

CS 571 - Data Visualization & Exploration

Overview and Introduction to Visualization

Instructor: Hamza Elhamdadi

UMassAmherst

- What is Visualization?
- Why Not Just Use Stats?
- Uses of Data Visualization
- Why does Visualization Work?
- Course Goals
- Course Logistics

What is Visualization?

**Data Visualization: A visual representation
of data (information) that can help a
person perform tasks more effectively**

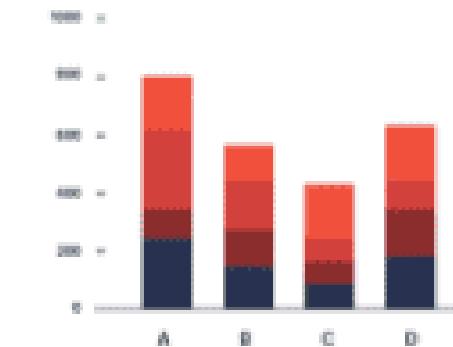
KPIs

180%

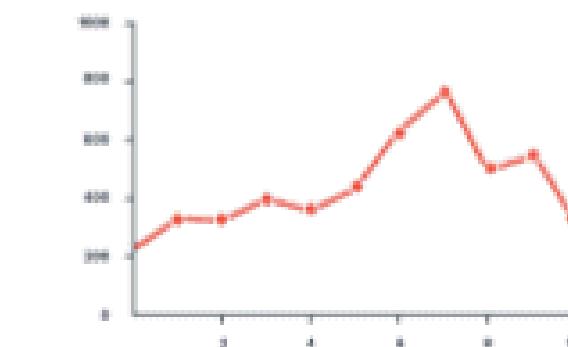
Tables

	A	B	C
X	\$40	240	48
Y	\$50	200	59
Z	\$60	310	79

Bar charts



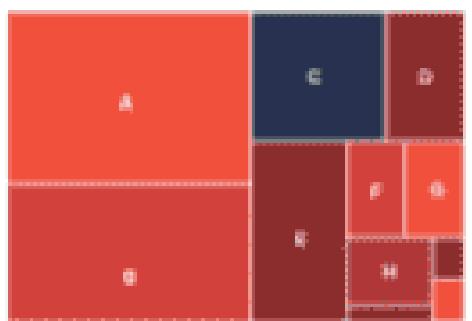
Line charts



Donut charts



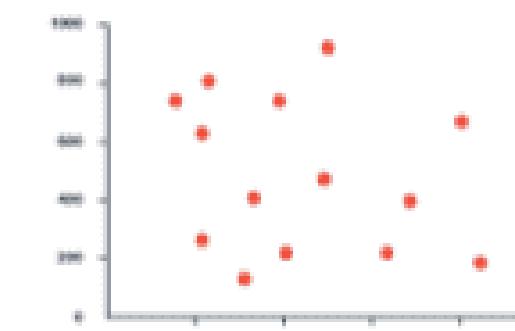
Tree Maps



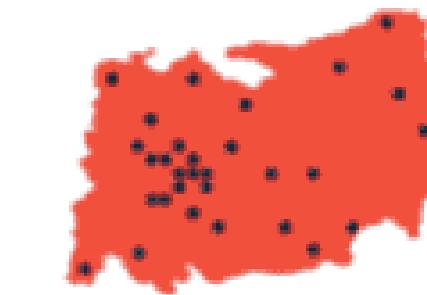
Bullet Charts



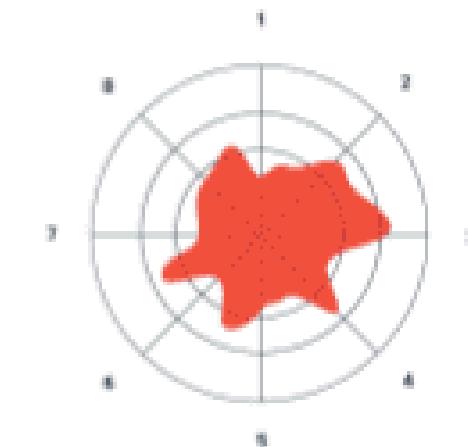
Scatter plots



Geo Maps



Radial charts



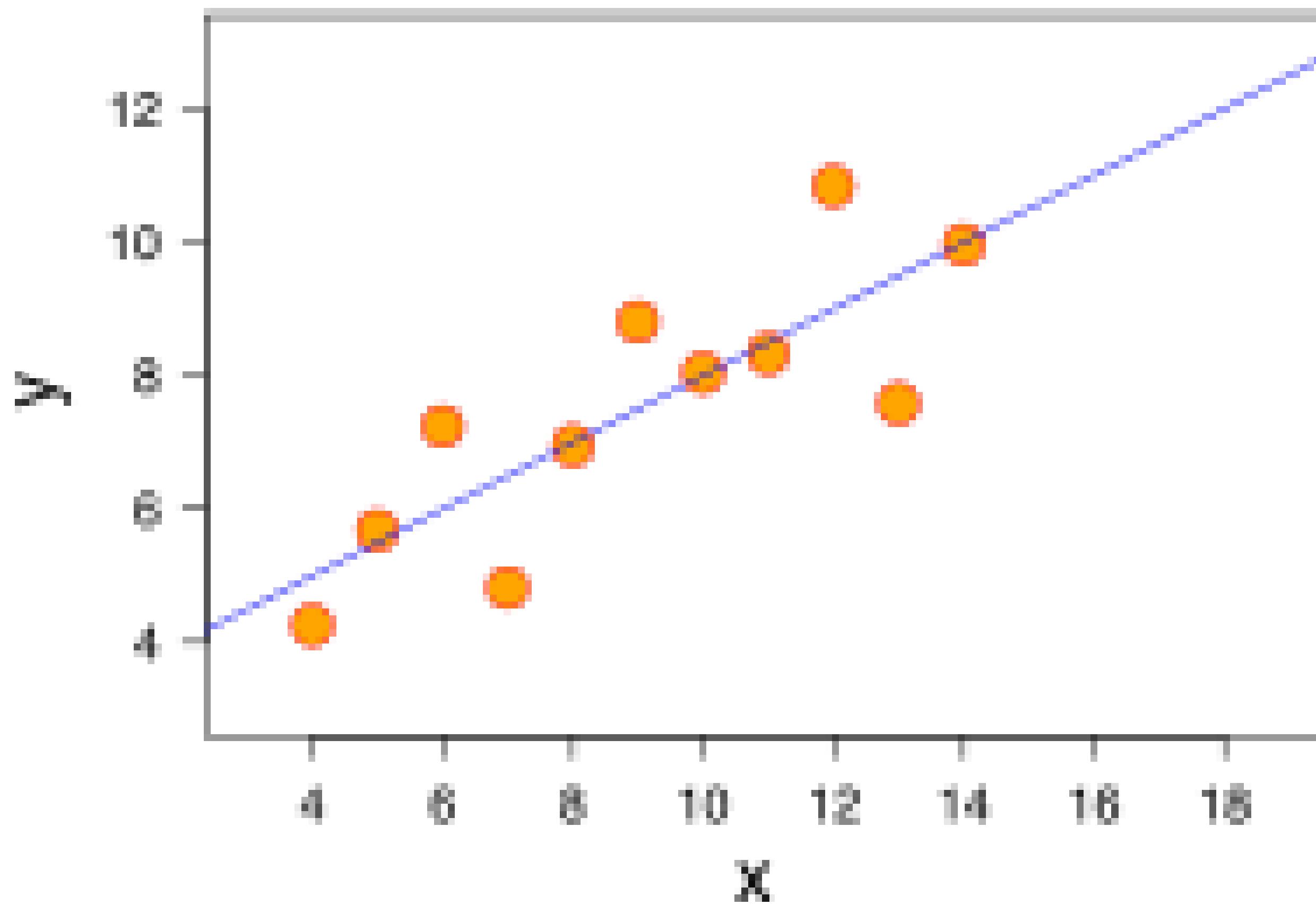
Why Not Just Use Stats?

Dataset I		Dataset II		Dataset III		Dataset IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

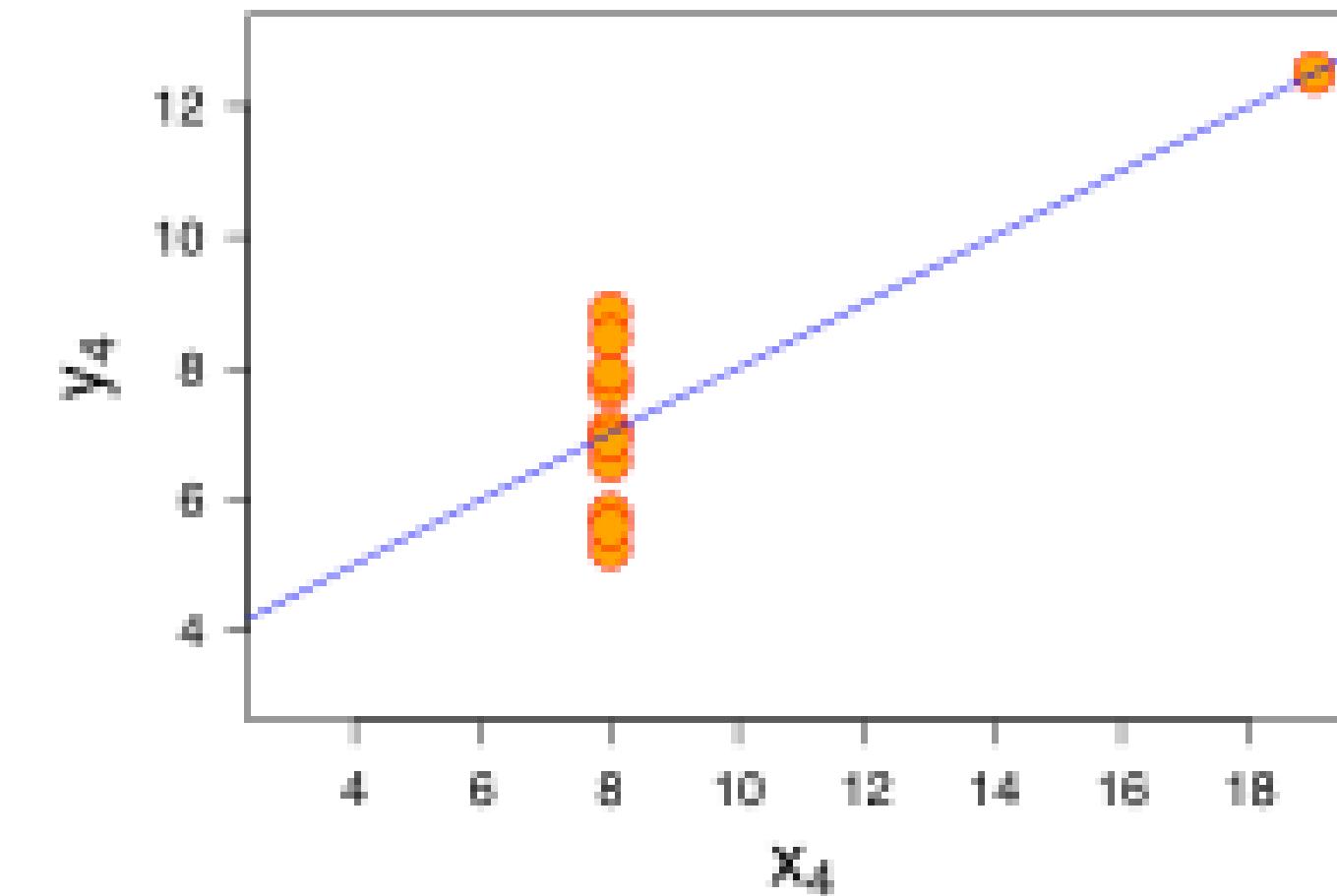
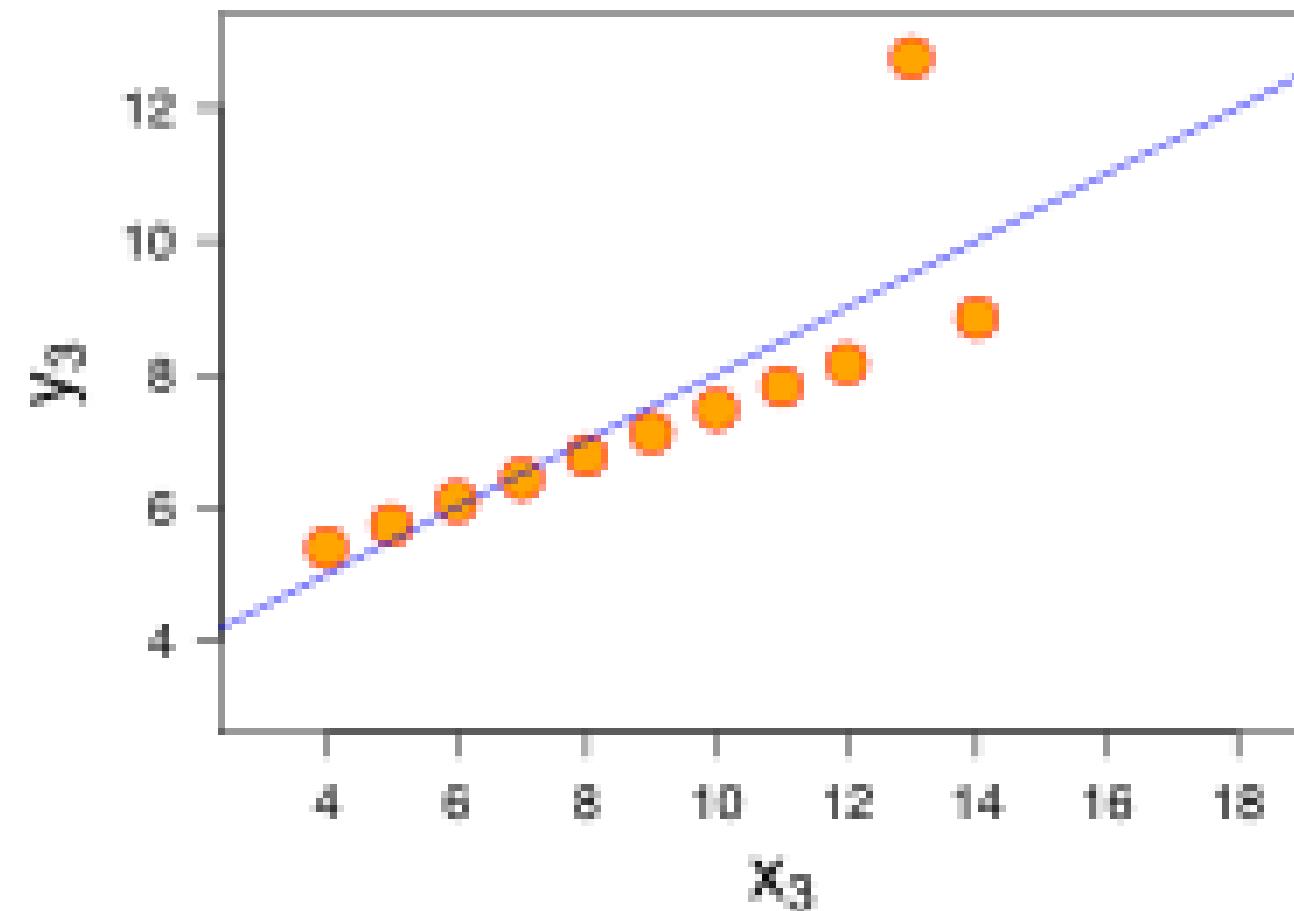
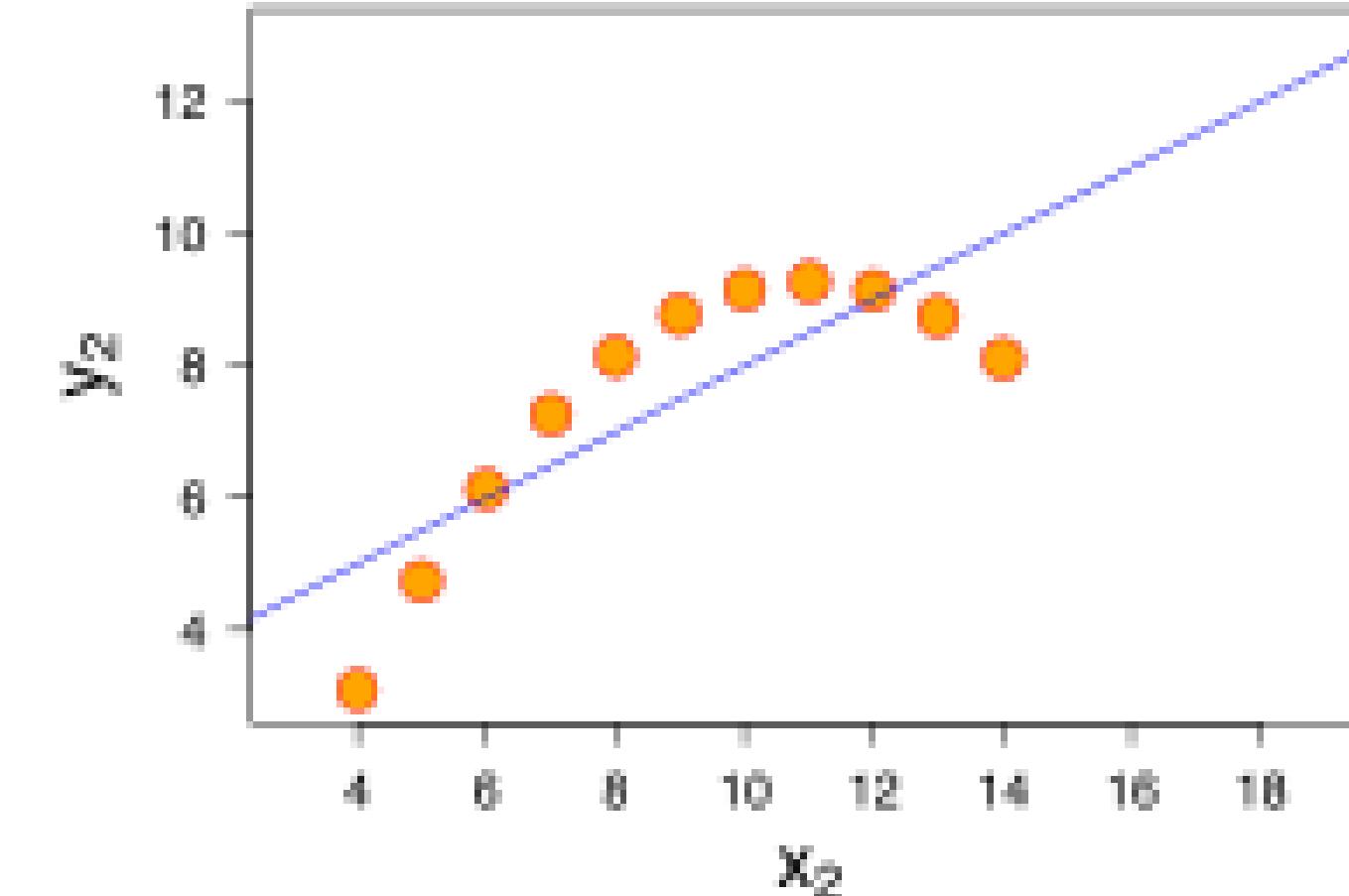
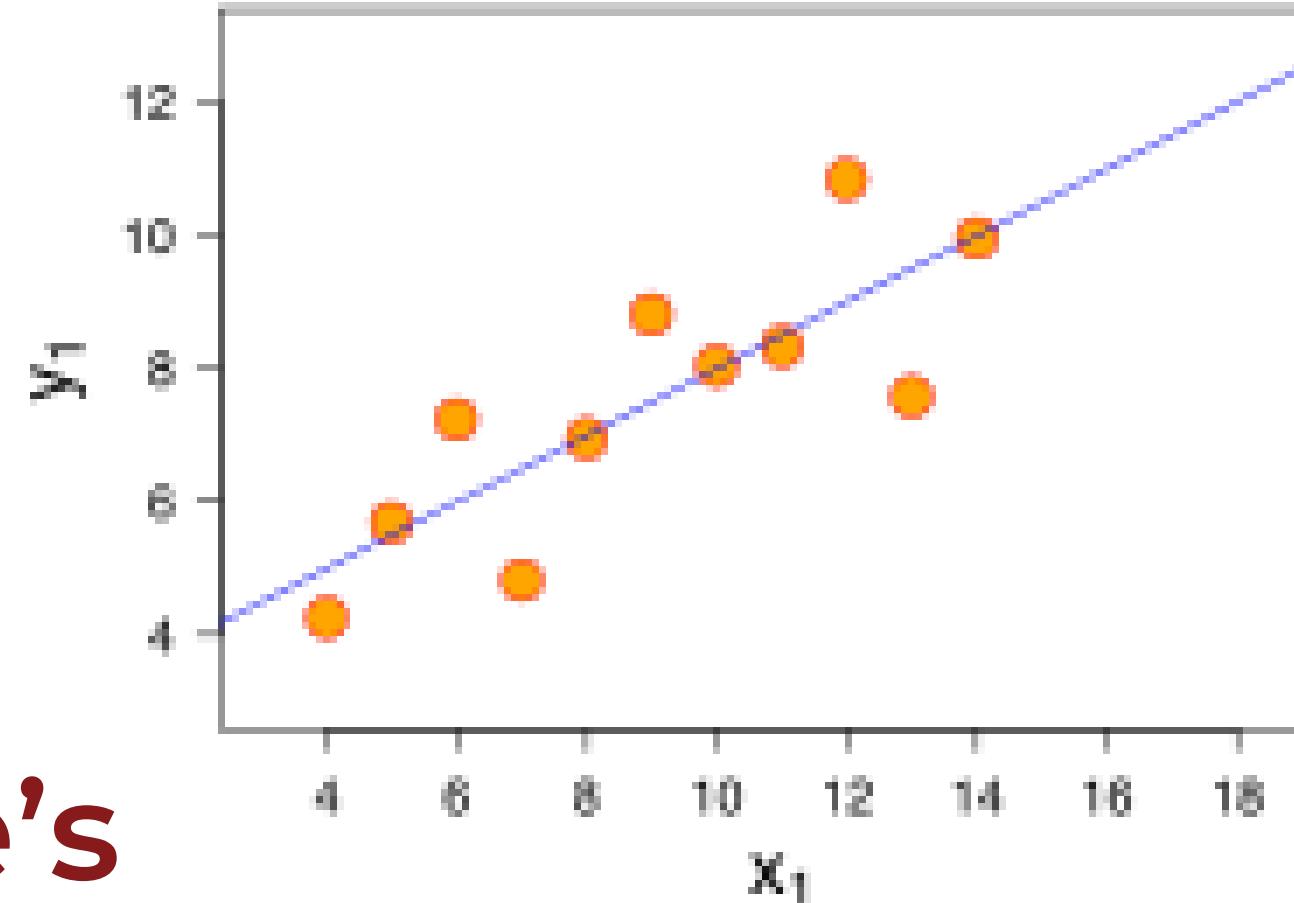
	Dataset I		Dataset II		Dataset III		Dataset IV	
	x	y	x	y	x	y	x	y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean								
Variance								
Correlation								

	Dataset I		Dataset II		Dataset III		Dataset IV	
	x	y	x	y	x	y	x	y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

	Dataset I		Dataset II		Dataset III		Dataset IV	
	x	y	x	y	x	y	x	y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	



Anscombe's Quartet



Uses of Data Visualization

Uses of Data Visualization

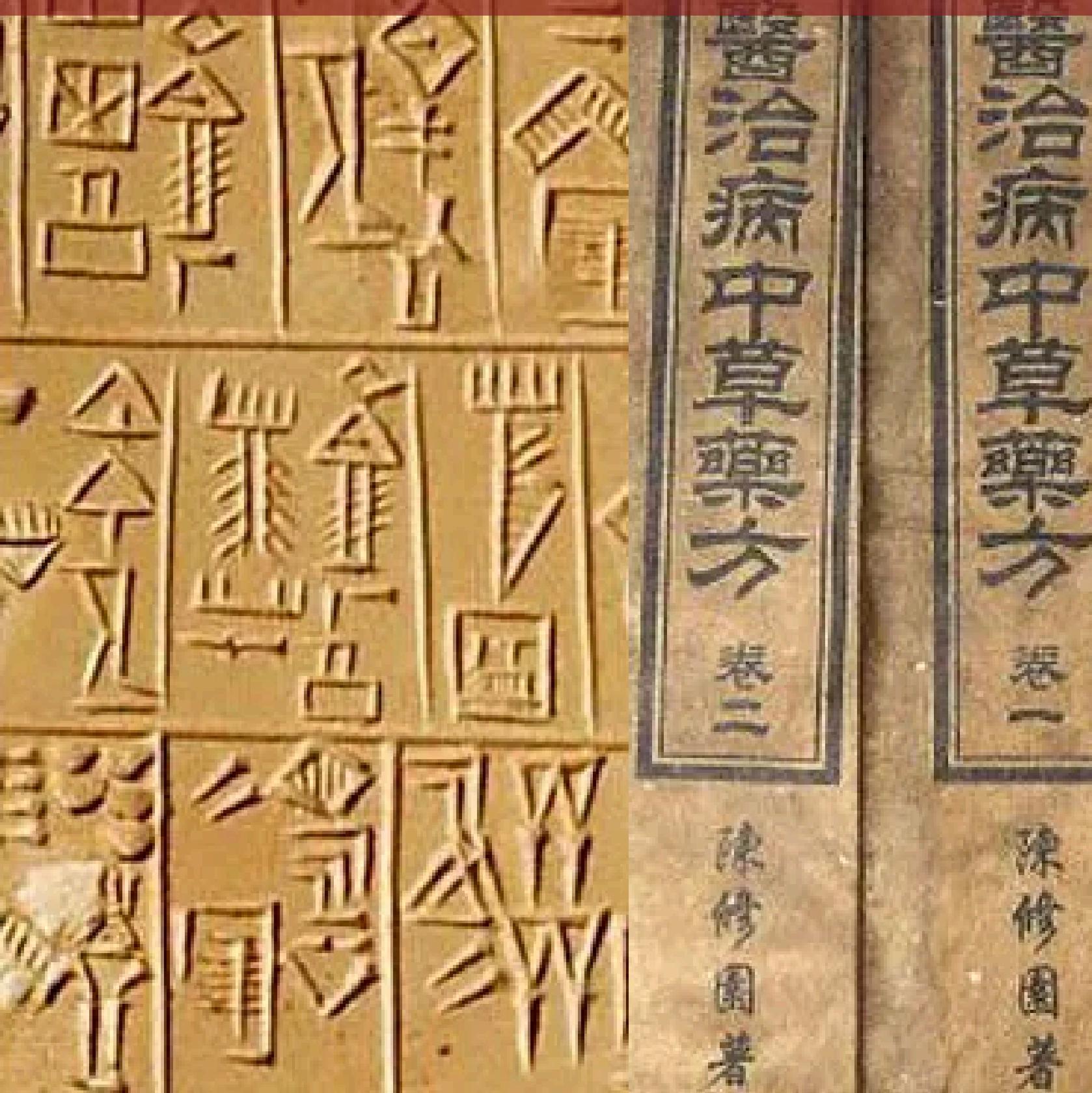
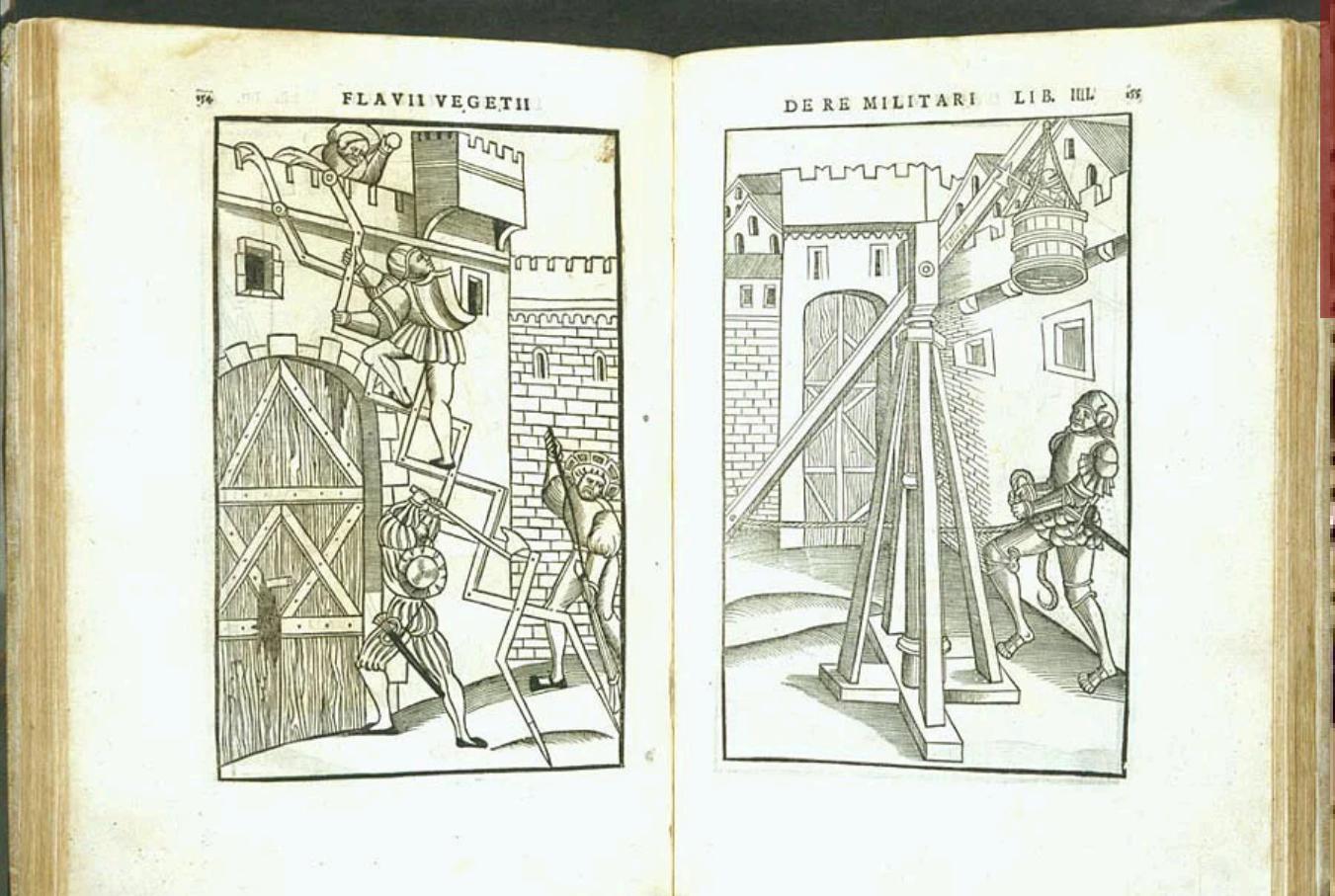
- Record Information
- Support Analysis & Reasoning
- Confirm Hypotheses
- Communicate Ideas

Record Information

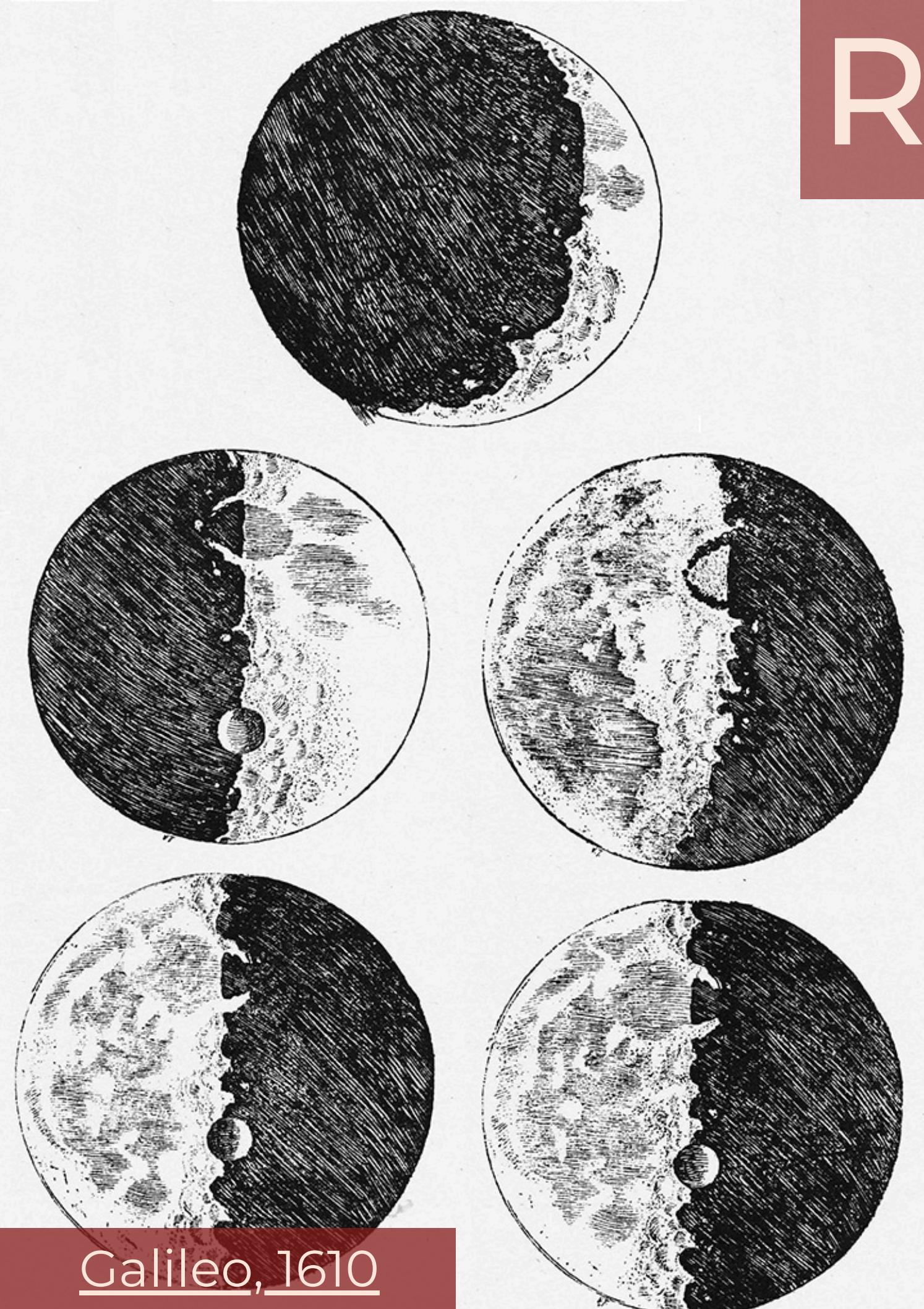
Record Information



Record Information



Record Information



Galileo, 1610



Muybridge, 1878



da Vinci, 1485

Record Information

Data Therapy, Emily and Rahul Bhargava



Data Therapy, Emily and Rahul Bhargava

<https://ejatlas.org>

Support Analysis & Reasoning

Rapid Transit & Frequent Bus Routes

Support Analysis
& Reasoning



Mapping Migration in the United States

Share full article



54

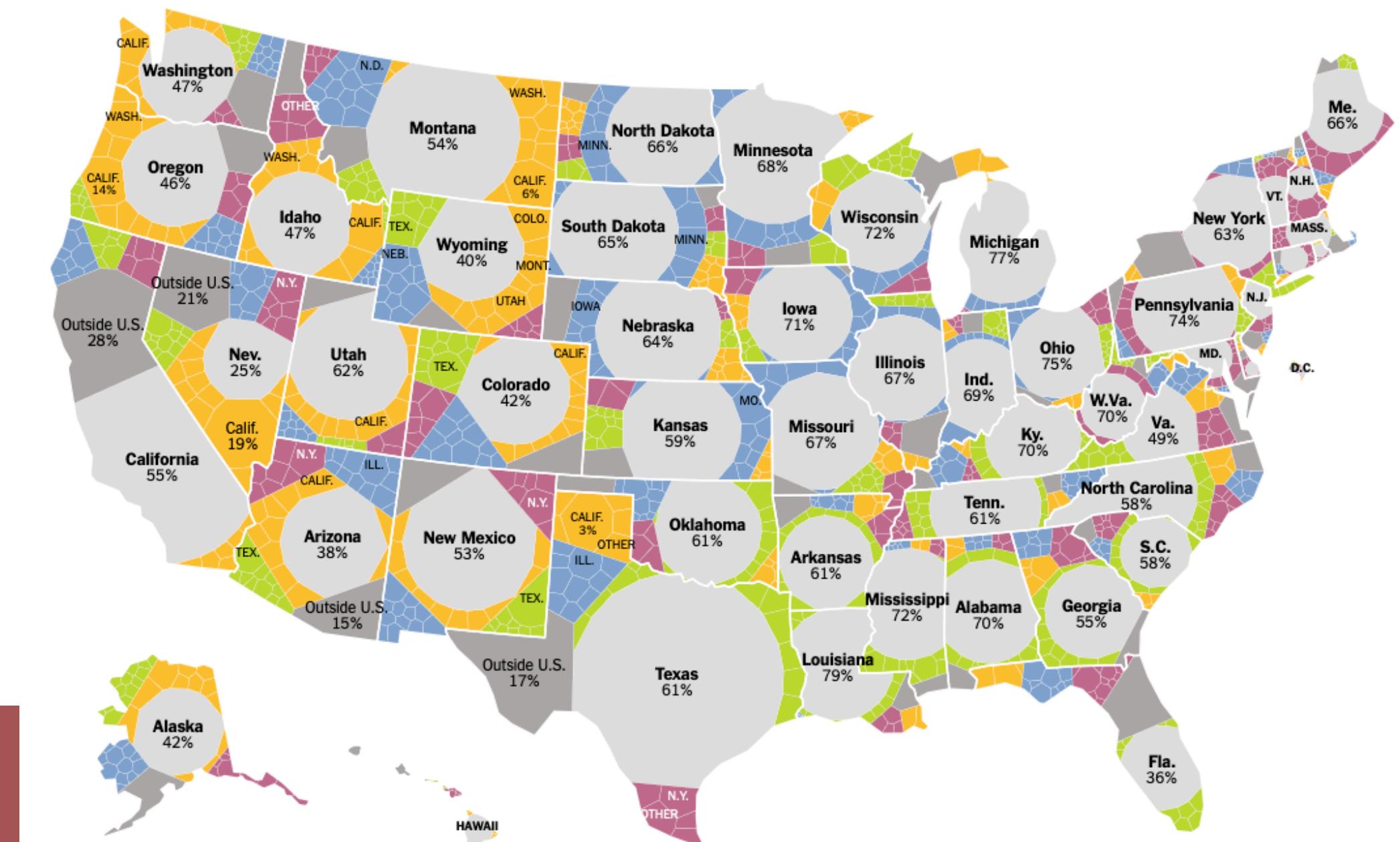
Where people who lived in each state in 2012 were born

Each shape represents where the people living in a state were born. Within a state, larger shapes mean a group makes up a larger share of the population.

SELECT A YEAR

1900 | 1950 | **2012**

■ Northeast ■ South ■ Midwest ■ West ■ Outside the U.S.*



Support Analysis & Reasoning

By Gregor Aisch and Robert Gebeloff

Aug. 15, 2014

New York Times

Confirm Hypotheses

Confirm Hypotheses

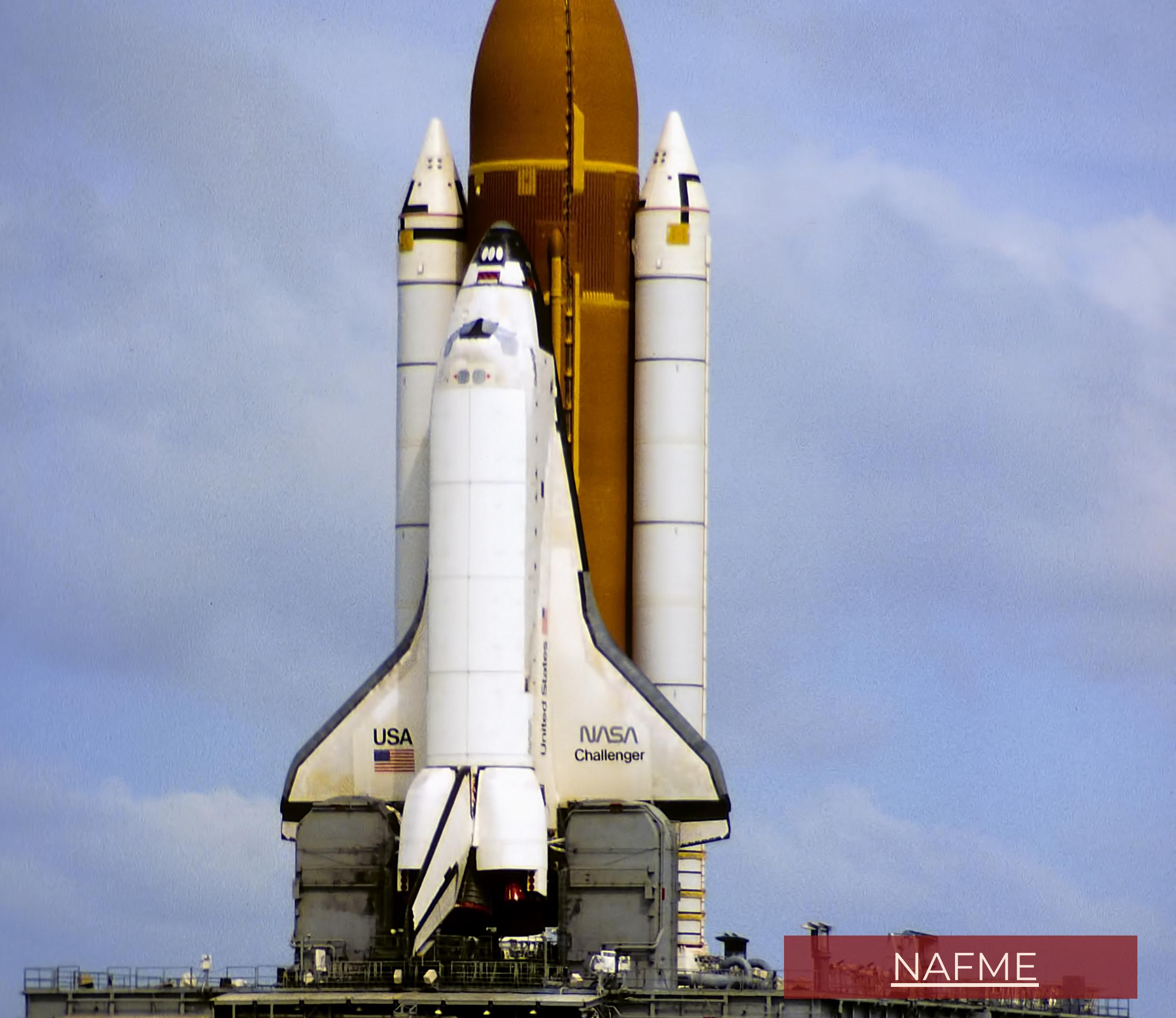
The Challenger
Disaster, 1986



[Wikipedia](#)

Confirm Hypotheses

The Challenger
Disaster, 1986



NAFME

Confirm

Hypotheses

The Challenger
Disaster, 1986

HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

MOTOR	O-RING	MFT	SRM No.	Cross Sectional View			Top View		Clocking Location (deg)
				Erosion Depth (in.)	Perimeter Affected (deg)	Nominal Dia. (in.)	Length Of Max Erosion (in.)	Total Heat Affected Length (in.)	
DM-4	47		22A	None	None	0.280	None	None	36° - 66°
DM-2	52		22A	NONE	NONE	0.280	NONE	NONE	338° - 18°
QM-3	48		15A	0.010	154.0	0.280	4.25	5.25	163
QM-4	51		15B	0.038	130.0	0.280	12.50	58.75	354
QM-4	51		15B	None	45.0	0.280	None	29.50	354
SRM-15	53		41D RH Forward Field	138	0.028	110.0	0.280	3.00	None
SRM-22	75		41C LH Aft Field*	11A	None	None	0.280	None	--
SRM-25	29		41B LH Forward Field	10A	0.040	217.0	0.280	3.00	14.50
SRM-25	27		STS-2 RH Aft Field	29	0.053	116.0	0.280	--	--
									90

*Hot gas path detected in putty. Indication of heat on O-ring, but no damage.
**Soot behind primary O-ring.
***Soot behind primary O-ring, heat affected secondary O-ring.

Clocking location of leak check port - 0 deg.

OTHER SRM-15 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY AND NO SOOT NEAR OR BEYOND THE PRIMARY O-RING.

SRM-22 FORWARD FIELD JOINT HAD PUTTY PATH TO PRIMARY O-RING, BUT NO O-RING EROSION AND NO SOOT BLOWBY. OTHER SRM-22 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY.

Blow By History

SRM-15 Worst Blow-By

- 2 CASE JOINTS (90°), (110°) Arc
- MUCH WORSE VISUALLY THAN SRM-22

SRM-22 Blow-By

- 2 CASE JOINTS (30-40°)

SRM-13A, 15, 16A, 18, 23A 24A

- NOZZLE Blow-By

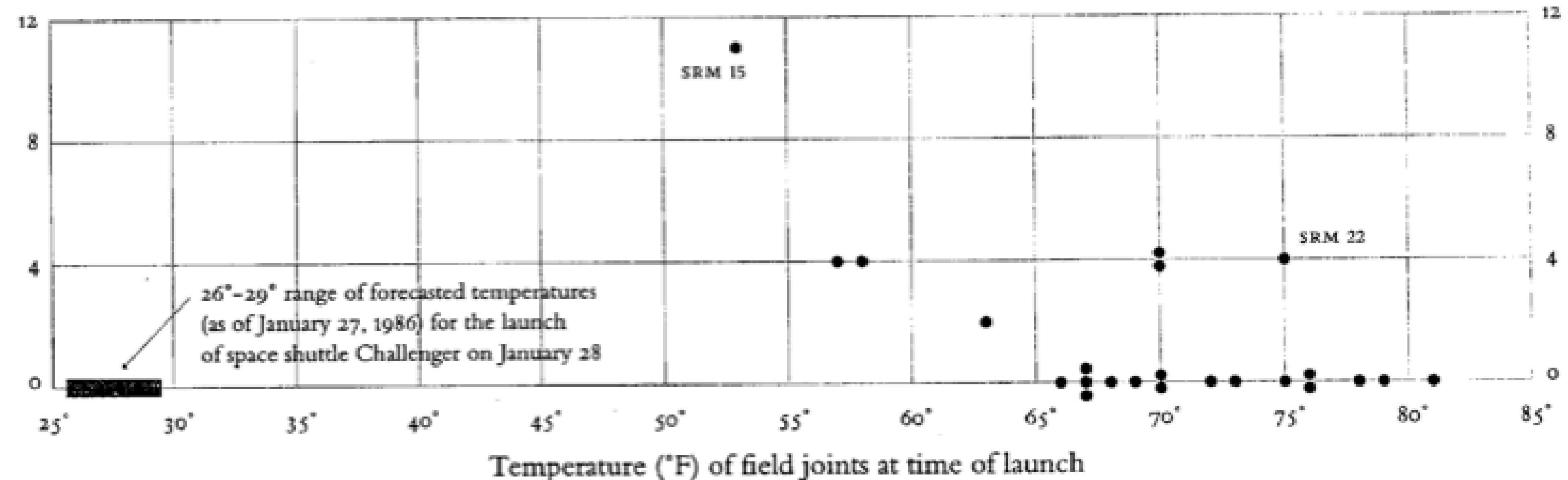
HISTORY OF O-RING TEMPERATURES (DEGREES - F)

MOTOR	MFT	AMB	O-RING	WIND
DM-4	68	36	47	10 MPH
DM-2	76	45	52	10 MPH
QM-3	72.5	40	48	10 MPH
QM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29	10 MPH
			27	25 MPH

Edward Tuft

Confirm Hypotheses

O-ring damage
index, each launch



The Challenger Disaster, 1986

[Edward Tuft](#)

Confirm Hypotheses

London Cholera
Outbreak, 1854



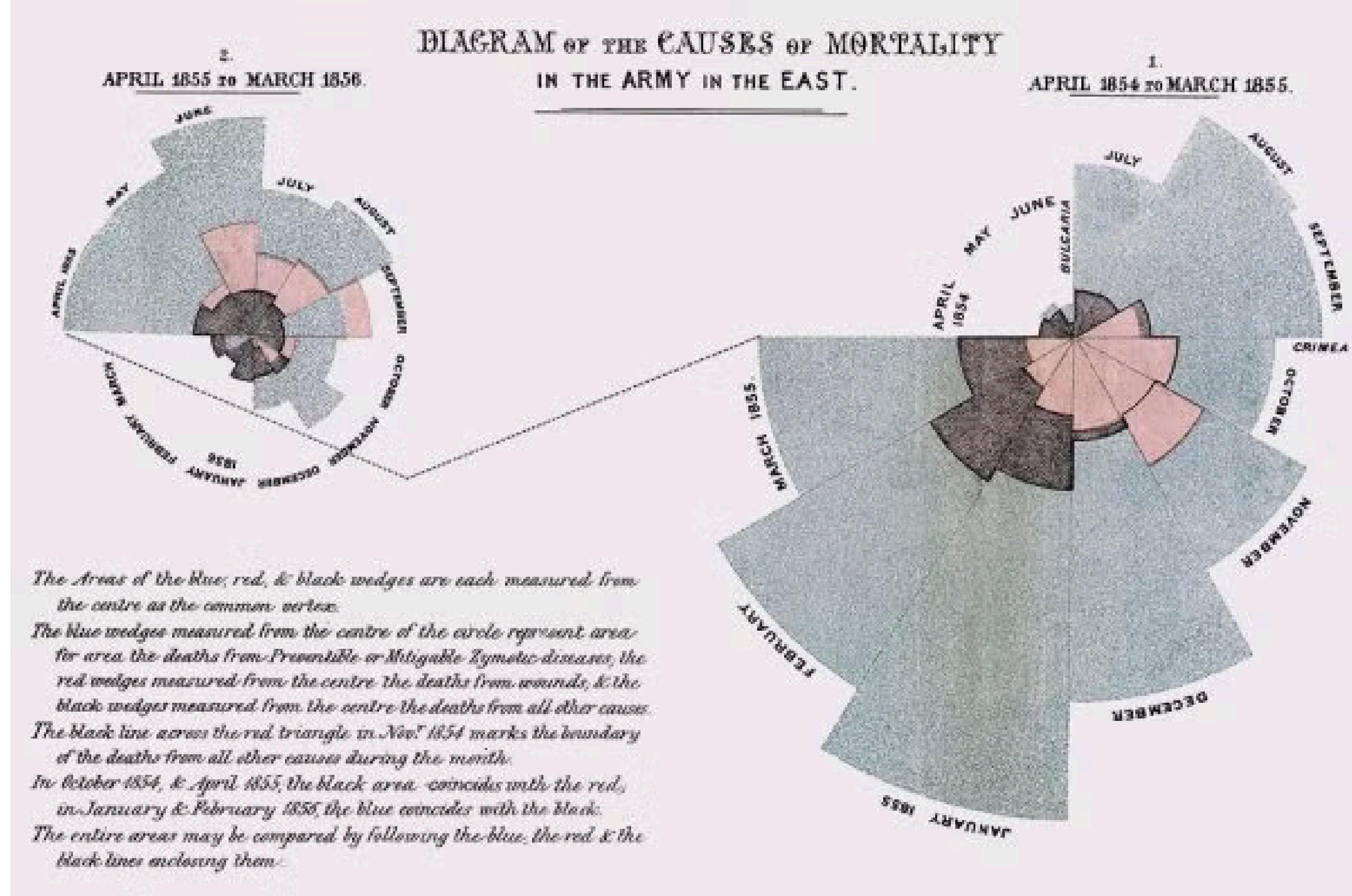
Confirm Hypotheses

London Cholera Outbreak, 1854



Communicate Ideas

DIAGRAM OF THE CAUSES OF MORTALITY
IN THE ARMY IN THE EAST.



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite

Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mohilow et qui rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

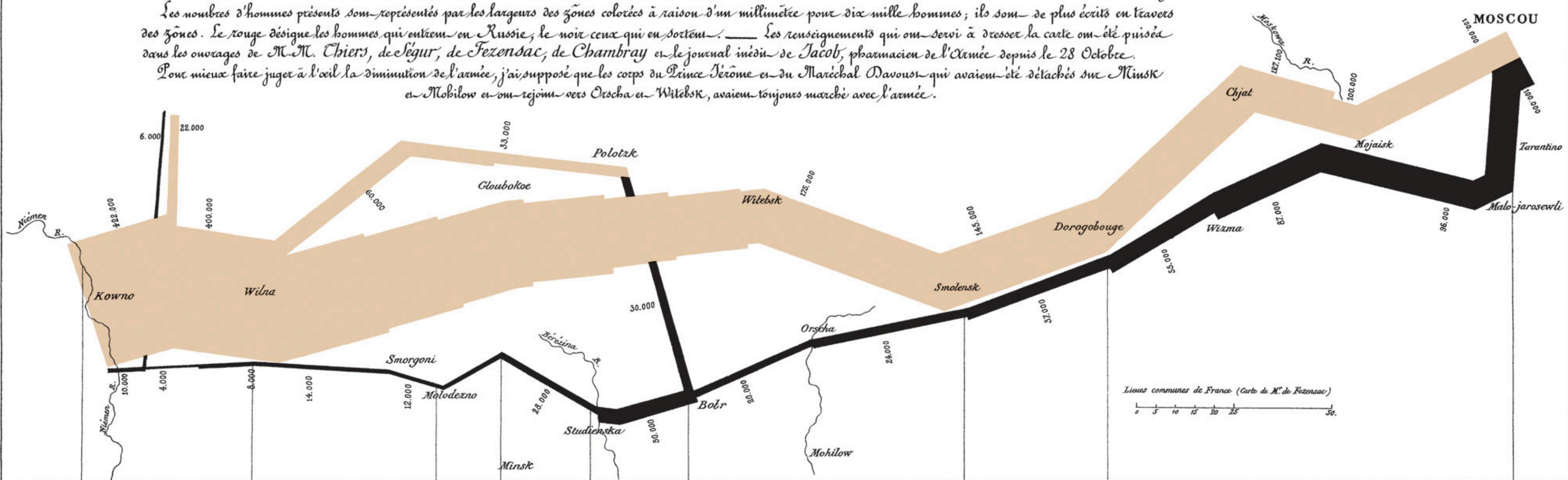
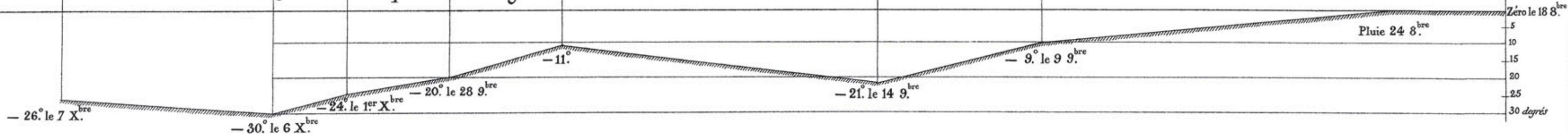


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop
le Niémen gelé.



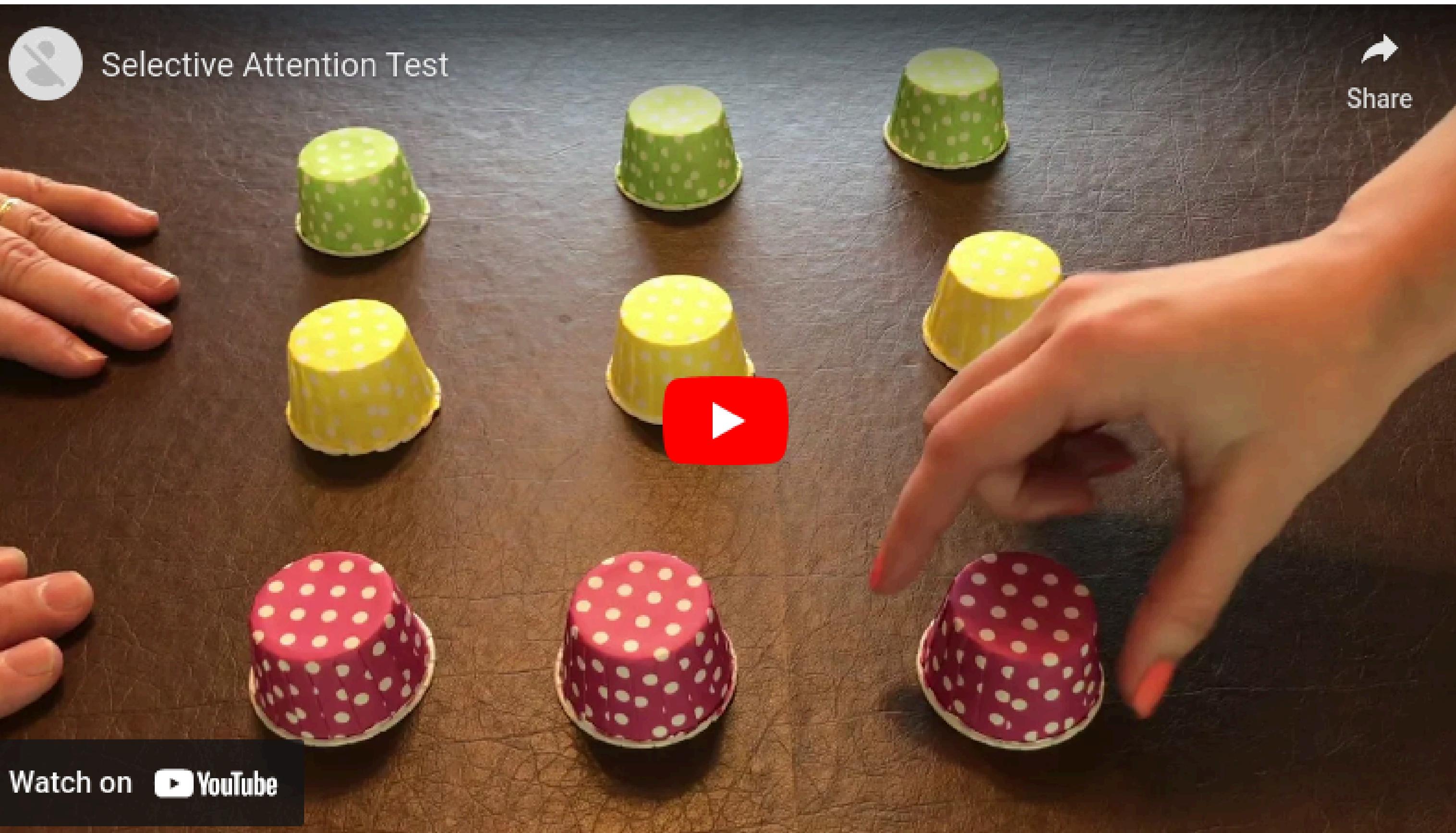
Autog. par Regnier, 8. Pas. St^e Marie St^e Gain à Paris.

Imp. Lith. Regnier et Dourdet.

Joseph Minard, 1869

Communicate Ideas

Why Does Visualization Work?



1. Perception & Cognition are Limited

We Can Use Visualization
to Augment Our
Perception & Cognition

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPSNVRCEDI
PSGHLMFHKMVPSTWYACEQTWRN

MTHI **V**I WY ADCE QGH KILK MTWYN
ARDCAIRE QGH L **V**KM FP STW YARN
GFP S **V**C E I L QG KMF PSN VR CE QDI
PS GHL MF HKM **V**P STW YACE QTWRN

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSSVCEILQGKMFPSNVRCEDQDI
PSGHLMFHKMVPSWYACEQTLWRN

But Visualization Can Also Deceive You

2. Memory is Limited

$$34 \times 28 = ?$$

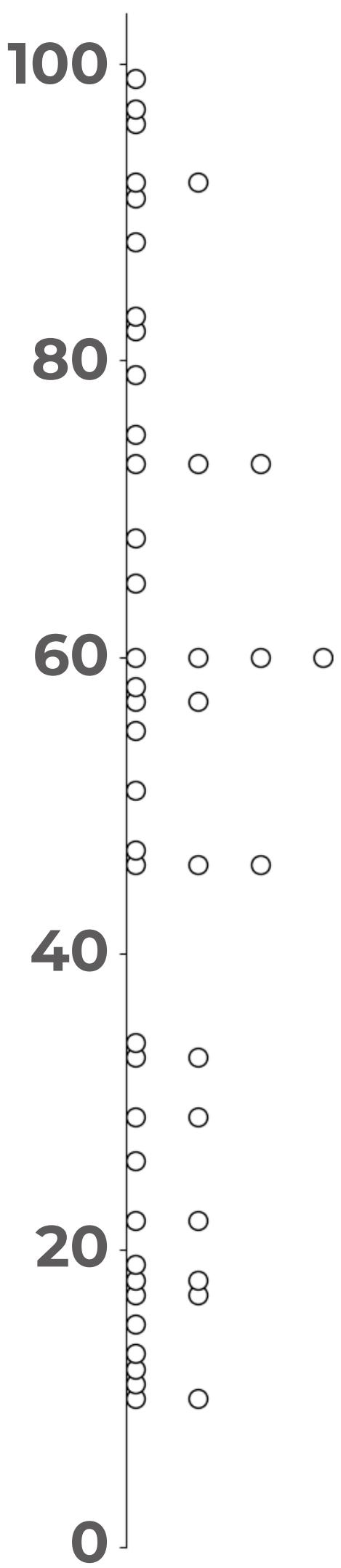
$$79 \times 16 = ?$$

$$34 \times 28 = ?$$

We Can Enhance
Our Working Memory
with Visualization

15	19	60
33	11	75
60	34	79
18	51	92
73	22	13
17	10	68
73	18	55
65	46	29
60	73	22
46	82	97
10	58	46
57	17	83
26	99	33
88	92	60
91	29	57
96	12	47

Which of these
numbers appears
most often?



Which of these
numbers appears
most often?

TRIGLYCERIDE LEVEL

Income Group	Men		Women	
	Under 65	65 or Over	Under 65	65 or Over
\$0-\$24,999	250	200	375	550
\$25,000+	430	300	700	500

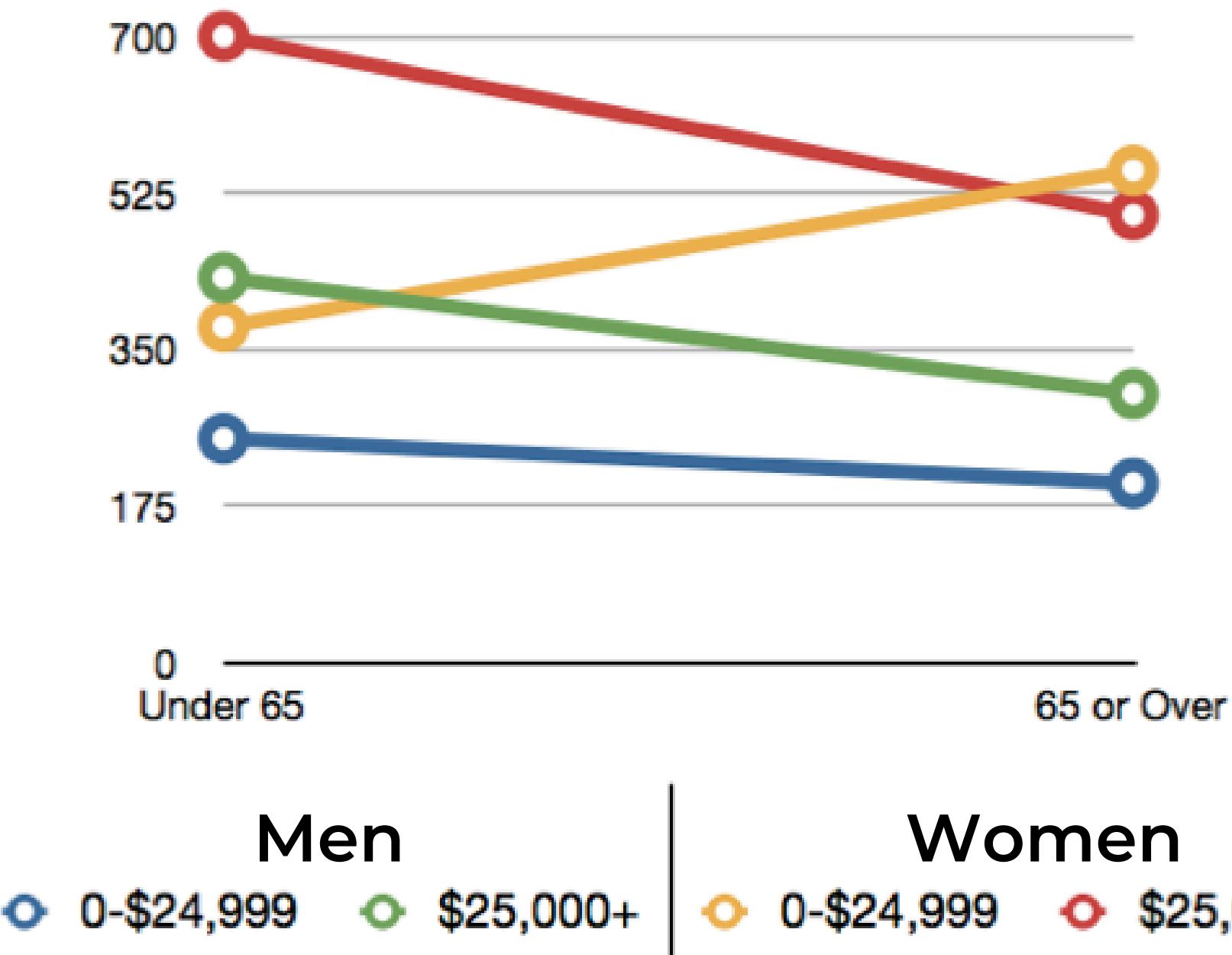
**For which gender and income group
does age affect triglyceride levels
differently?**

TRIGLYCERIDE LEVEL

Income Group	Men		Women	
	Under 65	65 or Over	Under 65	65 or Over
\$0-\$24,999	250	200	375	550
\$25,000+	430	300	700	500

**For which gender and income group
does age affect triglyceride levels
differently?**

TRIGLYCERIDE LEVEL



For which gender and income group
does age affect triglyceride levels
differently?

Course Goals

Course Goals

Design Effective Visualizations for Analysis

Course Goals

Design Effective Visualizations for Analysis

Techniques for a Broad Range of Data Types

Course Goals

Design Effective Visualizations for Analysis

Techniques for a Broad Range of Data Types

Develop Your Own Interactive Vis Tools

Disclaimer!

Knowledge of HTML, CSS, and Javascript Is Required

- We will cover these topics very briefly

Coding Assignments and Project will use D3js

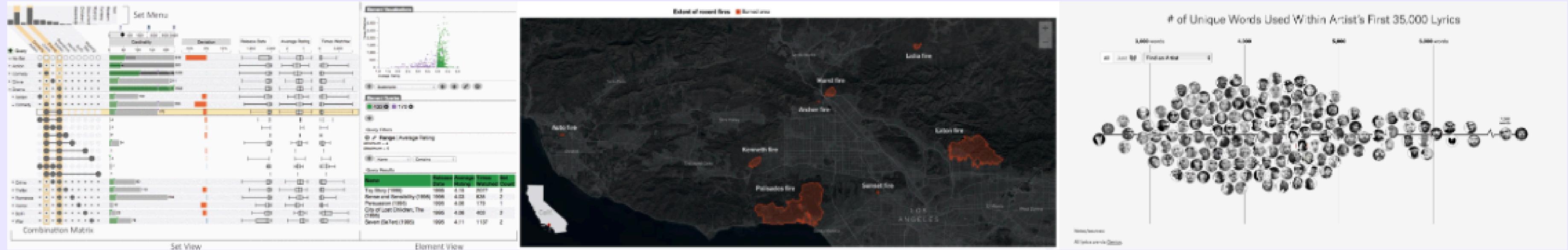
- Can have a steep learning curve

Course Logistics

Course Website

CS 571 - Data Visualization and Exploration
Spring 2025 - UMass Amherst

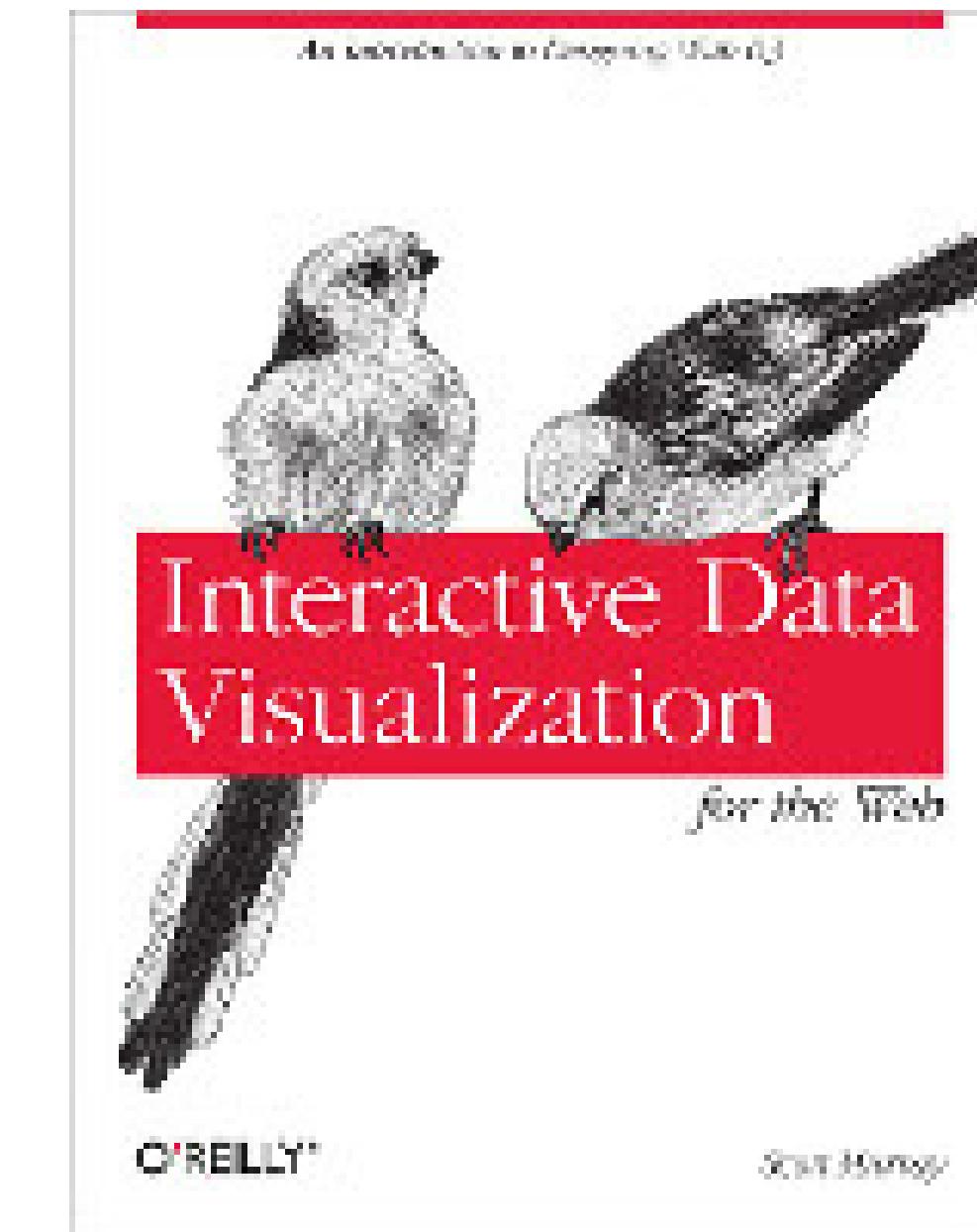
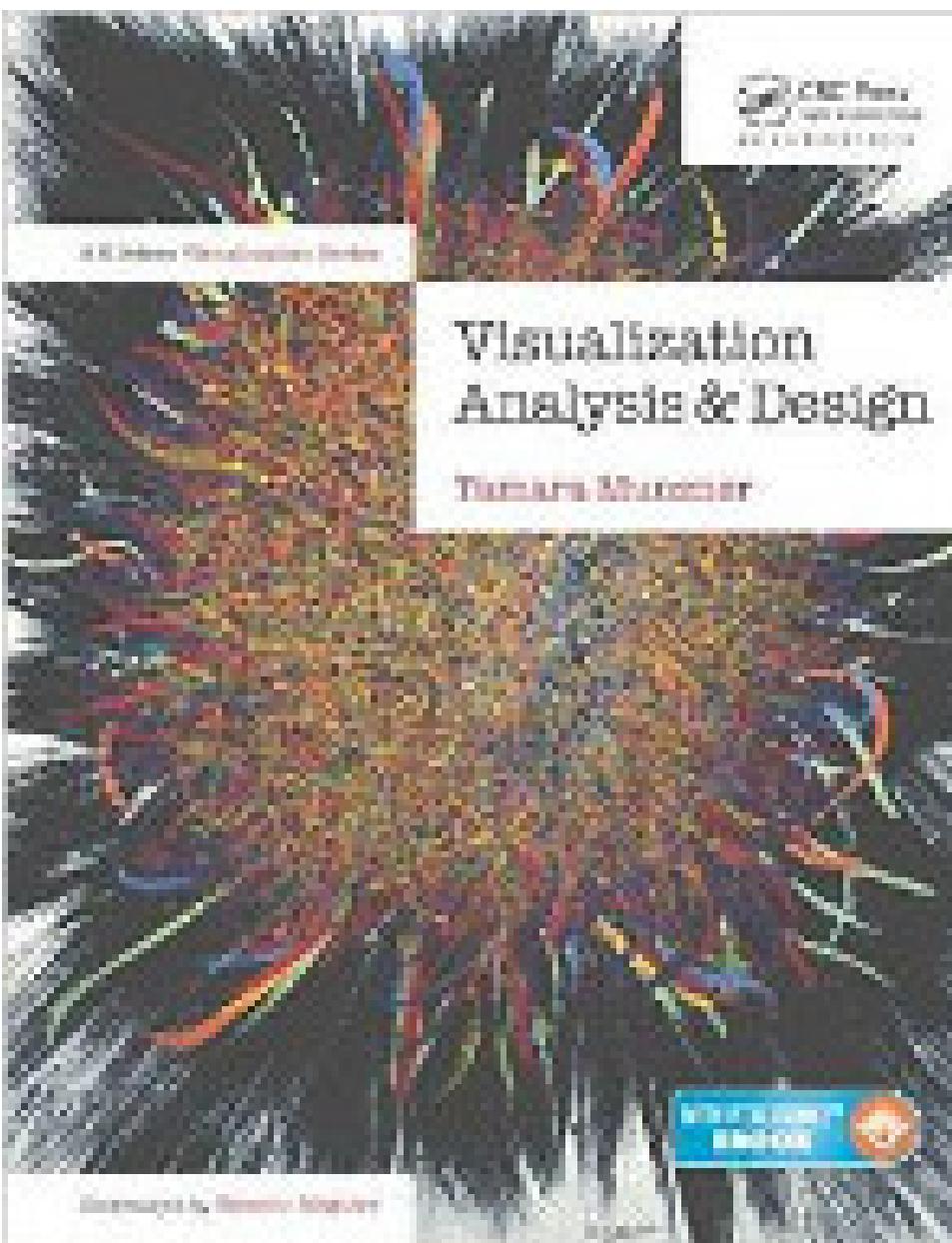
[Home](#) [Syllabus](#) [Schedule](#) [Project](#)



Data is increasingly ubiquitous, produced by scientists, businesses, engineers, and everyday people. We often use such data to help us make informed decisions and communicate effectively. Since data is often large and complex, visual representations of the data are often essential to facilitate our understanding. We call these visual representations "visualizations". But what makes a good visualization? Good visualizations not only visually represent data, but also improve our comprehension, communication, and decision-making.

In this course you will learn about the fundamentals of human perception, the theory of visualization, and good visualization design practices. You will also develop your own web-based visualizations using HTML5, CSS, JavaScript, SVG, and D3js.

Recommended Textbooks



Education Tools/Software

Piazza (Discussion Forum):

- Ask questions about content/assignments
- Find a team for the project



Gradescope:

- Submit assignments and project materials



Grade Breakdown

Project	40%
Programming & Design Assignments	35%
In-Class Group Activities	15%
Canvas Quizzes	10%

Programming & Design Assignments

There will be 5 assignments throughout the semester (not including Homework 0)

**Assignments will be due on Friday at 5:00PM
Eastern Time**

Your work must be your own!

Programming & Design Assignments

Late Policy:

- Up to 2 days late for each assignment
- Each day late reduces the maximum possible percentage by 1 point (out of 10)
- e.g., an assignment submitted
 - 1 day late can receive max of 9 points
 - 2 days late can receive max of 8 points

Programming & Design Assignments

Regrade Policy:

- You may submit a regrade request on Gradescope within 7 days of an assignment's due date
- Your regrade request must explain why you believe your answer is correct

Project

Projects will be completed in 3-person teams

You and your team will:

- Choose a topic that interests you
- Design a website with interactive visualizations that answers questions related to your topic

Project Timeline (dates may change)

Announce Your Project	Feb 7
Project Proposal	Feb 21
Project Review	Feb 24 - Mar 7
Project Milestone	Apr 4
Peer Feedback	Apr 16
Project Screencast	May 2
Final Project Submission + Group Feedback	May 12

Project Timeline (dates may change)

Announce Your Project	Feb 7
Project Proposal	Feb 21
Project Review	Feb 24 - Mar 7
Project Milestone	Apr 4
Peer Feedback	Apr 16
Project Screencast	May 2
Final Project Submission + Group Feedback	May 12

In-Class Group Activities & Attendance

I will not grade attendance for this class.
However, there will occasionally be graded in-class, group activities

You will form 2-3 person groups
(different groups each time) and submit a single Google Doc for these assignments

In-Class Group Activities & Attendance

**These assignments will be due by 5:00PM
Eastern Time the day that they are assigned**

**We will only grade submissions from students
who attended class that day!**

In-Class Group Activities & Attendance

If you must be absent for an excusable reason (per UMass policy), you may inform me by email before the start of the class, and:

In-Class Group Activities & Attendance

If you must be absent for an excusable reason (per UMass policy), you may inform me by email before the start of the class, and:

- If there will be a group activity that class, I will offer you the option to submit the assignment individually

In-Class Group Activities & Attendance

If you must be absent for an excusable reason (per UMass policy), you may inform me by email before the start of the class, and:

- If there will be a group activity that class, I will offer you the option to submit the assignment individually
- If there will be a group activity the next class, I will let you know

Canvas Quizzes

**At-Home Quizzes will be due on Canvas
approximately every 2 weeks**

Quizzes will have no more than 10 questions

**You will have unlimited attempts to take these
quizzes before the due date**

Canvas Quizzes

Unless otherwise specified, Quizzes will be:

- Released Thursday at 3:00PM Eastern Time
- Due Monday at 5:00PM Eastern Time

Reading Assignments (not graded)

We will post readings (from the textbook, online articles, or videos) each week

These readings will help you

- learn design & programming concepts,
- familiarize yourself with terminology, and
- notice which areas may need attention

Upcoming Due Dates

Homework 0 (Due Monday, Feb 3)

- Releases on Gradescope at 2:30PM today

Announce Your Project (Due Friday, Feb 7)

- Google Form will be released

Tuesday, Feb 4

- Start looking for teammates now