McMaster University

Family Medicine

David Braley Primary Care Research Collaborative

@McMasterFamMed

Just the basics: Learning about the essential steps to do some simple things in SPSS

Watch a recording of the session here

Melissa Pirrie, Larkin Lamarche

May 1st, 2017

Michael G. DeGroote SCHOOL OF MEDICINE

McMaster University
FAMILY MEDICINE

Just the basics: Learning about the essential steps to do some simple things in SPSS

Larkin Lamarche

Michael G. DeGroote
SCHOOL OF MEDICINE

Learning objectives

1. Be able to conduct a few screening tests on a dataset to detect data entry errors or abnormalities, and correct these errors
2. Be able to recode variables and calculate scores
3. Be able to run descriptive statistics for variables and understand the output from such tests
4. Understand the split file function
5. Be able to interrupt output from a *t*-test, ANOVA, and linear regression

The dataset

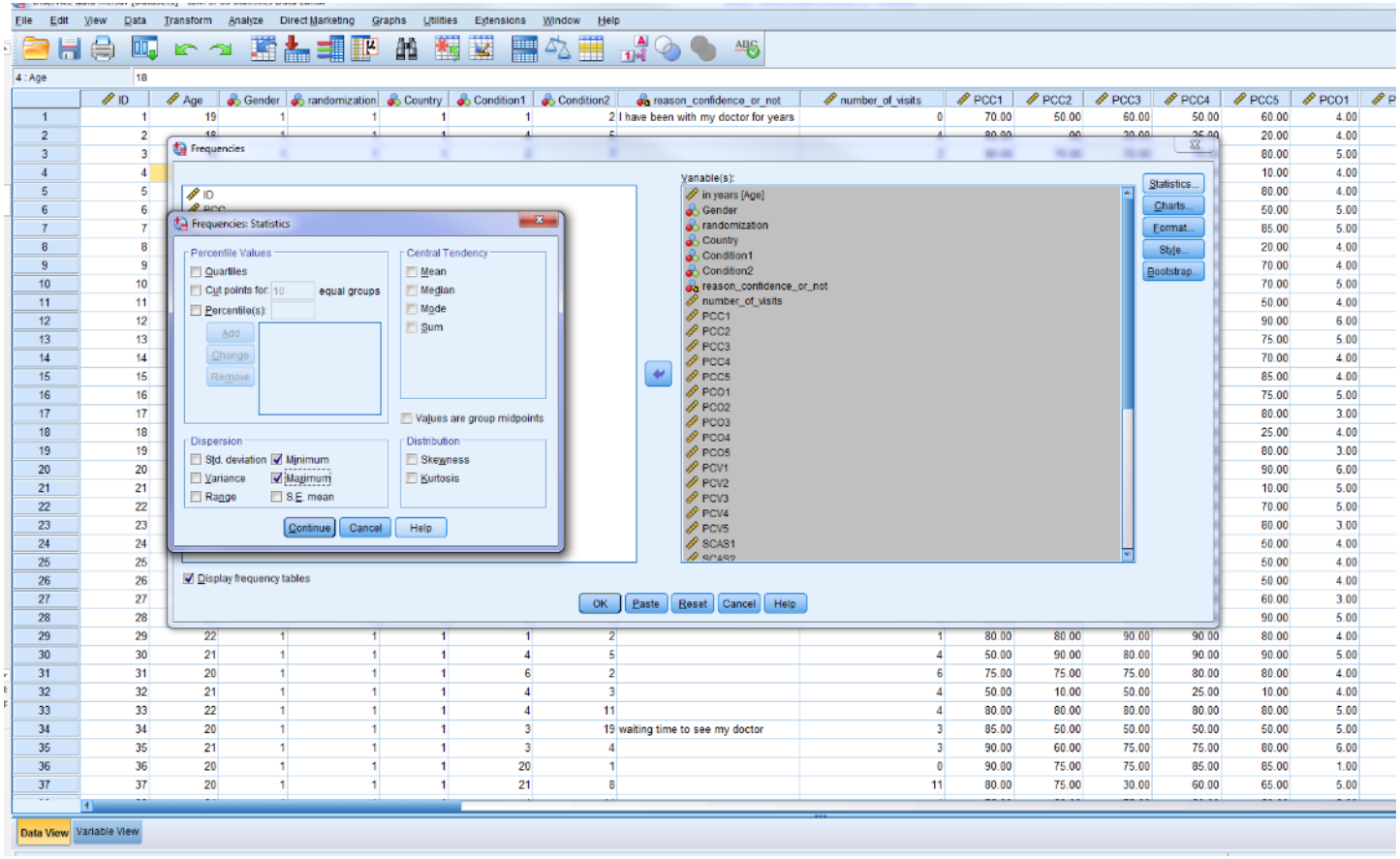
- Study measured primary care confidence and secondary care access
  - Primary care confidence survey has 3 subscales, higher scores=higher confidence
    - Primary care confidence expectancy (5 items, 0-100)
    - Primary care confidence outcome expectancy (5 items, 1-6)
    - Primary care confidence outcome value (5 items, 1-6)
  - Unnecessary secondary care access survey (9 items), higher scores=higher unnecessary access to secondary care
- Age, gender, country of birth and top two conditions for which a primary care provider was seen in the last 2 years, reason for confidence (or not being confident), number of primary care visits in last 2 years
- n=341
- Intervention was delivered to change unnecessary secondary care access

A few screening tests

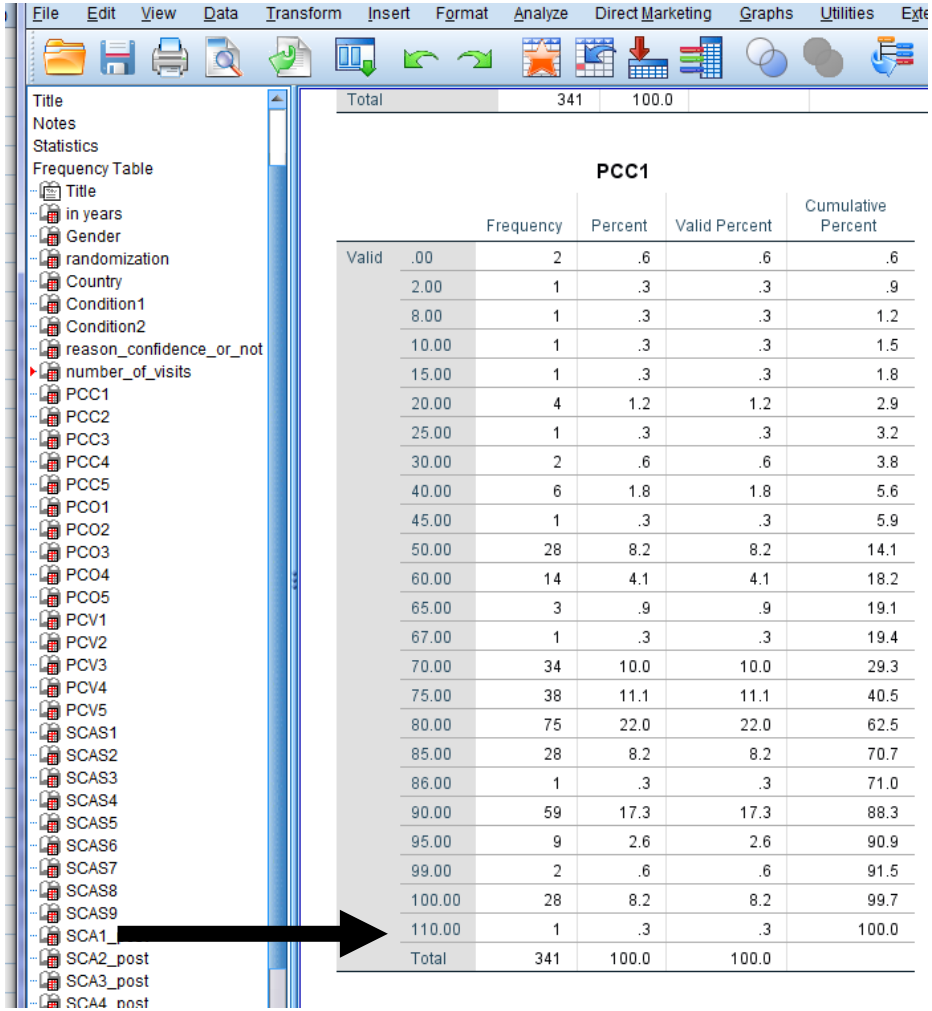
- ✓ Clean up string data
  - Spell check
- ✓ Check possible range of values
- ✓ Check missing data
  - Systematic or random?
  - Visual inspection

The clicks: Range and missing values

Analyze -> Descriptive Statistics -> Frequencies (or Descriptives)
(Statistics… Minimum, Maximum)



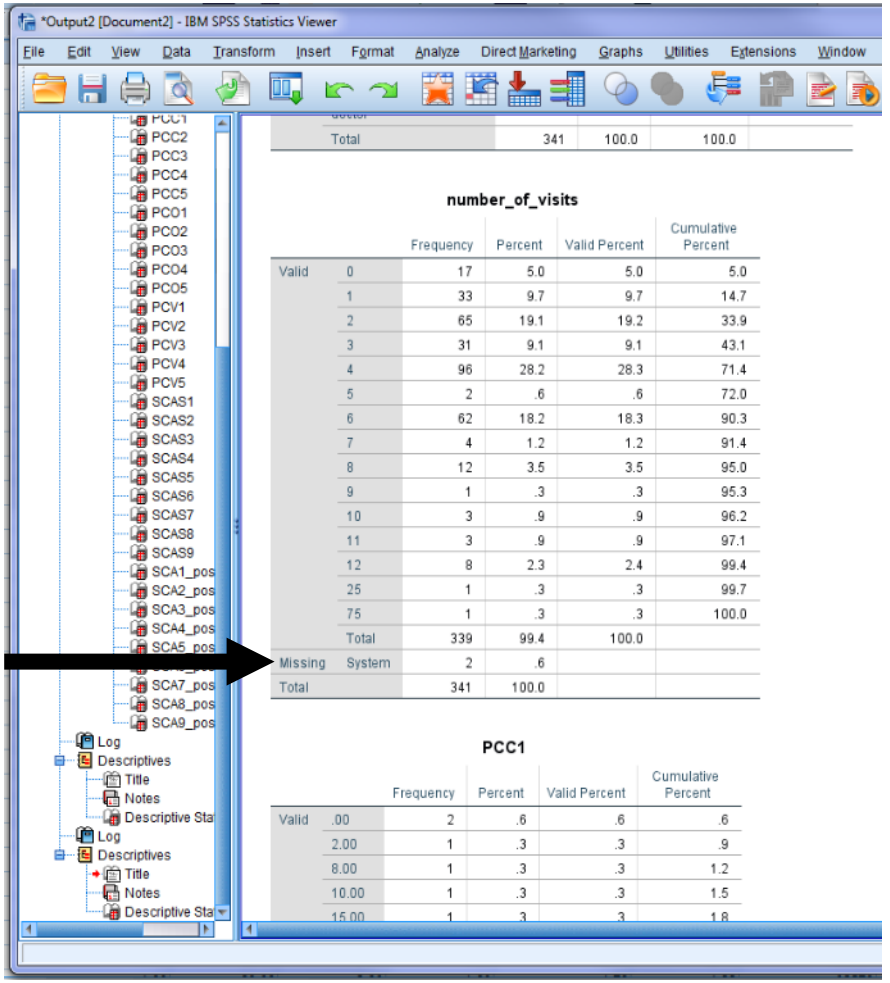
The Output: Checking range and missing values



Total 341 100.0

PCC1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	2	.6	.6	.6
	2.00	1	.3	.3	.9
	8.00	1	.3	.3	1.2
	10.00	1	.3	.3	1.5
	15.00	1	.3	.3	1.8
	20.00	4	1.2	1.2	2.9
	25.00	1	.3	.3	3.2
	30.00	2	.6	.6	3.8
	40.00	6	1.8	1.8	5.6
	45.00	1	.3	.3	5.9
	50.00	28	8.2	8.2	14.1
	60.00	14	4.1	4.1	18.2
	65.00	3	.9	.9	19.1
	67.00	1	.3	.3	19.4
	70.00	34	10.0	10.0	29.3
	75.00	38	11.1	11.1	40.5
	80.00	75	22.0	22.0	62.5
	85.00	28	8.2	8.2	70.7
	86.00	1	.3	.3	71.0
	90.00	59	17.3	17.3	88.3
	95.00	9	2.6	2.6	90.9
	99.00	2	.6	.6	91.5
	100.00	28	8.2	8.2	99.7
	110.00	1	.3	.3	100.0
	Total	341	100.0	100.0	



Total 341 100.0 100.0

number_of_visits

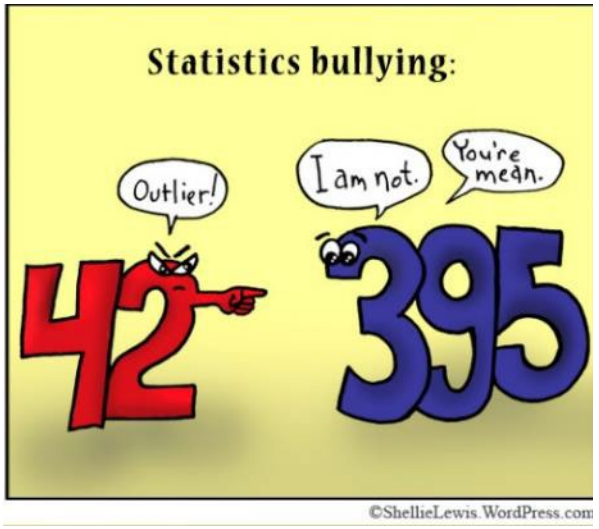
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	17	5.0	5.0	5.0
	1	33	9.7	9.7	14.7
	2	65	19.1	19.2	33.9
	3	31	9.1	9.1	43.1
	4	96	28.2	28.3	71.4
	5	2	.6	.6	72.0
	6	62	18.2	18.3	90.3
	7	4	1.2	1.2	91.4
	8	12	3.5	3.5	95.0
	9	1	.3	.3	95.3
	10	3	.9	.9	96.2
	11	3	.9	.9	97.1
	12	8	2.3	2.4	99.4
	25	1	.3	.3	99.7
	75	1	.3	.3	100.0
	Total	339	99.4	100.0	
Missing	System	2	.6		
Total		341	100.0		

PCC1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	2	.6	.6	.6
	2.00	1	.3	.3	.9
	8.00	1	.3	.3	1.2
	10.00	1	.3	.3	1.5
	15.00	1	.3	.3	1.8

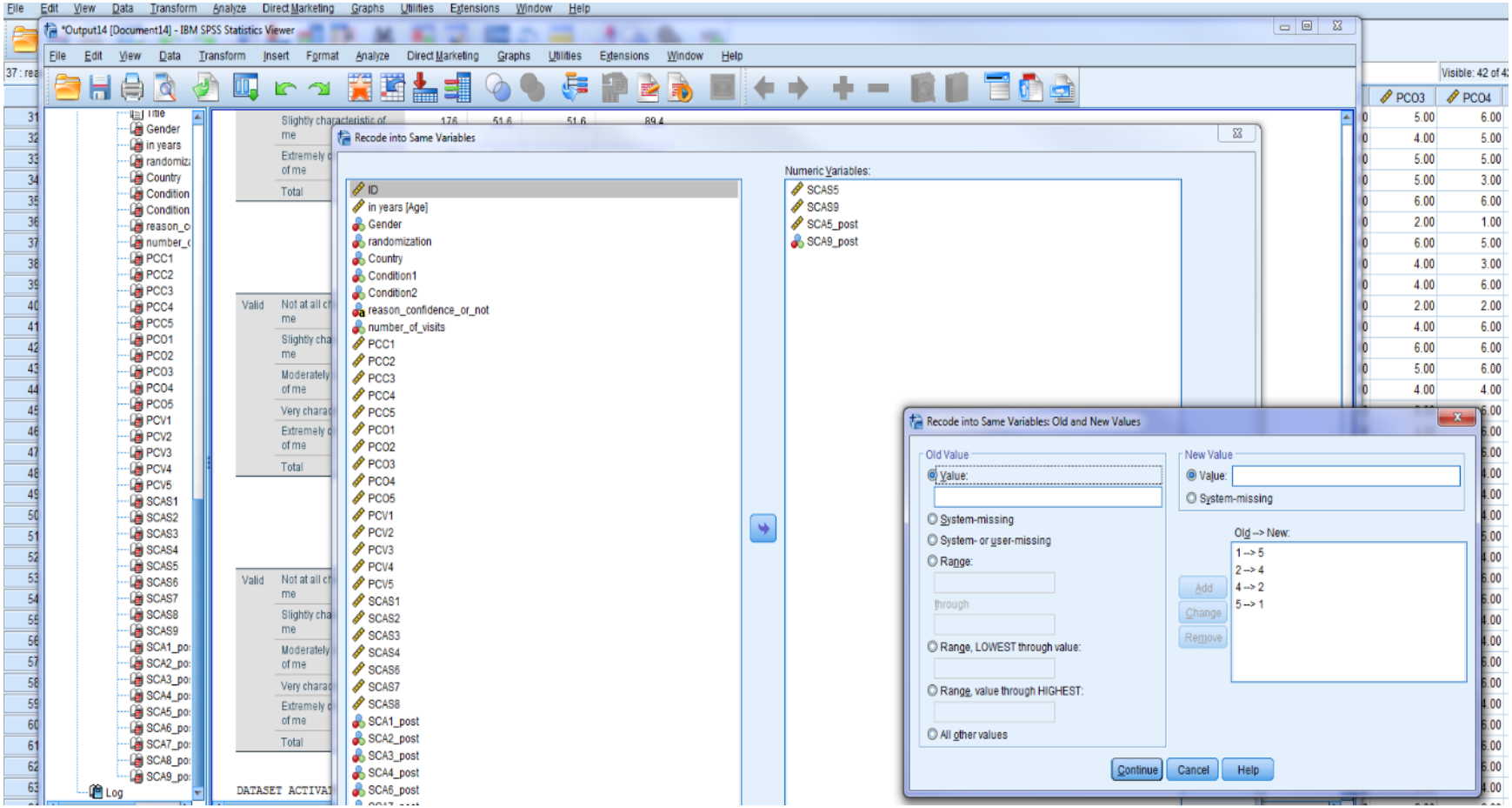
Recode variables and calculate scores

- Recode and calculate scores
  - Reverse worded items?
    - Items 5,9
  - Scores/subscales
  - Check for univariate outliers
    - z-score ±3.29
- Check internal consistency for scales
  - How closely related items are as a group
  - $\alpha \geq .70$ (>.90 too much?)



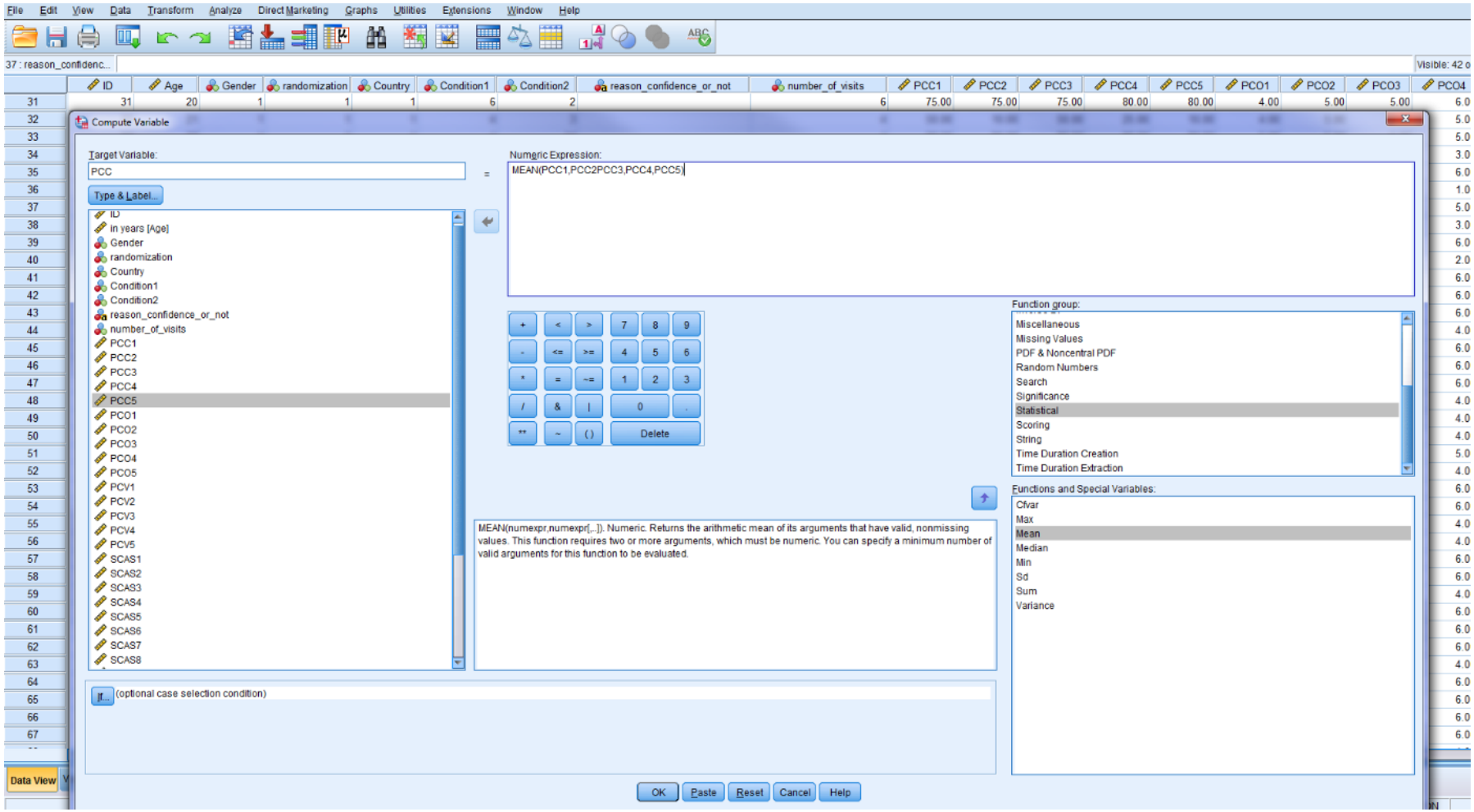
The clicks: Recode variables

Transform -> Recode into Same Variables (Old and New Values…)



The clicks: Calculate scores/subscales

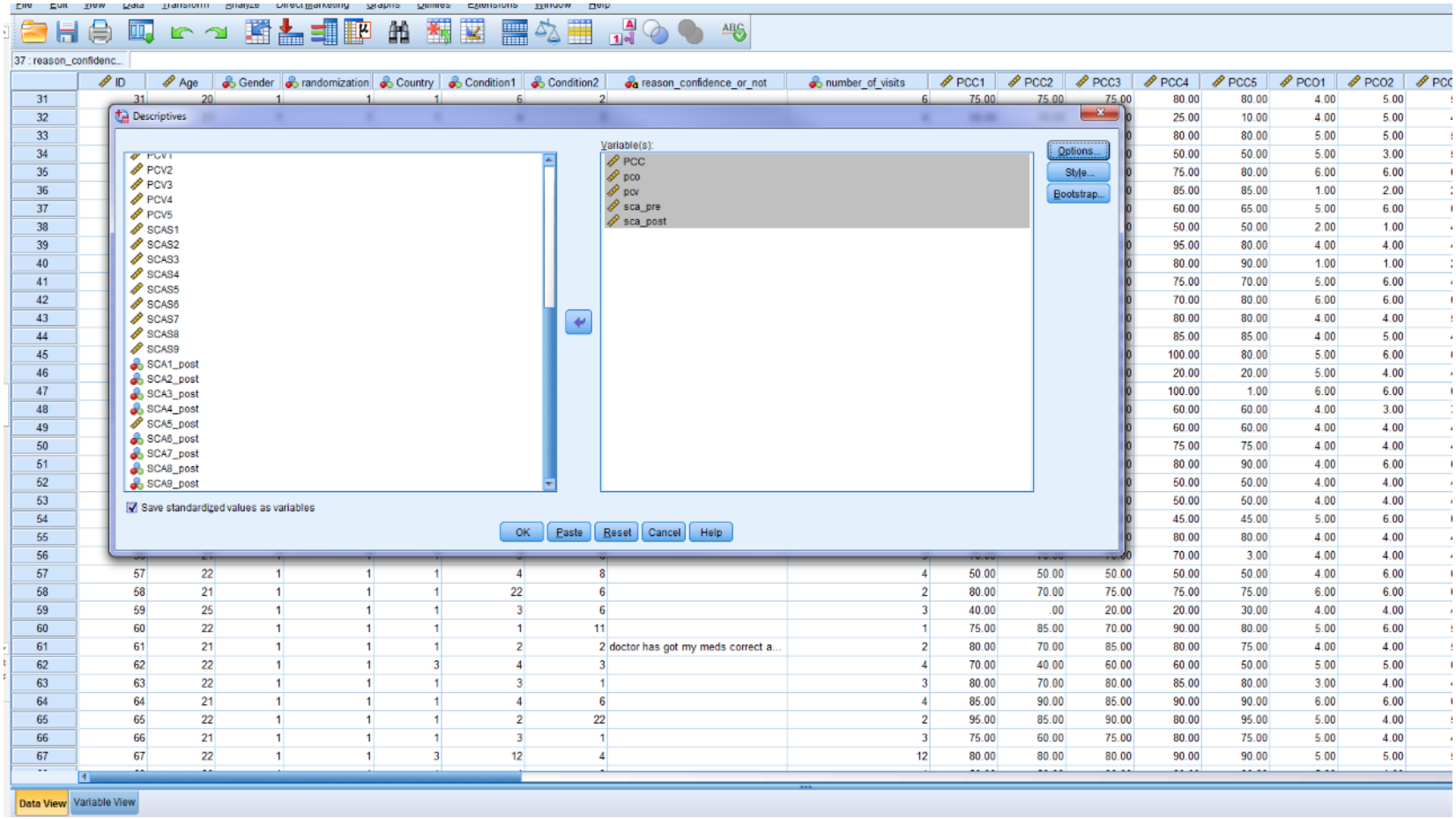
Transform -> Compute Variable (*mean* is under “Statistical” Function group)



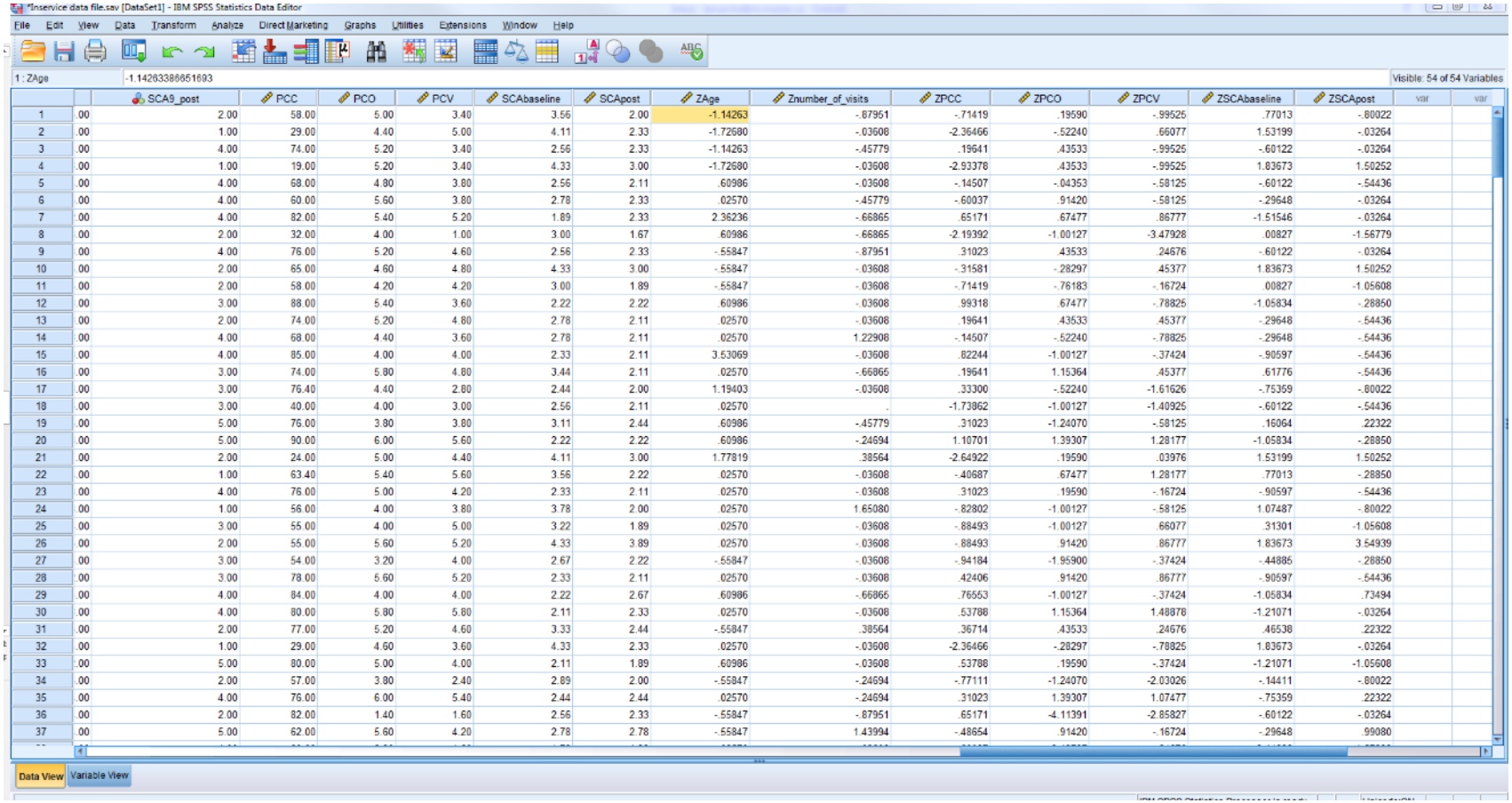
The clicks: Univariate outliers

Analyze -> Descriptives, select box “Save standardized values as variables”

Analyze -> Frequencies for z-scores



The Output: Scores and z-scores



*Inservice data file.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

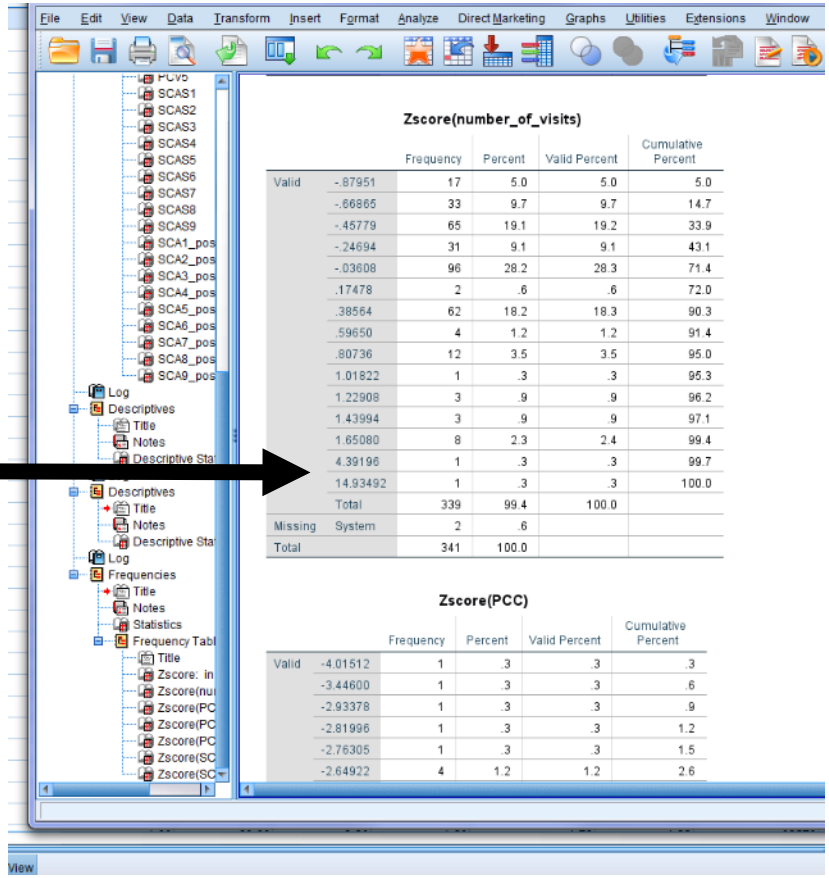
1 : ZAge -1.14263386651693 Visible: 54 of 54 Variables

		SCA9_post	PCC	PCO	PCV	SCAbaseline	SCApost	ZAge	Znumber_of_visits	ZPCC	ZPCO	ZPCV	ZSCAbaseline	ZSCApost	var	var
1	.00	2.00	58.00	5.00	3.40	3.56	2.00	-1.14263	-.87951	-.71419	.19590	-.99525	.77013	-.80022		
2	.00	1.00	29.00	4.40	5.00	4.11	2.33	-1.72680	-.03608	-2.36466	-.52240	.66077	1.53199	-.03264		
3	.00	4.00	74.00	5.20	3.40	2.56	2.33	-1.14263	-.45779	.19641	.43533	-.99525	-.60122	-.03264		
4	.00	1.00	19.00	5.20	3.40	4.33	3.00	-1.72680	-.03608	-2.93378	.43533	-.99525	1.83673	1.50252		
5	.00	4.00	68.00	4.80	3.80	2.56	2.11	.60986	-.03608	-.14507	-.04353	-.58125	-.60122	-.54436		
6	.00	4.00	60.00	5.60	3.80	2.78	2.33	.02570	-.45779	-.60037	.91420	-.58125	-.29648	-.03264		
7	.00	4.00	82.00	5.40	5.20	1.89	2.33	2.36236	-.66865	.65171	.67477	.86777	-1.51546	-.03264		
8	.00	2.00	32.00	4.00	1.00	3.00	1.67	.60986	-.66865	-2.19392	-1.00127	-3.47928	.00827	-1.56779		
9	.00	4.00	76.00	5.20	4.60	2.56	2.33	-.55847	-.87951	.31023	.43533	.24676	-.60122	-.03264		
10	.00	2.00	65.00	4.60	4.80	4.33	3.00	-.55847	-.03608	-.31581	-.28297	.45377	1.83673	1.50252		
11	.00	2.00	58.00	4.20	4.20	3.00	1.89	-.55847	-.03608	-.71419	-.76183	-.16724	.00827	-1.05608		
12	.00	3.00	88.00	5.40	3.60	2.22	2.22	.60986	-.03608	.99318	.67477	-.78825	-1.05834	-.28850		
13	.00	2.00	74.00	5.20	4.80	2.78	2.11	.02570	-.03608	.19641	.43533	.45377	-.29648	-.54436		
14	.00	4.00	68.00	4.40	3.60	2.78	2.11	.02570	1.22908	-.14507	-.52240	-.78825	-.29648	-.54436		
15	.00	4.00	85.00	4.00	4.00	2.33	2.11	3.63069	-.03608	.82244	-1.00127	-.37424	-.90597	-.54436		
16	.00	3.00	74.00	5.80	4.80	3.44	2.11	.02570	-.66865	.19641	1.15364	.45377	.61776	-.54436		
17	.00	3.00	76.40	4.40	2.80	2.44	2.00	1.19403	-.03608	.33300	-.52240	-1.61626	-.75359	-.80022		
18	.00	3.00	40.00	4.00	3.00	2.56	2.11	.02570	.	-1.73862	-1.00127	-1.40925	-.60122	-.54436		
19	.00	5.00	76.00	3.80	3.80	3.11	2.44	.60986	-.45779	.31023	-1.24070	-.58125	.16064	.22322		
20	.00	5.00	90.00	6.00	5.60	2.22	2.22	.60986	-.24694	1.10701	1.39307	1.28177	-1.05834	-.28850		
21	.00	2.00	24.00	5.00	4.40	4.11	3.00	1.77819	.38564	-2.64922	.19590	.03976	1.53199	1.50252		
22	.00	1.00	63.40	5.40	5.60	3.56	2.22	.02570	-.03608	-.40687	.67477	1.28177	.77013	-.28850		
23	.00	4.00	76.00	5.00	4.20	2.33	2.11	.02570	-.03608	.31023	.19590	-.16724	-.90597	-.54436		
24	.00	1.00	56.00	4.00	3.80	3.78	2.00	.02570	1.65080	-.82802	-1.00127	-.58125	1.07487	-.80022		
25	.00	3.00	55.00	4.00	5.00	3.22	1.89	.02570	-.03608	-.88493	-1.00127	.66077	.31301	-1.05608		
26	.00	2.00	55.00	5.60	5.20	4.33	3.89	.02570	-.03608	-.88493	.91420	.86777	1.83673	3.54939		
27	.00	3.00	54.00	3.20	4.00	2.67	2.22	-.55847	-.03608	-.94184	-1.95900	-.37424	-.44885	-.28850		
28	.00	3.00	78.00	5.60	5.20	2.33	2.11	.02570	-.03608	.42406	.91420	.86777	-.90597	-.54436		
29	.00	4.00	84.00	4.00	4.00	2.22	2.67	.60986	-.66865	.76553	-1.00127	-.37424	-1.05834	.73494		
30	.00	4.00	80.00	5.80	5.80	2.11	2.33	.02570	-.03608	.53788	1.15364	1.48878	-1.21071	-.03264		
31	.00	2.00	77.00	5.20	4.60	3.33	2.44	-.55847	.38564	.36714	.43533	.24676	.46538	.22322		
32	.00	1.00	29.00	4.60	3.60	4.33	2.33	.02570	-.03608	-2.36466	-.28297	-.78825	1.83673	-.03264		
33	.00	5.00	80.00	5.00	4.00	2.11	1.89	.60986	-.03608	.53788	.19590	-.37424	-1.21071	-1.05608		
34	.00	2.00	57.00	3.80	2.40	2.89	2.00	-.55847	-.24694	-.77111	-1.24070	-2.03026	-.14411	-.80022		
35	.00	4.00	76.00	6.00	5.40	2.44	2.44	.02570	-.24694	.31023	1.39307	1.07477	-.75359	.22322		
36	.00	2.00	82.00	1.40	1.60	2.56	2.33	-.55847	-.87951	.65171	-4.11391	-2.85827	-.60122	-.03264		
37	.00	5.00	62.00	5.60	4.20	2.78	2.78	-.55847	1.43994	-.48654	.91420	-.16724	-.29648	.99080		

Data View Variable View

The Output: z-scores

14.93492



Zscore(number_of_visits)

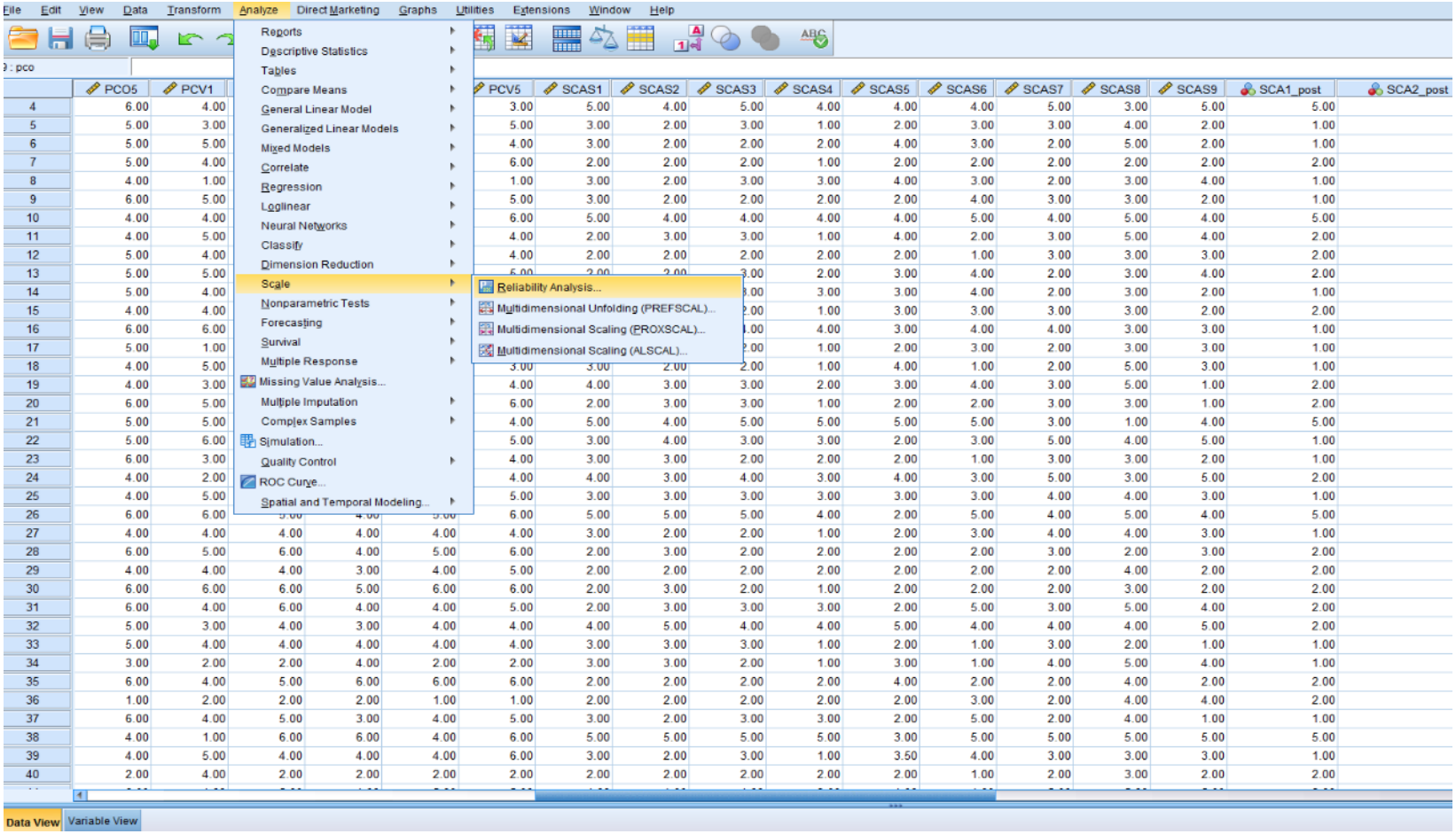
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-.87951	17	5.0	5.0	5.0
	-.66865	33	9.7	9.7	14.7
	-.45779	65	19.1	19.2	33.9
	-.24694	31	9.1	9.1	43.1
	-.03608	96	28.2	28.3	71.4
	.17478	2	.6	.6	72.0
	.38564	62	18.2	18.3	90.3
	.59650	4	1.2	1.2	91.4
	.80736	12	3.5	3.5	95.0
	1.01822	1	.3	.3	95.3
	1.22908	3	.9	.9	96.2
	1.43994	3	.9	.9	97.1
	1.65080	8	2.3	2.4	99.4
	4.39196	1	.3	.3	99.7
	14.93492	1	.3	.3	100.0
	Total	339	99.4	100.0	
Missing	System	2	.6		
Total		341	100.0		

Zscore(PCC)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-4.01512	1	.3	.3	.3
	-3.44600	1	.3	.3	.6
	-2.93378	1	.3	.3	.9
	-2.81996	1	.3	.3	1.2
	-2.76305	1	.3	.3	1.5
	-2.64922	4	1.2	1.2	2.6

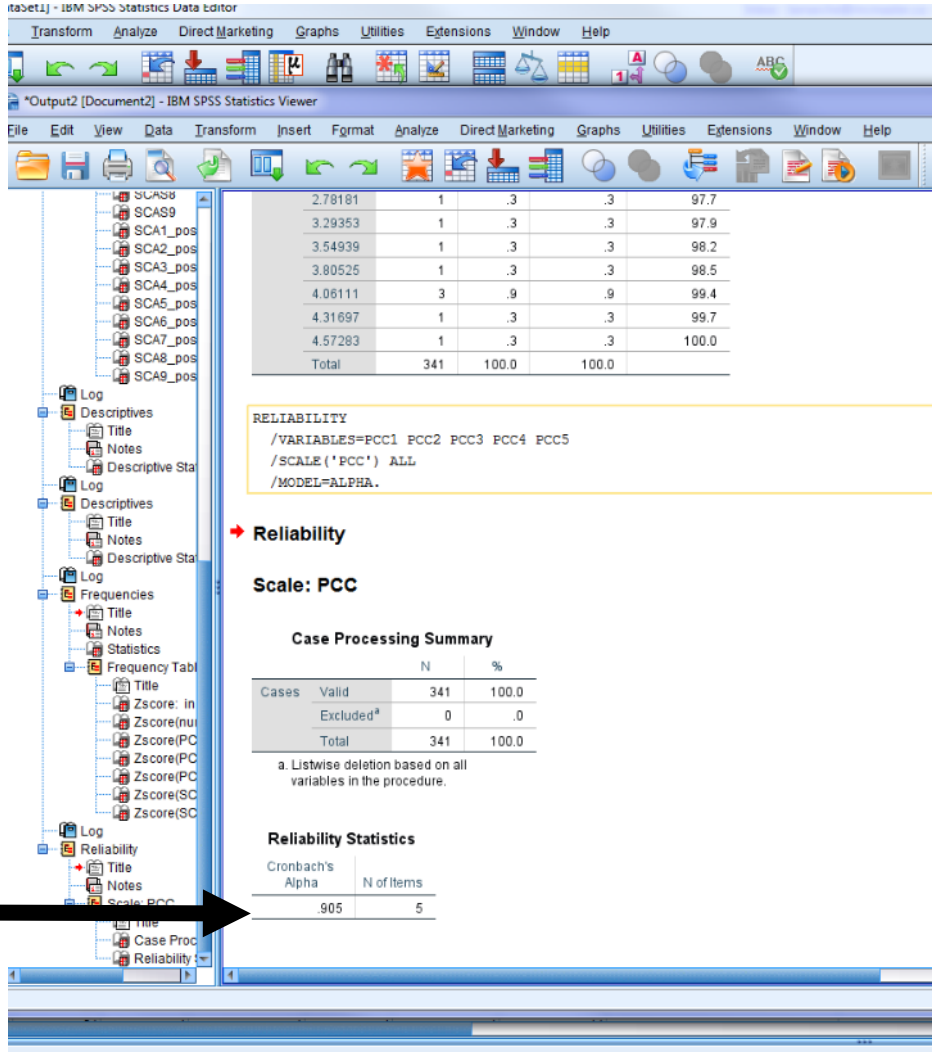
The clicks: Check internal consistency

Analyze -> Scale -> Reliability analyses



The Output: Internal consistency

α = .905

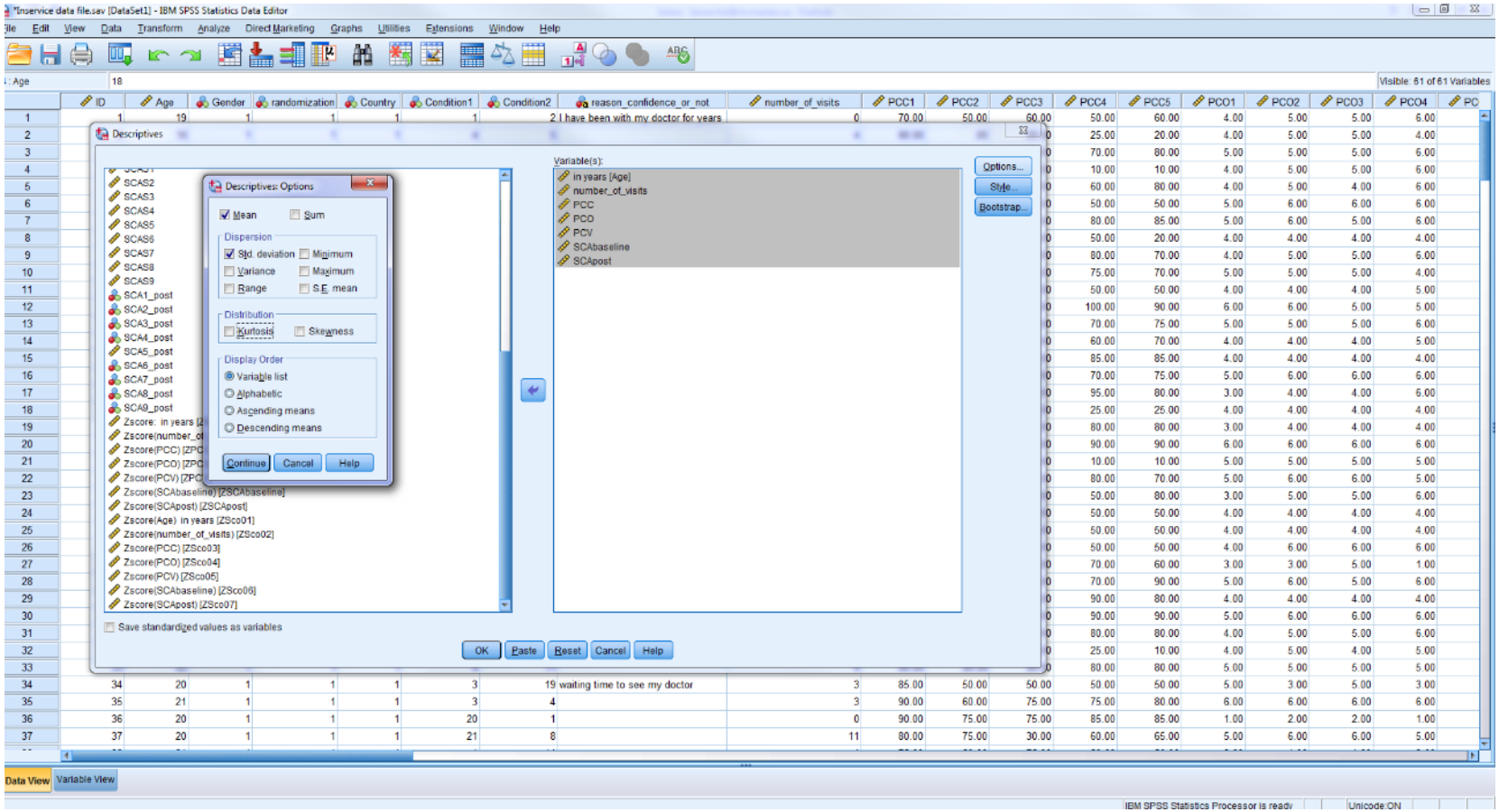


Descriptive statistics

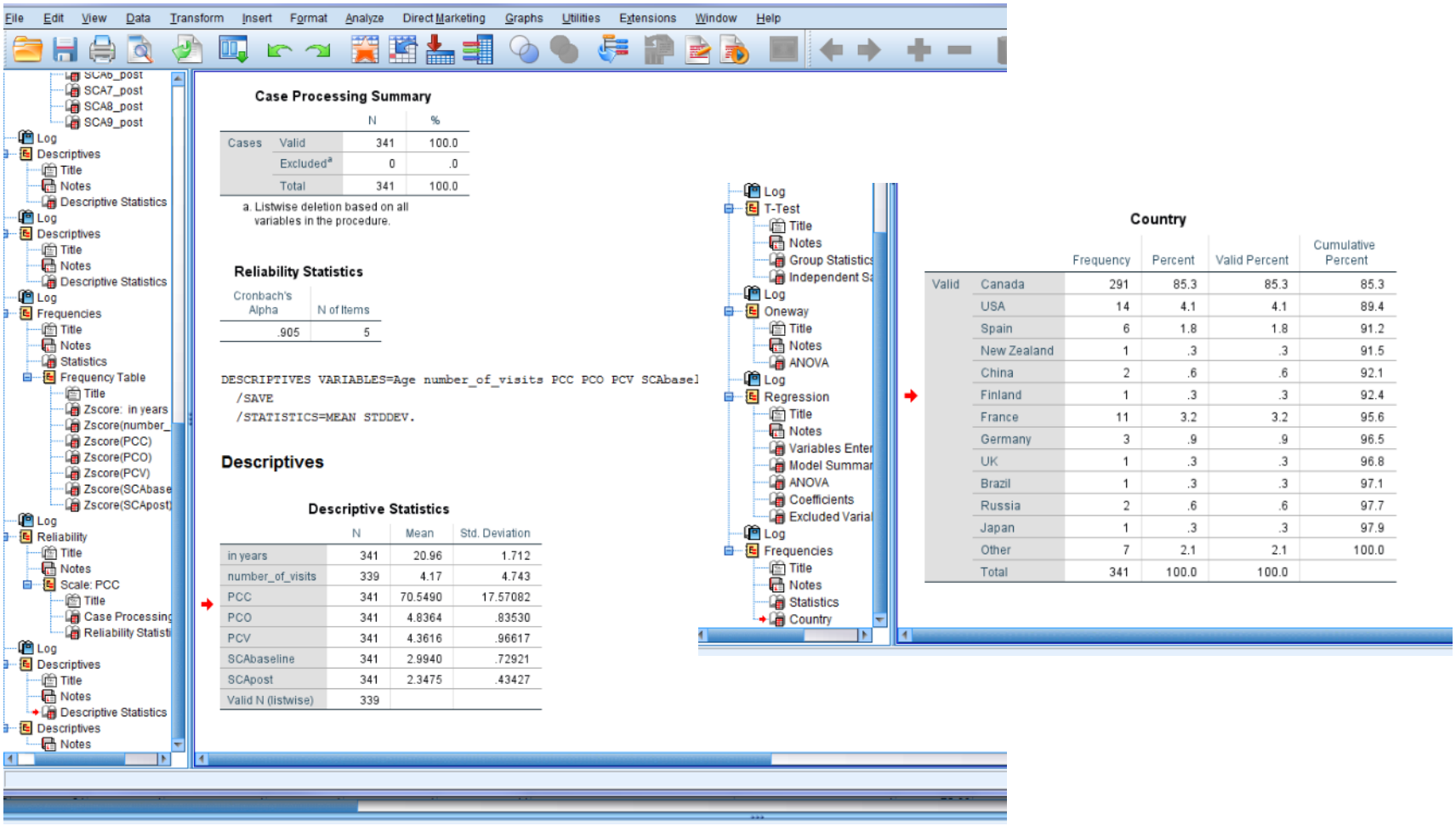
- Analyses that help to summarize data in meaningful way
  - Central tendency (mean, median, mode)
  - Dispersion or spread (range, variance, standard deviation)

The clicks: Descriptive statistics

Analyze -> Descriptive statistics (Frequencies for categorical data) -> Descriptives

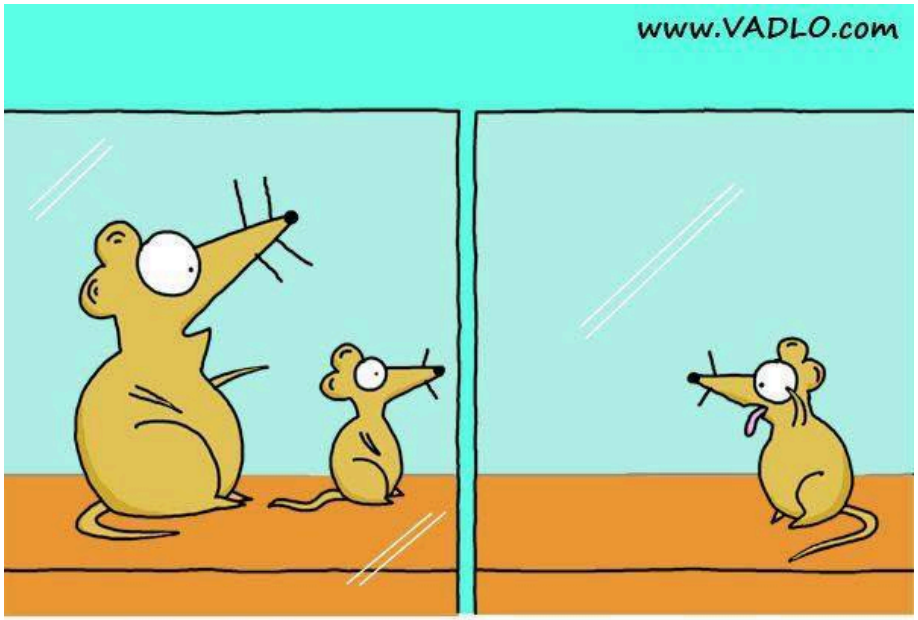


The output: Descriptive statistics



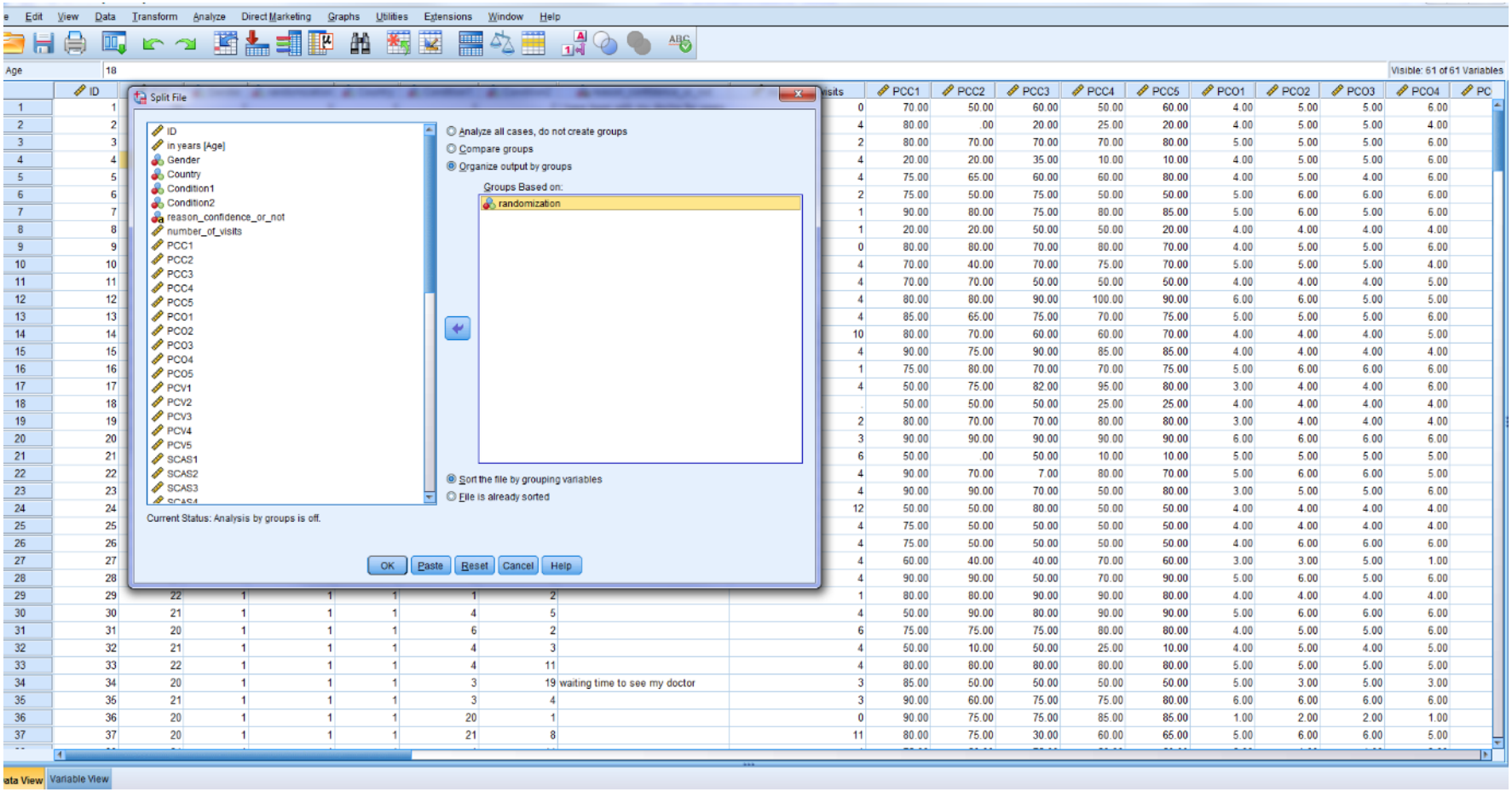
Understand the split file function

- Grouped data and descriptive statistics



Understand the split file function: The clicks

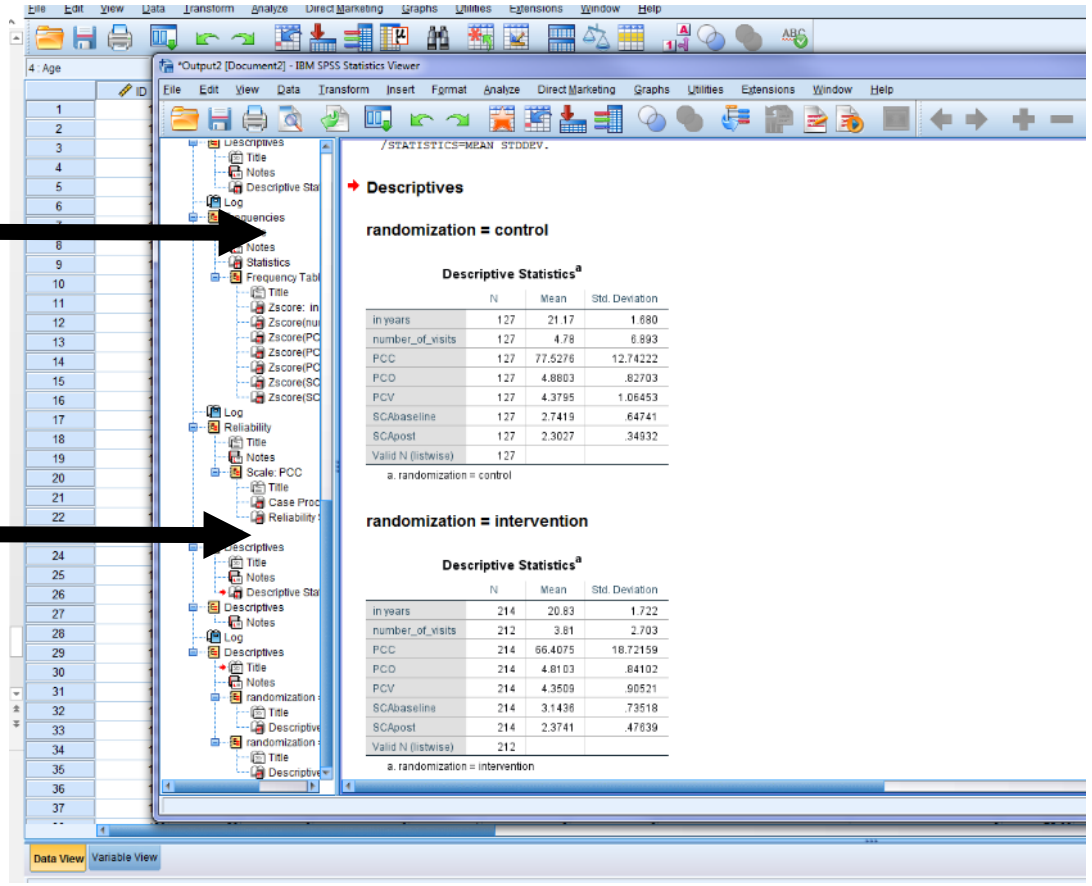
Date -> Split file…-> Organize output by groups (select grouping variable)



The output: Understand the split file function

Control

Intervention



The clicks: Statistical tests

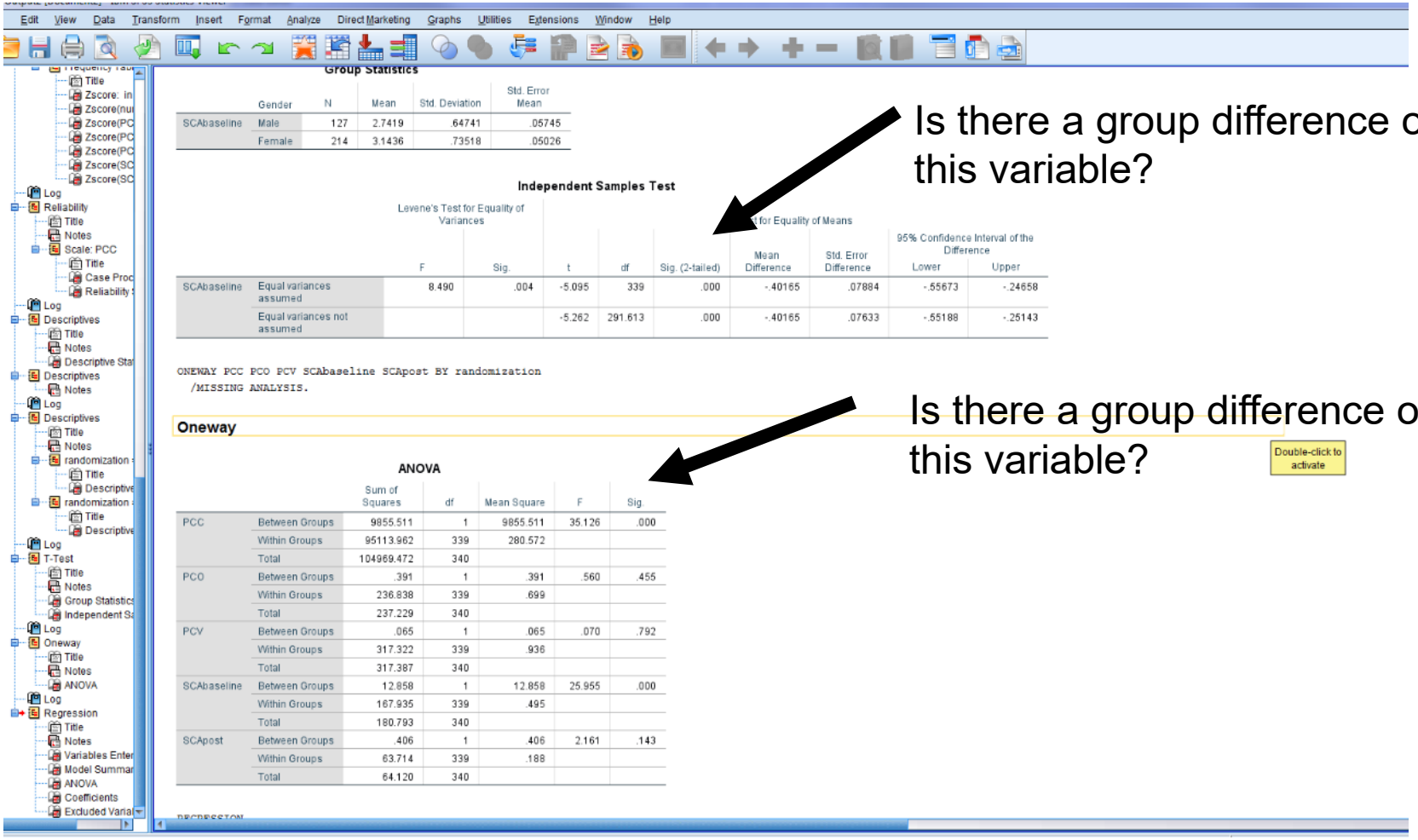
Are groups different? (Analyze -> compare means)

- t-test
- ANOVA

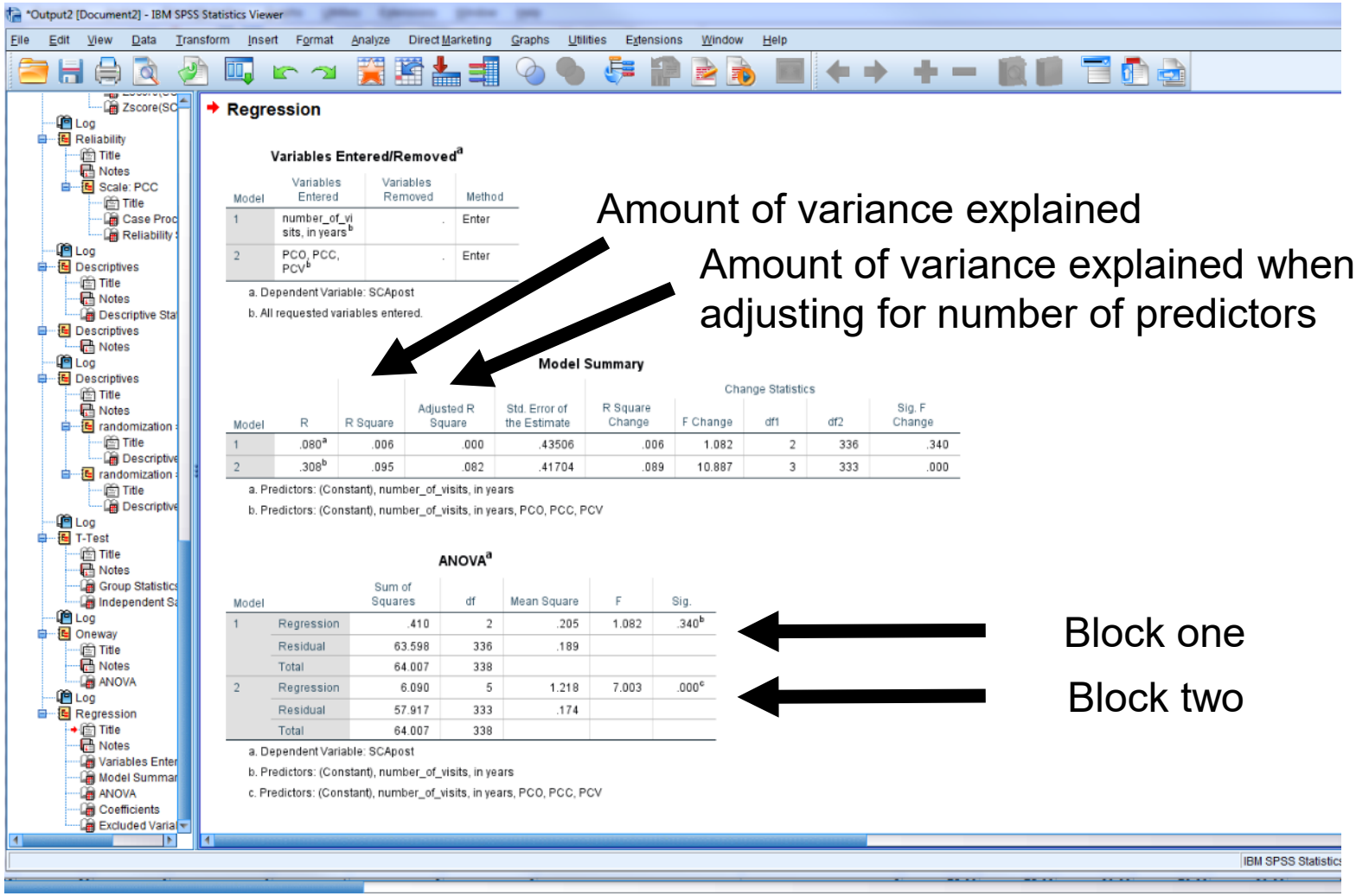
Prediction one variable from another (Analyze -> Linear regression)

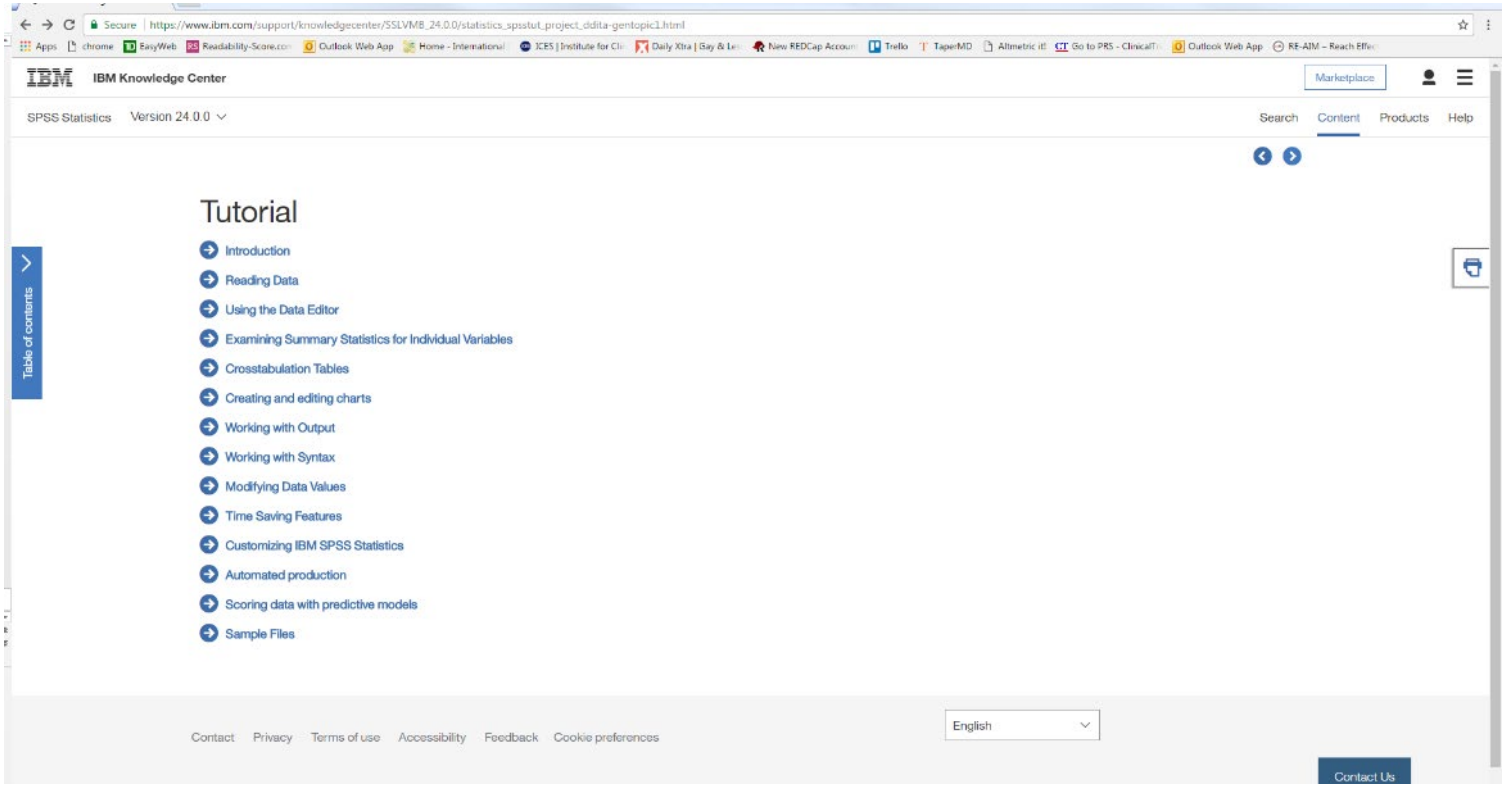
- Linear regression

The Output: Group differences



The Output: Regression



Wicked Resources

https://www.ibm.com/support/knowledgecenter/SSLVMB_24.0.0/statistics_spsstut_project_ddita-gentopic1.html