



THE UNIVERSITY OF  
**NEWCASTLE**  
AUSTRALIA

# How to analyse my data

## 4- 6 July 2018

---

### **Outlines**

- **Exploratory data analysis and visualising data**
  - **Formulating research questions**
  - **Data types and related statistical tests**
  - **How to interpret statistical results**
- 
- ◆ **Explanation of common statistical tests**
  - ◆ **Workbook with worked examples then hands on practice**
  - ◆ **Use statistical software to create output (SPSS)**
  - ◆ **SPSS software guide provided**
  - ◆ **Focus on understanding, concepts and interpretation of results**

**Instructor**

**Nic Croce**

**Statistical Support Service**



THE UNIVERSITY OF  
**NEWCASTLE**  
AUSTRALIA

Statistical Support Service

Statistics refresher seminar series

**What method do I use?**

13-June-2018

Presenter: Fran Baker

(Notes thanks to Kim Colyvas)

[statsupport@newcastle.edu.au](mailto:statsupport@newcastle.edu.au)

**Notes for this Seminar can be downloaded from:**

<http://www.newcastle.edu.au/about-uon/governance-and-leadership/faculties-and-schools/faculty-of-science-and-information-technology/resources/statistical-support-services>

**\*or search the UoN website for StatSS**

# What Method do I use?

1. The Research Question
2. The Variable Types – numeric / categorical
3. Commonly used Analysis Methods
  - Chi squared test
  - McNemar Test
  - Paired (dependent) t-test
  - Independent samples t-test
  - ANOVA (repeated measures)
  - Simple/Multiple Linear Regression
  - Factor Analysis



# Research questions

- When framing research questions they begin generally.
- The choice of a statistical method however is made easier by refining the question and making it specific.
- Eg Example 3 from later in the notes...  
‘We want to explore the idea that in children, exercise and medication can help with their asthma.’



# Specific research questions

<http://libguides.lmu.edu/content.php?pid=10084&sid=1965230>

‘We want to explore the idea that in children, exercise and medication can help with their asthma.’



# Specific research questions

<http://libguides.lmu.edu/content.php?pid=10084&sid=1965230>

‘We want to explore the idea that in children, exercise and medication can help with their asthma.’

## **A general version**

Do exercise and/or medication benefit asthmatic children?



# Specific research questions

<http://libguides.lmu.edu/content.php?pid=10084&sid=1965230>

‘We want to explore the idea that in children, exercise and medication can help with their asthma.’

## **A general version**

Do exercise and/or medication benefit asthmatic children?

**More specific** - measurable/countable variables

Does **mean** **Forced Expiratory Volume (1 sec)** **increase** by **increasing** exercise (...) and **increasing** medication (...) in asthmatic children **aged 7-12 yrs**?



# The Question determines the analysis

## Response -Y- (outcome, dependent) variable(s)

- List variable(s) measured on experimental unit.
- What is/are the response variable type(s)?
- If categorical how many levels within the category?
- If multiple response variables are they correlated?

## Explanatory -X- (predictor, independent) variable(s)

- List variables.
- What are variable types?
- Number of groups (samples) being studied (2 or multiple).
- Relationship between samples: Independent/Dependent
- Are multiple explanatory variables correlated?

## Some goals of analysis

- Are there differences due to a given variable(s)?
- Relationships – describe, predict effect of a variable(s)
- Determine the relative importance of variables.
- Create new variables or groupings



# Variable types → ★ key ★

- **Numeric** – measures and counts

- **Continuous** : temperature, weight, speed, distance  
(*Similar terms include scale and ratio*)
- **Discrete** : # of defects, result of die toss, item count

- **Categorical** – values based on categories

- **Ordinal** : Grades - FF, P, C, D, HD,  
\*Temperature - Low, Medium, High
- **Nominal** : gender – male/female  
colour - blue/green/yellow

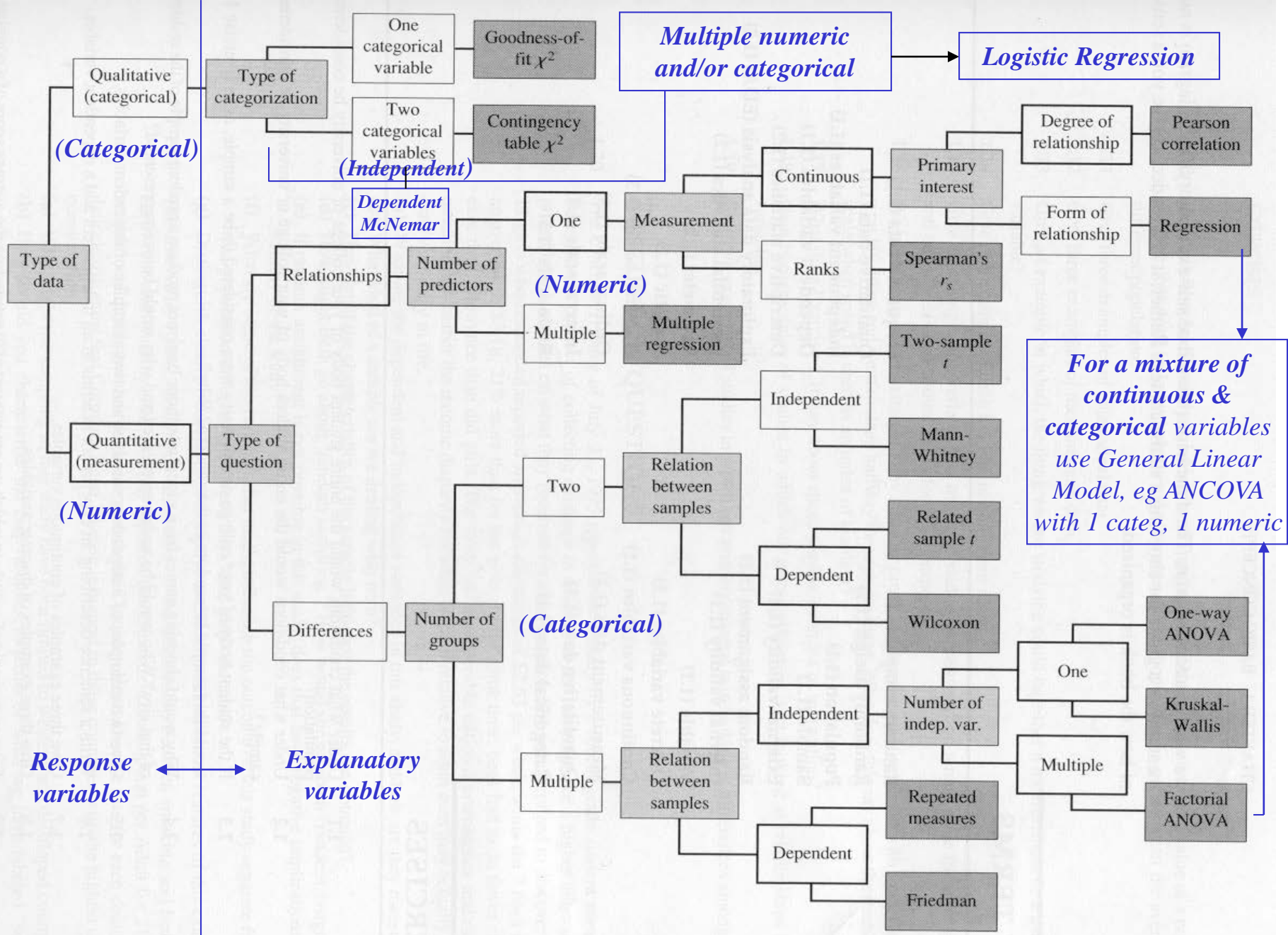


Diagram from Howell, D.C, Statistical Methods for Psychology, 5<sup>th</sup> Ed, 2001, p11, 11  
 Duxbury, Thomson Learning with additions *(in blue and italics)* by K. Colyvas (version 5-Jul-05).

## Method selection table (a)

Response Y	Explanatory X or A	Specific question(s)	Displays	Statistical method
<b>Singles samples</b>				
Categorical		How is data distributed?	Table Bar chart Pareto	Tables of %'s in categories, n
Numerical		How is data distributed?	Histogram	Tables of summary statistics, eg mean, sd, min, max, quartiles, n.

## Method selection table (b)

**See next slide**

<b>Y Response</b>	<b>X Explanatory</b>	<b>Specific question(s)</b>	<b>Displays</b>	<b>Statistical method</b>
Categorical	Categorical	How do proportions in response depend on <b>the levels of the explanatory variable?</b>	Tables	Chi-squared statistic
Categorical	Numerical	How does the proportion in response depend on the <b>values of the explanatory variable?</b>	Tables (X groups)	<i>Logistic regression</i> <i>Correlation (for a binary response only)</i>
Numerical	Categorical	How does mean level in response change with <b>the levels of the explanatory variable? If so how does it vary?</b>	Box plots Mean plots CI plots	t test (2 indep groups) ← ANOVA (3 or more groups)
Numerical	Numerical	How does mean level of response change with <b>values of the explanatory variable.</b>	Scatter plots	Correlation Regression
<b>Dependent (or paired) samples</b>				
Categorical	Categorical	Is there agreement between the matched levels? Is lack of agreement biased? Is there a difference in proportions?	Tables	Kappa (2x2 Agreement) McNemar's test (2x2 - bias in agreement or difference in proportions)
Numerical	Categorical	How does mean level in response change with the levels of the explanatory variable <b>WITHIN</b> e.g. subject	Box plots Mean plots Within CI plots	Paired t test ← Repeated measures ANOVA

**Multiple response variables  
(Multivariate)**

**Specific question(s)**

How do I reduce the  
number of variables?

Can the variables be  
categorised into  
cluster groups?

**Statistical Method**

Principal component  
analysis &  
Factor analysis.

Cluster analysis.

**NB No explanatory variables – only outcomes**

# How to determine the analysis method

## Put the info on the 'desk' – DSQ

### Data - Samples - Question

- **D:** variable types matter. Is the data counted or measured → **numerical** or in **categories**?
- **S:** how many samples, groups or categories? Are they **independent** or **dependent**?
- **Q:** is question about **independence**, **association** or **relationships**? Or are you being asked to **compare** something or find a **difference**?



# Example 1

*Is there a difference between topical treatments used on superficial sites in clearing MRSA (methicillin-resistant Staphylococcus aureus) infections within 14 days for hospitalised patients?*

**Response** Clear of MRSA in 14 days - Yes/No

Type ?

**Explanatory** Standard treatment, Tea tree oil

Type?

**Imagine research objective with a table or graph**

Treatment	Free of MRSA infection after 14 days		Total	% Clear
	No	Yes		
Standard	30	14	44	32%
Tea tree oil	26	32	58	55%
Total	56	46	102	45%

**Method?**





# Example 1 test

*Is there a difference between topical treatments used on superficial sites in clearing MRSA (methicillin-resistant Staphylococcus aureus) infections within 14 days for hospitalised patients?*

## Response variables

Type(s) Clear of MRSA in 14 days - Yes/No- **Categorical**

## Explanatory variables

Type(s) Topical Treatment – standard /tea tree oil - **Categorical**

Dependent? **Independent**

Imagine the result with a table or graph

Treatment	Free of MRSA infection after 14 days		Total	% Clear
	No	Yes		
Standard	30	14	44	32%
Tea tree oil	26	32	58	55%
Total	56	46	102	45%

Method - **Chi-squared statistic**

# Dependent (paired/matched) samples

- When repeated observations are made on
  - the same person, eg two different drugs to reduce blood pressure,
  - process or object, eg two different manufacturing lines, or two measuring devices on a sample of stream water, blood etc,
  - includes measurements closely spaced in time - ie longitudinal studies with short rather than long intervals in seeing patients **or** spatially – eg sun skin sun damage will be similar on different sites on a forearm, but very different to underarm areas
- Measurements on matched pairs, e.g. in a case/control study people matched on age and gender, or identical twins (genetic matching).
- \* Two measurements made on the same person/object/closely spaced in time or spatially are *likely to be closely related*.  
The response is likely to be similarly affected so we should NOT treat the measurements as independent.



# Example 2

*A poll of a random sample of 1600 British voters was taken and 944 indicated approval. Six months later they SAME 1600 people were asked again and 880 indicated approval.*

## Response & Type

## Explanatory & Type

Dependent?

(If so on what?)

First Survey	Second Survey		Total
	Approve	Disapprove	
Approve	794	150	944
Disapprove	86	570	656
total	880	720	1600

**Imagine research objective with a table or graph**

## Method

	First survey	Second survey
% Approve	59%	55%
	(944/1600)	(880/1600)

# Example 2 test

A poll of a random sample of 1600 British voters was taken and 944 indicated approval. Six months later the **SAME** 1600 people were asked again and 880 indicated approval.

**Response & Type - vote (A or D) categorical**

**Explanatory & Type - vote (A or D) categorical**

**Dependent**

(If so on what?)

**Same people**

**Imagine research objective with a table or graph**

First Survey	Second Survey		Total
	Approve	Disapprove	
Approve	794	150	944
Disapprove	86	570	656
total	880	720	1600

**Method**

**McNemar**

	First survey	Second survey
% Approve	59%	55%
	(944/1600)	(880/1600)



# Exercise 3: Paired or unpaired samples?

*a) A manufacturer has several hundred employees manufacturing ergonomic desks. Two new designs are being evaluated with respect to manufacturing time as all other the costs of production are similar. A random sample of 20 workers are drawn with the first 10 requested to manufacture desk type A and the second 10 workers to manufacture desk type B.*

What is the response? \_\_\_\_\_

What is the factor? \_\_\_\_\_

Are the two samples dependent or independent? \_\_\_\_\_

*b) Consider the situation above but now have all 20 people manufacturing desk type A and then the same 20 people manufacturing desk type B.*

Are the two samples dependent or independent? \_\_\_\_\_



# Exercise 3: Paired or unpaired samples?

*a) A manufacturer has several hundred employees manufacturing ergonomic desks. Two new designs are being evaluated with respect to manufacturing **time** as all other the costs of production are similar. A random sample of 20 workers are drawn with the first 10 requested to manufacture desk type A and the second 10 workers to manufacture desk type B.*

What is the response? **Time**

What is the factor? **Type**

Are the two samples dependent or independent? **Independent**

*b) Consider the situation above but now have all 20 people manufacturing desk type A and then the **same 20 people** manufacturing desk type B.*

Are the two samples dependent or independent? **Dependent**



# Imagine the results – 3b. dependent study

Assume mean time for A = 60 mins, typical variability **between** workers is  $\pm 10$  mins. Average improvement of B relative to A is 2

	Desk Type		
Worker	A	B	A-B
1			
2			
3			
4			
⋮			
20			

Mean  
Sd


# Paired or unpaired

## a. 2 sample t-test

Calculates the differences between the 2 sample means

## b. Paired t-test

Calculates the mean of the differences between pairs

\* Removal of the effect of variability between people allows the focus to be on the difference between the manufacture of the desk types. This gives more power.



# Example 3 – method comparisons

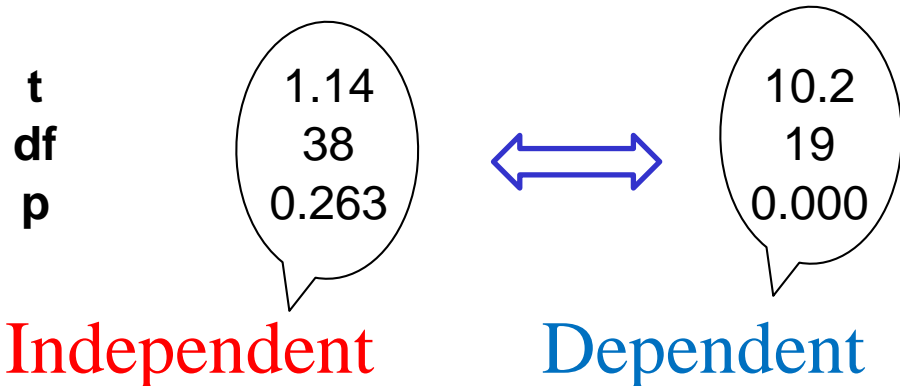
Using the two study designs from the paired/unpaired exercise

Worker	Desk Type		A-B
	A	B	
1	56.5	55.2	1.3
2	54.7	50.6	4.1
3	71.4	68.9	2.5
4	52.9	50.4	2.5
20	57.6	53.6	4.0

Imagine research objective with a table or graph

			<b>Mean diff</b>
<b>Mean</b>	59.7	57.4	2.27
<b>sd</b>	6.3	6.3	1.00
<b>diff of means</b>	<b>2.26</b>		

## Methods





# Obvious!

Easy to see that the paired t-test is better so why then do we see the 2 sample t-test being used???

## Study design!

=> Time and money – not always possible to set up the optimal study. Difficult to gain

- repeated observations on people/objects
- or closely spaced time/spatially
- or matched pairs / case - control

# Example 4

*A physiotherapist treated 17 patients with dizziness (diagnosed as cervical in origin) using sustained natural apophyseal glide (SNAG). A placebo group (n=17) was used for comparison. A Dizziness Handicap Inventory (DHI - larger is worse) was measured on both groups at baseline, post treatment and 6 wks and 12 wks post treatment.*

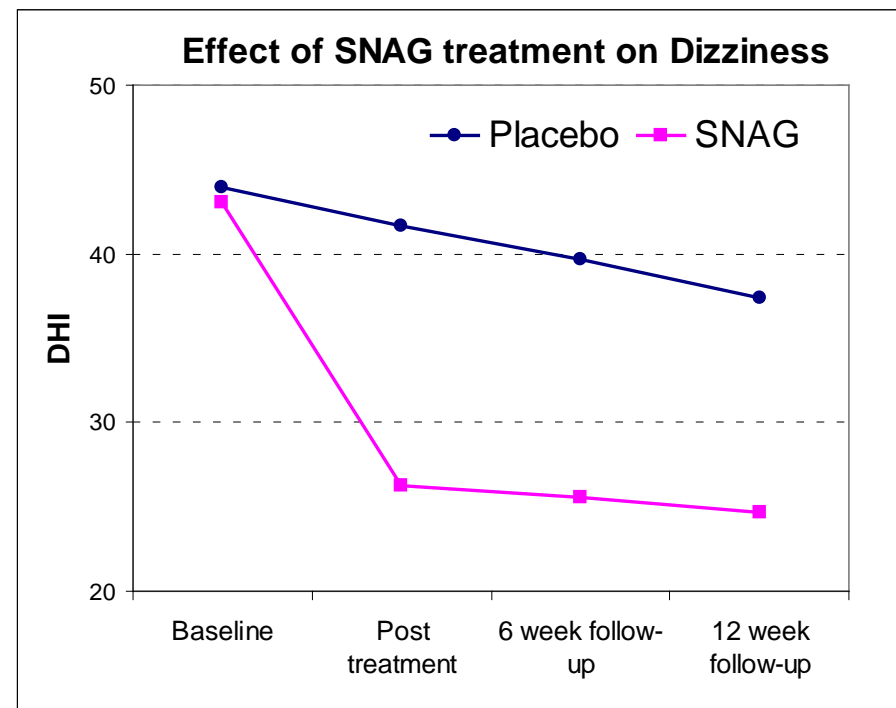
**Response & Type**

**Explanatory & Type**

Dependent? (On what?)

**Imagine research  
Objective**

**Method**



# Example 4 test

A physiotherapist treated 17 patients with dizziness (diagnosed as cervical in origin) using sustained natural apophyseal glide (SNAG). A placebo group (n=17) was used for *comparison*. A Dizziness Handicap Inventory (DHI - larger is worse) was measured on both groups at baseline, post treatment and 6 wks and 12 wks post treatment.

**Response & Type – DHI continuous/numerical**

**Explanatory & Type – treatment (2) + time (4) categorical**

**Dependent** (On what?)

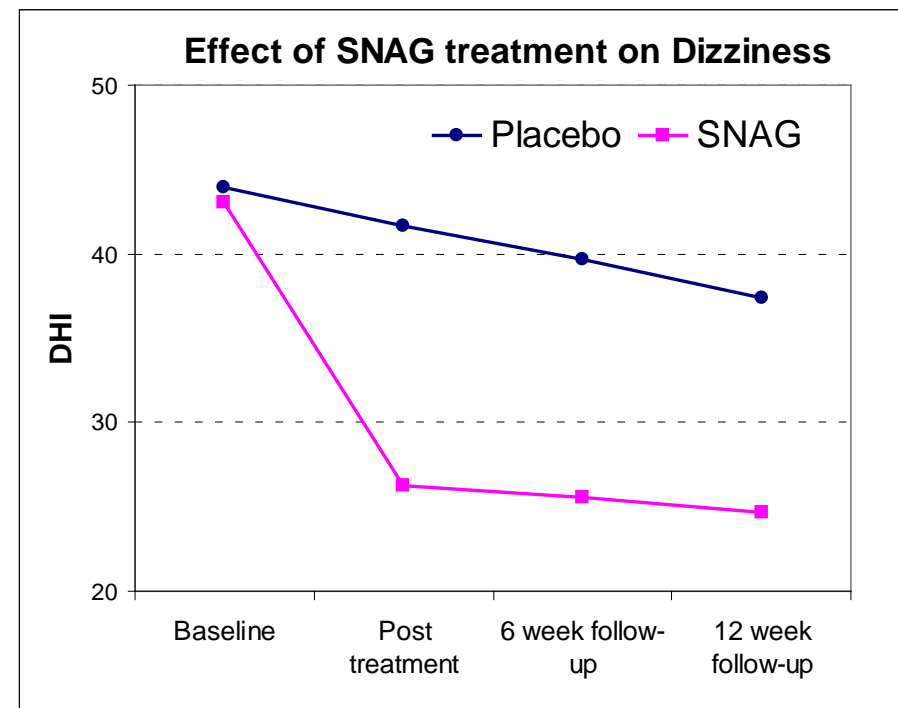
time – dependent on person

**Imagine research**

**Objective**

**Method**

**Repeated measures ANOVA**



# Example 5

*A large chain of motor inns that focuses on the business customer is looking for a better way to choose new locations for establishing a new motel. Data from 100 randomly chosen locations has been collected. The objective is to predict which sites are most profitable (Operating margin) using the variables total number motel + hotel rooms within 3 miles, number of miles to nearest competition, median household income, office space in surrounding community, distance to CBD.*

## Response

## Explanatory

Dependent?

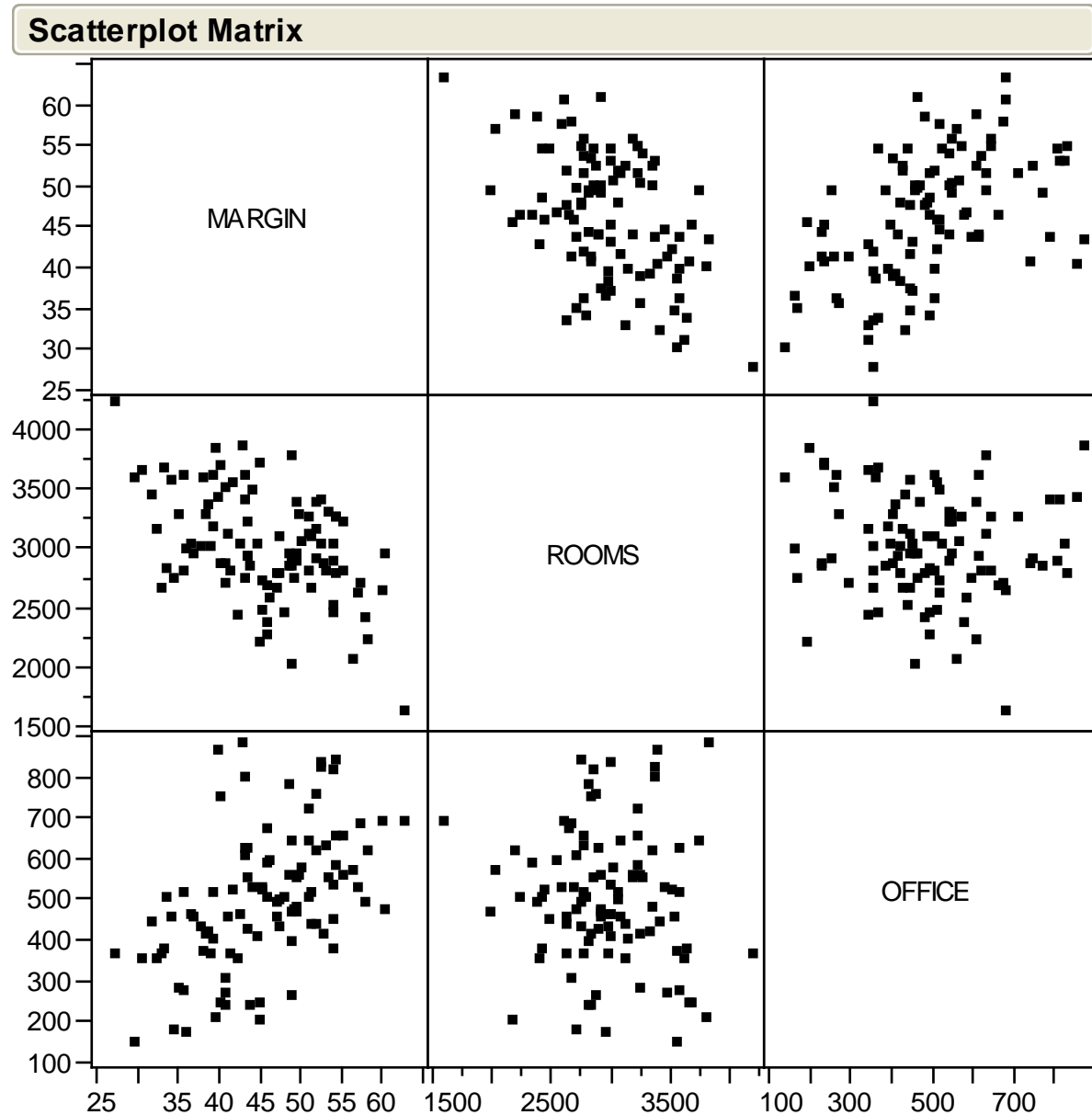
## Correlation matrix

Imagine graph  
and/or table

Method

	MARGIN	ROOMS	OFFICE
MARGIN	1.00	-0.47	0.50
ROOMS	-0.47	1.00	-0.09
OFFICE	0.50	-0.09	1.00

# Eg 5 scatterplot - correlation





# Example 5 test

relationship

*A large chain of motor inns that focuses on the business customer is looking for a better way to choose new locations for establishing a new motel. Data from 100 randomly chosen locations has been collected. The objective is to **predict** which **sites are most profitable** (Operating margin) using the variables total number motel + hotel rooms within 3 miles, number of miles to nearest competition, median household income, office space in surrounding community, distance to CBD.*

**Response - operating margin - numerical**

**Explanatory - motels, rooms, miles, income, office, CBD  
-numerical**

Dependent? **Indep - OM measured once**

## Correlation matrix

**Imagine graph  
and/or table**

**Method -**

**Regression**

	MARGIN	ROOMS	OFFICE
MARGIN	1.00	-0.47	0.50
ROOMS	-0.47	1.00	-0.09
OFFICE	0.50	-0.09	1.00

# **My project** - Description, hypothesis, aim etc

## **Response variables**

Type(s)

## **Explanatory variables**

Type(s)

Dependent? (On what)?

**Imagine the  
result with a  
table or graph**

**Method**



Break in lecture  
for 5-10 mins  
for class exercise

# Example 6

*People with back/neck pain were treated either with fake needles, control group (n=27), or real needles (n=25) to assess the effect of acupuncture on back/neck pain. Pain scores (questionnaire: 0=bad, 100=good) were recorded on each patient before and after treatment.*

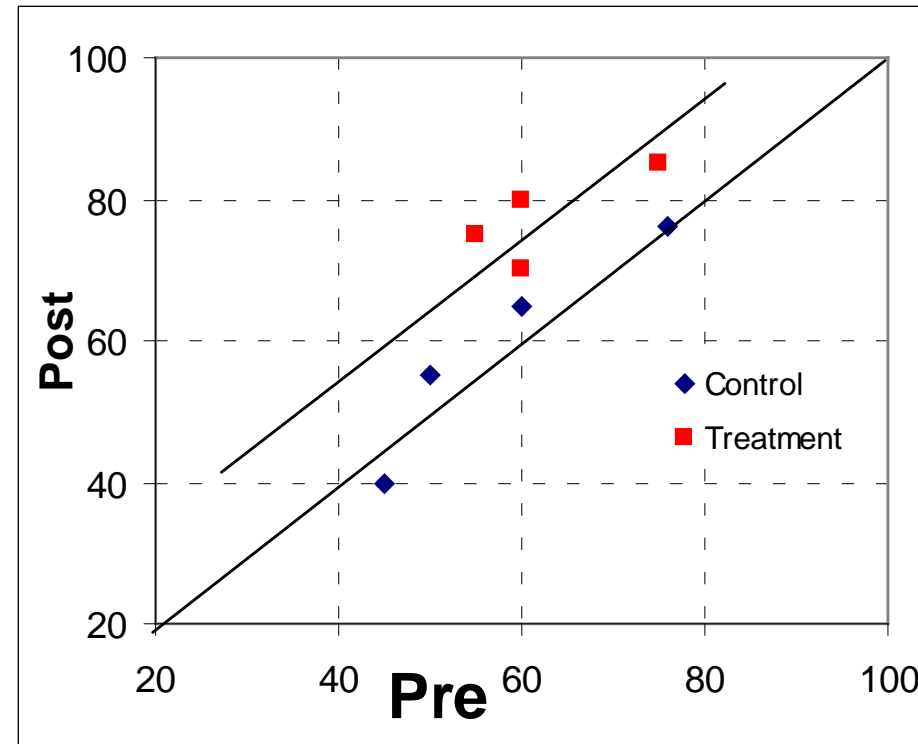
## Response

## Explanatory

Dependent? (On what?)

## Imagine research objective

## Method



# Example 6 test

relationship

*People with back/neck pain were treated either with fake needles, control group (n=27), or real needles (n=25) to assess the effect of acupuncture on back/neck pain. Pain scores (questionnaire: 0=bad, 100=good) were recorded on each patient before and after treatment.*

**Response** pain scores  
numerical

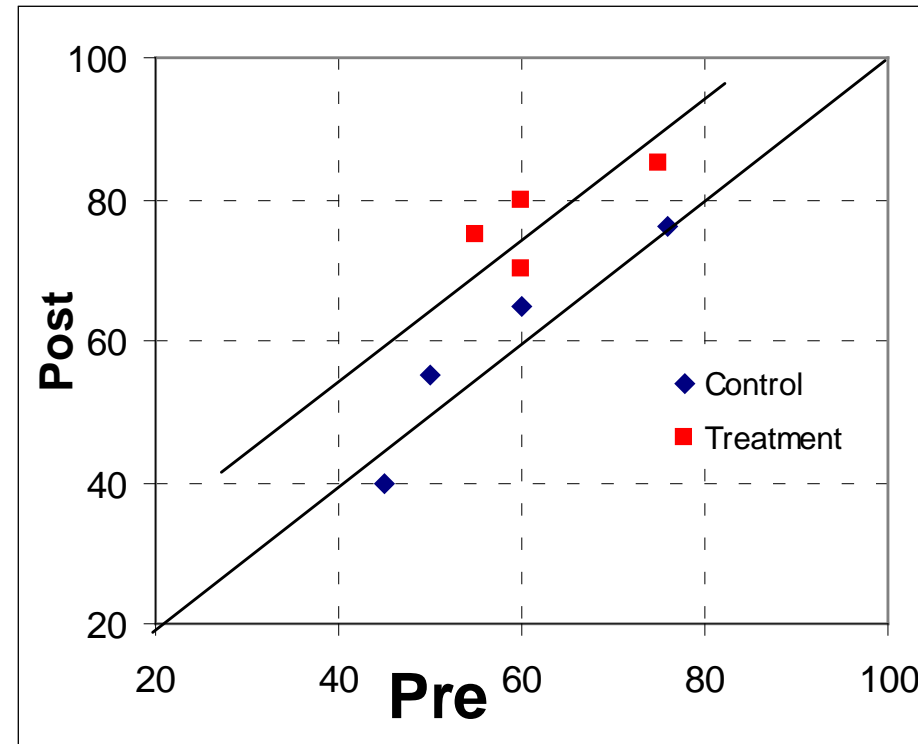
**Explanatory**

Dependent on individuals  
measured 2x (pre/post)

**Imagine research objective**

**Method**

Repeated measures regression



# Example 7

*The relative importance of genetic and environmental factors impacting the incidence of melanoma were examined. Families with high incidence of melanoma (3 or more members with melanoma) were studied by collecting information for each person in the extended family. The status of two suspect genetic mutations (mutation/no mutation) as well as skin colour, extent of sun skin damage, number of moles, number of freckles as other factors were measured.*

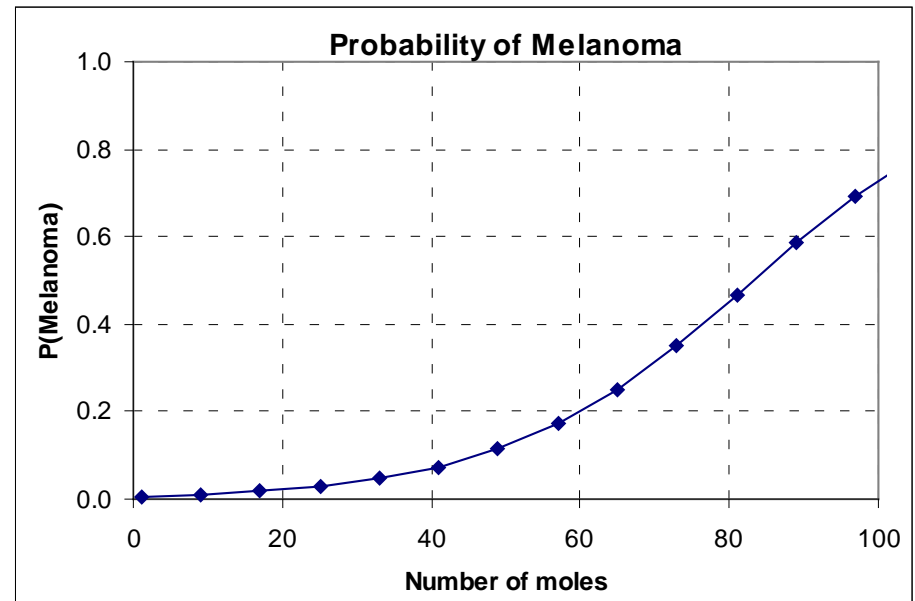
## Response

## Explanatory

Dependent? (on what?)

Imagine graph

## Method



# Example 7 test

relationship

*The relative importance of genetic and environmental factors impacting the incidence of melanoma were examined. Families with high incidence of melanoma (3 or more members with melanoma) were studied by collecting information for each person in the extended family. The status of two suspect genetic mutations (mutation/no mutation) as well as skin colour, extent of sun skin damage, number of moles, number of freckles as other factors were measured.*

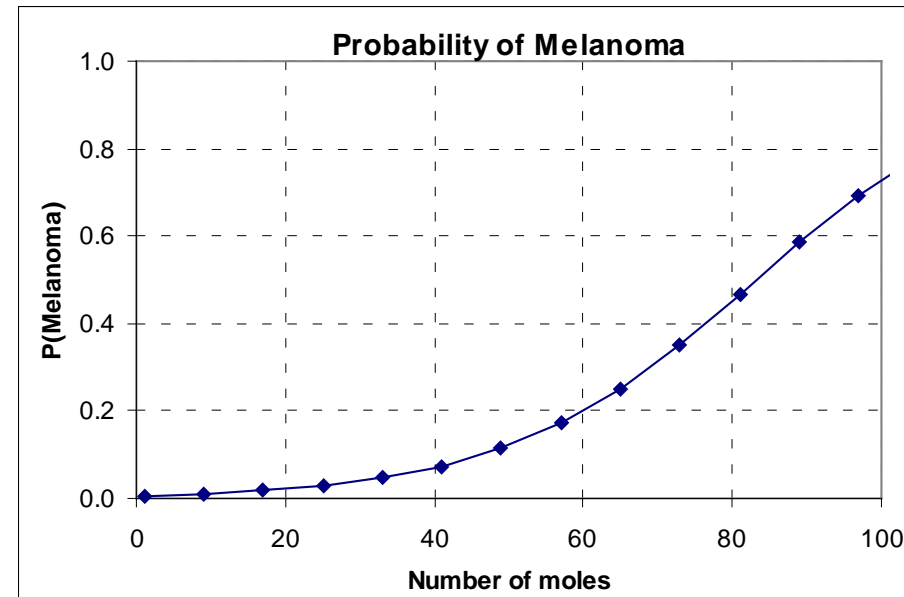
**Response melanoma Y/N categorical**

**Explanatory no. of moles numerical**

Dependent? (on what?) **indep**

**Imagine graph**

**Method logistic regression**



# Multivariate methods

- **Creating new variables (latent variables)**
  - have more desirable properties than an individual variable, eg psychological scales for depression, anxiety and stress,
  - Simplification – fewer variables summarise many
- **Classification (clusters)**
- **Relationships among/between variables**
  - unusual features in data, including outliers,
  - multivariate normality,
  - modelling relationships (SEM)
  - prediction of one group of variables by another group regression of Xs on Ys (PCR, PLS)

# T A B L E 1.1

## Cross-listing of multivariate methods and problem types

Problem type	Multivariate technique								PCR, PLS Sometimes & between variables
	PCA	FA	DA	CDA	CA	MANOVA	CVA	CCA	
Exploring relationships among variables	Sometimes	Definitely	Never	Never	Never	Never	Rarely	Sometimes	Definitely
Screening data	Definitely	Sometimes	Never	Never	Sometimes	Never	Never	Never	Does
Creating new variables	Does	Does	Does not	Does	Does not	Does not	Does	Does	Does
Predicting group membership	Does not	Does not	Does	Does	Does	Does not	Does not	Does not	Does
Comparing group means	Possibly	Possibly	Rarely	Rarely	Does not	Does	Does	Does not	Does
Comparing groups of variables	Possibly	Possibly	Never	Never	Never	Never	Never	Definitely	Definitely
Verifying clusters	Definitely	Possibly	Never	Never	Definitely	Never	Never	Never	Definitely
Reducing dimensionality	Definitely	Definitely	Never	Definitely	Never	Never	Definitely	Definitely	Definitely
Creating meaningful variables	Unlikely	Usually	Never	Possibly	Never	Never	Possibly	Unlikely	Unlikely

from Johnson, D.E, Applied Multivariate Methods for Data Analysis, 1998, Duxbury

PCA – Principal Components Analysis, FA – Factor Analysis, DA – Discriminant Analysis, CDA – Canonical Discriminant Analysis, CA – Cluster Analysis, MANOVA – Multivariate ANOVA, CVA – Canonical Variates Analysis, CCA – Canonical Correlation Analysis, PLS – Partial Least Squares

# Example 8 – Cluster and FA

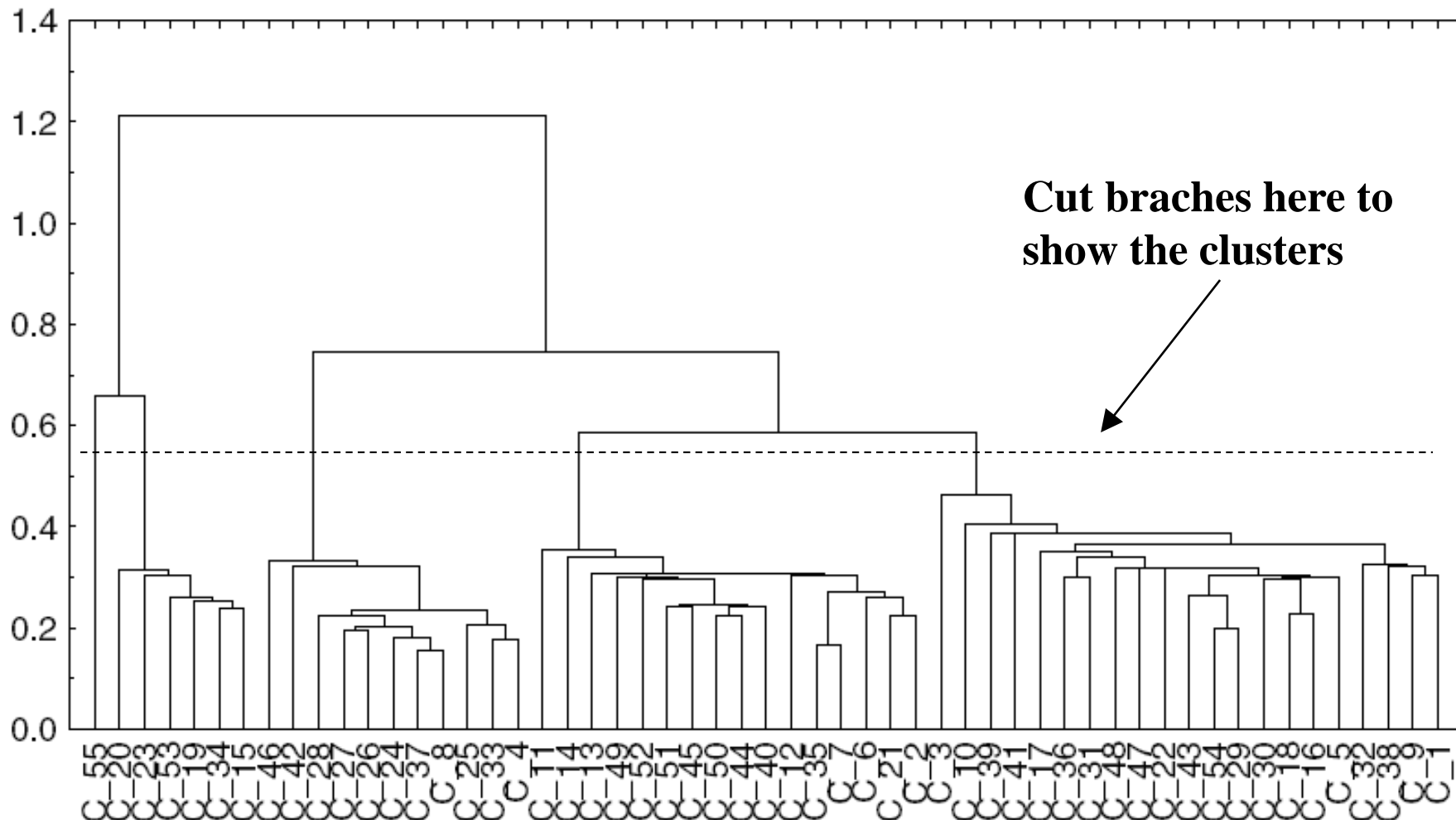
**Abstract** Radioisotope x-ray fluorescence (XRF) analysis has been utilized to determine the elemental composition of 55 archaeological pottery samples by the determination of 15 chemical elements. Fifty-four of them came from the Tel-Alramad site in Katana town, near Damascus city, Syria, and one sample came from Brazil. The XRF results have been processed using two multivariate statistical methods, cluster and factor analysis, in order to determine similarities and correlation between the selected samples based on their elemental composition. The methodology successfully separates the samples where four distinct chemical groups were identified.

E. H. Bakraji, Application of multivariate statistical methods to classify archaeological pottery from Tel-Alramad site, Syria, based on x-ray fluorescence analysis, *X-Ray Spectrom.* 2006; **35**: 190–194



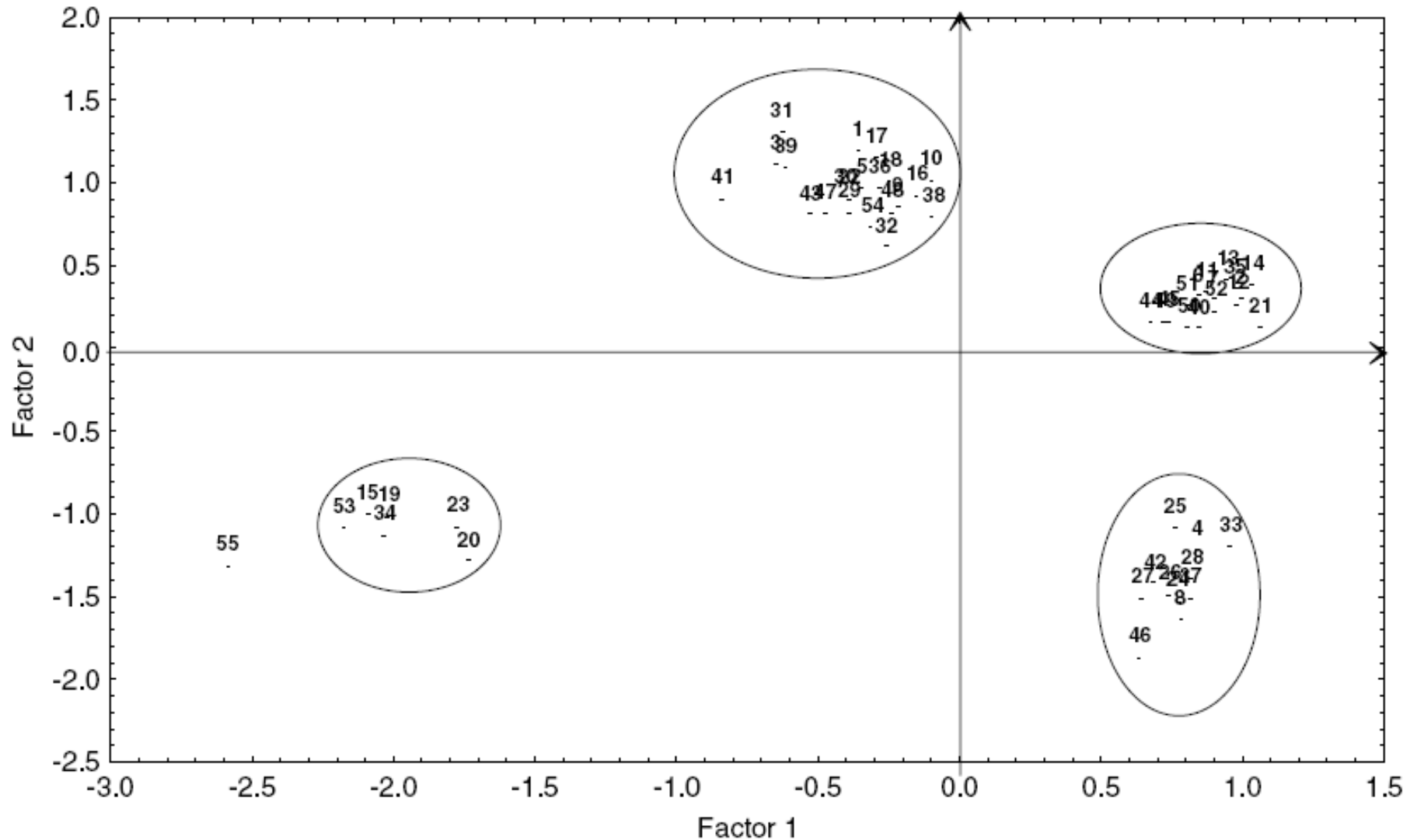


# Clustering – tree diagram – based on distance apart

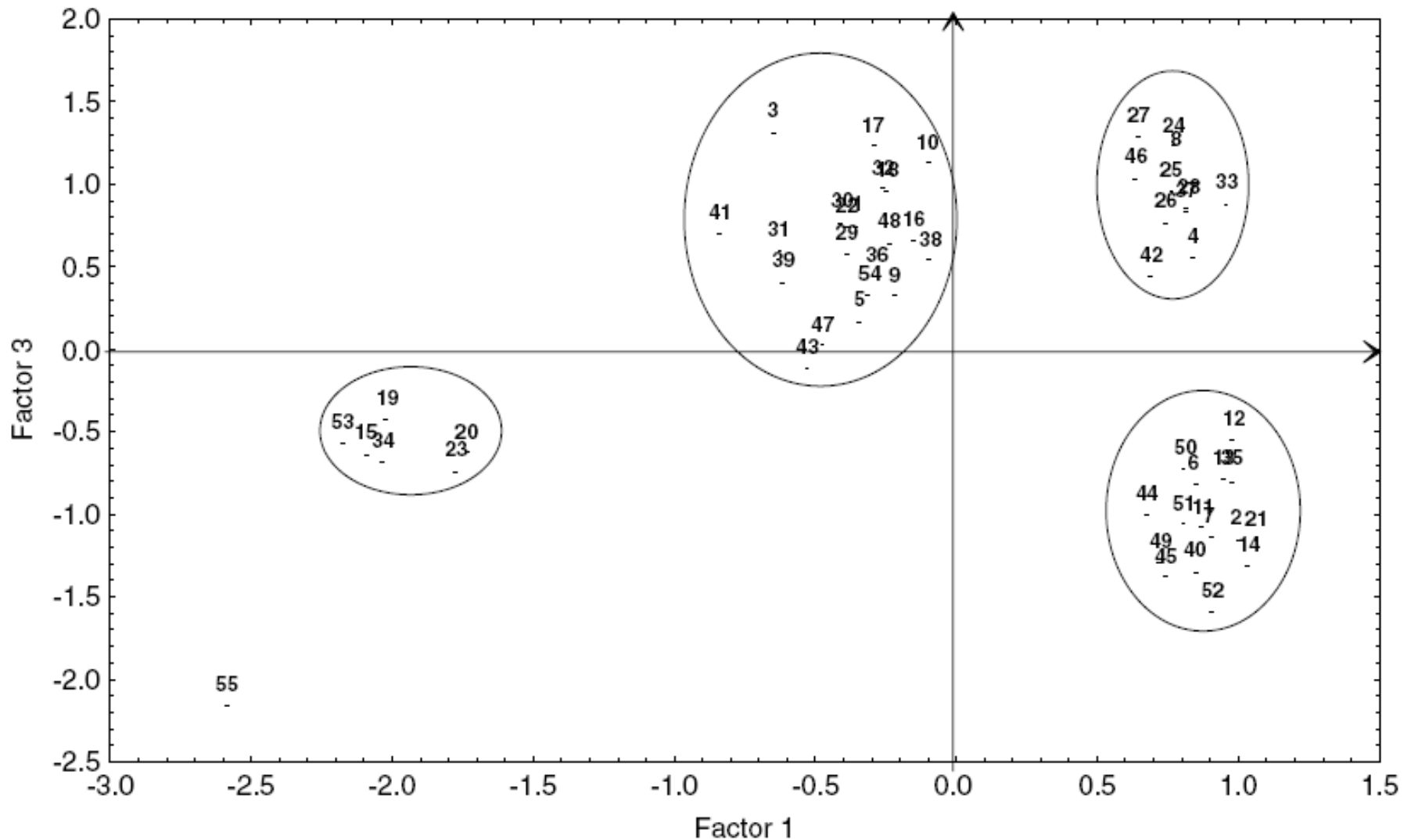


# Clusters based on factor analysis plot of factor scores

## FA1 vs FA 2



# Clusters based on factor analysis plot of factor scores FA1 vs FA 3



# Factor loadings (correlations)

Table 2. Factor loading for the samples data set. Fifteen elements, maximum likelihood factor analysis, varimax rotation

Elements	Factor 1	Factor 2	Factor 3	Communality
Br	-0.21	0.15	0.47	0.29
Ca	0.31	0.92	0.07	0.94
Cu	-0.70	-0.68	0.08	0.95
Fe	0.25	0.30	0.81	0.81
Ga	-0.73	-0.60	-0.20	0.93
K	0.13	-0.03	0.57	0.34
Mn	-0.30	0.28	0.38	0.31
Ni	-0.14	-0.80	-0.39	0.81
Pb	0.08	-0.87	-0.08	0.77
Rb	0.97	0.03	0.17	0.98
Sr	-0.34	0.80	0.30	0.85
Ti	0.02	0.81	0.02	0.65
Y	0.88	-0.14	0.17	0.83
Zn	-0.34	-0.53	-0.50	0.65
Zr	0.93	0.08	-0.11	0.89
Variance explained by factors (%)	36.4	26.3	10.6	73.3

Three factors explain most of the variation (73.3%), i.e. 15 variables simplified to 3.

Loadings identify the most important variables in each factor – the larger the absolute value of the loading the more important the variable.

eg Ga, Cu, Rb, Y and Zr in Factor 1



# Example 9

*A Pizza firm carried out a 15 question employee attitude survey with male and female kitchen staff, delivery drivers, shift supervisors and assistant managers participating.*

*(Next Slide for questions)*

**Response**

**Explanatory**

**Imagine the  
result with a  
table or graph**

**Method**

Table 1.5 Final interpretation of factors for non-supervisory staff

# Create 4 new variables from 15 - Factor Analysis

<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="writing-mode: vertical-rl; transform: rotate(180deg);">1 Company Pride</div> <div style="writing-mode: vertical-rl; transform: rotate(180deg);">2 Lack of commitment</div> <div style="writing-mode: vertical-rl; transform: rotate(180deg);">3 Willingness to work hard</div> <div style="writing-mode: vertical-rl; transform: rotate(180deg);">4 Attitude to employees</div> </div>				Short version of question	Full text of question
1			2 Great Company	2. I TALK UP PIZAZZ TO FRIENDS AS A GREAT COMPANY TO WORK FOR.	
1			4 Any job to stay	4. I WOULD ACCEPT ALMOST ANY TYPE OF JOB IN ORDER TO STAY WITH PIZAZZ.	
1			5 Values similar	5. MY VALUES AND PIZAZZ'S VALUES ARE SIMILAR.	
1			6 Proud tell others	6. I AM PROUD TO TELL OTHERS THAT I WORK FOR PIZAZZ.	
1			8 Inspires best	8. PIZAZZ REALLY INSPIRES ME TO DO MY BEST.	
1			10 Glad chose Pizzazz	10. I AM GLAD THAT I CHOSE PIZAZZ TO WORK FOR.	
1			13 Care about what happens	13. I REALLY CARE ABOUT WHAT HAPPENS TO PIZAZZ.	
1			14 Best company work for	14. FOR ME THIS IS THE BEST COMPANY TO WORK FOR.	
	2		3 Little Loyalty	3. I FEEL VERY LITTLE LOYALTY TO PIZAZZ.	
	2		7 Other Co. work similar	7. I COULD JUST AS WELL BE WORKING FOR ANOTHER COMPANY AS LONG AS THE WORK WAS SIMILAR.	
	2		9 Little change to leave	9. IT WOULD TAKE VERY LITTLE CHANGE IN MY WORK TO CAUSE ME TO LEAVE PIZAZZ.	
	2		11 Not much gained long time	11. THERE'S NOT MUCH TO BE GAINED BY STAYING AT PIZAZZ A LONG TIME.	
	2		15 Work Pizzazz a mistake	15. DECIDING TO WORK FOR PIZAZZ WAS A MISTAKE.	
		3	1 Work Harder	1. I AM WILLING TO WORK HARDER THAN MOST PEOPLE AT PIZAZZ.	
		4	12 Difficult agree attitudes	12. I OFTEN FIND IT DIFFICULT TO AGREE WITH PIZAZZ'S ATTITUDES TO ITS EMPLOYEES.	



# Example 9 test

*A Pizza firm carried out a 15 question employee attitude **survey** with male and female kitchen staff, delivery drivers, shift supervisors and assistant managers participating.*

**Response questions**      **categorical** response to survey

**Explanatory**      **none**

**Imagine the  
result with a  
table or graph**

**Method**      **Factor Analysis**

# Review & considerations – DSQ

## Data - Samples - Question

- Variable types matter – **numerical** or **categorical** :**Data**
- Variable types determine the method of analysis where there is a causal relationship
- Independent or dependent samples :**Samples**
- Multiple response variables – **relationships among the repeated measures**. May lead to variable reduction by grouping together variables by some common factor :**Question**



# Additional resources

- Forming a research question

<http://libguides.lmu.edu/c.php?g=323167&p=2173967>

- Statistics Learning Centre Videos and Resources on youtube

<http://www.youtube.com/watch?v=0r2o2Okpt3A>

# Thank you



THE UNIVERSITY OF  
**NEWCASTLE**  
AUSTRALIA

# How to analyse my data

## 4- 6 July 2018

---

- Outlines**
- **Exploratory data analysis and visualising data**
  - **Formulating research questions**
  - **Data types and related statistical tests**
  - **How to interpret statistical results**
- 
- ◆ **Explanation of common statistical tests**
  - ◆ **Workbook with worked examples then hands on practice**
  - ◆ **Use statistical software to create output (SPSS)**
  - ◆ **SPSS software guide provided**
  - ◆ **Focus on understanding, concepts and interpretation of results**

**Instructors**

**Nic Croce**

**Statistical Support Service**



THE UNIVERSITY OF  
**NEWCASTLE**  
AUSTRALIA

## Statistical Support Service

# Statistical advice, mentoring or analysis

- Free consultation is available for FSCI staff and students
- \$125 per hour for non-FSCI
- Paying? It may sound like a lot but early advice on your research can make it more efficient and effective saving you money, time, effort, frustration, anxiety ... !
- You can also learn how to do the analysis yourself. We can teach or guide you and teach you the software.

**Slide 3 has the StatSS link to the appointment form**