

データマネジメント試験（仮称） 科目Bのサンプル問題

問1 A社は、事務機器・オフィス用品販売会社である。データドリブン経営を推進するために、全社のデータを集約したデータ基盤を構築し、各部門による自律的なデータ活用を推進している。その実現のために、A社では、“データは生成した部門の資産であり、同時に全社の資産でもある”という考えの下、表1のデータ活用ルールを定めている。

表1 データ活用ルール

No.	ルール名	内容
1	データオーナーシップ	<ul style="list-style-type: none"> ・業務システムなどで最初にデータを生成した部門をデータオーナーとする。 ・値の加工を行わずにデータを連携・活用する場合、データ基盤上でもデータを生成した部門が、品質・意味に対する責任を負う。
2	派生データの責任	<ul style="list-style-type: none"> ・データ基盤のデータに対して、値の変換、クレンジング、統合などの加工を行って新たなデータを生成した場合、又は独自の修正を加えてレポートを作成した場合、実施した部門が成果物である当該データのデータオーナーとなり、その品質や内容に対する責任を負う。
3	カタログの管理責任	<ul style="list-style-type: none"> ・データカタログシステムの稼働維持はデータ基盤チームが行う。 ・各データの意味の記述内容の登録・更新・削除は、各データのデータオーナーが行う。
4	品質保証の原則	<ul style="list-style-type: none"> ・データの活用者は、活用するデータの品質に疑義がある場合、データカタログを参照し、データオーナーに問合せや改善要望を行う。

A社のマーケティング部のBさんは、来期のキャンペーン戦略立案のために、データ基盤にあるデータで分析することを考え、データカタログを検索した。分析対象とするデータ項目に関する情報を表2に示す。

表2 分析対象とするデータ項目に関する情報

データ項目名	連携元部門	データ基盤での加工仕様	意味
商談確度	営業部	なし。	S：受注目前 A：提案中
名寄せ後取引先名	営業部	あり。ルールベース処理で表記を統一。	分析用に表記ゆれを吸収した取引先名称

Bさんは、商談確度のデータを確認したところ、データカタログの意味にはない“C”や“引合中”といった値が多数存在し、一部のデータが未入力であることも判明した。

このままでは正確な分析ができず、キャンペーンのターゲット選定を誤るリスクがある。そこで、Bさんは、A社のデータ活用ルールに照らして、表3に示す四つのアクションを計画し、品質改善と責任の所在を明確にすることにした。

表3 品質改善アクションプラン

アクション	内容
1	意味欄に該当しない値、未入力などのデータ品質の問題を、【 a 】に報告する。
2	恒久対策には時間が掛かることから、来週の戦略会議に間に合わせる目的で、Bさんがやむを得ず、分析用データに対して独自の修正ルールを適用してレポートを作成する。この場合、レポートの数値に関する説明責任は【 b 】が負う。
3	今後、意味欄に該当しない値が発生しないよう、【 c 】が業務ルールの徹底などの恒久対策を主導する。
4	データの実態と意味欄の登録内容がかい離していることから、データカタログの意味欄を正しい状態に【 d 】が更新する。

設問 表3中の【 a 】～【 d 】に入れる字句の組合せはどれか。解答群のうち、適切なものを選べ。

解答群

	a	b	c	d
ア	営業部	営業部	営業部	営業部
イ	営業部	営業部	データ基盤チーム	データ基盤チーム
ウ	営業部	マーケティング部	営業部	営業部
エ	営業部	マーケティング部	営業部	データ基盤チーム
オ	営業部	マーケティング部	データ基盤チーム	営業部
カ	営業部	マーケティング部	データ基盤チーム	データ基盤チーム
キ	データ基盤チーム	営業部	営業部	データ基盤チーム
ク	データ基盤チーム	データ基盤チーム	営業部	営業部
ケ	データ基盤チーム	マーケティング部	データ基盤チーム	データ基盤チーム

問2 C社は、一般消費者向けのECサイトと実店舗の両方で商品を販売しており、それぞれで管理している顧客データを、データ基盤を使って統合し、分析することを目指している。当該分析はDさんが担当することになった。統合対象は、ECサイト顧客データと店舗会員データである。

Dさんはまず、両データの仕様を理解するために、データカタログで関連データの定義を確認した。データカタログの検索結果を表1に示す。

表1 データカタログの検索結果

項目	ECサイト顧客データ	店舗会員データ
データオーナー	・EC事業部	・店舗運営部
顧客ID	・数値型 ・顧客を一意に特定するキー ・制約：主キー（一意性制約あり，NULL不可）	・文字列型 ・顧客を一意に特定するキー ・制約：主キー（一意性制約あり，NULL不可）
電話番号	・文字列型 ・配送連絡用の本人認証済み携帯電話番号 ・形式：999-9999-9999 （9は0～9の半角数字を表す） ・制約：一意性制約あり，NULL不可	・文字列型 ・会員カード申込時の本人認証済み携帯電話番号 ・形式：999999999999 （9は0～9の半角数字を表す） ・制約：一意性制約あり，NULL不可
メールアドレス	・文字列型 ・顧客の会員登録時のメールアドレス ・制約：一意性制約あり，NULL不可	・文字列型 ・会員カード申込時のメールアドレス ・制約：一意性制約あり，NULL許容

Dさんは、個人情報を取り扱うために必要なC社内の申請を提出し、実データを参照することの許可を得た。その上で、データプロファイリングを行い、両データの全件を対象に、各項目の未入力（NULL）の有無や値の傾向を確認した。

その結果、ECサイト顧客データはいずれの項目も未入力はなかった。一方で、店舗会員データは、電話番号は全てのレコードについて入力されていたものの、メールアドレスは一部のレコードが未入力となっていた。

その後、具体的なデータ形式の確認のために、ECサイト顧客データと店舗会員データから同一顧客と推測される実データを抽出し、突き合わせた結果を表2に示す。

表 2 実データのサンプルの調査結果

EC サイト 顧客名 (参考情報)	EC サイト 顧客 ID	EC サイト 電話番号	EC サイト メール アドレス	店舗会員 顧客名 (参考情報)	店舗会員 顧客 ID	店舗会員 電話番号	店舗会員 メール アドレス
情報 一郎	10234	090-0000-0001	i.jouhou@example.jp	情報 一郎	S0010588	09000000001	i.jouhou@example.jp
基本 花子	25811	090-0000-0002	kihon.h@example.jp	基本 花子	S0023145	09000000002	kihon.h@example.jp
経済 次郎	31105	090-0000-0003	j.keizai@example.jp	経済 次郎	S0009872	09000000003	(NULL)

注記 1 データプロファイリングの結果、電話番号のハイフンの有無やメールアドレスの記述形式については、実データの全てにおいて、表 2 と同様の傾向であることが確認されている。

注記 2 EC サイト顧客名及び店舗会員顧客名は、確認のために各 ID にひも付く氏名情報を示したものであり、同姓同名の存在や手書き申込時の入力精度のばらつきなどがあることから、データ結合のキーとしては使用できないものとする。

経営陣から、優良顧客の特定のために、販売チャネル横断での顧客分析のレポートを至急作成するよう要求があった。D さんは、この要求に対して、このままでは同一顧客の EC サイトと店舗の購買実績をひも付けることができず、正確な優良顧客の特定ができないという根本的な課題が存在すると考えた。

D さんは、この状況における最も根本的な課題は【 a 】であると特定した。そして、経営陣からの緊急の要求に応えるための現実的な暫定対応として、【 b 】を検討した。

設問 本文中の【 a 】,【 b 】に入れる字句の組合せはどれか。解答群のうち、最も適切なものを選べ。

- (一) EC サイト顧客データと店舗会員データにおいて、電話番号のデータ形式が統一されていないこと
- (二) EC サイトと実店舗それぞれで顧客を一意に識別する顧客 ID が、販売チャネル横断の顧客識別子として機能しないこと
- (三) 店舗会員データのメールアドレスは、データカタログ上で NULL 許容とされており、実データにも未入力が含まれていること
- (四) EC サイトと実店舗それぞれで一意性が保証されている顧客 ID をそのまま結合キーとし、暫定的なレポートを作成すること
- (五) 一意性のある電話番号のデータ形式を整形し、電話番号をキーとして両データを結合して暫定的なレポートを作成すること
- (六) 一意性のあるメールアドレスをキーとして両データを結合し、暫定的なレポートを作成すること

解答群

	a	b
ア	(一)	(四)
イ	(一)	(五)
ウ	(一)	(六)
エ	(二)	(四)
オ	(二)	(五)
カ	(二)	(六)
キ	(三)	(四)
ク	(三)	(五)
ケ	(三)	(六)

データマネジメント試験（仮称） 科目Bのサンプル問題 解答例・出題趣旨

問番号	正解
問1	ウ
問2	オ

問番号	出題趣旨
問1	データ品質の問題を発見した際の一連のプロセス（報告，暫定対応，恒久対策，メタデータ管理）において，組織のデータマネジメントルールに基づき，各プロセスにおけるデータの責任主体を正しく判断する能力を問う。
問2	データカタログの情報と実データの調査結果を適切に解釈し，表面的な技術課題と，より本質的なガバナンス上の課題を切り分ける能力，そして，その状況下でビジネス要求に応えるための現実的な暫定対応を選択する能力を問う。