IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (3) Part (1): 1-9

UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN:2222-2995 E-ISSN:3079-3521
University of Kirkuk Journal For Administrative and Economic Science

Naser Hadeel Imad .Using hierarchical cluster analysis to classify factors affecting blood clotting in Kirkuk city. *University of Kirkuk Journal For Administrative and Economic Science* (2025) 15 (3) Part (1):1-9.

Using hierarchical cluster analysis to classify factors affecting blood clotting in Kirkuk city

Hadeel Imad Naser [1]

[1] *College of Administration and Economics-Kirkuk University, Kirkuk, Iraq*

hadeel.imad@uokirkuk.edu.iq [1]

Abstract: Blood clotting is a health problem that affects people of all ages and both sexes. It occurs due to several reasons, So it was necessary to study the factors that cause an increase in the incidence of the disease. The aim of this research is to apply statistical methods represented by hierarchical cluster analysis to classify the factors affecting blood clotting disease. These methods were applied to a sample size of (90) obtained from Kirkuk General Hospital in Kirkuk. The influencing factors were represented by (sex, age, blood pressure, diabetes, and vitamin K and D deficiency), Using the statistical program SPSS v. 25. After converting the data to standard values (due to the difference in the factors' scales) and after analysis, It was concluded that all the hierarchical methods used, namely (single linkage, complete linkage, median linkage, centroid linkage and ward's method), were very close in the results, as they all agreed that the clustering started from the factor (age and diabetes). This indicates that these factors were among the most influential factors in the incidence of the disease, followed by blood pressure, vitamin D deficiency, vitamin K deficiency and gender. The linkage methods (single, centroid and median linkage) showed the presence of the sex factor alone in the first cluster and the rest of the factors in the second cluster, which explains the similarity of the factors in the characteristics, while the linkage method (complete and ward) showed the joining of the two factors (sex and vitamin K deficiency) in the first cluster and the rest of the factors in the second cluster.

Keywords: Single linkage method, Complete linkage method, Median linkage method, Centroid linkage method, Ward's method, Blood clotting.

استعمال التحليل العنقودي الهرمي في تصنيف العوامل المؤثرة على مرض تخثر الدم في مدينة كركوك

م.م. هديل عماد ناصر [1]

[1] *كلية الإدارة والاقتصاد-جامعة كركوك، كركوك، العراق*

المستخلص: يعد مرض تخثر الدم مشكلة صحية يعاني منها مختلف الاعمار وكلا الجنسين والتي تحدث نتيجة عدة اسباب، لذا كان لابد من دراسة للعوامل التي تسبب في زيادة الاصابة بالمرض. ان الهدف من هذا البحث هو تطبيق الاساليب الاحصائية والمتمثلة بالتحليل العنقودي الهرمي لتصنيف العوامل المؤثرة على مرض تخثر الدم، تم تطبيق

هذه الاساليب على عينة حجمها (٩٠) تم الحصول عليها من مستشفى كركوك العام في كركوك والعوامل المؤثرة متمثلة ب (الجنس، العمر، ضغط الدم، السكري، ونقص فيتامين k و d)، باستعمال البرنامج الاحصائي spss v.25 وبعد تحويل البيانات الى قيم معيارية (نظراً لاختلاف مقاييس العوامل) وبعد التحليل تم التوصل الى ان جميع الطرائق الهرمية المستخدمة وهي طريقة (الربط المنفرد، الربط الشامل، الربط بالوسيط، الربط المركزية وطريقة ward) كانت متقاربة جداً في النتائج حيث اجمعت جميعها على ان التعنقد بدأ من عامل (العمر والسكري) وهذا يدل على ان هذه العوامل كانت من اكثر العوامل تأثيراً في الاصابة بالمرض ثم تليها العوامل ضغط الدم ، نقص فيتامين d ، نقص فيتامين k والجنس. بينت طرق الربط (المنفرد، المركزية والربط بالوسيط) وجود عامل الجنس وحده في العنقود الاول وبقية العوامل في العنقود الثاني الذي يفسر تشابه العوامل في الصفات، بينما بيّنت طريقة الربط (الشامل و wardعلى انضمام العاملين (الجنس ونقص فيتامين k) في العنقود الاول وبقية العوامل في العنقود الثاني.

الكلمات المفتاحية: طريقة الربط المنفرد، طريقة الربط الشامل، طريقة الربط بالوسيط، طريقة الربط المركزية، طريقة ward، تخثر الدم.

Corresponding Author: E-mail: hadeel.imad@uokirkuk.edu.iq

المقدمة

يعد مرض تخثر الدم من المشاكل الصحية التي تؤثر على جميع الافراد ولمختلف الاعمار وبالأخص التي تزيد اعمارهم عن ٤٠ سنة، حيث تعتبر عملية تخثر الدم من العمليات الحيوية والتي تعرف على انها مقاومة النزف ومنع فقدان الدم عند الاصابة بالجروح اذ يتكون الدم من جزأين: الجزء الاول سائل يعرف بالبلازما والجزء الثاني خلوي يضم ثلاث انواع من الخلايا منها البلازما الذي يشكل نسبة ٥٥٪ من الدم والذي يحتوي على نسبة ٩١٪ من الماء، والثاني هي المواد الصلبة وتشمل ثلاث بروتينات تشكل ٦-٧٪ من البلازما وغيرها من النواتج الخلوية ، عند حدوث تمزق او جرح فإن الجسم يعمل خثرة بشكل طبيعي لشفاء المنطقة المتضررة الا انه في بعض الحالات قد لا يتم تكوين الخثرة او تتكون خثرة الدم دون وجود جرح او نزيف مما يشكل خطر على حياة الشخص المصاب والتي تستدعي الى العلاج.

المبحث الأول: الاطار النظري

قامت عدة دراسات باستعمال التحليل العنقودي في بعض المجالات منها : استعملت تهاني واخرون عام (٢٠١٠) التحليل العنقودي في تحليل وتصنيف العوامل التي تؤثر على الاصابة بمرض القلب حيث استعمل الباحثين اربع من طرائق التعنقد على عينة متكونة من (٦٣٠) مريض بالقلب، وتوصلوا الى ان ثلاث منها اظهرت نفس النتائج بالنسبة للتعنقد وقام الباحثين بتجميعها في ثلاثة عناقيد حسب درجة تقاربها من الخصائص. قام فيصل عام (٢٠١٤) بإستعمال التحليل العنقودي لتصنيف المحافظات العراقية وفقاً لما تعرضت له من تأثيرات مسببة لمرض السرطان كما استعمل طريقة اخرى للتصنيف تعتمد على الانواع الاكثر انتشاراً من الامراض السرطانية وتوصل الى وجود اختلاف في التصنيف وذلك حسب درجة التشابه بالصفات وللفترة الزمنية (١٩٩٨-٢٠٠٦) [9]. في عام (٢٠١٥) قام ابراهيم واحمد بإستعمال التحليل العنقودي لتصنيف مجاميع المحافظات العراقية التي تعاني من مشكلات الحرمان لعام (٢٠٠٩) وذلك باعتمادهم على البيانات المتوفرة في الجهاز المركزي في الاحصاء- تقرير الاقتصاد العراقي لعام (٢٠٠٩) وتوصلوا الى ان المحافظات (بابل ، كربلاء ، المثنى ، النجف وذي قار) والمتمثلة بالعنقود الاول قد اعطت اعلى متوسط لديل الحرمان بالنسبة لمتغيري التعليم ووضع الاسرة الاقتصاد. قامت خلود وعمر في عام (٢٠١٦) بدراسة العوامل التي تؤثر على تزايد الولادات المشوهة في العراق من خلال قاعدة البيانات المتوفرة في دوائر وزارة الصحة في العراق باستثناء اقليم كردستان والتي تم جمعها عام (٢٠١٢) واستعملا التحليل العاملي لمعرفة اهم العوامل التي تؤثر على التشوهات الولادية في العراق و كذلك التحليل العنقودي لمعرفة أي محافظة تتعنقد بشكل مبكر على ولادة طفل مشوه وتوصلوا الى ان المحافظات كانت متقاربة جداً في النتائج اما محافظتي الانبار والموصل كانتا الاعلى من حيث التشوهات الولادية . قامت وفاء عبد الصمد و وفاء علي عام (٢٠٢٠) بإستعمال التحليل العنقودي لمعرفة الفرق في مستوى المؤشرات المالية لشركة التأمين الوطنية للفترة (٢٠٠٨-٢٠١٥)، استعملت الباحثتان الطريقة الهرمية (الربط المنفرد) وطريقة (k-mean) وتوصلتا الى ان عام (٢٠٠٨) شهد انخفاضاً في المركز المالي للشركة مقارنةً بعام (2015) الذي شهد ارتفاعاً فيه. استعملت الباحثة اسيا الطرائق الهرمية للتحليل العنقودي عام (٢٠٢١) المتمثلة بطريقة (الربط المنفرد، الربط الشامل، الربط المركزية، الربط بالوسيط والهرمية) لتصنيف المحافظات العراقية من خلال العوامل التي تؤثر على وفيات الامهات عند الولادة وتوصلت الى ان الطرق كانت متقاربة في النتائج فجميعها اجمعت على ان محافظة صلاح الدين كانت اكثر المحافظات تأثراً بالعوامل المؤدية الى وفيات الامهات .

اولاً: الاهداف

ان الهدف من البحث هو دراسة العوامل التي تؤثر على الاصابة بمرض تخثر الدم وتصنيف هذه العوامل (المتغيرات) في مجاميع (عناقيد) حسب درجة التشابه بينها وذلك باستعمال التحليل العنقودي الهرمي وباستعمال الطرائق المستخدمة في هذا التحليل نأمل للوصول الى افضل طريقة لتصنيف تلك العوامل الى عناقيد.

ثانياً: منهجية البحث

التحليل العنقودي cluster analysis : يعد من اساليب التصنيف للبيانات متعدد المتغيرات ويهدف الى تجميع الحالات او المتغيرات في مجموعات متمايزة تسمى العناقيد، يتم من خلال هذه العملية تقليل عدد المتغيرات. تعتمد عملية التجميع على مبدأ التشابه في الخصائص بين العناصر داخل العنقود الواحد والتباين الخارجي أي الاختلاف بين العناقيد المختلفة. تعد حسابات المسافات حجز الزاوية في عملية تكوين العناقيد أي في تجميع العناصر المتقاربة فالنقاط التي تكون المسافة بينها قليلة تتجمع ضمن نفس العنقود والتي تكون بينها المسافة بعيدة تتجمع في عناقيد اخرى وبهذا تكون العلاقة عكسية بين المسافة ودرجة التجانس، توجد عدة مقاييس في حساب المسافة بين نقاط المفردات اكثرها شيوعاً المسافة الاقليدية (Euclidean distance) والتي تحسب بالصيغة الاتية:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{m} (x_{ki} - x_{kj})^2} \qquad (1)$$

حيث ان x_{ki} تمثل قيمة المتغير x_k للعنصر i و ان x_{kj} تمثل قيمة المتغير x_k للعنصر j [12].

تحسب المسافة بين جميع ازواج (الحالات او المتغيرات) ووضعها في مصفوفة القرب (proximity matrix) والتي سيتم ضم زوج النقاط ذات المسافة الاصغر مع بعضها في كل خطوة [11][17].

التحليل العنقودي الهرمي Hierarchical cluster analysis : تبدأ عملية تكوين العناقيد من معالجة كل نقطة بيانات كعنقود منفرد ثم تتكاثف المجموعات تدريجياً عبر ادماج العناقيد الاكثر تقارباً في الخصائص حيث يتم دمج الازواج ذات المسافة الاقصر بينها اولاً ثم يتطور هذا التسلسل تصاعدياً حتى اكتمال اندماج جميع الوحدات في بنية هرمية التي يشار اليها بمخطط الشجرة.

ان للتحليل العنقودي الهرمي اسلوبين هما:

١- التحليل العنقودي الهرمي التقسيمي: وهي المنهجية التي يتم من خلالها تجزئة العناقيد الكبيرة الى تجمعات فرعية أصغر حجماً وفق معايير التشابه بينها ويتم تكرار عملية التجزئة تنازلياً حتى تتحول كل مشاهدة الى عنقود مستقل.

٢- التحليل العنقودي الهرمي التجميعي: وهي عكس منهجية التحليل العنقودي التقسيمي حيث يتم التجزئة بدايةً بعنقود واحد ثم تبدأ عملية دمج العناقيد المتشابهة مع بعضها تصاعدياً حتى يتم ضم جميع العناصر في عنقود واحد .

وللتحليل العنقودي الهرمي التجميعي عدة طرائق منها:

أ- **طريقة الربط المنفرد Single linkage method**: وهي الطريقة التي يتم عن طريقها دمج مجموعتين متقاربتين في الصفات اي تكون المسافة بينها قصيرة الى مجموعة جديدة حيث تستمر العملية وصولاً الى دمج اخر عامل او مجموعة حسب قصر المسافات بينها لتكَون مجموعة واحد تضم جميع العناقيد. لنفرض ان المسافة بين عنصرين y_i و y_j هي d_{ij} وان C_P و C_q يمثلان عنقودان يحتويان n_p و n_q من المفردات فتكون D(p,q) هي المسافة بين العناقيد وليكن C_r يمثل العنقود الجديد الذي يضم العنقودين عندها تحسب المسافة بين العنقودين C_P و C_q عن طريق الربط المنفرد خلال الصيغة :

$$D(C_p, C_q) = \min\{d_{ij}, i \in C_p, j \in C_q, p \neq q\} \qquad (2)$$

ب- **طريقة الربط المتكامل (الشامل) Complete linkage method** : وهي الطريقة التي تعمل عكس طريقة الربط المنفرد حيث يتم جمع مجموعتين ذات المسافة الابعد ودمجها في عنقود وتستمر العملية بتجميع العناقيد المنفردة لتشكل عنقود واحد، وتحسب المسافة بين العنقودين C_P و C_q عن طريق الربط الشامل من خلال الصيغة:

$$D(C_p, C_q) = max\{d_{ij}, i \in C_p, j \in C_q, p \neq q\} \qquad (3)$$

ت- **طريقة الربط بالوسيط Median linkage method**: لا يتم في هذه الطريقة حساب مسافة القرب بين عناصر المجاميع بأخذ اقرب او ابعد مسافة انما تُحدد مسافة التشابه بين العنقودين بمتوسط المسافة بينهم، حيث تكون الصيغة:

$$m(C_p, C_q) = \frac{1}{2}(\bar{Y}^p + \bar{Y}^q) \qquad (4)$$

حيث ان $\bar{Y}^q$ و $\bar{Y}^p$ مراكز العنقودين C_p, C_q .

ث- **طريقة الربط المركزية Centroid linkage method :** يتم من خلال هذه الطريقة تجميع العناقيد بالاعتماد على مراكزها اي ان المسافة بين العناقيد عبارة عن المسافة بين مراكزها ، تتم عملية دمج مجموعتين لهما اصغر مسافة بين مراكزها ، في كل خطوة وبعد دمج العنقودين C_p, C_q في عنقود جديد C_r الذي يحتوي على n_p ، n_q ، n_r حيث ان $(n_r=n_p+n_q)$ وتكون مراكزهم $\bar{Y}^q$ ، $\bar{Y}^p$ ، $\bar{Y}^r$ و المسافة تحسب عن طريق الربط المركزي حسب الصيغة:

$$\bar{Y}^r = \frac{1}{n_r}(n_p\bar{Y}^p + n_q\bar{Y}^q) \qquad (5)$$

ج- الطريقة الهرمية Ward's method: تعتمد هذه الطريقة على جدول تحليل التباين في حسابها حيث تظهر نتائج التحليل مجموع مربع المسافات الصغيرة داخل المجاميع ومجموع مربع المسافات الكبيرة بين المجاميع. بافتراض ان لدينا n من العينات موزعة داخل k من المجاميع حيث العنصر في العنقود C_t يرمز له Y_i^t وان n_t تمثل عدد العينات في العنقود C_t ، و $\bar{Y}^t$ يمثل مركز العنقود فإن مجموع المربعات داخل العنقود تكون :

$$S_t = \sum_{i=1}^{n_t} (Y_i^t - \bar{Y}^t)^T (Y_i^t - \bar{Y}^t) \qquad (6)$$

وان مجموع المربعات الكلي للمجاميع k تكون:

$$S = \sum_{t=1}^{k} \sum_{i=1}^{n_t} (Y_i^t - \bar{Y}^t)^T (Y_i^t - \bar{Y}^t) \qquad (7)$$

وان المسافة بين العنقودين C_p،C_q معرفة بمجموع المربعات بين العنقودين :

$$D^2(C_p, C_q) = S_r - S_p - S_q \qquad (8)$$

المبحث الثاني: النتائج والمناقشة

يتم في هذا الجانب تطبيق الطرائق المذكورة في الجانب النظري على عينة من مرضى تخثر الدم والتي تتألف من (٩٠) مريض من كلا الجنسين، تم الحصول عليها من مستشفى كركوك العام في كركوك لسنة (٢٠٢٣)، تضمنت البيانات العوامل التي تؤثر على زيادة نسبة الاصابة بمرض تخثر الدم، والعوامل هي:

١. الجنس.
٢. العمر حيث تتراوح فئات العمر بين (١٨-٧٣) سنة.
٣. نقص فيتامين k .
٤. ضغط الدم.
٥. السكري.
٦. نقص فيتامين d.

ولتحليل البيانات، أُتبعت الاجراءات الاحصائية التالية باستعمال البرنامج الاحصائي (spss v.25) :

- تحويل البيانات الى متغيرات معيارية نظراً لاختلاف وحدات قياس العوامل.
- فحص البيانات من وجود القيم المتطرفة فيها حيث تم استعمال مخطط الصندوق boxplot واظهر الرسم البياني على عدم وجود قيم متطرفة في البيانات.
- حساب معامل الارتباط بيرسون لتقييم العلاقات الخطية وللكشف عن وجود الارتباطات بين المتغيرات حيث اظهر الاختبار غياب الارتباط بين كل المتغيرات عدا متغير العمر والسكري حيث كانت قيمة معامل الارتباط ٠,٦٢ حيث هذه القيمة لا تدل على وجود تعدد خطي عالٍ.
- بالاعتماد على مقياس المسافة الاقليدية المربعة (Squared Euclidian distance) لحساب المسافة والاختلاف بين العوامل، ولتكوين العناقيد في التحليل العنقودي ولجميع الطرائق، كانت النتائج كالاتي:

أ- طريقة الربط المنفرد:

جدول (١) : مراحل التجميع للعناقيد بطريقة الربط المنفرد

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	5	41.794	0	0	2
2	2	4	110.858	1	0	3
3	2	6	117.494	2	0	4
4	2	3	130.896	3	0	5
5	1	2	159.284	0	4	0

يبين الجدول (١) التسلسل الهرمي لتكوين العناقيد بناءً على التشابه بين العوامل من خلال طريقة الربط المنفرد:

المرحلة الأولى: اندمج العاملان (٢) و(٥) المُتمثلان في (العمر والسكري) نتيجة المسافة الأصغر بينهما (٤١,٧٩٤)، مما يشير إلى علاقة ارتباطية قوية بينهم.

المرحلة الثانية: اتحدت المجموعة الناتجة (٢+٥) مع العامل (٤) (ضغط الدم) عند مسافة (١١٠,٨٥٨)، مما يعكس التفاعل بين العمر والأمراض المزمنة.

المرحلة الثالثة: انضم العامل (٦) (نقص فيتامين d) إلى التجميع السابق عند مسافة (١١٧,٤٩٤)، مُبرزاً دوره في اضطرابات التخثر المرتبطة بالأمراض المزمنة.
المرحلة الرابعة: اتحدت المجموعة ((٢+٥) + (٤+٦)) مع العامل (٣) (نقص فيتامين k) عند مسافة (١٣٠,٨٩٦) مؤكدً تأثيره كعامل خطر في المرض.
المرحلة الخامسة: اكتمل التجميع بانضمام العامل (١) (الجنس) عند المسافة الأكبر (١٥٩,٢٨٤)، مما يوضح طبيعته المنفصلة نسبياً.

جدول ٢: العناقيد النهائية للعوامل (عضوية العناقيد)

Case	2 Clusters
الجنس	1
العمر	2
نقص فيتامين k	2
ضغط الدم	2
السكري	2
نقص فيتامين d	2

الجدول (٢) اعلاه يبين التوزيع النهائي للعناقيد حيث يبين ان في حالة تصنيف العوامل الى عنقودين، ان عامل الجنس كان العامل الوحيد في (العنقود الاول) بينما العوامل الاخرى اجتمعت في نفس العنقود (العنقود الثاني) وهذا يفسر تشابه العوامل في العنقود الثاني وتقارب المسافة بينهم وتأثيرهم في المرض.

ب- طريقة الربط الشامل:

جدول (٣): مراحل التجميع للعناقيد بطريقة الربط الشامل

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	5	41.794	0	0	4
2	4	6	117.494	0	0	4
3	1	3	159.284	0	0	5
4	2	4	161.490	1	2	5
5	1	2	208.488	3	4	0

يبين الجدول (٣) التسلسل الهرمي لتكوين العناقيد بناءً على التشابه بين العوامل من خلال طريقة الربط الشامل:
المرحلة الأولى: اندمج العاملان (٢) و(٥) المُتمثلان في (العمر والسكري) نتيجة المسافة الأصغر بينهما (٤١,٧٩٤) مما يشير إلى علاقة ارتباطية قوية بينهم.
المرحلة الثانية: اندمج العاملان (٤) و(٦) المُتمثلان في (ضغط الدم ونقص فيتامين d) نتيجة المسافة الأصغر بينهما (١١٧,٤٩٤) مما يشير إلى علاقة ارتباطية جيدة بينهم.
المرحلة الثالثة: اندمج العاملان (١) و(٣) المُتمثلان في (الجنس ونص فيتامين k) نتيجة المسافة الأصغر بينهما (١٥٩,٢٨٤).
المرحلة الرابعة: اتحدت المجموعة (٢+٥) مع (٤+٦) عند مسافة (١٦١,٤٩٠) .
المرحلة الخامسة: اكتمل التجميع بدمج العوامل ((٢+٥) + (٤+٦)) مع العوامل (١+٣) وذلك لحصولهم على ابعد مسافة (٢٠٨,٤٨٨).

جدول (٤): العناقيد النهائية للعوامل (عضوية العناقيد)

Case	2 Clusters
الجنس	1
العمر	2
نقص فيتامين k	1
ضغط الدم	2
السكري	2
نقص فيتامينd	2

الجدول (٤) اعلاه يبين التوزيع النهائي للعناقيد حيث يبين ان عامل (الجنس ونقص فيتامين k) تم وضعهم في عنقود واحد (العنقود الاول) لقرب التشابه بينهم بينما العوامل الاخرى تم وضعها في العنقود الثاني وهذا يفسر تشابه العوامل في العنقود الثاني وتقارب المسافة بينهم وتأثيرهم في المرض.

ت- طريقة الربط بالوسيط:

جدول (٥): مراحل التجميع للعناقيد بطريقة الربط بالوسيط

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	5	41.794	0	0	3
2	4	6	117.494	0	0	3
3	2	4	102.393	1	2	4
4	2	3	128.395	3	0	5
5	1	2	126.137	0	4	0

يبين الجدول (٥) التسلسل الهرمي لتكوين العناقيد بناءً على التشابه بين العوامل من خلال طريقة الربط بالوسيط:
المرحلة الأولى: اندمج العاملان (٢) و(٥) المُتمثلان في (العمر والسكري) نتيجة المسافة الأصغر بينهما (٤١,٧٩٤)، مما يشير إلى علاقة ارتباطية قوية بينهم.
المرحلة الثانية: اندمج العاملان (٤) و(٦) المتمثلان في (ضغط الدم ونقص فيتامين d) عند بمسافة (١١٧,٤٩٤) وهذا يفسر ارتباط واضح بين العوامل.
المرحلة الثالثة: دمجت العوامل (٢+٥) مع (٤+٦) عند مسافة (١٠٢,٣٩٣).
المرحلة الرابعة: انضم العامل (٣) (نقص فيتامين k) مع العوامل المدمجة ((٢+٥)+(٤+٦)) عند مسافة (١٢٨,٣٩٥).
المرحلة الخامسة: اكتمل التجميع بانضمام العامل (١) (الجنس) عند المسافة الأكبر (١٢٦,١٣٧)، مما يوضح طبيعته المنفصلة نسبياً.

جدول (٦): العناقيد النهائية للعوامل (عضوية العناقيد)

Case	2 Clusters
الجنس	1
العمر	2
نقص فيتامين k	2
الدم ضغط	2
السكري	2
نقص فيتامين d	2

الجدول (٦) اعلاه يبين التوزيع النهائي للعناقيد حيث يبين ان عامل (الجنس) تم وضعه فقط في العنقود الاول بينما العوامل الاخرى تم تجميعهم في (العنقود الثاني) وهذا يفسر تشابه العوامل في العنقود الثاني وتقارب المسافة بينهم وتأثيرهم في المرض.

ث- الطريقة المركزية:

جدول (٧): مراحل التجميع للعناقيد بطريقة الربط المركزية

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	5	41.794	0	0	3
2	4	6	117.494	0	0	3
3	2	4	102.393	1	2	4
4	2	3	128.395	3	0	5
5	1	2	137.063	0	4	0

يبين الجدول (٧) التسلسل الهرمي لتكوين العناقيد بناءً على التشابه بين العوامل من خلال طريقة الربط المركزية:
المرحلة الأولى: اندمج العاملان (٢) و(٥) المُتمثلان في (العمر والسكري) نتيجة المسافة الأصغر بينهما (٤١,٧٩٤)، مما يشير إلى علاقة ارتباطية قوية بينهم.
المرحلة الثانية: اندمج العاملان (٤) و (٦) المتمثلان في (ضغط الدم ونقص فيتامين d) عند بمسافة (١١٧,٤٩٤) وهذا يفسر ارتباط واضح بين العوامل.

المرحلة الثالثة: دمجت العوامل (٢+٥) مع (٤+٦) عند مسافة (١٠٢,٣٩٣) .
المرحلة الرابعة: انضم العامل (٣) المتمثل (نقص فيتامين k) مع العوامل المدمجة ((٢+٥)+(٤+٦)) عند مسافة (١٢٨,٣٩٥).
المرحلة الخامسة: اكتمل التجميع بانضمام العامل (١) (الجنس) عند المسافة الأكبر (١٣٧,٠٦٣) ، مما يوضح طبيعته المنفصلة نسبياً.

جدول (٨): العناقيد النهائية للعوامل (عضوية العناقيد)

Case	2 Clusters
الجنس	1
العمر	2
نقص فيتامين k	2
ضغط الدم	2
السكري	2
dنقص فيتامين	2

الجدول (٨) اعلاه يبين التوزيع النهائي للعناقيد حيث يبين ان عامل (الجنس) تم وضعه في العنقود الاول بينما العوامل الاخرى تم تجميعهم في العنقود الثاني وهذا يفسر تشابه العوامل في العنقود الثاني وتقارب المسافة بينهم وتأثيرهم في المرض.

ج-طريقة Ward:

جدول (٩) : مراحل التجميع للعناقيد بطريقة الربط Ward

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	5	20.897	0	0	4
2	4	6	79.644	0	0	4
3	1	3	159.286	0	0	5
4	2	4	261.679	1	2	5
5	1	2	398.972	3	4	0

يبين الجدول (٩) التسلسل الهرمي لتكوين العناقيد بناءً على التشابه بين العوامل من خلال طريقة الربط ward:
المرحلة الأولى: اندمج العاملان (٢) و(٥) المُتمثلان في (العمر والسكري) نتيجة المسافة الأصغر بينهما (٢٠,٨٩٧)، مما يشير إلى علاقة ارتباطية قوية بينهم.
المرحلة الثانية: اندمج العاملان (٤) و (٦) المتمثلان في (ضغط الدم ونقص فيتامين d) عند بمسافة (٧٩,٦٤٤) وهذا يفسر ارتباط واضح بين العوامل .
المرحلة الثالثة: دمج العامل (١) مع (٣) المتمثلان (الجنس ونقص فيتامين k) عند مسافة (١٥٩,٢٨٦).
المرحلة الرابعة: دمجت العوامل (٢+٥) مع العوامل (٤+٦) عند مسافة (٢٦١,٦٧٩).
المرحلة الخامسة: اكتمل التجميع بدمج العوامل (١+٣) مع العوامل ((٢+٥) + (٤+٦)) عند المسافة الأكبر (٣٩٨,٩٧٢)، مما يوضح طبيعته المنفصلة نسبياً.

جدول (١٠): العناقيد النهائية للعوامل (عضوية العناقيد)

Case	2 Clusters
الجنس	1
العمر	2
نقص فيتامين k	1
ضغط الدم	2
السكري	2
نقص فيتامين d	2

الجدول (١٠) اعلاه يبين التوزيع النهائي للعناقيد حيث يبين ان عامل (الجنس ونقص فيتامين k) تم وضعهم في العنقود الاول بينما العوامل الاخرى تم وضعها في العنقود الثاني.

المبحث الثالث: الاستنتاجات

من خلال نتائج التحليل العنقودي الهرمي بطرائقه الخمس تم التوصل الى الاتي:

(١) كل الطرائق كانت متقاربة جداً في النتائج حيث اجمعت جميعها على ان التعاقد بدأ بدمج العاملين (٢ و٥) والذي بدوره يقودنا الى التوصل الى ان عامل العمر والسكري كانا اكثر العوامل تأثيراً في الاصابة بمرض تخثر الدم لذا يجب القيام بفحص دوري للسكري ولجميع الاعمار، ثم يليه العوامل المؤثرة بالتتابع وهي عامل ضغط الدم، نقص فيتامين d، نقص فيتامين k والجنس.

(٢) تشابه طريقة (الربط الشامل مع ward) وطريقة (الربط بالوسيط مع الربط المركزية) في مراحل تجميع العناقيد وبمسافات قريبة جداً.

(٣) اجمعت طرائق الربط (المنفرد، المركزية والربط بالوسيط) على ان العنقود الاول قد ضم عامل الجنس فقط وبقية العوامل اجتمعت في العنقود الثاني الذي يفسر تشابه العوامل في العنقود الثاني وتأثيرها على المرض، بينما اجمعت طريقة الربط (الشامل و ward) على ان العنقود الاول ضم العاملين (الجنس ونقص فيتامين k) وبقية العوامل في العنقود الثاني.

المصادر

المصادر العربية:

١- حسين، اسيا حمود (٢٠٢١). تصنيف المحافظات العراقية من خلال العوامل المؤثرة على وفيات الامهات عند الولادة باستخدام بعض الطرق الهرمية للتحليل العنقودي. مجلة الادارة والاقتصاد. العدد ١٣٠/ كانون الاول/ ٢٠٢١ الصفحات ٢١٣-٢٢٦.

٢- حنيش، ابراهيم سليمان و اسمييو ، خلود سليمان (٢٠١٩). المقارنة بين طرق التعاقد الهرمي واختيار أفضلها مع تطبيق عملي على بعض انواع الحليب المبيع في مدينة مصراته. عدد خاص بالمؤتمر السنوي الثالث حول نظريات وتطبيقات العلوم الاساسية والحيوية. ٧ سبتمبر ٢٠١٩. مجلة العلوم.

٣- خمو ، خلود يوسف ومحسن، عمر محي (٢٠١٦). العوامل المؤثرة على التشوهات الولادية في العراق لعام ٢٠١٢. **مجلة جامعة كركوك للعلوم الادارية والاقتصادية**. المجلد (٦) العدد (١).

٤- خمو، خلود يوسف ورؤوف، تمارة (٢٠١٧). انتشار ظاهرة التعليم الخصوصي في مؤسسات العراق الأسباب والعوامل. **مجلة جامعة كركوك للعلوم الإدارية والاقتصادية**. المجلد (٧) العدد (١).

٥- شعيبث ، ليلى جاسم (٢٠١٧). دراسة تأثير عامل التدخين والجنس والعمر في سرعة تجلط الدم والاصابة بالجلطات الدموية. مجلة القادسية للعلوم الصرفة. المجلد ٢٢. العدد ٣.

٦- عاشور، وفاء عبد الصمد وسلطان، وفاء علي (٢٠٢٠). تصنيف الواقع المالي لشركة التأمين الوطنية للفترة (٢٠٠٨-٢٠١٥) باستخدام التحليل العنقودي. مجلة الرافدين لعلوم الحاسوب والرياضيات، المجلد (١٤) العدد (٢).

٧- عباس، تهاني مهدي، حميد، سميرة مزهر ونايف، قتيبة نبيل (٢٠١٠). استخدام التحليل العنقودي cluster analysis في تحليل العوامل المؤثرة على مرض القلب. مجلة جامعة النهرين. المجلد ١٣ العدد (٣) ايلول ٢٠١٠ ص ٥٨-٦٥.

٨- كاظم، ابراهيم جواد ومحمد، احمد سلطان (٢٠١٥). تصنيف وتقييم دليل الحرمان في العراق لعام ٢٠٠٩ باستعمال بعض طرائق التحليل العنقودي. مجلة العلوم الاقتصادية والادارية، المجلد ٢١ العدد ٨٢ الصفحات ٢٩١-٤١١.

٩- نامق، فيصل ناجي (٢٠١٤). استخدام التحليل العنقودي لتصنيف مرض السرطان حسب الانواع والحالات في العراق للفترة (١٩٩٨-٢٠٠٦). **مجلة جامعة كركوك للعلوم الادارية والاقتصادية**. المجلد (٤) العدد (١).

ثانياً: المصادر العربية المترجمة

1- Abbas, Tahani Mahdi, Hamid, Samira Mazhar & Naif, Qutaiba Nabeel (2010). *Using Cluster Analysis in Analyzing the Factors Affecting Heart Disease.* Journal of Al-Nahrain University, Vol. 13, No. (3), September 2010, pp. 58–65.

2- Ashour, Wafaa Abdul-Samad & Sultan, Wafaa Ali (2020). *Classification of the Financial Situation of the National Insurance Company for the Period (2008–2015) Using Cluster Analysis.* Rafidain Journal of Computer Sciences and Mathematics, Vol. (14), No. (2).

3- Hanish, Ibrahim Suleiman & Ismiyyo, Khulood Suleiman (2019). *Comparison between Hierarchical Clustering Methods and Selecting the Best with a Practical Application on Some Types of Milk Sold in Misurata City.* Special Issue of the Third Annual Conference on Theories and Applications of Basic and Life Sciences, September 7, 2019. Journal of Sciences.

4- Hussein, Asia Hammoud (2021). *Classification of Iraqi Provinces through the Factors Affecting Maternal Mortality at Birth Using Some Hierarchical Methods of Cluster Analysis.* Journal of Administration and Economics, No. 130, December 2021, pp. 213–226.

5- Kazem, Ibrahim Jawad & Mohammed, Ahmed Sultan (2015). *Classification and Evaluation of the Deprivation Index in Iraq for 2009 Using Some Cluster Analysis Methods.* Journal of Economic and Administrative Sciences, Vol. 21, No. 82, pp. 291–411.

6- Khmou, Khulood Yousif & Mohsen, Omar Mohi (2016). *Factors Affecting Congenital Malformations in Iraq for the Year 2012.* Journal of Kirkuk University for Administrative and Economic Sciences, Vol. (6), No. (1).

7- Khmou, Khulood Yousif & Raouf, Tamara (2017). *The Spread of the Phenomenon of Private Tutoring in Iraqi Institutions: Causes and Factors.* Journal of Kirkuk University for Administrative and Economic Sciences, Vol. (7), No. (1).

8- Nameq, Faisal Naji (2014). *Using Cluster Analysis to Classify Cancer According to Types and Cases in Iraq for the Period (1998–2006).* Journal of Kirkuk University for Administrative and Economic Sciences, Vol. (4), No. (1).

9- Shuaibth, Layla Jasim (2017). *A Study of the Effect of Smoking, Gender, and Age on Blood Clotting Speed and the Incidence of Thrombosis.* Al-Qadisiyah Journal for Pure Sciences, Vol. 22, No. 3.

ثالثاً: المصادر الاجنبية

1- Bu, J., Liu, W., Pan, Z., & Ling, K. (2020). Comparative study of hydro chemical classification based on different hierarchical cluster analysis methods. International Journal of Environmental Research and Public Health, 17(24), 1–23. https://doi.org/10.3390/ijerph17249515.
2- Jarman, A. M. (2020). Hierarchical Cluster Analysis: Comparison of Single Linkage, Complete Linkage, Average Linkage and Centroid Linkage. International Journal of Georgia Soutern University, 1(1), 1–13. https://www.researchgate.net/publication/339443595.
3- Mohammed, M. I., Al-Zubaidy, K. M., & Al Obaidi, B. M.. (2021). Determination of Genetic Diversity Depending on Quantitative Characters among Genotypes of Triticale (X Trititcosecale Wittmac) Using Cluster Analysis. IOP Conference Series: Earth and Environmental Science, 910(1). https://doi.org/10.1088/1755-1315/910/1/012050.
4- Mohammed,S. N. & Jobouri, K. K. A. (2023). Cluster analysis for genotypes of bread wheat (Triticum aestivum L.). HIV Nursing, 23(1), 947–951.
5- Saqr, M. (2024). Learning Analytics Methods and Tutorials. In Learning Analytics Methods and Tutorials. https://doi.org/10.1007/978-3-031-54464-4.
6- Shetty, P., & Singh, S. (2021). Hierarchical Clustering: A Survey. International Journal of Applied Research, 7(4), 178–181. https://doi.org/10.22271/allresearch.2021.v7.i4c.8484.
7- Wierzchoń, S. T., & Kłopotek, M. A. (2018). Modern Algorithms of Cluster Analysis. In Studies in Big Data (Vol. 34). https://doi.org/10.1007/978-3-319-69308-8.
8- Yim, O., & Ramdeen, K. T. (2015). Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. The Quantitative Methods for Psychology, 11(1), 8–21. https://doi.org/10.20982/tqmp.11.1.p008.