



College of Contract Management  
United Kingdom

# Professional Diploma in Data Science



Syllabus

# Contents

## 1. Course Structure and Rules of Combination

## 2. Course Content

# 1. Course Structure and Rules of Combination

## 1.1 Rationale

This Professional Diploma in Data Science is the ideal course for learners who want to become a Data Analyst, Data Scientist, Director of Data Science or Chief Information Officer (CIO). Additionally, this course also provides learners with the ability to explore numerous industry specialities such as finance, health and social care and marketing. Through these lectures, regardless of what career path or industry you choose to work in, this Professional Diploma in Data Science will provide learners with the skills they need to succeed.

This Professional Diploma in Data Science is delivered via live online lectures. Our course lecturers have significant experience within the industry and will relate lecture content to real-life scenarios. In addition to this, lectures also include both practical examples and case studies. Through this delivery style, learners will be able to reflect on the practical challenges faced by professionals in the industry and establish an understanding of how to act in these situations in a manner that still works towards success.

## 1.2 Career Progression

This course will provide learners with the skills necessary to progress within the data science industry. Within the lectures, learners will increase their comprehension of both the theory and practical application of data science. Learners will gain the ability to analyse data, recognise key data structures and interpret findings. Within this course soft skills will be increased through general practice and hard skills will be established and developed through practical application.

## 1.3 Course Rules of Combination

The course can be completed in 7 months (approximately 28 weeks), and includes an assessment at the end of each module. Each module is worth 10 credits.

- DS701: Introduction to Data Science and AI
- DS702: Data Management and Preprocessing
- DS703: Practical Data Science

To achieve the Professional Diploma, candidates are required to complete all modules and pass their respective final assessments.

## 1.4 Entry Requirements

- Minimum 18 years of age **and**
- Relevant industry experience.
- An existing understanding in IT and Machine Learning

## 1.5 Module and Assessment Grades

The Assessor will award a grade for the achievement of each module (Fail, Pass, Merit or Distinction). Grades apply to overall performance in modules and assessments.

Indicative marking descriptors for differentiating between levels of achievement when marking assessments are provided below (Section 1.8).

The overall grade for a qualification is calculated using a points system. Each module grade attracts points as follows:

Fail	0 points
Pass	1 point
Merit	2 points
Distinction	3 points

## 1.6 Assessment

The assessment process is set by the College of Contract Management, defining the requirements learners are expected to meet in order to demonstrate that a learning outcome has been achieved. All learning outcomes must be achieved in order to gain attainment of credit for that module.

All completed assessments are marked and verified internally, and are subject to approval by our partner universities or awarding bodies.

The assessment criteria are based on 3 areas:

- 1. Task Achievement** - This is a measure of how well the candidate answers the task question(s) and identifies the important aspects of the task.
- 2. Technical Content** - This is a measure of how well the candidate identifies, describes and evaluates the technical aspects of the task.
- 3. Presentation** - This is a measure of how well the candidate presents the assessment, which includes the quality of the structure and paragraphing, the quality and relevance of visual or graphical content and the referencing used for quoted sources.

## 1.7 Assessment Policies

1. All submission of assessments must include:
  - a. a copy of the full brief given by the Examinations Officer or Course Administrator;
  - b. all source material must be cited in the text and a full bibliography of source material (including author, title, publisher, edition and page) listed at the end of the submission.
2. All submissions must be submitted into our system as instructed by the Examination Officer or Course Administrator.
3. All submissions under the student's name must only be the work of that student. All information sources must be acknowledged. There is the **possibility of failing the modules if the content of the assessment are deemed be plagiarised** as set out in the rules and regulations of the College.
4. All submissions should be in pdf format (unless software files are specified) and students must keep a copy of all submitted work for reference purposes. Receipt will be acknowledged by the College once the work is submitted via our online exam portal.
5. Whenever a candidate submits work after the approved deadline without an authorised extension, a maximum "Pass" grade will be awarded.
6. The Assessor will comment on the quality of the work for learning purposes.
7. Application for an extension must be requested prior to the submission deadline. Submissions must be made on the exam portal for each module extension request. A primary extension (two weeks) request can be made without the submission of any evidence or reasoning, any further extension requests will require submission of supporting documentation. All requests must be addressed to the Examination Officer or Course Administrator.

## 1.8 Indicative Marking Descriptors

**Note:** Please note that the bands below describe indicative characteristics only. An overall holistic approach is required when assessing a candidate's work and assigning a grade. Please read these grading bands in conjunction with the College of Contract Management Assignment Policy.

Grade	Task Achievement - The Relevance of the Response	Inclusion of Relevant Technical Knowledge in Content	Presentation/Coherence
<b>Distinction</b>			
70%+	The work demonstrates a comprehensive understanding of the task. All relevant information is included. The main issues are effectively identified and analysed. There is evaluation and some analysis of solutions to issues relevant to the task. The response shows control of content within the word count.	The work demonstrates a strong understanding of a wide range of technical issues relevant to the task. There is analysis of the advantages/disadvantages of possible choices, risks and potential outcomes.	The work is appropriately structured and the argument is developed coherently. There is a recognised form of source referencing which supports the points in the task. Paragraphing and titling are used effectively to assist the reader. The use of visual/graphical information is clear and effective in assisting the reader. The graphical information is relevant to the task and is accurate.
<b>Merit</b>			
60-69%	The work demonstrates a clear understanding of the main issues relevant to the task. The issues are explained effectively and potential solutions identified. There is some attempt to analyse the merits of the solutions to the task. The task is broadly achieved within the word count, if relevant to assignment.	The work demonstrates an understanding of the key technical issues of the task. There is clear description of relevant technical aspects with some attempt to evaluate the merits of these as appropriate to the task.	Demonstrates an awareness of presentation and an attempt to present the information with clarity and coherence. There is referencing of sources and use of paragraphing and titling to assist the reader. There is use of clear graphical information to support the assignment which has broad relevance to the task. There may be some limited inaccuracies/ omissions in these.
<b>Pass</b>			
40-59%	The work demonstrates an understanding of the task. The main points are identified and the task is achieved. There is no attempt to evaluate or analyse the solutions. There may be some inaccuracies, omissions and irrelevant content. There may be lack of control in relation to the word count.	The work demonstrates an understanding of the main technical issues which are identified. This may be limited to description with little evidence of evaluation. There may be some omissions and inaccuracies in the detail. There may be some irrelevant details.	There is an attempt to structure the information. There is evidence of paragraphing and titling which is not always appropriate. Some basic graphical information may be included which is of some assistance to the reader. There may be some omissions or inaccuracies. The work is generally coherent but there may be occasional lapses in coherence and structure.
<b>Fail</b>			
0-39%	The work shows a poor understanding of the task. Frequent inaccuracies. Failure to identify important aspects of the task. Much of the information is irrelevant to the task. There may be evidence of copy and paste from external sources. The response may be limited to lists of words with no attempt to explain the relevance/merits of these to the task. The assignment falls short of the word count.	The work demonstrates a lack of understanding of the technical aspects. There are omissions of important technical information. Errors are evident in the technical content. There is no attempt to explain the relevance of the technical content to the task.	Lacks structure and may be limited to lists of points which are not developed. Disorganised in structure causing difficulty for the reader to understand the points. The response is illegible or incoherent in places. No referencing of external sources. The graphical illustrations are of poor quality or absent. They may be irrelevant. There may be errors and a lack of clarity causing difficulty for the reader to understand.

## 1.9 Calculating Overall Qualification Grade

To calculate the overall qualification grade, the individual module grades should be added together and compared to the table below.

Candidates must pass all 3 modules of the course.

Total Points for all 3 Modules	Overall Grade
9	<b>Distinction</b>
8	
<b>Merit</b>	
7	<b>Merit</b>
6	
<b>Pass</b>	
5	<b>Pass</b>
4	
3	
<b>Fail</b>	
2 or fewer	<b>Fail</b>
Candidates must achieve at least a Pass in all 3 modules to be awarded the Professional Diploma.	

## 1.10 Mandatory Modules

Module Reference	Title	LH	Credit Value
DS701	Introduction to Data Science and AI	100	10
DS702	Data Management and Preprocessing	100	10
DS703	Practical Data Science	100	10

# DS701: Introduction to Data Science and AI

Learning outcomes: The learner will	Assessment criteria: The learner can
1. Understand the historical context of data science and artificial intelligence.	1.1 Describe the evolution and pivotal moments in the history of Data Science and AI.
2. Differentiate between data science and artificial intelligence.	2.1 Clearly distinguish between the core concepts, techniques, and objectives of Data Science and AI. 2.2 Identify areas of intersection and collaboration between the two fields.
3. Recognise data types and structures.	3.1 Classify various data types and explain their relevance in Data Science and AI projects. 3.2 Understand the significance of structured and unstructured data in different industry applications.
4. Appreciate real-world applications.	4.1 Enumerate key industries where Data Science and AI have made a significant impact. 4.2 Cite real-world case studies showcasing the transformative power of these technologies.
5. Identify key tools and technologies.	5.1 List and describe essential software, platforms, and technologies used in Data Science and AI. 5.2 Understand the significance of choosing the right tool for specific tasks.
6. Acknowledge industry barriers.	6.1 Understand the challenges industries face in adopting Data Science and AI. 6.2 Identify potential strategies to mitigate these challenges.

## Recommended Reading

1. Fawcett, T. and Provost, F. (2013) *Data Science for Business: What You Need to Know About Data Mining and Data-Analytical Thinking*. O'Reilly Media.
2. Russell, S. and Norvig, P. (2021) *Artificial Intelligence: A Modern Approach*. 4th ed. Pearson.
3. McKinney, W. (2022) *Python for Data Analysis*. 3rd ed. O'Reilly Media.
4. Burkov, A. (2022) *The Hundred-Page Machine Learning Book*. Andriy Burkov.
5. O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin.
6. Xu, Y. et al. (2021) 'Artificial Intelligence: A powerful paradigm for scientific research', *The Innovation*, 2(4). doi:10.1016/j.xinn.2021.100179.
7. Hassani, H. and Silva, E.S. (2023) 'The role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces are Revolutionizing the Field', *Big Data and Cognitive Computing*, 7(2), p. 62. doi:10.3390/bdcc7020062.



# DS702: Data Management and Preprocessing

Learning outcomes: The learner will	Assessment criteria: The learner can
1. Understand various data acquisition techniques.	1.1 Describe different sources of data, including databases, APIs, and web scraping. 1.2 Understand the challenges associated with gathering data from various sources. 1.3 Discuss the importance of data quality and its impact on analysis. 1.4 Evaluate the ethical considerations when acquiring data. 1.5 Understand Master data management and impact on organisations.
2. Master data cleaning and preprocessing techniques.	2.1 Identify common issues in datasets, such as missing values, outliers, and inconsistencies. 2.2 Apply various techniques to handle missing data. 2.3 Understand the importance of data consistency and standardisation. 2.4 Use tools and software to automate the data cleaning process.
3. Grasp the art of feature engineering.	3.1 Define what features are and their significance in data analysis. 3.2 Extract meaning features from raw data. 3.3 Transform features to enhance their predicative power. 3.4 Evaluate the importance of feature selection in model performance.
4. Understand data transformation techniques.	4.1 Differentiate between scaling, normalisation, and standardisation. 4.2 Apply various transformation techniques to prepare data for machine learning. 4.3 Evaluate the impact of data transformation on model performance.
5. Master data visualisation and Exploratory Data Analysis (EDA).	5.1 Describe the importance of EDA in understanding datasets. 5.2 Use various visualisation tools to explore data distributions, correlations, and patterns. 5.3 Interpret visualisations to derive meaningful insights from the data. 5.4 Apply statistical techniques to further understand datasets.
6. Data governance.	6.1 Define and explain the core principles and objectives of data governance. 6.2 Design and implement effective data governance strategies tailored to specific organisational needs. 6.3 Navigate the ethical considerations and regulatory requirements associated with data governance. 6.4 Recognise the role of data governance in the broader landscape of data management.

## Recommended Reading

1. Kazil, J. and Jarmul, K. (2016) *Data Wrangling with Python: Tips and Tools to Make Your Life Easier*. O'Reilly Media.
2. Briney, K. (2015) *Data Management for Researchers: Organize, maintain and share your data for research success*. Pelagic Publishing.
3. Garcia, S., Luengo, J. and Herrera, F. (2015) *Data Preprocessing in Data Mining*. Springer.
4. Zheng, A. and Casari, A. (2018) *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. O'Reilly.
5. Lin, J. and Dyer, C. (2010) *Data-Intensive Text Processing with MapReduce*. Morgan and Claypool Publishers.
6. Subramaniam, E., Kumar, B. and Bhowmik, R. (2010). *Data Acquisition Systems Current and future trends*. [online] Available at: <https://inspirehep.net/files/d3020f1715a9a12ff830899b75c9529a#:~:text=The%20near%20future%20also%20envisages>
7. Vogiatzis, D. (2024) *Your Guide to Data Transformation Techniques*, Coupler.io Blog. Available at: <https://blog.coupler.io/data-transformation-techniques/>

## DS703: Practical Data Science

Learning outcomes: The learner will	Assessment criteria: The learner can
1. Be able to analyse a dataset and understand the issues involved in using it to solve the chosen problem.	1.1 Understand the statistical properties of the variables and the correlations between them. 1.2 Identify any missing or corrupt data and apply appropriate remedies. 1.3 Identify any potential sources of bias or other ethical issues in the data and apply appropriate remedies. 1.4 Report on how these factors will affect the choice of model.
2. Be able to select an appropriate model for the problem, based on the properties of the dataset used and the business requirements.	2.1 Identify a set of potentially suitable candidate models. 2.2 Identify important hyperparameters of candidate models which may effect performance of models. 2.3 Identify a suitable set of metrics and other criteria with which to assess candidate models. 2.4 Identify the most suitable model and hyperparameters.
3. Be able to produce a working demonstration of the chosen model.	3.1 Identify best methods of presenting chosen model's outcomes to users. 3.2 Create a demonstration of the chosen system. 3.3 Implement a working model. 3.4 Report on the performance of the demonstration system and the steps necessary to develop a production model.