Presenting Empirical Results

General Practices in Reporting Quantitative Data

•Data can be presented in text, table, or chart form. When presenting data in all three forms, care should be taken to include only information and/or images that help to clarify points being made.

•For reference purposes, tables are usually the sensible option. Extensive tables should usually appear as appendices at the end of a report.

•In general, tables are better than graphs for giving structured numeric information. Graphs are better for demonstrating trends, making comparisons or showing relationships.

•Text alone should not be used to convey more than three or four numbers. Sets of numerical results should usually be presented as tables or pictures, rather than included in the text.

•When whole numbers are given in text, numbers less than, or equal to, nine should be written as words, while numbers from 10 upwards should be written in digits.

Continue

•When decimal numbers are quoted, the number of significant digits should be consistent.

•Tables and graphs should be self-explanatory. The reader should be able to understand them without detailed reference to the text. The title should be informative and rows and columns of tables or axes of graphs should be clearly labeled.

•On the other hand, the text should always include mention of the key points in a table or figure. If a table does not warrant discussion, it should not be there.

•Statistical information beyond means and frequencies (e.g., standard deviations, p-values, t-values), is usually required in formal scientific papers, but may not be necessary for a more general readership. When presented, care should be taken to do so in a way that does not obscure the main message of the table or graph.

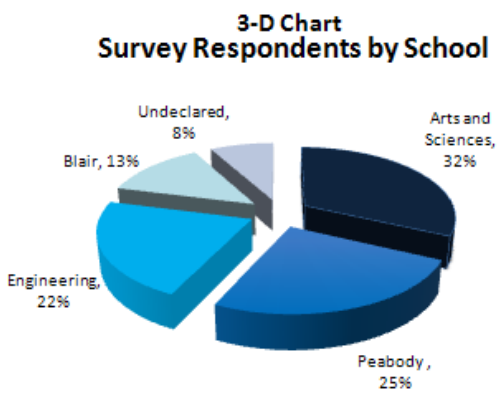
Pie Charts

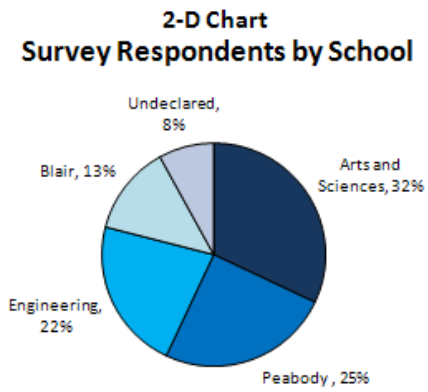
•Pie charts have limited utility. They can only be used to show parts of a whole (if all parts total 100%).

•Pie charts emphasize general findings, but do not make small differences apparent.

•Pie charts should only be used to represent categorical data with a relatively small number of values and should not consist of more than five or six slices.

•When presenting a pie chart, it is better not to use 3-D features, or break out the pieces, as this often makes it more difficult to compare the relative size of the slices.



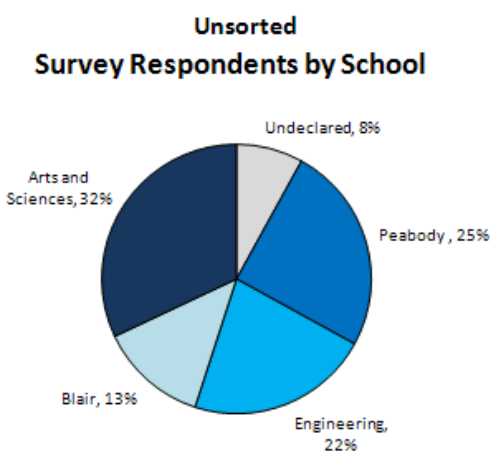


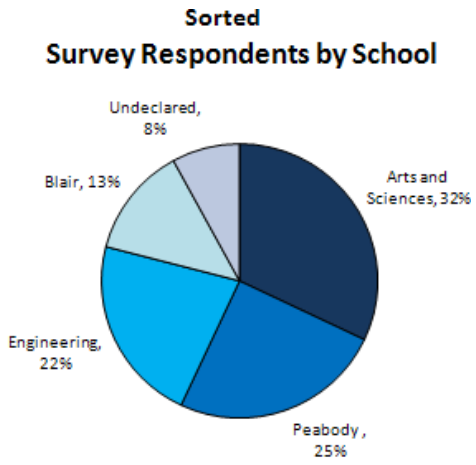
Continue

•It is always necessary to include category labels or a legend that describes which slice corresponds with which category. If labels are brief enough, it is better to place category labels directly next to the pie slices to which they correspond.

•It is good practice to include value labels (indicating the percentage of the pie represented by a given slice).

•It is also good to pre-sort data so that, clockwise or counter-clockwise, the relative size of pie slices is most apparent.



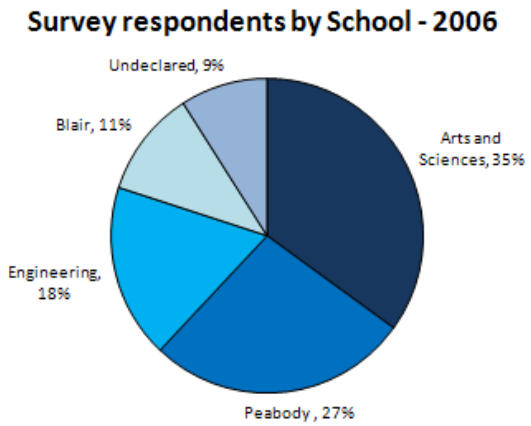


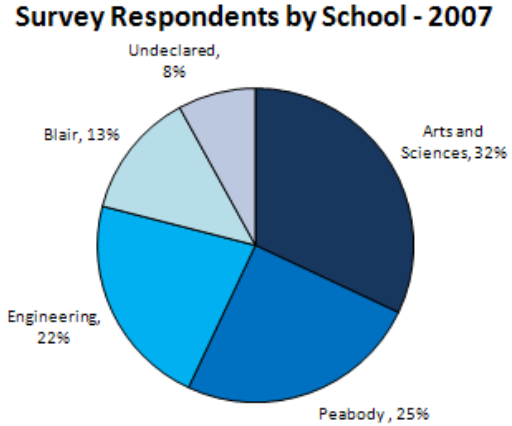
Continue

•The purpose of color in pie charts is to differentiate between pie slices to further facilitate comparison. When using color, it should be used thematically. The best use of color in a size-ordered pie chart is a progression of dark to light hues from the largest slice to smallest slice (see 2-D pie chart above).

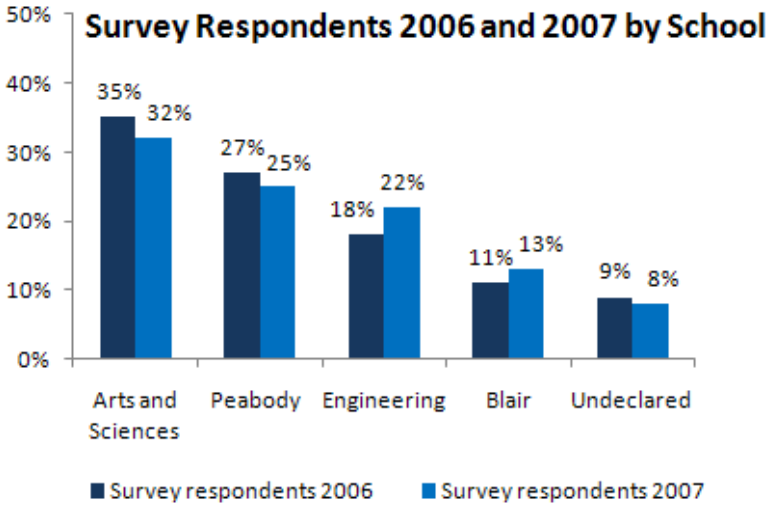
•A bar graph, rather than multiple pie graphs, is the better option if data need to be compared by more than one value. Pie graphs should not be used to represent more than one categorization of data.

Multiple Pie Graphs





A Single Bar Graph



Bar Graphs

•Bar graphs are used for direct comparison of data (e.g., student GPA's by class year).

•Bar graphs can also be used to show time series data when the number of time intervals is small.

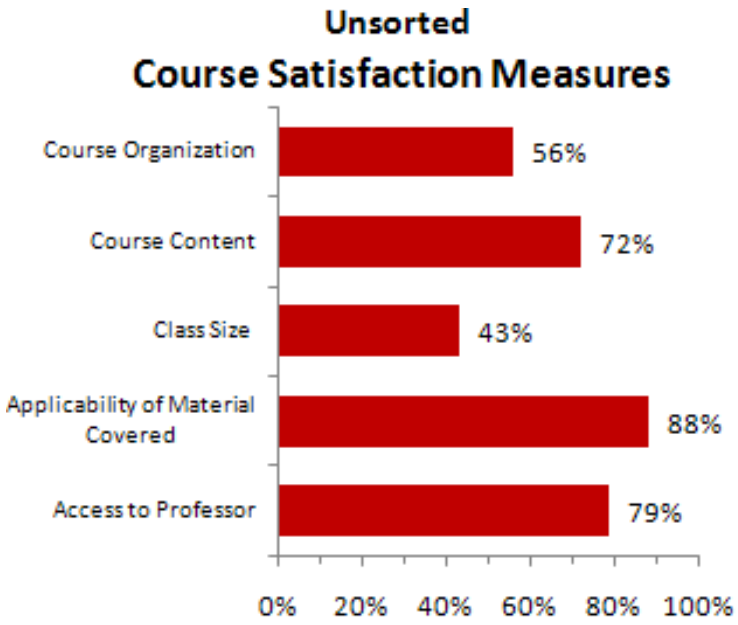
•If all values are positive integers, the scale should generally use 0 as a baseline. In the event that values include both positive and negative integers (e.g., in graphing differences in means), 0 should be the midpoint of the scale.

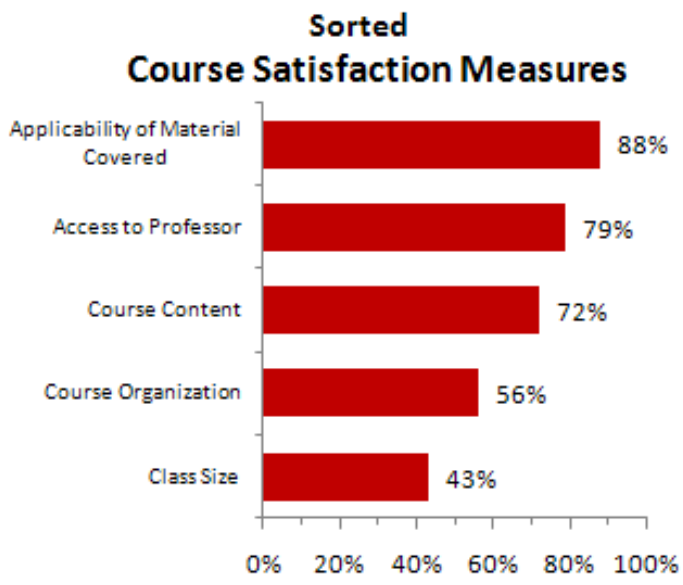
•Scale ranges should be standardized and not vary between graphs, when possible.

•Always try to avoid using 3-D features in a bar graph. The complexity of 3-D graphs makes them ineffective in conveying results to most audiences and there is usually a greater amount of data distortion that occurs.

•Bar graphs may be vertical or horizontal. The only difference between horizontal and vertical bars is that horizontal bar charts are seldom used to portray time series.

•To facilitate comparison and analysis, it is desirable that columns be sorted in some systematic order. The most common and visually effective schema is according to size of value.


Unsorted
Course Satisfaction Measures
Course Organization
56%
Course Content
72%
Class Size
43%
Applicability of Material Covered
88%
Access to Professor
79%
0%
20%
40%
60%
80%
100%

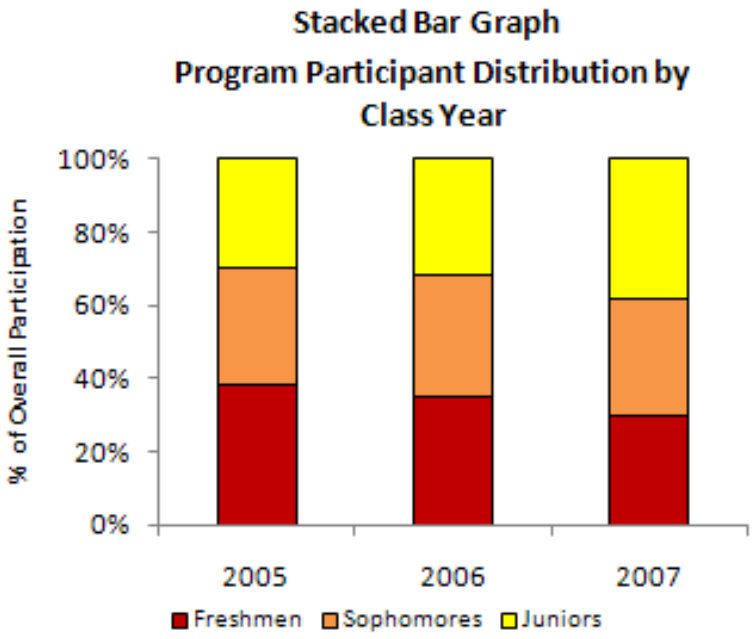

Sorted
Course Satisfaction Measures
Applicability of Material Covered
88%
Access to Professor
79%
Course Content
72%
Course Organization
56%
Class Size
43%
0%
20%
40%
60%
80%
100%

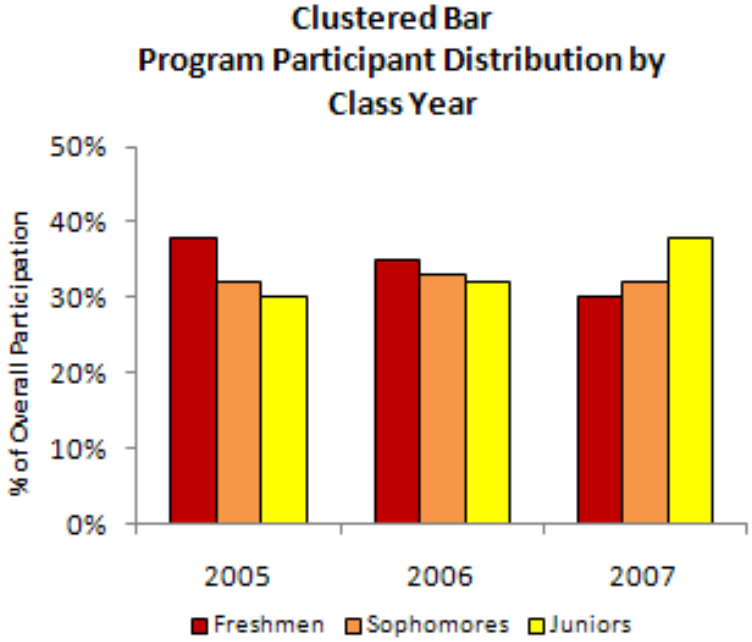
Continue

•It may also be desirable to order findings by a particular category such as class year (see Clustered Bar Graph below), where it is best to order sequentially from freshman to senior year or visa versa, or by grade achieved, where it is best to order by the standardized grade scale.

•Stacked bar graphs, which consist of one or more segmented bars where each segment represents the relative share of a total category, are generally not preferred because it is difficult to make comparisons among the second, third, or subsequent segments without a standard baseline.

•When graphing data from two or more different series, or different classes within the same series, it is preferable to create a bar graph that groups these values together, side by side (see below).


Stacked Bar Graph
Program Participant Distribution by Class Year
% of Overall Participation
100%
80%
60%
40%
20%
0%
2005
2006
2007
Freshmen
Sophomores
Juniors

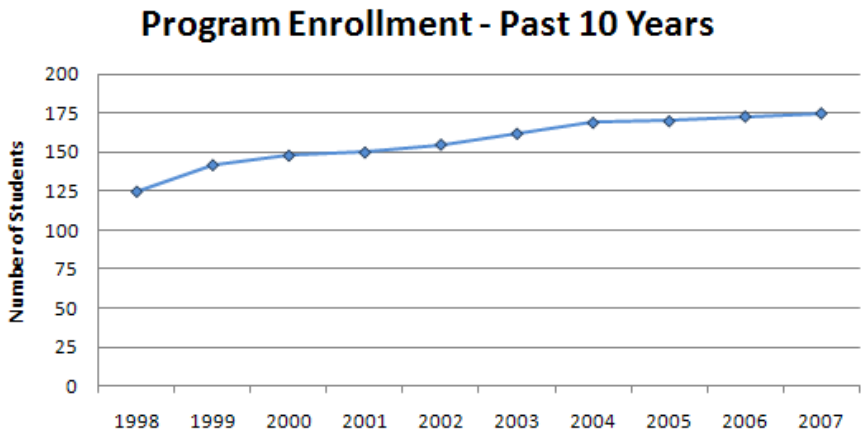

Clustered Bar
Program Participant Distribution by Class Year
% of Overall Participation
50%
40%
30%
20%
10%
0%
2005
2006
2007
Freshmen
Sophomores
Juniors

Line Graphs

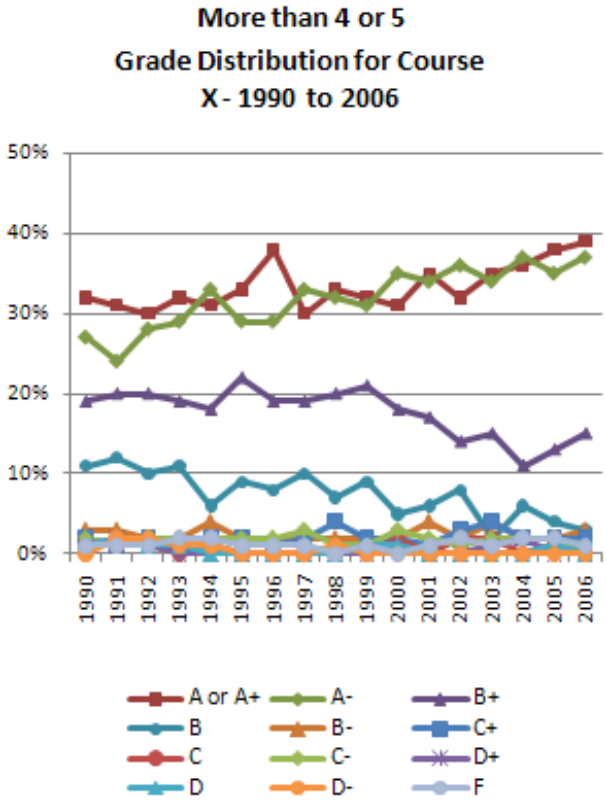
•Line graphs are most often used to display time series data (e.g., the average GPA of students in a starting cohort over their first eight semesters, or program enrollment over the past 10 years). See graph below.

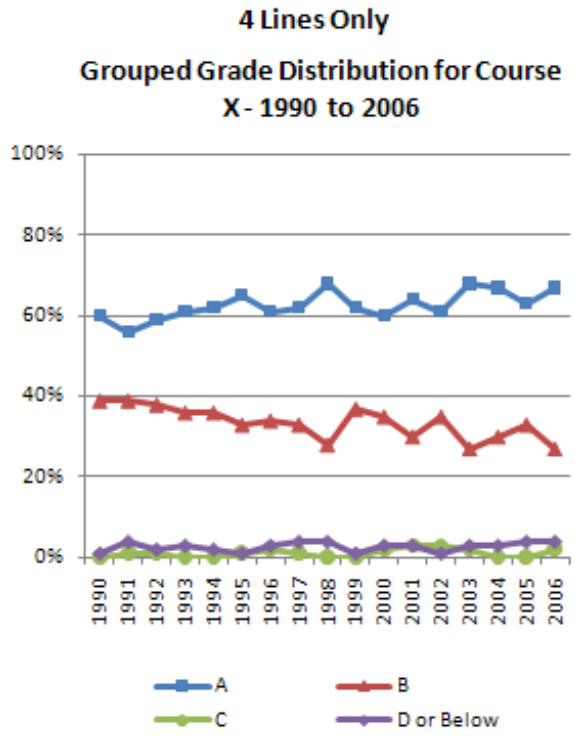
•Compared with bar graphs, line graphs are more effective in presenting five or more data points, but less effective in providing emphasis on differences over relatively few periods of time.

•When plotting time series data in a line graph, it is convention that the x-axis (horizontal) contains the categories of time (e.g., days of the week, months, years – depending on the data), and the y-axis (vertical) has frequencies of what is being measured (see graphs below).



Graphs with more than four or five lines tend to become confusing unless the lines are well separated.





Presenting Data in Tables

•Tables are the most effective way to present data for reference purposes.

•A table should always be given a meaningful, self-explanatory title.

•Each part of a table should be labeled clearly and abbreviations should be avoided.

•The number of digits and decimal places presented should be consistent and should be the minimum number that is compatible with the purpose of the table.

•It is usually better to convert counts into percentages, unless providing a simple frequencies table. More readers will care that 78% of students agreed with a statement rather than 325 students agreed.

•It is always important to include information in a table about the size of the sample from which a percentage is derived.

•A table should be constructed so that it is easy for readers to see differences and trends. If a table is presenting results from two or more different groups, years or survey cycles, it is good to include a column that indicates either the percentage change or the significance of differences observed.

Table 1

Comparison of Course Satisfaction Measures for Course X
2008 vs 2009

Satisfaction Measure	2009	2008
Applicability of Material Covered	88%	86%
Access to Professor	79%	68%
Course Content	72%	67%
Course Organization	56%	53%
Class Size	43%	44%

Improved Version of Table 1

Comparison of Course Satisfaction Measures for Course X
2008 vs 2009
(% "Satisfied" or "Very Satisfied")

Satisfaction Measure	2009 (N = 134)	2008 (N = 123)	Difference
Applicability of Material Covered	88%	86%	**2**
Access to Professor	79%	68%	**11**
Course Content	72%	67%	**5**
Course Organization	56%	53%	**3**
Class Size	43%	44%	**-1**

Note: Data taken from 2008 and 2009 end of term course evaluations for Course X

Continue

•It is best to present information in an order that makes sense to the reader by sorting from most frequently chosen response or highest score to lowest.

•A table should draw attention to the most salient points. Use boldface, italics, borders, and/or colors to draw attention to the most important figures, and put totals in boldface.

•Always note the source of data presented in a table.

•More complex tables that organize information by more than one level should be constructed to best reflect how data are grouped. It is best to merge cells that apply to more than one column in a table, rather than repeating the grouping information in more than one column. Shading can also provide greater organization and distinction between groups of data.

Table 2

2008 vs 2009 Divisional Enrollments by Gender

Academic Division	2009 Males	2009 Females	2008 Males	2008 Females
Physical Sciences	368	182	355	173
Natural Sciences	658	495	642	505
Humanities	352	435	375	415
Social Sciences	786	962	801	1002
Total	2164	2074	2173	2095

Improved Version of Table 2

2008 vs 2009 Divisional Enrollments by Gender

	2009		2008	
	Males	Females	Males	Females
Physical Sciences	368	182	355	173
Natural Sciences	658	495	642	505
Humanities	352	435	375	415
Social Sciences	786	962	801	1002
Total	**2164**	**2074**	**2173**	**2095**

Presenting discrete choice model results

	In vehicle Travel time (min)	Out of vehicle travel time (min)	Cost (CHF)	Monthly income (thousands CHF)	Young (below 25)	Old (over 65)	
Auto	-0.02 (-3.1)	-0.04 (-2.6)	-0.05 (-2.4)	0.00	-0.62 (-1.3)	-0.31 (-2.1)	
Rail	-0.03 (-3.4)	-0.04 (-2.6)	-0.07 (-2.1)	-2.34 (-1.7)	-0.02 (-0.7)	-0.04 (-1.6)	
Bus	-0.03 (-3.4)	-0.04 (-2.6)	-0.07 (-2.1)	-4.16 (-1.9)	0.00	0.00	
Alt n							