Microsoft Excel: Data Analysis & Graphing
Objectives

• Use relative, absolute, and mixed cell referencing
• Identify the types of graphs and their proper formatting
• Create and format charts in Excel
• Create and interpret trend lines
Review - Cell Reference: Relative Addressing

- Relative Cell Address – refer to the cell in the same relative position
- Ex: B12 – relative reference to value in cell B12, use fill handle to drag to cell C17
Review - Cell Reference: Absolute Addressing

- Absolute Cell Address – always refer to the same cell
- Ex: $B$12 – absolute reference to value in cell B12, use fill handle to drag to cell C17
Graphing

• A graph is often the best way to present data so that it is easily understood.

• There are several types of graphs available to present data. These graphs serve different purposes.
Graph Types: Categorical Data

Column Chart

Pie Chart

Bar Range Chart
Another Graph Type: Scatter Plot

Student Lab Report Scores

![Scatter Plot Diagram]

- Time spent (hours)
- Score

Another Scatter Plot

- Acceleration Noise (ft/sec/sec)
- Average Speed (mph)

Excel Data Analysis
Categorical Charts vs. Scatter Plots

- Categorical charts have **labels** on the x-axis
- Scatter plots have **values** on the x-axis
Example: Empirical Data

Motor Speed During Test

What happened?

Line Chart (Categorical)

Motor Speed During Test

Missing Data!

Scatter Plot
Interpolating & Predicting Data

• Excel graphing allows you to graph data and then create an equation that best fits that data.

• You can choose the type of equation for Excel to use, such as linear, exponential, logarithmic, etc.

• With a best-fit equation you may be able to interpolate or predict data not included in the worksheet. This must be done with caution.

• Note: You must use X-Y Scatter Plots
Example – Predicting Values

<table>
<thead>
<tr>
<th>Year</th>
<th>Births</th>
<th>Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>3,167,788</td>
<td>1,909,440</td>
</tr>
<tr>
<td>2003</td>
<td>3,612,258</td>
<td>1,989,841</td>
</tr>
<tr>
<td>2004</td>
<td>3,680,537</td>
<td>1,974,797</td>
</tr>
<tr>
<td>2005</td>
<td>3,669,141</td>
<td>2,039,369</td>
</tr>
<tr>
<td>2006</td>
<td>3,756,547</td>
<td>2,105,361</td>
</tr>
<tr>
<td>2007</td>
<td>3,909,510</td>
<td>2,167,999</td>
</tr>
<tr>
<td>2008</td>
<td>4,158,212</td>
<td>2,148,463</td>
</tr>
<tr>
<td>2009</td>
<td>4,110,907</td>
<td>2,169,518</td>
</tr>
<tr>
<td>2010</td>
<td>4,065,014</td>
<td>2,175,613</td>
</tr>
<tr>
<td>2011</td>
<td>4,000,240</td>
<td>2,268,553</td>
</tr>
<tr>
<td>2012</td>
<td>3,952,767</td>
<td>2,278,994</td>
</tr>
</tbody>
</table>

Question: How many births and deaths would we expect in 2016?
On your own, plot the data
What is missing?

US Population Statistics

- Births
- Deaths

Axis Labels (with units!)

Author, Date
That’s a lot of zeroes

US Population Statistics

- Births
- Deaths

Year:
- 2000
- 2002
- 2004
- 2006
- 2008
- 2010
- 2012
- 2014

Population (People):
- 0
- 500,000
- 1,000,000
- 1,500,000
- 2,000,000
- 2,500,000
- 3,000,000
- 3,500,000
- 4,000,000
- 4,500,000

BCM 5/11/12
Display Ordinate in Thousands
Properly Formatted Graph

US Population Statistics

Population (Thousands of People)

Year

Births

Deaths

BCM 5/11/12
Adding a Linear Trendline

1. Right click on any data point and select the ‘Add Trendline’ option.
2. Select ‘Linear’ as the regression type (note that there are other options available).
3. Select this option to show the equation of the line on the chart.
Trendline for Births

US Population Statistics

Note: You may need to re-format the trendline equation to improve resolution. You may also need to move the equation.
Formatting Trendline Label

1. Select trendline label (left-click)
2. Right-click on edge of label box
3. Select ‘Format Trendline Label’
Formatting Trendline Label

Select Number with 0 decimal places
Results

- Using these equations, we predict for 2016:
  - Births = 4,474,063
  - Deaths = 2,428,399
- Do you see any problem with these predictions?
- What can we do to improve our estimate?
Interpreting Charts

Were traffic fatalities stable?

Or were they cyclical?

TRAFFIC FATALITIES IN OHIO 1992-2002

TRAFFIC FATALITIES IN OHIO 1992-2002
Interpreting Charts

• Maybe we should plot the number of fatalities per 100 million miles driven.
• Which of the three is correct?
Formatting a Plot: One Data Set

- Title
- Axis Labels
  - Units!
- Even increments
- Author/Date

Motor Speed During Test

- Speed (% Maximum)
- Time (Hours)

BCM 6/20/12
Add Another Data Set

- Data markers different from each other
- Legend
Questions?

The way it made me feel was smart because I was asking good questions and giving good answers.