



SNS COLLEGE OF TECHNOLOGY

Coimbatore - 35



23BAT613 – Artificial Intelligence for Managers

Unit I – TECHNOLOGY OVERVIEW AND FUNDAMENTALS



Presented by

Ms.S.D.Shamini
Design Thinker

Redesigning Common Mind & Business Towards Excellence

1st Indian
Institution
to Implement
& Patent
Design
Thinking
FrameWork

sns
INSTITUTIONS
www.snsgroups.com

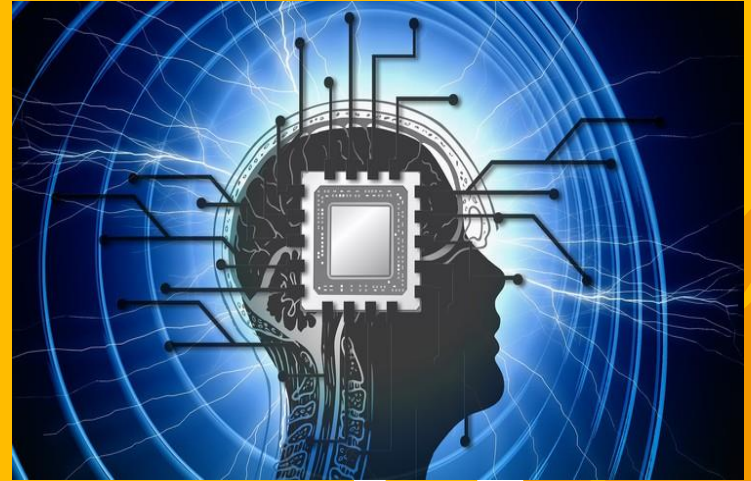
Sri SNS Charitable Trust
CELEBRATING
25
th Year
SILVER JUBILEE

Build an Entrepreneurial Mindset through our Design Thinking FrameWork



Recap

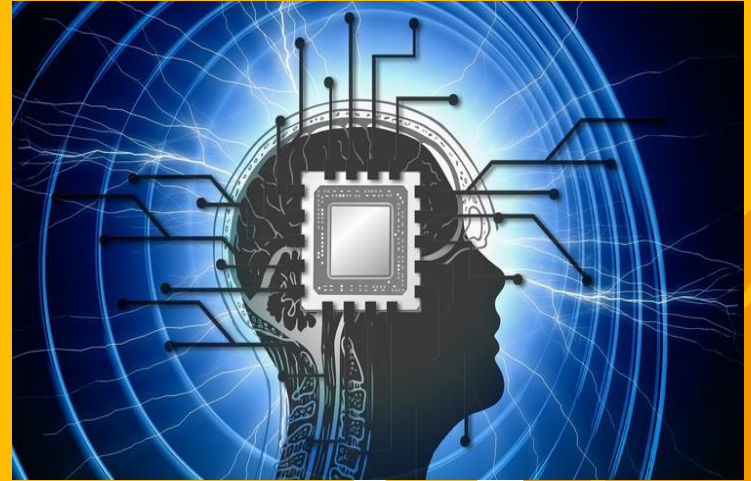
- Importance to Study Data
- Importance of Data Manipulation
- Techniques for Data Manipulation
- Importance of Attributes in DBMS
- Types of Attributes
- Example





Discussion about....

- Statistics Meaning
- Statistical Learning in AI
- Categories and Algorithms of ML
- Statistical Methods for Data Analysis





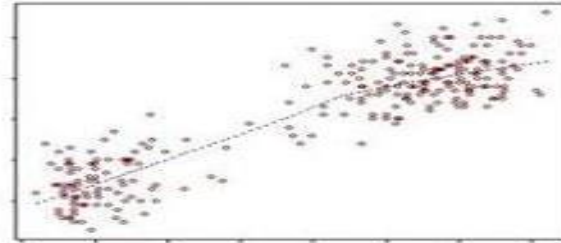
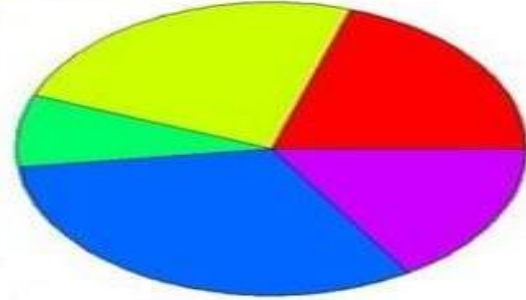
Reason to study Importance of Data





Statistics

“Statistics is the science of collecting, organizing, presenting, analyzing and interpreting numerical data to assist in making more effective decisions.”





Elementary Statistics

IMPORTANT TERMS

- Categorical/qualitative data** falls into groups or categories.
EX: State, type of pet, and gender.
- Quantitative data** represents counts or measurements.
EX: Profit, number of people in line, and lifetime of a product.
- Population:** The entire group studied.
EX: All U.S. registered voters.
- Sample:** A subset of a population from which data is collected.
EX: 1,000 randomly sampled, registered U.S. voters.
- Parameter:** A number that summarizes a population and is typically unknown.
- EX: The average price of gas in the U.S.
- Statistic:** A number that summarizes a sample.
EX: The average price of gas from 1,000 gas stations randomly selected from the U.S.

DATA TABLES

- Frequencies:** The number in each group.
- Relative frequencies:** The percentage in each group.

Data Table: Type of Pet		
Type of pet	Frequency	Relative frequency
Dog	10	.33
Cat	15	.50
Other	5	.17
Total	30	1.00

Data Table: Age Group		
Age group	Frequency	Relative frequency
Under 18	100	.24
18-30	200	.47
Over 30	125	.29
Total	425	1.00

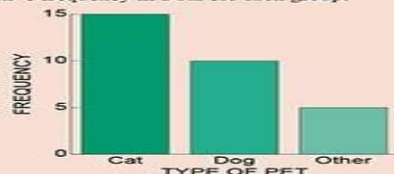
GRAPHS FOR SINGLE VARIABLE

The purpose of a graph is to show a visual of data.
PIE CHART
Pie chart: A graph of categorical data showing frequency or relative frequency in a circle with a slice for each group.



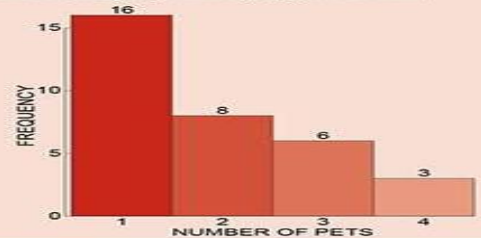
BAR GRAPH

Bar graph: A graph of categorical data showing frequency or relative frequency in a bar for each group.



HISTOGRAM

Histogram: A graph of quantitative data with the variable on the X-axis divided into groups (bars) and the frequency or relative frequency in each group on the Y-axis.



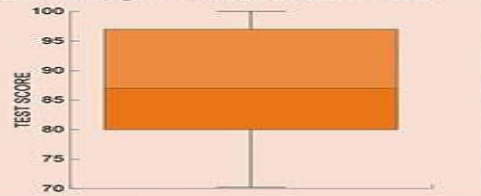
STEM-LEAF PLOT

Stem-leaf plot: A numerical graph of quantitative data with the last digit on the right of the line and the leading digits to the left of the line. You can recover the data set from a stem-leaf plot.
EX: Test scores with 70, 70, 71, 71, 76, and 79 as the lowest 6 scores.

Test Scores	
Stem	Leaf
7	001169
8	00447779
9	0123777799
10	00

BOXPLOT

- Boxplot:** A one-dimensional graph of quantitative data that shows the locations of the 5-number summary.
- 25% of the data lies in each section.
- A box contains the middle 50% of the data.
- The line in the box indicates the median.
- The lines coming out of the box end at the minimum and maximum.
- In this boxplot, the minimum = 70, Q1 = 80, Q2 = the median = 87.5, Q3 = 97, and the maximum = 100.



DESCRIPTIVE STATISTICS FOR SINGLE VARIABLE

Summarize quantitative data to indicate the center, variation, and relative standing.

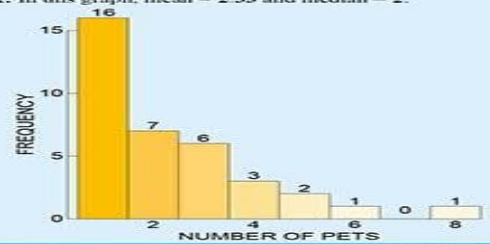
MEASURES OF CENTER

Measures of center indicate where the "middle" of the data is in different ways.

- Mean:** The average of the data set. Sample mean is $\bar{x} = \frac{1}{n} \sum x_i$, where:
 - x_1 = 1st data value
 - x_2 = 2nd data value
 - x_n = last data value (n = number of values)
 EX: Given 2, 4, and 5, the average is $\frac{2+4+5}{3} = 3.67$.
- Median:** Divides the ordered set in half.
 - If the data has an **odd** sample size, the median is the middle value.
EX: Given 1, 2, 3, 4, and 5, the median is 3.
 - If the data has an **even** sample size, the median is the average of the two middle values.
EX: Given 1, 2, 3, and 4, the median is $\frac{2+3}{2} = 2.5$.
- Mode:** The data value that occurs most often. This is not a good measure of center.

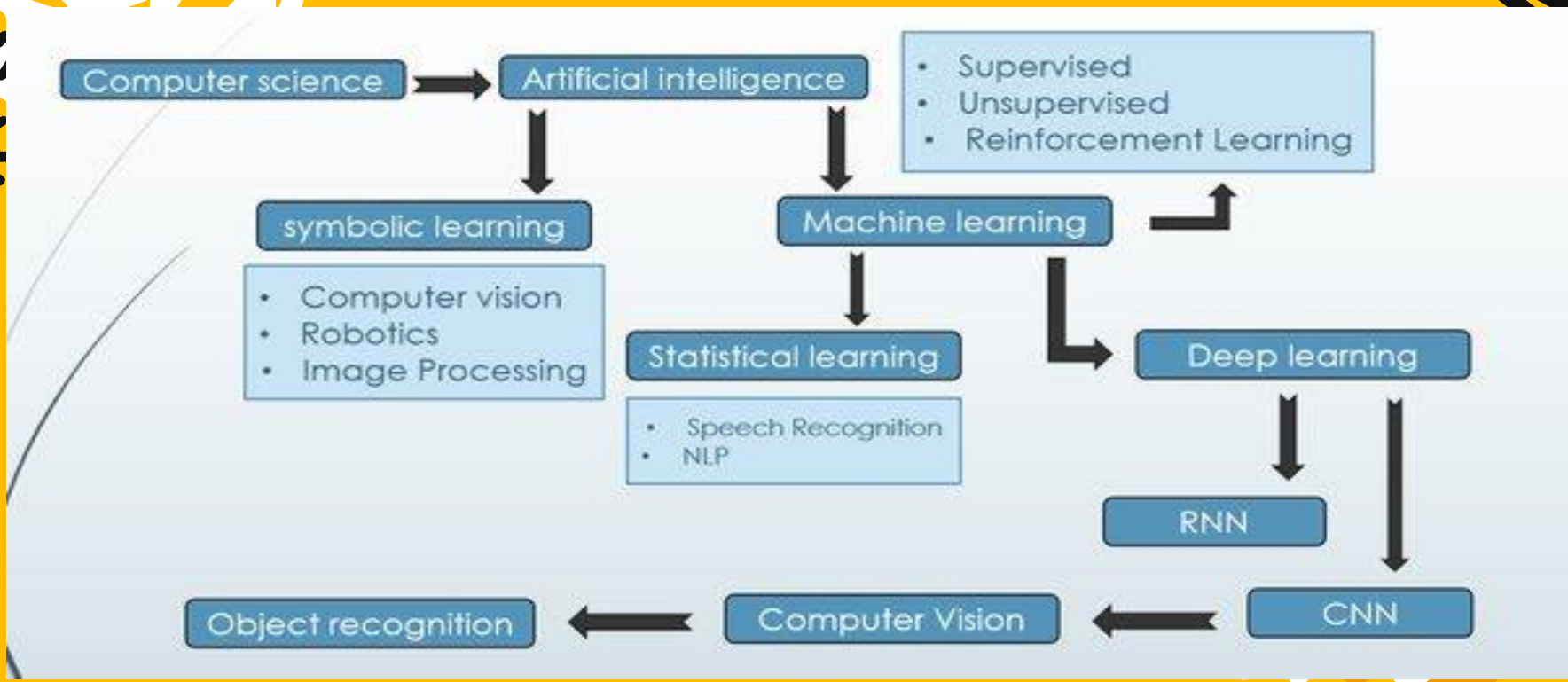
PROPERTIES

- Mean is affected by outliers; median is not.
- If mean > median, data are skewed right.
- EX: In this graph, mean = 2.33 and median = 2.



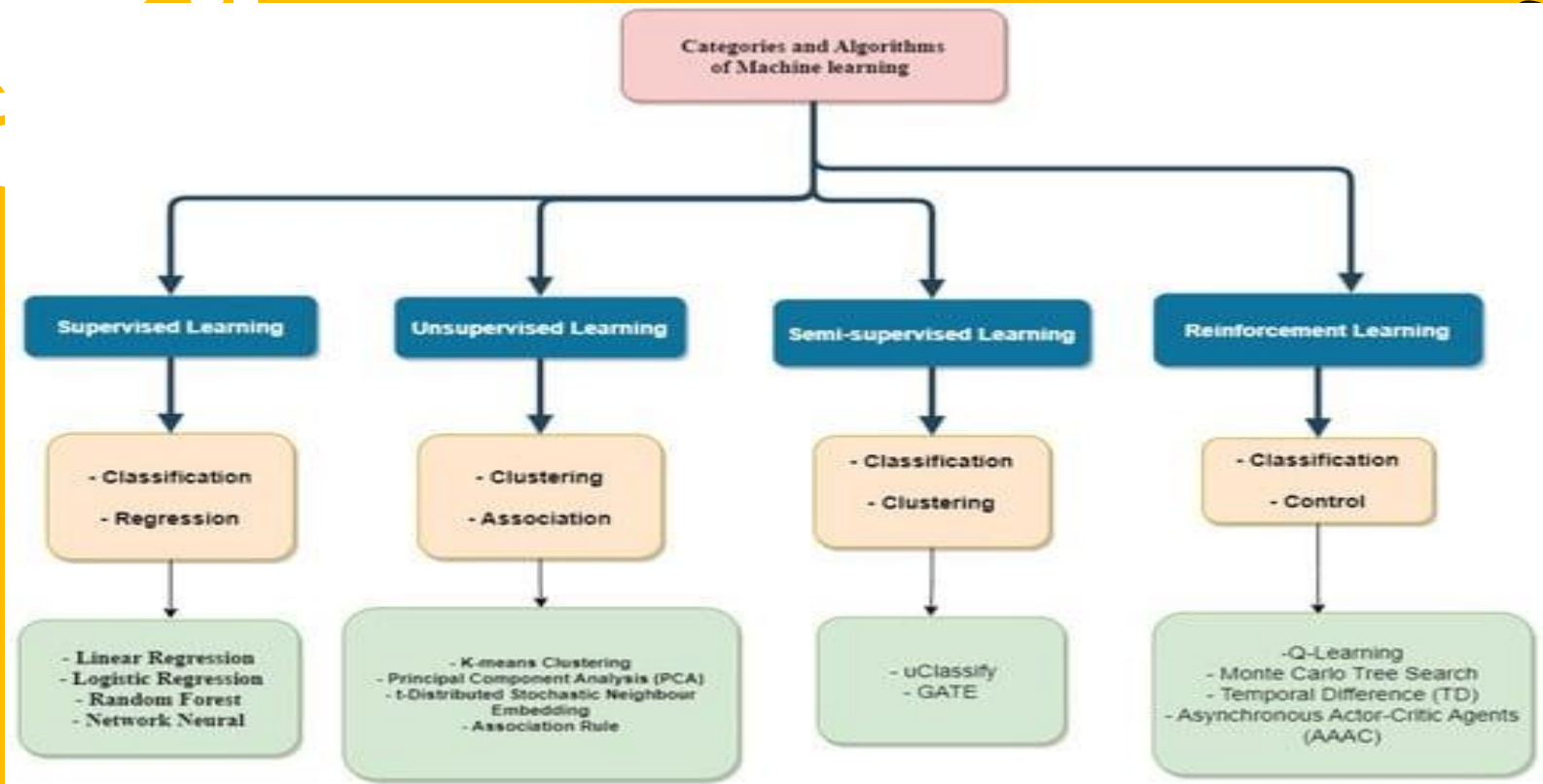


Statistical Learning IN AI

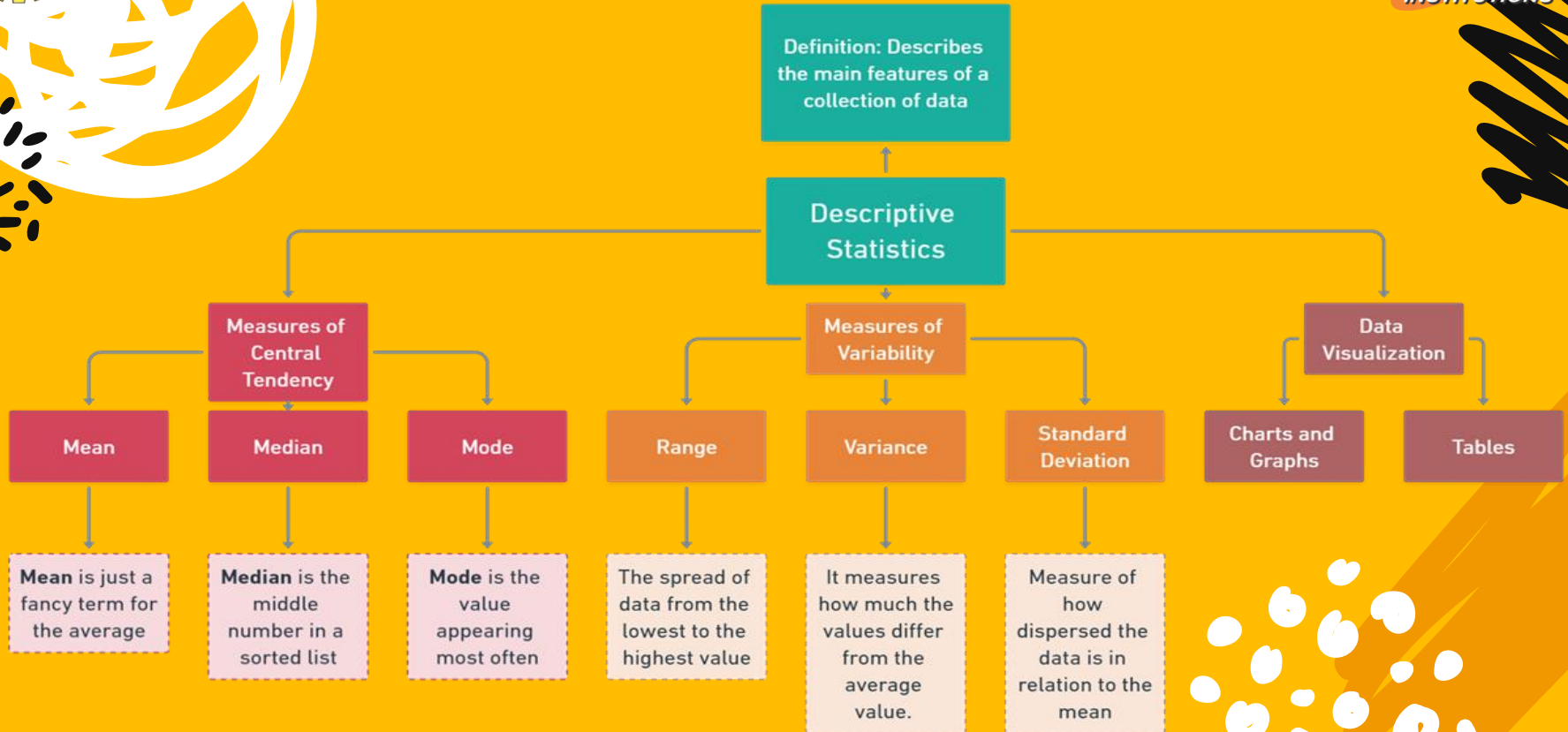




Categories and Algorithms of ML



Statistical Methods for Data Analysis





Knowledge Check

Which statistical technique is commonly used for unsupervised learning tasks such as clustering?

- X a) Logistic regression
- X b) K-means clustering
- X c) Decision trees
- X d) Support Vector Machines (SVMs).
- X **Answer: K-means clustering**



Summary

- Statistics Meaning
- Statistical Learning in AI
- Categories and Algorithms of ML
- Statistical Methods for Data Analysis





References

- <https://www.relyservices.com/blog/importance-data-processing-in-machine-learning>





REACH US



snsinstitutions



snsinstitutions



snsinstitutions



snsinstitutions



snsinstitutions

Thanks!

