

Visualizing Multi-dimensional Data



SETH HARRIGAN

**COMPUTER VISUALIZATION
FALL 2008**

Motivation



- Multi-dimensional datasets are common
 - Digital cameras
 - Wall-street stocks
 - Motor vehicles
 - Cellular telephones
- A mixture of interval, ordinal, and nominal data can be visualized well using a table

Motivation



| Features | See What's the Same or What's different | | | | | | |
|--------------------------------|--|--|--|--|-------------------|--|---|
| Lens aperture | F/3.0-5.8 | F/3.0-5.8 | F/3.0-5.8 | F/3.0-5.8 | F/2.8-5.2 | F/2.8-4.4 | F/2.8-4.9 |
| Equivalent 35mm focal length | 38 - 132 mm | 38 - 132 mm | 38 - 132 mm | 38 - 132 mm | 37.5 - 112.5 mm | 38 - 190 mm | 36 - 108 mm |
| Optical zoom | 3.4 x | 3.4 x | 3.4 x | 3.4 x | 3 x | 5 x | 3 x |
| Light sensitivity | ISO 200 ISO 800 ISO auto ISO 400 ISO 80 ISO 1600 ISO 100 | ISO 200 ISO 800 ISO auto ISO 400 ISO 80 ISO 1600 ISO 100 | ISO 200 ISO 800 ISO auto ISO 400 ISO 80 ISO 1600 ISO 100 | ISO 200 ISO 800 ISO auto ISO 400 ISO 80 ISO 1600 ISO 100 | ISO auto (64-800) | ISO 200 ISO 50 ISO 800 ISO auto ISO 400 ISO 1600 ISO 100 | ISO 200 ISO 800 ISO 400 ISO 80 ISO 1600 ISO auto (80-160) ISO 100 |
| Display (projector) technology | TFT active matrix | TFT active matrix | TFT active matrix | TFT active matrix | TFT active matrix | TFT active matrix | TFT active matrix |
| Memory storage capacity | 32 MB | 32 MB | 32 MB | 32 MB | 7 MB | Info unavailable | 32 MB |
| Video input type | Digital camera | Digital camera | Digital camera | Digital camera | Digital camera | Digital camera | Digital camera |
| Flash memory form factor | SD Memory Card | SD Memory Card | SD Memory Card | SD Memory Card | Integrated | Info unavailable | Integrated |
| Digital zoom | 4 x | 4 x | 4 x | 4 x | 4 x | 5 x | 4 x |
| Battery technology | Alkaline | Alkaline | Alkaline | Alkaline | Alkaline | Alkaline | Lithium ion |
| Weight | 5.8 oz | 5.8 oz | 5.8 oz | 5.8 oz | 4.1 oz | 6.1 oz | 4.4 oz |
| Effective sensor | 7,100,000 pixels | 7,100,000 pixels | 7,100,000 pixels | 7,100,000 pixels | 5,000,000 pixels | 8,100,000 pixels | 8,000,000 pixels |

Motivation



- Questionnaire surveys produce special dataset
 - Interval
 - Ordinal
- Usually compared only within-variable

Motivation



5. In the past 2 weeks, how many times have you viewed a previous day's lab material - this includes both while in the lab and from elsewhere.

| | | | |
|-------|--|----|------|
| 0 | | 4 | 4% |
| 1-2 | | 33 | 33% |
| 3-4 | | 35 | 35% |
| 4+ | | 27 | 27% |
| Total | | 99 | 100% |

6. Have you ever experienced any of the following problems while submitting an assignment/homework? (Check all that apply)

| | | | |
|---|--|----|-----|
| Forgetting the process for submission | | 27 | 27% |
| Forgetting to submit an assignment | | 27 | 27% |
| Being unsure whether you submitted an assignment | | 51 | 52% |
| Being concerned whether you submitted the correct file(s) | | 58 | 59% |
| I had none of these difficulties. | | 20 | 20% |
| Other, please specify View Responses | | 4 | 4% |

7. In lab, would you prefer to discuss questions with your classmates in person or in an electronic forum?

| | | | |
|----------------|--|----|------|
| In Person | | 68 | 69% |
| Electronically | | 31 | 31% |
| Total | | 99 | 100% |

8. If you could choose to add **only one** of the following features, which would it be?

| | | | |
|---|--|----|-----|
| Indicator for progress through a day's material | | 15 | 15% |
| Search through material from all labs | | 57 | 58% |

Motivation

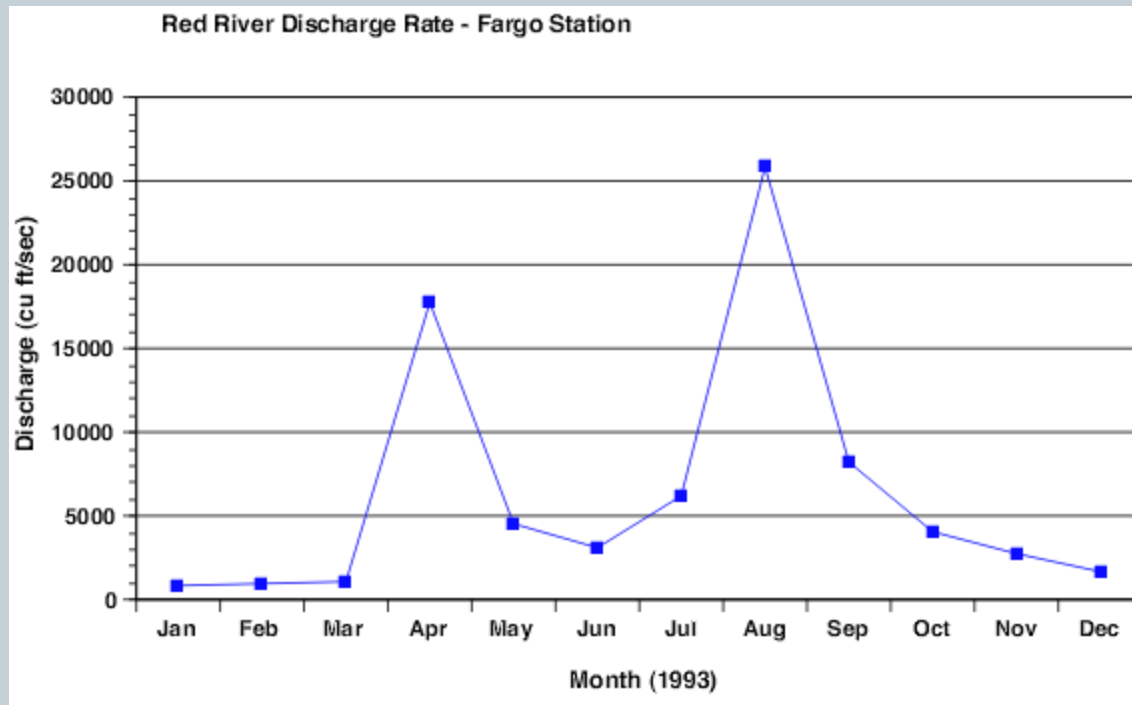


- Need a good way to relate variables to each other
- Need a good way to visualize multiple ordinal variables

Shortcomings



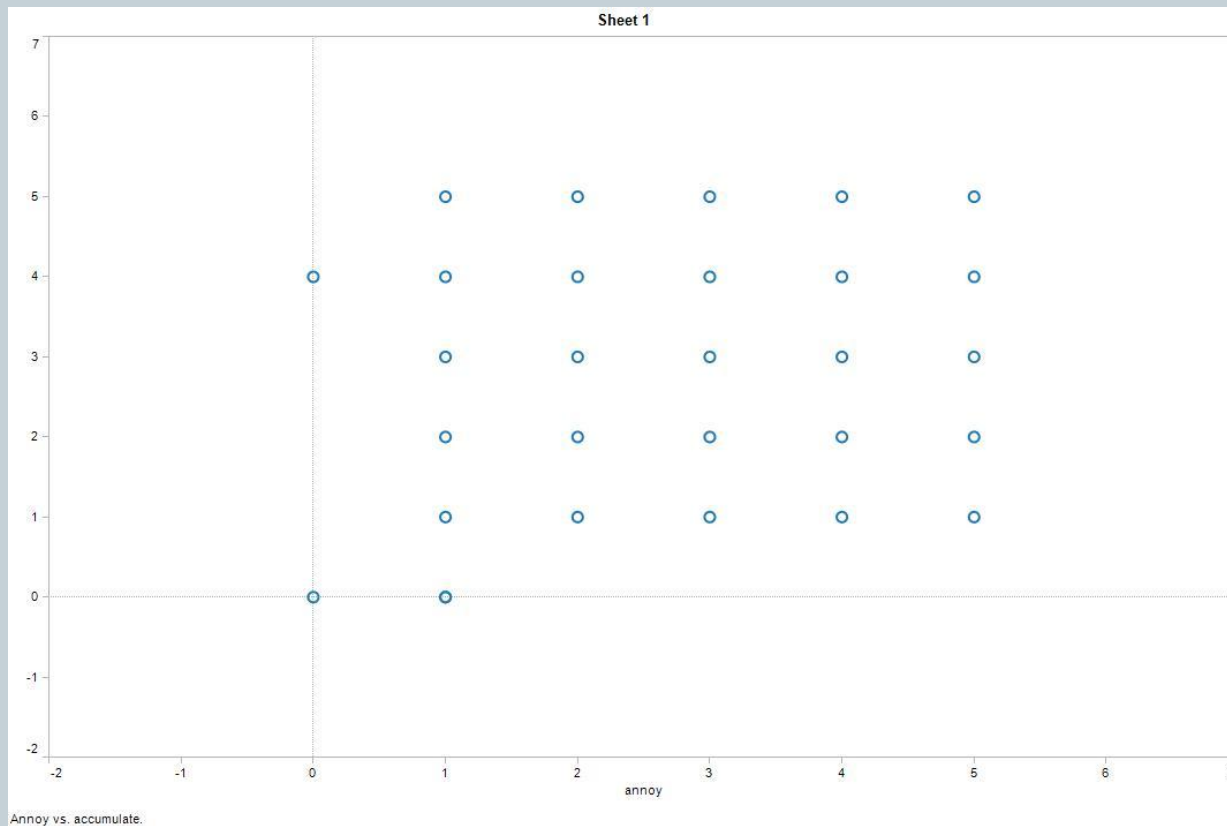
- Ordinal variables are usually graphed against interval variables



Shortcomings



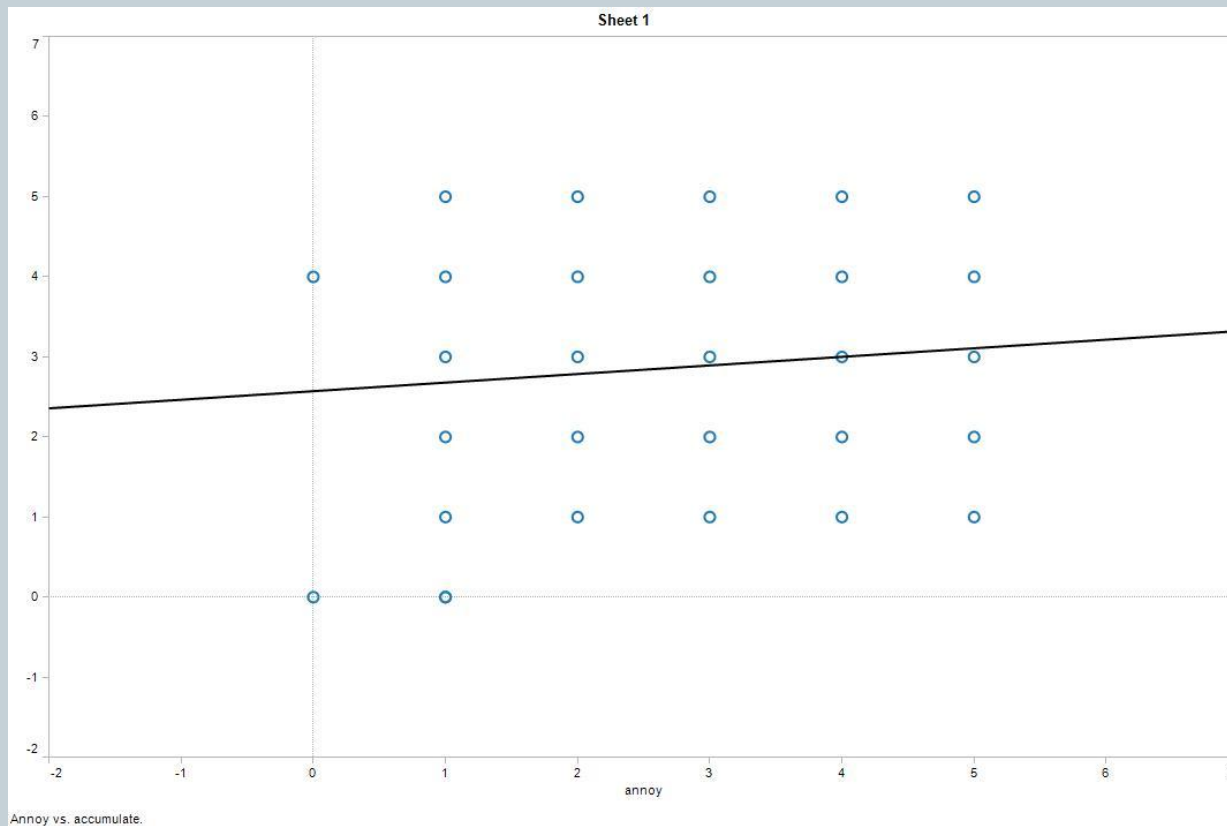
- Graphing ordinal against ordinal does not work well



Shortcomings



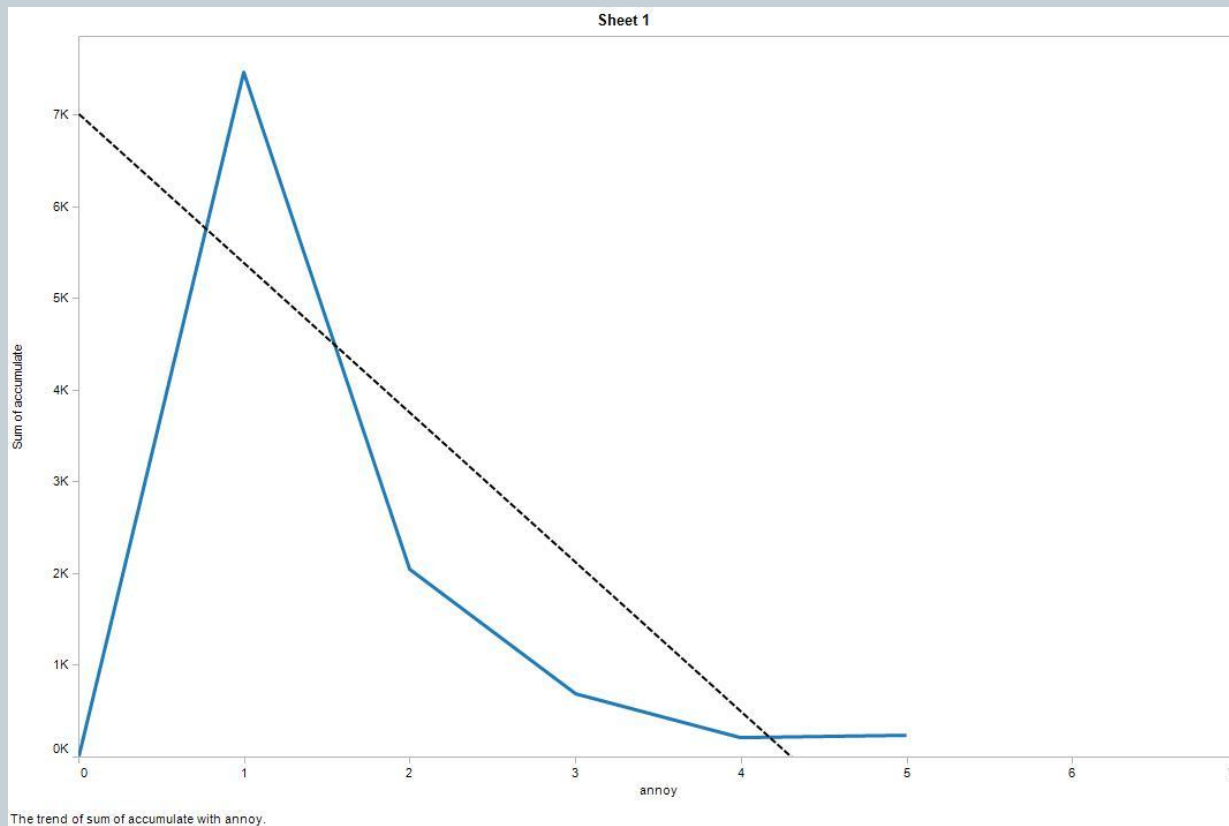
- Regression lines only help a little



Shortcomings



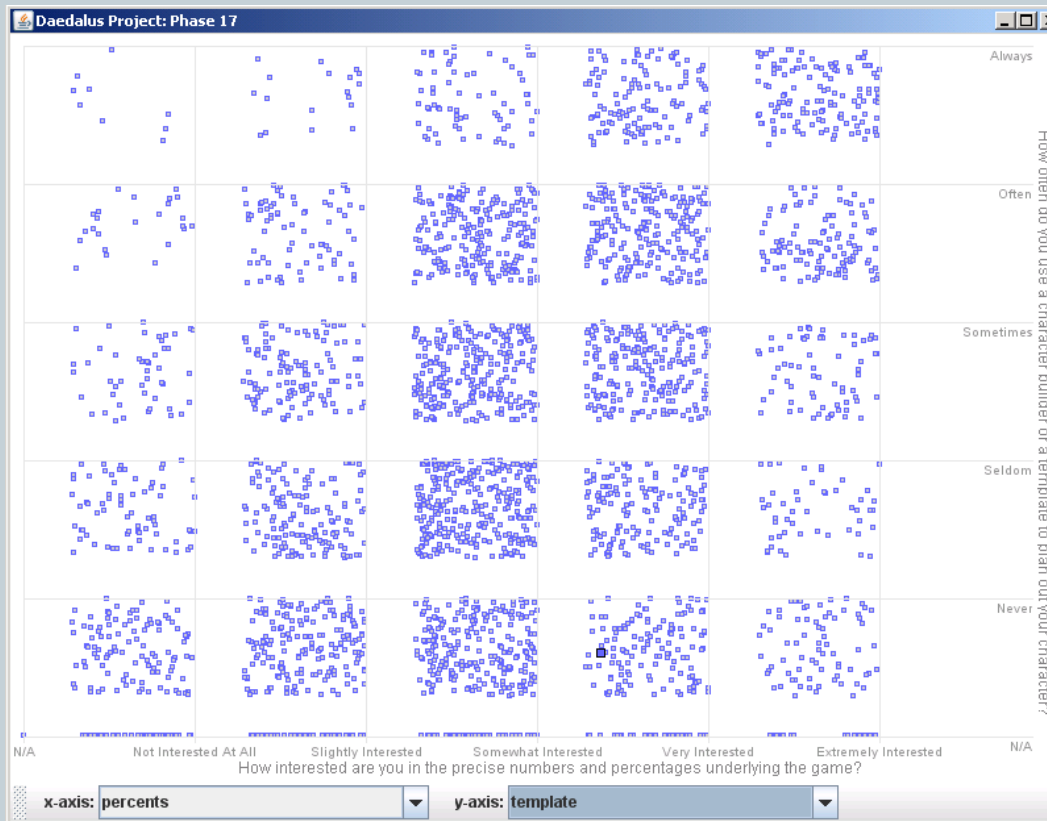
- Summing helps, but really encodes different data



Initial Concept



- Introduce random jitter



Initial Concept



- Multi-dimensional matrix
- Allow continuous rotation from viewpoint to viewpoint



Initial Concept



- ScatterDice

N. Elmqvist, P. Dragicevic, J.-D. Fekete. Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation. In *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2008)*, to appear, 2008. (Best paper award)

Previous Work



- **Geometrically transformed displays**

W. S. Cleveland. *Visualizing Data*. Hobart Press, 1993.

- **Iconic displays**

H. Chernoff. Using faces to represent points in k -dimensional space graphically. *Journal of the American Statistical Association*, 68:361–368, 1973.

- **Dense pixel displays**

D. A. Keim and H.-P. Kriegel. VisDB: Database exploration using multidimensional visualization. *IEEE Computer Graphics and Applications*, 14(5):40–49, Sept. 1994.

- **Dimensional stacked displays**

J. LeBlanc, M. O. Ward, and N. Wittels. Exploring N -dimensional databases. In *Proceedings of the IEEE Conference on Visualization*, pages 230–237, 1990.

Previous Work



- **Overview of methods**

D. A. Keim. Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics, 8(1):1–8, 2002.

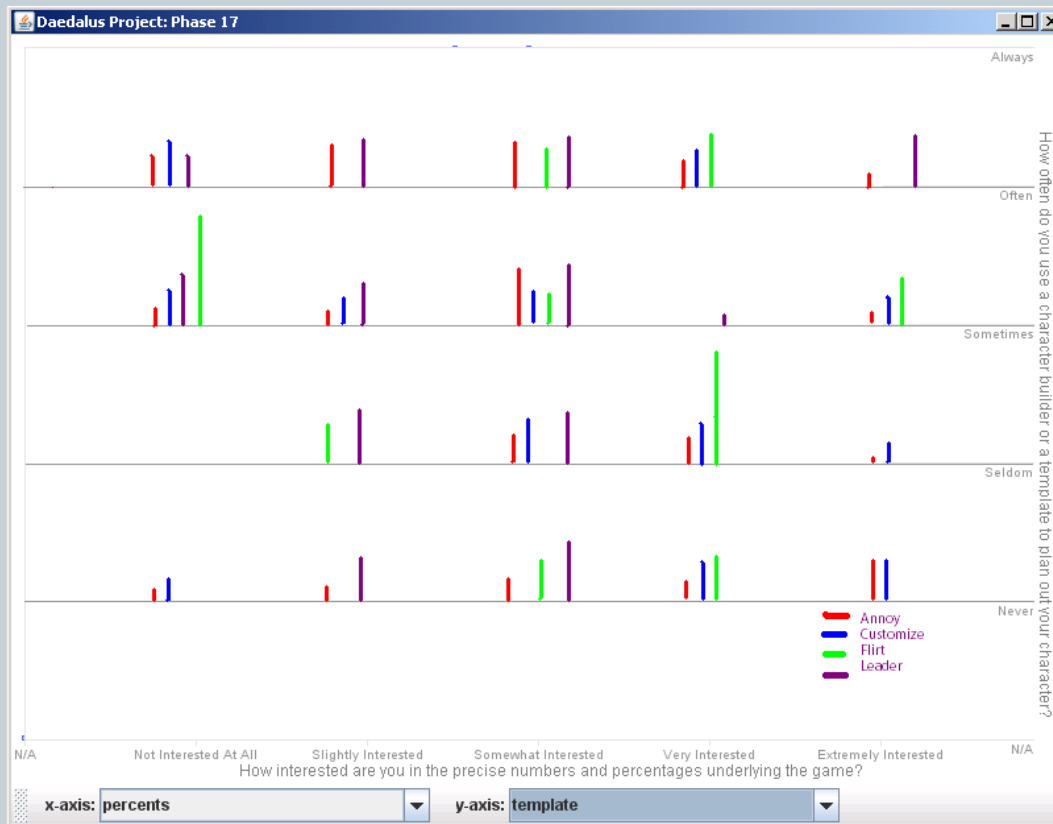
- **Encoding variables**

J. Bertin, *Graphics and Graphic Information Processing*, de Gruyter, Berlin, 1981.

Current Concept



- Color + Position + Size + Small multiples



Current Concept



- Much more difficult to interpret how the individual data points aggregate to the whole
- Allow many dimensions of data to be visualized using position
- Also considering how to specifically enhance a scatterplot to convey the data

Technical Challenges



- **Fitting many variables into small space**
 - Design through prototyping
- **Determining data encoding (colors? texture?)**
 - Reference previous research, experimentation
- **Maintaining part-to-whole relationships**
 - With each design, record what information is conveyed or lost
- **Building prototype**
 - Use existing knowledge and work in Prefuse and Flare

Milestones



- 10/31 - 5 other solution concepts
- 11/5 - Determine how, if at all to include interval data
- 11/5 - Create storyboards
- 11/10 - Determine color or other encoding
 - Create legend
- 11/20 - Build automatic optimal layout
- 11/25 - Design and build interaction
- 12/1 - Build Attribute-explorer style filters
- 12/10 - Create final presentation and paper