Effective Use of Column Charts

**Purpose**
This tool provides guidelines and tips on how to effectively use column charts to communicate research findings.

**Format**
This tool provides guidance on column charts and their purposes, shows examples of preferred practices and practical tips for column charts, and provides cautions and examples of misuse and poor use of column charts and how to make corrections.

**Audience**
This tool is designed primarily for researchers from the Model Systems that are funded by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR). The tool can be adapted by other NIDILRR-funded grantees and the general public.

The contents of this tool were developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90DP0012-01-00). The contents of this fact sheet do not necessarily represent the policy of Department of Health and Human Services, and you should not assume endorsement by the Federal Government.
Overview and Organization

- Simple Column Chart ........................................... 4
- Charting for Data Investigation .......................... 16
- Vertical Axis Scaling ........................................ 19
- 3-D Column Chart ........................................... 25
- Clustered Column Chart ..................................... 31
- Stacked Column Chart ....................................... 38
Overview

- Illustrations of types of column charts
- Guidance on the type of data and purposes for which column charts are well-suited
- Basic principles for presenting data visually using column charts
- Examples of preferred practices and practical tips for column charts
- Cautions and examples of misuse and poor use of column charts and how to correct such
- For some example charts, mock data was created to illustrate the concept and such mock examples should not be construed as real data
Simple Column Chart – Categorical Comparisons

- Data values displayed as vertical columns.
- Magnitude of each data element is represented by the height of the column.
- Can be used to display values for categorical items (percent voting by race and gender, death rates by cause of death etc).
- Shows comparisons among the categorical groups on the measure.
- Categories displayed on the horizontal axis.
Simple Column Chart – Categorical Comparisons

Age-Adjusted Death Rates Per 100,000 Standard Population, United States, 2012

Source: CDC, National Vital Statistics System, Mortality
Simple Column Chart – Trends

- Data values displayed as vertical columns.
- Magnitude of each data element is represented by the height of the column.
- Columns can also be used to display values of some data series over time (clients served by quarter, percent of households that are wireless-only, etc).
- Shows trends over time on the measure.
- Time displayed on the horizontal axis.
Simple Column Chart – Trends

Percent of Adults Living in Households that Do Not Have a Landline Telephone But Have At Least One Wireless Telephone - Wireless-Only Households - US

Source: National Health Interview Survey
Simple Categorical Column Chart: Cautions, Misuse, and Poor Use

Hospital Acquired Conditions [HAC] - Estimated Additional Inpatient Mortality
Deaths Per 1000 Cases - Per HAC - US - 2011

Unnecessary decimal detail on vertical axis.

Usually Better to Sort the Data, Value Descending

Gridlines generally unnecessary. Delete

Do not angle the categorical labels on horizontal axis – difficult to read.

If the category labels are lengthy, consider using a horizontal bar chart instead.
Simple Categorical Column Chart - Sorted Principles, Tips, Preferred Practices

Hospital Acquired Conditions [HAC] - Estimated Additional Inpatient Mortality Deaths Per 1000 Cases - Per HAC - US - 2011

- Central Line Associated Bloodstream Infections: 185.0
- Ventilator Associated Pneumonia: 144.0
- Postoperative Venous Thromboembolism: 104.0
- Pressure Ulcers: 72.0
- Falls: 55.0
- Surgical Site Infections: 28.0
- Catheter Associated Urinary Tract Infections: 23.0
- Adverse Drug Events: 20.0
- Obstetric Adverse Events: 1.5
Simple Bar Chart
Principles, Tips, Preferred Practices

Hospital Acquired Conditions [HAC] - Estimated Additional Inpatient Mortality Deaths Per 1000 Cases - Per HAC - US - 2011

- Central Line Associated Bloodstream Infections: 185.0
- Ventilator Associated Pneumonia: 144.0
- Postoperative Venous Thromboembolism: 104.0
- Pressure Ulcers: 72.0
- Falls: 55.0
- Surgical Site Infections: 28.0
- Cather Associated Urinary Tract Infections: 23.0
- Adverse Drug Events: 20.0
- Obstetric Adverse Events: 1.5

More readable as a horizontal bar chart

Source: CDC NCHS

<<Including source info is always helpful>>
Simple Categorical Column Chart: Cautions, Misuse, and Poor Use

City Proper - Population Estimate - 2014

- If you have more than a dozen or so categories, consider using a horizontal bar chart instead.
- Unnecessary precision in the column labels and on the vertical axis – reference in millions instead – e.g. 24.2
- Thin vertical axis itself is clutter – consider removing the axis line but retaining the axis interval labels
Simple Bar Chart
Principles, Tips, Preferred Practices

City Proper - Population Estimate - 2014

Shanghai 24.2
Karachi 23.5
Beijing 17.8
Delhi 17.1
Lagos
Istanbul
Guangzhou 14.2
Mumbai 12.7
Moscow 12.7
Dhaka 12.1
Cairo 12.0
Sao Paulo 11.9
Lahore 11.9
Shenzhen 11.3
Seoul 10.5
Jakarta 10.4
Kinshasa 10.0
Tianjin 9.7
Tokyo 9.1
Mexico City 8.9
Lima 8.7
Bengaluru 8.4
London 8.4
New York City 8.4
Bangkok 8.3

Population in Millions

More readable as horizontal bar chart.

Also, for interactive charting of data with many categories, viewers tend to judge magnitude more exactly with horizontal bar charts than vertical column charts.

Note: Legal political city boundaries. Does not include suburbs.
Simple Categorical Column Chart: Cautions, Misuse, and Poor Use

Generally should have no more than 9 value labels on the vertical axis.

External frames generally unnecessary clutter. Delete.

Don’t need legend for a single series chart. Delete.

Don’t need to repeat Axis Label if obvious or if addressed in title.
Generally add vertical and horizontal axis labels only when it is unclear from other chart information as to what the axes represent.

Make vertical and horizontal axis labels as brief as possible.

Horizontal axis label is useful in this example to differentiate CY FFY etc.
Simple Column Chart - Trends
Principles, Tips, Preferred Practices

Average Annual Percentage Change in Health Care Spending Per Capita - US

If you have provided a value for each data point, you could eliminate the labeling on the vertical axis entirely.

Source: CMS National Health Expenditures NHE
Interpretative Statement: Approx 956 Fatal Occupational Injuries Across all the Mondays in a Typical Year

Source: Bureau Labor Statistics.

Intermediate tick marks not needed. Generally, do not use any ticks on Categorical axes.

Why is Tuesday Seemingly Riskier?

Sat and Sun Lower as Expected
Charting for Data Investigation

Tuesday Fatal Occupational Injuries - By Year - US

The 9-11 terrorist attacks occurred on Tuesday Sep 11, 2001. 2,996 fatalities
Charting for Data Investigation

Total Fatal Occupational Injuries Per Day of Week of Incidence

Source: Bureau Labor Statistics.
Interpretative Statement: Approx 956 Fatal Occupational Injuries Across all the Mondays in a Typical Year. Excludes the fatalities from the 9-11 terrorist attacks.

Quick charts can often reveal anomalies, special cases, outlier data, as well as more prosaic data entry errors.
Vertical Axis Manipulation
Cautions, Misuse, and Poor Use

Potentially Preventable Inpatient Hospital Readmissions Within 30 Days of Discharge
Percent of Persons With Selected Target Discharge Diagnoses Readmitted Within 30 Days of Discharge - By Month of Discharge - Hospital System X [Mock Data]

At first glance, appears that hospital readmissions have been cut in half

Note: No numbers on the columns

Did not start vertical axis at zero
Vertical Axis Scaling
Principles, Tips, Preferred Practices

Potentially Preventable Inpatient Hospital Readmissions Within 30 Days of Discharge

Percent of Persons With Selected Target Discharge Diagnoses Readmitted Within 30 Days of Discharge - By Month of Discharge - Hospital System X [Mock Data]

- Started vertical axis at zero
- Better representation of actual improvement
  - Reduced readmissions from 18.0% to 16.3%
Vertical Axis Scaling
Principles, Tips, Preferred Practices

Potentially Preventable Inpatient Hospital Readmissions Within 30 Days of Discharge

Percent of Persons With Selected Target Discharge Diagnoses Readmitted
Within 30 Days of Discharge - By Month of Discharge - Hospital System X [Mock Data]

Start vertical axis at zero

Best representation of actual improvement. Reduced readmissions from 18.0% to 16.3%. But have not yet achieved goal of Under 15%.

Add Context.
Goal was to reduce readmissions to below 15% by Dec 2014.

Context [in this case, goal red line at 15%] is always useful
Conversely, you may have an indicator that is falling when it should be rising.

Manipulating the top end of the vertical axis range partially disguises the magnitude of the drop.

Note that in some circumstances it is preferable to top out a percentage vertical axis at 100%. So you will need to use your judgment on appropriate vs misleading axis scaling.
Contrast and Call Out

Median Visitors to Crabtree Mall - Evening Hours 6:00 PM Until Closing - By Day of Week During Period 22 Aug 2014 thru 14 Nov 2014

High School Football Home Games Reduce Mall Traffic by Half of Expected Volume on Friday Evenings During High School Football Season

![Bar chart showing median visitors on different days of the week.]

- Sunday: 3,000
- Monday: 4,050
- Tuesday: 4,200
- Wednesday: 4,050
- Thursday: 4,500
- Friday: 2,400
- Saturday: 5,250

Color contrast and interpretive statements often helpful
Column color and vertical separators help differentiate the impact of legislative changes in Legal Drinking Age and Blood Alcohol Content.
Column Charts – 3-D

- 3-D charts are generally a bad idea.
- The 3-D effect adds nothing, is more cluttered, creates confusion, and potentially obscures data.
- It’s harder to judge relative column height with 3-D columns.
- Readers are often confused as to whether the front edge of the 3-D column or the back edge of the 3-D column represents the magnitude of the value.
- Intent of all charts should be quick comprehension. 3-D distracts from the central message and slows the processing of the intended information.
If “Below 130” is the Desirable Range for LDL, which exercise groups are below 130?
3-D Column Chart
Cautions, Misuse, and Poor Use

LDL [Bad Cholesterol] Blood Level Among Men - By Level of Exercise - US

Even when you add value labels to a 3-D chart, often both the front and back edge of the columns don’t perceptually match the value.

Source: Mock Data
3-D Column Chart
Cautions, Misuse, and Poor Use

Adding skew to a 3-D chart further distorts the perception of the relative differences in the columns.
Without adding the value labels, can you tell if males or females have higher LDL levels and are there gender differences at various exercise levels?

Males have higher LDL levels, except at the Light Exercise level.
3-D Column Chart
Cautions, Misuse, and Poor Use

3-D columns have the potential to obscure other columns
Clustered Column Charts

- Clustered Column charts – also known as side-by-side column charts
- Useful for comparing multiple data series on one chart
- Vertical axis is numeric data (counts, percents etc)
- Compares values for different categories
- Columns are clustered side-by-side along the horizontal axis
- Generally the columns are color-coded by category to allow quick visual comparison by category
- Horizontal axis can be time-based or category-based
"Black Friday" headlines are now twice as common as "Thanksgiving" headlines during the month of November.
Clustered Column Charts – Legend Placement

Major National Newspaper Headlines Containing the Words: "Thanksgiving" vs "Black Friday"

Place legends to maximize the image of the chart per se. Place legends within the chart plot if possible, or above the chart or below the chart. But don’t place legend to the side [wastes chart space and throws off chart centering]
Generally, do not add unnecessary backgrounds to your charts – detracts from the message. Maximize message. Reduce noise.
Clustered Column Charts – Prevalence, Two Factors

Percent of Persons with Depression Symptoms Who Sought Help From a Mental Health Professional in the Prior 12 Months
By Depressive Symptom Severity and Race-Ethnicity, Age 12 and Older, US, 2009-2012

Across all Race-Ethnicity Groups, Persons with Increasingly Severe Symptoms Are More Likely to Seek Help
At Each Symptom Severity Level, White Non-Hispanics Are More Likely to Seek Help and Hispanics Are the Least Likely to Seek Help

Source: CDC NCHS
Clustered Column Charts – Prevalence, Two Factors

Percent of Persons with Depression Symptoms Who Sought Help From a Mental Health Professional in the Prior 12 Months
By Depressive Symptom Severity and Race-Ethnicity, Age 12 and Older, US, 2009-2012

Tip: When you wish the reader to compare differences (such as the differences in help-seeking by race-ethnicity within each level of depression), place the columns close together. Humans are better able to accurately judge and compare column heights when placed closer together.

Similarly, when visual separation is intended (such as among the three levels of depression, increase the gap width).
Clustered Column Charts – Distribution, Two Factors

Distribution of Persons Receiving Social Security Disability by Type of Beneficiary and by Disability Diagnostic Group - US - December 2013

For Disabled Workers and Widow(er)s, the Musculo-Skeletal Diagnostic Group is the Most Common.
For Disabled Adult Children, Intellectual Disability is the Most Common.

Source: Social Security

Generally, do not attempt to display more than 8 series in a clustered column chart
Stacked Column Charts

- Used to show relationship of components to the whole
- Vertical axis is numeric (counts, percents etc)
- Horizontal axis can be time-based or categorical (group-based)
Since 2000: Shotguns have remained approx stable. Rifles have almost doubled. Pistols have almost quadrupled.
Now can see that rifles and shotguns have decreased as a proportion of total firearms.

Source: Mock Data
Often difficult to see relationships clearly in a stacked column chart. Bottom layer of a stacked column and the total overall trend are easily interpreted in a stacked column chart. But the true trend in the additional layered columns is often difficult to reliably discern in a stacked column chart.

Source: Bureau of Labor Statistics
Unwanted births and substantially mistimed births are relatively less common as income rises.