

User Tasks & Analysis



CS 4460 – Intro. to Information Visualization
September 29, 2017
John Stasko

Learning Objectives



- Understand the importance of tasks, goals, and objectives for visualization
- Identify the common "low-level" tasks for visualizations
- Identify important "high-level" tasks for visualizations

What for?



- In order to build better visualizations, we need to understand what people might use them for
 - What tasks do they want to accomplish?

An Example



- search vs. browsing
- Value of Vis day (coming up):
 - Exploratory data analysis
 - Identifying better questions
 - Understanding, awareness, context, trust

Browsing vs. Search



- Important difference in activities
- Appears that information visualization may have more to offer to browsing

- But...browsing is a softer, fuzzier activity
- So, how do we articulate utility?
 - Maybe describe when it's useful
 - When is browsing useful?

Fall 2016

CS 7450

5

Browsing



- Useful when
 - Good underlying structure so that items close to one another can be inferred to be similar
 - Users are unfamiliar with collection contents
 - Users have limited understanding of how system is organized and prefer less cognitively loaded method of exploration
 - Users have difficulty verbalizing underlying information need
 - Information is easier to recognize than describe

Lin '97

Fall 2016

CS 7450

6

Thought



- Maybe infovis isn't about answering questions or solving problems... hmmm
- Maybe it's about asking better questions

Tasks



- OK, but browsing and search are very high level
- Let's be more specific...

Challenge



	A	B	C	D	E	F	G	H
1	Name	Rating	Country	Category	Price	ABV	Age	Brand
2	Tyrconnell 10 Year Old Single Malt W	100	Ireland	Single Malt	72	46	10	Tyrconnell
3	Dalmore 18 Year Old Single Highland	100	Scotland	Highlands	165	43	18	Dalmore
4	Powers 12 Year Old Irish Whiskey	99	Ireland	Blended	35	40	12	Powers
5	Sumtory The Yamazaki 18 Year Old Si	99	Japan	Single Malt	120	43	18	Sumtory
6	Glenmorangie 10 Year Old Single Ma	99	Scotland	Highlands	42	40	10	Glenmorangie
7	Glenmorangie Single Malt Scotch	99	Scotland	Highlands	80	46	10	Glenmorangie
8	Bunnahabhain 18 Year Old Single M	99	Scotland	Islay	92	46.3	18	Bunnahabhain
9	Laphroaig 18 Year Old Single Malt Sc	99	Scotland	Islay	107	48	18	Laphroaig
10	Cardhu 12 Year Old Single Malt Scot	99	Scotland	Highlands	45	40	12	Cardhu
11	Aberlour 18 Year Old Single Malt Sco	99	Scotland	Speyside	100	43	18	Aberlour
12	Balvenie 14 Year Old Single Malt Sco	99	Scotland	Speyside	60	43	14	Balvenie
13	Caol Ias Single Malt Scotch Distillers E	99	Scotland	Islay	70	43 *		Caol Ias
14	Kingdom 17 Year Old Scotch	99	Scotland	Blended	50	40	17	Kingdom
15	Balvenie 12 Year Old Doublewood Si	99	Scotland	Speyside	45	40	12	Balvenie
16	Glen Garioch Founders Reserve Scot	99	Scotland	Highlands	45	48 *		Glen Garioch
17	Bowmore 15 Year Old Single Malt Sc	99	Scotland	Islay	70	43	15	Bowmore
18	Rebel Yell Kentucky Straight Bourbon	99	USA	Bourbon	11	40 *		Rebel Yell
19	Pappy Van Winkle 15 Year Old Famil	99	USA	Bourbon	58	53.5	15	Pappy Van Winkle
20	Thomas H. Handy Kentucky Straight	99	USA	Rye	67	66.4	6	Thomas H. Handy
21	Ardbeg Uigeadail	99	Scotland	Islay	80	54.2 *		Ardbeg
22	Noah's Mill Bourbon	99	USA	Bourbon	60	57.15 *		Noah's Mill
23	Parker's Heritage Bourbon	99	USA	Bourbon	80	62	15	Parker's Heritage
24	Glenlivet 21 Year Old Single Malt Sco	97	Scotland	Speyside	123	40	21	Glenlivet
25	Macallan 21 Year Old Fine Oak Scotc	96	Scotland	Speyside	220	43	21	Macallan
26	George T. Stag Kentucky Straight Bl	96	USA	Bourbon	70	45 *		George T. Stag
27	Parker's Heritage Collection 10 Year	96	USA	Bourbon	80	63	10	Parker's Heritage
28	Rowan's Creek Bourbon	96	USA	Bourbon	50	50.05	12	Rowan's Creek
29	Woodford Reserve Master's Collecti	96	USA	Bourbon	80	46.2 *		Woodford Reserve
30	Lagavulin 21 Scotch	96	Scotland	Islay	300	56.5	21	Lagavulin
31	Highland Park 30 Scotch	96	Scotland	Islands	365	48.1	30	Highland Park
32	King Car Single Malt Whiskey	96	Taiwan	Single Malt	84	46 *		Kavalan
33	Rye Dog Whiskey	96	USA	Rye	65	50	0	Delaware Phoenix
34	Thirteen Colony Southern Corn WI	96	USA	Corn	30	47.5 *		Thirteen Colony
35	Glenfiddich 12 Year Old Single Malt S	96	Scotland	Speyside	43	40	12	Glenfiddich
36	Oban 15 Year Old Single Malt Scotch	96	Scotland	Highlands	89	43	15	Oban
37	Old Pulteney 30 Year Old Single Malt	96	Scotland	Highlands	400	44	30	Old Pulteney

Whiskeys

Come up with analytic queries, tasks, goals...

Fall 2016

CS 7450

9

Follow-on



- What are the (types of) tasks being done here?
 - Abstract away the domain
- Can you think of others?

Fall 2016

CS 7450

10

Task Taxonomies



- Number of different ones exist, important to understand what process they focus on
 - Creating an artifact
 - Human tasks
 - Tasks using visualization system
 - ...

One (Famous) Perspective



- Shneiderman proposed task \times data type taxonomy to understand what people do with visualization
- Mantra: “Overview first, zoom and filter, then details on demand”
 - Design paradigm for infovis systems

Taxonomy



- Data Types
 1. 1D
 2. 2D
 3. 3D
 4. Temporal
 5. ND
 6. Tree
 7. Network
- Tasks
 1. Overview
 2. Zoom
 3. Filter
 4. Details-on-demand
 5. Relate
 6. History
 7. Extract

Another Task Taxonomy



- Amar, Eagan, & Stasko – InfoVis '05

Background



- Use “commercial tools” class assignment from this class
- Students generate questions to be answered using commercial infovis systems
- Data sets:

Domain	Data cases	Attributes	Questions Generated
Cereals	78	15	107
Mutual funds	987	14	41
Cars	407	10	153
Films	1742	10	169
Grocery surveys	5164	8	126

- Generated 596 total analysis tasks

Fall 2016

CS 7450

15



Terminology



- *Data case* – An entity in the data set
- *Attribute* – A value measured for all data cases
- *Aggregation function* – A function that creates a numeric representation for a set of data cases (eg, average, count, sum)

1. Retrieve Value



General Description:

Given a set of specific cases, find attributes of those cases.

Examples:

- What is the mileage per gallon of the Audi TT?
- How long is the movie Gone with the Wind?

2. Filter



General Description:

Given some concrete conditions on attribute values, find data cases satisfying those conditions.

Examples:

- What Kellogg's cereals have high fiber?
- What comedies have won awards?
- Which funds underperformed the SP-500?

3. Compute Derived Value



General Description:

Given a set of data cases, compute an aggregate numeric representation of those data cases.

Examples:

- What is the gross income of all stores combined?
- How many manufacturers of cars are there?
- What is the average calorie content of Post cereals?

4. Find Extremum



General Description:

Find data cases possessing an extreme value of an attribute over its range within the data set.

Examples:

- What is the car with the highest MPG?
- What director/film has won the most awards?
- What Robin Williams film has the most recent release date?

5. Sort



General Description:

Given a set of data cases, rank them according to some ordinal metric.

Examples:

- Order the cars by weight.
- Rank the cereals by calories.

6. Determine Range



General Description:

Given a set of data cases and an attribute of interest, find the span of values within the set.

Examples:

- What is the range of film lengths?
- What is the range of car horsepowers?
- What actresses are in the data set?

7. Characterize Distribution



General Description:

Given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute's values over the set.

Examples:

- What is the distribution of carbohydrates in cereals?
- What is the age distribution of shoppers?

8. Find Anomalies



General Description:

Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers.

Examples:

- Are there any outliers in protein?
- Are there exceptions to the relationship between horsepower and acceleration?

9. Cluster



General Description:

Given a set of data cases, find clusters of similar attribute values.

Examples:

- Are there groups of cereals w/ similar fat/calories/sugar?
- Is there a cluster of typical film lengths?

10. Correlate



General Description:

Given a set of data cases and two attributes, determine useful relationships between the values of those attributes.

Examples:

- Is there a correlation between carbohydrates and fat?
- Is there a correlation between country of origin and MPG?
- Do different genders have a preferred payment method?
- Is there a trend of increasing film length over the years?

Discussion/Reflection



- Compound tasks
 - “Sort the cereal manufacturers by average fat content”
Compute derived value; Sort
 - “Which actors have co-starred with Julia Roberts?”
Filter; Retrieve value

Discussion/Reflection



- What questions were left out?
 - Basic math
 - “Which cereal has more sugar, Cheerios or Special K?”
 - “Compare the average MPG of American and Japanese cars.”
 - Uncertain criteria
 - “Does cereal (X, Y, Z...) sound tasty?”
 - “What are the characteristics of the most valued customers?”
 - Higher-level tasks
 - “How do mutual funds get rated?”
 - “Are there car aspects that Toyota has concentrated on?”
 - More qualitative comparison
 - “How does the Toyota RAV4 compare to the Honda CRV?”
 - “What other cereals are most similar to Trix?”

Fall 2016

CS 7450

31

Concerns/Limitations



- InfoVis tools may have influenced students' questions
- Graduate students as group being studied
 - How about professional analysts?
- Subjective – Not an exact science
- Data was really quantitative so may get a different set of tasks for relational/graph data
 - See Lee et al, BELIV '06

Fall 2016

CS 7450

32

Contributions



- Set of grounded low-level analysis tasks
- Potential use of tasks as a language/vocabulary for comparing and evaluating infovis systems

Another Perspective



- Taxonomy proposed
- "...used specifically for multidimensional visualizations, taking into account the generic objectives that a user has when using such techniques to perform exploratory analyses as a previous step of statistical analysis."

Task Taxonomy



- 7 tasks in 2 categories
 - User goals
 - Identify – Find, discover new information
 - Determine – Calculate, define a precise value
 - Compare – Compare data & values
 - Infer – Infer knowledge, generate hypotheses
 - Locate – Search and identify information
 - Intermediate level tasks to support analysis
 - Visualize – Represent the data a certain way
 - Configure – Normalize, filter, reorder, etc.

Fall 2016

CS 7450

35

More Details



- Each task has “parameters”
 - Identify
 - clusters
 - correlations
 - categories
 - properties
 - patterns
 - characteristics
 - thresholds
 - similarities
 - differences
 - dependencies
 - uncertainties
 - variations

Fall 2016

CS 7450

36

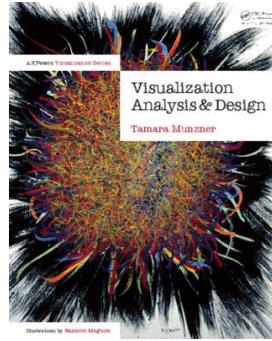
Abstract Tasks



Framework/Typology of abstract visualization tasks



Brehmer & Munzner
TVCG (InfoVis) '13



Chapter 3

Fall 2016

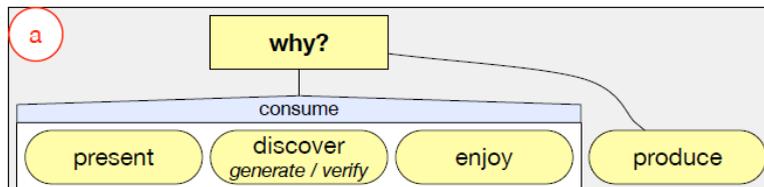
CS 7450

37

Why?



What are the top-level categories (answers) to the "Why?" question?

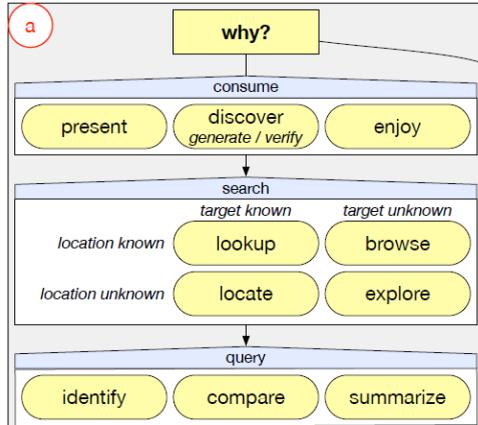


Fall 2016

CS 7450

38

Discover



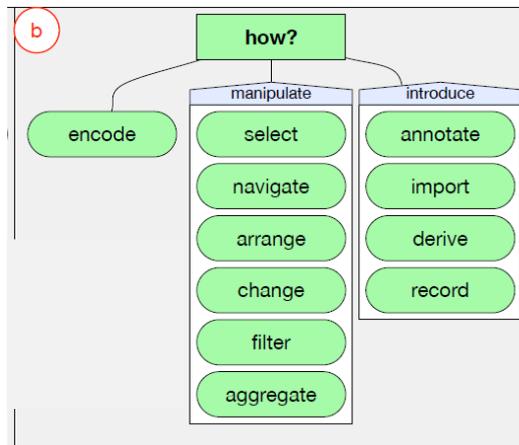
High to
low level

Fall 2016

CS 7450

39

How?



Fall 2016

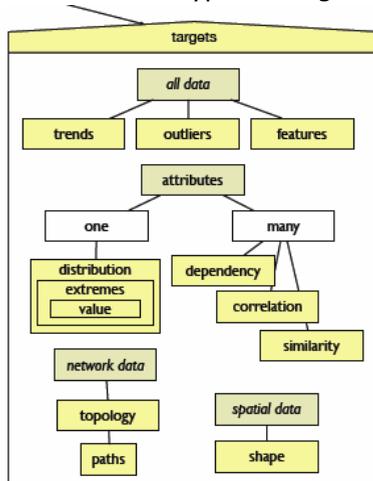
CS 7450

40

Targets



What are the types of targets?

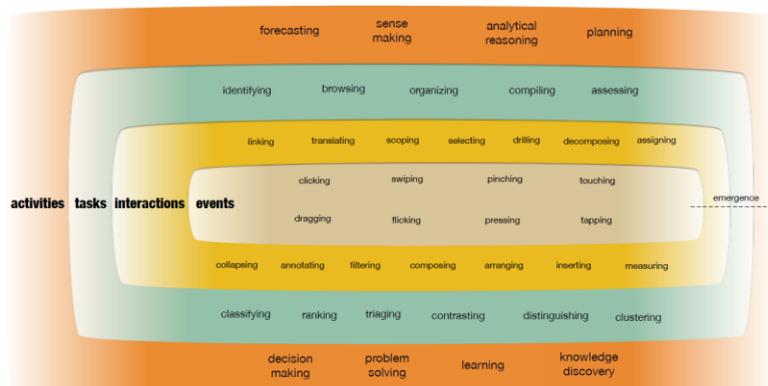


Fall 2016

CS 7450

41

Visual Analytic Activity



Sedig, Parsons, Babanski
JMPT'12

Fall 2016

CS 7450

42

Learning Objectives



- Understand the importance of tasks, goals, and objectives for visualization
- Identify the common "low-level" tasks for visualizations
- Identify important "high-level" tasks for visualizations

Fall 2016

CS 7450

43

HW 4



- Questions?
 - Get Tableau
 - Pick data set
- Due next Friday, 6th at 1pm (T-Square)

Fall 2016

CS 7450

44

Midterm Exam



- About 15 short answer questions
- Closed book/notes
- Examples
 - Concepts (majority)
 - Analyze a vis
 - Design a vis
 - Analyze some code

Fall 2016

CS 7450

45

Upcoming



- Midterm Exam
- **No (Scheller 300) class next Wed & Fri**
 - John T. and I are away all next week
- Value of Vis
 - Prep: Norman & Chang articles
 - Watch My EuroVis '14 Capstone talk

Fall 2016

CS 7450

46

References



- Spence & CMS texts
- All referred to papers