Levels of Measurement (Statistics, Data Science)

Levels of Measurement: Nominal, Ordinal, Interval and Ratio – Statology

Level of measurement - Wikipedia

Nominal- names

Ordinal- order

Interval- intervals, add and subtract

Ratio- ratios, multiply and divide

Don't worry about all these terms. Now we will talk about ordinal, which has names and order.

Ordinal measurement level has names and order.

Alphabetical order, months in the year, days of the week, all have names and order. *A B C...* *How happy are you?* *Jan, Feb, March...* *Mon, Tues, Wed...* *off L M H*

Numbers also have order. Ordinal numbers show relative order:

First, second, third... but they don't give an idea of quantity. Saying that someone is second in line, you don't know if they are also the last or if there are a million or more behind them.

For Ordinal levels of measurement, you can put them in order, but there is not a set interval between them.

Cardinal numbers have quantity and come in order. In the US, we use a base-ten number system and the number names also represent a specific amount, length, quantity, etc.

1   2   3

Counting numbers are the whole numbers that start at zero or one.

Mathcounting is a project to help spread the idea of using explicit number naming in English and other languages that don't use it.

Why mathcounting?

Our English counting system has names for the teens that are very confusing! Old English was a base twelve system.

Research and common sense both show that using explicit number names can make math thinking and learning easier.

The Mathcounting Project wants to:

(1) promote mathcounting and

(2) provide research opportunities to research the promotion, adoption, and science of mathcounting.

What is mathcounting?

|		
1 one	11 ten-one	10 ten
2 two	12 ten-two	20 two-ten
3 three	13 ten-three	30 three-ten
4 four	14 ten-four	40 four-ten
5 five	15 ten-five	50 five-ten
6 six	16 ten-six	60 six-ten
7 seven	17 ten-seven	70 seven-ten
8 eight	18 ten-eight	80 eight-ten
9 nine	19 ten-nine	90 nine-ten
10 ten	20 two-ten	100 one hundred is ten-ten

We use a place value system to write the numbers in base ten, with ten digits.

Since we say the largest numbers first, the numbers get higher in value to the left, because that is the direction that we write. Place value chart (math.net)

| Ten thousands	Thousands	Hundreds	Tens	Ones	.	Tenths	Hundredths
8	7	4	8	1	.	2	3

This tells us that 87481.23 has 8 ten thousands, 7 thousands, 4 hundreds, 8 tens, 1 one, 2 tenths, and 3 hundredths:

$$8(10000) + 7(1000) + 4(100) + 8(10) + 1(1) + 2(0.1) + 3(0.01) = 87481.23$$

You can practice with a place value chart to write numbers.

(c) Crainix
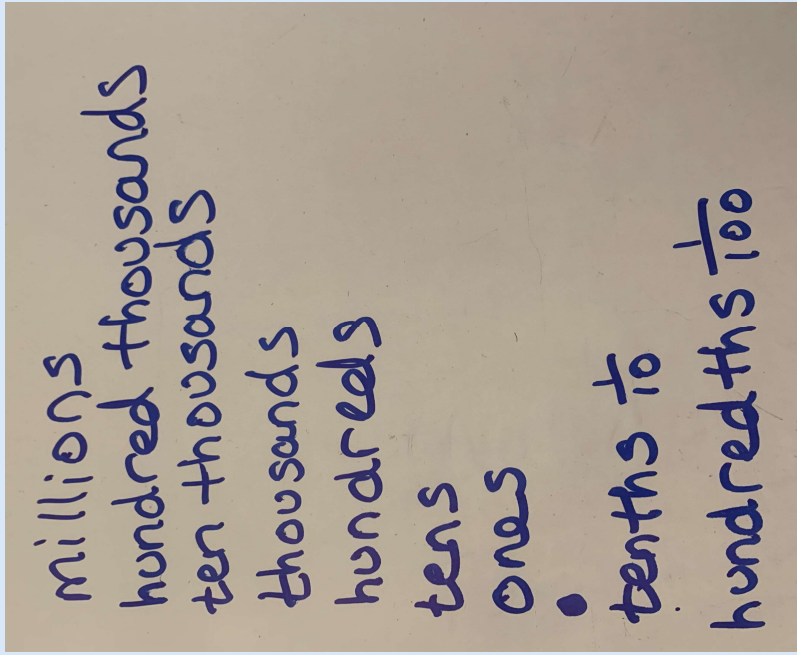
There are new names for bigger numbers! Scientists expand world's measuring unit systems for the first time this century - Los Angeles Times (latimes.com) cgpm – BIPM
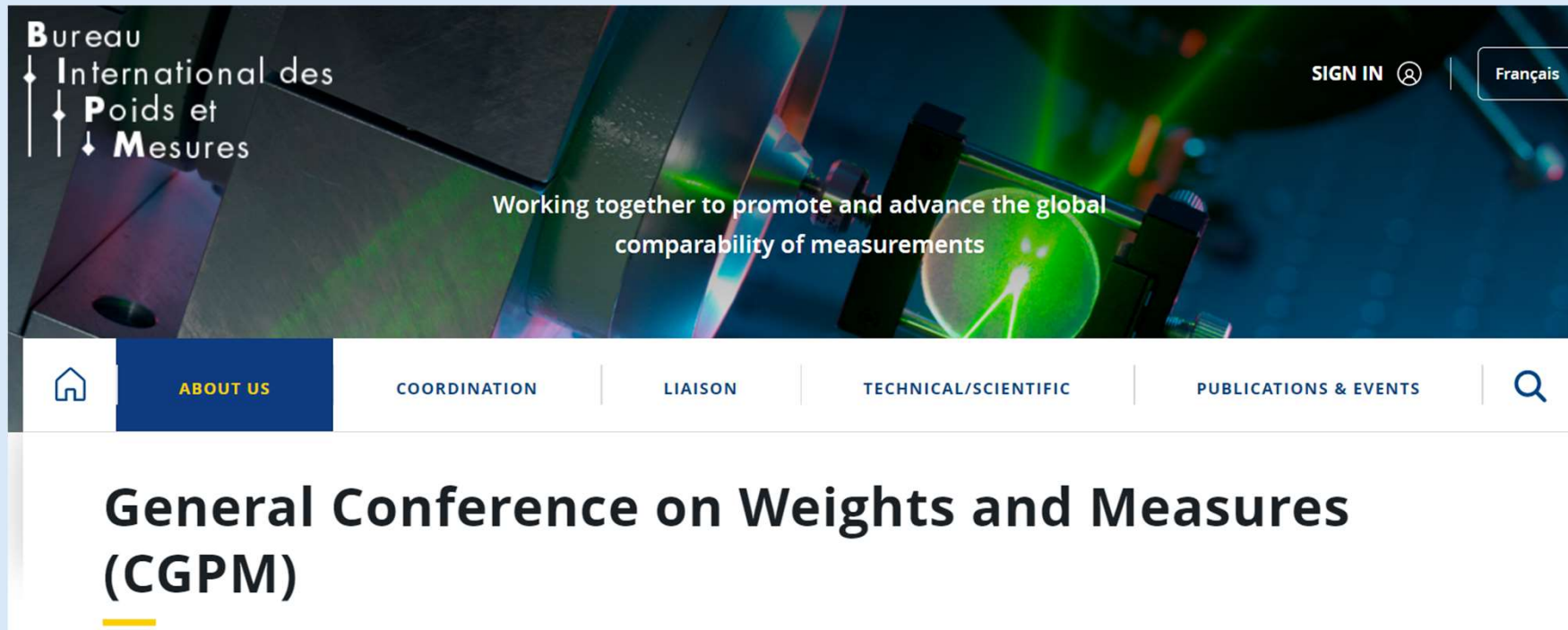
With ordinal data, which has order, you can sort. When you line up data in order, you can look at the number of data points in a category that has order.

Aury, Birdie, Doc, Jae, Simba, and Socrates.

I can order the names in alphabetical order as above. I can look at the Mode, from nominal data, which is the number of names in the starting letter category with the most elements.

Do you see what the mode is here?

In statistics, you can look at the median, which is the data in the middle.

Sometimes it doesn't make sense to talk about the median for nominal data. Since we have an even number of names, the median would be between Doc and Jae.

Aury, Birdie, Doc, Jae, Simba, and Socrates

If I pulled out Jae, and had Aury, Birdie, Doc, Simba, and Socrates, the name in the middle would be Doc, so that would be the median.

We can look at the median with number data.

1, 2, 2, 3, 4 This dataset has 5 counting or whole numbers. The mode is 2 as there are the most 2s. The median is also 2 as that is in the middle.

When there is an even number of datapoints, then we look at the two middle numbers and what is between them.

If we had 1,2,2,3 then the median for the numbers would still be two as we would average the 2s.

If we had 1,2,2,3,4,6 the median would be halfway between 2 and 3, which would be 2 ½ or 2.5 as the arithmetic average of the two middle numbers.

Here the average is arithmetic average or mean, which is the two middle numbers added and then divided by 2. We will talk more about this later with statistics.

1, 2, 2, 3, 4

Notice that we can't call this data a mathematical set {1,2 ,3, 4} because elements are repeated.

We can use the term dataset Data set – Wikipedia for collections of data that can have repetitions.

I will use the term datapoints instead of elements, since we are talking about datasets now.

1, 2, 2, 3, 4

Sometimes square brackets are used for datasets or sets that can have duplicates. In the Python programming language, these are called lists, so we can use that term interchangeably with the term dataset.

[1, 2, 2, 3, 4]

Ordered Pairs and Tuples

With a set, we didn't care about the order and there were no duplicates.

{1,2,3}

With a dataset or list, we can have duplicates with our datapoints. [1,2,2]

With ordinal data, we care about the order.

Tuples are ordered data and parentheses are used for tuples. There can be duplicates and the order of the data matters. (1,2,4) or (1,2,2) or (x,y,z)

Ordered pairs have two datapoints in order like (1, 3) or (-2, 5). In algebra we use (x,y) as our standard ordered pair.

Ordered triplets have three datapoints (x,y,z)

Beyond that, they are called n-tuples with n elements.

The Python programming language is used for Data Science and for programming.

There are four collection data types in the Python programming language:

• **List** is a collection which is ordered and changeable. Allows duplicate members. [1,2,2,3]

• **Tuple** is a collection which is ordered and unchangeable. Allows duplicate members. (1,2,2)

• **Set** is a collection which is unordered, unchangeable*, and unindexed. No duplicate members. {1,3,2}

from https://www.w3schools.com/python/python_lists.asp

That covers lots of ideas around ordinal data or measurement scales.

Order

Place value

Datasets and lists

     Median and mode of a dataset or list

Tuples