R で学ぶデータ解析とシミュレーション

〔付録〕

～ R Commander の概要 ～

メニュー

- **R Commander の概要 ←**
- データ「iris」の読み込み方法
- グラフの作成方法
- R Commander でのデータ解析例
- グラフのカタログ
- 各メニューの概要

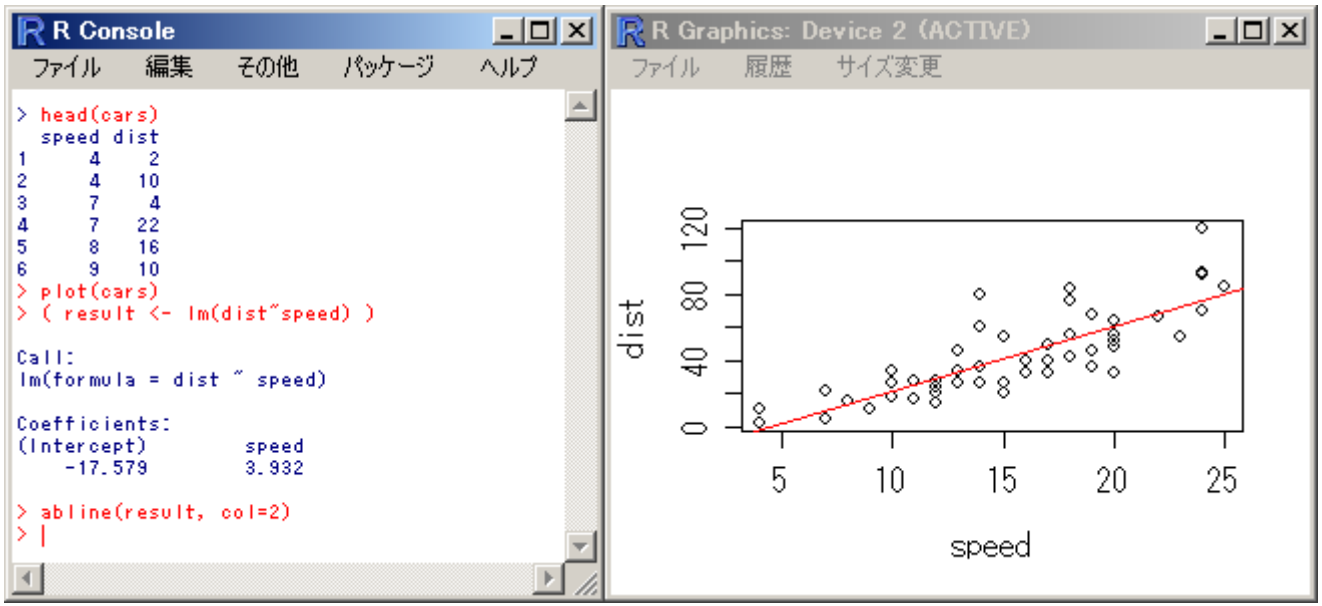
Graphic by (c)Tomo.Yun (http://www.yunphoto.net)

R の概要

- オープンソース＆フリーの統計解析用ソフト

【長所】

- 関数電卓，数値計算，プログラミング，統計解析，グラフィックスの機能があり，どの機能も充実している
- 機能拡張が容易に行える
- 使用人口が多いので，バグが少なく情報も豊富



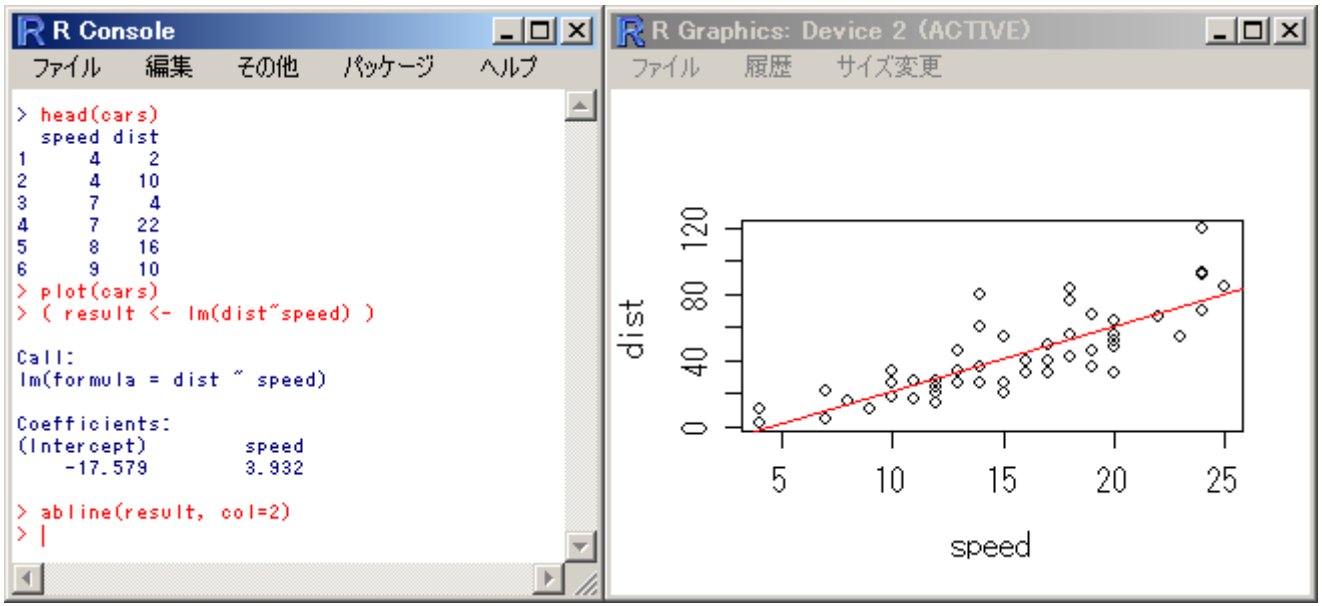
R の概要

- オープンソース＆フリーの統計解析用ソフト

【短所】

- 大規模なデータを扱う場合は多少骨が折れる
- EXCEL などの表計算ソフトに比べて GUI （マウス操作）の機能が劣っているので，R の命令をひとつひとつ覚えなければいけない…

⇒ そこで登場するのが GUI 版 Rである「R Commander」！



R Commander

R Commander

ファイル　編集　データ　統計量　グラフ　モデル　分布　ツール　ヘルプ

データセット: USJudgeRatings　データセットの編集　データセットの表示　モデル: <No active model>

スクリプトウィンドウ

```
t.test(USJudgeRatings$CFMG, alternative='two.sided', mu=0.0, conf.level=.95)
scatter3d(USJudgeRatings$CONT, USJudgeRatings$WRIT, USJudgeRatings$DECI, fit=c("linear","smooth"), bg=
scatter3d(USJudgeRatings$CFMG, USJudgeRatings$DECI, USJudgeRatings$DILG, fit=c("linear","smooth"), bg=
```

Submit

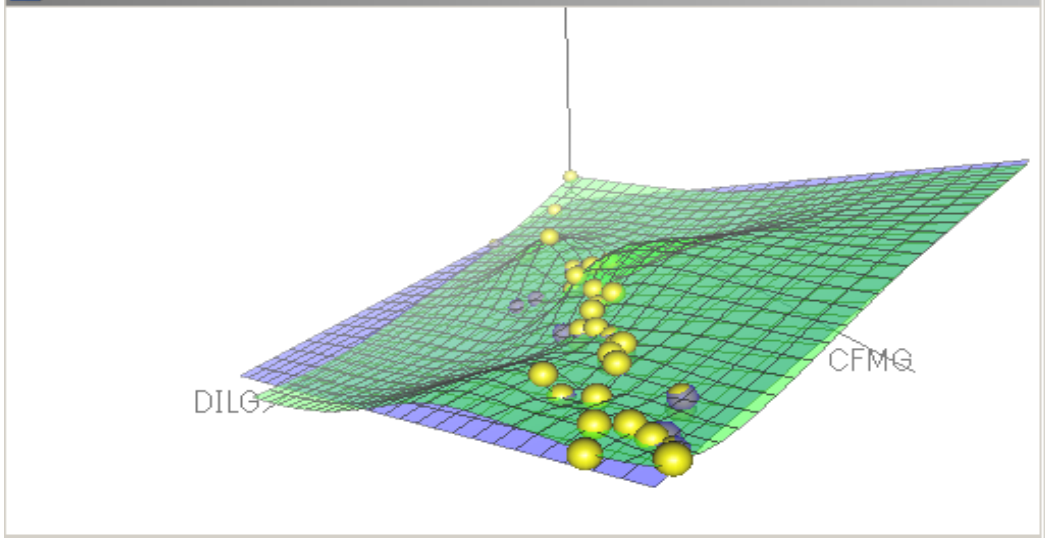
出力ウィンドウ

```
> t.test(USJudgeRatings$CFMG, alternative='two.sided', mu=0.0,

        One Sample t-test

data:  USJudgeRatings$CFMG
t = 57.0201, df = 42, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 7.214367 7.743773
```

RGL device 1 (active)



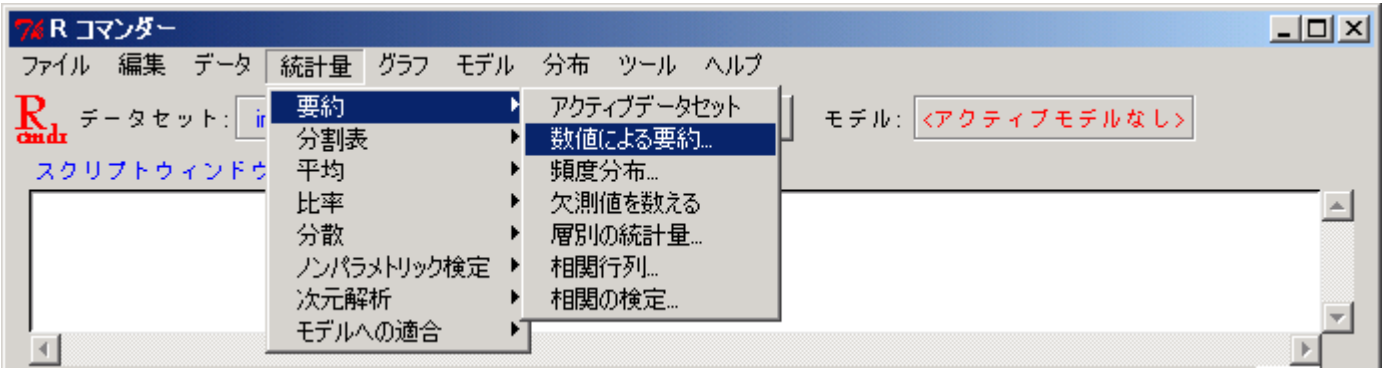
USJudgeRatings

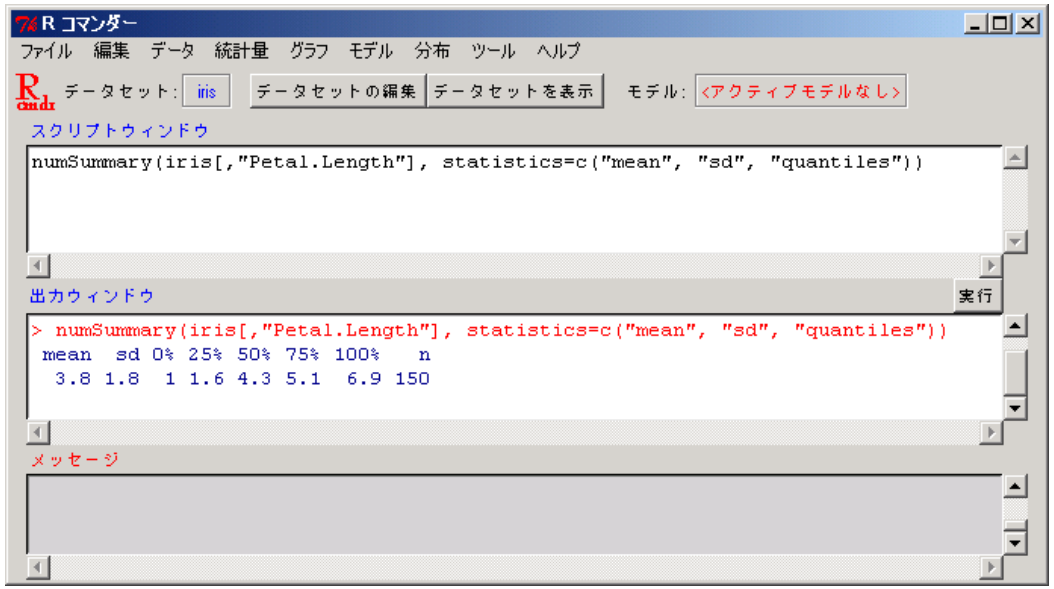
	CONT	INTG	DMNR	DILG	CFMG
HULL,T.C.	7.7	7.7	6.7	7.5	7.4
LEVINE,I.	8.3	8.2	7.4	7.8	7.7
LEVISTER,R.L.	9.6	6.9	5.7	6.6	6.9
MARTIN,L.F.	7.1	8.2	7.7	7.1	6.6
MCGRATH,J.F.	7.6	7.3	6.9	6.8	6.7
MIGNONE,A.F.	6.6	7.4	6.2	6.2	5.4
MISSAL,H.M.	6.2	8.3	8.1	7.7	7.4
MULVEY,H.M.	7.5	8.7	8.5	8.6	8.5
NARUK,H.J.	7.8	8.9	8.7	8.9	8.7
O'BRIEN,F.J.	7.1	8.5	8.3	8.0	7.9
O'SULLIVAN,T.J.	7.5	9.0	8.9	8.7	8.4
PASKEY,L.	7.5	8.1	7.7	8.2	8.0
RUBINOW,J.E.	7.1	9.2	9.0	9.0	8.4
SADEN.G.A.	6.6	7.4	6.9	8.4	8.0
SATANIELLO,A.G.	8.4	8.0	7.9	7.9	7.8
SHEA,D.M.	6.9	8.5	7.8	8.5	8.1
SHEA,J.F.JR.	7.3	8.9	8.8	8.7	8.4
SIDOR,W.J.	7.7	6.2	5.1	5.6	5.6
SPEZIALE,J.A.	8.5	8.3	8.1	8.3	8.4
SPONZO,M.J.	6.9	8.3	8.0	8.1	7.9
STAPLETON,J.F.	6.5	8.2	7.7	7.8	7.6
TESTO,R.J.	8.3	7.3	7.0	6.8	7.0
TIERNEY,W.L.JR.	8.3	8.2	7.8	8.3	8.4
WALL,R.A.	9.0	7.0	5.9	7.0	7.0
WRIGHT,D.B.	7.1	8.4	8.4	7.7	7.5
ZARRILLI,K.J.	8.6	7.4	7.0	7.5	7.5

R Commander とは？

- R Commander（アールコマンダー）は John Fox 教授（カナダ・McMaster 大学）が開発した GUI 版 R のこと
- マウス操作で R を使うことが出来る！（R の命令を覚えなくても R の出力が得られる！）
- 2005 年頃より、関西大学の荒木 孝治先生が主体となって R Commander のメッセージ翻訳がなされ…
- R Commander はバージョン 1.1-1 より本格的に日本語化された！

R Commander の概要





① メニューから機能を選択
② スクリプトウィンドウには実行したＲのコマンドが出力される
③ 出力ウィンドウには，実行結果が出力される
④ メッセージにはエラーや警告が出力される

メニュー

- R Commander の概要
- データ「iris」の読み込み方法 ←
- グラフの作成方法
- R Commander でのデータ解析例
- グラフのカタログ
- 各メニューの概要

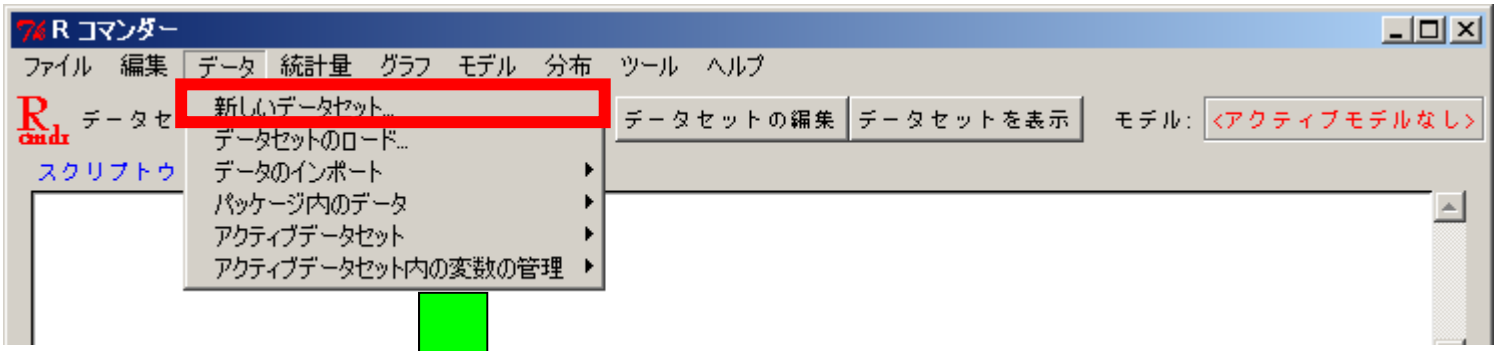
Graphic by (c)Tomo.Yun (http://www.yunphoto.net)

使用するデータ「iris」

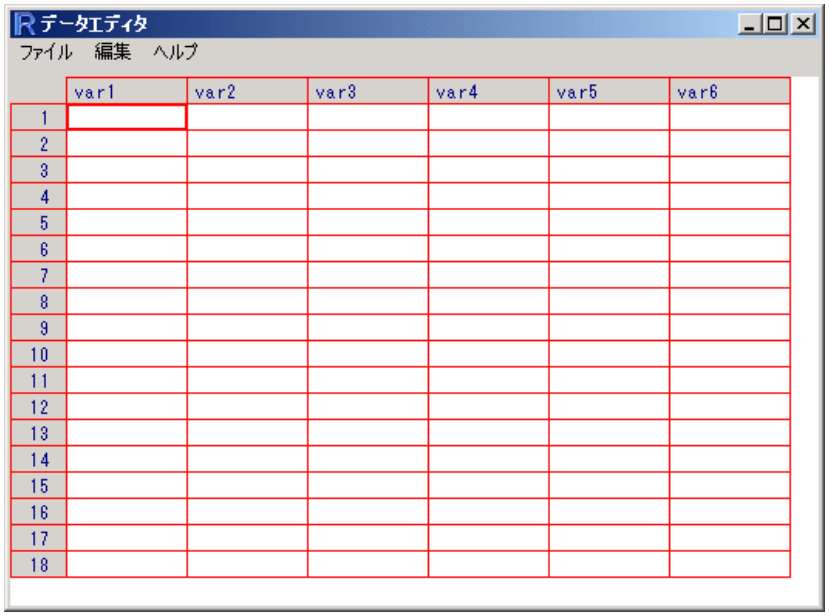
Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
・・・	・・・	・・・	・・・	・・・

- フィッシャーが判別分析法を紹介するために利用したアヤメの品種分類（Species：setosa，versicolor，virginica）に関するデータ
  ⇒ 以下の４変数を説明変数としてアヤメの種類を判別しようとした
  - アヤメのがくの長さ（Sepal.Length）
  - アヤメのがくの幅　（Sepal.Width）
  - アヤメの花弁の長さ（Petal.Length）
  - アヤメの花弁の幅　（Petal.Width）
- まずは R Commander にデータを読み込ませる方法をいくつか紹介

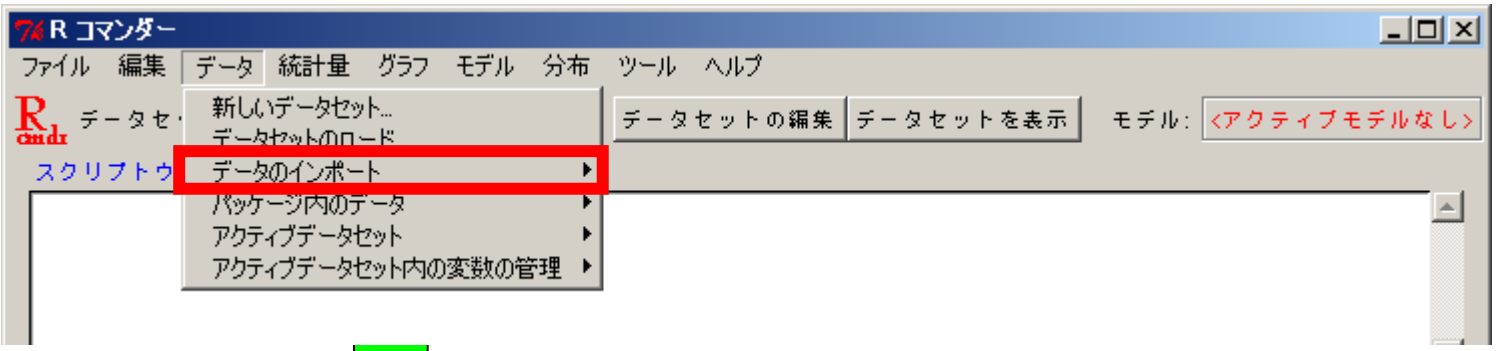
データ「iris」を手入力する場合



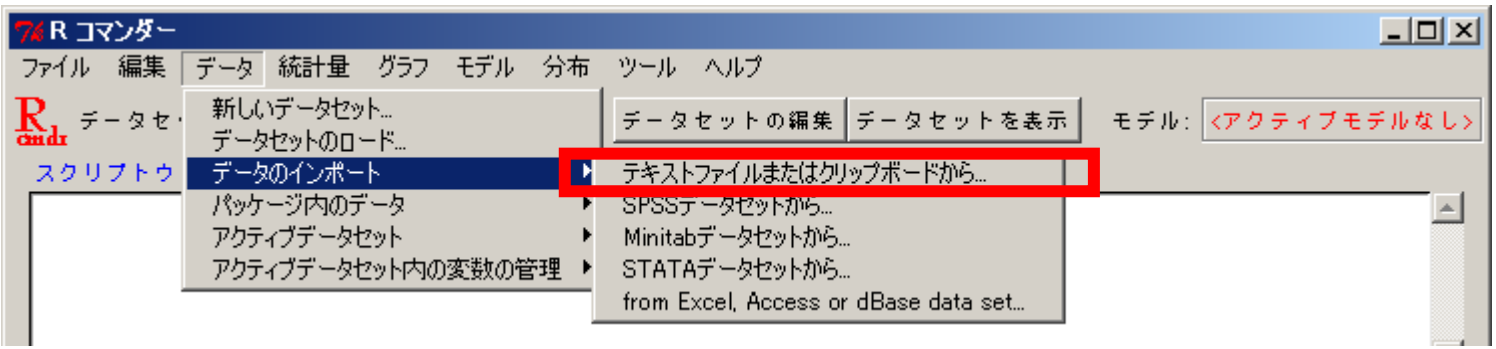
スプレッドシートに
データを入力する



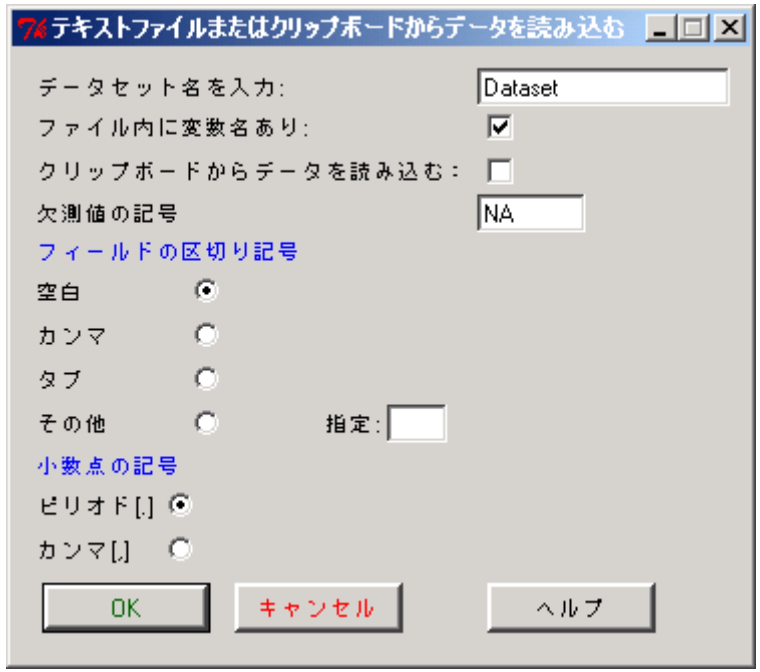
データ「iris」がテキストファイルである場合



「テキストファイルまたはクリップボードから...」を選択



データ「iris」がテキストファイルである場合

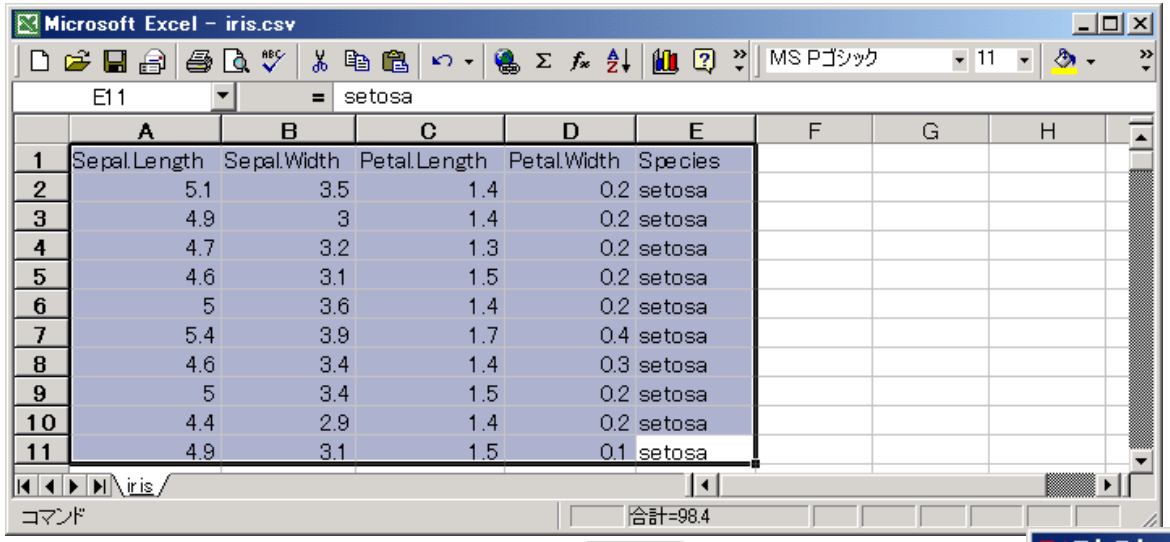


データの
読み込み

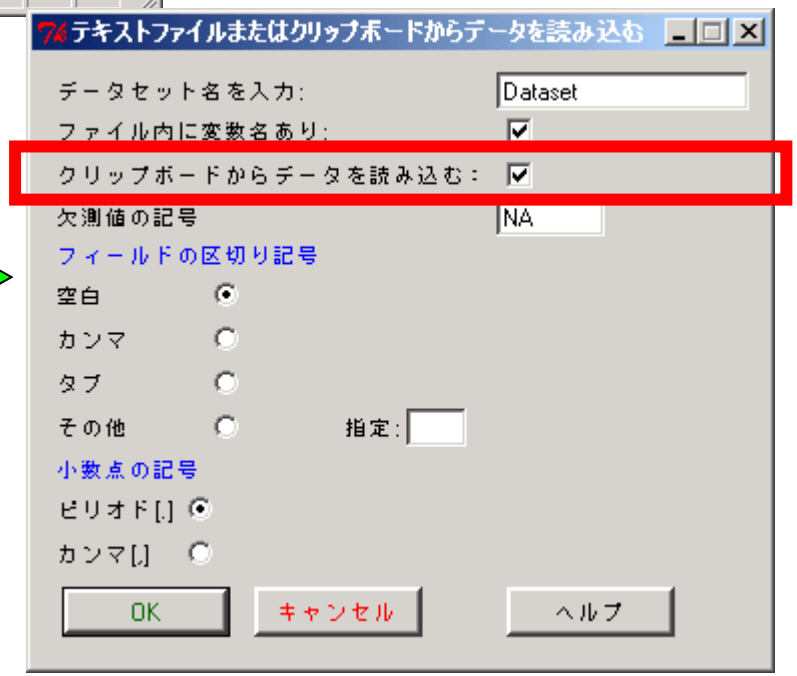
- データセット名を入力する
- ファイル内に変数名（列名）がある場合はチェック
- 欠測値の記号を指定する（通常は"NA"）
- フィールドの区切り記号（空白，カンマ，タブ，etc）を指定する
- 小数点の記号（ピリオド or カンマ）を指定する

⇒ 読み込むことが出来るデータセットの種類は豊富！

データ「iris」が Excel ファイルである場合①

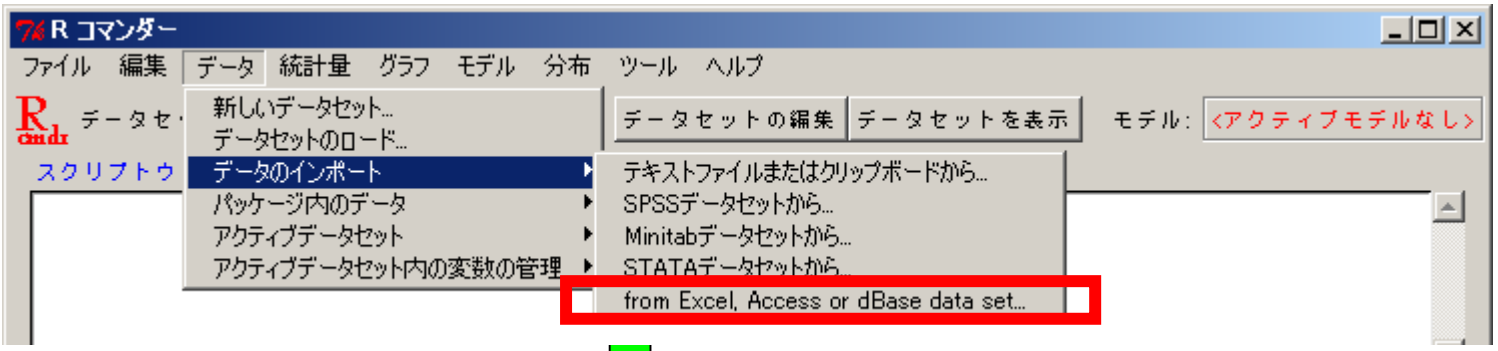


	A	B	C	D	E
1	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
2	5.1	3.5	1.4	0.2	setosa
3	4.9	3	1.4	0.2	setosa
4	4.7	3.2	1.3	0.2	setosa
5	4.6	3.1	1.5	0.2	setosa
6	5	3.6	1.4	0.2	setosa
7	5.4	3.9	1.7	0.4	setosa
8	4.6	3.4	1.4	0.3	setosa
9	5	3.4	1.5	0.2	setosa
10	4.4	2.9	1.4	0.2	setosa
11	4.9	3.1	1.5	0.1	setosa



EXCELファイルのデータをコピーした後，
そのデータを R Commander に読み込ませる
ことも可（「クリップボード…」）にチェック！

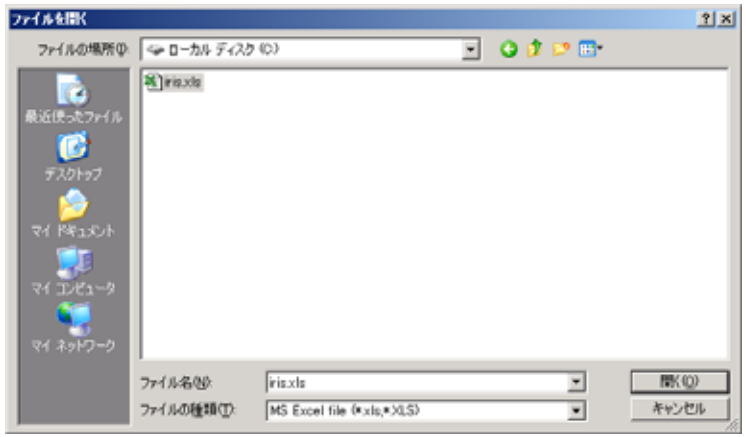
データ「iris」が Excel ファイルである場合②



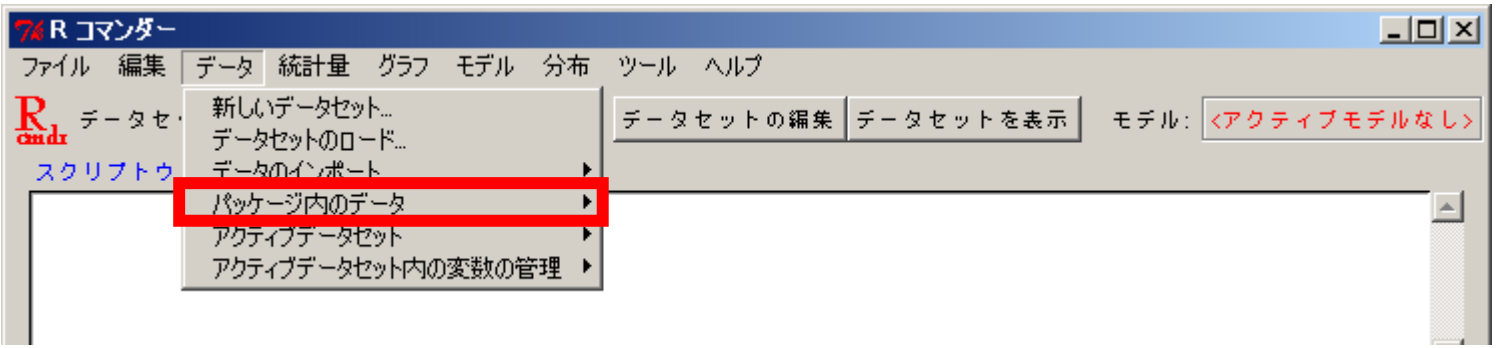
EXCELやAccessファイルの
データを読み込むときは，

① 上記画像のメニューを
選択してデータセット名を入力

② 読み込むファイル
を選択する

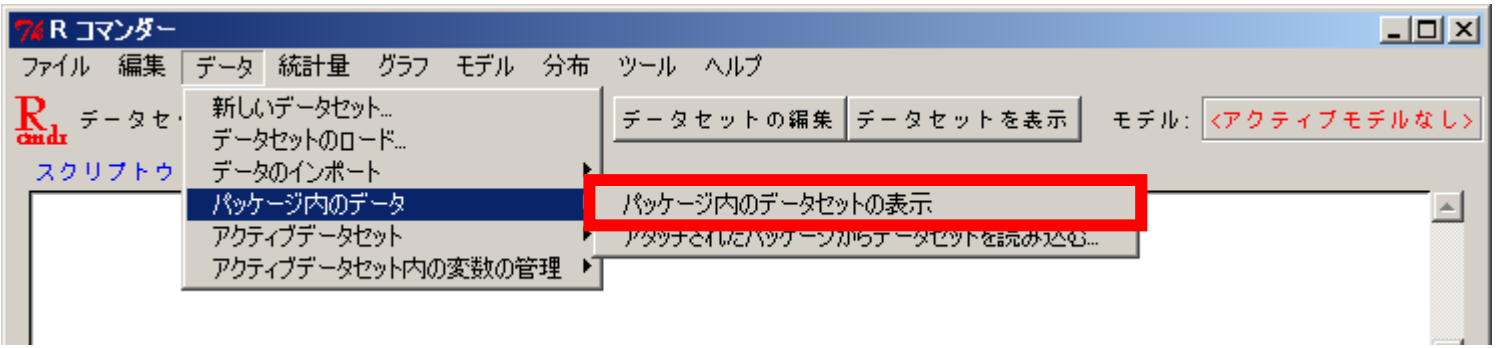


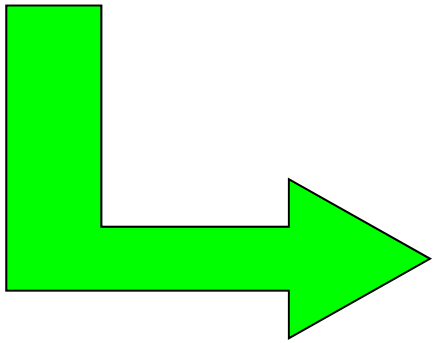
サンプルデータの読み込み

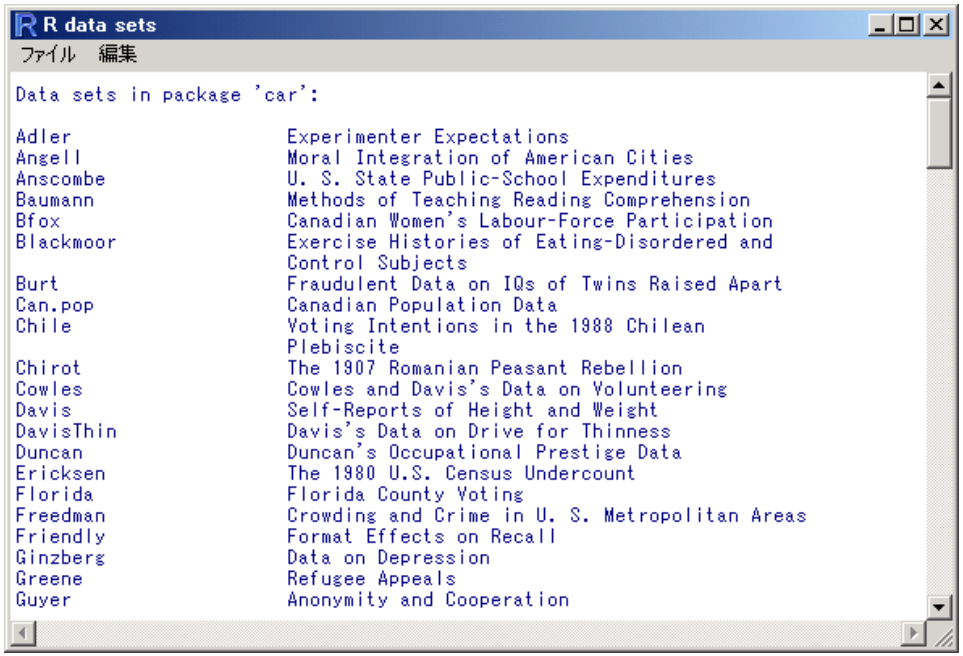


- パッケージ内のデータ
  - R には，サンプルデータセットが多数収録されている！
  - このメニューでサンプルデータの一覧を表示したり，サンプルデータを読み込むことが出来る
  - 本日はサンプルデータ iris を使用する ←

サンプルデータの読み込み

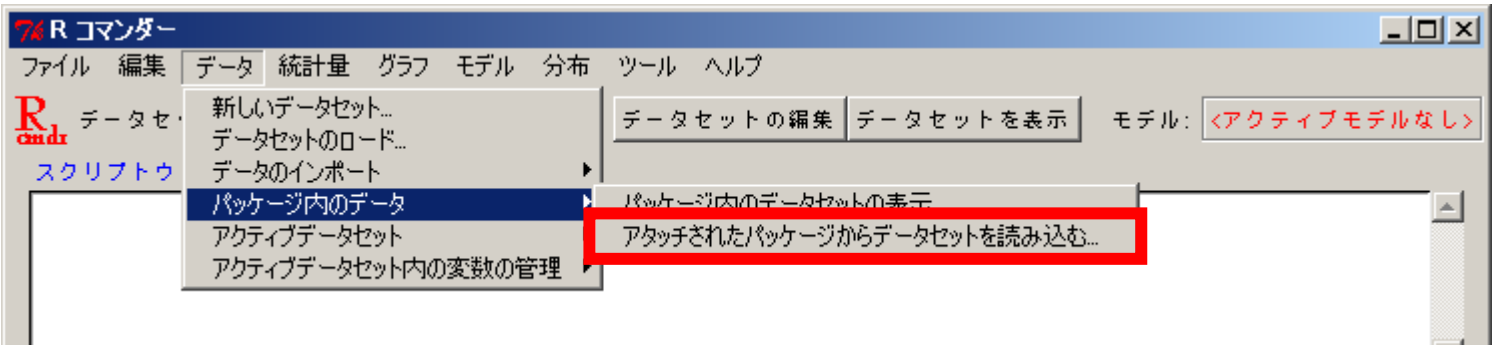






使うことが出来るサンプルデータの一覧（データの簡単な説明つき）が表示される

サンプルデータの読み込み



1. パッケージを選択
   （普通はdatasetsを選択）
2. 使用するデータセットの名前
   （ここでは **iris** ）をダブルクリック
3. データセット名を変更する場合は
   データセット名を修正

読み込んだデータの表示

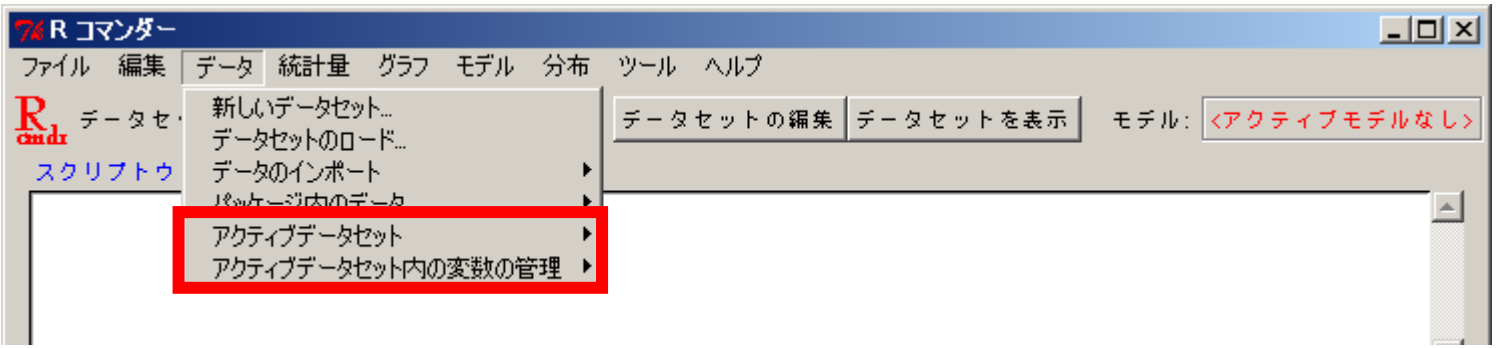
iris

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa
13	4.8	3.0	1.4	0.1	setosa
14	4.3	3.0	1.1	0.1	setosa
15	5.8	4.0	1.2	0.2	setosa
16	5.7	4.4	1.5	0.4	setosa
17	5.4	3.9	1.3	0.4	setosa
18	5.1	3.5	1.4	0.3	setosa
19	5.7	3.8	1.7	0.3	setosa
20	5.1	3.8	1.5	0.3	setosa
21	5.4	3.4	1.7	0.2	setosa
22	5.1	3.7	1.5	0.4	setosa
23	4.6	3.6	1.0	0.2	setosa
24	5.1	3.3	1.7	0.5	setosa
25	4.8	3.4	1.9	0.2	setosa
26	5.0	3.0	1.6	0.2	setosa
27	5.0	3.4	1.6	0.4	setosa
28	5.2	3.5	1.5	0.2	setosa
29	5.2	3.4	1.4	0.2	setosa
30	4.7	3.2	1.6	0.2	setosa

読み込んだデータセットを表示するときは，「データセットを表示」をクリック！

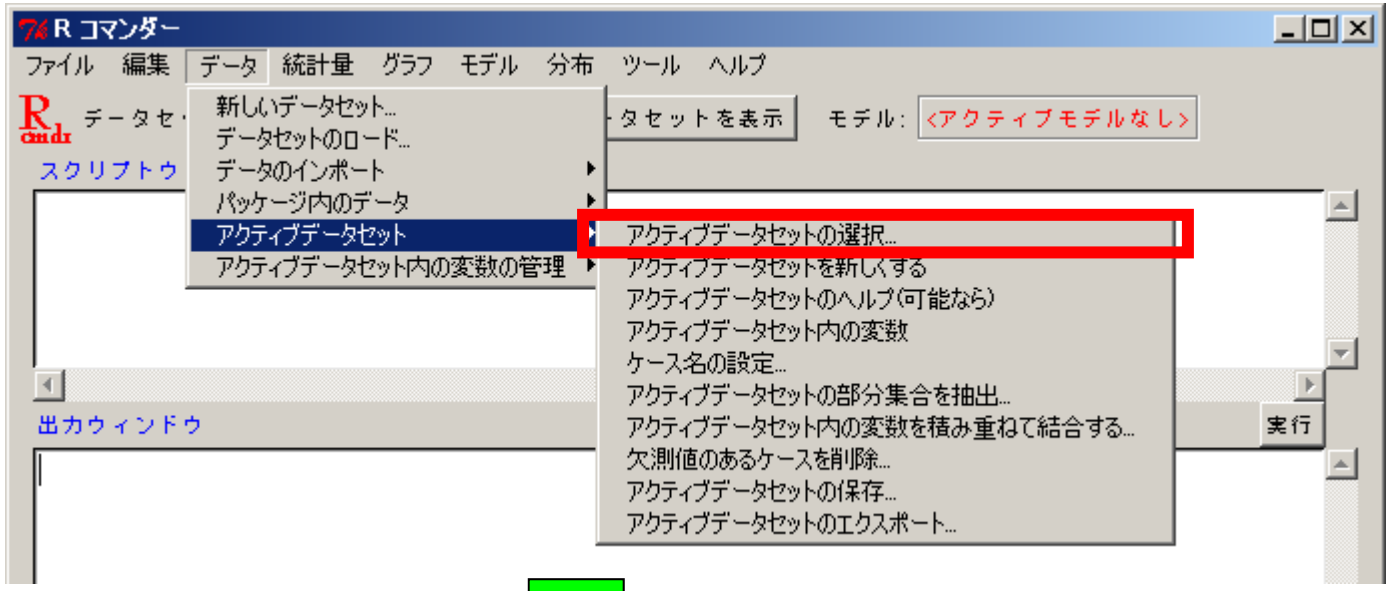
※ 読み込んだデータセットを表示するときは，「データセットの編集」をクリック

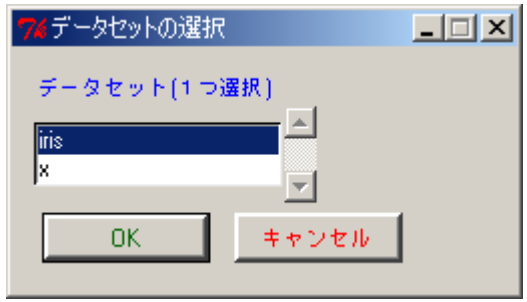
読み込んだデータの加工



- アクティブデータセット
- アクティブデータセット内の変数の管理
  - 解析に使うデータセットを選択する
  - 読み込んだデータセットに関する情報を見る
  - データの部分集合を切り取ったり，欠測値を除く処理を行う
  - データセットをファイルに出力する
  - 数値データをカテゴリデータに変換する
  - 変数名を変更したり，変数を削除する　　etc...

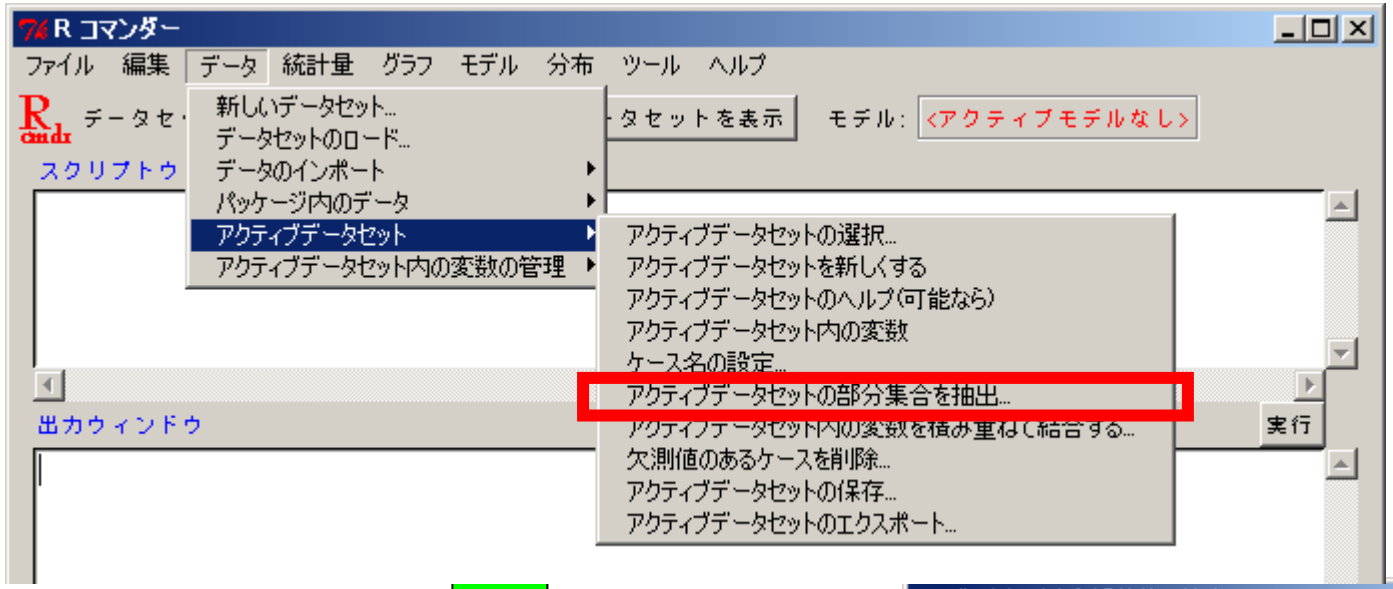
読み込んだデータの加工例





データを加工するときは
まず使用するデータセットを選択
（ここでは iris をダブルクリック）

読み込んだデータの加工例



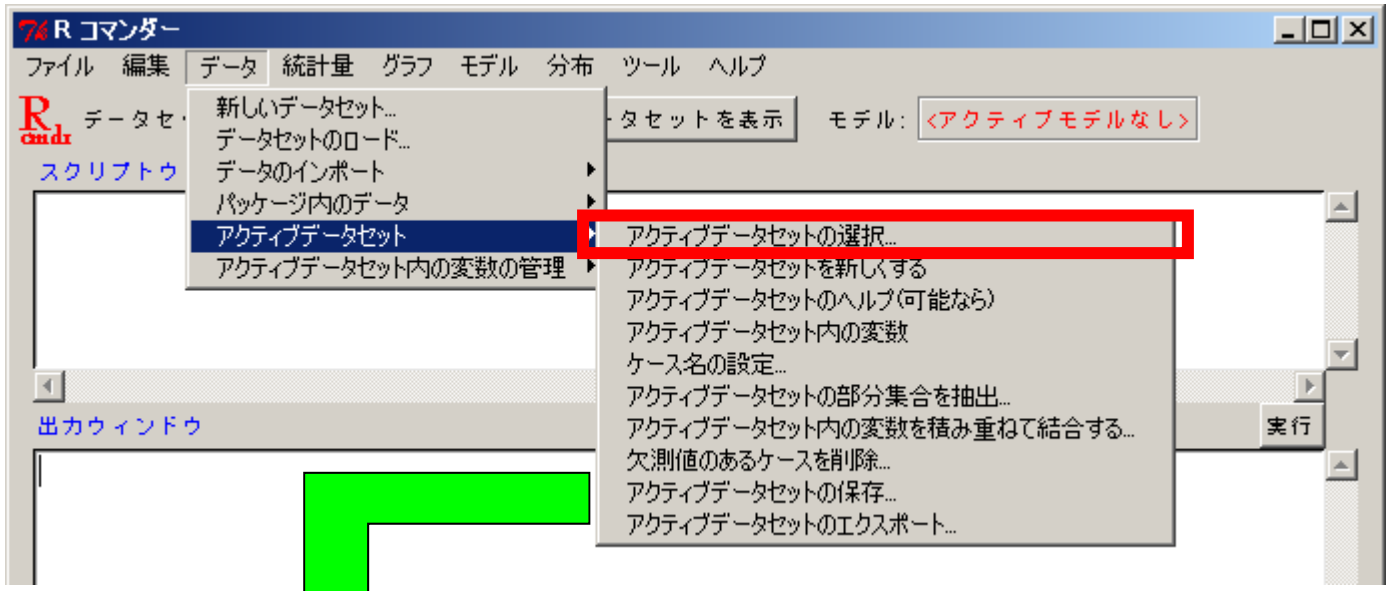
★ データセットから部分集合を切り出す

① 切り出す変数を選択する

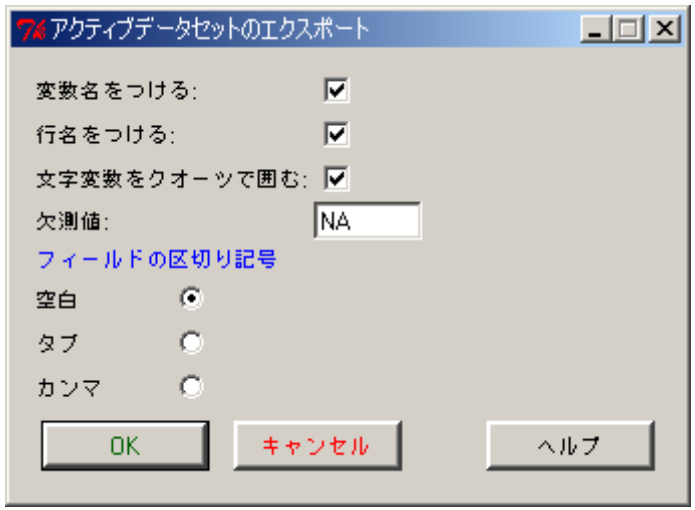
② 「部分集合の表現」に条件式を入力する

③ 切り出した後のデータセット名を入力する

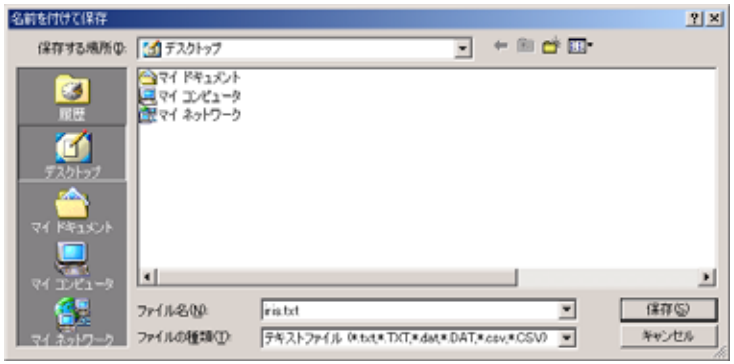
データの出力



諸設定をした後...



データセットをファイルに出力する
（テキストファイルのみ）

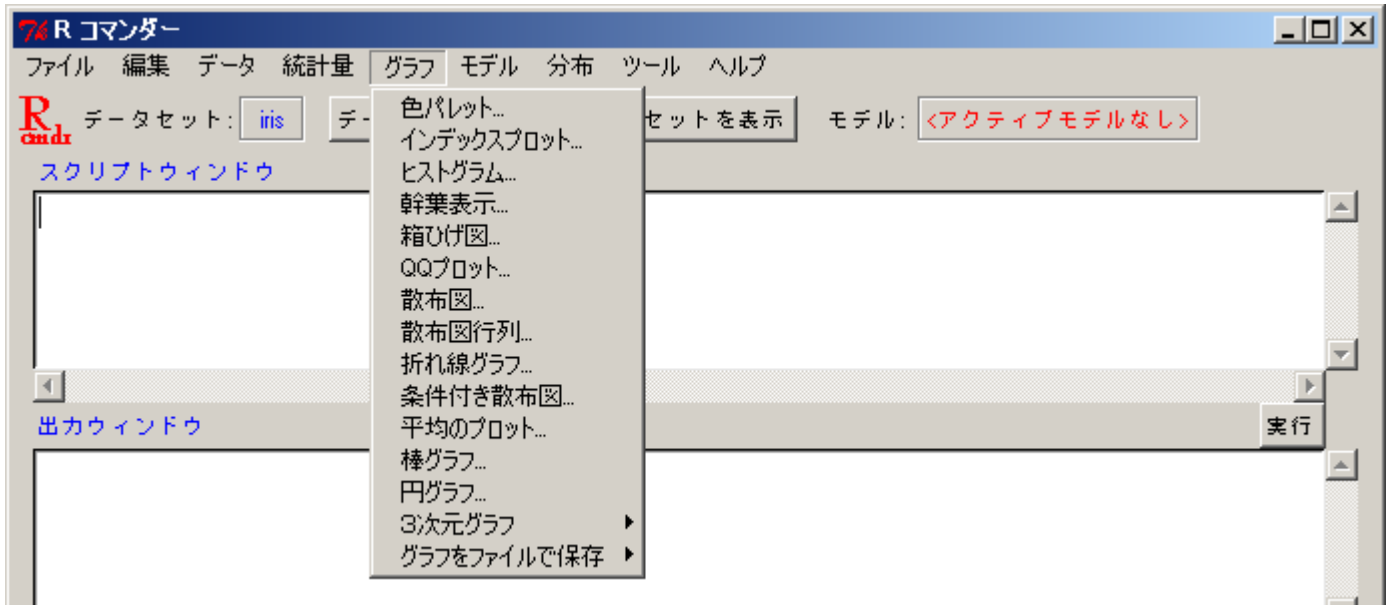


メニュー

- R Commander の概要
- データ「iris」の読み込み方法
- グラフの作成方法 ←
- R Commander でのデータ解析例
- グラフのカタログ
- 各メニューの概要

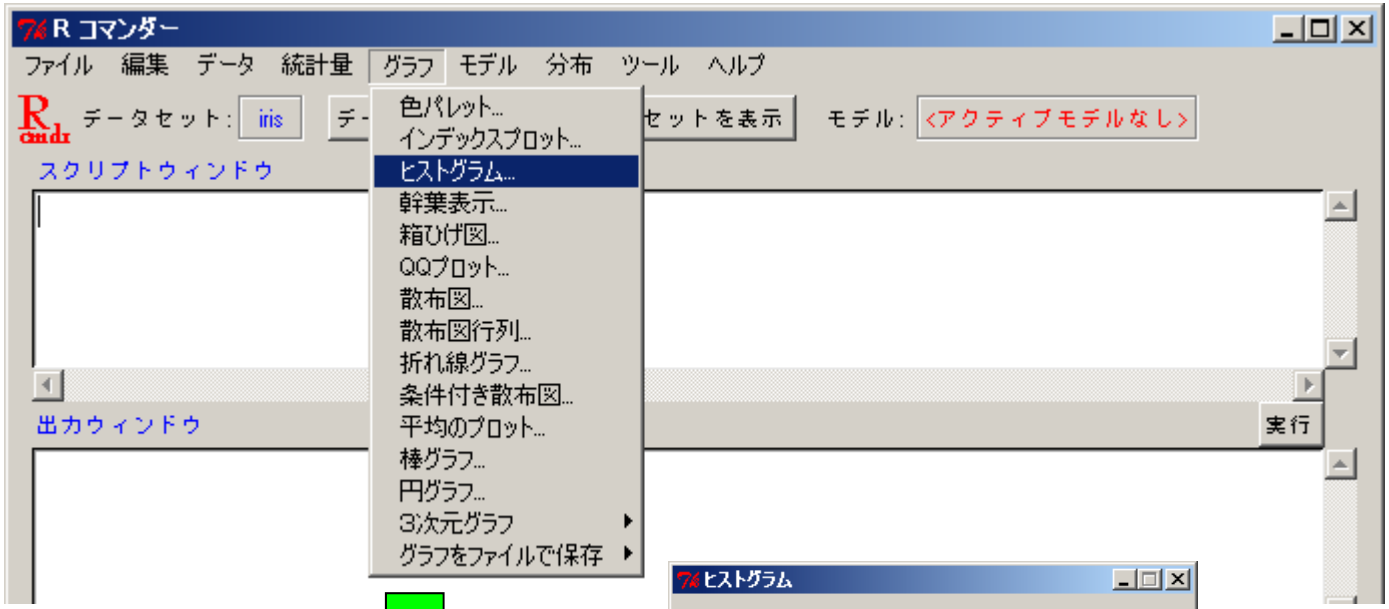
Graphic by (c)Tomo.Yun (http://www.yunphoto.net)

グラフに関するメニュー



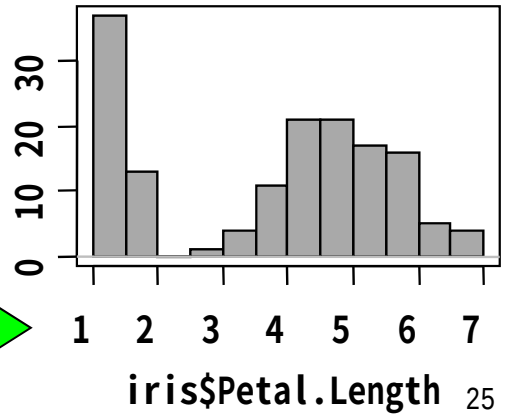
- グラフ → 様々な種類のグラフを描くことが出来る
  ⇒ インデックスプロット，ヒストグラム，幹葉表示（幹葉図），箱ひげ図，QQプロット，散布図，散布図行列，折れ線グラフ，条件付き散布図，平均のプロット，棒グラフ，円グラフ，3Dグラフ
- グラフをファイルに保存することも出来る
- 作成することができるグラフの一覧は後ほど紹介・・・

グラフの作成手順

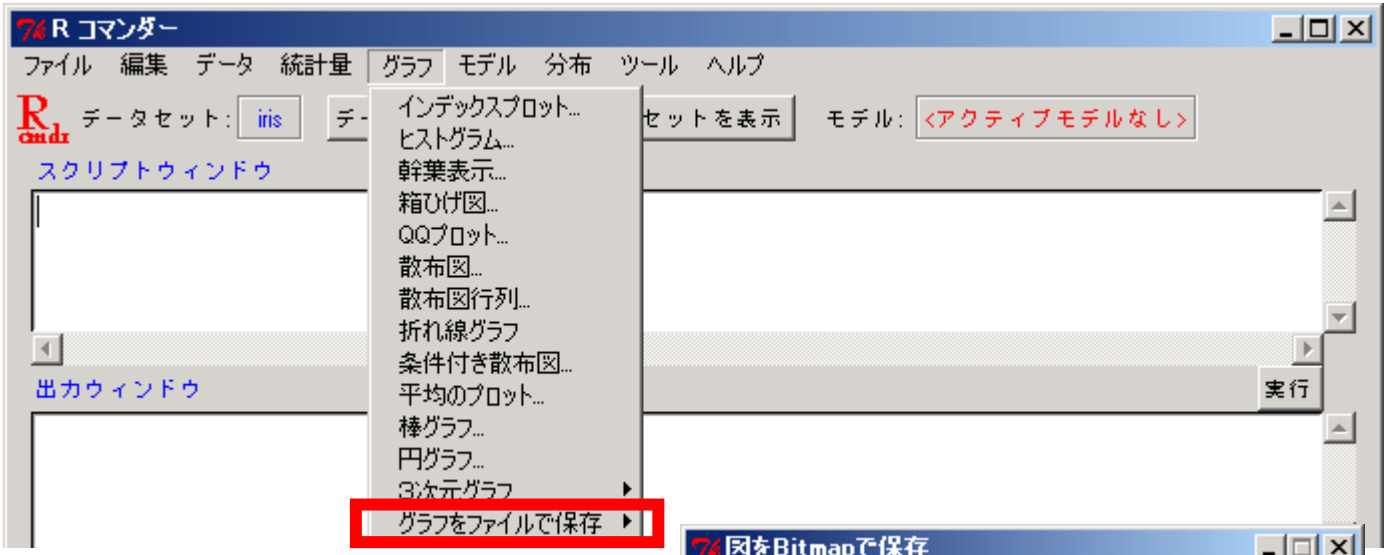


① 描きたいグラフ
(例:ヒストグラム)
のメニューを選択

②データの変数や
オプションを指定

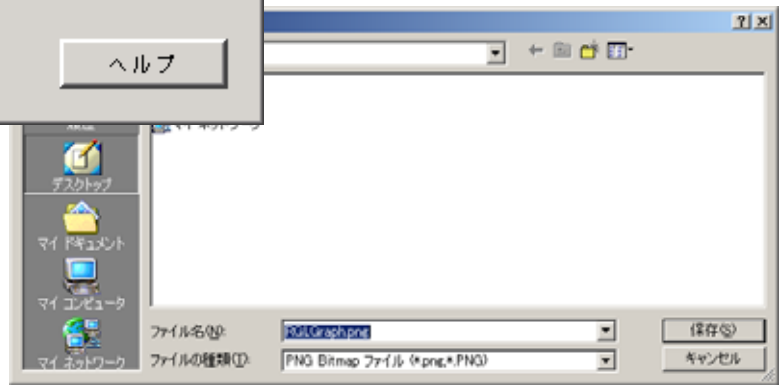


グラフの保存



描いたグラフを保存
することが出来る
(PNG，PDF，PS，EPS…)

グラフの形式やサイズを
指定して[OK]をクリック



メニュー

- R Commander の概要
- データ「iris」の読み込み方法
- グラフの作成方法
- **R Commander でのデータ解析例 ←**
- グラフのカタログ
- 各メニューの概要

Graphic by (c)Tomo.Yun (http://www.yunphoto.net)

データの表示

iris

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa
13	4.8	3.0	1.4	0.1	setosa
14	4.3	3.0	1.1	0.1	setosa
15	5.8	4.0	1.2	0.2	setosa
16	5.7	4.4	1.5	0.4	setosa
17	5.4	3.9	1.3	0.4	setosa
18	5.1	3.5	1.4	0.3	setosa
19	5.7	3.8	1.7	0.3	setosa
20	5.1	3.8	1.5	0.3	setosa
21	5.4	3.4	1.7	0.2	setosa
22	5.1	3.7	1.5	0.4	setosa
23	4.6	3.6	1.0	0.2	setosa
24	5.1	3.3	1.7	0.5	setosa
25	4.8	3.4	1.9	0.2	setosa
26	5.0	3.0	1.6	0.2	setosa
27	5.0	3.4	1.6	0.4	setosa
28	5.2	3.5	1.5	0.2	setosa
29	5.2	3.4	1.4	0.2	setosa
30	4.7	3.2	1.6	0.2	setosa

★ データを眺めても
よく分からない…
⇒ データを要約する！

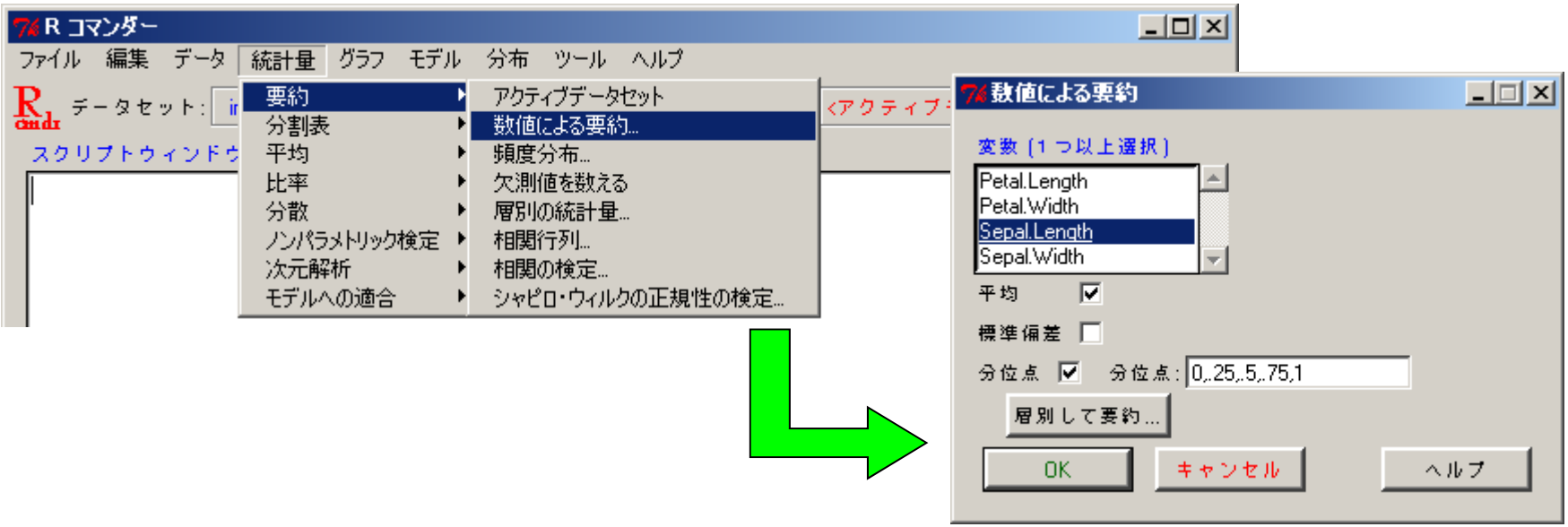
1 変数の要約

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
・・・	・・・	・・・	・・・	・・・

- アヤメのがくの長さ（Sepal.Length）の特徴をつかむには・・・

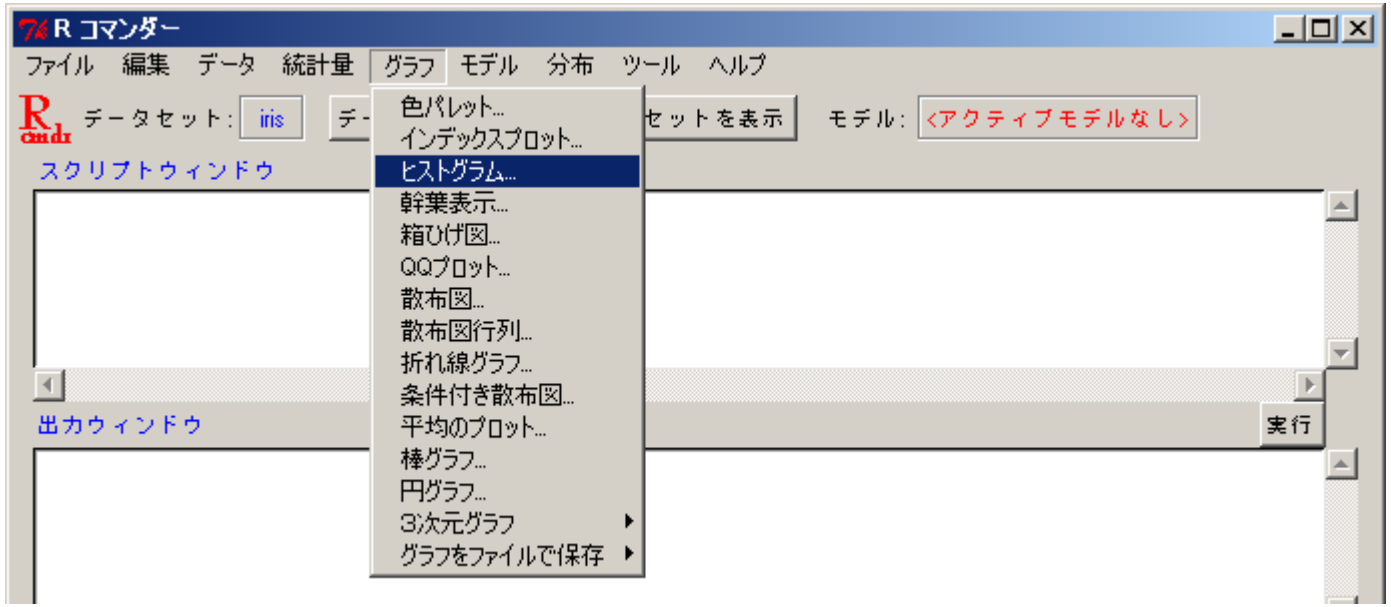
1. 数値による要約 ⇒ 要約統計量を求める
2. グラフによる要約 ⇒ ヒストグラムを作成する
3. 層別して要約統計量やヒストグラム

要約統計量の算出



mean	sd	0%	25%	50%	75%	100%	n
3.758	1.765298	1	1.6	4.35	5.1	6.9	150

ヒストグラムの作成



★ ヒストグラムを作成

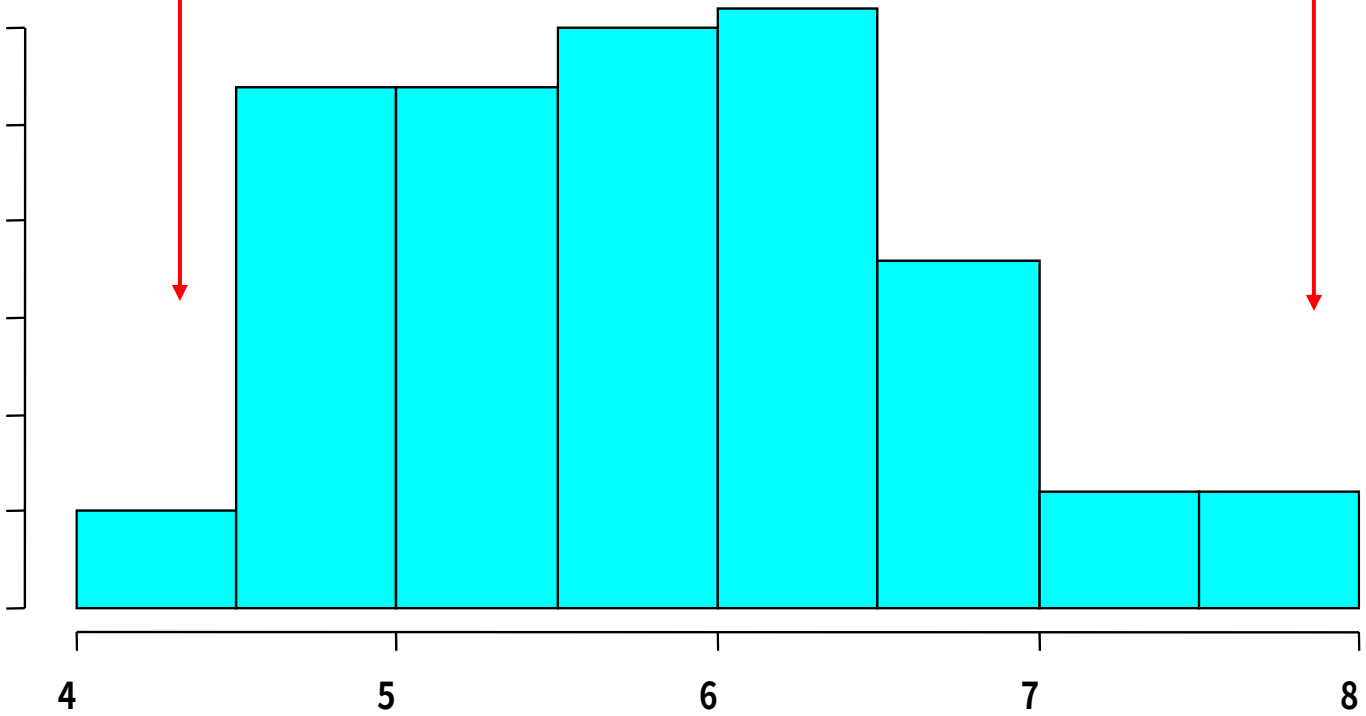
ヒストグラムの作成

0%	25%	50%	75%	100%
4.3	5.1	5.8	6.4	7.9

一番小さい値
（最小値）

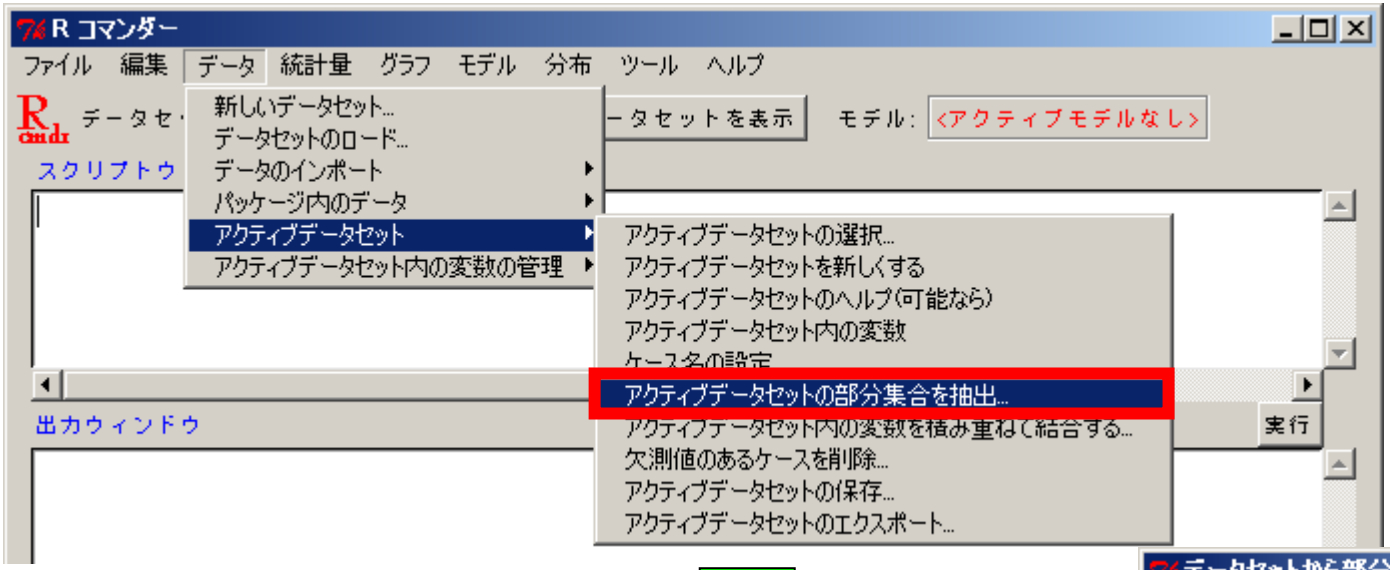
真ん中の値
（中央値）

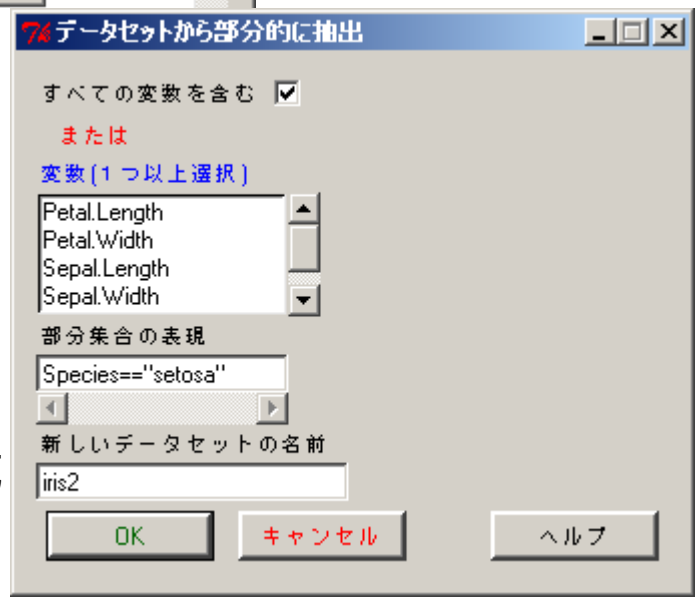
一番大きい値
（最大値）



一目瞭然！

層別にヒストグラムを作成





☆ アヤメの種類別にヒストグラムを描く

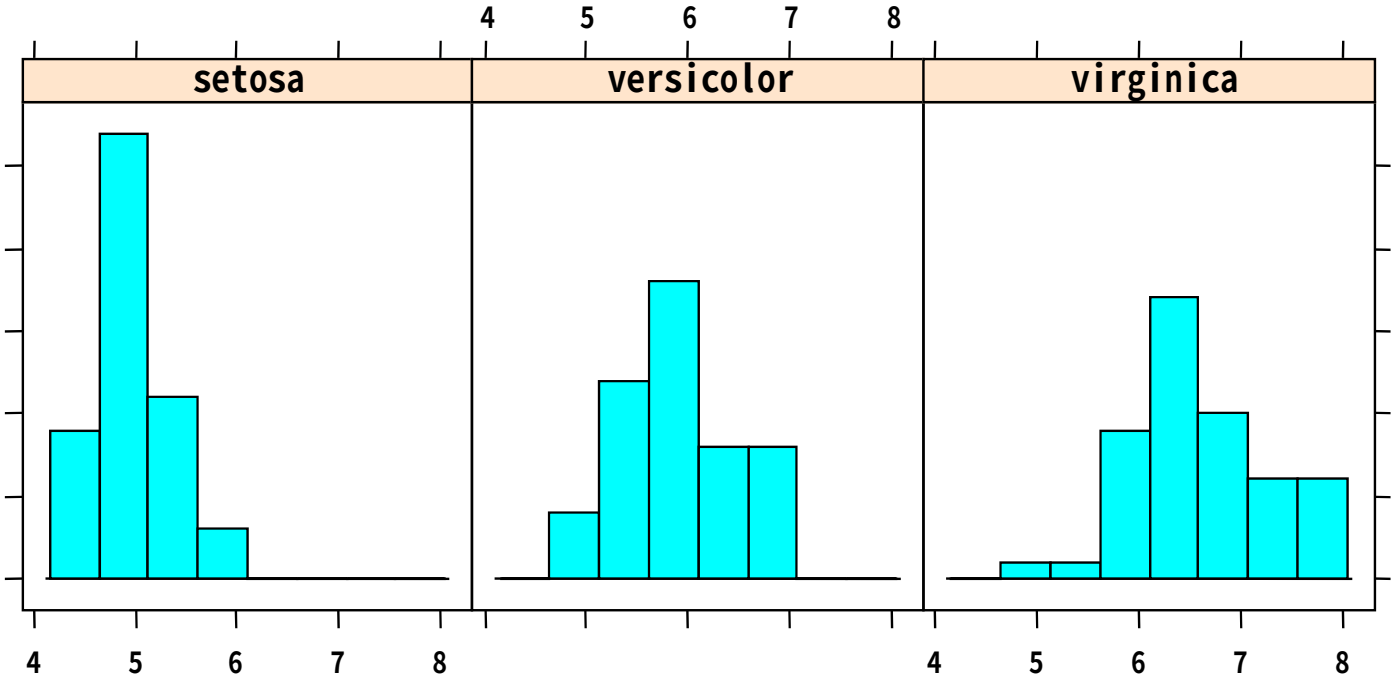
① データ **iris** からアヤメの種類（Species）が setosa となっているものだけを抽出

② 先ほど紹介した方法でヒストグラムを描く

③ versicolor，verginica についてもヒストグラム作成

層別にヒストグラムを作成

★ 層別にヒストグラムを作成



- Setosa　：がくが短い
- Verginica：がくが長い

⇒ 層別すると特徴が浮き出る！

```
> library(lattice)      # ちなみに R の命令はこんな感じ
> histogram(~ Sepal.Length | Species, data=iris)
```

一目瞭然！

2 変数の関係

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
・・・	・・・	・・・	・・・	・・・

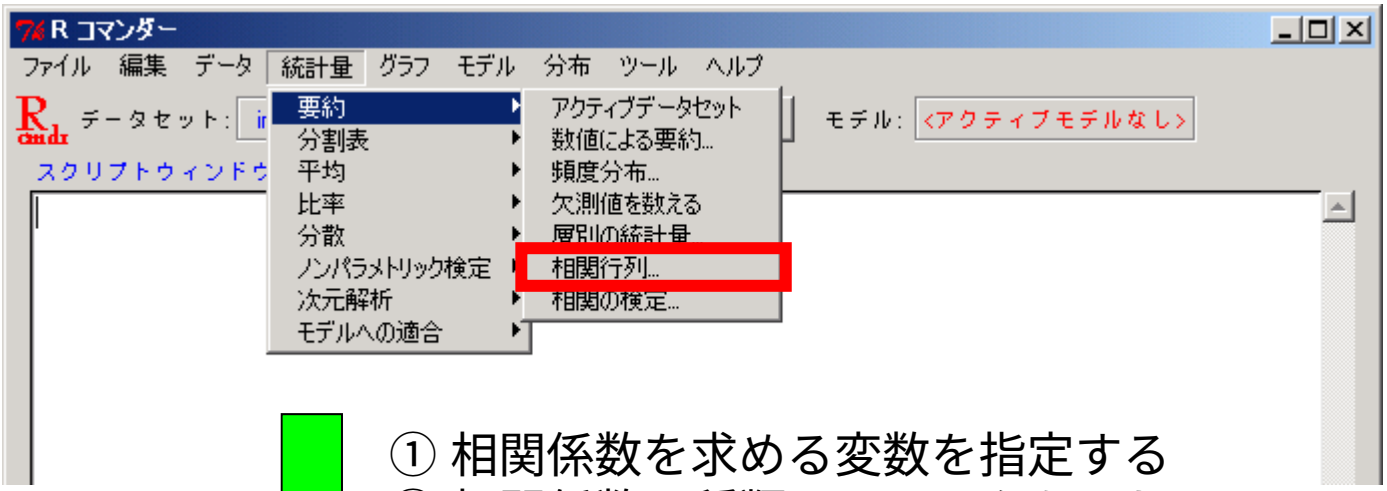
- アヤメの花弁の幅（Petal.Width）と花弁の長さ（Petal.Length）の関係を見る場合は・・・

1. 数値による要約 ⇒ 相関係数を求める

2. グラフによる要約 ⇒ 散布図を描く

3. 層別してグラフ（散布図）を描く

相関係数の算出

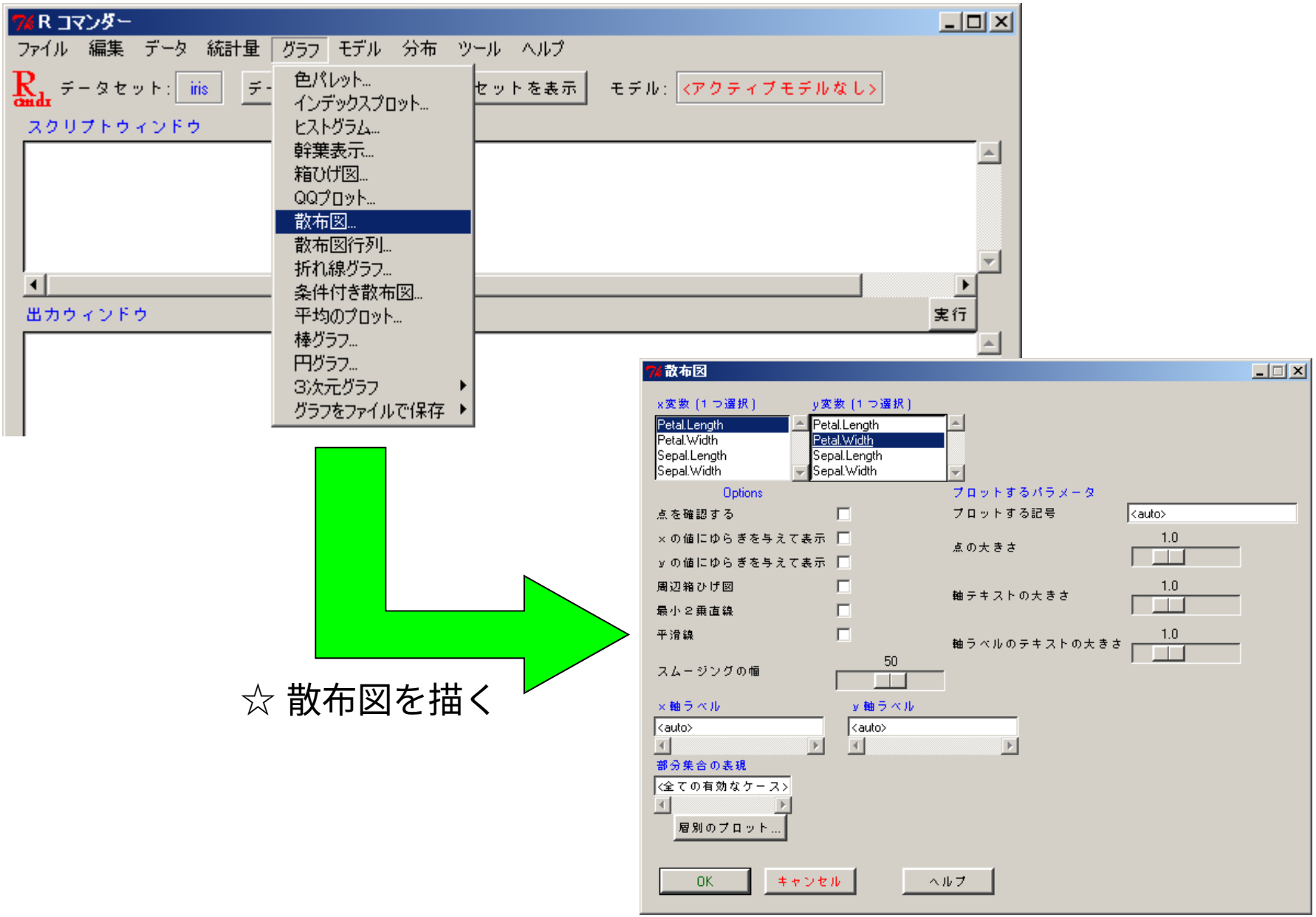


① 相関係数を求める変数を指定する
② 相関係数の種類にチェックを入れる

相関係数が求まる

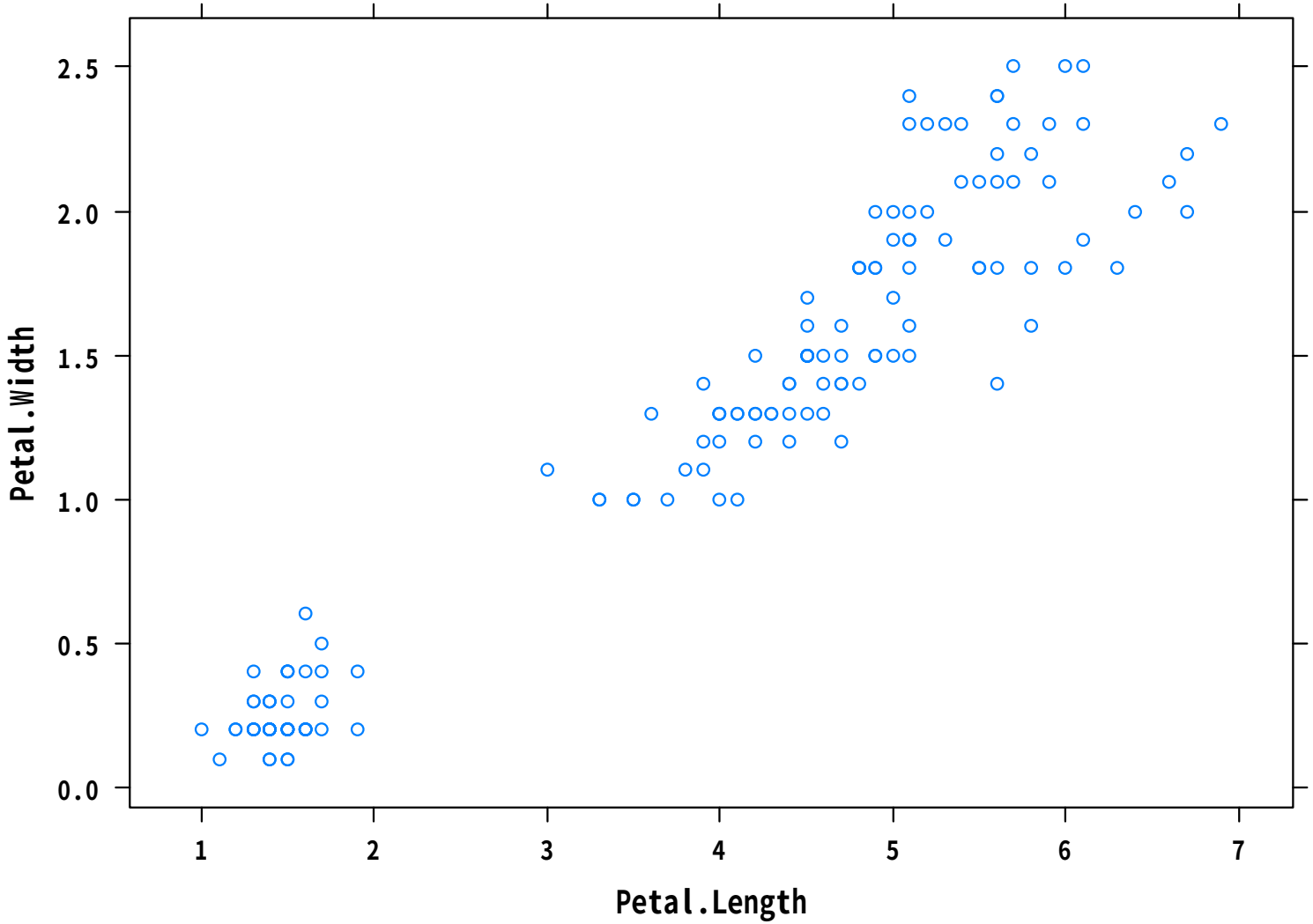
	Petal.Length	Petal.Width
Petal.Length	1.00	0.96
Petal.Width	0.96	1.00

散布図の作成



☆ 散布図を描く

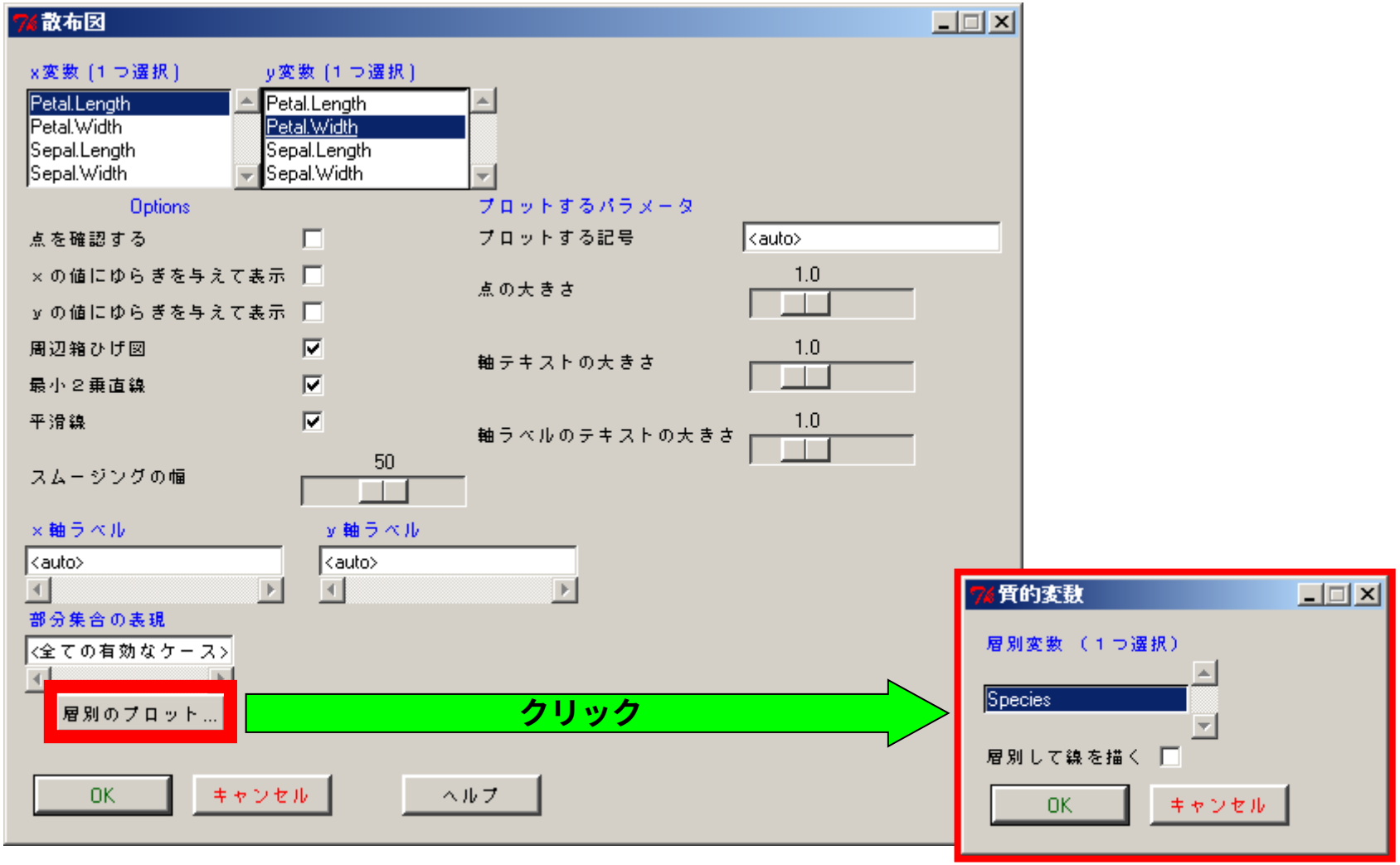
散布図の作成



一目瞭然！

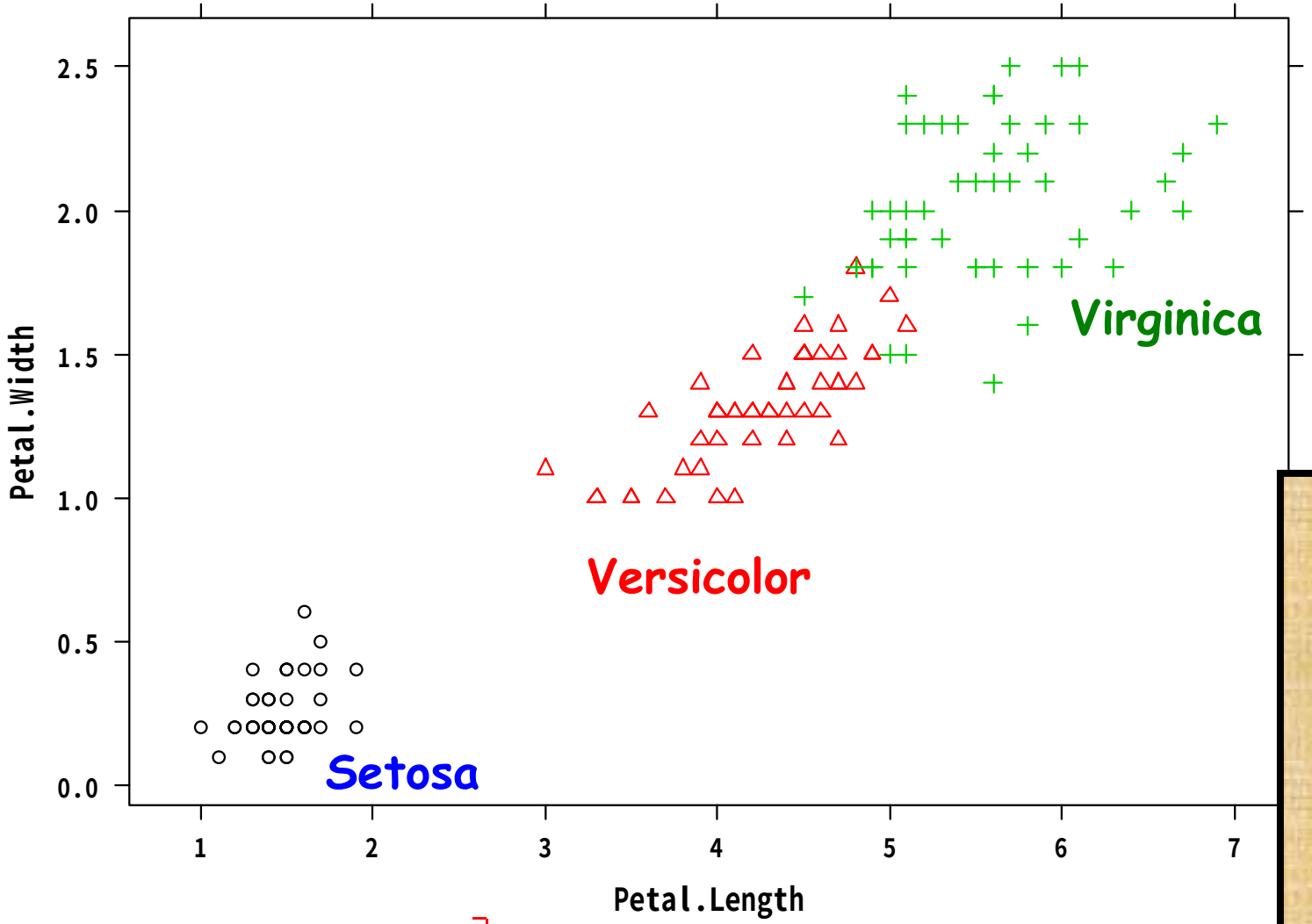
- Petal.Width と Petal.Length の関係は右肩上がり ⇒ ひと目で分かる！

層別散布図の作成



☆ 層別変数にアヤメの種類 Species を指定する

層別散布図の作成



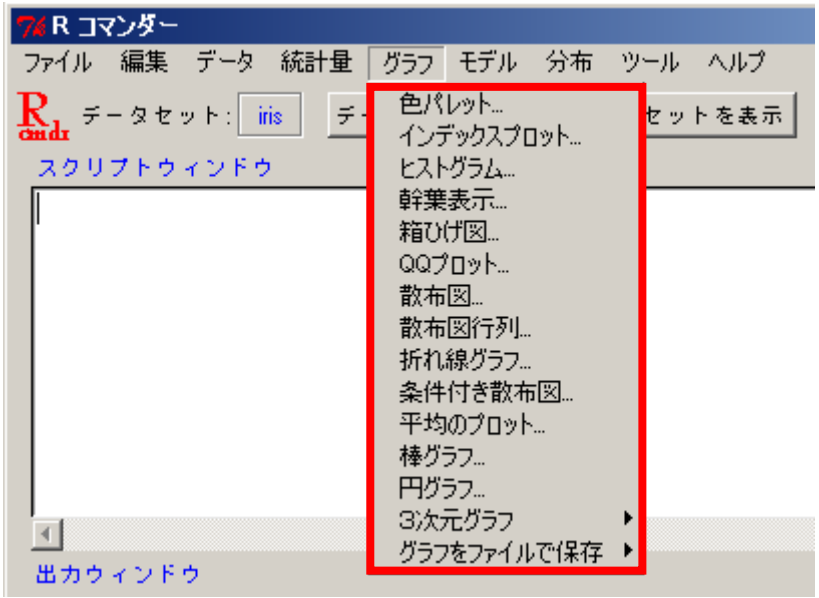
一目瞭然！

- Setosa　：左下に分布
- Verginica：右上に分布

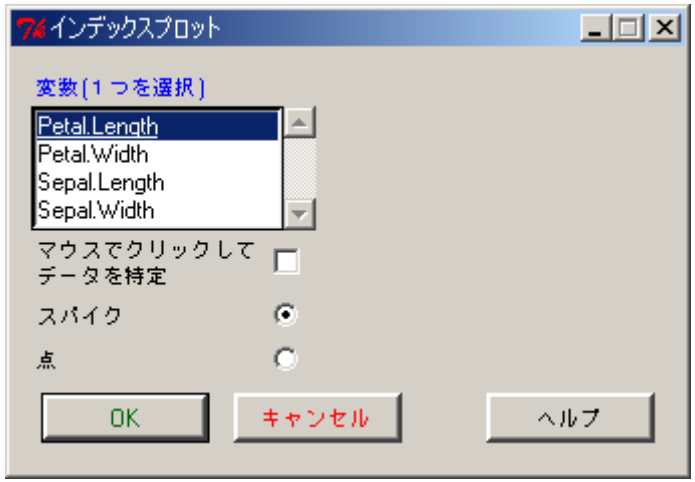
⇒ 層別すると特徴が浮き出る！

メニュー

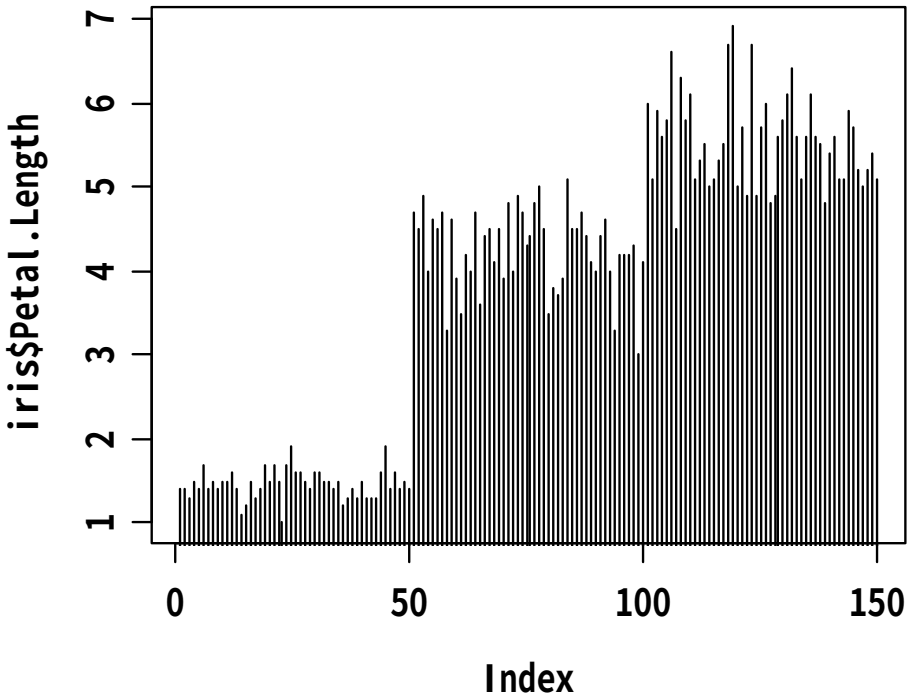
- R Commander の概要
- データ「iris」の読み込み方法
- グラフの作成方法
- R Commander でのデータ解析例
- グラフのカタログ ←
- 各メニューの概要



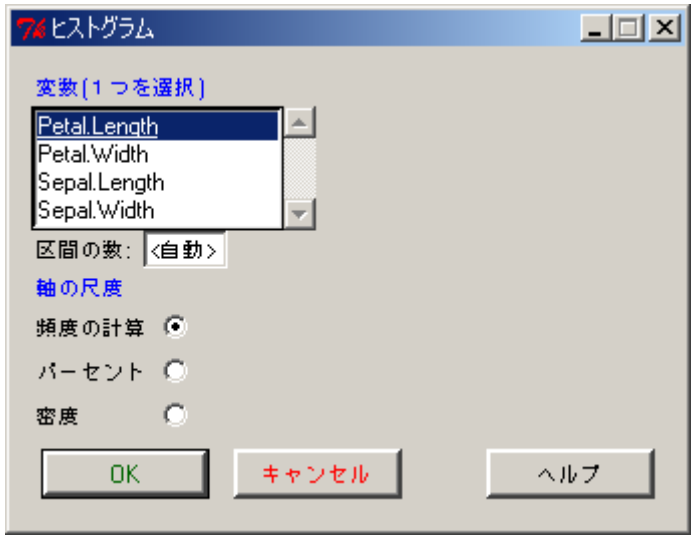
インデックスプロット



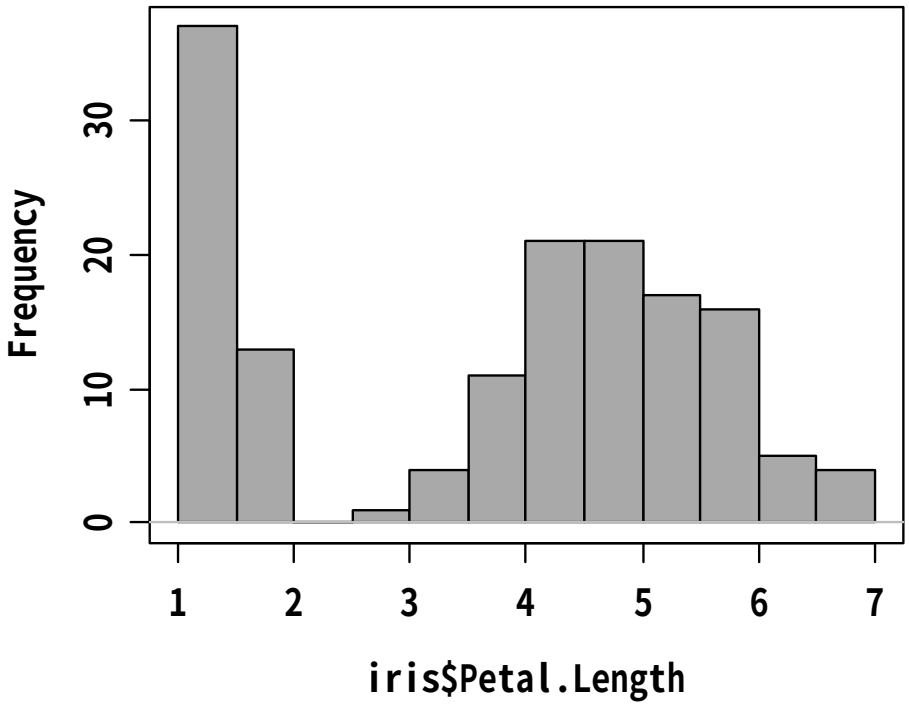
変数を１つ指定する



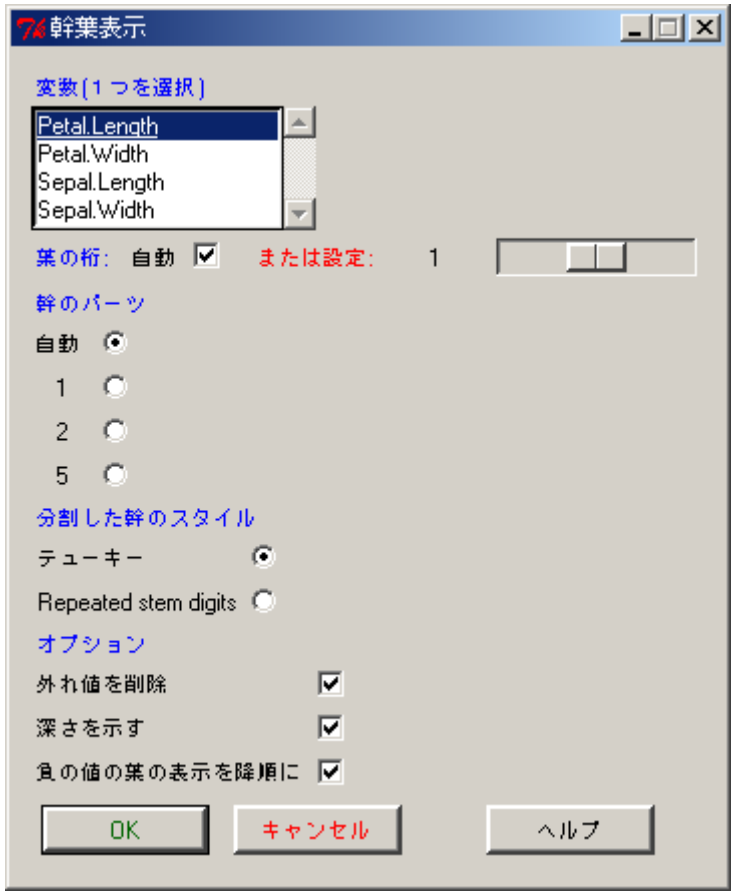
ヒストグラム



変数を１つ指定する



幹葉表示

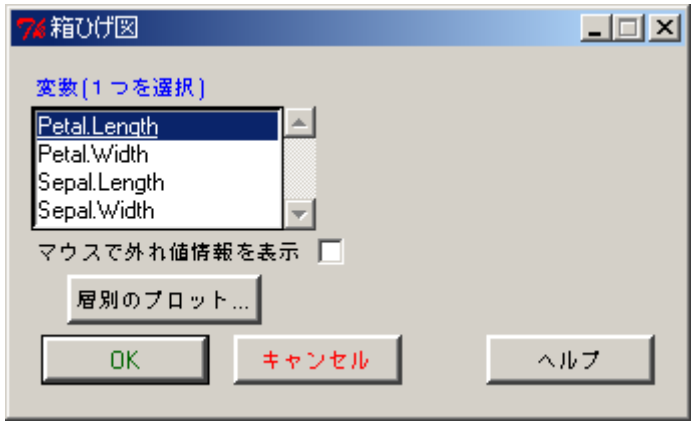


変数を１つ指定する

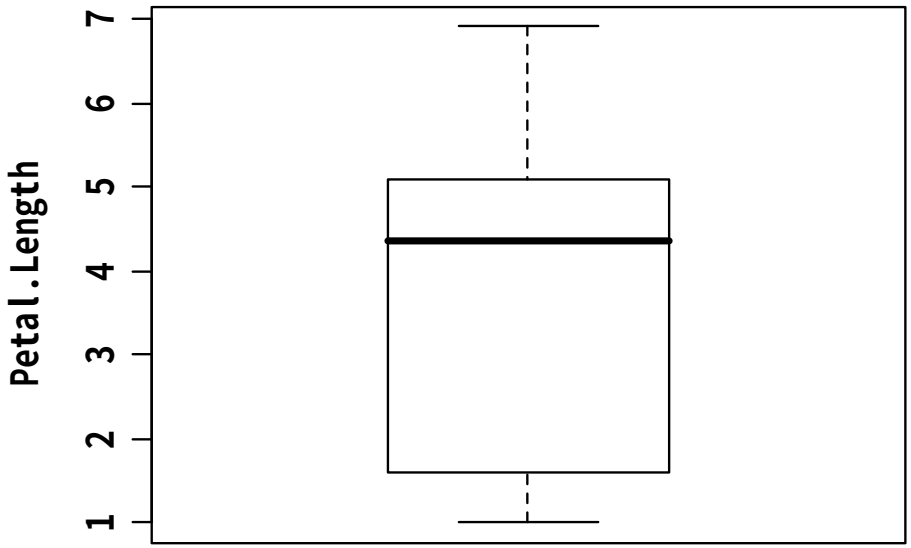
```
1 | 2: represents 1.2
 leaf unit: 0.1
         n: 150

   24     1* | 012233333334444444444444
   50     1. | 55555555555556666666777799
          2* |
          2. |
   53     3* | 033
   61     3. | 55678999
  (18)    4* | 000001112222334444
   71     4. | 5555555566677777888899999
   46     5* | 000011111111223344
   28     5. | 55566666677788899
   11     6* | 0011134
    4     6. | 6779
```

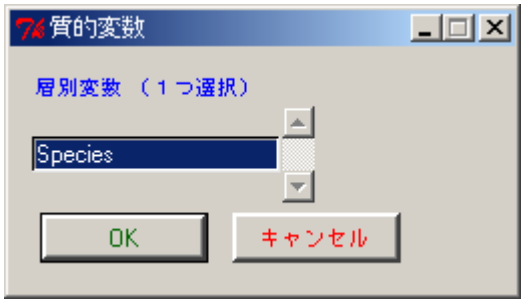
箱ひげ図



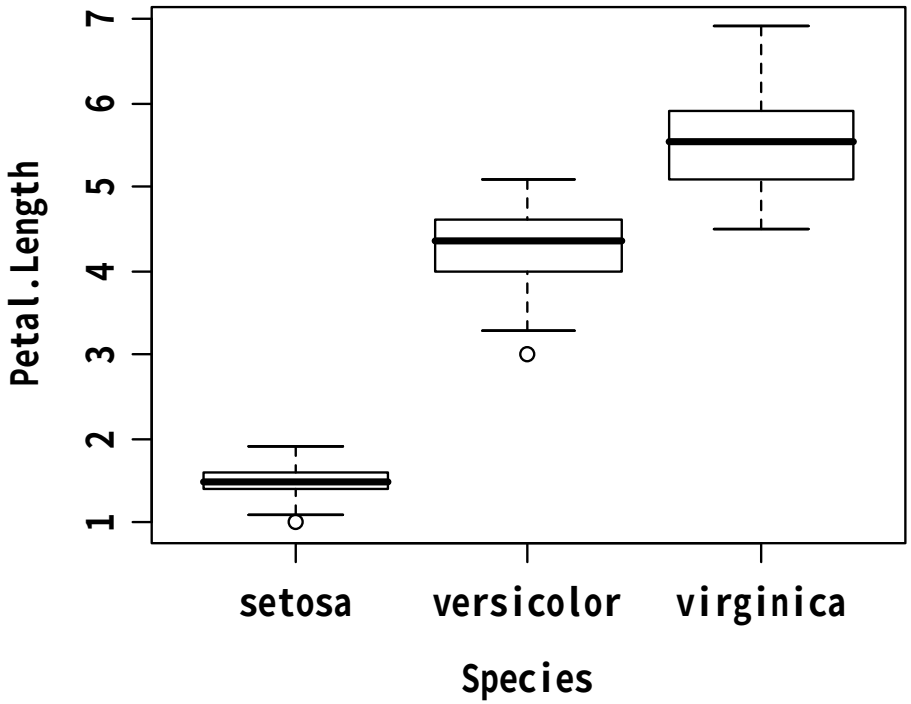
変数を１つ指定する



箱ひげ図（層別）

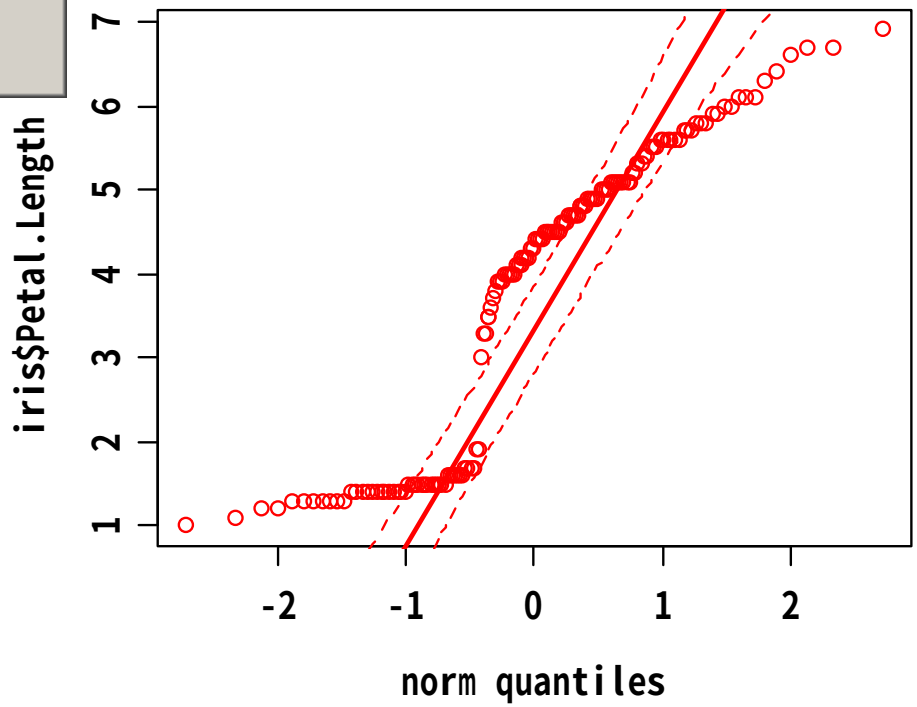


前のページの画面で「層別のプロット」
を選択することで
カテゴリ変数で層別したグラフを出力

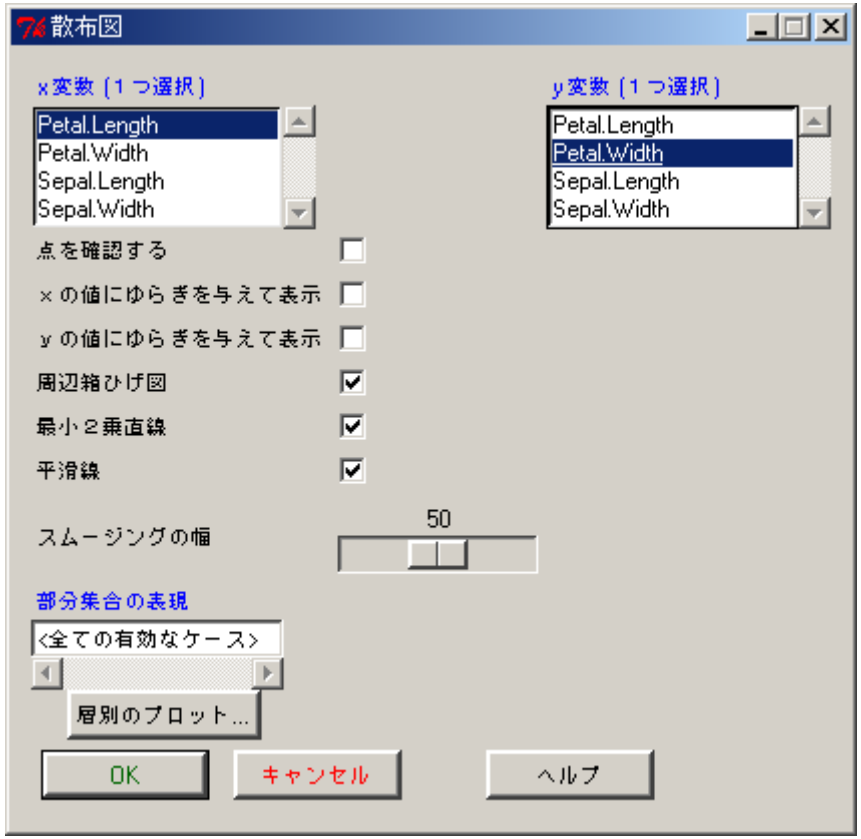


QQプロット

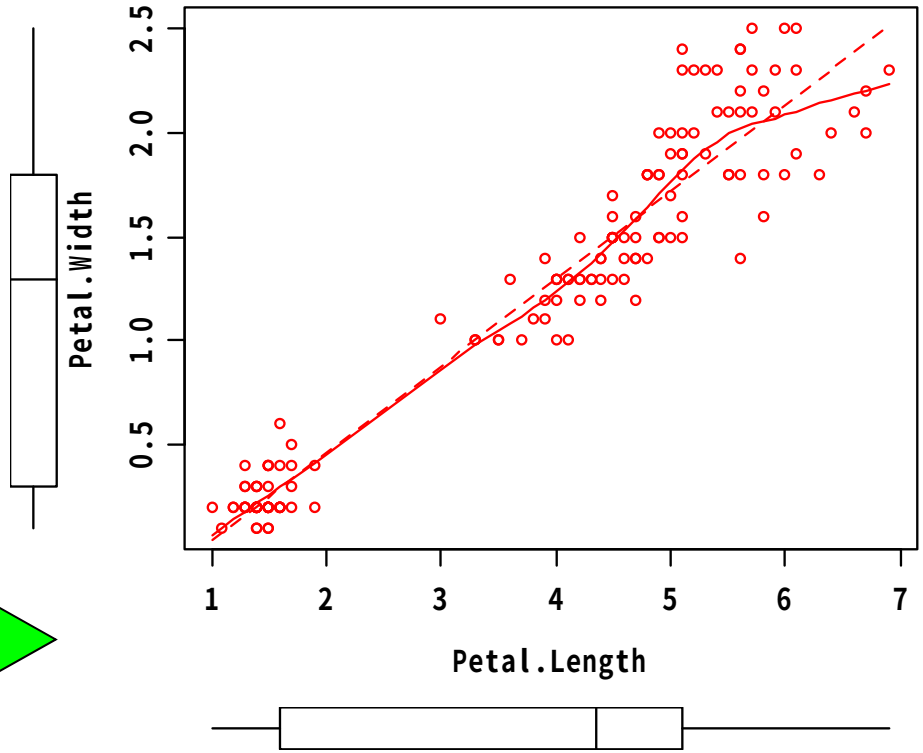
変数を１つ指定する



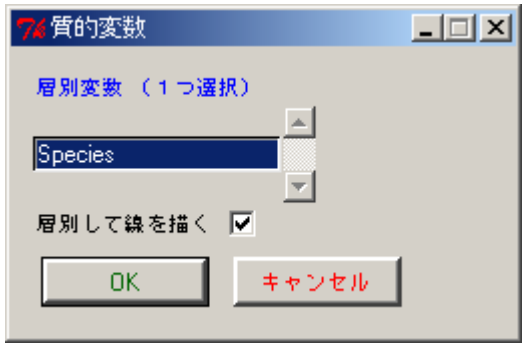
散布図



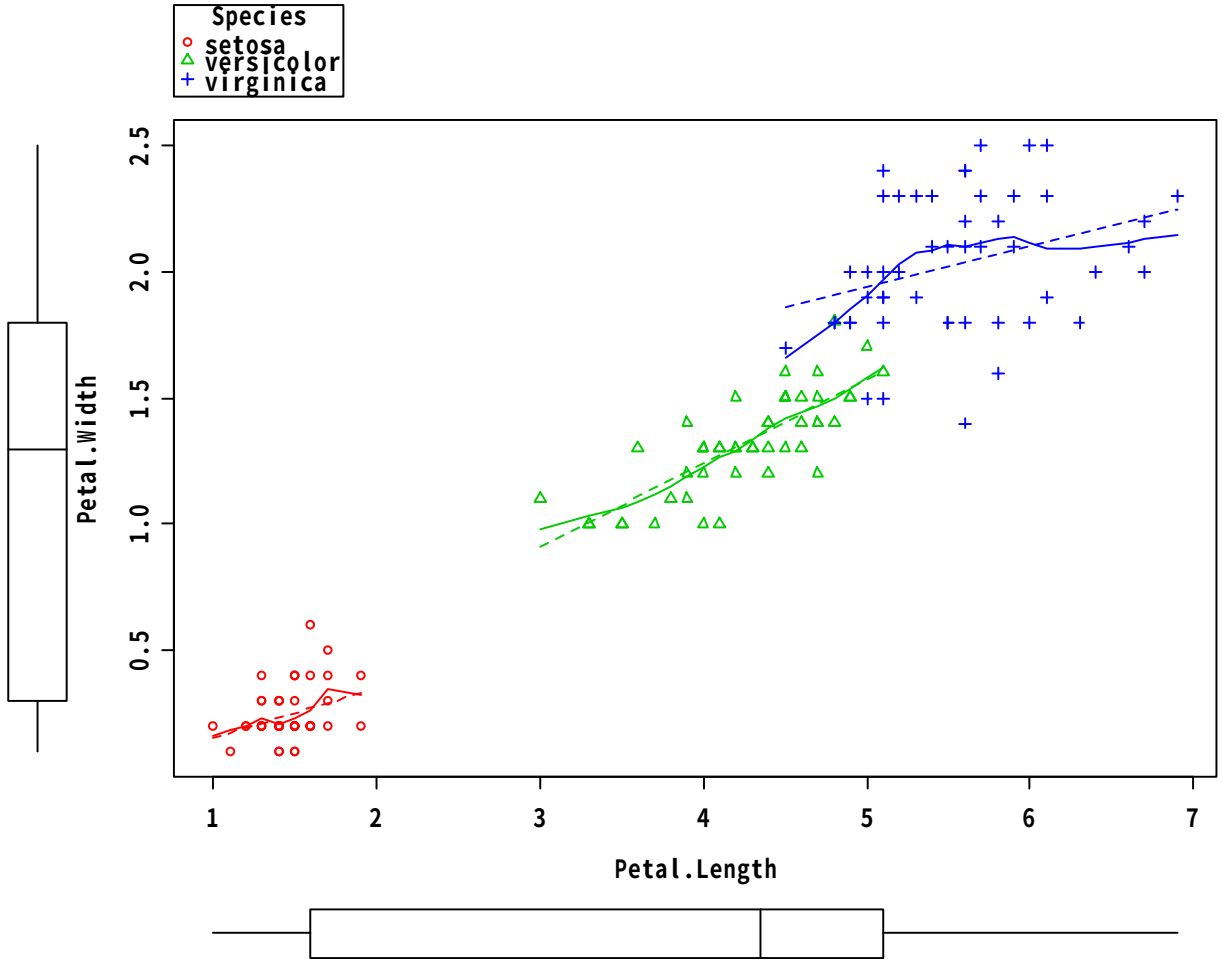
X軸の変数を１つ
Y軸の変数を１つ
指定する



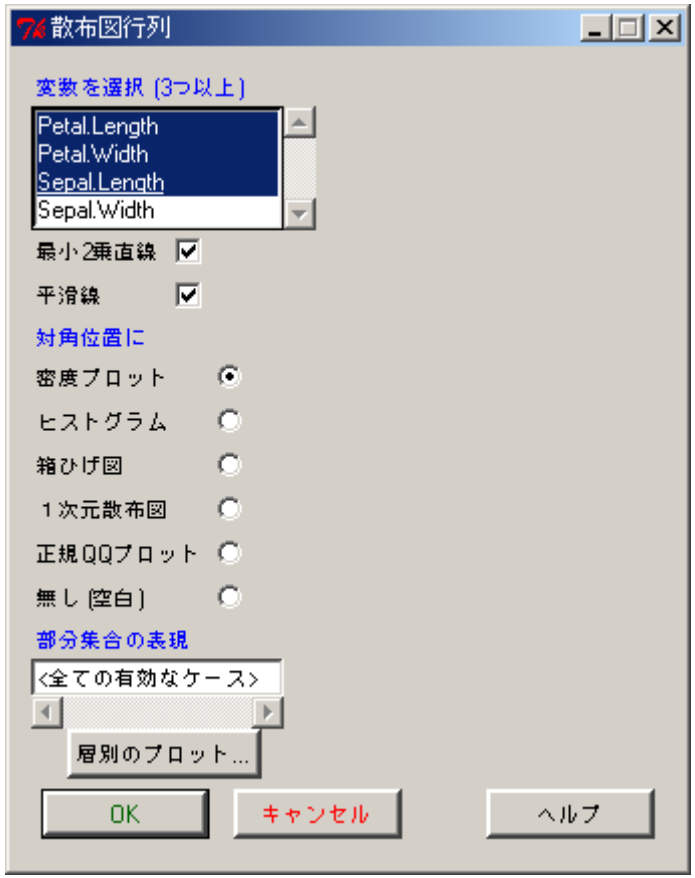
散布図（層別）



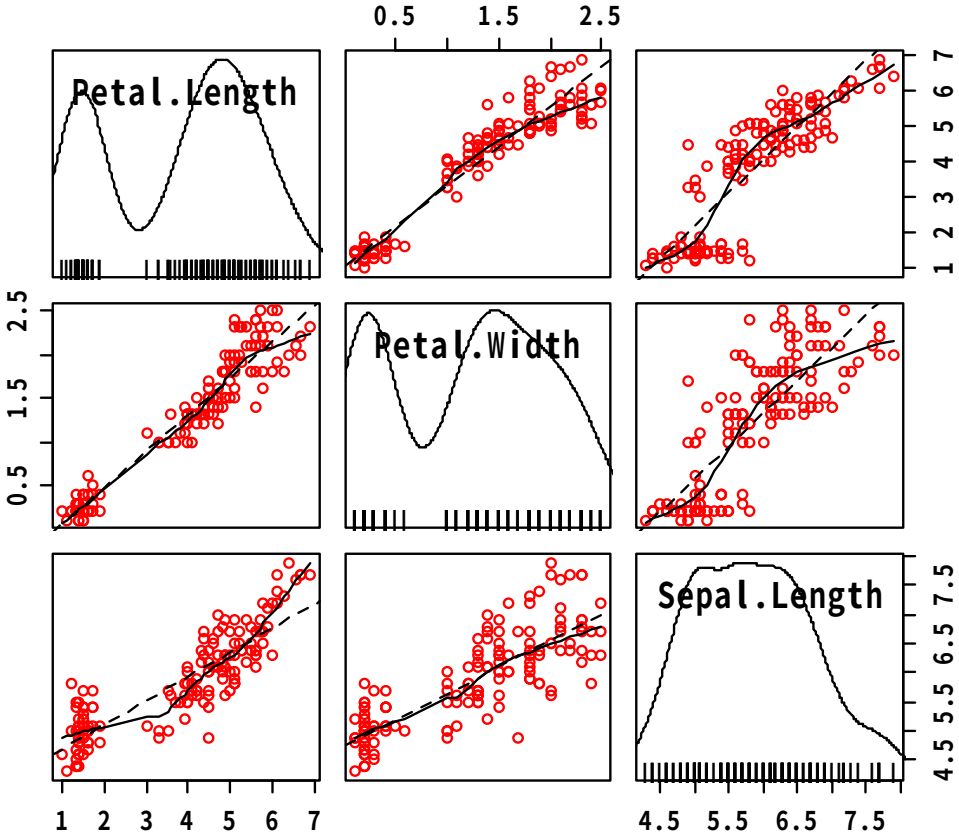
前のページの画面で
「層別のプロット」
を選択することで
カテゴリ変数で層別
したグラフを出力



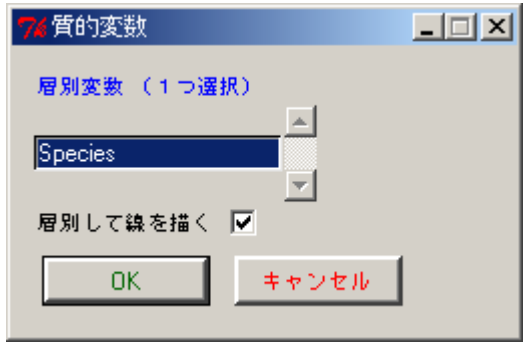
散布図行列

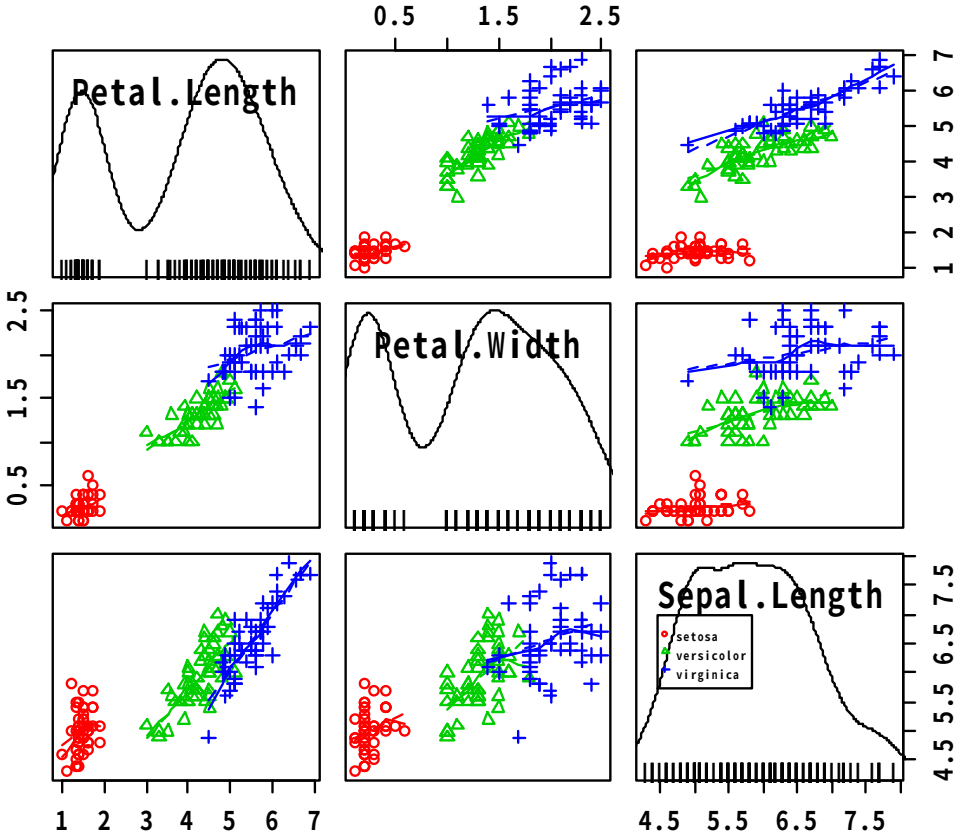


変数を３つ以上指定する



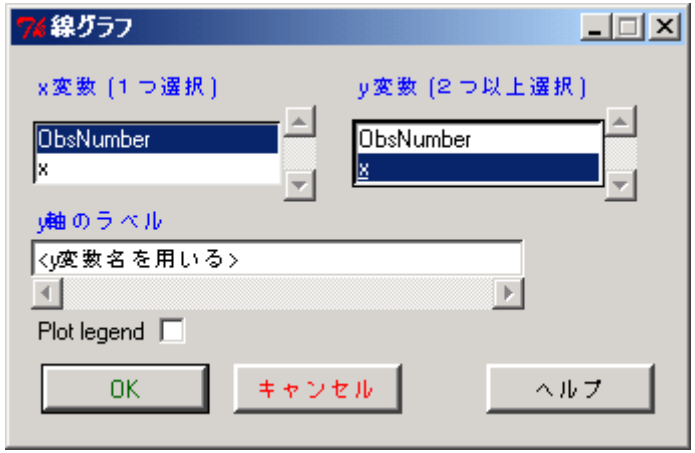
散布図行列（層別）

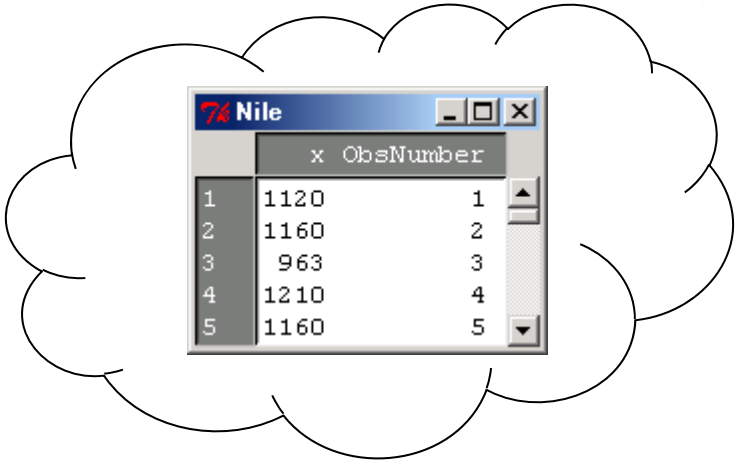




前のページの画面で「層別のプロット」を選択することで
カテゴリ変数で層別したグラフを出力

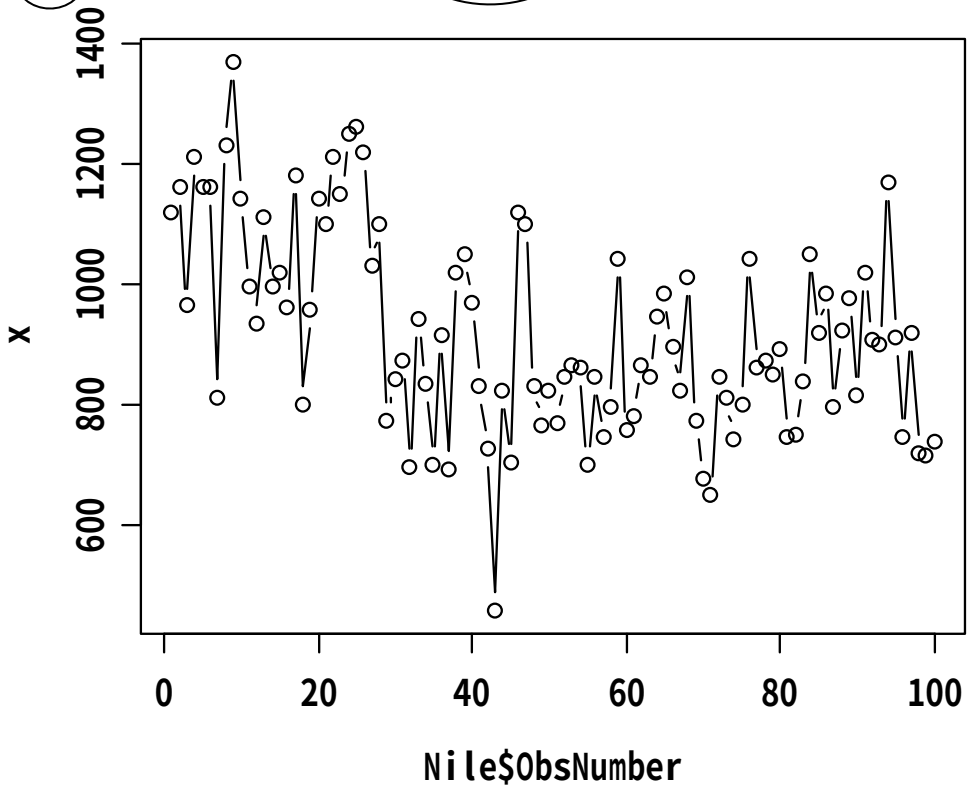
折れ線グラフ



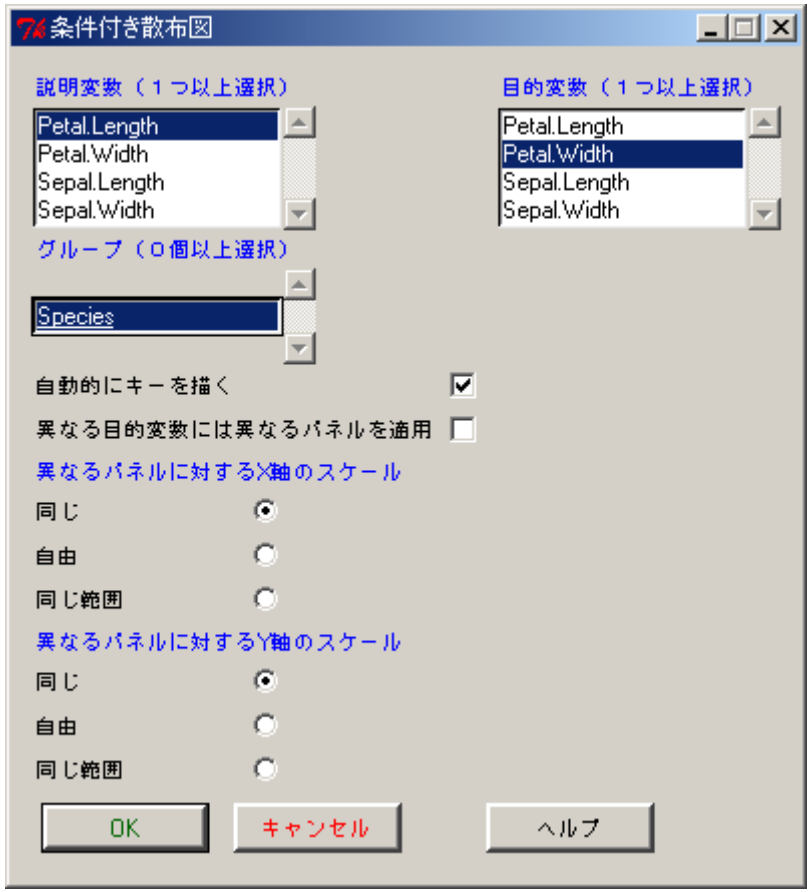



```
Nile <- as.data.frame(Nile)
```

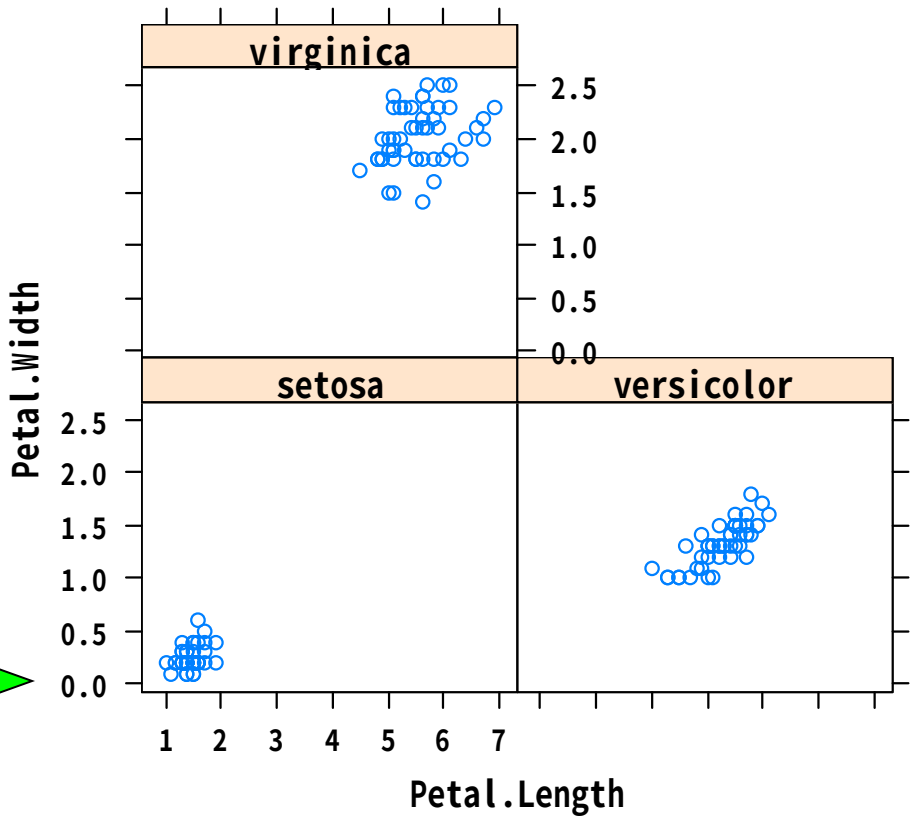
に観測値番号を追加したデータセット
「Nile」について折れ線グラフを作成する



条件付き散布図

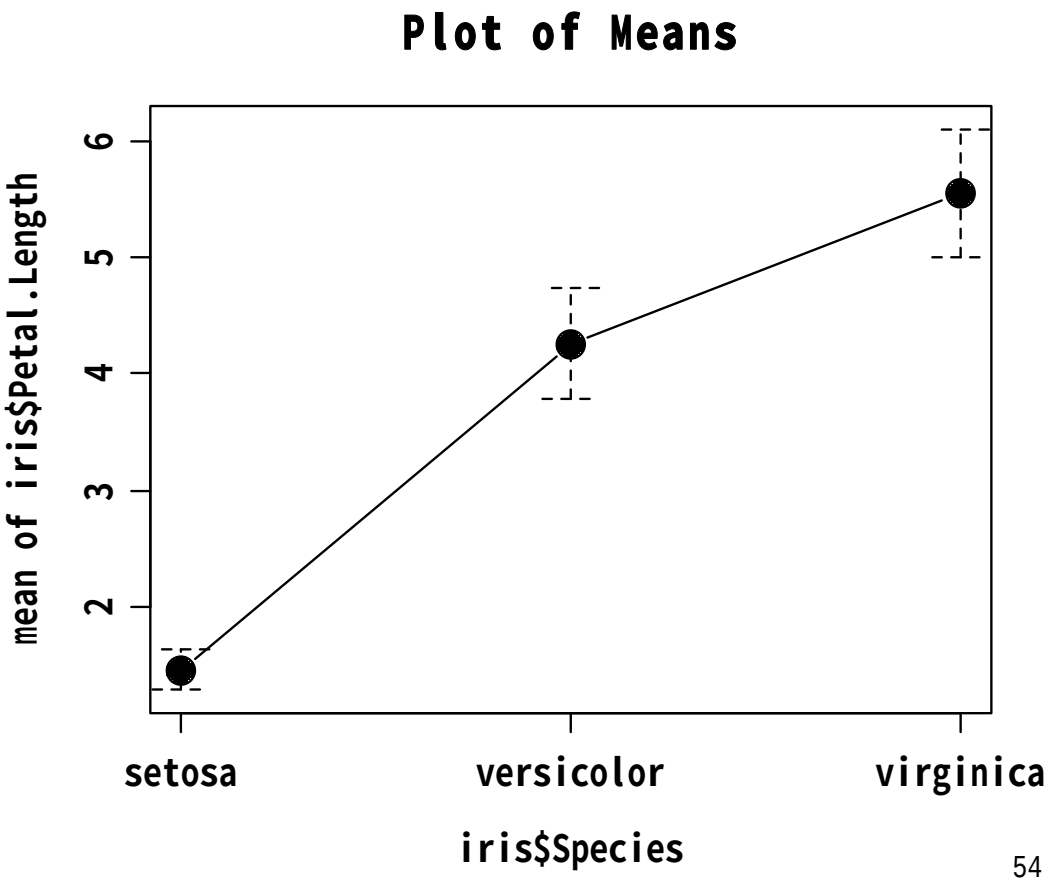


X軸変数を１つ
Y軸変数を１つ
層別変数を１つ
指定する



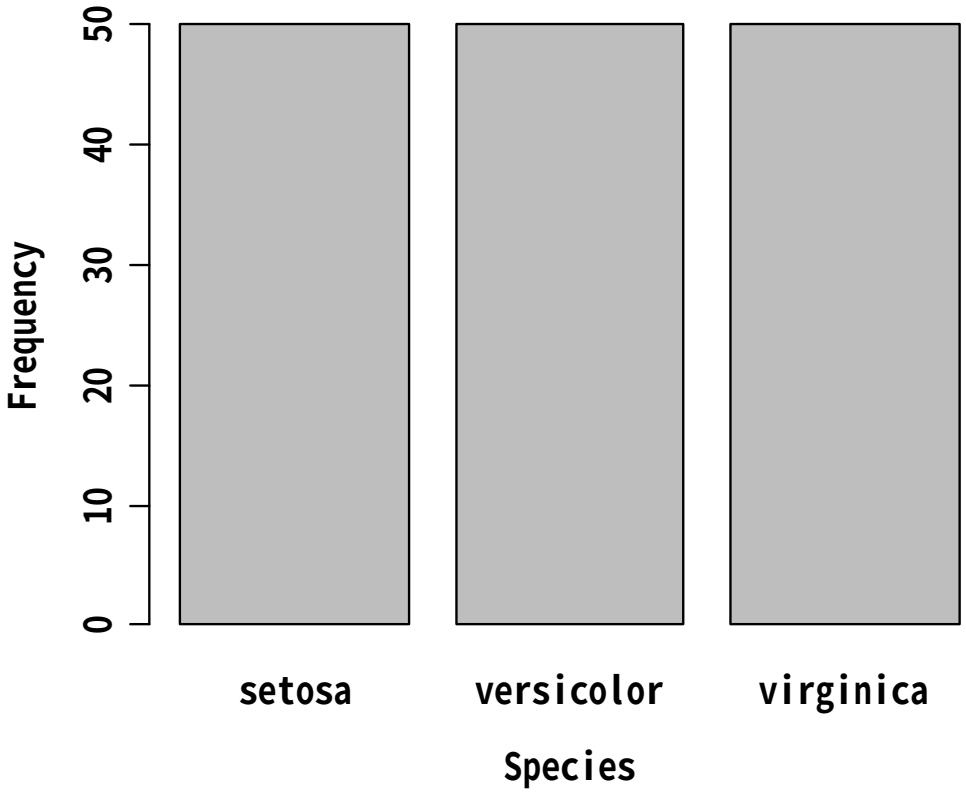
平均のプロット

カテゴリ（因子）変数と目的変数を指定する



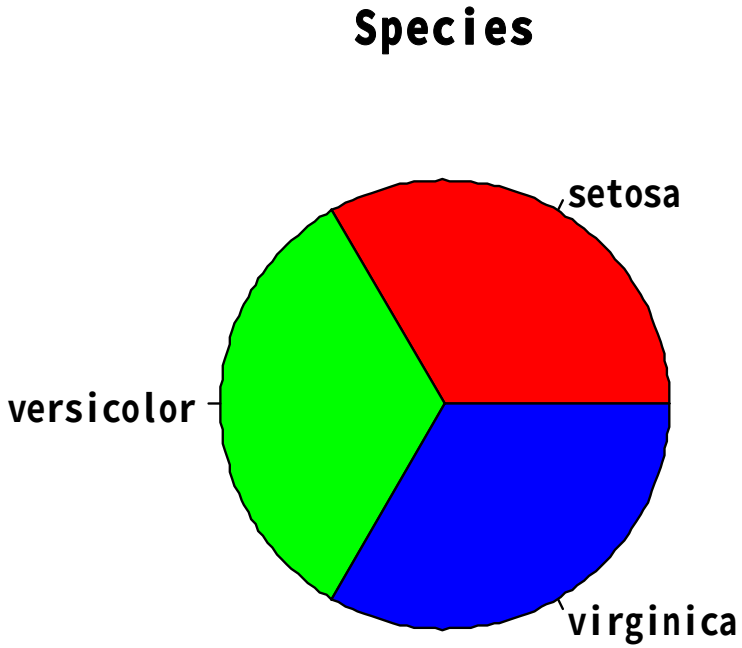
棒グラフ

変数を１つ指定する

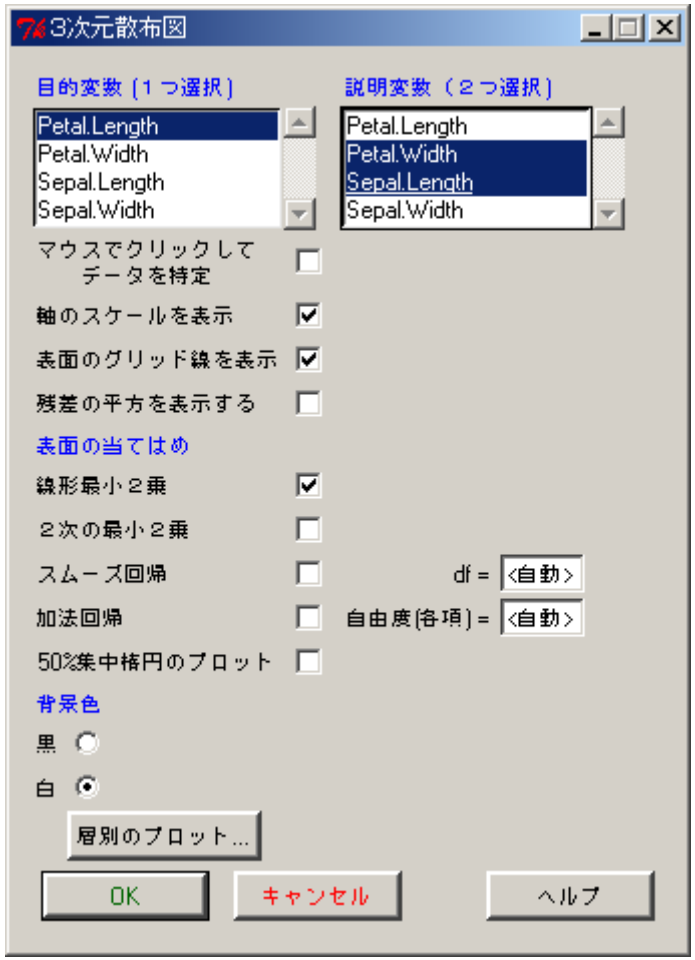


円グラフ

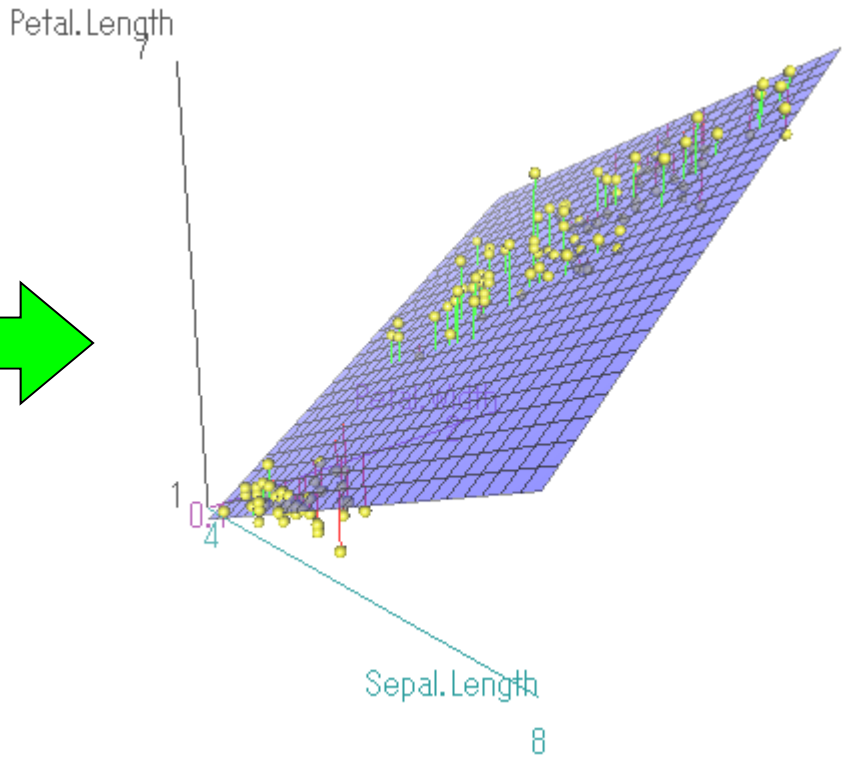
変数を１つ指定する



3Dプロット

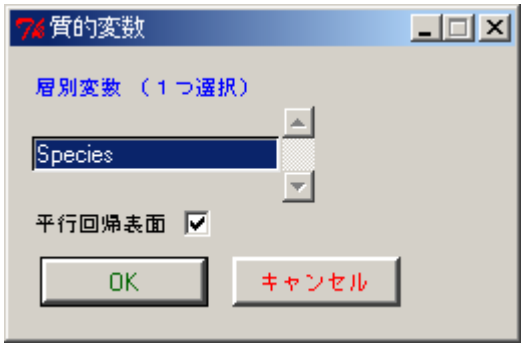


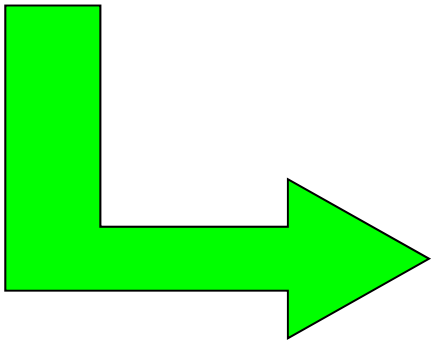
目的変数を１つ，説明変数を２つ指定する

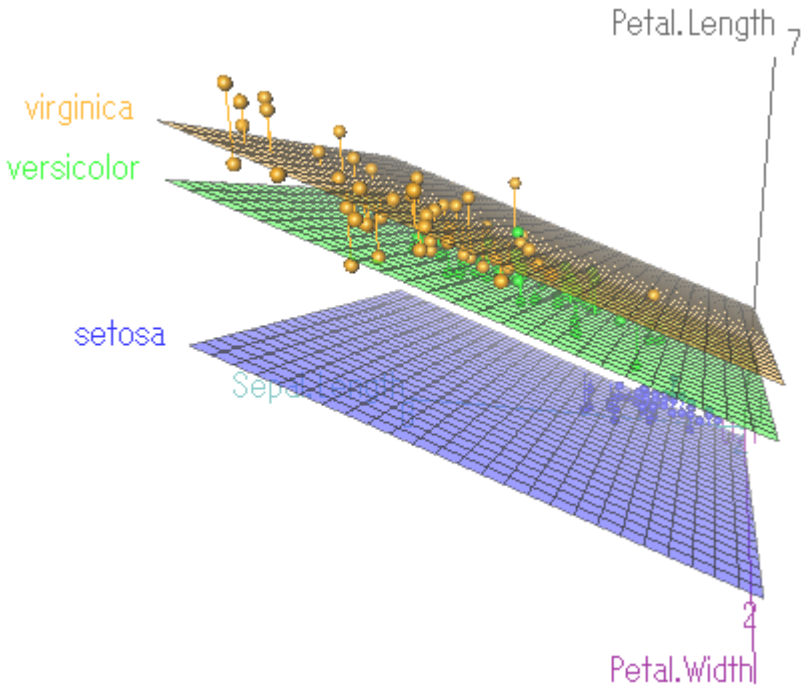


マウスでグラフを動かすことが出来る！

3Dプロット（層別）







前のページの画面で「層別のプロット」
を選択することで
カテゴリ変数で層別したグラフを出力

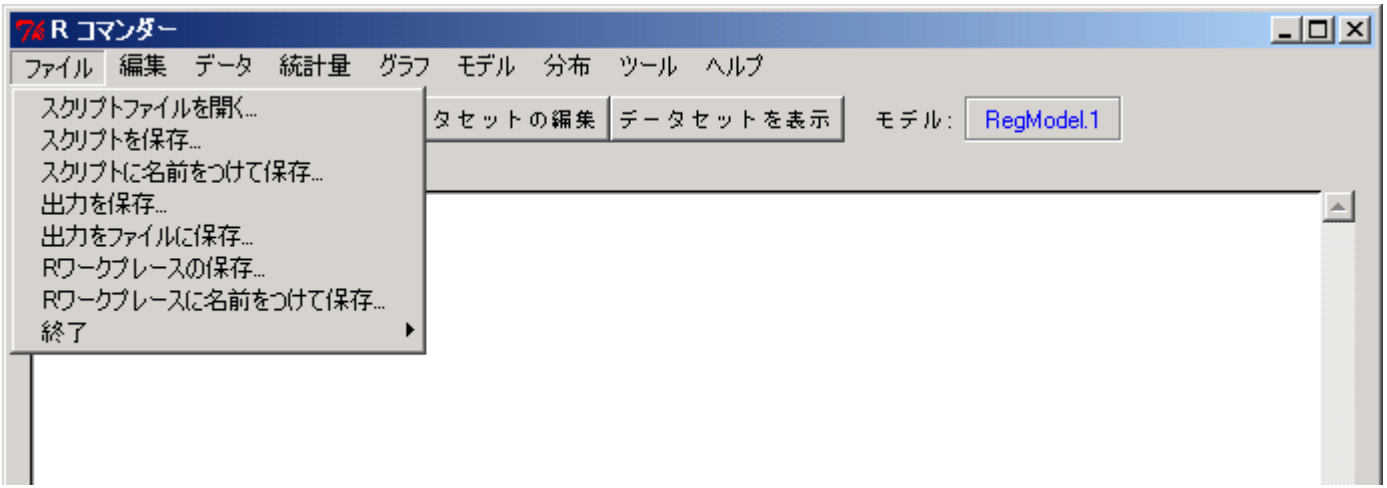
マウスでグラフを
動かすことが出来る！

メニュー

- R Commander の概要
- データ「iris」の読み込み方法
- グラフの作成方法
- R Commander でのデータ解析例
- グラフのカタログ
- 各メニューの概要 ←

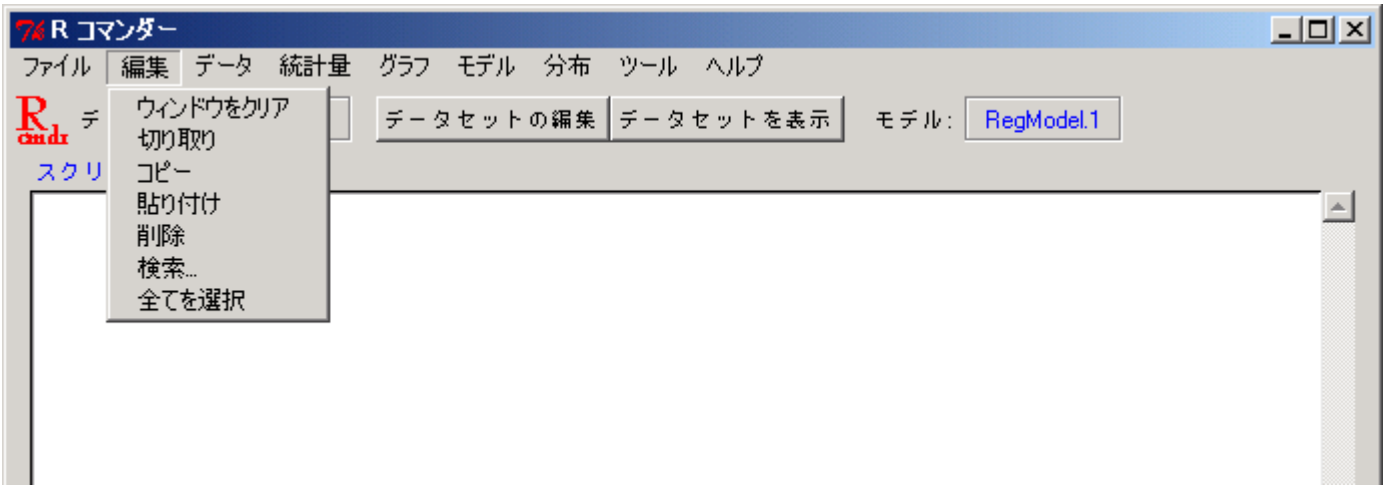
Graphic by (c)Tomo.Yun (http://www.yunphoto.net)

メニュー〔ファイル〕

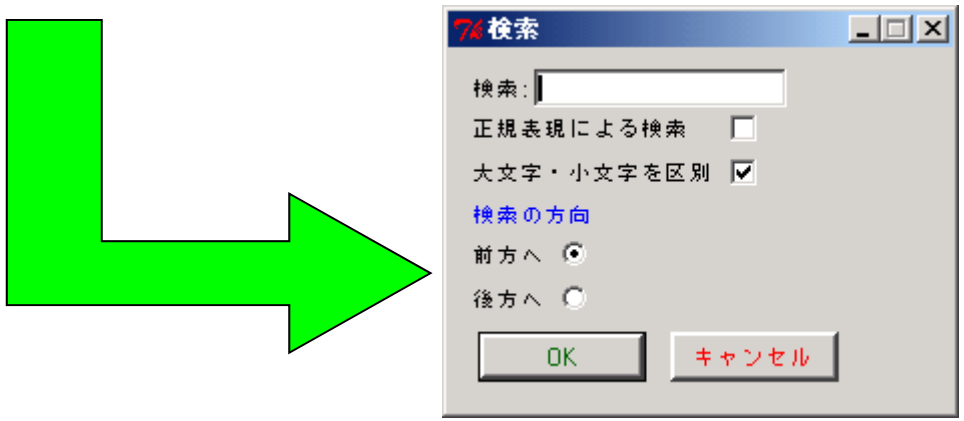


- **スクリプトファイルを開く**：R のプログラムファイルを開く
- **スクリプトを保存，スクリプトに名前をつけて保存**：スクリプトウィンドウの内容をファイルに保存
- **出力を保存，出力をファイルに保存**：ログウィンドウの内容をファイルに保存
- **R ワークプレースの保存， R ワークプレースに名前をつけて保存**：R の現在の作業内容（データ，関数など）をファイルに保存
- **終了**：R Commander を終了する（R 本体は終了しない）

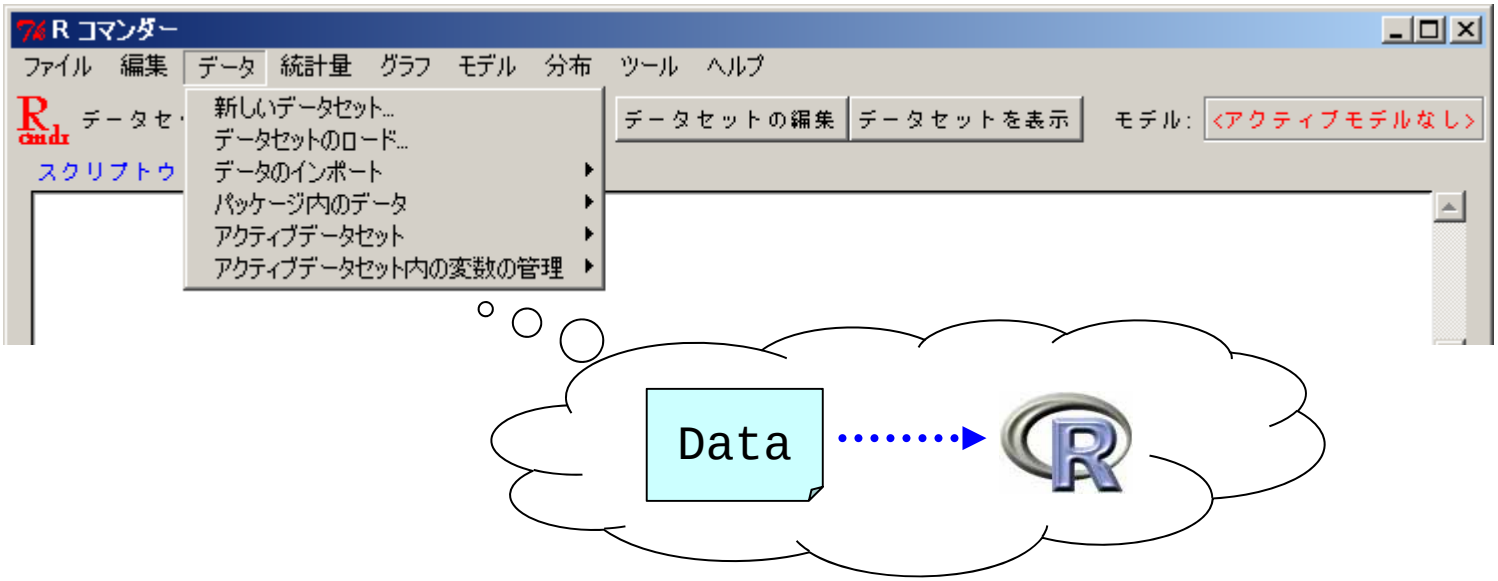
メニュー〔編集〕



- **ウィンドウをクリア：**スクリプトウィンドウまたはログウィンドウの内容を消去する（カーソルがある方のウィンドウが対象となる）
- **切り取り，コピー，貼り付け，削除，全てを選択：**（普通の編集機能）
- **検索：**カーソルがある方のウィンドウを対象として，文字列の検索を行う



メニュー〔データ〕のおさらい



- **新しいデータセット：**セル形式のウィンドウにデータを手入力する
- **データセットのロード：R** 上で作成したデータ（.rda）を読み込み
- **データのインポート：**txt，SPSS，Minitab，STATA，EXCEL，Access，dBase形式のデータファイルを読み込む
- **パッケージ内のデータ：R** に用意されているサンプルデータを読み込む
- **アクティブデータセット：**解析用データの選択やデータの加工を行う
- **アクティブデータセット内の変数の管理：**データの加工を行う

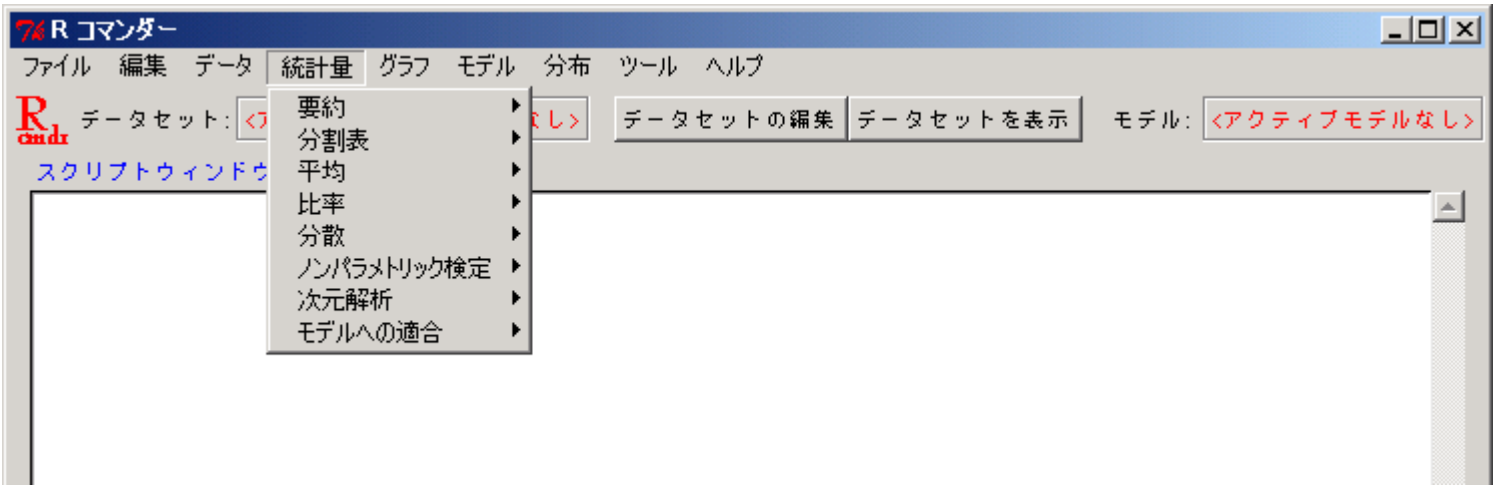
メニュー〔データ〕参考

■ R では以下の外部データを読み込む機能が用意されている

関数	用途
data.restore()	S3 Binary ファイルの読み込み
read.dbf()	DBF ファイルの読み込み
read.dta()	Stata バイナリファイルの読み込み
read.epiinfo()	Epi Info ファイルの読み込み
read.mtp()	Minitab Portable Worksheet の読み込み
read.octave()	Octave テキストデータファイルの読み込み
read.spss()	SPSS データファイルの読み込み
read.ssd()	SAS データセット（.sas7bdat）の読み込み
read.systat()	Systat ファイルの読み込み
read.xport()	SAS XPORT ファイルの読み込み
odbcConnect()※	EXCEL, ACCESS, … ファイルの読み込み

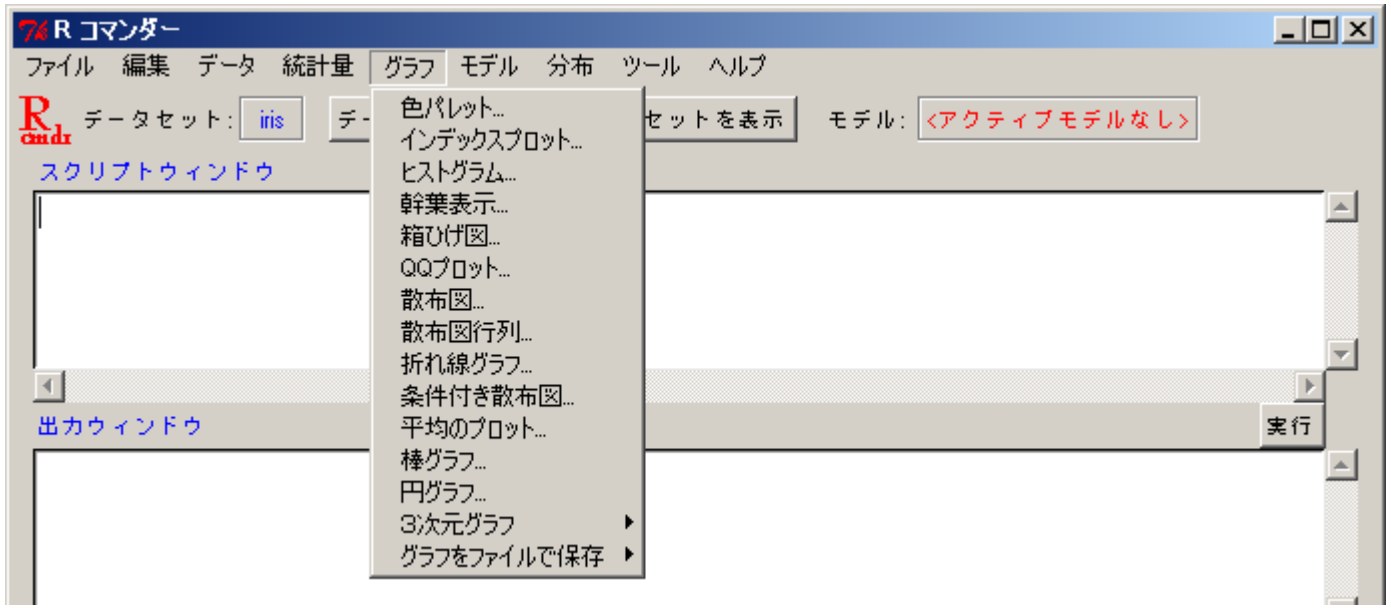
※ はパッケージ「RODBC」の関数，その他はパッケージ「foreign」の関数

メニュー〔統計量〕



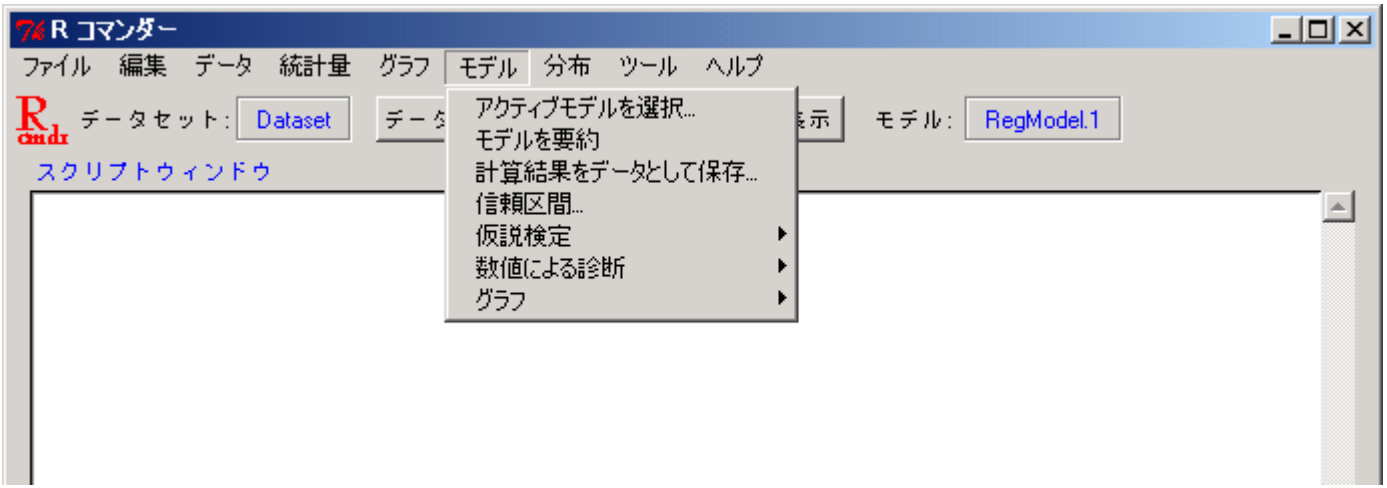
- **統計量** → 様々な統計量の算出や検定の実行，モデルの作成が出来る
  - 要約統計量，頻度集計，相関係数の算出
  - 分割表の作成，分割表に対する検定
  - 平均値に対する検定（ｔ検定，分散分析），比率データに対する検定
  - 分散についての検定（Ｆ検定，バートレットの検定）
  - ノンパラ検定（ウィルコクソン検定，クラスカル・ウォリス検定）
  - 次元解析（測定の信頼性，主成分分析，因子分析，クラスター分析）
  - モデルの作成
- 〔統計量のメニュー選択〕→〔データの変数等を指定〕という手順

メニュー〔グラフ〕



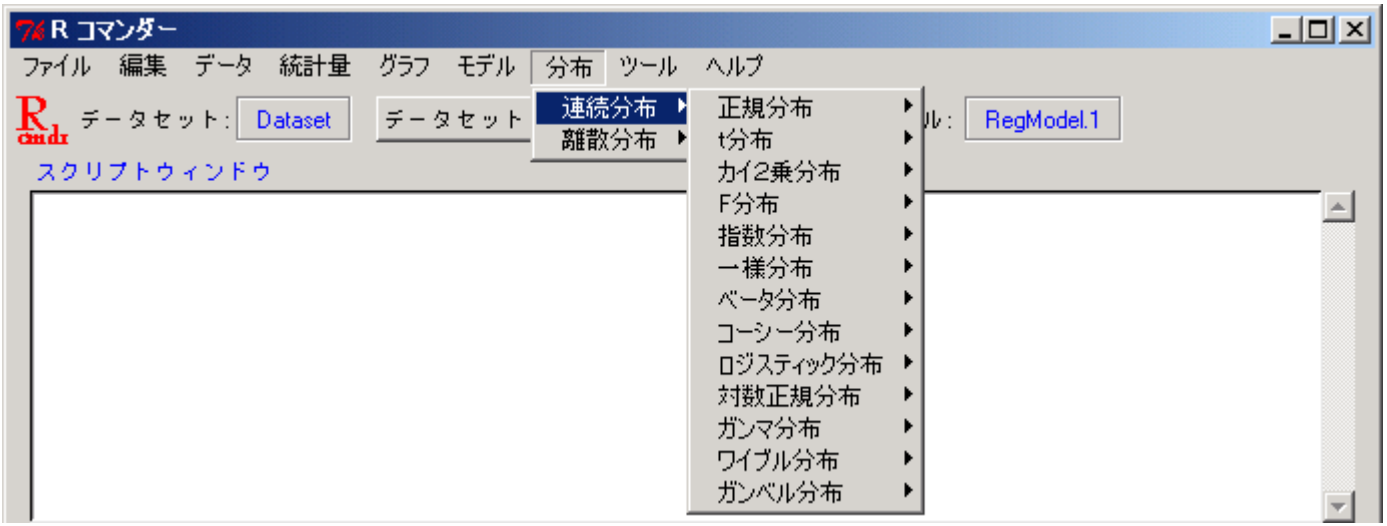
- グラフ → 様々な種類のグラフを描くことが出来る
  ⇒ インデックスプロット，ヒストグラム，幹葉表示（幹葉図），箱ひげ図，QQプロット，散布図，散布図行列，折れ線グラフ，条件付き散布図，平均のプロット，棒グラフ，円グラフ，3Dグラフ
- グラフをファイルに保存することも出来る

メニュー〔モデル〕



- モデル → メニューの「統計量」→「モデル」で作成したモデルについて詳細な検討を加えることが出来る
  - モデルの要約
  - 信頼区間の算出
  - 仮説検定
  - モデルの診断
  - モデルに関するグラフ描画

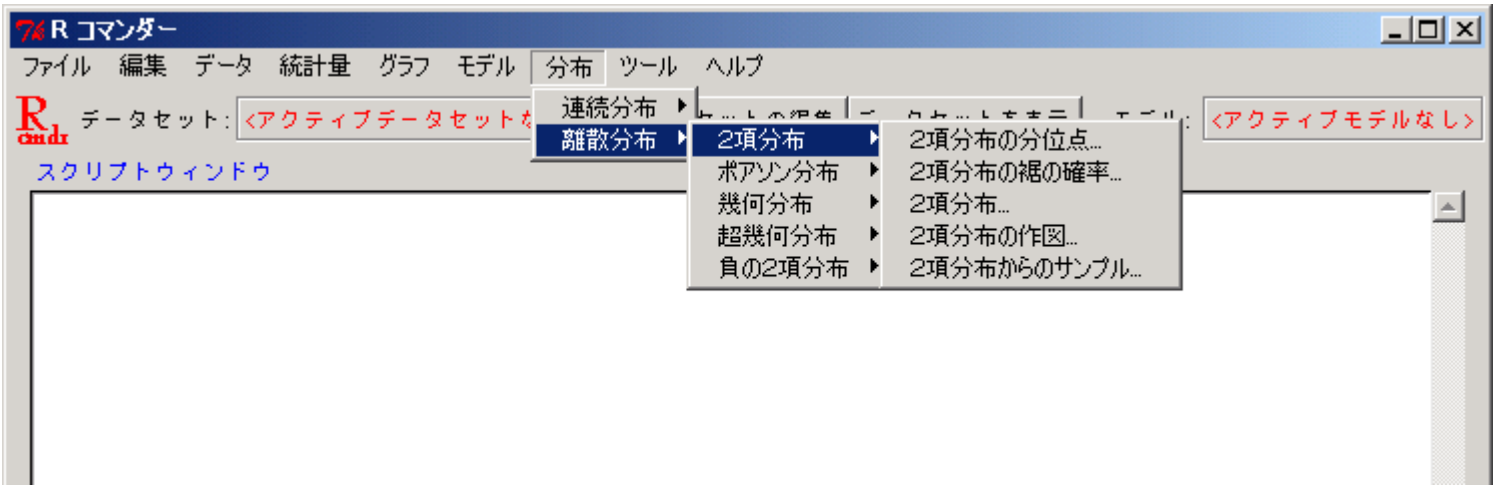
メニュー〔分布〕連続分布



- **正規分布，t 分布，χ2分布，F分布，指数分布，一様分布，ベータ分布，コーシー分布，ロジスティック分布，対数正規分布，ガンマ分布，ワイブル分布，ガンベル分布（二重指数分布）**について・・・

  ⇒ 累積分布の算出，確率点の算出，乱数の算出，グラフの描画を行う

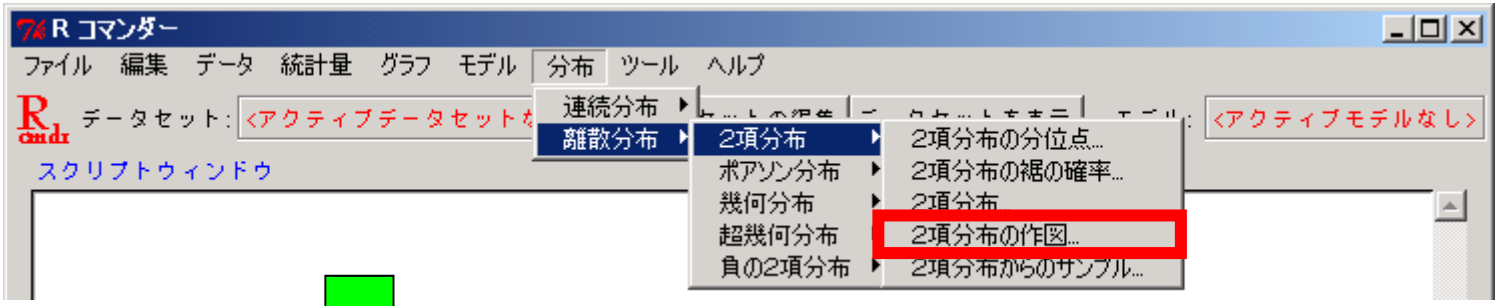
メニュー〔分布〕離散分布

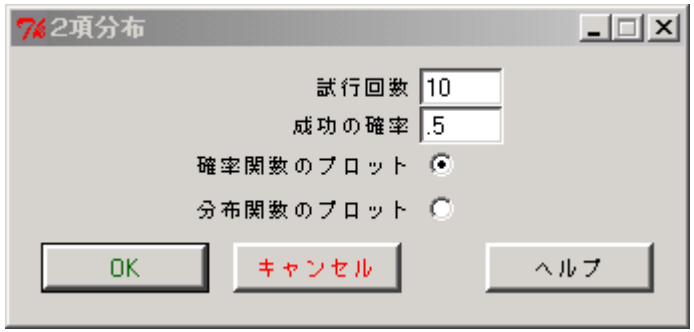


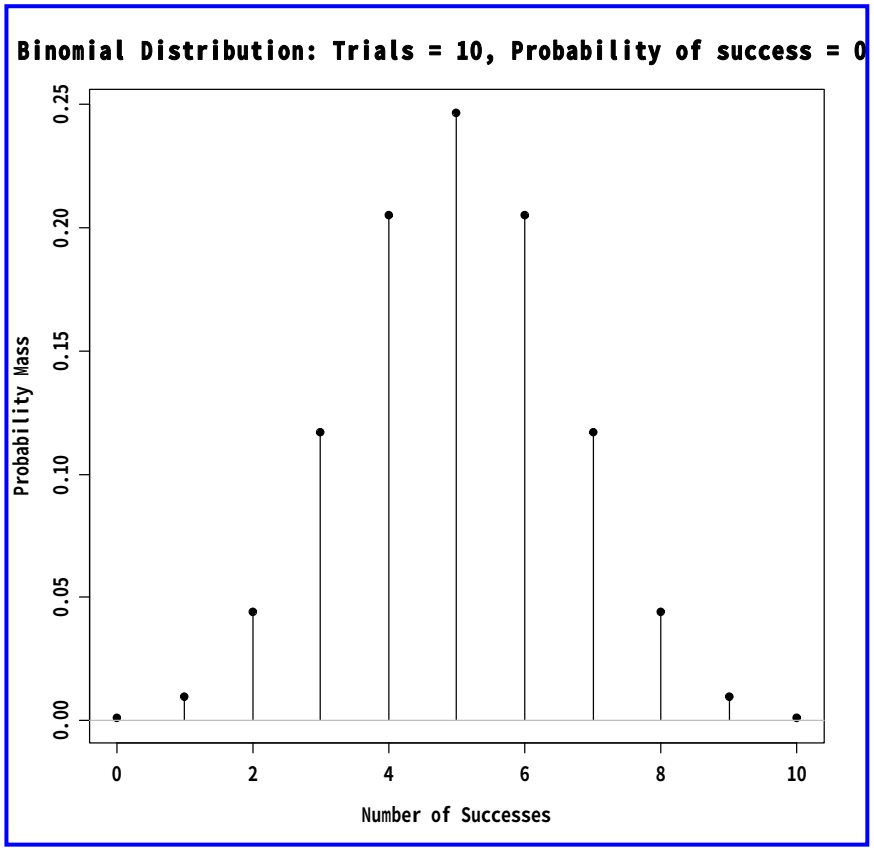
■ **２項分布，ポアソン分布，幾何分布，超幾何分布，負の２項分布**について・・・

⇒ 累積分布の算出，確率点の算出，確率，乱数の算出，グラフの描画を行う

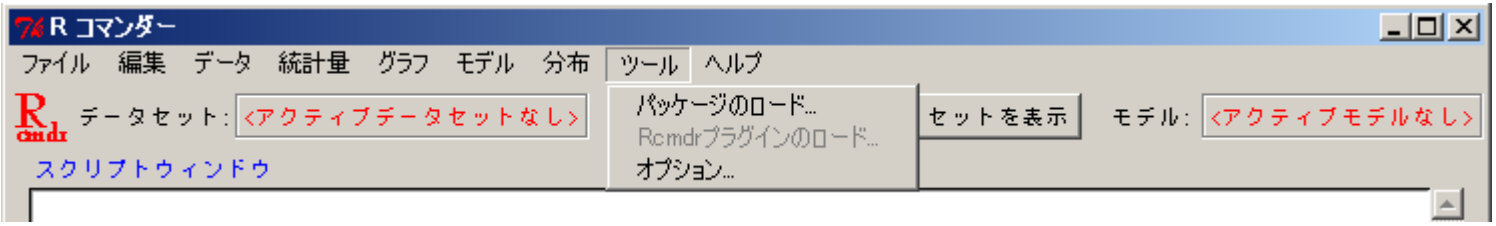
メニュー〔分布〕例：２項分布のグラフ描画



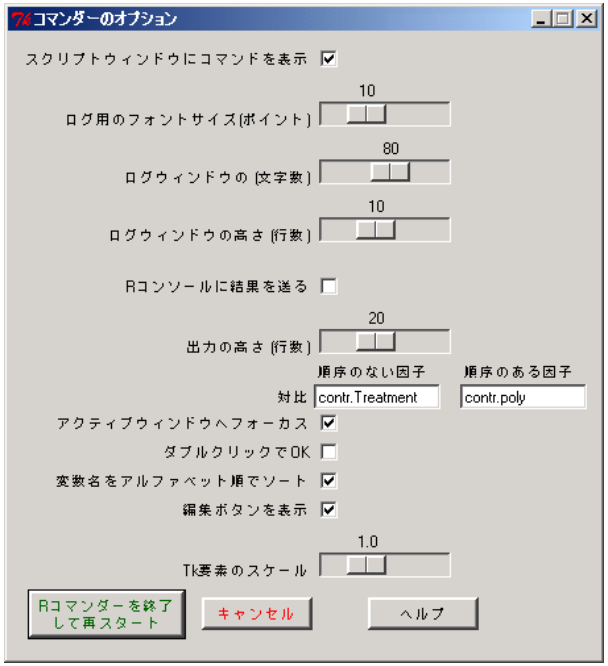




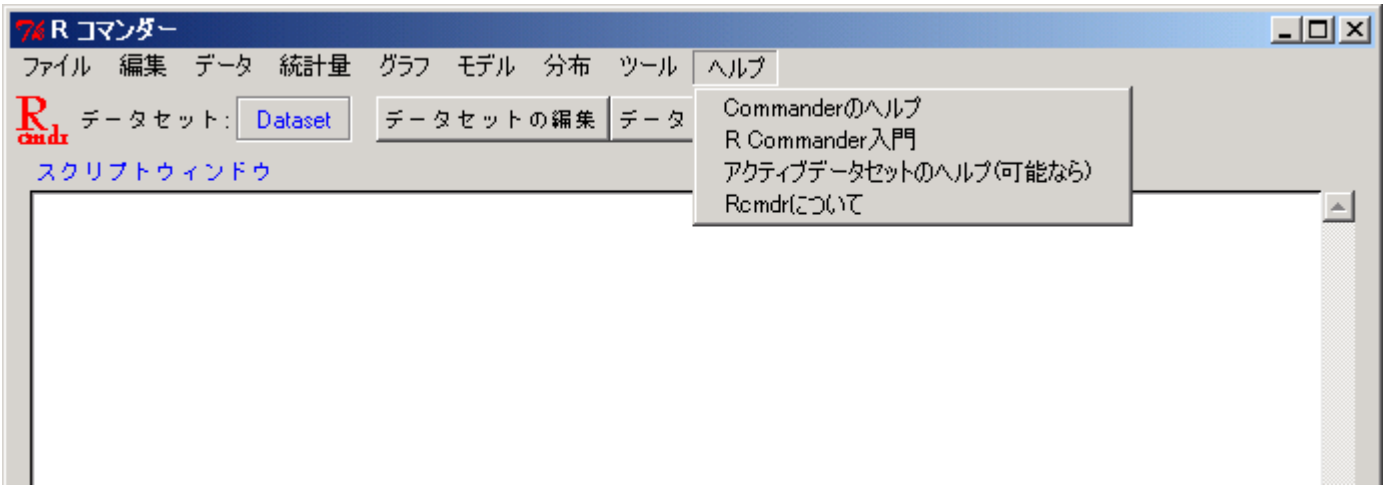
メニュー〔ツール〕



- **パッケージのロード：**
  R のパッケージを呼び出す
- **Rcmdr プラグインのロード：**
  R Commander 用のプラグイン（追加機能）を呼び出す
- **オプション：**R Commander のウィンドウの表示設定を変更する



メニュー〔ヘルプ〕



- **Commander のヘルプ**：R Commander のヘルプを表示
- **R Commander**：R Commander の作者・John Fox 氏の解説文書「Getting Started With the R Commander.」を表示
- **アクティブデータセットのヘルプ**：データセットのヘルプを表示（Rに用意されているデータセットを開いている場合）
- **Rcmdr について**： R Commander の概要を表示

参考文献〔R Commander 関係〕

- フリーソフトウェア R による統計的品質管理入門
  （荒木 孝治 編著，日科技連）
- R と R コマンダーではじめる多変量解析
  （荒木 孝治 編著，日科技連）
- Getting Started With the R Commander（John Fox）
  http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/
- R with Rcmdr: BASIC INSTRUCTIONS
  （Murray Logan）
  http://www.zoology.unimelb.edu.au/stats/Eworksheets/tutorials/RmanualScreen.pdf
- R Commander ハンドブック（舟尾 暢男 著，九天社）