CS2D7 Data Visualisation
Dr Claire Rocks & Dr Richard Kirk

July 2022
Lecture 4

What makes an Effective visualization?
Design Principles

• Where the process offers efficiency, design principles ensure effectiveness
• Most choices are relatively clear cut
• There are few universal rules
• But there are evidence-based useful suggestions about what you should and should not do
The Visualization Wheel, Cairo (2013)

Figure 3.11 Different professional backgrounds, different ways of facing projects.
Kirk’s Three Principles of Good Visual Design

1. Good data visualization is TRUSTWORTHY
2. Good data visualization is ACCESIBLE
3. Good data visualization is ELEGANT

Good data visualization is TRUSTWORTHY

- Without trust there is no opportunity to facilitate understanding
- Trust vs truth
- Multiple truths
- Trust is nurtured through accuracy and transparency
- Trust is fragile
Edward Tufte – Graphical Integrity (1)

• Visual representation consistent with the numerical representation
• What is the visual representation of the data? As physically measured or the perceived visual effect?
• Different people see the same areas differently, perceptions change with experience and perceptions are context dependent
• One way to overcome this is to use a table and tables usually outperform graphics on reporting on small data sets

The visual display of quantitative information / Edward R. Tufte (2001)
Edward Tufte – Graphical Integrity (2)

• Best we can hope for is uniformity in the graphics
  • The representation of numbers as physical measured on the surface of the graphic itself should be directly proportional to the numerical quantities represented
  • Clear, detailed and thorough labelling should be used to graphical distortion and ambiguity.

• Lie Factor = size of effect shown in graphic/size of effect in data

The visual display of quantitative information / Edward R. Tufte (2001)
Edward Tufte – Graphical Integrity (3)

• Show data variation not design variation e.g. scale moving in regular intervals
• The number of information carrying (variable) dimensions depicted should not exceed the number of dimensions in the data
• Graphics must not quote data out of context

The visual display of quantitative information / Edward R. Tufte (2001)
| Formulating your brief | • Some contexts lend themselves to exploiting the emotive qualities of the subject – others do not.  
  • You will need to make a judgement |
|------------------------|---------------------------------------------------------------------------------------------|
| Working with data      | • Trust is most at stake here  
  • You are responsible for being faithful to the data  
  • You need to be careful and transparent in how you handle the data |
| Developing your design | • Are you using a chart appropriately? Are pie charts representing >100%? Have you truncated the y axis? Are you using an appropriate aspect ratio? Or map projection?  
  • Are you scaling marks appropriately?  
  • Have you included all necessary annotations?  
  • Are the functional experiences reliable?  
  • Does the tone impact trust? Think colours, typeface and fontsize |
Gun deaths in Florida

Number of murders committed using firearms

2005
Florida enacted its ‘Stand Your Ground’ law

Source: Florida Department of Law Enforcement
C. Chan 16/02/2014
Kirk’s Three Principles of Good Visual Design

1. Good data visualization is TRUSTWORTHY
2. **Good data visualization is ACCESSIBLE**
3. Good data visualization is ELEGANT

Good data visualization is ACCESSIBLE

- Does your audience have access to useful understanding?
- Is it relevant?
- Is the design suitable?
- What effort for what understanding (minimum friction)?
  - What does the audience already know?
  - Is the subject complicated, complex or simple?
  - Do you simplify or clarify?
  - Is the chart complex?
  - How long does the audience have to process understanding?
  - How willing is the audience to engage?
Colour

• When used sparingly, this is one of the most powerful tools you have to draw attention
Colour

• RGB (screen) or CYMK (print) or HSL
• Legibility – making sure the differences between and the associations with colours used are readable and meaningful
• Quantitative data
  • Variation in lightness of hue is typically used for differentiating quantities
    • What range? Observed or observable?
  • Absolute judgements difficult with a continuous scale
  • Can be divided up into discrete classes
    • Are equal intervals the most appropriate?
### Country Level Sales Rank Top 5 Drugs

Rainbow distribution in color indicates sales rank in a given country from #1 (red) to #10 or higher (dark purple).

<table>
<thead>
<tr>
<th>Country</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUS</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>BRA</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>CAN</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>CHI</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>FRA</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>GER</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>IND</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>ITA</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>MEX</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>RUS</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>SPA</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>TUR</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>UK</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>US</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

### Top 5 Drugs: Country-level Sales Rank

<table>
<thead>
<tr>
<th>Country</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Brazil</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Canada</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>China</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>France</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Germany</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>India</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Italy</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Mexico</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Russia</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Spain</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Turkey</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>United States</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>
Colour

• Qualitative Data
  • Not aiming to imply order (nominal) or magnitude
  • Just want to differentiate categories
  • Variation in hue is typically the colour dimension to use
  • Once you exceed 12 categories, there are no longer sufficiently different hues available
    • Think about using filters, modifying the number of categories or use of texture
  • Ordinal colouring approaches align more with quantitative approaches
Colour

• Leverage colour selectively – this should be intentional – tempting to let the tool do it for you
• Use colour consistently
• Be thoughtful of the tone that colour conveys
  • Colour evokes emotion
• Be aware of cultural colour connotations
• Design with colourblind in mind
Design with colourblind in mind

• Roughly 8% of men and 0.5% women are colourblind
• Most frequently means difficulty in distinguishing between red and green
  • Red and green have meaning to us
• A number of sites and applications that let you see your visualisation through colourblind eyes
  • Vischeck.com
  • Checkmycolours.com
Colour – influencing factors

• Medium
• Purpose
• Accessible design
Kirk’s Three Principles of Good Visual Design

1. Good data visualization is TRUSTWORTHY
2. Good data visualization is ACCESSIBLE
3. Good data visualization is ELEGANT

Good data visualization is ELEGANT

• Eliminate the arbitrary – every decision you make, every dot, every pixel should be justifiable
• This does not mean minimal
• Agonise over the smallest of details
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)

- Demonstration of results
- Content expertise
- Local knowledge
- National reputation
- Affordability of services
- Previous work together
- Colleague recommendation

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.

% selecting given attribute

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)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>% selecting given attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstration of results</td>
<td><img src="demonstration_of_results_bar_graph" alt="Diagram" /></td>
</tr>
<tr>
<td>Content expertise</td>
<td><img src="content_expertise_bar_graph" alt="Diagram" /></td>
</tr>
<tr>
<td>Local knowledge</td>
<td><img src="local_knowledge_bar_graph" alt="Diagram" /></td>
</tr>
<tr>
<td>National reputation</td>
<td><img src="national_reputation_bar_graph" alt="Diagram" /></td>
</tr>
<tr>
<td>Affordability of services</td>
<td><img src="affordability_of_services_bar_graph" alt="Diagram" /></td>
</tr>
<tr>
<td>Previous work together</td>
<td><img src="previous_work_together_bar_graph" alt="Diagram" /></td>
</tr>
<tr>
<td>Colleague recommendation</td>
<td><img src="colleague_recommendation_bar_graph" alt="Diagram" /></td>
</tr>
</tbody>
</table>

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.

**FIGURE 3.14** Revamped summary of survey feedback
Good data visualization is ELEGANT

• Eliminate the arbitrary – every decision you make, every dot, every pixel should be justifiable
• This does not mean minimal
• Agonise over the smallest of details
• Decoration should be additive not negative
Edward Tufte – Graphical Excellence

• Graphical excellence is the well-designed presentation of interesting data -- a matter of *substance*, of *statistics*, and of *design*.

• Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency.

• 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.

• Graphical excellence is nearly always multivariate.

• Graphical excellence requires telling the truth about the data.
Maximise the data-ink ratio, eliminating any non-essential elements.

Primary importance is given to the exhibition of the data, not to the visual.
**Smooth Sales**
The online men's razor and blade market in the U.S. is growing, threatening traditional retail sales.

$111M  
2013

$189M  
2014

$141M  
2015, through May

---

**Dark Outlook**

Rising chocolate prices in China...
Retail prices for chocolate confectionery per kilogram

<table>
<thead>
<tr>
<th>Year</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>$9.40</td>
</tr>
<tr>
<td>'10</td>
<td>$9.90</td>
</tr>
<tr>
<td>'11</td>
<td>$11.10</td>
</tr>
<tr>
<td>'12</td>
<td>$11.90</td>
</tr>
<tr>
<td>'13</td>
<td>$12.60</td>
</tr>
<tr>
<td>'14</td>
<td>$13.30</td>
</tr>
</tbody>
</table>

...are weighing on cocoa demand...

Tons of cocoa beans produced in Asia, change from year earlier

<table>
<thead>
<tr>
<th>Year</th>
<th>Tons</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>10</td>
</tr>
<tr>
<td>2013</td>
<td>10</td>
</tr>
<tr>
<td>2014</td>
<td>10</td>
</tr>
<tr>
<td>2015</td>
<td>10</td>
</tr>
</tbody>
</table>

...are pushing down prices...

$c3,900 a metric ton

Cocoa futures

2013  | 3,000
2014  | 2,500
2015  | 2,000

...and prompting investors to pare back on bullish bets.

60,000 contracts

Net bullish bets on cocoa futures

Sources: Euromonitor (retail prices), Cocoa Association of Asia (beans processed), FactSet Futures; U.S. Commodity Futures Trading Commission (contracts)

---

*Source: Skim Intelligence*  
*Photo: F. Martin Rahman/The Wall Street Journal; styling by Anna Cardenas*
Edward Tufte - Aesthetics

• Excellence in visuals is attained when the complexity of the data matches the simplicity of the design
5 criteria (Mazza)

• **Functionality** – does the visualisation fulfil all of the requirements of the specification?

• **Effectiveness** – does the visualisation provide users with a better knowledge of the data?

• **Efficiency** - does the visualisation provide users with information more rapidly?

• **Usability** – is interaction with the visualisation simple and intuitive

• **Usefulness** – does the visualisation serve a purpose?
4 criteria (Beautiful Visualisation (2010))

• **Informative** – successfully conveys information
• **Efficient** – simple, focussed, clear, straightforward
• **Aesthetic** – Axes, layout, shape, colours, lines + typography used appropriately
• **Novel** – creative and interesting

Ben Fry - Principles

• Each project has unique requirements
  • A visualization should convey the unique properties of what it represents. The goal is to expose the fascinating aspect and make it self-evident.

• Avoid the all you can eat buffet
  • Often inclusion of overly specific details causes the viewer to miss what’s most important. Focussing on the question helps define minimum requirements

• Know your audience
  • What are their goals when approaching a visualization? What will they learn? Why are you doing it? Is is accessible to the audience?
Size Matters

- Relative size denotes relative importance
Position can matter

https://vanseodesign.com/web-design/3-design-layouts/
More flexible than it seems

• Choosing the right graphic form for each task is not easy
• Tempting to propose rules
• Reality is complex
• Ask yourself – what is the audience likely to do with this tool?
Things not to do

• Think carefully about pie charts and donut charts: we are not good at ascribing quantitative value to 2-D space
• Do not use 3D – the only exception is when you are plotting 3-dimensions and even then take care
• Secondary y-axis are not a good idea
Lab

• Text Analysis

This afternoon

• Evaluating Visualisations (workshop)