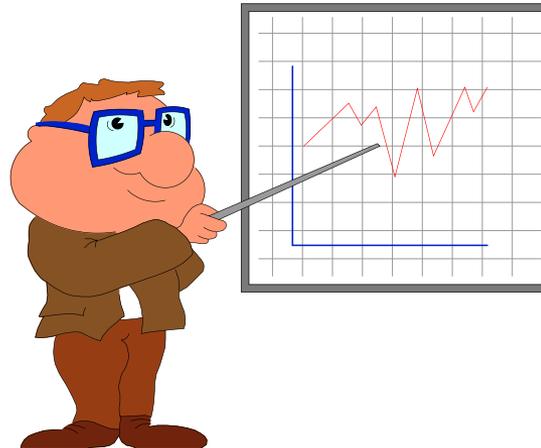TAIBAH UNIVERSITY

Faculty of Science

Department of Math.

جامعة طيبة
كلية العلوم
قسم الرياضيات

Probability and Statistics for Engineers

STAT 301

Teacher :

Lesson 2

Types of Data

Categorical Data

- **The objects being studied are grouped into categories based on some qualitative trait.**

- **The resulting data are merely labels or categories.**

Examples: Categorical Data

- **Eye color**

blue, brown, hazel, green, etc.
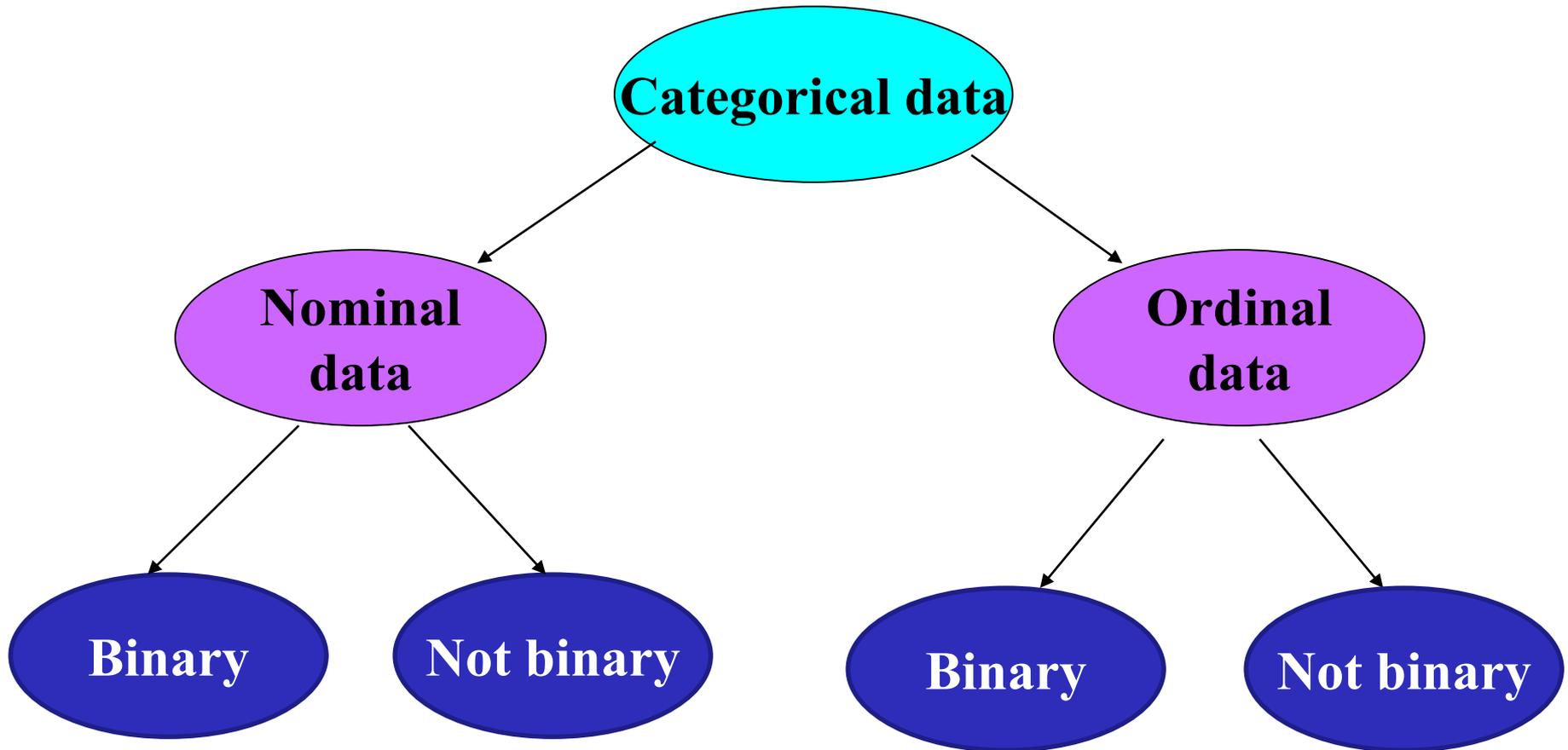
- **Gender:**

Male , Female.

- **Smoking status**

smoker, non-smoker

- **Attitudes towards the death penalty**

Strongly disagree, disagree, neutral, agree, strongly agree.

Categorical data classified as Nominal, Ordinal, and/or Binary

Nominal Data

- **A type of categorical data in which objects fall into *unordered* categories.**

Examples: Nominal Data

- **Gender**

  – **Male . Female .**

- **Nationality**

  – **French , Japanese, Egyptian, Chinese,… etc**

- **Smoking status**

  – **smoker, non-smoker**

Ordinal Data

- **A type of categorical data in which *order* is important.**

Examples: Ordinal Data

- **Class of degree**

  - **1st class, 2nd, 3rd class, fail**

- **Degree of illness**

  - **none, mild, moderate, acute, chronic.**

- **Opinion of students about stats classes**

  - **Very unhappy, unhappy, neutral, happy, ecstatic!**

Binary Data

- **A type of categorical data in which there are *only two categories*.**

- **Smoking status- smoker, non-smoker**

- **Attendance- present, absent**

- **Class of mark- pass, fail.**

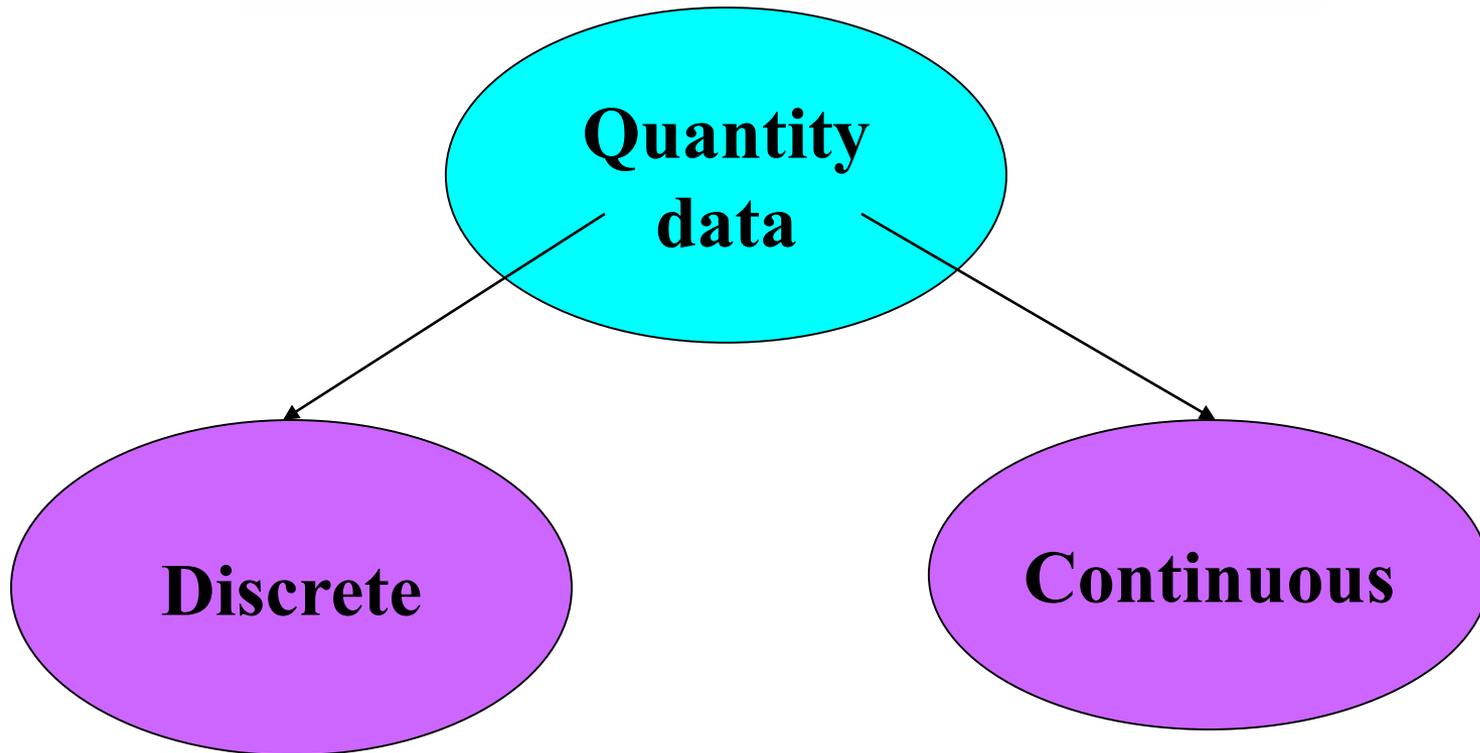- **Status of student- undergraduate, postgraduate.**

Quantity Data

- **The objects being studied are 'measured' based on some quantitative trait.**

- **The resulting data are set of numbers.**

Examples: quantity Data

- **Pulse rate**

- **Height**

- **Age**

- **Exam marks**

- **Time to complete a statistics test**

- **Family Size**

Quantity data can be classified as 'Discrete or Continuous'

Discrete Data

If the values / observations belonging to it may take only specific values[(integer) .

 There are gaps between the possible values).
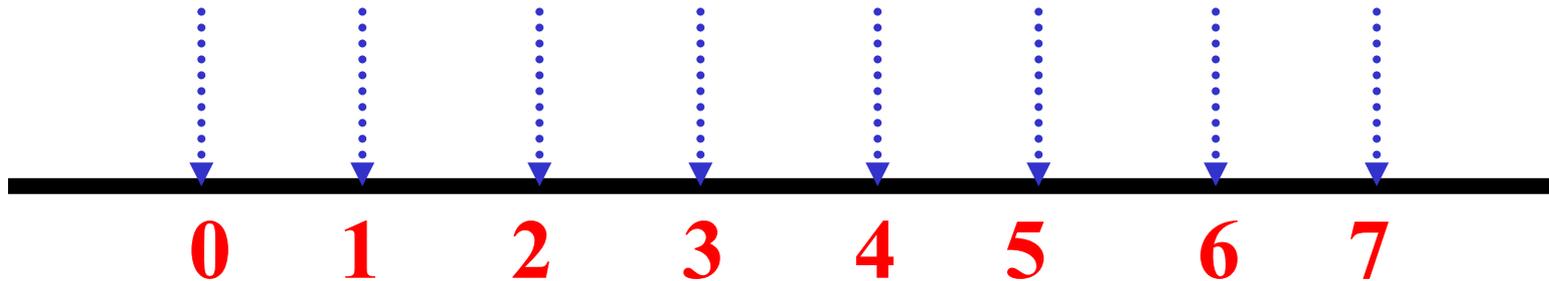
It does not containing fraction.

Implies counting.

Continuous Data

If the values / observations belonging to it may take on any value within a finite or infinite interval (real).

Can contain fraction.

Implies Measurement.

Discrete data -- Gaps between possible values- count

0 1 2 3 4 5 6 7

**Continuous data
no gaps between possible values- measure**

0   1000

Examples: Discrete Data

- **Number of children in a family**
- **Number of students passing a stats exam**
- **Number of crimes reported to the police**
- **Number of cars sold in a day.**

Generally, discrete data are counts.

We would not expect to find 2.2 children in a family or 88.5 students passing an exam or 127.2 crimes being reported to the police or half a bicycle being sold in one day.
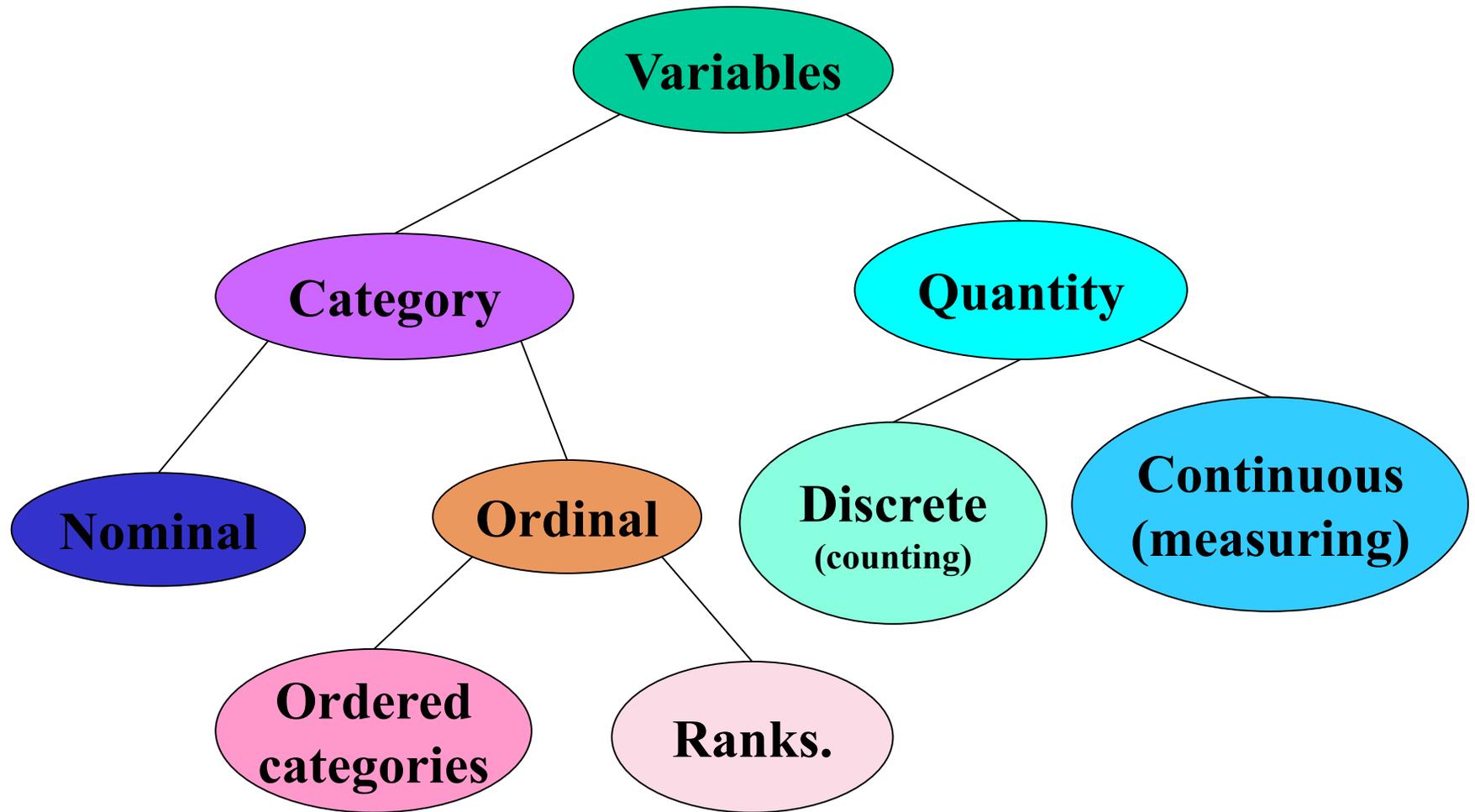
Examples: Continuous data

- **Weight**
- **Height**
- **Time to run 500 metres**
- **Age**

'Generally, continuous data come from measurements.

(any value within an interval is possible with a fine enough measuring device.).

Relationships between Variables.

Interval and ratio variables

- **Interval**:

  - **Numerical data**

  - **data can be ranked**

  - **Data has equal intervals between data points**

  - **There is no meaningful zero**

  - **ratios are meaningless.**

Interval and ratio variables

- **Ratio**:

  – **Numerical data**

  – **data can be ranked**

  – **Data has equal intervals between data points**

  – **True zero**

  – **True ratios exist between the different units of measure.**

Interval and ratio variables

- **Difference between interval and ratio usually not important for statistical analysis .**