

<演習問題>

1. Excel ファイルの読み込み

- Excel ファイル [EX1.xlsx] のシート [EX1] を読み込み、データセット「EX1」を作成してください。列の属性は、以下表を参照し指定してください。

NO	名前	ラベル	種類	長さ	出力形式
1	UKE_YMD	受付年月日	日付	8	YYMMDDS10.
2	MISE_CODE	支店コード	文字列	3	\$CHAR3.
3	AMOUNT	受付金額	数値	8	BEST12.

2. テキストファイルの読み込み

- CSV ファイル [EX2.csv] を読み込み、データセット「EX2」を作成してください。列の属性は、以下表を参照し指定してください。

NO	名前	ラベル	種類	長さ	出力形式
1	MISE_CODE	支店コード	文字列	3	\$CHAR10.
2	MISE_NAME	支店名	文字列	10	\$CHAR10.
3	CUST	管轄会員数	数値	8	BEST6.

- TXT ファイル [EX3.txt] を読み込み、データセット「EX3」を作成してください。列の属性は、以下表を参照し指定してください。

NO	名前	ラベル	種類	長さ	出力形式
1	YMD	日付	日付	8	YYMMDDS10.
2	EVENT	イベント名	文字列	16	\$CHAR16.
3	YYYYMM	年月	文字列	6	\$CHAR6.

3. 要約統計量の算出

- 要約統計量タスクを使用し、データセット「EX1」から以下 2 種類の平均・標準偏差・最小・最大・中央値・欠損値の数をレポート表示してください。

- 受付年月日ごとの受付金額の要約統計量
- 支店コードごとの受付金額の要約統計量

4. データのクレンジング

- データセット「EX1」を使用し、以下 A → C の順にクエリビルダを 3 つ作成してください。
 - 受付金額が欠損値のオブザベーションを削除
 - 支店コードが「010」「280」「440」のオブザベーションを抽出
 - 受付金額を「a.1万円未満」「b.1～5万円未満」「c.5～10万円未満」「d.10万円以上」の 4 つにセグメント分けし、「AMT_RANGE」として新規変数を作成

5. カテゴリ変数の頻度集計

- 演習 4 - C までで作成したデータセットをもとに、分割表分析（クロス集計表）を作成し、レポート表示してください。
対象の変数（カテゴリ）は、MISE_CODE（支店コード） × AMT_RANGE です。

(出力イメージ)

		分割表分析 結果 FREQ プロシジャ				
		表: MISE_CODE * AMT_RANGE				
		AMT_RANGE				
		a.1万円未満	b.1～5万円未満	c.5～10万円未満	d.10万円以上	合計
MISE_CODE(支店コード)	010	12	7	6	26	51
	度数	23.53	13.73	11.76	50.98	
280	度数	15	14	12	14	55
	行のパーセント	27.27	25.45	21.82	25.45	
440	度数	3	17	6	34	60
	行のパーセント	5.00	28.33	10.00	56.67	
合計	度数	30	38	24	74	166

6. データセットの横結合

- データセット「EX1」と「EX2」を、支店コードをキーに内部結合してください。
出力する変数は、「受付年月日」「支店コード」「支店名」「受付金額」としてください。

7. 集計表

- データセット「EX1」「EX2」「EX3」を使用し、各支店のイベント有無ごとに分けた受付金額合計を集計表タスクからレポート出力してください。

(出力イメージ)

		イベント有無		受付金額
		イベントあり	イベントなし	合計
支店コード	支店名			
010	札幌支店	1,945,000	3,007,000	4,952,000
115	東京支店	4,098,000	3,821,000	7,919,000
280	名古屋支店	2,535,000	722,000	3,257,000
345	大阪支店	2,244,000	1,487,000	3,731,000
440	福岡支店	2,755,000	3,136,000	5,891,000
合計 (ALL)		13,577,000	12,173,000	25,750,000