Data Visualization (DSC 530/CIS 568)

Data & Tasks

Dr. David Koop
JavaScript in one slide

- Interpreted and Dynamically-typed Programming Language
- Statements end with semi-colons, normal blocking with brackets
- Variables: var a = 0; let b = 2;
- Operators: +, -, *, /, [ ]
- Control Statements: if (<expr>) {...} else {...}, switch
- Loops: for, while, do-while
- Arrays: var a = [1,2,3]; a[99] = 100; console.log(a.length);
- Functions: function myFunction(a,b) { return a + b; } 
- Objects: var obj; obj.x = 3; obj.y = 5;
  - Prototypes for instance functions
- Comments are /* Comment */ or // Single-line Comment
Including JavaScript in HTML

• Use the script tag

• Can either inline JavaScript or load it from an external file

  - `<script type="text/javascript">
      a = 5, b = 8;
      c = a * b + b - a;
    </script>
    <script type="text/javascript" src="script.js"/>

• The order the javascript is in is the order it is executed

• Example: in the above, script.js can access the variables `a`, `b`, and `c`
JavaScript Features

• Any object can serve as an associative array
  
  states = {"AZ": "Arizona", "MA": "Massachusetts"};
  console.log("MA is" + states["MA"]);

• Array functions: map, filter, reduce, forEach
  
  Object.keys(states).filter(d => d.startsWith("A"));

• Function chaining is common (sometimes the original object is returned, others another object is returned)
  
  $("#myElt").css("color", "blue").height(200).width(320)

• Closures are functions that "remember their environments" [MDN]
  
  function makeAdder(x) {
      return function(y) {
          return x + y;
      };
  }

  var add5 = makeAdder(5);
Using Array Functions

- var a = [2, 4, 7, 11, 22, 84];

- **Named function:**
  ```javascript
  function isEven(d) {
    return (d % 2 == 0);
  }
  
  a.filter(isEven);
  ```

- **Anonymous function**
  ```javascript
  a.filter(function(d) { return (d % 2 == 0); });
  ```

- **Arrow function**
  ```javascript
  a.filter(d => (d % 2 == 0));
  ```
Assignment 1

- Due Monday, Feb. 11
- Questions?
Data Types

• Positions:
  - A position is a location in space (usually 2D or 3D)
  - May be subject to projections
  - e.g. cities on a map, a sampled region in an CT scan

• Grids:
  - A grid specifies how data is sampled both geometrically and topologically
  - e.g. how CT scan data is stored
Positions and Grids
Dataset Types

- **Tables**
  - Attributes (columns)
  - Items (rows)
  - Cell containing value

- **Networks**
  - Link
  - Node (item)

- **Fields (Continuous)**
  - Grid of positions
  - Cell

- **Geometry (Spatial)**
  - Position

- **Multidimensional Table**
  - Key 1
  - Key 2
  - Value in cell

- **Trees**
  - Value in cell

[Munzner (ill. Maguire), 2014]
# Tables

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>S</th>
<th>T</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order ID</td>
<td>Order Date</td>
<td>Order Priority</td>
<td>Product Container</td>
<td>Product Base Margin</td>
<td>Ship Date</td>
</tr>
<tr>
<td>3</td>
<td>10/14/06</td>
<td>5-Low</td>
<td>Large Box</td>
<td>0.8</td>
<td>10/21/06</td>
</tr>
<tr>
<td>6</td>
<td>2/21/08</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.55</td>
<td>2/22/08</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Small Pack</td>
<td>0.79</td>
<td>7/17/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Jumbo Box</td>
<td>0.69</td>
<td>7/17/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.65</td>
<td>7/18/07</td>
</tr>
<tr>
<td>35</td>
<td>10/23/07</td>
<td>4-Not Specified</td>
<td>Wrap Bag</td>
<td>0.52</td>
<td>10/24/07</td>
</tr>
<tr>
<td>35</td>
<td>10/23/07</td>
<td>4-Not Specified</td>
<td>Small Box</td>
<td>0.58</td>
<td>10/25/07</td>
</tr>
<tr>
<td>36</td>
<td>11/3/07</td>
<td>1-Urgent</td>
<td>Small Box</td>
<td>0.55</td>
<td>11/3/07</td>
</tr>
<tr>
<td>65</td>
<td>3/18/07</td>
<td>1-Urgent</td>
<td>Small Pack</td>
<td>0.49</td>
<td>3/19/07</td>
</tr>
<tr>
<td>66</td>
<td>1/28/05</td>
<td>5-Low</td>
<td>Wrap Bag</td>
<td>0.56</td>
<td>1/20/05</td>
</tr>
<tr>
<td>69</td>
<td>4/17/05</td>
<td>2-High</td>
<td>Small Pack</td>
<td>0.6</td>
<td>6/6/05</td>
</tr>
<tr>
<td>69</td>
<td>4/17/05</td>
<td>2-High</td>
<td>Wrap Bag</td>
<td>0.6</td>
<td>6/6/05</td>
</tr>
<tr>
<td>70</td>
<td>12/18/06</td>
<td>5-Low</td>
<td>Small Box</td>
<td>0.59</td>
<td>12/23/06</td>
</tr>
<tr>
<td>70</td>
<td>12/18/06</td>
<td>5-Low</td>
<td>Wrap Bag</td>
<td>0.82</td>
<td>12/23/06</td>
</tr>
<tr>
<td>96</td>
<td>4/17/05</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.55</td>
<td>4/19/05</td>
</tr>
<tr>
<td>97</td>
<td>1/29/06</td>
<td>3-Medium</td>
<td>Small Box</td>
<td>0.38</td>
<td>1/30/06</td>
</tr>
<tr>
<td>129</td>
<td>11/19/08</td>
<td>5-Low</td>
<td>Small Box</td>
<td>0.37</td>
<td>11/28/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.37</td>
<td>5/9/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.38</td>
<td>5/10/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.6</td>
<td>5/11/08</td>
</tr>
<tr>
<td>132</td>
<td>6/11/06</td>
<td>3-Medium</td>
<td>Medium Box</td>
<td>0.6</td>
<td>6/12/06</td>
</tr>
<tr>
<td>132</td>
<td>6/11/06</td>
<td>3-Medium</td>
<td>Jumbo Box</td>
<td>0.69</td>
<td>6/14/06</td>
</tr>
<tr>
<td>134</td>
<td>5/1/08</td>
<td>4-Not Specified</td>
<td>Large Box</td>
<td>0.82</td>
<td>5/3/08</td>
</tr>
<tr>
<td>135</td>
<td>10/21/07</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.64</td>
<td>10/23/07</td>
</tr>
<tr>
<td>166</td>
<td>9/12/07</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.55</td>
<td>9/14/07</td>
</tr>
<tr>
<td>193</td>
<td>8/8/06</td>
<td>1-Urgent</td>
<td>Medium Box</td>
<td>0.57</td>
<td>8/10/06</td>
</tr>
<tr>
<td>194</td>
<td>4/5/08</td>
<td>3-Medium</td>
<td>Wrap Bag</td>
<td>0.42</td>
<td>4/7/08</td>
</tr>
</tbody>
</table>

**Notes:**
- **item**: The specific product or order priority.
- **attribute**: The category or classification of the item.
- **cell**: The data entry point for each attribute.

*Source: D. Koop, DSC 530, Spring 2019*
Tables

Flat

- Data organized by rows & columns
  - row ~ item (usually)
  - column ~ attribute
  - label ~ attribute name
- Key: identifies each item (row)
  - Usually unique
  - Allows join of data from 2+ tables
- Compound key: key split among multiple columns, e.g. (state, year) for population

Multidimensional

- Split compound key
  - e.g. a data cube with (state, year)

[Munzner (ill. Maguire), 2014]
Table Visualizations

[Table Diagram]

[M. Bostock, 2011]
Networks

- Why networks instead of graphs?
- Tables can represent networks
  - Many-many relationships
  - Also can be stored as specific graph databases or files
Networks
Networks

[D. Koop, DSC 530, Spring 2019]

[Holten & van Wijk, 2009]
Fields

Scalar Fields

Vector Fields

Tensor Fields

Each point in space has an associated...
Fields

Scalar Fields
(Order-0 Tensor Fields)

Vector Fields
(Order-1 Tensor Fields)

Tensor Fields
(Order-2+)

Each point in space has an associated...

Scalar

\[ s_0 \]

Vector

\[
\begin{bmatrix}
  v_0 \\
  v_1 \\
  v_2 
\end{bmatrix}
\]

Tensor

\[
\begin{bmatrix}
  \sigma_{00} & \sigma_{01} & \sigma_{02} \\
  \sigma_{10} & \sigma_{11} & \sigma_{12} \\
  \sigma_{20} & \sigma_{21} & \sigma_{22} 
\end{bmatrix}
\]
Fields

- Difference between **continuous** and **discrete** values
- Examples: temperature, pressure, density
- **Grids** necessary to sample continuous data:

![Grid Types]

- uniform
- rectilinear
- structured
- unstructured

[Weiskopf, Machiraju, Möller]

- **Interpolation**: “how to show values between the sampled points in ways that do not mislead”
Spatial Data Example: MRI

[Image of an MRI machine and a grid-like 3D structure]

[Image of a series of MRI scans]

[Via Levine, 2014]
Scivis and Infovis

- Two subfields of visualization
- **Scivis** deals with data where the spatial position is given with data
  - Usually continuous data
  - Often displaying physical phenomena
  - Techniques like isosurfacing, volume rendering, vector field visualization
- In **Infovis**, the data has no set spatial representation, designer chooses how to visually represent data
SciVis

[Google Image Search for "scientific visualization", 2017]
InfoVis

[Google Image Search for "information visualization", 2017]
Sets & Lists

# OF UNIQUE WORDS USED WITHIN ARTIST'S FIRST 35,000 LYRICS

Notes/sources:

(1)(2) I used the first 5,000 words for 7 of Shakespeare’s works: Hamlet, Romeo and Juliet, Othello, Macbeth, As You Like It, Winter’s Tale, and Troilus and Cressida. For Melville, I used the first 35,000 words of Moby Dick.

All lyrics are provided by Rap Genius, but are only current to 2012. My lack of recent data prevented me from using quite a few current artists.

This data viz uses code by Amelia Bellamy-Royds’s in this jsfiddle.

[Daniels, http://experiments.undercurrent.com]
Attribute Types

- Categorical
  - Symbols:
    - 
    - 
    - 
    - 

- Ordered
  - Subtypes:
    - Ordinal
    - Quantitative
## Categorial, Ordinal, and Quantitative

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>S</th>
<th>T</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order ID</td>
<td>Order Date</td>
<td>Order Priority</td>
<td>Product Container</td>
<td>Product Base Margin</td>
<td>Ship Date</td>
</tr>
<tr>
<td>3</td>
<td>10/14/06</td>
<td>5-Low</td>
<td>Large Box</td>
<td>0.8</td>
<td>10/21/06</td>
</tr>
<tr>
<td>6</td>
<td>2/21/08</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.55</td>
<td>2/22/08</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Small Pack</td>
<td>0.79</td>
<td>7/17/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Jumbo Box</td>
<td>0.72</td>
<td>7/17/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.6</td>
<td>7/18/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.65</td>
<td>7/18/07</td>
</tr>
<tr>
<td>35</td>
<td>10/23/07</td>
<td>4-Not Specified</td>
<td>Wrap Bag</td>
<td>0.52</td>
<td>10/24/07</td>
</tr>
<tr>
<td>35</td>
<td>10/23/07</td>
<td>4-Not Specified</td>
<td>Small Box</td>
<td>0.58</td>
<td>10/25/07</td>
</tr>
<tr>
<td>36</td>
<td>11/3/07</td>
<td>1-Urgent</td>
<td>Small Box</td>
<td>0.55</td>
<td>11/3/07</td>
</tr>
<tr>
<td>65</td>
<td>3/18/07</td>
<td>1-Urgent</td>
<td>Small Pack</td>
<td>0.49</td>
<td>3/19/07</td>
</tr>
<tr>
<td>66</td>
<td>1/20/05</td>
<td>5-Low</td>
<td>Wrap Bag</td>
<td>0.56</td>
<td>1/20/05</td>
</tr>
<tr>
<td>69</td>
<td>6/4/05</td>
<td>4-Not Specified</td>
<td>Small Box</td>
<td>0.44</td>
<td>6/6/05</td>
</tr>
<tr>
<td>69</td>
<td>6/4/05</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.6</td>
<td>6/6/05</td>
</tr>
<tr>
<td>70</td>
<td>12/18/06</td>
<td>5-Low</td>
<td>Large Box</td>
<td>0.59</td>
<td>12/23/06</td>
</tr>
<tr>
<td>70</td>
<td>12/18/06</td>
<td>5-Low</td>
<td>Large Box</td>
<td>0.82</td>
<td>12/23/06</td>
</tr>
<tr>
<td>96</td>
<td>4/17/05</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.55</td>
<td>4/19/05</td>
</tr>
<tr>
<td>97</td>
<td>1/29/06</td>
<td>3-Medium</td>
<td>Medium Box</td>
<td>0.38</td>
<td>1/30/06</td>
</tr>
<tr>
<td>129</td>
<td>11/19/08</td>
<td>5-Low</td>
<td>Medium Box</td>
<td>0.37</td>
<td>11/28/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.37</td>
<td>5/9/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.38</td>
<td>5/10/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.6</td>
<td>5/11/08</td>
</tr>
<tr>
<td>132</td>
<td>6/11/06</td>
<td>3-Medium</td>
<td>Medium Box</td>
<td>0.6</td>
<td>6/12/06</td>
</tr>
<tr>
<td>132</td>
<td>6/11/06</td>
<td>3-Medium</td>
<td>Jumbo Box</td>
<td>0.69</td>
<td>6/14/06</td>
</tr>
<tr>
<td>134</td>
<td>5/1/08</td>
<td>4-Not Specified</td>
<td>Large Box</td>
<td>0.82</td>
<td>5/3/08</td>
</tr>
<tr>
<td>135</td>
<td>10/21/07</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.64</td>
<td>10/23/07</td>
</tr>
<tr>
<td>166</td>
<td>9/12/07</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.55</td>
<td>9/14/07</td>
</tr>
<tr>
<td>193</td>
<td>8/8/06</td>
<td>1-Urgent</td>
<td>Medium Box</td>
<td>0.57</td>
<td>8/10/06</td>
</tr>
<tr>
<td>194</td>
<td>4/5/08</td>
<td>3-Medium</td>
<td>Wrap Bag</td>
<td>0.42</td>
<td>4/7/08</td>
</tr>
</tbody>
</table>
### Categorial, Ordinal, and Quantitative

The following is a table showing data on order IDs, order dates, order priority, product container, product base margin, and ship dates. The data is labeled as quantitative, ordinal, and categorical.

<table>
<thead>
<tr>
<th>Order ID</th>
<th>Order Date</th>
<th>Order Priority</th>
<th>Product Container</th>
<th>Product Base Margin</th>
<th>Ship Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>10/14/06</td>
<td>5-Low</td>
<td>Large Box</td>
<td>0.8</td>
<td>10/21/06</td>
</tr>
<tr>
<td>6</td>
<td>2/21/08</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.55</td>
<td>2/22/08</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Small Pack</td>
<td>0.79</td>
<td>7/17/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Jumbo Box</td>
<td>0.72</td>
<td>7/17/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.6</td>
<td>7/18/07</td>
</tr>
<tr>
<td>32</td>
<td>7/16/07</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.65</td>
<td>7/18/07</td>
</tr>
<tr>
<td>35</td>
<td>10/23/07</td>
<td>4-Not Specified</td>
<td>Wrap Bag</td>
<td>0.52</td>
<td>10/24/07</td>
</tr>
<tr>
<td>35</td>
<td>10/23/07</td>
<td>4-Not Specified</td>
<td>Small Box</td>
<td>0.58</td>
<td>10/25/07</td>
</tr>
<tr>
<td>36</td>
<td>11/3/07</td>
<td>1-Urgent</td>
<td>Small Box</td>
<td>0.55</td>
<td>11/3/07</td>
</tr>
<tr>
<td>65</td>
<td>3/18/07</td>
<td>1-Urgent</td>
<td>Small Pack</td>
<td>0.49</td>
<td>3/19/07</td>
</tr>
<tr>
<td>66</td>
<td>1/20/05</td>
<td>5-Low</td>
<td>Wrap Bag</td>
<td>0.56</td>
<td>1/20/05</td>
</tr>
<tr>
<td>69</td>
<td>6/4/05</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.44</td>
<td>6/6/05</td>
</tr>
<tr>
<td>69</td>
<td>6/4/05</td>
<td>4-Not Specified</td>
<td></td>
<td>0.6</td>
<td>6/6/05</td>
</tr>
<tr>
<td>70</td>
<td>12/18/06</td>
<td>5-Low</td>
<td></td>
<td>0.59</td>
<td>12/23/06</td>
</tr>
<tr>
<td>70</td>
<td>12/18/06</td>
<td>5-Low</td>
<td></td>
<td>0.82</td>
<td>12/23/06</td>
</tr>
<tr>
<td>96</td>
<td>4/17/05</td>
<td>2-High</td>
<td></td>
<td>0.55</td>
<td>4/19/05</td>
</tr>
<tr>
<td>97</td>
<td>1/29/06</td>
<td>3-Medium</td>
<td></td>
<td>0.38</td>
<td>1/30/06</td>
</tr>
<tr>
<td>129</td>
<td>11/19/08</td>
<td>5-Low</td>
<td></td>
<td>0.37</td>
<td>11/28/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.37</td>
<td>5/9/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Medium Box</td>
<td>0.38</td>
<td>5/10/08</td>
</tr>
<tr>
<td>130</td>
<td>5/8/08</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.6</td>
<td>5/11/08</td>
</tr>
<tr>
<td>132</td>
<td>6/11/06</td>
<td>3-Medium</td>
<td>Medium Box</td>
<td>0.6</td>
<td>6/12/08</td>
</tr>
<tr>
<td>132</td>
<td>6/11/06</td>
<td>3-Medium</td>
<td>Jumbo Box</td>
<td>0.69</td>
<td>6/14/08</td>
</tr>
<tr>
<td>134</td>
<td>5/1/08</td>
<td>4-Not Specified</td>
<td>Large Box</td>
<td>0.82</td>
<td>5/3/08</td>
</tr>
<tr>
<td>135</td>
<td>10/21/07</td>
<td>4-Not Specified</td>
<td>Small Pack</td>
<td>0.64</td>
<td>10/23/07</td>
</tr>
<tr>
<td>166</td>
<td>9/12/07</td>
<td>2-High</td>
<td>Small Box</td>
<td>0.55</td>
<td>9/14/07</td>
</tr>
<tr>
<td>193</td>
<td>8/8/06</td>
<td>1-Urgent</td>
<td>Medium Box</td>
<td>0.57</td>
<td>8/10/06</td>
</tr>
<tr>
<td>194</td>
<td>4/5/08</td>
<td>3-Medium</td>
<td>Wrap Bag</td>
<td>0.42</td>
<td>4/7/08</td>
</tr>
</tbody>
</table>
Data Model vs. Conceptual Model

• Data Model: raw data that has a specific data type (e.g. floats):
  - Temperature Example: [32.5, 54.0, -17.3] (floats)

• Conceptual Model: how we think about the data
  - Includes semantics, reasoning
  - Temperature Example:
    • Quantitative: [32.50, 54.00, -17.30]
Data Model vs. Conceptual Model

• Data Model: raw data that has a specific data type (e.g. floats):
  - Temperature Example: [32.5, 54.0, -17.3] (floats)

• Conceptual Model: how we think about the data
  - Includes semantics, reasoning
  - Temperature Example:
    • Quantitative: [32.50, 54.00, -17.30]
    • Ordered: [warm, hot, cold]
Data Model vs. Conceptual Model

• Data Model: raw data that has a specific data type (e.g. floats):
  - Temperature Example: [32.5, 54.0, -17.3] (floats)

• Conceptual Model: how we think about the data
  - Includes semantics, reasoning
  - Temperature Example:
    • Quantitative: [32.50, 54.00, -17.30]
    • Ordered: [warm, hot, cold]
    • Categorical: [not burned, burned, not burned]

[via A. Lex, 2015]
Ordering Direction

- **Sequential**
- **Diverging**
- **Cyclic**

[Munzner (ill. Maguire), 2014]
Sequential and Diverging Data

• Sequential: homogenous range from a minimum to a maximum
  - Examples: Land elevations, ocean depths

• Diverging: can be deconstructed into two sequences pointing in opposite directions
  - Has a zero point (not necessary 0)
  - Example: Map of both land elevation and ocean depth

[Rogowitz & Treinish, 1998]
Cyclic Data

3.1. Mathematical description and types of spirals

A spiral is easy to describe and understand in polar coordinates, i.e. in the form \( r = f(\phi) \). The distinctive feature of a spiral is that \( f \) is a monotone function. In this work we assume a spiral is described by:

- **Archimedes’ spiral** has the form \( r = a\phi \). It has the special property that a ray emanating from the origin crosses two consecutive arcs of the spiral in a constant distance.

- **The Hyperbolic spiral** has the form \( r = \frac{a}{\phi} \). It is the inverse of Archimedes’ spiral with respect to the origin.

- More generally, spirals of the form \( r = a\phi^k \) are called Archimedean spirals.

- The **logarithmic spiral** has the form \( r = e^{k\phi} \). It has the special property that all arcs cut a ray emanating from the origin under the same angle.

For the visualization of time-dependent data, Archimedes’ spiral seems to be the most appropriate. In most applications data from different periods are equally important. This should be reflected visually in that the distance to other periods is always the same.

3.2. Mapping data to the spiral

In general, markers, bars, and line elements can be used to visualize time-series data similar to standard point, bar, and line graphs on Spiral Graphs. For instance, quantitative, discrete data can be presented as bars on the spiral or by marks with a corresponding distance to the spiral. However, since the \( x \) and \( y \) coordinate are needed to achieve the general form of the spiral their use is limited for the display of data values. One might consider to map data values to small absolute changes in the radius, i.e. \( \Delta r \). Yet, we have found this way of visualizing to be ineffective. We conclude that the general shape of the spiral should be untouched and other attributes should be used, such as:

- **colour**, 
- **texture**, including line styles and patterns.

Figure 1: Two visualizations of sunshine intensity using about the same screen real estate and the same color coding scheme. In the spiral visualization it is much easier to compare days, to spot cloudy time periods, or to see events like sunrise and sunset.

[Sunlight intensity, Weber et al., 2001]
“Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.”

— T. Munzner
Tasks

• Why? Understand data, but what do I want to do with it?
• Levels: High (Produce/Consume), Mid (Search), Low (Queries)
• Another key concern: Who?
  - Designer <-> User (A spectrum)
  - Complex <-> Easy to Use
  - General <-> Context-Specific
  - Flexible <-> Constrained
  - Varied Data <-> Specific Data
Tasks

What?

Why?

How?

Actions

- Analyze
  - Consume
    - Discover
    - Present
    - Enjoy

- Produce
  - Annotate
  - Record
  - Derive

Search

Target known | Target unknown
Location known | Look up | Browse
Location unknown | Locate | Explore

Query

- Identify
  - Compare
  - Summarize

Targets

- All Data
  - Trends
  - Outliers
  - Features

- Attributes
  - One
    - Distribution
  - Many
    - Dependency
    - Correlation
    - Similarity

- Network Data
  - Topology
    - Paths

- Spatial Data
  - Shape

Why?

How?

[D. Koop, DSC 530, Spring 2019]

[Munzner (ill. Maguire), 2014]
Actions: Analyze

→ Consume
  → Discover
  → Present
  → Enjoy

→ Produce
  → Annotate
  → Record
  → Derive

[Munzner (ill. Maguire), 2014]
Visualization for Consumption

• Discover new knowledge
  - Generate new hypothesis or verify existing one
  - Designer doesn’t know what users need to see
  - "why doesn't dictate how"

• Present known information
  - Presenter already knows what the data says
  - Wants to communicate this to an audience
  - May be static but not limited to that

• Enjoy
  - Similar to discover, but without concrete goals
  - May be enjoyed differently than the original purpose
Explore MTA Fare Data
Present Known Information

M. Stefaner, 2013
Enjoy Visualizations of Names

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 > An| Both □ Boys □ Girls

Names starting with 'AN' per million babies

“[W]e scientists now understand how important emotion is to everyday life, how valuable. Sure, utility and usability are important, but without fun and pleasure, joy and excitement, and yes, anxiety and anger, fear and rage, our lives would be incomplete.” —D. Norman (Emotional Design)
Measuring User Experience in Visualization

• Memorability: Capability of maintaining and retrieving information [J. Brown et al., 1977]

• Engagement: Emotional, cognitive and behavioral connection that exists, at any point in time and possibly over time, between a user and a resource. [S. Attfield et al., 2011]

• Enjoyment: Feeling that causes a person to experience pleasure. Pleasure is recognized with occurrent happiness and excitement, which can be explained in terms of belief, desire, and thought. [W. A. Davis, 1982]
Memorability

Figure 6.2: Policy shifts and interventions to enable wetland practices to accommodate multiple of ecosystem services and human health.

60% of Americans feel Romney performed better than Obama in debates.

Low Quality Description

High Quality Description

Memorable

Forgettable

[Source: M. Borkin et al., InfoVis 2015]
Memorability: Maps instead of Networks

Figure 1: We investigate the memorability of relational data represented with node-link (left-side) and map-based (right-side) visualizations; shown are a node-link and a map-based visualization with 200 nodes and 500 links from the LastFM dataset.

Abstract

We investigate the memorability of data represented in two different visualization designs. In contrast to recent studies that examine which types of visual information make visualizations memorable, we examine the effect of different visualizations on time and accuracy of recall of the displayed data, minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations of the same relational data: node-link diagrams and map-based visualization. We find significant differences in the accuracy of the tasks performed, and these differences persist days after the original exposure to the visualizations. Specifically, participants in the study recalled the data better when exposed to map-based visualizations as opposed to node-link diagrams. We discuss the scope of the study and its limitations, possible implications, and future directions.

1. Introduction

Researchers have long recognized that the visual display of information can be more effective than tables and numeric summaries [Ans73]. We also know that different visual designs offer significantly different reading precision [CM84]. In contrast, we do not understand nearly as well the memorability of the data that underlies the visualization. Is the design of a visualization responsible for how well users will remember its content? In this paper, we present evidence that different visual designs can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of different types of visualizations [BMG10, BARM12, MPWG12, VMTW12, IXTO11, BVB13]. These seminal studies focused on which types of visual information are memorable [BVB13]. To the best of our knowledge, no study has yet been performed to assess long-term memorability of the underlying data represented in these visualizations. In this paper, we focus on two alternative visualizations for relational data. Specifically, we compare node-link visualizations to map-based visualizations. Node-link visualizations date back to 1735 and are a standard way of depicting relational datasets. In node-link diagrams, entities are depicted as points (typically dots or circles) in low-dimensional space, and two related entities are connected with a curve (typically a straight-line segment). Cluster membership is typically indicated by filling each circle with a color that is unique for each cluster.

[B. Saket et al., EuroVis 2015]
Memorability: Maps instead of Networks

Abstract
We investigate the memorability of data represented in two different visualization designs. In contrast to recent studies that examine which types of visual information make visualizations memorable, we examine the effect of different visualizations on time and accuracy of recall of the displayed data, minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations of the same relational data: node-link diagrams and map-based visualization. We find significant differences in the accuracy of the tasks performed, and these differences persist days after the original exposure to the visualizations. Specifically, participants in the study recalled the data better when exposed to map-based visualizations as opposed to node-link diagrams. We discuss the scope of the study and its limitations, possible implications, and future directions.

1. Introduction
Researchers have long recognized that the visual display of information can be more effective than tables and numeric summaries[Ans73]. We also know that different visual designs offer significantly different reading precision[CM84]. In contrast, we do not understand nearly as well the memorability of the data that underlies the visualization. Is the design of a visualization responsible for how well users will remember its content?

In this paper, we present evidence that different visual designs can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of different types of visualizations[BMG10, BARM12, MPWG12, VMTW12, IXTO11, BVB13]. These seminal studies focused on which types of visual information are memorable[BVB13]. To the best of our knowledge, no study has yet been performed to assess long-term memorability of the underlying data represented in these visualizations.

In this paper, we focus on two alternative visualizations for relational data. Specifically, we compare node-link visualizations to map-based visualizations. Node-link visualizations date back to 1735 and are a standard way of depicting relational datasets. In node-link diagrams, entities are depicted as points (typically dots or circles) in low-dimensional space, and two related entities are connected with a curve (typically a straight-line segment). Cluster membership is typically indicated by filling each circle with a color that is unique for each cluster.

Figure 1: We investigate the memorability of relational data represented with node-link (left-side) and map-based (right-side) visualizations; shown are a node-link and a map-based visualization with 200 nodes and 500 links from the LastFM dataset.

[O. Saket et al., EuroVis 2015]
Memorability: Maps instead of Networks

We investigate the memorability of data represented in two different visualization designs. In contrast to recent studies that examine which types of visual information make visualizations memorable, we examine the effect of different visualizations of the same relational data: node-link diagrams and map-based visualization. We find significant differences in the accuracy of the tasks performed, and these differences persist days after the original interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.

In this paper, we focus on two alternative visualizations for standard way of depicting relational datasets. In node-link diagrams, entities are depicted as points (typically dots or circles) connected with a curve (typically a straight-line segment). Cluster summaries [B. Saket et al., EuroVis 2015]. To the best of our knowledge, no study has yet been performed to assess long-term memorability of the data that underlies the visualization.

Researchers have long recognized that the visual display of information can be more effective than tables and numeric summaries. To the best of our knowledge, no one has studied the memorability of data represented in two different visualization designs. In contrast to recent studies that examine which types of visual information make visualizations memorable, we examine the effect of different visualization designs on the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015]; we also know that different visual designs can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015], minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.

In this paper, we present evidence that different visualizations can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015], minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.

In this paper, we present evidence that different visualizations can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015], minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.

In this paper, we present evidence that different visualizations can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015], minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.

In this paper, we present evidence that different visualizations can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015], minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.

In this paper, we present evidence that different visualizations can impact the recall accuracy of the data being visualized. Several recent studies have tested the memorability of display data [B. Saket et al., EuroVis 2015], minutes and days after interaction with the visualizations. In particular, we describe the results of an evaluation comparing the memorability of two different visualizations on time and accuracy of recall of the displayed data.
ISOTYPE Visualizations

- Study [Haroz et al., 2015]
  - Want quick understanding and ease of remembering
  - Does ISOTYPE help?

- Results:
  - Stacked icons allow both length and quantity encoding
  - Icons are more memorable
  - Images that aren't used to show data are distracting

[Image by O. and M. Neurath, Study by S. Haroz et al., 2015]