, 12, 3, 4, 18, 13, 14, 34 , 29 , 7 , 8 , 27 , 19 , 33 , 9 , 16 , 15 , 17 , 26 , 25 , 32 , 35 , 38 , 43 , 40 , 39) بينما العنقود الثاني مكون من المشاهدات (10 , 11 , 23 , 22 , 24 , 30) .

أما المجموعة الثانية فتتكون كذلك من عنقودين وكل عنقود مؤلف من عدد من المشاهدات فمثلاً العنقود الأول مكون من المشاهدات (36 , 37 , 45 , 28 , 41 , 44 , 42 , 46 , 47) بينما العنقود الثاني مكون من المشاهدات (48) .



الشكل (2)

يوضح الهيكل الشجري للعناقيد المدمجة للمجموعة الثانية

الاستنتاجات

يمكن توضيح أهم الاستنتاجات بماياتي :

1. إن استخدام طريقة الربط التام في التحليل العنقودي أدى إلى تقسيم المشاهدات إلى مجاميع متجانسة .
2. قسمت مشاهدات كلتا المجموعتين إلى مجاميع متجانسة فيما بينها وبالنتيجة كانت مشاهدات المجموعة الأولى مكونة من أربعة عناقيد وكل عنقود مكون من عدة مشاهدات فضلاً عن ذلك قسمت مشاهدات المجموعة الثانية إلى أربعة عناقيد وكل عنقود مكون من عدد من المشاهدات المتجانسة فيما بينها .
3. يمكن القول بان مشاهدات المجموعة الأولى (الأندية التي ترشحت إلى المربع الذهبي) كانت أكثر تجانساً من مشاهدات المجموعة الثانية (الأندية خارج المربع الذهبي) وكما هو واضح من شكل الهيكل الشجري للمجموعتين وذلك لان الأندية التي ترشحت إلى المربع الذهبي تمتلك لاعبين يتميزون بمميزات عالية للاعبين كرة السلة مثلاً (طول الجسم , طول الذراع , وزن الجسم , طول الجذع مع الرأس ... الخ) ومن ثم لا بد أن تكون مشاهدات المجموعة الأولى أكثر تجانساً من مشاهدات المجموعة الثانية .

المصادر

أ. المصادر العربية

1. البرزنجي , نظيرة صديق كريم (2002): "دراسة إحصائية لتحليل التركيب الكيميائي لبعض الصخور الكربوناتيّة في كردستان العراق". رسالة ماجستير , كلية الإدارة والاقتصاد-جامعة صلاح الدين .
2. الطائي , زياد طارق سليمان (2004) : " البناء العاملي لقياسات الجسم - وظيفته ومدى مساهمته في كفاءة الأداء بكرة السلة " , رسالة دكتوراه . كلية التربية الرياضية , جامعة الموصل .
3. كريم , ثيان إحسان (1992) : "استخدام التحليل العنقودي لتقييم الأراضي للأغراض الزراعية" , رسالة ماجستير , كلية الإدارة والاقتصاد , جامعة صلاح الدين .
4. عمر , جمين أبو بكر (2003) : "استخدام المركبات الأساسية والتحليل العنقودي لتحديد اهم العوامل المؤثرة في اصابات جيوب اللثة (دراسة ميدانية) " , رسالة ماجستير , كلية الادارة والاقتصاد , جامعة صلاح الدين .

ب. المصادر الإنكليزية

5. Everitt , B .(1977) : " **Cluster Analysis** " , Cox & Wyman , London .
6. Mardia ,K.V , Kent ,T.J, Bibby , J. M (1979) , "**Multivariate analysis**" , Academic press , London .
7. Sas Institute Inc. (1985) : " Sas Users Guide :Statistics, Version 5 Edition.

الملاحق

يبين نتائج التحليل العنقودي (طريقة الربط التام) للمجموعة الأولى

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	47	48	20.250	0	0	21
2	6	26	20.500	0	0	14
3	7	11	25.500	0	0	15
4	28	40	34.000	0	0	33
5	3	27	34.250	0	0	28
6	5	10	35.250	0	0	19
7	38	46	35.750	0	0	10
8	34	42	36.250	0	0	24
9	37	45	39.750	0	0	22
10	14	38	51.500	0	7	26
11	32	35	51.500	0	0	30
12	36	44	54.500	0	0	25
13	4	23	54.500	0	0	23
14	6	20	61.000	2	0	33
15	7	13	66.500	3	0	17
16	30	43	74.250	0	0	26
17	7	22	87.500	15	0	27
18	25	31	97.500	0	0	28
19	5	8	103.250	6	0	40
20	9	24	109.250	0	0	31
21	41	47	112.500	0	1	35
22	12	37	115.750	0	9	31
23	4	19	126.000	13	0	39
24	18	34	136.250	0	8	29
25	1	36	137.500	0	12	27
26	14	30	155.750	10	16	38
27	1	7	167.250	25	17	35
28	3	25	172.250	5	18	37
29	18	21	172.750	24	0	34
30	32	33	173.250	11	0	39
31	9	12	186.500	20	22	43
32	29	39	215.500	0	0	42
33	6	28	226.000	14	4	38
34	16	18	277.250	0	29	44
35	1	41	286.250	27	21	41
36	2	17	290.250	0	0	41
37	3	15	297.500	28	0	40
38	6	14	311.500	33	26	42
39	4	32	333.250	23	30	43
40	3	5	549.000	37	19	46
41	1	2	587.500	35	36	45
42	6	29	634.000	38	32	44
43	4	9	944.250	39	31	45
44	6	16	1123.750	42	34	46
45	1	4	1328.250	41	43	47
46	3	6	1760.500	40	44	47
47	1	3	3162.250	45	46	0

يبين نتائج التحليل العنقودي (طريقة الربط التام) للمجموعة الثانية

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	20	20.500	0	0	5
2	38	43	29.360	0	0	16
3	6	12	30.500	0	0	8
4	3	4	39.750	0	0	25
5	2	21	51.500	1	0	21
6	41	44	53.290	0	0	24
7	8	27	53.610	0	0	19
8	5	6	57.000	0	3	21
9	22	24	57.250	0	0	27
10	26	32	59.610	0	0	29
11	9	16	62.000	0	0	32
12	14	34	65.000	0	0	26
13	19	33	72.500	0	0	28
14	1	31	84.500	0	0	31
15	17	25	88.290	0	0	22
16	38	40	91.990	2	0	23
17	10	11	94.250	0	0	34
18	37	45	112.000	0	0	36
19	7	8	125.260	0	7	28
20	47	48	128.090	0	0	46
21	2	5	137.750	5	8	25
22	15	17	138.500	0	15	33
23	38	39	159.250	16	0	39
24	41	42	161.540	6	0	35
25	2	3	166.000	21	4	31
26	13	14	166.500	0	12	30
27	22	30	167.500	9	0	37
28	7	19	186.000	19	13	32
29	26	35	200.250	10	0	39
30	13	29	205.250	26	0	41
31	1	2	205.500	14	25	38
32	7	9	209.360	28	11	33
33	7	15	265.000	32	22	42
34	10	23	266.000	17	0	37
35	28	41	278.060	0	24	40
36	36	37	330.490	0	18	44
37	10	22	368.750	34	27	45
38	1	18	477.000	31	0	41
39	26	38	479.590	29	23	43
40	28	46	542.360	35	0	44
41	1	13	562.500	38	30	42
42	1	7	749.250	41	33	43
43	1	26	869.340	42	39	45
44	28	36	942.740	40	36	46
45	1	10	1260.250	43	37	47
46	28	47	1533.090	44	20	47
47	1	28	2538.250	45	46	0