Rベースの統計解析ツールVisual R Platform

～オープンソースソフトウェアによる解析と特定用途向けソリューションのご提供～

株式会社NTTデータ数理システム
データマイニング部
関川 拓也

Global IT Innovator NTT DATA Group

NTT DATA Mathematical Systems Inc.

Copyright © 2019 NTT DATA Mathematical Systems Inc.

Visual R Platform (VRP) とは

NTT DATA Mathematical Systems Inc.

Copyright © 2019 NTT DATA Mathematical Systems Inc.

データサイエンティストに必要な3つの力

NTT DATA
NTT DATA Mathematical Systems Inc.

- **ビジネス力**：課題を解決する
- **データサイエンス力**：分析結果を読み取る
- **データエンジニアリング力**：統計、プログラム

最も大切な「ビジネス力」を持った人材が「データエンジニアリング力」を身に着ける有効な手段が

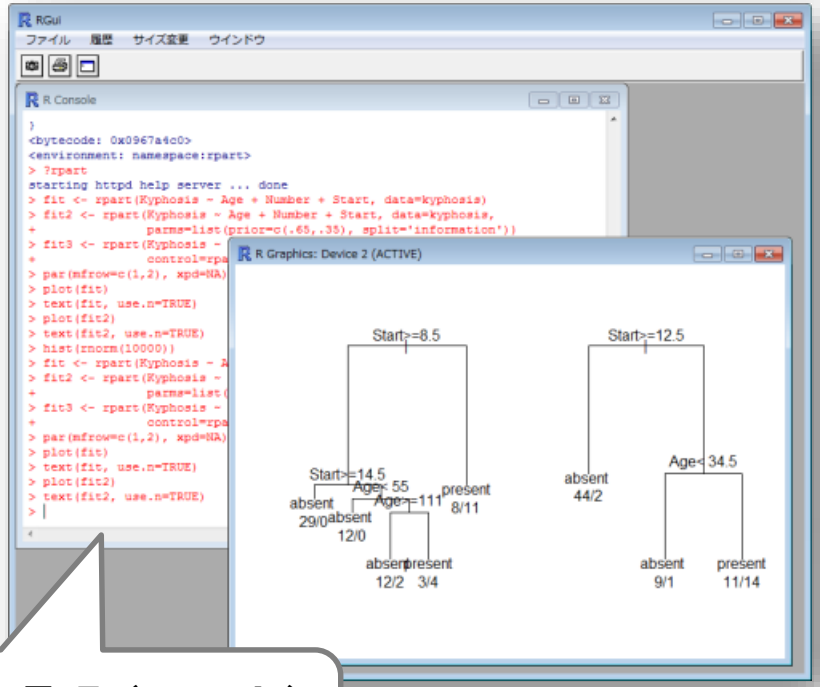
統計ソフトウェア

Copyright © 2019 NTT DATA Mathematical Systems Inc.

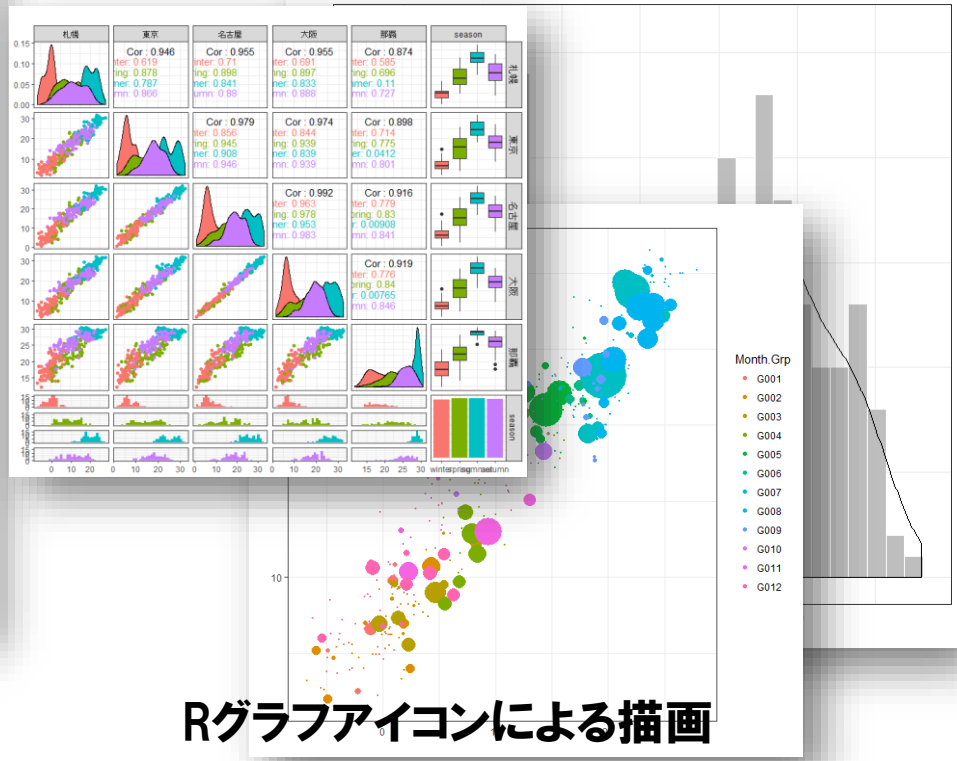
世界で最も広く使われている統計解析ツール R（アール）

NTT DATA
NTT DATA Mathematical Systems Inc.

フリーの統計ツール



「R言語」（コマンド）入力による実行



Rグラフアイコンによる描画

- **世界トップレベルのユーザー数**、米国政府機関等でも採用されています。
- **強力な機能**（統計解析、高速な関数群、豊富なグラフ機能）が利用できます。
- 膨大な拡張ライブラリ（CRANネットワーク）が日々強化され、**最新の分析アルゴリズム**を利用できます。

Copyright © 2019 NTT DATA Mathematical Systems Inc.

Rで統計！

NTT DATA
NTT DATA Mathematical Systems Inc.

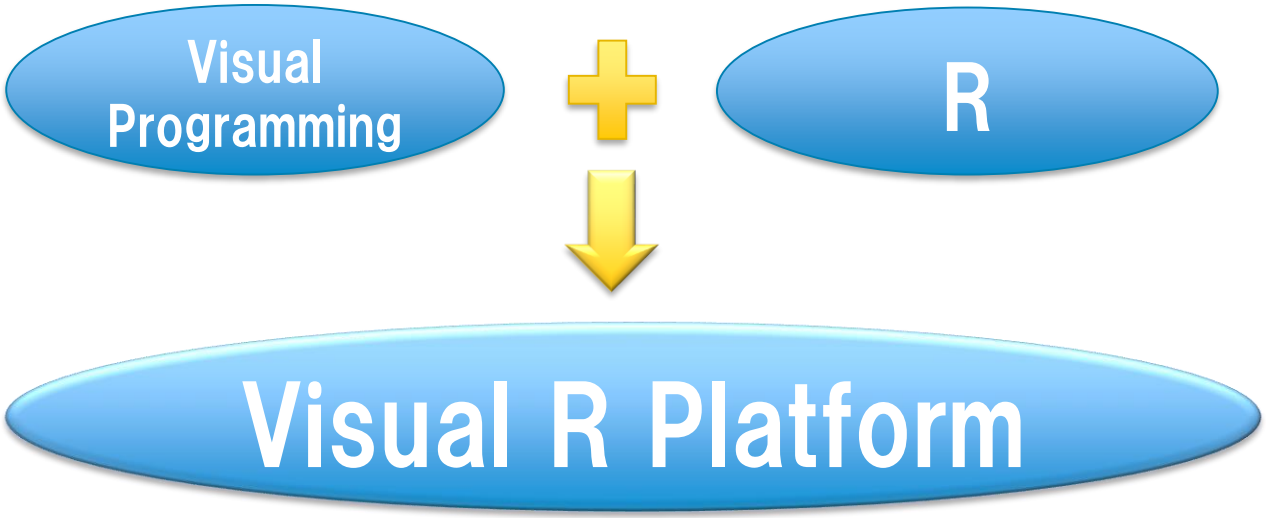
- 今や、統計ソフトのデファクトスタンダードとなったR
- しかし、Rは実は　プログラミング言語・・・
- もっと手軽に統計の分析・学習はできないものか・・？

Copyright © 2019 NTT DATA Mathematical Systems Inc.

Visual R Platformの特徴

NTT DATA Mathematical Systems Inc.

- Rベースのビジュアル分析ツール
  - 分析ツールの世界的標準Rの豊富な資産を利用
  - VMS等で定評あるビジュアルプログラミング環境

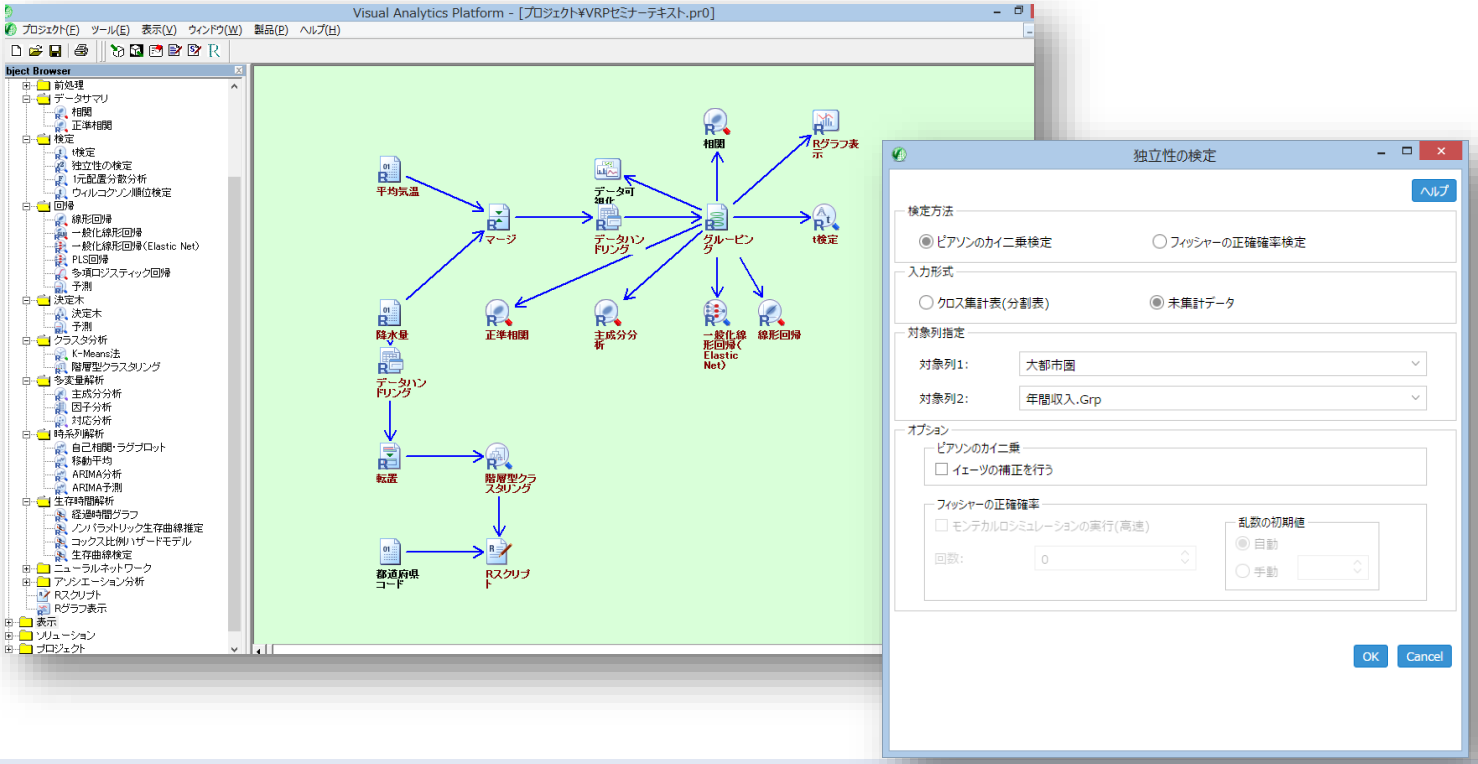


Copyright © 2019 NTT DATA Mathematical Systems Inc.

ビジュアルプログラミング環境

NTT DATA Mathematical Systems Inc.

- アイコンが分析の単位を表す
  - パラメータをダイアログで入力するだけで即実行
- アイコン間のリンクで処理フローを表現(コーディング不要)
- Rの関数を知らなくても分析が可能
  - 「どう書くか」ではなく「何を分析するか」に集中できる

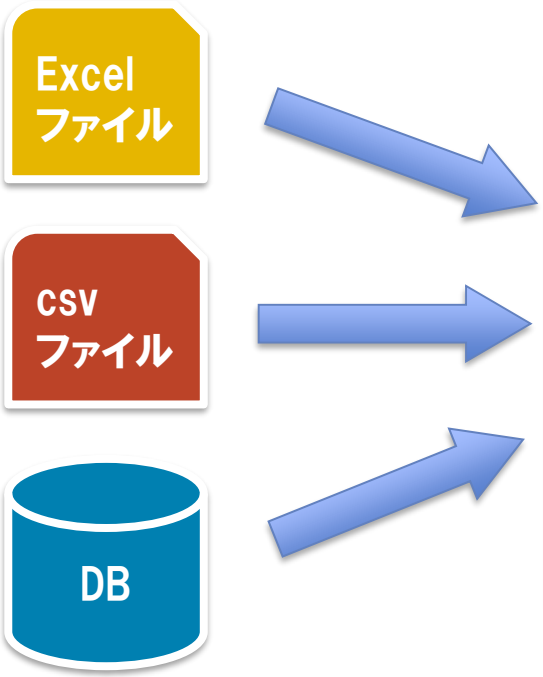


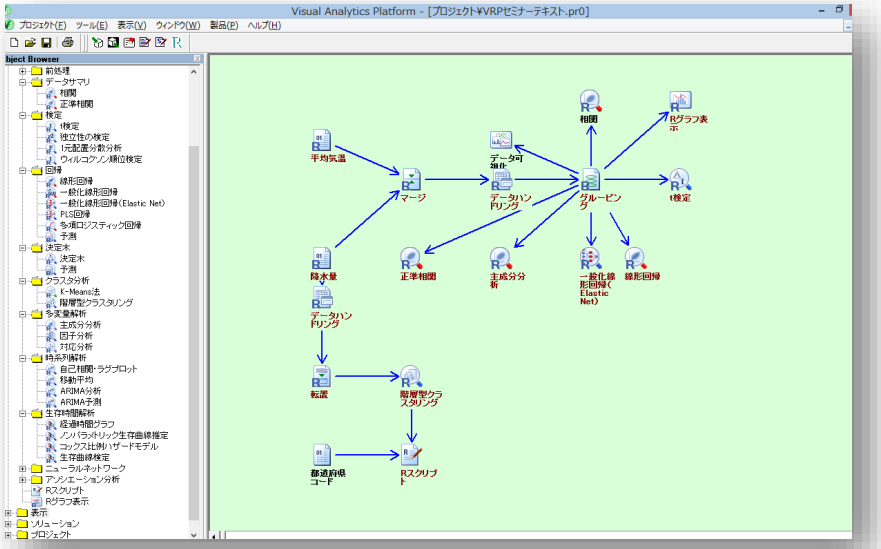
Copyright © 2019 NTT DATA Mathematical Systems Inc.

手軽にデータ取り込み

NTT DATA
NTT DATA Mathematical Systems Inc.

- データを簡単にVRPに取り込めます
  - ファイル（csv, excel他）をドラッグ＆ドロップ
  - 取り込み時のオプション指定もラクラク
- 各種DBからのデータ取り込みもサポート

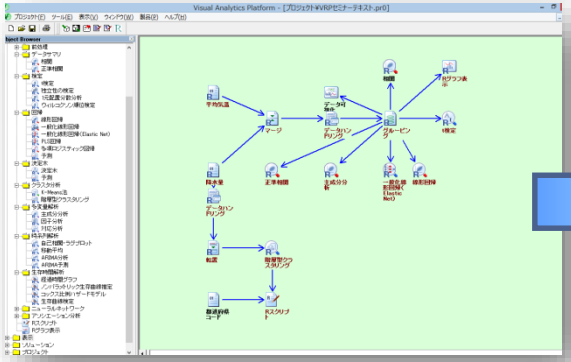


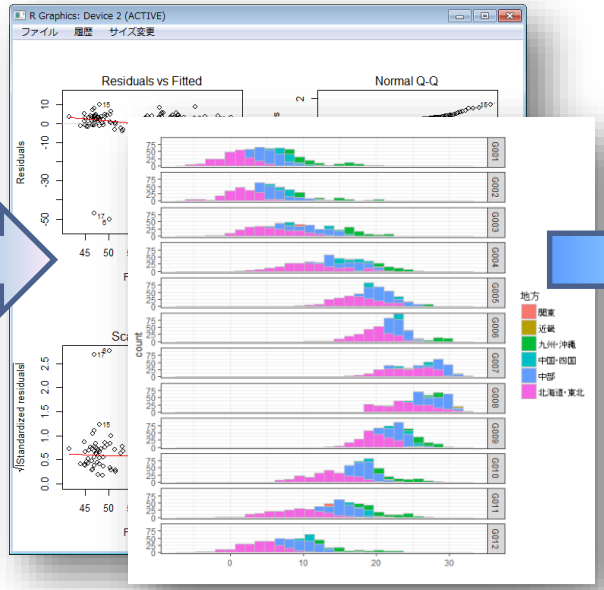


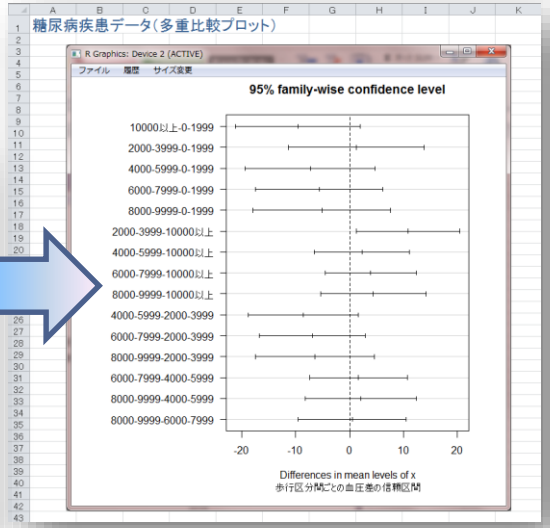
グラフ作成も簡単

NTT DATA Mathematical Systems Inc.

- ワンクリックで視覚化
  - Rグラフ表示アイコンで見栄えのするグラフを簡単に作成
  - オブジェクトビューによる分析結果の視覚化
- レポート、文書への貼り付け
  - 各種の画像ファイルに変換可能。資料作成にもご活用いただけます







Excel, Word, Powerpoint等

Copyright © 2019 NTT DATA Mathematical Systems Inc.

Rグラフ表示

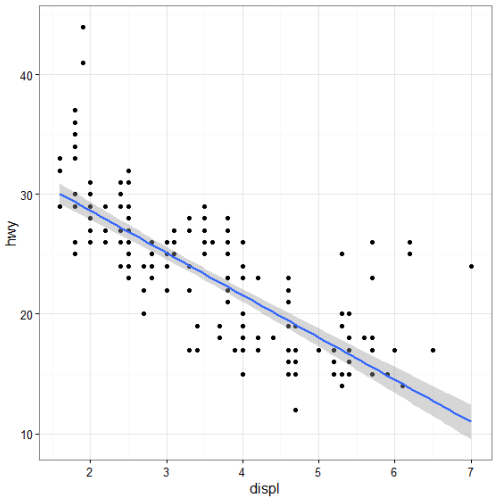
NTT DATA
NTT DATA Mathematical Systems Inc.

- VRP上のグラフ表示アイコン
  - 美しく、見栄えのするグラフを簡単に作成
  - 通常のVRPアイコンと同様に、Rソースエクスポート機能でR言語のソースコードに変換可能
- 加えて、Visual Analytics Platform独自のグラフも描画可能です

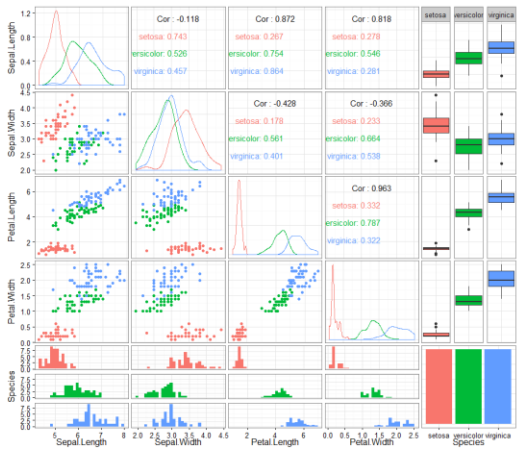
グラフの種類と主な機能	
散布図	散布図を描く. 回帰直線等のプロットも可能
折れ線	複数個の系列を同時にプロット
箱ひげ	カテゴリごとに分類してプロットする機能
対散布図	相関係数、箱ひげ図、ヒストグラムも同時に描画
ヒストグラム	カテゴリごとのプロット機能有り
棒グラフ	カテゴリごとのプロット機能有り
円グラフ	Pieチャートの描画
hi-lowグラフ	株価変動などに使われるグラフ

Copyright © 2019 NTT DATA Mathematical Systems Inc.

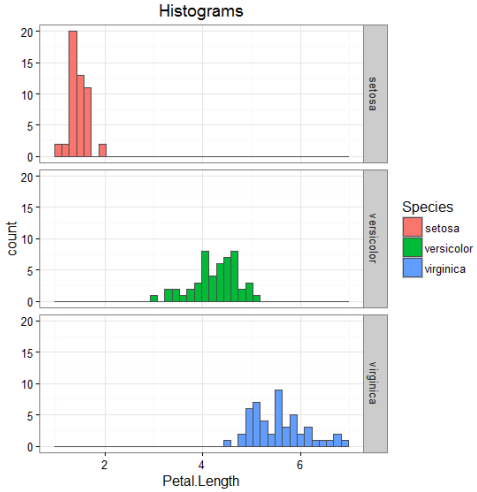
Rグラフ表示アイコンのプロット例



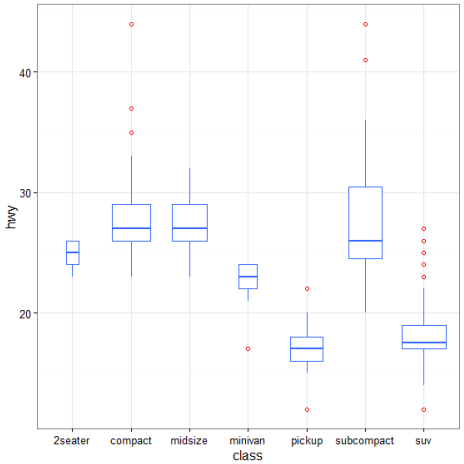
散布図(回帰直線付き)



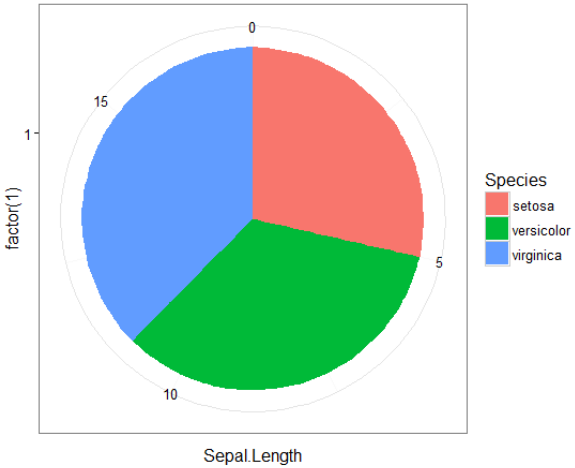
対散布図



ヒストグラム



箱ひげ図



円グラフ

Copyright © 2019 NTT DATA Mathematical Systems Inc.

Visual R Platform の統計解析機能

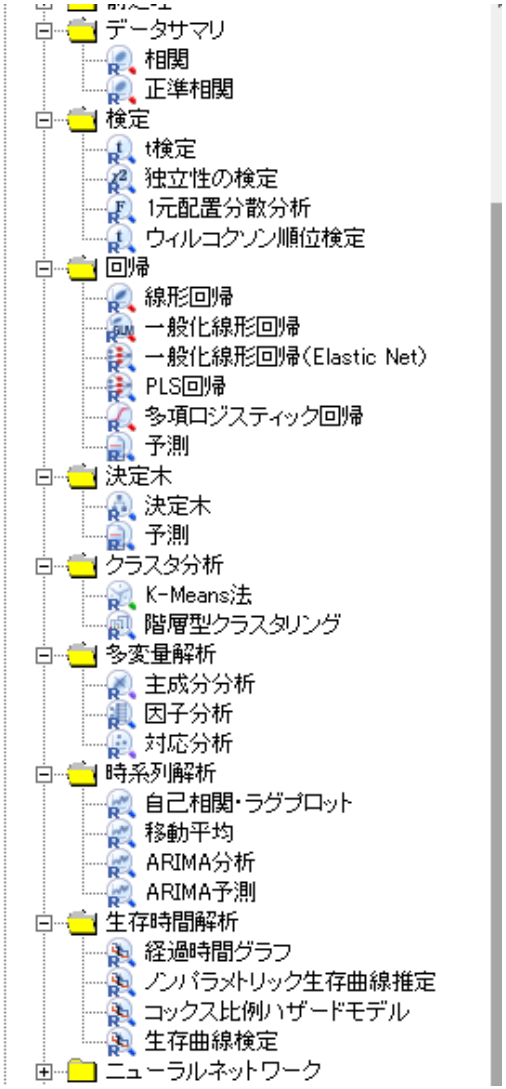
NTT DATA
NTT DATA Mathematical Systems Inc.

統計解析アイコン

- 基本的な統計解析機能は、既存のアイコンを利用
  - 検定、線形回帰、時系列解析など
- アイコン化されていない機能は、Rスクリプトアイコンを利用
  - 並列処理や要望に応じた機能追加（Rライブラリ、階層ベイズ、他機械学習）が利用可能
  - Rのグラフ関数も利用可能

R の持つ豊富な統計計算リソース、日々増加している機能やパッケージをフルに活用することができます

オブジェクトブラウザ内の統計解析アイコン一覧



Copyright © 2019 NTT DATA Mathematical Systems Inc.

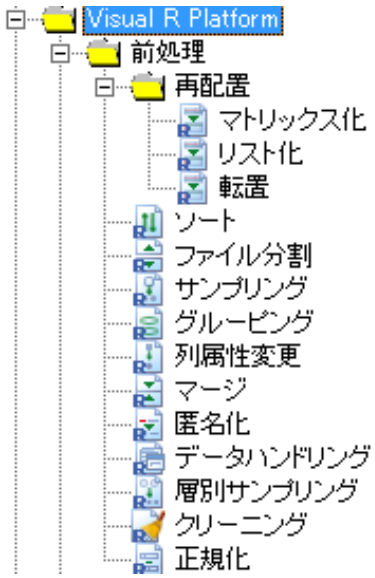
Visual R Platform のデータ処理機能

NTT DATA
NTT DATA Mathematical Systems Inc.

前処理アイコン

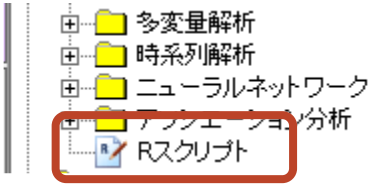
- 基本的な前処理加工、グラフ作成機能は、既存のアイコンを利用
  - マージやグルーピング、サンプリングなど
  - 平均や分散などの集計機能は「データハンドリング」アイコンで利用

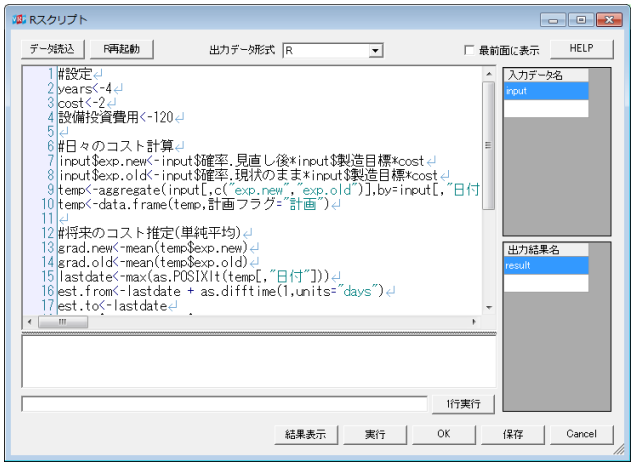
オブジェクトブラウザ内の前処理アイコン一覧



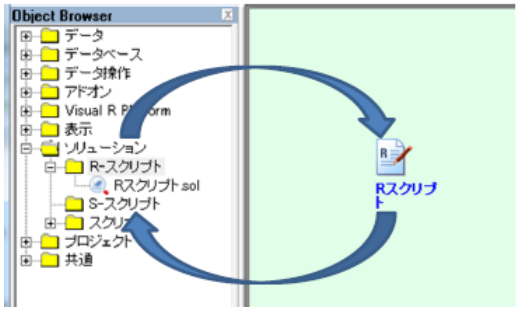
Rスクリプトアイコン

- アイコン以外の処理は、Rスクリプトアイコンを利用
  - Rスクリプトアイコンでカスタム対応も可能





作成したRスクリプトアイコンは、アイコン化して再利用可能！



Copyright © 2019 NTT DATA Mathematical Systems Inc.

Rスクリプト自動生成

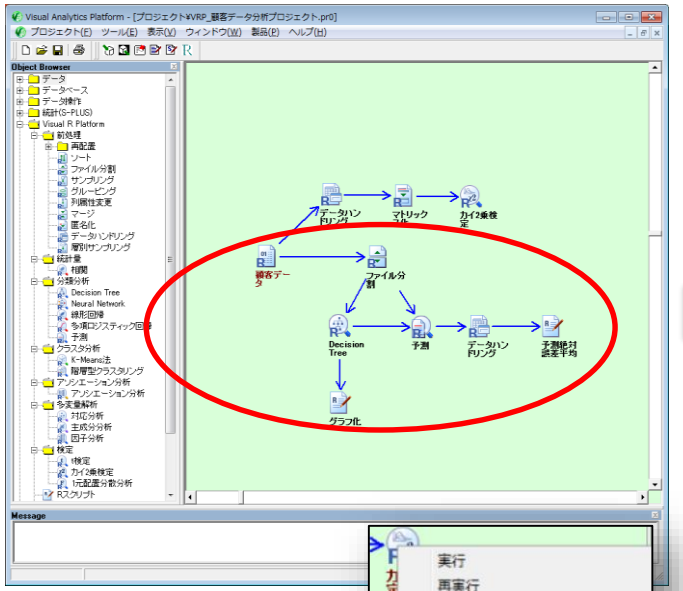
NTT DATA Mathematical Systems Inc.

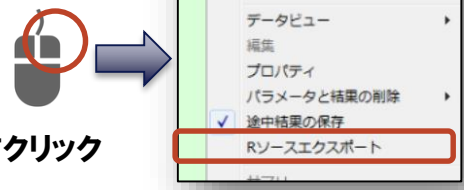
- Rソースエクスポート…プロジェクトの内容をR言語のスクリプトとして出力

右クリックメニューから R ソースエクスポートを実行

R スクリプトを自動生成

R がインストールされていれば ほとんどの環境で実行可能！（VRP 不要）



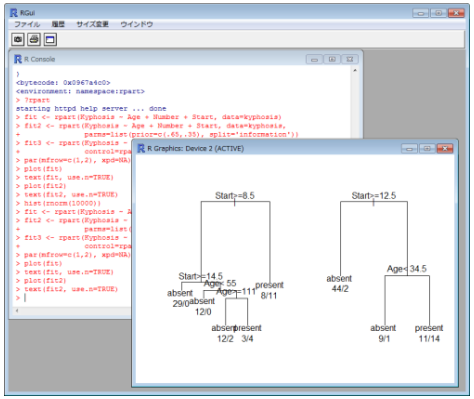


右クリック

```
# ****************************************
#  VRP専用スクリプトロード
# ****************************************
if (!exists("vrpEnv")) {
                                    vrpEnv=new.env()
                                    source("VRPlib.R", local=vrpEnv)
                                    attach(vrpEnv)
}

# ****************************************
#  実行部
# ****************************************
# R顧客データ
vrp1 = vrp.RData(rdataFile="R顧客データ_1.RData")
# サンプリング
vrp6 = vrp.RSampling(input=vrp1[["noName"]], target=c("性別", "年代", "職業", "利用時間
", "利用回数", "収益"), n.samp="100%", method="先頭から", key=NULL, seed=NA,
n.inter=0)
# ファイル分割
vrp4 = vrp.RDivide(input=vrp6[["result"]], rate=c(7,3), seed.auto=F, seed=0)
# 線形回帰
vrp5 = vrp.RLinearModel(x=vrp4[["learn"]], dependentColName = c("収益"),
independentColName = c("性別", "職業", "利用時間", "利用回数"), constantTerm = TRUE,
weightColName = NULL, dir = "both")

print("result is stored in vrp5")
```



活用方法の幅が広がる！

- ✓ 分析処理の記録
- ✓ スクリプトの共有
- ✓ Linux上での定型処理化

etc・・・

Copyright © 2019 NTT DATA Mathematical Systems Inc.

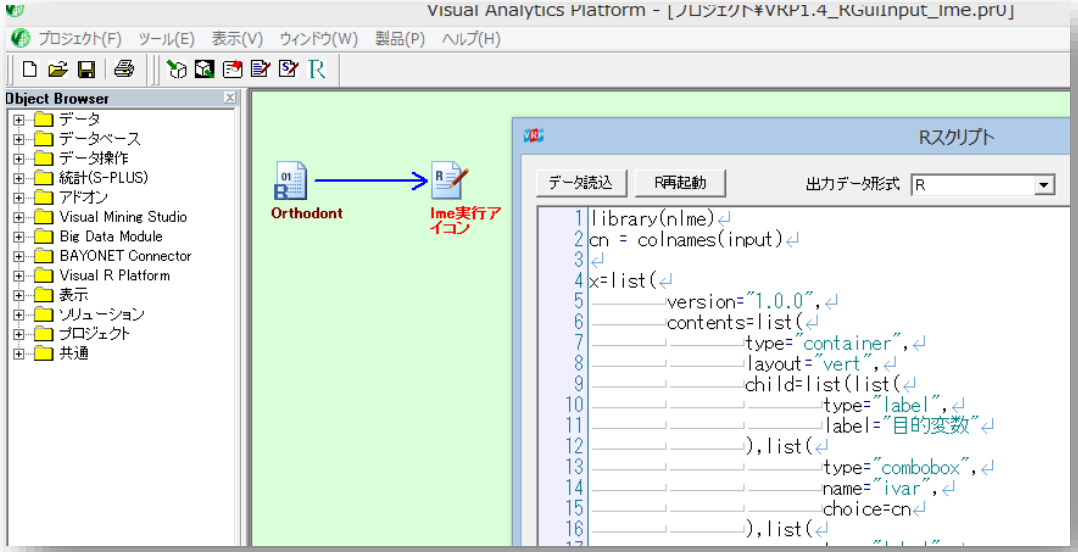
RGuiInput

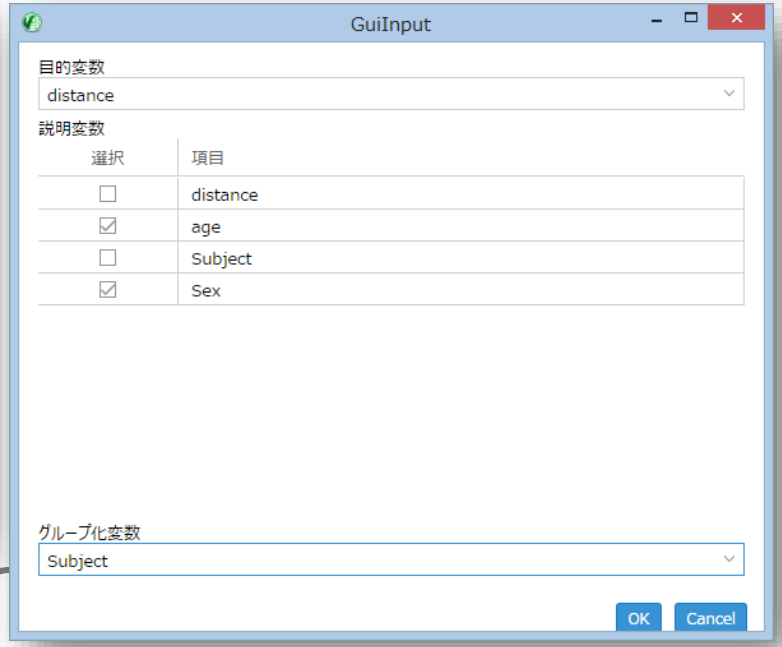
NTT DATA Mathematical Systems Inc.

- 簡易ユーザーインターフェース作成機能:簡易なスクリプト記述で標準アイコン同様のダイアログを表示、値の指定が可能に

スクリプトアイコンを編集し

実行すると、ユーザーインターフェース起動。パラメータ指定などがリスト、コンボボックスで実現





分析手法の幅が広がる！

✓ 専門性の高い分析用のカスタマイズを行ったアイコン作成が可能

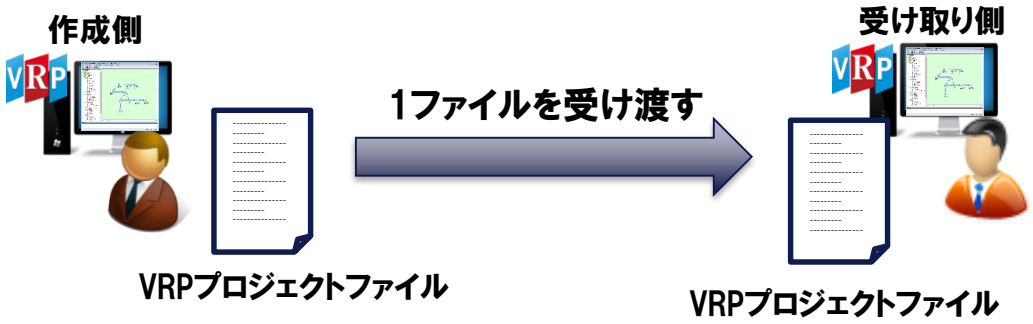
Copyright © 2019 NTT DATA Mathematical Systems Inc.

分析の共有

NTT DATA Mathematical Systems Inc.

分析環境の共有

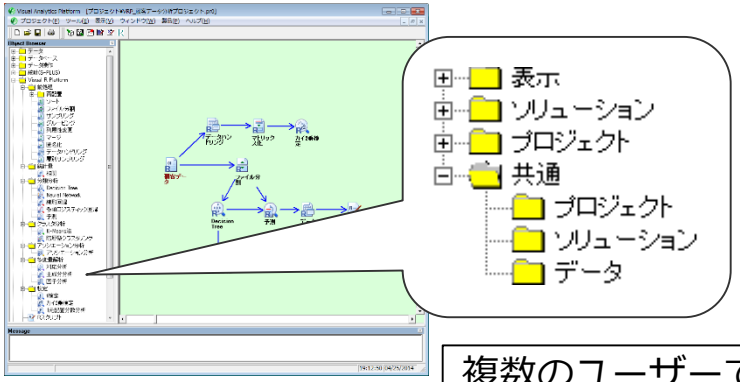
- 複数名でご利用の場合 その１…分析フローをファイル化して受け渡す



こんな時に有効！

- ✓ **作成したプロジェクトを異なるマシンの間で交換したい**
- ✓ **遠く離れた人に、添付ファイルの形でプロジェクトを渡したい**
- ✓ **作成したプロジェクトのバックアップを取りたい**

- 複数名でご利用の場合 その2…共有フォルダにプロジェクトファイルを保存する



複数のユーザーで共有可能な領域を作成
データも分析フローも簡単に共有できます

こんな時に有効！

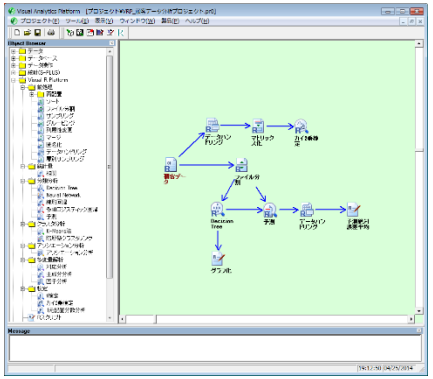
- ✓ **ユーザ共通のネットワークサーバに接続できる環境**
- ✓ **作成したデータ、スクリプトの交換を手間なく実施したい**

Copyright © 2019 NTT DATA Mathematical Systems Inc.

定型処理(バッチ)実行

NTT DATA Mathematical Systems Inc.

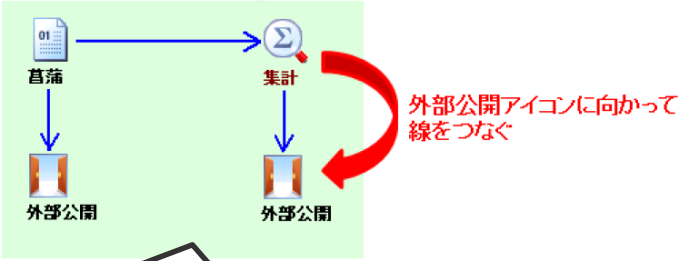
- 繰り返し頻繁に行う処理を定型（バッチ）処理化



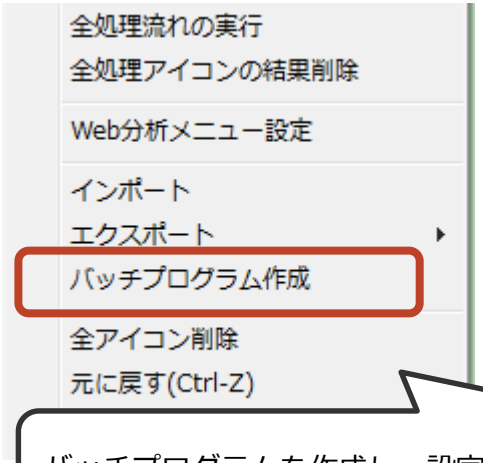
こんな時に有効！

- ✓ **他システムから計算エンジンとして呼び出したい**
- ✓ **日常的に繰り返す定型処理をスケジュールを組んで実行したい**
- ✓ **複数ファイルに対して同じ処理を実行しておき、結果だけを確認したい**

VRPプロジェクト作成後



プロジェクト上で右クリック「外部公開」を選択し、バッチ実行時に内容を変更したいアイコンと外部公開アイコンを矢印でつなぎます。



バッチプログラムを作成し、設定、実行します

日次バッチ、月次バッチなど、レポーティング機能としても活用可能！グラフ画像の出力も可能！

Copyright © 2019 NTT DATA Mathematical Systems Inc.

VRP 1.7の新機能

NTT DATA Mathematical Systems Inc.

Copyright © 2019 NTT DATA Mathematical Systems Inc.

特定分野向けソリューションのご提供

NTT DATA
NTT DATA Mathematical Systems Inc.

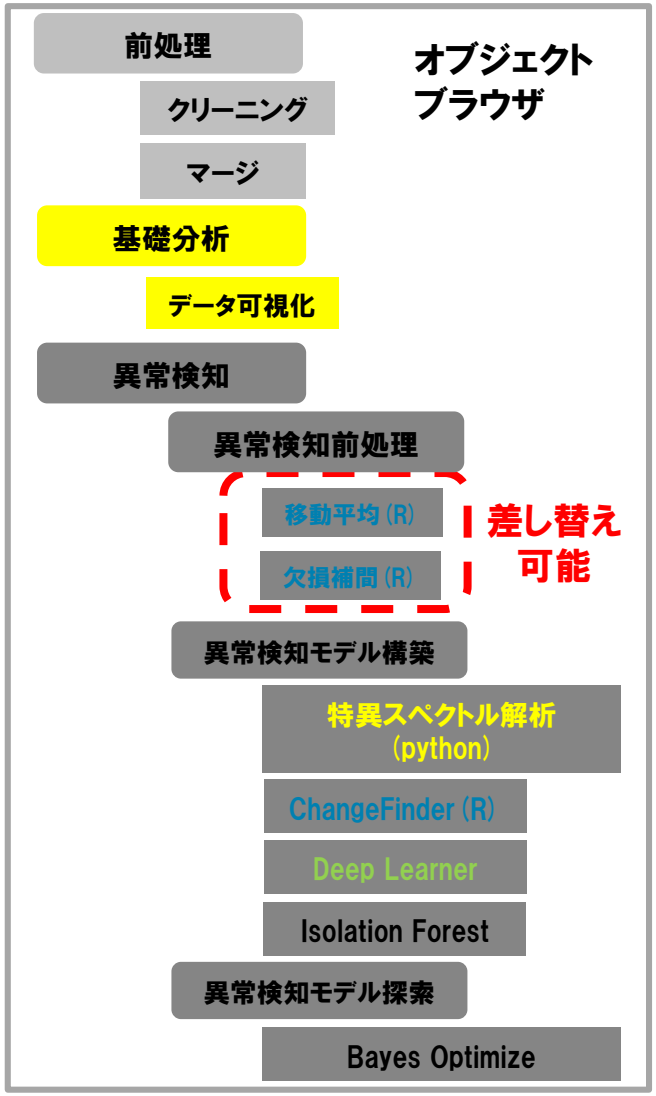
- VRPでは、従来の汎用的な分析機能をアイコンの形でご提供する他に、特定分野に特化したソリューションをVRPのプロジェクトファイルの形でご提供していく予定です.
- ＜特徴＞
  - 用途に応じた典型的な分析フローをご提供
  - 入力データを差し替えるだけで一連の分析が行えます
    - 前処理から基礎分析、モデル構築、結果の可視化まで
    - 表・グラフ・コメントを一括で管理可能
    - 最小の手間で精度向上を行えます
- 第1弾として、VRP 1.7では異常検知ソリューションに注力いたします.
- ＜想定適用分野の例＞
  - 生産ラインにおける異常検知（センサーデータからの異常判定）

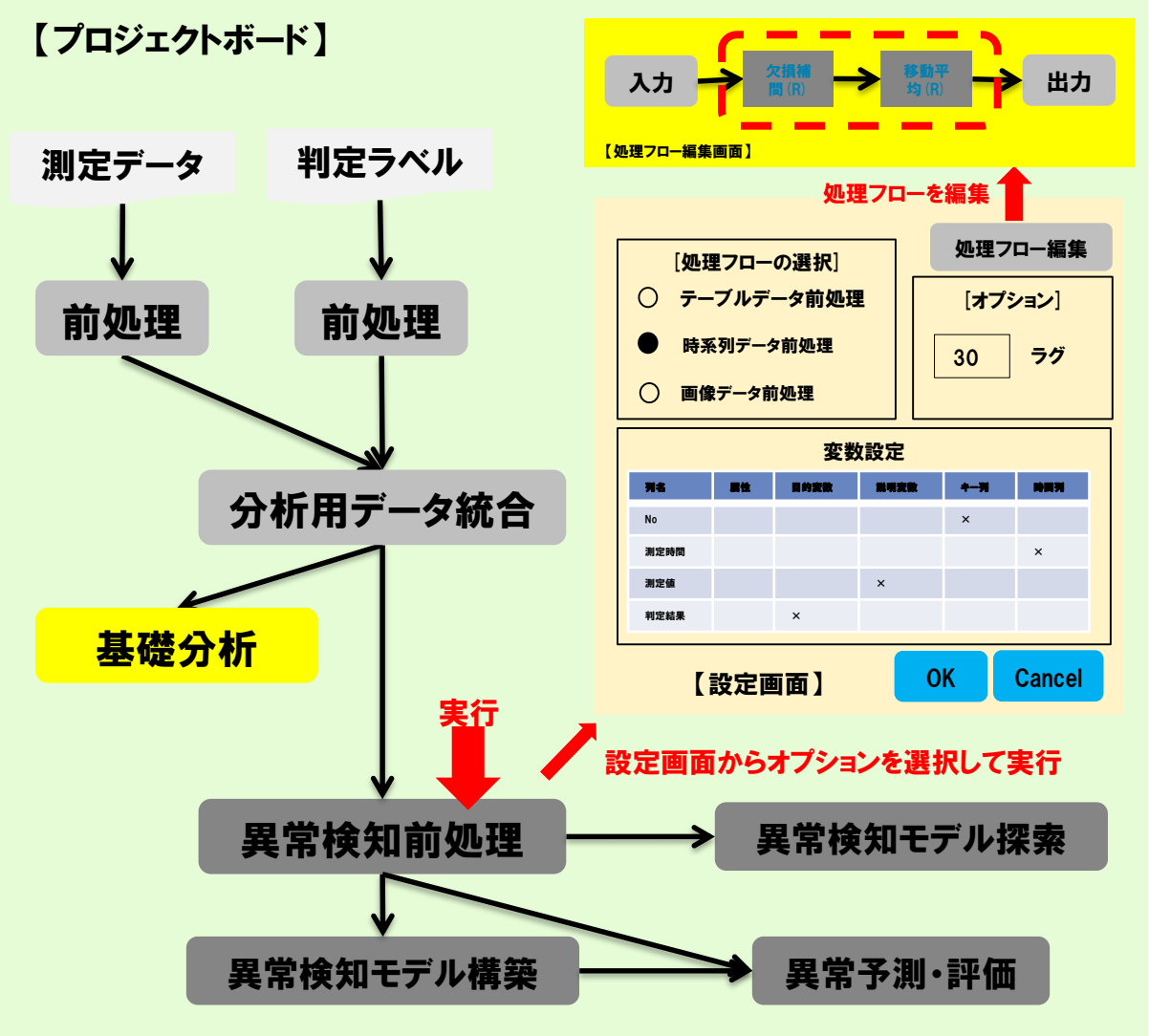
詳細は当日ご紹介します

Copyright © 2019 NTT DATA Mathematical Systems Inc.

ソリューション構成（イメージ図）

NTT DATA
NTT DATA Mathematical Systems Inc.





Copyright © 2019 NTT DATA Mathematical Systems Inc.

ソリューション事例

NTT DATA Mathematical Systems Inc.

NTTデータ数理システム
技術者インタビュー | Rユーザ向け分析プラットフォーム Visual R Platform

センサーから読み解く
製造工程の異常検知に活用

働き方改革や人口減少に伴う人手不足に対応するため、AIの活用を検討する企業が増えています。一方でAIといえども、使う人間に確かな知見や意思がないと狙った結果は得られません。また、多くの企業がツールの導入ではなく、活用で挫折することも報告されています。NTTデータ数理システムでは、AI活用への取り組みに不安がある方向けにサポートサービスを行っています。業務知識や自社のデータに対する知見はあるが分析経験がなく、しかし成果を求められている担当者に対して、その目的やデータに合わせ、幅広い経験や知識を持つ技術者が、親身に相談にのります。その技術者の1人、データマイニング部シニアアナリストの中道祐希は、センサーデータを利用した異常判別を実現した際に、Visual R Platform (VRP) を活用。その分析の進め方について聞きました。

NTTデータ数理システム
データマイニング部
中道 祐希

Interview

分析前のデータ整理も集計も。VRPで現状把握

相談者がどのようなことに困っていて、何を解決しようとしていたのかを教えてください。

中道 ご相談者は大手メーカーのICT部門に所属される方で、データを活用することで、さらなる品質向上につながる「気づき」が得られないか、という課題をお持ちでした。製造現場で記録されたデータの一部を持参されたので、お持ちになったデータについてヒアリングを行いました。
データは時系列に記録された100種類以上のセンサーデータで、一定の時間間隔で測定されており、1レコード1000時点前後あります。1レコード毎に、このようなテーブル形式で保存されていました。それぞれのレコードには、別途費用をかけて行ったテスト結果（合格、不合格）が紐づいていました。
分析の目的は「不合格となる原因が何か知りたい」「新たなセンサーデータから、テストをしなくても合否が判定できるようにしたい」ということでした。
ご相談者はすでに、ご自身で目星をつけたセンサーのグラフ化を行っていましたが、不合格に結びつくセンサーや規則性は発見できていませんでした。

問題の難しさを教えてください。

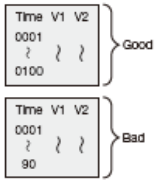
中道 通常の判別問題は1行に原因となり得るデータと、結果のデータが保存されている状態で分析を行い、結果に対する関連性を特定し、予測を行います。そのための手法はさまざま開発され、代表的なものにはDecision Tree（決定木）やニューラルネットワークがあります。

PROFILE

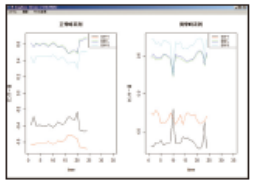
中道 祐希

2004年入社。異常検知を含む機械学習全般に関する受託分析に従事。近年は異常検知ツールの開発、コンサルティング、データサイエンティスト教育プログラムなどの自社ソフトを用いた顧客の支援を中心に取り組む。

分析データのイメージ



長さの異なる時系列データを結合し、欠損の値をつける



時系列のパターンをグラフで把握

［裏面に続く］

詳細は配布資料をご覧ください

Copyright © 2019 NTT DATA Mathematical Systems Inc.

数理システム製品間の連携

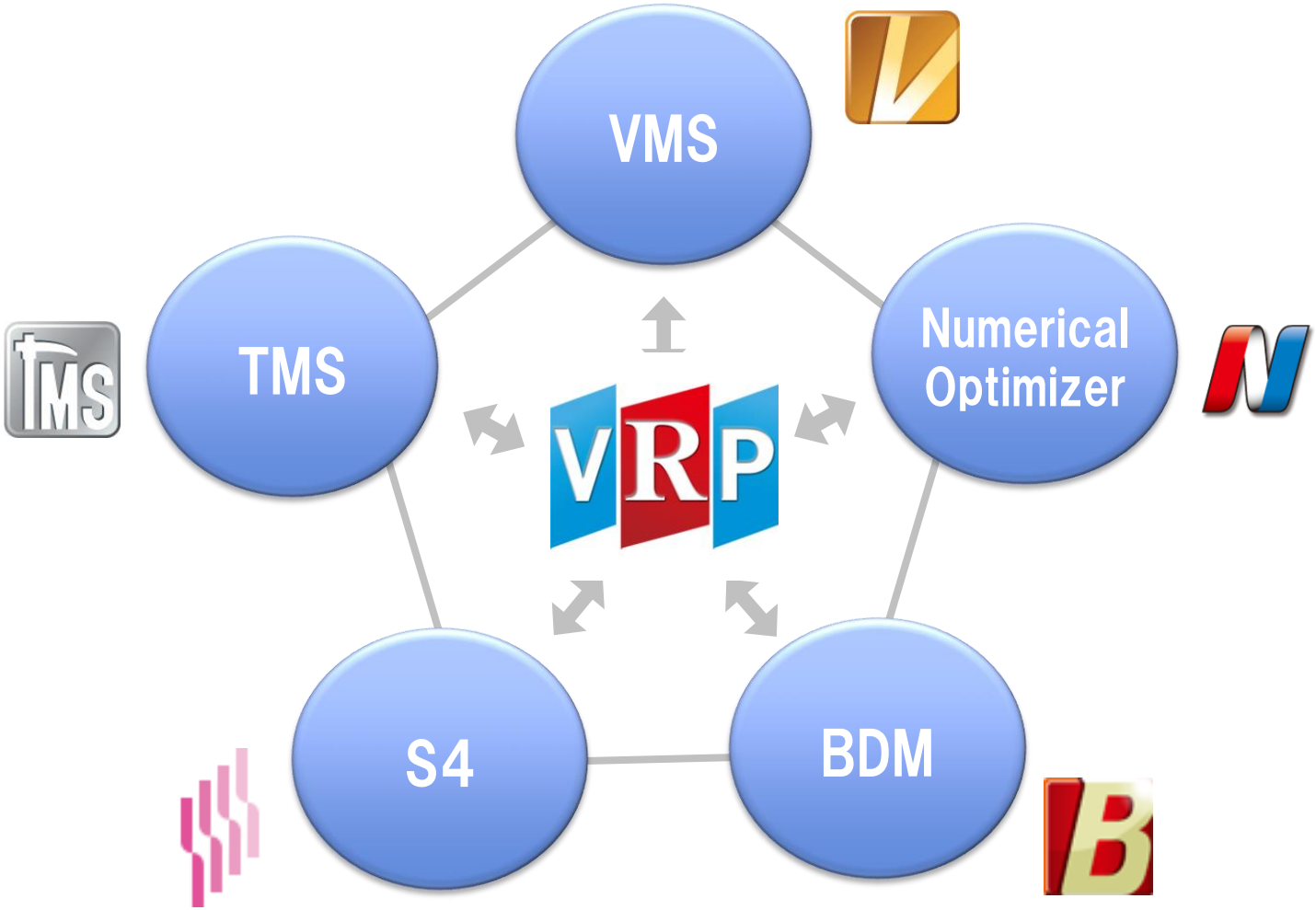
NTT DATA Mathematical Systems Inc.

Copyright © 2019 NTT DATA Mathematical Systems Inc.

数理システム製品群との連携

NTT DATA Mathematical Systems Inc.

- VMS, TMS等の数理システム他製品との連携が可能
  - 複数製品の連携で、さらに高度・多様な分析を



Copyright © 2019 NTT DATA Mathematical Systems Inc.

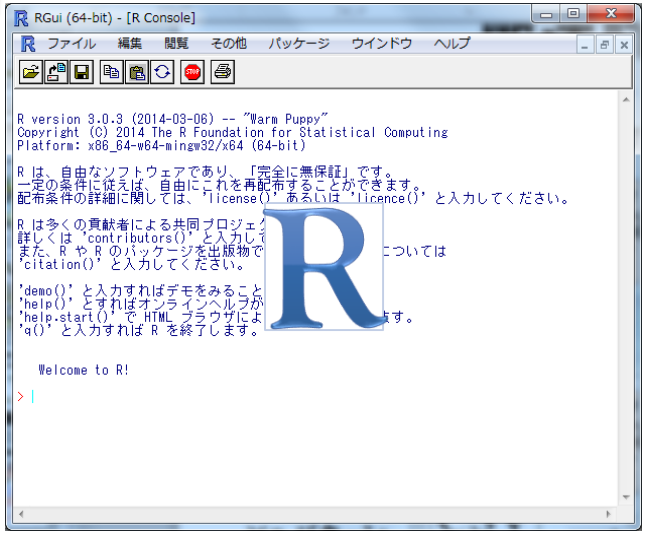
製品連携例（1）：RNUOPT + VRP

NTT DATA
NTT DATA Mathematical Systems Inc.

RNUOPTとは？
Numerical Optimizer **をエンジンとした**
最適化機能を提供する R **のパッケージ**

R **のデータ加工、統計解析機能と**
Numerical Optimizer **での最適化機能が融合**

Numerical Optimizer



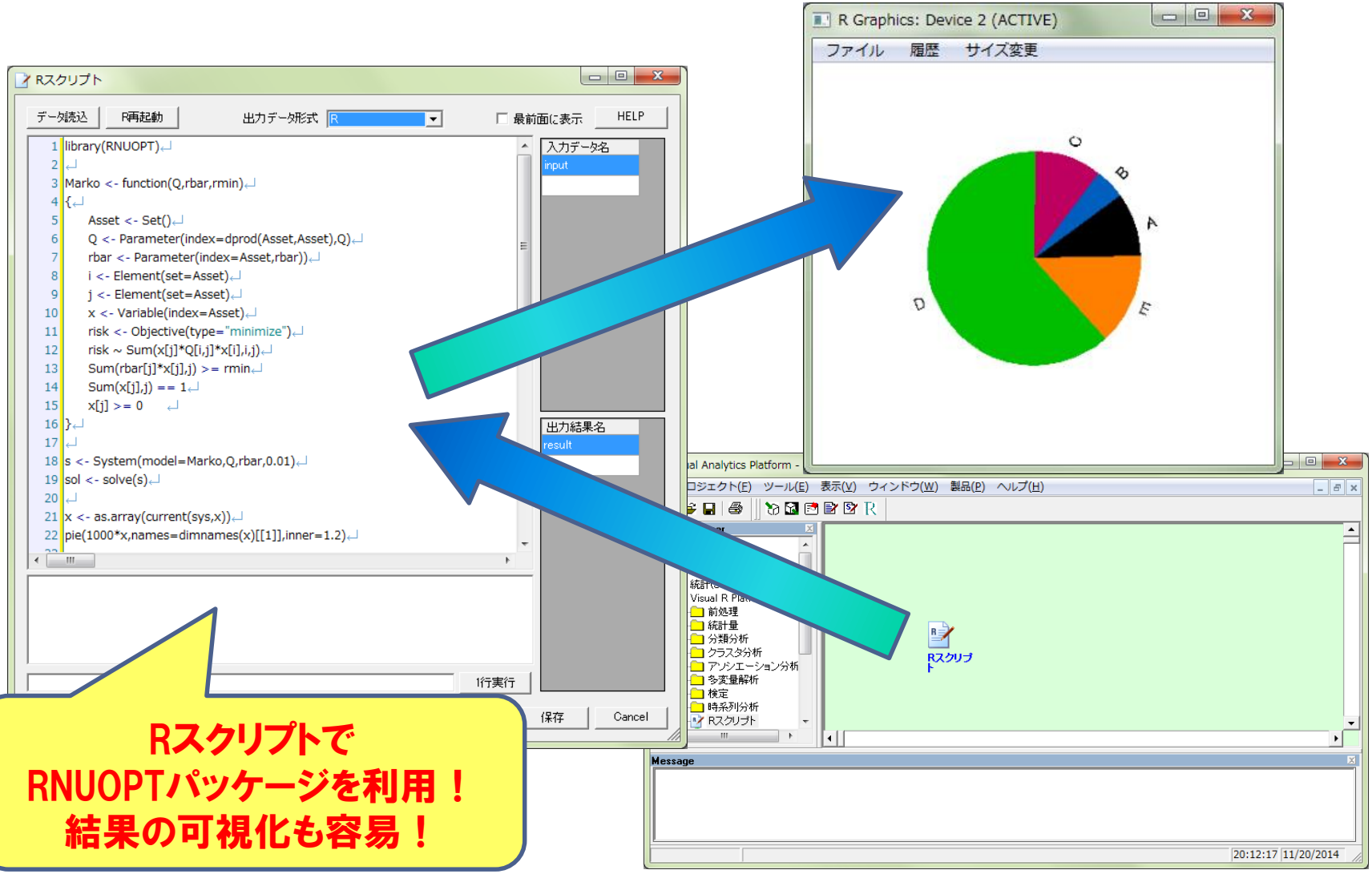
R **のインターフェースで**
Numerical Optimizer
を使用できます

もちろん
Visual R Platform
からも使用可能！

Visual R Platform

Copyright © 2019 NTT DATA Mathematical Systems Inc.

実行例



Rスクリプトで
RNUOPTパッケージを利用！
結果の可視化も容易！

お問い合わせ先

NTTデータ 数理システム
営業部またはデータマイニング部まで

毎月体験セミナー開催中！

R全般のサポート、
コンサルティング、Rによる
システム開発も承ります

何でもご相談ください！

NTT DATA
株式会社NTTデータ 数理システム

TEL : 03 - 3358 - 6681
FAX : 03 - 3358 - 1727
【URL】 http://www.msi.co.jp
【E-mail】 vrp-info@msi.co.jp

Copyright © 2019 NTT DATA Mathematical Systems Inc.