Information Visualization

Introduction

Dr. David Koop





What is Data Visualization?









Data Visualization











Data Visualization



RAND FUNK RAILROAD EMERSON, LAKE, & PALMER ALICE COOPER EDGAR WINTER (DAVID BOWIE) RAND FUNK RAILROAD EMERSON, LAKE, & PALMER PALMER ROBIN TROWED TRAFFIC TRAFFIC
BLACK SABBATH DEEP PURPLE
HEAVYMETAL COM
RAND FUNK RAL ROAD EMERSON, LAKE, & PALMER FOCUS ROBIN TROWER
RAND FUNK TH YES URIAH HEEP BACHMUNITURME OWEN
> RANGE JETHRO TULL
MOUNTAL STATION
CANTANA CANTEWART ON GENES BAND STEVE MUL
REVIVAL COCKER LEON RUSSELL LYNYRD SKYNY
REVISE COCKEN ALLMAN BROTHERS DOOBIE BROTHERS GREGG ALLMAN TEARS SHOCKING BLUE JANIS JOPLIN ALLMAN BROTHERS DOOBIE BROTHERS GREGG ALLMAN SHOCKING BLUE JANIS JOPLIN JAMES GANS STEELY DAN JOHN LENNON PLASTIC OND BAND JAMES GANS STEELY DAN JOHN LENNON PLASTIC OND BLUES JAMES GANS STEELY DAN JOHN LENNON PLASTIC OND BLUES JAMES GANS STEELY DAN DOODY BLUES JAMES GANS STEELY DAN DOODY BLUES JAMES GANS STEELY DAN DOODY BLUES JAMES GANS STEELY DAN
SHOCKING DECK PLASTIC OND BAND JAMES GANS STEELY DAN LOGGINS & MELLMAN
SHOCKING DEN PLASTIC OND DAND JAMES GANS STEELY DAN LOUGINS & MESSINA DANE DANE DANE DANE DANE DANE DANE DA
NORMAN GREENBAUM JOHN GILBERT O'SULLIVAN BETTE MIDLER BLUE SUENT
BLUE SUEDE
TONY ORLANDARY OSMONOS CHUCK BERRY MAUREEN M'GOVERN PAUL ANKA
ANNE RICK NELSON CHUCK DESLEY MAUREEN MOOVERN PAUL ANKA
HON FIRST EDITION HELEN REDOT RINGS DONNY OSMOND AND
LEBERT STORIES STORIES
WW CAN
VANITY FAMILY VANITY FAMILY VANITY FAME VANITY FAME VA
VANITY FAR DOLSE PARTRIDE ARC ARC VOUNG
EPISON LIGHTAGE JOAN MAC DAVIS
VANITY FAIR VANITY FAIR VANITY FAIR VANITY FAIR Series AROUS CAT STEVENS GRATEFUL DEND GRATEFUL DEND
GRATEFUL GOM STEVENSG GAT STEVENSG GAT STEVENSG GRATEFUL GAT STEVENSG GAT STEVENSG GAT STEVENSG GAT STATE FAUL SIMON FAUL SIMON ART GARFUNKLE MARIA MULDAUR ART GARFUNKLE MARIA MULDAUR GARFUNKLE MARIA MULDAUR GARFUNKLE GARFUNKLE MARIA MULDAUR GARFUNKLE GARFUNKLE MARIA MULDAUR GARFUNKLE
MELANIE CAN DENNER JIM CROCE ART GARFUNKLE, MARIA
BADFINGER KRIS KRISTOFFERSON JAMES TAYLOR KRIS KRISTOFFERSON JAMES TAYLOR JONI MITCHELL JONI MITCHELL JONATHAN EDWARD JAMES TAYLOR JONI MITCHELL JONI MITCHELL AMERICA JONI MITCHELL AMERICA JONI MITCHELL O'JAYS O'JAYS DOBIE GRAY R BROOK BENTON AL GREEN STYLISTICS Dauly
NASH JAMES LATEON JONI MITCHING AMERICO ISLET GRAY
R BROOK BENTON AL GREEN STYLISTICS BARRY WITHERS BILLY PAUL BLUE MAGIC
AL WILL
STAIRSTEPS DONNY HATHAWAY
NOMENTS EDWIN STARR BOBERTA FLACK JACKSON FIVE JEAN KNIGHT LUTHER INGRAM VS FREDA PAYNE BRENDA (TH TABULATIONS BETTY WRIGHT EDDIE KENDRICKS SISTER) VS DIALA DAYNE BRENDA (TH TABULATIONS BETTY WRIGHT EDDIE KENDRICKS LOVE UNLIMITED LOVE UNLIMITED
NOMENTS EDWIN STARR ROBERTA FLACK JACKSON FINE JEAN KNIGHT LUTHER INGRAM VS FREDA PAYNE BRENDA & TABULATIONS BETTY WRIGHT EDDIE KENDRICKS DIANA ROSS GLADYS KNIGHTS & TW PIPS LOVE JACKSON MEST
HONEY CONE PAR MICHAEL UNC
UKAMUTIKS SIDE
→ DRAMATICS MARTING PRESTON → STEVIE WONDER
I VER
STARS OR SUNGERS OR JOHNANA BILLY COBHAM
Son Income with a comment
1970 1971 1972 1973 1974
to Pay, R. Garofalo, 1977 (via Ti
to Pay, R. Garofalo, 1977 (via Ti











What is **Information** Visualization (InfoVis)?







Compared to Statistical Graphics?







Streamgraphs

- Visualize movie ticket sales by time
- Stack films that are in theaters on top of each other
- Area = the total sales
- "You can see Oscar contenders attract a smaller audience than the holiday and summer blockbusters and kind of slowly build an audience." — N. Yau, FlowingData



[Byron and Wattenberg, 2012]





Streamgraphs

- [Gelman & Unwin] Instead use two plots: 1. Total sales over time
 - 2. Trajectories for individual movies
- "Discussion burst out across the Web . . . that I am convinced would not have come about if instead of a Streamgraph, they used say, a stacked bar chart." — N. Yau



Northern Illinois University







"That Puzzle-Solving Feeling" — [Gelman & Unwin]







Nightingale's Coxcomb Diagram



- black lines enclosing them.

D. Koop, CSCI 628, Fall 2021



NIU

Northern Illinois University





Gelman and Unwin's Remake







D. Koop, CSCI 628, Fall 2021

Mortality rates in the Crimean War from April 1854 to March 1856

British Army Size in the Crimean War from April 1854 to March 1856



[Gelman and Unwin, 2014]

Northern Illinois University



Compared to Infographics?





America's Most Popular Charts



D. Koop, CSCI 628, Fall 2021





13

Decision Tree: The Obama-Clinton Divide



D. Koop, CSCI 628, Fall 2021

AMANDA COX/ THE NEW YORK TIMES







14

Infographics Embellish Boring Plots?

African Countries by GDP



GDP in billions of US dollars











Infographics Embellish Boring Plots?



African Countries by GDP

GDP in billions of US dollars

D. Koop, CSCI 628, Fall 2021

African Countries by GDP

TOP COUNTRIES BY GDP IN U.S. \$ BILLIONS

Gross domestic product (GDP) refers to the market value of all final goods and services produced within a country in a given period (2005 - 2009)

GDP CALCULATION

NIU

private consumption + gross investment + government spending + (exports – imports







Compared to Scientific Visualization?





SciVis



D. Koop, CSCI 628, Fall 2021













17

InfoVis







SciVis \rightarrow Fields





Scalar Fields (Order-0 Tensor Fields)



Each point in space has an associated...

 s_0

Scalar

D. Koop, CSCI 628, Fall 2021

Vector Fields (Order-1 Tensor Fields)





σ_{00}	σ_{01}	σ_{02}	
σ_{10}	σ_{11}	σ_{12}	
σ_{20}	σ_{21}	σ_{22}	
Tensor			





Visualization Taxonomy Structure

		Display Attributes	
	Given	Constrained	Chosen
Continuous	 Images (e.g., medical) Fluid / gas flow, pressure distributions Molecular structures (distributions of mass, charge, etc.) Globe – distribution data (e.g., elevation levels) 	Distortions of given / continuous ideas (e.g., flattened medical structures, 2D geographic maps, fish-eye lens views) Arrangement of numeric variable values	Continuous (high-dimensional) mathematical functions Continuous time-varying data, when time is mapped to a spatial dimension Regression analyses
Discrete	Classified data / images (e.g., segmented medical images) Air traffic positions Molecular structures (exact positions of components) Globe – discrete entity data (e.g., city locations)	Distortions of given / discrete ideas (e.g., 2D geographic maps, fish-eye lens views) Arrangement of ordinal or numeric variable values	Discrete time-varying data, when time is mapped to a spatial dimension Arbitrary entity-relationship data (e.g., file structures) Arbitrary multi-dimensional data (e.g., employment statistics)









Visualization Taxonomy Structure

SciVis		Display Attributes			
		Given	Constrained	Chosen	
	Continuous	 Images (e.g., medical) Fluid / gas flow, pressure distributions Molecular structures (distributions of mass, charge, etc.) Globe – distribution data (e.g., elevation levels) 	Distortions of given / continuous ideas (e.g., flattened medical structures, 2D geographic maps, fish-eye lens views) Arrangement of numeric variable values	Continuous (high-dimensional) mathematical functions Continuous time-varying data, when time is mapped to a spatial dimension Regression analyses	
	Discrete	Classified data / images (e.g., segmented medical images) Air traffic positions Molecular structures (exact positions of components) Globe – discrete entity data (e.g., city locations)	Distortions of given / discrete ideas (e.g., 2D geographic maps, fish-eye lens views) Arrangement of ordinal or numeric variable values	 Discrete time-varying data, when time is mapped to a spatial dimension Arbitrary entity-relationship data (e.g., file structures) Arbitrary multi-dimensional data (e.g., employment statistics) 	









Visualization Taxonomy Structure

Coil/io	C	Display Attributes			
SciVis		Given	Constrained	Chosen	
		 Images (e.g., medical) Fluid / gas flow, pressure distributions Molecular structures (distributions of mass, charge, etc.) Globe – distribution data (e.g., elevation levels) 	Distortions of given / continuous ideas (e.g., flattened medical structures, 2D geographic maps, fish-eye lens views) Arrangement of numeric variable values	Continuous (high-dimensional) mathematical functions Continuous time-varying data, when time is mapped to a spatial dimension Regression analyses	
Discroto		Classified data / images (e.g., segmented medical images) Air traffic positions Molecular structures (exact positions of components) Globe – discrete entity data (e.g., city locations)	Distortions of given / discrete ideas (e.g., 2D geographic maps, fish-eye lens views) Arrangement of ordinal or numeric variable values	 Discrete time-varying data, when time is mapped to a spatial dimension Arbitrary entity-relationship data (e.g., file structures) Arbitrary multi-dimensional data (e.g., employment statistics) 	











Kosara's Definition of Information Visualization

- result, it is not being visualized."
- the visualization must be able to stand on its own."

• It is based on (non-visual) data. "The data to be visualized must come from outside the program, and the program must be able (at least in principle) to work on different data sets. Also, visualization is not image processing or photography; if the source data is an image and is used as an image in the

• It produces an image. "Clearly, each visualization has the goal of producing one or more images from the data, and the visual must be the primary means of communicating the data. Other media can be part of a visualization, but















Kosara's Definition of Information Visualization

• The result is readable and recognizable. "There are many ways to transform data into images, most of which do not allow the viewer to else. The use of additional elements (or even "eye candy") is certainly visualization."

understand the underlying data. A visualization must produce images that are readable by a viewer, even if that requires training and practice. Visualization images must also be recognizable as such, and not appear to be something possible, but must not take precedence over the communication goals of the















Pragmatic <-> Artistic Visualization









Rules

- We saw many rules in CSCI 627 (Data Visualization)
- How do we use those to think about visualization?







Visualization Mistakes

Original

Left-click

Average number of likes per Facebook post 2016













Visualization Mistakes

Original

Left-click

Average number of likes per Facebook post 2016



D. Koop, CSCI 628, Fall 2021

Better

Left-click

Average number of likes per Facebook post 2016, '000











Can we break the rules?







Benefitting InfoVis with Visual Difficulties

Cognitive efficiency



Cognitive operations

Minimize the cognitive steps required to process visualization

Induce constructive, self-directed cognitive activity on the part of the user

ta-ink ratio Dat

Maximize the ratio of data to ink

Design representations that are most likely to engage a user to actively process the information

D. Koop, CSCI 628, Fall 2021

Visual difficulties









Benefitting InfoVis with Visual Difficulties

Organization	Choose the format which makes important information most visually salient	Cho dee data
Animation	Use animation to quickly and intuitively visualize important processes	Use inte cau case use
Labeling	Use labels rather than legends to optimize immediate clarity	Use on c

D. Koop, CSCI 628, Fall 2021

oose the format that best stimulates ep cognitive reflection on the important a

e static representations to induce erval visualization processing around usal mechanisms; consider animation in ses where mental animation lies beyond ers' capacities

e legends to stimulate deeper reflection data













Administrivia







About Me

- Research Interests
 - Visualization
 - Computational Provenance
 - Geospatial Analysis
- Research Projects
 - Dataflow Notebooks
 - Geospatial Trajectory Data
 - Provenance for Web Applications
- See my web page for more information
 - http://faculty.cs.niu.edu/~dakoop/







About You

- Research Papers?
- Data Visualization
- Tools? JavaScript, D3, Tableau, Others?
- Research Experience?
- What topics do you want to see covered?







About this course

- Course web page is authoritative:
 - http://faculty.cs.niu.edu/~dakoop/cs628-2021fa/
 - Schedule, Readings, Assignments will be posted online
 - Check the web site before emailing me
- Lectures: TuTh 9:30-10:45am in PM 252
- This is an Advanced (Tier 2) Graduate Course
 - Present and discuss cutting-edge topics
 - Work on research problems
- Requires participation: readings and discussions







About this Course

latest information visualization techniques developed by the research laboratory work."

 "Focus on advanced theory and methods for manipulating and visualizing the data of non-physical systems... Emphasis on the advanced study of the community. A computer programming background is required. Extensive









About this course

- Course Registration:
 - Make sure you have registered for the course
 - Email me if you are not registered but are interested in taking the course
- Review of course policies:
 - Plagiarism and academic honesty
- If you have any concerns or questions, please email me as soon as possible • If you are not sure if this course is a good fit, please email me or talk to me









Office Hours & Email

- Office hours will be held in person - Tu: 1:45-3pm, Th: 10:45am-12pm, or by appointment
- Please adhere to university regulations (Protecting the Pack)
- You do not need an appointment to stop by during scheduled office hours
- If you wish to meet virtually, please schedule an appointment
- If you need an appointment, please email me with **details** about what you wish to discuss and times that would work for you
- Many questions can be answered via email. Please consider writing an email before scheduling a meeting.







Expectations

- Be engaged:
 - Active participation
 - Constructive participation
- Work independently: self-directed and sustained
- Work collaboratively: learn from each other
- Put effort into the course:
 - Must put significant work in **each** week
 - Do not try to do everything before a deadline
 - Best effort (success or failure)







Next Class

- Review of Core Visualization Topics
- Topics Survey
- Bring Your Ideas



