Sampling strategies

The basic types of sampling

The Concept of Population

Sampling is used because it is not necessary to collect all the data to understand what the whole **population** is like. For example:

- In a sand dune ecosystem
  - We might be interested in the types of plants growing in each zone of the sand dune ecosystem. The population in this example is all of the plants in the ecosystem. To understand the whole population, we sample some plants in each zone.
- In a city
  - We might be interested in how the flow of traffic in two particular streets varies at different times of day. The population in this example is all of the traffic that travels down each street in a day. To understand the whole population, we sample traffic at key times of the day.

Population types

Heterogeneous population

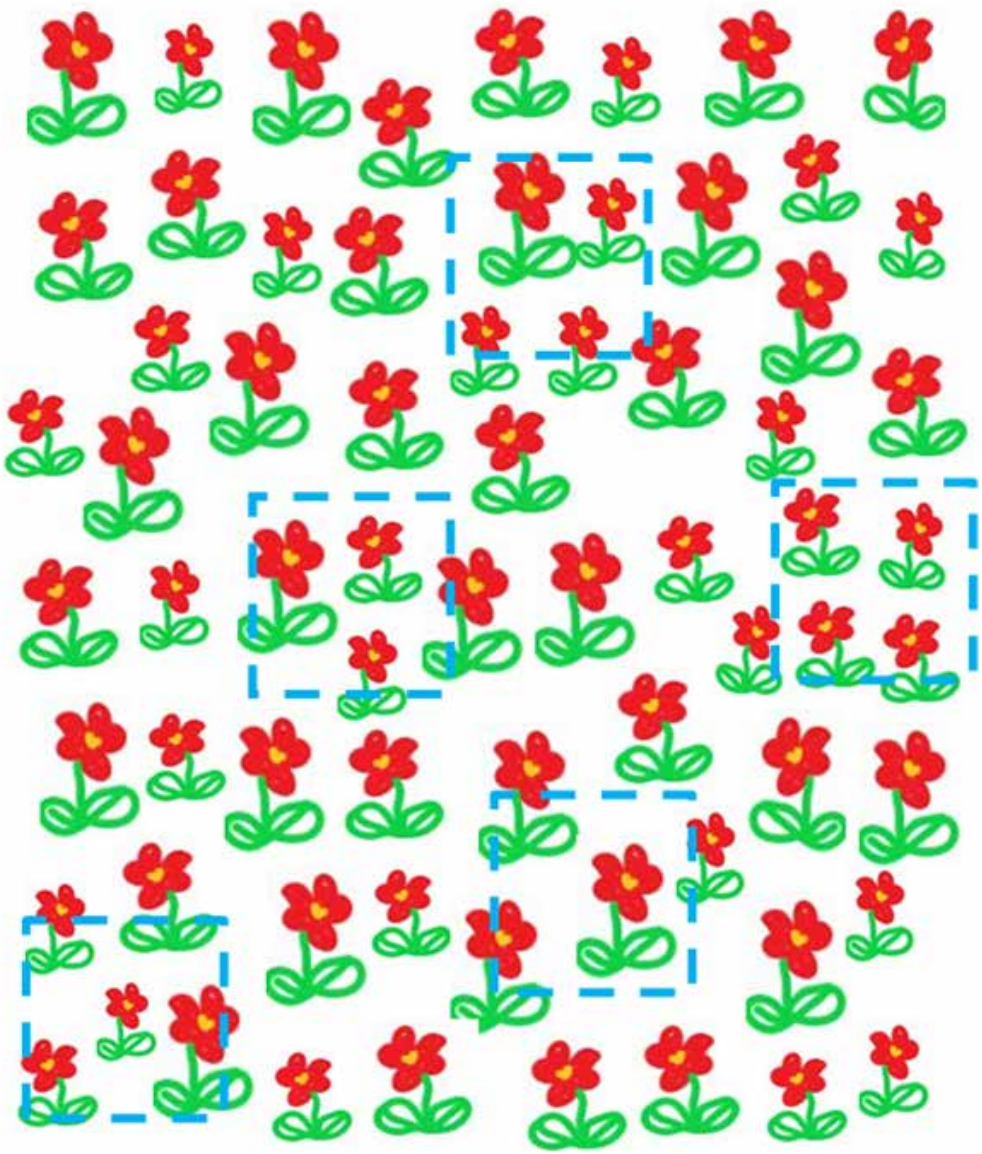
Homogeneous population

The concept of a representative sample

An effective sampling strategy is designed to collect a **representative sample**. This is a sample that accurately reflects variations in the whole population.

Probability sampling is used to ensure a representative sample is selected. The main types of probability sampling are:

- Systematic
- Random
- Stratified



If the population is **homogeneous** then a smaller sample should be representative.

If the population is **heterogeneous** (very varied) then a larger sample will be needed for it to be representative.

Random sampling

Random sampling

This involves using a **random** method (such as using an app to generate random numbers) to select a sample of data. For example:

- In a sand dune ecosystem
  - Pegging out a 10 metre by 10 metre area of the dunes and placing your quadrat according to co-ordinates generated by random numbers.
- In a city
  - Using a dice to decide whether to turn left, right, or go straight ahead, as you walk through the city, then flipping a coin to decide whether or not to collect data.

First random sample

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Second random sample

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Systematic sampling

Using a regular system to choose a sample of data, for example:

- In a sand dune ecosystem
  - Using a quadrat to record plant species every 25 metres along a line through the dunes.
  - Measuring wind speed at 20cm, 40cm, 60cm and 80cm above the sand.
- In a city
  - Recording land use every 100 metres from a city centre.
  - Using a Cleanliness Index every 10 metres from a bus stop, metro station or fast food take-away.

Review

Suggest how both random and systematic sampling could be used to investigate these sand dunes.

Ynyslas sand dunes, Ceredigion

Stratified sampling

Choosing a data set that is in proportion to segments of the whole **population**. For example:

- In a sand dune ecosystem
  - Using secondary data to estimate the percentage of the ecosystem covered by embryo dunes, mobile dunes, fixed dunes and dune slacks and then selecting a frequency of sample sites that match that proportion.
- In a city
  - Using secondary data to estimate the percentage of the population who are in specific age groups and then selecting sample sizes for a questionnaire that are in these proportions.

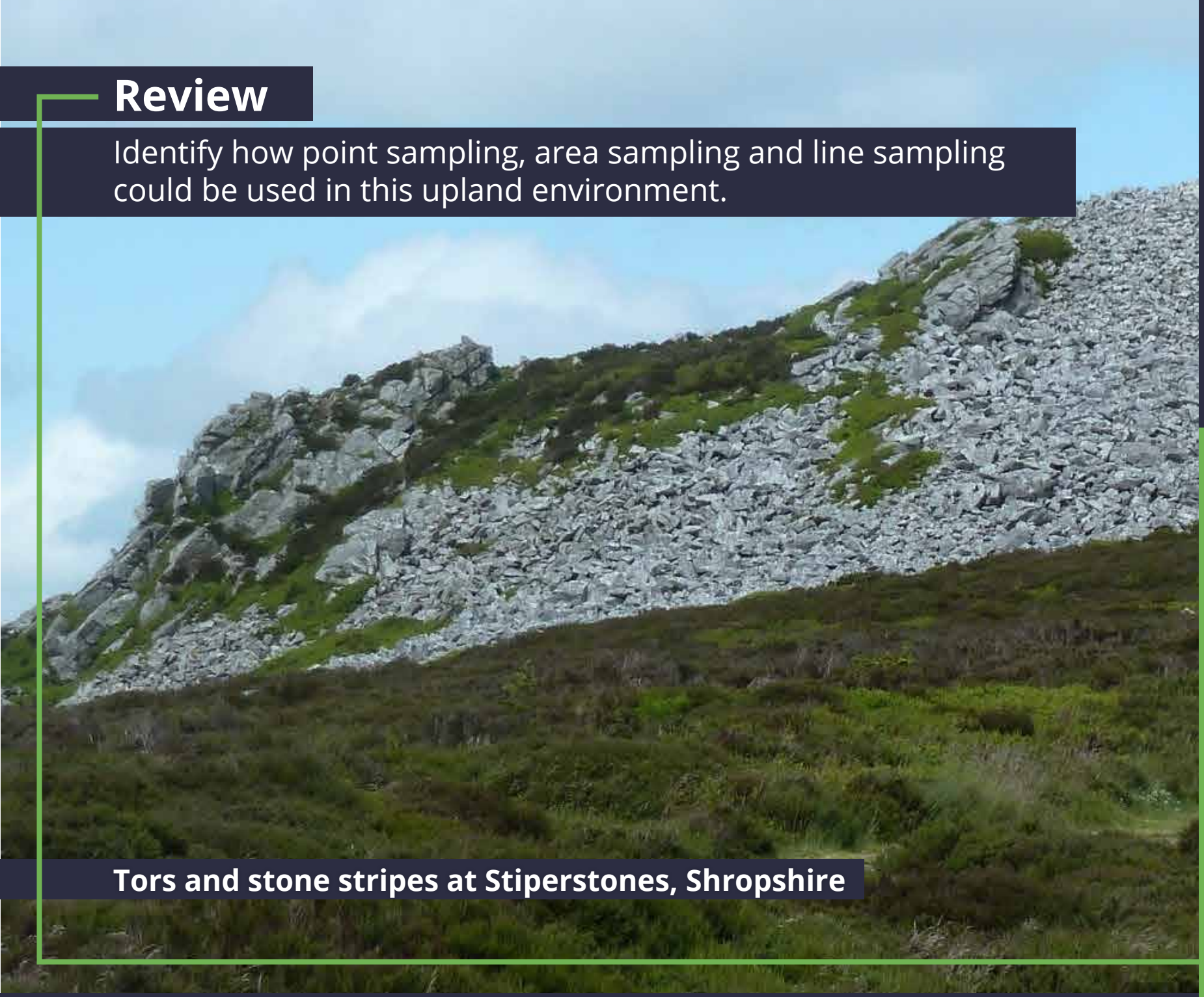
Points, areas and lines

We can organise a sample by collecting data at points, from an area, or along a line (or transect). For example:

	Points	Areas	Lines
Sand dune	Wind speed or soil pH at sample points.	Percentage cover of each plant by area.	Beach profile along a line from low tide to strand line.
City	Pedestrian flow at sample points.	EQI score for an area of 50 square metres.	Changing land use along a line radiating from the city centre.
Upland	Scree shape, size or orientation at sample points.	Percentage of a footpath that is trampled.	Slope angles along a cross section of a glaciated valley.

Review

Identify how point sampling, area sampling and line sampling could be used in this upland environment.



Tors and stone stripes at Stiperstones, Shropshire

Frequency

As well as deciding whether to use a systematic, random or stratified method of selecting the sample, you also need to consider sample size and **frequency**. Frequency is determined by the distance between sample points. For example:

- In a sand dune ecosystem
  - Setting a transect of 500 metres and sampling at a frequency of every 100 metres will give six sample points.
  - Setting a transect of 500 metres and sampling at a frequency of every 25 metres will give 21 sample points.

Frequency of systematic sampling

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64				68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Timing

In many investigations students also need to consider the timing of a sample. For example:

- In a sand dune ecosystem
  - Measuring wind speed every 10 minutes
- In a city
  - Counting pedestrian flows in a city centre for 5 minutes every hour

In each of these examples, the timing of the sample is systematic.

Timing 2

The timing of a sample needs to be considered when the data that is being collected varies over time. For example:

- In a sand dune ecosystem
  - Wind speed will fluctuate on a gusty day.
- In a city
  - Pedestrian and traffic flows vary on a diurnal pattern; i.e. flows vary depending on time of day due to rush hour commuting and lunch breaks.
  - Pedestrian and traffic flows also vary during the week.
- In an upland landscape
  - Visitor numbers to a honeypot are greater at weekends and Bank Holidays.

Review

Identify an investigation in this environment that would require both spatial and temporal sampling.