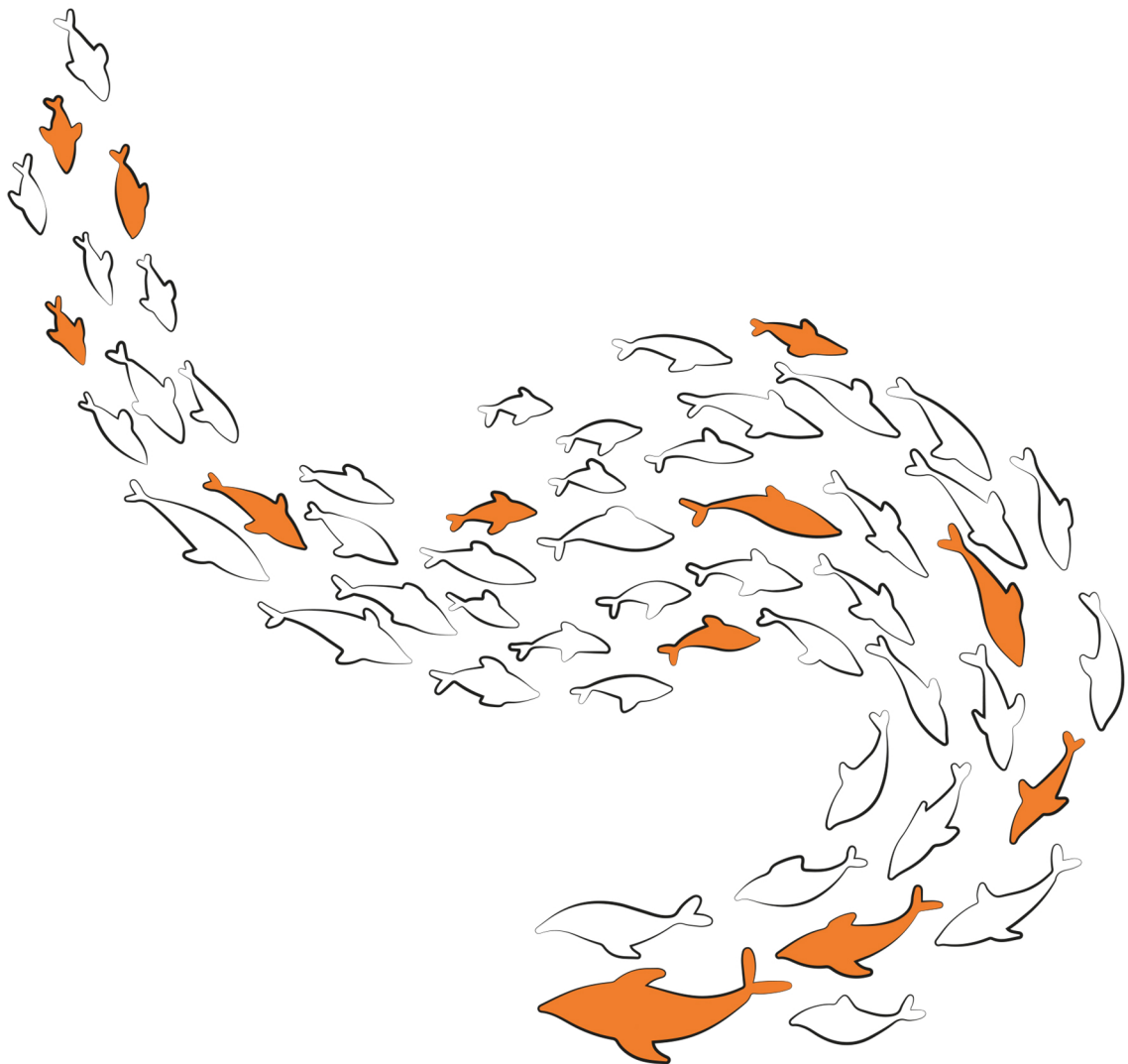


COURSEWARE

DATA MANAGEMENT FUNDAMENTALS (DMBOK BASED)

CDMP EXAM PREPARATION
(244 QUESTIONS + ANSWERS)



Data Management Fundamentals

(DMBOK based)

CDMP exam preparation

(244 questions + answers)

Colophon

Title: Data Management Fundamentals (DMBOK based)

Subtitle: CDMP exam preparation (244 questions + answers)

Authors: Paul Rakké

Publisher: Van Haren Publishing, 's-Hertogenbosch

ISBN Hard Copy: 978 94 018 1155 2

Edition: First edition, first print, May, 2024

Design: Van Haren Publishing, 's-Hertogenbosch

Copyright: © Van Haren Publishing 2024

For further information about Van Haren Publishing please e-mail

us at: info@vanharen.net or visit our website: www.vanharen.net

No part of this publication may be reproduced in any form by print, photo print, microfilm or any other means without written permission by the publisher.

Although this publication has been composed with much care, neither author, nor editor, nor publisher can accept any liability for damage caused by possible errors and/or incompleteness in this publication.

Publisher about the Courseware

The Courseware was created by experts from the industry who served as the author(s) for this publication. The input for the material is based on existing publications and the experience and expertise of the author(s). The material has been revised by trainers who also have experience working with the material. Close attention was also paid to the key learning points to ensure what needs to be mastered.

The objective of the courseware is to provide maximum support to the trainer and to the student, during his or her training. The material has a modular structure and according to the author(s) has the highest success rate should the student opt for examination. The Courseware is also accredited for this reason, wherever applicable.

In order to satisfy the requirements for accreditation the material must meet certain quality standards. The structure, the use of certain terms, diagrams and references are all part of this accreditation. Additionally, the material must be made available to each student in order to obtain full accreditation. To optimally support the trainer and the participant of the training assignments, practice exams and results are provided with the material.

Direct reference to advised literature is also regularly covered in the sheets so that students can find additional information concerning a particular topic. The decision to leave out notes pages from the Courseware was to encourage students to take notes throughout the material.

Although the courseware is complete, the possibility that the trainer deviates from the structure of the sheets or chooses to not refer to all the sheets or commands does exist. The student always has the possibility to cover these topics and go through them on their own time. It is recommended to follow the structure of the courseware and publications for maximum exam preparation.

The courseware and the recommended literature are the perfect combination to learn and understand the theory.

-- Van Haren Publishing

Other publications by Van Haren Publishing

Van Haren Publishing (VHP) specializes in titles on Best Practices, methods and standards within four domains:

- IT and IT Management
- Architecture (Enterprise and IT)
- Business Management and
- Project Management

Van Haren Publishing is also publishing on behalf of leading organizations and companies: ASLBiSL Foundation, BRMI, CA, Centre Henri Tudor, Gaming Works, IACCM, IAOP, IFDC, Innovation Value Institute, IPMA-NL, ITSqc, NAF, KNVI, PMI-NL, PON, The Open Group, The SOX Institute.

Topics are (per domain):

IT and IT Management

ABC of ICT
ASL®
CATS CM®
CMMI®
COBIT®
e-CF
ISO/IEC 20000
ISO/IEC 27001/27002
ISPL
IT4IT®
IT-CMF™
IT Service CMM
ITIL®
MOF
MSF
SABSA
SAF
SIAM™
TRIM
VeriSM™

Enterprise Architecture

ArchiMate®
GEA®
Novius Architectuur
Methode
TOGAF®

Business Management

BABOK® Guide
BiSL® and BiSL® Next
BRMBOK™
BTF
EFQM
eSCM
IACCM
ISA-95
ISO 9000/9001
OPBOK
SixSigma
SOX
SqEME®

Project Management

A4-Projectmanagement
DSDM/Atern
ICB / NCB
ISO 21500
MINCE®
M_o_R®
MSP®
P3O®
PMBOK® Guide
Praxis®
PRINCE2®

For the latest information on VHP publications, visit our website: www.vanharen.net.

About the author

Paul Rakké is an experienced independent consultant in the data management domain for over 30 years.

He held several positions from database manager, project manager to data consultant. The last years he acts more and more as a trainer. He teaches not only young students but also older employees in special trainingen on data management. DAMA-DMBOK framework is his specialism but he pays also attention to other frameworks and new developments in the fast changing data-driven, open society.

With focus on Data Governance and Big Data, but also on data ethics and GDPR related topics.

As a Certified Data Management Professional (CDMP) on the DAMA-DMBOK framework, he works for high schools and companies to accompany students and employees on the data journey, and helps with preparations for the CDMP-exam.

Not just from a technical perspective, but more from the perspective how to use data and with the message to be responsible when using data and to be ready for the transformation in a data-driven society.

Data Literacy is recently one of the extra focus points here and Paul integrated this with his vision on data management.

Table of Content

| | |
|--|-----------|
| About the author | 5 |
| Self-Reflection | 8 |
| Foreword | 11 |
| Introduction | 13 |
| The book | 13 |
| The questions | 14 |
| The exam | 15 |
| Part I: Sample questions per knowledge area/topic | 16 |
| Data Management (5) – Chapter 1 | 17 |
| Data Handling Ethics (5) – Chapter 2 | 18 |
| Data Governance (15) – Chapter 3 | 19 |
| Data Architecture (10) – Chapter 4 | 22 |
| Data Modeling and Design (12) – Chapter 5 | 24 |
| Data Storage and Operations (10) – Chapter 6 | 27 |
| Data Security (10) – Chapter 7 | 29 |
| Data Integration and Interoperability (10) – Chapter 8 | 31 |
| Document and Content Management (10) – Chapter 9 | 33 |
| Reference and Master Data (10) – Chapter 10 | 35 |
| Data Warehousing and Business Intelligence (10) – Chapter 11 | 38 |
| Metadata Management (13) – Chapter 12 | 40 |
| Data Quality Management (11) – Chapter 13 | 43 |
| Big Data & Data Science (5) – Chapter 14 | 45 |

| | |
|--|------------|
| Data Management Maturity Assessment (DMMA) (5) – Chapter 15 | 46 |
| Data Management Organization and Role Expectations (5) – Chapter 16..... | 47 |
| Part II: Sample questions per knowledge area/topic - Answers | 48 |
| Data Management (5) – Chapter 1 | 49 |
| Data Handling Ethics (5) – Chapter 2..... | 51 |
| Data Governance (15) – Chapter 3 | 53 |
| Data Architecture (10) – Chapter 4 | 58 |
| Data Modeling and Design (13) – Chapter 5..... | 61 |
| Data Storage and Operations (10) – Chapter 6..... | 65 |
| Data Security (10) – Chapter 7..... | 69 |
| Data Integration and Interoperability (10) – Chapter 8..... | 72 |
| Document and Content Management (10) – Chapter 9 | 75 |
| Reference and Master Data (10) – Chapter 10..... | 78 |
| Data Warehousing and Business Intelligence (10) – Chapter 11 | 82 |
| Metadata Management (13) – Chapter 12..... | 86 |
| Data Quality Management (11) – Chapter 13 | 90 |
| Big Data & Data Science (5) – Chapter 14..... | 93 |
| Data Management Maturity Assessment (DMMA) (5) – Chapter 15 | 95 |
| Data Management Organization and Role Expectations (2) – Chapter 16..... | 97 |
| Part III: Test exam Data Management Fundamentals (DMF)..... | 98 |
| Part IV: Test exam Data Management Fundamentals (DMF) - answers | 117 |

Self-Reflection of understanding Diagram

‘What you do not measure, you cannot control.’ – Tom Peters

Fill in this diagram to self-evaluate your understanding of the material. This is an evaluation of how well you know the material and how well you understand it. In order to pass the exam successfully you should be aiming to reach the higher end of Level 3. If you really want to become a pro, then you should be aiming for Level 4. Your overall level of understanding will naturally follow the learning curve. So, it’s important to keep track of where you are at each point of the training and address any areas of difficulty.

Based on where you are within the Self-Reflection of Understanding diagram you can evaluate the progress of your own training.

| <i>Level of Understanding</i> | <i>Before Training (Pre-knowledge)</i> | <i>Training Part 1 (1st Half)</i> | <i>Training Part 2 (2nd Half)</i> | <i>After studying / reading the book</i> | <i>After exercises and the Practice exam</i> |
|--|--|-----------------------------------|-----------------------------------|--|--|
| <i>Level 4 I can explain the content and apply it .</i> | | | | | |
| <i>Level 3 I get it! I am right where I am supposed to be.</i> | | | | | <i>Ready for the exam!</i> |
| <i>Level 2 I almost have it but could use more practice.</i> | | | | | |
| <i>Level 1 I am learning but don't quite get it yet.</i> | | | | | |

(Self-Reflection of Understanding Diagram)

Write down the problem areas that you are still having difficulty with so that you can consolidate them yourself, or with your trainer. After you have had a look at these, then you should evaluate to see if you now have a better understanding of where you actually are on the learning curve.

Troubleshooting

Problem areas:

Topic:

Part 1

Part 2

You have gone through the book and studied.

You have answered the questions and done the practice exam.

Foreword

Why should you certify as a Certified Data Management Professional? Who would be interested in your certification? Have you asked yourself these questions? I did when I originally encountered the Data Management Association (DAMA) International back in 2005. In these days I was working in the UK as a program manager on various projects. Delivering solutions in a complex multi-language, multi-cultural environment. In these international programs I was looking for the greatest common denominator: where can we find consensus on problem definitions, solution architectures and, ultimately, aligning the problem definition with the various proposed solutions. Theoretically schooled as a change manager, with a mathematically sound knowledge of business architecture and more than enough IT-knowledge supported by various certifications and hands-on experience. Why do we need Data Management? Why even certify in Data Management? I was once convinced that I had all the knowledge and skills to run these projects.

After running several programs I had noticed that most of the time I was solving the same issue over and over again: how do we create solutions that enable the business to collaborate across the business value chain. You would expect something with data here, wouldn't you? This is where it gets interesting. When discussing the value of the Data Management Body of Knowledge (DMBOK) I started to realize that collaboration should be supported by speaking the same language. That is exactly what DMBOK is about: speaking the same data language. And that is exactly the greatest common denominator between departments. Misunderstanding is not solved by setting up team meetings where all stakeholders refer to their own processes. Processes are not the language to cross bridges. I started to realize that data is the language that needs to be aligned. How do we communicate better between departments: we start by defining definitions, aka as data modeling; we define the context of the use of the definitions, aka metadata. From that point I adopted DMBOK and I have used it ever since in all my programs and projects.

In 2012, back in The Netherlands, I founded DAMA NL to tell all data management professionals what I had found in the UK. One of these people was Paul Rakké. Paul is one of those people where you immediately sense that he knows what he is talking about. With his profound experience in the data management arena, I found someone who had experienced a similar process to realize that data management was a missing linking pin in our project communication. We have started a journey to inform Dutch companies on the value of data and how to take advantage of that. This has resulted in tangible products for our clients and delivered us more insight in our own knowledge field: Data Management.

Paul and I would like to share our data management knowledge with you. I am now a lecturer at the Applied University of Utrecht. Paul has delivered practice exams to support you in becoming a Certified Data Management Professional. Where other practice exams excel at facts, Paul teaches you the context of the facts as well. His work dives deeper than the literal DMBOK. Paul encourages you to think, puzzle and explore the knowledge behind the book. He has succeeded in delivering a book that supports you in becoming a data management expert. The last step towards the certification is now only a formality.

More companies are embarking on the data journey to exploit the value of the data. Consequently, more data management professionals are needed to join this journey. For Paul and myself it was difficult to find professionals who are thinking alike, who speak the same data language. It would be much easier to look for a data management professional on LinkedIn. It would be nice to find professionals to share your knowledge and improve your skills. Certifying in Data Management makes you part of this data community. Therefore, we recommend that you get certified.

Why do you need to go for CDMP, the certification of DAMA-International? Peter Aiken, the president of DAMA-International when I was on the board, is best placed to answer this question. Here is his answer: "For over 30 years, DAMA has been the leading organization for data professionals by

developing a comprehensive body of data management standards and practices. Whether you are a data scientist, architect, engineer, integration specialist, or modeler, DAMA can enable your professional goals through internationally recognized, vendor-independent credentials.”

March 2024

Peter Vieveen
Lecturer Data Management
Applied University of Utrecht

Introduction

A lot is going on in the world of data. Not just because of dramatically increasing volumes, new enhanced technologies, discussions about regulations, (data) ethics and last, but not least, artificial intelligence and Chat GPT, but also because of a more demanding society and environment in the context of data.

Governments, organizations and individuals all play important yet changing roles. But to play these roles effectively, greater knowledge and changing attitudes are needed in order to meet the various requirements. Citizens and customers need to be serviced and supported by a much better, more controlled, and broader data landscape of higher quality in all its facets.

In this context the rise of digital literacy and data literacy with the objectives of leveling up capabilities and knowledge about data and its use for governments, organizations, and individuals is illustrative.

In this moving world of data, a structure is necessary to deal with the changing demands and requirements. That structure can be found by using a data management framework and acquiring knowledge to the appropriate level through following courses, reading books and get yourself certified to demonstrate your accomplishments and (re-)earn the trust about the use of data.

Be ready as a government, organization or individual and get yourself up to speed with knowledge about data and data management frameworks.

The book

This book has two objectives:

- To support you in learning more about the most used data management framework with its knowledge areas and associated topics all referred to in the DAMA DMBOK 2nd Edition textbook. By providing questions per topic, the reader can get a trigger to learn extra through using the textbook.
- To help you with the preparation for the Certified Data Management Professional (CDMP) exam. In this book you will find a complete set of newly created questions covering the topics. These are provided in a similar style to the actual exam, incorporating the different weights (different number of questions), to help you prepare yourself for the CDMP (Certified Data Management Professional) exam and to obtain a world-wide respected certificate.

Both, the sets of questions per topic (140 questions) and the complete test set of exam questions (100 questions), can help you in your preparation for your exam, or assist you in becoming more knowledgeable about this widely used data management framework.

Please note that these objectives can only be met when you have the textbook to hand are able to refer directly to it.

Note: The term 'topics' is used when following the chapters in the textbook, and not every chapter covers an already defined knowledge area.

The questions

The questions in this book are broken down into four parts:

- Part I: questions per topic, with the number of questions differing per topic dependent on the weight of the individual topic in the overall set of topics.
- Part II: answers and explanations of the questions from Part I.
- Part III: 100 questions, similar to a real exam, with similar weighting.
- Part IV: answers and explanations for these exam questions.

In the table below you can see how the questions are divided per topic:

| Chapter | Knowledge area/topic | # Questions per topic (Part I + Part II) | #Questions in the test exam (Part III+ Part IV) |
|---------|--|--|---|
| 1 | Data Management | 5 | 2 |
| 2 | Data Handling Ethics | 5 | 2 |
| 3 | Data Governance | 15 | 11 |
| 4 | Data Architecture | 10 | 6 |
| 5 | Data Modeling and Design | 13 | 11 |
| 6 | Data Storage and Operations | 10 | 6 |
| 7 | Data Security | 10 | 6 |
| 8 | Data Integration and Interoperability | 10 | 6 |
| 9 | Document and Content Management | 10 | 6 |
| 10 | Reference and Master Data | 10 | 10 |
| 11 | Data Warehousing and Business intelligence | 10 | 10 |
| 12 | Metadata Management | 13 | 11 |
| 13 | Data Quality Management | 11 | 11 |
| 14 | Big Data and Data Science | 5 | 2 |
| 15 | Data Management Maturity Assessment (DMMA) | 5 | N/A |
| 16 | Data Management Organization and Role Expectations | 2 | N/A |
| | Total | 144 | 100 |

The questions have the same possible options as in the real exam:

- Multiple choice.
- Five possible answers with just one correct answer.
- True/false questions.

The answers in part II and part IV refer to the DAMA DMBOK 2nd Edition textbook.

Note: 'Data Management Maturity Assessment' and 'Data Management Organization and Role Expectations' are separate chapters in the textbook. However, in the exam composition they are not separately mentioned, though a question is often included as part of the exam, for example as a Data Governance question.

The exam

As stated before, the exam consists of 100 questions conforming to the previously discussed weights per topic (in this situation we can use the term 'knowledge area' too).

- All the exams are open book exams, but keep in mind you are allowed just one book.
- You have 90 minutes to answer the 100 questions.
- If English is not your first language, you can purchase the ESL version which gives an additional 20 minutes.
- You will be informed immediately about the result.

Formally the exam is called the Data Management Fundamentals (DMF) exam and when you pass the exam you can get a CDMP certificate. Often the exam is also called the CDMP exam. The terms are used interchangeably, but they are not the same. In addition to the DMF exam, there are also specialist exams, and you must pass a specialist exam to become certified at a high-level.

More information about the exams and certificates can be found on several sites like <https://cmdp.info>, <https://www.dama.org> etc.

Part I:
Sample questions per
knowledge area/topic

Data Management (5) – Chapter 1

1. Which of the following statements is NOT a Data Management principle?
 - a) Data is an asset with unique properties.
 - b) Managing data means managing the quality of data.
 - c) It takes metadata to manage data.
 - d) Different types of data have different lifecycle characteristics.
 - e) A balance between strategic and operational needs is not required.

2. Value is the difference between the cost of a thing and the benefit derived from that thing, and for data is more complicated to calculate.
 - a) True
 - b) False

3. Which of the following statements is NOT correct in the context of the visuals of DAMA's Data Management Framework?
 - a) The DAMA Wheel defines the set of Data Management knowledge areas.
 - b) The environmental factors hexagon puts goals and principles in the center.
 - c) The Context Diagram presents the details within each knowledge area.
 - d) In the Data Management Framework, the relationship between the different knowledge areas is described.
 - e) The environmental factors hexagon shows the relationship between people, process and technology.

4. The data lifecycle is based on the product lifecycle.
 - a) True
 - b) False

5. Which of the following knowledge areas is described as 'the process of discovering, analyzing, representing, and communicating data requirements'?
 - a) Data governance.
 - b) Data architecture.
 - c) Document and content management.
 - d) Data modeling and design.
 - e) Data warehousing and business intelligence.

Data Handling Ethics (5) – Chapter 2

1. Which of the following is a core concept of Data Handling Ethics?
 - a) Impact on people.
 - b) Potential for misuse.
 - c) Economic value of data.
 - d) The data is expected to be complete and accurate.
 - e) All of the above.

2. Which of the following is a deliverable in the Data Handling Ethics context diagram?
 - a) Business strategy and goals.
 - b) Business culture.
 - c) Regulations.
 - d) Ethical data handling strategy.
 - e) Existing corporate policies.

3. Data minimization is described as 'Data must be kept in a form that permits identification of data subjects (individuals) for the purposes for which the personal data was processed.'
 - a) True
 - b) False

4. Which scenario of using data violates the ethical use of data?
 - a) Timing.
 - b) Misleading visualizations.
 - c) Unclear definitions or invalid comparisons.
 - d) Hunch and search.
 - e) All the above.

5. What does the ethical risk model in a sampling project NOT include?
 - a) Identification.
 - b) Behavior capture.
 - c) BI/Analytics/Data Science.
 - d) Data Governance.
 - e) Results.

Data Governance (15) – Chapter 3

1. Which of the following is an input in the Data Governance and Stewardship context diagram?
 - a) Data Governance scorecard.
 - b) Data maturity assessment.
 - c) Data strategy.
 - d) Business glossary.
 - e) Operating framework.

2. Which driver for Data Governance is focusing on improving processes (instead of reducing risks)?
 - a) Data security.
 - b) Privacy.
 - c) General risk management.
 - d) Regulatory compliance.
 - e) All of the above.

3. Which of the following statements describes a Data Governance Council?
 - a) The primary and highest authority organization for data governance in an organization.
 - b) Manages the data governance initiatives.
 - c) Ongoing focus on enterprise-level data definitions and data management standards across all knowledge areas.
 - d) Communities of interest focused on one or more specific subject areas or projects.
 - e) A departmental data governance council in large organizations.

4. The first activity to be performed in the implementation of Data Governance is?
 - a) Develop Data Governance strategy.
 - b) Perform readiness assessment.
 - c) Perform discovery and business alignment.
 - d) Define Data Governance for the organization.
 - e) Develop organizational touchpoints.

5. A data-centric organization values data as an asset and manages data through all phases of its lifecycle.
 - a) True
 - b) False

6. What is a synonym for a trustee?
- Data privacy officer.
 - Data steward.
 - Chief Information Security Officer (CISO).
 - Stakeholder.
 - Data owner.
7. Which of the following statements is NOT correct in the context of Data Governance?
- Data policies support standards.
 - Data policies vary widely across organizations.
 - Data policies describe the 'what'.
 - Data policies describe the 'how'.
 - Data policies are directives that codify principles and management intent into fundamental rules.
8. Which of the following type of data steward is described as 'having oversight of a data domain across business functions'?
- Enterprise data steward.
 - Chief data steward.
 - Executive data steward.
 - Business steward.
 - Technical data steward.
9. Which of the following is NOT included when performing a readiness assessment?
- Capacity to change.
 - Monitoring controls that detect unexpected events.
 - Data management maturity.
 - Collaborative readiness.
 - Business alignment.
10. A business case must be adjusted periodically to reflect the changing priorities and financial realities of the organization.
- True
 - False
11. Which principle for data asset accounting is described as 'value the data as an asset at a level that makes the most sense, or is the easiest to measure'?
- Value Principle.
 - Level of Valuation Principle.
 - Risk Principle.
 - Going Concern Principle.
 - Quality Principle.

12. Which of the following is NOT an example of touch points that support alignment and cohesiveness of enterprise data governance in areas outside the direct authority of the Chief Data Officer?
- a) Procurement and contracts.
 - b) Regulatory compliance.
 - c) Budget and funding.
 - d) System development lifecycle framework.
 - e) Business alignment assessment.
13. In a change management process, communications is key. Which of the following is NOT related to communications in this context?
- a) Implementing data management training.
 - b) Promoting the value of data assets.
 - c) Documenting stakeholder viewpoints and resolution alternatives.
 - d) Monitoring and acting on feedback about data governance activities.
 - e) Implementing new metrics and KPIs for realignment to support behaviors.
14. Which of the following is NOT a way to measure the value of data for an organization?
- a) Replacement cost.
 - b) Risk cost.
 - c) Selling data.
 - d) Market value.
 - e) Total cost of ownership.
15. When constructing an organization's operating model, the impact of regulation is not relevant.
- a) True
 - b) False

Data Architecture (10) – Chapter 4

1. Which column in the Zachman Framework refers to 'entities used to build the architecture'?
 - a) Where
 - b) Who
 - c) What
 - d) When
 - e) How

2. Data and enterprise architecture deal with complexity from two viewpoints: quantity-oriented and innovation-oriented.
 - a) True
 - b) False

3. Which of the following is NOT an Enterprise Data Architecture project-related activity?
 - a) Define scope.
 - b) Choose methodology.
 - c) Implement.
 - d) Design.
 - e) Understand business requirements.

4. Architecture designs can be aspirational or future-looking, implemented and active, or plans for retirement. Each topic must be well-documented. Which of the following refers to 'products expected to be available for use in the next 2+ years'?
 - a) Current
 - b) Deployment period
 - c) Strategic
 - d) Retirement
 - e) Containment

5. Which of the following development methodologies is described by 'learn and construct in gradual steps, creates a prototype based on vague overall requirements'?
 - a) Waterfall
 - b) Agile
 - c) DevOps
 - d) Lean
 - e) Incremental

6. Data architecture artifacts are usually captured within development projects, and then standardized and managed by data architects.
- a) True
 - b) False
7. Which of the following is NOT a Data Architecture governance activity?
- a) Overseeing projects.
 - b) Managing architectural designs, lifecycle and tools.
 - c) Defining standards.
 - d) Managing the availability of data throughout the data lifecycle.
 - e) Creating data-related artifacts.
8. Which of the following is an input in the data architecture context diagram?
- a) Implementation roadmap.
 - b) Data flows.
 - c) Enterprise data model.
 - d) Data value chains.
 - e) Enterprise Architecture.
9. Which of the following is NOT an architecture domain?
- a) Elements
 - b) Dependencies
 - c) Reliability
 - d) Roles
 - e) Purpose
10. Identify data storage and processing is one of the goals of data architecture.
- a) True
 - b) False

Data Modeling and Design (12) – Chapter 5

1. Which of the following is NOT a deliverable of the data modeling process?
 - a) Definitions.
 - b) Initial data requirements.
 - c) Diagram.
 - d) Issues and outstanding questions.
 - e) Lineage.

2. Forward engineering is the process of building a new application beginning with the requirements.
 - a) True
 - b) False

3. Which of the following is NOT a designing principle for a DBA in database design?
 - a) Resiliency
 - b) Performance and ease of use
 - c) Integrity
 - d) Security
 - e) Maintainability

4. In data modeling, what is the purpose of a surrogate key?
 - a) It assures that a value in a specific column in a table is unique.
 - b) It establishes relationships between tables/entities.
 - c) It is an additional key that can be used in addition to the primary key to locate specific data.
 - d) It serves a technical function in a table, not visible to end users.
 - e) It identifies the entity instance to which it belongs.

5. Data model validation can be done by using a Data Model Scorecard.
 - a) True
 - b) False

6. From which entity category are order, complaint, and return examples?
 - a) What
 - b) When
 - c) Where
 - d) Why
 - e) Measurement

7. Which of the following is NOT a characteristic of the data vault?
- a) Detail-oriented.
 - b) Uniquely linked set of normalized tables.
 - c) Time-based.
 - d) Best of breed between third normal form and star schema.
 - e) Low scalability.
8. The acronym SDLC stands for System Development Life Cycle.
- a) True
 - b) False
9. Which of the following normal forms is described by 'ensures each entity has the minimal primary key and that every attribute depends on the complete primary'?
- a) First normal form (1NF).
 - b) Second normal form (2NF).
 - c) Third normal form (3NF).
 - d) Fourth normal form (4NF).
 - e) Fifth normal form (5NF).
10. In supertype partition the supertype entity's attributes are included as nullable columns into a table.
- a) True
 - b) False
11. Which of the following is the first step when creating a conceptual data model (CDM)?
- a) Select notation.
 - b) Obtain sign-off.
 - c) Select scheme.
 - d) Complete initial corporate data.
 - e) Incorporate enterprise terminology.
12. Which of the following is NOT a notation form for a relational modeling scheme?
- a) Information Engineering (IE).
 - b) Integration Definition for Information Modeling (IDFIX).
 - c) Barker Notation.
 - d) Chen.
 - e) UML.

13. A compound key a composite key for which each attribute that makes up the key is a foreign key in its own right.
- a) True
 - b) False

Data Storage and Operations (10) – Chapter 6

1. Which of the following is a deliverable in the data storage and operations context diagram?
 - a) Data architecture.
 - b) Database technology evaluation criteria.
 - c) Data models.
 - d) Service level agreements.
 - e) Data requirements.

2. Which of the following statements concerning spatial databases is NOT correct?
 - a) A spatial database is optimized to store and query data that represents objects defined in a geometric space.
 - b) Spatial database systems use spatial indexes to quickly look up values and speed up database operations.
 - c) A spatial database can perform one or more spatial operations.
 - d) In a spatial database, spatial predicates allow true/false queries about spatial relationships between geometries.
 - e) Spatial database systems use key-value pairs.

3. Data retention planning should be part of the physical database design.
 - a) True
 - b) False

4. Which of the following database access controls is NOT part of DBAs primary responsibility for controls that enable access to the data.
 - a) Controlled environment for data assets that includes network roles and permissions management.
 - b) Designing user entitlement profiles that allow the business to function smoothly while following relevant restrictions.
 - c) Physical security of data assets and monitoring the defined protocols.
 - d) Continuous monitoring of the critical database servers.
 - e) Maintaining information security by access controls, database auditing etc.

5. In the context of Change Data Capture (CDC), data versioning, which evaluates columns that identify rows that have changed, is the only method to detect and collect changes.
 - a) True
 - b) False

6. Which of the following recovery types is described by 'restore the system as quickly as possible in order to minimize delays or shut downs'?
- a) Immediate recovery.
 - b) Critical recovery.
 - c) Non-critical recovery.
 - d) Instant data recovery.
 - e) None of the above.
7. Which of the following database terms is described as 'an execution of database software controlling access to a certain area of storage'?
- a) Database
 - b) Node
 - c) Database abstraction
 - d) Instance
 - e) Schema
8. Federated databases are best for heterogeneous and distributed integration projects.
- a) True
 - b) False
9. ACID is the diametric opposite of BASE-type systems. Which of the following is NOT a characteristic of a BASE-type system?
- a) Store dissimilar data.
 - b) Key-value stores.
 - c) Row/column.
 - d) Open-source.
 - e) Automatically spreads data across commodity servers (scaling).
10. Which of the following test environments is used to test functionality against requirements?
- a) Quality Assurance Testing (QA).
 - b) Integration testing.
 - c) User acceptance testing.
 - d) Performance testing.
 - e) Non-functional testing.

Data Security (10) – Chapter 7

1. Data Stewards are responsible for evaluating and determining the appropriate confidentiality level for data based on the organization's classification scheme.
 - a) True
 - b) False

2. Lack of automated monitoring represents serious risks. Which of the following is an example of such a risk?
 - a) Regulatory risk.
 - b) Risk of reliance on inadequate native audit tools.
 - c) Detection and recovery risk.
 - d) Administrative and audit duties risk.
 - e) All of the above.

3. Which of the following masking methods is described by 'remove data that should not be present in a test system'?
 - a) Substitution
 - b) Nulling or deleting
 - c) Encryption
 - d) Shuffling
 - e) Key masking

4. In penetration testing ('pen testing'), an ethical hacker, always hired from outside the organization, attempts to break into the system from outside, acting as if he/she was a malicious hacker.
 - a) True
 - b) False

5. Which of the following in the context of data security is described as 'the first line of defense against bad actors'?
 - a) Device security.
 - b) Facility security.
 - c) Credential security.
 - d) User ID standards for email systems.
 - e) Electronic communication security.

6. RACI matrices help clarify roles, the separation of duties, and responsibilities of different roles, excluding their security obligations.
 - a) True
 - b) False

7. Which of the following is a deliverable in the data security context diagram?
- a) Data security architecture.
 - b) Regulatory requirements.
 - c) Enterprise architecture standards.
 - d) Enterprise data model.
 - e) Business rules and processes.
8. Which of the following is NOT a processing step in the process of risk reduction?
- a) Identify and classify sensitive data assets.
 - b) Locate sensitive data throughout the enterprise.
 - c) Mask data to conform the requirements.
 - d) Determine how each asset needs to be protected.
 - e) Identify how this information interacts with business owners.
9. Which of the following is NOT a guiding principle for data security in an organization?
- a) Reduce risk by reducing exposure.
 - b) Collaboration.
 - c) Enterprise approach.
 - d) Clear accountability.
 - e) Build with reuse in mind.
10. Which of the following is NOT a confidentiality classification level?
- a) Internal use only.
 - b) Regulated use.
 - c) Confidential.
 - d) Restricted confidential.
 - e) Registered confidential.

Data Integration and Interoperability (10) – Chapter 8

1. Which of the following is NOT a part of Data Integration metrics?
 - a) Availability
 - b) Volume
 - c) Speed
 - d) Value
 - e) Costs

2. Which of the following statements is correct?
 - a) Two systems that are tightly coupled usually have an asynchronous interface.
 - b) Two systems that are loosely coupled usually have a synchronous interface.
 - c) When systems are tightly coupled, the availability is more assured than when systems are loosely coupled.
 - d) Where the systems are loosely coupled, replacement of systems in the application inventory can theoretically be performed without rewriting the systems with which they interact.
 - e) Service Oriented Architecture using an Enterprise Service Bus is an example of a tightly coupled data interaction design pattern.

3. Since most integration projects involve a significant amount of data, the most efficient means of conducting analyses is to use a data profiling tool.
 - a) True
 - b) False

4. Which of the following is NOT an example of transformation as part of the ETL process?
 - a) Format change.
 - b) Structure change.
 - c) Semantic conversion.
 - d) Sharding.
 - e) Re-ordering.

5. Which of the following is NOT a governance-driven management routine in the context of data integration and interoperability?
 - a) Mandated review of models.
 - b) Perform readiness assessment.
 - c) Auditing of metadata.
 - d) Gating of deliverables.
 - e) Required approvals for changes to transformation rules.

6. Data sharing agreements stipulate the responsibilities and acceptable use of data to be exchanged.
- a) True
 - b) False
7. Which of the following is an input in the Data Integration and Interoperability context diagram?
- a) Data semantics.
 - b) Thresholds and alerts.
 - c) Data exchange specifications.
 - d) Data services.
 - e) Data access agreements.
8. Batch data integration orchestration is usually triggered by an event.
- a) True
 - b) False
9. Which of the following is NOT a Data Integration and Interoperability concept?
- a) Enterprise service bus (ESB).
 - b) Enterprise application integration (EAI).
 - c) Service-based architecture (SBA).
 - d) Service-oriented architecture (SOA).
 - e) Complex event processing (CEP).
10. A delta is a set of data changed from a point of time.
- a) True
 - b) False

Document and Content Management (10) – Chapter 9

1. Which of the following statements is correct?
 - a) A flat taxonomy has relationships among the set of controlled categories.
 - b) A hierarchical taxonomy is a tree structure where nodes are related by a rule.
 - c) A facet taxonomy looks like a snowflake where each node is associated with another node.
 - d) A network taxonomy uses just facet structures.
 - e) None of the above.

2. Which characteristic is correct concerning well-prepared records?
 - a) Content
 - b) Context
 - c) Timeliness
 - d) Structure
 - e) All of the above

3. EDRM stands for Enterprise Discovery Reference Model.
 - a) True
 - b) False

4. What is the first phase in the EDRM process of e-discovery?
 - a) Identification
 - b) Processing
 - c) Presentation
 - d) Production
 - e) Preservation & collection

5. The Information Architecture and the content strategy together describe the 'what' (what content) will be managed in a system.
 - a) True
 - b) False

6. Which of the following is NOT an activity related to content lifecycle management?
 - a) Backup and recovery.
 - b) Publish and deliver content.
 - c) Manage versioning and control.
 - d) Capture records and content.
 - e) Manage retention and control.

7. Which of the following statements is correct?
- a) XML provides a language for representing both structured and unstructured data and information.
 - b) RDF (Resource Description Framework) uses JSON as its encoding syntax.
 - c) JSON is an open, light weight standard format with a language-dependent text format.
 - d) Unstructured data cannot be stored in a relational DBMS BLOB or in XML.
 - e) Either XML or JSON content cannot be returned using REST technology.
8. Which of the following is NOT an information governance maturity level as defined by ARMA Int. along with GARP?
- a) Sub-standard
 - b) In development
 - c) Managed
 - d) Proactive
 - e) Transformational
9. Which of the following is a deliverable in the document and the content management context diagram?
- a) Content repository.
 - b) Social media stream.
 - c) IT strategy.
 - d) Legal retention requirements.
 - e) Business strategy.
10. Which of the following is described by 'provides data content classifications for a given concept area'?
- a) Ontology
 - b) Thesaurus
 - c) Data model
 - d) Syntax
 - e) Taxonomy

Reference and Master Data (10) – Chapter 10

1. Which of the following activities is a basic step in planning for master data management?
 - a) Identify candidate sources that will provide a comprehensive view of the Master Data entities.
 - b) Develop rules for accurate matching and merging of entity instances.
 - c) Establish an approach to identify and restore inappropriate matched and merged data.
 - d) Establish an approach to distribute trusted data to systems across the enterprise.
 - e) All of the above.

2. Which of the following statements concerning entity resolution is NOT correct?
 - a) Entity resolution is the process of determining whether two references to real world objects refer to the same objects or to different objects.
 - b) Entity resolution refers to identifying and eliminating duplicate or redundant records within a certain dataset.
 - c) Entity resolution is a decision-making process.
 - d) Entity resolution is critical to MDM, as the process of matching and merging records enables the construction of the master data set.
 - e) Entity resolution includes a set of activities that enable the identity of entity instances and the relationship between entity instances.

3. Which of the following refers to the final phase (activity) in the reference data change request process?
 - a) Receive change request.
 - b) Identify impact.
 - c) Identify stakeholder.
 - d) Update and inform (if applicable).
 - e) Decide and communicate.

4. Two references that do not represent the same entity and are linked with a single identifier are called false-negatives.
 - a) True
 - b) False