



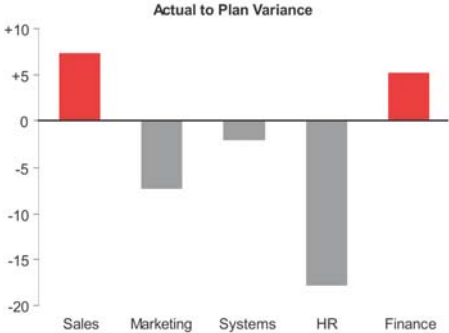
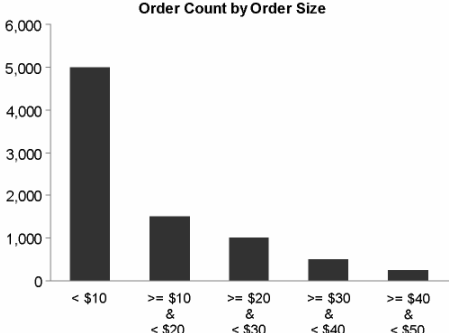
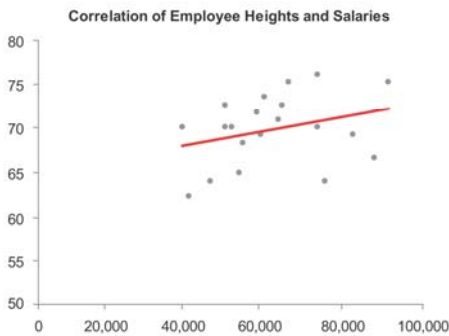
## **Selecting the Right Graph for Your Message**

Stephen Few

September 18, 2004

À  
À

Type/Description	Encoding Methods	Example																																							
<p><b>Nominal Comparison</b> A simple comparison of the categorical subdivisions of one or more measures in no particular order</p>	<ul style="list-style-type: none"> <li>• Bars only (horizontal or vertical)</li> </ul>	<p><b>Q1 2003 Calls by Region</b></p> <table border="1"> <caption>Q1 2003 Calls by Region</caption> <thead> <tr> <th>Region</th> <th>Calls</th> </tr> </thead> <tbody> <tr> <td>North</td> <td>3,000</td> </tr> <tr> <td>East</td> <td>4,500</td> </tr> <tr> <td>South</td> <td>2,200</td> </tr> <tr> <td>West</td> <td>4,800</td> </tr> </tbody> </table>	Region	Calls	North	3,000	East	4,500	South	2,200	West	4,800																													
Region	Calls																																								
North	3,000																																								
East	4,500																																								
South	2,200																																								
West	4,800																																								
<p><b>Time Series</b> Multiple instances of one or more measures taken at equidistant points in time</p>	<ul style="list-style-type: none"> <li>• Lines to emphasize overall pattern</li> <li>• Bars to emphasize individual values</li> <li>• Points connected by lines to slightly emphasize individual values while still highlighting the overall pattern</li> <li>• Always place time on the horizontal axis</li> </ul>	<p><b>2003 Sales</b></p> <table border="1"> <caption>2003 Sales (Estimated)</caption> <thead> <tr> <th>Month</th> <th>Grey Line Sales</th> <th>Orange Line Sales</th> </tr> </thead> <tbody> <tr><td>Jan</td><td>2,200</td><td>900</td></tr> <tr><td>Feb</td><td>2,500</td><td>1,000</td></tr> <tr><td>Mar</td><td>2,800</td><td>1,000</td></tr> <tr><td>Apr</td><td>2,500</td><td>900</td></tr> <tr><td>May</td><td>2,800</td><td>1,000</td></tr> <tr><td>Jun</td><td>3,000</td><td>1,000</td></tr> <tr><td>Jul</td><td>2,600</td><td>900</td></tr> <tr><td>Aug</td><td>2,800</td><td>500</td></tr> <tr><td>Sep</td><td>3,100</td><td>900</td></tr> <tr><td>Oct</td><td>2,900</td><td>900</td></tr> <tr><td>Nov</td><td>3,200</td><td>900</td></tr> <tr><td>Dec</td><td>3,600</td><td>1,000</td></tr> </tbody> </table>	Month	Grey Line Sales	Orange Line Sales	Jan	2,200	900	Feb	2,500	1,000	Mar	2,800	1,000	Apr	2,500	900	May	2,800	1,000	Jun	3,000	1,000	Jul	2,600	900	Aug	2,800	500	Sep	3,100	900	Oct	2,900	900	Nov	3,200	900	Dec	3,600	1,000
Month	Grey Line Sales	Orange Line Sales																																							
Jan	2,200	900																																							
Feb	2,500	1,000																																							
Mar	2,800	1,000																																							
Apr	2,500	900																																							
May	2,800	1,000																																							
Jun	3,000	1,000																																							
Jul	2,600	900																																							
Aug	2,800	500																																							
Sep	3,100	900																																							
Oct	2,900	900																																							
Nov	3,200	900																																							
Dec	3,600	1,000																																							
<p><b>Ranking</b> Categorical subdivisions of a measure ordered by size (either descending or ascending)</p>	<ul style="list-style-type: none"> <li>• Bars only (horizontal or vertical)</li> <li>• To highlight high values, sort in descending order</li> <li>• To highlight low values, sort in ascending order</li> </ul>	<p><b>Headcount</b></p> <table border="1"> <caption>Headcount by Department</caption> <thead> <tr> <th>Department</th> <th>Headcount</th> </tr> </thead> <tbody> <tr><td>Manufacturing</td><td>230</td></tr> <tr><td>Sales</td><td>160</td></tr> <tr><td>Engineering</td><td>60</td></tr> <tr><td>Operations</td><td>50</td></tr> <tr><td>Finance</td><td>40</td></tr> <tr><td>Info Systems</td><td>40</td></tr> <tr><td>Legal</td><td>20</td></tr> <tr><td>Marketing</td><td>10</td></tr> </tbody> </table>	Department	Headcount	Manufacturing	230	Sales	160	Engineering	60	Operations	50	Finance	40	Info Systems	40	Legal	20	Marketing	10																					
Department	Headcount																																								
Manufacturing	230																																								
Sales	160																																								
Engineering	60																																								
Operations	50																																								
Finance	40																																								
Info Systems	40																																								
Legal	20																																								
Marketing	10																																								
<p><b>Part-to-Whole</b> Measures of individual categorical subdivisions as ratios to the whole</p>	<ul style="list-style-type: none"> <li>• Bars only (horizontal or vertical)</li> <li>• Use stacked bars only when you must display measures of the whole as well as the parts</li> </ul>	<p><b>Regional % of Total Expenses</b></p> <table border="1"> <caption>Regional % of Total Expenses</caption> <thead> <tr> <th>Region</th> <th>% of Total Expenses</th> </tr> </thead> <tbody> <tr> <td>West</td> <td>33%</td> </tr> <tr> <td>East</td> <td>30%</td> </tr> <tr> <td>North</td> <td>21%</td> </tr> <tr> <td>South</td> <td>15%</td> </tr> </tbody> </table>	Region	% of Total Expenses	West	33%	East	30%	North	21%	South	15%																													
Region	% of Total Expenses																																								
West	33%																																								
East	30%																																								
North	21%																																								
South	15%																																								

<p><b>Deviation</b> Categorical subdivisions of a measure compared to a reference measure, expressed as the differences between them</p>	<ul style="list-style-type: none"> <li>• Lines to emphasize the overall pattern only when displaying deviation and time-series relationships together</li> <li>• Points connected by lines to slightly emphasize individual data points while also highlighting the overall pattern when displaying deviation and time-series relationships together</li> <li>• Bars to emphasize individual values, but limit to vertical bars when a time-series relationship is included</li> <li>• Always include a reference line to compare the measures of deviation against</li> </ul>	 <p><b>Actual to Plan Variance</b></p> <p>A bar chart showing variance for five departments: Sales, Marketing, Systems, HR, and Finance. The y-axis ranges from -20 to +10. Sales and Finance are positive (red bars), while Marketing, Systems, and HR are negative (grey bars). A horizontal reference line is at 0.</p> <table border="1"> <thead> <tr> <th>Department</th> <th>Variance</th> </tr> </thead> <tbody> <tr> <td>Sales</td> <td>+8</td> </tr> <tr> <td>Marketing</td> <td>-7</td> </tr> <tr> <td>Systems</td> <td>-2</td> </tr> <tr> <td>HR</td> <td>-18</td> </tr> <tr> <td>Finance</td> <td>+5</td> </tr> </tbody> </table>	Department	Variance	Sales	+8	Marketing	-7	Systems	-2	HR	-18	Finance	+5
Department	Variance													
Sales	+8													
Marketing	-7													
Systems	-2													
HR	-18													
Finance	+5													
<p><b>Frequency Distribution</b> Counts of something per categorical subdivisions (intervals) of a quantitative range</p>	<ul style="list-style-type: none"> <li>• Vertical bars to emphasize individual values (called a <i>histogram</i>)</li> <li>• Lines to emphasize the overall pattern (called a <i>frequency polygon</i>)</li> </ul>	 <p><b>Order Count by Order Size</b></p> <p>A histogram showing the number of orders in different size ranges. The y-axis is 'Order Count' (0 to 6,000). The x-axis shows order size ranges. A frequency polygon is overlaid on the bars.</p> <table border="1"> <thead> <tr> <th>Order Size Range</th> <th>Order Count</th> </tr> </thead> <tbody> <tr> <td>&lt; \$10</td> <td>5,000</td> </tr> <tr> <td>&gt;= \$10 &amp; &lt; \$20</td> <td>1,500</td> </tr> <tr> <td>&gt;= \$20 &amp; &lt; \$30</td> <td>1,000</td> </tr> <tr> <td>&gt;= \$30 &amp; &lt; \$40</td> <td>500</td> </tr> <tr> <td>&gt;= \$40 &amp; &lt; \$50</td> <td>200</td> </tr> </tbody> </table>	Order Size Range	Order Count	< \$10	5,000	>= \$10 & < \$20	1,500	>= \$20 & < \$30	1,000	>= \$30 & < \$40	500	>= \$40 & < \$50	200
Order Size Range	Order Count													
< \$10	5,000													
>= \$10 & < \$20	1,500													
>= \$20 & < \$30	1,000													
>= \$30 & < \$40	500													
>= \$40 & < \$50	200													
<p><b>Correlation</b> Comparisons of two paired sets of measures to determine if as one set goes up the other set goes either up or down in a corresponding manner, and if so, how strongly</p>	<ul style="list-style-type: none"> <li>• Points and a trend line in the form of a scatter plot</li> <li>• Bars may be used, arranged as a <i>paired bar graph</i> or a <i>correlation bar graph</i>, if scatter plots are unfamiliar</li> <li>• (Note: For descriptions of these graphs, see my book <i>Show Me the Numbers</i>.)</li> </ul>	 <p><b>Correlation of Employee Heights and Salaries</b></p> <p>A scatter plot showing the relationship between employee height (y-axis, 50-80) and salary (x-axis, 0-100,000). A red trend line indicates a positive correlation.</p>												

By understanding these seven types of quantitative relationships and the graphical methods that present them most effectively, you've already won half the battle. Knowing the best means to present data is the first big step; knowing how to design the separate components of a graph to communicate your message clearly, powerfully, and without distraction is the other big step, which we'll examine in the next article of this series.

(This article was originally published in *Intelligent Enterprise*.)