

# Plots & Charts

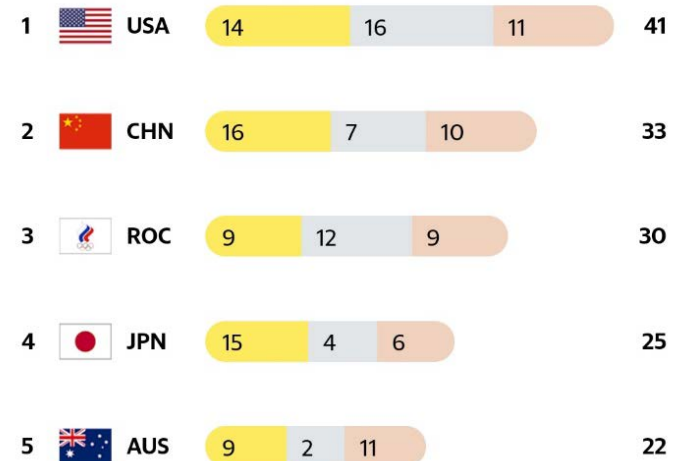
Goal: know some useful principles (or *guideline*) for generating effective plots and charts; know some basic types of plots/charts



# Why Should We Care?

- Everyone uses plots and/or charts (graphs)
- But...most people ignore or are unaware of some **simple principles**
- Default plotting tools (or default settings) are not always the best
- More importantly, **it is easy to lie or deceive people with bad plots**

Tokyo Games - Medal Race



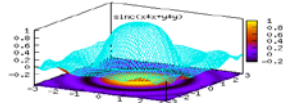


gnuplot homepage

FAQ  
Download  
Demos  
Documentation  
External Links  
Screenshots  
Contributed scripts and files  
Tutorials, learning, and help  
Building from CVS source  
More on patching and building



# Some examples of available plotting tools !



<https://www.r-project.org/>

<http://www.gnuplot.info/>

<https://www.tableau.com/>

<http://office.microsoft.com/en-us/excel/default.aspx>

<http://www.mathworks.com/>

<https://vega.github.io/voyager2/>

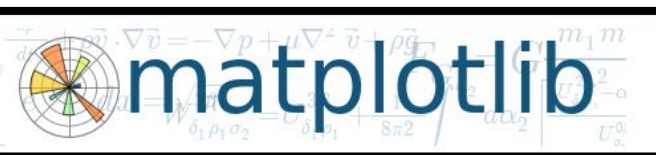
<https://vega.github.io/vega-lite/>

<http://www.sigmaplot.com/products/sigmaplot/sigmaplot-details.php>

<http://matplotlib.sourceforge.net/>

<http://www.wolfram.com/>

## And many more!!



# What Can Plots Do?

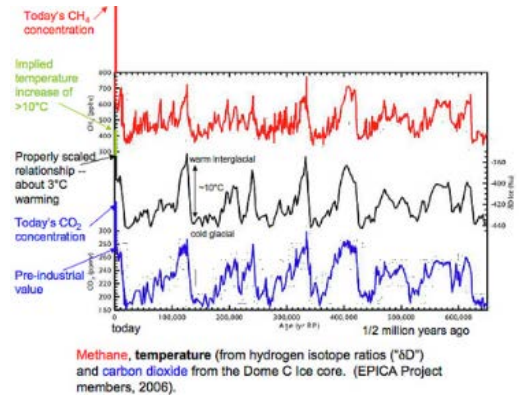
- **Data analysis and communication**

- **What plots do:** In a simplistic view, plotting reduces a large amount of information to a smaller form that is more easily understood via certain graphical representation.

Table 7-4 Direct global warming potentials of several well-mixed trace gases relative to CO<sub>2</sub>. The GWPs of the various non-CO<sub>2</sub> species are calculated for each of five time horizons (20, 50, 100, 200 and 500 years) using, as a IPCC, the carbon cycle model of Sagan/Pfister (1983). (Note that IPCC contained a typographical error which led to incorrect values for the direct GWP of methane.)

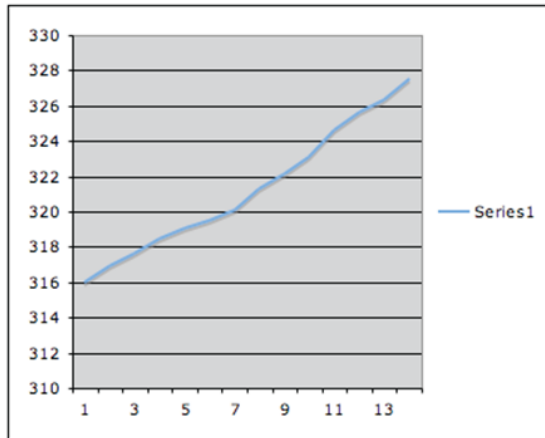
| Gas                | Lifespan (years) | Time Horizons |          |           |           |           |
|--------------------|------------------|---------------|----------|-----------|-----------|-----------|
|                    |                  | 20 years      | 50 years | 100 years | 200 years | 500 years |
| CO <sub>2</sub>    | 8                | 1             | 1        | 1         | 1         | 1         |
| CH <sub>4</sub>    | 10.5             | 35            | 29       | 11        | 7         | 4         |
| N <sub>2</sub> O   | 132              | 260           | 270      | 270       | 240       | 170       |
| CFC-11             | 55               | 4500          | 4100     | 3400      | 2400      | 1400      |
| CFC-12             | 116              | 7100          | 7400     | 7100      | 6200      | 4100      |
| HCFC-22            | 15.8             | 4200          | 2800     | 1600      | 970       | 540       |
| CFC-113            | 110              | 4600          | 4700     | 4500      | 3900      | 2500      |
| CFC-114            | 220              | 6100          | 6700     | 7000      | 7000      | 5800      |
| CFC-115            | 550              | 5500          | 6200     | 7000      | 7800      | 8500      |
| HCFC-123           | 1.71             | 330           | 130      | 90        | 55        | 30        |
| HCFC-124           | 6.9              | 1500          | 760      | 440       | 270       | 150       |
| HFC-125            | 40.5             | 5200          | 4500     | 3400      | 2300      | 1200      |
| HFC-134a           | 15.6             | 3100          | 1900     | 1200      | 730       | 400       |
| HCFC-141b          | 10.8             | 1800          | 980      | 580       | 350       | 200       |
| HCFC-142b          | 22.4             | 4800          | 2300     | 1800      | 1100      | 630       |
| HFC-143a           | 64.2             | 4700          | 4500     | 3800      | 2800      | 1600      |
| HFC-152a           | 1.8              | 530           | 250      | 150       | 89        | 49        |
| CCL <sub>4</sub>   | 47               | 1800          | 1600     | 1300      | 860       | 480       |
| CH <sub>3</sub> Cl | 6.1              | 360           | 170      | 100       | 62        | 34        |
| CF <sub>3</sub> Br | 77               | 5600          | 5500     | 4900      | 3800      | 2300      |

SAOD Table 7.2 (p. 7.6)

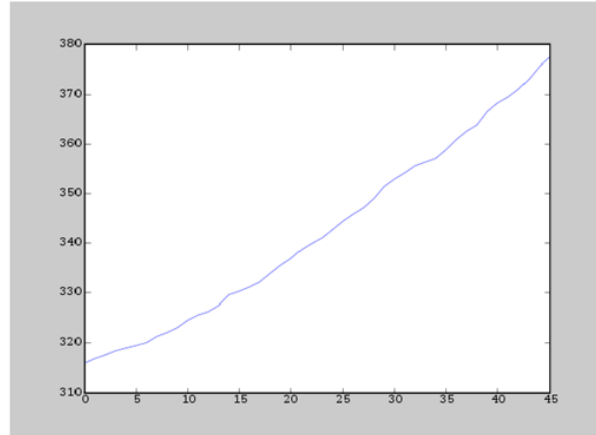


- **Benefit:** Reduction of the data to its simplest and cleanest form, such that the **relationships/patterns** inherent in the data (points) are easily perceived.

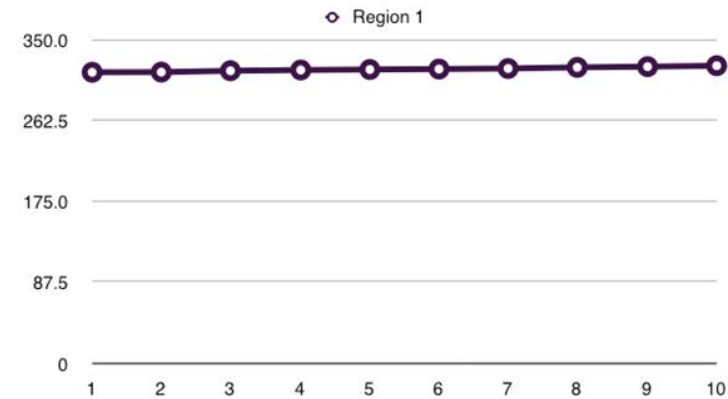
# Examples of plots generated by a few tools using their default setting



Default Excel Plot



Default Matplotlib/Matlab Plot



Default Pages Plot

**They look different visually!**

*The above examples demonstrate two important points:*

**First**, there is no obvious standard for what a plot should look like. This is easy to see by the differences in the axes and scale lines, the data rectangle inside the plot, and the actual representation of the data values.

**Second**, creating a plot is an iterative process. Different data and emphasis may need different iterative processes.

**Given these many types and styles of plots/charts, how to determine which one(s) are good or more effective?**

# Graphical Excellence

Graphical excellence is that which gives to the viewer the *greatest number of ideas* in the *shortest time* with the *least ink* in the *smallest space*.

And it requires **telling the truth about the data.**



# Summary of Tufte's Principles

## **1. Tell the truth (Expressiveness)**

Graphical integrity, authenticity, completeness...

## **2. Do it effectively (effectiveness) with clarity, precision, emphasis, ...**

The information provided here should be considered as guidelines

# PRINCIPLES OF PLOTTING

*Visualizing Data* [Cleveland 93] and *Elements of Graphing Data*  
[Cleveland 94] by William S. Cleveland

**There are other similar principles!!!!**

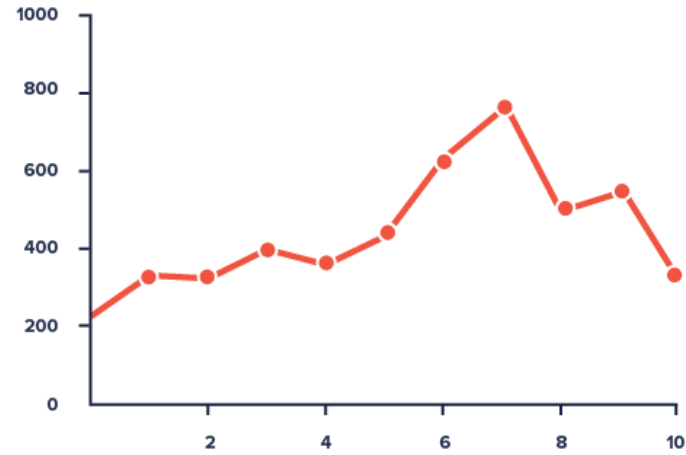
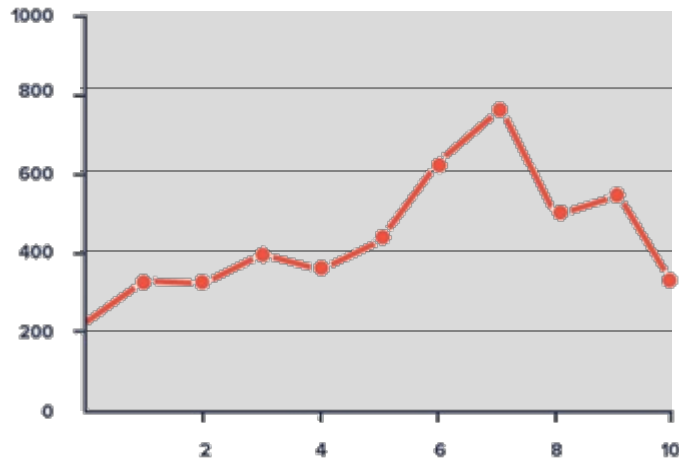
# Principles of Plotting

- Improving the **vision**
  - Improve the readability of the plot
- Improving the **understanding**
  - Ensure that the analysis of the plot is effectively communicated.

# Improving the Vision

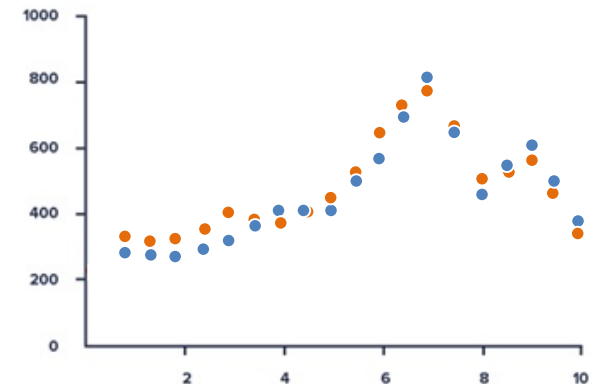
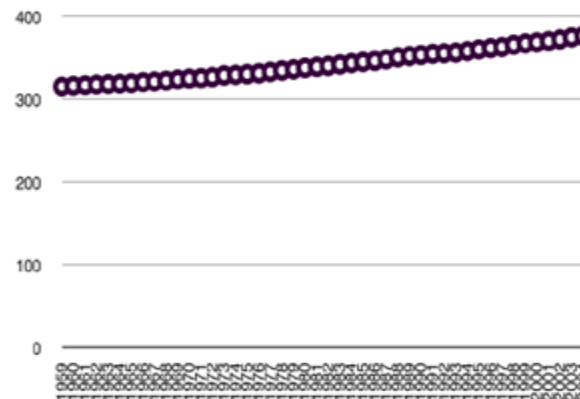
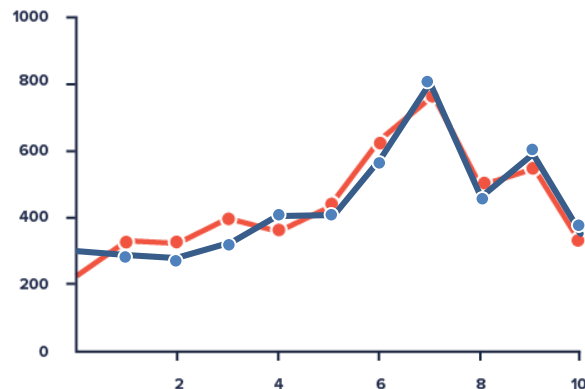
- Principle 1: Reduced clutter, Make data stand out
  - The main focus of a plot should be on the data itself, any superfluous elements of the plot that might obscure or distract the observer from the data needs to be removed.

**Less is more!!!!**



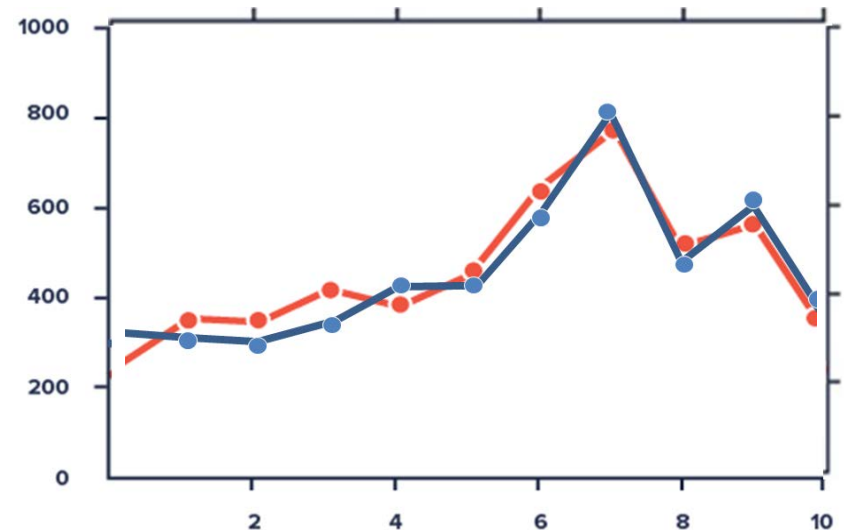
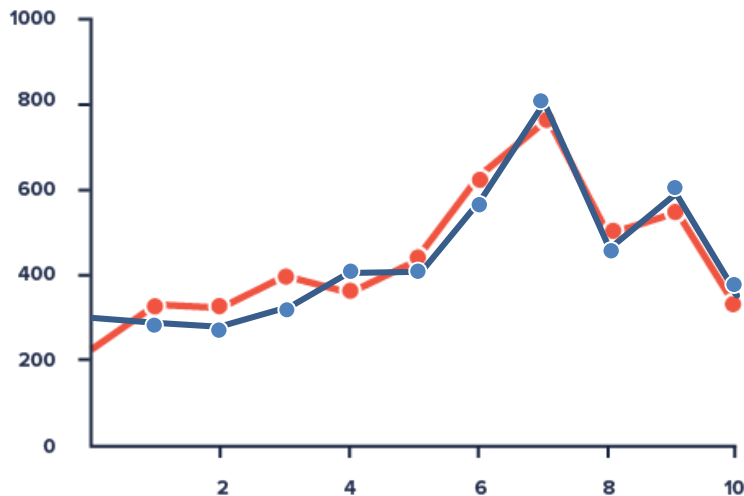
# Improving the Vision

- Principle 2: Use visually prominent graphical elements to show the data.
  - Connecting lines should never obscure points and points should not obscure each other.
  - If multiple data sets are represented in the same plot (superposed data), they must be visually separable.
  - If this is not possible due to the data itself, the data can be separated into adjacent plots that share an axis.



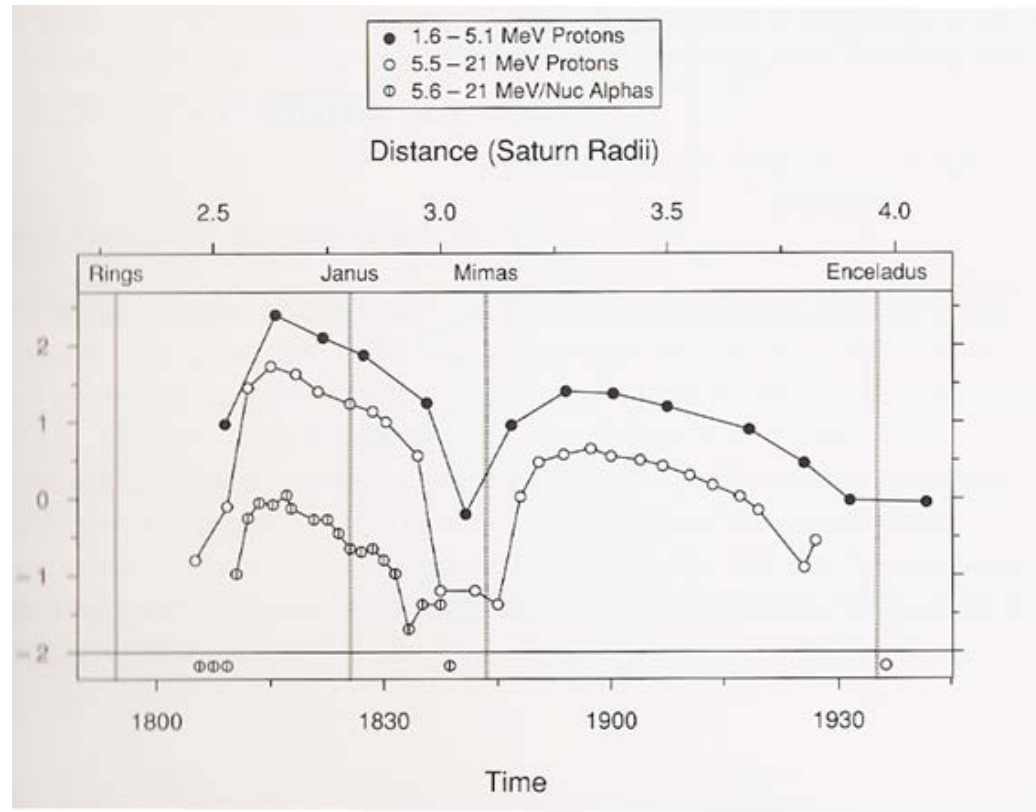
# Improving the Vision

- Principle 3: Use proper scale lines and a data rectangle
  - Two scale lines should be used on each axis (left and right, top and bottom) to frame to data rectangle completely (**optional, may against the closure principle**).
  - Add margins for data to make the plot prominent.
  - Tick-marks outs and **3-10** for each axis.



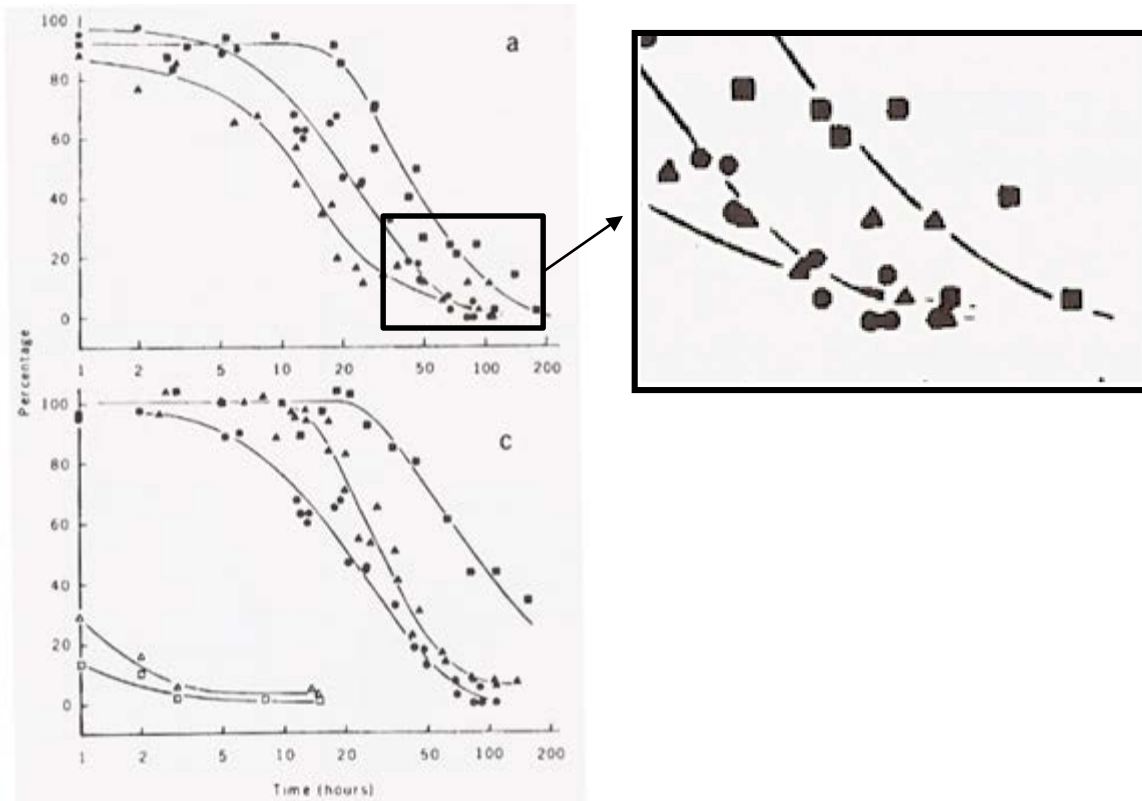
# Improving the Vision

- Principle 4: Reference lines, labels, notes, and keys (optional).
  - Reference lines are only used to show the thresholds within data.
  - Only use them sparsely when necessary and **don't let them obscure data**.



# Improving the Vision

- Principle 5: Superposed data set
  - Symbols should be separable and data sets should be easily visually assembled.



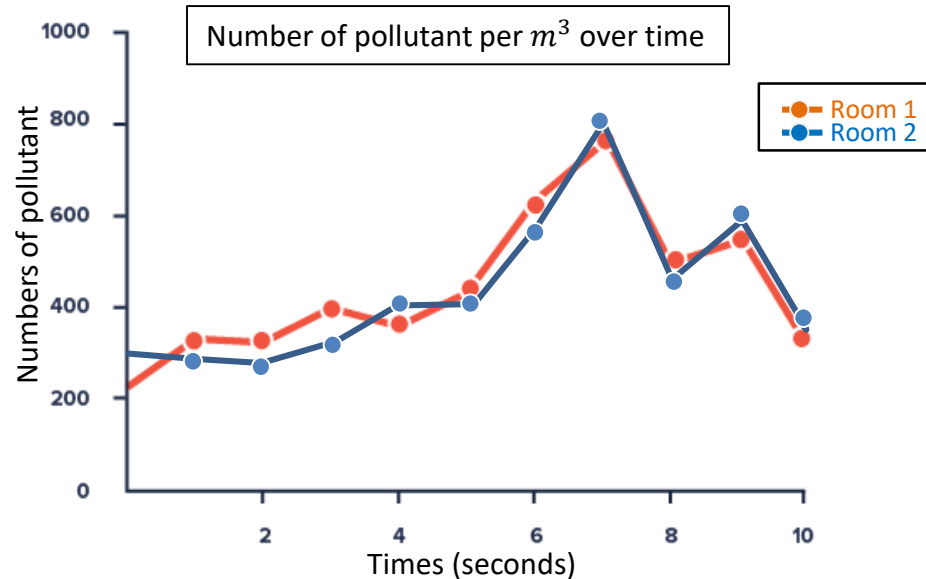


# Summary of Principles

- Improve **vision**
  1. **Reduced clutter**, make data stand out
  2. Use **visually prominent** graphical elements
  3. Use proper **scale lines and tick marks**
  4. Reference lines, labels, notes, and keys should not block data
  5. Superposed data set easily separable

# Improving the Understanding

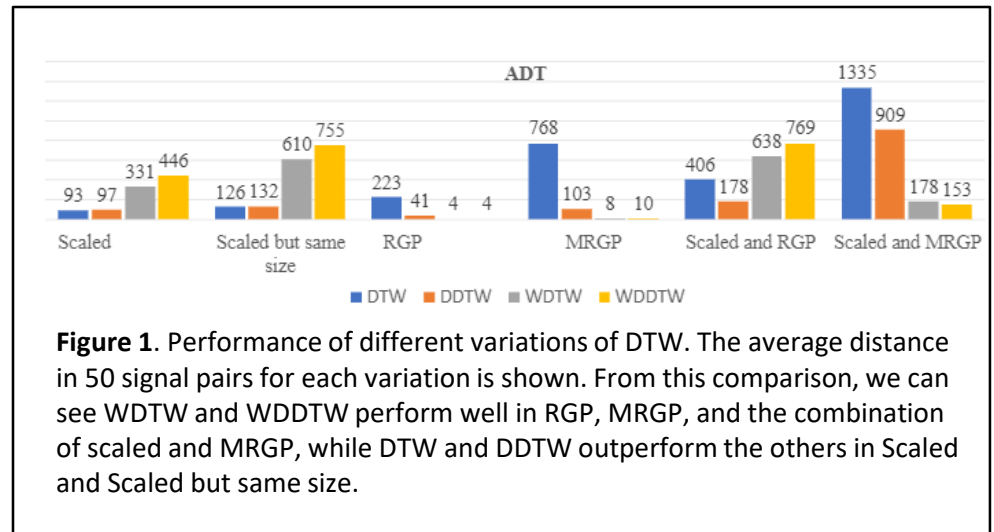
- Principle 1: Add plot title, axis labels!
  - Add a title to the plot to summarize what the plot is about.
  - Name the axes and provide units if possible
  - Add legends if needed



# Improving the Understanding

- Principle 2: Provide explanations and draw conclusions
  - A graphical representation is often the means in which a hypothesis is confirmed, or results are communicated.
  - Describe everything, draw attention to major features, describe conclusions

Explain everything in the plot.  
Do not let the observer guess.

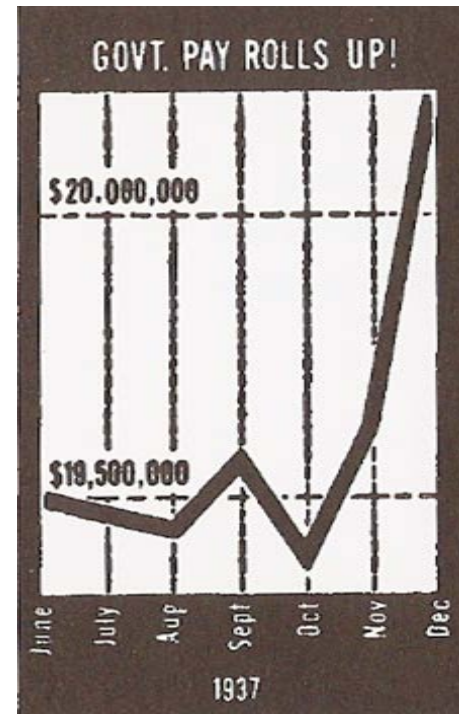
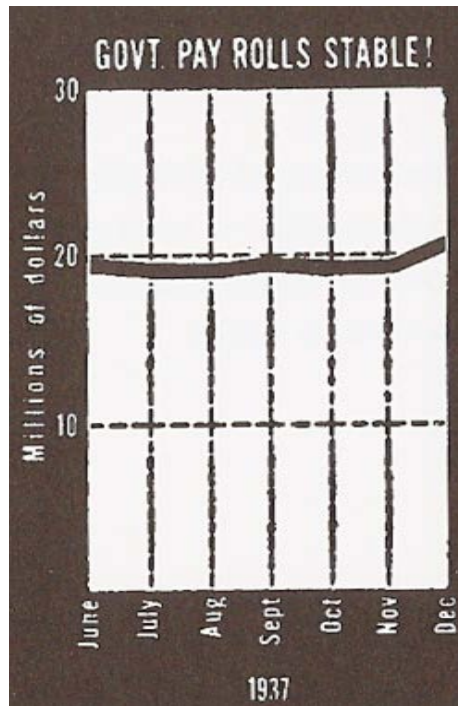


**Figure 1.** Performance of different variations of DTW. The average distance in 50 signal pairs for each variation is shown. From this comparison, we can see WDTW and WDDTW perform well in RGP, MRGP, and the combination of scaled and MRGP, while DTW and DDTW outperform the others in Scaled and Scaled but same size.

**Add figure caption!**

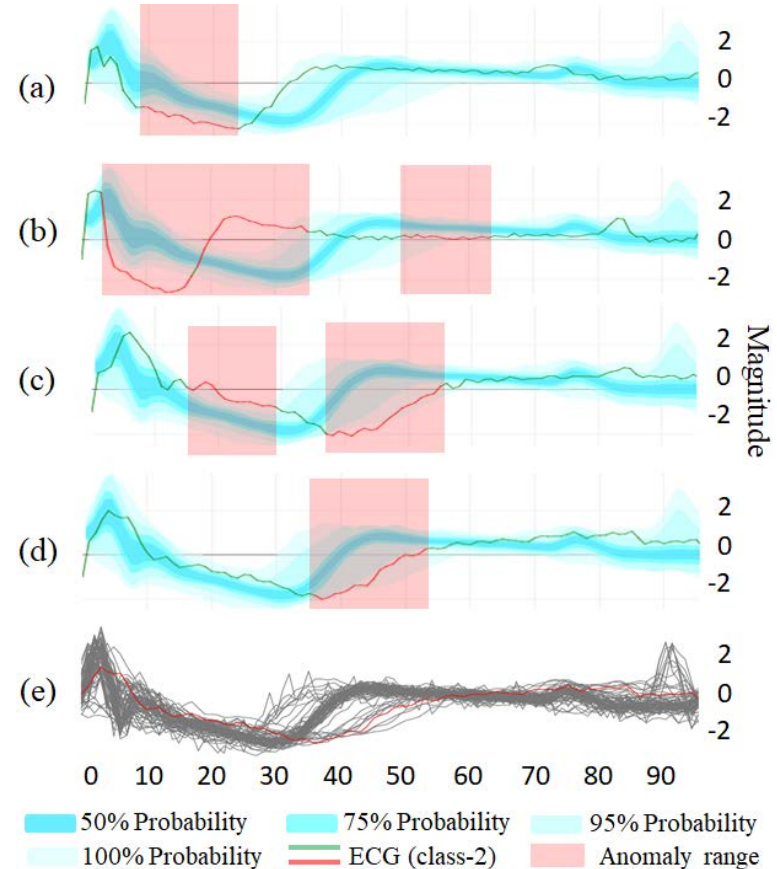
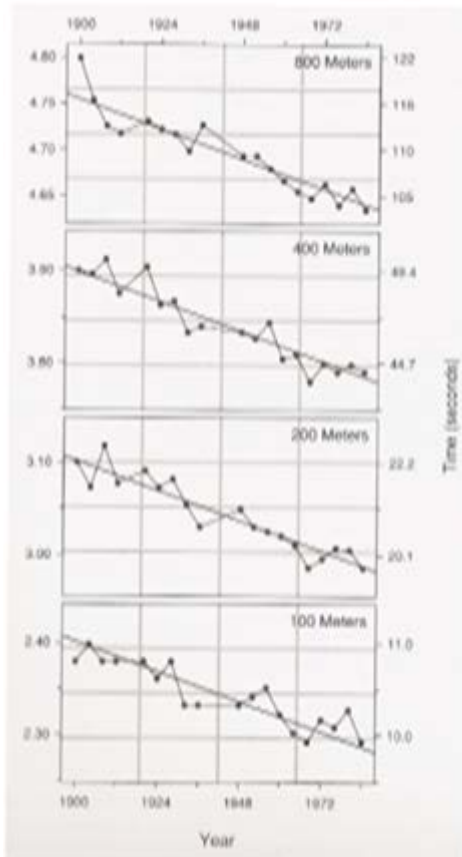
# Improving the Understanding

- Principle 3: Use all available space (Optional).
  - Fill the data rectangle as much as you can
  - Use absolute values for scientific data! But, you may use relative scale for comparison!



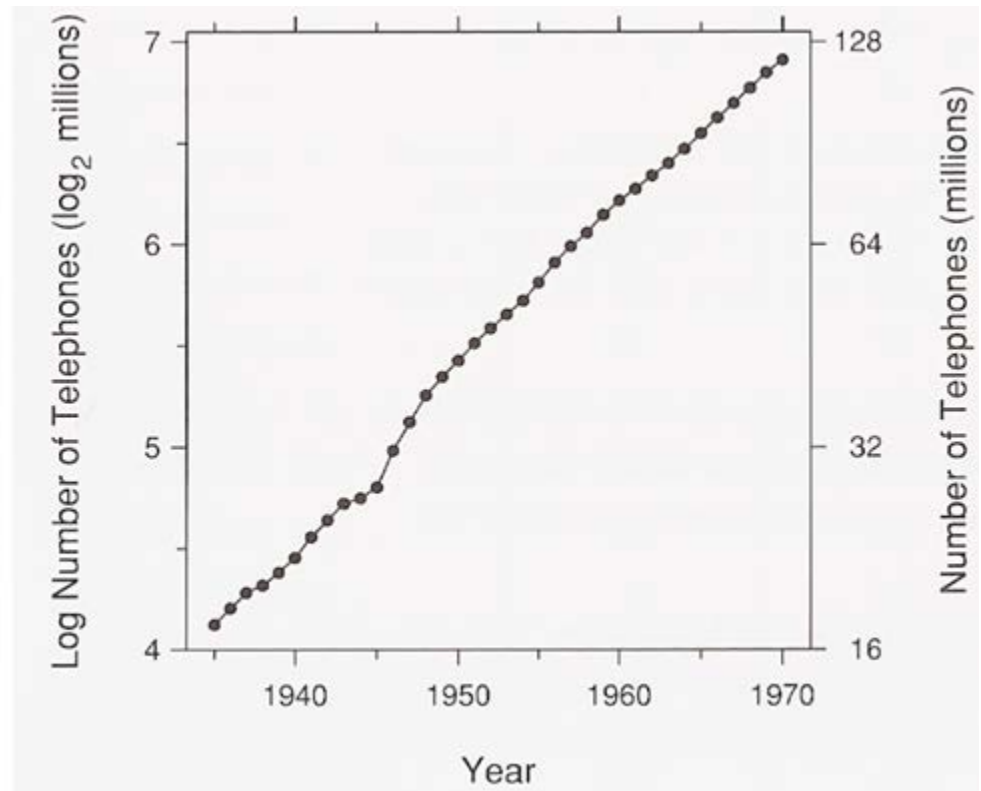
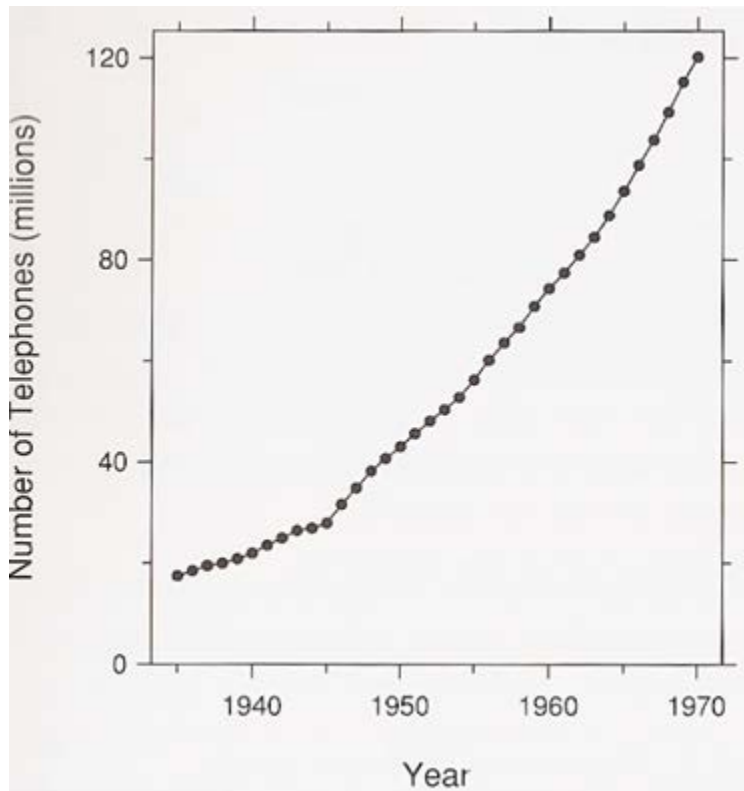
# Improving the Understanding

- Principle 4: **Align** juxtaposed plots
  - Make sure scales match and graphs are aligned



# Improving the Understanding

- Principle 5: Use log scales **when appropriate**
  - Used to show percentage change, multiplicative factors and skewness



# Summary of Principles

- Improve **vision**
  1. **Reduced clutter**, make data stand out
  2. Use **visually prominent** graphical elements
  3. Use proper **scale lines and tick marks**
  4. Reference lines, labels, notes, and keys should not block data
  5. Superposed data set easily separable
- Improve **understanding**
  1. **Provide title, axis labels, and legends**
  2. **Provide caption with explanations** and draw conclusions
  3. Use all available space when appropriate
  4. **Align** juxtaposed plots
  5. Use log scales when appropriate

# Other Useful Guidelines

- Show the data clearly
- Use simplicity in design of the graph
- Use alignment on a common scale
- Keep the visual encoding transparent
- Use standard forms that work

(<https://scc.ms.unimelb.edu.au/resources-list/data-visualisation-and-exploration/data-visualisation>)

## An efficient/iterative way to proceed while constructing a graph is to:

- 1) figure out the contexts and the message
- 2) figure out the way of presenting it, so the type, layout and style of graph to be used
- 3) construct the graph
- 4) check [Tables 1](#) and [2](#) and revise the graph accordingly
- 5) check whether the revised version conveys the message you want it to convey
- 6) check [Tables 1](#) and [2](#) and revise the graph accordingly
- 7) and so on...

Table 1 – Elements to check while constructing a graph. Refer to the text for details.

|                       |   |
|-----------------------|---|
| Data points           | symbol, size, color, overlap  |
| Lines                 | type, width, color, overlap   |
| Color                 | check whether needed at all; check whether all elements are easily distinguished  |
| Box and axes          | type of box, aspect, minimum and maximum value, data rectangle and scale-line rectangle, tick marks (number of, location, direction, length and width), tick marks' labels (font type, length, rotation, numbering style, abbreviations of text labels) |
| Legend                | check whether needed, position, box around, size, elements within (see above for "data points" and "lines", and "labels" within "box and axes")   |
| Background            | of the graph; of bars, cross-hatching, shading, color   |
| Text inside the graph | check whether needed; font style, size and color  |
| Captions              | check whether informative, explaining everything that is needed to understand the graph and the   |

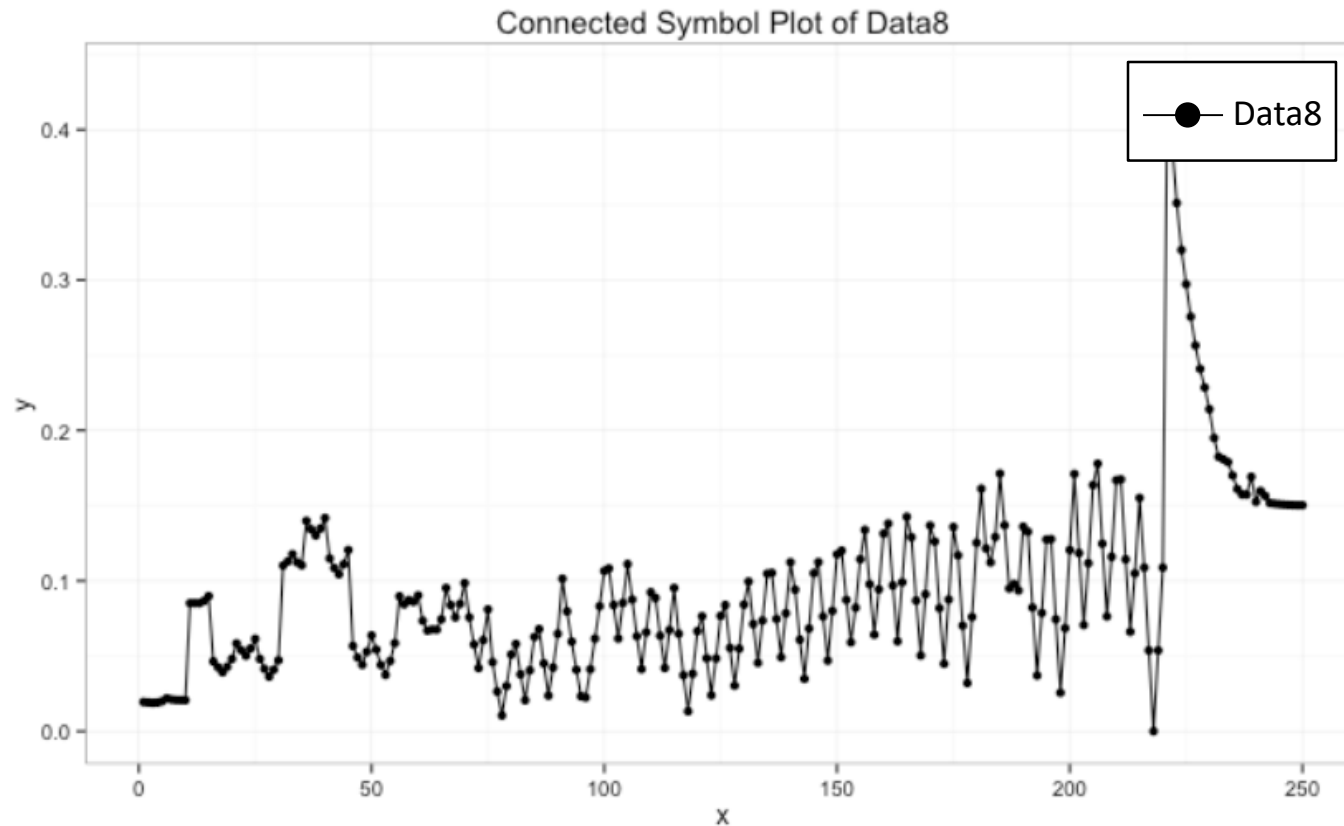
Table 2 – Clear vision principles given by Cleveland (1994).

1. Make the data stand out. Avoid superfluity.
2. Use visually prominent graphical elements to show the data.
3. Use a pair of scale lines for each variable. Make the data rectangle slightly smaller than the scale-line rectangle. Tick marks should point outward.
4. Do not clutter the interior of the scale-line rectangle.
5. Do not overdo the number of tick marks.
6. Use a reference line when there is an important value that must be seen across the entire graph, but do not let the line interfere with the data.
7. Do not allow data labels in the interior of the scale-line rectangle to interfere with the quantitative data or to clutter the graph.
8. Avoid putting notes and keys inside the scale-line rectangle. Put a key outside, and put notes in the caption or in the text.
9. Overlapping plotting symbols must be visually distinguishable.
10. Superposed data sets must be readily visually assembled.
11. Visual clarity must be preserved under reduction and reproduction.
12. Put major conclusions into graphical form. Make captions comprehensive and informative.

[https://www.scielo.br/scielo.php?pid=S0103-90162010000400017&script=sci\\_arttext&tlng=en](https://www.scielo.br/scielo.php?pid=S0103-90162010000400017&script=sci_arttext&tlng=en)

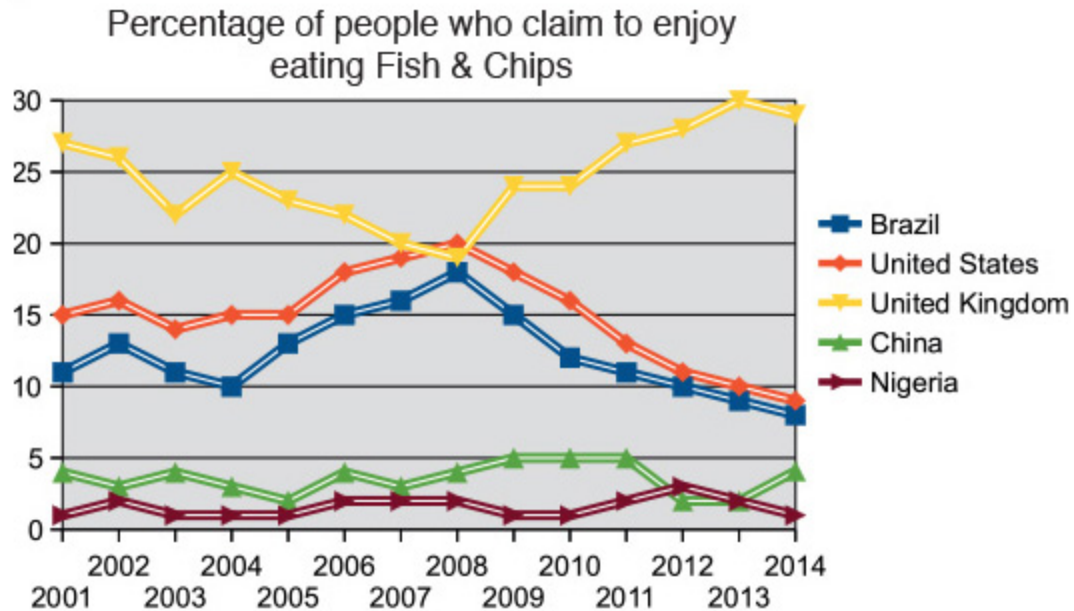


## Example for discussion



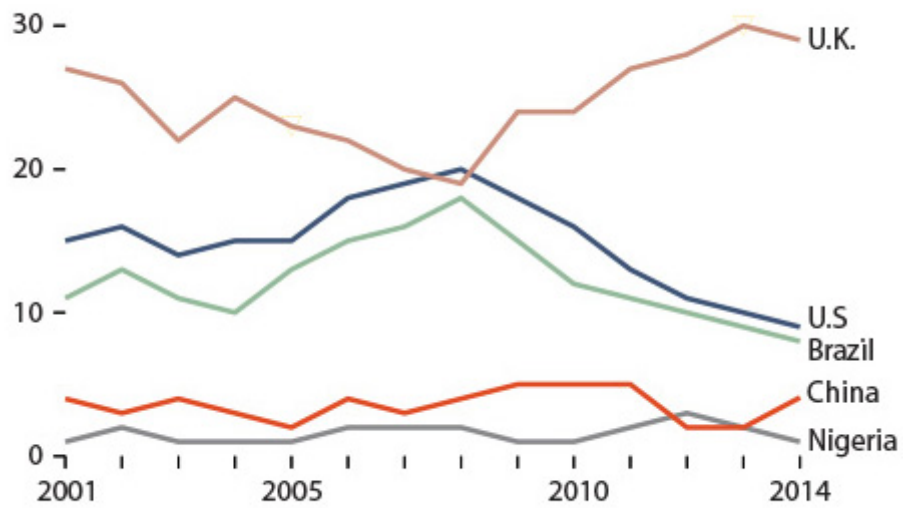
There are three parts in Data7, the first part is from 1 to 70, the second part is from 71 to 220, and the third part is from 221 to 250.

In the first two parts, there are two periodic fluctuating trends, and the frequencies are different. The third part is a sharply descending trend.



Source: CompletelyFakeData, Inc.

### People who claim to enjoy Fish & Chips (%)

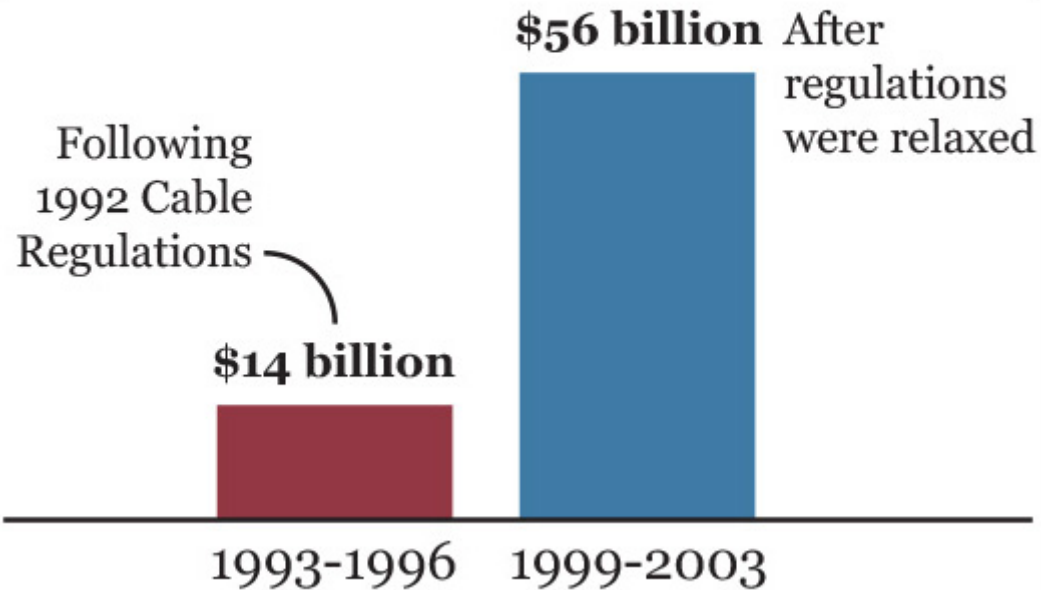


Source: CompletelyFakeData, Inc.

*From "The Truthful Art Data, Charts, and Maps for Communication" by Alberto Cairo*

**Less regulation =  
More industry investment**

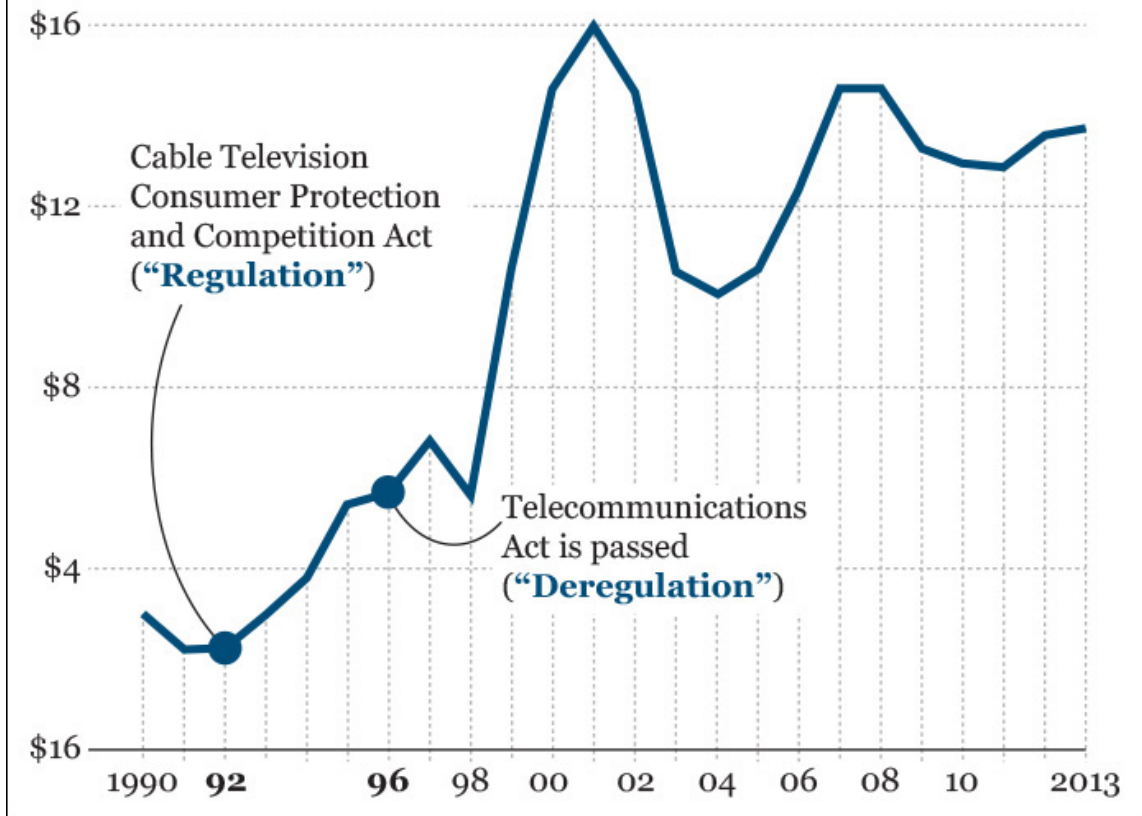
---



Any thing looks suspicious here??

## Cable Industry Infrastructure Expenditures

*In billions*

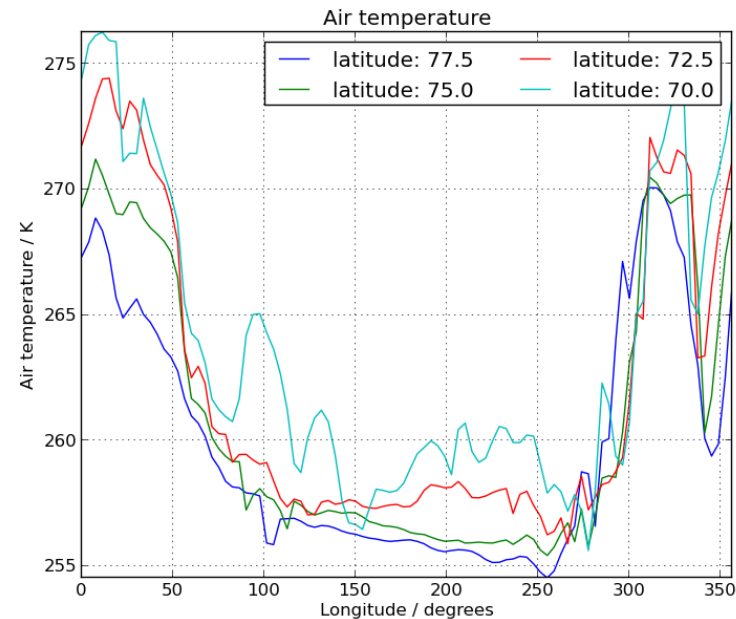
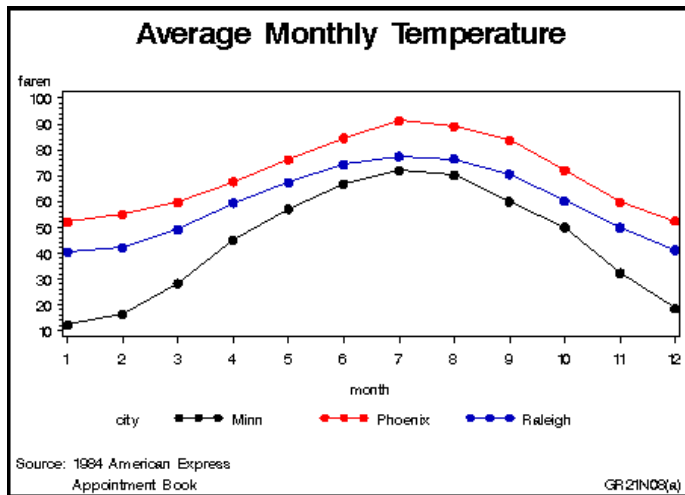


*From "The Truthful Art Data, Charts, and Maps for Communication" by Alberto Cairo*

# **SOME SIMPLE PLOTTING TECHNIQUES**

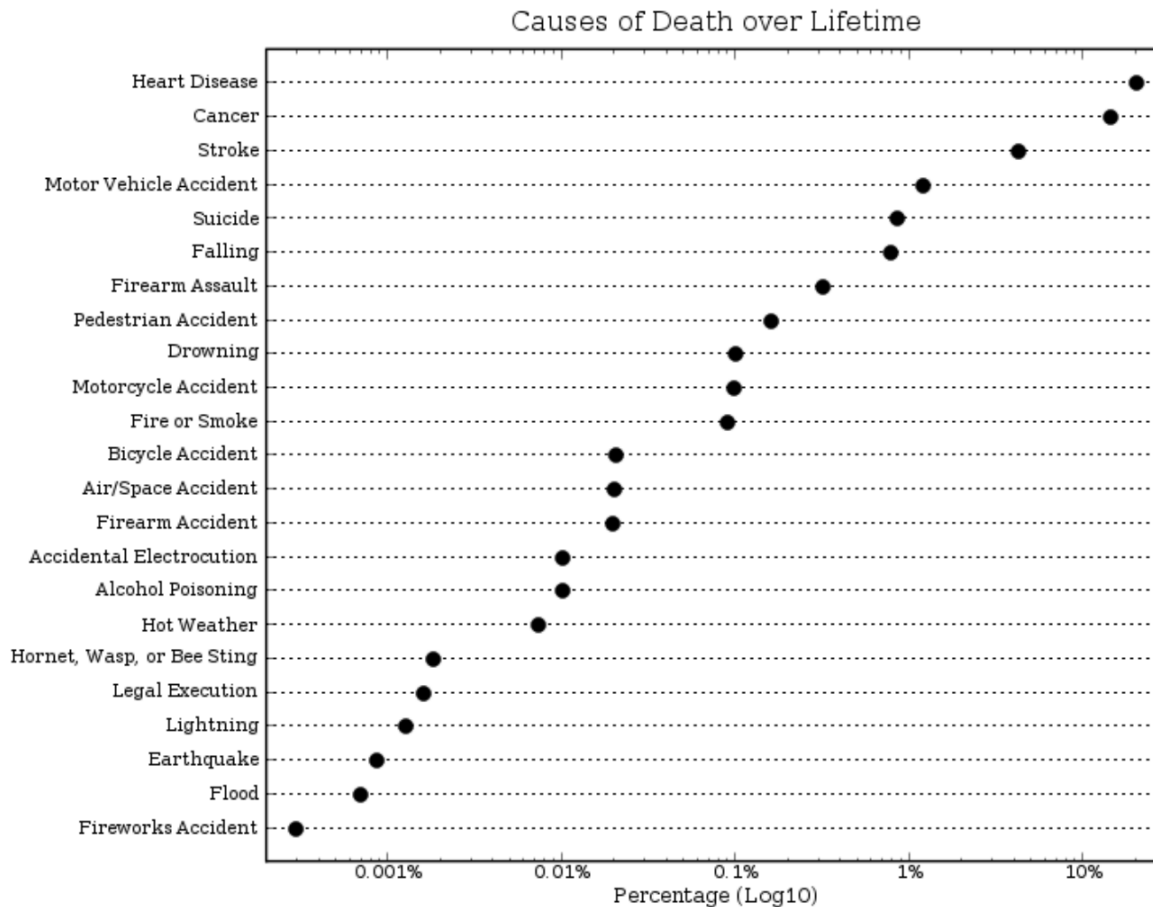
# Connected Symbol Plots (Line Plots)

- The most common plotting technique
- Used to plot time series or other 1D data with intrinsic order



# Dot Plots

- Similar in nature to bar charts or pie charts
- Should be used for quantitative labeled data



The data points do not have sequential relation!!

A dot plot showing the odds of dying.

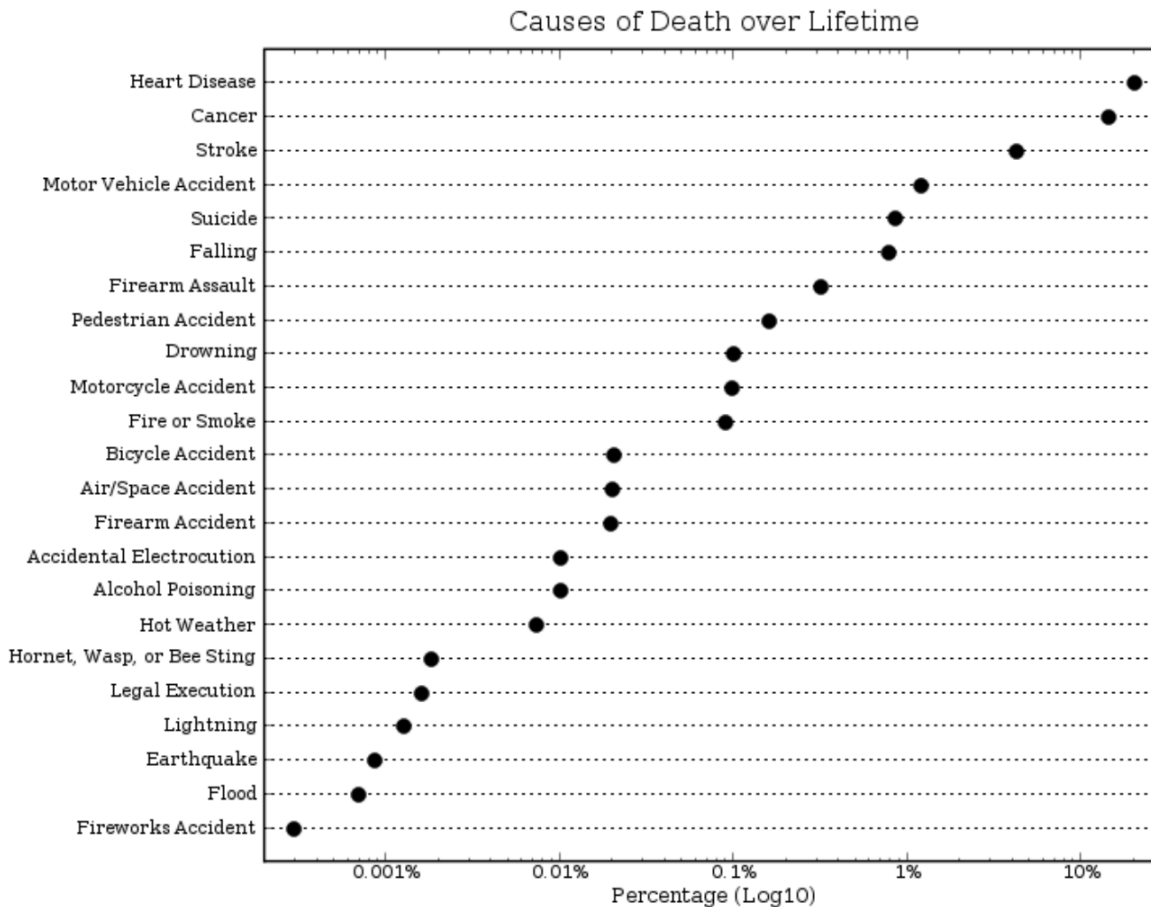


# Dot Plots

The values should normally be sorted such that the largest value is at the top.

- Exception: the data has an inherent order that must be preserved

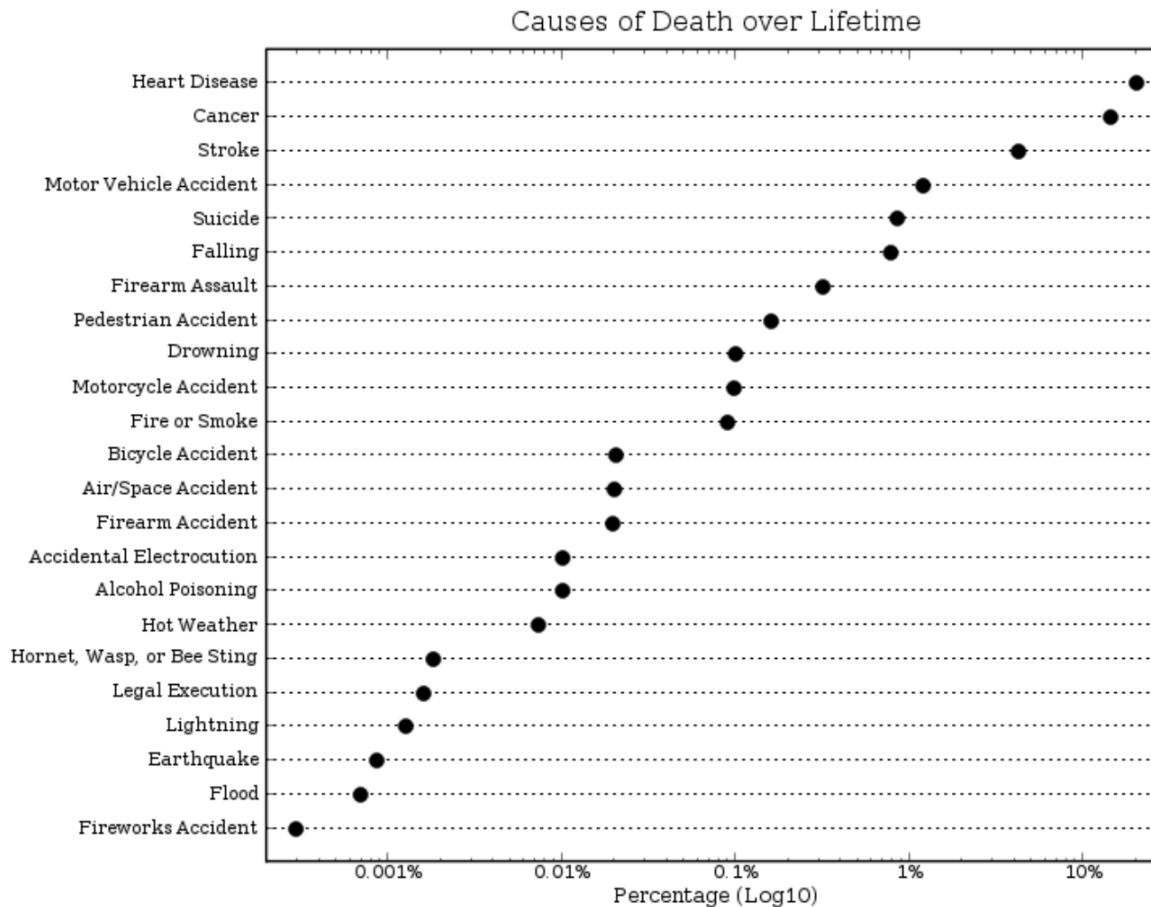
A log scale should be used to reduce skewness in the data



A dot plot showing the odds of dying.

# Dot Plots

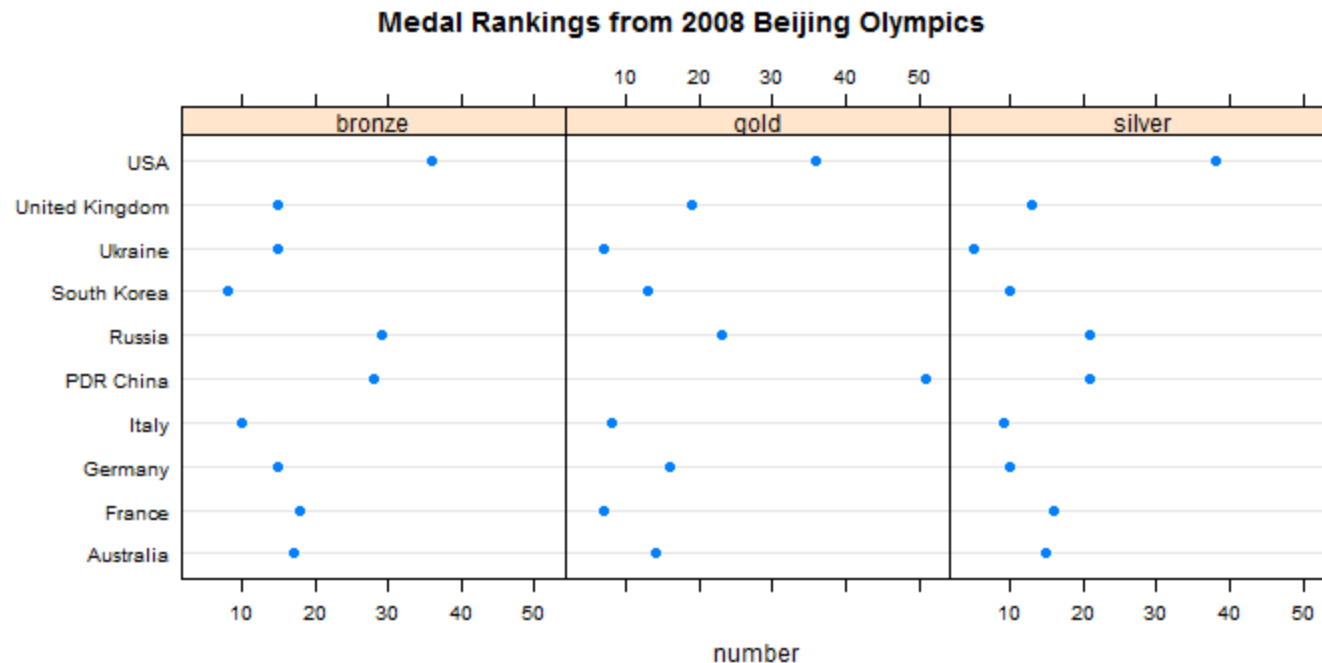
- Real world data is not always univariate.
- To represent multi-dimensional data, a multiway dot plot can be used



A dot plot showing the odds of dying.

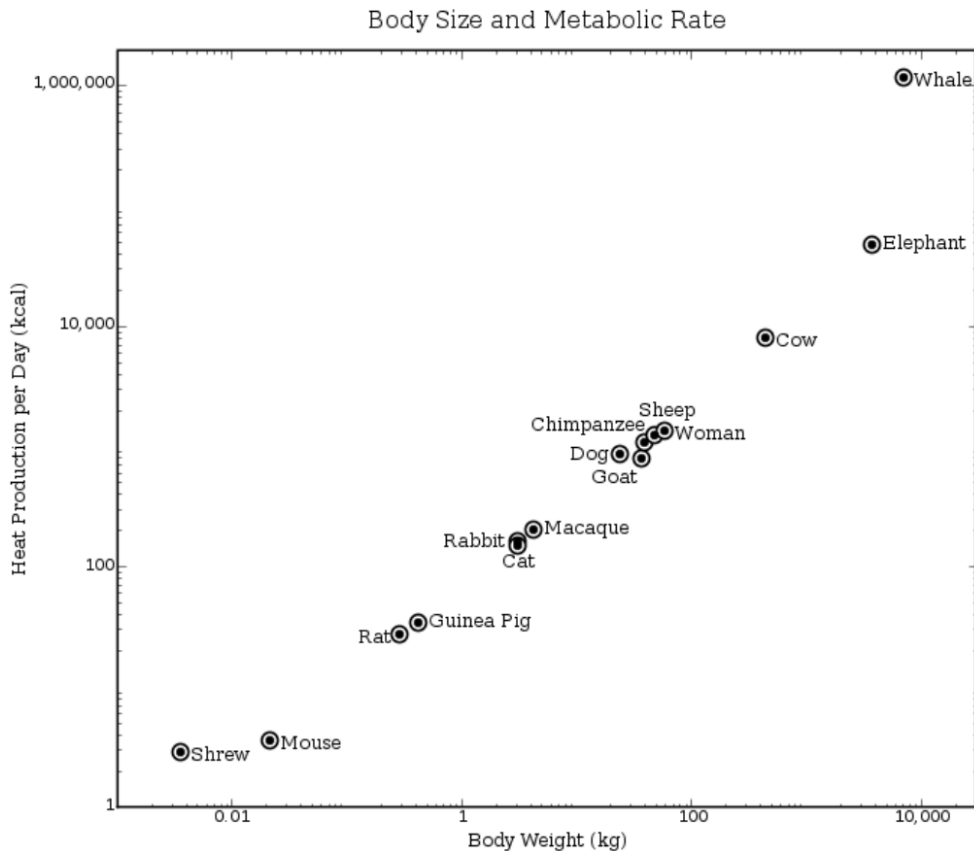
# Dot Plots

- A multiway dot plot is just several dot plots that share common labels and are juxtaposed such that they share an axis.



# Scatter Plots

- Scatter plots are used to show how one variable is affected by another, or correlated, in 2D data.

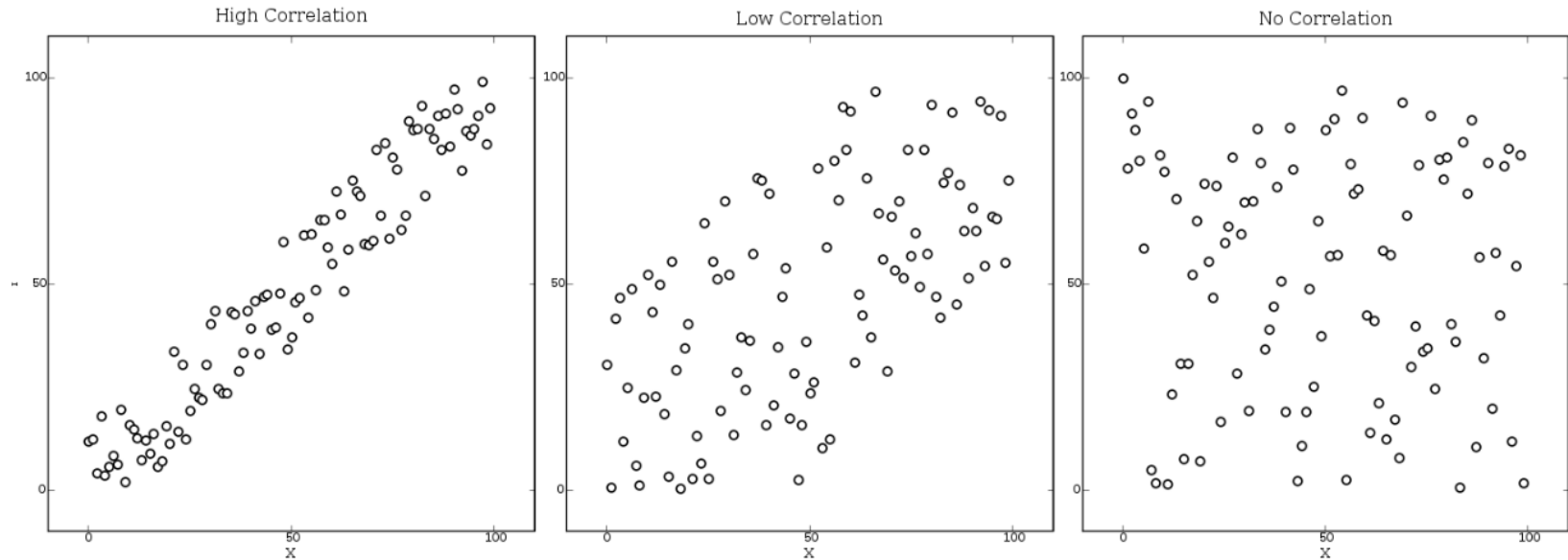


Need to make the symbols in the data stand out and keep the labels from obscuring the data and making the trend difficult to perceive

A scatter plot showing the biological principle of scaling for mammals. For each sample, the metabolic rate is plotted against the body mass to show a high correlation between the two variables. The points have also been labeled to provide additional information.

# Scatter Plots

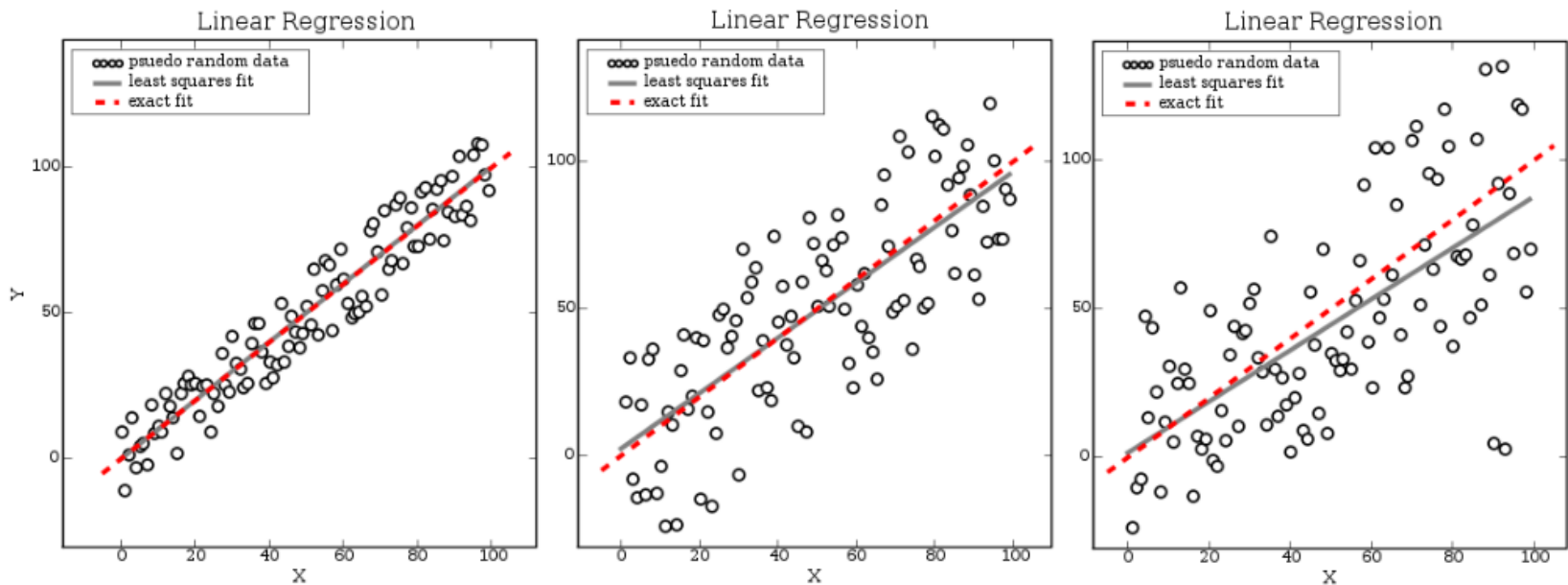
- If used properly, the correlation of the data can easily be discerned.



Scatter plots showing different levels (high, low, and no, respectively) of correlation for points generated with different magnitudes of randomness.

# Scatter Plots

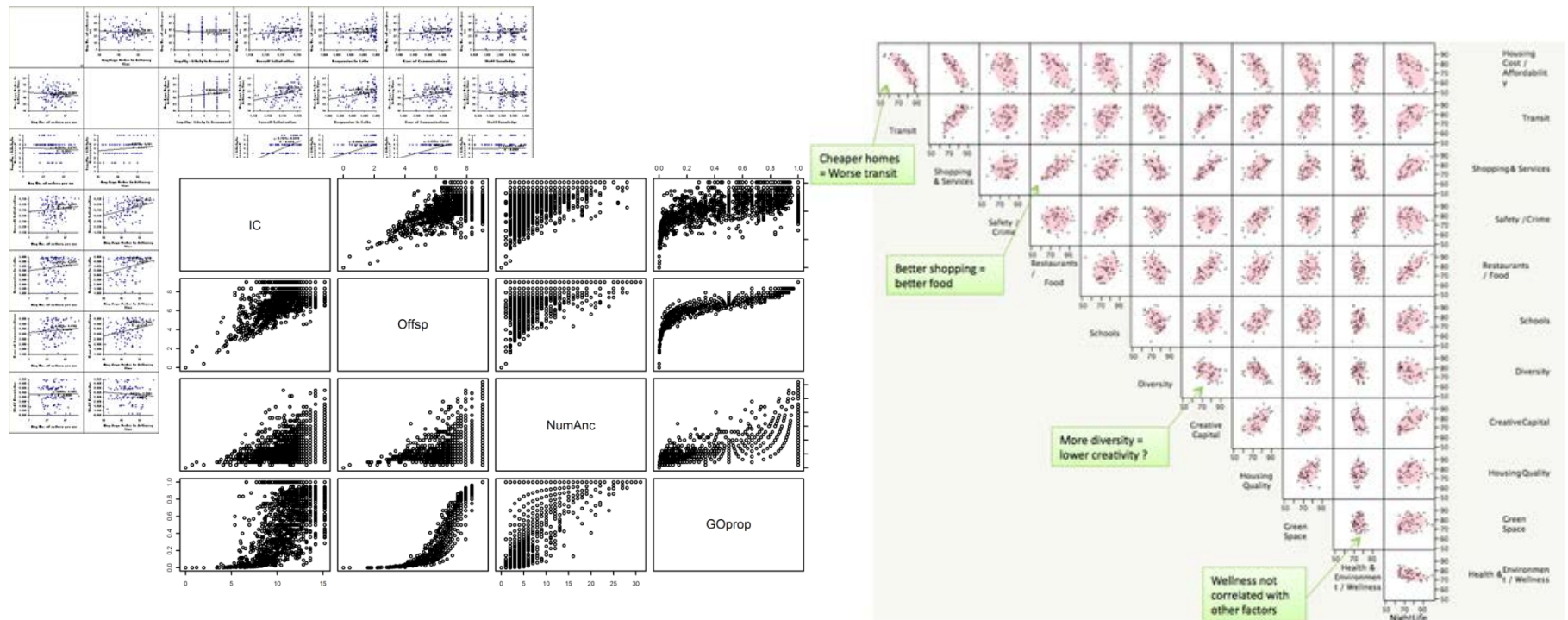
- It is often desirable to express the correlation as a line that provides the best fit for the data.



Linear regression using least squares fits a line to the data. The fit is good for high and low correlation (left and middle), but can result in problems in the case of outliers (right)

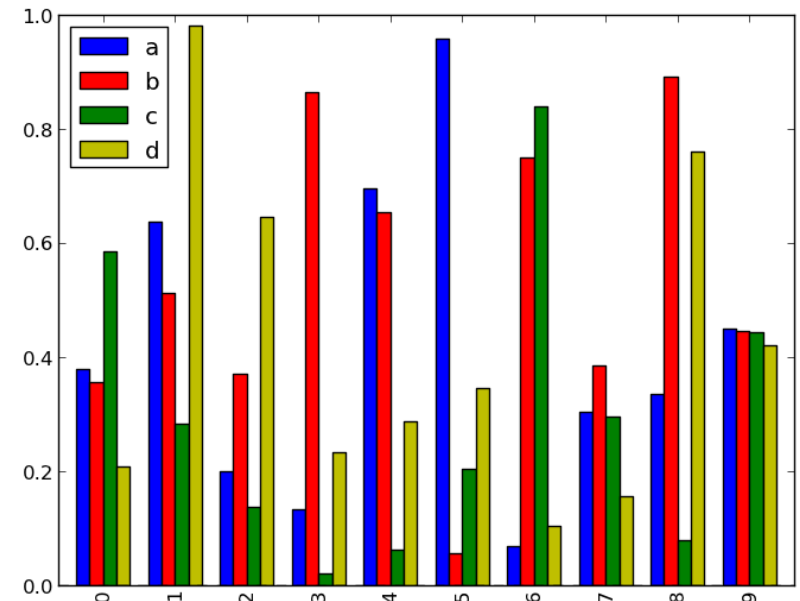
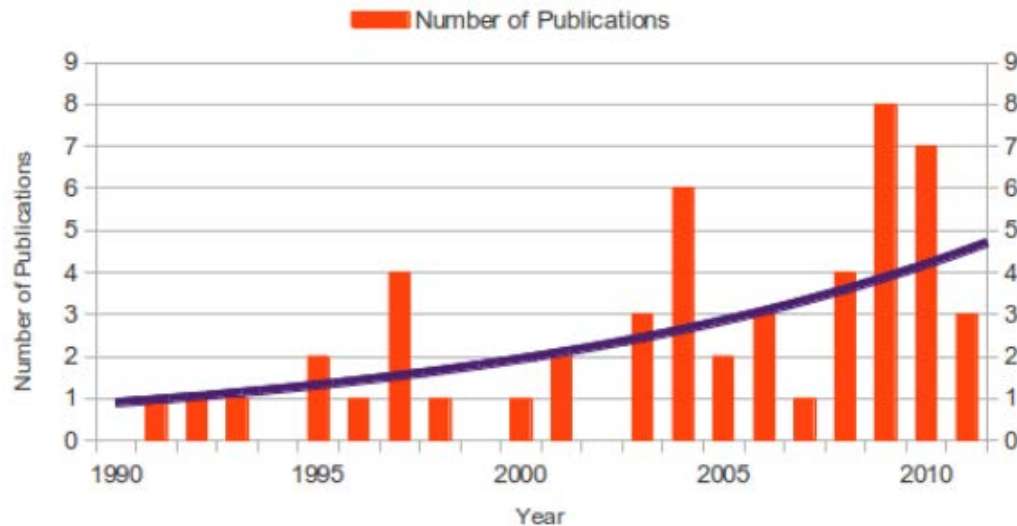
# Scatter Plots

- As with dot plots, scatter plots can be used to represent data in higher dimensions. This is frequently done with a scatter plot matrix.
- This assigns each dimension of the plot to a single row and column in the matrix. The variables are then plotted against each other as a standard scatter plot for each entry in the matrix.



# Histograms

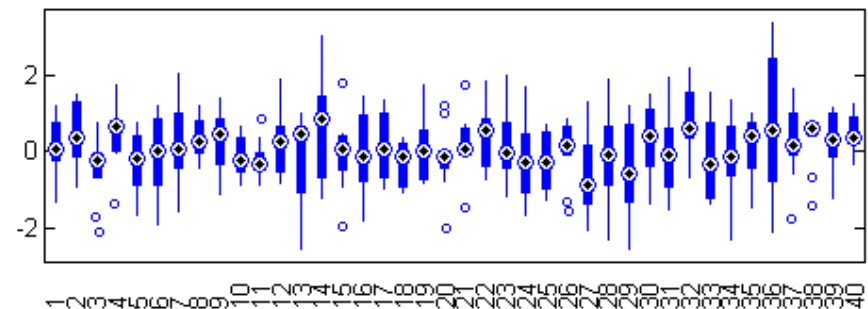
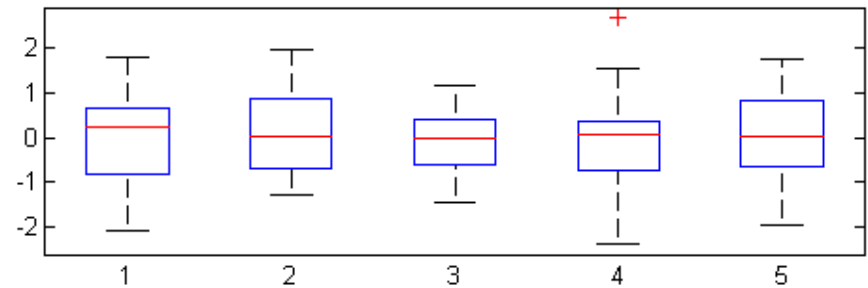
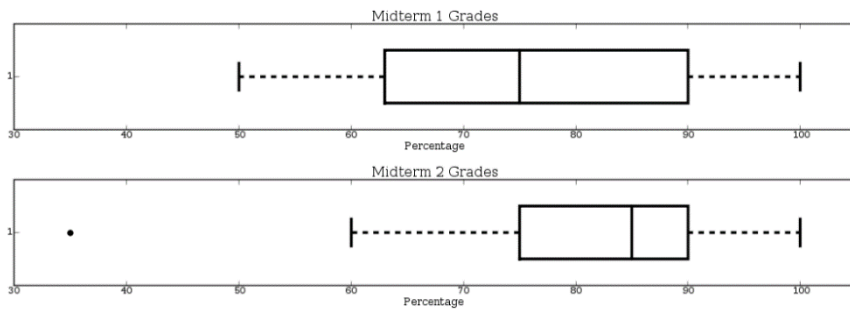
- Histograms are a special type of bar charts used for plotting distributions in data.
- The horizontal axis represents fixed intervals of the data and the vertical axis represents the number of values that lie within the intervals.





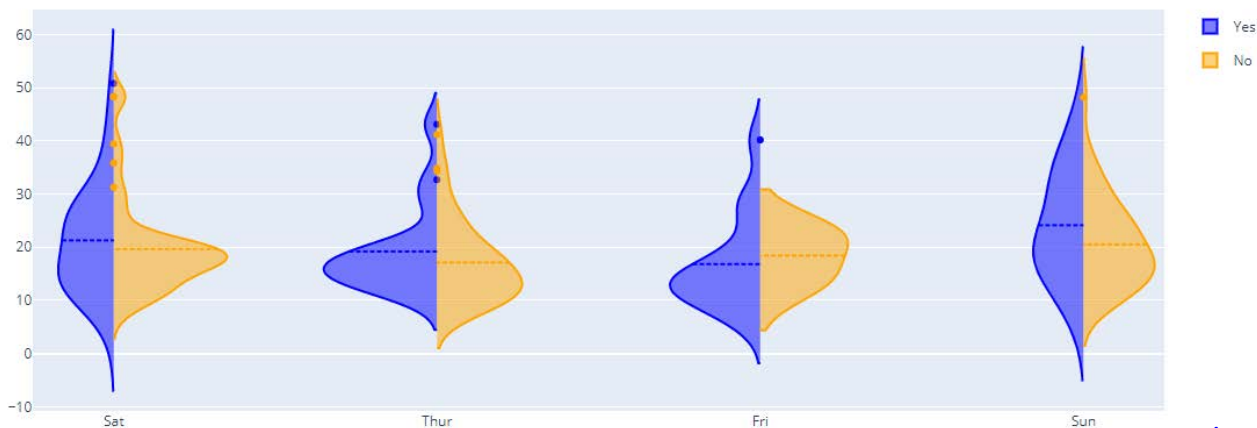
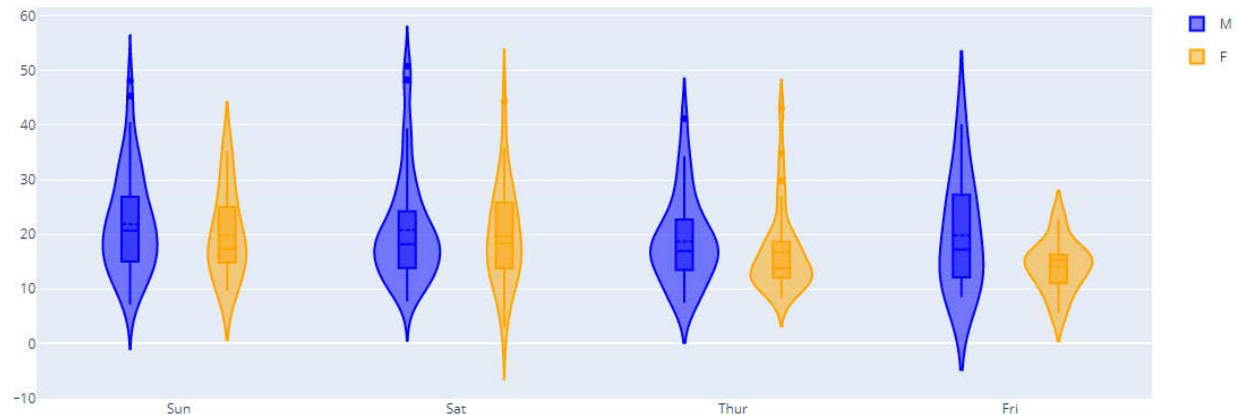
# Box Plots

- Box plots are typically used to represent the statistical variation in the data

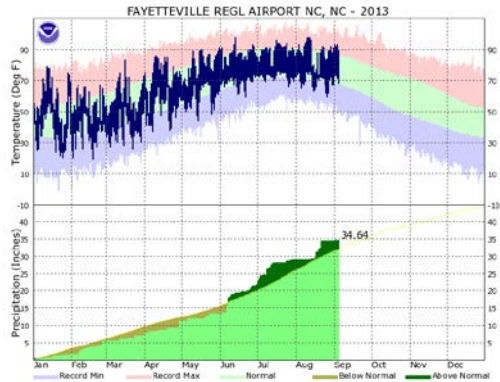
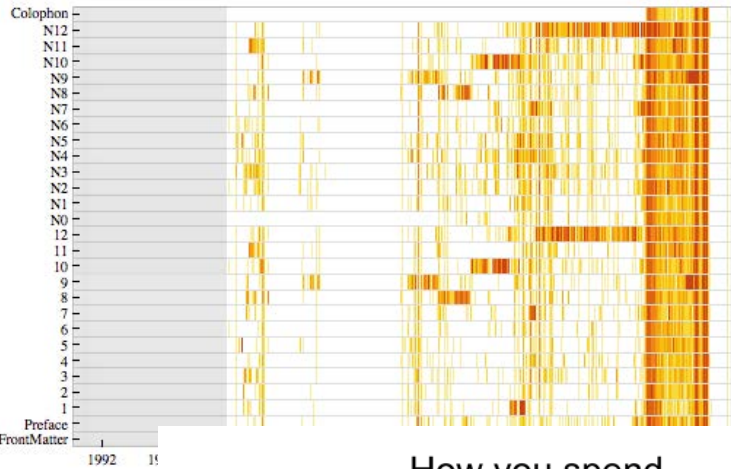


# Violin Plots

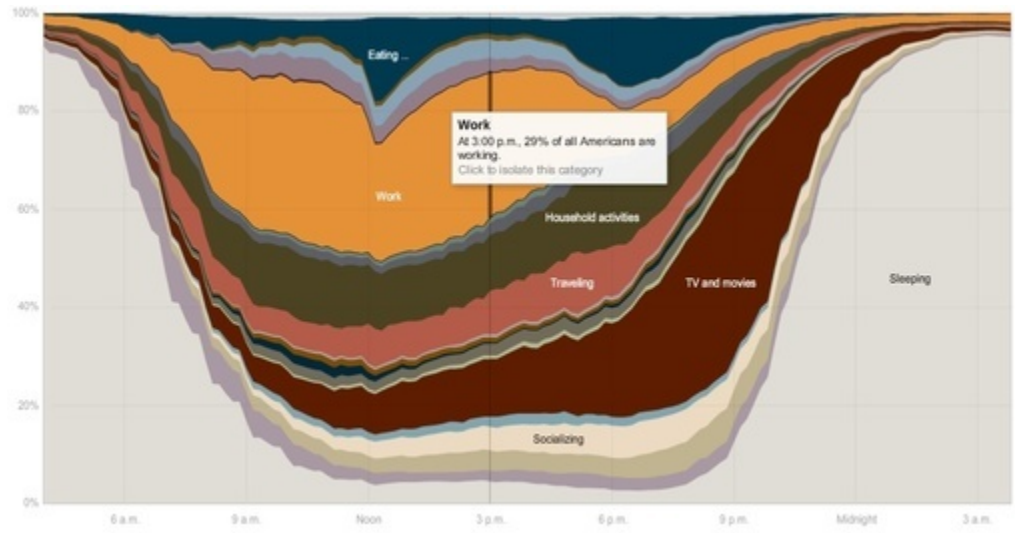
- similar to box plots, except that they also show the probability density of the data at different values



# Others























## How you spend your time:



SOURCE: Bureau of Labor Statistics, American Time Use Survey of Full Time University Students.

## Brief Overviews of Types of Graphs

### 2D Graphs









-  Bar/Column
-  Bar Dev
-  Bar Left Y
-  Bar Right Y
-  Bar Top
-  Bar X
-  Box
-  Detrended Probability
-  Half-Normal Probability
-  Hanging Bar Histograms
-  Histograms
-  Line
-  Pie Charts
-  Probability
-  Probability-Probability
-  Quantile-Quantile
-  Range
-  Scatterplots
-  Sequential/Stacked
-  Voronoi Scatterplot

### 3D XYZ Graphs







-  Contour
-  Deviation
-  Scatterplots
-  Space

-  Spectral
-  Trace






### 3D Sequential Graphs

-  Bivariate Histograms
-  Box
-  Range
-  Raw Data Contour/Discrete
-  Sequential Contour
-  Sequential Surface
-  Raw Data Spikes
-  Raw Data Surface

### 4D/Ternary Graphs

-  Scatterplots
-  3D Ternary
-  Contour/Area
-  Contour/Line
-  3D Deviation
-  3D Space




### 2D Categorized Graphs

-  Detrended Probability
-  Half-Normal Probability
-  Normal Probability
-  Probability-Probability
-  Quantile-Quantile

### 3D Categorized Graphs

-  Contour
-  Deviation
-  Scatterplots
-  Space
-  Spectral
-  Surface




### Ternary Categorized Graphs

-  Ternary Contour/Area
-  Ternary Contour/Line
-  Ternary Scatterplot

### nD/Icon Graphs

-  Chernoff Faces
-  Columns
-  Lines
-  Pies
-  Polygons
-  Profiles
-  Stars
-  Sun Rays

### Matrix Graphs

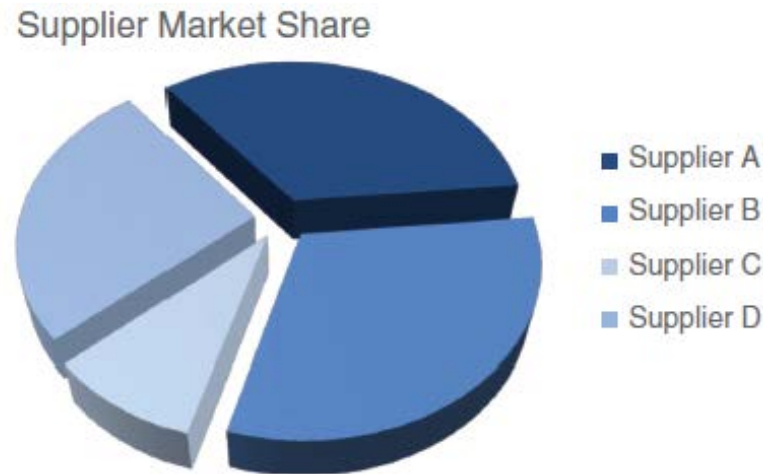
-  Columns
-  Lines
-  Scatterplot

# Recall the Gestalt Principles

- Proximity
- Similarity
- Enclosure
- Closure
- Continuity
- Connection
- Simplicity
- Common fate
- Symmetry
- Past experience

# To avoid

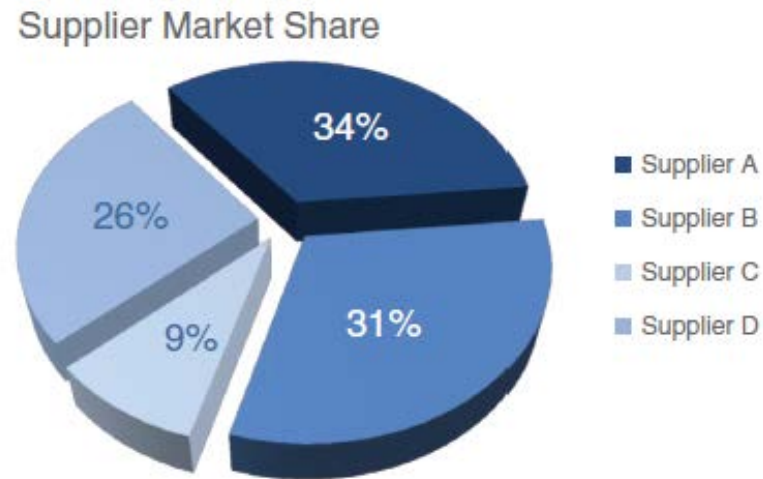
- Try not to use pie charts



Which supplier is the largest based on this visual?

# To avoid

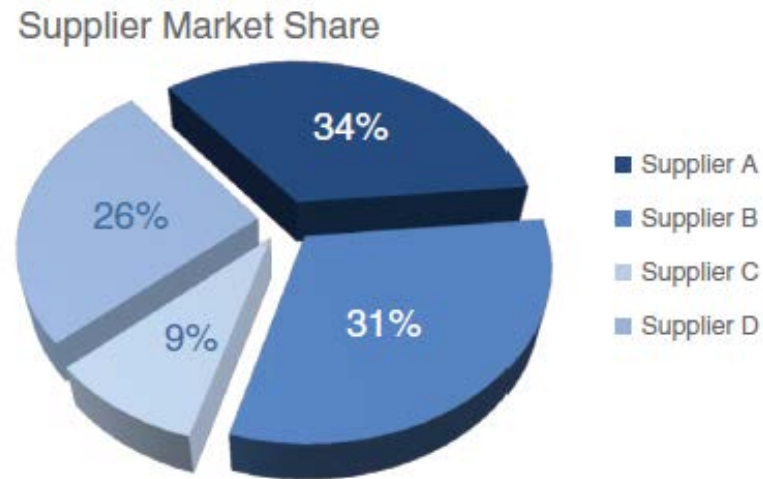
- Try not to use pie charts



Actual percentage!

# To avoid

- Try not to use pie charts



Actual percentage!

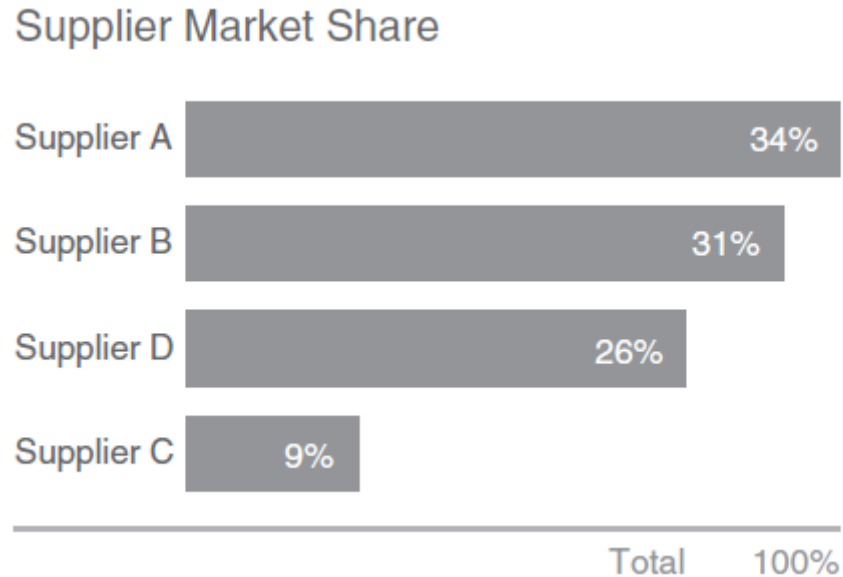
You may blame the 3D representation, but the more important issue is pie charts try to use area and angles to convey the difference in values!



# To avoid

- **Try not to use pie charts**

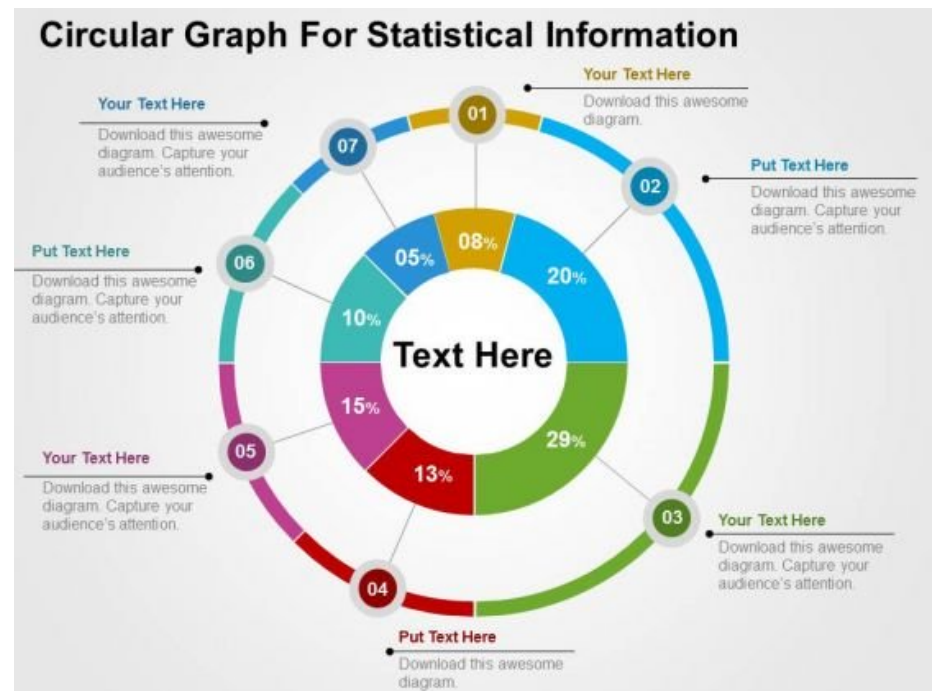
Most pie charts can be replaced by the following form of bar chart



# To avoid

- Try not to use pie charts

Alternatively, you can use the following circular graphic/diagram

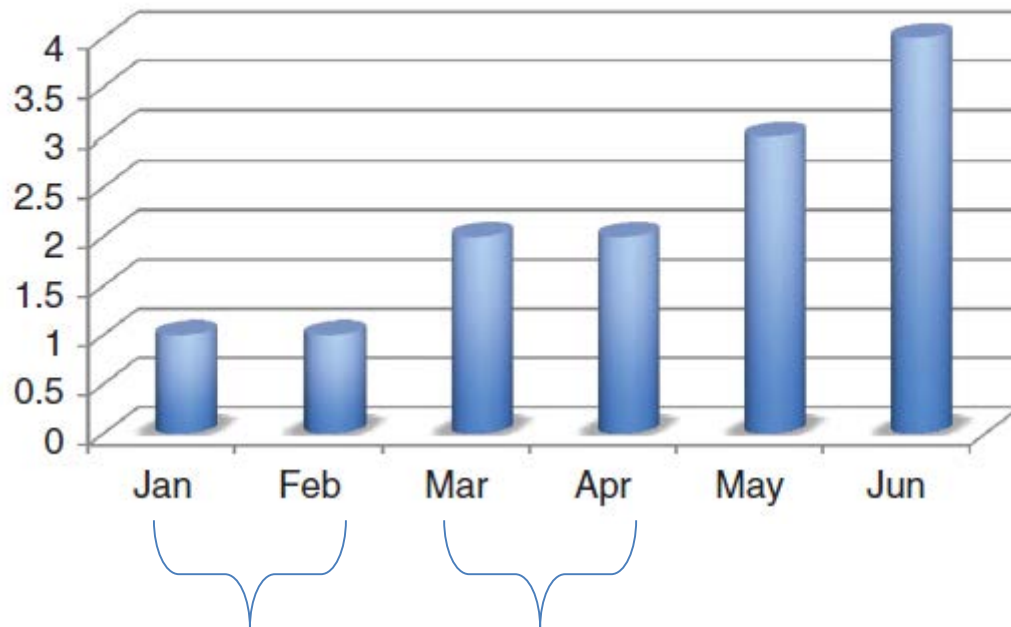


*From Google images*

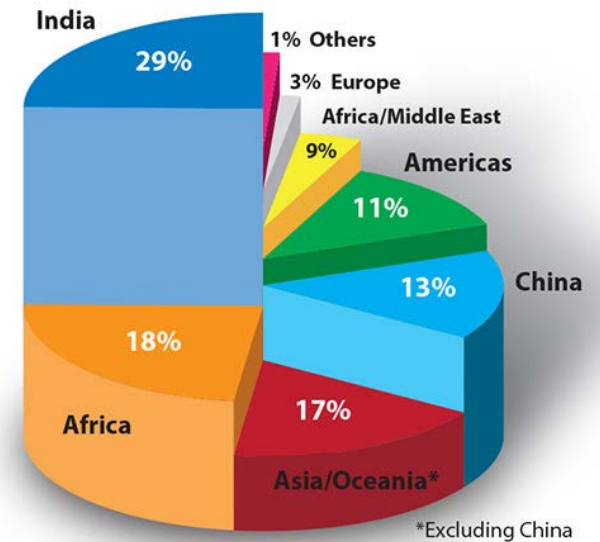
# To avoid

- **Never use 3D in your plots**

Number of issues



Share of worldwide urban population growth 2010-2050

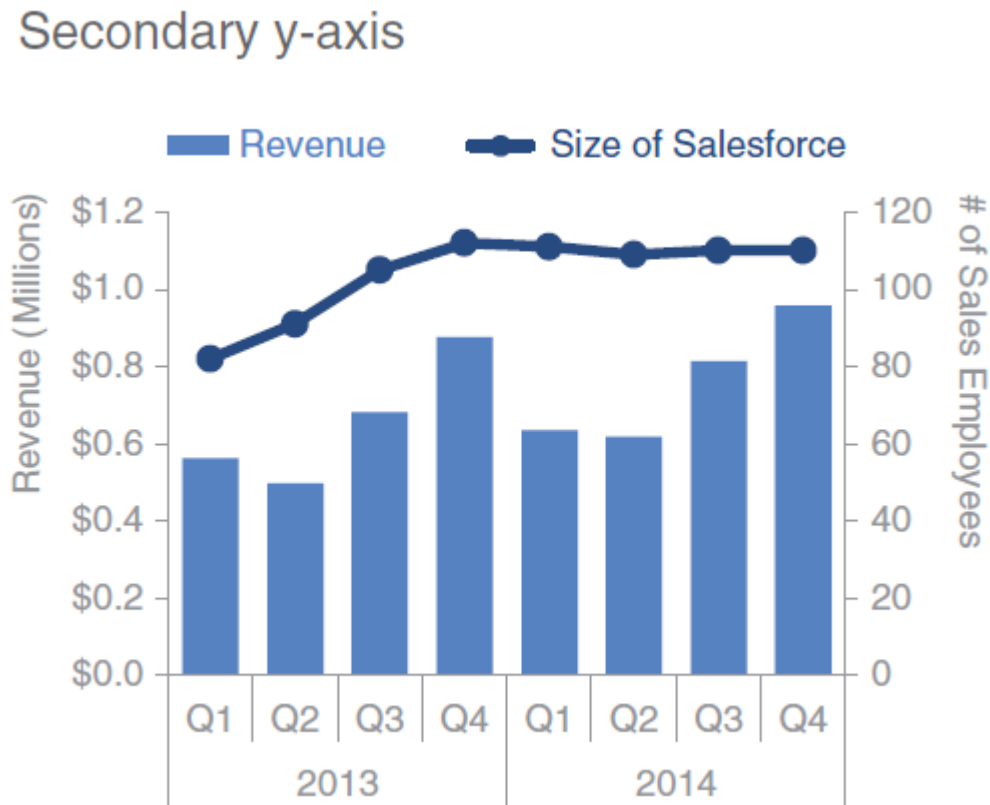


*From "Storytelling with data"*

*From "The Truthful Art Data, Charts, and Maps for Communication" by Alberto Cairo*

# To avoid

- **Secondary y-axis: generally not a good idea**



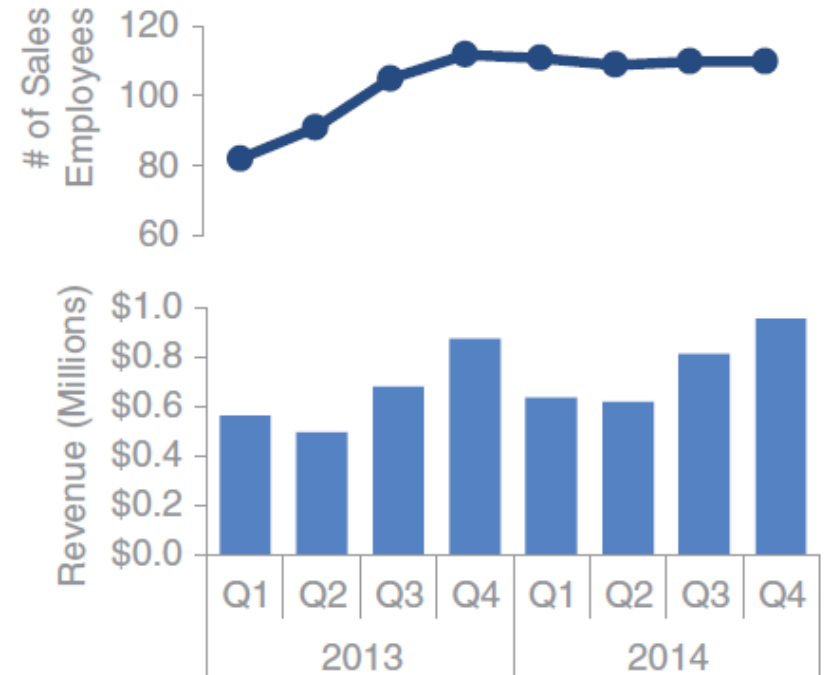
# To avoid

- **Secondary y-axis: possible alternatives**

Alternative 1: label directly



Alternative 2: pull apart vertically



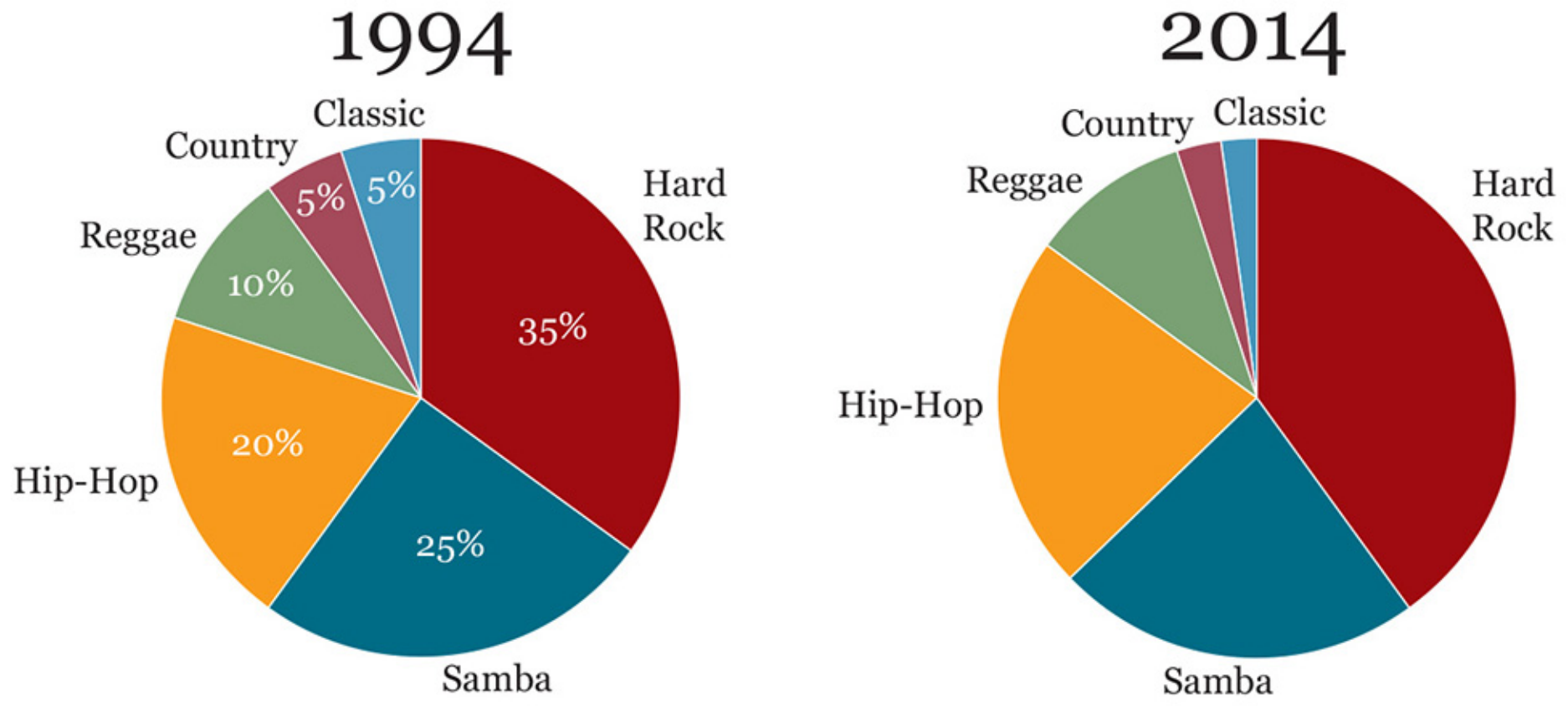
**How to use what we learned to improve our graphical representation?**

# How to make this visualization more effective?

## How Music Preferences Have Changed in Two Decades

Music styles preferred by University of Miami students. Survey based on interviews with 1,000 students.

**SOURCE:** WishfulThinkingData Inc.



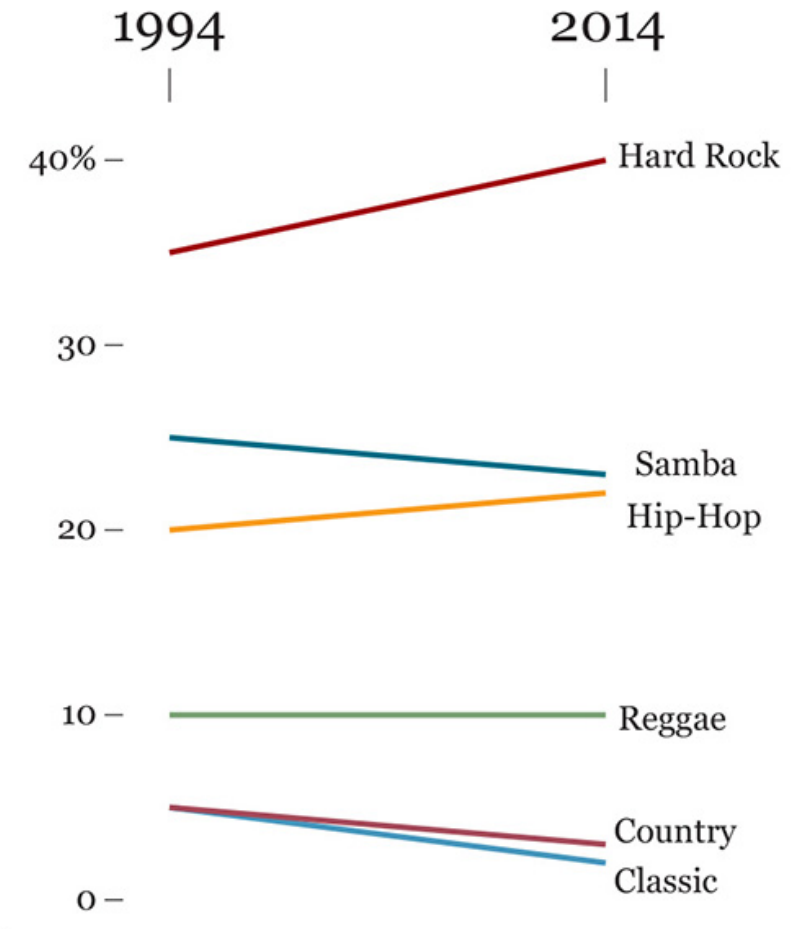
*From "The Truthful Art Data, Charts, and Maps for Communication" by Alberto Cairo*

# How about now?

## How Music Preferences Have Changed in Two Decades

Music styles preferred by University of Miami students. Survey based on interviews with 1,000 students.

**SOURCE: WishfulThinkingData Inc.**

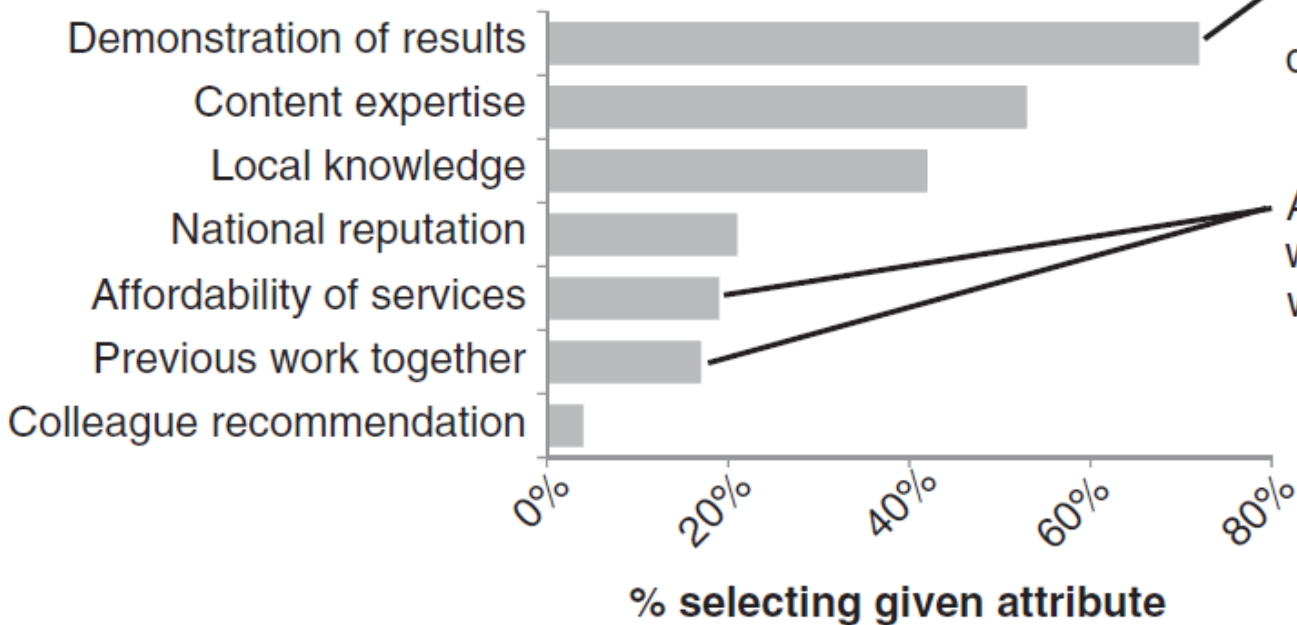


*From "The Truthful Art Data, Charts, and Maps for Communication" by Alberto Cairo*



# Demonstrating effectiveness is most important consideration when selecting a provider

In general, what attributes are the most important to you in selecting a service provider?  
(Choose up to 3)



Survey shows that demonstration of results is the single most important dimension when choosing a service provider.

Affordability and experience working together previously, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents. Note that respondents were able to choose up to 3 options.

## What are the issues?

*From "Storytelling with data"*

# An improved version

**Demonstrating effectiveness** is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

*(Choose up to 3)*



Survey shows that **demonstration of results** is the single most important dimension when choosing a service provider.

**Affordability** and **experience working together previously**, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents.  
Note that respondents were able to choose up to 3 options.

**Demonstrating effectiveness** is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

*(Choose up to 3)*



Survey shows that **demonstration of results** is the single most important dimension when choosing a service provider.

**Affordability** and **experience working together previously**, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents. Note that respondents were able to choose up to 3 options.

## Alignment

Use left-justified alignment rather than central alignment

**In many cases:** Without other visual cues, your audience will typically start at the top left of the page or screen and will move their eyes in a “z” shape (or multiple “z” shapes, depending on the layout) across the page or screen as they take in information.

Therefore, upper-left-most justifying/aligning the text (title, axis titles, legend) enables the audience to hit the details on how to read the data before reading the data.

## Demonstrating effectiveness is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

(Choose up to 3)



Survey shows that **demonstration of results** is the single most important dimension when choosing a service provider.

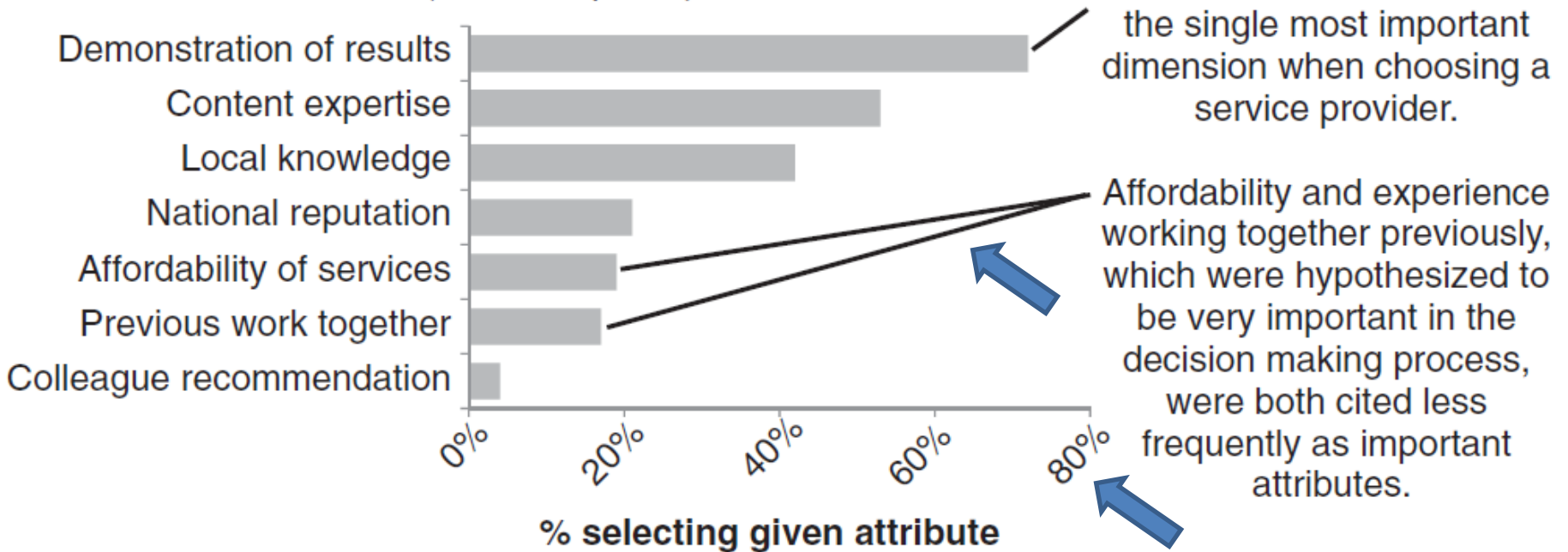
**Affordability** and **experience working together previously**, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents.  
Note that respondents were able to choose up to 3 options.

**Eliminate diagonal elements as much as possible.**

# Demonstrating effectiveness is most important consideration when selecting a provider

In general, what attributes are the most important to you in selecting a service provider?  
(Choose up to 3)



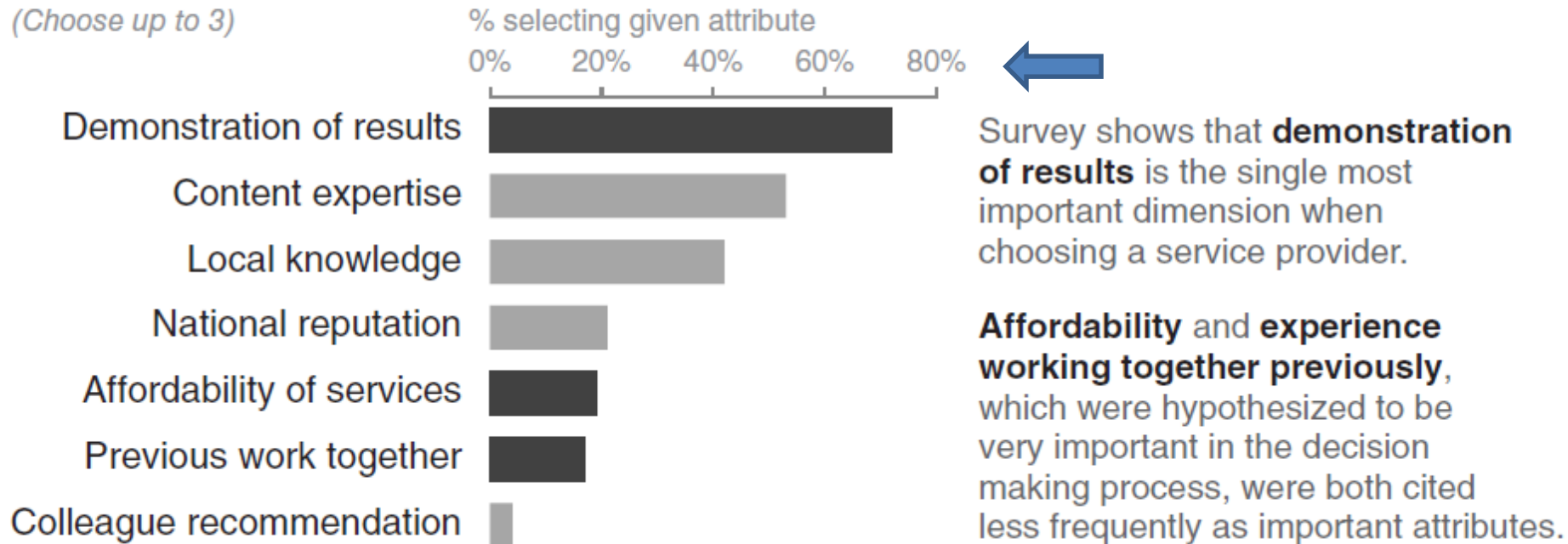
Data source: xyz; includes N number of survey respondents. Note that respondents were able to choose up to 3 options.

**Eliminate diagonal elements as much as possible.**

## Demonstrating effectiveness is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

(Choose up to 3)



Data source: xyz; includes N number of survey respondents.  
Note that respondents were able to choose up to 3 options.

**Eliminate diagonal elements as much as possible.**



## Demonstrating effectiveness is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

(Choose up to 3)



Survey shows that **demonstration of results** is the single most important dimension when choosing a service provider.

**Affordability** and **experience working together previously**, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents.  
Note that respondents were able to choose up to 3 options.

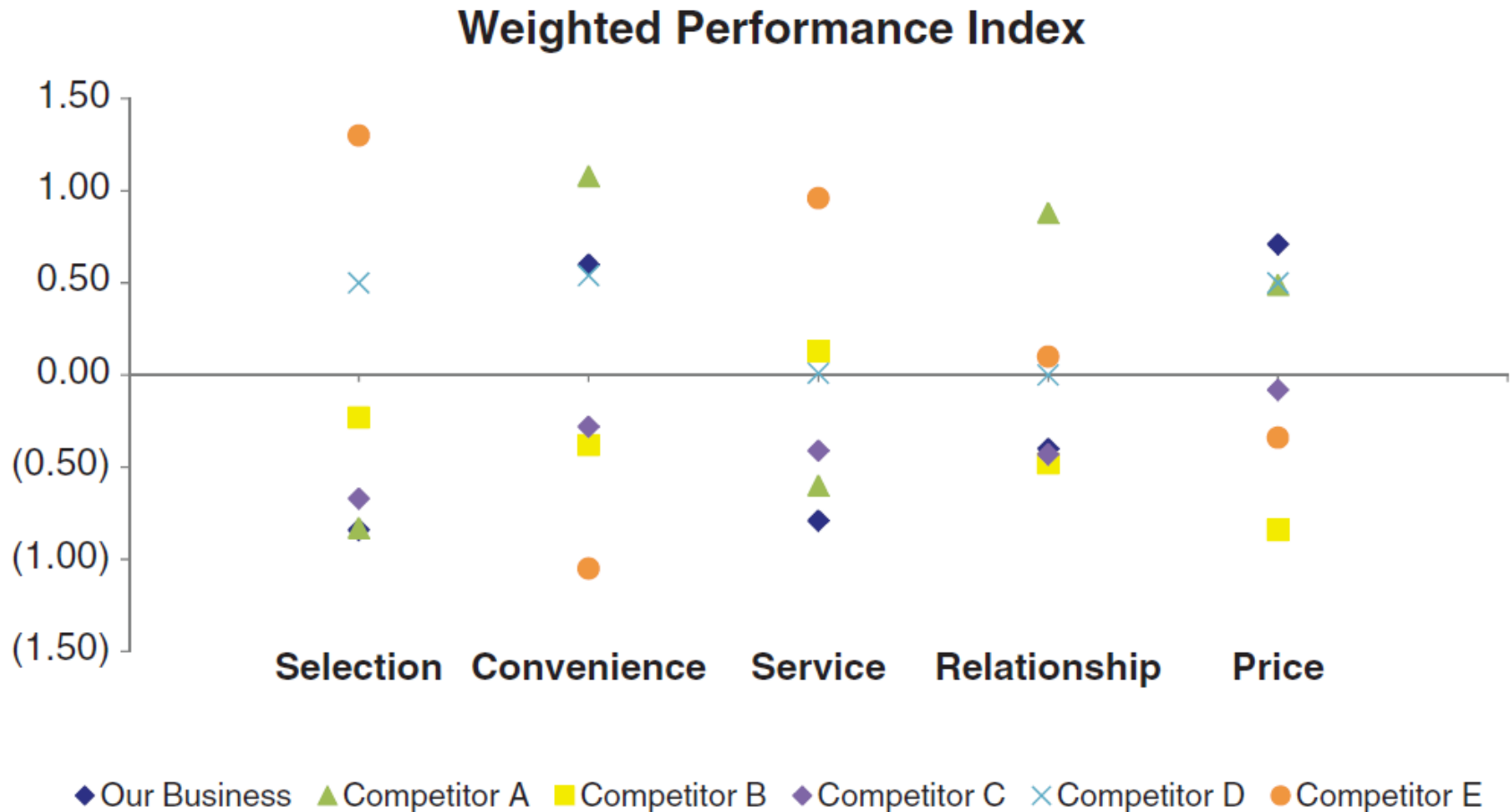
## Importance of **White Space**!!



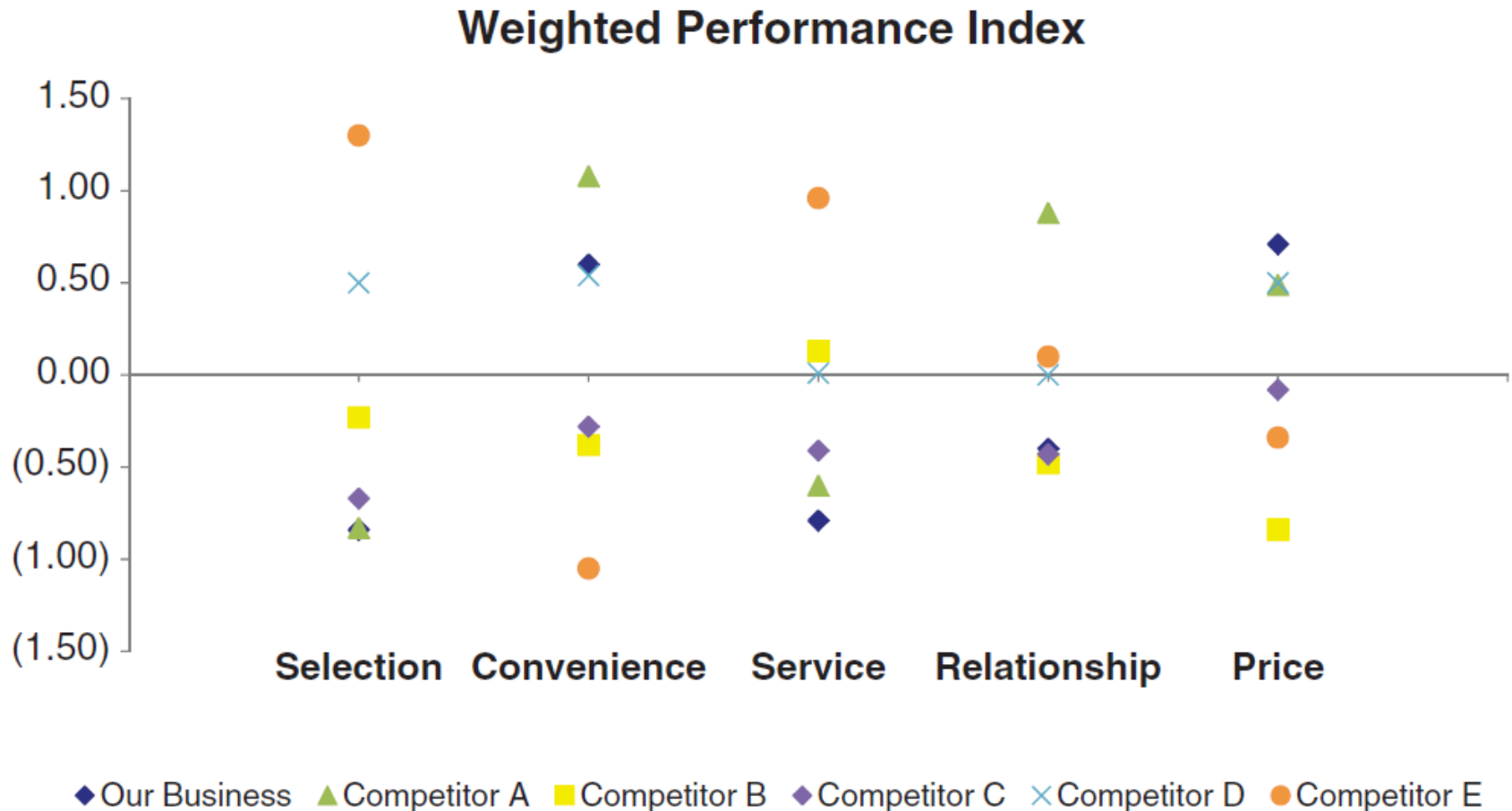
White space in visual communication is as important as *pauses* in public speaking.

*There is a speaker up in front of you and possibly due to nerves or perhaps because they're trying to get through more material than they should in the allotted time they are speaking a mile a minute and you're wondering how they're even able to breathe you'd like to ask a question but the speaker has already moved on to the next topic and still hasn't paused long enough for you to be able to raise your question*

# Another example



# Problem: Non-strategy use of contrast



*From "Storytelling with data"*

# An improved version

## Performance overview

### ■ Our business

- Competitor A
- Competitor B
- Competitor C
- Competitor D
- Competitor E



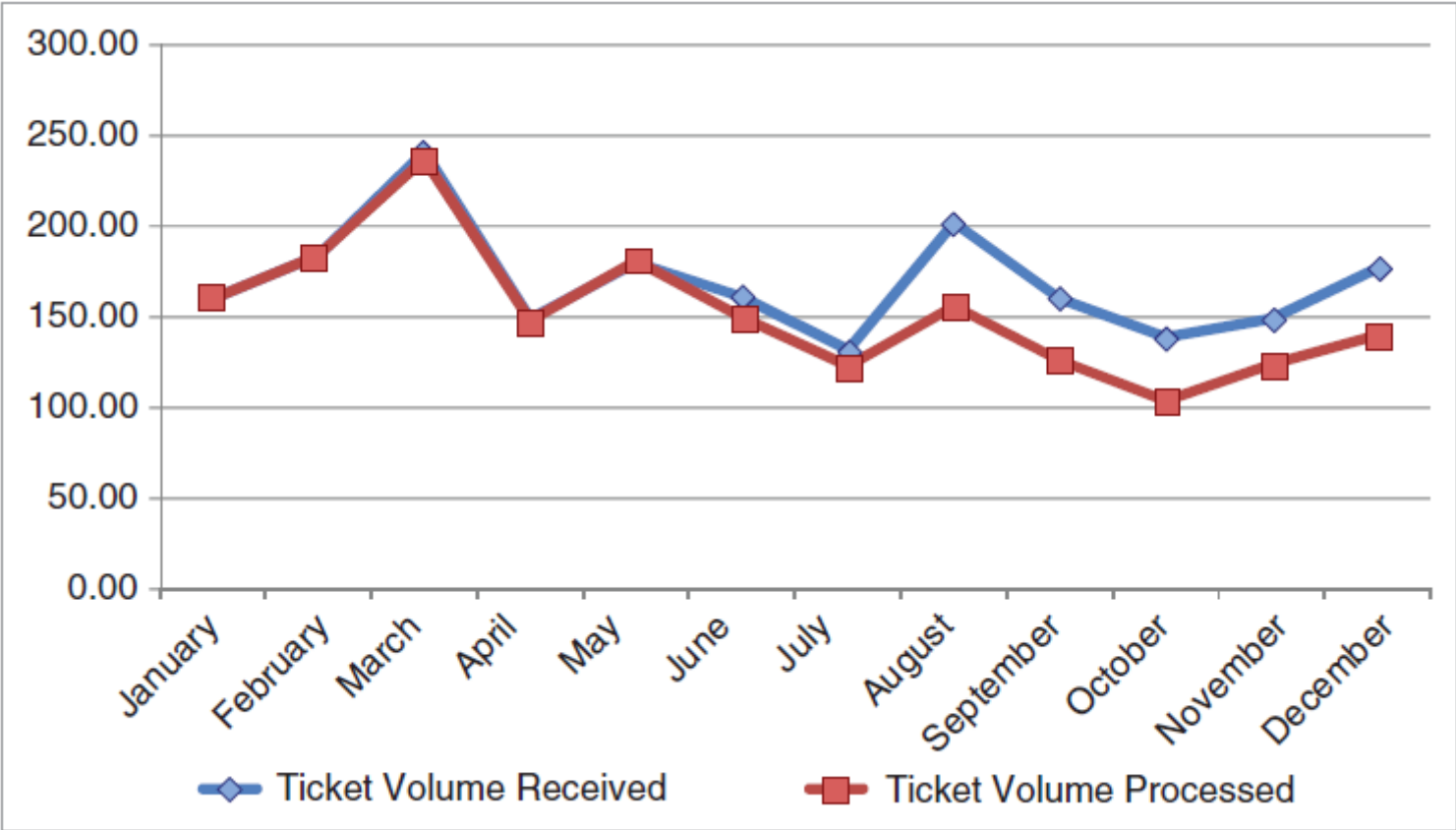
*From "Storytelling with data"*

# Another example

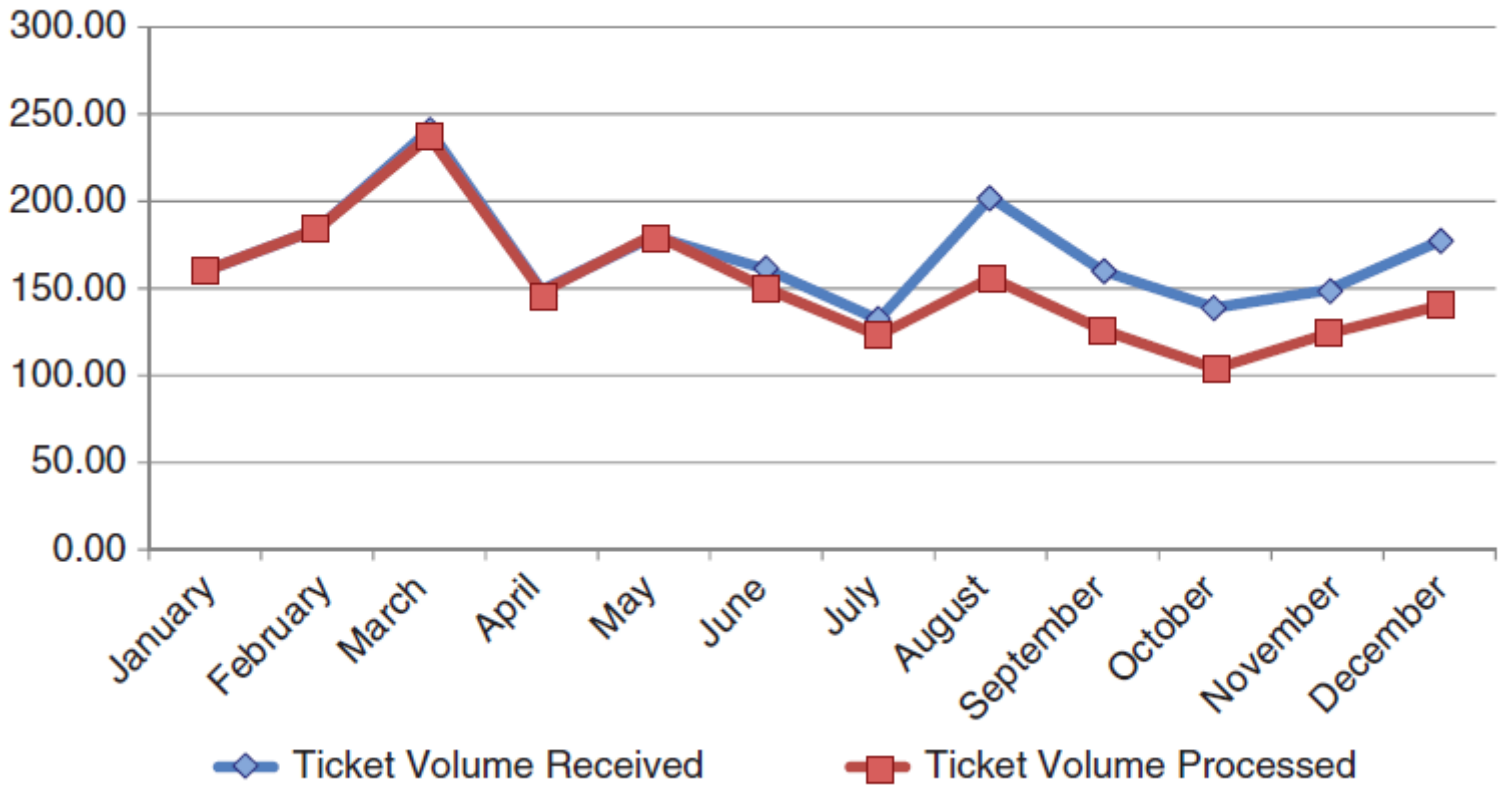
**Scenario:** Imagine that you manage an information technology (IT) team. Your team receives tickets, or technical issues, from employees. In the past year, you've had a couple of people leave and decided at the time not to replace them. You have heard a rumbling of complaints from the remaining employees about having to "pick up the slack." You've just been asked about your hiring needs for the coming year and are wondering if you should hire a couple more people.

You plot the monthly trend of incoming tickets and those processed over the past calendar year. You see that there is some evidence your team's productivity is suffering from being short-staffed and now want to turn the quick-and-dirty visual you created into the basis for your hiring request.

# Original plot

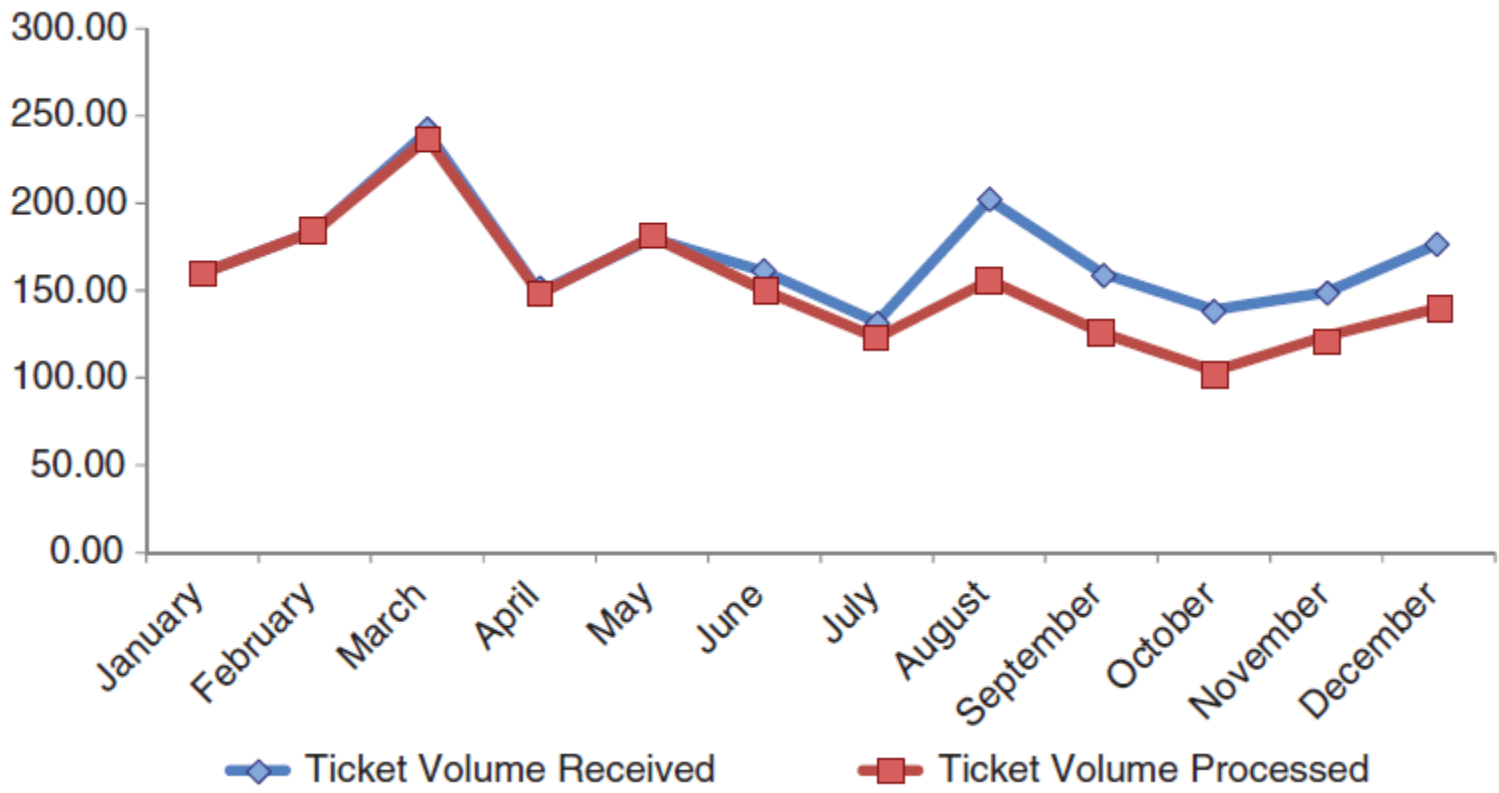


# Remove chart border

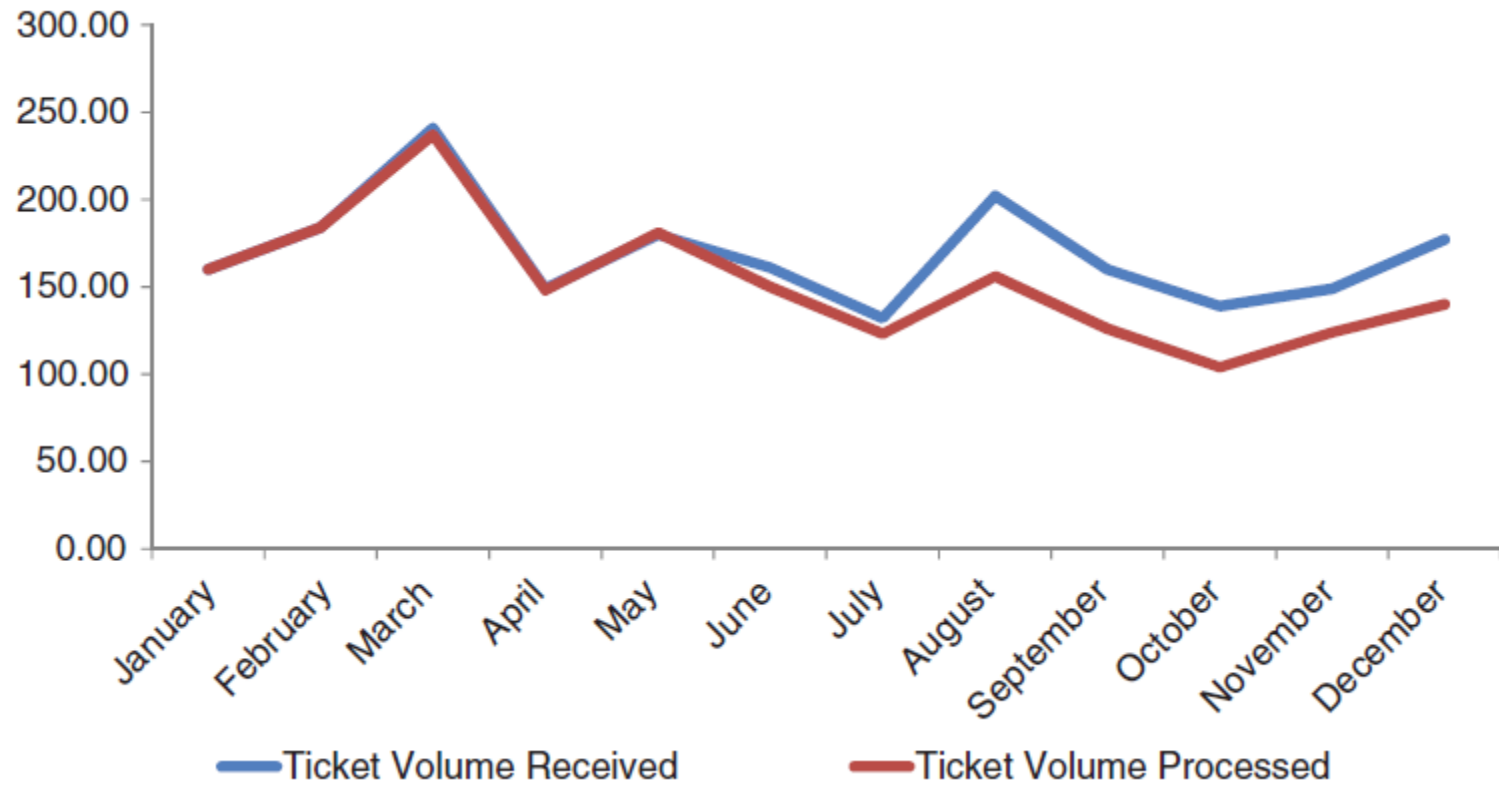




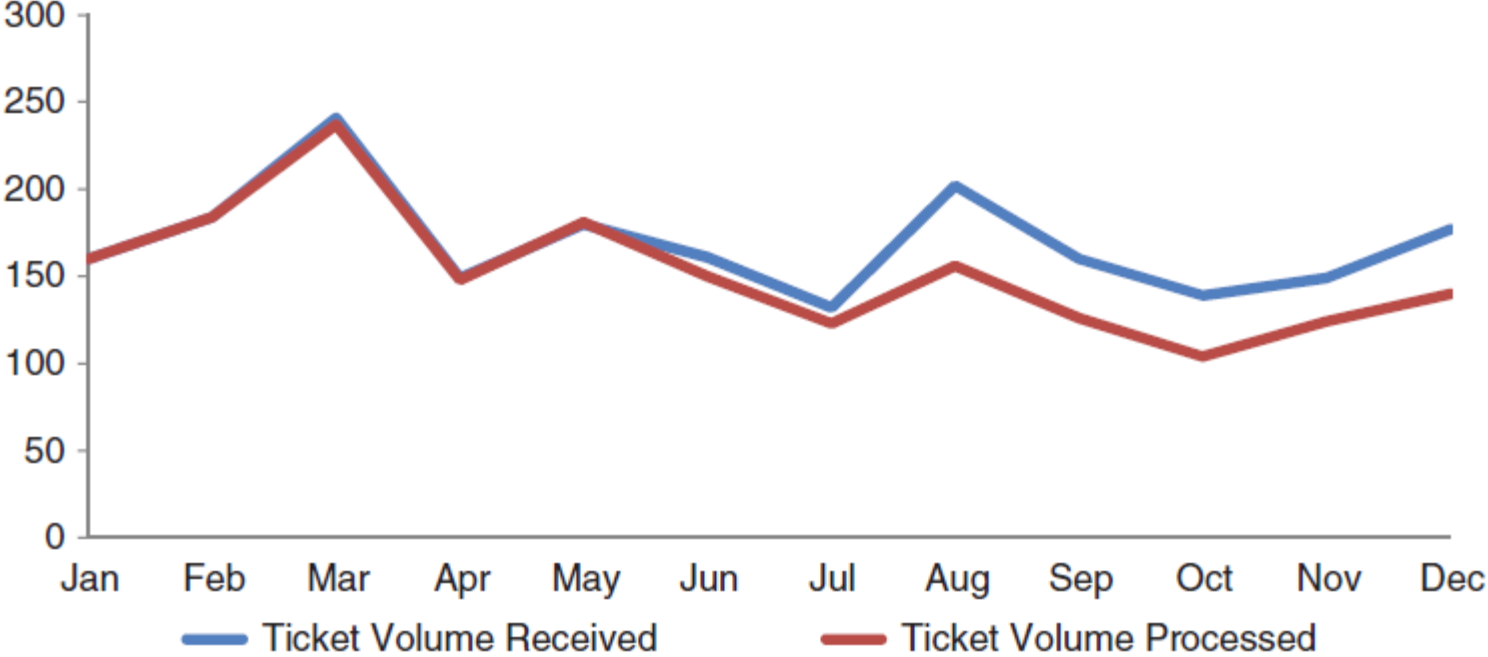
# Remove gridlines



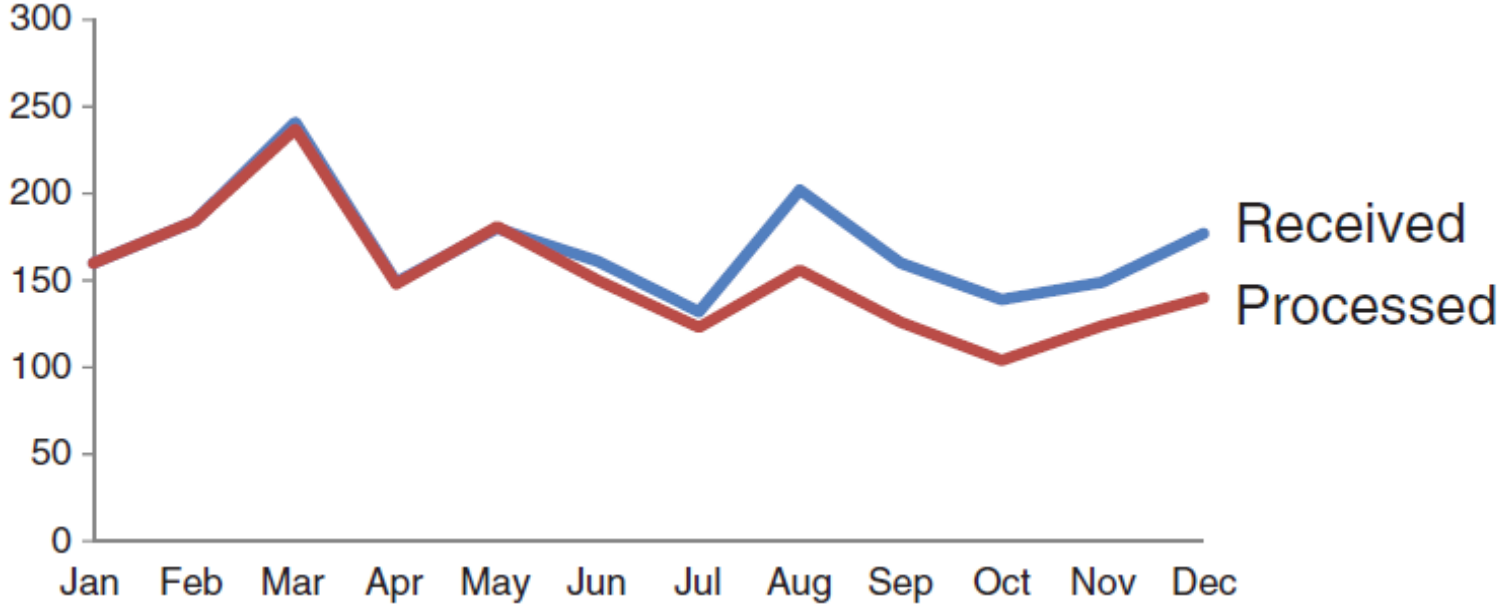
## Remove data markers



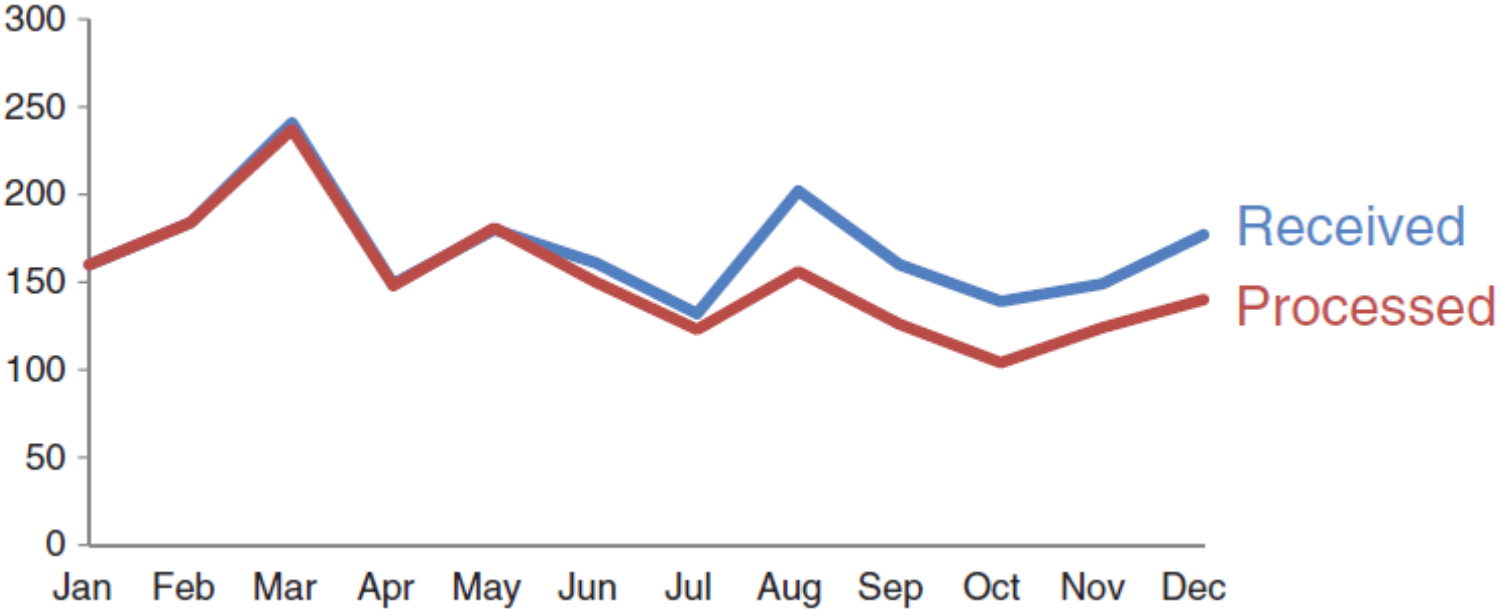
# Clean up axis labels



# Label the data directly



# Leverage consistent color



Note that the above de-cluttering process need not be suitable for other situations!

# Remove

to improve  
(the **data-ink** ratio)

Created by Darkhorse Analytics

[www.darkhorseanalytics.com](http://www.darkhorseanalytics.com)

<https://stat545.com/effective-graphs.html>

# Additional Reading

- Tufte's design principles
  - <http://classes.engr.oregonstate.edu/eecs/spring2015/cs419-001/Slides/tufteDesign.pdf>
- Bad graphs
  - <http://people.math.sfu.ca/~cschwarz/Stat-301/Handouts/node8.html>
- E. R. Tufte. The Visual Display of Quantitative Information, 2nd Edition. Graphics Press, Cheshire, Connecticut, 2001.
- Cole Nussbaumer Knaflic, Storytelling with data – a data visualization guide for business professionals, Wiley, 2015