

2012 (株)数理システムユーザーコンファレンス

# 手軽に統計解析/データマイニング環境を

## S-PLUS のご紹介

2012 (株)数理システムユーザーコンファレンス 2

# S-PLUS とは？

**S-PLUS**  
オブジェクト指向  
データ駆動型システム

## 世界的な 汎用データ解析/統計解析 ソフトウェア

- 開発元は 米国 Tibco Software Inc.
- ベル研究所によって開発され、データ解析に特化した独自の言語「S」がベース
- メニューを選択してデータ加工、グラフ作成、統計解析を実行するGUI機能を搭載
- (株)数理システムが20年以上にわたり日本総代理店として、日本語版の開発・サポート対応
- データ探索と統計的データ解析の機能を多く備え、世界中の研究者、プロフェッショナルにご利用いただいている

2012 (株)数理システムユーザーコンファレンス 3

# S-PLUS 導入実績

### 銀行、証券など金融系

Customers: 国内外の大手金融機関 (銀行、証券、生命保険、信託銀行)、カード会社

適用例:

- 経済指標間の相関分析
- ヒストリカルデータからの分析と予測
- ポートフォリオ構築 (数理計画化モジュールS+NUOPTの利用)
- 与信管理

### 医薬品、バイオ

Customers: 国内外の医薬品メーカー、研究機関、大学

適用例:

- 臨床統計
- ゲノム分析

### 通信

Customers: 通信会社、携帯キャリア

適用例:

- 発呼パターンの分析
- 不正アクセス分析
- 時系列の分析

**国内 4,000 サイト以上の実績**

### マーケティング

Customers: 商社、通信販売、耐久消費財マーケティング

適用例:

- キャンペーン効果分析
- ロイヤルカスタマーの分析

### 製造業

Customers: 大手半導体メーカー、自動車関連、重工業、電気電子

適用例:

- 製造工程分析
- 品質管理のチャート作成
- 実験計画のデザインと分析

2012 (株)数理システムユーザーコンファレンス 4

# S-PLUS の特長

標準で1,000以上の分析を搭載

オンラインマニュアル・ヘルプも充実

1. 統計計算が簡単に、GUIから行える

- 検定、回帰、実験計画法、決定木、etc.

2. グラフも描ける

- 散布図、ヒストグラム、バレット図、管理図 etc.

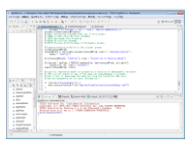
3. プログラミング & 自動実行

- 定型処理、帳票作成 etc.

2012 (株)数理システムユーザーコンファレンス 5

# S-PLUS Version 8.2 (最新版)

- Windows対応 64bit版の登場
  - 大規模データを解析
  - 搭載メモリ量にのみ依存
- S-PLUS の種類
  - GUI版 S-PLUS (32bit版のみ)
  - S-PLUS Batch (32bit版のみ)
  - コンソール版 S-PLUS
  - Workbench版 S-PLUS
    - ※32bit版と64bit版の両方を提供。一轄に利用可能 (両方提供)
- 動作環境
  - Windows 7 (32bit/64bit), Vista (32bit/64bit), XP (32bit)
  - Solaris (Sparc), Linux (RedHat、SUSE等) ※英語版



Workbench起動画面

2012 (株)数理システムユーザーコンファレンス 6

# 統計解析の必要性

せっかく集めたデータ、その分布や様子がわかれば  
経営・戦略に活用できる

データ収集

グループ 店舗  
北海道 北海道 133398  
東北 青森 126153  
東北 岩手 1427144  
東北 宮城 149111  
東北 秋田 142208

商品売上

データを眺めてみたもの...

データから特徴、傾向を見出す

宝の持ち腐れ  
どう処理すればいい?

経営・戦略に活用

有益情報の取得

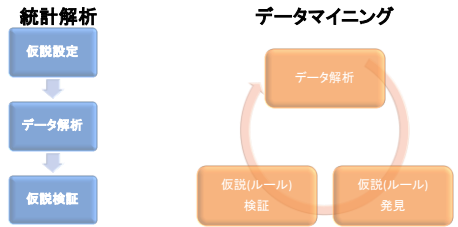
### データを取り巻く環境の変化

- 各種記録の電子データ化&大容量化
  - とりあえず何でも収集・蓄積->大容量化
  - 商取引(POSその他)、各種センサ、その他諸々のデータ
- データの蓄積、利用容易性(DB化)
  - DBMS整備->データの再利用が容易
- データの開放
  - 個人PC環境の向上、ネットワークの普及
  - (技術的には)誰でも容易にデータにアクセス

利用可能なデータは大量に存在するが  
本格的な分析などの高度利用は少ない場合が多い

↓  
データマイニングへの関心が高まる

### 統計解析 と データマイニングの違い



- 少量データから母集団の特性を推定、予測する
- あらかじめ仮説を設定する
- 検証には有意水準などの確率を用いる
- 目的に基づいてデータを収集することが多い
- 大量データからルールを発見する解析方法の総称
- あらかじめ仮説を設定しない
- 検証には全データを使用する
- データが自動的に蓄積され、大量に存在する

### データ解析の流れ



一連のプロセスや手法 => データマイニング

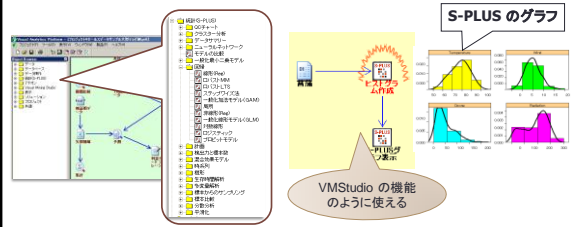
- 統計解析
- 知識工学・パターン認識
- ニューラルネットワーク 等

#### マイニングツール

<b>S-PLUS</b> データマイニング ソフトウェア	信頼性の高い統計解析を備えた、データ解析ソフトウェア
<b>Visual Mining Studio</b> Visual Mining ソフトウェア	ビジュアルプログラミングを主とした、データマイニングツール
<b>Text Mining Studio</b> Text Mining ソフトウェア	データマイニング+統計解析を使用した、テキストマイニングツール

### 汎用データマイニングツールVisual Mining Studioから S-PLUS をシームレスに連携

- Visual Mining Studio (VMStudio) は、数値システム自社開発
- S-PLUS の統計関数、グラフ機能、Sスクリプトを VMStudio のアイコンから利用可能
- データマイニングと統計解析を一環境で実行可能



VMStudio の機能のように使える

### 事例：施設保全会社様 施設劣化度の予測

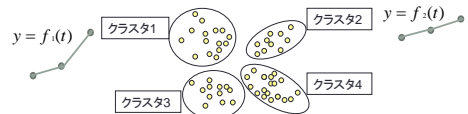
- ある公共施設の劣化について、修復や整備を担当
- 施設の劣化に影響する要因
  - 設備の利用率
  - 設備の状況(素材、場所、グレード等)
- 修復計画立案のため、翌年の劣化度(=修復費用)のより正確な予測が重要
- 全ての施設に同じ予測モデルがあてはまるとは限らない
- モデル適用単位(データの分割要因)がわからない

例)

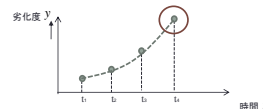


### データマイニングと統計解析で劣化度予測

- 施設の利用率・状況から、クラスタ分析を用いて似たもの同士のグループを作成 (VMStudio を利用)



- 各クラスタごとに最適な漸近式、非線形回帰モデルのパラメータを推定し、施設の劣化度を予測 (S-PLUS を利用)



2012 (株) 数値システムユーザコンファレンス 13

## WWWサーバソフト S-PLUS Enterprise Server

- S-PLUS を (Webなどの) サーバで利用
  - 大規模ネットワーク環境での解析サーバ
  - 解析ソフトをクライアントに置かない
  - 定型の高度解析処理を自動実行
- ブラウザ等で解析結果を表示
- 動作環境 (サーバ)
  - Windows Server 2008 (32bit/64bit), 2003 (32bit/64bit)
  - Linux, Solaris (英語)

2012 (株) 数値システムユーザコンファレンス 14

## S-PLUS Enterprise Server 導入メリット

- 誰でも、どこからでも利用できる
  - ブラウザとインターネットアクセスさえあれば良い
- サーバでの一元管理による、保守管理コストの低減
  - クライアントソフトの管理不要
  - データ管理の一元化
- 一般ユーザ (クライアント) と開発者の明確な分離可能
  - ユーザは利用に専念
- 少数ライセンスで全ネットワークに提供
  - 大規模組織でも同時使用ユーザは少ない

2012 (株) 数値システムユーザコンファレンス 15

## S-PLUS Enterprise Server のターゲットユーザ

2012 (株) 数値システムユーザコンファレンス 16

## 事例: 品質・精度管理

- シスメックス株式会社 様
  - 血球検査装置の世界トップメーカー
  - 世界中の顧客の検査機器をメンテナンス
  - できれば予防保守が望ましい
- SNCS (シスメックス・ネットワーク・コミュニケーション・システムズ)
  - オンラインで各地で稼働中の機器の精度管理、装置状態データを定期収集
- DBのデータを S-PLUS で自動的に可視化、分析
  - サーバの DB に各ヒストリカルデータを蓄積
  - 指定期間での品質管理図
  - さらに将来の故障予測へ

詳細は2011年ユーザコンファレンス論文集をご覧ください

2012 (株) 数値システムユーザコンファレンス 17

## 事例: 教育 e-Learning

- 立教大学様 e-Learning (社会調査士資格コース: 統計教育)
  - 解説ページ、ビデオ視聴、掲示板、小テストなど
- Web 上で対話的にデータの解析実習が可能
  - 自分でパラメータを変化させて、対話的確認が可能
  - 解析ソフトを直接使うような煩わしさは無い
- 「自分で操作している」感で学習効果アップを狙う

詳細は2010年ユーザコンファレンス論文集をご覧ください

2012 (株) 数値システムユーザコンファレンス 18

## まとめ

- 探索的なデータ解析のためのソフトウェア
- データの加工機能に優れ、高品質なグラフィックス
- データマイニングツールとの連携で手軽に幅広く分析が可能
- 統計Webアプリケーション構築に最適
- 長い実績、多数の既存ユーザ

30日間  
トライアル申込み  
受付中!

無料紹介セミナー  
毎月実施中!

各種カスタマイズ・コンサルティングも  
お受けしております。

【URL】 <http://www.msi.co.jp/splus/>  
【E-mail】 [splus-info@msi.co.jp](mailto:splus-info@msi.co.jp)