

問6 経営分析システムのためのデータベース設計に関する次の記述を読んで、設問 1～4 に答えよ。

P 社は、個人向けのカーシェアリングサービスを運営する MaaS（Mobility as a Service）事業者である。シェアリングのニーズが高い大都市の地区を中心に、500 駐車場で約 2,000 台の自動車（以下、車両という）を貸し出している。P 社には本社のほかに、各地区でのサービス運営を担当する支社が 10 社ある。本社はサービス全体を統括しており、新サービスの企画やマーケティングなどを行っている。支社は貸出管理システムを用いて現場で車両の貸出管理業務を行っている。

本社では、サービス運営状況を多角的な観点でタイムリーに把握して、適切な意思決定を行うために、貸出管理システムのデータをソースとする経営分析システムを構築することになった。本社の情報システム部の Q さんはデータエンジニアに任命され、データサイエンティストである R さんとプロジェクトを推進することになった。

〔データソースの調査〕

貸出管理システムには、貸出予約及び貸出実績のデータが過去 5 年間分蓄積されている。貸出管理システムのデータモデルの抜粋を図 1 に示す。

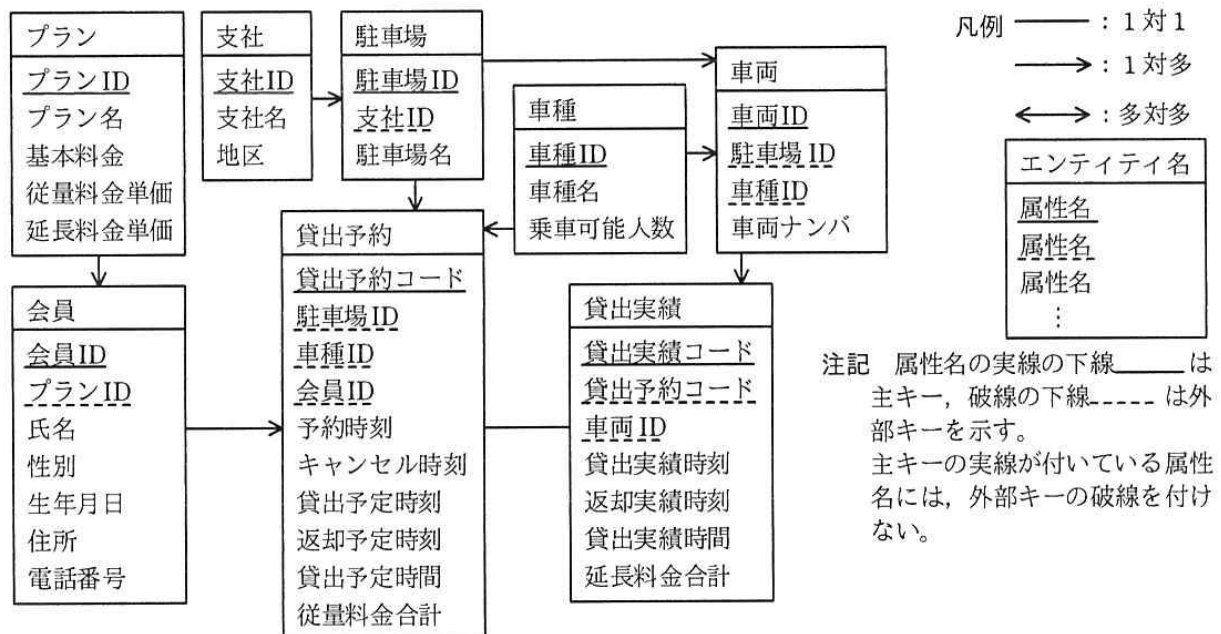


図 1 貸出管理システムのデータモデル（抜粋）

利用希望者はあらかじめ P 社の会員になり、いずれかのプランに加入しておく必要がある。プランごとに基本料金（月額）、従量料金及び延長料金（いずれも 10 分単位）の単価が決まっている。会員が車両を借りたいときは、P 社のホームページで借りたい日時や駐車場、車種などを選択し、貸出を予約する。貸出や返却の実績時刻が予約時の内容と異なる場合であっても、貸出予約の情報は修正しない。従量料金合計は予約時に指定された貸出予定時間を基に算出する。予約時に指定した返却予定時刻より早い時刻に返却しても、従量料金合計は減算しない。予約時に指定した返却予定時刻より遅い時刻に返却した場合は遅延返却として扱う。遅延返却は後の時間帯に予約している別の会員の迷惑となるので、超過した時間については従量料金よりも高い延長料金によって延長料金合計を算出する。これによって、遅延返却の発生件数（以下、遅延返却発生件数という）の低減を図っている。毎月末に当月の基本料金、従量料金合計及び延長料金合計を合算して、翌月に会員に請求する。

貸出管理システムのデータベースでは、データモデルのエンティティ名を表名にし、属性名を列名にして、適切なデータ型で表定義した関係データベースによって、データを管理している。時刻は TIMESTAMP 型、年月日は DATE 型で定義されている。

また、P 社では KPI の一つとして車両稼働率を重視している。車両稼働率とは、各車両における 1 日当たりの貸出実績時間の割合である。平均車両稼働率の目標データは、表計算ソフトのデータとして、年月日別・駐車場別・車種別に過去 3 年間分が蓄積されており、それ以前のデータは破棄されている。

〔業務要件の把握〕

P 社の経営企画部では、車両の追加整備計画の立案を検討している。Rさんは経営企画部にヒアリングを行い、経営分析システムの業務要件を把握した。業務要件の抜粋を図 2 に示す。

Qさんは、データソースの調査結果を踏まえて、図 2 の業務要件の実現可能性を評価した。その結果、①業務要件の一部は経営分析システムの運用開始直後には実現できないことが判明した。対応方針を経営企画部と協議した結果、業務要件は変更せず、運用開始直後の分析は、実現可能な範囲で行うことで合意した。

- ・地区別の人気車種、会員の性別・年代別の人気車種、駐車場別・車種別の平均車両稼働率、駐車場別・会員別の遅延返却発生件数を分析できること。なお、貸出実績の件数（以下、貸出実績件数という）が多い場合を人気車種であるとみなす。
- ・表計算ソフトのデータを用いて、平均車両稼働率の目標比や前年同期比を分析できること。
- ・これらのいずれにおいても、年別、月別、日別、週別、曜日別といった時間軸で傾向を分析できること。
- ・過去5年間について、分析対象期間を柔軟に変更して、期間による傾向の違いを分析できること。
- ・毎週月曜日の朝に最新のデータを確認できること。ただし、遅延返却発生件数については前日までの実績を翌営業日の朝に確認できること。
- ・貸出実績件数及び遅延返却発生件数は貸出予定の日付で集計すること。

図2 経営分析システムの業務要件（抜粋）

〔経営分析システムのデータモデル設計〕

次に、Qさんは図2の業務要件を基に、経営分析システムのデータモデルを多次元データベースとして設計した。多次元データベースの実装には、データモデルのエンティティ名を表名にし、属性名を列名にして、適切なデータ型で表定義した関係データベースを用いることにした。列指向データベースは用いず、データを行単位で扱う行指向データベースを用いることにした。問合せの処理性能を考慮して、データモデルの構造には a 構造を採用した。経営分析システムのデータモデルの抜粋を図3に示す。

経営分析システムには、最長で過去5年間分のデータを蓄積することにした。

年月日の週と曜日は、事前に定義したSQLのユーザ定義関数を用いて取得できる。

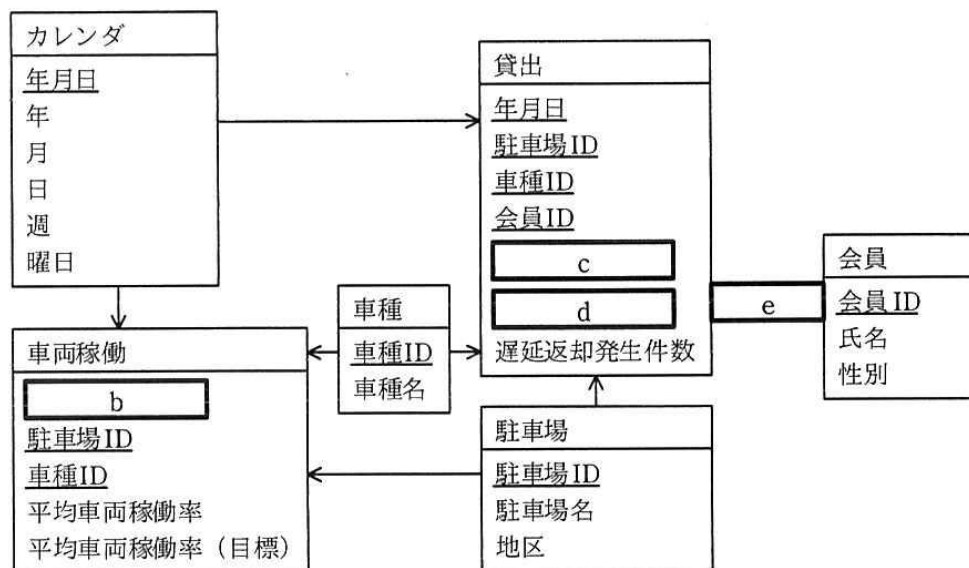


図3 経営分析システムのデータモデル（抜粋）

〔データ加工処理の開発〕

貸出管理システムのデータベースから経営分析システムのデータベースへのデータ連携時に、一部のデータを加工する必要がある。Qさんは、データ加工処理用のデータベースを用意し、データ加工を行うバッチ処理プログラムを開発した。図4のSQL文は、そこで用いられている図3の貸出表の遅延返却発生件数データを作成するためのものである。ここで、TIMESTAMP_TO_DATE関数は、指定されたTIMESTAMP型の時刻をDATE型の年月日に変換するユーザ定義関数である。

バッチ処理プログラムでは、図4のSQL文で作成したデータを貸出表に挿入する際、遅延返却発生件数が0件のレコードに対する処理も別途行うようになっている。

```
SELECT R.貸出予定年月日, R.駐車場ID, R.車種ID, R.会員ID, COUNT(*) AS 遅延返却発生件数
FROM (SELECT Y.貸出予約コード, Y.駐車場ID, Y.車種ID, Y.会員ID,
TIMESTAMP_TO_DATE(Y.貸出予定時刻) AS 貸出予定年月日, Y.返却予定時刻 FROM 貸出予約 Y) R
WHERE R.返却予定時刻 < J.返却実績時刻
```

図4 遅延返却発生件数データを作成するSQL文

〔分析のレスポンス性能の改善〕

性能検証を実施したところ、分析対象期間を過去複数年間、時間軸を月別として人気車種及び遅延返却発生件数を分析する場合、種々の分析に時間が掛かり過ぎるので改善してほしいという要望が経営企画部から挙がった。経営分析システムのデータベースのインデックスは既に適切に作成している。分析のレスポンス性能を改善するために、Qさんは②データマートとして集計表を追加した。

設問1 本文中の下線①について、実現できない業務要件を40字以内で具体的に答えよ。

設問2 〔経営分析システムのデータモデル設計〕について、(1), (2)に答えよ。

(1) 本文中の a に入れる適切な字句を解答群の中から選び、記号で答えよ。

解答群

ア 3層スキーマ

イ オブジェクト指向

ウ スタースキーマ

エ スノーフレイクスキーマ

オ 第3正規形

カ 非正規形

- (2) 図3中の ～ に入れる適切なエンティティ間の関連及び属性名を答えよ。なお、エンティティ間の関連及び属性名の表記は図1の凡例及び注記に倣うこと。

設問3 「データ加工処理の開発」について、(1)，(2)に答えよ。

- (1) 図4中の ， に入れる適切な字句を答えよ。なお、表の列名には必ずその表の相関名を付けて答えよ。

- (2) 図4のSQL文を実行すべき頻度を2字以内で答えよ。

設問4 本文中の下線②について、追加した集計表の主キーを答えよ。