

■ 模擬問題1

【設問】

Aさんは、コンビニエンスストアの新規出店に際して出店するかどうかを判別する機械学習のプロジェクトに携わっています。コンビニエンスストアから、表1のようなデータを預かりました。このとき、説明変数として用いる際にダミー変数に変換するデータはどれか？

店舗No	売上金額	店舗の面積	最寄り駅からの距離
1	3000	85.2	近い
2	4000	102.2	近い
3	5000	625.4	遠い
4	4000	842.3	遠い
5	7000	145.2	近い

【選択肢】

- 1.最寄り駅からの距離
- 2.売上金額
- 3.店舗面積
- 4.全て変換対象である

スキルカテゴリ：データ加工 **サブカテゴリ：**データクレンジング
スキル：名義尺度の変数をダミー変数に変換できる

■ 模擬問題2

【設問】

あるマンションの各部屋の電気使用量を1日単位で予測するモデルを作りたい。このとき必要な集計項目として最も適切でないものはどれか？

【選択肢】

- 1.マンションの1時間あたり日照時間
- 2.毎日の水道使用量
- 3.家庭の所得
- 4.家族構成

スキルカテゴリ：データの理解・検証 **サブカテゴリ：**データ理解
スキル：何のために集計しているか、どのような知見を得たいのか、目的に即して集計できる

■ 模擬問題3

【設問】

あなたは20,000件の債務者データから、住宅ローンを完済できるか分類するモデルの開発を担うことになりました。債務者データには、6つの変数[所得, 年齢, 家族構成, 職種, 当該債務者の過去の金融事故の有無, 完済したかどうかのフラグ]が含まれています。このとき、最初にやるべきこととして、最も適切でないことはどれか？

【選択肢】

- 1.分類アルゴリズムは複雑なものになると予測して、多層ニューラルネットワークにデータをインプットする
- 2.データの構造や性質を理解するために、生データを20件眺める
- 3.データの分布を確認するため、各変数に対してヒストグラムを作成する
- 4.データを完済できたものと、出来なかったものを分け、それぞれの変数の平均値と分散を比較する

スキルカテゴリ：データ可視化 **サブカテゴリ：**意味抽出

スキル：データの性質を理解するために、データを可視化し眺めて考えることの重要性を理解している

■ 模擬問題4

【設問】

下記記述は情報セキュリティの3要素の中のどの項目に該当するか？

「アクセス集中によるシステムダウンを避けるため、システムを二重化した」

【選択肢】

1. 可用性
2. 信頼性
3. 完全性
4. 安全性

スキルカテゴリ：ITセキュリティ **サブカテゴリ：**基礎知識

スキル：セキュリティの3要素（機密性、可用性、完全性）について具体的な事例を用いて説明できる

■ 模擬問題5

【設問】

下記に示す表2（売上テーブル）に対して下記の抽出条件での出力値はどれか？

```
SELECT COUNT(*) FROM 売上テーブル
WHERE 製品='商品B' AND (販売個数 >= 20 AND 販売個数 < 30) AND (営業店='大阪' OR 営業店='名古屋')
```

【選択肢】

1. 2
2. 3
3. 4
4. 5

販売Seq	製品	営業店	販売個数
1	商品A	東京	15
2	商品B	大阪	30
3	商品A	東京	27
4	商品B	東京	24
5	商品C	東京	30
6	商品B	大阪	15
7	商品B	大阪	25
8	商品A	大阪	20
9	商品C	名古屋	40
10	商品C	東京	30
11	商品B	名古屋	19
12	商品B	名古屋	22

スキルカテゴリ：データ加工 **サブカテゴリ：**フィルタリング処理

スキル：数十万レコードのデータに対して、条件を指定してフィルタリングできる（特定値に合致する・もしくは合致しないデータの抽出、特定範囲のデータの抽出、部分文字列の抽出など）

■ 模擬問題6

【設問】

Aさんは、再生可能エネルギーを開発・販売する企業でデータサイエンティストとして働くことになりました。はじめてのプロジェクトにおいては、現在、商品化を控えているエネルギーの営業を手助けすると上司から聞いており、いくつかデータも受け取っています。Aさんが行うべきこととして、最もふさわしくないものを選び

【選択肢】

1. 受け取ったデータを全て結合して機械学習を実行した
2. 営業の手助けとしてどのようなものがあるかについて同期の営業にヒアリングした
3. 顧客への営業活動から納品されるまでの業務プロセスを確認した
4. 競合である企業のIR情報を取得した

スキルカテゴリ：行動規範 **サブカテゴリ：**ビジネスマインド

スキル：「目的やゴールの設定がないままデータを分析しても、意味合いが出ない」ことを理解している