CERTIFIED ANALYTICS PROFESSIONAL

THE CAP-EXPERT (CAP-X) EXAM BLUEPRINT EXPLAINED

PRACTICE ANALYSIS 2024

This document represents a delineation of common or typical tasks performed by advanced Analytics Professionals across the complete analytics life cycle. The Blueprint defines the current knowledge, skills, and abilities (KSAs) that must be demonstrated by analytics professionals to conduct effective and successful analytics projects. INFORMS upholds stringent guidelines for the construction of our certification exams. A panel of subject matter experts (SMEs) with diverse analytics backgrounds collaborated to develop the lists of tasks performed by advanced-level analysts on the job, tasks that lead to analytics project success. The KSAs were validated by a larger audience of experts and rated for their frequency of use and importance to analytics project success. This "blueprint" outlines the domain knowledge tested on the INFORMS CAP®-Expert (CAP-X) examination.

We recognize that in the course of analytics work, these tasks are often performed multiple times as part of ongoing feedback loops, with modifications made with each iteration based on data, findings, and results. For clarity and simplicity, most of the feedback loops are not presented in this document. It is assumed and understood that they are a routine part of practice.

VALIDATION

The findings of this working group were then validated by a sample of practicing analytics professionals. Feedback from this survey resulted in slight modifications of the performance domains, tasks, and knowledge that comprise the test blueprint that determines the content of the CAP examination. The table below includes the final domains and their representation on the certification exam that were derived from the JTA and a review of validation survey recommendations.

EXAM BLUEPRINT

The exam covers tasks performed across the entire analytics life cycle. Tasks are categorized into the following seven primary domains of knowledge.

Domain	Examination Weight[1]
I: BUSINESS PROBLEM (QUESTION) FRAMING	16%
II: ANALYTICS PROBLEM FRAMING	17%
III: DATA	21%
IV: METHODOLOGY (APPROACH) SELECTION	13%
V: ANALYTICS / MODEL DEVELOPMENT	16%
VI: DEPLOYMENT	10%
VII: ANALYTICS SOLUTION LIFECYCLE MANAGEMENT	7%

[1] The domain weights shown are based on expert assessments of the importance of tasks and the frequency of their performance. Those weights influence the exam blueprint or the percentage of the exam devoted to each domain.

DOMAIN I: BUSINESS PROBLEM (QUESTION) FRAMING

The business problem (question) framing domain includes activities related to understanding the business problem and evaluating the scope of the problem.

Task 1.1: Develop an initial statement of a business problem (question).

- CAP-X 1.1.1: Identify and explain questions for obtaining information needed for creating a clear and complete business problem (question) statement.

Task 1.2: Identify all stakeholders and their perspectives.

- CAP-X 1.2.1: Identify and explain an effective method of conducting stakeholder analysis.
- CAP-X 1.2.2: Apply responsible, accountable, consulted, and informed criteria to a set of stakeholders.

Task 1.3: Determine if the business problem (question) is amenable to an analytics solution.

- CAP-X 1.3.1: Identify and explain changes to a problem statement and the reasons that the modified statement would make the problem amenable to an analytic solution.
- CAP-X 1.3.2: Identify and explain an issue that would likely have an impact on the feasibility of an analytics solution.

Task 1.4: Refine the initial business problem (question) statement until it is clear and concise.

- CAP-X 1.4.1: Identify revisions necessary to clarify and correct a business problem statement.

Task 1.5: Create an initial business case for a solution to the business problem (question).

- CAP-X 1.5.1: Identify the difference between types of business, societal, and analytical benefits.
- CAP-X 1.5.2: Identify the different types of business costs.
- CAP-X 1.5.3: Identify methods for determining the current state and operating procedures of business processes that are relevant to the business problem.
- CAP-X 1.5.4: Identify how to measure value in a potential solution.
- CAP-X 1.5.5: Identify the most effective approach of risk mitigation to address acceptable risks of implementing an analytics solution.
- CAP-X 1.5.6: Identify options for addressing unintended direct and indirect consequences of a potential analytics solution.

Task 1.6: Obtain sponsor agreement and stakeholder alignment on the business problem (question) statement.

- CAP-X 1.6.1: Identify how to respond to a concerned party who is either Responsible, Accountable, Supporting, Consulted, or Informed about the business problem (question) statement.

DOMAIN II: ANALYTICS PROBLEM FRAMING

The analytics problem framing domain includes activities related to understanding and framing the business problem (question) as an analytics problem.

Task 2.1: Reformulate the statement of the business problem (question) as an analytics problem statement.

- CAP-X 2.1.1: Identify what is missing from a business problem statement that would make it an analytics problem.
- CAP-X 2.1.2: Identify what information is necessary to translate a business problem statement into one or more analytics problem statement(s) and how the information should be obtained.
- CAP-X 2.1.3: Identify and explain the analytics problem statement that will result in meeting client requirements and achieving client success and project/program goals.

Task 2.2: Develop a proposed set of drivers/inputs and determine how they relate to outputs.

- CAP-X 2.2.1: Identify and explain the appropriateness of a proposed set of analytics inputs and outputs.
- CAP-X 2.2.2: Identify and explain the stated drivers or inputs and their relationship to outputs.

Task 2.3: State the set of assumptions related to the analytics problem.

- CAP-X 2.3.1: Identify explicit and implicit assumptions and the implications of those assumptions for appropriate analytical approach(es).
- CAP-X 2.3.2: Identify and explain the impacts if the assumptions for a project/program have not been met.
- CAP-X 2.3.3: Identify examples of the types of constraints that need to be addressed in an analytics problem statement.
- CAP-X 2.3.4: Identify limitations to an analytics problem statement with respect to a business problem.

Task 2.4: Define primary measures of success.

- CAP-X 2.4.1: Identify the tradeoff and relative importance of the primary measures of success as they relate to business and analytic problem statements.
- CAP-X 2.4.2: Identify and explain the process to verify if the primary measures of success have been met.

Task 2.5: Identify baseline performance of the current state.

- CAP-X 2.5.1: Identify the analytical problem statement that will have the most significant impact on the performance of the current state.

Task 2.6: Identify risks and mitigation strategies for an effective potential analytics solution implementation.

- CAP-X 2.6.1: Identify and explain the most effective approach for risk mitigation.
- CAP-X 2.6.2: Identify and explain the most effective approach for mitigating ethical risks.
- CAP-X 2.6.3: Identify the risks of using Generative AI as part of the analytics solution.

Task 2.7: Obtain sponsor agreement and stakeholder alignment on the Analytics Problem Framing.

- CAP-X 2.7.1: Identify how to gain alignment on the ananlytics problem statement from a party who is either Responsible, Accountable, Supporting, Consulted, or Informed.
- CAP-X 2.7.2: Evaluate the clarity and completeness in communicating the analytics problem statement including future concerns and sustainability for informed stakeholder alignment.

DOMAIN III: DATA

The data domain includes activities related to identifying what data is needed and available, manipulating data so as to render it usable to find the answer to an analytics problem, and the required documentation and reporting needs.

Task 3.1: Identify and prioritize data needs.

- CAP-X 3.1.1: Prioritize data needs, sources, and sequence for an analytics project statement.
- CAP-X 3.1.2: Identify the data requirements, limitations, and priorities when communicating data requirement to stakeholders and providers when conflicts exist.

Task 3.2: Identify and analyze data sources including data structures.

- CAP-X 3.2.1: Based on data attributes, data context, and metadata, identify appropriate courses of action; consider negotiating changes in business processes or different source systems.
- CAP-X 3.2.2: Based on an organization's data strategy, identify and explain how necessary roles should be defined and deconflicted.
- CAP-X 3.2.3: Identify data that is missing from a project and how it might be obtained.
- CAP-X 3.2.4: Identify how the analytics technology stack connects and how performance or other needs may shape architecture choices.
- CAP-X 3.2.5: Identify implications of data architectures for analytics projects/programs based on specific attributes.
- CAP-X 3.2.6: Identify advanced characteristics of a relational database.
- CAP-X 3.2.7: Identify the relevance of the 4 Vs (volume, veracity, variety, velocity) in a project.

Task 3.3: Create a data management plan.

- CAP-X 3.3.1: Identify how to work with stakeholders to develop a data strategy and data governance culture.

Task 3.4: Acquire data.

- CAP-X 3.4.1: Identify the trade offs and risks of different types of data sources and how to address potential issues.
- CAP-X 3.4.2: Identify methods for mitigating or managing data that contains potential risks and ethical implications.
- CAP-X 3.4.3: Identify appropriate methods for establishing lineage, traceability, and version control of data.

Task 3.5: Clean, harmonize, transform, merge/join, and validate data.

- CAP-X 3.5.1: Identify an appropriately designed solution to resolve conflicting needs from data in multiple sources.
- CAP-X 3.5.2: Identify problems related to cleaning and transforming data in a way to support future modeling.

Task 3.6: Assess data quality and identify relationships in the data.

- CAP-X 3.6.1: Identify how to correct issues with data that contains quality gaps including missing data, accuracy, completeness, consistency, timeliness, validity, uniqueness, and outliers.
- CAP-X 3.6.2: Identify patterns and characteristics of data and their implications when dealing with multivariate data and data profiling outputs.
- CAP-X 3.6.3: Identify and explain the clearest interpretation of a visual representation of data.
- CAP-X 3.6.4: Identify appropriate methods of evaluating data quality issues and how to mitigate them.

Task 3.7: Document and report data findings (e.g., data quality, impact analysis, results, data management plan.)

- CAP-X 3.7.1: Identify and explain valid conclusions about a data set's applicability for analytics.

Task 3.8: Validate and update the business and analytics problem statements.

- CAP-X 3.8.1: Identify appropriate changes to an analytic problem statement and the business implications of the changes based on observed data findings.

DOMAIN IV: METHODOLOGY (APPROACH) SELECTION

The methodology (approach) selection domain includes activities related to selecting the methods/techniques, software and tools that will enable analytics problem solutions.

Task 4.1: Determine available and appropriate methods/techniques for the identified problem.

- CAP-X 4.1.1: Understand and explain the business problem statement to identify the methods/techniques that should used.
- CAP-X 4.1.2: Understand and explain the analytics problem statement to identify which descriptive/diagnostic analytics methods/techniques should be used.
- CAP-X 4.1.3: Understand and explain the analytics problem statement to identify which predictive analytics methods/techniques should be used.
- CAP-X 4.1.4: Understand and explain the analytics problem statement to identify which prescriptive analytics methods/techniques should be used.
- CAP-X 4.1.5: Identify how Generative AI could be used to address an analytics problem statement and explain the benefits.

Task 4.2: Evaluate and select methods / techniques, given the resources and available data.

- CAP-X 4.2.1: Identify and explain technical costs, business costs, benefits, risks, and implications of choosing different methods/techniques when limitations on resources and data exist.

Task 4.3: Understand the technical solution architecture.

- CAP-X 4.3.1: Identify and explain an appropriate technical solution needed to employ in a solution set.

Task 4.4: Evaluate and select the technology stack.

- CAP-X 4.4.1: Identify strengths, weaknesses, and risks of the technology stack including databases, analytics software, networking, security, on-premise, cloud, open source vs. proprietary, platforms, etc.
- CAP-X 4.4.2: Identify alternatives to a spreadsheet analytics model and when they are appropriate.

DOMAIN V: ANALYTICS / MODEL DEVELOPMENT

The analytics/ model development domain includes activities related to identifying, using, calibrating, and integrating models as well as the documentation of the model and communication of findings.

Task 5.1: Design the model structure.

- CAP-X 5.1.1: Identify and explain appropriate inputs, outputs, and relationships for descriptive/ diagnostic analytics models.
- CAP-X 5.1.2: Identify and explain appropriate inputs, outputs, and relationships for predictive analytics models.
- CAP-X 5.1.3: Identify and explain appropriate inputs, outputs, and relationships for prescriptive analytics models.

Task 5.2: Build one or more appropriate models.

- CAP-X 5.2.1: Identify what does not belong or is missing from a predictive model and a set of model elements (dependent and independent variables).
- CAP-X 5.2.2: Identify what does not belong or is missing from a prescriptive model and a set of model elements (dependent and independent variables).
- CAP-X 5.2.3: Identify an error and how to rectify the error(s) for a predictive model.
- CAP-X 5.2.4: Identify an error and how to rectify the error(s) for a prescriptive model.

Task 5.3: Run, verify, and evaluate the model performance and outputs.

- CAP-X 5.3.1: Identify and explain the correct interpretation with a descriptive/ diagnostic analytics model output.
- CAP-X 5.3.2: Identify the correct ensemble modeling approach for a set of predictive analytics models.
- CAP-X 5.3.3: Identify an appropriate interpretation of the output from a predictive analytics model.
- CAP-X 5.3.4: Identify and explain the correct verification of the solution for a prescriptive analytics model output.
- CAP-X 5.3.5: Identify and explain which data and/or model issue is the cause of an observed biased or unethical outcome of a predictive or prescriptive model.

Task 5.4: Calibrate model and improve data applicability based on client input.

- CAP-X 5.4.1: Identify and explain the combination of data and model issues that are the cause of a client concern with the output of a predictive analytics model.
- CAP-X 5.4.2: Identify and explain the combination of data and model issues that are the cause of a client concern with the output of a prescriptive analytics model.

Task 5.5: If necessary, integrate multiple models.

- CAP-X 5.5.1: Identify and explain conflicts and synergies between multiple models.

Task 5.6: Document and communicate model findings including assumptions, limitations, and constraints.

- CAP-X 5.6.1: Identify and explain appropriate assumptions, limitations and constraints when communicating findings for an audience.

DOMAIN VI: DEPLOYMENT

The deployment domain includes activities related to the delivery of the analytics solution and requirements to support the deployment into the business.

Task 6.1: Perform business validation of the analytics solution.

- CAP-X 6.1.1: Identify how to perform business validation of the analytics solution and how it should be communicated to the client.
- CAP-X 6.1.2: Identify the business use of a model that has the most risk of adverse ethical analytics consequences.

Task 6.2: Deliver business validation report with findings.

- CAP-X 6.2.1: Identify a non-technical summary of technical findings that is most appropriate for a non-technical audience.

Task 6.3: Obtain sponsor agreement and stakeholder alignment on moving forward with deployment.

- CAP-X 6.3.1: Identify the characteristics of business impacts of implementing the analytics solution to ensure stakeholder alignment.

Task 6.4: Create requirements for the analytics solution including model, usability, system and business.

- CAP-X 6.4.1: Identify and explain the appropriate documentation of a production system for a business user.
- CAP-X 6.4.2: Identify characteristics of the model and report documentation necessary so that an analytics solution could be reused if the business circumstances should change.

Task 6.5: Actively support the implementation and testing of the production analytics solution/system.

- CAP-X 6.5.1: Identify the characteristics of deployment testing and why they are important.
- CAP-X 6.5.2: Identify the responsibilities of the analytics professional in the delivery of the production analytics solution/system.

Task 6.6: Actively support deployment validation and verification, including production data flows.

- CAP-X 6.6.1: Identify which data sources need to be synchronized during deployment when using multiple data sources.

DOMAIN VII: ANALYTICS SOLUTION LIFECYCLE MANAGEMENT

The analytics solution lifecycle management domain includes activities related to continuous oversight and calibration and training activities to ensure the analytics solution continues to function as planned and returns valid answers.

Task 7.1: Track analytics solution performance.

- CAP-X.7.1.1: Identify factors that may affect analytics solution performance over time.

Task 7.2: Recalibrate and maintain the analytics solution.

- CAP-X.7.2.1: Identify potential opportunities and risks for recalibration of the analytics solution.
- CAP-X.7.2.2: Identify and compare alternative solutions when there are changes in the external environment.

Task 7.3: Support training activities.

- CAP-X.7.3.1: Identify the type of training that is needed for business users in different roles.

Task 7.4: Validate the business case for the analytics solution over time.

- CAP-X.7.4.1: Identify the benefits and costs that are directly attributable to the analytics solution.

Task 7.5: Analyze side effects of the analytics solution over time.

- CAP-X.7.5.1: Identify an outcome from a predictive model that exhibits ethical issues but is not readily apparent in the data.

Task 7.6: Ensure documentation is complete and/or maintained.

- CAP-X.7.6.1: Identify the lifecycle activities that should be documented and monitored.