Preparation Guide

Edition 202310

Content

1. Overview

EXIN Data Analytics Foundation (DAF.EN)

Scope

EXIN Data Analytics Foundation certification confirms that the professional understands the life cycle of collecting, cleaning, analyzing and visualizing data for the purpose of gaining business insights.

The certification covers:

- turning data into insights
- collecting, organizing, managing
- cleaning
- analyzing
- visualizing

Summary

Data analytics has its roots in statistics, which dates back to Ancient Egypt. Globalization and technical facilitation have recently led to an ever-increasing amount of data. Organizations of all sizes see the benefit of systematically gathering and using data to inform business decisions. The vital role of data is reflected by the increasing number of functions that are related to it such as data analysts, data scientists, data engineers and data managers.

Data analytics entails storing structured and unstructured data and transforming data by cleaning it. Data analytics also includes aggregating data and drawing insights from it. However, it is not only for specialists. Most business-oriented roles include requesting or creating reports. Subsequently, it is crucial to understand the basic principles of how data is turned into insights for anyone who is systematically processing information or wants to submit efficient requests for reports to inform the decision-making process.

The EXIN Data Analytics Foundation certification helps professionals understand the basic concepts of data analytics, the activities in the data analysis process, and the different methods used to analyze data. Candidates who pass the exam also know how outputs from the analysis process can be visualized in charts, graphs, and plots.

Context

The EXIN Data Analytics Foundation certification is part of the EXIN Data Analytics qualification program.

Target group

The EXIN Data Analytics Foundation certification is developed for professionals who are aspiring a role where analyzing data is part of their core responsibilities. More specifically, the following roles could be interested:

- data/information analysts
- business intelligence (BI) analysts
- data administrators/(business) information managers
- data/analytics managers
- data scientists

This certification is also a good starting point for professionals involved in (business) operations and management in any domain, who are interested in the business benefits of data analysis and the techniques involved in it. These roles could include:

- (digital) marketing/media specialists
- marketing/market research analysts
- business unit/department managers
- business analysts
- finance professionals

Requirements for certification

- Successful completion of the EXIN Data Analytics Foundation exam.

Examination details

Examination type:	Multiple-choice questions
Number of questions:	40
Pass mark:	65% (26/40 questions)
Open book:	No
Notes:	No
Electronic equipment/aides permitted:	No
Exam duration:	60 minutes

The Rules and Regulations for EXIN's examinations apply to this exam.

Bloom level

The EXIN Data Analytics Foundation certification tests candidates at Bloom level 1 and 2 according to Bloom's revised taxonomy:

- Bloom level 1: Remembering – relies on recall of information. Candidates will need to absorb, remember, recognize and recall.
- Bloom level 2: Understanding – a step beyond remembering. Understanding shows that candidates comprehend what is presented and can evaluate how the learning material may be applied in their own environment. This type of questions aims to demonstrate that the candidate is able to organize, compare, interpret and choose the correct description of facts and ideas.

Training

Contact hours

The recommended number of contact hours for this training course is 14. This includes group assignments, exam preparation and short breaks. This number of hours does not include lunch breaks, homework and the exam.

Indication study effort

56 hours (2 ECTS), depending on existing knowledge.

Training organization

You can find a list of our Accredited Training Organizations at www.exin.com.

2. Exam requirements

The exam requirements are specified in the exam specifications. The following table lists the topics of the module (exam requirements) and the subtopics (exam specifications).

Exam requirements	Exam specifications	Weight
1. Turning data into insights		**10%**
	1.1 Concepts and the process	7.5%
	1.2 Business intelligence (BI)	2.5%
2. Collecting, organizing, managing		**32.5%**
	2.1 Data collection	10%
	2.2 Storing data	12.5%
	2.3 Variables	10%
3. Cleaning		**7.5%**
	3.1 Data scrubbing	7.5%
4. Analyzing		**35%**
	4.1 Statistics	5%
	4.2 Data mining	2.5%
	4.3 Machine learning	10%
	4.4 Algorithms	17.5%
5. Visualizing		**15%**
	5.1 Explanatory and exploratory	2.5%
	5.2 Charts and plots	7.5%
	5.3 Design and tools	5%
	Total	**100%**

Exam specifications

1 Turning data into insights

1.1 Concepts and the process

The candidate can…

1.1.1 explain concepts related to data analytics.

1.1.2 outline the steps in the process of data analysis.

1.1.3 describe risks in data analytics.

1.2 Business intelligence (BI)

The candidate can…

1.2.1 illustrate how business intelligence (BI) leads to business decisions.

2 Collecting, organizing, managing

2.1 Data collection

The candidate can…

2.1.1 outline channels for collecting data.

2.1.2 explain ways to source public data.

2.1.3 outline how to reach compliance with contemporary data laws.

2.2 Storing data

The candidate can…

2.2.1 distinguish between raw data, structured data, unstructured data, and big data.

2.2.2 differentiate between types of databases.

2.2.3 outline the utility of distributed file systems.

2.2.4 describe the advantages of using cloud solutions.

2.3 Variables

The candidate can…

2.3.1 distinguish between dependent and independent variables.

2.3.2 explain the common types of variables.

2.3.3 distinguish between continuous and discrete variables.

3 Cleaning

3.1 Data scrubbing

The candidate can…

3.1.1 describe methods for data scrubbing.

3.1.2 outline techniques for data scrubbing.

3.1.3 indicate the considerations for data retention.

4 Analyzing

4.1 Statistics

The candidate can…

4.1.1 explain what distinguishes descriptive statistics from inferential statistics.

4.2 Data mining

The candidate can…

4.2.1 explain data mining.

4.3 Machine learning

The candidate can…

4.3.1 summarize ways of using machine learning.

4.3.2 describe natural language processing (NLP).

4.3.3 describe key methods and techniques of NLP.

4.4 Algorithms

The candidate can…

4.4.1 describe how data and algorithms relate to each other.
4.4.2 clarify how regression analysis can be used to describe patterns in the data.
4.4.3 explain the use of classification models.
4.4.4 explain the use of cluster analysis.
4.4.5 explain what association analysis and sequence mining are and how they relate to each other.

5 Visualizing

5.1 Explanatory and exploratory

The candidate can…

5.1.1 describe explanatory and exploratory graphics.

5.2 Charts and plots

The candidate can…

5.2.1 illustrate types of charts.
5.2.2 illustrate types of plots.
5.2.3 illustrate the usage of heatmaps.

5.3 Design and tools

The candidate can…

5.3.1 summarize the purpose of aesthetic design in data visualization.
5.3.2 recall visualization tools.

3. List of basic concepts

This chapter contains the terms and abbreviations with which candidates should be familiar.

Please note that knowledge of these terms alone does not suffice for the exam. The candidate must understand the concepts and be able to provide examples.

accuracy
aesthetic design
algorithm
analysis
anomaly
artificial intelligence (AI)
big data
binning
Boolean
business intelligence (BI)
chart
classification
cloud
clustering
crowdsourcing
data analysis
data analytics
data cleaning
data collection
data lake
data management
data mining
data procurement
data retention
data science
data security
data validation
data quality
database
decision tree
distribution
duplicate
frequency
graphic
hyperparameter
k-means clustering
key-value store
machine learning
mean
median
metadata
model
natural language processing (NLP)
nearest neighbors clustering
one-hot encoding
outlier
percentile
plot
predictive model
probability
programming language
quartile
regression
relational database
schema
sequence mining
statistics
third-party libraries
variable
web scraping

4. Literature

Exam literature

The knowledge required for the exam is covered in the following literature:

A. Oliver Theobald
Data Analytics for Absolute Beginners: A Deconstructed Guide to Data Literacy (Introduction to Data, Data Visualization, Business Intelligence & Machine Learning)
Independently published (2019)
ISBN: 9781081762469 (hardcopy)

B. EXIN
Exam Literature Data Analytics
EXIN (2023)
Go to www.exin.com. Click on 'Professionals' and then on 'Certifications' to find the certification. The free download can be found under 'Required reading'.

Additional literature

C. Elizabeth Clarke
Everything Data Analytics - A Beginner's Guide to Data Literacy: Understanding the Processes That Turn Data Into Insights
Kenneth Michael Fornari (2022)
ISBN: 9781777967130 (hardcopy)

Comment

Additional literature is for reference and depth of knowledge only.

Literature matrix

Exam requirements	Exam specifications	Reference
1. Turning data into insights		
	1.1 Concepts and the process	B
	1.2 Business intelligence (BI)	A, Chapter 13
2. Collecting, organizing, managing		
	2.1 Data collection	A, Chapter 1
	2.2 Storing data	A, Chapter 2
	2.3 Variables	A, Chapter 3
3. Cleaning		
	3.1 Data scrubbing	A, Chapter 4
4. Analyzing		
	4.1 Statistics	A, Chapter 5
	4.2 Data mining	A, Chapter 5
	4.3 Machine learning	A, Chapter 5 & 11
	4.4 Algorithms	A, Chapter 6, 7, 8, 9 & 10
5. Visualizing		
	5.1 Explanatory and exploratory	A, Chapter 12
	5.2 Charts and plots	A, Chapter 12
	5.3 Design and tools	A, Chapter 12

Contact EXIN

www.exin.com