

***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

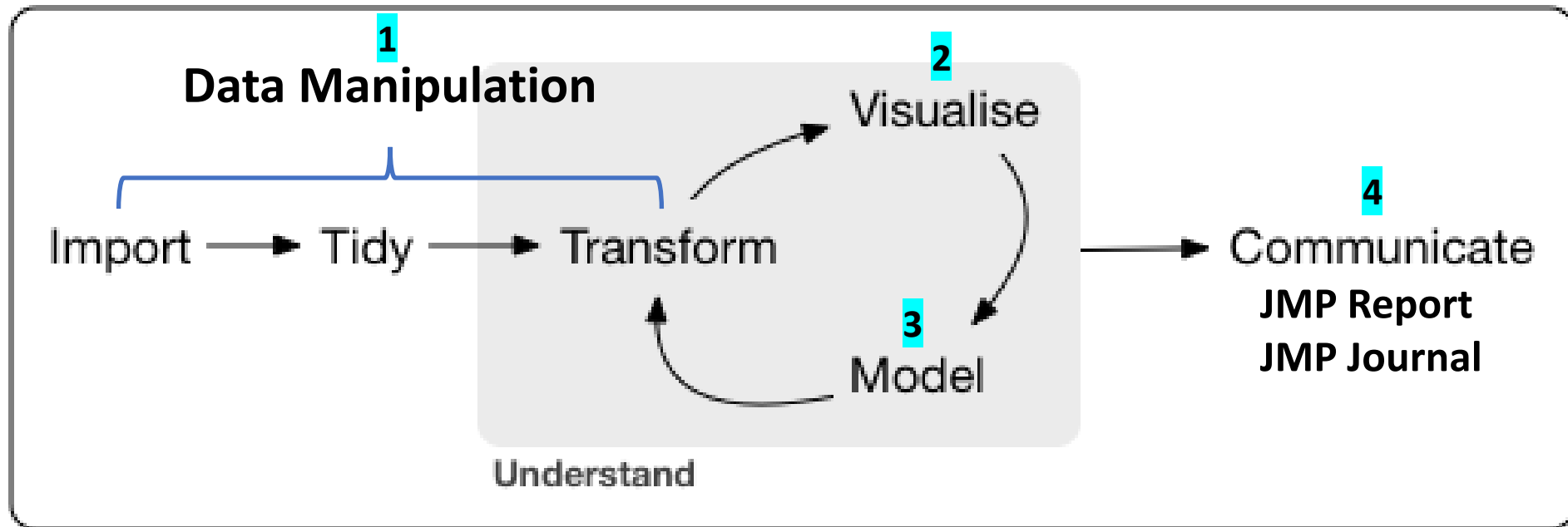
- What is JMP
- JMP basics (demos & learning by doing)
  - JMP basic navigations
  - Open/Import datasets
  - Manipulate data: a few simple tasks
  - Analyze data: a simple linear regression model
- Learning resources

# 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

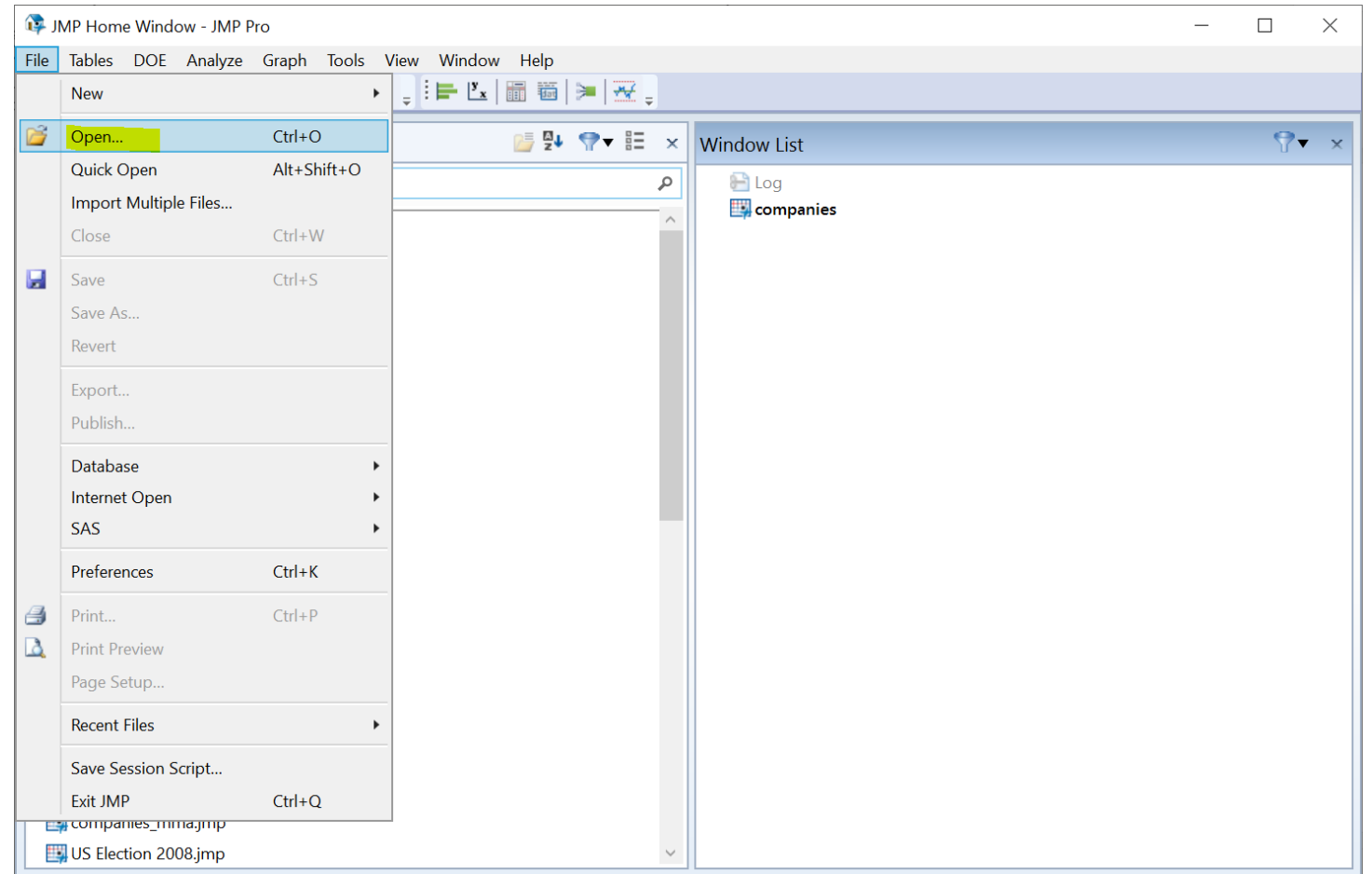


point-and-click & programming

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

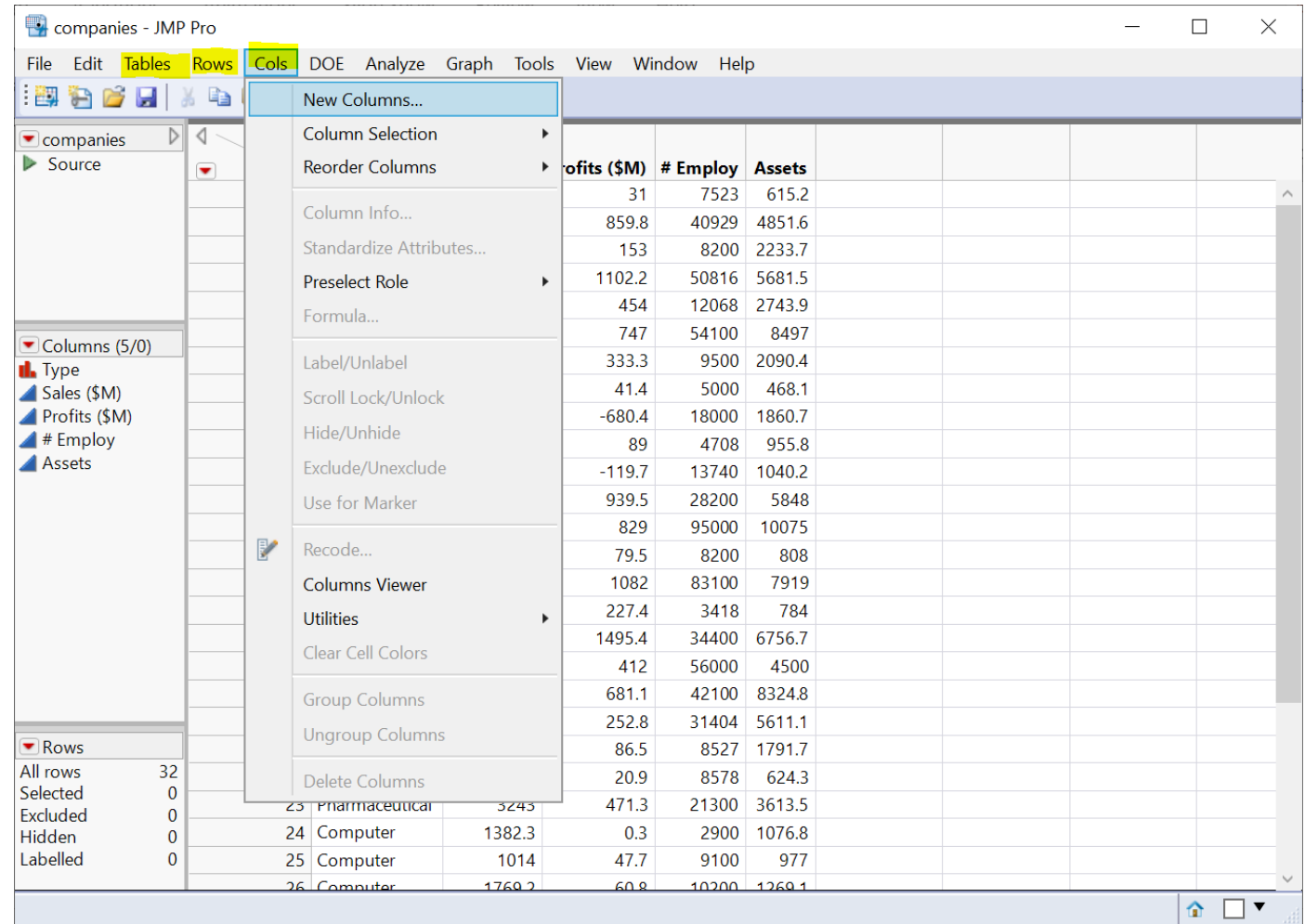
# JMP Navigations: Open / Import Data

- 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 software interface. The 'Analyze' menu is open, displaying options such as '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 has columns for '# Employ' and 'Assets', with rows numbered 1 through 26. The table data is as follows:

	# Employ	Assets
1	7523	615.2
2	40929	4851.6
3	8200	2233.7
4	50816	5681.5
5	12068	2743.9
6	54100	8497
7	9500	2090.4
8	5000	468.1
9	18000	1860.7
10	4708	955.8
11	13740	1040.2
12	28200	5848
13	95000	10075
14	8200	808
15	83100	7919
16	969.2	227.4
17	6698.4	1495.4
18	5956	412
19	5903.7	681.1
20	2959.3	252.8
21	1198.3	86.5
22	990.5	20.9
23	3243	471.3
24	1382.3	0.3
25	1014	47.7
26	1769.2	60.8

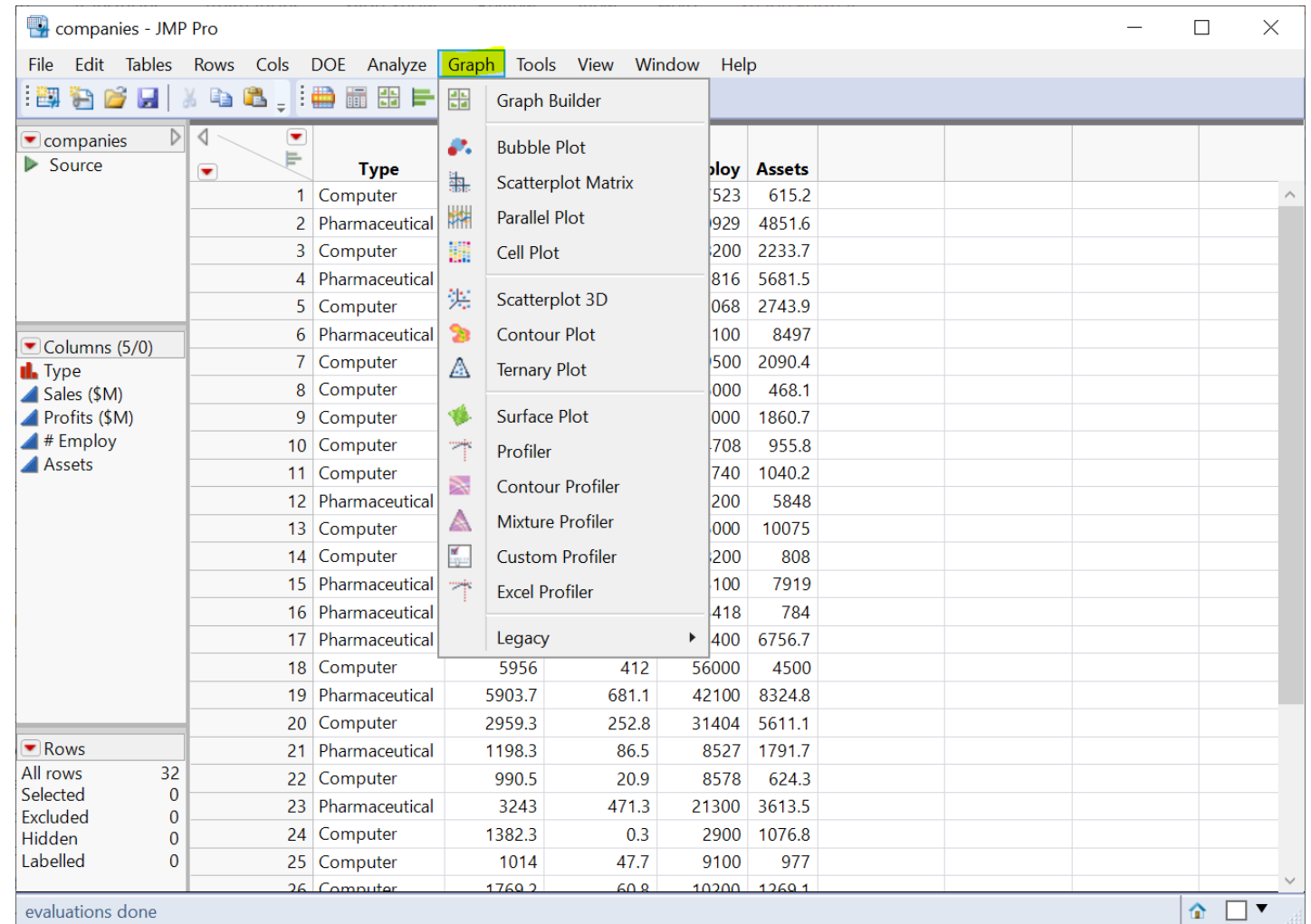
# JMP Navigations: Visualization & Reporting

- 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”

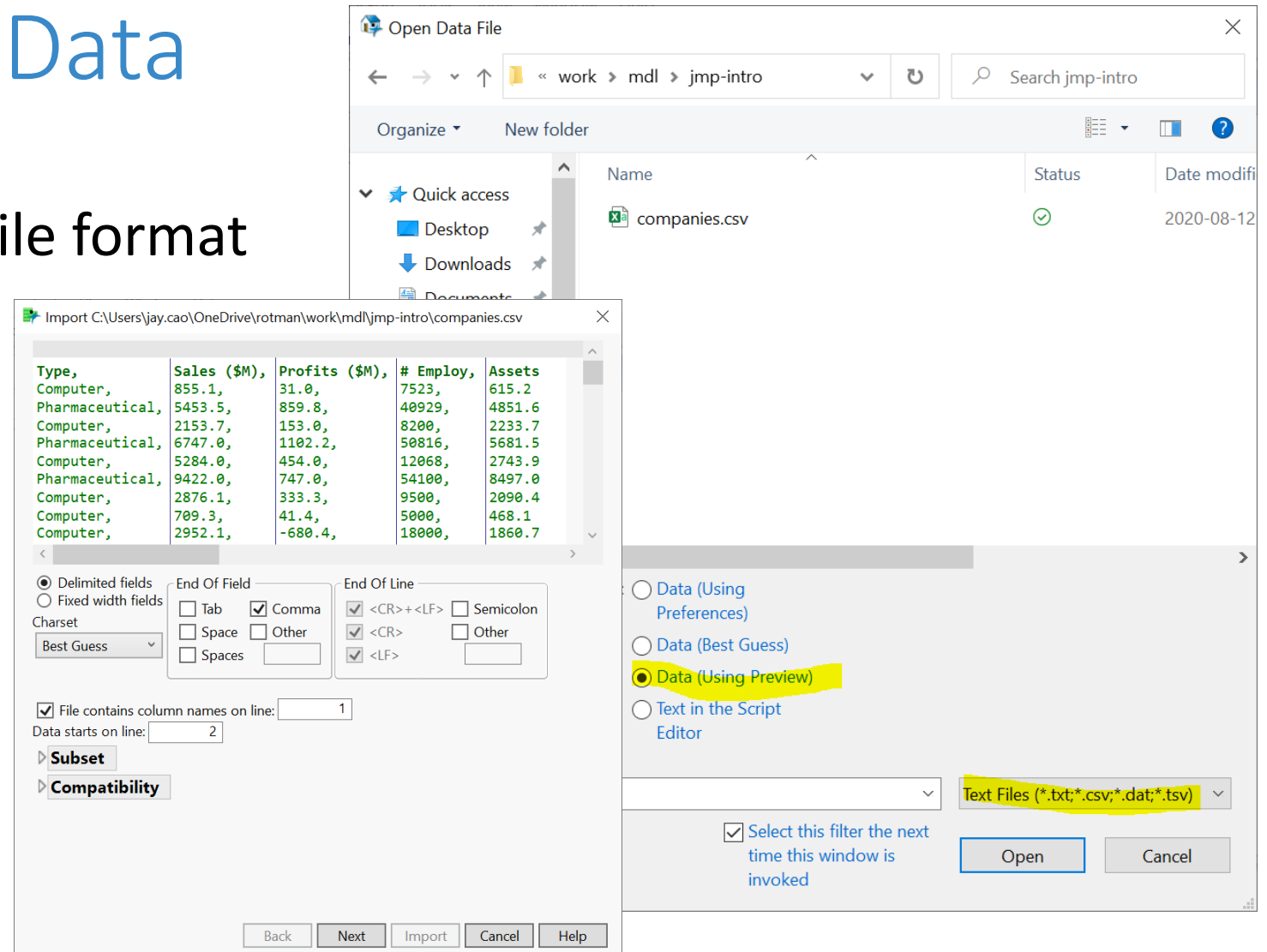
# Review: Import Data

- Import support many file format

- csv
- Excel
- json
- many more

- JMP native data file

- .jmp



<https://www.jmp.com/support/help/en/15.2/#page/jmp/import-your-data.shtml#>

# Review: Data Table

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

The screenshot displays the JMP Pro interface with the 'companies' data table open. The left sidebar shows three panels: 'Table (JMP Script)', 'Columns (5/1)', and 'Rows'. The 'Columns' panel lists the columns: Type, Sales (\$M), Profits (\$M), # Employee, and Assets. The 'Rows' panel shows 32 rows in total, with 0 selected, 0 excluded, 0 hidden, and 0 labelled. The main window shows the first 19 rows of the data table. The 'Column Info' dialog for the '# Employee' column is open, showing the column name, data type (Numeric), modeling type (Continuous), and format (Best, Width 12). The dialog also includes options for locking, using thousands separators, and a 'Column Properties' dropdown.

	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 (,)

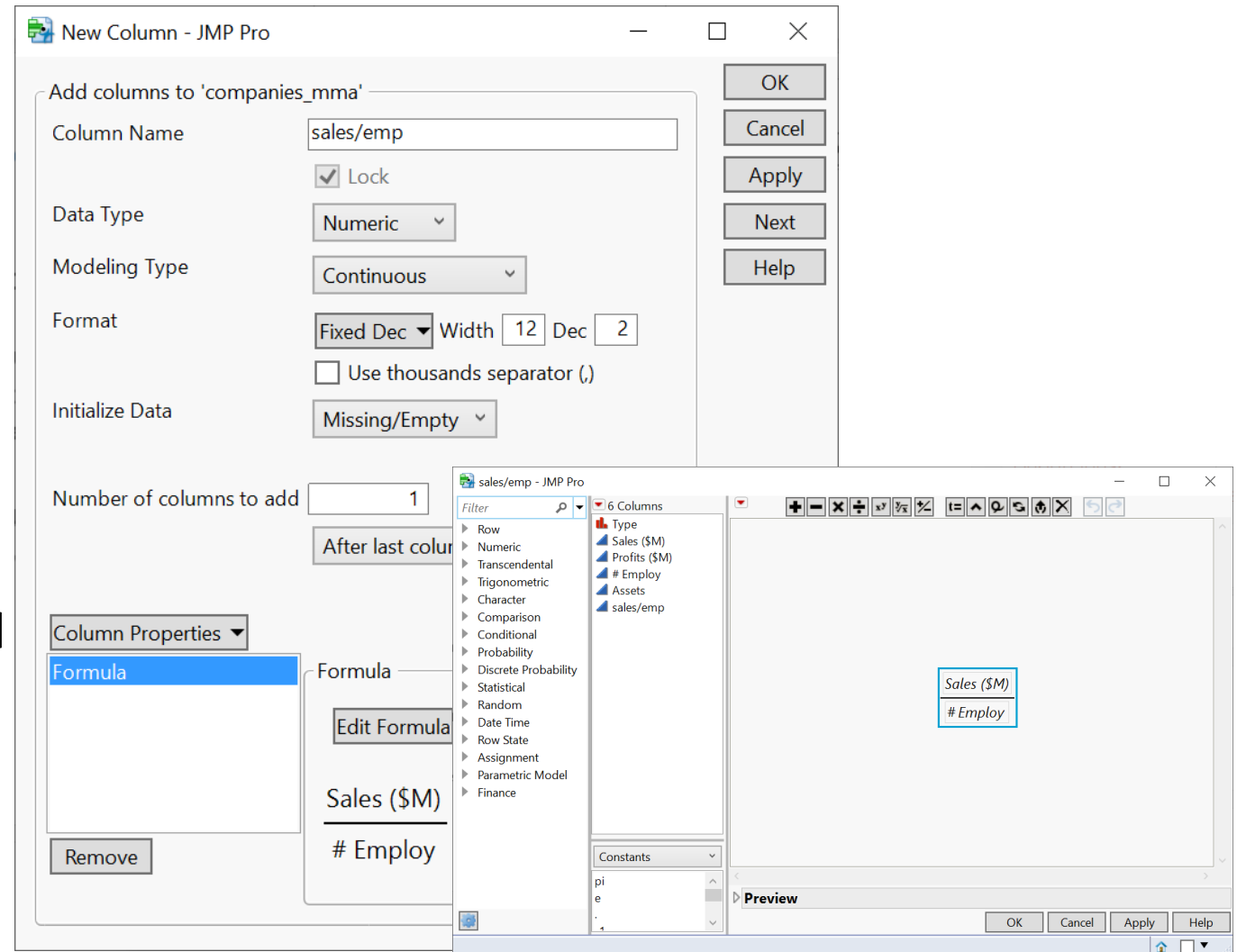
Column Properties

OK Cancel Apply Help

<https://www.jmp.com/support/help/en/15.2/#page/jmp/work-with-data-tables.shtml#>

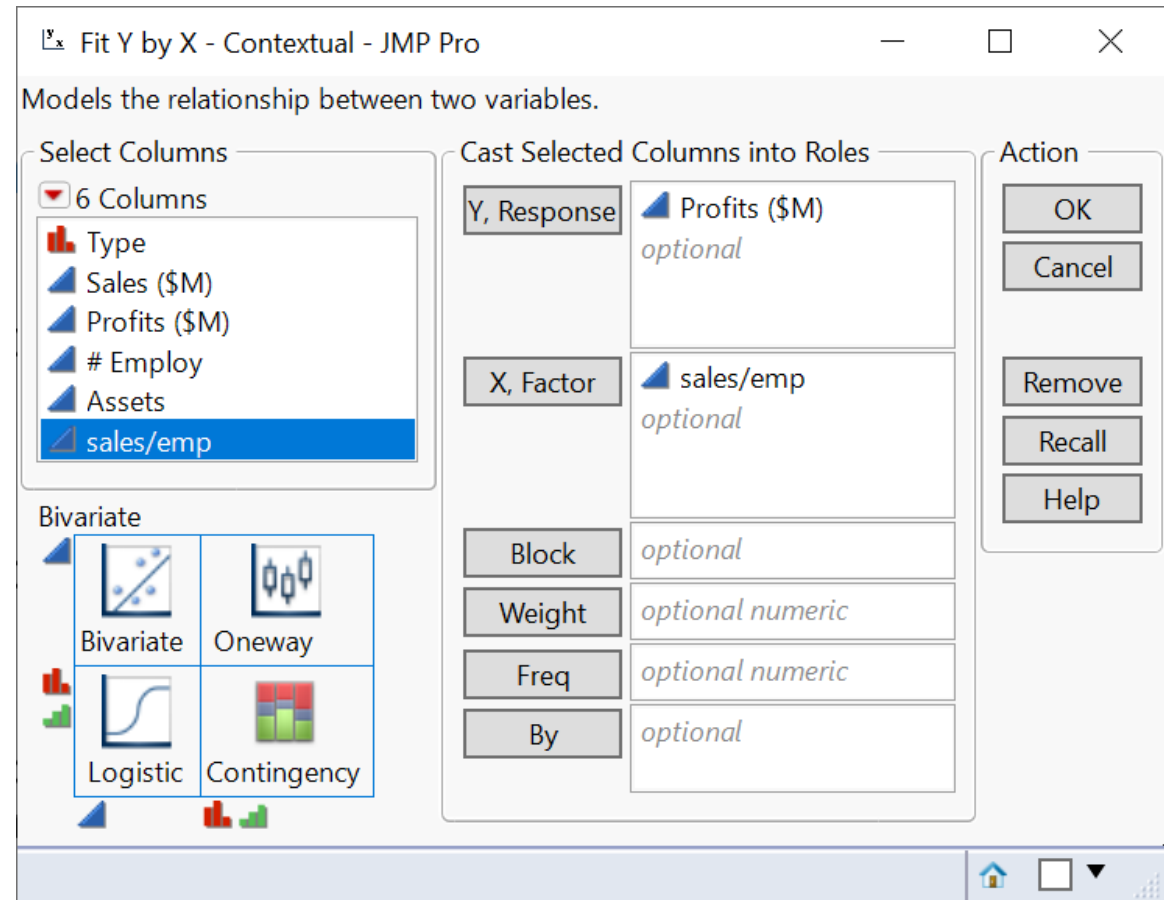
# Review: Create a New Column / Variable

- 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



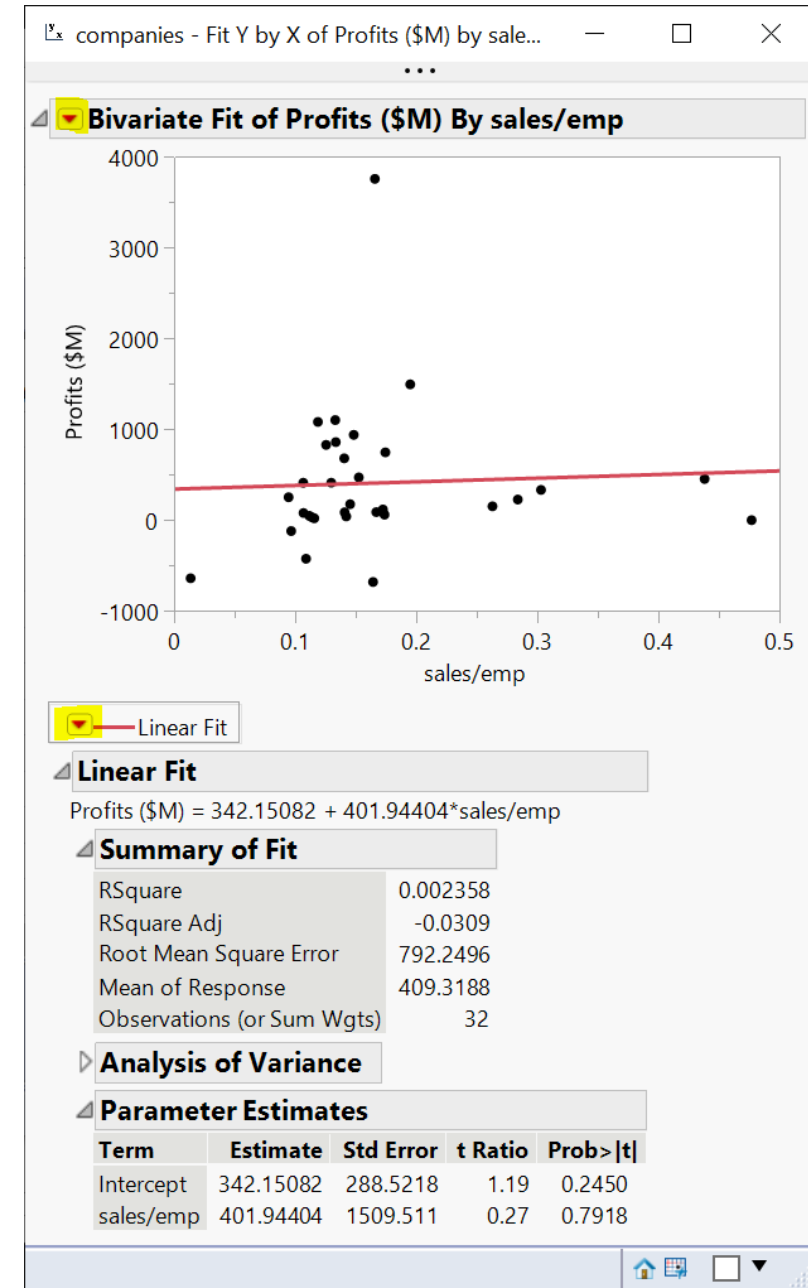
# Review: Fit Y by X Platform

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



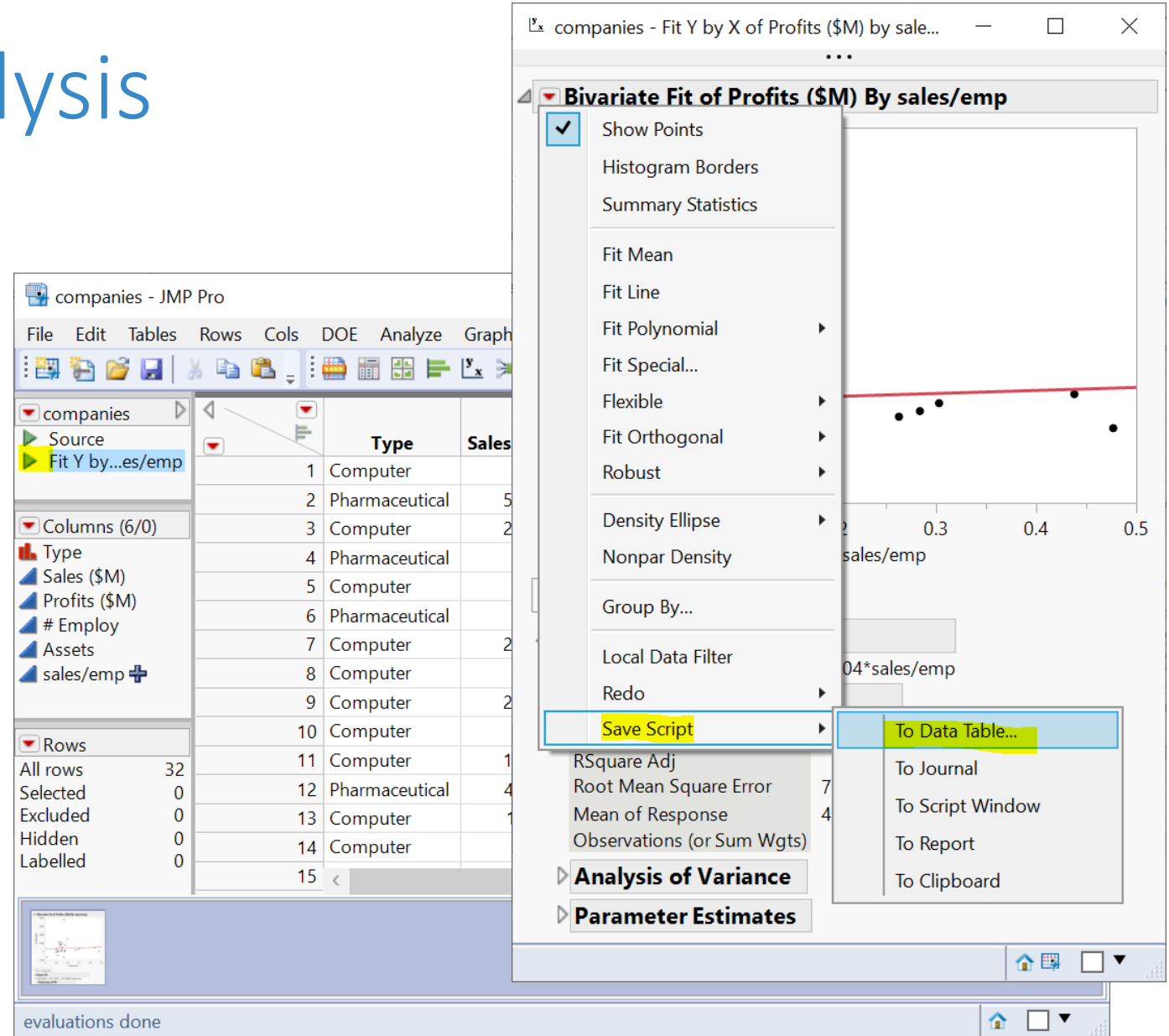
# Review: Analysis Report

- 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





# 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 screenshot displays the JMP Pro interface. The 'New Column - JMP Pro' dialog is open, showing the configuration for a new column named 'size'. The 'Data Type' is set to 'Character', the 'Modeling Type' is 'Nominal', and the 'Initialize Data' option is 'Missing/Empty'. The 'Column Properties' dropdown is set to 'Formula'.

The 'size - JMP Pro' window shows the formula editor for the 'size' column. The formula is: 
$$\text{If} \left( \begin{array}{l} \# \text{Employ} < 10000 \Rightarrow \text{"small"} \\ \text{else} \Rightarrow \text{"large"} \end{array} \right)$$

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

Fit Model - JMP Pro

**Model Specification**

Select Columns

7 Columns

- Type
- Sales (\$M)
- Profits (\$M)
- # Employ
- Assets
- sales/emp
- size

Pick Role Variables

Y: Profits (\$M) *optional*

Weight: *optional numeric*

Freq: *optional numeric*

Validation: *optional*

By: *optional*

Personality: Standard Least Squares

Emphasis: Effect Leverage

Help Run

Recall ☐ Keep dialog open

Remove

Construct Model Effects

Add Cross Nest Macros

Degree: 2


Attributes

Transform

☐ No Intercept

Sales (\$M)  
size

# 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](#)