

Monthly User Guide from JMP Korea

제 13호 (2018년 8월)

Excel 과 JMP, Excel 보다 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 lkju.shin@jmp.com 으로 연락 바랍니다

1. Excel 과 JMP

JMP 를 설치할 경우 CPU, 언어 지원 등과 함께 'Excel 추가 기능' 을 선택하게 되면, Excel 에서 JMP 로 바로 연결할 수 Platform 이 형성됩니다.

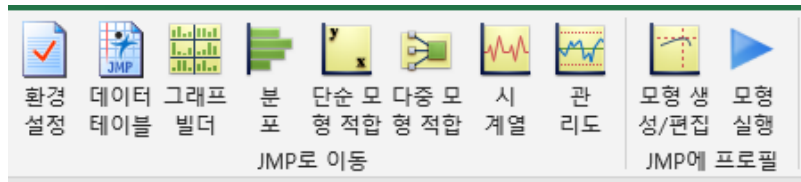
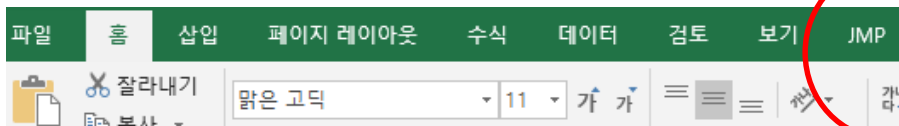
설치 옵션▼

이러한 옵션에 대한 자세한 내용은 www.jmp.com/jmp13install을 참조하십시오.

설치할 기능을 선택하고 설치하지 않을 기능은 선택을 취소하십시오.

- ☐ JMP 32비트 응용 프로그램
- ☒ JMP 64비트 응용 프로그램
- ☒ 언어 지원 파일(필수)
 - ☒ 영어
- ☒ 언어 지원 파일
 - ☐ 프랑스어
 - ☐ 독일어
 - ☐ 이탈리아어
 - ☐ 일본어
 - ☒ 한국어
 - ☐ 중국어 간체
 - ☐ 스페인어
- ☒ Excel 추가 기능
- ☐ SAS 소프트웨어 통합
 - ☐ Java Runtime Environment(필요한 경우)
 - ☐ 핵심 통합 파일(JMPtoJava 파일)

<선택 후 Excel 화면>



* JMP 13 version 이상의 경우 Excel 2016, 2013 및 2010 에서 Available 합니다

2. Excel에서의 JMP

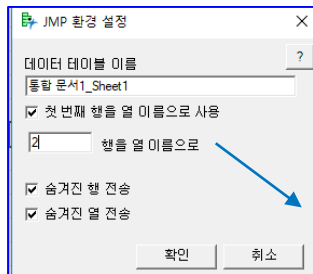
1. Excel에서 아래와 같이 data 입력 후

	A	B
1		
2		Y
3		23
4		13
5		12
6		24
7		32
8		

2. 상단의 'JMP' Click 한 후



3. '환경 설정' 아이콘 클릭



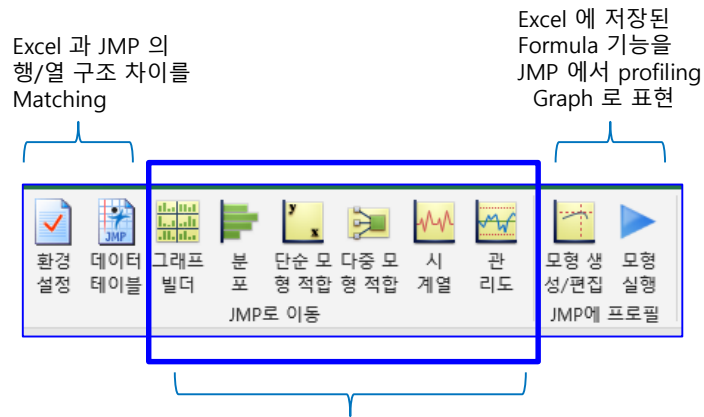
몇 번째 행을
Column name
으로
할 것인지 설정

4. '데이터 테이블' 아이콘 클릭
→ JMP로 바로 연결

	Column 1	Y
1	.	23
2	.	13
3	.	12
4	.	24
5	.	32

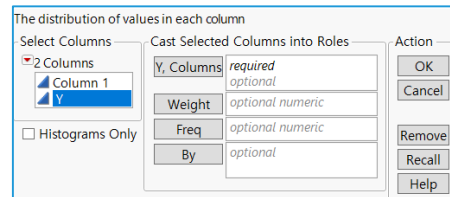
Excel에서
첫 번째 열이
비어 있을 경우
위와 같이 표시됨

5. Menu에 대한 설명



1. Excel 상태에서 자주 사용하는 JMP 기능
2. 해당 아이콘을 click 하면 JMP Data Table과 해당 기능의 window가 display됨
예) '분포' 아이콘을 클릭할 경우

	Column 1	Y
1	.	23
2	.	13
3	.	12
4	.	24
5	.	32



3. JMP 에서의 Excel Import

1. 대부분의 JMP User 들은 Excel 에서 JMP 로 link 하기보다는 JMP 에서 Excel 파일을 불러와서 사용

2. 아래와 같은 Excel Data 가 있다고 가정

	A	B	C	D
1				
2		년도	서울특별시	부산광역시
3		2008 년	24.2	26.1
4		2009 년	24.3	26.5
5		2010 년	23.2	24.6
6		2011 년	22.3	24.3
7		2012 년	22.2	23.5

3. JMP 에서 File / open 에 들어가서 Excel file 을 불러올 수 있음
* 참고 : JMP 에서 import 가능한 data file

All JMP Files (*.jmp;*.sas7bdat;*.xpt;*.stx;*.jst;*.jrn;*.jrp;*.jmx)
Data Files (*.jmp;*.sas7bdat;*.xpt;*.stx)
JMP Files (*.jmp;*.jst;*.jrn;*.jrp;*.jmpproj;*.jmpcust)
JMP Data Tables (*.jmp)
Excel Files (*.xls;*.xlsx;*.xlsm)
Text Files (*.txt;*.csv;*.dat;*.tsv;*.xml)
JMP Scripts (*.jst)
JMP Journals (*.jrn)
JMP Reports (*.jrp)
JMP Projects (*.jmpproj;*.jmpprojarchive)
JMP SQL Queries (*.jmpquery)
JMP Add-In Files (*.jmpaddin;*.jmpaddindef;*.def)
JMP Menu Files (*.jmpcust)
JMP Application Files (*.jmpappsource;*.jmpapp)
SAS Data Sets (*.sas7bdat;*.xpt;*.stx)
SAS Program Files (*.sas)
R Code (*.r)
MATLAB Code (*.m)
Other Source Code (*.py;*.c;*.cpp;*.sql;*.js)
HTML Files (*.htm;*.html)
FACS Files (*.fcs)
SPSS Data Files (*.sav)
Triple-S Survey Files (*.sss;*.xml)
JSON Data File (*.json)
xBase Data Files (*.dbf)
Shapefiles (*.shp)
Teradata database (*.trd)
HDFS Data File (*.h5)
SQLite database (*.sqlite;*.db;*.sqlite3;*.db3)
Minitab Portable Worksheet Files (*.mtp)

4. Excel file 을 open 하면 JMP Data file 이 바로 열리는 것이 아니라 아래와 같이 Excel 과 JMP 간의 행, 열 및 열 이름 차이 등을 조정 하기 위한 'Excel Import Wizard' 창이 열림

The dialog box 'Excel Import Wizard' is shown. The 'Data Preview' section displays a table with columns: '년도', '서울특별시', '부산광역시', '대구광역시', '인천광역시', and '광주광역시'. The 'Worksheets' section has 'Select sheets to open' and 'Custom setting' buttons. The 'Individual Worksheet Settings' section is highlighted with a red box and contains the following options:

- ☒ Worksheet contains column headers
- Column headers start on row: 1
- Number of rows with column headers: 1
- Data starts on row: 2
- Data starts on column: 1
- ☐ Concatenate worksheets and try to match columns
- ☐ Create column with worksheet name when concatenating
- ☒ Use for all worksheets

Buttons at the bottom include 'Restore Default Settings', 'Back', 'Next', 'Import', 'Cancel', and 'Help'.

6. 조정 후 'Import' 를 click 하면 아래와 같이 JMP Data Table 이 생성됨

	년도	서울특별시	부산광역시	대구광역시	인천광역시
1	2008 년	24.2	26.1	25	
2	2009 년	24.3	26.5	25.6	
3	2010 년	23.2	24.6	24.5	
4	2011 년	22.3	24.3	24.4	
5	2012 년	22.2	23.5	24.5	
6	2013 년	21.7	23.1	23.7	
7	2014 년	20.6	22.3	22.9	
8	2015 년	19.4	20.6	20.4	
9	2016 년	19.5	20.5	19.6	

5. 'individual Worksheet Settings' 에서 Excel 과 JMP 의 차이를 조정

* 네 가지 Option 의 뜻은 다음과 같음

- 1) Excel 의 몇 번째 Row 가 Column Header ?
- 2) 몇 개의 Row 가 Column Name ?
- 3) 몇 번째 Row 부터 Data ?
- 4) 몇 번째 Column 부터 Data ?

4. Excel 과 JMP 차이

Excel 과 JMP 는 다음과 같은 차이가 있습니다(Data Table 의 구조 측면)

Table Panel

Column Panel

Row Panel

Column Name

name	age	sex	height	weight
1 KATHY	12	F	59	99
2 LOUISE	12	F	61	123
3 JANE	12	F	55	74
4 JACLYN	12	F	66	145
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
19 CAROL	14	F	63	84
20 PATTY	14	F	62	85
21 FREDERICK	14	M	63	93
22 ALFRED	14	M	64	99

Data Grid

Items	Excel	JMP
Formula	개별 Cell 단위로 Formula 생성	Column 단위로 Formula 생성
Column Names	Column Name 은 Data Grid* 의 한 부분 Modeling Type 이 달라도 같은 Column 에 저장 가능	Column Name 은 Data Grid 의 부분이 아님 숫자(Numeric Data)와 글자(Character)가 함께 있으면 해당 Column 을 Character Data 로 인식
Tables & Worksheets	하나의 Spreadsheet 에 여러 개의 Table 및 Worksheet 포함 가능	JMP 는 Worksheet 의 개념이 없음 각각의 Data Table 은 XXX.jmp file 로 별도로 저장
Data Grid	Data 는 Data Grid 의 아무 곳이나 위치 가능	Data 는 Row 1 & Column 1 부터 시작
Analysis & Graph Reports	분석 결과 및 그래프는 Data Grid 안에 위치	분석 결과 및 그래프는 별도의 윈도우에 표시됨



5. Excel 보다 JMP

Excel 과 비교될 수 있는 JMP 의 몇 가지 기능을 살펴 보겠습니다.

1. 쉽게 표 만들기 및 통계량 계산하기
2. Formula
3. Categorization
4. 특정 Row 를 숨기거나 제외하는 기능

5. Excel 보다 JMP

1) 쉽게 표 만들기 및 통계량 계산하기

1. Sample data : big class.jmp

	name	age	sex	height	weight
1	KATIE	12	F	59	95
2	LOUISE	12	F	61	123
3	JANE	12	F	55	74
4	JACLYN	12	F	66	145

2. 분석(표 만들고자 하는) 하고자 하는 내용
: 성별, 나이별로 키와 몸무게의 평균, 표준편차,
최소값, 최대값에 대한 표

3. Analyze / Tabulate

Recall | Start Over | Done

5 Columns

name

age

sex

height

weight

Freq:

Weight:

Page Column:

N

Mean

Std Dev

Min

Max

Range

% of Total

N Missing

N Categories

Sum

Sum Wgt

Variance

Std Err

Drop zone for columns

Drop zone for rows

Resulting Cells

4. 변수(Sex, Age) 를 순차적으로 선택하여
'Drop Zone for Rows' 에 Drop

- 1) Age 선택 후 Drop, 그 다음 Sex 를 Drag 하여
a) 처럼 Drop 하면 변수 추가가 아닌 변수 변경,
b) 처럼 해야 변수 추가.

a)

age	N
12	8
13	7
14	12
15	7
16	3
17	3

c) 변수 추가 결과

sex	age	N
F	12	5
	13	3
	14	5
	15	2
	16	2
	17	1
M	12	3
	13	4
	14	7
	15	5
	16	1
	17	2

b)

age	N
12	8
13	7
14	12
15	7
16	3
17	3

5. 변수(Height, Weight) 를 선택한 다음 'N' 의
위치에 Drop.
구하고자 하는 통계량 선택 후 'Sum' 위치에 Drop

sex	age	height	weight
F	12	293	501
	13	177	286
	14	313	483
	15	126	204
	16	125	227
	17	62	116
M	12	172	291
	13	245	377
	14	457	727
	15	326	554
	16	68	128
	17	138	306

Mean
Std Dev
Min
Max

6. 소계(Sub Total) 및 합계 등을 구하고자 한다면
'Add Aggregate Statistics' 선택하면 아래와 같이 표가 완성됨

sex	age	height				weight			
		Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
F	12	58.6	5.413	52	66	100.2	33.73	64	145
	13	59.0	2.646	56	61	95.3	24.66	67	112
	14	62.6	1.517	61	65	96.6	25.64	81	142
	15	63.0	1.414	62	64	102.0	14.14	92	112
	16	62.5	3.536	60	65	113.5	2.121	112	115
	17	62.0	.	62	62	116.0	.	116	116
M	All	60.9	3.612	52	66	100.9	23.44	64	145
	12	57.3	5.508	51	61	97.0	26.96	79	128
	13	61.3	3.304	58	65	94.3	11	79	105
	14	65.3	2.289	63	69	103.9	10.68	92	119
	15	65.2	1.924	62	67	110.8	9.985	104	128
	16	68.0	.	68	68	128.0	.	128	128
	17	69.0	1.414	68	70	153.0	26.87	134	172
	All	63.9	4.308	51	70	108.3	21.1	79	172
All	All	62.6	4.242	51	70	105.0	22.2	64	172

7. 만들어진 표를 Excel 로 보내기 위해서는

- 1) ▼ Tabulate / make into data table 에 들어가서
data table 을 만든 다음
- 2) file / save as 에 들어가서 excel file 로 저장하면 됨

1. JMP 에도 Excel 처럼 강력한 formula 기능이 있음

2. Excel 과 JMP 차이점

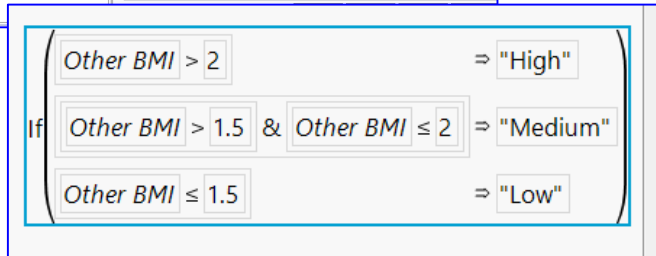
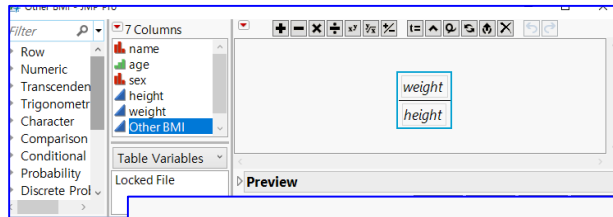
- 1) Excel : 개별 Cell 단위로 Formula 생성
- 2) JMP : Column 단위로 Formula 생성

3. 분석하고자 하는 내용(예시)

- 1) Sample file : [big class.jmp](#)
- 2) 비만도를 확인하기 위해
 - new column (Weight / Height) 를 계산하고 column name 에 'Other BMI' 입력
 - 또 다른 new column 에 'Other BMI' 값을 기준으로 High, Medium, Low 로 범주화

4. Other BMI 계산

- 1) cols / new column 또는 data table 의 new column 에서 우측 마우스 클릭한 후 column properties / formula 선택
column info 에 'Other BMI' 입력, formula 창에 Weight / Height 입력
- 2) 4-1) 과 new column 생성한 후, formula 창에 우측과 같이 입력 (column 명 : Categorization)



5. 결과

- 1) Formula 를 이용하여 new column 을 만든 결과
- 2) Formula 가 들어있을 경우 왼쪽 column table 의 column 명칭에 + (selection mark) 생성됨.
Selection mark 를 click 하면 해당 formula 창이 display 됨

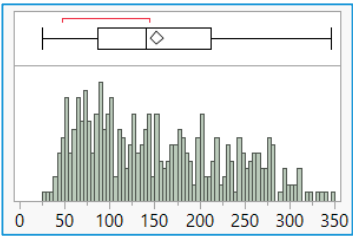
	name	age	sex	height	weight	Other BMI	Categorization
1	KATIE	12	F	59	95	1.6102	Medium
2	LOUISE	12	F	61	123	2.0164	High
3	JANE	12	F	55	74	1.3455	Low
4	JACLYN	12	F	66	145	2.1970	High
5	LILLIE	12	F	52	64	1.2308	Low
6	TIM	12	M	60	84	1.4000	Low
7	JAMES	12	M	61	128	2.0984	High
8	ROBERT	12	M	51	79	1.5490	Medium
9	BARBARA	13	F	60	112	1.8667	Medium
10	ALICE	13	F	61	107	1.7541	Medium

6. Formula 에 대한 세부적인 설명은 Help / books / 'Using JMP' 의 formula editor(14 version 기준 349 page ~) 부분 참조.

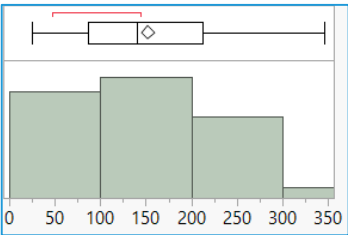
- 1. Categorization(범주화, Discretization)
 - 1) 연속형 data 를 이산형 data 로 만드는 것
 - 2) Continuous → Nominal or Ordinal
 - 3) Grouping 의 개념
- 2. Categorization 의 예 : 앞 페이지의 4 번
- 3. Categorization 사유
 - 1) Y(특히, 연속형 Y) 에 영향을 미치는 변수를 좀 더 쉽게 Screening 하고자 할 때
 - 2) 좀 더 단순한 통계 분석 방법을 사용하기 위해
 - 3) 1) 및 2) 의 사유로 Big Data 등을 다루는 Data Mining 분야에서 많이 활용
- 4. Categorization 방법
 - : 일반적으로 Excel 또는 분석 소프트웨어에서 제공하는 formula 기능 활용
- 5. JMP 에서도 Formula 기능을 이용하여 Categorization 가능하지만 **JMP 의 상호작용(Interaction) 기능을 이용하여 Graphical 한 측면에서 Categorization 하는 방법을 추천함**
- 6. sample file : diabetes.jmp

- 7. Analyze / Distribution 에 들어가서 Y 및 age ~ glucose 까지 선택, OK
- 8. 도구 모음의 Grabber 기능을 이용
Grabber 클릭 후 Y Data 에 대한 Histogram 위에서 마우스 왼쪽 눌러서 위/아래로 이동

1) 위로 이동했을 경우

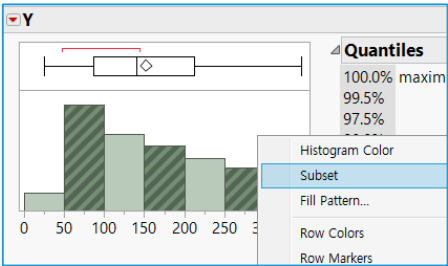


2) 아래로 이동했을 경우

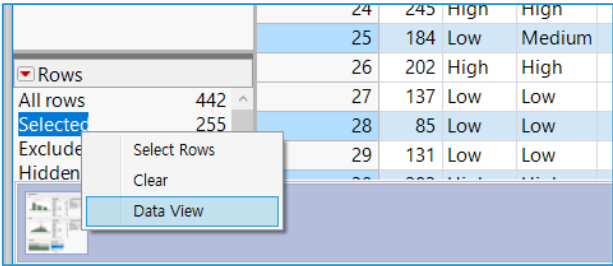


- 9. 8번의 방법을 활용하여 Histogram 의 Bin(막대 그래프) 개수를 일정 개수로 만든 다음, 제일 값이 큰 그래프와 작은 그래프 선택 후 (또는 큰 값, 중간 값 및 작은 값에 해당하는 그래프를 각각 선택 후) 해당 data 만 별도로 추출하여 분석할 수 있음

1) 히스토그램 위에서 마우스 우측 Click → Subset



2) Row Panel 의 Selected 에서 Data View Click



5. Excel 보다 JMP

4) 특정한 Row 를 숨기거나 제외하는 기능

- 1. JMP 에서는 특정한 Row 를 숨기거나 제외한 후 그래프를 그리거나 통계량 계산을 Excel 보다 쉽게 할 수 있습니다.
- 2. 특정 Row 를 선택 후 오른쪽 마우스를 click 하면 아래와 같습니다.
이 중에서 많이 사용하는 Hide, Exclude & Label 활용법에 대해 알아 보겠습니다

Hide and Exclude

Exclude/Unexclude

Hide/Unhide

Label/Unlabel

Colors

Markers

Color Rows by Row State

Select Matching Cells

Invert Selection

Clear Row States

Add Rows...

Delete Rows

3. 설명

구분	설명	Icon
Exclude	말 그대로 분석에서 제외하는 기능	
Hide	-원래 의미는 그래프에서 안 보이게 하는 기능 -통계량과 함께 표현될 경우에는 Display 됨 (예 : Histogram)	
Label	Graph 등으로 표현 시 Label (원래 Data) 를 Display	

4. 아래와 같이 Data 입력 후 Row 선택 → 오른쪽 마우스 Click 하여 Row 상태 설정

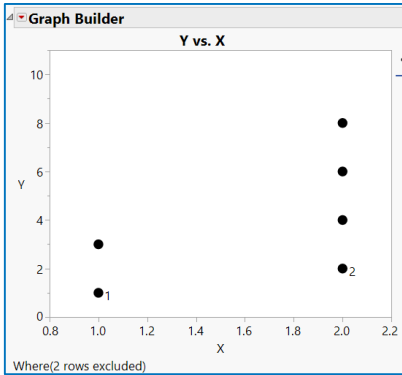
Hide Exclud...						
			Y	X	A	
			1	1	1	labeled
Columns (3/0)			2	2	2	labeled
Y			3	3	1	
X			4	4	2	
A			5	5	1	hidden,exclude
Rows			6	6	2	
All rows	10		7	7	1	excluded
Selected	1		8	8	2	
Excluded	2		9	9	1	hidden
Hidden	3		10	10	2	hidden
Labelled	2					

5. Excel 보다 JMP

4) 특정한 Row 를 숨기거나 제외하는 기능

5. Graph / Graph Builder

- 1) Row 1 & Row 2 : labeled
- 2) Exclude 및 Hide 된 4개를 제외하고 display
- 3) Exclude 된 2개는 Exclude 되었다고 표시



		Y	X	A
Hide Excl...		1	1	1 labeled
Columns (3/0)		2	2	2 labeled
Y		3	3	1
X		4	4	2
A		5	5	1 hidden, exclude
Rows		6	6	2
All rows	10	7	7	1 excluded
Selected	1	8	8	2
Excluded	2	9	9	1 hidden
Hidden	3	10	10	2 hidden
Labeled	2			

6. Analyze / Tabulate

- 1) Exclude 된 2개를 제외하고 합계 계산
- 2) A Column 포함 시 Missing 된 4개의 Row 및 Exclude 된 2개 Row 제외하고 계산

Y	Sum	43
X	Sum	13

2 rows have been excluded.

		A	
		hidden	labeled
Y	Sum	19	3
X	Sum	3	3

2 rows have been excluded.

7. Analyze / Distribution

- 1) Exclude 된 2개 row 제외하고 Histogram Plotting (Graph Builder 처럼 Hide Row 를 숨기지 않음)
- 2) Exclude 된 2개 row 제외하고 통계량이 계산됨

