



1. 「Data Set」でデータ名を指定
2. 「Source Column」で水準を変換したい因子の列名を指定
3. 「Target Column」で新しい水準を保存したい列名を指定

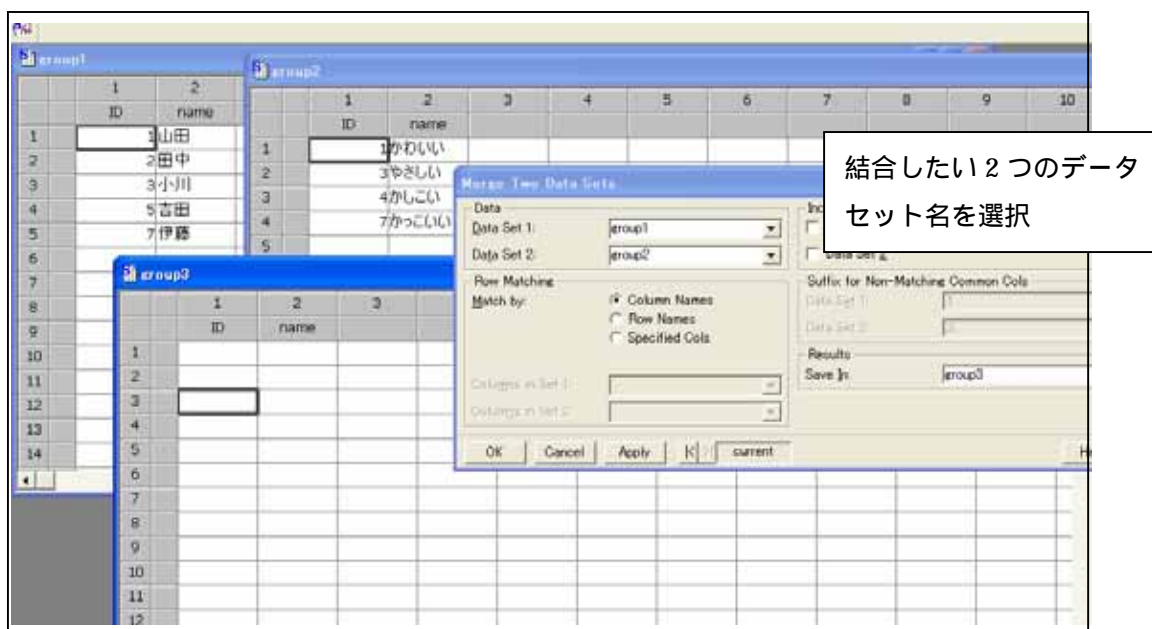
[Count]Level に表示されている数値と値はそれぞれ、因子水準と度数を表しています。結合したい因子水準をここから、CTRL キーを押しながら、選択します。

4. 「Combined Level」に新しい水準の値を指定（ここでは fruit）

## その2 . 2つのデータを結合したい！

ID をキーにして2つのデータを結合させたい

メニューバー「データ / マージ」

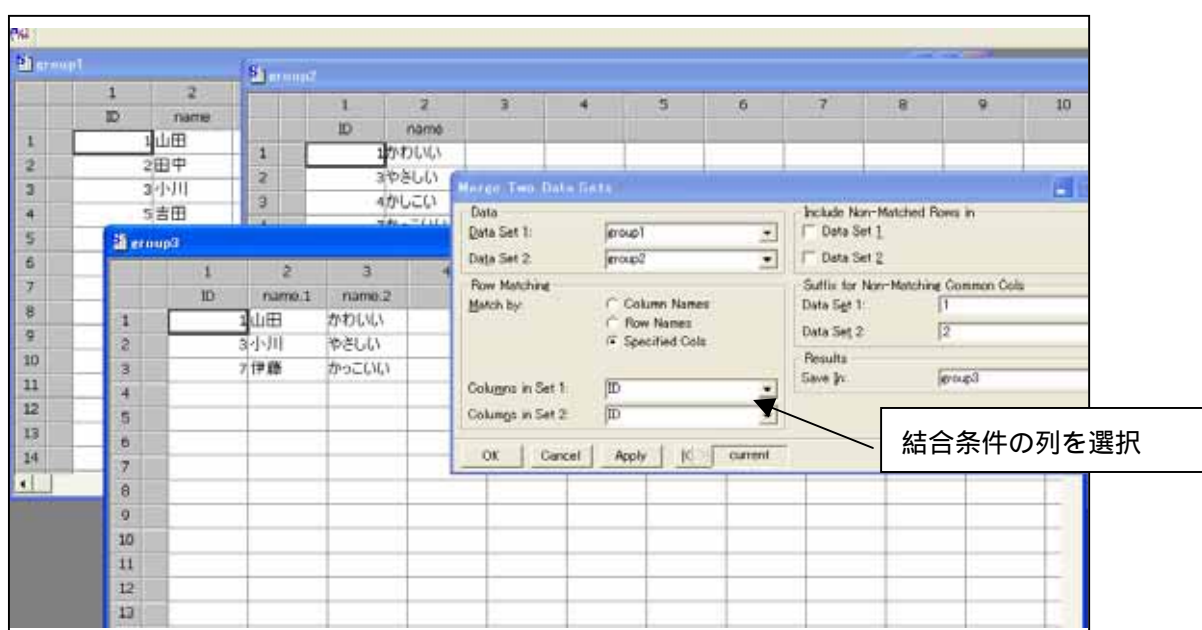


結合したいデータセットの名前を選択すると新しいデータセットは、ID と name という列名を持つ、空のデータセットになっています。これは、デフォルトで、結合条件を「同じ名前の列が

等しいデータを結合させる」となっているためです。上の例では、ID, name どちらも group1, group2 に存在するので、「ID が等しくてかつ、name が等しいデータを新しいデータとする」とされてしまいます。

ID が等しいデータを結合するには (name を結合条件にさせないためには)

1. 「Match by」で「Specified Cols」にチェック
2. 「Columns in Set1」から ID を選択
3. 「Columns in Set2」から ID を選択



どちらにも共通に存在するデータが新しいデータセットとなります。

どちらかのデータを元のテーブルとし、もう一方を参照テーブルとして、元のテーブルに含まれるデータは参照テーブル側にあってもなくても新しいデータにしたい場合は(Excelのvlookup相当)、「Include Non-Matched Rows in」で元のデータのほうにだけ、チェックを入れます。

## グラフィックス編

サンプル・データセット `fuel.frame` (車のデータ) を使い、いろいろなグラフを作成してみましょう。

### ・ 散布図と回帰直線

1. `Weight` 列、`Fuel` 列を CTRL キーを押しながら選択、「回帰直線」ボタンをクリック
2. 回帰直線をクリックして、右ボタン、メニューから「Curve Fitting」を選択
3. 「Poly. Order」を 1 から、好きな整数に変えることにより、回帰式の次数を変えることができます
4. タブ「By Conf Bound」で信頼区間の水準を選択、右側から線の種類を選ぶことにより、信頼区間を表示させることができます
5. メニューバー「挿入 / 曲線近似式」により、回帰式をグラフ上に表示させることができます

### ・ 対散布図

1. すべての列を選択して (左上ボタンをクリック) 「対散布図」ボタンをクリック
2. 点をクリックして右ボタン、「Smooth」を選択
3. 「Smoothing Type」を選ぶと、回帰直線やその他の近似線をひくことができます

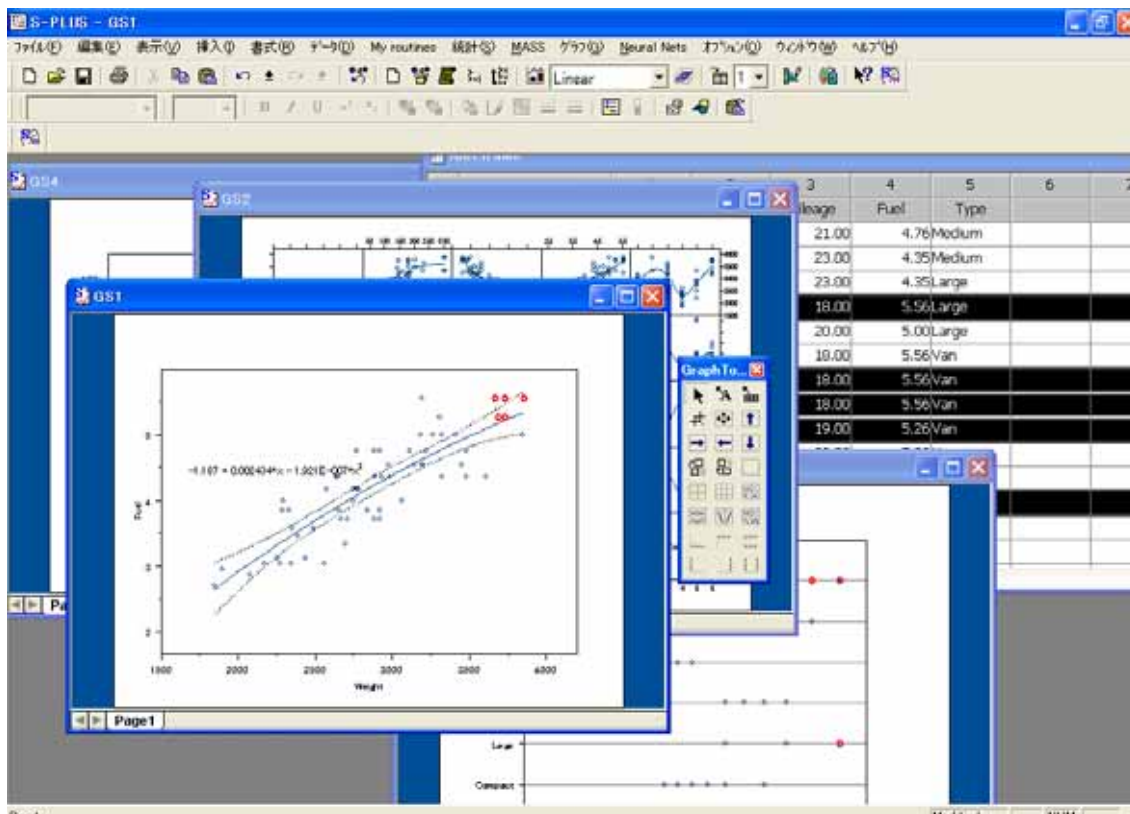
### ・ 点グラフ

1. `Fuel` 列、`Type` 列を CTRL キーを押しながら選択、「点グラフ」ボタンをクリック

## ・パレート図

### 1. Type 列選択、「パレート図」ボタンをクリック

作図したところで、ツールバー「グラフツール」ボタンをクリックし、開いたパレットの中から、「矩形内のデータ選択」ボタンをクリック、ひとつのグラフに移動し、マウスで点を囲んでドラッグ、ドラッグされた点は反転表示されますが、他のグラフ、ならびにグラフの元になっているデータシート上でも同様に反転表示されています。



## 統計解析編

主成分分析を取り上げ、便利な使い方をご紹介します。

グラフ作成に使った `fuel.frame` をここでも取り上げます。

1. メニューバー「統計 / 多変量解析 / 主成分分析」
2. 「Data Set」で `fuel.frame` を選択、あるいは入力
3. 「Variables」でとりあえず `<All>` を選択

これだけでも実行可能ですが、さらに、主成分負荷量 (Loadings) の表示と、バイプロットの作図も試します。

4. 「Results」タブをクリック、「Loadings」にチェックを入れる
5. 「Plot」タブをクリック、「Biplot」にチェックを入れる

結果、レポート・ウィンドウとグラフ・ウィンドウが開き、レポート・ウィンドウには各主成分の標準偏差、あるいは、その比率、各変量の負荷量 (係数) が表示されます。グラフ・ウィンドウには、第 1 主成分と第 2 主成分の散布図と、各変量の負荷の方向と大きさを表すバイプロットが表示されます。

負荷量を見ると、第 1 主成分で `weight` という変数のみが大きく得られています。これはごく自然な結果で、ここで取り上げたデータの変量はそれぞれ、単位が異なり、変数 `weight` は他の変量に対して、非常に大きな量をとっているからです。また、注目すべきは質的な変数 `Type` に対しても負荷量が求まっていることです。これは、因子水準に対して対比 (ダミー行列のようなもの) を作った結果です。

第 1 主成分以外を使って作図するには、「Plot」タブにある「Biplot Which Scores」でデフォルトの `c(1,2)` から、`c(2,3)` (第 2 主成分と第 3 主成分) のように変更することにより可能です。

では、主成分分析が元のデータの大きさに左右されないよう、デフォルトの分散共分散行列による解法ではなく、相関行列による解法を行うよう、指定します。

6. タブ「Model」の「Scaling」をデフォルトの「Covariance」から「Correlation」にする

7. 「Variable」から Type を除く

(このオプションを使う場合は、質的な変量、因子を除いてください)

その他、求まった主成分得点を Excel ファイルにエクスポートする方法などは S-PLUS のホームページ <http://www.msi.co.jp/splus/> にある「S-PLUS 利用での Q&A」にも一部のせています。

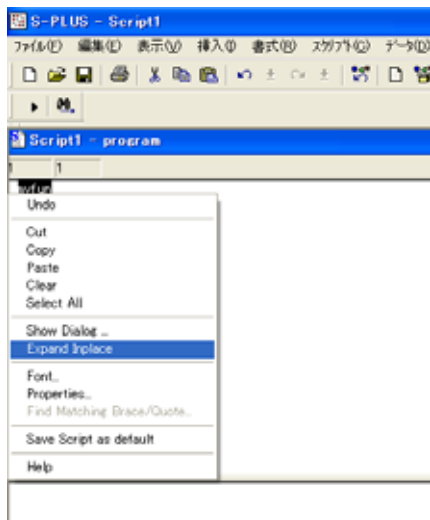
## スクリプト編

以前からの S-PLUS や S 言語のユーザーの方で、いまでもまったく GUI を使われない方もいらっしゃるかもしれません。関数を定義、あるいは修正する際にスクリプトを使われると非常に便利です。

・既存の関数定義を修正したい

1. メニューバー「ファイル / 新規ファイル」で「Script File」を選択
2. 修正したい関数の名前を入力、関数名をダブルクリックして（名前全体が反転表示される）右ボタン
3. 「Expand Inplace」を選択

すると、その関数の定義がスクリプト・ウィンドウ上に展開されます。



修正が必要な箇所を見つけ（ツールバーに「検索」ボタンがあります）、修正したら、ツールバー「実行」ボタンをクリック

これで、修正が反映されます。次回もこの方法で関数定義を修正されるのであれば、このスクリプト・ファイルは保存しておく必要がありません。

そのほか、スクリプト・ウィンドウ上で小規模なプログラムを書き、実行することにより、簡単なテストも可能です。その際、スクリプト実行中にエラーが出た場合、そのスクリプトで作成したオブジェクトはいずれも保存されません（変更は反映されません）。