[Draft only: 22 April 2020]

Commentary on SPSS files used in:

Earl Babbie, William E Wagner III and Jeanne Zaino

Adventures in Social Research: Data Analysis Using IBM SPSS Statistics
(10th edition, Sage 2019)

Page Contents

- 2 Book details
- 3 Getting the data
- 5 Opening the SPSS file
- 6 Making a copy of the file
- 9 Levels of Measurement
- 11 Names and Labels
- 15 Selecting the variables used in the book
- 21 Creating files for Gradpack users

Appendices

- 22 Variables in demo.sav
- 23 <u>Variables in exer.sav</u>
- 24 Variables common to both demo.sav and exer.sav.
- 24 Variables not in demo.sav or exer.sav.
- 26 Syntax to extract variables for demo.sav and exer.sav

Previously:

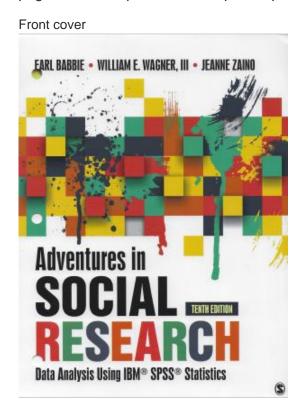
¹ [Retired]

Book details

Paperback is a hefty £65 from Sage UK: \$43 from Sage US. Slightly cheaper from <u>Amazon Kindle</u>, <u>Google Play</u>, <u>ebooks.com</u> and <u>Kobo</u>.

It comes in a large A4 format, 478 pages (two prefaces, 22 chapters, two appendices and an Index/glossary) and weighs 495 grams, Not easy to read in an armchair: you'll probably need a lectern!

There are 3 holes drilled right through the book and perforations along the edge of each page so that pages can be separated and kept in separate folder. This could be fun if students are sharing a copy!





Publisher's blurb on the rear cover says:

Inspire students to pursue their own adventures in social research with this practical, hands-on introduction to data conceptualization, measurement, and association through active learning.

Adventures in Social Research: Data Analysis Using IBM® SPSS® Statistics offers a practical, hands-on introduction to the logic of social science research for students in many disciplines. The fully revised Tenth Edition offers step-by-step instruction on data analysis using the latest version (24.0) of SPSS and current data from the General Social Survey. Organized to parallel most introductory research methods texts, this book starts with an overview of the social research process, then takes readers step by step through univariate, bivariate, and multivariate analysis using SPSS Statistics.

This text can be bundled with a Student Version of IBM® SPSS® software at a special price. Ask your SAGE sales representative for details.

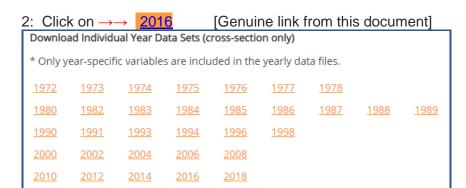
Getting the data

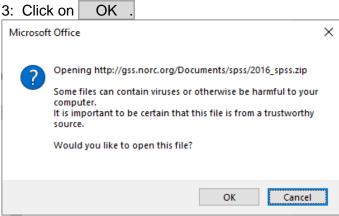
The authors use data from the NORC 2016 General Social Survey
The SPSS version of the data set is **GSS2016.SAV** and is freely downloadable from NORC.

NORC GSS page is Get the Data [Genuine link from this document]

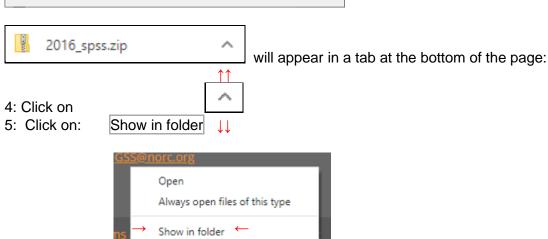


[Genuine link from this document]





Cancel



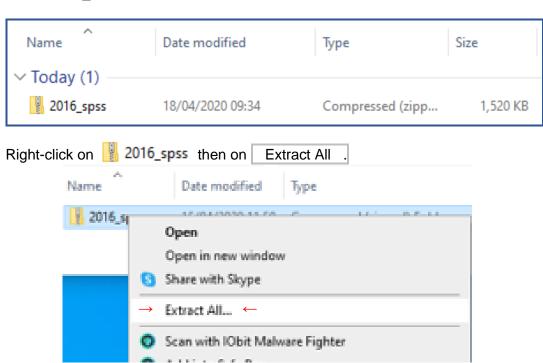
3

Zip folder 2016_spss appears in your **Downloads**

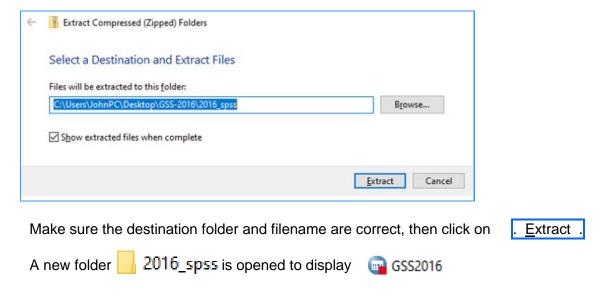
Make a new folder **GSS_2016** on your **Desktop** or other suitable location (eg. USB stick)



Folder GSS_2016

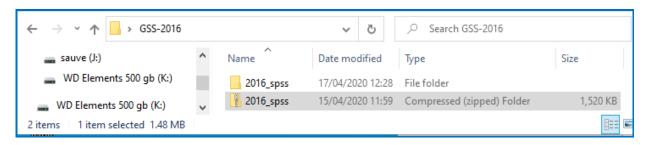


The following window appears:





There are also now two folders in GSS_2016:

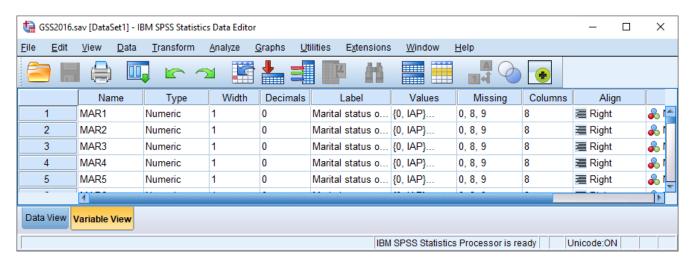


Opening the SPSS file

In folder ____ 2016_spss

Double click on GSS2016 to open the file:

File GSS2016.SAV opens in Variable View



[NB: There is no variable for case id]

Golden rule: Never work on an original file!

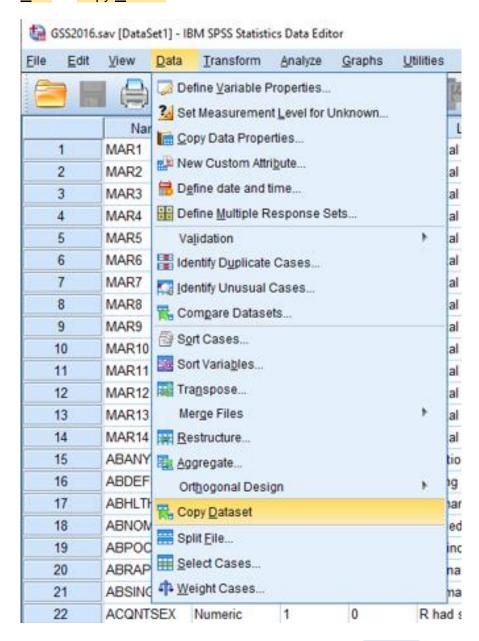
Make a copy!

... and work with that 2.

² These notes are primarily because I can't penetrate beyond the introduction to the companion website, so I've recreated the data sets from scratch. Luckily, the authors have used the same variables (from the 2008 GSS) except for two: **vote04** (Voted in 2004) is replaced by **vote12** (Voted in 2012) and **sei** (Socio-Economic Classification) by **sei10**.

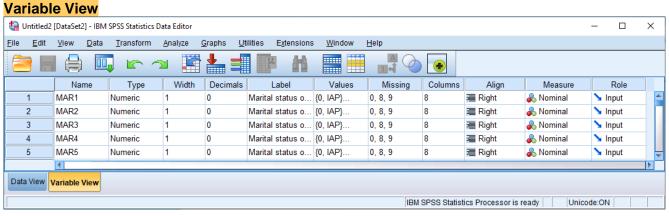
Making a copy of the file

Data >> Copy Dataset



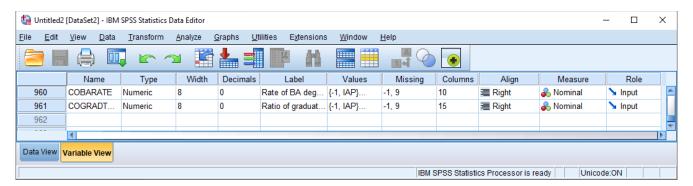
A new **Data Editor** is opened with the next available *Untitled name:

It immediately becomes the active file.

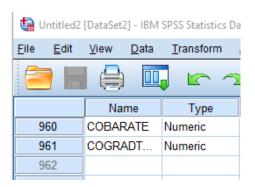


 $\downarrow\downarrow$

Use the vertical slide bar (on the right edge of the Data Editor) to scroll to the end of the file:

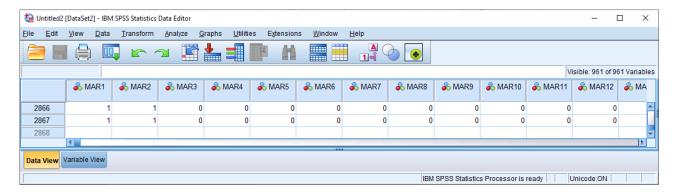


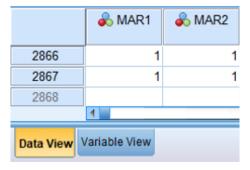
There are 961 variables:



Switch to **Data View**

Use the vertical slide bar (on the right edge of the Data Editor) to scroll to the end of the file:





There are 2867 cases.

[Commentary on SPSS files for Babbie et al 2019]

Levels of Measurement

The measurement levels for almost all variables are declared as (Nominal)

Correctly declared as (Nominal)

ABANY Abortion if woman wants for any reason

ABDEFECT Strong chance of serious defect

ABHLTH Woman's health seriously endangered

ABNOMORE Married--wants no more children

ABPOOR Low income--cant afford more children

ABRAPE Pregnant as result of rape

ABSINGLE Not married

COLATH Allow anti-religionist to teach **COLCOM** Should communist teacher be fired

COLHOMO Allow homosexual to teach

COLRAC Allow racist to teach
CONPRESS CONTV Confidence in press
CONTV Confidence in television
Whites hurt by aff. action

EVSTRAY Have sex other than spouse while married

GUNLAW Favor or oppose gun permits

MARITAL Marital status

PIKUPSEX R had sex with casual date last year

POSTLIFE Belief in life after death

RACDIF4 Differences due to lack of will

SEX Respondents sex

XMOVIE Seen x-rated movie in last year

Declared as (Nominal) but should be (Ordinal)

AFFRMACT Favor preference in hiring blacks **ATTEND** How often R attends religious services

CLASS Subjective class identification

INCOME
MARBLK
PARTYID
To help others
Total family income
Close relative marry black
Political party affiliation

POLVIEWS Think of self as liberal or conservative

POPULAR
PRAY
PREMARSX
RACE
RELIG
RINCOM16
To be well liked or popular
How often Does R pray
Sex before marriage
Race of respondent
R's religious preference
Respondents income

SEXFREQ Frequency of sex during last year

Items measured on Agree-Disagree scales declared as (Nominal) but should be (Ordinal)

FECHLD Mother working doesn't hurt children

FEFAM Amount of fees paid **FEHIRE** Fee given to get case

FEFAM Better for man to work, woman tend home

FEHIRE Should hire and promote women **FEPRESCH** Preschool kids suffer if mother works

RINCOM16 Respondents income

Items measured on other (0 - 10) scales declared as (Nominal) but should be (Ordinal) (Important - Not important; Far below - Far above etc.)

HELPOTH To help others (Rank 5 items)

HISPANIC Hispanic specified

HOMOSEX Homosexual sex relations

INCOME Total family income

MARBLK Close relative marry black

OBEY To obey

PARTNERS How many sex partner's R had in last year

THNKSELF To think for ones self

WORKHARD To work hard

Declared as (Nominal) but should be (Scale)

CHLDIDEL Ideal number of children [7=7 or more, 8 = As many as like]

EDUC Highest year of school completed

Declared as (Nominal) but could be anything.

PARTNERS How many sex partner's R had in last year **PARTNRS5** How many sex partner's R had in last 5 years

Correctly declared as (Scale).

AGE Age of respondent [97 = 97 or older]

MAEDUC Highest year school completed, mother

Highest year school completed, father

SIBS Number of brothers and sisters
TVHOURS Hours per day watching TV

Declared as (Scale) but should be (Nominal) 3

HISPANIC Hispanic specified

Declared as (Scale) but should be (Ordinal) 4

RINCOM16 Respondents income

This is due to the level, during automatic archiving, assigned by SPSS to variables with many values. Using the VARIABLE LEVEL command in syntax assigns correct levels.

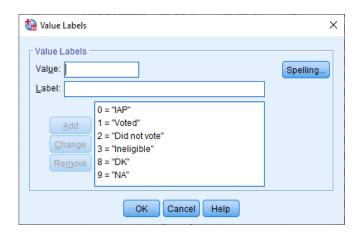
Names and Labels

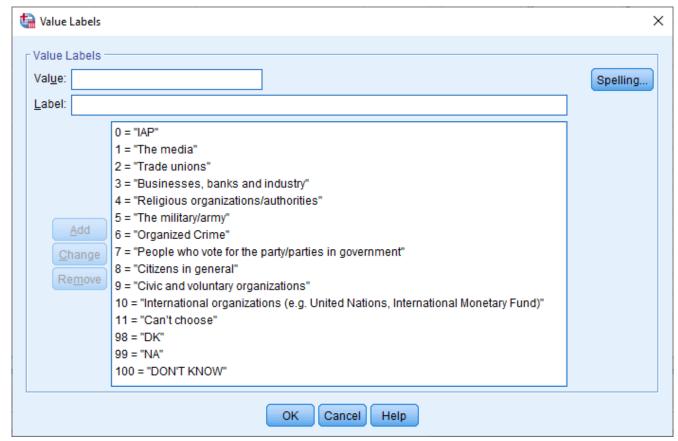
All **variable names** are in **UPPER CASE**.and restricted to 8-characters. This is a left-over from the early days of SPSS. Nowadays they can be **Mixed Case** and have up to 256 characters. However, 8-character names are much quicker to type if you are using syntax rather than the GUI! Names are case-insensitive in SPSS so you can type them in **lower case**.

All **variable labels** are thankfully now in **Mixed Case** which is much easier on the eye and better looking in output intended for publication.

Most value labels are in Mixed Case

VOTE12 Did R vote in 2012 election

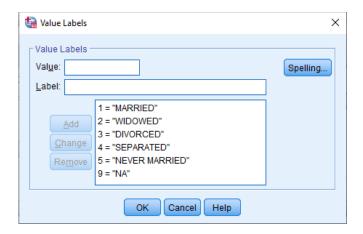




Off-scale responses such as IAP, DK, NA are still in UPPER CASE.

Other value labels are in **UPPER CASE** because these variables have been used over many waves and the (automated?) file creation software has simply been copied and re-run.

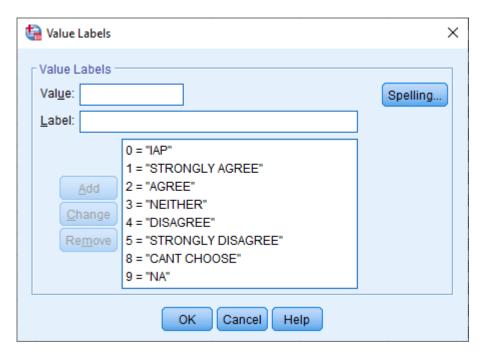
MARITAL Marital status



HAPPY General happiness

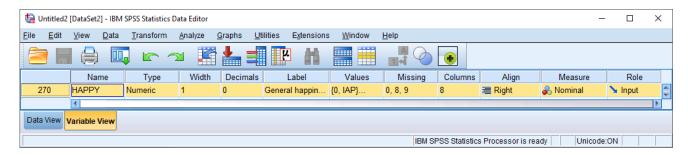


GOODLIFE Standard of living of R will improve



Missing Values

Missing values are declared for most variables as 0,8 and 9:



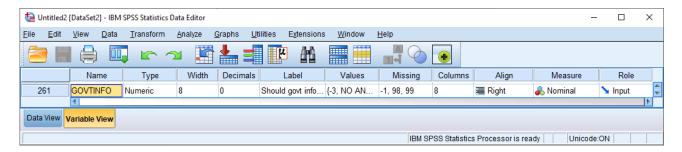
. .and are appropriately labelled:

HAPPY General happiness

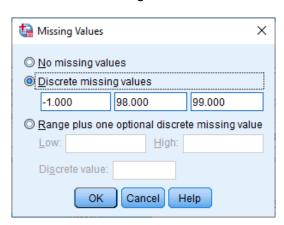


The missing values for some variables have negative values (and superfluous decimals):

GOVTINFO "Should govt info be public" -3, -2, -1



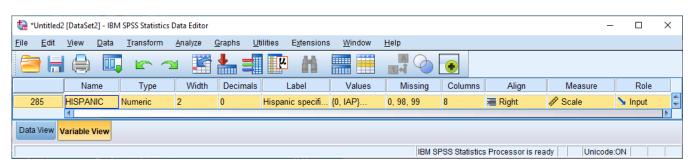
. . but two of the negative values, -3 and -2, are not declared as missing.

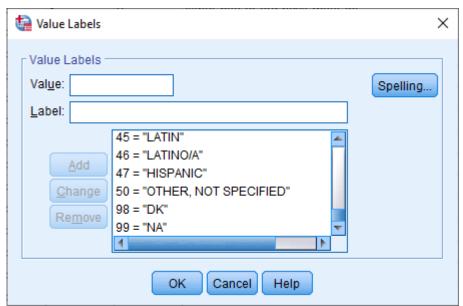




Missing values for variables with many values can be 98, 99, 998, 999 etc.

HISPANIC Hispanic specified 0, 98. 99





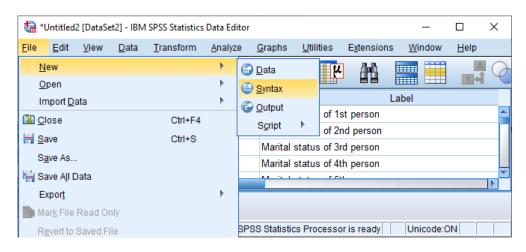
In the 2008 file **GSS2008.SAV** all variable and value labels files were in **UPPER CASE**, which I find unaesthetic, so I changed them all to **Mixed Case** using Python code supplied by Jon Peck. ⁵ I may still need to do this for some of the variables in **GSS2016.SAV**.

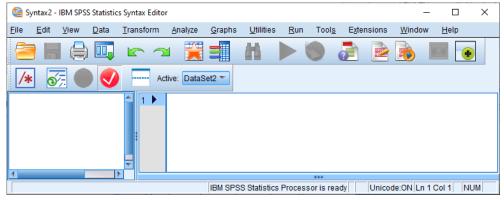
⁵ Retired Senior Software Engineer, IBM/SPSS: contributing author SPSS for Windows

Selecting the variables used in the book

We don't need all 691 variables in the file, just the ones used in the book:

File >> New >> Syntax





There is no variable to identify the cases, so type in:

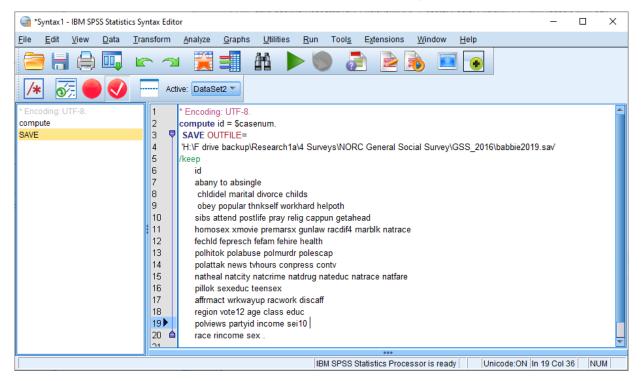
compute id = \$casenum.

Then type in (or copy/paste from here)

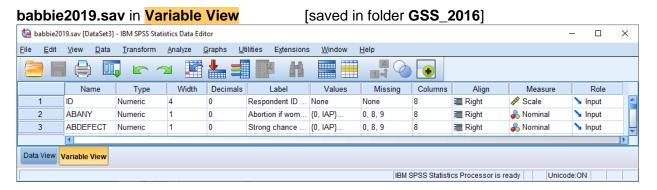
SAVE OUTFILE=

```
'H:\F drive backup\Research1a\4 Surveys\NORC General Social Survey\GSS_2016\babbie2019.sav'
  /keep
  id
  abany to absingle
   chldidel marital divorce childs
   obey popular thnkself workhard helpoth
  sibs attend postlife pray relig cappun getahead
  homosex xmovie premarsx gunlaw racdif4 marblk natrace
  fechld fepresch fefam fehire health
  polhitok polabuse polmurdr polescap
  polattak news tvhours conpress contv
  natheal natcity natcrime natdrug nateduc natrace natfare
  pillok sexeduc teensex
  affrmact wrkwayup racwork discaff
  region vote12 age class educ
  polviews partyid income sei10
  race rincome sex.
```

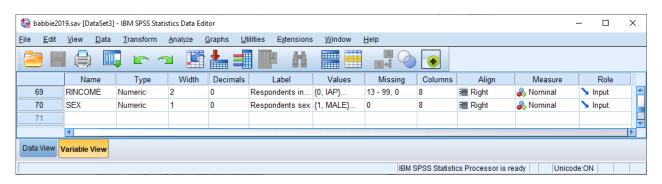
[Commentary on SPSS files for Babbie et al 2019]

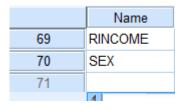


Ctrl + R or green arrow to get



Use the vertical slide bar (on the right edge of the Data Editor) to scroll to the end of the file:





There are 70 variables in the file

Variables in Babbie19.sav

Active file: babbie2019.sav

SPSS syntax: display labels.

Variable Labels

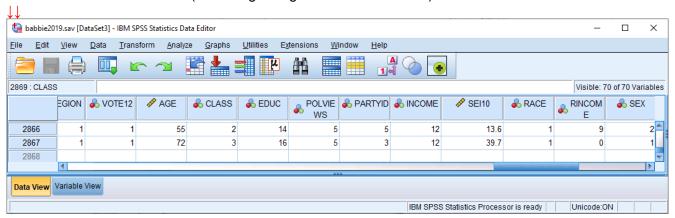
Variable	Position	Variable Labels Label		
ID		Respondent ID number		
ABANY		Abortion if woman wants for any reason		
ABDEFECT		Strong chance of serious defect		
ABHLTH		Woman's health seriously endangered		
ABNOMORE		Marriedwants no more children		
ABPOOR		Low incomecant afford more children		
ABRAPE		Pregnant as result of rape		
ABSINGLE		Not married		
CHLDIDEL		Ideal number of children		
MARITAL		Marital status		
DIVORCE		Ever been divorced or separated		
CHILDS		Number of children		
OBEY		To obey		
POPULAR		To be well liked or popular		
THNKSELF		To think for ones self		
WORKHARD		To work hard		
HELPOTH		To help others		
SIBS		Number of brothers and sisters		
ATTEND		How often R attends religious services		
POSTLIFE		Belief in life after death		
PRAY	21	How often Does R pray		
RELIG		R's religious preference		
CAPPUN		Favor or oppose death penalty for murder		
GETAHEAD		Opinion of how people get ahead		
HOMOSEX		Homosexual sex relations		
XMOVIE		Seen x-rated movie in last year		
PREMARSX	27	Sex before marriage		
GUNLAW	28	Favor or oppose gun permits		
RACDIF4		Differences due to lack of will		
MARBLK		Close relative marry black		
NATRACE	31			
FECHLD		Mother working doesn't hurt children		
FEPRESCH		Preschool kids suffer if mother works		
FEFAM		Better for man to work, woman tend home		
FEHIRE		Should hire and promote women		
HEALTH		Condition of health		
POLHITOK	37			
POLABUSE	38	Citizen said vulgar or obscene things		
POLMURDR		Citizen questioned as murder suspect		
POLESCAP		Citizen attempting to escape custody		
POLATTAK	41			
NEWS		How often Does R read newspaper		
TVHOURS		Hours per day watching TV		
CONPRESS		Confidence in press		
CONTV		Confidence in television		
NATHEAL		Improving & protecting nations health		
NATCITY	47			
NATCRIME		Halting rising crime rate		
NATDRUG		Dealing with drug addiction		
		U 1 19 11 11 11 11 11 11 11 11 11 11 11 1		

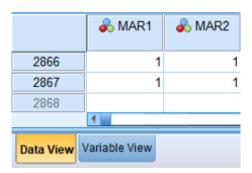
NATEDUC	50	Improving nations education system			
NATFARE	51	Welfare			
PILLOK	52	Birth control to teenagers 14-16			
SEXEDUC	53	Sex education in public schools			
TEENSEX	54	Sex before marriage:teens 14-16			
AFFRMACT	55	Favor preference in hiring blacks			
WRKWAYUP	56	Blacks overcome prejudice without favors			
RACWORK	57	Racial makeup of workplace			
DISCAFF	58	Whites hurt by aff. action			
REGION	59	Region of interview			
VOTE12	60	Did R vote in 2012 election			
AGE	61	Age of respondent			
CLASS	62	Subjective class identification			
EDUC	63	Highest year of school completed			
POLVIEWS	64	Think of self as liberal or conservative			
PARTYID	65	Political party affiliation			
INCOME	66	Total family income			
SEI10	67	R's socioeconomic index (2010)			
RACE	68	Race of respondent			
RINCOME	69	Respondents income			
SEX	70	Respondents sex			

Variables in the working file

Switch to **Data View**

Use the vertical slide bar ⁶ (on the right edge of the **Data Editor**) to scroll to the end of the file:





There are 2867 cases in the file.

Chapter 4 of the book has useful demonstrations (pp31ff) on scroll bars and other ways of navigating the Data Editor when it is displayed on-screen. SPSS

[Commentary on SPSS files for Babbie et al 2019]

Derived variables

The authors generate their derived variables using the GUI, but it's both quicker and easier in syntax:

[yet to be written]

Creating files for Gradpack users

The Gradpack version of SPSS is restricted to 50 variables and 500 cases.

Because files in the book have variables in **alphabetical** order, and to reduce the number of cases, I generated two files **DEMOGRAD.SAV** and **EXERGRAD.SAV**, each with the **same 500 cases** and with the variables in **alphabetic** order. ⁷

Active file: **GSS2016.SAV** SPSS syntax: **frequencies** sex.

SEX Respondents sex

					Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1 MALE	1276	44.5	44.5	44.5
	2 FEMALE	1591	55.5	55.5	100.0
	Total	2867	100.0	100.0	

Active file: **DEMOGRAD.SAV** SPSS syntax: **frequencies** sex.

SEX Respondents sex

					Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1 MALE	224	44.8	44.8	44.8
	2 FEMALE	276	55.2	55.2	100.0
	Total	500	100.0	100.0	

Active file: **EXERGRAD.SAV** SPSS syntax: **frequencies** sex.

SEX Respondents sex

					Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1 MALE	224	44.8	44.8	44.8
	2 FEMALE	276	55.2	55.2	100.0
	Total	500	100.0	100.0	

oo <u>ripporidix</u> re

⁷ See Appendix for the syntax

Variables in DEMOGRAD.SAV

Variable Labels

Variable Labels					
Variable	Position	Label			
ID	1	Respondent ID number			
ABANY	2	Abortion if woman wants for any reason			
ABDEFECT	3	Strong chance of serious defect			
ABHLTH	4	Woman's health seriously endangered			
ABNOMORE	5	Marriedwants no more children			
ABPOOR	6	Low incomecant afford more children			
ABRAPE	7	Pregnant as result of rape			
ABSINGLE	8	Not married			
AGE	9	Age of respondent			
ATTEND	10	How often R attends religious services			
CAPPUN	11	Favor or oppose death penalty for murder			
CHLDIDEL	12	Ideal number of children			
CLASS	13	Subjective class identification			
EDUC	14	Highest year of school completed			
EVSTRAY	15				
GETAHEAD	16	Opinion of how people get ahead			
GUNLAW	17	Favor or oppose gun permits			
HELPOTH	18	To help others			
HISPANIC	19	Hispanic specified			
HOMOSEX	20	Homosexual sex relations			
INCOME	21	Total family income			
MAEDUC	22				
MARBLK	23	Close relative marry black			
MARITAL	24				
OBEY	25	To obey			
PAEDUC	26	Highest year school completed, father			
PARTNERS	27	-			
PARTYID	28				
PIKUPSEX	29				
POLVIEWS		Think of self as liberal or conservative			
POPULAR	31				
POSTLIFE		Belief in life after death			
PRAY		How often Does R pray			
PREMARSX		Sex before marriage			
RACDIF4		Differences due to lack of will			
RACE	36				
RELIG	37				
RINCOM16	38	<u> </u>			
SEX	39	•			
SEXFREQ	40	·			
SIBS	41				
THNKSELF		To think for ones self			
WORKHARD		To work hard			
XMOVIE		Seen x-rated movie in last year			
AIVIO VIL	77	South A fatou filovio ili last your			

Variables in the working file

Variables in EXERGRAD.SAV

Variable Labels

Variable Labels					
Variable	Position	Label			
AFFRMACT	1	Favor preference in hiring blacks			
AGE	2	Age of respondent			
CLASS	3	Subjective class identification			
COLATH	4	Allow anti-religionist to teach			
COLCOM	5				
COLHOMO	6	Allow homosexual to teach			
COLRAC	7	Allow racist to teach			
CONPRESS	8	Confidence in press			
CONTV	9	Confidence in television			
DISCAFF	10	Whites hurt by aff. action			
EDUC	11	Highest year of school completed			
EVSTRAY	12				
FECHLD	13				
FEFAM	14				
FEHIRE	15				
FEPRESCH	16	·			
GRASS	17	Should marijuana be made legal			
HEALTH	18				
HISPANIC	19				
INCOME	20	Total family income			
MAEDUC	21	Highest year school completed, mother			
MARITAL					
	22				
NATORING	23	5 5			
NATERUG	24	0			
NATEDUC	25	Improving nations education system			
NATFARE	26				
NATHEAL	27	Improving & protecting nations health			
NATRACE	28	Improving the conditions of blacks			
NEWS	29				
PAEDUC	30	Highest year school completed, father			
PARTNERS	31	How many sex partner's R had in last year			
PILLOK	32	Birth control to teenagers 14-16			
POLABUSE	33	Citizen said vulgar or obscene things			
POLATTAK	34	Citizen attacking policeman with fists			
POLESCAP	35				
POLHITOK	36	Ever approve of police striking citizen			
POLMURDR	37	Citizen questioned as murder suspect			
PORNLAW	38	Feelings about pornography laws			
RACE	39	Race of respondent			
RACWORK	40	Racial makeup of workplace			
RINCOM16	41	Respondents income			
SEX	42	Respondents sex			
SEXEDUC	43	Sex education in public schools			
SEXFREQ	44	Frequency of sex during last year			
TEENSEX	45	Sex before marriage:teens 14-16			
TVHOURS	46	Hours per day watching TV			

Variables in the working file

Variables common to both DEMO.SAV and EXER.SAV.

AGE Age of respondent

CLASS Subjective class identification
EDUC Highest year of school completed

EVSTRAY Have sex other than spouse while married

HEALTH Condition of health
HISPANIC Hispanic specified
Total family income

MAEDUC Highest year school completed, mother

MARITAL Marital status

PAEDUC Highest year school completed, father **PARTNERS** How many sex partner's R had in last year

RACE Race of respondent
RINCOM16 Respondents income
Respondents sex

SEXFREQ Frequency of sex during last year

Variables in babbie20119.say not in either demo.say or exer.say.

Declared as (Scale) but should be (Nominal)

COIND10 Partner's industry code (2010)

COISCO08 Partner's occupation, 2010 census & 2008 ISCO code

HISPANIC Hispanic specified

OCC10 R's census occupation code (2010)
OTH16 Other protestant denominations
OTHER Other protestant denominations
OTHLANG1 What other languages Does R speak
OTHLANG2 What other languages Does R speak
PAIND10 Father's industry code (NAICS 2007)

PAISCO08 R's Father's occupation, 2010 census & 2008 ISCO code PAISCO88 R's Father's occupation, 1980 census & 1988 ISCO code

PAOCC10 Father's census occupation code (2010)
PAPRES10 Father's occupational prestige score (2010)

PAPRES105PLUS Father's occupational prestige score using threshold method (2010)

Correctly declared as (Nominal)

DIVORCE Ever been divorced or separated

Declared as (Nominal) but should be (Scale)

CHILDS Number of children [8 = 8 or more]

OVERSAMP Weights for black oversamples

AGE3 Age of person #3

AGEKDBRN R's age when 1st child born

COHORT Year of birth

COHRS1 Number of hours partner worked last week
No. of hours partner usually works a week
Weight deal with experimental randomization

INTAGE Age of interviewer

OLD1

to **OLD9** Ages of people in h/h

OVERSAMP Weights for black oversamples

Declared as (Nominal) but should be (Ordinal)

NEWS How often Does R read newspaper

PADEG Father's highest degree

ADVFRONT Sci Rsch is necessary and should be supported by federal govt

CHNGEWRK I would change my work for something different FeJOBAFF For or against preferential hiring of women

FINALTER Change in financial situation
GOODLIFE Standard of living of R will improve
WRKWAYUP Blacks overcome prejudice without favors

FINRELA Opinion of family income

FLEXHRS1 Importance of flexible hrs. at work
HAPCOHAB Happiness of relt with partner
HAPMAR Happiness of marriage

Happy General happiness
HEALTH Condition of health
HELPBLK Should govt aid blacks?
HELPNOT Should govt do more or less?
HELPORG1 Willing to work harder for the firm
HELPPOOR Should govt improve standard of living?

HELPFOOR
HELPSICK
HLPHITEC
HLPOTHS
HLPSOC
HLTHCARE
Should govt improve standard of living?
Should govt help pay for medical care?
Govt support for new products and tech
Importance of helping others in a job
Importance of social usefulness in a job
Govts resp: provide hith care for sick

HOPE1 I could get out of a jam

HOPE2 I am energetically pursuing my goals

HOPE3 There are many ways around problems i face

HOPE4 I see myself as successful

HOPE5 I can think of many ways to reach my current goals

HOPE6 I am meeting my current goals

NOUNEMP1
 NOUNEMP2
 NOUNEMP3
 NOUNEMP4
 R would accept position with lower pay to avoid unemployment
 R would accept temporary employment to avoid unemployment
 R would accept travel longer to work to avoid unemployment

NOUNEMP5 I would move within the usa for work
NOUNEMP6 I would move outside the usa for work

NUKEGEN Nuke power danger to envir

Syntax to extract variables for DEMO.SAV and EXER.SAV

* Encoding: UTF-8.

compute id = \$casenum.

* gss2016.sav.

*Get same sample for both exercises.

```
set seed 987654.
sample 1500 from 2867.
*demo.sav.
save out 'C:\Users\JohnPC\Desktop\GSS 2016\demo.sav'
       /keep id
       abany to absingle age attend
       cappun chldidel class
       educ evstray
       getahead gunlaw helpoth
       hispanic homosex
       income
       maeduc marblk marital
       obev
       paeduc partners partyid
       pikupsex polviews popular postlife pray premarsx
       racdif4 race relig rincom16
       sex sexfreq sibs
       thnkself
       workhard
       xmovie.
*exer.sav.
save out 'c:/users/johnpc/desktop/gss_2016/exer2.sav'
       affrmact age class colath colcom colhomo colrac conpress contv
       discaff educ evstray fechld fefam fehire fepresch
       grass
       health
       hispanic
       income
       maeduc marital
       natcrime natdrug nateduc natfare natheal natrace news
       paeduc partners pillok polabuse polattak polescap polhitok polmurdr pornlaw
       race racwork rincom16
       sex sexeduc sexfreq teensex tvhours.
```