



***Data Science with Python Foundation
Syllabus***



March 2020

1 Introduction

Data Science with Python qualifications are currently offered at the Foundation level.

This syllabus hence covers the Foundation level of examination. It is based on the Data Science with Python Foundation course first developed by GoDataDriven.

The course provides guidance on the principles and practice of loading, analysing, visualizing data with libraries such as pandas. It also teaches participants how predictive models work and how to use Scikit-learn can be used to train and fit models.

The training is divided into four knowledge chapters: Python and Jupyter basics, Data Analysis with Pandas and Matplotlib, Fundamentals of Machine Learning, and Practical Machine Learning with Scikit-learn.

2 Foundation Qualification

2.1 Purpose of the Foundation Qualification

The purpose of the Foundation qualification is to measure whether a candidate has sufficient knowledge and understanding of the fundamentals of Python, how to effectively use the pandas for data analysis, and the most important topics and concept of Machine Learning. The qualification also assesses if the candidate can translate the theoretical knowledge in code.

2.2 Target Audience

This qualification is aimed at data scientists who want to get a grip on the most popular and essential data science tools. These individuals require a working knowledge of the key principles of data science and are able to apply that on disparate data sets and business challenges.

2.3 High Level Performance Definition of a Successful Foundation Candidate

The candidate who meets this High Level Performance Definition should as a minimum be able to recall, recognize and demonstrate understanding of the theories, concepts, models and approaches outlined in the course.

Specifically (s)he should be able to demonstrate this understanding by being able to:

- Create interactive documents in Jupyter
- Organize Python project with attention to reproducibility and good coding practice
- Effectively broadcast Numpy objects
- Visualise data with Matplotlib and Pandas
- Perform data wrangling and aggregations in Pandas
- Organise code using pandas and Scikit-learn pipelines
- Use and understand side-effect free transformations
- Understand the most popular machine learning models
- Use scikit-learn transformers and estimators in pipelines to effectively perform grid search

3 Assessment Model

Each learning outcome in the High Level Performance Definition requires the candidate to demonstrate specific knowledge and skills. For each learning outcome a number of learning outcome measures are identified which are evaluated in the examination, in accordance with the Examination Design, to confirm that the learning outcome has been achieved. These learning outcome measures are shown as syllabus topics and define the scope of the standard required to achieve the qualification.

A classification widely used when designing assessments for certification and education is the Bloom's Taxonomy of Educational Objectives. This classifies learning objectives into six ascending learning levels, each defining a higher degree of competencies and skills. (Bloom et al, 1956, Taxonomy of Educational Objectives).

APMG have incorporated this into a Learning Outcomes Assessment Model that is then used to develop each qualification's Assessment Model. The model provides a simple and systematic means for assessing and classifying the learning outcome measures.

This structured approach helps to ensure:

- The appropriate level is identified for a qualification
- A clear delineation in learning level content between different qualifications
- Wording is standardized and syllabi are presented consistently across APMG's qualification portfolio
- Exam questions and papers are consistent in their design.

The Foundation qualification examines at levels 1 (recall) and 2 (understand). The Practitioner qualification tests at levels 2 (understand), 3 (apply) and 4 (analyse).

QUAL Assessment Model				
	1. Recall	2. Understand	3. Apply	4. Analyse
APMG Learning Level Definition	<i>remember previously learned information</i>	<i>grasp the meaning and make sense of information</i>	<i>use information to perform a skill or task</i>	<i>identify whether information has been used appropriately according to the rules and guidance</i>
Generic APMG Headers <i>For introducing the learning outcome measures (topics) in the Syllabus</i>	Recall terms and key facts about concepts, principles and procedures from the reference material	Understand key facts, concepts, principles and procedures from the reference material	Apply key facts, concepts, principles and procedures to a given scenario	Differentiate between appropriate and inappropriate use of the reference material in a given scenario
Qualification Example	Recall terms and key facts about definitions, techniques and tools relating to the syllabus area	Understand the definitions, techniques and their implementation to the syllabus area	Apply particular models or approaches relating to the syllabus area to a given scenario	Differentiate between appropriate and inappropriate use of particular approaches and models within the context of a given scenario

4 Qualification Scope

The definition of scope for each qualification is presented in the syllabus tables at the end of this document. Each syllabus area is a unit of learning that relates to the reference material or training course module.

The following syllabus areas are identified.

Syllabus Area Code	Syllabus Area Title
PY	Python and Jupyter basics
PD	Data analysis with Pandas and Matplotlib
ML	Fundamentals of Machine Learning
SK	Practical Machine Learning with SciKit-Learn

5 Syllabus Presentation

For each syllabus area the learning outcome measures are presented in order of learning level and are introduced by a standard header. There is only one header at each learning level for each syllabus area. The wording in this header is derived from the Assessment Model. Each measure is specific to a learning level.

The scope of each examination is shown by a tick in the respective column to the right of the topic description.

Practitioner qualification requirements are a summation of the Foundation and Practitioner learning outcome measures. All Foundation level requirements are required for Practitioner level but are assumed to have been met and are not directly assessed again, although Foundation level knowledge and understanding will be used when demonstrating Practitioner application and analysis learning outcomes.

Each of the syllabus areas is presented in a similar format as follows:

Syllabus Area Code		Syllabus Area : <i>QUAL Syllabus Area (XX) Theme</i> [1]	Foundation	Practitioner	Primary References
LC [2]					
Level	Topic				
Recall terms and key facts about the concepts, principles and procedures relating to <i>syllabus area</i>. [3] Specifically to recall:					
01 [4]	01 [5]	[6] The leadership styles and skills (Goleman): 1. Leadership styles that get results 2. Emotional competencies for leaders <i>Replace this with example from the qualification</i>	[7] ✓		[8] P175-178 P177-182
01	02				

Key to the Syllabus Area table

1	Syllabus Area	Unit of learning, e.g. course module, key activity area or section of the reference guide.
2	Syllabus Area Code	A unique 2 character code identifying the syllabus area.
3	Learning Level Header	Header introducing the syllabus topics (<i>learning outcome measures</i>) for a given learning level..
4	Level	Learning level of the learning outcome measure..
5	Topic Reference	Number of the topic within the learning level.
6	Topic Description (<i>Learning Outcome Measure</i>)	Precise and specific description of what is required of the candidate to demonstrate that a learning outcome has been achieved.
7	Foundation/Practitioner	Shows at which qualification level the topic is assessed . Note: A measure is only applied at one qualification level.
8	Primary Reference	The main reference supporting the learning outcome measure.

Syllabus Area Code		Syllabus Area :	Foundation	Practitioner	Primary References
PY		Python and Jupyter basics Theme			
Level	Topic				
Know facts, terms and concepts relating to the syllabus area.					
Specifically to recall:					
01	01	Python data types	Y		
01	02	Python functions	Y		
Understand the definitions, techniques and their implementation to the syllabus area.					
Specifically to identify how to:					
02	01	Index and slice	Y		
02	02	Manipulate strings	Y		
02	03	Manipulate lists	Y		
02	04	Write functions	Y		
02	05	Comprehensions	Y		

Syllabus Area Code		Syllabus Area :	Foundation	Practitioner	Primary References
PD		Data Analysis with Pandas and Matplotlib Theme			
Level	Topic				
Know facts, terms and concepts relating to the syllabus area.					
Specifically to recall:					
01	01	What is Pandas	Y		
01	02	What is Matplotlib	Y		
01	03	Matplotlib fundamentals	Y		
Understand the definitions, techniques and their implementation to the syllabus area.					
Specifically to identify how to:					
02	01	Slice and dice	Y		
02	02	Transform Data	Y		
02	03	Aggregate Data	Y		
02	04	Setup Pandas pipelines	Y		
02	05	Create figures	Y		
02	06	Plot data	Y		

Syllabus Area Code		Syllabus Area :	Foundation	Practitioner	Primary References
ML		Fundamentals of Machine Learning Theme			
Level	Topic				
Know facts, terms and concepts relating to the syllabus area.					
Specifically to recall:					
01	01	Machine Learning basic terminology	Y		
01	02	Types of learning	Y		
01	03	Supervised learning	Y		
01	04	Unsupervised learning	Y		
01	05	Cost functions	Y		
01	06	Optimization algorithms	Y		
01	07	Generalization power of models	Y		
01	08	Complexity of models	Y		

Syllabus Area Code		Syllabus Area :	Foundation	Practitioner	Primary References
ML		Fundamentals of Machine Learning Theme			
Understand the definitions, techniques and their implementation to the syllabus area.					
Specifically to identify how to:					
02	01	Choose the right algorithm	Y		
02	02	Model fitting practices	Y		

Syllabus Area Code		Syllabus Area :	Foundation	Practitioner	Primary References
SK		Practical Machine Learning with Scikit-Learn Theme			
Level	Topic				
Know facts, terms and concepts relating to the syllabus area.					
Specifically to recall:					
01	01	Estimators and Transformers	Y		
01	02	Pipelines	Y		
01	03	Hyperparameters selection	Y		
Understand the definitions, techniques and their implementation to the syllabus area.					
Specifically to identify how to:					
02	01	Model fitting with Estimators	Y		
02	02	Preprocessing data with Transformers	Y		
02	03	Hyperparameters tuning and cross-validation	Y		