

A row of five decorative letters, each with a thick black outline and a distressed, textured appearance. The letters are stacked slightly, creating a sense of depth. From left to right, they are: I, N, D, E, and X.

NAME: Deekshita Mishra Roll no.: 10 SUB.: Business
Intelligence
Teacher

Assignment - 1

Data warehousing and data mining are the foundation of business intelligence.

Syllabus

- Introduction
- Data Warehousing
- Business Analytics and Data Visualisation

Under introduction :

BI impacts :

- ↳ improving operational performance, quality, speed, dependability, quantity, flexibility, cost
- ↳ improving customer service; Feedback from customers to find loopholes and make amendments accordingly
- ↳ Identifying new opportunities; track the product

Capabilities of BI :

- ↳ Gathering of data (from various data sources)
- ↳ store
- ↳ analyze
- ↳ present the result in the form of results and graphs

Theories behind BI :

- ↳ OLAP and OLTP (databases)
(data cube)

Characteristics :

- ↳ easy way of handling the system

- ↳ should monitor the indicators (profit, loss)
 - ↳ reporting facilities (graphs or any visual representation of result or the analysis)
 - ↳ consistency, quality
- Implementation:
- ↳ Business Analysts (Power users)
 - ↳ Managers, Executives
 - ↳ IT team (Engineers)
 - ↳ Processes involved in making strategies and planning computerized.

'Under business analytics and Data Visualization'.

- ↳ Automated Decision Support (ADS): tools that covers business rules + predictive model for the use of business to take decision on repetitive problems.
- ↳ Web Intelligence: tools used for visual interaction report

- Data, Text and Web Mining
- Technologies Enabling Business Intelligence
- Management and future of Business Intelligence

What is Business Intelligence?

BI is an umbrella-term which comprises of strategies, processes, data application, techniques, technology, architecture which are used by an enterprise to gather data, understand data, perform processing, obtain results and present analysis of the result.

The key output of a BI is decision information, knowledge also)

Input is data (internal or external)
↳ Eg. business ranking of organization

BI must be able to handle structured as well as unstructured data.

Results presented via scorecards, dashboards etc

BI solution supports BI product
↳ makes use of BI tool.

BI tools are provided by BI vendors and BI solution is deployed in an organization.

Eg. A grocery chain collects detailed data of customers on purchase by using customer reward card and uses "BI sol" to categorize customers and to develop knowledge on how offers should be customized.

Real-time BI: real-time input to decision makers, whenever needed so that business processes are not shown.

Operational BI: focuses on operations such as quantity cost profit loss rather than planning or generating insights

* Factors driving BI

- Large amount of data
- Complicated decision
- Quick decisions
- Technological process

(1) Exploring Data Volume

- Organizations have accumulated large data volume about aspects such as sales, customers, complaints etc.
- BI solutions provide manager the ability to effectively utilize this larger data volume

- Increasingly complicated decision.

With increase in competition across countries. Availability of too much

information, unstructured data from various sources such as website, search, media etc.

Decision making has become complicated which has been made easy by BI solution.

→ Need for Quick reflexes.

Market and environmental influences may affect company capitalisation for which managers need to take decisions before window of opportunity closes.

BI solution helps to overcome delays in
→ converting data from various sources
→ integrating information across sources
→ making results available to decision makers.

→ Technological process is made possible by development in decision support system, resource planning system, data mining and text mining.

Business Intelligence and related technologies

* Knowledge Management (KM)

It is about doing what is needed to get the most out of the knowledge resource. It focuses on creating, sharing and applying knowledge.

Knowledge that can be imparted easily or Verbally → implicit

Knowledge that cannot be imparted easily → explicit (tacit)

KM system examples:

- Database of information that are retrieved from people who had already tackled a problem and found a solution.
- Contact details of a person who has expertise in a domain
- Feedback from customers and employees.

NOTE: BI focuses on internal and external data whereas data warehousing and data mining focuses internally primarily on internal data and structured data.

Table 1.1 in the book.

* Business Intelligence in Contemporary Organization :-

- (i) Contributions of BI :
- spreading of real time information in a user-friendly format by providing users a single point of access to important information.
 - creation of new knowledge based on past. Eg. Assessing at new insights based on customer behaviour pattern.
 - responsive and anticipative decision making decisions that have to be taken to compete with current market trends.
Eg. what kind of products an organization should focus on.
 - improving planning for the future.
Eg. recruitment of manpower or investment in R&D.

* Obstacles to BI

- Huge initial and ongoing investment on BI.
- Business events not being consistently defined throughout the organization.

Technological obstacles

- (i) BI tools are difficult to use.
- (ii) BI training is expensive.

Organiz
(i) Jit Profit
Technology
(iii)

Organizational drawbacks :-

- (i) (iii) Profits that are received is difficult to logically justify.
- (ii) (iii) Difficult to customize as per user needs.

Organizational.

- (ii) Finding single solution for all BI needs for an organization is difficult.

Chapter 2

Business Intelligence Capabilities

- (a) Organisational memory
- (b) Information Integration
- (c) Insight Creation
- (d) Presentation Capabilities

(a) It is a storage of information and knowledge in such a way that it can be accessed later.

(b) Ability to link structured and unstructured data from variety of sources

(c) Ability to develop new insight and use them in short term & long term to make better decisions.

(d) Ability to use appropriate reporting and balanced scorecard tools and thereby make BI more valuable to users.

* Features of BI :-

- (a) Benchmarking relative to competition and industry needs

(b) Intelligence in terms of ability to search and utilise data across various sources.

(c) Convenience in terms of customization and connectivity for presentation. *

* Factors leading to organizational memory capability :-

(a) Technological progress and cheap data storage have increased data volume in an organization and these data have to be stored for a longer period of time.

(b) Availability and use of unstructured data and explicit knowledge. #

* Organizational memory ^(OM) capability.

→ focuses on structured and unstructured data.

Structured data or information can be stored in database records.

Unstructured explicit knowledge can be stored in knowledge repositories.

→ focuses on intellectual resources of an organization.

→ data consistency and accuracy are the two important necessities of OM.

* Technologies enabling OM capabilities.

→ Any transactional systems (database systems)

→ Enterprise resource planning systems (ERP systems)

→ ERP systems - software which help business collect information from all core areas of an organization for decision-making.

→ technologies that support the capture of unstructured information.

Factors necessitating structured data and info may exist in a number of transactional databases that are mutually disconnected and often incompatible with each other. There are considerable amounts of data/information on websites, media etc.

* Capabilities of Information Integration

→ OM focuses on past data. BI focuses on latest data, information & knowledge.

→ Context from various sources needs to be synthesised (sort, cluster, classify etc) before it can be used for producing insights.

→ Therefore structured and unstructured data from a variety of sources with real-time data is information integration.

→ It is important to integrate the following :

- (1) Structured information from ERP
- (2) External information
- (3) External written sources & knowledge from text mining.

→ Factors leading to insight creation

- Need for quick reflexes
- Increase domain complexity

Insight Creation Capability :

If we have the description of what had happened, we can find new trends or patterns in prior to events and actions.

If we have an understanding of what had happened, models could

be developed. A prediction of future behaviour, which could depend on a model, that is developed based on understanding of the past, tested and refined using additional information over time.

* Technologies enabling insight creation

→ Data Mining / Knowledge Discovery

→ Business Analytics

→ Realtime Decision Support (use of models of DM or DA to support operational decisions in a real-time fashion)

Eg. Airlines use real-time decision support along with presentational capability to make decisions aimed at flight arrivals & departure on schedule.

* Factors necessitating capability

(1) Content : level of aggregating information

(2) Format : Use of tables, charts & videos.

The key factors that affect content & format are :

(1) Role of intended user.

(2) Task of intended user

(3) Preference

The only capability of presentational model should be user friendly based on role and task.

Technologies enabling presentation

(1) OLAP ~~the~~ Cubes : OLAP is a multi-dimensional data model which gives different views of data.

Three different OLAP architectures are - MOLAP, ROLAP and HOLAP.

(2) Digital Dashboards : It is a GUI that organises and presents information in a way that is easy to read and provides at a glance inside to what is actually happening in an organisation.

Dashboards enable top level executives to gain a view of business and identify strengths and weaknesses in real-time freeing them from traditional bottleneck of manual reporting.

(3) Scorecard : Strategic planning and management tool used by organisations to align their business activities to their organisation's vision and strategy.

They measure periodic results against a pre-defined goal.

Chapter 3

Major Theories and Characteristics of Business Intelligence

Changing Business Environment and Computerized Business Support.

The present business environment and need for computerised support for managerial decision-making leads to the model, the business pressure-response-support:

Remember fig 1.1

* Business Strategy is long term business planning.

* Globalisation → The process by which business or other organisation develops international influence or start operate on international scale

* Decision Support System

* Decision Making Process

- (1) Defining the Problem
- (2) Constructing a Model that describes the real-world problem.
- (3) Identifying possible solutions to the modeled problem and evaluating the solution.
- (4) Comparing, choosing and recommending a potential solution to the problem.

* Need for Computerised Decision Support for System.

- (1) Improved Communication and Collaboration.
- (2) Increased Productivity of Group Members.
- (3) Improved Data Management.
- (4) Managing Giant Data Warehouses becomes easy.
- (5) Overcoming cognitive limits in processing and storing information.

* Decision Support Matrix / Framework.

Remember Fig 1.2

* Types of Decision Structures

- (1) Structured Decision : These are taken on processes ^{is} that are routine and

typically repetitive for which standard solution exists.

(2) Unstructured Decisions are those that are taken on fuzzy processes and complex problems for which there is no clear-cut solution.

(3) Semi-structured : Decisions taken on processes that have structured and unstructured problems.

* Types of Control

(1) Strategic Planning : It is a process to develop and deliver and support strategies, plans and tactics.

(2) Managerial Control : The means by which an organisation resources are directed, monitored and measured (physical resources and intellectual properties)

(3) Operational Control : It is effective and efficient execution of specific tasks.

Phases of Decision Making.

(1) Intelligence Gathering : Searching for conditions in environment that calls for decision.

This step includes finding a problem, defining the problem, gathering data to solve the problems, constraints identification w.r.t the problem.

(2) Design phase : also known as solution discovery is about, finding out possible solutions.

(3) Choice : includes selection of best proposed solution

(4) Implementation : solving the real problem or finding out if the solution is working or can a better solution be proposed.

Conventional Management Information Systems (MIS) and Management Science (MS) tools are insufficient for semi-structured and unstructured problems, hence decision support systems (DSS) are used.

MS and ADSS is used for structured decisions.

* MIS : These tools are computerized reporting tools for manufacturing, marketing, transportation, human resource and accounting organized & programmed in such a way that it produces reports on operations for every level of management in a company.

Features of MIS :

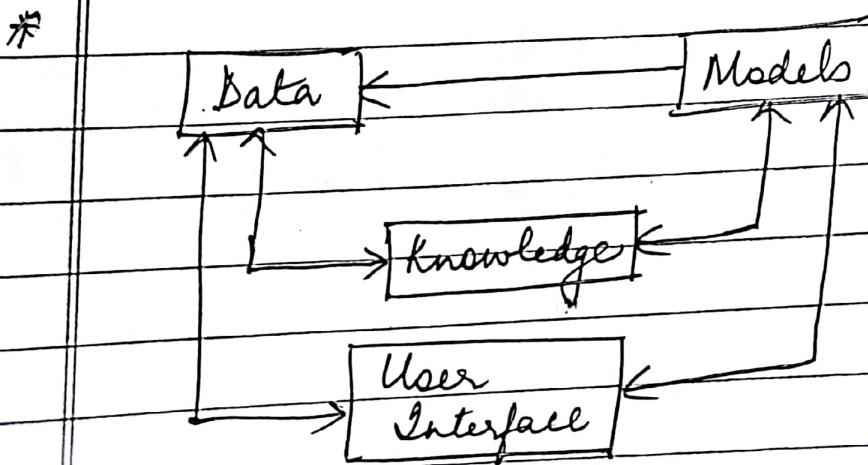
- (1) Produces reports on performances such as weekly, monthly and annual sales.
- (2) Reports can be viewed from web browsers and can be securely shared with others.
- (3) Reports can be exported in any format.
- (4) Suitable for any kind of business.

* Management Science

These tools are based on mathematical modelling or algebraic expressions that describe the problem.

- * Automated Decision Making : It is a rule-based system that provides a solution in one functional area to a specific repetitive managerial problem
- * Computer Support for Unstructured Decision : Decision for unstructured problem can be partially supported by standard computerized quantitative methods. It is usually necessary to develop customized solutions.
- * Computer Support for Semi-Structured Problems : This involves a combination of standard solution procedures and human judgement.

Remember fig 1.3
 (for phases of decision making)



High-level architecture of a Decision Support System

DSS is framework for the process of supporting managerial decisions - making by modeling and employing the models for solution analysis.

EXPLANATION OF FIG. 1.005 (DSS)

To solve any problem or to achieve a particular strategy laid out, data is important.

Models being the second component uses the data for manipulation. Knowledge denotes any intelligent system that is required to provide decision for unstructured or semi-structured problem.

User can interact with the DSS using user interface.

(Remember fig 1.6) → Evolution of BI.

* High-level architecture of BI (fig 1.7)

Data Warehouse → includes historical data along with real-time data, organized and summarized so that end users can easily view or manipulate data.

Business analytics tools :-

→ Reports and Queries : includes both

static and dynamic reporting, querying, multidimensional view etc.

(1)

→ Tools for advanced analytics includes statistical, financial and mathematical models.

→ Tools for data mining, web mining and text mining.

→ Business process management, also known as corporate performance management are the set of application and methodologies that contains used for monitoring and measuring performance.

(2)

(3)

User Interface: Dashboard, scorecards, digital scorecards.

(4)

Intelligent systems are customized purely on the business needs.

8/8/2017

Successful Business Intelligence

#

BI Maturity Level: Maturity level of an organization refers to ability of a particular organization for continuous improvement in a particular discipline.

Four broad categories of maturity level:

- (1) Descriptive → Reports and dashboard used to describe the measurements of any aspect of business which enables users to report on goals and actual values of any methods, metrics and networks.
- (2) Diagnostic → uses more advanced and analytical capabilities such as interactive visualisation to enable users to drill down more easily into data.
- (3) Predictive → discovering new insights or forecasting about future.
- (4) Prescriptive → insights from the predictive models are integrated into the business process to take corrective actions.

Successful Business Implementation

* Roadmap for BI at Organisational Level :-

- (1) Understand organisation's goals, strategies, decision-making process and performance metrics to help define

BI priorities:

- (2) Understand organisation's data (quantity, quality, availability)
 - (3) Review the current BI maturity level
 - (4) Review the existing IT infrastructure.
 - (5) Review the current BI tool sets.
 - (6) Work with stakeholders of organisation about goals and metrics.
- * (7) Drive performance management from top-down and bottom-up.
- (1) BI should not only serve to hold the operational staff accountable but, also put the same emphasis on accountability and responsibility on senior executives as well.
- * Use an agile modular approach.
→ focus on breaking down the performance management needs by

organisational priority and go about implementing BI using a highly iterative approach and shorter release cycles.

* ~~use~~ Design an easy to expand BI solution framework

→ The overall design of the BI solution framework should be easily expandable to accommodate features such as predictive and prescriptive analysis.

* ~~use~~ Leverage existing IT infrastructure and BI tools.

* ~~use~~ Focus on eight metrics with each release

→ Implementation of BI should be in such a way that the identified metrics of an organisation should be measured w.r.t the target.

* ~~use~~ Pay greater attention to data quality.

(1) Source of authentication

(2) Cleaning up of data

(3) Data transformation rules must be incorporated.

(4)

Implementing multiple levels of data validation in the loading process.

- * Keep the ~~of~~ presentation of data as simple as possible.
- * Launch early and often thereafter (same as agile)
- * Establish a centralized governance structure
- * Encourage feedback and proactively manage change

9/8/2017

Data Mining Architecture

(1) Enterprise Data Warehousing

* Pros : single environment for data warehousing.

* Cons : complex workload, querying, reporting, real-time access and loads will have to be done on a single data warehouse

Personal Requirement :

Complexity in managing the entire single environment needs more personnel.

Slow to adapt ^{to} technology.

Monolithic architecture does not take advantage of rapid advances in server hardware.

Eg. parallel performance.

Expensive Maintenance : Multiprocessing based data warehouses makes maintenance cost increasingly difficult.

Scaling becomes difficult.

Limited environment agility.

(2) Distributed Independent Mart

* Pros : Highly relevant to a particular department.

Appropriately sized, based on departmental needs.

* Cons : Data silos → implementing cross-organizational view of data silos requires significant time and expense.

Might result in data duplication.

In case of data duplication, consistency of data has to be always taken care of.

Proliferation of technologies : Multiple hardware and software technologies are deployed across departments.

Data management costs across different departments.

Advantages of Hub and Spoke Architecture

- (1) Data mart tables can be generated on a hub without impacting data mart users and then be published to one or many spokes.
- (2) Clear-cut separation of management and user workload.
- (3) Ability to size a database as per the specific needs.
- (4) Provides scalability

remember → fig 5.5a - 5.5c
to be referred from slides.

#

Selection of Data Warehouse Architecture

It depends on the following concerns:

(1) Information interdependence between organizational units.

(2) Top-level management's information needs.

(3) Urgency of need for a data warehouse.

(4) Nature of end users tasks.

(5) Constraints on IT resources.

(6) Strategic view of data warehouse prior to implementation.

(7) Compatibility with existing systems.

Integration, Extraction, Transformation and Loading.

(1)

Enterprise Application Integration

(2)

Enterprise Information Integration

(3)

Extraction, Transformation and Loading.

(1)

Enterprise Application Integration.

→ provides solution to load data from

different source systems into a different warehouse.

→ It also involves integrating applications functionalities and is focused on sharing functionalities across systems, thereby enabling flexibility and reuse.

→ EAI is accomplished using SOA (Service-Oriented Architecture) which is a style of S/w design where service is provided by application's components over a network.

② Enterprise Information Integration

→ It is a tool which performs real-time data integration from a variety of sources.

→ It displays data in a uniform representation on a data-monitoring interface, meaning integrated data appear relational to end users.

→ EII also uses XML which is structured and may be extended at a later time.

③ Extraction Transformation and Loading

- Uses of ETL tools :
- ① performs ETL process
 - ② maintains organizational data warehouse.
 - ③ helps in detecting and removing any anomalies in real-time data.
 - ④ should support documentation of how metadata changes as data move between source and target.

- During ETL process
- ① all input files are written to a staging area which are designed to facilitate the load process!
 - ② defining business rules such as how the data will be used (cardinality constraints)
 - ③ summarization rules containing dimension algorithms, i.e. what dimensions to include, what level within each dimension to include, data granularity
 - ④ standardization of encoded attributes
 - ⑤ consistency across columns or table names
 - ⑥ data quality issues must be dealt.
 - ⑦ business rules must be stored in central metadata repository.

→ Criteria in selecting ETL tools:

→ ability to read and write from large number of data source

→ Automating, capturing and delivery of metadata.

→ a history of conforming to open standard.

→ an easy to use interface for all ~~these are~~ types of users.

Remember fig 5.6

DATA WAREHOUSE DEVELOPMENT

* Direct benefits from a data warehouse

(1) Consolidated view of corporate data

(2) Allows users to perform extensive analysis

(3) Timely information with better representation

(4) Simplified access

* Indirect benefits of data warehouse

(1) Enhances business knowledge

(2) Competitive advantage with other industries

- (3)
- (4)
- (5)

Enhanced customer service and satisfaction
Facilitates decision making
Reforms business process.

* Crucial factors for the development of data warehouse:

- (1)
- (2)
- (3)

Support from the management and users
Time frame and budget of the project
Complexities of business objectives, goals and strategies.

* Guidelines that need to be considered when developing a vendor list:

- (1)
- (2)
- (3)
- (4)
- (5)
- (6)

Financial strength of the organisation
Qualified consultants
ERP linkages
Market share
Industry experience
Established partnerships.

* Data Warehouse Development Approaches:

- (1)

Enterprise data warehouse: Top-down development and follows a spiral approach.

(2)

Data mart approach: Development starts with the data marts of various departments followed by development of a centralized data warehouse at the organizational level if needed in a long term.

Remember Table - 5 - 2.

* Tasks performed during the development of data warehouse.

- (1) Determine the goals, objectives & business processes for an organization
- (2) Design a functional data warehouse dimensional model
- (3) Identify the data sources.
- (4) Identifying ETL best practices for development & maintenance.
- (5) Identify the need of ODS (operational data sources)
- (6) Identify the appropriate technologies for ETL, metadata and data warehouse
- (7) Launch a tiered deployment (DBA support, DB refresh, Cloning, cleansing, security etc)
- (8) Perform backups.

#

Best Practices for Implementing a Data Warehouse

- (1) The project must fit ^{with the} corporate strategy and business objectives.
- (2) Follow incremental development of data warehouse.
- (3) Adaptability to any environment.
- (4) Develop a business supplier as well as IT relationship.
- (5) They should be complete by end of the product.

#

links involved in data warehousing projects

(i) No mission or objectives identified clearly.

(ii) Quality of data unknown

(iii) Skills not in place

(iv) Inadequate budget

(v) Lack of supporting softwares

(vi) Source of data not identified properly

(vii) Too many changing requirements

(viii) Architectural and Design risks

(ix) Platform incompatibility

(x) Key IT people leaving the project

(xi) Geographically distributed environment

#

Issues w.r.t a successful data warehouse

(vi)

(i)

Starting with wrong sponsorship chain.

(vii)

(a) Need for an executive sponsor

(senior member of the project, i.e a chairperson who has influence over the necessary resources)

(viii)

(b) Need of an executive project driver, i.e someone who has earned the respect of other executives.

(ix)

(c) Need of an IT manager to head the project.

(ii)

Loading the warehouse with information just because it is available.

(iii)

Believing that data warehousing design and database design are same.

(iv)

Setting up expectations that you cannot meet.

(V)

Choosing a data warehouse manager who is technology-oriented rather than a user-oriented (it becomes an issue)

- (vi) Data should be made accessible to the right people at the right time in the right format.
- (vii) Believing that the problems are over when data warehouse is up and running.
- (viii) Data warehousing is a continuous process as BI keeps evolving.
- (ix) Focussing on ad hoc, periodic reporting instead of alerts.
→ Alert systems: These are BI tools which send notifications to the user that a certain event or condition has happened. Eg. when daily sales figure falls under a certain percentage of the target
- (x) Delivering data with overlapping or confusing definitions
- # (xi) User participation in data warehouse development
- (i) Data modelling (tasks performed)
→ define business rules
→ decide on aggregations

(ii) Axis modelling (how data are to be retrieved from a data warehouse)

(iii) Indexing (for faster access)

19/8/2017

Massive data warehouses and scalability.

as size of a data warehouse grows when :

→ amount of data in warehouse increases

→ complexity of user queries increases

→ as business expands

→ as number of concurrent users increases.

so test

Scalability of a data warehouse is said to be good when queries and data access function grow linearly with size of the data warehouse and data warehouse grows horizontally and vertically.

Real-time data warehousing

This is also called active data warehousing or zero latency data warehousing.

Need for real-time data warehousing:

- increased data volume and acceleration of update speeds.
- to make fast and consistent decisions
- traditional data warehouses are not business critical as data is updated on a weekly basis.

Definition of a real-time data warehouse

It provides integrated information repository to drive strategic and tactical decision support within an organization.
Eg. flight management dashboard application.

Comparison between traditional and real-time data warehouses.

Traditional	Active
→ strategic decisions → strategic and only.	→ tactical decisions
→ results sometimes hard to measure	→ results measured with operations.

- daily, weekly, monthly → data updation
updation is every minute.
- moderate user concurrency → high number of users accessing and querying the system simultaneously.
- highly restrictive reporting; often uses pre-developed summary tables → flexible ad hoc reporting.
- power users and internal users → operational staffs, call centres, external users.

Critical issues in real-time data warehousing

- (1) Not all data should be updated continuously
- (2) Real-time reports are different when produced at slightly different times.

Data warehousing administration and security issues.

- (1) Knowledge on high-performance software and hardware, network technologies, business insights, decision-making process, data warehouse design and

good communication skill

* Security issues of a data warehouse

- Establishing effective corporate security policies which should be communicated among all levels in an organization.
Eg. non-disclosure policies
- Implementing logical security techniques to restrict access.
Eg. user authentication, encryption.
- Limiting physical access to data center environment
- Establishing an effective internal control review process with an emphasis on security and privacy

22/8/2017

Business Analytics Field and Overview

BA is a broad category of applications and techniques for gathering, storing, analyzing and providing access to data to help enterprise to users make better

business and strategic decisions.

BA is also termed as analytical processing, BI tool, BI application. Eg. an analytic application used for credit scoring for a loan applicant might

- (i) automatically accept or deny the loan application.
- (ii) select a loan limit or
- (iii) to select which other type of loan to approve.

Tools and techniques of BA.

Broad category of BA.

① Information and Knowledge

Discovery (OLAP querying, data mining, text mining, web mining).

② Decision Support and Intelligence

System (ADS, ESS, EIS, BIS, Web Analytics)

③ Visualisation (dashboard, scorecard)

* Microstrategy is classification of BA tools.

Enterprise reporting → These are pixel perfect report format.

Cube analysis → used to provide simple OLAP, multidimensional slice and dice analytical capabilities.

ad hoc queries and analysis → perform on MOLAP to query a database for analysis. Statistical analysis and data mining: used to perform predictive analysis or to discover correlation between two metrics.

Alert system: to notify users whenever an event occurs.

* SAP's classification of BA tools :

- (1) Operational
- (2) Managerial
- (3) Strategical

Executive Information and Support System

EIS :

(i) Computer-based system that provides timely and relevant information to the top executives to monitor organization's performance.

(ii) Provides exception reporting

(iii) Provides drill down operations

(iv) ESE: goes beyond EIS to include analyses support communications, office automation and intelligent support.

Capabilities of EIS and ESS

- drill down
- helps to measure critical success factors
 - Eg. customer satisfaction, quality of food w.r.t restaurant.
- measures key performance indicators:
these are measures which help know whether a business is moving in the direction of its set goals and objectives. Eg. profit, sale, loss.
- helps produce status reports : A report on current affairs or the latest data available on key performance indicator (KPI).
- helps do ad hoc analysis, i.e analysis done at any time.
- helps provide exception reporting
- provides slicing and dicing

#

Online Analytical Processing

These are activities performed by end users in online systems for generating reports, graphs, statistical analysis and query.

*

Difference between OLTP and OLAP.

OLTP

OLAP

- | | |
|---|--|
| → Operational workers | → Top-level executives |
| → Repetitive transactions | → Activity can be ad hoc. |
| → Simple queries | → Complex queries. |
| → Answers user queries | → Pattern analysis, trend analysis etc. |
| → Nature of design is application-oriented. | → Nature of design is subject-oriented. |
| → Answers closed ended questions. | → Answers open ended questions. |
| → Structured query language. | → DMQL |
| → Nature of data is relational | → Nature of data is historical, summarized and multidimensional. |

*

Types of OLAP

MOLAP, ROLAP, Desktop OLAP,
Web OLAP → OLAP data accessible from
web browser.

Desktop OLAP → Multidimensional analysis
is done in a local machine.

* Characteristics of OLAP tools

(1) Multidimensional conceptual view
for formulating queries.

(2) Transparency to the user

- Underlying data repository should
- Source of data (nature) be made
- Computing architecture known to user.

(3) Easy accessibility

- Batch and online access should be made available.

Eg. ETC batch processing.

(4) Consistent reporting performance.

- Performance to run a report should not degrade as size of the database increases.

(5) Client - server architecture

(6) Generic dimensionality
• Treat each dimension equally on the basis of operation capabilities.

(7) Dynamic sparse matrix handling
• OLAP tools must be able to handle missing values in the incoming data in case of real-time data warehouses.

(8) Multi-user support
(9) Unrestricted cross dimensional operation
• Should allow for drill down, roll-up and any other data manipulations across dimensions

(10) Intuitive data manipulating
• Avoid menu or multiple trips to UI
• Enable consolidated path.

(11) Flexible reporting
• Provide capabilities to the business users to arrange columns or rows or any cells in a manner that facilitates easy analysis and generation of information.

(12) Appropriate number of dimensions and aggregation.
• Accommodate 15^{to} 20 dimensions within an analytical model and should have suitable levels.

* 4 types of Analytical Processing
~~Operations~~ Performed by Analysts

- (1) Categorical : Static analysis based on historical data. Eg. disease prediction
- (2) Exegetical : Drill down operation performed on historical data.
Eg. drill down operation performed on revenue of entire business will give revenue per business unit or in a quarter.
- (3) Contemplative : Analysis done by changing a single value to determine its impact.
- (4) Formulaic : Allows for changes to multiple variables.

26/8/2017

Reports and Queries

- (1) Routine Reports : Reports which are automatically and periodically distributed to subscribers. Eg. weekly sales report sent to managers.
- (2) Ad hoc / On-demand reports : Reports generated whenever needed and data in the reports depend upon user demand.

* Multilingual support in which reports generated may be in certain language and can be translated to any other language.

(3) Representative types of report.

(i) Finance / Accounting : Eg. accounts receivable and payable reporting, balance sheet reporting, profit and loss income statement analysis.

(ii) Human Resource : Eg. career path reporting and analysis, employee expense reporting, time and labor management reporting

(iii) Sales forecasting and marketing : Eg. channel analysis, sales pipeline reporting and analysis.

(iv) Customer relationship management : Eg. customer acquisition retention and attrition . Eg. customer loyalty analysis.

(v) Product sales : Eg. discount and promotion analysis, inventory analysis.

(vi) Operations and Supply Chain Management : Eg. call centre analysis, quality control analysis

(vii) Website analysis: Web Navigation Analysis and web traffic analysis.

* Features that a report could have
(additional features)

(1) Report distribution to any touchpoint; report content can be carried by emails, file servers, portals and in any format.

(2) Self-subscription as well as administrative based distribution: An individual personally gets subscribed to some reports to have knowledge on other projects.

(3) Delivery on-demand on-schedule or on-event

(4) Automatic content personalisation: Reports that provide only that information which is requested.

Queries.

- (1) Ad hoc queries : Any query that cannot be determined prior to the moment the query is issued. Eg. a user may decide to place a query after he/she gets a report; on-the-fly SQL query is generated.

Once reports are generated, managers would want to go for detailed investigation of the displayed data in the report. Such an analysis can be facilitated by multi-dimensionality

Multi-dimensional databases support multi-dimensionality in which data are organized ^{specifically} to support easy and quick access.

Limitations of dimensionality

- (1) MDB can significantly take up more computer storage than relational database
- (2) Multidimensional products are costlier.

- (3) Loading MSB takes more time and resources.
- (4) Maintenance space is costlier.

Advanced Business Analytics

- (1) Data Mining
- (2) Predictive Analysis

③ Loading MSB takes more time and resources.

④ Maintenance space is costlier.

Advanced Business Analytics

① Data Mining

② Predictive Analysis

29/8/2017

* Data Visualization : Refers to tools and technologies that support visualisation and interpretation of data and information at several points along data processing chain.

Technologies used for data visualisation are

- any digital image
- geographical information systems
- data analytical tools
- virtual reality

* Spreadsheets for visualisation :

Features of spreadsheets

- Charts
- Sparkline
- Conditional formatting
- Pivot Table (summarize large datasets using 'n' dimensions)
- Power map (helps visualise geographic data on a map)

* Geographic Information System (GIS)

- Site selection to open up new outlet of business
- Use GIS to answer questions such as "How many" and "What if".
- GIS can help find demographic profile of a store's neighbourhood to determine the kind of products that can be kept in the store.
- Helps assess damage due to natural calamity.
- Helps identify clusters of call disconnects and dispatch technicians accordingly.

* Application of GIS in banks :

- Determine branch and ATM locations
- Analysing customer demographics such as age, education, income etc will help them come up with new products of marketing campaigns.

- Helps managers analyse 'what if' questions.
- Determining market potential for bank activities.

* Features of GIS

- Creation of maps based on geographic information.
- Toolkits such as clip tool, buffer tool, continuous and discrete raster.

* GIS functions

(1) Data capture : Data for GIS can come from many sources such as scanning for area photographs, paper maps (undergoes manual digitization, i.e conversion of paper maps into numerical digits that can be stored in computer in the form of points or lines), GPS.

(2) Data compilation

Involves cleaning up of errors and noisy data followed by assigning data to their respective attributes.

- ③ Data storage : Data storage is based on two modes : Vector and Raster.
- ④ Data manipulation : Involves co-ordinate change, projections, edge detection, measuring area, perimeter.
- ⑤ Analysis :

ADS.

30/8/2017

Automated Decision Support Systems
These are rule-based ~~solutions~~ systems that provide solutions to repetitive managerial problems.

ADS is also called as enterprise decision management. ADS systems are closely related to BI and BA in the following ways:

- (a) BA models operate on business tools.
- (b) Business rules can be used to trigger automatic decisions that can be a part of BI application.
- * Business Rules : These are embedded in a rule-driven work-flow engine.

There are four types of business rules:

(a) Analysis rules → These are used to calculate performance metrics from business transaction data.

(b) Context Rules → These enable performance metrics to be tied to business goals and forecast.

(c) Exception Rule → Sending an alert to a business user when a metric exceeds a threshold defined in exception rule.

(d) Actions Rule → Manual decision making.

* Benefits / Characteristics of AIs :-

- (1) Rapidly builds rule-based applications;
- (2) Injecting predictive analysis is also possible.
- (3) Providing AIs to legacy systems is also possible.

* Implementing AOS .

- (1) Rule Engine ~~at all~~
- (2) Mathematical and statistical algorithms
- (3) Industry specific packages
- (4) Workflow Applications : Once a rule engine makes a decision, the workflow system moves the business process through required steps.

* Competitive Intelligence :-

Monitoring the activities of the competitors in the market CI can drive business performance by increasing market knowledge and also improve strategic planning.

Technologies used in ~~as~~ CI are -

→ Optical Character Recognition (OCR):
This helps identify printed characters from business magazines or articles. ~~using~~ so that it can be stored in knowledge management system.

→ Radio Frequency Identification

→ Intelligent Agents are the softwares which helps business monitor competitor's website or even monitor the entire internet for a specific information.

→ Internet : helps business know the feedback about their products or even offer competitive prices for the products.

Another aspect of CI is industrial espionage (IE) who exist in industries look for confidential marketing plans or strategic plans of the competitors. It is considered unethical and illegal

Sessional Partion till here

2/9/2017

Web analytics

Application of business analytics is BI . BI is a method for decision making and web analytics is a solution for BI . through which data can be collected for business growth.

Web Analytics is about gathering data from web and analyzing it !