BUSINESS ANALYSIS PRACTICE ADVANCEMENT | IIBA® PUBLICATION

INTRODUCTION TO BUSINESS DATA ANALYTICS

A PRACTITIONER VIEW

SUPPORTING DATA-DRIVEN DECISION MAKING THROUGH EFFECTIVE BUSINESS ANALYSIS

Introduction to Business Data Analytics: A Practitioner View

International Institute of Business Analysis, Toronto, Ontario, Canada.

© 2019 International Institute of Business Analysis and Agile Alliance. All rights reserved.

This document is provided to the business analysis community for educational purposes. IIBA® does not warrant that it is suitable for any other purpose and makes no expressed or implied warranty of any kind and assumes no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information contained herein.

Permission is granted to IIBA members to reproduce this document for your own personal, professional, or educational use. IIBA® members may not transfer ownership of their complimentary copy.

IIBA®, the IIBA® logo, BABOK® and Business Analysis Body of Knowledge® are registered trademarks owned by International Institute of Business Analysis. CBAP® is a registered certification mark owned by International Institute of Business Analysis. Certified Business Analysis Professional, ECBA, EEP, and the EEP logo are trademarks owned International Institute of Business Analysis

No challenge to the status or ownership of these or any other trademarked terms contained herein is intended by the International Institute of Business Analysis.

Any inquiries regarding this publication, requests for usage rights for the material included herein, or corrections should be sent by email to info@iiba.org.

Table of Contents

Chapter 1: Introduction to Business Data Analytics: A Practitioner View

1.1 What is Business Data Analytics? 1
1.2 Business Data Analytics Objectives 3
1.3 Business Analysis and Business Data Analytics 4

Chapter 2: Business Data Analytics Domains and Tasks

2.1 Identify the Research Questions 6
2.2 Source Data 10
2.3 Analyze Data 16
2.4 Interpret and Report Results 20
2.5 Use Results to Influence Business Decision-Making 23

Contributors 25

1 Introduction to Business Data Analytics: A Practitioner View

The Introduction to Business Data Analytics: A Practitioner View introduces business analysis concepts, activities, tools, techniques, skills and how they're applied when performing business data analytics related work.

Business data analytics has become an area of great interest for organizations, as it has been recognized as a means by which organizations can obtain valuable insights from data; supporting more informed business decision-making. As a result, more organizations are investing in business data analytics as a means to deliver on their strategic imperatives, innovate, and obtain competitive advantages in their marketplace. Such investments are driving the demand for more skilled professionals with business data analytics knowledge and experience.

This Introduction to Business Data Analytics: A Practitioner View explores the relationship of business data analytics to business analysis, emphasizing how analysis experience coupled with business analysis techniques and competencies can support business data analytics initiatives across the organization.

1.1 What is Business Data Analytics?

As a broad definition, business data analytics is a practice by which a specific set of techniques, competencies, and procedures are applied to perform continuous exploration, iteration, and investigation of past and current business data for the purposes of obtaining insights about a business that can lead to improved decision-making. Business data analytics can be defined more specifically through several perspectives.

These perspectives include, but are not limited to business data analytics as a:

- movement,
- capability,
- data-centric activity set,
- decision-making paradigm, and
- set of practices and technologies.

1.1.1 Business Data Analytics as a Movement

Business data analytics as a movement involves a management philosophy or business culture of evidence-based problem identification and problem-solving. In this perspective, evidence through data is the driver of business decisions and improvement. When this philosophy is in place, evidence is not chosen to support a preconception or point of view; instead, all available applicable evidence is used to make informed business decisions.

1.1.2 Business Data Analytics as a Capability

As a capability, business data analytics includes the competencies possessed by the organization and its employees. Business data analytics competency is not solely limited to the ability of an organization to complete analytical activities. It also includes capabilities such as innovation, culture creation, and process design. The capability or lack thereof may define or constrict what the organization is actually capable of achieving through business data analytics.

1.1.3 Business Data Analytics as a Data-centric Activity Set

As an activity set, business data analytics includes the actions required for an organization to use evidence-based problem identification and problem solving. Business data analytics has been defined by expert practitioners as involving six core data-centric activities:

- accessing,
- examining,
- aggregating,
- analyzing
- interpreting, and
- presenting results.

1.1.4 Business Data Analytics as a Decision-making Paradigm

As a decision-making paradigm, business data analytics is a means for informed decision-making. Through this lens, business data analytics is considered the tool of making decisions through the use of evidence-based problem identification and problem-solving.

1.1.5 Business Data Analytics as a Set of Practices and Technologies

Business data analytics is also considered a set of practices and technologies required to perform the analytics work itself. These practices can be discussed in the context of 5 business data analytics domains:

- Identify Research Questions,
- Source Data,
- Analyze Data,
- Interpret and Report Results, and
- Use Results to Influence Business Decision-Making.

1.2 Business Data Analytics Objectives

Organizational leaders frequently make business decisions based on personal expertise and instinct. Business data analytics removes cognitive and personal biases from the decision-making process by using data as the primary input for decision-making. When performed well, business data analytics can create a competitive advantage for the organization.

For example, algorithms based on weather, soil, and other conditions have been found to be more accurate in predicting the price and quality of red wine after it has been aged compared to the wine experts who influence the decision-making based on their own cognitive biases as to what they enjoy and do not enjoy in a wine.

In a broad sense, the objective of business data analytics is to explore and investigate business problems or opportunities through a course of scientific inquiry. The specific objectives of business data analytics are dependent on the type of analysis that is being performed.

There are four types of analytics methods:

- **Descriptive**: provides insight into the past by describing or summarizing data. Descriptive analytics aims to answer the question "What has happened?"
- **Diagnostic**: explores why an outcome occurred. Diagnostic analytics is used to answer the question "Why did a certain event occur?"
- **Predictive**: analyzes past trends in data to provide future insights. Predictive analytics is used to answer the question "What is likely to happen?"
- **Prescriptive**: utilizes the findings from different forms of analytics to quantify the anticipated effects and outcomes of decisions under consideration. Prescriptive analytics aims to answer the question "What should happen if we do ...?"

TIME / QUESTION TYPE	WHAT	WHY
PAST	DESCRIPTIVE What happened?	DIAGNOSTIC Why did it happen?
PAST/FUTURE	PREDICTIVE What is likely to happen based on past trends?	
FUTURE	PRESCRIPTIVE What should happen if we take a certain path? What is the best outcome given the uncertainty?	

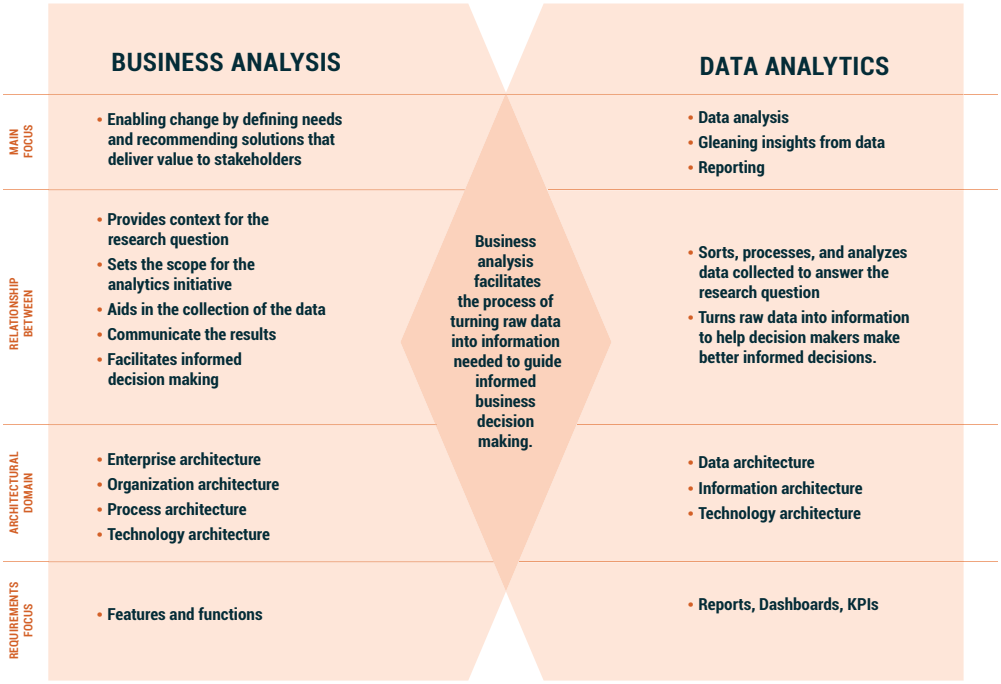
1.3 Business Analysis and Business Data Analytics

The terms business data analytics and business analysis are frequently used interchangeably. However, there are significant differences between the two terms. Business analysis is the practice of enabling change in an enterprise by defining needs and recommending solutions that deliver value to stakeholders. Comparatively, business data analytics is focused on the process of data analysis.

Business analysis provides the business context for business data analytics. Business analysis defines the focus for the research questions being asked and sets the scope before data is collected. Business analysis also aids in the collection of data and the implementation of the data collection processes. Business data analytics is used to sort, process, and analyze the data once assembled.

Once the analysis of the collected data is complete, business analysis activities are performed to interpret the results obtained from analytics; transforming information into business decisions. Business analysis activities are performed to communicate the results of business data analytics and facilitate the implementation of informed business decisions made as a result of what is learned from analyzing the data collected.

Some consider business data analytics as a specialty or subset of business analysis; one that is focused on data analysis. This viewpoint is taken since many skills and competencies often discussed when defining business analysis are equally important when performing business data analytics work. In this Introduction to Business Data Analytics, we treat the two disciplines separately and elaborate on how business analysis is utilized to perform business data analytics effectively. Even those who do not identify themselves as a business analyst but who are responsible for business data analytics within their organizations will benefit from improving their business analysis skills and competencies.



2 Business Data Analytics Domains and Tasks

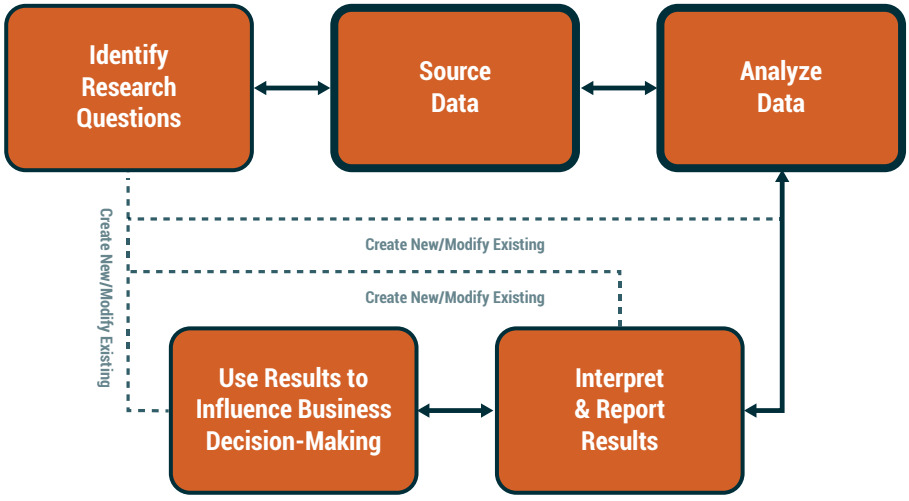
Business Data Analytics Domains and Tasks presents the practices and activities that are commonly considered business data analytics work. Emphasis is placed on identifying areas where business analysis skills are important to perform the business data analytics tasks and not on identifying the job title who would take responsibility for performing the work.

Business data analytics responsibilities can be assigned in a variety of ways and should be delegated to the resources having the best skill set to complete the work, regardless of their job title. While many of the tasks can be performed by those with the title business analyst, on an analytics initiative, anyone with business analysis skills from business analysts, data analysts, data scientists, business architects, and business SMEs may complete the work.

The term domain is used to represent a group of logically (but not sequentially) related tasks performed to achieve a specific objective.

The tasks discussed here are grouped and presented in the context of 5 business data analytics domains:

- Identify the Research Questions
- Source Data
- Analyze Data
- Interpret and Report Results
- Use Results to Influence Business Decision-Making



2.1 Identify the Research Questions

Identify the Research Questions includes the business data analytics activities such as

- defining the business problem or opportunity,
- assessing the current state,
- defining the future state, and
- framing the research question that business data analytics will be leveraged to answer.

Such questions and the subsequent data analysis performed to obtain insights supports effective business decision-making.

2.1.1 Define Business Problem or Opportunity

Although the tasks in business data analytics are iterative and not a sequential list of steps in a methodology, defining the business problem or opportunity is often the first step performed in any business data analytics initiative. It is a task where those having strong business analysis skills can assist with the work.

The analyst facilitates discussions with stakeholders to elicit, observe, and analyze through a process of continuous discovery, any and all relevant information that will help the team understand the context of the situation.

Sometimes the business is experiencing a problem they believe business data analytics can help solve, such as understanding why there is a sudden decrease in internet sales. In other situations, the business may be interested in using business data analytics to uncover opportunities - as would be the case of a manufacturing company looking to collect maintenance and performance data on its machinery to determine how to predict and avoid equipment outages. In either scenario, the analyst utilizes various business analysis elicitation techniques to obtain the necessary information required to define the business problem or opportunity that analytics might address.

When defining the business problem or opportunity, analysts utilize a number of elicitation techniques such as interviews, job shadowing, surveys, and workshops and possess competencies in facilitation, communication, and leadership. They should possess adequate business and organizational knowledge to facilitate discussions.

2.1.2 Assess Current State

Business data analytics is used to enable organizations to make informed decisions. Understanding the current state of the organization or context of the proposed change is fundamental to informed decision-making. Information obtained from a current state assessment provides contextual information so that the results of data analysis can be better interpreted.

Analyzing the current state involves understanding the business need and how it relates to the way the organization currently functions. The results of the current state analysis set a baseline and context for making a change. Whether discussing the changes associated to the implementation of a new customer relationship management (CRM) system or the process changes proposed after gaining insightful information from the results of a business data analytics effort - analyzing the current state is a necessary step.

A current state assessment can include understanding the business value chain or how data and information flow throughout the organization.

The analyst may uncover insights such as whether the organization has an appetite for analytics or budget and expertise to perform the work. They will need to become the business domain expert and understand trends and evolving business models.

When conducting a current state assessment, analysts utilize a number of techniques to elicit, analyze, and visually depict the current state of the organization such as business model canvas, organizational, scope, and process modeling. Conceptual and systems thinking along with business acumen and solution knowledge are also useful.

2.1.3 Define Future State

According to *BABOK® Guide* version 3, all purposeful change must include a definition of success. Defining the future state creates a vision of the desired outcome of the change. Defining success for a business data analytics initiative is as important as any other change initiative.

Defining the future state includes:

- ensuring the future state is clearly defined and understandable,
- that it is achievable with the resources available,
- that key stakeholders have a shared vision, developed by consensus, of the outcome being sought, and
- setting measurable objectives to ensure the desired vision is met.

To establish measurable objectives, the analyst facilitates discussions between business stakeholders to determine the types of metrics to consider. Working collaboratively, the decision-makers select the most appropriate measures to assess using business data analytics. These measures may be a combination of strategic and operational key performance indicators (KPIs). Some KPIs may focus on assessing performance for a specific geography or for a target audience. There may be industry specific metrics such as ARPU (average revenue per user) which is used in telecom or store footfall which is used in retail.

Another important aspect of defining the future state is establishing the scope for the analytics effort. Establishing scope entails understanding what areas of the business are participating in the analytics effort and determining what stakeholders have questions to raise and information to provide. A

future state, with respect to a business data analytics initiative, could also include setting a vision about the length and breadth of data analytics capabilities. For example, tracking more KPIs, increasing the frequency of reports being generated from monthly to daily/weekly, automating reporting functionality, or having data available real-time.

The desired output from defining the future state is a clear understanding of the business objectives and a firm understanding of the value the business is seeking to obtain from the business data analytics effort.

Analysts utilize metrics and key performance indicators (KPIs) and use different models to visually communicate the future state. This includes scope models to understand boundaries and stakeholder maps to identify those who might be impacted by this work. Having conceptual thinking skills helps to understand the big picture and provide the context for the analytics work. Interaction skills, communication skills, analytical thinking, and problem-solving skills are useful when leading discussions to identify metrics and establish objectives.

2.1.4 Formulate Research Question

Before any of the detailed analytics work is performed, such as capturing and analyzing data, the research question that the analytics will be used to answer is formulated with business stakeholders. Formulating the research question involves facilitating discussions to identify the different questions that could be explored, specifying the questions in easily understood language, and bringing the team to consensus as to the best set of research questions to answer.

At this point, the business is making a conscious decision to move away from traditional decision-making based on instinct and gut feel. There is rationale for why business data analytics work is needed and a desire to use the results of analytics to derive insights to guide business decision-making. The questions formed focus on those that can be answered with data.

The analyst requires the skills to identify the right problem or opportunity and to focus the team on the right question to ensure the analytics work is guided properly. Discussions move beyond brainstorming a list of ideas and target producing a concrete list of specific research questions the team believes are worth pursuing. On occasion, the team may need to work out what data are available before determining which ideas are achievable with analytics. The question, once formed, guides the scope and drives the activities of the analytics team.

The results of the analysis obtained when defining the business problem or opportunity, analyzing the current state, and defining the future state provides context when formulating the research questions. The analytics team, including business stakeholders, may start with a long list of questions and require ongoing collaboration to reduce the list identifying the highest valued questions to use. Technical resources or the analyst, based on their understanding of the data and the business problem or opportunity, may suggest research questions that could be explored.

Good research questions are clearly stated and do not use technical language. The final research question is reviewed with all stakeholders to ensure consensus. At this point, the objective is to obtain stakeholder consensus on a question that clearly articulates what it is the business is looking to answer through analytics and to do so with clear business language. In the Perform Data Analysis task, the data scientist will restate the research question using more mathematical language.

There are situations where it is easier for an analytics project to address a group of questions rather than individual analytics projects asking questions one at a time.

When formulating the research question, analysts utilize a variety of elicitation techniques to facilitate discussions with stakeholders, decision models to help the team reach consensus, and templates to guide the development of the question. Strong facilitation, leadership, and negotiation skills are useful when facilitating consensus among stakeholders.

2.1.5 Plan Business Data Analytics Approach

Planning the business data analytics approach defines how the analytics work will be performed.

When planning a business data analytics approach, analysts:

- determine the capabilities and capacity of the organization to perform analytics so the team understands what is realistically possible.
- identify 'quick wins' versus longer-term efforts.
- determine the type of analytics being asked for. For example descriptive, diagnostic, predictive, or prescriptive.

Planning is an iterative process, and changes to the approach are made as new knowledge is gained. Each domain includes an element of planning which may influence the overall approach to analytics.

There is no right or wrong answer as to the degree of formality of the business data analytics approach. Some organizations may choose to formally document the decisions made when defining their approach by using a business data analytics planning template, while other teams may choose to build more visual models to capture the decisions and include the information on shared wikis and within the team's workspace.

When planning a business data analytics approach, analysts utilize techniques such as brainstorming to quickly identify a list of activities needing to be performed, functional decomposition to break down high-level concepts into lower level tasks, and estimation to assess how long it may take to complete various activities. Analysts planning a business data analytics approach use facilitation, leadership, and negotiation skills to obtain stakeholder consensus.

2.2 Source Data

The Source Data domain includes the business data analytics tasks performed to

- determine what data is needed to answer the research question,
- identify what data is already collected and what new data is required,
- make decisions on how and when to collect the data,
- assess the quality of the data, and
- make changes to the data acquisition approach as needed.

The tasks within the Source Data domain are performed by individuals who possess strong technical skills. In some organizations, this may be a data analyst, data scientist, or a business analyst when they possess sufficient technical expertise. While data scientists see datasets as a set of variables, it is the business analyst who brings the insight to determine whether a dataset might be useful to explore within a particular business context. Business analysts understand the meaning behind data variables, in essence, the importance of the data to the business. Because of these differences in viewpoints, a well-structured data analytics team includes resources who collectively can provide both business and data science skills when sourcing data.

2.2.1 Plan Data Collection

Before data can be sourced, research and analysis is performed to determine which data is available for analytics. Some data may not be available due to privacy rules while other data may only be available during specific time-frames. Planning data collection involves the planning efforts to take into consideration

- which data is needed,
- the availability of the data,
- determining when and how the data will be collected, and
- how it will be validated once collected.

An analyst supports the data scientist with data sourcing - identifying the data required to answer the research question. This work includes determining the data that is currently collected (whether used or not) and that which is currently not collected but would be helpful in answering the question. Data sourcing involves determining which sources to use (for example, sales, financial, inventory system or data lake, mart, vault, or warehouse.) Some data may be available from multiple sources, then the task involves determining the 'best' source to use. Data sourcing often involves collaboration with the architecture team who can share valuable insights into recommended sources as well as compliance with legal regulations, data privacy, and architecture principles.

When planning data collection, consideration is also given to non-functional requirements. This includes privacy, security, retention, volume, timing, integration, and frequency requirements along with any constraints imposed by data availability and existing service level agreements. Analysts look for situations where the data may have both short-term and long-term effects on business decision-making and determine how this influences the frequency of data collection. When the frequency and timing needs for the business data analytics efforts are greater than what is currently happening, an assessment of costs to obtain the data at a more regular interval needs to occur.

Consideration is given to the level of effort required to obtain the data. Data sourced internally may be easier and cost less to obtain than data obtained from external sources. How much the data needs to be manipulated once obtained may influence sourcing decisions as well. For example, if there is a choice between obtaining data directly from a centrally managed data warehouse or pulling data from a peripheral secondary source where the data has already been manipulated into a more usable form, an assessment of data quality may be needed to help determine the best source. A direct pull of data and subsequent data manipulation may mean a little more work and overhead cost, but that might be acceptable if the pre-massaged data is questionable from a quality perspective.

Analysts also determine how much data will be structured versus unstructured and determining how much of each type is feasible to use.

- **Structured data** is data that is organized, well thought out, and formatted such as data residing in a database management system (DBMS). Structured data is easily accessed by initiating a query in a query language such as SQL (standard query language).
- **Unstructured data** is the exact opposite of structured data as it exists outside of any organized repository like a database. Unstructured data takes on many forms and sources such as text from word processing documents, emails, and social media sites or image, audio, or video files.

There is considerable more work to organize unstructured data for analysis; therefore consideration has to be made on whether the unstructured data will be useful and how it will be used. While unstructured data might be more complex, the challenges here can be minimized depending on whether the team has the necessary tools to work in this space as well as the experience and skills.

Once a data collection plan is created, stakeholders who are impacted or who possess some ownership over the data review the plan along with the analytics team who will perform the work. The analyst takes responsibility for facilitating the team to consensus in order to obtain approval of the data collection approach.

When planning data collection, analysts utilize various elicitation techniques to acquire the information necessary to build the data collection plan. Brainstorming provides a quick list of data sources to consider. Document analysis is used to identify data sources through the review of existing architecture models. Skills such as organization and solution knowledge

provide context and insights when developing a data collection approach. Problem-solving is used when identifying data sources and decision-making is used when facilitating discussions with those who must approve the data collection plan.

2.2.2 Determine the Data Sets

Determining data sets involves performing a review of the data expected from the data sources and determining specifics such as data dimensions to consider, sample size, and relationships between different data elements. It involves making a determination about which datasets need to be collected as a whole versus partial datasets, for example determining whether to use the entire spreadsheet versus specific rows within it. When source data cannot be identified, determining data sets also involves identifying data gaps. Data gaps occur when data doesn't exist or is missing due to errors such as a failure in the data collection process.

The analyst may need to collate and assess data by establishing relationships between different data and identifying data linkages between data from various sources. They may use data discovery tools or database querying tools to assess data availability.

There may also be an assessment of the five Vs (volume, velocity, variety, veracity, and value). Each of these characteristics may influence decision-making with regards to which datasets to consider.

- **Volume**: is determined by the amount of data being produced and the size of the data sets needing to be processed
- **Velocity**: is determined by the speed at which data is generated and the frequency by which the data needs to be collected
- **Variety**: is determined by the variety of data sources, formats, and types needing to be processed
- **Veracity**: refers to the trustworthiness of the data, and that which presents uncertainties and inconsistencies in the data.
- **Value**: refers to the necessity of driving any data analytics exercise from real, valuable business goals

Non-functional requirements and existing service level agreements may constrain the availability of data. For example, privacy or security considerations may deem a dataset unfit for use.

Analysts possess a firm understanding of the lexicon used by the different business units and are capable of drawing comparisons and relationships between different data sets having the same meaning. Analysts also possess strong visualization skills and contribute to creating architectural diagrams that depict the data sources and frequency of the data feeds. Such models are essential when facilitating discussions about data sourcing with stakeholders and facilitating approvals.

Analysts also support data scientists by analyzing the cost versus the benefits of different data sets. It is ideal for the analytics team to collect their own data

from scratch, but frequently there are not enough resources to do so. The analyst advises on the advantages and disadvantages of using different data sets from a cost, value, timing, risk, and feasibility perspective. This is especially important when the data needed for business data analytics must be acquired from an external 3rd party. Certain research questions may need to be dropped when it is determined too expensive to obtain the data needed to answer it.

When determining data sets, analysts utilize a variety of techniques to help them work with and understand the data before building their analytical models. Data profiling is used to assess data quality. Data sampling is used when breaking a large source of data into a smaller more manageable set of data. Sampling helps an analyst reduce the amount of data they have to work with as it provides a means to use a representative subset of the larger population. Skills such as creative thinking and conceptual thinking are useful when formulating ideas about which data to use as is having sufficient business acumen to determine which data sets may be best to use based on the current business situation.

2.2.3 Collect Data

Collecting data involves the activities performed to support the data scientist with data setup, preparation, and data collection. The degree of involvement analysts have with data collection depends on how the organization structures the analytics team as well as the technical abilities of the analyst.

Before the data scientist begins collecting large amounts of data, it may be necessary to test the data collection approach by using a small number of observations. If the data collection method is a survey, this task might involve piloting the survey with a small population of participants before performing the survey with the larger population.

When collecting data, analysts:

- determine if the data will be coming from different files/sources,
- identify where the data is going to be collected from (for example, database, spreadsheet, and other sources.) and,
- understand the history of data in order to assess data quality. This is referred to as data lineage.

When data is collected from different sources, analysts determine if the disparate sources represent the same data in the same way. For example, if data source A uses numeric codes to specify gender and data source B uses alpha codes, the need for reconciling data elements across sources needs to be identified.

The file format for the output produced from each source is also identified. Further analysis determines if the data needs to be formatted prior to merging it into a single file. For example, will spaces need to be removed when moving data from a text file to a spreadsheet? Will data formats need to change so data is consistent between sources?

As data is collected, it should be analyzed to identify potential problems with the data collection approach.

When collecting data, analysts leverage techniques such as surveys and experiments. Data collection is performed using automated tools over manual processes. Data analysis skills determine what data to use, how to collect it, and with understanding its relevance and relationship to what is being analyzed. Demonstrating skills such as trustworthiness and ethics helps to build trust and rapport with stakeholders who may be needed to gain access to data or participate in elicitation activities. Business acumen is necessary during the testing of the data approach and when profiling data.

2.2.4 Validate Data

Validating data involves validating that the planned data sources can and should be used and when accessed, the data obtained are providing the types of results expected. Since detailed analysis of the data has yet to be performed, the objective of validation at this point is high-level.

Business validation involves having the business stakeholders approve the data sources and establish the acceptance criteria that define the parameters for assessing the accuracy of the data. It also includes validating any relevant requirements. For example, if the outcome of data analysis is expected to be a report, business validation involves validating the format and data elements to be included in the report.

Technical validation involves the technical testing and validation to assess data quality. There are a number of characteristics reflected in high-quality data, such as:

- **Accuracy**: means that the data is correct and represents what was intended by the source. Accurate data is not misleading. Accuracy might be assessed by comparing numbers displayed by a front-end system with data retrieved from the database.
- **Completeness**: means that the data is comprehensive and includes what is expected and that nothing is missing. Completeness might be assessed by ensuring required fields do not include null values.
- **Consistency**: means how reliable the data is. Data values are consistent when the value of a data element is the same across sources. Consistency might be assessed by ensuring only date values are being displayed in date fields.
- **Uniqueness**: data that is unique will be valuable to the organization. Uniqueness might be assessed by determining whether any duplicates exist in the data.
- **Timeliness**: data that is fresh and current is more valuable than data that is out of date. Timeliness might be assessed by determining whether the data being received is for the period being requested.

Data validation may be performed by a data analyst, data scientist, or business analysis practitioner with sufficient skills to use the necessary tools to access data and the underlying competencies to analyze the results.

Business validation is performed by key business stakeholders with the authority to approve data sources for use in analytics initiatives.

When validating data, analysts utilize techniques such as data mapping and business rules analysis. Data mapping is used to create a source-to-target data map to define the mapping between the data sources being used. Business rules analysis provides the understanding of the business rules governing the data providing the guidance as to what should be validated.

Conceptual thinking skills help make sense out of the large sets of disparate data sets under analysis and to draw relationships and understanding from the data. Business knowledge provides context to the data being validated, helping the analyst determine if the data is accurate and complete.

2.3 Analyze Data

Analyzing data involves deciding how data analysis will be performed, including which models and mathematical techniques will be used.

Analyzing date involves

- preparing the data for analysis,
- performing the data analysis,
- determining whether the analytical solution/results are helping to answer the business question, and
- making adjustments to the approach when it does not.

When analyzing data, the business analysis practitioner is more likely to support the data scientist than to be responsible for running the analytical models themselves. A business analyst brings the required domain knowledge for the area under analysis to the team, providing context to the problem or opportunity the analytics effort is aiming to address. The data scientist requiring a strong understanding of the business considerations, may not possess that knowledge themselves. Strong collaboration between the data scientist and the business analysis practitioner ensures that the analytics work is performed within the correct business context.

2.3.1 Develop Analysis Plan

The analysis plan may be formal or informal. The objective is to ensure sufficient time to plan the data analysis activities required for the initiative.

When developing the analysis plan, the analyst answers determines:

- which mathematical techniques the data scientist plans to use,
- which models are expected for use? (Such as regression, logistics regression, decision trees/random forest, neural nets),
- which data sources will be used and how data will be linked/joined, and
- how will data be preprocessed and cleaned.

The business analysis practitioner provides insights into the plan or may draft the initial plan for review by the data scientist. It is the data scientist who possesses deep technical expertise to decide how the data analysis will be conducted. Business analysis skills are applied by ensuring sufficient information about the business domain is provided to the data scientist, so an effective approach to data analysis is structured.

When developing the analysis plan, if the plan is formally documented, templates ensure consistency of plans across initiatives and guide the planning decisions needing to be made.

The metrics and key performance indicators for the initiative are required for the data scientist to determine if the outcomes from data analysis are producing the results required to address the business need.

Organizational knowledge helps business analysis practitioners provide the context for the data scientist's work.

2.3.2 Prepare Data

This task entails obtaining access to the planned data sources and establishing the relationships and linkages between sources in order to create a coherent dataset that can be used to achieve the objectives of the data analysis effort. The data scientist identifies how different datasets are related, considers whether the data can be linked in theory, and decides whether it can happen in practice.

Preparing data involves understanding the relationships that exist between data. For example, do two tables have a 0 to 1, 1 to 1, or 1 to many relationship? Preparing also involves establishing the joins or linkages between sources, normalizing data to reduce data redundancy, and converting data. Sometimes the data collected is unrecognizable and must be transformed to lend value to the analytics effort. Data cleansing is a process by which data is transformed to correct or remove bad data. The data scientist identifies the rules for consolidating data, performs the consolidation, and then validates the results to see if the business rules are being adhered to. Any mechanisms the data scientist builds to automate the data acquisition or preparation processes can be elevated for use by other analytics teams.

Data scientists leverage a host of techniques when preparing data. Weighting is one technique applied to data to correct bias. Sample weights can be applied to address the probability of unequal samples and survey weights applied to address bias in surveys. For example, survey weights may be applied when there is a need to correct an uneven representation of subpopulations in a survey population (for example, more women have answered the survey, than men.) Weighting is used to correct for the uneven representation. CRUD (Create, Read, Update, and Delete) matrices provide an easy method for specifying business rules surrounding how data can be managed. Data scientists utilize strong technical skills and knowledge of statistics when preparing data for use in an analytics initiative.

2.3.3 Explore Data

Exploring the data involves performing initial exploratory analysis to ensure the data being collected is what was expected from the data sources. It provides a form of quality check to ensure the right type and quality of data is being obtained prior to executing more detailed data analysis work. Data exploration is primarily the responsibility of the data scientist, but the work is most effectively performed when paired with an understanding of the business domain; an area where a business analysis practitioner can lend much assistance.

Exploratory analysis involves obtaining a subset of data and identifying initial trends and relationships to develop a fair understanding of the value the data is providing. The data scientist looks for data gaps or data redundancy that signal data may need to be cleansed or data outliers (noise) that signal data may need to be excluded. Missing data from a survey could mean that a person is missing from the dataset or that a person might have only answered certain questions on a survey. The data scientist makes an assessment of the

data quality to determine the course of action using the following check points:

- **Data integrity**: Can the data be trusted? For example, is the required data present?
- **Data validity**: Is the data truly representative of an underlying construct? For example, is IQ a good measure of a person's intelligence?
- **Data reliability**: If data is collected more than once, will the same results be obtained? For example, will a survey respondent answer a question differently on different days of the week?
- **Data bias**: Does the data portray an accurate picture of a given situation? For example, are employees over-estimating the quality of their work or do we have a situation where the survey participants are not a representative sample?

Where possible and when required, data scientists transform data, removing unrecognized data characters or converting data to a consistent data format when disparate data formats exist. If the data collection processes are not providing a sufficient amount of good data, the data scientist determines a new approach to sourcing the data. This may involve establishing new joins or relationships between data or identifying completely new data sources. It might be necessary to go back and consider whether there are other datasets that could be collected if the first dataset is not usable. Once there is an assurance that the data sources are providing the right data, what is learned from the exploratory analysis can be used to guide the approach taken to perform the detailed data analysis.

When exploring data, analysts utilize data mining to identify information or patterns that require further investigation. Data scientists utilize a host of data discovery and profiling tools to mine data. They use models to determine data quality. For example, histograms can be used to understand the distribution of values across variables. Feedback loops are used to allow for adjustments to be made about what techniques and models best fit the data. Ongoing collaboration between the data scientist and the business analysis practitioner pairs the industry and business domain knowledge possessed by the business analyst with the analysis results produced by the data scientist to determine whether the results are helping to answer the business question.

2.3.4 Perform Data Analysis

Data analysis involves the extensive deep analysis performed once the data quality issues are resolved through exploratory analysis. Performing data analysis involves the application of mathematics and the completion of extensive mathematical analyses related to answering the research questions for different stakeholders. Where exploratory analysis tested the dataset, performing data analysis involves using the results of exploratory analysis to determine the best mathematical methods and approaches to utilize and then conducting the in-depth data analysis required to answer the research question. The original question raised in business language is transformed into a mathematical question, which is loaded into a model to perform the deeper analysis.

When performing data analysis, data scientists utilize technical techniques requiring extensive mathematical skills to use. Some techniques are leveraged to find associations or to cluster data, which is helpful when identifying patterns (for example, association rule learning, classification tree analysis, and K-means clustering.) Many techniques, such as the use of machine learning and artificial intelligence, advance the data scientist's analysis capabilities. Data scientists use regression analysis to predict and forecast. Simulation is used to play out a series of actions or behaviors.

Analyzing data requires extensive application of statistics. Data scientists use creative thinking skills to determine different approaches for answering the research question, especially when the data results are not helping to achieve the stated objectives. As with exploratory analysis, the data scientist uses industry and business domain knowledge, and when not present, these skills can be augmented by leveraging the skills of the business analysis practitioner.

2.3.5 Assess the Analytics and System Approach Taken

Assessing the analytics and system approach taken involves collaborating as an analytics team to determine whether the results from data exploration or data analysis are helping to answer the business question. Assessing the analytics approach is performed iteratively with Explore Data and Analyze Data.

When issues arise with data sourcing or with the results of data analysis, the approach to analytics adapts. For example, when data exploration uncovers issues with data quality or determines the wrong data is being collected, or data gaps are an issue, there may need to be adjustments made to how and where the data is being collected. If the results of data exploration are acceptable it is still possible that the results from data analysis fail to answer the questions being asked. The results from data analysis may not produce results that help meet the objectives of the initiative.

In these scenarios, data exploration and data analysis tasks are repeated. Iteration occurs between the data exploration and data analysis tasks until the data scientist is comfortable with the data sources being used. Their assessment is based on the quality of data being obtained and its value toward answering the research questions.

When assessing the analytics and system approach taken, business analysis practitioners require basic skills in statistics and a basic understanding of data science tools and technologies. They should possess sufficient business acumen to provide context to the analysis space. Business analysis practitioners answer questions the data scientist may pose related to the business.

Adaptability is necessary to adjust the analysis approach as more data is uncovered, new insights learned, or different levels of stakeholders are involved. Trustworthiness is important as in some industries, having access to certain types of data comes with a great deal of responsibility, often with legal implications. It is important to know what is acceptable data use and what is not and what can be accessed or viewed and what cannot.

2.4 Interpret and Report Results

Interpreting and reporting results involves utilizing the results obtained from data analysis to glean insights from the data collected and determining how best to communicate and report the outcome from business data analytics to interested stakeholders. The outputs from Interpret and Report Results are used to influence business decision-making.

2.4.1 Identify and Understand the Stakeholders

Identifying and understanding stakeholders allows analysts to effectively engage and collaborate with the variety of stakeholders engaged in an analytics initiative. Each stakeholder group

- articulate different needs and objectives,
- pose different types of research questions,
- are interested in different volumes and timings of analytics results,
- hold different skill-sets for interpreting those results, and
- possess different levels of education in, and experience with analytics.

Understanding the unique characteristics of each stakeholder group increases the analysts' effectiveness with each group.

Before an initiative starts, analysts seek to answer the following questions::

- Who are the stakeholders?
- What is their level of knowledge about analytics?
- What aspect of the project is of interest to them?
- What communication methods and techniques are appropriate?
- When should stakeholders be communicated to?

When identifying and understanding stakeholders, analysts utilize techniques to facilitate the identification of stakeholders such as brainstorming, interviews, or the process of reviewing process flows or organizational charts. They use models to depict aspects of their stakeholders such as stakeholder matrices and onion diagrams, and create or review personas. Analysts build models to relate the enterprise strategic goals to the organizational goals and objectives and the stakeholders or stakeholder groups impacted. Using facilitation and communication skills along with knowledge about the business or specific organizations will support an analyst when performing stakeholder identification.

2.4.2 Plan Stakeholder Communication

Planning stakeholder communication identifies what needs to be communicated, to whom, how it needs to be communicated, and when it needs to be communicated. Planning stakeholder communication on a business data analytics initiative is similar to most other initiative-level communication planning. It requires the analyst to know and understand who the stakeholders are and the communication preferences of individual stakeholders and stakeholder groups.

2.4.3 Determine Communication Needs of Stakeholders

Determining the communication needs of stakeholders increases the effectiveness of communication by allowing for the customization of communications to individual stakeholders so that the message is clearly understood.

Understanding the characteristics of the stakeholders, an output from stakeholder analysis provides guidance when planning and determining a communication approach. Stakeholder communication requirements may include stakeholder preferences

- with regards to what information they wish to see,
- how stakeholders wish to receive information, and
- how often they wish to be informed.

Each stage of the analytics process is communicated in a variety of ways, particularly with regards to the research proposal and the resulting conclusions. The analyst uses their understanding of the stakeholders to determine how best to communicate.

When planning stakeholder communication, analysts utilize a variety of elicitation techniques to identify the communication needs of stakeholders and to define the best approach to share the results from analytics. Retrospectives or lessons learned are ways by which information with regards to what is working or has worked can be leveraged to inform the communication plan. Using facilitation and communication skills along with knowledge about the business or specific organizations enable the development of a well thought out communication approach.

2.4.4 Derive Insights from Data

Analysts utilize a variety of models to derive insights from the data collected. These models are developed with a variety of data visualization tools. Analysts select the appropriate type of chart for visualization to perform comparisons, to depict distribution patterns, relationships, and compositions. A variety of visualizations such as graphs, charts, and reports are used to assure that meaningful, usable analytics for the business is attained.

When deriving insights from data, analysts read and interpret graphs and models. They use data visualization to communicate the summarized results to stakeholders. Organizational skills are important in order to process information and review and assemble the results in an organized fashion. Analysts also require the ability to view results from a holistic viewpoint.

2.4.5 Document and Communicate Findings from Completed Analysis

Documenting findings from the completed analysis involves presenting the results back to the stakeholders who initiated and had interest in the research. It includes identifying how to best package and communicate the data analysis results, making decisions about the level of summarization required, and how best to group information for optimal understanding.

Analysts highlight the main themes, synthesizing results in order to build a narrative that can be understood by the intended recipients. Depending on the communication needs of the stakeholders, they may also produce executive reports and analytics dashboards.

The main aspects to consider when reporting results are:

- What are the most important aspects of the conclusions for each stakeholder?
- Is there a graph or other form of visual representation that can communicate the information more effectively?
- What method of communication is going to be most effective to display the results in a meaningful way?
- Is there a way to make the communication more engaging (for example, a video or visualization rather than a pure text report)?

An important aspect to consider when documenting and communicating the findings from an analytics initiative is to let the data drive the conclusions. Any insights reached should be based on the data collected. In other words, let the data speak for itself.

The findings include the results as well as an explanation of the methods used in the analysis, the process followed to derive the results, and any limitations or weaknesses in the data or methods used. When building the narrative, questions such as “Where are the data gaps?“, “What does this mean for the business?“, “How can the business improve?“ are addressed.

When documenting and communicating findings from completed analysis, analysts utilize such techniques as data storytelling and data visualization.

Data visualization provides visual models in order to communicate data relationships and results. The objective is to pictorially communicate information that would otherwise be too complex to relay in the written word. Through visualization tools, static graphs, and charts are turned into dynamic models that decision-makers can use to view resulting analytics information from different perspectives and level of granularity.

Analytical storytelling involves the development of a narrative about the results of data analysis and writing it in a manner where the data is used to tell a story about the patterns, trends, and behaviors observed. The story provides context to the situation being investigated through analytics with the objective of providing supporting information for organizational decision-making. It is here that the fundamental value proposition for business data analytics is demonstrated; as the organization replaces their decision-making process based on instinct with one that is built on evidenced-based decision-making.

Data storytelling and data visualization work together to enable clear, concise, and visually appealing communication. These techniques are best performed by those who are visual thinkers and have effective writing skills.

2.5 Use Results to Influence Business Decision-Making

Use Results to Influence Business Decision-Making involves using the results of business data analytics to influence and make informed decision-making.

2.5.1 Recommend Action

Before an analyst can recommend changes to address the business need, evaluation to determine how successful the analysis was is conducted. Did the outcome of the analytics work answer the research question? How well did the analysis address the business need?

The activities performed within the 5 domains of business data analytics are iterative. When the outcome is not what was expected, if the data doesn't deliver the kind of insights required, if there is no feasible solution yet to address the business need, then the business data analytics cycle repeats, starting with the formation of a new research question.

If the analysis was sufficient to provide valuable insights to drive business change, then the effort switches to using the results to drive conversations about how the changes will be made. These possibilities are referred to as solution options. Solution options proposed may include elements of process, tool, resource, or IT system changes.

Analysts elicit the types of solution options the business might consider to address the business need, rating and ranking the options, and proposing a recommendation back to the decision-makers based on the analysis and insights gleaned from the analytics efforts. Business analysis practitioners are skilled at identifying solutions that align to the strategic direction of the organization, are valuable, provide a return for the needed investment, and address stated KPIs. This is typically an area where data scientists need support to make the connection back to the business.

Changes resulting from a business data analytics initiative are prioritized, funded, and initiated like all other change within the organization. Analysts play an important role in explaining the options and setting into motion the work required to move forward on making the recommended organizational changes.

When recommending solution options, analysts utilize financial analysis techniques to determine the potential value of the various options. Focus groups are used to obtain feedback from participants with regards to the options under consideration. Many types of models whether they are depicting processes, scope, or various elements of the organization are used when making a recommendation or explaining a particular solution. Creative thinking, problem-solving, systems and conceptual thinking are all skills used by analysts when recommending actions.

2.5.2 Develop Implementation Plan

An implementation plan outlines the implementation strategy and includes a list of the tasks needing to be completed to ensure successful implementation of a change. Implementation plans for an analytics initiative look and are used no differently than implementation plans for other types of projects the analyst may work on.

The plan will include tasks, sub-tasks, resources, high-level estimates provided by the stakeholders responsible for completing the tasks, and a sequence showing flow and task dependencies. Constraints, assumptions, risks, and dependencies are also identified and discussed.

When developing an implementation plan, analysts break down the work that needs to be completed to implement the proposed changes. Functional decomposition is a technique that is used to drill down high-level tasks into lower level tasks and activities. This often takes the form of a work breakdown structure. Brainstorming is used early on to identify an initial list of tasks for the plan as can any number of elicitation techniques to obtain the information needed to develop this plan. Skills in facilitation to lead planning discussions are helpful; when developing an implementation plan.

2.5.3 Manage Change

In business data analytics, the analyst supports the change management team or possibly holds the role of change manager to oversee the transformation of the analysis results into implemented policies and procedures within the organization. Implementing change is the end goal, and it is where the organization will realize the value from its analytics efforts. Business analysis practitioners are skilled in fulfilling the role of change manager as they ensure the continuity between the analytics work and implementation.

Before implementing changes, business stakeholders must agree on which changes to make. Similar to other types of projects, there is a level of effort required to analyze the options, understand the constraints, risks, assumptions, costs, and value proposition for each option before decision-makers decide on implementing some type of change, including changing procedures or policies. The analyst plays an important role in facilitating discussions, explaining options, and driving the decision-makers to a decision.

When managing change, analysts utilize various types of models to help communicate existing processes and workflow within the current organization. The same models can then be used to show any proposed changes being suggested. Organization models, process models, sequence diagrams are a few of these. Systems thinking skills are helpful to understand the people, processes, and technologies and how best to make changes based on analytics results. Analysts leverage teaching skills to communicate proposed procedural changes effectively. Decision-making skills aid facilitating agreements on the types of changes to be made.

Contributors

Authors

- Anne Tixier
- Laura Paton, MBA, CBAP, IIBA-AAC, CSM (Chair)
- Leelyn Cruddas
- Dr. Mark Griffin

Reviewers

- Angela Weller, CSM, CSPO
- Anna Sloan
- Charlotte DeKeyrel
- Darcey Leischner, CSM
- Jodie Kane, CPBA, CSM, CSPO, CPBI
- JoJo John
- Kunal Joshi, PMP, CBAP
- Melanie Lee, MSc, CBAP, CSPO
- Melody Hicks
- Parvathi Ramesh, CA, CBAP, CISA
- Ramanpal Singh Anand
- Sruti Chandra, MBA, CBAP
- Swaroop Oggu

About IIBA

International Institute of Business Analysis™ (IIBA®) is a non-profit professional association founded in 2003 to support the recognition and advancement of the business analysis profession and discipline.

Through a global network, IIBA connects Members, Chapters, Corporations and Partners around the world.

As the voice of the profession, IIBA works to maintain global standards of practice, certifications, professional development and networking opportunities.

To become an IIBA member, visit iiba.org /membership/become-a-member.

To find your local Chapter, visit iiba.org/ membership/chapters.

IIBA Publications

IIBA publications offer a wide variety of knowledge and insights into the profession and practice of business analysis for the entire business community.

Standards such as *A Guide to the Business Analysis Body of Knowledge*® (*BABOK*® *Guide*), *The Agile Extension to the BABOK*® *Guide*, and the Global Business Analysis Core Standard represent the most commonly accepted practices of business analysis around the globe.

Whitepapers and introductions extend the practice of business analysis beyond the global standards and explore new and evolving areas of practice where business analysis is making a significant impact.

Thought Leadership papers provide leading edge guidance and best practices that help to deliver better business outcomes.

Reports and research inform and challenge traditional thinking on wide variety of topics that are of interest to the community.

For more information on IIBA certifications, visit www.iiba.org.

IIBA® International Institute of Business Analysis™