Enterprise Data Science, Data Analytics & Hospital Operations Management:

Concepts, Values & Interconnections.

July, 2014

Alexander Kolker, Data Scientist

- **What do Data Science and Data/Business Analytics strive to accomplish?**

"Data is the energy source (fuel) of business transformation and Data Science/Business Analytics is the engine for its delivery."

Data focus:

How to capture (massive) amount of raw data in different formats (numerical values, descriptive texts, videos, audio, etc.), how to store, retrieve, and manage it.

Data / Business Analytics focus:

What data to collect, what metrics are needed to solve a business problem, how to clean the raw data to make it usable, what analytic method (algorithm) to use, how to turn analysis into actionable managerial decision.

Take-away:

Data Science and Data/Business Analytics strive to transform the raw data into **actionable** business decisions using a set of quantitative methods (algorithms), processes, and computer technologies.

Key points:

- Business analytics is NOT an extension or a side effect of collecting, keeping and presenting / querying data in general.
- Data have business value only in a **specific business context**.
- Currently there is a bias to looking at data itself instead of focusing on business context and data analysis methods and techniques.

Data Analytics:

- **Seeks to provide operational observations into issues that we**
  - **(i) either know that we know, or**
  - **(ii) know we don't know**

  by means of the simultaneous application of statistical methods, and computational algorithms to quantify system / process performance and forecast its future outcome.

Examples:

Descriptive analytics quantitatively describes the current main features of a collection of data.

Predictive analytics strives to predict / forecast relationships between known random variables or sets of data in order to identify (project) how an event will occur in the future.

Data Science, on the other hand, seeks to provide strategic actionable insights in the situations where we don't know what we don't know.

- The primary means through which actionable insights can be found are:
  - Predictive analytics in the area of causation
  - Prescriptive analytics, i.e. predictive + decision science
  - Machine learning

Thus, the overall value of **Data Science** is exponentially increased when it is coupled with its close 'cousin', **Data Analytics**, and both are integrated into an end-to-end **Enterprise Value Chain.**

Advanced Business Analytics

Predictive Analytics:

is the application of mathematical techniques and computational algorithms to predict data that are not currently available using some currently available data.

Predictive analytics supports business decision-making but it does not make decisions itself.

Example:
Given my current census, what will my census be tomorrow?

Predicting the tomorrow's census is helpful for making staffing decision but it does not make staffing decision in itself.

Advanced Business Analytics

Prescriptive Analytics:

Recommends business decisions based on
(i) business decision rules, or (ii) mathematical optimization models.

Example:
Given predicted tomorrow's census and acuity, how many nurses should I staff based on this information?

Decisions are made using:

Business rules, e.g. if patient : nurse ratio mandated to 3 then staff at the level: census/3.

or

Optimization, e.g. given a set of constraints, find the staffing level that minimizes the overall staffing cost within this set of constraints

Examples of Business Analytics Applications

- Time series analysis of the past / historical data to answer the question:

"What happened- trends, seasonality, auto-correlation, and how to extend the time series into the future?"

- Multivariate statistical analysis to answer the question:

"What factors played the most important role: significant vs. insignificant factors?"

- Predictive modeling / forecasting to answer the question:

"What will happen next?"

- Data Science / Prescriptive analytics to answer:

"Where are the opportunities- factors, resources, new directions?
What actions should be made to steer in the desired business direction?"

Examples of Business Analytics Problems

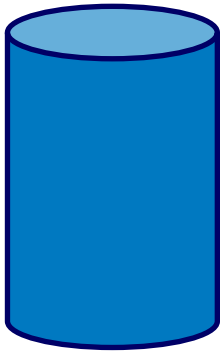
Business Problem	The most appropriate methodology
Long-term budgeted staffing: Given variable census and the cost of under- and over-staffing per nurse, what is the optimal staffing that minimizes the total staffing cost?	The "Newsvendor framework" with nurse: patient ratios, and/or Patient Classification System (PCS) *(Kolker, A., API Presentation, 2013)*
Capacity: Given variable patient arrivals, how many beds, procedure / operating rooms, or pieces of equipment are needed for different services?	Discrete Event Simulation *(Kolker, A., Healthcare Management Engineering: What Does This Fancy Term Really Mean? Springer_Briefs Series in Healthcare Management & Economics. Springer, NY, 2012)*
Staffing: Given variable patient arrivals, how many nurses, physicians and other providers are needed for a particular shift in order to best achieve operational and service performance objectives?	Discrete Event Simulation/Linear optimization *(Kolker, A., Healthcare Management Engineering: What Does This Fancy Term Really Mean? Springer_Briefs Series in Healthcare Management & Economics. Springer, NY, 2012)*
Scheduling: Given variable patient census, what is the minimal-cost staff schedule that includes mandatory days-off, full- and part-time, and shift rotations?	Linear optimization *(Kolker, A., Healthcare Management Engineering: What Does This Fancy Term Really Mean? Springer_Briefs Series in Healthcare Management & Economics. Springer, NY, 2012)*
Patient flow: How to organize patient move through hospital system without undue delay in order to achieve the system patient throughput goals?	Discrete Event Simulation *Kolker, A., Interdependency of Hospital Departments and Hospital-Wide Patient Flows. Chapter 2, pp. 43-63. In: Patient Flow: Reducing Delay in Healthcare Delivery. 2-nd Edition, R. Hall (Ed), Springer, NY, 2013.*

Examples of Business Analytics Problems (Cont.)

Business Problem	The most appropriate methodology
Resource allocation: Is it more efficient to use specialized resources or pooled (interchangeable) resources (operating/procedure rooms, beds, equipment, and staff)? Does it make economic sense to keep some patient service lines, or drop them at all?	Linear Optimization *(Kolker, A., Healthcare Management Engineering: What Does This Fancy Term Really Mean? Springer_Briefs Series in Healthcare Management & Economics. Springer, NY, 2012)*
Forecasting: How to forecast the future patient volumes, unit census (demand), or transaction volumes for short- and long-term budget planning and other planning purposes?	Markov chains; Time series analysis (stationary and non-stationary). Autocorrelations. Fractal Analysis. *(Kolker, A., Healthcare Management Engineering: What Does This Fancy Term Really Mean? Springer_Briefs Series in Healthcare Management & Economics. Springer, NY, 2012)*
Performance / Productivity scoring: Given multiple performance inputs and outputs for the unit, estimate its productivity score relative to other similar units, and decide on the factors that can be leveraged to improve the productivity score of this unit.	Data Envelopment Analysis (DEA) based on linear optimization. *(Kolker, A., API Presentation, 2014)*
Discharge Planning: Given variable patient length of stay (LOS) by acuity and/or other factors, predict the expected number of discharges in the next time period (say, 24 hours)	Estimation of conditional probabilities; hazard functions. *(Kolker, A., Healthcare Management Engineering: What Does This Fancy Term Really Mean? Springer_Briefs Series in Healthcare Management & Economics. Springer, NY, 2012)*

Overall Input – Output Process Flow

Unified Data Base as “Fuel” for all products



Engine for conversion “fuel” into justified actionable business decisions

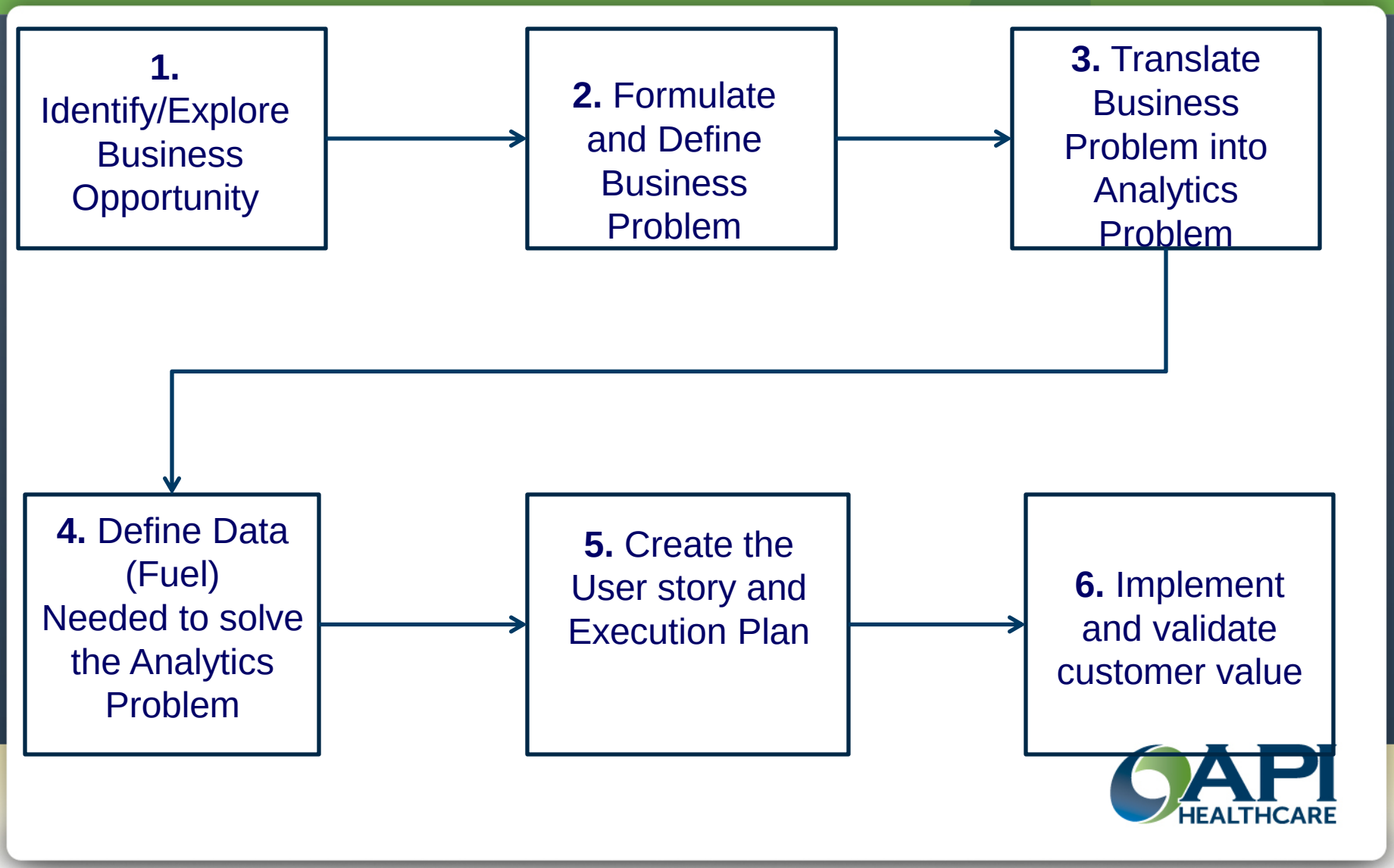
Data Science

- Predictive analytics
- Prescriptive analytics

Output

Actionable justified decisions

Main Steps in Business Decision-Making Flow

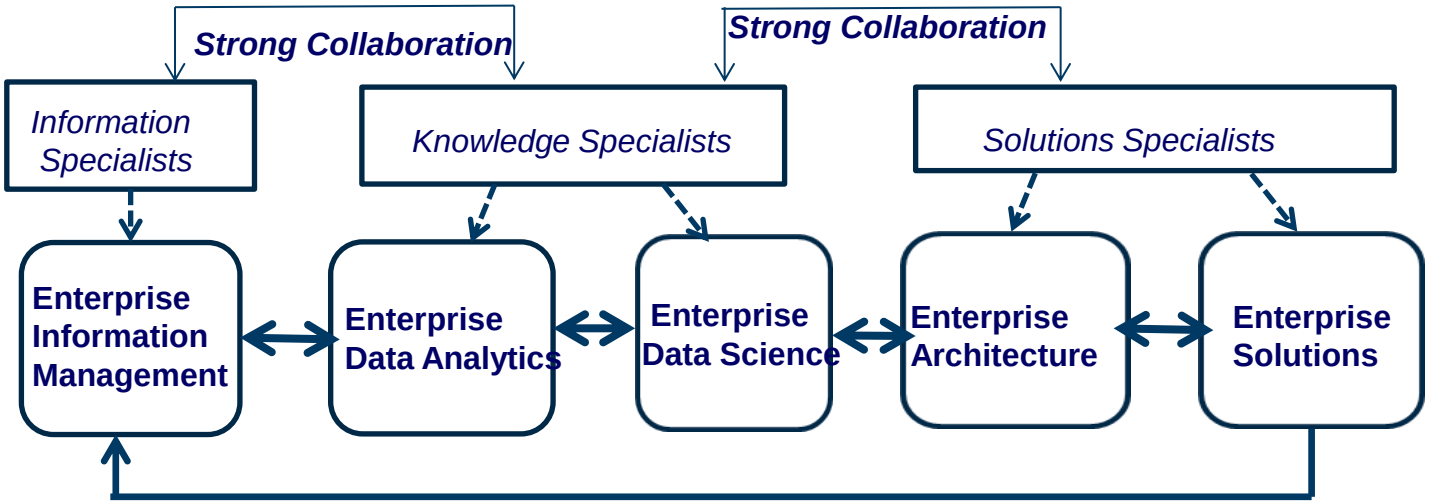


Example

- **1. Business opportunity:**
  Helping the customer to plan ahead in determining the optimized budgeted staffing
- **2. Business problem:**
  Given the unit variable shift census, what is the optimal budgeted staffing that minimizes the total costs of staff call-on (understaffing) and call-off (overstaffing)?
- **3. Translation into Analytics Problem**:
  "Newsvendor" framework that minimizes the total costs of under- and overstaffing
- **4. Define data (fuel) needs**:
  - PCS estimation for patient total hours of care per shift, or patient : nurse ratio broken by acuity
  - unit variable census
  - the costs of understaffing and overstaffing over the base rate

- **5. User Application for Optimized Staffing**
  - Recall Census & Acuity Data
  - Input Cost of Overstaffing and Understaffing
  - Calculating the Cost Ratio
  - Convert PCS hours of care into variable required FTE
  - Calculating the empirical cumulative distribution staffing function
  - Solving equation for the optimal long-term budgeted staffing

Enterprise solution data value chain



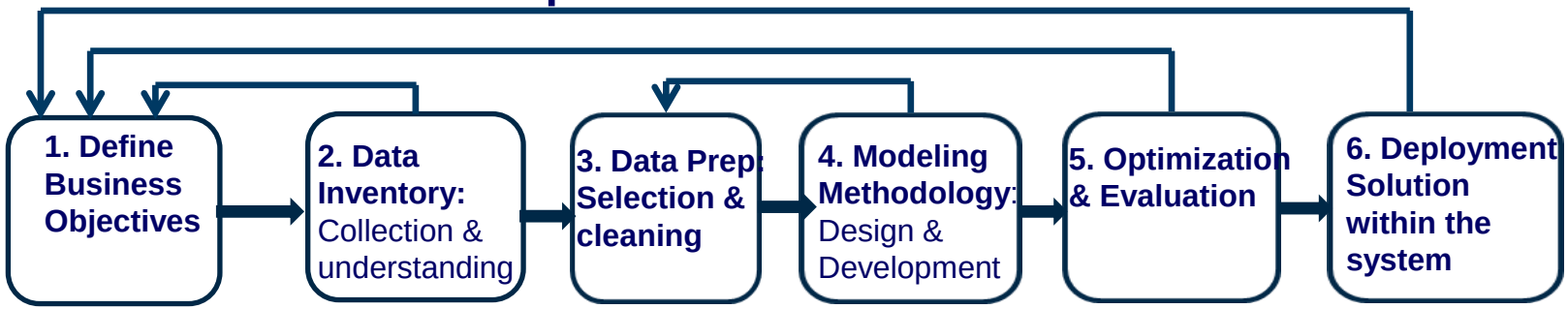
Capabilities for:
Managing current large scale data Assets.
Includes
-relational data base
-data warehouse
-data marts
-emergence of big data

Leverages data assets for day-to-day operational insights:
-staffing
-scheduling
-capacity
-forecasting
-patient flow
-productivity
-LOS

Seeks to exploit Information and analytic capabilities for developing enterprise strategic decisions

Seek to fuse operationally Into the capabilities of the larger enterprise through Architecture and Solutions

Enterprise Data Science



Determine type & volume of data required to meet business objective

-Define data selection rational
-Determine data availability
-If not available, develop data collection plan
-summarize available/ collected data
-verify data quality: how critical are missing values

-Select data to clean
-Integrate & format data for the next step

-Select Modeling techniques
-Generate test & validation designs
-Build model parameters
-Assess model performance

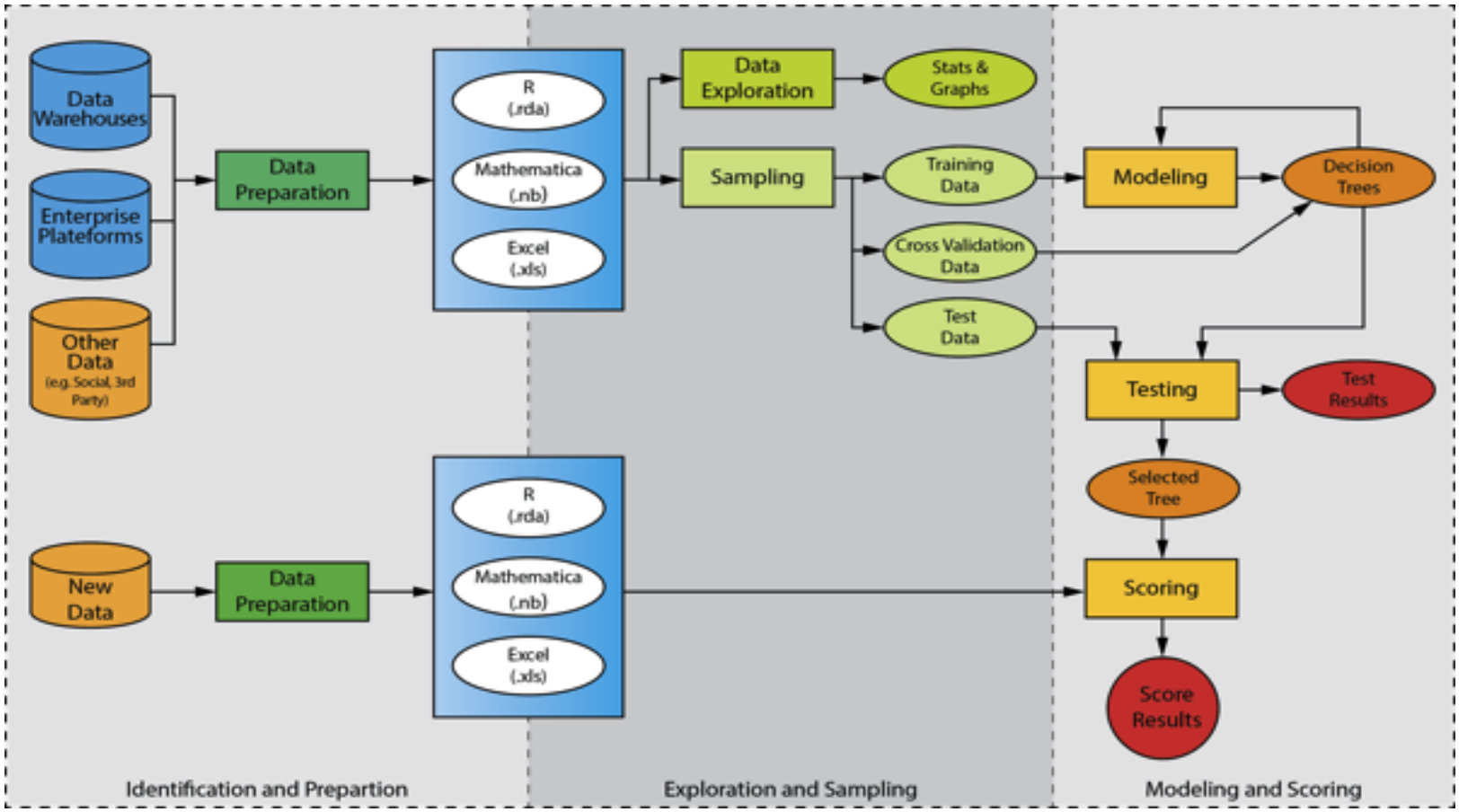
-Integrate individual modeling modules
-Evaluate results stability
- Approve model

-Assess deployment Environment
-Create model integration strategy
-Plan model deployment, monitoring & maintenance

Enterprise Data Value Circular Diagram

- The sequence of the phases is not strict and moving back and forth between different phases is typical.
- The arrows in the process diagram indicate the most important and frequent dependencies between phases.
- The outer circle in the diagram symbolizes the cyclic nature of the process itself:
  - It continues after a solution has been deployed.
  - The lessons learned during the process can trigger new, often more focused business questions and subsequent data analysis will benefit from the experiences of previous ones.

Data Science processes for an overall enterprise solution



Overall Key Points:

- Current Business Analytics is usually reporting or querying, i.e. presenting data in the convenient condensed form, such as graphs, dash-boards, etc. across various applications (clinical, revenue cycles, time & attendance, etc.).

- Reporting / querying provides some basic business information, but it does not directly help in business decision-making (except for simplest straightforward situations)

- Next Generation Business Analytics should provide a means for business decision-making based on predictive and prescriptive quantitative models using quantitative Data Science/Computer simulation methodologies.

- **Advanced Predictive and Prescriptive business analytics is a main component of the next generation Hospital Operations Management.**