Codebook

Exploring Data 2nd edition: The Teaching Datasets

The five datasets documented here contain much more information than was used in the text or the exercises of the second edition of 'Exploring Data'. It is hoped that they will provide interesting material for you to explore using the SPSS package. The data has been taken from several different sources and all the datasets are available in machine-readable form for teaching purposes from the ESRC Data Archive at the University of Essex. Brief documentation about the variables in each dataset is provided here but fuller information about the studies from which the data has been extracted are given in the chapter in which the dataset is first used and on the website for the book http://polity.co.uk/exploringdata. Readers are referred to the original sources for the complete documentation of the data.

Many of the datasets that are available to researchers from the ESRC Data Archive contain literally hundreds, and sometimes thousands, of variables. This can make them unwieldy and daunting for students and junior researchers to use. The datasets documented here have therefore been specifically designed for students in that they include only a small subset of the original variables. However there has been no sub-setting of the total number of cases in the datasets so the results obtained from analysing them will be the same as results obtained when analysing the full datasets available from the archive.

List of teaching datasets

GHS05_ind_teach.sav	**17 variables**	**page 2**	**Chapter 1, 2, 7**
NCDS_ExpData_teach.sav	**32 variables**	**page 7**	**Chapter 3, 8, 9, 13**
YCS11_teach.sav	**10 variables**	**page 19**	**Chapter 6**
WorldBank_2000_Teach.sav	**13 variables**	**page 22**	**Chapter 10**
BSA2004_Teach.sav	**49 variables**	**page 24**	**Chapter 12**

Summary of datasets used in each of the chapters of the book	
Chapter 1	GHS05_ind_teach.sav
Chapter 2	GHS05_ind_teach.sav
Chapter 3	NCDS_ExpData_teach.sav
Chapter 4	No SPSS exercises
Chapter 5	No SPSS exercises
Chapter 6	YCS11_teach.sav
Chapter 7	GHS05_ind_teach.sav
Chapter 8	NCDS_ExpData_teach.sav
Chapter 9	NCDS_ExpData_teach.sav
Chapter 10	WorldBank_2000_Teach.sav
Chapter 11	No SPSS exercises
Chapter 12	BSA2004_Teach.sav
Chapter 13	NCDS_ExpData_teach.sav

GHS05_ind_teach.sav

This dataset includes seventeen variables drawn from the General Household Survey 2005 (see Exploring Data 2e appendix to Chapter 1 on the website that accompanies the book at http://polity.co.uk/exploringdata) and there are a total of 30069 cases. The sample includes individuals of all ages from just a few months old to 100 years. However the majority of the variables listed here contain data that has only been collected from adult respondents. For example, children have not been asked about their working hours or alcohol consumption. The main GHS consists of a household questionnaire, completed by the Household Reference Person, and an individual questionnaire, completed by all adults aged 16 and over resident in the household. The individual questionnaire includes data from the household dataset, and additional sections on employment, pensions, education, health, smoking, drinking, family information, financial situation, and income. The variables included in this dataset have been taken from the individual questionnaire.

The original data set from which this teaching dataset was extracted is deposited at the UK Data Archive as the following dataset:

SN 5640 -General Household Survey, 2005

Variables

```
Name (Position) Label

persno (1) Person number
    Measurement Level: Nominal

pid (2) Personal identifier
    Measurement Level: Nominal

age (3) Age
    Measurement Level: Scale

sex (4) Sex
    Measurement Level: Ordinal
          Value     Label

              1     Male
              2     Female

drating (5) estimated weekly units: all drinks
    Measurement Level: Ordinal
    Missing Values: -6, -8, -9

          Value     Label

            -9 M   DNA
            -8 M   NA
            -6 M   Child/proxy/NI
             0     Abst/none last year

drinkamt (6) Amount of alcohol respondent drinks
```

```
      Measurement Level: Ordinal
      Missing Values: -8, -9

              Value     Label

                -9 M   DNA
                -8 M   NA
                 1     hardly drink at all
                 2     drink a little
                 3     drink a moderate amount
                 4     drink quite a lot
                 5     or drink heavily?

nssec5 (7) NS-SEC - Five-class version
      Measurement Level: Ordinal
              Value     Label

               -9.0     DNA
               -8.0     NA
               -6.0     CHILD/NO INT/PROXY
                1.0     Managerial and professional occs
                2.0     Intermediate occupations
                3.0     Small employers and own account workers
                4.0     Lower supervisory and technical occupations
                5.0     Semi-routine occupations
               96.0     Never worked and long term unemployed
               97.0     Not classified


nssec8 (8) NS-SEC - Eight-class version
      Measurement Level: Ordinal

              Value     Label

               -9.0     DNA
               -8.0     NA
               -6.0     CHILD/NO INT/PROXY
                1.1     Large employers and higher managerial occs
                1.2     Higher professiona occs
                2.0     Lower Managerial and professional occs
                3.0     Intermediate occupations
                4.0     Small employers and own account workers
                5.0     Lower supervisory and technical occupations
                6.0     Semi-routine occupations
                7.0     Routine occupations
                8.0     Never worked and long term unemployed
               97.0     Not classified

jobtim3 (9) time in current job grouped
      Measurement Level: Ordinal
              Value     Label

              -9.00     DNA
              -8.00     NA
              -6.00     CHILD/NO INT/PROXY
               1.00     less than 2yrs
               2.00     2<5yrs
```

```
            3.00     5 or more years


genhlth (10) Health on the whole in last 12 months
     Measurement Level: Ordinal


            Value     Label

             -9       DNA
             -8       NA
             -6       NO INT/PROXY/CHILD
              1       Good
              2       Fairly Good
              3       Not Good



workhrs (11) Total work hours
     Measurement Level: Ordinal


            Value     Label

            -9.00     DNA
            -8.00     NA
            -6.00     CHILD/NO INT/PROXY


ftptwk (12) Working full or part-time
     Measurement Level: Ordinal


            Value     Label

             -9       DNA
             -8       NA
             -6       CHILD/NO INT/PROXY
              1       full time
              2       or part time?


hiqual (13) Highest qualification
     Measurement Level: Ordinal


            Value     Label

             -8       NA
             -6       CHILD/OUT AGE/NO INT
              1       Higher Degree
              2       NVQ Level 5
              3       First Degree
              4       Other Degree
              5       NVQ Level 4
              6       Diploma in Higher Education
              7       HNC/HND BTEC Higher etc
              8       Teaching - Key Stage 2 (7-11)
              9       Teaching - Key Stage 1 (5-7)
             10       Teaching - Foundation Stage (3-5)
             11       Teaching - level not stated
             12       Nursing etc
             13       RSA Higher Diploma
             14       Other higher education below degree
             15       NVQ Level 3
```

```
            16      GNVQ/GSVQ Advanced
            17      A Level/Vocational A Level or equivalent
            18      RSA Advanced Diploma
            19      OND/ONC/BTEC/SCOTVEC
            20      City & Guilds Advanced Craft /Part3
            21      Certificate of Sixth Year Studies (CSYS) (Scotland)
            22      Highers
            23      A/S Level or equivalent
            24      Trade Apprenticeship
            25      NVQ Level 2 or equivalent
            26      GNVQ full intermediate
            27      RSA Diploma
            28       City & Guilds Craft/Part 2
            29      BTEC/SCOTVEC First or Gen Diploma etc
            30      O Level, GCSE Grase A*-C or equivalent
            31      NVQ Level 1 or equivalent
            32      GNVQ/GSVQ Part One Foundation
            33      CSE below grade 1, GCSE below grade C
            34      BTEC/SCOTVEC first ot gen certificate
            35      SCOTVEC modules
            36      RSA Other
            37      City & Guilds Foundation/Part 1
            38      YT/YTP certificate
            39      Otherqualification
            40      No qualifications
            41      Dont Know
            42      Teaching - Key Stage 3 (11-14)
            43      Teaching - Key Stage 4 (14-16)
            44      Teaching - Further Education
            45      Foundation degree
            46      GNVQ/GSVQ Full Foundation
            47      GNVQ/GSVQ Part One Intermediate

edlev00 (14) Education Level - 2000
     Measurement Level: Ordinal


            Value     Label

            -9      Never attended school
            -8      NA
            -6      CHILD/OUT AGE/NO INT
             1       Higher Degree
             2      First Degree
             3      Teaching qualification
             4      Other higher qualification
             5      Nursing qualification
             6      GCE A level in two or more subjects
             7      GCE A level in one subject
             8      GCSE/Olevel, standard grades, 5+
             9      GCSE/Olevel 1-4
            10      CSE below grade 1, GCSE below grade C
            11      Apprenticeship
            12      Other qualification
            13      no qualification

edlev7 (15) Education Level - 2000 (3 groups)
     Measurement Level: Ordinal
```

```
           Value     Label

            -9       Never attended school
            -8       NA
            -6       AGEOUT,MSPR
             1       Higher Education
             2       Other qualification
             3       No qualification

drweek (16) estimated weekly units: all drinks
     Measurement Level: Scale
```

This variable is a simplified version of drating as it provides the number of units drunk per week rounded to the nearest integer. This was carried out using the following syntax in SPSS:

```
compute drweek=rnd(drating).
```

```
drcat (17) Weekly alcohol consumption level
     Measurement Level: Ordinal

           Value     Label

            .00      None
           1.00      Moderate drinking
           2.00      Heavy drinking
```

This variable is derived from the variable drweek to indicate an individual's weekly alcohol consumption in broad categories. This variable has been created separately for men and women because the guidelines about alcohol consumption are different for men and women. The syntax used to create the variable is as follows:

```
do if sex=1.
recode drweek (0=0) (1 thru 21=1) (22 thru hi=2) into drcat.
end if.

do if sex=2.
recode drweek (0=0) (1 thru 14=1) (15 thru hi=2) into drcat.
end if.
```

NCDS_ExpData_teach.sav

This dataset includes 32 variables extracted from the first eight sweeps of a cohort of individuals born in Britain in one week of 1958 (known as the National Child Development Study). More information is provided about the study in Chapter 3. The data is longitudinal in that it has been collected from the same individuals over the first forty-six years of their lives. There are a total of 18558 cases in the dataset but some individuals have died, emigrated or dropped out of the study over the years and so the sample with valid data in adulthood is considerably smaller than this. For example there are 9533 cases with valid data at age 46.

The original datasets from which this teaching dataset was created are deposited at the UK Data Archive as follows:

SN 5565 -National Child Development Study: Childhood Data, Sweeps 0-3, 1958-1974

SN 5566 -National Child Development Study: Sweep 4, 1981, and Public Examination Results, 1978

SN 5567 -National Child Development Study: Sweep 5, 1991

SN 5578 -National Child Development Study: Sweep 6, 1999-2000

SN 5579 -National Child Development Study: Sweep 7, 2004-2005

```
Variables

Name (Position) Label

ncdsid (1) NCDS serial number
    Measurement Level: Nominal
    Column Width: 7  Alignment: Left
    Print Format: A7
    Write Format: A7

sex (2) 0-3D Sex of child
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -1

          Value     Label

            -1 M  Not known
             1     Male
             2     Female

n492 (3) Social class mother's husband at child's birth (GRO 1951)
    Measurement Level: Ordinal
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: 9 thru 999, -1
```

```
          Value     Label

            -1 M  NA,no husband
             1    Unemployed,sick
             2    I
             3    II
             4    III
             5    IV
             6    V
             9 M  Students
            10 M  Dead or away
            11 M  Retired
            12 M  Single,no husbnd
           999 M  Not in survey at birth


_



n1840 (4) 1T Draw-a-man test score at age 7
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -999 thru -1

          Value     Label

           -999 M  Not in survey at age 7
           -888 M  Other missing
             -2 M  No educational
             -1 M  NA,no score

n920 (5) 2T Total score on general ability test at age 11
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -999 thru -1

          Value     Label

           -999 M  Not in survey at age 11
           -888 M  Other missing
             -1 M  NA

n923 (6) 2T Reading comprehension test score at age 11
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8
    Write Format: F8
    Missing Values: -999 thru -1

          Value     Label

           -999 M  Not in survey at age 11
```

```
                -888 M  Other missing
                  -1 M  NA


n926 (7) 2T Mathematics test score at age 11
     Measurement Level: Scale
     Column Width: 8  Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

               Value     Label

                -999 M  Not in survey at age 11
                -888 M  Other missing
                  -1 M  NA

_



n851 (8) 2S Fathers' interest in childs education at age 11
     Measurement Level: Ordinal
     Column Width: 8  Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

               Value     Label

                -999 M  Not in survey at age 11
                -888 M  Other missing
                  -2 M  No dad,inapplic
                  -1 M  NA
                   1    Over concerned
                   2    Very interested
                   3    Some interest
                   4    Little interest
                   5    Cant say

n852 (9) 2S Mothers' interest in childs education at age 11
     Measurement Level: Ordinal
     Column Width: 8  Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

               Value     Label

                -999 M  Not in survey at age 11
                -888 M  Other missing
                  -2 M  No mum,inapplic
                  -1 M  NA
                   1    Over concerned
                   2    Very interested
                   3    Some interest
                   4    Little interest
                   5    Cant say
```

—

n2384 (10) 3P Social class father,male head (GRO 1970) at age 16
     Measurement Level: Ordinal
     Column Width: 8 Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

             Value   Label

              -999 M Not in survey at age 16
              -888 M Other missing
                -1 M NA,NMH,forces
                 1   I
                 2   II
                 3   III non-manual
                 4   III manual
                 5   IV non-manual
                 6   IV manual
                 7   V
                 8   Unclear

n2928 (11) 3T Test 1-reading comprehension at age 16
     Measurement Level: Scale
     Column Width: 8 Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

             Value   Label

              -999 M Not in survey at age 16
              -888 M Other missing
                -1 M NA

n2930 (12) 3T Test 2-mathematics comprehension at age 16
     Measurement Level: Scale
     Column Width: 8 Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

             Value   Label

              -999 M Not in survey at age 16
              -888 M Other missing
                -1 M NA

—

n0region (13) Region at PMS (1958) - Birth
     Measurement Level: Scale

```
Column Width: 8  Alignment: Right
Print Format: F2
Write Format: F2
Missing Values: -2

        Value    Label

          -2 M  Not in PMS
          -1    No imformation
           1    North
           2    North West
           3    E & W.Riding
           4    North Midlands
           5    Midlands
           6    East
           7    South East
           8    South
           9    South West
          10    Wales
          11    Scotland

n3region (14) Region at NCDS3 (1974) - 16 years
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F2
    Write Format: F2
    Missing Values: ** thru -1

        Value    Label

          ** M  Not in survey at age 16
          ** M  Other missing
          -2 M  Not in PMS
          -1 M  No imformation
           1    North
           2    North West
           3    E & W.Riding
           4    North Midlands
           5    Midlands
           6    East
           7    South East
           8    South
           9    South West
          10    Wales
          11    Scotland

_



dvht16 (15) 3D Height in metres at 16 years
    Measurement Level: Scale
    Column Width: 10  Alignment: Right
    Print Format: F10.3
    Write Format: F10.3
    Missing Values: -999.000 thru -1.000
```

11

```
              Value     Label

          -999.000 M  Not in survey at age 16
          -888.000 M  Other missing
            -1.000 M  Not recorded


newghsq (16) 4D 1981 ghs qualif classification-revised at age 23
     Measurement Level: Ordinal
     Column Width: 8  Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

              Value     Label

            -999 M  Not in survey at age 23
            -888 M  Other missing
               1    HIGHER  DEGREE
               2    DEGREE
               3    TEACHERS
               4    HTECHNCD ETC
               5    NURSES
               6    2+A S
               7    A OR ONC OR TEC
               8    5+O-S OR CRAFT
               9    O S+
              10    O-S ONLY
              11    NO O-S  -SMETHNG
              12    CSE S
              13    APPRENT-ICESHIP
              15    OTHER
              16    NONE


_



earngsn (17) 4D Net earnings per week from all jobs at age 23
     Measurement Level: Scale
     Column Width: 10  Alignment: Right
     Print Format: F10.2
     Write Format: F10.2
     Missing Values: -999.00 thru -1.00

              Value     Label

          -999.00 M  Not in survey at age 23
          -888.00 M  Other missing

n4269ph (18) 4D Gross earnings per hour at age 23
     Measurement Level: Scale
     Column Width: 8  Alignment: Right
     Print Format: F8
     Write Format: F8
     Missing Values: -999 thru -1

              Value     Label
```

```
                      -999 M  Not in survey at age 23
                      -888 M  Other missing
                        -1 M  HOURS VARY


dvht23 (19) 4D Height in metres at 23 years
      Measurement Level: Scale
      Column Width: 10  Alignment: Right
      Print Format: F10.3
      Write Format: F10.3
      Missing Values: -999.000 thru -1.000

              Value      Label

          -999.000 M  Not in survey at age 23
          -888.000 M  Other missing
            -1.000 M  Missing


mal (20) 4D Total malaise inventory score < na at age 23
      Measurement Level: Nominal
      Column Width: 8  Alignment: Right
      Print Format: F8
      Write Format: F8
      Missing Values: -999 thru -1

              Value      Label

               -999 M  Not in survey at age 23
               -888 M  Other missing
                 -1 M  Missing


_


n503913 (21) CMI:74 F1 General health at age 33
      Measurement Level: Ordinal
      Column Width: 8  Alignment: Right
      Print Format: F8
      Write Format: F8
      Missing Values: -999 thru -1

              Value      Label

               -999 M  Not in survey at age 33
               -888 M  Other missing
                  1     Excellent
                  2     Good
                  3     Fair
                  4     Poor

hqual33 (22) DV:Highest qual gained at age 33
      Measurement Level: Scale
      Column Width: 8  Alignment: Right
      Print Format: F2
      Write Format: F2
      Missing Values: ** thru -1
```

```
            Value     Label

             ** M  Not in survey at age 33
             ** M  Other missing
             -1 M  No information
              0    No qualification
              1    CSE 2-5/equiv NVQ1
              2    O Level/equiv NVQ2
              3    A Level/equiv NVQ3
              4    Higher qual NVQ4
              5    Degree/higher NVQ5,6
```

_

hlth42 (23) How is your health generally at age 42
    Measurement Level: Ordinal
    Column Width: 7 Alignment: Right
    Print Format: F1
    Write Format: F1
    Missing Values: * thru *

```
            Value     Label

              * M  Not in survey at age 42
              * M  Other missing
              * M  Not answered
              * M  Dont know
              1    Excellent
              2    Good
              3    Fair
              4    Poor
              8    Dont know
              9    Not answered
```

sc42 (24) (Current Job) Social Class at age 42
    Measurement Level: Ordinal
    Column Width: 4 Alignment: Right
    Print Format: F4.1
    Write Format: F4.1
    Missing Values: -999 thru -1.0

```
            Value     Label

            -999 M  Not in survey at age 42
            -888 M  Other missing
            -9.0 M  Not answered
            -8.0 M  Dont know
             1.0    I Professional
             2.0    II Managerial-technical
             3.1    IIINM Skilled non-manual
             3.2    IIIM Skilled manual
             4.0    IV Partly skilled
             5.0    V Unskilled
             6.0    Others
```

14

—

n6gor (25) NCDS6 Government Office Region at age 42
    Measurement Level: Scale
    Column Width: 8 Alignment: Right
    Print Format: F2
    Write Format: F2
    Missing Values: ** thru -1

            Value   Label

              ** M Not in survey at age 42
              ** M Other missing
              -9 M Not answered
              -8 M Dont know
               1   North East
               2   North West
               3   Yorkshire and the Humber
               4   East Midlands
               5   West Midlands
               6   East of England
               7   London
               8   South East
               9   South West
              10   Wales
              11   Scotland
              12   Not recorded'
              13   Not app in IoM Ch Is

n7gor (26) Government Office Region at Interview at age 46
    Measurement Level: Scale
    Column Width: 8 Alignment: Right
    Print Format: F4
    Write Format: F4
    Missing Values: -999 thru -1

            Value   Label

            -999 M Not in survey at age 46
            -888 M Other missing
              -1 M Unknown
               1   North East
               2   North West
               3   Yorkshire and The Humber
               4   East Midlands
               5   West Midlands
               6   East of England
               7   London
               8   South East
               9   South West
              10   Wales
              11   Scotland

—

15

```
sc46 (27) (Current Job) Social Class at age 46
    Measurement Level: Scale
    Column Width: 3  Alignment: Right
    Print Format: F5.1
    Write Format: F5.1
    Missing Values: -999 thru -1.0

            Value      Label

             -999 M  Not in survey at age 46
             -888 M  Other missing
             -9.0 M  Refusal
             -8.0 M  Don't Know
             -7.0 M  Other missing
             -3.0 M  insufficient information to code
             -1.0 M  Not applicable
              1.0     I Professional
              2.0     II Managerial-technical
              3.1     IIINM Skilled non-manual
              3.2     IIIM Skilled manual
              4.0     IV Partly skilled
              5.0     V Unskilled
              6.0     Others

hlth46 (28) Self-assessment of health at age 46
    Measurement Level: Scale
    Column Width: 7  Alignment: Right
    Print Format: F2
    Write Format: F2
    Missing Values: ** thru -1

            Value      Label

               ** M  Not in survey at age 46
               ** M  Other missing
               -9 M  Refusal
               -8 M  Don't Know
               -7 M  Other missing
               -1 M  Not applicable
                1     Excellent
                2     Good
                3     Fair
                4     Poor
                5     Very poor


_


nd7ns8 (29) (Current Job) Social Class at age 46 ns-sec
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F2
    Write Format: F2
```

```
    Missing Values: ** thru -1

          Value     Label

            ** M  Not in survey at age 46
            ** M  Other missing
            -9 M  Refusal
            -8 M  Don't know
            -7 M  Other missing
            -3 M  insufficient information to code
            -1 M  Not applicable
             1    Higher managerial and professional occupations
             2    Lower managerial and professional occupations
             3    Intermediate occupations
             4    Small Employers and Own account workers
             5    Lower supervisory and technical occupations
             6    Semi-routine occupations
             7    Routine occupations
             8    Never worked and long-term unemployed

ht7cm (30) Height in cm without shoes at age 7
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8.2
    Write Format: F8.2
    Missing Values: -999.00 thru -1.00

          Value     Label

        -999.00 M  Not in survey at age 7
        -888.00 M  Other missing

sc23 (31) Current or last job 1980 social class at age 23
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
    Print Format: F8.2
    Write Format: F8.2
    Missing Values: -999.00 thru -1.00

          Value     Label

        -999.00 M  Not in survey at age 23
        -888.00 M  Other missing
           1.00    Professional
           2.00    Intermediate
           3.00    Skilled non-manual
           4.00    Skilled manual
_


           5.00    Semi-skilled
           6.00    Unskilled manual
           7.00    Armed forces etc

sc33 (32) CURRENT/LAST JOB: RGs Social Class 1991 at age 33
    Measurement Level: Scale
    Column Width: 8  Alignment: Right
```

```
Print Format: F8.2
Write Format: F8.2
Missing Values: -999.00 thru -1.00

       Value     Label

   -999.00 M  Not in survey at age 33
   -888.00 M  Other missing
      1.00     Professional
      2.00     Managerial\tech
      3.00     Skilled non-manual
      4.00     Skilled manual
      5.00     Partly skilled
      6.00     Unskilled


_
```

YCS11_teach.sav

This dataset is based on data extracted from the eleventh cohort of the Youth Cohort Study which started in 2002. Further details about the study are provided in the appendix to Chapter 6. The study is longitudinal in that it has contacted the same sample of individuals on an annual basis since spring 2002 when respondents were in year twelve, and age 16 or 17. The majority of the 10 variables included in this teaching dataset have been extracted from the first survey (and have the prefix s1). The main activity of individuals at age 19 in spring 2005 is included (s4act1). There are a total of 16707 cases in the dataset but not all of these have remained in the study so that there is only information about main activity at age 19 for 6180 individuals.

The original data set from which this teaching dataset was extracted is deposited at the UK Data Archive as the following dataset:

SN 5452 -Youth Cohort Study of England and Wales, 2002-2005; Cohort Eleven, Sweep One to Four

```
Variables

Name (Position) Label

sn (1) Serial Number
    Measurement Level: Ordinal
          Value     Label


            -9     Not answered (999999)
            -1     Item not applicable

s4weight (2) sweep 4 core weight
    Measurement Level: Scale

s4act1 (3) Main activity at sweep 4
    Measurement Level: Ordinal
    Missing Values: -9.00 thru -1.00

          Value     Label

          -9.00 M  Not answered
           1.00     GST
           2.00     FT Education
           3.00     FT job
           4.00     Part-time job
           5.00     Looking after home or family
           6.00     Looking for work/ Unemployed
           7.00     Taking a break
           8.00     Doing something else


s1truan1 (4) Year 11 truancy
    Measurement Level: Ordinal

          Value     Label
```

```
            -9      Not stated (9)
            -1      Item not applicable
             1      Persistent
             2      Weeks at a time
             3      Several days at a time
             4      Occasional
             5      Particular days or lessons
             6      For the odd day or lesson
             7      Never
```

s1gor (5) Government Office Region
      Measurement Level: Nominal

```
            Value     Label

             1        North East
             2        North West
             3        Yorks and Humber
             4        E Midlands
             5        W Midlands
             6        East of England
             7        Inner London
             8        Outer London
             9        South East
            10        South West
            11        Wales
```

dadsec (6) father's grouped ns-sec
      Measurement Level: Ordinal

```
            Value     Label

            1.00      Large employers and higher professionals
            2.00      Lower professional and higher technical occupations
            3.00      Intermediate
            4.00      Lower supervisory occupations
            5.00      Semi routine and routine occupations
            6.00      Other
```

momsec (7) mother's grouped ns-sec
      Measurement Level: Ordinal

```
            Value     Label

            1.00      Large employers and higher professionals
            2.00      Lower professional and higher technical occupations
            3.00      Intermediate occupations
            4.00      Lower supervisory occupations
            5.00      Semi routine and routine occupations
            6.00      Other
```

famsec (8) family grouped ns-sec
      Measurement Level: Ordinal

```
            Value     Label
```

20

```
            1.00     Large employers and higher professionals
            2.00     Lower professional and higher technical occupations
            3.00     Intermediate
            4.00     Lower supervisory occupations
            5.00     Semi routine and routine occupations
            6.00     Other

s1pared (9) parent's education
     Measurement Level: Ordinal

             Value     Label

            1.00     At least one parent with degree
            2.00     At least one parent with A level
            3.00     Neither parent with A level

s1sex (10) Respondent's gender
     Measurement Level: Nominal
             Value     Label

             -1     Item not applicable
              1     Male
              2     Female
```

WorldBank_2000_Teach.sav

This dataset is somewhat different from the other teaching datasets included for use with this book because the unit of analysis is a country rather than an individual person. Further information about the data can be found in Chapter 10. There are a total of 13 variables included here which have been selected because of their relevance to the Millennium Development Goals discussed in Chapter 10. A total of 208 countries are included in the dataset but there is a considerable amount of missing data due to the difficulties of collecting comparable data from all countries in the World. The reliability and validity of the data for different countries varies enormously so not too much reliance should be placed on the figure for any one variable in any one country. Further support on accessing and analysing this type of data is available from ESDS International, their website is at: http://www.esds.ac.uk/international.

Variables

```
Name (Position) Label

Countrygp (1) Country Group
    Measurement Level: Nominal


        Value      Label

         1.00      South and East Asia
         2.00      Africa
         3.00      South America and Caribbean
         4.00      Middle East
         5.00      Western Europe etc.
         6.00      Ex-State Socialist

Country (2) Country
    Measurement Level: Nominal

CO2 (3) CO2 emissions per capita
    Measurement Level: Scale

GNIPCAP (4) GNI per capita Atlas method US$
    Measurement Level: Scale

Immun (5) Immunization, measles (% of children ages 12-23 months)
    Measurement Level: Scale

LIFEXP (6) Life expectancy
    Measurement Level: Scale


MALNUT (7) Malnutrition prevalence, weight for age (% of children under 5)
    Measurement Level: Scale

INFMORT (8) Mortality rate, infant (per 1,000 live births)
    Measurement Level: Scale

UND5MORT (9) Mortality rate, under-5 (per 1,000)
```

Measurement Level: Scale

POV (10) Poverty headcount ratio at $1 a day (PPP) (% of population)
        Measurement Level: Scale

FMTED (11) Ratio of female to male enrollments in tertiary education
        Measurement Level: Scale

FMPED (12) Ratio of female to male primary enrollment
        Measurement Level: Scale

FMSED (13) Ratio of female to male secondary enrollment
        Measurement Level: Scale

BSA2004_Teach.sav

This dataset includes 49 variables from the 2004 British Social Attitude Survey. Further information about this study can be found in the appendix to Chapter 12. The total sample size is 3199 individuals. However, three different versions of the questionnaire were used in the fieldwork. Respondents were randomly assigned to one of the versions. Modules may have been included on one, two or three versions of the questionnaire (thus giving different sample sizes). The 2004 questionnaire covered a broad range of topics including attitudes to social housing, redistribution, views of the welfare state, questions about citizenship, expectations of and satisfaction with the National Health Service, ways of improving primary and secondary education, management-employee relations and life-cycle events.

The original dataset from which this teaching dataset was extracted is deposited at the UK Data Archive as the following dataset:

SN 5329 -British Social Attitudes Survey, 2004

```
List of variables

Name (Position) Label

Serial (1) Serial Number                        Q1
    Measurement Level: Nominal

SPoint (2) Sample point  <167-299>              Q9
    Measurement Level: Nominal

StRegion (3) Standard Region [11 categories] dv   Q11
    Measurement Level: Nominal

        Value     Label

           1      Scotland
           2      North East
           3      North West
           4      Yorkshire&Humberside
           5      West Midlands
           6      East Midlands
           7      Eastern
           8      South West
           9      South East
          10      London
          11      Wales


Region (4) Region          [6 categories]dv     Q12
    Measurement Level: Nominal

        Value     Label

           1      Scotland
           2      NE+NW+York&Humb
           3      E+W Midlands
```

```
                4       Wales
                5       SW+SE+Eastern
                6       London

Country (5) Country: England, Scotland or Wales? Q28
        Measurement Level: Nominal

                Value     Label

                1       England
                2       Scotland
                3       Wales

RSex (6) Sex of Respondent                    Q39
        Measurement Level: Nominal

                Value     Label

                1       Male
                2       Female

RAge (7) Respondent's age in years           Q40
        Measurement Level: Scale
        Missing Values: 99

                Value     Label

                97      97+
                98      Don't   know
                99 M    Not     answered


PeopTrst (8) People can be trusted/can't be too careful?A2.13
        Measurement Level: Ordinal
        Missing Values: *, *

                Value     Label

                * M     Skp,B+C version
                * M     No self-compltn
                1       Almost always trust
                2       Usually be trusted
                3       Usually cantbe2careful
                4       Almost always cantbe2careful
                8       Can't   choose
                9       Not     answered

RNSSEC (9) NS-SEC - long version              Q519
        Measurement Level: Ordinal

                Value     Label

                -1.0    Skp,nvr had job
                 1.0    Employers in large organisations
                 2.0    Higher managerial occupation
                 3.1    Higher prof: traditional employees
                 3.2    Higher prof: new employees
```

25

```
     3.3     Higher prof: traditional self-employed
     3.4     Higher prof: new'self-employed
     4.1     Lower prof&higher tech: traditional employees
     4.2     Lower prof&higher tech: new employees
     4.3     Lower prof&higher tech: traditional self-employed
     4.4     Lower prof&higher tech: new self-employed
     5.0     Lower managerial occupations
     6.0     Higher supervisory occupations
     7.1     Intermed: clerical and administrative
     7.2     Intermed: sales and services
     7.3     Intermed: technical and auxiliary
     7.4     Intermed: engineering
     8.1     Small employers: non-professional
     8.2     Small employers: agriculture
     9.1     Own account workers: non-professional
     9.2     Own account workers: agriculture
    10.0     Lower supervisory occupations
    11.1     Lower techn: craft
    11.2     Lower techn: process operative
    12.1     Semi-routine: sales
    12.2     Semi-routine: service
    12.3     Semi-routine: technical
    12.4     Semi-routine: operative
    12.5     Semi-routine: agriculture
    12.6     Semi-routine: clerical
    12.7     Semi-routine: childcare
    13.1     Routine: sales and service
    13.2     Routine: production
    13.3     Routine: technical
    13.4     Routine: operative
    13.5     Routine:  agriculture
    14.1     Never worked
    14.2     Long-term unemployed
    15.0     Full-time students
    16.0     Not stated or inadeq described
    17.0     Not classifiable for other reason
    98.0     Don't   know
    99.0     Not     answered

REconAct (10) Respondent's main economic activity last week? Q539
     Measurement Level: Nominal

          Value    Label

             1     In full-time education (not paid for by
             2     On government training/employment progra
             3     In paid work (or away temporarily) for a
             4     Waiting to take up paid work already acc
             5     Unemployed and registered at a benefit o
             6     Unemployed, not registered, but acti
             7     Unemployed, wanting a job (of at least 1
             8     Permanently sick or disabled
             9     Wholly retired from work
            10     Looking after the home
            11     (Doing something else) (WRITE IN)
            98     Don't   know
            99     Not     answered
```

```
Tea (11) Terminal education age<categorised> Q766
     Measurement Level: Ordinal


           Value     Label

              1      15 or under
              2      16
              3      17
              4      18
              5      19 or over
              6      Still at school
              7      Still at college or university
             97      Other Answer(Write In)
             98      Don't   know
             99      Not     answered


Tea2 (12) How old was Respondent when left Full Time Education?Q763
     Measurement Level: Ordinal
           Value     Label

             95      at school
             96      at college/univ
             97      Other answer
             98      Don't   know
             99      Not     answered


HEdQual (13) Highest educational qualification obtained   Q829
     Measurement Level: Ordinal


           Value     Label

              1      Degree
              2      Higher educ below degree
              3      A level or equiv
              4      O level or equiv
              5      CSE or equiv
              6      Foreign or other
              7      No qualification
              8      DK/Refusal/NA



HEdQual2 (14) Highest educational qualification obtained [postgrad]Q830
     Measurement Level: Ordinal


           Value     Label

              1      Postgraduate degree
              2      First degree
              3      Higher educ below degree
              4      A level or equiv
              5      O level or equiv
              6      CSE or equiv
              7      Foreign or other
              8      No qualification
              9      DK/Refusal/NA
```

27

HHincome (15) How much household income last year Q972
    Measurement Level: Ordinal

            Value   Label

              1   Q ls thn   3999
              2   T 4000-   5999
              3   O 6000-   7999
              4   K 8000-   9999
              5   L 10000-   11999
              6   B 12000-   14999
              7   Z 15000-   17999
              8   M 18000-   19999
              9   F 20000-   22999
             10   J 23000-   25999
             11   D 26000-   28999
             12   H 29000-   31999
             13   A 32000-   37999
             14   W 38000-   43999
             15   G 44000-   49999
             16   N 50000-   55999
             17   E 56000 or more
             97   Refused information
             98   Don't   know
             99   Not   answered

UnionSA (16) Is Respondent a member of a Trade Union/S.A.?Q604
    Measurement Level: Nominal

            Value   Label

              *   Skp,nvr had job
              1   Yes, trade union
              2   Yes, staff association
              3   No
              8   Don't   know
              9   Not   answered

TUSAEver (17) Has Respondent ever been a member of a TU/SA Q605
    Measurement Level: Nominal

            Value   Label

              *   Skp,Now member
              *   Skp,nvr had job
              1   Yes, trade union
              2   Yes, staff association
              3   No
              8   Don't   know
              9   Not   answered

NghBrHd (18) How long Respondent lived in neighbourhood? Q632
    Measurement Level: Scale

            Value   Label

28

```
                    98    Don't   know
                    99    Not     answered




ResPres (19) Respondent live in city/suburbs or country?  Q697
     Measurement Level: Nominal

             Value     Label

                 1    a big city
                 2    the suburbs/outskirts of a big city
                 3    a small city or town
                 4    a country village
                 5    a farm or home in the country
                 7    (Other answer (WRITE IN))
                 8    Don't   know
                 9    Not     answered

Voted01 (20) Did Respondent vote 2nd June 2001 in G.E.?   Q834
     Measurement Level: Nominal

             Value     Label

                 1    Yes, voted
                 2    No
                 8    Don't   know
                 9    Not     answered

PdJobSC (21) Respondent currently in paid work?A2.48B2.22C2.23
     Measurement Level: Nominal
     Missing Values: *, *

             Value     Label

               * M  No self-compltn
                 1    Yes, 10+hrs
                 2    No
                 8    Don't   know
                 9    Not     answered

CharOft (22) Respondent give money to charity how often?  B619
     Measurement Level: Ordinal
     Missing Values: *

             Value     Label

               * M  Skp, A+C version
                 1    Never
                 2    Occasionally but less often than once a year
                 3    Once/twice a year
                 4    every few months
                 5    Once/twice a month
                 6    Once/week or more
                 8    Don't know
```

```
                    9    Not answered

CharAmt (23) Respondent gives how much to charity per yearB620
    Measurement Level: Ordinal
    Missing Values: *

            Value    Label

            * M  Skp, A+C version
            *    Skp,never gives
            1    Less than 5gbp
            2    5.00-12gbp
            3    12.01- 50gbp
            4    50.01- 120gbp
            5    120.01- 500gbp
            6    More than 500gbp
            8    Don't know
            9    Not answered


PolitSC (24) How interested Respondent in politics?      A2.10
    Measurement Level: Ordinal
    Missing Values: *, *

            Value    Label

            * M  Skp,B+C version
            * M  No self-compltn
            1    Very interested
            2    Fairly interested
            3    Not very interested
            4    Notatall interested
            8    Can't   choose
            9    Not     answered

GovRight (25) Trust people in government to do what is rightA2.11a
    Measurement Level: Ordinal
    Missing Values: *, *

            Value    Label

            * M  Skp,B+C version
            * M  No self-compltn
            1    Agree strongly
            2    Agree
            3    Neither agree nor disagree
            4    Disagree
            5    Disagree strongly
            8    Can't   choose
            9    Not     answered

GovSelf (26) Politicians for what get out of itA2.11b
    Measurement Level: Ordinal
    Missing Values: *, *

            Value    Label
```

```
                    *  M   Skp,B+C version
                    *  M   No self-compltn
                    1      Agree strongly
                    2      Agree
                    3      Neither agree nor disagree
                    4      Disagree
                    5      Disagree strongly
                    8      Can't   choose
                    9      Not     answered


DemoNow (27) How well GB democracy work today? A2.25a
    Measurement Level: Ordinal
    Missing Values: -1, -2

            Value     Label

            -2 M  Skp,B+C version
            -1 M  No self-compltn
             0    Very poorly
            10    Very well
            98    Can't   choose
            99    Not     answered


GrpMemb1 (28) Respondent would/did join a political party?A2.6a
    Measurement Level: Ordinal
    Missing Values: *, *

            Value     Label

            *  M   Skp,B+C version
            *  M   No self-compltn
            1      Belong+active
            2      Belong+not participate
            3      Used to belong
            4      Never belonged
            8      Can't   choose
            9      Not     answered


GrpMemb2 (29) Respondent would/did join TU,profession association?A2.6b
    Measurement Level: Ordinal
    Missing Values: *, *

            Value     Label

            *  M   Skp,B+C version
            *  M   No self-compltn
            1      Belong+active
            2      Belong+not participate
            3      Used to belong
            4      Never belonged
            8      Can't   choose
            9      Not     answered


GrpMemb3 (30) Respondent would/did join church, religious associationA2.6c
    Measurement Level: Ordinal
    Missing Values: *, *
```

```
          Value     Label

          * M  Skp,B+C version
          * M  No self-compltn
          1    Belong+active
          2    Belong+not participate
          3    Used to belong
          4    Never belonged
          8    Can't   choose
          9    Not     answered




GrpMemb4 (31) Respondent would/did join sports/cultural groupA2.6d
     Measurement Level: Ordinal
     Missing Values: *, *

          Value     Label

          * M  Skp,B+C version
          * M  No self-compltn
          1    Belong+active
          2    Belong+not participate
          3    Used to belong
          4    Never belonged
          8    Can't   choose
          9    Not     answered

GrpMemb5 (32) Respondent would/did join voluntary associationA2.6e
     Measurement Level: Ordinal
     Missing Values: *, *

          Value     Label

          * M  Skp,B+C version
          * M  No self-compltn
          1    Belong+active
          2    Belong+not participate
          3    Used to belong
          4    Never belonged
          8    Can't   choose
          9    Not     answered


IncDiffs (33) Income differences GB to large?A2.39aB2.13a
     Measurement Level: Ordinal
     Column Width: 8  Alignment: Right
     Print Format: F1
     Write Format: F1
     Missing Values: *, *

          Value     Label

          * M  Skp, C  version
          * M  No self-compltn
```

```
                1     Strongly agree
                2     Agree
                3     Neither agree/disagree
                4     Disagree
                5     Strongly disagree
                8     Can't choose
                9     Not      Answered

IncDiff (34) Government should reduce income differences.?A2.39bB2.13b
     Measurement Level: Ordinal
     Column Width: 7  Alignment: Right
     Print Format: F1
     Write Format: F1
     Missing Values: *, *

               Value    Label

                * M   Skp, C  version
                * M   No self-compltn
                1     Strongly agree
                2     Agree
                3     Neither agree/disagree
                4     Disagree
                5     Strongly disagree
                8     Can't choose
                9     Not      Answered


IDCards (35) Every adult should have an ID card? A2.61d
     Measurement Level: Ordinal
     Missing Values: *, *

               Value    Label

                * M   Skp,B+C version
                * M   No self-compltn
                1     Agree    strongly
                2     Agree
                3     Neither
                4     Disagree
                5     Disagree strongly
                8     Don't    know
                9     Not      answered

MoreWelf (36) Government should spend > on poor      A2.63bBC2.44b
     Measurement Level: Ordinal
     Missing Values: *, *

               Value    Label

                * M   No self-compltn
                1     Agree    strongly
                2     Agree
                3     Neither agree/disagree
                4     Disagree
                5     Disagree strongly
                8     Don't    know
```

```
                 9    Not     Answered


SocHelp (37) People not deserve any help   A2.63dBC2.44d
     Measurement Level: Ordinal
     Missing Values: *, *

            Value    Label

            * M  No self-compltn
            1    Agree    strongly
            2    Agree
            3    Neither agree/disagree
            4    Disagree
            5    Disagree strongly
            8    Don't   know
            9    Not     Answered

DeathApp (38) Death penalty appropriate for some crimesA2.66cB2.47cC2.46c
     Measurement Level: Ordinal
     Missing Values: *, *

            Value    Label

            * M  No self-compltn
            1    Agree strongly
            2    Agree
            3    Neither
            4    Disagree
            5    Disagree strongly
            8    Can't   choose
            9    Not     answered



WrongLaw (39) Always obey law even when law is wrongA2.66eB2.47eC2.46e
     Measurement Level: Ordinal
     Missing Values: *, *

            Value    Label

            * M  No self-compltn
            1    Agree strongly
            2    Agree
            3    Neither
            4    Disagree
            5    Disagree strongly
            8    Can't   choose
            9    Not     answered

PartyId2 (40) Party political identification (compressed)    dvQ211
     Measurement Level: Nominal

            Value    Label

            1    Conservative
            2    Labour
```

```
                3       Liberal Democrat
                4       Other party
                5       None
                6       Green Party
                8       Other/DK/Ref

agegp1 (41) Age grouped (1)
     Measurement Level: Ordinal

            Value     Label

             1.00     45 and under
             2.00     46 and over

This is a derived variable created using the following syntax
```

recode Rage (lo thru 45=1) (46 thru 97=2) into agegp1.
variable labels agegp1 'Age grouped (1)'.
value labels agegp1 1 '45 and under' 2 '46 and over'.

```
agegp2 (42) Age grouped (2)
     Measurement Level: Ordinal

            Value     Label

             1.00     Under 40
             2.00     40-59
             3.00     60 plus

This is a derived variable created using the following syntax
```

recode rage (lo thru 39=1) (40 thru 59=2) (60 thru 97=3) into agegp2.
variable labels agegp2 'Age grouped (2)'.
value labels agegp2 1 'Under 40' 2 '40-59' 3 '60 plus'.

```
incomegp (43) Household income last year grouped
     Measurement Level: Ordinal

            Value     Label

             1.00     Less than £20,000
             2.00     £20,000 or more

This is a derived variable created using the following syntax
```

recode hhincome (1 thru 8=1) (9 thru 17=2) into incomegp.

```
incomeg1 (44) Household income last year grouped
    Measurement Level: Ordinal

            Value    Label

            1.00     Less than £29,000
            2.00     £29,000 or more

This is a derived variable created using the following syntax
```

recode hhincome (1 thru 11=1) (12 thru 17=2) into incomeg1.

variable labels incomeg1 'Household income last year grouped'.
value labels incomeg1 1 'Less than £29,000' 2 '£29,000 or more'.

```
trust (45) Attitudes towards trusting people
    Measurement Level: Nominal

            Value    Label

            .00      Can't be too careful
            1.0      People can be trusted

This is a derived variable created using the following syntax
```

recode peoptrst (1 2=1) (3 4=0) into trust.

Value labels trust 1 'People can be trusted' 0 "Can't be too careful".
variable labels trust 'Attitudes towards trusting people'.

```
volgp (46) Ever been a member of a voluntary group?
    Measurement Level: Nominal
    Missing Values: -1.00

            Value    Label

            .00      No
            1.0      Yes
```

This is a derived variable created using the following syntax

```
recode grpmemb5 (1 2 3=1) ( 4 8 9=0) (else=-1) into volgp.
missing values volgp (-1).
value labels volgp 0 'No' 1 'Yes'.
variable labels volgp 'Ever been a member of a voluntary group?'
```

volgp1 (47) Currently a member of a voluntary group?
    Measurement Level: Nominal
    Missing Values: -1.00

            Value   Label

              .00   No
             1.0   Yes

This is a derived variable created using the following syntax

```
recode grpmemb5 (1 2 =1) ( 3 4 8 9=0) (else=-1) into volgp1.
missing values volgp1 (-1).
value labels volgp1 0 'No' 1 'Yes'.
variable labels volgp1 'Currently a member of a voluntary group?'
```

qual1 (48) Highest level of Qualifications (1)
    Measurement Level: Ordinal

            Value   Label

             1.00   A-level equiv or higher
             2.00   Olevel eqiv or lower

This is a derived variable created using the following syntax

```
recode hedqual (1 2 3=1) (4 5 6 7=2) into qual1.

value labels qual1 1 'A-level equiv or higher' 2 'Olevel eqiv or lower'.
Variable labels qual1 'Highest level of Qualifications (1)'.
```

qual2 (49) Highest level of Qualifications (2)
    Measurement Level: Ordinal

```
         Value     Label

         1.00      Degree or higher
         2.00      Alevel eqiv or lower

This is a derived variable created using the following syntax
```

recode hedqual (1 2 =1) (3 4 5 6 7=2) into qual2.
value labels qual2 1 'Degree or higher' 2 'Alevel eqiv or lower'.

Variable labels qual2 'Highest level of Qualifications (2)'.