Research Methods

Royal College of Art
Postgraduate Art & Design

Christina Choi, Ph.D.
24 October 2024

Today’s presentation

- Last lecture
  - Background research
  - Identifying potential topics to investigate
- Today’s lecture
  - Research methods and testing
  - Connecting variables to data
  - Analyzing data to answer questions
- Next lecture
  - Defining design research questions
  - Planning activities for your study
  - Engaging users/stakeholders

What is Research?

What is Research

- The systematic collection of data, facts and information for the advancement of knowledge

What are research methods

Strategies and techniques used to gather data that will be analyzed to understand something.

- They are the tools that help you answer some question

- They help you to:
  - Describe actions or situations
  - Interpret events
  - Explore how or why something occurred

- You will need to
  - Have some question to investigate
  - Understand the variables (what you need to measure)
  - Have a way to analyze the data from your measurements (to provide an answer)

Components of a Research Study

Research methods are used within the context of a study setup, so it is important to understand them.

- Groups
- Independent Variable
- Dependent Variable
- Measurement/Observation
- Treatment / Manipulation of variables
- Time

Types of Studies

- Descriptive, Relationship, Differences
  - Descriptive – Observational studies
  - Relationship
    - Correlation designs
    - Regression designs
  - Differences (these are used to support or reject a hypothesis)
    - Simple experimental designs

      (groups assigned using single criteria)
    - Factorial designs

      (groups assigned using multiple-criteria)
    - Time series (single subject) designs

      (dependent variable changes differently over time)

Research Design Notation

- Tic-tac-toe notation:

  R Group1 O X O

  R Group 2 O O

  R = randomly assigned

  O = observation

  X = treatment/variable manipulation

Descriptive Study

- What percentage of people do not wash their hands after going to the bathroom?

Event	1
Group 1	O

- More sophisticated, observe if they wash their hands, then ask how often

Event	1
Group 1 (yes)	O
Group 2 (no)	O

Relationship: Correlation Design

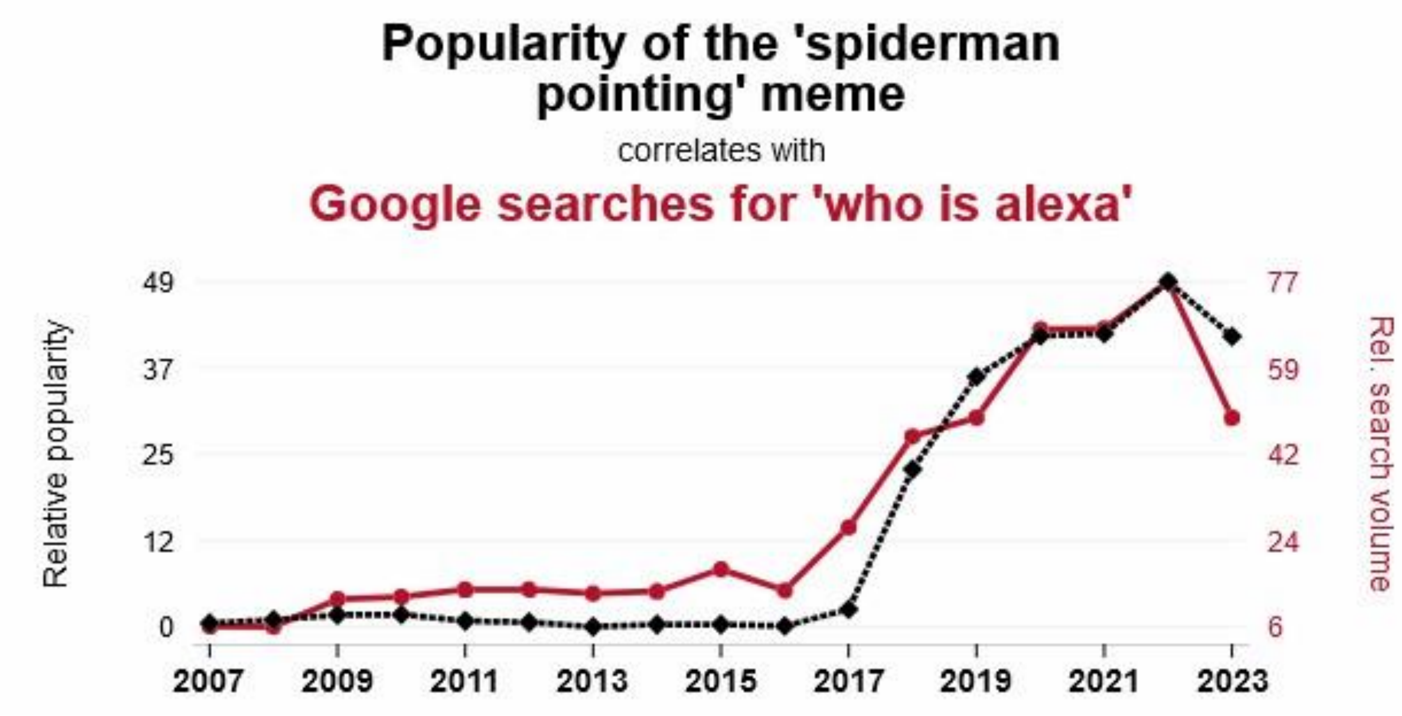
- Correlation designs look for linear dependence between two variables
- A correlation design may consist of many variables, but they are observed all at once

Event	1
R Group 1	O

- May find that one variable increases or decreases relative to another variable
- Does not show cause
- Can be used to identify related variables to be investigated in a Differences study to determine causality

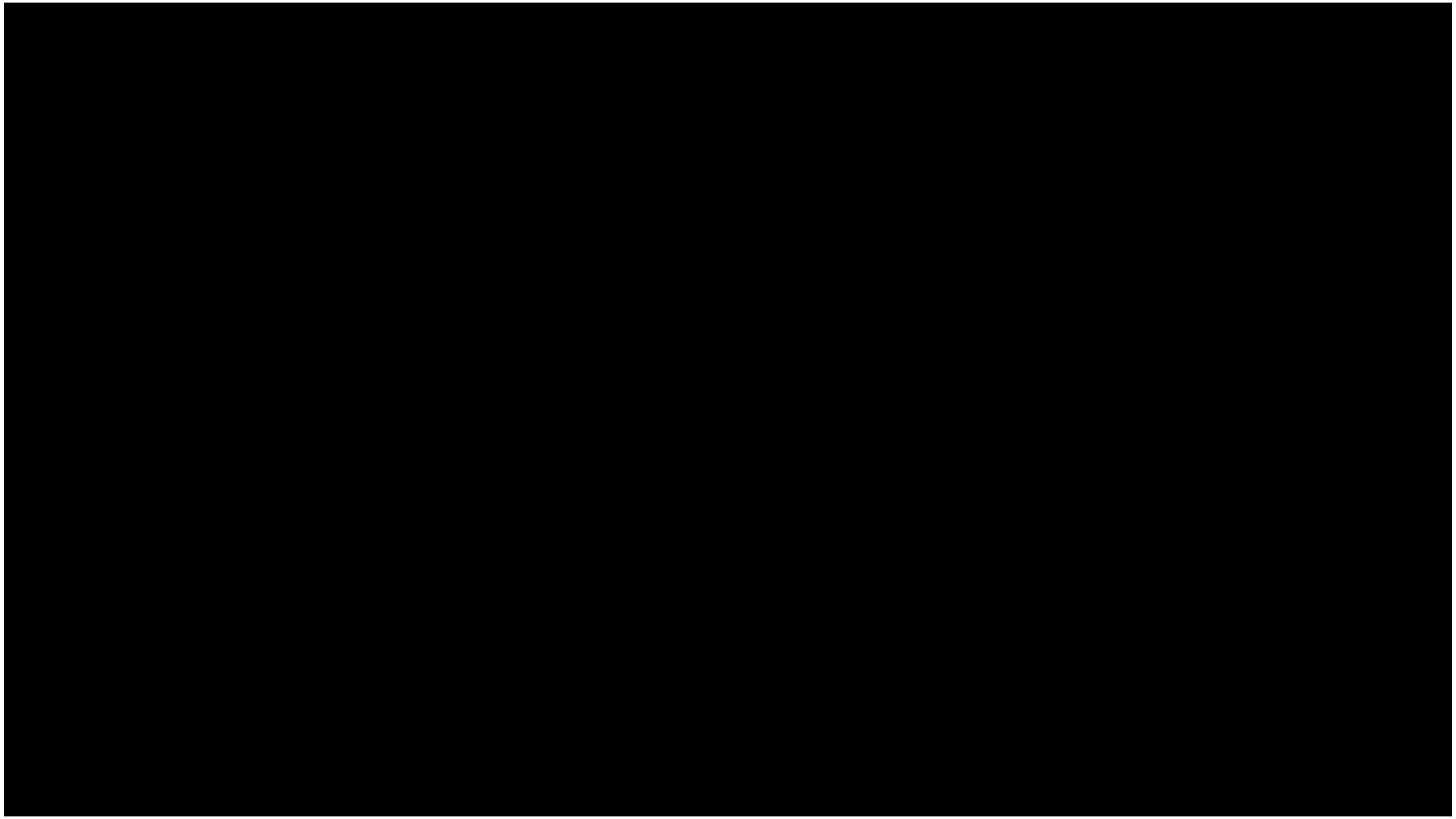
Correlation and Causation

- Correlation can’t tell you why something happened
- Decisions based only on correlation data can be very wrong



https://www.tylervigen.com/spurious-correlations

Ice Cream Kills!

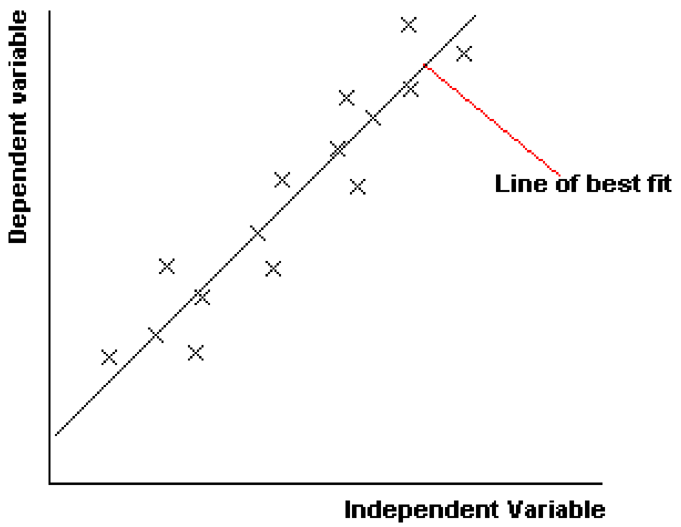


Video: https://youtu.be/VMUQSMFGBDo

Relationship: Correlation Design

- Data from correlation studies look similar to:

Subject	V1	V2	V3	etc
1	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0
.	.	.	.	.
n	0.0	0.0	0.0	0.0



- A graph (scatter plot) of two variables can be positive (uphill) or negative (downhill)

Relationship: Regression Design

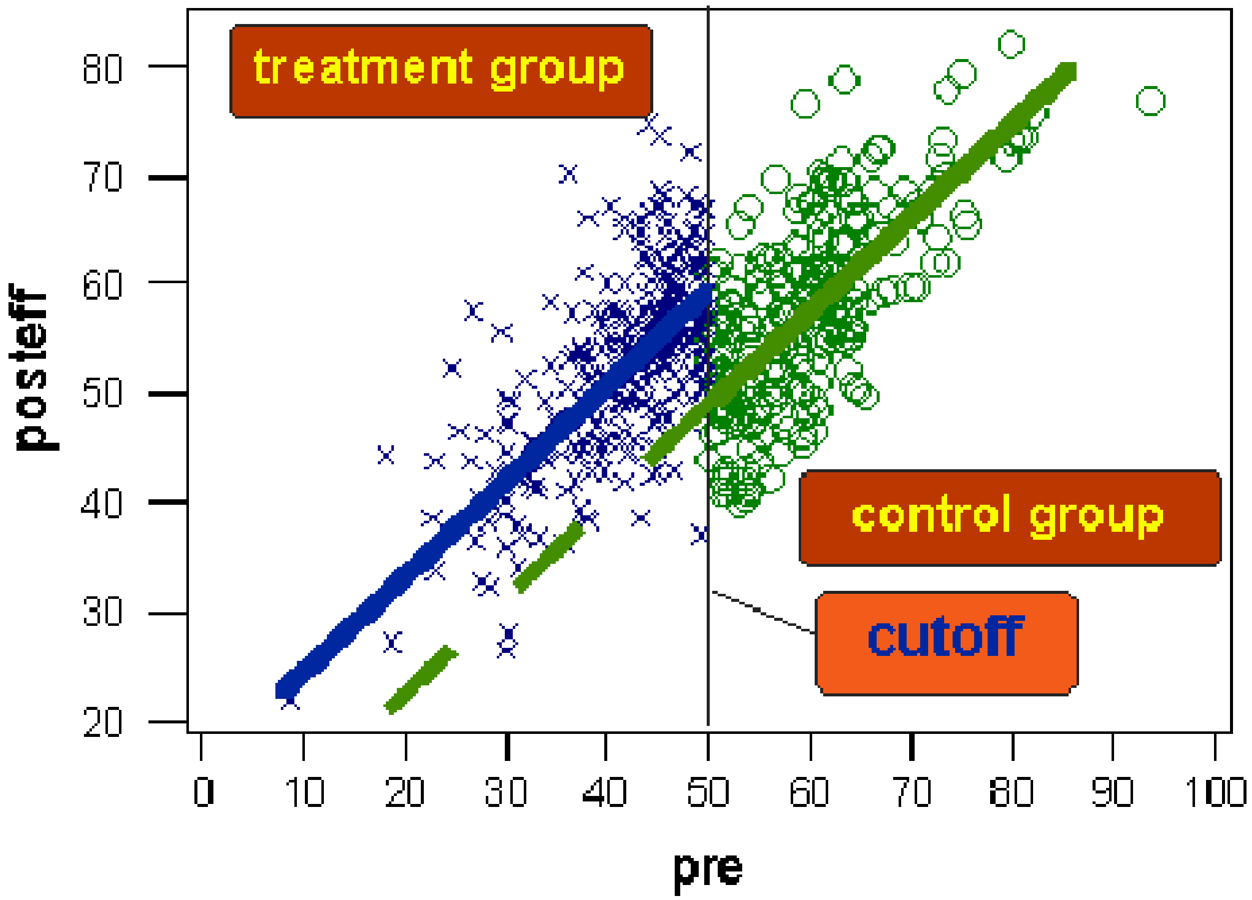
- Treatment applied to one group based on a cutoff
- Pre and post treatment measurements taken:

Event	1	2	3
R Group 1	O	X	O
R Group 2	O		O

- Variables are correlated to observe differences between the groups

Relationship: Regression Design

- Treatment is applied to one group based on cutoff
- Looking for ‘discontinuity’ between regression lines



Differences: Experimental Design

- Use of random assignment is an experimental design
- Use of non-random assignment is quasi-experimental
- Independent variables are categorical
- Dependent variables are either categorical or continuous

Experimental Design

Event		1	2	3
Group 1	O1	X	O2	
Group 2	O1		O2	

- Groups are independent from one another
- Differences in Group 1 compared to the control Group 2 are due to the treatment X

Differences: Factorial Design

- Used when there are one or more independent variables and you wish to find out if there is an interaction between them

Event	1	2
R Group 1	X1A	O
R Group 2	X1B	O
R Group 3	X2A	O
R Group 4	X2B	O

X1 = treatment 1 X2 = treatment 2

A = category 1 B = category 2

Factorial Design

- Factorial designs look for three conditions:
  - Are there differences between treatments X1 and X2 (are there effects from X1 and X2?)
  - Are there differences between group A and group B
  - Are there any interactions between groups A and B (because of X1 and X2)

Differences: Time Series Design

- Data collected at regular intervals to look for changes over time
  - Before a treatment
  - During a treatment
  - After a treatment
- Experimental designs measure effects after all treatments have been done, time series designs observe behavior during each individual stage of treatment
- Extension of classic medical case study. An example would be if we want to know how well a diabetic patient can regulate their glucose level before, during and after treatment

Types of Statistical tests

- Non parametric statistics
  - Tests data that does not belong to any particular distribution

    or
  - Used when data has ranking (order) but doesn't have numeric interpretation
- Parametric statistics
  - Assumes that data comes from certain types of distributions
  - If the data meets the test's assumptions, they can make more accurate estimates than non-parametric statistics.

Categories of Data

- The four possible scales of measurement are:
  - Nominal
  - Ordinal
  - Interval
  - Ratio

Nominal Scale

- Nominal measurements are classification categories
- There is no concept of order or distance between different categories
- Example
  - You might have a survey with different response categories of:
    - No response
    - Male
    - Female
    - Car owner
    - Home owner

Ordinal Scale

- Data in an ordinal scale have rankings with respect to one another. It has Order only
- Ordinal scales do not have a concept of distance
- Example – A survey that asks people to give their opinion on a scale of 1-10 provides Ordinal data.

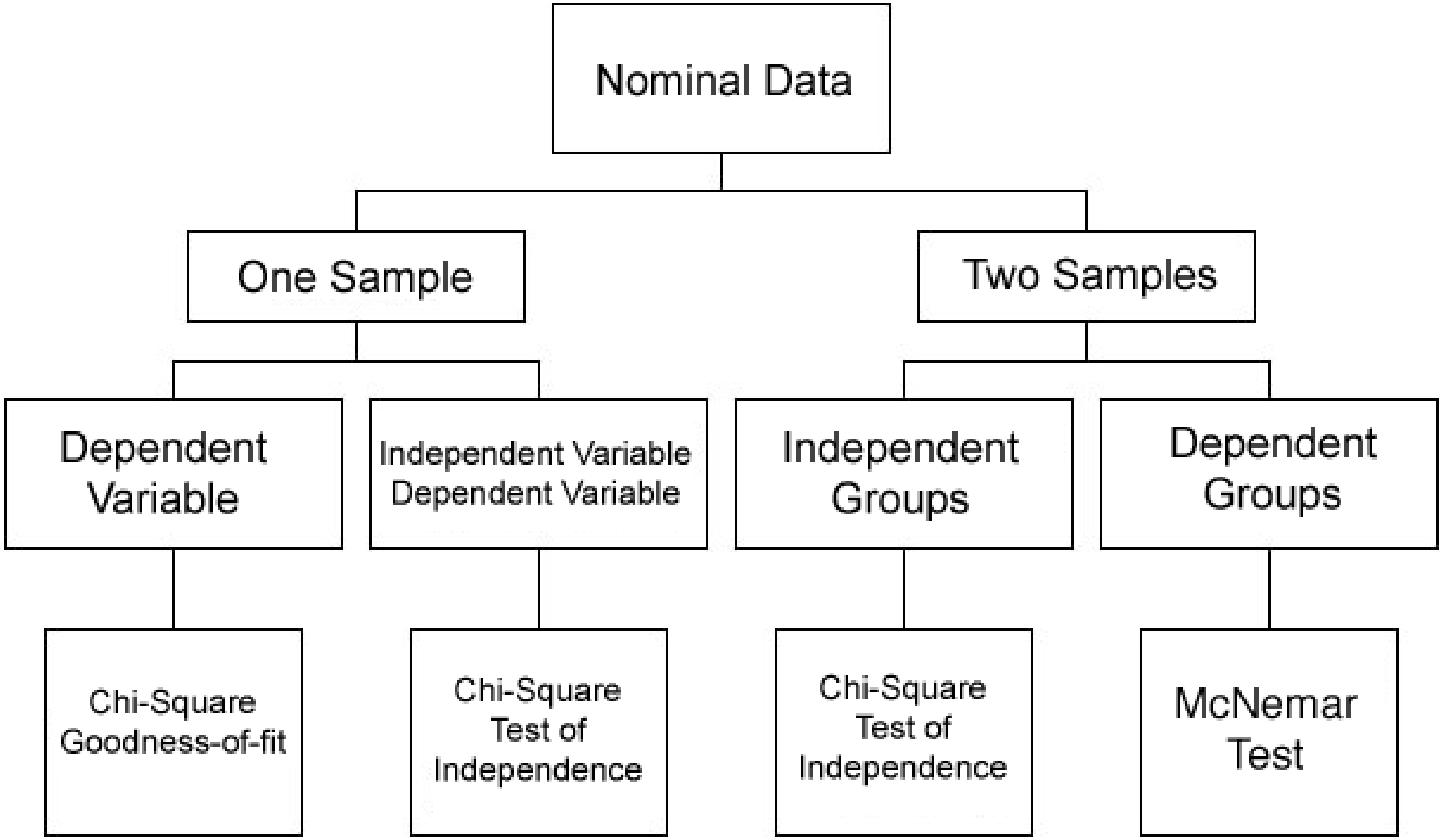
Interval Scale

- An interval scale is a measurement scale
  - There are ordered points along the scale
  - There is a defined distance between each point
  - There is no concept of a zero point on an interval scale to indicate the absence of something
- Example – The Fahrenheit and Celsius temperature scales are interval scales

Ratio Scale

- A ratio scale has certain distance defined between each ordered point on the scale
- A value of 0 indicates the absence of the thing being measured
- Example – The Kelvin temperature scale would be a ratio scale since a 0 value indicates the complete absence of temperature

Tests for Nominal Data

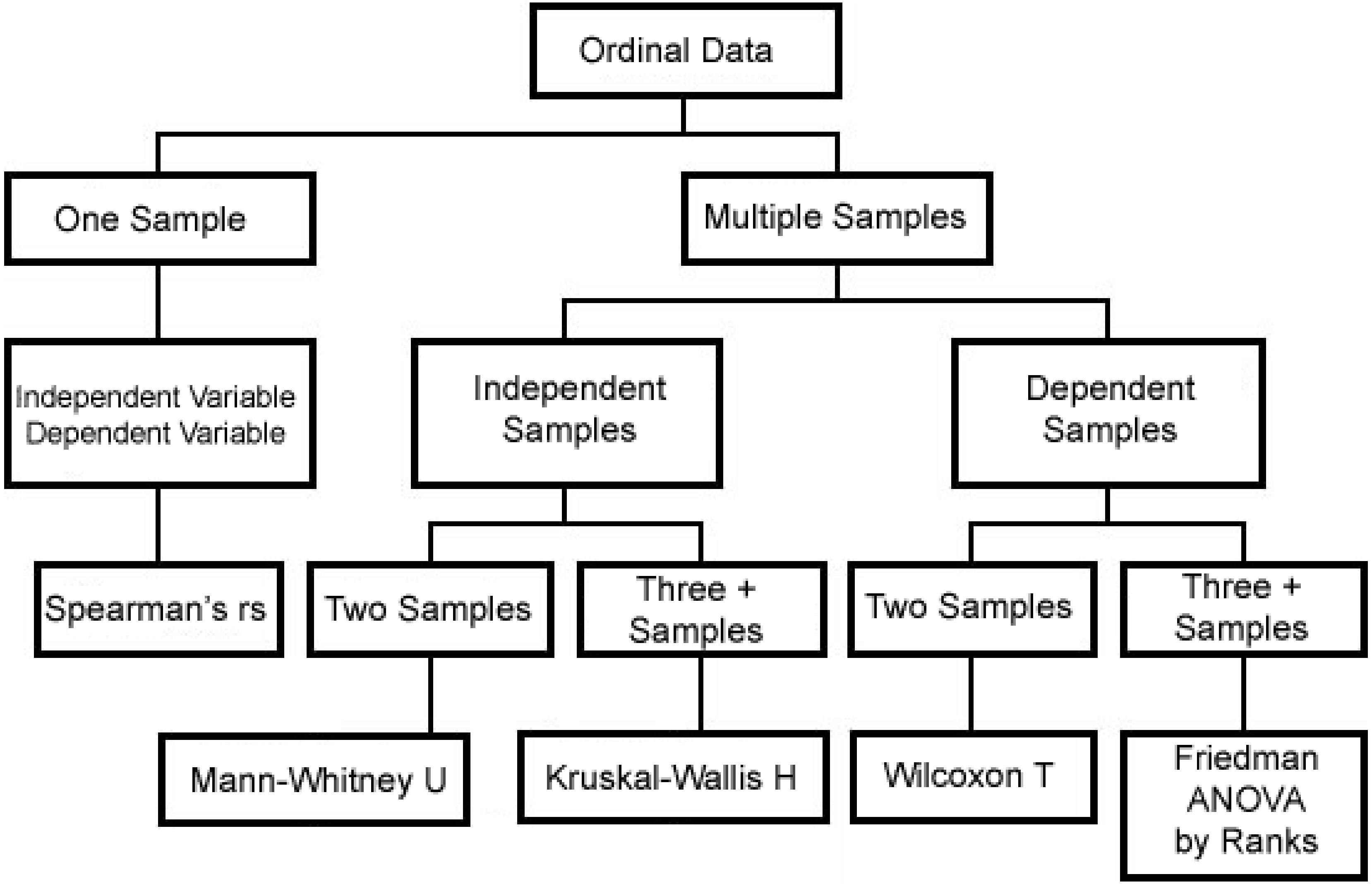


http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_23.html

Chi square test

- Used when data is all nominal
- Tells you if there is any significant differences between what you observe compared with what you expect

Tests for Ordinal Data



http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_24.html

Spearman’s Rank Sum

- Also called Spearman’s rho (denoted by Greek letter r)
- Non-parametric test for correlation
- Values can range from -1 to +1
    - -1 means a one to one negative correlation between variables
    - 0 means that there is no correlation between variables
    - +1 means a one to one positive correlation between variables

Mann-Whitney

- Non parametric test to determining if two independent samples of data are similar
  - Two sample test for significance
  - Similar to the parametric two sample t-test
  - Used on ordinal or higher data
- Can be used on samples with the same or different number of observations

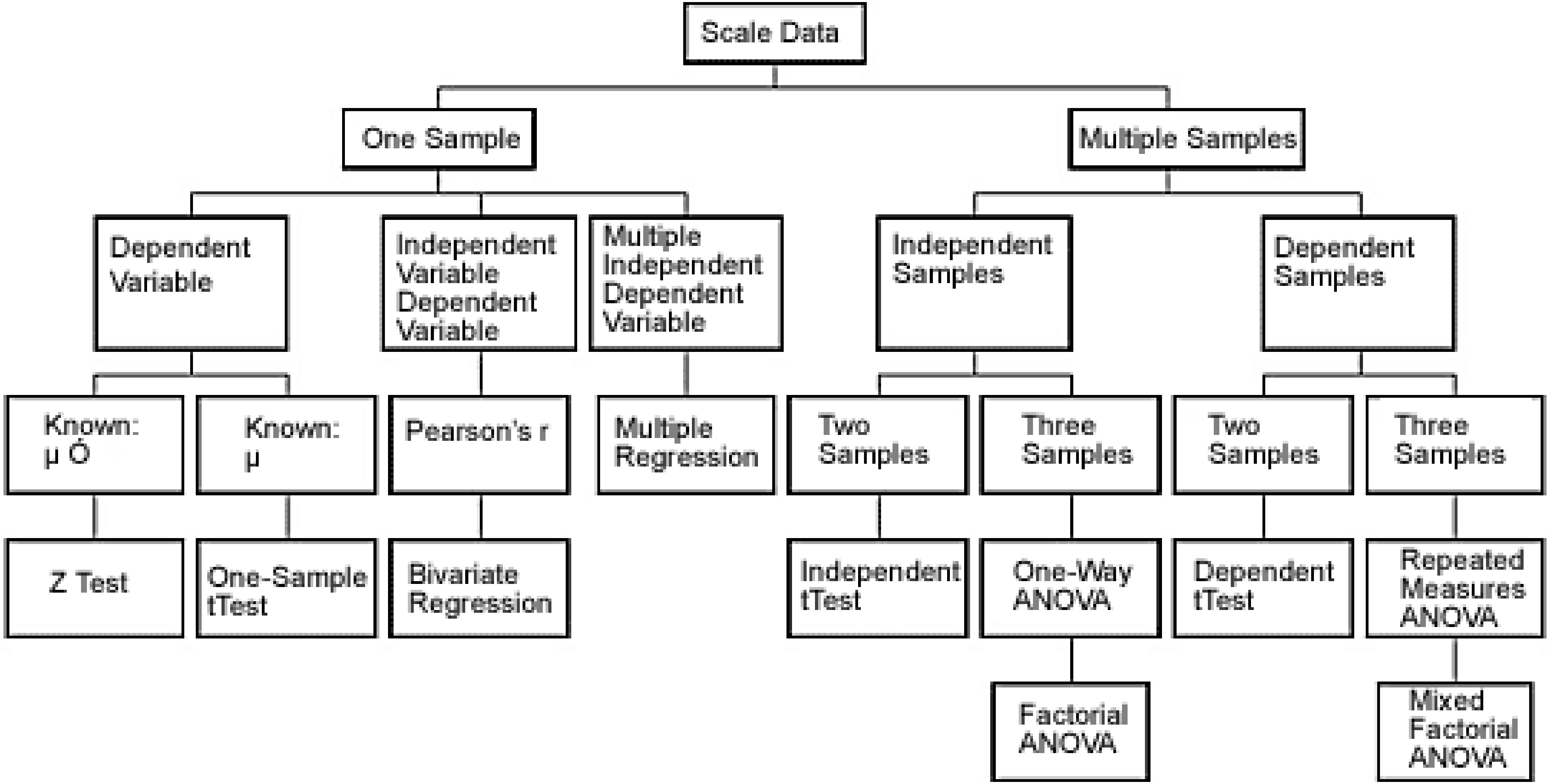
Kruskal-Wallis

- Non parametric method for evaluating differences between multiple independent groups of data
  - Does not assume a normal data distribution
  - It does assume that the shape of the distribution between different groups is the same
- Similar to the parametric one way analysis of variance test

Wilcoxon T

- Used for comparing 2 dependent samples
  - Paired data
  - Repeated measures
- Non parametric
  - Does not rely on a normal data distribution
  - Must be interval or ratio data

Tests for Interval and Ratio data

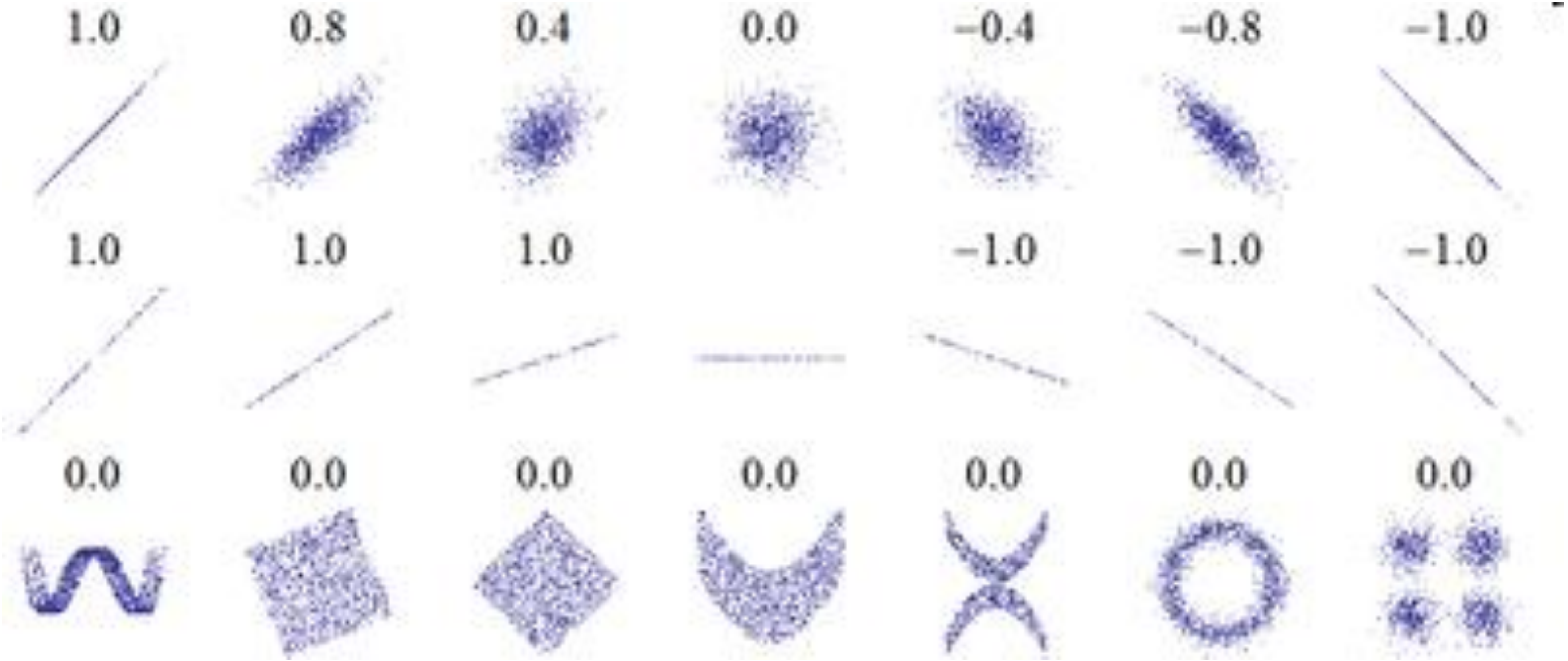


http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_25.html

Pearson's r

- Measure of correlation between two variables
- Values can range from -1 to +1
  - -1 means a one to one negative correlation between variables
  - 0 means that there is no correlation between variables
  - +1 means a one to one positive correlation between variables
- It is a parametric test and assumes a linear relationship between variables

Pearson’s r



t-test

- Parametric test for testing for differences
  - Several versions
    - One sample
    - Multiple samples
    - Independent observations
    - Dependent observations

Analysis of Variance (ANOVA)

- Analysis of variance is a parametric test that is used to look for (mean) differences between three or more sets of data
  - One way ANOVA used to compare 2 or more independent groups
  - Factorial ANOVA used to compare 2 or more treatments upon (multi-dimensional) groups of data
  - Repeated Measures ANOVA uses to analyze cases where the same subjects are used in each measurement (ex. changes over time)

Data Collection – connection to variables

- Measuring Variables
  - Directly measurable
    - duration, errors, temperature, distance, etc
  - Indirectly measurable
    - usability, self-esteem, quality of life, stress level, etc
- Theoretical (or research) Construct

A measurement of a concept that is inferred based on a combination of other, measurable variables

Construct Validity and Reliability

- Construct validity – how accurately does the construct actually measure what it is intended to measure?
- Construct reliability – how consistent is the construct at measuring over time and in different situations?

- Background research
  - What tools have already been developed to measure abstract concepts
  - Are they commonly used in similar situations
  - Are they already accepted as reliable and valid
  - If you make up your own, you have to show that it really measures what you say it does

Example

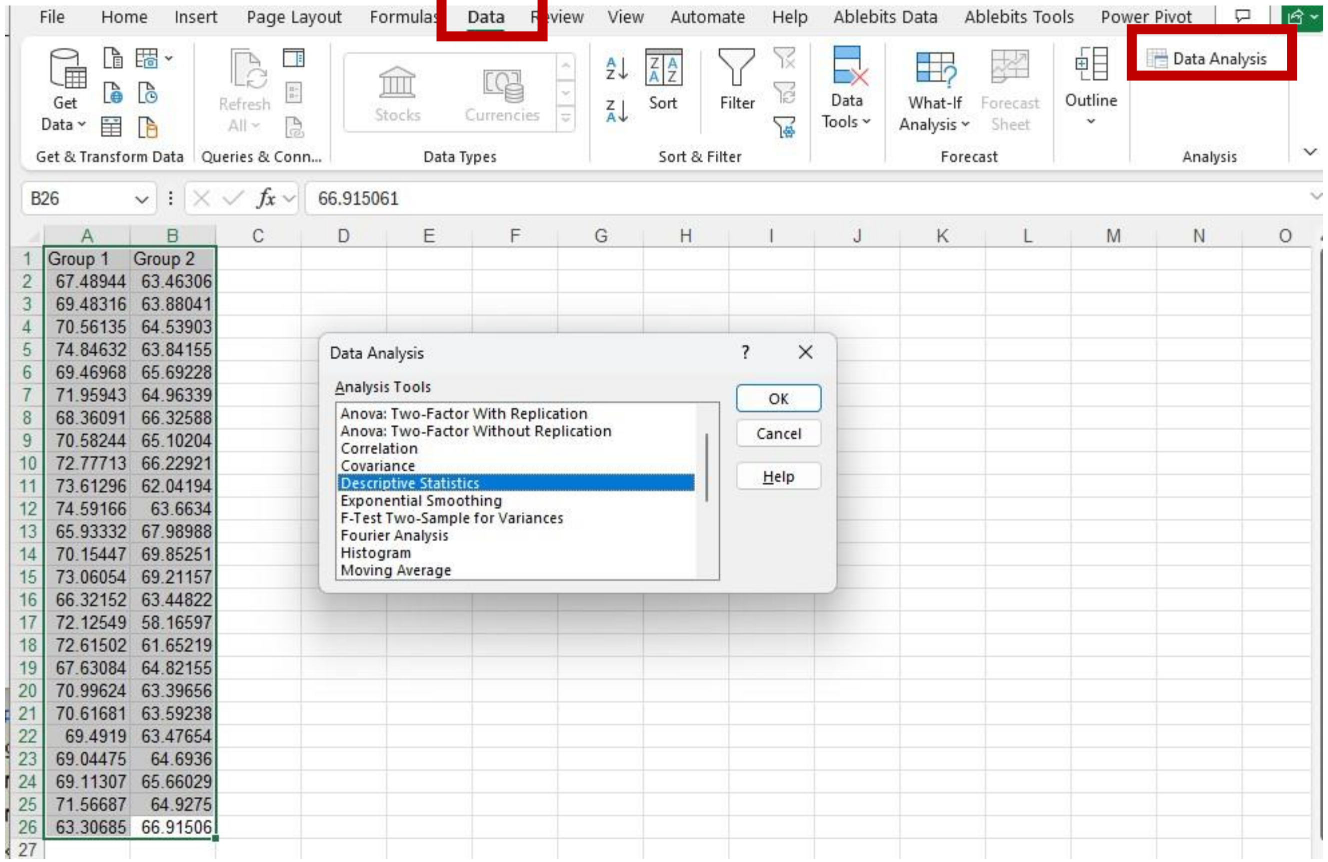
Group 1	Group 2
67.489439	63.463062
69.483160	63.880407
70.561353	64.539034
74.846320	63.841551
69.469678	65.692283
71.959434	64.963393
68.360909	66.325883
70.582437	65.102038
72.777127	66.229205
73.612962	62.041943
74.591664	63.663395
65.933320	67.989878
70.154467	69.852506
73.060535	69.211567
66.321518	63.448222
72.125492	58.165974
72.615020	61.652194
67.630836	64.821550
70.996237	63.396557
70.616807	63.592375
69.491898	63.476537
69.044748	64.693599
69.113072	65.660290
71.566874	64.927502
63.306848	66.915061

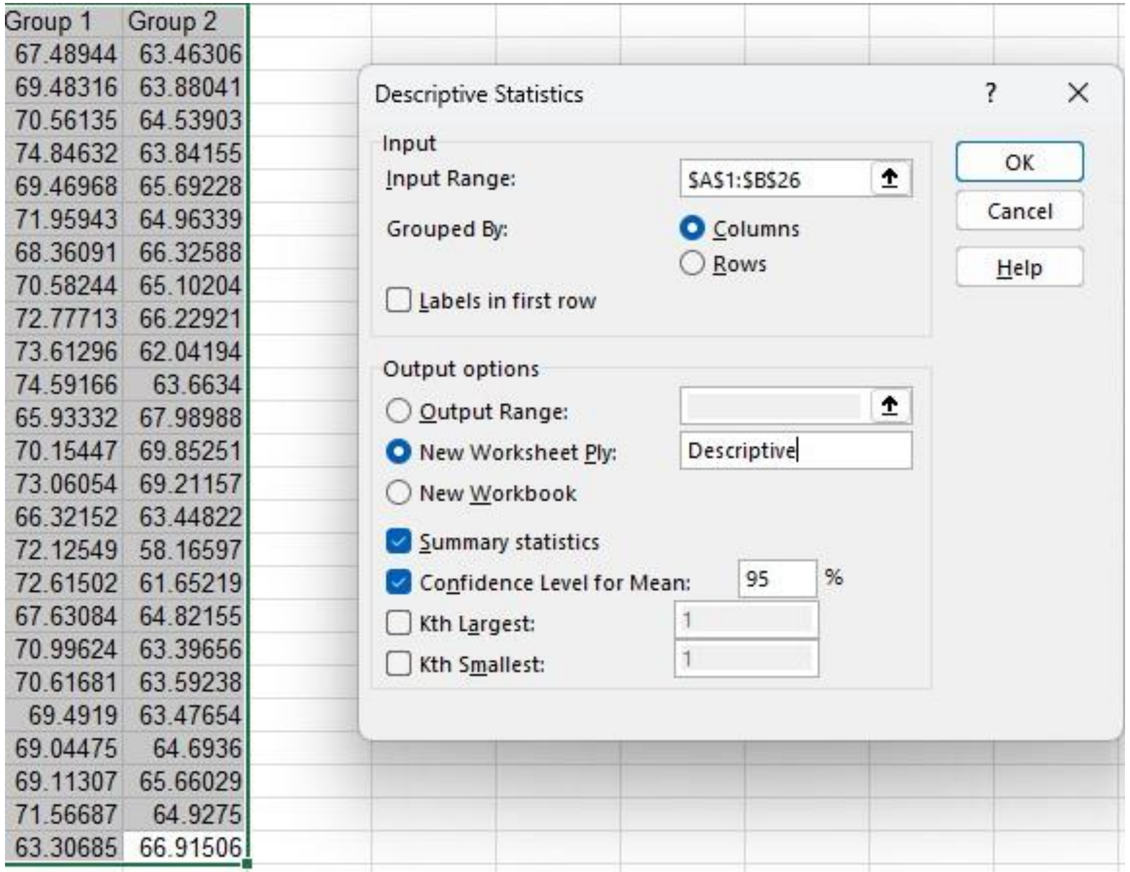
The data shown is from two groups of time measurement data from an investigation comparing whether the controls of a new design allow a task to be done more quickly than the original.

Group 1 is data collected from the original control configuration and Group 2 is data collected from the new configuration.

We want to know, is the new configuration better than the original?

Generating Descriptive stats

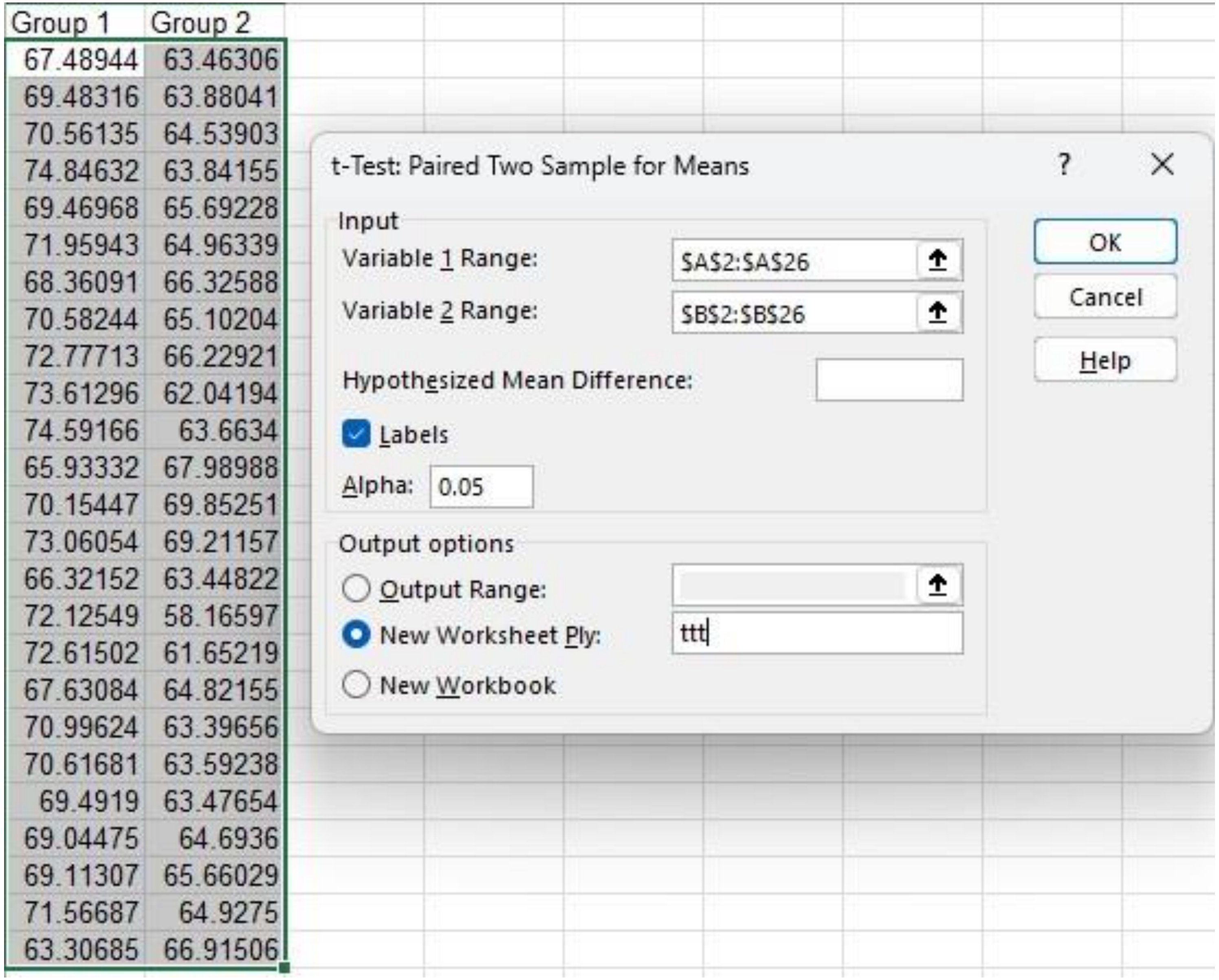




Descriptive stats - Output

Column1		Column2	
Mean	70.2284862	Mean	64.70184
Standard Error	0.555390366	Standard Error	0.481136
Median	70.561353	Median	64.6936
Mode	#N/A	Mode	#N/A
Standard Deviation	2.77695183	Standard Deviation	2.405681
Sample Variance	7.711461466	Sample Variance	5.787301
Kurtosis	0.275201715	Kurtosis	1.774828
Skewness	-0.488985843	Skewness	-0.17387
Range	11.539472	Range	11.68653
Minimum	63.306848	Minimum	58.16597
Maximum	74.84632	Maximum	69.85251
Sum	1755.712155	Sum	1617.546
Count	25	Count	25
Confidence Level(95.0%)	1.146269378	Confidence Level(95.0%)	0.993016

Running a t-test



Group 1	Group 2
67.48944	63.46306
69.48316	63.88041
70.56135	64.53903
74.84632	63.84155
69.46968	65.69228
71.95943	64.96339
68.36091	66.32588
70.58244	65.10204
72.77713	66.22921
73.61296	62.04194
74.59166	63.6634
65.93332	67.98988
70.15447	69.85251
73.06054	69.21157
66.32152	63.44822
72.12549	58.16597
72.61502	61.65219
67.63084	64.82155
70.99624	63.39656
70.61681	63.59238
69.4919	63.47654
69.04475	64.6936
69.11307	65.66029
71.56687	64.9275
63.30685	66.91506

t-test results

t-Test: Paired Two Sample for Means		
	67.489439	*63.463062*
Mean	70.34261317	64.753456
Variance	7.706960716	5.969422321
Observations	24	24
Pearson Correlation	-0.319322245	
Hypothesized Mean Difference	0	
df	23	
t Stat	6.452338981	
P(T<=t) one-tail	6.93989E-07	
t Critical one-tail	1.713871528	
P(T<=t) two-tail	1.38798E-06	
t Critical two-tail	2.06865761	

Activities

Activity 1 : Analyzing Data

Data Set 1	Data Set 2
162	168
170	136
184	147
164	159
172	143
176	161
159	143
170	145

Instructions

- Generate a descriptive statistics and perform a t-test on the following two data sets to determine if there is a significant difference between them.

Activity 1 : results

Data Set 1		Data Set 2	
Mean	169.625	Mean	150.25
Standard Error	2.853178	Standard Error	3.903981
Median	170	Median	146
Mode	170	Mode	143
Standard Deviation	8.070006	Standard Deviation	11.04213
Sample Variance	65.125	Sample Variance	121.9286
Kurtosis	0.072339	Kurtosis	-1.06908
Skewness	0.515471	Skewness	0.519181
Range	25	Range	32
Minimum	159	Minimum	136
Maximum	184	Maximum	168
Sum	1357	Sum	1202
Count	8	Count	8
Confidence Level(95.	6.746694	Confidence Level(95.0%)	9.231449

t-Test: Paired Two Sample for Means

	Data Set 1	Data Set 2
Mean	169.625	150.25
Variance	65.125	121.9286
Observations	8	8
Pearson Correlation	-0.17674777	
Hypothesized Mean Difference	0	
df	7	
t Stat	3.706873373	
P(T<=t) one-tail	0.003792994	
t Critical one-tail	1.894578605	
P(T<=t) two-tail	0.007585988	
t Critical two-tail	2.364624252	

Activity 2 : Study Setups

Instructions

- Break up into groups of 3 people
- Devise a study design for one of the scenarios
- Define the following:
  - What question will be answered
  - What kind of study design will you use
  - What are the variables (what will you actually measure. It may help to sketch out what the products might look like)
  - How will the data be collected (directly measured, survey, etc)
  - How would the data be analyzed in order to answer the research question
- Some teams will present their designs

Setup a study

Topic 1: You have designed a washer/drier set and you want to find out if the new design is more efficient and causes less physical stress for the user.

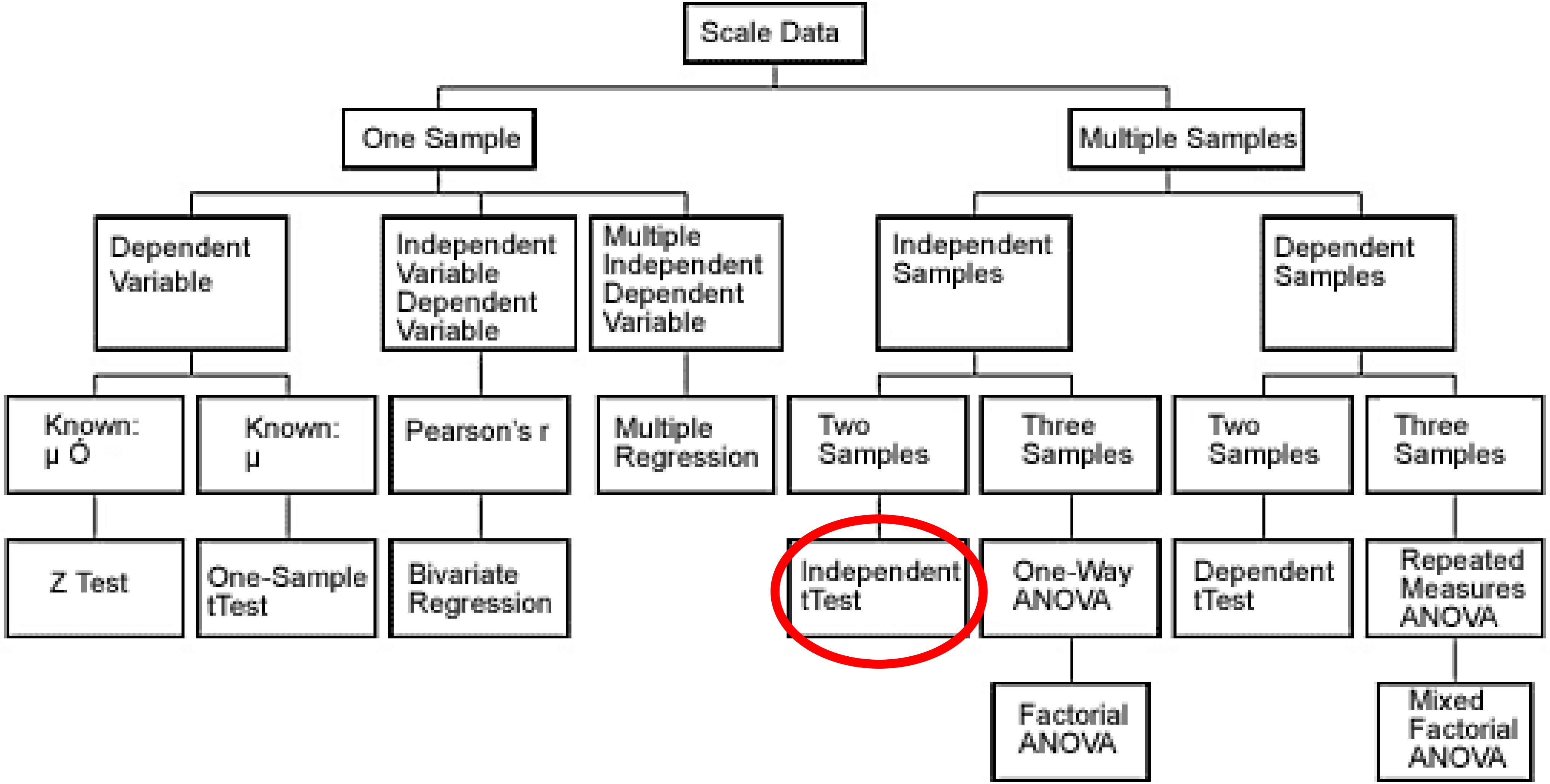
Topic 2: You have designed a new control panel layout that will be used to control an automobile assembly line and you want to find out if the new layout results in fewer operator errors.

Topic 3: You have 4 competing designs for a new cooktop. Each one has a different arrangement between the controls and the heating elements. You want to find out which arrangement that users most correctly associate between each control and heating element.

Topic 1

- Question: The new washer/drier design allows faster washer/drier transfer with less lower back stress.
- Study Design: Differences (simple experimental)
- Variables:
  - Washer Design (independent)(nominal data)
  - Transfer time (dependent) *(ratio data)*
  - Low back stress (dependent) *(interval data)*
- Data:
  - Transfer time-measured in seconds needed to move a standard load from washer to drier
  - Low back stress - calculated by observing lifting motion and utilizing the NIOSH lifting equation
- Analysis : Independent t-test

Tests for Interval and Ratio data



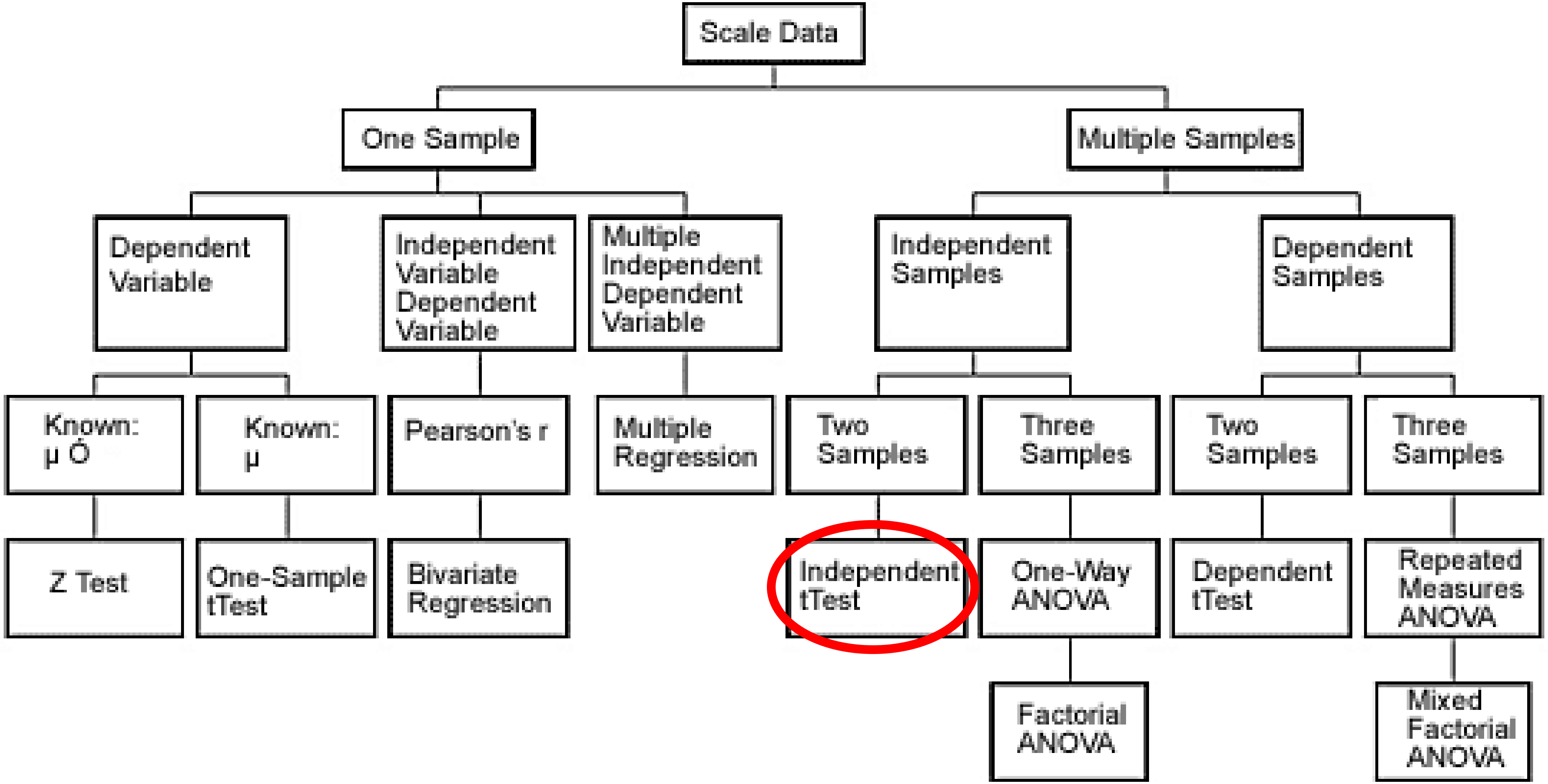
http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_25.html

Topic 2

- Question: The new control panel layout will lead to fewer operator errors when performing a task.
- Study Design: Differences (simple experimental)
- Variables:
  - control panel design (independent) (nominal data)
  - number of errors (dependent) *(ratio data)*
- Data:

  number if errors would be counted while a task is performed
- Analysis: Independent t-test

Tests for Interval and Ratio data



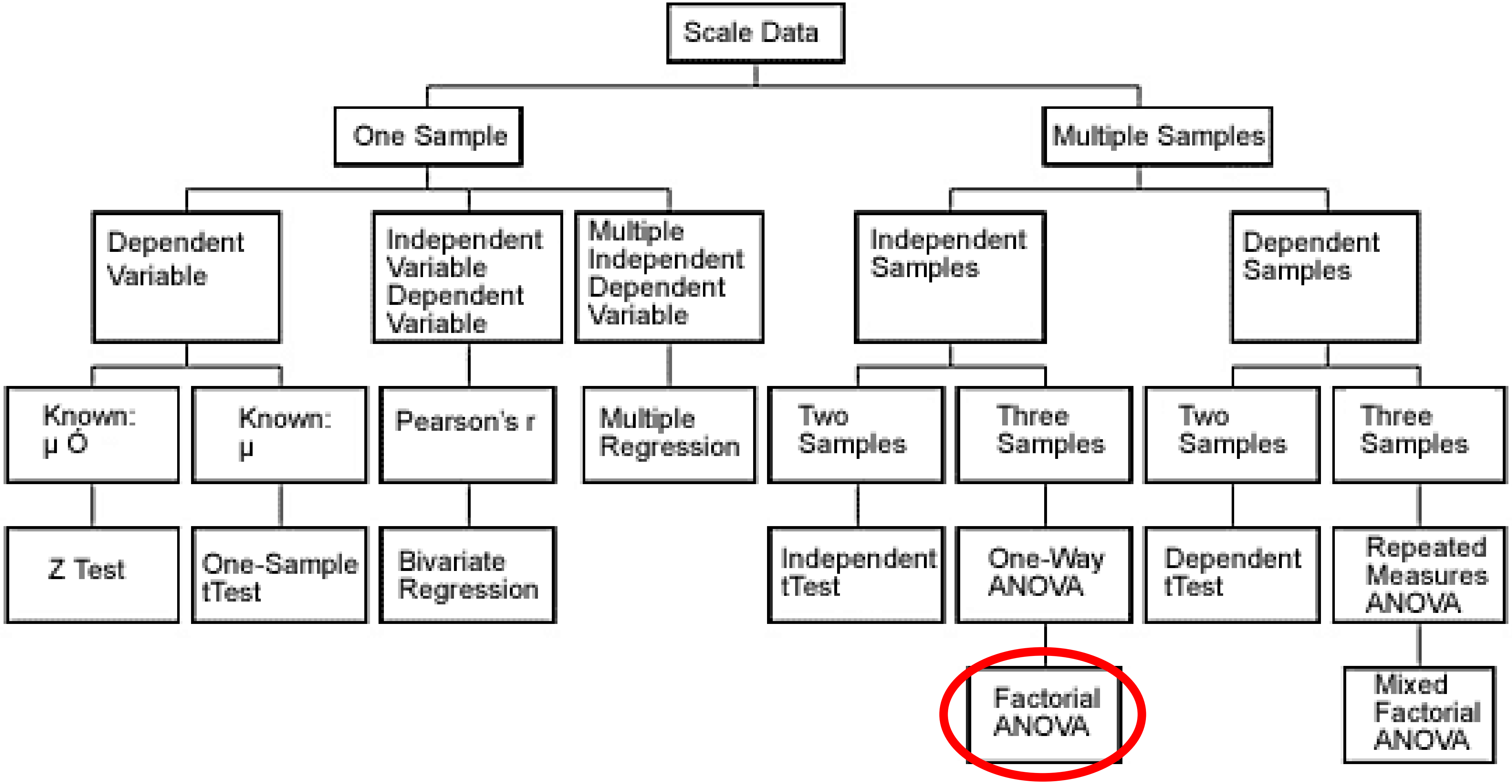
http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_25.html

Topic 2 - alternate

- Question: The new control panel layout will lead to fewer operator errors when performing multiple defined tasks.
- Study Design: Differences (factorial design)
- Variables:
  - control panel design (independent) (nominal data)
  - the assigned task (independent) (nominal data)
  - number of errors (dependent) *(ratio data)*
- Data:

  number if errors would be counted while each task is performed
- Analysis: Factorial ANOVA

Tests for Interval and Ratio data



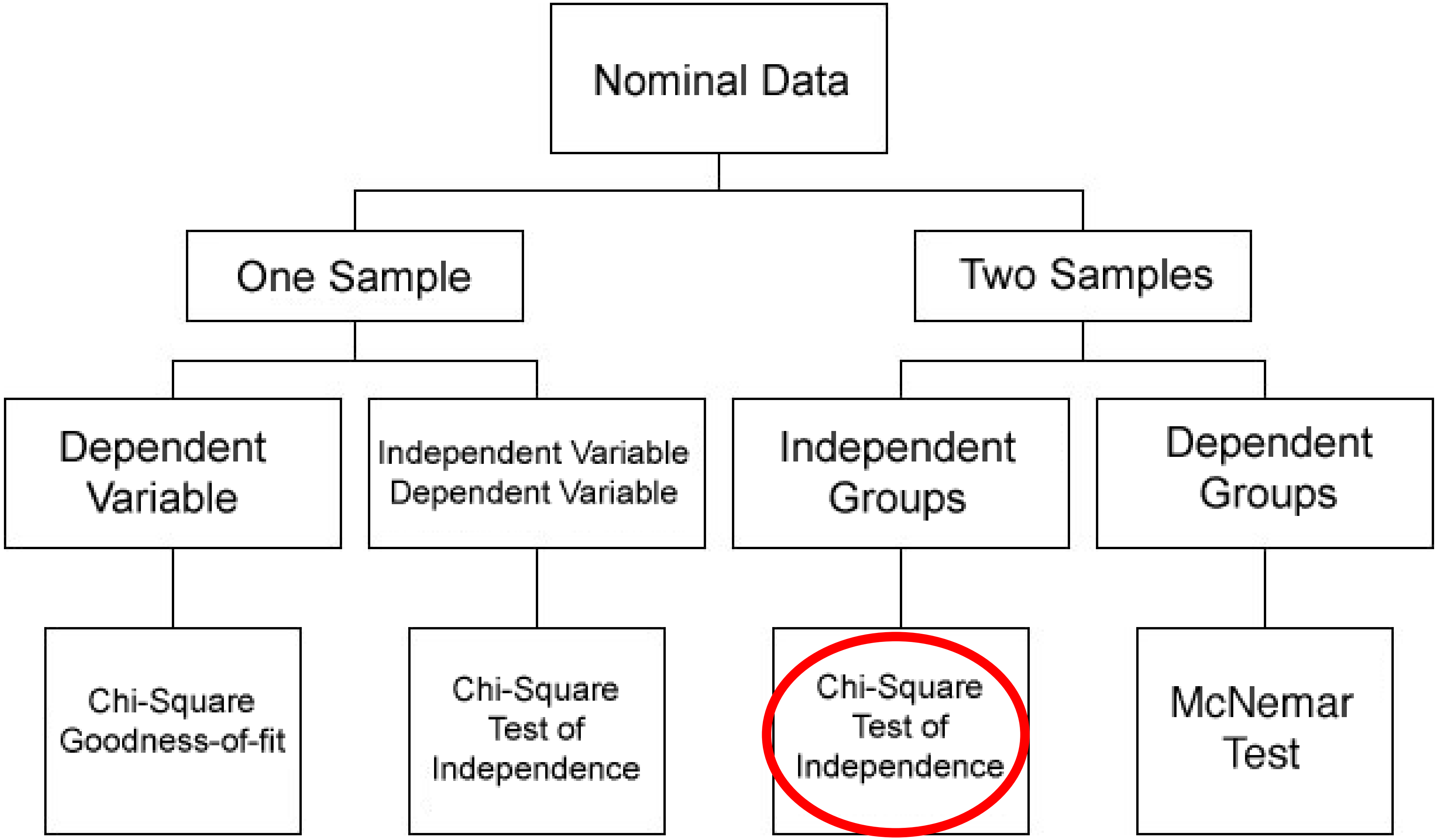
http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_25.html

Topic 3

- Question: Users more intuitively associate the correct control to burner functionality of one of the new cooktop designs compared to the old?
- Study Design: Differences (simple experimental)
- Variables:
  - Cooktop/control layout (independent) (nominal data)
  - User assumption (dependent) *(nominal data)*
- Data:

  Survey users to discover which control they assume operates each burner
- Analysis: Chi square test for independence

Tests for Nominal Data



http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_23.html

Recap

- What you don’t need to worry about
  - You don’t need to be a stats expert
  - You don’t need to memorize or know how or when to use all the kinds of tests
- You **should**
  - understand how data relates to the variables you want to measure
  - understand whether or not those variables are actually measurable
  - understand how to setup an appropriate investigation to answer a question.
    - Description
    - Correlation
    - Differences
- This will help you:
  - Answer specific questions
  - Understand how data can support your design decisions
  - Justify and support your decisions and outcomes
  - Consider how you can approach your eventual IRP project
- Resources
  - A copy of this will be on Moodle
  - Sample: http://uxpajournal.org/wp-content/uploads/sites/7/pdf/JUS_Choi_August2019.pdf


JUS
Journal of Usability Studies
Vol. 14, Issue 4, August 2019 pp. 187–200

Applying Tangible Augmented Reality for Product Usability Assessment

Young Mi Choi
Associate Professor
School of Industrial Design, Georgia Institute of Technology
247 4th St. NW
Atlanta, GA 30332
USA
christina.choi@gatech.edu

Abstract

When developing a new product, it is common for designers to feel that they do not have enough information about users’ needs. This is especially true at the front end of a new product design process. An important component of this process is testing and validating potential design concepts. The aim of this study was to explore the validity of augmented reality (AR) and tangible augmented reality (TAR) as tools for evaluating the usability of a product. For this study, 70 college students were recruited to perform a usability evaluation of a space heater product and equivalent AR and TAR representations of it. The results indicate that overall TAR can be a reliable method for evaluating the usability of a fully realized product, especially for products with physical interface controls. However, TAR was not found to be reliable with respect to Ease of Use. Overall, AR was not found to be as reliable as TAR or with respect to any specific aspect of usability that was measured. Applications, limitations, and areas for further study are discussed in this article as well.

Keywords

Tangible Augmented Reality, Augmented Reality, usability assessment, product design evaluation