Ch - 1

Q) What is data ?

→ Data is a raw fact information and statistics which can be in various forms such as numbers, text, sound, images, videos or any other forms.

Q) Why data mining ?

→ The ~~explorus~~ explosive growth of data (from terabytes to petabytes) that means production of data is too much in various areas. Due to increase in size of database, increase of computerized growth in the society, automatic analysis overcome on the manual analysis which shows the need of data mining.

3/9/25

Evolution of Data :-

1960 → collect data or create database
1970 → database management system (DBMS) (using SQL)
1980 → advancement database (RDBMS)
19920 - 2000 → Data mining.

Q) What is data mining ?

→ • Data mining refers to extracting or mining knowledge from large amount of data.
• Another term used for data mining is knowledge discovery from data. (KDD)

Teacher's Signature

Date: 3/9/25

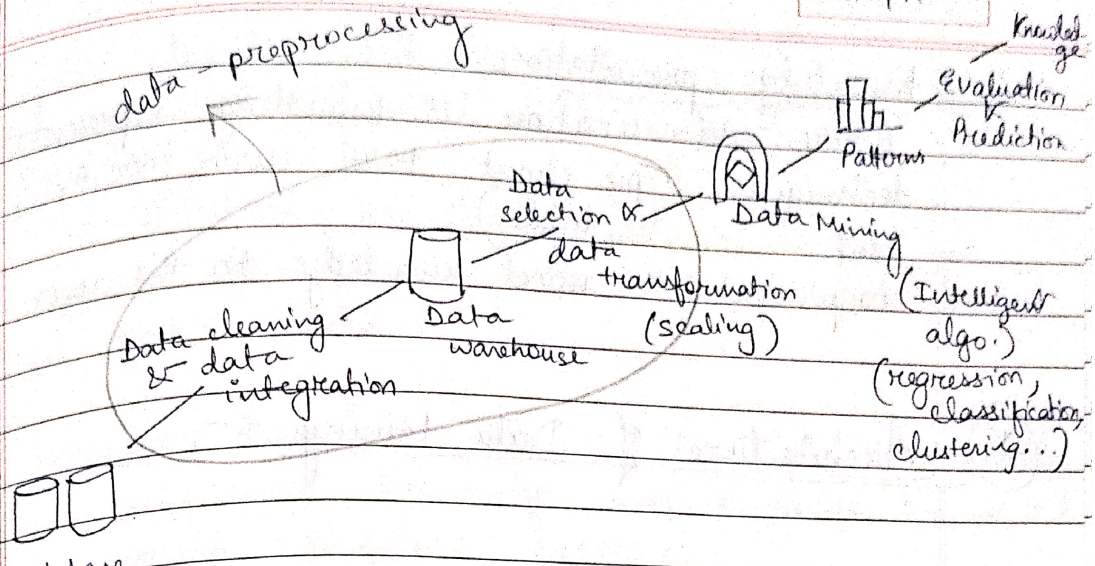


fig : Steps of KDD

KDD → Extraction of interesting info. or patterns from data, in large database is KDD. It consists of 7 steps :-

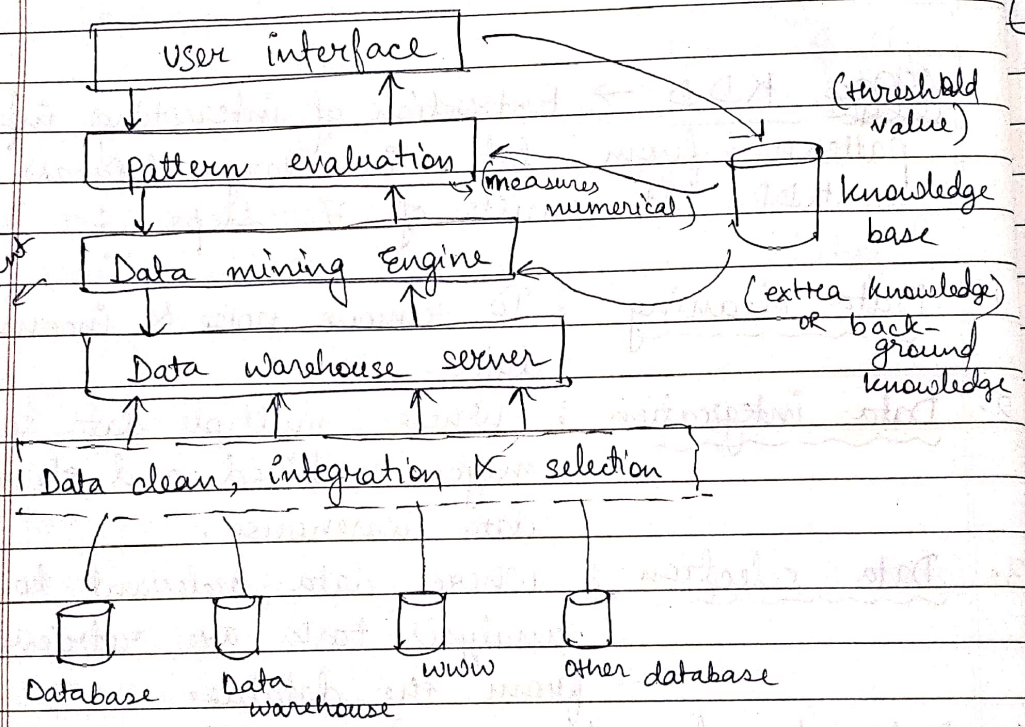
1. Data cleaning : To remove noise & inconsistent data.
2. Data integration : Where multiple data sources maybe combined and stored in data warehouse.
3. Data selection : Where data relevant to the analysis task are retrieved from the database.
4. Data transformation : Where data are transformed or consolidated into appropriate forms for mining by performing scaling.
5. Data Mining : An essential process where intelligent methods are applied in order to extract data patterns.
6. Pattern evaluation : To identify truly interesting patterns representing knowledge based on measures.

Teacher's Signature

Date: 5/9/25

7. Knowledge presentation :
Where visualization & knowledge representation techniques (pie chart, bars, curve {ROC, AUC}, tables) used to present the mined knowledge to the user.

(*) Architecture of Data Mining :-



(1) Database, Data warehouse, WWW, other database :

This is set of database, datawarehouse spreadsheet or other kind of information repository.

Teacher's Signature

(2) Data cleaning and data integration techniques may be performed on the data.

(3) The database or data warehouse server is responsible for fetching the relevant data based on user's request.

(4) Knowledge base – This is the domain knowledge or background knowledge used to guide the search or evaluate resulting patterns.

(5) Data Mining Engine – This is essential to the datamining system & ideally consist of set of modules for task prediction & evaluation analysis.

(6) Pattern evaluation – This component typically employ interestingness measures and interact with data mining modules to focus the search.

(7) User interface – This module communicates between user & data mining system, allow the user to interact with the system.

(*) Predictive Analytics :-

It is the process of extracting information from large dataset in order to make prediction & estimate about future outcomes.

(1) Eg – SBI, is the largest public sector bank in India. Around 70% of the banks customer access through multiple channels and requisite data warehousing facility.

Teacher's Signature _____

Date: 5/9/25

E.g (1) Loan default prediction :

↳ Data Mining techniques
Analysing past records, the banks find the customers with high credit utilisation & irregular payment history are likely to default.

↳ Predictive Analytics
When a new customer apply for a loan, the bank uses this partner to predict likelihood of default.

E.g (2) : Healthcare :-

↳ Data Mining knowledge
Hospital studies patient history, and discovers that people high BP, obesity & smoking habits are more prone to heart disease.

↳ Predictive Analytics
A doctor uses this patterns in a predictive model to identify patients who are at risk of developing heart disease in the future.

(*) Data Mining on what kind of Data?

Here we will discuss the number of different data repositories on which mining can be performed.

Teacher's Signature

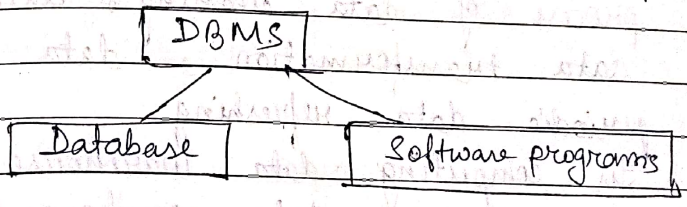
GENIUS
Date 6/9/25

The data repositories are :-

1. Relational database
2. Data warehouse
3. Transactional database
4. Advanced database system
   a) Object relation database (ORD)
   b) Temporal database, Sequence Database
   c) Spatial database, Spatio temporal Database
   d) Time - series Database
   e) Text database & Multimedia database
   f) Hetergenous database & Legacy database
5. WWW

(1) Relational Database :-

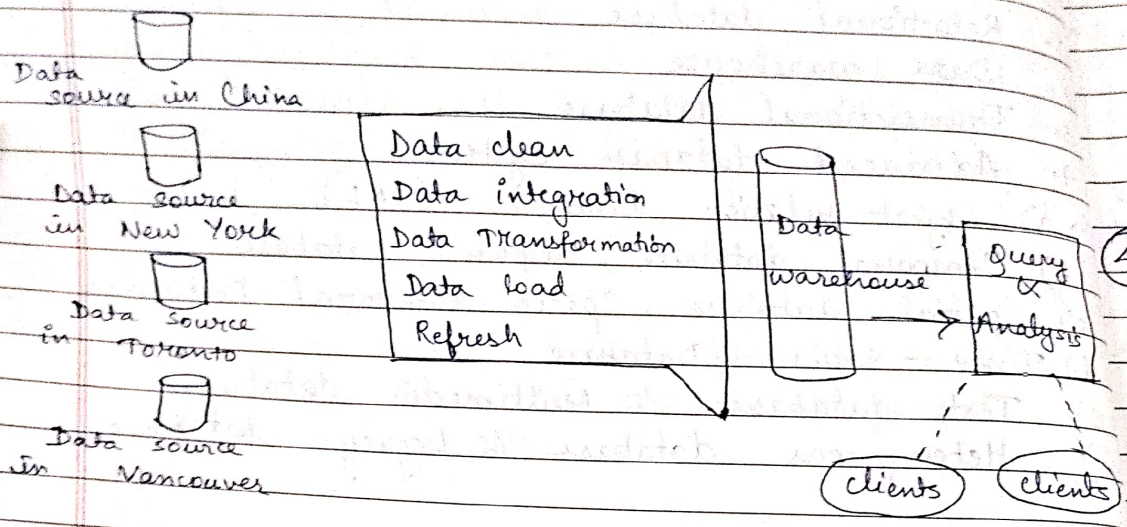
- DBMS consists of database and set of software programs.



- Database consist of a collection of interrelated data.
- Software programs are used to manage & access the data stored in the database.
- The software programs involve mechanisms like database structures, share or distributed ~~star~~ data access and for ensuring the consistency and security of the information.
- A relational database is a collection of tables each of which is assigned a unique name.
- Each table consist of set of attribute & a large set of tuples.

Teacher's Signature

(2) Data Warehouse :-



- A data warehouse is a repository of information collected from multiple sources stored under unified scheme.
- Data warehouse are constructed via a process of data cleaning, data integration, data transformation, data loading & periodic data refreshing.
- In computing, data warehouse also known as enterprise data warehouse (EDW) is a system used for reporting and data analysis.

(3) Transactional database :-

- Transactional database consists of a file where each record represents a transaction.
- A transaction includes a unique transaction identity number (trans ID) and a list of the items making up the transaction.

- The transactiona
regarding the
transaction,
number of
sale occurred

(A) Advanced
It includes
→ object
→ spatial d
→ temporal
& time-
→ text dat
→ heterogenou

i) ORD :-
- In ORD
each ob
set of
- ORD su

ii) Tempor
→ It st
related
→ The
to pa

Seque
→ A s
event
time
clic

Teacher's Signature

- The transactional database also stores information regarding the sale such as date of the transaction, the customer ID number, ID number of salesperson and branch at which sale occured & so on.

(4) Advanced database systems :-

It includes

→ object relational databases (ORD)
→ spatial database & spatio temporal databases
→ temporal databases, sequence databases, & time-series databases.
→ text databases, multimedia databases.
→ heterogeneous database, legacy databases

i) ORD :-

- In ORD each entity consider as object & each object associated with a set of variables, set of messages, set of methods.
- ORD supports objects, classes & inheritance.

ii) Temporal Database, sequence :-

→ It stores relational data that include time-related attribute.
→ The attributes which stores information relating to past, present and future times.

iii) Sequence Database :-

→ A sequence database stores sequences of ordered events, with or without a concrete notion of time. E.g - Customer shopping sequences, web click streams.

Teacher's Signature _____

(iii) Spatial Databases :-

→ A spatial database is a database optimized for storing & quering data that represents objects defined in a geometric space. E.g - points, lines and polygons.

→ It is used in geography, remote sensing, urban planning & natural resource management.

- Spatial Temporal Database :-

→ It manages both space & time information E.g - Tracing of moving objects.

→ A spatial database that stores spatial objects that change with time is called spatio-temporal database.

(iv) Time-series Databases (TSDB) :-

→ It is a computer system that is designed to store and retrieve data records that is a part of a "time series" which is a set of data points that are associated with time-stamps.

(v) Text databases & Multimedia databases :-

→ Text databases are databases that contain word descriptions for objects.

→ The words are not simple keywords but rather long sentences or paragraphs, such as product specifications, errors, warning messages, etc.

- Multimedia → It stores images, audio & video data.
- Heterogenous DB :-

→ It consists of a set of interconnected autonomous component databases.

→ The component communicate in order to exchange info. & answer queries.

- Legacy DB → It is a group of heterogenous DB that combines diff. kinds of datasystems such as relational or object oriented DBs, hierarchical DB, n/w databases, spreadsheets, multimedia DB or file system.

(5) WWW (World wide web) :-
Web mining can be defined as the method of utilizing data mining techniques & algo. to extract useful info. directly from the web.

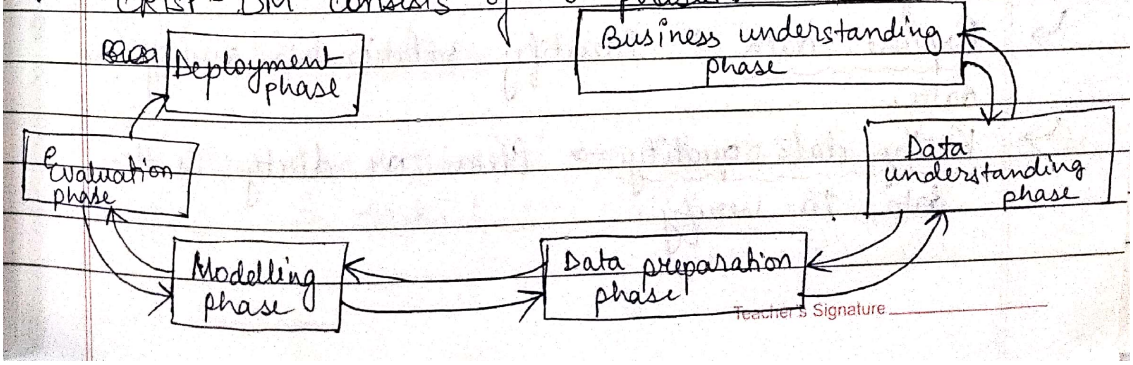
Date: 9/9/25

Q) Differentiate between Data Mining & Predictive Analytics.

Data Mining	Predictive Analytics
1) It is a process of extracting knowledge from large dataset.	1) Apply the knowledge to predict future outcome.
2) Goals : To explore & understand data.	2) Goals : To predict & forecast future values.
3) Approach : Descriptive (find out the hidden patterns).	3) Approach : Predictive (focus on building models from past data to unseen data.
4) Techniques : Clustering, regression, classification, etc.	4) Techniques : Decision tree, random forest, neural network.
5) Output : Patterns	5) Output : Probabilities
6) Data used : Present as well as past data.	6) Data used : Only past data.

(*) CRISP - DM :-
Cross Industry Standard Process Data Mining

- CRISP-DM is a standardised data mining process used for fitting data mining into the general problem solving strategy of a business or a research.
- CRISP-DM consists of 6 phases:



Teacher's Signature

GENIUS
Date: 9/9/2

a) Business Understanding phase :-
This phase focus on understanding the objectives & requirements of the project. This phase having 4 task :

↳ Determine business objectives →
You should first thoroughly understand what the customer really want.

↳ Assess situation →
Determine resources availability, project requirement, assess risk & contigency and conduct a cost benefit analysis.

↳ Determine data mining goals →
In addition to defining business objectives, what success data mining technical perspective should be used.

↳ Product project plan →
Select technologies & tools & define details plan for each project phase.

b) Data Understanding phase :-

It drives the focus to identify, collect & analyse the dataset. This phase also have 4 task :

↳ Collect initial data → Acquire the necessary data & load it into your tools.

↳ Describe the data → Examine the data & document like data format, no. of records or fields.

↳ Explore data → Identify relationship among the data.

↳ Verify data quality → Clean or dirty is the data to verify.

Teacher's Signature

c) Data Preparation phase :-

It prepares the final dataset for modelling. It is having 5 tasks :

↳ Select data → Determine which dataset will be used.
↳ Clean data → To correct, impute or remove erraneous values.
↳ ~~Cont~~ Construct data → Derive new attributes that will be helpful.
↳ Integrate data → Create new dataset by combining data from multiple source.
↳ Reformat data (if necessary).

d) Modelling phase :-

In this phase build and assess various models based on different modelling techniques.

1. Select modelling techniques
2. Generate test design → For modelling approach split the data into training & testing.
3. Build model → Execute few lines of code
4. Assess model → Multiple models are competing against & each other to interpret the model results, based on domain knowledge.

e) Evaluation phase :-

It focus on technical model assessment. It is having 3 tasks :

Teacher's Signature

↳ Evaluate result
↳ Review process (check if properly executed)
↳ Determine next phase.

f) Deployment phase :-

A model is not particularly useful unless a customer can access its results.
Tasks include :

1. Plan deployment
2. Plan monitoring & maintainence
3. Produce final report
4. Review project

Fallacies of DM (Misconception)

1) Data mining tools are automated tools that can be deployed on data repositories to find answers to our problems.

Reality : There are no automatic data mining tools which will automatically solve your problems. There are methodologies available like CRISP-DM which streamlines data mining process into overall business plan of action.

2) Data mining process is autonomous requiring no human intervention.

Reality : Without skilled human intervention (experts), blind use of data mining software will only provide with a wrong answer to

Teacher's Signature

wrong question. So, human intervention is required, to update the model.

3) Data mining pays for itself quickly.

Reality : Benefits take time. Cost include startup cost, data collection, data warehouse preparation cost, software & skilled experts.

4) Data mining software packages are intuitive and easy to use

Reality : Needs to understand the statistical & mathematical assumption behind the model.

5) Data mining will identify the causes or problems.

Reality : KDD process will help to show hidden patterns but human judgement & domain expertise needed.

6) Data mining will automatically clean up messy database.

Reality : Still needs to handle missing values, outliers & inconsistency data manually.

7) DM always provides positive result.

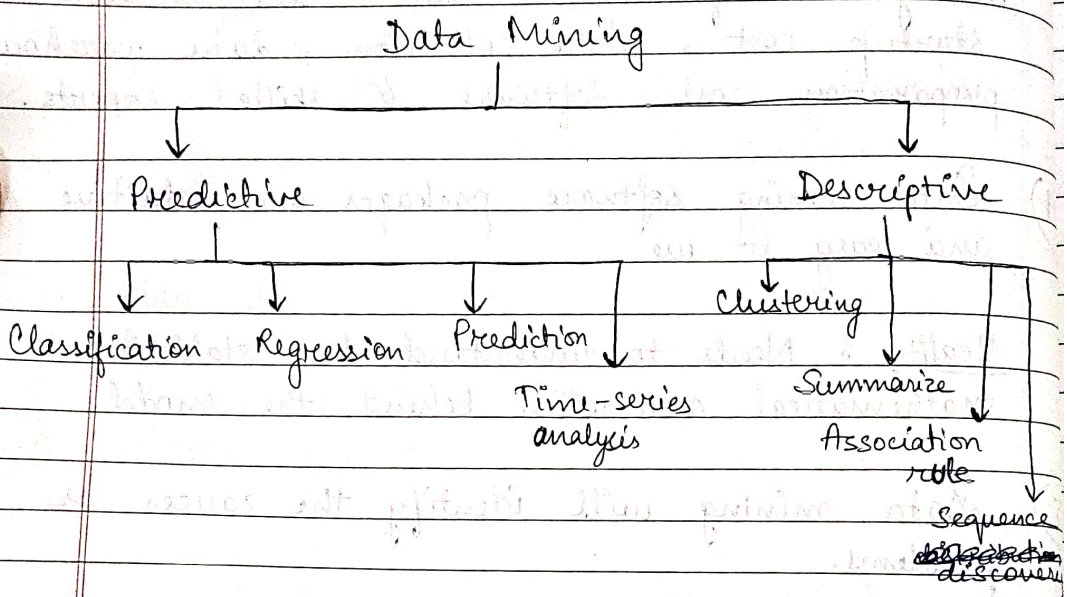
Reality : Not guaranted. Poor data assumption can lead to misleading outcomes.

Teacher's Signature _____

Date: 12/9/25

Task in Data Mining accomplish

Data mining task involve, finding patterns & useful information from large dataset.



1) Classification :-

- It is a supervised learning technique in data mining that involve categorization or classification of data object into pre-defined classes based on their features or attributes.

- In supervised learning, labelled data can be used to build a model that can predict the class of new unseen data. It is of 2 types — binary classification & multi-class classification.

- Steps to build classification :

a) Data preparation
b) Feature selection
c) Split for train or test

Teacher's Signature _____

GENIUS
Date: 12/9/25

d) Model selection
e) Model training
f) Model evaluation
g) Model tuning
h) Essemble learning (optional)
i) Model deployment

2) Regression :-

- Regression is a supervised learning technique used to predict a continuous numerical value by analyzing the relationship b/w dependent variable & one or more independent variable using past data.

- Types of regression :
a) Linear regression
b) Logistic
c) Polynomial
d) Lasso
e) Ridge

Q) We are having the dataset, hours of study (x) to obtain exam score (y).

x	y
1	2
2	4
3	5

→ The regression line is represented as $\hat{y} = a + bx$

Step 1 : Calculate mean of x & mean of y.

$\bar{x} = 2$; $\bar{y} = 3.67$

Step 2 : Calculate b then a

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

Teacher's Signature _____

$\bar{x} = 2$; $\bar{y} = 3.67$

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
1	2	-1	-1.67	1.67	1
2	4	0	0.33	0	0
3	5	1	1.33	1.33	1
				$\Sigma = 3$	

$\therefore b = \frac{3}{2} = 1.5$

$a = \bar{y} - b\bar{x} = 3.67 - (1.5)(2)$
$= 3.67 - 3 = 0.67 \Rightarrow a = 0.67$

$\therefore \hat{y} = a + bx = 0.67 + 1.5x$

Let $x = 4$, (when a student learns for 4 hrs)

$\therefore \hat{y} = 0.67 + (1.5)(4) = 6.67$

- Regression is also known as estimation

3) Prediction :-

Prediction is similar to classification & estimation (regression), except that for prediction the result lie in future.

E.g : Estimating house price based on location, size & features.

4) Time-series analysis :-

It is a way of analysing a sequence of data points collected over an interval of time. It is composed of 4 main elements :

a) Trend → A long term movement of data.

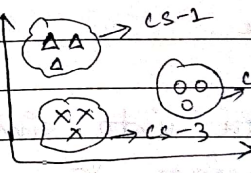
b) Seasonality → A predictable recurring patterns of fluctuation repeats over a fixed, regular interval such as date, week, month or year

c) Cycle → Long term

d) Irregularity → Ran... fluctuation in the ... by trend component.

5) Clustering :-

- Clustering refers ... observations in
- Clustering differs ... there is no t...
- Clustering is u...
- If the data ... both then bi... distance is m...



6) Association ...

- Association t... finding, w...

E.g - Mar... finding a... are purchase... purchased ...

- This r... set occurs

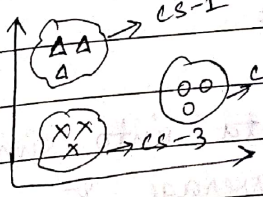
Teacher's Signature

c) Cycle → Long term movement or fluctuations.

d) Irregularity → Random or unpredictable fluctuation in the data that are not explained by trend, seasonal or cyclic component.

5) Clustering :-

- Clustering refers to grouping of records, observations into a similar objects.
- Clustering differs from classification, in that there is no target variables for clustering.
- Clustering is used in un-supervised learning.
- If the data points having less distance, both then belongs to same cluster whereas distance is more belongs to different cluster.



6) Association :-

- Association task for data mining is the job of finding, which attributes go together.
  E.g - Market basket analysis, where finding out which items in a super market are purchased together & which items are never purchased together.
- This rule shows how frequently an item set occurs in a transaction.

Teacher's Signature

Date: 13/9/25

7) Sequence Discovery :-

It is a data mining technique used to identify frequently occuring order patterns in a sequence.

8) Summarization :-

Summarization is the process reducing large dataset into shorter, more understandable format that highlights patterns, trends & relations.

Types :

a) Descriptive → Uses statistical measures to describe the features or numerical data.
E.g - Mean, median, mode.

b) Aggregation → combining data into simpler summaries like sum, average & group count.

c) Sampling → Select a representive subset of the data.

d) Text-summarization → Uses NLP to reduce large text documents into shorter.

BENIUS

Date: 13/9/25

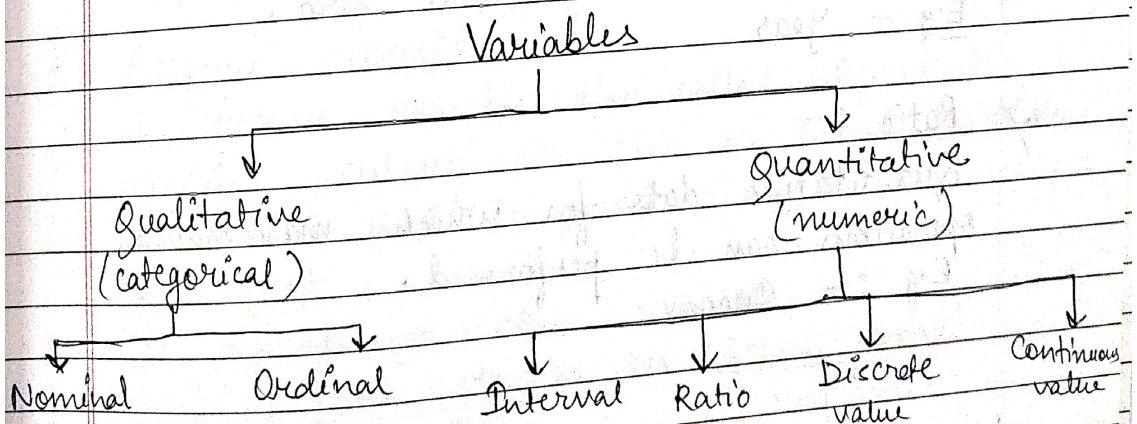
Appendix

Data Summarization & Visualization.

Dataset → Is a collection of related data usually represented in tabular form.

Applicant	Marital status	Mortgage	Income	Rank	Year	Risk
1	single	Y	38k	2	2009	Good
2	married	Y	32k.	7	2010	Good
3	other	N	25k	9	2011	Good
4	other	N	36k	3	2009	Good
5	other	Y	33k	4	2010	Good
6	other	N	24k	10	2008	Bad
7	Married	Y	25k	8	2010	Good
8	Married	Y	48k.	1	2007	Good
9	Married	Y	32k	6	2009	Bad
10	Married	Y	32k	5	2010	Good

- **Rows** → records / samples / observations / objects / subjects / instance
- **Columns** → attributes / features / variables / dimensions / field
- **Cell.** → values which contains actual data



Teacher's Signature

1) Qualitative :-
It describes qualities or categories of data. It is also called categorical values.
E.g - marital status, mortgage, risk, rank

2) Quantitative :-
It takes numerical values & allow arithmetic operation. It is also called numerical variables.
E.g - Income, year

3) Nominal :-
- It is used for names, labels or categories.
- It categorizes without order.

E.g - Marital status, mortgage, risk

4) Ordinal :-
This is having a particular order with naming.
Arithmetic operations cannot be done.
E.g - rank, income, grade (like O, A, B, ...)

5) Interval :- ~~& Ratio~~
It is quantitative data defined on an interval without a natural zero.
E.g - year

6) Ratio :-
Quantitative data for which mathematical operations can be performed.
E.g :- Income.

BENIUS

Date: 16/9/25

7) Discrete variable :-

- It is quantitative or numerical variable that can take finite or countable number including 0.
- It is obtained by counting.
- It cannot take values in between 2 no.s
- Eg - 2.5 X, 1.7 X (not allowed)
- Gaps exist in between possible values.
- Graphical representation → Barchart

8) Continuous :-

- It is a quantitative variable that can take any value within a given range.
- Get from measurable.
- No gaps between the points.
- Graphical representation → Histogram

(*) Predictor variable :-

- A predictor variable is a variable, whose value is used to help predict the value of response variable.
- The predicted variable in table are all variables except risk.

(*) Response variable :-

- A response variable also called dependent variable, output variable or target variable which is used to predict, explain or measure the effect.
- It depends on other variable.

Eg i) Effect of hour studies on exam score.
(predictor → response)

ii) Effect of medicine dose on blood pressure
(predictor → response)

Teacher's Signature

Date: 5/9/2

Basic statistical description of data :-

Central tendency - It is used to measure the location of the middle or center centre of data distribution.

E.g : Mean
Median
Mode

1) **Mean** → Arithmetic average.

E.g : dataset = 2, 4, 6, 8, 10

$$\text{mean} = \frac{2+4+6+8+10}{5} = 6$$

Advantage : ◦ Easy to compute
◦ Include all the data pt-s

Disadvantage : ◦ Sensitive to outliers.

2) **Median** → The middle value when data is arranged in ascending or descending order.

E.g : $d = 3, 5, 7, 9, 11 \rightarrow 7$

$d = 3, 5, 7, 9 \rightarrow \frac{5+7}{2} = 6$

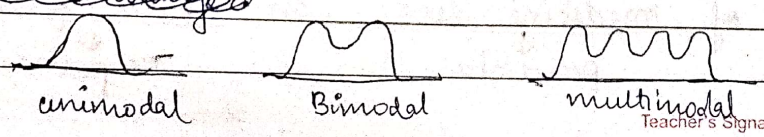
Advantage : Not affected by outlier

Disadvantage : Ignores the magnitude values

3) **Mode** → Most frequently appearing value on the dataset.

~~Advantage~~ ⇒ $d = \{2, 3, 9, 9, 8, 8, 8, 8, 10\}$

~~Disadvantage~~ mode = 8.



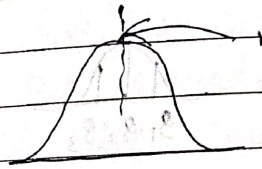
Teacher's Signature

Mid range :-
It is the average of largest & smallest values in the set. It is applicable for numeric data.

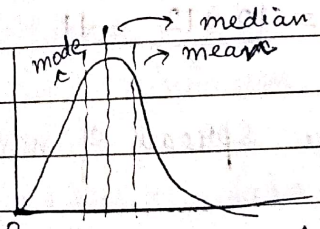
	Measure	Best used	Sensitive to Outliers	Data type
1.	Mean	Numeric & symmetric data	Yes	Interval/ Ratio
2.	Median	Skewed data or with outlier	No	Ordinal, interval/ ratio
3.	Mode	Most frequently accessed items	No	Nominal, ordinal, interval/ratio

(*) Measuring the discortion of data :-

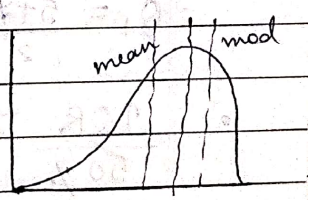
- In a unimodal frequency curve with perfect symmetric data distribution, the mean, median, mode are all at the same central value



Symmetric

- 

positive skewed (if mode < median).



negative skewed (mode > median).

In real application, data are not symmetric.

Teacher's Signature

Date: 17/9/25

(*) Measures of variability :-

→ Range
→ Quartile (Q_1, Q_2, Q_3)
→ Interquartile range (IQR) [$Q_3 - Q_1$]
→ Five number summary (min, Q_1, Q_2, Q_3, max)
→ Box plot
→ Z-score
→ Variance
→ Standard deviation

- **Range** → Range is the difference between maximum & minimum value in the dataset.

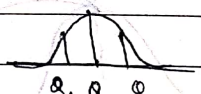
  E.g : 5, 7, 9, 10, 12

  Range = 12 - 5 = 7

- **Quartile** → Q_1 is known as 1st quartile which represent 25%ile of data below it.
  Q_2 is known as 2nd quartile or median which show 50%ile of data below it.
  Q_3 ; third quartile which shows 75%ile of data below it.

  E.g : 5, 7, 9, 10, 12

  ↓ Q_2

  $Q_1 = \frac{5+7}{2} = 6$; $Q_3 = \frac{10+12}{2} = 11$

- **IQR** → IQR measures the spread of middle 50% of the data.

  IQR = $Q_3 - Q_1$

Common rule for identifying suspected outliers is to single out values falling atleast 1.5 * IQR (capping) above the 3rd quartile or below the 1st quartile.

Teacher's Signature

- Five number summary
  statistical summary
  minimum, Q_1, Q_2, Q_3

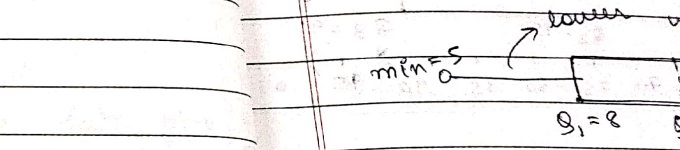
  E.g : 5, 7, 8, 12,

  min = 5, max = 2

  Q_2 = [illegible]

  Q_1 = 8 ; Q_3

- Box - plot (whisker

  lower outlier = Q_1 -
  upper outlier = Q_3

  Box plot is a
  five number
  Box extends
  A line inside
  * Whisker - lin
  the minimu
  outliers.
  * Outliers - d

  Since all v
  no outliers

- Five number summary → It is a concise statistical summary of data consisting of minimum, Q_1, Q_2, Q_3 & maximum.

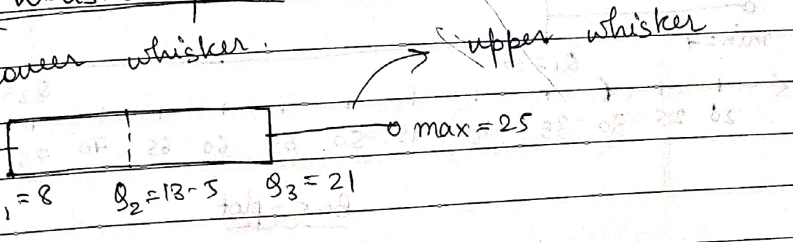
Q) E.g. : 5, 7, 8, 12, 13, 14, 18, 21, 23, 25

min = 5, max = 25

$Q_2 = \frac{13+14}{2} = 13.5$

$Q_1 = 8$; $Q_3 = 21$

- Box-plot (whisker plot) :-



lower outliers = $Q_1 - 1.5 \times IQR$ (= -11.5)

upper outliers = $Q_3 + 1.5 \times IQR$ (= 40.5)

Box plot is a graphical respresentation of five number summary of a dataset.
Box extends from Q_1 to Q_3 known as IQR.
A line inside the box shows median Q_2.

* Whisker - line extending from the box to the minimum & maximum value excluding outliers.
* Outliers - data points that lie beyond.

Since all values lies between -11.5 & 40.5, so no outliers (derived from above example)

Teacher's Signature

Q) Collect math mark out of 100 for 10 student from a class and the records are:

11, 7, 35, 55, 64, 90, 86, 88, 95 & 97

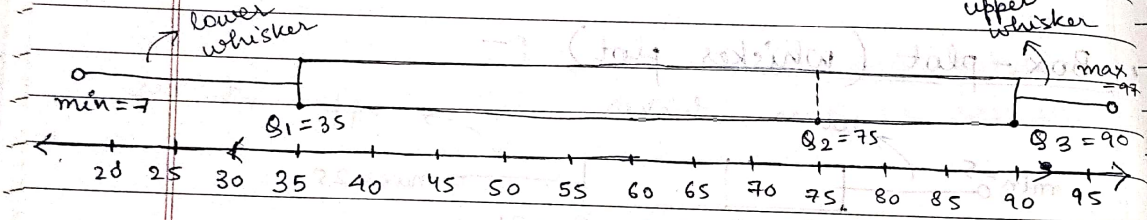
Order:

7, 11, 35, 55, 64, 86, 88, 90, 95, 97

min = 7 ; max = 97

median = $\frac{64+86}{2} = \frac{150}{2} = 75$

$Q_2 = 75$; $Q_1 = 35$; $Q_3 = 90$



Box-plot

Lower outlier = $Q_1 - 1.5 \times IQR$

$= 35 - (1.5 \times 55) = 35 - 82.5$

$= -47.5$

Upper outlier = $Q_3 + 1.5 \times IQR = 90 + 82.5$

$= 172.5$

∴ No outliers.

- Z-score → Z-score measures how many standard deviations, a data point is from the mean.

$$Z = \frac{x - \mu}{\sigma}$$

$x \rightarrow$ observed value

$\mu \rightarrow$ mean of dataset

$\sigma \rightarrow$ standard deviation

$Z = 0$, value is exactly at the mean

$Z > 0$, value is above the mean

$Z < 0$, value is below the mean

At higher absolute from the mean. Outliers — data often considered

Q) In an exam Standard dev 85. Calcul

$Z = \frac{85 - 7}{10}$

$\therefore Z > 0$

- Variance →

It measure

Q) {2, 4, 6}

$n = 3$

$\sum_{i=1}^{3} (x_i -$

$\therefore \sigma^2 =$

- Standard is the

$\sigma =$

Date: 19/9/25

At higher absolute z-score, indicate a point further from the mean.

Outliers — data points with $|z| \geq 3$ are often considered outliers.

Q) In an exam, avg score is 70, the standard deviation is 10 & student score is 85. Calculate z-score

$$Z = \frac{85 - 70}{10} = \frac{15}{10} = 1.5$$

$\therefore Z > 0$, so above the mean.

- Variance → $$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2$$

It measures the average square deviation.

Q) $\{2, 4, 6\}$

$n = 3$, $\bar{x} = \frac{2+4+6}{3} = 4$.

x_i	$x_i - \bar{x}$
2	-2
4	0
6	2

$$\sum_{i=1}^{3} (x_i - \bar{x})^2 = 4 + 4 = 8.$$

$\therefore \sigma^2 = 8/3 = 2.66$

- Standard deviation → Standard deviation is the square root of variance.

$$\sigma = \sqrt{\sigma^2}$$

Teacher's Signature