Lancaster University

# PSYC214: Statistics
# Lecture 1 – Measurement, variance and inferential statistics

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

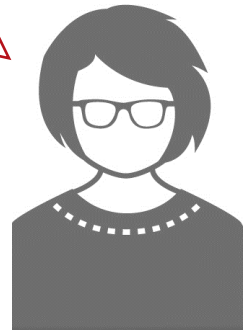
# Lecture 1 – Measurement, variance and inferential statistics

## Agenda/Content

- Experimental science
- Variables
- Descriptive statistics
  - Levels of measurement
  - Measures of central tendency
  - Measures of variability
- Distributions
- Inferential statistics and hypotheses
- Within and between participant designs

# Controlled experiment

A scientific investigation in which both the control group and experimental group(s) are kept under similar conditions apart from the factor under study, so that the effect of influence of that factor can be identified or determined.



# Experimental science

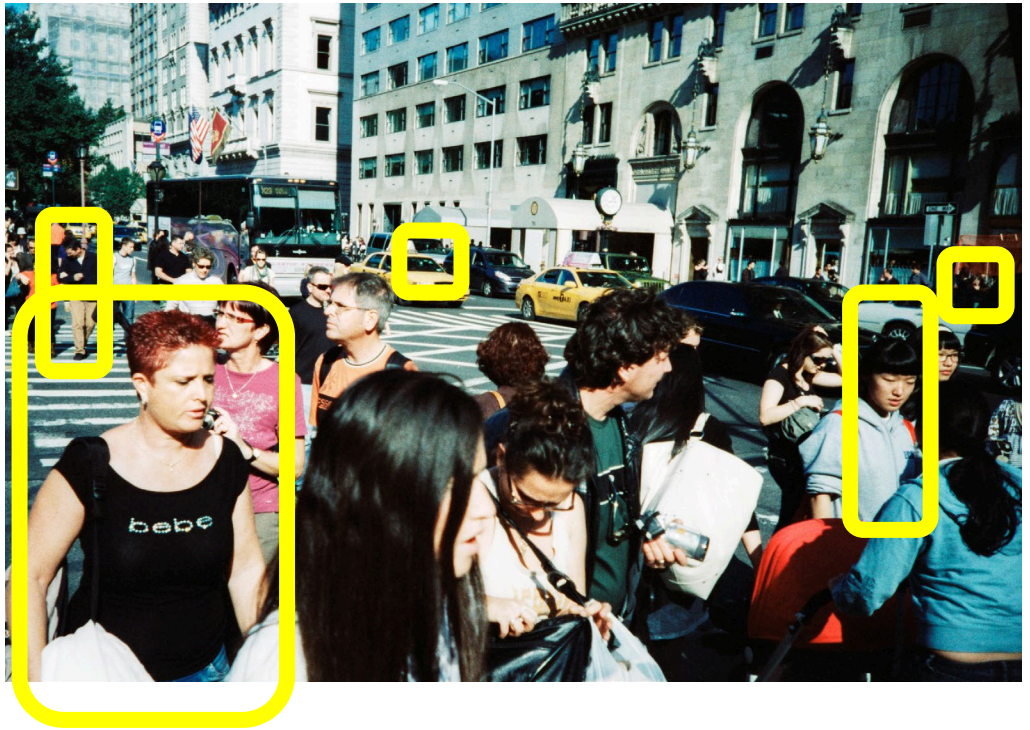
*Population versus sample*

- Population is every individual you are interested in

# Experimental science

*Population versus sample*

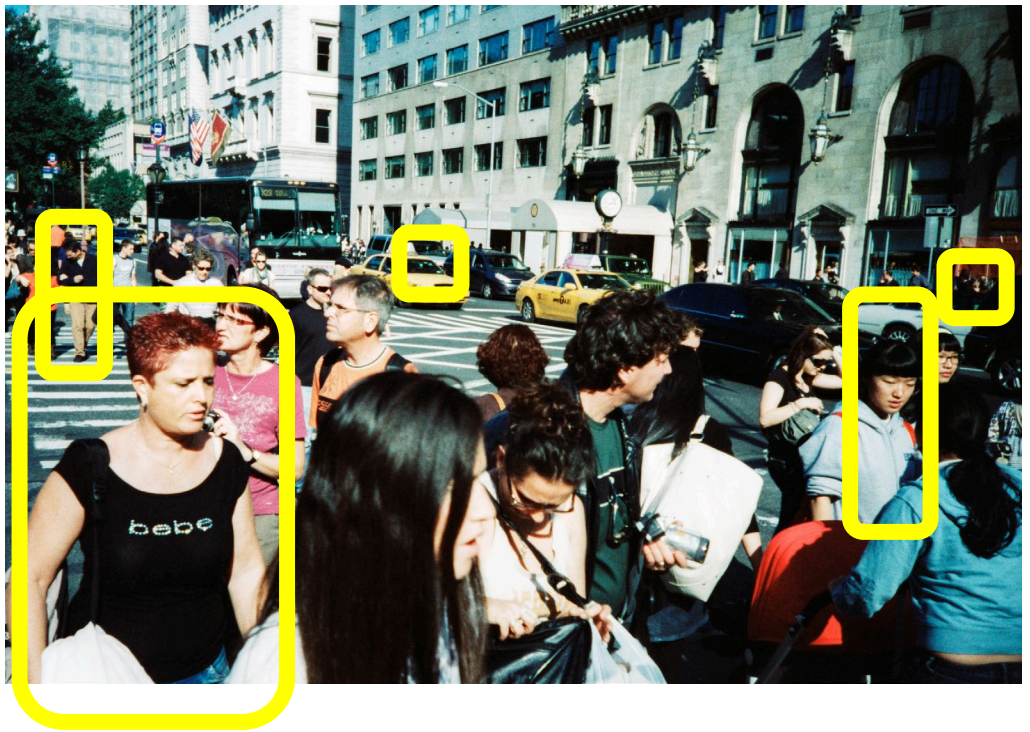
- Population is every individual you are interested in
- The sample is a subset of your population of interest. We examine samples because it is typically impossible to sample everyone in the population



# Experimental science

*Population versus sample*

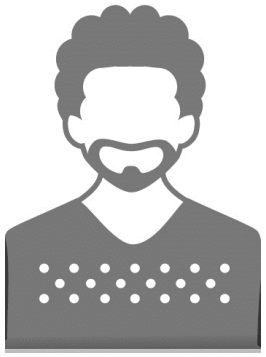
- You should always opt for random sampling, where you pick your sample randomly
- However, in reality, we often use opportunity sampling where we recruit who we have access to



# Variables

*Independent Variable*

- The variable (FACTOR) the experimenter manipulates or changes, which may be assumed to have a direct effect (i.e., influences change) on the dependent variable.



*Dependent Variable*

- The outcome of interest. It is the variable being tested and measured in an experiment. It is 'dependent' on the effect (i.e., influence) of the independent variable.

# Independent variable

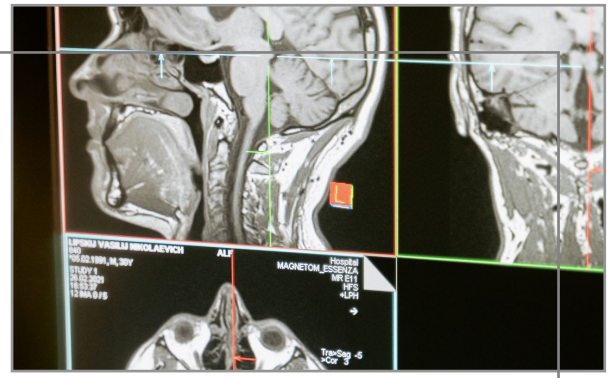
Type of Treatment

Treatment factor

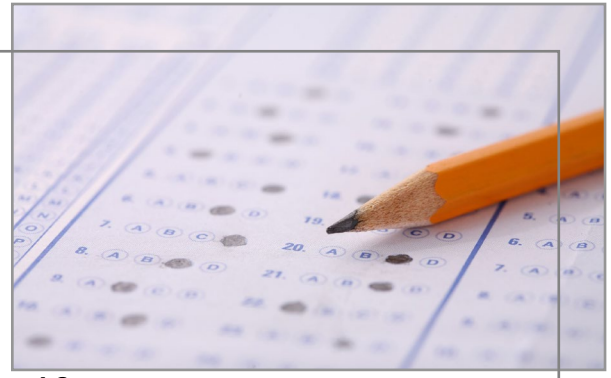
Experimental manipulation

# Dependent variable – i.e., outcome

Behavioural measure



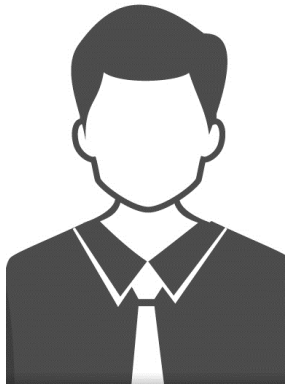
Physiological measure



Self-report measure

# Statistics

- Use descriptive statistics to describe characteristics and tendencies of your sample
- Use inferential statistics to determine whether the performance and characteristics of your sample generalizes to the population



Lancaster University

# PSYC214: Statistics
# Lecture 1 – Measurement, variance and inferential statistics – Part 2

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

# Descriptive statistics

1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

# Descriptive statistics

Lancaster University

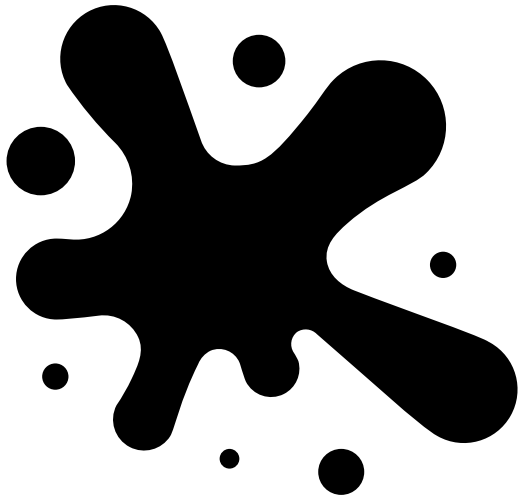
1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

# 1. Levels of measurement

*Nominal, Ordinal, Interval, Ratio*

# 1. Levels of measurement

*Nominal, Ordinal, Interval, Ratio*



# 1. Levels of measurement - Types

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | ● | | | |
| Rank or order | | | | |
| Known and proportionate intervals | | | | |
| True zero | | | | |

# 1. Levels of measurement - Types

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | ● | ● | | |
| Rank or order | | ● | | |
| Known and proportionate intervals | | | | |
| True zero | | | | |

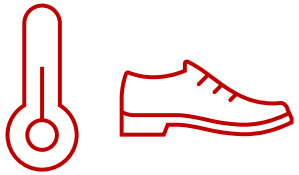
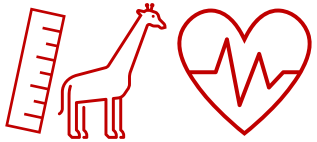
# 1. Levels of measurement - Types

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | ● | ● | ● | |
| Rank or order | | ● | ● | |
| Known and proportionate intervals | | | ● | |
| True zero | | | | |

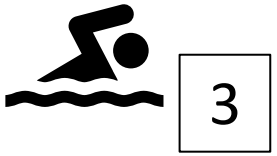
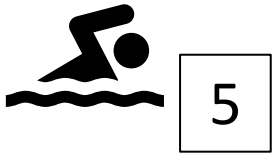
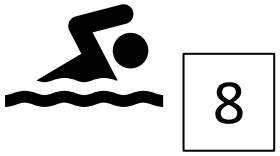
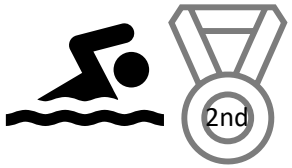
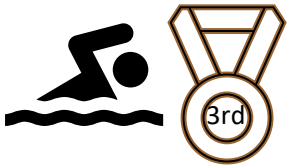
# 1. Levels of measurement - Types

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | ● | ● | ● | ● |
| Rank or order | | ● | ● | ● |
| Known and proportionate intervals | | | ● | ● |
| True zero | | | | ● |

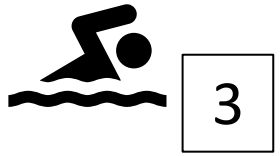
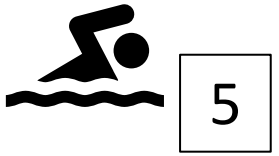
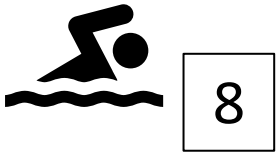
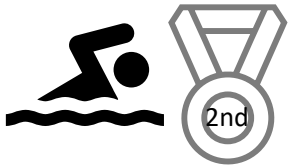
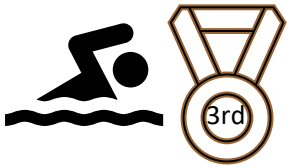
# 1. Levels of measurement - Examples

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names |  | | | |
| Rank or order | |  | | |
| Known and proportionate intervals | | |  | |
| True zero | | | |  |

# 1. Levels of measurement - Examples

| | | | | |
|---|---|---|---|---|
| Nominal | Numbers assigned to an athlete |   |   |   |
| Ordinal | Rank order of winners |   |   |   |
| Interval | Performance rating on a scale (0-10) | 9.8 | 9.6 | 8.6 |
| Ratio | Time to finish in seconds | 21.07 | 21.14 | 21.79 |

# 1. Levels of measurement - Examples

| | | | | | |
|---|---|---|---|---|---|
| Categorical | Nominal | Numbers assigned to an athlete |   |   |   |
| | Ordinal | Rank order of winners |   |   |   |
| Quantitative | Interval | Performance rating on a scale (0-10) | 9.8 | 9.6 | 8.6 |
| | Ratio | Time to finish in seconds | 21.07 | 21.14 | 21.79 |

# 1. Levels of measurement - Examples

Lancaster University

| | | | | | |
|---|---|---|---|---|---|
| Categorical | Nominal | Numbers assigned to an athlete |  |  |  |
| | Ordinal | Rank order of winners |  |  |  |
| Quantitative | Interval | Performance rating on a scale (0-10) | 9.8 | 9.6 | 8.6 |
| | Ratio | Time to finish in seconds | 21.07 | 21.14 | 21.79 |

Lancaster University

# PSYC214: Statistics
# Lecture 1 – Measurement, variance and inferential statistics – Part 3

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

# Descriptive statistics

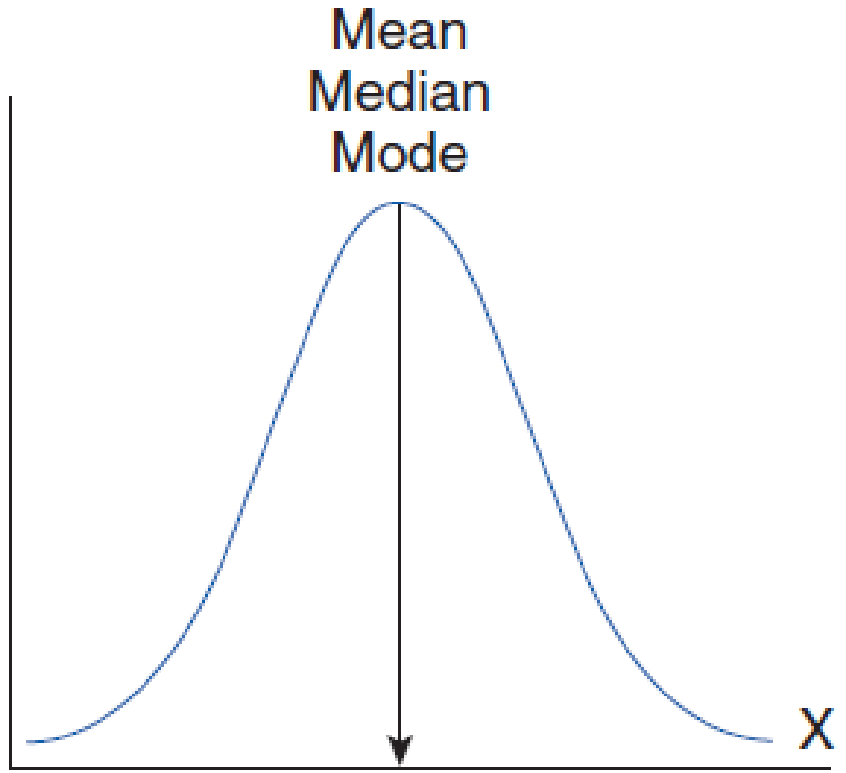
Lancaster University

1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

# 2. Measures of central tendency

*A single value that describes the way in which a group of data clusters around a central value, i.e., the centre of the data set*

- There are three measures of central tendency
  - Mode
  - Median
  - Mean



# 2. Measures of central tendency

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | Mode, % frequencies | | | |
| Rank or order | | | | |
| Known and proportionate intervals | | | | |
| True zero | | | | |

# 2. Measures of central tendency

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | Mode, % frequencies | Mode, % frequencies | | |
| Rank or order | | Median, percentile | | |
| Known and proportionate intervals | | | | |
| True zero | | | | |

# 2. Measures of central tendency

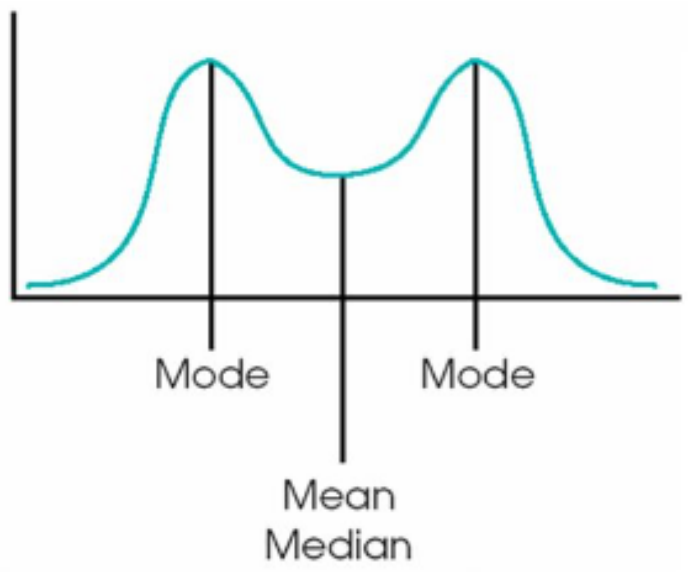
Lancaster University

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Categories, Names | Mode, % frequencies | Mode, % frequencies | Mode, % frequencies | Mode, % frequencies |
| Rank or order | | Median, percentile | Median, percentile | Median, percentile |
| Known and proportionate intervals | | | Mean, standard deviation | Mean, standard deviation |
| True zero | | | | All above |

# 2. Measures of central tendency - Mode

*The most frequent score/data*

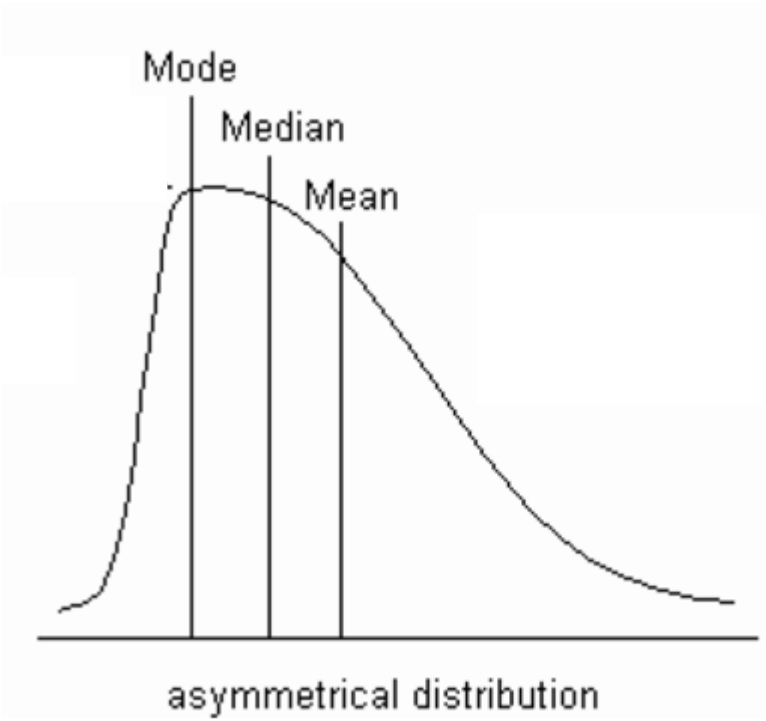
- Level of measurement: Nominal, ordinal or interval/ratio
- Shape of distribution: Bimodal or multimodal



# 2. Measures of central tendency - Median

*The middle number when data are ordered*

- Level of measurement: Ordinal or interval/ratio
- Shape of distribution: Highly skewed



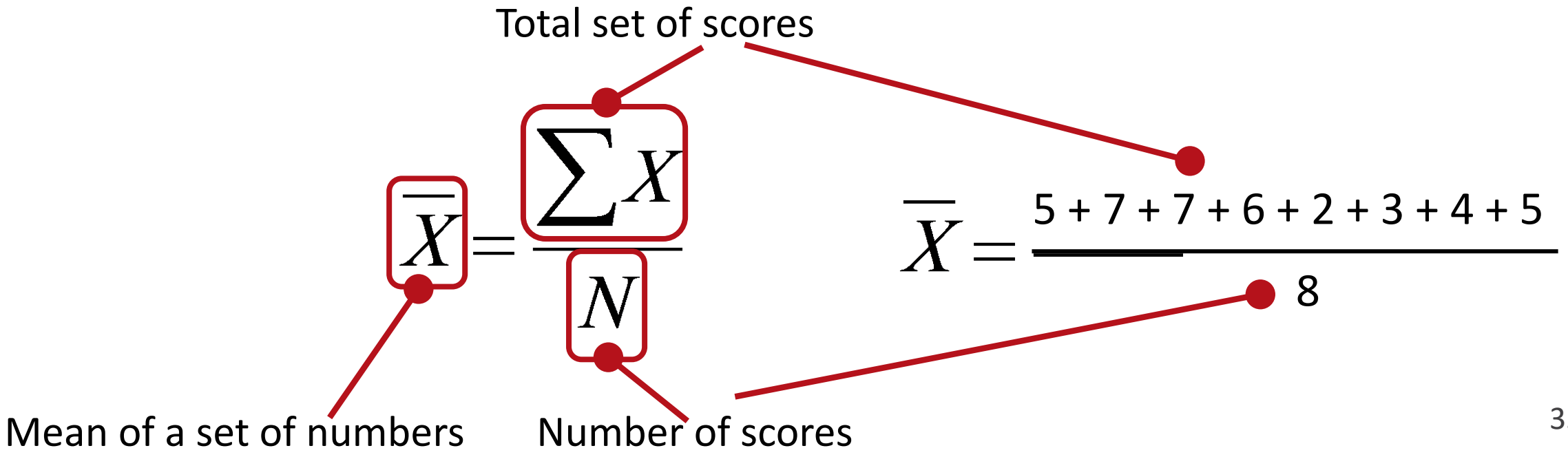
# 2. Measures of central tendency - Mean ($\bar{X}$)

*The average, is the sum (Σ) of all scores (x) divided by the number of scores (N)*

$$\overline{X} = \frac{\sum X}{N}$$

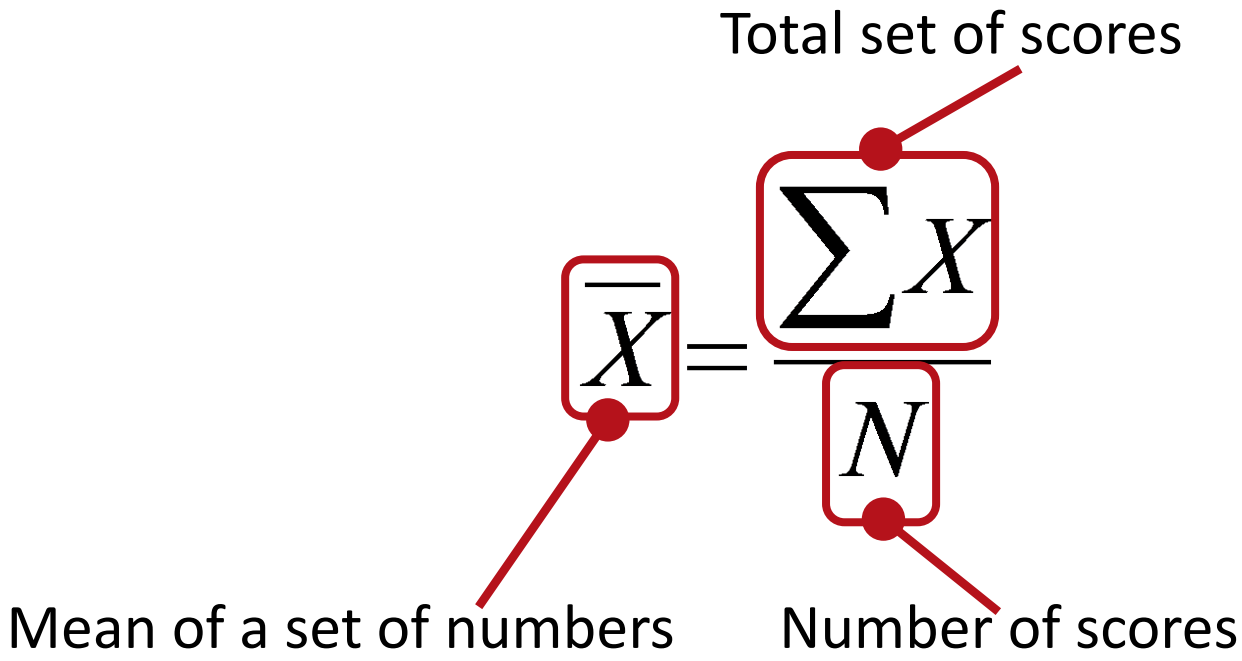
# 2. Measures of central tendency - Mean ($\bar{X}$)

*The average, i.e., the sum (Σ) of all scores (x) divided by the number of scores (N)*



# 2. Measures of central tendency - Mean ($\bar{X}$)

*The average, i.e., the sum (Σ) of all scores (x) divided by the number of scores (N)*




$\overline{X} = 4.875$

# PSYC214: Statistics
# Lecture 1 – Measurement, variance and inferential statistics – Part 4

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

# Descriptive statistics

Lancaster University

1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

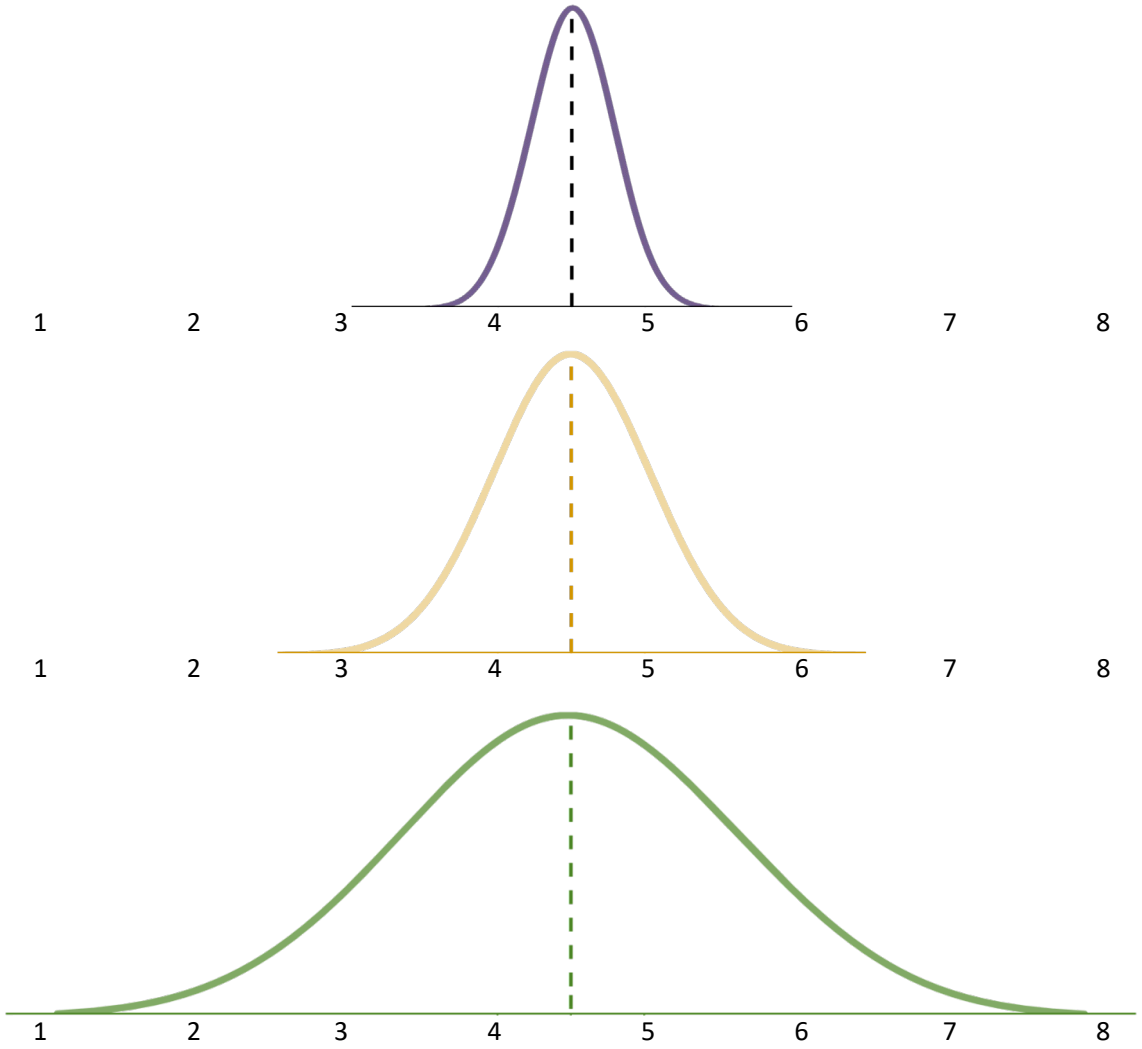
# 3. Measures of variability

Lancaster University

*The spread or dispersion of scores in relation to the midpoint of data.*

- Range
- Sum of squares
- Variance
- Standard deviation

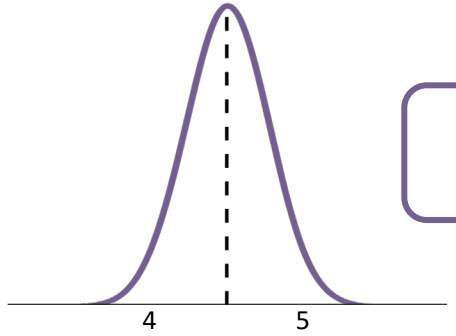
# 3. Measures of variability - why care?



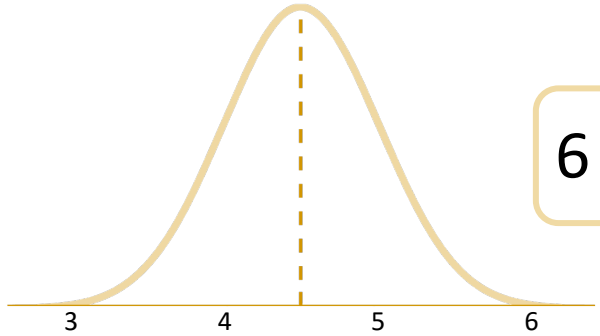
# 3. Measures of variability - range

*The difference between the highest and lowest score*

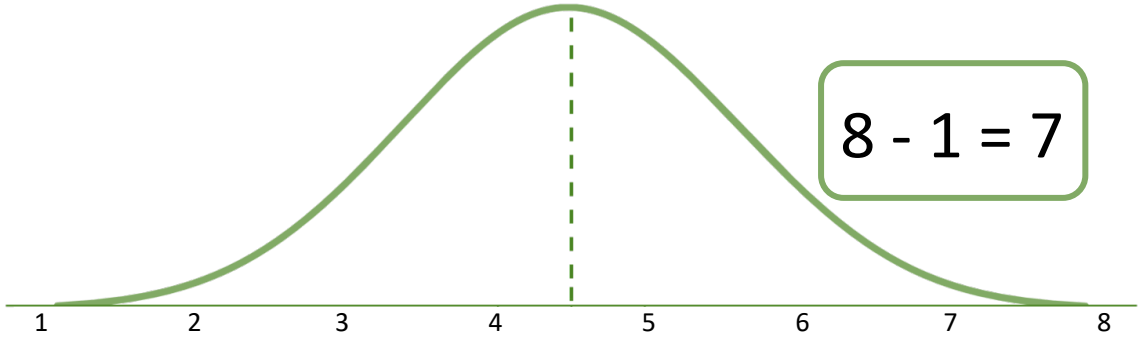
- Subtract the lowest value in the distribution by the highest value




5.5 - 3.5 = 2



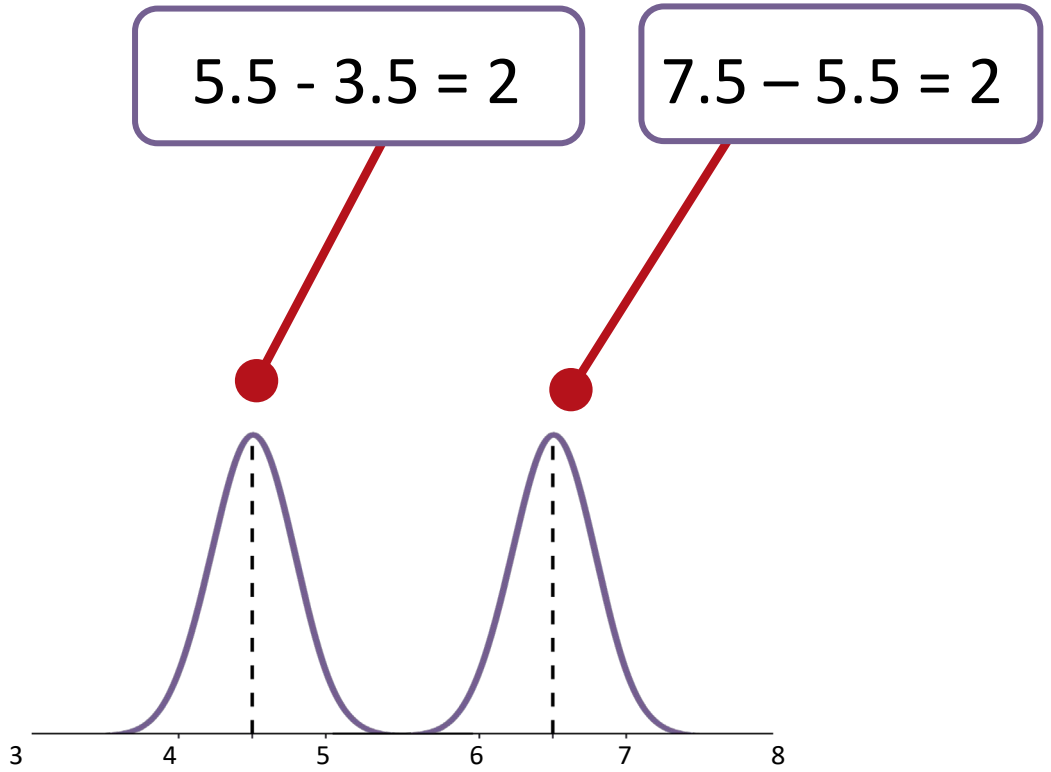

6 - 3 = 3




8 - 1 = 7

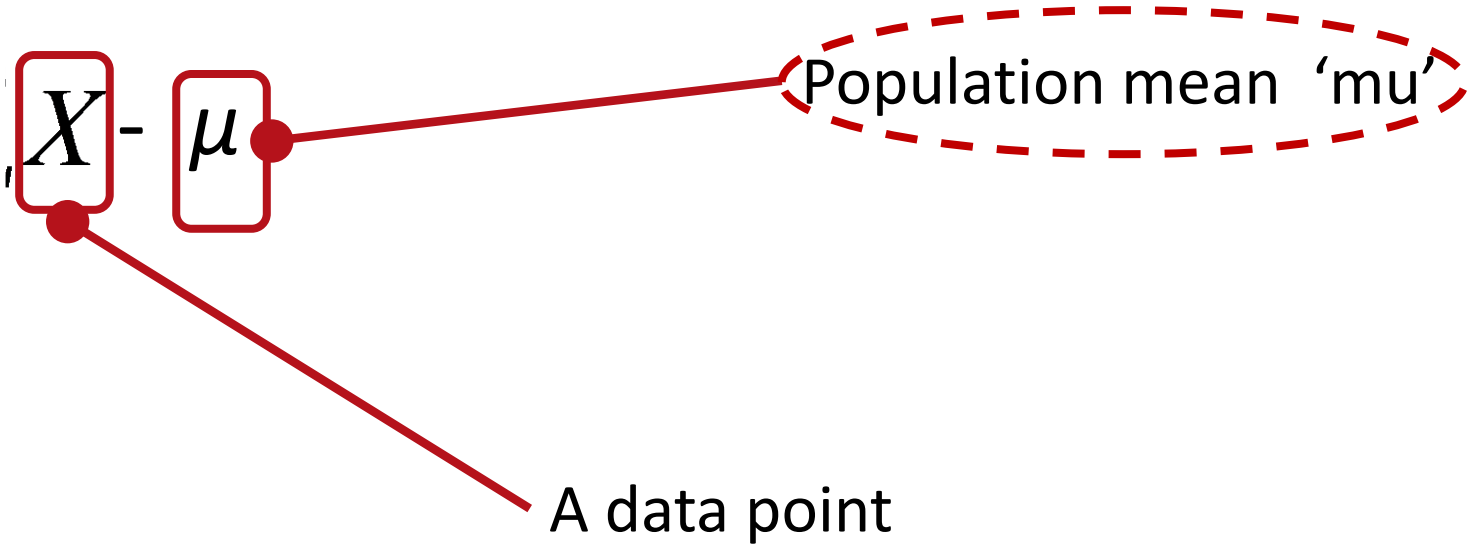
# 3. Measures of variability - range

*When is it not useful?*



# 3. Measures of variability - sum of squares

1. Calculate difference

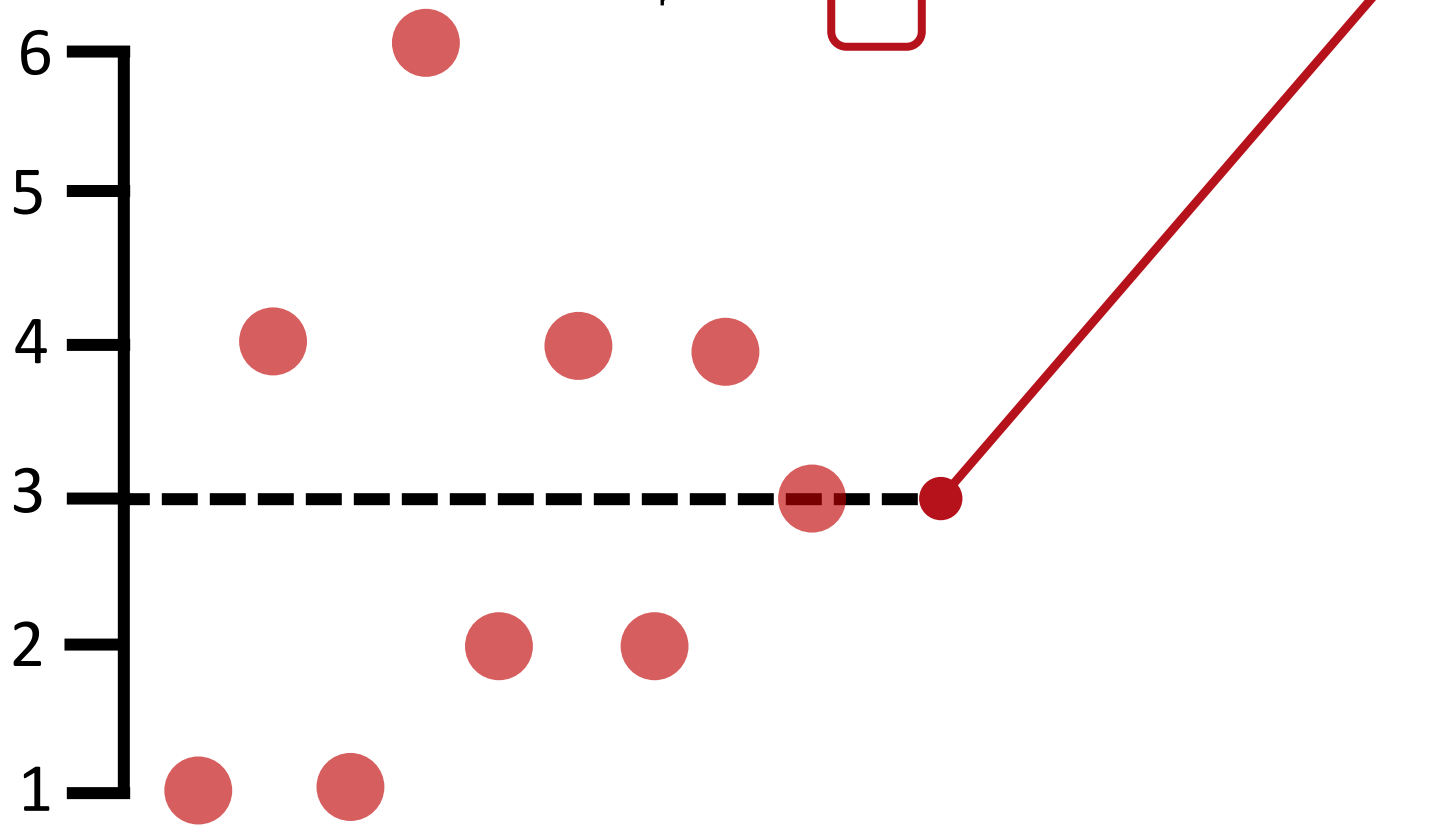


# 3. Measures of variability - sum of squares

1. Calculate difference

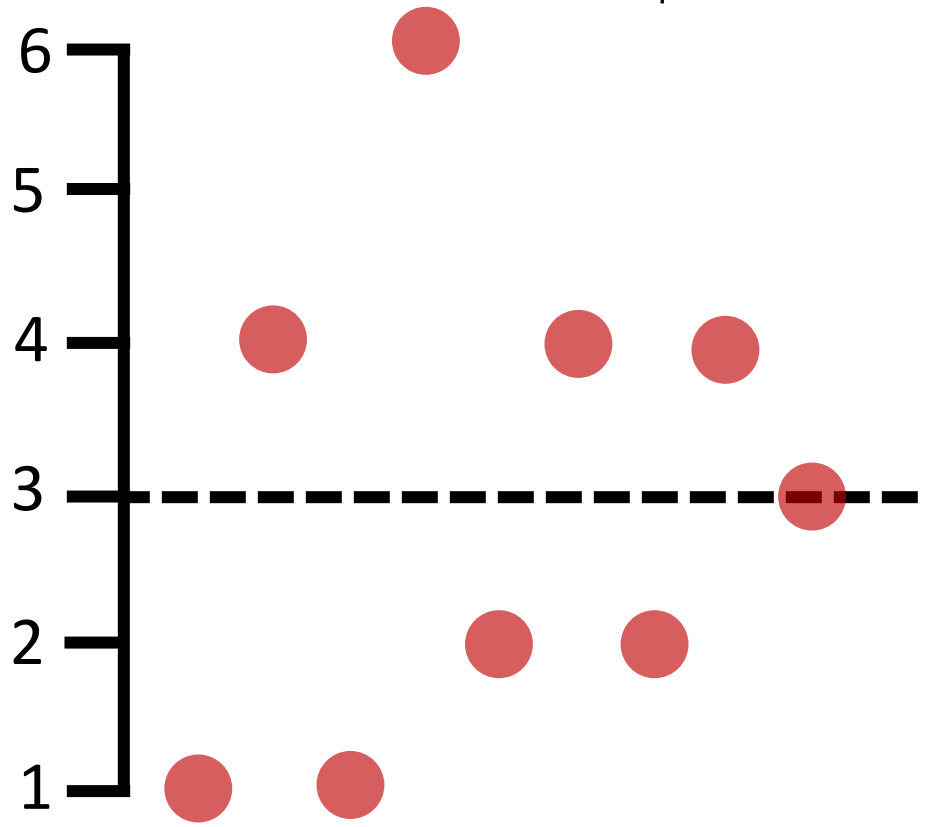
$X - \mu$

Population mean



# 3. Measures of variability - sum of squares

1. Calculate difference $X - \mu$




| Data point | χ - μ |
|---|---|
| χ1 | -2 |
| χ2 | 1 |
| χ3 | -2 |
| χ4 | 3 |
| χ5 | -1 |
| χ6 | 1 |
| χ7 | -1 |
| χ8 | 1 |
| χ9 | 0 |
| Total | 0 |

# 3. Measures of variability - sum of squares

Lancaster University

1. Calculate difference $X - \mu$
2. Calculate the sum of squares

$$\text{Sum of squares (SS)} = \sum(\mu - x_i)^2$$

is the sum of all data

is the population mean

is each data point

| Data point | χ - μ | (χ - μ)² |
|---|---|---|
| χ1 | -2 | 4 |
| χ2 | 1 | 1 |
| χ3 | -2 | 4 |
| χ4 | 3 | 9 |
| χ5 | -1 | 1 |
| χ6 | 1 | 1 |
| χ7 | -1 | 1 |
| χ8 | 1 | 1 |
| χ9 | 0 | 0 |
| Total | 0 | 22 |

# 3. Measures of variability - variance

- Variance: Average deviation around the mean of a distribution (average of sum of squares)

$$\text{Variance } (\sigma^2) = \frac{\sum(\mu - x_i)^2}{n-1}$$

Sum of squares

Degrees of freedom

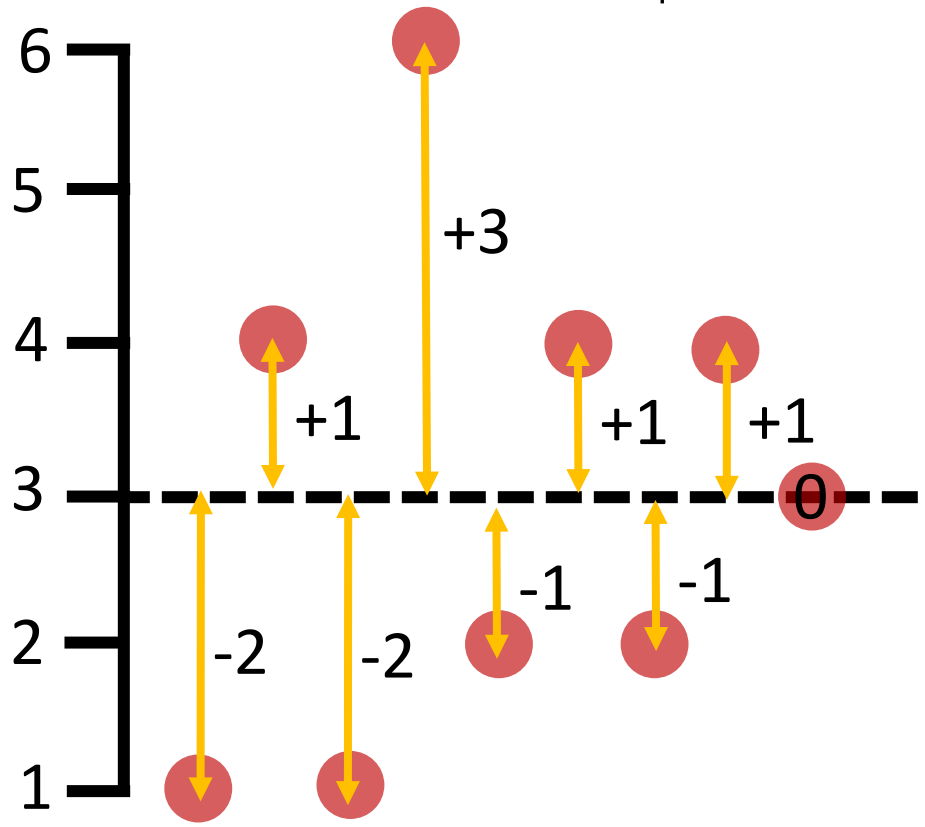
Where $\mu$ is the mean

$x_i$ is each data point

$n$ is the number of data points

# RECAP – find the difference

1. Calculate difference $X - \mu$




| Data point | χ - μ |
|---|---|
| χ1 | -2 |
| χ2 | 1 |
| χ3 | -2 |
| χ4 | 3 |
| χ5 | -1 |
| χ6 | 1 |
| χ7 | -1 |
| χ8 | 1 |
| χ9 | 0 |
| Total | 0 |

# RECAP - sum of squares

1. Calculate difference $X - \mu$
2. Calculate the sum of squares

$$\text{Sum of squares (SS)} = \sum(\mu - x_i)^2$$

is the sum of all data

is the population mean

is each data point

| Data point | χ - μ | (χ - μ)² |
|---|---|---|
| χ1 | -2 | 4 |
| χ2 | 1 | 1 |
| χ3 | -2 | 4 |
| χ4 | 3 | 9 |
| χ5 | -1 | 1 |
| χ6 | 1 | 1 |
| χ7 | -1 | 1 |
| χ8 | 1 | 1 |
| χ9 | 0 | 0 |
| Total | 0 | 22 |

# RECAP – variance (to take into account the number of data points!)

- Variance: Average deviation around the mean of a distribution (average of sum of squares)

$$\text{Variance } (\sigma^2) = \frac{\sum(\mu - x_i)^2}{n-1}$$

Where $\mu$ is the mean

$x_i$ is each data point

$n$ is the number of data points

Sum of squares

Degrees of freedom

# 3. Measures of variability – standard deviation

- Standard deviation (σ): Measure of the typical deviation from the mean. It is the squared root of the variance

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{\sum(\mu - x_i)^2}{n-1}}$$

Variance

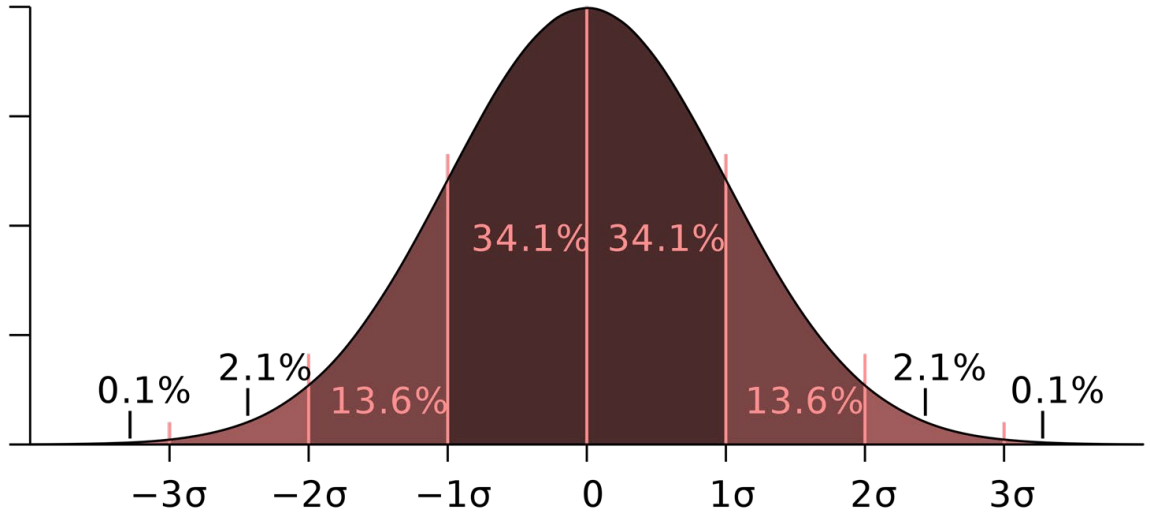
Where $\mu$ is the mean

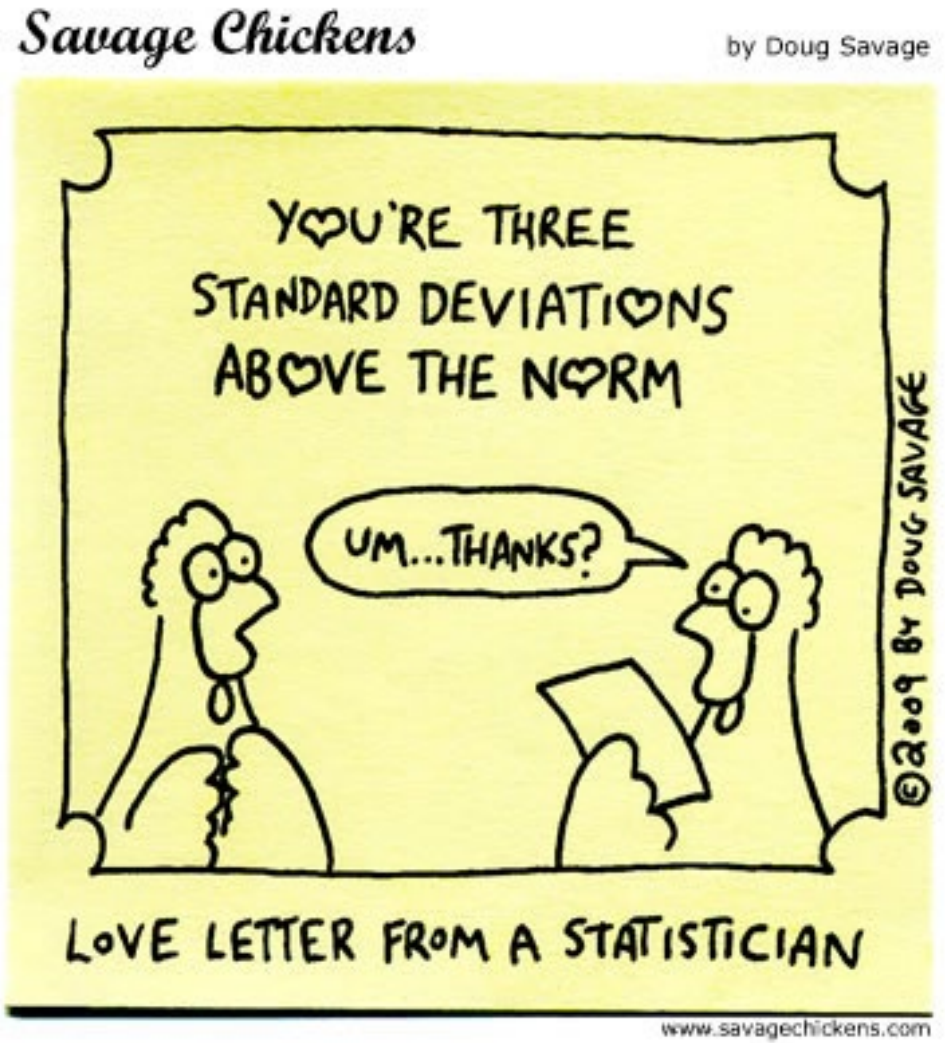
$x_i$ is each data point

$n$ is the number of data points

# 3. Measures of variability – standard deviation

- Standard deviation (σ): Measure of the typical deviation from the mean. It is the squared root of the variance




Savage Chickens
by Doug Savage
YOU'RE THREE STANDARD DEVIATIONS ABOVE THE NORM
UM...THANKS?
©2009 BY DOUG SAVAGE
LOVE LETTER FROM A STATISTICIAN
www.savagechickens.com

Lancaster University

# PSYC214: Statistics
# Lecture 1 – Measurement, variance and inferential statistics – Part 5

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

# Inferential statistics

Lancaster University

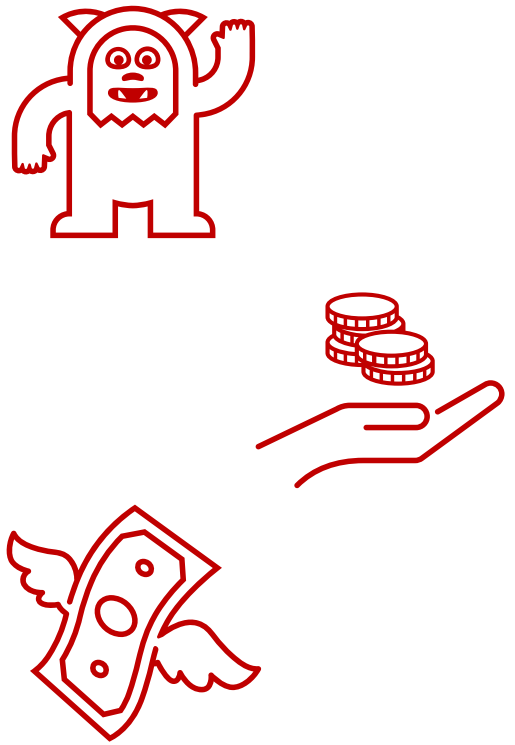
1. Allow you to draw conclusions based on extrapolations
2. Use data from the sample of participants in the experiment to compare the treatment groups and make generalizations about the larger population of participants
3. Provide a quantitative method to decide if the null hypothesis ($H_0$) should be rejected

# Inferential statistics - comparing groups

*Often, a researcher is interested in gathering information about different populations in order to compare them*

- What is the effect of our treatment/manipulation on an outcome of interest?

- Compare anxiety levels in different age groups
- Compare charitable behaviour before and after Christmas
- Compare Pre and Post consumer behaviour of Covid-19



# Inferential statistics - Hypotheses

Lancaster University

## $H_0$ the Null Hypothesis

- $H_0$: there is no significant difference between the conditions/groups and the null hypothesis is accepted.
- Under $H_0$, the samples come from the same population.

## $H_1$ the Experimental Hypothesis

- $H_1$: there is a significant difference between the conditions/groups and the null hypothesis is rejected.
- Under $H_1$, the samples come from the different populations.

# Inferential statistics - (Non)parametric tests

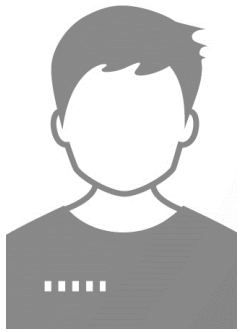
Lancaster University

- Statistical tests can be separated into:
  - Parametric
  - Non-parametric

While parametric tests are the norm in psychology and are generally more powerful than non-parametric tests, they require that the scores be an interval or ratio measure and there needs to be homogeneity of variance

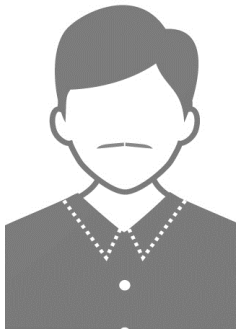
# Example set 1

*George observed a group of patients before, during and after the administration of a drug X to evaluate the effectiveness of the treatment.*



*Fozia measured participants' scores on a Psychological test of creativity in the morning, noon and afternoon in order to see whether there are any differences throughout the day.*

*Kenji measured the visual acuity of a single group of observers. He asked each subject to complete a vision test after they wore each of the five different brands of contact lenses under investigation.*



# In all cases

Lancaster University

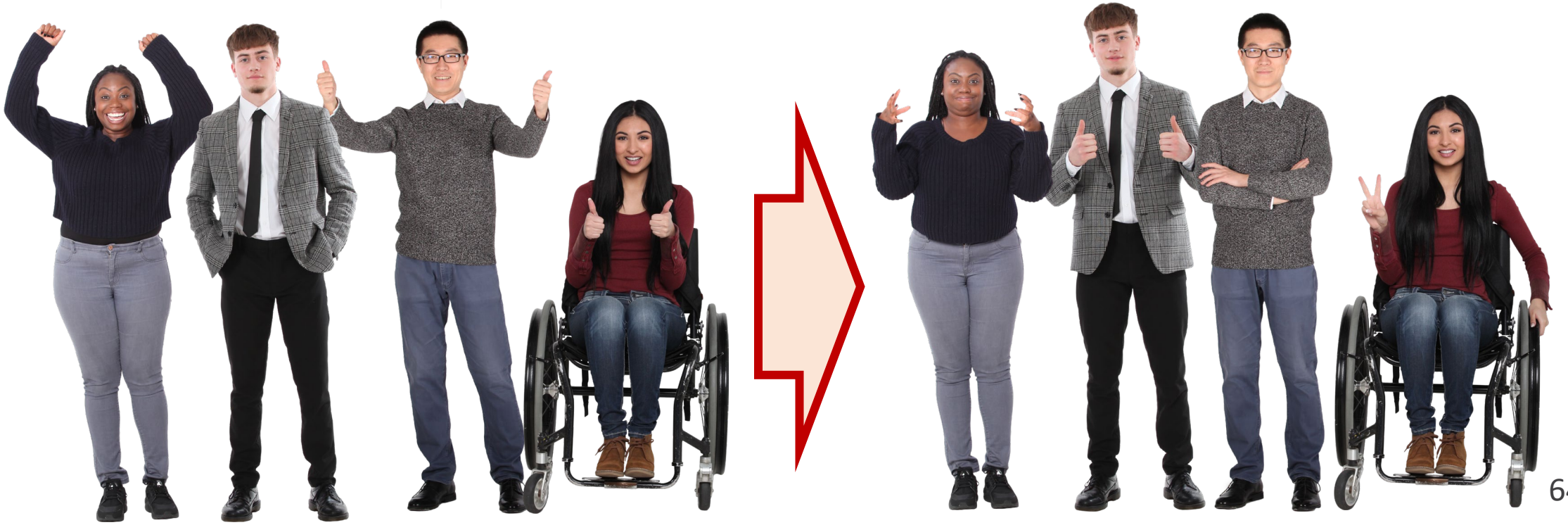
The same participant (used to be called **subject**) is being tested in different conditions:

- Before, during and after treatment
- Morning, noon and afternoon
- Five different brands of contact lenses

As each participant (subject) provides scores on the different conditions, we say that the measures are **related** and **correlated**

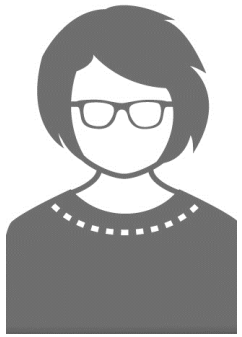
# In all cases

In these three cases, the Independent Factor is said to be a WITHIN-subject factor as it is altered within each subject.

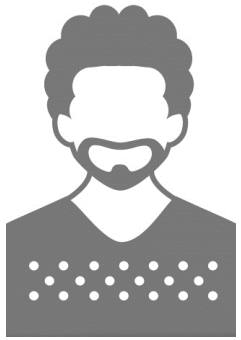


# Example set 2

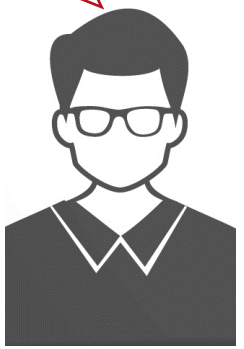
*Lucy is interested in age differences in mental toughness. She recruits 20 young adults, 20 middle-aged adults and 20 older adults, and asks them to complete a Hardiness Test.*



*Manuel is studying whether statistics lectures are more effective in the morning or in the afternoon. He administers a pop quiz to the morning and afternoon classes and compares the performance.*



*Mo wants to examine differences in personality traits between students from different universities. He recruits students from Lancaster, York and Bath and asks them to complete a sociability questionnaire.*



# In all cases

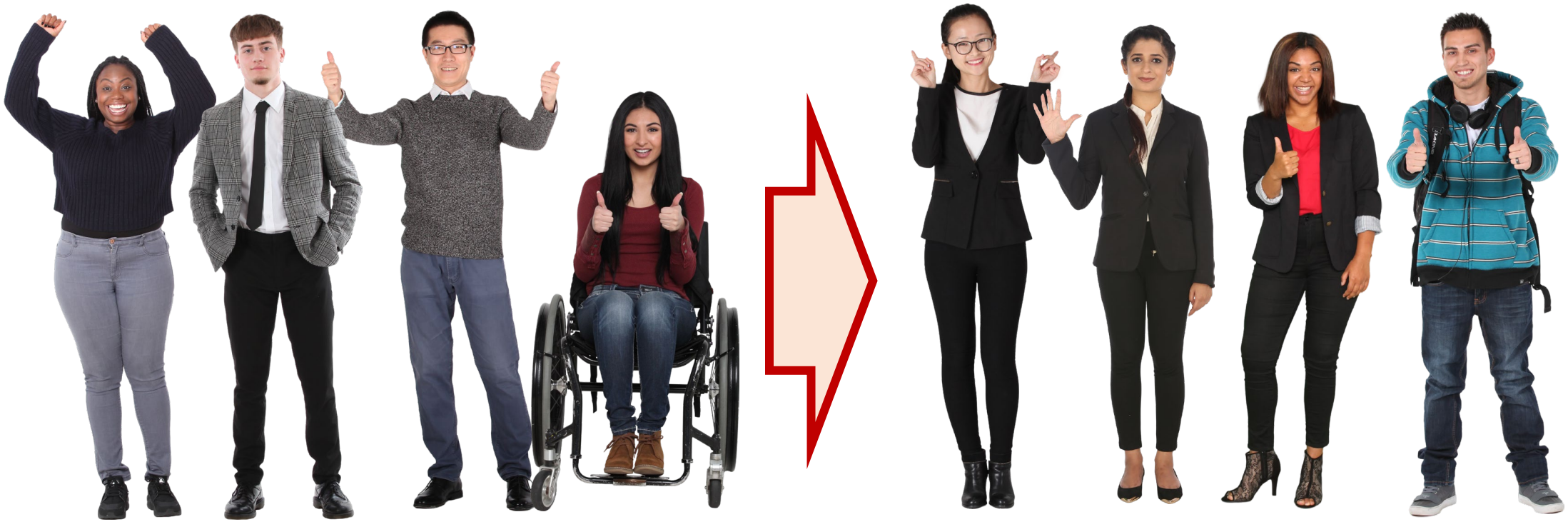
The different subjects are being tested in different conditions

- Young, Middle-Aged and Older Adults
- Morning class and Afternoon class
- Lancaster, York and Bath Universities

Because different observer provides scores on the different conditions, we say that the measures are **unrelated** and **uncorrelated**

# In all cases

In this case, the Independent factor is said to be a BETWEEN-subject factor as it is altered between each subject.



# Lecture 1 – Measurement, variance and inferential statistics

Lancaster University

## Review

- Experimental science
- Variables and levels of measurement
- Descriptive statistics
  - Levels of measurement
  - Measures of central tendency
  - Measures of variability
- Distributions
- Inferential statistics and hypotheses
- Within and between participant designs

Lancaster University

# Thank you for attention!