Business Intelligence for Business Analysts

Methods, Model and Techniques for Project Success

Sean David Christopher

bookboon
The eBook company

SEAN DAVID CHRISTOPHER

BUSINESS INTELLIGENCE FOR BUSINESS ANALYSTS

METHODS, MODEL AND TECHNIQUES FOR PROJECT SUCCESS

Business Intelligence for Business Analysts: Methods, Model and Techniques for Project Success
1st edition

ISBN 978-87-403-1896-8

CONTENTS

ABOUT THE AUTHOR

With close to 20 years experience in the information technology field, Sean has led many IT projects including those in the workspace collaboration, database management and reporting domains, spanning software development, data modelling and business process improvement initiatives. Contributing a broad knowledge base over the last 13 years to business analysis activities, Sean has assumed an innovative approach to modernization of IT systems, applications and web services. Recently, Sean has helped the Standards Council implement a business intelligence framework on leave reporting for the Human Resources division, has positioned the SCC to implement a new expense management solution and a renewed standards collaboration management application, in addition to gearing up to present his findings for a new Standards Information Lifecycle Management system.

Sean possesses an honours degree from Carleton University of Ottawa and masters degree in business administration from Delta International University of New Orleans. He is also a certified specialist in SQL Server Business Intelligence, and holds a first degree black belt in Tae Kwon Do.

Sean has a passion for teaching. Having founded Smoothcube Business Analytics Corporation in 2014, the company attained endorsed education provider in 2017 from the International Institute of Business Analysis (IIBA), the parent international organization that promotes the business analysis industry around the world, and to which Sean is a member in good standing.

Sean's leisure moments include gardening, golf and having found a renewed passion for keelboat sailing, having attained his basic sailing standard certification. His dream is to sail from Halifax, Nova Scotia to Nassau, Bahamas. While preparing those logistics in the meantime, and readying for the keelboat cruising course, he can be seen on the lakes, rivers and bays of Ottawa in a kayak from time to time, testing the latest GPS navigation system, usually a Magellan or a Garmin device, and working on rope knots.

He lives in Ottawa with his beautiful wife Donna of 24 years and three gorgeous children Emily, Jacob and Meghan. His favorite quote is from Winston Churchill: "Success is not final. Failure is not fatal. The courage to continue on is what matters most." Mahatma Gandhi is also a source of motivation for Sean.

Sean can be contacted on the following social media outlets:
Twitter: s_d_christopher
Email: seandavidchristopher@gmail.com
LinkedIn: seandavidchristopher
Web: smoothcubeanalytics@gmail.com
Facebook: smoothcube

INTRODUCTION

The advances in business intelligence opens up unique opportunities for small, medium and large organizations. Poll any leadership team about why their company uses business intelligence and you could receive a wide range of answers. Some of those answers may include a need to discover new markets, identify trends in existing data, support better decision making in line with strategic goals or objectives, achieve more confidence in forecasting and budgeting with financial data, or simply to mine data to get to the level of granularity that will help uncover problems with business processes using key performance indicators to therefore help plan improvements or changes to those processes.

Data presented to end users in the context of business intelligence allows them to take informed actions with the results. Until the advent of business intelligence software and related tools, users and the technology team supporting the business units depended, and perhaps still depend on static, simple reporting as the norm for looking at data and making decisions based on that data. While these reports feature traces of advanced functions like pivot tables and charting methods, less than optimal decisions are made with static reports as issues like trend analysis, perspectives and key performance indicators are missing to satisfy the analytical demand made by management and leadership teams.

Often, reports are generated based on existing transactional system data which contain inconsistent data formats, bad data and a mixed array of alphanumeric text that either cannot be easily analyzed or leaves the user with little ability to assess the real intent of the data. Today, looking at cleaned data and drilling down into that data yields a wider spectrum of decision making capabilities using a system designed for analytical processing.

The business analyst plays a key role in the elicitation, documentation, analysis, modelling and presentation of requirements on a BI project. Acting as ambassadors of change with the goal to help organizations achieve greater success, the business analyst can leverage an assortment of tools and techniques to assess change and ultimately, affect change through the application of tested analytical principles at their disposal.

Our journey will look at how to support BI projects through the lens of the business and what requirements, methods, models and concepts to think about by taking a look at data at a strategic level, then move on to how analysis of needs are documented and augmented by modelling tools, then take a peek at transitional thinking and concluding with key messages and takeaways.

1 STAKEHOLDER ANALYSIS

Before we can begin to elicit requirements from key business stakeholders about strategy and to undertake the activity of an analysis of the requirements gathered from having talks with stakeholders, the task falls to the business analyst to identify the key stakeholders and their roles.

One of the methods that can help the analysis of stakeholders is the study of corporate documentation such as organizational charts, corporate plans and annual reports. While the analysis is not limited to these types of formal documentation, they do represent a good starting point to uncovering the key people involved in a business intelligence project.

Upon conclusion of an initial review of the documentation, we can begin to categorize our list of stakeholders. The categorization of stakeholders will assist with our business analysis approach when interacting with them to draw out their data requirements. We can also begin to draft our business analysis approach to communicating with them as well.

1.1 STAKEHOLDER CATEGORIES

Stakeholders can be grouped in three categories: executive, management and operational.

Designing the approach to each targeted stakeholder will require some critical thinking, formulation of a line of questioning and different elicitation techniques in order to draw out the requirements and focus the discussion on the main aspects of their roles in the organization.

Below is a table that identifies the key stakeholder categories and the descriptions of the focus, questioning and elicitation technique upon which to concentrate the interactive approach to engage with people slotted into those categories.

Stakeholder Category/Approach	Executive	Management	Operational
Focus	High level business and strategic approach	Lines of business spanning multiple users, typically broken down by divisions or branches.	Staff that support the lines of business and interact with data the most, prepare reports for management and executive staff.
Questioning	Ask specific questions about data to which they need access in order to help them make executive decisions that affect their business.	Geared towards data they need to make decisions about process improvement, product or service delivery level adjustments.	Span the nature of the data, the types of data, the input and output process facilitated by applications and user interfaces.
Technique(s)	Interviews	Interviews, Job Shadowing, Document Review	Workshops, Focus Groups, Interviews, Interface Analysis

Table 1 – Stakeholder Category Approaches

1.2 THE TACTICS EXPLAINED

Executives take a strategic approach to a business. Planning strategy and tactics is a key part of their work in addition to oversight of their direct reports and supporting staff. While time is a premium for all staff, it is very important that good use of an executive's time is kept in mind and planned accordingly before meeting to discuss data requirements and to maximize the output of stated data requirements.

Management can be responsible for collecting reports from a wide variety of direct operational staff. Typically, managers and directors who conduct the activities of specific divisions and branches take a process approach, ensure standards, policies and procedures are carried out and that access to information is available, ready for distribution, processing and analysis.

Operational staff has a more direct, intimate knowledge of data collected from a variety of sources such as application user interfaces and other data entry mechanisms, such as forms used to capture purchase requisition requests or technical problems with a piece of hardware or software. Generally staff are tasked with preparing reports using a variety of reporting tools at their disposal, such as Excel or Microsoft Access. Often reports are generated for a higher purpose such as informing management on how their business units are performing against stated corporate objectives and whether specific business targets are being met throughout specific points in the year.

Ultimately, executives need to be informed too and require access to data to help them make decisions about process improvement, strategy changes or tactical shifts in how operations are carried out in order for the organization to either stay on track for meeting given goals and objectives or to get back on track if there is a deviation. Advanced reporting will help all personnel make adjustments and take action to change, update or modify how things are done.

Face to face meetings with stakeholders will also help the business analyst understand the limitations and boundaries that exist within the organization in terms of data culture.

Limitations and boundaries may be related to people, processes and technology.

1.3 LIMITATION ASSESSMENT

The analyst should attempt to draw out requirements that can be written to address issues related to limitations about:

- Staff understanding of the purpose and intent of the data with which they interact.
- The data collection process including the specific technologies or applications used to collect, store and manage data.
- Understanding about data sources used from outside the organization. Sources of data from outside the organization include "open data"[1], such as demographic or GDP data from government bodies and agencies like the United Nations
- Specific data improvement feedback mechanisms in place. Data improvement processes provide users with a technical support structure to identify issues with data formatting, quality and integrity.

- Knowledge transfer between supporting business units like Information Technology (IT) and staff in other operational business units about data reporting capabilities with current technology. Often, users are unaware of the capabilities of a given reporting mechanism and the solutions that can be leveraged within the existing data management environment. While the technical capability is not the limitation, the limitation is knowledge about the capability.
- Resource allocation models for people creating reports and people consuming reports including historical issues concerning getting data out to people on time and in the correct format.
- The data culture such as whether people understand analytics and what it is designed to do for a business. This is particularly important for companies that work in high intensity industries where competition is fierce and decisions made can either lead to a desirable situation such as increased market penetration, increased sales, and more customers or an undesirable one where the company could go out of business.
- Perceptions about the current state and functional limitations of data repositories, availability of data and data quality in general.

1.4 BOUNDARY ASSESSMENT

Limitations can occur within specific business boundaries. Limitations act as constraints within a boundary space, often dictated by motivation, money and capacity. Boundaries are the outer limits of a specific business space. Business are always pushing boundaries in order to achieve greater things, but can face different types of limitations within that business space.

Boundary assessment can yield insights into:

- The process for collecting and reporting on data and whether other people within the organization know about the ownership of data;[2]
- Ownership of generating reports as part of daily responsibilities;
- Whether reports are shared across divisions or branches, operational units or even between executives;
- Assessing whether an attempt in the past of the organization trying to take a holistic view of their data and come up with an enterprise approach to data availability, including the adoption of the principle of "single source of truth" for data. Single sources of truth remove duplication of data across applications and systems and helps to streamline access to data using one single source of truth;

- Determine if there is a culture of data hoarding or hijacking. The adage of "information is power" holds true here. Often, business cultures let hoarding of data happen whereby the sharing of information may pose a front to a staff members' unique perspective of a business and therefore motivate the staff to yield little information or nothing at all to other staff. Hijacking in this case is the taking over of another business units data with the intent to hoard. Such a situation may arise if divisions have cross functional responsibilities. This type of activity may occur in an organization as politics plays a huge role often in internal decision making. The psychology of human behaviour suggests that power is a huge motivation for some people and the knowledge or perceived gains about controlling information or "data points" contributes to that power. The point here is that the business analyst understands the emotional and human factors that contribute to decision making within the divisional boundaries of an organization; and
- Timing of access to data and issues that have arisen in the past related to getting access to data in a timely manner.

The graphic below shows what limitations would look like within a set boundary. Boundaries can be expanded in some situations where the limitations can be lifted, thus expanding the limitation sphere and once enough of the limitations are lifted, the boundary would start to shift outward as well.

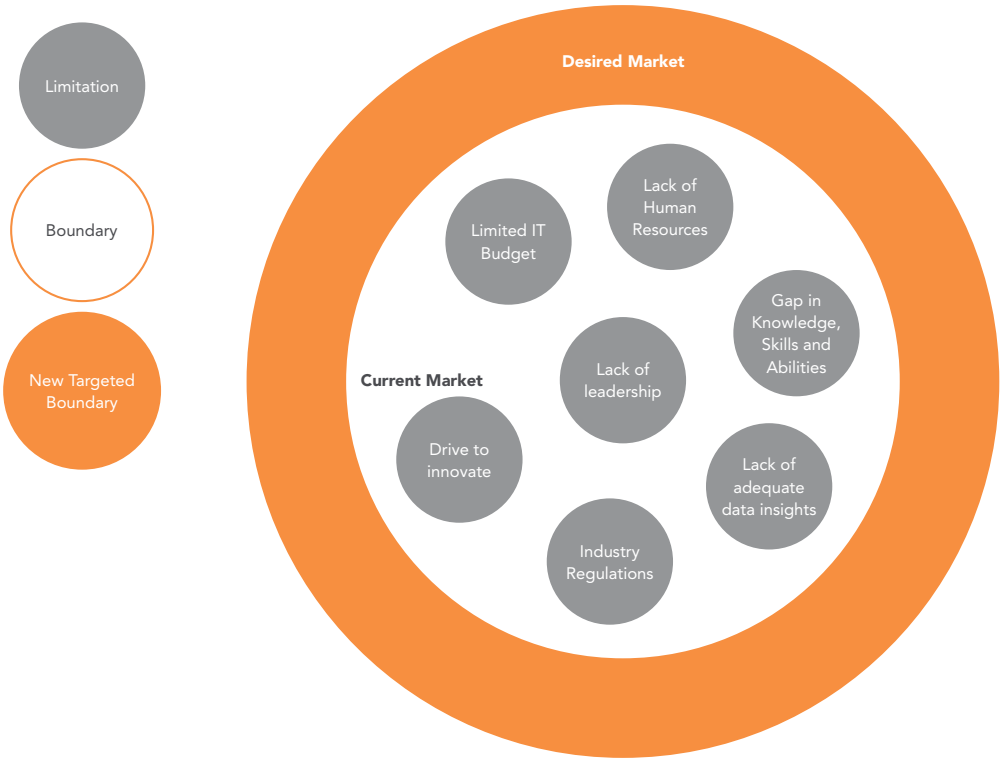

Figure 1 – Sample Limitations and Boundaries graphic

2 ASSESSING THE ENTERPRISE ARCHITECTURE

Armed with an assessment of potential stakeholders, the next activity is to engage in enterprise analysis (EA), whose tactics and methods can be applied to any project. EA is particularly important to the analysis of data and business intelligence because it provides the business analyst with the concepts needed to conduct a high-level review of an organization's business makeup.

EA consists of high level activities, approaches and tasks designed to help a BA assess a company using a holistic view. The result of these activities are outputs that can be used to dive deeper into the core data needs of an organization.

The next few pages will highlight some of the finer points of enterprise analysis. But first, consider the graphic on the next page.

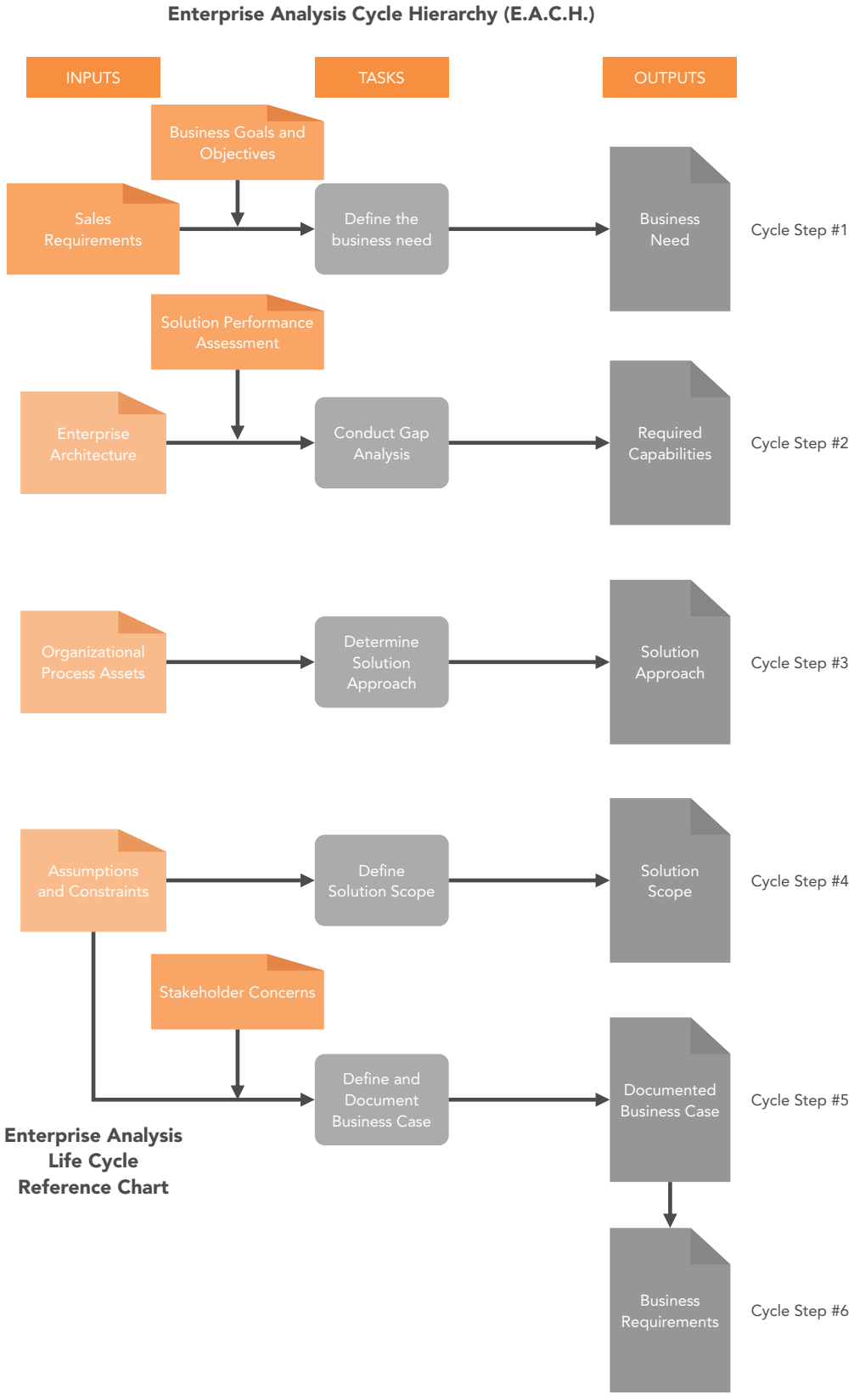

Figure 2 – Enterprise Analysis Cycle Hierarchy

The EACH model of enterprise analysis will help you in your journey to learning the requirements phases to support a business intelligence project. Each phase is explained as it relates to the BI project. After this chapter, you will be able to:

- Understand the six stages of enterprise analysis;
- Connect business requirements to strategic business components; and
- Learn the delineation between business need, business case and business requirements.

The EACH model has 3 components: a series of inputs, the activity involved in the EA and the outputs.

Inputs (in blue or light gray in the EACH template) at this point of our data project may include stated requirements, which are requirements that have been put down in writing or on some other documentation mechanism as provided by our stakeholders. The analysis may also possess some assessments about past data projects or if the BI project is an enhancement to an existing solution, the performance and functional observations of the existing solution. Other inputs include documented IT architecture, stakeholder concerns, organizational process assets like business flows or business process management diagrams and / or assumptions and constraints. Constraints are discussed in a later chapter.

The cycle identifies the processes (in yellow) to:

- Investigate the business direction and need;
- Study the gaps in terms of where the business is today (The "As Is" state) and where the business wants to go (the "To Be" state). This type of study is also called situational analysis assessment;
- Validate the solution approach;
- Focus the solution scope;
- Formulate the business case; and
- Write the business requirements.

At the end, all of the inputs and cycle activities yield outputs as identified in green in the EACH template.

2.1 INVESTIGATE THE CORPORATE DIRECTION AND NEED (EACH CYCLE STEP #1)

A study of the business direction can yield insights into:

- Process changes, technical implementations, training or education gaps, business opportunity or strategic shift;
- Monitoring changes in the legal or regulatory environment;
- Understand an organization's industry;
- Assess the competitive environment;
- Techniques include:
    - Identify business needs by talking to management, strategists, key stakeholders as we have discussed in Chapter 1. EA activities may also yield additional stakeholders as a benefit to the analysis; and
    - Root cause analysis and problem restatement. I like using the "5 whys" of business analysis. This includes taking a stakeholder's answer and reformulating it to ask why again, and so on. These types of questions can yield different answers to get to the true cause of a problem and often guides the analysis down a whole different path of inquiry.
        - Example
            - **Analyst**: Why do you need better data analytics? **Stakeholder**: Data analytics will help us determine which European markets to penetrate for our products.
            - **Analyst**: Why do you need to penetrate European markets? **Stakeholder**: Our existing analysis has shown we have saturated our current market.
            - **Analyst**: Why has the current market reached saturation? **Stakeholder**: Our products have outgrown the current market demographic.
            - **Analyst**: Why have your products outgrown the current market demographic? **Stakeholder**: Our customers have indicated a desire to use other innovative products.
            - **Analyst**: Why have your customer indicated a desire to use other products? **Stakeholder**: We failed to reach out on a routine basis with them and we have lost their loyalty.
            - In the above example, while the company may have a strategic goal of moving into new markets, the real cause has to due with a **business process failure** to conduct outreach with existing customers about their products!

- The breakdown of a need into manageable parts which will help focus requirements gathering activities. This technique is called functional decomposition.
- Acquire and study documents such as corporate plans, strategic plans, and business plans.

2.2 STUDY THE GAPS (EACH CYCLE STEP #2)

Gap analysis helps with identifying the actions or changes needed in order for a business to move from its "As Is" state to its desired "To Be" state.

Capability gaps can emanate from an analysis of:

- Human resource knowledge, skills and abilities. In other words, does the organization's staff possess the required skills available now? Or does the business need to acquire the skills from outside the company? Is training needed to bring the skill set up to the level needed to help move the business along its strategic path?
- Technical environment, including existing software features and functionality in an assessment of outdated, or legacy systems.
  - Does functionality exist within the current solution?
  - Are features not "turned on" that could be leveraged?
- Internal business processes.
- Create the "gap assessment" to augment the "As Is"/"To Be", situational or positional state
- Techniques:
  - Study all possible documentation related to the needs assessment.
  - Document strengths, weaknesses, opportunities and all threats associated with the business. A SWOT analysis contributes immensely to identifying problem areas or capability shortcomings, what a business does well and the nature of its competition.

2.3 VALIDATE THE SOLUTION APPROACH (EACH CYCLE STEP #3)

Good business analysis will always include a process for ensuring that a solution approach is not only understood by stakeholders but validated, meaning everyone understand the intent and purpose of the approach.

Some of the tasks or activities in the process include:

- Explore options about how the solution could be implemented.
- Avoid falling into discussions about how the organization will satisfy the business need, technically or resourcefully. Focus on the "what".
- List options that all agree to, including:
    - Technical:
        - Open Source (OS), Off the Shelf (OTS)
        - In house development, Outsourcing
        - Software as a Service
        - Service Oriented Architecture (SOA)
    - Resource
        - Train and educate
        - Procure outside skill sets
        - Assess partnership opportunities
        - Leverage existing knowledge, skills and abilities
    - Business
        - Process renewal, change or update
        - Organizational structure changes or shifts

2.4 FOCUS THE SCOPE (EACH CYCLE STEP #4)

Requirements scope and its management are crucial factors for project success.

Some of the tasks and activities include:

- Focus on the business analysis area of assessment that will be in view that will help to deliver an eventual solution that will satisfy the business need.
    - Tip: Keep the sphere of focus as narrow as possible but within boundaries of the solution's new capabilities
    - Tip: Explain scope phasing to stakeholders if necessary as a business intelligence solution implementation may require stages.
- Technique: scope modelling and user stories
- The challenge is in the term "solution"
    - Avoid confusing the term with a final deliverable product.
    - Solution scope is a conceptual term and is used to assess the work involved in requirements elicitation, documentation, management, and prioritization.
- Scope can be small or large but needs to be manageable.
- Tip: Get the documented scope approved but have a change management plan in place.
    - Scope will determine the sphere of elicitation for functional, qualitative and other requirements needed to support the BI project.

2.5 FORMULATE THE CASE (EACH CYCLE STEP #5)

Making the case for any project is critical to advancing the any data initiative forward.

Some of the tasks and activities included in this process are:

- The sponsor needs to know:
- Costs, Net Present Value, Recovery period, Discounted rate, Rate of Return
    - What are the tangible and intangible benefits to an organization engaging in a business intelligence initiative?
    - What is the problem being solved and how will it be solved?
    - What are the associated technical, business, and implementation risks?
- Mitigation strategy or plan
    - Know your audience for the business case
        - Tip: seek advice from Finance, Risk Managers
- Is a project already underway?
    - What if we don't proceed with the project?
- What resources are required?
    - Time, People, Equipment, Money
- Does the project's need align with the overall business strategy?
    - If so, demonstrate. If not, review and validate.
        - Tip: Obtain case templates if available from other projects; talk to project managers

2.6 WRITE THE BUSINESS REQUIREMENTS (EACH CYCLE STEP #6)

Once the need is found, the gaps are identified, the approach and scope are conceptualized and the case is written and approved, we have business requirements in which to feed other business analysis activities

Business requirements are what the business needs to get to its desired state and may consist of stated stakeholder requirements. Business requirements are not in their entirety the functional, qualitative, transitional, technical, stakeholder or system requirements of the project as these will be identified as inputs are assessed and models are created to help explain data needs for the BI project. Business requirements are assumed to mean stated, validated, modelled and approved requirements.

3 STRATEGIC REVIEW

A key component of enterprise analysis, as we discussed in Chapter 2 is strategic review, part of EACH cycle step #1. BI projects should include time spent on analyzing the overall direction of the company as strategic directives often yield insights into the higher-level data needs of an organization.

Research conducted by Thomas Davenport, Jeanne Harris and Robert Morison, knowledge leaders in the field of business analytics, found that forty percent of decisions made by managers are taken from experience or gut feeling and not by looking deeper into what the data is telling them.[3]

Some of these reasons, based on personal technical experience include insufficient facts influencing management to avoid using reports as background, corruption of data leading to a lack of faith in the integrity of the data source(s), erosion of trust in the reporting mechanism or simply the reliance on similar, past decisions as indicators of success.

Exercise
From a data perspective, document your reasons as to why business leaders make gut decisions. What is your evaluation of the potential gains and losses from making decisions using "gut feeling"? To understand these reasons and to weed out contributing factors, a schedule of meetings with key stakeholders is needed since we would have identified our data stakeholders. The goal of these meetings is to develop an understanding of the business requirements by first analyzing the strategic goals and objectives of the organization, helped by an initial enterprise analysis. Another technique that we can employ is document review. The outputs of these two activities is to map the goals and objectives to higher level data needs that will contribute to the measurement of not only how their business is doing at various stages of their operational life cycle but also, and more importantly to improve the way decisions are made with data at their disposal.

Table 2 – Exercise on reasons leaders make gut decisions

3.1 DOCUMENT REVIEW

A good place to start to assess high level goals and objectives is to review the corporation plan of an organization. The corporate plan generally contains not only the strategic objectives, or the quantitative measures of the organization but also the goals which are referred to as the qualitative dimensions describing specific business targets.

Practical Example

South African Airways has 5 main objectives as stated on its corporate web site. One of those objectives is to "Achieve and maintain commercial sustainability." While our analysis may consider it a goal, as it describes a qualitative dimension of their strategic plan, a qualitative objective is "a reduction in CO2 emissions of 50% by 2050, relative to 2005 level".[4]

Table 3 – Real world example of a stated objective

Ideally, our analysis should factor in what the business wants to achieve not only in the next 1 to 3 years, but in the next 5 to 10 years if possible. As seen from the South African Airways example, our analysis already yields a stated objective of reducing CO2 emissions by 50%, 30+ years hence. We have, as analysts a data metric that can be used as a benchmark or milestone as the organization takes measurements of its corporate performance over time, such as quarterly or semiannually.

One challenge to strategic review is the inertia an organization may have whereby they have trouble or difficulty in projecting stated goals and objectives. While business analysis could literally spend a lot of time on strategy analysis, as an extension to an enterprise analysis (the corporate objectives) an opportunity presents itself for our analytical toolkit to be opened to ask objective questions to strategic planners. We can use techniques such as benchmarking studies, market trend analysis, or industry white papers as sources of information to ignite the discussion to identify strategic goals and objectives, or if already identified, make changes to them if necessary.

As such, the identification of what data staff need access to, how staff use data now, where staff get reports, how staff interact with data, what staff use data for, what staff do with reports once reviewed and to whom staff communicate these reports can be a first line of questioning. We must also keep in mind that leaders want access to a fast and effective reporting system with results. In today's dynamic working environments, whether needing good insights to improve business processes, assess market competitiveness, find trends in customer demand or whatever the goal may be for an organization, the need for business to access an analytical environment on demand and supported by accurate data is the new business modus operandi.

The questions we need to ask of each category of stakeholder in order to understand how data management can achieve strategic goals and objectives is dependent on the following practical observations from experience:

1. The types of data to which each line of business needs access is dependent on the operational lines of business under the scope of analysis including financial, marketing, sales, staff performance (Human Resources), and customer complaints and inquiries, to name a few.
2. The current use of data yields insights into not only data needs but also how staff interact with data. For example, do they use slicing and dicing, pivot tables or "what if" scenarios? Or to they make decisions using static, descriptive reports where little to no analytical processing is available?
3. The level of depth of IT involvement in creating, generating and distributing advanced data reports is a factor is measuring not only the readiness of an organization to shift to an analytical environment but also yields insights into opportunities for self service solution development where users can run their own reports, putting less pressure on IT to build reports all of the time. The term self serve is used often in discussions about service oriented architecture (SOA). SOA was pointed out in Chapter 2, under validation approaches. SOA is a concept that puts more services, like data reporting into the hands of the user to satisfy on demand requests. In SOA, more IT services are exposed to end users to facilitate on demand data access.
4. The decisions that business executives make on a regular basis reveals a rich pool of potentially innovative methods to alleviate data shortcomings and highlight existing benefits, costs and risks to having bad data, untimely data, old data, or missing data cropping up in existing data services and reports.
5. The frequency with which reports are required can contribute to the planning of a self serve environment in the future. Multiple demands on IT to produce reports may start to show cracks in the process as competing priorities for IT time take over from routinely satisfying reporting demands by staff. In addition, lack of urgency identification to most help desk tickets created by users asking for reports in general often creates perceptions of little to no importance given to the needs for reports, thus the shelving of a ticket based on this notion can contribute to the lack of timely reports being generated by IT.

Practical example

The on demand nature in today's business environment suggests that organization's want to increase production in areas for which they are operationally responsible and accountable. According to the website maximizer.com[5], the number one reason to employ business intelligence is just for that reason: to boost productivity. Maximizer asserts that "through traditional data gathering methods, users need to compile and analyze data and write related reports. This can be incredibly time consuming, especially for small businesses that may not have the employees to do it. With a BI program, [staff] can pull data and create the reports at the click of a button thus freeing up time and resources allowing employees to be more productive on their own tasks."

Table 4 – Reasons why business intelligence employed in companies

3.2 KEY DATA PROCESSES

Information Technology Group

The processes for how data is collected, stored and accessed using information technology solutions may be documented or modelled within an organization's centralized file repository or quality management system. Theory is one thing. Reality is another. User's like convenience so asking questions about all possible sources of where processes are documented is a key skill to bring to the BI project.

Aside from the potential challenges of diverse sources of information, the benefits of discussing data processes with management of the IT staff, including managers, directors and Chief Information Officers (CIO) include the ability to uncover existing technical solutions used to satisfy current data needs. Discussions of this nature can help a business analyst:

1. Obtain data flow models;
2. Show where current reports are stored, how many reports are available and to whom the reports are delivered, either manually through mechanisms like email or dynamically through a report generating solution within an application or via some independent database that exposes data for reporting purposes to staff;
3. Obtain samples of the structure of reports as needed for the scope of the data analysis
4. Explain the process for receiving, prioritizing, creating or editing reports for staff;
5. Identify limitations with current technology as to delivering key data to staff in formats they desire; and
6. Identify IT business plans, goals and objectives that support corporate strategy

Exercise

High Level Assessment of Data Categories Per Business Process

It is time to think about the business. Explore the following organization and using brainstorming, put a checkmark under the key business process for which you believe, at a high level, data is required. Think of each need for each business process and determine if a data category could be identified as being a source of data that could be required, or potentially used by that business process.[6]

	Business Processes					
Data Category	Finance	Administration	Communications	Sales	Marketing	HR
Supplier Information						
Customer Information						
Product Revenue						
Customer Complaints						
Customer Demographics						
Market Trends						
Staff Leave						
Client Satisfaction Results						
Assembly Line Statistics						

Table 5 – Identifying data dimension targets by business process

One key outcome from this exercise is the potential to identify "single sources of truth", highlighted by how many checks are placed across business processes for a specific data category. Another outcome could yield which business processes use similar data that could identify potential sources of data duplication, and therefore contribute to data deduplication efforts where focus groups can assist the analysis with the goal to streamline the sharing of data across business domains. This is where the concept of master data management can inform the analysis by having all stakeholders participate in a discussion about "single sources of truth"; in other words, a data source that acts as the central repository for the distribution of data to a wider audience.

Once we obtain, document and assess the strategic drivers, motivational factors and high-level operational demands behind data needs, the next step is to begin a deeper analysis using these inputs, arming our analysis to next learn more about other key business processes, not just that of IT as described above.

Below are 10 questions that could be asked of a key member of the management team using the interviewing business analysis technique while drawing out requirements for a business intelligence project, at a high level. A rationale is provided as to why ask the question.

10 questions to ask of senior management about their data use	
Question	**Rationale**
What is the primary modus operandi of your branch or division?	Determine the scope of business. Knowing the type of business in which a leader is responsible for directing staff and operations sets a foundation for the rest of the interview.
What are your strategic objectives for the next year and in the following years?	Remember your audience. Members of an executive team have little time for gossip or rambling stories. You need to think objectively like they do. This question gets them talking about a subject for which they are keen on talking and a domain for which they have a high level of responsibility.
What is the nature of the data in the reports you currently review?	Dive right in. Get to the point. The types of reports add value to the primary modus operandi question. Don't worry if the reports come from staff, the Internet or some other location. The point is to talk about the reports, not the people preparing the reports. You will have a chance to ask questions of the preparers later in a workshop or interview.

10 questions to ask of senior management about their data use	
Question	**Rationale**
What level of granularity can you achieve with the reports?	Asking this question gets to the nitty gritty of how deep the executive can go into being able to drill down to specific levels of data to see information at a very fine level, like customer location, or customer purchases for a given day of the week. While understandably a somewhat ambiguous question, the point is to get to the root cause of some of the reporting limitations and boundaries an executive may have. If they can't get to the right data at the right level of detail at the time they desire, they risk making a decision that could go counter to a desired effect or perceived outcome.
Do you have direct access to the reports using a dashboard?	How quickly can the executive get to the information they need? If they depend on staff to send them reports via email using Excel as the attachment, then there could be a lag in time between sender and receiver, which is an undesirable event. Answers to this question yield insights into the method of communication delivery and access later for a BI project, as part of a solution approach.
How have decisions in the past using data been affected positively or negatively?	Never an easy question. You want to refrain from trapping them into a corner about "negative" outcomes from decisions. But it needs to be asked in order to assess the conditions – and missing data – that led to the alternative outcome.
What delivery methods do you need in order to access data when you need it?	Think mobile. Think iPad. Think Blackberry, iPhone, Samsung, or laptop. Think anything other than desktop PC tethered to a wall! The method of delivery will determine the technologies that make access to data simple, effective and of the very highest of value.
Is the business culture analytical?	Traditionally, reports come from information technology staff, or staff that can get data from sources like Excel, Access or some other staff accessible product. The point here is to assess the readiness of staff to think along analytical terms. Most staff are not expected to be analytical. In analytical environments, staff think about data in different ways and help determine data needs more easily as they think strategically, rather than operationally. It is fine to think operationally. Today, most organizations get their objectives directly from the CEO, down to the Vice President, on through to middle management, and funnelled to operational staff. If analytics is part of the culture, people are thinking about data, not about getting reports ready that just contain data.

10 questions to ask of senior management about their data use	
Question	**Rationale**
Is your branch or division a consumer of the data generated within its operational line of business or do you share data across operational lines?	If this data is captured, stored, generated and shared only within a specific branch, them you have identified the business scope of any BI project. If, on the other hand, the data is shared, them you need to assess how the other executives use the data as inputs into their decision making, and ask similar questions.
Do you need to augment data, like customer sales, with additional data from sources you don't have?	Helping the executive to think beyond their current operations may yield insights into data that they desire but do not possess at that moment in time. This helps identify any demand driven data needs; needs that are only fulfilled by creating new pools of data within an organization or capturing the data from outside sources, like open data portals that make data metrics available for public consumption, and whose information can augment facts to enrich the decision-making process.

Table 6 – Example so questions to ask executives about data usage and needs

Armed with high level and perhaps more detailed strategic information and requirements, the analysis can start on a road to inquiring deeper into the data needs of an organization. From there, the analysis can start looking at gaps and the process can begin to develop a solution approach that creates scoped requirements.

We continue now with a journey into the Business Intelligence Framework Core.

4 THE BI FRAMEWORK CORE

Armed with a list of stakeholders, strategic needs and a working knowledge of enterprise analysis, the business analyst can leverage a distinct set of stages that makes up the deeper analytical phases of the business intelligence framework. Inputs are assessed and a conceptual solution can start to take shape.

Figure 3 illustrates the core stages constituting the BI framework.

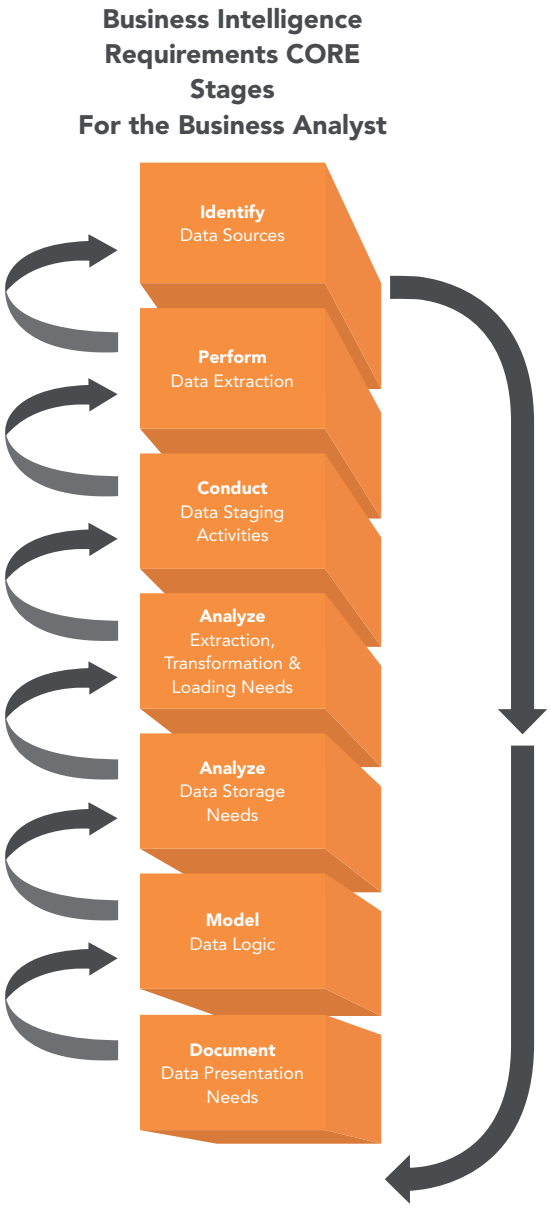

Figure 3 – The business intelligence framework

There are six keys stages to a BI project. The business analyst will be working within a specific stage as the project progresses along. A lot of work up to this point, including the stakeholder analysis, enterprise analysis and strategic review will yield insights into problems a business faces with respect to their data. Each of the stages are explained below.

4.1 DATA SOURCES

Sources of information and data abound within an organization, often stored in repositories unknown to IT. A business analysis can use the following tasks and activities to identify data sources:

- Identify and document operational data from sources;
- Identify and document market data the organization can leverage, such as demographics or surveys that can be pulled off of the web from open data sources to serve potential requirements for benchmarking or to augment operational data, identify trends and make comparisons; and
- Prepare a list of the different formats of the source data (text, relational databases, Excel).

4.2 DATA EXTRACTION FROM SOURCES

Analyzing data from a multitude of data sources can be inefficient. From a time saving perspective, importing data into an environment disconnected from production source systems and be useful. The following business analysis tasks and activities can assist with data extraction:

- Import data from sources;
- Identify data refresh rates; and
- Identify critical points of extraction.

4.3 DATA STAGING

Placing data in an environment where true analysis of data can be undertaking without affecting the production systems is key to continuing efforts in preparation for analyzing data. Data analysts call this stage data staging, and business analysis can include the following tasks and activities:

- Data imported stored on a staging server to prepare for extraction, transformation and loading into a data mart or warehouse database. Data at this stage is still in transactional, not analytical form;
- Extraction, Transformation and Loading (ETL);
- Transform data from staging into data mart environment. Data is in analytical form at this point;

- Format and standardize data;
- Identify the scope of the transformation including retention policies and how far back in time wants to extract data;
- Identify data refresh rates;
- Identify requirements on what users need to select from the data mart;
- Identify enterprise wide data; and
- Identify business specific data (perspectives).

4.4 DATA STORAGE

Good planning of any data warehouse, data mart or data project includes knowing where post-transactional data will be stored. The tasks and activities for a business analysis include:

- Identify the system that will host the data warehouse or data mart environment for analytical processing to occur; and
- Identify the system characteristics required to support a data warehousing (i.e., business intelligence) initiative. While a domain of IT specialists, there is no reason why business analysis cannot be conducted here as the cost of such systems will need to be included in any business case development, which is a business analysis task.

4.5 DATA LOGIC

A key technique in any data project for a business analyst to master is data modelling and data flows. Key tasks and activities include:

- Identify business rules;
- Apply logic to the queries used to select data; and
- Identify enterprise wide data

4.6 DATA PRESENTATION

User will be more likely to use a business intelligence solution if the data is presented as required from the perspective of users. Key tasks and activities include:

- Identify the tools to be used for analytical processing of data in reports;
- Identify the methods for sharing data;
- Identify frequency of distributing reports; and
- Assist users with how to use tools and understand the analytical part of the data presented to them. Users will only be able to select data from a data mart or warehouse. Transactional data using inserts, updates and deletes will continue to be processed in sourced data repositories.

As explained earlier, a key aspect of the business requirements elicitation and analysis life cycle involves looking at management processes (those that support the strategic arm of the business), operational processes such as Sales or Marketing and supporting processes such as Information Management/Information Technology (IM/IT), Administration and Finance. A process review will often yield further sources and needs of data that will add to the inputs into the design of the BI solution.

With strategy typically set by an upper management leadership team, processes for managing how effectively and efficiently an organization is moving along the path to meeting its goals and objectives generates a need for data by those at the management level (the managers and directors of operational units); data that is often fed to the senior management team. Armed with having the data needs at a higher level from senior management, an analyst can leverage this input to ask deeper questions of the leadership team closest to the staff members who carry out key operations. In turn, this input is added to validating and discovering data most often handled by operational units.

While the depth and breadth of the types of data requirements needed at the operational level is diverse, suffice it to write that staff on the lines work with data day in and day out. For example, Finance inputs expenses and revenues, a pharmacist's assistant dispenses medicine at a cost, marketers query data from open data sources, sales staff execute the marketing plans and acquire new customers or make repeat sales whose metrics get entered into a database. Logistics staff, the drivers of trucks and the handlers of packages all process data at some point during the day. The data inserts, updates, and deletes are added to transactional systems every second, minute, hour and day, making the need to extract that data in its current form the responsibility of the BI solution providers.

Like any gap analysis, the process starts with what is happening now. Often, the current form of support to users is provided by IT usually on a case by case, or on demand basis. The inefficiency with this is staleness of the data from the time the report is provided to a user to the time the user makes a request for updated data. The process of making a connection to a transactional system, loading data into a reporting medium (like a spreadsheet) and giving it to users who are on their own to work with the data contributes to an ongoing process of inefficiency and ineffectiveness.

Typical reporting environments require tens, if not hundreds of queries developed by an IT team, or technical expert to satisfy the diverse reporting needs of specific lines of business. It involves an intricate knowledge of table joins and knowledge of how to build advanced and complex structured query language (SQL) queries, which take time and effort to render these assets as usable outputs by business users. Using generic queries often doesn't satisfy the need to answer key questions related to strategic goals and objectives, or provide data in such a way as to answer those deep, insightful questions executives, managers and operational support staff ask of the data. Transactional systems are not designed for fast querying and analysis. They are designed to process selects, inserts, updates, and delete transactions on records.

While an analysis of current processes and procedures can lend a huge amount of information into what a new BI environment should provide to users in the future, reports based on transactional data merits a need for a solution that is more than just a neat representation of static data with the occasional chart popped in for visual effect; the analysis is also about exploring new ways of getting to data, processing it and taking action with little IT intervention after an initial development and deployment of the solution.

With a grip on strategic plans and situational analysis of internal processes, we can start to identify all sources of data, the first stage of our BI framework. Some of these sources are common to everyday uses such as database management systems, text based sources, metadata on files stored in a document repository. Even microfiche is a candidate as a data source. The sphere of data is not limited to the records in tables and other tuple based containers like databases but also metadata, the data that describes data. Alternatively, the term "metadata" can best be described as the information that sheds light on the data inside a data source. An example of this for a database (and it depends on the database) includes number of rows, create date, who created the database, its size, location, number of users, status and last backup to name a few of the metadata attributes potentially available to you as the analyst.

While we are investigating the sources of data and beginning the process of understanding the data, we can revisit answers to several questions of our stakeholders that we identified earlier. Validation questions might take the form as follows:

- What does the current data look like?
- Are there formatting and integrity issues?
- Is there duplication of data across sources?
- What data is popular and why?
- What data is least used or requested and why?
- How do users think about data?
- Have they thought of using data in a specific way?
- How do current staff roles drive data needs?

Answers to these questions can yields insights into whether data is missing from their current needs which should be developed to meet future demands. Lines of questioning can be facilitated through interviews to ask stakeholders the kinds of questions they ask themselves when they look at a specific data domain, like Sales. In addition, users often inform the analysis through the revelation of the level of granularity they need in their data reports such as whether they need reports down to the hour, minute or second. Sometimes, we want to dig into each stakeholder answer about their operations from the data they are reviewing to explore the subject of how decisions made in the past may have been affected not just by access to good data but also from the existence and use of bad data.

At this stage, some aspects of the data extraction, transformation and loading (ETL) process will come out during a review of data sources by asking some of the questions above. Answers may uncover issues with data formats, data types, standardized date and time usage, duplication of data, and how data will need to be transformed from online transactional processing (OLTP) source systems to online analytical processing (OLAP) systems can be documented.[7] These insights will help develop the logical data model that will help with the development of the information cube(s) that will eventually represent the physical implementation of the analytical system comprised of fact data and dimensional data, discussed shortly.

Another key area to discuss with stakeholders is how data is to be aggregated. While aggregation is achievable in static reports, the time and effort to manually create these aggregations through queries may warrant the exploration to a shift to an analytical system that provides these functions automatically. We cannot assume that users will be given an interface to their data using advanced spreadsheets that come with these functions. While spreadsheets are popular, they might not always to be an option of choice as reporting interfaces.

Keep in mind also that even if aggregation is provided, the static data reports may be stale after a certain period of time as they often come from transactional systems. OLAP environments live in powerful systems and are designed to do just one thing: provide analytical foundations for fast, deep and broad data processing in real time or close to real time as possible, with presentation and visualization support based on a foundation of well defined data structures, calculations and aggregations as requirements elicited by the analyst. In OLAP environments, users connect to the analysis system to get the latest updates on demand. This pull concept is more efficient than the traditional push concept where IT dumps data and provides it to users.

As our repository of data sources, needs and supporting requirements grows, we can next focus on how to conceptualize the data in an analytical environment to provide the logical foundation for building a physical BI solution, which introduces the topic of multidimensional cubes.

5 CUBES, FACTS AND DIMENSIONS

When we speak of multidimensional cubes, we are referring to the dynamic blocks of information that contain facts and supporting data. The review of cubes, fact table and dimension tables will evolve around their business meaning.

To help conceptualize our information cube(s), we can leverage a couple of tools. The concept of entity relationship modelling (ERM) is important. ERM can help us answer questions that will lead to the formation of the fact data and their relationship to dimensional data inside the cube framework. The purpose of this type of modelling method is to help us with logical concepts and Ralph Kimball, a proponent of data marts noted that dimensional modelling "is a logical design technique for structuring data so that it's intuitive to business users and delivers fast query performance."[8]

Business intelligence is where data is presented to users in information cubes, or just cubes for short. Cubes represent data in terms understandable by business users.

Another aspect of cube analysis is to identify what aggregations are required. Aggregation is the precalculation of data into summarized or group form. Examples of this include overall total sales, sales by region, sales by city, products sold by customer and similar types of numeric aggregations and data expansion points a user might need to either drill deeper into their analysis of the corporate data by expanding data to a more granular level of reporting or see things at a higher level by collapsing the data at points of aggregation.

Another important point is to find out how frequent the cubes are to be updated from data pulled from the data sources. This information will help the developers of the BI solution with resource planning and how frequent to extract data from the data sources that feed the staging area where transformation of the data occurs before the data is loaded into the BI solution.

This point in time is also an opportunity to determine the roles and permissions that the BI solution must support. In most BI projects, users will only need to have read access in order to query an information cube. Our stakeholder list can help identify data modellers, power users, business analysts, data stewards, developers, database administrators and any executive sponsors or BI project steering committee members which are job categories that can assist in our roles and permissions conceptualization.

Fact tables store the numeric details of our business domains under analysis in the cube. In the BI world, we call these numeric data "measures". Most fact tables are just that, facts. They contain only numeric data. Our job is to ask the right questions of each line of business about what facts they need to have captured in the BI solution. Facts are often validated using our analysis of operational business processes and data sources as part of our situational ("as is", "to be") analysis.

To note further, fact tables also have keys that not only uniquely identify a record in the fact table, called a surrogate key, but also point to data elements found in other tables making up the BI solution, called foreign keys. This establishes the physical fact to dimension relationships we conceptualize during an ERM exercise.

Dimension tables give context to the fact tables. We must further elicit requirements from the respective business units about the methods in which facts will be sliced, diced and aggregated. Slicing and dicing is often used when talking about functions found in spreadsheets.

Slicing and dicing is mechanism for viewing data in a multidimensional cube in specific ways. These methods are often misunderstood. They also have different meanings. Slicing is a technique where a user needs to analyze measures (the facts) down to a specific level of detail within a specific dimension. An example of this is a user wanting to see how many sales (facts) were made by a regional office (dimension). Dicing is a technique where a user wants to view information in a sliced cube but at level of detail that spans dimensions. An example of dicing is a user wanting to see how many sales were made by a regional office via a specific sales channel that came from a specific country during specific months. We slice the sales data cube (by region) and break the cube down further by dicing it into smaller cubes of detail (channel, country, and time).

The example above is a way to position a discussion with stakeholders about future needs by which they want to view their data and at what level of details. I often start a conversation with "So you need to report on your data by…" that feeds the dimension discussion and modelling activity. The "by" term helps identify data dimensions.

Example
Analyst: What types of data is needed by which you need to make decisions?
Stakeholder: I need to assess our sales by sales person, by region and by product category.

Table 7 – Example question to determine dimensional data needs

In the above example, the "by sales person", "by region" and "by product category" would represent the data dimensions that augment and support the sales facts. So the requirements in this example include sales numbers, people, geographic regions and categorized products.

Also remember that as part of our data dimension analysis, we are not limited to using data and defining data attributes that only exist from within the organization. We should also ask questions to the business regarding appetite for augmenting factual data using open data from other sources such as Google, the United Nations or the Government of Canada's Open Data portal and other leading organizations that make available data for research and other purposes to the public. The requirements elicitation process might yield a need for obtaining census data, information on demographics, or survey data that is either not available within an organization at the time but is available from other sources outside the business that specialize in collecting, storing, processing and providing open data. The data that could be integrated from outside of a BI solution will depend on what the goals and objectives are of the business based on a strategic review.

Knowing about which dimensions to make available for which business processes can also be identified through a simple charting exercise. What this chart consists of is a listing of all of the business areas under consideration, such as Sales, and a list of all of the potential dimensional data that could apply to this business area (time, date, country, etc.). Below is an example of an output from such as exercise. The processes will yield areas where facts can be gathered. The dimensions that could apply to these processes will yield what context data we would need to support the facts.

Dimensions (Context)/ Process (Facts)	Product	Employee	Country	Customer
Orders	✓	✓	✓	✓
Complaints	✓		✓	✓
Marketing			✓	✓

Table 8 – Mapping dimensions (business context) to processes (fact sources)

Once we have a good understanding of the data that will need to be loaded into our BI solution, we have some fun with stakeholders about how the data is to be transformed before it is served up to the end users. Transformation is the technical process applied to data where data is cleaned up and formatted to a consistent standard before it gets put into the BI environment. Often the data extracted from transactional systems is populated into a staging area where transformation scripts work on the data which is then ready for loading into the analytical system. What those transformation standards are becomes our responsibility during analysis. We need to identify any corporate policies or procedures that can help us with our transformation requirements elicitation, and if there aren't any, we need to elicit, document, validate and get approval on those transformation requirements. Transformation requirements are not to be confused with transitional requirements, explained later. Transformation applies to data; transition applies to the shift from one cultural data aspect to another, from transactional thinking to analytical thinking.

As previously noted, transactional systems often have bad data in them. It is our job to assist stakeholders identify this bad data and making it better, cleaner and structured to a consistent format that will make the end users value the BI rollout. Bad data, or lack of sufficient data altogether that could easily be captured and processed, can lead to bad decisions and could result, for example in losing market share or shifting to a different strategic tactic whose need may have been premature and consumed valuable resources unnecessarily. Good data may have informed the decision otherwise. Bad data extracted from a transactional system and loaded into an analytical system in a format that is equally bad is a pointless exercise and a waste of valuable time and effort for all resources involved in a BI project.

Our work up to now will consist of many requirements, often at a volume where scope management is required in order to deliver a BI solution. Scope typically starts out with a large array of needs and wants and without some project governance or controls, becomes unwieldy. Starting out working with a small subset of the scope, such as one business process like Sales, works best. Future data needs can be factored in later using a change management scheme. The key here is to identify the data requirements that provide the highest beneficial returns to the organization. Start with these requirements to initiate the process where you dive deeper into ensuring they are clear and concise before working on the other requirements. These can be used to make up the business requirements value/feasibility matrix, discussed later in this chapter.

5.1 DATE AND TIME DIMENSIONS

Most analytical environments use dates for registering when specific transactions have occurred in an OLTP system, such as when a sale has been made of a specific product from a certain category, when a customer has registered a complaint with a company help desk line, or when a user signs up for a specific event held in a specific city.

Date and time dimensions enhance the value proposition of providing numeric details associated with a particular transaction. In an OLTP environment, dates and times are stored as signifiers of a particular event in time. In an OLAP environment, where users query data extracted, transformed and loaded into an analytical system, data and time dimensions augment the facts; that is they add additional informational benefits to a specific event in terms of the range of capabilities associated with slicing and dicing facts based on those dates and times. Used with other dimensional data, dates and times can reveal powerful insights to a keen analytical eye.

Below is a typical date and time used in a Microsoft SQL Server database environment: 2017/03/22:12:15:30. We can all safely assume that means March 22, 2017 at fifteen minutes past noon, or more specifically fifteen and a half minutes past noon. From an online transaction processing perspective, this is a very useful piece of information and has a fair bit of detail. From a standard, static reporting vantage point, this information may suffice for users who need to know just the date and time when transactions occurred. While the date component may also be universally recognized, meaning that someone in Canada and France understand it as March 22, the subtlety lies in the fact that a user in France may want the textual name of the month and day written in their language: French.

With a little extra work setting up the date and time dimension table in an analytical environment, user experience and preference requirements (i.e., perspectives) can be achieved with some high value propositions built into the data presented in a more advanced, dynamic way.

Knowing the user requirements for an advanced, or business intelligence project dependent on dates and times is key to understanding what users need in terms of not only what data is presented but also how the date is presented to them.

Language requirements are critical for companies that have offices spanning the globe or have a reach beyond a single country; critical in the sense that users in any country should be able to work; or in this case view data in the language of their choice. In this case the realm of language requirements goes beyond just the interface in which the data is presented to users, whether that interface is Excel, Power BI, or SQL Server Reporting Services. The realm broadens to providing data in specific languages too.

Dimensions for dates and times can be as wide and deep as needed; wide in terms of the number of attributes or columns need to store dimensional details about dates and deep in terms of how far back and ahead the date dimensions should go. While the task of creating date and time dimensional data in the future is easy, the business analysis activity should determine via data reviews how far back to create older dimensional data. For example, if a company is analyzing product sales only and their first sale ever was 1988, then you most likely only need to provide date dimensional details starting as far back as January 1, 1988. Recalling our single source of truth ideal, a single date and time dimensional table in an analytical environment can be the source for a multitude of fact tables within a single BI realm. Start where needed, but plan for updating the dimension as more and more data is targeted for loading into the analytical environment.

5.2 CONCEPTUALIZING TIME

Modelling the requirements that drive a date and time dimensional table are keys to not only understanding business parameters but properly documenting the data needs. A date model may look like this if hashed out using a whiteboard or doodle pad:

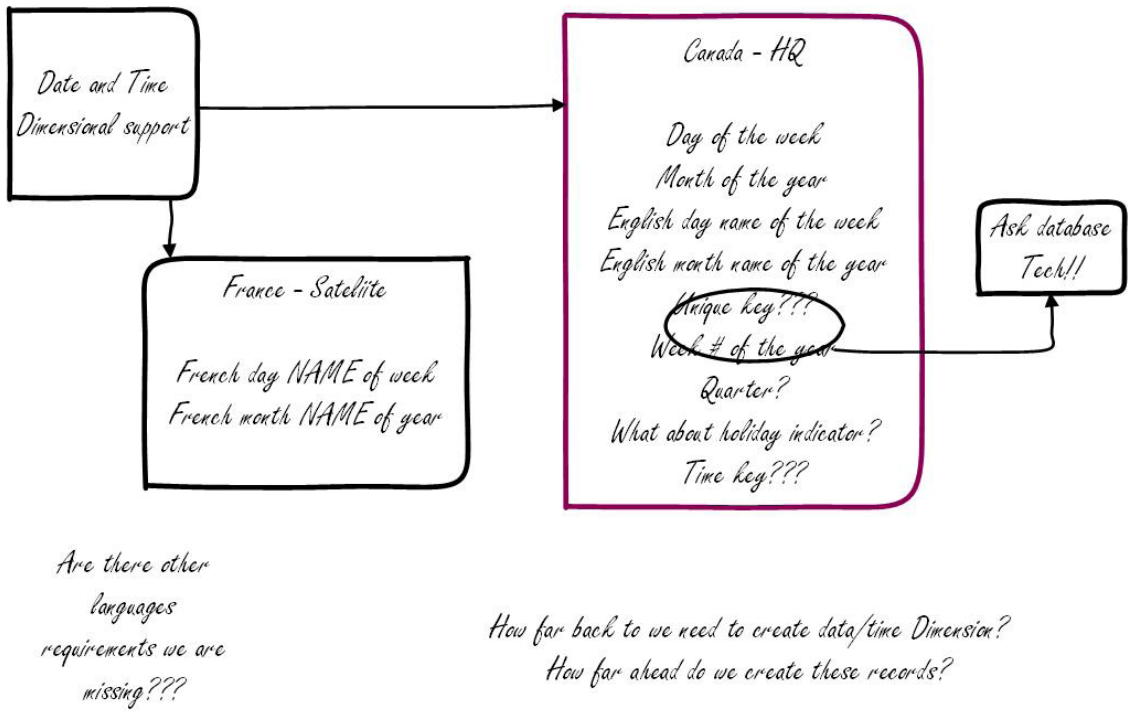

Figure 4 – Whiteboard rendition of conceptual date dimension

5.3 MODELLING THE EXAMPLE USING THE DATE DIMENSION

Next, we might spend a few hours modelling the requirements using Unified Modelling Language (UML) data modelling, a class based modelling language. We assume that we have confirmed and documented as stated requirements:

- That no other language requirements other than English and French is needed;
- The database technician has stated that a unique key is required to support distinct records in the date dimension table;
- We need to create date attributes as far back as 1988; and
- For now, creating date dimensions into 2020 will suffice. Note: the database technician will know how to do this. For now, the requirement is to be able to create date dimensions for the next 3 years or so which can be achieved automatically using functions in any database management system.

Below is a Visio rendition of an initial draft of the UML date dimension model.

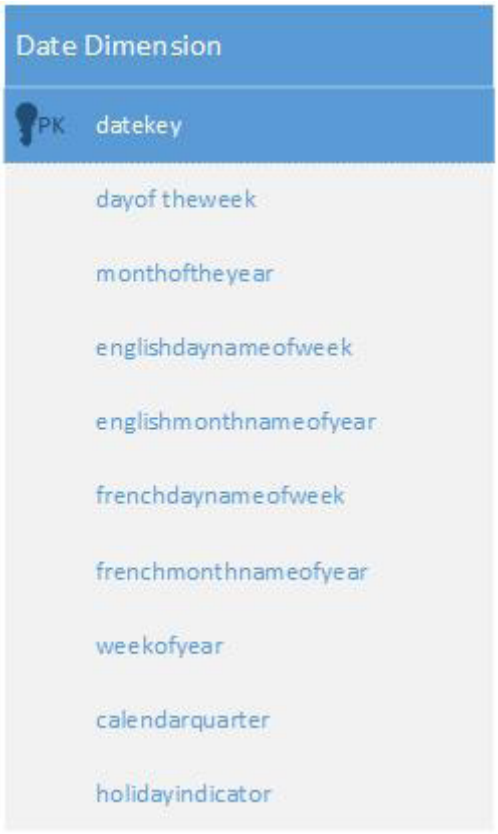

Figure 5 – Translating whiteboard requirements into UML data model

A few things to notice in the UML diagram. Visio automatically made the "datekey" attribute a primary key. The primary key is the unique identifier in the date dimension table. Also, time is invested in determining the data types and their lengths in the model. In this case, the focus is on the requirement, leaving the types and length to the database technician to implement. There is no need here to consume valuable business analysis resources on this aspect of the model.

The physical date dimension table with the cross section of it shown below from a sample of year 2000 records:

Results | Messages

	datekey	dayoftheweek	monthoftheyear	englishdaynameofweek	englishmonthnameofyear	frenchdaynameofweek	frenchmonthnameofyear	weekofyear	calendarquarter	holidayindicator
1	20000101	7	1	Saturday	January	samedi	janvier	1	4	NULL
2	20000102	1	1	Sunday	January	dimanche	janvier	2	4	NULL
3	20000103	2	1	Monday	January	lundi	janvier	2	4	NULL
4	20000104	3	1	Tuesday	January	mardi	janvier	2	4	NULL
5	20000105	4	1	Wednesday	January	mercredi	janvier	2	4	NULL
6	20000106	5	1	Thursday	January	jeudi	janvier	2	4	NULL
7	20000107	6	1	Friday	January	vendredi	janvier	2	4	NULL

Figure 6 – Physical data model from conceptual logical requirement

5.4 THE TIME DIMENSION

Up to this point we have spent time on the date dimension. What about the time dimension? Having a time dimension depends on the business. Some people may not care about time dimensions while others may be fanatical about them. Ask the key stakeholders about this need. For now we will assume that time dimension is required, as our stakeholders are analytical enthusiasts! Next are the questions we should be asking of the key stakeholders, including the technical team:

- Why is a time dimension required in the first place? In other words, what value does a time dimension bring to the analytical environment?
- Do we use the 12 or 24-hour system?
- How do we conceptualize and model a time-based dimension?
- What is the impact on the date dimension in terms of the number of records needed to be created within the date dimension table to support records that have different timestamps but the same date?
- Is there some kind of link that we can establish to keep the data and time records separate?

5.5 THE VALUE OF A TIME DIMENSION

When thinking about using time dimensions, consider the business drivers for it, as any enterprise or strategic review would yield or may be discovered during an analysis of data needs at an operational level. Ask executive, management and operational users how they see time dimensions being used as part of their analytical environment. Even if the discussion leans towards not using, consider ways to introduce the benefits of using a time dimension. Consider the case study below.

Example

The marketing team has put together what they consider a game changing program in support of a new product being shipped by their company, ABC. The advertising team spends a large pool of money in support of the marketing program. The CIO wants to know not only when the products gets purchased but within how many seconds, minutes or hours customers buy the product from the moment the CEO makes the "go live" decision with the program. The CIO wants this information to assess the length of time between marketing a product and consumers buying it, at a deeper level than simply the date on which their customers make the buy.

Table 9 – Analyzing real world usage of time dimensions

5.6 WORKING WITH TIME IN BOTH FORMATS AND CONCEPTS

Going back to our example, we might have source dates extracted from our transactional system that look like this, assuming a 12 hour format:

- 2017/03/22:12:15:30 PM
- 2017/03/22:12:18:40 PM
- 2017/03/22:18:16:10 PM

In our date dimension table, the date would be represented as 20170322 and is used at the date key. Here we have 3-time stamps, which mean 3 different records in the time dimension. These time records could be represented in our analytical system like this:

- 121530,
- 121840
- 181610

If we put on our critical thinking caps, we know the time dimension falls between 00:00:01 and 23:59:59 so the number of records is bounded by this range. So we conceptualize and model our time dimension the same way as we do for the date dimension, only we have a narrower pool of records with which we need to work.

Working with a 12-hour calendar shortens the number of records thus calling for a parameter to identify AM or PM. While the 24-hour table would contain almost 236,000 records, its volume might deter database administrators from using up vital disk space. The business requirements elicitation process will yield guidance on what to do, revealing either stated requirements or a policy decision from the CIO or key database administrator.

Our time dimension fact table might look like this:

Unique ID	Time
121530	121530
121480	121480
181610	181610

Table 10 – Example of time dimension table

In this example, it's good practice to use the same number for the unique identifier and time attribute.

The extraction process from the transactional system would split out the date from the time part. In terms of the date dimension table there is little impact from our earlier modelling. But we need to be able to associate a date with a time in the analytical environment. Part of the extraction, transformation and loading is assumed to have created this association. In the analytical environment, we link two-dimension tables together using a factless fact table. In this case the factless table, as built by the database administrator or BI professional will consist of the date unique key and time unique key. Let's use an example using the same timestamp but a different date for our case study.

Factless ID	Date Unique Key	Time Unique Key
1	20170322	121530
2	20170322	121840
3	20170322	181610
4	20170323	121530
5	20170323	121920

Table 11 – Associating dates with time using relationship table

I added a few extra records to show how the date unique key can be associated with multiple time unique keys. And as you can see, the time key is the same for a purchase made on March 22nd and 23rd. What are the odds of this happening? Rare but the situation could arise if the site has a lot of user traffic and ecommerce purchasing volume. Think of Fedex or one of the other logistics companies. Online transactions happen almost every second. For now, we are illustrating the concept.

The impact on the date dimension is nil. The same number of records remains the same. The factless fact table facilitates the time to date relationship linkage.

Date and time dimensions play an important role in the business intelligence project. The data is structured in such a way that it can provide leaders with the insights into what is happening with their programs and operations that allows them to make better decisions with the data they have at hand.

5.7 THE VALUE TO FEASIBILITY MATRIX

The value to feasibility matrix can help prioritize requirements. A prioritized list of requirements is subject to an agreement among the BI project team, the operational lines of business, the senior management team and the project's executive sponsor.

The matrix is easy to setup. The hard part is getting agreement on what priority level to assign a BI requirement. To get there involves objectivity, diplomacy and good judgement from all users having a stake in the project.

The process goes like this: for each requirement written, ask each business unit their perceived value of that requirement to their operations. Not all requirements will be High priority nor can they be. In the real world, all requirements having a high priority makes it difficult to execute a solution development plan according to schedule.[9] Through business analysis techniques like negotiation, discussion and rationalization, requirements can be put on a list and subjected to a litany of questions about their need to implement requirements at certain times.

What helps with that implementation schedule is asking the IT staff who will be responsible for building the BI solution how feasible it would be to implement that requirement. Document the findings. Not every requirement can be implemented with ease. Once all of the requirements have been reviewed and a feasibility level has been applied to each of those requirements, you can create and populate the matrix that identifies the requirements that fall somewhere within the value to feasibility spectrum.

The suggested industry approach is to implement those BI requirements that are of high value to the business and very feasible to implement by the IT solution providers. These, if all parties agree, could be the scope of the initial BI project. Those requirements that the business considers of lower value but IT believes is not feasible to implement, meaning it is difficult to implement and requires extra resources to get it done, might be put aside, ignored or dropped. Requirements that fall in between can be either put aside for implementation once the initial requirements are met or can be documented in a log for further discussion because of the uncertainty as to their technical feasibility or business importance.

Table 12 on the next page documents an example.

Example	
Requirement 1	The business intelligence tool must auto calculate the difference between two dates.
Requirement 2	The business intelligence tool must also allow a user to change the chart type associated with viewing specific data.
Requirement 3	The business intelligence tool must provide a "what if" scenario tool.
Requirement 4	The business intelligence tool must allow the user to issue a voice command to view different data presented on the screen.

Table 12 – Example requirements to help with prioritization of requirements

Using 4 sample BI requirements above, let us assume the following after the business analyst has met with both the business user and the IT solution provider:

- Requirement 1 is highly feasible with a high business priority
- Requirement 2 is highly feasible with a low business priority
- Requirement 3 has a low feasibility with a high business priority
- Requirement 4 has a low feasibility with a low business priority

A highly feasible requirement is one the solution provider can implement with relative ease; relative to a requirement that is more complex or requires more resources in order to make a solution available (the low feasibility aspect).

Table 13 below shows the value/feasibility of the sample requirements.

Feasibility / Value	Low	High
High	Requirement 3: The business intelligence tool must provide a "what if" scenario tool.	Requirement 1: The business intelligence tool must auto calculate the difference between two dates.
Low	Requirement 4: The business intelligence tool must allow the user to issue a voice command to view different data presented on the screen.	Requirement 2: The business intelligence tool must also allow a user to change the chart type associated with viewing specific data.

Table 13 – Assigning requirements to quadrants based on feasibility and value

In this example, Requirement 1 would be implemented first. Requirement 2 would be implemented if resources permit. Requirement 3 requires more discussion between the solution provider, the business analyst and the business user as there may be some points of ambiguity or confusion prohibiting the solution being developed to meet the requirement. Requirement 4, given it lowly status would have a solution deferred or postponed.

6 PRESENTATION REQUIREMENTS

At this stage of the analysis, the requirements for how data is to be presented to ends users starts to take shape. There are many types of ways that data can be presented to end users. In some cases, it depends on the type of stakeholder that is looking at the data. In other cases, it may not matter.

Below is a table that shows the stakeholder category and their potential perspective with a rationale as to why this might be the case. Naturally, the analyst will validate these assumptions using techniques like interviews, workshops, focus groups or information sessions where feedback is encouraged from users.

Stakeholder Category	Presentation Requirement (Assumed)	Rationale
Executive	Dashboard Automated Decision Data	Executives need quick access to data, summarized at a level where they can glean important trends or issues without having to drill down into too much detail. Remember our stakeholder approach: time is important to these employees. The way the data is presented should mimic this factor.
Middle Management	Conditional Alerts Interactive Dashboards	Middle managers are responsible for their business processes. From time to time, being notified of specific conditions that warrants immediate action may drive the requirement for conditional reporting.
Operational Staff	Ad-hoc queries Interactive Dashboards Cyclical Reports	Often, staff need to generate quick ad-hoc reports like number of customers registered since last quarter. They may also need to check on how things have progressed since a specific point in time in the past to help with determining trends; trends that can be fed up the chain to middle management on to the executive staff member.

Table 14 – Methods to determine data presentation requirements

The nature in which data is presented will also depend on how the stakeholder wishes to access the information from a preferred device.

Executives are typically on the move, in meetings, travelling to other offices within regions or across the globe. Mobile devices are increasingly being relied upon to provide functionality to senior management specifically. The mobile nature of an executive's work will drive the types of presentation needs. An Excel file will not render well on a mobile device as a standalone file but it might if tailored for a responsive dashboard.

The same principles go for middle management and staff. Middle managers may travel, or they might not. Staff usually spend a majority of their day in the office but may, in some business circumstances require mobility and therefore a responsive system in order to view data from a device serving their mobility needs. Think of a massive warehouse with staff running around, like an aviation assembly centre. Imagine the need for responsive data from the right device at the right time in such a business environment.

Presentation of data and the requirements that will fill the need for presenting the data well can be the difference between a solution being accepted with small adjustments made in future iterations of the application and less acceptance with bigger changes needed later, requiring more resources in terms of time, money and people.

7 THE IMPORTANCE OF DATA PROFILING

One key element in the process of ensuring that data being analyzed is of good quality, is formatted correctly and whose integrity is of consistently high value is the performance of data profiling.

Data profiling has often been the domain of database administrators and data scientists. And these continue to be vital stakeholders in the need for addressing data that, through experience has been of bad quality, inconsistently stored, and stored using the wrong types of data for which the data's intent would suggest otherwise.

The business analyst can contribute a great deal to the realm of data profiling which is an exercise is obtaining statistics about the data. In this case, we are looking beyond the realm of metadata, discussed shortly and peering into the business meaning of metrics about the types of data, their lengths, maximum and minimum values, uniqueness, the frequency of null values and empty data.

The crucial point of data profiling as an analyst is to not only work with the business community to understand their data using common business language but also to help the database administrator, who will create the physical database model, understand the constraints, limitations and boundaries of data that will need to be stored in a data warehouse. The data warehouse will be used to support the analytical environment upon which users will come to depend in order to make better decisions using data at their disposal and to help them understand and interpret the results.

Data profiling can be used in parallel with any and all logical data models that an analyst would create to conceptualize the way data will be represented physically in actual database management system, like Microsoft SQL Server, MySQL, MongoDB, Frontbase, Postgres, Oracle, and Sybase just to name a few of the database solutions available on the market today.[10]

7.1 METADATA COMPARED TO DATA PROFILING

As an analyst, the term "metadata" might have been spread around data management discussions, and means "data about data". Using this perspective, the data being referred to is different from data profiling, discussed shortly.

Metadata can be narrowed to the statistics that represents the database management system in which data is stored, queried, updated, deleted and otherwise extracted in order to perform data analysis.

The statistics that is covered by the metadata realm include the following (and we will use the simple example of a sample database stored within a database management system that could literally host hundreds of databases):

- The size of the database in Kilobytes, Megabytes, Gigabytes or in some business situations Terabytes;
- How many tables make up a particular database;
- When the database was created;
- The person who created the database. This person may or may not be someone from the Information Technology Department;
- The date the database was first backed up;
- The date the database was last backed up;
- The size of the tables that make up the database;
- The number of views stored inside the database. Views represent a window into data stored within tables. Views typically show data from queries, often complex in nature to show specific data to users based on specific reporting needs, as the analysis earlier would have discovered;
- Default collation;[11]
- Default character set which is important for internationalization;[12]
- For specific tables:
    - How many rows; and
    - Average row length (a bit of data profiling but not quite the full array of profiling options); and
- Name and type of any stored procedures used in the database. A stored procedure is a piece of database code that can be reused to perform a function over and over, which may include data formatting, for example.

The purpose of using metadata statistics under data profiling activities is to plan to new databases, including databases for business intelligence that will be used to provide an advanced analytical environment, enriched with high value data attributes, facts and dimensions to augment the decision-making process and cultivate new ways to think about, and interact with data.

Some of the goals of metadata activities may include:

1. Use taxonomy (the use of specific terms related to the nature of the content) and controlled vocabulary for content enrichment;
2. Use social media features for personal classification and improved findings;
3. Use taxonomy and metadata ("data about the data") to improve application navigation and user browsing experience;
4. Use taxonomy and metadata to improve search and discovery of information;
5. Share content types (specific kinds of data attributes) across collections of documents.
6. Use existing data retention rules to help analyse and assess the lifecycle management of information;
7. Administer taxonomy term management; and
8. Import existing taxonomy using a metadata import file to speed up the availability of terms used to categorize specific data and related content.

The activities carried out to plan for the eventual migration of data into an ETL environment include:

1. Creation of database views for assets such as people, organizations, activities, financials, visits, groups, locations, and other generic tables to provide insights into data supporting business applications;
2. Creation of information schemas for tables which provide column names, default values, data types and sizes;
3. Creation of a master list of applications to identify business use, primary branch users, supporting application services, and databases supporting the applications;
4. Creation of a master list of lookup tables supporting applications;
5. Identification of candidates for database archiving (old databases not in use or idle);
6. Creation of master data profiles for tables in use by existing legacy applications (these help identify, for each data type, precision and scale, minimum and maximum numeric and date values, distinct values, null and blank counts (with percentage of total records), minimum, maximum and average string lengths;
7. The capturing of user interface screen shots to assist with discussions with business units about how they use data and to identify opportunities for improvement;
8. The conducting of cardinality, pattern, and frequency distribution analysis on specific tables.
9. Creation of a consolidated view (pivot table) of all tables supporting the varios data categories (e.g., people, organizations, activities, etc) to get a sense of frequency distribution and use of different columns;
10. Development of a migration paper to assist technical experts with understanding how databases store and make data available to its application and presentation layer; and
11. The conducting of an outward looking analysis of data used to support outbound communications like phone numbers, emails address; the goal is to identify data quality issues using a single interface for the analysis prior to any transformation being carried out on the data.

7.2 DATA PROFILING

Below is a practical example, taken from the author's experience with data profiling on the logical steps an analyst could take to assess data before data is migrated to an analytical environment.

1. Using architecture and concept plans, determine the data sources in scope.
   a. Determine retention needs based on corporate retention policy within a quality management system
   b. Apply retention rules on data sources as selected for data profiling
2. Categorize the data sources into groups, like people, organizations, products and so forth.
3. Import data into a separate database environment from unstructured environments, like flat file data repositories.
4. Set up individual data profiling projects using database profiling tools for categorized groups like:
   a. Customers,
      i. Including other terms used to describe a customer relationship management system;
   b. Vendors;
   c. Suppliers;
   d. Discussions/Collaborations;
      i. Including all document sharing forums;
   e. Inventory;
   f. Financial,
      i. Including procurement systems, or similar terms used to describe the electronic capture of purchase requisitions or expense claims tracking;
   g. QMS, or quality management system;
   h. Users;
   i. Groups;
   j. Helpdesks,
      i. Includes Travel, General Enquiries or other general request forums;
   k. Products/Services;
5. Estimate time effort for each discovery phase, including:
   a. Structure;
   b. Data; and
   c. Relationships
6. Run data profile projects;

7. Perform analysis and check for:
   a. Duplication;
   b. Date formatting;
   c. Data ranges;
   d. Domains;
   e. Uniqueness;
   f. Standardization of data, i.e., ISO country codes;
   g. Address formatting;
   h. Missing values;
   i. Null values; and
   j. Business rule violation
8. Identify business units in scope for data cleansing
9. Identify resources to help with data cleansing
   a. Identify skill sets needed to perform cleansing activities by business units
10. Perform data cleansing on legacy systems

Example

Many modernization projects involve the transition of legacy applications and data to a new system. One example is HR system modernization, including access to key reports in a timely manner. During the process, the opportunity for users to disclose the processes that work well, which processes could use improvement and the limitations and boundaries of how data is collected, used and disseminated can yield huge insights into supplementing the design of a new system. One particular aspect of modernizing an HR system is the analysis of requirements related to staff training requests. Below are some of the issues that may be identified during an requirements workshop:

- Employees fill out an online form to submit training requests:
  - The form includes the type of training request and approximate cost.
- The submitted form is routed to an employee manager for approval. The manager submits a response that changes the status of the requests using a business rules engine within the software application.
- Human Resources is notified using email.
- An HR staff member processes the request.
- HR informs the analysis team that training requests are not linked to staff performance objectives. A manual process is in place to verify training defined for an employee and matched against training taken at the end of the year. Training information is managed using an Excel file.
- Information is extracted from the financial management system to verify actual training costs against estimated costs in order for branches to manage their training budgets. A report is manually created by inputting the numbers once again in Excel.
- Training courses are added to the employee profile along with the dates of training.
- Employees download their training certificates from their training provider and manually upload them into their employee profile.
- Reports on training are contained in a budget book. Creating this book is time consuming and error tracking is labour intensive due to the amount of manual input.
- Travel associated with training taken also must be tracked against budget actuals, as sometimes reports on travel expenses are incomplete.
- IT has informed the team that the building of key reports for HR using data from both the HR application and financial system is a low priority.

Drivers of change
The main drivers are: • The length of time it takes to create the queries in the legacy application and to create the interface to those queries that presents the information on the screen. Time consuming, lack of timely delivery of information to satisfy on demand reporting needs, the heavy customization required to set up and allow for reports to be generated are some of the factors that contributes to other options for data management being used, like Excel. • The lack of integration between a true HR system and the organization's financial management system for data input and subsequent reporting outputs. • Dependency and reliance on manual processes increases risk of bad data entering the system, which contributes to errors and makes the system prone to other issues such as formatting and inconsistent use of data. Data duplication is inevitable with little or no spot checks on data entry validation. • Data profiling of training requests, for example can yield benefits about what data looks like, how often data may be replicated where replication is not needed, how well, or not so well the data is formatted such as whether specific data fields should allow a mix of alphanumeric characters where they appear.

Table 15 – Example of importance data profiling can lend to a data project

Data profiling is an activity that takes data and studies issues with the data like formatting, inconsistent use of data perceived to be standard, duplication, different versions of the same text and so forth. Profiling also pins down the true data needs because it helps uncover data that has never been used or used sparingly and contributes to discussions with stakeholders about whether the data is needed. It could also identify business processes dependent on specific data and can also help validate assumptions about the true intent of the data for a specific process. Finally, profiling helps determine a consistent, standardized and cleaner use of data stored in the analytical environment.

8 THE TRANSITIONAL FACTOR

Managing a shift from an environment accustomed to working with static, transaction based reports to an analytical one is a challenging task. Moving from an OLTP mindset to and OLAP one is achievable with patience, time, commitment, resources and "buy in" from the executive sponsor. To assist with identifying and gathering requirements from a transitional perspective, an organizational readiness assessment can gauge an organization's preparedness and appetite to change from a static, generic reporting culture to one that supports a dynamic business intelligence framework based on analytical foundations.

Readiness assessments can often yield, in addition to the details about an organization's analytical culture, transitional requirements that will help an analyst understand what training, learning and further knowledge needs are required to get an organization from its current state to its future state.

Not all organizations are analytical in nature in terms of data and nor do they all need to be. In other words, not everyone within an organization immediately thinks strategically and analytically and they all don't necessarily need to do so. It can be a significant cultural shift to get an organization from thinking in terms of the traditional reports they get to "analyze" to being analytical by nature or just thinking analytically to help you draw out requirements for a BI project. The process of eliciting, documenting, validating and modelling transformational and transitional requirements can assist immensely with what is needed to have a successful shift from the traditional method of reporting to an environment adorned with advanced analytical methods to enrich data reporting and improve decision making within the organization.

Organizational readiness is meant to gauge how prepared the organization is culturally and technically to move to an analytical environment. There are some techniques that a business analyst can use in order to document the requirements to ensure that a transition, while not always smooth, can be executed according to plan using specific tactics.

Consider below the following readiness categories, the potential risks and the mitigating factors an analyst could assess. Analysts may need to put on their Maslowian hat for this exercise.[13] Maslow's hierarchy of needs is considered a key measure to assess aspirational aspects of human life but can also be leveraged for assessing the human aspect of any organizational transition.[14]

8.1 PEST FACTORS

Influencers of an organization's readiness include political, economic, societal and technological factors. A rich mixture of these factors can lead to a smooth transition or a much harder transition. In addition to assessing the readiness categories listed above, these PEST factors should also be considered in general as the analyst assesses the organization's cultural environment.

8.1.1 BEHAVIOURAL

People hold specific beliefs and values close to them, even at work. Beliefs are much easier to change than are values. Behaviours at work have a direct and indirect relationship with the culture within an organization. Business analysis should consider how staff react to changes introduced to their environment by observing how they react, what they say or what they do not say. Body language reveals much by saying little or nothing.

No changes in behaviour may impede an organization from moving to a new level where an organization needs to get to in order to achieve strategic objectives.

In meetings with stakeholders, assess the way staff interacts with each other. How much respect, integrity and professionalism comes into play? Determine if tensions exist between staff by observing tone, posture, silence or need for monopolizing a conversation. In confidence, assess what staff are telling you about the nature of their work. Do they enjoy their work? Do they have a sense of ownership of the work? Is there work satisfying to them? What are their complaints? What are their successes? Does the organization celebrate success? Are mistakes used as inputs for punishment or as personal and professional growth drivers through the application of lessons learned?

Politics is everywhere. As in government, politics exists also in the corporation. Politics comes to the fore often when people want to protect specific work domains or assets and resist given them up. This is true for projects involving data sharing and modernization of data reporting. Service oriented architecture creates more possibilities where users can self-serve their data needs. An SOA architecture might mean that specific lines of business give up their "ownership" of data and yield it to a centralized model where staff across business lines can have access to it where once only specific business units had it in the past. Politics can interfere with efforts to meet these types of transition goals.

8.1.2 CULTURAL

As mentioned, behaviours can be a direct result of the culture inculcated within an organization. If management treats customers as second-class clients, staff may do so and this transfers to the behaviours staff exhibit within and without the organization. Often, introducing change is difficult if business has been done a certain way, with some success, for many years, or if management has been in place for a long time without any challenge to the way business is conducted by senior management and executives.

Rigid cultures may contribute to longer transition times, even for data projects. Staff who are trained to perform in certain ways may have trouble understanding a shift or may not fully comprehend why there needs to be a shift in the way they carry out their work. In addition, if executives, middle management and operational staff have not been shown the evidence as to why a shift can be successful, there may be restraint in moving forward. This is why a business case is made as part of any enterprise analysis.

Return on investments can shed light on why as shift is necessary. Benefits of moving can also influence and calm the cultural resistance to change. Explaining to staff the benefits of employing new ways of using and interacting with data can bring to their work will add value to the transitional planning and execution when the time comes to move forward with the data project.

8.1.3 TECHNICAL

Staff who use a particular piece of software or application become attached to it. Familiarity with software may lead to resistance to change, especially if the analysis leads to a conclusion that new software or new, more innovative ways to using tools disrupts the familiarity attributes of working with an existing software implementation. On the flip side, if IT in the past has been resistant to making change based on business needs and failed to work with business units to effect change, stoic relations can also lead to slowdowns in adoption of new technology or software, especially tools that are used to create advanced reports and business intelligence from data pools. Stoicism is a much harder stakeholder to unravel.

Using software that is either unsupported or has reached the end of its life cycle can expose an organization to unnecessary risk of software either not working and therefore requiring an expensive fix, or the corporation missing out on new and innovative features and functions available in software that is supported. The other aspect is that if the software is developed in house, there is a risk of the knowledge base being unavailable should employees leave an organization, or are removed involuntarily. The cost to onboard new IT staff up to bring them up to speed can be expensive with little or not return guaranteed.

A new solution that is being introduced into an organization should be accompanied by good after sales support agreement or service by the vendor that is charged with delivering the software to an organization. Proper training plans should incorporate what training is being provided, to whom the training is targeted and for how long. If the software is to be developed in house, the same principles would apply. Proper transition requirements would detail what specific aspects of the software requires general training and extensive training, to whom and during what windows of time for training.

8.1.4 FINANCIAL

The size of information technology budgets or capital budgets that are created to feed software projects influence the types of software that can be used to satisfy the business requirements of business intelligence projects. Appetite for open source or off the shelf software can govern which vendors to work with, and the cost of using their software.

The tighter the purse strings, the narrower the potential solutions that fit an organization's data needs, or the more creative a development team needs to be in order to sustain value over time in the solution being provided to the business. The other aspect is a vendor not willing to provide support or services as agreed to. This is possible and should fall to the procurement team or legal team to address but this is another example of financial risk, usually covered under the auspices of a contracting domain.

Make sure the contracts are adequately detailing the requirements and delivery options that are agreed to by both parties.

8.1.5 ECONOMIC FACTORS

In addition to the financial risk associated with transitions, economics means all other budgetary aspects of the transition to include aspects like return on investment, cost/benefit analysis of using various options. These inputs can be used as part of the transition analysis, including how transitions went in the past if they occurred. How successful were past projects involving similar needs? Is there a means to put a value on the costs and benefits of past projects that can feed the current transition assessment team as part of the preparatory phase.

Get creative and work with the team to come up with a plan to provide high value for the budget that is available. A culture of creativity, innovation, motivation and sharing of ideas can go a long way in providing solutions with stretched resources, like money.

8.1.6 TRAINING

Transition requirements and the more staff need to learn in terms of software use, process changes, and technological improvements will bolster the training requirements, time and resources needed to fulfill the inevitable high learning curve an organization and its staff will discover as new, analytical ways are introduced within an organization.

Staff may not be ready for training when the new software, solution or application is ready for rollout. Some segments of the organization may be ready, while others, because of the nature of their work are either all not available or part of the functional group is available. Issues like this require deeper planning, more time to coordinate who is getting training and when. The risk is that training will be stretched out beyond a reasonable amount of time as the learning curve may be high and the loss of knowledge bigger if staff cannot use the training material right away.

The volume of the transition requirements will dictate the scope of training. Like any project or requirement scope, think more is less. Assess which parts of the organization will benefit the highest from an immediate training session and schedule the other parts of the organization when they are ready, more available or generally in a position to take on the training and get full value from the training in order to use what they learn right away on their jobs. In addition, roll out software components, if possible, for those that will use it right away and are ready to get started.

9 CONCLUSION

The size of the BI project will depend on the scope and complexity of the work involved. Part of an enterprise analysis will consist of working with stakeholders to gauge whether a BI solution is to cover the whole organization or a subset of that organization. This is critical because the answer to that question will drive whether we analyze a larger or smaller data source set, and will determine whether we target a broad or narrow subject area and whether our task will be complex or simple.

As analysts, we need to ask questions at the enterprise level to associate and strengthen the connection between data and how it complements the need to measure progress towards meeting strategic enterprise initiatives.

We also need to understand that we must try to help an organization rethink traditional methods of reporting by laying out the benefits of a business intelligence concept; benefits that are in tune with the nature of the business environment under analysis. We need to continue to act as ambassadors of change, the drivers of a retooling of how data is to be perceived, used and understood by each stakeholder in an organization.

Furthermore, we need to assess whether data under analysis is meant for all stakeholders or is a subset of the organization. Determining the breadth of the data requirements will assist with scope management and level of effort required to effect and manage change to a shift to an analytical culture, or if the culture is already mature, get them to the next level of analytics.

Ideally we should augment our requirements repository with the support of tools like entity relationship models, data models and dimension models. Stress points, bottlenecks and gaps in data needs can be revealed through the use of an "as is/to be" assessment. An organization will always want to move from point A to point B. It is our business to assess the width of the gaps in terms of resources and solutions required to fill them.

Next, work with less. Use the priority matrix as a guide to understanding which data requirements should be implemented first. Don't make life complicated by accepting a list of requirements as all needing to be implemented. Negotiations and further insights will govern which requirements get implemented first. There is always room for consensus and compromise in any project.

Above all, have fun with the work. It can be challenging but rewarding, especially with the knowledge gains that come from working on a BI project and the notion that business analysts provide key support to a critical process that contributes to the achievement of an organization's overall corporate objectives using good analytical data as the foundation for effective decision making.

10 REFERENCES

International Institute of Business Analysis (2015). *A Guide to the Business Analysis Body of Knowledge*. Version 3. IIBA.

The Productivity Puzzle – Why Productivity Goes Beyond Just Technology. June 24, 2015. Retrieved from https://www.maximizer.com/blog/the-productivity-puzzle-why-productivity-goes-beyond-just-technology/

Davenport (Thomas), Harris (Jeanne) & Morison (Robert) (2010). *Analytics at Work: Smarter Decisions, Better Results*.

SAA & The Environment. 2017. Retrieved from https://www.flysaa.com/about-us/leading-carrier/social-responsibility/saa-the-environment

UN Data. A world of information. Retrieved from http://data.un.org

Kimbal (Ralph), Ross (Margy) & Thornthwaite (Warren). *The Data Warehouse Lifecycle Toolkit*. John Wiley & Sons. 2011.

Collation. Retrieved from https://mariadb.com/kb/en/sql-99/22-sql-collations-collation/

Maslow's hierarchy of needs. Retrieved from http://www.businessballs.com/maslow.htm

ENDNOTES

1. A good starting point for open UN data is the following web site: http://data.un.org/.
2. While data is owned by an organization, not with specific users, the context here implies that specific staff are generally responsible for administering data to support processes for interacting with, collecting, storing and extracting data from applications and database management systems in order to fulfill responsibilities to upper management and executives who need the information, using reporting tools as a foundation for assessment.
3. For more, please read "Analytics at Work: Smarter Decisions, Better Results" by Thomas Davenport, Jeanne Harris and Robert Morison (2010).
4. Publicly available information courtesy of South African Airways, http://www.flysaa.com.
5. Please see the corporate web site https://www.maximizer.com/, a company with interests in the Customer Relationship Management (CRM) business domain.
6. The exercise represents a list of potential data categories and business processes to illustrate the concept of where business data could be shared across operational units, or identify where data is specifically needed by a business domain.
7. OLAP and OLTP are database specific terms. They key is that OLTP, or transactional systems perform many queries and inserts, updates and deletes. OLAP, or analytical systems support queries, or selects only.
8. Ralph Kimball and Bill Inmon are two thought leaders on the concept of data marts (Kimball) and data warehousing (Inmon). The quote comes from "The Data Warehouse Lifecycle Toolkit" by Ralph Kimball, Margy Ross and Warren Thornthwaite (John Wiley & Sons, 2011).
9. If all requirements are high priority, then all of the requirements have no priority.
10. Web sites can be found at http://www.microsoft.com, http://www.mysql.com, http://www.mongodb.com, http://www.frontbase.com, http://www.postgresql.com, http://www.oracle.com and http://infocenter.sybase.com respectively.
11. Collation is a database specific term and means "a set of rules that determines the result when characters from a Character set are compared". Quote courtesy of https://mariadb.com/
12. Internationalization is a "design process that ensures a product can be adapted to various languages and regions."
13. For more on Maslow's hierarchy of needs, please visit http://www.businessballs.com/maslow.htm
14. Business analysts are not expected to know psychological terms, but concepts about human needs can play a role in understanding emotional intelligence which can be a factor in any change in any business environment.