الفصل الثالث

تبويب البيانات الجغرافية

Tabulation Of Geographic Data

إن المنهج الكمي مرتبط أساساً بجمع البيانات وتحليلها بقصد التوصل إلى نتائج وحلول لمشكلات هي قيد الدراسة، وإن نظرة على البيانات الجغرافية تعطي صورة عن مدى ملاءمة هذه البيانات لهذا النوع من مفاهيم الدراسة. كما أن هذا المنهج يقدم أفضل الوسائل الإحصائية والرياضية الحديثة التي تعطي أفضل النتائج وأدقها.

إن البيانات في الجغرافيا تأخذ أشكالاً عدة منها: الارتفاعات وإحصاءات الأمطار والحرارة، ومقدار التصريف النهري، ودرجة الانحدار، وأرقام السكان، وحركة المرور، وقيمة الأراضي، ولا شك أن القائمة إن لم تكن لا نهائية فإنها طويلة للغاية.

والبيانات بشكل عام هي حقائق علمية وفي الوقت نفسه أجزاء من المعلومات التي تشكل المادة الخام للموضوع الذي يتعلق بها، وهذه البيانات على أية حال تتميز بميزتين هامتين هما:

١- إن هذ البيانات دقيقة وتمثل حقائق رقمية، وتتميز المعلومات التي تعطيها بأنها كمية أكثر منها وصفية؛ وكنتيجة لذلك فإنها تخضع للدراسة الإحصائية، وبالإمكان خزنها بسهولة سواءٌ على صورة قوائم على الورق أو أشرطة أو في ذاكرة الحاسوب.

٢- إن هذه البيانات تجمع لغرض خاص ومحدد، إذ إنها تمثل نتيجة قياسات وفحوصات مستهلكة للوقت وباهظة التكاليف في معظم الحالات، وفي العادة يبرر جمعها استخدامها في اختبار نظرية أو فرضية وضعت مسبقاً.

إن البيانات الجغرافية ليس لها ميزات خاصة تميزها عن البيانات التي تستخدمها العلوم الأخرى، والحقيقة أن البيانات ليست الشيء الذي يفرق ويميز بين العلوم المختلفة، بل هي الطريقة التي تفسر وتستعمل بها البيانات والحقائق. وبصورة عامة نستطيع القول إن أية بيانات يمكن أن تكون جغرافية إذا كانت تستطيع أن توضح علاقات أو تحل مشكلات لها طبيعة جغرافية.

ولا يهم بالطبع فيما إذا كانت البيانات نفسها تستقطب انتباه المشتغل بالأرصاد الجوية أو الهيدرولوجي أو الاقتصادي أو المخطط، وبالرغم من ذلك فإن هناك بعض أنواع من البيانات التي لها فوائد للجغرافي أكثر من أي شخص آخر وما ذلك إلا لما لها من ميزات خاصة.

وبالرغم من تنوع البيانات الجغرافية فإن الميزة أو الخاصية التي تجمع معظم أشكالها هي أنها موزعة مكانياً، أي إن قيم هذه البيانات تختلف من مكان لآخر، وهذا واضح في القيم الخاصة بالنقاط (مثل: الارتفاع، هطول المطر، ودرجة الانحدار، معدل تدفق المرور)، أو الخطوط (مثل: كثافة المرور)، أو المسطحات (كثافة السكان، ومقدار إنتاج المحاصيل). إلا أن بعض البيانات تحتوى على قيم تخص المكان نفسه، وفي هذه الحالة تحدث الاختلافات بين القيم عبر الزمان (الإحصاءات المناخية، التصريف

النهري، الإنتاج المعدني وغيرها)[1].

إن البيانات الجغرافية شأنها شأن أي بيانات إحصائية أخرى يمكن إدراجها تحت نوعين أساسيين من أنواع البيانات الإحصائية هما:

١- البيانات النوعية أو الاسمية *Naminal Data*:

وهي البيانات الوصفية غير القابلة للقياس العددي، ولكن يمكن تصنيفها حسب النوع أو الدرجة، كتقسيم الأشياء إلى: صالحة وغير صالحة، أو تقسيم الأفراد إلى أغنياء وفقراء، وأكثر ما تستخدم هذه البيانات في الدراسات البشرية مثل: الحالة الزواجية، والحالة التعليمية، والمهنة، والنشاط الاقتصادي... إلخ.

فمثلاً: الحالة الزواجية تنحصر في الحالات التالية: لم يتزوج أبداً- متزوج- مطلق- أرمل، وكذلك الحال بالنسبة للحالة التعليمية التي يمكن تصنيفها إلى متعلم وغير متعلم. ويمكن تصنيف العمالة إلى عامل وغير عامل، وكذلك المهنة إلى مزارع، فني، تاجر، إلخ...

ويمكن التوسع في هذه التصنيفات حسب غرض الدراسة، والمهم هو أن هذه البينات يتم وضعها في جداول تظهر فيها الصفات مع عدد المشاهدات التي تنتمي لكل صفة (انظر جدول ٣ : ١).

٢- البيانات الرتبية: *Ordinal Rank Data*

تمتاز البيانات الرتبية عن البيانات النوعية بأن الفئات التي يتكون منها مجتمع الدراسة تكون في هيئة رتب، وليس مجرد فئات تميز بين الأنواع

(1) Peter David, "Data Description and Presentation Vol. 3 of Science in Geography, pp. 1-2.

فقط، وأكثر ما تستخدم البيانات الرتبية لقياس درجة الاستجابة بالموافقة أو المعارضة لقضية ما، أو تدرج الخصائص والمميزات، بحيث ترتب بيانات القيم على أساس تسلسلها إما تنازلياً أو تصاعدياً، فمثلاً: لو طرح السؤال التالي على مجموعة من الطلبة لمعرفة رأيهم في استعمال الحاسوب في الجغرافية الكمية وهو: هل توافق على استخدام الحاسوب في الجغرافية الكمية؟

يمكن أن تصنف الإجابات على الشكل التالي:

١- لا أوافق بشـــدة ٢- لا أوافق ٣- لا أعـــرف ٤- أوافق ٥- أوافق بشدة.

وهذا التصنيف للإجابات يحمل في طياته درجات الاستجابة، التي تتراوح بين المعارضة الشديدة والموافقة الشديدة.

ومن الأمثلةالأخرى على البيانات الرتيبة: تصنيف الرواسب النهرية بحسب أحجامها إلى:

١- دقيقة ٢- متوسطة ٣- غليظة. أو تصنيف المنحدرات إلى:

١- خفيفة ٢- متوسطة ٣- شديدة. أو تصنيف المستوى التعليمي إلى:

١- أمي ٢- ابتدائي ٣- ثانوية ٤- جامعي.

جدول (١:٣)

سكان إمارة مكة للعمر ١٢ سنة فأكثر

موزعين حسب الجنس والحالة الزواجية

الحالة الزواجية	الذكور ٪	الإناث ٪	الإجمالي ٪
لم يتزوج أبداً	٤٢	٢٥	٣٥
متزوج	٥٤	٦٠	٥٧
أرمل	٢	١٢	٦
مطلق	٢	٣	٢
المجموع	١٠٠ ٪	١٠٠ ٪	١٠٠ ٪

المصدر: مصلحة الإحصاءات العامة ـ التعداد العام للسكان ١٩٧٤م في منطقة مكة المكرمة، الجدول رقم ٦٥ (النسب من حساب الباحث).

البيانات الرقمية:

وهي البيانات الكمية التي تأخذ قيماً عددية، ويطلق عليها البعض اسم بيانات الفترة Intervalley - Scaled Data وتكون قابلة للقياس العددي مثل: بيانات السن والدخل وعدد أفراد الأسرة وحجم الحيازات الأرضية ودرجات الحرارة وكميات الأمطار وتكون على صورتين:

أ- بيانات كمية غير مبوبة: هي البيانات الفردية التي تمثل قيماً معينة

ومحددة لكل جزء من العينة المأخوذة، كالبيانات التي تجمع رأساً من الميدان، ويمكن أن تكون استمارات حصر السكان وإحصاءاته قبل تقسيمها إلى فئات أو مجموعات نموذجاً لمثل هذه البيانات. ويمثل الجدول (٣-٢) قراءات لكميات الأمطار الساقطة على مدينة ما لمدة ١٠٠ يوم بالمليمتر المكعب.

جدول (٣-٢)

كميات الأمطار الساقطة على مدينة ما خلال ١٠٠ يوم (بالمليمتر المكعب)

٦٠	٣٠	٤٤	٥١	٧٥	٥٧	٥٤	٤٠	٥٠	٥٩
٢١	٥٠	٧١	٧٨	٥٥	٧٢	٥٦	٤١	٥١	٧٠
٦٣	٣١	٧١	٥٣	٥١	٣٦	٧٩	٥٢	٦٧	٣٢
٥٦	٥٧	٤٨	٥٦	٤٥	٤٨	٦٣	٤٣	٦٥	٦٠
٦٥	٣٣	٦٩	٥١	٤١	٤٤	٥٨	٤٨	٥٧	٢٠
٥٨	٣٨	٤٦	٥٠	٨٢	٥٦	٤٧	٤٢	٦٣	٥٤
٦٧	٣٩	٥٦	٦٩	٥٩	٤٢	٣٠	٤٥	٦٢	٥٣
٦٠	٥٢	٨١	٣٢	٥٤	٥٣	٥٨	٤٩	٥٤	٦٦
٥٥	٣٦	٤١	٨٠	٦٨	٢٥	٧١	٤٦	٥٢	٥٣
٢٩	٣٥	٨٥	٤٦	٥٩	٥٨	٥٠	٤٤	٦١	٦١

ملاحظة: الأرقام السابقة توضيحية فقط وليست حقيقية.

إن هذه البيانات لا تحتوي على تكرارات. ولا يمكن استخدامها على حالتها الحاضرة دون معالجة خاصة لاستخراج المقاييس الإحصائية اللازمة.

ب) بيانات كمية مبوبة Grouped Data: إذا كانت البيانات بالصورة الأولية التي جمعت من الميدان كالمعلومات الموجودة في الجدول السابق، لا بد من تصنيفها وتقسيمها إلى فئات متساوية أو متفاوتة المدى لتسهيل دراستها وفهمها.

ولتبويب البيانات الكمية في جداول، تقسم البيانات إلى مجموعات متشابهة تسمى فئات، ونضع في كل فئة المفردات التي تنتمي إليها (التكرارات)، ويسمى التوزيع الناتج من عملية تفريغ البيانات الأولية بالتوزيع التكراري Frequency Distribution، ويسمى الجدول الذي يضم هذا التوزيع بالجدول التكراري Frequencey Table، ويمثل الجدول (٣-٤) بيانات كميات الأمطار الساقطة على مدينة ما الموجودة في جدول (٣-٢) بعد تقسيمها إلى مجموعات تكرارية متماثلة.

وهنا لا بد من التفريق في البيانات الكمية بين نوعين من الظاهرات المدروسة أو المتغيرات وهما:

أ- البيانات المتصلة: وتسمى بالمتغيرات السيارة Continuous: وهي المتغيرات التي تأخذ جميع القيم بين حدي التغير، وإذا زاد من رقم إلى الرقم التالي له يمر بكل الكسور الممكنة بينهما، فمثلاً: الزمن يعتبر متغيراً سياراً أو متصلاً؛ لأنه يمكن أن يقاس بالثواني أو أجزاء من الثانية مهما بلغت الدقة.

وأيضاً درجة الحرارة والرطوبة وكمية المطر أو مبالغ المشتريات والمبيعات أو صافي الربح وكذلك أطوال وأوزان وأعمار الظاهرات الطبيعية والبشرية.

ب- البيانات المنفصلة: وتسمى بالمتغيرات الوثابة Discrete وهي المتغيرات التي لا تمكنها طبيعها من أن تأخذ جميع القيم بين حدي المتغير بل تتحرك عند إعداد معينة دون سواها. فمثلاً عدد أفراد الأسرة يأخذ أرقاماً صحيحة ١، ٢، ٣، .. إلخ، ولكنه لا يأخذ أرقاماً كسرية ولا يعقل أن يكون عدد الأفراد ٥، ٣ فرداً مثلاً. وما يقال عن الأفراد يقال عن عدد غرف المسكن، وعدد العمال في مصنع، وعدد الحوادث على الطرق وغيرها من المتغيرات الوثابة.

٣- تصنف البيانات الكمية بحسب أبعادها الزمانية والمكانية إلى مجموعتين رئيسيتين هما:

أ- البيانات الزمانية Temporal Data: وتشمل جميع البيانات التي تكون منتظمة في إطار زماني معين: كدرجات الحرارة خلال فترة معينة، أو كميات الأمطار الساقطة في مواسم معينة، أو مقدار إنتاج البترول في بلد ما خلال فترة زمنية محددة. ومن هنا ترتبط البيانات بعنصر الزمن، ويطلق الجغرافيون على هذا النوع من البيانات اسم: (السلاسل الزمنية)؛ لأن البيانات تنتظم ضمن عدد محدد من السنوات، أو الأيام، أو المواسم أو غير ذلك من الأمور المرتبطة بعنصر الزمن.

ب- البيانات المكانية Spatial Data: وتشمل البيانات التي تكون منتظمة في إطار مكاني معين، بحيث لا يمكن فهمها وتفسيرها إلا في

حدود ذلك الإطار، وتأخذ البيانات المكانية ثلاثة أشكال متميزة:

١- البيانات النقطية Point Data: وهي التي تمثل مواقع الظاهرات الجغرافية التي تمثل نقاطاً على الخرائط، مثل: موقع مدينة أو موقع بئر، أو موقع محطة بترول، وكذلك الحال في مواقع المدارس والمصانع والقرى... إلخ.

٢- البيانات الخطية Line Data: وهي البينات التي تأخذ شكل الخط على الخرائط مثل: خطوط شبكات المواصلات، وخطوط الحدود السياسية، وخطوط شبكات المياه: كالأنهار والروافد وجميع أنواع الظاهرت التي يمكن أن يكون لها طول معين.

٣- البيانات المساحية Areas وهي البيانات المرتبطة بمساحة الظاهرات المكانية. ومن الجدير بالذكر أن المساحات يمكن تحديدها بخطوط مثل: الأقاليم الزراعية، أو المناطق العمرانية، أو الوحدات الإدارية، أو المساحة التي يمتد عليها مطار معين أو بحيرة أو حديقة ... إلخ.

جدولة البيانات:

من الطبيعي أنه لا توجد طريقة واحدة لتقسيم البيانات إلى مجموعات متدرجة وتصنيفها تبعاً لذلك. وعملية تحديد عدد الفئات وطول كل منها تتوقف على الهدف الذي يرمي إليه الباحث من هذا التصنيف، والطريقة المتبعة في الغالب هي طرح أصغر قيمة من أكبر قيمة في البيانات المراد تلخيصها، فنحصل على ما يسمى بالمدى المطلق للبيانات، وهو المدى الذي تنتشر فيه بيانات البحث، ثم نقسمه إلى عدد مناسب من الفئات آخذين في الاعتبار ما يلي:

١ - ألا يكون طول الفئة كبيراً وبالتالي عدد الفئات صغيراً فتضيع الظاهرة المدروسة .

٢- ألا يكون طول الفئة صغيراً وبالتالي عدد الفئات كبيراً، فينتفي الهدف من تلخيص البيانات في فئات .

إن عدد الفئات يتوقف على عدة عوامل أهمها :

١ - المدى المطلق للظاهرة المدروسة، أي الفرق بين أصغر القيم وأكبر القيم في البيانات، فكلما اتسع مدى التغير كلما احتاج إلى عدد أكبر من الفئات .

٢- عدد التكرارات : وكلما زاد عدد التكرارات كلما احتجنا إلى فئات أكثر .

٣- اختلاف القيم عن بعضها : فإذا كانت المفردات متقاربة فيما بينها احتاجت إلى عدد أقل من الفئات، والعكس صحيح .

تبويب البيانات في جداول:

إن الهدف من تبويب البيانات هو إعطاء صورة سهلة وواضحة للظاهرة المدروسة، ومعنى الجدول هو أننا نسعى لوضع كل المفردات التي لدينا معاً في نظام مختصر يتناسب مع إمكانات الباحث الزمنية والمالية والفنية .

ولتوضيح طريقة تصنيف البيانات في جداول معينة نأخذ المعلومات الواردة في الجدول (٣ : ٢) والذي يبين قراءات كمية المطر الساقطة على مدينة معينة خلال ١٠٠ يوم . فلتصنيف هذه البيانات نتبع الخطوات الآتية :

١ - نحسب المدى المطلق للبيانات وهو ٨٥-٢٠=٦٥ مليمتر .

٢- نختار طولاً مناسباً للفئة وليكن ١٠ .

٣- بقسمة المدى المطلق على طول الفئة نحصل على عدد الفئات التي يشملها الجدول وبقسمة ٦٥ على ١٠ يكون عدد الفئات التي لدينا سبع فئات متساوية طول كل منها ١٠ .

- تشمل الفئة الأولى كمية الأمطار من ٢٠ إلى أقل من ٣٠ .

- وتشمل الفئة الثانية كمية الأمطار من ٣٠ إلى أقل من ٤٠ .

وهكذا حتى الفئة الأخيرة وتشمل كمية الأمطار من ٨٠ إلى أقل من ٩٠ .

ويمكن كتابة الفئات السابقة بعدة صور مختلفة، فقد تكون على النحو المبين أدناه :

٢٠-٣٠

٣٠-٤٠

٤٠-٥٠

. . . وهكذا .

وهذه الصورة معيبة إذا كانت هناك قراءة قيمتها ٣٠ ملم بالضبط، فهذه قد تدخل الفئة الأولى وقد تدخل الفئة الثانية، وتلافياً للبس لا نستخدم هذه الطريقة . وإذا استخدمت وجب تحديد أكثر : كأن يذكر في الجدول من ٢٠ إلى ٣٠ بحيث تحدد الفئة الأولى بأنها تستوعب حجم الأرقام التي تقل عن ٣٠ مثل : ٢٩,٦٠ أو ٢٩,٩٩ حتى إذا ما وصلت القيم إلى ٣٠ وضعت في الفئة الثانية .

وتلافياً لذلك يقترح الإحصائيون صيغة أخرى وهي :

٢٠-٢٩ .

٣٠-٣٩ .

٤٠-٤٩ .

. . . وهكذا .

وهذه الصورة قد تكون صحيحة وقد تكون معيبة ، فهي صحيحة إذا كنا لا نتصور أن الظاهرة المدروسة يمكن أن تأخذ قيماً كسرية ، أما إذا كانت الظاهرة تأخذ قيماً كسرية فهي غير مناسبة ، ففي مثالنا السابق يمكن أن تكون إحدى القراءات هي ٢٩,٥ ملم ، فهذه القراءة لا تدخل ضمن الفئة الأولى ؛ لأن نهايتها ٢٩ ولا تدخل ضمن الفئة الثانية لأن بدايتها هي ٣٠ .

إن صيغة عرض البيانات السابقة ذات الفئات ٢٠ - ٢٩ ، ٣٠-٣٩ ، . . إلخ . تصلح فقط في البيانات ذات المتغيرات الوثابة باعتبار أن قيمها لا تشتمل على كسور عددية ، أما المتغيرات السيارة فلا يصلح لها هذا النوع من تبويب البيانات ، والطريقة التي تصلح في كل الأحوال هي التي تعتمد على ذكر الحد الأدنى للفئة (وهو أصغر قيمة داخل الفئة) وترك الحد الأعلى (وهو أكبر قيمة للفئة) ، أو العكس ذكر الحد الأعلى وترك الحد الأدنى .

فإذا ذكر الحد الأدنى كان الجدول على النحو التالي :

-٢٠

-٣٠

-٤٠

وإذا ذكر الحد الأعلى كان الجدول على النحو التالي:

- ٣٠

- ٤٠

- ٥٠

. . . وهكذا.

إن الصيغة الأولى تقرأ من ٢٠ إلى أقل من ٣٠، ونلاحظ أن هذه الصيغة تجعل الفئة الأولى تسمح بوجود ٢٠ و ٢١ و ٢٢ . . . و ٢٩ وجميع كسور الرقم ٢٩ مهما علت هذه الكسور بشرط ألا تصل إلى ٣٠، فإذا وصلت إلى ٣٠ ضمت إلى الفئة الثانية.

أما الصيغة الثانية فتقرأ أكثر من ٢٠ إلى ٣٠: وأكثر من ٣٠ إلى ٤٠ وهكذا، وهذه الصيغة تجعل الفئة الأولى تسمح بوجود ٢٠، ٢١، ٢٢ . . . و٢٩ و ٣٠ فإذا زادت عن الثلاثين ضمت إلى الفئة الثانية، والخلاف بين الصيغة الأولى وهذه الصيغة هي أن الرقم ٣٠ يضم إلى الفئة الأولى في هذه الصيغة في حين أنه في الصيغة الأولى يضم إلى الفئة الثانية.

تفريغ البيانات في جداول:

نرسم جدولاً تفريغياً، مكوناً من ثلاثة أعمدة: الأول ليضم الفئات المختلفة حسب ترتيبها التصاعدي، والثاني ليحتوي على علامات التفريغ،

obeikandl.com

والثالث ليشتمل على التكرارات على النحو المبين في جدول (٣ : ٣)، ثم نرجع إلى الأعداد المدرجة في جدول (٣ : ٢) ونأخذها مفردة مفردة. ونستمر في هذه العميلة إلى أن يتم أخذ جميع مفردات المجموعة، مع ملاحظة تكوين حزم من أربعة خطوط رأسية يجمعها خط أفقي لتسهيل عملية عد الخطوط التي تمثل في الواقع التكرارات المناظرة للفئات المختلفة، حتى يتسنى ترجمتها إلى أعداد وادخالها في العمود الثالث من الجدول (٣ : ٢) الذي يسمى في العادة بالجدول التفريغي.

جدول (٣-٣)

جدول تفريغي يوضح توزيع كمية الأمطار الساقطة

على مدينة ما خلال ١٠٠ يوم

فئات الأمطار (ملم)	العلامات	عدد الأيام (التكرارات)
٢٠ -		٤
٣٠ -		١١
٤٠ -		٢٠
٥٠ -		٣٦
٦٠ -		١٧
٧٠ -		٨
٨٠ - ٩٠		٤
الجملة	-------	١٠٠

٤- نأخذ العمودين الأول والثالث من الجدول التفريغي السابق

فنحصل على الجدول التكراري كما يتضح من الجدول (٣-٤).

جدول رقم (٣-٤)

توزيع كمية الأمطار الساقطة على مدينة ما خلال ١٠٠ يوم

فئات كمية الأمطار بالمتر المكعب	عدد الأيام (التكرار)
٢٠ -	٤
٣٠ -	١١
٤٠ -	٢٠
٥٠ -	٣٦
٦٠ -	١٧
٧٠ -	٨
٨٠ - ٩٠	٤
الجملة	١٠٠

إن هذا الجدول التكراري الذي يبين كميات الأمطار الساقطة مرتبة ترتيباً تصاعدياً مع ما يقابلها من عدد الأيام التي سقطت فيها كميات متماثلة من الأمطار يمتاز بخصائص هامة هي:

١- إن الفئات التكرارية متساوية الأطوال، وهذا بالطبع سيسهل عملية الرسم البياني للجدول السابق، كما يسهل إجراء بعض العمليات الحسابية لتقدير بعض المقاييس الإحصائية، غير أن هناك حالات أخرى كثيرة لا تكون الفئات التكرارية فيها متساوية الأطوال لاعتبارات مختلفة، فإذا وجدنا أن بعض الفئات ذات تكرارات قليلة، فقد يفضل ضم فئتين أو

أكثر، كما أن الحاجة إلى تفصيل أكثر شمولاً في بعض أجزاء التوزيع دون أجزائه الأخرى قد تبرر استخدام فئات غير متساوية الأطوال، فمثلاً في الجدول التالي (٣-٥) عرض التوزيع العمري للسكان في فئات طولها سنة واحدة، ثم ازداد طول الفئة إلى (٥) عند الأعمار المتقدمة، ثم تركت الفئة الأخيرة مفتوحة النهاية فكتبت ٦٥ سنة فأكثر؛ وذلك لأهمية السنوات الأولى بالنسبة للدراسات السكانية والصحية مما يستوجب تفصيلها بعرضها في فئات قصيرة، ثم لقلة عدد السكان في الأعمال المتقدمة (٦٥ فأكثر) مما لا يستدعي عرضها تفصيلياً في فئات منتظمة مثل الأعمار السابقة لها، وبذلك يختار لها فئة غير محددة النهاية لتمثيلها.

٢- إن الفئات في الجدول (٣-٤) مقفلة أو بمعنى آخر التوزيع مقفل عند بدايته ونهايته، فالفئة الأولى تبدأ بالرقم ٢٠، والفئة الأخيرة تنتهي بالرقم ٩٠، أي إن النهاية الصغرى والعظمى للظاهرات المدروسة معروفة ومحددة، غير أننا نجد أنه قد يتعذر تحديد البداية أو النهاية أو كليهما معاً، وفي هذه الحالة يكون تقدير بعض المقاييس الإحصائية بدقة تامة أمراً مستحيلاً (مثل: الوسط الحسابي والتباين وغيرها). ولكننا مع هذا قد نضطر إلى استخدام توزيع مفتوح من أحد طرفيه إذا تعذر تحديد النهاية أو البداية. وهذه الظاهرة تُلاحظ كثيراً عند التعبير عن التوزيع العمري للسكان، فتذكر الفئة الأولى على أنها أقل من سنة مثلاً والفئة الأخيرة على أنها ٦٥ سنة فأكثر كما هو الحال في الجدول (٣-٥) المفتوح الطرفين.

جدول (٣-٥)

عدد السكان الأجانب في منطقة مكة المكرمة

موزعين حسب الجنس وفئات السن

فئات السن	الذكور	الإناث	الإجمالي %
أقل من سنة	٤,٩٣٩	٤,٦٢٦	٩,٥٦٥
١ - ٤	٢١,٥٠٠	٢٠,٦١٥	٤٢,١١٥
٥ - ٩	٢٢,٨٧٣	٢١,٨٤٠	٤٤,٧١٣
١٠ - ١٤	١٨,٨٧٢	١٤,٢١٦	٣٣,,٠٨٨
١٥ - ١٩	٢٣,٦٤١	١٠,٩٩٩	٣٤,٦٤٠
٢٠ - ٢٤	٢٩,٣٢١	١٠,٩٨٤	٤٠,٣٠٥
٢٥ - ٢٩	٢٨,٩٧٧	١١,٩٨٩	٤٠,٩٦٦
٣٠ - ٣٤	٢٢,٤٧٦	١٠,٢٤٢	٣٢,٧١٨
٣٥ - ٣٩	١٨,٢٨٨	٧,٥٧٦	٢٥,٨٦٤
٤٠ - ٤٤	١٣,٤٢١	٦,٢٧٢	١٩,٦٩٣
٤٥ - ٤٩	٨,١٢٧	٣,٥٥٧	١١,٦٨٤
٥٠ - ٥٤	٧,٠٠٢	٤,١٣٧	١١,١٣٩
٥٥ - ٥٩	٣,٠٢٠	١,٦٣٧	٤,٦٥٧
٦٠ - ٦٤	٣,٦٢٣	٢,٨٢٣	٦,٤٤٦
٦٥ فأكثر	٣,٨٨٧	٣,٦٢٥	٧,٥١٢
الإجمالي	٢٢٩,٩٦٧	١٣٥,١٣٨	٣٦٥,١٠٥

مصفوفة البيانات الجغرافية: هي الشكل الذي ترتب البيانات من خلاله، وهي في العادة مستطيل أو مربع مقسم إلى عدد من الأعمدة وعدد آخر من الصفوف، ويختلف عدد هذه الأعمدة أو الصفوف بحسب حجم البيانات المدخلة ونوعها.

obeikandl.com

تستخدم الأعمدة عادة لإدخال البيانات الخاصة بالظواهر المدروسة، فمثلاً: إذا كانت لدينا بيانات عن عدد السكان ونسبة المواليد والوفيات والدخل ونسبة سكان المدن لعدد من الدول العربية (انظر جدول ٣-٦) ونريد إدخالها في مصفوفة البيانات، نخصص عموداً لإعداد السكان، وعموداً آخر لنسبة المواليد، وثالثاً للوفيات، ورابعاً لنسبة سكان المدن، وخامساً للدخل، فهذه الأعمدة تمثل الظواهر المدروسة، وتعرف هذه الظواهر باسم: المتغيرات (Variables)؛ لأن قيمها تختلف بين منطقة وأخرى وبين زمن وآخر؛ ولذا عرفت باسم المتغيرات.

أما الصفوف أو الأسطر (Rows) فتمثل المشاهدات Observations وهي الحالات (Cases) التي تجرى عليها الدراسة، أو التي ندرس من خلالها الظواهر الجغرافية.

والمشاهدات أو الحالات في الجدول السابق هي: الدول المختارة للدراسة. حيث تمثل كل حالة دولة معينة، وعلى هذا تكون المصفوفة الجغرافية الموجودة في جدول رقم (٣-٦) هي مصفوفة مستطيلة مكونة من ثمانية صفوف وخمسة أعمدة، فإذا رمزنا لهذه المصفوفة بالرمز (س) فإنه يمكن أن تختصر لتقرأ (س ٨×٥) وهذا يعني أن عدد المشاهدات (الأسطر) هو ثماني حالات، وعدد الظواهر المدروسة (المتغيرات) هو خمس متغيرات.

يختلف شكل المصفوفة وحجمها تبعاً لعدد الحالات المدروسة، وكذلك الظاهرات المراد دراستها، فقد تكون المتغيرات طبيعية أو بشرية أو

اقتصادية، وقد تكون الحالات مناطق أو أقاليم إدارية أو سياسية، أو مدن أو محطات رصد أو أشخاص أو غير ذلك، وعلى العموم يمكن إدخال آلاف المتغيرات وكذلك آلاف الحالات أو المشاهدت إلى قائمة المصفوفة.

جدول (٣-٦)

مصفوفة البيانات الجغرافية

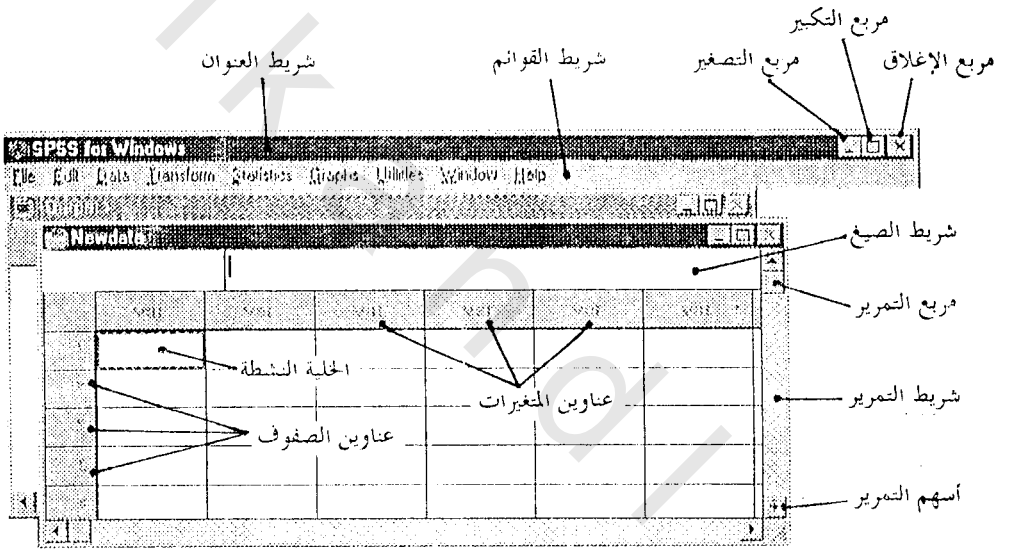
	المتغير الأول عدد السكان بالمليون	المتغير الثاني المواليد بالألف	المتغير الثالث الوفيات بالألف	المتغير الرابع سكان المدن ٪	المتغير الخامس الدخل بالدولار
الحالة الأولى السعودية	١٩	٣٦	٤	٧٩	٧٢٤٠
الحالة الثانية الكويت	٢	٢٦	٢	٩٨	١٩٠٠٠
الحالة الثالثة الأردن	٤	٣٢	٦	٧٨	١٣٩٠
الحالة الرابعة الجزائر	٢٩	٣٠	٦	٥٠	١٦٩٠
الحالة الخامسة مصر	٦٤	٣٠	٧	٤٤	٧١٠
الحالة السادسة ليبيا	٥	٤٥	٨	٦٥	غير متوفر
الحالة السابعة المغرب	٢٨	٢٩	٦	٤٧	١١٥٠
الحالة الثامنة السودان	٢٧	٤٢	١٢	٢٧	غير متوفر

وإزاء تعدد حالات الدراسة ومتغيراتها تصبح عملية التحليل شبه مستحيلة بالوسائل العادية؛ ولهذا لا بد من استخدام الحاسوب بهذا الخصوص.

استخدام الحاسوب لإدخال البيانات: لعل من أفضل البرامج المستخدمة لمعالجة البيانات الإحصائية هو برنامج الحزمة الإحصائية للعلوم الاجتماعية (SPSS) الذي يحوي على العديد من المميزات الخاصة بإدخال البيانات وتخزينها ومعالجتها إحصائياً، وعمل التحليلات المناسبة، وإظهار الأشكال البيانية المختلفة.

في نظام الـ (SPSS) تخزن البيانات في ملفات خاصة تعرف بملفات البيانات (DATA FILES)، ويلزم لإنشاء ملفات البيانات إدخال هذه البيانات على هيئة مشاهدات خاصة بمتغيرات معينة، ثم تحديد مواقعها ومسمياتها بطرق محددة من خلال البرنامج الإحصائي (SPSS).

تدخل البيانات إلى برنامج (SPSS) من خلال اللوحة الأولى في البرنامج (انظر شكل ٣-١) وتعرف هذه اللوحة باسم لوحة البيانات Spread Shect أو (ورقة العمل)وعندما تبدأ تشغيل الـ (SPSS) تظهر نافذة SPSS FOR WINDOWS. وإذا دققت النظر في الشكل ستلاحظ أن (SPSS) فتح صفحة أسماها: (New Data) وعندما تكتب البيانات في الشكل رقم (٣-٢) حل محل (New Data) اسم: Pop. sav وهو اسم البرنامج المقترح إدخاله.



شكل (١:٣) لوحة البيانات أو ورقة العمل في (Spss)

obeikan.com

تشمل الشاشة الافتتاحية للبرنامج على مجموعة من العناصر هي :

١- الأعمدة : ويصل عددها ٢٥٦ عمود .

٢- الصفوف أو السطور : ويصل عددها إلى أكثر من ١٣ ألف صف .

٣- الخلية : وهي نقطة تقاطع أي عمود مع أي سطر على ورقة العمل .

٤- الخلية النشطة : هي الخلية التي يحيط بها برواز يسمى مؤشر الخلية ، ويظهر عنوانها دائماً في شريط الصيغ ، وهذه الخلية هي التي تستقبل المدخلات من لوحة المفاتيح .

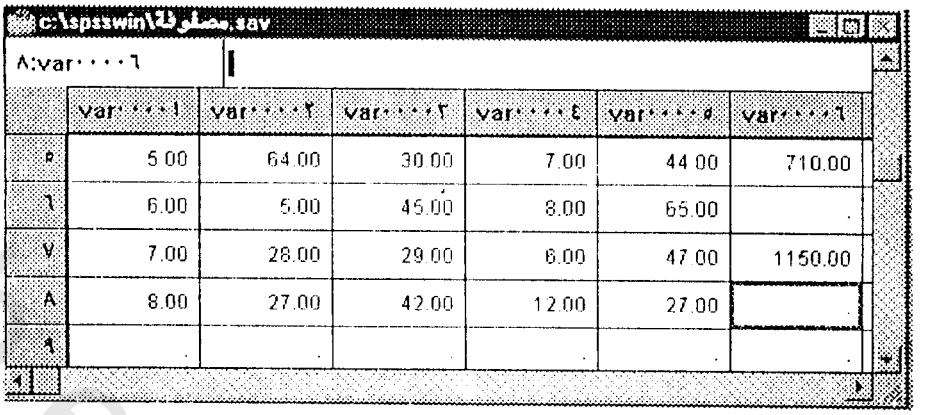
٥- شريط العنوان (Title Bar) : ويشتمل على اسم البرنامج واسم الملف .

٦- شريط القوائم (Menu Bar) : ويشتمل على أسماء القوائم . وكل قائمة تشتمل على أوامر واختيارات متعددة تتحكم في ورقة العمل وطريقة التعامل معها .

٧- أشرطة التمرير Scroll Bars : تستخدم للتصفح داخل ورقة العمل ، أي الانتقال من مكان لآخر ، وتوجد على الحافة اليمنى والسفلى ، ويوجد على طرفي كل شريط سهم يستخدم للتحريك أيضاً حسب اتجاه السهم ، كما يوجد على كل شريط مربع صغير يستخدم للغرض نفسه .

٨- شريط الصيغ Formula Bar : تظهر فيه محتويات الخلية النشطة ، كما تظهر فيه البيانات التي تدخل إلى الخلية أثناء إدخالها .

٩- أزرار التحكم الثلاث وهي [X] [▫] [-] والأول لإغلاق النافذة ، والثاني لتكبيرها ، والثالث لتصغير النافذة بصورة عنوان صغير في أسفل الشاشة .

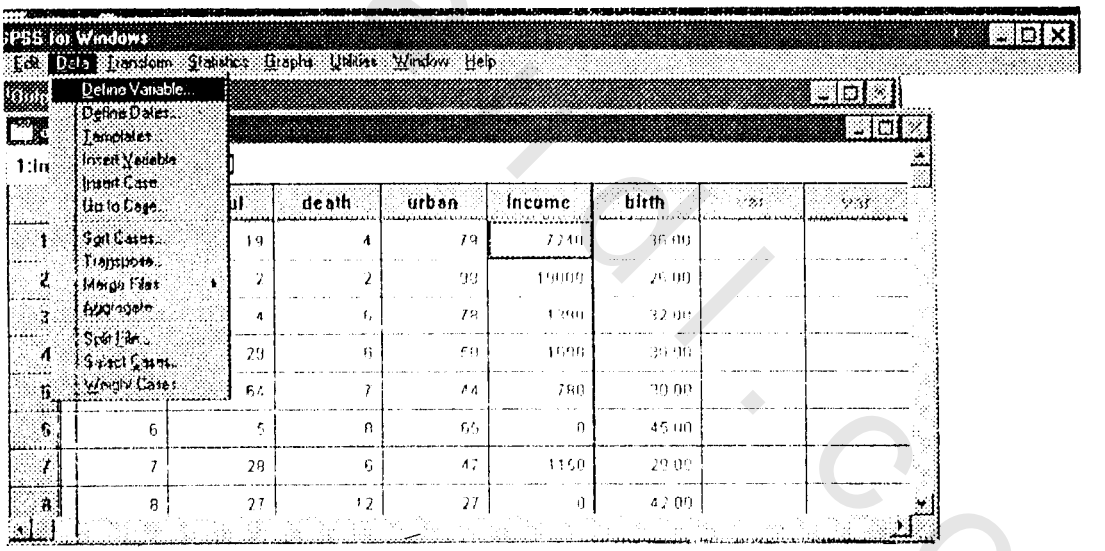


c:\spsswin\مصطفى.sav

A:var٠٠٠٦

	var٠٠٠١	var٠٠٠٢	var٠٠٠٣	var٠٠٠٤	var٠٠٠٥	var٠٠٠٦
٥	5.00	64.00	30.00	7.00	44.00	710.00
٦	6.00	5.00	45.00	8.00	65.00	.
٧	7.00	28.00	29.00	6.00	47.00	1150.00
٨	8.00	27.00	42.00	12.00	27.00	.
٩	.	.	.	.	.	.

شكل (٢:٣) مصفوفة بيانات سكانية لبعض الدولة العربية



شكل (٣:٣) الأوامر الخاصة بالتعامل مع المتغيرات والمشاهدات

إدخال البيانات: تدخل مصفوفة البيانات إلى برنامج (SPSS) من خلال ورقة العمل الأولى المسماة: New Data والمكونة من عدد الخلايا الناجمة عن تقاطع الأعمدة مع الصفوف. وكل عمود معنون ب: (Var) وهي اختصار لكلمة Varible أو متغير، أما المشاهدات أو الحالات فتدخل إلى الخلايا على شكل صفوف مرقمة (١، ٢، ٣، ٤، ٥، . . . إلخ).

تسمية المتغيرات: لا بد من تحديد أسماء المتغيرات من أجل التعرف عليها عند استعمالها، والخطوات المتبعة في تسمية المتغيرات خطوات سهلة وبسيطة. وتبدأ الخطوط باختيار الأمر (Data) أي بيانات من شريط القوائم الموجود في أعلى اللوحة، فتظهر قائمة أخرى تتضمن عدداً من الأوامر الخاصة بالتعامل مع المتغيرات والمشاهدات (شكل رقم ٣-٢) نأخذ الاختيار الأول من تلك القائمة وهو تسمية المتغيرات Define Variable، فتظهر لوحة جديدة خاصة بتسمية المتغيرات ٣-٤ نكتب الاسم الذي نختاره لذلك، ثم نختار الأمر (OK) فيظهر الاسم في العمود المخصص له في لوحة البيانات.

تحديد طبيعة المتغيرات : هناك مجموعة من الأوامر الخاصة بتحديد طبيعة المتغيرات من حيث :

١ – هل المتغير هو رقمي أو نصي .

٢ – ما هو عدد الأرقام الداخلة في المتغير .

٣– هل المتغير يحتوي على كسور عشرية ، وما هو عدد الخانات اللازمة للكسر .

٤ – هل هناك بيانات غير متوفرة (Missing Data) .

تخزين الملفات : تخزن مصفوفة البيانات بعد الانتهاء من إدخالها في ملف خاص (File) يتكون من مقطعين ، يمثل المقطع الأول أي اسم نختاره على أن لا يزيد طوله عن ثمانية مقاطع ، بينما يتكون المقطع الثاني من المقطع (Sav) وهذا المقطع خاص بأنواع الملفات في نظام الـ (Spss) .

تتبع الخطوات التالية لتخزين المعلومات :

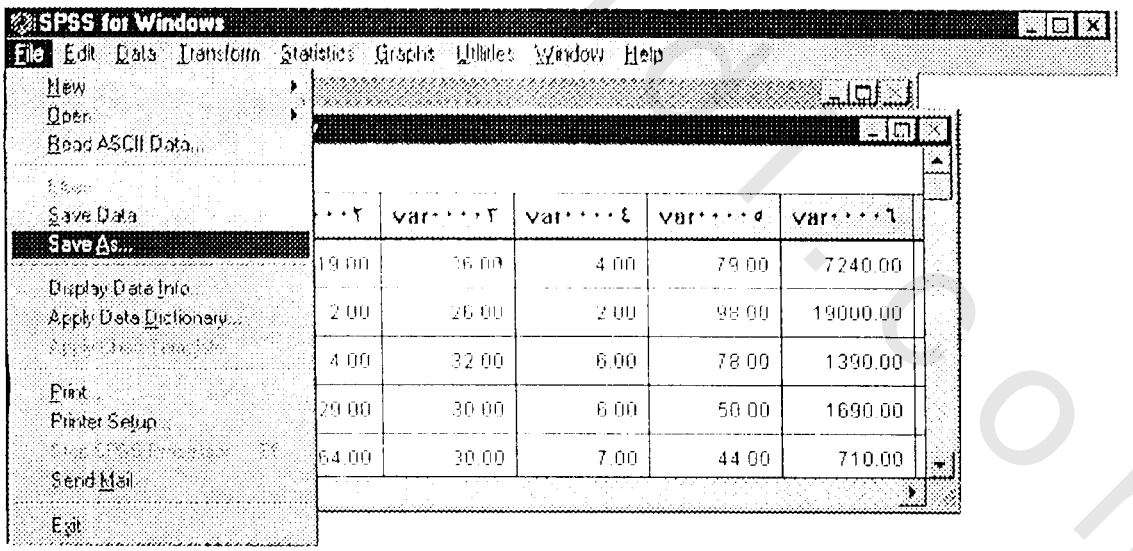
١ – من قائمة الأوامر نختار أمر (File) ، ونختار من القائمة المنسدلة أمر تخزين باسم (Save As) . (انظر شكل ٣–٥) .

٢ – نطبع اسم الملف الذي نختاره وليكن (Pop. Sav) ، وهذا اختصار لكلمة Population ، واسم الملف وسيلة هامة لتذكر الموضوع وتمييزه عن غيره من الملفات االأخرى ، وهذا الاسم يستخدم عند طلب الملف لإجراء تعديلات عليه ، أو لإجراء أي عمليات إحصائية .

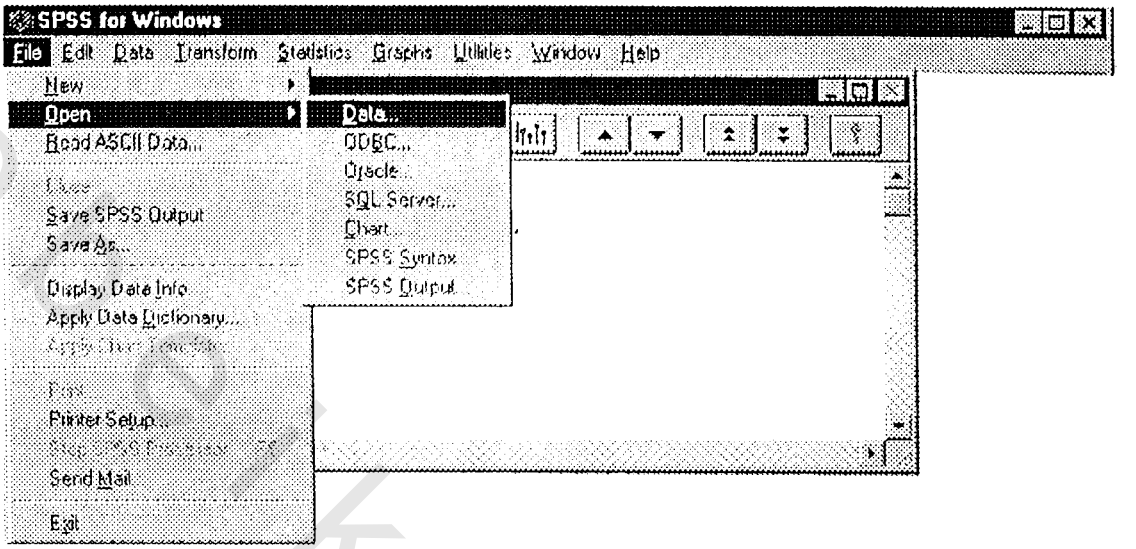
٣– إذا أجريت أي تعديلات أخرى على الملف ، تخزن بعد كل مرة يجري فيها التعديل ، وذلك باستخدام أمر خزن (Save) وليس باستخدام الأمر السابق (Save As) إلا إذا كنا نريد تخزين الملف باسم جديد .

طلب الملفات: نطلب أي ملف بيانات (Data File) مخزن في برنامج (SPSS) في حالتين : إذا كنا نرغب في استخدامه ، أو في إدخال تعديلات عليه ، ثم

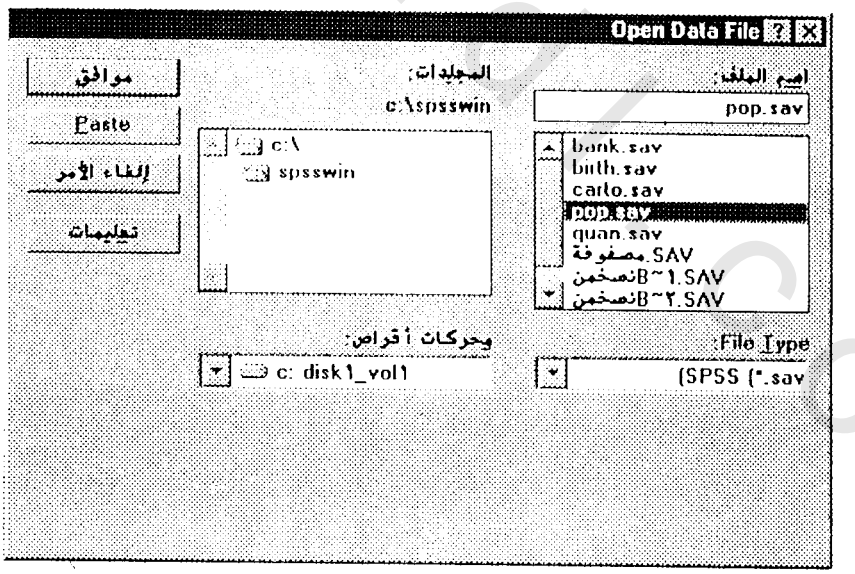
إعادة تخزينه من جديد، ونطلب الملف الذي نريده باستخدام أمر File من شريط القوائم، وبالنقر على File تظهر قائمة أخرى تحوي عدداً من الأوامر، نختار منها فتح (Open) فتظهر قائمة أخرى بأنواع الملفات الخاصة ببرنامج (Spss) المحفوظة (شكل ٣-٦)، يهمنا من تلك القائمة النوع الأول من الملفات (Data)، والنوع الأخير الخاص بملفات النتائج (Spss out put). فإذا كنا نطلب ملف البيانات ننقر على الأمر الأول (Data) فتظهر شاشة جديدة تتضمن قائمة بملفات البيانات المخزنة في (SPSS)، نستطيع أن نفتح الملف المطلوب بالنقر على اسم الملف مرتين متتاليتين يظهر الملف ويكون جاهزاً للاستعمال. (شكل ٣-٧).



شكل (٣:٥) طريقة حفظ الملفات في (Spss)



شكل (٦:٣) طلب الملفات في برنامج (Spss)



شكل (٧:٣) قائمة الملفات في (Spss)

جدولة البيانات بالحاسوب: قدمت أجهزة الحاسوب تسهيلات كثيرة في مجال إدخال البيانات وجدولتها، ويمكن استخدام برنامج (Spss) لإعداد الجداول التكرارية بأنواعها المختلفة، ويجب أن يسبق تكوين أي جدول تحديد بعض العناصر الأساسية ومنها:

١- الحد الأدنى والحد الأعلى للقيم.

٢- عدد الفئات المطلوبة.

٣- حدود الفئات.

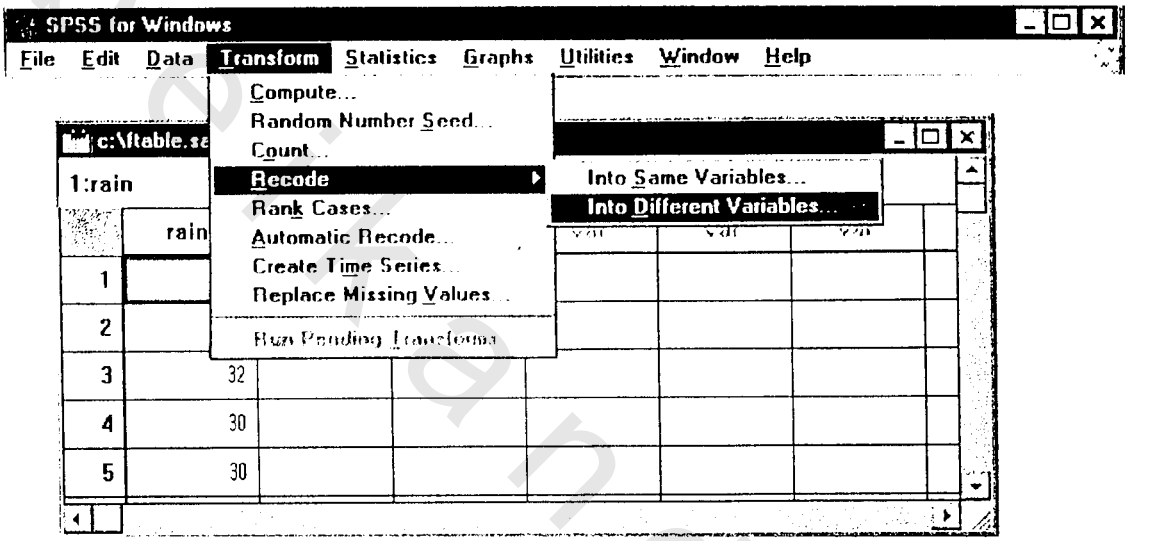
وعملية التحديد هذه شبيهة بما قمنا به عند تفريغ البيانات في الجدول التكراري المعمول يدوياً.

سنأخذ المثال االسابق الخاص بكميات الأمطار الساقطة على مدينة ما لمدة مائة يوم، والتي عملنا لها جدولاً تكرارياً بالطريقة اليدوية (انظر الجداول رقم ٣: ٢، ٣: ٣، ٣، : ٤) التي تبين المراحل التي استخدمناها لإنشاء الجدول التكراري، والخطوات التالية تبين طريقة عمل الجدول التكراري بالحاسوب.

١- ندخل بيانات الجدول السابق رقم (٣: ٢) في العمود الأول من ورقة العمل في (spss) ونخزنها في ملف خاص يتكون من متغير واحد هو المطر (rain)، ويمكن أن نسمي هذا الملف باسم الجدول التكراري .(Ftable Save).

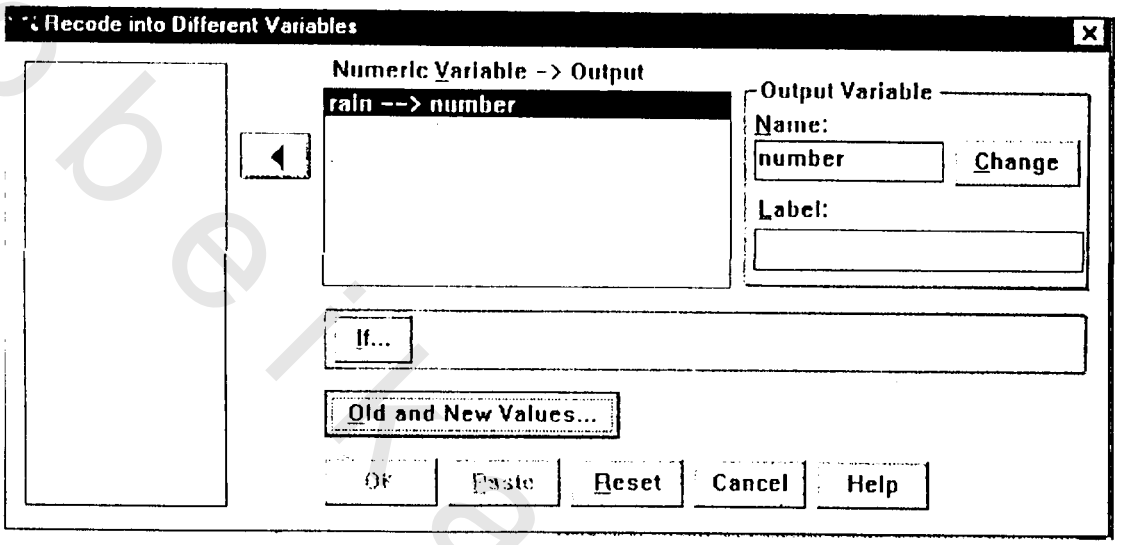
٢- من شريط القوائم نختار أمر التحويل (Transform) لتحويل المشاهدات الكمية للمتغير (Rain) إلى مشاهدات رمزية، تمثل أرقام الفئات

التي تنتمي إليها تلك المشاهدات، تظهر فور اختيارنا للموضوع (Trans-form) الشاشة الثانوية المنسدلة التي تحوي العديد من الأوامر. (شكل ٣-٨).



شكل (٣:٨) أوامر تبويب البيانات

٣- نختار من تلك الشاشة الأمر (Recode) وهو عملية الترميز، أي إعطاء قيم رمزية بديلة لقيم المشاهدات، فتظهر شاشة ثانوية أخرى تتضمن خيارين بتسمية المتغير الرمزي الجديد باسم المتغير الأصلي (Recode into same variable)، أو إنشاء متغير جديد (Recode into different vari-able)، نختار الخيار الثاني بغية المحافظة على المشاهدات الأصلية وعدم إتلافها، ونسمي المتغير الجديد: (Number). (شكل ٣-٩).

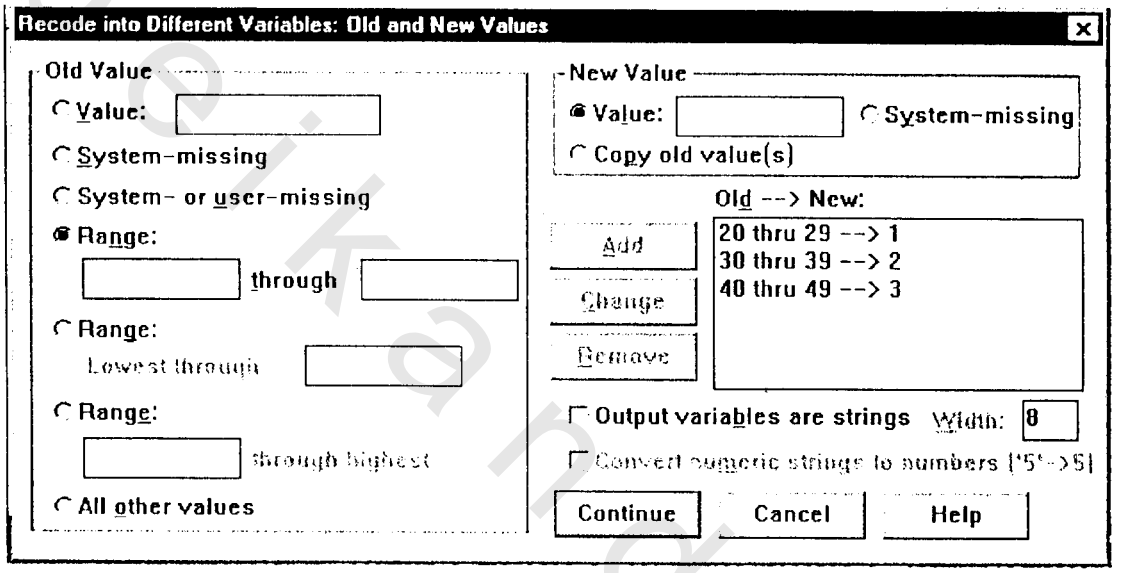


شكل (٩:٣) لوحة تبويب البيانات في متغير آخر

٤- تظهر بعد الانتهاء من عملية الاختيار شاشة جديدة بعنوان Recode into Different Variable شكل رقم (٣-٩)، ننقل المتغير (Rain) إلى مربع المتغيرات المدخلة والخارجة (Variable Input Variable - Output) ونكتب (Number) في المستطيل المعنون (Name) الواقع أسفل عبارة (Output Vara-ble) ثم نختار الأمر (Change) للقيام بعملية التغير.

٥- ننقر بفارة الحاسوب على المفتاح (Old and New Values) فتظهر شاشة جديدة بعنوان: Recode into Different Variables: Old New Values تنقسم تلك الشاشة كما هو مبين في شكل رقم (٣-١٠) إلى قسمين: قسم

مخصص لكتابة حدود الفئات وعنوانه : (Old Value) والقسم الآخر وعنوانه : (New Value) مخصص لتحديد الرموز التي ستعطي للمشاهدات الواقعة ضمن كل واحدة من تلك الفئات .



شكل (١٠:٣) ترميز البيانات من أجل تبويبها

نبدأ بكتابة حدود الفئة الأولى، فنختار الكلمة الأولى (Range) ونكتب الحد الأدنى لتلك الفئة وهو (٢٠) في المكان المخصص لذلك من تلك اللوحة، ثم نكتب حدها الأعلى وهو (٢٩) في المكان الثاني الذي تفصل بينه وبين المكان الأول كلمة (Through)، ننقل فارة الحاسوب فور الانتهاء من ذلك إلى النصف الثاني من اللوحة المخصصة لتحديد الرموز التي ستعطى للمشاهدات الواقعة ضمن كل واحدة من تلك الفئات، ونكتب

رمز تلك الفئة في المكان المخصص لذلك مقابل كلمة Value، نختار بعد الانتهاء من ذلك (ADD) فتظهر الفئة الأولى والرمز الذي يقابلها في المكان المخصص لذلك والمبين في اللوحة نفسها، وعنوانه: "Old---New" الفئة الأولى، ورمزها الجديد (١). نعيد هذه العملية مع بقية الفئات. (شكل ٣-١٠).

٦- نختار فور الانتهاء من ذلك الأمر (Continue) الموجود في أسفل اللوحة، فتظهر اللوحة السابقة شكل رقم (٣: ٩) نختار من تلك اللوحة الأمر (OK) فتظهر لوحة بيانات Spss ثانية تتضمن المتغير الجديد (Number) الذي يمثل الفئات التي تنتمي إليها المشاهدات.

٧- نعود إلى شريط القوائم من جديد، ونختار الأمر "Statistics" أي: «إحصاء» فتظهر شاشة ثانوية تتضمن قائمة بمختلف الأساليب الإحصائية، نضغط الفأرة عند الأسلوب Summarize أي «اختصار» فتظهر شاشة ثانوية أخرى تتضمن أوامر متعددة لاختصار البيانات في جداول تكرارية متنوعة، نختار من تلك اللوحة الأمر (Frequencies) أي «تكرارات» فتظهر شاشة أخرى تتضمن تحت عنوان Variables المتغيرين (Number), (Rain) ننقل المتغير (Number) من موقعه إلى الموقع الجديد ونضغط الفأرة عند الأمر OK فيقوم، الحاسوب بعمل الجدول التكراري المبين في جدول رقم (٣: ٧) حيث تظهر أرقام الفئات في العمود الأول، والتكرارات المقابلة لتلك الفئات في العمود الثاني، والنسبة المئوية للتكرارات في العمود الثالث، أما العمود الرابع فيشمل التكرارات المعدلة فيما إذا كان هناك قيم مفقودة، فإذا

لم يكن هناك قيم مفقودة من البيانات ظهر هذا العمود بصورة مماثلة للعمود الثاني الذي قبله . والعمود الأخير يظهر النسبة التراكمية أو المتجمع الصاعد للتكرارات .

٨- نخزن ملف النتائج باسم ftable وذلك باستخدام الأمر Save SPSS output وليس باستخدام الأمر Save as الذي استخدمناه لتخزين البيانات ونستطيع أن نستعيد ذلك الملف من جديد بطلبه من بين ملفات النتائج SPSS output وليس من ملفات البيانات Data files .

20 Jun 98 SPSS for MS WINDOWS Release 6.0

NUMBER

Value Label	Value	Frequency	Percent	Valid Percent	Cum Percent
	1.00	4	4.0	4.0	4.0
	2.00	11	11.0	11.0	15.0
	3.00	20	20.0	20.0	35.0
	4.00	36	36.0	36.0	71.0
	5.00	17	17.0	17.0	88.0
	6.00	8	8.0	8.0	96.0
	7.00	4	4.0	4.0	100.0
	Total	100	100.0	100.0	

Valid cases 100 Missing cases 0

جدول رقم ٧:٣

عمل الجدول التكراري بالحاسوب

obeikandl.com

obeikandl.com

أسئلة وتطبيقات

س١ : ما الفرق بين المتغيرات المنفصلة والمتغيرات المتصلة؟

اعط أربعة أمثلة لكل منهما؟ وما هي الأسس التي يقوم عليها هذا التقسيم؟

س٢ : ما هي عيوب التوزيع التكراري المفتوح؟

س٣ : مع أي المتغيرات تصلح الصيغة التالية لتوزيع الفئات ولماذا؟

١٠-١٤

١٥-١٩

٢٠-٢٤

س٤ : البيانات التالية تمثل نتائج(٥٠) طالباً في مادة الكمية والإحصائية؟

٩٥، ٥٩، ٨٨، ٧٦، ٩٣، ٦٨، ٧٢، ٨٥، ٧٧، ٦٨

٧٩، ٦٢، ٦٧، ٩٧، ٨٨، ٧٨، ٥٣، ٧٤، ٧٣، ٦٠

٦٦، ٧٨، ٨٢، ٧٥، ٩٤، ٧٧، ٦٩، ٦٨، ٦٠، ٨٣

٧٥، ٩٤، ٦٦، ٧٧، ٩٠، ٨٥، ٧٨، ٦٣، ٥٠، ٦٠

٧٧، ٧٤، ٧٥، ٧١، ٥٩، ٦٤، ٩٠، ٨٦، ٨٩، ٧٤

أجب عما يلي :

١- أوجد أدنى علامة وأعلى علامة؟

obeikandl.com

٢- استخرج المدى؟

٣- بوب البيانات السابقة في جدول تكراري طول فئاته ١٠؟

٤- استخرج علامات الطلاب الخمسة الأوائل في هذه المادة؟

٥- استخدم الحاسوب في تبويب هذ البيانات في جدول تكراري طول فئاته ٢٠ .

س٦ : ماذا نقصد بمصطلح «البيانات الجغرافية»؟ وهل يمكن تمييز هذا النوع من البيانات عن غيرها من البيانات الأخرى؟ ما هي الخصائص العامة التي تشترك فيها البيانات الجغرافية؟

س٧ : قارن بين كل مما يلي :

١- البيانات التصنيفية والبيانات الكمية .

٢- البيانات الزمانية والبيانات المكانية .

س٨ : عرف كلاً مما يلي : شريط القوائم، مصفوفة البيانات، تسمية المتغيرات في Spss .