الدرس 5: ترميز البيانات

تمهيد:

قبل الشروع في التعامل مع محرر البيانات في برنامج Spss ينبغي القيام بترميز البيانات وإعدادها للإدخال في هذا البرنامج، وهو خطوة جد هامة من خطوات تحليل البيانات، ويتم من خلالها تهيئة البيانات سواءا كانت استمارة أو مقابلة أوملاحظة...أو غيرها، وذلك لكي يتمكن البرنامج من فهمها والتعامل معها.

ترميز البيانات Coding data:

ويقصد به عملية تحويل بيانات كل سؤال أو فقرة من أسئلة أو فقرات أداة جمع البيانات إلى أرقام أو أحرف ليسهل إدخالها إلى الحاسب الآلي، وغالبا ما نستخدم الأرقام لتمثيل البيانات، وذلك بأن يعطى كل متغير ترميزا معينا يعني مؤشرا معينا للبرنامج. ويجب التفريق هنا بين أنواع المتغيرات المختلفة. ويسري الترميز على كل الأداة بحيث تصحح جميع الاستمارات المراد إدخالها مثلا وترقم حسب أفراد العينة. وحسب برنامج Spss فإن الأشخاص الذين قاموا بالإجابة على أسئلة الاستمارة يطلق عليهم اسم الحالات Cases (الأفراد) وكل سؤال في الاستمارة هو بمثابة متغيرة Variable، وتشكل الإجابات قيم المتغيرة Valeur de variable.

ويفضل أولا استخدام أوراق الترميز لإدخال البيانات من الاستمارات، ثم القيام بعد ذلك بإدخال البيانات إلى الحاسب لإزالة الإرباك خاصة إذا كان عدد الأسئلة كبيرا.

دليل الترميز:

هو البيان العام الذي يتضمن اسئلة استمارة المعلومات وفقراتها ومتغيراتها في صورة اختصارات حرفية أو رقمية ترمز إلى وصف المتغيرات أو الأسئلة في صياغتها التامة المصاغة بها. وظيفته مساعدة الباحث على تذكر السؤال أو المتغيرات بعد اختصارها حتى لا يقع في أخطاء التحليل، ويرجع إليه الباحث لتفسير الاختصارات التي ترمز إلى البيانات بعد إدخالها إلى الحاسب الآلي.

طريقة ترميز الاستمارة:

تختلف طريقة الترميز حسب نوع السؤال، وفيما يلي بعض أنواع الأسئلة الأكثر استخداما في البحوث الاجتماعية، وكيفية التعامل معها لترميزها على برنامج Spss[1]:

1- سؤال مغلق يسمح باختيار إجابة واحدة: مثل:

-الجنس: ذكر ☐ أنثى ☐

عندما يكون لدينا سؤال يسمح فيه باختيار إجابة واحدة فقط فإنه يكفي صياغته في متغيرة واحدة، وفي هذا المثال لدينا متغير الجنس مكون من ذكور وإناث، يمكن صياغته على برنامج Spss بإضافة متغيرة جديدة باسم الجنس ونقوم بترميز الذكور بالرقم 1 والإناث بالرقم 2.

2- سؤال عبارة عن متغيرة كمية: مثل:

-السن:.....

تتطلب الإجابة على هذا السؤال تحديد عمر الشخص وفي هذه الحالة تعتبر المتغيرة كمية متصلة ولا تحتاج إلى أي ترميز. وفي بعض الأحيان يمكن تحويلها إلى فئات وهنا ترمز كل فئة بمركزها؛ مثلا الفئة [0- 10[ترمز بمركز هذه الفئة وهو 5.

3- سؤال يتضمن أكثر من متغير كمي: مثل:

عدد الإخوة: ..... الذكور: ..... الإناث: .....

في هذا السؤال يجب معرفة عدد الإخوة الذكور وعدد الإخوة الإناث، ويشكلان معا متغيرتين كميتين، وفي نفس السؤال مطلوب معرفة المجموع الذي يشكل متغيرا ثالثا ضمن هذا السؤال. ولربح الوقت يمكن حسابه من خلال جمع عدد الإخوة الذكور والإناث وذلك للاختصار ولربح الوقت في إدخال البيانات.

4- سؤال مفتوح متغيرة اسمية: مثل:

-المهنة: .....

يصاغ السؤال في متغيرة واحدة اسمية، ولترميز هذا السؤال يمكن تحديد بعض المهن وترميزها بأرقام، ولاختيار أي المهن سيتم الاعتماد عليها فإنه يجب الأخذ بعين الاعتبار عدة عوامل منها:

المهن الأكثر تكرارا في الاستمارة

الهدف من الاستمارة

الدقة المطلوبة

[1]: جمال شعوان: **مدخل لدراسة وتحليل البيانات الإحصائية تطبيقات على برنامج Spss**، (pdf)، 2014، ص. ص12 .16 .

5- سؤال منطقي يسمح باختيار إجابة واحدة فقط: مثل:

هل ترغب بمتابعة الدراسة بالجامعة: نعم ☐ لا ☐

يعتبر هذا النوع من الأسلة شبيها بالنوع الأول إلا أن الأمر يختلف في طبيعة المتغيرة، حيث أن الأسلة المنطقية التي تتطلب الإجابة بنعم أو لا غالبا ما يتم دراستها من الناحية الكمية وبالتالي فهي تختلف عن متغيرة الجنس من حيث طبيعة القياس، ونستخدم قيمة 1 و0 في ترميزها حيث 1= نعم و0= لا، وتأخذ وحدة القياس صيغة كمية Scale.

6- سؤال على شكل جدول: مثال:

أماكن الاتصال بالأنترنت

مكان الاتصال	صعب	نادر	دائم
المنزل			
نادي الأنترنت			
مكان العمل			

المطلوب في هذه السؤال تحديد درجة الاتصال بالأنترنت من الأماكن الثلاثة المذكورة، وبالتالي سيتم التعامل مع هذا السؤال على أساس أنه ثلاث أسئلة، وستحدد ثلاث متغيرات هي: (الاتصال من المنزل- الاتصال من نادي الأنترنت- الاتصال من مكان العمل).

وكل متغيرة ستأخذ ثلاث قيم رتبية يجب ترميزها بأرقام يراعى فيها احترام التراتبية، مثلا (1=صعب، 2= نادر، 3= دائم).

ويجب تحديد المقياس الرتبي للمتغيرات.

7- سؤال رتبي: مثل:

-هل توافق أن يكون تسجيل الطالب في الجامعة عن طريق الأنترنت فقط:

موافق بشدة ☐ موافق ☐ محايد ☐ معارض ☐ معارض بشدة ☐

هذا السؤال يشبه السؤال السابق؛ فهو عبارة عن سؤال واحد يقبل جوابا واحدا، ويمكن ترميزه على شكل متغيرة واحدة ذات أجوبة اختيارية رتبية، مثلا (1= معارض بشدة، 2= معارض، 3= محايد، 4= موافق، 5= موافق بشدة)

8- سؤال يسمح بأكثر من إجابة: مثل:

-ماهي أهم الهوايات التي تمارسها:

القراءة ☐ الرياضة ☐ السفر ☐ أخرى ☐

في هذا السؤال نلاحظ أن الشخص يمكن أن يجيب على أكثر من اختيار واحد، لذلك فإن متغيرا واحدا لا يكفي لتمثيل السؤال. في هذه الحالة يفضل إنشاء أربعة متغيرات، كل متغير له احتمال إجابتين (نعم / لا)، حيث (1= نعم ، 0 =لا)

9- سؤال عبارة عن فئات يقبل إجابة واحدة: مثل:

-الدخل الشهري:

أقل من 50.000 ☐ 100.000-50.000 ☐ 200.000-100.000 ☐

أكثر من 200.000 ☐

المتغيرة المدروسة في هذه الحالة هي متغيرة الدخل الشهري وهي متغيرة كمية صيغت على شكل فئات، ولترميز الفئات نستخدم مركز الفئة.

مركز الفئة هو المتوسط الحسابي لطرفي الفئة، بمعنى مجموع الطرفين مقسوما على 2، فمثلا مركز الفئة 200.000-100.000 هو 15.000.

بالنسبة للفئة التي تتضمن عبارة "أقل من" فإن مركز الفئة يساوي الحد الأعلى للفئة ناقص نصف طول الفئة التي بعدها، أما بالنسبة للفئة "أكثر من" فإن مركزها يساوي الحد الأدنى للفئة زائد نصف طول الفئة التي قبلها.

10- أسئلة تحديد الأهمية: ويطلب فيها من المبحوث وضع ترتيب للإجابات حسب أهميتها مثل:

-حدد أي رابط من الروابط التالية تفضل الانتماء إليها قبل غيرها (رتبها حسب الأهمية):

الروابط	الأول	الثاني	الثالث	الرابع	الخامس
الأسرة					
القبيلة					
القرية					
المدينة					
الوطن					
الوطن العربي					

يتم تحويل كل فقرة من فقرات السؤال إلى متغير مستقل ثم يعرف كل متغير على حدة، وتحدد قيمته ويتم وصفها وترميزها وتصبح على النحو التالي:

الأسرة: س1 س2 س3 س4 س5 س6

القبيلة: ق1 ق2 ق3 ق4 ق5 ق6

وهكذا