THE Open GROUP

The Open Group® Professional Certification Program

Conformance Requirements for the Data Scientist Profession (Open CDS)

Version 1.0
December 2018

The Open Group® Professional Certification Program:
Conformance Requirements for the Data Scientist Profession (Open CDS)

Document Number: X1822

Published by The Open Group, December 2018.

Comments relating to the material contained in this document may be submitted to:

The Open Group, 800 District Avenue, Suite 150, Burlington, MA 01803, United States

or by electronic mail to:

ogspecs@opengroup.org

Contents

1. Introduction

The Open Group Professional Certification Program (the Program) is designed to validate the existence of those qualities and skills in a professional that enable the effective development, implementation, and operation of Business or IT-related specializations. The Program is skills and experience-based and goes beyond validating the mastery of any specific knowledge base.

The Program covers multiple Professions. This document is for the Data Scientist Profession, which leads to certification as an Open Certified Data Scientist (Open CDS).

The Open Group supports two different routes to certification:

- The first route is Direct Certification by The Open Group
- The second route is Indirect Certification through third-party programs accredited by The Open Group

The Conformance Requirements for each of the Professions in the Program apply equally to Direct and Indirect Certification.

This document is intended for individuals who wish to pursue Direct Certification in the Data Scientist Profession as Open Certified Data Scientists, and for organizations that wish to run Accredited Certification Programs internally.

1.1 Conformance Requirements

This document defines the requirements for certification as an Open Certified Data Scientist (Open CDS), and may also be used as the foundation of a Data Scientist Profession framework.

These Conformance Requirements define those qualities and skills in a professional that enable the effective practice of Data Science. The requirements are skills and experience-based and do not define or require any specific knowledge base.

The document may be used on its own or, in conjunction with other related documents from The Open Group, as a guide for individual career development as well as a framework for Data Scientist Profession programs within members of The Open Group and other public and private sector organizations. The Conformance Requirements are designed to be flexible and extensible so that the framework may be adapted and extended to meet the needs of any industry, country, or organization.

1.2 Key Documents

The Data Scientist Profession is based upon three key documents:

- The Open Group Professional Certification Program Certification Policy, which sets out the policies and processes by which a Data Scientist may achieve certification
- The Open Group Professional Certification Program Conformance Requirements for the Data Scientist Profession (Open CDS) (this document), in which are documented the skills and experience that a Data Scientist must possess to achieve certification
- The Open Group Professional Certification Program Configuration Document for the Data Scientist Profession (Open CDS), which outlines the specific certification policies and processes for the Data Scientist Profession

Practical information about the certification process is available through the Open CDS FAQ and other Open CDS documentation on the Certification Authority's website.

1.3 Levels of Certification

The Program recognizes three levels of certification:

Level 1:	A professional who is able to perform with assistance/supervision with a wide range of appropriate skills as a contributing professional.
Level 2:	A professional who is able to perform independently and take responsibility for delivery of solutions as lead professional.
Level 3:	A professional who has significant breadth and depth of impact on the business through the application of their Profession.

Candidates applying for Level 3 certification must have been previously certified at Level 2 in the Profession – such certification need not be current and may be to a previous version of the Program.

1.4 This Document

This document is the first version of the Open CDS Conformance Requirements.

1.5 Migration and Change History

This section details changes made to the Open CDS Conformance Requirements.

Version No.	Date	Change
1.0	December 2018	First publication.

2. Data Scientist Roles and Responsibilities (Informative)

According to the Program a Data Scientist works with business leaders to solve problems by understanding, preparing, and analyzing data to predict emerging trends and provide recommendations to optimize business results.

Data Scientists typically have academic training in a quantitative discipline such as statistics, operations research, machine learning, or econometrics and use a variety of data and analytics tools and languages. Business acumen is an important skill for Data Scientists; for example, in understanding the business problem, in influencing strategic choices through data, and in deploying the solution. To effectively communicate their findings to business leaders, Data Scientists need strong communication, visualization, and storytelling skills.

2.1 Characteristics of the Data Scientist

The key skill and contribution Data Scientists bring to their pursuits is the ability to integrate statistical, economic, and computer science techniques to address business problems.

Effective Data Scientists typically possess and exhibit the following characteristics:

Skills and experience producing analytic models	Data Scientists develop statistical models; a combination of postulates, data, and inferences designed to solve a business problem. In order to accomplish this they must be proficient in the techniques that go into the formulation of models, including business understanding, data discovery and transformation, identification and assessment of alternative modeling approaches, model selection, validation, and formulation of solution context for deployment.
Appropriate skills and experience, including Data Science breadth	Data Scientists require business acumen with practical skills and experience using many modeling techniques for solving different business problems, experience with multiple technologies, and integration methods. While often relying on other professionals to set up the modeling environment and helping to integrate with existing processes, the Data Scientist must have broad enough skills and experience to be able to successfully communicate with both executives and technical staff. Beyond that base of Data Science breadth, effective Data Scientists usually possess additional skills in multiple industries.
Disciplined, method-driven execution	The Data Scientist uses formalized methods to guide and drive the development of solutions, the management of their work, and the production of their deliverables.
Full lifecycle experience	In the development of Data Science solutions that address business problems, the Data Scientist's work spans all phases of the business cycle, from product identification through service delivery. Thus, the Data Scientist must be able to clearly scope their work effort to fit within the bounds of the business environment. This may be a short consulting engagement to make recommendations to improve an individual process issue or it may involve capturing real-time data to alert with issues or provide recommendations to a manufacturing team.
Leadership	An effective Data Scientist is a leader, providing knowledge, technical, and team leadership skills in their work, to their clients, and for their teams.

Strong personal and professional skills	The Data Scientist must have a high level of communication, consulting, and client relationship skills. The Data Scientist must be able to clearly communicate complex technical and business concepts to clients (internal or external) and team members, and to negotiate change when needed. Problem-solving of client business and technical issues is a key role of the Data Scientist and he or she must be capable of effectively identifying and framing problems, extracting and transforming elements of information, and integrating this information to produce timely and well-considered decisions.

Wherever the word "client" appears in this document it is intended to be read as meaning either an internal (in-house) client, or an external client as part of a consulting engagement.

3. Level 1 and Level 2 Conformance Requirements (Normative)

The Conformance Requirements for a Level 1 or Level 2 Open Certified Data Scientist are broken down as follows:

- Core Basic skills
- Data Scientist Basic skills
- Experience Profile requirements
- Professional Development requirements
- Community Contribution requirements
- Experience requirements

3.1 Skill Levels

For the Core Basic skills and Data Scientist Basic skills, Candidates must meet or exceed the minimum skill level defined for each of the skills.

Table 1 lists the definition of skill levels and associated proficiency levels.

Table 1: Skill Level Definitions

Skill Level	Proficiency	Experience
Limited	Limited or no knowledge	None
General	General conceptual knowledge only	Limited – read about it, some education
Applied	Applied knowledge	Performs with supervision or mentoring
Deep	In-depth knowledge	Mastered the current state-of-the-art and is able to perform without supervision
Expert	Expert knowledge	Advances the state-of-the-art

3.2 Core Basic Skills

Table 2 lists the Core Basic skills for Level 1 and Level 2.

To achieve certification Candidates must be able to document that they have demonstrated these skills at the required level (or higher) repeatedly and successfully.

The term <discipline-specific> refers to the Data Scientist Profession.

Table 2: Core Basic Skills for Level 1 and Level 2

Ref.	Skill	Description	Rationale	Skill Level 1	Skill Level 2
CBS01	Communicate in Writing	Good written communications of <discipline-specific> topics, including the use of proper grammar, spelling, document organization, clarity, and use of appropriate content for the audience to meet its purpose.	Open Certified professionals need to be able to effectively communicate their <discipline-specific> topics – topics that are critical for the continuation of the work.	Applied	Deep
CBS02	Communicate Verbally	Good verbal communications, with responsiveness to questions, ability to stay on subject, use of good feedback, and follow-up questions, etc., leading to effective two-way communication. Culturally-appropriate body language is expected in face-to-face meetings and video teleconferencing.	Open Certified professionals need to be able to effectively communicate their <discipline-specific> topics – topics that are critical for the continuation of the work.	Applied	Deep
CBS03	Leading Teams	Given a scope of <discipline-specific> work to be accomplished, plan the work, form a team to perform the work, and guide the team and its members in performing the work to completion.	Open Certified professionals must be able to take on a leadership role leading to results in the scope of the work and therefore must exhibit leadership skills.	Applied	Deep

Ref.	Skill	Description	Rationale	Skill Level 1	Skill Level 2
CBS04	Mediate Equitable Solutions	Given a conflict that jeopardizes the integrity of the solution, mediate differing stakeholder opinions to arrive at equitable resolutions that ensure successful and stable outcomes.	Open Certified professionals must be capable of maintaining the integrity of their work products while simultaneously serving the needs of multiple stakeholders of diverse needs.	Applied	Deep
CBS05	Understand Business Aspects	Understand the stakeholders' business needs, how they relate to their business and mission, and to the <discipline-specific> activities.	Open Certified professionals must have business insight into how <discipline-specific> activities and work products serve the business needs of a variety of stakeholders and how they relate to the larger business context.	Applied	Deep
CBS06	Develop Solutions	Given one or more business or technical requirements, create the <discipline-specific> structures of a system or solution that can be validated to meet those requirements while adhering to business and/or technical constraints.	Open Certified professionals must be skillful in creating solutions that can be demonstrated to solve problems while adhering to business and/or technical constraints.	Applied	Deep
CBS07	Manage Discipline-specific Risks within a Project	Given a project plan, identify those elements of the plan that put the integrity of the <discipline-specific> aspects of the plan/timeline at risk. Manage those elements through to completion as agreed by the client/project manager.	Open Certified professionals must be able to work closely with clients/project managers and address issues in project plans that put their work at risk. They must be able to communicate the likelihood and impact of risks and come to a mutual agreement with clients/project managers.	Applied	Deep

3.3 Data Scientist Basic Skills

Table 3 lists the Data Scientist Basic skills for Level 1 and Level 2.

To achieve certification Candidates must be able to document that they have demonstrated these skills at the required level (or higher) repeatedly and successfully.

Table 3: Data Scientist Basic Skills for Level 1 and Level 2

Ref.	Skill	Description	Rationale	Skill Level 1	Skill Level 2
DSS01	Characterize Business Problems	Characterize the business problem into quantifiable form.	Open Certified Data Scientists must be able to collaborate with business stakeholders to identify intended beneficiaries, quantifiable improvements in business conditions, key performance indicators, and other factors needed to define the nature and scope of the business problem.	Applied	Deep
DSS02	Formulate Research Question and Build Hypotheses	Formulate the business problem as a research question with associated hypotheses, and determine what data is needed to answer the question in a way that provides value to the business.	Open Certified Data Scientists must be able to formulate the business problem as a research question, characterizing it as hypotheses to be tested to establish whether the problem can be solved satisfactorily and to determine whether extra data is needed to ensure completeness of analysis and deliver value to the business.	Applied	Deep
DSS03	Tailor Method	Given a business problem, tailor an industry standard method (such as CRISP-DM) to understand what is expected and deliver the best business value.	Open Certified Data Scientists must be able to work closely with clients/project managers on the definition of phases/tasks/outputs highlighting the role of business understanding, data understanding, data preparation, modeling, evaluation, and deployment without which analytic projects are most likely to fail.	Applied	Deep

Ref.	Skill	Description	Rationale	Skill Level 1	Skill Level 2
DSS04	Identify and Collect Data	Identify and collect initial data, explore data, and describe data quality level.	Open Certified Data Scientists must be able to identify and collect structured and unstructured data from many sources – e.g., databases (SQL/NoSQL), web pages, .csv, JSON, surveys – and manipulate the data through reshaping, summarization, combining, or sub-setting to produce a data description report that defines the meaning, range, and data type to be used in the performance of Data Science modeling. The professional must also be able to measure the completeness, quality, and timeliness of the data collected.	Applied	Deep
DSS05	Construct Usable Data	Construct usable data sets from multiple structured and unstructured data sources.	Open Certified Data Scientists must be able to prepare data for use in modeling and visualization. This includes data selection, imputation of missing values (if necessary), feature generation and transformations (mathematical function or categorization), derived attributes, merging and reformatting as needed.	Applied	Deep
DSS06	Using Techniques	Given requirements, identify and define/refine requirements and analytic techniques for turning functional requirements into a predictive or prescriptive analytics or AI solution that addresses business problems.	Certified Open Data Scientists understand the pros/cons of different data analysis (e.g., statistical process control, exploratory data analysis) and modeling techniques (e.g., unsupervised, supervised, classification, regression) and can describe why they are appropriate for a given method of approach. Ensure variable reduction techniques can be applied as appropriate during solution development.	Applied	Deep

Ref.	Skill	Description	Rationale	Skill Level 1	Skill Level 2
DSS07	Use Data Science Platforms and Collaboration Tools	Given business results, make effective use of Data Science platforms and collaboration tools to document, share, and improve quality and communication of results.	Open Certified Data Scientists train, test, and validate a wide variety of machine learning algorithms that are workable solutions to business problems. They collaborate with other Data Science professionals for peer review and validation of their modeling efforts to ensure quality and completeness of the work, using appropriate collaboration tools that enable practitioners to share assumptions, results, code, and model development.	Applied	Deep
DSS08	Validate Model	Given a business problem validate that the model meets its business requirements and its intended use.	Open Certified Data Scientists must be able to validate the results of modeling work using clear criteria and through exploration, experimentation, and triangulation. Model assumptions, test design, parameter settings, as well as success criteria (e.g., precision, recall, F-statistic) must all be validated prior to model deployment.	Applied	Deep
DSS09	Translate Insight into Business Value	Translate complex quantitative or qualitative analysis into compelling and easily understandable and actionable business insights through visualization and other appropriate techniques.	Open Certified Data Scientists must be able to make their results consumable by non-Data Scientist professionals. This means having a clear understanding of the appropriate visualization for the data available and the method(s) used in the analysis. Additionally, a good Data Scientist should understand the parameters of graphical excellence as well as the differences between friendly and unfriendly visualizations.	Applied	Deep

Ref.	Skill	Description	Rationale	Skill Level 1	Skill Level 2
DSS10	Deploy Solutions	Given validated model results, deploy the model and ensure that it is being used properly on an ongoing basis, and that any decline in model performance will be detected.	Open Certified Data Scientists must be able to deploy validated models in an operational environment, including monitoring model factors, establishing a retraining schedule, and determining when a model has expired.	Applied	Deep

3.4 Experience Profile Requirements

An Experience Profile is a coherent written description of a project or Data Scientist engagement that provides a Candidate with the opportunity to show how they perform as a Data Scientist, and enables a Peer Review Board to understand and question the Candidate's thought processes and decisions.

Each Experience Profile must describe a project completed not more than eight (8) years preceding the submission of the Milestone Application Form to the Certification Authority. Projects over two (2) years long may be used for multiple Experience Profiles under either of the following conditions:

- **The project had clearly-defined work efforts which took place in parallel, each with their own solution development and design activities and their own deliverables**
- **The project had clearly-defined phases that were executed in succession, each with its own solution development and design activities and deliverables**

  Note that a second project phase that constructs and implements the solution developed by the first phase does not meet this requirement.

In either case, each profiled project entity must meet all of the Experience Profile criteria defined in Table 4 below.

Each Experience Profile must include:

- A description of the business purpose of the project
- A concise description of the project
- The Candidate's role
- The Candidate's period of involvement

Table 4 defines the attributes that must be present within Experience Profiles for Level 1 or Level 2 certification and against which the Experience Profiles will be evaluated.

Table 4: Required Attributes for Experience Profiles for Level 1 and Level 2

Ref.	Experience	Description	Rationale	Level 1	Level 2
EXP01	Successfully Developing Data Science Solutions	Demonstrated success.	Data Scientist professionals are expected to be successful in developing Data Science solutions that have been used by the business.	Candidates must have acted in the role of Data Scientist on a Data Science engagement or project that met its acceptance criteria. Meeting acceptance criteria means that a validated model that solved the problem was accepted or adopted.	Candidates must have acted in the role of Data Scientist on a Data Science engagement or project that met its acceptance criteria. Meeting acceptance criteria means that a validated model that solved the problem was accepted or adopted.
EXP02	Leading	Perform as a lead Data Scientist.	Data Scientist professionals are expected to lead the development of Data Science deliverables.	Not applicable to this level of certification.	Candidates must have performed as a lead Data Scientist in the development of a major project or subsystem.

3.5 Professional Development Requirements

Table 5 lists the Professional Development requirements for Level 1 and Level 2.

To achieve certification Candidates must be able to demonstrate that they have met the following requirements.

Table 5: Professional Development Requirements for Level 1 and Level 2

Ref.	Development	Description	Rationale	Level 1	Level 2
PD01	Training in the Discipline of Data Science	Candidates must have completed formal training in the discipline of Data Science (such as statistics, operations research, machine learning, or econometrics) either through attendance at a taught course, or through self-study.	Open Certified Data Scientists are expected to have undergone at least a basic level of training in their discipline. The Program is intentionally not specific about the nature of the training expected because of the evolving nature of the body of knowledge and the Profession.	Attendance at a taught course, or through self-study.	Attendance at a taught course, or through self-study.
PD02	Knowledge of Technology Trends and Techniques	Candidates are required to develop and maintain their knowledge of the technology, trends, and techniques that are relevant to developing solutions for solving business problems.	Open Certified Data Scientists are expected to develop and maintain an understanding of technology trends and techniques so that they can leverage that body of knowledge into feasible solutions.	Develop and maintain personal knowledge.	Maintain personal knowledge.
PD03	Knowledge of Vertical Industry Sectors (e.g., telecoms, financial, etc.)	Candidates are required to develop and maintain an understanding of their client's business as it pertains to their client's vertical industry (e.g., telecoms, financial, etc.).	Open Certified Data Scientists are expected to develop and maintain an understanding of their client's business as it pertains to their client's vertical industry (e.g., telecoms, financial, etc.).	Develop and maintain personal knowledge.	Maintain personal knowledge.

Ref.	Development	Description	Rationale	Level 1	Level 2
PD04	Skills and Knowledge in Data Science	Candidates are required to develop and maintain their skills and knowledge in Data Science.	Open Certified Data Scientists are expected to continue to develop their skills and to stay up-to-date with the development of their Profession.	Maintain personal skills and knowledge.	Maintain personal skills and knowledge.

3.6 Community Contribution Requirements

Table 6 lists the Community Contribution requirements for Level 1 and Level 2.

To achieve certification Candidates must be able to demonstrate that they have met the following requirements.

Table 6: Community Contribution Requirements for Level 1 and Level 2

Ref.	Contribution	Description	Rationale	Level 1	Level 2
CC01	Contribution to the Data Scientist Profession	Candidates must make contributions to the Data Scientist Profession; for example, publications, teaching, mentoring, research collaboration, or participation in professional organizations.	Open Certified Data Scientists are expected to contribute to the growth and vitality of their Profession.	Contribute to the Profession. Mentoring people in the Profession is required.	Contribute to the Profession. Mentoring people in the Profession is required.
CC02	Generation of Assets for Reusability	Candidates must demonstrate the creation of reusable assets, such as process documentation, packages, and libraries.	Open Certified Data Scientists must be able to develop assets to enhance the productivity of other Data Scientists through reuse.	Demonstrated experience of using reusable assets created by others or by providing assets for reuse. Two or more examples.	Demonstrated experience of creating reusable assets by providing 5 (five) or more examples of reuse assets (process documentation, libraries, packages, etc.).

3.7 Experience Requirements

Table 7 lists the Experience requirements for Level 1 and Level 2.

To achieve certification Candidates must be able to demonstrate that they have at least the following experience.

Table 7: Experience Requirements for Level 1 and Level 2

Ref.	Experience	Description	Rationale	Level 1	Level 2
EC01	Performance as a Data Scientist	Candidates must perform in the role of a Data Scientist and demonstrate experience using a variety of Data Science techniques during projects over a specified period of time.	Open Certified Data Scientists must demonstrate an understanding of multiple modeling techniques such as regression, simulation, machine learning, etc. over a minimum period of time.	Demonstrate the use of at least two (2) techniques, possibly with supervision, in at least one (1) project.	Demonstrate five (5) or more Data Science techniques over a period of at least 36 months.
EC02	Use Data and Analysis Tools/ Languages Across a Spectrum of the Data Science Workflow	Experience using a variety of data types (e.g., structured, unstructured) and analytics tools (e.g., SPSS, SAS) and languages (e.g., Python, R).	Open Certified Data Scientists must be able to use many different kinds of data (structured/ unstructured) and tools/languages (open source, proprietary) to perform their work.	Demonstrate the use of at least one (1) tool or language to manage structured or unstructured data.	Demonstrate the use of two (2) or more tools or languages to manage structured or unstructured data.
EC03	Communicate the Business Value of Insights and Limitations of the Analysis	Candidates must demonstrate experience of communicating insights and their limitations effectively to stakeholders using appropriate visualization techniques.	Open Certified Data Scientists must be able to communicate results effectively to a range of stakeholders and understand any limitations of the analysis.	Demonstrate the ability to convey insights and limitations to at least one (1) stakeholder group.	Demonstrate the ability to convey insights and limitations to two (2) or more stakeholder groups.

Ref.	Experience	Description	Rationale	Level 1	Level 2
EC04	Demonstrate Breadth of Business Problem Solved	Candidates must demonstrate the ability to solve a variety of Data Science business problems.	Open Certified Data Scientists must be able to solve a variety of business problems with Data Science methodologies.	Demonstrate two (2) or more.	Demonstrate three (3) or more.

4. Level 3 Conformance Requirements (Normative)

The Conformance Requirements for a Level 3 Open Certified Data Scientist are broken down as follows:

- Core Foundation skills
- Experience Profile requirements
- Professional Development requirements
- Community Contribution requirements
- Experience requirements

4.1 Core Foundation Skills

Table 2 lists the Core Foundation skills for Level 3.

To achieve certification Candidates must be able to document that they have demonstrated these skills at the required level (or higher) repeatedly and successfully.

Table 8: Core Foundation Skills for Level 3

Ref.	Skill	Description	Rationale	Skill Level 3
DDS01	Employ Collaborative Influence	Facilitate the implementation of an important business initiative by promoting teaming and cross-organizational participation.	Distinguished Data Scientists achieve results that require support and collaboration of disparate groups with potentially conflicting interests.	Deep
DDS02	Employ Cross-Organizational Leadership	Initiate, lead, and influence multi-disciplinary initiatives across organizational boundaries coordinating the activities necessary to succeed.	Distinguished Data Scientists lead successful business transformations that involve multiple disciplines across organizational boundaries.	Deep
DDS03	Manage Risks	Guide an organization's strategy to recognize and mitigate the weaknesses or exposures in their plans and implementations in a way that secures successful and sustainable outcomes.	Distinguished Data Scientists detect and mitigate risks that jeopardize the business's initiatives, compliance, and/or organizations at an enterprise scale.	Deep
DDS04	Develop Strategic Plans	Identify and drive strategic decisions and plans for an enterprise.	Distinguished Data Scientists are responsible for driving decisions and plans that affect the strategy of an enterprise.	Deep
DDS05	Manage Cross-Organizational Projects	Allocate project activities and Data Science assignments from multiple projects to multiple teams of Data Scientists, across multiple organizational units.	Distinguished Data Scientists can manage complex projects that involve multiple organizational units and multiple teams of Data Scientists.	Deep

Ref.	Skill	Description	Rationale	Skill Level 3
DDS06	Communicate on an Executive Level	Communicate Data Science vision and strategies to business and/or project/program executives in a way that is appropriate and gains their commitment.	Distinguished Data Scientists communicate in a manner that convinces at the executive level.	Deep
DDS07	Advocate Stakeholders' Interests	Simultaneously advocate multiple stakeholders' interests.	Distinguished Data Scientists understand and advocate the various, and potentially conflicting, interests and views of multiple stakeholders.	Deep
DDS08	Engage Thinking at Strategic Levels	Apply strategic thinking to vision, mission, and strategy to deliver positive impact and results to the business.	Distinguished Data Scientists apply holistic and strategic thinking for the enterprise in order to identify opportunities that deliver significant positive business impact.	Deep
DDS09	Provide Solution Optimization	Assess, analyze, and provide recommendations for optimizing solutions.	Distinguished Data Scientists advise on the selection of data modeling techniques and provide inputs to optimize the solution in terms of cost, quality, and time.	Deep
DDS10	Innovate Breakthrough Solutions	Provide breakthrough innovation in the use of theoretical results or technology to deliver greater business value.	Distinguished Data Scientists innovate in the creation and use of theoretical results or technology to deliver business value to their clients.	Deep
DDS11	Apply Governance Solutions	Apply and maintain processes and policies for governance in those programs or projects for which the Data Scientist is responsible.	Distinguished Data Scientists apply and maintain processes and policies for governing data in the projects and programs for which they are responsible.	Deep
DDS12	Set Metrics for Compliance	Establish metrics for validating the conformance of an implementation to a solution.	Distinguished Data Scientists select and/or create metrics that are appropriate for assessing compliance of implementations to their solution.	Deep

4.2 Experience Profile Requirements

An Experience Profile is a coherent written description of a project or Data Scientist engagement that provides a Candidate with the opportunity to show how they perform as a Distinguished Data Scientist, and enables a Peer Review Board to understand and question the Candidate's thought processes and decisions.

Each Experience Profile must describe a project completed not more than eight (8) years preceding the submission of the Milestone Application Form to the Certification Authority. Projects over two (2) years long may be used for multiple Experience Profiles under either of the following conditions:

- **The project had clearly-defined work efforts which took place in parallel, each with their own solution development and design activities and their own deliverables**
- **The project had clearly-defined phases that were executed in succession, each with its own solution development and design activities and deliverables**

  Note that a second project phase that constructs and implements the solution developed by the first phase does not meet this requirement.

In either case, each profiled project entity must meet all of the Experience Profile criteria defined in Table 9 below.

Each Experience Profile must include:

- A description of the business purpose of the project
- A concise description of the project
- The Candidate's role
- The Candidate's period of involvement

Table 9 defines the attributes that must be present within Experience Profiles for Level 3 certification and against which the Experience Profiles will be evaluated.

Table 9: Required Attributes for Experience Profiles for Level 3

Ref.	Experience	Description	Rationale	Skill Level 3
DEXP01	Leading Successful Significantly Complex Projects	Distinguished Data Scientists have repeated experience successfully directing significantly complex Data Science initiatives.	Distinguished Data Scientists have experience with developing solutions that typically involve different stakeholders, numerous multiple data types, multiple data sources, multiple integrated techniques, and require highly critical accuracy.	The Candidate must explain why the project was significantly complex, and how their involvement was essential to the success of the project.

Ref.	Experience	Description	Rationale	Skill Level 3
DEXP02	Leading Successful Projects with Significant Business Impact	Distinguished Data Scientists have repeated experience successfully directing projects that significantly, positively, and measurably affect the business.	Distinguished Data Scientists have experience with developing solutions that result in significant positive business value and/or impact.	The Candidate must explain why the project was successful and how it impacted the business significantly, and how their involvement was essential to the success of the project.
DEXP03	Driving Significant Strategic Initiatives	Distinguished Data Scientists have repeated experience in setting and driving Data Science goals and objectives for significant strategic initiatives.	Distinguished Data Scientists provide thought leadership towards directing and delivering the mission of the organization.	The Candidate must explain how they drove a strategic initiative and how their involvement was essential to the success of the initiative.

4.3 Professional Development Requirements

Table 10 lists the Professional Development requirements for Level 3.

To achieve certification Candidates must be able to demonstrate that they have met the following requirements.

Table 10: Professional Development Requirements for Level 3

Ref.	Development	Description	Rationale
DPD01	N/A	N/A	N/A
DPD02	Knowledge of Technology Trends	Candidates are required to continually develop and maintain their knowledge of the technology trends that are relevant to developing solutions for solving business problems.	Distinguished Data Scientists continually develop and maintain their understanding of the technology trends so that they can leverage that body of knowledge into feasible solutions.
DPD03	Knowledge of Vertical Industry Sectors	Candidates are required to develop and maintain an understanding of their client's business as it pertains to their client's vertical industry (e.g., telecoms, financial, etc.).	Distinguished Data Scientists develop and maintain an understanding of their client's business as it pertains to their client's vertical industry (e.g., telecoms, financial, etc.).
DPD04	Skills and Knowledge in Data Science Disciplines	Candidates must continually develop their skills and knowledge of the latest data science-related trends and techniques.	Distinguished Data Scientists continue to develop their skills and to stay up-to-date with the development of their Profession.

4.4 Community Contribution Requirements

Table 11 lists the Community Contribution requirements for Level 3.

To achieve certification Candidates must be able to demonstrate that they have met the following requirements.

Table 11: Community Contribution Requirements for Level 3

Ref.	Contribution	Description	Rationale
DCC01	Contribution to the Data Scientist Profession	Candidates must make contributions to the Data Scientist Profession; for example, publications, teaching, contributions to open source, research collaboration, or participation in professional organizations.	Distinguished Data Scientists are expected to contribute to the growth and vitality of their Profession inside and outside their organization.
DCC02	Development of Professional Data Scientists	Candidates are expected to develop professional Data Scientists through teaching, serving on review boards, coaching, and/or mentoring.	Distinguished Data Scientists transfer knowledge and experience to the Data Scientist community.

4.5 Experience Requirements

Table 12 lists the Experience requirements for Level 3.

To achieve certification Candidates must be able to demonstrate that they have at least the following experience.

Table 12: Experience Requirements for Level 3

Ref.	Experience	Description	Rationale
DEC01	Establish Data Science Vision	Establish Data Science vision for significantly complex projects.	Distinguished Data Scientists establish strategies for projects that are of significant scope and/or scale. Significant scope and/or scale here refers to engagements that involve, for example, different stakeholders, numerous multiple data types, multiple data sources, multiple integrated techniques, and require highly critical accuracy.
DEC02	Innovative Leadership	Lead with impact, successfully introducing innovations by re-imagining existing problems through the application of new Data Science capabilities.	Distinguished Data Scientist apply advanced Data Science capabilities such as deep learning to re-imagine and solve existing problems that open up new opportunities.

Ref.	Experience	Description	Rationale
DEC03	Demonstrate Business Impact	The Candidate must have demonstrated positive significant impact on the business through a significantly complex project or initiative.	Distinguished Data Scientists demonstrate measurable positive significant impact on the business outcome through significantly complex projects or initiatives. The Candidate must have demonstrated this through financial or other KPIs.
DEC04	Influence Use of Reusable Assets	Candidates must encourage the implementation and reuse of assets across project domains.	Distinguished Data Scientists must be able to advocate the reuse of available assets.
DEC05	Generation of Assets for Reusability	Candidates must demonstrate their contribution to establishing the environment which supports the creation of reusable assets.	Distinguished Data Scientists must be able to advocate and encourage the development of reusable assets.

Acknowledgements

The contributions to this document by the following individuals are gratefully acknowledged:

- James de Raeve, The Open Group
- Andy Ellis, DXC Technology
- Neeraj Madan, IBM
- Maureen Norton, IBM
- George Stark, IBM
- Andras Szakal, IBM
- Krishna Venkatraman, IBM
- Cristina Woodbridge, The Open Group