

## Specialist Digital Tech Credential: Data Analytics

### 1. Data Management and Governance

- a. Awareness of definitions of key terms of data analysis-**knowledge** - Level 1
- b. Explain the important issues around data governance, including change management purpose and structure - **knowledge** - Level 2
- c. Understand data profiling; explain data summarization, sanity checks, and validation - **knowledge** - Level 2
- d. Differentiate common data typologies, including structured vs. unstructured, numeric vs. text, root vs. derived-**knowledge** - Level 2
- e. Explain the difference between a logical and physical data diagram - **knowledge** - Level 2
- f. Recognize and explain the importance of Key Performance Indicators and Metrics and how to describe and communicate data for KPIs - **ability** - Level 2
- g. Recognize the basic data governance issues from data life cycles- **ability** - Level 3

### 2. Data Storage

- a. Identify the main concepts, key technologies, strengths, and architectures of analytic solutions - **knowledge** - Level 1
- b. Understand how your data is stored and how that impacts data analytics and performance - **knowledge** - Level 2
- c. Explain the pros and cons of on-premises vs cloud-based analytics solutions - **knowledge** - Level 2
- d. Describe the implications of data architecture on data processing- **knowledge** - Level 2
- e. Explain the different data organizational schemas (e.g., 1st normal form vs. 3rd normal form vs. Star Schema, etc.) and how they relate to database management systems - **knowledge** - Level 2
- f. Explain the differences between batch analytics and streaming analytics, where each would be applied, and the constraints of each - **knowledge** - Level 3
- g. Properly use and develop a data dictionary for publicly shared data - **skill** - Level 3

### 3. Data Quality

- a. Identify steps to ensure data quality at each level - **knowledge** - Level 2
- b. Explain strategies for dealing with data quality issues - **knowledge** - Level 2
- c. Understand how to curate data, how data will be used, and ensure that data is fit for purpose and use-**skill** - Level 3
- d. Implement common information retrieval and filtering applications in databases and data systems, specifically in SQL, R, or another common language – **skill** - Level 3
- e. Organize and integrate relevant data from various sources – **skill** - Level 3

- f. Identify data quality issues (e.g., missing values, duplication of records, outliers) - **skill** - Level 4
- g. Perform data summarization, sanity checks, and validation using appropriate tools and software - **skill** - Level 4
- h. Assess data in terms of reliability and appropriateness to the possible solutions – **ability** - Level 6

#### 4. Data Manipulation

- a. Explain the purpose and benefits of an API- **knowledge** - Level 2
- b. Explain potential uses and applications of data manipulation tools given a source and type of data - **skill** - Level 2
- c. Perform basic data manipulation using appropriate tools and software - **skill** - Level 3
- d. Perform data summarization, sanity checks, and validation using appropriate tools and software - **skill** - Level 3
- e. Create and edit simple data structures and storage - **ability** - Level 3

#### 5. Probability and Descriptive and Inferential Statistics

- a. Demonstrate **knowledge** of probability and standard statistical distributions - **knowledge** - Level 2
- b. Explain core statistical inference concepts (for example, deriving relevant hypotheses, evaluating the hypotheses, and prediction with uncertainty) - **knowledge** - Level 2
- c. Differentiate among data analytic approaches (e.g., descriptive vs. diagnostic vs. predictive vs. prescriptive analytics) - **knowledge** - Level 2
- d. Demonstrate and explain the role and importance of model validation and accuracy metrics in analytics projects, hypothesis testing, and information retrieval- **knowledge** - Level 2
- e. Explain core probability concepts (e.g., random variables, key distributions, conditional probability, Bayes theorem) - **knowledge** - Level 2
- f. Explain sampling methods (for example, stratified sampling, simple random sampling, and cluster sampling) - **knowledge** - Level 2
- g. Articulate the limits of statistical inference and statistical measurement - **knowledge** - Level 2
- h. Interpret results from analysis in the context of the original problem - **skill** - Level 3
- i. Choose appropriate mathematical methods and apply towards data analysis - **skill** - Level 3
- j. Leverage the appropriate sampling methodologies - **skill** - Level 4
- k. Explain how to ask the right questions that lead to actionable analytics - **ability** - Level 5
- l. Draw insights from results of analysis in the context of the original problem - **ability** - Level 6

#### 6. Data Visualization and Communication

- a. Explain the role of data visualization in discovery, communication, and decision-making - **knowledge** - Level 2
- b. Explain key tools and technologies for creating visualizations - **knowledge** - Level 2

- c. Evaluate data visualization options for proper application in various situations - **skill** - Level 3
- d. Visualize data using various types of displays including tables, dashboards, graphs, maps, and trees- **ability** - Level 5
- e. Create effective static and interactive data visualizations or narratives that employ analytics and visualization software and strategies for various audiences- **ability** - Level 6
- f. Properly define a problem in context, use appropriate data, and deliver a compelling visualization to explain or answer a question- **ability** - Level 6

## 7. Foundations of Systems Development Life Cycle

- a. Describe key activities in systems development and the role of modeling- **knowledge** - Level 2
- b. Explain the purpose of using SDLC on certain projects- **knowledge** - Level 2
- c. Identify and describe several SDLC models (e.g., waterfall, Agile) - **knowledge** - Level 2
- d. Explain the analytics lifecycle and the CRISP-DM model - **knowledge** - Level 2

## 8. Data Ethics

- a. Identify how global legal, policy and/or ethical constraints might impact data analyses-**knowledge** - Level 2
- b. Identify the established ethical and legal issues in data management facing organizations-**knowledge** - Level 2
- c. Explain how ethical, compliance, and legal issues should/must be considered in data driven decision making- **knowledge** - Level 2
- d. Discuss the importance of provenance, transparency, and explainability in data analysis and the ability to build trust - **knowledge** - Level 2
- e. Present real world examples of data bias and the unintended consequences of using analytics, machine learning, and AI in making decisions - **knowledge** - Level 2
- f. Recognize the importance that data protection plays in managing data and building trust - **knowledge** - Level 2
- g. Demonstrate an understanding of the way your ML algorithms are vulnerable and can be manipulated - **knowledge** - Level 2
- h. Understand the limitations of using data analytics tools - **knowledge** - Level 2
- i. Demonstrate awareness of personal privacy issues related to the collection and usage of data-**knowledge** - Level 3
- j. Explain the limitations and potential unintended effects of data analysis when such algorithms encounter new scenarios- **knowledge** - Level 3

**BLOOM's TAXONOMY:**

**Level 1** - Recall /regurgitate facts without understanding. Exhibits previously learned material by recalling facts, terms, basic concepts and answers.

**Level 2** - To show understanding finding in-formation from the text. Demonstrating basic understanding of facts and ideas

**Level 3** -To use in a new situation. Solving problems by applying acquired knowledge, facts, techniques and rules in a different way

**Level 4** - To examine in detail. Examining and breaking information into parts by identifying motives or causes; making inferences and finding evidence to support generalisations

**Level 5** - To change or create into something new. Compiling information together in a different way by combining elements in a new pattern or proposing alternative solutions

**Level 6** -To justify. Presenting and defending opinions by making judgements about information, validity of ideas or quality of work based on a set of criteria

---

**KSA DEFINITIONS:**

**Knowledge:** subjects, topics, and items of information that an employee should know at the time he or she is hired or moved into the job

**Skills:** technical or manual proficiencies which are usually learned or acquired through training; measurable and observable

**Abilities:** demonstrate capacity to apply knowledge and skills simultaneously in order to complete a task or perform an observable behavior, or help a person do a job