



# جامعة الناصر AL-NASSER UNIVERSITY

تصنيف وتمييز المحافظات اليمنية بحسب مصادر الدخل الفردي باستخدام أسلوب  
التحليل العنقودي والتحليل التمييزي

د. فؤاد عبده إسماعيل المخلافي

قسم الإحصاء

كلية التجارة والاقتصاد

جامعة صنعاء

AUTHORIZED BY AL-NASSER UNIVERSITY'S RESEARCH OFFICE

جميع حقوق النشر محفوظة لمكتب البحوث والنشر بجامعة الناصر

### ملخص البحث

تهتم الجمهورية اليمنية كغيرها من الدول بتقليص التفاوت في توزيع الدخل بين السكان وتحقيق أكبر عدالة في توزيع الدخل والقضاء على الفقر، لذا قامت الجمهورية اليمنية بإجراء العديد من المسوح لميزانية الأسرة (الدخل والإنفاق) والسعي لعمل دراسات وأبحاث تتناول هذا الموضوع، وعلى الرغم من قلة الدراسات في هذا المجال ومع ذلك فهي تقدم لنا بعض المؤشرات عن الدخل والإنفاق، وفي هذا البحث تم تطبيق أساليب التحليل الإحصائي متعدد المتغيرات حيث تم استخدام أسلوب التحليل العنقودي متعدد المتغيرات من أجل معرفة التفاوت في توزيع الدخل الفردي بين المحافظات اليمنية وقد تم التوصل إلى أن هناك تقارباً في مصادر الدخل الفردي بين (17) محافظة شكلت العنقود الأول وهي المحافظات ذات الدخل المنخفض وهي أبين، لحج، إب، تعز، ريمه، ذمار عمران، صعده، الجوف، الضالع، صنعاء، الحديدة، مأرب، حجة، المحويت، البيضاء، حضرموت، أيضاً كان هناك تقارباً في مصادر الدخل الفردي بين أربع محافظات شكلت العنقود الثاني وهي المحافظات ذات الدخل المرتفع وهي المهرة، الامانة، عدن، شبوة.

كما تم تطبيق أسلوب التحليل التمييزي متعدد المتغيرات لتمييز مصادر الدخل التي تساهم بدرجة كبيرة في هذا التفاوت وتبين أن المتغيرات X1 (الاجور والمرتبات)، X3 (المباع من الأسماك)، X4 (العائد من أنشطة الأعمال الخاصة)، X5 (القيمة التقديرية لإيجار المسكن المملوك للأسرة)، X6 (المنح الدراسية والمساعدات المالية) هي مصادر الدخل الفردي التي ساهمت بدرجة كبيرة في التمييز بين المحافظات.

Classification and discrimination The Yemen's Governorates by income sources using  
Cluster analysis and discriminant analysis methods

D. Fuad Abdo Esmail AL-Mekhlafi  
Department of Statistics  
Faculty of Commerce and Economics  
Sana'a University

### **Abstract**

Interested in the Republic of Yemen, like other countries to reduce the disparity in income distribution among the population and to achieve greater fairness in the distribution of income and poverty eradication, so the Republic of Yemen to conduct several surveys to the household budget (income and expenditure) and strive to work studies and research on this subject, In spite of the lack of studies in this area, however, they give us some indication of income and expenditure, and this research has been applied methods of multivariate statistical analysis, where the method was used cluster analysis to see the disparity in the distribution of per capita income between The Yemen's Governorates has been reached that there is a convergence in per capita income sources between (17) Governorate formed the first cluster of low-income Governorates which Abyan , Laheg, Ibb, Taiz, Reymah, Dhamar , Amran, Sa'adah ,Al-Jawf ,Al-Daleh, Sana'a , Al-Hodeidah, Mareb, Hajjah, Al-Mahweet, Al-Baida , Hadramout , also there was a convergence of per capita income sources between the four Governorates formed a second cluster of high-income Governorates which, Aden, Al-Maharah , Sana'a City , Shabwah.

It also has been applied discriminant analysis method to distinguish the sources of income that contribute significantly to this disparity shows that the variables X1(wages and salaries), X3 (the sold from fish), X4 (revenue from private business activities),

X5(estimated value of the rent of the housing that the household owned) and X6 (scholarships and financial aids) are the sources of per capita income, which contributed significantly to the distinction between the Governorates.

### الفصل الأول : الإطار المنهجي للبحث

#### 1-1 المقدمة

إعتاد علماء الاقتصاد قياس رفاة البلدان إعتياداً على الدخل الفردي الحقيقي، فكلما ازداد دخل الفرد أمكنه الحصول على كمية أكبر من السلع والخدمات، وبالتالي حقق مستوى معيشة أفضل ، ولا خلاف على أن ثمة عوامل أخرى مهمة تدخل في قياس مستوى رفاة الإنسان من قبيل توافر الخدمات الاجتماعية الأساسية مثل التعليم والصحة، ولعل هذا ما دفع برنامج الأمم المتحدة الإنمائي قبل عقدين إلى صياغة مؤشر التنمية البشرية بدمج البعدين التعليمي والصحي إلى جانب الدخل في قياس رفاة الإنسان، ومع أن هذا المؤشر وبالنظر إلى البيانات الإحصائية المتاحة والقابلة للمقارنة دولياً يقدم صورة أدق لمستوى الرفاه إلا أنه لا يخلو من عيوب أهمها إغفال مستوى التفاوت بين سكان البلد الواحد، ومن أجل تدارك هذا الواقع عمد برنامج الأمم المتحدة الإنمائي في آخر تقاريره السنوية إلى تصحيح مؤشره بدمج مدى عدالة توزيع الدخل في ترتيب البلدان ضمن سلم التنمية البشرية وقد أسفر هذا التصحيح عن تراجع مؤشر بعض الدول نظراً إلى سوء توزيع الدخل في هذه البلدان مقارنة بالمتوسط العالمي، وتعد مشكلة توزيع الدخل في الدول النامية من المشاكل المهمة التي تثير اهتمام الكثير من الاقتصاديين وذلك لاتصالها الوثيق بمستويات المعيشة لكل فئات المجتمع وبخاصة الفئات ذات الدخل المنخفض والتي تشكل نسبة كبيرة منها من جهة واعتبارها هدفاً من أهداف النمو الاقتصادي من جهة أخرى.

تهتم الجمهورية اليمنية كغيرها من الدول بتقليص التفاوت في توزيع الدخل بين السكان وتحقيق أكبر عدالة في توزيع الدخل والقضاء على الفقر، لذا اهتمت الجمهورية اليمنية بإجراء العديد من المسوح لميزانية الأسرة (الدخل والإنفاق) والسعي لعمل دراسات وأبحاث تتناول هذا الموضوع، وعلى الرغم من قلة الدراسات في هذا المجال ومع ذلك فهي تقدم لنا بعض المؤشرات عن الدخل والإنفاق ، وفي هذا البحث سيتم تطبيق أساليب التحليل الإحصائي متعدد المتغيرات حيث سيتم

استخدام اسلوب التحليل العنقودي متعدد المتغيرات من أجل معرفة التفاوت في توزيع الدخل الفردي بين المحافظات اليمنية كما سيتم تطبيق اسلوب التحليل التمييزي متعدد المتغيرات لتمييز مصادر الدخل التي تساهم بدرجة كبيرة في هذا التفاوت، وقد تم تقسيم البحث إلى أربعة فصول، الفصل الأول : الإطار المنهجي للبحث، الفصل الثاني : الجانب النظري، الفصل الثالث : الجانب التطبيقي، الفصل الرابع : الاستنتاجات والتوصيات.

### 2-1 مشكلة البحث

إن معظم الدراسات والأبحاث التي تقوم بدراسة التفاوت في توزيع الدخل اعتمدت على منحنى لورنس ومعامل جيني وغيرها ، وهذه المقاييس تعتمد على توزيع الافراد داخل فئات الدخل فإذا اردنا عمل مقارنة بين المحافظات المختلفة لابد من حساب هذه المقاييس لكل محافظة على حده ، كما ان هذه المقاييس تعتمد فقط على اجمالي الدخل الفردي ولا تهتم بمصادر الدخل المختلفة من اجل اظهار الاهمية النسبية لكل مصدر على حده ، لذا سنقوم باستخدام اسلوب التحليل العنقودي لمعرفة التفاوت في توزيع الدخل الفردي بين المحافظات اليمنية كما سيتم استخدام اسلوب التحليل التمييزي لمعرفة مصادر الدخل التي تساهم بدرجة كبيرة في هذا التفاوت.

### 3-1 اهداف البحث

يهدف هذا البحث إلى استخدام التحليل العنقودي متعدد المتغيرات في تصنيف المحافظات اليمنية إلى مجموعات متجانسة بحسب متوسط دخل الفرد السنوي من خلال مصادر الدخل المختلفة وتحديد مدى التباعد والتقارب بين هذه المحافظات وإلى استخدام التحليل التمييزي متعدد المتغيرات لتمييز مصادر الدخل المتسببة في مثل هذا التفاوت.

### 4-1 اهمية البحث

تكمن أهمية البحث في استخدام اساليب التحليل الاحصائي متعدد المتغيرات وهي من الاساليب الاحصائية المتقدمة حيث سيتم استخدام اسلوب التحليل العنقودي لتصنيف المحافظات اليمنية بحسب مصادر الدخل مما يسهل عملية المقارنة بينها، كما سيتم استخدام اسلوب التحليل التمييزي لتمييز مصادر الدخل الفردي وبالتالي معرفة المصادر ذات الأهمية النسبية الكبيرة والتركيز عليها لمعالجة الاختلالات الموجودة في توزيع دخل الفرد السنوي بين المحافظات.

### 5-1 منهجية البحث

إعتمد البحث في جانبه النظري على الكتب والمراجع العلمية والدوريات المتخصصة والدراسات المختلفة المتعلقة بأدبيات البحث أما في جانبه التطبيقي فقد تم اعتماد اسلوب التحليل العنقودي متعدد المتغيرات في تصنيف المحافظات اليمينية الى مجموعات متجانسة بحسب متوسط دخل الفرد السنوي من خلال مصادر الدخل المختلفة وتحديد مدى التباعد والتقارب بين هذه المحافظات وإلى استخدام اسلوب التحليل التمييزي متعدد المتغيرات لتمييز مصادر الدخل المتسببة في مثل هذا التفاوت.

#### 6-1 الدراسات السابقة

هناك العديد من الدراسات الأجنبية والعربية التي استخدمت اسلوب التحليل العنقودي وأسلوب التحليل التمييزي ، على سبيل الذكر منها:

دراسة (الحنيطي واخرون،2004 ) ، حيث تم استخدام التحليل التمييزي لتمييز الأسر الفقيرة من غير الفقيرة في المناطق النائية التابعة لإقليم جنوب الاردن.

دراسة (AL-osaimy & Bamakhramah, 2004) ، حيث تم استخدام التحليل التمييزي لتمييز أداء البنوك الاسلامية في المملكة العربية السعودية.

دراسة ( Gwo & Chun,2006 ) ، حيث تم استخدام التحليل العنقودي والتحليل التمييزي في تحليل العوامل الهيدرولوجية لمصادر المياه في تايوان.

دراسة (الجاعوني وغانم، 2007) ،حيث تم استخدام التحليل التمييزي في توصيف وتوزيع الأسر داخل الهيكل الاقتصادي الاجتماعي في المجتمع السوري.

دراسة (مصطفى ، 2007 ) ، حيث تم استخدام بعض طرق التحليل العنقودي في التصنيف مع تطبيق عملي على مجموعة من الدول العربية في القرن الافريقي والعراق.

دراسة (Fazzo& et al,2008) ، حيث تم تطبيق التحليل العنقودي في تحليل معدلات الوفيات والتشوهات في محافظة نابلي الايطالية.

دراسة (نامق، 2010) ، حيث تم استخدام اسلوب التحليل العنقودي لتصنيف الانفاق على السلع والخدمات الاساسية وفقا للمستوى البيئي (حضر وريف ) في العراق.

دراسة (Gumparthi & Manicka,2010) ، حيث تم تصنيف المخاطر في المنشآت الصغيرة والمتوسطة في الهند بالاعتماد على التحليل التمييزي.

دراسة (ذنون ، 2012) ، حيث تم استخدام طريقتي التحليل العنقودي والتحليل التمييزي في التصنيف مع التطبيق على نتائج الدرجات العملية التي تمثل نتائج طلبة قسم أنظمة الحاسبات (المرحلة الاولى) في المعهد التقني في نينوى في العراق.

دراسة (Muntaner & et al.,2012) ، حيث تم استخدام التحليل العنقودي المتسلسل لتصنيف البلدان ذات الدخل المتوسط والدخل المنخفض بحسب لوائح سوق العمل والاعتماد عليها في تمييز مؤشرات الصحة السكانية.

#### 7-1 فرضيات البحث

تم صياغة الفرضيتين التاليتين:

1- لا يوجد تجانس في توزيع الدخل الفردي بين المحافظات اليمينية.

2- يتأثر عدم التجانس في توزيع الدخل بمصادر الدخل الفردي المختلفة.

#### 8-1 حدود البحث

تم الاعتماد على مخرجات مسح ميزانية الاسرة 2006 الصادر عن الجهاز المركزي للإحصاء وهو آخر مسح تم تنفيذه حتى كتابة هذا البحث علماً أن مثل هذه المسوح يتم تنفيذها كل 8 إلى 10 سنوات.



## الفصل الثاني : الجانب النظري

### 1-2 التحليل العنقودي (Cluster Analysis)

التحليل العنقودي يعتبر احد فروع التحليل الاحصائي متعدد المتغيرات, والتحليل العنقودي عبارة عن إجراءات تهدف إلى تصنيف مجموعة حالات أو متغيرات بطرق معينة وترتيبها داخل عناقيد بحيث تكون الحالات المصنفة داخل عنقود معين متجانسة فيما يتعلق بخصائص محددة وتختلف عن حالات أخرى موجودة في عنقود آخر (جودة ، 2008، ص29).

### 1-1-2 المفاهيم الأساسية في التحليل العنقودي

#### 1-1-1-2 العنصر (Element) (الجبوري وحمزة ، 2000، ص 820)

العنصر  $X_i$  هو متجه في فراغ المقياس ل  $n$  من الأبعاد.

$$X_i = (X_{i1}, X_{i2}, \dots, X_{in}) \quad (1-2)$$

والعناصر هي قيم عددية لكميات ممكنة القياس.

#### 2-1-1-2 المسافة (D) (Distance)

هي الحيز أو الفضاء الفاصل بين عنصرين، والعلاقة بين التشابه والمسافة هي علاقة عكسية ويمكن إجراء التحليل العنقودي بالاستناد على أي منهما وهناك صيغ متعددة لقياس هذه المسافة منها صيغة منكوسكي (Minkowski metric) وهي كما يلي ( Rencher, 2002, P453):

$$d(x,y) = \left[ \sum_{i=1}^p |x_i - y_i|^r \right]^{\frac{1}{r}} \quad (2-2)$$

عندما  $r=1$   $d(x,y)$  تقيس مسافة نواصي المدن بين نقطتين في فراغ ذي  $p$  من الأبعاد.

عندما  $r = 2$   $d(x,y)$  تقيس المسافة الاقليدية بين نقطتين في فراغ ذي  $p$  من الابعاد. وعموما فان تغيير قيمة  $r$  يؤدي الى تغيير الاوزان المعطاة للفروق الكبيرة والفروق الصغيرة.

#### 3-1-1-2 العنقود (Cluster) (Timm,2002,P515)

هو عبارة عن مجموعة من العناصر المتجانسة الى حد ما ولوصف ما داخل العنقود الواحد والمختلفة عن العناصر داخل العناقيد الاخرى.

#### 4-1-1-2 الشجرة البيانية (Tree diagram or dendogram) (Timm, 2002,P523)

هي الشكل الهرمي الناتج بعد إجراء عملية العنقدة ( تكوين العناقيد) ويمكن الوصول إليها وفق طريقتين:

##### 1- طريقة التجميع (agglomerative) (Hardle&Simar,2003,P308) ، حيث يبدأ التحليل بعنقود واحد لكل حالة ثم

يتم تجميع العناقيد المتشابهة تدريجيا حتى نصل الى العدد المطلوب من العناقيد.

##### 2- طريقة التجزئة (divisive) (Rencher,2002, P479)، في هذا النوع يتم اعتبار ان جميع الحالات تتجمع في

عنقود واحد وبعد ذلك يتم تصنيف الحالات في عناقيد أصغر فأصغر .

وفي كلا الحالتين فإن النتائج التي تظهرها الطريقتين هي شجرة تسلسل هرمي وأن بداية الشعبة يطلق عليها الجذر ونقاط

التفرعات يطلق عليها العقد ، والعقدة النهائية أو الاخيرة على الشجرة ليس لها تفرعات يطلق عليها الاوراق وهي تمثل

العناصر التي اجتمعت مع بعضها وكل واحدة من العقد في الشجرة وبضمنها الجذر تمثل مجموعة نوعية لكافة الاشياء

التي يمكن الوصول اليها في تلك العقدة باتجاه المقدمة ومن خلال الشجرة.

#### 2-1-2 خطوات التعتقد (Clustering Steps) (Timm , 2002 , p523)

1- حساب مصفوفة المسافة أو مصفوفة الارتباط أو مصفوفة التشابه.

2- يتم ربط العنصرين اللذين تكون المسافة بينهما أقصر المسافات ضمن المصفوفة المحسوبة في الخطوة السابقة ، وفي

حالة وجود مسافات متساوية فبالإمكان إجراء عملية الربط لأكثر من عنصرين في مرحلة واحدة (لعنصرين معاً).

3- يتم حساب مصفوفة المسافة الجديدة التي تأخذ بنظر الاعتبار التغيرات التي حصلت في الخطوة الثانية.

4- يتم الاستمرار بعملية الربط حتى الوصول الى شجرة العناقيد.

الجدير بالذكر بأنه قد يتم إجراء تحويل للبيانات مثل أخذ اللوغاريتم أو التحويل إلى الدرجة المعيارية قبل إجراء العملية أعلاه وبالذات عند اختلاف وحدات القياس.

### 3-1-2 طرق التحليل العنقودي الهرمي (المتسلسل) (Hierarchical Clustering Methods)

هناك طرق عديدة للتحليل العنقودي ، لكل طريقة هناك خصائص معينة تتوفر فيها تختلف عن الطرق الأخرى بعضها يعتمد طريقة التجميع والآخر يعتمد طريقة التجزئة (فهومي ، 2005، ص813 )، وسنتناول بالشرح أكثر الطرق استخداما وهي طرق التحليل العنقودي الهرمي (المتسلسل)، وهذه الطرق لا تتطلب المعرفة المسبقة بعدد العناقيد التي سيتم تصنيف الحالات على أساسها كما أنها تتناسب العينات الصغيرة نسبياً ، وفي التحليل العنقودي المتسلسل توجد طرق مختلفة وسنركز على طرق الإدماج المتسلسل وخصوصا ما يسمى بطرق الربط ( Linkage methods ) وهذه الطرق تلائم تجميع الحالات كما تلائم تجميع المتغيرات وهذا مالا يتحقق لطرق التجميع الأخرى (عزام ، 1998، ص ص 865-866).

#### 1-3-1-2 طريقة الربط المنفرد (Single Linkage) وتسمى أيضاً بطريقة الجوار الأقرب (nearest neighbor)

تعتمد هذه الطريقة بالأساس على اعتبار أن العنصرين الأكثر تشابهاً بين العناصر يشكل نواة العنقود، ثم تضاف باقي الوحدات إلى هذه النواة بالتسلسل وحسب درجة الشبه مع عناصر نواة العنقود إذ يضاف الأكثر شبهاً ثم الأقل وبالتدرج وفي حالة ربط مجموعة من العناقيد مع بعضها فإن ذلك يتم بالاستناد على اقرب المسافات أو معاملات التماثل بين أزواج العناصر وحسب الصيغة التالية:

$$D_{IJ} = \text{Min}(D_{ij}) \quad (3-2)$$

حيث  $J, I$  تمثل العناصر في العناقيد  $J, I$  على التوالي.

#### 2-3-1-2 طريقة الربط الشامل (Complete Linkage) وتسمى أيضاً بطريقة الجوار الأبعد (farthest neighbor)

(Rencher, 2002, PP459- 460)

تعتمد هذه الطريقة بالأساس على اعتبار أن العنصرين الأكثر تشابهاً بين العناصر يشكل نواة العنقود، ثم بعد ذلك تعمل هذه الطريقة بشكل معاكس تماماً لمبدأ عمل الطريقة السابقة ، فالعنصر المرشح للدخول الى العنقود يجب أن تكون المسافة بينه وبين أي من عناصر العنقود هي اكبر من المسافة بين اي عنصر آخر وبين أي من عناصر العنقود ويتم الربط حسب

$$D_{IJ} = \text{Max}(D_{ij}) \quad (4-2) \quad \text{الصيغة الاتية:}$$

حيث J، أ تمثل العناصر في العناقيد J، I على التوالي.

كما يوجد طريقة الربط المتوسط (Average Linkage) وطريقة مجموع المربعات الاضافية ( Ward's )

(Incremental Sum of Squares)، ولمعرفة المزيد عن هذه الطرق يمكن الرجوع الى المصدر

(Timm,2002,PP528-529).

## 2-2 التحليل التمييزي (Discriminant Analysis)

التحليل التمييزي هو ايضاً أسلوب إحصائي لتحليل البيانات متعددة المتغيرات ،حيث يهتم بمسألة التمييز بين مجموعتين أو اكثر والتي تكون متشابهة في كثير من الصفات على أساس عدة متغيرات من خلال استخدام الدالة المميزة والتي هي عبارة عن تركيب خطي للمتغيرات المستقلة ، ويختلف التحليل التمييزي عن التحليل العنقودي في أن فكرة التحليل العنقودي تبدأ دون توفر معرفة مسبقة بعدد المجاميع أو اياً من المفردات تنتمي لهذه المجموعة أو تلك ( Rencher,2002,P451 )، كما ان التحليل التمييزي يختلف عن تحليل الانحدار في أن المتغير التابع في التحليل التمييزي هو متغير إسمي وهو من المتغيرات النوعية بينما المتغير التابع في تحليل الانحدار هو في الغالب متغير مستمر وهو من المتغيرات الكمية (الراوي ، 1987 ،ص 510 ).

أما عملية التصنيف (Classification) فهي العملية اللاحقة بعد تكوين الدالة المميزة حيث يتم الاعتماد على هذه الدالة بالتنبؤ وتصنيف المفردة الجديدة لإحدى المجموعات قيد الدراسة بأقل خطأ تصنيف ممكن ، ويشترط تساوي التباينات للمجموعات المبحوثة ، وهناك تمييز خطي في حالة مجموعتين ، وتمييز خطي في حالة أكثر من مجموعتين ، أما التمييز غير الخطي فيستخدم في حالة عدم تساوي التباينات (الجبوري وحمزة ، 2000 ).

## 1-2-2 الدالة المميزة الخطية في حالة مجموعتين (Linear Discriminant Function –Two Groups)

إن دالة التمييز هي نموذج يمكن صياغته اعتماداً على مؤشرات العينة التي تم اختيار مفرداتها ووضعت في مجموعتين مختلفتين ، وبواسطة هذه الدالة نستطيع أن نختبر المفردة ونحدد عائديتها إلى أي مجموعة.

فلو فرضنا أن مجال العينة هو  $W$  سوف يقسم إلى قسمين ( $R$ ) يعود إلى المجموعة الأولى و ( $W-R$ ) يعود إلى المجموعة الثانية أما الحد الفاصل بين المجموعتين فيمكن أن يعود إلى أية مجموعة من هاتين المجموعتين. إذا كان لدينا المتغيرات ( $X_1, X_2, \dots, X_p$ ) فان الصيغة العامة للدالة المميزة هي كالآتي (الراوي، 1987، ص 511):

$$Y_i = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad , \quad i=1,2,\dots,r \quad (5-2)$$

حيث:

$P$  : عدد المتغيرات الداخلة في الدالة .

$\beta$  : معاملات الدالة المميزة المعيارية.

$r$  : عدد الدوال المميزة.

ولتحديد الاختلافات بين المجموعتين فإنه من المناسب استخراج الأوساط الحسابية لهاتين المجموعتين.

إن عملية التقدير للمعاملات ( $\beta, S$ ) والتي تجعل الدالة تعطي أفضل تمييز بين المجموعتين لا بد أن يتم من خلال جعل مربع الفرق بين متوسطي المجموعتين إلى التباين المشترك للمجموعتين أكبر ما يمكن ( Rencher, 1995, p297 ) أي

$$Q = \frac{[\bar{Y}_1 - \bar{Y}_2]^2}{\sum_{i=1}^2 \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2} = \frac{[\beta'(\bar{X}_1 - \bar{X}_2)]^2}{\beta' S \beta} \quad (6-2) \quad \text{أن :}$$

حيث نقدر معاملات الدالة المميزة من خلال تعظيم النسبة  $Q$  باشتقاقها جزئياً ومساواتها بالصفر ونحصل على:

$$\beta^{\wedge} = S^{-1}(\bar{X}_1 - \bar{X}_2)' \quad (7-2)$$

في حالة وجود مجموعتين يكون لدينا دالة تمييز واحدة فقط ، وفي حالة وجود ثلاث مجاميع يكون لدينا دالتين تمييزيتين.

وبعد استخراج المعاملات  $\beta^{\wedge}$  تصنف المفردة إلى إحدى المجموعتين بالاعتماد على نقطة وسط المجموعتين ( $L$ ) التي

تجعل احتمال التصنيف الخاطئ أقل ما يمكن (عزام ، 1998 ، ص 742 ).

$$L = \frac{\bar{Y}_1 + \bar{Y}_2}{2} \quad (8-2)$$

تصنف المشاهدة إلى المجموعة الأولى إذا كانت  $\hat{Y} > L$

تصنف المشاهدة إلى المجموعة الثانية إذا كانت  $\hat{Y} < L$

تصنف المشاهدة عشوائياً إلى المجموعة الأولى أو الثانية إذا كانت  $\hat{Y} = L$

حيث :

$$\hat{Y} = (\bar{X}_1 - \bar{X}_2)' S^{-1} X \quad (9-2)$$

### 1-1-2-2 إختبار معنوية الدالة المميزة الخطية

عندما يراد التمييز بين مجموعتين ، فإنه يمكننا أن نختبر الفرضية التي تنص على تساوي متوسطات المجموعتين:

$$H_0: M_1 = M_2$$

$$H_1: M_1 \neq M_2$$

إحصاءة الاختبار المستخدمة في حالة التمييز بين مجموعتين هي  $T^2$  (Hotelling) وصيغته كما يلي

:( Rencher,1995,P140)

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} D^2 \quad (10-2)$$

حيث :  $D^2$  يمثل (Mahalanobis Distance) وصيغته كما يلي:

$$D^2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \quad (11-2)$$

ويستخدم اختبار (F) وصيغته كما يلي :

$$F = \frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} T^2 \quad (12-2)$$

بدرجة حرية  $(P, n_1 + n_2 - p - 1)$  ، نرفض  $H_0$  بمستوى معنوية  $F_\alpha$  إذا كانت:

$$F_{cal} > F_{\alpha, (P, n_1 + n_2 - p - 1)}$$

ونقبل  $H_1$  وهذا يدل على أن متوسطات المجموعات غير متساوية وأنه توجد فروق معنوية بين المجموعتين وهذا معناه أن الدالة المميزة الخطية قابلة للتمييز بدرجة عالية ، ايضاً يمكن استخدام مقياس ولكس (Wilks-Criteria) وفق الصيغة

$$\Lambda = \frac{|W|}{|T|} \quad (13-2) \quad \text{التالية (Rencher, 1995, P180):}$$

حيث :

T: مصفوفة التباين والتغاير الكلي للمجموعات.

W : مصفوفة التباين والتغاير داخل المجموعات.

وتتراوح قيمة  $(\Lambda)$  بين الصفر والواحد ، فإذا كانت قريبة أو مساوية للواحد فان ذلك يشير إلى أن متوسطات المجموعات متساوية وبذلك لا يوجد تمييز بين المجموعات ، اما اذا كانت قيمتها قريبة من الصفر فان ذلك يدل على قوة التمييز.

كما يمكن استخدام مقياس  $(\chi^2)$  ويعد هذا المقياس أكثر دقة من مقياس  $(\Lambda)$  وصيغته كما يلي :

$$\chi^2 = - \text{Log} (\Lambda) \quad (14-2)$$

بدرجة حرية  $P(K-1)$ ، حيث P عدد المتغيرات ، K عدد المجموعات.

#### 2-1-2-2 إحتمال خطأ التصنيف

هناك نوعان من احتمال خطأ التصنيف هما:

1-احتمال خطأ التصنيف  $P_{12}$  وهو احتمال تصنيف المفردة إلى المجموعة الثانية وهي أصلاً تعود إلى المجموعة الأولى.

2- احتمال خطأ التصنيف  $P_{21}$  وهو احتمال تصنيف المفردة الى المجموعة الأولى وهي أصلاً تعود إلى المجموعة الثانية.

وبذلك سوف يكون تقدير احتمال التصنيف كما يلي(Hardle& Simar,2003,p332):

$$P_{12} = P_{21} = \phi(-D/2) \quad (15-2)$$

حيث  $\phi$  : تمثل دالة التوزيع الطبيعي القياسي.

D: هو جذر مقياس مهالانوبيس  $D^2$  (Mahalanobis Distance).

### الفصل الثالث: الجانب التطبيقي

في هذا الفصل سيتم تطبيق أسلوب التحليل العنقودي على البيانات بعد أن تم التصنيف حسب المحافظات اليمنية ويهدف إيجاد تجمعات من المحافظات تكون متجانسة فيما بينها من أجل التصنيف والمقارنة في آن واحد كما سيتم تطبيق أسلوب التحليل التمييزي لتمييز مصادر الدخل المتسببة في حالة عدم التجانس بين المحافظات، وستكون نتائج التحليل والمقارنة بين الحالات (المحافظات) قيد الدراسة بالاعتماد على المتغيرات (مصادر الدخل الفردي)، الجدول في الملحق جدول (م1)، ، وسيتم استخدام البرنامج الإحصائي (SPSS version 20) في كل مراحل التحليل.

#### 1-3 تحليل نتائج التحليل العنقودي

تم استخدام طريقة الربط الشامل لإيجاد العلاقة بين المحافظات بالاعتماد على مقياس التشابه المستخدم وهو معامل منكوسكي (Minkowski metric) بتطبيق الصيغة (2-2) عندما  $r=1$  والنتائج موضحة بالجدول (1-3) والتي تمثل مصفوفة القربية (Proximity matrix) والتي تقرر التشابه أو عدم التشابه بين المحافظات ويعبر عنها بالمسافات المشتقة بين كل محافظة مع المحافظات الأخرى .



من خلال الجدول (3-1) نلاحظ أن محافظتي أبين ولحج شكلتا أول تعنقد وقد كان أقل معامل اقتراب (6756.000) لتشكل عنقوداً جديداً في مرحلة جديدة وهي المرحلة اللاحقة (5) يلي ذلك تعنقد في محافظتي إب وتعز بمعامل اقتراب (11015.000) لتشكل عنقوداً جديداً في المرحلة (5) ثم تلاها تعنقد في محافظتي حجة و المحويت بمعامل اقتراب (18655.000) لتشكل عنقوداً جديداً في المرحلة (12) واستمرت عملية التعنقد من أبعد الأقل الى أبعد الأكبر بشكل تصاعدي حيث كان أكبر معامل اقتراب هو (228824.000) مابين محافظتي إب والأمانة ، أما شكل الشجرة البيانية (Dendrogram) فتضم قياسات (مستويات) تمتد إلى (25) وحدة قياس حيث يشير طول الخط إلى زيادة درجات عدم التشابه وتوجد عدة عقد (Nodes) موجودة في الشجرة تمثل كل عقدة اندماج حالتين أو أكثر كما هو واضح في الشكل (3-1) ، وفي نهاية تحليل البيانات تم وضع النتائج في عناقيد (Clusters) من حيث التشابه في مصادر الدخل للمحافظات وكما هو واضح في الشكل (3-1) أن هناك (17) محافظة تعنقدت عند مسافات متقاربة وهي أبين ، لحج ، إب ، تعز ، ريمه ، ذمار ، عمران ، صعده ، الجوف ، الضالع ، صنعاء ، الحديدة ، مأرب ، حجة ، المحويت ، البيضاء ، حضرموت ويعكس ذلك التشابه في مصادر الدخل الفردي والتجانس الكبير بين هذه المحافظات وبالتالي فهي تشكل العنقود الاول ذات الدخل المنخفض بينما تعنقدت أربع محافظات عند مسافات متباعدة وهي شبوة ، عدن ، الامانة ، المهرة ، وبالتالي فهي تشكل العنقود الثاني ذات الدخل المرتفع.

### 2-3 تحليل نتائج التحليل التمييزي

تم الاعتماد في تكوين دالة التمييز الخطية على عدة متغيرات ( مصادر الدخل الفردي) جمعت عن كل مفردة (محافظة) ، وتم استخدام المجاميع المستخلصة من نتائج التحليل العنقودي في التحليل التمييزي ، وحيث أن عدد المجاميع هي مجموعتين لذا سيتم استخدام التحليل التمييزي في حالة مجموعتين :

أ- المتغير التابع والذي يمثل حالة المحافظة (1 يمثل محافظات العنقود الاول ذات الدخل المنخفض) ، (2 يمثل محافظات العنقود الثاني ذات الدخل المرتفع).

ب- المتغيرات المستقلة تمثلت بمصادر الدخل الفردي وهي X1 (الاجور والمرتبات) ، X2 (المبايع من كمية الانتاج الزراعي والحيواني) ، X3 (المبايع من الأسماك) ، X4 (العائد من أنشطة الاعمال الخاصة) ، X5 (القيمة التقديرية

لإيجار المسكن المملوك للأسرة) ، X6 (المنح الدراسية والمساعدات المالية) ، X7 (مصادر الدخل الأخرى النقدية والعينية) ، X8 (قيمة الاستهلاك من انتاج الأسرة والمهداة).

### 1-2-3 إختبار معنوية المتغيرات في الدالة المميزة

نحتاج إلى احتساب المتوسطات الحسابية للمتغيرات الداخلة في المجموعات وهي كما في الجدول (2-3) لكي يتم اختبار معنوية جميع المتغيرات وذلك لمعرفة أهمية كل متغير في الدالة المميزة ، ومدى تأثيره في تحليل النتائج باستخدام تحليل التباين الأحادي ، وكانت النتائج كما هو في الجدول (3-3) ، ونلاحظ من خلال الجدول (3-3) أن المتغيران X1 (الاجور والمرتبات) وX5 (القيمة التقديرية لإيجار المسكن المملوك للأسرة) يمتازان بمعنوية عالية ، ولهما تأثير كبير من حيث التفرقة بين مجموعتي المحافظات يليهما المتغير X7 (مصادر الدخل الأخرى النقدية والعينية) ثم X4 (العائد من أنشطة الاعمال الخاصة) و X3 (المباع من الأسماك) و X2 (المباع من كمية الانتاج الزراعي والحيواني) أما المتغيرين X8 (قيمة الاستهلاك من انتاج الأسرة والمهداة) و X6 (المنح الدراسية والمساعدات المالية) فليس لهما تأثير معنوي.

### 2-2-3 إختبار معنوية الدالة المميزة الخطية

لاختبار معنوية الدالة المميزة الخطية تم الحصول على المقاييس الموضحة في الجدول (4-3) ، يتضح من هذا الجدول أن التباين بين المجموعتين لكل المتغيرات المميزة أمكن تفسيره بدالة تمييزية وفسرت الدالة التمييزية 100% من التباين ، ومعامل ارتباطها القانوني بلغ (0.955) ، كما تم استخدام مقياس ولكس (Wilks-Criteria) وفق الصيغة (2-13) حيث نلاحظ أن قيمة هذا المقياس قريبة من الصفر ، ايضا مقياس  $(\chi^2)$  وفق الصيغة (2-14) حيث نلاحظ معنوية هذا المقياس وعليه فإنه يمكن القول أن الدالة التمييزية تكفي لتفسير التباين بين المجموعتين وهذا يثبت إمكانية المعادلة التمييزية على تصوير نموذج خطي أمثل مكون من مجموعة من المتغيرات (مصادر الدخل الفردي) قادر على التمييز بين المجموعتين.

### 3-2-3 تفسير المعاملات التمييزية المعيارية

تم تقدير معاملات الدالة المميزة المعيارية وفقاً للصيغة (2-7) وكانت قيمها كما هو موضح في الجدول (3-5) ، وللمعاملات التمييزية المعيارية أهمية تحليلية كبيرة ، حيث يعبر المعامل التمييزي المعياري للمتغير عن مقدار مساهمته النسبية في المعادلة التمييزية وهذا يعني أن مساهمة المتغير في المعادلة التمييزية تكون كبيرة إذا ما كانت القيمة المطلقة لمعامله كبيرة وتدل إشارة المعامل التمييزي المعياري على مساهمة المتغير سالبة كانت او موجبة.

وبالنظر إلى الجدول (3-5) الذي يبين الأوزان التمييزية المميزة للمتغيرات في الدالة التمييزية نلاحظ أن المتغير X3 (المباع من الاسماك) ذو تأثير عال قياساً بالمتغيرات الأخرى وكان يساهم مساهمة سلبية في تمييز المحافظات يليه المتغير X1 (الاجور والمرتببات) حيث يساهم إيجابياً في تمييز المحافظات ويليهما المتغيرات X6 (المنح الدراسية والمساعدات المالية) ، X5 (القيمة التقديرية لإيجار المسكن المملوك للأسرة) و X4 (العائد من أنشطة الاعمال الخاصة ) بينما كانت أقل مساهمة للمتغيرات X2 (المباع من كمية الانتاج الزراعي والحيواني) ، X7 (مصادر الدخل الأخرى النقدية والعينية) و X8 (قيمة الاستهلاك من انتاج الأسرة والمهداة).

### 3-2-4 حساب احتمال التصنيف الصحيح

سوف نتحقق من تصنيف المحافظات باستخدام التحليل التمييزي فيما إذا كانت تقع فعلاً ضمن المجموعة المصنفة إليها أم أن المحافظة تعود إلى المجموعة الأخرى أي كان تصنيف المحافظة خاطئاً ، وبالاعتماد على متوسطي المجموعتين ( $M_1 = -1.484$  ،  $M_2 = 6.305$ ) تم حساب نقطة وسط المجموعتين ( $L$ ) التي تجعل احتمال التصنيف الخاطئ أقل ما يمكن بتطبيق الصيغة (2-8) ، واعتماداً على هذه النقطة تم الحصول على نتائج التصنيف للمحافظات كما هو موضح في الجدول (3-6) ، ويتضح من الجدول (3-6) أن نسبة التصنيف الصحيح لكل الجامعات هي ( 100% ) أي أن (21) محافظة من أصل (21) قد صنفت بشكل صحيح إلى المجموعة التي تنتمي إليها.

الجداول والأشكال

جدول (1-3) تجميع المحافظات حسب مراحل التحليل العنقودي

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2 أبين	15 لحج	6756.000	0	0	5
2	1 إب	5 تعز	11015.000	0	0	5
3	7 حجة	17 المحويت	18655.000	0	0	12
4	10 ذمار	19 عمران	22770.000	0	0	7
5	1 إب	2 أبين	24626.000	2	1	10
6	4 البيضاء	9 حضرموت	27108.000	0	0	12
7	10 ذمار	12 صعدة	27406.000	4	0	14

8	6 الجوف	20 الضالع	28400.000	0	0	11
9	8 الحديدية	16 مأرب	36439.000	0	0	13
10	1 إب	21 ريمة	36669.000	5	0	16
11	6 الجوف	13 صنعاء	42818.000	8	0	14
12	4 البيضاء	7 حجة	46697.000	6	3	13
13	4 البيضاء	8 الحديدية	58798.000	12	9	18
14	6 الجوف	10 ذمار	60212.000	11	7	16
15	11 شبوة	14 عدن	66399.000	0	0	17
16	1 إب	6 الجوف	69525.000	10	14	18
17	3 الأمانة	11 شبوة	87119.000	0	15	19
18	1 إب	4 البيضاء	97771.000	16	13	20
19	3 الأمانة	18 المهرة	141533.000	17	0	20
20	1 إب	3 الأمانة	228824.000	18	19	0

جدول (2-3) المتوسطات الحسابية لمصادر الدخل للمجموعتين

المتوسط	X1	X2	X3	X4	X5	X6	X7	X8
المجموعة (1)	43925.9	11503.5	688.2	17215.9	6116.7	110.2	9222.4	13541.2
المجموعة (2)	99109.8	2586.3	6233.8	36787.0	15012.8	223.0	21505.0	8573.0

جدول (3-3) إختبار F لكل متغير في الدالة المميزة الخطية

Variables	Wilks' Lambda	F	df1	df2	Sig.
X1	.388	29.989	1	19	.000
X2	.769	5.708	1	19	.027

X3	.763	5.916	1	19	.025
X4	.728	7.094	1	19	.015
X5	.354	34.635	1	19	.000
X6	.937	1.280	1	19	.272
X7	.548	15.652	1	19	.001
X8	.885	2.475	1	19	.132

جدول (3-4) المقاييس المستخدمة في اختبار الدالة المميزة

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
	10.339 <sup>a</sup>	100.0	100.0	.955
Test of Function(s)	Wilks' Lambda	Chi-square	Df	Sig.
	.088	36.423	8	.000

جدول (3-5) معاملات الدالة المميزة المعيارية

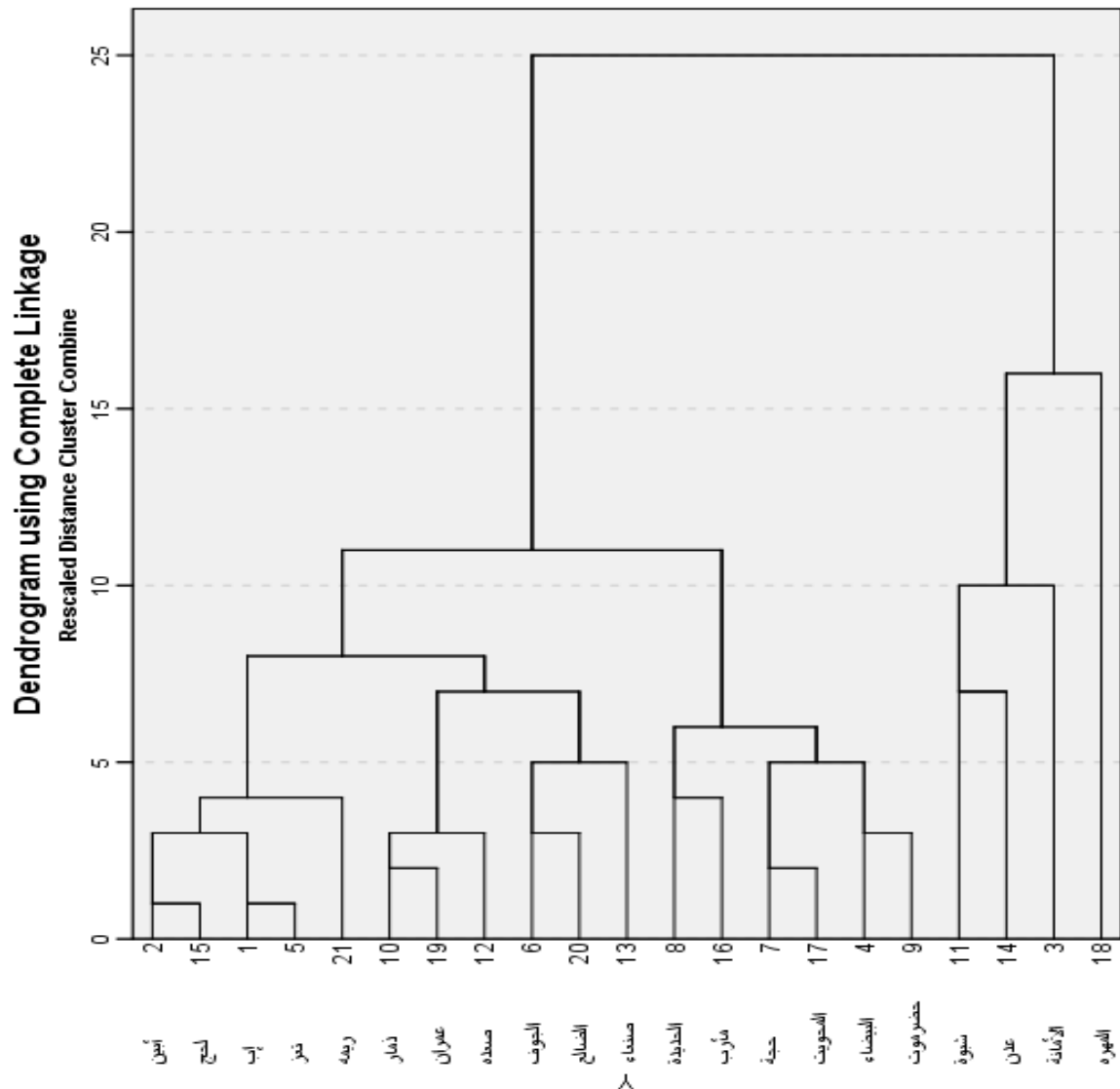
Variables	X1	X2	X3	X4	X5	X6	X7	X8
coefficients	2.086	.524	-2.248-	.776	1.253	-1.524-	.361	-.103-

جدول (3-6) نتائج التصنيف

Group		Predicted Group Membership		Total
		1.00	2.00	
Original	Count	1.00	2.00	
		17	0	17
		0	4	4
	%	100.0	.0	100.0
		2.00	1.00	
		.0	100.0	100.0

100.0% of original grouped cases correctly classified

شكل (3-1) الشجرة البيانية لطريقة الربط الشامل للربط بين المحافظات



المصدر : مخرجات البرنامج الاحصائي SPSS

#### الفصل الرابع : الاستنتاجات والتوصيات

من خلال الدراسة التطبيقية لغرض تطبيق أسلوب التحليل العنقودي متعدد المتغيرات في تصنيف المحافظات اليمينية إلى مجموعات متجانسة بحسب متوسط دخل الفرد السنوي بالاعتماد على مصادر الدخل المختلفة وتحديد مدى التباين والتقارب بين هذه المحافظات وإلى استخدام التحليل التمييزي متعدد المتغيرات لتمييز مصادر الدخل المتسببة في مثل هذا التفاوت توصل الباحث إلى جملة من الاستنتاجات والتوصيات.

#### 1-4 الاستنتاجات

- 1- من خلال استخدام اسلوب التحليل العنقودي في تصنيف المحافظات اليمينية إلى مجموعات متجانسة بحسب مصادر الدخل الفردي أمكن تصنيفها إلى عنقودين متجانسين.
- 2- هناك تقارباً في مصادر الدخل الفردي بين (17) محافظة شكلت العنقود الاول وهي المحافظات ذات الدخل المنخفض وهي أبين ،لحج ،إب ،تعز ،ريمه ، ذمار ،عمران ،صعده ،الجوف ،الضالع ،صنعاء ،الحديدة ،مأرب ،حجة ،المحويت ،البيضاء ،حضر موت.
- 3- هناك تقارباً في مصادر الدخل الفردي بين أربع محافظات شكلت العنقود الثاني وهي المحافظات ذات الدخل المرتفع وهي شبوة ،عدن ،الامانة ،المهرة.
- 4- عند استخدام التحليل التمييزي في تمييز مصادر الدخل الفردي المتسببة في عدم التجانس بين المحافظات تم استخدام المجاميع المستخلصة من نتائج التحليل العنقودي في التحليل التمييزي.
- 5- تبين أن المتغير X3(المباع من الأسماك) ذو تأثير عال قياساً بالمتغيرات الأخرى وكان يساهم مساهمة سلبية في تمييز المحافظات يليه المتغير X1 (الاجور والمرتبآت) حيث يساهم إيجابياً في تمييز المحافظات ويليها المتغيرات X6(المنح الدراسية والمساعدات المالية ) و X5(القيمة التقديرية لإيجار المسكن المملوك للأسرة) و X4 (العائد من أنشطة الأعمال الخاصة ) بينما كانت اقل مساهمة للمتغيرات X2 (المباع من كمية الانتاج الزراعي والحيواني) و X7(مصادر الدخل الاخرى النقدية والعينية) و X8 (قيمة الاستهلاك من انتاج الأسرة والمهداة).
- 6- نسبة التصنيف الصحيح لكل المجاميع هي ( 100% ) أي أن (21) محافظة من أصل (21) قد صنفت بشكل صحيح إلى المجموعة التي تنتمي إليها.

#### 2-4 التوصيات

- 1- حيث أن المتغيرات X1 (الاجور والمرتبآت)، X3 (المباع من الأسماك)، و X4 (العائد من أنشطة الأعمال الخاصة )، X5 (القيمة التقديرية لإيجار المسكن المملوك للأسرة)، X6 (المنح الدراسية والمساعدات المالية ) هي مصادر الدخل الفردي التي ساهمت بدرجة كبيرة في التمييز بين المحافظات، يوصي الباحث بأن يتم تقليل الفوارق في مصادر الدخل بين المحافظات من خلال ايجاد فرص عمل متساوية لكل المحافظات وتوفير أساليب الاصطياد الحديثة في المحافظات



- الساحلية أيضاً تشجيع الاستثمارات الخاصة في المحافظات ذات الدخل المنخفض فضلاً عن توزيع المنح الدراسية والمساعدات المالية بشكل عادل بين المحافظات.
- 2- استخدام أسلوب التحليل العنقودي في تصنيف المحافظات والبلدان بشكل عام بحسب مقياس التنمية البشرية.
- 3- استخدام أسلوب التحليل العنقودي والتحليل التمييزي في تصنيف كثير من الظواهر الاقتصادية والاجتماعية والصحية وغيرها من الظواهر.
- 4- يوصي الباحث القائم في الجهاز المركزي للإحصاء باستخدام الأساليب الإحصائية المتقدمة لدراسة متوسط دخل الفرد والأسرة ونفقاتها وغيرها من الدراسات المهمة.
- 5- تحديث هذه الدراسة في حالة تنفيذ مسح حديث لميزانية الأسرة.

#### المصادر العربية

- 1- الجاعوني، فريد و غانم، عدنان "التحليل الإحصائي متعدد المتغيرات (التحليل التمييزي) في توصيف وتوزيع الأسر داخل الهيكل الاقتصادي الاجتماعي في المجتمع"، مجلة جامعة دمشق للعلوم الاقتصادية والقانونية، المجلد (23) ،العدد(2)، ص ص (313-331)،2007.

- 2- الجبوري، شلال وحمزة، صلاح "تحليل متعدد المتغيرات"، دار الكتب لجامعة بغداد، بغداد، العراق، 2000.
- 3- الجهاز المركزي للإحصاء، مخرجات مسح ميزانية الاسرة متعددة الأغراض (2006).
- 4- الحنيطي، دوخي وآخرون "تميز الأسر الفقيرة من غير الفقيرة في المناطق النائية التابعة لإقليم جنوب الأردن"، مجلة التنمية والسياسات الإقتصادية، المجلد (7)، العدد (1)، ص ص (1-32)، 2004.
- 5- الراوي، خاشع "المدخل الى تحليل الانحدار"، وزارة التعليم العالي والبحث العلمي، جامعة الموصل، الموصل، العراق، 1987.
- 6- جودة، محفوظ "التحليل الاحصائي المتقدم باستخدام SPSS"، دار وائل للنشر، الطبعة الأولى، عمان، الأردن، 2008.
- 7- خواجه، خالد زهدي "أساليب تحليل بيانات دخل ونفقات الأسرة" المعهد العربي للتدريب والبحوث الاحصائية، بغداد، العراق، 2005.
- 8- نون، يونس نون "استخدام طريقتي التحليل العنقودي والتحليل التمييزي في التصنيف مع تطبيق على نتائج الدرجات العالمية"، مجلة تكريت للعلوم الإدارية والاقتصادية، جامعة تكريت، المجلد (8)، العدد (25)، ص ص (160-172)، 2012.
- 9- عزام، عبد المرضي حامد "التحليل الاحصائي للمتغيرات المتعددة من الوجة التطبيقية" كتاب مترجم، دار المريخ للنشر، الرياض، المملكة العربية السعودية، 1998.
- 10- فهمي، محمد شامل بهاء الدين "الاحصاء بلا معاناة، المفاهيم والتطبيقات باستخدام برنامج (SPSS)"، دار الطباعة والنشر العامة، معهد الادارة العامة، الرياض، المملكة العربية السعودية، 2005.
- 11- مصطفى، نزار "استخدام بعض طرق التحليل العنقودي في التصنيف مع تطبيق عملي"، مجلة التقني، الكلية التقنية الادارية، بغداد، المجلد (20)، العدد (2)، ص ص (32-43)، 2007.

12- نامق، فيصل ناجي "اسلوب التحليل العنقودي لتصنيف الاتفاقيات على السلع والخدمات الاساسية وفقا للمستوى البيئي (حضر وريف ) للسنوات 1971 - 2007"، مجلة كلية بغداد للعلوم الاقتصادية الجامعة ، العدد (25) ، ص ص (331 -352)، 2010.

المصادر الأجنبية

- 1- AL-Osaimy M.H & Bamakhramah A.S "An Early Warning System for Islamic Banks Performance" , J.KAU, Islamic Econ., Economics Department , Faculty of Economics and Administration, King Abdul-Aziz University , Jeddah , Saudi Arabia, Vol. (17), No. (1), pp. (3-14),2004.
- 2- Fazzo,L. et al "Cluster analysis of mortality and malformations in the Provinces of Naples and Caserta (Campania Region)", Ann Ist Super Sanità , Rome, Italy, Vol. (44), No. (1), PP(99-111),2008.
- 3- Gumparthi,S. & Manicka,V "Risk Classification Based on Discriminant Analysis For Smes" ,International Journal of Trade, Economics and Finance, Department of Corporate Secretaryship, Alagappa University,Karaikudi, India, Vol.(1), No.(3) , pp(242-246),2010.
- 4- Gwo,F. L &, Chun,M. W "Performing cluster analysis and discrimination analysis of hydrological factors in one step", Advances in Water Resources, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan ,Vol.(29) , pp (1573-1585),2006.
- 5- Hardle, W.& Simar, L "Applied Multivariate Statistical Analysis" Berlin and Louvain-la-Neuve, Germany,2003.
- 6- Muntaner ,C. et al "Hierarchical cluster analysis of labour market regulations and population health: a taxonomy of low- and middle-income countries" BMC Public Health

,Department of Healthcare Management, Korea University, Seoul, Republic of Korea, pp(1–15),2012.

7– RENCHER, A. C "Methods of Multivariate Analysis", John Wiley & sons, New York , USA,1995.

8– RENCHER, A. C "Methods of Multivariate Analysis" , Second Edition, John Wiley & sons, New York , USA,2002.

9– Timm, N.H "Applied Multivariate Analysis" Springer–Verlag ,New York, Inc. USA,2002.

### الملحق

جدول (1م) متوسط الدخل السنوي للفرد بحسب مصادر الدخل المختلفة لمحافظة الجمهورية اليمنية

قيمة الأستهلاك من انتاج الأسرة والمهداه	مصادر الدخل الأخرى النقدية والعينية	المنح الدراسية والمساعدات المالية	القيمة التقديرية لايجار المسكن المملوك للأسرة	العائد من انشطة الاعمال الخاصة	المباع من الأسماك	المباع من كمية الانتاج الزراعي والحيوانات	الاجور والمرتبات	المحافظات
15,080	17,168	120	6,028	23,112	0	5,524	40,539	إب
10,329	14,675	11	6,740	11,536	57	5,016	44,569	أبين
8,440	24,700	792	18,075	73,430	0	4,701	84,531	الأمانة
6,454	11,012	206	6,605	11,442	0	14,919	61,010	البيضاء
13,732	15,716	77	6,313	16,436	5	5,694	39,503	تعز
6,528	4,942	41	5,425	6,317	0	15,334	29,569	الجوف
9,927	2,653	52	5,004	5,937	56	11,658	60,029	حجة
8,405	5,825	24	7,907	35,757	7,732	10,497	59,134	الحديدة
6,866	13,611	157	8,158	16,979	3,850	1,993	61,192	حضر موت
24,341	8,293	44	5,578	22,037	0	10,984	33,211	نمار
7,033	9,598	11	5,674	29,959	0	3,744	88,112	شبو
24,609	4,076	32	4,740	31,451	0	18,557	36,859	صعدة
20,355	4,609	120	7,558	13,247	0	31,445	26,164	صنعاء
5,148	22,102	15	19,606	13,646	5,475	6	75,564	عدن
12,564	14,365	80	6,748	9,810	0	7,053	44,255	لحج
11,407	5,201	181	6,162	31,244	0	8,739	42,226	مأرب
14,961	5,199	422	5,172	8,349	0	10,602	67,042	المحويت
13,671	29,620	74	16,696	30,113	19,460	1,894	148,232	المهرة
14,759	5,714	12	5,268	23,888	0	16,899	30,710	عمران
8,664	14,167	227	6,440	13,619	0	17,634	23,333	الضالع
21,220	9,554	68	4,137	11,509	0	3,011	47,395	ريمة

المصدر : الجهاز المركزي للإحصاء