STATISTICALLY BASED REPORTS
ANSWERS

Key Terms and Causation

Explanatory, Response, and Confounding Variables
STOP AND CHECK (PAGE 7)

- Ice cream:
    - Explanatory variable: Day of the week.
    - Response variable: Ice creams sold.
    - Confounding variable: Time of the year.
- Health foods:
    - Explanatory variable: The type of food eaten.
    - Response variable: Health levels.
    - Confounding variables: Existing health conditions; other foods eaten; exercise habits.
- Facial hair:
    - Explanatory variable: Moustache, or no moustache.
    - Response variable: Income.
    - Confounding variables: Field of work; age.

Observational Studies vs. Experiments
STOP AND CHECK (PAGE 8)

- An observational study involves analysing something that is already occurring in the world. An experimental study is where the researcher manipulates certain variables to better infer causal relationships.
- It would be an observational study, as the researcher does not manipulate the number of males and females who use the cafeteria. Additionally, the researchers cannot change the amount of money each person spends.

Causation and Correlation

STOP AND CHECK (PAGE 11)

- A causal claim can be justified when confounding variables are controlled for, and the study is randomised and experimental, not observational.
- Participants require random assignment to groups to better ensure they are representative of the population as a whole. Randomisation also reduces the change of researcher bias.

Correlation

STOP AND CHECK (PAGE 12)

- A correlation describes an **association** between two variables. A causal claim infers that one variable **influences** the other.
- Other correlated variables may be the number of hours people sleep and their age, or the number of hours spent outside and the time of year.

Checking Causation vs. Correlation in a Problem

STOP AND CHECK (PAGE 14)

- A causal claim can only be made when it is an experimental study. Additionally, all confounding variables need to be controlled for and the participants have to be representative of the population being studied.

Key Terms and Causation

QUICK QUESTIONS (PAGE 14)

- There is some statistical evidence to support this claim as it is an experimental study. The independent variable is mindfulness habits, with the two levels being 'completing a reflection' and 'not completing a reflection'. There are sample sizes of over 30, and were randomly selected.

- A causal claim cannot be made as the study is observational in nature. Researchers did not control for the caffeine consumed by students beforehand. Consequently, they cannot reliably analyse the impact of caffeine levels as they only recorded it after the fact. Additionally, the outcome variable (time taken) is not a clear indicator of anything other than speed. For example, it does not consider the accuracy of the answers, nor if students answered everything.

Samples

Samples

STOP AND CHECK (PAGE 17)

- 30
- A representative sample is needed to be able to generalise the findings of the study to the wider population.
- A biased sample would involve some level of decisions being made, i.e., it is not randomised. Examples of this include separating groups by gender, age, ethnicity, IQ, income, or another demographic variable. The only exception to this is if that variable is under study, and therefore, the separation is essential to the study.

Random Sampling Methods

STOP AND CHECK (PAGE 20)

- The four types of random sampling methods are:
    - Simple random sampling.
    - Systematic random sampling.
    - Stratified random sampling.
    - Cluster sampling.

Non-random Sampling Methods

STOP AND CHECK (PAGE 23)

- The three types of non-random sampling methods are:
    - Quota sampling.
    - Person-in-the-street sampling.
    - Self-selected sampling.

Samples

QUICK QUESTIONS (PAGE 23)

- Population: High school students.
    - Sampling method: Cluster sampling.
- Population: Eligible voters.
    - Sampling method: Self-selected sampling.
- Population: New Zealand drivers.
    - Sampling method: Simple random sampling.

Sources of Error

Sampling and Non-sampling Errors

STOP AND CHECK (PAGE 33)

- Sampling errors refer to errors arising due to using a sample, not the whole population. Non-sampling errors arise due to everything else which creates variance in a sample.
- The 5 types of non-sampling errors are:
    - Excluded groups.
    - Self-selection.
    - Iffy/poorly worded questions.
    - Peer/social pressure.
    - Non-response error.

Confounding Variables

STOP AND CHECK (PAGE 29)

- When thinking about possible confounding variables, consider other factors or variables which could influence the relationship being analysed.
- Confounding variables cause error as it can lead a researcher to falsely attribute a relationship which may actually be due to a third (or more) variables. Confounding variables undermine the validity and reliability of an experiment, and mean researchers cannot confidently claim anything.

Extending Findings

STOP AND CHECK (PAGE 30)

- We cannot guarantee that the findings of one population will translate to another as we cannot be sure they share all of the same elements/variables. In other words, the findings of one population are not guaranteed to be generalizable to another.

Blast through Graphs

STOP AND CHECK (PAGE 32)

- The upper and lower 25% of data in a box-and-whisker graph is the **lower** and **upper quartiles**.
- If a relationship between two variables is positive, the regression line will slope upwards as it moves to the right.

Misleading Graphs

STOP AND CHECK (PAGE 35)

- A graph could be misleading if:
    - It has incorrect proportions.
    - It has incomplete data.
    - Has two vertical axes.

- Is 3D skewed.

Confidence Intervals

Theoretical Probability

STOP AND CHECK (PAGE 42)

- For example, the theoretical probability of randomly picking a green ball out of a bag containing 4 green balls and 5 pinks.

The Margin of Error & the 95% Confidence Interval

STOP AND CHECK (PAGE 38)

- The margin of error is a small allowance made to accomodate for unavoidable sampling variability. It tells us how far the true value may be from the sample value.
- The margin of error is added and minused from the sample proportion to get the confidence interval.
- The rule of thumb method should not be used when the sample proportions are outside the range of 0.3-0.7

Using the Confidence Interval for Claims

STOP AND CHECK (PAGE 39)

- A majority claim can be justified if over 50% of the data supports the statement.

Finding Sample Size from MoE

STOP AND CHECK (PAGE 41)

- To find the sample size from the MoE:
  - Plug the MoE into the MoE equation.

- Bring the $\sqrt{n}$ up by multiplying it on both sides.
- Divide both sides by the MoE
- Square both sides.

Dependent Proportions

STOP AND CHECK (PAGE 43)

- Proportions are dependent if they relate to two different answers to the same questions from the same sample.
- We need to find the confidence interval for the difference between dependent proportions to be able to make an exact claim between any possible differences.

Independent Proportions

STOP AND CHECK (PAGE 46)

- Comparing statistics provides evidence of risk (the likelihood of something occurring in comparison to something else), as well as being able to measure the increase or decrease of a variable over time.
- To find the MoE for the difference between two dependent proportions, use:

$$\text{MoE of difference} = 2 \times \text{MoE of the study}$$

- To find the MoE for the difference between two independent proportions, first find the MoE's of each sample. Then, use:

$$\text{Moe of difference} = \pm 1.5 \times \frac{MoE_1 + MoE_2}{2}$$

To find the difference, this value is then applied to the difference between the proportions. If the range contains 0, it means there is a chance that the difference is 0, i.e., there is **no difference**.

Confidence Intervals

QUICK QUESTIONS (PAGE 46)

- MoE = $\pm \frac{1}{\sqrt{n}}$

  MoE = $\pm \frac{1}{\sqrt{500}}$

  MoE = 0.0447 (3 d.p.), or 4%

  Confidence interval (CI) = Sample proportion $\pm$ MoE

  CI = 41 $\pm$ 0.0447

  CI = 40.955 to 41.045 (3 d.p.)

- MoE of difference = $\pm$ 2 $\times$ MoE

  MoE of difference = $\pm$ 2 $\times$ 4

  MoE of difference = $\pm$ 8%

  Sample difference = 55 - 45

  Sample difference = 10

  Confidence interval = Sample difference $\pm$ MoE

  CI = 10 $\pm$ 8

  CI = 2% to 18%

- Is it unlikely that the majority of people without tattoos and piercings feel that a tattoo makes a person less attractive, as only 45% of those without (26.5% of the total population) reported thinking this.

