

統計処理パッケージANALYSTについて

* 桜井 尚子, ** 西原 純,
 *** 内本 佳彦, *** 阪上 直美, *** 山口 正道

— 目 次 —

1. はじめに	72
2. 適用分野	72
3. 特徴	
(1) 会話形式による統計処理	72
(2) 豊富なデータ処理機能	72
(3) 統計グラフィックス	73
4. 取り扱えるデータ	73
5. データ処理機能	
(1) 集計・作表サブシステム (ANALYST/TABULATE)	76
(2) 統計解析サブシステム (ANALYST/ANALYZE)	76
(3) 統計グラフサブシステム (ANALYST/GRAPHICS)	77
(4) J E F オブション (ANALYST/JEF)	77
6. ANALYSTシステムと機能	79
7. 使用方法	
(1) データの準備	84
(2) ANALYSTシステムへのデータの入力と定義	85
(3) データバンク	87
(4) 多重回答について	91
(5) マクロ機能について	92
8. 具体的データ処理例の紹介	
(1) ジョブ制御文	95
(2) マクロライブラリ Pnnnn, MACRO, DATA (HOKEN) の作成方法	97
(3) ANALYSTパッケージの活用の仕方	100
(4) グラフ出力	102
① 散布図・折れ線グラフ (GPLOTコマンド)	103
② 棒グラフ (BARコマンド)	104
③ 円グラフ (PIEコマンド)	106
④ 多角形グラフ (STARコマンド)	108
⑤ 立方体棒グラフ (BLOCKコマンド)	109
9. おわりに	111

* 富士通株式会社九州支店システム部
 ** 教育学部地理学
 *** 情報処理センター

1. はじめに

経済分野、経営分野、社会分野、医学・薬学・生物分野工学分野及び自然科学分野など社会の至るところで様々な活動が行われ、その結果として、大量の情報が発生している。これらの情報を記録したものがデータである。データを有効に活用する手段としてデータ解析がある。

データ解析は、主に統計手法などを用いてデータからより多くの情報を引き出すことを目的としている。ANALYSTはデータ解析を目的としたソフトウェアの一つである。

ANALYST (ANALyzer for Statistical data)は、統計情報の管理から各種集計や帳票出力などの基本的な統計処理、多変量解析などの本格的な統計解析、さらには統計グラフ処理までを行える統計データ処理パッケージで、会話処理、バッチ処理のいずれの処理形態でも利用できるほか、日本語処理機能も備えている。またPLANNER (柔軟なリレーショナルデータベースを核とした対話型エンドユーザ志向の計画管理情報システム) とのインタフェース、F9450IIのEPOCファミリとのデータ交換機能等も整備されている。

又、統計パッケージとして有名な"SPSS"や、"BMDP"と比較し、ANALYSTは項目(変数)を作成しての多様な集計が容易にでき"SAS"と同じような新しい機能(グラフ出力機能)を有している。

2. 適用分野

- (1) 臨床データ処理
- (2) 品質データ処理
- (3) 調査分析 (アンケートデータ処理など)

3. 特徴

(1) 会話形式による統計処理

端末を用いて会話形式で統計処理を行うことができる。また、大量のデータの処理を行う場合には、会話形式と同一仕様のコマンドを用いてバッチ処理を行うことができる。

(2) 豊富なデータ処理機能

- ・データのチェック/修正/加工の一時処理
- ・集計/作表から統計解析まで幅広い統計処理
- ・統計データ処理への配慮. . . データを層別化した処理
欠測値に対する処理
多重回答データの処理

(3) 統計グラフィックス

- ・円グラフ、棒グラフ、折れ線グラフ、散布図、立体棒グラフ、および多角形の作成
- ・回帰直線や t 検定結果などの統計解析結果の付加
- ・グラフの保存と編集、グラフの体裁修正

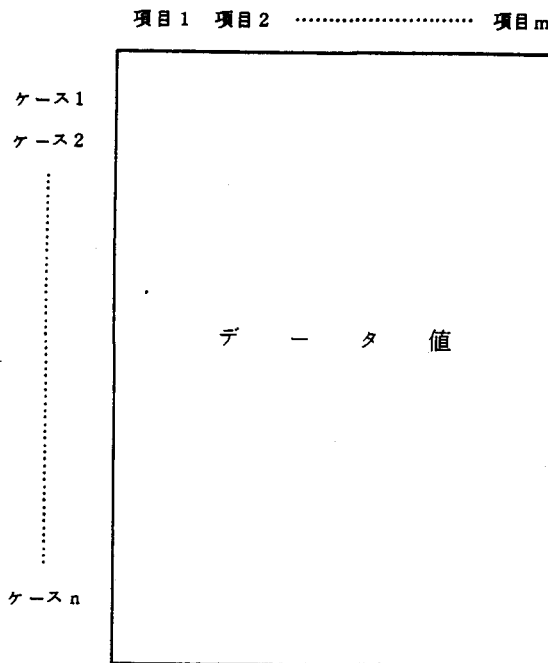
グラフィックス機能を応用すると回帰分析の予測結果や残渣など統計解析結果をグラフで表示することができる。

4. 取り扱えるデータ

ANALYSTは図1に示すような素データ行列を取り扱う。

行列は、列方向にデータ項目をとり、行方向に各ケース（サンプル）をとった多項目のデータである。ANALYSTでは、データ項目を変数と呼び一つのケースm個の変数のデータ値から構成される。m個のケースが集まって、一つの素データ行列を構成する。

図1 項目×ケースの素データ行列



取り扱えるデータの型は数値で表現された数量データ（量的データ）と数値や文字で表現された定性データ（質的データ）がある。

表1 データの統計的分類

データの種類		内 容	許される演算	例
質的 (定性的)	名義尺度	単に同質、異質の決定を行うために各目的に付けられた分類を示す特性	同じ属性を持つ標本の数の集計	性別 男…1 女…2 職業など
	順序尺度	大小関係のみが存在して絶対的な値が意味を持っていない特性	順位付けや順序統計量の計算	満足度 満足…3 普通…2 不満…1
量的 (定量的)	間隔尺度	大小関係が意味を持ち、しかも等間隔な目盛りを持った特性	和、差をもとにした統計量の計算	温度 (摂氏, 華氏)
	比率尺度 (距離尺度)	間隔尺度の性質の上に更に絶対的の原点を持った特性	加減乗除をもとにした統計量の計算	身長, 体重, 収入, 絶対 温度

5. データ処理機能

表2 データ処理機能一覧

機能項目		機能概要
基本 処 理	データ値のチェック と修正、演算機能	各種演算の組合せによるチェック、修正 など 四則演算、関数演算（平方根、対数 指数など）、条件式演算（大小比較、 一致比較、論理演算） 同一パターンのチェックや演算 多重回答データの形式変換 文字、数値データのコード化
	データのソート	複数のキーによるデータのソート
	データの表示と ファイル出力	データの定義内容の表示 条件検索と検索データの値の表示 条件検索したデータのBCDファイル出力
集 計 ・ 作 表	基本統計と簡易作図	平均、標準偏差、最大、最小などの基本 統計量の計算 散布図の作成 各ケースの2次元プロット
	ク ロ ス 表	クロス表の作成と検定 クロス統計表の作成
	集 計	基本集計表の作成 テーブル集計表の作成 内訳図の作成 ヒストグラムの作成 グループ別の統計量計算とファイル出力
統 計 解 析	多 変 量 解 析	重回帰分析（変数選択法を含む） 判別分析（ χ^2 ） 因子分析（主因子法・最尤法・直交回 転・斜交回転） 主成分分析 正準相関分析
	数 量 化 理 論	第I類、第II類（量質混合モデルの取 扱い可能） 第III類（量データの取扱い可能） 第IV類
	変 数 間 の 関 連 分 析	相関係数と偏相関係数の計算 2分属性同士の連関係数の計算 順位相関係数の計算
	検 定	汎用統計・検定表の作成 平均値の差の検定 分布の適合度の検定 対応のある変数間のノンパラメトリック 検定 グループ間のノンパラメトリック検定

これらの機能はANALYST コマンドとして、図2、表3に示す様に用意されている。

(1) 集計・作表サブシステム (ANALYST/TABULATE)

集計・作表サブシステムは、統計データの集計と作表が中心になっている。

◎クロス表の作成と検定機能

クロス集計を行い、柔軟な形式でのクロス表の表示と χ 値などによるクロス表の検定を行う。

多重回答データの集計も行える。

(2) 統計解析サブシステム (ANALYST/ANALYZE)

統計解析サブシステムは、多変量解析、数量化理論などの統計解析処理が中心となっている。ただし、クラスター分析は処理機能の中にはない。

◎数量化理論第I類

数量化理論第I類の分析処理を行う。

数量化理論第I類は外的基準が数量の場合の質的データのモデル分析であり、量的データの重回帰分析に対応する。

量質混合モデルを取り扱える拡張したアルゴリズムも備えている。

◎平均値の差の検定

グループ別に各変数の平均値を計算し、それぞれのグループ間で平均値の差があるかどうかのt検定を行う。

多変数の平均値 (ベクトル) をまとめて、各グループ間で差があるかどうかのホテリングのT検定も行える。

◎ノンパラメトリック検定(1)

対応のある変数間の代表値の差の検定を行う。

符号検定、ウィルコクソンのT検定、フリードマンの2変数分散分析、ケンドールの一致係数の検定機能を備えている。ノンパラメトリック検定のほかにt検定も行える。

◎ノンパラメトリック検定(2)

グループ間の代表の差の検定を行う。

マン・ウィットニーのU検定、コルモゴロフ・スミルノフ検定、クラスカル・ウォーリスのH検定の検定機能を備えている。

(3) 統計グラフサブシステム (ANALYST/GRAPHICS)

統計グラフサブシステムは、グラフィックディスプレイ、XYプロッタ、日本語ラインプリンタなどへの統計グラフの出力が中心になっている。

◎グラフ体裁の変更

グラフの体裁について、様々な変更が可能である。

1画面への複数グラフの表示、項目のソート、色や文字サイズの設定、タイトルの設定、ハッチングパターンの設定、表示シンボルや線種の設定、スケーリングの変更などのほか、細かいグラフ体裁の変更が画面を見ながら行える。

◎各種媒体への表示

F9430シリーズ、Tektronix4010シリーズ、F6653、F6658などのグラフィックディスプレイへの表示が行える。

また、XYプロッタへの出力や日本語ラインプリンタへの出力も可能である。

◎グラフの保存

作成したグラフをデータバンクに保存しておくことができる。

◎グラフの再表示

データバンクに保存したグラフを縮小、拡大したり、あるいは、複数のグラフを画面上にレイアウトして表示することができる。

◎補助機能

グラフ作成のための各種の補助機能を備えている。

ハードコピー、画面の消去、処理の一定時間の休止、ヘッディングやフットィングの定義などの機能がある。

(4) JEFオプション (ANALYST/JEF)

集計、作表処理に日本語処理機能を付加する。

◎日本語情報の定義/入力

定義/入力可能な日本語情報としては次のものがある。

a. ヘッディング及びフットィング

統計処理結果リスト中に印刷されるヘッディング及びフットィング

b. ラベル

変数やコード値に対する説明文

c. データ値

主に氏名などデータ識別用に利用する。

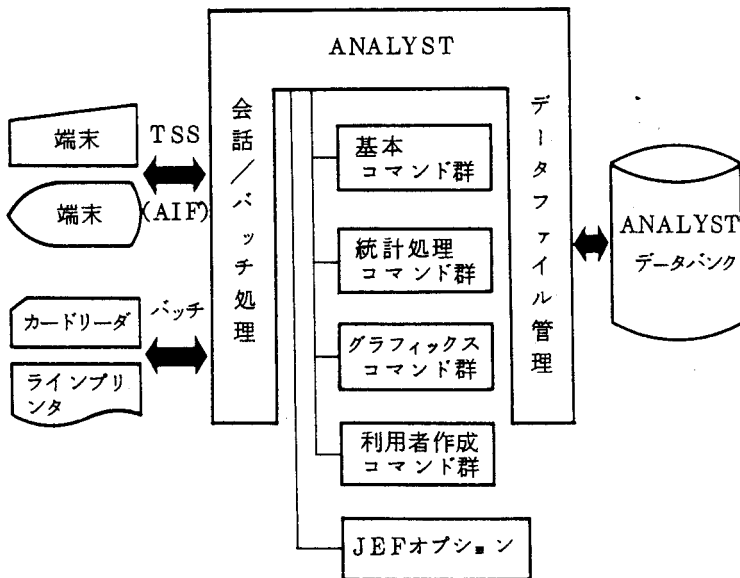
文字データと同じような扱いを受ける。

日本語情報を定義/入力する方法として、次のような各種変換が使用できる。

- ・ K I Sコード変換
- ・ 16進コード変換
- ・ カタカナひらがな変換
- ・ 英数字カナコード変換
- ・ J E F漢字コードの直接入力（日本語直定数）
- ・ 日本語ファイルによるコード変換
- ・ 日本語項目コード変換

◎日本語による統計処理結果の表示

処理結果リストに、日本語のヘッディング及びフットィングが印刷できるほか、集計・作表処理、統計解析処理、統計グラフ処理では日本語による処理結果の表示が行える。



↔ : コマンド, 結果表示, データの流れを示す。

6. ANALYSTシステムと機能

ANALYSTはTSS環境下、あるいはバッチ環境下で使用でき、独立した各コマンド処理プログラムからなる。

会話型 (TSS) でANALYSTを利用する場合は下記の通りにコマンドを入力する。

コ マ ン ド	オ べ ラ ン ド
%ANALYST	[BANK (データセット名)] [BNEW] [MACRO (データセット名)] [MNEW]

(オペランド: BANK, MACROについては後述)

LOGON TSS F1234 S(2048)

:

READY

%ANALYST

ANALYST_ 各種ANALYSTコマンド (図2, 表3) 入力

バッチでANALYSTを利用する場合には下記のカタログドプロシジャを使用してジョブ制御文を作成する。

プロシジャ名	記 号 パ ラ メ タ
ANALYST	

//F1234A JOB, CLASS=F, REGION=2048K

// EXEC ANALYST

:

:

図2 ANALYST/BASEのコマンド対系

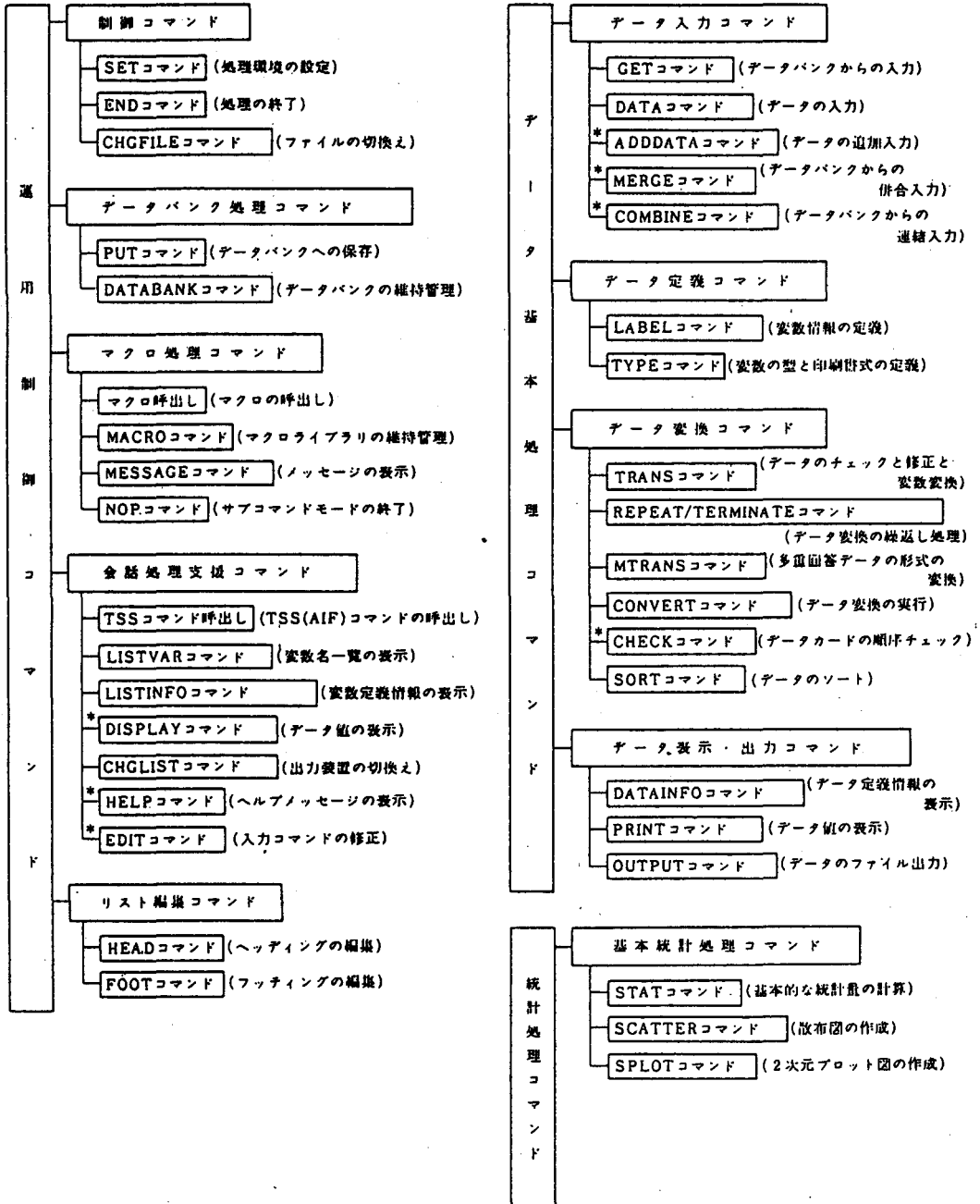


表3 コマンド一覧 (続く)

運用制御コマンド

種類	コマンド名	説明
制御コマンド	SET	処理全体に関するパラメータや条件や環境などの設定や変更を行うコマンド
	END	全処理の終了を指示するコマンド
	CHGFILE	データバンクやマクロライブラリなどのdd名(アクセス名)を切り換えるコマンド
データバンク処理コマンド	GET	データバンクからデータを入力するコマンド
	PUT	データバンクにデータを保存するコマンド
	MERGE	データバンクから複数のデータを変数方向に併合して入力するコマンド
	COMBINE	データバンクから複数のデータをケース方向に併合して入力するコマンド
	JOIN	データバンクから二つのデータを一キーに従って併合して入力するコマンド
	GETLABEL	データバンクからラベル情報を入力してマージするコマンド
	PUTLABEL	データバンクにラベル情報を保存するコマンド
	DATABANK	データバンクの維持管理を行うコマンド
マクロ処理コマンド	マクロ呼出し	マクロを呼び出すコマンド
	MESSAGE	マクロ実行中にメッセージを表示するコマンド
	NOP	マクロの中でサブコマンドモードを終了させるコマンド
	MACRO	マクロの登録や更新など、マクロライブラリの維持管理を行うコマンド
会話処理支援コマンド	TSS (AIF) コマンドの呼出し	TSS (AIF) コマンドを ANALYST の処理中に呼び出すコマンド
	CONTENTS	入力データの概要を表示するコマンド
	LISTVAR	変数名の一覧を表示するコマンド
	LISTINFO	ラベル、コード値などの変数定義情報を表示するコマンド
	DISPLAY	データ値を表示するコマンド
	CHGLIST	結果出力を端末かホスト計算機のプリンタに切り換えるコマンド
	WAIT	端末を待ち状態にするコマンド
ERASE	端末の画面を消去するコマンド	
リスト編集コマンド	HEAD	ヘッディングを編集するコマンド
	FOOT	フッタを編集するコマンド

種類	コマンド名	説明
PLANNER インタフェース コマンド	OPENDB	PLANNER データベースを開くコマンド
	CLOSEDB	PLANNER データベースを閉じるコマンド
	GETDB	PLANNER データベースからテーブルを入力するコマンド
	PUTDB	PLANNER データベースへテーブルを出力するコマンド
	SHOWDB	PLANNER データベース上のテーブル名の一覧を表示するコマンド
	PLANNER	PLANNER のコマンドを直接呼び出すコマンド
F9450-II データ伝送 コマンド	IMPORT	ANALYST からデータを F9450-II の EPOC 形式のファイルに伝送するコマンド
	EXPORT	F9450 II の EPOC 形式のファイルを ANALYST に入力データとして伝送するコマンド

データ基本処理コマンド

種類	コマンド名	説明
データ入力コマンド	DATA	端末、カード、ファイルなどからデータを入力するコマンド
	ADDVAR	入力済みのデータに、端末、カード、ファイルなどから、新規変数のデータを追加入力するコマンド
	ADDCASE	入力済みのデータに、端末、カード、ファイルなどから、新規ケースのデータを追加入力するコマンド
	CHECK	データカードの並びをチェックするコマンド
データ定義コマンド	LABEL	変数ラベル、コード値、カテゴリラベルの変数定義情報を定義するコマンド
	TYPE	各変数のデータ値の型や印刷書式を定義するコマンド
	RENAMEVAR	変数名を変更するコマンド
データ変換コマンド	TRANS	四則演算、関数演算、条件式演算を組み合わせて、データのチェックや修正や変換を行うコマンド
	REPEAT/ TERMINATE	多数の変数や定数に関する同一パターンの変換処理を行うためにTRANSコマンドの繰り返し処理を指示するコマンド
	MTRANS	多重回答データの形式の変換を行うコマンド
	CONVERT	TRANSコマンド、REPEATコマンド、MTRANSコマンドなどで指示されたデータ変換を実際に行うコマンド(データ変換の実行は統計処理コマンドの中でも自動的に行われる)。
	SORT	データをソートするコマンド
データのコード化コマンド	CODE	データ値からすべての異なる文字列(又は数値)を抽出し、コード化を行うコマンド
	MCODE	複数組のデータ値(文字列)をコード化して、多重回答形式データを自動的に作成するコマンド
	NCODE	数値データを区間に分割して、コード化を行うコマンド
データ表示・出力コマンド	DATAINFO	ラベルやコード値などの変数定義情報を表示するコマンド
	PRINT	データ値を表示するコマンド
	OUTPUT	データ値をファイルに出力するコマンド

統計処理コマンド

種類	コマンド名	説明
基本統計処理コマンド	STAT	平均、標準偏差などの基本統計量を計算するコマンド
	SCATTER	散布図を作成するコマンド
	SPLOT	各ケースを2次元でプロットするコマンド
集計・作表コマンド	MCROSS	クロス表を作成するコマンド
	CSTAT	クロス統計表を作成するコマンド
	AGGREGATE	層別統計量を産出するコマンド
	FREQ	基本集計表を作成するコマンド
	TFREQ	テーブル集計表を作成するコマンド
	FCHART	内訳図を作成するコマンド
	HIST	ヒストグラムを作成するコマンド
MHIST	拡張ヒストグラムを作成するコマンド	
統計解析コマンド	REGRES	重回帰分析を行うコマンド
	DISCRIM	判別分析を行うコマンド
	FACTOR	因子分析を行うコマンド
	PCA	主成分分析を行うコマンド
	CANCORR	正準相関分析を行うコマンド
	QUANT1	数量化理論第I類の分析を行うコマンド
	QUANT2	数量化理論第II類の分析を行うコマンド
	QUANT3	数量化理論第III類の分析を行うコマンド
	QUANT4	数量化理論第IV類の分析を行うコマンド
	CORR	相関係数、偏相関係数を計算するコマンド
	BINOM	連関係数を計算するコマンド
	RANKCORR	順位相関係数を計算するコマンド
	GTEST	汎用統計・検定表を作成するコマンド
	MEANTEST	平均値の差を検定するコマンド
FITTEST	分布の適合度を検定するコマンド	
NONPAR1	対応のある変数間のノンパラメトリック検定を行うコマンド	
NONPAR2	グループ間のノンパラメトリック検定を行うコマンド	

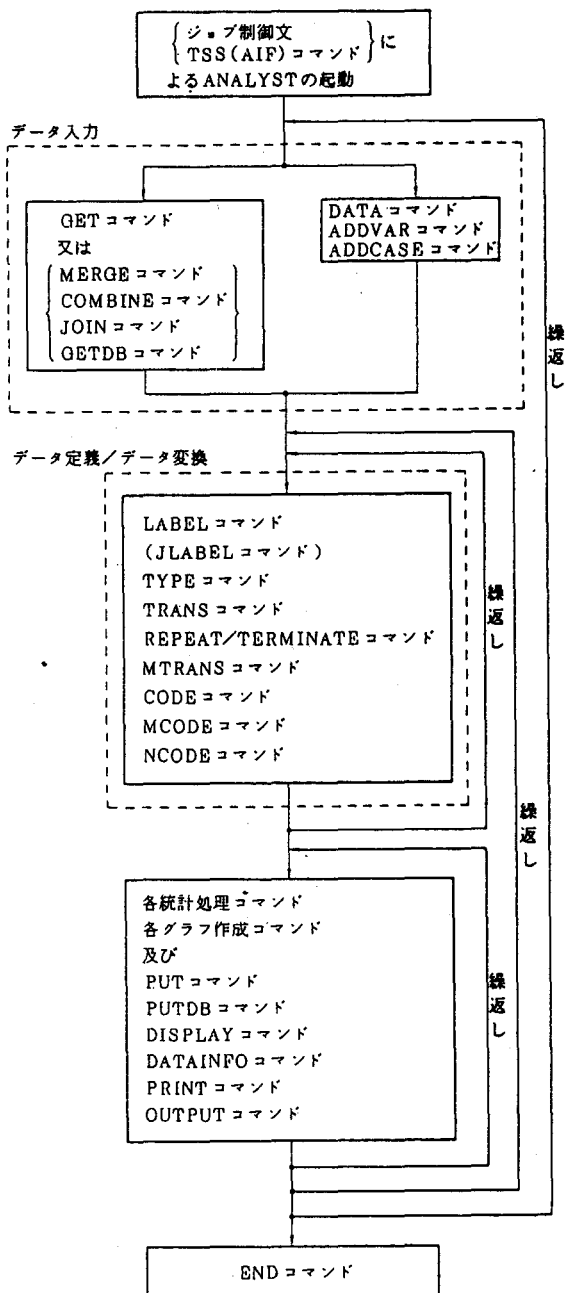
グラフィックスコマンド

種類	コマンド名	説明
グラフィックス環境設定コマンド	GRAPHICS	グラフィックス処理の環境を設定するコマンド
	GRAPHEND	グラフィックス処理の環境を解除するコマンド
グラフ作成コマンド	PIE	円グラフを作成するコマンド
	BAR	棒グラフを作成するコマンド
	GPLOT	折線グラフ、散布図を作成するコマンド
	BLOCK	立体棒グラフを作成するコマンド
	STAR	多角形グラフを作成するコマンド
グラフ作成補助コマンド	GSET	グラフの体裁の設定や変更を行うコマンド
	GHEAD	グラフのヘッディングを定義するコマンド
	GFOOT	グラフのフッタを定義するコマンド
	GEDIT	データバンクに保存したグラフをレイアウトして再表示するコマンド
	NLP	グラフの日本語ラインプリンタ出力モードを設定するコマンド

日本語処理支援コマンド

種類	コマンド名	説明
日本語情報入力コマンド	JHEAD	日本語のヘッディングを編集するコマンド
	JFOOT	日本語のフッタを編集するコマンド
	JLABEL	日本語の変数ラベル、コード値、日本語のカテゴリラベルの変数定義情報を定義するコマンド

図3 コマンド呼び出し順序の基本形



7. 使用方法

使用方法について下記の順に説明していく。

- ① データの準備
- ② ANALYSTシステムへのデータの入力と定義
- ③ データのバンク
- ④ 多重回答
- ⑤ マクロ機能について

(1) データの準備

データの構成

ANALYSTで取り扱うデータは図1のような縦方向にケース、横方向に変数がとられた長方形型のデータであり、素行列と呼ぶ。

データのコード化

統計データの中には、質的データと量的データがあり質的データはあらかじめANALYSTに入力前にコンピュータに入力できるような形にコード化しておく必要がある。

例「男」．．． 1

「女」．．． 2

但し、コードは必ずしも数値ばかりでなく、文字値を用いてもよい。入力データの制限値はIタイプで7ケタ以下、文字タイプでは4文字以下の制限がある。

データの作成

コード化したデータを、ANALYSTに入力するためには、端末を使いテキスト編集エディタにより80欄イメージのデータファイルを作成したり、直接ANALYST起動中にデータを入力する方法がある。

ANALYSTデータの編成

固定形式（各変数をどのケースにおいても一定の欄に作成、各ケースは必ず新たな行からつくる）

データファイルからの入力

DATAコマンドにより、以下に示す入力媒体、ファイルからデータを入力することが出来る。

カード入力

データファイル. . . BCDファイル

FORTRANバイナリファイル

COBOL作成ファイル

PL/I作成ファイル

他パッケージで作成したファイル. . . SDAファイル

クロスセクションファイル

READY

%ANALYST

ANALYST DATA

DATA. . . . (DATAサブコマンド入力)

DATAコマンドは各種機能をサブコマンドに分割している。

表4 DATAコマンドのサブコマンド一覧

サブコマンド	機能
INPUT	データ入力の媒体指定
VARIABLE	入力データの変数名の定義
FORMAT	入力データのデータ形式の指定
READ	データ入力開始の指示
TRANS*	基本的なデータ変換の指示
REPEAT* / TERMINATE*	データ変換の繰り返し処理の指示
MTRANS*	多重回答データの変換指示
LABEL*	変数情報の定義
TYPE*	変数の型と印刷書式の定義
LISTVAR*	変数名一覧の表示
LISTINFO*	変数情報の表示

(3) データバンク

データバンクは、ANALYSTのデータ変換コマンドでさまざまな処理を施された入力データとその定義情報（データ定義コマンドで定義した情報）を保存・管理することができ、また統計処理コマンドの統計処理結果の一部も保存・管理することが出来る。

データバンクの機能 . . . データの保存

データの取り出しと編集

データの管理

データの保存

データバンクへのデータの保存（出力）としては、入力して加工編集されたデータのPUTコマンドによる保存や一部の統計処理コマンドでの処理結果の保存などがある。また、このほかに特殊なものとして、グラフの保存もある。

データバンクに保存出力する変数やケースを選択することも、サブコマンドで指示することにより可能である。

例えば、A, B, Cの3変数を保存する場合には次のように指示する。

```
PUT CUSTOMER  
SELECT A B C
```

データバンク作成例

素データの57桁目のデータを読みデータバンク (ABANK, DATAのメンバJOBKENSU) へ加工して保存

図4 素データ

```

-----1-----2-----3-----4-----5-----6-----7-----
00010 B00J08288504470084111584111515460784111515428411151546ROAA
00020 B00J08178804450084111984111919121684110918198411191912ROLL
00030 B00J08176504450084110984110917504384110917388411091750ROBB
00040 B00J08307204430084111684111614431384111614418411161443ROBB
00050 B00J08130704430084113084113014081584113014078411301408ROBB
00060 B00J08130904430084113084113014101784113014098411301410ROBB
00070 B00J08131104430084113084113014224384113014118411301422ROBB
00080 B00J08334904380084111984111911261084111911208411191126ROCC
00090 B00J08281304380084111584111514144184111514088411151414ROFF
00100 B00J08281504380084111584111515135184111514118411151513ROCC
00110 B00J08281804380084111584111515140284111514128411151514ROBB
00120 B00J08286104380084111584111515243084111515138411151524ROFF

```

READY

%ANALYST BANK(ABANK, DATA) BNEW

ANALYST DATA_NAME(JOBKENSU)..... ①

DATA INPUT USER..... ②

DATA VARIABLE CLASS..... ③

DATA LABEL CLASS='JOB KENSU'(A='A JOB'/B='B JOB'/C='C JOB'/D='D JOB'/-
R='E JOB'/F='F JOB'/H='H JOB'/L='L JOB')... ④

DATA FORMAT FIXED('56X,A1,273X')... ⑤

DATA READ

ANALYST PUT JOBKENSU..... ⑥

説明

- ① クラス情報の出力先としてデータバンク ABANK, DATAのメンバ JOBKENSUUを割り当てる。
- ② 素データは利用者ファイルである。
- ③ 読み込むデータの変数名はCLASSと定義して⑤で指定される形式で入力する。
- ④ 変数CLASSに対して、JOBKENSUという変数ラベルをつけ、ジョブクラス A-LにそれぞれA JOB, B JOB, ... L JOBというカテゴリラベルを定義する。
- ⑤ 素データを入力する形式

- ⑥ データバンク ABANK. DATA のメンバ JOBKENSU へデータを加工し保存する。

データの取り出し

データバンクからのデータの取り出しは、保存しているデータをそのままの形で取り出す "GET コマンド" による取り出しと複数のデータを連結/併合しながら取り出す "COMBINE コマンド"、"MERGE コマンド" および "JOIN コマンド" による取り出しがある。

また、いったん取り出されたデータに対しては、データ変換やデータ定義による再編集が行えるほか、ケースや変数の追加という編集処理も行える。

"CUSTOMER" というデータ名のデータを入力するためには、次のような GET コマンドを与えるだけでよい。

```
GET CUSTOMER
```

まず、ケースの連結を行いながらデータ入力を行うコマンドとして、COMBINE コマンドがある。

同一の変数を持つ複数のデータがケース方向に結合されるわけである。

結合するデータのデータ名を X1, X2,, X9 とすると、COMBINE コマンドで次のように指定すればよい。

```
COMBINE X1 X2 . . . . X8 X9
```

次に、変数の併合を行いながらデータ入力を行うコマンドとして、MERGE コマンドがある。

対応する同一個数のケースを持つ複数のデータが変数方向に結合される。

結合するデータのデータ名を Y1, Y2,, Y7 とすると、MERGE コマンドで次のように指定すればよい。

```
MERGE Y1 Y2 . . . Y6 Y7
```

次に、同じキー変数値を持つケースどうしを一つのケースにまとめ二つのデータを併合して入力を行うコマンドとして、JOIN コマンドがある。

この例では変数Aがキー変数であり、その値が等しいケースどうしを結合している。
対応のないケースは対応のない部分に欠測値を詰めている。

併合するデータのデータ名がZ1, Z2とすると、JOINコマンドで次のように指定すればよい。

```
JOIN Z1 Z2 A
```

又、バンクどおしの処理だけでなく、ケースや変数のデータの追加も行うことができる。そのコマンドが"ADDVAR及びADDCASEコマンド"である。

データの管理

データバンク上のデータを管理するために、次のような機能を用意している。

- ・データバンクのアンロード (退避保存)
- ・データバンクのロード (復元入力)
- ・データ (メンバ) の削除
- ・データバンクの圧縮
- ・データ名 (メンバ名) の変更
- ・データ (メンバ) 一覧の表示
- ・データ管理情報の表示

これらの機能のためのコマンドとして、"DATABANKコマンド"がある。

各機能は、DATABANKコマンドのサブコマンドで処理されるようになっている。

例えば、データ名"CUSTOMER"のデータを削除し、データバンクのデータ名一覧を表示し、データバンクの圧縮を行う処理は、次のようなDATABANKコマンド及びサブコマンドを与えればよい。

```
DATABANK  
DELETE CUSTOMER  
LISTD  
CONDENSE
```

(4) 多重回答について

アンケート調査などでよく用いられる質問の形式として、多重回答式（多肢選択式、マルチアンサーとも呼ばれる）の質問がある。

この形式の質問によるデータは、ANALYSTで通常使用するデータとは、かなり違った形をしており、ANALYSTにおいても、特別の取り扱いをしなければならない。

ANALYSTでは、2つの形式で多重回答に対応できる。一つはダミー変数であり、他方は多重回答用変数である。どちらの変数形式でも一貫した統計処理に適応可能であるが、処理の性格によっては適合・不適合が考えられるので、MTRANSコマンドを用いて形式の相互変換を行うとよい。

コマンド	オ ペ ラ ン ド
MTRANS	MULTI DUMMY

(5) マクロ機能について

ANALYSTでは利用者がコマンドを入力する手続きを簡略化できるようにコマンドのマクロ化の機能を備えている。マクロとは任意のコマンドの並びを一まとめにしたものであり、これをマクロライブラリに登録しておけば、自由に何度でも呼び出して使用することが出来る。マクロライブラリは1レコード80バイトの区分データセットであり、メンバ名をマクロ名として使用する。ANALYSTのマクロライブラリは通常の一般的な区分編成ファイルであり、TSSのエデッタ (EDITコマンド) またはPFDEデッタで取り扱うことができる。ANALYSTのMACROコマンドでもEDITコマンドにより、エディタを呼び出し、マクロの更新処理を行うことができる。

```
MACRO
EDIT AAMAC
..      エディタのサブコマンドで更新を行う
..
*END SAVE
```

マクロ機能を使ったデータバンクの作成例

<マクロ名: PUTBANKの内容>

```
DATA NAME (JOBKENSU)
INPUT USER
VARIABLE CLASS
LABEL CLASS='JOB KENSU'(A='A JOB'/B='B JOB'/C='C JOB'/D='D JOB'/-
          E='E JOB'/F='F JOB'/H='H JOB'/L='L JOB')
FORMAT FIXED(' (56X, A1, 273X) ')
READ
PUT JOBKENSU
```

以上のサブコマンドを%ANALYSTコマンドのMACROコマンド"PUTBANK"として使用する。

<マクロ名"PUTBANK"を使い素データのデータバンクへの出力>

%ANALYSTコマンドを使い素データをデータバンクに出力し保存する。

(素データ:TOKEI. DATA, データバンク:ABANK. DATA, マクロ名:AMAKURO. DATA)

```
READY
ALLOC F(MYDATA) DA(TOKEI. DATA) SHR...... ①
READY
%ANALYST BANK(ABANK. DATA) BNEW MACRO(AMACRO. DATA)...... ②
ANALYST DATABANK---ABANK. DATA
ANALYST MACRO LIBRARY ---AMACRO. DATA..... ③
ANALYST START V10/L10 .....
ANALYST PUTBANK---
1 VARIABLES WILL BE INPUT WITH SPECIFIED FORMAT.
1 CARDS (RECORDS) WILL BE READ PER CASE. EFFECTIVE CASES .... 4952.. ④
DATA (JOBKENSU) HAS BEEN PUT ON DATA BANK..... ⑤
ANALYST END...... ⑥
ANALYST END V10/L10
READY
```

説明

- ① 素データTOKEI. DATAをDD名MYDATAで割り付けておく。この場合D名MYDATAは固定である。
- ② データバンクとしてABANK. DATA、マクロライブラリとして MACRO. DATAを指定する。データバンクを新たに作成する場合はBNEWパラメータを指定し既存の場合は省略する。
- ③ あらかじめ作成したMACROコマンドPUTBANKの実行で素データをデータバンクに加工して保存する。
- ④ 読み込まれた有効データ数(4952個)のメッセージ
- ⑤ データバンクへのデータの保存を確認するメッセージ
- ⑥ 統計処理を終了する。

バッチ使用例

- ① データバンクおよびマクロライブラリを新規に作成して実行、また入力する素データ (MYDATA, DATA) を指定する。

```
// EXEC ANALYST
//MACRO DD DSN=FXXXX, MAC, DATA, DISP=(NEW, CATLG),
// SPACE=(TRK, (10, 10, 10)), UNIT=PUB,
// DCB=(LRECL=80, BLKSIZE=3120, RECFM=FB)
//BANK DD DSN=FXXXX, BANK, DATA, DISP=(NEW, CATLG),
// SPACE=(TRK, (20, 20, 10)), UNIT=PUB, DCB=(LRECL=6140, BLKSIZE=6144, RECFM=VBS)
//MYDATA DD DSN=FXXXX, MYDATA, DATA, DISP=SHR
//SYSPRINT DD SYSOUT=A
//FT06FO01 DD SYSOUT=A
//SYSIN DD *
DATA NAME(TEST01)
VARIABLE A B C
FORMAT FREEFIELD
READ
PUT TEST01 REP
END
/*
//
```

- ② 作成済みのデータバンクとマクロライブラリを使用して実行する。

```
// EXEC ANALYST
//MACRO DD DSN=FXXXX, MAC, DATA, DISP=SHR
//BANK DD DSN=FXXXX, BANK, DATA, DISP=SHR
//MYDATA DD DSN=FXXXX, MYDATA, DATA, DISP=SHR
//SYSPRINT DD SYSOUT=A
//FT06FO01 DD SYSOUT=A
//SYSIN DD *
DATA NAME(TEST02)
VARIABLE A B C
FORMAT FREEFIELD
READ
PUT TEST02 REP
END
/*
//
```


8. 具体的データ処理例の紹介

富士通提供の統計処理パッケージANALYSTを使って、保健管理センターの健康診断ファイルをデータ処理し、簡単な操作で基本統計処理、及び検索結果のグラフ化ができるなど有益な結果が得られたので、ここに紹介したい。

ANALYSTによる統計処理を行う場合は、まず素データを加工して、ANALYSTのデータバンクに登録する必要がある。今回使用した素データは、健康診断ファイルで、レコード長435バイト、全データ項目数320で、ファイルレイアウトは、図1のようになっている。以下、その登録処理をバッチで行う方法を説明する。

カラム	1	2	3	4	9	10	11	16	17	18	37	40	41	44	45	435
項目名	検 診 年 度	年 次	学 籍 番 号	性 別	生 年 月 日					身 長	体 重					

図1. 素数データ (健康診断ファイル)

(1) ジョブ制御文

```
//PnnnnA JOB, CLASS=C .....①
// EXEC ANALYST
//MYDATA DD DSN=Fnnnn, REPORT59, DATA, DISP=SHR .....②
//BANK DD DSN=Fnnnn, BANK, DATA, DISP=(NEW, CATLG), .....③
// SPACE=(6144, (20, 20, 10)), UNIT=PUB,
// DCB=(LRECL=6140, BLKSIZE=6144, RECFM=VBS)
//MACRO DD DSN=Fnnnn, MACRO, DATA, DISP=SHR.....④
//SYSIN DD *
HOKEN .....⑤
PUT HOKEN .....⑥
/*
//
```

- ① Fnnnnは、課題番号で各ユーザが設定すればよい。
- ② ANALYST統計処理の入力データとなる素データ
- ③ ANALYSTデータバンクのデータセット名で区分データセットである。
- ④ 素データをどう加工するか、コマンドを処理の手順に従って並べたマクロライブラリで区分データセットである。
- ⑤ マクロライブラリのメンバ名
- ⑥ ANALYSTデータバンクFnnnn, BANK, DATAに登録するメンバ名

ジョブを実行させ、正常に終了(完了コードCD=0000または、0004)するまで、この処理を繰り返す。処理が正常に終了したら次のコマンドでデータバンクが作成されているか確認する。

READY

LISTDS BANK,DATA,SP,M

Fnnnn, BANK, DATA

-- RECFM - LRECL - BLKSIZE - DSORG

VBS 6140 6144 PG

-- VOLUMES --

PUB012

-- TOTAL - - - EXT - - - 2NDARY - - - EMPTY - - - DIR (ALLOC USED)

(TRK) (NO.) (TRK) (TRK) (BLK) (BLK)

18 2 10 0 10 1

-- MEMBERS --

HOKEN-----Fnnnn, BANK, DATAのメンバ名HOKENが確かに作成されている。

完了コードCD=0004のときは、データがないか、読み取り形式の不一致である。その場合でもそのデータが欠測値となるだけで、処理は正常に終了したものと見なしてよい。

別の素データFnnnn, HOKENGO, DATAをANALYSTデータバンクFnnnn, KEN, DATAのメンバ名KENSAとして新規に登録するときは、以下のようなジョブ制御文になる。

```

//FnnnnA JOB, CLASS=C
// EXEC ANALYST
//MYDATA DD DSN=Fnnnn, HOKEN60, DATA, DISP=SHR
//BANK DD DSN=Fnnnn, BANK, DATA, DISP=SHR
//MACRO DD DSN=Fnnnn, MACRO, DATA, DISP=SHR
//SYSIN DD *
      HOKEN
      PUT KENSA
/*
//

```

但し、マクロライブラリは、Fnnnn, MACRO, DATAのメンバ名HOKENを使用している。

(2) マクロライブラリFnnnn, MACRO, DATA (HOKEN)の作成方法

マクロとは、任意のコマンドの並びを一まとめにしたもので、これをマクロライブラリのメンバ名で登録しておけば、自由に何回でも呼び出して使用できる。

例えば、Fnnnn, MACRO, DATAは、マクロライブラリ名で、HOKENは、そのメンバ名である。データ属性は、1レコード80バイト、ブロックサイズ3120バイトの区分データセットである。次にTSSによる処理の例を示す。

TSSセッションを開設するため次のコマンドを入力する。

```
LOGON TSS Fnnnn S(2048)
```

PASSWORD?

パスワードの入力

READY

```
Edit MACRO, DATA (HOKEN)
```

以下に示すマクロライブラリデータの入力を行う。

このときのマクロライブラリのデータセット属性は、1レコード80バイト、ブロックサイズ3120バイトにシステムによって自動的に指定される。

以下にマクロライブラリの具体例を示す。

```
DATA NAME (HOKEN) .....①
INPUT USER.....②
VARIABLE  NEN GBAN1 SEX BIRT LENG WEIG BMAX BMIN TG TC HDLC WBC RBC
           HGB HCT HIMAN TCHDLC AI.....③
FORMAT FIXED(' (2X, 2I1, 5X, I1, I1, 20X, 2F4. 1, 37X, 2I3, 141X, 3I3, 9X, 2I3, 2F3. 1, -
           137X, F4. 1, I3, F3, 2) ') .....④
LABEL NEN=GAKUNEN (1=1NENJI/2=2NENJI/3=3NENJI/
           4=4NENJI/) .....⑤
```

- ① データ入力の開始宣言
- ② 入力するデータにNAMEオペランドで名前をつけておくとよい。但し、省略可能
- ③ 入力するデータはユーザファイルであることを指示。
ユーザファイルは、この場合Fnnnn, REPORT59, DATAで、DD名がMYDATAで割り付けられる。
変数名の定義をする。使用できる文字は、8文字以内の英数字、カナ名である。
- ④ 入力するデータ形式を指定するために用いる。データ形式は、FORTRANのFORMAT文と同一の記述法で入力書式を記述して指定する。
- ⑤ 変数情報の定義をする。カテゴリ変数は、必ずこのコマンドで全コード値を定義しておく。
この場合、NEN:変数名 GAKUNEN:変数ラベル 1, 2, 3, 4, 5, 6:コード値 NENJI:カテゴリラベルである。

読み込んだデータをある区間で区切り、コード化するためには、繰り返し処理をする REPEAT/TERMINATE コマンド及び変数変換を行う TRANS コマンドを使用する。

例えば、身長150cm以下の人は LENGCODE=1、150cm以上155.0cm以下の人は LENGCODE=2、……、身長180.1cm以上の人は LENGCODE=8 とコード化するには、

```
REPEAT  I1=0, 150. 1, 155. 1, 160. 1, 165. 1, 170. 1, 175. 1, 180. 1/.....①
        I2=150. 0, 155. 0, 160. 0, 165. 0, 170. 0, 175. 0, 180. 0, 200. 0/
        I3=1:8
TRANS  IF (LENG>=I1 AND LENG <=I2) LENGCODE=I3 .....②
TERMINATE.....③
        LABEL LENGCODE='SHINTYOU BETU'.....④
                1=('UNDER 150.0 CM'/
                2='150.1-155.0 CM'/
                3='155.1-160.0 CM'/
                4='160.1-165.0 CM'/
                5='165.1-170.0 CM'/
                6='170.1-175.0 CM'/
                7='175.1-180.0 CM'/
                8='180.1-200.0 CM')
```

の様に記述する。

- ① 繰り返し処理の開始を指示するとともに、繰り返し処理に用いる変数および定数を指定する。
- ② 新しい変数 LENGCODE に IF 文の論理式を満足するようなコード値を代入する。IF 文の形式は、FORTRAN とほとんど同じである。
- ③ 繰り返し処理の終了を指示。
- ④ 新しい変数 LENGCODE にラベルをつけ、各コード値の意味の説明。

以下、同様にして各変数を定義し、最後に、READ サブコマンドで Fnnnn, MACRO, DATA (HOKEN) は、完成である。

(3) ANALYSTパッケージの活用の仕方

ANALYSTパッケージは多くの機能をもっており、その機能全部を説明することはできないが、その中でも基本的な統計量を計算し、その結果をグラフ化してディスプレイに表示する機能について説明する。

まず TSSセッションを開設する。

```
LOGON TSS Pnnnn S(2048)
```

```
PASSWORD ?
```

```
パスワードの入力
```

```
READY
```

(これで準備完了)

次にANALYSTコマンドを入力する。

```
%ANALYST BANK(BANK, DATA) MACRO(MACRO, DATA)
```

```
ANALYST DATABANK.....BANK, DATA
```

```
ANALYST MACRO LIBRARY.....MACRO, DATA
```

```
ANALYST START V10/L10 01/14/86 14:15:16
```

システムからメッセージが表示され、ANALYSTモードになる。このときディスプレイには「ANALYST_」と表示される。

まずANALYSTデータバンクPnnnn, BANK, DATAから、メンバ名HOKENのバンクを指定する。

```
ANALYST_ GET HOKEN
```

```
DATA(HOKEN) HAS BEEN GOT FROM DATA BANK
```

```
ANALYST_
```

これでANALYSTを使う準備ができ、STATコマンドで基本的な統計量の計算を行うことができる。

STATコマンドで、計算できる統計量としては、和、平均、最小値、最大値、レンジ、標準偏差、分散、標準誤差、尖度、歪度、幾何平均、調和平均などである。

例:身長(LENG)についてすべての統計量を計算する。

ANALYST_ STAT LENG

STAT_ OPTION ALL

STAT_

VARIABLE	LENG
WHOLE CASES	2824
VALID CASES	2025
SUM	338354.201
MEAN	167.088
MINIMUM	142.300
MAXIMUM	198.000
RANGE	55.700
STD. DEVIATION	7.914
VARIANCE	62.631
STD. ERROR	0.176
KURTOSIS	2.652
SKEWNESS	-0.234
POSITIVE CASES	2025
GEOMETRIC MEAN	166.697
HARMONIC MEAN	166.721

WHOLE CASES:総ケース数

VALID CASES:有効ケース数

SUM:和

MEAN:平均

MINIMUM:最小値

MAXIMUM:最大値

RANGE:レンジ

STD. DEVIATION:標準偏差

VARIANCE:標準誤差

KURTOSIS:尖度

SKEWNESS:歪度

POSITIVE CASES:正の値をもつケース数

GEOMETRIC MEAN:幾何平均

HARMONIC MEAN:調和平均

例:身長(LENG)について基本統計量を、性別(SEX)に計算する。

ANALYST_ STAT LENG

STAT_ OPTION GROUP(SEX)

STAT_

VARIABLE	LENGCODE	SHINTYOU	BEU	
GROUPED BY	SEX	SEIBETU		
CATEGORY LABEL	TOTAL	OTOKO	ONNA	
WHOLE CASES	2824	1945	879	
VALID CASES	2025	1448	577	
SUM	9898.00	8132.00	1766.00	
MEAN	4.89	5.62	3.06	
MINIMUM	1.00	2.00	1.00	
MAXIMUM	8.00	8.00	6.00	
RANGE	7.00	6.00	5.00	
STD. DEVIATION	1.58	1.12	1.00	
VARIANCE	2.51	1.26	0.99	
STD. ERROR	0.04	0.03	0.04	

(4) グラフ出力

ANALYSTには、数値データやカテゴリデータを集計して、結果を統計グラフとして各種図形表示媒体(NLP、WDS、GDP等)に出力する機能をもっている。

準備として、ANALYSTモードでグラフィック環境設定のコマンドをキーインする必要がある。

ANALYST_ GRAPHICS

以後グラフ作成コマンドが入力できる。グラフモードを終了したいときは、GRAPHENDとキーインすればよい。

但し、グラフ出力が出来るのは、センター内では、ワープロ端末のWDS(F6658)と特殊端末室に設置しているグラフィックディスプレイGDP(F9434B、F9431C)だけである。

グラフは、①散布図・折れ線グラフ(GPLOTコマンド)

- ②棒グラフ (BARコマンド)
- ③円グラフ (PIEコマンド)
- ④多角形グラフ (STARコマンド)
- ⑤立体棒グラフ (BLOCKコマンド)

がある。

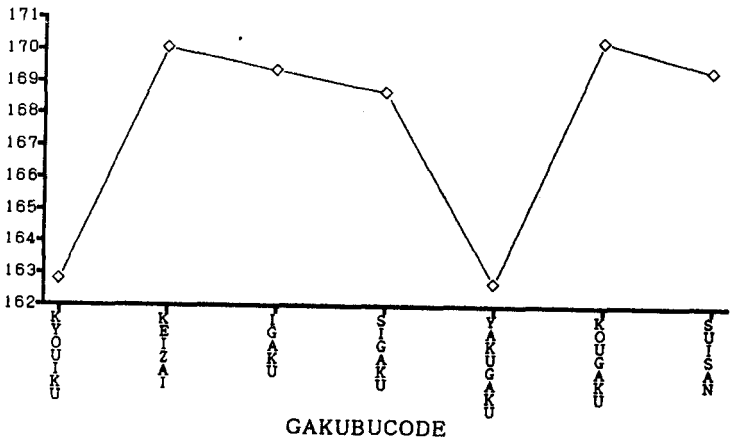
グラフ出力する場合、プロットシンボル、ハッチングパターンを選択、出力形態(例えば、昇順、降順等)をどうするか、又各セグメントラベルの表示をどうするかなど、グラフ出力の体裁を整える必要がある。そのためにグラフ体裁の変更コマンドとして、GSETコマンド/サブコマンドが用意されている。ここではそのコマンドの詳しい説明は省略しますが、次に説明するグラフ出力例では、単位を表示するため、GRAPHICSモードにおいて、あらかじめGSET JUNIT(1) (単位:「人」と表示する)が入力されているものとする。

グラフ出力例を次に示す。尚、コマンド及びオペランドの詳細な内容については、「ANALYST/GRAPHICSコマンド説明書」を参照されたい。

①散布図・折れ線グラフ

例:学部別(GAKUBU)を横軸、身長(LENG)を縦軸にして、平均値をプロットし、線で結ぶ。

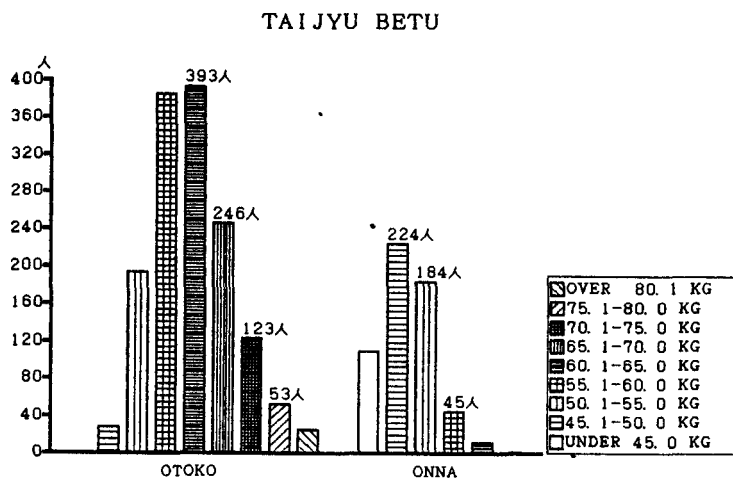
```
ANALYST_ G P L O T G A K U B = L E N G
G P L O T _ O P T I O N L I N E P L O T
G P L O T _
```



②棒グラフ

例:体重別(WEIGCODE)の人数を性別(SEX)でグループ集計して、連立棒グラフを作成する。

ANALYST_ BAR WEIGCODE
 BAR_ OPTION GROUP (SEX)
 BAR_

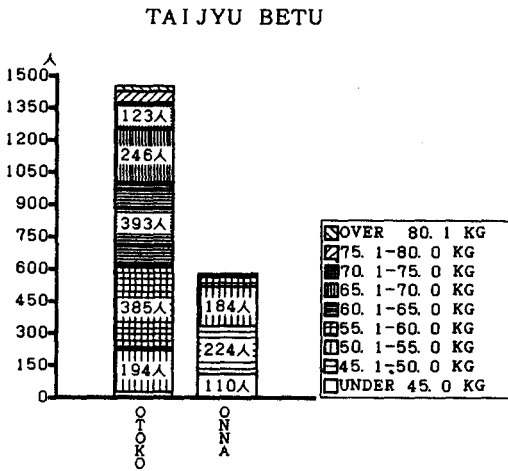


例:体重別(WEIGCODE)の人数を性別(SEX)でグループ集計して、積み上げ棒グラフを作成する。

ANALYST_ BAR WEIGCODE

BAR_ OPTION GROUP(SEX) STACK

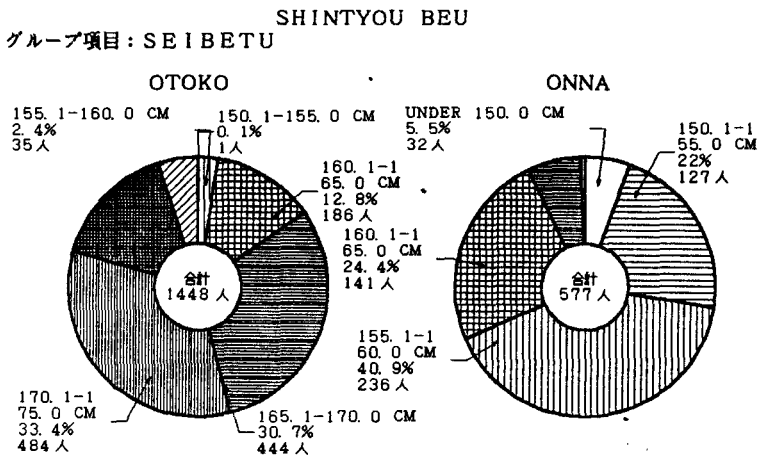
BAR_



③円グラフ

例:身長別(LENGCODE)人数を性別(SBX)でグループ集計して、2つの単純円グラフを作成する。

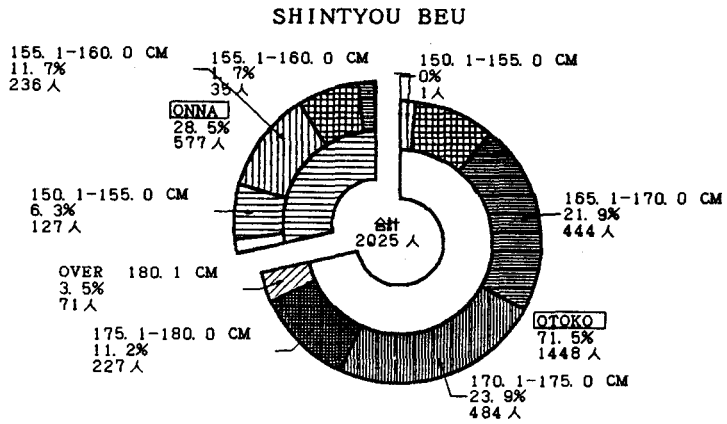
ANALYST_ PIE LENGCODE
 PIE_ OPTION GROUP(SBX)
 PIE_



例:身長別(LENGCODE)人数を性別(SEX)でグループ集計して、階層円グラフを作成する。

ANALYST_ PIE LENGCODE
PIE_ OPTION GROUP(SEX) LEVEL
PIE_

グループ項目: SEIBETU



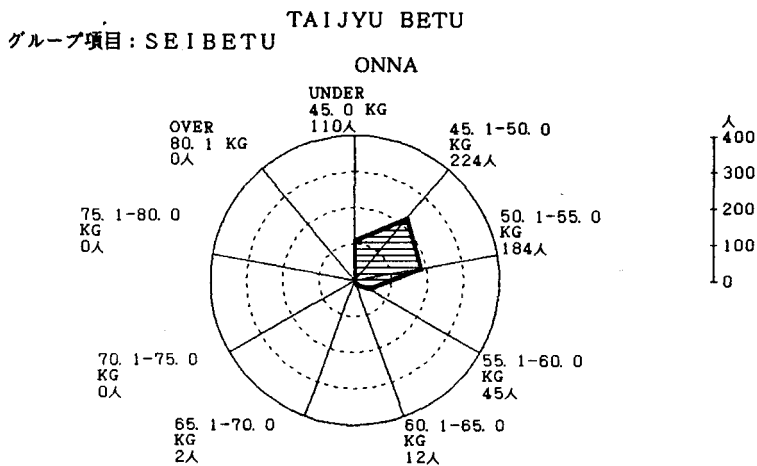
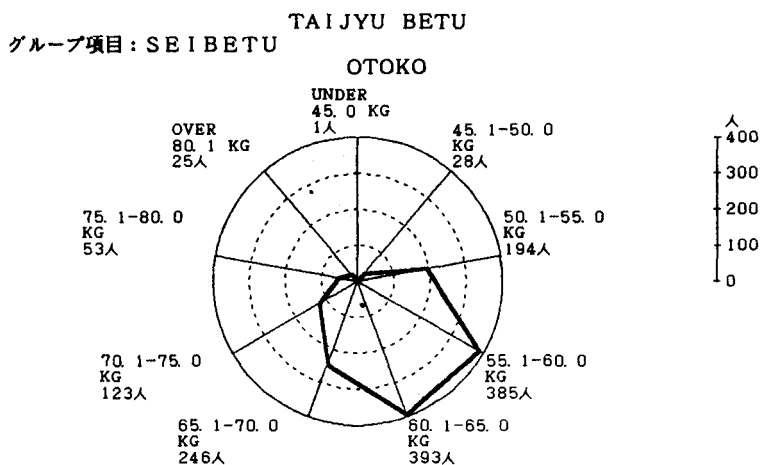
④多角形グラフ

例:体重別(WEIGCODE)人数を性別(SEX)でグループ集計して、レーダチャートで作成する。

ANALYST_ STAR_ WEIGCODE

STAR_ OPTION RADAR GROUP (SEX)

STAR_



⑤立体棒グラフ

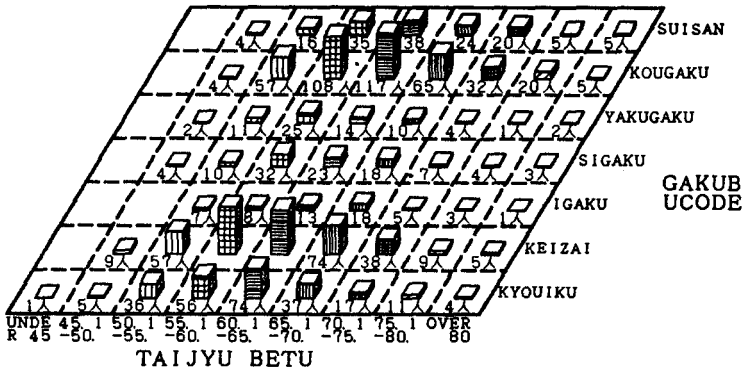
例:体重別(WEIGCODE)人数と学部別(GAKUBU)人数を、性別(SEX)にクロス度数集計して立体棒グラフを作成する。

ANALYST_ BLOCK GAKUBU/WEIGCODE

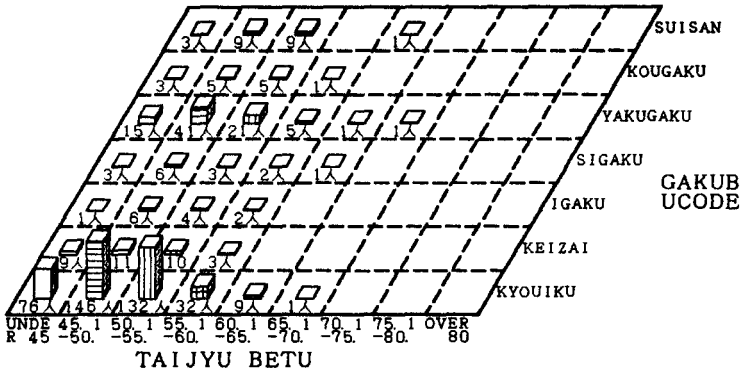
BLOCK_ OPTION GROUP(SEX)

BLOCK_

グループ項目: SEIBETU OTOKO



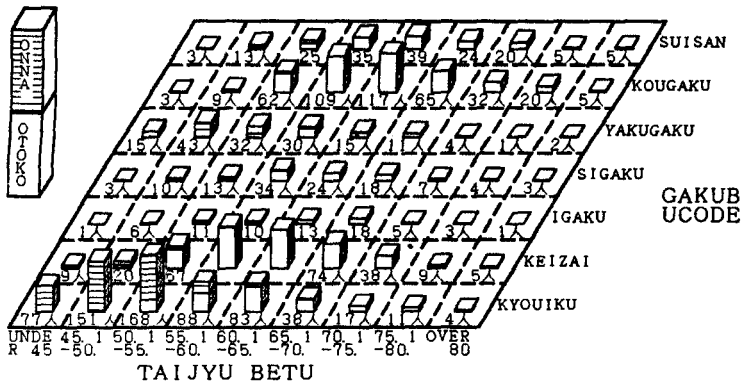
グループ項目: SEIBETU ONNA



例:体重別(WEIGCODE)人数と学部別(GAKUBU)人数を、性別(SEX)にクロス度数集計して立体棒グラフを積み上げて作成する。

ANALYST_ BLOCK GAKUBU/WEIGCODE
 BLOCK_ OPTION GROUP(SEX) STACK
 BLOCK_

グループ項目: SEIBETU



9. おわりに

以上、「統計処理パッケージANALYST」の概要である。

ANALYSTを用いて分析する場合には、次の4つのケースが考えられる。

- ① 変数名、変数ラベル、カテゴリーラベルの定義、変数変換、欠測値指定、処理コマンド、データなどすべて会話型で逐次入力する。
- ② データ入力のみファイルから入力し、他は①と同様に会話型で逐次入力をする。
- ③ 変数名、変数ラベル、カテゴリーラベルの定義、変数変換、欠測値指定、処理コマンドなどをあらかじめテキスト編集エディタ（英数字のみの場合EDITコマンド、日本語も用いる場合PFD）で、データセットに作成しておき、それをANALYST独自のMACRO機能として利用する。もちろん、データもファイルからの一括入力を行う。
- ④ 分析をバッチ処理で行うこととし、③とほぼ同様にテキスト編集エディタを用いて、コントロールカードとしてエキスト編集エディタで作成し、SUBMITする。

ただし、ANALYSTで分析しなければならないのは、ほとんどが変数やカテゴリー値が多い場合である。その際、①、②のように会話型で逐次入力するのは事実上不可能である。使用形態としては、特に最初の段階では③、④の形態が現実的である。

なお、「ANALYST」の詳細な内容については、下記マニュアルを参考されたい。

—— 参考マニュアル ——

- FACOM OSIV ANALYST解説（統計データ処理パッケージ）
- FACOM OSIV ANALYST/BASEコマンド説明書（70AR0821）
- FACOM OSIV ANALYST/TABULATEコマンド説明書（70AR0822）
- FACOM OSIV ANALYST/ANALYZEコマンド説明書（70AR0823）
- FACOM OSIV ANALYST/JEFコマンド説明書（70AR0824）