

# DY0-001試験の準備方法 | 効果的なDY0-001問題例試験 | 完璧なCompTIA DataX Certification Exam日本語独学書籍



## CompTIA DataX Certification Exam Objectives

EXAM NUMBER: DY0-001



CompTIA

P.S. JpshikenがGoogle Driveで共有している無料かつ新しいDY0-001ダンプ: <https://drive.google.com/open?id=1zWDWPHHYNkBF-tZ436ynSd3mhgmLy2bq>

Jpshikenは初めて試験を受けるあなたが一回で試験に合格して、認証資格を取ることを保証します。Jpshikenが提供して差し上げたのは高品質のCompTIAのDY0-001「CompTIA DataX Certification Exam」模擬問題集で、あなたがステップバイステップで試験に準備する手順を指導しています。JpshikenのCompTIAのDY0-001試験問題集は絶対あなたに成功をもたらすことを保証します。JpshikenのCompTIAのDY0-001認定試験に準備するために色々な方法がありますが、

## CompTIA DY0-001 認定試験の出題範囲:

トピック	出題範囲
トピック 1	<ul style="list-style-type: none"><li>モデリング、分析、そして成果: このセクションでは、データサイエンス コンサルタントのスキルを評価し、探索的データ分析、特徴量特定、そしてオブジェクトの挙動と関係性を解釈するための可視化技術に焦点を当てます。データ品質の問題、特徴量エンジニアリングや変換といったデータエンリッチメントの実践、そして反復処理やパフォーマンス評価といったモデル設計プロセスについても考察します。また、実験結果に基づいてモデル選択の正当性を示す能力、そして適切な可視化ツールを用いて多様なビジネスオーディエンスに効果的にインサイトを伝える能力も評価されます。</li></ul>

トピック 2	<ul style="list-style-type: none"> <li>数学と統計: このセクションでは、データサイエンティストのスキルを測定し、仮説検定、回帰分析、確率関数など、データサイエンスで用いられる様々な統計手法の応用について学びます。また、統計分布、データ欠損の種類、確率モデルに関する理解度も評価します。受験者は、データの操作と分析に関連する基本的な線形代数と微積分の概念を理解し、ARIMAなどの時間ベースモデルと予測や因果推論に用いられる縦断的研究を比較することが求められます。</li> </ul>
トピック 3	<ul style="list-style-type: none"> <li>運用とプロセス: この試験セクションでは、AI</li> <li>ML運用スペシャリストのスキルを測定し、データサイエンスワークフローにおけるデータ取り込み手法、パイプラインオーケストレーション、データクリーニング、バージョン管理の理解度を評価します。受験者は、様々なデータタイプとフォーマットに対するインフラストラクチャのニーズを理解し、クリーンコードプラクティスを管理し、ドキュメント標準に準拠することが求められます。また、このセクションでは、継続的デプロイメント、モデルパフォーマンスマニタリング、クラウド、コンテナ、エッジシステムなどの環境全体へのデプロイメントなど、DevOpsとMLOpsの概念についても考察します。</li> </ul>
トピック 4	<ul style="list-style-type: none"> <li>データサイエンスの専門応用: この試験セクションでは、シニアデータアナリストのスキルを評価し、制約付き最適化、強化学習、エッジコンピューティングといった高度なトピックを扱います。テキストのトークン化、埋め込み、感情分析、法学修士（LLM）といった自然言語処理の基礎も網羅します。さらに、物体検出やセグメンテーションといったコンピュータービジョンのタスクについても学習し、グラフ理論、異常検出、ヒューリスティックス、マルチモーダル機械学習の理解度も評価されます。これにより、データサイエンスが複数の分野やアプリケーションにまたがってどのように展開していくかが示されます。</li> </ul>
トピック 5	<ul style="list-style-type: none"> <li>機械学習: このセクションでは、機械学習エンジニアのスキルを評価し、過学習、特徴選択、アンサンブルモデルといった機械学習の基礎概念を網羅します。教師あり学習アルゴリズム、ツリーベース手法、回帰分析手法などが含まれます。また、CNN、RNN、Transformerといったディープラーニングのフレームワークとアーキテクチャ、そして最適化手法についても解説します。さらに、教師なし学習、次元削減、クラスタリングモデルについても取り上げ、受験者が最新の分析で使用される幅広い機械学習アプリケーションと手法を理解するのに役立ちます。</li> </ul>

>> DY0-001問題例 <<

## DY0-001日本語独学書籍 & DY0-001試験合格攻略

我々はDY0-001問題集の英語版と日本語版を開発しています。英語版と日本語版の内容が同じですが、言葉だけ違います。DY0-001問題集に英語試験と日本語試験を準備する受験者たちは気楽に試験に合格することができます。それに、我々のCompTIAのDY0-001日本語版問題集を購入するなら、英語版をおまけにさし上げます。

### CompTIA DataX Certification Exam 認定 DY0-001 試験問題 (Q47-Q52):

#### 質問 #47

Which of the following compute delivery models allows packaging of only critical dependencies while developing a reusable asset?

- A. Virtual machines
- B. Containers
- C. Thin clients

- D. Edge devices

正解: **B**

解説:

# Containers (e.g., Docker) allow developers to package an application along with only the necessary runtime, libraries, and critical dependencies. This makes the asset lightweight, reusable, and portable across environments. Unlike virtual machines, containers share the host OS kernel and are far more efficient in packaging only what's essential.

Why the other options are incorrect:

- \* A: Thin clients refer to client-server models with minimal local processing - not relevant to dependency packaging.
- \* C: Virtual machines include an entire OS, leading to more overhead than necessary for reusable assets.
- \* D: Edge devices are hardware-based deployments typically used in IoT scenarios, not packaging tools.

Official References:

- \* CompTIA DataX (DY0-001) Official Study Guide - Section 5.2: "Containers enable consistent development environments by packaging applications and only critical dependencies, making them ideal for portability and reuse."
- \* Docker Documentation: "Containers package code and dependencies into a single unit of software, ensuring consistency across environments while minimizing overhead."

#### 質問 #48

A movie production company would like to find the actors appearing in its top movies using data from the tables below. The resulting data must show all movies in Table 1, enriched with actors listed in Table 2.

□ Which of the following query operations achieves the desired data set?

- **A. Perform a LEFT JOIN on Table 1 using column Movie, with Table 2 using column Acted\_In.**
- B. Perform a UNION between Table 1 using column Movie, and Table 2 using column Acted\_In.
- C. Perform an INNER JOIN between Table 1 using column Movie, and Table 2 using column Acted\_In.
- D. Perform an INTERSECT between Table 1 using column Movie, and Table 2 using column Acted\_In.

正解: **A**

解説:

# A LEFT JOIN ensures all rows from Table 1 (Top Movies) are preserved, even if there's no matching actor data in Table 2. This matches the requirement to show all movies, enriched with actor information when available.

Why the other options are incorrect:

- \* A: INNER JOIN would exclude movies without matching actor entries.
- \* B: UNION combines distinct rows - not appropriate for matching columns between two tables.
- \* C: INTERSECT shows only common movies - excludes unmatched top movies.

Official References:

- \* CompTIA DataX (DY0-001) Study Guide - Section 5.2: "LEFT JOINs are used when all records from one table (primary) must be retained, even if there are no matching rows in the secondary table."

#### 質問 #49

A data scientist is merging two tables. Table 1 contains employee IDs and roles. Table 2 contains employee IDs and team assignments. Which of the following is the best technique to combine these data sets?

- A. left join on Table 1 with Table 2
- B. right join on Table 1 with Table 2
- **C. inner join between Table 1 and Table 2**
- D. outer join between Table 1 and Table 2

正解: **C**

解説:

# An inner join returns only those records that have matching keys (employee IDs in this case) in both tables.

Since each table provides a different attribute for the same entity (employee), an inner join is the most efficient and accurate method when focusing on employees present in both tables.

Why the other options are less ideal:

\* B & C: Left or right joins would include unmatched data, which may lead to nulls.

\* D: An outer join brings in all records from both tables and fills nulls where no matches exist, which may introduce irrelevant or incomplete entries.

Official References:

\* CompTIA DataX (DY0-001) Official Study Guide - Section 5.2:"Inner joins are most appropriate when combining datasets with matching keys to retain only relevant, intersecting records."

\* SQL for Data Analysts, Chapter 3:"Use inner joins when combining tables on a common key to include only matched data for analysis."

## 質問 #50

A data scientist needs to:

Build a predictive model that gives the likelihood that a car will get a flat tire.

Provide a data set of cars that had flat tires and cars that did not.

All the cars in the data set had sensors taking weekly measurements of tire pressure similar to the sensors that will be installed in the cars consumers drive.

Which of the following is the most immediate data concern?

- A. Multivariate outliers
- B. Lagged observations
- C. Insufficient domain expertise
- D. **Granularity misalignment**

正解: D

解説:

# Granularity misalignment refers to a mismatch between the level of detail in the predictor variables and the event being predicted. In this case, flat tires are likely discrete, infrequent events, while tire pressure is measured weekly. If the prediction model is trying to link a specific tire pressure value to a binary outcome (flat tire: yes/no), and the timing doesn't align precisely, the predictor variable (pressure) may not be granular enough to accurately associate with the event.

Why the other options are incorrect:

\* B: While outliers can exist, they are not the most immediate concern given the time-series nature of the data.

\* C: While domain expertise is helpful, it doesn't directly address the data structure issue.

\* D: Lagged observations can be engineered in modeling but aren't the primary problem here.

Official References:

\* CompTIA DataX (DY0-001) Official Study Guide - Section 3.1 (Data Granularity):"Granularity misalignment occurs when the temporal or spatial resolution of features does not align with the prediction target."

\* Data Science Process Guide, Section 2.3:"Predictive performance can suffer when temporal mismatch exists between observations and outcomes. Granularity issues must be resolved prior to modeling."

## 質問 #51

Under perfect conditions, E. coli bacteria would cover the entire earth in a matter of days. Which of the following types of models is the best for explaining this type of growth?

- A. **Exponential**
- B. Logarithmic
- C. Polynomial
- D. Linear

正解: A

解説:

# Bacterial growth under ideal conditions follows exponential behavior: the population doubles at regular intervals. This results in a rapid increase that aligns with the formula:  $N(t) = N_0 e^{rt}$

ちなみに、Jpshiken DY0-001の一部をクラウドストレージからダウンロードできます: <https://drive.google.com/open?id=1zWDWPHIYNkBF-tZ436ynSd3mhgmLy2bq>