

DATA STORY CASE STUDIES

Examining, presenting, and critiquing example data analyses from the web.

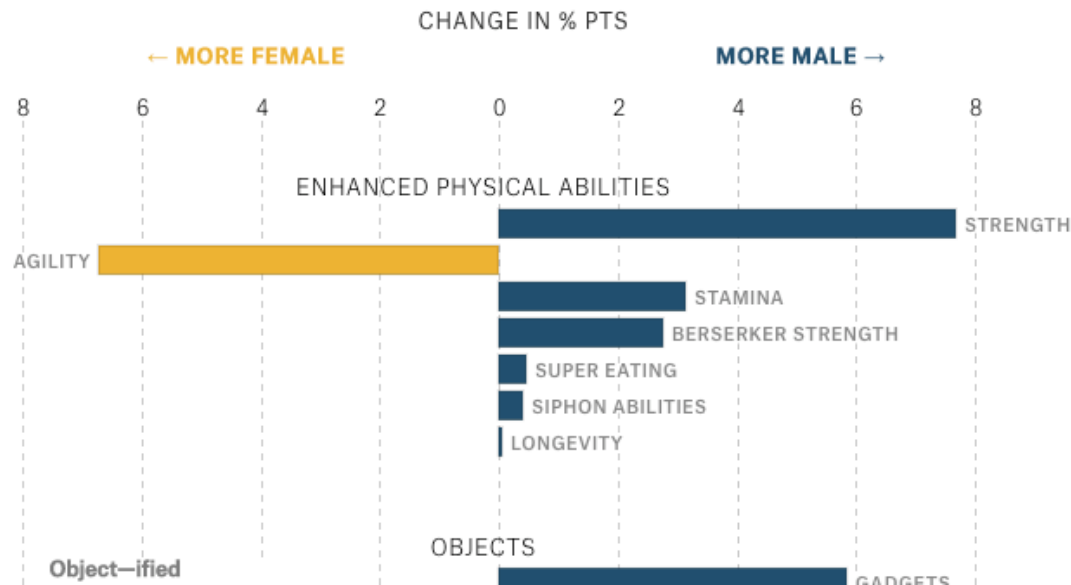


Analyzing the Gender Representation of 34,476 Comic Book Characters (The Pudding)

Strong men; thoughtful women

Three out of five comic book characters have at least one superpower, regardless of gender. When we categorize these powers, we find that there are some clear gender imbalances.

The difference in male and female powers



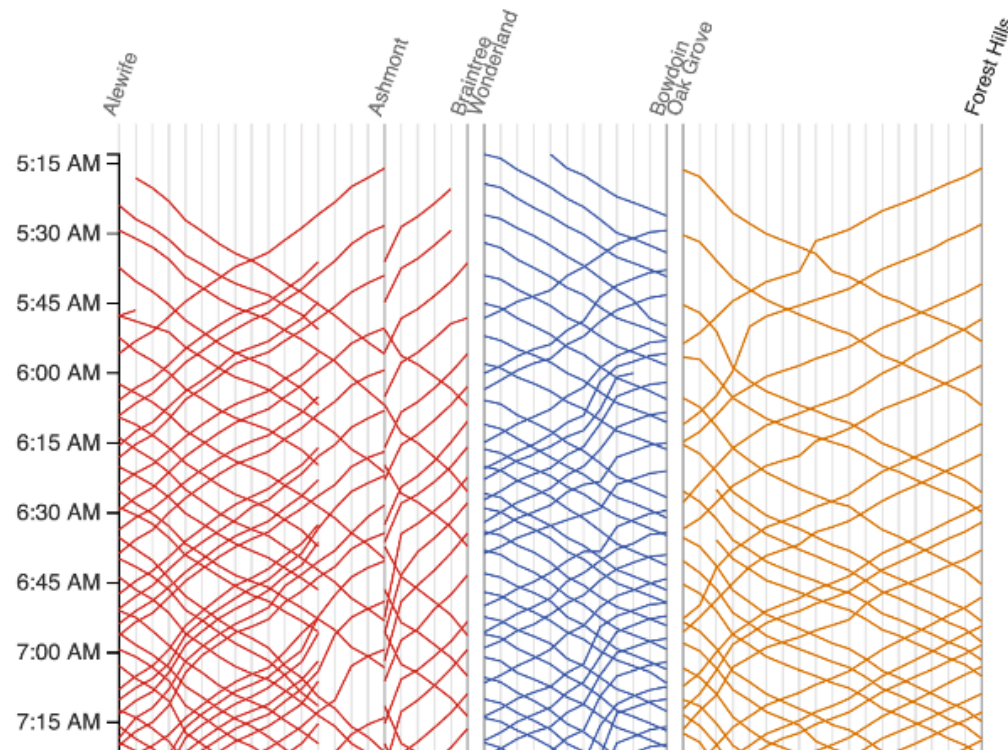
WHY DO THIS?

A chance to see, critique, and learn from a bunch of really nice analysis examples of data journalism and professional data science.

The Marey Chart: Visualising time, location and direction

Our next step was to incorporate multiple dimensions into the exploratory analysis.

We were inspired by the Marey Chart, which was featured in Edward Tufte's vaunted 1983 classic The Visual Display of Quantitative Information. More recently, it was used by Mike Barry and Brian Card for their extensive visualisation project on the Boston subway system:



Screenshot 2: Taken from <http://mbtaviz.github.io/>

[How the Circle Line rogue train was caught with data \(data.gov.sg\)](http://data.gov.sg)

WHY DO THIS?

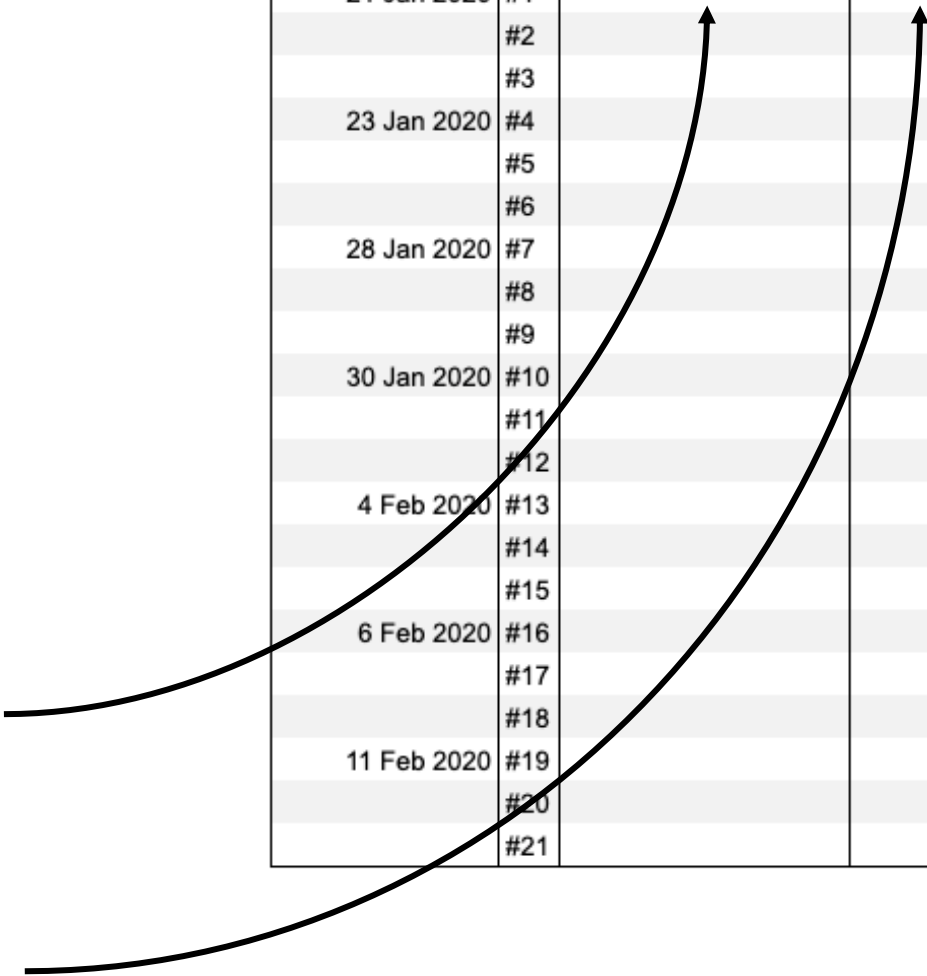
We will be asking you to create stories kind of like these the main deliverable for your final project.

WHAT TO DO?

3 Presentations each course meeting starting in **Week #3**

- **Sign up for a date** on the sign-up sheet.
- **Pick a story.** We have some suggestions, but there are lots more out there.
(Check with us if you want to propose a different one.)

Class		Presenter Name	Data Story
21 Jan 2020	#1		
	#2		
	#3		
23 Jan 2020	#4		
	#5		
	#6		
28 Jan 2020	#7		
	#8		
	#9		
30 Jan 2020	#10		
	#11		
	#12		
4 Feb 2020	#13		
	#14		
	#15		
6 Feb 2020	#16		
	#17		
	#18		
11 Feb 2020	#19		
	#20		
	#21		



IF YOU TAKE ONE OF
THESE, FLAG IT.

IF YOU WANT TO PICK
A DIFFERENT ONE,
ASK US AND ADD IT
TO THE LIST.

(THERE ARE LOTS OF GREAT SOURCES!)

● **TheUpshot**

FiveThirtyEight

ThePudding

Etc.

Example Data Stories

Taken

Essay

[How we caught the circle line rogue train with data - Datagov Singapore.](#)
[Call me Adele. How I turned my sorrows into a data set - Wipawe Sirikolkarn.](#)
[The Rythm of Food - Moritz Stefaner](#)
[New York's Noisiest Neighborhoods - Ben Wellington](#)
[How Machines Destroy \(And Create!\) Jobs, In 4 Graphs - Quoc Trung Bui](#)
[How We Analyzed the COMPAS Recidivism Algorithm - Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin.](#)
[The NYPD Was Systematically Ticketing Legally Parked Cars for Millions of Dollars a Year-Open Data Just Put an End to It - I Quant NY.](#)
[You'll Never Guess the Cleanest Fast Food Joint in NYC - I Quant NY.](#)
[Half of Manhattan is Within 4 Blocks of a Starbucks - I Quant NY.](#)
[How Software in Half of NYC Cabs Generates \\$5.2 Million a Year in Extra Tips - I Quant NY.](#)
[How to Fix NYC's No-Cabs-At-4PM Problem - I Quant NY.](#)
[Success: How NYC Open Data and Reddit Saved New Yorkers Over \\$55,000 a Year - I Quant NY.](#)
[Why Are so Many Babies Born around 8:00 A.M.? - Zan Armstrong](#)
[Hollywood's Gender Divide And Its Effect On Films - Polygraph](#)
[Analyzing the Gender Representation of 34,476 Comic Book Characters - Pudding](#)
[Film Dialogue from 2,000 screenplays, Broken Down by Gender and Age - Polygraph](#)
[My Google Search History - visualized - Lisa Charlotte Rost](#)
[A Data Driven Exploration of Kung Fu Films - Jim Vallandingham](#)
[Quantifying and Visualizing "Deep Work" - Enrico Bertini](#)
[Baby Names Exploratory analysis - Michal Hron](#)
[A Tale of Twenty-Two Million Citi Bike Rides: Analyzing the NYC Bike Share System - Todd W. Schneider.](#)
[Analyzing 1.1 Billion NYC Taxi and Uber Trips, with a Vengeance - Todd W. Schneider.](#)
[Improving the efficiency of food inspections in Chicago - city of Chicago](#)
[#Dataviz tutorial: Mapping San Francisco home prices using R - Ken Steif](#)
[LEGO color themes as topic models - Nathaneal Aff](#)
[The Simpsons by the Data - Todd W. Schneider](#)
[Analyzing 40 Years Of Star Wars Box Office Success And The Future Of Cinema - Stephen Tracy](#)
[The Most Popular High School Plays And Musicals - Elissa Nadworny](#)
[Bussed out: How America moves its homeless - The Guardian](#)
[The Trump Effect \(Energy + Environment\)](#)
[Here's every total solar eclipse happening in your lifetime. Is this year your best chance? - Denise Lu](#)
[How do you draw a circle? We analyzed 100,000 drawings to show how culture shapes our instincts - Thu-Huong Ha, Nikhil Sonnad](#)
[Cracking the Mystery of Egg Shape - Sarah Crespi, Jia You](#)
[An Interactive Visualization of Every Line in Hamilton - Shirley Wu](#)
[Terriers Were Once The Greatest Dogs In The World - Oliver Roeder](#)
[The Suno Matchup Centuries In The Making - Benjamin Morris](#)

DATA STORY CASE STUDY PRESENTATIONS

In A 10-Minute Presentation:

- Show us the piece.
- Explain **Who** created it? What **data**? What **tools**?)
- Explain the overall **message** and **takeaways**.
- Highlight its **strengths**
 - Impressive data collection, analysis methods, visualizations, etc.
 - Interesting or exciting observations
- Critique its **weaknesses**
 - Can you identify any problems with the data, analysis, visualizations, etc.?
 - What could they do better?

QUESTIONS?