Supplementary material for JMP-compatible exercises

How to install JMP Pro

JMP PRO

- Analysis package JMP is now available for use at our universities (Univ. license)
  - This package can be downloaded from MMC and installed on individual PCs for use on both Win and Mac computers (requires annual renewal).
  - JMP is another excellent software package from SAS. It has a lot of features that are easier to use than SAS,
    - such as machine learning tools, experimental design (DOE), Structured Equation Modeling (SEM), etc.
    - The user interface has been wonderfully improved with each new version..

Please give it a try.

- JMP Pro covers basic and advanced analysis.
- Compared to SAS, JMP Pro is superior in data visualization.
- JMP Pro is useful for creating SAS data sets and for data cleaning, so it should be used in conjunction with SAS.

JMP can be downloaded to your PC by SUMS students, faculty, and staff.
Download instructions are given below.

Go to the SUMS Multimedia Center website.
URL:https://www.shiga-med.ac.jp/mmc/

Download JMP Pro" from the various services

1.。

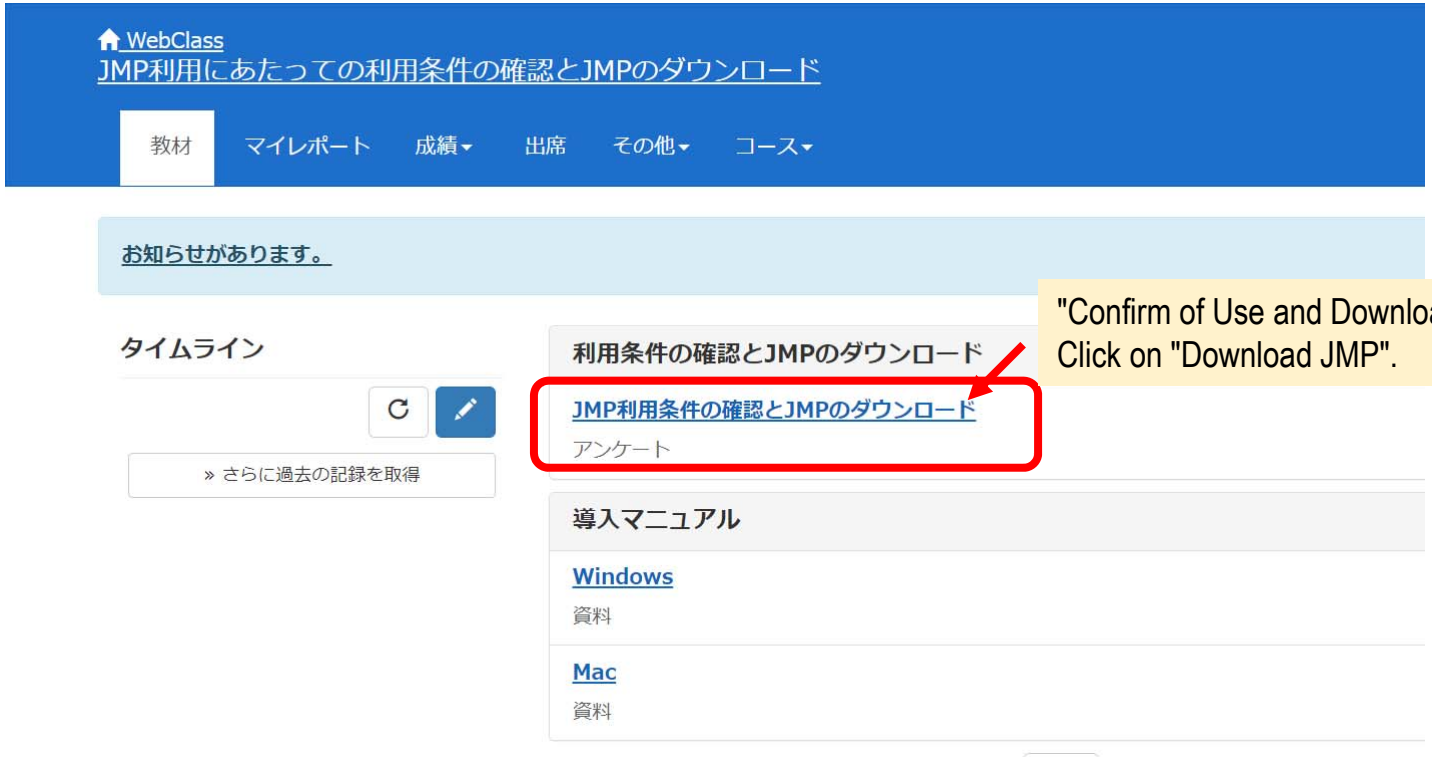
Download JMP files and register as a user

「利用方法・利用申請について」
How to use and application for use

SUMS e-leaningの「JMP利用にあたっての利用条件の確認とJMPのダウンロード」
の
コースを受講する
をクリック

Take the SUMS e-leaning course on "Download JMP and check the terms and conditions for using JMP.
Click on "Download JMP Pro".

Webclass



"Confirm of Use and Download JMP".
Click on "Download JMP".

教材　終了

> JMP利用にあたっての利用条件の確認とJMPのダウンロード

JMP利用条件の確認とJMPのダウンロード

原田　亜紀子

前のページ　[次のページ]

終了

問 1

> 問 2

Download the JMP from the following link

以下のリンクからダウンロードください。 Select WIN or MAC to download

Windows版： Click on “Download JMP” （WINDOWS）

https://gakunai.shiga-med.ac.jp/mmc/service/software/jmp/zip/SAS_Software_Depot_win.zip

Mac版： Click on “Download JMP” （MAC）

https://gakunai.shiga-med.ac.jp/mmc/service/software/jmp/zip/SAS_Software_Depot_mac.zip

学内からのみダウンロードできます。 Downloading is only available on campus.

前のページ　[次のページ]

終了

When the download is complete, click the "Finish" button.

The following files will be downloaded

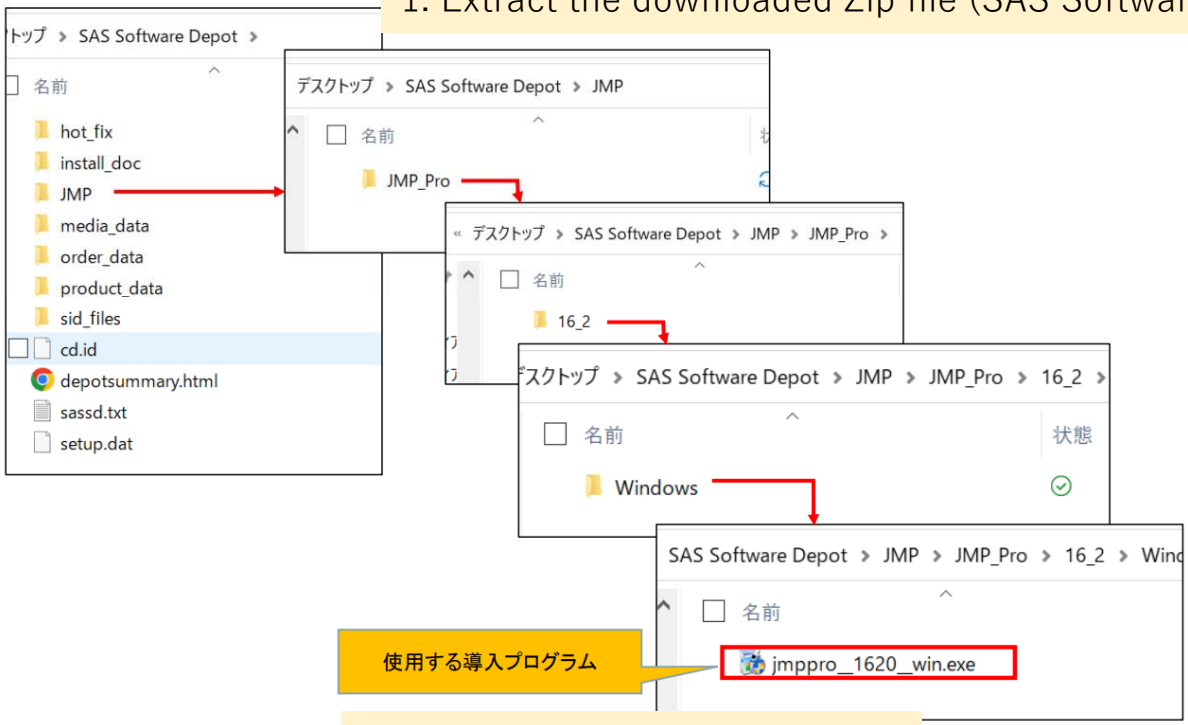
SAS_Software_Depot_mac.zip

SAS_Software_Depot_win.zip

(Win)

1. ダウンロードしたZipファイル（SAS Software Depot_Win.zip）を解凍します。

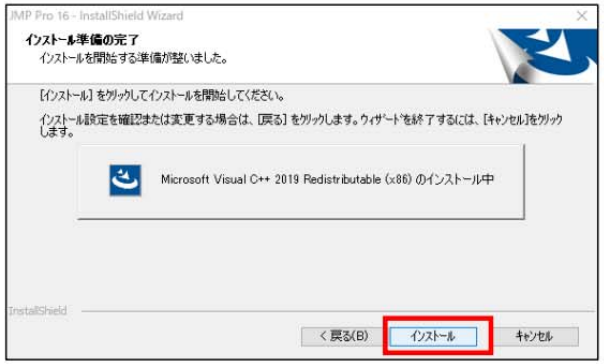
1. Extract the downloaded Zip file (SAS Software Depot_Win.zip).

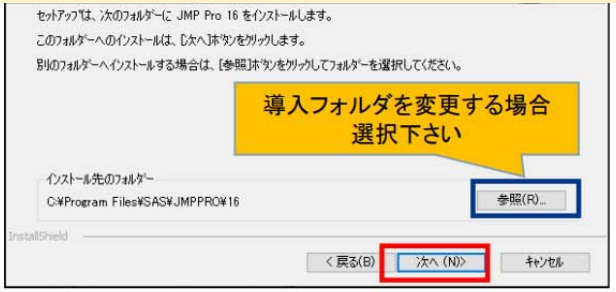


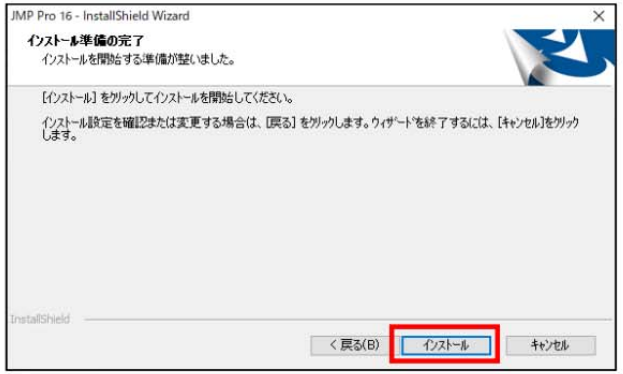
This is the program you will use.

2. SAS SoftwareDepot/JMP/JMP_Pro/16_2/Windows/jmppro_1620_win.exe を実行し、導入を進めてください。

SAS SoftwareDepot/JMP/JMP_Pro/16_2/Windows/jmppro_1620_win.exe and proceed with the installation.

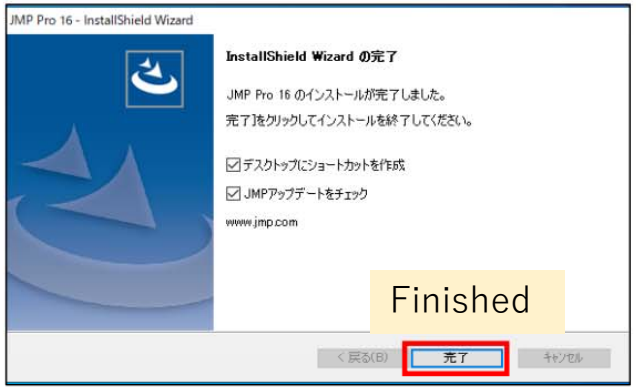






If you want to change the installation folder, click on and select folder.

2．導入（続き） Continued



Finished

Agree

Do not register now

ユーザ登録は必要ありません

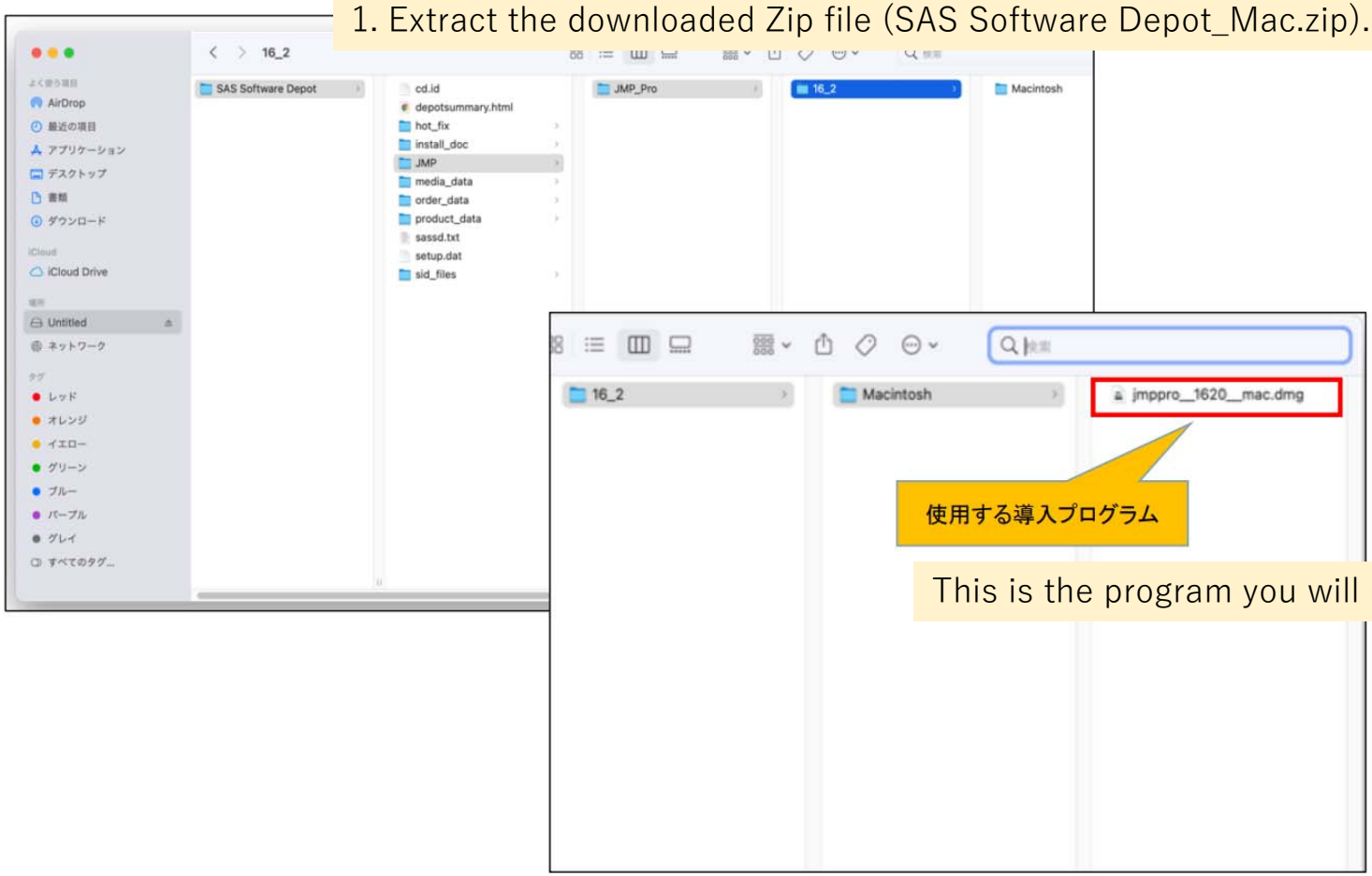
導入完了

Installation Complete

(Mac)

1. ダウンロードしたZipファイル（SAS Software Depot_Mac.zip）を解凍します。

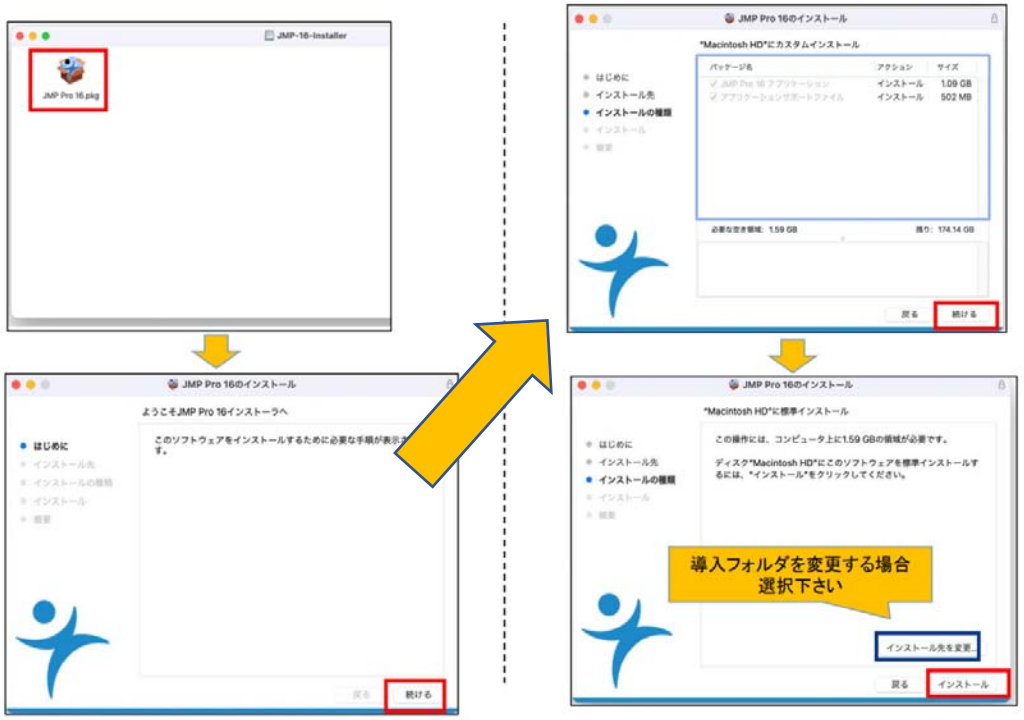
1. Extract the downloaded Zip file (SAS Software Depot_Mac.zip).



This is the program you will use.

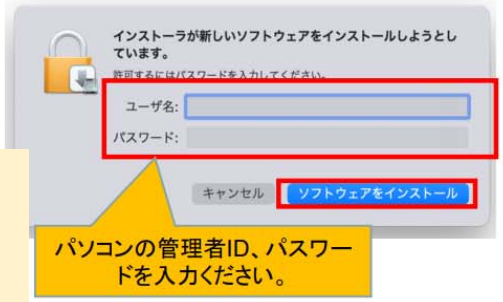
SAS SoftwareDepot/JMP/JMP_Pro/16_2/Windows/jmppro_1620_win.exe and proceed with the installation.

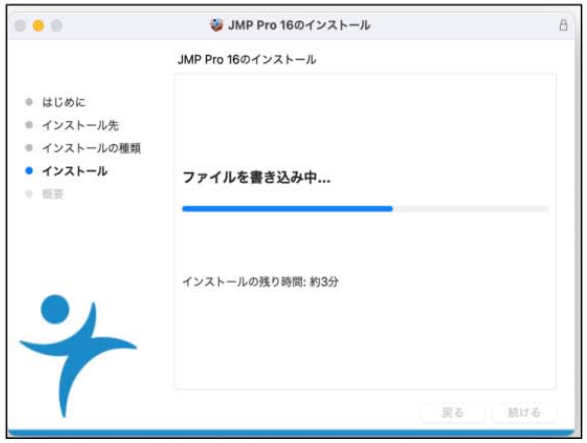
2. SAS SoftwareDepot/JMP/JMP_Pro/16_2/Macintosh/jmppro_1620_mac.dwg を実行し、導入を進めてください。

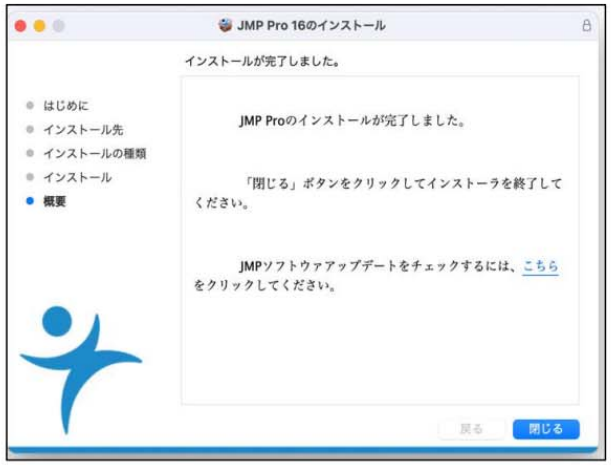


2. 導入（続き）

Enter the administrator ID and password of your computer.



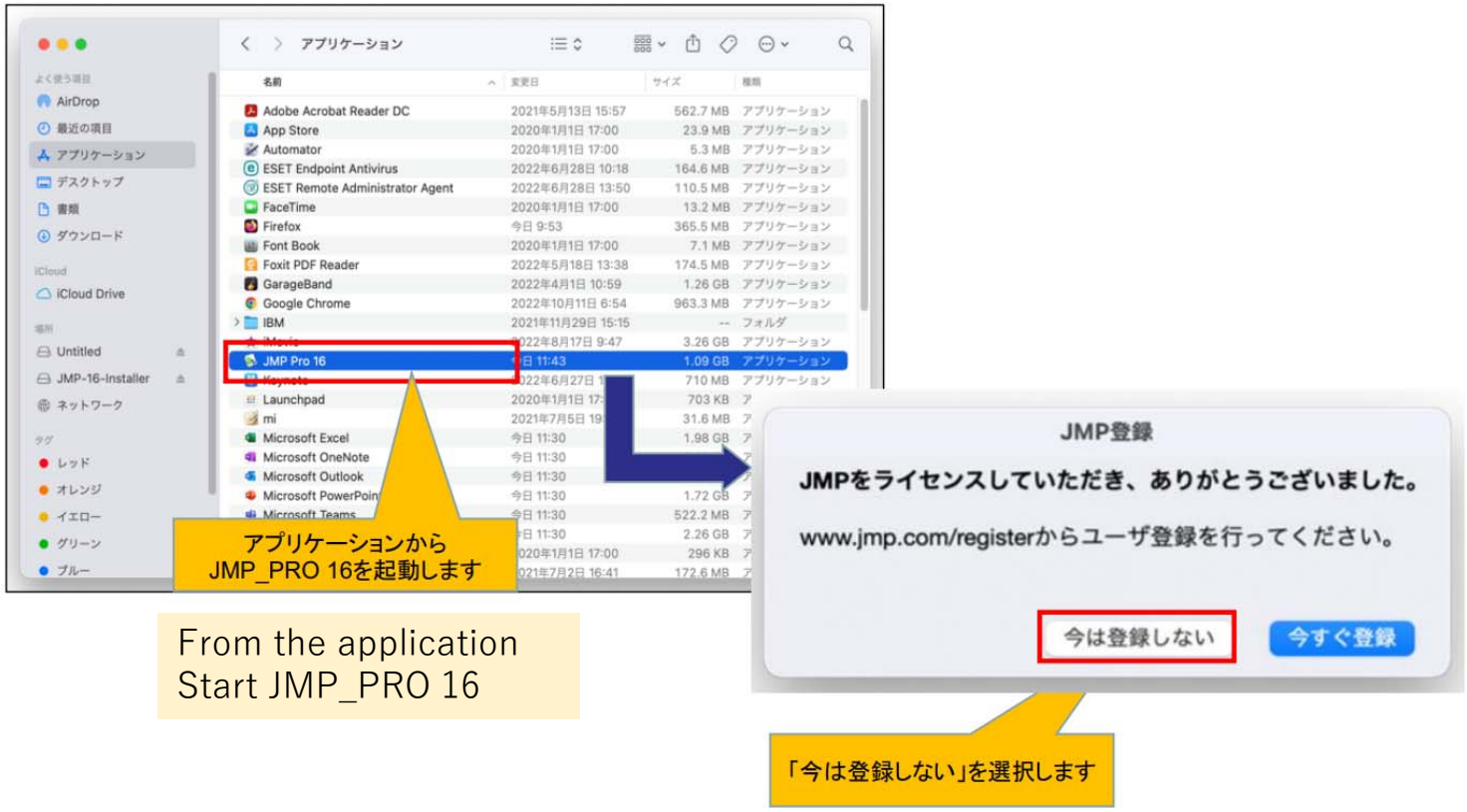




導入完了

Installation Complete

3. 導入後のJMP登録は必要ありません



アプリケーションから
JMP_PRO 16を起動します

From the application
Start JMP_PRO 16

「今は登録しない」を選択します

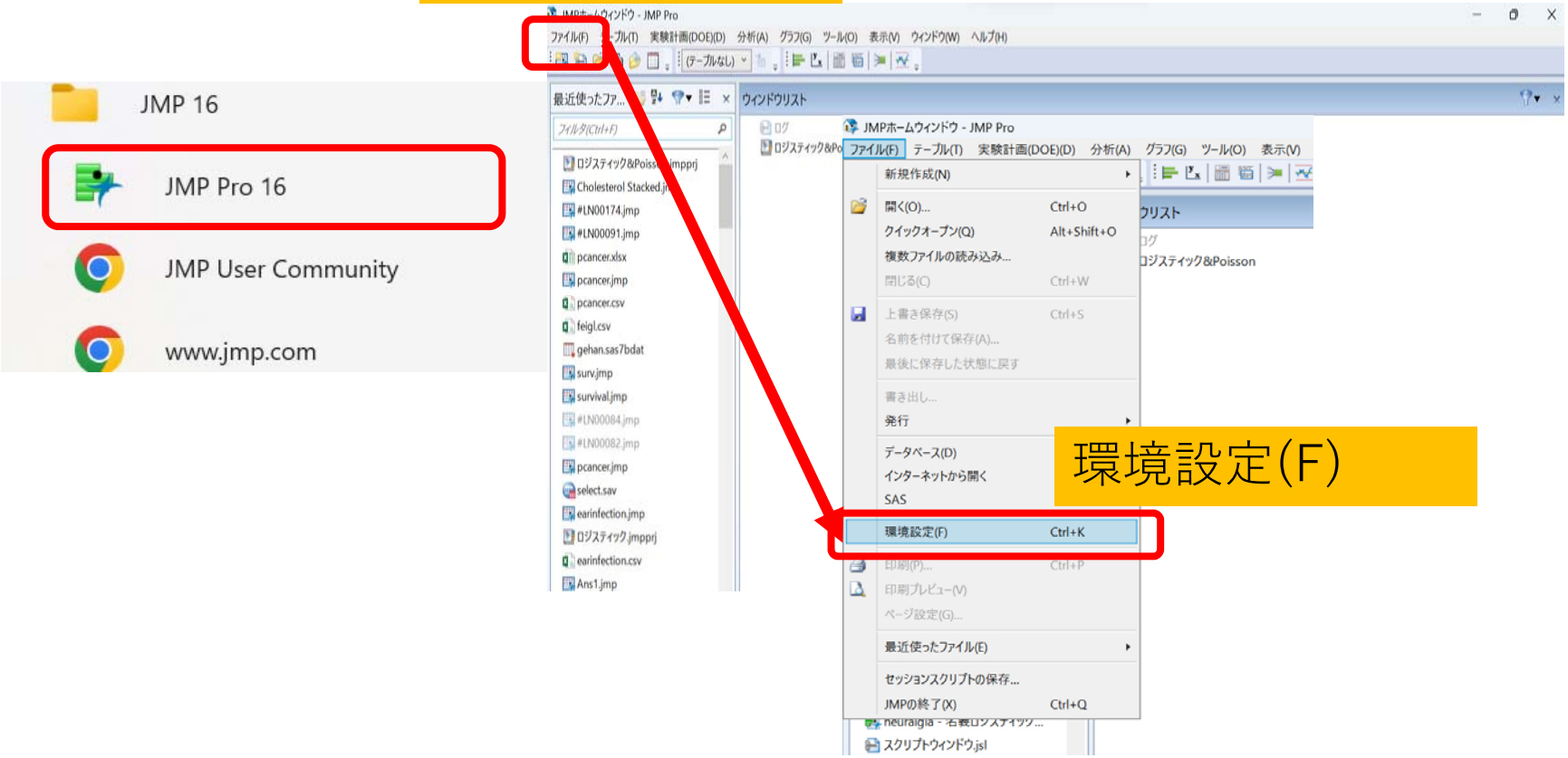
Do not register now

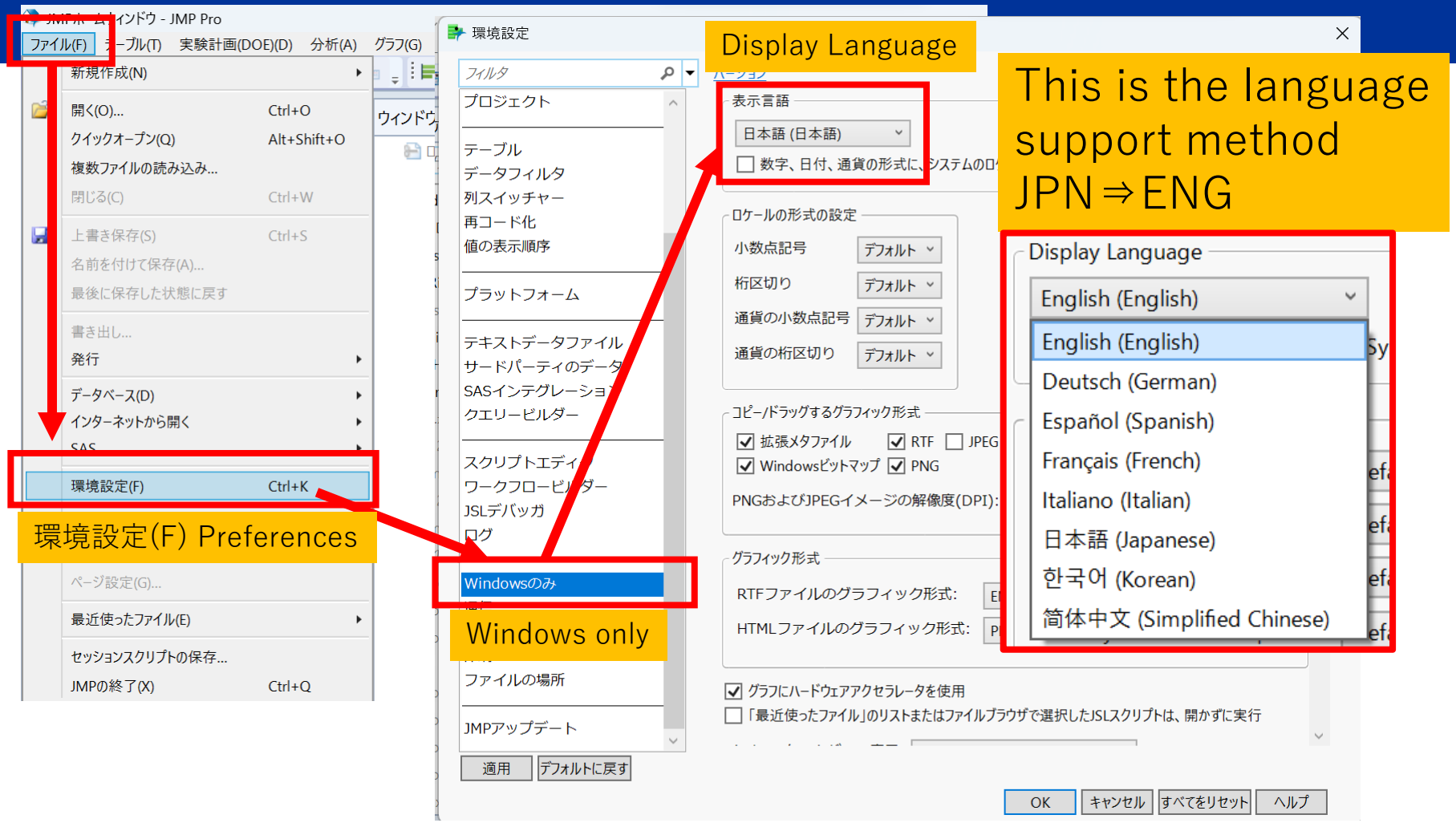
Software Language Support

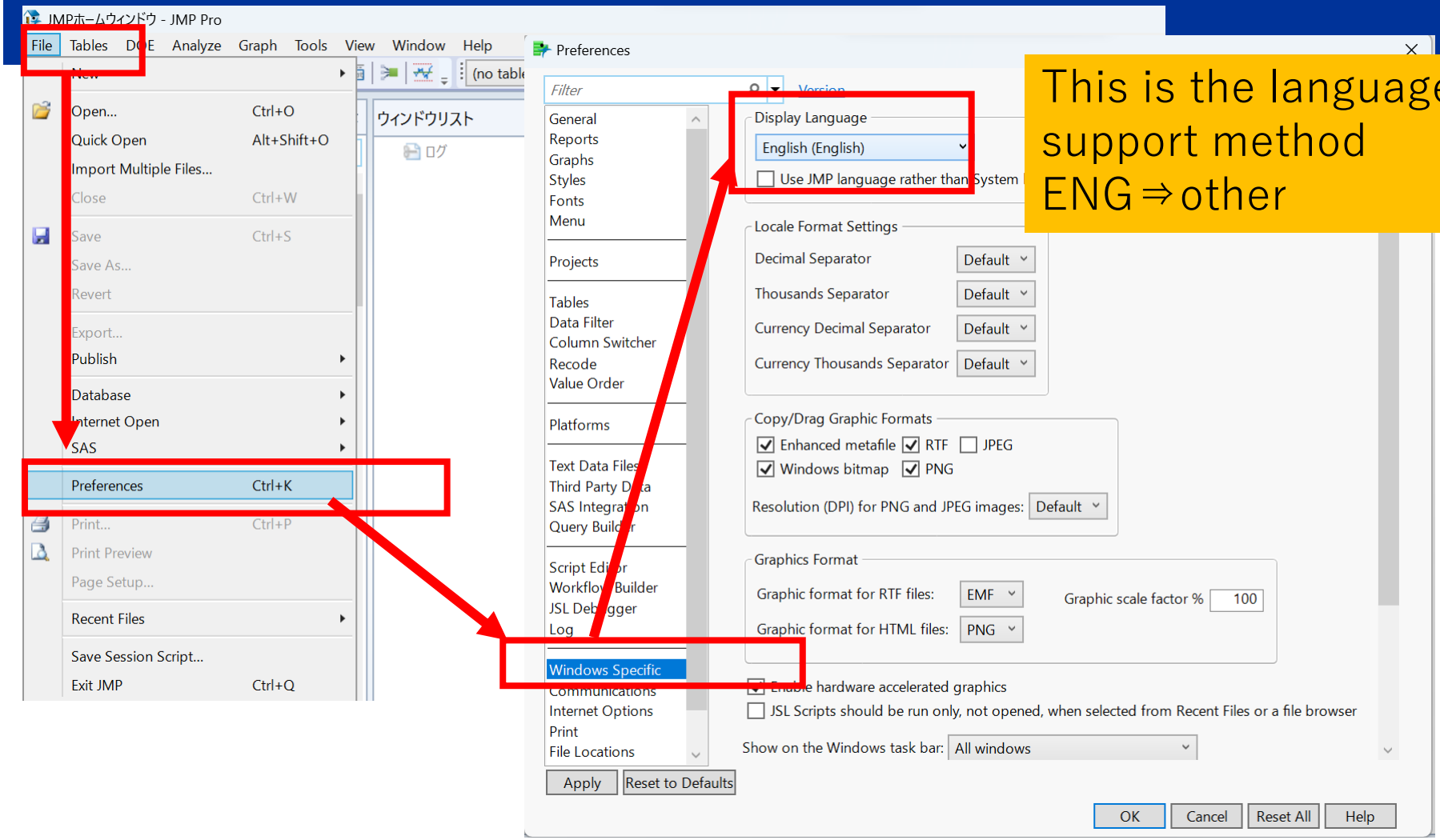
- Windows needs to be changed
- Mac supports PC languages. No need to change language settings

Windows

ファイル　File




JMP ホームウィンドウ - JMP Pro
ファイル(F)
テーブル(T)
実験計画(DOE)(D)
分析(A)
グラフ(G)
新規作成(N)
開く(O)... Ctrl+O
クイックオープン(Q) Alt+Shift+O
複数ファイルの読み込み...
閉じる(C) Ctrl+W
上書き保存(S) Ctrl+S
名前を付けて保存(A)...
最後に保存した状態に戻す
書き出し...
発行
データベース(D)
インターネットから開く
環境設定(F) Ctrl+K
ページ設定(G)...
最近使ったファイル(E)
セッションスクリプトの保存...
JMPの終了(X) Ctrl+Q
環境設定(F) Preferences
環境設定
フィルタ
プロジェクト
テーブル
データフィルタ
列スイッチャー
再コード化
値の表示順序
プラットフォーム
テキストデータファイル
サードパーティのデータ
SASインテグレーション
クエリービルダー
スクリプトエディタ
ワークフロービルダー
JSLデバッガ
ログ
Windowsのみ
Windows only
ファイルの場所
JMPアップデート
適用
デフォルトに戻す
Display Language
表示言語
日本語 (日本語)
数字、日付、通貨の形式に、システムのロケール
ロケールの形式の設定
小数点記号 デフォルト
桁区切り デフォルト
通貨の小数点記号 デフォルト
通貨の桁区切り デフォルト
コピー/ドラッグするグラフィック形式
拡張メタファイル
RTF
JPEG
Windowsビットマップ
PNG
PNGおよびJPEGイメージの解像度(DPI):
グラフィック形式
RTFファイルのグラフィック形式:
HTMLファイルのグラフィック形式:
グラフにハードウェアアクセラレータを使用
「最近使ったファイル」のリストまたはファイルブラウザで選択したJSLスクリプトは、開かずに実行
OK
キャンセル
すべてをリセット
ヘルプ
This is the language support method
JPN ⇒ ENG
Display Language
English (English)
English (English)
Deutsch (German)
Español (Spanish)
Français (French)
Italiano (Italian)
日本語 (Japanese)
한국어 (Korean)
简体中文 (Simplified Chinese)



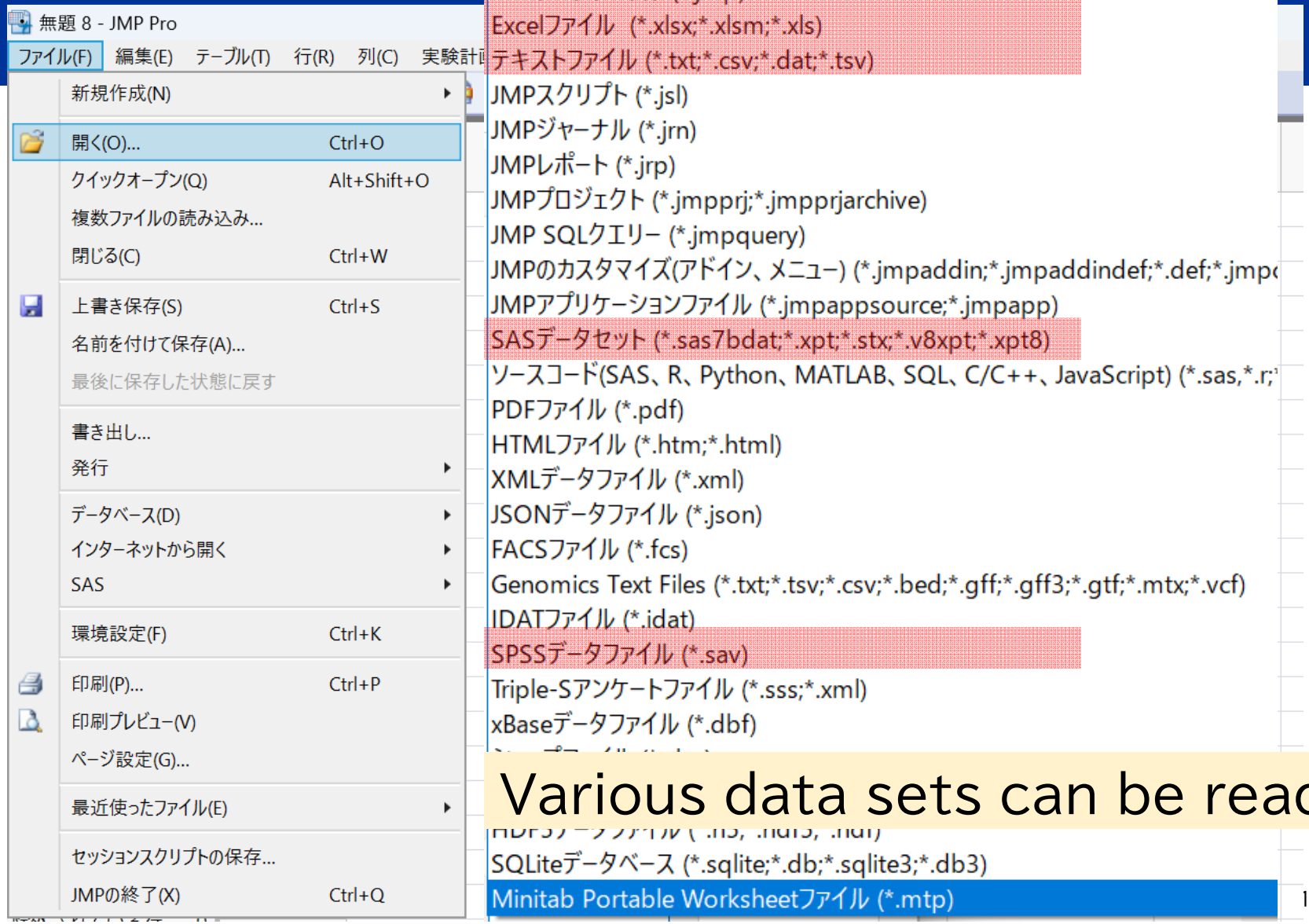
This is the language support method
ENG ⇒ other

SAS has prepared the JMP Learning Library for learning purposes, please refer to it if necessary

- JMP Manual
  - https://www.jmp.com/en_us/home.html
- JMP Learning Library
  - https://www.jmp.com/en_us/learning-library.html
- On-Demand Webinars
  - https://www.jmp.com/en_us/events/ondemand/webcasts/by-application-area/all-application-areas.html

Supplementary material for JMP(SPSS)-compatible exercises

File Import



Various data sets can be read

Import

Solubility - JMP Pro

ファイル(F) 編集(E) テーブル(T) 行(R) 列(C) 実験計画(DOE)(D) 分析(A) グラフ(G) ツール(O) 表示(V) ウィンドウ(W) ヘルプ(H)

Solubility

ロックされたファイル C:¥Progra

ノート Chemical compounds we

出展 Koehler, M. G., Grigoras, S

- 一変量の分布
- 三次元散布図
- 多変量の相関
- 散布図行列とグラフレット

列(7/0)

- 溶質
- 1-オクタノール
- エーテル
- クロロフォルム
- ベンゼン
- 四塩化炭素
- ヘキサン

	溶質	1-オクタノール	エーテル	クロロフォルム	ベンゼン	四塩化炭素	ヘキサン
1	METHANOL	-0.770	-1.150	-1.260	-1.890	-2.100	-2.800
2	ETHANOL	-0.310	-0.570	-0.850	-1.620	-1.400	-2.100
3	PROPANOL	0.250	-0.020	-0.400	-0.700	-0.820	-1.520
4	BUTANOL	0.880	0.890	0.450	-0.120	-0.400	-0.700
5	PENTANOL	1.560	1.200	1.050	0.620	0.400	-0.400
6	HEXANOL	2.030	1.800	1.690	1.300	0.990	0.460
7	HEPTANOL	2.410	2.400	2.410	1.910	1.670	1.010
8	ACETIC_ACID	-0.170	-0.340	-1.600	-2.260	-2.450	-3.060
9	PROPIONICACID	0.330	0.270	-0.960	-1.350	-1.600	-2.140
10	BUTYRICACID	0.790	0.610	-0.270	-0.960	-0.970	-1.760
11	HEXANOICACID	1.920	1.950	1.150	0.300	0.570	-0.460
12	PENTANOICACID	1.390	1.000	0.280	-0.100	-0.420	-1.000
13	TRICHLOROAC…	1.330	1.210	-0.690	-1.300	-1.660	-2.630
14	DICHLOROACE…	0.920	1.310	-0.890	-1.400	-2.310	-2.720
15	CHLOROACETI…	0.220	0.370	-1.920	-1.600	-2.560	-3.140
16	METHYLACETA…	0.180	0.430	1.160	0.530	0.320	-0.260
17	ETHYLACETATE	0.730	0.930	1.800	1.010	0.950	0.290
18	ACETONE	-0.240	-0.210	0.240	-0.050	-0.300	-0.910
19	ETHYLAMINE	-0.300	-1.180	-0.350	-1.300	-1.270	-1.770
20	PROPYLAMINE	0.280	-0.540	0.260	-0.520	-0.590	-1.000
21	TRIMETHYLAM…	0.270	-0.260	0.540	-0.290	-0.090	-0.480
22	N-BUTYLAMINE	0.740	0.110	0.560	-0.080	-0.040	-0.620
23	DIETHYLAMINE	0.570	-0.070	0.810	-0.050	0.030	-0.480
24	PYRIDINE	0.650	0.080	1.430	0.410	0.230	-0.210

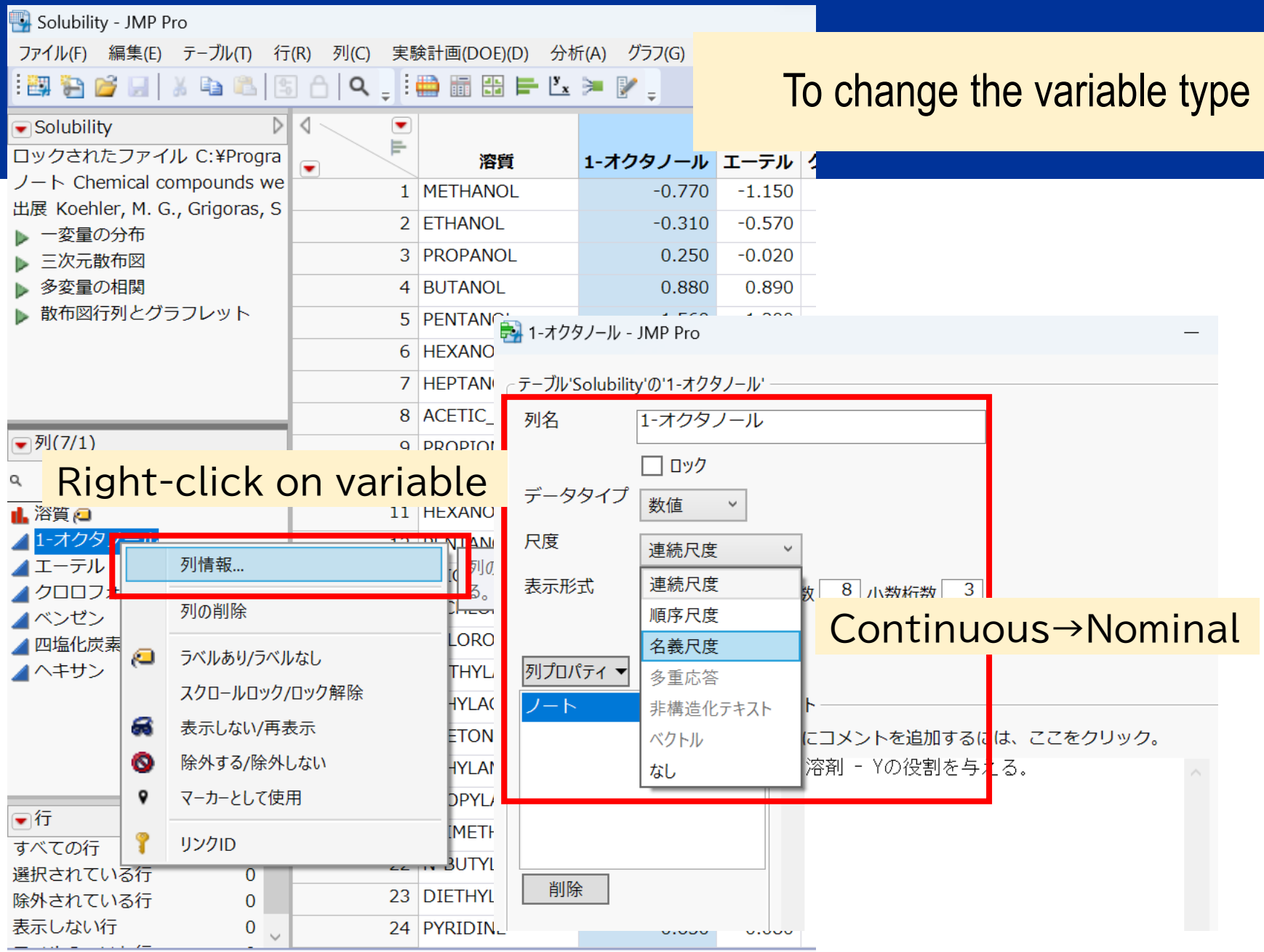
List of variable

- Continuous
- Ordinal
- Nominal

Number of data read

すべての行	72
選択されている行	0
除外されている行	0
表示しない行	0

To change the variable type



The distributions of numeric variables

Using JMP, try to output the following indicators

Position: mean, median, percentile

Variance: range, interquartile range, CV

Distribution: skewness, angularity, normality

Association between two variables Correlation

Example

- Health check-up data　N=421
  - Age 40-60yrs.
  - Variable
    - SEX, AGE, Systolic Blood Pressure (SBP), Diastolic blood pressure (DBP)
    - Urinary Sodium (SALT)
  - Dataset name=health

Check the shape, location, and distribution indices of the distribution

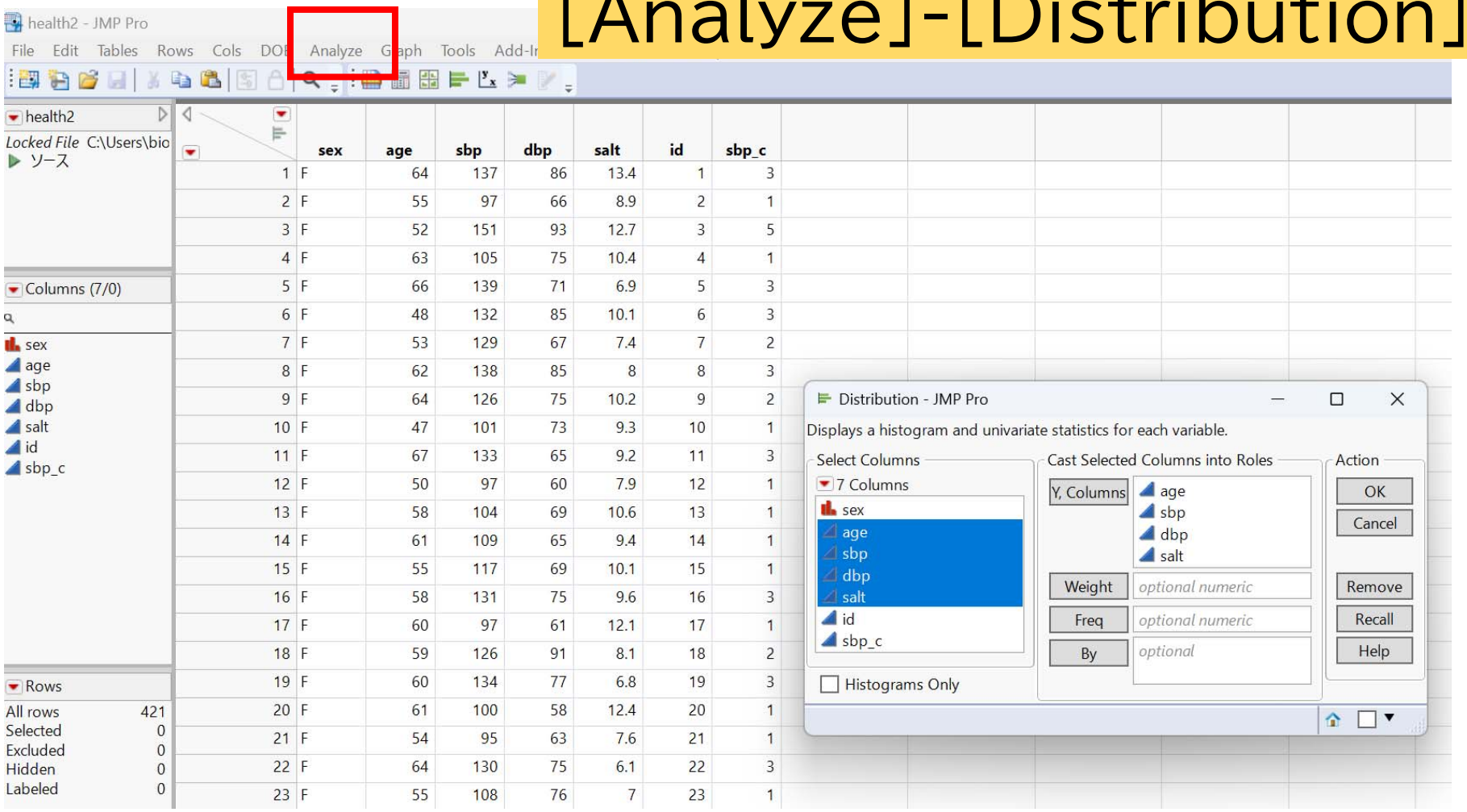
1. Location
   1)Mean
   2)Median
   3)Mode
   4)Percentile

2. Variability dispersion
   1) Variance
   2) SD
   3) Range
   4) Inter-quartile range

3. Shape
   1) Skewness 歪度
   2) Kurtosis 尖度

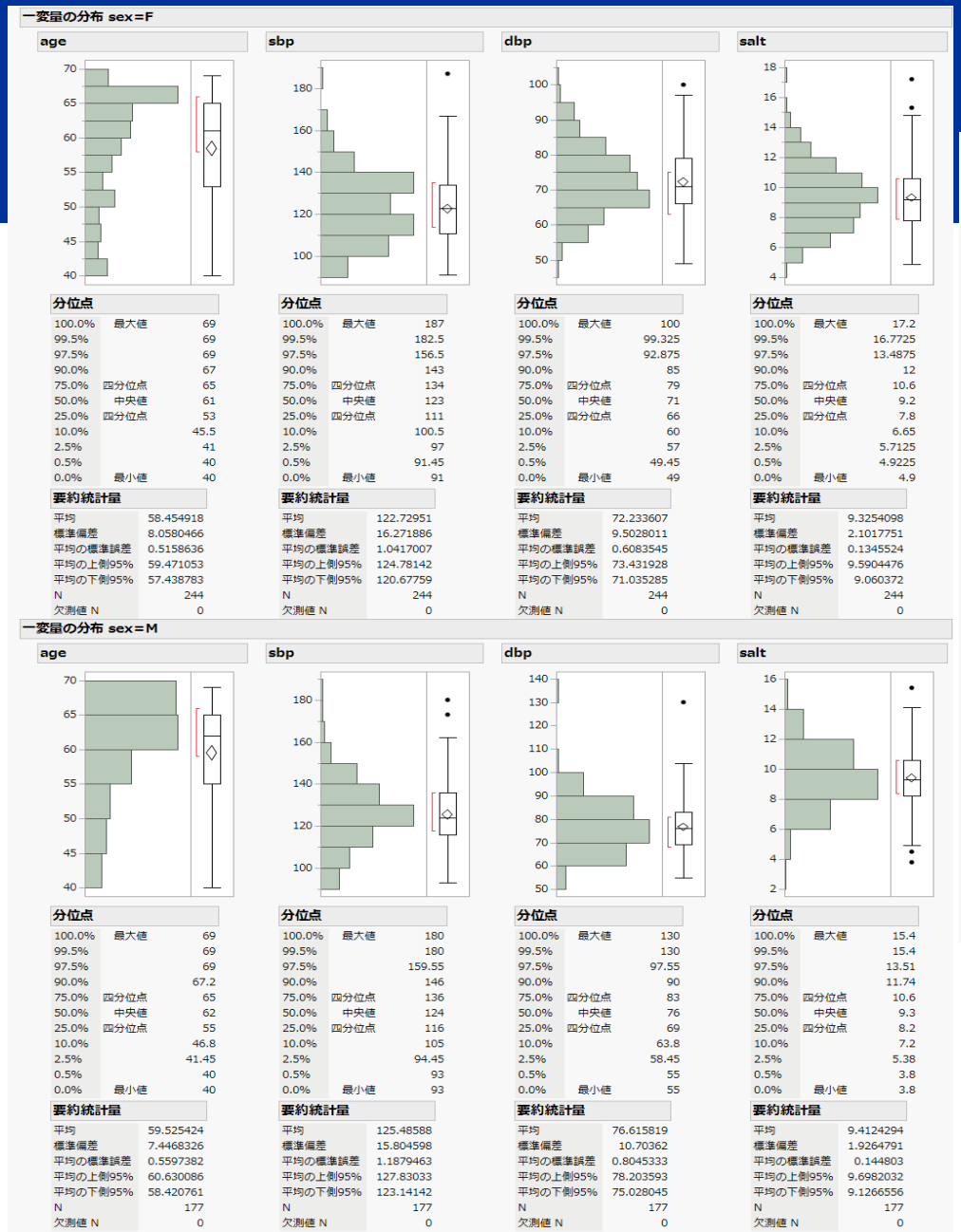
Distribution

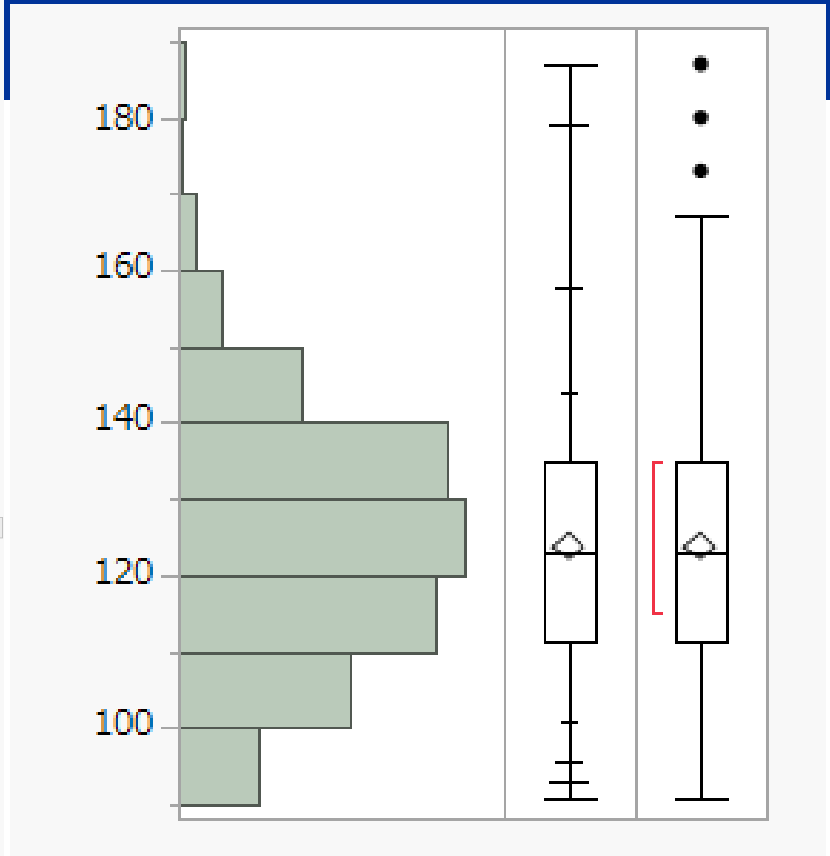
[Analyze]-[Distribution]



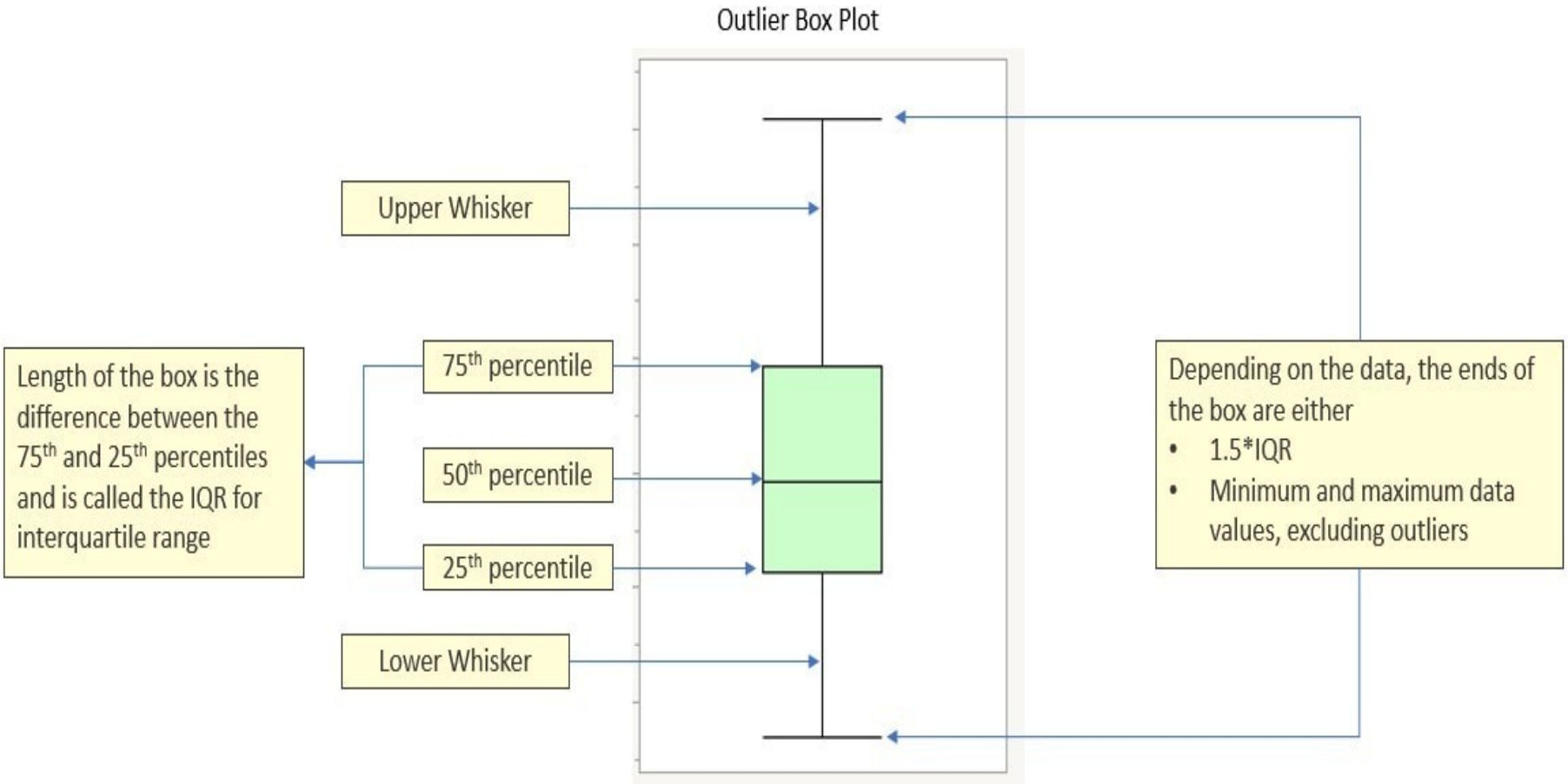
	sex	age	sbp	dbp	salt	id	sbp_c
1	F	64	137	86	13.4	1	3
2	F	55	97	66	8.9	2	1
3	F	52	151	93	12.7	3	5
4	F	63	105	75	10.4	4	1
5	F	66	139	71	6.9	5	3
6	F	48	132	85	10.1	6	3
7	F	53	129	67	7.4	7	2
8	F	62	138	85	8	8	3
9	F	64	126	75	10.2	9	2
10	F	47	101	73	9.3	10	1
11	F	67	133	65	9.2	11	3
12	F	50	97	60	7.9	12	1
13	F	58	104	69	10.6	13	1
14	F	61	109	65	9.4	14	1
15	F	55	117	69	10.1	15	1
16	F	58	131	75	9.6	16	3
17	F	60	97	61	12.1	17	1
18	F	59	126	91	8.1	18	2
19	F	60	134	77	6.8	19	3
20	F	61	100	58	12.4	20	1
21	F	54	95	63	7.6	21	1
22	F	64	130	75	6.1	22	3
23	F	55	108	76	7	23	1

[Distribution]

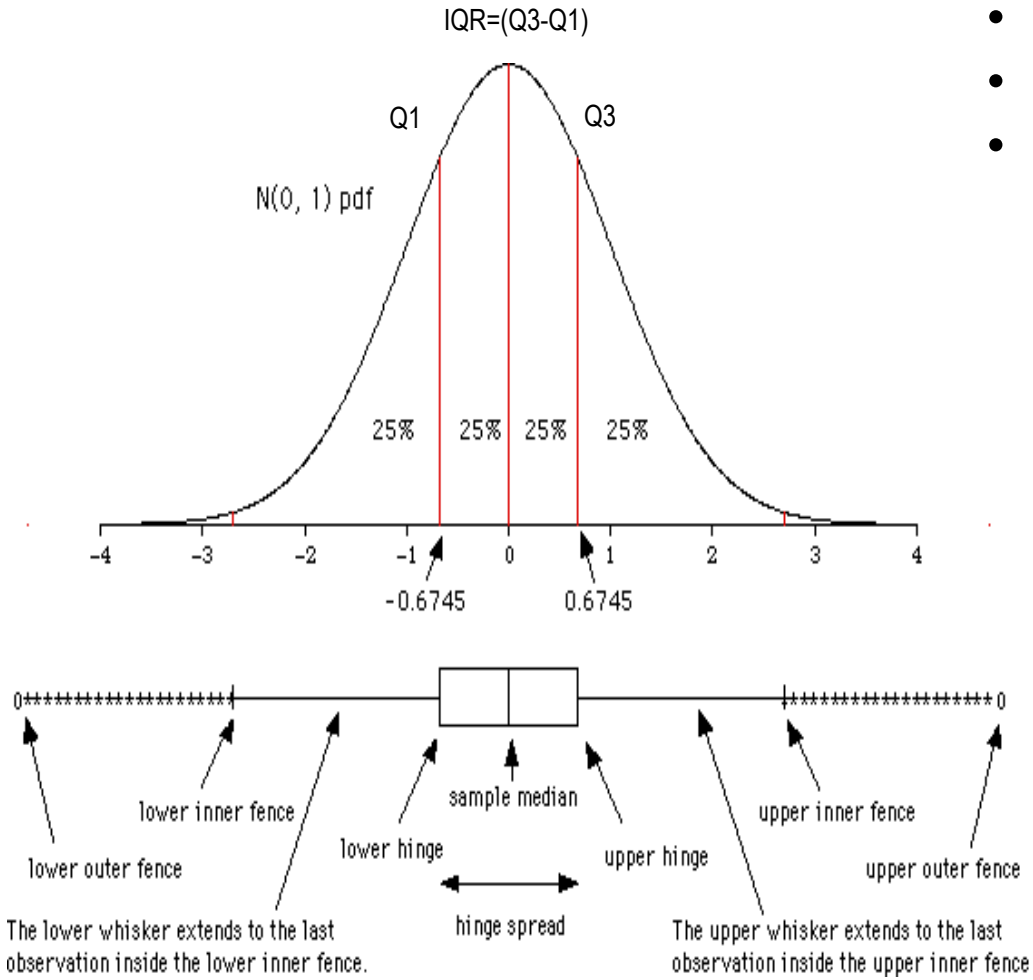




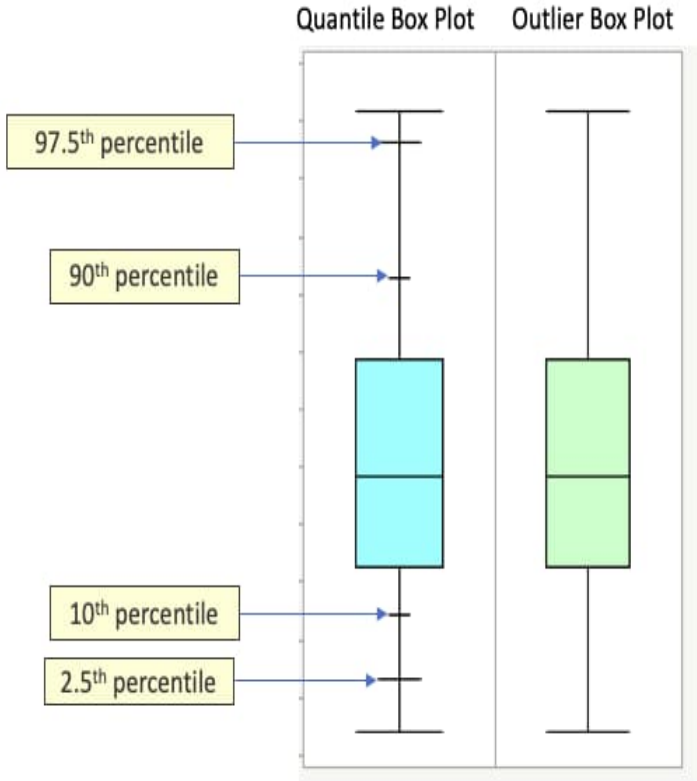
Boxplot


Outlier Box Plot
Upper Whisker
75th percentile
50th percentile
25th percentile
Lower Whisker
Length of the box is the difference between the 75th and 25th percentiles and is called the IQR for interquartile range
Depending on the data, the ends of the box are either
• 1.5*IQR
• Minimum and maximum data values, excluding outliers

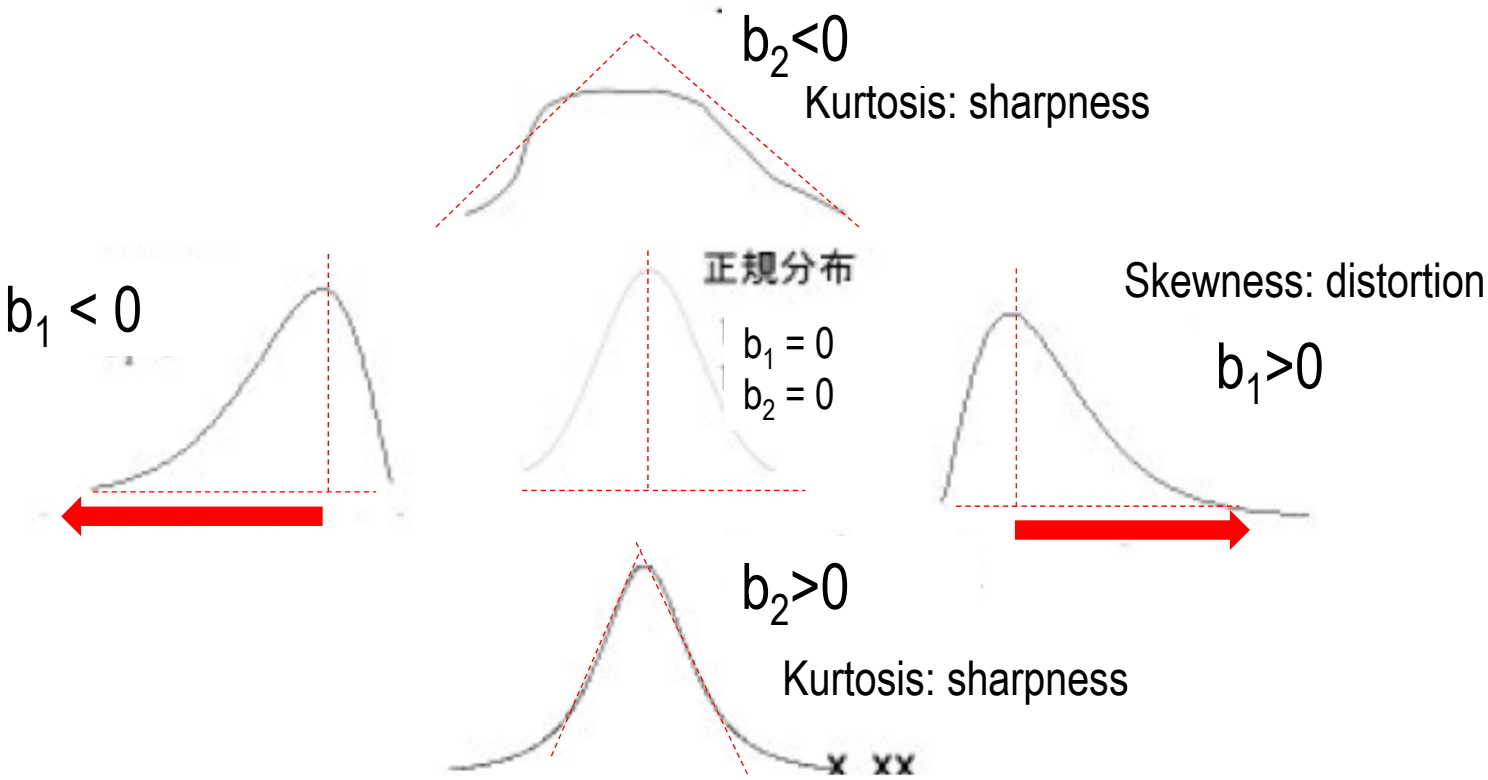
Boxplot



- Median is center of box
- 25th percentile is bottom of box
- 75th percentile is top of box



Shape
Skewness(b_1), Kurtosis(b_2)

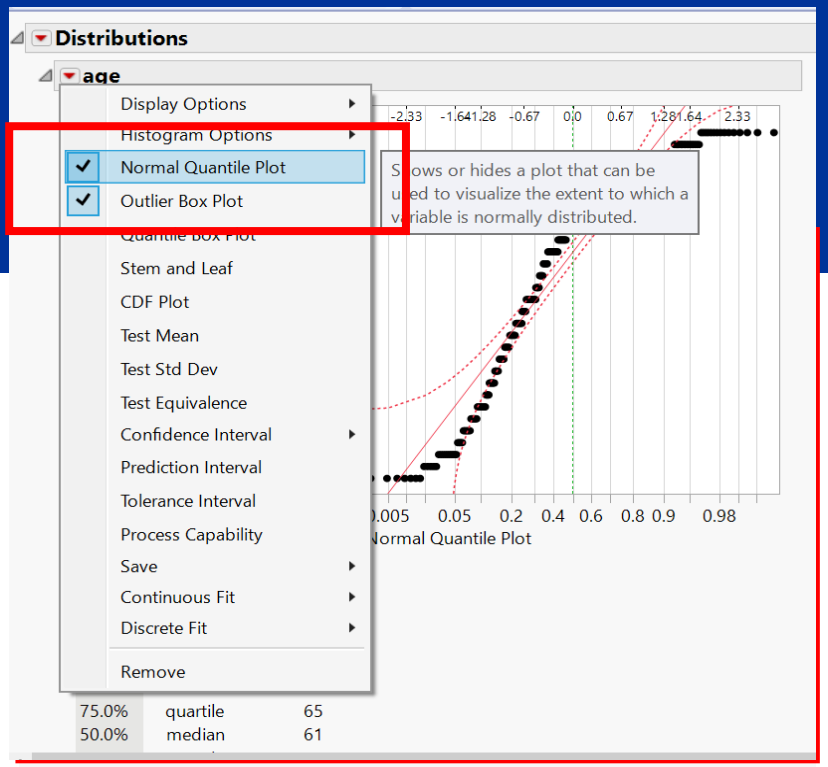


Skewness: distortion

Kurtosis: sharpness

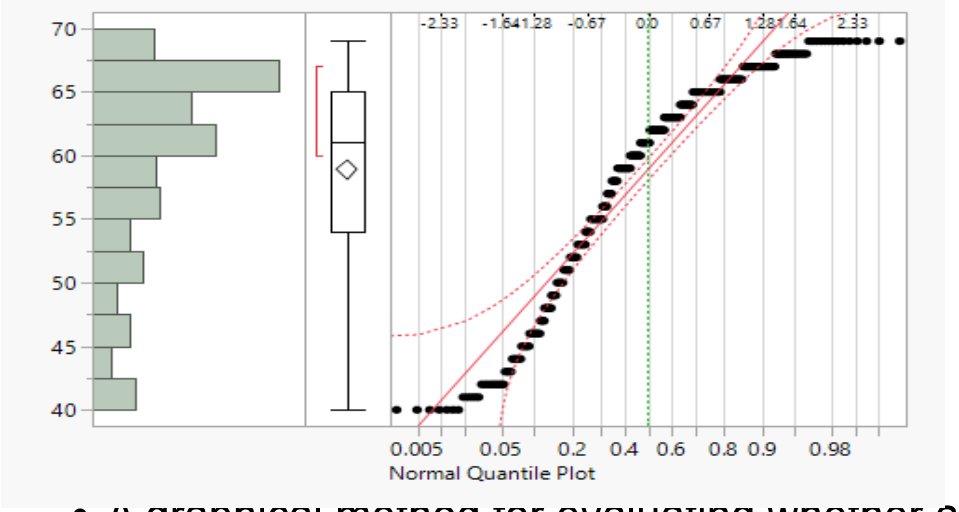
$$b_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^3/(n-1)}{SD^3}, \quad b_2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^4/(n-1)}{SD^4} - 3$$

Skewness(b_1), Kurtosis(b_2)
Normal probability plot (Q-Qplot)



分位点		
100.0%	最大値	187
99.5%		179.23
97.5%		157.9
90.0%		144
75.0%	四分位点	135
50.0%	中央値	123
25.0%	四分位点	111.5
10.0%		101
2.5%		95.55
0.5%		93
0.0%	最小値	91

要約統計量	
平均	123.88836
標準偏差	16.115783
平均の標準誤差	0.7854352
平均の上側95%	125.43223
平均の下側95%	122.34449
N	421
歪度	0.3612619
尖度	0.3200493

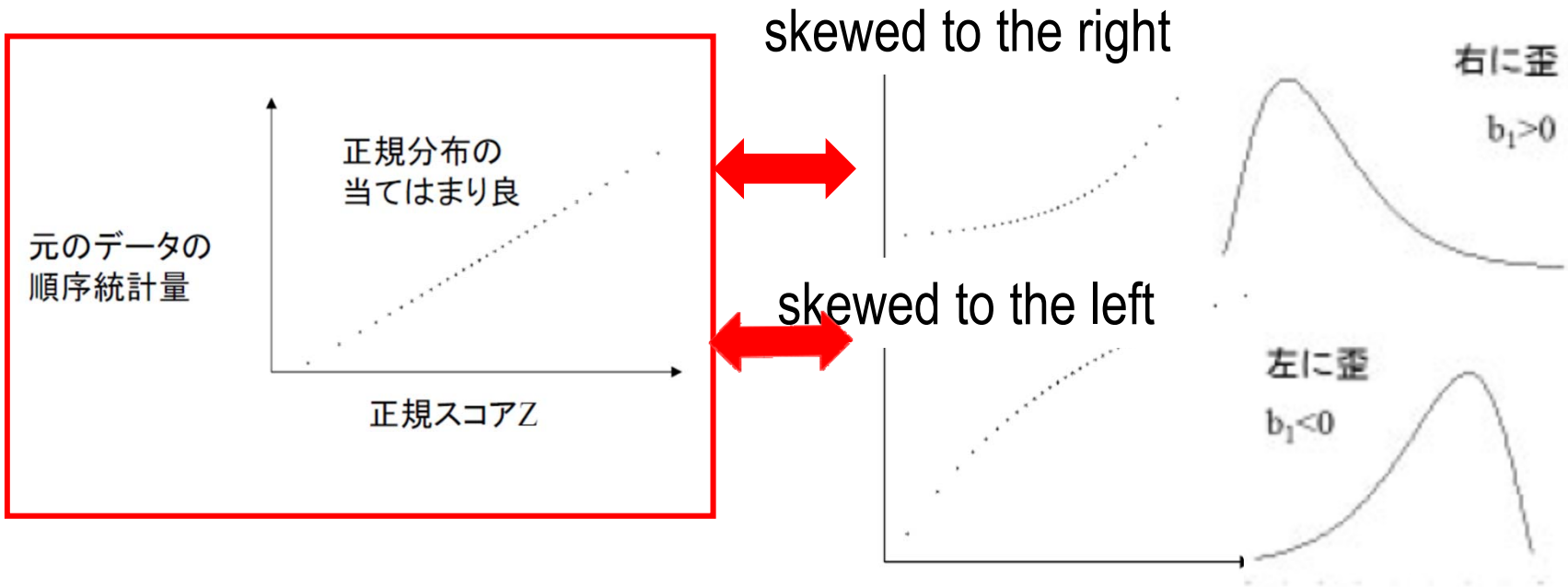


- A graphical method for evaluating whether a data set is approximately normally distributed
- The normal probability plot is constructed as follows
  - Vertical axis: ordered response values
  - Horizontal axis: median value of the normal ordered statistic

Normal probability plot

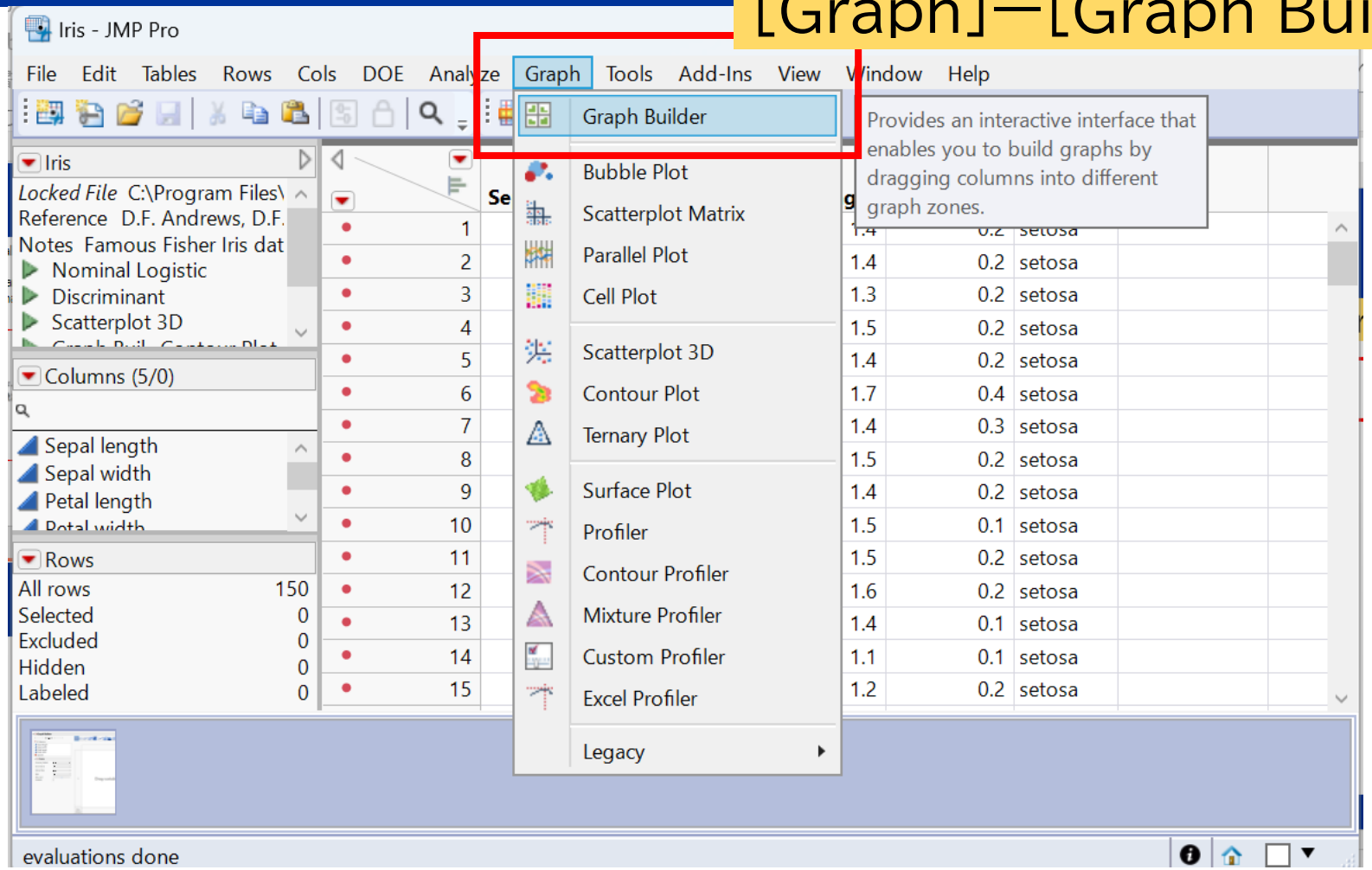
1.Normal probability plot to determine if the data are normally distributed

2.What are the deviations from normality (data skewed, tails shorter than expected, tails longer than expected)?

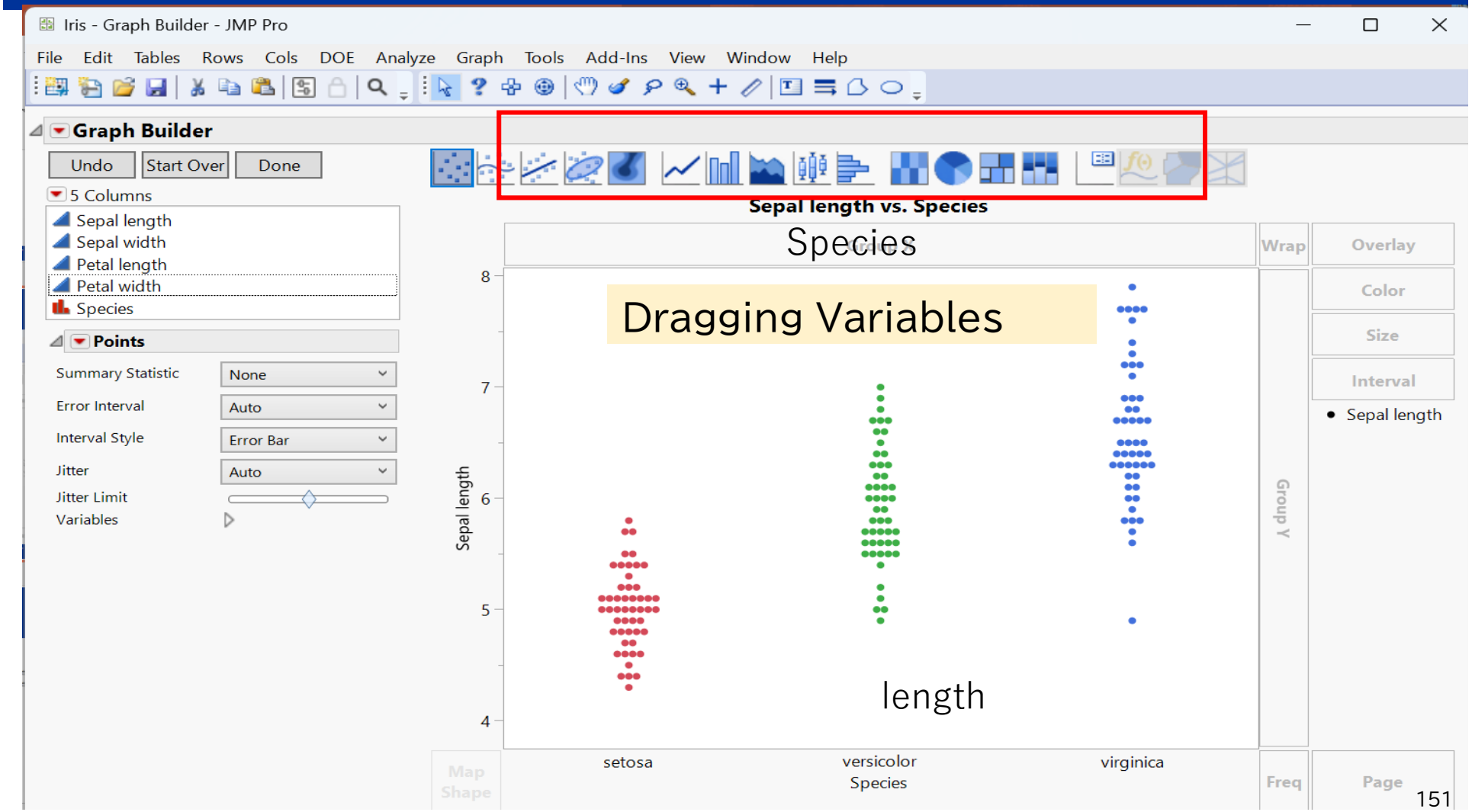


Graph Builder [Iris.jmp]

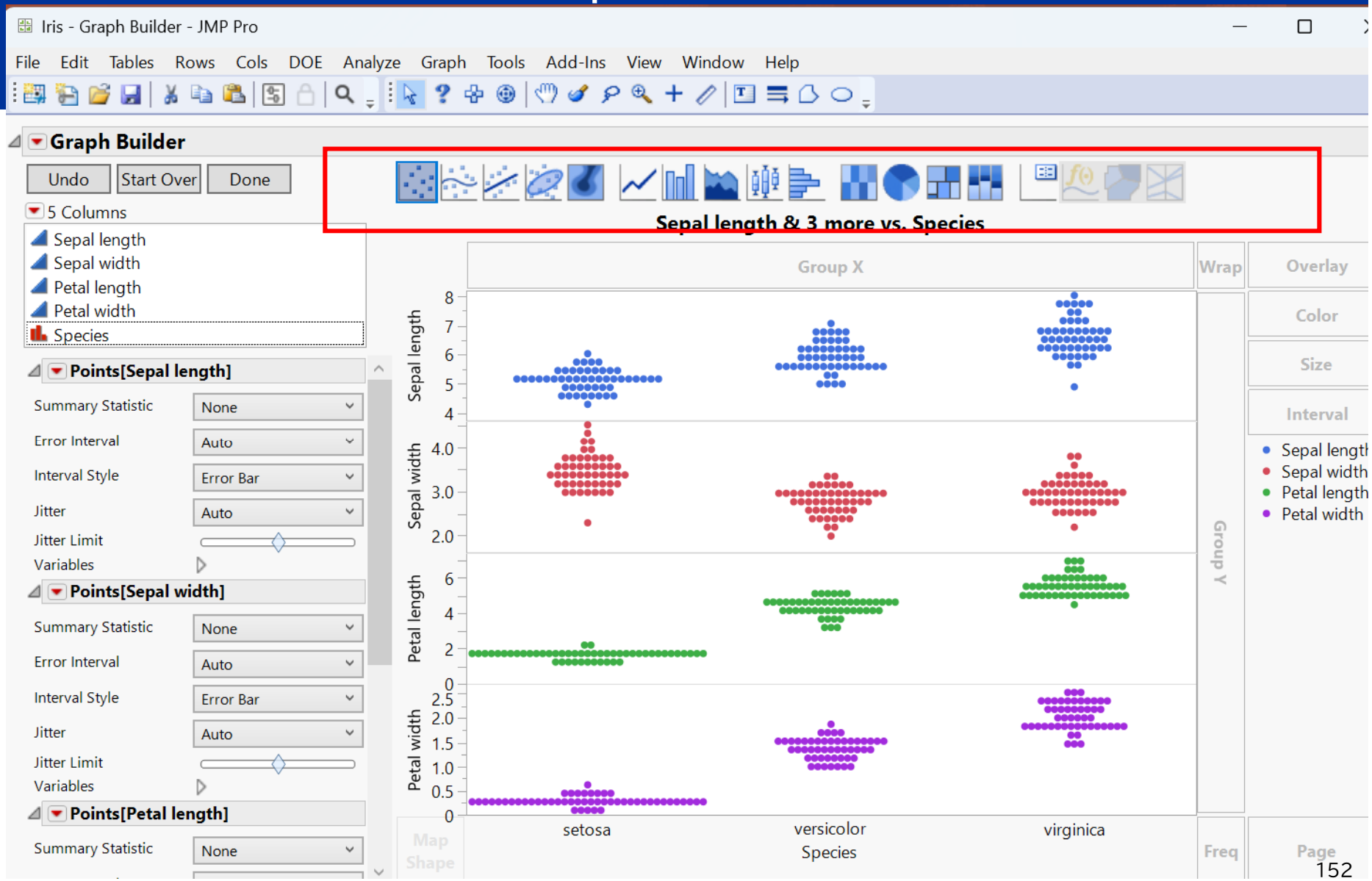
[Graph]—[Graph Builder]



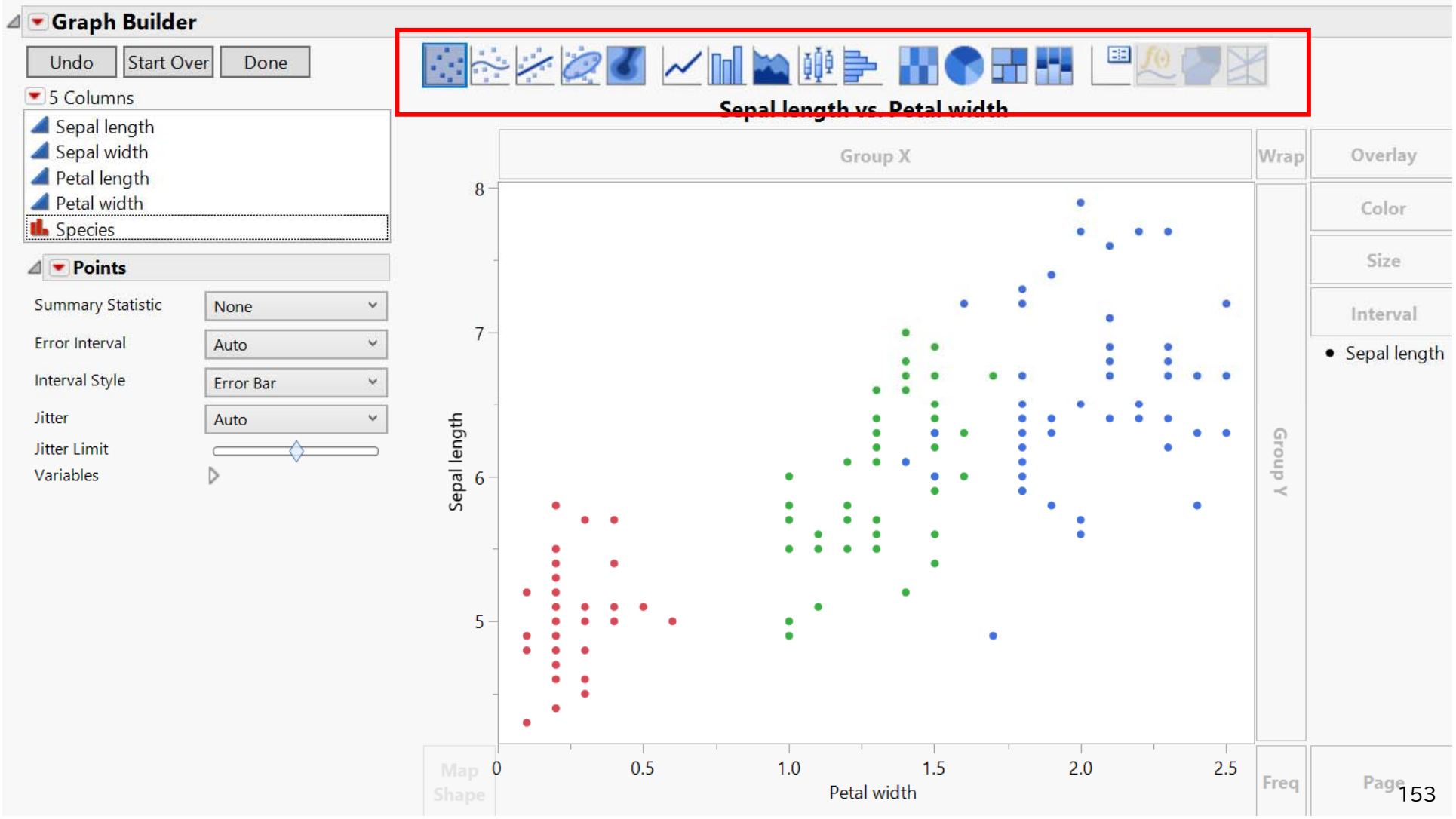
Graph Builder



Graph Builder



Graph Builder



Today's overview

1. Usage of Statistical Software (package)
   1) Brief Introduction to SAS
   2) Reading Data

2. Description of Data
   1) Location: mean, median, percentile
   2) Dispersion: range, inter-quartile range, CV
   3) Distribution: skewness, kurtosis, Normality
   4) Association between two variables: Correlation

Today's overview

1. Usage of Statistical Software (package)
   1) Brief Introduction to SAS
   2) Reading Data

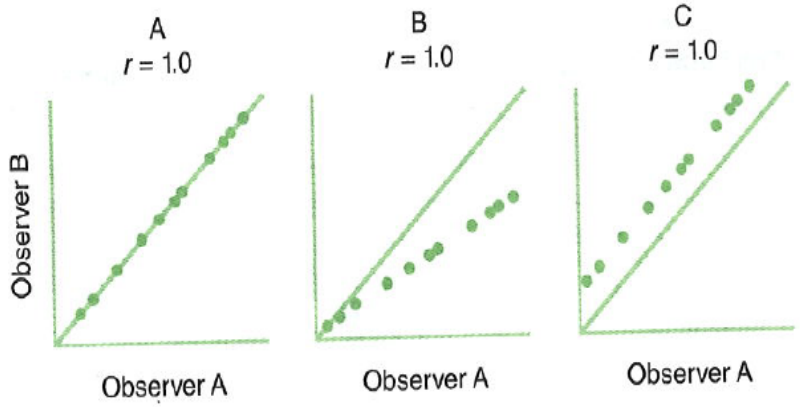
2. Description of Data
   1) Location: mean, median, percentile
   2) Dispersion: range, inter-quartile range, CV
   3) Distribution: skewness, kurtosis, Normality
   4) Association between two variables: Correlation

Correlation

- Correlation is a bivariate analysis that measures the strength of association between two variables and the linear association

  → The correlation coefficients of A, B and C are all 1.0

  → It is not necessarily a good measure of agreement.



- The value of the correlation coefficient varies between +1 and -1.
- As the correlation coefficient value goes towards 0, the relationship between the two variables will be weaker.

Pearson product-moment correlation coefficient (usually denoted by "r")

$$\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}} = Cov(x,y)/\sqrt{V(x)V(y)}$$

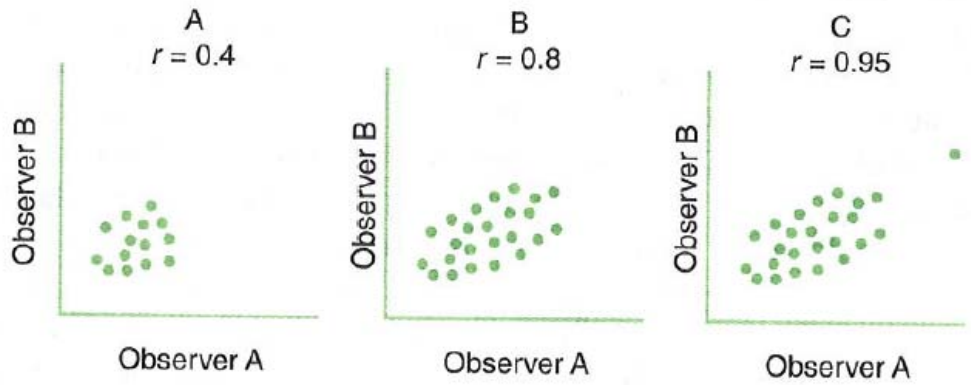


FIGURE 8-11 Hypothetical examples illustrating the high sensitivity of Pearson's correlation coefficient to the range of values and to the presence of outliers (extreme values).

Spearman correlation coefficient or rank correlation coefficient (denoted by "ρ")

- Non-parametric test that measures the strength of dependence between two variables.

$$\rho = 1 - \frac{6}{n^3 - n}\sum_{i=1}^{n}(R_i - R_i')^2$$

ρ= Spearman rank correlation
Ri,Ri'= rank
Ri-Ri'= the difference between the ranks of corresponding variables
n= number of observations

- (x,y) translate as rank order data (1 to n) and calculate Pearson correlation coefficient. If there are tie data, these are given as mean orders.

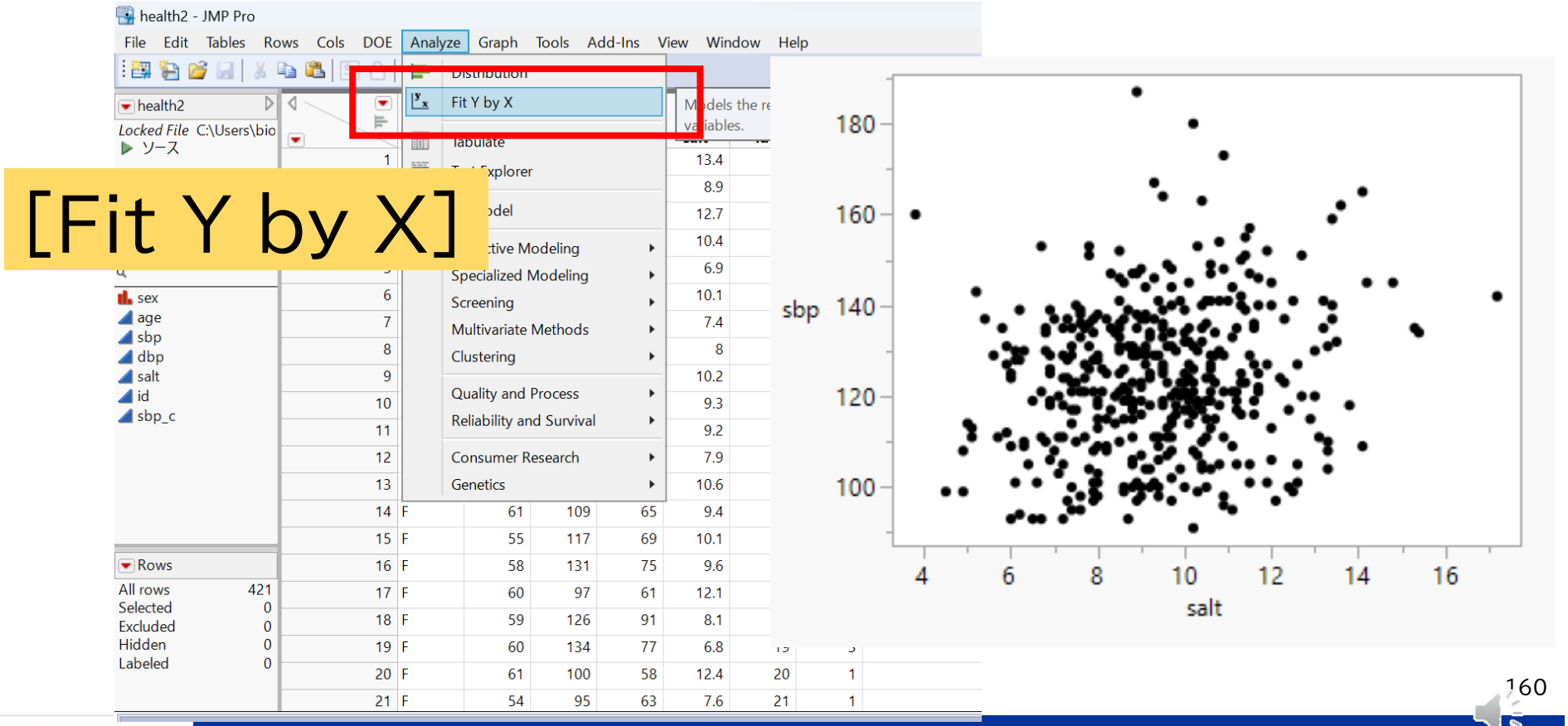
A note on correlation coefficients

- A large correlation coefficient does not indicate the magnitude of the causal relationship
  Be careful about making judgments based on tests and significance probabilities
  It only indicates that "correlation is not zero".
  (It is meaningless to compare the correlation coefficients by significance probability when n is different.)

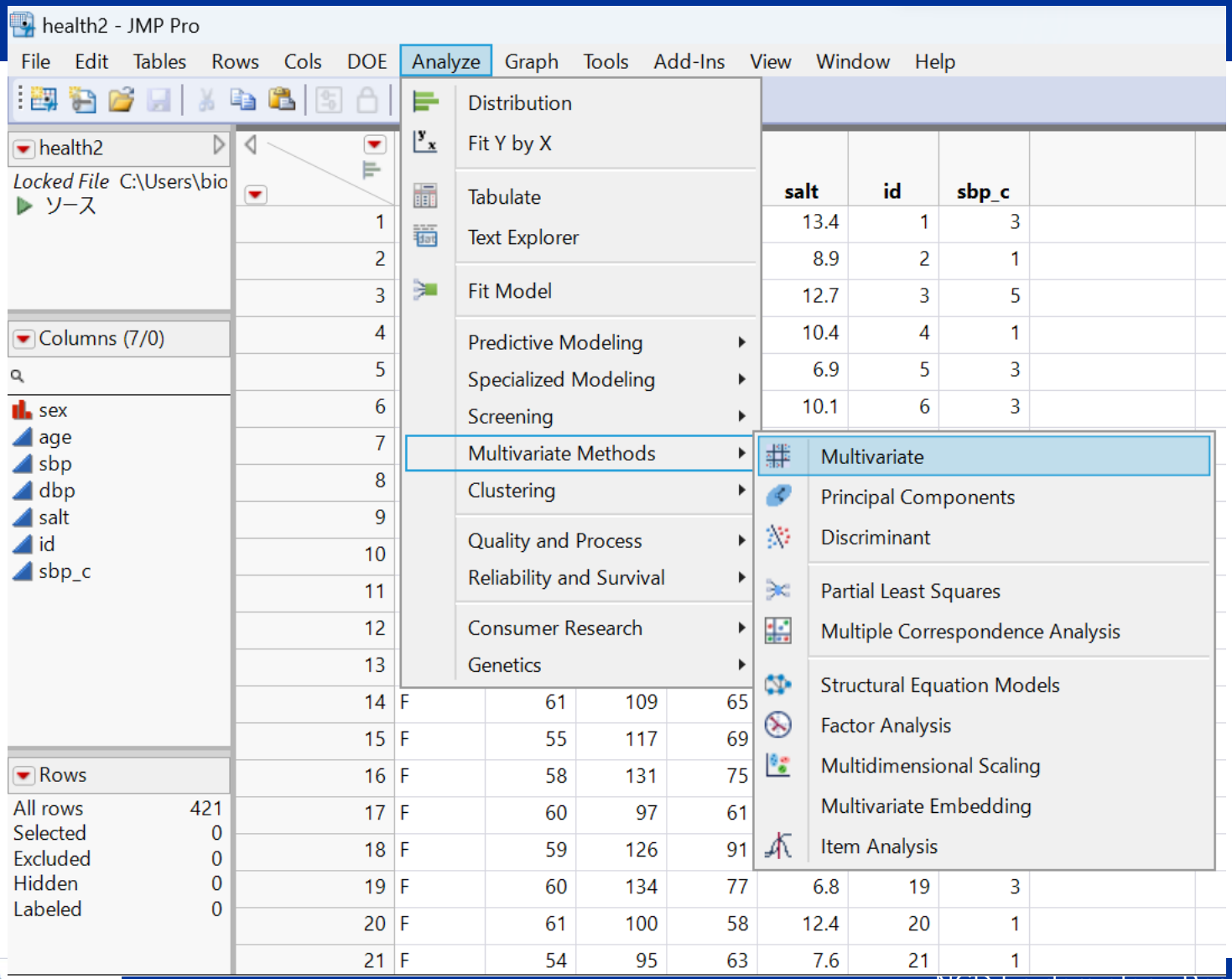
- Does it follow a normal distribution?
  For relationships such as floor and ceiling effects, use the "rank correlation coefficient".
- Beware of outliers
- Beware of mixed distributions

Correlation

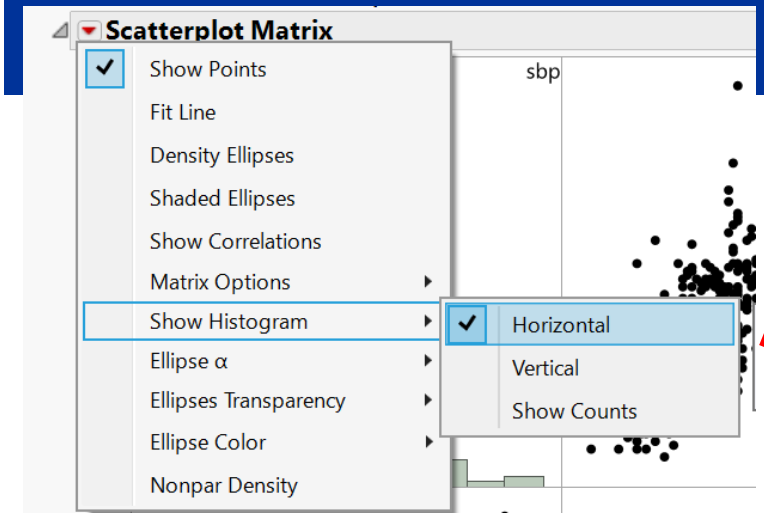
Correlation coefficient between blood pressure and salt intake using the health2 dataset

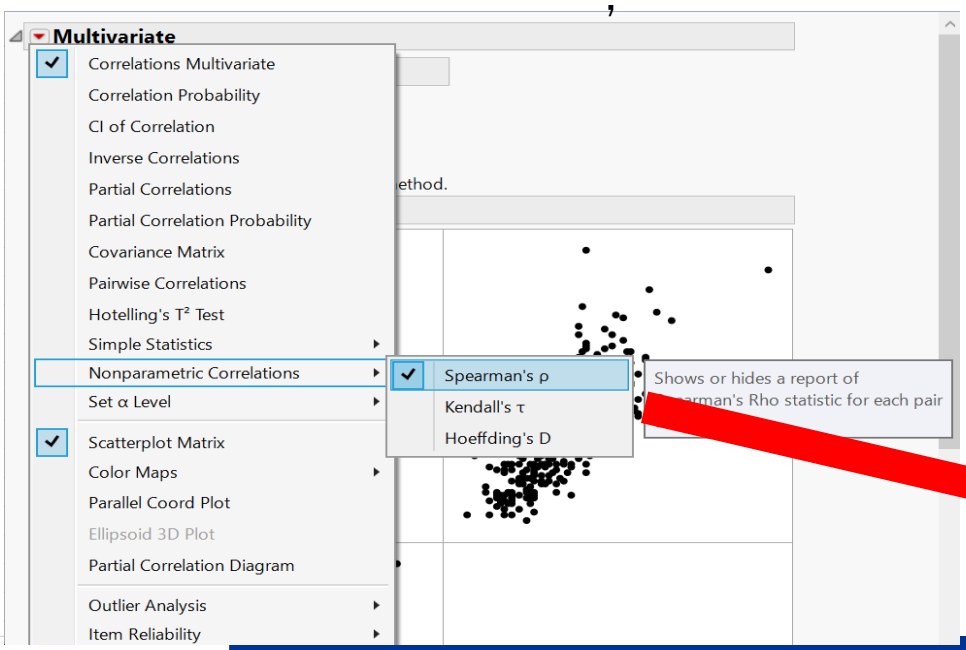


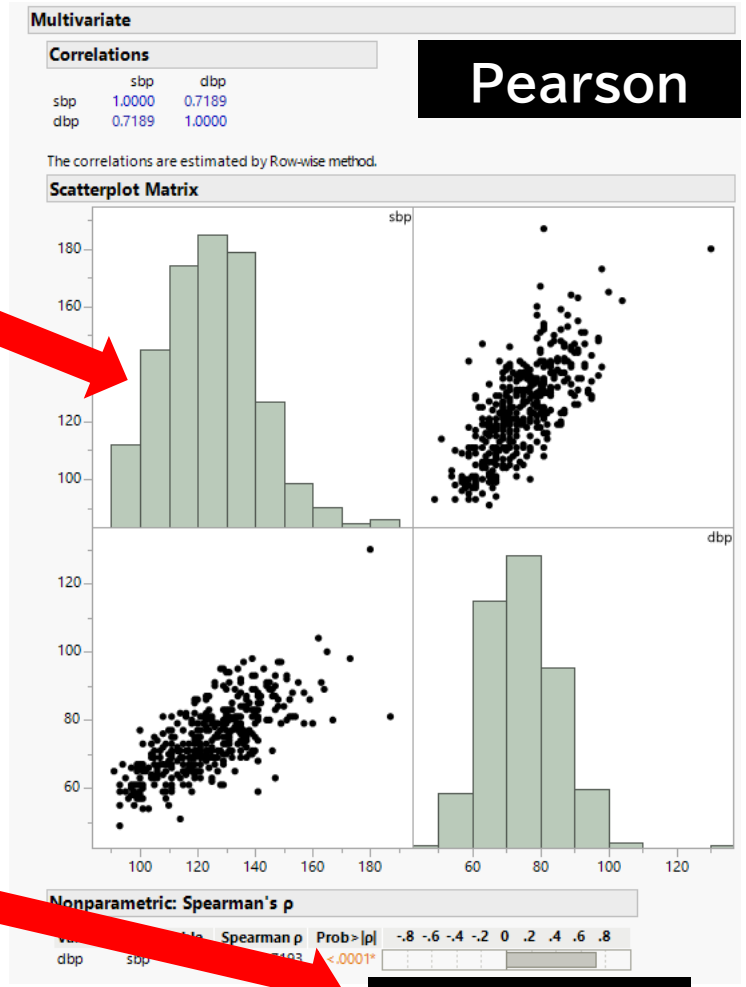
[Analyze]—[Multivariate]—[Multivariate]



Correlation（JMP）



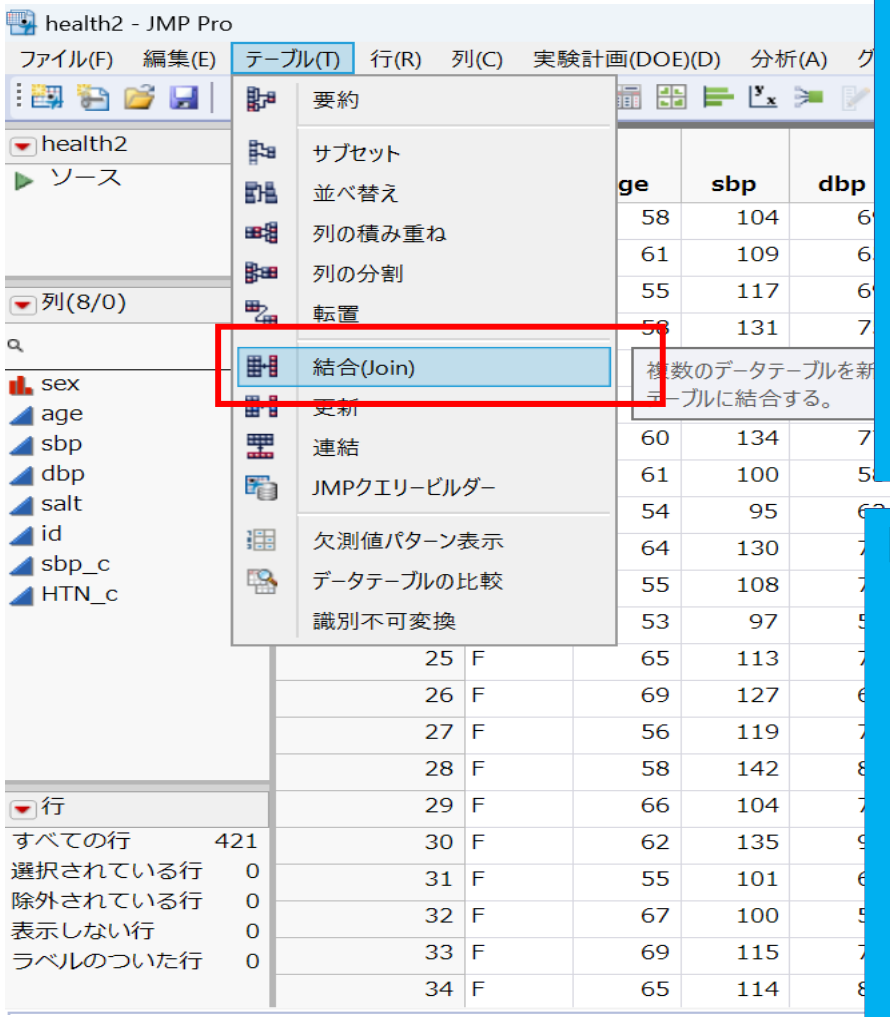


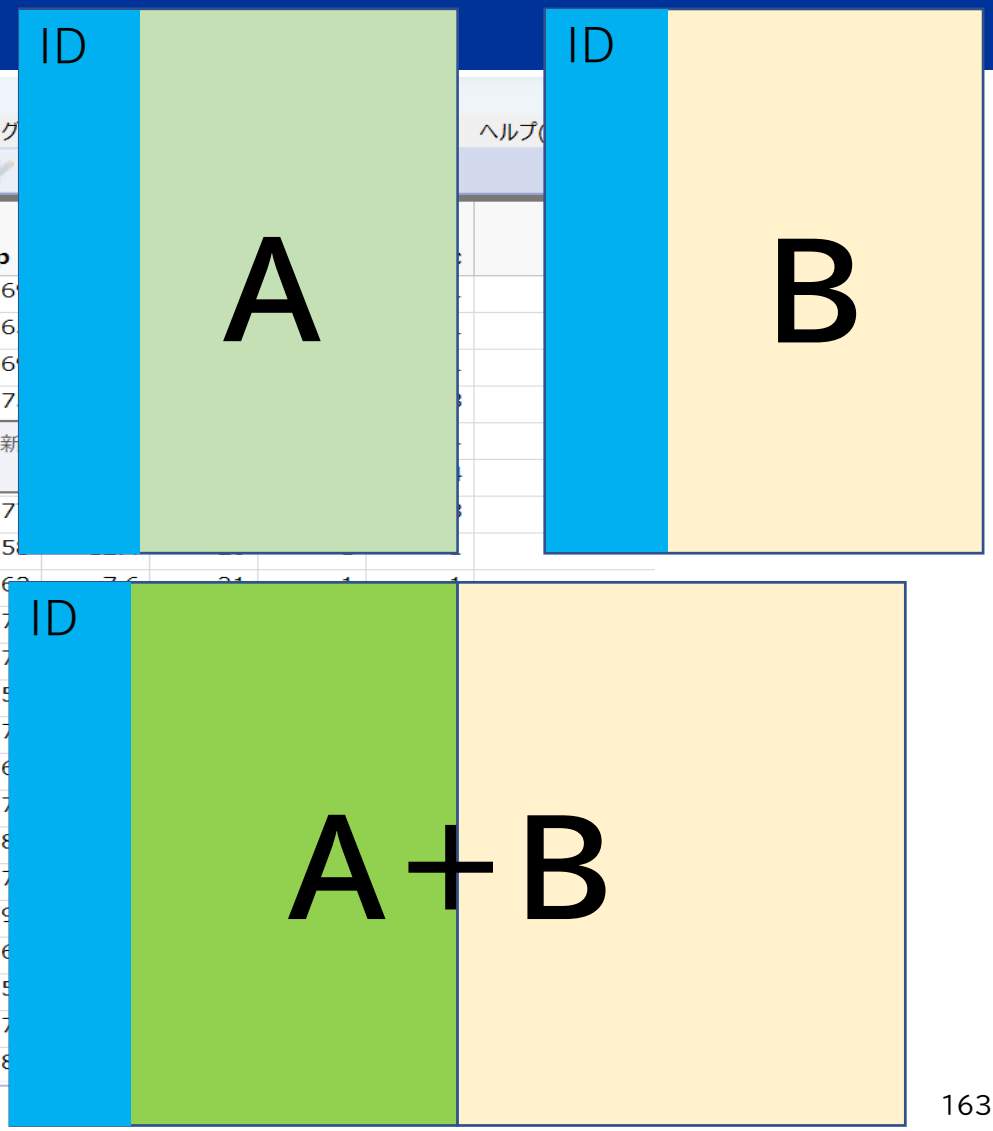


Pearson

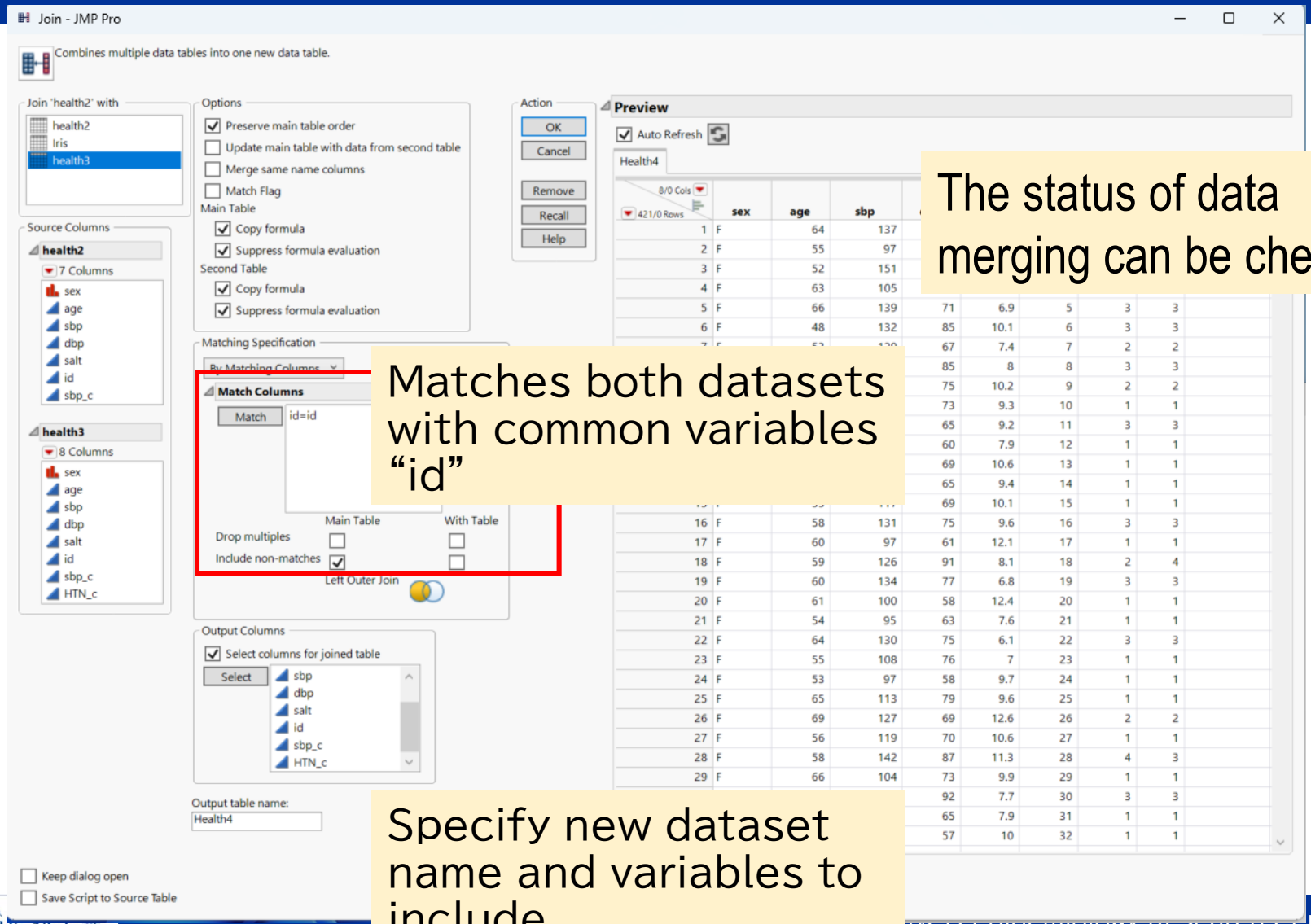
Spearman

Dataset “join”





Dataset “join”
Health2 & Health3 →Health4

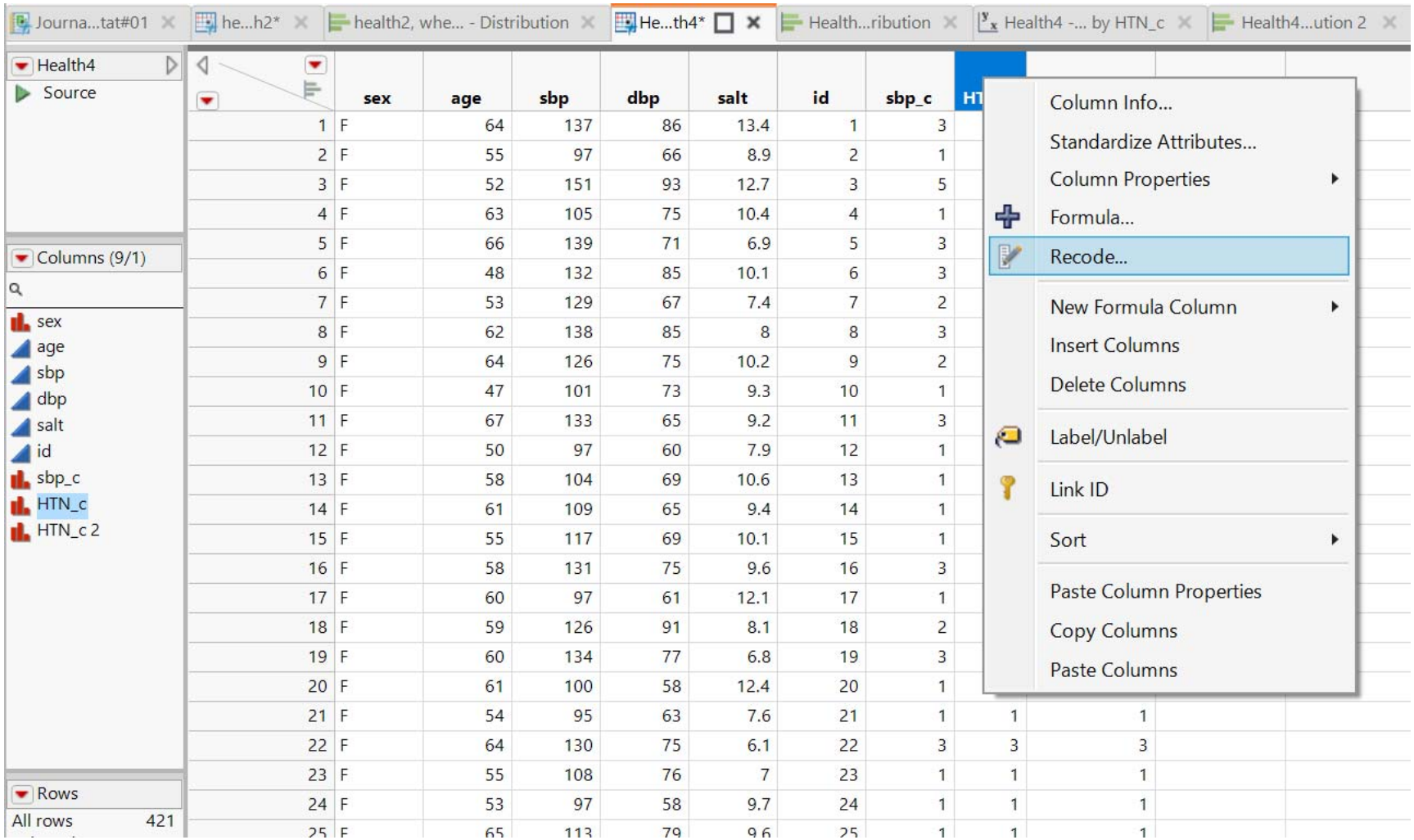


The status of data merging can be checked.

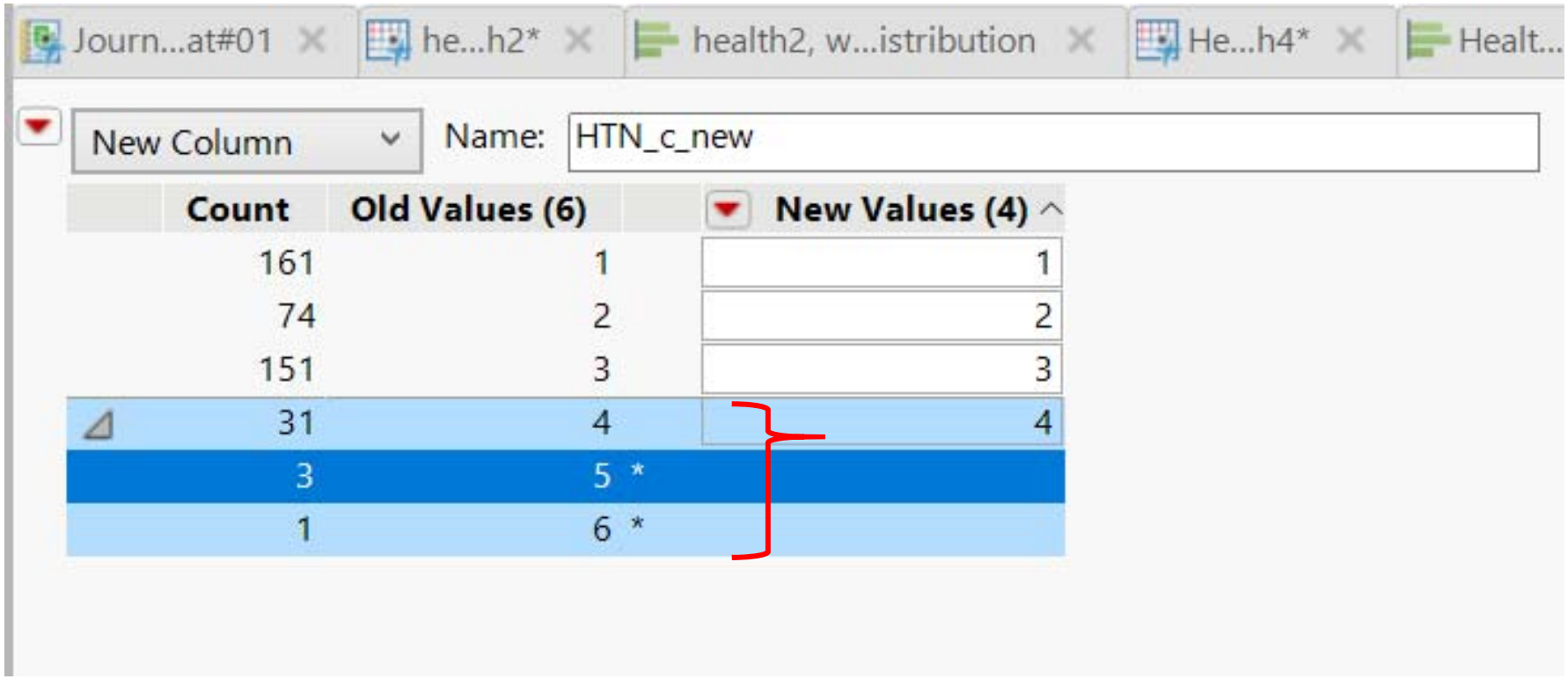
Matches both datasets with common variables “id”

Specify new dataset name and variables to include

Recode
HTN_c's 5 and 6 want to match 4 because of the small number of people



Rcode


Journ...at#01
he...h2*
health2, w...istribution
He...h4*
Healt...
New Column
Name:
HTN_c_new
Count
Old Values (6)
New Values (4)
161 1 1
74 2 2
151 3 3
31 4 4
3 5 *
1 6 *

Recode

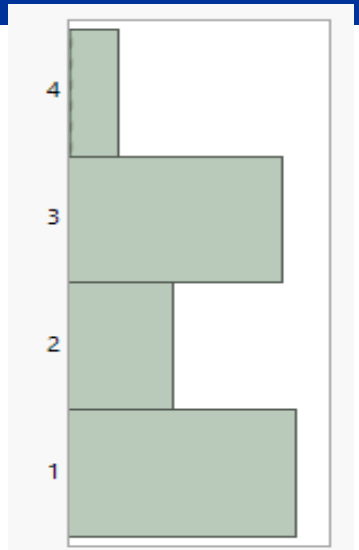
Jour...at#01 | he...h2* | health2, ...istribution | He...h4* | Healt...ution | Health...y HTN_c | Healt...uti

Health4
Source

Columns (9/1)
sex
age
sbp
dbp
salt
id
sbp_c
HTN_c
HTN_c_new

Rows
All rows 421
Selected 3
Excluded 0
Hidden 0

	sex	age	sbp	dbp	salt	id	sbp_c	HTN_c	HTN_c_new
1	F	64	137	86	13.4	1	3	3	3
2	F	55	97	66	8.9	2	1	1	1
3	F	52	151	93	12.7	3	5	4	4
4	F	63	105	75	10.4	4	1	1	1
5	F	66	139	71	6.9	5	3	3	3
6	F	48	132	85	10.1	6	3	3	3
7	F	53	129	67	7.4	7	2	2	2
8	F	62	138	85	8	8	3	3	3
9	F	64	126	75	10.2	9	2	2	2
10	F	47	101	73	9.3	10	1	1	1
11	F	67	133	65	9.2	11	3	3	3
12	F	50	97	60	7.9	12	1	1	1
13	F	58	104	69	10.6	13	1	1	1
14	F	61	109	65	9.4	14	1	1	1
15	F	55	117	69	10.1	15	1	1	1
16	F	58	131	75	9.6	16	3	3	3
17	F	60	97	61	12.1	17	1	1	1
18	F	59	126	91	8.1	18	2	4	4
19	F	60	134	77	6.8	19	3	3	3
20	F	61	100	58	12.4	20	1	1	1
21	F	54	95	63	7.6	21	1	1	1
22	F	64	130	75	6.1	22	3	3	3
23	F	55	108	76	7	23	1	1	1
24	F	53	97	58	9.7	24	1	1	1
25	F	65	113	79	9.6	25	1	1	1
26	F	69	127	69	12.6	26	2	2	2
27	F	56	119	70	10.6	27	1	1	1

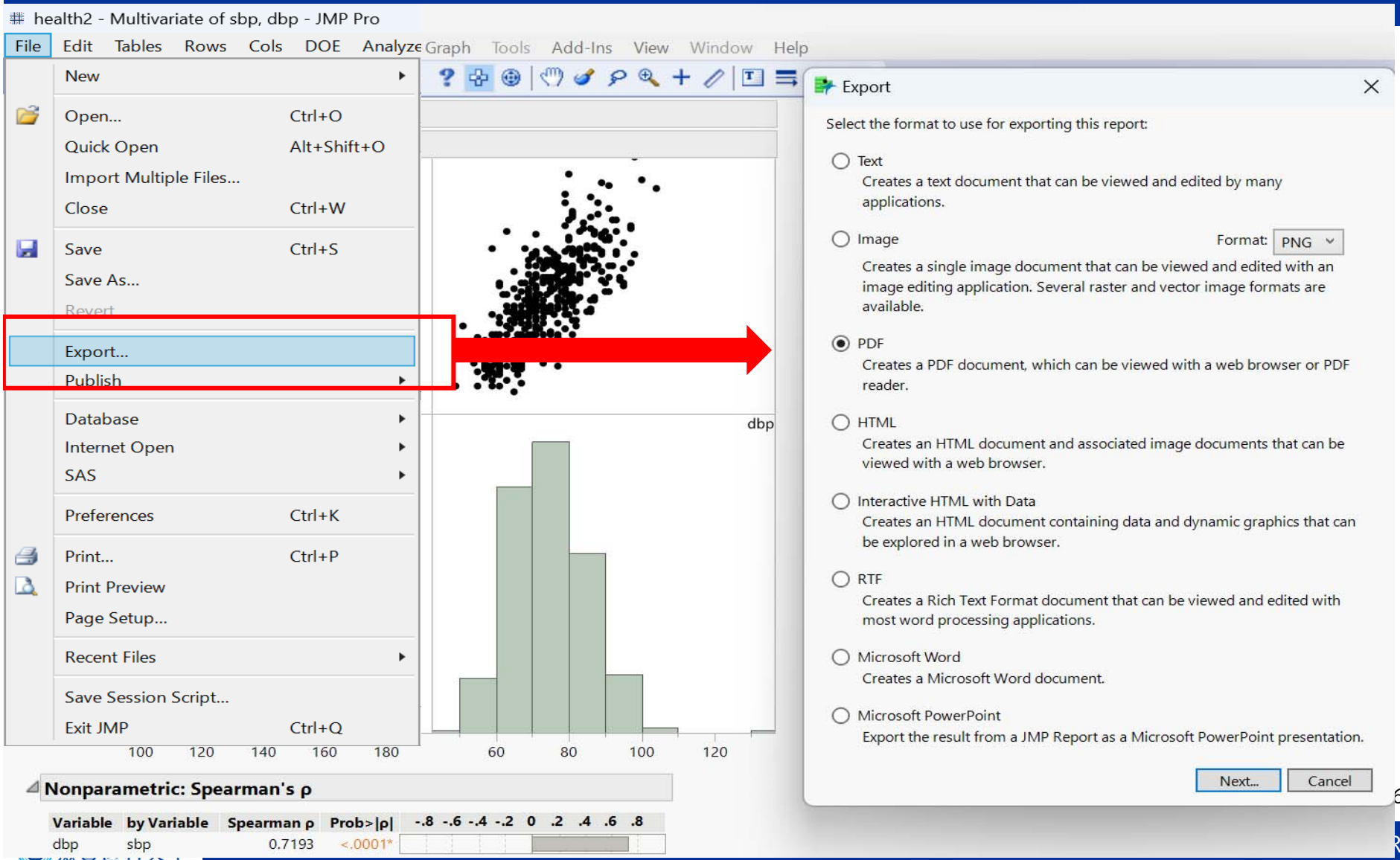


Frequencies

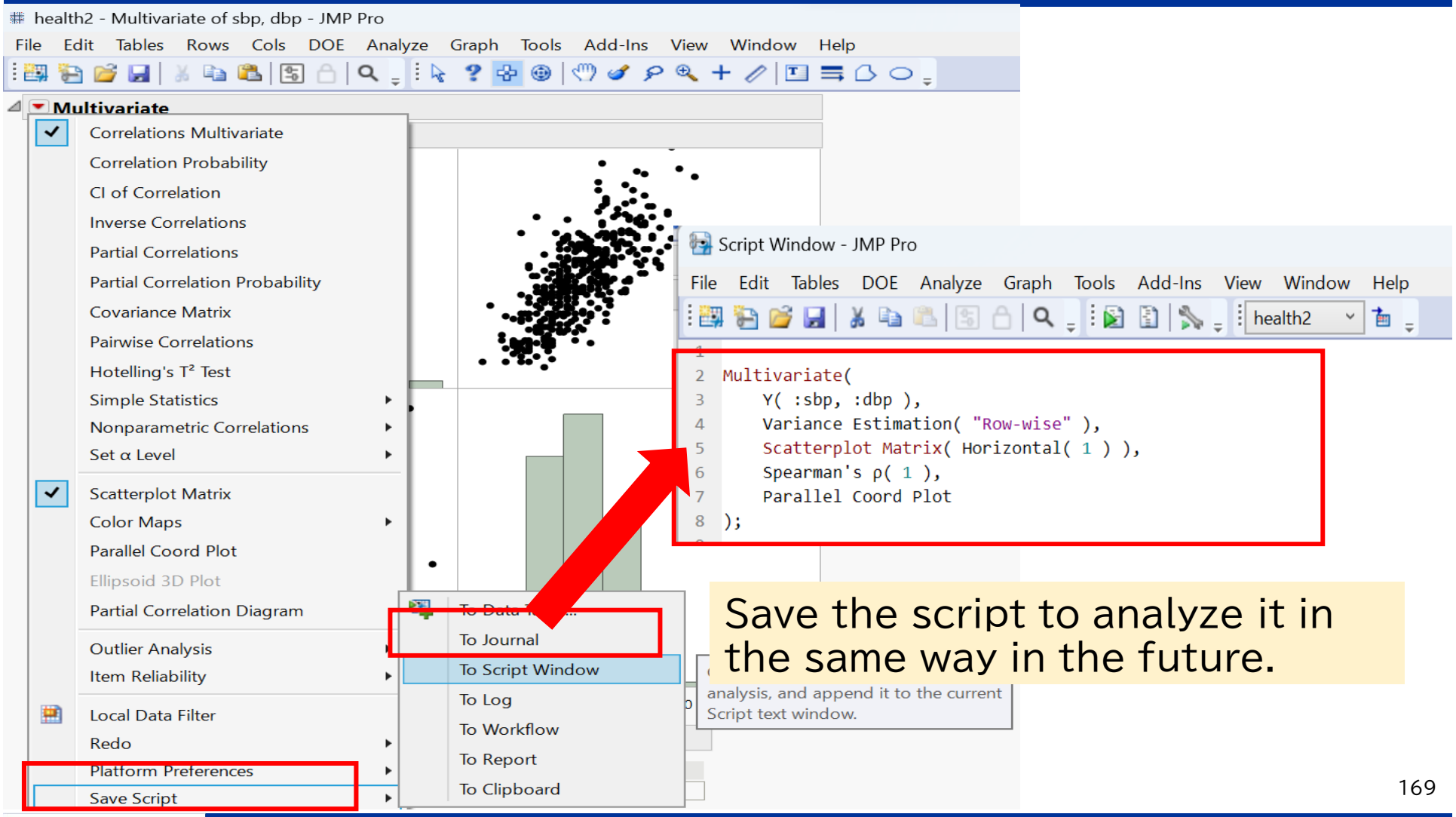
Level	Count	Prob
1	161	0.38242
2	74	0.17577
3	151	0.35867
4	35	0.08314
Total	421	1.00000

N Missing 0
4 Levels

Output of results, etc.

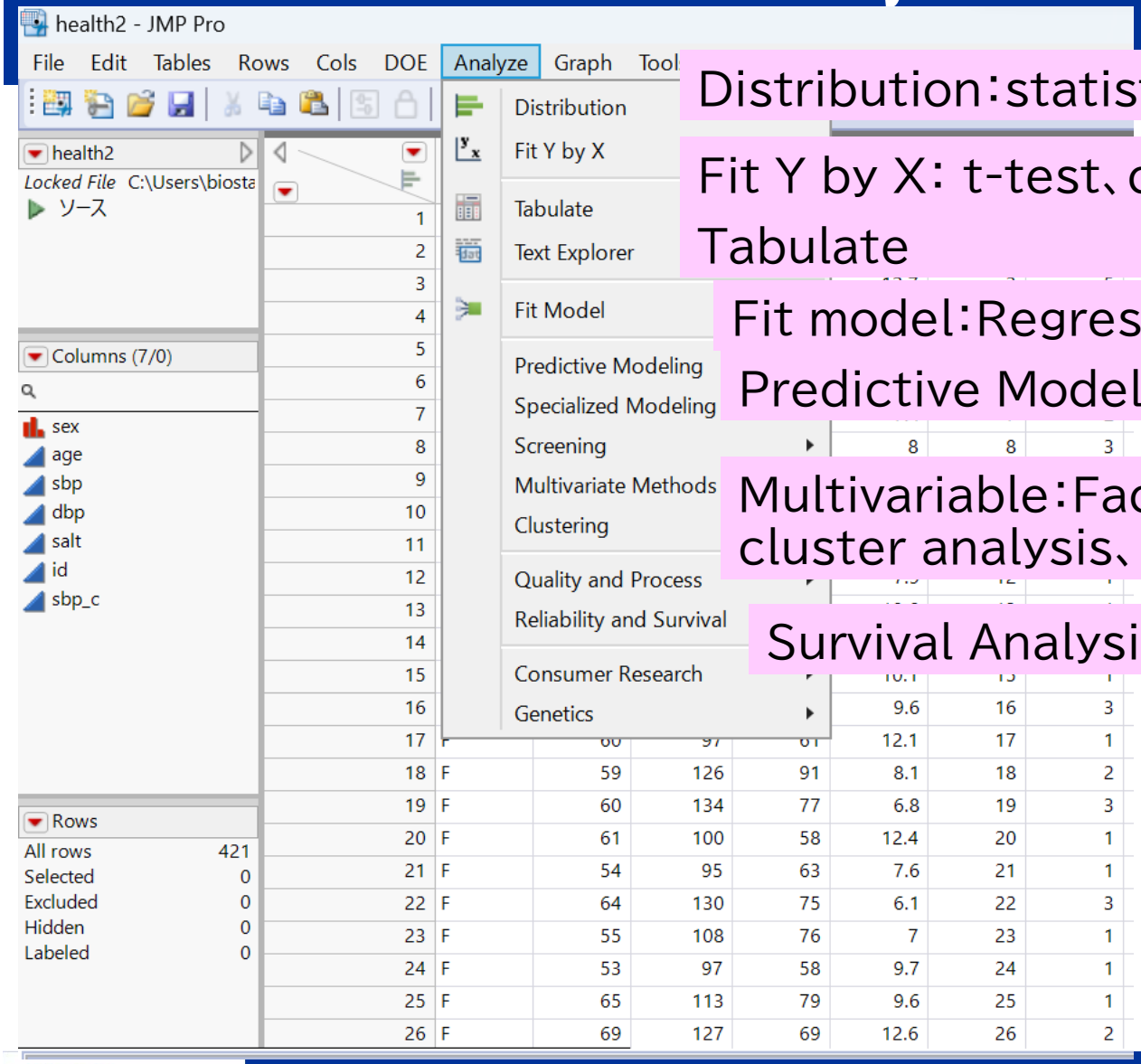


Save Analysis Results as “Script”



Save the script to analyze it in the same way in the future.

Analysis



Distribution：statistics、histogram

Fit Y by X: t-test、chi-square、ANOVA

Tabulate

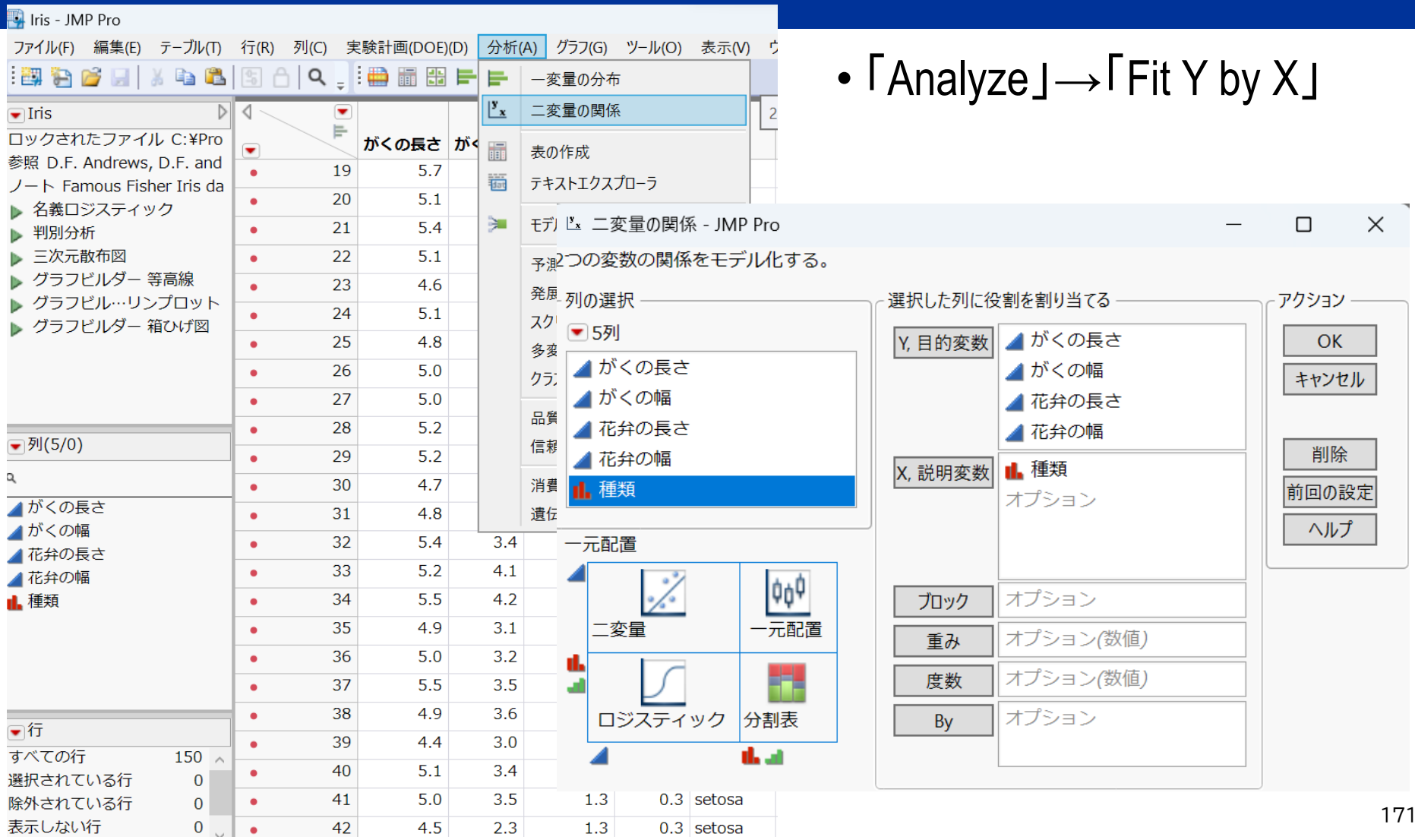
Fit model：Regression

Predictive Modeling：ML

Multivariable：Factor analysis、cluster analysis、LCA、SEMなど

Survival Analysis

Comparison of sample data
Two-sample test, (t-test, Wilcoxon test, χ2 test, etc.)



- 「Analyze」→「Fit Y by X」

Tabulate

Tabulate

To add to the table, drag and drop columns or statistics into the column header or row label area of the table.

Undo | Start Over | Done

7 Columns

- sex
- age
- sbp
- dbp
- salt
- id
- sbp_c

Freq

Weight

ID

Page Column

Mean
Std Dev
Min
Max
Range
% o...otal
N ...ssing
N C...ries
Sum
Sum Wgt
Variance
Std Err
CV
Median
Me...tion
Ge...ean
Inte...nge
Quantiles
Mode
Col...n %
Row %
All

☐ Include missing for grouping columns

☐ Order by count of grouping columns

☐ Add Aggregate Statistics

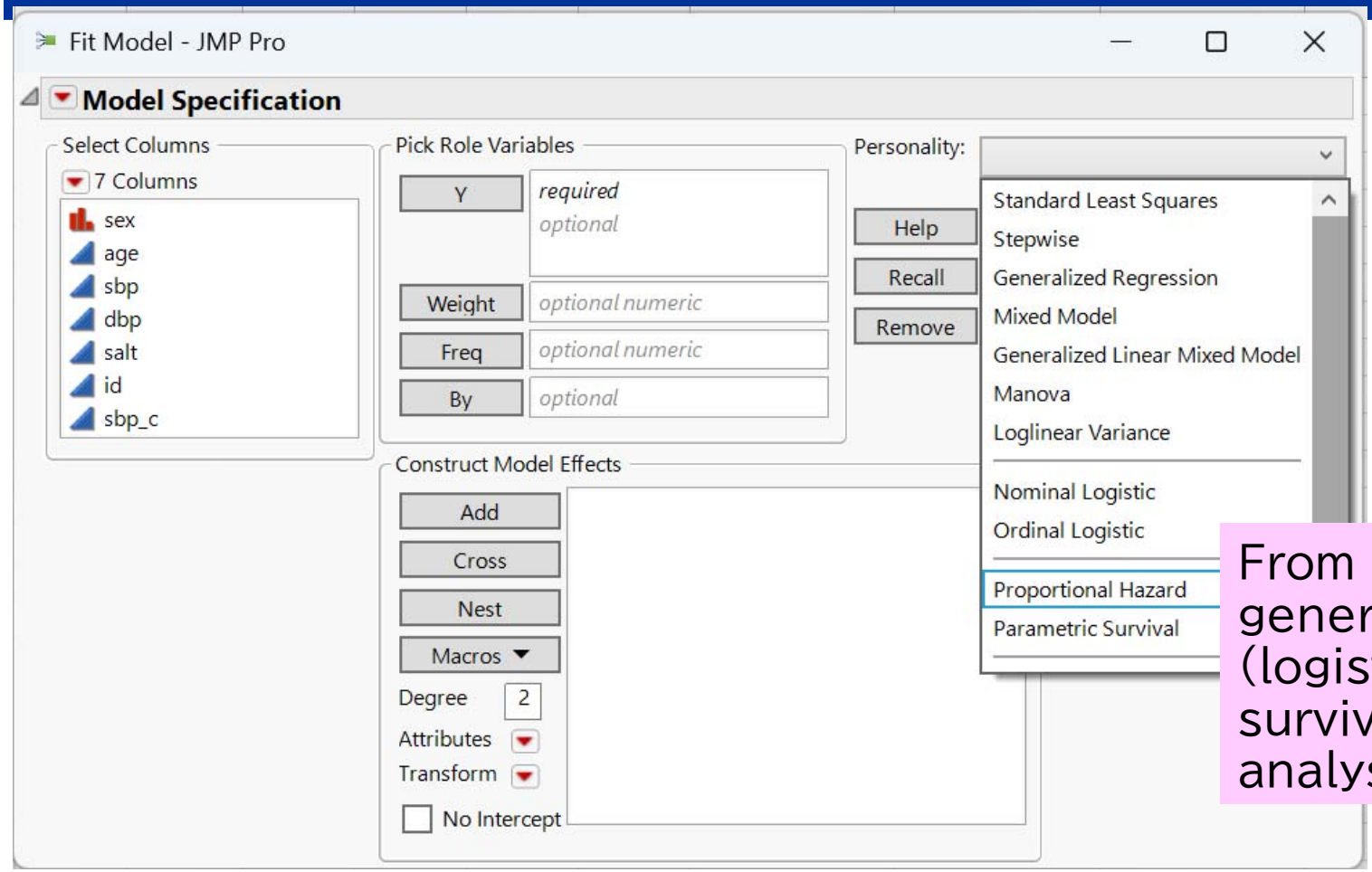
Default Statistics

Change Format

	sex							
	F				M			
	age	sbp	dbp	salt	age	sbp	dbp	salt
N	244	244	244	244	177	177	177	177
Mean	58.454918033	122.7295082	72.233606557	9.3254098361	59.525423729	125.48587571	76.615819209	9.4124293785
Std Dev	8.0580466387	16.271885868	9.5028010656	2.1017751257	7.4468326066	15.804597772	10.703619743	1.9264791474
Min	40	91	49	4.9	40	93	55	3.8
Max	69	187	100	17.2	69	180	130	15.4

Drop and add variables and statistics

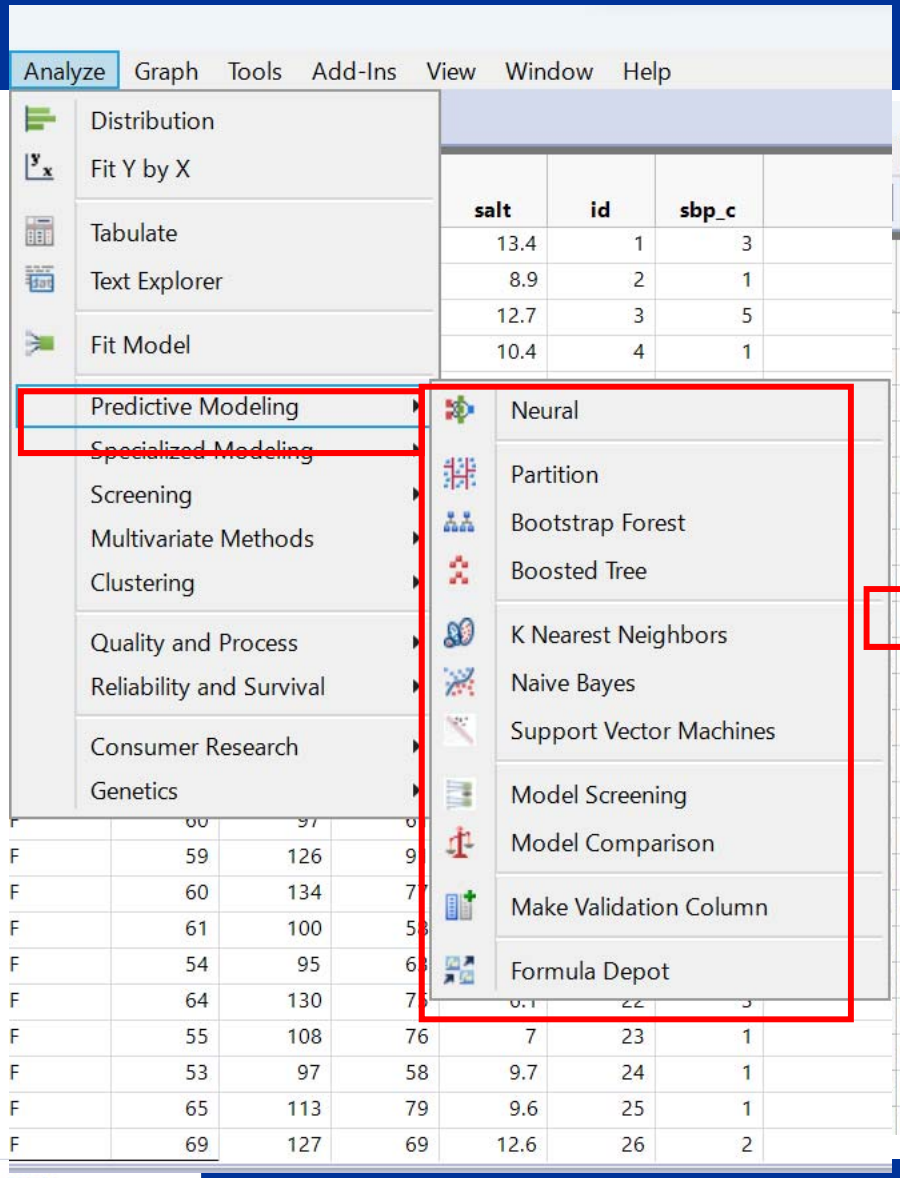
Fit model:Regression

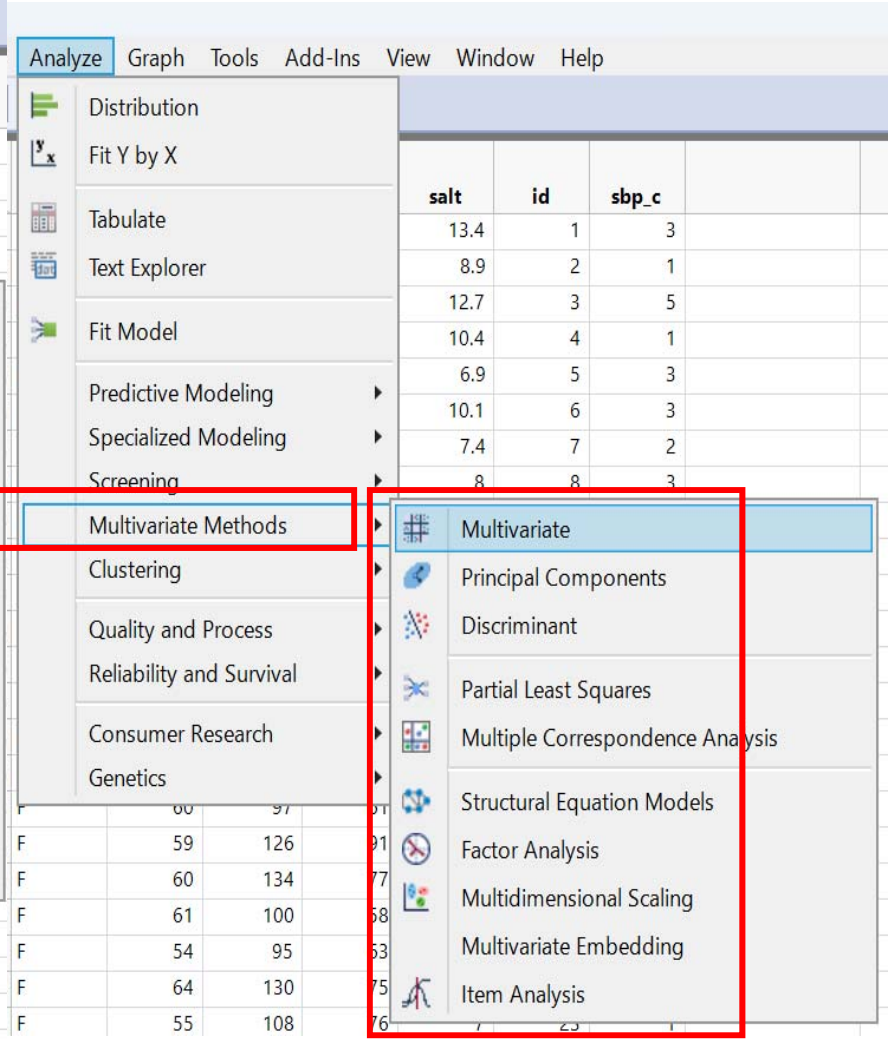


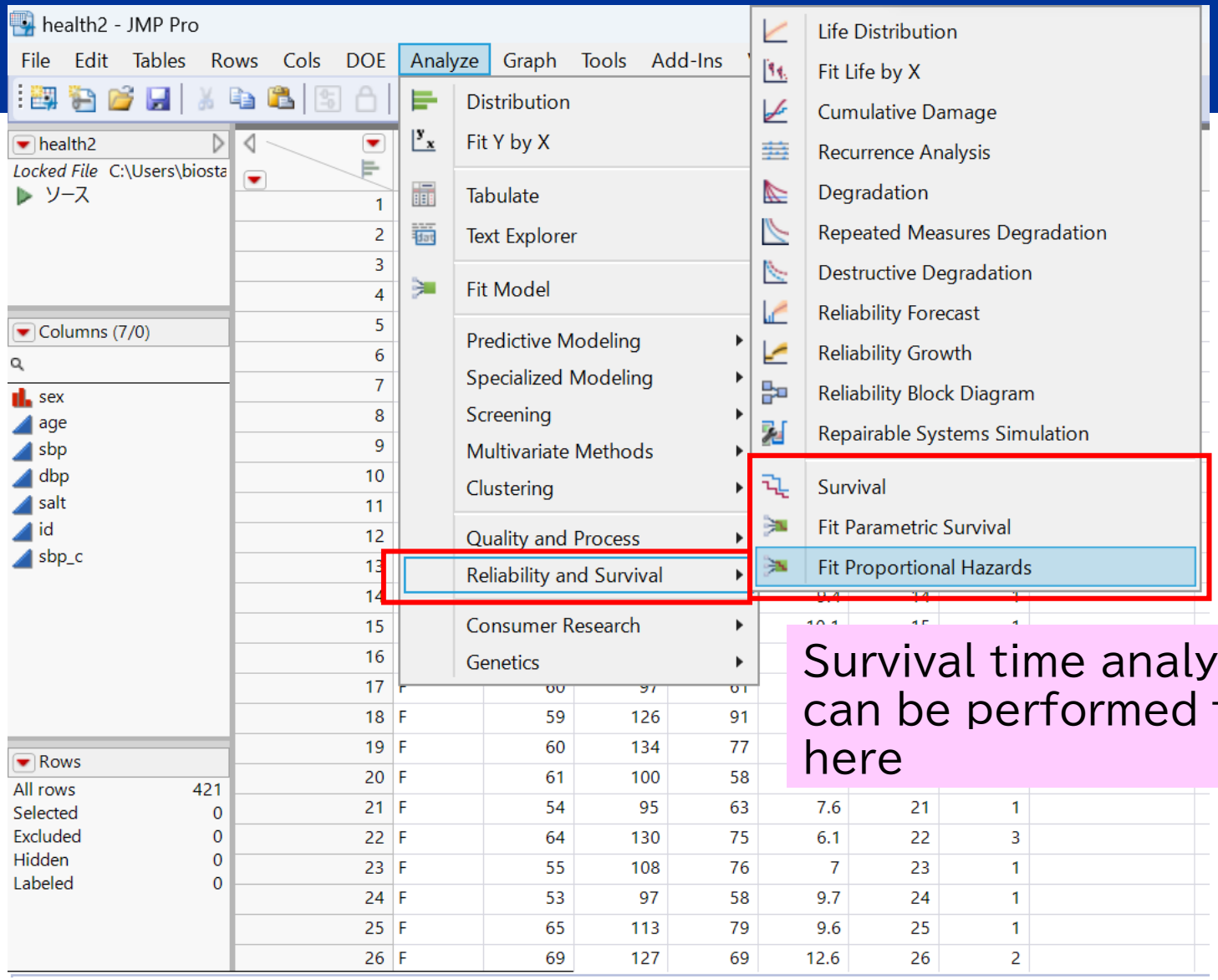
From here, generalized linear (logistic, Poisson), survival time analysis is possible

Predictive Modeling：ML

Multivariable：Factor analysis、cluster analysis、LCA、SEMなど



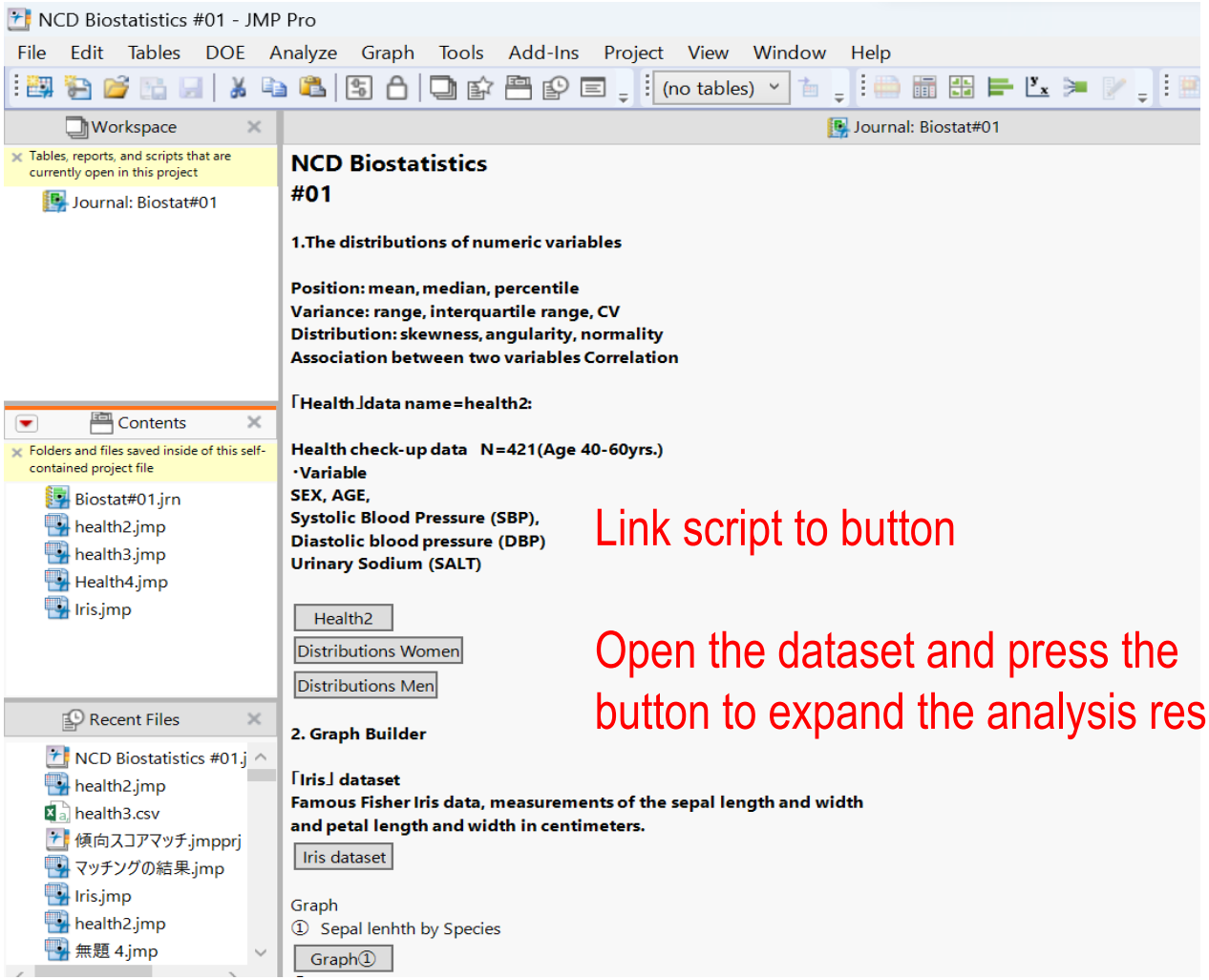




Survival time analysis can be performed from here

Journal and Project functions

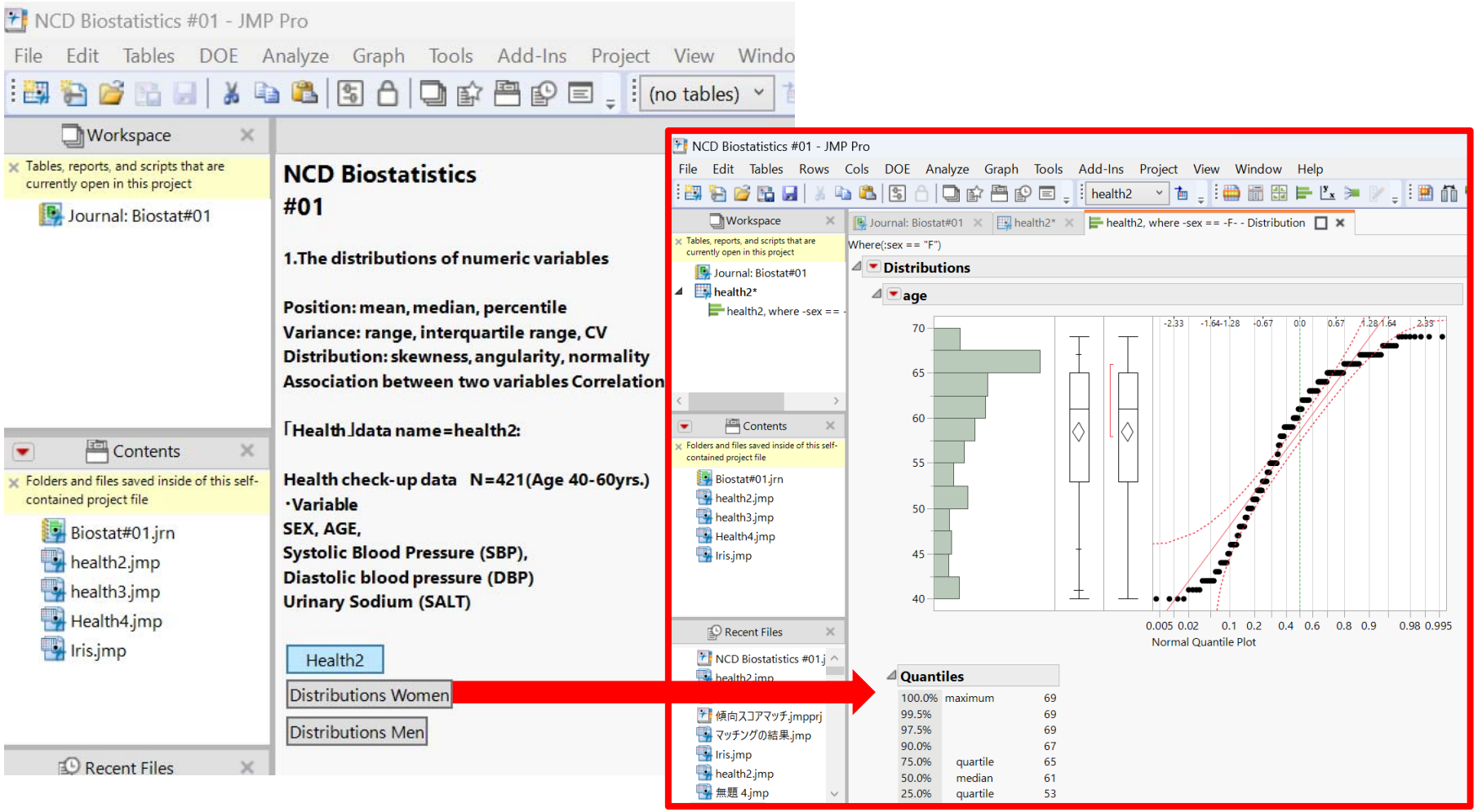
- A series of analysis processes can be organized and saved



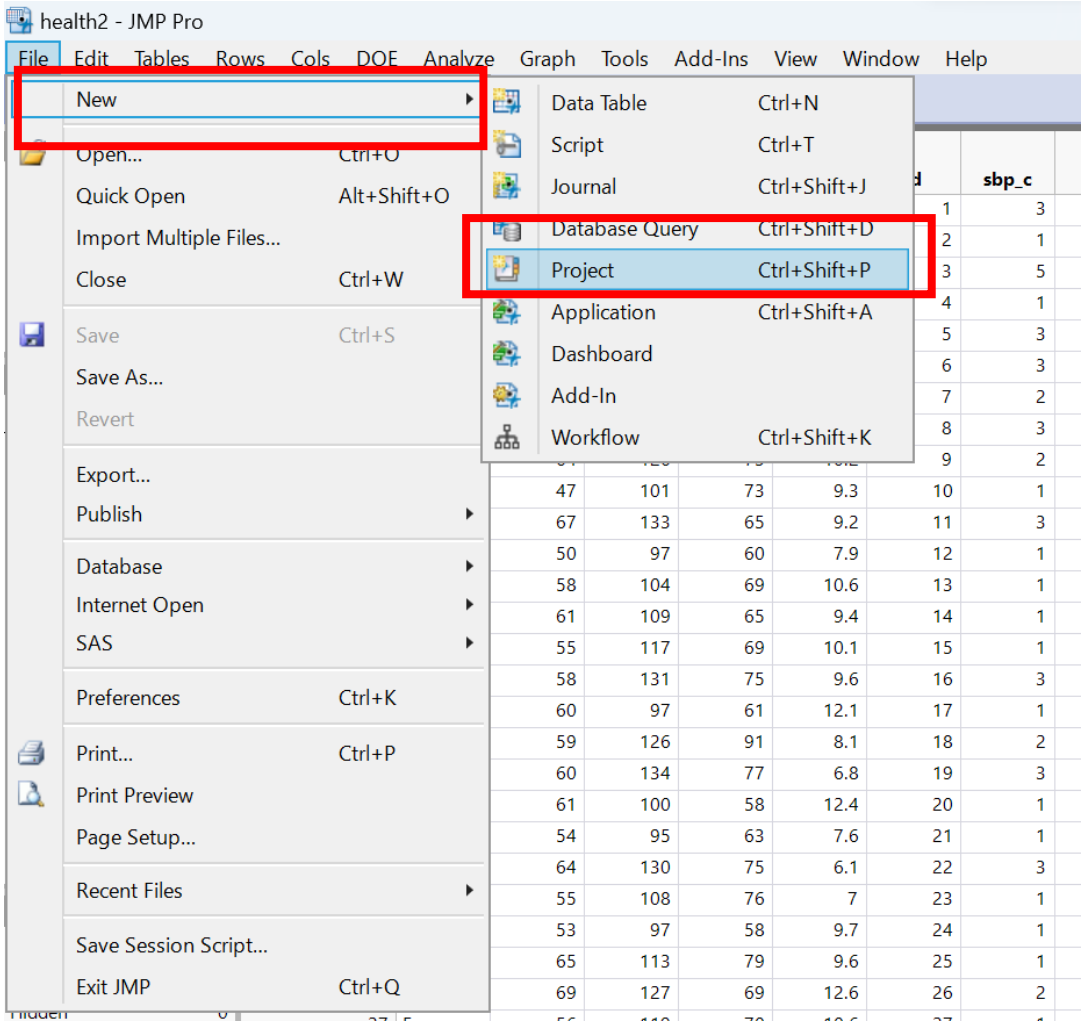
Link script to button

Open the dataset and press the button to expand the analysis results

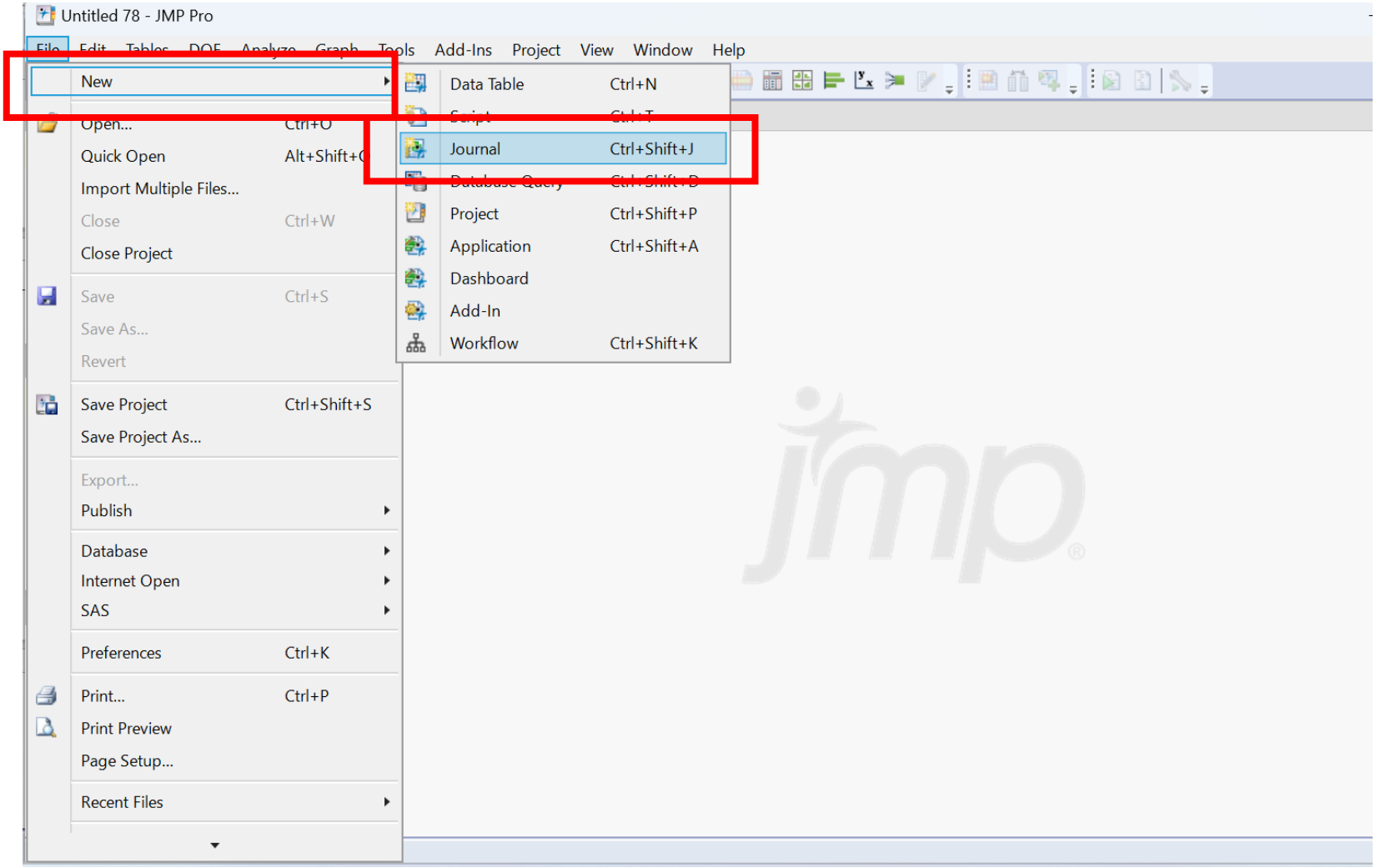
Dataset & Results are expanded



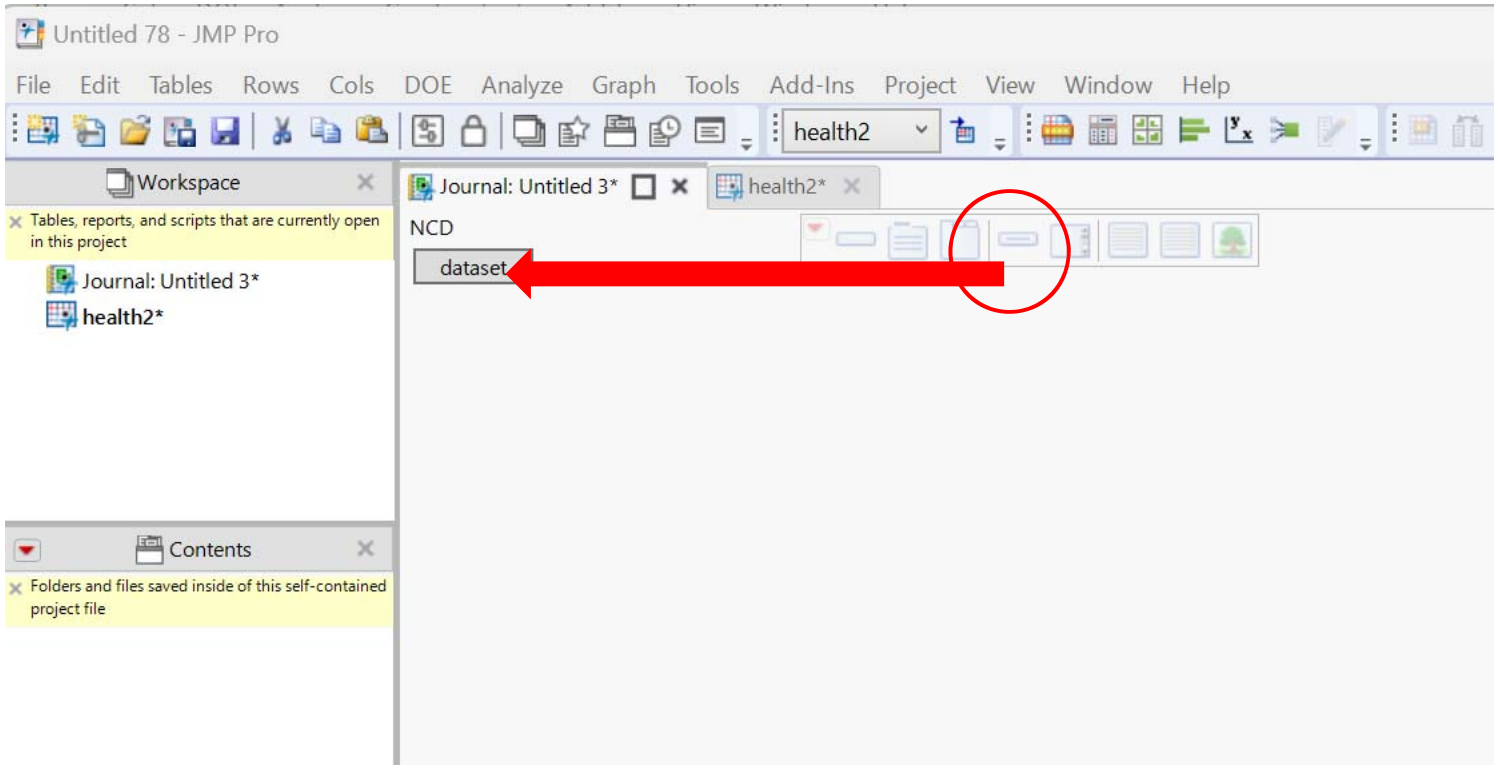
How to build “Project”



Build Journal in Project screen

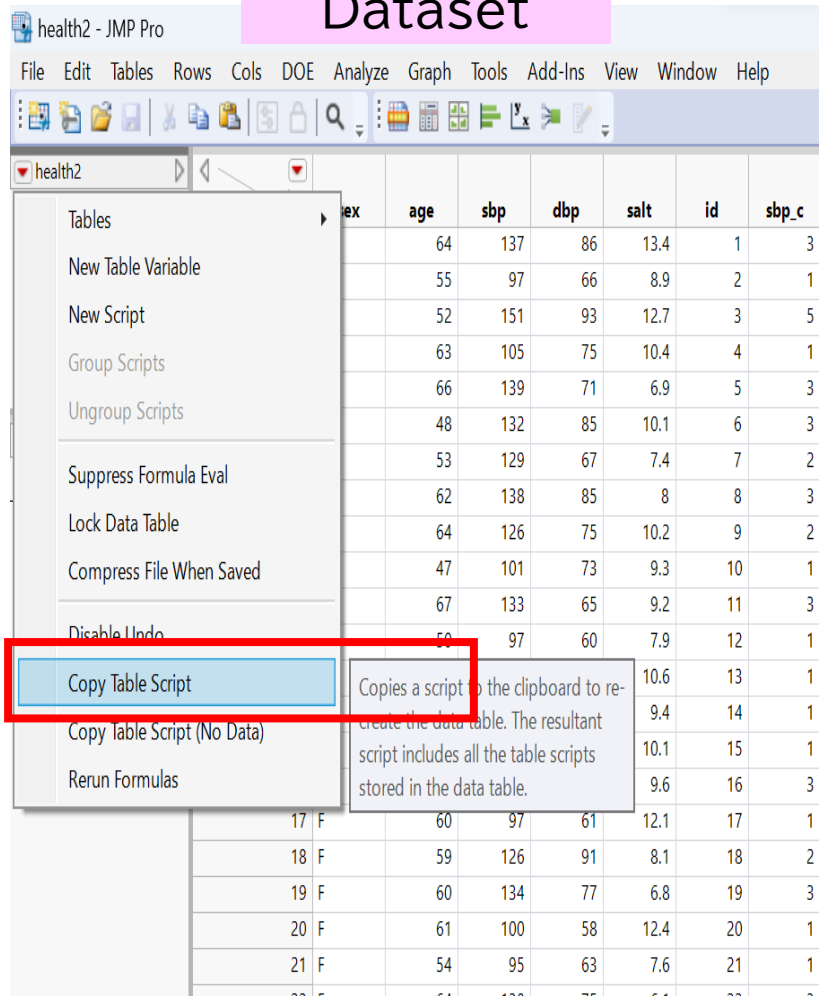


Build Text & Button

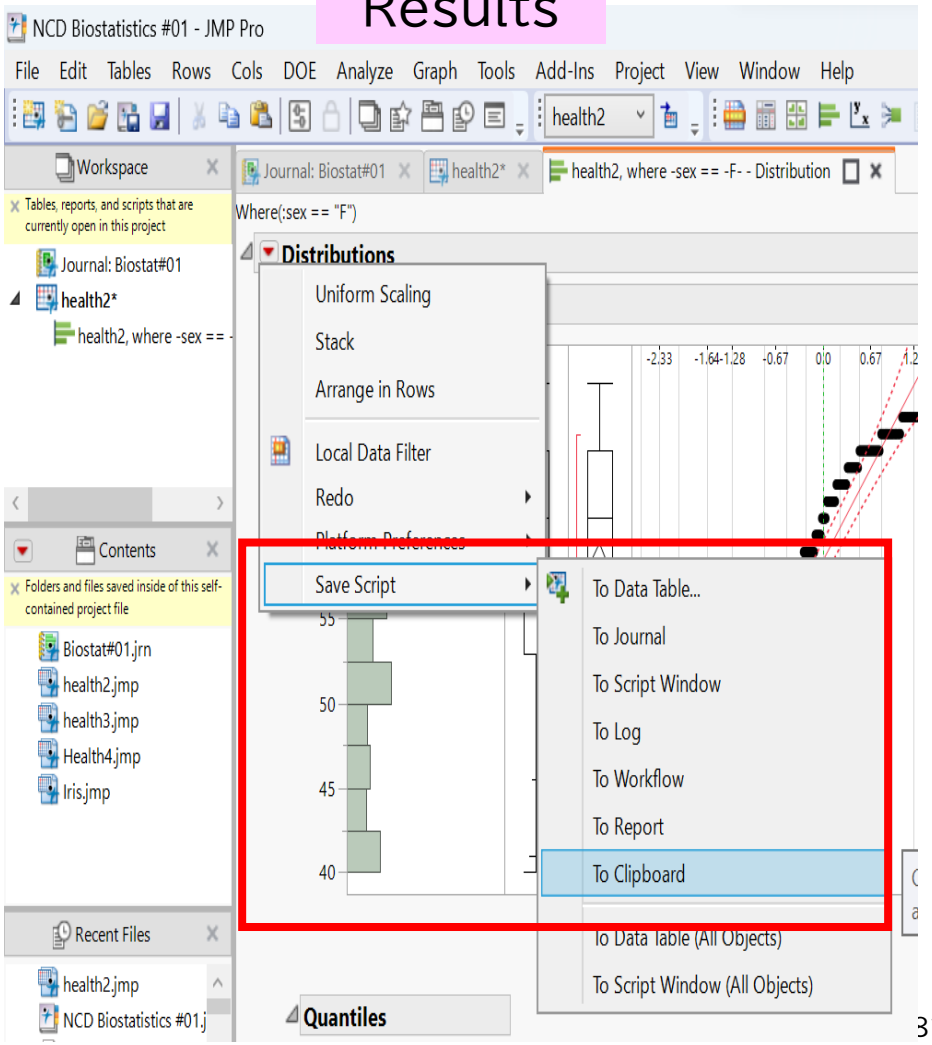


Copy Script

Dataset



Results



Set script

Untitled 78 - JMP Pro

File Edit Tables DOE Analyze Graph Tools Add-Ins Project

Workspace

Tables, reports, and scripts that are currently open in this project

Journal: Untitled 3*

Contents

Folders and files saved inside of this self-contained project file

Recent Files

health2.jmp
NCD Biostatistics #01.jmpprj
health3.csv
傾向スコアマッチ.jmpprj

NCD

dataset

Set Button Name
Style
Set Script...
Encrypt Script...
Decrypt Script...
Show Properties
Edit

Edit Script

Name: dataset

Script:

Paste the copied script here

OK Cancel Help

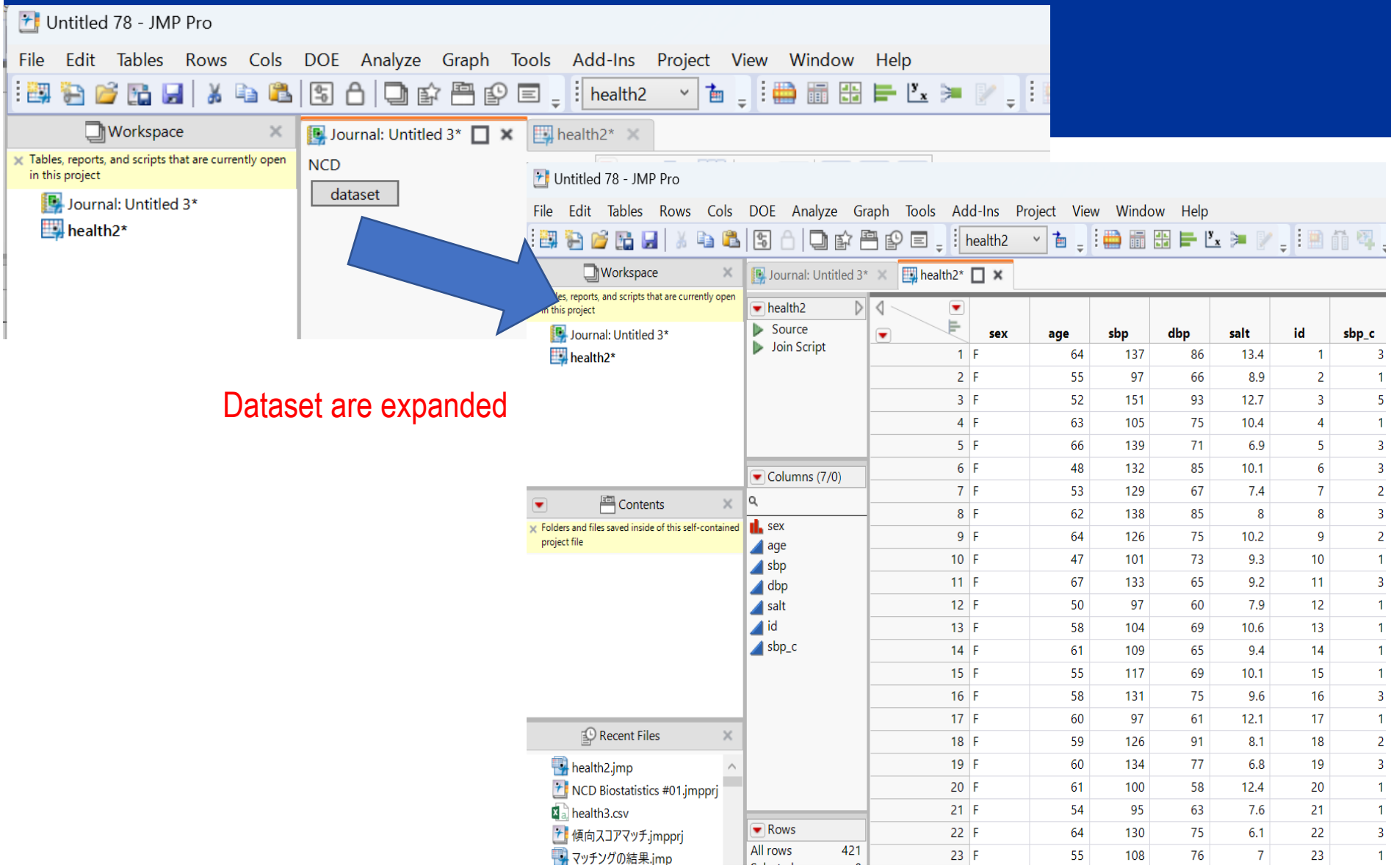
Edit Script

Name: dataset

Script:

```
New Table( "health2",
    Add Rows( 421 ),
    New Script(
        ["en" => "Source", "ja" => "ソース"],
        Open(
            "/E:/■■■大学/◎◎滋賀医大/■■講義関連/だい/20
            columns(
                New Column( "sex", Character, "Nominal"
                New Column( "age", Numeric, "Continuous
                New Column( "sbp", Numeric, "Continuous
                New Column( "dbp", Numeric, "Continuous
                New Column( "salt", Numeric, "Continuou
                New Column( "id", Numeric, "Continuous"
                New Column( "sbp_c", Numeric, "Continuo
                New Column( "c", Numeric, "Continuous",
                New Column( "HTN_c", Numeric, "Continuo
                New Column( "HTN", Character, "Nominal"
                New Column( "_c", Character, "Nominal"
            ),
            Import Settings(
                End Of Line( CRLF, CR, LF ),
```

OK Cancel Help



Dataset are expanded

Report

- For the following Q1-Q4 assignments shown in the lecture slides, check that the results can be reproduced using a statistical package JMP.

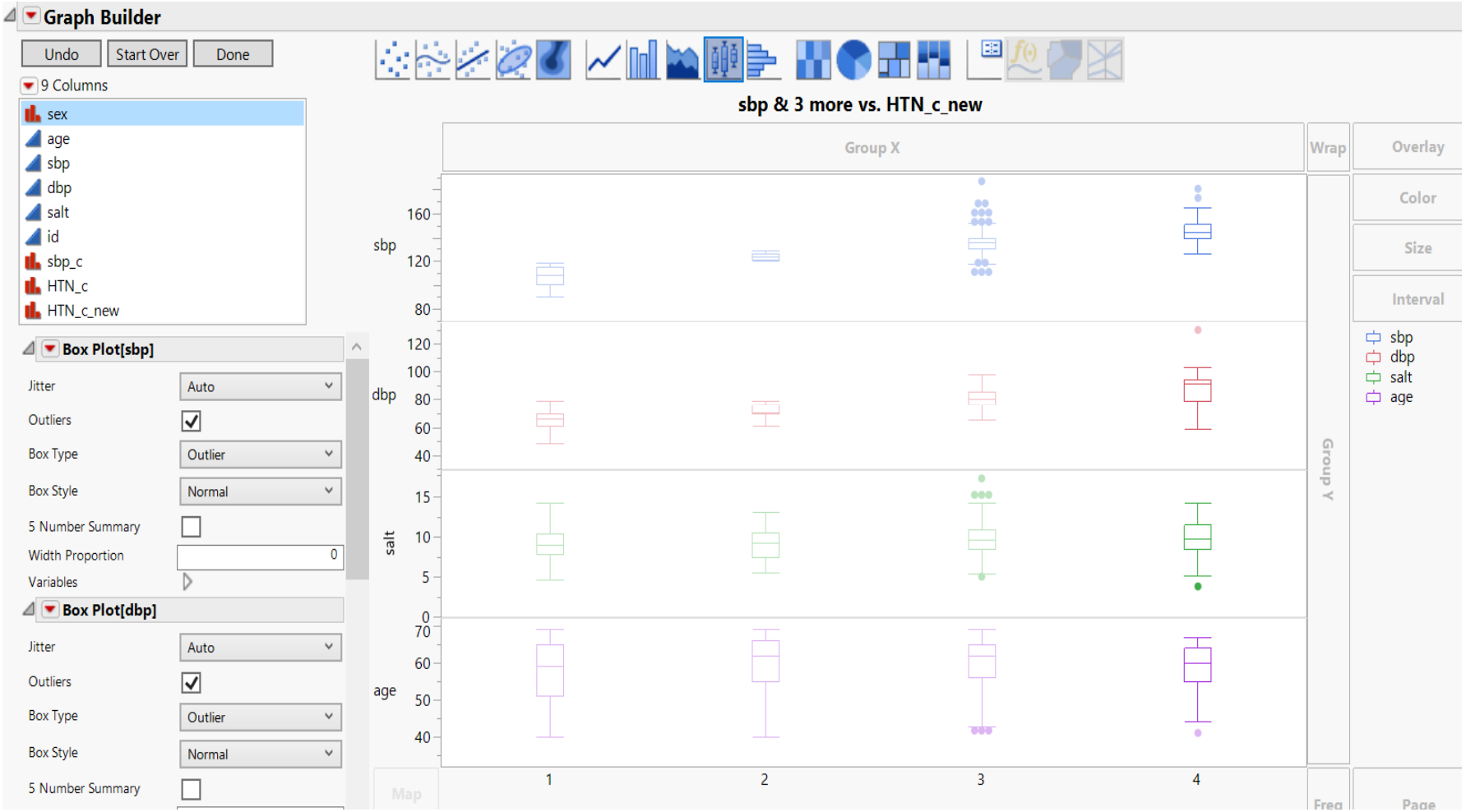
  - Q1 Check the shape, location, and distribution (var= SBP, DBP, SALT).
  - Q2 Using “health2” dataset, calculate correlation coefficients between blood pressure and salt intake.
  - Q3 The ”Health4” dataset created in the course of the exercise has a blood pressure category variable (HTN_c) added to it (HTN_c should Recode 4-6 as 4).

    This was used to

    1) Create a box-and-whisker diagram of SBP, DBP, and SALT for each blood pressure category.
    →Use Graph builder.

    2) Tabulate the number of subjects, mean, SD, max, and min values of SBP, DBP, and SALT for each blood pressure category.
    →Use tabulate,

Hint


Graph Builder
Undo
Start Over
Done
9 Columns
sex
age
sbp
dbp
salt
id
sbp_c
HTN_c
HTN_c_new
Box Plot[sbp]
Jitter
Auto
Outliers
Box Type
Outlier
Box Style
Normal
5 Number Summary
Width Proportion
0
Variables
Box Plot[dbp]
Jitter
Auto
Outliers
Box Type
Outlier
Box Style
Normal
5 Number Summary
sbp & 3 more vs. HTN_c_new
Group X
Wrap
Overlay
Color
Size
Interval
sbp
dbp
salt
age
Group Y
Map
Freq
Page
1
2
3
4