silk data

machine learning
excellence delivered

How visualization improves AI interaction with workforce

silk data

FORUM. DIGITAL AI

AI understanding…

silk data

… is not necessary

- Nobody knows how natural intelligence works
- In many cases understanding / interpretation is redundant
  - OCR
  - Speech recognition
  - Image classification
  - ~~Self-driving cars~~ ???
  - Mass-marketing
- Caveat: reasonable error level

BUT: Driving is social

Source: https://techcrunch.com/2016/08/30/drive-ai-uses-deep-learning-to-teach-self-driving-cars-and-to-give-them-a-voice/

AI understanding…

silk data

… is definitely necessary

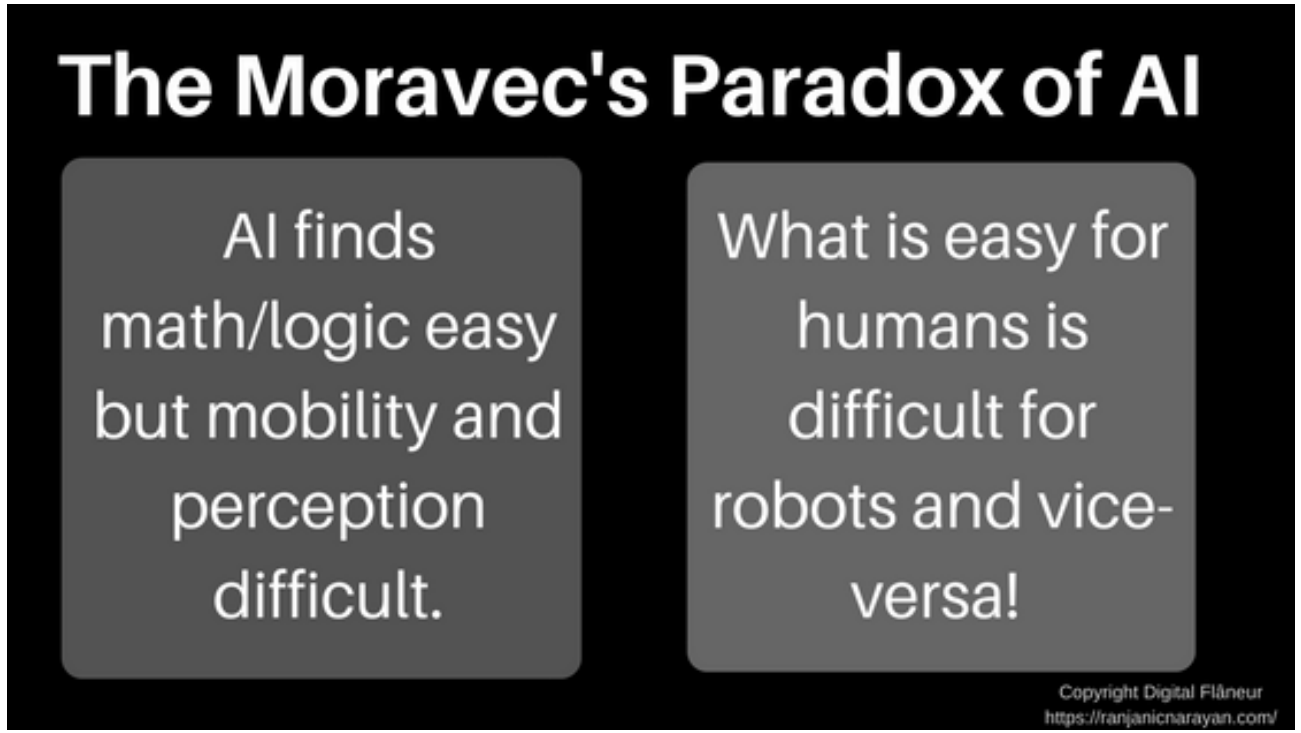
- Social-based tasks
  - Human interaction
  - Business-critical decisions
  - Regulated areas (banks, health)

“AI paradox”

Early 1980s:

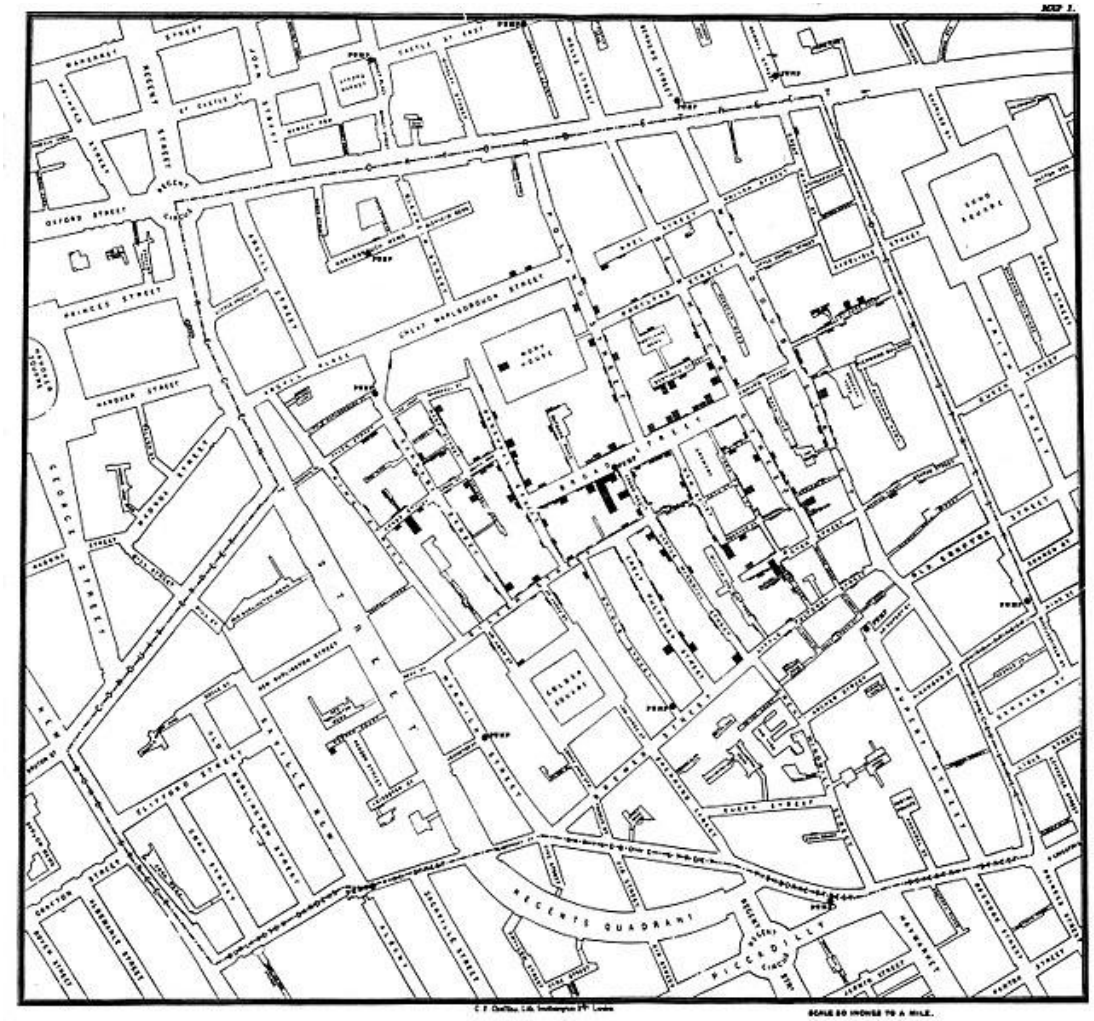
“it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility.”

https://en.wikipedia.org/wiki/Moravec%27s_paradox



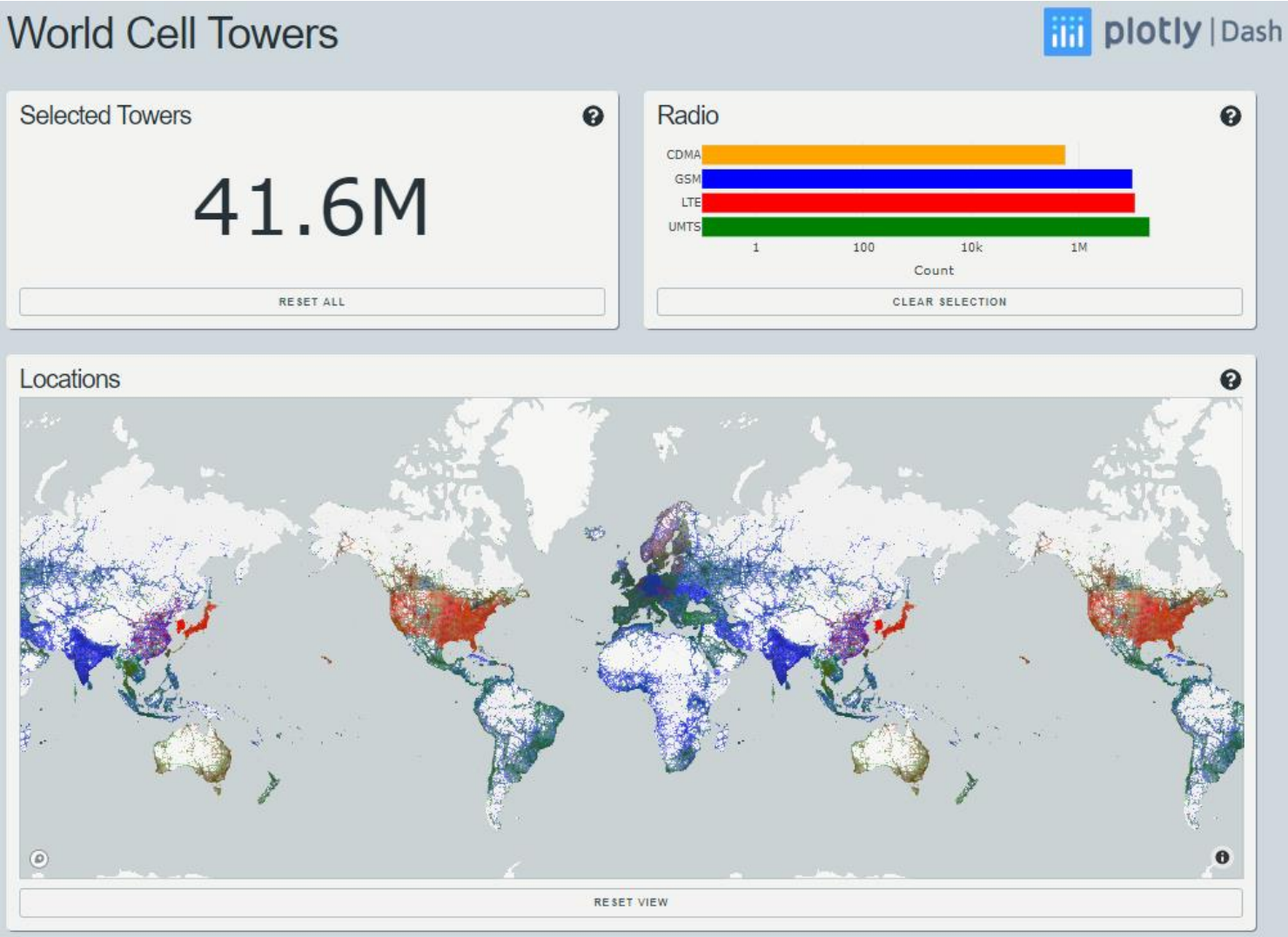
Source: https://ranjanicnarayan.com/2018/01/28/the-ai-paradox/

Snow’s cholera map



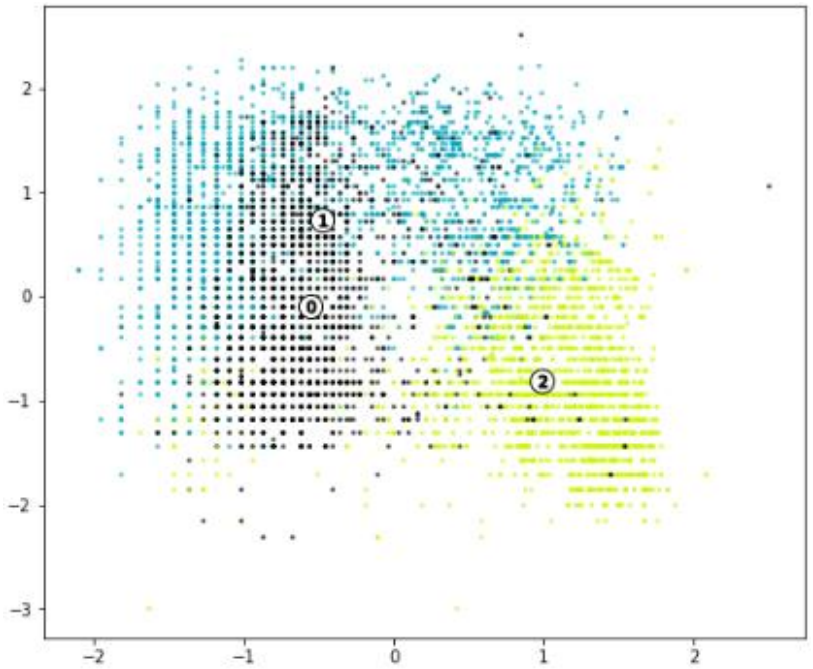
Source: https://en.wikipedia.org/wiki/GIS_and_public_health

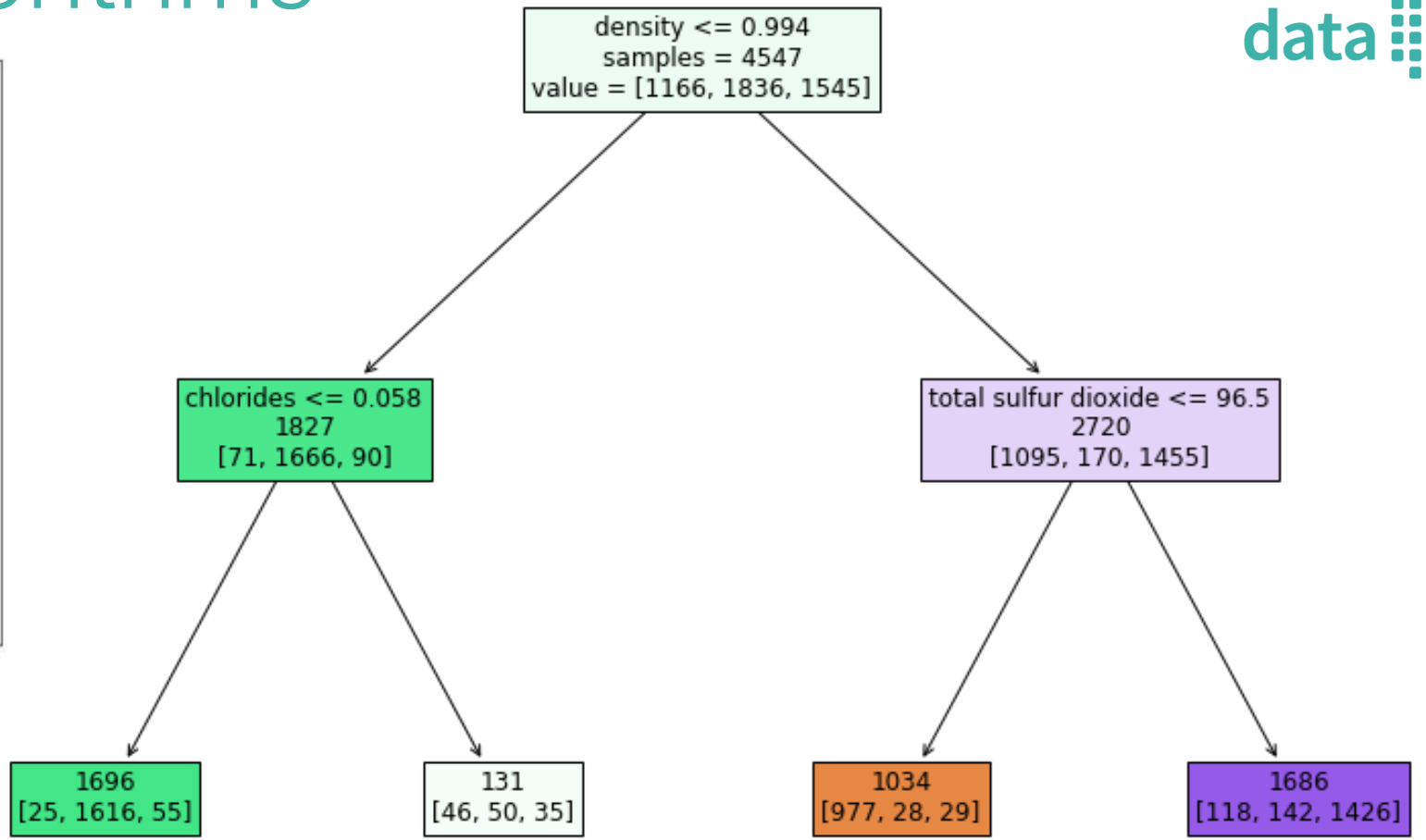
Dashboards



Source: https://dash-dask-py.plotly.host/dash-world-cell-towers/

‘Explainer’ algorithms





Wine quality dataset

fixed acidity, volatile acidity, citric acid, **residual sugar**, chlorides, free sulfur dioxide, **total sulfur dioxide**, **density**, pH, sulphates, alcohol, quality

Visualization in NLP: similarity

Animal Farm	Snowball's Chance
Old Major the old boar on the Manor Farm calls the animals on the farm for a meeting where he compares the humans to parasites and teaches the animals a revolutionary song Beasts of England When Major dies two young pigs Snowball and Napoleon assume command and turn his dream into a philosophy The animals revolt and drive the drunken and irresponsible Mr Jones from the farm renaming it Animal Farm They adopt Seven Commandments of Animal ism the most important of which is All animals are equal Snowball attempts to teach the animals reading and writing food is plentiful and the farm runs smoothly The pigs elevate themselves to positions of leadership and set aside special food items ostensibly for their personal health Napoleon takes the pups from the farm dogs and trains them privately Napoleon and Snowball struggle for leadership When Snowball announces his plans to build a windmill Napoleon has his dogs chase	The story begins with the death of Napoleon the original antagonist of Animal Farm The animals of the farm fearing what will become of them learn that Snowball is alive and well and Snowball returns to the farm to encourage capitalism 0.483 second windmill is soon built alongside the first and the two are thenceforth known as the Twin Mills allegorical of the Twin Towers of the World Trade Center and the animals all learn to walk on their hind legs something hitherto forbidden by Old Major shortly before the expulsion of Jones from the farm Also in place of the original Seven Commandments Snowball adopts a single slogan for the animals to live by All animals are born equal what they become is their own affair As time passes the animals under the leadership of Snowball realise the properties of monetary gain and begin to file lawsuits against neighbouring farms allowing Animal Farm to gain land and wealth The revitalised farm also attracts animals

Visualization: Summarization

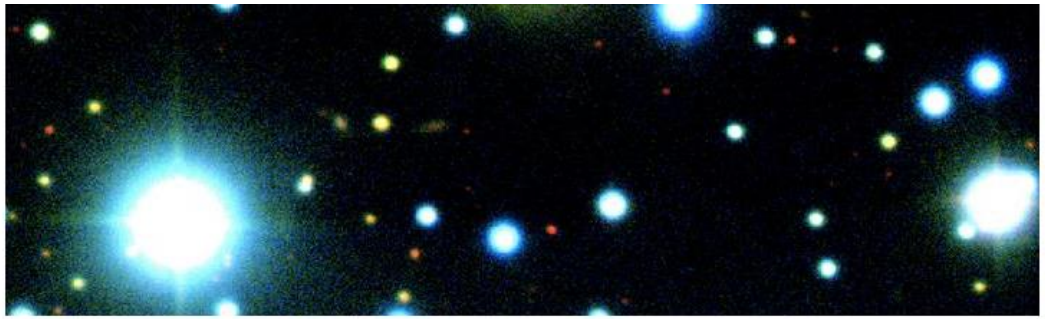


Фото: Shriharsh Tendulkar / Gemini Observatory

«Среда, из которой пришел сигнал, разительно отличается от тех, что содержат ранее локализованные источники быстрых радиовсплесков. Это полностью ломает все предыдущие предположения о природе этих загадочных внегалактических событий», — пишут авторы исследования.

Быстрые радиовсплески являются одной из величайших тайн в астрономии. Происхождение этих коротких вспышек в радиодиапазоне, типичная энергия которых эквивалентна выбросу в космическое пространство энергии, испускаемой Солнцем в течение нескольких десятков тысяч лет, остается загадкой. И хотя эти события длятся всего одну тысячную доли секунды, с момента их первого обнаружения в 2007 году телескопами зафиксировано более сотни сигналов, однако локализовать источники до этого момента удалось только в четырех случаях.

«Многократные сигналы, которые мы улавливали в 2016 году от первого из известных повторяющихся быстрых радиовсплесков, приходили из очень специфических и экстремальных условий внутри крошечной карликовой галактики. С одной стороны, его открытие стало важным шагом на пути к разгадке головоломки о природе этих событий,

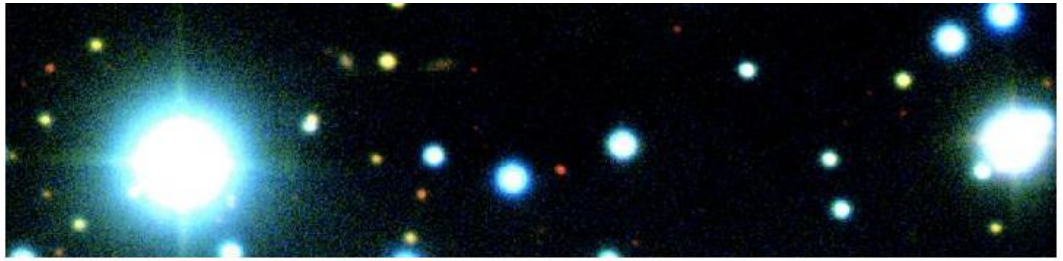


Фото: 🙂Shriharsh 🙂Tendulkar / 💎Gemini 💎Observatory

«Среда, из которой пришел сигнал, разительно отличается от тех, что содержат ранее локализованные источники быстрых радиовсплесков. Это полностью ломает все предыдущие предположения о природе этих загадочных внегалактических событий», — пишут авторы исследования.

Быстрые радиовсплески являются одной из величайших тайн в астрономии. Происхождение этих коротких вспышек в радиодиапазоне, типичная энергия которых эквивалентна выбросу в космическое пространство энергии, испускаемой 🙂Солнцем в течение нескольких десятков тысяч лет, остается загадкой. И хотя эти события длятся всего одну тысячную доли секунды, с момента их первого обнаружения в 2007 году телескопами зафиксировано более сотни сигналов, однако локализовать источники до этого момента удалось только в четырех случаях.

«Многократные сигналы, которые мы улавливали в 2016 году от первого из известных повторяющихся быстрых радиовсплесков, приходили из очень специфических и экстремальных условий внутри крошечной карликовой галактики. С одной стороны, его открытие стало важным шагом на пути к разгадке головоломки о природе этих событий, а с другой — подняло больше вопросов, чем дало ответов», — рассказывает 🙂Бенито 🙂Маркот, ведущий автор исследования из Объединенного института VLBI.

https://www.silkdata.ai/products/summarize-text

Visualization: Text Topics

- [x] [100%] without_topic
- [x] [50.2%] php, mysql, html_cs, lamp, php_mysql, mvc, framework, mvc_framework, javascript_jquery, html
- [ ] [13.1%] code, tdd, framework, unit, continuous_integration, apis, practice
- [ ] [9.7%] tester, startup, manual, like
- [x] [8.2%] marketing, seo, campaign, search, strategy, digital_marketing, social_medium, online, ppc, executive

Essential Qualifications:

- Strong PHP and JavaScript programming skills.
- Extensive experience in HTML and CSS.
- Experience working with Wordpress.
- Nice to have experience with Bootstrap.
- Experience working with Google Analytics.
- Up-to-date knowledge of modern design.
- Familiarity with the latest marketing tools such as AB Testing tools, Google Tag Mananger, Google Web Master Tools, Optimizely, Crazy Egg.
- Up-to-date knowledge of Search Engine Optimization techniques.
- Self-motivated, with ability to learn quickly and demonstrate a strong technical aptitude.
- Good spoken English is a must.

Data source: https://www.kaggle.com/c/job-salary-prediction

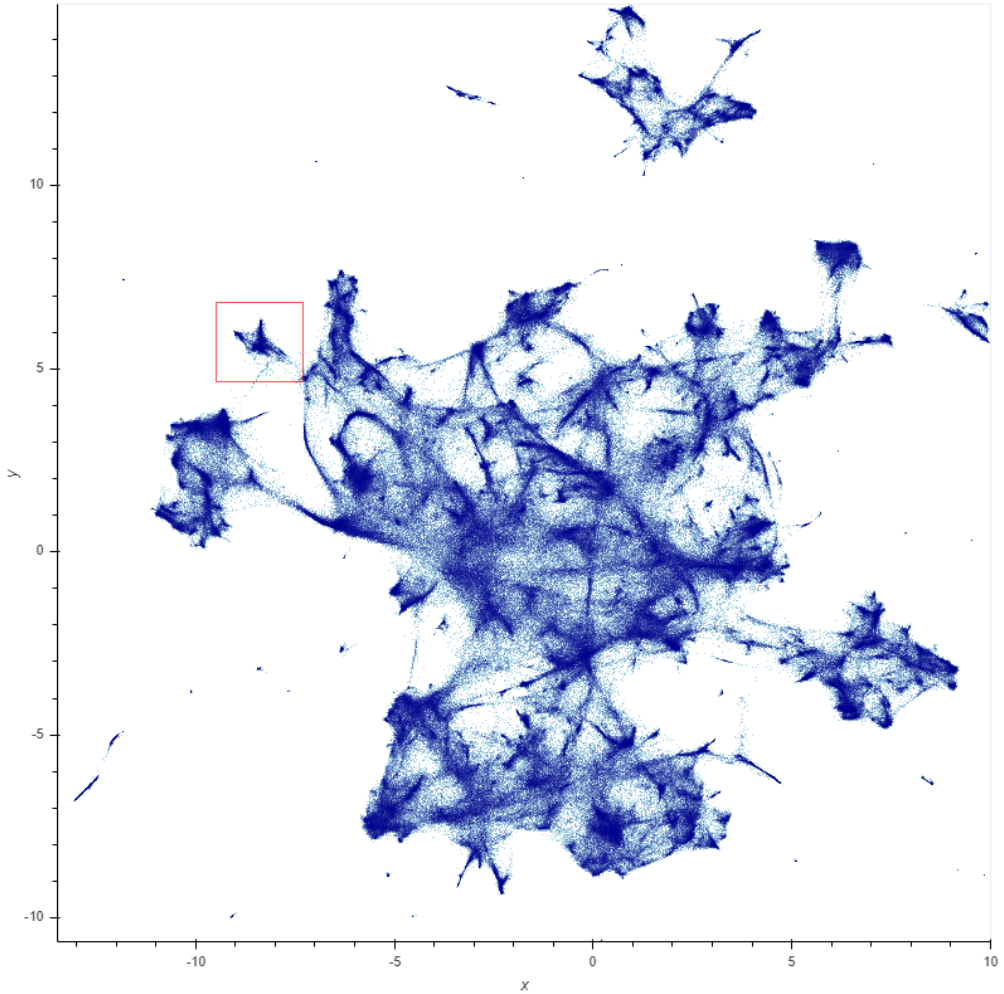
‘Explainer’ algorithm: NLP

Deep learning and sentiment analysis:

True Label	Predicted Label	Attribution Label	Attribution Score	Word Importance
pos	pos (0.99)	pos	1.54	It was a fantastic performance ! pad
pos	pos (0.71)	pos	1.69	Best film ever pad pad pad pad
pos	pos (0.95)	pos	1.49	Such a great show ! pad pad
neg	neg (0.18)	pos	-1.19	It was a horrible movie pad pad
neg	neg (0.22)	pos	-1.70	I 've never watched something as bad
neg	neg (0.38)	pos	-1.29	It is a disgusting movie ! pad

Source: https://captum.ai/tutorials/IMDB_TorchText_Interpret

SilkData: Wikipedia navigation (semantic cloud)



Related Wikipedia articles

#	Score	Title
0	27	role-playing video game
1	24	history of video games
2	19	list of commercial failures in video g
3	18	mario party
4	18	video game console
5	16	platform game
6	16	nintendo
7	15	tactical role-playing game
8	14	music of the final fantasy series
9	13	greyhawk
10	12	sega saturn
11	12	wii
12	12	the legend of zelda
13	12	handheld game console
14	11	dragon quest (video game)
15	11	final fantasy vii
16	11	deus ex (video game)
17	11	playstation 3
18	11	halo (franchise)
19	11	dreamcast
20	11	real-time strategy
21	10	super smash bros.
22	10	super nintendo entertainment syste
23	10	final fantasy vi
24	10	final fantasy

Words in selection

#	Score	Word
0	1,890	playstation
1	1,618	gameplay
2	1,118	xbox
3	1,112	nintendo
4	876	multiplayer
5	761	role-playing
6	686	wii
7	675	single-player
8	628	gamecube
9	532	visuals
10	392	dreamcast
11	371	minigames
12	358	roguelike
13	354	turn-based
14	351	in-game
15	343	ign
16	326	reviewer
17	318	console
18	314	enjoyable
19	312	sega
20	306	gamepro
21	283	unlock
22	282	action-adventure
23	280	metacritic
24	267	famicom

https://pegasus.silkcodeapps.de/modeling/modeling

https://www.silkdata.ai/semanticmap

Thank you for attention!

visit: silkdata.ai mail to: hello@silkdata.ai

www.silkdata.ai
hello@silkdata.ai
+48 452 380 167

Silk Data Sp. z o.o.

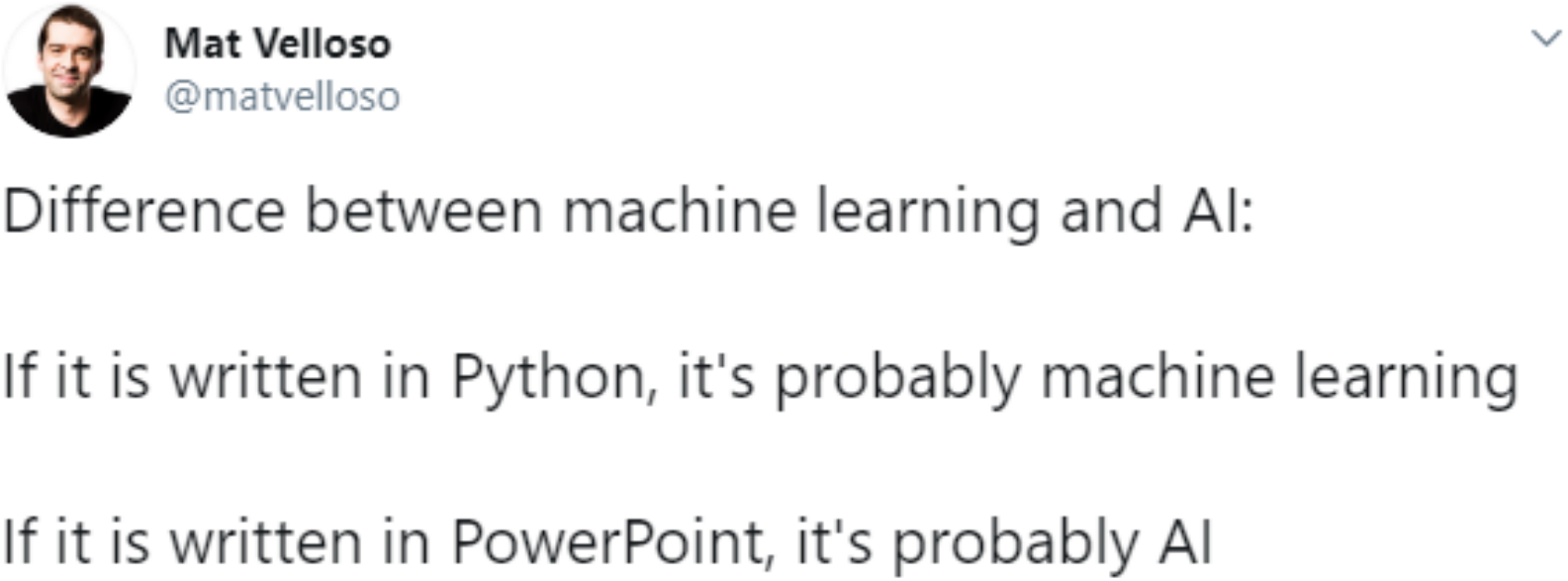
Domaniewska 17/19, off. 133
02-672 Warsaw,
Poland
+48 452 380 167

SilkCode GmbH

Luisenstraße 62
D-47799 Krefeld
Germany
+49 (2151) 387 3531

Extra slides

AI vs. ML



By Mat Velloso, an adviser to Satya Nadella at Microsoft

Source: https://twitter.com/matvelloso/status/1065778379612282885
https://hub.packtpub.com/so-you-want-to-learn-artificial-intelligence-heres-how-you-do-it/

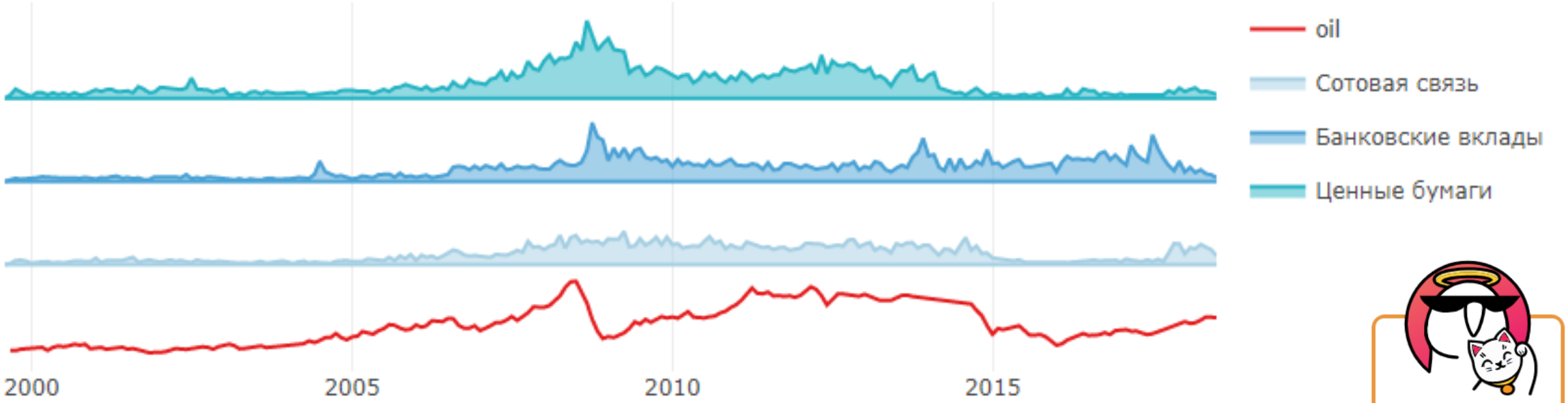
Visualization

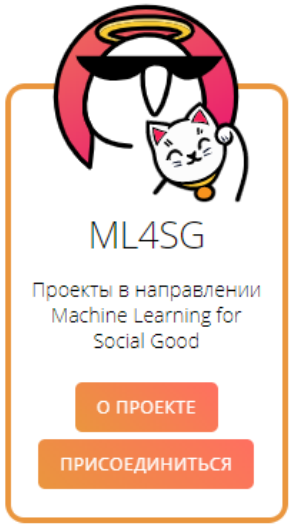
Warning!

- For non-mathematicians, this is going to be tougher than the previous material.
  - You may have to spend a long time studying the next two parts.
- If you are not used to thinking about hyper-planes in high-dimensional spaces, now is the time to learn.
- To deal with hyper-planes in a 14-dimensional space, visualize a 3-D space and say "fourteen" to yourself very loudly. Everyone does it.
  - But remember that going from 13-D to 14-D creates as much extra complexity as going from 2-D to 3-D.

Source: https://twitter.com/drbeef_/status/1005887240407379969

Visualization: topic dynamics



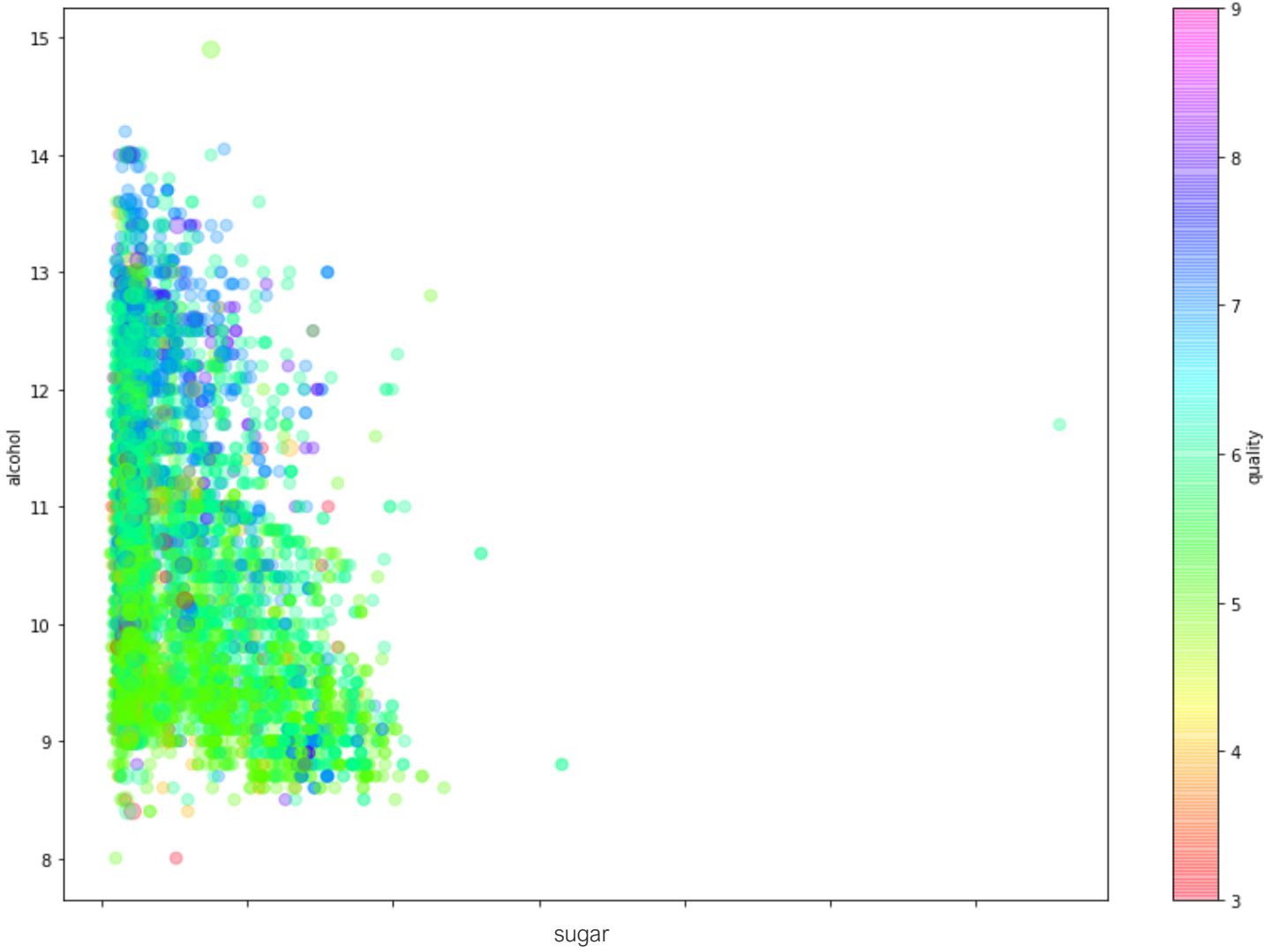


Source: http://51.15.75.134:8080/
https://ods.ai/

‘Explainer’ algorithms

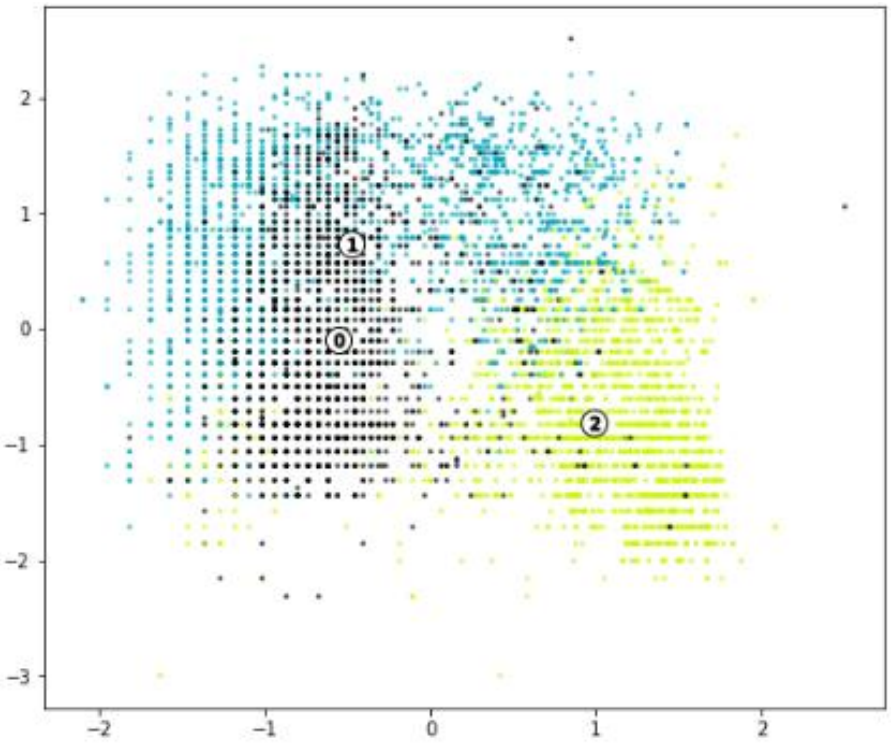
Wine quality dataset

Parameters:
'fixed acidity’,
'volatile acidity’,
'citric acid’,
'residual sugar',
'chlorides',
'free sulfur dioxide’,
'total sulfur dioxide',
'density’,
'pH’,
'sulphates’,
'alcohol',
'quality'

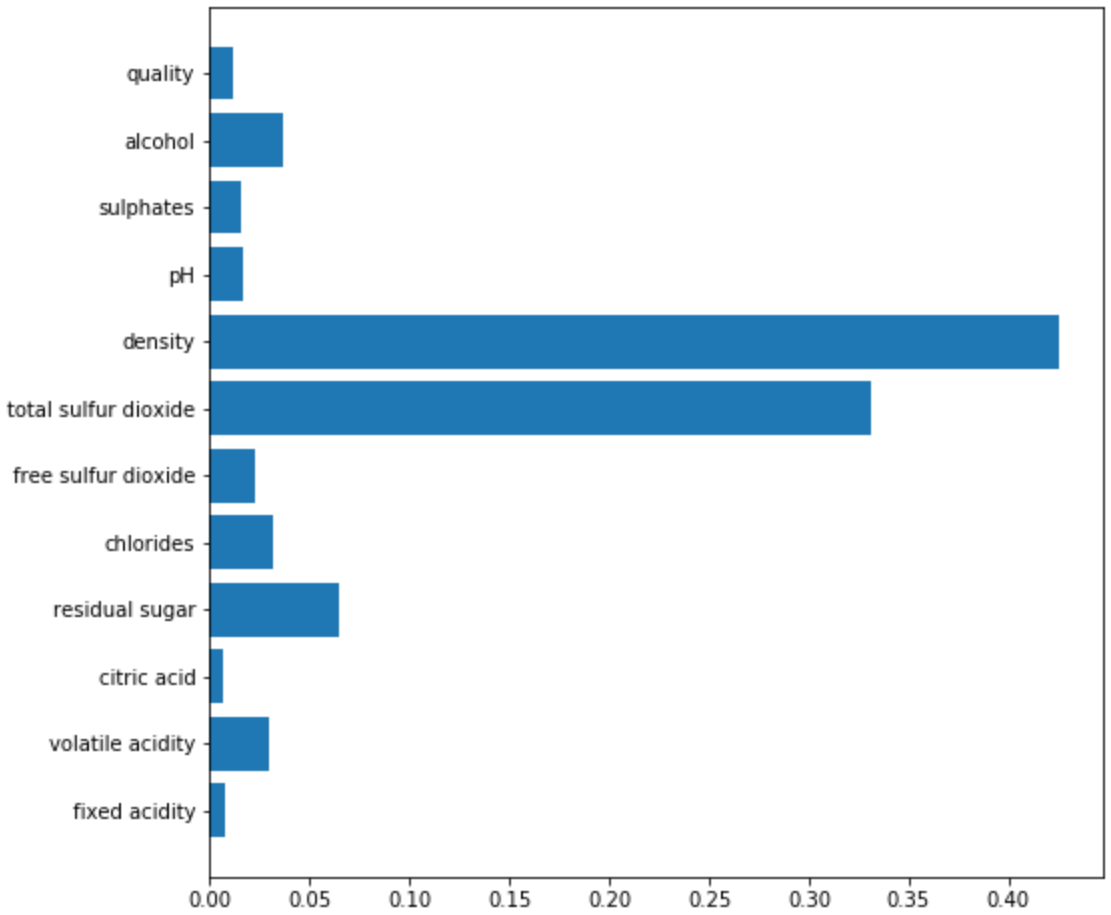


Dataset source: https://archive.ics.uci.edu/ml/datasets/Wine+Quality

‘Explainer’ algorithms



Parameter importance:



Wine quality dataset

Dataset source: https://archive.ics.uci.edu/ml/datasets/Wine+Quality

‘Explainer’ algorithm: NLP

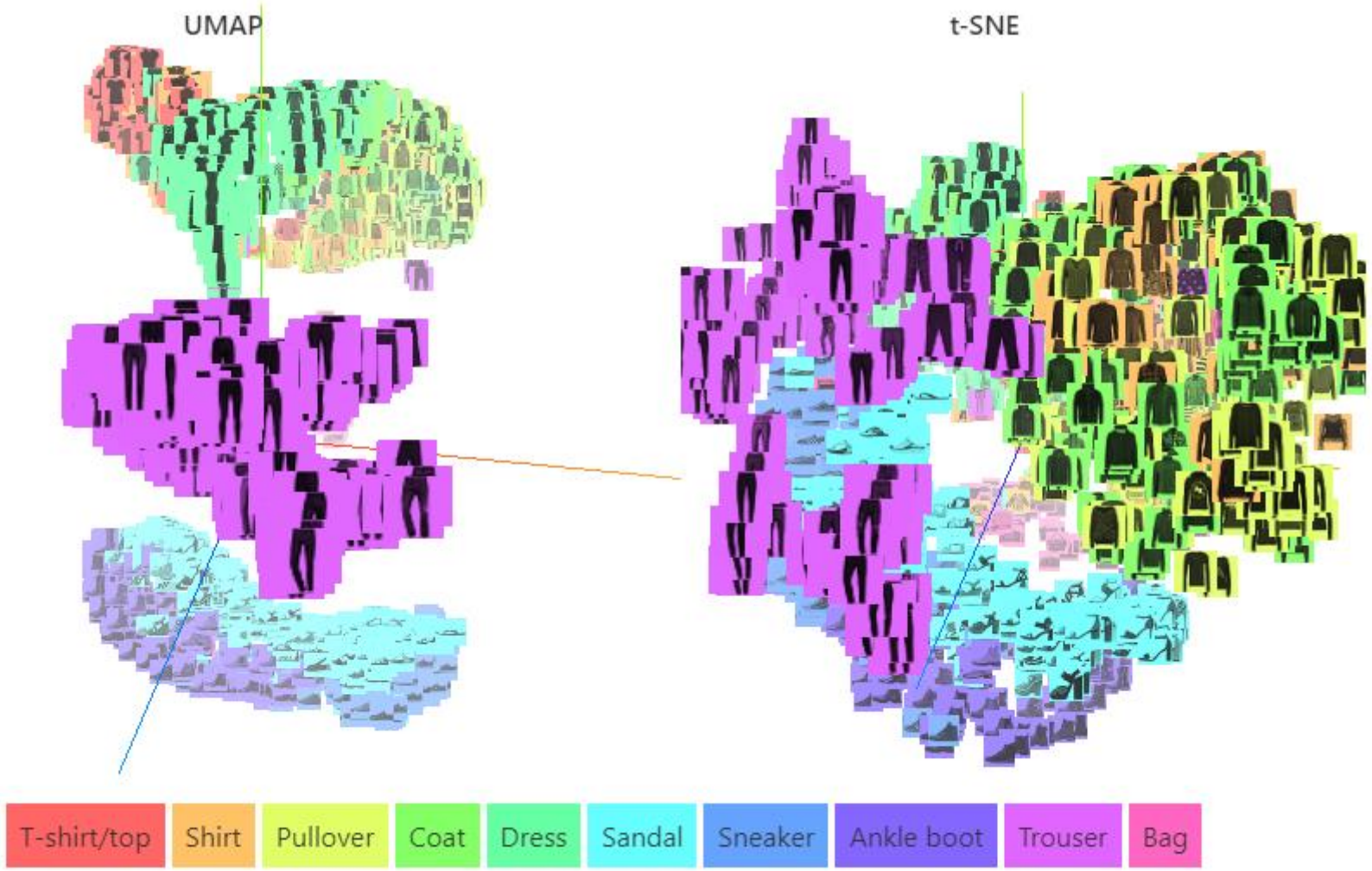
y=sci.med (probability **0.576**, score **0.621**) top features

Contribution?	Feature
+0.972	Highlighted in text (sum)
-0.351	<BIAS>

as i recall from my bout with kidney stones, there isn't any medication that can do anything about them except relieve the pain. either they pass, or they have to be broken up with sound, or they have to be extracted surgically. when i was in, the x-ray tech happened to mention that she'd had kidney stones and children, and the childbirth hurt less

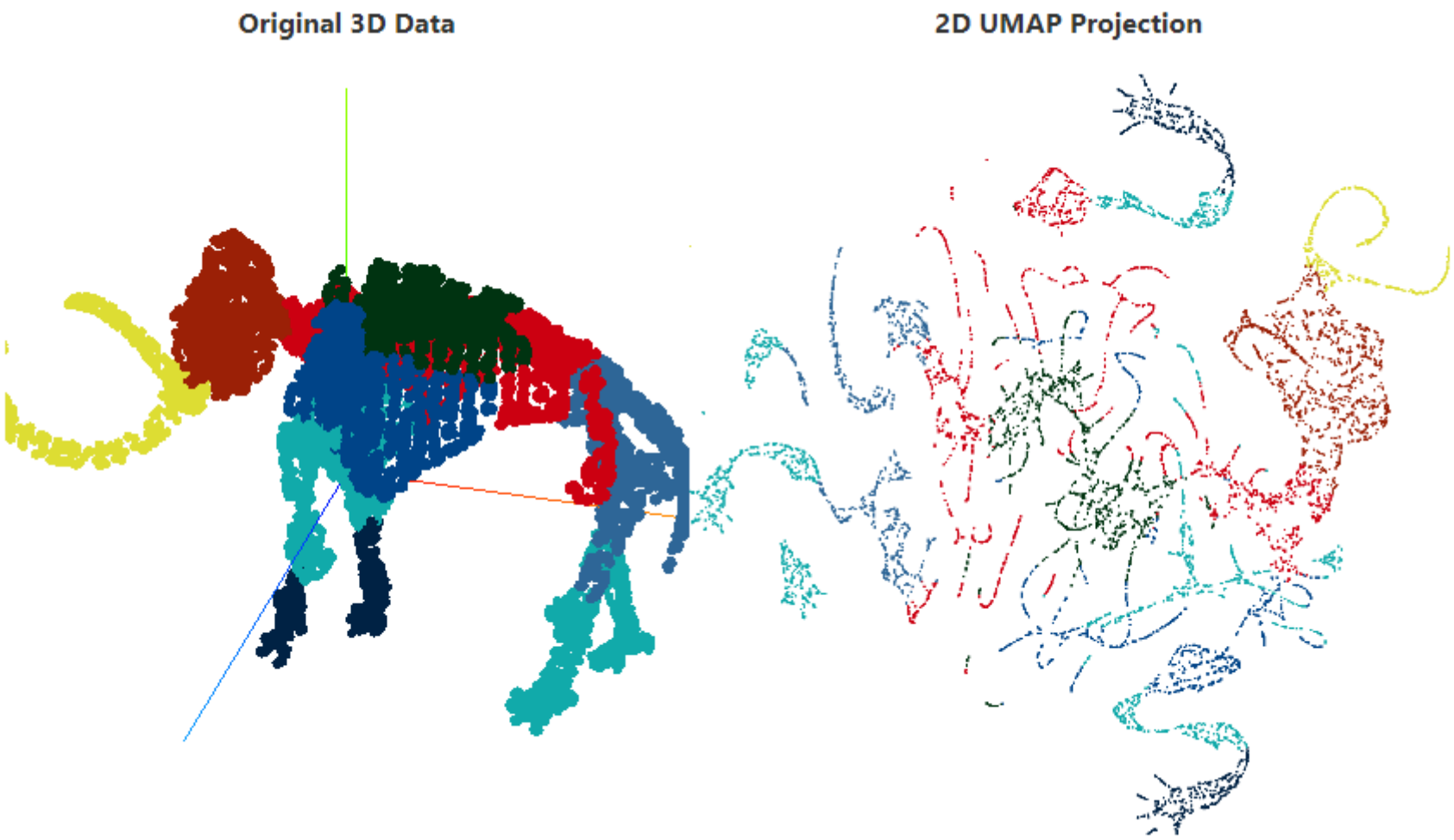
https://github.com/TeamHG-Memex/eli5

UMAP vs. t-SNE



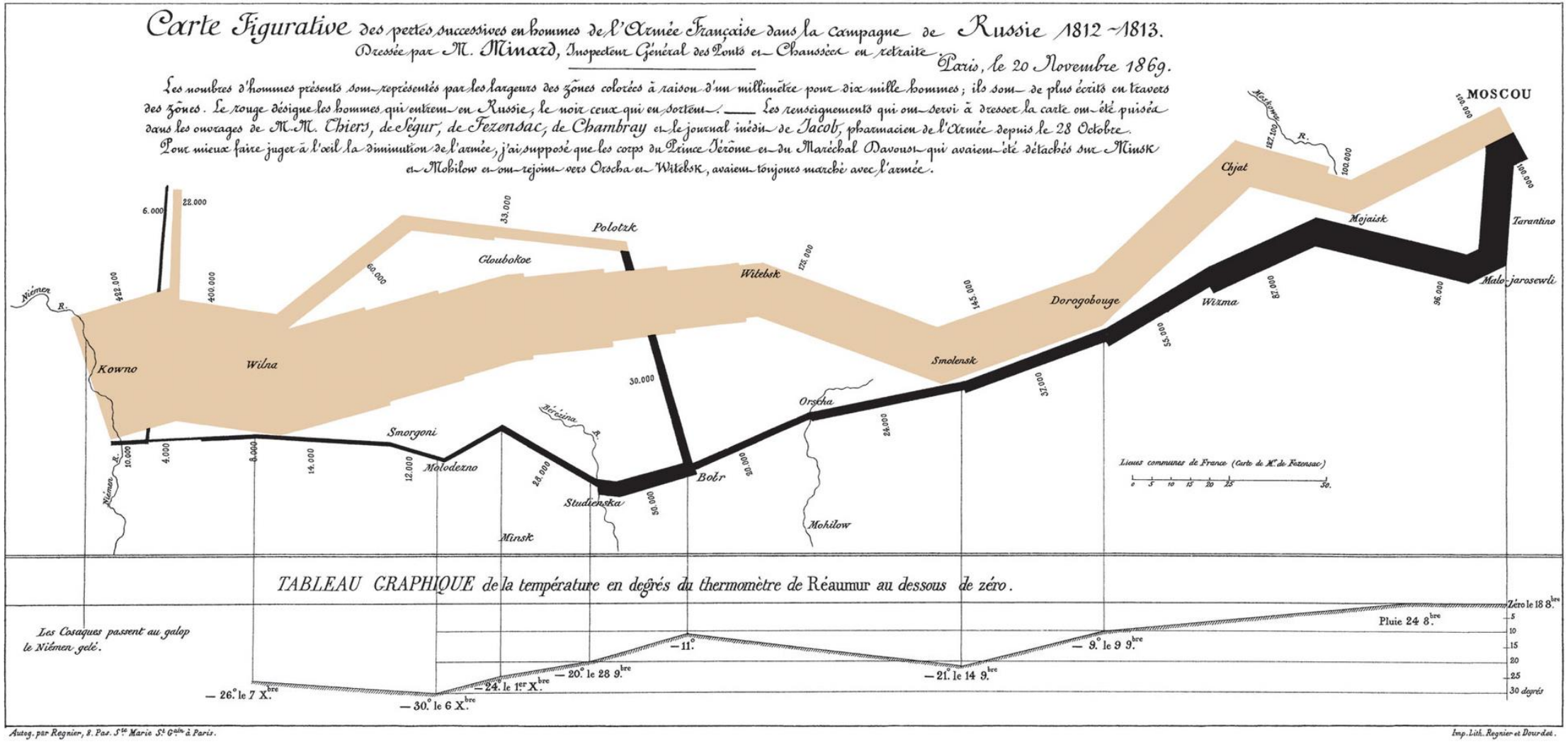
Source: https://pair-code.github.io/understanding-umap/

UMAP



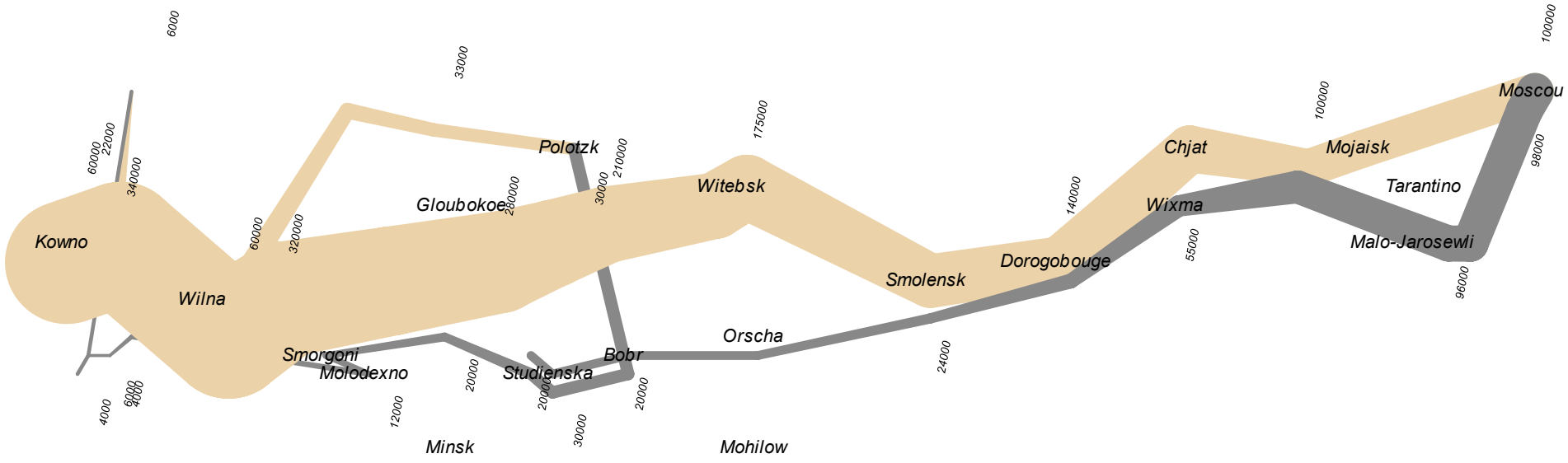
Source: https://pair-code.github.io/understanding-umap/

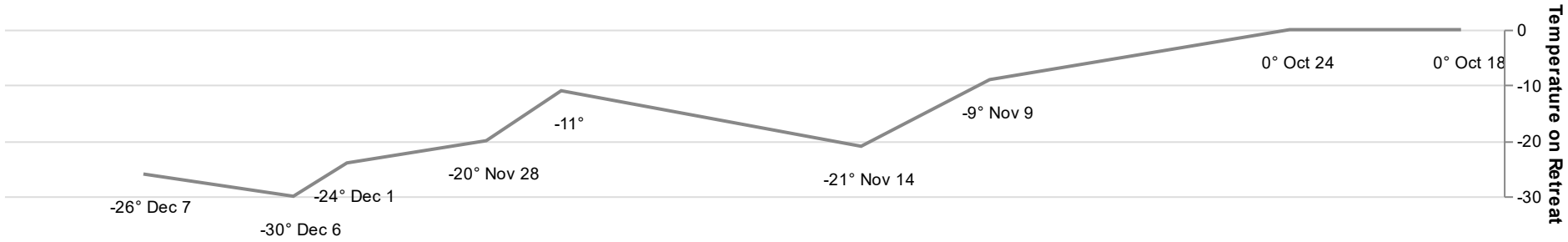
Charles Minard's map



Source: https://en.wikipedia.org/wiki/Charles_Joseph_Minard

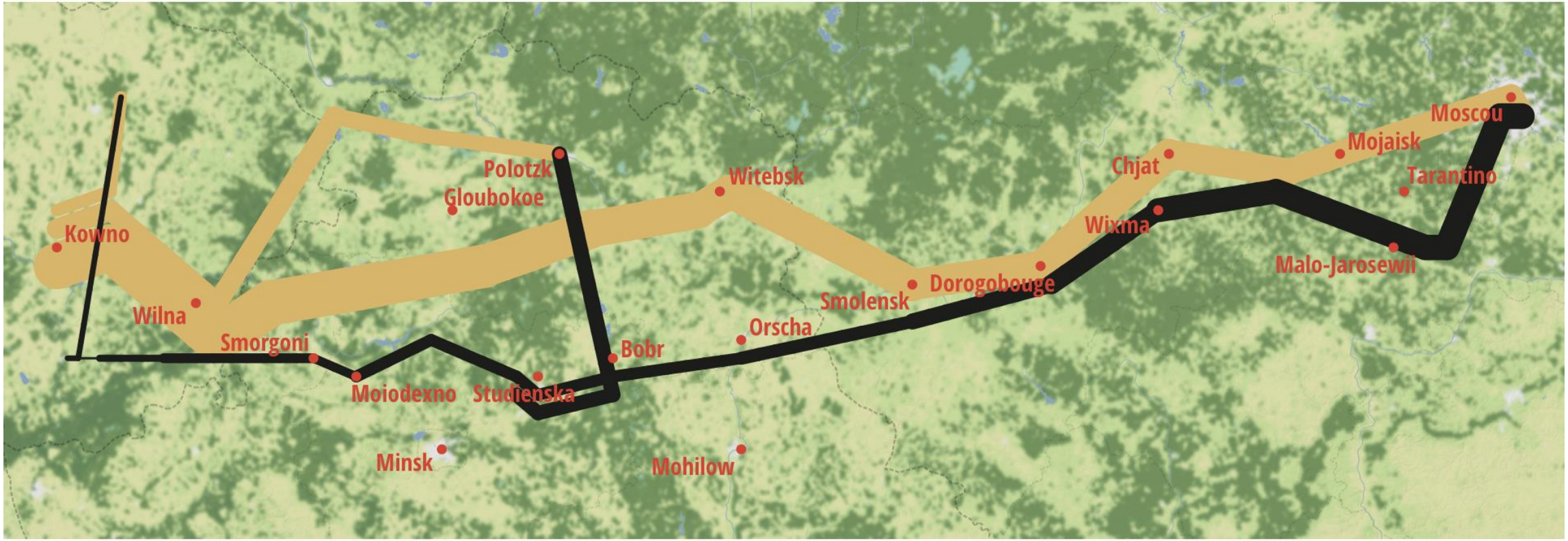
Charles Minard's map: Altair





Source: https://github.com/blue-yonder/grammar-of-graphics-minard-example

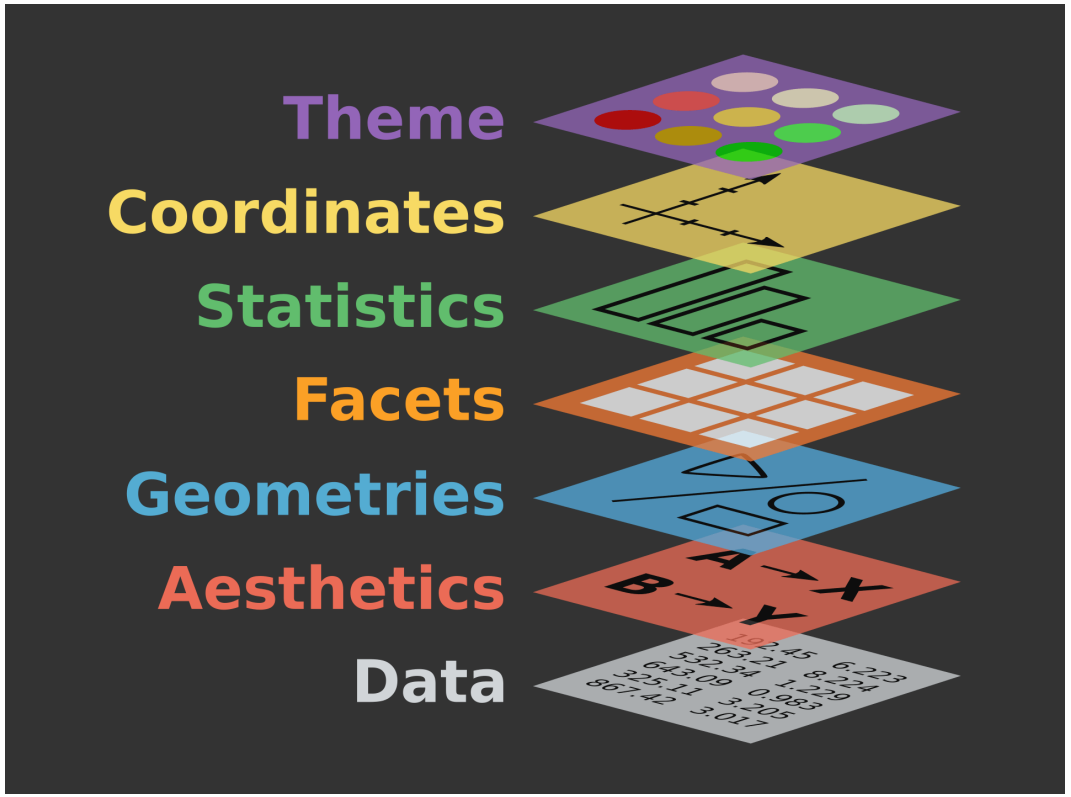
Charles Minard's map: ggplot2

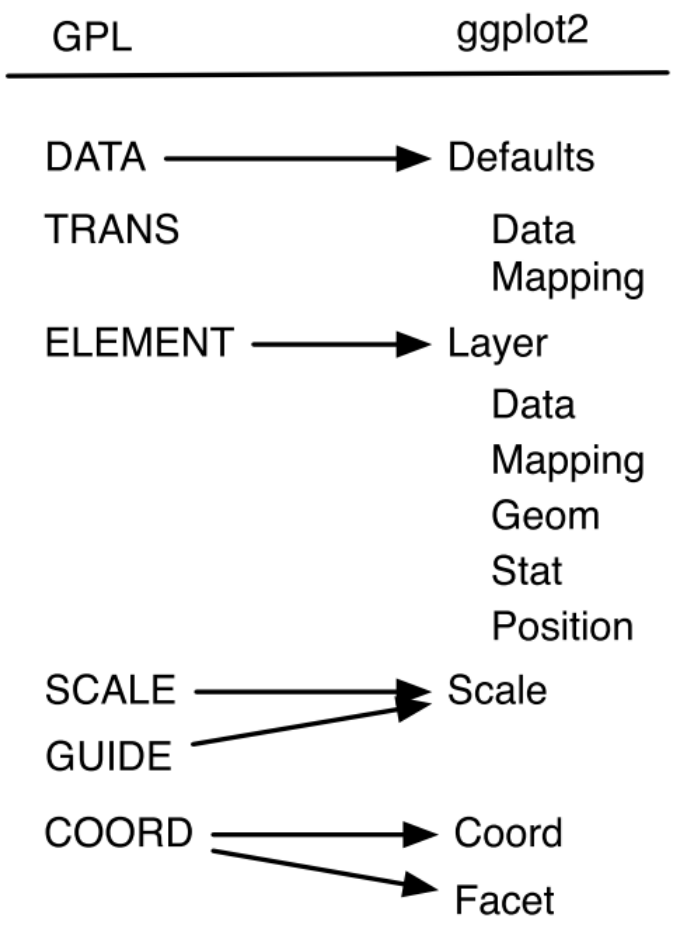


Source: https://github.com/andrewheiss/fancy-minard

Layered Grammar of Graphics

- A Layered Grammar of Graphics, Hadley WICKHAM (2010)
- Grammar: “the fundamental principles or rules of an art or science”
- ggplot2 R package

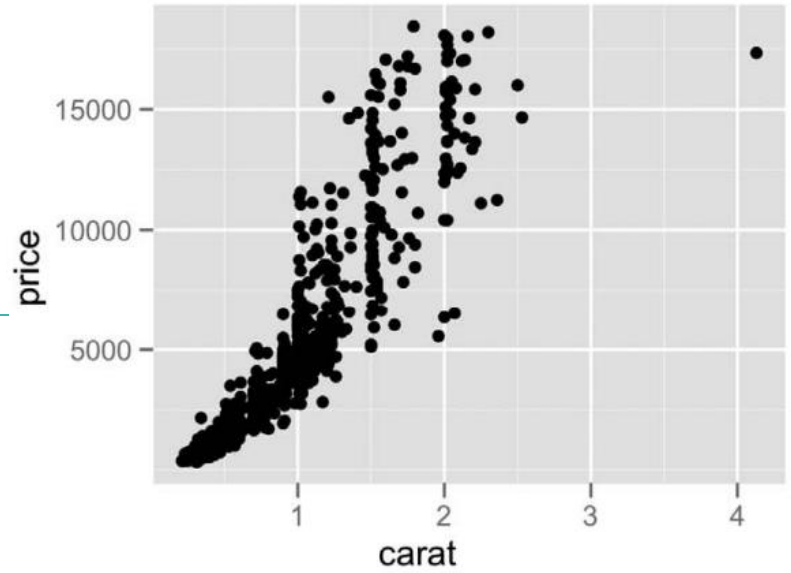


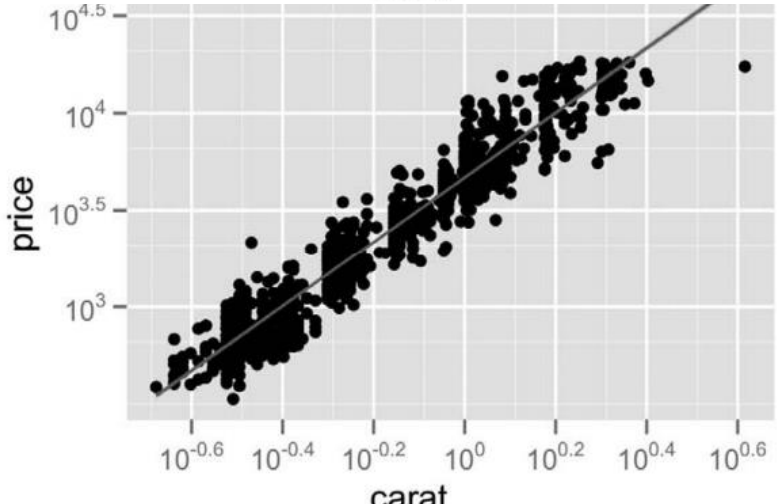


Example

```
ggplot() +
  layer(
    data = diamonds, mapping = aes(x = carat, y = price),
    geom = "point", stat = "identity", position = "identity“
  ) +
scale_y_continuous() +
scale_x_continuous() +
coord_cartesian()

… stat_smooth(method = lm) +
scale_x_log10() +
scale_y_log10()
```





Packages

- Java
  - GPL (not available)
- R
  - ggplot2
- Java Script
  - Vega + Vega-Lite
- Python
  - plotnine
  - altair (Vega + Vega-Lite)