



## Data Management Course Descriptor

|                         |                           |                 |                               |
|-------------------------|---------------------------|-----------------|-------------------------------|
| Course Title            | Data Management           | Faculty         | EDGE Innovation Unit (London) |
| Course code             | NCHMV467                  | Course Leader   | Marta Rodriguez-Martinez      |
| Credit points           | 15                        | Teaching Period | 6 weeks                       |
| FHEQ level              | 4                         | Date approved   | March 2021                    |
| Compulsory/<br>Optional | Compulsory                |                 |                               |
| Prerequisites           | Data Science Fundamentals |                 |                               |

### COURSE SUMMARY

This course develops learners' knowledge and skills in data management. Learners will work within a Python environment to develop processes and code to extract, transform and load data through automated pipelines. Learners will study approaches to data integration, data migration and data synchronisation, in addition to testing strategies (Technical Acceptance Testing, User Acceptance Testing, Performance Stress testing). Learners will apply their knowledge through a written report investigating their organisation's data architecture.

### COURSE AIMS

- Train learners to apply ETL using Python to combine different datasets.
- Train learners to evaluate an organisation's data management.
- Train learners to apply testing strategies to ensure correct data integration, migration or synchronisation.

### LEARNING OUTCOMES

On successful completion of the course, learners will be able to:

#### KNOWLEDGE AND UNDERSTANDING

- K1a Explain the difference between data integration and data migration.
- K2a Describe best practice integration approaches using ETL techniques.
- K3a Describe different data testing strategies.

### SUBJECT SPECIFIC SKILLS

- S1a Manipulate data from three sources into a single dataset.
- S2a Identify relevant testing strategies to ensure that the ETL code produces correct unified data sets.
- S3a Investigate and describe the organisation's data architecture.

### TRANSFERABLE AND PROFESSIONAL SKILLS

- T1a Research a topic and present findings.
- T2a Make conceptual connections between different data sources.
- T3a Ability to think analytically and critically about methods of data integration.

### TEACHING AND LEARNING

The course learning and teaching hours will be structured as follows:

- Seminar = 15 hours
- Applied Seminar = 7.5 hours
- Coaching session = 1 hour
- Academic Drop-ins = 4 hours
- Reading and Assessment Prep = 32.5 hours
- Application of learning = 90 hours

Total = 150 hours

### ASSESSMENT

#### FORMATIVE

Learners will be formatively assessed during the course through a range of techniques. During seminars, applied learning seminars and one to one coaching, questioning and quizzing are used as formative assessment. The purpose of the formative assessment is to provide learners with developmental feedback to support their learning.

#### SUMMATIVE

| AE | Assessment Type    | Weighting | Online submission | Duration                                     | Length             |
|----|--------------------|-----------|-------------------|--|--------------------|
| 1  | Set exercises      | 75%       | Yes               | Requiring on average 25-35 hours to complete | 2500 words +/- 10% |
| 2  | Written assignment | 25%       | Yes               | -  | 1500 words +/- 10% |

The summative assessment will be assessed in accordance with the assessment aims set out in the Programme Specification.

## FEEDBACK

Learners will receive formal feedback in a variety of ways. For summative assessments, feedback is written and provided within 20 days of submission, and for formative assessment it could include verbal or written feedback. Learners will also have a progress review each quarter, see above, with their Multiverse Coach and their Line Manager. These reviews are an opportunity to discuss progress to date both in their role and on the degree programme.

## INDICATIVE READING

### BOOKS

Note: Comprehensive and current reading lists for courses are produced annually in the Course Syllabus or other documentation provided to learners; the indicative reading list provided below is used as part of the approval/modification process only.

Kimball, R. and Caserta, J., 2013. *The data warehouse ETL toolkit*. John Wiley & Sons.

Inmon, W.H., 2019, *Data Architecture: A Primer for the Data Scientist*, Elsevier Science.

## INDICATIVE TOPICS

Learners will study the following topics:

- ETL in Python
- Best practices for data integration and data migration
- Organisational data architecture
- Testing strategies applied to ETL

| <b>Title: NCHMV467 Data Management Course Descriptor</b> |               |                |  |                    |                           |
|--|---------------|----------------|--|--------------------|---------------------------|
| <b>Approved by: Academic Board</b>                       |               |                |  |                    |                           |
| Version number   | Date approved | Date published | Owner                                    | Location           | Proposed next review date |
| 1.0  | March 2021    | March 2021     | Marta Rodriguez-Martinez                 | Multiverse website | March 2026                |
| <b>Modifications (As per AQF4)</b>                       |               |                |  |                    |                           |
| Version number   | Date approved | Date published | Modification (including category number) |                    |                           |
|  |               |                |  |                    |                           |