DATA VISUALIZATION

Graphs
WHAT IS A GRAPH, REALLY?
1. **graph**

A slang term for a sketching of what someone plans on tagging.

**graffiti.**

*boy 1: Hey check out this graph i drew in science class.*

*boy 2: It looks good call me when you want to hit up.*

by Eamz February 08, 2007

👍 27 👎 18
2. **Graph**

A word used in the north west around the area of Manchester and Liverpool

It means "bad" or something that is "annoying" as in "mate that was well graph" or "stop being so graph"

"that was so graph"
"stop being so graph"

by oli hughes November 07, 2012
Graph - Mythical 1337 Creature. The peak of gaming excellence.

see leet uber
Antonyms non-factor

That was SO Graph.
He killed that thing with Graph like Skills.

by BoBO1 June 10, 2005
A diagram representing a system of connections or interrelations among two or more things by a number of distinctive dots, lines, bars, etc.
#1 Infographic

#2 Scientific Journals

#3 News media

#4 Government Reports

Borkin et al. What makes a Visualization Memorable? 2013 IEEE
A visualization was more memorable if... it contains human recognizable objects.
A visualization was more memorable if... it is distinct.
A visualization was more memorable if...

it is a distinct visualization type.
A visualization was more memorable if...

*it is colorful.*
A visualization was more memorable if... *it is visually dense.*
A visualization was more memorable if...

it has a low data-to-ink ratio.
Graphs aren’t always necessary.
HOW DO YOU FEEL ABOUT DOING SCIENCE?
<table>
<thead>
<tr>
<th>Interest</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excited</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td>Kind of interested</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>OK</td>
<td>40</td>
<td>14</td>
</tr>
<tr>
<td>Not great</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Bored</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

Data courtesy of Cole Nussbaumer
How do you feel about doing science?

<table>
<thead>
<tr>
<th>Interest</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excited</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td>Kind of interested</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>OK</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Not great</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Bored</td>
<td>40</td>
<td>6</td>
</tr>
</tbody>
</table>

Example courtesy of Hanspeter Pfister
How do you feel about doing science?

<table>
<thead>
<tr>
<th>Level [%]</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>10-20</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>20-30</td>
<td>40</td>
<td>14</td>
</tr>
<tr>
<td>30-40</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>40-50</td>
<td>19</td>
<td>38</td>
</tr>
</tbody>
</table>

Interest:
- Excited
- Kind of interested
- OK
- Not great
- Bored

Example courtesy of Hanspeter Pfister
How do you feel about doing science?

Before the program, the majority of children felt just OK about science. After the program, more children were Kind of interested and Excited about science.

Example courtesy of Hanspeter Pfister
Opinion change to the question: How do you feel about doing science?

- Excited: -26.00
- Kind of interested: 5.00
- OK: 0.00
- Not great: -1.00
- Bored: 1.00

Example courtesy of Hanspeter Pfister
How do you feel about doing science?

Example courtesy of Hanspeter Pfister
After the pilot program, 68% of kids expressed interest towards science, compared to 44% going into the program.

Example courtesy of Hanspeter Pfister
tomaytoe
Solanum lycopersicum
$64
PANTONE 1795 C
tdTomato 554ex 581em
$64
DATA DEFINITIONS
QUANTITATIVE | QUALITATIVE
QUANTITATIVE  QUALITATIVE

RATIO  INTERVAL  ORDINAL  NOMINAL
VALUES

ORDINAL

CATEGORICAL

RELATIONAL
SO WHAT?
VALUES
GRAPHS: WORKING PARTS
Title of this Graph
A description of the data or something worth highlighting to set the stage.

Source: Somewhere reputable

Nathan Yau, Data Points
Title of this Graph
A description of the data or something worth highlighting to set the stage.

100 units

0


Source: Somewhere reputable

Visual Cues
Visualization involves encoding data with shapes, colors, and sizes. Which cues you choose depends on your data and your goals.

Coordinate System
You map data differently with a scatterplot than you do with a pie chart. It’s x- and y-coordinates in one and angles with the other; it’s cartesian versus polar.

Nathan Yau, Data Points
Title of this Graph
A description of the data or something worth highlighting to set the stage.

Source: Somewhere reputable

Scale
Increments that make sense can increase readability, as well as shift focus.

Context
If your audience is unfamiliar with the data, it’s your job to clarify what values represent and explain how people should read your visualization.

Nathan Yau, Data Points
GRAPHING: VISUAL CUES

PROXIMITY
VALUES
COLOR
PROXIMITY
VALUES
QUANTITATIVE  QUALITATIVE

ASSOCIATION  DIFFERENTIATION
GRAPHING: COORDINATE SYSTEMS
VISUAL CUES + COORDINATE SYSTEMS
## Visual Cues

<table>
<thead>
<tr>
<th>Coordinate Systems</th>
<th>Position</th>
<th>Length</th>
<th>Angle</th>
<th>Direction</th>
<th>Shapes</th>
<th>Area or Volume</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartesian</td>
<td><img src="image1.png" alt="Example" /></td>
<td><img src="image2.png" alt="Example" /></td>
<td><img src="image3.png" alt="Example" /></td>
<td><img src="image4.png" alt="Example" /></td>
<td><img src="image5.png" alt="Example" /></td>
<td><img src="image6.png" alt="Example" /></td>
<td><img src="image7.png" alt="Example" /></td>
</tr>
<tr>
<td>Polar</td>
<td><img src="image8.png" alt="Example" /></td>
<td><img src="image9.png" alt="Example" /></td>
<td><img src="image10.png" alt="Example" /></td>
<td><img src="image11.png" alt="Example" /></td>
<td><img src="image12.png" alt="Example" /></td>
<td><img src="image13.png" alt="Example" /></td>
<td><img src="image14.png" alt="Example" /></td>
</tr>
<tr>
<td>Geographic</td>
<td><img src="image15.png" alt="Example" /></td>
<td><img src="image16.png" alt="Example" /></td>
<td><img src="image17.png" alt="Example" /></td>
<td><img src="image18.png" alt="Example" /></td>
<td><img src="image19.png" alt="Example" /></td>
<td><img src="image20.png" alt="Example" /></td>
<td><img src="image21.png" alt="Example" /></td>
</tr>
</tbody>
</table>

Nathan Yau, Data Points
Graphing: Scale

Linear
Values are evenly spaced

Categorical
Discrete placement in bins

Percent
Representing parts of a whole

Nathan Yau, Data Points
Graphing: Scale

Logarithmic
Focus on percent change

1 10 100 1,000 10,000

Ordinal
Categories where order matters

Horrible Bad Okay Good Great

Time
Units of months, days, or hours

Jan. Feb. March April May

Nathan Yau, Data Points
Figure 1. chicken
1 | DATA
2 | WORKING PARTS
3 | GRAPH TYPES
4 | WHAT ARE WE TRYING TO SHOW?
5 | EXAMPLES & EXPLANATIONS
GRAPH TYPES: HEAVER HITTERS
GRAPH TYPES:
OTHER
GRAPH TYPES:
EXOTIC
Patterns and Shape

Increase
Decrease
Combination
Outlier
Noise
<table>
<thead>
<tr>
<th>Patterns</th>
<th>Position</th>
<th>Length</th>
<th>Angle</th>
<th>Direction</th>
<th>Area or Volume</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase</td>
<td><img src="image1" alt="Pattern" /></td>
<td><img src="image2" alt="Length" /></td>
<td><img src="image3" alt="Angle" /></td>
<td><img src="image4" alt="Direction" /></td>
<td><img src="image5" alt="Area or Volume" /></td>
<td><img src="image6" alt="Color" /></td>
</tr>
<tr>
<td>Decrease</td>
<td><img src="image7" alt="Pattern" /></td>
<td><img src="image8" alt="Length" /></td>
<td><img src="image9" alt="Angle" /></td>
<td><img src="image10" alt="Direction" /></td>
<td><img src="image11" alt="Area or Volume" /></td>
<td><img src="image12" alt="Color" /></td>
</tr>
<tr>
<td>Combination</td>
<td><img src="image13" alt="Pattern" /></td>
<td><img src="image14" alt="Length" /></td>
<td><img src="image15" alt="Angle" /></td>
<td><img src="image16" alt="Direction" /></td>
<td><img src="image17" alt="Area or Volume" /></td>
<td><img src="image18" alt="Color" /></td>
</tr>
<tr>
<td>Outlier</td>
<td><img src="image19" alt="Pattern" /></td>
<td><img src="image20" alt="Length" /></td>
<td><img src="image21" alt="Angle" /></td>
<td><img src="image22" alt="Direction" /></td>
<td><img src="image23" alt="Area or Volume" /></td>
<td><img src="image24" alt="Color" /></td>
</tr>
<tr>
<td>Noise</td>
<td><img src="image25" alt="Pattern" /></td>
<td><img src="image26" alt="Length" /></td>
<td><img src="image27" alt="Angle" /></td>
<td><img src="image28" alt="Direction" /></td>
<td><img src="image29" alt="Area or Volume" /></td>
<td><img src="image30" alt="Color" /></td>
</tr>
</tbody>
</table>
Chart Suggestions—A Thought-Starter

**Comparison**
- What would you like to show?

**Relationship**
- Two Variables
- Three Variables
- Among Items

**Distribution**
- Few Variables
- Over Time

**Composition**
- Single Variable
- Many Data Points
- Static

**Chart Types**
- Variable Width Column Chart
- Table or Table with Embedded Charts
- Bar Chart
- Column Chart
- Circular Area Chart
- Line Chart
- Column Chart
- Line Chart
- Scatter Chart
- Bubble Chart
- 3D Area Chart
- Stacked 100% Column Chart
- Stacked Column Chart
- Stacked 100% Area Chart
- Stacked Area Chart
- Pie Chart
- Waterfall Chart
- Stacked 100% Column Chart with Subcomponents

Juice Analytics – Graph Chooser | Extreme Presentations.com
CATEGORICAL DATA: COMPARISON/COMPOSITION
Categories

When your data is straightforward, with a value for each category, these are easy to read and create.

Bar graph

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>

With length as visual cue, useful for straightforward comparisons

Symbol plot

A  B  C

Can be used in place of bars, but can be hard to see small differences

Nathan Yau, Data Points
Parts of a whole

The categorical breakdown within a population can be interesting, and you might want to keep the groups together, although often not essential.

**Pie chart**

Parts add to 100 percent, typically sorted clockwise for readability

**Stacked bar chart**

Often used to show poll results and can also be used for raw counts

Nathan Yau, Data Points
Subcategories

Data can have a hierarchical structure, which can be important in data interpretation and it often allows for different points of view.

**Treemap**

A

B

C

D

E

Shows hierarchical structure in a compact space, area often combined with color

**Mosaic plot**

A

B

C

D

E

Allows comparison across multiple categories in one view
65% of the market is controlled by companies B and C
RELATIONAL DATA
CORRELATION
Should You Trust That Doctor?

**Trust with your life:**
- The Doctor
- Dr. Seuss
- Dr. Strange
- Dr. Quinn
- Dr. Huxtable
- Doogie Howser, MD

**Possesses a professional doctorate in medical studies:**
- Doc Brown
- J.D.
- Dr. Bunsen Honeydew
- Dr. Fishman
- Dr. Oz
- Dr. Hannibal Lecter

**No professional training:**
- Spin Doctors
- Dr. Dre
- Doctor Venture
- Dr. Horrible
- Doctor Evil
- Dr. Robotnik

**Run, run very far:**
- Zoidberg
- Dr. Phil
- Dr. Spaceman
- Dr. Octopus
- Dr. Pepper
## Nathan Yau, Data Points

<table>
<thead>
<tr>
<th>Games played</th>
<th>Usage</th>
<th>Field goal %</th>
<th>Free throw %</th>
<th>3-point %</th>
<th>Turnovers</th>
<th>Turnover ratio</th>
<th>Minutes per</th>
<th>Points per</th>
<th>Rebounds per</th>
<th>Assists per</th>
</tr>
</thead>
</table>

![Data Points](image)
<table>
<thead>
<tr>
<th>Games played</th>
<th>Usage</th>
<th>Field goal %</th>
<th>Free throw %</th>
<th>3-point %</th>
<th>Turnovers</th>
<th>Turnover ratio</th>
<th>Minutes per</th>
<th>Points per</th>
<th>Rebounds per</th>
<th>Assists per</th>
</tr>
</thead>
</table>

Nathan Yau, Data Points
Nathan Yau, Data Points
DISTRIBUTIONS
**Distribution Summary**

You can visualize data at different granularities with the charts above. These show key values for a less specific view of distributions.

**Box plot**

![Box plot diagram](image)

Shows range, median and quartiles

**Violin plot**

![Violin plot diagram](image)

A combination of a box plot and density plot
Distribution of one variable

You can see where data is clustered and see any outliers by keeping track of where they sit on a value axis.

Histogram

Value axis on the horizontal and density on the vertical

Density plot

Like histogram but continuous instead of bins

Nathan Yau, Data Points
Distribution of multiple variables
Sometimes values come as pairs, and it makes sense to show both values at the same time.

Heat map
Density on a 2-D plane, using color as visual cue

Surface plot
Shows same patterns as heat map, but uses height instead of color
100 IMAGINARY PEOPLE
Nathan Yau, Data Points
**One-inch bins**
Small bins show variations at higher granularity.

**Two-inch bins**
You see less variation, but the distribution around the median is more obvious.

**Half-foot bins**
You can see distribution around the median, but you can only see some variation.

**One-foot bins**
The spread of the data isn’t as obvious, because the larger bins show less detail.
1 | DATA
2 | WORKING PARTS
3 | GRAPH TYPES
4 | WHAT ARE WE TRYING TO SHOW?
5 | SPECIFIC EXAMPLES & EXPLANATIONS
EXAMPLES
WHAT IS A GRAPH, REALLY?
A graph is a means to an end. You need to understand the end to find the best means...

Jackie, hot air