



The role of statistics in research

Martin Kidd

Centre for Statistical Consultation
(CSC)



Phases of a project

- Planning
- Execution
- Data capturing
- Data analysis
- Reporting

Types data

- Nominal or classification data
 - Identify groups of subjects (male/female)
 - No order
- Ordinal data
 - Order important
 - Discrete data (1,2,3,4 or low, medium, high)
- Continuous data
 - Measurements



Planning

- Aims & objectives will determine the statistical content
- Proper statistical planning essential
- Statistician should preferably be involved
- Define the population
- Possible experimental design
 - Experimental vs control
 - Pre vs post testing for interventions

Questionnaire design

- Use existing instruments!
- Base design on previous studies
- Think about questionnaire validity:
 - Content validity
 - Reliability
 - Discriminant validity
 - Etc.
- Too many people ask these questions after the survey

Questionnaire design

- Responses to questions
 - List of choices: only one can be selected
 - Multiple selection: more than one selected
 - Continuous data
 - If an accurate number can be filled in, don't use categories
 - Eg age: fill in exact age, not age categories
 - Open ended
 - If list of choices type, try not to leave open ended
 - Time consuming to analyze

Questionnaire design

- Types of responses
 - Likert scales:
 - How many options? 4pt, 5pt, 7pt scale?
 - Be careful of the wording
 - Make sure its ordinal !!!

1 Completely disagree

2 Disagree

3 Unsure

4 Agree

5 Completely agree

1 Completely disagree

2 Disagree

3 Neutral

4 Agree

5 Completely agree

Sample size

- In 99% of the cases governed by time and money
- Sample size calculation some times a requirement
- Never easy because it requires information that's hard to come by or does not exist

Sample size

- Effect size, variation in the data and power plays a critical role.
- Usually not known before start of data collection
 - Use literature
 - Pilot study

Execution

- Pilot study in most cases essential
- Sort out procedures for data capturing
- Statistician can (and should) give advice
- Possible interim stats for more accurate sample size calculations

Data capturing

- Please make sure data capturing is done correctly
- Strict guidelines:
 - Cases in rows
 - Variables in columns
 - Etc...
- I spend 90% of my time cleaning data, 10% on statistical analysis!

Data capturing

- Most commonly done in Excel
- Variables in columns
- Subjects in rows
- Each variable must have a unique name
- Names in first row

Data capturing

- Data for control & experimental on the same sheet
- Descriptive names for variables
- Descriptive names for levels of nominal variables:
 - Gender: male, female
- Use numbers for ordinal variables:
 - Effectiveness: 1 – bad, 2 – medium, 3 – good

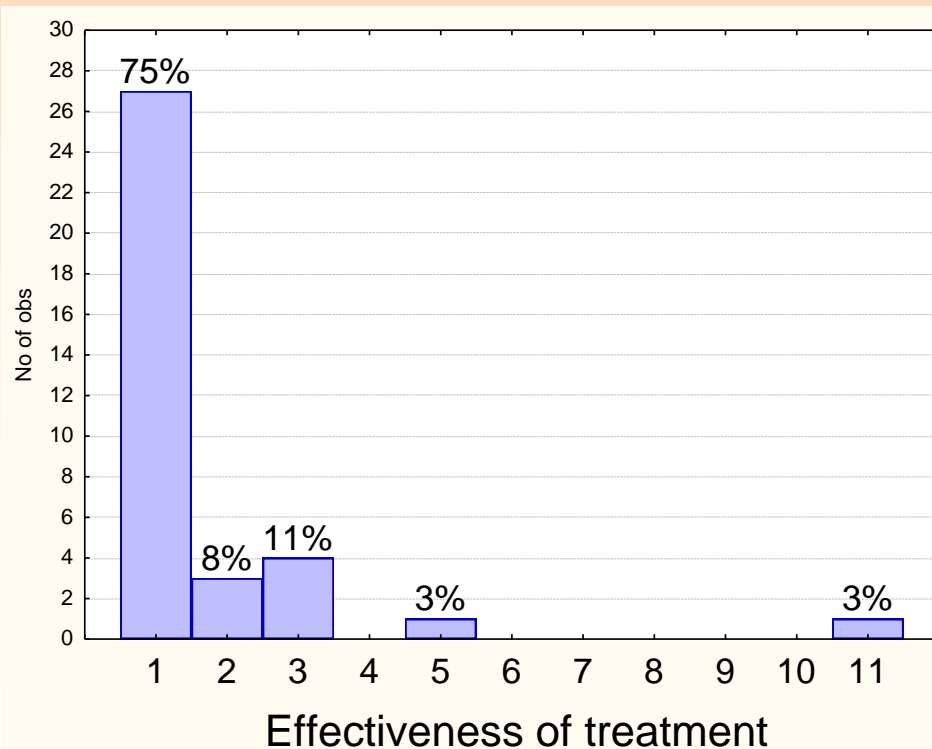


Data capturing

- Excel spreadsheets can be customized to make data entry easier
- We don't do capturing ourselves
- We can refer to people that do it for remuneration
- Probably the most important step:
 - One can easily get lost in sea of data
 - Mistakes: analyses must be re-done!

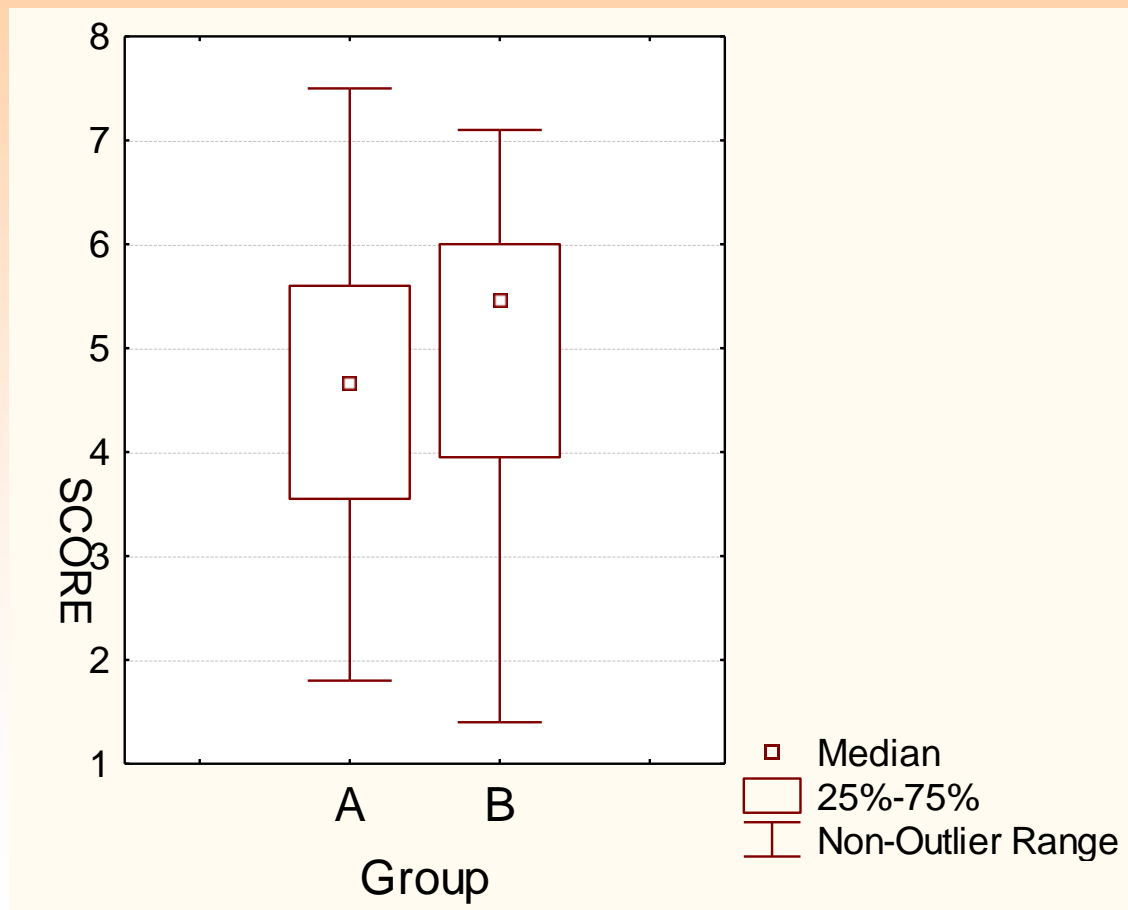
Data analysis

- Most of the time be handled by statistician
- Summary statistics:
 - Frequency tables, means sd's etc
 - Important 1st step for identifying problems and guiding further analyses



Data analysis

- Summary statistics: more examples



Data analysis

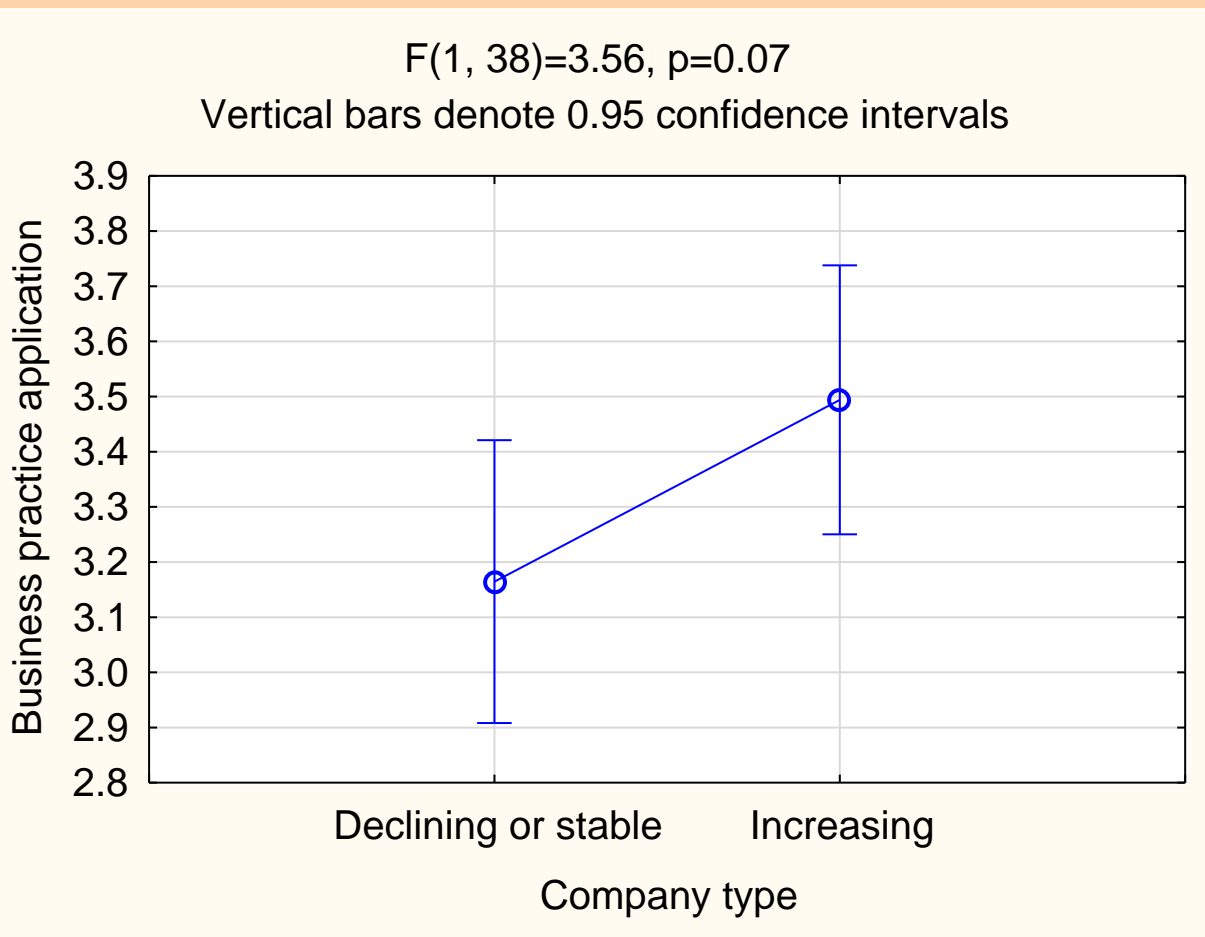
- Post summary statistics:
 - Statistician must be guided by researcher
 - Formulate list of questions based on aims
 - Statistical hypotheses are derived from these questions
 - Statisticians are not experts in your field so we don't understand your jargon

Data analysis

- The specific techniques used are determined by the questions, not the other way round!
- Appropriate techniques are dependent on hypotheses and types of data
- Examples:
 - Comparing means between groups: ANOVA
 - Comparing categorical variables: Cross tabulation with the χ^2 test
 - Many more

Data analysis

- ANOVA example



Data analysis

- Statistical packages:
 - CSC uses Statistica for 90% of analyses
 - SPSS
 - SAS
 - R (becoming more and more popular)
 - Various specialized packages

Reporting

- Responsibility of the researcher
- Statistician will assist and guide with interpretation
- Actual writing should be done by the researcher
- Statistician can read results sections and make comments

Reporting

- Co-authorship:
 - Should the statistician be co-author?
 - Decision of the principal researcher
 - If we are co-authors, we want insight into the article before its submitted

Centre for Statistical Consultation

- Primary aim is to assist students/researchers with stats
- Personnel
 - Prof Martin Kidd mkidd@sun.ac.za
 - At USB campus on Tuesdays
 - Prof Daan Nel (Director) dgnel@sun.ac.za

Centre for Statistical Consultation

- Funding:
 - CSC 50% funded centrally by University
 - Tariff for 2017: R280/hour
 - Paid by dept/study leader research funds
 - Two faculties makes funds available:
 - Economic & Management Sciences
 - Education
 - We do not invoice students