مقارنة أساليب التكتل الهرمى
واختيار أفضل الطرق لتنميط مجموعة
من المتغيرات التى تسهم فى خروج الطفل إلى الشارع*

عبير صالح**

تهدف هذه الدراسة إلى مقارنة الطرق المختلفة لربط التجمعات الجزئية وفقًا لطريقة التكتل الهرمى. وسوف تعتمد المقارنة على معامل ارتباط Cophenetic، بهدف تحديد الطريقة الأفضل لتنميط مجموعة من المتغيرات التي تسهم فى خروج الطفل إلى الشارع ومن ثم تعرضه للتعامل مع المخدرات سواء فى صورة تعاطى أو توزيع أو اتجار. وقد انتهت الدراسة من واقع البيانات الخاصة بأطفال الشوارع الذين يتعاملون مع المخدرات فى صورها المختلفة، وعلى وجه الخصوص الأطفال الموجودون وقت الدراسة فى المؤسسات الحكومية والجمعيات الأهلية وأيضًا المقيمون بالشارع إلى وجود أربعة عوامل رئيسة تؤدى إلى ظهور وتنامى هذه المشكلة وهى: تأثير النظراء، والتفكك الأسرى، والقسوة والعنف وإيذاء الطفل، وأخيرًا البيئة السكنية للطفل. هذه العوامل مجتمعة تدفع الطفل للهروب من البيت حيث يعتبر الشارع فى هذه الحالة من عوامل الجذب.

مقدمة

كثيرًا ما يواجه الباحث أعدادًا كبيرة من المشاهدات الخاصة بظاهرة ما، ويكون من الصعوبة دراستها ما لم يتم وضعها فى مجموعات متجانسة أو أنماط مميزة يمكن التعامل معها بوصفها وحدات وبالتالى دراستها بصورة مبسطة. وتحليل التجمعات Cluster Analysis هو الأسلوب الإحصائى المتقدم الذى يحقق هذا

---

* تم إعداد هذه المقالة فى إطار بحث" أطفال الشوارع وجرائم المخدرات"، الذى قام به المركز القومى للبحوث الاجتماعية والجنائية بالتعاون مع المجلس القومى لمكافحة وعلاج الإدمان، إشراف الأستاذة الدكتورة عزة كريم، وعضوية كل من الأستاذة الدكتورة سهير عبدالمنعم الباحث الرئيسى، والأستاذة الدكتورة سعاد عبدالرحيم، والأستاذة الدكتورة ماجدة عبدالغنى، ومنى بدير، ووفاء الشعار، ومريم سرور.

** خبير إحصاء بالمركز القومى للبحوث الاجتماعية والجنائية.

الهدف؛ حيث يقوم بتجميع مفردات أو تجميع متغيرات في عدد من التجمعات Clusters المحدودة، وغير المعروفة مسبقًا، وأيضًا المتباينة فيما بينها والمتجانسة داخليًا، وفى هذه الدراسة سوف نهتم بتحليل التجمعات للمتغيرات وليس للمفردات.

ولتحليل التجمعات تطبيقات متعددة منها على سبيل المثال، فى مجال التخطيط الإقليمي للمجتمع، تم استخدامه فى تنميط مناطق المجتمع تبعًا لأولويات الرعاية الصحية، بهدف توزيع مراكز الرعاية الصحية على مناطق المجتمع الأكثر احتياجًا[١]. أيضًا فى مجال الزراعة، تم استخدامه لدراسة التفاوت في إمكانيات التنمية الزراعية في محافظات مصر[٢]. وفى تنميط الاختلافات المظهرية والمحصولية لبعض أصناف النخيل المزروعة بالمملكة العربية السعودية[٣]. وفى مجال العلوم الاجتماعية، تم استخدامه في تنميط أهم استراتيجيات مصادر العيش لدى الأسر السورية الفقيرة والميسورة الحال[٤]. وفى تنميط التفضيلات والاحتياجات الثقافية للشباب المتردد على معرض الكتاب المصرى[٥]، وأيضًا تنميط المحافظات المصرية وفقًا للنشاط الإجرامي للمسجل خطر[٦]. وفيما يتعلق بالبيئة، تم استخدامه لتنميط الخصائص البيئية والعمرانية للتجمعات الريفية فى المملكة الأردنية الهاشمية[٧].

ويبدأ دور تحليل التجمعات بعد تحديد متغيرات الدراسة المراد تصنيفها في تجمعات، ومن هذه البداية يتم حساب معاملات التشابه أو المسافة بين أزواج المتغيرات المختلفة. ولمفهوم المسافة والتشابه أهمية خاصة فى تحليل التجمعات، حيث إن المسافة Distance مقياس لمدى بُعد متغيرين عن بعضهما، أما التشابه Similarity فهو مقياس لمدى قرب متغيرين من

بعضهما. وتختلف المعاملات باختلاف طبيعة المتغيرات؛ حيث توجد معاملات تستخدم فى حالة البيانات المتصلة، ومعاملات تستخدم فى حالة البيانات التكرارية، ومعاملات أخرى تستخدم للبيانات الثنائية وهى المتغيرات محل الدراسة الحالية.

وتفترض المرحلة التالية في تحليل التجمعات أن كل متغير من متغيرات الدراسة يصف مجموعة جزئية خاصة به ويتم ضم هذه التجمعات الجزئية فى تجمعات أكبر بالاعتماد على المسافات (أو التشابه) بينها[٨]، وذلك بناءً على طريقتين للضم: طريقة المتوسطات k-means Cluster، وطريقة التجميع الهرمى Hierarchical Cluster؛ حيث تتطلب طريقة المتوسطات معرفة عدد التجمعات مسبقًا، بينما طريقة التجميع الهرمى تتطلب تصنيف متغيرات من النوع نفسه، ثنائية التصنيف مثلا، ولا تحتاج لمعرفة عدد التجمعات مسبقا، ولذا فسوف تستخدم هذه الطريقة فى الدراسة، إذ أنه فى طريقة التجميع الهرمى يتم ضم التجمعات الجزئية فى تجمعات أكبر من خلال أسلوبين يعرف الأول باسم أسلوب التوزيع Divisive، والثانى بأسلوب التكتل Agglomerative. وطريقة التكتل الهرمي هى أكثر الطرق شيوعًا واستخدامًا فى تكوين التجمعات، وسوف نتبناها في هذه الدراسة[٩].

ووفقًا لطريقة التكتل الهرمى يتم ضم أى اثنين من التجمعات الجزئية ومن ثم ربطهما معًا فى تجمع أكبر بناءً على المسافة بينهما، وتوجد طرق عديدة لحساب هذه المسافة تختلف باختلاف طبيعة المتغيرات المكونة لهذه التجمعات، وهذه الطرق فى حالة بيانات الدراسة وهى البيانات ثنائية التصنيف: طريقة الربط المفرد، والربط الكامل، ومتوسط الربط سواء الربط بين

المجموعات أو الربط داخل المجموعات[١٠]. وهذه الطرق المختلفة مبررة رياضيًا ويمكن استخدام أى منها فى تكوين التجمعات ولكنها تعطى نتائج مختلفة، ومهمة اختيار الطريقة المستخدمة فى التحليل ليست عملية مباشرة، إذ يتعين اللجوء إلى معيار للمفاضلة بينهم. وفى حالة التكتل الهرمي يطبق معامل ارتباط Cophenetic، حيث كلما زادت قيمة المعامل كانت مؤشرًا على جودة الطريقة المستخدمة فى عملية التكتل.

لذلك تهدف هذه الورقة إلى دراسة ومقارنة تناقض الأربعة طرق المختلفة من طرق ربط التجمعات الجزئية وفقًا لطريقة التكتل الهرمى، وسوف تعتمد المقارنة على معامل ارتباط Cophenetic بهدف تحديد الطريقة الأفضل لتنميط مجموعة من المتغيرات التى تسهم فى خروج الطفل إلى الشارع ومن ثم تعرضه للتعامل مع المخدرات سواء فى صورة تعاطى أو توزيع أو اتجار، ولتحقيق هذا الهدف سوف تتناول الورقة النقاط التالية:

أولًا: تكوين التجمعات.

ثانيًا: أساليب ضم التجمعات الجزئية فى طريقة التكتل الهرمى.

ثالثًا: معايير جودة تحليل التجمعات.

رابعًا: تطبيق تحليل التجمعات.

أولًا: تكوين التجمعات

بعد حساب مقاييس المسافة أو التشابه بين أزواج المتغيرات المختلفة، يتم ضم هذه المتغيرات فى تجمعات جزئية أكبر بالاعتماد على هذه المسافات. ومن طرق ضم المتغيرات: طريقة التجميع الهرمى، وطريقة المتوسطات وهما على النحو التالى.

١ - طريقة التجميع الهرمى Hierarchical Clustering

طريقة التجميع الهرمي هي أكثر طرق تكوين التجمعات استخدامًا، وهى لا تتطلب معرفة مسبقة بعدد التجمعات. ويتم هذا التجميع من خلال أسلوبين يعرف الأول باسم أسلوب التكتل Agglomerative، والثانى باسم أسلوب التوزيع Divisive. وهما كما يلى[١١]:

أ – أسلوب التكتل الهرمى Agglomerative Hierarchical Technique

ويجرى أسلوب التكتل الهرمي وفقًا للخطوات التالية:

- افتراض أن كل متغير هو مجموعة جزئية، ثم يتم تكوين مصفوفة متماثلة تسمى مصفوفة معاملات التقارب Proximities Matrix ، تمثل قيمها مقاييس المسافة أو الاختلاف Measures of Distances or Dissimilarity بين أزواج التجمعات الجزئية المختلفة.
- بناءً على معاملات المصفوفة السابقة، يتم تجميع مجموعتين جزئيتين لتكوين أول تجمع في مجموعة جزئية أكبر. وكما أن هناك طرقًا عديدة لحساب المسافات أو الاختلاف بين المتغيرات، توجد أيضًا طرق عديدة لحساب المسافة أو الاختلاف بين أى اثنين من التجمعات سوف نعرضهم لاحقًا.
- إعادة حساب مصفوفة التشابه (أو الاختلاف) بين التجمع الجديد وباقى التجمعات.
- يتم تكرار هذه العملية وعليه إما أن تُضم مفردتان معًا لتكون تجمع آخر، أو تضم مفردة إلى التجمع الناتج من الخطوة الثانية لينتج تجمعًا أكبر يتكون من ثلاث مفردات وهكذا، وبالطريقة نفسها فى الخطوات

المتتالية تتم عملية الضم للخروج بتجمع جديد أو تجمع أكبر لحين الوصول إلى مجموعة جزئية واحدة تشمل جميع المتغيرات.

ب - أسلوب التوزيع Divisive Technique

يبدأ هذا الأسلوب بافتراض وجود مجموعة واحدة للمتغيرات يتم تقسيمها إلى تجمعين جزئيين، وهذان التجمعان الجزئيان يتم أيضًا تقسيمهما إلى مجموعات جزئية أصغر، وهكذا حتى تتكون لكل متغير مجموعة خاصة به، أو حتى نصل إلى عدد التجمعات المطلوبة. ويتم تجزيء التجمع إلى تجمعين بناءً على عدة طرق منها ما يعرف Bisecting K-means والتى تتم وفق الخطوات التالية:

الخطوة الأولى: اعتبار وجود مجموعة واحدة للمتغيرات ولتكن M حيث $M = [X_1, X_2, ..., X_N]$، وتحديد مركزها وليكن W. لتقسيمها إلى تجمعين M_L و M_R نختار قيمة مبدئية لمركز التجمع L بطريقة عشوائية وليكن C_L. ثم نحسب مركز التجمع C_R كالآتى:

$$C_R = W - (C_L - W).$$

الخطوة الثانية: توزيع المتغيرات على التجمعين السابقين بناءً على القاعدة التالية:

$$\begin{cases} X_I \in M_L & if \quad \|X_I - C_L\| \leq \|X_I - C_R\| \\ X_I \in M_R & if \quad \|X_I - C_L\| \succ \|X_I - C_R\| \end{cases}$$

الخطوة الثالثة: إعادة حساب متوسطات التجمعات W_L, W_R مرة أخرى.

الخطوة الرابعة: تكرار هذه العملية حتى يتحقق الشرط التالى:

$$C_L = W_L, C_R = W_R.$$

ويتم اختيار التجمع الواجب تقسيمه وفقًا لعدة معايير منها التجمع الأكثر عددًا، أو التجمع صاحب أعلى تباين بالنسبة لمتوسطه.

٢ - **طريقة المتوسطات** K-Means Clustering

تعتمد هذه الطريقة على تحديد عدد التجمعات مسبقًا وليكن عددها k ، بعد ذلك يتم ضم كل متغير إلى التجمع وفقًا لأقل مسافة بينه وبين متوسطات التجمعات المختلفة. وتتم طريقة المتوسطات تبعًا للخطوات الآتية[١٢]:

- اختيار قيم مبدئية لمراكز التجمعات بطريقة عشوائية.
- توزيع المتغيرات على التجمعات بناءً على أقل مسافة بينها وبين مراكز التجمعات.
- إعادة حساب متوسطات التجمعات مرة أخرى.
- إعادة خلط المتغيرات وتوزيعها من جديد.
- تكرار هذه العملية حتى تستقر المتغيرات فى تجمعات معينة، وهو يحدث عندما تصبح الزيادة فى مراكز التجمعات المحسوبة المتتالية طفيفة.

وبهذا تختلف هذه الطريقة عن طريقة التجميع الهرمى فى أنها لا تحتاج إلى حساب مقاييس التشابه أو المسافة بين أزواج المتغيرات المختلفة وفى المقابل تتطلب معرفة عدد التجمعات مسبقًا، كما أن نتائجها تتأثر باختيار القيم المبدئية لمراكز التجمعات وكذلك بالقيم المتطرفة فى البيانات، علاوة على أن المتغير الواحد يمكن ضمه لأكثر من تجمع أثناء عملية التحليل على عكس طريقة التجميع الهرمى، أيضًا لا تعطى الطريقة عددًا من التجمعات الممكنة أو

مدى من الحلول ما لم تتم إعادة التحليل مرة أخرى وهو ما تقدمه طريقة التجميع الهرمى.

ولذا فسوف تستخدم طريقة التجميع الهرمى بأسلوب التكتل الهرمى فى هذه الدراسة.

ثانياً: أساليب ضم التجمعات الجزئية فى التكتل الهرمى

فى تجمع التكتل الهرمى، يتم ضم أي عنصرين من المتغيرات أو (التجمعات) ولتكن A و B وفقًا لمقاييس التشابه أو المسافة بينهما. وتوجد عدة طرق للربط خاصة بالمتغيرات الثنائية موضع الدراسة وهى على النحو التالى:

١ - الربط المفرد[١٤] Nearest neighbor (Single linkage)

تحسب المسافة بين التجمعين A و B بناء على المسافة بين أقرب عضوين x و y من أعضاء التجمعين، كما يلى:

$$\min\{\, d(x,y) : x \in A, y \in B \,\}.$$

٢ - الربط الكامل[١٥] Furthest neighbor (Complete linkage)

تحسب المسافة بين التجمعين A و B بناء على المسافة بين أبعد عضوين x و y من أعضاء التجمعين، كما هو موضح:

$$\max\{\, d(x,y) : x \in A, y \in B \,\}.$$

٣ - متوسط الربط بين المجموعات[١٦] Average linkage between Groups

تحسب المسافة بين التجمعين A و B بناء على متوسط المسافات بين كل الأزواج الممكنة من أعضاء التجمعين.

٤ - **متوسط الربط داخل المجموعات**[١٧] Average linkage within groups

تحسب المسافة بين التجمعين A و B بناء على متوسط المسافات بين كل الأزواج الممكنة من أعضاء التجمع الناتج.

ويتضح أن متوسط الربط بنوعيه يستفيد من كل المسافات البينية بين كل زوجين من المتغيرات أعضاء التجمعين ولا يعتمد على مسافة واحدة ويهمل باقي المعلومات عن المسافات الأخرى. وسوف تهتم الدراسة الحالية بالمقارنة بين هذه الطرق الأربع واختيار أفضل طريقة من خلال معيار الجودة المناسب.

أما باقي الطرق[١٨] مثل Ward's Method, Centroid Method و Median Method فهي طرق تستخدم Squared Euclidean Distances وهو خاص بالمتغيرات المتصلة، وبالتالى فلن نتعرض لها.

ثالثاً: معايير جودة تحليل التجمعات

يقوم تحليل التجمعات بتصنيف البيانات (مفردات أو متغيرات) بصرف النظر عن مدى قابلية البيانات لعملية التصنيف، فحتى لو كانت تلك البيانات لا يجمعها أي نمط حقيقي فسوف يقوم التحليل بتصنيفها، وبالتالي يصبح لزامًا أن يكون تقييم جودة تحليل التجمعات جزءًا من عملية التحليل نفسها. وتوجد ثلاث طرق للحكم على جودة توفيق نتائج تحليل التجمعات[١٩]، وهى كما يلى.

١ - **معيارا التماسك أو الانفصال** Cohesion or Separation Coefficients

يتم الحكم على مدى جودة التصنيف بناءً على معيارين هما: معيار التماسك Cohesion وهو يعبر عن مدى تماسك كل تجمع أو تشابه مفرداته، أو معيار الانفصال Separation وهو يعبر عن مدى اختلاف أو انفصال كل تجمع عن

التجمعات الأخرى، ويمكن حساب كل معيار بناءً على مصفوفة التماثل أو الاختلاف كما يلى[٢٠]: بفرض أن نتائج إجراء تحليل التصنيف كان عدد k من التجمعات هى C_i، وكان m_i يمثل حجم كل تجمع (عدد المتغيرات التى يضمهم معًا)، حيث $i = 1,2,..,k$، فإن:

$$cohesion\ (C_i) = \sum_{\substack{x \in C_i, \\ y \in C_i}} proximity\ (x, y),$$

$$separation\ (C_i, C_j) = \sum_{\substack{x \in C_i, \\ y \in C_j}} proximity\ (x, y).$$

تقاس جودة النتائج الإجمالية لتحليل التصنيف كما يلى:

$$overall\ validity = \sum_{i=1}^{k} w_i\ validity\ (C_i),$$

حيث

$$validity\ (C_i) = cohesion\ (C_i)$$

$$or \qquad\qquad = separation\ (C_i, C_j).$$

و w_i هى الأوزان أما تساوى الواحد أو مقلوب حجم التجمع $1/m_i$.

٢ – معيار Silhouette

وهو معيار يجمع بين معيارى التماسك والانفصال معًا، ويحسب كالآتى: بفرض المتغير i فى التجمع A فإن معيار Silhouette للمتغير i، S_i يعرف كالتالى:

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}}$$

حيث $a(i)$ هو متوسط الاختلافات بين المتغير i و باقى المتغيرات فى التجمع A وهو التجمع الذى ينتمى إليه المتغير، و $b(i)$ متوسط الاختلافات بين المتغير i والمتغيرات فى التجمع B وهو التجمع الأقرب للتجمع A. وبناءً على القيم التى يأخذها هذا المعيار والتى تتراوح ما بين (-١ و ١)، يمكن الحكم على جودة التصنيف، حيث القيمة السالبة وهى غير مرغوب بها لأنها تعنى أن $(a_i \succ b_i)$ بمعنى أن الاختلافات بين المتغير i وباقى المتغيرات فى التجمع التى تنتمى إليه أكبر من الاختلافات بينها وبين المتغيرات فى التجمعات الأخرى، بينما القيمة الموجبة التى تقترب من (١) فهى معيار على جودة توفيق نموذج تحليل التجمعات. ويحسب متوسط هذا المعيار لكل تجمع بأخذ متوسطات المعيار لكل متغير ينتمى إليه، وبتكرار هذا الحساب عبر التجمعات المختلفة الناتجة يكون معيارًا إجماليًّا للحكم على جودة التصنيف[٢١].

٣ - معامل ارتباط Cophenetic

يُطبق المعياران السابقان فى حالة معرفة عدد التجمعات مسبقًا، أما فى حالة التكتل الهرمى فيتم حساب معامل ارتباط Cophenetic[٢٢] وهذا المعامل يقيس الارتباط بين مقاييس المسافة أو التشابه بين أى اثنين من المتغيرات كما جاء فى مصفوفة التشابه الذى تم على أساسه ربط المتغيرات وفقًا لشكل Dendrogram. ولحساب هذا المعامل سوف نفترض أن المتغيرات موضع الاهتمام هى $\{X_i\}$ وأن التمثيل الشجرى الناتج بعد إجراء تحليل التجمعات هو $\{T_i\}$، يعرف معامل ارتباط Cophenetic بالمعادلة التالية:

$$C = \frac{\sum_{i<j}(x(i,j)-x)(t(i,j)-t)}{\sqrt{\sum_{i<j}[(x(i,j)-x)^2][\sum_{i(j}(t(i,j)-t)^2]}}$$

حيث $x(i,j)=|X_i - X_j|$ هى المسافة بين المتغير رقم i والمتغير رقم j وفقًا لمصفوفة التشابه، و $t(i,j)$ هى أقل مسافة تم ربط المتغير رقم i والمتغير رقم j معًا وفقًا لشكل Dendrogram. وجدير بالذكر أنه كلما زادت قيمة المعامل كان هذا مؤشرًا على جودة الطريقة المستخدمة فى عملية التكتل.

رابعاً: تطبيق تحليل التجمعات

من المشكلات والأخطار التى تواجه أطفال الشوارع فى مصر تعرضهم للعنف وتعاطى مواد الإدمان والاستغلال من تجار المخدرات والمجموعات الإجرامية المنظمة؛ حيث تحترف هذه العصابات استقطاب الوافدين الجدد من الأطفال واستغلال صغر سنهم وإرهابهم لاستخدامهم باعتبارهم أدوات مساعدة فى ترويج وتوزيع المخدرات، مع إمكانية تعرضهم للقتل للتخلص من آثار تلك الجرائم. ومن هنا كانت أهمية التعرف على العوامل التى تؤدى إلى ظهور وتنامى المشكلة كما جاءت فى البحث الميدانى، ونظرًا لتعدد هذه العوامل يكون من الأفضل تلخيصها فى عدد أقل من المعلومات التى تساعدنا على فهم ما يدفعهم إلى العيش فى الشوارع. ولتحقيق ذلك تم تطبيق الأسلوب الإحصائى "تحليل التجمعات" وتمت خطوات التحليل على النحو الآتى:

١ - تحديد متغيرات الدراسة

تم الاعتماد على البيانات الميدانية التى تم جمعها من أفراد العينة محل الدراسة والبالغ عددهم (٤٠٠) طفل باستخدام أداة البحث، والمتمثلة فى استمارة استبار، تم تطبيقها على عينة من أطفال الشوارع الذين لهم علاقة بالمخدرات والموجودين فى المؤسسات الحكومية والجمعيات الأهلية بكل من محافظتى القاهرة والجيزة التي تتعامل مع أطفال الشوارع، وأيضًا تم التطبيق على بعض من الأطفال الموجودين بالشارع. وفى ضوء هدف هذه الدراسة تم تحديد الأسئلة أو المتغيرات على النحو التالى:

١. الحالة التعليمية للطفل: أمي	(Q5)
٢. الخروج للشارع في سن صغيرة (قبل سن التعليم الأساسي)	(Q7)
٣. يعيش في أسرة غير مكتملة	(Q10)
٤. يعيش مع أسرة عدد أفرادها أكبر من ٦ أفراد	(Q11)
٥. غياب الأب (سواء بالوفاة أو لا يُعلم عنه شيء)	(Q14.1)
٦. غياب الأم (سواء بالوفاة أو لا يُعلم عنها شيء)	(Q14.2)
٧. وجود زوجة أب	(Q15.1)
٨. وجود زوج أم	(Q15.2)
٩. الحالة التعليمية للأب: أمي	(Q16.1)
١٠. الحالة التعليمية للأم: أمية	(Q16.2)
١١. وجود مشكلات مع الأهل	(Q24)
١٢. سوء معاملة الأب	(Q29.1)
١٣. سوء معاملة الأم	(Q29.2)
١٤. سوء معاملة زوجة الأب	(Q29.3)
١٥. سوء معاملة زوج الأم	(Q29.4)
١٦. الخلافات مع الإخوة	(Q29.5)
١٧. التشجيع من الغير على ترك البيت	(Q29.6)
١٨. الشقة غير مستقلة	(Q31)
١٩. لا توجد مجارٍ	(Q32.4)
٢٠. سمع عن المخدرات من الأصحاب	(Q37.1)
٢١. يحب شرب المخدرات مع الأصحاب في الشارع	(Q45.2)
٢٢. يحصل على المخدرات من الأصحاب	(Q46)
٢٣. أحد الأصحاب هو أول من علمه أخذ المخدرات	(Q48)

٢- مصفوفة معاملات التماثل

تلى مرحلة تحديد المتغيرات وطبيعتها، مرحلة قياس درجة التماثل أو المسافة بين كل متغيرين من متغيرات الدراسة، وتختلف هذه المقاييس باختلاف طبيعة المتغيرات، وتعتمد فى حالة البيانات الثنائية وهى المتغيرات محل الدراسة على معاملات الاقتران المحسوبة من جداول الاقتران ٢×٢ وهو يوضح التوزيع المشترك لعينة الدراسة وفقًا لهذين المتغيرين وذلك عندما يكون كلّ من المتغيرين له صفتان فقط (1, 0) على النحو التالى:

المتغير الأول	المتغير الثانى	
	1	0
1	A	B
0	C	D

وتعبر A عن عدد الحالات التي يكون فيها المتغير الأول والمتغير الثانى = 1، وتعبر B عن عدد الحالات التى يكون فيها المتغير الأول = 1 والمتغير الثانى = 0 ، وأما C فتعبر عن عدد الحالات التى يكون فيها المتغير الأول = 0 والمتغير الثانى = 1، وأخيرًا تعبر D هن عدد الحالات التى يكون فيها المتغيران الأول والثانى = 0.

وتختلف مقاييس التشابه فيما بينها من حيث الوزن النسبي الذي يعطيه كل مقياس لكل خلية من خلايا الجدول الأربع[٢٣]، ومن هذه المقاييس مقياس Simple Matching Coefficient، وهو كالاتى: $\frac{A+D}{A+B+C+D}$ وهو يعطى أوزانًا متساوية لكل خلية، وقد تم اختياره لأنه لا يوجد سبب مسبق يدفعنا

لإعطاء أوزان نسبية لبعض الخلايا أكبر من الوزن النسبى المعطى للخلايا الأخرى.

وبناءً على مقياس التشابه المختار كانت مصفوفة معاملات التشابه الموضحة فى الجدول رقم (١) وذلك باستخدام حزمة البرامج[٢٤] SPSS.

جدول رقم (١)

Similarity Coefficient Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	.663																					
3	.533	.470																				
4	.588	.670	.480																			
5	.515	.473	.658	.343																		
6	.558	.470	.680	.315	.693																	
7	.448	.385	.660	.340	.433	.695																
8	.445	.393	.638	.308	.680	.448	.673															
9	.503	.405	.390	.380	.393	.530	.660	.548														
10	.548	.413	.388	.488	.510	.288	.468	.635	.643													
11	.438	.575	.625	.745	.373	.410	.455	.398	.430	.473												
12	.483	.390	.455	.300	.498	.565	.680	.638	.660	.563	.415											
13	.478	.365	.420	.210	.673	.640	.635	.703	.665	.568	.290	.770										
14	.463	.345	.530	.230	.608	.700	.790	.708	.690	.543	.325	.785	.825									
15	.463	.345	.495	.200	.723	.650	.675	.793	.640	.608	.295	.720	.885	.835								
16	.453	.370	.475	.275	.678	.665	.620	.653	.600	.543	.350	.730	.820	.800	.805							
17	.553	.370	.475	.280	.598	.605	.645	.653	.640	.583	.330	.715	.760	.765	.760	.720						
18	.490	.515	.585	.450	.678	.620	.505	.548	.530	.578	.490	.550	.590	.590	.590	.570	.590					
19	.573	.433	.573	.398	.650	.638	.553	.570	.548	.520	.473	.598	.628	.608	.633	.643	.578	.703				
20	.615	.630	.565	.775	.358	.370	.375	.323	.400	.453	.735	.300	.220	.240	.220	.235	.305	.445	.418			
21	.503	.623	.558	.738	.395	.423	.368	.340	.408	.450	.688	.298	.238	.243	.258	.298	.313	.448	.440	.793		
22	.503	.580	.620	.695	.403	.380	.395	.388	.380	.458	.690	.370	.250	.255	.235	.270	.360	.510	.448	.780	.723	
23	.575	.668	.563	.793	.355	.368	.348	.320	.378	.425	.733	.288	.188	.218	.193	.228	.273	.433	.415	.858	.830	.798

٣ - المقارنة بين أساليب الربط واختيار الأسلوب الأفضل

المرحلة الثالثة فى تحليل التجمعات، هى توزيع المتغيرات فى تجمعات بالاعتماد على معاملات التشابه فى مصفوفة التشابه، وهنا يوجد عدد من الخيارات اخترنا منها طريقة التكتل الهرمى، التى لا تتطلب معرفة مسبقة بعدد المجموعات، وفى إطار هذه الطريقة هناك أربعة أساليب للربط بين المجموعات الجزئية فى كل خطوة من خطوات طريقة التجميع، وهى ربط مفرد، وربط كامل، ومتوسط الربط بين المجموعات، ومتوسط الربط داخل المجموعات، وهى لا تعطى النتائج نفسها كما أنه لا توجد طريقة بينها أفضل من الطرق الأخرى، وبالتالى يتعين للمفاضلة بينها اللجوء إلى حساب معامل ارتباط Cophenetic والذى يتم تبعًا للخطوات التالية:

أ - تكوين مصفوفة معاملات التشابه بناءً على مقياس Simple Matching Coefficient.

ب - تكوين مصفوفة أخرى مناظرة لكل طريقة من طرق الربط الأربع، وهى مصفوفة متماثلة تسمى: Cophenetic Distance Matrix، تعبر عناصرها عن معامل التشابه بين كل مفردتين، ويقصد بالتشابه فى هذه المصفوفة المعامل الذى تم على أساسه ضمهما لأول مرة فى تجمع جزئى وفقًا لشكل Dendrogram والذى سيتم عرضه لاحقًا، وتوضح الجداول من رقم ٢ إلى رقم ٥ هذه المصفوفات الأربع المناظرة لكل أسلوب من أساليب الربط وهى ربط مفرد وربط كامل ومتوسط الربط بين المجموعات ومتوسط الربط داخل المجموعات.

جـ - حساب معامل الارتباط بين المصفوفتين.

جدول رقم (٢)

Cophenetic Distance Matrix for Single Linkage Method

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	.663																					
3	.625	.625																				
4	.663	.670	.625																			
5	.625	.625	.680	.625																		
6	.625	.625	.680	.625	.7																	
7	.625	.625	.680	.625	.723	.7																
8	.625	.625	.680	.625	.723	.7	.790															
9	.625	.625	.680	.625	.690	.690	.690	.690														
10	.625	.625	.680	.625	.643	.643	.643	.643	.643													
11	.625	.625	.680	.625	.648	.678	.678	.678	.678	.643												
12	.625	.625	.680	.745	.625	.625	.625	.625	.625	.643	.678											
13	.625	.625	.680	.625	.723	.7	.765	.723	.690	.643	.678	.625										
14	.625	.625	.680	.625	.625	.7	.785	.765	.690	.643	.643	.625	.765									
15	.625	.625	.680	.625	.625	.7	.785	.7	.690	.643	.678	.625	.765	.785								
16	.625	.625	.680	.625	.643	.7	.790	.7	.690	.643	.678	.625	.765	.785	.835							
17	.625	.625	.680	.625	.625	.7	.790	.7	.690	.643	.643	.625	.765	.785	.835	.833						
18	.625	.625	.680	.625	.643	.7	.790	.7	.690	.643	.678	.625	.765	.785	.835	.820	.820					
19	.625	.625	.680	.625	.678	.703	.703	.678	.703	.643	.703	.625	.703	.703	.703	.703	.703	.703				
20	.663	.670	.625	.793	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625			
21	.663	.670	.625	.793	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.663		
22	.663	.670	.625	.793	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.663	.663	
23	.663	.670	.625	.793	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.625	.858	.885	.633

جدول رقم (٣)

Cophenetic Distance Matrix for Complete linkage Method

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	.663																					
3	.308	.308																				
4	.308	.308	.308																			
5	.433	.433	.433	.433																		
6	.433	.433	.433	.433	.433																	
7	.433	.433	.433	.433	.433	.433																
8	.433	.433	.433	.433	.680	.433	.188															
9	.188	.188	.188	.188	.188	.188	.188	.520														
10	.188	.188	.188	.188	.188	.188	.188	.520	.520													
11	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520												
12	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308											
13	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308	.188										
14	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308	.188	.710									
15	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308	.188	.720	.720								
16	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308	.188	.720	.720	.825							
17	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308	.188	.720	.720	.885	.825						
18	.188	.188	.188	.188	.188	.188	.188	.520	.520	.520	.308	.188	.720	.720	.800	.800	.800					
19	.188	.188	..550	.703	.188	.188	.188	.705	.703	.703	.708	.188	.520	.520	.550	.550	.703	.550				
20	.308	.308	.308	.308	.433	.433	.433	.188	.188	.188	.188	.688	..188	..188	.188	.188	.703	.188	.188			
21	.308	.308	.308	.308	.433	.433	.433	.188	.188	.188	.188	.688	.188	.188	.188	.188	.188	.188	.188	.793		
22	.308	.308	.308	.308	.433	.433	.433	.188	.188	.188	.188	.688	.188	.188	.188	.188	.188	.188	.188	.723	.723	
23	.308	.308	.308	.308	.433	.433	.433	.188	.188	.188	.188	.688	.188	.188	.188	.188	.188	.188	.188	.723	.793	.723

جدول رقم (٤)

Cophenetic Distance Matrix for Between-Group linkage Method

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	.663																					
3	.383	.383																				
4	.595	.595	.383																			
5	.383	.383	.591	.383																		
6	.383	.383	.670	.383	.591																	
7	.383	.383	.670	.383	.591	.695																
8	.383	.383	.591	.383	.591	.591	.591															
9	.383	.383	.549	.383	.591	.591	.549	.549														
10	.383	.383	.549	.383	.591	.591	.549	.549	.545													
11	.383	.383	.549	.383	.664	.591	.591	.591	.383	.545												
12	.595	.383	.591	.718	.591	.383	.383	.383	.545	.545	.383											
13	.383	.595	.383	.383	.591	.591	.591	.691	.545	.545	.383	.383										
14	.383	.383	.591	.383	.591	.591	.591	.591	.545	.545	.591	.383	.744									
15	.383	.383	.591	.595	.591	.591	.591	.691	.545	.545	.591	.383	.751	.744								
16	.383	.383	.591	.383	.591	.591	.591	.691	.545	.545	.591	.383	.751	.744	.830							
17	.383	.383	.591	.383	.591	.591	.591	.691	.545	.545	.591	.383	.751	.744	.885	.830						
18	.383	.383	.591	.383	.591	.591	.591	.691	.545	.545	.591	.383	.751	.744	.808	.808	.808					
19	.383	.383	.591	.383	.664	.591	.591	.591	.545	.545	.703	.383	.591	.591	.591	.591	.591	.591				
20	.595	.595	.383	.768	.595	.383	.383	.383	.383	.383	.383	.718	.383	.383	.383	.383	.383	.383	.383			
21	.595	.595	.383	.768	.595	.383	.383	.383	.383	.383	.383	.718	.383	.383	.383	.383	.383	.383	.383	.811		
22	.595	.595	.383	.768	.595	.383	.383	.383	.383	.383	.383	.718	.383	.383	.383	.383	.383	.383	.383	.749	.749	
23	.595	.595	.383	.768	.595	.383	.383	.383	.383	.383	.383	.718	.383	.383	.383	.383	.383	.383	.383	.858	.749	.749

جدول رقم (٥)

Cophenetic Distance Matrix for Within-Group linkage Method

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2	.685																					
3	.515	.515																				
4	.685	.720	.515																			
5	.515	.515	.677	.515																		
6	.515	.515	.677	.622	.622																	
7	.515	.515	.622	.622	.622	.622																
8	.515	.515	.622	.622	.622	.622	.743															
9	.515	.515	.622	.622	.622	.622	.713	.713														
10	.515	.515	.622	.622	.622	.622	.685	.685	.685													
11	.515	.515	.622	.515	.515	.515	.652	.652	.652	.515												
12	.685	.720	.515	.622	.622	.622	.515	.515	.515	.652	.515											
13	.515	.515	.622	.622	.622	.622	.743	.743	.713	.652	.652	.515										
14	.515	.515	.622	.622	.622	.622	.743	.743	.713	.652	.652	.515	.780									
15	.515	.515	.622	.622	.622	.622	.743	.743	.713	.652	.652	.515	.798	.780								
16	.515	.515	.622	.622	.622	.622	.743	.743	.713	.652	.652	.515	.798	.780	.848							
17	.515	.515	.622	.622	.622	.622	.743	.743	.713	.652	.652	.515	.798	.780	.885	.848						
18	.515	.515	.622	.622	.622	.622	.743	.743	.713	.652	.652	.515	.798	.780	.828	.828	.828					
19	.515	.515	.622	.622	.622	.622	.703	.703	.713	.703	.703	.515	.703	.703	.703	.703	.703	.703				
20	.685	.72	.515	.622	.515	.515	.515	.515	.515	.515	.515	.758	.515	.515	.515	.515	.515	.515	.515			
21	.685	.72	.515	.622	.515	.515	.515	.515	.515	.515	.515	.758	.515	.515	.515	.515	.515	.515	.515	.827		
22	.685	.72	.515	.622	.515	.515	.515	.515	.515	.515	.515	.758	.515	.515	.515	.515	.515	.515	.515	.778	.778	
23	.685	.72	.515	.622	.515	.515	.515	.515	.515	.515	.515	.758	.515	.515	.515	.515	.515	.515	.515	.858	.827	.778

ويوضح جدول رقم (٦) معاملات الارتباط المناظرة لكل طريقة من طرق الربط التي تم استخدامها فى الدراسة. وقد أسفرت النتائج عن أن طريقة الربط داخل المجموعات هى صاحبة أعلى معامل ولذلك سوف نستخدمها فى تحليل التجمعات بطريقة التكتل الهرمى.

جدول رقم (٦)

الأسلوب	معامل ارتباط cophenetic
الربط المفرد	.696
الربط الكامل	.526
متوسط الربط بين التجمعات	.817
متوسط الربط داخل التجمعات	.846

وبالتالى تم تحليل التجمعات للمتغيرات باستخدام طريقة التكتل الهرمى بأسلوب الربط داخل التجمعات وكانت النتائج على النحو التالى.

٤ - تكوين التجمعات وفقًا لأسلوب الربط المختار وهو متوسط الربط داخل التجمعات

تتلخص نتائج تحليل التجمعات وفقًا لطريقة متوسط الربط داخل التجمعات فى بيانات الجدول رقم (٧)، والذى يشير إلى الخطوات المتتالية لما ينتج من التجمعات وأرقام المتغيرات (أو أرقام التجمعات) التى تندمج فى كل خطوة.

السطر الأول من الجدول يمثل الخطوة الأولى، حيث تم ضم المعاملة السيئة من الأم (رقم ١٣) والمعاملة السيئة من زوج الأم (رقم ١٥) فى تجمع واحد، وهذا ما يشير إليه العمودان الثانى والثالث من جهة اليسار والمعنون

بـ Cluster Combined، ويشير العمود الرابع تحت عنوان Coefficient إلى معامل التماثل بين هذين المتغيرين. ولأن هذه الخطوة هى الأولى فى خطوات التحليل، فإن المعامل بين هذين المتغيرين هو أكبر معامل فى جدول رقم (١) والخاص بمعاملات التماثل وهو (0.885) أما قيم العمود السابع فإنها تشير إلى رقم الخطوة التالية التى يتم فيها دمج متغير آخر أو تجمع آخر إلى التجمع الناتج فى الخطوة الحالية. فعلى سبيل المثال نجد أن قيمة هذا العمود فى السطر الأول هى ٣ ، وهذا يعنى أن الخطوة رقم ٣ قد تم ضم أحد المتغيرات، وهو المعاملة السيئة من زوجة الأب (رقم ١٤) إلى التجمع الأول ليصبح تجمعًا مكونًا من ٣ متغيرات. ثم نجد قيمة هذا العمود فى الخطوة (أو السطر) رقم ٣ هى القيمة ٤ فنجد فى الخطوة رقم ٤ انضمام الخلافات مع الإخوة (رقم ١٦) إلى هذا التجمع ليصبح تجمعًا مكونًا من ٤ متغيرات، وهكذا. أما قيم العمودين الخامس والسادس، فإنها تشير إلى أى من الخطوات السابقة التى تم فيها انضمام المتغيرين المشار إليهما فى العمودين الثانى والثالث فى تجمعات سابقة، وإذا كانت القيمة صفرًا فإنها تدل على أن المتغير منفرد ولم يتم ضمه إلى أى تجمع سابق.

جدول رقم (٧)

Agglomeration Schedule using within Average Linkage

Stage	Cluster combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	13	15	.885	0	0	3
2	20	23	.858	0	0	5
3	13	14	.848	1	0	4
4	13	16	.828	3	0	7
5	20	21	.827	2	0	6
6	4	20	.798	0	5	9
7	13	17	.798	4	0	8
8	12	13	.780	0	7	11
9	4	22	.778	6	0	10
10	4	11	.758	9	0	13
11	8	12	.754	0	8	12
12	7	8	.734	0	11	14
13	2	4	.720	0	10	17
14	7	9	.713	12	0	18
15	18	19	.703	0	0	20
16	5	6	.693	0	0	19
17	1	2	.685	0	13	22
18	7	10	.685	14	0	20
19	3	5	.677	0	16	21
20	7	18	.652	18	15	21
21	3	7	.622	19	20	22
22	1	3	.515	17	21	0

٥ - التمثيل الشجرى لخطوات تكوين التجمعات

الخطوة التالية من خطوات التحليل هى عرض الخطوات المتتالية لتكوين التجمعات باستخدام الرسم البيانى. وهناك نوعان من التمثيل البيانى، الأول هو التمثيل البيانى الأفقى للتجمعات Vertical Icicle Complete، والثانى هو التمثيل الشجرى للتجمعات Dendrogram. ويتميز التمثيل الشجرى على التمثيل الأفقى فى أنه يأخذ فى الاعتبار قيم معاملات التشابه عند التمثيل البيانى، ولذا سوف نكتفى بعرض التمثيل الشجرى للتجمعات.

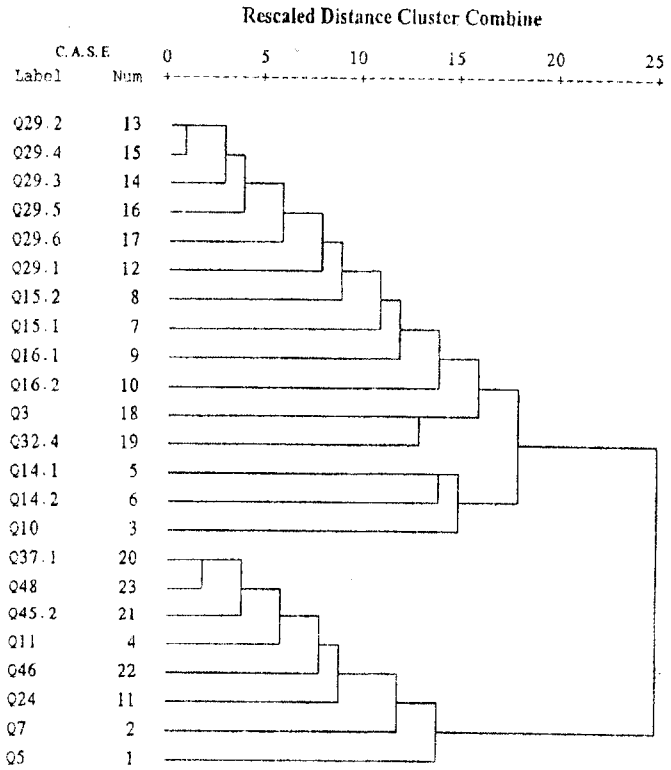
فى تحليل التجمعات باستخدام حزمة برامج "SPSS" لا يتم تمثيل القيم الفعلية لقيم معاملات التشابه، ولكن تتم إعادة قياسها بحيث تتراوح القيمة بين الصفر والقيمة ٢٥.

ويمثل الشكل رقم (١) الرسم البيانى الشجرى لتحليل التجمعات للمتغيرات، وحيث إنه قد تمت إعادة القياس للمعاملات بحيث تتراوح بين الصفر و٢٥، فإن ذلك يعنى أن أكبر قيمة لمعامل التماثل فى الجدول رقم (٧) وهى 0.885 تمثل القيمة (١) على الرسم وأصغر قيمة لمعامل التماثل وهى 0.515 تناظرها القيمة ٢٥، وتمثل باقى قيم المعاملات بين هاتين القيمتين. وبفحص الرسم نجد أن أول تجمع كان بين المتغيرين (١٣، ١٥)، حيث تم ربطهما معًا بخط رأسى يبعد عن خط البداية بمقدار الوحدة من وحدات القياس على الرسم. كذلك نلاحظ ربط المتغيرين (٧، ٣) بخط رأسى يبعد عن خط البداية بمقدار أكبر، وهكذا تتم خطوات الربط بين التجمعات بخطوات رأسية تبعد بمسافات تتناسب مع قيم المعاملات المناظرة لها فى الجدول رقم (٧) إلى

أن يتم الربط فى آخر تجمع رأسى يبعد ٢٥ عن خط البداية، وهو ما يناظر المعامل 0٠515.

شكل رقم (١)

Dendrogram using Average Linkage (within Group)



عدد التجمعات

من المخرجات التى تنتج عند إجراء تحليل التجمعات باستخدام طريقة التكتل الهرمى جميع الأعداد الممكنة من الأنماط، بداية من أن كل متغير يمثل نمطًا منفردًا إلى حالة النمط الواحد الذى يضم جميع المتغيرات. ويشير الجدول رقم (٨) إلى تسعة نماذج من التجمعات، وأيضًا المتغيرات التى تنتمى إلى كل تجمع.

ويتم تحديد العدد الأمثل من التجمعات بناءً على الهدف من إجراء التحليل؛ بحيث لا يكون بالعدد الكبير الذى يصعب فيه تحديد خصائص تلك التجمعات، ولا بالعدد القليل الذى تختفى معه تلك الخصائص.

وقد يتم تحديد عدد التجمعات وفقًا لمحك معامل التشابه[٢٥] – وهى التى تم استخدامها فى هذه الدراسة– حيث يتم التوقف عن الخطوة التى يكون فيها الفرق فى معامل التشابه بينها وبين الخطوة السابقة صغيرًا نسبيًّا. وعلى ذلك يمكن اعتبار أن عدد الأنماط هو (٤). وقد تم توضيح الأنماط والمتغيرات التى تنتمى إلى كل منها فى شكل رقم (٢).

جدول رقم (٨)

Cluster Membership of Variables using within Average Linkage

Label	Case	10 clusters	9 clusters	8 clusters	7 clusters	6 clusters	5 clusters	4 clusters	3 clusters	2 Clusters
Q5	1	1	1	1	1	1	1	1	1	1
Q7	2	2	2	2	2	1	1	1	1	1
Q10	3	3	3	3	3	2	2	2	2	2
Q11	4	2	2	2	2	1	1	1	1	1
Q14.1	5	4	4	4	4	3	3	2	2	2
Q14.2	6	5	5	5	4	3	3	2	2	2
Q15.1	7	6	6	6	5	4	4	3	3	2
Q15.2	8	6	6	6	5	4	4	3	3	2
Q16.1	9	7	6	6	5	4	4	3	3	2
Q16.2	10	8	7	7	6	5	4	3	3	2
Q24	11	2	2	2	2	1	1	1	3	1
Q29.1	12	6	6	6	5	4	4	3	3	2
Q29.2	13	6	6	6	5	4	4	3	3	2
Q29.3	14	6	6	6	5	4	4	3	3	2
Q29.4	15	6	6	6	5	4	4	3	3	2
Q29.5	16	6	6	6	5	4	4	3	3	2
Q29.6	17	6	6	6	5	4	4	3	3	2
Q31	18	9	8	8	7	6	5	4	3	2
Q32.4	19	10	9	8	7	1	5	4	3	1
Q37.1	20	2	2	2	2	1	1	1	1	1
Q45.2	21	2	2	2	2	1	1	1	1	1
Q46	22	2	2	2	2	1	1	1	1	1
Q48	23	2	2	2	2	1	1	1	1	1

جدول رقم (٩)

- التجمع الأول
  - الطفل غير متعلم
  - خرج للشارع قبل سن التعليم الأساسى
  - يعيش مع أسرة عدد أفرادها أكبر من ٦ أفراد
  - وجود مشكلات مع الأهل
  - سمع عن المخدرات من الأصحاب
  - يحب شرب المخدرات مع الأصحاب
  - يحصل على المخدرات من الأصحاب
  - أحد الأصحاب هو أول من علمه أخذ المخدرات

- التجمع الثالث
  - وجود زوجة أب أو وجود زوج أم
  - الأب أمي أو الأم أمية
  - معاملة سيئة من الأب أو الأم
  - معاملة سيئة من زوجة الأب
  - معاملة سيئة من زوج الأم
  - خلافات مع الإخوة
  - التشجيع من الغير على ترك المنزل

- التجمع الثانى
  - يعيش فى أسرة غير مكتملة
  - غياب الأب
  - غياب الأم

- التجمع الرابع
  - شقة غير مستقلة
  - لا توجد مجارٍ

نخلص مما سبق إلى وجود أربعة أنماط لأسباب ظاهرة أطفال الشوارع، ومن ثم تعرضهم للتعامل مع المخدرات سواء فى صورة تعاطى أو توزيع أو إتجار. ويمكن التعبير عن كل تجمع على النحو التالى:

١ - التجمع الأول "تأثير النظراء"

أثبتت الدراسة أن كبر حجم الأسرة عن الحد الذي يعجز فيه الآباء عن توجيه الأبناء وتلبية احتياجاتهم وإلحاقهم بالتعليم الأساسى، يدفع الطفل للخروج للشارع في سن صغيرة هربًا أو للعمل لمساعدة أسرهم. ومع المشكلات المستمرة مع الأهل واختفاء الجو الأسرى يزداد تأثير النظراء؛ حيث ينقاد الطفل إلى رغباتهم ويصبحون هم الجماعة المرجعية له، حيث إنه سمع عن المخدرات من الأصحاب، بل إنهم أول من علموه تعاطى المخدرات وهم الذين يجلبونها له.

٢ - التجمع الثانى "التفكك الأسرى"

إذا كان للأسرة دورها الإيجابى فى الرعاية والحماية والتنشئة وتوفير الاحتياجات الأساسية، فإنه يمكن أيضًا أن يكون لها دور سلبي إذا ما سادت فيها أجواء التوتر والاضطراب والتفكك. ومن أهم أسباب التفكك الأسرى غياب الأب أو الأم بسبب الطلاق، أو وفاة أحد الوالدين حيث يؤدى ذلك إلى ضعف التوجيه والتربية والرقابة مما يسهم فى خلق مشكلات أكثر خطورة ويكون إحدى آليات الهروب منها هو الخروج للشارع.

٣ - التجمع الثالث "القسوة والعنف ضد الأبناء"

قد ينعكس ضعف المستوى التعليمي لأرباب الأسر على سلوكياتهم، متمثلًا ذلك فى العنف والقسوة على الأبناء، فلا يستطيعون الفصل بين التربية بأساليب تنشئة سليمة ومعتدلة، وبين توقيع العقاب العنيف القاسى، هذه القسوة المبالغ فيها مع الخلافات والمشاحنات بين الإخوة تجعل الطفل ينفر من أجواء البيت، وتدفعه إلى الهروب من البيت للعيش بحرية بعيدًا عن قسوة الأهل وضغوطهم.

٤ - التجمع الرابع "عدم توافر المسكن الملائم"

يمثل عدم تناسب السكن وافتقاده المرافق والخدمات الأساسية مع اختفاء استقلالية الجو الأسرى مناخًا طاردًا للأطفال إلى الشارع؛ حيث يعتبر الشارع في هذه الحالة من عوامل الجذب، تخلصًا من بيئة فقيرة ماديًا ومعنويًا.

الخلاصة

فى هذه الدراسة تم اختيار مجموعة من المتغيرات التى تسهم فى خروج الطفل إلى الشارع، ومن ثم تعرضه للتعامل مع المخدرات سواء فى صورة تعاطى أو توزيع أو اتجار، وذلك من واقع بيانات بحث "أطفال الشوارع وجرائم المخدرات"، الذى قام به المركز القومى للبحوث الاجتماعية والجنائية على أفراد عينة من أطفال الشوارع الذين لهم علاقة بالمخدرات والموجودين فى المؤسسات الحكومية والجمعيات الأهلية بكل من محافظتى القاهرة والجيزة التى تتعامل مع أطفال الشوارع، وأيضًا تم التطبيق على بعض الأطفال الموجودين بالشارع. وقد تم تنميط هذه المتغيرات فى مجموعات متجانسة أو أنماط مميزة باستخدام تحليل التجمعات Cluster Analysis وهو الأسلوب الإحصائى المتقدم الذى يحقق هذا الهدف.

وقد تم أيضًا استخدام طريقة التكتل الهرمى لضم هذه المتغيرات فى تجمعات، وفى إطار هذه الطريقة هناك أربعة أساليب للربط بين المجموعات الجزئية فى كل خطوة من خطوات طريقة التجميع، وهى ربط مفرد، وربط كامل، ومتوسط الربط بين المجموعات، ومتوسط الربط داخل المجموعات، وهى لا تعطى النتائج نفسها كما أنه لا توجد طريقة بينها أفضل من الطرق الأخرى. وقد تم إجراء التحليل فى هذه الدراسة حسب كل من الطرق المختلفة

لربـط التجمعـات الجزئيـة، كمـا تـم حساب معامـل ارتبـاط Cophenetic للمفاضلة بينها. وقد أسفرت النتائج عن أن طريقة الربط داخل المجموعات هى صاحبة أعلى معامل، ولذلك تم استخدامها فى تحليل التجمعات. كذلك أسفرت نتائج التحليل أيضًا عن وجود أربعة عوامل رئيسة تؤدى إلى ظهور وتنامى مشكلة خروج الطفل إلى الشارع وهى: تأثير النظراء، والتفكك الأسرى، والقسوة والعنف وإيذاء الطفل، وأخيرًا البيئة السكنية للطفل. هذه العوامل مجتمعة تدفع الطفل للهروب من البيت.

المراجـــع

1 - www. swideg jeeran.com/ geography/../4/1035369.htm.

٢ – عادل سلطان، منصور مغاورى، التفاوتات فى إمكانية التنمية الزراعية فى محافظات الجمهورية (تصنيف الموارد الزراعية)" ، المؤتمر السنوى الأول ، المركز القومى للبحوث الاجتماعية والجنائية : المهام – المجالات، المجلد الأول، مايو ١٩٩٩، ص ص ٣٤٩ – ٣٨٧.

3 - www.digital-library.ksu.edu.sa/V.13M160R66.DOC.

4 - www.napcsyr.org/dwnld.files/policy../21_nara_rh_ar_pdf.

٥ – ماجده عبدالغنى، تنميط التفضيلات والاحتياجات الثقافية للشباب المتردد على معرض الكتاب باستخدام أسلوب تحليل التجمعات، المجلة الجنائية القومية، المجلد الرابع والأربعون، العدد الثالث، سبتمبر ٢٠٠٧، ص ص ٨٩-١١٤.

٦ – ماجدة عبدالغنى، تنميط محافظات الجمهورية وفقًا للنشاط الإجرامى للمسجل الخطر، المجلة الجنائية القومية، المجلد الثامن والأربعون، العدد الثالث، نوفمبر ٢٠٠٥، ص ص٤١-٤.

7 - www.iesr.shams.edu.eg/jourtnal/volume. 3_part1.pdf.

8 - Richard, A.J. and Dean, W., Applied Multivariate Statistical Analysis, 3rd Ed. New Delhi, 2002, pp. 573-602.

9 - www.users.cs.umn.edu/~kumar/dmbook/ch8.pdf.

10 - Richard, A. J. and Dean, W., op. cit., p.584.

11 – Ibid., pp. 585-596.

12 - www. cs.umd.edu/ ~mount/ projects/ K Means /.

13 - www.statisticalinnovations.com/products/twostep.pdf.

14 - Richard, A. J. and Dean, W., op. cit., pp.586-598.

15 – Ibid., pp.590-592.

16 – Ibid., p.594.

17 – Ibid., p.595.

18- www.public.asu.edu/~jyeo2/CLUSTER/Fall2007/.../Basiccluster.ppt.

19 - www.cs.kent.edu/~jin/DM08/Cluster Validation. pdf.

20 - Handl, J. K., Multi Objective Clustering and Cluster Validation, Springer Series on Computational Intelligence, 16, 2007, pp. 21-47.

21 - Moh'd B., and Mohammad Al Rawi, An Efficient Approach for Computing Silhouette Coefficient, Journal of Computer Science, 3, 2008, pp. 252-255.
22 - Miguel, S., and Eduardo G., Cluster Analysis: A Case Study of Environmental Data, Journal of Engineering and Computer Innovations. Vol.2, 2011, pp.22-27. Available online at http://www.academicjournals.org/JECI.
23 - Holmes F., and Huynh, Comparison of Similarity Measures in Cluster Analysis with Binary Data. Paper Presented at the Annual Meeting of the American Educational Research Association, 2000.
24 - Norusis, M. L., SPSS Advanced Statistics User's Guide, SPSS., Inc., 1990, pp. 349-363.
25 – Ibid., pp. 355-356.

COMPARISON OF HIERACHICAL CLUSTERING METHODS AND SELECTING THE BEST WAYS OF PATTERNING A SET OF VARIABLES THAT CONTRIBUTE IN CHILD RESORT TO THE STREET

Abeer Saleh

The best ways for the clustering of a set of variables that Contribute to the exit the child out to the street There are numerous ways in which clusters can be formed. Hierarchical clustering is one of the most straightforward methods. Many methods can be used to decide which cases or clusters should be combined at each step. All are based on either a matrix of distances or a matrix of similarities between pairs of cases. The methods differ in how they estimate distances between clusters at successive steps. Since the merging of clusters at each step depends on the distance measure, different distance measures can result in different cluster solutions for the same clustering method.

This paper aims to study the most frequently used methods for clustering the set of variables that contribute to the exit the child out to the street.