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

د.فؤاد عبده إسماعيل المخلافي

قسم الإحصاء

كلية التجارة والاقتصاد

جامعة صنعاء

ملخص البحث

تهتم الجمهورية اليمنية كغيرها من الدول بتقليص التفاوت في توزيع الدخول بين السكان وبتحقيق 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 الاحصائي 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 (١٧) محافظة شكلت العنقود الاول وهي المحافظات ذات الدخل المنخفض وهي أبين، لحج ، إب ، تعز ، ريمه ، ذمار عمران ، صعده ، الجوف ، الضالع ، صنعآء ،الحديدة ، مأرب ،حجة ، المحويت ، البيضاء ، حضرموت ، ايضاً 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 انشطة 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

D. Fuad Abdu Ismail Al-Mikhlafi

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

١-١ المقدمة

إعتاد علماء الاقتصاد قياس رفاه البلدان إعتماداً 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 رفاه 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 رفاه 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 سلّم 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 الدخول بين السكان وبتحقيق 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 الاحصائي 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.

١-٢ مشكلة البحث

إن معظم الدراسات والأبحاث التي تقوم بدراسة التفاوت في توزيع الدخول اعتمدت على منحنى لورنس ومعامل جيني وغيرها ، وهذه المقاييس تعتمد على توزع 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.

١-٣ 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.

١-٤_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.

١-٥_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 اليمنية 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.

١-٦ الدراسات السابقة

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

دراسة (الحنيطي واخرون،٢٠٠٤) ، 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 ، ٢٠٠٧) ، 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، ٢٠١٠) ، 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 ، ٢٠١٢) ، حيث تم استخدام طريقتي التحليل العنقودي والتحليل التمييزي في التصنيف مع التطبيق على نتائج الدرجات العاملية التي تمثل نتائج طلبة قسم انظمة 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.

١-٧ فرضيات البحث

تم صياغة الفرضيتين التاليتين:

١- لا يوجد تجانس في توزيع الدخل الفردي بين المحافظات اليمنية.

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

١-٨ حدود البحث

تم الاعتماد على مخرجات مسح ميزانية الاسرة ٢٠٠٦ الصادر عن الجهاز المركزي للإحصاء وهو آخر مسح تم تنفيذه حتى كتابة هذا البحث علماً أن مثل هذه المسوح يتم تنفيذها كل ٨ إلى ١٠ سنوات.

الفصل الثاني : الجانب النظري

٢-١ 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 أخر (جودة ، ٢٠٠٨، ص29).

٢-١-١ المفاهيم الأساسية في التحليل العنقودي

٢-١-١-١ العنصر (Element)(الجبوري وحمزة ، ٢٠٠٠،ص ٨٢٠)

العنصر X_i هو متجه في فراغ المقياس ل n من الابعاد.

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

والعناصر هي قيم عددية لكميات ممكنة القياس.

٢-١-١-٢ المسافة (D) (Distance)

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

(٢-٢)

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

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

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.

٢-١-١-٣ 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.

٢-١-١-٤ الشجرة البيانية (Tree diagram or dendogram)(Timm, 2002,P523)

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

١- 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.

٢- 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 فأن 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.

٢-١-٢ خطوات التعنقد(Clustering Steps)(Timm , ٢٠٠٢ , p523)

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

٢- يتم ربط العنصرين اللذين تكون المسافة بينهما أقصر المسافات ضمن المصفوفة المحسوبة في الخطوة السابقة ، وفي حالة وجود مسافات متساوية فبالإمكان أجراء عملية الربط لأكثر من عنصرين في مرحلة واحدة (لعنصرين معاً).

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

٤- 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.

٢-١-٣ 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(فهمي، 2005،ص) 813، 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 تناسب العينات الصغيرة نسبياً ، وفي التحليل العنقودي المتسلسل توجد طرق مختلفة وسنركز على طرق الإدماج المتسلسل وخصوصا ما يسمى بطرق الربط (methods Linkage)وهذه الطرق تلائم تجميع الحالات كما تلائم تجميع المتغيرات وهذا مالا يتحقق لطرق التجميع الأخرى (عزام ،١٩٩٨ ،ص ص ٨٦٥-٨٦٦).

٢-١-٣-١ طريقة الربط المنفرد (Single Linkage) وتسمى ايضاً بطريقة الجوار الأقرب (neighbor nearest)

تعتمد هذه الطريقة بالأساس على اعتبار أن العنصرين الاكثر تشابهاً بين العناصر يشكل نواة العنقود، ثم تضاف باقي الوحدات إلى هذه النواة بالتسلسل وحسب درجة الشبه مع عناصر نواة العنقود إذ يضاف الاكثر شبهاً ثم الاقل وبالتدريج وفي حالة ربط مجموعة من العناقيد مع بعضها فأن ذلك يتم بالاستناد على اقرب المسافات أو معاملات التماثل بين ازواج العناصر وحسب الصيغة التالية: $D_{IJ}=Min(D_{ij})$ (٣-٢)

حيث i,j تمثل العناصر في العناقيد I,J على التوالي.

٢-١-٣-٢ طريقة الربط الشامل (Complete Linkage) وتسمى ايضاً بطريقة الجوار الأبعد) (Rencher,2002,PP459- 460)farthest neighbor)

تعتمد هذه الطريقة بالأساس على اعتبار أن العنصرين الاكثر 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 اياً من عناصر العنقود هي اكبر

من المسافة بين اي عنصر آخر وبين اياً من عناصر العنقود ويتم الربط حسب الصيغة الاتية:

$$D_{IJ}=Max\ (D_{ij}) \qquad (٤-٢)$$

حيث i,j تمثل العناصر في العناقيد I,J 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).

٢-٢ التحليل التمييزي (Discriminant Analysis)

التحليل التمييزي هو ايضاً أسلوب إحصائي لتحليل البيانات متعددة المتغيرات ،حيث يهتم بمسألة التمييز بين مجموعتين أو اكثر والتي تكون متشابهه 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 ، ١٩٨٧ ،ص٥١٠).

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 ،٢٠٠٠).

٢-٢-١ 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($X1, X2, \ldots\ldots, Xp$) فان الصيغة العامة للدالة المميزة هي كالآتي (الراوي، ١٩٨٧،ص٥١١):

$$Y_i = \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots + \beta_p X_p \quad , \quad i = 1,2,\ldots r \qquad (٢-٥)$$

حيث:

P : عدد المتغيرات الداخلة في الدالة .

β : معاملات الدالة المميزة المعيارية.

r : عدد الدوال المميزة.

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

إن عملية التقدير للمعلمات (β,S) والتي تجعل الدالة تعطي أفضل تمييز بين المجموعتين لابد أن يتم من خلال جعل مربع الفرق بين متوسطي المجموعتين إلى التباين المشترك للمجموعتين أكبر ما يمكن(Rencher,1995,p297) أي أن : (٢-٦)

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

حيث نقدر معلمات الدالة المميزة من خلال تعظيم النسبة Q باشتقاقها جزئيا ومساواتها بالصفر ونحصل على:

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

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

وبعد استخراج المعاملات $\hat{\beta}$ تصنف المفردة إلى إحدى المجموعتين بالاعتماد على نقطة وسط المجموعتين (L) التي تجعل احتمال التصنيف الخاطئ أقل ما يمكن (عزام ، ١٩٩٨ ، ص ٧٤٢).

$$L = \frac{\bar{Y}_1 + \bar{Y}_2}{2} \qquad (٨-٢)$$

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 :

(٩-٢)

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

٢-٢-١-١ 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):

$$T^2 = \frac{n1\ n2}{n1+n2} D^2 \qquad (10-2)$$

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

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

(١١-٢)

ويستخدم اختبار (F) وصيغته كما يلي :

$$F = \frac{n1+n2-p-1}{(n1+n2-2)p} T^2 \qquad (12-2)$$

بدرجة حرية (P,n1+n2−p−1) ، نرفض H_0 بمستوى معنوية F_α اذا كانت:

$$F_{cal} > F_{\alpha,(P,n1+n2-p-1)}$$

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 التالية(Rencher, 1995, P180):

$$\Lambda = \frac{|W|}{|T|} \qquad (١٣-٢)$$

حيث :

T: مصفوفة التباين والتغاير الكلي للمجموعات.

W : مصفوفة التباين والتغاير داخل المجموعات.

وتتراوح قيمة (Λ) 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 (χ^2) ويعد هذا المقياس أكثر دقة من مقياس (Λ) وصيغته كما يلي :

$$\chi^2 = - Log(\Lambda) \qquad (14-2)$$

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

٢-٢-١-٢ 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.

٢- 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):

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

حيث ϕ : 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 وبهدف إيجاد تجمعات من المحافظات تكون متجانسة فيما بينها من أجل التصنيف والمقارنة في آن واحد كما سيتم تطبيق اسلوب 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 (م١)، ، وسيتم استخدام البرنامج الإحصائي(SPSS version 20) في كل مراحل التحليل.

٣-١ تحليل نتائج التحليل العنقودي

تم استخدام طريقة الربط الشامل لإيجاد العلاقة بين المحافظات بالاعتمادعلى مقياس التشابه المستخدم وهو معامل منكوسكي (Minkowski metric) بتطبيق الصيغة (٢-٢) عندما r=1 والنتائج موضحة بالجدول (3-1) والتي تمثل مصفوفة القرابة (Proximity matrix) والتي تقرر التشابه أو عدم التشابه بين المحافظات ويعبر عنها بالمسافات المشتقة بين كل محافظة مع المحافظات الاخرى .

من خلال الجدول (٣-١) نلاحظ أن محافظتي أبين ولحج شكلتا أول تعنقد وقد كان أقل معامل اقتراب (6756.000) لتشكل عنقوداً جديداً في مرحلة جديدة وهي المرحلة اللاحقة (٥) ويلي 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 (١٢) 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 (Dendorogram) 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 (٣-١) ، وفي نهاية تحليل البيانات تم وضع النتائج في عناقيد (Clusters) من حيث التشابه في مصادر الدخل للمحافظات وكما هو واضح في الشكل (٣-١) أن هناك (١٧) محافظة تعنقدت عند مسافات متقاربة وهي أبين ، لحج ، إب ، تعز ، ريمه ، ذمار، عمران ، صعده ،الجوف ،الضالع ،صنعآء ، 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.

٣-٢ 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 (١ يمثل محافظات العنقود الاول ذات الدخل المنخفض) ، (٢ يمثل محافظات العنقود الثاني ذات الدخل المرتفع).

ب– المتغيرات المستقلة تمثلت بمصادر الدخل الفردي وهي X1(الاجور والمرتبات) ، X2 (المباع من كمية الانتاج الزراعي والحيواني) ، X3 (المباع من الأسماك) ، X4 (العائد من انشطة الاعمال الخاصة) ، X5 (القيمة التقديرية لإيجار المسكن المملوك للأسرة) ، X6 (المنح الدراسية والمساعدات المالية) ، X7 (مصادر الدخل الاخرى النقدية والعينية) ، X8 (قيمة الاستهلاك من انتاج الأسرة والمهداة).

٣-٢-١ إختبار معنوية المتغيرات في الدالة المميزة

نحتاج إلى احتساب المتوسطات الحسابية للمتغيرات الداخلة في المجموعات وهي كما في الجدول(٣–٢) 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 (٣–٣)، ونلاحظ من خلال الجدول(٣–٣) 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 المتغيرX7(مصادر الدخل الاخرى النقدية والعينية) ثم X4(العائد من انشطة الاعمال الخاصة) و X3(المباع من الأسماك) وX2 (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 (٣-٤) ، 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(٢-١٣) حيث نلاحظ أن قيمة هذا المقياس قريبة من الصفر ، ايضا مقياس (χ^2) وفق الصيغة (٢-١٤) 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.

٣-٢-٣ تفسير المعاملات التمييزية المعيارية

تم تقدير معاملات الدالة المميزة المعيارية وفقا للصيغة(٢-٧) وكانت قيمها كماهو موضح في الجدول (٣-٥) ، 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 اوموجبة.

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 والمرتبات) حيث يساهم إيجابياً في تمييز المحافظات ويليهما المتغيرات X6(المنح الدراسية والمساعدات المالية) ،X5(القيمة التقديرية لإيجار المسكن المملوك للأسرة) وX4 (العائد من انشطة الاعمال الخاصة) بينما

كانت أقل مساهمة للمتغيرات X2 (المباع من كمية الانتاج الزراعي والحيواني) ،X7(مصادر الدخل الاخرى النقدية والعينية) و X8 (قيمة الاستهلاك من انتاج الأسرة والمهداة).

٤-٢-٣ 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

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	19عمران	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 الحديدة	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 الحديدة	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

خاطئاً ، وبالاعتماد على متوسطي المجموعتين (M_1= −1.484− ، M_2= 6.305) تم حساب نقطة وسط المجموعتين (Z) التي تجعل احتمال التصنيف الخاطئ أقل ما يمكن بتطبيق الصيغة (٢−٨) ، واعتماداً على هذه النقطة تم الحصول على نتائج التصنيف للمحافظات كما هو موضح في الجدول (٣−٦) ، ويتضح من الجدول(٣−٦) أن نسبة التصنيف الصحيح لكل المجاميع هي (١٠٠%) أي أن (٢١) محافظة من أصل (٢١) قد صنفت بشكل صحيح إلى المجموعة التي تنتمي إليها.

الجداول والأشكال

جدول (٣−١) تجميع المحافظات حسب مراحل التحليل العنقودي

جدول(٣−٢) المتوسطات الحسابية لمصادر الدخل للمجموعتين

المتوسط	X1	X2	X3	X4	X5	X6	X7	X8
المجموعة(١)	43925.9	11503.5	688.2	17215.9	6116.7	110.2	9222.4	13541.2
المجموعة(٢)	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 (٣-٤) المقاييس المستخدمة في اختبار الدالة المميزة

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

حدول (3-5) معاملات الدالة المميزة المعيارية

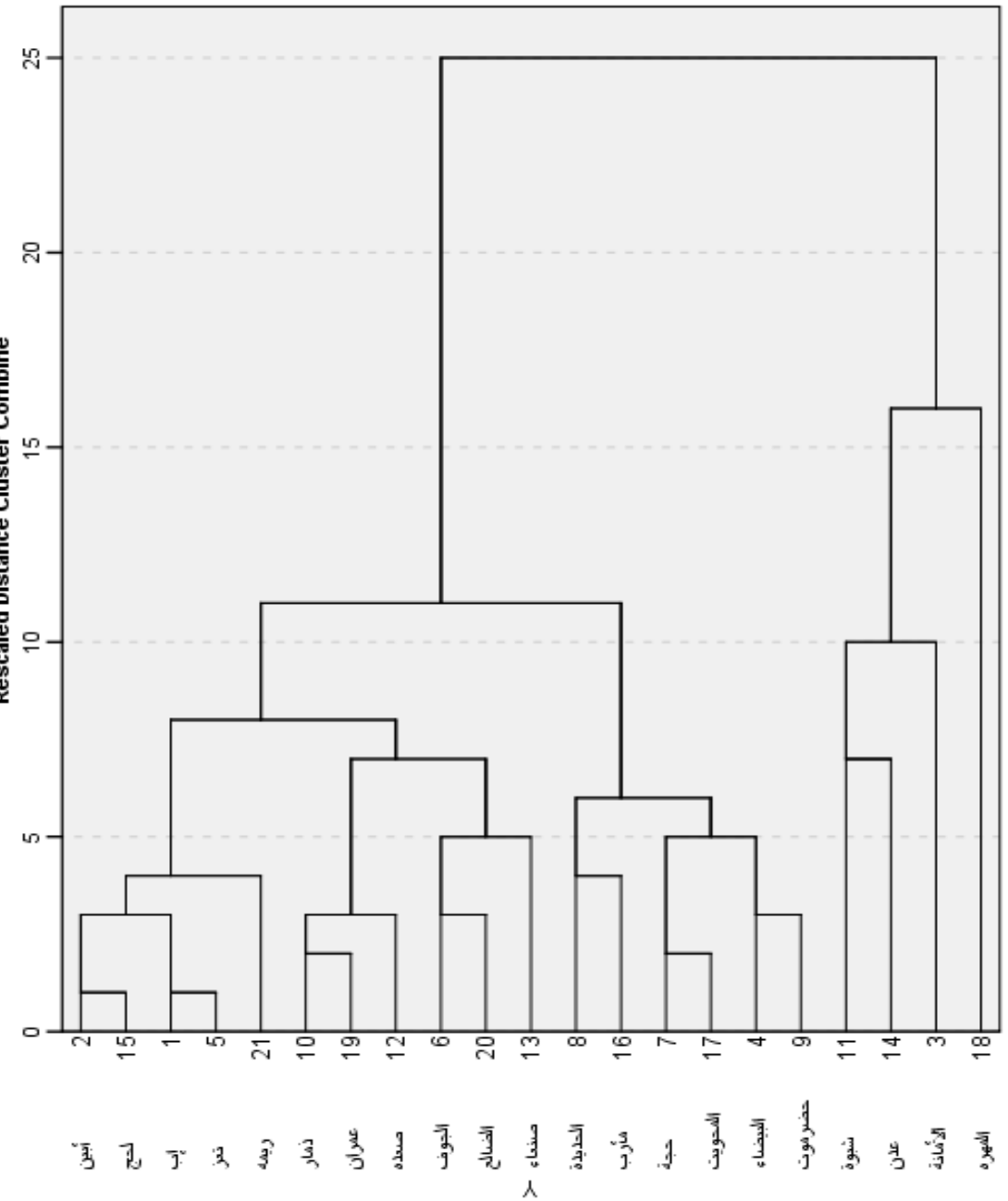
Variables	X1	X2	X3	X4	X5	X6	X7	X8
coefficients	2.086	.524	-2.248-	.776	1.253	-1.524-	.361	-.103-

جدول (٣-٦) نتائج التصنيف

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

100.0% of original grouped cases correctly classified

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


Rescaled Distance Cluster Combine
0
5
10
15
20
25
2 15 1 5 21 10 19 12 6 20 13 8 16 7 17 4 9 11 14 3 18
أمين
لحج
إب
تعز
ريمه
ذمار
عمران
صعده
الجوف
الضالع
صنعاء
الحديدة
مأرب
حجة
المحويت
البيضاء
حضرموت
شبوة
عدن
الأمانة
المهره
Y

المصدر : مخرجات البرنامج الاحصائي SPSS

الفصل الرابع : الاستنتاجات والتوصيات

من خلال الدراسة التطبيقية لغرض تطبيق أسلوب التحليل العنقودي متعدد المتغيرات في تصنيف المحافظات اليمنية إلى مجموعات متجانسة بحسب متوسط دخل الفرد السنوي بالاعتماد على مصادر الدخل المختلفة وتحديد مدى التباعد 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.

٤-١ الاستنتاجات

١- من خلال استخدام اسلوب التحليل العنقودي في تصنيف المحافظات اليمنية إلى مجموعات متجانسة بحسب مصادر الدخل الفردي أمكن تصنيفها إلى عنقودين متجانسين.

٢- هناك تقارباً في مصادر الدخل الفردي بين (١٧) محافظة شكلت العنقود الاول وهي المحافظات ذات الدخل المنخفض وهي أبين ،لحج ،إب ،تعز ،ريمه ، ذمار ،عمران ،صعده ،الجوف ،الضالع ،صنعآء ،الحديدة ،مأرب ،حجة ،المحويت ،البيضاء ،حضرموت.

٣- هناك تقارباً في مصادر الدخل الفردي بين أربع محافظات شكلت العنقود الثاني وهي المحافظات ذات الدخل المرتفع وهي شبوة ،عدن ،الامانة ،المهرة.

٤- 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.

٥- تبين أن المتغير X3(المباع من الأسماك) ذو تأثير عال قياساً بالمتغيرات الأخرى وكان يساهم مساهمة سلبية في تمييز المحافظات يليه المتغير X1 (الاجور والمرتبات) حيث يساهم إيجابياً في تمييز المحافظات ويليهما المتغيرات X6(المنح الدراسية والمساعدات المالية) و X5(القيمة التقديرية لإيجار المسكن المملوك للأسرة) و X4 (العائد من انشطة الأعمال الخاصة)

بينما كانت اقل مساهمة للمتغيرات X2 (المباع من كمية الانتاج الزراعي والحيواني) وX7(مصادر الدخل الاخرى النقدية والعينية) وX8 (قيمة الاستهلاك من انتاج الأسرة والمهداة).

٦- نسبة التصنيف الصحيح لكل المجاميع هي (١٠٠%) أي أن (٢١) محافظة من أصل (٢١) قد صنفت بشكل صحيح إلى المجموعة التي تنتمي إليها.

٤-٢ التوصيات

١- حيث أن المتغيرات X1 (الاجور والمرتبات)،X3 (المباع من الأسماك)، وX4 (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.

٢- إستخدام أسلوب التحليل العنقودي في تصنيف المحافظات والبلدان بشكل عام بحسب مقياس التنمية البشرية.

٣- إستخدام أسلوبي التحليل العنقودي والتحليل التمييزي في تصنيف كثير من الظواهر الاقتصادية والاجتماعية والصحية وغيرها من الظواهر.

4- يوصي الباحث القائمين في الجهاز المركزي للإحصاء باستخدام الأساليب الإحصائية المتقدمة لدراسة متوسط دخل الفرد والأسرة ونفقاتها وغيرها من الدراسات المهمة.

٥- تحديث هذه الدراسة في حالة تنفيذ مسح حديث لميزانية الأسرة.

المصادر العربية

١- 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(٢)، ص ص(٣١٣-٣٣١)،٢٠٠٧.

٢- الجبوري، شلال وحمزة، صلاح "تحليل متعدد المتغيرات"، دار الكتب لجامعة بغداد، بغداد، العراق،٢٠٠٠.

٣- الجهاز المركزي للإحصاء، مخرجات مسح ميزانية الاسرة متعددة الأغراض(٢٠٠٦).

٤- الحنيطي، دوخي واخرون "تمييز الأسر الفقيرة من غير الفقيرة في المناطق النائية التابعة لإقليم جنوب الأردن" ،مجلة التنمية والسياسات الإقتصادية، المجلد (7) ،العدد(١) ،ص ص(١- ٣٢)،٢٠٠٤.

٥- الراوي ، خاشع "المدخل الى تحليل الانحدار"، وزارة التعليم العالي والبحث العلمي ،جامعة الموصل ،الموصل ،العراق،١٩٨٧.

٦- جودة , محفوظ "التحليل الاحصائي المتقدم باستخدام SPSS " ، دار وائل للنشر، الطبعة الأولى ، عمان، الأردن،٢٠٠٨.

٧- خواجه، خالد زهدي "أساليب تحليل بيانات دخل ونفقات الأسرة" المعهد العربي للتدريب والبحوث الاحصائية، بغداد، العراق،٢٠٠٥.

٨- ذنون ، يونس ذنون "استخدام طريقتي التحليل العنقودي والتحليل التمييزي في التصنيف مع تطبيق على نتائج الدرجات العاملية" ،مجلة تكريت للعلوم الإدارية والاقتصادية ، جامعة تكريت، المجلد(٨)،العدد(25)،ص ص(١٦٠-١٧٢)،٢٠١٢.

٩- عزام، عبد المرضي حامد "التحليل الاحصائي للمتغيرات المتعددة من الوجهة التطبيقية" كتاب مترجم، دار المريخ للنشر، الرياض ،المملكة العربية السعودية ،.١٩٩٨

١٠- فهمي ، محمد شامل بهاء الدين "الاحصاء بلا معاناة، المفاهيم والتطبيقات باستخدام برنامج (SPSS)" ، دار الطباعة والنشر العامة ، معهد الادارة العامة،الرياض، المملكة العربية السعودية،٢٠٠٥.

١١- مصطفى ، نزار "استخدام بعض طرق التحليل العنقودي في التصنيف مع تطبيق عملي" ،مجلة التقني ، الكلية التقنية الادارية،بغداد، المجلد(٢٠)، العدد (٢)، ص ص(٣٢-٤٣)،٢٠٠٧.

١٢- نامق، فيصل ناجي "اسلوب التحليل العنقودي لتصنيف الانفاق على السلع والخدمات الاساسية وفقا للمستوى البيئي (حضر وريف) للسنوات ١٩٧١- ٢٠٠٧"، مجلة كلية بغداد للعلوم الاقتصادية الجامعة ، العدد (25) ،ص ص(٣٣١-٣٥٢)،٢٠١٠.

المصادر الأجنبية

١- 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.

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

الملحق

جدول(م١) متوسط الدخل السنوي للفرد بحسب مصادر الدخل المختلفة لمحافظات الجمهورية اليمنية

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

المصدر : الجهاز المركزي للإحصاء