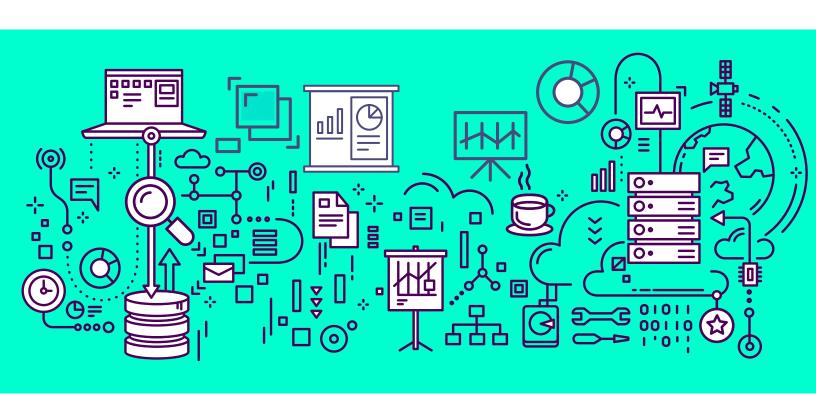# CDS™

CANDIDATE GUIDE

2023

Association of Data Scientists (ADaSci)

Incorporated in Bangalore, India in 2019, Association of Data Scientists (ADaSci) is the premier global professional body of data science & machine learning professionals. We aim to foster the development, dissemination and implementation of knowledge in data science & machine learning

Content

Introduction

The Chartered Data Scientist (CDS™) has set the global standard for data science. Developed by the world's leading data science practitioners, the designation signifies a mastery of the skills and knowledge needed to help organizations succeed in today's rapidly changing landscape. Its curriculum is updated annually by a group of distinguished professionals and leading academics of diverse backgrounds, ensuring that the designation meets the evolving demands of the global industry. By achieving your CDS certification, you indicate to potential employers that you are requisite knowledge of the field, providing you with an edge in your career and professional development.

**BECOMING A
CERTIFIED CDS**

The first step to becoming an CDS is passing a rigorous exam. To enroll in the program, register for the CDS Exam online via the ADASCI website. CDS Exam is offered throughout the year. After achieving a passing score on CDS Exam, candidates must demonstrate a minimum of two years of full-time work experience in data science or a related field to complete their certification. If you have questions about whether your work experience qualifies, please contact **info@adasci.org**

○ **CDS EXAM**
   THREE HOURS/150 QUESTIONS

○ **PROFESSIONAL WORK EXPERIENCE**
   TWO YEARS, FULL-TIME

WHO SHOULD CONSIDER THE CDS CERTIFICATION

DATA SCIENTISTS

The CDS helps professionals in data science, analytics, artificial intelligence or other critical areas to broaden their knowledge and enhance their skill set. Those who are just beginning their careers benefit from the breadth of the curriculum, which exposes candidates to the major strategic aspects of data science. More established practitioners often choose to become certified CDS in order to ensure that they are apprised of the latest trends in data science, or because they want to challenge themselves by testing their knowledge against an elite pool of data scientists from around the world.

CAREER CHANGERS

Whether you work in data science or are interested in transitioning to a data science role, becoming an CDS can help accelerate your career. Professionals from non- data science roles become CDS in order to develop specialized, practical knowledge that can be applied in the industry. Undertaking the rigorous course of study to become an CDS signals a commitment to an area of technology that is growing rapidly across the globe.

STUDENTS

Students with an interest in data science may elect to sit for the CDS Exam during or immediately after completing their studies. The CDS curriculum can complement their prior coursework or help them develop a foundation of specialized knowledge that goes beyond their academic curriculum. Since the CDS Exam is practitioner- driven, earning the CDS designation demonstrates to future employers that they are able to master complex real-world challenges.

Using the CDS to
Accelerate Your Career

The CDS certification is by far the best known and most respected designation for data science. As an CDS, you'll have a competitive advantage that can help you stand out to employers. By earning your certification, you'll demonstrate that you possess the knowledge and tools necessary to assess and manage the challenges associated with the dynamic financial services industry.

GLOBAL RECOGNITION
Mastering the CDS curriculum represents a significant commitment to the data science profession. Because of this, CDSs are considered industry leaders around the world, and are known for their strong dedication to professional excellence.

BENEFITS OF THE CDS
Demonstrate your knowledge Earning your certification shows that you have mastered the knowledge bases necessary to effectively assess and manage risk.

HIGHLIGHT YOUR EXPERIENCE
The designation testifies that you have worked in the field for at least two years, signaling to employers that you have the experience necessary to succeed.

JOIN AN ELITE GROUP
You'll be part of a network of like-minded professionals, allowing you to expand your career opportunities.

UNDERSCORE YOUR RELIABILITY
All CDSs are expected to adhere to the ADASCI Code of Conduct, letting employers know that CDSs will help safeguard their firms' reputations.

Exam Development,
Structure and Content

EXAM DEVELOPMENT

The CDS Program is developed under the guidance of the ADASCI Committee, which is comprised of prominent global data science experts and leaders. The CDS Committee establishes the topic areas tested on the Exam on an annual basis. To further align with industry needs and calibrate our understanding of the demands of the global risk management community, we also conduct formal surveys designed to determine the knowledge, skills and abilities required of effective data scientists. This process helps ensure that successful candidates are prepared to effectively contribute to their organizations.

EXAM STRUCTURE

The CDS consists of one computer based, multiple-choice exams. The CDS Exam consists of 150 equally weighted questions. Candidates are allotted three hours to complete the Exam. The Exams are offered in person in American English throughout the year. The Exams are comprehensive, practice-oriented assessments that cover the fundamental tools and techniques used in data science, their underlying theories.

EXAM CONTENT TOPICS AND WEIGHTS

Probability Theory, Statistics and Linear Algebra (weightage 12%)

Counting, Random variables, distributions, quantiles, mean-variance, p-Value, Confidence Interval, Hypothesis testing, t-test, z-test, Chi-Square test, Analysis of Variance (ANOVA), Conditional probability, base rate fallacy, Joint distributions, covariance, correlation, independence, Central limit theorem, Frequentist significance tests and confidence intervals, Maximum Likelihood Estimation, Bayes' theorem and Bayesian statistics, Scalars, Vectors, Matrices, and Tensors. Multiplying Matrices and Vectors, Eigen decomposition, Singular Value Decomposition.

Data Engineering and Databases (weightage 8%)

Relational databases, Non-relational databases, Graphical databases, batch processing, stream processing, in-memory processing, data management, SQL Databases, NoSQL Databases, data warehouses, data lakes, ETL and Data pipelines, data access, governance and integration, operations and security.

Exploratory Data Analysis (weightage 8%)

Data understanding, Data visualization, Visualization techniques, Univariate analysis, Multivariate analysis, EDA and visualization tools.

Supervised and Unsupervised Learning (weightage 15%)

Linear and Non-linear Models, Classification, Regression, K-Nearest Neighbours, Naïve Bayes, Clustering, K-Means Clustering, Hierarchical Clustering, Various learning errors, regularization, estimator bias-variance trade-off, active learning, Support vector machine (SVM) and kernels, Model selection and model selection criteria, Ensemble learning - bagging and boosting, Expectation-Maximization (EM) algorithm, Hidden Markov models, Bayesian networks, Probabilistic inference, Association Rule Learning, Reinforcement Learning, Time-Series Analysis, Cross-Validation.

Neural Networks and Deep Learning (weightage 11%)

Feedforward Networks, Backpropagation Learning, Gradient Descent, Optimization techniques for neural networks, Regularization techniques, Neural net training algorithms, Convolutional Networks, Recurrent and Recursive Neural Networks, Representation Learning, Autoencoders, Deep Generative Models, Factor Analysis, t-Distributed Stochastic Neighbour Embedding (t-SNE), Transfer learning frameworks.

Natural Language Processing (weightage 8%)

Text extraction and preprocessing, Text Classification, Sentiment analysis, Information retrieval, Parsing, Tokenization, Vectorization, Part of Speech (POS) Tagging, Sequence modelling, Word embedding, Word2Vec, Language Modelling, NLP Transformers, Generative models for NLP.

Computer Vision (weightage 8%)

Image processing, Image transformations, Image segmentation, Image classification, Object recognition, Image reconstruction, Image augmentation, Image tagging, Video analytics, Transformers in computer vision, Generative models for computer vision.

Deployment and Model management (weightage 8%)

Deployment of machine learning model, CI/CD and deployment pipelines, tracking model quality, reporting and visualization mechanisms for model performance, MLOps, Data drift and concept drift.

Programming Frameworks for Data Science (weightage 10%)

Python Lists, Dictionaries, Arrays, Conditional statements, Loops, Data Frame, Functions, Object-oriented programming, Files, Exceptions, Sci-Kit Learn, Keras, TensorFlow, PyTorch, Python tools for deep learning.

Business and Data Science (weightage 12%)

Identifying stakeholders, Handling data privacy concerns, Determining problem-data science fit, defining problem statements for multiple stakeholders, understanding constraints and scope of data science projects, Defining and communicating business benefits, identifying data sources and creating initial reports, Decision Modelling.

Payments
and Fees

The CDS Exam requires a program enrollment fee of USD **250.** Enrolling of CDS exam also gives you a one-year

membership of **Association of Data Scientists (ADaSci).**

-

Reference Textbooks

CDS provides Reference Textbooks, practice exams and more to help you get ready for the Exam. Due to the sizable amount of material covered, we suggest that you use a weekly study schedule. Preparation time will vary based on your prior professional experience, academic background and familiarity with the curriculum's concepts. Preparing for the Exam at the last minute is strongly discouraged.

REFERENCE TEXTBOOKS:

Section 1:

A Course in Probability Theory, Kai Lai Chung, Academic Press. An Introduction to Statistical Learning: With Applications in R, Daniela Witten, Gareth James, Robert Tibshirani, and Trevor Hastie, Springer Publication. Introduction to Probability Models, 9th Edition, Sheldon M.

Section 2:

Database System Concepts Textbook by Avi Silberschatz, Henry F. Korth, and S. Sudarshan, McGraw Hill Publication. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, Martin Kleppmann, O'Reilly Publication.

Section 3:

Practical Statistics for Data Scientists: 50 Essential Concepts, Peter Bruce and

Andrew Bruce, O'Reilly Publication.

Section 4:

Pattern Recognition and Machine Learning, Christopher Bishop, Springer Publication. Machine Learning, Tom M. Mitchell, McGraw Hill Publication. The Elements of Statistical Learning: Data

Mining, Inference and Prediction, 2nd Edition, T Hastie, R Tibshirani and J Friedman, Springer Series in Statistics, Springer Publications.

Section 5:

Deep Learning Book by Aaron C. Courville, Ian Goodfellow, and Yoshua Bengio, MIT Press. Machine Learning A Probabilistic Perspective, Kevin P.

Murphy, MIT Press. Neural Networks and Learning Machines, 3rd Edition, Simon Haykin, Pearson Publication.

Section 6:

Foundations of Statistical Natural Language Processing, Christopher D. Manning and Hinrich Schutze, The MIT

Press. Natural Language Processing with Python, Steven Bird, Ewan Klein and Edward Loper, O'Reilly Publication.

Section 7:

Computer Vision: Algorithms and

Applications, Richard Szeliski, Springer Publication.

Section 8:

Evaluating Machine Learning Models, Alice Zheng, O'Reilly Publication. Building Machine Learning Powered Applications, Emmanuel Ameisen, O'Reilly Publication.

Section 9:

Python Cookbook: Recipes for Mastering Python 3, 3rd Edition, David Beazley & Brian K. Jones, O'Reilly Publication. Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts, Tools and Techniques to Build Intelligent Systems, 2nd Edition, Aurelien Geron, O'Reilly Publication. R for Data Science: Import, TIDY, Transform, Visualize, and Model Data, Hadley Wickham and Garrett Grolemund, O'Reilly Publication.

Section 10:

Laursen GHN, Thorlund J (2016) Business Analytics for Managers: Taking Business Intelligence Beyond Reporting, 2nd ed. (John Wiley & Sons, Hoboken, NJ). Business Analytics, 2nd Edition, James Evans, Pearson Publication.

Chartered Data Scientist™ (CDS)
Exam Terms and Conditions

Every aspirant for the Chartered Data ScientistTM (CDSTM) exam is required to accept and adhere to the following terms and conditions and privacy policy agreements.

1. Important Note

1.1 These terms and conditions are entered between the Candidate (You as a prospective awardee of Chartered Data Scientist) and the ADaSci (Association of Data Scientists) as on the date when the candidate has checked "I accept ADaSci Terms and Conditions" option at the time of registration.

1.2 The ADaSci has the right to change, and/or update these terms and conditions from time to time.

2. Purpose of the Award

2.1 The following are the purpose of the award of Chartered Data Scientist given by ADaSci:

2.1.1 It is a recognition of skill sets, knowledge and expertise in the field of data science of the person who owns this charter.

2.1.2 It is not an academic/professional degree or diploma or equivalent.

3. Eligibility Criteria

3.1 The minimum age of the candidate to appear for the CDS exam must be 18 years on the date of registration.

3.2 If the age of the candidate is between 13 to 18 years, there must be an undertaking signed by the parents or the guardian.

3.3 The age must not be less than 13 years.

3.4 The candidate must have at least two years of relevant experience as a Data Scientist for the award of this charter. However, the candidate can appear in the exam and receive the result, but the award of Chartered Data Scientist will be put on hold until the candidate attains two years of experience.

4. Data and Information

4.1 The candidate must provide correct information related to Name, Date of Birth, Address, Parents, Email Address, Mobile Number, Qualifications, Experience, etc at the time of registration.

4.2 The ADaSci has the right to verify the information provided by the candidate at the time of registration.

4.3 If, at any point in time, the information provided by the candidate is found incorrect, the ADaSci has the right to cancel the registration and if the charter is awarded, it has the right to revoke the charter from the candidate.

5. Exam Result and Report

5.1 The candidate must take the exam within one year after the registration of the CDS exam.

5.2 The candidate must understand that the ADaSci has the right to set the minimum passing criteria for the CDS exam.

5.3 ADaSci has the right to change the minimum passing criteria for the CDS exam from time to time in

order to maintain the quality of the exam and the purpose of the award.

5.4 ADaSci has the right not to divulge the minimum passing criteria.

5.5 On successfully passing the exam based on the passing criteria, the candidate will be provided with a result report.

6. Award of Charter

6.1 The award of **Chartered Data Scientist** will be provided to the candidate by ADaSci on successfully passing the exam along with the following conditions:

6.1.1 Verification of all the information including experience details provided by the candidate during registration.

6.1.2 If the candidate has not attained the two years of experience at the time of the exam and when he or she provides the required experience details later.

7. Re-Evaluation

7.1 As per the current regulations, there is not any provision for the re-evaluation of the exam.

8. Ownership of the Charter

8.1 The individual person who appeared for the exam and was awarded the charter will be the sole owner of the **Chartered Data Scientist** designation.

8.2 Any organization or institution cannot claim its name as the owner of this charter on behalf of the individual person. If done, it will be considered as a violation of terms and conditions, guidelines and ethics.

9. Validity

10.1 The awarded charter will have lifetime validity.

10. Attempts

10.1 An individual can take any number of attempts to pass this exam throughout the year and there is not any limit on attempts.

11. Revocation of the Charter

11.1 In case of any violation of the terms and conditions, especially points 4.3 and 9.2, the ADaSci has the right to revoke the awarded charter from the person who owns it.

11.2 If, at any time in the future, it comes to the notice of ADaSci any misconduct by the candidate during the exam, the awarded charter will be revoked from the person who owns it.

11.3 If, at any time in the future, the owner of this charter is found not adhering to guidelines and ethics as defined by ADaSci, the awarded charter will be revoked from the person who owns it.

12. Other Notes

12.1 The ADaSci does not provide the following guarantee to the owner of this charter:

12.1.1 To get a job in any organization or provide any type of job assistance.

12.1.2 To get admission to any institution for study.

ADaSci | THE ASSOCIATION
OF DATA SCIENTISTS

#280, 2ND FLOOR, 5TH MAIN, 15TH A CROSS
RD, SECTOR 6, HSR LAYOUT, BENGALURU,
KARNATAKA 560102

WWW.ADASCI.ORG