

Foundations of Datawarehousing and Business Intelligence

Sajin R Nair

Assistant Professor, Department of CSE-Data Science, Malla Reddy College of Engineering, Hyderabad, Telangana, India.

Dr. J. Gladson Maria Britto

Professor & HOD, Department of CSE- Data Science, Malla Reddy College of Engineering, Hyderabad, Telangana, India.

Anju Gopi

Assistant Professor, Department of CSE – Artificial Intelligence & Machine Learning, Malla Reddy College of Engineering, Hyderabad, Telangana, India.

Akash Dey

Assistant Professor, Department of CSE-Data Science, Malla Reddy College of Engineering, Hyderabad, Telangana, India.

Published by



**Centivens Institute of
Innovative Research**
Private Limited

Foundations of Datawarehousing and Business Intelligence

Copyright © 2026 by CIIR

All rights reserved. Approved reprint of the edition published by CIIR. No part of this book may be reproduced in any form without the written permission of the publisher.

Limits of Liability/Disclaimer of Warranty: The authors are solely responsible for the contents of the paper in this volume. The publishers or editors do not take any responsibility for the same in any manner. Errors, if any, are purely unintentional and readers are required to communicate such errors to the editors or publishers to avoid discrepancies in future. No warranty may be created or extended by sales or promotional materials. The advice might not be suitable for right for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Additionally, readers should note that internet websites referenced in this work may have changed or disappeared since publication.

CIIR also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

ISBN 978-81-990577-3-9



Authors

Sajin R Nair

Dr. J. Gladson Maria Britto

Anju Gopi

Akash Dey

Price: ₹499

Centivens Institute of Innovative Research

307, 5th Street Extension, Gandhipuram,

Coimbatore - 641012, Tamil Nadu, India.

E-mail: publications@centivens.com

Website: www.centivens.com

Acknowledgement

First and foremost, I thank, the Almighty God for endowing his immense blessing that helped me in each step of my progress towards the successful completion of this text book “Foundations of Datawarehousing and Business Intelligence”.

I owe a warm debt of gratitude to my co-authors for their guidance, novel ideas, inspiration and brilliant discussions at every stage of this work. From my bottom of heart I thank him for made this as a good research work.

It is my privilege to thank the management and our Principal Dr. Venkata Krishnan for their constant support and cooperation to complete the text book.

Most importantly, none of this would have been possible without the love and patience of our family. Words alone cannot describe the encouragement, affection and our parent, wife and children; I thank them for all the help, care and moral support provided throughout this work.

Preface

In today's digital world, data has become a vital asset for organizations across all sectors. The ability to transform data into meaningful information is essential for effective decision-making and strategic planning. This textbook, *Foundations of Data Warehousing and Business Intelligence*, is designed to provide students with a strong foundation in the concepts, techniques, and applications of data warehousing and business intelligence.

The book presents the subject in a clear, systematic, and learner-friendly manner. It covers the fundamentals of data warehousing, data modeling, ETL processes, OLAP, data mining, and business intelligence tools. Each chapter is organized to build conceptual understanding while encouraging practical application through examples, illustrations, and exercises.

The content has been developed in accordance with current academic curricula and is suitable for undergraduate and postgraduate students, educators, and aspiring data professionals. It aims to develop analytical thinking, problem-solving abilities, and decision-making skills that are essential in today's data-driven environment.

We hope this book will serve as a valuable learning resource for students, a useful teaching guide for instructors, and a practical reference for professionals.

Dedication



This text book is devoted to Almighty God for making it all possible, our family members and our Gurus for their loving, encouraging and everlasting support.

Author Profile



Mr. Sajin R. Nair is an Assistant Professor in the Department of Computer Science and Engineering at Malla Reddy College of Engineering. With three years of teaching experience, he is dedicated to fostering academic excellence and advancing research in emerging technologies. He holds an M. Tech degree in Computer Science and Engineering and is currently pursuing his Ph.D. His research interests focus on Deep learning, Bigdata and Artificial Intelligence. He has actively contributed to the academic community by presenting research papers at conferences and participating in various workshops and Faculty Development Programs (FDPs). Mr. Nair has also completed the Coursera certification in Deep Learning with PyTorch: Generative, build a computer vision app with Azure Cognitive Services, Machine Learning Pipelines with Azure ML Studio and Deep Learning with PyTorch: Image Segmentation demonstrating his commitment to continuous learning and industry-relevant expertise. He is a professional member of AICTSD, IFERP, IAEPT, IAENG and Internet Society. In recognition of his innovative contributions, he has been awarded a design grant patent for an AI-based pest infestation monitoring camera, and UK design patent Scanner for object recognition with deep and machine learning integration showcasing his dedication to applied research and technological innovation.

Author Profile



Dr. J. Gladson Maria Britto is working as a Professor & Head in the Department of CSE-Data Science at Malla Reddy College of Engineering, Hyderabad. He obtained his Bachelor's degree in CSE, Master's in CSE and Doctoral Degree in CSE from various state government universities. He has 18 years of teaching experience and 4 years of industry experience. He published 28 research papers in various international conferences and reputed journals and also published 4 patents, 6 books, 3 book chapters. His interested areas are Artificial Intelligence & Machine Learning, and Data Science applications. He is instrumental in organizing technical symposiums, workshops, Seminars, Guest Lectures, Technical Talks, Motivational Talks, Short Term Training Programs (STTPs), Expert Talks, Technical Quiz and FDPs to improve the student's technical skills in various dimensions. He received AICTE- Vishwakarma Awards 2020, CSI INAPP AWARDS 2021 and various recognition from reputed bodies like CSI. He is optimistic, passionate and enthusiastic individual who enjoys working with positive people who can share his enthusiasm for teaching and learning with Motivation as well as support.

Author Profile



Mrs. Anju Gopi is an Assistant Professor at Malla Reddy College of Engineering with three years of teaching experience in the field of Computer Science and Engineering. She holds an M. Tech degree in Computer Science and Engineering and is currently pursuing her Ph.D. Her academic interests are centered on Artificial Intelligence and Generative AI. She has actively participated in academic activities by presenting research papers at conferences and attending various workshops and faculty development programs. Mrs. Anju Gopi has completed the Infosys Springboard certification in Generative AI, reflecting her commitment to continuous learning and industry-aligned skills. She is a professional member of AICTSD and IAENG. She is also the recipient of a design grant patent for an AI-based pest infestation monitoring camera, highlighting her contribution to applied research and innovation.

Author Profile



Mr. Akash Dey is an Assistant Professor at Malla Reddy College of Engineering with a strong academic foundation in Computer Science and Engineering. He completed his B. Tech in CSE in 2020 and earned his M. Tech in CSE with a specialization in the Internet of Things (IoT) in 2022. With 3 years of teaching experience, Akash is dedicated to providing a blend of theoretical knowledge and practical skills, preparing students for the dynamic tech industry. His areas of interest include IoT and emerging technologies, where he strives to inspire innovation and learning.

TABLE OF CONTENTS

S. NO	TITLE	PAGE NO
1	What is Data?	1
	1.1 Data Warehouse Definition	2
	1.1.1 Subject-Oriented	2
	1.1.2 Benefits of Data Warehouse	3
	1.1.3 Need for Data Warehouse	4
	1.1.4 Database Vs Datawarehouse	5
	1.2 Data Warehouse Modeling	5
	1.2.1 Difference between OLTP and OLAP	8
	1.3 OLAP Operations	8
	1.3.1 OLAP Operations	9
	1.4 Data Cube Computation Methods	15
	1.4.1 Materialization of Cube: Full, Iceberg, Closed and Shell Cubes	15
	1.4.2 Roll-up/Drill-down	16
	1.4.3 Slice-and-Dice	17
	1.4.4 Grouping Sets	17
	1.4.5 Online Analytical Processing (OLAP)	17
	1.4.6 SQL Queries	18
2	Data Ware Housing and Business Intelligence	19
	2.1 Business Intelligence Introduction	19
	2.1.1 BI Methods	20
	2.1.2 The BI Process	21
	2.2 Leveraging Data and Knowledge for BI	21
	2.2.1 Data Sources for Business Strategy	22
	2.3 BI Components	25
	2.3.1 BI involves	25
	2.4 Bi Dimensions	28

S. NO	TITLE	PAGE NO
	2.4.1 Information Space	28
	2.4.2 BI Functional Space	28
	2.5 Business Intelligence and Business Analytics	30
	2.5.1 Operational Systems	30
	2.5.2 Management Information System (MIS)	30
	2.5.3 Decision Support System (DSS)	30
	2.5.4 Business Analytics (BA)	31
	2.5.5 Difference Between BI and BA	31
	2.5.6 Business Analytics (BA)	32
	2.5.7 Descriptive vs Predictive Analysis	32
	2.5.8 Managers vs Analysts	33
	2.5.9 Reporting vs Applying	34
	2.5.10 New Analytics Strategy vs Existing Analytics Strategy	34
	2.5.11 Current Events vs Future Possibilities	34
	2.5.12 Ease of Operations	35
	2.5.13 Tools in BI and BA	36
	2.5.14 Applications	36
	2.6 Information Hierarchy	37
	2.6.1 Data Layer	37
	2.6.2 Information Layer	37
	2.6.3 Knowledge Layer	38
	2.6.4 Wisdom Layer	38
	2.6.5 BI Life Cycle	38
	2.7 Data For Bi	40
	2.7.1 Data Issues and Data Quality for Bi	40
	2.7.2 Types of Data Used in BI	41
	2.7.3 Key Aspects of Data Quality	41
	2.7.4 Why is Data Quality Important?	41

S. NO	TITLE	PAGE NO
	2.7.5 Data Issues For BI	42
3	BI Implementation	44
	3.1 Data Ware Housing and Business Intelligence	44
	3.1.1 Data Ware Housing and Business Intelligence	45
	3.1.2 What Benefits Does Successful BI Implementation Bring in Its Wake?	45
	3.2 Key Drivers Of BI	47
	3.2.1 Single Version of the Truth	47
	3.2.2 Current and Accurate Information	47
	3.2.3 Rapidly Changing Information Needs	47
	3.2.4 Customer Service Excellence	48
	3.2.5 New Service Delivery Channels	48
	3.2.6 Multiple Internal Databases	48
	3.2.7 Purchased Packages	48
	3.2.8 Increasing Complexity of Systems	49
	3.2.9 Application System Evolution	49
	3.2.10 Computer Networks and External Databases	49
	3.3 Key Performance Indicators and Performance Metrics	50
	3.3.1 Financial Metrics	51
	3.3.2 Marketing Metrics	51
	3.3.3 Project Management Metrics	51
	3.3.4 Customer Service Metrics	52
	3.3.5 What is Business Intelligence Architecture?	52
	3.3.6 Business Intelligence Architecture Layers	54
	3.3.7 Business Intelligence Architecture Components	56

	3.3.8 Business Intelligence Architecture Diagram	59
	3.3.9 Business Intelligence Architecture Components and Diagram	59
	3.3.10 BI Architecture Tools	63
	3.4 Business Decision Making	64
	3.4.1 4 Phases of the Decision-Making Process	64
	3.4.2 A Cyclic Process of Intelligence Creation	67
	3.4.3 Styles Of BI	39
	3.4.4 The Value of Business Intelligence:	70
	3.4.5 Effective Decision-Making	71
	3.4.6 Sales & Marketing	71
	3.4.7 Customer Experience	71
	3.4.8 Boosting Productivity	72
	3.4.9 Data Accuracy and Compliance	72
	3.4.10 Tips for Choosing the Right Business Intelligence Tools	72
4	Advanced BI	74
	4.1 Big Data and BI	75
	4.1.1 BI Big Data	78
	4.2 Social Networks	80
	4.2.1 In other words, Big Data is Characterized By	82
	4.2.2 How To Create a Good Strategy	82
	4.2.3 Measuring Results	82
	4.2.4 Social Networks and Big Data: Get the Strategy Right for an Incredible Experience	83
	4.3 Mobile BI	85
	4.3.1 What Is Mobile Business Intelligence (BI)?	86
	4.4 Emerging Trend	88
	4.4.1 DQ and MD Management	89

	4.4.2 Creating a Data-Driven Culture	90
	4.4.3 Data Governance	90
	4.4.4 Data Discovery and Visualization	90
	4.4.5 Real-Time Data & Analytics Gain More Traction	91
	4.4.6 Artificial Intelligence	91
	4.4.7 Data Automation	91
	4.4.8 Eye-Opening Business Intelligence Statistics in 2023	92
	4.4.9 How Can Business Intelligence (BI) Help Businesses during the Crisis	93
	4.4.10 Discover Opportunities for Boosting Growth	93
	4.4.11 Find Areas for Saving Costs	93
	4.4.12 Make Informed Decisions to Minimize Risks	93
	4.4.13 Speed Up Normal Operations	93
	4.5 Description Of Different, BI-Tools (Pentaho, KNIME)	94
	4.5.1 Pentaho BI What is Pentaho BI?	94
	4.5.2 KNIME	96
	4.5.3 SAP Business Objects	97
	4.5.4 Datapine	98
	4.5.5 Microsoft Power BI	98
	4.5.6 Looker	99
	4.5.7 Clear Analytics	100
	4.5.8 Tableau	100
	4.5.9 Oracle BI	101
	4.5.10 Domo	102
5	Business Intelligence and Integration Implementation	103
	5.1 Pitch to Key Players	103
	5.1.1 Choose Tools and Create a Team	103

	5.1.2 Develop a Strategy	104
	5.1.3 Set Up Data Integration	104
	5.1.4 Choose an End-User Interface	104
	5.1.5 End-User Training	104
	5.2 Privacy and ethics social networking and BI	105
	5.2.1 The Goldmine of Social Media Data	105
	5.2.2 What Are Social Media Insights?	105
	5.2.3 What Is Social Media Business Intelligence?	106
	5.2.4 Key Metrics of social media	106
	5.2.5 How Business Intelligence Fits In	107
	5.2.6 Key Benefits of BI in social media	108
	5.2.7 Social Media Business Intelligence Can Help Businesses	108
	5.2.8 Monitor your Competition	109
	5.2.9 Analyze Social Media Data	109
	5.2.10 Stay Up-To-Date on Industry Trends	109

UNIT- I

1. What is Data?

Data is a collection of information gathered by observations, measurements, research or analysis. They may consist of facts, numbers, names, figures or even description of things. Data is organized in the form of graphs, charts or tables. There exist data scientist who does data mining and with the help of that data analyse our world.

You must have watched the weather forecast reports on news channels. They list the minimum temperature, the maximum temperature, rainfall predictions and measurements. The tabular representation is indicated below.

Table 1: Tabular Representation

City	Min. Temp. (in Degrees)	Max. Temp. (in Degrees)	Rain
Mumbai	25	40	22%
Delhi	32	45	16%
Bangalore	23	35	28%
Kolkata	30	42	20%
Chennai	33	48	21

Classification of Data

Data is classified into

- Qualitative: It describes the quality of something or someone. It is descriptive information. For example, the skin colour, eye colour, hair texture, etc. gives us the qualitative information about a person.
- Quantitative: It provides numerical information. Example, the height and weight of a person.

1.1 Data Warehouse Definition

A Data Warehouse (DW) is a relational database that is designed for query and analysis rather than transaction processing. It includes historical data derived from transaction data from single and multiple sources.

A Data Warehouse provides integrated, enterprise-wide, historical data and focuses on providing support for decision-makers for data modeling and analysis.

A Data Warehouse is a group of data specific to the entire organization, not only to a particular group of users.

It is not used for daily operations and transaction processing but used for making decisions

1.1.1 Subject-Oriented

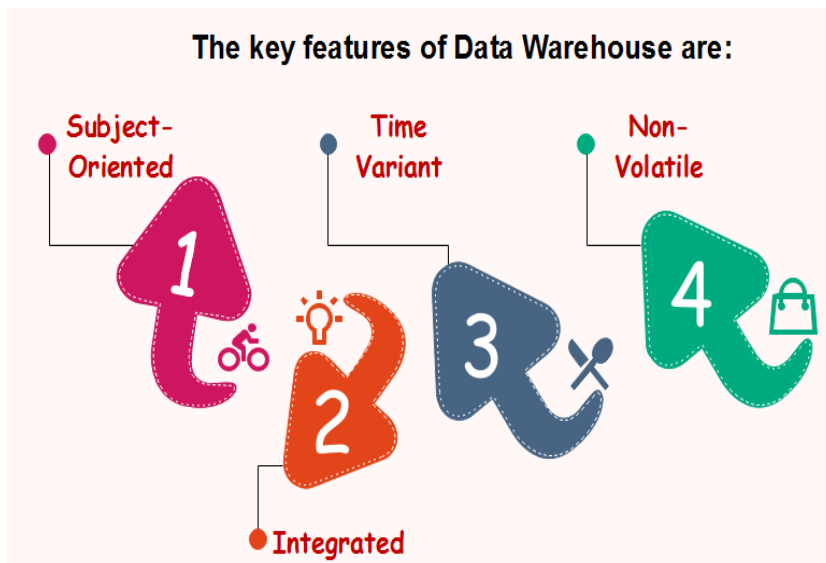


Figure 1.1: The Key Features of Data Warehouse

A data warehouse targets the modeling and analysis of data for decision-makers. Therefore, data warehouses typically provide a concise and straightforward view around a particular subject, such as customer,

product, or sales, instead of the global organization's ongoing operations. This is done by excluding data that are not useful concerning the subject and including all data needed by the users to understand the subject.

1.1.1.1 Integrated

A data warehouse integrates various heterogeneous data sources like RDBMS, flat files, and online transaction records. It requires performing data cleaning and integration during data warehousing to ensure consistency in naming conventions, attributes types, etc., among different data sources.

1.1.1.2 Time-Variant

Historical information is kept in a data warehouse. For example, one can retrieve files from 3 months, 6 months, 12 months, or even previous data from a data warehouse. These variations with a transactions system, where often only the most current file is kept

1.1.1.3 Non-Volatile

The data warehouse is a physically separate data storage, which is transformed from the source operational RDBMS. The operational updates of data do not occur in the data warehouse, i.e., update, insert, and delete operations are not performed. It usually requires only two procedures in data accessing: Initial loading of data and access to data.

1.1.2 Benefits of Data Warehouse

1. Understand business trends and make better forecasting decisions.
2. Data Warehouses are designed to perform well enormous amounts of data.
3. The structure of data warehouses is more accessible for end-users to navigate, understand, and query.
4. Queries that would be complex in many normalized databases could be easier to build and maintain in data warehouses.

5. Data warehousing is an efficient method to manage demand for lots of information from lots of users.
6. Data warehousing provides the capabilities to analyze a large amount of historical data.

1.1.3 Need for Data Warehouse



Figure 1.2: Need of Data Warehouse

- 1. Business User:** Business users require a data warehouse to view summarized data from the past. Since these people are non-technical, the data may be presented to them in an elementary form.
- 2. Store historical data:** Data Warehouse is required to store the time variable data from the past. This input is made to be used for various purposes.
- 3. Make strategic decisions:** Some strategies may be depending upon the data in the data warehouse. So, data warehouse contributes to making strategic decisions.
- 4. For data consistency and quality:** Bringing the data from different sources at a commonplace, the user can effectively undertake to bring the uniformity and consistency in data.

5. High response time: Data warehouse has to be ready for somewhat unexpected loads and types of queries, which demands a significant degree of flexibility and quick response time.

1.1.4 Database Vs Datawarehouse

Table 2: Database vs. Data Warehouse

Feature	Database	Data Warehouse
Primary Function	Store and manage data for day-to-day operations	Analyze historical data for trends and insights
Data Structure	Optimized for fast retrieval and modification (CRUD – Create, Read, Update, Delete)	Optimized for complex queries and analysis (OLAP – Online Analytical Processing)
Data Currency	Primarily current data	Primarily historical and integrated data from various sources
Schema	Highly normalized to minimize redundancy	Often denormalized to improve query performance for analysis
Updates	Frequent updates as transactions occur	Periodic updates (batch processing)
Users	Operational applications, individual users	Business analysts, data scientists, executives
Security	Focuses on data integrity and access control for specific users	Focuses on data governance and access control for analytical purposes
Complexity	Simpler to design and manage	More complex to design, implement, and maintain due to data integration and transformation
Cost	Lower cost due to smaller size and simpler infrastructure	Higher cost due to larger storage requirements and processing power

1.2 Data Warehouse Modeling

Data warehouse modeling is the process of designing the schemas of the detailed and summarized information of the data warehouse. The goal of data warehouse modeling is to develop a schema describing the reality, or at least a part of the fact, which the data warehouse is needed to support.

Data warehouse modeling is an essential stage of building a data warehouse for two main reasons. Firstly, through the schema, data warehouse clients can visualize the relationships among the warehouse data, to use them with greater ease. Secondly, a well-designed schema allows an effective data warehouse structure to emerge, to help decrease the cost of implementing the warehouse and improve the efficiency of using it.

Data modeling in data warehouses is different from data modeling in operational database systems. The primary function of data warehouses is to support DSS processes. Thus, the objective of data warehouse modeling is to make the data warehouse efficiently support complex queries on long term information.

In contrast, data modeling in operational database systems targets efficiently supporting simple transactions in the database such as retrieving, inserting, deleting, and changing data. Moreover, data warehouses are designed for the customer with general information knowledge about the enterprise, whereas operational database systems are more oriented toward use by software specialists for creating distinct applications.

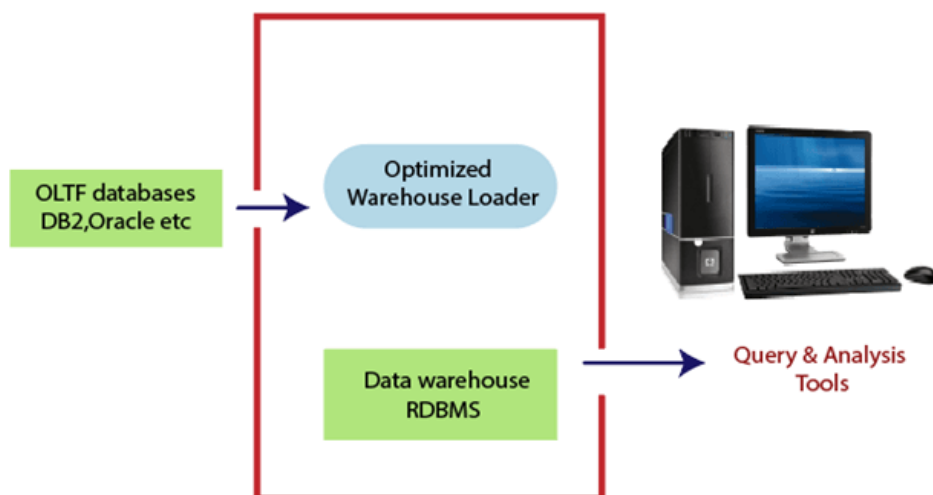


Figure 1.3: Data Warehouse Model

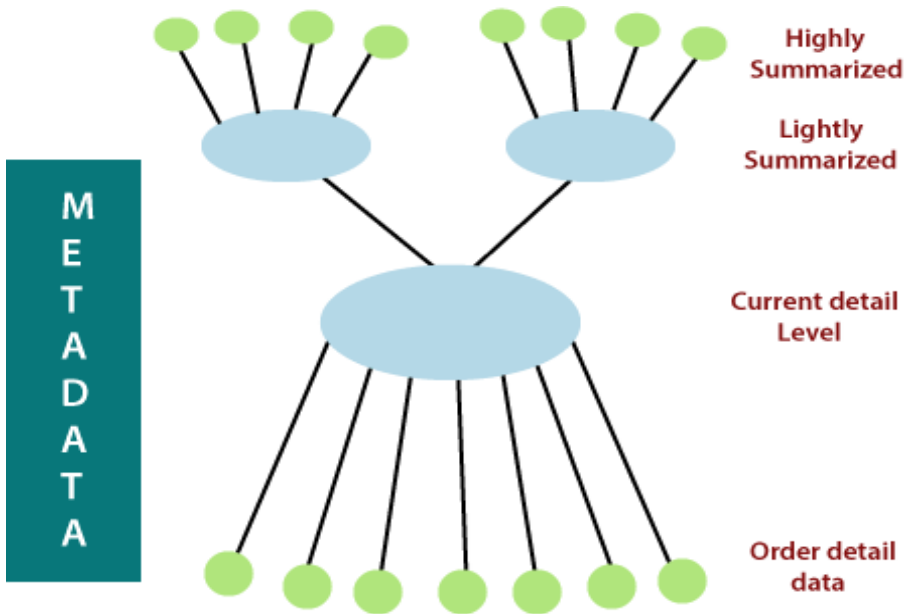


Figure 1.4: The Structure of Data Inside the Data Warehouse

- Reflects the most current happenings, which are commonly the most stimulating.
- It is numerous as it is saved at the lowest method of the Granularity.
- It is always (almost) saved on disk storage, which is fast to access but expensive and difficult to manage.

Older detail data is stored in some form of mass storage, and it is infrequently accessed and kept at a level detail consistent with current detailed data.

Lightly summarized data is data extract from the low level of detail found at the current, detailed level and usually is stored on disk storage. When building the data warehouse have

to remember what unit of time is summarization done over and also the components or what attributes the summarized data will contain.

Highly summarized data is compact and directly available and can even be found outside the warehouse.

1.2.1 Difference between OLTP and OLAP

Below is the difference between OLAP and OLTP in Data Warehouse

Table 3: Difference Between OLAP & OLTP in Data Warehouse

Parameters	OLTP	OLAP
Process	It is an online transactional system. It manages database modification.	OLAP is an online analysis and data retrieving process.
Characteristic	It is characterized by large numbers of short online transactions.	It is characterized by a large volume of data.
Functionality	OLTP is an online database modifying system.	OLAP is an online database query management system.
Method	OLTP uses traditional DBMS.	OLAP uses the data warehouse.
Query	Insert, Update, and Delete information from the database.	Mostly select operations
Table	Tables in OLTP database are normalized.	Tables in OLAP database are not normalized.
Source	OLTP and its transactions are the sources of data.	Different OLTP databases become the source of data for OLAP.
Data Integrity	OLTP database must maintain data integrity constraint.	OLAP database does not get frequently modified. Hence, data integrity is not an issue.
Response time	It's response time is in millisecond.	Response time in seconds to minutes.
Data quality	The data in the OLTP database is always detailed and organized.	The data in OLAP process might not be organized.
Usefulness	It helps to control and run fundamental business tasks.	It helps with planning, problem-solving, and decision support.
Operation	Allow read/write operations.	Only read and rarely write.

1.3 OLAP Operations

OLAP stands for Online Analytical Processing (OLAP) could be a innovation that's utilized to organize expansive business databases and

back business intelligence. OLAP databases are separated into one or more cubes, and each cube is organized and designed by a cube administrator to fit the way simply recover and analyze data so that it is less demanding to form and utilize the PivotTable reports and PivotChart reports that and just require.

1.3.1 OLAP Operations

There are five basic analytical operations that can be performed on an OLAP cube:

Drill down: In drill-down operation, the less detailed data is converted into highly detailed data. It can be done by:

- Moving down in the concept hierarchy
- Adding a new dimension

In the cube given in overview section, the drill down operation is performed by moving down in the concept hierarchy of Time dimension (Quarter -> Month).

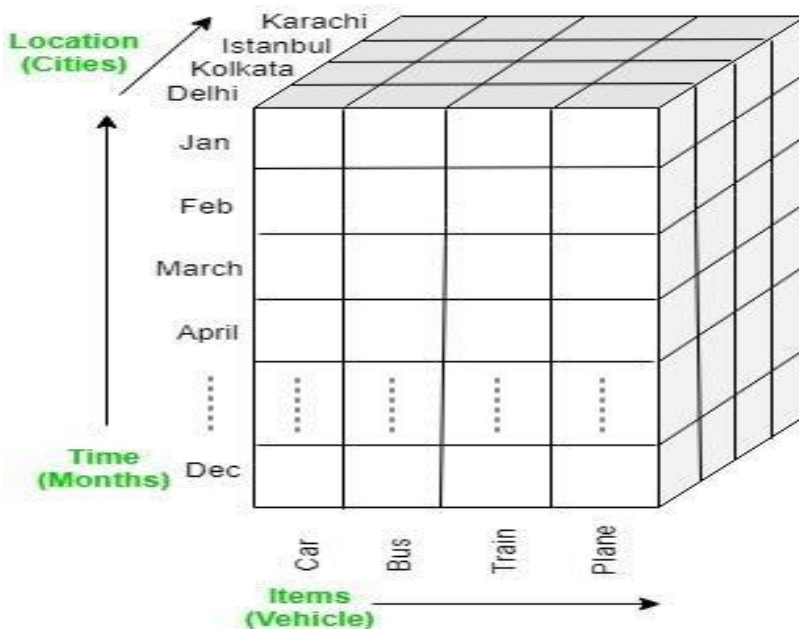


Figure 1.5: OLAP Operation

Roll up: It is just opposite of the drill-down operation. It performs aggregation on the OLAP cube. It can be done by:

- Climbing up in the concept hierarchy
- Reducing the dimensions

In the cube given in the overview section, the roll-up operation is performed by climbing up in the concept hierarchy of *Location* dimension (City -> Country)

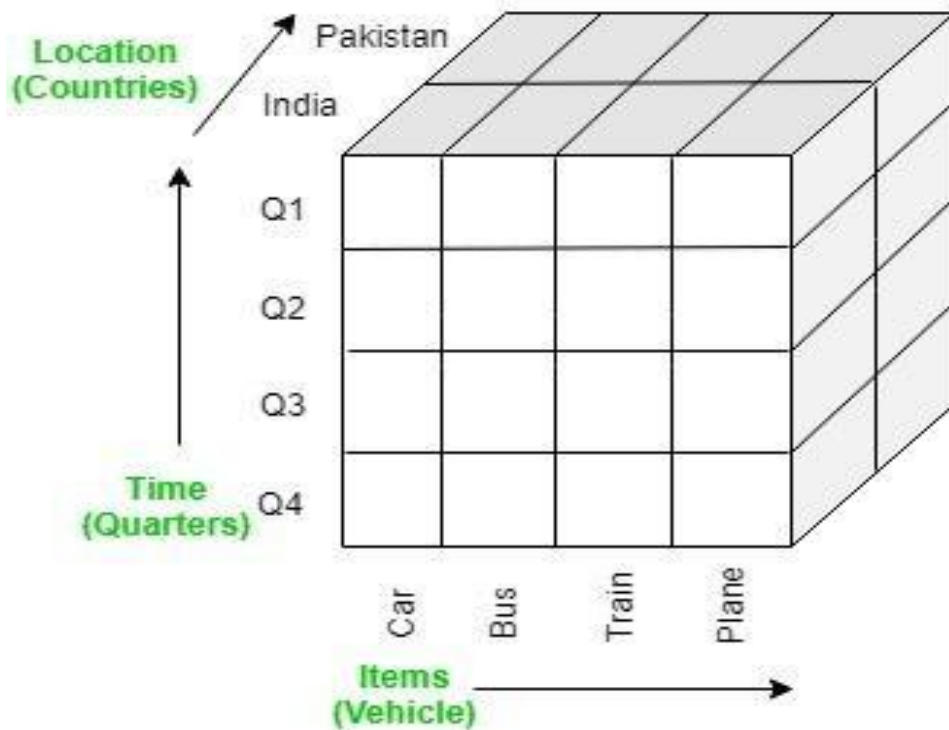


Figure 1.6: Roll-up Operation

Dice: It selects a sub-cube from the OLAP cube by selecting two or more dimensions. In the cube given in the overview section, a sub-cube is selected by selecting following dimensions with criteria:

- Location = “Delhi” or “Kolkata”
- Time = “Q1” or “Q2”
- Item = “Car” or “Bus”

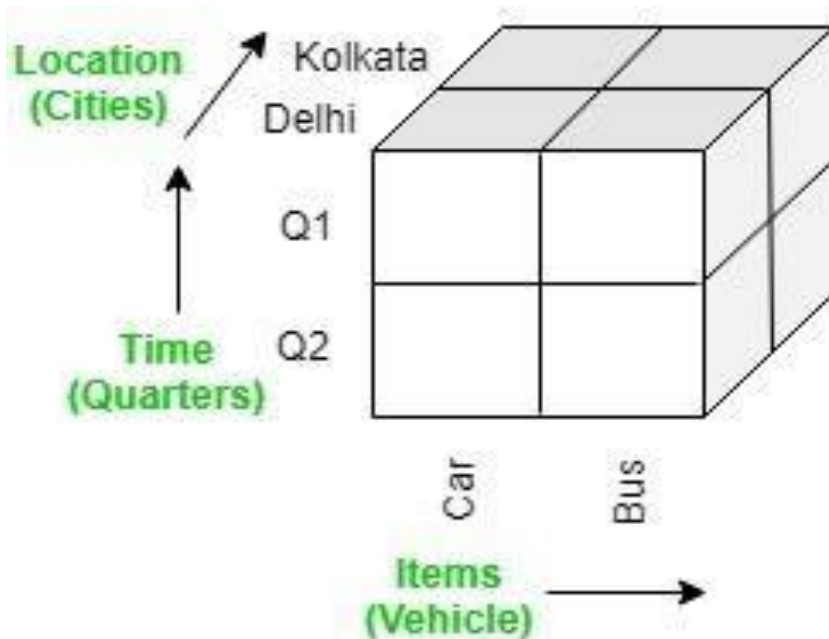


Figure 1.7: Dice Operation

Slice: It selects a single dimension from the OLAP cube which results in a new sub-cube creation. In the cube given in the overview section, Slice is performed on the dimension

Time = "Q1"



Figure 1.8: Slice Operation

Pivot: It is also known as rotation operation as it rotates the current view to get a new view of the representation. In the sub-cube obtained after the slice operation, performing pivot operation gives a new view of it.

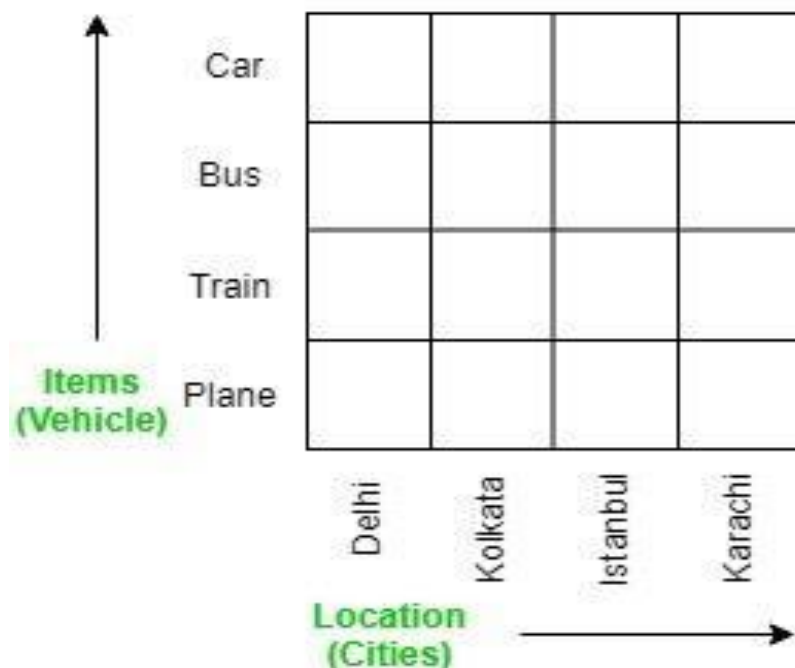


Figure 1.9: Pivot Operation

1.3.1.1 What is a Data Cube?

A data cube is a multidimensional representation of data intended to facilitate effortless retrieval and structured analysis of the underlying data. When organized in a cube rather than a network of relational tables, it becomes easier for the user to establish relationships between data that could otherwise be challenging to figure out. This directly results in enhanced in-depth analysis and advanced drill down. Every face of the cube can be programmed to represent a particular category and users can pivot the cube to look at the same data from a unique perspective.

While data cubes can exist as a simple representation of data, without any extensive capabilities to analyze large volumes, OLAP data cubes are particularly valuable for complex data analysis, including business intelligence as they provide a comprehensive view of information across

different dimensions, such as time, products, locations, or customer segments. For example, if you are looking at a sales data cube, different dimensions can show you data by year, product category, locations, customers, etc. So, whenever we mention a data cube from here on, we will be referring to an OLAP model.

1.3.1.2 What are Data Cube Classifications?

OLAP emerged as a response to the limitations of relational databases for analytical and multidimensional data processing. OLAP databases are optimized for complex queries, multidimensional analysis, and fast retrieval of processed data. They allow users to interact with data from different angles and hierarchies. OLAP has developed into three major classifications:

- **ROLAP:** Relational OLAP stores and processes data in a relational database. ROLAP servers typically query the relational database to generate reports and analyze data. Although scalable to large volumes of data, ROLAP can encounter scalability issues after a certain point due to its dependency on the underlying database. It integrates well with existing relational databases which makes it easier to implement and maintain.
- **MOLAP:** Multidimensional OLAP stores and processes data in a multidimensional format conceptualized like a cube. This structure makes it easier to establish relationships between data points and optimize the data for complex analytical queries. MOLAP cubes are pre-calculated and stored in a separate database from the source data. With many operations done beforehand, MOLAP does not require as much processing power from the relational database server.
- **HOLAP:** Hybrid OLAP combines the best features of ROLAP and MOLAP. HOLAP stores some data in a relational database and some data in the multidimension format. Therefore, HOLAP provides both

the scalability of ROLAP and the performance of MOLAP for organizations with large and complex data sets.

1.3.1.3 What are the Data Cube Operations?

Data cubes support various operations that allow users to examine and analyze data from different perspectives. Here is an overview of some key data cube operations:

- **Roll-up:** This operation adds up all the data from a category and presents it as a singular record. It is like zooming out of the cube and looking at the data from a broader perspective.
- **Drill-down:** While trying to access a transaction on the point-of-access, users need to descend into a dimension hierarchy. For instance, drilling down on the product dimension in the sales data cube would provide detailed sales figures for each product within each region.
- **Slicing:** When users want to focus on a specific set of facts from a particular dimension, they can filter the data to focus on that subset. Slicing a sales data cube to focus on “Electronics” would restrict the data to sales of electronic products only.
- **Dicing:** Breaking the data into multiple slices from a data cube can isolate a particular combination of factors for analysis. By selecting a subset of values from each dimension, the user can focus on the point where the two dimensions intersect each other. For example, dicing the product dimension to “Electronics” and the region dimension to “Asia” would restrict the data to sales of electronic products in the Asian region.
- **Pivoting:** Pivoting means rotating the cube to view the data from a unique perspective or reorienting analysis to focus on a different aspect. Pivoting the sales data cube to swap the product and region

dimensions would shift the focus from sales by product to sales by region.

1.4 Data Cube Computation Methods

1.4.1 Materialization of Cube: Full, Iceberg, Closed and Shell Cubes

A three-dimensional data cube with dimensions A, B, and C and an aggregate measure M. And can think of a data cube as a lattice of cuboids. Each cube is meant to symbolize a group. The fundamental cuboid, encompassing all three dimensions, is the ABC.

The aggregate measure (M) is calculated for each permutation of the three dimensions. There are six different cuboids that make up a data cube, with the base cuboid being the most specific. The apex cuboid is the most generalized cuboid. It stores a single number - the sum of all the tuples' measures in the base cuboid's measure M. From the topmost cuboid of the data cube example, we can descend into the lattice to access deeper levels of information.

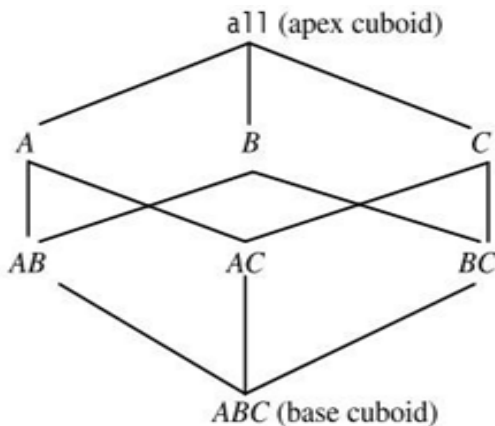


Figure 1.10: Materialization of Cube

1.4.1.1 Base Cell in Data Cube

The term "base cell" refers to a cell in the base cuboid. Aggregate cells are cells that are not based on a cube. Each dimension that is aggregated

in an aggregate cell is represented by a "" in the cell notation. Let's pretend we're working with an n -dimensional data cube.

Let each cell of the cuboids that make up the data cube be denoted by $a = (a_1, a_2, \dots, a_n, \text{measurements})$. If there are m ($m \leq n$) values of a , b , c , d , e , f , g , h , i , j , k , l , m , n , and o that are not "", then we say that a is an m -dimensional cell (that is, from an m -A is a base cell if and only if $m = n$; otherwise, it is an aggregate cell (where $m \leq n$)).

1.4.1.2 Cuboids in Data Cube

As an intriguing compromise between storage requirements and response times for OLAP, data cubes that are partially materialized are a viable option. We cannot compute the entire data cube, but rather only parts of it, called cuboids, each of which is composed of a subset of the cells in the full cube.

1.4.1.3 Minimum Support and Iceberg in Data Cube

Not only does this result in a more effective utilization of resources (namely, CPU time and disc space), but it also makes it possible to conduct more accurate analysis. There is a good chance that the non-passing cells are not important enough to warrant further investigation.

Cubes that only partially materialize are referred to as iceberg cubes, and this phrase is used to characterize such cubes. The term "minimum support," also abbreviated as "min sup" for short, describes the criteria that are the absolute minimum acceptable. It is common practice to refer to the effect of materializing only a fraction of the cells in a data cube as the "tip of the iceberg." In this context, "iceberg" refers to the entire cube including all cells.

1.4.2 Roll-Up/Drill-Down

This method involves aggregating data along one or more dimensions to create a summary of the dataset. It can be used to drill-down into specific areas of interest within the data. Roll-up/Drill-down is useful for

quickly summarizing large datasets into manageable chunks while still maintaining important information about each dimension.

For example, if and have sales data for multiple products across several regions, and could use roll- up/drill-down to see total sales across all regions or drill-down into sales numbers for one particular product in one region.

1.4.3 Slice-and-Dice

This method involves selecting subsets of data based on certain criteria and then analyzing it using different dimensions. It is useful for identifying patterns that may not be immediately apparent when looking at the entire dataset.

Slice-and-Dice allows users to select subsets of data based on specific criteria such as time period or customer demographics which can then be analyzed using different dimensions like product categories or geographic locations. This helps identify patterns that may not be immediately apparent when looking at the entire dataset.

1.4.4 Grouping Sets

This method involves grouping data by multiple dimensions at once, allowing for more complex analysis of the dataset. Grouping Sets are useful when analyzing large datasets with multiple dimensions where users want to group by two or more dimensions at once. For example, grouping sets could show total revenue broken down by both product category and region simultaneously.

1.4.5 Online Analytical Processing (OLAP)

This method uses a multidimensional database to store and analyze large amounts of data. It allows for quick querying and analysis of the data in different ways. OLAP databases are specifically designed for analyzing large amounts of multi-dimensional data quickly through pre-aggregated

values stored in memory making it ideal for real-time decision- making scenarios like stock market analysis.

1.4.6 SQL Queries

SQL queries can be used to compute data cubes by selecting specific columns and aggregating them based on certain criteria. This is a flexible method that can be customized based on the needs of the user.

UNIT-II

2. Data Ware Housing and Business Intelligence

2.1 Business Intelligence Introduction

Business intelligence, or BI, is a technology-driven data analysis process that helps an organization's executives, managers and workers make informed business decisions. As part of the BI process, relevant data is collected and prepared for analysis. Queries are run against the data, and the analytics results are used to support operational decision-making and strategic planning.

Traditional Business Intelligence, capital letters and all, originally emerged in the 1960s as a system of sharing information across organizations.



Figure 2.1: The Modern Analytics Workflow

Businesses and organizations have questions and goals. To answer these questions and track performance against these goals, they gather the necessary data, analyze it, and determine which actions to take to reach their goals. On the technical side, raw data is collected from business systems. Data is processed and then stored in data warehouses, the cloud,

applications, and files. Once it's stored, users can access the data, starting the analysis process to answer business questions.

BI platforms also offer data visualization tools, which convert data into charts or graphs, as well as presenting to any key stakeholders or decision-makers.

2.1.1 BI Methods

Much more than a specific “thing,” business intelligence is an umbrella term that covers the processes and methods of collecting, storing, and analyzing data from business operations or activities to optimize performance. All of these things come together to create a comprehensive view of a business to help people make better, actionable decisions. Over the past few years, business intelligence has evolved to include more processes and activities to help improve performance. These processes include:

- **Data mining:** Using databases, statistics, and machine learning (ML) to uncover trends in large datasets
- **Reporting:** Sharing data analysis to stakeholders so they can draw conclusions and make decisions
- **Performance metrics and benchmarking:** Comparing current performance data to historical data to track performance against goals, typically using customized dashboards
- **Descriptive analytics:** Using preliminary data analysis to find out what happened
- **Querying:** Asking the data-specific questions, BI pulling the answers from the data sets
- **Statistical analysis:** Taking the results from descriptive analytics and further exploring the data using statistics such as how this trend happened and why.

- **Data visualization:** Turning data analysis into visual representations such as charts, graphs, and histograms to more easily consume data
- **Visual analysis:** Exploring data through visual storytelling to communicate insights on the fly and stay in the flow of analysis
- **Data preparation:** Compiling multiple data sources, identifying the dimensions and measurements, and preparing it for data analysis.

2.1.2 The BI Process

Although business intelligence is utilized in different ways and for different purposes by individual companies, the process is fairly uniform throughout all industries and typically unfolds as follows:

Step 1: Data from various sources — including internal company data and external market data — is collected, integrated, and then stored; because “big data” is commonly used, data is commonly stored in what’s called a data warehouse, created by a data engineer

Step 2: Data sets are created and prepared for data analysis, often by creating data analysis models **Step 3:** Data analysts run queries against the data sets or models

Step 4: The results of queries are used to produce visualizations in the form of charts, graphs, histograms, or other visual representations, along with BI dashboards and reports

Step 5: Decision-makers utilize the data visualizations and reports to help them in making decisions; they may also use their BI dashboard to probe further into the data for more information.

2.2 Leveraging Data and Knowledge for BI

Data is everywhere and companies are collecting more of it than ever before. While companies are surrounded by a sea of data points and equipped with powerful tools designed to leverage that data, many organizations struggle to unlock the power hidden in all that information.

Leveraging data might be the secret to success for everything from marketing to product development, but it's not that simple. Not all businesses use data in the same way or for the same purpose. Many organizations struggle to simply collect, analyze and leverage data to build a solid business strategy.

Leveraging data involves the strategic use of data in order to improve business and gain a competitive advantage. It involves analyzing data from a variety of sources to measure performance and process, set strategic goals, and guide improvement. Effectively leveraging data not only involves collecting and analyzing data points, it also requires companies to act on the insights from that data.

Data collected by a business can be used in many ways, including to:

- **Identify customer preferences and trends:** This could enable the tailoring of products and services to better meet market demands.
- **Inform marketing strategies and the allocation of resources:** This might include segmenting audiences into different groups based on their purchase or usage data.
- **Understand user behavior:** User behavior, including user experience data, could provide insight into how customers interact with a product or service.
- **Using Data to Build Business Strategy:** A successful business strategy should encompass all parts of the business. This means that you must also use data that represents all parts of the business, but that data might serve a different purpose in the overall picture of your organization's strategy for success.

2.2.1 Data Sources for Business Strategy

Leveraging data can be a daunting task for any organization. Data can come from a variety of sources, from customers to competitors and just about everywhere in between. This data can be analyzed separately from

other sources or integrated to create a more nuanced picture of your business's place within the market.

It's important to explore the types of data you can derive from internal and external sources as well as the benefits and drawbacks connected with using that data.

*(i) **Leveraging Internal Data Sources***

- **Customer Data-** Customer data can take a variety of forms, including customer purchase history, demographics, and feedback data (such as CSAT or NPS scores). Using customer data helps drive personalized marketing and audience segmentation, improves customer experience, and can drive retention. However, this comes with concerns about privacy and data accuracy, as well as obtaining consent for data collection and usage. It's important to follow all privacy laws when handling consumer data.
- **Financial Data-** Financial data relates to the company's finances, including debts and income. This may include income statements, cash flow statements, or financial forecasts. Leveraging financial data can support strategic financial planning, budgeting, and risk management for your company. Some drawbacks to using financial data include data accuracy and limited access to comprehensive data. This could impact the accuracy of financial forecasts and may present obstacles to complying with financial regulations.
- **Employee Data-** This may include employee performance records, HR surveys, payroll data, or training history data. Companies can leverage employee data to support human resources decision-making, manage and retain talent, and engage employees. There are some disadvantages to collecting and using employee data. Specifically, companies should always protect employee privacy and anonymity, and make sure that they are complying with labor laws.

- **Social Media Data-** Social media data can be gathered from on-platform content such as comments on your company's posts, organic posts created by other users, reviews, and more. With social media, you can leverage real-time customer engagement to analyze sentiment, monitor mentions of your brand, and execute targeted advertising. Although it can be difficult to identify relevant conversations and manage the high volume of interactions on social media, it can provide a rich source of data to leverage.
- **Web Analytics Data-** Web data can include website traffic, conversion rates, and user experience data. Deriving insights from web analytics data can support website optimization, enable companies to segment customer experiences on their website, and reinforce the effectiveness of digital marketing. There are some drawbacks to using web analytics data. It can be difficult to track user interactions between sites, or as users come to the website from different marketing channels.
- **IoT (Internet of Things) Data-** IoT refers to when physical objects are equipped with sensors, software, and other technology that connects and exchanges data. IoT data could include sensor data, machine performance metrics, GPS data, or temperature readings. This enables real-time monitoring of products, predictive maintenance, and process optimization. Although IoT enables rich business data, managing such large volumes of data, storing, and processing that data, and ensuring data security and privacy can be a challenge.

(ii) Leveraging External Data Sources

- **Market Data-** Market data can include surveys, focus groups, industry reports, and competitor research data. Using market data can inform strategic decision-making and fuel competitive intelligence. It can also help your company proactively identify market trends. One drawback to using market research data is that it is costly and time-consuming to collect, which makes it difficult to

stay up to date on market trends. Additionally, possible bias from the sources and methodologies can make it difficult to accurately interpret research findings.

- **Competitor Data-** Like market data, competitor data is essential to any business strategy. Companies can leverage competitor data such as pricing information, product features, market share reports, and customer reviews to benchmark their own performance, conduct competitive analysis, and identify market opportunities based on unaddressed market problems.
- **Supply Chain Data-** Supply chain data may include shipping records, inventory levels, supplier performance metrics, and demand forecasts. Leveraging supply chain data can improve efficiency, drive cost reductions, and enhance demand forecasting. On the other hand, supply chain data collection from suppliers may be inconsistent or delayed. Additionally, it may be difficult to integrate complex supply chain data from a variety of sources.
- **Government and Public Data-** Public data can include economic indicators such as GDP or employment rates, government census data, and regulatory filings. Leveraging public data can inform market entry or exit decisions, regulatory compliance, and economic trend analysis. Some drawbacks to using government and public data include the risk of outdated or incomplete data, as well as challenges to validating the accuracy of the data or difficulties of customizing the data.

2.3 BI Components

Business intelligence processes analyze data and deliver actionable insights that help managers and executives make informed decisions.

2.3.1 BI involves

- Collecting information from internal and external sources.

- Preparing data for analysis.
- Running queries.
- Creating intuitive visualizations, reports and dashboards.

2.3.1.1 Main Components of Business Intelligence System

1. Data Source
2. Data Mart / Data Warehouse
3. Data Exploration
4. Corporate Performance Management
5. Data Mining
6. Optimization
7. Decisions

Data Source

To begin, the first step is gathering and consolidating data from an array of primary and secondary sources. These sources vary in origin and format, consisting mainly of operational system data but also potentially containing unstructured documents like emails and data from external providers.

Data Mart / Data Warehouse

Through the utilization of extraction and transformation tools, also known as extract, transform, load (ETL), data is acquired from various sources and saved in databases designed specifically for business intelligence analysis. These databases, commonly known as data warehouses and data marts, serve as a centralized location for the gathered data.

Data Exploration

The third level of the pyramid offers essential resources for conducting a passive analysis in business intelligence. These resources include query and reporting systems, along with statistical methods. These techniques are referred to as passive because decision makers must first develop ideas

or establish criteria for data extraction before utilizing analysis tools to uncover answers and confirm their initial theories. For example, a sales manager might observe a decrease in revenues in a particular geographic region for a specific demographic of customers. In response, she could utilize extraction and visualization tools to confirm her hypothesis and then use statistical testing to validate her findings based on the data.

Corporate Performance Management

Dashboards weave several metrics related to business performance together to convey a data story. For instance, stock traders keep an eye on financial dashboards to track market fluctuations, analyze short and long-term trends in stock prices, and create investment opportunities.

Data Mining

The fourth level, known as active business intelligence methodologies, focuses on extracting valuable information and knowledge from data. Part II of this book will delve into various techniques such as mathematical models, pattern recognition, machine learning, and data mining. Unlike the tools discussed in the previous level, active models do not rely on decision makers to come up with hypothesis but instead aim to enhance their understanding.

Optimization

As you ascend the pyramid, you'll encounter optimization models that empower you to choose the most optimal course of action among various alternatives, which can often be quite extensive or even endless. These models have also been effectively incorporated in marketing and logistics.

Decisions

At last, the pinnacle of the pyramid reflects the ultimate decision made and put into action, serving as the logical end to the decision-making process. Despite the availability and effective utilization of business intelligence methodologies, the decision still lies in the hands of the decision makers, who can incorporate informal and unstructured

information to fine-tune and revise the suggestions and outcomes generated by mathematical models.

2.4 BI Dimensions

Dimensions and measures are the physical aspects of data aggregation. The most important BI dimensions discussed in this topic are:

- The degree of freedom along the most important BI implementation features.
- BI dimensions are largely set by business dimensions for a given business entity.
- The existing variety of business dimensions and BI options creates a multidimensionality of choices that encompasses both information space and BI functional space.

2.4.1 Information Space

Information sources, accessibility, efforts, and cost.

2.4.2 BI Functional Space

Methods, models, procedures, and software products.

The Most Important BI Dimensions Discussed in the Chapter

2.4.2.1 Internal Focus

Within business operations such as:

- Maintenance
- Personnel
- Finance

2.4.2.2 External Focus

Market-related aspects including:

- Customers
- Competitors
- Statistics from surveys
- Research and feedback

2.4.2.3 Centralized Deployment

Provides

- Greater data accuracy
- Veracity
- Protection (high level)

2.4.2.4 Decentralized Deployment

Storage and management of an organization's business data are distributed among various departments.

2.4.2.5 Question Complexity: From Simple to Complex

- Simple: Fully knowledgeable and predictable.
- Complex: Models provide only reasonable predictions within boundaries.

2.4.2.6 Functional Scope

Functional scope ranges from wide to narrow.

2.4.2.7 Level of Automation of BI Functions

Level 1

Data in structured format.

Level 2

Automated reporting.

Level 3

Automated insights with anomaly detection.

Level 4

Automated recommendations (Next Best Action).

Level 5

Automated action – the final level of automation where the system is empowered to undertake decisions without human guidance.

- Data-driven to insight-driven.
- Real-time to right-time.

- A compromise approach to select a BI system configuration.

2.5 Business Intelligence and Business Analytics

The terms Business Intelligence (BI) and Business Analytics (BA) are often confused but differ from each other. It is important to understand each concept and how they help a company.

2.5.1 Operational Systems

- Usually operated directly by shop-floor workers or front-line staff.
- Provide key data required to support management and operations.

2.5.2 Management Information System (MIS)

- Used by middle managers to help ensure the smooth running of the organization in the short-to-medium term.

Examples

- Sales Management System
- Budgeting System
- HRM System

2.5.3 Decision Support System (DSS)

- Analyzes existing structured information.
- Helps senior managers project the potential effects of decisions in the future

Business Intelligence refers to the process of gathering and analyzing data to make better decisions. Individuals can use BI at all levels of an organization, including managers, executives, and even individual employees.

2.5.3.1 Features of Business Intelligence

- BI focuses on collecting and analyzing data from multiple sources to identify trends and patterns.

- It helps organizations make better-informed decisions about the company's future directions.

2.5.3.2 BI Tools and Usage

Business analysts use BI tools such as:

- Dashboards
- Reports
- Scorecards
- Visualizations

These tools help in performing job duties such as:

- Creating reports
- Managing projects
- Analyzing information gathered from different departments within an organization.

2.5.4 Business Analytics (BA)

Business Analytics is a subset of Business Intelligence that focuses on applying statistical analysis techniques to gain insights into how customers use products or services.

As a result, Business Analytics professionals typically have more advanced degrees in mathematics or statistics than those who work with BI.

Example

- BI might tell business leaders what their current customers look like.
- BA might tell them what their future customers are doing.

Some experts use BA as a term to describe a set of predictive tools used within the realm of Business Intelligence.

2.5.5 Difference Between BI and BA

- BI (Business Intelligence) refers to the processes used to generate data reports.

- In contrast, BA (Business Analytics) refers to the processes used to apply data-driven insights to make decisions or take action.
- BI focuses mainly on reporting.
- Whereas BA focuses on using information to make better decisions.

2.5.5.1 New Analytics Strategy vs Existing Analytics Strategy Business Intelligence (BI)

- BI is concerned with aggregated data collected from various sources (like databases) and analyzed for insights about business performance.
- BI tools help users make better decisions by providing data-driven insights in various areas, including:
 - Finance
 - Marketing

2.5.6 Business Analytics (BA)

- BA involves an analytical approach to solving problems within a business context rather than just looking at aggregated data analytics.

2.5.6.1 Business Analytics (BA) Tools

BA Tools Are Employed for Many Functions, Including

- Correlational analysis
- Regression analysis
- Factor analysis
- Forecasting analysis
- Image analytics
- Others

2.5.7 Descriptive vs Predictive Analysis

The main difference between BI and BA is:

- BI focuses on descriptive analysis.
- BA focuses on predictive analysis.

2.5.7.1 Business Intelligence (BI)

BI is used to analyze historical data to predict future trends.

It is typically used for:

- Sales forecasts
- Product recommendations

2.5.7.2 Business Analytics (BA)

- BA uses predictive modeling to predict how different variables affect each other.

Example

If a retailer wants to know how price affects sales volume, a business analyst can help predict the effect that raising prices would have on the company's bottom line.

2.5.8 Managers vs Analysts

One of the biggest differences between BI and BA is that:

- Managers deal with data to make decisions.
- Analysts deal with data to answer questions.

2.5.8.1 Managers

Managers typically manage a company's operations and finances, including:

- Marketing
- Sales
- Production

They make decisions based on data collected from across the company.

2.5.8.2 Analysts

- Analysts work on producing reports for management or other stakeholders.
- As part of their responsibilities, they may also conduct research or analysis on behalf of a client or company.

2.5.9 Reporting vs Applying

Reporting and applying are two ways in which the fields of Business Intelligence and Business Analytics differ.

- BI (Business Intelligence) refers to the processes used to generate data reports.
- In contrast, BA (Business Analytics) refers to the processes used to apply data-driven insights to make decisions or take action.
- BI focuses on reporting.
- Whereas BA focuses on using information to make better decisions.

2.5.10 New Analytics Strategy vs Existing Analytics Strategy

2.5.10.1 Business Intelligence (BI)

BI is concerned with aggregated data collected from various sources (such as databases) and analyzed for insights about business performance.

BI tools help users make better decisions by providing data-driven insights in various areas, including:

- Finance
- Marketing

2.5.10.2 Business Analytics (BA)

- BA involves an analytical approach to solving problems within a business context rather than simply looking at aggregated data analytics.

2.5.11 Current Events vs Future Possibilities

- BI uses real-time information to generate new insights into how a business operates and helps take action based on those insights.

2.5.11.1 Business Intelligence (BI)

- BI focuses mostly on looking at the past and present to develop forecasts based on historical data.

- It uses past data to help create a picture of what things might be like in the future.
- BI can also be used to analyze what happened when certain changes were made in a business.
- This helps organizations make better decisions based on previous outcomes.

2.5.11.2 Business Analytics (BA)

- BA is more focused on looking at real-time data and analyzing it to find patterns or trends that help predict what might happen in the future.

2.5.12 Ease of Operations

Business Intelligence (BI) systems make it easy for businesses to:

- Store data
- Access data
- Analyze data

People with little technical knowledge can use BI software to access necessary information without writing code or creating detailed reports, which might otherwise be difficult.

2.5.12.1 Business Analytics (BA)

BA is more technical in nature.

It typically requires an understanding of:

- Statistics
- Programming languages such as:
 - Python
 - R

BA involves analyzing large amounts of data from multiple sources.

2.5.13 Tools in BI and BA

2.5.13.1 Business Intelligence (BI) Tools

BI uses various tools to collect, analyze, and report data. These tools include:

- Databases (such as SQL)
- Data warehouses (such as Hadoop)
- Business Intelligence applications (such as Tableau)
- Visualization tools (such as Microsoft Power BI)

2.5.13.2 Business Analytics (BA) Tools

BA analytics uses predictive models to forecast future trends.

These models use statistical methods such as:

- Regression analysis
- Time-series analysis

These techniques help predict future outcomes based on historical data points.

BA Tools Include

- RStudio
- SAS Institute
- SPSS

2.5.14 Applications

2.5.14.1 BI Applications Include

- Sales force Sales Cloud helps companies collect customer data from multiple sources and use it to create personalized customer experiences.
- BA is a subset of BI that uses statistical and mathematical techniques to analyze large datasets from which a business can extract meaningful information.
- BA helps companies make better decisions by identifying patterns and trends in existing data.

2.6 Information Hierarchy

The Information Hierarchy graph below organizes data, information, knowledge, and wisdom in layers, with an increasing level of abstraction and addition of knowledge, starting from the bottom-most data layer. Various analytical systems help to transform content from one layer to a higher one so as to be better comprehended by analysts.

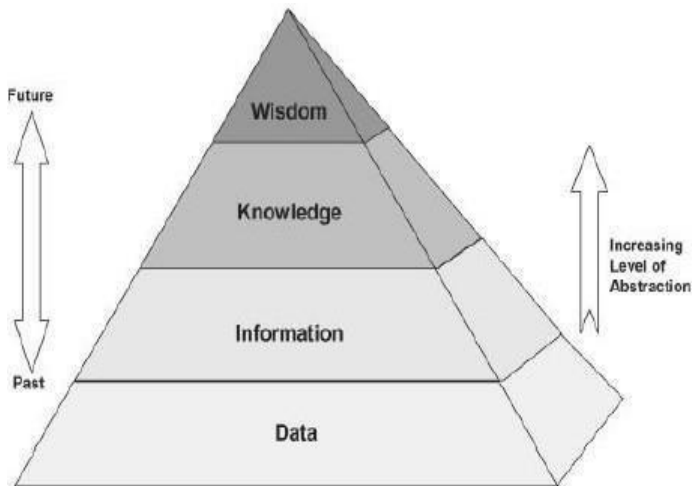


Figure 2.12: Information Hierarchy

2.6.1 Data Layer

Data are transactional, physical, and isolated records of activity (e.g., business transactions, customer interactions, facts or figures obtained from experiments or surveys). Data are, for example, numbers, texts, images, videos, and sounds, in a form that is suitable for storage or processing by a computer. Data are the most basic level and by themselves have little purpose and meaning.

2.6.2 Information Layer

Information is the semantic interpretation of data, and may represent relationships among data with meaning and purpose (e.g.: a customer called for a specific type of support). Information there for has been

organized and presented in a systematic fashion to clarify the underlying meaning.

2.6.3 Knowledge Layer

Knowledge is the general awareness or possession of information, facts, ideas, truths, or principles. Knowledge is generally personal and subjective. Knowledge can also be extracted from underlying information by applying some form of induction.

2.6.4 Wisdom Layer

Wisdom is the knowledge of what is true or right coupled with just judgment as to action. Wisdom requires a specific kind of knowledge and experience to make the right decisions and judgments in actions.

2.6.5 BI Life Cycle

Business Intelligence is the process of analyzing unprocessed data and turning it into knowledge that the company can use to make decisions. Organizations can better understand their marketing strategies, earnings and losses, effective development and management, and market trends and consumer behavior by using business intelligence.

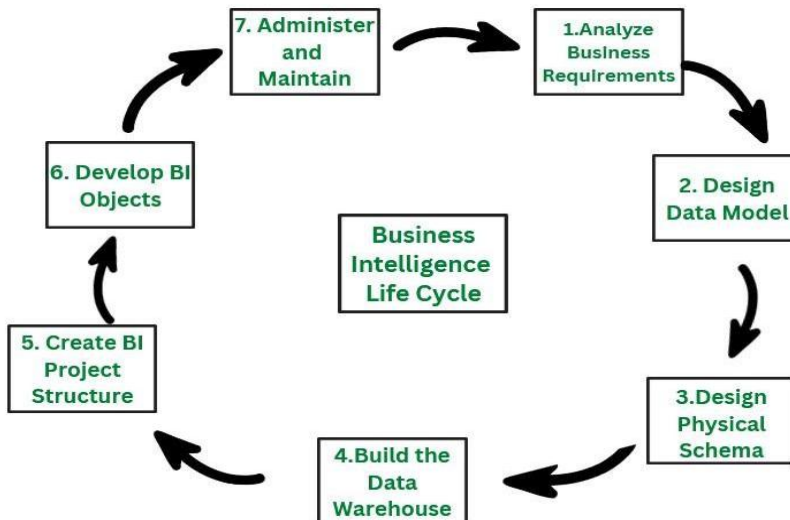


Figure 2.3: Business Intelligence Life Cycle

2.6.5.1 Phase1: Analyze Business Requirements

The first step in the Business Intelligence life cycle is to analyze the business requirements. The user identifies the business requirements in order to determine the type of analysis that the user then needs to perform. Identifying the requirements, let the user decides the further action to be performed.

For example, any retail company can analyze the sales data to figure out the products that are top- selling and the products that least sell.

2.6.5.2 Phase2: Design Data Model

Once the requirements are identified the user needs to design the logical model according to the requirements.

This logical model help the user to analyze the relationships that exist within the data entities. For example, for any retail company, the data model consists of products, their customers, and the sales data

2.6.5.3 Phase3: Design the Physical Schema

Once the logical model is prepared the next step is to design the physical schema using the data model. The physical schema describes the structure and the content of the data warehouse.

For example, in any retail company, physical schema consists of sales-related facts, product- customer relationships, and the sales transactions

2.6.5.4 Phase4: Build the Data Warehouse

Once the logical and physical schema is designed, the next step is to build the data warehouse. The design of a data warehouse depends on the physical and logical schema. After the design of the data warehouse, the data and the content from the source system are loaded into the data warehouse for further steps.

For example, for the retail system, designing the data warehouse consists of developing a database that would store the details of customers, products, and other requirements for the business.

2.6.5.5 Create the Project Structure (Metadata)

The next step after designing the data warehouse is to create a project structure also known as metadata. With the help of this created project structure, the mapping of the tables and data in the data warehouse is easier. Creating the project structure describes the further steps and types that need to be implemented.

For example, the project structure of the retail company consists of the attributes of the data, the design, and the working flow of the system. This project structure or metadata gives a brief idea about the working of the system.

2.6.5.6 Develop the BI Objects

The next step is to develop the BI objects such as metrics, attributes, dashboards, reports, and facts. This step consists of developing the reports and dashboards that can be used to analyze the data in the data warehouse. For example, the retail company can develop reports and statistics charts that can describe the profit and loss margins.

2.6.5.7 Administer and Maintain the Project

The last step is to administer and maintain the project continuously as it undergoes changes. The project needs to be monitored to maintain the changes, security, and performance of the system.

For example, the retail company needs to monitor the reports and statistics accordingly to increase the profit of the sales.

2.7 Data For Bi

2.7.1 Data Issues and Data Quality For Bi

BI data can include both historical and real-time data gathered from a combination of internal IT systems and external sources. Before it's used in BI applications, raw data from different source systems usually must be integrated, consolidated and cleansed to ensure it's accurate and consistent.

2.7.2 Types of Data Used in BI

- **Structured Data:** This is data organized in a predefined format, like in databases, spreadsheets, or tables.
- **Unstructured Data:** This data lacks a predefined format, like text documents, emails, or social media posts.
- **Semi-structured Data:** Data that has some structure but isn't as rigid as structured data, such as JSON or XML files.
- **Real-time Data:** Data that is collected and analyzed as it happens.
- **Historical Data:** Data from past periods, used for trend analysis and forecasting.

2.7.3 Key Aspects of Data Quality

- **Accuracy:** Data should be free from errors and reflect the real-world situation.
- **Completeness:** All necessary data points should be present, and no crucial information should be missing.
- **Consistency:** Data should be uniform and free from contradictions across different sources and systems.
- **Validity:** Data should conform to predefined standards, rules, and constraints.
- **Uniqueness:** Data should be free from duplicates or redundant entries.
- **Timeliness:** Data should be available when needed and reflect the current state of affairs.
- **Relevance:** Data should be meaningful and applicable to its intended purpose.
- **Integrity:** Data should remain reliable, accurate, and consistent throughout its lifecycle.

2.7.4 Why is Data Quality Important?

Data quality refers to the degree to which data meets established standards for accuracy, completeness, consistency, reliability, and timeliness,

ensuring it's trustworthy and suitable for analysis, decision-making, and other data-driven activities.

- **Better business decisions:** High quality data allows organizations to identify key performance indicators (KPIs) to measure the performance of various programs, which allows teams to improve or grow them more effectively. Organizations prioritize data quality will undoubtedly have an advantage over their competitors.
- **Improved business processes:** Good data also means that teams can identify where there are breakdowns in operational workflows. This is particularly true for the supply chain industry, which relies on real-time data to determine appropriate inventory and location of it after shipment.
- **Increased customer satisfaction:** High data quality provides organizations, particularly marketing and sales teams, with incredible insight into their target buyers. They are able to integrate different data across the sales and marketing funnel, which enable them to sell their products more effectively. For example, the combination of demographic data and web behavior can inform how organizations create their messaging, invest their marketing budget, or staff their sales teams to service existing or potential clients.

2.7.5 Data Issues For BI

Data quality issues can arise from various sources and affect the BI outcomes in different ways. In this article, we will explore some of the common data quality issues that affect BI outcomes and how to prevent or resolve them.

2.7.5.1 Incomplete Data

It refers to missing or insufficient information in the data sources. This can happen due to human errors, system failures, data integration issues, or lack of standards. Incomplete data can affect the BI outcomes by reducing the coverage, reliability, and validity of the analysis.

For example, if some customer records are missing or incomplete, the BI reports may not reflect the true customer behavior, preferences, or satisfaction. To avoid incomplete data, and should define clear data requirements, validate and monitor the data sources, and implement data quality rules and checks.

2.7.5.2 Inconsistent data

It refers to data that does not match or conform to the expected format, structure, or values. This can happen due to different data definitions, formats, or standards across the data sources, or due to data entry errors, duplication, or manipulation. Inconsistent data can affect the BI outcomes by creating confusion, ambiguity, and errors in the analysis.

For example, if some product names are spelled differently or have different codes in different data sources, the BI reports may not show the correct sales, inventory, or profitability. To avoid inconsistent data, should establish and enforce common data definitions, formats, and standards, and use data cleansing and deduplication tools.

2.7.5.3 Invalid Data

It refers to data that does not meet the business rules, logic, or expectations. This can happen due to data entry errors, data corruption, data manipulation, or data quality issues in the external data sources. Invalid data can affect the BI outcomes by producing inaccurate, misleading, or irrelevant results.

For example, if some sales transactions have negative or zero values, or some customer records have invalid email addresses or phone numbers, the BI reports may not show the correct revenue, conversion, or retention.

UNIT-III

3. BI Implementation

3.1 Data Ware Housing and Business Intelligence

Today, even small locally-dealing firms prefer to collect information on clients' demographics along with their tastes and preferences as to the products or services the vendor offers. For bigger organizations, collecting business data is a vital necessity that is highly instrumental in analyzing historical activities and the current stand of their venture as well as in divining market trends symptomatic of their niche for months and even years to come.

Purposing to outsell their rivals, blue-chip companies spend outrageous sums on similar activities turning business intelligence (BI for short) into their major competitive differentiator.

3.1.1 Is BI Really That Valuable?

The term “business intelligence” subsumes a whole range of procedures aimed at collecting, hoarding, and analyzing business operations information. It can be subsequently utilized to make knowledgeable data-driven decisions based on the comprehensive picture this overview creates.

What activities does an average business intelligence implementation plan include?

- **Data drilling.** Specific AI and machine-learning techniques are employed to discover and collect relevant data to trace trends in the obtained datasets.
- **Data arrangement.** By amassing information from multiple data sources and aligning measurements and dimensions used there, the information is prepared for further analysis.

- **Benchmarking.** On customized dashboards, historical data is compared against the current performance metrics to see how they dovetail into the preliminary set goals.
- **Querying.** In the collected database, you can seek answers to very specific business-related questions.
- **Data representation.** Charts, graphs, tables, and histograms can help visualize numbers to get them more “digestible” for further consumption.
- **Descriptive and statistical analysis.** At first, you find out what happened, then dig into the process and causes of the development under scrutiny.
- **Reporting.** When the analysis is completed, all stakeholders receive the results and plan their actions correspondingly. Once you implement business intelligence, you will be able to leverage all these summands to the maximum and drive much value for your business.

3.1.2 What Benefits Does Successful BI Implementation Bring in Its Wake?

3.1.2.1 Real-Time Data Tracking

In the modern volatile world, if you wait for reports to be prepared for a week or two, the information they contain may well be outdated when it reaches the decision-makers. Or your reaction will be inadequate and overlate. Or you will fail to notice some ominous development and prevent its adverse consequences. BI implementation provides instant access to all critical data and allows you to be proactive rather than reactive in your business choices.

3.1.2.2 Data Harmonization

Very often, fragmented pieces of data different departments possess prevent them from correctly observing the entire picture. By business intelligence implementation you will get a single version of the truth, which

is crucial for the data-driven decision-making CEOs are engaged in on behalf of the whole company.

3.1.2.3 Enhanced Visibility

Perusing hundreds of pages with statistics is tiresome and bewildering. Successful business intelligence implementation enables you to get a clear vision of all that happens in your organization (workflow-wise) and highlight areas and processes that you wish to focus your attention on.

3.1.2.4 Boosted Efficiency of Your Organization

Realization of problem areas is only the first stage. You should make conclusions as to where your employees underperform, get rid of redundant roles, and take other steps to optimize or even drastically improve your operations. Moreover, you will be able to see the implications of your actions, realize whether they have the desired effect, and fine-tune or even revamp your strategy accordingly.

3.1.2.5 Deeper Customer Insights

Better understanding your clients' behavior and purchasing motifs that form certain patterns is a key to improving both your products and your customer service.

3.1.2.6 Consistent Sales Insights

Having real-time as well as historic sales and marketing data at your disposal, you will be able to spot persistent trends in the niche, predict the success of certain sectors, and thus detect ways to increase revenues riding the tide.

3.1.2.7 A Competitive Edge

Business intelligence can be directed not only at the processes inside your organization but outside as well. You can compare your metrics with the competitors' indices and see where you lag behind and what they are doing better. Armed with this awareness, you can start planning a campaign to win their customers to your side.

3.2 Key Drivers Of BI

There are many business drivers in play today that are motivating companies to establish data warehouses. Current, consistent and accurate business information, they believe is critical for strategic and tactical decision making.

Some of the business drivers are summarized below:

3.2.1 Single Version of the Truth

Fragmented, inconsistent and outdated data in multiple databases does not permit good strategic and tactical decision making. Companies require that business intelligence be consolidated and presented in a suitable format for decision making. Inconsistent information from disparate information systems is no longer acceptable.

Data Warehouses help companies to achieve a single version of the truth by consolidating the most accurate and current data from the most reliable systems.

3.2.2 Current and Accurate Information

In a highly competitive market place, businesses need to quickly identify problems and opportunities in order to respond to events expeditiously and appropriately. Up-to-date information on sales, profits, inventories and customers can help identify problems early and leverage opportunities that could otherwise be missed. Most application systems are too narrowly scoped and operate on cycles that don't support real- time or near real-time information access. A data warehouse, however, can be designed to deliver up-to-date accurate information to decision makers.

3.2.3 Rapidly Changing Information Needs

It is very difficult for businesses to anticipate future information needs. Application systems often seem rigid and unable to adapt to evolving management information needs. Businesses need the flexibility to slice and dice data in many ways in order to identify and analyze changes in the

market place or in the business itself. Data Warehouses are designed for online, analytical purposes and provide great flexibility.

3.2.4 Customer Service Excellence

It is often said that 10% of a business's customers account for 90% of the business's profits. Identifying the good customers and providing them with excellent service helps retain good customers. Data Warehousing can help identify a company's best customers using a any number or criterion.

3.2.5 New Service Delivery Channels

It is no longer sufficient to provide customers with just 9:00 AM to 5:00 PM in-store service. Customers want to do business 7 days a week, 24 hours per day using alternate service delivery channels such as via the Internet or telephone. By examining all customer transactions, regardless of the channel used, businesses can better understand their customers and serve them better. Data Warehousing is critical for profiling customers and their transactions, regardless of the channel used. There are many technical drivers in play that are motivating companies to establish data warehouses for online queries and analytics. These are summarized below.

3.2.6 Multiple Internal Databases

Most medium and large businesses operate dozens, if not hundreds of un-integrated application systems. Individual departments in companies often focus on their own narrow system and information needs and don't see the corporate value of integrating data. When silos of un-integrated data exist, data soon gets out of synch. Companies have a need for database that reflects a "single version of truth". Data Warehouses can help do that.

3.2.7 Purchased Packages

"Out of the Box" purchased applications sometimes use underlying concepts and definitions that differ from those used by the business in existing custom-built applications. For example, a "customer" in one system could encompass all current and past customers plus potential

future customers. In another system, a customer might be defined more narrowly as someone who has purchased a product and service during the past 12 months.

Such inconsistencies create problems from an analytical perspective. A count of customers done in the first database differs from a count done in the second. Companies have a need to align concepts and terminology. Data Warehouses help do this alignment.

3.2.8 Increasing Complexity of Systems

The underlying data structures of application systems are often very complicated. To create what would intuitively might appear to be a simple query often requires complex programming logic that involves navigating multiple database tables and or applications systems. Writing reports or queries can consequently take time and money. Companies have a need for a reporting environment that allows reports and queries to be generated quickly, inexpensively and without expensive IT skills. Data Warehouses can simplify the reporting environment.

3.2.9 Application System Evolution

Businesses are highly dynamic and applications systems constantly need to be enhanced to support new business requirements. When systems are changed, reports and queries that access any changed tables must also be updated. This maintenance work can be very costly. Businesses have a need to trim their application support costs. Data Warehouses can help shelter reports and queries from system changes that occur in "front end" operational systems.

3.2.10 Computer Networks and External Databases

The rapid growth of computer networks has allowed companies to exchange data with their suppliers, consumers, government bodies and other groups. Businesses often have a need to integrate data from internal and external databases. Data Warehouse can be designed to integrate corporate data with external data for reporting purposes.

3.3 Key Performance Indicators and Performance Metrics

Key performance indicators are data that show and just how good and are at attaining and business goals. Meanwhile, metrics track the status of and business processes. With KPIs, and will know if and are hitting and overall business targets, while metrics focus on the performance of specific business processes.

KPI stands for key performance indicator, a quantifiable measure of performance over time for a specific objective. KPIs provide targets for teams to shoot for, milestones to gauge progress, and insights that help people across the organization make better decisions.

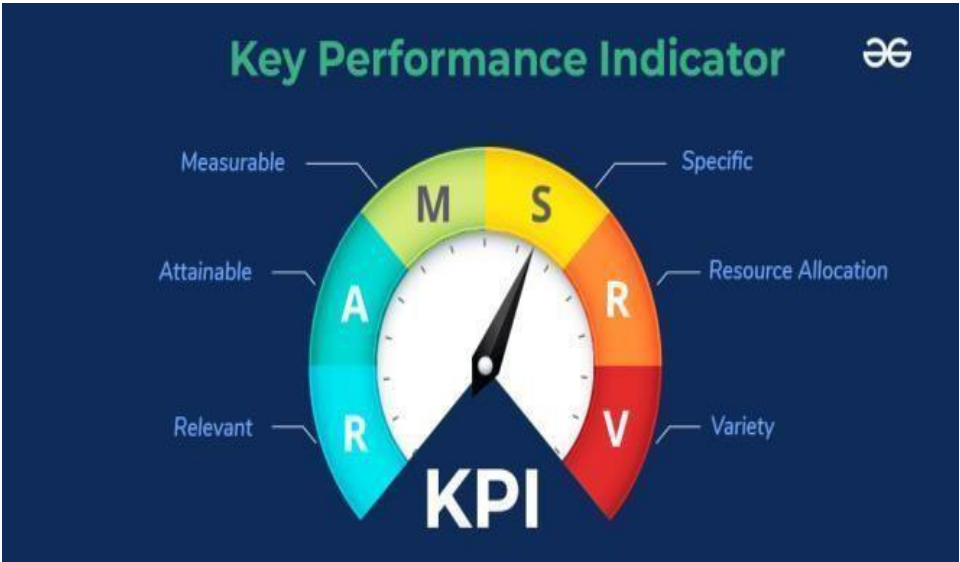


Figure 3.1: Key Performance Indicator

The best performing businesses work their way through numbers, and there’s no reason why and shouldn’t give more flesh to and business decisions as they do. Business intelligence key performance indicators or KPIs allow and toggle and the overall health of company, any one of departments or even how and customers perceive and company. And now a days, don’t have to tread the business intelligence game manually. And

just have to invest in an excellent business intelligence tool to weave the magic numbers for and, a feat that is severely limited or altogether forbidding in the past. If and are resourceful enough, should know which business intelligence tools is best for and unique needs.

3.3.1 Financial Metrics

From a tool like and accounting software, and should be looking at and cash flow, and balance sheet and income statement to derive and financial metrics. One look at any of these metrics should tell and if and business is healthy moneywise, which means and are generating revenue and astutely handling and finances. If and are planning to steer and company into new growth direction or to spark interest from potential investors, and will be presenting these financial KPIs to them for proof of investment value.

3.3.2 Marketing Metrics

Business-wise, next only to financial metrics in the ladder of importance. Marketing metrics reveal the numbers to inform whether your latest marketing campaigns are delivering the rates have set to reach. Capable marketing software tools should give the values, for example, on new content approach riding latest marketing campaigns across multiple channels.

3.3.3 Project Management Metrics

i) Productivity

Productivity is a straight forward computation of how much you are getting from all those that you put into a project. By dividing all the input units with the output units, you get a good picture of how efficient you are using the resources available to you for each project.

ii) Return on Investment (ROI)

Return on investment tells you how much you get in return for each dollar you invested in a project. It's a simple equation that involves the

project net benefits divided by the costs, then multiplying the result by 100.

3.3.4 Customer Service Metrics

i) Net Promoter Score (NPS)

Sourced from a single-question survey, this gold standard customer experience metric gives you a singular picture of customer loyalty. Customers respond with values ranging from 0 to 100, where you will be aiming for the higher end of the range. Get promoters from those who score 9 or 10, passive from those who score 8 or 9 and detractors from those who give a score of 0 to 6.

3.3.5 What is Business Intelligence Architecture?

Business intelligence architecture consists of a collection of principles and guidelines that empower business organizations to leverage computer-based methods and technologies for generating meaningful insights. These methodologies are significantly for business intelligence design systems that streamline online data visualization, reporting and analysis.

The fundamental purpose of business architecture is to close the gap that exists between an organization's strategic objectives and its operational realities. It entails comprehending the purpose, vision, and goals of the company and converting them into real-world systems and procedures that facilitate efficient execution. Business architecture strives to improve operations, boost efficiency, and spur innovation by coordinating the many parts of an organization.

3.3.5.1 Advantages of Business Intelligence Architecture

Business intelligence architecture can be deployed by business companies at any point to evaluate, handle, and depict business information.

Here are some of the advantages of Business Intelligence Architecture:

- 1. Enhanced Decision-Making:** Businesses can access, analyze, and store data using a systematic process thanks to business intelligence architecture. This makes it easier to make well-informed decisions based on up-to-date, pertinent, and correct information, which produces more successful plans of action.
- 2. Enhanced Operational Efficiency:** Organizations may simplify their operational procedures by using business intelligence architecture. It minimizes human mistakes by automating the data integration, cleansing, and transformation processes. This raises the quality of the data and increases the effectiveness of the reporting, analysis, and forecasting processes. Organizations may more efficiently distribute resources, lower expenses, and boost overall productivity by optimizing their processes.
- 3. Better Business Insights:** Business Intelligence Architecture enables organizations to glean insightful information from massive data sets. It reveals patterns, trends, and correlations within the data by using sophisticated analytics techniques, including data mining, statistical analysis, and predictive modeling. This enables organizations to identify market opportunities, understand customer behavior, optimize sales and marketing efforts, and make data-driven decisions to stay ahead of the competition.
- 4. Enhanced Data Visualization and Reporting:** Business Intelligence Architecture makes it easier to create data visualizations and reports that are both logical and aesthetically pleasing. Stakeholders may readily understand complicated information by using interactive dashboards, charts, and graphs. By doing this, the organization as a whole is better able to share knowledge on important metrics, trends, and performance indicators.
- 5. Enhanced Competitive Advantage:** Organizations may obtain a competitive edge by utilizing business intelligence architecture. They

can proactively keep an eye on consumer preferences, market trends, and business advancements. They can react quickly to shifting business dynamics and make data-driven decisions before their rivals if they have access to real-time or almost real-time data. Organizations can take advantage of new possibilities, adjust to market changes, and promote innovation thanks to their agility and reactivity.

6. A Stronger Focus on Data Governance and Compliance: Business intelligence architecture encourages data governance and compliance inside organizations. In order to guarantee data accuracy, security, and privacy, it creates standards, rules, and processes for data management. By enabling organizations to follow industry best practices and legal standards, it lowers the risk of data breaches and non-compliance.

7. Scalability and Flexibility: The business intelligence architecture is designed to accommodate the evolving needs of organizations. It provides scalability to handle increasing data volumes, user demands, and analytical complexity. The architecture can adapt to changes in technology, data sources, and business requirements. This scalability and flexibility allow organizations to grow and evolve their analytics capabilities over time, ensuring the architecture remains aligned with their expanding needs.

3.3.6 Business Intelligence Architecture Layers

A number of layers make up the business intelligence architecture, which combines them smoothly to turn raw data into insightful information. In the total pipeline for data processing and analysis, each layer is essential.

Let's delve further into these layers:

1. Data Source Layer: The Business intelligence architecture's underlying structure is the Data Source Layer. It includes a variety

of internal and external data sources, including cloud services, databases, spreadsheets, and APIs. Structured, semi-structured, or unstructured data may be present in these sources. This layer's main goal is to collect data from many sources while maintaining its accuracy, integrity, and security. The data integration layer plays a crucial role in consolidating and standardizing data obtained from diverse sources. Its primary objective is to merge, cleanse, and harmonize data from multiple origins. This layer encompasses operations such as data extraction, data transformation, and data loading (ETL). Its purpose is to ensure the reliability, accuracy, and analysis readiness of the data.

- 2. Data Storage Layer:** The Data Storage Layer is in charge of storing the combined and processed data. Usually, a data lake or warehousing strategy is used. While data lakes contain unprocessed, raw data in its original state, data warehouses organize data into organized schemas that are best for analysis and querying. This layer makes it easier to retrieve data quickly and offers a framework for additional analysis.
- 3. Data Modeling Layer:** To facilitate effective data analysis, the Data Modeling Layer entails the development of logical models and schemas. It involves dimensional modeling, which divides data into dimensions and metrics (such as time, product, and location). This layer supports ad-hoc searches and reporting while making it simple to navigate, aggregate, and slice and dice data.
- 4. Business Analytics Layer:** This layer concentrates on deriving valuable insights from the data. To find patterns, trends, and correlations, it makes use of a variety of analytical approaches, including statistical analysis, data mining, and machine learning. Users may do complex analyses, produce reports, and build interactive visualizations using this layer to aid decision-making processes.

5. Presentation Layer: Stakeholders communicate with the Business Intelligence system through the Presentation Layer’s user interface. It has tools for data visualization, reporting, and dashboards that make the analyzed data easy to understand and use. This layer promotes a data-driven culture inside the organization by empowering users to explore data, get insights, and make wise decisions.

3.3.7 Business Intelligence Architecture Components

Business intelligence architecture is made up of a number of parts that work together to help organizations gain useful insights from their data. Each element plays a significant role in the overall design, supporting data processing, analysis, and decision-making.



Figure 3.2: Business Intelligence Architecture Components

Let’s examine each of these elements in more detail:

- 1. Data Sources:** The core components of a business intelligence architecture are data sources. They comprise both internal systems (such as databases, ERP, and CRM systems) and external sources (such as social media, market research data, and web analytics). These sources offer the raw data needed for analysis and decision-making. The format and structure of the

data may vary, necessitating the use of proper extraction and integration procedures.

2. Unstructured, partially structured, or structured data may be present in the data sources. Semi-structured and unstructured data are more flexible in nature and may need extra processing and normalization, whereas structured data refers to well-defined and organized data with a preset format. Effective data integration and analysis require a thorough understanding of the properties and formats of the data sources.
3. **Data Integration Tools:** Data Integration Tools are responsible for extracting, transforming, and loading (ETL) data from various sources into the Business Intelligence system. These technologies make it possible to clean, combine, and integrate data, guaranteeing its consistency and quality. They manage data transformations, automate the gathering process, and provide easy data loading into the data retention layer. Data quality checks and validations are carried out throughout the integration process to find and fix any discrepancies or mistakes. Data enrichment, in which extra data from outside sources is added to improve the already-existing datasets, may also be a part of it. The data must be appropriately formatted, standardized, and connected for effective analysis and reporting, which is accomplished during the integration phase.
4. **Data Warehouse:** A data warehouse functions as a centralized repository that houses integrated and well-structured data derived from various sources. Its primary purpose is to support analytical processing, enabling efficient querying and reporting. To enhance data retrieval and analysis, data warehouses employ techniques such as indexing, segmentation, and data

compression. They offer a historical perspective and serve as a foundational element for data modeling and research endeavors.

5. **OLAP Cubes and Data Modeling:** Data modeling involves the creation of logical representations for the data stored within the data warehouse. Popular techniques for dimensional modeling include star schemas and snowflake schemas. These models establish relationships between variables, such as time, product, and customer, and outcomes, such as sales and revenue. Data models serve as the foundation for constructing multidimensional structures known as OLAP (Online Analytical Processing) Cubes, which facilitate efficient and interactive analysis of the data.
6. **Business Analytics Tools:** The term “business analytics tools” encompasses a diverse array of software applications employed for data analysis and generating valuable insights. These resources encompass predictive modeling software, platforms for machine learning, and programs for statistical analysis. They empower users to explore data, identify patterns, perform complex calculations, and develop models to support decision-making processes.
7. **Reporting and Visualization Tools:** Reporting and Visualization Tools offer an intuitive user interface for aesthetically appealing and meaningfully displaying analyzed data. They consist of reporting systems, interactive visualization software, and dash boarding tools. With the help of these tools, users may produce personalized reports, interactive charts, and visual dashboards that improve decision-making and enable data exploration.
8. **Data Security and Governance:** Vital elements of the business intelligence architecture are data security and governance. They entail putting policies into place to guarantee the availability,

confidentiality, and integrity of data. This comprises data encryption, data masking, access restrictions, and auditing tools. Additionally, data governance practices ensure compliance with regulatory requirements and establish data quality standards.

3.3.8 Business Intelligence Architecture Diagram

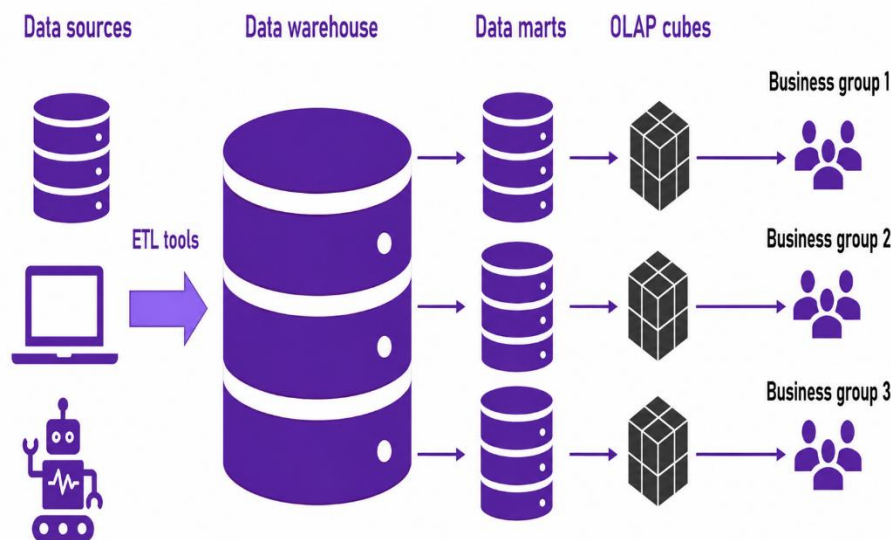


Figure 3.3: Business Intelligence Architecture Diagram

3.3.9 Business Intelligence Architecture Components and Diagram

A BI architecture can be deployed in an on-premises data center or in the cloud. In either case, it contains a set of core components that collectively support the different stages of the BI process from data collection, integration, data storage and analysis to data visualization, information delivery and the use of BI data in business decision-making.

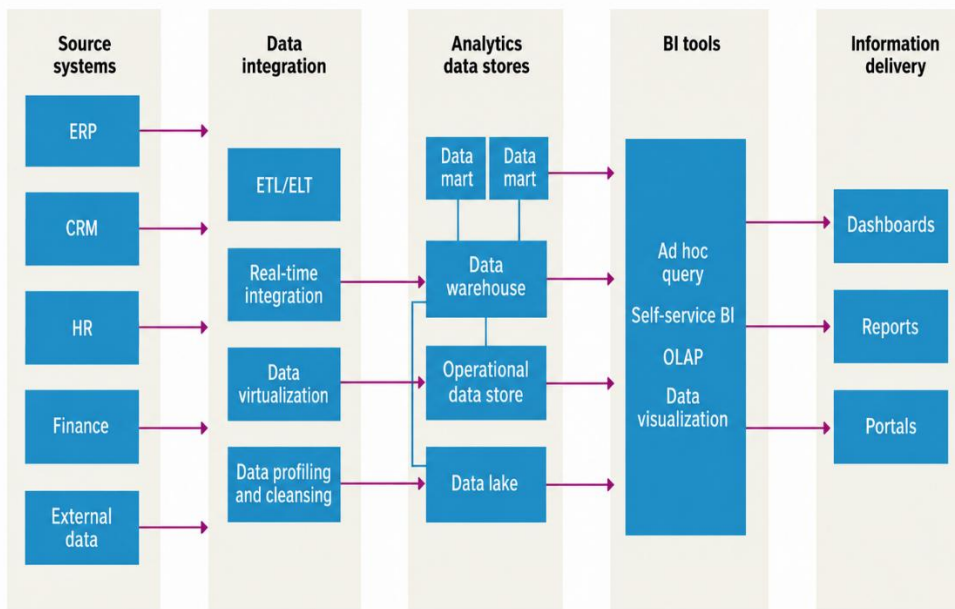


Figure 3.4: Sample Diagram of a Business Intelligence Architecture

This shows the various technologies used to run BI and analytics applications.

The core components of a BI architecture include the following:

Source Systems: These are all of the systems that capture and hold the transactional and operational data identified as essential for the enterprise BI program. For example, this can include enterprise resource planning, customer relationship management, flat files, application programming interfaces, finance, manufacturing and supply chain management systems as well as secondary sources, such as market data and customer databases from outside information providers. As a result, both internal and external data sources are often incorporated into a BI architecture.

Important criteria in the data source selection process include data relevancy, data currency, data quality and the level of detail in the available data sets. In addition, a combination of structured, semi-structured and unstructured data types might be required to meet the data analysis and decision-making needs of executives and other end users.

Data Integration and Cleansing Tools: To effectively analyze the collected data for a BI program, an organization must integrate and consolidate different data sets to create unified views of them. The most widely used data integration technology for BI applications is extract, transform and load (ETL) software, which pulls data from source systems in batch processes. A variant of ETL is extract, load and transform, a technology in which data is extracted and loaded as-is and transformed later for specific BI uses. Other methods include real-time data integration, such as change data capture and streaming integration to support real-time analytics applications, and data virtualization, which combines data from different source systems virtually.

A BI architecture typically also includes data profiling and data cleansing tools that are used to identify and fix data quality issues. They help BI and data management teams provide clean, consistent data that's suitable for BI uses.

Analytics Data Stores: This encompasses the various repositories where BI data is stored and managed. The primary repository is a data warehouse, which usually stores structured data in a relational, columnar or multidimensional database and makes it available for querying and analysis. An enterprise data warehouse can also be tied to smaller data marts set up for individual departments and business units with data that's specific to their BI needs.

In addition, BI architectures often include an operational data store (ODS) that's an interim repository for data before it goes into a data warehouse. An ODS can also be used to run analytical queries against recent transaction data. Depending on the size of a BI environment, a data warehouse, data mart and an ODS can be deployed on a single database server or separate business intelligence systems.

A data lake running on a Hadoop cluster or other big data platform can also be incorporated into a BI architecture as a repository for raw data of

various types. The data can be analyzed in the data lake itself or filtered and loaded into a data warehouse for analysis. A well-planned architecture should specify which of the different data stores is best suited for particular BI uses.

BI and Data Visualization Tools: The tools used to analyze data and present information to business users include a suite of technologies that can be built into a BI architecture -- for example, ad hoc query, data mining and online analytical processing software. In addition, the growing adoption of self-service BI tools enables business analysts and managers to run queries themselves instead of relying on the members of the BI team to do that for them.

BI software also includes data visualization tools that can be used to create graphical representations of data in the form of charts, graphs and other types of visualizations designed to illustrate trends, patterns and outlier elements in data sets.

Dashboards, Portals and Reports: These information delivery tools give users visibility into the results of BI and analytics applications with built-in data visualizations and, often, self-service capabilities to do additional data analysis. For example, BI dashboards and online portals can be designed to provide real-time data access

with configurable views and give users the ability to drill down into data. Reports tend to present data in a more static format.

Other components that increasingly are part of a business architecture include data preparation software used to structure and organize data for analysis and a metadata repository, a business glossary and a data catalog, which can help users find relevant data and understand its lineage and meaning.

3.3.10 BI Architecture Tools

BI architecture tools facilitate the centralization of data collection as well as data analysis and visualization. These tools play an integral role in empowering businesses to make informed decisions and extract insights from extensive data sets.

Some examples of BI tools on the market include the following:

1. **Datapine-** Datapine lets users access, view, analyze and share their company data on a single analytics platform. Users can perform data analysis, create interactive business dashboards and obtain new business insights through a simple drag-and-drop interface.
2. **Domo-** The Domo cloud-based platform unifies data, systems and people for seamless business operations. It provides enterprise tools for data aggregation, analytics, dashboards and reporting for organizations looking to maximize data value.
3. **Dundas BI-** This enterprise-level BI tool lets users create and customize interactive dashboards and reports. The software can either act as a central data hub or integrate into existing websites for customized BI capabilities.
4. **Good Data-** As part of the Good Data platform, this tool offers an enterprise-level option for data analytics and business intelligence. It helps users analyze data coming from multiple sources and create reports.
5. **Infor Birst-** Infor Birst is a cloud-based platform that uses a networked approach and modern enterprise- class architecture with a focus on multi-tenancy. Birst ensures that a company's data remains connected by centralizing both decentralized and centralized data.
6. **Microsoft Power BI-** Users can run analytics either in the cloud or in a reporting server. The tool comes with built-in artificial intelligence features and offers end-to-end encryption features.

7. **Oracle Business Intelligence-** This integrated set of tools lets users gather, store, analyze and report data for smart decision-making. In addition, it includes a scalable BI server, dashboards, a content library, web- based reporting and analytics tools.
8. **SAS Business Intelligence-** This collection of tools lets corporate users conduct self-service analytics. Its two components -- Enterprise Business Intelligence and Business Visualization -- provide interactive visualizations and analytics to aid with data analysis and decision-making.
9. **Tableau-** In addition to data visualization features, this tool offers live visual analytics and supports most databases and numerous data sources.
10. **Zoho Analytics-** This self-service BI and data analytics software lets users analyze data, generate data visualizations and uncover insights quickly and easily. This tool is accessible to both small and large-sized organizations.

3.4 Business Decision Making

One of the most concrete definitions of digital transformation we use is-- digital transformation is IT for business. Basically, application of digital capabilities and technologies in business processes, products, and models. Moreover, It is a cultural change. It is a change that challenges organizations to be innovative, adaptive and customer oriented.

Today, if we look at the area Herbert Simon had an impact on, we will see that there is no digital transformation without business intelligence (which among other things, consists of information processing and decision-making) and organizational changes.

3.4.1 Phases of the Decision-Making Process

Simon's model defines four phases of decision-making process:

- Intelligence Phase

- Design Phase
- Choice Phase
- Implementation Phase

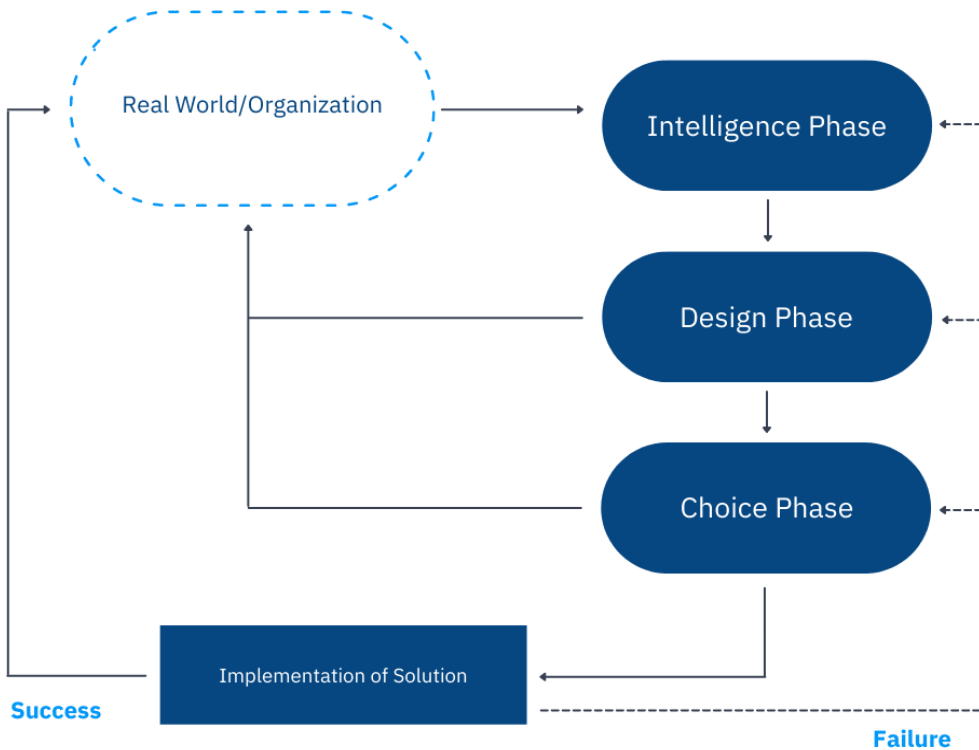


Figure 3.5: Phases of the Decision-Making Process

It's important to note that first Simon's model contained only three phases—implementation phase was added later. Also, we could add a fifth phase—monitoring, but that can be viewed as an input for the intelligence phase.

3.4.1.1 Intelligence Phase

Firstly, the decision-making process starts with the intelligence phase. In the first phase, decision makers examine reality and try to identify problems or opportunities correctly. This phase is not only related to the Simon's decision-making process, but also to other fields and other methodologies. For example, we like to practice Lean Startup methodology

which emphasizes the importance of right problem definition before building anything (product or business).

Additionally, one of the pillars of digital transformation is the data. Organizations need to become data driven. That means proper usage and implementation of Business Intelligence(BI) systems. Business Intelligence implementations are considered successful only if you have clear business needs and see real benefits from it. Business Intelligence is not just about data. It should be connected with organizational goals and objectives

Therefore, intelligence phase includes actions like:

- Defining organizational objectives
- Data collection
- Problem identification and classification

The intelligence phase can last really long. But, since decision-making process starts with this phase, it should be to be done properly. This is a key ingredient in every business success.

3.4.1.2 Design Phase

The main goal of the design phase is to define and construct a model which represents a system, by defining relationships between collected variables. Once we validate the model, we define the criteria of choice and search for several possible solutions for the defined problem (opportunity). We wrap up the design phase by predicting the future outcomes for each alternative.

3.4.1.3 Choice Phase

In this phase we are actually making decisions. The end product of this phase is a decision. Decision is made by selecting and evaluating alternatives defined in previous step. If we are sure that the decision we made can actually be achieved – we are ready for the next phase.

3.4.1.4 Implementation Phase

All the previous steps we’ve made (intelligence, design, and choice) are now implemented. Implementation can be either successful or not. Successful implementation results with a solution to the defined problem. On the other hand, failure brings us back to the earlier phase.

We described Simon’s model which, even today, serves as the basis of most models of decision- making process. A process is described as a series of events that precede final decisions. It is important to say that, at any point, the decision maker may choose to return to the previous step for additional validation.

Even Simon’s model was sometimes criticized as being general, that is why we need to be aware of the importance of decision-making. This model is a concept, a framework of how organizations and managers make decisions.

3.4.2 A Cyclic Process of Intelligence Creation

Business Intelligence is the process of analyzing unprocessed data and turning it into knowledge that the company can use to make decisions. Organizations can better understand their marketing strategies, earnings and losses, effective development and management, and market trends and consumer behavior by using business intelligence.

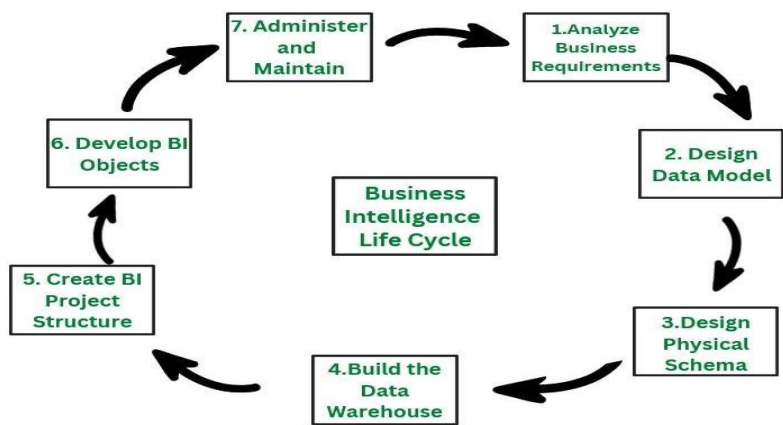


Figure 3.6: Business Intelligence Life Cycle

3.4.2.1 Phase 1: Analyze Business Requirements

The first step in the Business Intelligence life cycle is to analyze the business requirements. The user identifies the business requirements in order to determine the type of analysis that the user then needs to perform. Identifying the requirements, let the user decides the further action to be performed.

For example, any retail company can analyze the sales data to figure out the products that are top- selling and the products that least sell.

3.4.2.2 Phase 2: Design Data Model

Once the requirements are identified the user needs to design the logical model according to the requirements.

This logical model help the user to analyze the relationships that exist within the data entities. For example, for any retail company, the data model consists of products, their customers, and the sales data.

3.4.2.3 Design the Physical Schema

Once the logical model is prepared the next step is to design the physical schema using the data model. The physical schema describes the structure and the content of the data warehouse.

For example, in any retail company, physical schema consists of sales-related facts, product- customer relationships, and the sales transactions

3.4.2.4 Build the Data Warehouse

Once the logical and physical schema is designed, the next step is to build the data warehouse. The design of a data warehouse depends on the physical and logical schema. After the design of the data warehouse, the data and the content from the source system are loaded into the data warehouse for further steps.

For example, for the retail system, designing the data warehouse consists of developing a database that would store the details of customers, products, and other requirements for the business.

3.4.2.5 Phase5: Create the Project Structure (Metadata)

The next step after designing the data warehouse is to create a project structure also known as metadata. With the help of this created project structure, the mapping of the tables and data in the data warehouse is easier. Creating the project structure describes the further steps and types that need to be implemented.

For example, the project structure of the retail company consists of the attributes of the data, the design, and the working flow of the system. This project structure or metadata gives a brief idea about the working of the system.

3.4.2.6 Develop the BI Objects

The next step is to develop the BI objects such as metrics, attributes, dashboards, reports, and facts. This step consists of developing the reports and dashboards that can be used to analyze the data in the data warehouse. For example, the retail company can develop reports and statistics charts that can describe the profit and loss margins.

3.4.2.7 Administer and Maintain the Project

The last step is to administer and maintain the project continuously as it undergoes changes. The project needs to be monitored to maintain the changes, security, and performance of the system.

For example, the retail company needs to monitor the reports and statistics accordingly to increase the profit of the sales.

3.4.3 Styles Of BI

This paper looks at the historical development of BI applications and BI technology, and concludes that five common Styles of BI have evolved

during the past decade - each style representing a certain characteristic usage and function by end users. These 5 Styles of BI are:

- **Enterprise Reporting** - Broadly deployed pixel-perfect report formats for operational reporting and scorecards/dashboards targeted at information consumers and executives.
- **Cube Analysis** - OLAP slice-and-dice analysis of limited data sets, targeted at managers and others who need a safe and simple environment for basic data exploration within a limited range of data.
- **Ad Hoc Query and Analysis** - Full investigative query into all data, as well as automated slice and- dice OLAP analysis of the entire database - down to the transaction level of detail if necessary. Targeted at information explorers and power users.
- **Statistical Analysis and Data Mining** - Full mathematical, financial, and statistical treatment of data for purposes of correlation analysis, trend analysis, financial analysis and projections. Targeted at the professional information analysts.
- **Alerting and Report Delivery** - Proactive report delivery and alerting to very large
- Populations based on schedules or event triggers in the database. Targeted at very large user populations of information consumers, both internal and external to the enterprise

3.4.4 The Value of Business Intelligence:

As the growth of e-commerce continues to saturate the market, the importance of BI becomes even more apparent than ever. Business owners have to make smart decisions regarding how they wish to see their marketing spend, as anything a consumer now wants is only a click away.

There are many compelling reasons behind this: improving performance, boosting sales, long-term customer relations, etc., – all built through better customer experiences. Integrating business intelligence into your

operations helps your company by delivering value through the following ways:

3.4.5 Effective Decision-Making

The sole reason behind the implementation of BI is to convert raw company data into analyzable, well-structured insights that enable the organizational executives to implement strategic decision- making.

Great business intelligence means having all your business data in a unified dashboard to include all the relevant data from different areas such as finance, sales, and many others – all that aim to provide a holistic view of the business. At the end of it all, business decisions will be made based on facts rather than assumptions.

3.4.6 Sales & Marketing

Incorporating BI data allows a company to boost a current marketing campaign's performance, increasing sales in the long run. Through BI, the sales department can get the right tools to help measure consumer trends through improved visibility. You can also get specialized features that help track and measure sales & marketing campaigns, providing the relevant data that would support future marketing initiatives.

3.4.7 Customer Experience

Business intelligence is also helpful by delivering the necessary information to help companies understand how their customers interact with their business. Data accuracy is improved when one can access all customer information from a single dashboard, enabling the businesses to enhance customer support, engagement, and experience.

Moreover, BI helps analyze customer insights to improve the targeting and segmentation of the different categories of customers. This helps identify which resources need to be applied for the businesses to attract only valuable customers to achieve particular goals.

3.4.8 Boosting Productivity

The automation of routine tasks through BI helps an organization to refine its operational processes. While there are many project management software that can do the same, Business intelligence introduces ways to seamlessly improve inventory control and reduce inefficient constrictions within the organizational structure. Easy-to-access centralized data also cuts down on the administration time and efforts, simultaneously boosting data integrity and productivity.

3.4.9 Data Accuracy and Compliance

The centralized nature of BI data boosts transparency in organizational operations. As it offers you a holistic view of your business and customers, BI also seeks to expose the errors that might lead to lots of downtime and wasted resources.

Companies are now tasked with heavier responsibilities when it comes to the protection of personal data. There is a keen focus on organizations to adhere to data protection regulations should they wish to store customer information. Implementing BI tools ensures that businesses can address the issue of data governance and integrity.

3.4.10 Tips for Choosing the Right Business Intelligence Tools

Going for the right BI tools is as important as collecting the data itself. There are many businesses intelligence software that one can go for, but how will you pick the one that best fits your organizational needs?

Here are three important factors to consider:

3.4.10.1 Integration

Before implementing any BI tools, every company has its own reporting processes. Ensure that the tools you go for are easily integrated with

existing structures and can easily incorporate the data received from different sources.

3.4.10.2 Identify Your Immediate Goals

One of the first steps to undertake before getting a business intelligence tool is identifying the goals you wish to achieve. Setting parameters from the initial stages enables you to glean the right data.

3.4.10.3 User-friendliness

Before settling on a tool, ensure that it has an intuitive interface that's easy to use by all the approved users. Nothing is worse than any system that's clunky and hard to operate. This means that the BI tool has to be easy to access, operate, and translate the information it provides.

UNIT - IV

4.Advanced BI

A solution to better manage your business

Advanced BI is a management information solution that meets business requirements and provides reports and interactive dashboards that facilitate the quick and assertive interpretation of current and historical information.

The executives are provided with indicators to monitor production metrics and control profitability, guiding financial institutions towards profitable growth and the achievement of goals.

Key Functionalities

1. Business Indicators

Quick access to management and profitability indicators calculated automatically by pre-defined processes.

2. Data Templates

Data organized according to the analyses to be carried out. Wide range of templates with customizable components depending on the organization.

3. Availability of Historical Information

Historical information allowing for business monitoring.

4. Dashboards and Reports

Interactive and flexible reports and Dashboards with easy-to-read graphical overviews. Allows end-users to create reports and Dashboards.

Advantages

1. Reduced time-to-market
2. Easy to implement solution that provides a simple analysis layer.

3. Better Decisions
4. It offers information in the appropriate format for reflection and decision-making.
5. Autonomy
6. Users are able to perform exploratory data analysis and make their own findings
7. Versatility

Flexible and customizable solution, allowing the production of analysis instruments adapted to each organization and to the different user profiles.

4.1 Big Data and BI

Business intelligence and big data sound similar but they are not the same.

In general, business intelligence (BI) refers to structured and readily usable information that impacts profitability and competitive advantage. On the other hand, big data — as you might expect — refers to enormous heaps of digital information scattered all over the place, with practitioners typically focusing more on unstructured data.

Both fields involve crunching information to generate insight and make things happen. But they differ in terms of the volume and nature of information each focuses on, as well as the tools they use to process data. Their specific aims and outcomes sometimes coincide but not always. You don't need big data to build a decent business intelligence system, for example, but big data dramatically enhances BI capabilities at scale.

Business intelligence helps companies make smart, revenue-boosting decisions. Enterprises purposely use BI to improve process, planning, and profits. Meanwhile, big data can perform the same functions but can do so faster and at scale. Big data also helps organizations achieve many

remarkable feats: design the perfect bra, fight cancer, protect national security, and enhance athletic performance. Just to name a few.

Over the years, think tanks and business leaders have attempted to update the meaning of big data and business intelligence as economic and technological contexts evolve. Here are two of the most widely-cited definitions:

“Business intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance.”

“Big data is a term applied to datasets whose size or type is beyond the ability of traditional relational databases to capture, manage, and process the data with low-latency. And it has one or more of the following characteristics – high volume, high velocity, or high variety. Big data comes from sensors, devices, video/audio, networks, log files, transactional applications, web, and social media - much of it generated in real time and in a very large scale.

Business intelligence practitioners generally handle structured data while big data professionals feel at home processing humongous volumes of unstructured data at lightning speeds. Both can provide the fourth and most important V (i.e., value) in the form of descriptive, predictive, and prescriptive analysis/reporting.

Finally, each field uses a different set of enabling technologies, with the data science toolbox generally more sophisticated than the one for BI, although they might share common tools such as SQL and Python. Both BI and Big data goal is to help the business to make good decisions by analyzing the huge datasets to expand the business and optimizing the cost.

This data analysis not only enables decision making but also involves an active part in the development of strategies and methods that make sure

the success of organizations. This data analysis can be called “Business Intelligence”, whereas “Big Data” is a relatively new term for Business intelligence.

Since the times of BI, the volumes of data sets become incredibly large, the best example we can consider is social media. As the result, more effort and strategies should be applied to tackle with them and make them useful for successful business.

Business Intelligence helps in finding the answers to the business questions we know, whereas Big Data helps us in finding the questions and answers that we didn't know before.

Although Business Intelligence and Big Data are two technologies used to analyze data sets to help organizations in the decision-making process, there are differences present between them. They both differ in the way they analyze the data.

Business Intelligence is based on the principle of combining all business data sets into a central server, this data will be analyzed in offline mode, after saving the information in a platform or environment called Data Warehouse. The data sets are structured in a relational database with additional indexes and forms of access to the tables in the warehouse.

Whereas in the Big Data environment, data is stored on a distributed file system (e.g. HDFS), rather than storing on a central server. Data will be distributed across the worker nodes for easy processing. Distributed File System is much safer and flexible.

BI solutions carry the data to the processing functions, whereas Big Data solutions take the processing functions to the data sets. Since the analysis is positioned around the information (Data), it is simpler to handle larger amounts.

BI solutions are more towards the structured data, whereas Big Data tools can process and analyze data in different formats, both structured and unstructured.

Big Data solutions can process the historical data and also data coming from real-time sources, whereas in Business Intelligence, it processes the historical data sets.

Big Data technology uses parallel processing concepts (Map reducing algorithm), which improves the speed of analyzing and processing the data sets by distributing jobs into several parallel execution processes, at the end the results are combined and shown, this makes analyzing the large volumes easier.

4.1.1 BI Big Data

Table 1: BI Big Data

Purpose	The purpose of Business Intelligence is to help the business make better decisions. BI helps in delivering accurate reports by extracting information directly from the data source.	The main purpose of Big Data is to capture, process, and analyze the data, both structured and unstructured, to improve customer outcomes.
Ecosystem/Components	Operation systems, ERP databases, Data Warehouse, Dashboard etc.	Hadoop, Spark, R Server, Hive, HDFS etc.
Tools	Below is the list of tools used for Business Intelligence. These tools enable a business to collate, analyze and visualize data, which can be used in	Below is the list of tools used in Big Data. These tools or frameworks store a large amount of data and process them to get insights from data to

	making better business decisions and to come up with good strategic plans.	make good decisions for the business.
	- Tableau	- Hadoop
	- Qlik Sense	- Spark
	- Online analytical processing (OLAP)	- Hadoop
	- Sisense	- Polybase
	- Data Warehousing	- Presto
		- Cassandra
		- Plotly
		- Cloudera
Characteristics / Properties	Big data can be described by some characteristics such as Volume, Variety, Variability, Velocity, and Veracity.	Below are the six features of Business Intelligence: Location intelligence, Executive Dashboards, “What if” analysis, Interactive reports, Metadata layer, and Ranking reports.
Benefits	Below is the list of benefits of Business Intelligence:	Below is the list of benefits of Big Data:
	- Helps in making better business decisions	- Better Decision making
	- Faster and more accurate reporting and analysis	- Fraud detection
	- Improved data quality	- Storage, mining, and analysis of data

	- Reduced costs	- Market prediction & forecasting
	- Increase revenues	- Improves the service
	- Improved operational efficiency	- Helps in implementing the new strategies
		- Keep up with customer trends
		- Cost savings
		- Better sales insights, which help in increasing revenues etc.
Applied Fields	Social media, Healthcare, Gaming Industry, Food Industry etc.	The banking sector, Entertainment, and social media, Healthcare, Retail, and wholesale etc.

4.2 Social Networks

Social Networks and Big Data: the best strategy for getting to know your customers big data fits into social media networks like a glove.

This is because the volume of daily data produced in these virtual environments is a real gold mine for companies prepared to prospect for it.

We will reflect more on this thought in this article. Keep reading to understand how you can benefit from the combination of Social Networks + Big Data.

Social Networks: the gold mine of data

In the age of Digital Transformation, social media is ubiquitous. To give you an idea, Facebook has about 2.5 billion users across the globe.

The fact is: there are social networks for all tastes, from those that like entertainment, like TikTok, to the most professional ones, like LinkedIn. Thus, it is practically impossible to find a person who has access to the internet and does not interact on these platforms.

And the amount of data from social networks continues to explode. About 510,000 Facebook comments are written per minute.

Another 347,000 tweets are sent in the same timeframe.

However, many companies still do not know what to do with this information – or even measure it correctly.

Superficial monitoring and the lack of an outside opinion are two reasons for this. The data should help companies gain in-depth knowledge about their customers.

To think about this more clearly, keep in mind that data needs to answer the following: Who are my customers besides their email and mailing addresses?

What are they interested in?

What concerns them, and how does my company relate to their interests, desires, and concerns?

→This whirlwind of information needs to be adequately captured, organized, processed and analyzed. And Big Data is the tool for the job.

→You might also like: Business Analytics – The data age has begun! Big Data and unlimited data inputs

Let's quickly recall the concept of Big Data.

It is a set of “high volume, high speed and/or high variety information assets that require innovative and economical ways of processing, allowing for improved insights, decision making, and process automation.”

4.2.1 In Other Words, Big Data is Characterized by:

Large amounts of data being analyzed in real-time External sources – such as social networks – are consulted

A good Big Data strategy has to manage and analyze the exponential volume of information from social networks. This can only be done with the proper people and tools.

It's not very common, but it is urgent. Organizations miss out on many opportunities only by superficially evaluating social media data – or just using it to measure their branding efforts and distribute content.

4.2.2 How to Create a Good Strategy

1. Collect data from networks and social media to analyze them in an automated and systematic way. The procedure usually starts with the identification of the relevant platforms and information.
2. The data is then collected and compressed.
3. The extracted information can be evaluated according to several aspects.

→The results are reports and indexes that support decision-making processes.

By working with infinite data inputs using Big Data methodologies, companies get a better view of their customers' opinions, allowing them to develop their image or brand better.

In addition to marketing, customer service is also supported by the results of this strategy. For example, a detailed analysis of the discussion facilitates the identification of weaknesses.

4.2.3 Measuring Results

Information used for social media analysis can be Likes

Tweets Shares Links

Facebook comments on posts, forums, and other platforms

Advanced Big Data platforms and methods go beyond these metrics to be more “objective.” They can even perform analysis of feelings and other intrinsic characteristics that are not verbalized adequately by users.

4.2.4 Social Networks and Big Data: Get the Strategy Right for an Incredible Experience

Big Data makes social network analysis more focused on predictions and the effects of current strategies in the future. Predictive analytics is gaining popularity among marketing, sales, and competitive intelligence analysts, among others, and Big Data has a significant role to play.

Professionals dedicated to this strategy are not only able to predict consumer behavior. Still, they can also use their data analysis conclusions in various other methods of marketing, sales, logistics, etc. It is in this way that the path towards surprising innovations is carved.

Now, where to start? Below we have a checklist that should help you stay on track. It should help you get to know the current and potential customers in-depth to create incredible experiences and enhance your company’s results.

Identify what you want: what are your Big Data and Social Media strategy going to achieve? Take advantage of a proven Big Data strategy.

In this case, it can be:

Social analysis: briefly refers to the measurement of non-transactional data on various social media and websites, such as Facebook, Twitter, and Instagram. This strategy is based on the analysis of conversations and criticisms that arise on these platforms. It presents three main analytical advantages: awareness, engagement, and word of mouth. Social analysis is also useful in predicting the peak demand for specific products.

Data exploration: an approach that makes heavy use of research and data mining to find solutions and correlations that are not easily detectable with internal data. It is currently used by

companies focused on Inbound Marketing to generate insights into potential customers' behavior on the site. In addition, it helps to identify new data segments and bring information about customer behavior and preferences.

Invest in the right technological tools: you will need to invest in Analytics and data visualization tools, but you may also need Business Intelligence and CRM. It all depends on the objectives of your strategy.

Identify the necessary changes in the infrastructure: if the company's old data was stored in traditional formats, this might not facilitate the execution of complex algorithms and analysis. Different departments (marketing, sales, board, etc.) may need integration to collect and optimize data in a more usable format. Integration between various departments is essential to implement change at scale. If your existing infrastructure is not connected correctly, you will need to prepare for changes.

Work on the customer experience as a guide for your strategy: the primary use of Big Data is to generate insights that can help company's better serve their customers. Customer-oriented marketing is a new way to approach the market and generate revenue. At the end of the day, you need to communicate to your customer that you are there to solve a problem and not just to make a profit. Big Data provides information about the customer's mindset that can improve and even change current marketing practices. As we've already said: Another thing you need to focus on is creating a fine line between data collection and abuse of privacy – your customers must feel that they are not spied on. As much as you work with cloud computing tools, it will be necessary to facilitate tracking everything that is done with data duly authorized by users.

Be agile: when implementing disruptive technologies, many obstacles can arise, including some that no one imagined initially. You need to adjust your budget, stakeholders, and ideologies based on the circumstances and ideas you gather. It is best to start with a high-level plan and make the necessary changes. In the end, you can propose an action plan that is nowhere near the initial idea, but it will be worth it.

→Finally, recognize that you will initially need to seek outside help. A good consulting company can help you to structure your Social Networks and Big Data strategy. Do not hesitate to find a company that already has a background in the subject and qualified talents.

4.3 Mobile BI

Mobile is changing the way that we work and consume information. We are hardly “Offline” anymore, which means that users need to have access to their analytics to make decisions faster based on insights on the fly. In other words, dashboards need to be consumed by your mobile device and be able to provide end-users with data information that they can easily absorb. This becomes tricky due to the amount of real estate you have on a phone vs on your desktop.

The definition of mobile BI refers to the access and use of information via mobile devices. With the increasing use of mobile devices for business – not only in management positions – mobile

BI is able to bring business intelligence and analytics closer to the user when done properly. Whether during a train journey, in the airport departure lounge or during a meeting break, information can be consumed almost anywhere and anytime with mobile BI.

Mobile BI – driven by the success of mobile devices – was considered by many as a big wave in BI and analytics a few years ago. Nowadays, there is a level of disillusion in the market and users attach much less importance to this trend.

4.3.1 What Is Mobile Business Intelligence (BI)?

You're probably familiar with the term business intelligence (BI), but if not, business intelligence refers to the technology-driven process of analyzing data and providing actionable insight resources that can be used to make intelligent business decisions. Or is otherwise known that BI is about having the right data at the right time to make the right decision.

Consequently, Mobile Business Intelligence (BI) refers to the ability to access and perform BI-related data analysis on mobile devices and tablets. It makes it easier to display KPIs, business metrics, and dashboards. Mobile BI empowers users to use and receive the same features and capabilities as those in the desktop / app-based BI software.

4.3.1.1 Why Is Mobile BI Important?

Businesses these days possess an abundance of data. In this fast-paced environment, everyone needs real-time data access to make data-driven decisions anytime and anywhere.

The number of organizations using mobile apps like SaaS applications for critical business processes is increasing every day. Whether you are a CEO, a sales person, digital marketer, department manager, or another employee, mobile BI can help you increase productivity, improve the decision-making process, and overall boost your business.

4.3.1.2 Advantages of Mobile BI Accessibility

Having company' insights at your fingertips is the most valuable advantage of mobile BI. You are not limited to one computer in one location, but instead, you can access important data information on your mobile device at any time and from any location. Having real-time data insights always available helps improve the overall productivity of your daily operations.

4.3.1.3 Improved Decision-Making

Mobile BI apps speed up the decision-making process. When decisions must be made, or when actions must be taken on the spot and at the moment, mobile BI provides up-to-the-minute insights based on data to help users when they need it the most.

4.3.1.4 Stay Ahead of Your Competitors

Access to real-time data helps in seeing business opportunities sooner, reacting to changes in market conditions in a timely manner, as well as increasing the opportunity to up-sell and cross-sell. Deploying a mobile BI solution makes you more flexible and more adaptable to business shifts.

4.3.1.5 Mobile BI solutions

Every mobile device can be used to display and work with data. But of course, there are some significant differences – the size of the screen is obviously difference #1. For example, some data visualizations require more screen space, while others can easily fit in the width of a smartphone or a tablet.

There are a couple of ways to implement content on mobile devices with those being the most common:

Web page – Every mobile device includes a web browser that can access almost any web page. Depending on some factors, the quality of the accessed page could be terrible, acceptable, or great. In this scenario, BI developers need to check how their content renders on mobile devices when creating reports and other data visualizations. This means that they should design the application specifically for mobile use.

HTML5 site – Same as the web page, but with some improvements. HTML5 enables RIA (Rich Internet Application) content to be displayed across all types of mobile devices without relying on proprietary standards. The advantages of HTML5 are functions such as zooming,

double-tapping, and pinching, as well as that the user doesn't have to install the app on its device to be able to use it.

Native app – Those are web applications that can be downloaded and installed on mobile devices. The application's software is tailored to the OS of the mobile devices. It is the most difficult and most expensive solution for manufacturers to support mobile BI, but it enables interactive and enhanced use of analytics content.

4.4 Emerging Trends

Nowadays, we produce more data daily than ever before, and decision-making is an all-important aspect of business and operations management. While some business people rely on their intuition, savvy entrepreneurs use solid data as a basis. But contrary to popular belief, analytics don't follow your experience-based instincts. Data can either support intuition or not, but it helps you determine the next steps and directions.

By 2025, the world's data will have exceeded 180 zettabytes coming from over 41 billion connected devices.

Data management is no longer an ideal; it's an expectation in today's business world. Over the last few years, spreadsheets have finally given way to informative real-time data visualizations and interactive business dashboards.

Business intelligence (BI) helps companies collect, analyze, present, and integrate data to receive valuable information for business growth. Over time, the business intelligence industry has morphed into a billion-dollar enterprise. More and more companies undergo digital transformation in response to industry demands.

The future of BI is being defined, and everyone who wants to be data-driven wants to keep an eye on new trends. Let's have a look at current and future BI trends.

The Current Business Intelligence Trends in 2022

The amount of information in all industries is growing fast. Today, organizations are continuously generating volumes of data as the speed, variety, and complexity of data are also increasing. But only those companies that know how to handle and use data properly can find useful insights and create real profit.

Let's have a look at top Business intelligence trends. In 2022, companies are still working to position themselves for the long term: addressing the root causes of their problems, e.g. data quality, and working to create a data-driven culture holistically. Data management has become more important, too. Data discovery and visualization, plus self-service analytics, are important to organizations.

4.4.1 DQ and MD Management

Qualitative data is at the core of the strategic, tactical, and operational management of every company. Data from The BI Survey shows that companies have always strived to provide a high level of data quality. In 2022, when information is increasingly becoming a vital factor of production, the need for high-quality data is growing, too. All in all, poor-quality data results in false facts and wrong decisions at companies.

Rapidly growing volumes of information from internal and external sources make databases keep expanding. Why is using qualitative data considered one of the latest BI trends?

- Organizations must be able to respond as flexibly as possible to dynamically changing market demands.
- Organizations risk missing out on valuable business opportunities, whereas competitors can take advantage of this situation.
- Organizations can only unlock the full economic potential of their data if the underlying data is well managed and delivered in high quality.

4.4.2 Creating a Data-Driven Culture

Recently, many well-known outlets have written about this new trend in business intelligence. The concept of a data-driven culture at a company envisions information as a key resource for implementing ideas in every department. The key goal is to empower every employee to actively use data to bolster their daily work.

This way, they can use the company's potential and data to make effective decisions, analyze current trends, and make critical competitive advantage more visible. The ultimate goal is to create a cultural framework that helps the company's members collaborate to move data to the centre of decision-making: from the owner to the employee (marketer, financier, manager, HR, etc.)

4.4.3 Data Governance

Data governance is a recent trend in business intelligence and is vital to enabling organizations to respond quickly. Data governance encompasses the people, processes, and technology required to manage and protect the company's data assets. For that reason, corporate information must be clearly understandable, correct, complete, trustworthy, secure, and available for use.

Here are the main objectives:

- Minimizing risks
- Setting up rules for information use
- Improving internal and external communication
- Making data more valuable
- Saving time and costs

4.4.4 Data Discovery and Visualization

Data discovery is a business user-centric process designed for discovering insights by visually navigating data or using guided advanced

analytics. This process can get divided into three steps: data preparation, visual analysis, and managed advanced analytics.

The goal is to leverage the full value of information not only to improve decision-making but also to impact business process optimization and support new business models directly.

Interactive and novel types of visualizations enable decision makers to instantly see key trends and identify outliers. Users are better able to find insights and detect outliers when data gets presented as charts and graphs on a single page rather than hidden in tables that take up multiple pages.

2023 Future Business Intelligence Trends: Predictions You Want to Be Thinking About.

4.4.5 Real-Time Data & Analytics Gain More Traction

Recently, organizations have seen a growing need for real-time data, and this trend will continue in 2023. Up-to-date information is becoming more important than ever before, and as the world has changed, companies also need to adapt. Since the onset of the pandemic, we have seen the need for accurate, real-time updates is critical to developing the right strategies to respond to adverse situations like that. Besides, interactive dashboards will help companies immediately access relevant information related to their business and respond if any potential issues arise.

4.4.6 Artificial Intelligence

Artificial intelligence (AI) and machine learning are future trends in business intelligence. AI is a branch of computer science to make machines do what sophisticated human intelligence would normally do. AI and ML revolutionize how we interact with our analytics and data management.

4.4.7 Data Automation

Business intelligence topics would be incomplete without data automation (analysis). This emerging trend is all about activities where businesses automate as many processes as possible using multiple tools

and technologies such as AI, machine learning, low-code and no-code tools, etc.

4.4.8 Eye-Opening Business Intelligence Statistics in 2023

According to business intelligence statistics, any business effectively using information extracted from market data will surely benefit. Of course, big companies in any industry know this. We understand the value of BI to your company's growth. All in all, we have hand-picked some of the most relevant business intelligence statistics for 2022.

Here are some top statistics in business intelligence to help you understand the importance of it all:

- By 2023, over 33% of large-sized businesses will have practised decision intelligence.
- Data analytics makes decision-making 5x faster for businesses.
- Self-service BI is crucial to 60% of research and development departments.
- The global BI market will have grown to \$33.3 billion by 2025.
- Most businesses adopted BI in 2020.
- The top three latest trends in BI are data visualization, data quality management, and self-service business intelligence (BI).
- Seven out of ten businesses rate data discovery as very important.
- 71% of buyers prefer Microsoft Azure as their number one vendor.
- Only 37.8% of firms report building a data-driven organization.
- Only 26.8% consider themselves successful in building a data culture.
- Reporting, dashboards, data integration, data warehousing and data processing are among the top five technologies and initiatives of strategic importance for BI.

4.4.9 How Can Business Intelligence (BI) Help Businesses during the Crisis

The crisis has made one point obvious: the traditional wait-and-watch strategy for optimizing operations won't work. Companies must make data-driven decisions to devise strategies to empower employees, navigate through the crisis smoothly to earn clients' trust and loyalty, stabilize operations, and optimize costs—none of which is possible without BI.

Let's have a look at these top trends in business intelligence in more detail:

4.4.10 Discover Opportunities for Boosting Growth

The main business intelligence industry trend is removing barriers between siloed data sources. BI can help you match information from thousands of datasets and provide forecasts that are not only accurate but also timely. Such information can enable historical data analysis, helping companies uncover growth opportunities—even as sectors plunge.

4.4.11 Find Areas for Saving Costs

By identifying unnecessary costs and redundant processes, BI can list business areas to yield the best results in cost savings while highlighting avenues of business that require investments, such as digital workspaces, business continuity, or remote workforce.

4.4.12 Make Informed Decisions to Minimize Risks

By connecting to a wide array of data sources, BI can provide the figures, information, and facts to help companies make informed business decisions and mitigate potential risks.

4.4.13 Speed Up Normal Operations

BI provides a fast and simple overview of data from across the company—that can easily go unnoticed. By incorporating sophisticated data analysis tools and technologies, BI can maximize correlations, patterns, and trends and helps employees make data-driven decisions.

Learn more about “How to be a data-driven decision-making company in times of crisis”

4.5 Description Of Different, BI-Tools (Pentaho, KNIME)

Business intelligence tools (BI tools) are all about helping you understand trends and deriving insights from your data so that you can make tactical and strategic business decisions. They also help you identify patterns in the mountains of data your business builds up. But the question is... how do you know which business intelligence tool is best for your business? And which tools will help you achieve your online goals?

Why use business intelligence tools?

For starters, data discovery, which used to be limited to the expertise of advanced analytics specialists, is now something everyone can do using these tools. And not only that, these tools give you the insights you need to achieve things like growth, resolve issues that are urgent, collect all your data in one place, forecast future outcomes and so much more.

In this article we will explain the top 15 Business Intelligence tools in 2021 and hopefully put you on the right path towards selecting you're a tool fit for your business.

Keep in mind: these business intelligence tools all vary in robustness, integration capabilities, ease-of-use (from a technical perspective) and of course, pricing.

4.5.1 Pentaho BI What is Pentaho BI?

Pentaho is a Business Intelligence tool that offers a wide range of data solutions to its customers.

The main features of this tool are reporting, data integration, data mining, data analysis that account for the improvement of the business. A Pentaho suite enhances the overall performance of the business by

generating informative reports in varied formats like text, XML, HTML, CSV, Excel, PDF, etc.

4.5.1.1 Why should you go for the Pentaho BI?

Pentaho Business Intelligence tool suite is a set of tools that offers several benefits to businesses at an affordable cost and fast speed in terms of data management. Compared to other BI tools like SAP, SAS BIA, and IBA, Pentaho BI offers exceptional technical support to the customers. It is highly scalable and offers large volume support to process data up to billion terabytes in size.

The scope of the Pentaho BI suite is vast supporting all kinds of data and data sources that furnish limitless visualization options. It supports an unlimited amount of data be it big data or existing data in the business IT. The tool works on several core engines that work independently and is administered by a dedicated community. It can be used across different platforms that process hybrid data (text, graphics, visuals GIFs, etc.) like mobile apps, cloud apps.

4.5.1.2 Pentaho Features

Pentaho BI offers multiple features for the smooth workability of the business, such as:

- High-end data analysis through well-defined ETL (Extract, Transform, Load) capabilities
- Expertise in products across varied domains
- Comprehensive report designer taking care of business needs
- Additional reports or sub-reports along with the main detailed report
- High scope for newer additions and updates
- 24*7 technical support by the Pentaho Community
- Unmatched reporting and query handling capabilities
- Amplified functionality and efficient systems
- Short integration TAT

- Exceptional data source accommodability with high runtime metadata support

4.5.2 KNIME

What is KNIME?

KNIME is an open-source end-to-end data analytics solution. It utilizes visual workflows with drag-and-drop functionality and thousands of nodes to lessen the data analytics learning curve data, with more than 1,800 prebuilt default workflows for streamlined setup.

It allows for data ingestion, preparing, cleansing, analyzing and visualizing. It can be scaled for deeper analytics through integrations with sophisticated data modeling capabilities. It can be hosted on premise or in the cloud through Microsoft Azure.

4.5.2.1 Benefits and Insights

Key differentiators & advantages of KNIME

- **Open-Source:** Join a network of thousands of users, enabling collaboration and support. The source code is free to download and access.
- **Free To Use:** Save money by getting access to all of the platform's features for free. Licensed productivity and collaboration extensions are available at a cost.
- **Increased Business Intelligence:** Get digestible, actionable data to make informed business decisions. Aggregating large datasets into predictive and prescriptive models via comprehensive visualizations and summary statistics gives users projections for the best course of action.
- **Scalable:** Obtain access to big data by scaling up the project in-platform. Integrations to distributed and multi-threaded data processing allow projects to grow.

- **End-To-End Analytics:** It is capable of handling some tasks from start to finish without integrations. Additional integrations may be required for increasing scale and completing more sophisticated analytics.

4.5.3 SAP Business Objects

SAP Business Objects is a business intelligence software which offers comprehensive reporting, analysis and interactive data visualization. The platform focuses heavily on categories such as Customer Experience (CX) and CRM, digital supply chain, ERP and more. What's really nice about this platform is the self-service, role-based dashboards it offers enabling users to build their own dashboards and applications. SAP is a robust software intended for all Website: www.sap.com

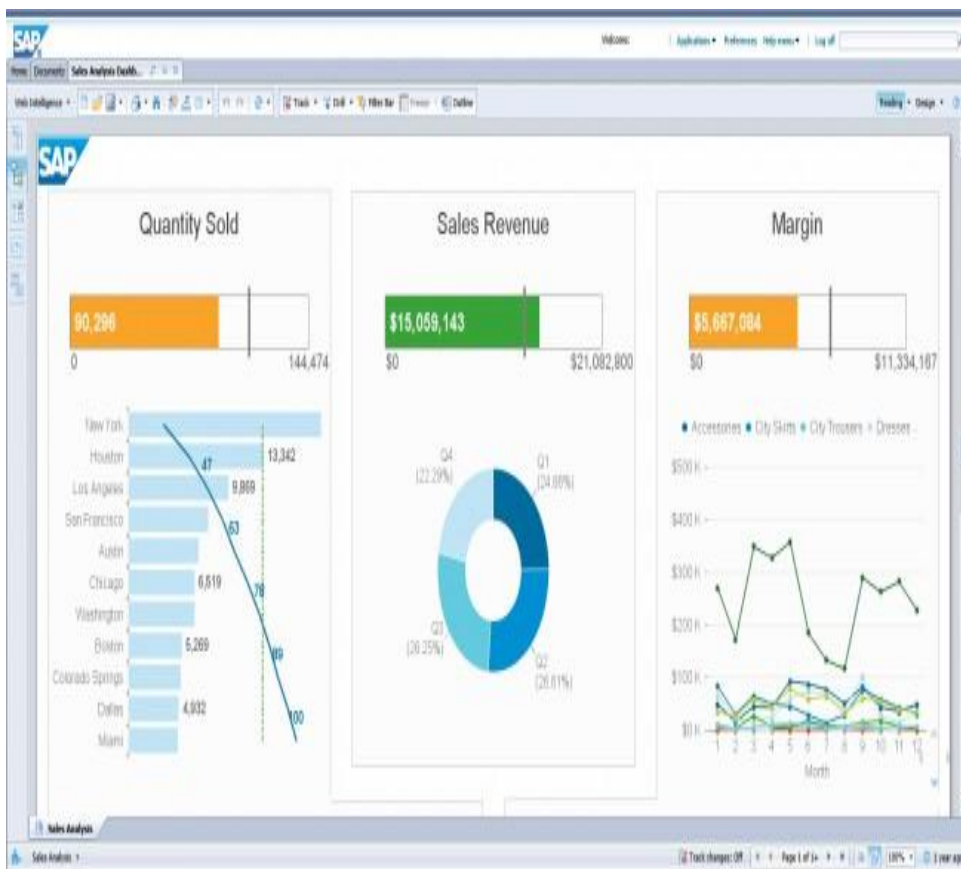


Figure 4.1: SAP Business Objects

4.5.4 Dapine



Figure 4.2: Dapine

Dapine is an all-in-one business intelligence platform that facilitates the complex process of

data analytics even for non-technical users. Thanks to a comprehensive self-service analytics approach, datapine’s solution enables data analysts and business users alike to easily integrate different data sources, perform advanced data analysis, build interactive business dashboards and generate actionable business insights.

Website: www.datapine.com

4.5.5 Microsoft Power BI



Figure 4.3: Microsoft Power BI

Microsoft Power BI is a web-based business analytics tool suite which excels in data visualization. It allows users to identify trends in real-time and has brand new connectors that allow you to up your game in campaigns. Because it's web-based, Microsoft Power BI can be accessed from pretty much anywhere. This software also allows users to integrate their apps and deliver reports and real-time dashboards.

Website: www.powerbi.microsoft.com

4.5.6 Looker



Figure 4.4: Looker

Data discovery app, Looker is another business intelligence tool to look out for! Now part of Google Cloud, this unique platform integrates with any SQL database or warehouse and is great for startups, midsize-businesses or enterprise-grade businesses. Some benefits of this particular tool include ease-of-use, handy visualizations, powerful collaboration features (data and reports can be shared via email or USL as well as integrated with other applications), and reliable support (tech team).

Website: www.looker.com

4.5.7 Clear Analytics



Figure 4.5: Qlik Analytics

This is for all of the Excel-lovers out there...This BI tool is an intuitive Excel-based software that can be used by employees with even the most basic knowledge of Excel. What you get is a self-service Business Intelligence system that offers several BI features such as creating, automating, analyzing and visualization your company’s data. This solution also works with the aforementioned Microsoft Power BI, using Power Query and Power Pivot to clean and model different datasets.

Website: www.clearanalyticsbi.com

4.5.8 Tableau

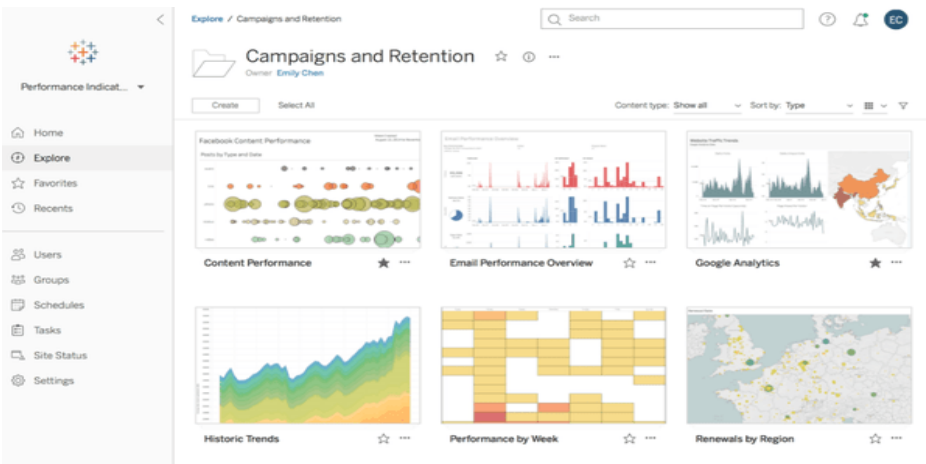


Figure 4.6: Tableau

Tableau is a Business Intelligence tool specialized in data discovery and data visualization. With the software you can easily analyses, visualize and share data, without IT having to intervene. Tableau supports multiple data sources such as MS Excel, Oracle, MS SQL, Google Analytics and Salesforce. Users will gain access to well-designed dashboards that are very easy to use. Additionally, Tableau also offers several standalone products including Tableau Desktop (for anyone) and Tableau Server (analytics for organizations), which can be run locally, Tableau Online (hosted analytics for organizations) and many more.

Website: www.tableau.com

4.5.9 Oracle BI



Figure 4.7: Oracle BI

Oracle BI is an enterprise portfolio of technology and applications for business intelligence. This technology gives users pretty much all business intelligence capabilities, such as dashboards, proactive intelligence, ad hoc, and more. Oracle is also great for companies who need to analyse large data volumes (from Oracle and non-Oracle sources) as it is a very robust solution. Additional key features include data archiving, versioning, a self-service portal and alerts/notifications.

Website: www.oracle.com

4.5.10 Domo

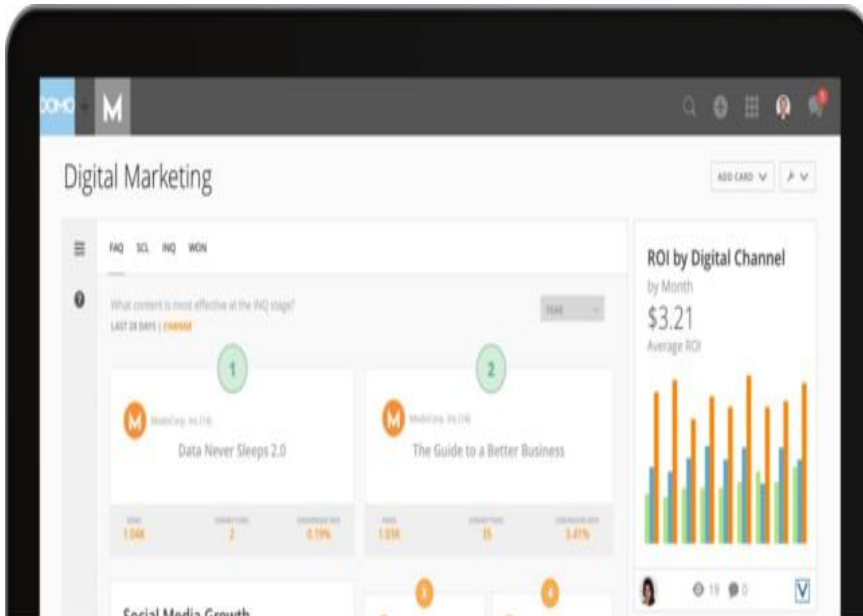


Figure 4.8: Domo

Domo is a completely cloud-based business intelligence platform that integrates multiple data sources, including spreadsheets, databases and social media. Domo is used by both small companies and large multinationals. The platform offers micro and macro level visibility and analyses (including predictive analysis powered with Mr. Roboto, their AI engine). From cash balances and listings of your best-selling products by region to calculations of the marketing return on investment (ROI) for each channel. The only setbacks with Domo are the difficulty in downloading analyses from the cloud for personal use and the steep learning curve.

Website: www.domo.com

UNIT-V

5. Business Intelligence and Integration Implementation

Business Intelligence is the practice of collecting and analyzing data and transforming it into useful, actionable information. In order to make good business decisions, leaders need accurate insights into both the market and day-to-day operations. Business Intelligence uses methods and tools like machine learning to take massive, unstructured swaths of data and turn them into easy-to-use reports.

But how exactly to implement BI into a company? What kinds of BI tools are available? This article aims to outline the process.

5.1 Pitch to Key Players

Because employees of many different departments will be involved in managing the data in business intelligence, everyone needs to be on the same page before moving forward. Because it is so important to have a shared understanding, have a look at BI software solutions together to make sure all of the key people that will be involved have a full understanding.

5.1.1 Choose Tools and Create a Team

The next step is preparation. In order to understand what sort of tool will be needed, defining the requirements for the new BI system is necessary. Smaller businesses are usually able to utilize a BI system as-is, while larger companies may need to look at custom solutions. Another important task at this stage is to gather a team to work on a Business intelligence strategy.

It is important to have a representative from each department who will be involved in its implementation so that they can simplify communication and provide their own department-specific insights from the beginning.

You need to make sure that all departments are data-friendly and in sync with each other.

5.1.2 Develop a Strategy

Depending on the industry, a BI strategy will require a variety of aspects. Most will include documentation of data sources, the KPIs of the specific industry, the kind of reporting necessary, and whether or not the data flow will require automation. Laying out these components will be helpful down the line.

5.1.3 Set Up Data Integration

This step is one of the most involved and will require quite a bit of time, as well as a lot of work from the IT department. Data warehouses, a database that keeps the information in a processed and defined format, cannot connect directly to information sources, so data integration tools must process the raw data ahead of time to allow it to be usable. Businesses with large volumes of data may also need additional technology to keep the processes from running too slow.

5.1.4 Choose an End-User Interface

This stage is where the data gets turned into usable information for the end-users. Previously, BI systems only presented statistical reports, but nowadays, interactive dashboards are available with customizable information. These interfaces allow for real-time reporting to find specific information easily. Real-time, customizable reports are called “ad-hoc reporting” and can prove invaluable.

5.1.5 End-User Training

The last step in integrating a system into a company is to make sure all employees understand how to use it. It is highly recommended to implement training sessions for all potential users, whether it is an automated process or in-person training sessions with a manager or other team member.

Although business intelligence has been around for a couple of decades, the landscape of available technologies is ever-changing. Rather than just supplying statistical reports, customizable, interactive interfaces allow for much simpler and more effective collaboration between analysts. The technology continues to evolve, including cloud-based and mobile platforms

5.2 Privacy and ethics social networking and BI

For this topic, refer ppt..... link is given

BusinessIntelligencepptdownload

<https://slideplayer.com/slide/16201498/#.Y44xfdKfAq8.whatsapp>

In today's data-driven world, businesses have a wealth of information at their fingertips, and one of the richest sources of customer insights is social media. Business Intelligence (BI) techniques are increasingly being employed to harness the power of social media data. In this article, we'll explore the synergy between Business Intelligence and social media and how companies can effectively extract valuable customer insights to drive informed decision- making.

5.2.1 The Goldmine of Social Media Data

Social media platforms are not just spaces for connecting and sharing; they are treasure troves of customer data. From Facebook and Twitter to Instagram and LinkedIn, users generate vast amounts of information through posts, comments, likes, shares, and more. This data can provide critical insights into customer behavior, preferences, sentiments, and trends.

5.2.2 What Are Social Media Insights?

Social media insights are the valuable data and information we glean from social media platforms. Through social media business intelligence,

this data can give businesses a greater understanding of their customers, competition, and industry trends.

5.2.3 What Is Social Media Business Intelligence?

Social media business intelligence (SMBI) uses social media data to inform and guide business decisions. SMBI tools and techniques can collect, store, analyze, and visualize social media data.

Social media business intelligence can help you understand your customers, improve your marketing strategy, and make better business decisions. Social media insights are the data and information you can glean from social media platforms.

This can include demographics, interests, behaviors, and other data to help you understand your target audience. Business intelligence is a process that uses data and analytics to help businesses make better decisions.

It allows businesses to track performance, understand customers, and make predictions. Social media business intelligence is using social media data to gain insights into your business. This can include understanding your customers, improving your marketing strategy, and making better business decisions

5.2.4 Key Metrics of Social Media

Like panning for gold, not everything that falls into your funnel is going to be useful. To ensure you collect the right kind of data, you should identify which key performance indicators (KPIs) are relevant to your needs and interests. Some key metrics for social media data collection are:

- Number of followers or page likes
- Response rate
- Demographics and location
- Post reach
- Post engagement

- Comments
- Mentions
- Retweets
- Shares
- Likes
- Active followers
- Traffic

5.2.5 How Business Intelligence Fits In

Business Intelligence involves the use of advanced analytics, data mining, and reporting tools to transform raw data into actionable insights. When applied to social media data, BI can help businesses:

1. **Sentiment Analysis:** Determine how customers feel about their products, services, and brand. BI tools can analyze social media comments and reviews to gauge sentiment, identifying positive and negative trends.
2. **Demographic Profiling:** Understand the demographics of social media followers and users. BI can help segment and target specific customer groups effectively.
3. **Competitor Analysis:** Monitor and analyze competitor activities and customer reactions on social media, gaining a competitive edge.
4. **Trend Identification:** Detect emerging trends and topics in real-time. This can inform marketing strategies and product development.
5. **Customer Feedback:** Gather valuable feedback from customers through social media comments and messages, identifying areas for improvement.

5.2.6 Key Benefits of BI in Social Media

1. **Informed Decision-Making:** BI empowers businesses to make informed decisions based on data-driven insights rather than intuition alone.
2. **Improved Customer Engagement:** By understanding customer preferences and sentiment, companies can tailor their social media content and engagement strategies for better results.
3. **Enhanced Marketing Strategies:** BI can reveal which social media channels and campaigns are most effective, allowing businesses to allocate resources strategically.
4. **Risk Mitigation:** Early identification of negative trends or customer dissatisfaction can lead to proactive problem-solving and reputation management.

5.2.7 Social Media Business Intelligence Can Help Businesses

- Better understand their customers
- Track and analyze customer sentiment
- Identify social media influencers
- Monitor competitors
- Spot industry trends

Many social media BI tools are available, some of which are free and some paid. Some popular social media BI tools include:

- **Hootsuite Insights:** Hootsuite's social media analytics tool provides users with valuable insights into their social media data.
- **Sprout Social:** Sprout's social media reporting and analytics features help businesses to track, measure, and improve their social media performance.
- **Buzz Sumo:** Buzz Sumo is a tool that helps users analyze social media content and identify influencers.

- **Meltwater:** Meltwater's social media monitoring tool tracks online conversations and provides users with valuable insights.
- **Brand watch:** Brand watch is a social media monitoring tool that provides users with valuable insights into social media data.

There are several ways to use social media for business intelligence. Here are some tips: Listen to what your customers are saying

To do so, use social media listening tools and track what your customers say about your brand, products, and industry. This data will improve your marketing strategy, product development, and customer service.

5.2.8 Monitor your Competition

Use social media to keep an eye on what your competitors are doing. What kind of content are they sharing? What social media platforms are they available and responsive on? What strategies seem to be working well for them?

5.2.9 Analyze Social Media Data

Use social media analytics tools to gather data about your and your competitor's social media activity. This data can improve your social media strategy and measure the success of your campaigns.

5.2.10 Stay Up-To-Date on Industry Trends

Use social media to stay abreast of industry news and trends. This data can keep you informed of your business decisions and strategies.

Make social media part of your business intelligence strategy

Social Media Analytics

Social media analytics studies social media data to evaluate and optimize business outcomes.

There are four types of social media analytics which include descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics. Analytics is the digging tool for social media data.

1. **Descriptive analytics:** social media data that helps answer the question of “what happened?”
2. **Diagnostic analytics:** social media data that helps answer the question of “why did it happen?”
3. **Predictive analytics:** social media data that helps answer the question of “what will happen next?”
4. **Prescriptive analytics:** social media data that helps answer the question of “what should we do?”

Social media data can help businesses better understand customers, monitor sentiment, track competitor activity, and spot industry trends.

What Are the Seven Layers of Social Media Analytics?

There are seven layers of social media analytics:

- **Data collection:** This is gathering social media data from various online platforms.
- **Data processing:** This is the process of cleaning and organizing social media data for analysis.
- **Data mining:** This is the process of extracting valuable insights from social media data.
- **Data visualization:** This presents social media data in an easy-to-understand format.
- **Data interpretation:** This is the process of making sense of social media data and extracting meaning from it.
- **Data analysis:** This is the process of analyzing social media data to gain insights into your business.
- **Data reporting:** This is presenting social media data in a report format.

Each layer of social media analytics is essential in its own right. Understanding all seven layers allows you to use social media data more effectively to improve your business.

What Analytics Are Important For Social Media?

Some important social media analytics include:

- **Sentiment analysis:** This is the process of analyzing social media data to understand the emotions and opinions of your customers.
- **Engagement rate:** This is the percentage of people interacting with your social media content. It measures the success of your marketing campaigns.
- **Reach:** This is the number of people who see your social media content. It helps to measure the success of your marketing campaigns.
- **Click-through rate:** This is the percentage of people who click on a link in your social media content. It can be used to measure the effectiveness of your marketing campaigns.

Social intelligence (not to be confused with social BI, which involves the sharing of reports and other visualizations generated by business intelligence platforms) involves collecting data from people who view or interact with your business on social media. This data can consist of demographics, location, number of times visitors view the page before making a purchase, etc.

By collecting this data, users can glean insights, interpret trends and make predictions based on said data in order to make data-driven decisions about their business practices. This can be anything from marketing changes in order to cater to key demographics, brand alignment with certain causes, price changes, etc.

According to the most recent studies, we generate 2.5 quintillion bytes of data every day, and that number is expected to be closer to 463 exabytes by 2025 — that's 463 with 20 zeros behind it. There are 3.48 billion unique social media users worldwide contributing to that data creation, and YouTube, Facebook and Instagram all hold positions in the top 10 most-visited sites on the Internet. Social media data is rich and there for the taking, so any organization that isn't using it is behind the times.

Text Books

1. Data Mining – Concepts and Techniques - JIAWEI HAN & MICHELINE KAMBER, Elsevier, 4th Edition.
2. Rajiv Sabherwal “Business Intelligence” Wiley Publications, 2012.

Reference Books

1. Efraim Turban, Ramesh Sharda, Jay Aronson, David King, Decision Support and Business Intelligence Systems, 9th Edition, Pearson Education, 2009.
2. David Loshin, Business Intelligence - The Savy Manager's Guide Getting Onboard with Emerging IT, Morgan Kaufmann Publishers, 2009.
3. Philo Janus, Stacia Misner, Building Integrated Business Intelligence. Solutions with SQL Server, 2008 R2 & Office 2010, TMH, 2011. R22 B.Tech. CSE (AI and ML) Syllabus JNTU Hyderabad Page 97 of 147
4. Business Intelligence Data Mining and Optimization for decision making [Author: Carlo-Verellis] [Publication: (Wiley)]
5. Data Warehousing, Data Mining & OLAP- Alex Berson and Stephen J. Smith- Tata McGrawHill Edition, Tenth reprint 2007
6. Building the Data Warehouse- W. H. Inmon, Wiley Dreamtech India Pvt. Ltd.
7. Data Mining Introductory and Advanced topics – Margaret H Dunham, PEA.