

Data Science Practitioner (DSP) Examination Overview and Syllabus for Candidates

IABAC.org

Copyright © IABAC 2015-2017.

IABAC®, IABAC logo®, DSF, DSP and CDS are all registered trade marks of IABAC Pvt Limited.
Reproduction of this material requires the permission of IABAC

All rights reserved. EN_DSP_2017_CandidateSyllabus_V1.6

Introduction

This document is intended to provide information to prospective Data Science Practitioner (DSP) candidates and indeed anyone interested in the format and content of the DSP Examination.

It includes:

- Data Science Practitioner (DSP) Examination Overview – general information and overview of the exam format
- Data Science Practitioner (DSP) Syllabus for Candidates – detailing:
 - o the format of the exam and the learning outcomes from the delegates' course of study that will be assessed in each section
 - o the specific criteria by which each learning outcome is assessed

Data Science Practitioner Examination Overview

Pre-requisite qualifications: Candidates pursuing this examination needn't hold any prerequisites.

Training: Though formal training is not mandatory, attendance at Registered Education Partner (REP) course is strongly recommended.

Material permitted: The examination is 'closed book'. No material permitted

Examination duration: 90 mins.

Format:

1. 15 questions of Multiple Choice Questions.
2. Each question carries 5 Marks / 10 Marks
3. No Negative marking.

Pass mark: Candidates are expected to achieve a score of 55% (55 marks) or higher in order to pass the examination and be awarded certification.

Syllabus for Candidates

Exam Sections, Syllabus and Bloom's Index

Knowledge Area	Syllabus Details	Exam Weightage	Bloom's Index
KAG1-DSDA: Data Analytics group including Machine Learning, statistical methods, and Business Analytics	<p>Theory of computation Design and Analysis of Algorithms Machine Learning Theory</p> <p>Mathematics of computing Discrete Mathematics and Graph Theory Probability & Statistics Probabilistic reasoning Statistical methods, including descriptive statistics, exploratory data analysis (EDA) and confirmatory data analysis (CDA) Information theory Mathematical analysis Mathematical software and tools</p> <p>Computing methodologies Artificial Intelligence Natural Language Processing Knowledge Representation and Reasoning Data mining and knowledge discovery Text analysis, Data mining Text analytics including statistical, linguistic, and structural techniques to analyse structured and unstructured data Machine Learning theory and algorithms Classification methods</p> <p>Information systems (to support Data Science applications) Decision Analysis and Decision Support Systems Data warehousing and Data Mining Data Analysis and statistics Multimedia information systems Data Mining Predictive analytics and predictive forecasting</p> <p>Big Data Technologies and Systems Big Data algorithm for large scale data processing</p>	70%	4

Copyright © IABAC 2015-2017.

IABAC®, IABAC logo®, DSF, DSP and CDS are all registered trade marks of IABAC Pvt Limited. Reproduction of this material requires the permission of IABAC

All rights reserved. EN_DSP_2017_CandidateSyllabus_V1.6

	<p>Big Data Analytics Big Data systems Big Data algorithms for data ingest, pre-processing, and visualisation Big Data analytics platforms and tools (including Hadoop, Spark, and cloud based Big Data services) Big Data systems for application domains</p>		
<p>KAG2-DSENG: Data Science Engineering group including Software and infrastructure engineering</p>	<p>Computer systems organisation for Big Data applications (including high performance networks) Parallel and Distributed Computer Architecture Computer networks: architectures and protocols Computer networks for high-performance computing and Big Data infrastructure</p> <p>Big Data software organisation and engineering Software (systems) architectures Requirements engineering and software systems development Large and ultra-large scale software systems organisation Cloud enabled applications development</p> <p>Modelling and simulation Modelling and simulation theory and techniques (general and domain oriented) Large scale modelling and simulation systems</p> <p>Big Data systems organisation and management Enterprise information systems Large scale data storage and data management systems Collaborative and social computing systems and tools</p> <p>Big Data (Data Science) applications design Programming languages for Big Data analytics: R, python, others Models and languages for complex interlinked data presentation and visualisation</p>	5%	4

<p>KAG3-DSDM: Data Management group including data curation, preservation and data infrastructure</p>	<p>Data management systems Database management systems Database design and models Data Modelling, Databases and Database Management Systems Data Models and Query Languages Database administration</p> <p>Digital libraries and archives Digital libraries and archives organisation Information Retrieval Data curation and provenance Search Engines technologies</p> <p>Data Management and Enterprise data infrastructure Data management, including Reference and Master Data Data Warehousing and Business Intelligence Data storage and operations Data archives/storage compliance and certification Metadata, linked data, provenance Data infrastructure, data registries and data factories Data security and protection Data governance, data quality, data Integration and Interoperability Data Management Planning Responsible data use, data privacy, ethical principles, legal issues</p> <p>Data Management and organisation Research data infrastructure, Open Science, Open Data, Open Access, ORCID Data infrastructure compliance and certification Ethical principle and data privacy</p>	<p>5%</p>	<p>2</p>
<p>KAG4 - DSRM: Scientific and Research Methods group</p>	<p>Scientific/Research Methods Research methodology, paradigms and research cycle . Modelling and experiment planning Data selection and quality evaluation Use cases analysis: research infrastructures and projects . Research data management plan and ethical issues</p>	<p>5%</p>	<p>2</p>
<p>KAG6 - DSDK: Data Science Domain Knowledge group includes domain specific knowledge</p>	<p>Applied Data Science case cases in Domains, HR, Retail, Fraud Analytics, Finance Trends, Health Care, Infrastructure Management.</p>	<p>15%</p>	<p>2</p>

Copyright © IABAC 2015-2017.

IABAC®, IABAC logo®, DSF, DSP and CDS are all registered trade marks of IABAC Pvt Limited. Reproduction of this material requires the permission of IABAC

All rights reserved. EN_DSP_2017_CandidateSyllabus_V1.6

Bloom's Taxonomy Reference

Bloom's Learning Index	Description
1	Remembering: Recall or retrieve previous learned information.
2	Understanding: Comprehending the meaning, translation, interpolation, and interpretation of instructions and problems. State a problem in one's own words.
3	Applying: Use a concept in a new situation or unprompted use of an abstraction. Applies what was learned in the classroom into novel situations in the work place.
4	Analyzing: Separates material or concepts into component parts so that its organizational structure may be understood. Distinguishes between facts and inferences.
5	Evaluating: Make judgments about the value of ideas or materials.
6	Creating: Builds a structure or pattern from diverse elements. Put parts together to form a whole, with emphasis on creating a new meaning or structure.

Copyright © IABAC 2015-2017.

IABAC®, IABAC logo®, DSF, DSP and CDS are all registered trade marks of IABAC Pvt Limited. Reproduction of this material requires the permission of IABAC

All rights reserved. EN_DSP_2017_CandidateSyllabus_V1.6