Data Visualization (CSCI 490/680)

Design

Dr. David Koop





Set and Cluster Visualization

- Set and cluster visualization not covered in depth in the textbook
- Nice summary of set visualization in the following paper:
 - Visualizing Sets and Set-typed Data, B. Alsallakh et al., 2014
 - Also: <u>Web Resources</u>









Set-typed Data - Characteristics

- Set Algebra
 - Set operations, Cartesian product, power set, ...
- Set similarities
 - Similarity measures (Jaccard, Tversky, etc.)
- Element degree
 - exclusive set membership
- Dimensionality
 - 2ⁿ possible combinations
 - $2^{(2^n)}$ possible queries

































































Find elements based on their set memberships



Find elements with a specific set membership degree





Filter out elements based on their set membership degrees

Create a new set that contains certain elements

Find elements belonging to a specific set

Find sets containing a specific element

Filter out elements based on their set memberships



Find elements belonging to a specific set

Find elements based on their set memberships

Find elements with a specific set membership degree

Filter out elements based on their set memberships

Filter out elements based on their set membership degrees

Create **a new set** that contains certain elements











Find elements based on their set memberships



Find elements with a specific set membership degree

Filter out elements based on their set membership degrees



Create **a new set** that contains certain elements

Find elements belonging to a specific set

Filter out elements based on their set memberships



Find elements belonging to a specific set



Find sets containing a specific element





Find elements based on their set memberships

Find elements with a specific set membership degree

Filter out elements based on their set memberships

Filter out elements based on their set membership degrees



Create a new set that contains certain elements





Find sets containing a specific element



Find elements based on their set memberships



Find elements with a specific set membership degree

Filter out elements based on their set memberships



Create a new set that contains certain elements

Find elements belonging to a specific set

Filter out elements based on their set membership degrees



#

Find the **#** of sets in the set family



Analyze **Mo**







Analyze **exclusion relations**



Analyze https://www.intersection.relations



Find intersections between k sets

Find sets involved in certain intersection







Find the set with largest pairwise set intersections



Analyze set & set intersection cardinalities



Analyze and compare set similarities



Analyze and compare set exclusiveness



Highlight specific sets, subsets, or set relations









Find the set with largest pairwise set intersections



Analyze set & set intersection cardinalities



Analyze and compare set similarities



Analyze and compare set exclusiveness



Highlight specific sets, subsets, or set relations













set-theoretical operations



#

Find the **#** of sets in the set family



Analyze **Mo**







Analyze **exclusion relations**



Analyze https://www.intersection.relations



Find intersections between k sets

Find sets involved in certain intersection







Find the set with largest pairwise set intersections



Analyze set & set intersection cardinalities



Analyze and compare set similarities



Analyze and compare set exclusiveness



Highlight specific sets, subsets, or set relations





#

Find the **#** of sets in the set family



Analyze **Mo**







Analyze **exclusion relations**



Analyze https://www.intersection.relations



Find intersections between k sets

Find sets involved in certain intersection







Find the set with largest pairwise set intersections



Analyze set & set intersection cardinalities



Analyze and compare set similarities



Analyze and compare set exclusiveness



Highlight specific sets, subsets, or set relations



Find the attribute value of a certain element

Find the distribution of an attribute in a certain set or subset

Compare the attribute values of sets or set intersections

Analyze the set memberships for elements having certain attribute values

Create a new set out of elements having certain attribute values

Element Attributes



Compare the attribute values of sets or set intersections

Analyze the set memberships for elements having certain attribute values

> Create a new set out of elements having certain attribute values

Element Attributes

)

Find the attribute value of a certain element





Compare the attribute values of sets or set intersections

Analyze the set memberships for elements having certain attribute values

> Create a new set out of elements having certain attribute values



Element Attributes

)

Find the attribute value of a certain element





Compare the attribute values of sets or set intersections

Analyze the set memberships for elements having certain attribute values

> Create a new set out of elements having certain attribute values





Element Attributes

Find the attribute value of a certain element









Compare the attribute values of sets or set intersections

Analyze the set memberships for elements having certain attribute values

> Create a new set out of elements having certain attribute values



Element Attributes

Find the attribute value of a certain element









Venn Diagram











What about cardinality?

Area encoding



D. Koop, CS 490/680, Fall 2019

Using glyphs



[B. Alsallakh et al., 2014]



Northern Illinois University



Euler Diagram Variants

use edges



D. Koop, CS 490/680, Fall 2019



use a concentric layout









KelpFusion



(a) Bubble Sets



(b) Kelp Diagrams

D. Koop, CS 490/680, Fall 2019

(c) LineSets



(e) KelpFusion (medium)



[Meulemans et al., 2013]



Biset edge bundling (and grouping)



D. Koop, CS 490/680, Fall 2019

[M. Sun et al., 2016]















Project Design

- Start working on turning your visualization ideas into designs
- Feedback to Blackboard today
- Sketch (talk about today)
- Options:
 - Try vastly different options
 - Refine an initial idea





<u>Assignment 4</u>

- Create Choropleth Maps
 - Deal with projections and GeoJSON Data
 - Select appropriate colormaps
- [CS 680 Only] Part 3 is using other libraries, you only need to do one option
- Example image at the right is **not** a solution to Part 3, needs proper colormapping!





Guidelines for Visualization Design









WTF Visualizations (<u>wtfviz.net</u>)

3D Category Scatter



Stage

D. Koop, CS 490/680, Fall 2019







Northern Illinois University



Tufte: "The da Vinci of Data" —<u>NYTimes</u>



D. Koop, CS 490/680, Fall 2019

BEAUTIFUL EVIDENCE

The Visual Display of Quantitative Information SECOND EDITION

VISUAL EXPLANATIONS

[https://www.edwardtufte.com/tufte/, 2017]



Northern Illinois University







Bad: Data magnitude <≠> Mark magnitude









Good: Data magnitude <=> Mark magnitude







Starting Scales at Zero?

Median household income in 2010 inflation adjusted dollars



D. Koop, CS 490/680, Fall 2019

——Non-Zero Based









Wavy baselines for non-zero starts









Cherry-picking data



D. Koop, CS 490/680, Fall 2019

[Fox News via Media Matters, 2012]







Show all the data











Tufte's Lie Factor









Tufte's Lie Factor

- Size of effect = (2nd value 1st value) / (1st value)
- Lie factor = (size of effect in graphic) / (size of effect in data)
- In the graphic:

Lie Factor =





D. Koop, CS 490/680, Fall 2019

5.3 - 0.6 0.6 14.8 27.5 - 18









(Some of) Tufte's Integrity Principles

- Show data variation, not design variation
- Clear, detailed, and thorough labeling and appropriate scales
- Size of the graphic effect should be directly proportional to the numerical quantities ("lie factor")









Avoid Chartjunk











Avoid Chartjunk













Avoid Chartjunk













Avoid Chartjunk?









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

Low QUALITY + DESCRIPTION













Data-to-Ink Ratio (Also Unjustified 3D)







D. Koop, CS 490/680, Fall 2019

Northern Illinois University







Maximize Data-to-Ink Ratio



D. Koop, CS 490/680, Fall 2019

Z Z Z Z

No Unjustified 3D

D. Koop, CS 490/680, Fall 2019

matplotlib gallery

Excel Charts Blog

No Unjustified 3D

- Occlusion hides information
- Perspective distortion dangers
- Tilted text isn't legible
- Can help with shape perception

Eyes Beat Memory

- Reduce cognitive load (using up working memory)
- Animation versus side-by-side views
- Change blindness

"Computer-based visualization systems provide visual tasks more effectively."

D. Koop, CS 490/680, Fall 2019

representations of datasets designed to help people carry out

– T. Munzner

Design Iteration

Design Iteration

New York Giants Indianapolis Colts San Diego Chargers **Baltimore Ravens New England Patriots Green Bay Packers New Orleans Saints Atlanta Falcons New York Jets** Cincinnati Bengals **Houston Texans Carolina Panthers Denver Broncos Arizona Cardinals Jacksonville Jaguars Detroit Lions** ampa Bay Buccaneers **Dallas Cowboys**

Daudan Manaina					Andrewsland
Peyton Manning					Andrew Luck
Drew Brees Drew Brees	Philip Rivers				
Kyle Boller S	teve McNair	Joe Flacco			
Tom Brady		Matt Cassel	Tom Brady		
Brett Favre		Aaron Rodgers		Aaron Rodgers	Aaron Rodgers
Aaron Brooks D	rew Brees		Drew Bree	s	
Michael Vick Michael V	ick	Matt Ryan	Matt Rya	n	
	Chad Penningtor	Brett Favre M	ark Sanch : Mark Sa	nchez	
Carson Palm Carson Palmer		Ryan Fitzp Carson	Palmer	Andy Dalton	
David Carr		Matt Schar	Jb		Matt Schaub
Jake Delhomme		Jake Delhomme		Cam Newto	on
Jake Plummer	Jay Cutler		Kyle Orton	Tim Tebow	Peyton Manning
	Matt Leina · Kurt	Warner			
Byron Leftwich	David Garrard	David Garrard		Blaine Gab	bert
Joey Harrington	lon Kitna			Matthew Staf	ford
Chris Simr	ns Bruce Gra (Jeff Garcia		Josh Freema	n	Josh Freeman
Drew Bledsoe	Tony Romo	Tony Romo		Tony Romo	

Design Iteration

Design

- Unlike a math problem, there are many different approaches for the visualization of some data
- Need to have some way to discuss how to determine whether a visualization is doing what we want
- Validation: Understand why a design is effective
 - What problems can be effective
 - Do this at different levels

Four Nested Levels of Design

D. Koop, CS 490/680, Fall 2019

Visual encoding/interaction idiom

Potential problems at each level

Domain situation You misunderstood their needs

Data/task abstraction You're showing them the wrong thing

Wisual encoding/interaction idiom The way you show it doesn't work

Algorithm Your code is too slow

D. Koop, CS 490/680, Fall 2019

Validation at each level

D. Koop, CS 490/680, Fall 2019

- Ineffective encoding/interaction idiom
- Validate Test on target users, collect anecdotal evidence of utility Validate Field study, document human usage of deployed system

Five Design-Sheet Methodology

D. Koop, CS 490/680, Fall 2019

Five Stages

- 1. Meet with client and consider task; or contemplate task on own.
- 2. Ideate and sketch small ideas.
- 3. Sketch and plan three alternative designs.
- 4. Consider solutions with client; or deliberate on own.
- 5. Generate realization sheet, and implement prototype. Discuss with client and re-iterate if necessary.

Five Stages

D. Koop, CS 490/680, Fall 2019

The Five Sheets

Ideas	Sheet 1	Layout
Filter		
Categorize		
Combine & Refine		Focus / F
Question		
Ideation		Alterna

D. Koop, CS 490/680, Fall 2019

ative Designs

Example: University Access for Disabled Students

Sheets 2-4

D. Koop, CS 490/680, Fall 2019

1	TITLE A DOST WE BADROYADZLITY
4.0/	AUTHER:
XI	DATE : 16/11/2012
	SHEET : 2 - FOS 3
	TASK : BAR - CHART REPRESENTATION
	RE THE OSA, EMPLOYABILITY & LEANERS ONTO
M	· HOVERENG OVER EACH AVORAGED UNEVERSETY BAR DRAWS THE YEAR BY YEAR BREAKDOWN.
	· (LICKENG ON AN INDEVEDUAL
	YEAR WOULD SET ALL DARAGROUND "INATH" BARS TO THE FIGURES
	OF THAT YEAR ALONE FOR
OUNI-4 O	·CLECKENG ON THE MINUS
0	SYMBOL WOULD MEMEMBE
	THAT UNE WETH THE REMARKED
	WHES EXAGNOING TO PERC
	DATA CRE CADED
	MONTEN TO CHANGE OD MAN
LAVOUT	HOLD COLOURS FOR DECERDINE
Luiout	USERS PREFERENCES?
FOCUS	
	OPERATION S
3	DISCUSSION
	· DOBS 'MENTINTISTIG' UNE'S AND
	ANY USEFULI INTERALTIONS
	· DUATING THE HOUSE AND
\leq	BREAK DOWN OF DATA SHOW O
	THE YEAR GE DESPLAYED IN
	TOXT FOR FASE OF ENDERDATION
	· SHOW ALL BUTTON COULD BE
1	WARKUIL BUT IT MAY CLUTTER THE DATA.
	. VIDUALISATION ASSUMES THE
TO SCALLS	DIVLY CORECATIONS WILL BE NUMBRIDGE
Lfar	· COULD OTHER DATA OF OTHER AND
	ON THE GARS? PERCENTAGES OF
	DEGREE CLASSEPECATEON BY TYPE?

[J. Roberts et al., 2016]

Sheet 5

D. Koop, CS 490/680, Fall 2019

(a) (a) (a)	TTTLA . ASCHTIONSHERS BETWEEN DA
ianes veitw	AUTHOR:
2004	DATE : 17/11/2012
BUSTWITH B	SHOUT : 4 - FOS 3
ICEAF 20 MOAGAN C NDWR 38 INSPA C	TASK : SENAL DESEAN CONCEPT WENG 3 STATUS, MAP, RUASEDIT MAP 0- DAR CRAPH
and the sound	DEVELOPMENT OFFICEPOTO RECON FOSZ, FOSZ & CLEVITY COMMYECATION
LOCATE MULE OCATEDIN	· All DATA DRG LOADED
REPERTY A	· SWETCH VEEN" BUTTON AROWS USER 20
	SWATCH OF TWEEN STATE & FTATE 2
web view 🗠	(LYDH THANSETERYS TO STATE 3,
2004	A FUIL BREAKDOWN OF THE LINES
KYSTWYTH D	DATA BY YEAR.
MORGAN B	"RETURN" BUTTON WOULD RETURN THE
NOWR B	USAR TO STATE 2.
AL 8 LICATS	· UNIS CAM BE EXCLUDED / INCLUDED
LONPLEFEE EASTER	LA THE VOW (GRATES 162) BY
MAPURELATOX	· YEAR SULECTABLE BY USER
OF LOLATEO	
Pomien 2	
neruers .	OPPOATTONS
an allow	OPERATION
RCOAD CADIN	DETAIL
- CLARC' DATA DAL	. TEME TO BUELD ESTEMATED AT
BAR ALL LUP WALLAR	16 HOURS
or serio Aun o P	· DATA-SETS AQUIRED HEDM
werth assister and	SCHID WHEES
CALLE TO	THE RANGE AS THE PARTY
LAYOUT	IS FRAITEONIS OF %'S.
FOCUS	· COLOUR CHOICE BASED ON
	SUNCEY DATA, HOWEVER WOULD NOT
	MAY CHOSE YOULD PURPLE P. DUCE
, AFTER 1ST YEAR	TO SUET ORITERANDMALY & PROTANOPY
	COLOMA BLENDNESS.
	ALBESGG - YELLOW,
ANGG DEPOYADING	0 B33552 - PURPLE,
W STATE, MARN	· DATA-SET WILL MOED SAME CLOWKING
OH THE BI-POCAR	DATA. ALL UNES HAVE ALL VEARS
DATION AND	MAY NOT DICLUDE STATE THE
THE DUTION	L STATE THREE BASED ON FOR
BE DEFENSIONT	
FACH PATA - ETRA	
E THE SAME UPDER	- 1
its idenation	1

Northern Illinois University 50

Prototype

D. Koop, CS 490/680, Fall 2019

Northern Illinois University 51

