

問1

文を完成させるには、回答領域で適切なオプションを選択してください。

記述的分析により、[] がわかります。

【選択肢】

- A. 将来起こる可能性が最も高いもの
- B. 過去に何が起こったのか
- C. 結果に影響を与えるために実行できるアクション
- D. 過去に何かが起こったのか

問1（解答）

- B. 過去に何が起こったのか

【解説】

◆正解の理由

記述的分析 (Descriptive Analytics) は、データ分析の最も基本的な形式であり、過去のデータを調べて集計し、何が起こったのかを要約することを目的としています。

例えば、過去の売上高、顧客数、ウェブサイトのアクセス数などを把握し、「先月何が起こったか？」という疑問に答えるものです。

◆不正解の選択肢の補足説明

A. 将来起こる可能性が最も高いもの: これは 予測的分析 (Predictive Analytics) の目的です。過去のパターンに基づいて、未来の結果（例：次月の売上、顧客の離脱確率）を予測します。

C. 結果に影響を与えるために実行できるアクション: これは 処方的分析 (Prescriptive Analytics) の目的です。何が起こるかを予測し（予測的分析）、その結果を最適化するために「何をすべきか？」というアクションを推奨します。

D. 過去に何かが起こったのか: 選択肢Bと似ていますが、データ分析の分類では、「何が起こったのか」という実績の要約が 記述的分析 に該当します。

問2

次の各ステートメントについて、ステートメントが true の場合は
[はい] を選択します。それ以外の場合は、「いいえ」を選択します

- ①正規化には、データベース テーブル間の関係を削除することが含まれます。
- ②データベースを正規化すると、データの冗長性が減ります。
- ③正規化によりデータの整合性が向上します。

【選択肢】

- A. はい
- B. いいえ

問2 (解答)

①:いいえ

②:はい

③:はい

【解説】

①いいえ: 正規化は、テーブル間の関係を削除するのではなく、適切にデータ構造を分割することで、テーブル間の論理的な関係（リレーションシップ）を定義し、強化するプロセスです。テーブルを主キーや外部キーで関連付けます。

② データの冗長性が減る: 正規化の主な目的の1つは、情報を繰り返し保存することを避けることです（例：顧客の住所を注文テーブルごとに何度も保存しない）。

③ データの整合性が向上する: 冗長性が減ることで、同じデータに対して複数の矛盾した値が存在するリスクが減り、更新や削除の異常を防ぎ、データの正確性と一貫性が保たれます。

問3

文を正しく完成させる答えを選択してください。

抽出、変換、ロード (ETL) プロセスには [] が必要です

【選択肢】

- A. データ ソースとデータ ターゲット内の一致するスキーマ
- B. データを変換するのに十分な強力なターゲット データ ストア
- C. ターゲット データ ストアにロードされる前に完全に処理されたデータ
- D. データターゲットがリレーショナルデータベースであること

問3 (解答)

- C. ターゲットデータストアにロードされる前に完全に処理されたデータ

【解説】

◆正解の理由

ETL (Extract, Transform, Load) プロセスでは、ソースから抽出されたデータは、ロード先のターゲットシステムに書き込まれる前に、専用の処理領域（ステージング領域など）でクレンジング、集計、結合などの必要な変換 (Transform) を完了する必要があります。これがELTとの大きな違いです。

◆不正解の選択肢の補足説明

- A. データソースとデータターゲット内に一致するスキーマ:スキーマが一致していれば変換 (T) は不要ですが、ETLは異なるスキーマを持つソースデータをターゲットに合わせるために使われることが一般的です。
- B. データを変換するのに十分な強力なターゲットデータストア: 変換 (T) は、通常、ETLツールまたは中間サーバーで行われ、ターゲットデータストア (L) の処理能力には依存しません。
- D. データターゲットがリレーショナルデータベースであること: ETLのターゲットは、データウェアハウス（リレーショナルが多い）だけでなく、NoSQLデータベースやデータレイクなどの様々なデータストアタイプであることもあります。

問4

文を正しく完成させる答えを選択してください。

バッチ処理では、[]

【選択肢】

- A. データは常に一度に 1 行ずつ挿入されます。
- B. データはリアルタイムで処理されます。
- C. データ処理結果を配信する際の遅延は許容範囲内です。
- D. 処理はシリアルにのみ実行できます。

問4 (解答)

C. データ処理結果を配信する際の遅延は許容範囲内です。

【解説】

◆正解の理由

バッチ処理 (Batch Processing) は、蓄積された大量のデータを一定間隔でまとめて処理する手法です。処理の開始から結果が得られるまでに、数分から数時間といった顕著な遅延 (Latency) が許容されます。これは、リアルタイム処理が必要なストリーミング処理と対比されます。

◆不正解の選択肢の補足説明

A. データは常に一度に 1 行ずつ挿入されます: バッチ処理は、効率向上のため、通常、データをまとめて (バッチで) 挿入します。

B. データはリアルタイムで処理されます: これはストリーム処理の特徴です。

D. 処理はシリアルにのみ実行できます: バッチ処理は、大規模なデータセットを効率的に処理するために、しばしば 超並列処理 (MPP) を利用して並列に実行されます。

問5

文を正しく完成させる答えを選択してください。

オーディオ ファイルの文字起こしは [] 分析の一例です。

【選択肢】

- A. 認知
- B. 説明的な
- C. 予測的な
- D. 規範的な

問5 (解答)

A. 認知

【解説】

◆正解の理由

認知分析 (Cognitive Analytics) は、AIや機械学習の技術を用いて、非構造化データ（音声、画像、テキストなど）を理解し、構造化された情報に変換することを目的とします。

オーディオファイルの文字起こし（音声認識）は、非構造化データである音声からテキストを抽出する典型的なタスクであり、認知分析に分類されます。

◆不正解の選択肢の補足説明

B. 説明的な (Descriptive) 分析: 過去のデータで「何が起こったか」を要約する分析です。

C. 予測的な (Predictive) 分析: 過去のパターンから「将来何が起こるか」を予測する分析です。

D. 規範的な (Prescriptive) 分析: 最適な結果を得るために「何をすべきか」を推奨する分析です。

問6

ビジネス上の質問に答えるために使用できる分析の種類を一致させます。

各分析タイプは、1回だけ使用することも、複数回使用することも、まったく使用しないこともあります。

- ①先月の売上が増加したのはなぜですか？
- ②さまざまな在庫アイテムを購入するために予算をどのように割り当てればよいですか？
- ③会社のビジネス文書にはどのような人物が記載されていますか？

【選択肢】

- A. 認知
- B. 診断
- C. 説明的な
- D. 予測的
- E. 規範的

問6 (解答)

①:B. 診断

②:E. 規範的

③:A. 認知

【解説】

◆正解の理由

① 先月の売上が増加したのはなぜですか? → B. 診断 (Diagnostic) 分析: 「なぜそれが起こったのか?」という原因を特定するために、記述的分析の結果をさらに深掘りする分析です。

② さまざまな在庫アイテムを購入するために予算をどのように割り当てればよいですか? → E. 規範的 (Prescriptive) 分析: 「何をするべきか?」という最適な行動を推奨する分析です。この質問は、結果を最適化するためのアクションを求めていきます。

③ 会社のビジネス文書にはどのような人物が記載されていますか? → A. 認知 (Cognitive) 分析: 文書 (非構造化データ) から、人物名などのエンティティを抽出・認識するタスクであり、AIや機械学習を利用する認知分析に該当します。

◆不正解の選択肢の補足説明

C. 説明的な (Descriptive) 分析: 「何が起こったのか?」 (例: 売上高はいくらか?) に答えます。

D. 予測的 (Predictive) 分析: 「将来何が起こるか?」 (例: 来月の売上はいくらになるか?) に答えます。

問7

次のページにJSON ドキュメントがあります。各ステートメントを完成させる回答の選択肢を選択します。

- ①顧客は [] です。
- ②アドレスは[]です。
- ③ソーシャルメディアは []です

【選択肢】

- A. ネストされた配列
- B. ネストされたオブジェクト
- C. ルートオブジェクト

問7

```
"customer" : {  
    "first name" : "Ben",  
    "last name" : "Smith",  
    "address" : {  
        "line 1" : "161 Azure Ln",  
        "line 2" : "Palo Alto",  
        "zip code" : "54762"  
    },  
    "social media": [  
        {  
            "service" : "twitter",  
            "handle" : "@bensmith"  
        },  
        {  
            "service" : "linkedin",  
            "handle" : "bensmith"  
        }  
    "phone numbers": [  
        {  
            "type" : "mobile",  
            "number" : "555-555-555"  
        }  
}
```

問7 (解答)

- ①:C. ルートオブジェクト
- ②:B. ネストされたオブジェクト
- ③:A. ネストされた配列

【解説】

① 顧客は [C. ルートオブジェクト] です

ドキュメント全体は、キー "customer" と、その値である波括弧 {...} で囲まれたデータ構造で構成されています。

JSON ドキュメントの最上位 (ルート) にあるこの単一の構造体が、ルートオブジェクトです。このオブジェクト全体が、特定の顧客の情報を保持しています。

② アドレスは [B. ネストされたオブジェクト] です

customer オブジェクトの内部に、キー "address" と、その値である別の波括弧 {...} で囲まれたデータ構造があります。

この address オブジェクトは、"line 1", "line 2", "zip code" など、複数の構造化された要素をグループ化しています。

オブジェクトの中に入れ子になっているオブジェクトであるため、ネストされたオブジェクトに該当します。

③ ソーシャルメディアは [A. ネストされた配列] です

customer オブジェクトの内部に、キー "social media" と、その値である角括弧 [...] で囲まれたデータ構造があります。

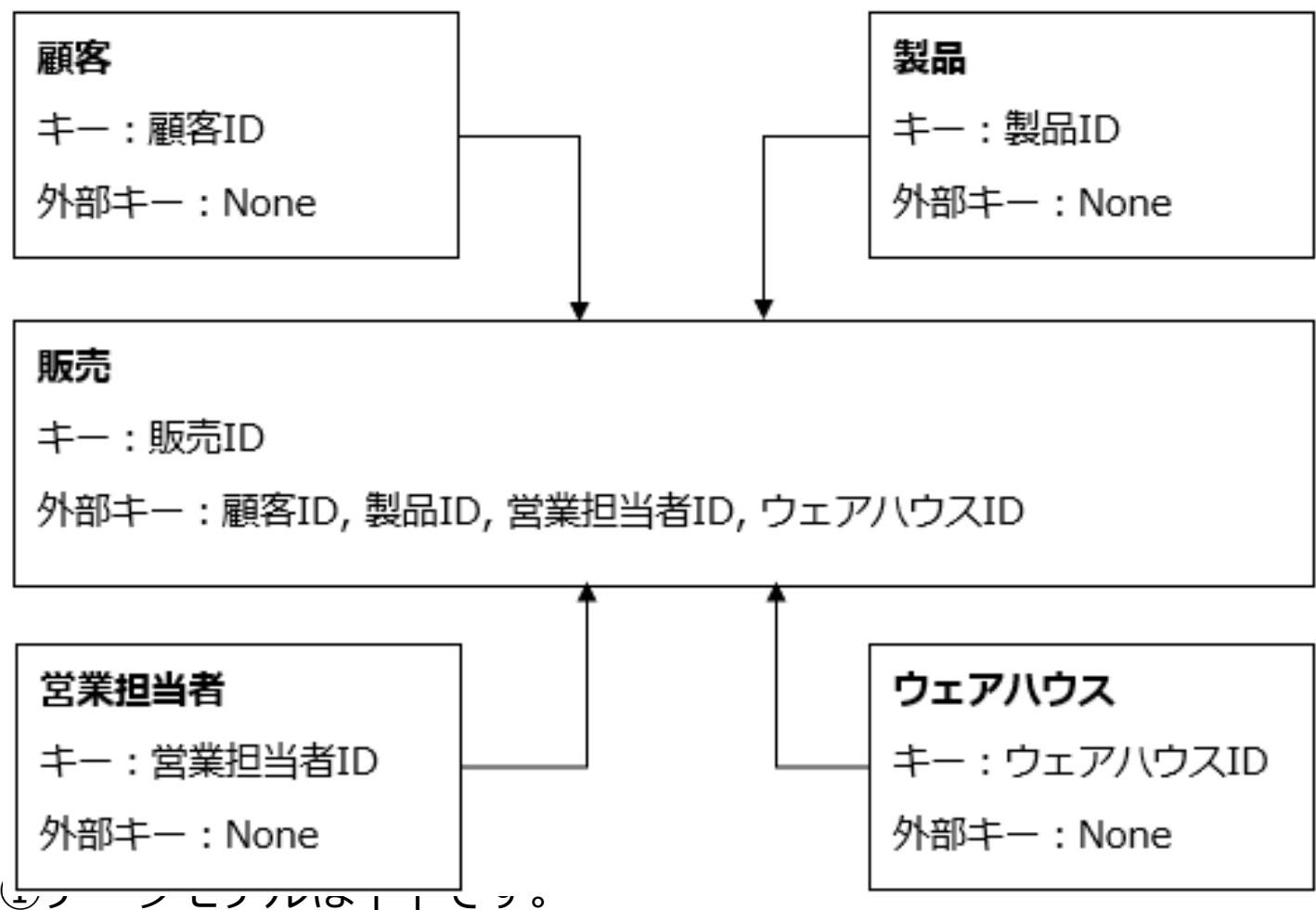
角括弧 [] は、順序付けられたデータのリスト (配列) を表します。この配列の中には、TwitterとLinkedInの2つのソーシャルメディアアカウントがリストとして格納されています。

オブジェクトの中に入れ子になっている配列であるため、ネストされた配列に該当します。

問8

次の資料に示されているデータ モデルを確認しています。

各ステートメントを完成させる回答の選択肢を選択します。



②顧客は [] テーブルです。

【選択肢】

- A. トランザクションモデル
- B. スタースキーマ
- C. スノーフレークスキーマ
- D. ファクト
- E. ディメンション
- F. ブリッジ

問8 (解答)

①:B. スタースキーマ

②:E. ディメンション

【解説】

◆正解の理由

① データモデルは [B. スタースキーマ] です

スタースキーマは、データモデリングにおいて最もシンプルかつ一般的な形式です。中央に配置された1つの大きなファクトテーブルと、それを直接取り囲む複数のディメンションテーブルで構成され、その関係性は星型に見えます。

② 顧客は [E. ディメンション] テーブルです

顧客は、ビジネスの主体に関する属性（顧客名、住所など）を格納するエンティティであるため、ディメンションテーブルに該当します。

◆不正解の選択肢の補足説明

A. トランザクションモデル

これは、分析用データモデルではなく、オンライントランザクション処理 (OLTP) システム（例：CRM、ERP）で使用されるモデルです。データの書き込み、更新、整合性を最適化するために、高度に正規化されています。

C. スノーフレークスキーマ

スタースキーマと似ていますが、ディメンションテーブルがさらに**正規化（細分化）**され、別のディメンションテーブルを持つ構造です（例：顧客ディメンションが地域ディメンションを持つ）。これによりテーブル数は増えますが、分析クエリの複雑性が増す可能性があります。

D. ファクト

ファクトテーブルは、売上数量や金額といった測定データを保持するテーブルです。「顧客」というエンティティ情報自体を保持するテーブルではないため、②の選択肢としては不適切です。

F. ブリッジ

ブリッジテーブルは、ディメンション間の多対多の関係を解決するために導入される特殊な中間テーブルです。主要なデータモデルのタイプや主要なテーブルの分類（ファクト/ディメンション）ではありません。

問9

文を正しく完成させる答えを選択してください。

Azure Synapse Analytics の超並列処理 (MPP) エンジンは []

【選択肢】

- A. 処理を計算ノード全体に分散します。
- B. 制御ノード間で処理を分散します。
- C. クライアント接続を計算ノード間でリダイレクトします。
- D. クライアント接続を制御ノード間でリダイレクトします。

問9 (解答)

A. 処理を計算ノード全体に分散します。

【解説】

◆正解の理由

Azure Synapse Analytics の専用SQLプールは、超並列処理 (MPP) アーキテクチャを使用しています。

MPPは、大規模なクエリを複数の独立したコンピューティングリソース（計算ノード）に分割し、処理を並列で実行することで、高速なデータ処理を実現します。

◆不正解の選択肢の補足説明

B. 制御ノード間で処理を分散します：制御ノードはクエリのオーケストレーションと結果の集約を担当しますが、処理の大部分は計算ノードで実行されます。

C. / D. クライアント接続をリダイレクトします：制御ノードがクライアント接続を処理し、クエリプランを計算ノードに配布します。接続を計算ノード間で直接リダイレクトすることは主要な機能ではありません。

問10

文を正しく完成させる答えを選択してください。

[] は、キーに基づいてテーブル内のデータ行を並べ替えて格納するテーブルに関連付けられたオブジェクトです。

【選択肢】

- A. クラスター化インデックス
- B. ファイルテーブル
- C. 外部キー
- D. ストアド プロシージャ

問10 (解答)

A. クラスター化インデックス

【解説】

◆正解の理由

クラスター化インデックス (Clustered Index) は、テーブル内のデータ行の物理的な格納順序を決定します。

インデックスキーの順序と、ディスク上のデータの順序が一致しているため、データ検索の効率が大幅に向上します。リレーショナルテーブルは、通常、1つ（またはゼロ）のクラスター化インデックスのみを持つことができます。

◆不正解の選択肢の補足説明

B. ファイルテーブル (Filetable): SQL Serverの特殊なテーブルで、ファイルシステム内のファイルとディレクトリを構造化データとして格納できます。

C. 外部キー (Foreign Key): 2つのテーブル間の関係を強制し、参照整合性を保証する制約です。データの物理的な並べ替えは行いません。

D. ストアド プロシージャ (Stored Procedure): データベース内に保存されたSQLコードのブロックであり、データそのものではありません。