Why focus on tasks and effectiveness?
- tasks serve as constraint on design (as does data)
  - idioms do not serve all
  - Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Further reading
- optional textbook for following up on visualization foundations lectures
- optional papers/books
  - links and references posted on course page
- if DL links, use library EZproxy from off campus

Week 1: Intro, Tasks and Data, Marks and Channels
Tamara Munzner
University of British Columbia

Why have a human in the loop?
Computer-based visualization systems provide visual representations if datasets designed to help people carry out tasks more effectively.
- don't need vis when fully automatic solution exists and is trusted
- don't just replace people with computational decision-making methods.

Why analyze?
- imposes structure on huge design space
- scaffold to help you think systematically about choices
- increasing chance of finding good solutions if you understand full space of possibilities

Why show the data in detail?
- summaries lose information
- confirmed and find unexpected patterns
- assess validity of statistical model

Why use an external representation?
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.
- external representation: replace cognition with perception
- domain situation
  - who are the target users?
- task abstraction
  - computational limits
  - processing time
  - system memory
  - human limits
  - attention and memory
- display limits
  - pixels are precious resources, the most constrained resource
  - information density ratio of space used to encode info vs unused whitespace
  - tradeoff between clutter and viewing space, find sweet spot between design and space

Why focus on tasks and effectiveness?
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.
- task abstraction
  - computational limits
  - processing time
  - system memory
  - human limits
  - attention and memory
- display limits
  - pixels are precious resources, the most constrained resource
  - information density ratio of space used to encode info vs unused whitespace
  - tradeoff between clutter and viewing space, find sweet spot between design and space

Why analyze?
- imposes structure on huge design space
  - scaffold to help you think systematically about choices
  - analyzing existing as stepping stone to designing new
- most possibilities ineffective for particular task/data combination

Why show the data in detail?
- summaries lose information
- confirmed and find unexpected patterns
- assess validity of statistical model

Why analyze?
- imposes structure on huge design space
  - scaffold to help you think systematically about choices
  - analyzing existing as stepping stone to designing new
- most possibilities ineffective for particular task/data combination

Analysis framework: Four levels, three questions
- domain situation
  - are the target users?
- abstraction
  - translate from specifics of domain to vocabulary of vis
  - what is shown? data abstraction
  - why is the user looking at it? task abstraction
- idiom
  - how is it shown?
  - visual encoding idiom: how to draw
  - interaction idiom: how to manipulate
- algorithm
  - efficient computation
Dataset and data types

- Data and Dataset Types
  - Tables
  - Networks & Fields
  - Geometry
  - Cluster, Trees
  - Bins
  - Attributes

- Data Types
  - Items
  - Positions

- Dataset Availability
  - Static
  - Dynamic

Derived

- don’t just draw what you’re given!
  - decide what the right thing to show is
  - create it with a series of transformations from the original dataset
  - draw data:
    - one of the four major strategies for handling complexity

Attribute types

- Attribute Types
  - Categorical
  - Ordered
  - Quantitative

- Ordering Direction
  - Sequential
  - Diverging
  - Cyclic

Actions: Search, query

- what does user know?
  - target, location

- how much of the data matters?
  - one, some, all

- independent choices for each of these three levels
  - analyze, search, query

Analysis example: Derive one attribute

- slider number
  - centrality metric for streets/networks
  - derived (quantitative attributes)
  - draw up $5 of $20K for good skeletons

Definitions: Marks and channels

- marks: positional primitives
  - geometric centers

- channels: visual appearance of marks
  - color

Encoding visually

- analyze symbol structure
  - as combination of marks and channels

Three major datatypes

- Data/Task abstraction
- Data/tasks
- What?
- Why?
- Attributes
- Data/Task Types
- Data and Task Types
- Tables
- Networks
- Spatial (Continuous)
- Geometry (Spatial)

- visualization vs computer graphics
- geometry vs design decisions

Actions: Analyze

- consumers

- discover vs present
  - classic split
  - also explore vs explain

- produce
  - newsletter
  - also causal, social

- produce
  - annotate, record
  - derive
  - crucial design choice

Derive

- new distributed
  - feature models
  - feature vectors
  - feature typology

Encoding visually with marks and channels

- analyze symbol structure
  - as combination of marks and channels

- marks: positional primitives

- channels: visual appearance of marks

Why is validation difficult?

- different ways to get it wrong at each level

- solution: use methods from different fields at each level

- problem-driven work

- data-driven work

Three major datatypes

- Data/Task abstraction
- Data/tasks
- What?
- Why?
- Attributes
- Data/Task Types
- Data and Task Types
- Tables
- Networks
- Spatial (Continuous)
- Geometry (Spatial)

- visualization vs computer graphics
- geometry vs design decisions

Actions: Analyze

- consumers

- discover vs present
  - classic split
  - also explore vs explain

- produce
  - newsletter
  - also causal, social

- produce
  - annotate, record
  - derive
  - crucial design choice

Derive

- new distributed
  - feature models
  - feature vectors
  - feature typology

Encoding visually with marks and channels

- analyze symbol structure
  - as combination of marks and channels

- marks: positional primitives

- channels: visual appearance of marks

Why is validation difficult?

- different ways to get it wrong at each level

- solution: use methods from different fields at each level

- problem-driven work

- data-driven work

Three major datatypes

- Data/Task abstraction
- Data/tasks
- What?
- Why?
- Attributes
- Data/Task Types
- Data and Task Types
- Tables
- Networks
- Spatial (Continuous)
- Geometry (Spatial)

- visualization vs computer graphics
- geometry vs design decisions

Actions: Analyze

- consumers

- discover vs present
  - classic split
  - also explore vs explain

- produce
  - newsletter
  - also causal, social

- produce
  - annotate, record
  - derive
  - crucial design choice

Derive

- new distributed
  - feature models
  - feature vectors
  - feature typology

Encoding visually with marks and channels

- analyze symbol structure
  - as combination of marks and channels

- marks: positional primitives

- channels: visual appearance of marks

Why is validation difficult?

- different ways to get it wrong at each level

- solution: use methods from different fields at each level

- problem-driven work

- data-driven work

Three major datatypes

- Data/Task abstraction
- Data/tasks
- What?
- Why?
- Attributes
- Data/Task Types
- Data and Task Types
- Tables
- Networks
- Spatial (Continuous)
- Geometry (Spatial)

- visualization vs computer graphics
- geometry vs design decisions

Actions: Analyze

- consumers

- discover vs present
  - classic split
  - also explore vs explain

- produce
  - newsletter
  - also causal, social

- produce
  - annotate, record
  - derive
  - crucial design choice

Derive

- new distributed
  - feature models
  - feature vectors
  - feature typology

Encoding visually with marks and channels

- analyze symbol structure
  - as combination of marks and channels

- marks: positional primitives

- channels: visual appearance of marks

Why is validation difficult?

- different ways to get it wrong at each level

- solution: use methods from different fields at each level

- problem-driven work

- data-driven work

Three major datatypes

- Data/Task abstraction
- Data/tasks
- What?
- Why?
- Attributes
- Data/Task Types
- Data and Task Types
- Tables
- Networks
- Spatial (Continuous)
- Geometry (Spatial)

- visualization vs computer graphics
- geometry vs design decisions

Actions: Analyze

- consumers

- discover vs present
  - classic split
  - also explore vs explain

- produce
  - newsletter
  - also causal, social

- produce
  - annotate, record
  - derive
  - crucial design choice

Derive

- new distributed
  - feature models
  - feature vectors
  - feature typology

Encoding visually with marks and channels

- analyze symbol structure
  - as combination of marks and channels

- marks: positional primitives

- channels: visual appearance of marks
Discriminability: How many usable steps?

- must be sufficient for number of attribute levels to show
  - line width: few bins

Separability vs. Integrity

- perceptual system mostly operates with relative judgements, not absolute
  - that's why accuracy increases with common frame/scale and alignment
  - Weber's Law: ratio of increment to background is constant
  - filled rectangles differ in length by 1:9, difficult judgement

Relative vs. absolute judgements

- perception of luminance is contextual based on contrast with surroundings

Relative luminance judgements

- color constancy across broad range of illumination conditions

Further reading

- Chap 1, What's Vis, and Why Do It?
- Chap 2, What Data Abstraction
- Chap 3, Why Such Abstraction
- Chap 4, Analyze Your Lenses for Visualization
- Chap 5, Marks and Channels

- Perception in Vis: web page with demos, Christopher Healey.
Demo 3: Vancouver Crime

- **Tableau Lessons**
  - multiple pills on a shelf, pill ordering
  - show filters
  - undo
  - duplicate & rename tabs

- **Big Ideas**
  - underlying causes can be tricky to understand

---

Demo 4: Back to the Future

- **Tableau Lessons**
  - simple analytics: totals
  - more disaggregation practice
  - Show Me

- **Big Ideas**
  - beyond simple bars
  - challenges of missing data

---

Assignment

- **Music Sales**
  - work through workbook on your own
  - submit finished version (in workbook .twbx format)

- **Vancouver Crime**
  - analyze further on your own
  - write up brief news story (submit in PDF format)
  - < 500 words
  - up to 2 screenshots from Tableau
  - write up reflections (submit in PDF format)
  - discuss dead ends
  - include Tableau screenshots

- submit before next class (Thu Tue Sep 20)
  - email tmm@cs.ubc.ca and caitlin@discoursemedia.org with subject JOURN Week 1

---