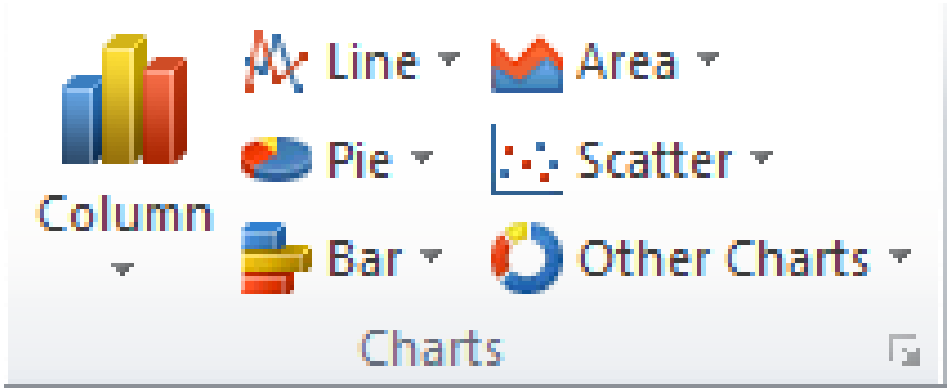
Northeastern University
College *of* Computer and Information Science

CS1100: Computer Science and Its Applications

Creating Graphs and Charts in Excel

Charts

- Data is often better explained through visualization as either a graph or a chart.
- Excel makes creating charts easy:
  - Column Charts
  - Pie Charts
  - Bar Graphs
  - Line Graphs
  - Area Graphs
  - Scatter Plots



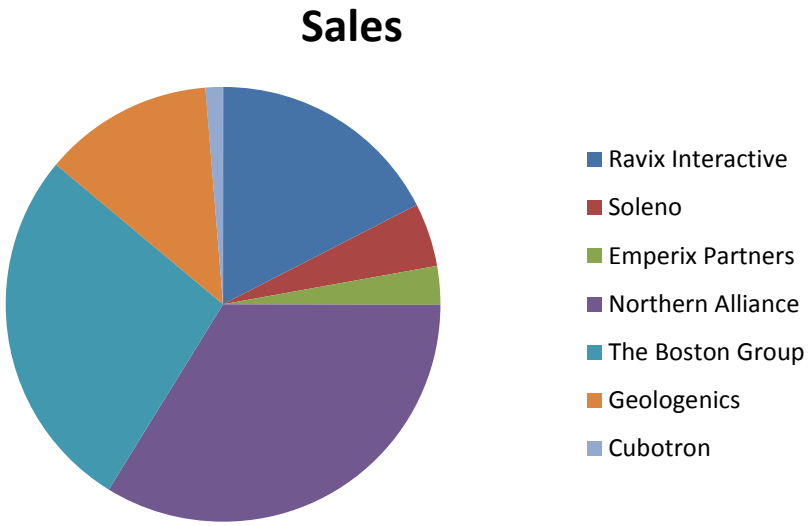
Sample Data

- Here's some sales data that we would like to visualize:

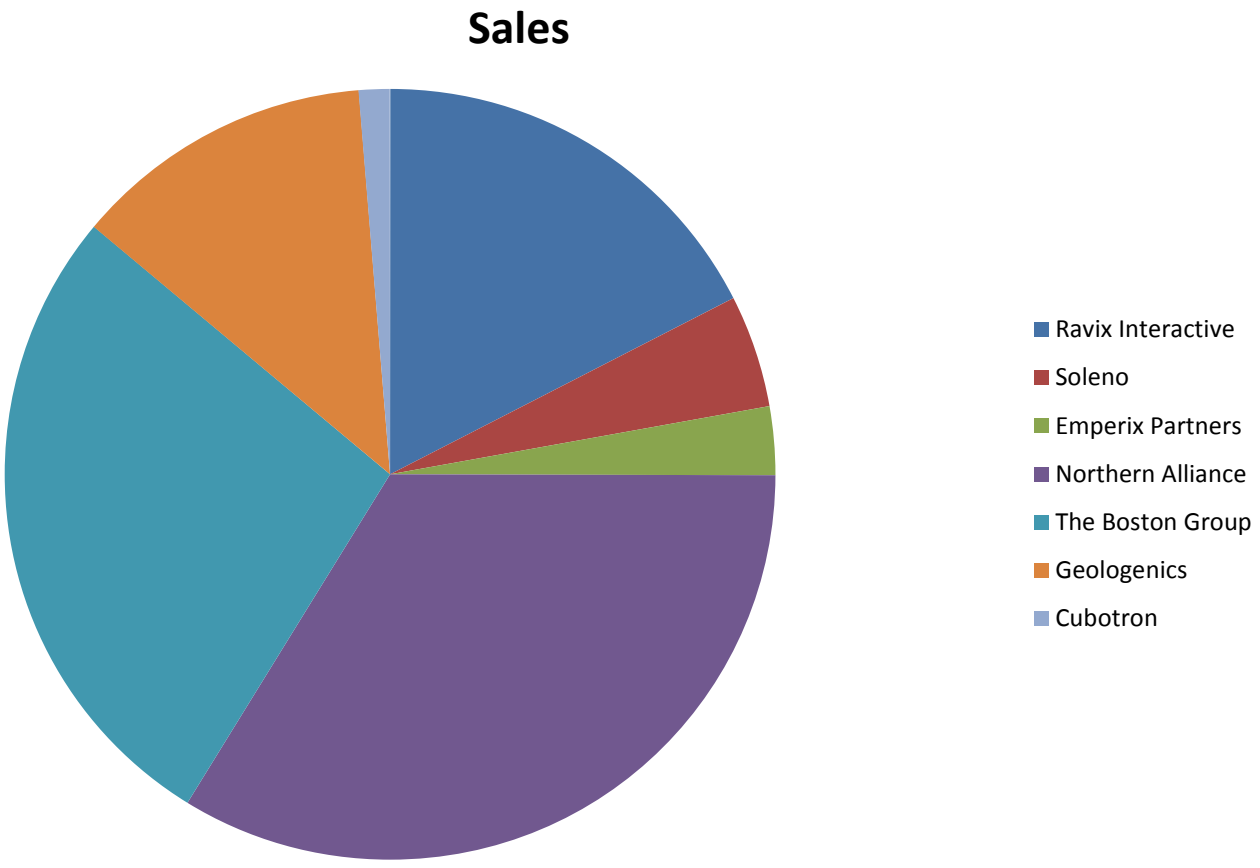
	A	B
1	Customer	Sales
2	Ravix Interactive	$ 120,900
3	Soleno	$ 32,800
4	Emperix Partners	$ 19,870
5	Northern Alliance	$ 234,000
6	The Boston Group	$ 189,000
7	Geologenics	$ 87,500
8	Cubotron	$ 9,000

Pie Charts

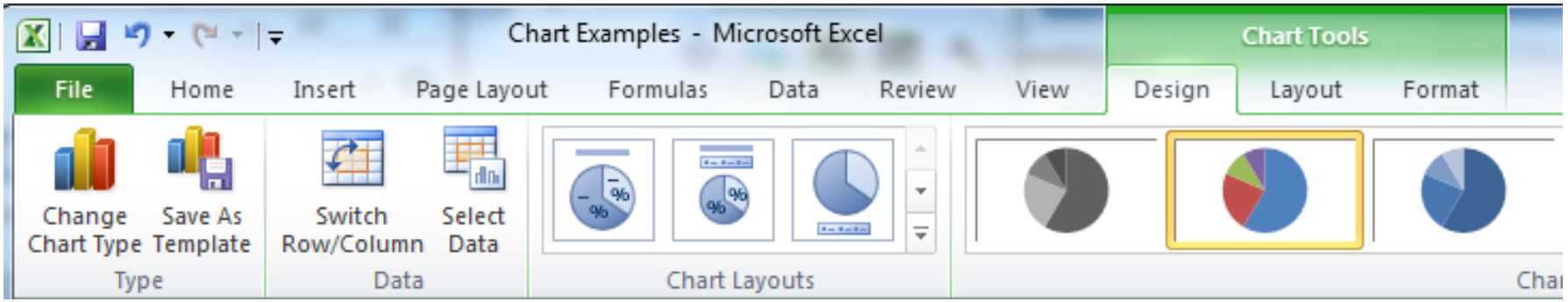
- A pie chart is useful when you are trying to show proportions.
- How much of the sales revenue comes from each client?
- Who are our largest clients?

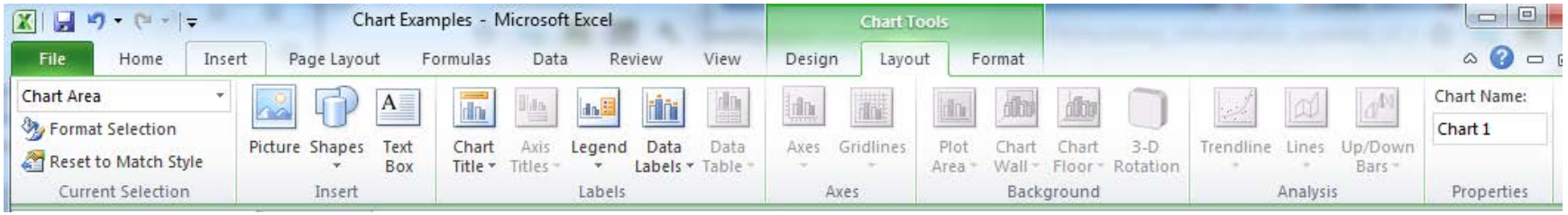


The Chart Layout

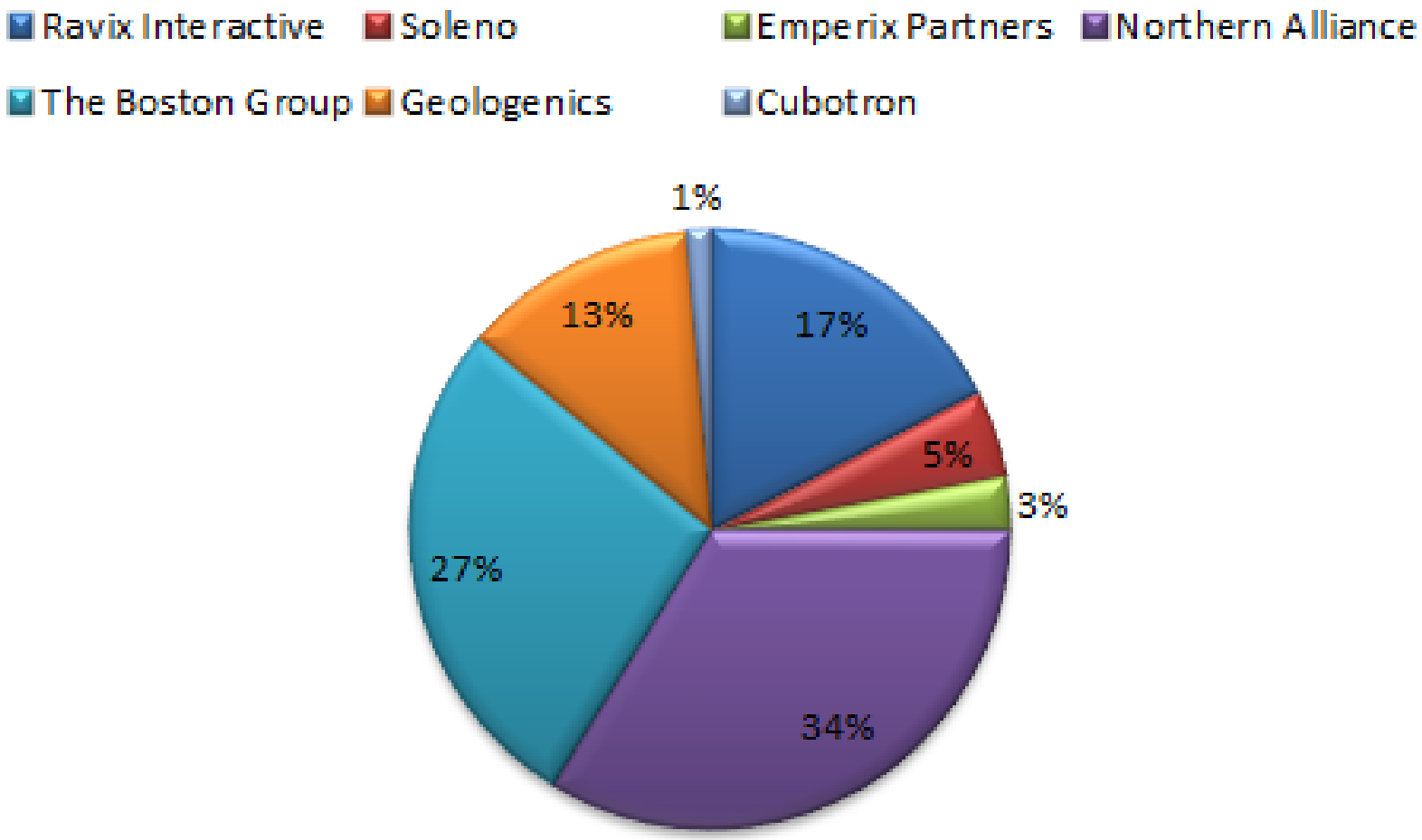


Customizing a Chart





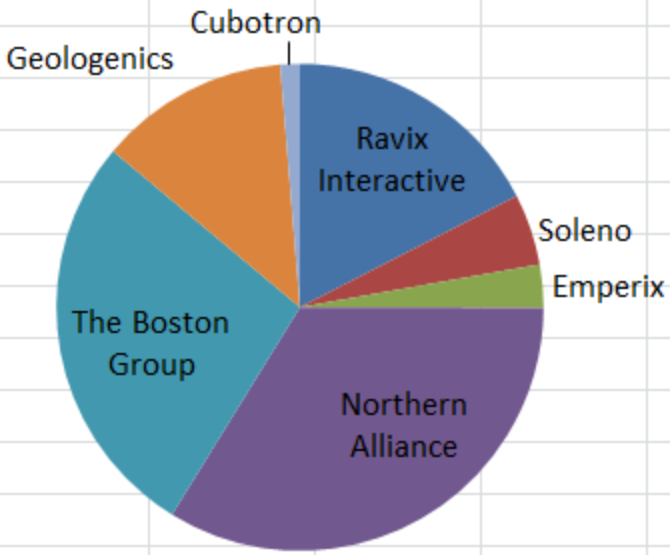
Sales



Transparency to Create a Minimal Display

- Useful for creating a worksheet display that minimizes chart details and simply shows a small graphic to support a set of numbers

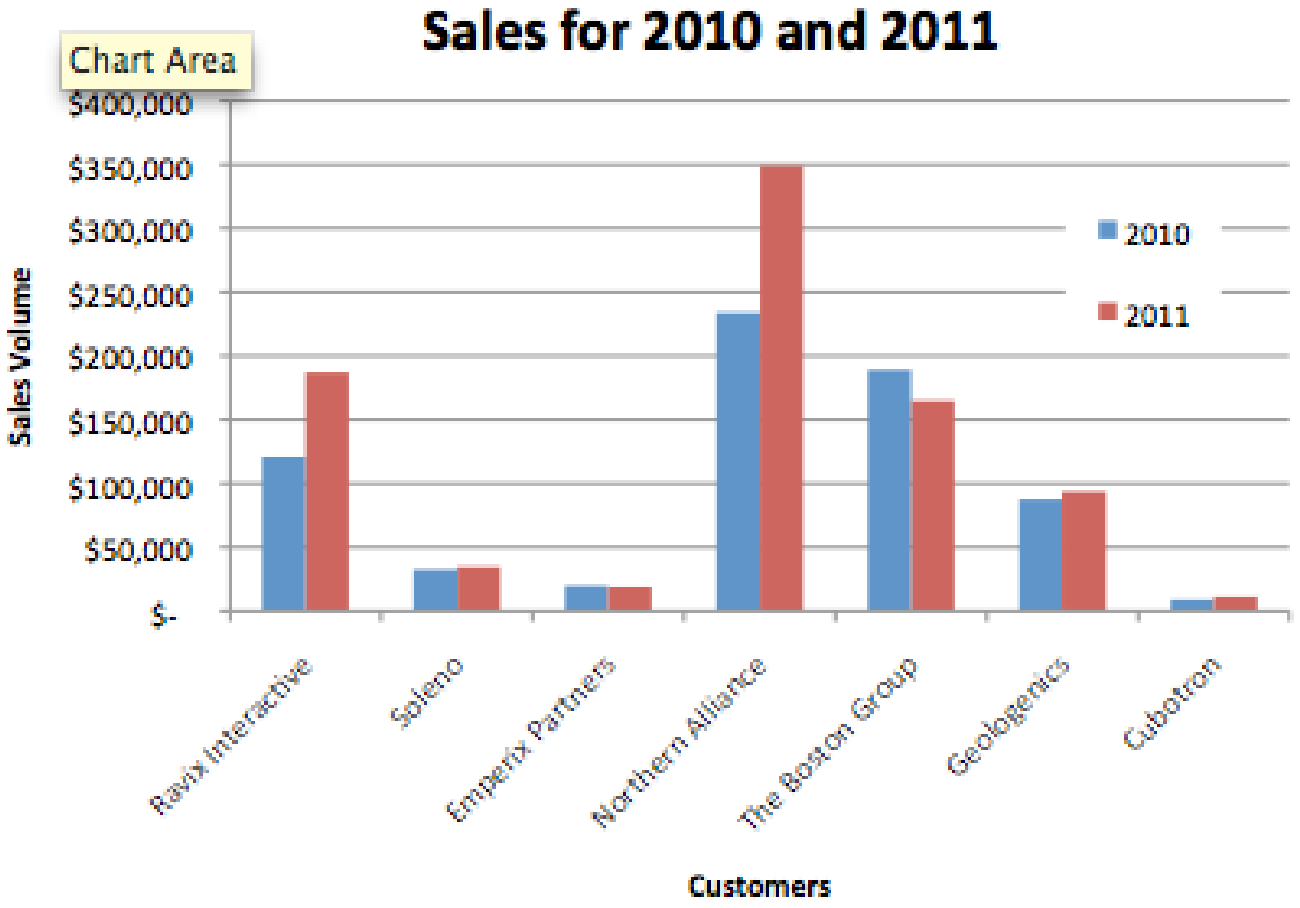
Customer	Sales
Ravix Interactive	$ 120,900
Soleno	$ 32,800
Emperix	$ 19,870
Northern Alliance	$ 234,000
The Boston Group	$ 189,000
Geologenics	$ 87,500
Cubotron	$ 9,000



Column Chart

- Also known as a bar chart, with rectangular bars of lengths usually proportional to the magnitudes or frequencies of what they represent.
- The bars are vertically oriented in a column chart
- Useful for showing data changes over a period of time, or illustrating comparisons
- Categories organized on horizontal axis
- Values on vertical axis

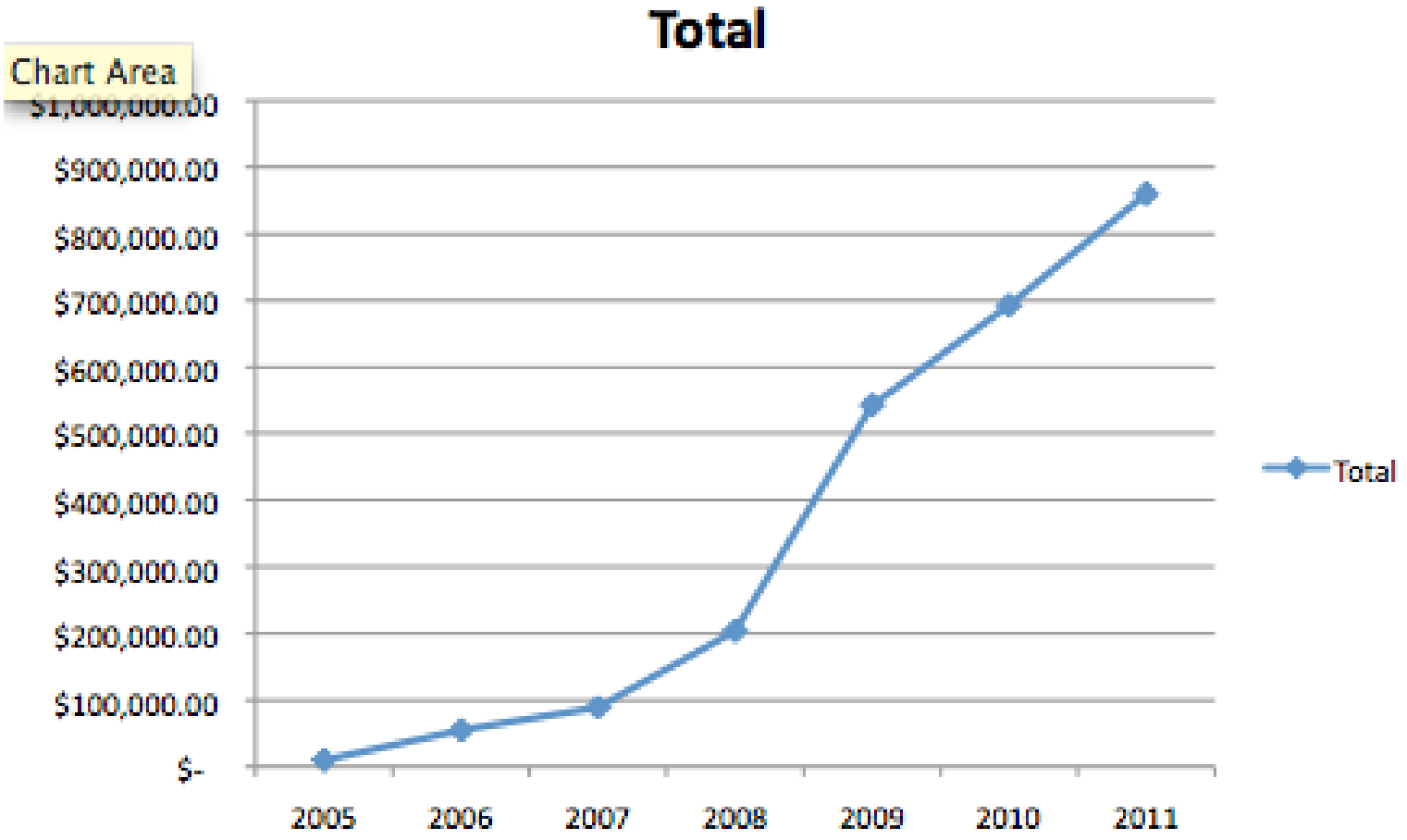
Column chart



Line Graph

- Often used to plot changes in data over time such as weekly temperature changes or stock market prices
- If plotting changes over time:
  - Time is plotted along the horizontal or x-axis
  - Data is plotted as individual points along the vertical axis

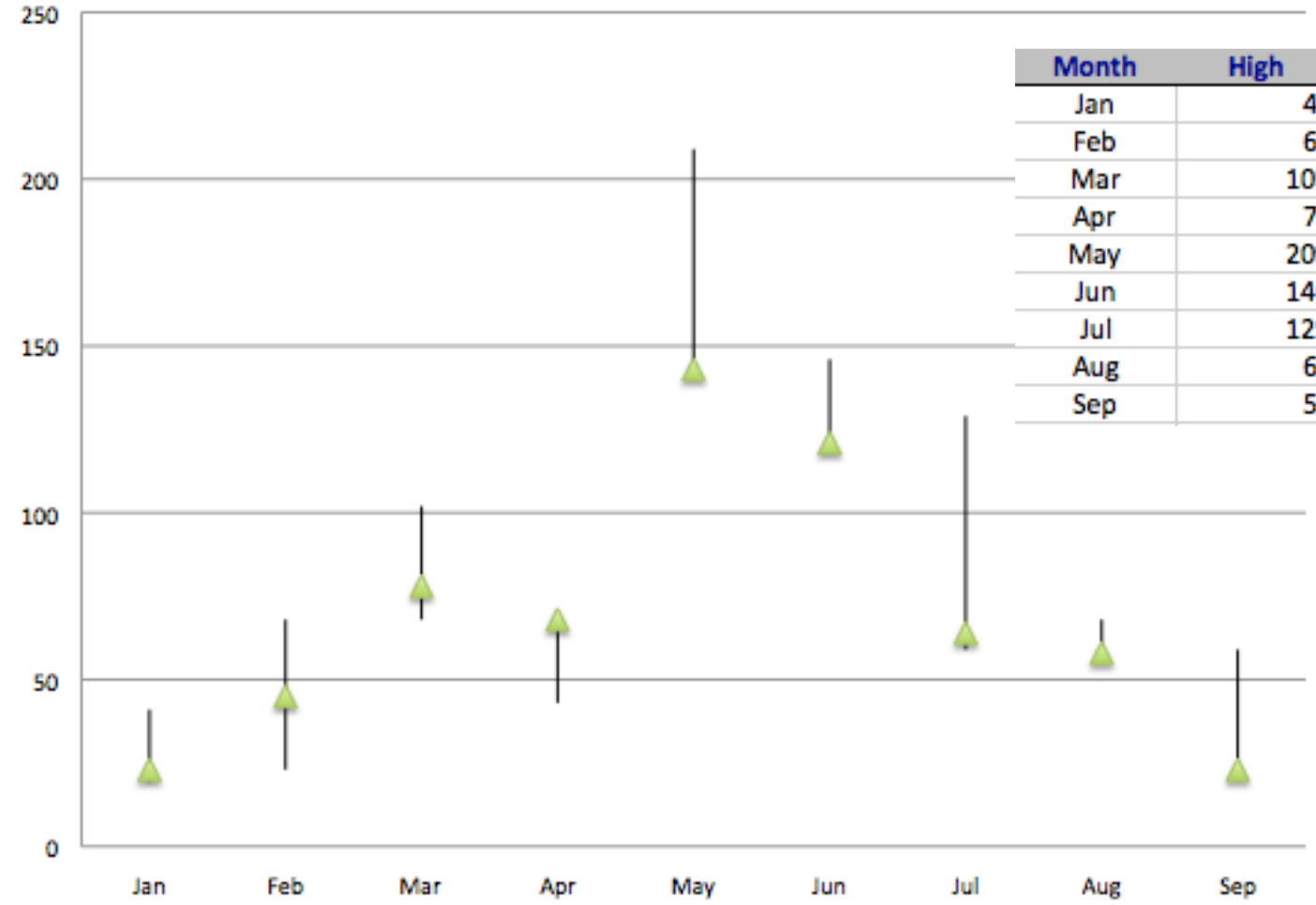
Line Graph



High Low Close Graph

- Used to illustrate the fluctuation of stock prices or for scientific data
- The data should be arranged with stock names as row headings, and High, Low and Close entered as column headings
- In “Stock” Charts in Excel

High Low Close



Month	High	Low	Close
Jan	41	19	23
Feb	68	23	45
Mar	102	68	78
Apr	71	43	68
May	209	142	143
Jun	146	119	121
Jul	129	59	64
Aug	68	55	58
Sep	59	21	23

X/Y Scatter Plot

- Useful for determining how things relate to one another e.g. profits vs. expenditures, height vs. weight, etc.
- Each data point has more than one attribute
  - Person (height, weight)
  - Quarter (profit, expenditure)
- Each attribute on single axis

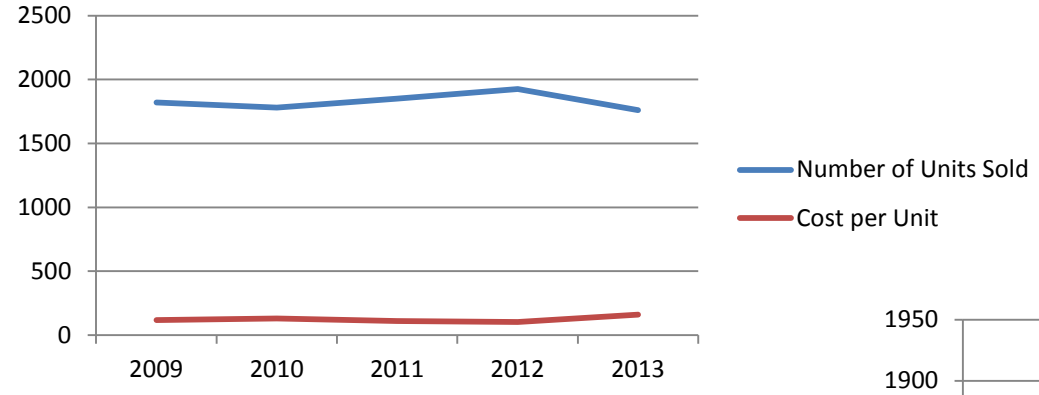
X/Y Scatter Plot

Assigning a Series to a Secondary Axis

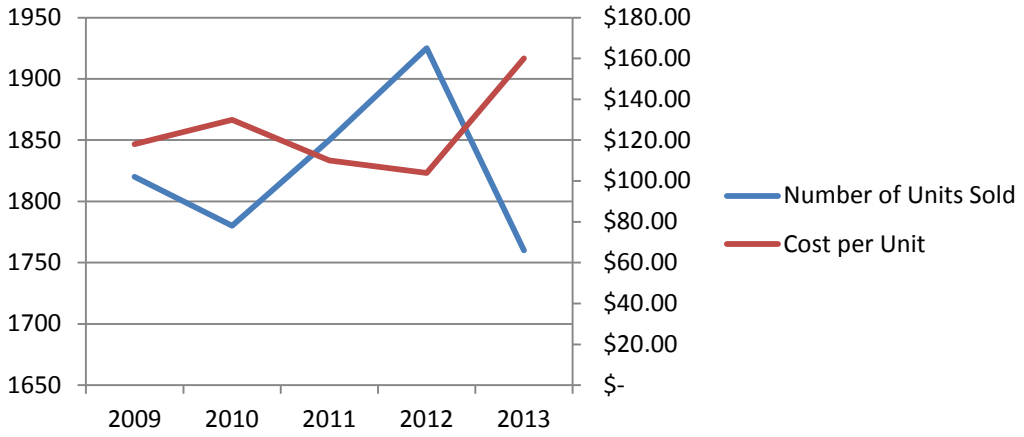
- A secondary value axis can make it easier to compare data series that have deviating ranges.
- Example: a series showing number of units sold per year has a range that is much higher than cost per unit per year that it's hard to see how they relate to each other. Putting one of the series on a secondary axis makes it possible to compare

Assigning a Series to a Secondary Axis

- The line graph on the left shows two data series with widely differing ranges, so it's hard to compare them.
- The graph on the right plots one series on a secondary axis making it much easier to compare.
- To move a series to a secondary axis, right-click on the series, click **Format Data Series**, select **Series Options** then select **Secondary Axis**.

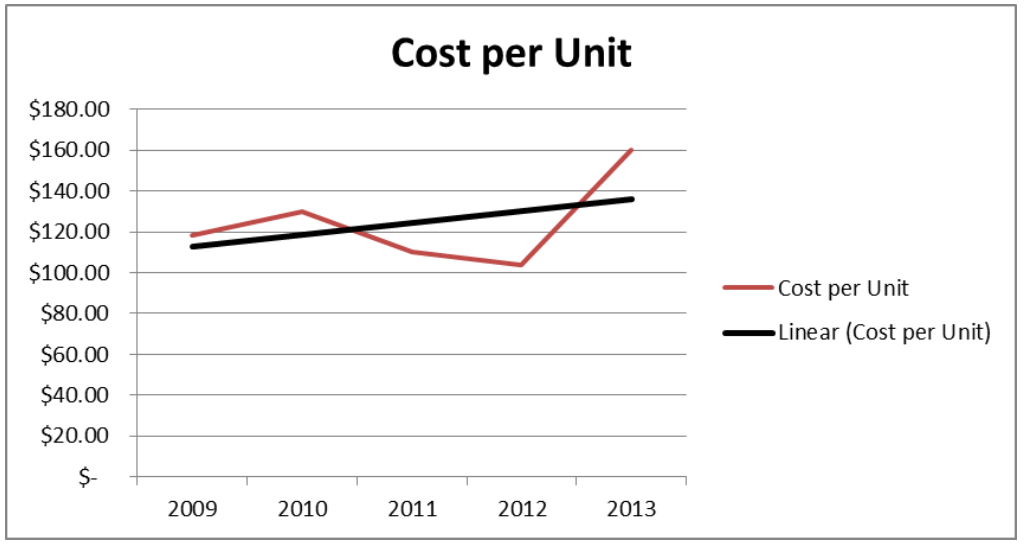


Column1	Number of Units Sold	Cost per Unit
2009	1820	$ 118.00
2010	1780	$ 130.00
2011	1850	$ 110.00
2012	1925	$ 104.00
2013	1760	$ 160.00



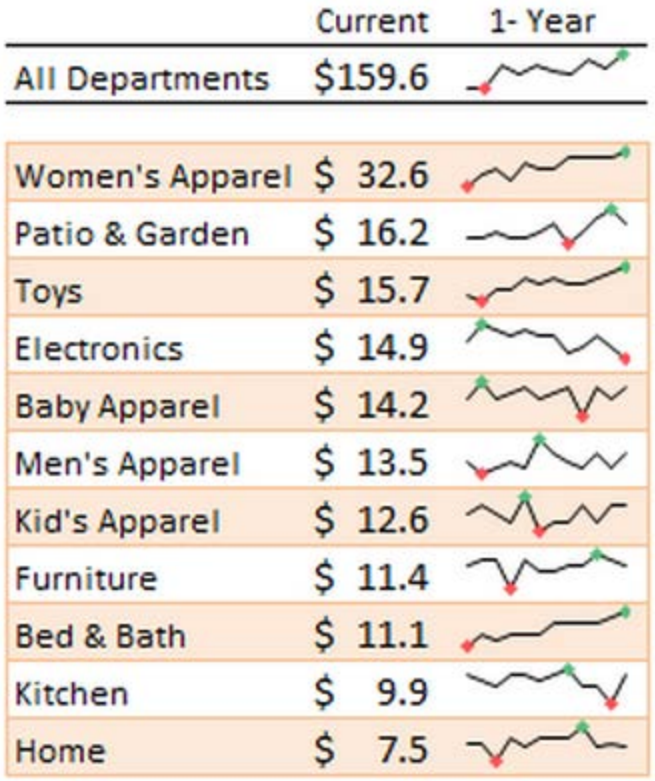
Trendlines, Error Bars, etc.

- Excel also provides statistical analysis tools via the Layout tab / Analysis section.
  - Trendlines show the "best fit" for the data.
  - Error bars show "confidence intervals" around data points.



Sparklines

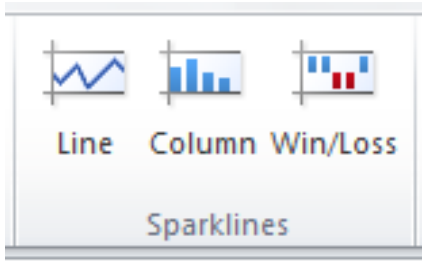
- New to Excel 2010, we can also create charts or graphs that live within one cell
- Their inventor, Edward Tufte, describes them as “intense, simple, word-sized graphics”
- Meant to be embedded into what they are describing
- Presents the general shape of variation in some measurement, in a simple and highly condensed way



To Create Sparklines:

- Select the cell where you want the Sparkline to appear
- Click the Insert tab and

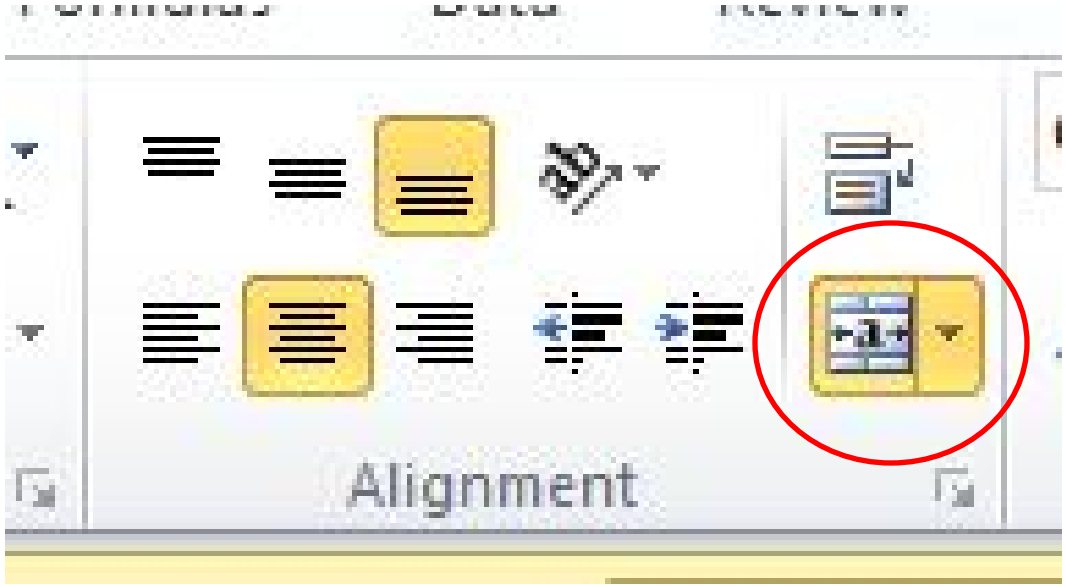
  look for the Sparklines group



- Choose the data range and the location for the Sparkline.

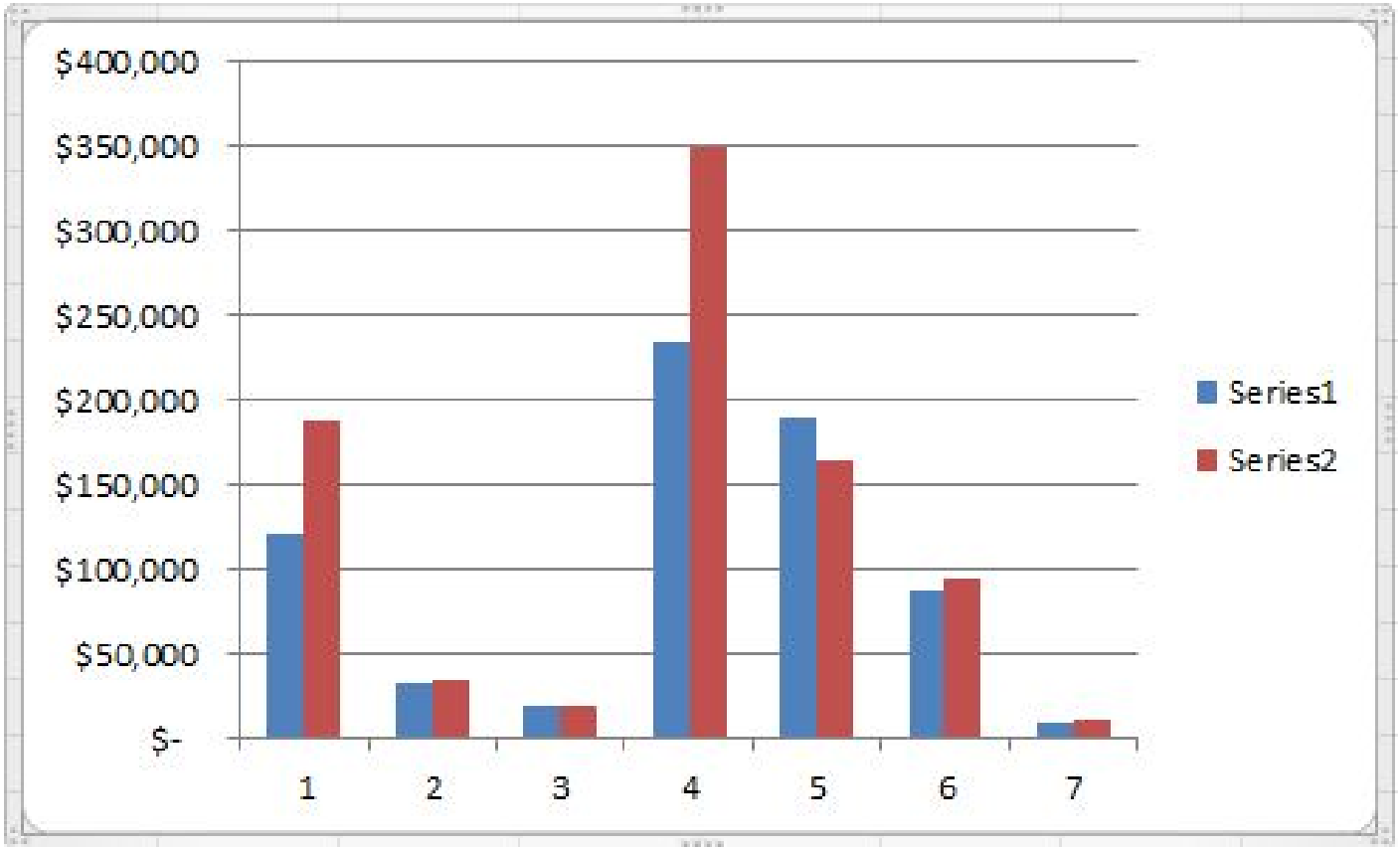
Merging Cells

- To make sparklines bigger, you can merge multiple cells into a single cell.
  - In the home tab:



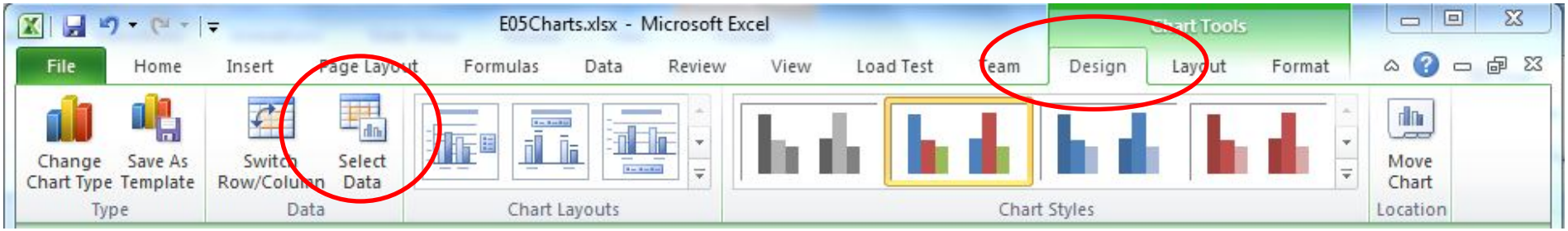
Common Issues: data labels

- Data labeled “Series1”



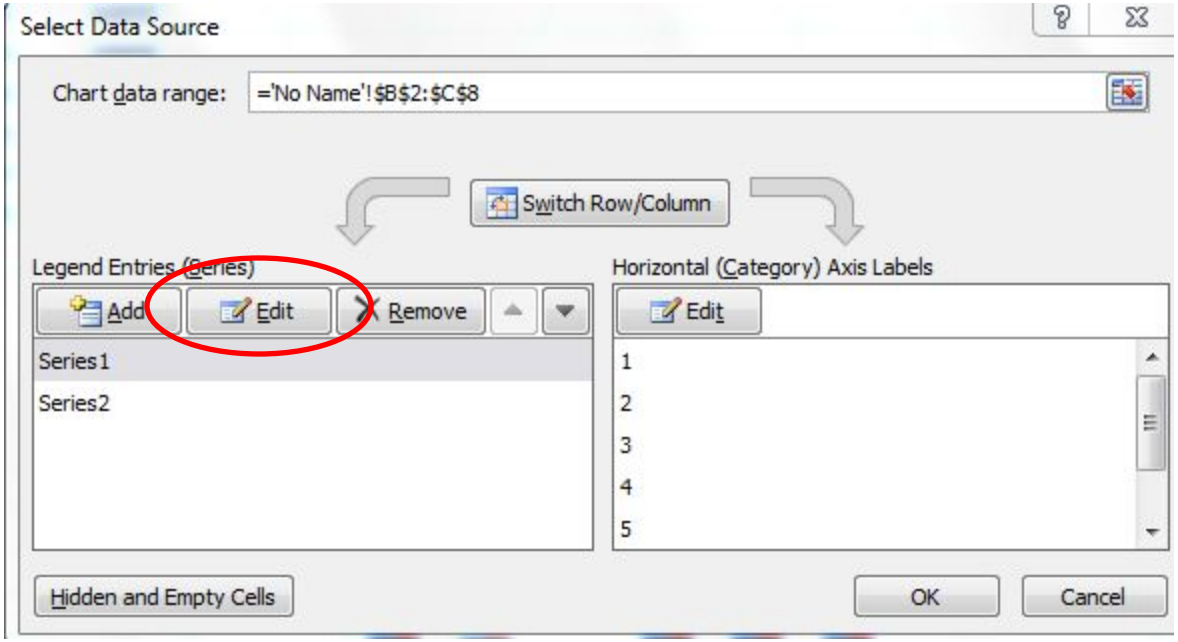
Common Issues: data labels

- Data labeled “Series1”
- To fix it: Select Data



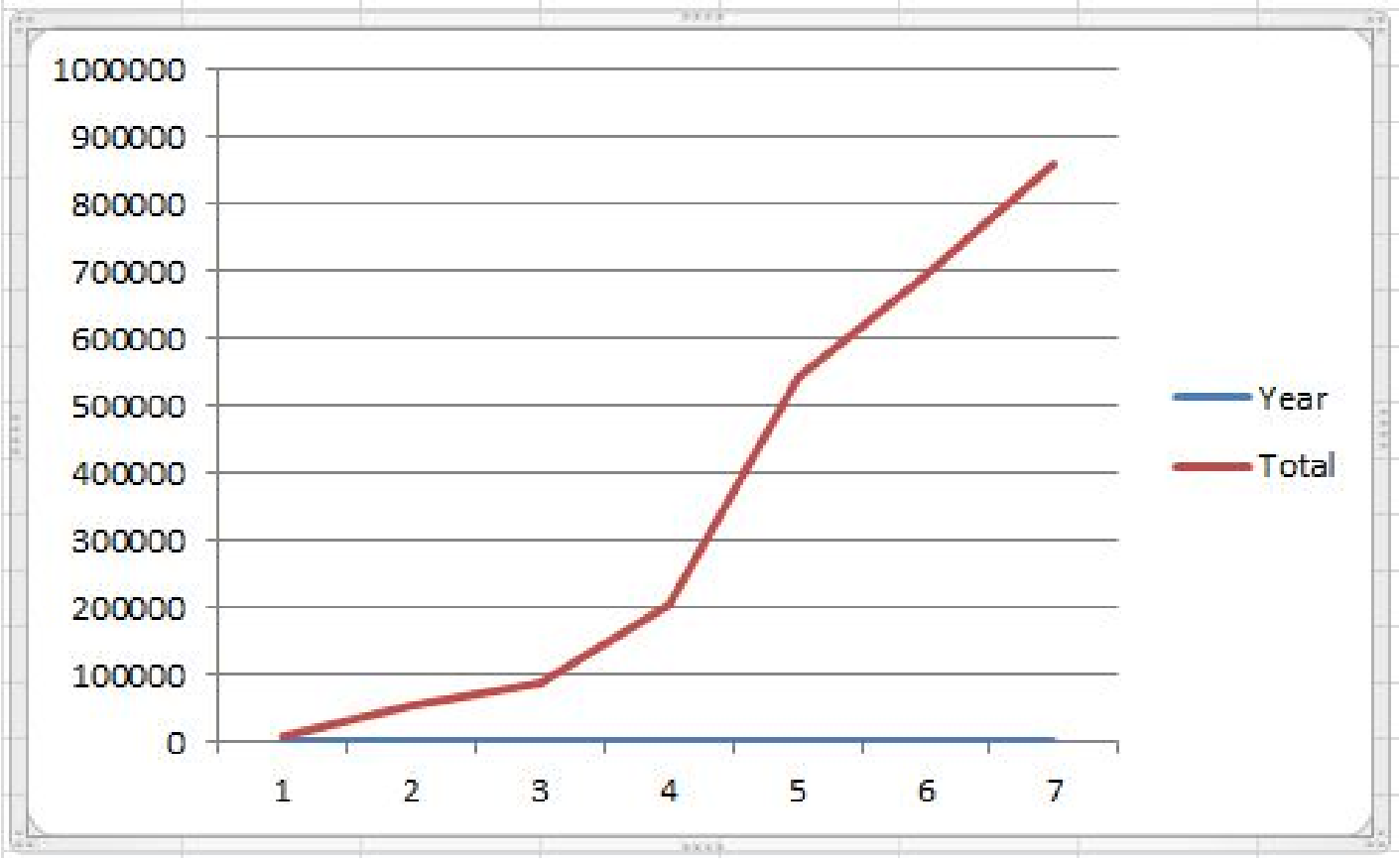
Common Issues: data labels

- Data labeled “Series1”
- To fix it: Select Data
  - Edit Series Name



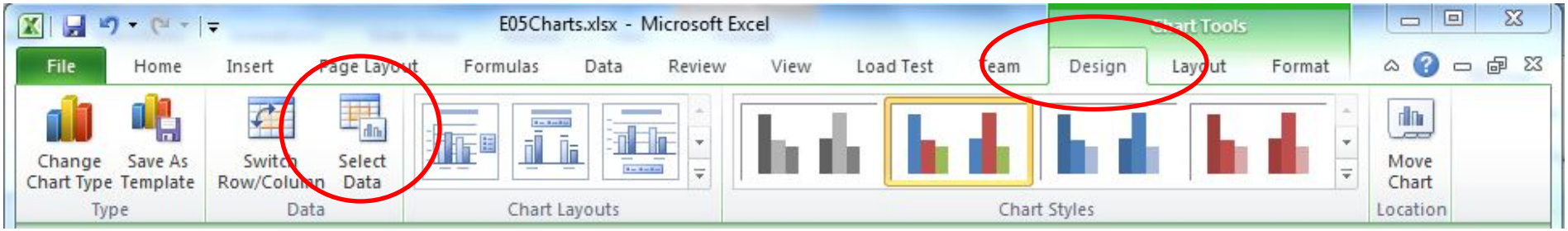
Common Issues: axis labels

- Axis labels plotted instead



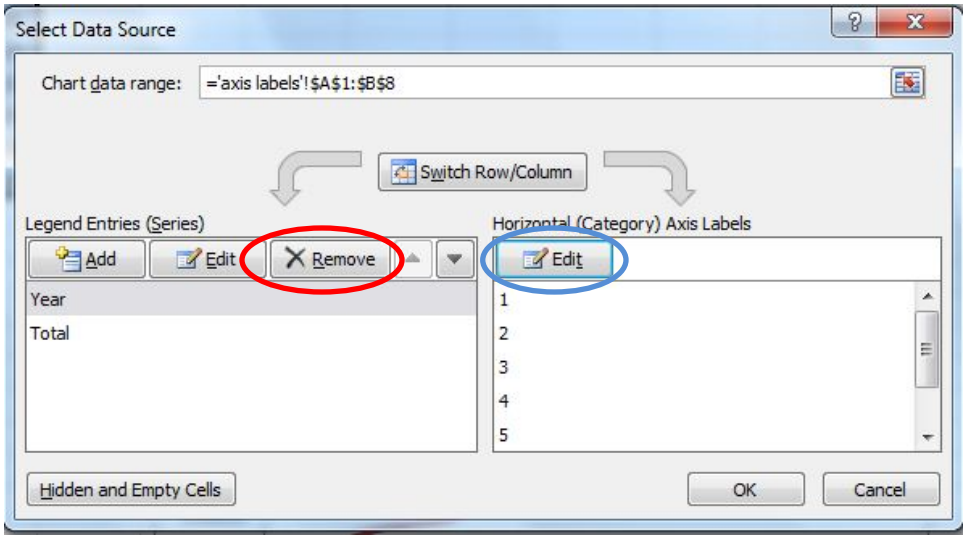
Common Issues: axis labels

- Axis labels plotted instead
- To fix it: Select Data



Common Issues: axis labels

- Axis labels plotted instead
- To fix it: Select Data
  1. Remove axis series
  2. Edit Axis Labels



Histograms

- Histograms are a specialized type of bar graph used to summarize groups of data.
- In some cases, you may collect a large number of data points for a single level of an independent variable.
  - That is, you take the same measurement over and over again. For example, when a lack of precision in measuring process does not give a good estimate of the true value with only a single measurement.

Binning

- How to summarize the results of these measurements?
- One way might be to simply calculate the average of all these measurements.
  - This would not, however, give you a good feel for how the data is *distributed.*
- A distribution graph, or **histogram**, allows you to see how many measurements fall within set ranges, or **bins**, of the dependent variable.
  - usually depicted as a bar chart, with one bar representing the count of how many measurements fall in a single bin.

Set up bins

- Find the minimum and maximum values of your data and the total range.
- Pick the number of bins that you want to use. Think small.
- Calculate the bin size: (Max – Min)/#bins and Rounddown to get a whole number

Min	2
Max	75
# Bins	5
Bin Size	14

Set up bins

- Set up the first bin in the first Bin Array cell
  - This will be the Min value of the data
- The next bin will be Min + Bin Size
- The next bin will be the previous Bin + Bin Size

Bin Array	Frequency
2	
16	3
30	7
44	7
58	2
72	2
86	2
	0

Set up bins

Base Case

- Set up the first bin in the first Bin Array cell
  - This will be the Min value of the data
- The next bin will be Min + Bin Size

Inductive Case

- The next bin will be the previous Bin + Bin Size

Bin Array	Frequency
2	
16	3
30	7
44	7
58	2
72	2
86	2
	0

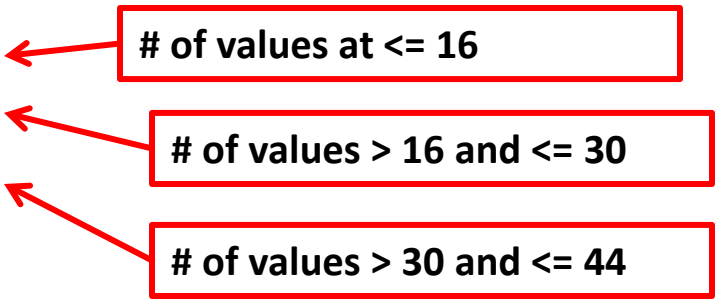
Compute Frequencies

- Use the FREQUENCY array function to fill in the data column. **(Order of the steps is important)**
- **First:** select the range for the Frequency plus one extra cell (extra cell for values that are greater than the highest interval in the data_array.)
- **Second:** type in the frequency function, =FREQUENCY(data_array, bin_array)
- **Third:** press CTRL-SHIFT-ENTER for Windows, or CMD-ENTER on Mac

Bin Array	Frequency
2	
16	3
30	7
44	7
58	2
72	2
86	2
	0

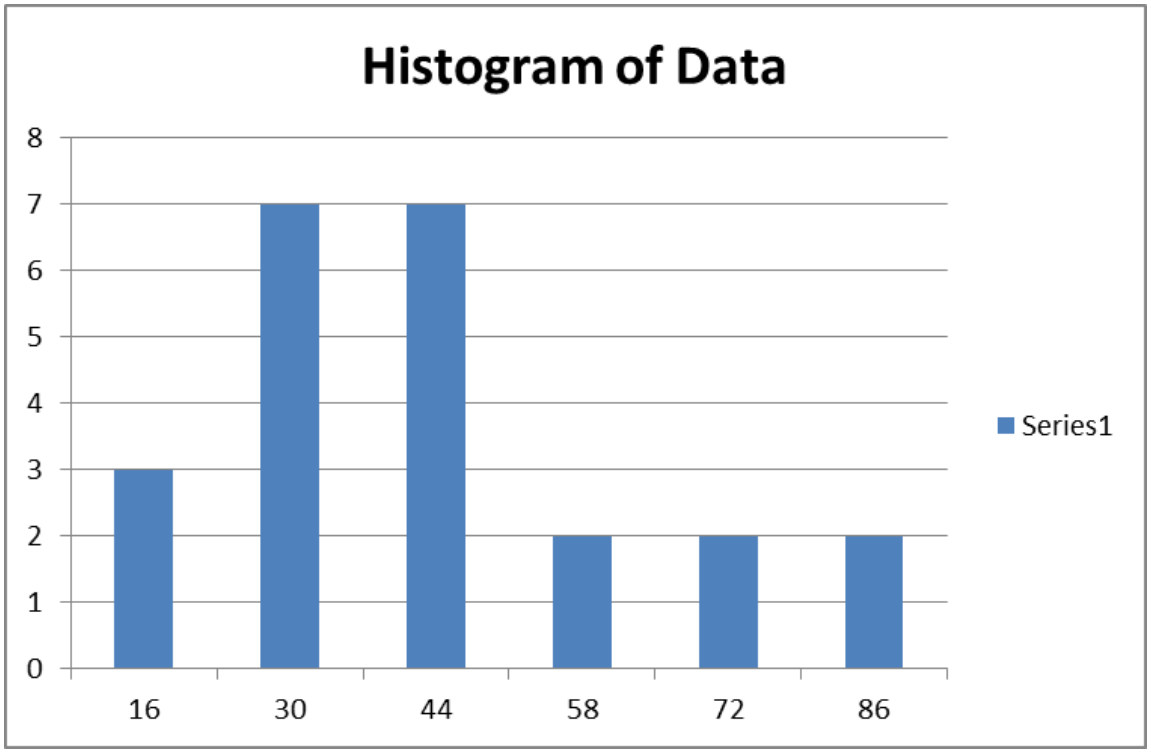
Compute Frequencies

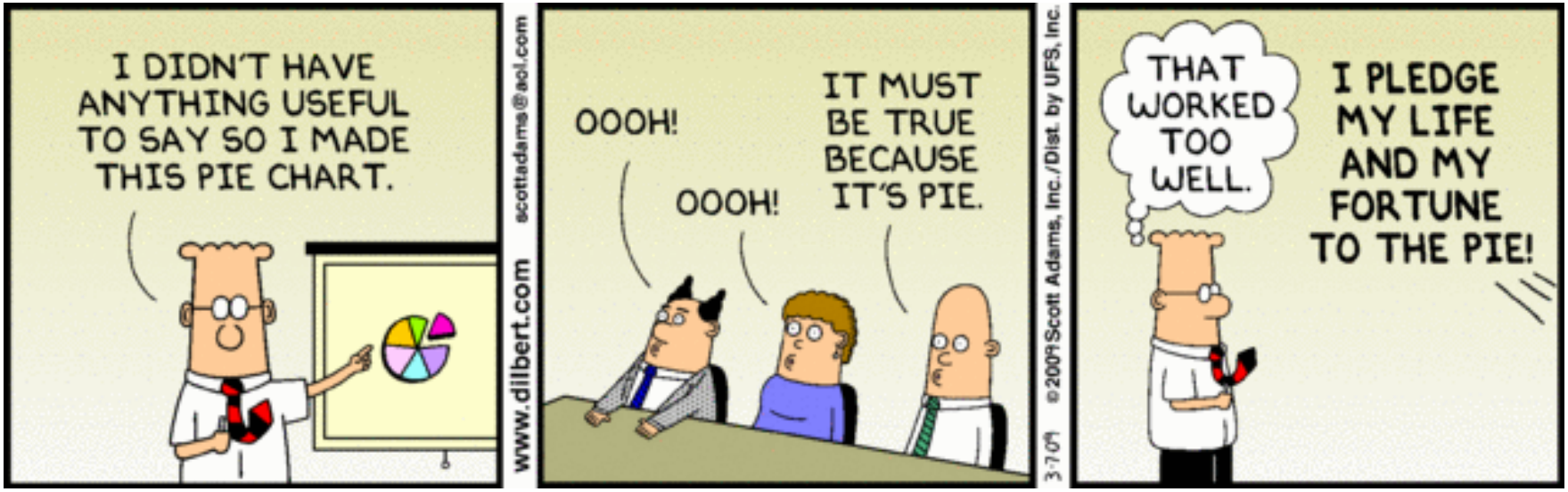
Bin Array	Frequency
2	
16	3
30	7
44	7
58	2
72	2
86	2
	0



Plot Histogram - Frequency vs. Bin Data

- Highlight the bin array and frequency numbers.
  Click on the icon for **Column Chart**.
  Series: X values are bin values, Y is the frequency.
  Add titles.





Any Question?