SOME BASIC DEFINITIONS AND TERMINOLOGIES

DATA

The first step statistics deal with is the collection of numerical data. Statistical data may be classified into two types:

1: Primary/ Raw/ Ungrouped Data.
2: Secondary/ Grouped Data.

Data which are collected for the first time for a specific purpose and are original in nature are know as Primary data while those used in an investigation after doing some arangements/tabulation is called Grouped data.

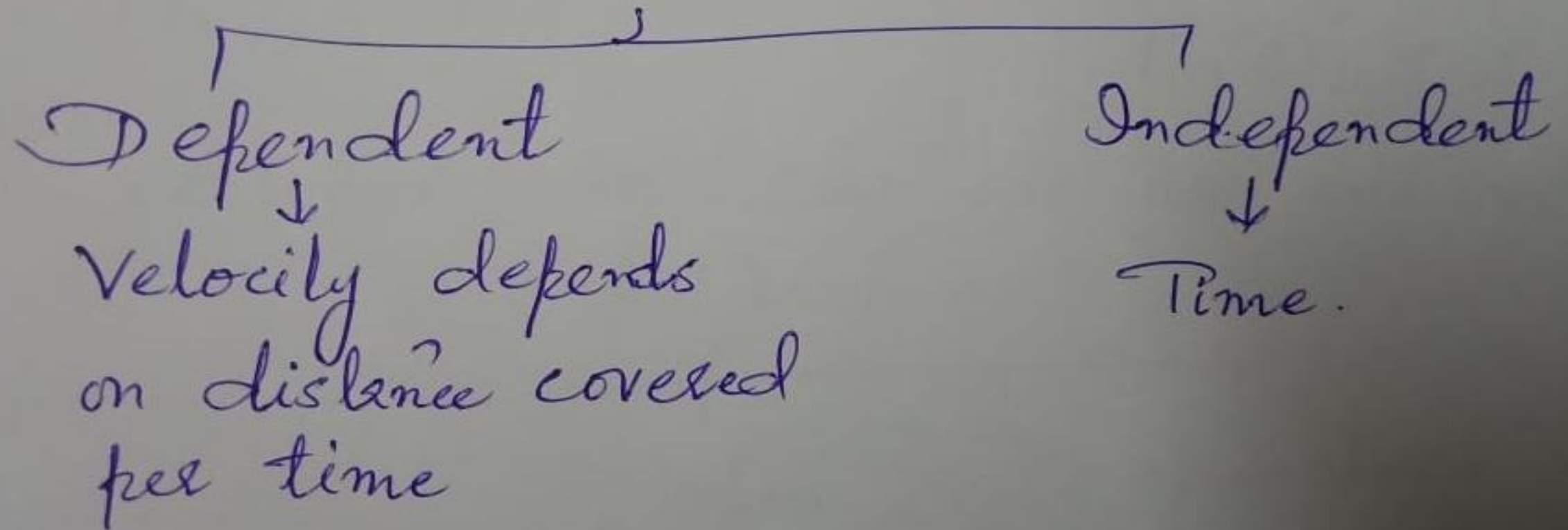Before defining further classification of data lets recall some concepts of basic algebra for generalization.

VARIABLE

Notation used to represent any value of some same collection of objects (sets). Commonly we use x, y, z for denoting a variable. (something that can varry/change)

e.g: $\mathbb{R} \rightarrow$ Set of real numbers

If we want to represent $\mathbb{R}$ with a single notation, we write $x \in \mathbb{R}$ (x is a member of $\mathbb{R}$/ x belongs to set of Real numbers)

CLASSIFICATION:

Variable.

Dependent
↓
Velocity depends
on distance covered
per time

Independent
↓
Time.

We can further divide them into two.
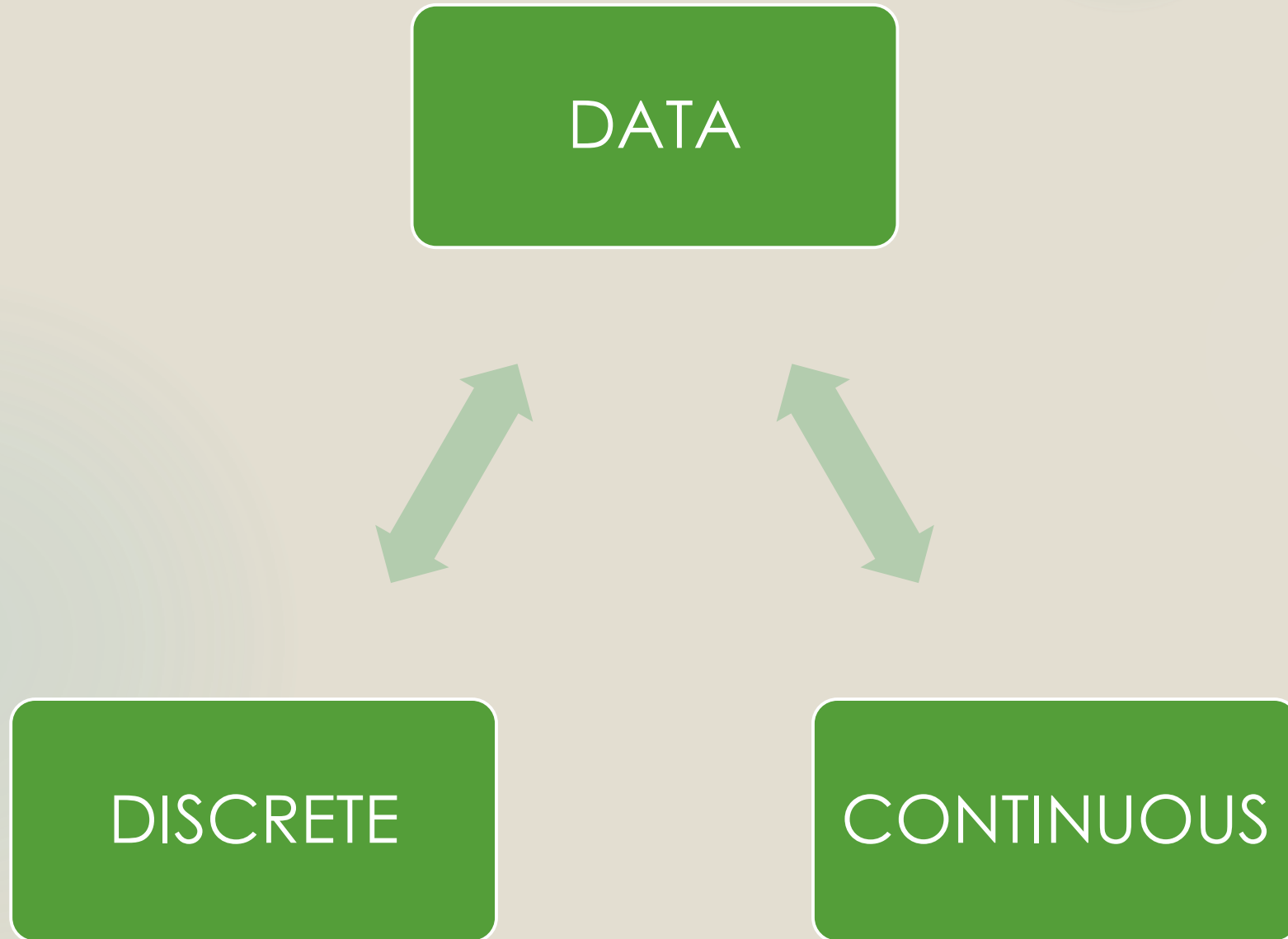
Discrete
↓
Countable like 1, 2, 3
But sometimes uncountable
things are also discrete.
for example: hair on
ur head.

Continuous
↓
Measurable
like real numbers
we have infinite
many real nos
b/w 0 and 1

Therefore, we ca further classify our data into two:

Data describe by discrete variable is called Discrete while data describe by continuous variable is called Continuous.

For example, number of students in class is discrete while height or weight of students is continuous data.

Task

in next 10 minutes, CR of section D and F will tell me their **smart** samples of students of their sections for the data of their academic record (CGPA) in 1st year. Remaining students will classify their data

PRESENTATION OF DATA

There are two methods which may be used for the presentation of data.

1: Frequency Distribution

2: Graphical Representation

Frequency Distribution is statistical table which shows the arrangements of data according to its size with their corresponding number of values side by side while Graphical Representation is some graphs showing data.

STEPS FOR CONSTRUCTING FREQUENCY DISTRIBUTION

1: Determine number of classes:

There is no hard n fast rule for this. The number of classes actually depends on the size of data. Therefore, we should select them with our judgement.

2: Determine class width:

$$h = \frac{\max.value - \min.value}{No.of\ classes} = \frac{Range}{m}$$

3: Deicide starting limit for 1ˢᵗ class and end limit for last class:

Starting point should be minimum value or some value less than min. value and end limit must be maximum value or some value greater than max. value.

Tip: choose integers for limits in case of discrete data and one point less (greater) in case of continuous data.

Note: for next class u have to add h in lower limit of ur 1ˢᵗ class.

Construct a frequency distribution of weights of 50 students of a college, then find mid-values and class boundaries.

$$114 \quad 120 \quad 103 \quad 110 \quad 124 \quad 135 \quad 132 \quad 125 \quad 124 \quad 119 \quad 116$$

$$105 \quad 122 \quad 125 \quad 145 \quad 87 \quad 136 \quad 117 \quad 117 \quad 131 \quad 119 \quad 134$$

$$104 \quad 102 \quad 109 \quad 126 \quad 119 \quad 131 \quad 127 \quad 146 \quad 122 \quad 120 \quad 118$$

$$114 \quad 119 \quad 129 \quad 124 \quad 99 \quad 92 \quad 120 \quad 143 \quad 138 \quad 119 \quad 118$$

$$96 \quad 140 \quad 126 \quad 121 \quad 95 \quad 112$$

Solution:

Maximum value = 146

Minimum value = 87

Range = 59

Let us take 8 classes, then $h = \dfrac{59}{8} = 7375 = 8$ and we construct a frequency table

Class Intervals	Tally Marks	Frequency
85 - 92	//	2
93 - 100	///	3
101 - 108	////	4
109 - 116	//// /	6
117 - 124	//// //// //// ////	19
125 - 132	//// ///	8
133 - 140	////	5
141 - 148	///	3
Total		50

Class Boundaries: For class boundaries find

$$\frac{upper \text{ limit of } 1\text{st class} - lower \text{ limit of } 2\text{nd class}}{2}$$

Add this value in upper limit of each class and subtract in lower limit for getting class boundaries.

Note: no ratta of add/subtract 0.5