The Elusive Data Scientist:
Real-world analytic competencies

Gregory S. Nelson, MMCi, CPHIMS
Monica Horvath, PhD
ThotWave Technologies, Chapel Hill, NC

Abstract

You've all seen the job posting which looks more like an advertisement for the ever-elusive unicorn. They begin by outlining the required skills that include a mixture of tools, technologies and masterful "things that you should be able to do." Unfortunately, many such postings begin with restrictions to those with advanced degrees in math, science, statistics, or computer science and experience in your specific industry. They must be able to perform predictive modeling, natural language processing and, for good measure, candidates should only apply if they know artificial intelligence, cognitive computing, and machine learning. The candidate should be proficient in SAS, R, Python, Hadoop, ETL, real-time, in-cloud, in-memory, in-database and must be a master storyteller. I know of no-one who would be able to fit that description and still be able to hold a normal conversation with another human.

In our work, we have developed a competency model for analytics which describe nine performance domains that encompass the knowledge, skills, behaviors, and dispositions that today's analytic professional should possess in support of a learning, analytically-driven organization. In this paper, we will describe the model and provide specific examples of job families and career paths that can be followed based on the domains which best fit your skills and interests. We will also share with participants a self-assessment tool where they can see where the stack up!

Introduction

Despite the volume of readily available health data, healthcare is still learning how to systematize data-guided conversations. Even data-appreciative executives struggle with demonstrating immediate ROI of analytics. But evolving reimbursement policies have shown that the risks are grave for organizations who cannot leverage data among payers, providers, and patients in the clinical ecosystem. As a result, the analytic capabilities must include those processes which turn that historical data into predictive and prescriptive interventions which guide the trajectory of the organization in its pursuit of improved patient care. Developing staff to achieve future capabilities on an organizational level requires examination of the desired analytic functions, skills, and abilities that guide analytic investigations to their end state, that is, delivering actionable insight to the business.

The Analytics Lifecycle

Analytics is a lot more than just assembling a group of data scientists. There is a whole sphere of activities that analytic leaders must be concerned with in order to meet their business goals. Project managers, product managers, process architects, business analysts, and technical developers are all key to achieving analytic success. Organizational capabilities must include the following elements which comprise the analytics life-cycle:

- **Business Analysis -** prioritize requests, clearly articulate the problem, capture and document requirements, and assess the potential solutions using data and advanced analytics

- **Data Exploration –** identify what data is required to answer a question, acquire the data and harmonize, rescale, clean, and prepare data for analytics, as well as explore and characterize the data

- **Quantitative and Qualitative Analysis –** use a variety of techniques that include data visualizations, descriptive and inferential statistics, and advanced analytics. Support the data storytelling to help solve existing problems or anticipate the unexpected

- **Communication of Results –** as analytic and data insights leave the "laboratory" we must champion the results through understanding and action where we anticipate the challenges and consider how the results can be acted upon and operationalized

- **Data Product Life-Cycle Management –** view analytic models as "data products" requiring the design, implementation, testing and deployment as a professional responsibility. Product management includes the proactive management of knowledge, change, quality processes, project execution, program evaluation and team management

In Figure 1 we show our perspective of the analytics lifecycle. It begins with a definition of a question that addresses a business problem. The loop is closed when analytic insights are operationalized into the healthcare business workflow in some way. It is important to understand that each phase of moving through analytics is not a clean, sequential step. One may need to extract and explore data before going back to refine the question. Exploration and analysis can reveal that you do not have the right data sets at your fingertips.

But if this process is followed appropriately, then it should be rare that the operationalization stage fails due to a requirement missed during the earlier steps.
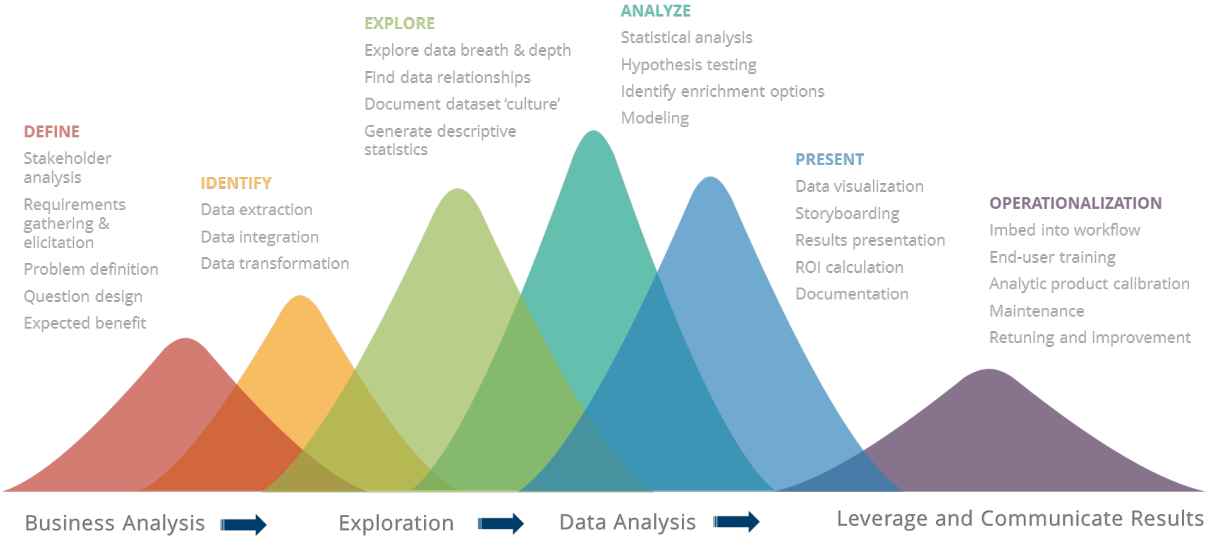

DEFINE
Stakeholder analysis
Requirements gathering & elicitation
Problem definition
Question design
Expected benefit

IDENTIFY
Data extraction
Data integration
Data transformation

EXPLORE
Explore data breath & depth
Find data relationships
Document dataset 'culture'
Generate descriptive statistics

ANALYZE
Statistical analysis
Hypothesis testing
Identify enrichment options
Modeling

PRESENT
Data visualization
Storyboarding
Results presentation
ROI calculation
Documentation

OPERATIONALIZATION
Imbed into workflow
End-user training
Analytic product calibration
Maintenance
Retuning and improvement

Business Analysis ➡   Exploration ➡   Data Analysis ➡   Leverage and Communicate Results

Figure 1: The healthcare analytics lifecycle

Analytic Capabilities and Organizational Design

The goal of linking analytic capabilities to organizational design is to ensure that there is a framework in which specific roles, jobs and teams are defined and configured. The purpose of this process is to answer the question: how do we need to be organized for the changes to work? It encompasses all the building blocks of a business – formal and informal structures, internal processes and systems, relationships, people, and knowledge.

There are three primary drivers of organization strategy, design, and development: culture, strategy, and value. Organizations must be designed to reflect not only where the company is now relative to strategy, philosophy, and the value propositions of its customers but also, where it will need to be to achieve a competitive advantage in the future. We typically do this in four steps:

1. Adjudicate organization design principles
2. Design organization structure
3. Define roles, jobs and teams
4. Estimate the organization size and composition

Note that these steps will highly depend on where an organization is with regard to their structure. Some organizations are in an early state without a clearly defined structure while others have an existing framework for analytics. For the latter, focus should be placed on ensuring clarity in roles and responsibilities. The following table summarizes the key outputs of the Organization Design processes.

3

Tool	Description
Career Map	The Career Map provides interrelates competency levels across job families and provides a path for employee advancement by being multifunctional.
Job Definitions	The Job Definition deliverable documents groups of roles into individual jobs and then map competencies to those jobs.
Job Descriptions (and Job Families)	The Job Descriptions deliverable documents the next level of detail from the Job Definitions deliverable. Each organization is likely to have its own template but a Job Description typically includes the following elements: job purpose, responsibilities, associated competencies, education and experience requirements and reporting relationships. Job Descriptions should map directly to Job Families that are used in the Team Design.
Job Gap Analysis	The Job Gap Analysis deliverable provides a mapping of current (as-is) to future (to-be) jobs together with a competency gap analysis.
Skills and Competency Matrix	The Skills and Competency Matrix provides a maps of skills needed to key competencies required in order to meet the future state organizational capabilities desired.

Table 1: Organizational Design Outputs

Job Families and Roles

Critical to our strategy is the notion that organizational (analytic) leaders want to improve the competencies of their people. In our experience as consultants, educators, and managers of analytic teams, we have identified five job families that can cover the variety of roles and responsibilities needed to perform the analytics life cycle.

BUSINESS ANALYSIS

Business Analysis refers to the core capability of achieving organizational goals by combining business knowledge, clinical workflow, and data analysis within a continuing improvement mindset. Within this family, there are those roles with a greater technical emphasis as well as those that focus more on information synthesis. All roles share a strong command of basic requirements analysis as well as quantitative skills either in the analysis or management of data. Important competencies include a strong command of business workflow, knowledge management, feasibility assessment, data-driven change leadership, and business impact assessment. Typically, specific Business Analysis roles have other functions that overlap with Statistical Analysis, Technical Data Analysis, or Project Management job families.

STATISTICAL ANALYSIS

Statistical Analysis refers to the core capability of analyzing data for insights and solutions that address business challenges. Typically, this is done using advanced knowledge of statistics, data visualization, and some algorithmic programming. Roles in this job category are expected to be highly consultative with business owners across the enterprise as they produce information to be consumed by wider audiences include senior leaders, researchers, frontline care staff, and even patients. Core competencies include exemplary analytic

thinking, visualization, and storytelling. Specific roles that may be more technical in nature and include a data programming component will overlap the competencies emphasized in the Technical Data Analysis job family.

TECHNICAL ANALYSIS

Technical Analysis spans a variety of technical roles where quantum data and data products are cleaned, manipulated, modeled, and transformed into substrate that can be leveraged by those who seek insight from enterprise data.   While tools to accomplish this come and go, the ability to adopt new methods quickly as well as move between tools is important for the modern analytics team. Roles in this job family understand the implications of technology frameworks on the ability to organize, retrieve, and share data insights. Core competencies include data wrangling, data profiling, tool agility, and systems thinking.

LEADERSHIP

The Leadership family include both line managers and director-level leaders that guide analytic teams. They assist other parts of the organization in consuming data and analytic products as to impact decisions regarding how the health business functions. At their appropriate role level, they bring together business, quality, technical, and analytic interests of the enterprise to drive collaboration, best practice sharing, and deployment of shared intellectual assets to achieve strategic goals. Leaders must have significant knowledge of healthcare culture as well as the workflow of those staff that they lead. They also must have exemplary capabilities in design thinking, data-driven decision making, analytics evangelism, and the maintenance of strategic alignment.

PROJECT MANAGEMENT

The Project Management job family is a catch-all to describe those roles that focus on developing, managing, and enforcing process around products, projects, and portfolios. They are essential to the Analytic Lifecycle and govern many of the processes that are core to turning insights into action. They will scope projects, maintain project plans, set team priorities, and even mentor teams in effectively using good processes.   Some individuals in this role also manager small teams. They typically also have strong healthcare domain knowledge and excel at consensus building as part of aligning projects to organizational strategy. These managers increasingly use agile methodologies in their development activities.

Knowledge Domains

Knowledge Areas, Competencies and Objectives

'Knowledge,' 'skills,' and 'competencies' are words used without precision when people speak of talent development. But for this discussion, it is important to define our terms clearly so that the process of aligning analytics to desired organizational capabilities is clear. With this in mind, Table 2 lists the concepts important to this discussion.

Tool	Description
Knowledge area	Body of information that a person carries with them that allows them to perform competently in completing a certain job function
Skill	The actions taken to perform an activity that can be readily measured by a performance assessment
Competency	The collective knowledge, skills, and behaviors that affords an individual the ability to adequately perform a function
Ability	The power to perform a specific activity at a specific point in time
Objective	A specific goal to meet that contributes to the demonstration of competency at a certain level of proficiency
Proficiency	A specific level of achievement that is attainable given a certain competency
Capabilities	The process that can be deployed and through which individual competencies are applied
Assessment	An evaluation of a learner's ability to meet an objective by demonstrating a particular set of skills

Table 2: Learning and development concepts

Healthcare Analytics Competencies

A major challenge in building an analytics team is defining the blend of skills that suit team mission and the enterprise culture. Even when positions are filled, there is a continued need to train and fine-tune the staff blend of greenhorns, experienced hands, communicators, and programmers who all share the analyst title. And even if a health organization successfully lures an established data scientist, the chances are good that they require intensive mentoring and education to understand the business environment that creates patient care data. Another barrier to health analyst effectiveness is communication. A good analyst uses the language of science to evaluate a problem in a structured manner, but business rarely employs such terms.   They want to 'improve care' or 'reduce costs', and their lack of precision must be resolved by the analyst as to turn it into something that can be precisely programmed or modeled. At the same time, the analyst must have the confidence to question a clinical leader's directives if they feel that the approach is flawed. This all occurs in the context of another challenging reality of healthcare, that is, the data can have grave quality issues. The best analysts have strong people skills and patience to persevere through the data slop, which some leaders say is harder to find than the technology experience. Finally, analytics expertise is expensive, and leaders need to balance quick wins with deep dives into projects that serve as a team learning experience. Without planning to show a return on effort, c-suite support for analytics can dry up quickly.

In order to understand how to develop staff to achieve future capabilities, a competency model is required that maps analytic functions, skills, and competencies to specific organizational roles (Figure 2).   Once the gap between current and future competencies are identified on both the organizational and individual levels, learning plans can be developed tailored to specific needs.
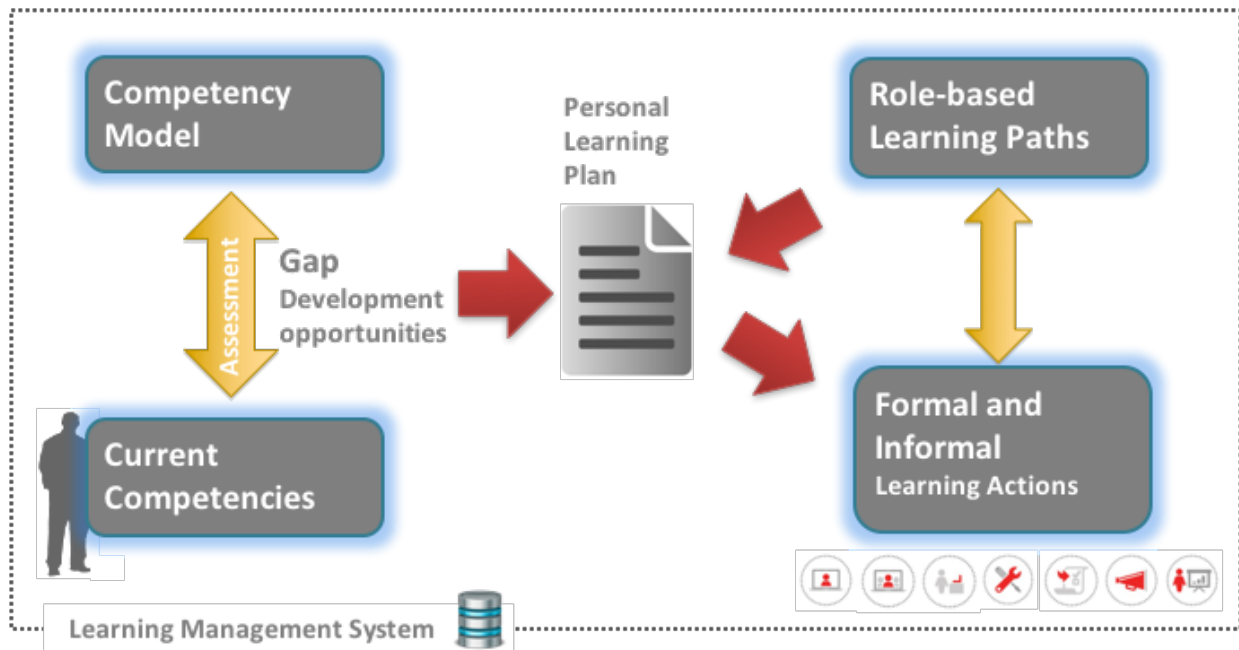
Figure 2: Competency-Based Learning

In defining this gap, there are a number of options available in the literature but none unique to healthcare. We therefore developed our ThotWave Healthcare Analytics Competency Model® to help teams understand their current capabilities, potential future states, as well as the gaps to bridge in order to get there. It has been mapped to two well-known external standards/ certification bodies including INFORMs Certified Analytics Professional and AHIMA's Certified Health Data Analyst (CDHA®). Developed through a process of workplace analysis and expert knowledge, our model includes nine domains of knowledge, skills, and behaviors that need to be demonstrated within a healthcare analytics team (Figure 2). It is noteworthy that many of the competencies we recognize as being critical for analytics have a number of non-statistical and non-technical features. This is because our model seeks to address the entire analytics lifecycle. We utilize diagnostic, formative, summative, and evaluative assessments aligned to our competency model to map an individual's initial mindset, skillset, and toolset and then monitor the growth.

Figure 2: Nine Competency Domains of Health Analytic Knowledge

C<small>OMPETENCIES FOR</small> J<small>OB</small> F<small>AMILIES</small>

Figure shows the relative composite skill level required for each job family outlined in the Introduction according to our competency model. In looking at the relatedness of the different plotted competencies for each job family, one could imagine moving between career paths based on the shared skills between different ladder branches. The value of this for organizations is that we can use this model to specify the desired roles in the ideal future state and then craft a plan for both how staff can expect to grow into those roles as well as evolve beyond them.
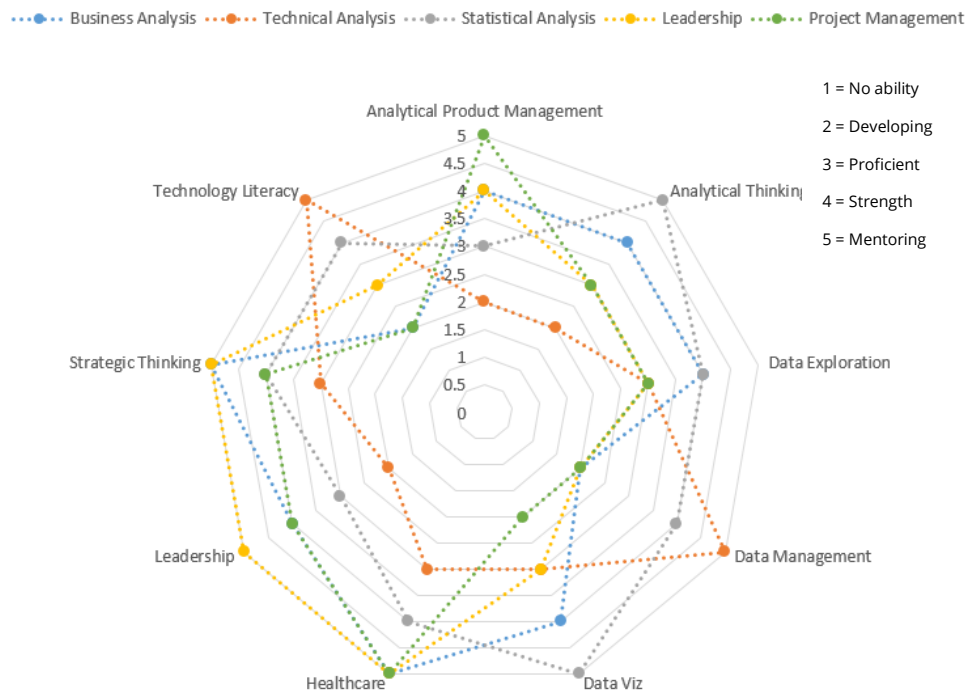

Figure 3: Healthcare analytics competencies for analytic job families

Developing Analytic Competencies

Analytic Talent Management

Talent management is a critical success factor for modern healthcare organizations. An estimated 31% of employees in healthcare voluntarily leave their position in the first 12 months[1] and an it can take many organizations up to 18 months to become proficient in an organization's unique technology and data environment. Consequently, we must face the fact that classical training approaches are no longer adequate to continuously develop a data-enabled workforce.

---

[1]   Source: https://www.talx.com/benchmarks/turnover/

Personal communication and http://www.medpagetoday.com/practicemanagement/practicemanagement/54541

A well-nurtured analytics team is one where staff feel a sense of mastery, autonomy, and purpose. Mastery implies that one works at the top of their license while their skills are continually grown. Professional development plans are roadmaps for this process, but roughly 88% of employees don't make time for learning and 64% of managers don't enable or follow-up on learning and development plans (Zappa, 2016). Part of the problem is that training too often equates to conference attendance or technical workshops.

Autonomy is the second piece of staff nurturing and hinges on role clarity and accountability, which are known drivers of organizational health. Leading teams usually assign analysts specific service lines to cover and act as diplomats between business users and those involved in the technical execution. Providence Health leadership takes this approach and trusts staff to deliver without the need for micromanagement while enforcing hard deadlines.

The third and most important component of well-nurtured teams is that employees feel a strong sense of purpose in what they do. Poor management is a lethal destroyer of meaning and an attrition accelerant. Researchers note that meaningfulness is more important to staff than other aspects of work, including pay, promotions, or working conditions. One simple way for leaders to nurture meaning is always to thank them personally for their specific effort they bring to the table.

To retain talent, optimize the use of information and resources, hasten the translation of research insights into bedside innovation, and ensure that people are operating at expected levels of competency, many Academic Medical Centers are creating initiatives to enhance the capabilities of key data and analytics stakeholders across the enterprise. Establishing clear competency goals, talent development plans, and appropriately-designed curriculum are critical to ensuring that strategic initiatives can be supported through a knowledgeable, well-prepared analytics community. An effective talent management strategy means that the organization has the ability to effectively acquire, develop, and retain high performing individuals to deliver the benefits of the analytics strategy. Effective talent management processes include ensuring that the organization has the right resources assigned to the critical roles to drive the change. Talent Management should take place once the Analytics Capabilities and Organizational Design are defined (previous section) and the gaps to getting there have been characterized.

Individual and Team Skills Assessment

Assessing the aggregate training needs of teams is best done using a gap analysis exercise. In such a scenario, we would compare the assessed competencies of teams to that of an ideal future state. Each staff member completes the electronic ThotWave Health Analytics Self-Assessment that evaluates both their interest and capabilities in the nine core health analytic competencies. This assessment asks an individual about their perceived strengths in executing a number of activities such as querying relational database systems or designing a quality metric to support an internal customer. Results are benchmarked against the ideal score expected for a given organizational role. This allows us to define a learning and development plan that addresses individuals' unique competency needs. Advantages of this approach include:

- Provides executive leadership with substantial detail about the capabilities of different analytic teams

with resolution to the individual staff member level.

- Uses a comparative benchmark that can be specific to the perceived competency imperatives of the assessed organization.

- Highlights deficiencies in knowledge among and within teams, which can be used to tailor both messaging and content surrounding analytics services.

- Supports consistency and accuracy of job classifications for staff.

- Clarifies the competencies needed for career progression as well as the paths that should be taken to develop talent.

- Provides information regarding what may be the best use cases for instructor-led training.

- Ensures a foundation in the tools and methodologies to achieve competency analysis moving forward as staff grows and changes.

While we have traditionally used this assessment when working with organizations, our methodology is available to individuals in health analytics who may not currently be on the career path they desire. To support these individuals, we have developed the ThotWave Healthcare Analytics Interest Inventory which gives each individual the opportunity to reflect on their personal experiences and expertise within each of the 9 Healthcare Analytics Knowledge Domains. They also can complete the Skills Assessment and receive gap reports to understand strengths and opportunities for improvement based on the stated career development goals

Competency-Based Curriculum for Analytics

A professional development plan is a bit of a mythical creature in that most individuals know what it is but have not seen one in action. Even if staff have contributed to them as part of seasonal reviews, it is rare that their recommendations are followed. Much of this stems from healthcare's current reactive nature. Few from leadership down to the rank-and-file feel that they have the time to craft a development plan much less devote the time to executing it. But a solid training plan is essential to avoid the ambiguous 'go learn' directives and ensure that development receives protected time in an employee's work week. As we help organizations improve their overall analytics maturity, one of the key elements of achieving this vision is to execute an effective training process.

So one thing to remember in developing a curriculum for training your data champions is that it's a journey, not just a few courses. Most of the time, people think of training as these steps in blue in Figure 4 where it is delivered. But to have the education really have the desired impact, the prepare and sustain phases are essential. The preparation phase includes the skills assessment process mentioned earlier as well as a variety of change management-facing activities that prepare staff for this shift in professional develop. The sustain phase provides ongoing reinforcement with a variety of short educational activities built into professional development plans. This could include the creation of lunch and learn groups; tool-based users groups; or

providing access to curated content subscriptions to help staff stay up to date in topics around the competencies you are emphasizing.

Prepare
Deliver
Sustain

Change Management Support · Technology Demos · Skills Assessment · Core Competency Definition · Training Modality Selection · Team Meetings

Town Halls · eLearning · Practicum Projects · Instructor-Led Training · Hackathons · Hands-On Workshops · Lunch-and Learn Groups

Peer Mentoring · Writing & Blogging · Curated Content · Tool-Based User Groups · 'Ask the Expert' Office Hours · Performance Assessments
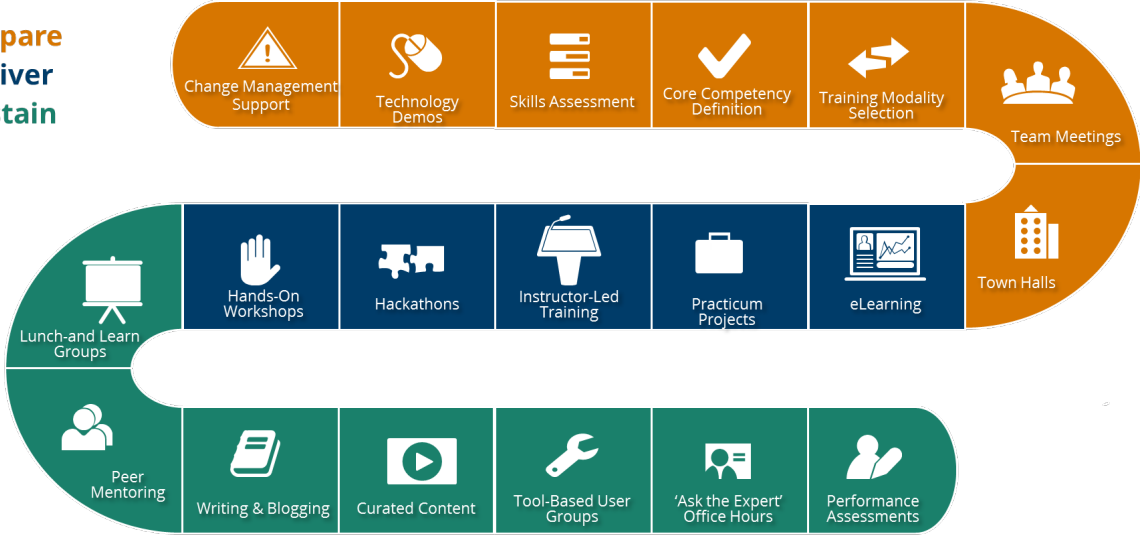
Figure 4: Three phases of the training journey

When picking the specific curriculum to deliver to an analytics organization, you want to ensure that you have a variety of training that appeals to the different types of learning domains. Learning objectives, which map to competencies, can be classified according to Bloom's Taxonomy which reflects the learning being absorbed in either a cognitive, affective, or sensory manner. Figure 5 shows a learning path provided to a healthcare organization seeking to redesign their internal approach to analytics training.

		Business Analysis	Analytical product Management	Statistical Analysis	Leadership	Technical Analysis
Awareness and Understanding	Security and Privacy	●	●	●		●
	Operational Workflows And Healthcare Provider And Care Team Roles	●				
	Data Governance For Analytics		●		●	
	Analytics Life Cycle		●			
	Platforms and Infrastructure for Analytics					●
Skill Development	Applying EHR Workflow To Analytical Products	●	●		●	
	Healthcare Data Visualization With <Tool Specific>	●		●		
	Essentials Of Question Design And Data Collection	●		●		
	Stats 101	●				
	Advanced Topics In Statistical Analysis			●		
	Agile Project Management		●			
	Enterprise Portfolio Management		●			
	Healthcare Data Analytics With <Tool>			●		
	Exercises In Profiling And Characterizing Data					●
Skill Integration	Data Visualization Best Practices	●		●		
	Storytelling	●		●		
	Data Exploration	●		●		
	Operationalizing Analytic Insights		●			●
	Combining And Integrating Data Sources					●
	Applying Systems Thinking To Analytics		●			
	Data Modeling For Analytics					●
	Data Wrangling					●
	Exercises In Data Integration And Data Quality					●
	Exercises in Tool Agility					●
Leadership and Mentoring	Managing Analytics Teams And Processes		●		●	
	Leading Analytics Change: Exercises In Collaboration And Influence		●		●	
	Data Strategy And Data Governance				●	
	Mentorship Leading And Evangelizing				●	
	Strategic Alignment Of Healthcare Analytics				●	
	Exercises In Coaching And Developing Analytic Staff				●	

Figure 5: A curriculum map for the 5 healthcare analytic job families

Once the final performance objectives and behaviors are agreed upon, it is important to incorporate/update the job descriptions with the future performance objectives and behaviors. This will allow the organization's Human Resources to adjust job description details that may impact recruiting, compensation, and/or benefit criteria. This also will enable future employees who hold the job to have a clear understanding of the objectives and behaviors required to support the business.

Career Progression

The reality of the modern workplace is that individuals grow skills from other job ladders lateral to their own. Healthcare margins are tight and budgets thin, meaning positions aren't always backfilled when attrition occurs. Existing staff learn new skills and broaden their capabilities to keep the business running. *Unfortunately, traditional career paths in healthcare have not awarded this individual agility.* Existing job titles can fail to reflect what an individual actually does during their day. Other industries are eschewing the career ladder in favor of the 'career lattice', which is a diagonal framework that relates lateral experiences, skill acquisition, and peer networking to show employees how to move to a variety of positions. In applying the competency model to the career lattice, we can show staff the gap between current and aspirational skills for a desired position (Figure 6). While many of these roles may exist in another form within existing human resources classification, analytics leadership will want to consider how their families interrelate into a career lattice in making hiring, talent

retention, and professional development decisions. Where the interrelation is not there, our model is intended as a starting place as they rethink their organizational design.
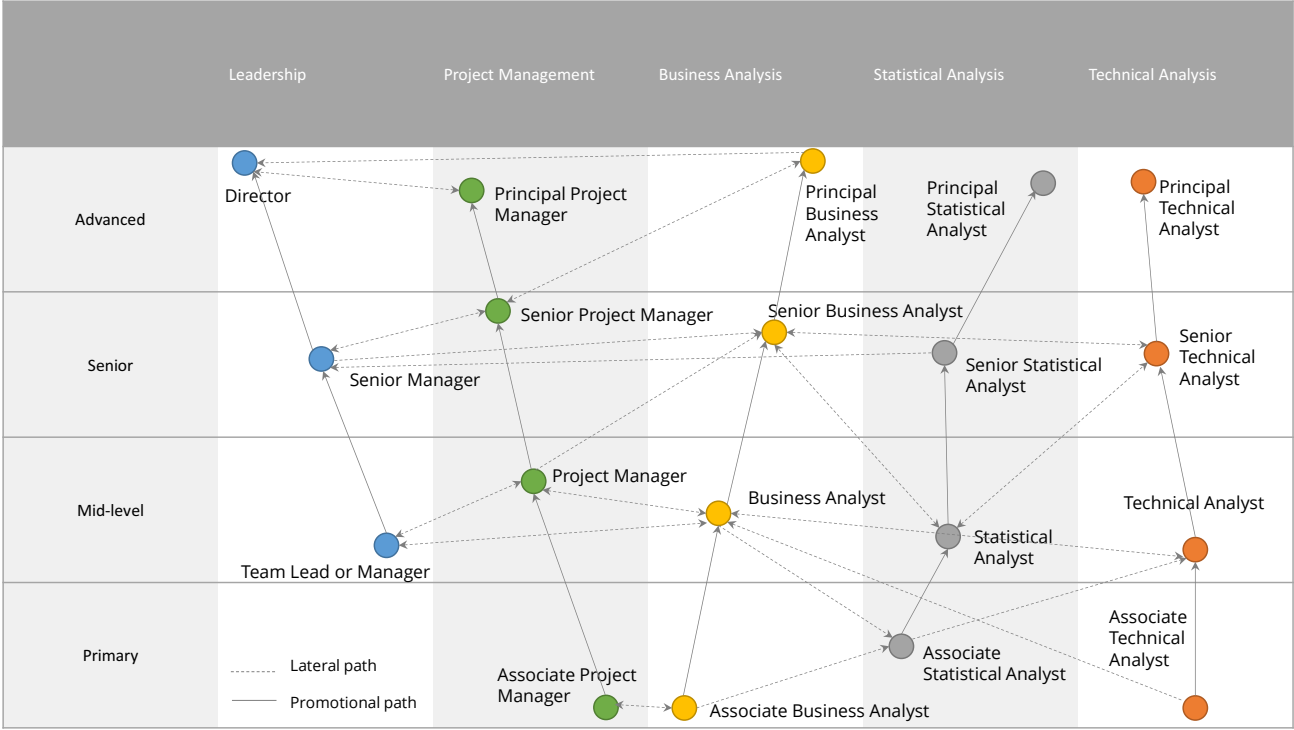

Figure 6: Healthcare analytics competencies for analytic job families

Summary

We developed our ThotWave Healthcare Analytics Competency Model® to help teams understand their current capabilities, potential future states, as well as the gaps to bridge in order to get there. Our model is specifically designed to support success in healthcare analytics. Individual competencies and learning paths are aligned to the knowledge, skills, and abilities needed to achieve organizational capabilities. The assessment documents the employee knowledge and skills with a view to understanding the gap between current and future capabilities. By focusing on individual competencies, performance plans can be built for each individual. Our perspective is that meaningful educational opportunities will nurture the growth and active participation in analytics in healthcare and position an organization to increase its overall analytic maturity.

Biography

Greg Nelson, President and CEO, ThotWave Technologies, LLC.

Greg is a global healthcare and Business Intelligence (B.I.) executive with over two decades of experience and leadership in the field. Greg is a prolific writer and speaker interested in healthcare analytics and the strategic use of information technology.

He received his BA in Psychology from the University of California at Santa Cruz and advanced his studies toward a PhD in Social Psychology and Quantitative Methods at the University of Georgia. Recently, Greg completed his Masters degree from Duke University in Clinical Informatics from the Fuqua School of Business. His academic and professional interests include helping organizations mature their analytic capabilities. Founder, President, and CEO of ThotWave Technologies, a niche consultancy specializing in healthcare analytics, Greg is particularly interested in how lessons from across other industries can be applied to help solve the challenges in healthcare.

With certifications in Healthcare IT, Project Management, Six Sigma and Balanced Scorecard, Greg is also a prolific writer and has presented over 200 professional and academic papers in the United States and Europe. He won the Best Paper Award in 2013 at the Pharmaceutical SAS Users Group Conference and sits on the board of the SAS Global Users Group. In 2011, Greg was selected by SAS into their loyalty partner group. "This program acknowledges and supports individuals who are recognized experts in their fields and have a long-term relationship with SAS."

Married to wife Susan and living on a small "farmlet" in rural North Carolina, Greg is an avid woodworker, enjoys photography, rides a Harley-Davidson Motorcycle, and strives to be a lifelong learner.

Contact information

Your comments and questions are valued and encouraged. Contact the authors at:

Greg Nelson   greg@thotwave.com

ThotWave Technologies, LLC

1289 Fordham Boulevard #241

Chapel Hill, NC 27514 (800) 584 2819

http://www.thotwave.com

 SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

thinking data® is registered trademark of ThotWave Technologies, LLC.

Other brand and product names are trademarks of their respective companies.

References

Zappa, J. a. S., Sean. (2016). A Holistic View of Learning: Aligning Talent Development, Learning & Development, and Employee Education Benefits. Retrieved from http://www.trainingindustry.com/webinars/a-holistic-view-of-learning-aligning-talent-development-learning-and-development-and-employee-education-benefits.aspx