Roll No: 41258

Unit test

DMW

2) Three tier data warehousing architecture

The following the three tiers of the data warehouse architecture

A) Bottom tier :- the bottom tier is the database warehouse server. It is the relational database system. The backend tools & utilities are used feed data into the bottom tier. They preform extract, clean, load and refresh functions

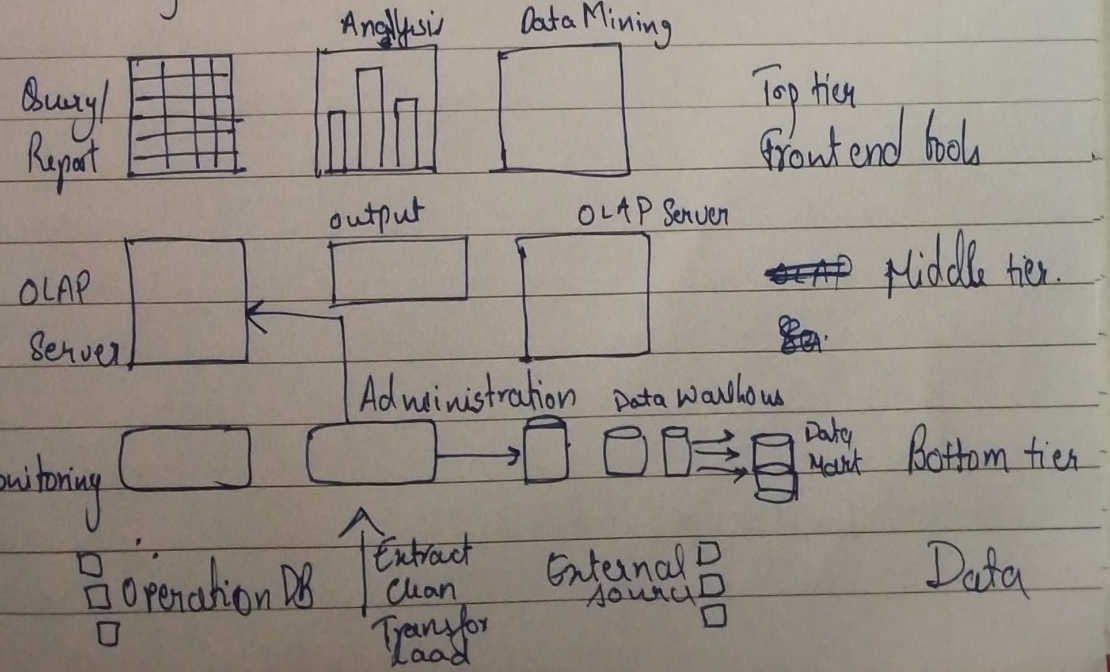
B) Middle tier :-

In the middle tier, we have the OLAP server that can be implemented in following ways:-

i) By relational OLAP (ROLAP) which is an extended relational database management system.

ii) By multidimensional OLAP (MOLAP) model, which directly implements the multidimensional data & operations.

C) Top-Tier :- This is the front-end client layer. This layer holds the query tools & reporting tools, analysis tools & data mining tools.



DATE : / /

Categories of measures used in multidimensional data:-

1) Distributive.
2) Algebraic
3) Holistic

~~1) Distributive~~

Rainbow
PAGE:
DATE: / /

1) Data

13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70

i) Mean of data - 29.96
ii) Median of data - 25 (14th term)
iii) Mode of Data - 25 & 35
iv) Data modality - Bimodal
v) five point summary →
minimum - 13, Quartile 1 - 20.5
Median 25, Quartile 3(Q3) - 35, Maximum 70

Data reduction techniques:-

1) Data Cube aggregation :-
i. Aggregate data in similar form.

2) Dimension reduction:-
i Reduce the size as it eliminates outdated or redundant features.
a) Step Wise forward selection :-
selection starts with an empty set of attributes and later on we decide best of the original attribute on the set Base on their relevance to other attributes.
b) Step wise Backward selection:-
The selection begins with complete attributes in the original data and at each point, it eliminates the worst remaining attribute in the set.
c) Combination of forwarding & Backward selection:-

3) Data Compression:-
i. reduces the file size of the files using different encoding mechanism.

a) Lossless Compression
b) Lossy Compression

4) Numerosity Reduction:-
5) Discretization & Concept Hierarchy operation
   a) Top-down discretization
   b) Bottop up discretization.

---

3) Types of attributes:-

Attribute is a objects property or characteristics.

Types:-

1) Nominal Attribute:-
   Only provide enough attribute to differentiate between two objects eg Roll No.
2) Ordinal Attribute:-
   Sufficient information to order the objects eg Height, Grade.
3) Binary Attribute:-
   0 is absence of any feature & 1 is the inclusion of the characteristics.
4) Numeric Attribute:-
   it is quantitative, measured and represented in integers or real value.
5) Ratio Scale Attribute:-
   difference and ratios are significant flue eg age, length.

Cosine similarity of the document using frequency vector

D1: The sun in the sky is bright.
D2: We can see the shinning sun, the bright sun.