Chapter 1: Data Collection

The chapters of Stats Year 1 could be broadly organised as follows:

Experimental

i.e. Dealing with collected data.

Chp1: Data Collection

Methods of sampling, types of data, and populations vs samples.

Chp2: Measures of Location/Spread

Statistics used to summarise data, including mean, standard deviation, quartiles, percentiles. Use of linear interpolation for estimating medians/quartiles.

Chp3: Representation of Data

Producing and interpreting visual representations of data, including box plots and histograms.

Chp4: Correlation

Measuring how related two variables are, and using linear regression to predict values.

Theoretical

Deal with probabilities and modelling to make inferences about what we 'expect' to see or make predictions, often using this to reason about/contrast with experimentally collected data.

Chp5: Probability

Venn Diagrams, mutually exclusive + independent events, tree diagrams.

Chp6: Statistical Distributions

Common distributions used to easily find probabilities under certain modelling conditions, e.g. binomial distribution.

Chp7: Hypothesis Testing

Determining how likely observed data would have happened 'by chance', and making subsequent deductions.

1:: Populations vs samples

"Suggest why we would not test all the light bulbs."
"Identify the sampling frame."

2:: Random Sampling

Describe the disadvantages of systematic sampling.

3:: Non-Random Sampling

Describe how a stratified sample would be conducted, including strata sizes.

4:: Types of data

Continuous vs discrete, terms such as class intervals, class boundaries, class width.

5:: Edexcel's 'Large Data Set'

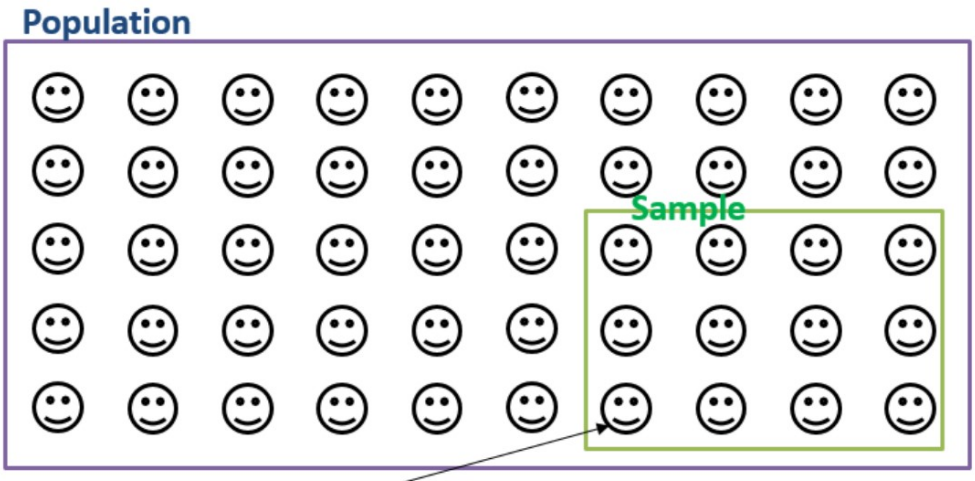
What you're expected to know about the 'large data set' of weather data, and how to use it.

Populations and Samples

A **population** is:	the <u>whole</u> set of items that are of interest.
A **sample** is:	some <u>subset of the population</u> intended to <u>represent the population</u>.

You're probably used to a 'population' meaning all humans/animals within a country/ecosystem. But a population could be "*all the lightbulbs in a factory*" or "*all the cars in the UK*".

Sampling key terms



Each individual thing in the population that can be sampled is known as a **sampling unit**.

Often sampling units of a population are individually named or numbered **to form a list** called the **sampling frame**.

Populations vs Samples

We could collect data either from a sample, or from the entire population.
Data collected from the entire population is known as a **census**.

	Advantages	Disadvantages
Census	Should give completely accurate result.	• **Time consuming** and **expensive**. • Cannot be used when testing involves **destruction** *– for example, finding out number of biscuits in a machine filled box/bag* • **Large volume of data** to process.
Sample	• **Cheaper**. • **Quicker**. • **Less data** to process.	• Data may **not** be **accurate**. • Data may **not** be large enough to **represent small sub-groups.**

Example: A supermarket wants to test a delivery of avocados for ripeness by cutting them in half.

a. Suggest a reason why the supermarket should not test all the avocados in the delivery.
b. The supermarket tests a sample of 5 avocados and finds that 4 of them are ripe. They estimate that 80% of the avocados in the deliver are ripe. Suggest one way that the supermarket could improve their estimate.

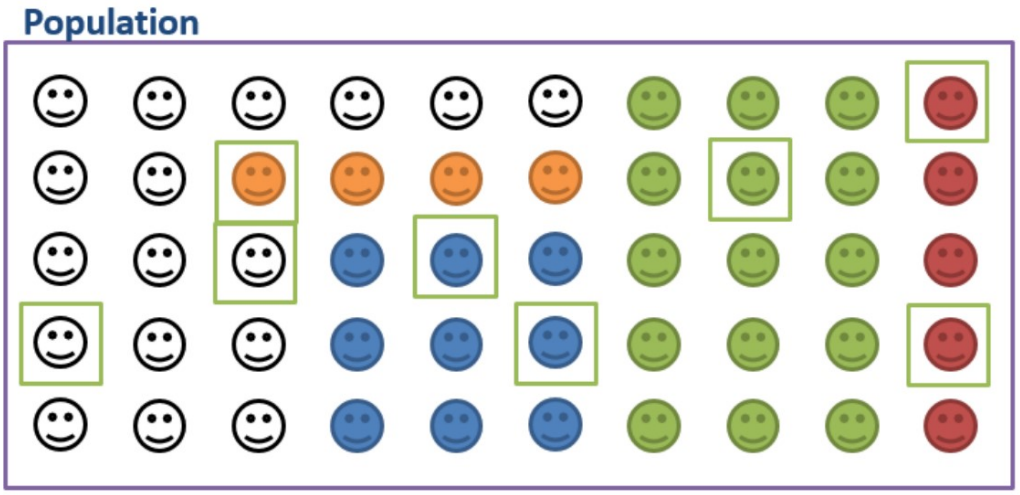
Ex 1A

Types of Sampling

We'll be investigating different types of sampling, what they are, how to carry them out, and their advantages and disadvantages. I've included a blank version of this at the back for you to use to test yourself during revision.

	Type	What it is	How to carry out	Advantages	Disadvantages
Random Sampling	**Simple Random Sampling**				
	Systematic Sampling				
	Stratified Sampling				
Non-random Sampling	**Quota Sampling**				
	Opportunity/ Convenience Sampling				

Random Sampling – 3 types



We want each **sampling unit** in our **sampling frame** to have an **equal chance of being chosen**, in order to **avoid bias**.

This is known as **random sampling**.
There are a few ways of doing this...

i) Simple Random Sampling

Type	What it is	How to carry out
Simple Random Sampling	Every sampling unit in the sampling frame has an equal chance of being selected	In sampling frame each item has identifying number. Use random number generator, or 'lottery sampling' (names in a hat)

Advantages	Disadvantages
• Bias free • Easy and cheap to implement • Each number has a known equal chance of being selected	• Not suitable when population size is large • Sampling frame needed

There are 64 girls and 56 boys in a school. Explain briefly how you could take a random sample of 15 pupils using a simple random sample. **(3)**

Allocate a number between 1 and N (or equiv) to each pupil.

Use random number tables, computer or calculator to select 15 different numbers between 1 and 120 (or equiv).

Pupils corresponding to these numbers become the sample.

M1 ← Mark for allocating identifier to each sampling unit.

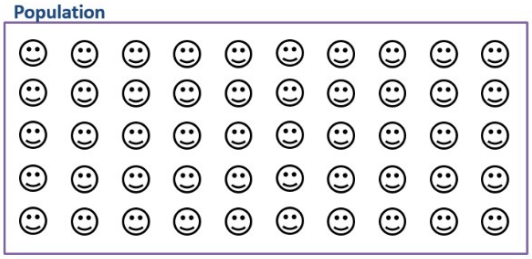
B1 ← Mark for one (bias-free) method to select such a number.

B1 ← Mark for explicitly mentioning how that number is actually used.

ii) Systematic Sampling

Type	What it is	How to carry out
Systematic Sampling	Required elements are chosen at regular intervals in ordered list	i.e. Take every kth elements where: $k = \frac{pop\ size\ (N)}{samp\ size\ (n)}$ starting at random item between 1 and k

Population



Advantages	Disadvantages
• Simple and quick to use • Suitable for large samples/ populations	• Sampling frame again needed • Can introduce bias if sampling frame not random

A telephone directory contains 50 000 names. A researcher wishes to select a systematic sample of 100 names from the directory. Explain in detail how the researcher should obtain such a sample. **(2)**

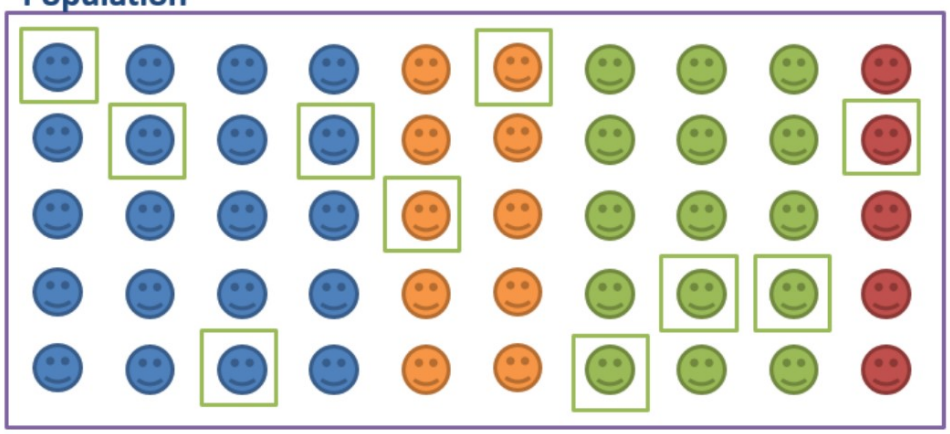
Randomly select a number between 00 and 499 (001 and 500) **B1**
select every 500th person **B1**

We need a random first item.

iii) Stratified Sampling

We want to sample 20% of the population. If the population were divided into distinct groups (e.g. age ranges), known as '**strata**', we could randomly sample 20% from each group, ensuring each group is equally represented.

Population



Type	What it is	How to carry out
Stratified Sampling	Population divided into groups (strata) and a <u>simple random sample carried out in each group</u> Used when sample is large and population naturally divides into groups	Same proportion $\frac{samp\ size\ (n)}{pop\ size\ (N)}$ sampled from each strata

Advantages	Disadvantages
• Reflects population structure. • Guarantees proportional representation of groups within population	• Population must be clearly classified into distinct strata • Selection within each stratum suffers from same disadvantages as simple random sampling

A school has 15 classes and a sixth form. In each class there are 30 students. In the sixth form there are 150 students. There are equal numbers of boys and girls in each class. There are equal numbers of boys and girls in the sixth form. The head teacher wishes to obtain the opinions of the students about school uniforms. Explain how the head teacher would take a stratified sample of size 40. **(7)**

Total in School = $(15 \times 30) + 150 = 600$	B1
random sample of $\frac{30}{600} \times 40$ (Use of $\frac{40}{their 600}$)	M1
= 2 from each of the 15 classes	A1
random sample of $\frac{150}{600} \times 40$ Either	
= 10 from sixth form;	A1
Label the boys in each class from 1 – 15 and the girls from 1 – 15.	B1
use random numbers to select 1 girl and 1 boy	B1
Label the boys in the sixth form from 1 – 75 and the girls from 1 – 75. use random numbers to select 5 different boys and 5 different girls.	B1

The 100 members of a yacht club are listed alphabetically in the club's membership book.
The committee wants to select a sample of 12 members to fill in a questionnaire.

a) Explain how the committee could use a calculator or random number generator to take a simple random sample of the members.
b) Explain how the committee could use a lottery sample to take a simple random sample of the members.

A factory manager wants to find out what his workers think about the factory canteen facilities.
The manager decides to give a questionnaire to a sample of 80 workers. It is thought that different age groups will have different opinions.
There are 75 workers between ages 18 and 32.
There are 140 workers between ages 33 and 47.
There are 85 workers between ages 48 and 62.

a) Write down the name of the method of sampling the manager should use.
b) Explain how he could use this method to select a sample of workers' opinions.

Non-Random Sampling – 2 types

i) Quota Sampling

Consider the following scenario: You wish to conduct a survey in the UK **on whether being left-handed affects IQ**. We need to choose people to assess.

Why would random sampling be problematic?
Because <u>we don't know the sampling frame</u>, i.e. don't have a list of all left-handed (and non-left-handed) people in the UK.

For this scenario we'd likely use **quota sampling**, i.e.

1. As with stratified sampling, divide population into groups according to characteristic of interest, then determine size of each group in sample to reflect proportions within the population.
2. But instead of random sampling within each group, we actively choose people within each group via suitable means (e.g. advertising), **until the 'quota' for each group is filled**.
   - e.g. an interviewer could meet people, assess the group they are in, and allocate them into the appropriate quota (i.e. left or right handed people)
   - If a person refuses to be interviewed, or the quota is already full, move onto the next person

'*quota*': a fixed share/number of something

Type	What it is	How to carry out
Quota Sampling	Population divided into groups according to characteristic. A quota of items/people in each group is set to try and reflect the group's proportion in the whole population	<u>Interviewer selects the actual sampling units</u>

Advantages	Disadvantages
• Allows small sample to still be representative of population • No sampling frame required • Quick, easy, inexpensive • Allows for easy comparison between different groups in population	• Non-random sampling can introduce bias • Population must be divided into groups, which can be costly or inaccurate • Increasing scope of study increases number of groups, adding time/expense • Non-responses are not recorded

A lake contains 3 species of fish. There are estimated to be 1400 trout, 600 bass and 450 pike in the lake. A survey of the health of the fish in the lake is carried out and a sample of 30 fish is chosen.

(a) Give a reason why stratified random sampling cannot be used. **(1)**

(b) State an appropriate sampling method for the survey. **(1)**

(c) Give one advantage and one disadvantage of this sampling method. **(2)**

(d) Explain how this sampling method could be used to select the sample of 30 fish. You must show your working. **(4)**

(a) Sampling frame within each species of fish in the lake impossible to obtain. B1 **(1)**

(b) Quota sampling B1 **(1)**

(c) Advantages:
Sample can be obtained quickly
Costs are kept to a minimum B1
Administration of survey is easy
Disadvantages:
Not possible to estimate sampling errors B1
Process not random
Surveyor may not be able to identify species of fish easily **(2)**

(d)

Species	Quota
Trout	$\frac{1400}{2450}\times 30 = 17.14$
Bass	$\frac{600}{2450}\times 30 = 7.35$
Pike	$\frac{450}{2450}\times 30 = 5.51$

Fish are caught from the lake until the quota of 17 trout, 7 bass and 6 pike are reached. B1B1B1
If a fish is caught and the species quota is full, then this is ignored. B1 **(4)**

ii) Opportunity Sampling

A variant of quota sampling is **opportunity/convenience sampling**.

This is where we find people **at the same time the survey is being carried out** (e.g. exit polls at polling stations). This is not a suitable method for the left-handed example, because giving the likely time-consuming nature of assessment coupled with resources required, we'd likely arrange with the people taking part before the actual assessment tasks took place.

Type	What it is	How to carry out
Opportunity/ Convenience Sampling	Sample taken from people who are available at time of study, who meet criteria	Interviewer selects the actual sampling units according to the set criteria

Advantages	Disadvantages
• Easy to carry out • Inexpensive	• Unlikely to provide a representative sample • Highly dependent on individual researcher

	Type	What it is	How to carry out	Advantages	Disadvantages
Random Sampling	**Simple Random Sampling**	Every sampling unit in the sampling frame has an equal chance of being selected	In sampling frame each item has identifying number. Use random number generator, or 'lottery sampling' (names in a hat)	• Bias free • Easy and cheap to implement • Each number has a known equal chance of being selected	• Not suitable when population size is large • Sampling frame needed
	Systematic Sampling	Required elements are chosen at regular intervals in ordered list	i.e. Take every k^{th} elements where: $k = \frac{pop\ size\ (N)}{samp\ size\ (n)}$ starting at random item between 1 and k	• Simple and quick to use • Suitable for large samples/ populations	• Sampling frame again needed • Can introduce bias if sampling frame not random
	Stratified Sampling	Population divided into groups (strata) and a simple random sample carried out in each group Used when sample is large and population naturally divides into groups	Same proportion $\frac{samp\ size\ (n)}{pop\ size\ (N)}$ sampled from each strata	• Reflects population structure. • Guarantees proportional representation of groups within population	• Population must be clearly classified into distinct strata • Selection within each stratum suffers from same disadvantages as simple random sampling
Non-random Sampling	**Quota Sampling**	Population divided into groups according to characteristic. A quota of items/people in each group is set to try and reflect the group's proportion in the whole population	Interviewer selects the actual sampling units	• Allows small sample to still be representative of population • No sampling frame required • Quick, easy, inexpensive • Allows for easy comparison between different groups in population	• Non-random sampling can introduce bias • Population must be divided into groups, which can be costly or inaccurate • Increasing scope of study increases number of groups, adding time/expense • Non-responses are not recorded
	Opportunity/ Convenience Sampling	Sample taken from people who are available at time of study, who meet criteria	Interviewer selects the actual sampling units according to the set criteria	• Easy to carry out • Inexpensive	• Unlikely to provide a representative sample • Highly dependent on individual researcher

Simple Random Sampling	Systematic Sampling	Stratified Sampling	Quota Sampling	Opportunity Sampling

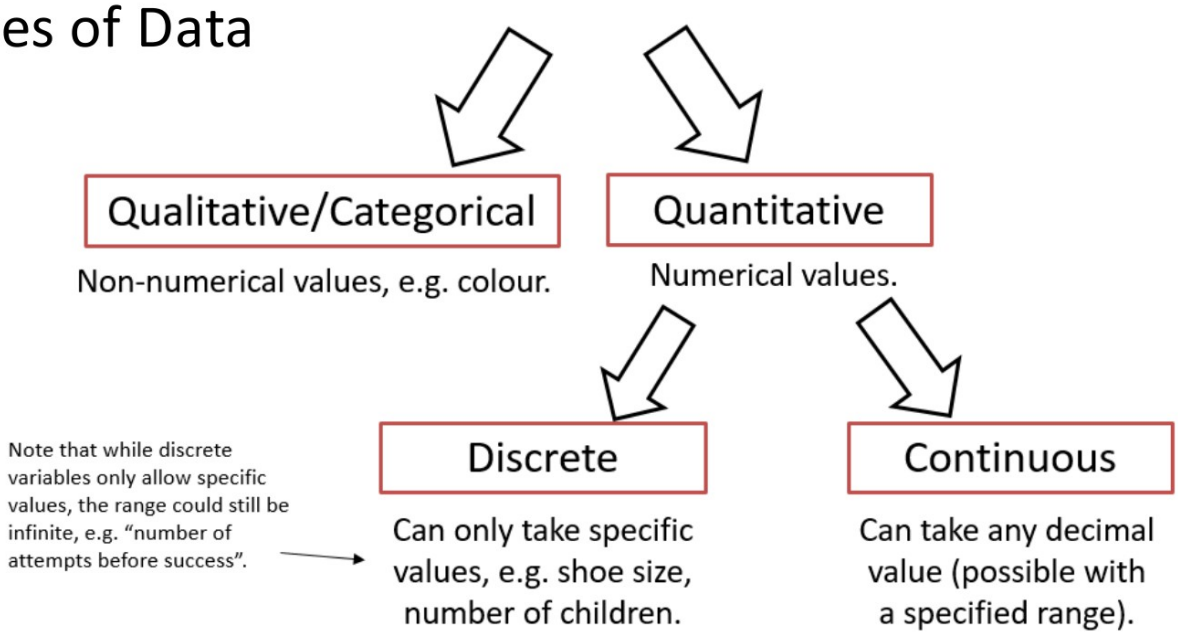
Suggest a suitable sampling method.

"You wish to test lightbulbs produced by a factory in a daily batch."

"You wish to survey consumer opinion on your new drink *FizzGuzz* released in the UK."

"You wish to determine students' favourite TV programmes in your school, that is fairly representative of each year group."

Types of Data



Eddie carries out a survey about the pet dogs his classmates own.
Decide what type of data the following are:

(a) How many dogs each person owns

(b) The colour of the dogs

(c) The type of dog

(d) The name of each dog

(e) The age of each dog

(f) The mass of each dog

Grouped Data

Weight w (kg)	Frequency
$0 \leq w < 20$	3
$20 \leq w < 70$	4

Data can be **grouped** for conciseness, at the expense of losing the exact original values.

$20 \leq w < 70$

This is known as a **class interval**.

Lower class boundary

Midpoint **= 45**

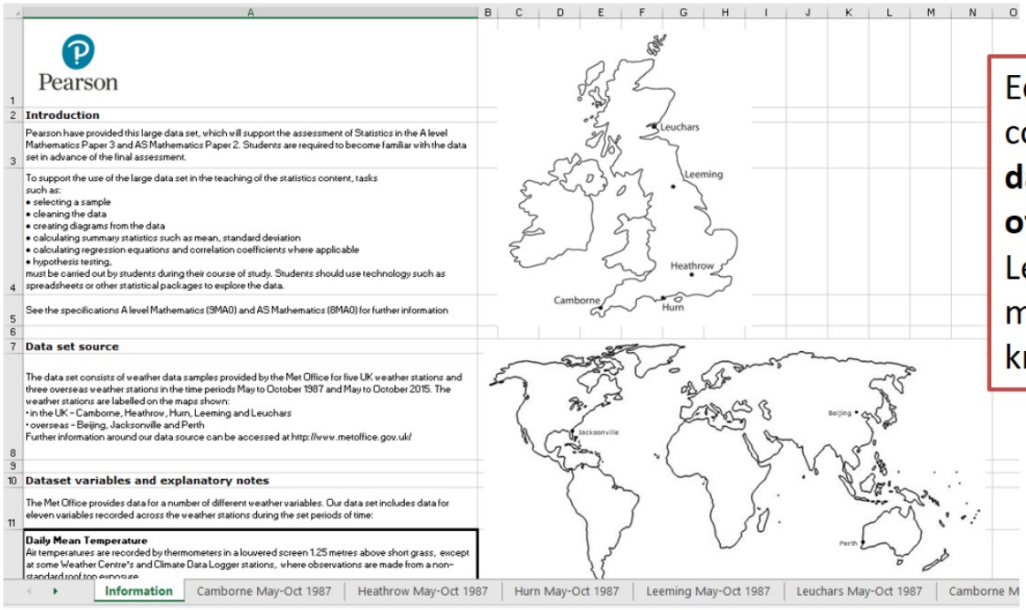
Upper class boundary

Class width = $70 - 20 = \mathbf{50}$

Large Data Set (LDS)

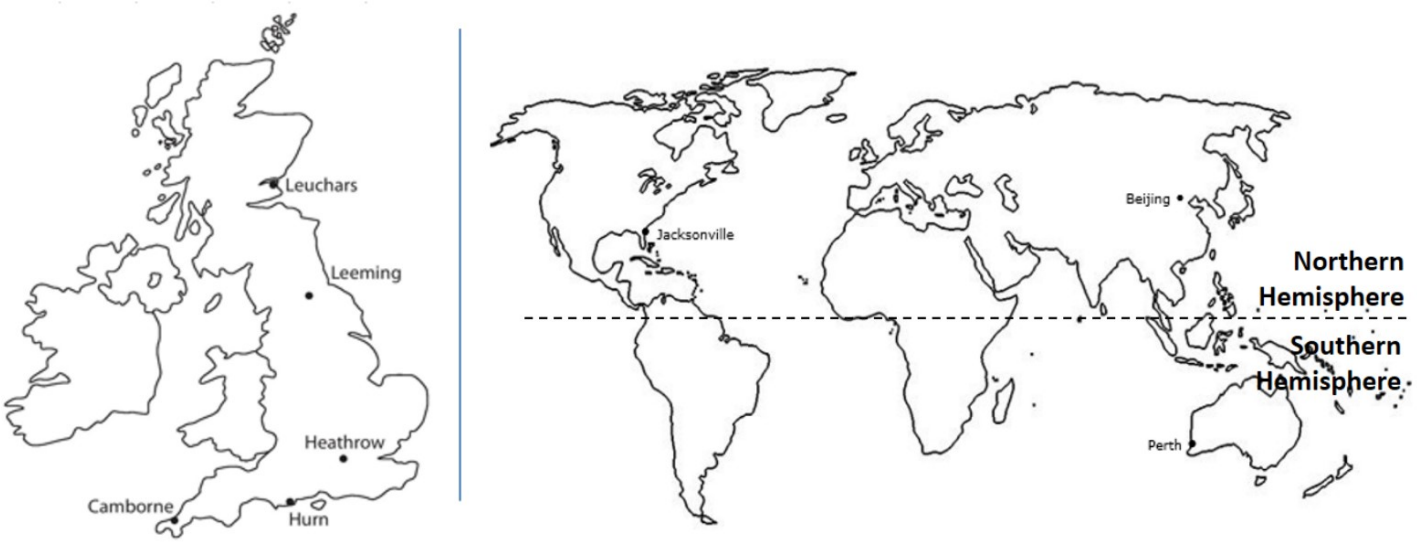
All A Level exam boards are obligated to provide a 'large data set'. Data in exam questions will often be from this set, and you are encouraged to explore this data (which is publicly available) in Microsoft Excel.

It is important to note that you are expected to be familiar with this data set before you go into your exam, including some basic knowledge of geography.



Edexcel's data set concerns **weather data from a number of weather stations**. Let's explore what you might be expected to know...

The Large Data Set.xls



1	You should know the names and rough locations of the 5 UK weather stations, as well as the 3 international weather stations.

The data was recorded for:

- May-Oct 1987
- May-Oct 2015

All the following are daily…

2 You should be familiar with the variables involved and their respective units.

Total rainfall
(in mm)

tr/trace means less than 0.05mm
You can treat these as 0 in calculations!

Mean Windspeed

kn/knot is "nautical mile per hour". $1kn = 1.15\ mph$
Windspeed also given on **Beaufort Scale**:

0 = Calm	$< 1kn$
1-3 = Light	1-10kn
4 = Moderate	11-16kn
5 = Fresh	17-21kn

Mean Visibility
How far (in metres) can be seen into the horizon during daylight hours.

Wind Direction

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	CAMBORNE				© Crown Copyright Met Office 2015								
2	NGR = 1627E 4067N												
3	Altitude = 87 metres												
4	Latitude = 50:22N Longitude = 05:33W												
5													
6	Date	Daily Mean Temperature (0900-0900) (°C)	Daily Total Rainfall (0900-0900) (mm)	Daily Total Sunshine (0000-2400) (hrs)	Daily Mean Windspeed (0000-2400) (kn)	Daily Mean Windspeed (0000-2400) (Beaufort conversion)	Daily Maximum Gust (0000-2400) (kn)	Daily Maximum Relative Humidity %	Daily Mean Total Cloud (oktas)	Daily Mean Visibility (Dm)	Daily Mean Pressure (hPa)	Daily Mean Wind Direction (o)	Cardinal Direction
7	01/05/1987	10.7	3.1	n/a	n/a	n/a	n/a	100	7	2000	1018	360	N
8	02/05/1987	8.9	0.1	n/a	n/a	n/a	n/a	91	3	3200	1020	320	NW
9	03/05/1987	8.1	0	n/a	n/a	n/a	n/a	77	5	3600	1029	350	N
10	04/05/1987	8.2	0	n/a	n/a	n/a	n/a	83	5	4100	1036	350	N
11	05/05/1987	9.8	0	n/a	n/a	n/a	n/a	86	5	2700	1036	10	N
12	06/05/1987	9.3	0	n/a	n/a	n/a	n/a	100	1	1000	1033	330	NNW
13	07/05/1987	10.9	0	n/a	n/a	n/a	n/a	100	3	600	1031	350	
14	08/05/1987	10.5	tr	n/a	n/a	n/a	n/a	89	1	2400	1025	110	
15	09/05/1987	10.9	0	n/a	n/a	n/a	n/a	95	3	900	1017	360	
16	10/05/1987	9.9	0	n/a	n/a	n/a	n/a	79	4	4100	1018	10	
17	11/05/1987	8.8	6	n/a	n/a	n/a	n/a	95	7	2500	1017	270	W
18	12/05/1987	10.2	tr	n/a	n/a	n/a	n/a	97	5	2400	1009	310	NW
19			2.2	n/a	n/a	n/a	n/a	77	4	4600	1016	340	NNW
20			tr	5.9	16	Moderate	35	95	7	3100	1008	290	WNW
21			0	12.3	13	Moderate	27	77	4	4500	1012	10	N
22			tr	11.6	6	Light	16	92	4	3700	1015	290	WNW

	I	J	K	L	M	N
6	Daily Mean Total Cloud (oktas)	Daily Mean Visibility (Dm)	Daily Mean Pressure (hPa)	Daily Mean Wind Direction (o)	Cardinal Direction	Daily Max Gust Corresponding Direction (o)	Cardinal Direction
7	7	2000	1018	360	N	20	NNE
8	3	3200	1020	320	NW	330	NNW
9	5	3600	1029	350	N	350	N
10	5	4100	1036	350	N	350	N
11	5	2700	1036	10	N	10	N
12	1	1000	1033	330	NNW		
13	3	600	1031	350			
14	1	2400	1025	110			
15	3	900	1017	360			
16	4	4100	1018	10			
17	7	2500	1017	270	W	260	W
18	5	2400	1009	310	NW	310	NW
19	4	4600	1016	340	NNW	340	NNW
20	7	3100	1008	290	WNW	270	W
21	4	4500	1012	10	N	10	N
22	4	3700	1015	290	WNW	290	WNW

Mean Pressure
In hectopascals (hPa)

Mean temperature
(in $°C$)

Total sunshine
(nearest $\frac{1}{10}$ of an hour)

Maximum Gust
(in kn) is highest instantaneous wind speed.

Humidity
is the % of air saturation with water vapour. 100% is the maximum % water content air can contain.

Mean Cloud Cover
Oktas means the number of $\frac{1}{8}$ths of the sky covered.

3 You should have a vague idea of the range of values for each location.

UK Location (2015)	Temp Range	Wind Speed Range
Camborne	10-20	3-18
Heathrow	8-29	3-19
Hurn	6-24	2-19
Leeming	4-23	3-17
Leuchars	4-19	3-23

World Location (2015)	Temp Range	Wind Speed Range
Beijing	8-33	2-9
Jacksonville	15-31	1-12
Perth	8-25	4-14

Beijing temp range relatively large.
Min Jacksonville temp high.
Perth similar to UK.

Mean wind speed in UK across full period was roughly 9 nm. But 4 nm in Beijing (i.e. lower), 5 in Jacksonville (again lower), 8 in Perth (similar to UK).

4	You should have a vague idea of the range of values for each variable for the data set as a whole.

Variable	Typical value(s)
Gust (UK only)	8 – 52 nm
Rainfall	0 – 60 mm in UK, but more extreme maximums elsewhere (e.g. 102mm in Perth)
Pressure	988 – 1038 hPa
Wind Speed on Beaufort scale	Max is 'fresh' (5). Most Light or Moderate.
Sunshine (UK only)	0 – 16 hrs
Cloud Cover	0 – 8 oktas (i.e. full spread)

The Large Data Set

Locations

5 UK weather stations

3 overseas

Time Periods

May – October 1987 (6 months)
May – October 2015 (6 months)

Seasons

May/June are the end of spring
July-Sept is summer
October is autumn

Perth (Australia) is in the southern hemisphere, so July-Sept is winter

UK Great Storm

The night of 15-16th October 1987
Gusts up to 100 knots recorded

Florida hurricanes

12 October 1987 Hurricane Floyd
1-2 October 2015 Hurricane Joaquin

Variables Recorded

Daily ~~Maximum~~ mean Temperature
℃

Daily Total Rainfall
mm

Daily Total Sunshine
hours

Daily Maximum Relative Humidity
%; mist and fog if > 95%

Daily Mean Windspeed; Daily Maximum Gust
knots (1kn = 1.15mph)
and Beaufort scale

Daily Mean Wind Direction; Daily Maximum Gust Direction
bearing (°)
and cardinal direction

Cloud Cover
oktas (eighths); 0 – 8

Visibility
Dm (decametres)
1 Dm =10 m

Pressure
hPa (hectoPascal)
1 hPa = 100 Pa

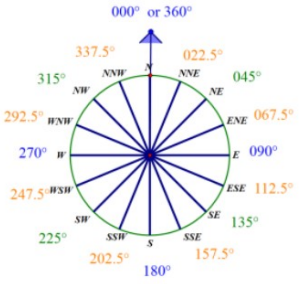
n/a
reading not available

tr (trace)
rainfall < 0.05mm

Beaufort Scale

Discrete, scale of 13 values:
0 (calm, < 1kn)
12 (hurricane, 64kn+)

Cardinal Directions



Oktas

Eighths of the sky covered by cloud
Discrete, scale of 9 values:
0 (clear sky)
8 (completely overcast)

Sources
Maps: *Pearson*
Compass: *mathsmutt.co.uk*

@sxpmaths

LDS Overview.pdf

HURN					© Crown Copyright Met Office 1987	
Date	Daily mean temperature (°C)	Daily total rainfall (mm)	Daily total sunshine (hrs)	Daily mean windspeed (kn)	Daily mean windspeed (Beaufort conversion)	Daily maximum gust (kn)
01/6/1987	15.1	0.6	4.5	7	Light	19
02/6/1987	12.5	4.7	0	7	Light	22
03/6/1987	13.8	tr	5.6	11	Moderate	25
04/6/1987	15.5	5.3	7.8	7	Light	17
05/6/1987	13.1	19.0	0.5	10	Light	33
06/6/1987	13.8	0	8.9	19	Fresh	46
07/6/1987	13.2	tr	3.8	11	Moderate	27
08/6/1987	12.9	1	1.7	9	Light	19
09/6/1987	11.2	tr	5.4	6	Light	19
10/6/1987	9.2	1.3	9.7	4	Light	n/a
11/6/1987	12.6	0	12.5	6	Light	18
12/6/1987	10.4	0	11.9	5	Light	n/a
13/6/1987	9.6	0	8.6	5	Light	15
14/6/1987	10.2	0	13.1	5	Light	18
15/6/1987	9.2	3.7	7.1	4	Light	25
16/6/1987	10.4	5.6	8.3	6	Light	25
17/6/1987	12.8	0.1	5.3	10	Light	27
18/6/1987	13.0	7.4	3.2	9	Light	24
19/6/1987	14.0	tr	0.4	12	Moderate	33
20/6/1987	12.6	0	7.7	6	Light	17

(a) Describe the type of data represented by daily total rainfall.

Alison is investigating daily maximum gust. She wants to select a sample of size 5 from the first 20 days in Hurn in June 1987. She uses the first two digits of the date as a sampling frame and generates five random numbers between 1 and 20.

b) State the type of sample selected by Alison.
c) Explain why Alison's process might not generate a sample of size 5.

HURN					© Crown Copyright Met Office 1987	
Date	Daily mean temperature (°C)	Daily total rainfall (mm)	Daily total sunshine (hrs)	Daily mean windspeed (kn)	Daily mean windspeed (Beaufort conversion)	Daily maximum gust (kn)
01/6/1987	15.1	0.6	4.5	7	Light	19
02/6/1987	12.5	4.7	0	7	Light	22
03/6/1987	13.8	tr	5.6	11	Moderate	25
04/6/1987	15.5	5.3	7.8	7	Light	17
05/6/1987	13.1	19.0	0.5	10	Light	33
06/6/1987	13.8	0	8.9	19	Fresh	46
07/6/1987	13.2	tr	3.8	11	Moderate	27
08/6/1987	12.9	1	1.7	9	Light	19
09/6/1987	11.2	tr	5.4	6	Light	19
10/6/1987	9.2	1.3	9.7	4	Light	n/a
11/6/1987	12.6	0	12.5	6	Light	18
12/6/1987	10.4	0	11.9	5	Light	n/a
13/6/1987	9.6	0	8.6	5	Light	15
14/6/1987	10.2	0	13.1	5	Light	18
15/6/1987	9.2	3.7	7.1	4	Light	25
16/6/1987	10.4	5.6	8.3	6	Light	25
17/6/1987	12.8	0.1	5.3	10	Light	27
18/6/1987	13.0	7.4	3.2	9	Light	24
19/6/1987	14.0	tr	0.4	12	Moderate	33
20/6/1987	12.6	0	7.7	6	Light	17

Calculate:

a) The mean daily maximum temperature for the first five days of June in Hurn in 1987.
b) The median daily total rainfall for the week of 14th June to 20th June inclusive.
c) The median daily total rainfall for the same week in Perth was 19.00mm. Karl states that more southerly countries experience higher rainfall during June. State with a reason whether your answer to part (b) supports this statement.

1. Sara is investigating the variation in daily maximum gust, t kn, for Camborne in June and July 1987.

She used the large data set to select a sample of size 20 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique Sara used.

(1)

(b) From your knowledge of the large data set, explain why this process may not generate a sample of size 20.

(1)

Ex 1E

Blank maps to test your knowledge of the locations

Blank table to test your knowledge of sampling types

	Type	What it is	How to carry out	Advantages	Disadvantages
Random Sampling	Simple Random Sampling				
	Systematic Sampling				
	Stratified Sampling				
Non-random Sampling	Quota Sampling				
	Opportunity/ Convenience Sampling				