Statistics in Data Science

Class # 02

Today’s Agenda

1. Random Variable
2. Kinds of Variable
3. Variable Measurement Scale
4. Measure of Central Tendency (Mean , Median , Mode)

Random Variable:

A random variable is a property that can take any value

Example: Weight , Height etc

Kinds Of Variable

1. Qualitative Variable/ Categorical Variable
2. Quantitative Variable

Qualitative Variable:

We have categories in Qualitative or categorical variable.

Here we can not do any measurement Like Addition Subtraction etc.

Example: Gender , Blood Group etc

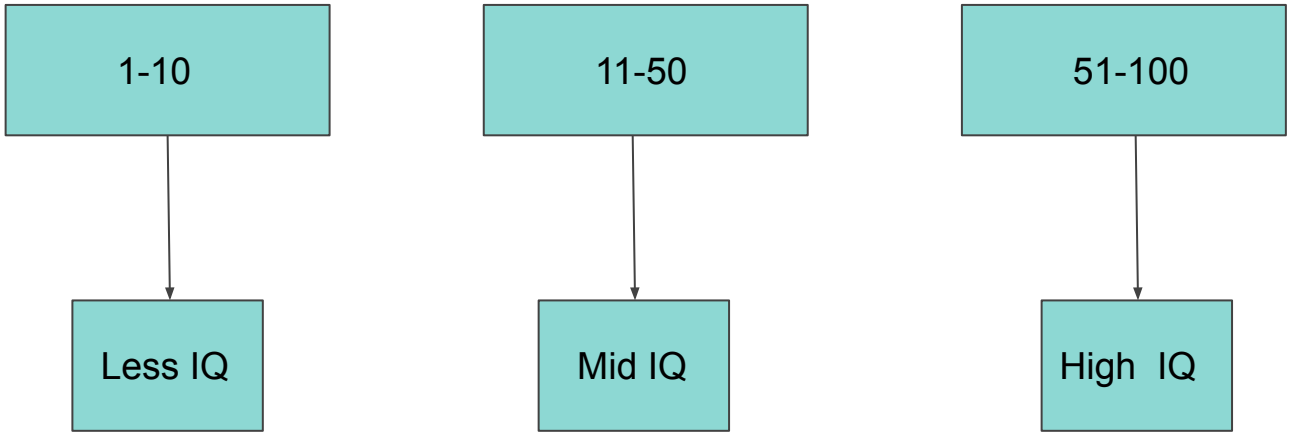
Quantitative Variable:

It can be measured numerically Like Addition Subtraction etc

Example : Height Weight etc

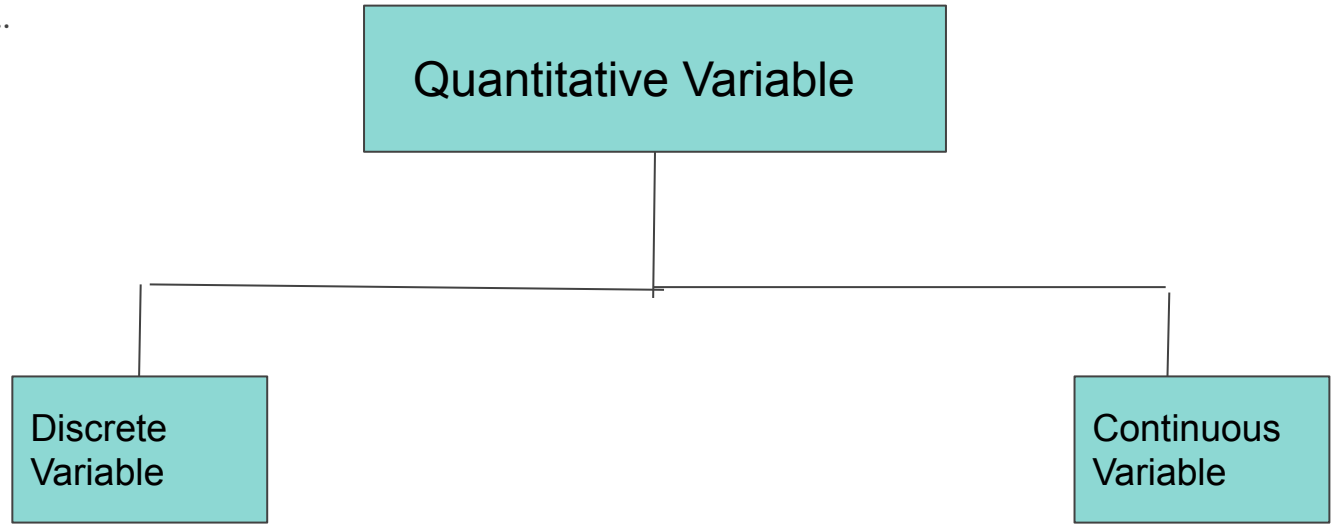
Example Scenario Qualitative Variable:

Let’s we have some IQ classes



Quantitative Variable:

...



Examples of Discrete and Continuous Variable

Discrete Variable:

Bank Accounts , No of Children {2,3,4,5} etc

Continuous Variable:

Height , Weigh , Amount of Rainfall (in inches) {7.6 , 5.2}

Questions:

1. What Kind of Variable Color is:
2. What Kind of Variable River Length is:
3. What Kind of Variable Population in city

Variable Measurement Scale:

There are Four Types of Variable measurement Scale

1. Nominal
2. Ordinal
3. Internal
4. Ratio

Continue....

Nominal Data:

These are basically categorical Data Like Gender , Color etc

Ordinal Data:

Order of the Data Matters but the value does not

Continue...

Example:

Suppose we have a data like:

Marks	Rank
85	1
76	2
54	4
65	3
44	5

Continue...

Explanation

So in this case we doesn't care about marks we know the rank.. In this way we doesn't care about the value we care about order

Continue...

Interval Data:

Order matters value also matters but zero is not Present

Continue....

Suppose we have Data of Time:

12pm 1am 3pm 8am

But if we say 0pm or 0am it does not have any meaning. So in this case value matters order matters but Zero is not present

Continue...

Ratio Data

Assignment

Measure of Central Tendency

What is Central Tendency:

It refers to the measure used to determine the centre point of the data. Whenever we want to find out the centre part of the data we use mean median and mode

Continue….

Mean

Continue….

Median

Continue....

Mode