Welcome to Pie Charts Anonymous

Our Motto:

“In God we trust, all others bring data.”
W. Edwards Deming

Business Intelligence
Transforms beyond the data
What is Data Visualization?
And why is it important?

How many times does the digit 7 appear?

5 2 8 3 6 1 9 3 6 2 5 3 7 4 3 8 3
8 5 8 9 6 2 1 4 4 3 9 3 6 5 2 4 9
1 0 2 7 5 2 8 3 6 1 6 2 9 3 8 3 8
5 8 4 7 2 0 3 7 3 5 4 7 1 8 2 0 1
2 5 3 6 4 3 9 1 0 8 9 5 7 3 4 5 3
2 7 5 2 8 3 6 1 6 2 9 3 8 3 8 5 8
4 7 2 0 3 7 3 5 4 7 1 8 2 0 1 9 6
2 1 4 4 3 9 3 6 5 2 4 9 1 0 2 7 5
2 8 3 6 1 6 2 9 3 8 3 8 5 8 4 7 2
0 3 7 3 5 4 7 1 8 2 0 1 2 5 3 6 4
3 9 1 0 8 9 5 7 3 4 5 3 2 7 5 2 8
3 6 1 6 2 4 6 2 7 5 9 1 5 2 6 3 6

5 2 8 3 6 1 9 3 6 2 5 3 7 4 3 8 3
8 5 8 9 6 2 1 4 4 3 9 3 6 5 2 4 9
1 0 2 7 5 2 8 3 6 1 6 2 9 3 8 3 8
5 8 4 7 2 0 3 7 3 5 4 7 1 8 2 0 1
2 5 3 6 4 3 9 1 0 8 9 5 7 3 4 5 3
2 7 5 2 8 3 6 1 6 2 9 3 8 3 8 5 8
4 7 2 0 3 7 3 5 4 7 1 8 2 0 1 9 6
2 1 4 4 3 9 3 6 5 2 4 9 1 0 2 7 5
2 8 3 6 1 6 2 9 3 8 3 8 5 8 4 7 2
0 3 7 3 5 4 7 1 8 2 0 1 2 5 3 6 4
3 9 1 0 8 9 5 7 3 4 5 3 2 7 5 2 8
3 6 1 6 2 4 6 2 7 5 9 1 5 2 6 3 6
Business Intelligence

Moving from data to action...

- Data
- Information
- Knowledge

How many...?  How much...?  What are...?  Why are...?  How do we...?  What can we...?

Plan to guide actions or key decisions

Visualization and Data Presentation

Visual Perception
Visualization and Data Presentation

Compare And Contrast
**Visualization and Data Presentation**

**Pie Charts**

Likely Outcomes of Owning a Real Lightsaber

- Save the galaxy
- Sell it to the highest bidder
- Become a ruthless killer
- Immediate, accidental self-dismemberment

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Save the galaxy</td>
<td>3</td>
</tr>
<tr>
<td>Sell it to the highest bidder</td>
<td>5</td>
</tr>
<tr>
<td>Become a ruthless killer</td>
<td>2</td>
</tr>
<tr>
<td>Immediate, accidental self-dismemberment</td>
<td>3</td>
</tr>
</tbody>
</table>

- 5 2 8 3 6 1 9 3 6 2 5 3 7 4 3 8 3
- 8 5 8 9 6 2 1 4 4 3 9 3 6 5 2 4 9
- 1 0 2 7 5 2 8 3 6 1 6 2 9 3 8 3 8
- 5 8 4 7 2 0 3 7 3 5 4 7 1 8 2 0 1
- 2 5 3 6 4 3 9 1 0 8 9 5 7 3 4 5 3
- 2 7 5 2 8 3 6 1 6 2 9 3 8 3 8 5 8
- 4 7 2 0 3 7 3 5 4 7 1 8 2 0 1 9 6
- 2 1 4 4 3 9 3 6 5 2 4 9 1 0 2 7 5
- 2 8 3 6 1 6 2 9 3 8 3 8 5 8 4 7 2
- 0 3 7 3 5 4 7 1 8 2 0 1 2 5 3 6 4
- 3 9 1 0 8 9 5 7 3 4 5 3 2 7 5 2 8
- 3 6 1 6 2 4 6 2 7 5 9 1 5 2 6 3 6
This is how a Pie Chart Represents the Data

Try to quickly compare the totals of the digits.

This is how a Bar Chart Represents the Data

Easier comparison even without color encoding!
General Rules for Pie Charts

1. Don’t Use Pie Charts
   If you must break Rule #1 then:
2. Make sure it adds up 100%
3. Only a few categories
4. Start at noon and move clockwise
5. Largest to Smallest Values
6. Add Labels for %
7. Avoid 3D
8. Keep it Simple
The most common visualizations

Bar Chart - category comparison (with target line)
Line Chart - time series data
Flow Chart – process flow (also Swimlane diagram)
Bullet Graph
100% Stacked Bar Chart
Scatter Plot – relationship/correlation
Box Plot – grouping with summaries
Control Charts
Area Chart

Visualization and Data Presentation

Poor Choices:
- Wrong Chart Type for the data
- 3D chart
- Use of gradient and shadows
- Rotated Text
- Hard to Read
- Over use of gridlines
- Over use of axis and data labels
- Bad Color Choices
Visualization and Data Presentation

Really Poor Choice:

Visualization and Data Presentation

Even a bad 3D Pie Chart is better:
Visualization and Data Presentation

Acceptable use of the Pyramid:

Maslow’s Hierarchy of Needs

- Physiological: breathing, food, water, sex, sleep, homeostasis, excretion
- Safety: security of body, of employment, of resources, of morality, of the family, of health, of property
- Love/Belonging: friendship, family, sexual intimacy
- Esteem: self-esteem, confidence, achievement, respect of others, respect by others
- Self-actualization: morality, creativity, spontaneity, problem solving, lack of prejudice, acceptance of facts

Visualization and Data Presentation

Poor Choices:
- Avoid Pyramids, Funnels and Cones
- Cluttered Dashboard
  - Too Much Information (Crammed on One Page)
  - Lots of Scroll
- Over use of Graphs
- Gauges
- Flashing Cells
- Wrong Message or No Message at all

Visualization and Data Presentation

Best Practices:
- Books on Data Visualization
  - Jacques Bertin
    - Semiology of Graphics
  - Edward Tufte
    - The Visual Display of Quantitative Information
    - Envisioning Information
  - Stephen Few
    - Information Dashboard Design
    - Show Me the Numbers
    - Now You See it

Visualization and Data Presentation

Data Visualization by the Experts
On Minard’s graph:

“This is one of the worse graphs ever made. He [Tufte]’s very happy because it shows five different pieces of information on three axis and if you study it for fifteen minutes it really is worth a thousand words. I don’t think that’s what graphs are for. I think they try to make a point in two seconds for people who are too lazy to read the forty words underneath.”

- Seth Godin
Every Few Months: 42%
Never: 29%
Every Few Weeks: 19%
Every Few Days: 8%
Daily: 2%

71% of occurrences are caused by low frequency.
The KISS Principle:

- Keep it Simple, Stupid!

“It seems that perfection is reached not when there is nothing to add, but when there is nothing left to take away.”

- Antoine de Saint Exupéry
The Shaffer 4 C’s of Data Visualization

**Clear** - easily seen; sharply defined
- who’s the audience? what’s the message?
- clarity more important than aesthetics

**Clean** - thorough; complete; unadulterated
- labels, axis, gridlines, formatting, right chart type, color choice, etc.

**Concise** - brief but comprehensive;
- not minimalist but not verbose

**Captivating** - to attract and hold by beauty or excellence
- does it capture attention? is it interesting? does it tell the story?

Visualization and Data Presentation

**Figure out the message**
**Then apply the tools**
Percentage of chart that looks like PacMan

- 85%
- 15%

Percentage of Chart which Looks like PacMan

- Looks Like PacMan: 85%
- Does Not Look Like PacMan: 15%

<table>
<thead>
<tr>
<th>% of Chart that looks like Pacman</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Looks Like Pacman</td>
<td>15%</td>
</tr>
<tr>
<td>Does not look like Pacman</td>
<td>85%</td>
</tr>
</tbody>
</table>

How Ingenious a Pringles Can Is

- Full container: This is awesome
- 3/4 container: Hmm, this is getting parked
- 1/2 container: Crop, my hand got stuck
- 1/4 container: Great, I just didn't like them

Amount of Pringles Left
How Ingenious a Pringles Can is:

- Full Container
- 3/4 Container
- 1/2 Container
- 1/4 Container

Ingeniousness
- This is awesome!
- Hmm, this is getting awkward.
- Crap, my hand got stuck!
- @#$% it, I'll just drink them.

Amount of Pringles Left

Source: www.HamiltonCountyAuditor.org
Visualization and Data Presentation

What is the Message?

Food expenditure for one week

- Germany: $500.07
- Ecuador: $31.55
Germany: The Melander family of Bargteheide
Food expenditure for one week: $500.07
Source: "Hungry Planet"
By Peter Menzel and Faith D'Aluisio

Ecuador: The Ayme family of Tingo
Food expenditure for one week: $31.55
Source: "Hungry Planet"
By Peter Menzel and Faith D'Aluisio

Elle and Nina
A Data Visualization of your checking account

Current Balance

Checking Acct #123456789: $843.62
Checking Acct #123456789: -$43.17

Visualization and Data Presentation

Sparklines (Tuft):
Visualization and Data Presentation

Sparklines:

- 2,922 Visits
- 1,694 Absolute Unique Visitors
- 4,975 Pageviews
- 1.70 Average Pageviews
- 00:02:21 Time on Site
- 71.59% Bounce Rate
- 57.97% New Visits

Visualization and Data Presentation

Sparktweets

Fresh data on world primary energy use by %: Oil 34.7%, Coal 25.3%, NG 25.9%, Nuke 5.47, Hydro 6.63 | spark: @pkadrosky @alexkerin

@indius record 2011 to date:

@timhaines — how hard I was laughing before and after reading that tweet #SparkTweet

`'11 US Budget ($=820B): Mass/Transit _
Airports  Highways  _  Defense

Invented by Stephen Few (January 2013)

Visualization and Data Presentation

Trellis Charts (Small Multiples):

- "Illustrations of postage-stamp size are indexed by category or a label, sequenced over time like the frames of a movie, or ordered by a quantitative variable not used in the single image itself."
Protanomaly (1 out of 100 males) – “red weak”
Deuteranomaly (5 out of 100 males) – “green weak”
Dichromasy (2 out of 100 males) - no perceptible difference between red, orange, yellow, and green
Protanopia (1 out of 100 males) - use light/dark
Deuteranopia (one out of 100 males) – “green blind”

= 10% of males are Color Blind
Main issue is Red and Green
“Don’t be Ugly By Accident!”
OkCupid (blog.okcupid.com)

- Collected 552,000 examples of users pictures
- Asked people to rate their attractiveness
- Compared with the data from the camera
Visualization and Data Presentation

What did we learn from OkCupid?

1. Someone went rogue with Microsoft Excel

2. Take your profile picture with an SLR camera without the flash at 6:30am or 5pm for your best chance of being considered attractive.

Visualization and Data Presentation

Showing performance vs. target (ex. actual to forecast)

- Bar Chart with Target line
  - Easy to build
  - Good when single target for all categories

[Graph example]

100% Stacked Bar Chart
Visualization and Data Presentation

**Bullet Graph:**

Invented by Stephen Few

**Visualization and Data Presentation**

Two Variations on the Bullet Graph:

- Remove performance bands
  - Now it’s simply actual to target
  - Easier to understand
- Dual Axis Bullet Graph
  - Shows both value and %

**Geospatial:**

- Location can relate disparate data
- A Map is worth a thousand cells
- Easy to see problem areas
- By State, County, Zip, Block
- Even down to Address

![Unemployment rates by state, seasonally adjusted, December 2008](image)
Visualization and Data Presentation

Sequential
Colors can be ordered from low to high

Diverging
Two sequential schemes extending out from a critical midpoint value

Categorical
Lots of contrast between each adjacent color
Foreclosures (July 2008)

Source: RealtyTrac (www.RealtyTrac.com)

Breaking News: Missouri invades Arkansas, Mississippi and Alabama Land Swap

Kroger Pharmacy Growth?

What the data actually shows:

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacies</td>
<td>1,920</td>
<td>1,964</td>
<td>1,973</td>
<td>1,967</td>
<td>1,966</td>
<td>1,948</td>
</tr>
<tr>
<td>Stores</td>
<td>2,468</td>
<td>2,486</td>
<td>2,481</td>
<td>2,468</td>
<td>2,458</td>
<td>2,435</td>
</tr>
</tbody>
</table>


The Real Story

Total Number of Kroger Stores

Total Number of Kroger Pharmacies
What is the real message?

1. From 2007 to 2010 Kroger closed stores that don’t have pharmacies thus increasing the % of stores with a pharmacy.

2. Kroger has increased Pharmacy Sales from $6.5B to $7.3B with fewer stores and fewer pharmacies

They missed a very compelling Story
12% increase in Sales with Lower Costs!

Interactive Visualizations

Word Clouds
(tag clouds)

www.Wordle.net

Bush State of the Union - 2002
Without the Clutter of Unimportant Words: Descriptive Keyphrases for Text Visualization

Jason Chuang, Christopher D. Manning, Jeffrey Heer
Visualization and Data Presentation

Hamilton County vs. Auditor's Office

Same data but a completely different story
Since 2008, Hamilton County has reduced 1,442 full-time employees while the Auditor's Office has only eliminated 13.
Visualization and Data Presentation

Additional Materials:
- PerceptualEdge.com
- JuiceAnalytics.com/writing/
  - Blog
- vis.stanford.edu
- EdwardTufte.com
- FlowingData.com
- Storytellingwithdata.com
- vizwiz.blogspot.com
- Datarevelations.com