MONASH University

Information Technology

FIT5190 Introduction to IT Research Methods

Lecture 6

The Nature of Evidence

Slides prepared by

Frada Burstein, David Green, Jacques Steyn, Geoff Webb, Chung-Hsing Yeh

Learning objectives

- Be aware of
  - Basic issues in dealing with evidence
  - Common types of data
  - The roles of quantitative evidence (e.g. scales) and qualitative evidence (e.g. classifications).
- Understand
  - The basics of scale construction
  - Basic methods of quantifying data
  - The nature of surveys
  - The advantages of online resources

Overview

- This lecture introduces some basic ideas regarding evidence in research.
- Topics covered include the following:
  - Justification in research;
  - Causation and correlation;
  - Scales and measurement;
  - Surveys;
  - Online data resources.

Justification in scientific writing

- EVERY statement must be justified by one of:
  - Direct support from evidence
  - Deduction from earlier justified statements
  - Reference to literature
- Example:

  "**Previous research** has established that the PQS method is valid in general **(Jones, 1995)**. However, **as Table 3 shows**, it can fail when certain conditions are not satisfied."

Research design

Evidence takes many forms

Justification via evidence

- Existence proof
  - e.g. implement program to prove that your model works
- Counter example
  - e.g. find a case that contradicts a general theory
- Characterisation
  - e.g. test key properties of a system
- Comparison
  - e.g. speed of new algorithm vs. old
- Indicators
  - e.g. user numbers as indicator of quality
- Experiments
  - e.g. test hypothesis on actual cases

Existence proof

- Common in computing
  - Implement an idea to show it works
  - Generate a solution
- Some common sub-questions
  - Performance
  - Characterisation
  - Implications

Causation

- A and B are independent: **A B**
- A causes B: **A → B;**
- Feedback: **A←→B**
- A contributes to B: **A →C→B←D**
- A and B have a common cause: **C→A C→B**
- Side effects: **A+C→B**
- Placebo effect

Causation – some false inferences

- Confusing correlation with causation
  - I touched the keyboard and the computer crashed
- “B follows A” does not mean “A causes B”
  - After opening King Tut’s tomb, the archaeologist died.
- Events often have more than a single cause
  - Blaming one person for an accident when many factors contribute
- Biased perception / confirmation bias
  - Every time I visit the bank I get stuck in the slow queue (ignores times when the queue was NOT slow)

Causation – some false inferences

- False generalisation
  - Cats have 4 legs
    My dog has 4 legs, so my dog is a cat
  - Some one named Fred owes me money
    This man is named Fred, so he must owe me money
  - I visited a restaurant in Paris and the waiter was rude
    Therefore every restaurant in Paris is rude to visitors
  - When I visited Finland the weather was hot
    Therefore Finland is a hot country

Debugging as hypothesis testing

- My computer screen does not work
  - Either
    (A) the screen is faulty
    (B) the video driver is faulty
    (C) the cable is faulty
  - Assume A,
    - then a new screen will work
    - changing cable or computer will not get it to work.

Correlation and causation

What can you conclude?

Correlation and causation

What can you conclude?

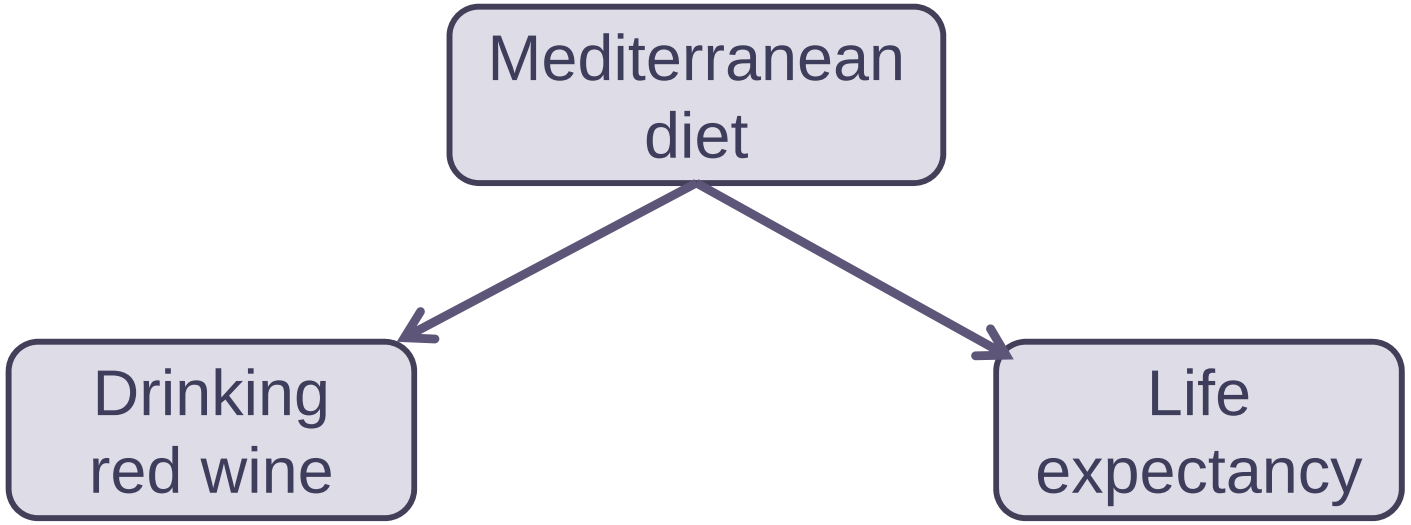
- Drinking red wine improves life expectancy?

Correlation and causation

What can you conclude?

- People who live longer are more inclined to drink red wine

Correlation and causation



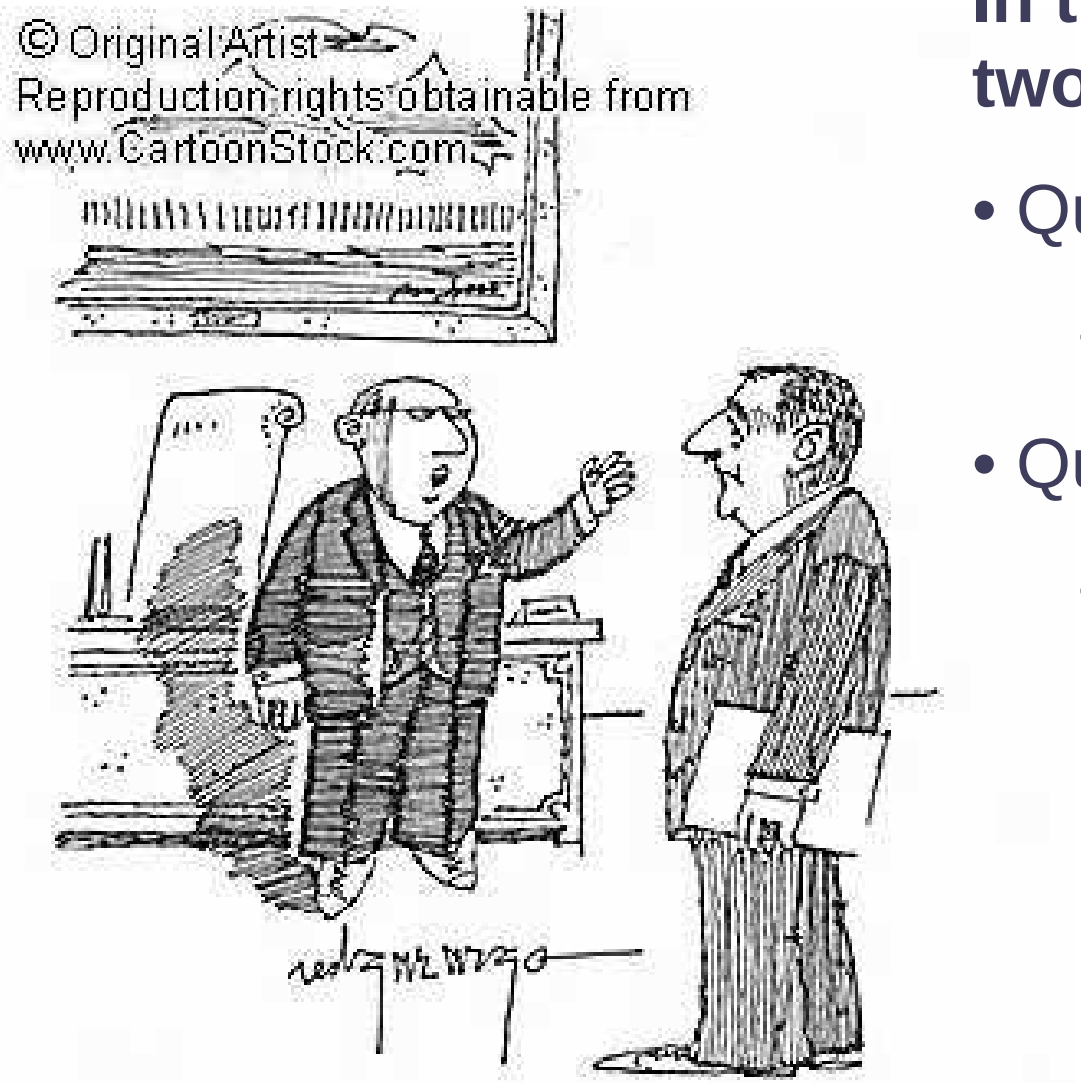
What can you conclude?

- There are common causal factors

Exercise - correlation and causation

- Variables may be correlated for several reasons:
  - Chance
  - Structural similarity
  - Indirect causation
  - Direct causation
- Which of the above might lead to these correlations?
  - Cases of pneumonia & football ticket sales
  - Cases of pneumonia & umbrella sales
  - Cases of pneumonia & use of new herbicide (weedkiller)
  - Cases of pneumonia with race winnings
  - Weight with height (for humans)
  - Weight with annual income

Evidence



In this lecture we look at two kinds of evidence:

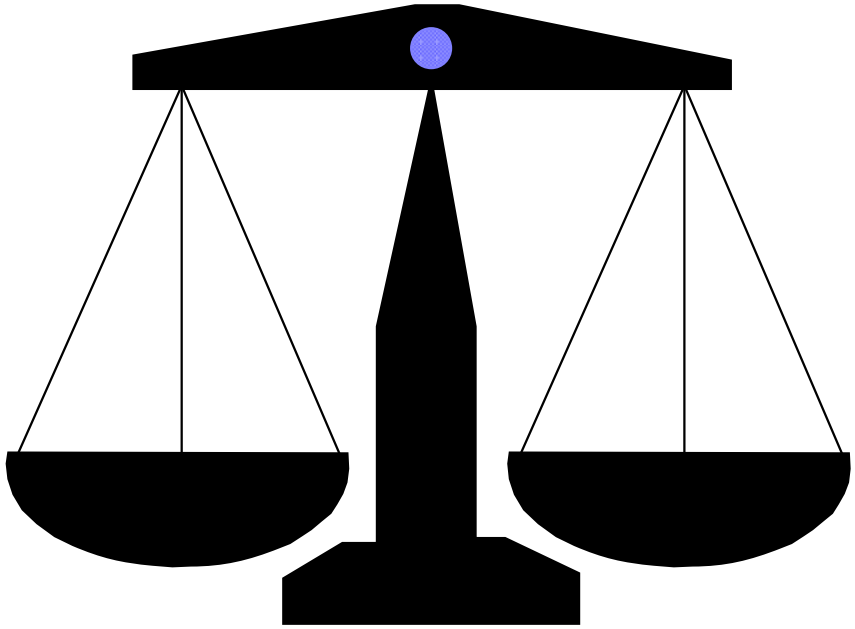
- Quantitative / Numeric
  - e.g. scales
- Qualitative / Categorical
  - e.g. classifications

Data models

- Define goals of the study
  - Express goals as questions
- How do you identify what data is required?
  - Understand the intended use of the data
  - Balance possible against desirable
- Collect raw data, establish how to create derived data
  - e.g. height = 1.9m, rather than height = “tall”
- Qualitative or quantitative or both?

Scales and measurement

- Qualitative (Categorical)
  - Nominal
  - Ordinal
- Quantitative (Numeric)
  - Interval
  - Numerical (Ratio)



Nominal scale - qualitative

- Categories, e.g.
  - Colour: Red, Yellow, Green, Blue, ...
  - Blood type: A, B, AB, O
  - Yes and No
- No ordered relationship to one another
- Possible to assign numeric values to different categories but not relevant mathematically.
- No zero value
- Must be possible to categorise every subject

Ordinal scale - qualitative

- Order: 1st < 2nd < 3rd etc.
- Illustrates the relative order of magnitude
- Cannot determine the distance between them
- Basically a ranking scale
- Usually assume the distance between the intervals in unequal
- e.g. Grade: Fail, Pass, Credit, Distinction, High Distinction

Interval scale - quantitative

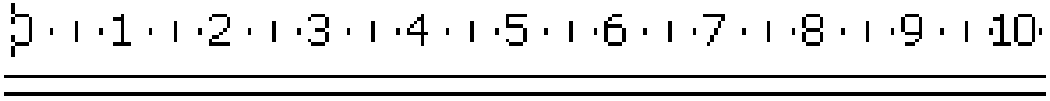
- Both order of magnitude and distance between values in the scales.
  - Ordered, differences between values have mathematical meaning
- Do not have a true zero point, i.e. no meaningful zero
  - e.g. temperature
- Not possible to say 30^{o}C is twice as hot as 15^{o}C
- Possible to perform addition and subtraction
  - Differences make sense but not ratios
- Examples
  - IQ
  - Fahrenheit

Numeric (ratio) scale - quantitative

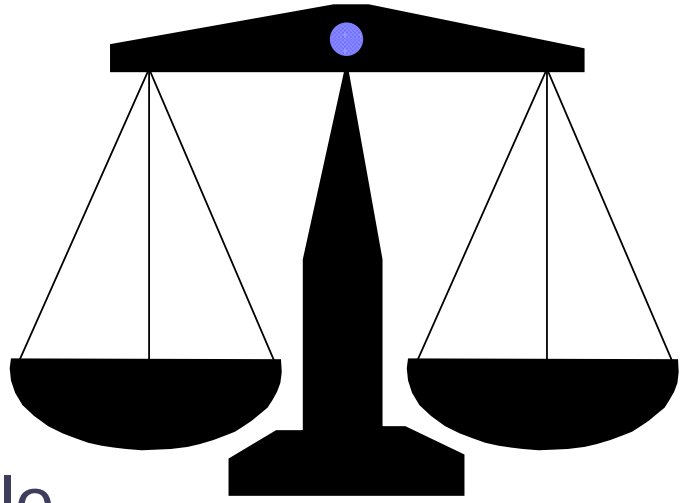
- Values are real numbers
- Ordered with meaningful zero
  - Value of zero can be implied
- Both differences and ratios make sense
- Allows for mathematical calculations +, -, x, /
- Use of numbers in a scale
- Need to define units
- e.g. Weight, processing time

Principles of scale construction

- Definition
  - A **measure** is a formalised procedure to obtain values that express some property or feature.
- Advantages
  - Minimises observer bias
  - Repeatable
  - Ability to combine with other observations



Example – scale



- Mass
  - Numeric scale
  - Arithmetic calculations are possible
  - Zero mass and a standard weight are defined
  - Procedure based on a balance

What is a worthwhile scale?

- It must be *valid*
  - The feature must really change like the scale
- It must be *useful*
  - It must help explain, predict or control
- Remember:
  - We define a variable to represent some property of the real world.
  - A scale is used to measure the variable.
  - The scale must be able to measure the variable and should be valid over time.

How to define a scale

- Specify the underlying model precisely
  - Specify the attribute for property to be modelled.
  - Assign variables to measure the attribute
- Define standards
  - Need to fix values within the scale
  - [if possible] define criteria for 0, 1 (allows for concept of <, >)
- Fix the system of measurement
  - define the possible mathematical relationships
- Look for some good sources for measurement scales published in related literature

Considerations / Implications

- There are some scales which are standard
  - e.g. Richter (earthquakes)
- Some scales are widely used, but not standard
  - e.g. Body Mass Index (BMI) kg/m2: a measure used to determine body fat
- Be careful of competing scales
  - If the scale is not comprehensive it may not measure what you want
  - It should state the necessary conditions under which it can be used
  - There are statistical methods to test the validity of the measures (Construct Validity – e.g.: https://explorable.com/construct-validity;
  - Or http://www.youtube.com/watch?v=2gdorgugwK8 Parts 1, 2, & 3

Case study - Intelligence scales

- What is intelligence?
  - Need to be able to compare intelligence
  - Related to intelligence quotient (IQ)
- So what are the models?
  - Comparison of a subject's performance on a standard series of tasks against the whole population.
    I.Q. = 100 x (mental level) / age
  - Other scales take into account other factors.

The uses of IQ scales

- Designed to identify learning difficulties
- Misuses abound
  - Uncertainties over what tests really measure
    - Cultural bias
  - Pigeon-hole individuals
  - Controversy over misleading differences in measured IQ between sexes, races, cultures
  - Controversial claims about IQ and genetics

- http://psychology.about.com/od/intelligence/a/intelligence.htm

Statistical survey

- A method for collecting quantitative information about items in a population
- May be either exploratory or designed to test a hypothesis

Types of statistical surveys

- Observational
  - Observers collect the required information about the **subjects** (e.g. people who are the subject of surveys)
  - e.g. Traffic survey
- Questionnaire
  - Pre-arranged set of questions that the subject answers
  - Many modalities
    - Phone, internet, mail, personal contact
- Interview
  - Researcher interacts with the subjects
  - Often structured like a questionnaire, but researcher follows-up on answers in an ad hoc manner

Unit of analysis

- “This is the unit—for example individual, household, corporation, or whatever—about which information is required in a research project.”
- “Students often confuse the unit of enquiry (sometimes also called the *unit of analysis*) with the sampling unit. The two need not be the same for any particular study.
- One may sample households, and then collect information (from one or more of its members) about all individual residents. Here, the sampling unit is the household, and the unit of enquiry is the household member”.

Gordon Marshall (1998). “Unit of enquiry." *A Dictionary of Sociology*. Retrieved April 04, 2010 from Encyclopedia.com: http://www.encyclopedia.com/doc/1O88-unitofenquiry.html

Surveys and standards

- Data standards define ...
  - What data are collected
  - How data are collected
  - Associated metadata
  - Format for storing data
  - Quality criteria

Example – medical survey data

- Demography
  - Age, sex, source of notification to register
- Attack
  - Place, transport used
- Timing
  - Delays : calling aid, examination, treatment; reaching hospital
- Medical condition (present)
  - Angina, chest pain, tiredness, breathlessness, palpitation
- Medical history
  - Angina, diabetes, hypertension, infarction, smoker , stroke?
- Clinical details
  - Shock, heart rate, heart rhythm, respiratory rate, blood pressure, weight, height, source of information

Medical example - reviews

- State of patient
  - Alive or dead? If alive, degree of recovery.
- Rehospitalisation
  - How frequently hospitalised? Discharge diagnoses?
  - Days spent in coronary care?
  - Commenced systematic rehabilitation?
- Complications
  - Congestive cardiac failure? Shock? Thromboembolism? Cardiac arrest? Angina? Infarction? other?
- Diagnostic criteria
  - History, electrocardiogram, serum enzymes, autopsy findings, alternative diagnosis made

Issues with survey research

- Sampling
  - Hard to get an unbiased sample
  - Even if you contact a random sample, there is likely to be bias in who responds
- Response rate
  - Hard to get people to complete the survey
- Questions that measure the constructs you seek to investigate can be hard to find and justify
- Priming
  - The order in which questions are asked may influence answers
- Answers are not always accurate
  - Subject may have false beliefs
  - Subject may be dishonest
  - Desirability effects
- Time and labour intensive

Questionnaire modalities

	Phone	Internet	Mail	Personal contact
Time required	fast	fast	slow	slow
Chance for subjects to seek clarification	yes	no	no	yes
Labour required	high	low	low	high
Response rates	declining	low	low	high
Sampling	need phone numbers	need contact method (e.g. email addresses)	need addresses	need to go to subjects
Anonymity	moderate	high	high	low

Questionnaire design

- Be clear
  - Make a plan for your study before you create a questionnaire
  - Avoid ambiguity
  - Make it simple
  - Keep it short
  - Do not use technical terms, jargon
  - Include clear concise instructions
- Be courteous
- Be impersonal
- If you do not expect the respondent to know the exact response, specify so.
  - How much do you estimate you spend on lunch on average?

Questionnaire design

- Avoid
  - Loaded or emotionally charged questions
  - Unnecessary personal questions
  - Leading questions
  - Double-barrelled questions

Loaded / leading / suggestive questions

- Contain implicit assumptions
- Trick a respondent to answer in specific way
- Suggests the answer the interviewer is looking for
- Contains the information the interviewer is looking for
- Contains an implicit assumption of an alleged fact

Examples of loaded / leading / suggestive questions

- Don't you think it is wrong to skip a red traffic light?
- Do you agree that it is wrong to skip a red traffic light?
- Have you stopped beating your partner?
- So you still smoke?
- So you still beat your partner?
- You are a student at Uni B, aren't you?
- Should we stop causing global warming by building brown coal power stations?
- http://www.youtube.com/watch?v=G0ZZJXw4MTA

Questionnaire design

- Organise questionnaire in a logical sequence
- When moving to a new topic – include a transitional sentence
- Begin with a few interesting and non-intimidating items
- Consider first impressions
  - Short
  - Concise
  - Professionally presented

Questionnaire design considerations

- Have a clear research objective
- Test
  - Each question
  - The questionnaire as a whole
  - Send to the pilot settings
- Think ahead
  - what sort of data do you need?
  - how will they be analysed?
  - how much time/resources do you have to collect and analyse these data

Questionnaire design

- Explain the purpose of the study to prospective subjects (people)
- Offer results of your study to respondents
  - and do it!

Sampling

- Researchers want to assess the whole population
- Rare to survey entire population
  - Costly
  - Sampling is very efficient.

Sampling

- Sampling - A method for collecting information from a group [of people] and drawing inferences about a larger population or universe, from the analysis of only part thereof, the sample.
    - Gordon Marshall (1998) "sampling." *A Dictionary of Sociology*. Retrieved April 04, 2010 from Encyclopedia.com: http://www.encyclopedia.com/doc/1O88-sampling.html
- Probability sampling:
  - Random sampling
- Non-probability sampling:
  - Accidental
  - Purposive
  - Snowball sampling

Sampling

- Non-random sampling
  - The researcher cannot ensure that each element in the population will be represented in the sample
  - Makes no pretence of being representative of a population
  - Voluntary samples (only responses of volunteers are used)
  - Quota samples
  - Expert samples
  - Convenience or accidental sampling

Sampling

- Quota sampling
  - Variant of convenience sampling.
  - Respondents from different identified sub-populations in the same proportion as they are found in the general population

    e.g. If there are twice as many Law students at Monash University as IT students, choose 100 Law students and 50 IT students
- Random sampling
  - Utilises some form of random selection
  - Requires a procedure that assures that the different subjects in your population have equal probabilities of being chosen

How big a sample do you need?

- If you are doing statistical hypothesis tests
  - Depends on calculations of statistical power
  - The probability that you will fail to reject the null hypothesis when it is false
- If you are doing exploratory investigations
  - Depends on the error you want to tolerate and the size of the population
  - Use an online sample size calculator

Question types

- Open ended
  - Provide the most information
  - Difficult to interpret
  - Best when exploring new ideas
  - How could this unit be improved?
- Closed
  - Specific type of response required
    - What is your age in years?
    - What is your address?

Multiple choice questions

- Form of closed response
- Minimises data entry errors
- Fast to answer
- Fast to extract answers
- Make sure all options are listed
  - Include ‘other’ option if any chance of additional answers
    - Other
    - Don't know
    - None of the above
    - Often useful to ask respondent to specify with text
- Makes sure that options are mutually exclusive

Numbers

- Values are often coded as numbers

  1 = strongly disagree

  2 = disagree

  3 = neither agree nor disagree

  4 = agree

  5 = strongly agree

- Qualitative values should not be treated like numbers, *even if they are coded as such*
  - It is not meaningful to calculate averages unless values are at least on an interval scale

Likert scale

- Method of eliciting opinions and feelings
- Often used in psychology/management/IS questionnaires
- Typically involve offering a list of ordered responses that range from Strongly Disagree to Strongly Agree, or similar categories
- Ideally symmetric with neutral option in the middle

Agreement	Support	Like
Strongly agree	Strongly support	Like a lot
Agree	Somewhat support	Like a little
Neither agree nor disagree	Neither support nor oppose	Neither like nor dislike
Disagree	Somewhat oppose	Dislike a little
Strongly disagree	Strongly oppose	Dislike a lot

- Forced choice variant removes neutral option
  - http://www.socialresearchmethods.net/kb/scallik.php

Interpretation of Likert scales

- Difficult to make sure that the scale is an *interval* scale
  - i.e., is the difference between *Strongly agree* and *Agree* identical to the difference between *Agree* and *Neither agree nor disagree*
- If not an interval scale then
  - Invalid to take average scores
  - Should analyse the number of responses for each option (Mode, not Median score)

Interpretation of Likert scales (continued)

- Central tendency bias
  - Subjects tend to avoid extremes
- Acquiescence bias
  - Subjects tend to agree
  - Can be avoided by expressing each question twice with meaning of positive and negative responses reversed
- Social desirability bias
  - Subjects tend to provide answers that present themselves favourably

Final thoughts about surveys

- While you might not use surveys for formal research
  - Useful tool for gathering information in many contexts
    - Especially given online tools
    - e.g. http://www.surveymonkey.com/
  - Understanding survey methods helps you assess survey results published in the research papers

New opportunities in data collection

- Advances in IT are changing many aspects of scientific method
  - Data rich environment
  - New kinds of studies are possible
    - e.g. data mining
- Data repositories
  - Aggregate data from many sources
  - “Recycling” data
  - Require uniform data formats and standards
  - Require metadata to identify data

Online data resources

- Online data repositories are now common
- Increasingly important in many fields
- Some studies now rely on such data alone
- Many fields now require data sharing and reuse
  - Compulsory for grants and publication in some fields
- Need to check and conform to standards in data gathering, storage and release

Online data resources

- Examples
- Astronomy
  - International star atlas
  - Compiles all observations
- Biotechnology
  - Compile DNA and protein sequences
  - Enables cross-references of new sequences
  - Entirely new kinds of study are possible

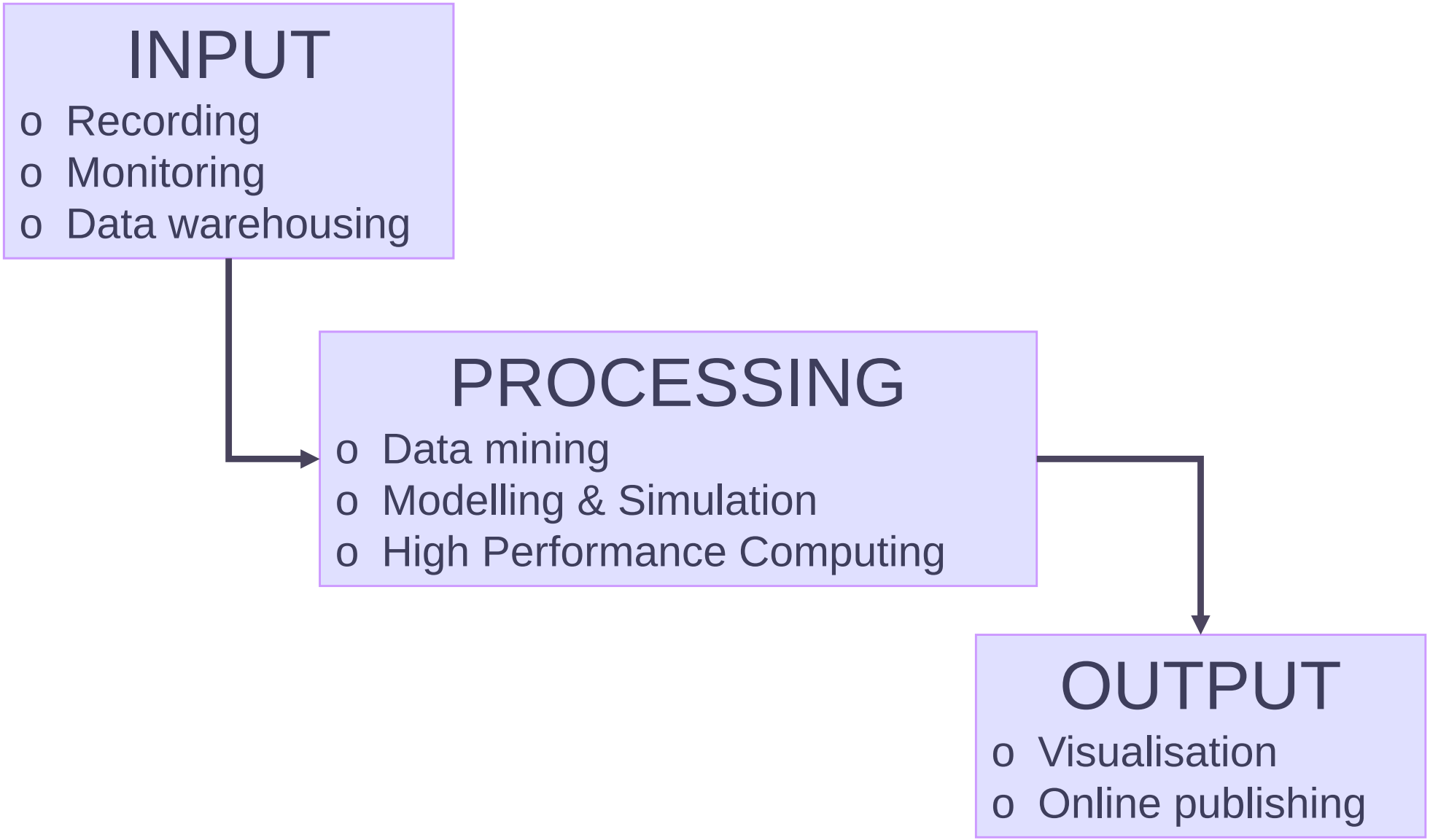
Examples of online repositories

- Australian Bureau of Statistics
  - http://www.abs.gov.au/
- Australian Department of the Environment
  - http://www.environment.gov.au/
- U. of C. (Irvine) Machine Learning Repository
  - http://archive.ics.uci.edu/ml/
- Genbank
  - http://www.ncbi.nlm.nih.gov/Genbank/index.html
- Statistical Reference Datasets
  - http://www.itl.nist.gov/div898/strd/
- The Stony Brook Algorithm Repository
  - http://www.cs.sunysb.edu/~algorith/
- SourceForge
  - http://sourceforge.net/

e-Research

- New paradigm affecting all fields of research
- Using Information and Communication Technology (ICT) to underpin research activity
- Concerns all stages and aspects of research
- Enabling developments
  - Increasing processor speed
  - Storage capacity
  - Networking and communications
  - Input devices
  - Software

e-Research activities



e-Research @ Monash

- http://www.monash.edu.au/eresearch/
- Examples of Projects
  - TARDIS: online data management
  - MASSIVE - Multi-modal Australian ScienceS Imaging and Visualisation Environment
  - Geographic Information Systems
  - Australian Video Art Archive
  - General Equilibrium Modelling Software
  - Bioreactor Modelling
  - Securing Wireless Sensor Networks
  - Bioinformatics in Protein Crystallography
  - Molecular Dynamics Simulation

Some national e-Research resources

- SAAP National Data Collection Agency
  - http://www.aihw.gov.au/housing/sacs/ndca/index.cfm
- National Statistical Service
  - http://www.nss.gov.au/
- A National Data Architecture for Australian Research
  - > http://www.mmv.vic.gov.au/Assets/536/1/ANationalDataArchitectureforAustResearch.pdf
- Australian Research Online
  - http://research.nla.gov.au/
- Australian government public information datasets
  - http://data.australia.gov.au/
- Government data
  - > USA - www.data
  - > the European Union data portal
    - https://open-data.europa.eu/en/data/

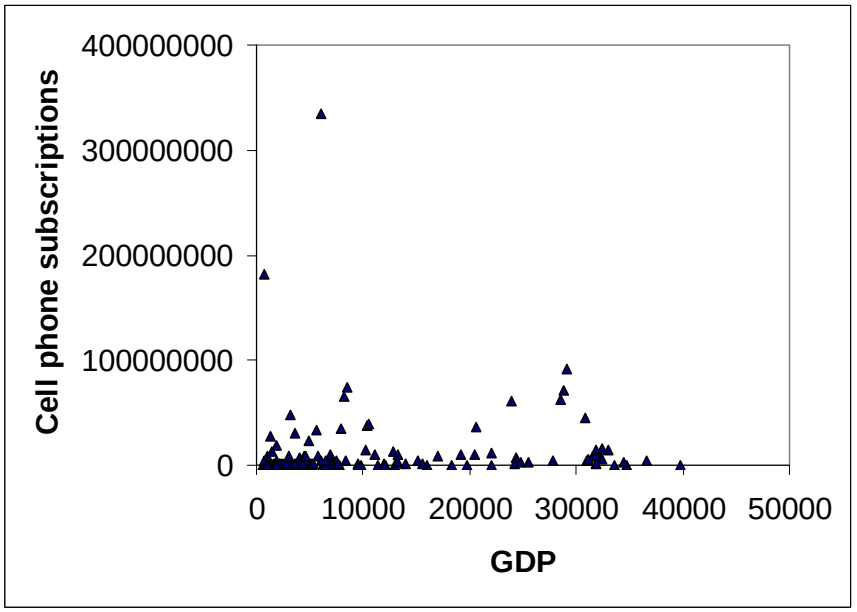
Some national data repositories

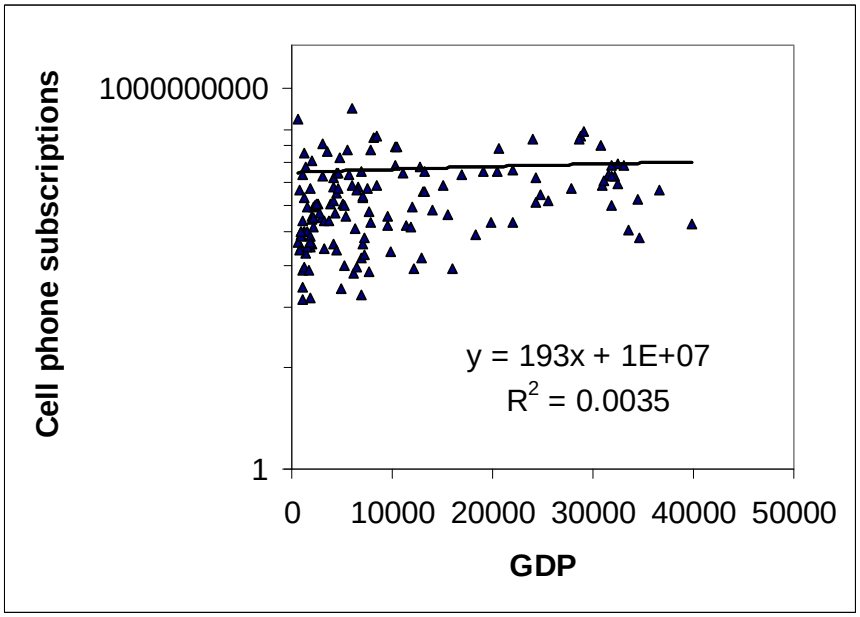
- ABS Australian Bureau of Statistics
  - http://www.abs.gov.au/
- Stock Exchange
  - http://www.asx.com.au/
- Bureau of Meteorology
  - http://www.bom.gov.au/
- Most Government departments

Using data resources – a simple example

- Research Question
  - Do countries with higher Gross Domestic Product (GDP) have more cell (mobile) phone subscribers?
- Approach
  - Test for correlation between totals for different countries
  - Data: United Nations (144 countries in 2004)
  - Use a spread sheet to plot data and fit a regression line

Plots of cell phone subscriptions against GDP for 144 countries





Raw Data

Logarithmic Scale

Conclusion: No significant correlation found

Question: How else might this study be done?

Useful references

- Fowler, F.J. (1993). *Survey Research Methods.* 2nd ed. Sage Publications, Newbury Park, CA.
- Galliers, R., ed. (1991). ‘ Choosing information systems research approaches’. In Robert Galliers (Ed.) (1992) *Information systems research: Issues, methods and practical guidelines.* (Ch.8). Oxford: Blackwell Scientific.
- Leedy, P.D. with contributions by Newby, T.J. and Ertmer, P.A. (1997). *Practical Research: Planning and Design.* 6th ed. Upper Saddle River, NJ: Merrill. [Chapters on survey method].
- Zikmund, W. G. (1997) *Business Research Methods.* 5th ed. Dryden Press, Fort Worth, TX.

Some recommended reading

- The Skeptics Dictionary
  - http://skepdic.com/ , especially entries on
- Science & Philosophy
  - http://www.skepdic.com/tiscience.html
- Critical Thinking
  - http://www.skepdic.com/ticriticalthinking.html
- Essay on Critical Thinking
  - Facione, P. (2007). Critical Thinking: What It Is and Why It Counts. Insight Assessment. http://www.insightassessment.com/pdf_files/what&why2006.pdf