

Rotman

QUICK INTRO TO JMP

Tutorial / <https://tdmdal.github.io/jmp-intro-2021/>

November 17, 2021 Prepared by Jay / [TDMDAL](#)



Rotman School of Management
UNIVERSITY OF TORONTO

Goal for Today – Get You Started with JMP

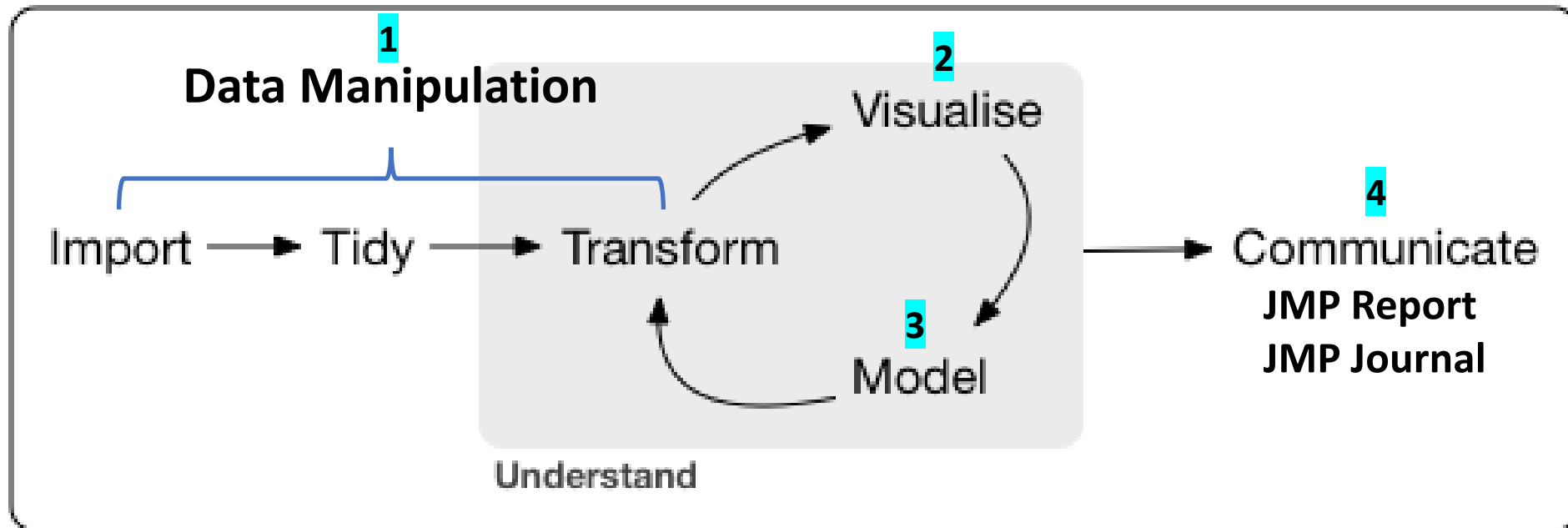
- What is JMP
- JMP basics (demos & learning by doing)
 - JMP basic navigations
 - Open/Import datasets
 - Manipulate data: a few simple tasks
 - Analyze data: a simple linear regression model
- Learning resources

What is JMP

- Predictive analytics software from SAS Institute
 - data manipulation
 - [visualization](#)
 - statistical & machine learning modeling
 - reporting
- Intuitive beginner-friendly **point-and-click interface**
 - Analytics with zero/minimum coding
- Flexible and extensible for advanced users
 - JMP Scripting Language (JSL) for automating or extending point-and-click functionality
 - Connect to the richness of SAS: retrieve SAS data and submit SAS code
 - Connect to Database engine, Matlab, R, Python, Excel, Web API, etc.

See key features of JMP Pro at [JMP Pro website](#).

A Typical Workflow

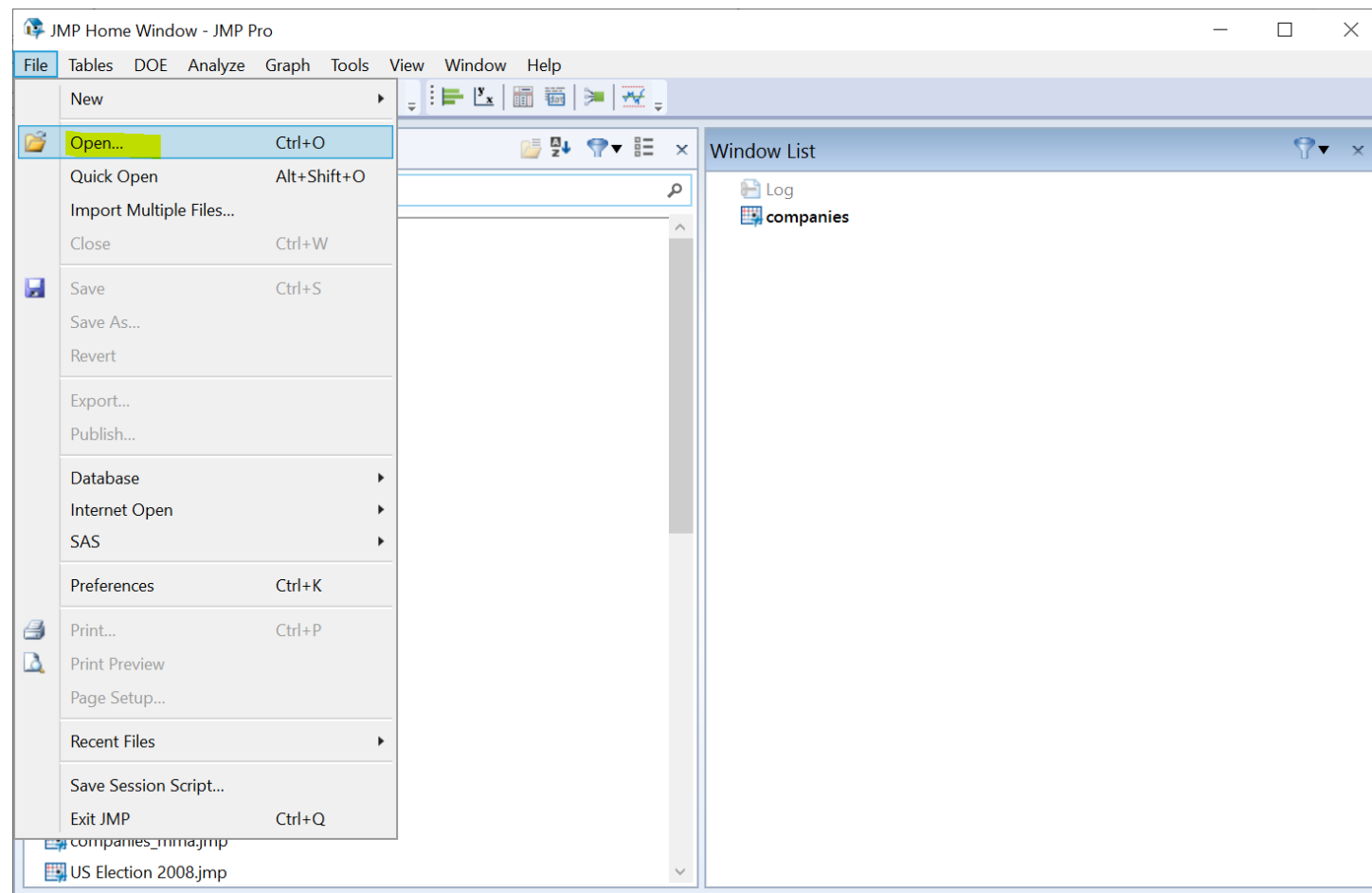


point-and-click & programming

Concept and graph adapted from <https://r4ds.had.co.nz/introduction.html>

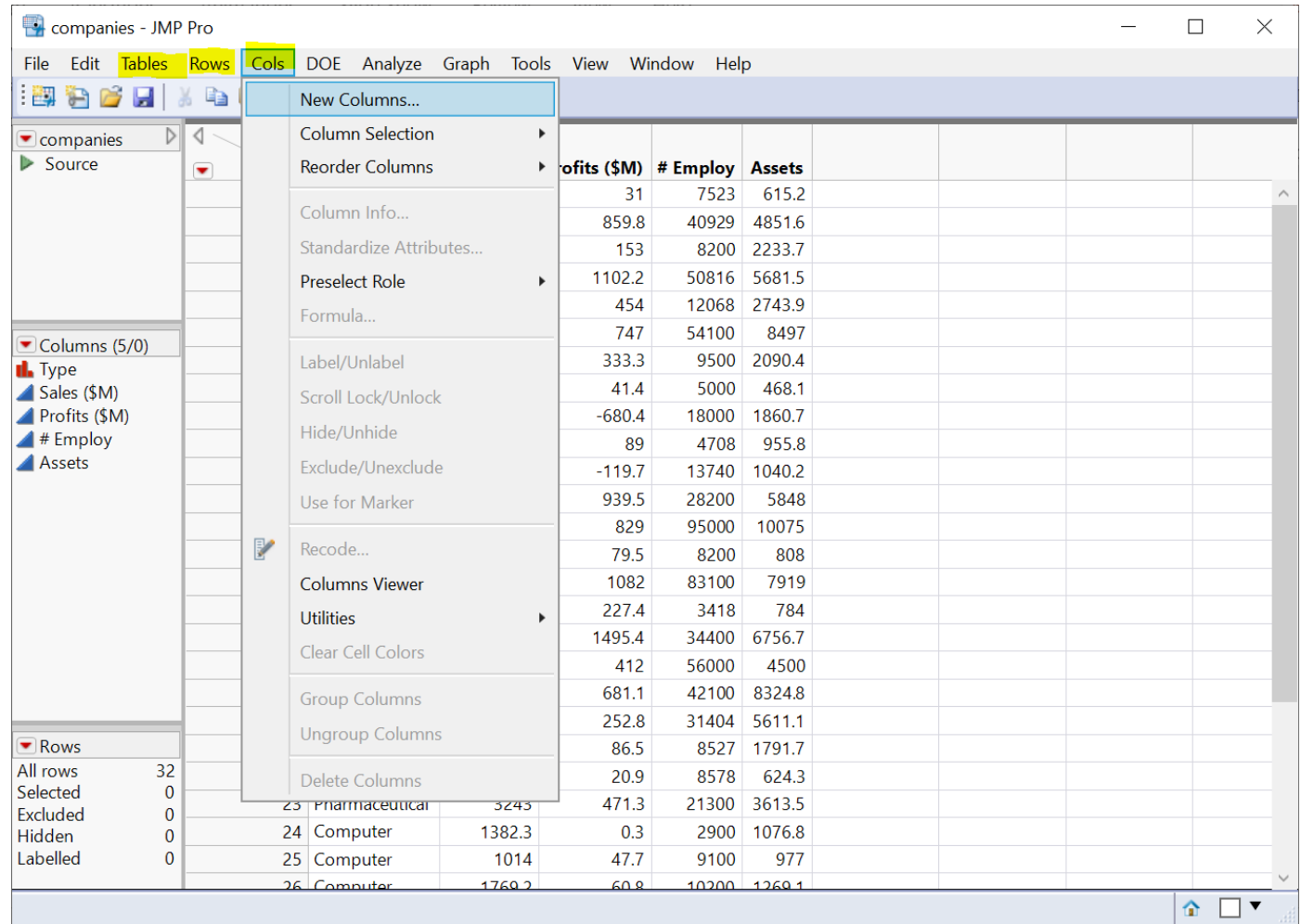
JMP Navigations: Open / Import Data

- Open native JMP data file
 - .jmp
- Import data files in other format
 - .csv files
 - Excel files
 - ... many more



JMP Navigations: Data Manipulation

- Column operations
 - E.g., create a new col
- Row operations
 - E.g., filter rows
- Table operations
 - E.g., merge two tables



JMP Navigations: Modeling

- Crosstab analysis
 - Simple: Fit Y by X
 - Sophisticated (e.g. multiple responses, nested X): Consumer Research -> Categorical
- Regressions
 - Simple: Fit Y by X
 - Multiple: Fit Model
- K-means clustering
 - Clustering -> K Means Cluster
- Multidimensional Scaling
 - Multivariate Methods -> Multidimensional Scaling

The screenshot shows the JMP Pro interface with the 'Analyze' menu open. The menu options include: Distribution, Fit Y by X, Tabulate, Text Explorer, Fit Model, Predictive Modeling, Specialized Modeling, Screening, Multivariate Methods, Clustering, Quality and Process, Reliability and Survival, and Consumer Research. The background data table is as follows:

			# Employ	Assets					
1	Com		7523	615.2					
2	Pharm		40929	4851.6					
3	Com		8200	2233.7					
4	Pharm		50816	5681.5					
5	Com		12068	2743.9					
6	Pharm		54100	8497					
7	Com		9500	2090.4					
8	Com		5000	468.1					
9	Com		18000	1860.7					
10	Com		4708	955.8					
11	Com		13740	1040.2					
12	Pharm		28200	5848					
13	Com		95000	10075					
14	Com		8200	808					
15	Pharmaceutical	5044	1002	83100	7919				
16	Pharmaceutical	969.2	227.4	3418	784				
17	Pharmaceutical	6698.4	1495.4	34400	6756.7				
18	Computer	5956	412	56000	4500				
19	Pharmaceutical	5903.7	681.1	42100	8324.8				
20	Computer	2959.3	252.8	31404	5611.1				
21	Pharmaceutical	1198.3	86.5	8527	1791.7				
22	Computer	990.5	20.9	8578	624.3				
23	Pharmaceutical	3243	471.3	21300	3613.5				
24	Computer	1382.3	0.3	2900	1076.8				
25	Computer	1014	47.7	9100	977				
26	Computer	1769.2	60.8	10200	1269.1				

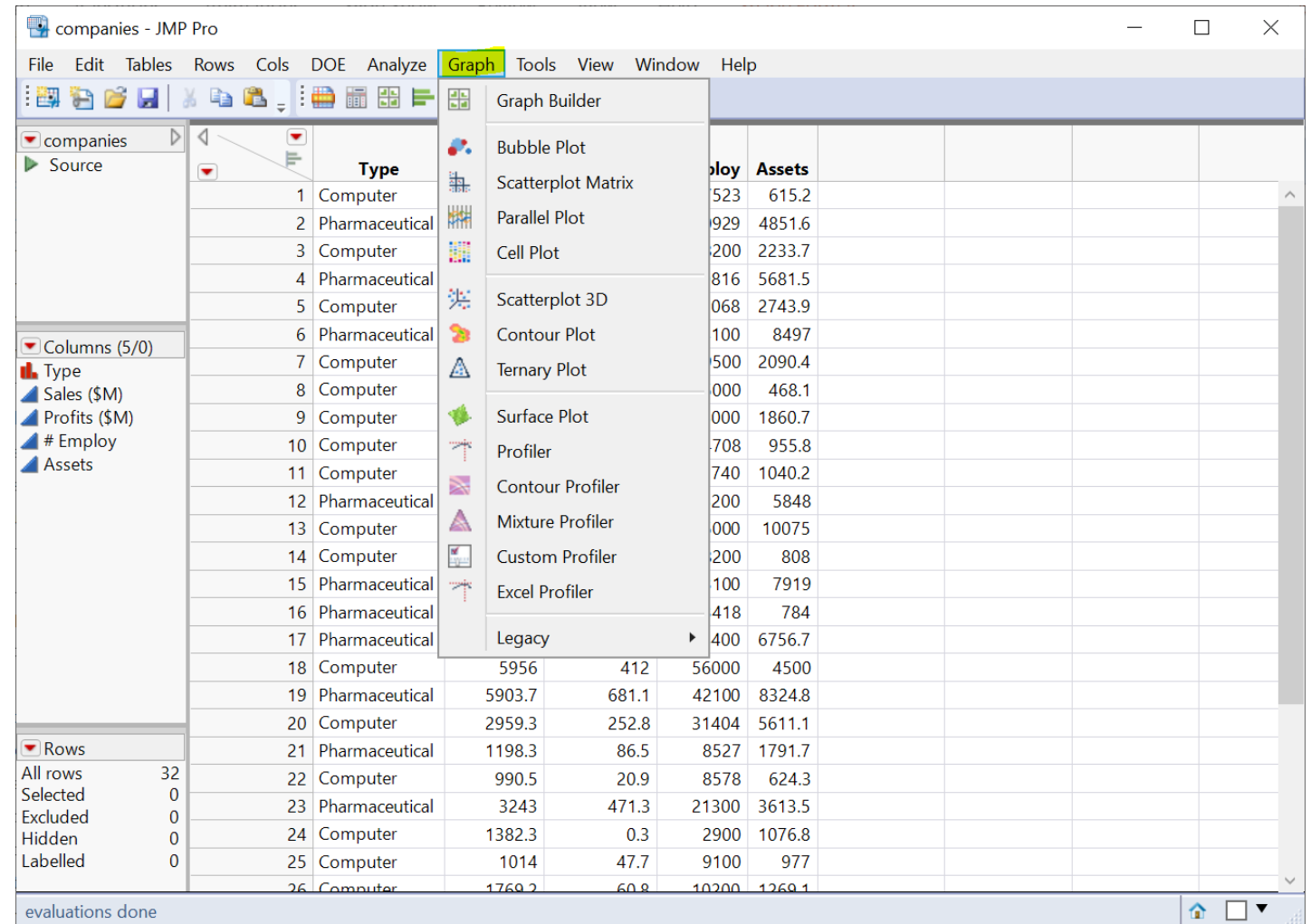
JMP Navigations: Visualization & Reporting

- Visualization

- Analysis graphs (comes with modeling)
- Specialized graphs

- Reporting

- Analysis reports (comes with modeling)
- JMP journal for presentation



JMP Navigations: Demo

- A simple example
 - import data
 - fit y by x: Profit by Sales per employee (a simple linear model)
- Data (companies.csv)

```
Type, Sales ($M), Profits ($M), # Employ, Assets
Computer, 855.1, 31.0, 7523, 615.2
Pharmaceutical, 5453.5, 859.8, 40929, 4851.6
Computer, 2153.7, 153.0, 8200, 2233.7
...
```

Your Turn (Hands-on; 10 mins)

- Repeat the demo I just did
 - import the dataset to a JMP data table
 - fit y by x (Profit by Sales per employee)
 - save the analysis script in the data table
- Challenge: run a multiple linear regression
 - predict Profits using two variables: Sales and Size
 - Size is a categorical variable defined as
 - if # Employ < 10000, then size = “small”
 - Otherwise, Size = “large”

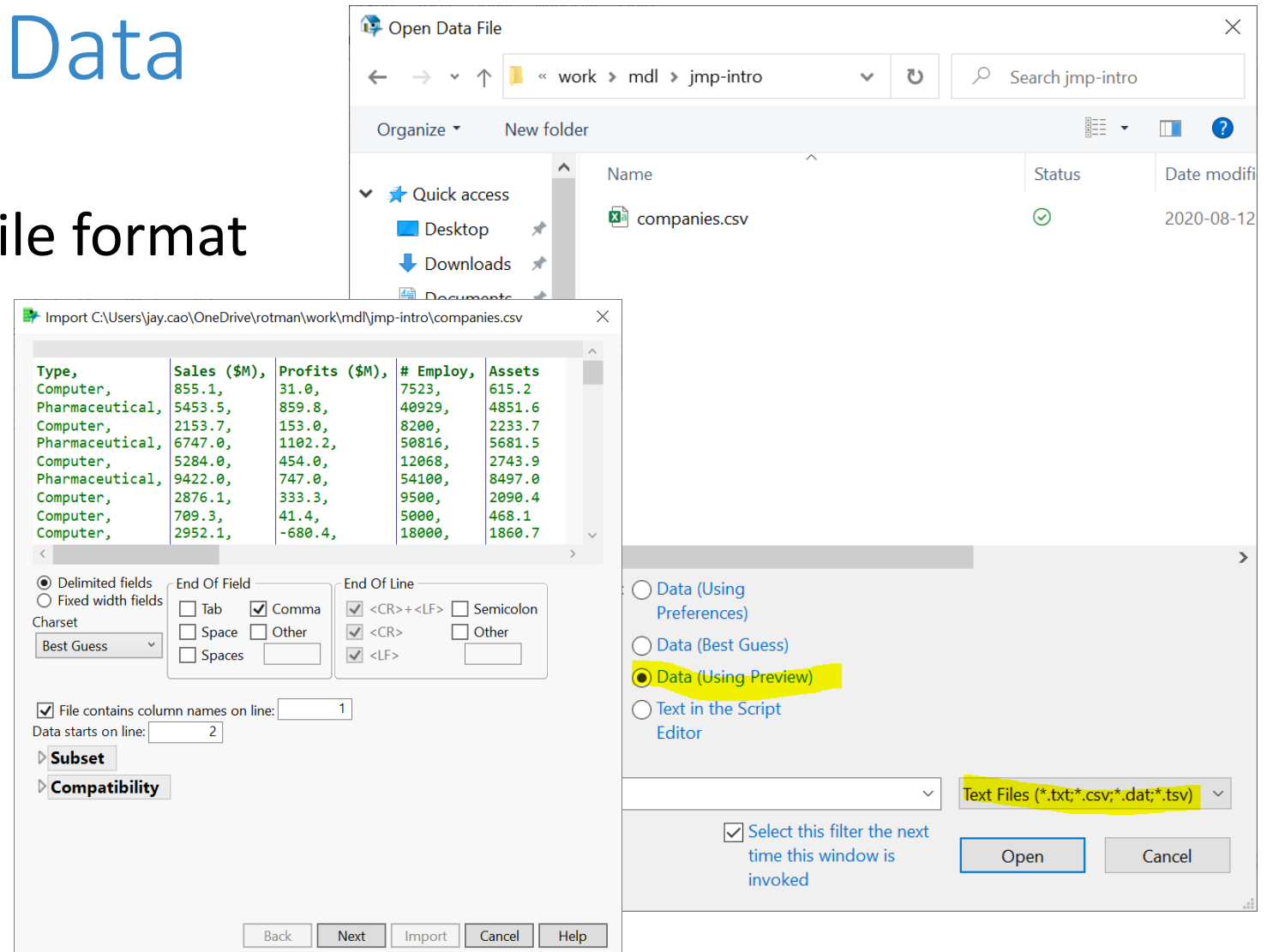
Review: Import Data

- Import support many file format

- csv
- Excel
- json
- many more

- JMP native data file

- .jmp



Review: Data Table

- Three panels on the left
 - Table (JMP Script)
 - Columns
 - Rows
- Column Info

The screenshot displays the JMP Pro interface. The main window shows a data table with columns: Type, Sales (\$M), Profits (\$M), # Employee, and Assets. The # Employee column is selected. A dialog box titled '# Employee - JMP Pro' is open, showing the column's properties. The dialog includes fields for Column Name, Data Type (set to Numeric), Modeling Type (set to Continuous), and Format (set to Best with a width of 12). There are also checkboxes for Lock and Use thousands separator (,).

	Type	Sales (\$M)	Profits (\$M)	# Employee	Assets
1	Computer	855.1	31	7523	615.2
2	Pharmaceutical	5453.5	859.8	40929	4851.6
3	Computer	2153.7	153	8200	2233.7
4	Pharmaceutical	6747	1102.2	50816	5681.5
5	Computer	5284	454	12068	2743.9
6	Pharmaceutical				
7	Computer				
8	Computer				
9	Computer				
10	Computer				
11	Computer				
12	Pharmaceutical				
13	Computer				
14	Computer				
15	Pharmaceutical				
16	Pharmaceutical				
17	Pharmaceutical				
18	Computer				
19	Pharmaceutical				

Employee - JMP Pro

'# Employee' in table 'companies'

Column Name: # Employee

Lock

Data Type: Numeric

Modeling Type: Continuous

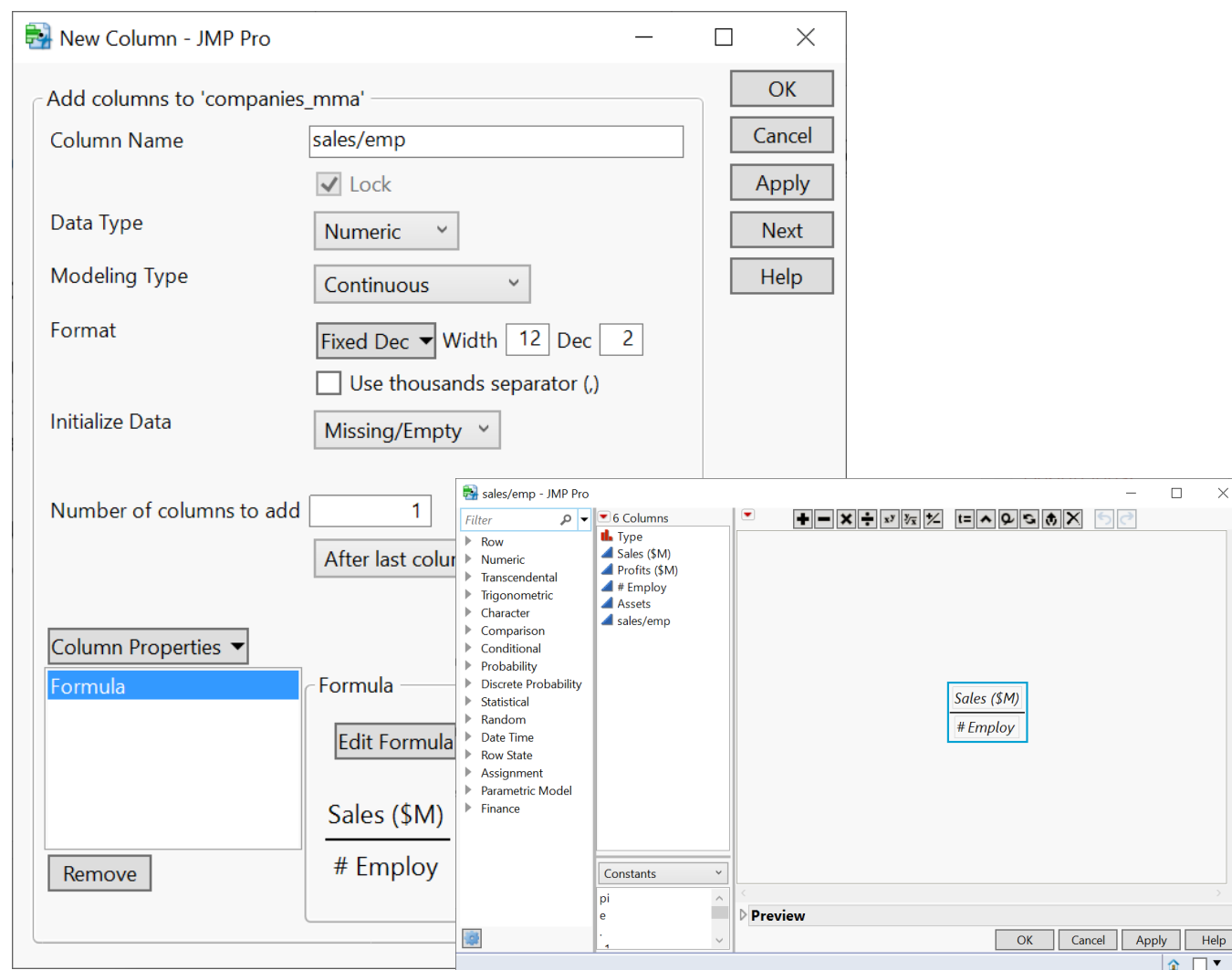
Format: Best Width: 12

Use thousands separator (,)

Column Properties

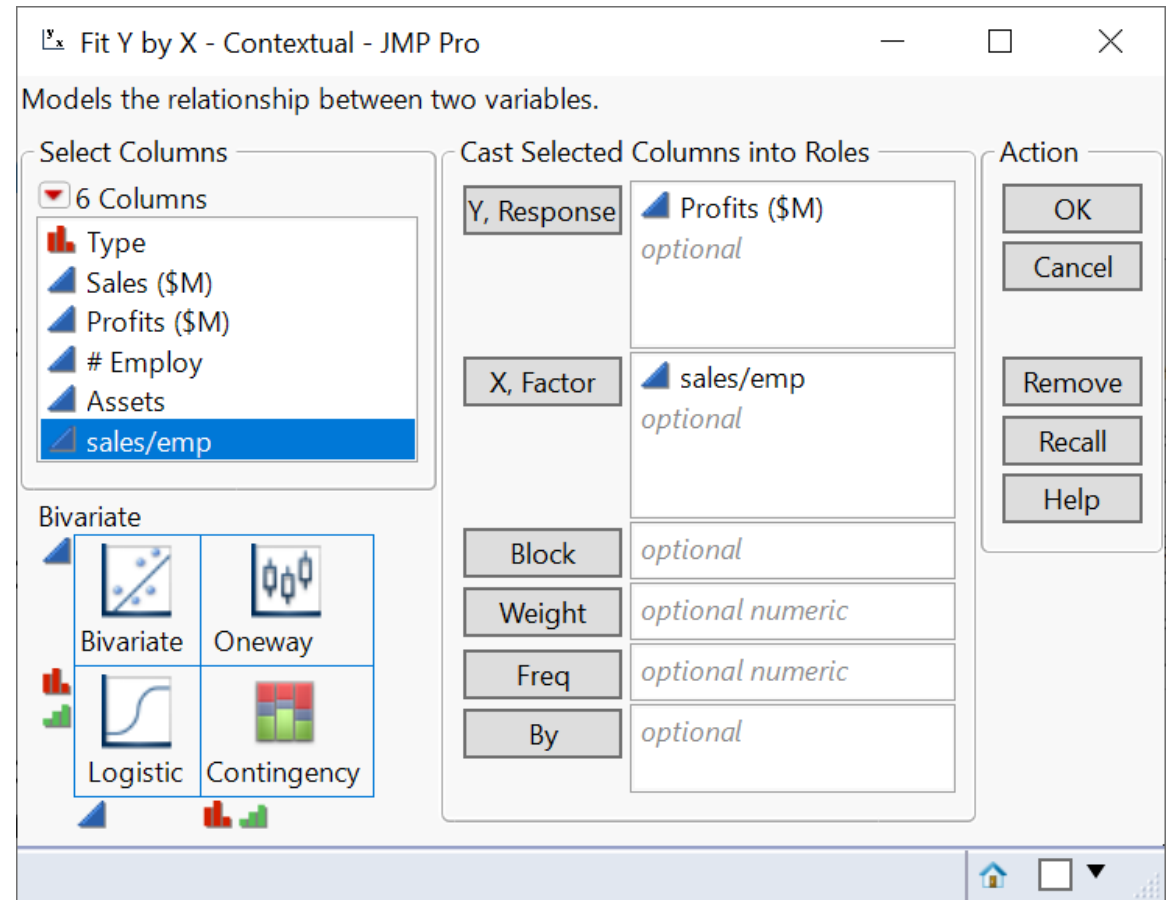
Review: Create a New Column / Variable

- Create a new column
 - Col name: sales/emp
 - Data type: Numeric
 - Col property: Formula
 - Formula: Sales (\$M) / # Employ
- A note on column names
 - JMP is flexible with col names
 - In general, special symbols in col names is not a good idea
- Formula editor



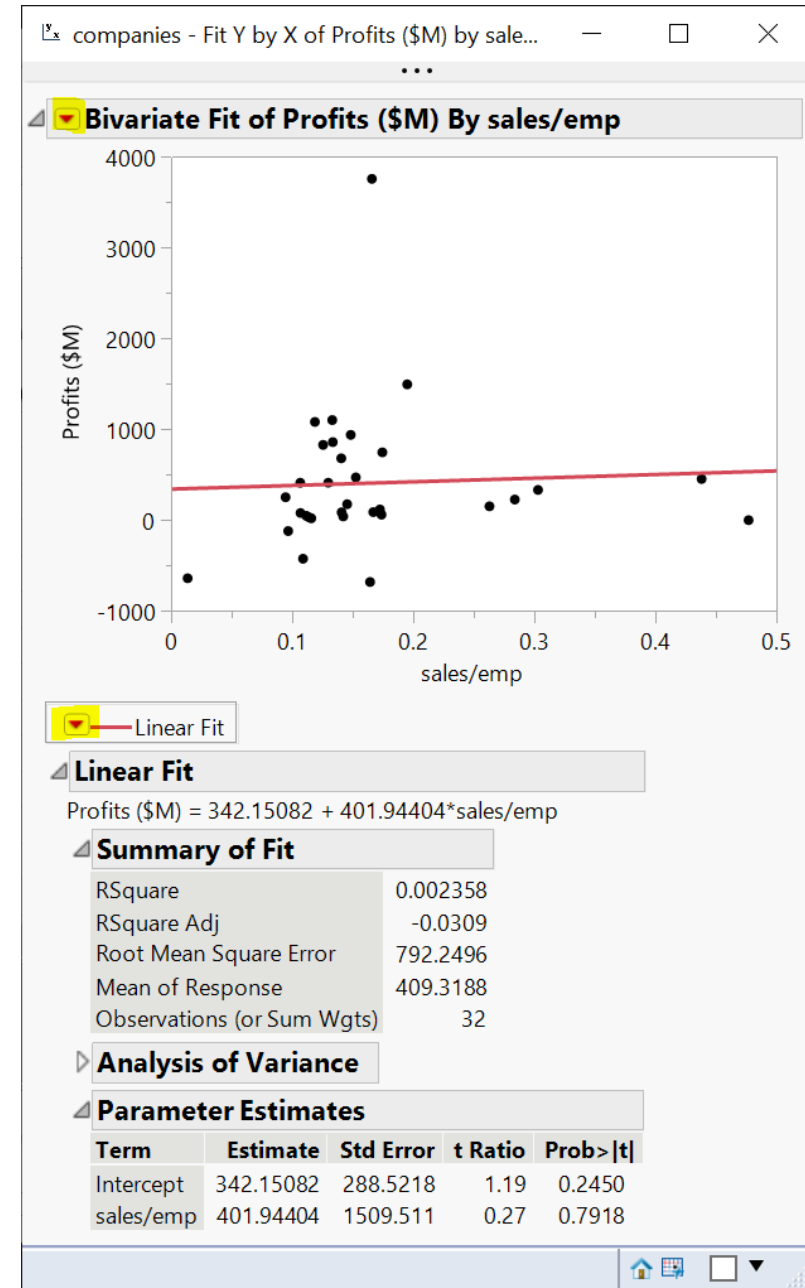
Review: Fit Y by X Platform

- Relationships between **two** variables
- Model choice depends on Y and X variable types
 - four model choices under bivariate model



Review: Analysis Report

- Many actions are available under the red triangle buttons
 - Fit line
 - Redo -> Redo Analysis
 - Redo -> Automatic Recalc
 - Save Script -> To Data Table...



Review: Save Analysis

- Save data table
- Capture the script for analysis report

The screenshot shows the JMP Pro interface with a data table and a bivariate fit analysis. The data table has the following columns:

Type	Sales
1 Computer	
2 Pharmaceutical	5
3 Computer	2
4 Pharmaceutical	
5 Computer	
6 Pharmaceutical	
7 Computer	2
8 Computer	
9 Computer	2
10 Computer	
11 Computer	1
12 Pharmaceutical	4
13 Computer	1
14 Computer	
15	

The bivariate fit analysis window shows the following options:

- Show Points (checked)
- Histogram Borders
- Summary Statistics
- Fit Mean
- Fit Line
- Fit Polynomial
- Fit Special...
- Flexible
- Fit Orthogonal
- Robust
- Density Ellipse
- Nonpar Density
- Group By...
- Local Data Filter
- Redo
- Save Script (highlighted)
- RSquare Adj
- Root Mean Square Error
- Mean of Response
- Observations (or Sum Wgts)
- Analysis of Variance
- Parameter Estimates

The 'Save Script' sub-menu is open, showing the following options:

- To Data Table... (highlighted)
- To Journal
- To Script Window
- To Report
- To Clipboard

Review: Create a Categorical Variable/Col

- Cols -> New Columns
 - Col Name: size
 - Data Type: Character
 - Model Type: Nominal
 - Col Property: Formula
- It's a categorical variable
 - if # Employ < 10000, then size = "small"
 - else, size = "large"

The image shows two overlapping windows from the JMP Pro software. The top window is titled 'New Column - JMP Pro' and is used to create a new column. The 'Column Name' is 'size', the 'Data Type' is 'Character', the 'Modeling Type' is 'Nominal', and the 'Initialize Data' is 'Missing/Empty'. The bottom window is titled 'size - JMP Pro' and shows the column's properties. The 'Formula' field contains the following conditional statement:

$$\text{If} \left(\begin{array}{l} \# \text{Employ} < 10000 \Rightarrow \text{"small"} \\ \text{else} \Rightarrow \text{"large"} \end{array} \right)$$

The 'Preview' section shows a data sample with the following results:

Formula result	# Employ
"small"	7523
"large"	40929
"small"	8200
"large"	50816
"large"	12068
"large"	54100

Review: Fit Model - Multiple Linear Reg

- Analyze > Fit Model

The screenshot shows the 'Fit Model - JMP Pro' dialog box, specifically the 'Model Specification' section. The dialog is titled 'Fit Model - JMP Pro' and has standard window controls (minimize, maximize, close). The 'Model Specification' section is expanded, showing the following options:

- Select Columns:** A list of 7 columns is shown: Type, Sales (\$M), Profits (\$M), # Employ, Assets, sales/emp, and size.
- Pick Role Variables:** The 'Y' role is assigned to 'Profits (\$M)'. The 'Weight' role is set to 'optional numeric', 'Freq' to 'optional numeric', 'Validation' to 'optional', and 'By' to 'optional'.
- Construct Model Effects:** The 'Add' button is highlighted. The list of effects includes 'Sales (\$M)' and 'size'. The 'Degree' is set to 2, 'Attributes' and 'Transform' are set to 'None', and the 'No Intercept' checkbox is unchecked.
- Personality:** The 'Standard Least Squares' personality is selected.
- Emphasis:** The 'Effect Leverage' emphasis is selected.
- Buttons:** 'Help', 'Run', 'Recall', 'Remove', and 'Keep dialog open' (checkbox) are visible.

Optional: Join Two Tables

- What if the variables you need for the analysis are in two tables?
- Join two tables

The image shows two overlapping windows from the JMP Pro software. The top window, titled 'companies_part1 - JMP Pro', displays a data table with the following columns: ID, Type, Profits (\$M), # Employ, and Assets. The bottom window, titled 'companies_part2 - JMP Pro', displays a data table with the following columns: ID and Sales (\$M). Both tables have 19 rows.

ID	Type	Profits (\$M)	# Employ	Assets
1	Computer	31	7523	615.2
2	Pharmaceutical	859.8	40929	4851.6
3	Computer	153	8200	2233.7
4	PH			
5	Co			
6	PH			
7	Co			
8	Co			
9	Co			
10	Co			
11	Co			
12	PH			
13	Co			
14	Co			
15	PH			
16	PH			
17	PH			
18	Co			
19	PH			

ID	Sales (\$M)
1	855.1
2	5453.5
3	2153.7
4	6747
5	5284
6	9422
7	2876.1
8	709.3
9	2952.1
10	784.7
11	1324.3
12	4175.6
13	11899
14	873.6
15	9844
16	969.2
17	6698.4
18	5956
19	5903.7

Optional: Inner Join

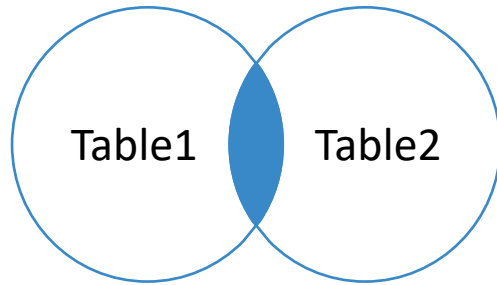


Table1

pk	t1c1
1	a
2	b

Table2

fk	t2c1
1	c
1	d
3	e

pk	t1c1	fk	t2c1
1	a	1	c
1	a	1	d

Join - JMP Pro

Join rows from several sources by matching value.

Join 'Table1' with

- Table2
- Table1

Options

- Preserve main table order
- Update main table with data from second table
- Merge same name columns
- Match Flag

Main Table

- Copy formula
- Suppress formula evaluation

Second Table

- Copy formula
- Suppress formula evaluation

Matching Specification

By Matching Columns ▾

Match Columns

Match pk=fk

Drop multiples Main Table With Table

Include non-matches Main Table With Table

Inner Join

Output Columns

- Select columns for joined table

Action

OK Cancel

Remove Recall Help

Output table name:

Keep dialog open

Save Script to Source Table

Optional: Inner Join

companies_part1 - JMP Pro

File Edit **Tables** Rows Cols DOE Analyze Graph Tools View Window Help

Summary
Subset
Sort
Stack
Split
Transpose
Join
Update
Concatenate
JMP Query Builder
Missing Data Pattern
Compare Data Tables
Anonymize

	Type	Profits (\$M)	# Employ	Assets
Computer	31	7523	615.2	
Pharmaceutical	859.8	40929	4851.6	
Computer	153	8200	2233.7	
Pharmaceutical	1102.2	50816	5681.5	
Computer	454	12068	2743.9	
Pharmaceutical	747	54100	8497	
Computer	333.3	9500	2090.4	
Computer	41.4	5000	468.1	
Computer	-680.4	18000	1860.7	
Computer	89	4708	955.8	
Computer	-119.7	13740	1040.2	
Pharmaceutical	939.5	28200	5848	
Computer	829	95000	10075	
Computer	79.5	8200	808	
Pharmaceutical	1082	83100	7919	
Pharmaceutical	227.4	3418	784	
Pharmaceutical	1495.4	34400	6756.7	
Computer	412	56000	4500	
Pharmaceutical	681.1	42100	8324.8	

evaluations done

Join - JMP Pro

Join rows from several sources by matching value.

Join 'companies_part1' with

- companies_part1
- companies_part2

Options

- Preserve main table order
- Update main table with data from second table
- Merge same name columns
- Match Flag

Main Table

- Copy formula
- Suppress formula evaluation

Second Table

- Copy formula
- Suppress formula evaluation

Matching Specification

By Matching Columns

Match Columns

Match ID=ID

Drop multiples Main Table With Table

Include non-matches Main Table With Table

Inner Join

Output Columns

- Select columns for joined table

Action


OK Cancel Remove Recall Help

Output table name:

Keep dialog open

Save Script to Source Table

Learning Resources (jmp.com)

- [JMP Academic \(Students\)](#) 
 - [JMP Basics](#) (great beginner videos!)
 - [Go Deeper](#) (one-page guide, short videos, webinars, and a full course.)
- [JMP Documentation Library](#) (many examples on included datasets)
 - Getting started document: [Discovering JMP](#)
 - Basics
 - [Using JMP](#) (data table)
 - [Basic Analysis](#)
 - [Essential Graphing](#)
 - Specific topics
 - [Fitting Linear Models](#)
 - [Predictive and Specialized Modeling](#)