DP-700 Microsoft Fabric Data Engineer Practice Exam (Sample)

Study Guide

BY EXAMZIFY

Everything you need from our exam experts!

Table of Contents

Introduction

Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.

This guide was designed to help you do exactly that.

Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.

The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:

- **Practice answering questions under realistic conditions,**
- **Improve accuracy and speed,**
- **Review explanations to strengthen weak areas, and**
- **Approach the exam with greater confidence.**

We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.

Remember: successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.

Let's begin.

How to Use This Guide

This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here's how to get the most out of your Examzify study guide:

1. Start with a Diagnostic Review

Skim through the questions to get a sense of what you know and what you need to focus on. Your goal is to identify knowledge gaps early.

2. Study in Short, Focused Sessions

Break your study time into manageable blocks (e.g. 30 – 45 minutes). Review a handful of questions, reflect on the explanations.

3. Learn from the Explanations

After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.

4. Track Your Progress

Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.

5. Simulate the Real Exam

Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.

6. Repeat and Review

Don't just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning. Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.

There's no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly, adapt the tips above to fit your pace and learning style. You've got this!

Questions

1. Which option best describes the Lakehouse data store in Fabric?

A. Optimized for unstructured data only

B. Combines data warehousing and data lake capabilities

C. Stores only raw CSV

D. Is the same as OneLake

2. Which Fabric experience is used to move and transform data?

A. Data Warehousing

B. Data Science

C. Data Factory

D. Data Integration

3. In Fabric Real-Time Intelligence, where can ingested data reside for subsequent analysis?

A. Data lake only

B. On-premises data store

C. Lakehouse or Eventhouse

D. SQL data warehouse only

4. What is data lineage, and how is it captured in Fabric?

A. Data lineage is the process of archiving datasets for compliance.

B. Data lineage tracks data origin and movement within source systems only.

C. Data lineage tracks data origin, movement, and transformations across pipelines and lakehouse; Fabric captures lineage via pipeline definitions, dataflows, and table metadata.

D. Data lineage is only stored in notebooks.

5. You want to include data in an external Azure Data Lake Store Gen2 location in your lakehouse, without the requirement to copy the data. What should you do?

A. A Data pipeline that uses a Copy Data activity to load the external data into a file

B. Create a Dataflow Gen2 that extracts the data and loads it into a table

C. Create an external table

D. Create a shortcut

6. What is a data pipeline?

A. A sequence of activities to orchestrate a data ingestion or transformation process

B. A special folder in OneLake storage where data can be exported from a lakehouse

C. A saved Power Query

D. A dataset for BI reporting

7. What data retention strategies are typical in Fabric?

A. Define retention policies per dataset, archive or purge stale data, support point-in-time restore, and implement lifecycle management to balance cost and compliance.

B. Retain all data forever with no policy.

C. Purge everything after a fixed period regardless of usage.

D. Retention is not configurable in Fabric.

8. How would you describe a lakehouse table stored on OneLake enabling BI-friendly SQL with ACID transactions?

A. A NoSQL key-value store

B. A relational database managed separately

C. A CSV file

D. A Delta Lake table on OneLake enabling BI-friendly SQL queries with ACID transactions

9. What do temporal window transformations enable you to do?

A. Automatically enable and disable streaming at scheduled times.

B. Aggregate event data in a stream based on specific time periods.

C. Delete expired data after a specified retention period.

D. Filter events by a static property value.

10. How does Fabric's data catalog facilitate data discovery?

A. It indexes datasets, schemas, lineage, access policies, and usage metadata, enabling search, discovery, and governance across assets.

B. It only indexes datasets.

C. It indexes external data sources but not lineage.

D. It stores metadata but does not support search.

Answers

1. B
2. C
3. C
4. C
5. D
6. A
7. A
8. D
9. B
10. A

Explanations

1. Which option best describes the Lakehouse data store in Fabric?

A. Optimized for unstructured data only

B. Combines data warehousing and data lake capabilities

C. Stores only raw CSV

D. Is the same as OneLake

The Lakehouse in Fabric is designed to merge the strengths of data lakes and data warehouses. It provides scalable storage for diverse data types—structured, semi-structured, and unstructured—while also offering SQL-based querying, metadata management, schema enforcement, and ACID transactions for reliable analytics. This combination lets you store raw data and curated, optimized data in one place and analyze it efficiently with governance features. So, it best describes a system that combines data warehousing and data lake capabilities. It's not limited to unstructured data, not restricted to raw CSV files, and it isn't just the same as OneLake—the Lakehouse sits on top of storage like OneLake to enable analytics.

2. Which Fabric experience is used to move and transform data?

A. Data Warehousing

B. Data Science

C. Data Factory

D. Data Integration

This question centers on how data is moved and transformed within Fabric. In Fabric, the Data Factory experience is built for data movement and transformation. It lets you create pipelines that copy data from various sources to destinations and apply transformations through data flows as part of the workflow, effectively handling ETL/ELT tasks. Other options describe different aspects: Data Warehousing focuses on storing and organizing data for analysis; Data Science is about building models and extracting insights; Data Integration is a broader concept of merging data from multiple sources, but the explicit tool designed for orchestrating movement and transformation in Fabric is Data Factory.

3. In Fabric Real-Time Intelligence, where can ingested data reside for subsequent analysis?

A. Data lake only

B. On-premises data store

C. Lakehouse or Eventhouse

D. SQL data warehouse only

In Fabric Real-Time Intelligence, ingested data isn't limited to a single storage type—you can land it in either a Lakehouse or an Eventhouse for analysis. A Lakehouse combines the scalability and breadth of a data lake with the structured query and governance features of a data warehouse, making it suitable for both raw and curated data used in analytics. An Eventhouse is tailored for streaming, low-latency event data, enabling real-time processing and immediate insights. This flexibility lets you choose the storage target that best fits your workflow: long-term analytics on diverse data in a lakehouse, or rapid real-time analysis on streaming data in an eventhouse. The other options are too narrow: a data lake alone doesn't capture the combined analytics capabilities of a lakehouse, on-premises storage isn't the cloud-native model for Fabric, and a SQL data warehouse only doesn't cover the streaming/event use case.

4. What is data lineage, and how is it captured in Fabric?

A. Data lineage is the process of archiving datasets for compliance.

B. Data lineage tracks data origin and movement within source systems only.

C. Data lineage tracks data origin, movement, and transformations across pipelines and lakehouse; Fabric captures lineage via pipeline definitions, dataflows, and table metadata.

D. Data lineage is only stored in notebooks.

Data lineage is the end-to-end visibility of where data comes from, how it moves through systems, and how it is transformed along the way, all the way from source to its final destination in the lakehouse. In Fabric, this lineage is captured by tying together the facts from pipeline definitions, dataflows, and the metadata of tables. That means you can trace a piece of data from its source, through each transformation or step in a pipeline, to where it resides in a table, with the transformations clearly reflected in the lineage graph. This end-to-end view supports impact analysis, governance, and auditing, since you can see not just where data originated but every point it passed through and how it was changed. The other ideas don't fit because archiving datasets is about storing data for retention, not tracking its origin and movement. Limiting lineage to the source systems misses the downstream steps and transformations that occur after ingestion. And storing lineage only in notebooks isn't how Fabric organizes or surfaces end-to-end provenance; lineage is captured in pipeline definitions, dataflows, and table metadata.

5. You want to include data in an external Azure Data Lake Store Gen2 location in your lakehouse, without the requirement to copy the data. What should you do?

A. A Data pipeline that uses a Copy Data activity to load the external data into a file

B. Create a Dataflow Gen2 that extracts the data and loads it into a table

C. Create an external table

D. Create a shortcut

The concept being tested is bringing external data into a lakehouse without duplicating it by using a lightweight pointer. Creating a shortcut to the Azure Data Lake Storage Gen2 location serves exactly this need: it provides a metadata reference to data that already lives in ADLS Gen2, so the lakehouse can access and query it directly without copying the data into the lakehouse storage. The data remains in its original location, preserving storage costs and avoiding duplication, while still making it available for analysis through the lakehouse interface and governance surface. Other options involve moving or duplicating data: a data pipeline or dataflow would copy data into a table or file in the lakehouse, which contradicts the requirement to avoid copying. An external table could reference external data, but shortcuts are the intended, simpler mechanism in this scenario to integrate external ADLS Gen2 data into the lakehouse without copying.

6. What is a data pipeline?

A. A sequence of activities to orchestrate a data ingestion or transformation process

B. A special folder in OneLake storage where data can be exported from a lakehouse

C. A saved Power Query

D. A dataset for BI reporting

A data pipeline is an automated sequence of tasks that orchestrates data ingestion and transformation from source systems to a destination. It coordinates multiple steps—such as extracting data, applying transformations, performing quality checks, and loading results—with dependencies, scheduling, and error handling to ensure data flows reliably. This is why describing it as a sequence of activities to orchestrate a data ingestion or transformation process best captures what a data pipeline does. The other options point to things that are not about the workflow of moving and processing data: a storage location for exporting data, a single saved query, or a dataset used for BI reporting.

7. What data retention strategies are typical in Fabric?

A. Define retention policies per dataset, archive or purge stale data, support point-in-time restore, and implement lifecycle management to balance cost and compliance.

B. Retain all data forever with no policy.

C. Purge everything after a fixed period regardless of usage.

D. Retention is not configurable in Fabric.

In Fabric, data retention is handled through policy-driven controls that apply to each dataset, balancing cost, compliance, and recoverability. Defining retention policies per dataset lets different datasets follow the specific regulatory or business needs they have. Archiving older data moves it to cheaper storage while keeping it accessible if needed, and purging stale data helps control costs and meet privacy rules. Point-in-time restore provides a way to recover data from a specific moment, which is important after mistakes or ransomware within the retention window. Lifecycle management automates moving and deleting data across storage tiers to maintain ongoing cost and compliance balance. The other options don't fit because keeping everything forever has no policy, purging everything after a fixed period overlooks value and usage, and retention is indeed configurable in Fabric.

8. How would you describe a lakehouse table stored on OneLake enabling BI-friendly SQL with ACID transactions?

A. A NoSQL key-value store

B. A relational database managed separately

C. A CSV file

D. A Delta Lake table on OneLake enabling BI-friendly SQL queries with ACID transactions

Using a Delta Lake table on OneLake gives you a lakehouse that supports BI-friendly SQL with ACID transactions. Delta Lake adds ACID guarantees, scalable metadata management, and data versioning on top of cloud storage, so you can perform reliable, concurrent reads and writes and run SQL-based queries for business intelligence without worrying about partial writes. On OneLake, this means you query the data with familiar SQL tools while preserving transactional integrity, combining the openness of a data lake with the reliability and queryability you expect from a database. The other options don't fit because a NoSQL key-value store typically lacks full SQL access and strong ACID behavior at scale; a plain CSV file has no transactional guarantees or optimized BI querying; and a relational database managed separately isn't leveraging the lake storage pattern and the Delta Lake capabilities.

9. What do temporal window transformations enable you to do?

A. Automatically enable and disable streaming at scheduled times.

B. Aggregate event data in a stream based on specific time periods.

C. Delete expired data after a specified retention period.

D. Filter events by a static property value.

Temporal window transformations focus on grouping streaming data by time intervals so you can compute metrics over consistent periods. By applying a window, you divide the continuous stream into time buckets (like 5-minute or 1-hour windows). Within each bucket you can perform aggregations such as counts, sums, or averages, giving you time-based insights like total events per minute or average latency per hour. This time-bounded aggregation is what makes window transformations powerful for real-time analytics. Other options describe different capabilities that aren't about grouping by time intervals: scheduling the streaming itself, deleting data based on retention, or filtering events by a static value. Window-based aggregation specifically enables the time-based grouping and calculation across a stream.

10. How does Fabric's data catalog facilitate data discovery?

A. It indexes datasets, schemas, lineage, access policies, and usage metadata, enabling search, discovery, and governance across assets.

B. It only indexes datasets.

C. It indexes external data sources but not lineage.

D. It stores metadata but does not support search.

Fabric's data catalog enables data discovery by indexing a broad set of metadata for each asset—datasets, schemas, data lineage, access policies, and usage metadata—so you can search across the catalog and understand how assets connect and are governed. This means you can find data by topic or domain, inspect the data's origin and its flow to downstream systems, verify who can access it and what rules apply, and see how the data is being used and how fresh it is. That combination of rich, interconnected metadata and searchable context makes discovery practical and supports governance across assets. The other options are too narrow, since indexing only datasets misses schema, lineage, and policy context, excluding lineage loses provenance, and storing metadata without search prevents quick discovery.

Next Steps

Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.

As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.

If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at hello@examzify.com.

Or visit your dedicated course page for more study tools and resources:

https://dp700msfabricdataengineer.examzify.com

We wish you the very best on your exam journey. You've got this!