

Sampling and Data Collection

Topic Test

1 a) Snowball sampling

(respondents share the survey with others)

b) e.g. Easy/quick to set up

or Social media is used by a lot of people so there is a high chance of a larger sample

c) e.g. Biased to those who have access to the internet / social media

or Respondents may not be truthful.

2 a) Judgemental sampling

(the researcher judges the shopping centre is representative)

b Advantage: Easy/quick to collect the data

or people at the shopping centre are more likely to have an opinion about public toilets

Disadvantage: Biased to those in the shopping centre that day
or may not be representative of all the population

3 e.g. Go to 5 running clubs and take a random sample of athletes in each club.

Record the running club and the running time for the 100m sprint for each athlete.

4 • Generate 8 random numbers in the range 1 to 25

← range must be specified

• Do not ignore repeats

← must be explicit

• Select the audience members assigned the numbers generated.

(otherwise you just have a list of numbers)

5 a)

- Number the students from 1 to 1200
- Generate a random number in the range 1 to 20
- Select the student with that number and every 20th student after that

$$\left(\frac{1200}{60} = 20 \right)$$

b) Yes because the sample size divides the population size

c) No because not all samples of size 60 can be selected
e.g. students 1 and 2 can never be in the same sample

A random sample must have

- every member of the population has an equal chance of being selected
- every sample of size n has an equal chance of being selected

6

a) • Generate random numbers in the range 1 to 659

(explicit)

- • Ignore repeats until 20 different numbers are generated
- Select the teeth assigned these numbers

↑
(otherwise you just have a list of numbers)

← (range must be specified)

b) The sampling method is random so the method itself is not biased

c) The population of teeth are only extracted molar teeth, so the data collected is biased to extracted / molar teeth.

<u>7</u>	Population	Sample
Smoke/Vape	60	6
Don't smoke or vape	660	66
Total	720	72

$\div 10$

- Assign the smokers/vapers a number from 1 to 60
- Generate 6 different random numbers in the range 1 to 60 ignoring repeats
- Select the smokers/vapers assigned the numbers.

← range must be stated (here)

← (explicit)

(must do this)

- Repeat for non smokers/vapers for the range 1 to 660 and generate 66 different random numbers

← (correct ranges and numbers specified)

8 a) Suggestion A: Cluster sampling

Suggestion B: Simple random sampling

b) Suggestion A: No

Suggestion B: Yes

← (if restaurants are not picked at random, some may never be considered)

c) i) ^{e.g.} Only need a list of employees for the 5 restaurants chosen

ii) e.g. The sample obtained will be a random sample.

9 a (i)

Suggestion A: Cluster sampling

Suggestion B: Random sampling

Suggestion C: (proportional) stratified sampling

Suggestion D: Systematic sampling

← (Not enough information to say simple or unrestricted)

(ii)

Suggestion A: No

Suggestion B: Yes

Suggestion C: Yes

Suggestion D: Yes

← (because it is proportional)

← (because $\frac{90000}{150}$ is an integer)

b No because (e.g.) it is not possible to obtain a sample of electors all from urban areas.

↑ For a sample to be random, all samples of size n have the same chance of selection

c Yes. The advantage is that the sample of electors will fairly represent the dwelling types of the population.

d Suggestion C may not be useful since only half of the sample are electors are from urban areas, but at least there will be some representation from urban areas.

Suggestion B may also not be useful because we cannot guarantee anybody in the sample will be from urban areas (or the whole sample may be).