Data Visualization (CSCI 627/490)

Introduction

Dr. David Koop





What is Data Visualization?









How is it different from Computer Graphics?







The purpose of computing is about insight, not numbers

D. Koop, CSCI 627/490, Fall 2023

– R. W. Hamming











The purpose of visualization is about insight, not pictures

D. Koop, CSCI 627/490, Fall 2023

- Shneiderman











Why do we visualize data? (vs. looking at tables?)







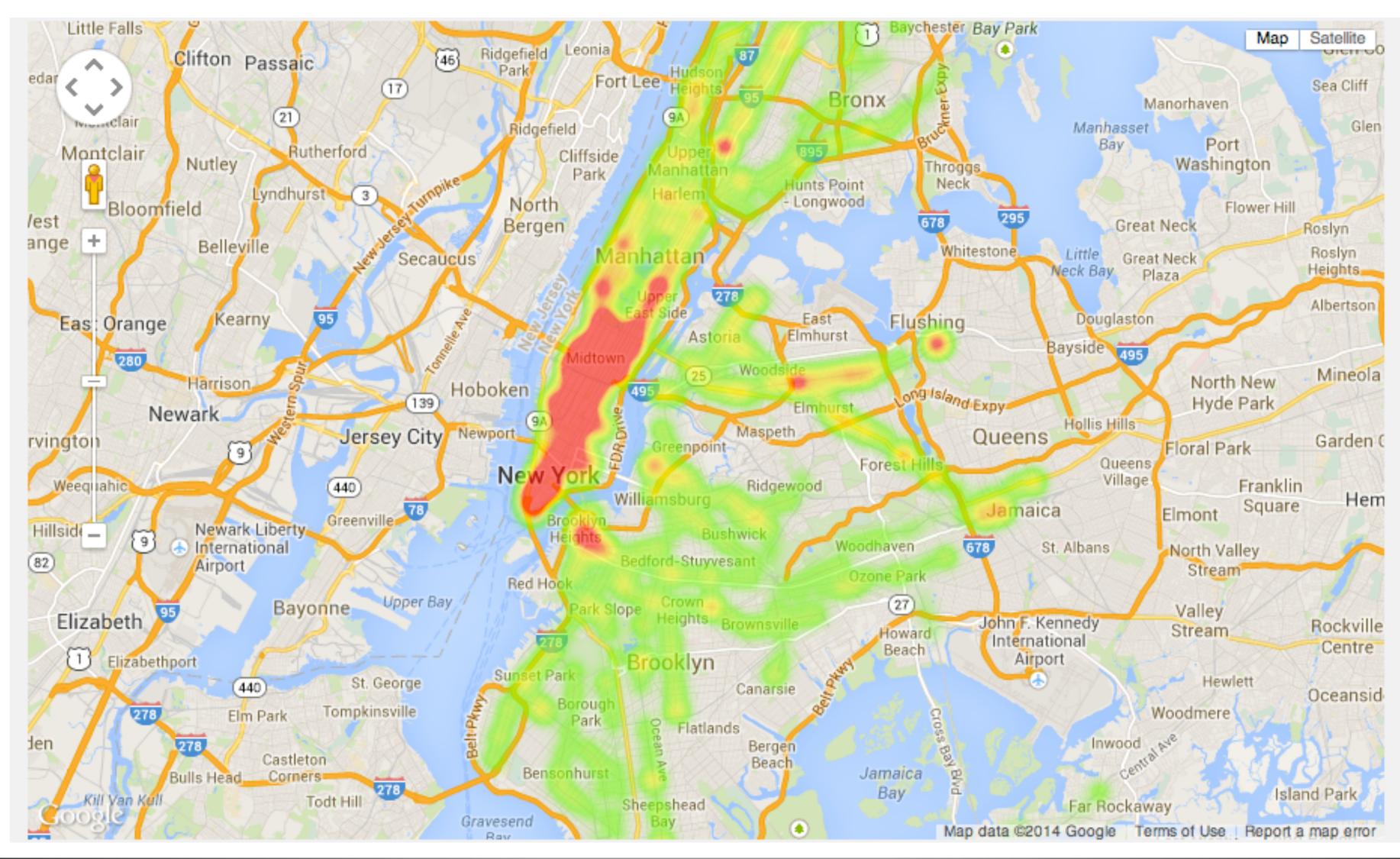
MTA Fare Data Table

	REMOTE	STATION	FF V	SEN/DIS	7-D AFAS UNL	D AFAS/RMF I	JOINT RR TKT	7-D UNL	30-D UNL
1 R	R011	42ND STREET & 8TH AVENUE	00228985	00008471	00000441	00001455	00000134	00033341	0007125
2 R	R170	14TH STREET-UNION SQUARE	00224603	00011051	00000827	00003026	00000660	00089367	0019984
3 R	2046	42ND STREET & GRAND CENTRAL	00207758	00007908	00000323	00001183	00003001	00040759	0009661
4 R	R012	34TH STREET & 8TH AVENUE	00188311	00006490	00000498	00001279	00003622	00035527	0006748
5 R	293	34TH STREET - PENN STATION	00168768	00006155	00000523	00001065	00005031	00030645	0005437
6 R	2033	42ND STREET/TIMES SQUARE	00159382	00005945	00000378	00001205	00000690	00058931	0007864
7 R	2022	34TH STREET & 6TH AVENUE	00156008	00006276	00000487	00001543	00000712	00058910	0011046
8 R	R084	59TH STREET/COLUMBUS CIRCLE	00155262	00009484	00000589	00002071	00000542	00053397	0011396
9 R	R020	47-50 STREETS/ROCKEFELLER	00143500	00006402	00000384	00001159	00000723	00037978	0009074
10 R	R179	86TH STREET-LEXINGTON AVE	00142169	00010367	00000470	00001839	00000271	00050328	0012525
11 R	R023	34TH STREET & 6TH AVENUE	00134052	00005005	00000348	00001112	00000649	00031531	0007504
12 R	R029	PARK PLACE	00121614	00004311	00000287	00000931	00000792	00025404	0006536
13 R	R047	42ND STREET & GRAND CENTRAL	00100742	00004273	00000185	00000704	00001241	00022808	0006821
14 R	R031	34TH STREET & 7TH AVENUE	00095076	00003990	00000232	00000727	00001459	00024284	0003867
15 R	R017	LEXINGTON AVENUE	00094655	00004688	00000190	00000833	00000754	00020018	0005506
16 R	R175	8TH AVENUE-14TH STREET	00094313	00003907	00000286	00001144	00000256	00038272	0007466
17 R	R057	BARCLAYS CENTER	00093804	00004204	00000454	00001386	00001491	00039113	0006811
18 R	138	WEST 4TH ST-WASHINGTON SO	00093562	00004677	00000251	00000965	00000127	00031628	0007445





MTA Fare Data Visualization



D. Koop, CSCI 627/490, Fall 2023



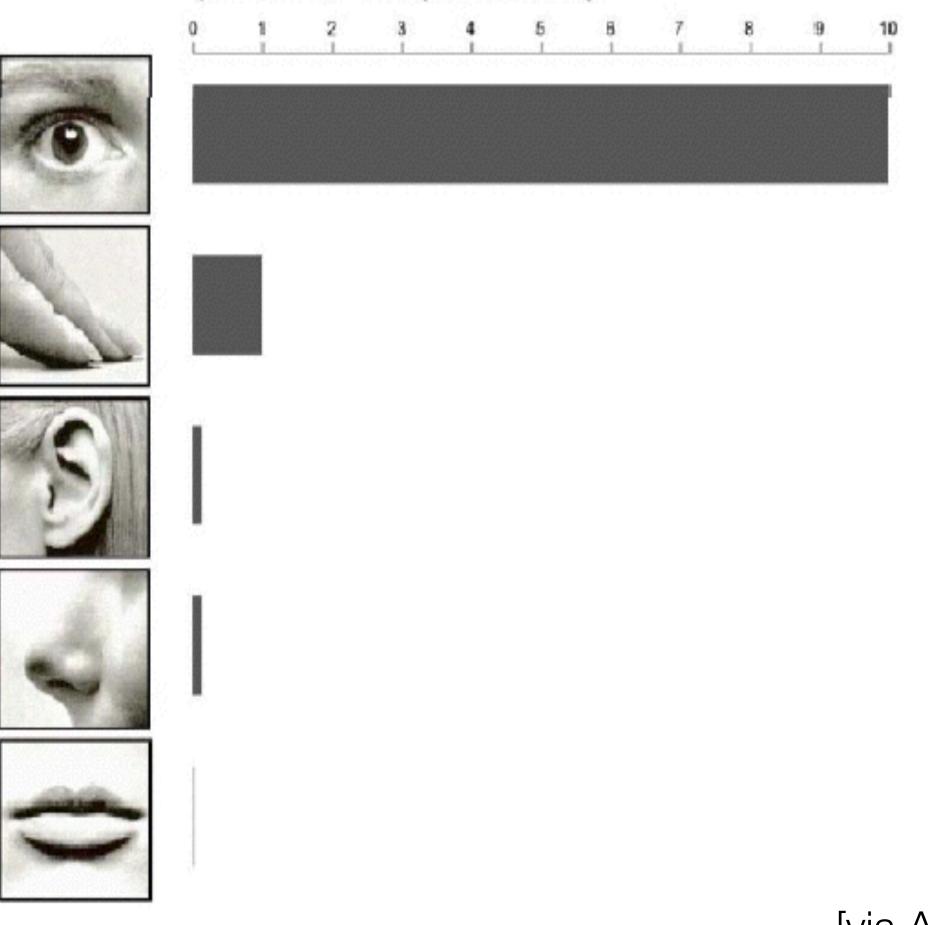


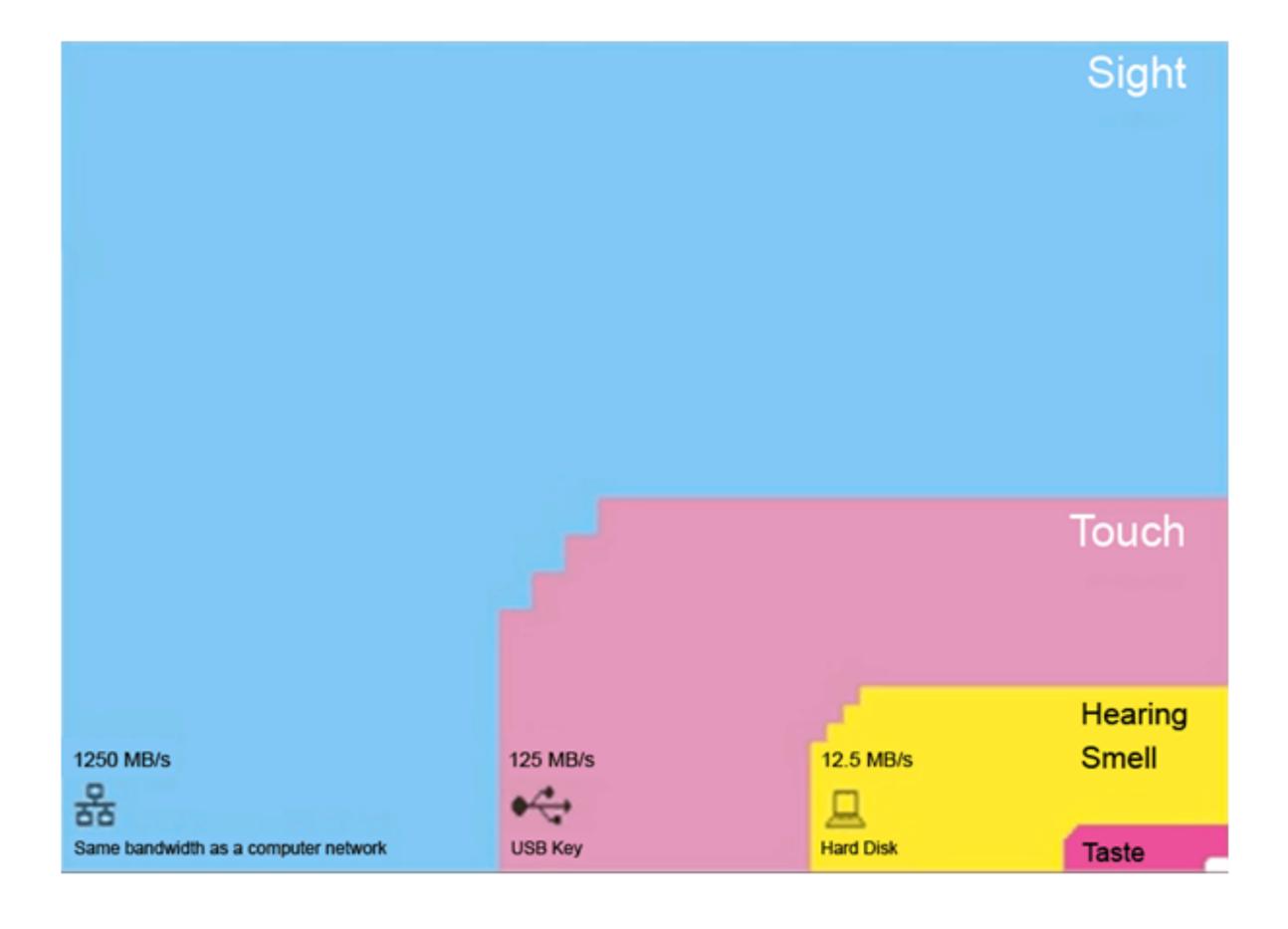
8

Why do we visualize data?

Total Bandwidth

(millions of bits per second)













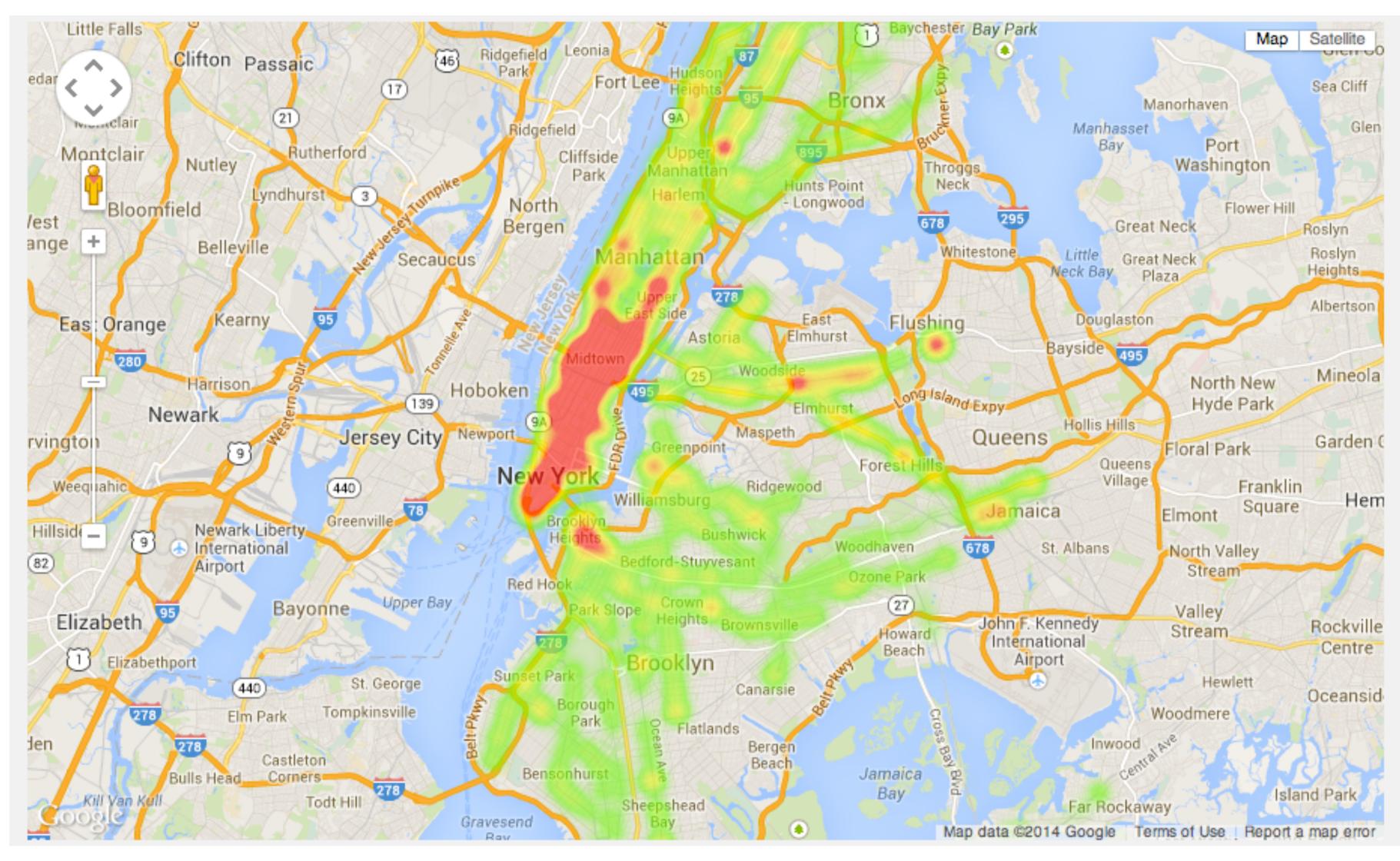


What are the purposes for visualization?





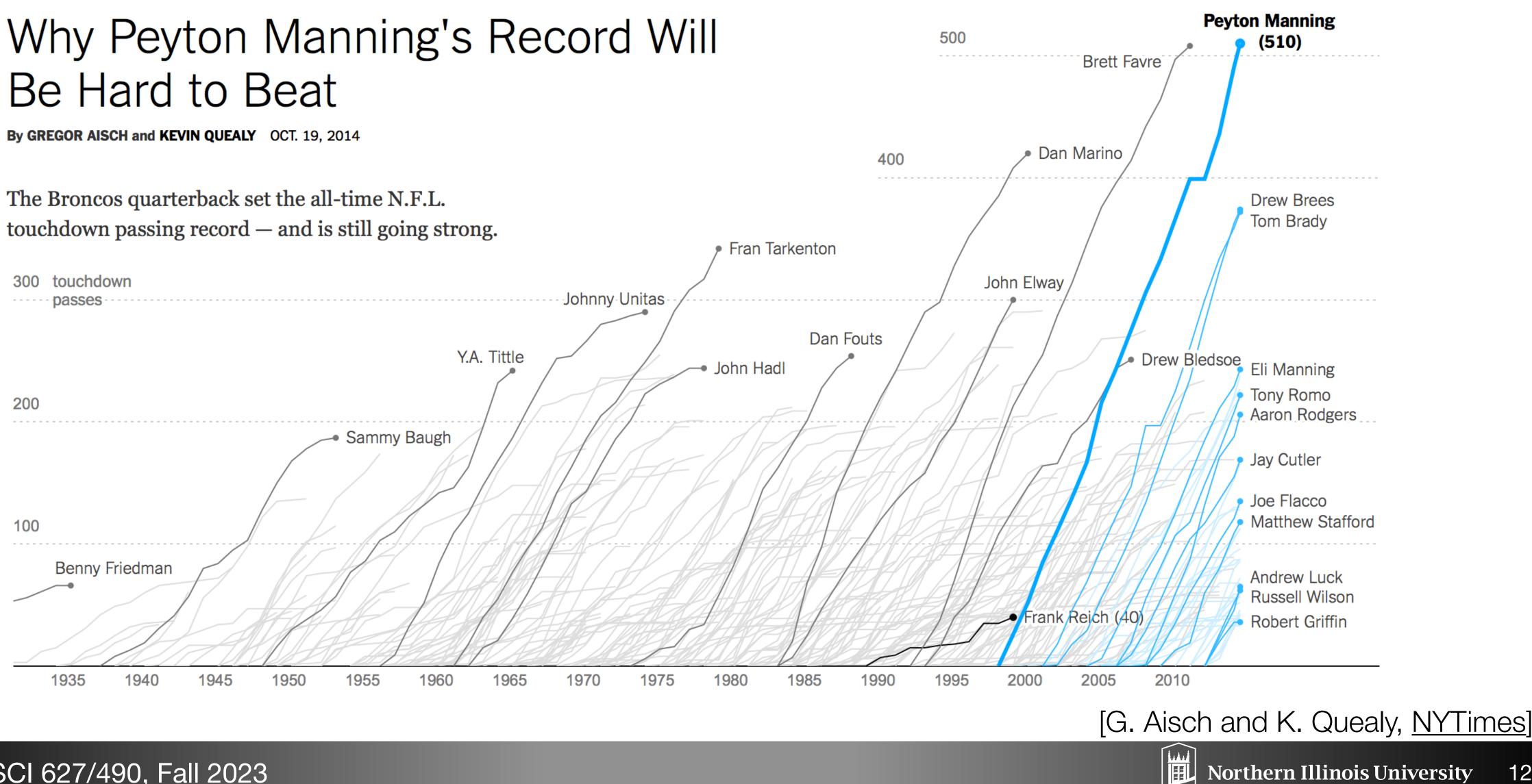
Analysis: Subway Ridership Density







Communication: Peyton's Records



D. Koop, CSCI 627/490, Fall 2023





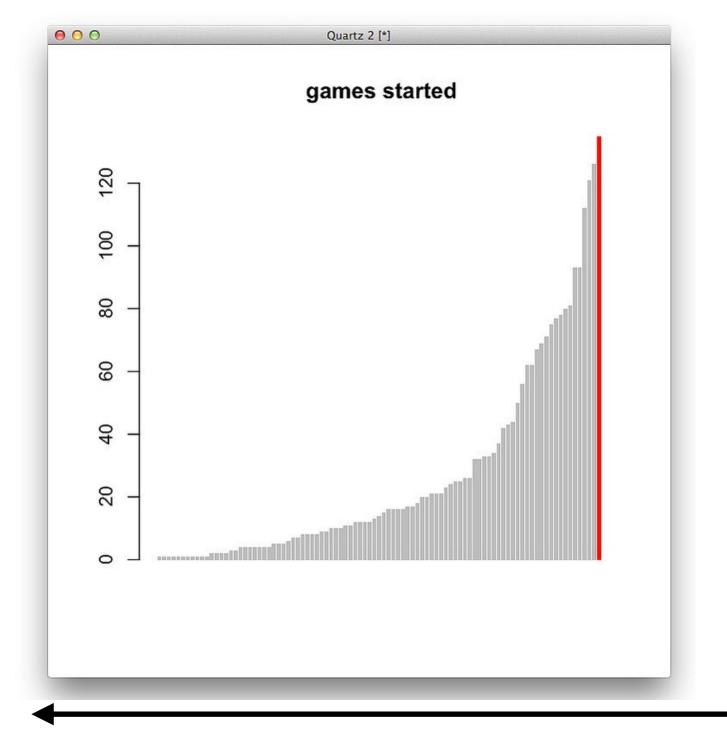
NIU



Exploration <-> Communication Spectrum

Consecutive Starts by a Quarterback for a Single Team

Each streak ------ shows

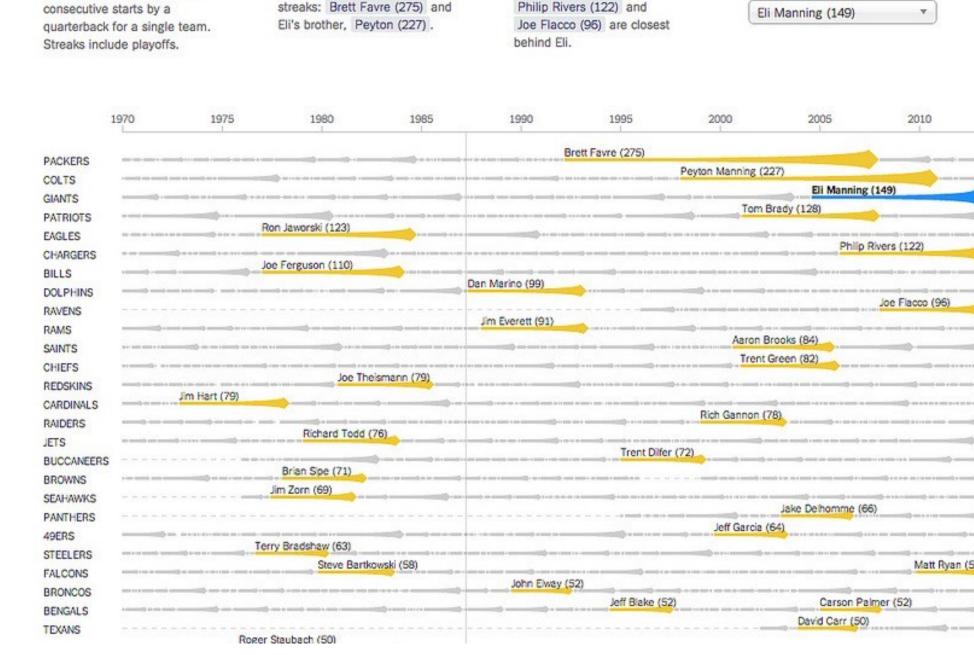


Exploration



Questions

D. Koop, CSCI 627/490, Fall 2023



Only two players have longer

streaks: Brett Favre (275) and

Among active players,

Philip Rivers (122) and

Confirmation

Communication

Answers/Persuasion





Find a quarterback



13



What types of data can we visualize?





Types of Data

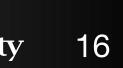
- Tables
- Networks (Graphs)
- Spatial Data
 - Geography
 - Physical (e.g. Scientific, Medical)
- Text
- Sets



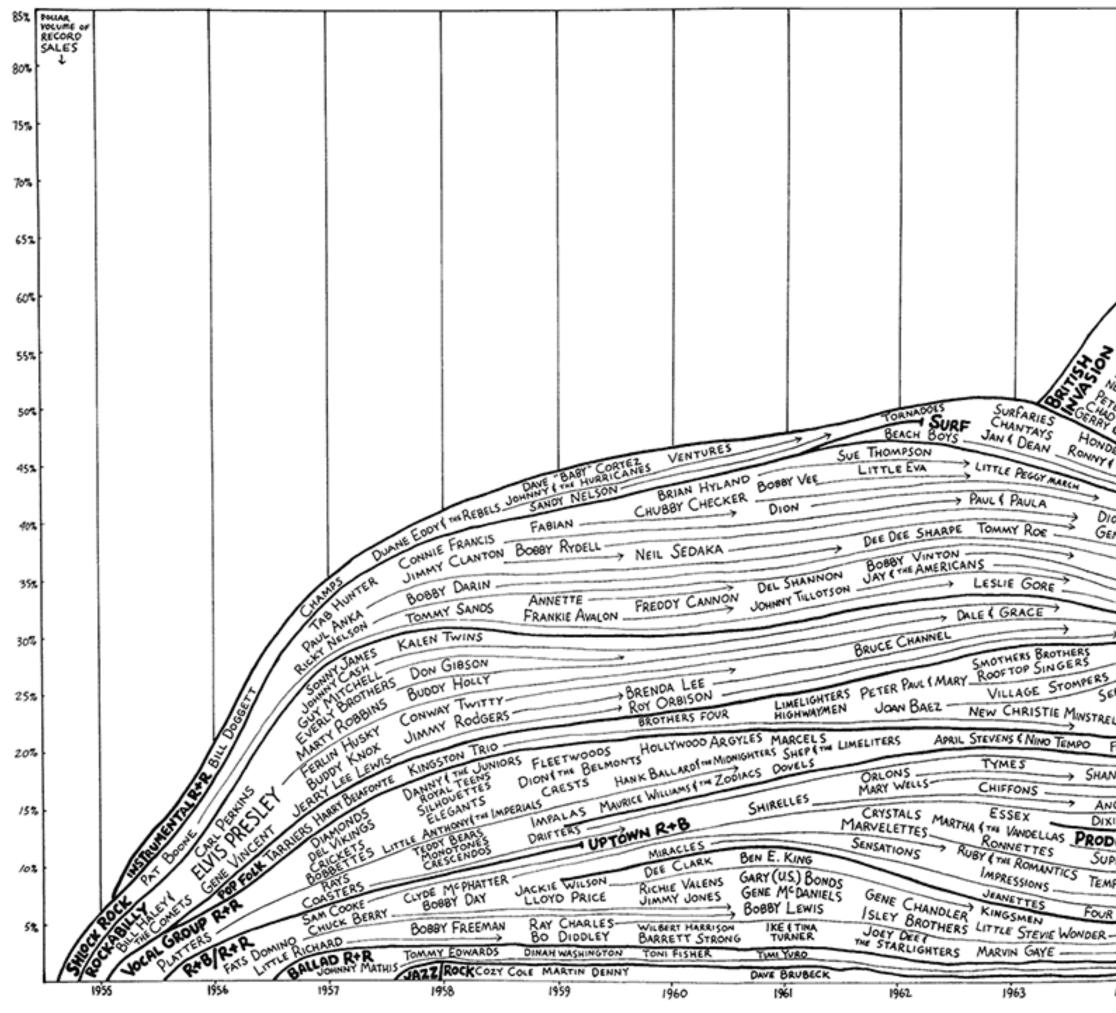


Where have you seen visualizations?

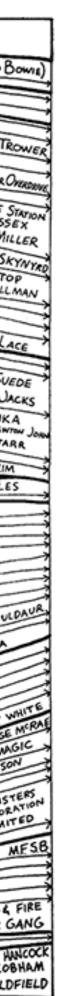




Books/Posters



			GRAND FUNK RAILROAD		ALICE COOPER	EDGAR WINTER	(DAVID BOWIE)
				GIT BLACK SABBAT	(State of the sta	DEEP PURPLE -	
			HEAV	YMEINE C	ARGENT	PINK FLOYD -	
			EUNK RAULROAD	EMERSON, LAKE, & PA	Ver >	Focus A	ROBIN TROWER
			GRAND FUNK		162	URIAH HEEP BACHMAI	witten a
	RE & THE RAIDERS VAN MORRISON KINGSMEN PROCUL HARUM	CREAM BLIND F		GEORGE HARRISON		Bee	ASTOKNER OVERSKI
	RE THE RAIDERS VAN MORRISON	IRON BUTTERFLY LED Z	EPPELIN MOUNTAIN	, , , , , , , , , , , , , , , , , , ,	MALO J		SVILLE CY.
Bur Reve	KINGSMEN PROCUL HARUM	STEPPENWOLF-	WHO SANTANA		LEON RUSSELL	OF L.C.	SERVICE AND A
C TONG TO TO A	SSOCIATION TURTLES	STEPPENWOLF GUESS	ARWATER REVIDE COCH	JANIS JOPLIN A	LLMAN BROTHERS		NYRD SKYNYR
ROLUNKS ZOMBIES			CUOLE .	JANIS JOPLIN PLASTIC OND BAND MODDY BLUES	James Cure S	DOBIE BROTHERS	REGG ALLMAN
DAVIALS VIDORIROS	STANDELLS VANILLA BIO	BROTHER (THE HOLDING COMPAN	Y NICHT JOHN LENNON	M0001 010-0-0	Grand CANS L	OGGINS & MESSINA	
	TROGGS ACIDROCKJIMI HENDI	THREE D		GIL GIL	BERT O'SULLIVAN B	SETTE MIDLER	APER LACE
S BEATLES MOODY BLUES	HOLLIES POPPort	RPLANE	TAIRCASE NORMAN GREEN TAIRCASE NORMAN GREEN MARY TONY ORLANDON MARY ANNE MURRAY ANNE MURRAY CARPENTERS	DAWN		The second se	BLUE SUEDE
DUNINE WINN HERMAN'S HERMITS	VOGUES POPROCKBOX TOPS DUSTY SPRINGFIELD BEE GEES CYRCLE HAPPENING		ALRCASE ORLANDRA	OSMONOS	CHUCK BERRY MAN	DEEN Microw	LERRY JACKS
TER GORATS TREODY & THE DREAMERS	LEFT BANK BUCKINGS	SpiRAL Out	MARY ANNE ANTERS	RICK NELSON	CHUCK BERRY MAN		UL ANKA
MY 6 JEREMEY HERMAN'S HERMITS !	NEIL DIAMOND BUCKINGHAMS	CLASSICS IN PETER.TA	WINNON FIRST EDITION	HELEN REDUT-	-> WINGS		NGO STARR
NDELLS PACEMAKERS PETULA CLARK	NANCY SINATRA BOARD	VISS ROOTS	ESTO THUMPE	NY		CT08165	NOY KIM
CATIONAS CLARK-	SCOTT M	LEMON GAP LEMON PIPERS JACKIE U LETTERMEN TON JO O.C. SMITH GLEN MENUSPORT	NESELL ENGLEBERT HUMT	CLIMAN	RASPBERRIES	TY GRITTY DIAT BAND	EAGLES
DIONNIC	SCOTT MCK	O.C. SMITH GLEN	Chi Port Alle	PARI	RASPOLINE NIT		
DIONINE WARNICK HERB ALPERT & TJB GENE PITNEY	BOB LIND FIFTH DI	MARY HOPKIN JOHN	CASH POPPY FAIR R VANITY FAIR VANITY FAIR	PARTRID	NEIL YOUNG		
RUPPI FOUN SONNY TOWN	MULAMES (THE SHONDELLULU	BOBBY GOLDSBORD ARC	TES EDISONES MALES	LIGHTFOOT	MAC DAVIS	\rightarrow	
GARY LEVING & THE A	BOB LIND FIFTH DIMENSION MONKEES STRAWBERRY ALARM CL	DOBBY GOIDSBOAN	CASH POPPY CAN R VANITY FAIR NES VANITY ENSON NES CONSON LIGHTHOUSE NES CONSON LIGHTHOUSE NES CONSON LIGHTHOUSE NES CONSON DEAD NES CON	JOAN BAEL JOAN LIGHTFOOT GORDON LIGHTFOOT CAT STEVENS JU CAROLE KING PA	UL SIMON		
CHART LEWIS & THE PE	ATBOYS ALARM CL	1910 FRUITGUM CO. BOND	CH GRANNIE	Cr.	M CROCE	GARFUNKLE MU SEALS & CROFTS	ARIA MULDAUN
WE FIVE		B DYLAN EXPRESS BROOK	No AV	LOUN DENVEN U	DON MELEAN ART	SEALS & CROFTS	Abba
U ROUNN NEW	VAUDVILLE PAULO BO	B DYLAN OY ROL BOD	GBLOODS BREAM	CARLINGEFERSON	DON MECHANIS	THERS	
RS BORNDIPITY FOLK BOB DYLANDIPITY FOLK BOB DYLANDIPITY FOLK BORN SPOONFUL LOVIN' SPOONFUL STATUTION FAITHFUL BARRY MK GUIRE STATUTION FOLK	BARRY SADLER JANIS JAN	GFIELD JUDY COLLINS	JAMES TAYLOR	JONI MITCHELL JO	AMERICA ISL	EY BROTHERS	
RELS BARRY MC GUIRE S	DONOVAN BUFFALO SPRIN				HI LITES		BARRY WHITE BEORGE MERA
FOUR SEASONS		CRUIL LEBRY	BUTLER BROOK BENTON	AL GREEN ST	TYLISTICS	Quin	OF WAGIC
HANGRI-LAS	Trees	CROSON, CROSON	STAIRSTEPS	BILL WITHERS			AL WIE
AUGULE ROTHERS	MARVIN GAVE & TAMMY TERRET	POP SOUL FRIEN	NCTION EDWIN STARR	ROBERTA FLACK	INGRAM	NE KENDRICKS	TER SISTERS
ANGELS DIXIE CUPS RIGHTEOUS BROTHERS		FOUN	DATIONS FREDA PAYNE	JEAN KNIGHT LO	S BETTY WRIGHT	EDOIL POINT	CORPORATION
DUCED SOUL BARBAN			DIANA ROSS	GLADYS KNIGHTS THE PIP	S	5 101	MESE
	AGE TOO TRAVELU	ARTHUR BROWN UR		HONEY CONE A HONEY CONE DRIVELIUS BROSS SISTER ROSE MAL	MICHAEL OM	TOWER OF POWER	
EMPTATIONS	ARLA THOMAS ARTHUR CONNELLY JOE TEX	ARTHUR BROWN UR.		> DRAMATICS MAI	DOSSTON		WWW FIRE
VK TOPS D	LOCY SLEDGE SAME UNIC	ARTHUR BROWN UR. ISAAC OTIS REDDING ISAAC	AYES STEVIE WONDER -	BI	LLY PREST	IS MAYFIELD BART	A THE GANG
Maunic Resources	CAPITOLS Soul SURVIVORS	CHAMBERS I	THE ALLSTARS			A WANN	HERBIE HANCOCH
JAMES BROWN	ATCH RYDERI ESQUIRES	ARCHIE BELLE WATTS IOS			STAPLE SINGERS CA	CANTANA.	BILLY COBHAM
1964 RAMSEY LEWIS	1966 1967	1968 1969	1970	1971	1972	1973	1974
							-
	[Rock 'N'	Roll is Ha	re to Pav	R Gar	ofalo 1	977 h	via Ti
				🔟 North	ern Illino	bis Univ	ersity
				NIU			

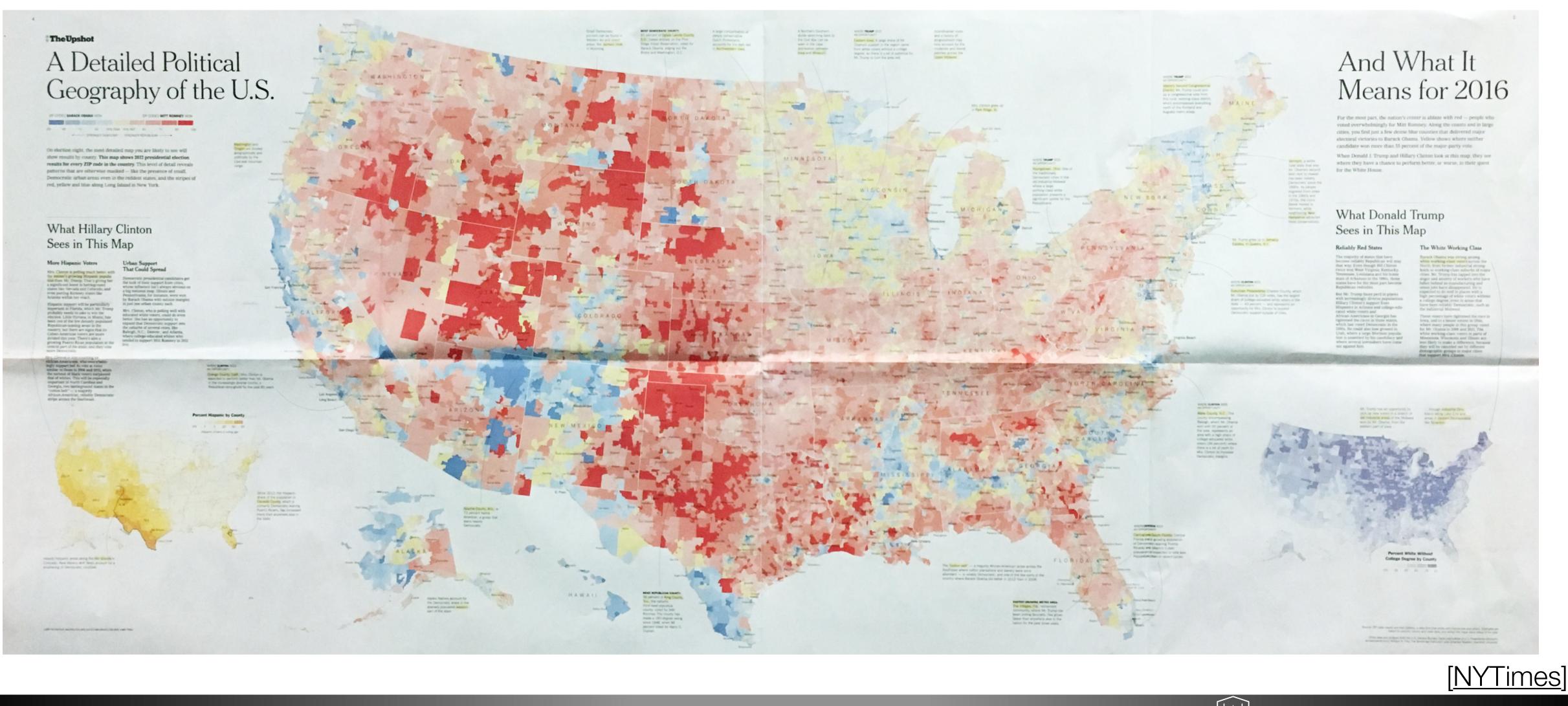








Newspapers

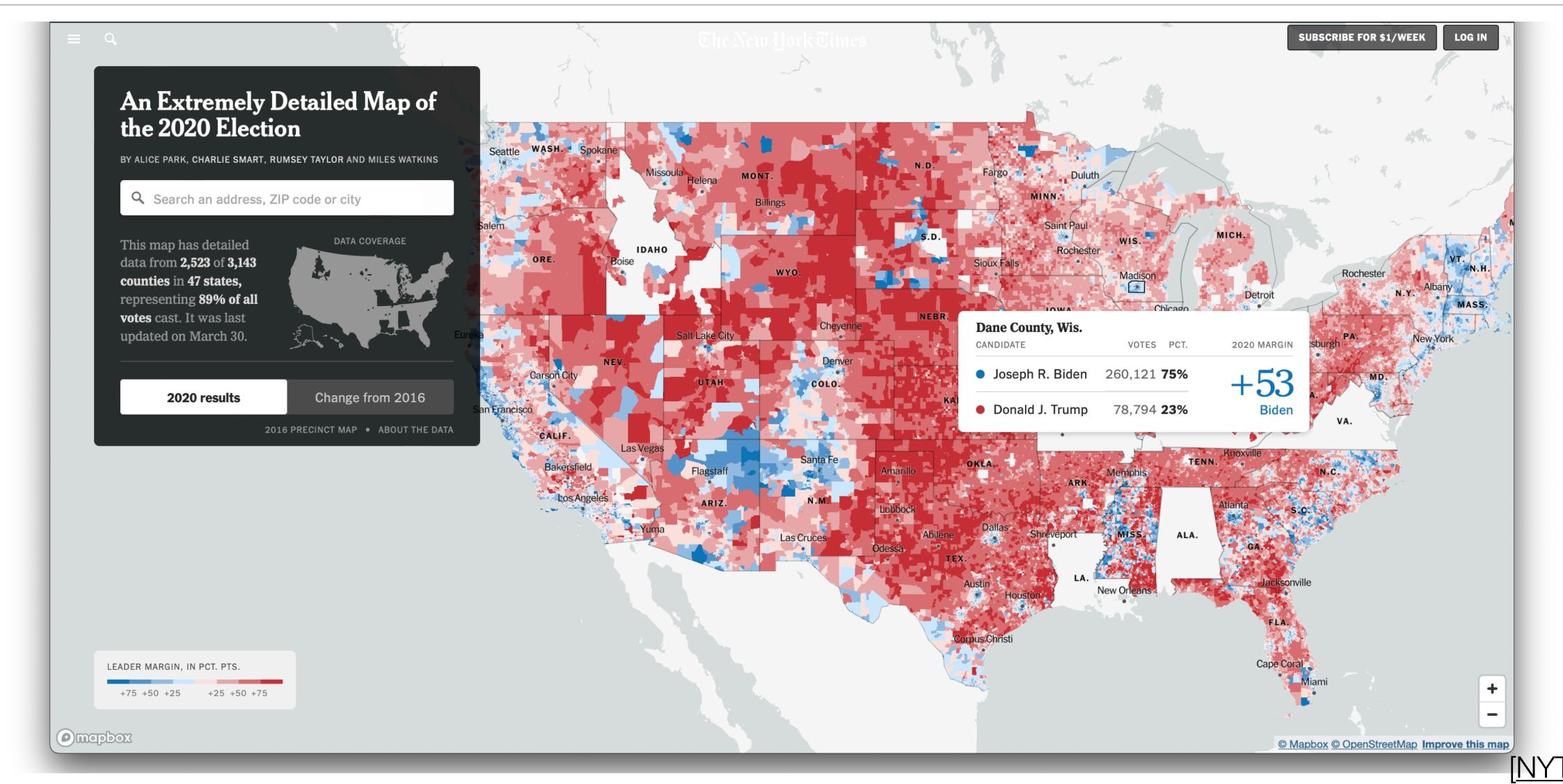








Web



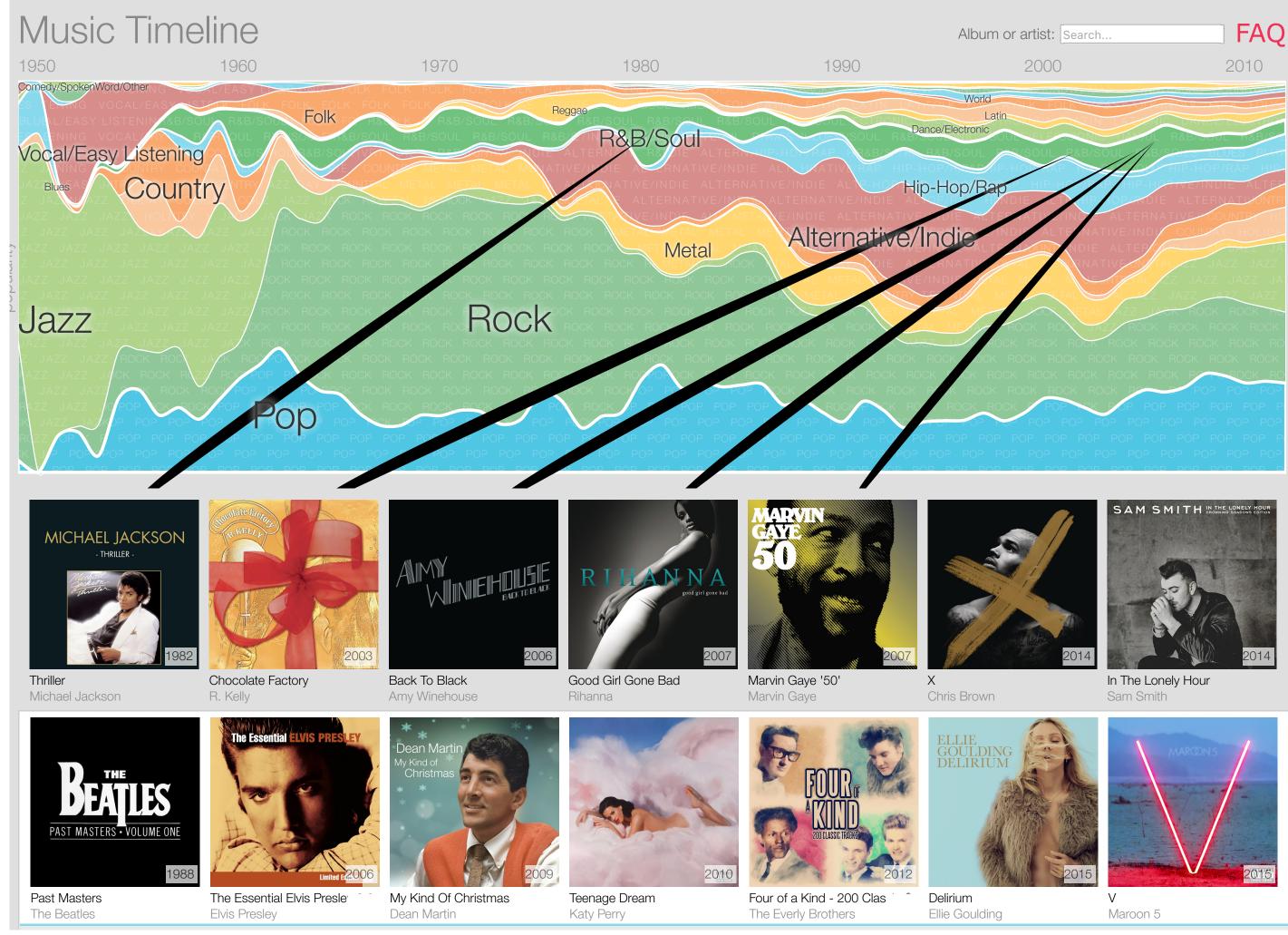








Web



D. Koop, CSCI 627/490, Fall 2023

[Music Timeline, Google Research (No Longer Working)]





What is the advantage of the second version?







Interaction







How do we create modern visualizations?







Tools

- ...

- Desktop Applications:
 - Excel (see <u>excelcharts.com</u>)
 - Tableau
- Grammars:
 - Vega-Lite
- Programming Frameworks
 - d3.js

- Observable Plot, plot.ly, deck.gl

D. Koop, CSCI 627/490, Fall 2023

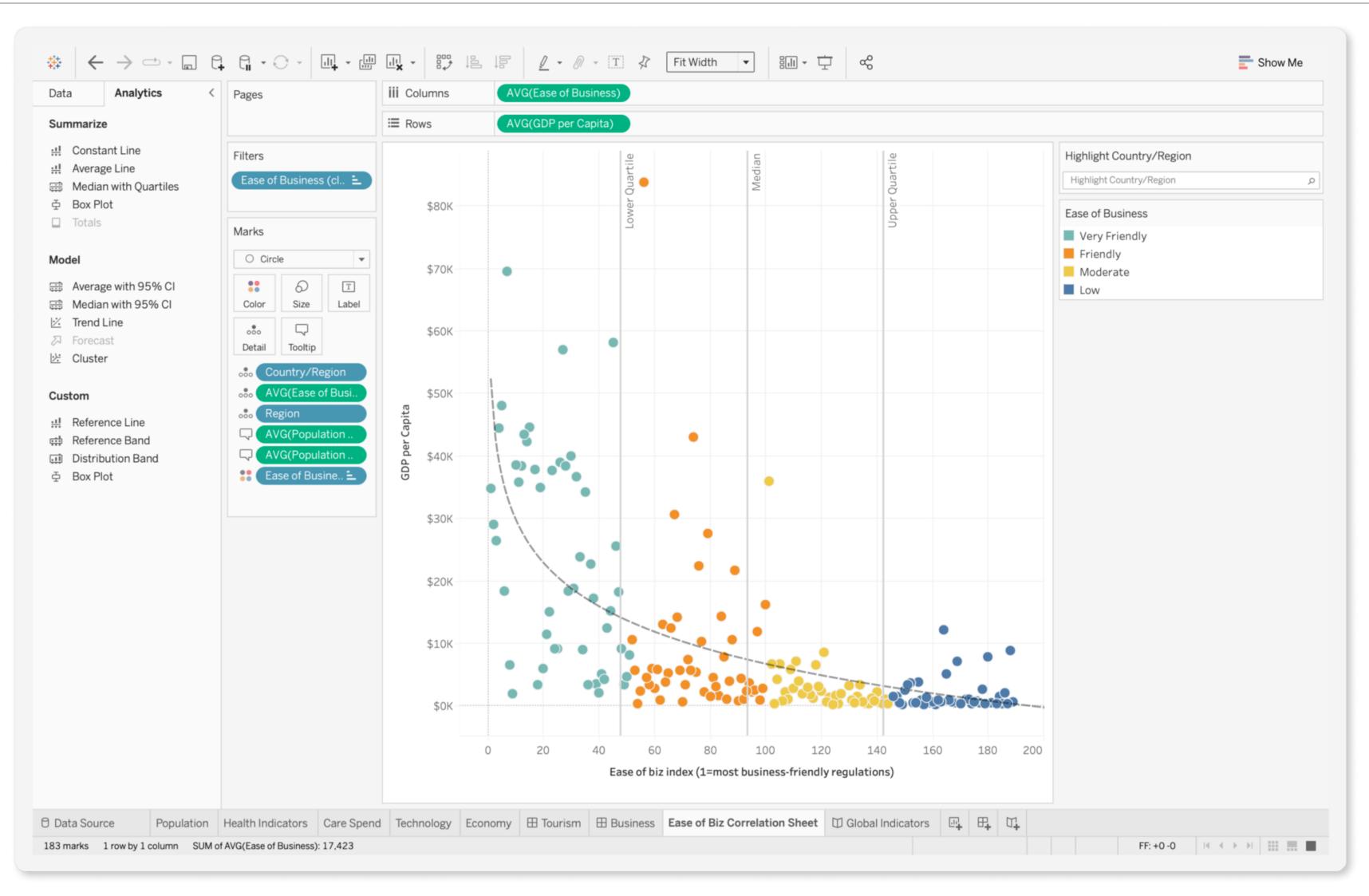
- Tradeoffs
 - Speed
 - Customization
 - Understanding
 - Dissemination





24

Tableau







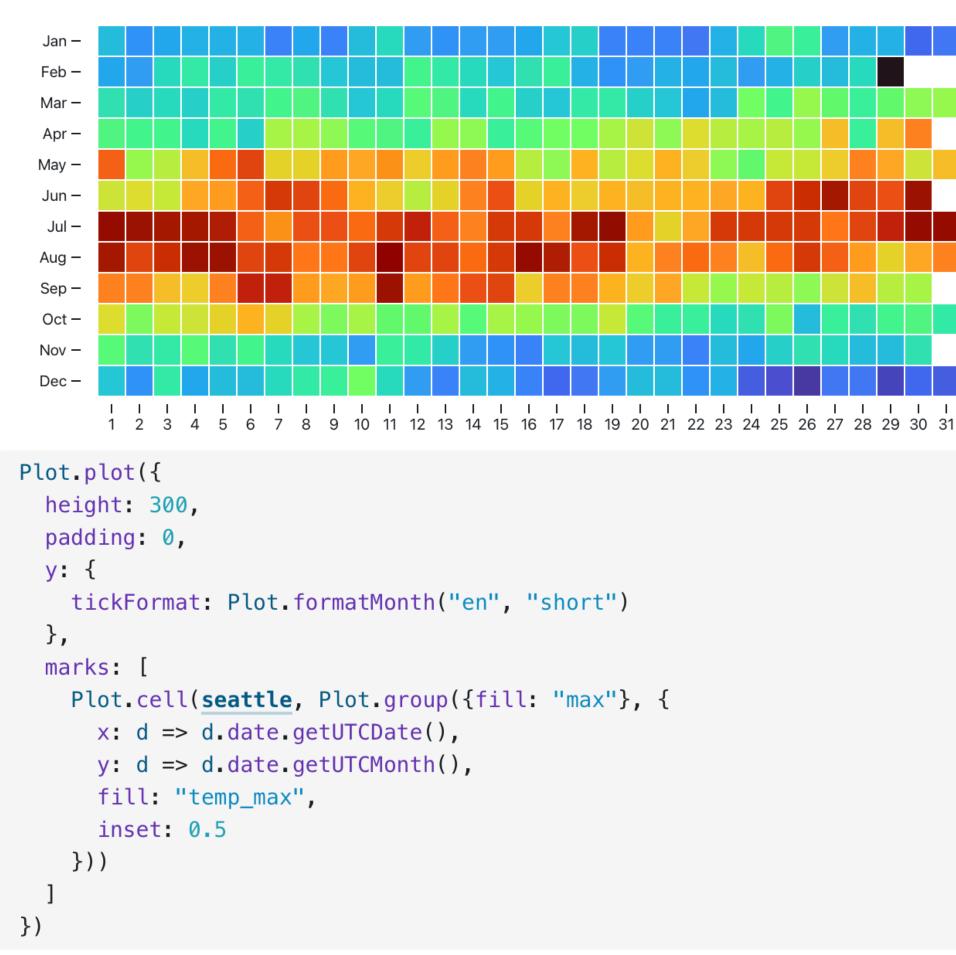






Observable Plot

seattle = > Array(1461) [Object, Object, Objec seattle = FileAttachment("seattle-weather.csv").csv({typed: true})













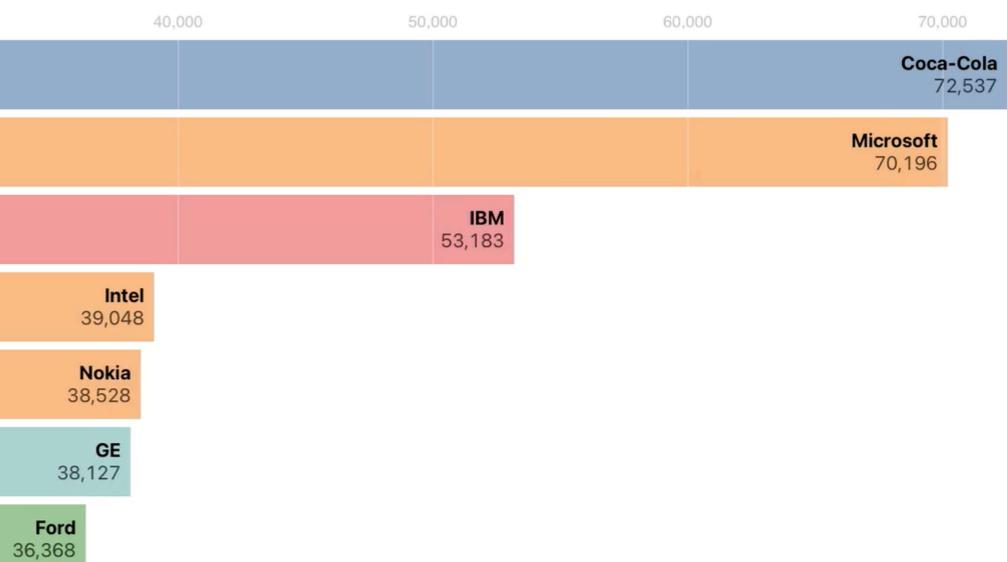


<u>d3.js</u>

Best Global Brands

Value in \$M; color indicates sector. Data: Interbrand

10	,000 20,0	00	30,000
			Disney 33,553
		McDona 27,8	Id's 359
		AT&T 25,548	
	Marlı 22,	boro ,110	
	Mercedes-Ben 21,10	z 4	













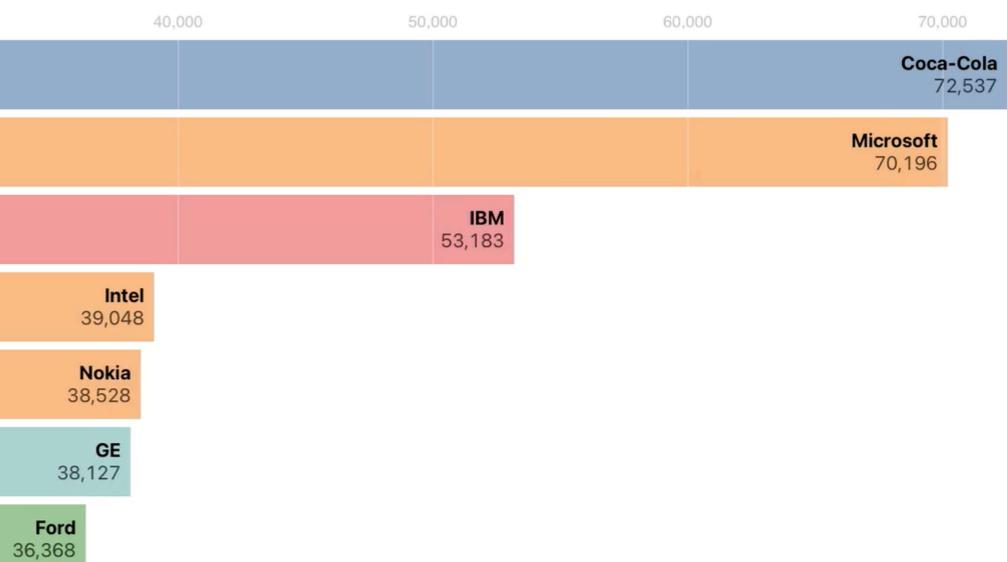


<u>d3.js</u>

Best Global Brands

Value in \$M; color indicates sector. Data: Interbrand

10	,000 20,0	00	30,000
			Disney 33,553
		McDona 27,8	Id's 359
		AT&T 25,548	
	Marlı 22,	boro ,110	
	Mercedes-Ben 21,10	z 4	















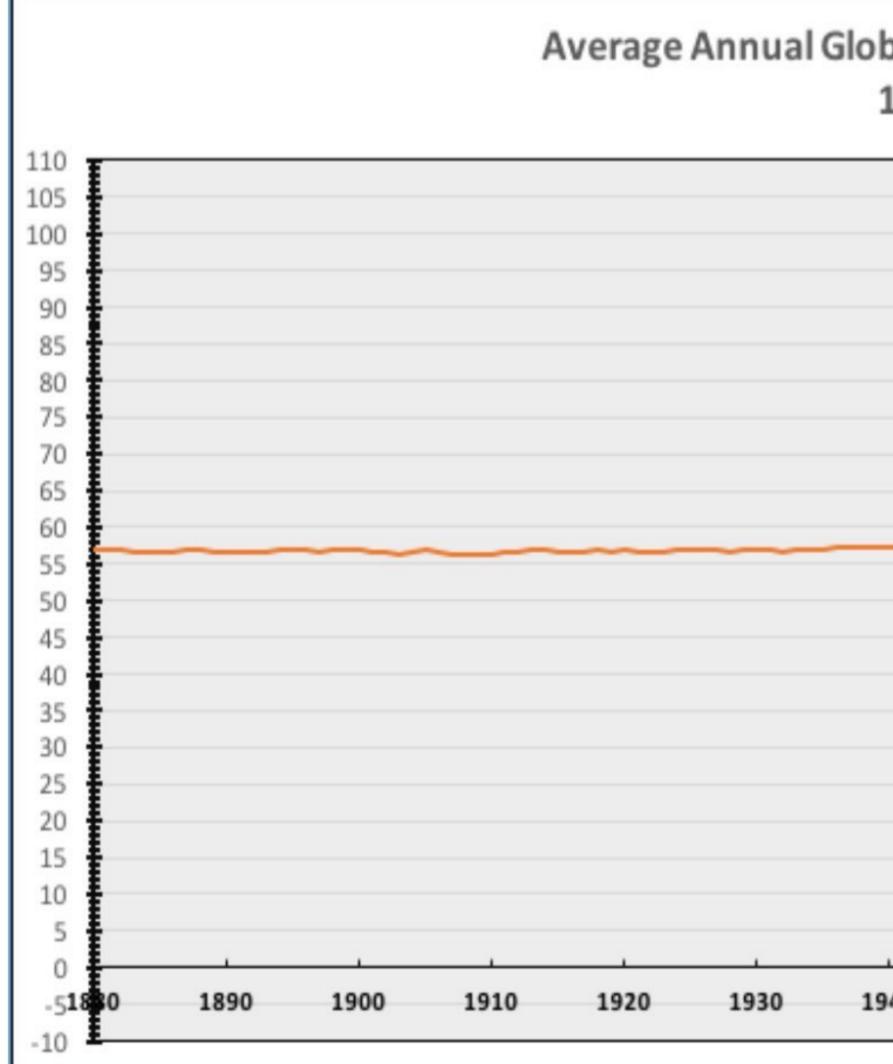
Why do we care about the design of visualizations?







Design: Focus on only the y-axis



D. Koop, CSCI 627/490, Fall 2023

Average Annual Global Temperature in Fahrenheit 1880-2015

								, I	
								1	
								1	
								1	
								1	
								1	
40	1950	1960	1970	1980	1990	2000	2010		
							Гс		
								∋.₽	laywa

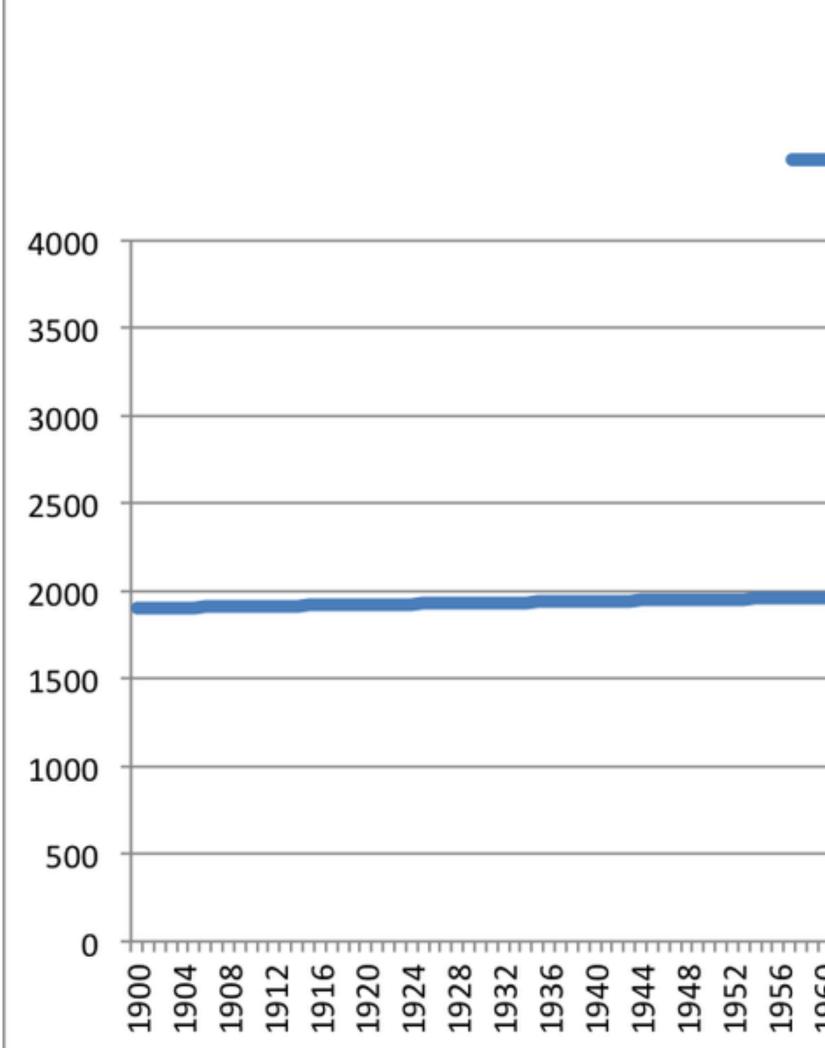








Design: Year on the y-axis



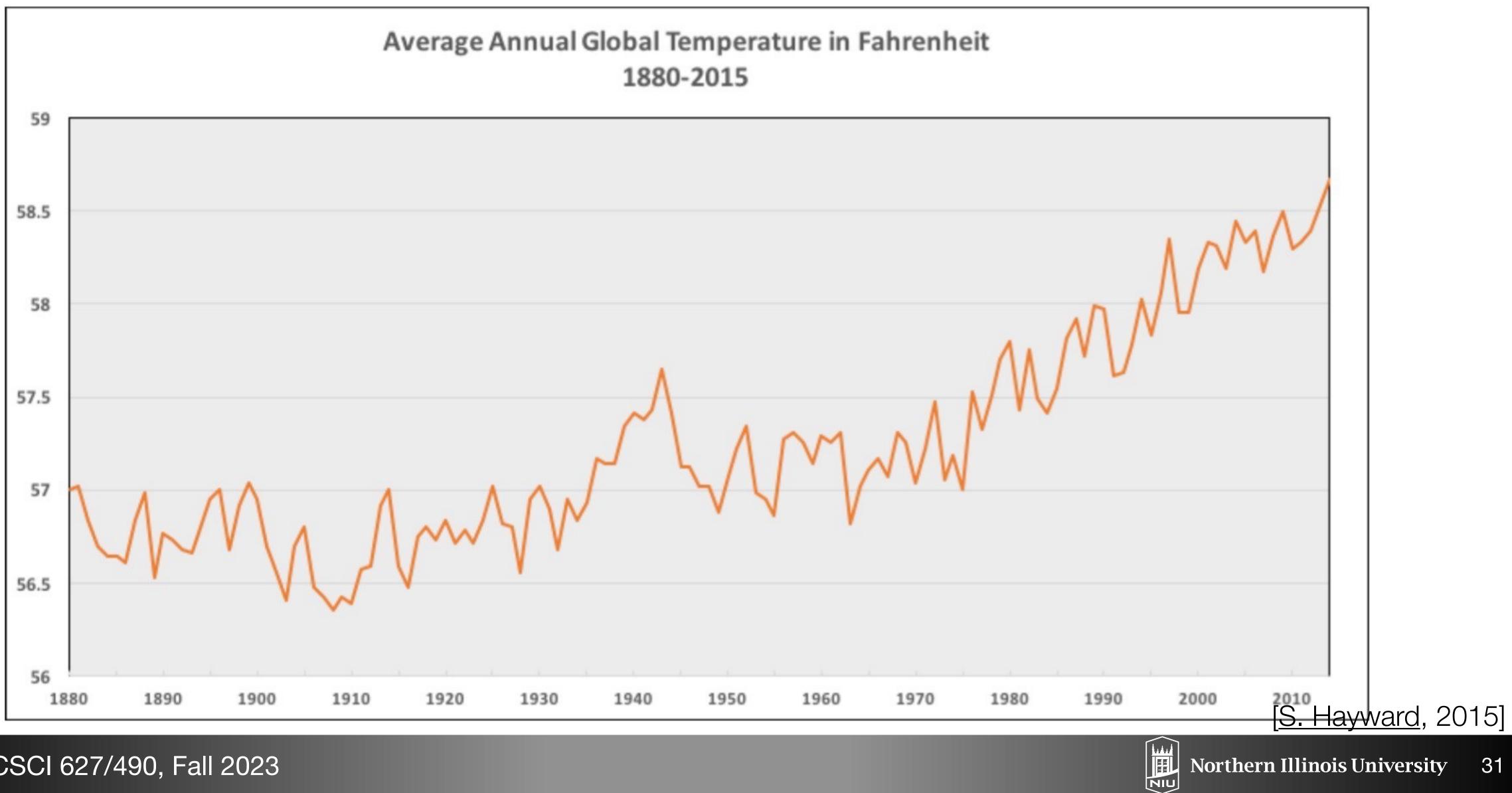
Year	
Year (A.D.)	
\frown \rightarrow \frown	
1960 1964 1968 1976 1976 1976 1976 1976 1976 1976 1976	ଁ କୁଁ <u>[@bizweekgraph</u>
	Iorthern Illinois University







Design: Different y-axis



D. Koop, CSCI 627/490, Fall 2023

Northern Illinois University

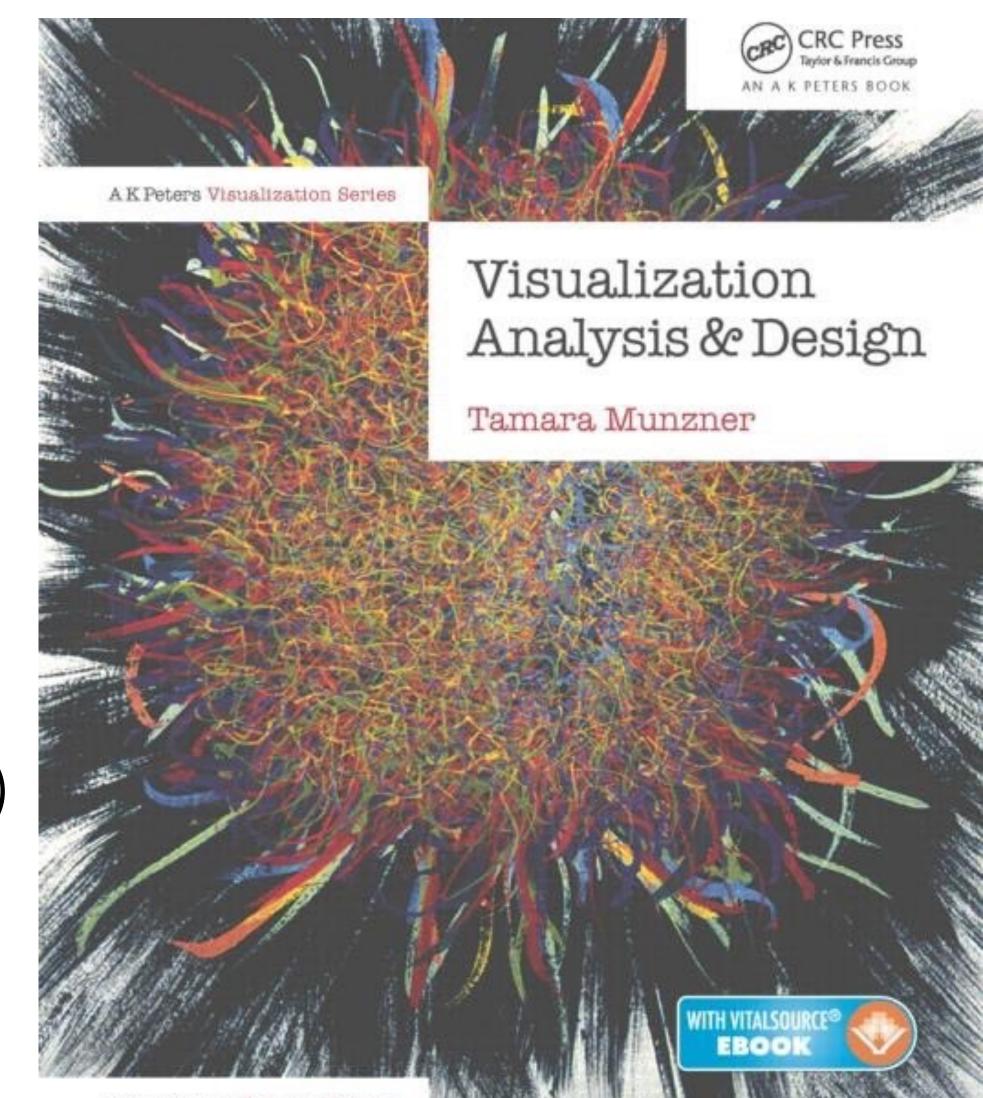






Administrivia

- <u>Course Web Site</u>
- Syllabus
 - Plagiarism
 - Accommodations
- Textbook:
 - Required: <u>Munzner</u> (VAD)
- Assignments
- Exams: Midterm (Oct. 23) and Final (Dec. 11)







Administrivia

- Undergraduate (CSCI 490) and Graduate (CSCI 627)
 - Graduate: Extra reading, exam questions, project emphasis
- Research Topics:
 - Also investigate some topics in depth
 - Research papers as assigned reading (CSCI 627)
- Project: Create an interactive visualization (or vis research)
 - Design
 - Data analysis
 - Insight
 - Presentations: Last week of class









Exams, Quizzes, and Missed Classes

- All quizzes and exams are in-person, closed-book, and closed-notes
- No makeup quizzes or exams will be given*
- You are responsible for all material presented in class and assigned readings. If you miss a lecture, be sure to obtain class notes from a classmate before the next class meeting









Office Hours & Communication

- Office hours will be held in person
- Scheduled office hours are open to all students
 - M: 1:45-3pm, W: 10:45am-12pm, or by appointment
- You do not need an appointment to stop in during scheduled office hours
- If you need an appointment outside of those times, please email me with details about what you wish to discuss
- Many questions can be answered via email. Please consider writing an email before scheduling a meeting.
- Do not send me screenshots of code! (send code or Observable links)







Do not cheat!







Do not plagiarize

- It is **Academic Misconduct**
- Do your own work, do not copy anyone else's work, text, sentences, ...
 - Anyone = another student, an internet source, book, blog, ...
- Never quote text unless there is a specific need.
 - Usually, only famous quotes or very specific definitions
 - "I think there is a world market for maybe five computers." -Thomas Watson (1874-1956), Chairman of IBM, 1943)
- Cite sources that back up your claims or reflect the origin of an idea
 - Vertex cover is an NP-Complete problem [1]. ... [1] Garey, M. R., and Johnson, D. S., "Computers and intractability: a guide to NP-completeness." 1979.







Do not cheat

- Cheating on assignments, projects, and exams is not allowed • You will receive a zero for any assignment/exam/etc. where cheating has
- occurred
- You will **fail** the course if you cheat more than once
- Misconduct is reported through the university's system
- You may discuss problems and approaches with other students
- You may not copy or transcribe code from another source (includes generative AI)







Do ask questions!







Do ask questions

- If you are stuck on a specific issue with an assignment: - Do consult books, online documentation, tutorials
- - Do discuss that specific issue with a classmate
 - Do email me with **specific** questions
- If you are asked about a question:
 - Do not share your code
 - If the questioner is trying to cheat, walk away - If you see an obvious mistake, kindly point it out - Suggest a specific function or library that may be useful





Questions?



