Classifying Data

Friday, 1 February 2019 7:25 PM

By the end of the lesson I would hope that you have an understanding and be able to apply to questions the following concepts:

- Know the different between data and variables
- Know the different type of variables:
 - Categorical variables
 - Nominal variables
 - Ordinal variables
 - Numerical variables
 - Discrete variables
 - Continuous variables
- · Know how to compare numerical and categorical variables

RECAP:

This is the start of a whole new series of videos for Further Mathematics!

Welcome to this the start of the Core Data Section.

Further Mathematics is an awesome subject which uses the CAS calculator to help you look at data.

As with most things in Mathematics, we need to make sure we understand the different words and terms which are used in the subject. This lesson will be the start to learn some of the language which will be used.

What is data?

Always love a good definition!

data

/'dertə/ 🕪

noun



"there is very little data available

synonyms: facts, figures, statistics, details, particulars, specifics, features; More

 the quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.

· PHILOSOPHY

things known or assumed as facts, making the basis of reasoning or calculation.

It's always good with some examples isn't it?!

Height (cm)	Weight (kg)	Age (years)	<i>Sex (M male, F female)</i>	Fitness level (1 high, 2 medium, 3 low)	Pulse rate (beats/min)
173	57	18	M	2	86
179	58	19	M	2	82
167	62	18	(M)	1	96
/ 195	84	18	F	1	71
¥ 173	64	18	М	3	90
184	74	22	F	3	78
175	60	19	F	3	88
140	50	34	Μ	3	70

Source: Cambridge Further Mathematics Textbook

What do we notice about the differences?

Some of the data has numbers, other seems to have letters.

Some are whole number values and other seem to be rounded to the nearest whole number.

What is a variable?

variable

/ˈvɛːrɪəb(ə)l/ ♠)

adjective

adjective: variable



- short for variable star.
- the region of light, variable winds to the north of the north-east trade winds or (in the southern hemisphere) between the south-east trade winds and the westerlies.

Sometimes we get rather more than we were expecting!

🤾 Let's say a variable is a quality or quantity about which we record information

What are the variables in the table below?

Height (cm)	Weight (kg)	Age (years)	<i>Sex (M male, F female)</i>	Fitness level (1 high, 2 medium, 3 low)	<i>Pulse rate (beats/min)</i>
173	57	18	M	2	86
179	58	19	Μ	2	82
167	62	18	Μ	1	96
195	84	18	F	1	71
173	64	18	Μ	3	90
184	74	22	F	3	78
175	60	19	F	3	88
140	50	34	Μ	3	70

There are 6 of them!

Barry is at it with the language

It's often useful to be able to describe variables as a certain type. The two main types are Categorical and Numerical.

Let's start with numerical!

English actually tries to sneak in some help with Mathematics. The following two words have the same first three letters ...





Numerical data: Deals with numbers. More specifically things we can count or measure.

Numerical data can be split into two sub types!

Discrete Numerical Variables

Discrete numerical variables can take whole number values only. They are generally things we can count.

Examples might include:

- The number of people in the classroom
- The size of shoes
- The number of TV sets in a house



The best way to think of discrete data is ... ask yourself if you can have a decimal value? For examples:

- can I have 3.4 students in a room?
- Can I have 6.471 shoes? ٠
- Can I have 100.3 televisions? ٠

Continuous numerical variables

Following on from the above examples which can't have decimal values, it would make sense that there must be things which we can measure which take numbers, but involve decimals. These are generally things which we need to measure.

Examples might include:

- ٠ Time
- Distance
- ٠ Weight
- Heights







Running a race, we know that it would make no sense to measure all the competitors times in seconds!



Categorical Variables

Going back to the table, we note that there was a column which didn't have numerical values:

Height (cm)	Weight (kg)	Age (years)	fex (M male, F female)	Fitness level (1 high, 2 medium, 3 low)	Pulse rate (b eats/ min)
173	57	18	M		86
179	58	19	М	2	82
167	62	18	м	1	96
195	84	18	F	1	71
173	64	18	Μ	3	90
184	74	22	F	3	78
175	60	19	F	3	88
140	50	34	Μ	3	70

The column which related to Sex, listed only the letters F and M. Can we place the letters in an order of importance? Can we work out the average of a letter?

Data which uses words or letters is described as **categorical**. This is data which characterises or qualities of people or things. Where no order can be implied, we call the variables as **NOMINAL**.

Examples of categorical data might include:

- the colours of peoples eyes,
- the states people live in, or
- the gender someone is.



Core Data Analysis CH1 Page 4

- the genuer someone is.



Above I talked about categorical data not being able to order people. There are certain types of categorical data which can be used to describe and group objects, but also give an indication of order of importance. This data is called **ORDINAL VARIABLES.**

HINT: Think ORDER and ORDINAL

For example.

We can describe levels of fitness in the following ways:

