Information Visualization
Crash Course

(AKA Information Visualization 101)

Chad Stolper
Assistant Professor
Southwestern University
(graduated from Georgia Tech CS PhD)
What is Infovis?
Why is it Important?
Human Perception
Chart Basics
(If Time, Some Color Theory)
The Shneiderman Mantra
Where to Learn More
What is Information Visualization?
Information Visualization

“The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.”

Card, Mackinlay, and Shneiderman 1999
Communication

Exploratory Data Analysis (EDA)
Communication

(gone wrong)
EDWARD R. TUFTET
VISUAL EXPLANATIONS
IMAGES AND QUANTITIES, EVIDENCE AND NARRATIVE
Edward Tufte

An American statistician and professor emeritus of political science, statistics, and computer science at Yale University.

He is noted for his writings on information design and as a pioneer in the field of data visualization.

-Wikipedia
Space Shuttle Challenger

January 28, 1986
Morning Temperature: 31°F
Less than 1 second after ignition, a puff of smoke appeared at the aft joint of the right booster, indicating that the O-rings burned through and failed to seal. At this point, all was lost.

On the launch pad, the leak lasted only about 2 seconds and then apparently was plugged by fuel and insulation as the shuttle rose, flying through rather strong cross-winds. Then 58.7 seconds after ignition, when the Challenger was 6 miles up, a flicker of flame emerged from the leaky joint. In 3 seconds, the flame grew and engulfed the fuel tank (containing liquid hydrogen and liquid oxygen) that tank ruptured and exploded, destroying the shuttle.

As the shuttle exploded and broke up at approximately 73 seconds after launch, the two booster rockets crisscrossed and continued flying wildly. The right booster, identifiable by its failure plume, is now to the left of its non-defective counterpart.

The flight crew of Challenger 51-L. Front row, left to right: James D. (Jim) Irwin, pilot; Francis R. (Dick) Scobee, commander; Ronald E. McNair, payload specialist. Back row: Ellison S. Onizuka, S. Christa McAuliffe, Gregory C. Jarvis, and Judith A. Resnik.
Rubber O-rings, nearly 38 feet (11.6 meters) in circumference; 1/4 inch (6.4 mm) thick.

The field joint that leaked.

Most Watched Science Experiment

Richard Feynman, Physics Nobel laureate explained how rubber became rigid in cold temperate

YouTube video: https://youtu.be/6Rwcbsn19c0

Video originally from: http://www.FeynmanPhysicsLectures.com
How did this happen?
Engineers at **Morton Thiokol**, the rocket maker, presented on the day before and recommended **not to launch.**

<table>
<thead>
<tr>
<th>Temperature Concern on SRM Joints</th>
</tr>
</thead>
<tbody>
<tr>
<td>27 Jan 1986</td>
</tr>
</tbody>
</table>
Conclusions:

- Temperature of O-ring is not only parameter controlling blow-by.
  - SRM 15 with blow-by had an O-ring temp at 53°F.
  - SRM 22 with blow-by had an O-ring temp at 75°F.
  - Four development motors with no blow-by were tested at O-ring temp of 47° to 52°F.
  - Development motors had putty packing which resulted in better performance.

- At about 50°F blow-by could be experienced in case joints.

- Temp for SRM 25 on 1-28-84 launch will be 29°F at 9 AM and 38°F at 2 PM.

- Have no data that would indicate SRM 25 is different than SRM 15 other than temp.

Recommendations:

- O-ring temp must be ≥ 53°F at launch.
  - Development motors at 47° to 52°F with putty packing had no blow-by.
  - SRM 15 (the best simulation) worked at 53°F.

- Project ambient conditions (temp & wind) to determine launch time.
## History of O-Ring Damage in Field Joints (Cont)

<table>
<thead>
<tr>
<th>SRM No.</th>
<th>O-Ring Temp (°F)</th>
<th>SRM No.</th>
<th>O-Ring Temp (°F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66°</td>
<td>13</td>
<td>70°</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>14</td>
<td>67°</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>15</td>
<td>69°</td>
</tr>
<tr>
<td>2</td>
<td>70°</td>
<td>16</td>
<td>68°</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>17</td>
<td>72°</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>18</td>
<td>73°</td>
</tr>
<tr>
<td>3</td>
<td>69°</td>
<td>19</td>
<td>70°</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>20</td>
<td>57°</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>21</td>
<td>E</td>
</tr>
<tr>
<td>4</td>
<td>80°</td>
<td>22</td>
<td>63°</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>23</td>
<td>H</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>24</td>
<td>78°</td>
</tr>
</tbody>
</table>

* No Erosion

---

**Morton Thiokol, Inc.**

Wasatch Operations

*Information on this page was prepared to support an oral presentation and cannot be considered complete without the oral discussion.*

[Ref. 2/26-2 2 of 3]
<table>
<thead>
<tr>
<th>Flight</th>
<th>Date</th>
<th>Temperature °F</th>
<th>Erosion incidents</th>
<th>Blow-by incidents</th>
<th>Damage index</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>51-C</td>
<td>01.24.85</td>
<td>53 °</td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>Most erosion any flight; blow-by; back-up rings heated. Deep, extensive erosion.</td>
</tr>
<tr>
<td>41-B</td>
<td>02.03.84</td>
<td>57 °</td>
<td>1</td>
<td></td>
<td>4</td>
<td>O-ring erosion on launch two weeks before Challenger. O-rings showed signs of heating, but no damage. Coolest (66°) launch without O-ring problems.</td>
</tr>
<tr>
<td>61-C</td>
<td>01.12.86</td>
<td>58 °</td>
<td>1</td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>41-C</td>
<td>04.06.84</td>
<td>63 °</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>04.12.81</td>
<td>66 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>04.04.83</td>
<td>67 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>51-A</td>
<td>11.08.84</td>
<td>67 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>51-D</td>
<td>04.12.85</td>
<td>67 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>11.11.82</td>
<td>68 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>03.22.82</td>
<td>69 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>11.12.81</td>
<td>70 °</td>
<td>1</td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>11.28.83</td>
<td>70 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>41-D</td>
<td>08.30.84</td>
<td>70 °</td>
<td>1</td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>51-G</td>
<td>06.17.85</td>
<td>70 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>06.18.83</td>
<td>72 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>08.30.83</td>
<td>73 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>51-B</td>
<td>04.29.85</td>
<td>75 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>61-A</td>
<td>10.30.85</td>
<td>75 °</td>
<td></td>
<td>2</td>
<td>4</td>
<td>No erosion. Soot found behind two primary O-rings.</td>
</tr>
<tr>
<td>51-I</td>
<td>08.27.85</td>
<td>76 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>61-B</td>
<td>11.26.85</td>
<td>76 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>41-G</td>
<td>10.05.84</td>
<td>78 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>51-J</td>
<td>10.03.85</td>
<td>79 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>06.27.82</td>
<td>80 °</td>
<td></td>
<td></td>
<td>?</td>
<td>O-ring condition unknown; rocket casing lost at sea.</td>
</tr>
<tr>
<td>51-F</td>
<td>07.29.85</td>
<td>81 °</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
O-ring damage index, each launch

26°–29° range of forecasted temperatures (as of January 27, 1986) for the launch of space shuttle Challenger on January 28

Temperature (°F) of field joints at time of launch
So, communication is extremely important.

Visualization can help with that – communicate ideas and insights.
Hans Rosling:

The best stats you've ever seen

TED2006 · 19:50 · Filmed Feb 2006
Subtitles available in 48 languages

http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_veever_seen.html
Visualization can also help with Exploratory Data Analysis (EDA)

But why do you need to explore data at all???
“There are three kinds of lies: lies, damned lies, and statistics.”
Mystery Data Set
## Mystery Data Set

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean( x )</td>
<td>9</td>
</tr>
<tr>
<td>variance ( x )</td>
<td>11</td>
</tr>
<tr>
<td>mean( y )</td>
<td>7.5</td>
</tr>
<tr>
<td>variance ( y )</td>
<td>4.122</td>
</tr>
<tr>
<td>correlation ( x,y )</td>
<td>0.816</td>
</tr>
<tr>
<td>Linear Regression Line</td>
<td>y = 3 + 0.5x</td>
</tr>
</tbody>
</table>
Anscombe’s Quartet

https://en.wikipedia.org/wiki/Anscombe%27s_quartet
Anscombe’s Quartet

Sanity Checking Models

Outlier Detection
Data visualization leverages human perception
Name the five senses.
<table>
<thead>
<tr>
<th>Sense</th>
<th>Bandwidth (bits/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sight</td>
<td>10,000,000</td>
</tr>
<tr>
<td>Touch</td>
<td>1,000,000</td>
</tr>
<tr>
<td>Hearing</td>
<td>100,000</td>
</tr>
<tr>
<td>Smell</td>
<td>100,000</td>
</tr>
<tr>
<td>Taste</td>
<td>1,000</td>
</tr>
</tbody>
</table>
A (Simple) Model of Human Visual Perception
A (Simple) Model of Human Perception

Stage 1
Parallel detection of basic features into an iconic store

Stage 2
Serial processing of object identification and spatial layout
Stage 1: Pre-Attentive Processing

Rapid
Parallel
Automatic

(Fleeting = lasting for a short time)
Stage 2: Serial Processing

Relatively Slow
(Incorporates Memory)
Manual
Stage 1: Pre-Attentive Processing

The eye moves every 200ms
(so this processing occurs every 200ms-250ms)
Example

1281768756138976546984506985604982826762
9809858458224509856458945098450980943585
9091030209905959595772564675050678904567
8845789809821677654876364908560912949686
Example

1281768756138976546984506985604982826762
98098584582224509856458945098450980943585
9091030209905959595772564675050678904567
8845789809821677654876364908560912949686
A few more examples from Prof. Chris Healy at NC State
Raise your hand if a **RED DOT** is present...

(On the left or on the right?)
Color (hue) is pre-attentively processed.
Raise your hand if a RED DOT is present...
Shape is pre-attentively processed.
Determine if a RED DOT is present...
Hue and shape together are NOT pre-attentively processed.
Pre-Attentive Processing

- length
- width
- size
- curvature
- number
- terminators
- intersection
- closure
- hue
- lightness
- flicker
- direction of motion
- binocular lustre
- stereoscopic depth
- 3-D depth cues
- lighting direction
<table>
<thead>
<tr>
<th>Group</th>
<th>Attribute</th>
<th>Length</th>
<th>Width</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form</td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Curvature</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enclosure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blur</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color</td>
<td>Hue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intensity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Position</td>
<td>2-D Position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spatial Grouping</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motion</td>
<td>Direction</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pre-Attentive $\rightarrow$ Cognitive
Gestalt Psychology

Berlin, Early 1900s
Gestalt Psychology

Goal was to understand pattern perception

Gestalt (German) = “seeing the whole picture all at once” instead of a collection of parts

Identified 8 “Laws of Grouping”

Gestalt Psychology

1. Proximity
2. Similarity
3. Closure
4. Symmetry
5. Common Fate
6. Continuity
7. Good Gestalt
8. Past Experience
How many groups are there?
Proximity
How many groups are there?
Similarity
How many shapes are there?
Closure
How many items are there?
How many sets are there?
Common Fate
How many objects are there?
Continuity
How many objects are there?
Good Gestalt
What is this word?
FLIGHT
Past Experience

FLIGHT
Pre-Attentive Processing

Gestalt Laws
Detect Quickly
Detect quickly does NOT mean detect accurately.

Ideally you want both.
More accurate

Less accurate

Position

Length

Angle

Slope

Area

Volume

Color

(Density

Mackinlay, 1986
<table>
<thead>
<tr>
<th>Precision of Quantitative Perception</th>
<th>Attribute</th>
<th>Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very precise</td>
<td>Length</td>
<td></td>
<td>Longer = greater</td>
</tr>
<tr>
<td></td>
<td>2-D Position</td>
<td>● ● ● ● ●</td>
<td>Higher or farther to the right = greater</td>
</tr>
<tr>
<td>Not very precise</td>
<td>Width</td>
<td></td>
<td>Wider = greater</td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td>● ● ● ● ●</td>
<td>Bigger = greater</td>
</tr>
<tr>
<td></td>
<td>Intensity</td>
<td>● ● ● ● ●</td>
<td>Darker = greater</td>
</tr>
<tr>
<td></td>
<td>Blur</td>
<td>● ● ● ● ●</td>
<td>Clearer = greater</td>
</tr>
</tbody>
</table>
What does this tell us?
Barcharts, scatterplots, and line charts are really effective for quantitative data.
(and for statistical distributions)
Tukey Box Plots
Outliers

Largest < Q3 + 1.5 IQR

Largest < Q3

Median

Smallest > Q1

Smallest > Q1 - 1.5 IQR
Tufte’s Chart Principles

Edward Tufte
Tufte’s Chart Principles

DO NOT LIE!
Tufte’s Chart Principles

**DO NOT LIE!**

Maximize Data-Ink Ratio
Minimize Chart Junk
Subsea Oil Collection

- Avg circa 2,000 bbl per day
- Total of 13,500 bbis collected

Cumulative Oil Collected, bbis

- Cumulative Oil
- Subsea Area
“Cumulative”
Subsea Oil Collection

- Avg: circa 2,000 bbls per day
- Total: 13,600 bbls collected

Cumulative Oil Collected, bbls

Oil Collected Daily

Barrels

The Month of May, 2010

http://www.perceptualedge.com/blog/?p=790
PET PEEVE #208:

GEOGRAPHIC PROFILE MAPS WHICH ARE BASICALLY JUST POPULATION MAPS

http://xkcd.com/1138/
Tufte’s Chart Principles

DO NOT LIE!
Maximize Data-Ink Ratio
Minimize Chart Junk
DIAMONDS WERE A GIRL’S BEST FRIEND
Average price of a one-carat D-flawless

$60,000
$50,000
$40,000
$30,000
$20,000


TIME Chart by Nigel Holmes
Source: The Diamond Registry
Please...
No pie charts.
No 2.5D charts.
PLEASE DON’T EVER DO THIS!
But otherwise...
Barcharts, scatterplots, and line charts are really effective for quantitative data.
Anyone else bored by my color choices?
In fact, grayscale can be risky...
In fact, grayscale can be risky...
Color is Powerful
Color

Call attention to information
Increase appeal
Increase memorability
Another dimension to work with
Have you heard of RGB?

Additive color model: colors create by mixing red, green, blue light
We see in RGB, but we don’t interpret in RGB...
HSV Color Model

Source: color picker in Affinity Designer
Hue

Post & Greene, 1986
**Actual color names if you’re a girl ...**

- red
- magenta
- purple
- blue
- pink
- hot pink
- salmon
- orange
- yellow
- light green
- lime green
- neon green
- green
- aqua
- teal
- blue

**Actual color names if you’re a guy ...**

- red
- magenta
- purple
- blue
- pink
- hot pink
- salmon
- orange
- yellow
- light green
- lime green
- neon green
- green
- aqua
- teal
- blue

---

2010/05/03/color-survey-results/
Hue and Colorblindness

10% of males and 1% of females are Red-Green Colorblind
NOAA's Latest High Resolution Weather Model is Released

Can you order these (low→hi)?
Colormaps

- Categorical limits: noncontiguous
  - 6-12 bins hue/color
  - Far fewer if colorblind
    - 3-4 bins luminance, saturation
- Size heavily affects salience
  - Use high saturation for small regions, low saturation for large

---

http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html via Munzner
Color Brewer for Picking Color Scales

Overview
Zoom+Filter
Details on Demand

Shneiderman Mantra
(Information-Seeking Mantra)

https://www.mat.ucsb.edu/g.legrady/academic/courses/11w259/schneiderman.pdf
NameVoyager: Explore baby names and name trends letter by letter

Looking for the perfect baby name? **Sign up for free** to receive access to our expert tools!

Baby Name > Cha

Names starting with 'CHA' per million babies

Charles

Charlie

Chad

Charlotte

Rank in 2011: 27

Click a name graph to view that name. Double-click to read more about it.

http://www.babynamewizard.com/voyager
Where to learn more?
CS 7450
Information Visualization
Every Fall
Visualization @GeorgiaTech
vis.gatech.edu
How to Make Good Charts

• Edward Tufte’s One-Day Workshop
  – http://www.edwardtufte.com/tufte/courses

• Edward Tufte, *Visual Display of Quantitative Information*
  – http://www.edwardtufte.com/tufte/books_vdqqi

• Stephen Few, *Show Me the Numbers: Designing Tables and Graphs to Enlighten*
Visualization Theory “Books”

• Tamara Munzner VIS Tutorial and Book
• Colin Ware, *Information Visualization: Perception for Design*
• Stephen Few, *Now You See It*
• Edward Tufte, *Envisioning Information*
• Edward Tufte, *Visual Explanations*
• Edward Tufte, *Beautiful Evidence*
• Tamara Munzner, *Visualization Analysis & Design*
Perception and Color Websites

• Chris Healy, NC State
  – http://www.csc.ncsu.edu/faculty/healey/PP/index.html

• Color Brewer
  – http://colorbrewer2.org/

• Maureen C. Stone (Color Links, Blog, Workshops)

• Subtleties of Color by Robert Simmon of NASA
Visualization Blogs

• Flowing Data by Nathan Yau

• Information Aesthetics by Andrew Vande Moere

• Information is Beautiful by David McCandless

• Visual.ly Blog

• Indexed Comic by Jessica Hagy
Infographics

Visual.ly/view

(wtfviz.net)
Thanks!

Chad Stolper
chadstolper@gatech.edu
Questions?

Chad Stolper
chadstolper@gatech.edu