Business Intelligence and Analytics A Comprehensive Overview

Presentation · August 2018

CITATIONS
0

READS
4,131

1 author:

Jack G Zheng
Kennesaw State University
40 PUBLICATIONS **332** CITATIONS

SEE PROFILE

Business Intelligence and Analytics

A Comprehensive Overview

IT 4713/6713 BI

Jack G. Zheng

Fall 2019 (since V1 2012)

http://jackzheng.net/teaching/it4713/
http://jackzheng.net/teaching/it6713/
https://www.researchgate.net/publication/327578485
https://www.edocr.com/v/r4dg6mjr/

Kennesaw State University

Overview

- What is business intelligence (BI) and analytics?
  - BI/Analytics as an information and *decision process*
  - BI/Analytics as an *information system*
- BI/Analytics process
  - Data, information, decision
- BI/Analytics systems and tools
  - Values, capabilities, and components
  - Technologies, architectures, platforms
  - Application areas
  - Products, industries, and markets
- BI evolution and trend: traditional BI and modern BI
- BI/Analytics learning and career

Types of Information Processing

For a more detailed comparison of **OLTP and OLAP**:
http://www.slideshare.net/fmhyudin/oltp-vs-olap-23317601

Transactional Processing

- Focus on data item processing (data insertion, modification, deletion), transmission, and non-analytical query

- Change product price
- Increase customer credit limit
- Who has not paid bills?

Analytical Processing

- Focus on reporting, analysis, transformation, and decision support

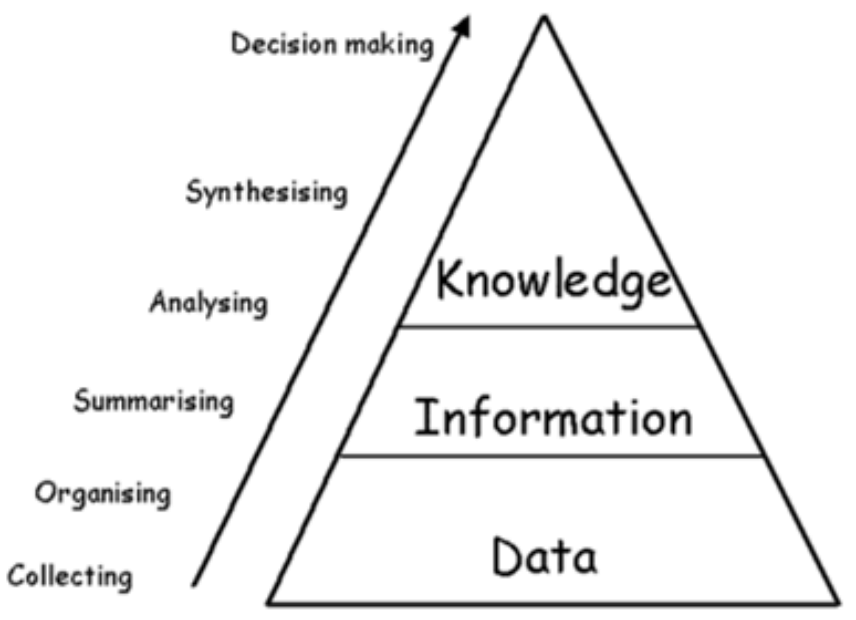
- Is there a significant increase of operational cost?
- What are the top 10 most profitable products?

DIKW

- The DIKW hierarchy depicts relationships between data, information, knowledge (and wisdom).
  - Data: raw value elements or facts
  - Information: the result of collecting and organizing data that provides context and meaning
  - Knowledge: the concept of understanding information that provides insight to information, thus useful and actionable
- The model can be loosely relate to the levels of transactional processing (OLTP) and analytical processing (OLAP)

Analytical Processing

Transactional Processing



For more extensive reading: http://en.wikipedia.org/wiki/DIKW_Pyramid
Different opinion: https://hbr.org/2010/02/data-is-to-info-as-info-is-not

Examples of Analysis

- Non-analytical query (like a simple list based on certain conditions)
  - Get a list of students enrolled in in the IT 6713 class.
- Descriptive analysis (summarizing)
  - How many students are enrolled in online IT graduate courses for the past year?
- What if analysis
  - If inventory levels are reduced by 10%, what is the new cost of inventory storage?
- Reasoning and correlation
  - What is the reason for a decrease of total sales this year?
  - How do advertising activities affect sales of different products bought by different type of customers, in different regions? (synthesizing)
- Fuzzy decision
  - What new advertising strategies need to be undertaken to reach our customers who can afford a high priced product?
  - Should we invest more on our e-business?

What is Business Intelligence?

Business Intelligence is a set of methods, processes, architectures, applications, and technologies that gather and transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making.

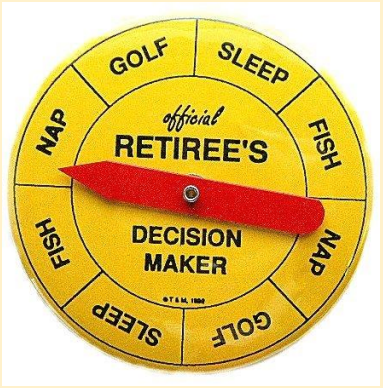
Adapted from Forrester Report
"Topic Overview: Business Intelligence", 2008
https://www.forrester.com/report/Topic+Overview+Business+Intelligence/-/E-RES39218
More BI from Forrester
https://www.forrester.com/business-intelligence

Data

- Different types of data
  - Numeric vs. textual
  - Structured vs. unstructured
  - Standard format vs. proprietary format
  - Internal vs. external data, system stored vs. file based data
  - Raw fact data vs. simulated/forecast/estimated data
  - Simple fact data vs. calculated metrics data

- Common data problems
  - Structured, unstructured, semi-structured
    - Information and knowledge management is the management of both structured data (15% of information) and unstructured data (85% of information), according to the Butler Group.
    - 80 percent of business is conducted on unstructured information (Gartner Group).
  - Information overloading
    - too much data and information with varied formats and structure
    - difficulty of data organization for effective access and retrieval
    - difficult to find useful information (knowledge) from them
    - Multiple copies of data exists sometimes with conflicts
  - Big data
    - Variety, Velocity, Volume, Veracity https://www.ibmbigdatahub.com/infographic/four-vs-big-data
  - Data everywhere
    - Data in separate systems and different sources; internal and external
    - Problem of spreadmart http://en.wikipedia.org/wiki/Spreadmart
    - Over 43 percent of organizations have more than six content stores. (Forrester Research).
  - Difficulty of access
    - We may have that data but we cannot access it (or difficult to get it), because of technical issues or administrative issues.
  - Don't have that data
    - The data is simply not available.
    - The collection of data may need additional process and is costly.

Decision Making



- Decisions can be made based on
  - Facts, or data
  - Simulation (models)
  - Intuition, perception, sense
  - Group negotiation
- Traditionally BI has been also understood as Decision Support System (DSS) – known as data driven DSS (data directly contributes to decision without intensive and advanced analytical techniques).

Extended reading: a brief history of DSS
http://dssresources.com/history/dsshistory.html

- Problems in decision making
  - A gap between data and knowledge (useful information leading to a decision).
  - Management/operation by intuition
  - Lack of effective feedback and alignment systems, no improvement cycles
  - Need good analytical processing and models
- Evolving analytical needs in decision support
  - Real-time, most recent data
  - Business user driven, agile, instant
  - Exploratory and interactive

Additional Notes about BI

- BI is the an umbrella term for a set of methods, processes, applications, and technologies used to
  - gather, provide access to, analyze, and report data and information
  - support understanding and decision making
  - A common goal in BI is to drive performance
- The evolution of BI resides both in “business” and “intelligence”
  - The term “business” is more general and represents the application domain; not just related to profit driven businesses.
  - Traditionally BI is related to business or corporate operations, but can also extend to other types of organizational contexts, like non-profits, governments, institutions, etc.
  - Intelligence represents the resource and the techniques or methods
- Narrowly speaking, intelligence comes from data (facts). Traditional BI normally does not directly address other content types and formats (which usually falls under artificial intelligence).
  - In this sense, BI focuses on *analytical data processing*.
- Broadly speaking, intelligence, or knowledge, comes from human experience and tacit knowledge, in various format like text, image, video, etc.
  - In this sense, BI is also related to *knowledge management* (either BI under KM or vice versa) http://capstone.geoffreyanderson.net/export/19/trunk/proposal/research/Knowledge_management.pdf

Evolution of BI

The search for the perfect “business insight system”, from Performance Dashboard, by Wayne Eckerson
http://download.101com.com/pub/tdwi/files/performancedashboards.pdf

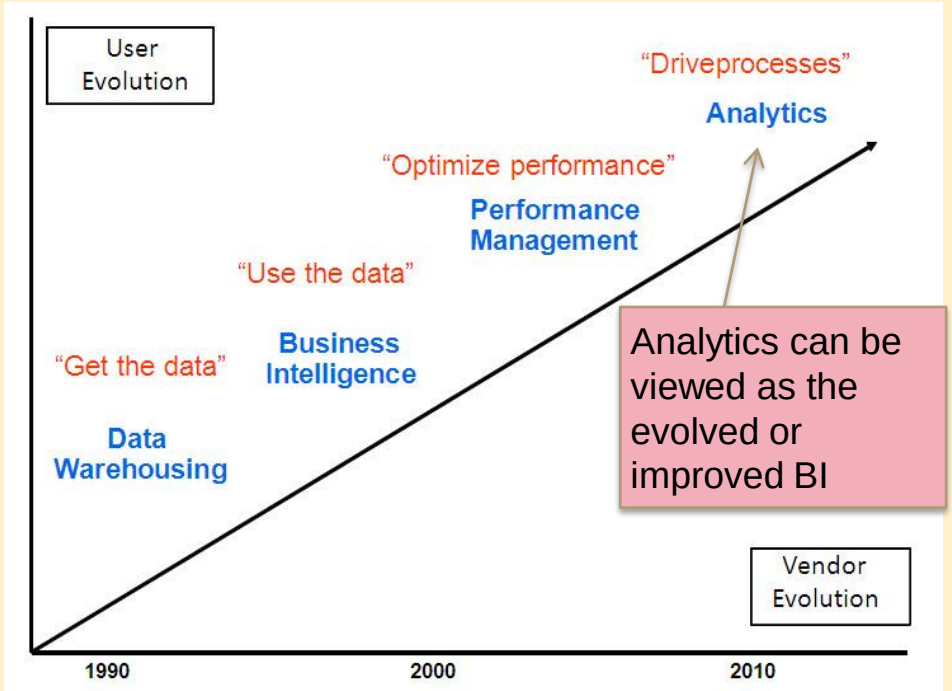
1980s	Executive information systems (EIS), decision support systems (DSS)
1990s	Data warehousing (DW), business intelligence (BI)
2000s	Dashboards and scorecards, performance management
2010+??	Analytics, big data, personal BI, data science, augmented BI, …

“With each new iteration, capabilities increased as enterprises grew ever-more sophisticated in their computational and analytical needs and as computer hardware and software matured.”

Solomon Negash (2004), Business Intelligence, CAIS (13)
https://www.researchgate.net/publication/228765967_Business_intelligence

Analytics

- Analytics has emerged as a catch-all term for a variety of different business intelligence (BI) and application-related initiatives. … Whatever the use cases, “analytics” has moved deeper into the business vernacular.
  - https://www.gartner.com/it-glossary/analytics/
- Analytics refers to a more systematical, automated, and flexible process of data analysis for revealing insights and decision support in more extensive application areas, e.g. sports, disease, network traffic, etc.
  - http://pestleanalysis.com/differences-between-business-analytics-and-business-analysis/
- Analytics initially referred to advanced statistical modeling using tools like SAS and SPSS. … Now, analytics refers to the entire domain of leveraging information to make smarter decisions. In other words, reporting and analysis.
  - The Evolution of BI Semantics http://www.b-eye-network.com/blogs/eckerson/archives/2011/02/whats_in_a_word.php



The Evolution of BI Semantics
http://www.b-eye-network.com/blogs/eckerson/archives/2011/02/whats_in_a_word.php

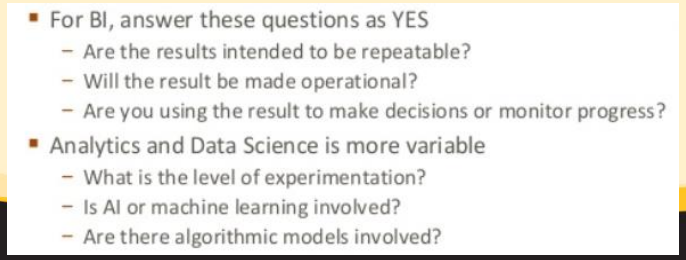
BI and Other Related Terms

- Big data
  - "Big Data is not a system; it is simply a way to say that you have a lot of data. https://www.linkedin.com/pulse/big-data-silver-bullet-tomas-kratky
  - Big data covers non-structure and various data formats including text, blob, multimedia, etc.

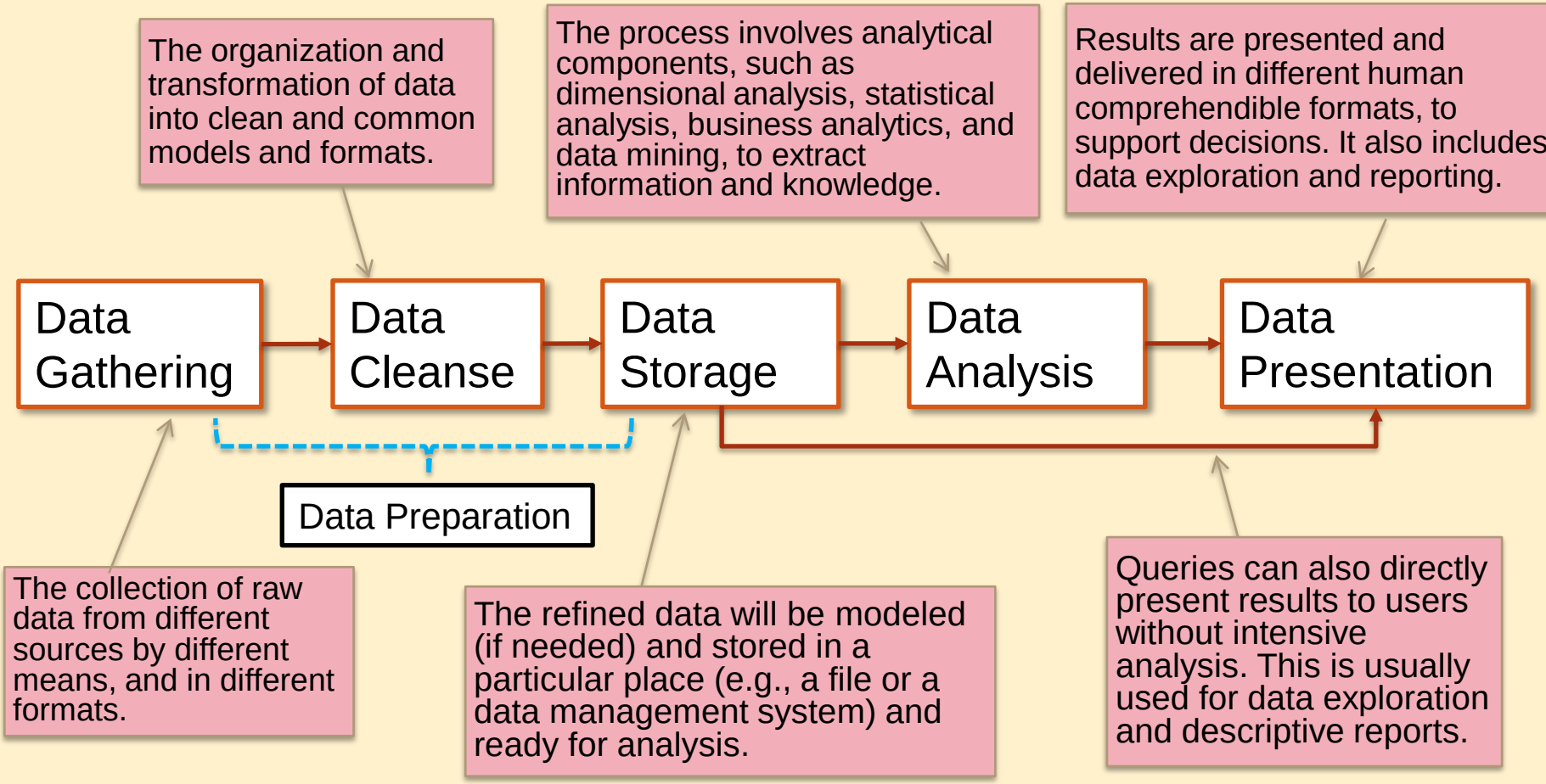
- Data science
  - An interdisciplinary field about processes and systems to extract knowledge or insights from data in various forms
  - Focus on advanced analytics and presentation models and methods
  - Using autonomous or semi-autonomous techniques and tools, typically beyond traditional BI to discover deeper insights, make predictions, or generate recommendation.
  - A good data scientist = data hacker + programmer+ analyst+ coach+ story teller+ artist (http://analyticsindiamag.com/data-science-the-most-desirable-job-in-the-21st-century/)
  - "In some ways, data science is an evolution of BI." https://www.linkedin.com/pulse/data-science-business-intelligence-whats-difference-david-rostcheck/

- All these new terms try to differentiate them from the (traditional) BI. However, if one considers BI is a dynamic and evolving field, then all these new terms are just extensions/expansions of BI; they all still fall under the umbrella of the general BI.
  - "In its more comprehensive usage, BI is all of the systems, platforms, software, technology, and techniques that are essential for the collection, storage, retrieval, and analysis of data assets within a given organization." – Dataversity 2015 Report on BI vs Data Science

- More perspectives from the industry
  - http://www.dataversity.net/distinguishing-analytics-business-intelligence-data-science/ and https://www.slideshare.net/Dataversity/analytics-business-intelligence-and-data-science-whats-the-progression
  - https://solutionsreview.com/business-intelligence/data-science-vs-data-analytics-whats-the-difference/
  - https://www.betterbuys.com/bi/business-intelligence-vs-business-analytics/

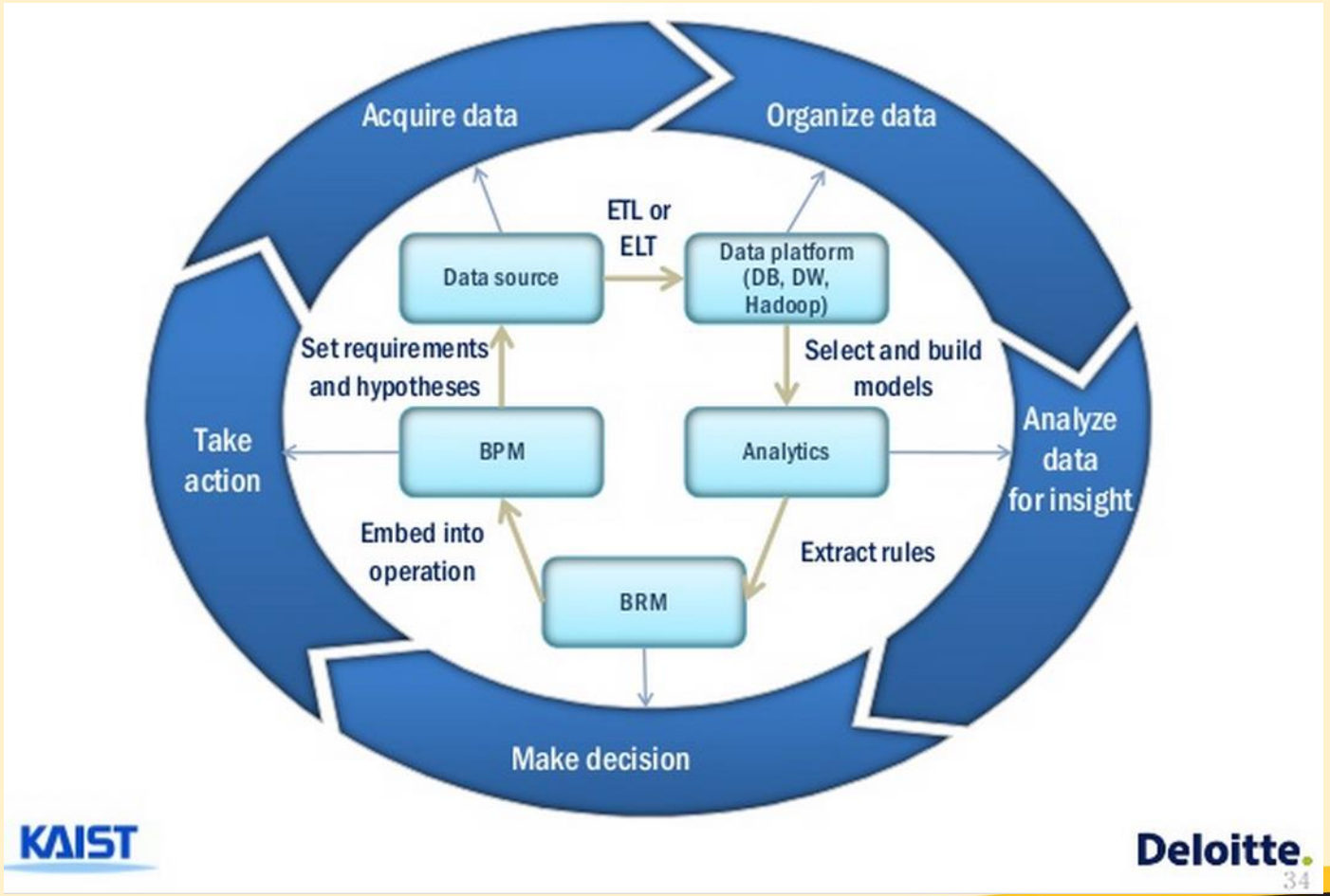


BI/Analytics: A General Process



BI in the Decision Process

Another view from the business decision perspective
http://www.slideshare.net/junesungpark/business-process-based-analytics



BI Capabilities

This is consistent with the general BI or analytics process but more from an information behavior angle.

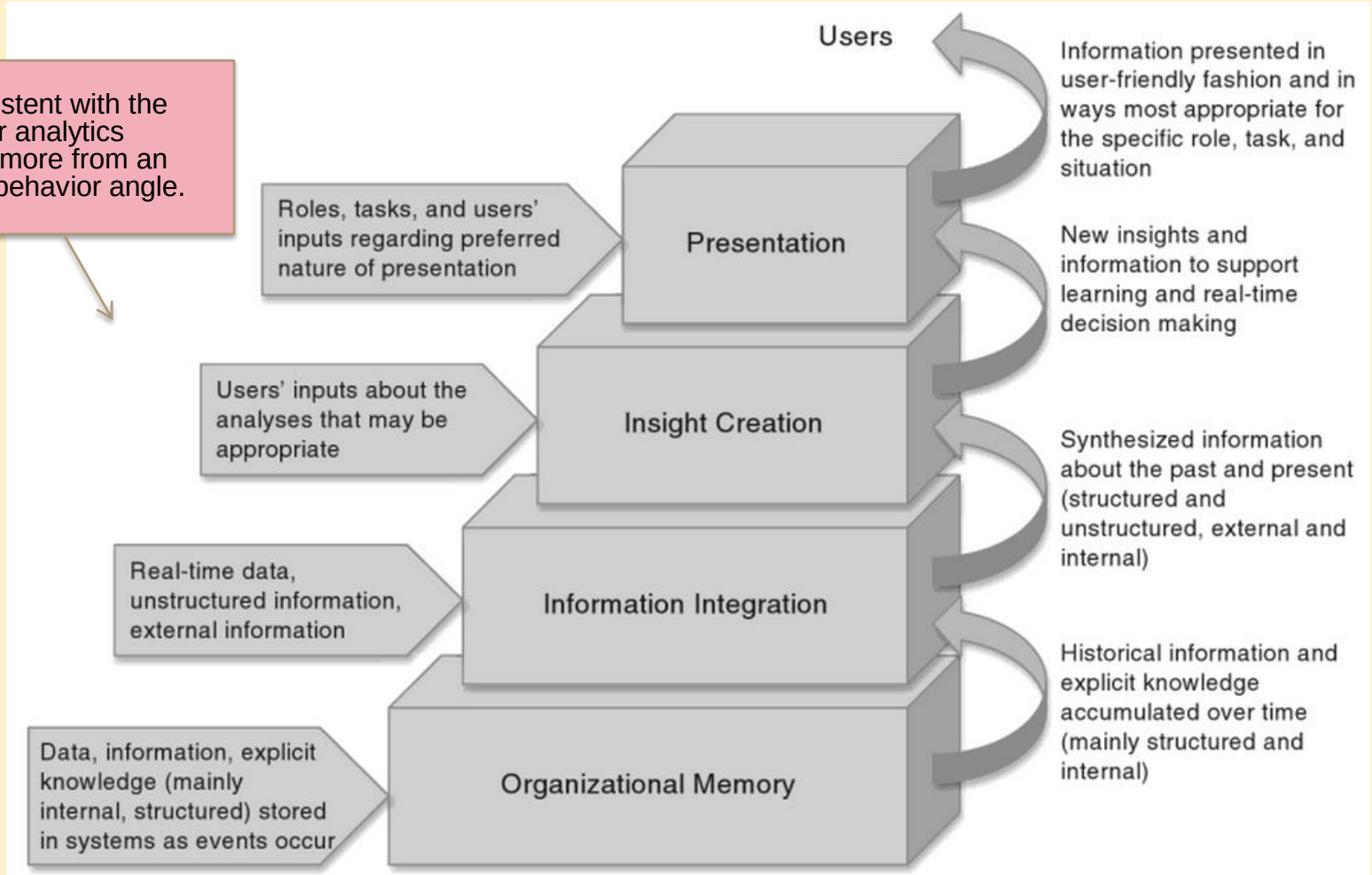


FIGURE 2.5 Inputs and Outputs of the Four Business Intelligence Capabilities

Figure from: Business Intelligence, Rajiv Sabherwal, Irma Becerra-Fernandez, John Wiley & Sons, 2011
http://books.google.com/books?id=T-JvPdEcm0oC

BI Systems and Platforms

- A BI system is a computer information system that implements (part or whole) BI capabilities and processes
- The values of BI Systems
  - Provide an integrated data (analytical) processing platform
  - Enable easy and fast access of data and information at all levels (raw data, analysis results, metrics, etc.)
  - Streamline a controlled and managed process of data driven decision making
- An enterprise level BI system emphasizes more on control and performance. While a more user-oriented analytics platform enables nontechnical users to autonomously execute full-spectrum analytic workflows from data access and preparation to interactive analysis and the collaborative sharing of insights.

BI System (Components) at a Glance

Data Storage and Management

- Relational database
- Data warehouse
- Data lake
- Data modeling
- Data governance
- Data integration
- ETL
- Data quality
- Metadata
- Master Data
- Data virtualization

Analytical Processing

- Query
- OLAP
- Business analytics
- Statistics
- Data mining
- Text mining
- Advanced analytics

Presentation

- Reports
- Data visualization
- Dashboard
- Scorecards
- Strategy map
- Visual analytics
- Free form results

Delivery

- Local files
- Website
- Reporting server
- Application server
- BI server
- Portal
- Excel services

Applications

- Performance management
- Benchmarking
- Market research
- CRM
- Strategic management
- Web site analytics

Users with applications (browser, desktop app, mobile app, email, etc.) and devices (computer, tablet, phone, print-outs, etc.)

Critical Capabilities of a BI and Analytics Platform

Gartner Magic Quadrant Report 2018/2019

- **Infrastructure**
  - **BI Platform Administration.** Capabilities that enable scaling the platform, optimizing performance and ensuring high availability and disaster recovery.
  - **Cloud BI.** Platform-as-a-service and analytic-application-as-a-service capabilities for building, deploying and managing analytics and analytic applications in the cloud, based on data both in the cloud and on-premises.
  - **Data Source Connectivity.** Capabilities that allow users to connect to the data contained within various types of storage platforms.

- **Data Management**
  - **Governance and Metadata Management.** Tools for enabling users to share the same systems-of-record semantic model and metadata. These should provide a robust and centralized way for administrators to search, capture, store, reuse and publish metadata objects, such as dimensions, hierarchies, measures, performance metrics/key performance indicators (KPIs) and report layout objects, parameters and so on.
  - **Self-Contained ETL and Data Storage.** Platform capabilities for accessing, integrating, transforming and loading data into a self-contained storage layer, with the ability to index data and manage data loads and refresh scheduling.
  - **Self-Service Data Preparation.** The drag-and-drop, user-driven data combination of different sources, and the creation of analytic models such as user-defined measures, sets, groups and hierarchies.
  - **Scalability and Data Model Complexity.** The degree to which the in-memory engine or in database architecture handles high volumes of data, complex data models, performance optimization and large user deployments.

- **Analysis and Content Creation**
  - **Advanced Analytics.** Enables users to easily access advanced analytics capabilities that are self-contained within the platform itself or available through the import and integration of externally developed models.
  - **Analytic Dashboards.** The ability to create highly interactive dashboards and content, with visual exploration and embedded advanced analytics.
  - **Interactive Visual Exploration.** Enables the exploration of data via the manipulation of visual properties and visual forms representing aspects of the dataset being analyzed. These tools enable users to analyze the data by interacting directly with a visual representation of it.
  - **Augmented Data Discovery**: Automatically finds, visualizes and narrates important findings such as correlations, exceptions, clusters, links and predictions in data that are relevant to users without requiring them to build models or write algorithms.
  - **Mobile Exploration and Authoring.** Enables organizations to develop and deliver content to mobile devices in a publishing and/or interactive mode, and takes advantage of mobile devices' native capabilities, such as touchscreen, camera, location awareness and natural-language query.

- **Sharing of Findings**
  - **Embedding Analytic Content.** Capabilities including a software developer's kit with APIs and support for open standards for creating and modifying analytic content, visualizations and applications, embedding them into a business process, and/or an application or portal. These capabilities can reside outside the application (reusing the analytic infrastructure), but must be easily and seamlessly accessible from inside the application without forcing users to switch between systems.
  - **Publish and collaborate Analytic Content.** Capabilities that allow users to publish, deploy and operationalize analytic content through various output types and distribution methods, with support for content search, storytelling, scheduling and alerts.

- **Overall:** Ease of Use, Visual Appeal and Workflow Integration.

A Practical System Architecture in MSBI

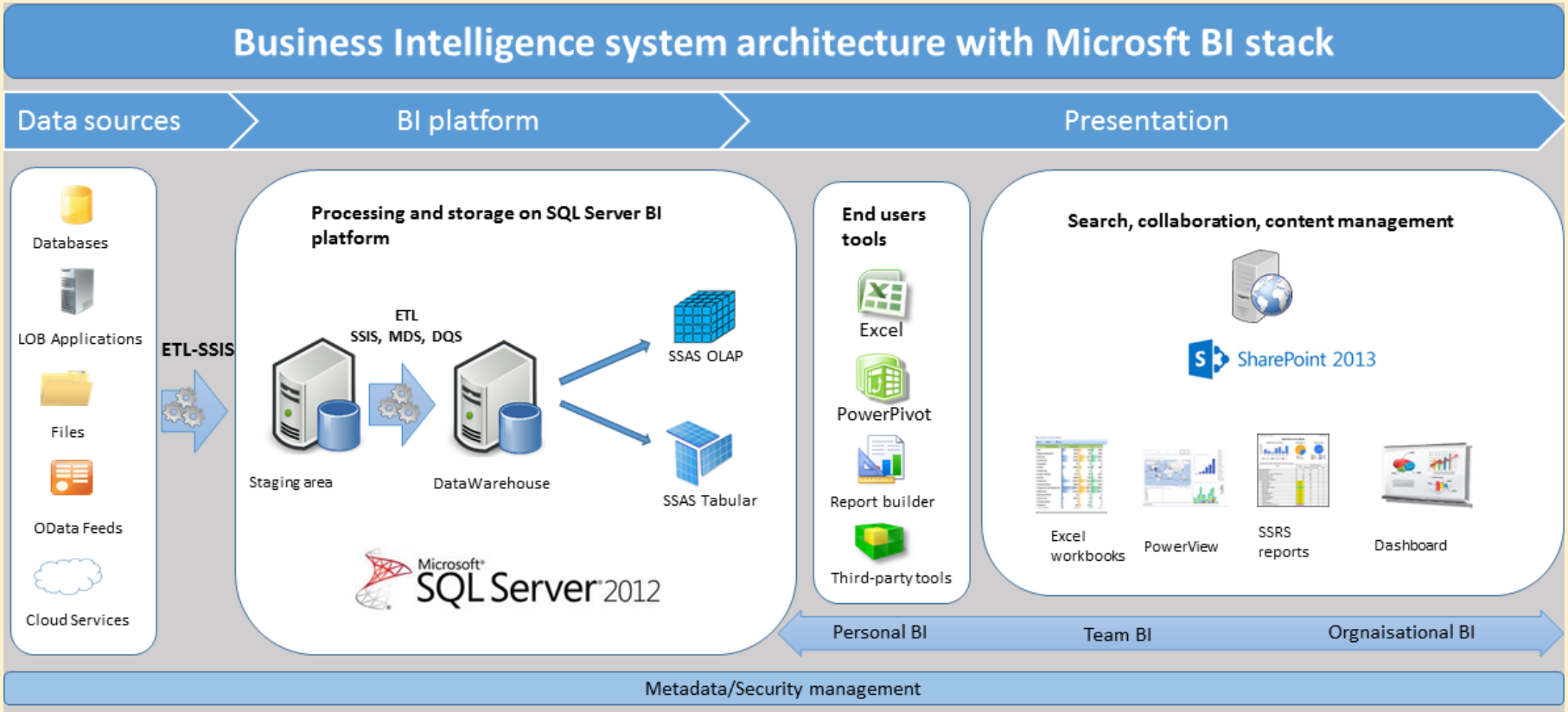


Image from https://bipointblog.wordpress.com/2014/05/28/implementation-of-a-bi-system-using-microsoft-bi-stack-introduction/

Data Management/Storage

- In traditional BI, a special database system called data warehouse or data mart is often used to store enterprise data
  - The purpose of a data warehouse is to organize lots of stable data for ease of analysis and retrieval.
- Traditional (operational) relational databases facilitate data management and transaction processing. They have two limitations for data analysis and decision support
  - Performance
    - They are transaction oriented (data insert, update, move, etc.)
    - Not optimized for complex data analysis
    - Usually do not hold historical data
  - Heterogeneity
    - Individual databases usually manage data in very different ways, even in the same organization (not to mention external data sources which may be dramatically different).
- The data warehouse approach is a centralized and structured approach for analytical data management. For more recent personal BI/analytics, data is also kept locally for easy access and manipulation, without much technical support.

Data Gathering and Integration

- Enterprise level data are coming from multiple different sources, but need to be combined and associated
  - Operational databases
  - Spreadsheets
  - Text, CSV
  - PDF, Paper

Data is never clean!
You will spend most of your time cleaning and preparing data!

- The need to bring together different data/information
  - Autonomous (may not have the control and management of data)
  - Distributed (from different systems and places)
  - Different (in data model, format, or platform)

- General processing steps - ETL
  - **Extraction**: accessing and extracting the data from the source systems, including database, flat files, spreadsheets, etc.
  - **Transformation**: data cleanse, change the extracted data to a format and structure that conform to the destination data.
  - **Loading**: load the data to the destination database, and check for data integrity

- Traditional BI focuses on upfront separate ETL processes that load the data in a centralized storage. In modern BI and analytics, data cleanse and transformation may happen just-in-time with analysis.

Analysis Techniques

- Descriptive reporting
  - Structured and fixed format reports
  - Based on simple and direct queries
  - Usually involves simple descriptive analysis and transformation of data, such as calculating, sorting, filtering, grouping, and formatting
  - Ad hoc query and reporting
- OLAP (Online Analytical Processing)
  - A multi-dimensional analysis and reporting application for aggregated data
  - Great for discovering details from large quantities of data
- Business analytics
  - Business analytics (BA) is the practice of iterative, methodical exploration of an organization's data with emphasis on statistical analysis.
- Advanced and computation intensive: data mining, deep learning, etc.
  - Data mining techniques are a blend of statistics and mathematics, and artificial intelligence and machine-learning.

OLAP

- OLAP is a function/operation that is optimized to answer queries that are multi-dimensional
  - OLAP solutions traditionally heavily rely on backend processing and dedicated IT personnel
- Multi-dimensional queries
  - A dimension is a particular way (or an attribute) of describing and categorizing data
  - Such queries are usually arithmetic aggregation operations (sum, average, etc.) on records grouped by multiple dimensions (attributes) at different aggregation levels.
  - A pivot table or crosstab is usually used for OLAP result view (aggregated data)

Descriptive and operational report

- Example analysis
  - "What is the total sales amount grouped by product line (dimension 1), location (dimension 2), time (dimension 3) and … (other dimensions)?"
  - "Which segment of business provides the most revenue growth?"

More open and exploratory analysis

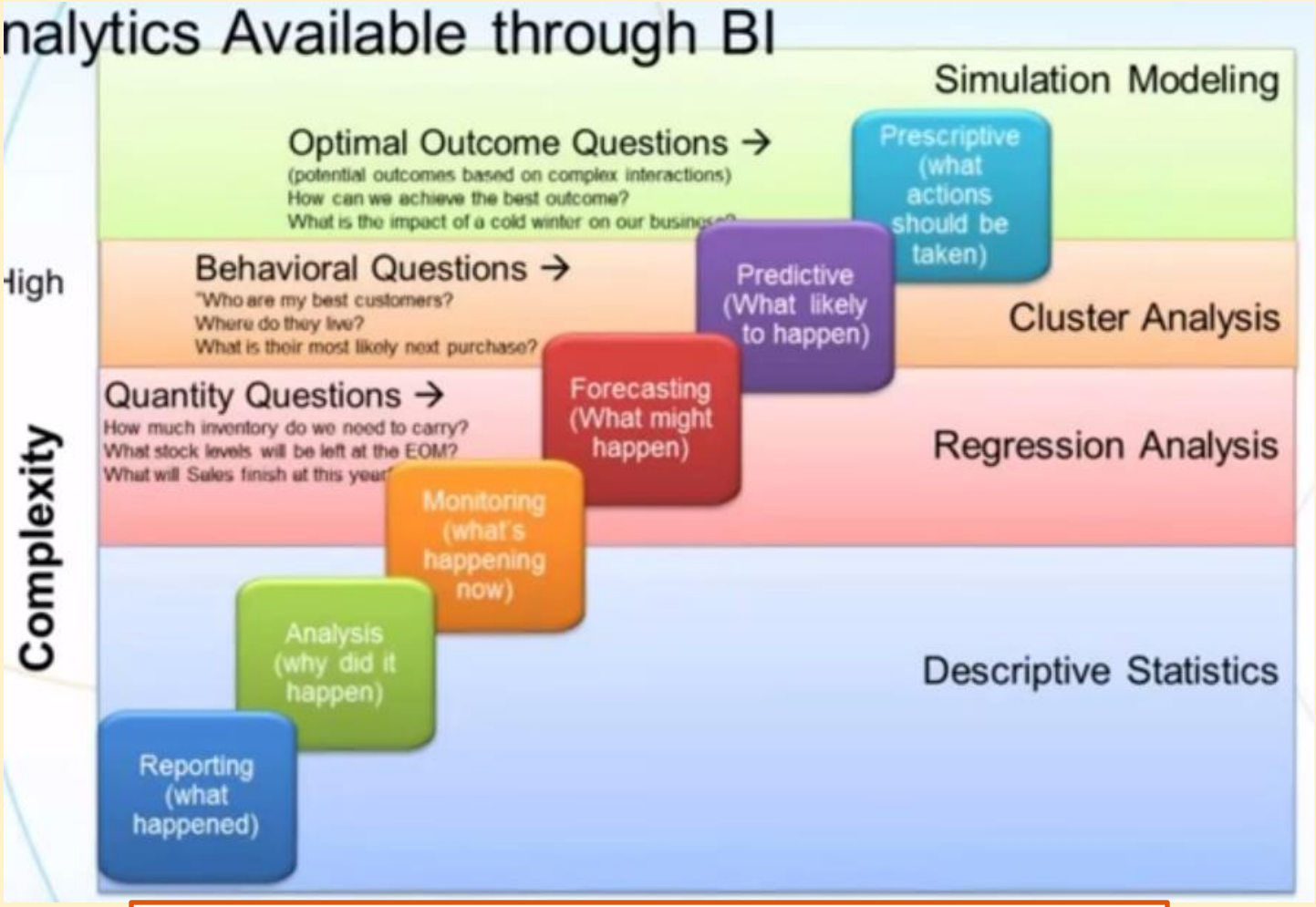
Basic Techniques in Business Analytics

- Regression
  - Reasoning, estimating the relationships among variables
- Forecasting
  - Trend analysis, based on extrapolation of historical data
- Correlation
  - Relationship discovery between factors (but not causal relationship)
- Factor analysis
  - Determine impacting variables and their variability

Data Mining

- Data mining (or, knowledge discovery in database, KDD)
  - Processes and techniques for seeking knowledge (relationship, trends, patterns, etc.) from a large amount of data
  - Non-trivial, non-obvious, and implicit knowledge
  - Extremely large datasets
- Data mining applications use sophisticated statistical and mathematical techniques to find patterns and relationships among data
  - Classification, clustering, association, estimation, prediction, trending, pattern, etc.
- Common techniques
  - Neural network, genetic algorithm, machine learning

Levels of Analytical Processing



Advanced Analytics and Business Intelligence
https://www.youtube.com/watch?v=oNNk9-tmsZY

Presentation

Presentation is key – be a master of power point.

- The last mile of BI is the presentation of data or analysis to human users
- Data presentation is the method by which people summarize, organize and communicate information using a variety of tools, including tables, diagrams/charts, and other visualization techniques
- Multiple ways to present results
  - Regular/periodical reports
  - Live and real time dashboard
  - Free form ad hoc results
  - Edited PowerPoint
- Presentation commonly utilizes data visualization techniques to assist interpreting and presenting data in a visual way.

Data Visualization

- Data visualization is the visual and interactive exploration and graphic representation of data of any size, type (structured and unstructured) or origin.
- Visualizations help data comprehension and enhance problem solving capabilities
  - Provide a high level overview of complex data sets
  - Extract/provoke additional (implicit) perspectives and meanings
  - Ease the cognitive load of information processing
  - Recall or memorize data
  - Enable perceptual inference operations and detection of patterns
- Visualization in BI
  - Data visualization is an important part of understanding for information seeking and decision making.
  - Visualization tools have become increasingly important to business intelligence, in which people need technology support to make sense of and analyze complex data sets and all types of information.
  - As organizations seek to empower non-technical users to make data-driven decisions, they must consider the prowess of data visualization in delivering digestible insights.
- Visualization can also be part of the analysis process (visual analytics)

Get more details at

- https://www.edocr.com/v/yqwmqeba/jgzheng/Business-Data-Visualization
- https://www.researchgate.net/publication/321804138_Data_Visualization_for_Business_Intelligence

Basic Visualization Forms/Styles

Form/Style	Description	Typical Types and Examples
Embedded visual	It is embedded in, or directly on top of, texts and other forms of data presentation (table, graphic, etc.).	• Conditional formatting (visual cues) • Inline chart (Sparkline)
Block visual	It is displayed as an independent visual unit and occupies a larger space. It is often a part of a report or dashboard, appearing together with other content. But sometimes it can become a standalone visual with many data points or enough complexity.	• Chart • Illustrational diagram • Map (smaller) • Data table (usually with embedded visuals)
Standalone visual	It is a standalone application and is not mixed with other types of content or tool. Most interactions are within the visual. It may consist of a combination of different types of visuals.	• Dashboard • Visual analysis tool (or an analytical dashboard) • Map (bigger or full screen)

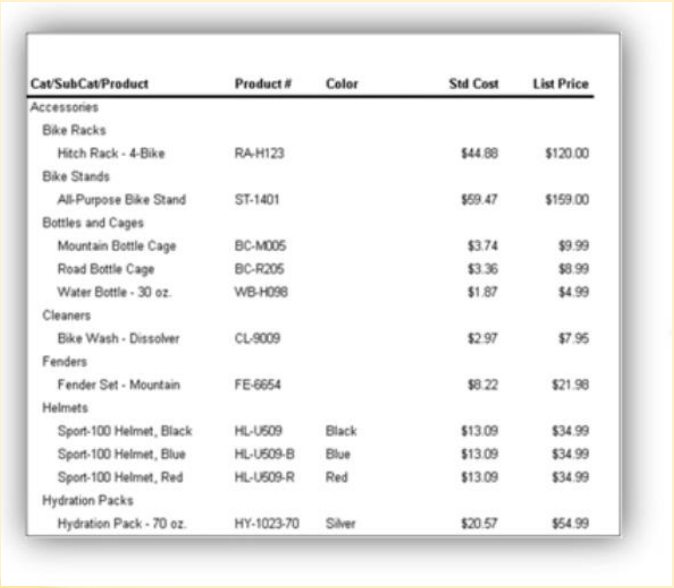
See more details at https://www.edocr.com/v/yqwmqeba/jgzheng/Business-Data-Visualization

Reports

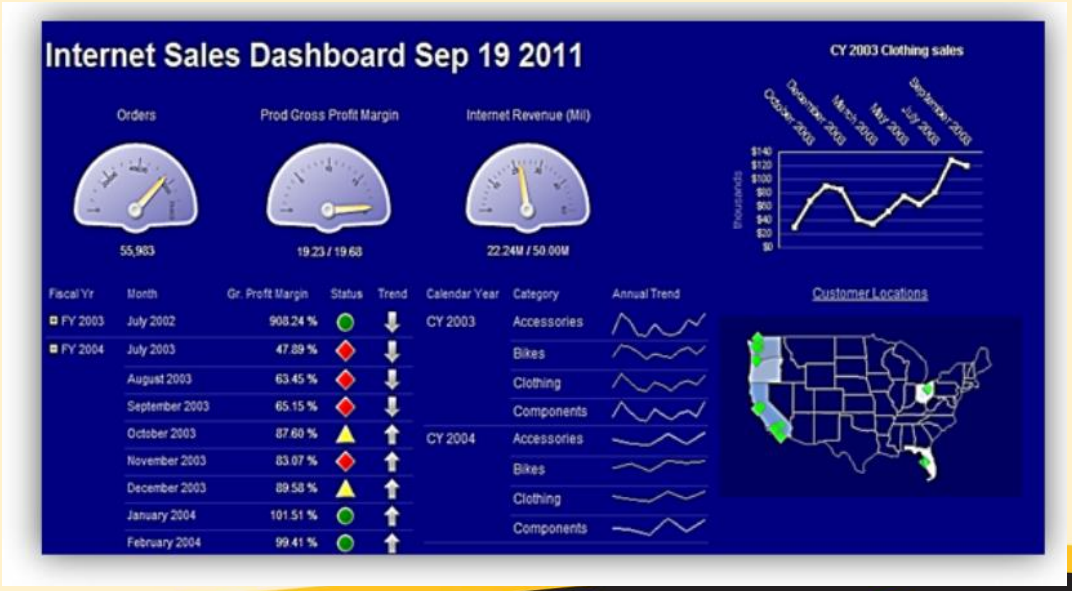
- Reports
  - A report is the presentation of detailed data arranged in defined layouts and formats
  - Based on simple and direct queries: usually involves simple analysis and transformation of data (sorting, calculating, filtering, filtering, grouping, formatting, etc.)

- Traditional reports contain detailed data in a tabular format and typically display numbers and text only.
  - Its purpose is mainly for printing (with styling) or exporting (raw data).
  - It is geared towards people who need data rather than a direct understanding of data.

- Modern reports can be interactive and visual but the focus is still on detailed data. Sometimes the distinction is a bit blurred with dashboards in some practical cases.
  - http://www.crazybikes.com/mrc/CRAZYBIKES.R00090s
  - A report style “dashboard” (or more like a visual intensive interactive report): https://www.itdashboard.gov/drupal/summary/006



Cat/SubCat/Product	Product #	Color	Std Cost	List Price
Accessories				
Bike Racks				
Hitch Rack - 4-Bike	RA-H123		$44.88	$120.00
Bike Stands				
All-Purpose Bike Stand	ST-1401		$59.47	$159.00
Bottles and Cages				
Mountain Bottle Cage	BC-M005		$3.74	$9.99
Road Bottle Cage	BC-R205		$3.36	$8.99
Water Bottle - 30 oz.	WB-H098		$1.87	$4.99
Cleaners				
Bike Wash - Dissolver	CL-9009		$2.97	$7.95
Fenders				
Fender Set - Mountain	FE-6654		$8.22	$21.98
Helmets				
Sport-100 Helmet, Black	HL-U509	Black	$13.09	$34.99
Sport-100 Helmet, Blue	HL-U509-B	Blue	$13.09	$34.99
Sport-100 Helmet, Red	HL-U509-R	Red	$13.09	$34.99
Hydration Packs				
Hydration Pack - 70 oz.	HY-1023-70	Silver	$20.57	$54.99

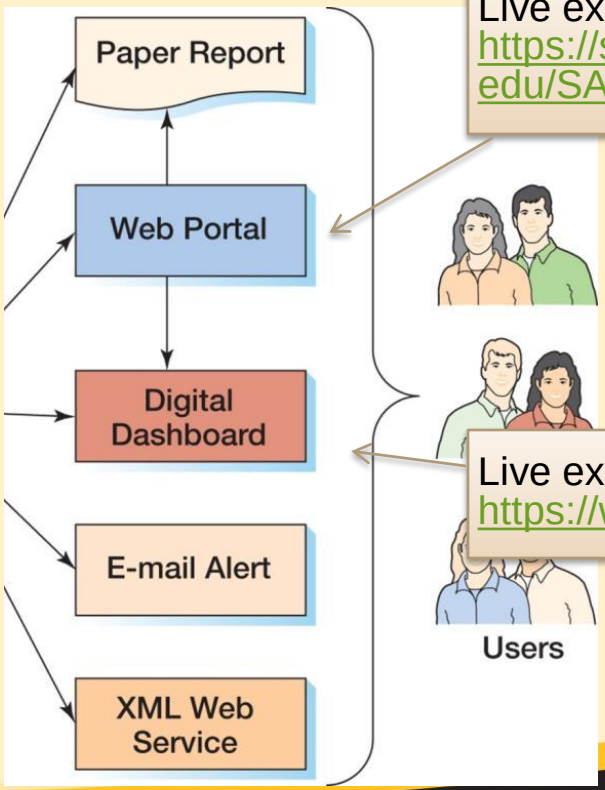


Dashboard

- A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.
  - Dashboard Confusion, Stephen Few, http://www.perceptualedge.com/articles/ie/dashboard_confusion.pdf
- A set of visualization or presentation of data views organized in a single screen/page
  - The data is generally KPIs and shows trends, breakdowns, and comparisons against a forecast or historical data
  - A dashboard generally contains a variety of different views of data: charts, diagrams, tables, standalone numbers, interactive controls (such a filters)
- Dashboard vs. report, visual analysis tool, and scorecard
  - http://www.dashboardinsight.com/articles/digital-dashboards/fundamentals/what-is-a-dashboard.aspx
- The Values of Dashboard
  - Allow decision makers to see a variety of relevant data that affects their divisions or departments
  - Quickly understand data and respond quickly at one place; save time over running multiple reports
  - More: http://www.bidashboard.org/benefits.html

Delivery Medium

- Delivery is about managing and delivering data and analysis results to users
  - Modern: social sharing, cloud hosting,



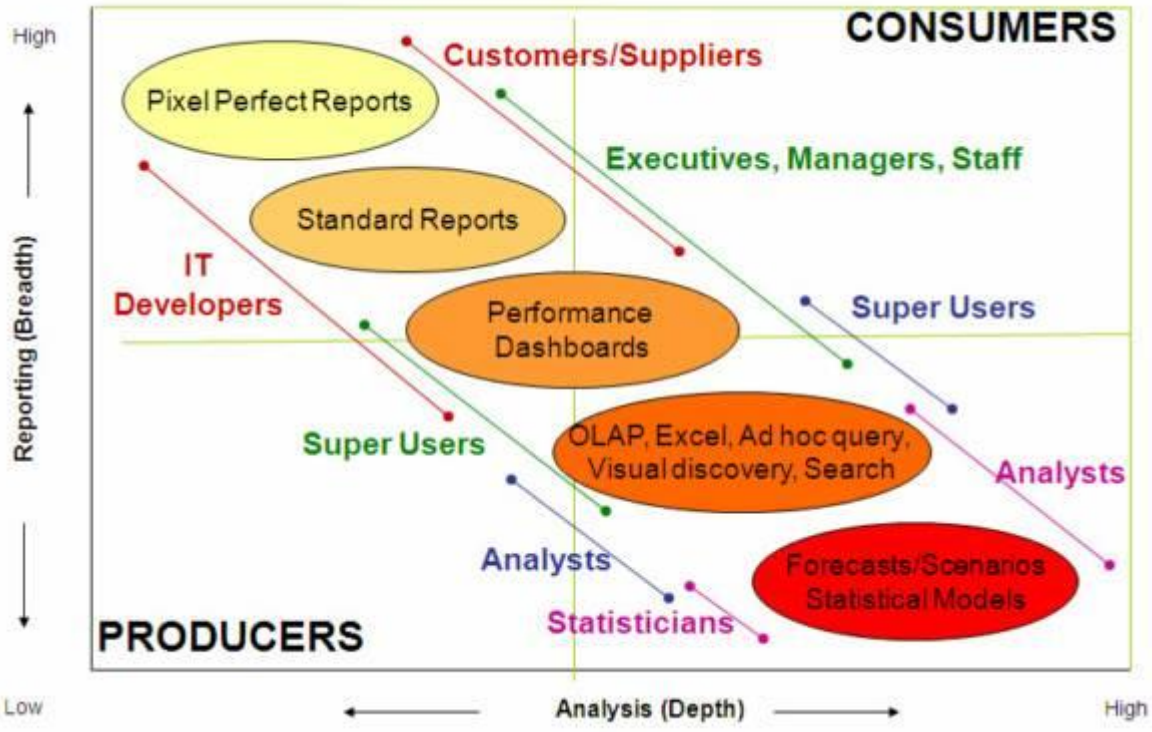
Live example: https://sasweb1.kennesaw.edu/SASPortal/public

Live example: https://www.itdashboard.gov/

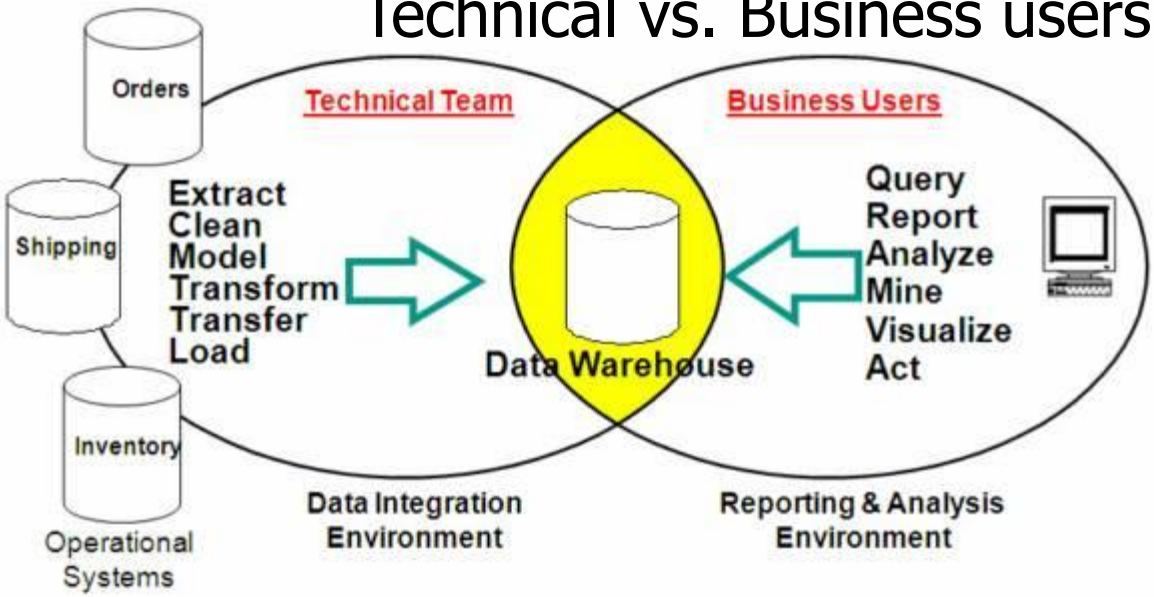
Figure from Database Processing13th Edition, by David Kroenke and David Auer

BI Users

Producers
vs.
Consumers
(at different levels)

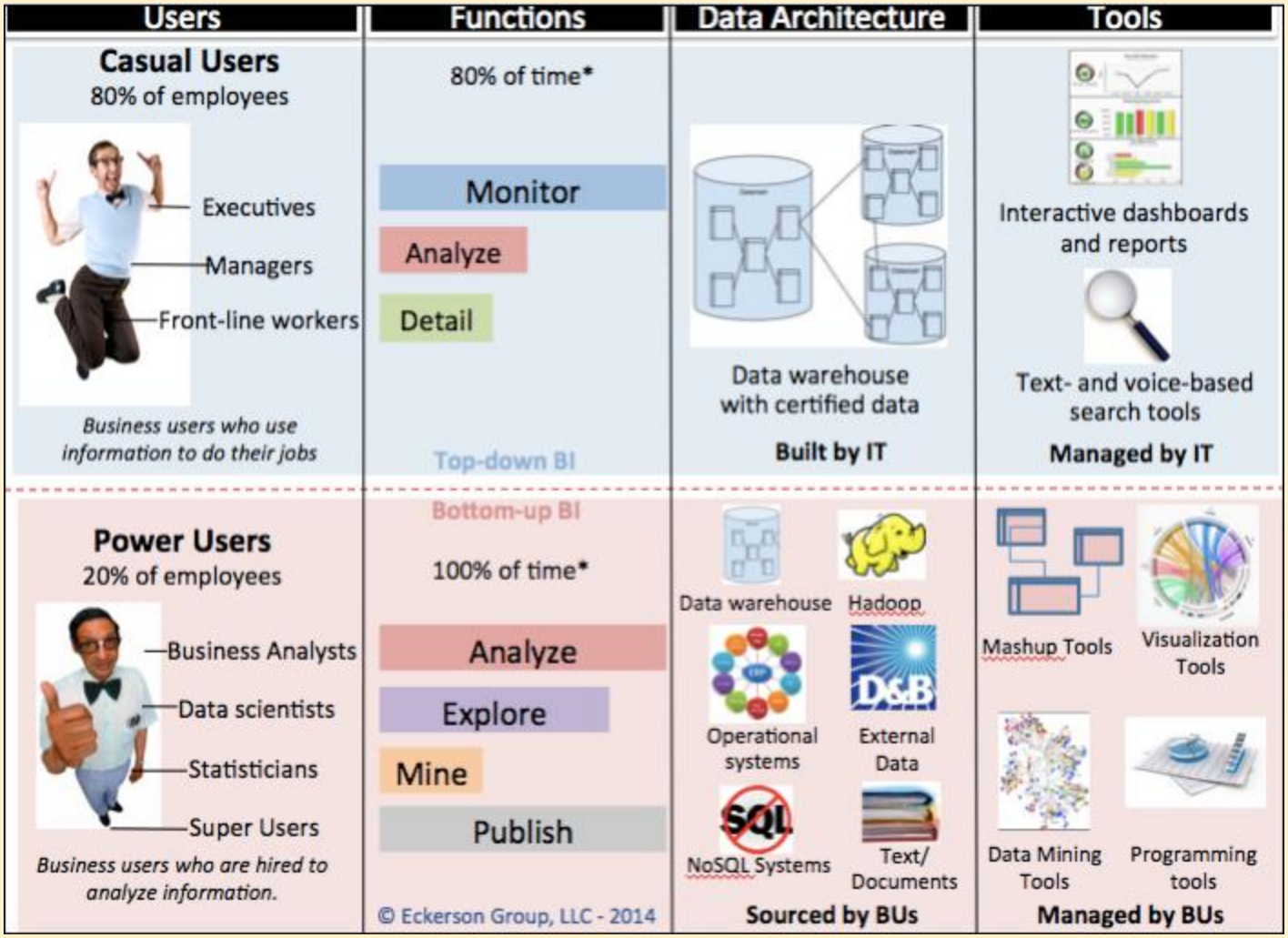


Technical vs. Business users



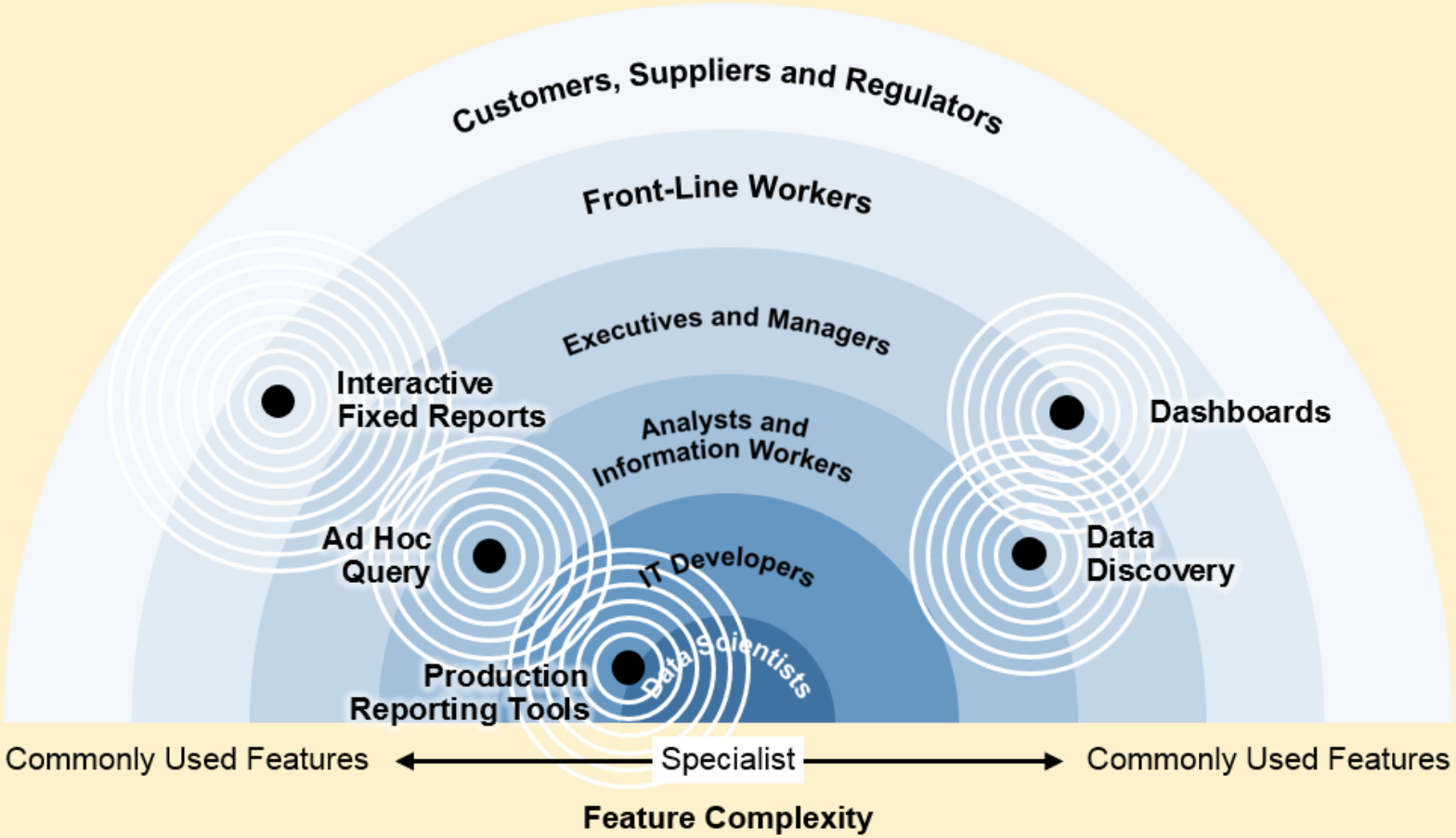
Figures originally from
http://www.bileader.com/Dashboards.html

Users Have Different Needs



http://eckerson.com/articles/part-iv-seven-keys-to-a-united-bi-environment
https://vimeo.com/68143902

The Fit between Tools and Users



Gartner Report,
Select the Right Business Intelligence and Analytics Tool for the Right User
Published: 23 May 2016 **Analyst(s):** Cindi Howson

BI/Analytics Application Areas

- BI/Analytics can be applied in all “businesses” (industries, functional areas, or domains) to drive “business” performance
  - Companies (for profit) and financially related
    - Retail, manufacture, real-estate, financial, sports, media, advertising, entertainment, healthcare, publication, energy, etc.
  - Public (non-profit)
    - Organization, institution, association, community, etc.
  - Government: citizen service, city planning, crime, immigration, etc.
  - Personal: personal health, exercise, learning, eating, power consumption, etc.
- BI can be applied at different levels
  - Strategic: focused on high level organizational strategies and directions
  - Tactic: focused on goals of a organization unit
  - Operational: focused on streamlining day-to-day operations.
  - https://www.business2community.com/business-intelligence/the-four-sides-of-business-intelligence-0548311#ycaoYFUR04W76YiY.97

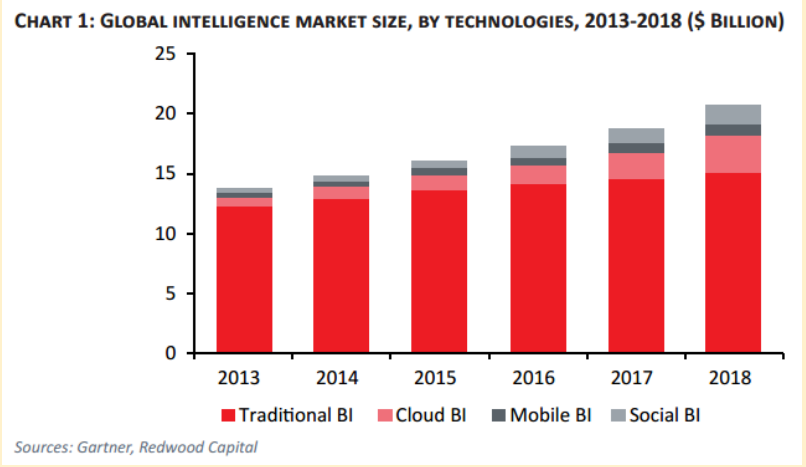
Sample BI/Analytics Applications

- Business management
  - Strategic planning
  - Performance management
  - Process intelligence
  - Competitive intelligence
- Marketing and sales
  - CRM
  - Customer behavior analysis
  - Targeted marketing and sales strategies
  - Customer profiling
  - Campaign management
  - Inventory management
- Human resource/capital
  - HR analytics
  - Talent management
- Project and program management
- Power and energy management
- Healthcare management
- IT management
  - Web analytics
  - App analytics
  - Security management
- Supply chain and Logistics
  - Supplier and vendor management
  - Shipping and inventory control
- Insurance
- Government
  - City planning
  - Traffic management
  - Urban Analytics
  - Power usage
- Education
  - Learning analytics
  - Student engagement and success
  - Institutional effectiveness
- Social analytics
- Sports and games analytics

BI Market

Figure from http://www.forbes.com/sites/louiscolumbus/2014/06/24/roundup-of-analytics-big-data-business-intelligence-forecasts-and-market-estimates-2014/

The global Business intelligence market size is expected to grow from USD 17.09 Billion in 2016 to USD 26.88 Billion by 2021, at a Compound Annual Growth Rate (CAGR) of 9.5%.
https://www.marketsandmarkets.com/Market-Reports/social-business-intelligence-bi-market-1048.html



Major vendors

Worldwide Business Analytics Software Market Shares, 2015, IDC July 2016

Worldwide Business Intelligence and Analytics Tools Software Revenue by Vendor, 2013-2015 ($M)

	Revenue ($M)			Share (%)			Growth (%)	
	2013	2014	2015	2013	2014	2015	2013–2014	2014–2015
SAP	2,163.1	2,305.0	2,013.5	17.0	17.0	14.4	6.6	-12.6
Microsoft	1,255.3	1,356.5	1,488.2	9.9	10.0	10.7	8.1	9.7
SAS	1,240.4	1,298.1	1,376.2	9.8	9.6	9.9	4.6	6.0
IBM	1,606.3	1,529.7	1,374.1	12.7	11.3	9.8	-4.8	-10.2
Oracle	1,045.5	1,059.2	1,038.4	8.2	7.8	7.4	1.3	-2.0
Tableau Software	225.2	399.3	630.6	1.8	2.9	4.5	77.3	57.9
Qlik	431.3	503.0	556.5	3.4	3.7	4.0	16.6	10.6
Total	12,696.5	13,561.1	13,961.8	100.0	100.0	100.0	6.8	3.0

Major Vendors/Products

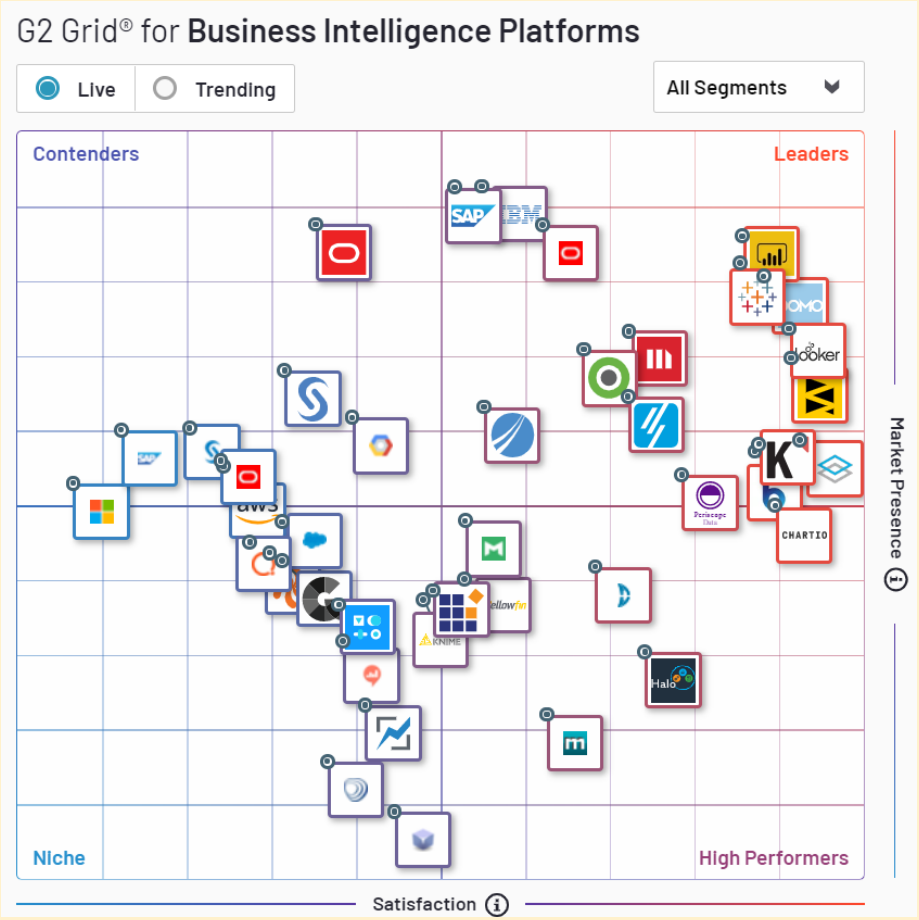
- Mega vendors provide complete solutions that cover full spectrum of BI processes.
  - Microsoft: SQL Server, Power BI, SharePoint, Excel
    - https://www.microsoft.com/en-us/cloud-platform/bi-analytics
  - SAP: SAP BusinessObjects BI
    - https://www.sapbi.com
  - IBM: Cognos, Watson
    - http://www.ibm.com/analytics/us/en/technology/business-intelligence/
  - Oracle: Oracle BI 12c
    - https://www.oracle.com/solutions/business-analytics/business-intelligence/index.html
  - SAS: SAS Enterprise BI
    - http://www.sas.com/en_us/software/business-intelligence.html
- More Other top BI tools, including Tableau, Qlik, etc.
  - https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms
  - http://www.capterra.com/business-intelligence-software/#infographic
  - https://www.softwareadvice.com/bi/
  - https://www.g2.com/categories/business-intelligence
  - https://www.bitool.net
- Open source tools, including BIRT, Pentaho, etc.
  - https://blog.capterra.com/top-8-free-and-open-source-business-intelligence-software/

Vendor Positioning

Notice this year Gartner put analytics before BI.

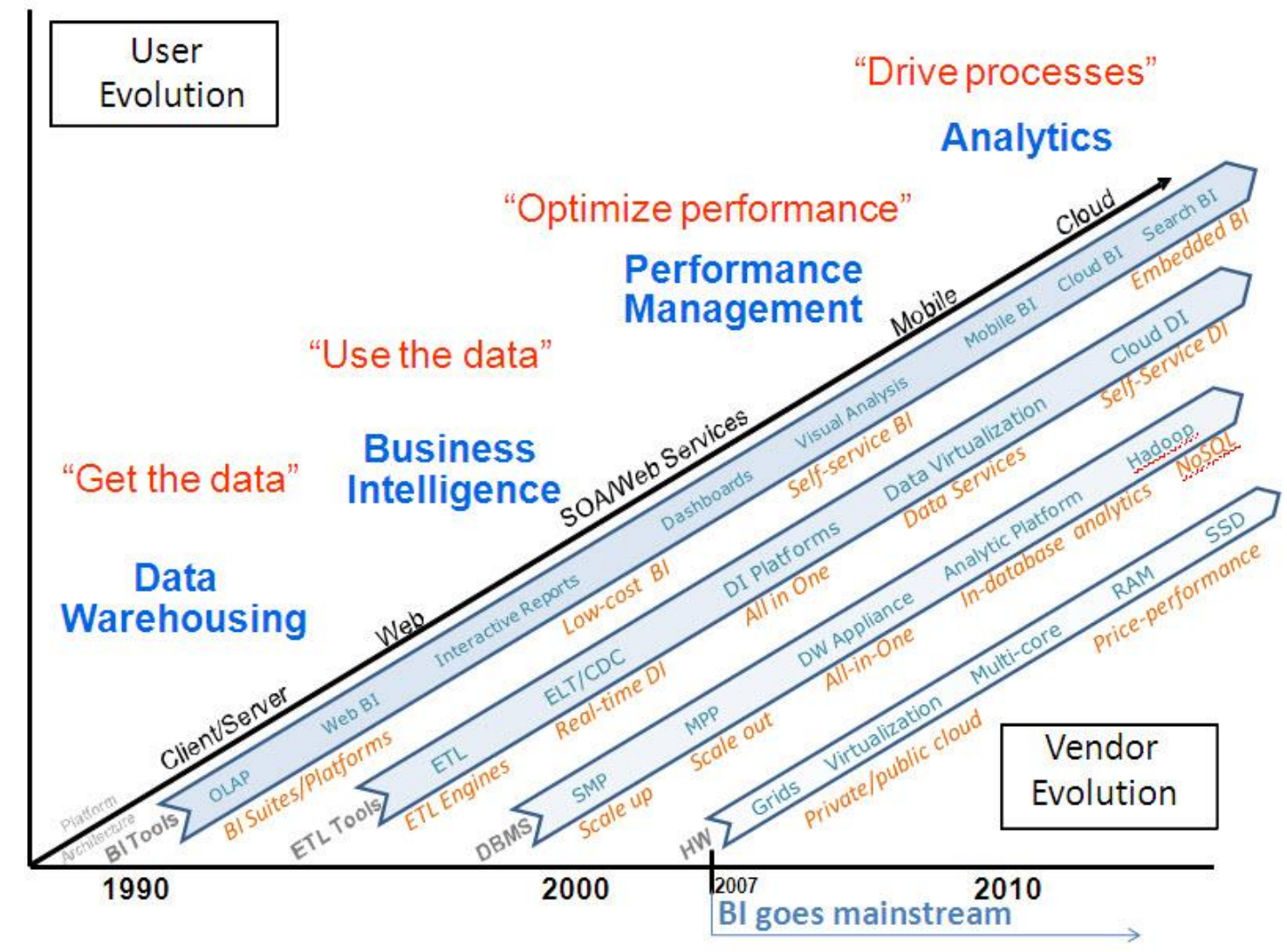
Gartner Magic Quadrant for Analytics and Business Intelligence and Platforms

- https://www.atscale.com/blog/analyzing-gartner's-2019-magic-quadrant-for-analytics
- https://www.atscale.com/blog/magic-quadrant-for-analytics-and-business-intelligence-platforms-2018



https://www.g2.com/categories/business-intelligence-platforms

BI Trends



http://www.b-eye-network.com/blogs/eckerson/archives/2011/03/bi_market_evolu.php

The Modern/New BI

- A modern BI platform supports IT-enabled analytic content development. It is defined by a self-contained architecture that enables nontechnical users to autonomously execute full-spectrum analytic workflows from data access, ingestion and preparation to interactive analysis and the collaborative sharing of insights. It moves from passive collection and use of data (reporting driven) to proactive generation of data (business development driven).

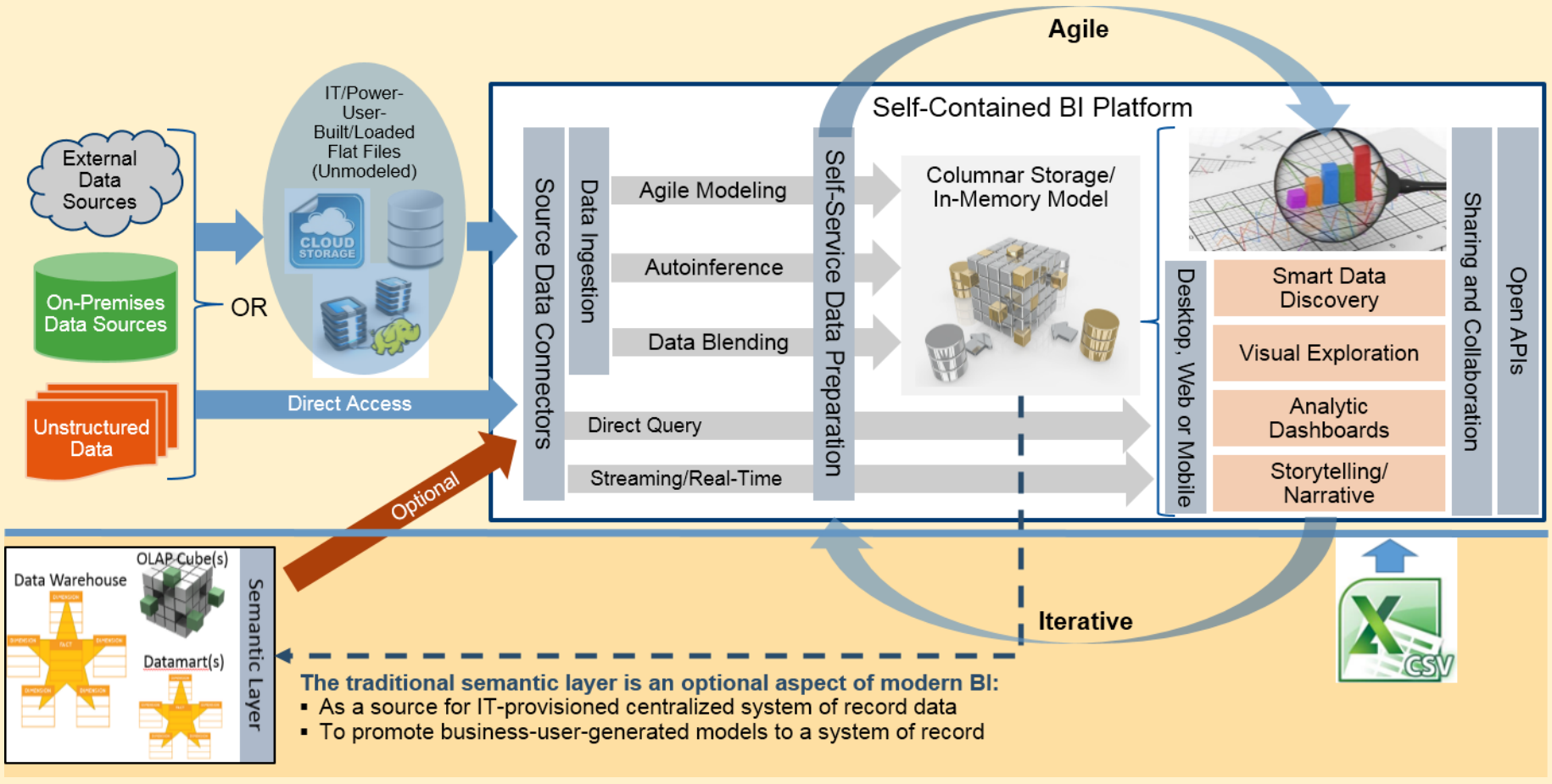
- By contrast, traditional BI platforms are designed to support modular development of IT-produced analytic content, and specialized tools and skills and significant upfront data modeling, coupled with a predefined metadata layer, are required to access their analytic capabilities.

- https://www.slideshare.net/Dataversity/analytics-business-intelligence-and-data-science-whats-the-progression

Technology Insight for Modern Business Intelligence and Analytics Platforms
Gartner Report, October 2015

Analytic Workflow Component	Traditional BI Platform	Modern BI Platform
Data source	Upfront dimensional modeling required (IT-built star schemas)	Upfront modeling not required (flat files/flat tables)
Data ingestion and preparation	IT-produced	IT-enabled (business-led)
Content authoring	Primarily IT staff, but also some power users	Business users;
Analysis	Predefined and regular reporting, based on predefined model	Free-form exploration, ad hoc analytics
Insight delivery	Distribution and notifications via scheduled reports or portal; passive collection and use of data (reporting driven).	Sharing and collaboration, storytelling, open APIs

A Changing BI Platform



The traditional semantic layer is an optional aspect of modern BI:

- As a source for IT-provisioned centralized system of record data
- To promote business-user-generated models to a system of record

Technology Insight for Modern Business Intelligence and Analytics Platforms
Gartner Report, October 2015

Some Notable Trends/Features

- Personal (self service) BI/Analytics
  - Self service enables nontechnical users to autonomously execute full-spectrum analytic workflows from data access, ingestion and preparation to interactive analysis and the collaborative sharing of insights
- Embedded analytics
  - Use of reporting and analytic capabilities directly in transactional business applications http://www.gartner.com/it-glossary/embedded-analytics/
- Search Driven BI
  - Build a report and charts on the fly, using web search style.
  - Most BI and analytics tools require users use a tool to “build” a query or visualization. When users type in a search phrase, such as "sales for New York 2016," and either an existing report or a visualization appears.
  - A search driven environment uses the power of Google-like search and build report and visualizations on the fly.
  - Some product examples: ThoughtSpot, SAP BusinessObjects Explorer, Microsoft Power BI Q&A, IBM Cognos Go Search, Oracle Endeca, Information Discovery and Information Builders Magnify.
  - https://www.youtube.com/watch?v=868-pR-cxZo
  - http://www.thoughtspot.com/what-is-search-driven-analytics
- Other trends
  - In-memory processing (in-memory OLAP): emerging technology for processing of data stored in an in-memory database. http://www.bi-dw.info/in-memory-olap.htm
  - Mobile BI/Cloud BI: new delivery method
  - Visual BI or visual analytics - http://www.perceptualedge.com
  - Augmented analytics and natural language processing
  - Advanced analytics (machine learning, deep learning, AI, etc.)
  - http://www.zdnet.com/article/is-the-business-intelligence-market-finally-maturing/
  - https://www.slideshare.net/TableauSoftware/top-10-business-intelligence-trends-for-2017

Personal (Self-Service) BI

- Self-service BI refers to BI products that are primarily configured and designed to be used by non-technical business users such as managers and business analysts.
  - Typical products: Tableau, Qlik, etc.

- Some features
  - Shifting focus from IT back to user
  - Supporting ad hoc analytic needs, hence more interactive and explorative
  - It still has fundamental BI components and provides BI capabilities, but they are more integrated than separated
  - Using self-contained in-memory processing
    - Easy and efficient data models (or even without one)
    - Quick query and analysis
    - Flexible data manipulation
    - Simple delivery and sharing
  - Independent but very often work with enterprise systems
  - Best for individuals or non-corporate environments

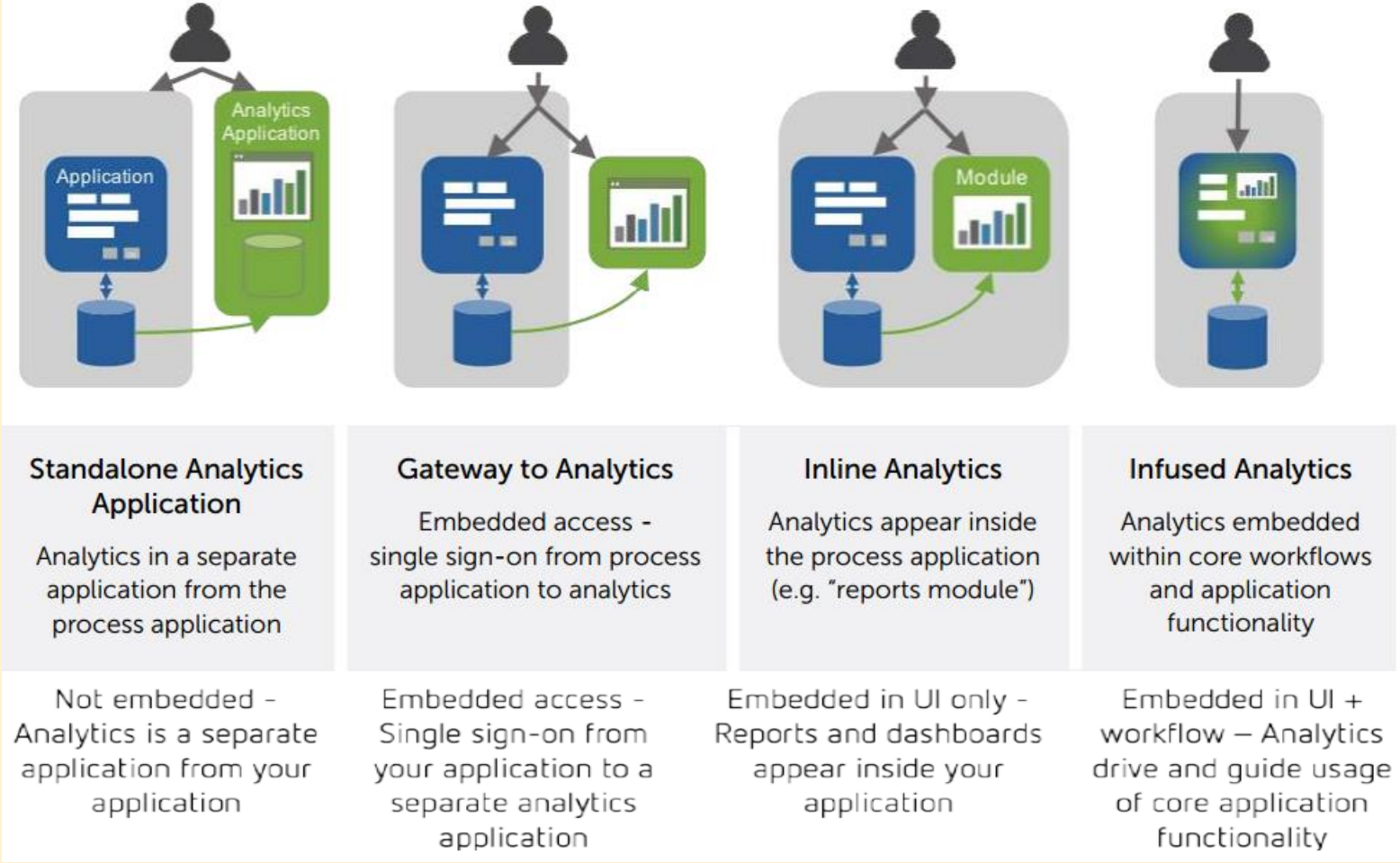
- Different levels of self-service
  - Started from client oriented report building and data visualizations, and eventually extended to analysis models, and finally to data discovery, preparation, and cleanse.
  - https://www.eckerson.com/articles/part-2-one-size-does-not-fit-all-customizing-self-service-analytics-for-business-users

- Dashboards, reporting, end-user self-service, and advanced visualization are the top four most important technologies and initiatives strategic to BI in 2018.
  - https://www.forbes.com/sites/louiscolumbus/2018/06/08/the-state-of-business-intelligence-2018/#b2fca2878289

- The global self-service business intelligence market to grow from USD 3963.04 million in 2016 to USD 10992.96 million by 2023, at a CAGR of 15.69%.
  - http://www.nbc-2.com/story/38414064/global-self-service-business-intelligence-market-2018-size-share-growth-trends-type-application-analysis-and-forecast-by-2023

Embedded Analytics

Integrating components into existing enterprise applications through APIs



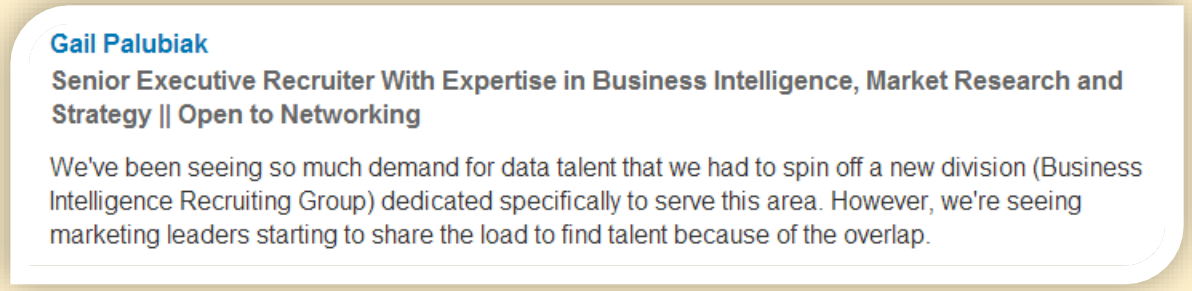
Built-in Analytics

Analytics is provided by the application itself

2016 State of Embedded Analytics Report by Logi Analytics
http://www.logianalytics.com/report/state-embedded-analytics-2016/

BI/Analytics Careers

- Typical BI positions
  - BI solution architects and integration specialists
  - Business and BI analysts
  - BI application developers and testers
  - BI system support specialists
  - Data warehouse specialists
  - Database analysts, developers and testers



Gail Palubiak

Senior Executive Recruiter With Expertise in Business Intelligence, Market Research and Strategy || Open to Networking

We've been seeing so much demand for data talent that we had to spin off a new division (Business Intelligence Recruiting Group) dedicated specifically to serve this area. However, we're seeing marketing leaders starting to share the load to find talent because of the overlap.

- BI jobs in Atlanta
  - https://www.dice.com/jobs?q=Business+intelligence&l=Atlanta%2C+Ga+Metro+Area

Critical Knowledge and Skills

- Three competencies
  - Technical, Business (management), Analytical
- Technical knowledge
  - Knowledge of database systems and data warehousing technologies
  - Ability to manage database system integration, implementation and testing
  - Ability to manage relational databases and create complex reports
  - Knowledge and ability to implement data and information policies, security requirements, and state and federal regulations
  - Knowledge of client tools used by business users
  - Knowledge of data models
  - Knowledge of programming tools used in analytics
- Solution development and management
  - Working with business and user requirements
  - Capturing and documenting the business requirements for BI solution
  - Translating business requirements into technical requirements
  - BI project lifecycle and management
- Business and Customer Skills and Knowledge
  - Effective communication and consultation with business users
  - Understanding of the flow of information throughout the organization
  - Ability to effectively communicate with and get support from technology and business specialists
  - Ability to understand the use of data and information in each organizational units
  - Ability to train business users in information management and interpretation
- https://www.datapine.com/blog/bi-skills-for-business-intelligence-career/

Sample Roles (from real world job ads)

Business Intelligence Specialist

- Maintain or update business intelligence tools, databases, dashboards, systems, or methods.
- Provide technical support for existing reports, dashboards, or other tools.
- Create business intelligence tools or systems, including design of related databases, spreadsheets, or outputs.

Business Intelligence Analyst

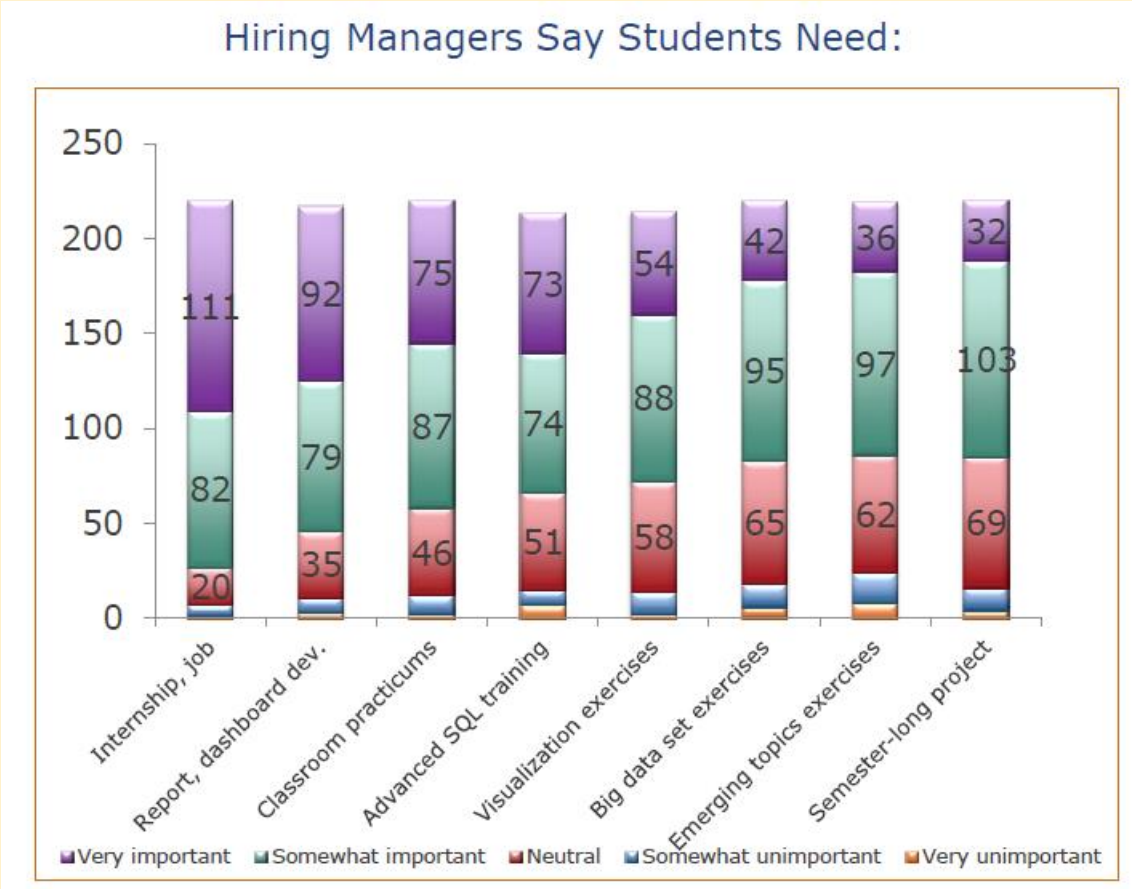
- Technical skill requirements
  - Works with business users to obtain data requirements for new analytic applications, design conceptual and logical models for the data warehouse and/or data mart.
  - Develops processes for capturing and maintaining metadata from all data warehousing components.
- Business skills requirements
  - Transform data into analytical insight and desire to leverage the best technique to arrive at the right answer.
  - Generate standard or custom reports summarizing business, financial, or economic data for review by executives, managers, clients, and other stakeholders.
  - Analyze competitive market strategies through analysis of related product, market, or share trends.
  - Collect business intelligence data from available industry reports, public information, field reports, or purchased sources.
  - Maintain library of model documents, templates, or other reusable knowledge assets.

https://dzone.com/articles/five-data-tasks-that-keep-data-engineers-awake-at

Business Intelligence Developer

- Business Intelligence Developer is responsible for designing and developing Business Intelligence solutions for the enterprise.
- Key functions include designing, developing, testing, debugging, and documenting extract, transform, load (ETL) data processes and data analysis reporting for enterprise-wide data warehouse implementations.
- Responsibilities include:
  - working closely with business and technical teams to understand, document, design and code ETL processes;
  - working closely with business teams to understand, document and design and code data analysis and reporting needs;
  - translating source mapping documents and reporting requirements into dimensional data models;
  - designing, developing, testing, optimizing and deploying server integration packages and stored procedures to perform all ETL related functions;
  - develop data cubes, reports, data extracts, dashboards or scorecards based on business requirements.
- The Business Intelligence Report Developer is responsible for developing, deploying and supporting reports, report applications, data warehouses and business intelligence systems.

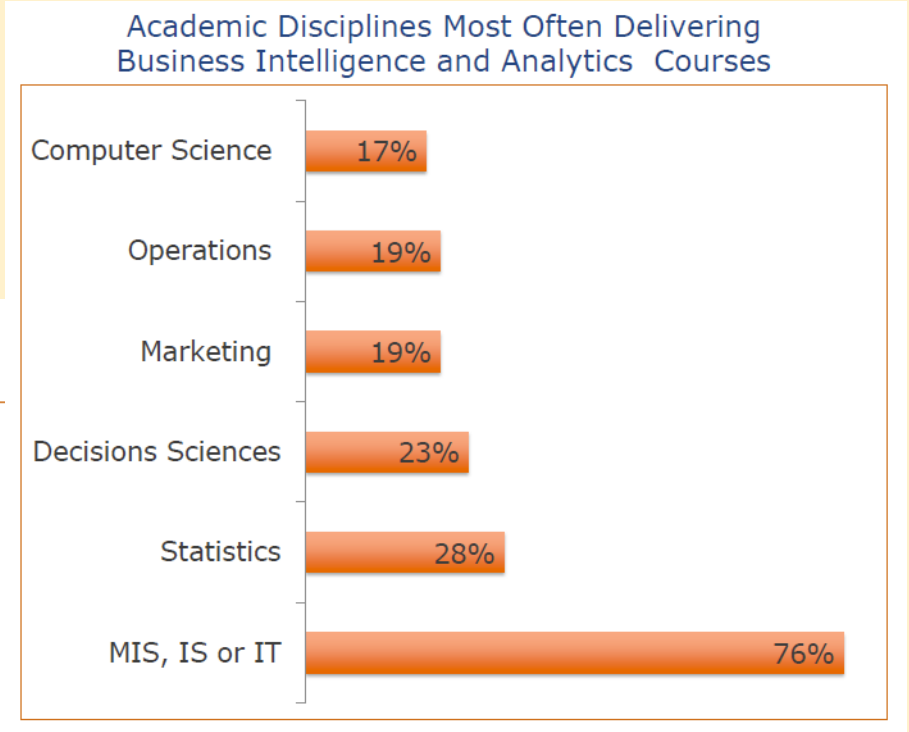
What are Employers Looking For?

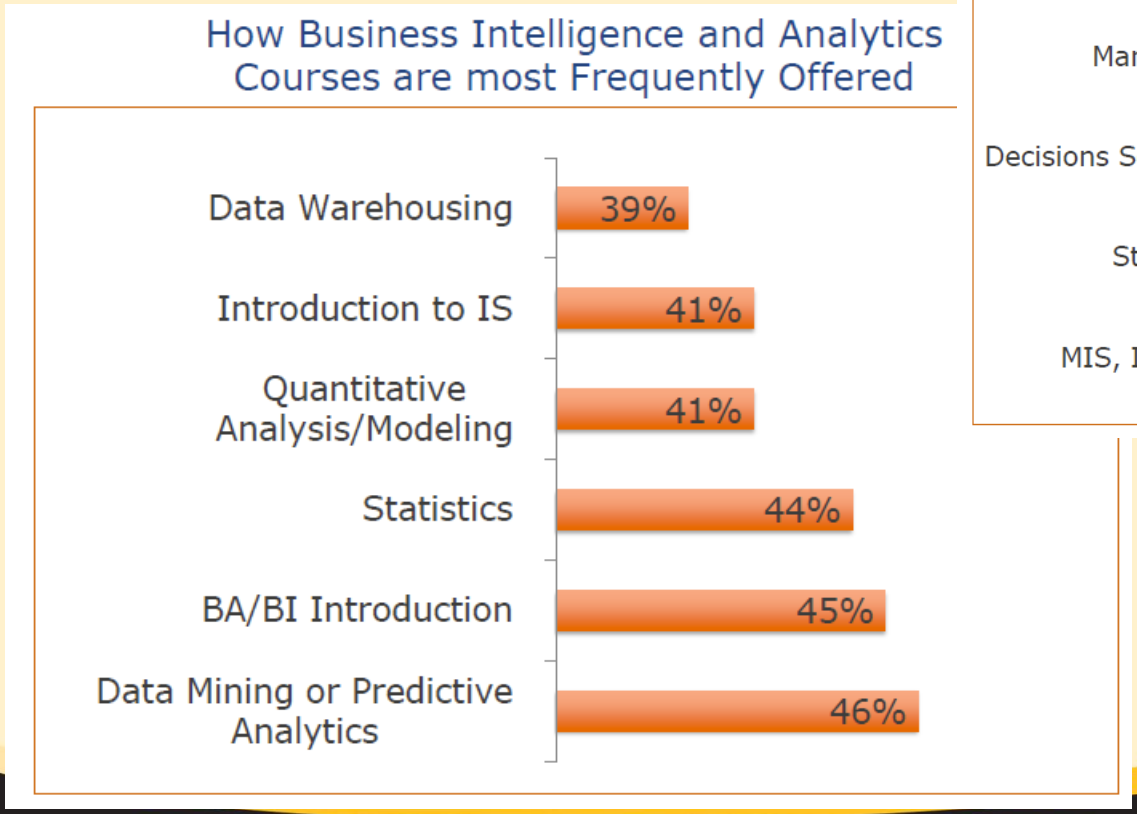


Data from BI Congress III Survey 2013

Where to Take BI Courses

Data from BI Congress III Survey 2013





BI/Analytics Education at KSU

- MSIT/BSIT
  - IT 6713 Business Intelligence http://jackzheng.net/teaching/it6713/
  - IT 7113 Data Visualization http://jackzheng.net/teaching/it7113/
  - Certificate on data management and analytics http://ccse.kennesaw.edu/it/programs/cert-dm.php
- BSIT
  - The new track on “data analytics and technology”
  - IT 4713 Business Intelligence http://jackzheng.net/teaching/it4713/
- Other departments
  - Ph.D. in Analytics and Data Science https://datascience.kennesaw.edu
  - ACS 8310 Data Warehousing
  - IS 8935 Business Intelligence - Traditional and Big Data Analytics
  - Certificate in High Performance Cluster Computing http://ccse.kennesaw.edu/cs/programs/cert-hpcc.php
- Lecture notes on BI and Data Visualization
  - https://www.edocr.com/user/jgzheng

Good Readings

- A quick, more conceptual and practical introduction of BI by Jared Hillam (Intricity): http://www.youtube.com/watch?v=LFnewuBsYiY
- BI intro video by LearnItFirst (focused more on the traditional BI; there are some good points which I do agree): https://www.youtube.com/watch?v=LhZX0MAYKp8
- Distinguishing Analytics, Business Intelligence, Data Science: https://www.dataversity.net/distinguishing-analytics-business-intelligence-data-science/
- An Overview of (traditional) BI Technology from CACM (premium magazine from ACM): http://cacm.acm.org/magazines/2011/8/114953-an-overview-of-business-intelligence-technology/fulltext
- Others
  - http://wps.prenhall.com/wps/media/objects/2519/2580469/addit_chmatl/TURBMC04_0131854615App.pdf
  - A Brief History of Decision Support Systems by D.J. Power: http://dssresources.com/history/dsshistory.html
  - Advanced Analytics and Business Intelligence: https://www.youtube.com/watch?v=oNNk9-tmsZY
  - History of BI (casual video with wacky visuals): https://www.youtube.com/watch?v=_1y5jBESLPE
  - https://www.datapine.com/blog/bi-skills-for-business-intelligence-career/
  - https://www.1keydata.com/datawarehousing/datawarehouse.html

Good General BI/Analytics Resources

- General BI web sites
  - BI and DW resource directory: http://www.bi-dw.info
  - BeyeNetwork: http://www.b-eye-network.com
  - Dataversity: http://www.dataversity.net/
  - Business intelligence resources: http://www.businessintelligence.com/
  - https://solutionsreview.com/business-intelligence/
  - DSS Resources: http://dssresources.com/
  - ACM techpack: http://techpack.acm.org/bi/
  - Data warehouse world: http://www.dwhworld.com
  - http://businessintelligence.com/
  - http://blog.capterra.com/learn-about-business-intelligence-resources/
  - https://www.itprotoday.com/business-intelligence
- General learning resources
  - https://www.1keydata.com/datawarehousing/datawarehouse.html
- Organizations and communities
  - The Data Warehousing Institute: http://tdwi.org
- Paid industry reports: you may get some free reprints from some vendors after registration.
  - Gartner annual report on “Magic Quadrant for Analytics and Business Intelligence Platforms”
  - Gartner report “Technology Insight for Modern Analytics and Business Intelligence Platforms”
  - The Forrester Wave™: Enterprise BI Platforms (two versions, one for on-premise and one for cloud)
  - Forrester Playbook: https://www.forrester.com/playbook/The+InsightsDriven+Business+Playbook/-/E-PLA940
- Industry experts and influencers
  - Howard Dresner: http://dresneradvisory.com
  - Wayne Eckerson: https://www.eckerson.com/blogs/the-new-bi-leader
  - Gregory Piatetsky: http://www.kdnuggets.com