

Research Skills for Mental Health

Managing and Reporting Data

Stacey Coleman

Session Outline

- **Types of data**
- **Descriptive statistics**
 - Central tendency
 - Introduction to normal distribution
- **Frequencies**
- **Displaying data**

2

RECAP: Variable

- The term **VARIABLE** is used to describe the empirical object or phenomenon we are observing
- Two main types of variable in quantitative research:
 - Independent variable
 - Dependent variable

3

IV's and DV's

- **Dependent variable:**
- The factor/variable that is assumed to be directly affected by changes in the IV in an experiment
- The dependent variable DEPENDS UPON the independent variable

4

IV's and DV's

- **Independent variables:**
- Variable which the experimenter manipulates in order to determine the effect on a given behaviour (i.e., the effect on the chosen dependent variable)
- NOT influenced by any other variable in your study

5

IV's and DV's

When choosing IV's and DV's, we need to consider two main things:

1. The **design** of the study
2. The type of data we therefore need to collect

6

Design: A basic introduction

- Study design refers to a number of different things
 - Think back to week 1 – types of quantitative design include non-experimental and experimental
- With experimental designs, we (more often than not) use IV's and DV's.
- Categorical IV's need to be defined in terms of how they are being classified/ categorised...

7

Design: A basic introduction

- Two main study design types for you to be concerned with...
- **BETWEEN SUBJECTS**
- **WITHIN SUBJECTS**

8

Design: A basic introduction

BETWEEN SUBJECTS

- Different people in all groups/conditions
- Any one participant can **ONLY** belong to one group/condition

9

Design: A
basic
introduction

BETWEEN SUBJECTS

- E.g., IV of gender; male OR female
- Any participant selected can ONLY belong to the male OR the female group

10

Design: A
basic
introduction

WITHIN SUBJECTS

- Participants take part in all groups/conditions
- ALL participants belong to ALL groups/conditions

11

Design: A
basic
introduction

WITHIN SUBJECTS

- E.g., IV of response to stimuli; medical pictures AND medical terms
- ALL participants would see the medical pictures AND the medical terms

12

TYPES OF DATA

Quantitative Methods

- Quantitative methods use numbers as data
- *WHY do we, and HOW can we, use numbers.....?*

14

Why use numbers?

- Numbers add value to what we are saying...
- Which of the below sounds better:
 1. Many students from UWTSD have a part time job
 2. 75% of students from UWTSD have a part time job

15

Can we trust numbers?

- Many people have strong opinions about numbers, especially statistics...
- Numbers don't "PROVE" anything..
 - Correct – they don't..
 - But as psychologists we don't *want* to PROVE anything, either...

16

Where do numbers fit?

- When we examine numerical data, its relevant largely depends on what is being asked of your participants...
- Ask yourself:
 - Does this provide a useful insight?
 - Does it indicate outliers that may need further investigation?
 - Does it help me spot any trends?
 - Does it affect how you should think about a topic area?

17

TASK!

What can you think of that can be best represented in numerical terms?



Types of data

- Two broad types:
- Categorical**
 - E.g., gender, personality type
 - Limited number of groups to which people can belong to, or can only belong to one group
- Continuous**
 - E.g., age, scores on a scale, seconds taken to complete a task
 - Can take a vast range of values, often from zero to an infinite number

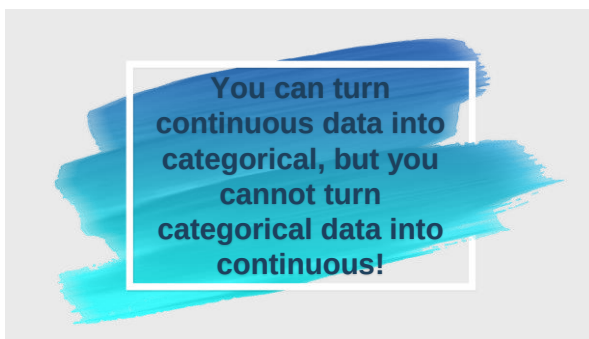
Groups

Numbers

19



TYPES OF DATA TASK



You can turn continuous data into categorical, but you cannot turn categorical data into continuous!

Descriptive Statistics

RECAP: Report assessment

- You must calculate and report appropriate descriptive statistics for the data provided, and write up your final report in accordance with APA conventions for report writing in psychology.

23

Experiments and Studies

- **We measure a sample of a population on some factor**
 - Intelligence, cancer remission, face recognition; anything you can think of!
- **It's usually impractical to measure an *entire* population!**
 - Too big, too expensive, not really needed

24

Descriptive Statistics

Using the individual pieces of data from each participant

- Scores on individual tasks for each participant
 - Number of questions answered correctly on IQ test
 - Level of anxiety felt
 - Reported narcissism scores

25

Descriptive Statistics

- Produces 'raw' data
 - And sometimes A LOT of it!
 - We need a way to reduce this to a manageable level

26

Descriptive Statistics: Presentation

Participants (in your method section)

- Number of, age range, number of males/females, average age of groups
 - This is standard for ANY study!

E.g., the sample used for this study consisted of 30 (10 male, 20 female) psychology students who ranged in age from 18 to 41 ($M = 21.50$, $SD = 5.60$)

27

**Descriptive
Statistics:
Presentation**

Results

- Descriptive statistics of the dataset
 - Central tendency of the variable being measured (if/where applicable)
 - Remember: no measure of central tendency for categorical data!
 - Dispersion of scores around that mean
 - Variance and standard deviation from the average score (if applicable)

28

**CENTRAL
TENDENCY??**

**Central
Tendency**

- People will score differently on almost ANY measure
 - Individual differences; error; fatigue

30

Central Tendency

- If I test a group of 100 people for their IQ, some will score more or less than others. This could be because they are more or less intelligent, more or less tired, or have more or less things playing on their mind!

31

Central Tendency

- If I test a group of 100 people for their IQ, some will score more or less than others. This could be because they are more or less intelligent, more or less tired, or have more or less things playing on their mind!
- So, we need to find the 'average' of all the participants in the group

32

Central Tendency

- *Three main ways to measure central tendency:*
- 1) **MODE**
 - The most commonly occurring score
- 2) **MEDIAN**
 - The middle score when lined up in rank order (smallest to largest)
- 3) **MEAN**
 - The sum of all scores divided by the number of participants

33

The Mode

Take all eligible responses (scores, reaction times, etc)

10, 5, 8, 12, 8, 3, 9, 18, 7

The most commonly occurring score:

10, 5, 8, 12, 8, 3, 9, 18, 7

Mode = 8

34

The Mode

Issues with the mode?

10, 3, 8, 12, 8, 3, 9, 18, 7

Non-modal!

Bimodal!

10, 3, 8, 12, 6, 2, 9, 18, 7

35

The Mode

Issues with the mode?

10, 3, 8, 12, 8, 3, 9, 18, 7

Only useful for certain types of data!

10, 3, 8, 12, 6, 2, 9, 18, 7

36

The Median

- Put in rank order

10, 5, 8, 12, 8, 3, 9, 18, 7

3 5 7 8 8 9 10 12 18

Central tendency is the middle number

Median = 8

37

The Median

- But what about even numbers of data?*

3 5 7 8 8 9 10 12 14 18

Add the numbers and divide by two
 $(8+9) / 2 = 8.5$

Median = 8.5

38

The Median

- Major problem with the median...*

1 1 1 2 5 7 8 9 18

- Takes no account of the other scores! Purely about the centre

39

The Mean

- The measure of arithmetic mean provides us with a powerful description of a large set of data
- Takes into account each individual 'data point' within the sample
- BUT – can be a weakness if a very extreme score is present

40

The Mean

E.g., 10 12 1 9 8 500

(mean = 133)

- Larger samples reduce this problem by weakening the effect of each data point individually.

41

The Mean

Summed total divided by the number within the sample (N)

- For a single participant (over several trials)
- For a number of participants (sample)
- We most often look for the mean of all of participants scores in a group

42

Research and Data

How many participants???

- A representative sample...
- We can have an extreme participant in a sample
- The effect of an extreme participant is reduced by a larger sample

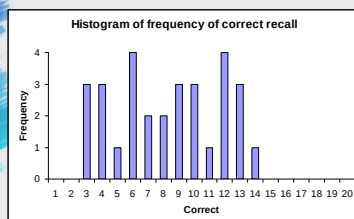
43

Displaying Data

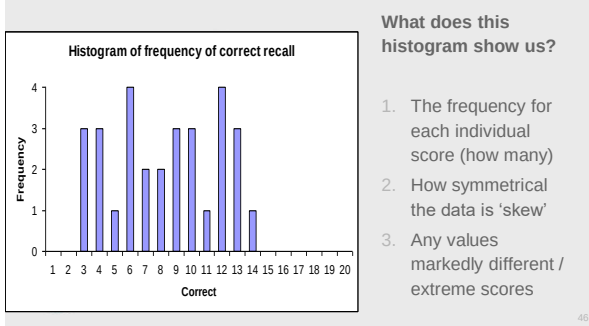
- First descriptives can consider the 'frequencies' of certain scores
- How many times each score occurs in a sample
- Can represent this graphically by plotting out the 'frequency' for each score in a 'histogram'

44

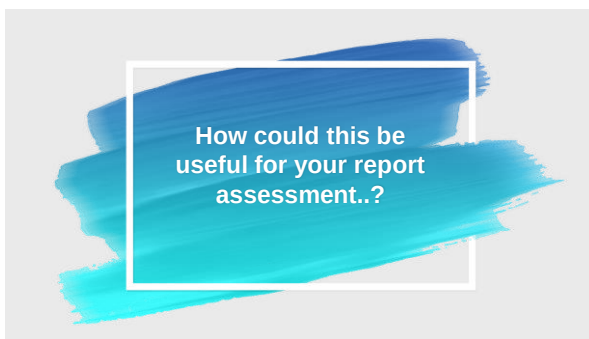
Histogram



45





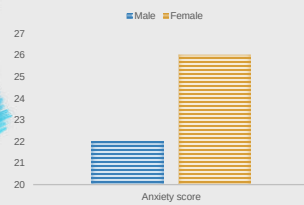


Displaying Data

- Can also use other forms of chart to display numerical data...
- Bar chart,
- Scatter graph,
- Pie chart

49

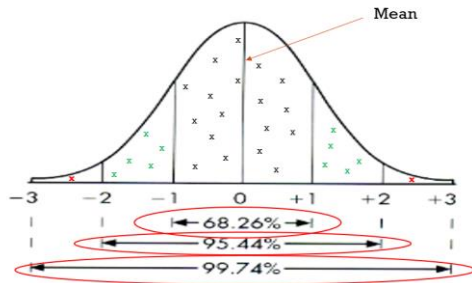
Displaying Data



Visual representation of average scores obtained

50

Normal Distribution



52

Means

- Just knowing the mean of a group is not enough
- Group A: $M = 11$
- Group B: $M = 10$

Is the difference between these two groups down to chance? Or is it a reliable effect of something?

- How many people have we tested?
- What is the variation of the scores in the groups?

53

ANY
QUESTIONS?
