

Business Industry Leadership Team (BILT)

Data Analytics KSAs November 2024

This material is based upon work supported by the National Science Foundation under Grant No. 2300188. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

This prioritized Knowledge, Skills, and Abilities (KSA) list was developed by the National IT Innovation Center (NITIC) in collaboration with IT industry subject matter experts (SMEs). This list is intended to help faculty and administrators align curriculum with industry needs so graduates are "workforce ready" when they graduate.

SMEs convened online to rank each KSA item one by one – a ranking of "4" meant that item was essential for entry-level IT workers, while a "1" meant that item could be removed from program curriculum. By default, items with an average vote of 2.6 or lower were turned pink to signal that this item may not be worth keeping.

After the vote, the SMEs discussed the results as a group. This discussion led to some revisions in the KSA descriptions and rankings. Some "pink" items the SMEs together deemed important despite the lower average and were kept, while some "green" items with a higher average the SMEs recommended be removed from the list. Those items can be found at the bottom of each KSA section under "Recommended Removal."

The "To Be Included in the Next Vote" section at the bottom includes items the SMEs suggested that were not included in the original list vote.

This KSA prioritization process is a cornerstone of the successful Business and Industry Leadership Team (BILT) model which puts businesses in a co-leadership role.

Learn more about how the BILT works – and how you can implement it with your own program – by using these resources:

NITIC orientation video; <u>www.tiny.cc/BILTorient</u> Convergence Technology Center BILT handbook: <u>www.tiny.cc/implementingBILT</u>

This content has been built on the work of the National Convergence Technology Center (DUE 1700530), the Pathways to Innovation grant (DUE 2039395, and the IT Skill Standards grant (DUE 1838535).

Data	Analytics Knowledge & Skills - updated November 2024	ed November 2024 # votes (4 = most important)			green cell > 2.60	
		4	3	2	1	Avg
Knowledge Knowledge focuses on the understanding of concepts. It is theoretical and not practical. An individual may have an understanding of a topic or tool or some textbook knowledge of it but have no experience applying it. For example, someone might have read hundreds of articles on health and nutrition, many of them in scientific journals, but that doesn't make that person qualified to dispense advice on nutrition.						
K-1	Knowledge of risk management processes as part of Software Development Life Cvcle.	3	6	5	2	2.63
K-2	Knowledge of computer algorithms.	3	8	3	2	2.75
K-3	Knowledge of computer programming principles.	5	8	3	0	3.13
K-4	Knowledge of data administration and data standardization policies.	6	8	1	1	3.19
K-5	Knowledge of data mining and data management principles.	6	9	1	0	3.31
K-6	relationships, and views.	10	6	0	0	3.63
K-7	Knowledge of a broad and just-in-time understanding of mathematical concepts (e.g., logarithms, trigonometry, linear algebra, calculus, statistics, and operational analysis).	4	3	8	1	2.63
K-8	Knowledge of programming language structures and logic.	6	8	2	0	3.25
K-9	Knowledge of query languages such as SQL (structured query language) and NOSQL.	10	4	2	0	3.50
K-10	Knowledge of sources, characteristics, and data assets.	5	6	5	0	3.00
K-11	Knowledge of the various technologies for organizing and managing information (e.g., databases, bookmarking engines).	4	11	1	0	3.19
K-14	Knowledge of how to utilize current popular frameworks and languages.	4	8	4	0	3.00
K-15	Awareness of machine learning and Al.	10	5	1	0	3.56
K-16	Knowledge of Personally Identifiable Information (PII) data security standards and how to mask the data.	8	4	4	0	3.25
K-17	Knowledge of the principal methods, procedures, and techniques of gathering information and producing, reporting, and sharing information.	8	7	1	0	3.44
K-18	Knowledge of data mining techniques.	7	5	4	0	3.19
K-19	Knowledge of relational database framework and why relation databases are used.	9	5	2	0	3.44
K-20	Knowledge of how to extract, analyze, and use metadata.	10	5	1	0	3.56
K-21	Awareness of a variety of machine learning techniques (clustering, decision tree learning, artificial neural networks, etc.) and their real-world advantages/drawbacks.	7	6	2	0	3.33
K-22	Knowledge of advanced statistical techniques and concepts and when to apply them.	5	6	5	0	3.00
K-23	Knowledge of the underlying theory and concepts of Relational Databases (e.g., Microsoft SQL Server, Oracle, Teradata MySQL)	6	6	3	1	3.06
K-24	Knowledge of data analysis concepts.	9	6	1	0	3.50
K-25	Knowledge of how to identify and document potential ethical concerns for application of model outputs	6	6	4	0	3.13
K-26	Knowledge in implementing/developing basic data modules using existing tools.	4	8	4	0	3.00
K-27	Knowledge in Regression Analysis (e.g., Hierarchical Stepwise, Generalized Linear Model, Ordinary Least Squares, Tree-Based Methods, Logistic)	5	6	5	0	3.00
K-28	Knowledge of refining the problem statement and delineate.	3	11	1	1	3.00
K-29	Knowledge of tuning models and data.	4	8	4	0	3.00
K-30	Knowledge of how to clearly articulate information requirements into well-formulated research questions and data tracking variables for inquiry tracking purposes.	6	5	5	0	3.06
K-31	Knowledge of ethics as it applies to data analytics and how to apply ethical iudgment when policies are not well-defined.	5	7	4	0	3.06
K-32	Knowledge in determining if a difference in values is significant (statistical and common sense) or not.	3	9	4	0	2.94
K-33	Knowledge of the landscape of BI tools (e.g., Power BI, Google, Tableau) and data preparation tools and understanding of the data platform associated with each.	8	6	2	0	3.38
K-34	Knowledge of identifying basic common coding flaws at a high level.	1	8	7	0	2.63
	Recommended Removal - Knowledge					
K-12	Knowledge of command-line tools (e.g., mkdir, mv, ls, passwd, grep, PowerScript).	3	2	7	4	2.25

Data Analytics Knowledge & Skills - updated November 2024		# vo most	green cell > 2.60		
	4	3	2	1	Avg
K-13 Knowledge of interpreted and compiled computer languages.	0	9	5	2	2.44
Skills The capabilities or proficiencies developed through training or hands-on experience. Skills are the practical application of theoretical knowledge. Someone can take a course on investing in financial futures, and therefore has knowledge of it. But getting experience in trading these instruments adds skills.					
S-1 Skill in conducting queries and developing algorithms to analyze data structures.	8	5	3	0	3.31
S-2 Skill in creating and utilizing mathematical or statistical models.	3	8	5	0	2.88
S-3 Skill in data mining techniques (e.g., searching file systems) and analysis.	5	9	2	0	3.19
S-4 Skill in using and contributing content to data dictionaries and documentation.	3	10	3	0	3.00
5-5 Skill in generating queries and reports.	11	5	0	0	3.09
S-6 R).	8	5	3	0	3.31
S-7 Skill in data pre-processing (e.g., imputation, dimensionality reduction, normalization, transformation, extraction, filtering, smoothing).	8	4	4	0	3.25
S-8 Skill in identifying patterns or relationships.	8	7	1	0	3.44
S-9 Skill in performing sentiment analysis.	2	8	5	0	2.80
S-10 Skill in using basic descriptive statistics and techniques (e.g., normality, model	5	9	1	0	3.27
S-11 Skill in using data analysis tools (e.g., Excel, Python)	12	4	0	0	3.75
S-12 Skill in using data mapping tools.	5	9	2	0	3.19
S-13 Skill in using outlier identification and removal techniques.	3	9	4	0	2.94
S-14 Skill in writing scripts in contemporary/popular languages.	5	9	2	0	3.19
S-15 Skill to identify sources, characteristics, and uses of the data assets.	2	9	4	0	2.87
S-16 Skill in conducting information searches.	3	9	4	0	2.94
S-17 Sill in developing or recommending analytic approaches or solutions to problems and situations for which information is incomplete or for which no precedent exists.	2	8	6	0	2.75
S-18 Skill in evaluating information for reliability validity and relevance.	7	5	4	0	3.19
S-19 Skill in preparing briefings and creating presentations.	6	2	8	0	2.88
S-20 Skill in tailoring analysis to the necessary levels (e.g., classification and	6	7	2	1	3.13
S-21 Skill in validating the data that comes up after using search engines (e.g., Google, Yahoo, LexisNexis, DataStar).	5	7	4	0	3.06
S-22 Skill in utilizing feedback to improve processes, products, and services.	6	6	4	0	3.13
S-23 Skill in performing data analysis including applying statistics.	7	9	0	0	3.44
S-24 Skill in using statistical / other popular computer languages and frameworks to	4	9	3	0	3.06
S-25 Skill in Visualization using R. Python, or other languages and frameworks.	7	6	3	0	3.25
S-26 Skill in problem-solving skills and critical thinking ability.	12	4	0	0	3.75
S-27 Skill in collaboration and communication skills within and across teams.	6	10	0	0	3.38
S-28 Skill in analytics problem framing (e.g., define geometric sets).	4	8	4	0	3.00
S-29 Skill in dissecting a problem and examining the interrelationships between data that	7	5	4	0	3.19
S-31 Skill in using popular/contemporary data visualization tools.	8	5	3	0	3.31
S-32 Skill in sourcing data used in information, assessment, and/or planning products.	4	6	6	0	2.88
S-33 Skill in evaluating information for reliability, validity, and relevance.	5	7	4	0	3.06
S-34 Skill in focus research efforts to meet the customer's decision-making needs.	4	4	5	2	2.67
S-35 Skill in identifying information gaps.	2	11	3	0	2.94
S-36 Skill in recognizing and assisting in cognitive biases which may affect analysis.	3	7	6	0	2.81
S-37 Skill in recognizing and assisting in deception in reporting and analysis.	3	7	5	1	2.75
S-38 Skill in understanding objectives and effects.	3	10	3	0	3.00
S-39 Skill in understanding and using the databases and tools to run queries to solve the business problem.	7	8	1	0	3.38
S-40 Skill in identifying patterns.	6	8	2	0	3.25
S-41 Skill in presenting and telling the story with data.	9	5	2	0	3.44
S-42 Skill in driving contidence in numbers they are presenting by indicating probabilities of the numbers being accurate.	5	7	4	0	3.06
To Be Included in Next Vote					
Knowledge of prompt opgingering/Conserting Al					
Knowledge of Agentic Al					

Data Analytics Knowledge & Skills - updated November 2024		# votes (4 = most important)			
	4	3	2	1	Avg
Knowledge of data platforms, (e.g., Databricks/Snowflake).					
Knowledge of data stacks (Google, Azure, AWS).					
Knowledge of the different data architectures (e.g., database, datamart, data lake,					
data warehouses=).					