SPSS: Up and running for academic research

iT Centre Learning

iT services

UNIVERSITY OF OXFORD

The small print

Prerequisites

Time in the workshop is precious – it is an opportunity for you to interact with the workshop leader and other participants through questions and discussions and to share your experiences and concerns. To make the most of this time we sometimes ask you to carry out learning activities ahead of the workshop so that everyone comes into the class with the same basic knowledge. We keep this prior learning to a minimum and often make use of online videos. Online videos provided through LinkedIn Learning can be accessed free of charge by University members anytime, anywhere, through a browser or app.

Your course booking will tell you if any prior learning activity is required. If you don't have an environment where you can do this learning, you can come along to one of our LinkedIn Learning sessions. These are a quiet space where you can work through videos or other workshop resources.

If you arrive for a workshop without having done the prior learning, the workshop leader may suggest that you come back on another session.

Copyright

This version of the course materials is adapted from teaching slides and handbook created by Kerstin Frie and subsequently edited by Rachel Pechey. Jacqueline Murphy updated the teaching slides (text edits and screenshots of the latest version of SPSS) and revised the handbook structure.

Jacqueline Murphy makes this booklet and the accompanying slides available under a Creative Commons licence (BY-NC-SA: Attribution-NonCommercial-ShareAlike).

The Oxford University crest and logo and IT Services logo are copyright of the University of Oxford and may only be used by members of the University in accordance with the University's branding guidelines.

About the workshop designer

Jacqueline has expertise using SPSS, R and Stata software for statistical analysis, and has been working in quantitative medical research since 2010, including as a researcher and statistician for University of Oxford and Queen Mary University, London. Jacqueline has previously been a teaching assistant (demonstrator) in statistics for MSc level courses alongside her work in Oxford and is an Associate Fellow of the Higher Education Academy.

Revision history

Version	Date	Author	Comments
1.0	2019	Kerstin Frie	Created original teaching slides
2.0	2020/21	Rachel Pechey	Edited teaching slides and handbook
3.0	October 2022	Jacqueline Murphy	Revised handbook structure and edited teaching slides (text edits and updated SPSS screenshots)

About this workshop

SPSS is a statistics and data analysis program. This course aims to provide participants with the skills to handle data and perform statistical analyses using SPSS. The course is at beginner level and is aimed at researchers and others who want to use intuitive and flexible statistical software that does not require writing code.

What you will learn

During the course, we will look at the SPSS interface (including output file vs data file; variable view vs data view; how to create a variable, define a variable, add data).

You will learn to organise your data well in SPSS, sorting and selecting, managing duplicates and case, and controlling variables. With descriptive analysis, you will use frequencies, custom tables and graphs. SPSS offers tools for performing and interpreting tests for associations (bivariate and partial correlations and regressions) and tests for means and variances (t-tests and ANOVAs).

The emphasis is on using SPSS to perform statistical analyses, rather than teaching statistical theory. Further resources for learning statistical analysis using SPSS will be signposted in the course slides.

What you need to know

Prerequisites for the course: A basic understanding of statistical concepts (such as mean, median, standard deviation, p-value) is expected.

If you need to review this knowledge, LinkedIn Learning is a great place to get guidance. There is an activity with relevant videos in the IT Learning Portfolio: visit skills.it.ox.ac.uk/it-learning-portfolio and select or search for "SPSS".

The resources you need

The computer, software, and course data files are provided. For participants using the IT Learning Centre devices the course files will be made available directly on the computers before you arrive.

Alternatively, you can bring your own with the latest version of the SPSS software already installed (read our guidance first). You will need to be able to copy the provided data files from the classroom computers onto your own device (i.e. using a USB memory stick). There will be limited time for troubleshooting technical (e.g. installation) problems with personal devices during the course so you are advised to use the classroom computers unless you are familiar with setting up statistical software on your own device.

Course Objectives

This workshop has the following objectives which will be addressed through teacher-led demonstrations, exercises for participants to complete, and a take-home course handbook:

Objective One: Be familiar with the layout and structure of SPSS

Objective Two: Organise data in SPSS

Objective Three: Become familiar with SPSS outputs for a range of statistical tests

Objective Four: Become familiar with SPSS outputs for a range of statistical tests

Further information

Getting extra help

The IT Learning Centre offers bookable clinics where you can get pre- or post-course advice. Contact us using courses@it.ox.ac.uk.

Study Videos from LinkedIn Learning

On our website, you will find our collection of self-service courses and resources. This includes providing LinkedIn Learning video-based courses free to all members of the University. Visit skills.it.ox.ac.uk/linkedin-learning and sign in with your Single Sign-On (SSO) credentials.

Some courses recommend pre- and/or post-course activities to support your learning. You can watch the online videos anywhere, anytime, and even download them onto a tablet or smartphone for off-line viewing.

About the IT Learning Portfolio online

Many of the resources used in the IT Learning Centre courses and workshops are made available as Open Educational Resources (OER) via our Portfolio website at skills.it.ox.ac.uk/it-learning-portfolio and select or search for "SPSS".

Find the pre-course activity for this course in the IT Learning Portfolio: visit skills.it.ox.ac.uk/it-learning-portfolio and select or search for "SPSS".

A copy of the SPSS data sets used during the course is provided at: https://skills.web.ox.ac.uk/spss-up-and-running-for-academic-research-course-pack

About the IT Learning Centre

The IT Learning Centre delivers over 100 IT-related teacher-led courses, which are provided in our teaching rooms and online, and we give you access to thousands of on-line self-service courses through LinkedIn Learning.

Our team of teachers have backgrounds in academia, research, business and education and are supported by other experts from around the University and beyond.

Our courses are open to all members of the University at a small charge. Where resources allow, we can deliver private courses to departments and colleges, which can be more cost-effective than signing up individually. We can also customize courses to suit your needs.

Our fully equipped suite of seven teaching and training rooms are usually available for hire for your own events and courses.

For more information, contact us at courses@it.ox.ac.uk.

About IT Customer Services

The IT Learning Centre is part of the Customer Services Group. The group provides the main user support services for the department, assisting all staff and students within the University as well as retired staff and other users of University IT services. It supports all the services offered by IT Services plus general IT support queries from any user, working in collaboration with local IT support units.

The Customer Services Group also offers a data back-up service; an online shop; and a computer maintenance scheme. Customer Services is further responsible for desktop computing services – for staff and in public/shared areas – throughout UAS and the Bodleian Libraries.

SPSS: Up and running for academic research

Jacqueline Murphy

With thanks to Kerstin Frie and Rachel Pechey

SPSS

- SPSS is a widely used statistical software package produced by IBM
- Allows you to perform the most common statistical calculations
- No coding needed, intuitive point and click interface
- SPSS is available for Windows and Mac, works in exactly the same way on both systems

Structure of this Course

- The course is structured around a series of Research Questions (RQs)
- In-person sessions:
  - Instructor-led demonstrations
  - Give-it-a-go exercises with trouble-shooting help
- Handbook:
  - Contains all computing steps for the demonstrations and exercises
  - Glossary of statistical concepts

https://skills.web.ox.ac.uk/spss-up-and-running-for-academic-research-course-pack

Research Questions	Organising Data	Descriptive Analysis	Graphing	Simple Stats	Associations	ANOVA	Output Usage
RQ1: Frequency sweet vs savoury breakfast	Compute variables Recode variables		Bar chart				
RQ2: Units breakfast	Compute variables	Descriptives		One-sample t-test			
RQ3: Effect of food types on satiety				Paired-samples t-test			
RQ4: Differences in height by gender	Compute variables			Independent-samples t-test			
RQ5: Relationship age and height	Splitting files		Scatterplot		Bivariate correlation		
RQ 6: Relationship general pain and age, controlled for health			Pie chart		Partial correlation		
RQ 7: Reaction time to emotional facial expressions	Sorting cases Selecting cases	Frequencies	Boxplot		Linear regression		
RQ 8: Effect of supermarket promotion on sales	Duplicate cases					One-way ANOVA	
RQ 9: Effect of new pain treatment		Custom Tables				Two-way ANOVA	Formatting graphs
RQ 10: Effect of weight loss programme	Merge files					Repeated /mixed ANOVA	Cleaning output Formatting tables Exporting graphs and tables
Additional Features	Find and Replace	Explore Data Crosstabs					

Before we start…

- Please feel free to ask clarifying questions during demonstrations
- Please do ask for help during the give-it-a-go exercises

- All data are adapted from SPSS example data sets or fabricated
- Please feel free to save the example data on a USB drive so you can use it at home (tip: keep a "master" copy of the original unedited data)

Statistical concepts we will be using

- Mean/Standard Deviation/Variance
- Between- vs Within-Subjects Variables
- Significance level and p-values
- Confidence Intervals
- t-tests
- Correlations
- ANOVAs
- Regressions

Getting started in SPSS

- Using Data in SPSS
  - Import the Import.xlsx file into SPSS
  - Open the Breakfast.sav file in SPSS
- Data vs. Output
  - Take a look at the Data and Output windows
- Variable view
  - In the *Variable View* of the data file, create a new variable named age
  - Define the variable age as numeric, with 0 decimals, at scale measure

RQ1: Frequency Sweet vs Savoury Breakfast

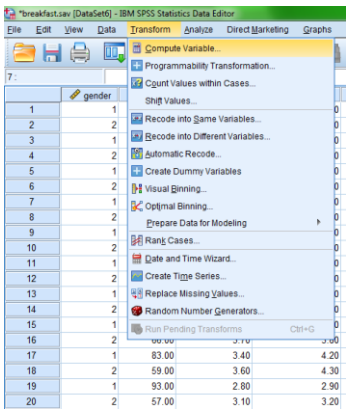
- Data set of 31 participants' breakfast consumption over 40 days
- 11 variables: gender, satiety of sweet & savoury breakfast, breakfast units split up by breakfast items: 3 x savoury, 5 x sweet

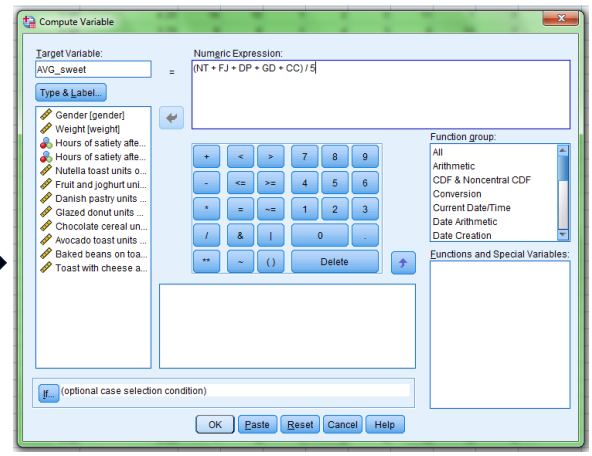
☛ Do participants consume predominantly sweet or savoury foods for breakfast?

- Compute averages sweet vs. savoury foods
- Compute difference sweet vs. savoury foods
- Recode difference score into a categorical variable
- Bar chart: Frequency predominantly sweet vs. savoury foods

RQ1: Frequency Sweet vs. Savoury Breakfast
Compute variables: average sweet items

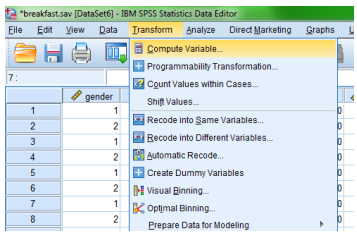
Go to *Transform* and select *Compute Variable*...

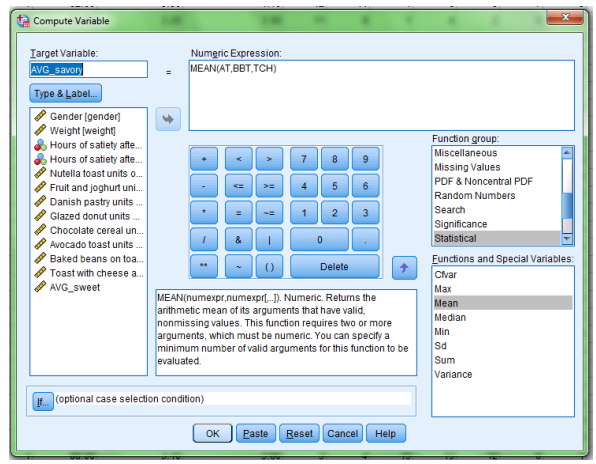




- Enter a target variable name
- Double-click on the five sweet breakfast variables to add them to the numeric expression box
- Enter "+" between the variables to add the values and put brackets around all variables, then divide by 5
- Click *OK*

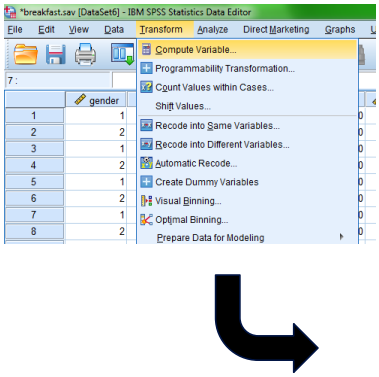
RQ1: Frequency Sweet vs. Savoury Breakfast
Compute variables: average savoury items

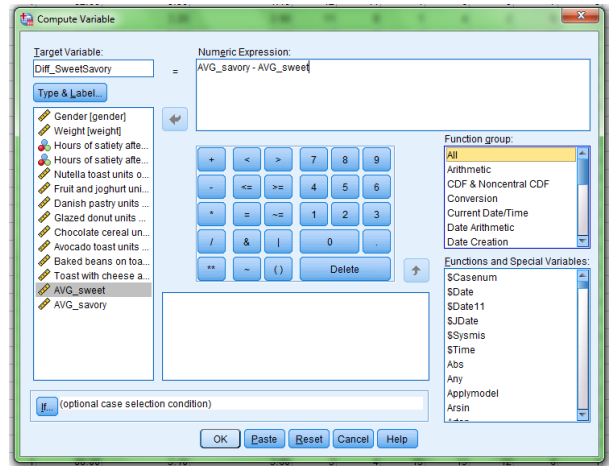




- Select *Compute Variable…* from the tab *Transform*
- Enter a target variable name
- Find the *Mean* function in the function group *Statistical*
- Double-click on all the variables you want to include in the calculation, make sure they are entered between the brackets, separated by commas
- Click *OK*

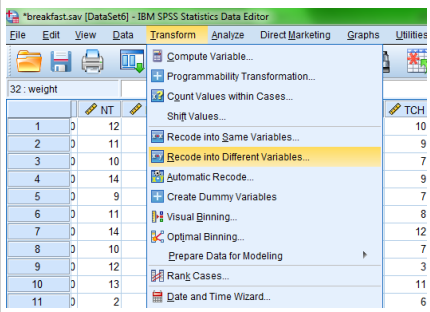
RQ1: Frequency Sweet vs. Savoury Breakfast
Compute variables: difference sweet vs. savoury



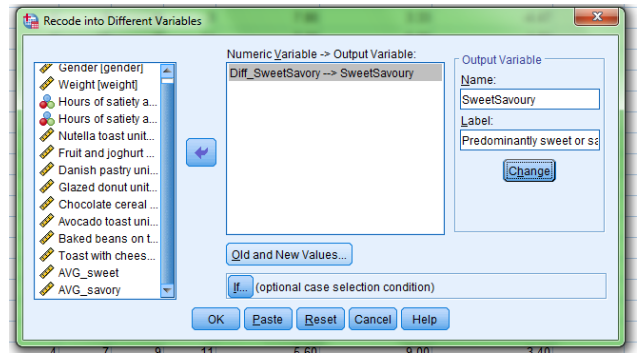


- Select *Compute Variable…* from the tab *Transform*
- Enter a target variable name
- Double-click on the variables you want to include in the calculation
- Add "-" between the variables to substract one value from another
- Click *OK*

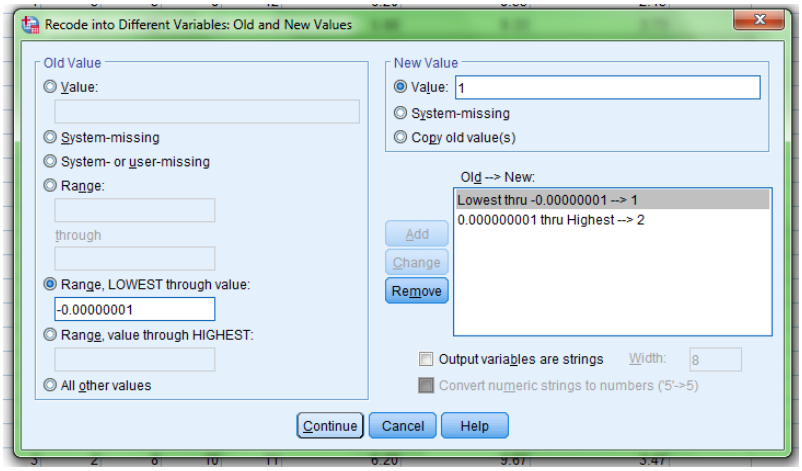
RQ1: Frequency Sweet vs. Savoury Breakfast
Recode variables: sweet vs. savoury



- Select *Recode into Different Variables…* from the tab *Transform*
- Add the variable you want to recode to the middle box
- Choose a name for the output variable (recoded values)
- Optional: add a label to remember what the variable consists of
- Click *Change*
- Click *Old and New Values…* to define the recoding

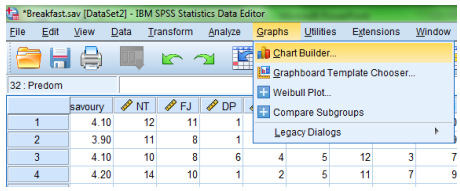


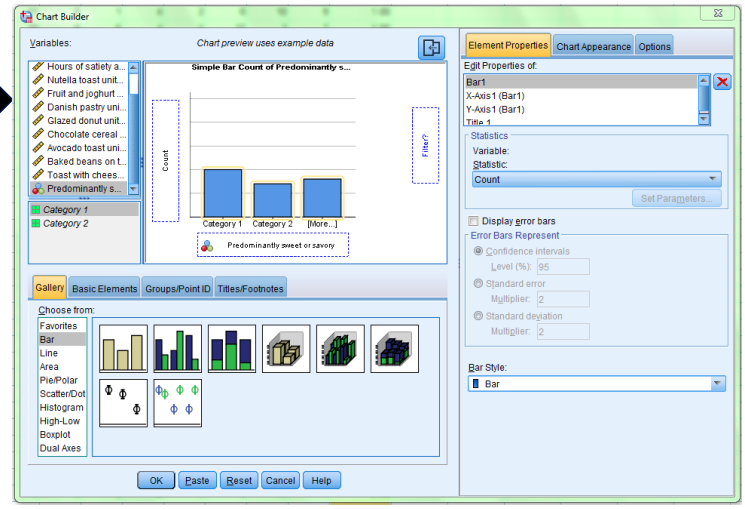
RQ1: Frequency Sweet vs. Savoury Breakfast
Recode variables: sweet vs. savoury



- A new box will have appeared
- Select *Range, LOWEST through value* and enter a negative value close to zero
- Enter "1" into the box *Value* in the section *New Value* and then click *Add* to confirm this recoding
- Select *Range, value through HIGHEST* and enter a positive value close to zero
- Enter "2"(or any other value but 1) into the box *Value* and add this recoding to the box
- Click *Continue* to get back to the previous dialogue
- Click *OK* to start the recoding

RQ1: Frequency Sweet vs. Savoury Breakfast
Bar chart: Frequency sweet vs savoury





- Select *Chart Builder* from the tab *Graphs*
- In the bottom section, choose the category *Bar* in order to see the different bar chart options
- Double-click on the simple bar chart (top left)
- Drag the recoded variable from the variable list to the x-axis
- The y-axis should automatically change to Count
- Click "OK"

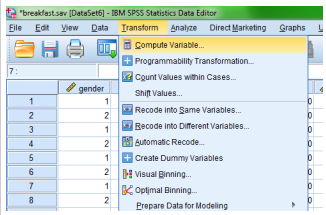
RQ2: Breakfast Units

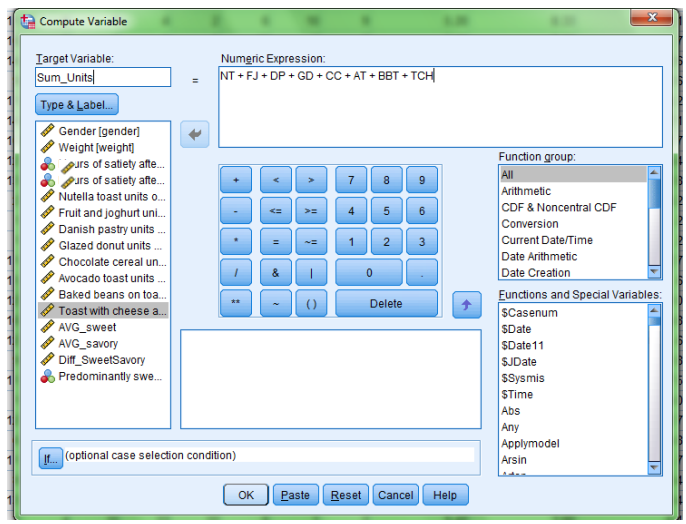
- Same data set as RQ1

☛Do participants consume more than one unit of breakfast per day?

- Compute variables: sum of breakfast items
- Get to know the data: Descriptives of the variable in question
- One sample t-test: Did the participants consume more than one unit of breakfast a day (i.e. 40 units over 40 days)?

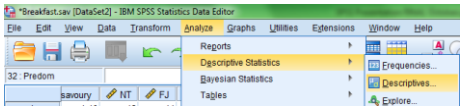
RQ2: Breakfast Units
Compute variables: sum of breakfast items

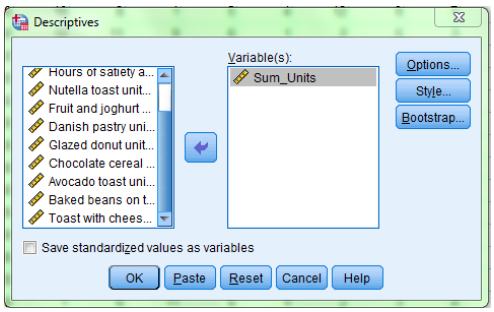


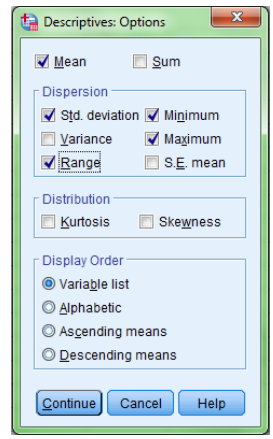


- Select *Compute Variable…* from the tab *Transform*
- Enter a target variable name
- Double-click on all breakfast variables to add them to the numeric expression box
- Enter "+" between the variables to add the values
- Click *OK*

RQ2: Breakfast Units
Descriptive Statistics

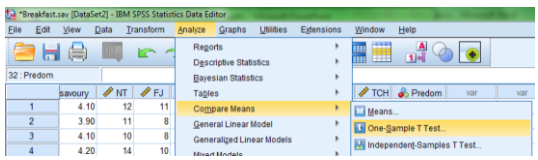


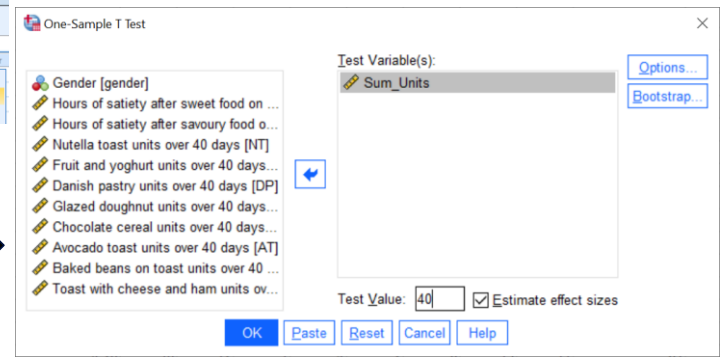




- Select *Descriptives…* under *Descriptive Statistics* in the tab *Analyze*
- Add the variable Sum to the *Variable(s)* box
- Click on *Options…* Select the statistics of interest, then click on *Continue*
- Click *OK*

RQ2: Breakfast Units
One-sample t-test





- Select *One-Sample T Test…* under *Compare Means* in the tab *Analyze*
- Double-click on the sum variable to add it to the box
- Enter the value against which you want to test the data in the *Test Value* box, in this case 40
- Click *OK*

RQ2: Breakfast Units
One-sample t-test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Sum_Units	31	56.03	6.585	1.183

One-Sample Test

Test Value = 40

	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper
Sum_Units	13.555	30	<.001	<.001	16.032	13.62	18.45

One-Sample Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Sum_Units	Cohen's d	6.585	2.435	1.722	3.136
	Hedges' correction	6.756	2.373	1.679	3.057

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

- SPSS returns a table with statistics, a table with the test results and a table with effect sizes
- The second table gives the p-value for this two-tailed test in the column *Significance (Two-sided p)*, if this value is below 0.05, it is considered significant = the sample's mean is significantly different from the value against which it was tested
- In this case, we find that the participants ate significantly more than 40 breakfast items across the 40 days of the study

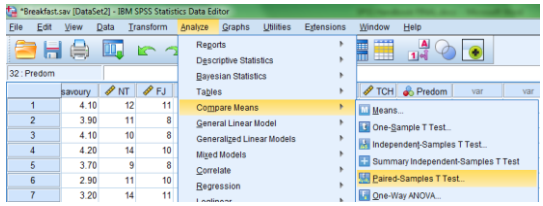
RQ3: Effect of Food Types on Satiety

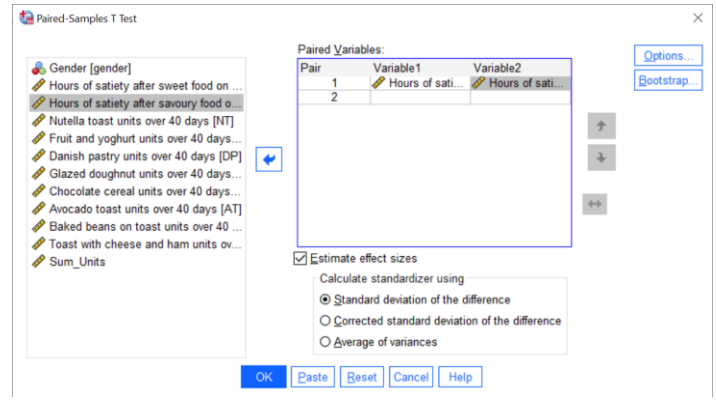
- Same data set as for RQ1 and RQ2
- Data on satiety of sweet vs. savoury breakfast

☛Do sweet or savoury foods fill you up for longer?

- Paired-samples t-test

RQ3: Effect of Food Types on Satiety
Paired-samples t-test





- Select *Paired-Samples T Test…* under *Compare Means* in the tab *Analyze*
- Double-click on the two satiety variables to add them to the two columns of the box
- Click *OK*

RQ3: Effect of Food Types on Satiety
Paired-samples t-test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Hours of satiety after sweet food on average	2.9806	31	.50360	.09045
	Hours of satiety after savoury food on average	3.5452	31	.56618	.10169

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	Hours of satiety after sweet food on average & Hours of satiety after savoury food on average	31	.838	<.001	<.001

- SPSS returns four tables, one with statistics, one with correlation results, one with the test results and one with effect sizes
- The main results are in the third table *Paired Samples Test* (see next slide)

RQ3: Effect of Food Types on Satiety
Paired-samples t-test (cont.)

Paired Samples Test

		Paired Differences					t	df	Significance	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			One-Sided p	Two-Sided p
Pair 1	Hours of satiety after sweet food on average - Hours of satiety after savoury food on average	-.56452	.31043	.05575	-.67838	-.45065	-10.125	30	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Pair 1	Hours of satiety after sweet food on average - Hours of satiety after savoury food on average	Cohen's d	.31043	-1.819	-2.390	-1.235
		Hedges' correction	.31438	-1.796	-2.360	-1.219

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

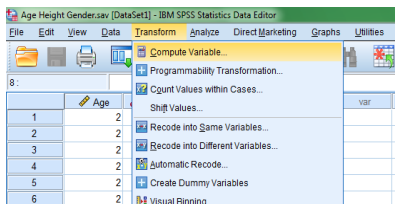
- SPSS gives the p-value for this two-tailed test in the column *Two-Sided p*, if this value is below 0.05, it is considered significant = the two samples tested are different from each other
- In this case, we find a significant difference between the satiety after savoury vs sweet food, whereby the savoury food kept participants full for longer

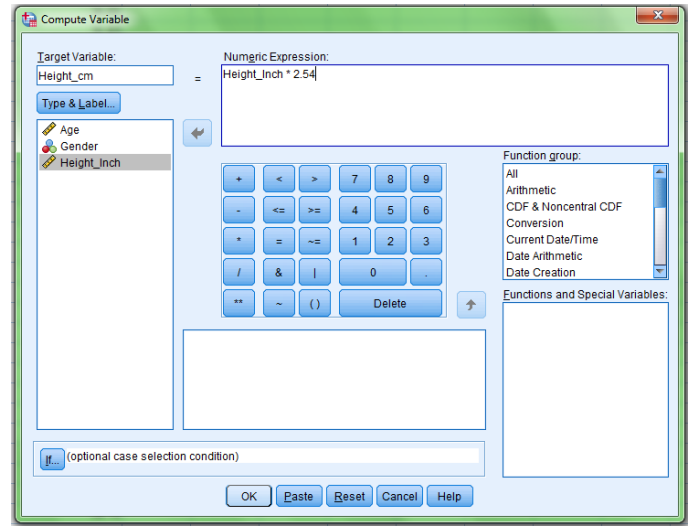
RQ4: Effect of Gender on Height

- Data set of 104 boys and girls aged 2-9
- 3 variables: Age, Height and Gender
- Height measured in inches

☛ Is there an effect of gender on height (in cm)?
- Convert height measurements from inches to centimeters
- Independent samples t-test

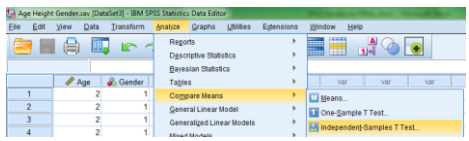
RQ4: Effect of Gender on Height
Convert Measurement

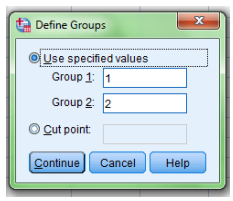


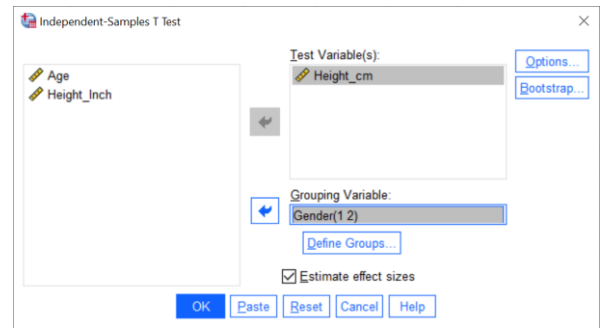


- Select *Compute Variable…* from the tab *Transform*
- Enter a target variable name
- Double-click on *Height_Inch* variable to add it to the numeric expression box
- Multiply with 2.54
- Click *OK*

RQ4: Effect of Gender on Height
Independent Samples t-test







- Select *Independent-Samples T Test…* under *Compare Means* in the tab *Analyze*
- Select the *Height_cm* variable and add it to the *Test Variable(s)* box
- Select the *Gender* variable and add it to the *Grouping Variable* box
- Click on *Define Groups...* to open the dialogue on the right, enter the values defining the two groups, in this case 1 and 2
- Click *Continue* and then *OK*

RQ4: Effect of Gender on Height
Independent Samples t-test

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Height_cm	Male	58	108.5517	15.26281	2.00411
	Female	46	115.9565	16.26579	2.39826

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Height_cm	Equal variances assumed	.289	.592	-2.387	102	.009	.019	-7.40480	3.10233	-13.55826	-1.25134
	Equal variances not assumed			-2.369	93.714	.010	.020	-7.40480	3.12539	-13.61059	-1.19901

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
Height_cm	Cohen's d	15.71319	-.471	-.862	-.078
	Hedges' correction	15.82993	-.468	-.856	-.077
	Glass's delta	16.26579	-.455	-.851	-.055

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

- SPSS runs the Levene's Test for Equality of Variances, since this test is not significant (*Sig.* value above 0.05), we use the t-test outcomes of the upper row (*equal variances assumed*)

- SPSS gives the p-value for this two-tailed test in the column *Significance (Two-Sided p)*, if this value is below 0.05, it is considered significant = the two samples tested are different from each other
- In this case, we find a significant difference in height between the groups, whereby the girls are on average taller than the boys

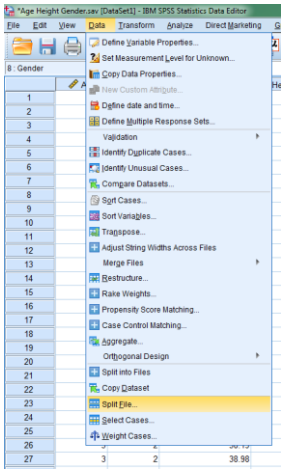
RQ5: Relationship of Age & Height, by gender

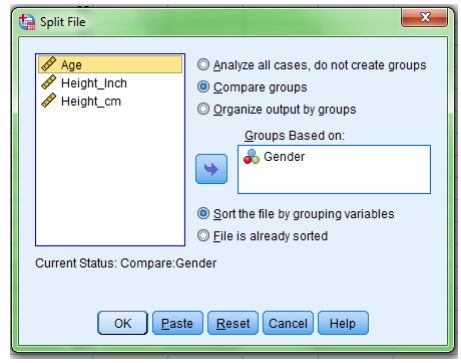
- Same data set as RQ4

☛ Is there a correlation of age and height? Is this dependent on gender?

- Splitting file: by gender
- Bivariate correlation
- Scatterplot

RQ5: Relationship of Age & Height, by gender
Splitting file by gender

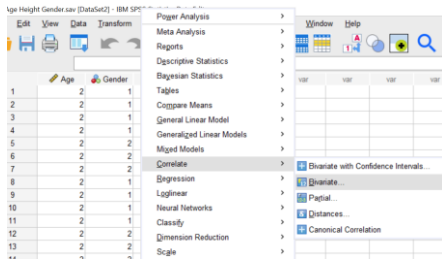


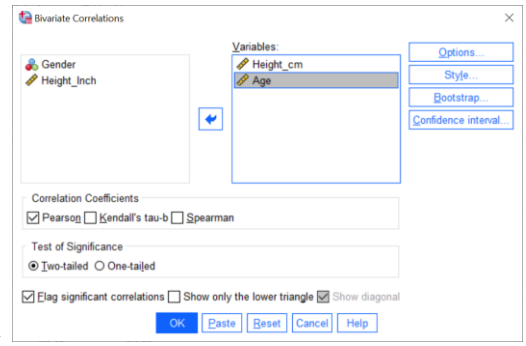


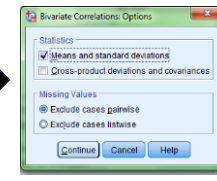
- Select *Split File…* in the tab *Data*
- Select the option *Compare groups*
- Double-click on the gender variable to add it to the box
- Click *OK*

- Lower right corner of the data window should now read "Split by gender"

RQ5: Relationship of Age & Height, by gender
Bivariate correlation







- Select *Bivariate…* under *Correlate* in the tab *Analyze*
- Add the variables Height_cm and Age to the box on the right
- Optional: Click on *Options…* to add statistics to the output, then click on *Continue*
- Select the correlation coefficients you want to have calculated, in this case Pearson
- Click *OK*

RQ5: Relationship of Age & Height, by gender
Bivariate correlation

Correlations

Descriptive Statistics

Gender		Mean	Std. Deviation	N
Male	Height_cm	108.55	15.263	58
	Age	5.31	2.280	58
Female	Height_cm	115.96	16.266	46
	Age	5.43	2.344	46

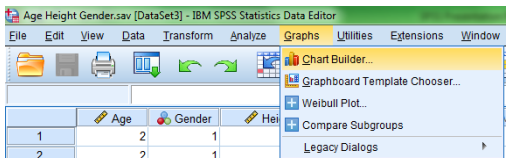
Correlations

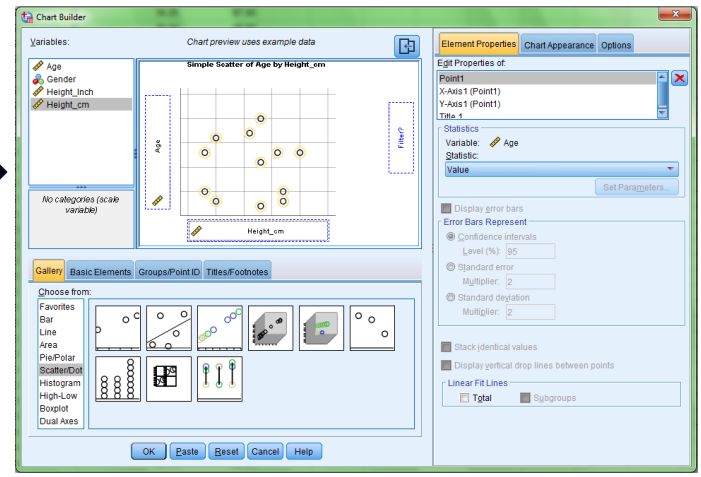
Gender			Height_cm	Age
Male	Height_cm	Pearson Correlation	1	.988**
		Sig. (2-tailed)		.000
		N	58	58
	Age	Pearson Correlation	.988**	1
		Sig. (2-tailed)	.000	
		N	58	58
Female	Height_cm	Pearson Correlation	1	.994**
		Sig. (2-tailed)		.000
		N	46	46
	Age	Pearson Correlation	.994**	1
		Sig. (2-tailed)	.000	
		N	46	46

**. Correlation is significant at the 0.01 level (2-tailed).

- SPSS returns two tables, the first shows descriptive statistics, the second presents the correlation results
- The tables are split by gender, SPSS returns correlation coefficients for males and females separately
- The first row *Pearson Correlation* is the Pearson correlation coefficient, it indicates the direction and strength of the association
- The second row *Sig. (2-tailed)* shows the p-value
- The third row *N* shows the sample size for each calculation
- In this case, we find that the significantly positive association of height and age exists for both boys and girls

RQ5: Relationship of Age & Height, by gender
Scatterplot





- Select *Chart Builder* from the tab *Graphs*
- In the bottom section, choose the category *Scatter/Dot* in order to see the different scatterplot options
- Double-click on the simple scatterplot(top left)
- From the list of variables, drag the variables age and Height_cm to the chart preview
- Click *OK*

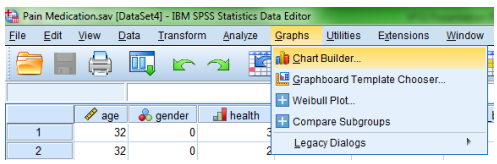
RQ6: Association of pain & age, controlling for health

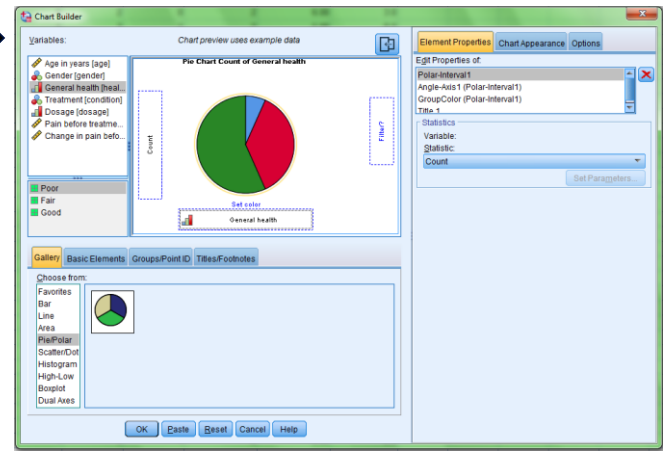
- Data set of 60 participants from a pain treatment study
- 7 variables: age, gender, general health, condition (control or treatment), dosage (low or high), pain at baseline, pain change after treatment

☛ Is there an association between pain and age, while controlling for general health?

- Pie chart: general health
- Partial correlation: age and baseline pain, controlling for general health

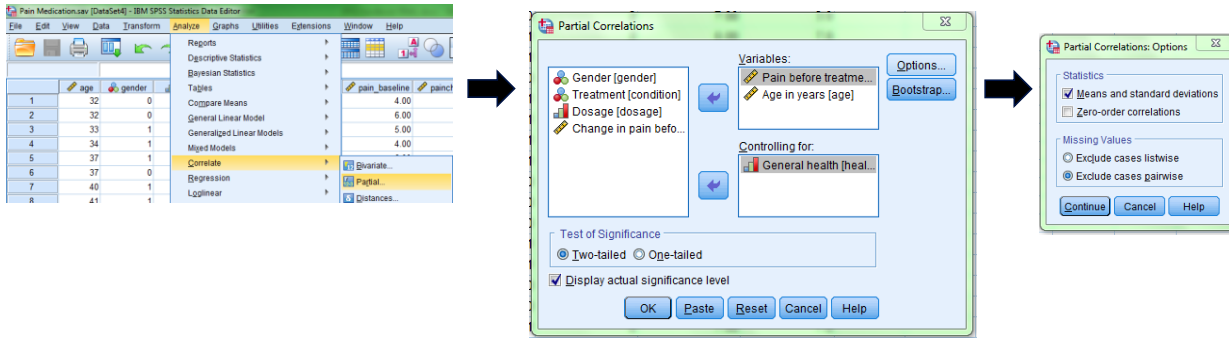
RQ6: Association of pain & age, controlling for health
Pie Chart





- Select *Chart Builder* from the tab *Graphs*
- In the bottom section, choose the category *Pie/Polar*
- Double-click on the depicted pie chart
- From the list of variables, drag the variable General health to the chart preview
- Click *OK*

RQ6: Association of pain & age, controlling for health
Partial Correlation



- Select *Partial…* under *Correlate* in the tab *Analyze*
- Add the variables Pain before treatment and Age in years to the *Variables* box
- Add the General health variable to the *Controlling for* box
- Optional: Click on *Options...* to add statistics to the output, then click on *Continue*
- Click *OK*

RQ6: Association of pain & age, controlling for health
Partial Correlation

Partial Corr

Descriptive Statistics

	Mean	Std. Deviation	N
Pain before treatment	7.0167	1.46706	60
Age in years	54.2333	10.76461	60
General health	2.3500	.73242	60

Correlations

Control Variables			Pain before treatment	Age in years
General health	Pain before treatment	Correlation	1.000	.511
		Significance (2-tailed)	.	.000
		df	0	57
	Age in years	Correlation	.511	1.000
		Significance (2-tailed)	.000	.
		df	57	0

- SPSS returns two tables, the first shows descriptive statistics, the second presents the correlation results
- The first row of the correlations table shows the correlation coefficient, it indicates the direction and strength of the association
- The second row *Significance (2-tailed)* shows the p-value
- The third row *df* shows the sample size for each calculation
- In this case, we find a significant positive relationship between age and pain, while controlling for general health

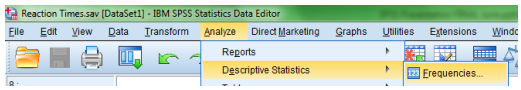
RQ 7: Reaction time to emotional facial expressions

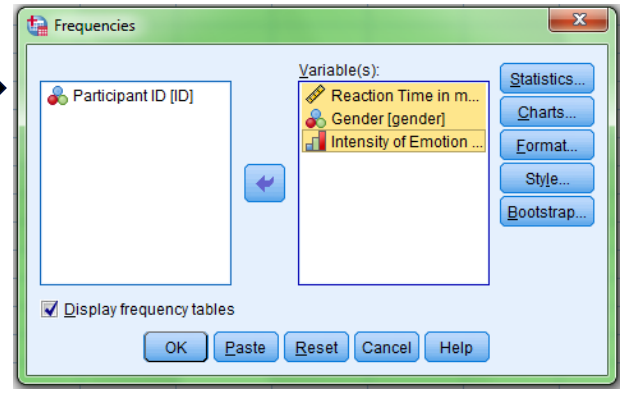
- Data set of 141 participants: reaction time study
- Participants had to identify emotional expressions of different intensities, reaction times were measured
- 3 variables: gender, reaction time and intensity of emotion

☛ Do gender and emotion intensity predict reaction time?

- Getting to know the data set: Frequencies
- Boxplot: Finding outliers (less than 250ms/less or more than 1.5 IQR away from upper or lower quartile)
- Sorting and Selecting Cases
- Linear regression, outcome: reaction time, predictors: emotion intensity and gender

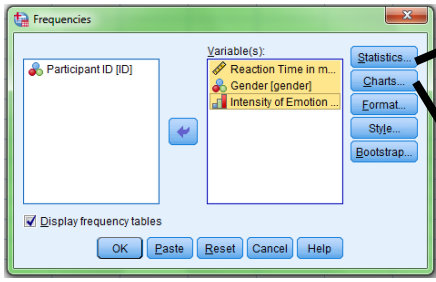
RQ 7: Reaction time to emotional facial expressions
Frequencies

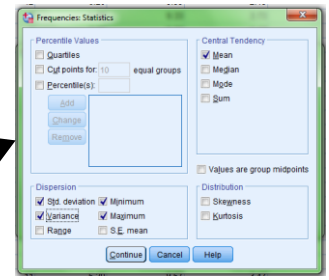


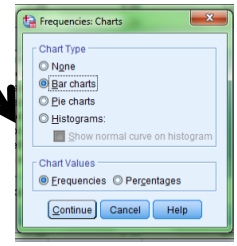


- Select *Frequencies* under *Descriptive Statistics* in the tab *Analyze*
- Double-click on the variables to add them to the box
- Open options for *Statistics* and *Charts* by clicking on the respective buttons (see next slide)
- When finished with selecting options, click *OK*

RQ 7: Reaction time to emotional facial expressions
Frequencies



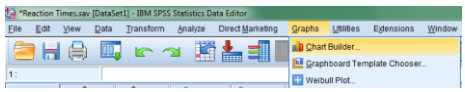




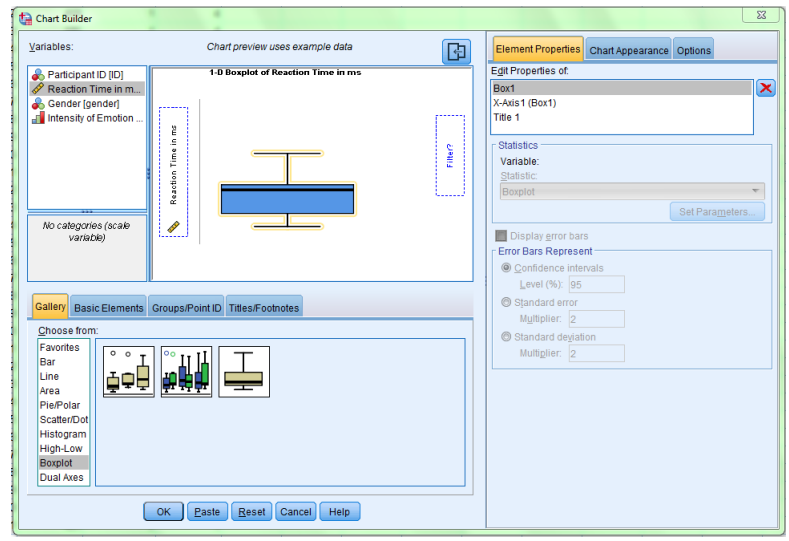
- Select the statistics you want to receive for the selected variables
- Click *Continue*

- Select the type of chart you want to receive for the chosen variables
- Choose whether the values should be presented as frequencies or percentages
- Click *Continue*

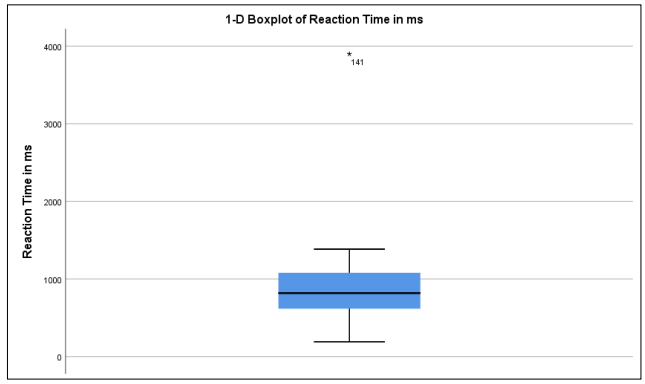
RQ 7: Reaction time to emotional facial expressions
Boxplot



- Select *Chart Builder* from the tab *Graphs*
- In the bottom section, choose the category *Boxplot*
- Double-click on the boxplot chart type on the right
- From the list of variables, drag the variable Reaction time to the y-axis
- Click *OK*

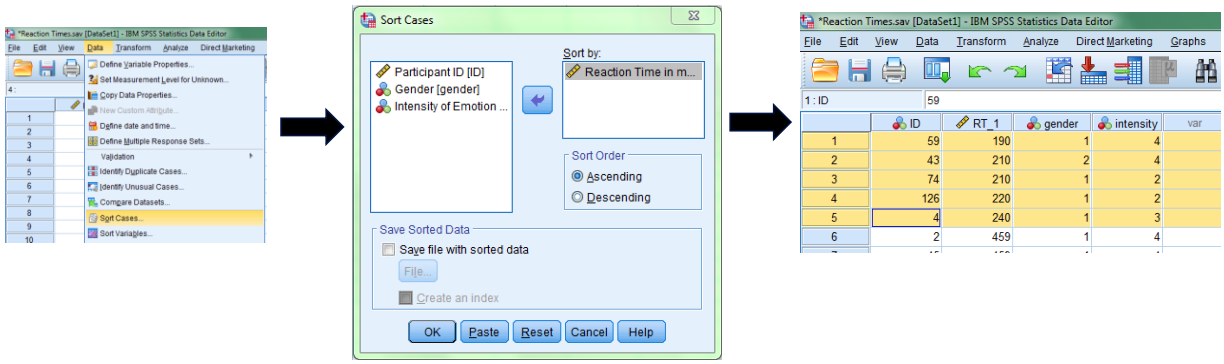


RQ 7: Reaction time to emotional facial expressions
Boxplot



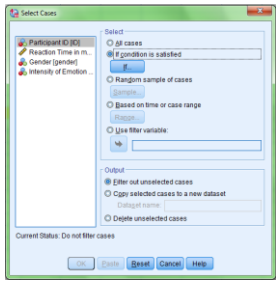
The boxplot shows there is 1 outlier in the data, around 4000ms

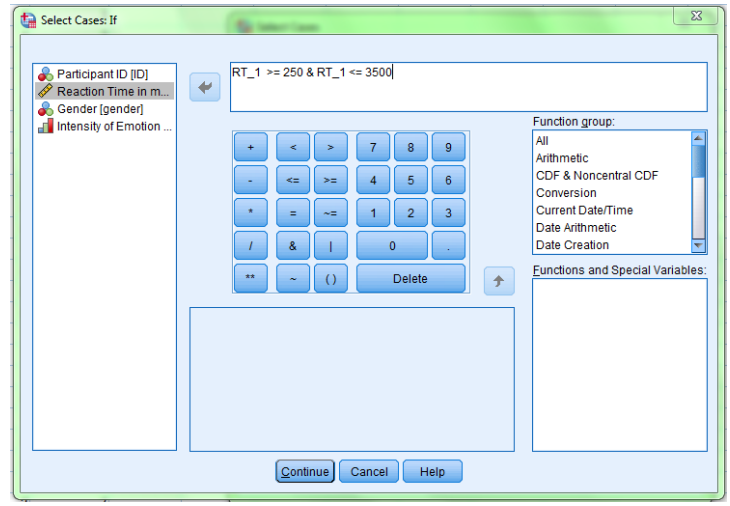
RQ 7: Reaction time to emotional facial expressions
Sorting Cases



- Select *Sort Cases…* in the tab *Data*
- Add the variable Reaction Time to the *Sort by* box and select ascending sort order
- Click *OK*
- The data set should now be sorted by reaction time, we can see that 5 data points are below 250ms → outliers

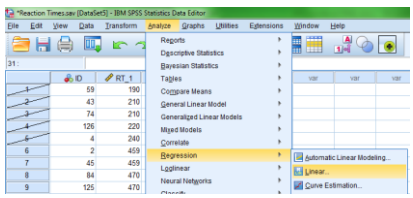
RQ 7: Reaction time to emotional facial expressions
Selecting Cases

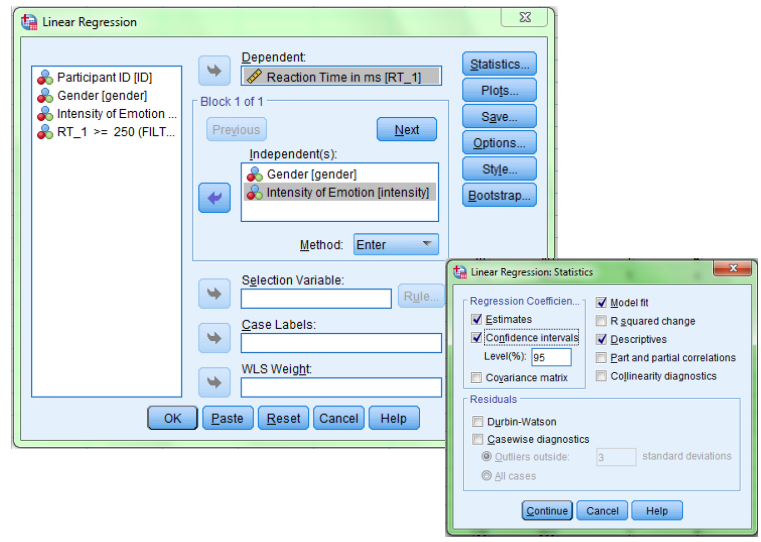




- Select *Select Cases…* in the tab *Data*
- Select the option *If condition is satisfied* and click on *If…* to define the condition
- Add the Reaction time variable to the numeric expression box and define that the value should be >=250 & <=3500 to exclude all outliers, then click *Continue*
- Click *OK*

RQ 7: Reaction time to emotional facial expressions
Linear regression





- Select *Linear…* under *Regression* in the tab *Analyze*
- Add the Reaction Time variable to the *Dependent* box and the variables Gender and Intensity of Emotion to the Independent variables box
- Click on *Statistics…* and select *Estimates, Confidence Intervals, Model fit* and *Descriptives*, then click Continue
- Click *OK*

RQ 7: Reaction time to emotional facial expressions
Linear regression

Descriptive Statistics

	Mean	Std. Deviation	N
Reaction Time in ms	857.57	257.368	135
Gender	1.53	.501	135
Intensity of Emotion	2.61	1.099	135

Correlations

		Reaction Time in ms	Gender	Intensity of Emotion
Pearson Correlation	Reaction Time in ms	1.000	.861	-.964
	Gender	.861	1.000	-.849
	Intensity of Emotion	-.964	-.849	1.000
Sig. (1-tailed)	Reaction Time in ms	.	.000	.000
	Gender	.000	.	.000
	Intensity of Emotion	.000	.000	.
N	Reaction Time in ms	135	135	135
	Gender	135	135	135
	Intensity of Emotion	135	135	135

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.967[a]	.936	.935	65.846

a. Predictors: (Constant), Intensity of Emotion, Gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8303655.547	2	4151827.773	957.599	.000[b]
	Residual	572307.535	132	4335.663		
	Total	8875963.081	134			

a. Dependent Variable: Reaction Time in ms

b. Predictors: (Constant), Intensity of Emotion, Gender

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound
1	(Constant)	1246.249	56.364		22.111	.000	1134.756	1357.742
	Gender	79.316	21.448	.154	3.698	.000	36.889	121.742
	Intensity of Emotion	-194.931	9.777	-.833	-19.937	.000	-214.272	-175.590

a. Dependent Variable: Reaction Time in ms

- SPSS returns several tables, the first shows descriptive statistics, the second presents correlation results, after that the model summary and regression results are listed
- The first row of the *ANOVA* table shows whether the overall regression is significant
- The second row and third row of the Coefficients table show the results for the two predictors, the *Standardized Coefficients Beta* indicates direction and strength of the effect
- In this case, we find that the overall model is significant. Moreover, both gender and intensity of emotion significantly predict reaction time

RQ 8: Effect of supermarket promotion on sales

- Data from 133 supermarkets, effect of supermarket promotion on sales
- 3 promotion conditions: control, price reduction, price reduction + advertisement
- 2 variables: condition of promotion, average daily sale

☛ Did the two types of promotion have an effect on sales?

- Duplicate cases
- One-way ANOVA: Did the two types of promotion have an effect on sales?