

Monthly User Guide from JMP Korea

제 14호 (2018년 9월)

알쓸신잡 (알아 두면 쓸 데 있는 신비한 잡학 사전) of JMP

Small but quite Useful Tips of JMP

* 본 Guide 는 매월 세 번째 수요일에 발행됩니다

(2018년 7월호부터는 JMP 14 Version 기준입니다)

** Monthly User Guide 지난 호는 다음 Site 를 참조하세요(https://www.jmp.com/ko_kr/newsletters.html)

*** 본 Guide 의 내용과 관련한 문의는 Ikju.Shin@jmp.com 으로 연락 바랍니다



알쓸신잡 of JMP(Small but quite Useful Tips of JMP)

이번 호에서는 JMP 사용자들이 알아 두면 유용한 10가지의 Tip 을 정리해 보았습니다

1. 분석 결과를 저장하는 방법
2. Data Table 에 다른 Data Table 을 저장하는 방법
3. 손쉽게 도움말을 이용하는 방법
4. 보다 실무적으로 Subset(Data 추출) 하는 방법
5. Graph 에서 순차적으로 Hide/Unhide 하는 방법
6. Variability Chart 에서 Axis(축) 표현 방법 차이
7. 공정 능력 분석에서 Spec 을 한꺼번에 입력하는 방법
8. 일정 범위 이상의 Data 를 Missing(결측치) 으로 처리하고 싶다
9. 특정 변수를 항상 표현되게(Display 되게) 하고 싶다
10. 가장 유용한 단축키 2가지

1. 분석 결과를 저장하는 방법

1. 아래와 같은 data 가 있다고 가정하고

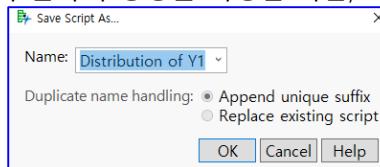
Operator	Y1	Y2
1 A	35	54
2 A	46	56
3 A	54	35
4 A	64	56
5 A	56	89
6 B	46	54
7 B	47	45
8 B	57	35
9 B	58	45
10 B	47	76

2. Analyze / Distribution 에서 Y1 을 선택, OK 한 다음
그 결과를 저장하는 방법은

▼ Distributions / save script 아래에 몇 가지 저장 방법이 있는 데
그 중 data table 에 저장하는 방법이 가장 일반적입니다

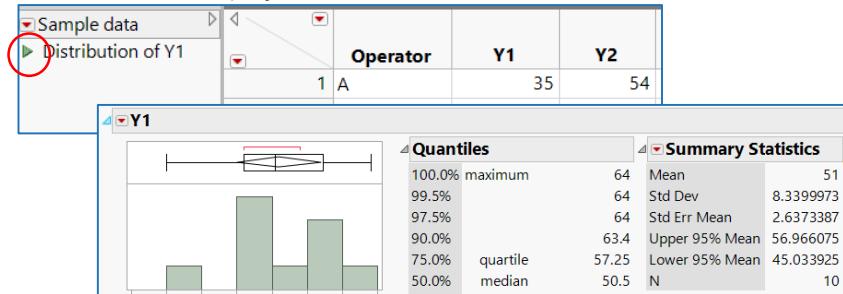


3. ▼ Distributions / save script / to data table 을 클릭한 후
저장할 분석 결과의 명칭을 지정한 다음, OK 를 클릭하면

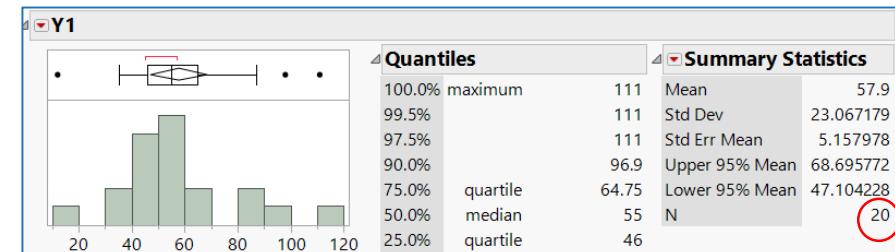


4. 아래와 같이 Table Panel 에 저장됨.

- 저장된 명칭 왼쪽 옆의 연두색 삼각형을 click 하거나
- 저장된 명칭 위에서 오른쪽 마우스 클릭, Run Script' 를 클릭하면
분석 결과가 Display 됨



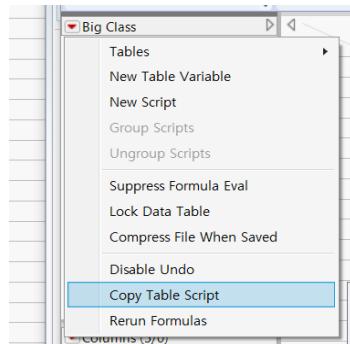
5. Data 가 Update(추가되었다) 하더라도, 위의 4번처럼 하면
Update(또는 추가)된 분석 결과가 Display 됨.



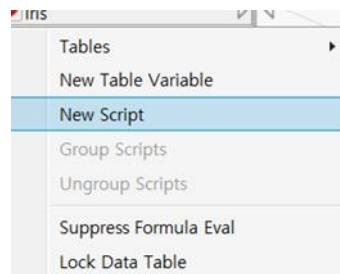
2. Data Table 에 다른 Data Table 을 저장하는 방법

1. A, B 두 개의 Data Table 이 있고 B Data Table 을 A Data Table 에 저장한다고 가정

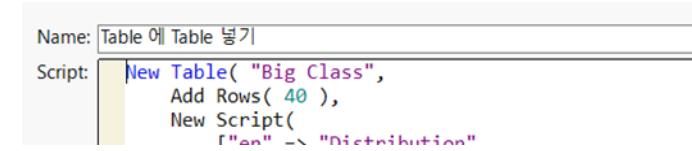
2. 저장할 B Data table, 왼쪽 위 Table panel 의 왼쪽 빨간 색 역삼각형에서 copy table script 선택한 후



3. A Data Table 의 왼쪽 Table panel 의 왼쪽 빨간 색 역삼각형에서 new script 선택

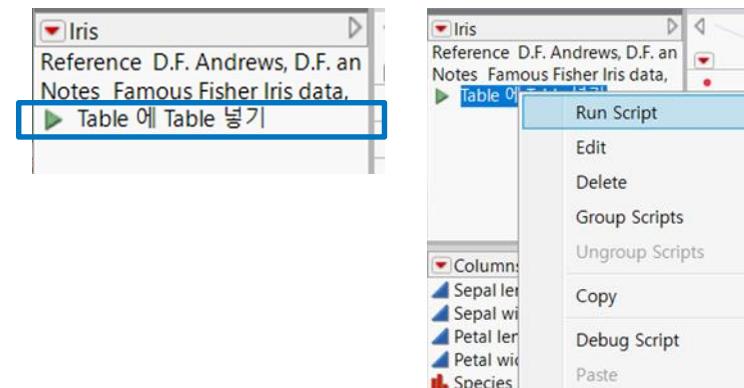


4. Script Name 입력 후 OK 클릭



5. 이렇게 하면 아래와 같이 저장됩니다.

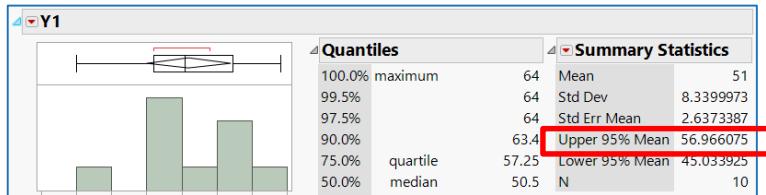
-저장된 명칭 왼쪽 옆의 연두색 삼각형을 click 하거나
-저장된 명칭 위에서 오른쪽 마우스 클릭,
'Run Script' 를 클릭하면 저장된 data table 이 open 됩니다



3. 손쉽게 도움말을 이용하는 방법

1. 1번(분석 결과를 저장하는 방법)과 같은 Data 가 있다고 가정하고

2. Analyze / Distribution 에서 Y1 을 선택하여 실행한 결과 중 'Upper 95% Mean' 의 의미에 대해 학습하고자 한다면



3. 대부분의 경우는 Help / JMP Help 에 들어가서, 'Upper 95% Mean' 를 입력하여 검색함. 이럴 경우 아래와 같이, 꽤 많은 결과가 검색되어서 살펴보자 하는 것을 정확히 찾기가 쉽지 않음.

Search
Upper 95% Mean
Submit

Multiple Comparisons
Use this option to obtain tests and confidence levels that compare means defined by levels of your model. The goal of multiple comparisons methods is to determine whether group means differ, while controlling the risk of reaching an incorrect conclusion. The Multiple Comparisons option...

Effect Details
The Effect Details report provides details, plots, and tests for individual effects. It consists of separate reports for each effect, with the emphasis that you select in the Fit Model launch window.

Oneway Platform Options
The Fit Group menu appears if you have specified multiple Y or X variables. Menu options allow you to fit or order them by RSquare. See Fit Group Options in the Fitting Linear Models book for more information.

Display Options
Using Display Options, you can add or remove elements from a plot. Some options might not appear until you run the report.

The Summary Statistics Report
For continuous variables, the Summary Statistics report displays the mean, standard deviation, and other statistics. You can control which statistics appear in this report by selecting Customize Summary Statistics from the triangle menu next to Summary Statistics.

Linear Fit and Polynomial Fit Reports
The Linear Fit and Polynomial Fit reports begin with the equation of fit.

4. 손쉽게 도움말을 이용하는 방법

1) 상단의 도구 모음에서 ? 를 click 한 후, 살펴보자 하는 'Upper 95% Mean' 위에서 한 번 더 click 하면



2) 아래와 같이 해당되는 도움말 Page 로 바로 이동함.

Description of the Summary Statistics Report

Mean

Estimates the expected value of the underlying distribution for the response variable of the nonmissing values divided by the number of nonmissing values.

Std Dev

The normal distribution is mainly defined by the mean and standard deviation. This becomes large:

- 68% of the values are within one standard deviation of the mean
- 95% of the values are within two standard deviations of the mean
- 99.7% of the values are within three standard deviations of the mean

Std Err Mean

The standard error of the mean, which estimates the standard deviation of the distribution.

Upper and Lower Mean Confidence Limits

Are 95% confidence limits about the mean. They define an interval that is very likely to contain the true mean.

N

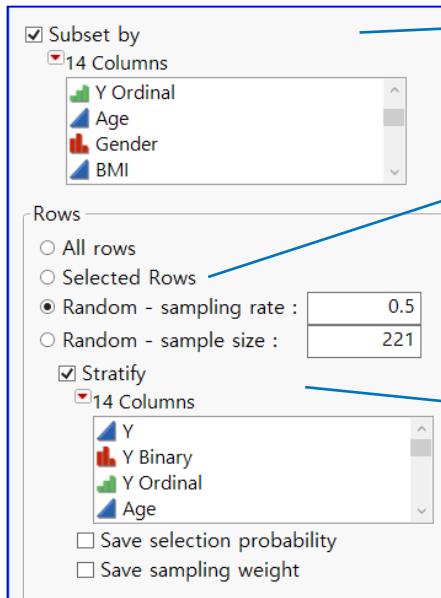
Is the total number of nonmissing values.

4. 보다 실무적으로 Subset(Data 추출) 하는 방법

- 전체 Data에서 일부를 추출하고자 할 경우에는

Sample data : diabetes.jmp

- 일반적으로는 Tables / Subset에서 실행함



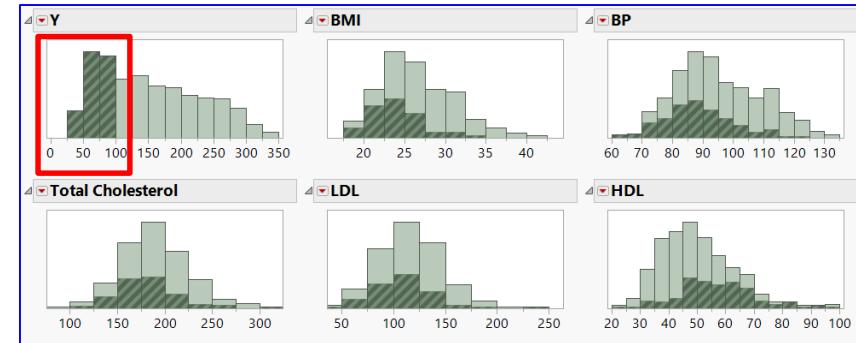
(총별) 변수별로 Subset
만약 'Gender'를 선택하면
남녀별로 분리된 2개의
data table이 생성됨

Data table에서 선택한
rows만 subset
(row panel의 'selected'에서
우측 마우스 클릭 후
data view를 클릭한 결과와
같음

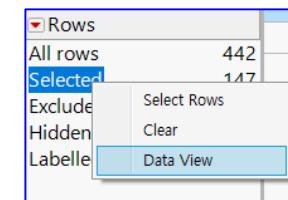
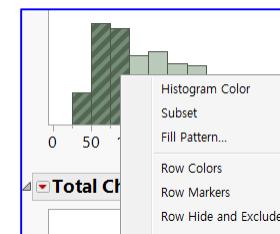
층화(Stratified) 랜덤 샘플링

<보다 실무적인 방법>

- 관심있는 모든 변수에 대해 Histogram을 그림
(Analyze / Distribution)
- 관심있는 변수의 관심 있는 영역을 선택함
(여기서는 반응치 Y의 100 이하 영역)



- Histogram에서 우측 마우스 클릭 후 subset을 선택하거나
row panel의 'selected'에서 우측 마우스 클릭 후 data view를
클릭하면 관심 있는 부분만 추출됨



5. Graph 에서 순차적으로 Hide/Unhide 하는 방법

상황 : Graph(예를 들어 산점도)에서 특정 영역을 단계적으로 선택한 후 영역별로 hide/unhide 등을 할 수 있는 방법

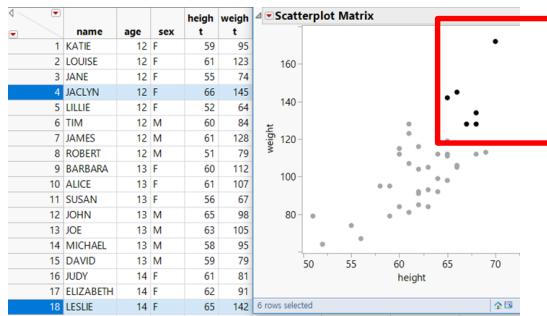
Solution : 두 가지 방법 중 하나로 가능

1. Data table에서 row state를 저장하거나
2. Data filter를 활용하는 방법(좀 더 나은 방법인 듯)

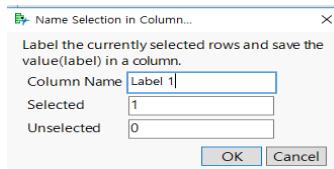
Sample data : Big class.jmp

<방법 1. Data table에서 row state를 저장>

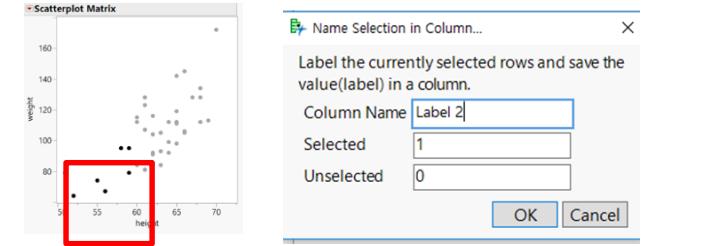
1. Fit Y by X 또는 Graph builder 등의 graph에서 hide 할 특정 영역 선택 후



2. 그레프 위에서 우측 마우스 click / name selection in column에서 hide(select)/unhide(unselect)에 대한 label 값 입력 및 column name 입력



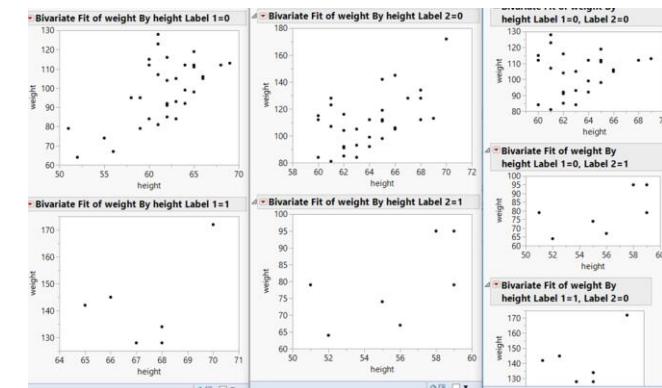
3. 2단계로 hide 할 있다면 위와 동일하게 실행



4. data table에서 아래와 같이 표시됨

	name	age	sex	height	weight	Label 1	Label 2
1	KATIE	12	F	59	95	0	1
2	LOUISE	12	F	61	123	0	0
3	JANE	12	F	55	74	0	1
4	JACLYN	12	F	66	145	1	0
5	LILLIE	12	F	52	64		
6	TIM	12	M	60	84		
7	JAMES	12	M	61	128		
8	ROBERT	12	M	51	79		
9	BARBARA	13	F	60	112		
10	ALICE	13	F	61	107		
11	SUSAN	13	F	56	67		
12	JOHN	13	M	65	98		
13	JOE	13	M	63	105		
14	MICHAEL	13	M	58	95		
15	DAVID	13	M	59	79		
16	JUDY	14	F	61	81		
17	ELIZABETH	14	F	62	91		
18	LESLIE	14	F	65	142		

5. 분석할 menu (여기서는 fit y by x)에 들어가서 by에 label 1 and/or label 2를 입력하면 됨



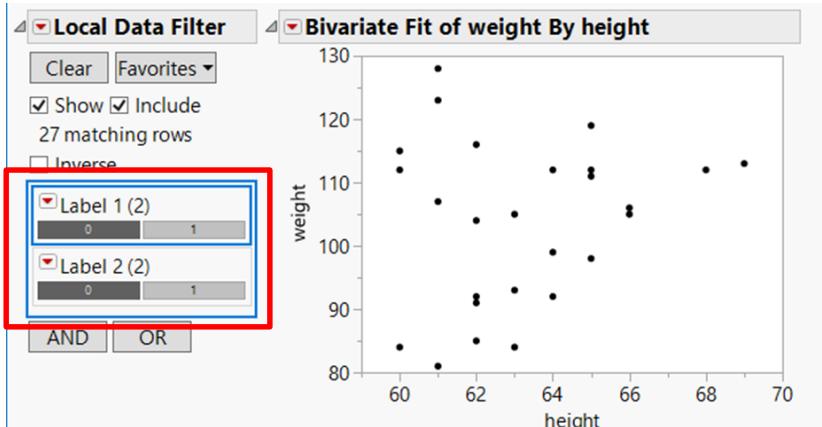
5. Graph 에서 순차적으로 Hide/Unhide 하는 방법

<방법 2. Data filter 를 활용하는 방법>

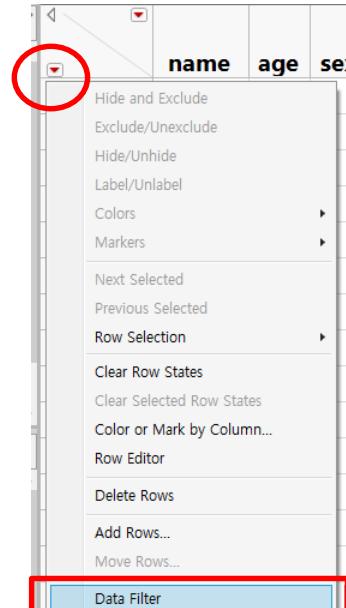
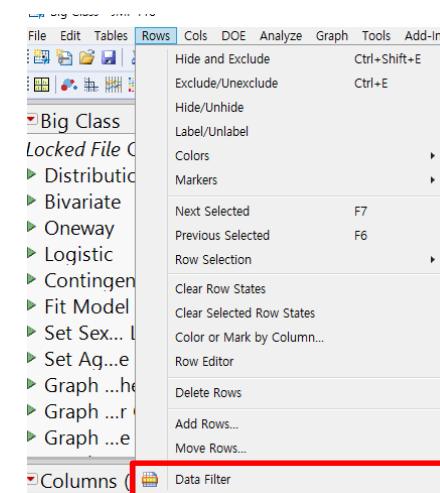
1. 방법 1에서 예로 든 것처럼 fit y by x 에 들어가서 y(weight), x(height) 만 선택 후 OK

Bivariate ~ 왼쪽 옆의 빨간색 역삼각형 click 후
local data filter 에서 label 1, label 2 선택

2. Data filter 창에서 label 1, 2 를 순차적으로 선택 가능
label 이 여러 개 있고, 그것들의 조합하더라도 사용 가능



3. Rows / data filter 에서도 가능



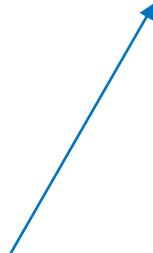
or

6. Variability Chart에서 Axis(축) 표현 방법 차이

Variability Chart에서 Axis(축) 표현 방법의 차이에 대해 살펴 보고자 함

1. 1번(분석 결과를 저장하는 방법)과 같은 Data 가 있다고 가정하고

Operator	Y1	Y2
1 A	35	54
2 A	46	56
3 A	54	35
4 A	64	56
5 A	56	89
6 B	46	54
7 B	47	45
8 B	57	35
9 B	58	45
10 B	47	76



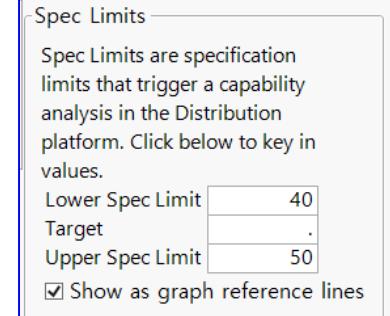
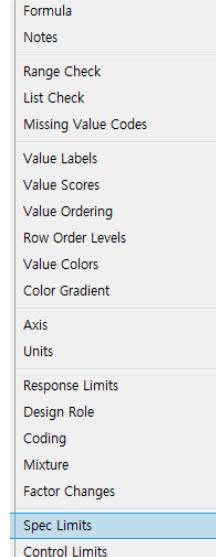
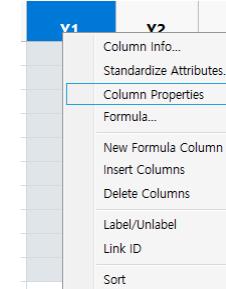
2. Y1, Y2 각각에 Spec 을 입력

(Y1의 Spec 은 40~50, Y2의 Spec 은 10~100)

1) Spec 입력 방법 : Column Name에서 우측 마우스 클릭

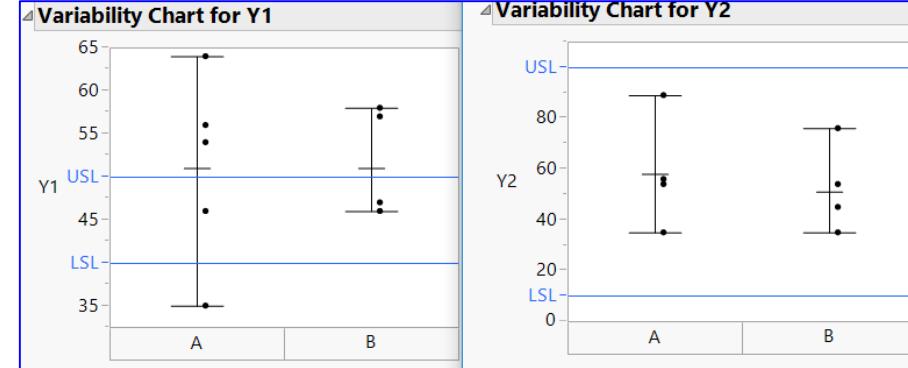
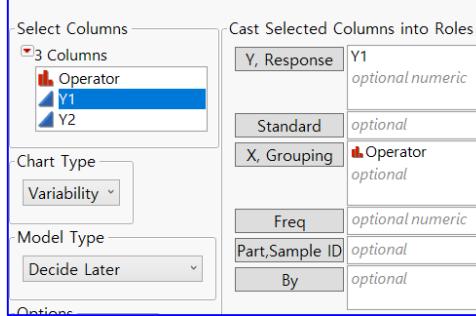
Column Properties / spec limits 클릭

2) Show as graph reference lines 선택



6. Variability Chart에서 Axis(축) 표현 방법 차이

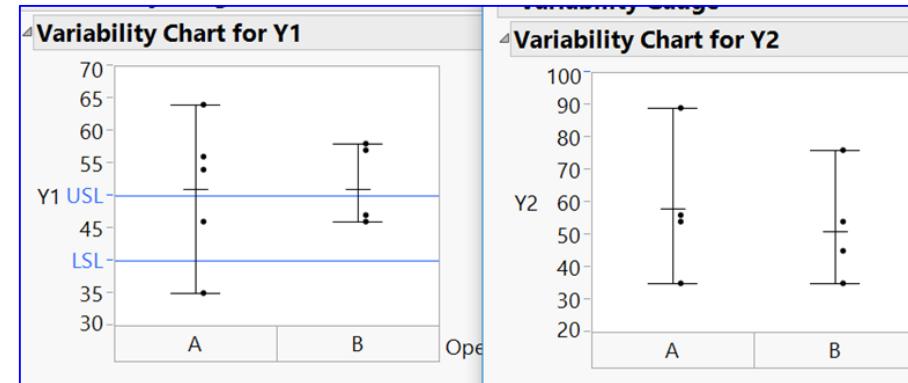
2, Analyze / Quality and Process / Variability / Attribute Gauge Chart에서 아래와 같이 입력(Y1, Y2 모두 실행)



JMP 13

3. JMP 13 Version에서는 Spec 과 실제 Data 범위를 모두 표현할 수 있도록 Y 축의 범위가 표현되지만, JMP 14 Version에서는 Spec 과는 무관하게 실제 data 값을 기준으로 Y 축의 범위가 표시됨을 알 수 있다
-Y2 의 경우 실제 Data 값 대비 Spec 의 범위가 넓으므로 Default 로 표현되는 Y 축에는 Spec 이 표시되지 않는다
-마우스를 이용하여 축(Axis) 범위 조정 가능

4. Graph Builder 등 다른 그래프에서는 Spec 및 실제 Data 범위를 모두 표현할 수 있도록 Display 됨



JMP 14

7. 공정 능력 분석에서 Spec 을 한꺼번에 입력하는 방법

1. 공정 능력 분석을 하기 위해서는

- 1) Data table 에서 해당 변수별로 spec 을 모두 입력하거나
- 2) Spec 만을 입력한 별도의 data table 을 만들어야 하는 데,
여기서는 이 방법을 설명함

Sample data : diabetes.jmp

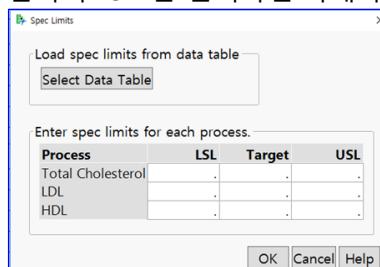
2. Total Cholesterol, LDL, HDL 3가지 변수에 대해
공정능력 분석을 한다고 가정

3. 공정 능력 분석 및 Spec 입력 방법

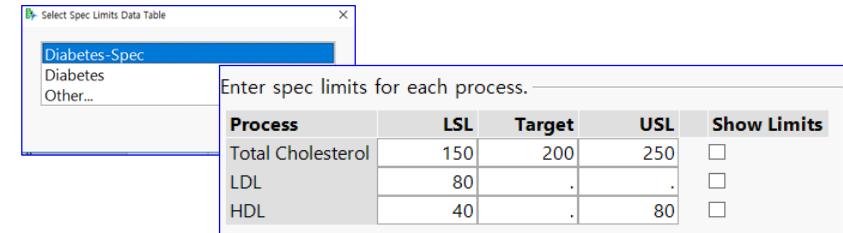
- 1) data file 에서 3개 변수 선택

3P	Total Cholesterol	LDL	HDL	TCh
101	157	93.2	38	
87	183	103.2	70	
93	156	93.6	41	

2) Analyze / Quality and Process / Process Capability 에 들어가서
변수 선택 후 OK 를 클릭하면 아래와 같은 Window 가 Display 됩니다



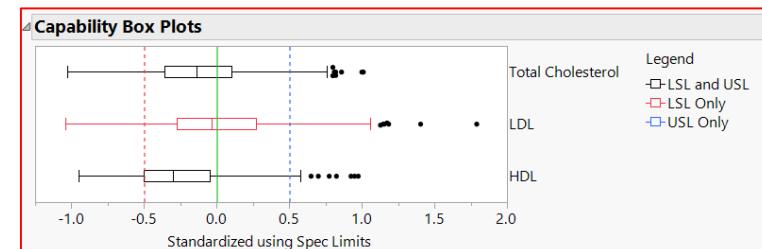
3) Spec 을 직접 입력하거나 select data table 클릭한 후
spec 이 들어있는 파일을 선택



4) Spec 은 양측 Spec 또는 단측 Spec 일 경우 아래와 같이 입력하면 됨

Process	LSL	Target	USL
1 Total Cholesterol	150	200	250
2 LDL	80	.	.
3 HDL	40	.	80

5) 양측 Spec, 단측 Spec 을 구분할 수 있도록 결과가 Display 됨



7. 공정 능력 분석에서 Spec 을 한꺼번에 입력하는 방법

4. 또 다른 방법(JMP 14 Version 에서만 가능)

Analyze / Quality and Process / Manage Spec Limits

1) Analyze / Quality and Process / Manage Spec Limits 에서 아래와 같이 3개 변수 선택 후 OK 클릭하면

SP	Total Cholesterol	LDL	HDL	TCH
101	157	93.2	38	
87	183	103.2	70	
93	156	93.6	41	

2) 아래와 같은 Window 가 Display 됨.

-Spec 을 직접 입력하거나

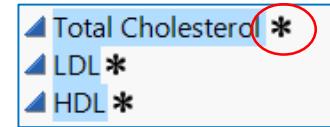
-'Load from limits Table' 을 click 하여 앞 페이지처럼 실행하면 됨

Manage Spec Limits

Column	LSL	Target	USL	Show Limits
Total Cholesterol	150	200	250	<input checked="" type="checkbox"/>
LDL	80	.	.	<input checked="" type="checkbox"/>
HDL	40	.	80	<input checked="" type="checkbox"/>

Show Limits -> Show as Graph Reference Lines

3) Save to Column Properties : Column 속성에 Spec 이 저장되고 Column Panel 에 그 결과가 표시됨



4) Save to Tall Spec Limits Table : Spec 의 별도 Table 에 입력하는 방법

	Variable	LSL	USL	Target	Show Limits
1	Total ...	150	250	200	1
2	LDL	80	.	.	1
3	HDL	40	80	.	1

5) Save to Wide Spec Limits Table : Spec 의 별도 Table 에 입력하는 방법

	Limit	Total Cholesterol	LDL	HDL
1	LSL	150	80	40
2	USL	250	.	80
3	Target	200	.	.

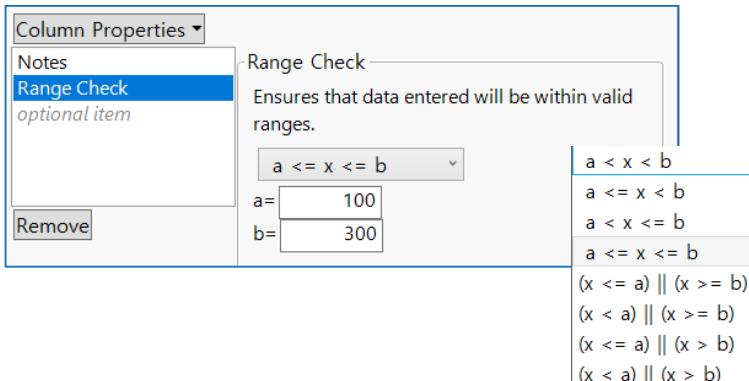
8. 일정 범위 이상의 Data 를 Missing(결측치) 으로 처리하고 싶다

- 일정 범위 이상의 data 를 missing 으로 처리하고 싶을 경우

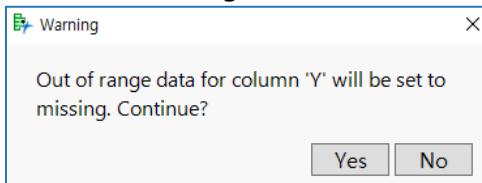
Sample data : diabetes.jmp

- Column(Y) 값 중 100 미만, 300 초과를 Missing 으로 처리하고 하고자 한다면

- 1) column name에서 우측 마우스 클릭, column properties / range check, 아래와 같이 입력



- 선택 범위 바깥이 missing 으로 처리된다는 경고 창이 뜸



- Yes 를 클릭하면 아래처럼 해당 data 가 전부 Missing 으로 처리됨

Y	Y Binary	Y Ordinal
151	Low	Medium
•		
141	Low	Low
206	High	High
135	Low	Low
•		
138	Low	Low
•		
110	Low	Low
•		
101	Low	Low
•		
179	Low	Medium
185	Low	Medium
118	Low	Low

- Y Data 를 이용하여(formula) 만든 Y Binary, Y Ordinal 도 함께 missing 으로 처리됨

- JMP 에서 Missing Value 에 대한 표현 방법

- Numeric : .
- Text : blank(빈 칸)

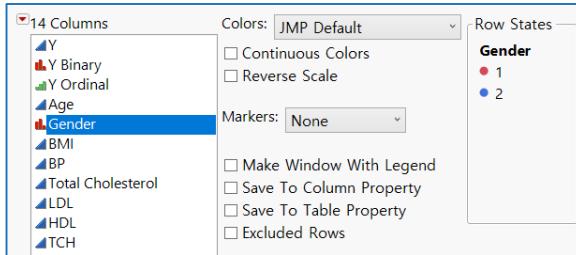
9. 특정 변수를 항상 표현되게 하고 싶다

- 특정 변수(Column) 를, 중요한 층별 변수처럼 어떤 그래프를 그리든, 항상 Display 되게끔 하고자 할 경우

Sample data : diabetes.jmp

- 성별(Gender) 를 항상 표현되게끔 하고 싶을 경우

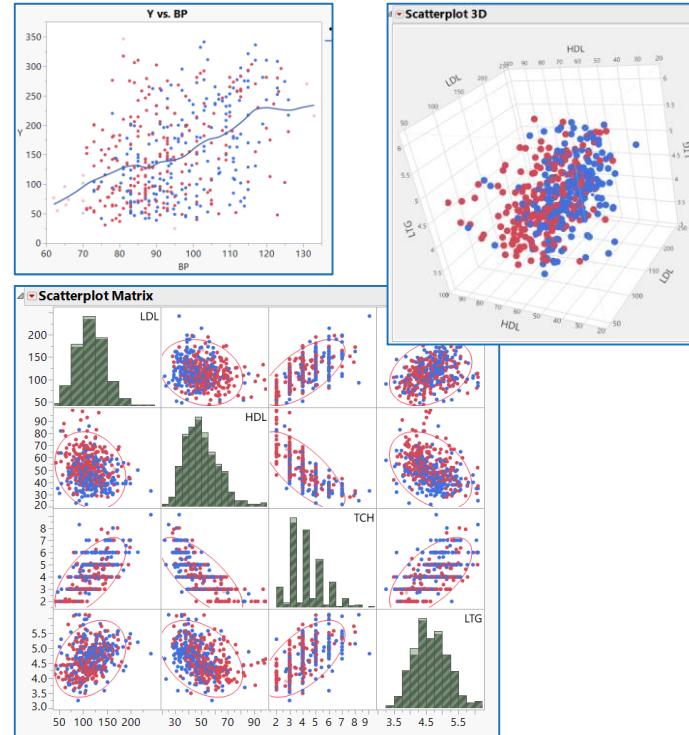
Rows / Color or Mark by Column
아래와 같이 선택 후 OK



- Data Table에서 아래와 같이 표현됨

	Y	Y	Gender	BMI	
	Binary	Ordinal	Age		
•	1	151	Low	Medi...	59 2 32.1
•	2	75	Low	Low	48 1 21.6
•	3	141	Low	Low	72 2 30.5
•	4	206	High	High	24 1 25.3
•	5	135	Low	Low	50 1 23.0
•	6	97	Low	Low	23 1 22.6
•	7	138	Low	Low	36 2 22.0

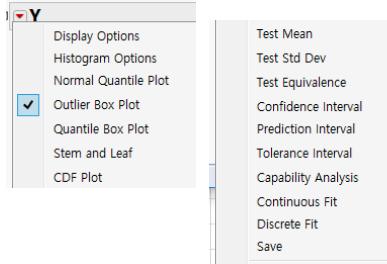
- 어떤 그래프에서도, Gender에 따라 표현됨을 알 수 있음



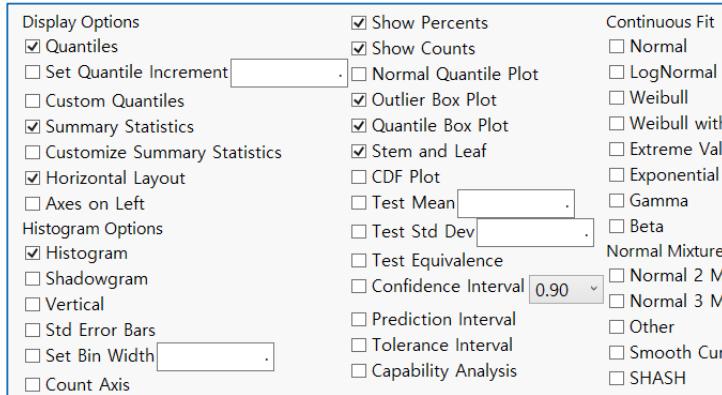
- 해제하는 방법 : rows / clear row states

10. 가장 유용한 단축키 2가지

1. JMP 의 모든 Menu 는 기본 분석 후, 추가 분석을 하기 위해서는 기본 분석 이름의 왼쪽 옆에 있는 빨간색 역삼각형을 클릭하여야 함



2. 실행하고 싶은 추가 분석이 여러 가지일 경우, 하나씩 실행한다면 매우 번거로움. 이럴 경우 **Alt Key** 를 누른 상태에서 빨간색 역삼각형을 클릭하면, 아래와 같이 모든 하위 Menu 가 Display 됨. 필요한 하위 Menu 선택하여 실행하면 됨



3. 가장 중요하고도, 필요한 단축 키는 ?

실행 취소(undo)

Ctrl Z

