CS-5630 / CS-6630 Uisualization for Data Science

Alexander Lex <u>alex@sci.utah.edu</u>







pictures visualization The purpose of computing is insight, not numbers.

- Richard Wesley Hamming - Card, Mackinlay, Shneiderman

M. acuminata Banana Date P. dactylifera Cress Arabidopsis thaliana Rice Oryza sativa Sorghum Sorghum bicolor Brachypodium distachyon Brome



(

Sorghum bico 34,496 / 27,39

Phoenix dactylifera 28,889 / 19,027

Dryza sativa 10,612 / 27,049

Arabidopsis thaliana 27,169 / 21,950



[D'Hont et al., Nature, 20



)12



vi · su · al · i · za · tion
I. Formation of mental visual images
2. The act or process of interpreting in visual terms or of putting into visible form

Visualization Definition

Visualization is the process that transforms (abstract) data into interactive graphical representations for the purpose of exploration, confirmation, or presentation.

Good Data Visualization

... makes data accessible ... combines strengths of humans and computers ... enables insight ... communicates

Uisualization

"Visualization is really about external cognition, that is, how resources outside the mind can be used to boost the cognitive capabilities of the mind."



Stuart Card

Why Visualize?

To inform humans: Communication

How is ahead in the election polls?

When questions are not well defined: Exploration

What is the structure of a terrorist network?

Which drug can help patient X?

Purpose of Visualization





Confirmation



U.S. JOB LOSS (DEC. 2007 TO JAN. 2010)

Communication

Example Communication

Be Hard to Beat

The Broncos quarterback set the all-time N.F.L. touchdown passing record - and is still going strong.



[New York Times]

Example Exploration: Cancer Subtypes



[Caleydo StratomeX]

_



Why Graphics?

Figures are richer; provide more information with less clutter and in less space.

Figures provide the gestalt effect: they give an overview; make structure more visible.

Figures are more accessible, easier to understand, faster to grasp, more comprehensible, more memorable, more fun, and less formal.

list adapted from: [Stasko et al. 1998]

Total Bandwidth

(millions of bits per second)













the public ochoold note on a contraction of the OVcity's main public hospital was a wreck, for and the city's public-housing projects were shuttered. are Campanella then switched to an the identically constructed map, only this mtime based on 2010 census data, and etin bits and pieces on the screen there ve was a simple and arresting picture of he riwhat Katrina meant. In the neighborhoods that were once a dense black, res many of the little squares had thinned ite and turned gray. The sharp lines that m. once separated the teapot from Central City were now blurry: the white gareas of the city were pushing north, ke into the vacuum left by the exodus. rh The Bywater was graying, as it genre trified still further. "Before Katrina, an ne American Community Survey estin mate of New Orleans Parish populag tion was four hundred and fifty-five d thousand, and about sixty-eight per y cent black," Campanella said. "Now the latest estimate is three hundred and eighty-four thousand, and it's about 0

15

Textual description of a map of the effects of hurricane Katrina on New Orleans. New Yorker, posted by Alberto Cairo



When not to visualize? When to automate?

Well defined question on well-defined dataset

Which gene is most frequently mutated in this set of patients?

What is the current unemployment rate?

Decisions needed in minimal time

High frequency stock market trading: which stock to buy/sell? Manufacturing: is bottle broken?



The Ability Matrix



Insight is generated by the human – not the computer!

Planning Diagnosis Prediction

> Cognition Common Knowledge Creativity

Why Use Computers?

Scale

Drawing by hand (or Illustrator) infeasible inflexible (updates!) How to draw an MRI scan?



[Bruckner 2007]

Why Use Computers?

Interaction

Interaction allows to "drill down" into data

Integration

Integration with algorithms

Make visualization part of a data analysis pipeline



[Sunburst by John Stasko, Implementation in Caleydo by Christian Partl]





Why User Computers?

Efficiency

Re-use charts / methods for different datasets

Quality

- Precise data driven rendering
- Storytelling
 - Use time

Tell Stories





Why not just use Statistics? IV Ι III II Х X V X V X V 8 6.5 10 9.1 10 7.4 10 8.0 8 6.9 8 5.7 8 8.1 8 6.7 13 7.5 13 12. 13 8.7 87.7 98.8 98.7 97.1 8.8.8 11 8.3 8 8.4 11 9.2 11 7.8 14 9.9 87.0 14 8.1 14 8.8 6 6.1 6 7.2 8 5.2 66.0 4 4.2 4 3.1 19 12. 45.3 12 10. 12 9.1 8 5.5 12 8.1 7 6 1 772 87.9 4.8 5 5 **Mean x: 9 y: 7.50** 6.8 Variance x: 11 y: 4.122 **Correlation x – y: 0.816** Linear regression: y = 3.00 + 0.500x

Anscombe's Quartett



Mean x: 9 y: 7.50 Variance x: 11 y: 4.122 Correlation x – y: 0.816 Linear regression: y = 3.00 + 0.500x





Simulated Annealing, CHI 2017, Justin Matejka, George Fitzmaurice

Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through

Visualization =

Human Data Interaction

Human-Data Interaction

Data

Visualization in the Data Science Process



Figure 2-2. The data science process

Big Data

2010: 1,200 exabytes, largely unstructured Google stores ~10 exabytes (2013) Hard disk industry ships ~8 exabytes/year



Funniest Dog Videos https://www.youtube.com ...

15 Exabytes in Punch Cards:4.5 km over New England





798 Instagram photos uploaded in 1 second 🕠



1,277 Tumblr posts in 1 second





Example: Personal Data

1-25 DM



Big Data in Science and Engineering

"Big Data" hasn't just transformed industry! have changed the way science and engineering are done.

Examples:

- Large physics experiments and observations
- Cheaper and automated genome sequencing
- Smart buildings / cities (blyncsy)
- Geophysical imaging

Controversy: Hypothesis or data driven methods

- It's also transformed science and engineering. Cheap sensors (e.g. imaging)



Example: CERN Large Hadron Collider Data

CERN has publicly released over 300TB of data: <u>CERN Open Data Portal</u>

How much is that?

- At 15 GB of storage a piece, you'd need **20,000 Gmail** accounts to store the whole shebang. If you wanted to send that much data at the max attachment size of 25 MB, it would take you **12** million emails.
- A DVD-R holds 4.7 GB. You'd need 63,830 of them to hold 300 TB.
- Your Blu-ray collection wouldn't need to expand quite so much. 6,000 discs ought to hold it.
- It takes Pandora about a day and a half to burn through a gig of mobile data. So if the CERN data was an album, you could stream it in just over 1,230 years.
- At 350 MB per hour for 4K video streaming, so if the CERN data was a 4K movie it'd probably be about 857,142 hours, or about 98 years long.
- But it ain't no thing compared to what the National Security Agency works with. Going by 2013 figures the agency released, the NSA's various activities "touch" 300 TB of data every 15 minutes or so

(Popular Mechanics Article)

Example: Genomics



542



23andV

Protostomes protogram and protocol in the second se



Global Ice Ages Oceans Rust Earth Birth 1000 2000 3000 Millions of Years Ago 3000 2000 7009 542 440 370 250 200 65 Today

All the major and many of the minor living branches of life are shown on this diagram, but only a few of these that have gone extinct are shown. Example: Dinosaurs - extinct



NSA Utah Data Center (Bluffdale, Utah)

Storage Capacity?

estimates vary, but <u>Forbes</u> <u>magazine</u> estimates 12 exabytes (12,000 petabytes or 12 million terabytes)



"The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it that's going to be a hugely important skill in the next decades, ... because now we really do have essentially free and ubiquitous data."

> Hal Varian, Google's Chief Economist The McKinsey Quarterly, Jan 2009

Humans! Human Data Interaction

Why Humans?

- Leveraging human capabilities
 - Pattern Discovery: clusters, outliers, trends
 - **Contextual Knowledge:** expectations for dataset, explanations for patterns
 - **Action:** humans learn and take action
- But: we also have to design for Humans and their limitations
Not everything that can be drawn can be read!





Limits of Cognition



Daniel J. Simons and Daniel T. Levin, Failure to detect changes to people during a real world interaction, 1998

How did we get here? A bit of history

"It is things that make us smart"



Donald A. Norman

erte l'erte que una mil farther. 1 - part. (45 A) on I the short the redaction of All and the method of the second seco Farmer and the second s after ber fing an bour preis strengen . Tan merna pas hounth bonte nolumpi Bridnany es Blande 10-14 F.Lu. Con age of the Inthe St. mbacom il untere epoten-tary base depeton ali anti unter epitin la pare energion ali bai e ana di bi basari inagnobiji bi ali bai e ana di bi basari inagnobiji bi ali bai e ana di bi basari inagnobiji bi anti gaor angus Ner C. Nedunovi prutte ab est antam, pet a comino des tre re amarnie Questione tane de anten 4. f pretipram justam re prapit abeier mite. 1 the second cant . and is the way as The second defendence of the second matte mittermite mate. Dera mindt. Of tott/ peccater minds and it some and the sector and .

The first Brak of 10 For DGf=DE, redDAg= DC. When-fore AGA=CEA=BCI=AG. Black not Ars.det. g. mayler, h. j. may. h. my det. The patting of the point Awathia or without the line B C varies the cafes, but the coalisation, and the democalisation, are every where abks. Schol. The Lise A G might be taken with a pair of teen-paffers but the folding antivers to no puttulare , as Prodes will interacts. PHON. III. Bus def.
B D fail out of B B^D = BD s = A d= BE.
adde and character of a Data PROP. IV. A D If two triangles B A Co E D F , bette two film of the ant BA. A C equal rates of the other ED. DF reak as his correspondent file (thes is EA == ED.

<text><text><text><text><text>

PROP. V.

The ampier A E C, A C B, at also have of an Epidecien triangle A B C, one repeat can to the miser, and of the repeat of the AB, A C be producted, the ample CB D, B C E, under the have builte result one to the effer. "Take A E = A D; and f join at the C D, and B E. C D; and B E. Eccuate, in the triangles ACDs styr

C A Desardo taries a la ser A Gand A E simple a A Desardo de angle A commentantien borb, servi-inderstare in the angle A service and then borb, servi-inderstare in the angle A be made to the CD and the angle A De and the angle A be made to the CD and the angle A De and the angle A be made to the CD and the angle A De and the angle A be made to the CD and the C Tase. F mail be the angle ECR = DRC. Evidence of the Alle the angle ECR = DRC. Evidence of the Alle the angle ECR = DRC. Evidence of the Alle the angle ECR = DRC. Evidence of the angle A E D = A C Detendence the angle A B C angle A E D = A C Detendence the angle A B C angle A C B P bies out set in from (a-



Record



Konya town map, Turkey, c. 6200 BC



Anaximander of Miletus, c. 550 BC

Milestones Project

Record



Leonardo Da Vinci, ca. 1500



Galileo Galilei, 1616 Donald Norman



William Curtis (1746-1799)

The History of Visual Communication The Galileo Project, Rice University

Record



Eadweard J. Muybridge, 1878



Analyze



Planetary Movement Diagram, c. 950



Analyze

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.









The Upright divisions are Ten Thousand Lounds each . The Black Lines are Exports



proportions of the Turkish Empire located in Asia, Europe and Africa before 1789



Find Patterns



John Snow, 1854

Communicate



Aung. per Regnier, S. Pas. 5" Karis S. Ort & Paris.

C.J. Minard, 1869

Ing Sid. Repair & Dowites .

Communicate



London Subway Map, 1927



An Overhaul of an Underground Icon

Next month, the Metropolitan Transportation Authority will unveil a remore than a decade. Related Article »



Next month, the Metropolitan Transportation Authority will unveil a resized, recolored and simplified edition of the well-known map, its first overhaul in

Interact



Ivan Sutherland, Sketchpad, 1963





Doug Engelbart, 1968



Modern Examples

Analyze



M. Wattenberg, 2005

Communicate



Hans Rosling, TED 2006

Who is CS-5630 / CS-6630?

http://alexander-lex.net Alexander Lex Assistant Professor, Computer Science Before that: Lecturer, Postdoctoral Fellow, Harvard PhD in Computer Science, Graz University of Technology



<u>@alexander lex</u>





visualization design lab





Miriah Meyer

Alexander Lex



http://vdl.sci.utah.edu/



Pascal Goffin



Ethan Kerzner



Alex Bigelow



Nina McCurdy



Jennifer Rogers



Sam Quinan



Jimmy Moore



Cameron Waller



Carolina Nobre

We're looking for PhD **Students!**





Miriah Meyer

Alexander Lex





visualization design lab





SCI Institute Scientific Computing and Imaging Institute Scientific Computing **Biomedical Computing** Scientific Visualization Information Visualization Image Analysis



http://sci.utah.edu



The SCI Institute is an internationally recognized leader in visualization, scientific computing, and image analysis. Our overarching research objective is to create new scientific computing techniques, tools, and systems that enable solutions to problems affecting various aspects of human life.

	SCI Home The Institute Research Centers Medi
--	--

News from the SCI Institute Research News | Software News | Honors/Awards | Publications | All SCI News





SCIRun 5.0 Released



Login Admin

Search...





Seg3D 2.2.0 Now Available Jul 01, 2015



Jun 25, 2015

Big Scientific Data Made Simple Jun 23, 2015

SCI Events

<			Aug 20	15		>
s	М	Т	w	Т	F	S
26	27	28	29	30	31	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

RSS 2.0 FEED

View all SCI Events

Upcoming SCI Events

Large, Multivariate (Biological) Networks





Collapse rows		Connectiv	ity Matrix				+											
state -																		
Collapse		+ degree	city	state	market			•			3	~	.,		-			.
nala		L				ic		ž i	6 Q.	86	Ξ	¥Ξ	₩8	1	23	а	Ξ.	
Cula	- 1 -	000 0	C	05														
none 🔻		10	Oakland	05	39457	OAK				-								
		I	CL os Angeles	05	32575	LAX												
		1 0	Gan Diego	05	03570	GAN												
Sort order		φ	Monterey	05	34599	MEN												
		Q.	Conta Barbara	05	34669	GDA												
Select+		1 0	Gan Franciaco	05	0P457	810												
		ρ	Ontario	05	02575	ONT												
Call		Φ	Gan Luis Obiapo	05	04906	SBP												
- Cont		10	Gan Jose	05	02452	8.0												
encoding		10	Georgeoto	05	33182	GME												
onlorm ar-		0	Ereano	04	91690	GAT .												
colonnap		ц С	Orescent City	04	30800	ano												
		p	Burbeck	05	02575	nue												
Path metric		φ C	Carlshad	04	01041	GI D												
		P	Palm Springs	04	34969	P3P												
path count	t	10	Ganta Ana	05	32575	GNA												
		ф С	Artata/Funeka	05	30157	AOV												
		Ψ.	Dekerafield	05	33561	nin -												
Legend		9	Mammoth Lakes	05	33060	MMU												
260	105	ě.	Long Deach	04	02575	LGD												
200	103	4	Hedding	04	31782	HIDD												
559	209	ě o	Genta Mana	05	06805	OWK			۰.									
838	313			08														
1117	417	6	Portland	08	00007	MCD												
		å	weatord Development	08	0.000	NU H												
1396	52]	ů.	Concernanced	00	01000	0.00												
1675	625	. n n	1 UBono	COR .	11003	1.000												
1955	729	á		THE	04060	000												
- 222	129	-4.	PARCO/MEDDEWIC	SECONDCC.	100.00	P 0.0												









Set Visualization

Multidimensional Data

Multivariate Rankings

5			I		World Univer	rsity Rank	king				Wor	d Unive	rsity Rank	ing
				17.99%	32.94%	19.63%	19.63%	4, 4,			40.00%	10.00%	20.00%	
Relat	Rank	School Name	Country	Acade	Employer repu	Faculty/	Citation	I I	Separator	Rank	Academic reputation	Empl	Faculty/st	:ud
Date	1.	Massachusetts Insti	United States						A A	1.				
1995	2.	Harvard University	United States							2.				
) 1995	З.	University of Camb	United Kingdom							3.				
1995	4.	Imperial College L	United Kingdom						\sim	4.				
1995	5.	University of Oxfor	United Kingdom							5.				
1995	б.	UCL (University Col	United Kingdom						\sim	6.				
	7.	Stanford University	United States							7.				
	8.	Yale University	United States							8.				
	9.	Princeton Universit	United States						·	9.				
	10.	University of Chica	United States							10.				
	11.	ETH Zurich (Swiss F	Switzerland						/	11.				
	12.	Columbia Universit	United States						$\sim \gamma$	12.				
	13.	University of Penns	United States							13.				
	14.	Cornell University	United States							14.				
	15.	University of Edinb	United Kingdom						$\sim T$	15.				
	16.	Ecole Polytechniqu	Switzerland							16.				
	17.	King's College Lond	United Kingdom		93.7 (0.94)					17.				
	18.	University of Toron	Canada							18.				
	19.	McGill University	Canada						\sim	19.	94.6 (0.95)	89	8.9 (0.9)	79.
	20.	National University	Singapore						\sim	20.				
	21.	University of Michi	United States						$A \sim 1$	21.				
	22.	University of Califo	United States						$K \times \gamma$	22.				
	23	California Institute	United States							23.				
	24	University of Bristol	United Kingdom							24.				
	25.	Duke University	United States						r i	25.				







Cancer Subtypes / Omics Clustering and Stratification



Genomic Data

Alternative Splicing / mRNA-seq





5555555555 acolec estimate (TPV) oup GBM Offic2552-833a-4e8f-9... IC2a0737-7c27-46b8._

Genealogies & Clinical Data





								(True)	ression	orespior	(a)	alai	- colar				2.8
				1	ନ୍ତ୍ର	ased	Se one	ein noise	Jeep. Diag	Jeropar	(Thu all BH	P ^D Disq.B	94°		æ	White O	the . use deal
		KindredID	RelativeID	ser	Se.	SUL	Oex	PS Tatal:32	Total:60	24	ي ي Total:30	Total:60	MaxBMI Totab69	AgeMaxBMI Total:58	<u>98%</u>	(BD	CBL
	7017			44%	60%	10%	52%	h.,	<u> </u>	48%	ահո	L	ي سالي			0.0%	
	2	0190.99	∇_{Δ}	<u>-</u>	-	-	-	0 65	i <u> </u>	$\simeq 1$	- <u>-</u>	0 20	`13 <u>4</u> 5	0 30	V A	V .A.	∇_{Δ}
		913938	#12789	Ξ.	1	Ξ.	Ξ.			71					- 21	_	aunebot waun
	3	903988	#10/09		Ξ.	Ξ.	Ξ.			÷.,		a a0	10 4	<i>d</i> a	10	_	gunanot wount.
	11-	903988	#48154	Ē.			1			Ē. 1						_	
		903988	#54315	ī.						Ē. 1						-	
	1	903988	#67703					_	_		_	_		_		-	
	1	903988	#9278													-	
	1	903988	#26695					—	—		—	_				-	asphyxia
	7	903988		н.				- 4/4			m+#	144	- 58	46.0 1		-	
a	1	903988	#40329						0			Ð	8			-	carbon monox
	1	903988	#5170													-	
	1	903988	#43983													-	
	1	903988	#5181													-	
	1	903988	#42479						_		_	_				-	
	1	903988	#139						8							-	
	1	903988	#1967					_	_		—					-	cardiac arre
	1	903988	#5172													-	
	1	903988	#5171													-	
		903988	#10841	а.						5.5					- 21		
	1	903988	#56710	2	-		-	0					0	-0	- 21	-	
_		903988	#5436	Ξ.	Ξ.	Ξ.	21	_	_	÷.,					- 21		
		903988	#65130	÷.,	21	Ξ.	21			20					- 21		strangulatio
		903988	#65129	21	21		21			÷.,					- 21	-	
		9/3938	#0428				1			11						_	
	10.0	903988	#65491				Ξ.	_	_	71	_	_	0	0	- 61	-	
		903988	#39045	1	Ξ.		Ξ.			11							
		903988	#18524													-	
		903988	#4556													-	
		903988	#4551						0							-	
	1	903988	#52034													-	
	1	903988	#62036													-	
	1	903988	#17581													-	
	1	903988	#47083													-	
	1	903988	#47087													-	
	1	903988	#61319													-	
	1	903988	#61318													-	

Course Staff



Carolina Nobre Teaching Mentee



TBA Teaching Assistant

Pranav Dommata

Teaching Assistant

Flbout You

Structure & Goals

Course Goals. You will learn:

How to efficiently visualize data **Evaluate** and **critique** visualization designs **Apply** fundamental principles & techniques **Design** visual data analysis solutions **Implement** interactive data visualizations Web development skills

Course Components Lectures: introduce theory **Design Critiques:** develop "an eye" for vis design, critique, learn by example Labs: short coding tutorials, examples Based on a published script on website Strongly related to homework assignments **Homeworks** help practice specific skills Final Project gives you a chance to go through a complete vis project

Course Components

Design Lecture Design Studios



Theory

Lecture Reading Discussion

> Labs D3 reading Self-study **Office hours**

Design Skills - Coding Skills

<!DOCTYPE html> <meta charset="utf-8"> <style>

text { font: 10px sans-serif;

</style> <body> <script src="http://d3js.org/d3.v3.min.js"></script> <script>



Schedule

Lectures: Tuesday and Thursday 2:00-3:20 pm, L101 WEB Labs: Wednesday, 6:00-7:30 pm, L110 WEB (scheduled on demand)

Online Students:

YouTube Channel

Three Parts:

I. Technical Foundations HTML, Javascript, D3

II. Visualization Fundamentals

Perception, Visual encodings, Design Guidelines, Tasks..

III. Abstract Data Visualization

Tables, Graphs, Maps

Schedule

CS 5630/6630

Today		August 2017	•				Print Week	Month	Ag
	Mon	Tue		Wed	Thu	Fri	Sat	5	Sun
	31		Aug 1	2	3	4	5	5	
	7		8	9	10	11	12	2	
	14	•	15	16	17	18	19)	
	21	14:00 VIS Lectu	22 Ire	23 14:00 Alex Lex Office	24 14:00 VIS Lecture	25 Homework 0 Due	26	5	
	28	14:00 VIS Lectu	29 Ire	30 14:00 Alex Lex Office	31 14:00 VIS Lecture	Sep 1 Add/Drop Deadline Homework 1 Due	2	2	
Events	shown in time :	zone: Mountain Ti	ime					+ G <mark>oo</mark> gle	<mark>e</mark> Ca

Subject to change

Week 1

Lecture 1: Introduction

Tuesday, August 22, 2017

What is visualization? Why is it important? Who are we? Course overview.

Recommended reading

 A Tour through the Visualization Zoo. Jeffrey Heer, Michael Bostock, Vadim Ogievetsky. Communications of the ACM, FO 07 1 0010


Information <u>http://dataviscourse.net</u>

Uisualization for Data Science cs-5630 / cs-6630

Home Syllabus Schedule Project Resources Fame



The amount and complexity of information produced in science, engineering, business, and everyday human activity is increasing at staggering rates. The goal of this course is to expose you to visual representation methods and techniques that increase the understanding of complex data. Visualization for data discovery and communication is an important part of the data science pipeline. Good visualizations not only present a visual interpretation of data, but do so by improving comprehension, communication, and decision making.

In this course you will learn about the fundamentals of perception, the theory of visualization, good design practices for visualization, and how to develop your own web-based visualizations using HTML5, CSS, JavaScript, SVG, and D3.

The course begins by bootstrapping your web development skills, moves on to fundamentals of perception, introduces data types you will encounter, and then focuses on visualization techniques and methods for a broad range of data types. An integral component of the course are regular design critiques and redesigns that will hone your skills in understanding, critiquing and developing visualization techniques.

The course is offered in the fall term 2017 at the University of Utah in two variants: CS-5630 for undergraduates and CS-6630 for graduate students, with a special section of CS-6630 (002) designated for data certificate students. Classes start on Tuesday.



UpSet visualizing intersecting sets | Wind map | How states have shifted

Companion Course: Visualization for Scientific Data



CS 5635 / CS 6635 Chris Johnson Spring 2018





Communicate

Slack

- http://dataviscourse2017.slack.com/
- **Please use slack for all general questions code, concepts, etc. Only use e-mail for personal inquiries**

Canvas

- https://utah.instructure.com/courses/448047
- **Homework submissions, Grades**
- **Office Hours**
 - **Alex: Wednesdays at 2pm, WEB 3887**
 - **TAs: starting next week**
- **E-Mail**
 - alex@sci.utah.edu

Required Books

An Introduction to Designing With D3

Interactive Data Visualization

for the Web

O'REILLY[®]

Scott Murray





Tamara Munzner



Programming









Data-Driven Documents

Is this course for me ???



Prerequisites

Programming experience C, C++, Java, Python, etc. Willingness to think about user-centered design This is not your average CS course! We care about the human in the loop! Willingness to learn new software & tools This can be time consuming You will need to build skills by yourself! **Engineering vs Computer Science**

Formalities

- How are you graded? 6 Homework Assignments: 40% Varying value, 2%-10%, depending on length/difficult Start early! Will take long if you don't know JS/D3 yet Due on Fridays, late days: -10% per day, up to two days. Final Project: 40% Teams, proposal and two milestones Exams: 20%
 - Two exams: last class before fall break and end of term

Cheating

You are welcome to **discuss** the course's ideas, material, and homework with others in order to better understand it, but the work you turn in must be your own (or for the project, yours and your teammate's). For example, you must write your own code, design your own visualizations, and critically evaluate the results in your own words. You may not submit the same or similar work to this course that you have submitted or will submit to another. Nor may you provide or make available solutions to homeworks to individuals who take or may take this course in the future.

Will automatically check for plagiarism in all your submissions

No Device Policy

No Computers, Tablets, Phones in lecture hall except when used for exercises Switch off, mute, flight mode Why? It's better to take note by hand Notifications are designed to grab your attention Applies to Theory lectures, coding along in technical lectures encouraged

This Week

HWO, including course survey Lecture on Perception Readings D3 Book, Chapters 1-3 VDA Book, Chapter 1

Preface. 1. Introduction. Why Data Visualization? Why Write Code? Why Interactive? Why on the Web? What This Book Is Who You Are What This Book Is Not Using Sample Code Thank You Introducing D3. What It Does What It Doesn't Do Origins and Context Alternatives Easy Charls Graph Visualizations Geomapping Almost from Scratch Three-Dimensional Tools Built with D3 3. Technology Fundamentals. The Web HTML Content Plus Structure



Next Week

HW1 due Introduction to Git, HTML, CSS Office hours start!

https://github.com/dataviscourse/2017-dataviscourse-homework/

lataviaaaura		Pull request	s Issues Gist	🦨 +- 🎯
Jataviscourse	ə / 2016-dataviso	ourse-homework	💿 Unwa	atch = 8 🚖 Star 0 💱 Fork 1
> Code ① Is	sues 0 🔅 Pull re	quests o 🖂 Wiki 👍 Puls	se 🔟 Graphs 🔅 Settings	
mework mater	ial for 2016 Data vie	sualization course. — Edit		
② 12 commits ② 1 branch		○ 0 releases	L2 3 contributors	
anch: master •	New pull request		Create new file Upload	files Find file Clone or download -
alexsb committ	ed on GitHub Removed	github username sentence.		Latest commit 19c7871 a day ago
I hw0	Remo	ved github username sentence.		a day ago
i hw1	minor	updates to HW1		4 days ago
README.md	addec	redame info; wip on hw0 revision	ns	4 days ago
More inform	nation can be found	on the course website		
We sugges		ository:		
	t you clone this rep			
git clone	t you clone this rep	/dataviscourse/2016-dataviscou	urse-homework	
git clone To receive d dataviscour:	t you clone this rep https://github.com updates and add ne se-homework directo	/dataviscourse/2016-dataviscou wly released homeworks upo ry and running:	unse-homework date youre repository by od-ing	into the 2016-
git clone To receive a dataviscour: git updat	t you clone this rep https://github.com updates and add ne se-homework director	/dataviscourse/2016-dataviscou wly released homeworks upo ry and running:	unse-homework date youre repository by od-ing	into the 2016-
git clone To receive a dataviscour git updat	t you clone this rep https://github.com updates and add ne se-homework directo	/dataviscourse/2016-dataviscou ewly released homeworks upo ry and running:	unse-homework date youre repository by od-ing	into the 2015-
git clone To receive a dataviscour git updat	t you clone this rep https://github.com updates and add ne se-homework directo	/dataviscourse/2016-dataviscou ewly released homeworks upo ry and running:	unse-homework date youre repository by od-ing	into the 2015-

New Track: Human Centered Computing **NON-CS COURSES** Design

REQUIRED COURSES

CS 6540 - HCI (humans + interfaces)

- CS 6xxx Advanced HCI (humans + things)
- CS 6630 Visualization for Data Science (humans + data)

ED PSY 6010: Introduction to Stats and Research Design PSY 6140 - Cognitive Neuroscience Approaches to Research (methods) PSY 6420 - Methods in Social Psychology

ELECTIVES

Pre-approved course list from within CS and across campus

Up to 3 electives can be taken from outside CS

DES 5320 - Typographic Communication

DES 5370 - Digital Fabrication

DES 5710 - Product Design and Development

Ed Psych

ED PSY 6030 - Introduction to Research Design

Psych

PSY 6120 - Advanced Human Cognition

PSY 6700 - Neuropsychology

Anthropology

ANTH 6169 - Ethnographic Methods

Sociology

SOC 6110 - Methods of Social Research

EAE

EAE 6900 - Games User Research

EAE 6900 - A.I. For Games