

BIOS6606: LectureSep04

# Introduction to Prism and Good Data Practices

Kathleen Torkko, PhD, MSPH, MS

## **Some housekeeping**

We have are still working on TA office hours and location  
New TA Kaitlin Olson

Start attendance  
Please initial sheet

Suggestion “bags” are on door handles

# Aims of this Lecture

- Give you an introduction to the use of GraphPad Prism
- Learn about some good data practices

# Help is available to learn Prism

[http://www.graphpad.com/guides/prism/8/user-guide/index.htm?how\\_to\\_learn\\_prism.htm](http://www.graphpad.com/guides/prism/8/user-guide/index.htm?how_to_learn_prism.htm)

The screenshot shows the GraphPad Prism 8 User Guide website. The browser address bar displays the URL: [https://www.graphpad.com/guides/prism/8/user-guide/how\\_to\\_learn\\_prism.htm](https://www.graphpad.com/guides/prism/8/user-guide/how_to_learn_prism.htm). The website has a purple header with the GraphPad logo and navigation links: Statistics Guide, Curve Fitting Guide, Prism Guide (selected), and Resources. A left sidebar contains a search icon and a list of links: Welcome to Prism 8 User Guide, View the Statistics or Regression guide, **How to learn Prism** (selected), Prism Labs, What's new in Prism 8?, What was new in earlier versions?, How Prism is unique, Getting started with GraphPad Prism, Differences between Windows and Mac versions of Prism, Getting around in Prism, Prism files, Work efficiently. Repeat analyses and graphing, Organizing Prism projects, Data tables, Importing and exporting data, Analyzing data with Prism, Graphs, Page layouts, Adding text, drawings, results and pictures to graphs or layouts, Exporting images from Prism, Printing, Scripts, Standardizing Prism use in a lab, and Support and license. The main content area is titled 'How to learn Prism' and includes the following sections:

- Browse the Statistics and Regression guides**

This Prism Guide is only one of three guides. Check out the other two. Both do more than explain how to use Prism. They also explain the concepts of data analysis.

  - [GraphPad Statistics Guide](#)
  - [GraphPad Curve Fitting Guide](#)
- Step by step tours**

Read this [one page tour](#), which will take about five minutes to follow.

Then follow longer detailed tutorials focused on [statistical analysis](#) and [curve fitting](#) with step-by-step instructions.
- Work with tutorial data sets**

Choose these from the Welcome (File..New) dialog. After choosing one of the eight kinds of tables, choose "Start with sample data to follow a tutorial". Each sample comes with brief instructions, and links to longer explanations.

A screenshot of the 'Welcome to GraphPad Prism' dialog box is shown, illustrating the 'Start with sample data to follow a tutorial' option. The dialog box has a title bar 'Welcome to GraphPad Prism' and a subtitle 'XY tables: Each point is defined by an X and Y coordinate'. It features the GraphPad Prism logo (Version 8.00.200) and a 'New table & graph' button. Below this, there are options to 'Enter or import data into a new table' (selected) and 'Start with sample data to follow a tutorial' (highlighted with a red box). A 'Select a tutorial data set:' section lists various data sets, including 'Error bars in XY tables', 'Correlation & regression (linear and nonlinear)', and 'Pharmacology'.

To use Prism effectively, you need to understand these essential concepts:

### **Prism data tables are organized. It is essential to choose the right kind of data table.**

Unlike Excel or most statistics programs, Prism's data tables are organized with a structure or format. To use Prism effectively, therefore, you must learn the [distinction between the eight types of data tables](#). Don't try to choose the kind of table based on the kind of graph to make. That approach doesn't always work. Choose a table based on the organization of your data and the analyses you wish to perform.

### **A Prism project can contain many data tables and graphs**

A Prism project (another term for a Prism file) is not limited to a single graph or data table, but rather can contain up to 500 data tables, graphs, analyses, and page layouts.

### **Prism can automatically plot error bars from raw data**

One of Prism's strengths is its ability to handle replicates and error bars. For XY and Grouped tables, replicates are placed in side-by-side subcolumns. For Column tables, replicates are stacked in each column. If you enter replicate values, Prism can plot either individual replicates or error bars. It can plot error bars from the replicates you entered automatically. You don't have to specify any calculations.

### **When you edit or replace data, analyses and graphs are automatically updated**

Prism remembers the logical links between data tables, info sheets, results tables, graphs and layouts. When you edit or replace data, Prism automatically recomputes linked analyses and redraws linked graphs.

### **Analyses can be chained**

Results tables can be analyzed further. From any results table with green grid lines, click the Analyze button. This lets you, say, first transform your data, and then fit a curve (nonlinear regression) to the results.

### **Results of analyses can be presented on multiple analysis tabs**

Analysis tabs at the top of the results page let you see different parts of the results.

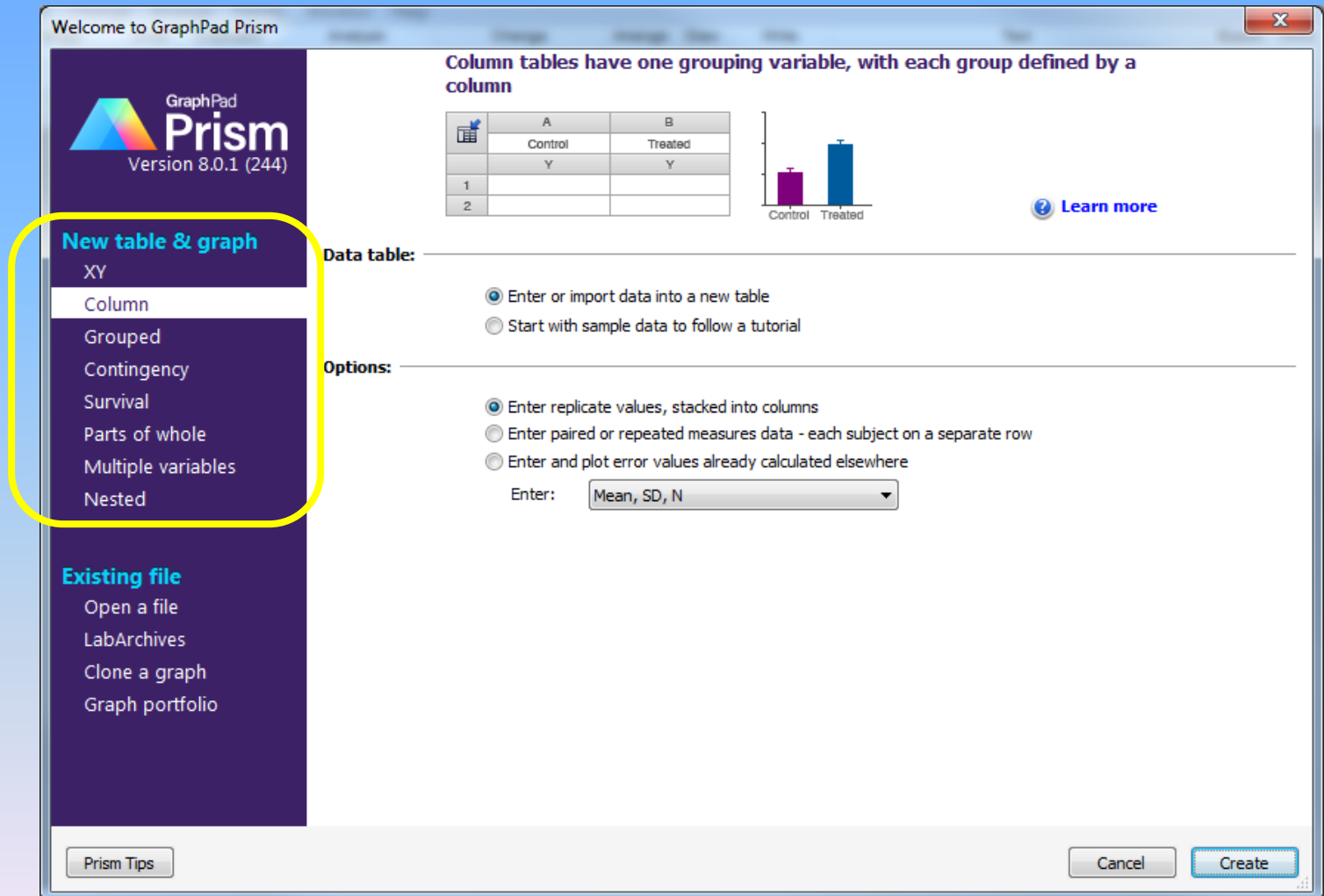
### **One table ≠ one graph**

By default, Prism creates one graph for each data table you create, but this should not limit you. You can plot a table on multiple graphs. And you can plot data from multiple data tables on one graph.

### **Recycle your work**

Once you've analyzed and graphed one data table, it is easy to repeat your work with new data by [duplicating a family](#) or [cloning a graph](#).

When you start Prism, you'll see the Welcome dialog. You'll see 8 tabs representing the 8 kinds of data tables.



# The six different data tables

## New table & graph

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

XY tables - An XY table is a graph where every point is defined by both an X and a Y value. This kind of data are often fit with linear or nonlinear regression.

Table format: XY		X	A		
		minutes	Test group A		
	X	X	A:Y1	A:Y2	A:Y3
1	Title	0	0.0	0.0	0.0
2	Title	2	391.0	384.0	543.0
3	Title	4	562.0	478.0	584.0
4	Title	6	746.0	798.0	715.0
5	Title	8	823.0	754.0	669.0
6	Title	10	736.0	846.0	742.0
7	Title	12	832.0	855.0	799.0
8	Title	14	923.0	750.0	816.0
9	Title	16	801.0	854.0	826.0
10	Title	18	811.0	795.0	864.0
11	Title	20	942.0	831.0	938.0

Analyses performed with XY data

- Linear regression
- Nonlinear regression
- Correlation (Pearson or Spearman)

Column tables - Use column tables if your groups are defined by one scheme, perhaps control vs. treated, or placebo vs. low-dose vs. high-dose. Each column defines one group.

	A	B
	Placebo	Active drug
	Y	Y
1	45	67
2	23	46
3	56	113
4	76	79
5	81	123
6	87	
7	99	

Analyses performed with Column data

- t test (one-sample, paired and unpaired)
- Mann-Whitney / Wilcoxon
- Column statistics (means, medians, SD, etc.)
- One-way ANOVA (followed by post tests)
- Kruskal-Wallis
- ROC curves

Also will do correlations

# The six data tables

## New table & graph

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Grouped tables** - The idea of two-way variables is best understood by example. One grouping variable (male vs. female in the example below) is defined by rows; the other grouping variable (control vs. treated) is defined by columns.

		A			B		
		Control			Treated		
		A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3
1	Men	34.5	32.9	43.3	87.5	321.5*	81.7
2	Women	42.3		45.9	109.4	111.2	115.4
3	Title						

Analyses performed with Grouped data

- Two-way ANOVA
- Repeated-measures two-way ANOVA
- Bonferroni multiple comparisons test

**Contingency tables** - Contingency tables are used to tabulate the actual number of subjects (or observations) that fall into the categories defined by the rows and columns of a table.

		A	B
		Graft Patent	Graft Obstructed
		Y	Y
1	Standard treatment	45	5
2	Experimental treatment	49	1
3	Title		

Analyses performed from a contingency table

- Fisher's exact test
- Chi-square test
- Odds ratios and relative risk

**Survival tables** - Survival tables are used to enter information for each subject. Prism then computes percent survival at each time, and plots a Kaplan-Meier survival plot (and also compares survival with the log-rank and Gehan-Wilcoxon tests).

- Kaplan-Meier / Log-rank

**Parts of whole tables** - A Parts of whole table is used when it makes sense to ask: What fraction of the total is each value? This table is often used to make a pie chart. The example below shows the number of students who received grades of A, B, C, etc. The sum of the five values in the table (61) is the total number of students.

Table format: Parts of whole		A
		Number of Students
		Y
1	A	23
2	B	29
3	C	7
4	D	2
5	E	0
6	Title	

Analyses performed on parts of whole data

- Chi-square goodness of fit



## New table & graph

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

# New in Prism 8

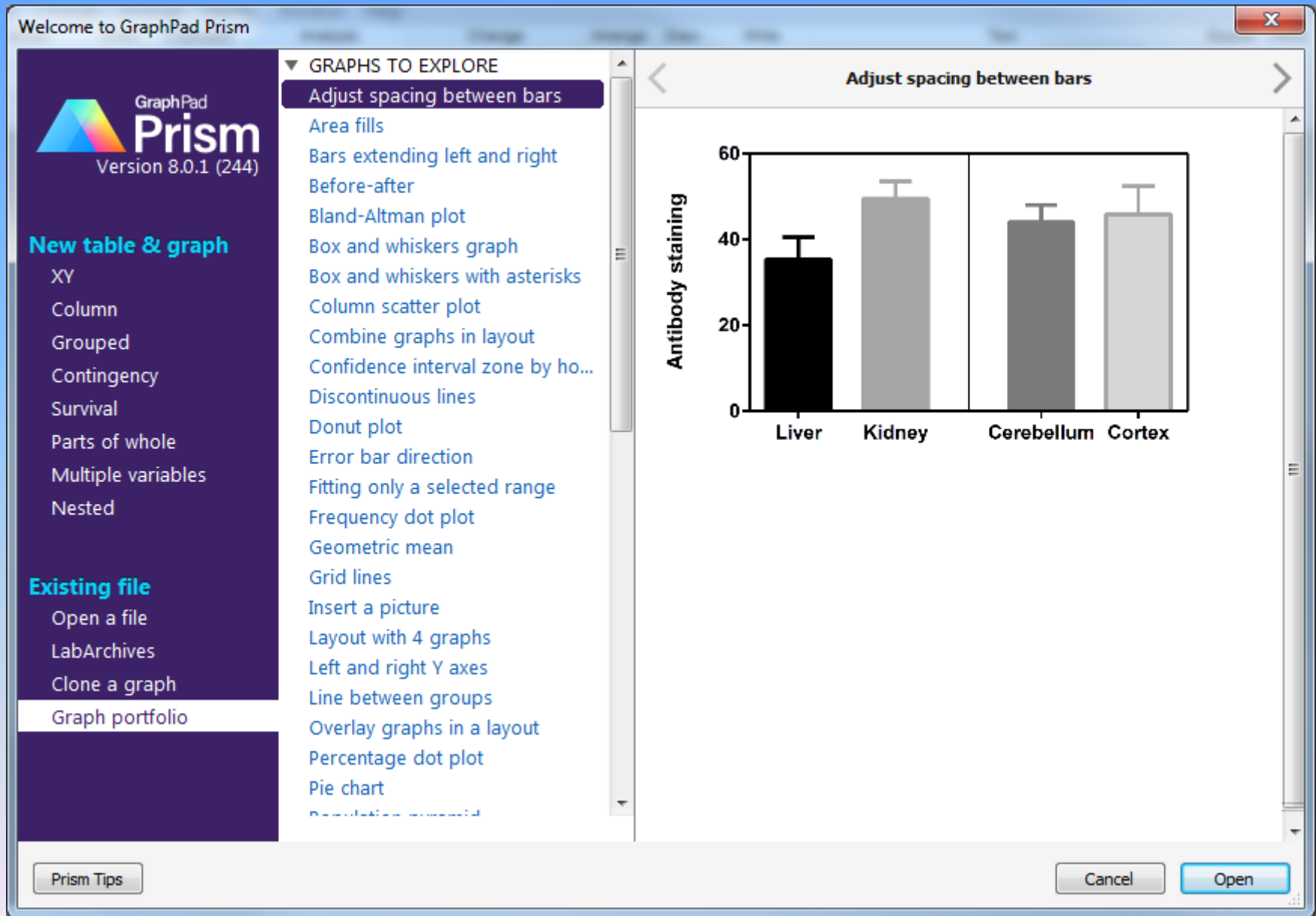
**Multiple variable tables.** A multiple variable table is arranged the same way most statistics programs organize data. Each row is a "case" (experiment, animal, etc.). Each column is a different variable. There are no subcolumns in multiple variable tables.

Table format:		Variable A	Variable B	Variable C	Variable D	Variable E	Variable F	Variable G	Variable H	Variable I	Variable J
Multiple variables		Glycosylated hemoglobin %	Total cholesterol	Glucose	HDL	Age in years	Male?	Height in inches	Weight in pounds	Waist in inches	Hip in inches
1	Title	4.309999943	203	82	56	46	0	62	121	29	38
2	Title	4.440000057	165	97	24	29	0	64	218	46	48
3	Title	4.639999866	228	92	37	58	0	61	256	49	57
4	Title	4.630000114	78	93	12	67	1	67	119	33	38
5	Title	7.719999790	249	90	28	64	1	68	183	44	41

**Nested tables.** A nested table is used when there are two levels of nested or hierarchical replication. In this example, there are two treatments given to three rats each and the outcome variable was measured four times in each rat.

	Group A			Group B		
	Teaching method A			Teaching method B		
	Room 1	Room 2	Room 3	Room 4	Room 5	Room 6
1	21	18	35	26	38	31
2	26	25	28	34	44	41
3	33	26	32	27	34	34
4	22	24	36		45	35
5		21	38		38	38
6		25				46
7						

# Welcome Dialog: Graph Portfolio



# Tutorial Datasets

Welcome to GraphPad Prism



## New table & graph

### XY

- Column
- Grouped
- Contingency
- Survival
- Parts of whole
- Multiple variables
- Nested

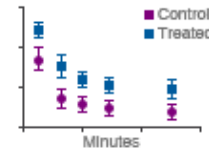
## Existing file

- Open a file
- LabArchives
- Clone a graph
- Graph portfolio

Prism Tips

## XY tables: Each point is defined by an X and Y coordinate

	X	A			B		
	Minutes	Control			Treated		
	X	A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3
1	Title						
2	Title						
3	Title						



[Learn more](#)

Data table:

- ☐ Enter or import data into a new table
- ☒ Start with sample data to follow a tutorial

Select a tutorial data set:

- ☒ **Error bars in XY tables**
  - Entering replicate data
  - Entering mean (or median) and error values
- ☒ **Correlation & regression (linear and nonlinear)**
  - Correlation
  - Simple linear regression
  - Linear regression - Compare slopes
  - Interpolate unknowns from a linear standard curve
  - Nonlinear regression - One phase exponential decay
  - Eliminating outliers during nonlinear regression
  - RIA or ELISA - Interpolate unknowns from sigmoidal curve
- ☒ **Pharmacology**
  - Dose-response - X is log(dose)
  - Dose-response - X is dose
  - Dose-Response - Ambiguous until constrained
  - Dose-response - EC50 shift by global fitting

- Binding - Saturation binding to total and nonspecific
- Binding - Saturation binding, specific binding only
- Binding - Compare two models - One vs two site competition
- Schild competitive antagonist - Global nonlinear regression
- Operational model of agonist action - Global nonlinear regression

- ☒ **Enzyme kinetics**
  - Enzyme kinetics - Michaelis-Menten
  - Enzyme kinetics - Competitive inhibition

- ☒ **Special uses of XY tables**
  - Entering dates into the X column
  - Entering elapsed times into the X column
  - Making an XY bar graph
  - XY Frequency distribution

**New table & graph**

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Existing file**

Open a file

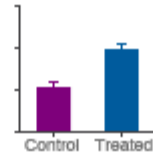
LabArchives

Clone a graph

Graph portfolio

Column tables have one grouping variable, with each group defined by a column

	A	B
	Control	Treated
	Y	Y
1		
2		

[Learn more](#)**Data table:**

- ☐ Enter or import data into a new table
- ☒ Start with sample data to follow a tutorial

**Select a tutorial data set:**

- ☐ **Error bars in column tables**
  - Entering replicate data
  - Entering mean (or median) and error values
- ☐ **T tests**
  - t test - Unpaired
  - t test - Paired
  - t test - One sample
- ☐ **One-way ANOVA**
  - Ordinary one-way ANOVA
  - Repeated measures one-way ANOVA
- ☐ **Special uses of column tables**
  - Analyze a stack of P values
  - ROC curve
  - Bland-Altman method comparison
  - Metanalysis (Forest) plot
  - Descriptive statistics

Prism Tips

Cancel

Create

**New table & graph**

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Existing file**

Open a file

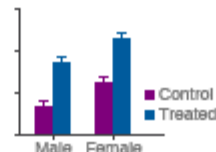
LabArchives

Clone a graph

Graph portfolio

Grouped tables have two grouping variables, one defined by columns and the other defined by rows

Table format		A			B		
		Control			Treated		
		A:Y1	A:Y2	A:Y3	B:Y1	B:Y2	B:Y3
1	Male						
2	Female						

[Learn more](#)**Data table:**

- ☐ Enter or import data into a new table
- ☒ Start with sample data to follow a tutorial

**Select a tutorial data set:****Error bars in grouped tables**

- Entering replicate data
- Entering mean (or median) and error values

**Two-way ANOVA**

- Ordinary - two data sets
- Ordinary - three data sets
- Repeated measures - matched values stacked
- Repeated measures - matched values in same row

**Three-way ANOVA**

- Three-way ANOVA 2 x 2 x 2
- Three-way ANOVA 2 x 2 x K

**Special uses of grouped tables**

- Multiple t tests
- Heat map

Prism Tips

Cancel

Create

**New table & graph**

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Existing file**

Open a file

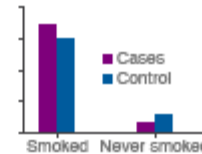
LabArchives

Clone a graph

Graph portfolio

**Contingency tables:** Each row defines a treatment or exposure, each column defines an outcome, and each value is an exact count of objects or events

Table format		A	B
Contingency		Cases	Control
		Y	Y
1	Smoked		
2	Never smoked		

[? Learn more](#)**Data table:**

- ☐ Enter or import data into a new table
- ☒ Start with sample data to follow a tutorial

**Select a tutorial data set:**

- ☒ Chi-square test of prospective data (aspirin and MI)
- ☐ Fishers exact test of retrospective data (smoking and cancer)
- ☐ Sensitivity and specificity (HIV)
- ☐ Chi-square test for trend

Prism Tips

Cancel

Create

**New table & graph**

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Existing file**

Open a file

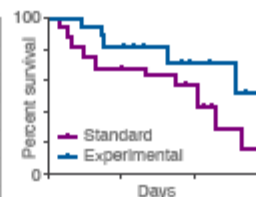
LabArchives

Clone a graph

Graph portfolio

**Survival tables: Each row tabulates the survival or censored time of a subject**

Table format	X	A
Survival	Days	Standard
	X	Y
1	Title	
2	Title	
3	Title	
4	Title	

[? Learn more](#)**Data table:**

- ☐ Enter or import data into a new table
- ☒ Start with sample data to follow a tutorial

**Select a tutorial data set:**

- ☒ Comparing two groups
- ☐ Three groups

Prism Tips

Cancel

Create

**New table & graph**

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Existing file**

Open a file

LabArchives

Clone a graph

Graph portfolio

**Parts of whole tables: Each row defines a mutually exclusive category**

Table format		A
Parts of whole		Cases
		Y
1	Male	
2	Female	

 [Learn more](#)**Data table:**

- ☐ Enter or import data into a new table
- ☒ Start with sample data to follow a tutorial

**Select a tutorial data set:**

- ☒ Distribution of student grades (compute fractions of total)
- ☐ Chi-square to compare observed and expected distributions of Mendel's peas

Prism Tips

Cancel

Create



**New table & graph**

XY  
Column  
Grouped  
Contingency  
Survival  
Parts of whole

Multiple variables

Nested

**Existing file**

Open a file  
LabArchives  
Clone a graph  
Graph portfolio

**Multiple variable tables:** Each column represents a different variable. Each row represents a different individual or experimental unit.

Table Format Multiple variables		Variable A	Variable B	Variable C
		Title	Title	Title
1				

[Learn more](#)

**Data table:**

- ☐ Enter or import data into a new table  
☒ Start with sample data to follow a tutorial

**Select a tutorial data set:**

- ☒ Multiple regression  
☐ Correlation matrix

Prism Tips

Cancel

Create

**New table & graph**

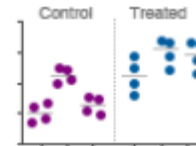
XY  
Column  
Grouped  
Contingency  
Survival  
Parts of whole  
Multiple variables  
Nested

**Existing file**

Open a file  
LabArchives  
Clone a graph  
Graph portfolio

**Nested data tables:** Hierarchical or nested is a data design when each treatment was tested in experimental replicates and each experimental replicate was assessed a number of times (technical replicates)

	Group A			Group B		
	Control			Treated		
	Rat 1	Rat 2	Rat 3	Rat 4	Rat 5	Rat 6
1						
2						

[Learn more](#)**Data table:**

- ☐ Enter or import data into a new table  
☒ Start with sample data to follow a tutorial

**Select a tutorial data set:**

- ☒ Nested one-way ANOVA  
☐ Nested t test

[Prism Tips](#)[Cancel](#)[Create](#)

### **Analyses performed with XY data**

- Nonlinear regression (curve fit)
- Linear regression
- Fit Spline/LOWESS
- Smooth, Differentiate or Integrate Curve
- Area Under Curve
- Deming (Model II) Linear regression
- Correlation matrix
- Correlation XY
- Interpolate a Standard Curve

### **Analyses performed with Column data**

- Unpaired t test
- Paired t test
- Ratio-paired t test
- Mann-Whitney test
- Wilcoxon test
- Kolmogorov-Smirnov test
- Ordinary one-way ANOVA
- Brown-Forsythe and Welch ANOVA
- RM one-way ANOVA
- Kruskal-Wallis test
- Friedman test
- One sample t and Wilcoxon test
- Descriptive Statistics
- Normality and Lognormality Tests
- Frequency Distribution
- ROC Curve
- Bland-Altman
- Identify Outliers
- Analyze a stack of P values

### **Analyses performed with Grouped data**

- Two-way ANOVA (and mixed model)
- Three-way ANOVA (and mixed model)
- Row means with SD or SEM
- Multiple t tests - one per row

### **Analyses performed from a contingency table**

- Chi-square and Fisher's exact test (also computes odds ratios and relative risk)
- Fraction of total

### **Analyses performed with survival data**

There is one survival analysis in Prism that gives you choices for any of these analyses:

- Kaplan-Meier
- Log-rank test
- Wilcoxon-Gehan test

### **Analyses performed on parts of whole data**

- Fraction of total
- Chi-square goodness of fit (compare observed distribution with theoretical distribution)

### **Analyses performed on multiple variable data**

- Correlation matrix
- Multiple linear regression
- Extract and rearrange portions of the data onto a new table
- Transform and select
- Identify outliers
- Descriptive statistics

### **Analyses performed from nested tables**

- Nested t test
- Nested one-way ANOVA
- Descriptive statistics (separate results for each subcolumn)
- Normality (and lognormality) tests (separate results for each subcolumn)>
- Outlier tests (separate results for each subcolumn)
- One-sample t test (separate results for each subcolumn)

Project1:Column: Entering replicate data - GraphPad Prism 8.0.1 (244)

File Edit View Insert Change Arrange Family Window Help

Prism File Sheet Undo Clipboard Analysis Change Import Draw Write Text Export Print Send LA Help

Search...

▼ Data Tables

- Column: Entering replicate data
- + New Data Table...

▼ Info

- Project info 1
- + New Info...

▼ Results

- + New Analysis...

▼ Graphs

- Column: Entering replicate data
- + New Graph...

▼ Layouts

- + New Layout...

Family

- Column: Entering replicate data
- Column: Entering replicate data

	Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H	Group I	Group J	Group K
	Untreated	Placebo	Treated	Title	Title	Title	Title	Title	Title	Title	Title
1	3.4										
2		2.3	4.2								
3	4.3	5.2	7.8								
4	3.0	4.5	5.9								
5	3.9	3.1	6.4								
6	4.1	5.0	7.6								
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											

**How the data are organized**  
The three columns define three groups. The replicate values are stacked in each column

**How to graph these data**  
Look in the navigator to the left and find the Graphs folder. Click on the graph in bold. Prism will present several choices for graphing these data. With small numbers of values (like here), it is best to create a scatter plot that shows every value.

**Notes**

- Unlike many statistics programs, Prism does not define groups using a grouping variable. Instead, the groups are defined by which column each value is entered into.
- When you make a graph, you can choose to plot a column scatter graph (showing each value) or various kinds of graphs that summarize the data (bar graphs, box-and-whisker graphs, violin plots).

[? More about Column tables](#)

Column: Entering replicate data Row 1, A: Untreated



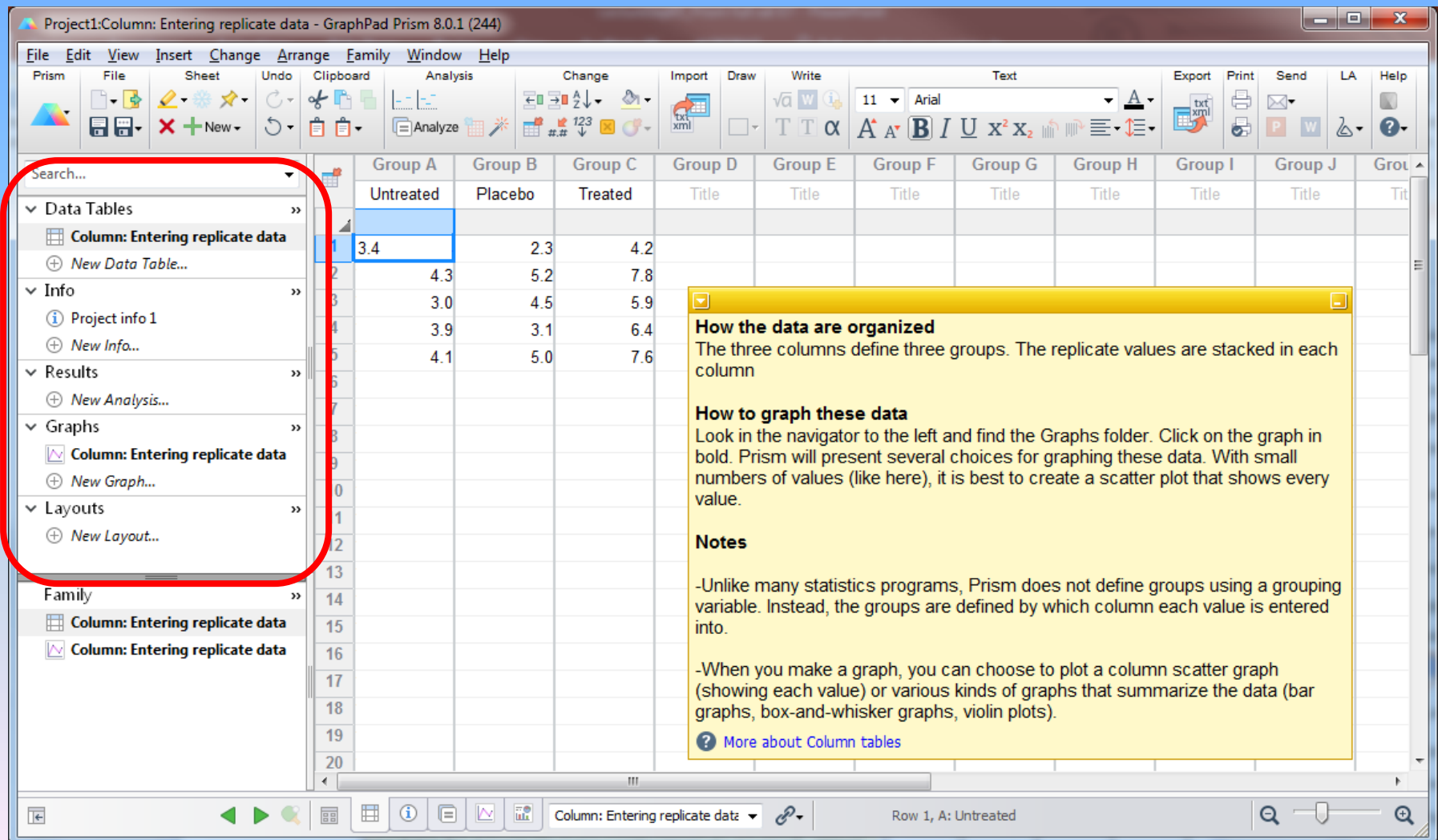
## Prism

Control	Treated	Treated+Antagonist
54	87	45
23	98	39
45	64	51
54	77	49
45	89	50
47		55

## Most other programs

Weight	Treatment Group
54	Control
23	Control
45	Control
54	Control
45	Control
47	Control
87	Treated
98	Treated
64	Treated
77	Treated
89	Treated
45	Treated+Antagonist
39	Treated+Antagonist
51	Treated+Antagonist
49	Treated+Antagonist
50	Treated+Antagonist
55	Treated+Antagonist

A Prism project file (or PZF file) contains all the pieces of a project. Each Prism project has basic five sections: [Data tables](#), [Info sheets](#), [Results](#), [Graphs](#), and [Layouts](#). Each section can have up to 500 sheets. The term 'sheet' refers to an individual data table, info page, set of results, graph or page layout.



Project1:Column: Entering replicate data - GraphPad Prism 8.0.1 (244)

File Edit View Insert Change Arrange Family Window Help

Prism File Sheet Undo Clipboard Analysis Change Import Draw Write Text Export Print Send LA Help

Search...

- ▼ Data Tables
  - Column: Entering replicate data
  - + New Data Table...
- ▼ Info
  - Project info 1
  - + New Info...
- ▼ Results
  - + New Analysis...
- ▼ Graphs
  - Column: Entering replicate data
  - + New Graph...
- ▼ Layouts
  - + New Layout...

Family

- Column: Entering replicate data
- Column: Entering replicate data

	Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H	Group I	Group J	Group K
	Untreated	Placebo	Treated	Title	Title	Title	Title	Title	Title	Title	Title
1	3.4		4.2								
2	4.3	2.3	7.8								
3	3.0	5.2									
4	3.9	4.5	7.6								
5	4.1	3.1									
6		5.0									
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											

**How the data are organized**  
The three columns define three groups. The replicate values are stacked in each column

**How to graph these data**  
Look in the navigator to the left and find the Graphs folder. Click on the graph in bold. Prism will present several choices for graphing these data. With small numbers of values (like here), it is best to create a scatter plot that shows every value.

**Notes**

- Unlike many statistics programs, Prism does not define groups using a grouping variable. Instead, the groups are defined by which column each value is entered into.
- When you make a graph, you can choose to plot a column scatter graph (showing each value) or various kinds of graphs that summarize the data (bar graphs, box-and-whisker graphs, violin plots).

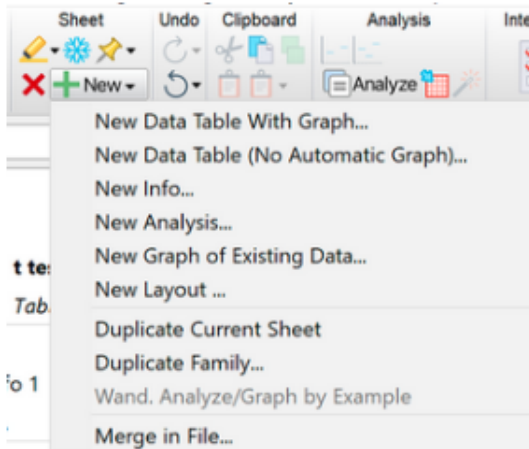
? More about Column tables

Column: Entering replicate data Row 1, A: Untreated

[https://www.graphpad.com/guides/prism/8/user-guide/adding\\_new\\_sheets\\_to\\_your\\_proj.htm?q=Adding+new+sheets+to+your+project](https://www.graphpad.com/guides/prism/8/user-guide/adding_new_sheets_to_your_proj.htm?q=Adding+new+sheets+to+your+project)

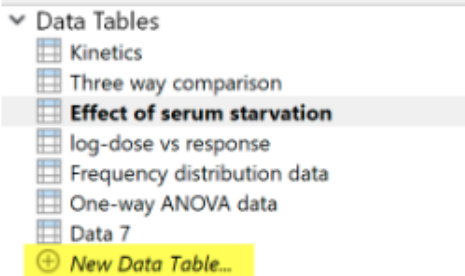
## Adding new sheets to your project

Prev Next



Click the New button on the Sheet toolbar and choose what kind of sheet you want to create. The first choice is used the most. It adds a new empty [data table and linked graph](#) to your project. You can also add a New Data Table (No Automatic Graph), a [new Info sheet](#), a [new graph of an existing data table](#), or a [new layout](#). To be complete, this menu also lets you add a new analysis, but you'll find it easier to start at a data table or graph and [click the Analyze button](#).

You can also add a new table, analysis, graph or layout by a link in the navigator.



You can also add to your project by [duplicating a family of sheets](#) or by [cloning a graph](#).





[http://www.graphpad.com/guides/prism/8/user-guide/index.htm?tips\\_for\\_using\\_prism.htm](http://www.graphpad.com/guides/prism/8/user-guide/index.htm?tips_for_using_prism.htm)

## Tips for using Prism

Expand all tips

Collapse all tips

### Getting started tips

- ▣ [Before using Prism for the first time...](#)
- ▣ [Pay attention to the distinctions between the eight kinds of data tables](#)
- ▣ [Learn about the various approaches to making bar graphs](#)
- ▣ [Recycle your work](#)

### Navigation tips

- ▣ [View all sheets in the gallery](#)
- ▣ [Jump between related sheets using the Family navigator](#)
- ▣ [Ping pong \(flip back and forth\) between two sheets](#)
- ▣ [Highlight or annotate sheets you want to return to](#)
- ▣ [Search for sheets](#)

### Graphing Tips

- ▣ [Use the Format Graph and Format Axes dialogs](#)
- ▣ [Change the color \(or size..\) of just one symbol or bar](#)
- ▣ [Use Prism Magic to make graphs consistent](#)
- ▣ [Change colors by picking a color scheme](#)
- ▣ [Prism can plot error bars directly from raw data](#)
- ▣ [To add data to a graph, drag a table from the Navigator and drop onto the graph](#)

### Data analysis tips

- ▣ [Don't run the analysis again -- change the analysis choices](#)
- ▣ [Use the 'Learn' and 'Analysis checklist' buttons](#)
- ▣ [Paste results onto your graph](#)
- ▣ [Create chains of analyses, without copying and pasting](#)

### Exporting tips

## 📖 Differences between Windows and Mac versions of Prism

- 📄 File compatibility between Windows and Mac
- 📖 Unique to Prism Windows
  - 📄 Features in Windows but not Mac
  - 📄 Working with multiple files in Prism Windows
  - 📄 Automatic back up of Prism files
- 📖 Unique to Prism Mac
  - 📄 Features in Mac but not Windows
  - 📄 Working with multiple files in Prism Mac
  - 📄 Overview: Resume and Versions in Prism Mac
  - 📄 Details: Resume and Versions in Prism Mac
  - 📄 File..Save As command in Mac

## Excel issues with Mac?

### Features in Windows but not Mac

⏪ Prev ⏩ Next

- [Object linking and embedding](#). This means you can paste a Prism graph into another program and double-click to edit in Prism. We wish we could allow this to work on Prism Mac, but Apple and Microsoft simply do not provide the "hooks" Prism would need.
- [Excel linking and embedding](#). This means you can paste Excel data into Prism, and double click to edit within Excel. We also wish this were possible with the Mac.
- Import and export EMF and WMF images. These formats are only used on Windows computers.
- Send to Word with a [single click](#).
- High contrast printing, with no grays and no colors (choice on Preferences dialog).
- [Visual Basic for Applications \(VBA\) commands](#) to launch Prism and start a script.

# Getting Data into Prism



## Copy and paste

### Pasting data from Excel (Windows)

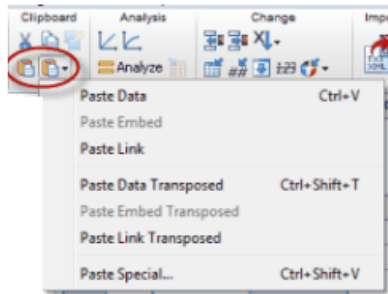
Prev Next

Select a range of Excel data and copy it to the clipboard. Switch to Prism, move the insertion point to the cell of a data table that will become the upper-left corner of the pasted data.

#### Pasting Excel data

Then click on the Paste  or Paste Special  button in the Prism Clipboard toolbar.

If you use the Paste button or the keyboard shortcut (Ctrl+V), you will only paste the values that have been copied to the clipboard. If you choose Paste Special, you have several options about how you want to insert your data into Prism's data table. These options are only available with Windows. If you use a Mac, skip to [this page](#).



- Paste Data -- Prism retains no link back to the Excel spreadsheet. The values will be pasted and nothing else. The only advantage to using this method is that it is conceptually simpler.
- Paste Embed -- You paste the selected data into the data table and also paste a copy of the entire spreadsheet file into your Prism project, even if it has several worksheets (tabs). This means you don't need to worry about keeping track of multiple files. You don't need to save the spreadsheet file separately (except as a backup). You can open Excel from within Prism, edit the data, and instantly update the analyses and graphs in Prism. A potential disadvantage of embedding is that you will end up with multiple copies of the same data. Another disadvantage is that your Prism files will be larger.
- Paste Link -- Paste the values into the data table but also create a link back to the Excel file. The link has two functions. It lets you trace (and document) the source of the data so you stay organized. It also is a live link. If you edit or replace the data within Excel, Prism will update the analyses and graphs. The Paste Link command is available only if you paste from an Excel workbook that has a name. If you are working with a new Excel file, you must first save it (thus giving it a name) before you can link its data into Prism. When in doubt, use Paste Link.
- Paste Data Transposed -- Converts data in Excel rows into columns in Prism, and vice versa.

I mostly use  
ctrl-c and ctrl-v

# Getting Data into Prism



## Copy and paste

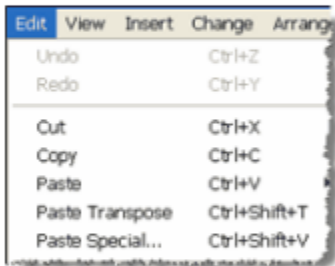
### Pasting data from Excel (Mac)

Prev Next

#### Pasting data from Excel

Select a range of Excel data and copy it to the clipboard. Switch to Prism and move the insertion point to the cell of the data table that will become the upper-left corner of the pasted data.

Then click on the Paste  or Paste Special  button in the Clipboard toolbar.



- Paste -- You will only paste the values that have been copied to the clipboard.
- Paste Transpose -- Converts data in Excel rows into columns in Prism, and vice versa.
- Paste Special -- You'll then see the [Import and Paste Special](#) dialog, where you can choose whether to filter and rearrange the data (such as transposing).

Note that it is not possible to Paste Embed or Paste Link to Excel files on the Macintosh. Those features only work on Windows.

let's play



## Dataset: Cholesterol and Personality Type

In California, between 1960-61, 3,154 middle aged men were studied

Is behavior pattern associated with heart disease?

Cholesterol measured in 40 of the heaviest men

- 20 Type A (aggressive and ambitious)

- 20 Type B (relaxed and non-competitive)

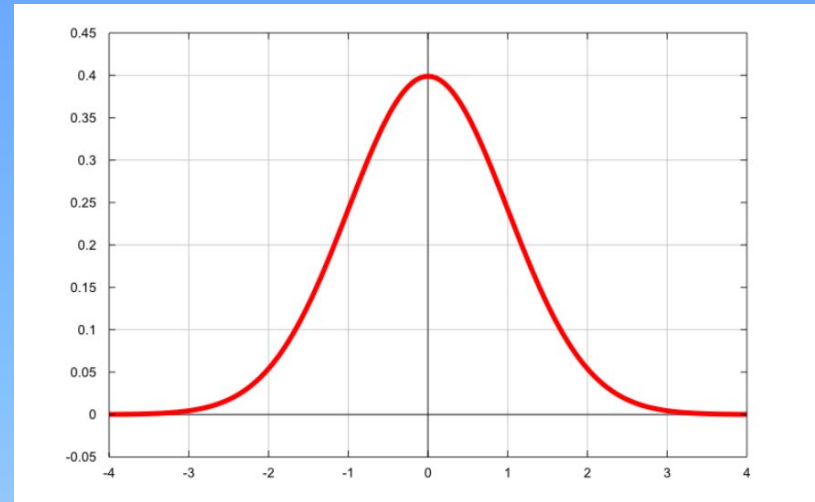
The best first step is to graph the data (after you get the data into your statistical program)

## Why graph the data?

Check for data distribution

Is it normal (symmetrical)?

## The Normal Distribution

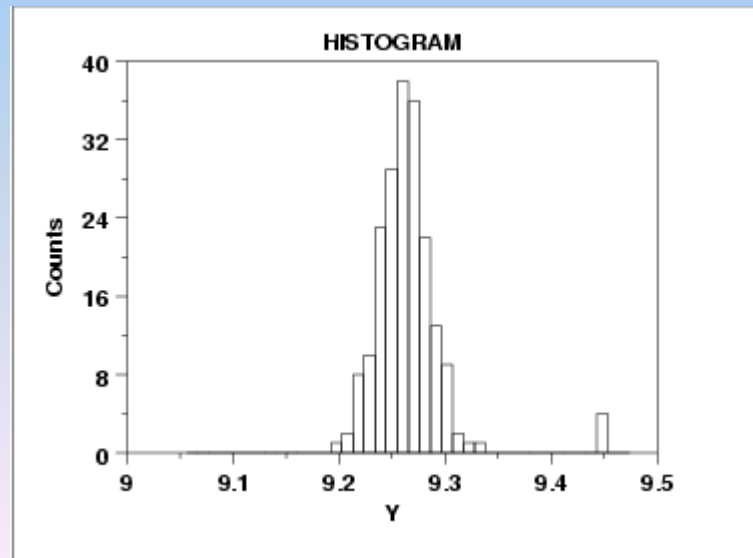


We will talk about the normal distribution, the assumption of normality, the CLT, *etc.* later

Check for “outliers”

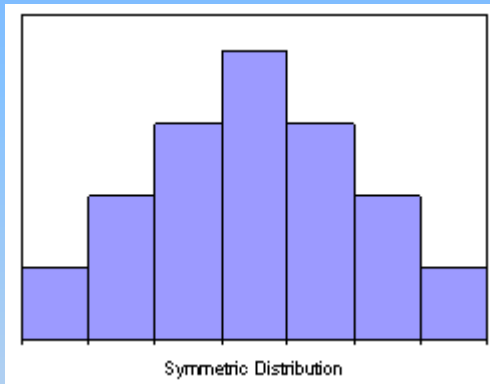
Are they real data?

Are they errors?



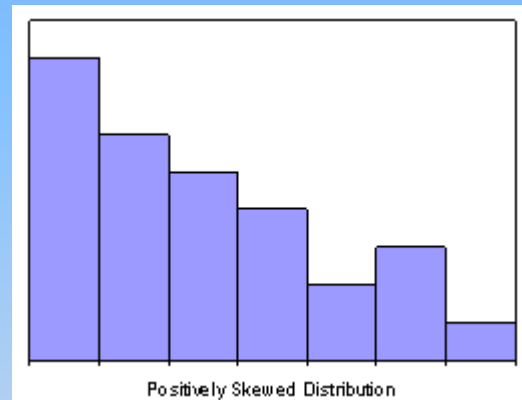
# Data Distributions

Frequency histograms are useful for assessing whether distributions are symmetrical or skewed (but be careful of bin size)



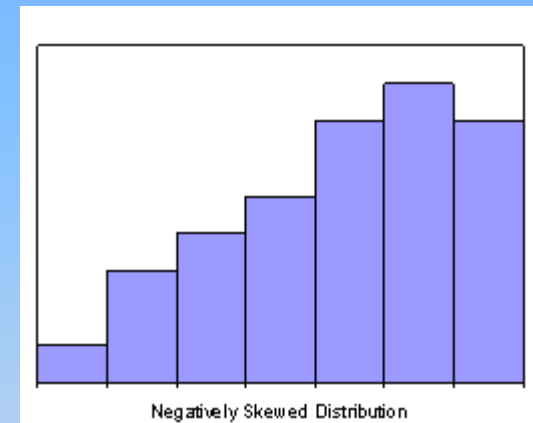
No Skew

Symmetrical distribution



Right or Positively  
Skewed

Long tail to the right

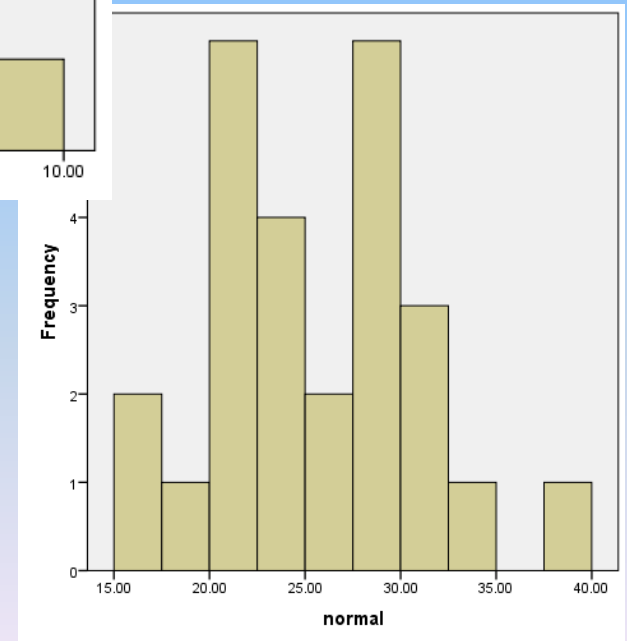
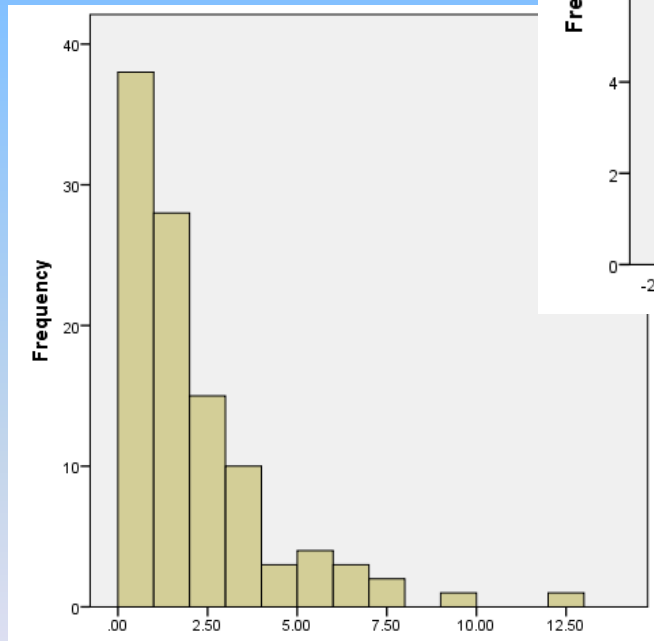
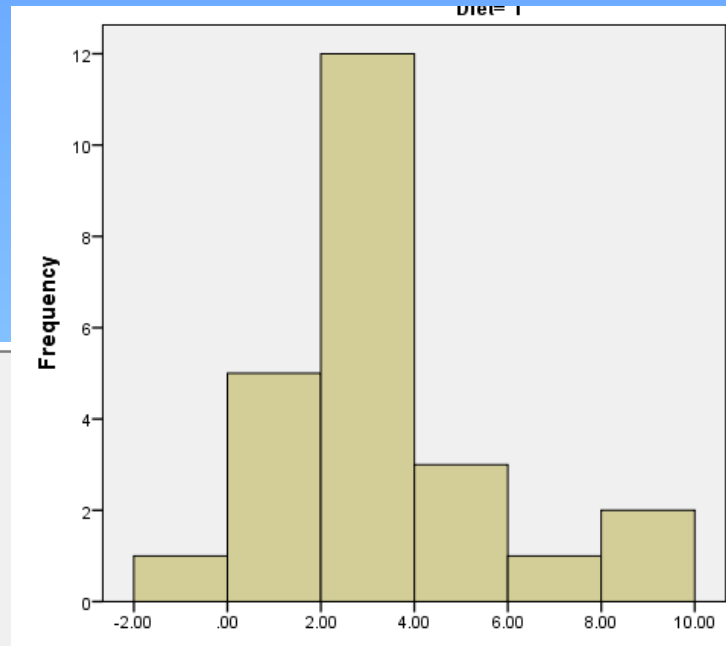


Left or Negatively  
Skewed

Long tail to the left

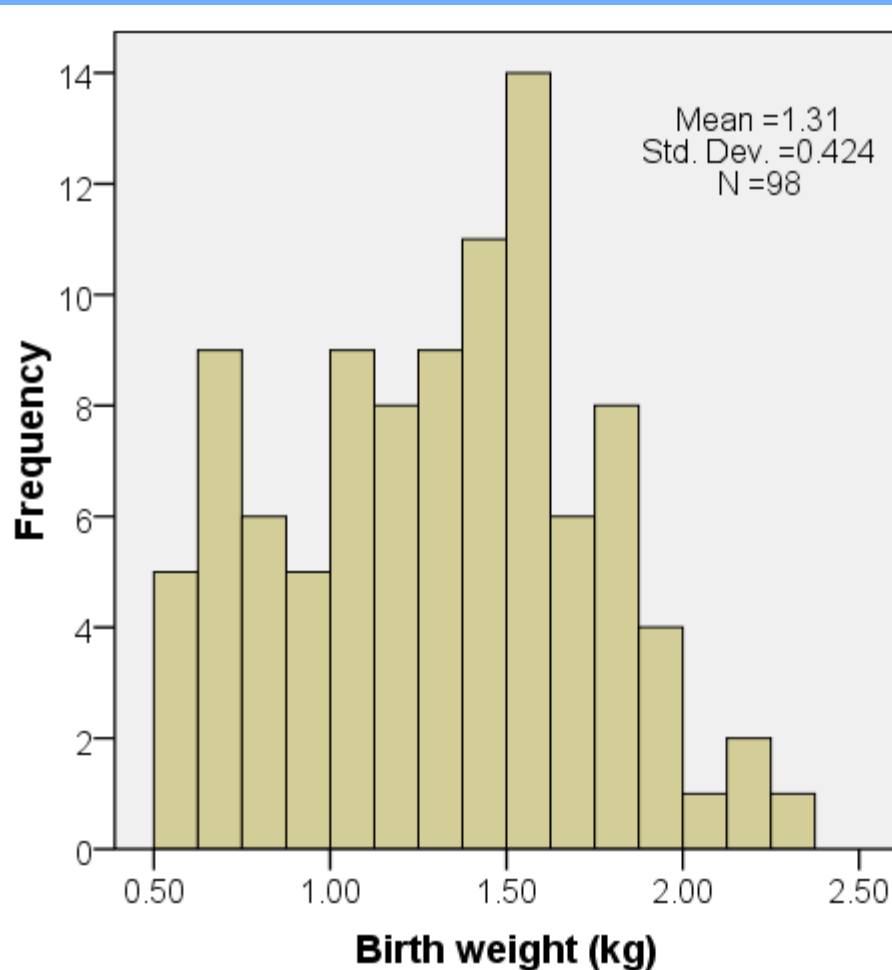


Do these look symmetrical?



# Frequency Histograms are Used for Continuous Variables

## Birth Weights (kg)



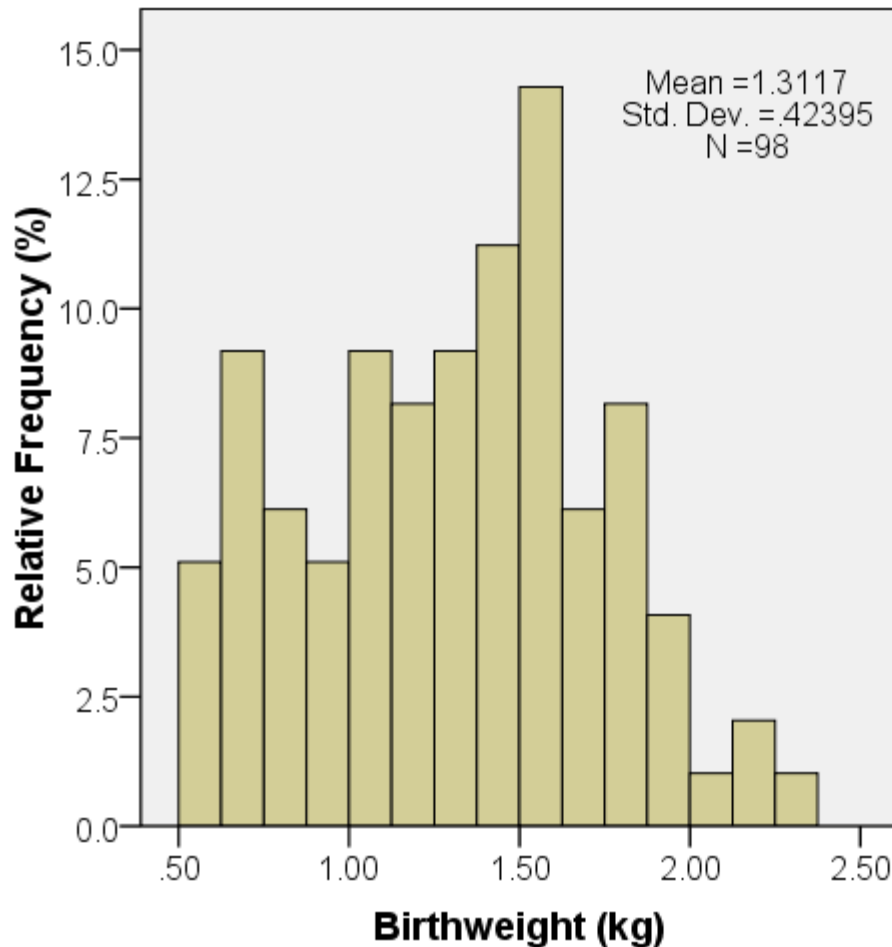
Easy to construct

The underlying continuous variable is divided into a number of categories of equal width (bin width).

The order of the bars cannot be changed.

# Frequency vs. Relative Frequency

## Birth Weights (kg)



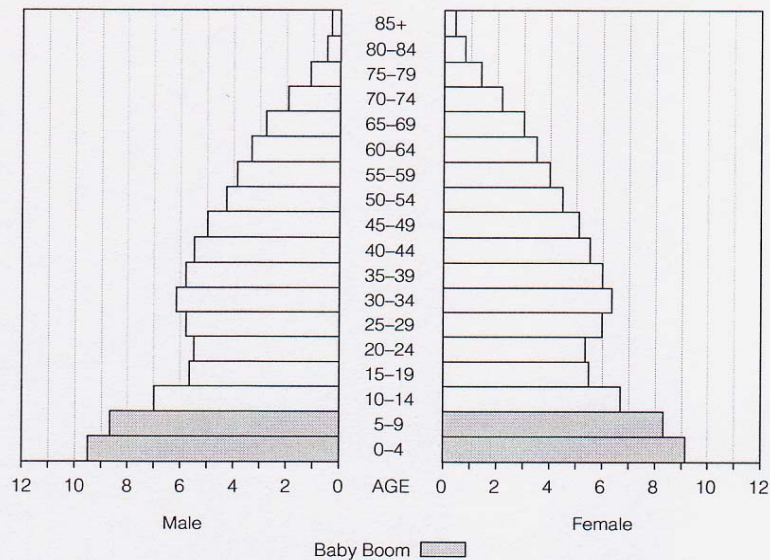
The Y-axis can show a *count* of cases falling into each category

Or, it can show the percentage of all cases falling into each category, as in this graph.

The percentage (or proportion) of cases falling into each category is called the relative frequency.

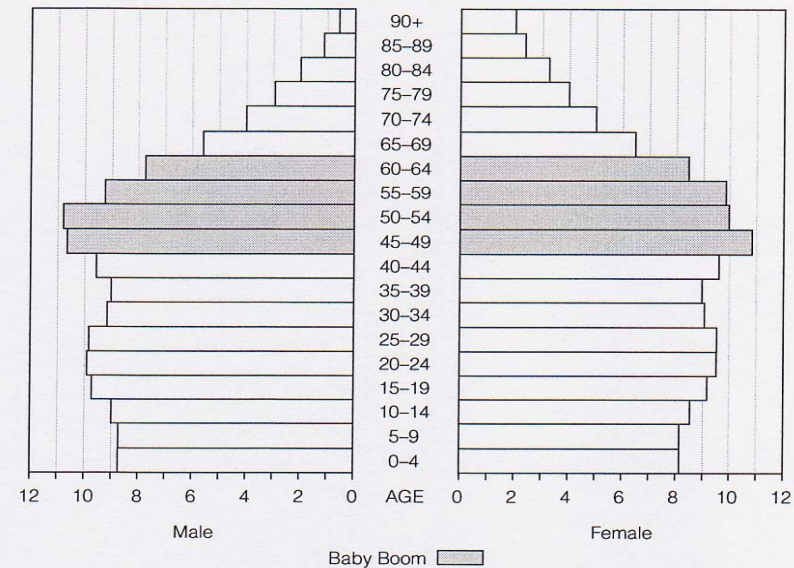
# An Example: US Population by Age and Gender

Figure 3.10 **U.S. Population by Gender and Age, 1955 (in millions)**

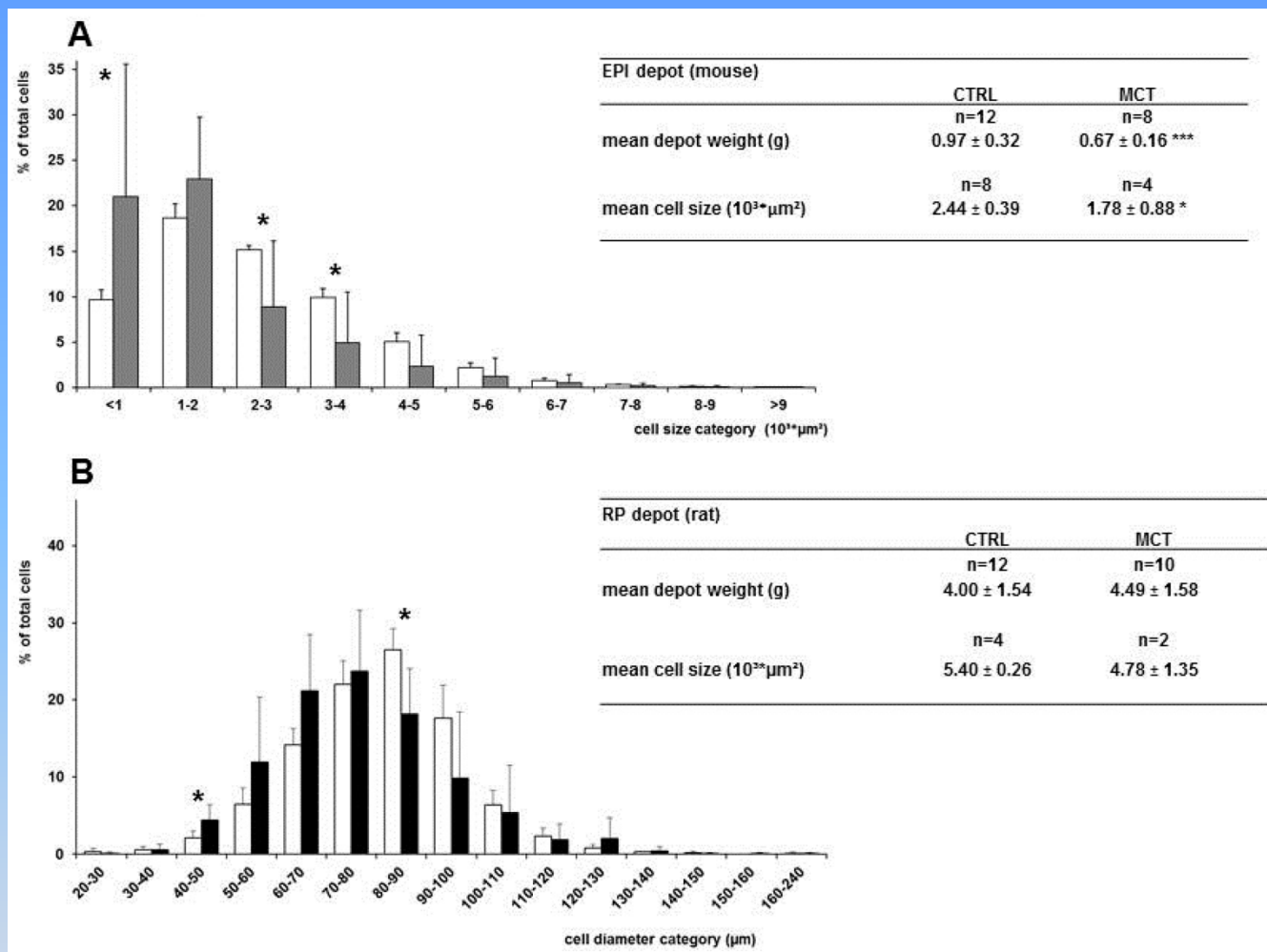


Source: U.S. Bureau of the Census, *Current Population Reports*, 1992, P23-178.

Figure 3.11 **U.S. Population by Gender and Age, 2010 (in millions)**



Source: U.S. Bureau of the Census, *Current Population Reports*, 1992, P23-178.



*Outliers?*

**Figure 4.** Panel (A) Frequency distribution of EPI adipocyte cell size in CTRL (n = 8; white bars) and MCT (medium-chain triglycerides, n = 4; grey bars) mice on PN day 98; Panel (B) Frequency distribution of RP adipocyte cell size in CTRL (n = 4; white bars) and MCT (n = 2; black bars) rats on PN day 98. The tables show the mean depot weight and cell size. Values are means ± SD; \* p < 0.05; \*\*\* p < 0.001 vs. CTRL; EPI, epididymal fat; RP, retroperitoneal fat.

## Cholesterol and personality type

## Create a frequency histogram by type and all types together

We will play with bin width and graph formats

	A	B	C	D
1	num	TypeA	TypeB	
2	1	233	200	
3	2	291	185	
4	3	312	263	
5	4	250	246	
6	5	246	224	
7	6	197	212	
8	7	268	188	
9	8	224	250	
10	9	239	148	
11	10	239	169	
12	11	254	226	
13	12	276	175	
14	13	234	242	
15	14	181	252	
16	15	248	153	
17	16	252	183	
18	17	202	137	
19	18	218	202	
20	19	212	194	
21	20	325	213	

**New table & graph**

XY

Column

Grouped

Contingency

Survival

Parts of whole

Multiple variables

Nested

**Existing file**

Open a file

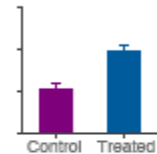
LabArchives

Clone a graph

Graph portfolio

Column tables have one grouping variable, with each group defined by a column

	A	B
	Control	Treated
	Y	Y
1		
2		

[Learn more](#)**Data table:**

- ☒ Enter or import data into a new table  
☐ Start with sample data to follow a tutorial

**Options:**

- ☒ Enter replicate values, stacked into columns  
☐ Enter paired or repeated measures data - each subject on a separate row  
☐ Enter and plot error values already calculated elsewhere

Enter:

Mean, SD, N



Prism Tips

Cancel

Create

Project1:Data 1 - GraphPad Prism 8.2.0 (435)

FileEditViewInsertChangeArrangeFamilyWindowHelp

PrismFileSheetUndoClipboardAnalysisChangeImportDrawWriteTextExportPrintSendLAHelp

Search...

▼Data Tables

Data 1

New Data Table...

▼Info

Project info 1

New Info...

▼Results

New Analysis...

▼Graphs

Data 1

New Graph...

▼Layouts

New Layout...

Family

Data 1

Data 1

Group AGroup BGroup CGroup DGroup EGroup FGroup GGroup HGroup IGroup JGroup KGroup LGroup M

TitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitleTitle

12345678910111213141516171819

12345678910111213141516171819

◀▶🔍🔍

📊📋📄📈📉

Data 1

Row 1, Column A

🔍🔍

Decrease text size





# But we also wanted to do a distribution for all cholesterol combined

Project1:Data 1 - GraphPad Prism 8.2.0 (435)

File Edit View Insert Change Arrange Family Window Help

Prism File Sheet Undo Clipboard Analysis Change Import Draw Write Text Export Print Send LA Help

Search...

▼ Data Tables

- Cholesterol by type
- + New Data Table...

▼ Info

- Project info 1
- + New Info...

▼ Results

- + New Analysis...

▼ Graphs

- Cholesterol by type
- + New Graph...

▼ Layouts

- + New Layout...

Family

- Cholesterol by type
- Cholesterol by type

	Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H	Group I	Group J	Group K	Group L	Group
	TypeA	TypeB	Title	Title	Title	Title	Title	Title	Title	Title	Title	Title	Title
1	233	200											
2	291	185											
3	312	263											
4	250	246											
5	246	224											
6	197	212											
7	268	188											
8	224	250											
9	239	148											
10	239	169											
11	254	226											
12	276	175											
13	234	242											
14	181	252											
15	248	153											
16	252	183											
17	202	137											
18	218	202											
19	212	194											

Cholesterol by type

Row 1, A: TypeA

Drop down menu: Choose “New Data Table with Graph”

Project1:All Cholesterol - GraphPad Prism 8.2.0 (435)

File Edit View Insert Change Arrange Family Window Help

Prism File Sheet Undo Clipboard Analysis Change Import Draw

Search...

▼ Data Tables »

- Cholesterol by type
- All Cholesterol**
- + New Data Table...

▼ Info »

- Project info 1
- + New Info...

▼ Results »

- + New Analysis...

▼ Graphs »

- Cholesterol by type
- All Cholesterol**
- + New Graph...

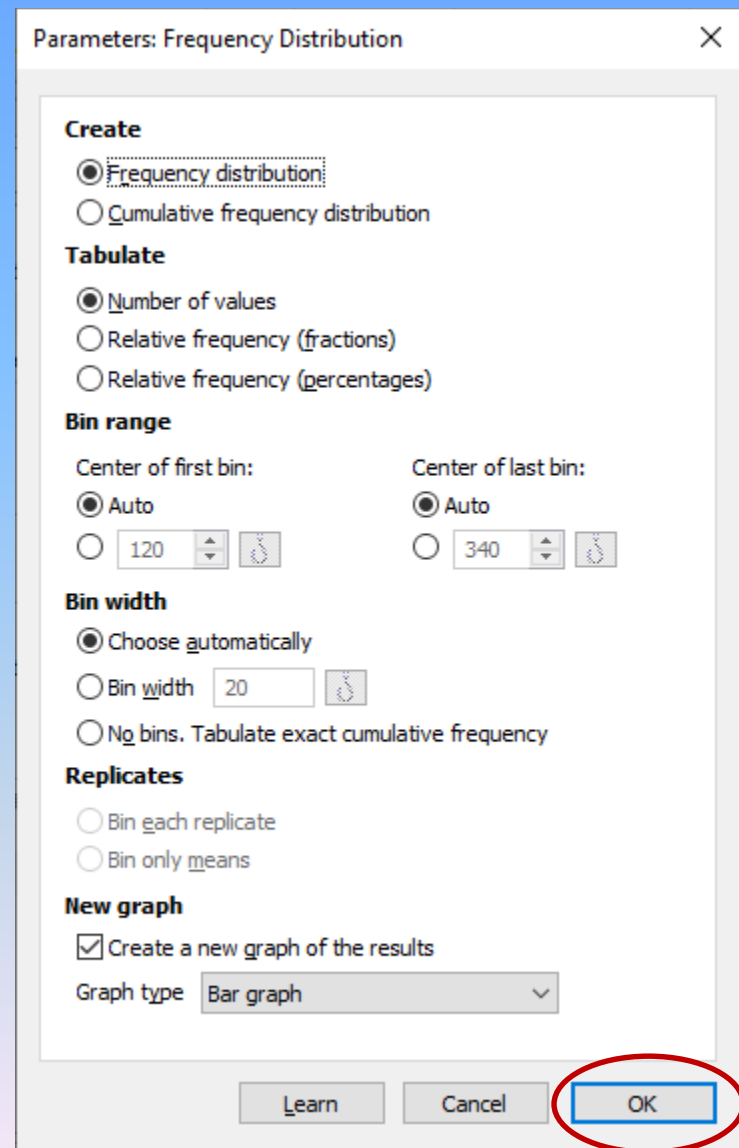
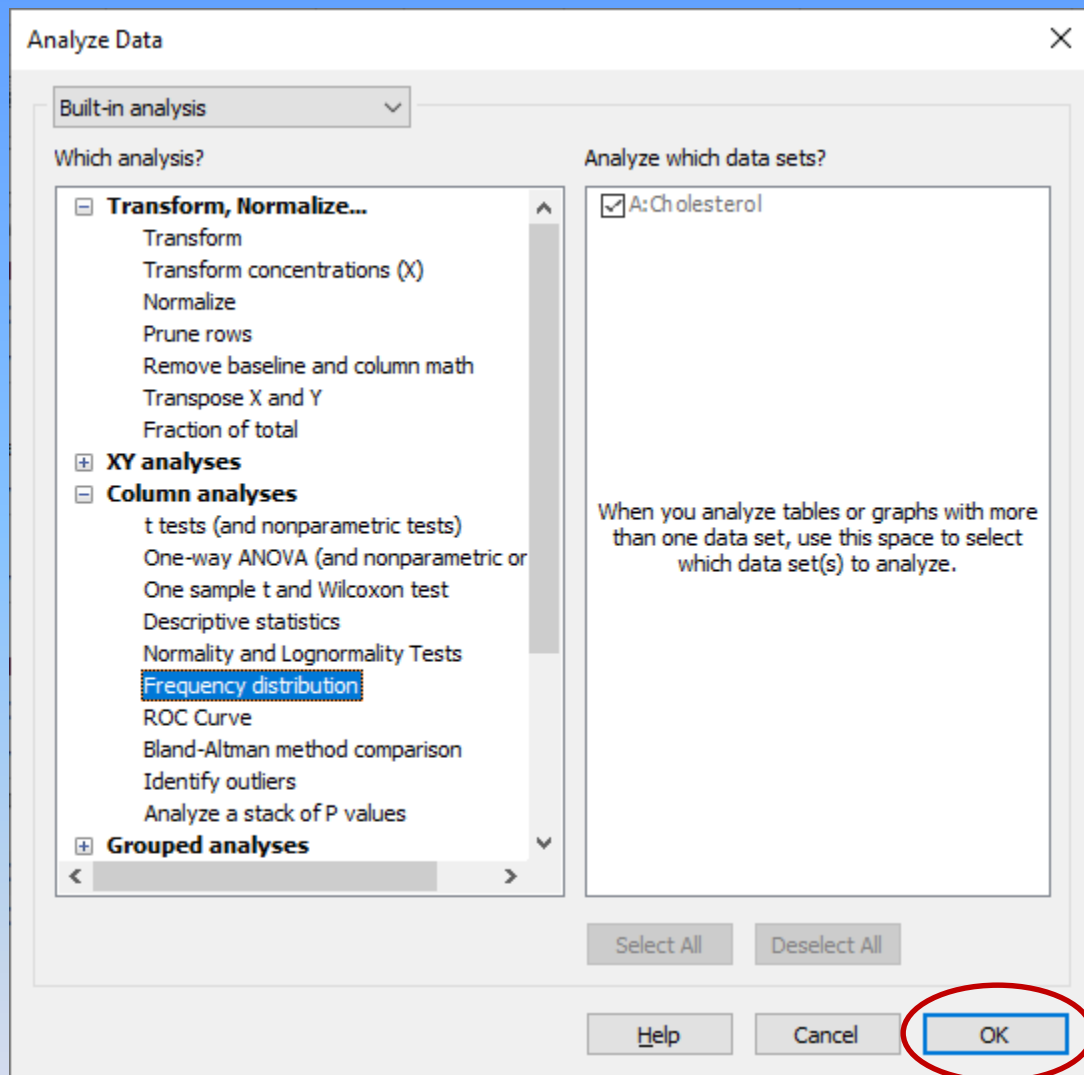
Family »

- All Cholesterol**
- All Cholesterol

Cholesterol

	Group A	Group B	Group C	Group D	Group E
	Cholesterol	Title	Title	Title	Title
1	233				
2	291				
3	312				
4	250				
5	246				
6	197				
7	268				
8	224				
9	239				
10	239				
11	254				
12	276				
13	234				
14	181				
15	248				
16	252				
17	202				
18	218				
19	212				

← → 🔍 📊 ⓘ 📄 📈 📉 All Cholesterol 🔗



# Analyze Data



Built-in analysis

Which analysis?

- ☐ Transform, Normalize...
  - Transform
  - Transform Concentrations (X)
  - Normalize
  - Prune rows
  - Remove baseline and column math
  - Transpose X and Y
  - Fraction of total
- ☒ XY analyses
- ☐ Column analyses
  - t tests (and nonparametric tests)
  - One-way ANOVA (and nonparametric)
  - Column statistics
  - Frequency distribution**
  - ROC Curve
  - Bland-Altman method comparison
  - Correlation
  - Identify outliers
  - Analyze a stack of P values
- ☒ Grouped analyses
- ☒ Contingency table analyses
- ☒ Survival analyses
- ☒ Parts of whole analyses
- ☒ Generate curve
- ☒ Simulate data
- ☒ Recently used

Analyze which data sets?

☒ A:Cholesterol

When you analyze tables or graphs with more than one data set, use this space to select which data set(s) to analyze.

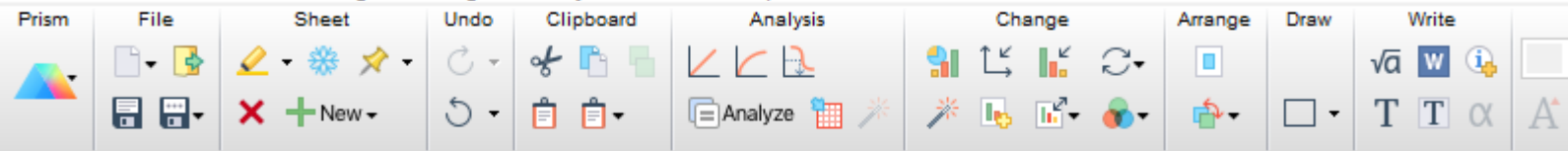
Select All

Deselect All

Help

Cancel

**OK**



Search...

+ New Data Table...

▼ Info

i Project info 1

+ New Info...

▼ Results

Histogram of All Cholesterol

+ New Analysis...

▼ Graphs

Cholesterol by type

All Cholesterol

Histogram of All Cholesterol

+ New Graph...

▼ Layouts

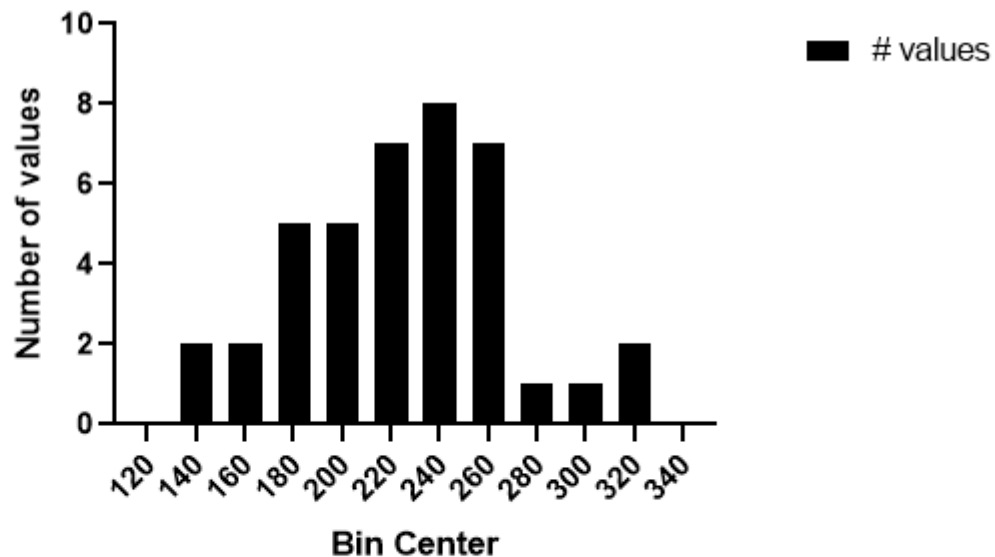
Family

All Cholesterol

Histogram

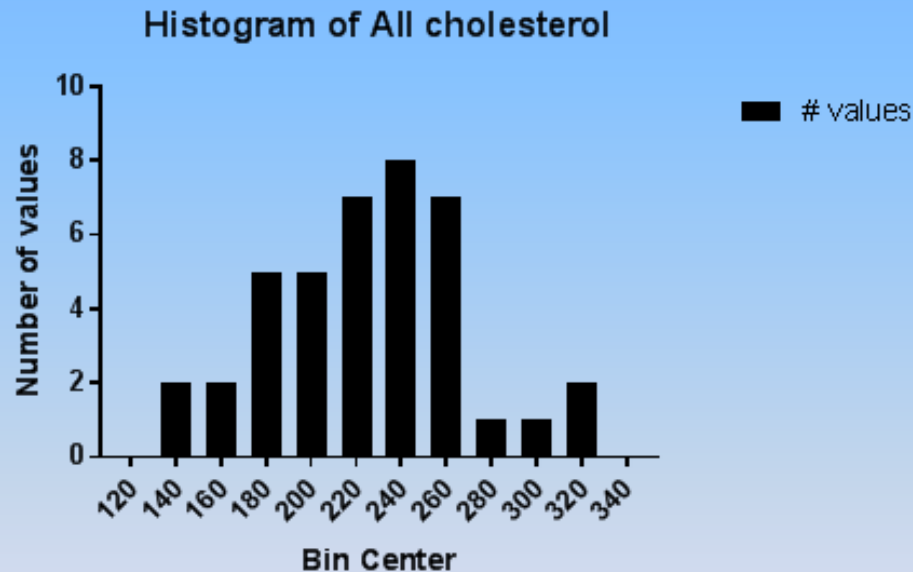
Histogram of All Cholesterol

## Histogram of All Cholesterol



Right click on graph to “Copy Graph” or press “Control + c”. Then paste (or “Control + v”) the graph into Word.

Or use the “Send to PowerPoint” button



# Go back to the Data Tables to do the Distribution by Type

Project1:Cholesterol by type - GraphPad Prism 8.2.0 (435)

File Edit View Insert Change Arrange Family Window Help

Prism File Sheet Undo Clipboard Analysis Change Import Draw Write

Search...

▼ Data Tables

- Cholesterol by type
- All Cholesterol
- New Data Table...

▼ Info

- Project info 1
- New Info...

▼ Results

- Histogram of All Cholesterol
- New Analysis...

▼ Graphs

- Cholesterol by type
- All Cholesterol

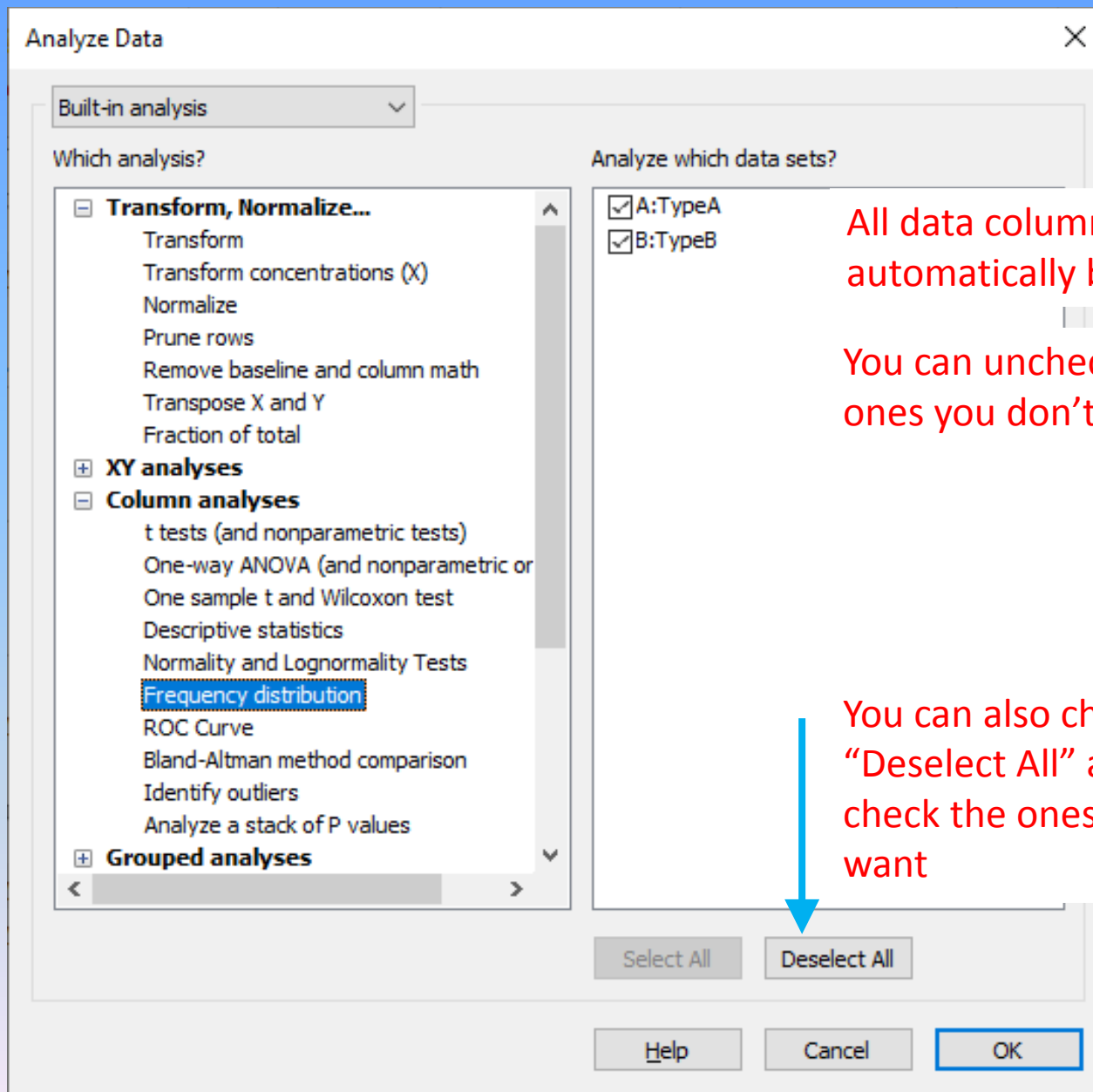
Family

- Cholesterol by type
- Cholesterol by type

	Group A	Group B	Group C	Group D	Group E	Group F	Gr
	TypeA	TypeB	Title	Title	Title	Title	
1	233	200					
2	291	185					
3	312	263					
4	250	246					
5	246	224					
6	197	212					
7	268	188					
8	224	250					
9	239	148					
10	239	169					
11	254	226					
12	276	175					
13	234	242					
14	181	252					
15	248	153					
16	252	183					
17	202	137					
18	218	202					
19	212	194					

Cholesterol by type Row 1,





All data columns (sets) will automatically be checked

You can uncheck the ones you don't want

You can also choose "Deselect All" and check the ones you want

Parameters: Frequency Distribution

**Create**

- ☒ Frequency distribution  
☐ Cumulative frequency distribution

**Tabulate**

- ☒ Number of values  
☐ Relative frequency (fractions)  
☐ Relative frequency (percentages)

**Bin range**

Center of first bin:

Center of last bin:

- ☒
- Auto

- ☒
- Auto

☐ 120☐ 340**Bin width**

- ☒ Choose automatically  
☐ Bin width 20  
☐ No bins. Tabulate exact cumulative frequency

**Replicates**

- ☐ Bin each replicate  
☐ Bin only means

**New graph**

- ☒ Create a new graph of the results

Graph type Bar graph. Separated

Learn

Cancel

OK

Parameters: Frequency Distribution

**Create**

- ☒ Frequency distribution  
☐ Cumulative frequency distribution

**Tabulate**

- ☒ Number of values  
☐ Relative frequency (fractions)  
☐ Relative frequency (percentages)

**Bin range**

Center of first bin:

Center of last bin:

- ☒
- Auto

- ☒
- Auto

☐ 120☐ 340**Bin width**

- ☒ Choose automatically  
☐ Bin width 20  
☐ No bins. Tabulate exact cumulative frequency

**Replicates**

- ☐ Bin each replicate  
☐ Bin only means

**New graph**

- ☒ Create a new graph of the results

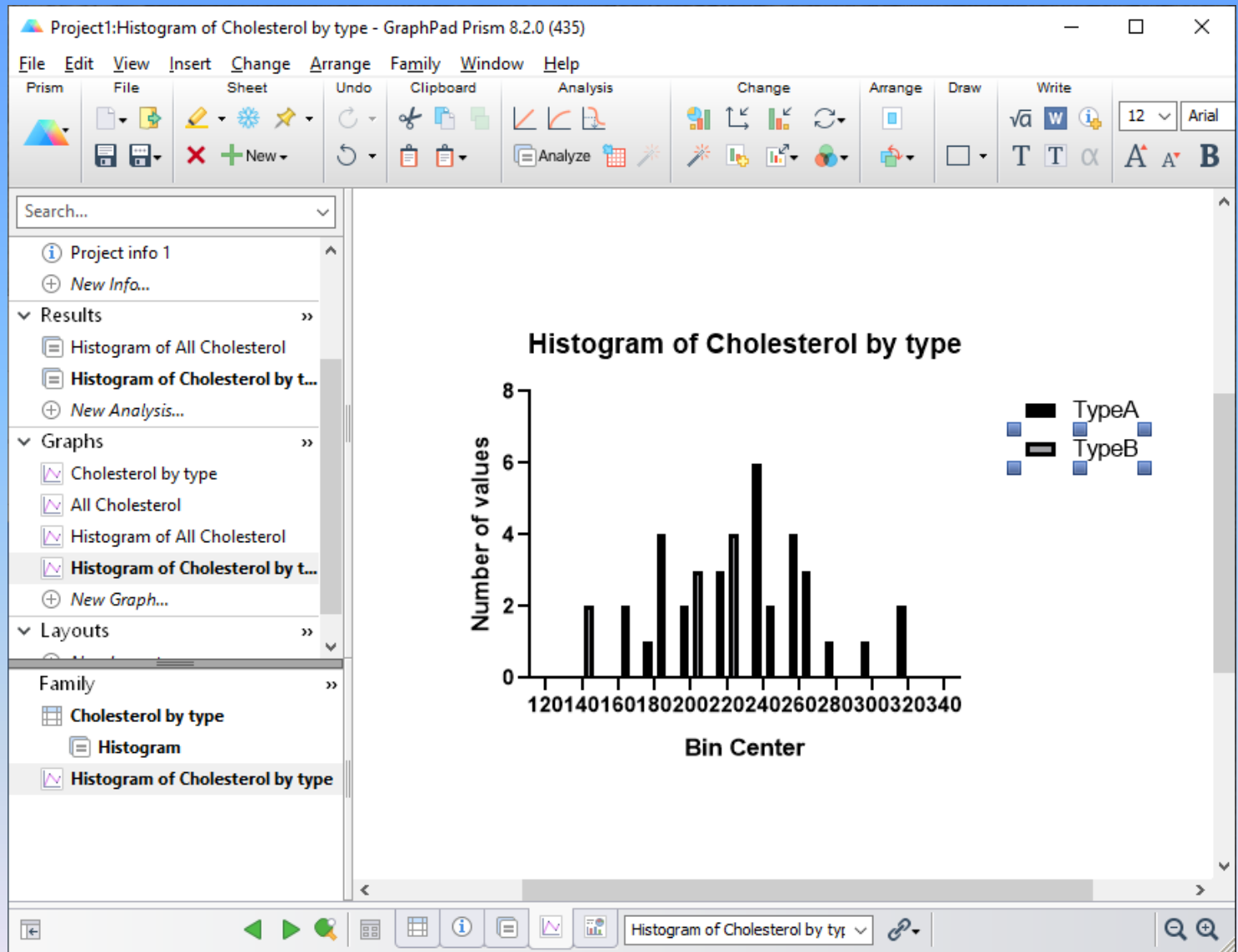
Graph type Bar graph. Interleaved

Learn

Cancel

OK

# Interleaved



Double-click on the TypeB to start customizing the graph

Format Graph

Appearance | Data Sets on Graph | Graph Settings | Annotations

Data set: Histogram of Cholesterol by type:Frequency distri < > Global

Style

Appearance: Bar Plot: Mean

Bars and boxes

Fill: Border: 2 pt Border color:

☐ Fill pattern Color:

☐ Symbols

Color: Shape: Border color:

Size: Border thickness:

☒ Error bars

Color: Style: Dir.: Thickness: Auto (1/2 pt)

☐ Lines

Color: Thickness: Auto (1 pt) Line and error go:

Style: Pattern: Length:

Additional options

Plot data on: ☒ Left Y axis ☐ Right Y axis

☒ Show legend

☐ Revert legend to column title

☐ Label each point with its row title

Color: ☒ Auto

Change Graph Type... Help Cancel Apply OK

Format Graph

Appearance | Data Sets on Graph | Graph Settings | Annotations

Data set: Histogram of Cholesterol by type:Frequency distri < > Global

Style

Appearance: Bar Plot: Mean

Bars and boxes

Fill: Border: 2 pt Border color:

☐ Fill pattern Color:

☐ Symbols

Color: Shape: Border color:

Size: Border thickness:

☒ Error bars

Color: Style: Dir.: Above Thickness: Auto (1/2 pt)

☐ Lines

Color: Thickness: Auto (1 pt) Line and error go:

Style: Pattern: Length:

Additional options

Plot data on: ☒ Left Y axis ☐ Right Y axis

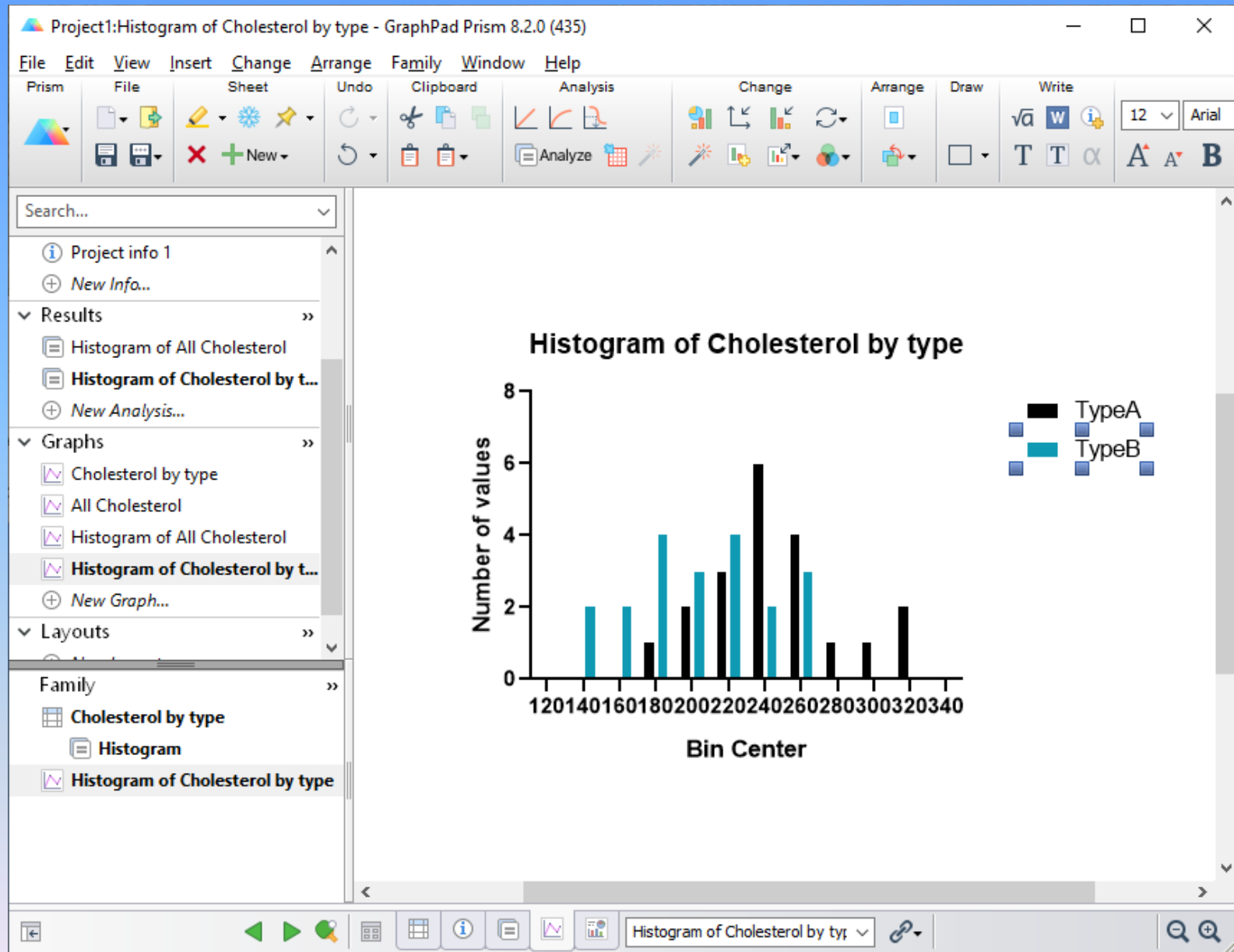
☒ Show legend

☐ Revert legend to column title

☐ Label each point with its row title

Color: ☒ Auto

Change Graph Type... Help Cancel Apply OK



Double click on bars in graph to continue customizing

Format Graph

Appearance Data Sets on Graph **Graph Settings** Annotations

**Direction**

☒ Vertical ☐ Horizontal

**Baseline**

☒ Automatic

☐ Bars begin at Y=

☐ Hide baseline

**Spacing (as percent of space used for data)**

Blank/missing cells:  % (0 % = no gap)

Between adjacent data:  % Before the first column:  %

Additional gap between groups:  % After the last column:  %

**Discontinuous axis**

☒ When the axis is discontinuous, also make the bars (or connecting lines) discontinuous

**Scatter plot appearance**

☐ Standard (Width of distribution of points proportionate to the number of points at that Y value. Best to represent data distribution)

☐ Classic (Prioritizes minimizing overlap of individual data points over representing shape of data distribution. May cause "smiles" in the scatterplot)

☐ Expanded (Prioritizes minimizing overlap of individual data points over representing shape of data distribution. Doesn't cause visual pattern)

**Format of individual bars**

To format a bar differently than the rest, exit this dialog, right click on the bar and choose "Format This Bar".

☐ Remove all the individual formatting. Revert all bars to the format of their data sets.

**Shape of legend key**

☒ Rectangle ☐ Square

Change Graph Type... Help Cancel Apply OK

Format Graph

Appearance Data Sets on Graph **Graph Settings** Annotations

**Direction**

☒ Vertical ☐ Horizontal

**Baseline**

☒ Automatic

☐ Bars begin at Y=

☐ Hide baseline

**Spacing (as percent of space used for data)**

Blank/missing cells:  % (0 % = no gap)

Between adjacent data:  % Before the first column:  %

Additional gap between groups:  % After the last column:  %

**Discontinuous axis**

☒ When the axis is discontinuous, also make the bars (or connecting lines) discontinuous

**Scatter plot appearance**

☐ Standard (Width of distribution of points proportionate to the number of points at that Y value. Best to represent data distribution)

☐ Classic (Prioritizes minimizing overlap of individual data points over representing shape of data distribution. May cause "smiles" in the scatterplot)

☐ Expanded (Prioritizes minimizing overlap of individual data points over representing shape of data distribution. Doesn't cause visual pattern)

**Format of individual bars**

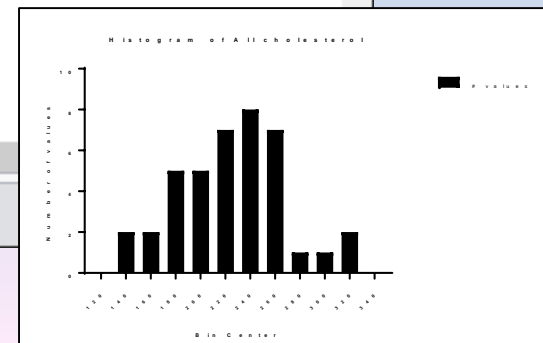
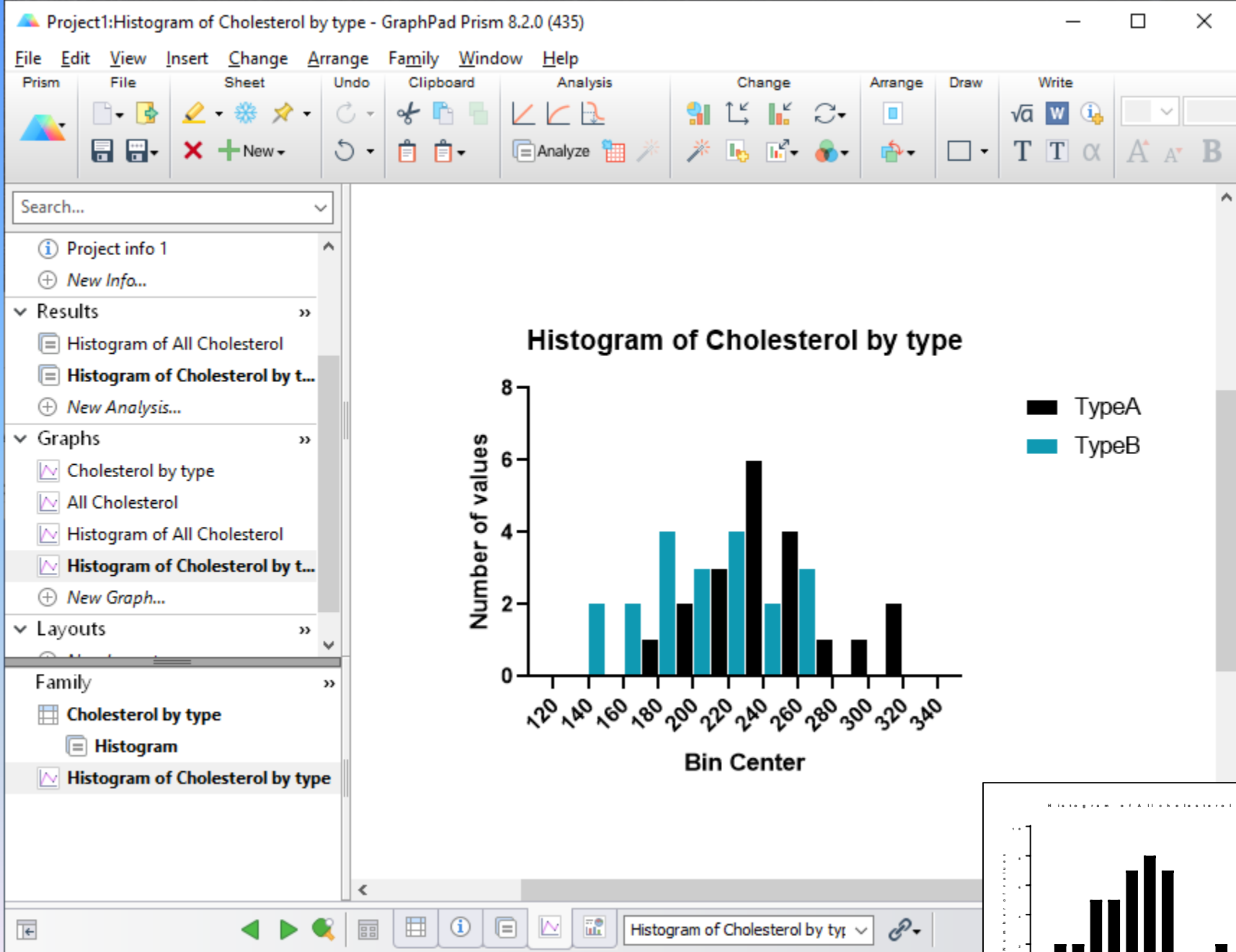
To format a bar differently than the rest, exit this dialog, right click on the bar and choose "Format This Bar".

☐ Remove all the individual formatting. Revert all bars to the format of their data sets.

**Shape of legend key**

☒ Rectangle ☐ Square

Change Graph Type... Help Cancel Apply OK



# GOOD DATA MANAGEMENT



# A bad example....

[illegible]

# Usual Data Structure in Excel

One variable per  
column

	A	B	C	D
	Subject ID	Gender	Year of study	Height
1	1	Male	1	170
2	2	Female	2	160
3	3	Female	3	165
4	4	Male	PG	175
5	5	Female	3	168

One row per subject  
One record

# Good Data Entry Practices

- Variable names should be:
  - Short but informative
  - Have consistent nomenclature
- Missing data should be left blank (create an explanatory column)
  - DO NOT use “99” or NA for missing data
- Do not mix text and numbers in the same column
- Make sure the dataset is complete before sending it off to be analyzed.
- Provide a key along with the data
  - Define numerical coding such as race categories or gender
- Make a copy of the original data and work on a copy
  - You can always go back the original data if you corrupt the data you are working on

NIHfundingByDiseaseBurden1999 - Microsoft Excel

File Home Insert Page Layout Formulas Data Review View JMP Acrobat SAS ?

Clipboard Font Alignment Number Styles Cells Editing

H29 fx 0.7

	A	B	C	D	E	F	G	H	I	J	K
1	num	conditiondisease	NIHfunds	incidence	prevalence	hospitaldays	mortality	yll	daly		
2	1	AIDS	1,410,925	138	1,303	329	42.0	979	1,267		
3	2	Breast cancer	381,880	523	2,262	525	44.0	426	1,421		
4	3	Dementia	304,411	959	7,082	841	34.0	99	2,866		
5	4	Diabetes mellitus	298,920	2,308	37,850	3,181	57.0	407	2,357		
6	5	Ischemic heart disease	269,100	2,216	8,976	11,815	481.0	2,973	8,876		
7	6	Alcohol abuse	256,600	11,085	18,092	2,563	7.0	98	4,690		
8	7	Injuries	198,700			15,218	149.0	3,109	8,608		
9	8	Dental and oral disorders	187,100			158	0.1	0.9	870		
10	9	Cirrhosis	169,800	168	1,238	869	31.0	380	1,584		
11	10	Depression	143,800	20,622	12,785	8,409	8.0	23	8,393		
12	11	Lung cancer	127,796	430	874	1,536	149.0	1,158	2,987		
13	12	Stroke	120,280	1,282	9,467	6,450	153.0	746	4,977		
14	13	Schizophrenia	111,479	162	7,164	4,129	0.4	2	2,249		
15	14	Colorectal cancer	105,525	499	1,926	1,600	58.0	391	1,626		
16	15	STD	102,583			69	0.3	28	404		
17	16	Prostate cancer	92,661	452	2,020	585	35.0	153	574		
18	17	Multiple sclerosis	82,800	16	461	149	2.0	28	236		
19	18	Asthma	81,600	3,594	15,919	1,820	5.0	65	1,236		

data code notes dataoriginal

Ready 100%

NIHfundingByDiseaseBurden1999 - Microsoft Excel

File Home Insert Page Layout Formulas Data Review View JMP Acrobat SAS

Clipboard Font Alignment Number Styles Cells Editing

B9 fx

	A	B	C	D	E	F	G	H
1		<b>Field</b>	<b>measure</b>	<b>note</b>				
2		NIHfunds	in thousands of dollars					
3		incidence	in thousands					
4		prevalence	in thousands					
5		hopitaldays	in thousands					
6		mortality	in thousands					
7		yll	in thousands	years of life lost				
8		daly	in thousands	disability-adjusted life-years	One disability-adjusted life-year is def			
9								
10								
11								
12		Estimates of incidence and prevalence were obtained from the Global Burden of Disease Study and apply to economies in 1990. All other estimates are specific to the United States for 1994						
13								
14								
15								
16								
17								
18								

data code notes dataoriginal

Ready 100%

# Assignment

Create data frequency (distribution) bar graphs using data from the SilvestreRoig paper (pdf of paper in Files/Assignment/Week02Sep04, the week the assignment is given out)

The Excel file with data for the assignment is in Files/Datasets/Week02Sep04

Assignment is in Files/Assignment/Week02Sep04

Due in Canvas next Wednesday before class

## General Comments on Homework 2 and Graphing for Assignments in General:

- Read the question: questions will often have multiple parts, make sure you understand what you are asked to do before you start making graphs. Point values of questions will tend to give you an idea of how much work is necessary.
- Make the colors in the graph so that if they were printed in black and white, you could still distinguish between different categories. Typically, one light and one dark color is a good idea.
- When graphing data, look at the end result to make sure it makes sense. Don't compare two different variables on the same axes. Make sure your x-axis and y-axis scales are logical to the data you are looking at.
- Watch the bin widths you use. Typically, the default is the best, but if you do change the bins make sure the bins chosen show the overall trend of the data in the best way. Look at the raw data to see trends.
- X, Y, and Main labels should describe all variables fully but in a concise way.