McMaster University

Family Medicine

David Braley Primary Care Research Collaborative

@McMasterFamMed

Using Statistical Package for the Social Sciences (SPSS): Part I

Watch a recording of the session here

Melissa Pirrie
Larkin Lamarche

January 26th, 2021

Michael G. DeGroote SCHOOL OF MEDICINE

McMaster University

Family Medicine

Using Statistical Package for the Social Sciences (SPSS): Part I

Melissa Pirrie

Larkin Lamarche

@McMasterFamMed

Michael G. DeGroote SCHOOL OF MEDICINE

Caveats

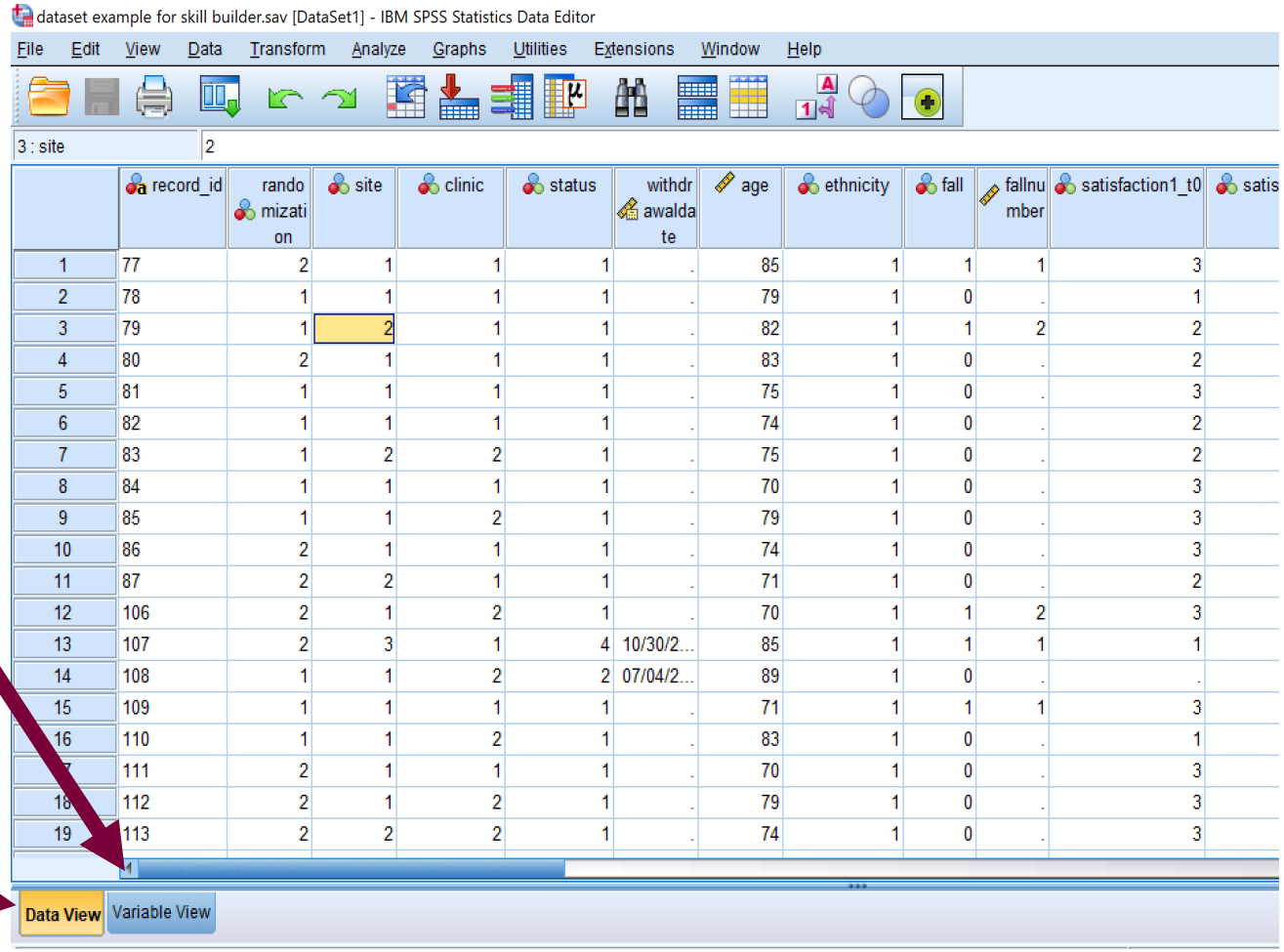
- SPSS is only one software to store and analyze data
- There was a skill builder done in 2017
  - X:\Training and Resources\Knowledge and Skills Builder Sessions\Software\SPSS June 2017

Learning objectives

1. Name the two ‘views’ of SPSS and describe their function
2. Identify 3 ways to get data into SPSS
3. Identify the parts of database structure in SPSS
4. Understand how to check the integrity of your database structure
5. Understand some simple ways to check the quality of your data
6. *Apply these methods of checking to data output*

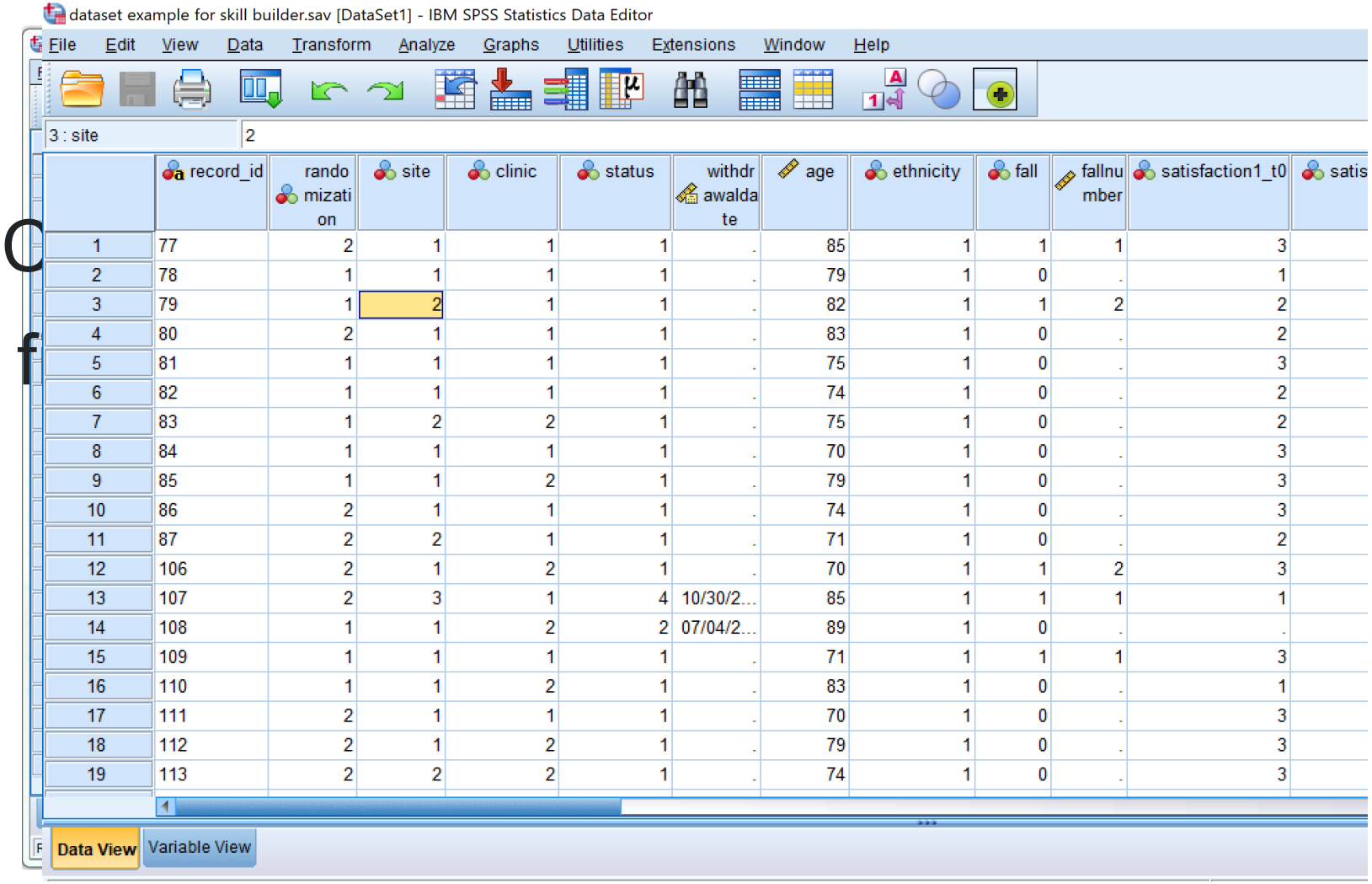
Two ‘views’

1. Variable view
   - Is done first
   - Sets up the database structure to tell data how to behave
2. Data view
   - Houses the data
   - ‘how’ it behaves is directed by variable view



Getting data into SPSS

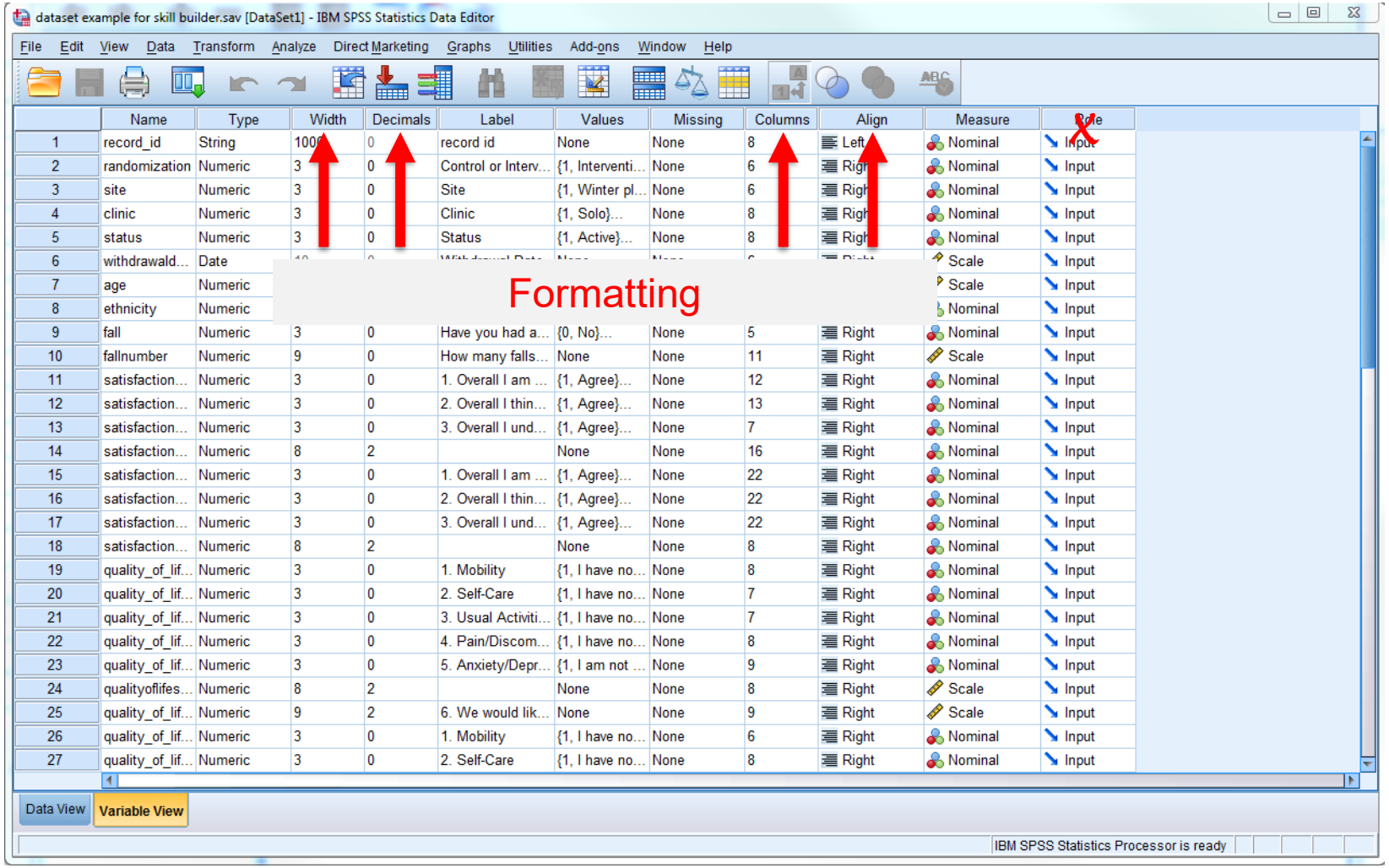
1. Direct entry
2. Import from REDC
3. Import from CSV f



	record_id	randomization	site	clinic	status	withdrawaldate	age	ethnicity	fall	fallnumber	satisfaction1_t0	satis
1	77	2	1	1	1	.	85	1	1	1	3	
2	78	1	1	1	1	.	79	1	0	.	1	
3	79	1	2	1	1	.	82	1	1	2	2	
4	80	2	1	1	1	.	83	1	0	.	2	
5	81	1	1	1	1	.	75	1	0	.	3	
6	82	1	1	1	1	.	74	1	0	.	2	
7	83	1	2	2	1	.	75	1	0	.	2	
8	84	1	1	1	1	.	70	1	0	.	3	
9	85	1	1	2	1	.	79	1	0	.	3	
10	86	2	1	1	1	.	74	1	0	.	3	
11	87	2	2	1	1	.	71	1	0	.	2	
12	106	2	1	2	1	.	70	1	1	2	3	
13	107	2	3	1	4	10/30/2...	85	1	1	1	1	
14	108	1	1	2	2	07/04/2...	89	1	0	.	.	
15	109	1	1	1	1	.	71	1	1	1	3	
16	110	1	1	2	1	.	83	1	0	.	1	
17	111	2	1	1	1	.	70	1	0	.	3	
18	112	2	1	2	1	.	79	1	0	.	3	
19	113	2	2	2	1	.	74	1	0	.	3	

Database structure

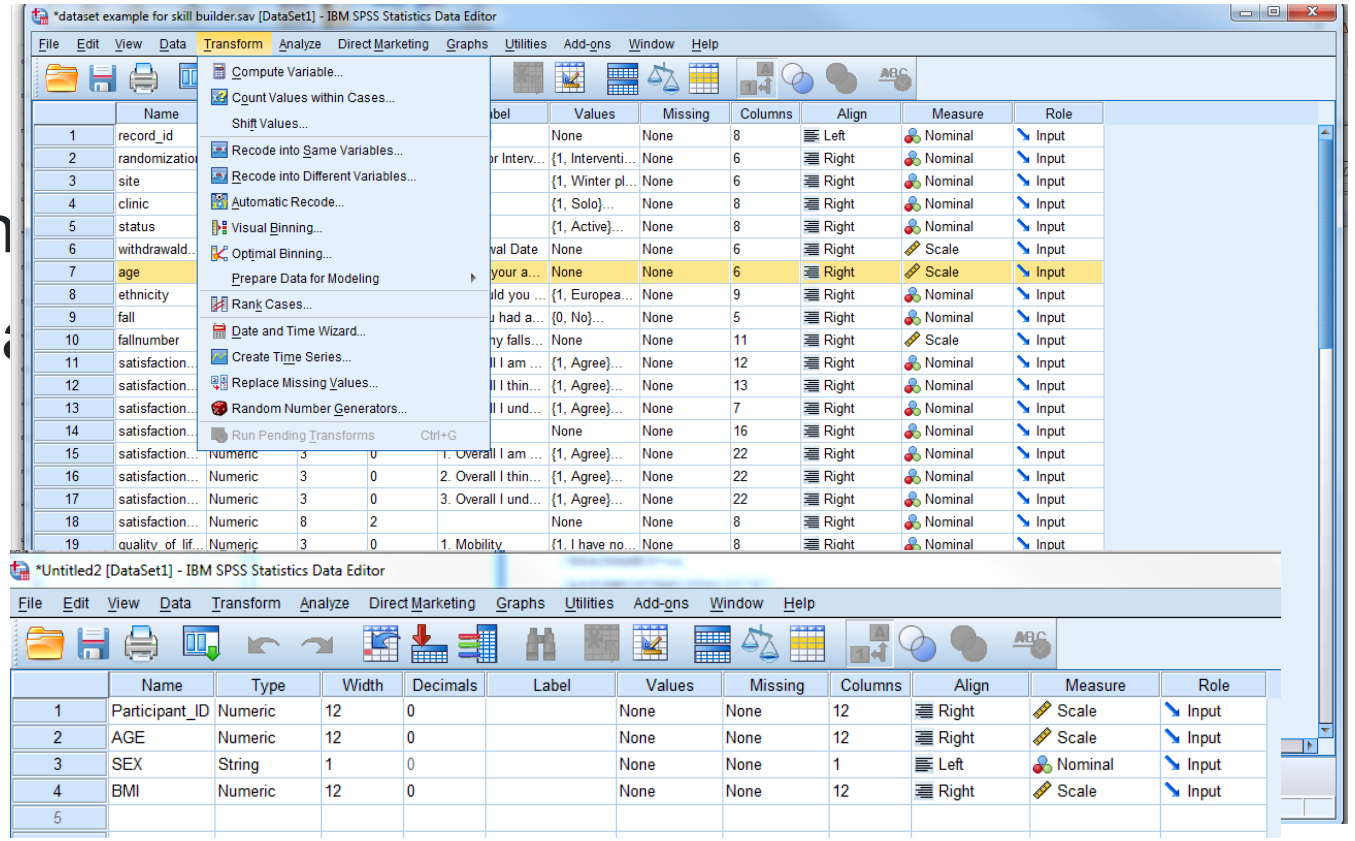
- In ‘variable view’
- Parts of the database
  - Name
  - Type
  - Width
  - Decimals
  - Label
  - Values
  - Missing
  - Columns
  - Align
  - Measure
  - Role



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	record_id	String	100	0	record id	None	None	8	Left	Nominal	Input
2	randomization	Numeric	3	0	Control or Interv...	{1, Interventi...	None	6	Right	Nominal	Input
3	site	Numeric	3	0	Site	{1, Winter pl...	None	6	Right	Nominal	Input
4	clinic	Numeric	3	0	Clinic	{1, Solo}...	None	8	Right	Nominal	Input
5	status	Numeric	3	0	Status	{1, Active}...	None	8	Right	Nominal	Input
6	withdrawald...	Date								Scale	Input
7	age	Numeric								Scale	Input
8	ethnicity	Numeric								Nominal	Input
9	fall	Numeric	3	0	Have you had a...	{0, No}...	None	5	Right	Nominal	Input
10	fallnumber	Numeric	9	0	How many falls...	None	None	11	Right	Scale	Input
11	satisfaction...	Numeric	3	0	1. Overall I am ...	{1, Agree}...	None	12	Right	Nominal	Input
12	satisfaction...	Numeric	3	0	2. Overall I thin...	{1, Agree}...	None	13	Right	Nominal	Input
13	satisfaction...	Numeric	3	0	3. Overall I und...	{1, Agree}...	None	7	Right	Nominal	Input
14	satisfaction...	Numeric	8	2		None	None	16	Right	Nominal	Input
15	satisfaction...	Numeric	3	0	1. Overall I am ...	{1, Agree}...	None	22	Right	Nominal	Input
16	satisfaction...	Numeric	3	0	2. Overall I thin...	{1, Agree}...	None	22	Right	Nominal	Input
17	satisfaction...	Numeric	3	0	3. Overall I und...	{1, Agree}...	None	22	Right	Nominal	Input
18	satisfaction...	Numeric	8	2		None	None	8	Right	Nominal	Input
19	quality_of_lif...	Numeric	3	0	1. Mobility	{1, I have no...	None	8	Right	Nominal	Input
20	quality_of_lif...	Numeric	3	0	2. Self-Care	{1, I have no...	None	7	Right	Nominal	Input
21	quality_of_lif...	Numeric	3	0	3. Usual Activiti...	{1, I have no...	None	7	Right	Nominal	Input
22	quality_of_lif...	Numeric	3	0	4. Pain/Discom...	{1, I have no...	None	8	Right	Nominal	Input
23	quality_of_lif...	Numeric	3	0	5. Anxiety/Depr...	{1, I am not ...	None	9	Right	Nominal	Input
24	qualityoflifes...	Numeric	8	2		None	None	8	Right	Scale	Input
25	quality_of_lif...	Numeric	9	2	6. We would lik...	None	None	9	Right	Scale	Input
26	quality_of_lif...	Numeric	3	0	1. Mobility	{1, I have no...	None	6	Right	Nominal	Input
27	quality_of_lif...	Numeric	3	0	2. Self-Care	{1, I have no...	None	8	Right	Nominal	Input

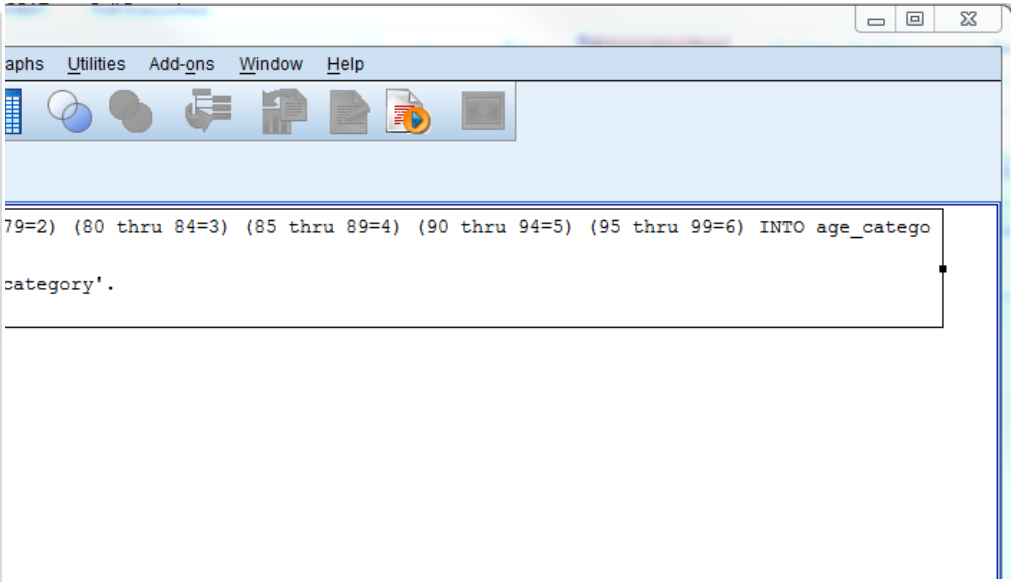
Database re-structuring

- Correct any errors
  - variable types (numeric/strin
  - measure type (scale, nomina
- Enter missing information
  - Variable and value labels
- Recode variables
  - Imported as string
  - Require manipulation to be interpreted



Time Saving Tip!

- Syntax shows up in the output window
- File -> New -> Syntax



Checking data quality: Visual checks

1. Visual checks
   - In ‘data view’
   - Literally scrolling through your data to check for:
     - Missing data patterns (random versus systematic)
     - Logic of data
     - Odd values
     - Sample size/group sizes
       - » Data completeness

Checking data quality: ‘Run’ checks

1. ‘Run’ checks
   - Analyze -> Descriptive or Frequencies
     - Are data within the possible range of scores?
     - What does the variability look like?
   - Cronbach’s alpha for survey
     - Does the survey show good internal consistency? (α ≥ .70)
       - Is it too high? (α ≥.90)
     - Analyze -> Scale -> Reliability analysis
2. Check assumptions for analyses

Common oopsies and fixies

Oopsy	Fixy (most likely)
• Frequency or other basic descriptives won't run	• Variable type may have set itself to "string" from your imported data
	• Variable "measure" may be incorrectly set to nominal (if scale) or scale (if nominal)
• Both mean and range are much higher than expected	• Your data has "99" or "999" for missing values and you haven't shared this info with SPSS
• Mean seems okay but range is not	• Data entry error or issue with recoding
• You have output that no one else can interpret	• Add value labels!

The data structure has to be set before entering data – you can do simple checks to ensure the integrity of the structure and the data before analysis

Learning objectives

1. Name the two ‘views’ of SPSS and describe their function
2. Identify 3 ways to get data into SPSS
3. Identify the parts of database structure in SPSS
4. Understand how to check the integrity of your database structure
5. Understand some simple ways to check the quality of your data
6. *Apply these methods of checking to data output*

SPSS Basics Part 2 will be in four weeks…

McMaster University

Family Medicine

Department of Family Medicine
Michael G. DeGroote School of Medicine
Faculty of Health Sciences

fammedmcmaster.ca
@McMasterFamMed