Assignment 3

Marissa

2/10/2020

Effectively Communicating Numbers: Selecting the Best Means and Manner of Display

Executive Summary

- Ability to display data graphically is not intuitive
  - Requires set of visual design skills(* Quantitative information: the numbers that measure performance, identify opportunities, forecast the future

Introduction

- Six fundamental skills:
  - Skills 3 - 6 are for if one or more graphs are needed
    1. Determine your message and identify the data necessary to communicate it
    2. Determine if a table, graph, or combination of both is needed to communicate your message
    3. Determine the best means to encode the values
    4. Determine where to display each variable
    5. Determine the best design for the remaining objects
    6. Determine if particular data should be featured above the rest, and if so, how

General Concepts and Practices

Tables Versus Graphs

- Table:
  - Data are expressed in the form of text
  - Data are arranged in columns and rows
  - Best: to look up individual values or quantitative values must be precise
- Graph:
  - Data are expressed graphically (as picture)
  - Data are displayed in relation to one or more axes along which run scales that assign meaning to the values
  - Best: message to communicate resides in the shape of the data (patterns, trends, exceptions)

Quantitative Versus Categorical Data

- Quantitative information: consists of numbers and data that identifies what the numbers mean
  - Quantitative Data: the numbers
  - Categorical Data: the labels that tells us what the numbers measure
- Three Types of Categorical Scales
  - Nominal
    - Discrete items that belong to a common category, but don't relate to one another in any particular way
    - Items have no particular order
    - Items do NOT represent quantitative values
  - Ordinal
    - Items have an intrinsic order
    - Items do NOT represent quantitative values
  - Interval
    - Items have an intrinsic order
    - Items represent quantitative values

The Seven Common Relationships in Quantitative Business Data

- Number become meaningful only when compared to related numbers
- Seven Relationship Types:
  - Time-Series Relationships
  - Series of measures taken across equal intervals of time
    - Most common relationship in quantitative business data
    - Time can be divided into varying intervals of varying duration (years, quarters, months, weeks, days, hours)
    - Reveals trends and patterns to make informed decisions
  - Ranking Relationships
  - Values are sequenced by size, large to small or small to large (Performance of sales, expense of departments)
    - Easier to compare values by placing those that are most similar near one another
  - Part-to-Whole Relationships
  - Reveal the portion that each value represents to some whole
    - Useful to see how something is divided into parts; percentage of each part to the whole
  - Deviation Relationships
  - Displayed to feature how one or more set of values differ from some reference set of values
    - Business: how some actuals deviate form predefined target
  - Distribution Relationships
    - How values are spread across their entire range
    - Learn the shape of the distribution (normal, skewed, gaps, concentrations)
  - Correlation Relationships
    - Pairs of values, each measuring something different about an entity are displayed to reveal if there is significant relationship between them
    - Help predict, take advantage of, or avoid particular behaviors
  - Nominal Comparison Relationships
    - Displays nothing but a nominal comparisons

The Best Means to Encode Quantitative Data in Graphs

- Two-dimensional graphs work well:
  - Line lengths
  - 2-D position
- Four types of objects that work best for encoding quantitative values:
  - Points
    - Smallest of the objects
    - Shape: dots, squares, triangles, Xs, dashes
    - Strengths: (1) used to encode values along two quantitative sales simultaneously , (2) can be used when the quantitative scales do not begin at zero
    - Emphasize individual values
  - Lines
    - Connect individual values in a series, emphasizing the shape of the data
    - Trends, patterns, and exceptions
    - Only use lines to encode data along an interval scale
  - Bars
    - Encode data that emphasizes individual values powerfully
    - (1) 2-D position of the bar's endpoint in relation to quantitative scale
    - (2) length of the bar
    - Quantitative scale MUST include zero
  - Boxes
    - Both ends encode quantitative values
    - Box-And-Whisker Plot

The Best Practices for Formatting Graphs to Remove Distractions

- Anything that does NOT contribute in an essential way to the meaning of a graph is a distraction that harms communication
  - Solution: remove it

A Step-By-Step Graph Selection and Design Process

- Order of steps doesn't have to followed strictly

Determine Your Message and Identify Your Data

- Determine what you want to say first
- Before you can communicate data, must know what the data means and know what's important based on the needs of the audience

Determine If a Table, Graph, or Both is Needed to Communicate Your Message

- Look up and compare values: table
- Message in shape: graph

Determine the Best Means to Encode the Values

- Nominal Comparison: bars, points
- Time-Series: lines, bars, points
- Ranking: bars, points
- Part-To-Whole: bars, stacked bars
- Deviation: lines, points connected with lines
- Frequently Distribution: bars, lines
- Correlation: points and trend line

Determine Where to Display Each Variable

- Categorical Scales: X-axis
- Quantitative Scales: Y-axis
- Horizontal Bars when these two conditions exist:
  - Text labels are associated with that bars are long
  - There are many bars
- Using small multiples to support an additional variable

Determine the Best Design for the Remaining Objects

- Determine the Range of the Quantitative Scale
  - Bars MUST start at zero
  - Others may have a narrow scale (little below lowest data value and little above highest)
    - Make obvious that you narrowed the scale
- If a Legend is Required, Determine Where to Place It
  - Best: label data directly
- For Each Axis, Determine If Tick Marks Are Required and How Many
  - Tick marks are only necessary on quantitative scales
    - No real purpose on categorical scales
    - Too many = clutter
    - Too few = not enough detail to interpret data values
- Determine the Best Location for the Quantitative Scale
  - Y-Axis: left side, right side, or both sides of graph
  - X-Axis: top, bottom, or both
  - Places scale nearest to the values you wish to emphasize
- Determine If Grid Lines are Required
  - Gridlines are useful when:
    - Values cannot be interpreted with the necessary degree of accuracy
    - Subset of points in multiple related scatter plots must be compared
  - Make grid lines barely visible
- Determine What Descriptive Text is Needed
  - Text often needed:
    - Descriptive title
    - Axis titles
  - Include notes to describe the graph, what should be observed, how to read the graph

Determine If Particular Data Should Be Featured, and If So, How

- Highlighting particular data if it is more important than the rest:
  - Encode these items using bright or dark colors
  - Bars: place border around desired bar
  - Lines: makes lines thicker
  - Points: larger points or stand-alone color

Conclusion

- Order to communicate information:
  - 1. Determine your message and identify the data necessary to communicate it
  - 2. Determine if a table, graph, or combination of both is needed to communicate your message
  - 3. Determine the best means to encode the values
  - 4. Determine where to display each variable
  - 5. Determine the best design for the remaining objects
  - 6. Determine if particular data should be featured above the rest, and if so, how