DATATYPES

There are Basically Two Types of data:

1.Quantative

2.Qualitative

Quantitative data deals with numbers and things you can measure objectively: dimensions such as height, width, and length. Temperature and humidity. Prices. Area and volume.

Qualitative data deals with characteristics and descriptors that can't be easily measured, but can be observed subjectively—such as smells, tastes, textures, attractiveness, and color.

Qauntative data Are broadly classififed into two:

1.Continous

2.Discreate

Continous:-Continuous Data can take any value (within a range)

Examples:

* A person's height: could be any value (within the range of human heights), not just certain fixed heights,
* Time in a race: you could even measure it to fractions of a second,
* A dog's weight,
* The length of a leaf,

Discreate:Discrete Data can only take certain values.that is intermediate values are not possible

Example: the number of students in a class

We can't have half a student!

Example: the results of rolling 2 dice

Only has the values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12

Or

Discrete data is a count that can't be made more precise. Typically it involves integers. For instance, the number of children (or adults, or pets) in your family is discrete data, because you are counting whole, indivisible entities: you can't have 2.5 kids, or 1.3 pets.

Continuous data, on the other hand, could be divided and reduced to finer and finer levels. For example, you can measure the height of your kids at progressively more precise scales—meters, centimeters, millimeters, and beyond—so height is continuous data.

Discreate is further divide into :

1.Count

2.Categorial

Qualitative Data Types are Brodly classified s:

1.Nominal

2.Ordinal

3.Ratio

4.interval

Nominal values represent discrete units and are used to label variables, that have no quantitative value. Just think of them as „labels“. Note that nominal data that has no order. Therefore if you would change the order of its values, the meaning would not change.

Ordinal values represent discrete and ordered units. It is therefore nearly the same as nominal data, except that it’s ordering matters.

Interval values represent ordered units that have the same difference. Therefore we speak of interval data when we have a variable that contains numeric values that are ordered and where we know the exact differences between the values

The problem with interval values data is that they don’t have a „true zero“. That means in regards to our example, that there is no such thing as no temperature. With interval data, we can add and subtract, but we cannot multiply, divide or calculate ratios. Because there is no true zero, a lot of descriptive and inferential statistics can’t be applied.

Ratio values are also ordered units that have the same difference. Ratio values are the same as interval values, with the difference that they do have an absolute zero. Good examples are height, weight, length etc.

#RANDOM VARIABLE

Random variables describes the probabilty for uncertain future numerical outcome of random process

It is variables because it can take one of the several values

And it is called random because there are some chances associated with te possible value

For example : if we roll a dice proabblity of each no will be ⅙ ,⅙,⅙,⅙,⅙,⅙

#Probablity

Probablity is a chance of something will happen -how likely is that some event will happen

Probability of an event happening P(E) = Number of ways it can happen n(E)/ Total number of outcomes n(T)

Probability is the measure of the likelihood that an event will occur. Probability is quantified as a number between 0 and 1, where 0 indicates impossibility and 1 indicates certainty.

Why probability is important?

Uncertainty and randomness occur in many aspects of our daily life and having a good knowledge of probability helps us make sense of these uncertainties. Learning about probability helps us make informed judgments on what is likely to happen, based on a pattern of data collected previously or an estimate

How Probability is used in Data Science?

Data science often uses statistical inferences to predict or analyze trends from data, while statistical inferences uses probability distributions of data. Hence knowing probability and its applications are important to work effectively on data science problems.

What is Conditional Probability?

Conditional probability is a measure of the probability of an event (some particular situation occurring) given that (by assumption, presumption, assertion or evidence) another event has occurred.