Fachhochschule Kiel
University of Applied Sciences

Data Visualization and Visual Analytics
Introduction

Study Program Data Science
Prof. Dr. Tillmann Schwörer


Why do we visualize data?

“We are drowning in information, while starving for wisdom”

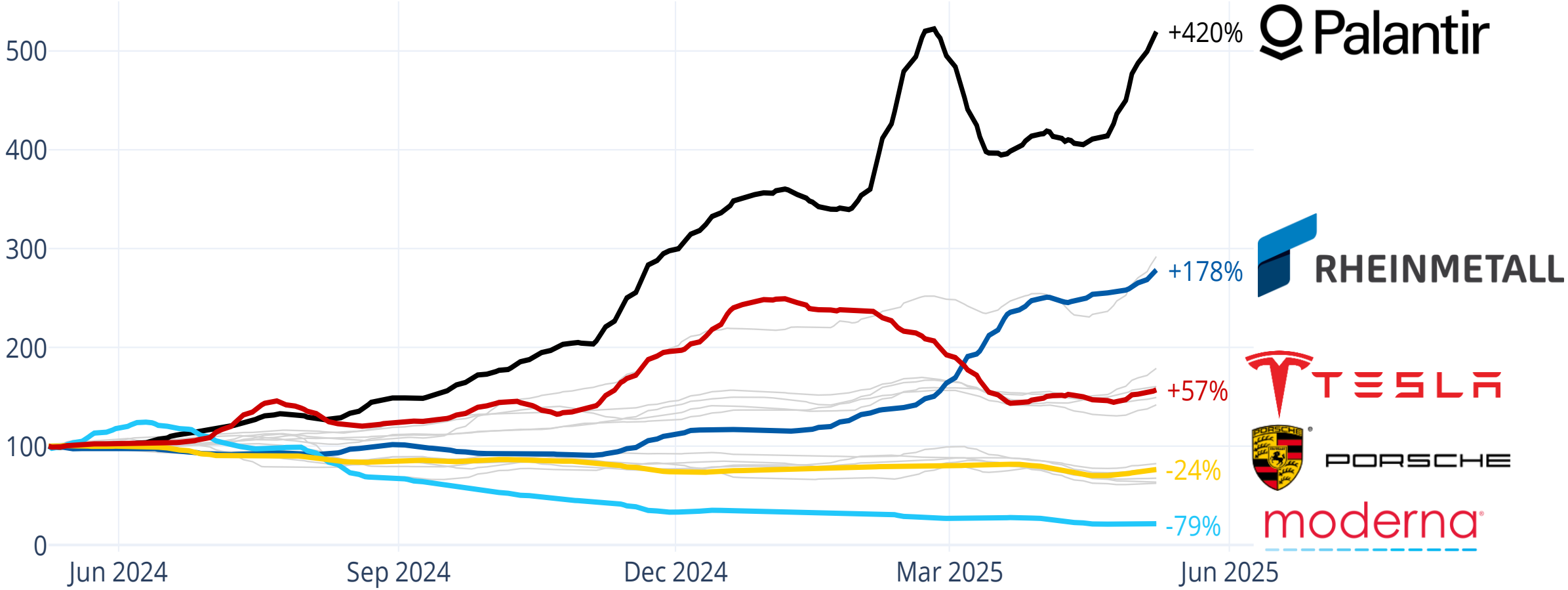
Edward Osborne Wilson, Biologist

Our brain is not trained in interpreting text and numbers

Ticker	Airbnb	Siemens Energy	IBM	Intel	Mercedes-Benz	LVMH Moët Hennessy Louis Vuitton	Moderna	Netflix	Nike	Porsche	Palantir	Rheinmetall	SAP	TESLA	Walmart
Date															
2024-05-09	147.05	23.30	162.46	29.90	62.91	769.96	122.69	612.09	91.68	47.05	21.14	535.88	175.56	171.97	59.97
2024-05-10	146.32	24.28	163.32	29.66	62.76	769.37	117.31	610.87	89.27	46.75	20.60	529.94	173.86	168.47	60.01
2024-05-13	149.13	24.19	163.72	30.32	62.92	763.71	125.67	616.59	91.02	47.42	20.94	530.34	174.47	171.89	59.94
2024-05-14	146.70	24.14	163.52	30.85	63.81	772.79	128.32	613.66	91.09	47.91	21.44	515.90	172.50	177.55	59.39
2024-05-15	145.80	25.71	164.40	31.07	63.50	769.96	129.06	613.52	89.99	47.47	21.67	515.40	173.90	173.99	59.37
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2025-05-02	125.26	72.60	245.55	20.62	49.05	495.85	27.61	1156.49	58.59	36.76	124.28	1576.00	266.30	287.21	98.75
2025-05-05	124.89	73.34	249.18	20.27	49.63	492.25	27.84	1134.06	57.35	36.45	123.77	1627.00	267.35	280.26	99.33
2025-05-06	121.67	72.92	249.12	19.94	49.81	489.10	24.43	1137.69	57.04	36.66	108.86	1645.00	264.35	275.35	98.55
2025-05-07	123.12	70.98	253.37	20.31	49.69	483.50	24.06	1155.41	58.62	36.64	110.48	1633.50	261.50	276.22	98.83
2025-05-08	126.39	73.34	254.14	21.00	50.87	487.30	24.38	1144.43	58.91	37.04	119.15	1701.00	262.80	284.82	97.43

...but in visual patterns

Stock Price (Indexed to 100)



Patterns in raw data are easily overlooked

Mannschafts-Pässe 2024-2025 Bundesliga

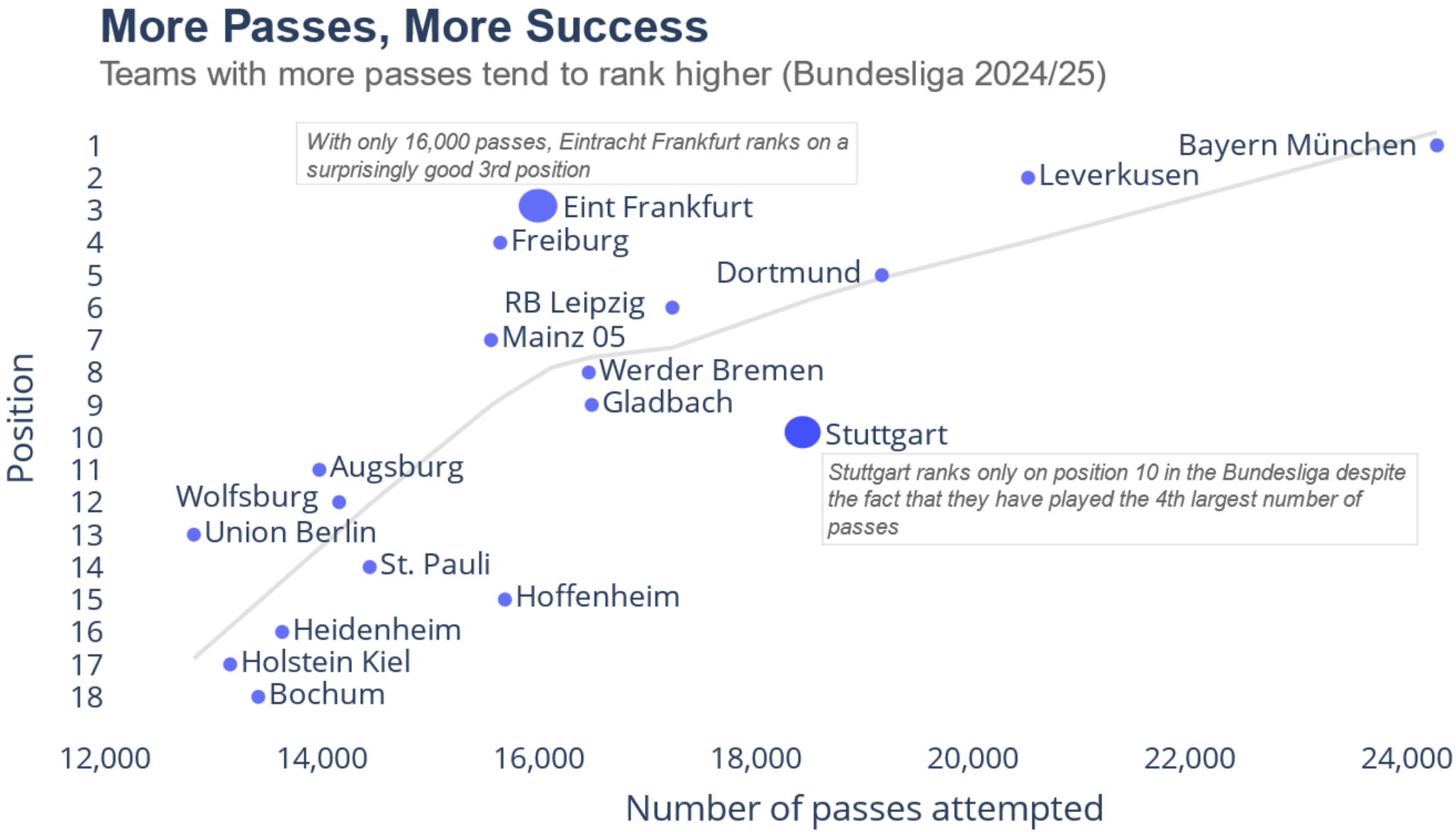
Spielerstatistiken ansehen | Glossar | Per90-Statistiken-Konvertierung

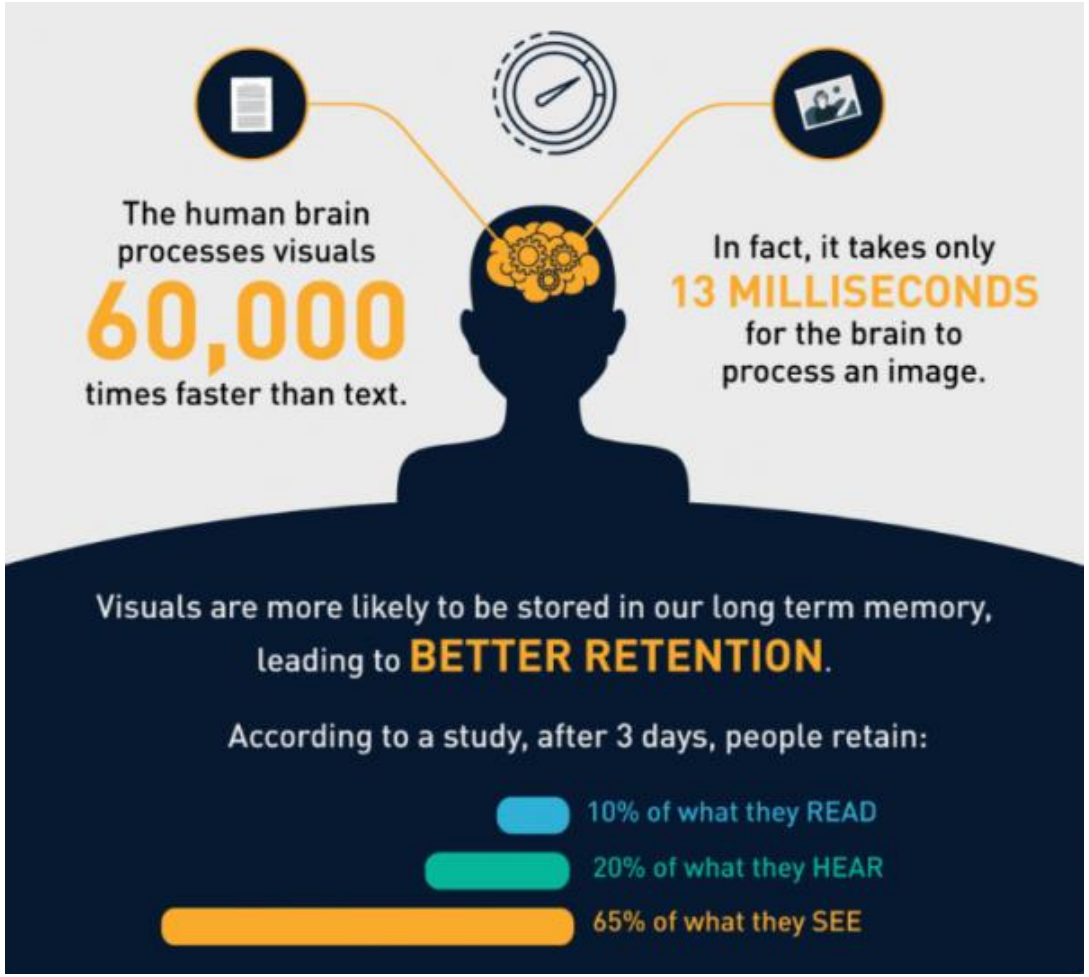
Mannschafts-Statistiken | Gegnerische Statistiken

			Gesamt					Kurz			Mittel			Lang					Erwartet					
Verein	# Sp	90	Vol ▼	Ver	Vol%	GesEntf	ProgEntf	Vol	Ver	Vol%	Vol	Ver	Vol%	Vol	Ver	Vol%	Vor	xAG	xA	A-xAG	WP	1/3	PPA	FlkSr	PrgP
Bayern München	29	32.0	21459	24276	88,4	354654	110768	10675	11442	93,3	8966	9750	92,0	1457	2197	66,3	61	56.2	70.5	+4.8	479	1997	433	67	1992
Leverkusen	23	32.0	17421	20508	84,9	285096	95299	8957	9740	92,0	6694	7574	88,4	1333	2238	59,6	50	39.6	49.3	+10.4	367	1554	304	79	1627
Dortmund	28	32.0	16073	19159	83,9	278282	93401	7311	8033	91,0	7115	7904	90,0	1323	2328	56,8	51	45.2	46.9	+5.8	353	1248	309	84	1468
Stuttgart	28	32.0	15434	18543	83,2	256148	91697	7689	8469	90,8	6226	7077	88,0	1170	2163	54,1	41	41.0	48.5	0.0	337	1221	345	83	1412
RB Leipzig	29	32.0	13886	17229	80,6	234831	86055	6420	7317	87,7	6117	7104	86,1	1029	1899	54,2	35	33.5	37.7	+1.5	296	1077	286	51	1271
Gladbach	26	32.0	13314	16486	80,8	230923	81753	6120	6841	89,5	5727	6613	86,6	1180	2178	54,2	39	35.4	32.4	+3.6	284	794	259	87	1040
Eint Frankfurt	26	32.0	13102	16121	81,3	218676	80927	6318	6973	90,6	5382	6201	86,8	1028	2069	49,7	42	41.8	41.2	+0.2	334	944	289	50	1249
Werder Bremen	23	32.0	13102	16459	79,6	233638	84753	5444	6234	87,3	5869	6839	85,8	1343	2350	57,1	36	34.4	39.9	+1.6	300	891	313	88	1263
Freiburg	26	32.0	12304	15643	78,7	218996	79334	5286	6068	87,1	5545	6421	86,4	1140	2266	50,3	32	28.5	34.6	+3.5	270	918	240	70	1080
Hoffenheim	34	32.0	12202	15686	77,8	214558	81414	5438	6235	87,2	5232	6248	83,7	1158	2306	50,2	28	31.4	37.5	-3.4	307	992	284	104	1157
Mainz 05	25	32.0	11836	15558	76,1	211018	81001	5375	6199	86,7	4973	5940	83,7	1213	2517	48,2	33	35.4	41.8	-2.4	292	975	288	85	1213
St. Pauli	28	32.0	11076	14440	76,7	196716	75143	4784	5508	86,9	4924	5830	84,5	1069	2267	47,2	21	25.4	27.5	-4.4	266	841	193	68	1024
Wolfsburg	27	32.0	10870	14158	76,8	205346	76679	4441	5249	84,6	4842	5646	85,8	1339	2510	53,3	36	30.6	34.0	+5.4	289	935	230	49	1138
Augsburg	29	32.0	10796	13976	77,2	199243	77921	4549	5317	85,6	4853	5658	85,8	1188	2239	53,1	25	23.3	31.4	+1.7	275	844	237	73	1097
Heidenheim	27	32.0	10254	13633	75,2	199584	77772	3793	4591	82,6	4889	5774	84,7	1349	2536	53,2	22	26.8	29.6	-4.8	274	770	217	79	905
Holstein Kiel	27	32.0	9938	13154	75,6	190244	75012	3835	4496	85,3	4752	5641	84,2	1189	2403	49,5	28	25.1	26.0	+2.9	253	646	171	64	832
Bochum	26	32.0	9679	13413	72,2	191306	80978	3630	4384	82,8	4497	5542	81,1	1365	2691	50,7	21	28.2	33.6	-7.2	292	811	223	74	1028
Union Berlin	28	32.0	9283	12819	72,4	166565	70509	4199	4980	84,3	3669	4618	79,4	1087	2262	48,1	24	23.0	31.0	+1.0	283	729	231	73	971

https://fbref.com/de/wettbewerbe/20/Bundesliga-Statistiken

… but can be made apparent using visualizations





Visualizations are faster perceived and longer memorized than text.

Interpreting statistical output

Descriptive statistics

	Average	Std Deviation
Income	3000 €	812 €
Tenure	5.0 years	3.3 years

Correlation(Tenure, Income) = 0.81

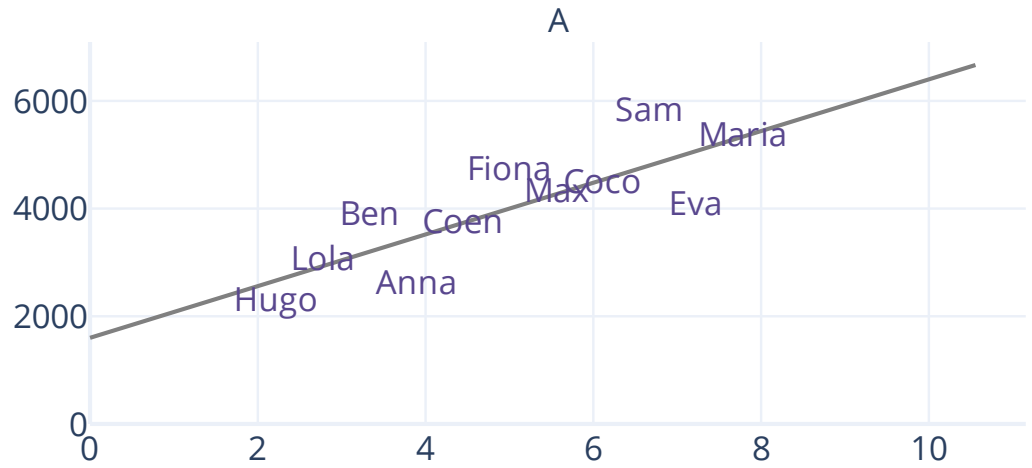
Linear Regression:

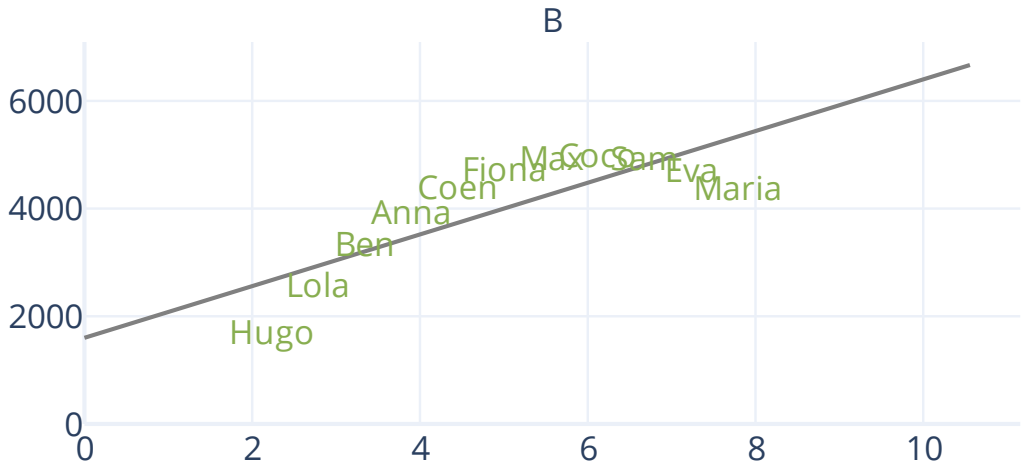
Income = 1600 + 480 * Tenure
(±600) (±113)

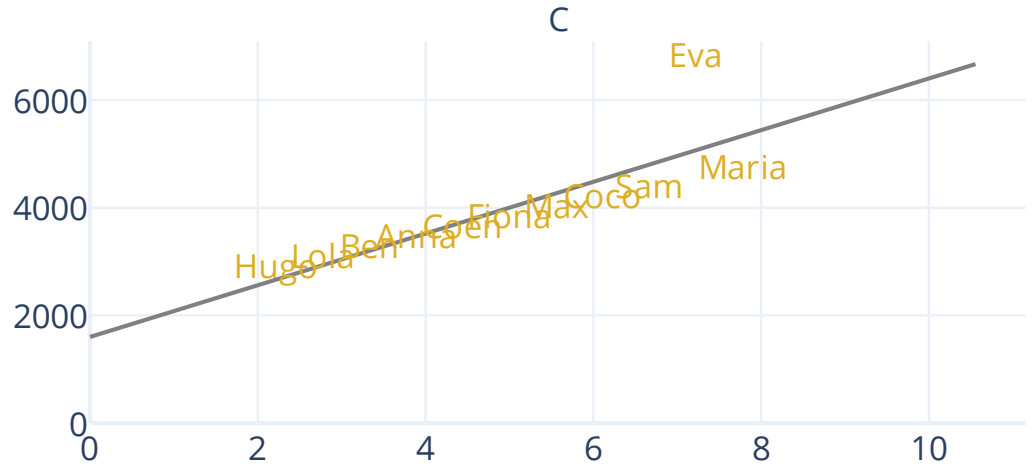
with N = 11 and R^2 = 0.66

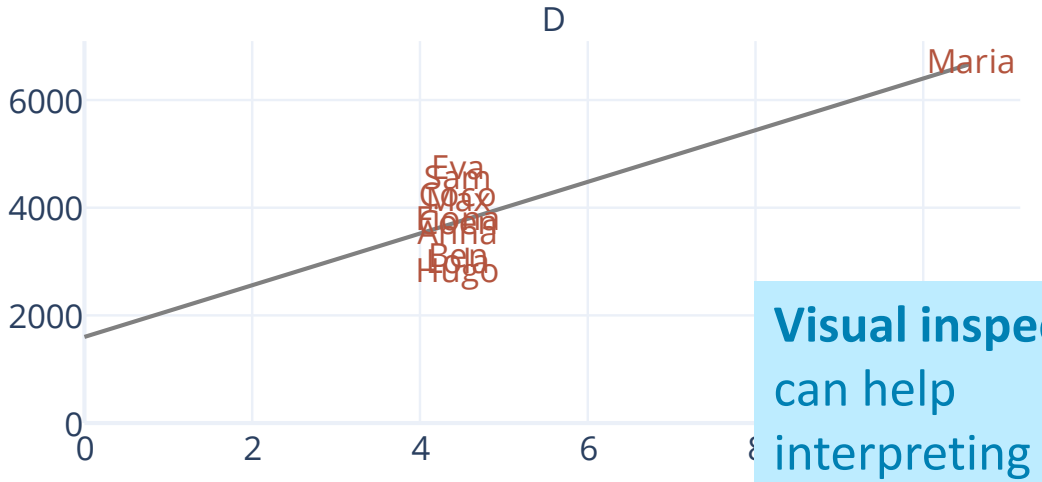
… can be misleading

All of the following visualizations are consistent with these statistics!





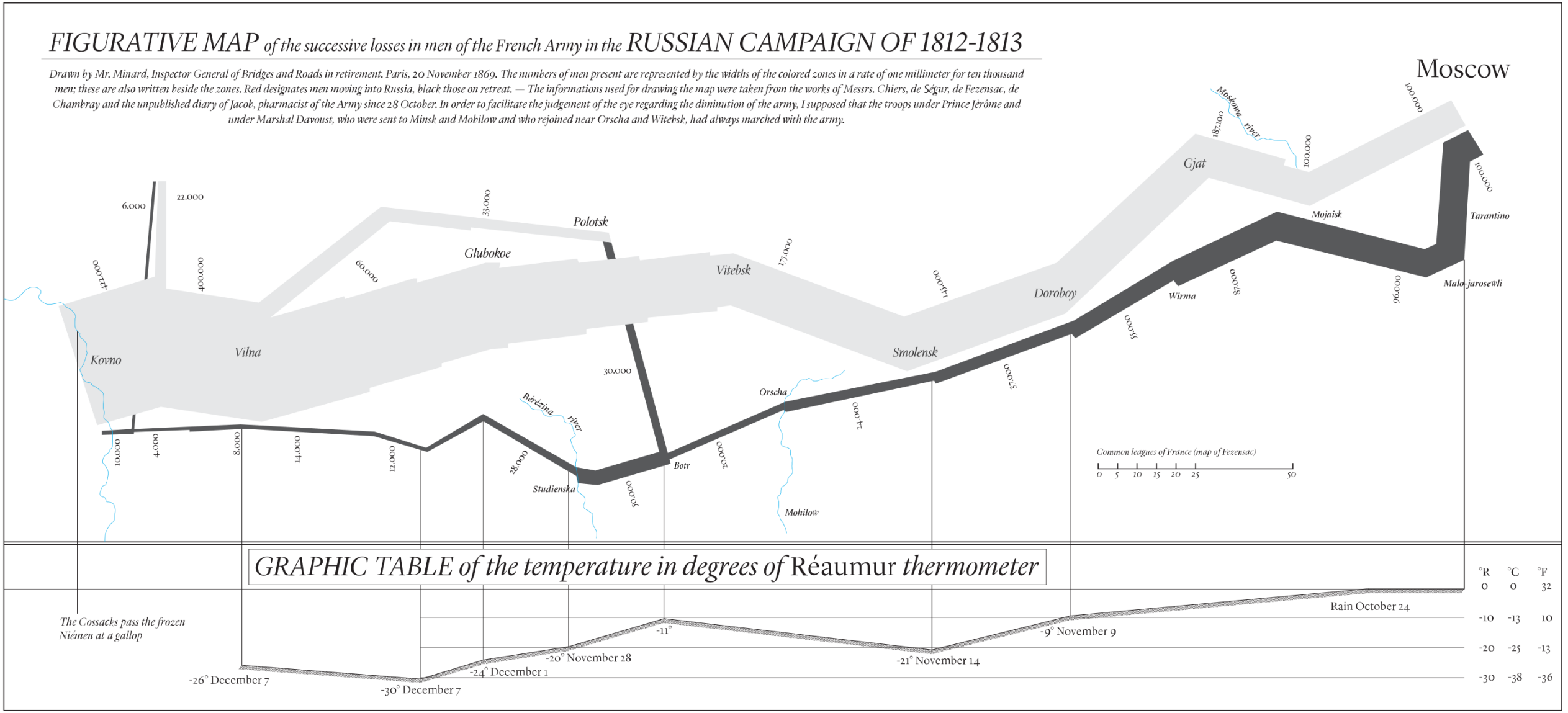




Visual inspection can help interpreting the data correctly!

A visualization worth a 1000 words

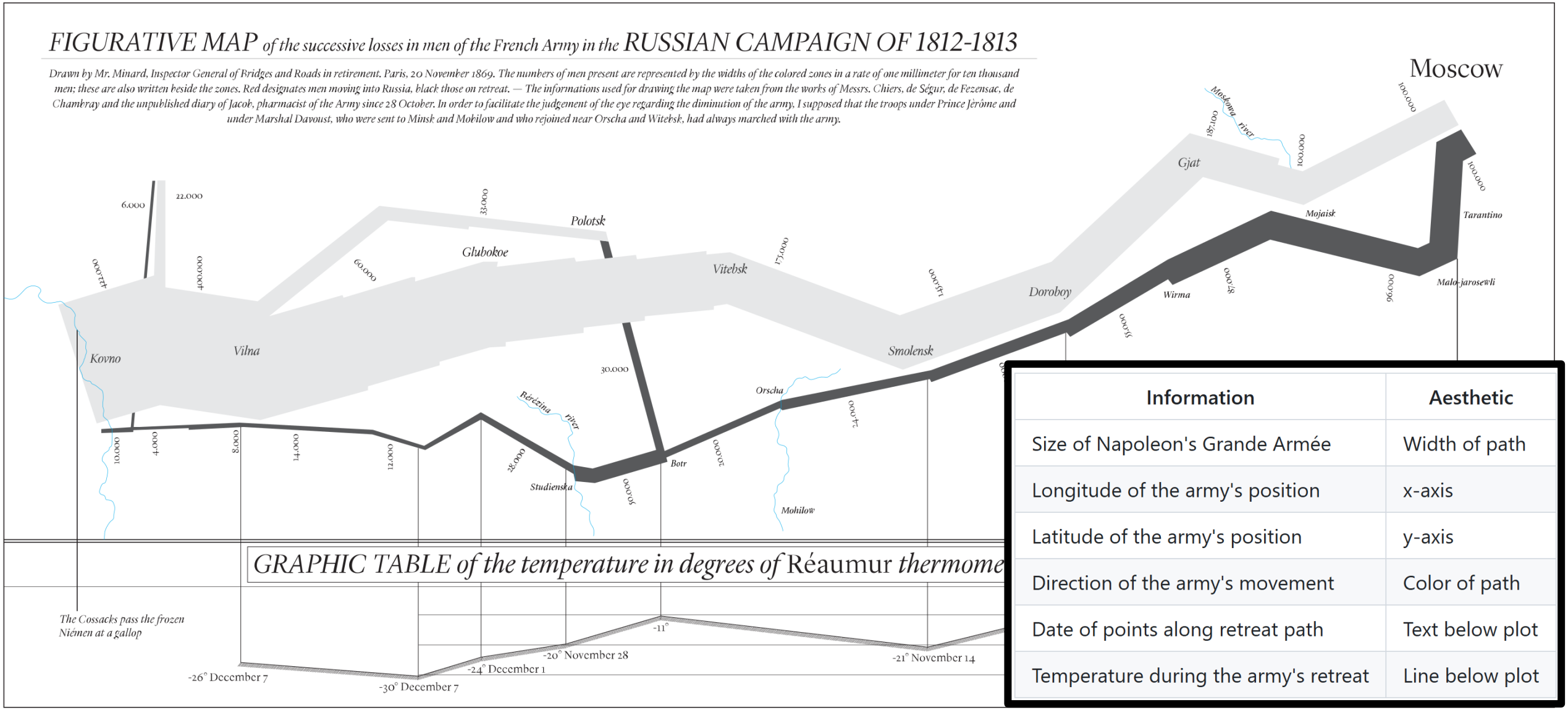
Modern Remake of Charles Minard's map of Napoleon's disastrous Russian campaign of 1812



Source: https://en.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Redrawing_of_Minard's_Napoleon_map.svg

A visualization worth a 1000 words

Modern Remake of Charles Minard's map of Napoleon's disastrous Russian campaign of 1812



Information	Aesthetic
Size of Napoleon's Grande Armée	Width of path
Longitude of the army's position	x-axis
Latitude of the army's position	y-axis
Direction of the army's movement	Color of path
Date of points along retreat path	Text below plot
Temperature during the army's retreat	Line below plot

Source: https://en.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Redrawing_of_Minard's_Napoleon_map.svg

Visual Storytelling

https://www.youtube.com/watch?v=jbkSRLYSojo

Course Organization

Course website on Moodle

https://learn.fh-kiel.de/course/view.php?id=16720

- All course materials
- Exam information
- Announcements

Lecture Date and Location

Time and place

- Monday, 12:45 – 4:00 PM in O15-3.47
- Wednesday, 12:45 – 4:00 PM in O15-3.47

Lecture Period

- 10.11. - 19.12.2025
- 17.11.2025:
  - 12.45 AM – 2.00 PM: *Final Test of the Programming course*
  - 2.15 PM – 4.00 PM: Regular DataViz course

Zoom

- We use Zoom meetings in the lectures to easily share the screen and discuss code examples and solutions from students
- Zoom Meeting Link: see course Moodle page

Literature

The main materials are the ones uploaded to the course Moodle page

Additional literature

- Nussbaumer Knaflic: Storytelling with Data: A Data Visualization Guide for Business Professionals (2015)
- Jonathan Schwabish, Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks, 2021
- Wilke: Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (2019)

Workload

The course will be intense and may require significant time investments

- regular lecture times
- work on exercises
- work on capstone project
- self-study time

Anzahl der SWS	4 SWS
Leistungspunkte	5,0 Leistungspunkte
Präsenzzeit	48 Stunden
Selbststudium	102 Stunden

Examination Registration

You do not need to formally register for course participation, but you do need to register for the examination via qis.fh-kiel.de

Registration period: 10.11. - 23.11.2025 (!!!)

Late submission policy

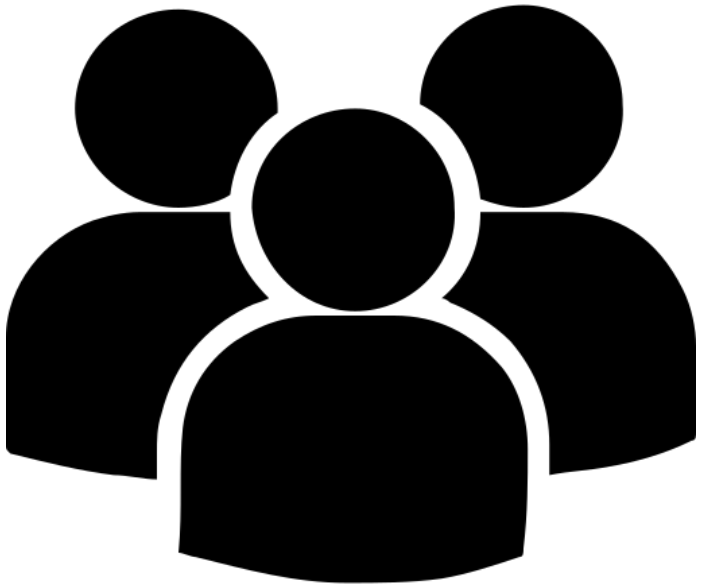
Exercises

- Some of the exercises are for practice only (**practice exercises**), others count towards your grade and need to be submitted (**submission exercises**).
- Submission exercises are provided on Mondays or Wednesdays. Your solutions must be submitted until Sunday 11.59 PM → allows us to review and discuss solutions in the following week
- Late submissions are accepted, but you will lose points.

In case of special circumstances (e.g. illness), exceptions from the general rule are possible — please contact me if such a situation arises.

Teamwork

- Collaboration is allowed at the level of discussing and exchanging ideas and approaches
- But each student must produce and submit their own results.
- Directly copying another student's solution (even with minor edits) is not permitted



Generative AI Policy (I)

The use of Generative AI (e.g., ChatGPT, Gemini, Copilot etc.) in Data Science can be both a blessing and a curse:

- A blessing if you use it to learn more effectively: find and understand errors; explore different solution strategies; accelerate feedback; broaden your horizon.
- A curse if you use it merely to get results without learning: if you copy AI code without understanding it; if you skip over the process of struggling, thinking, and making mistakes; if you treat AI as a shortcut rather than a tool for growth.

Data Science is best learned through trial and error. Making mistakes and fixing them is a central part of the learning journey.

→ The Generative AI Policy is also available in Moodle

Generative AI Policy (II)

Rules on the Usage of Generative AI

1. You are allowed to use generative AI, but only in a supportive manner (e.g., brainstorming, understanding of concept and methods, debugging, etc.). The main parts of your submissions (code, reasoning, reflection) must be written on your own. And all submitted work must reflect your own understanding and effort.
2. I may schedule individual assessment sessions (in-person or Zoom) where you walk me through your solutions for any of the submissions. You must be able to explain your code and thought process, and why you proceeded in this way. If you cannot demonstrate ownership, your work may not be credited.

→ The Generative AI Policy is also available in Moodle

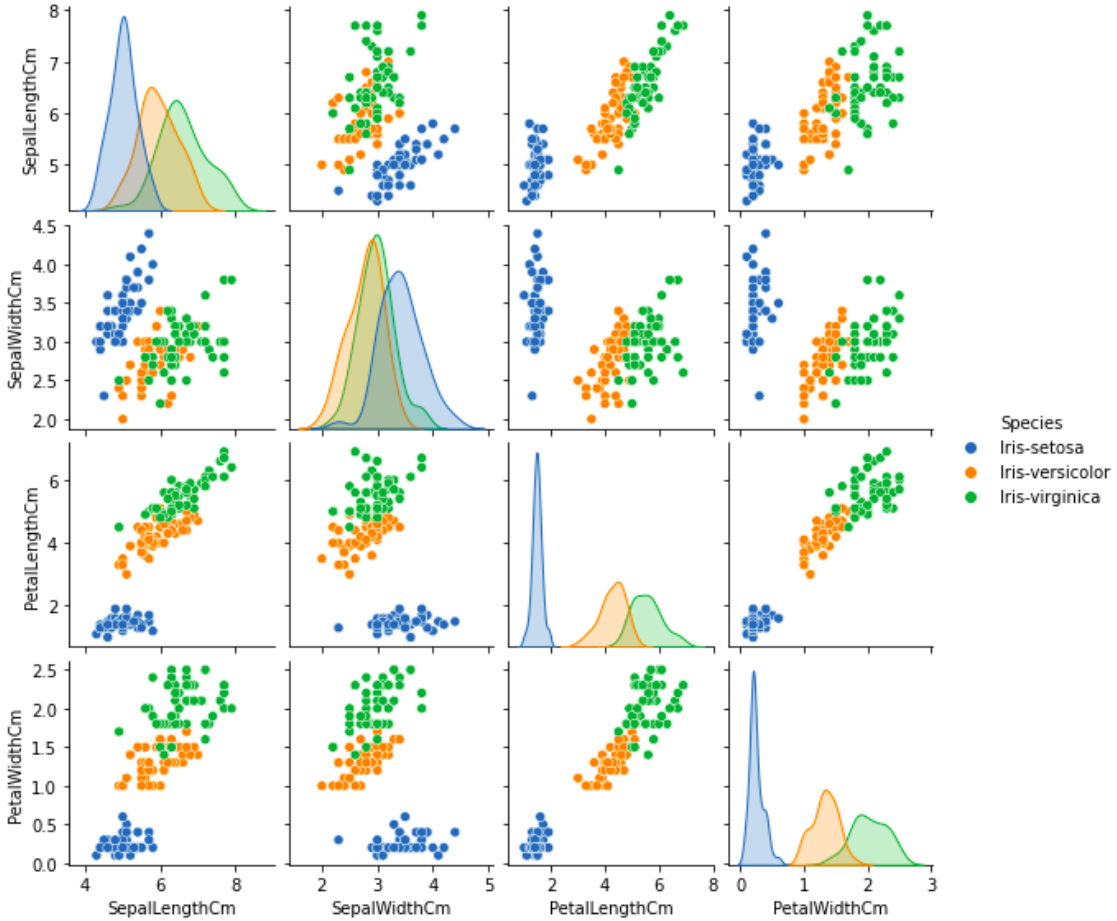
Portfolio examination: task overview

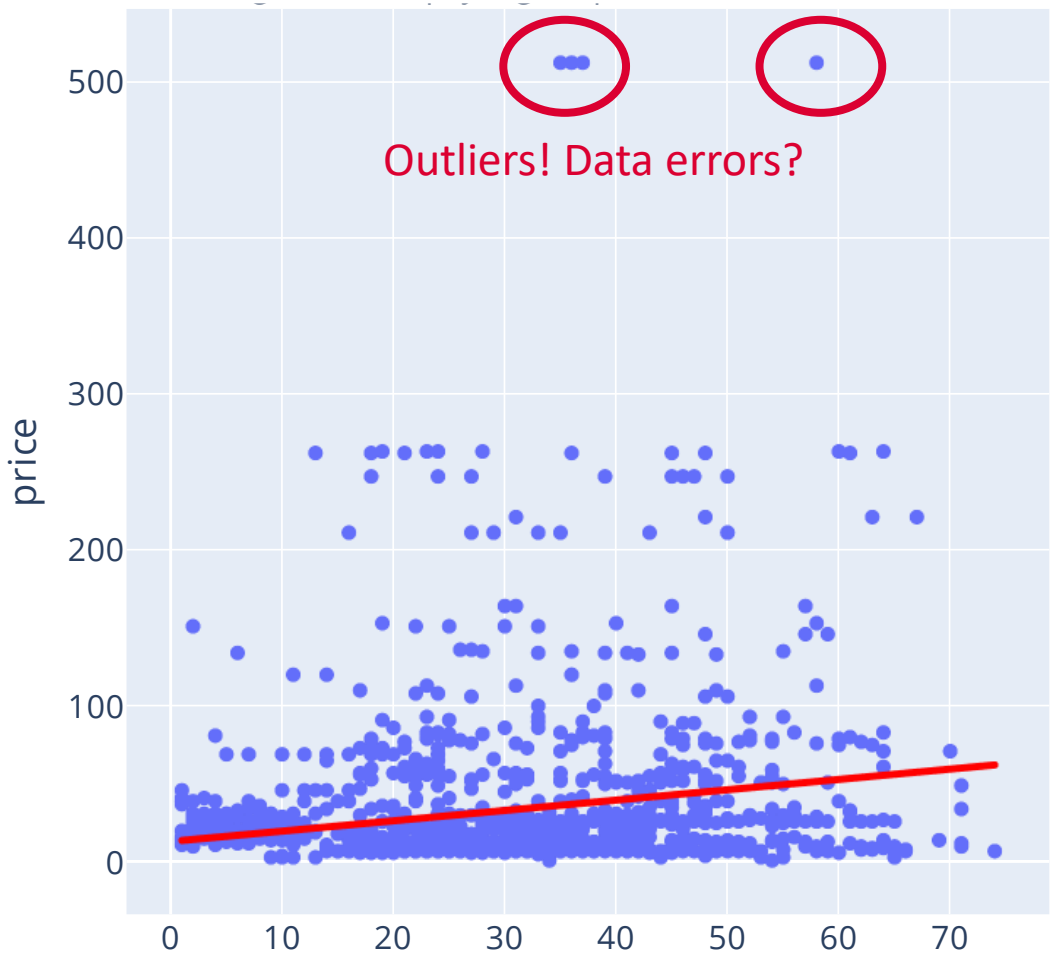
Task	Weight	(Due) date	Teamwork allowed?	GenAI allowed?
Exercises	30%	Sundays, 11.59 PM	yes, but individual submissions	supportive mode only*
Miniproject - Interactive Visualization	20%	18.01.2026, 11.59 PM	yes, but individual submission	supportive mode only*
Capstone Project - Visual Storytelling	50%	18.01.2026, 11.59 PM	no	supportive mode only*

I highly value active participation and contributions to the lecture. Outstanding participation during the semester may be honored with bonus points


What this course is (and is not) about

Visual Data Exploration?





Interesting and important, but not the main course target!

Explanation and Visual Storytelling

Exploration

Making sense of data *for yourself*

- Find patterns or problems in data
- Iterative, exprimental, messy

Explanation

Making sense of data *for others*

- Communicate patterns, insights and implications effectively
- Deliberate design choices
- „Visual Storytelling“

Prerequisite

Main goal of this lecture!

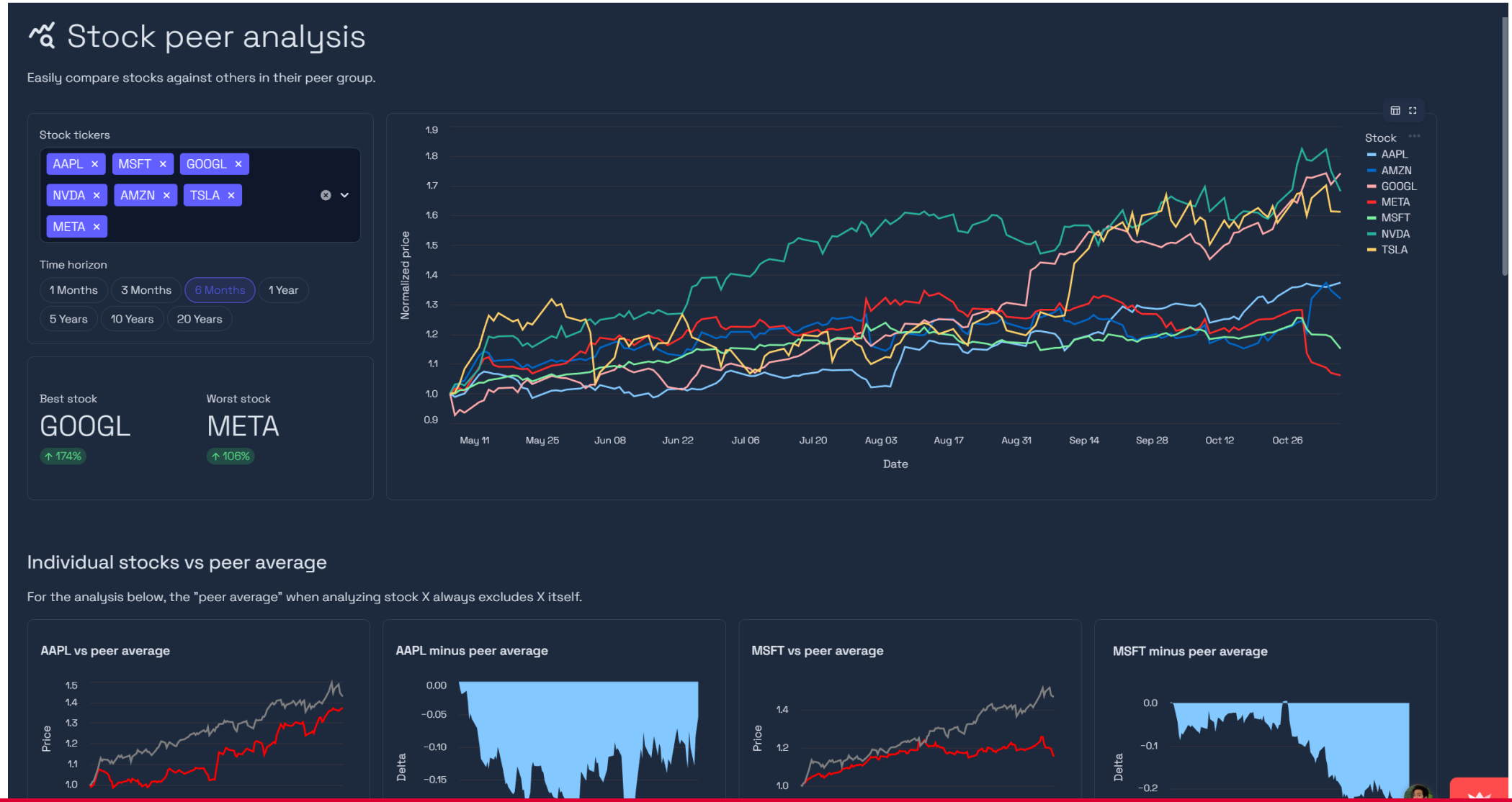
Exploration vs. Explanation

"We might have to **open 100 oysters** (test 100 different hypotheses or look at the data in 100 different ways) to find perhaps **two pearls**.

When we're at the point of **communicating our analysis to our audience**, we really want to be in the explanatory space, meaning **you have a specific thing you want to explain**, a specific story you want to tell—probably about those two pearls."

Nussbaumer Knaflic, „Storytelling with Data: A Data Visualization Guide for Business Professionals (2015)

Interactive Dashbords



Power BI

+ableau

Streamlit

Towards the end of the lecture

Lecture Roadmap

Data Domains

Comparing Categories | Relationships | Geospatial | Time |
Part-to-whole | Distributions | Uncertainty | …

Storytelling

Perception + Visualization Design

Python + Tools

Interactive Visualization


What makes a good data visualization?

Graphical excellence is the **well-designed** presentation of **interesting data**—a matter of substance, of statistics, and of design … [It] consists of complex ideas **communicated with clarity**, precision, and efficiency. … And graphical excellence requires telling the **truth** about the data.“

Edward Tufte, The Visual Display of Quantitative Information, p. 51

Criteria for good visualizations

Criteria	What it means
Truthful	Reliable data source Faithful representation Transparent and reproducible process

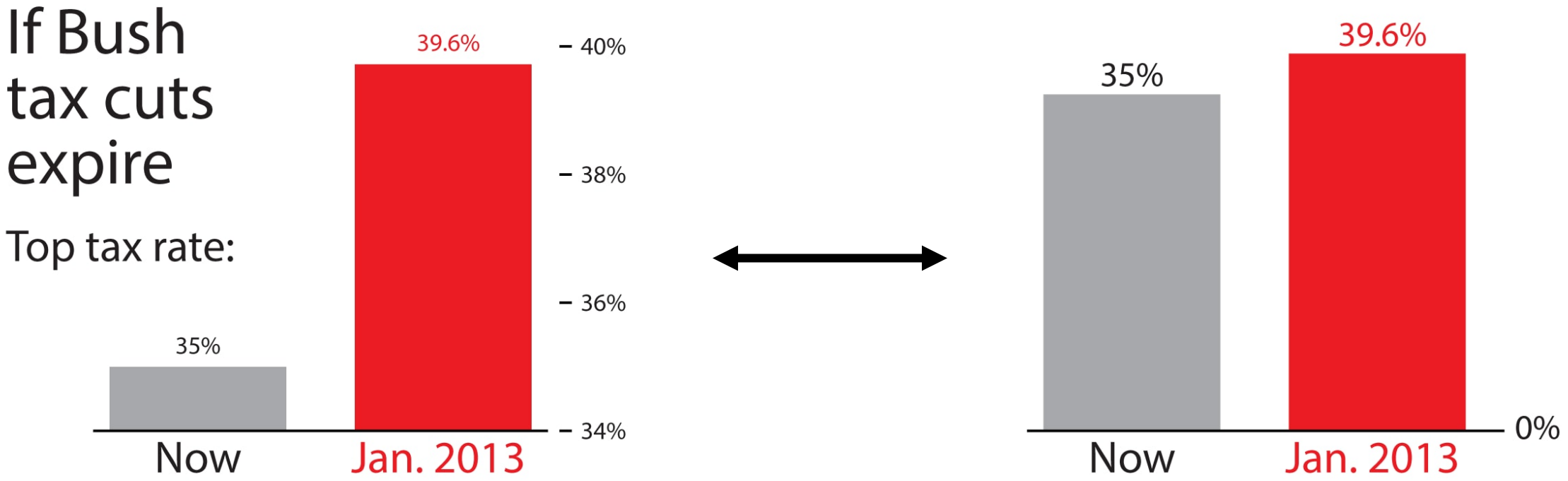
Criteria for good visualizations

Criteria	What it means
Truthful	Reliable data source Faithful representation Transparent and reproducible process
Useful	Insightful and (ideally) actionable Effectively communicated Understood accurately by audience and without efforts

Criteria for good visualizations

Criteria	What it means
Truthful	Reliable data source Faithful representation Transparent and reproducible process
Useful	Insightful and (ideally) actionable Effectively communicated Understood accurately by audience and without efforts
Beautiful	Draw attention Engage audience Make it memorable

Truthful?

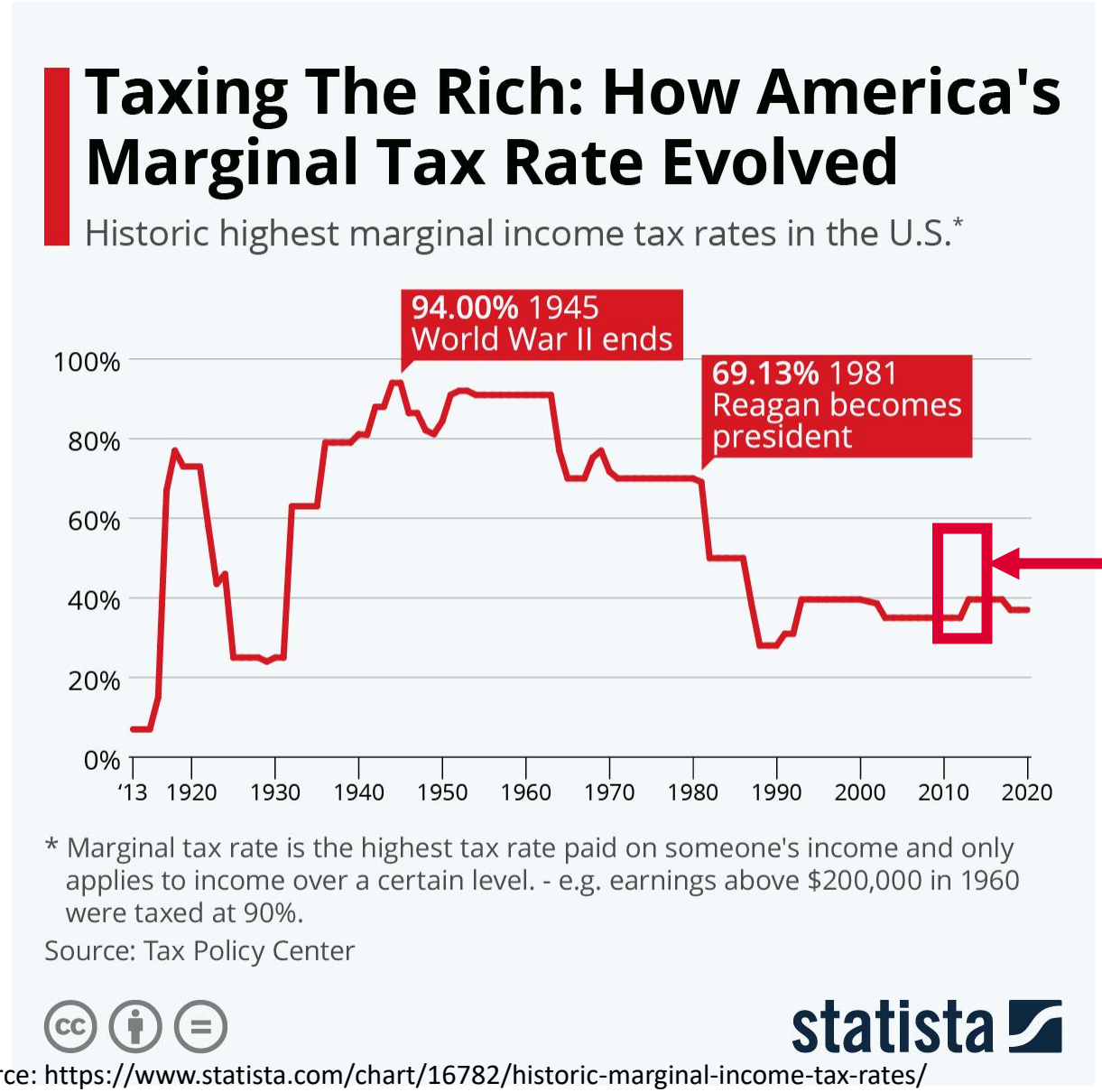


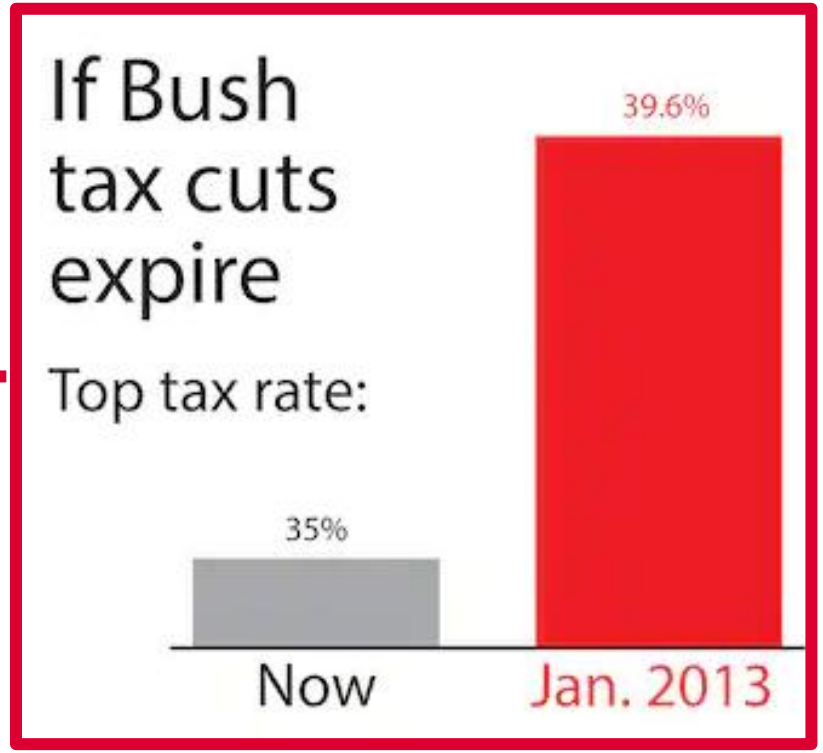
Adaptation of a Chart from Fox Business in 2011

specific year -> lets zoom out

Source: https://www.washingtonpost.com/business/2019/10/14/youve-been-reading-charts-wrong-heres-how-pro-does-it/

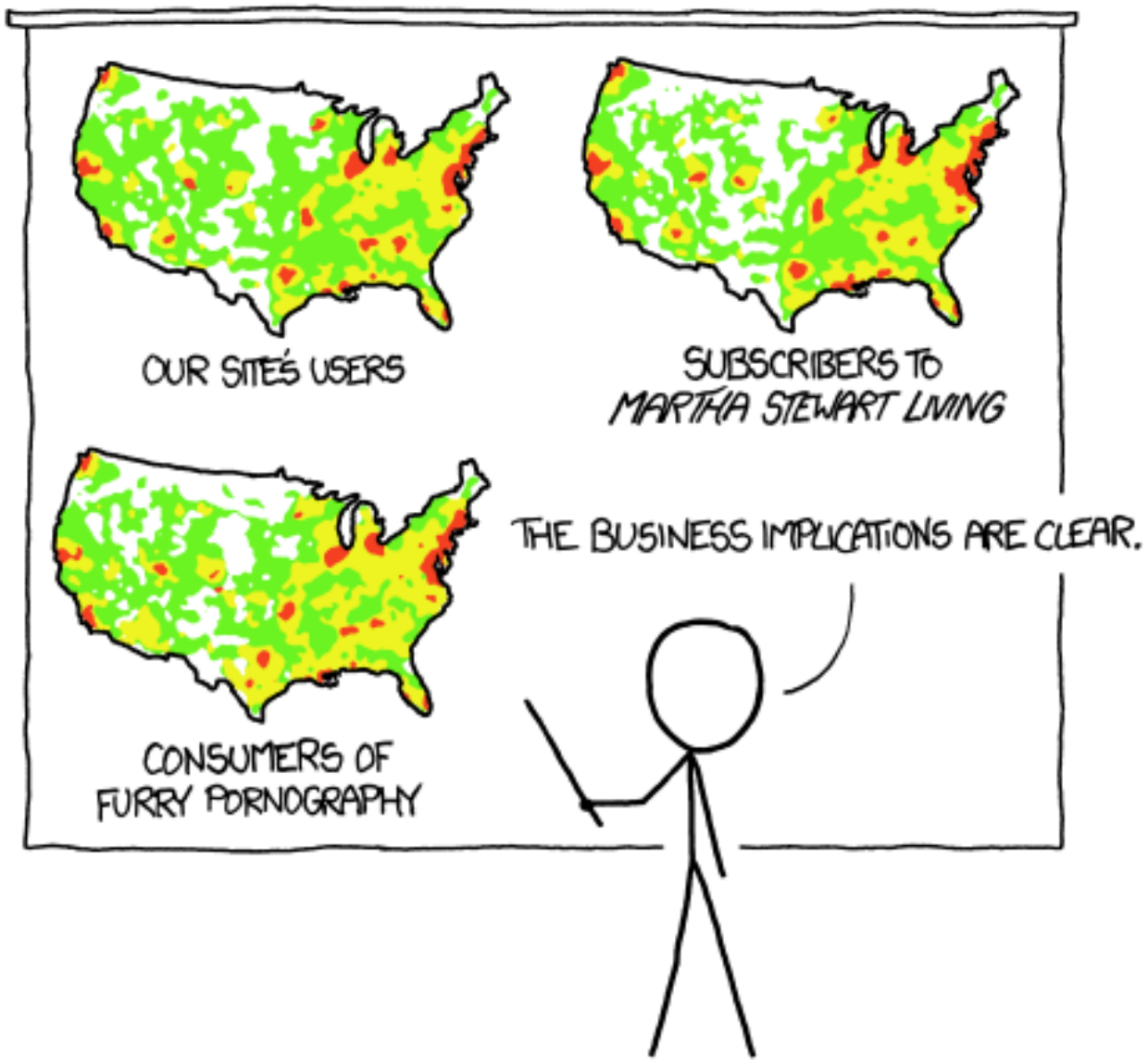
Truthful?





Source: https://www.statista.com/chart/16782/historic-marginal-income-tax-rates/

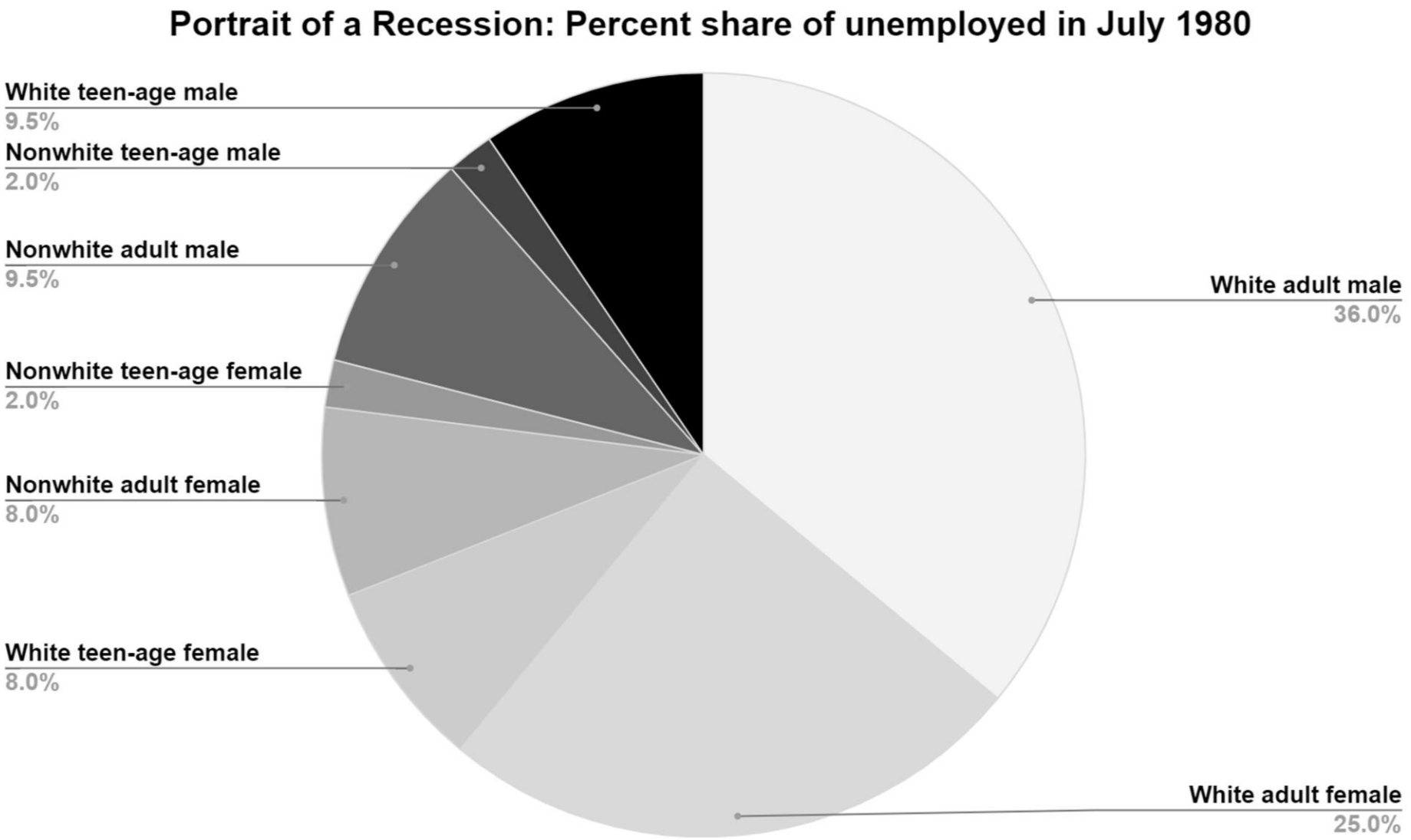
Useful?



thats basically just a map of where alot of people live

Sources: https://xkcd.com/1138/

Memorable?


Portrait of a Recession: Percent share of unemployed in July 1980
White teen-age male
9.5%
Nonwhite teen-age male
2.0%
Nonwhite adult male
9.5%
Nonwhite teen-age female
2.0%
Nonwhite adult female
8.0%
White teen-age female
8.0%
White adult male
36.0%
White adult female
25.0%

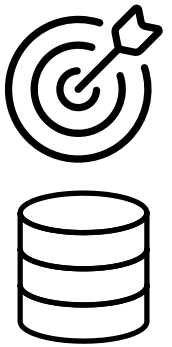

An exemplatory visual storytelling process

Visual Storytelling Process

Stage	Description
1. Context	• Clarify purpose, topic/questions, audience, medium

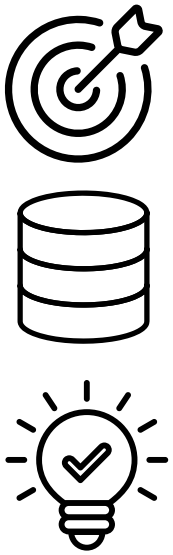
Visual Storytelling Process

Stage	Description
1. Context	• Clarify purpose, topic/questions, audience, medium
2. Data	• Identify and prepare data • Explore for patterns and insights



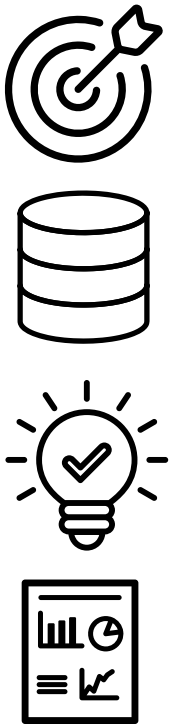
Visual Storytelling Process

Stage	Description
1. Context	• Clarify purpose, topic/questions, audience, medium
2. Data	• Identify and prepare data • Explore for patterns and insights
3. Story	• Decide upon the main message • Narrative structure



Visual Storytelling Process

Stage	Description
1. Context	• Clarify purpose, topic/questions, audience, medium
2. Data	• Identify and prepare data • Explore for patterns and insights
3. Story	• Decide upon the main message • Narrative structure
4. Design	• Choose visual forms and refine • Assemble the final output



1. Context

Our use case: we want to gain insights into stock price movements of DAX companies

- **Audience**: general public
- **Questions**: no specific ones, just interest for salient patterns
- **Medium**: static, single visualization

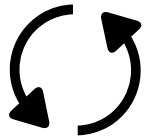
2. Data

Identify suitable (trustworthy) data sources and prepare data for analysis

- **Stock prices**: Yahoo Finance, Deutsche Börse, etc.
- Potentially enrich using further data sources
  - Macroeconomic indicators: GDP, employment, exchange rates
  - Company fundamentals: profit, revenues, …
  - News sentiment

2. Data (aquisition and preparation)

Yahoo Finance API
Mapping of company names
Reshape data
Normalize data

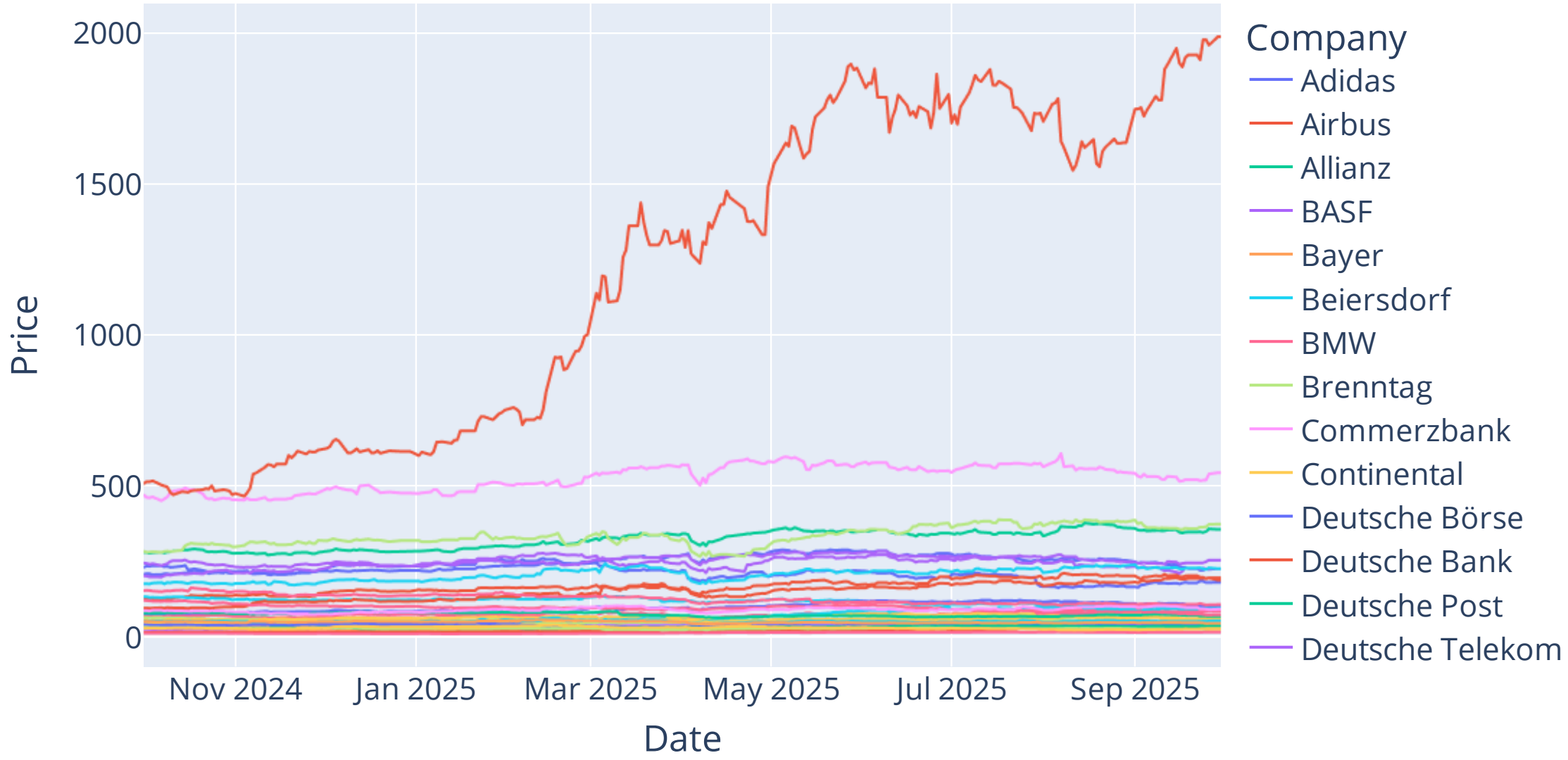


First exploratory plots

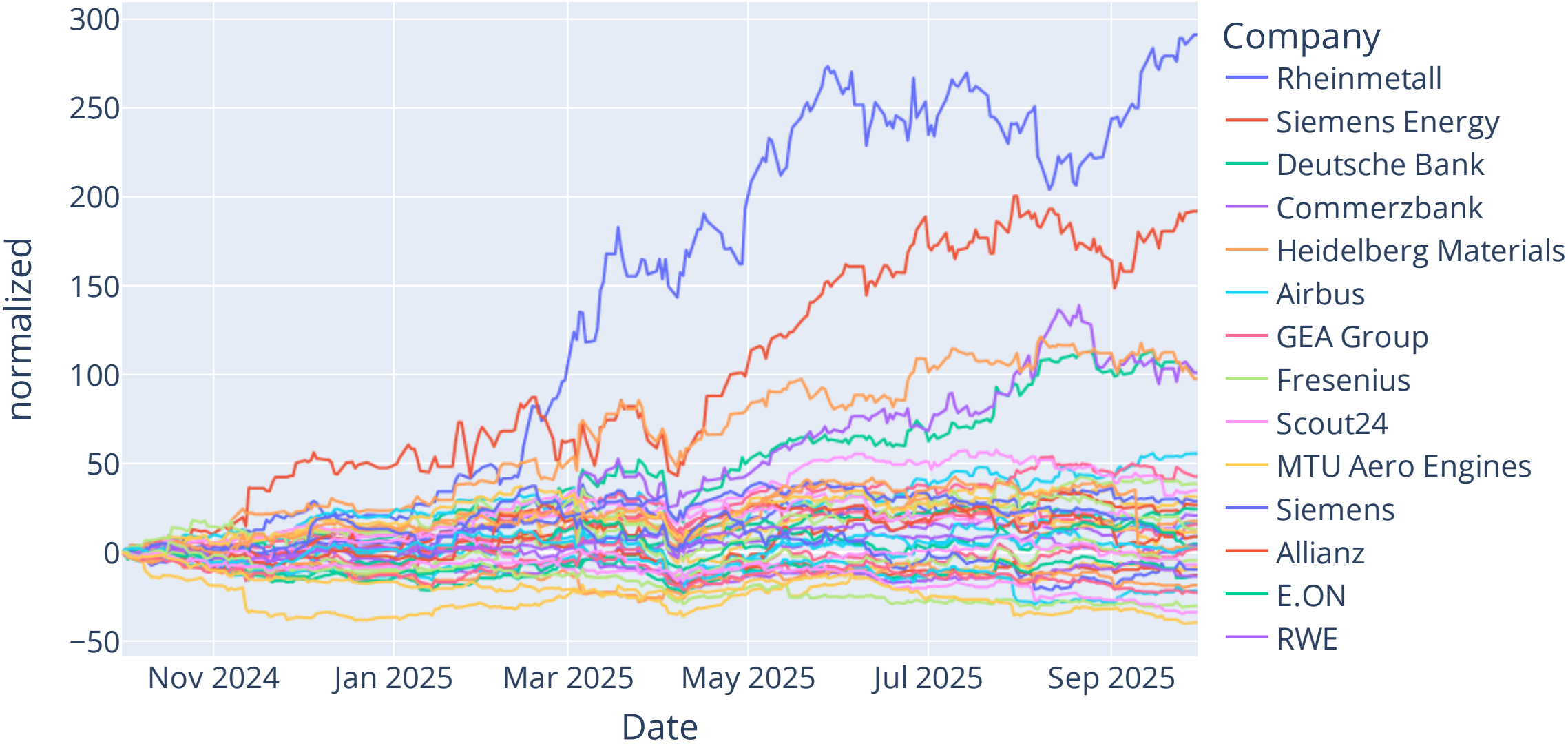
Ticker	ADS.DE	AIR.DE	ALV.DE	BAS.DE	BAYN.DE	BEI.DE	BMW.DE	BN
Date								
2024-10-01	230.67	125.93	281.36	44.53	30.30	133.42	73.76	
2024-10-02	235.03	124.54	279.73	44.85	30.25	133.12	73.65	
2024-10-03	234.34	123.74	277.72	45.07	29.72	130.79	72.51	
2024-10-04	236.02	124.07	279.73	45.79	29.66	131.19	73.78	
2024-10-07	236.12	124.66	280.21	45.88	29.63	130.20	73.85	
...	...	...	...	...	...	...	...	
2025-09-24	184.75	193.78	346.00	42.44	27.61	88.38	84.82	
2025-09-25	180.30	193.92	349.30	42.09	27.25	87.30	84.82	

	date	ticker	price	company	normalized
0	2024-10-01	ADS.DE	230.67	Adidas	0.00
1	2024-10-02	ADS.DE	235.03	Adidas	1.89
2	2024-10-03	ADS.DE	234.34	Adidas	1.59
3	2024-10-04	ADS.DE	236.02	Adidas	2.32
4	2024-10-07	ADS.DE	236.12	Adidas	2.36
...	...	...	...	...	...
10115	2025-09-24	ZAL.DE	27.43	Zalando	-5.41
10116	2025-09-25	ZAL.DE	26.95	Zalando	-7.07
10117	2025-09-26	ZAL.DE	26.28	Zalando	-9.38
10118	2025-09-29	ZAL.DE	26.22	Zalando	-9.59
10119	2025-09-30	ZAL.DE	26.22	Zalando	-9.59

2. Data (exploration)



2. Data (exploration)



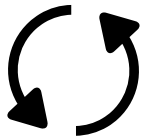
3. Story

Collecting potential stories

- Top and bottom performers
- Arms and energy companies benefit from war in Ukraine
- Shock to the DAX in April 2024
- Banks show co-movement

Possible directions to dig deeper

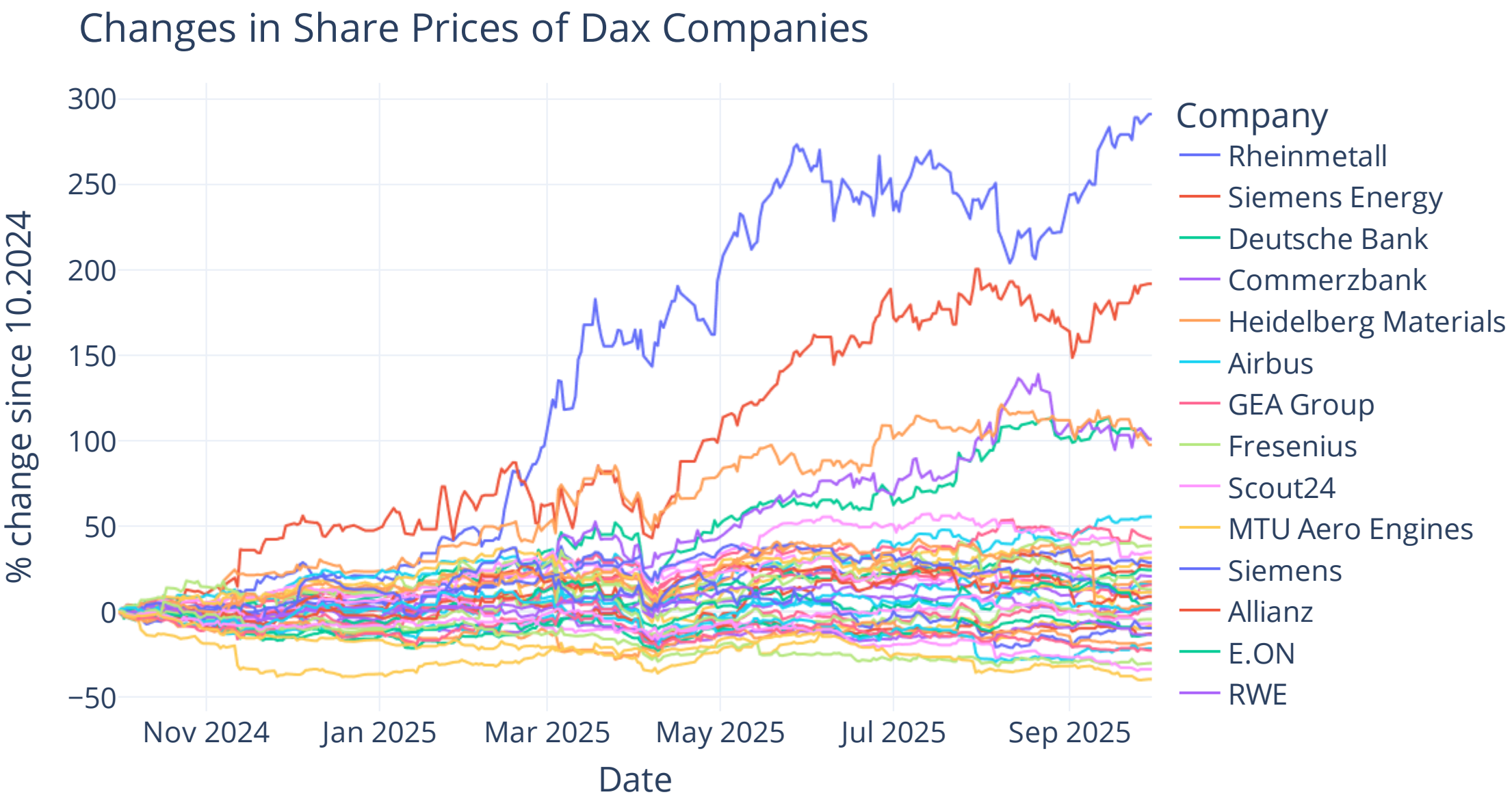
- Can we classify companies (banks, arms, energy, …)?
- What happened in April 2024?
- Longer time horizon?
- Compare German with US companies?



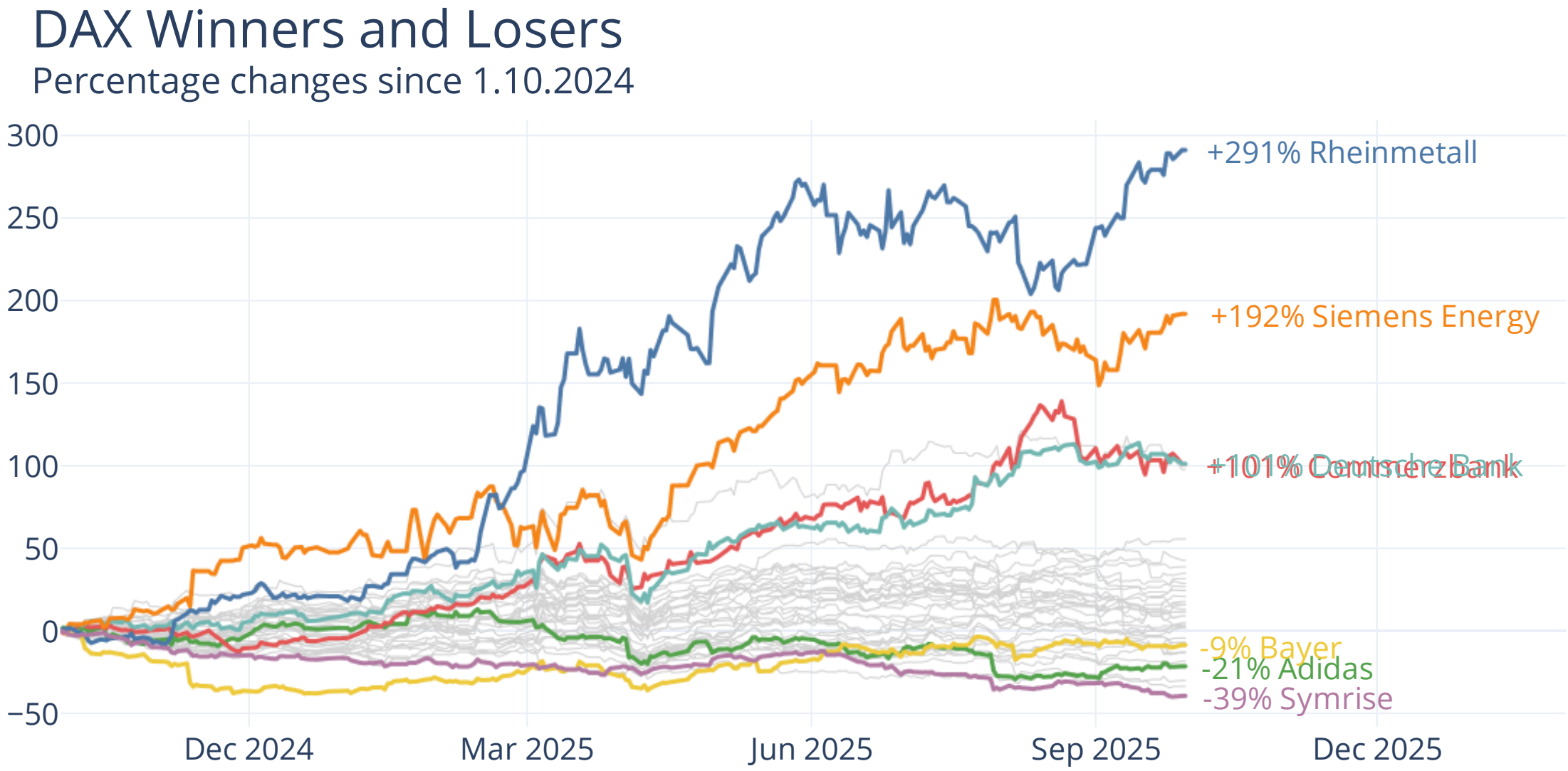
News investigation
Data acquisition
Data exploration

Choose key message: Top and bottom performers + (April shock)

4. Design: de-clutter

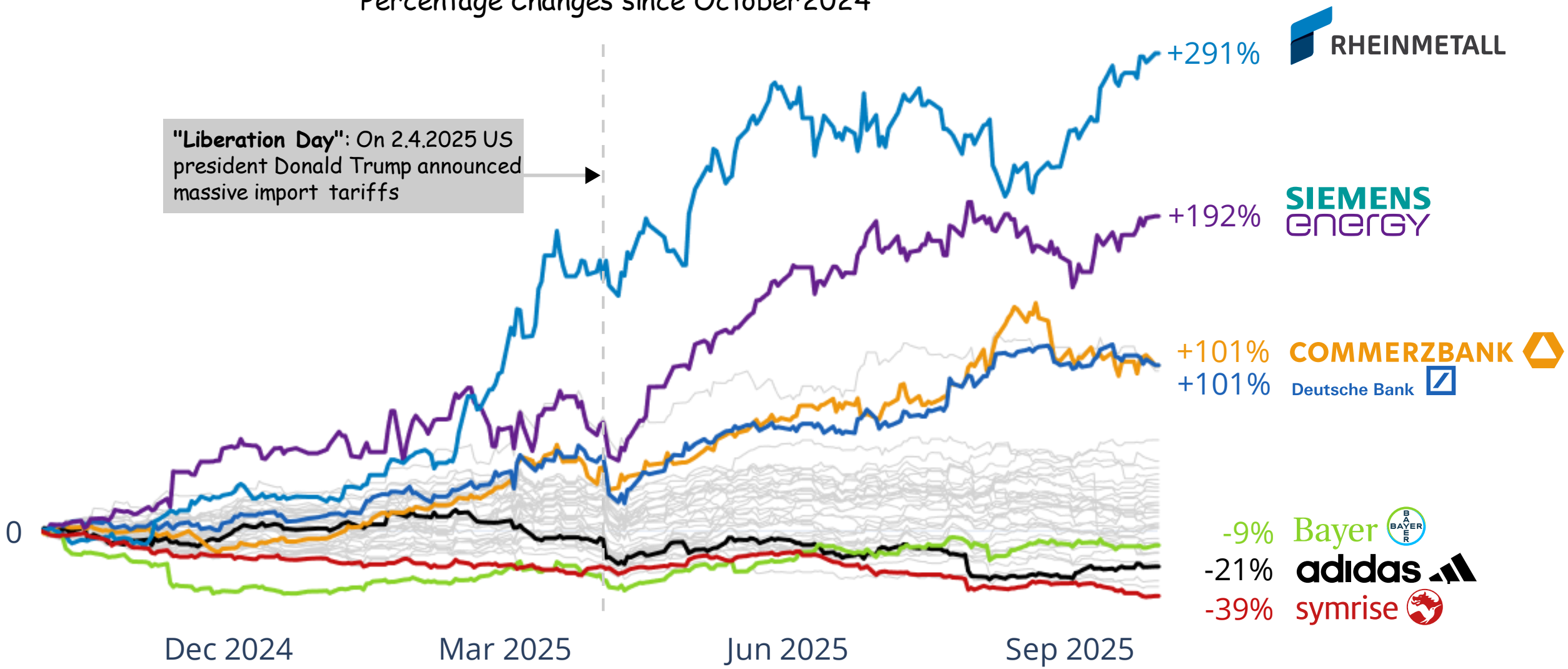


4. Design: highlight data



DAX Winners and Losers

Percentage Changes since October 2024



4. Design

- De-clutter
- Color highlighting
- Labels and reference lines
- Explainers with key messages

Most refinements are carried out with **Python Plotly**

Export as vector graphic

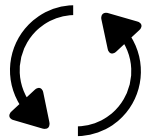
- Overlapping labels
- Company Logos
- Final touches

Manual finetuning with **Inkscape** (or Adobe Illustrator)

4. Design

Good design works at two levels:

- **Micro-design***:* individual chart
- **Macro-design***:* composition of full story



Criteria for both:

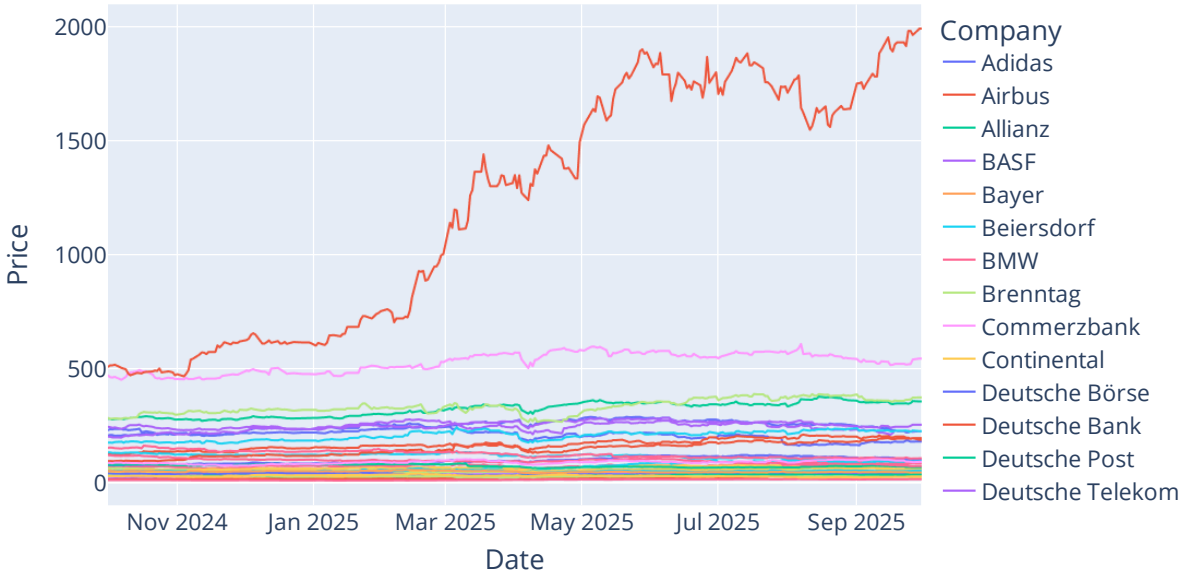
- Design supports story
- Logical flow of information
- Consistent choices
- Visually attractive

Extend the story:

- Zoom into Rheinmetall
- Compare sectors
- Study effects of Trump
- Compare foreign companies
- Longer time horizon
- …

Recap

Before



After