Have no fear: Just the math you need to know for monitoring

Lerna Ekmekcioglu

Sr Solutions Architect, AWS

Thoughts and opinions are my own and do not represent that of my employer

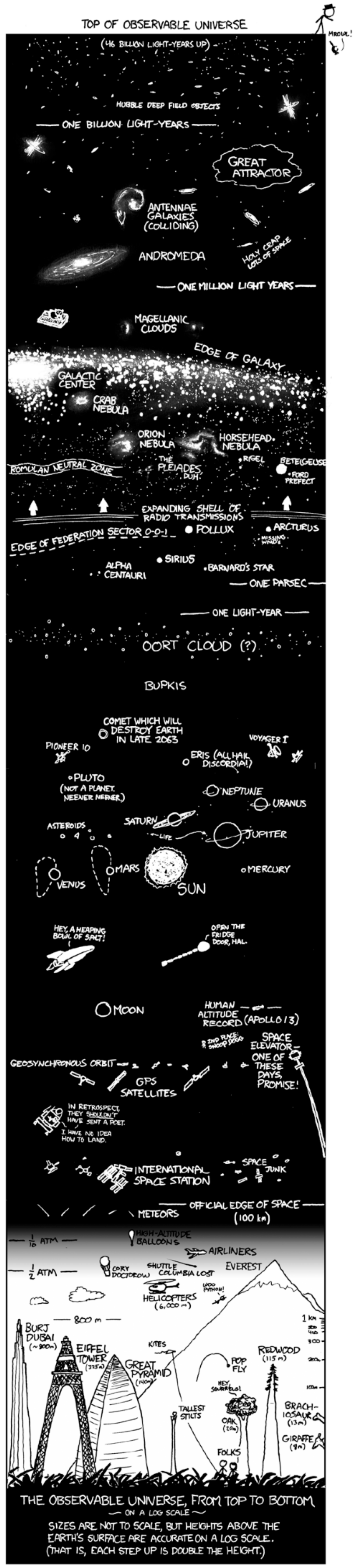
Agenda

- Visualizing the data: Logarithmic scale
- Summarizing the data:
  - Center
  - Spread
  - Relative standing

Visualizing the data: Logarithmic scale

Logarithms

https://xkcd.com/482: Height



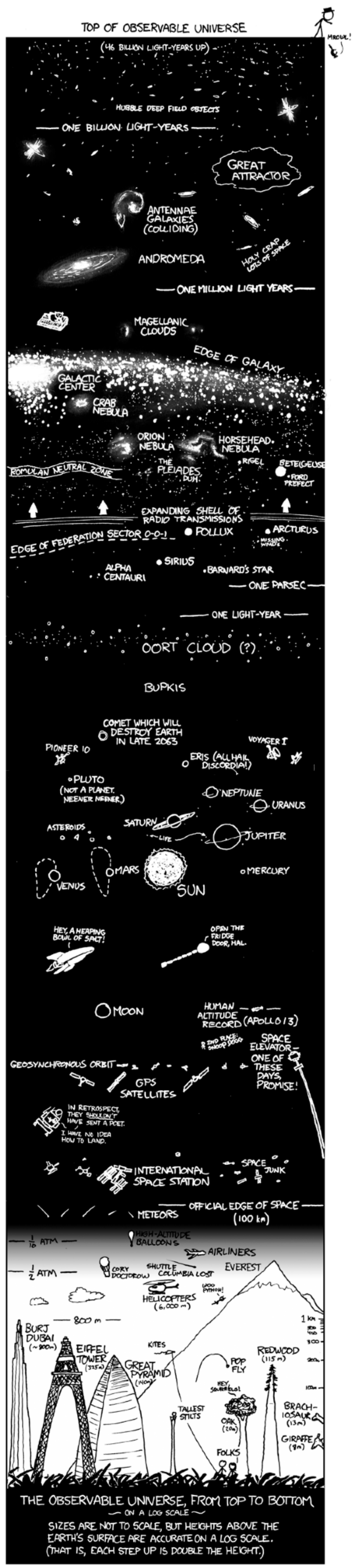
https://xkcd.com/482: Height, CC BY-NC 2.5
https://www.explainxkcd.com/wiki/index.php/482:_Height

Logarithms

The observable universe from top to bottom
~ on a log scale ~

…**heights** above the earth's surface are
accurate on a log scale.

https://xkcd.com/482: Height, CC BY-NC 2.5
https://www.explainxkcd.com/wiki/index.php/482:_Height



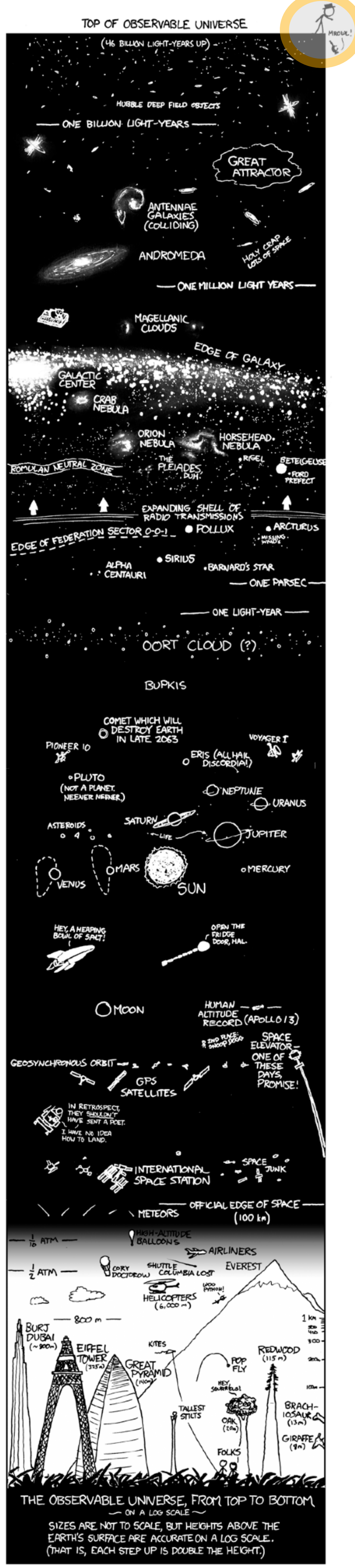
Why logarithms?

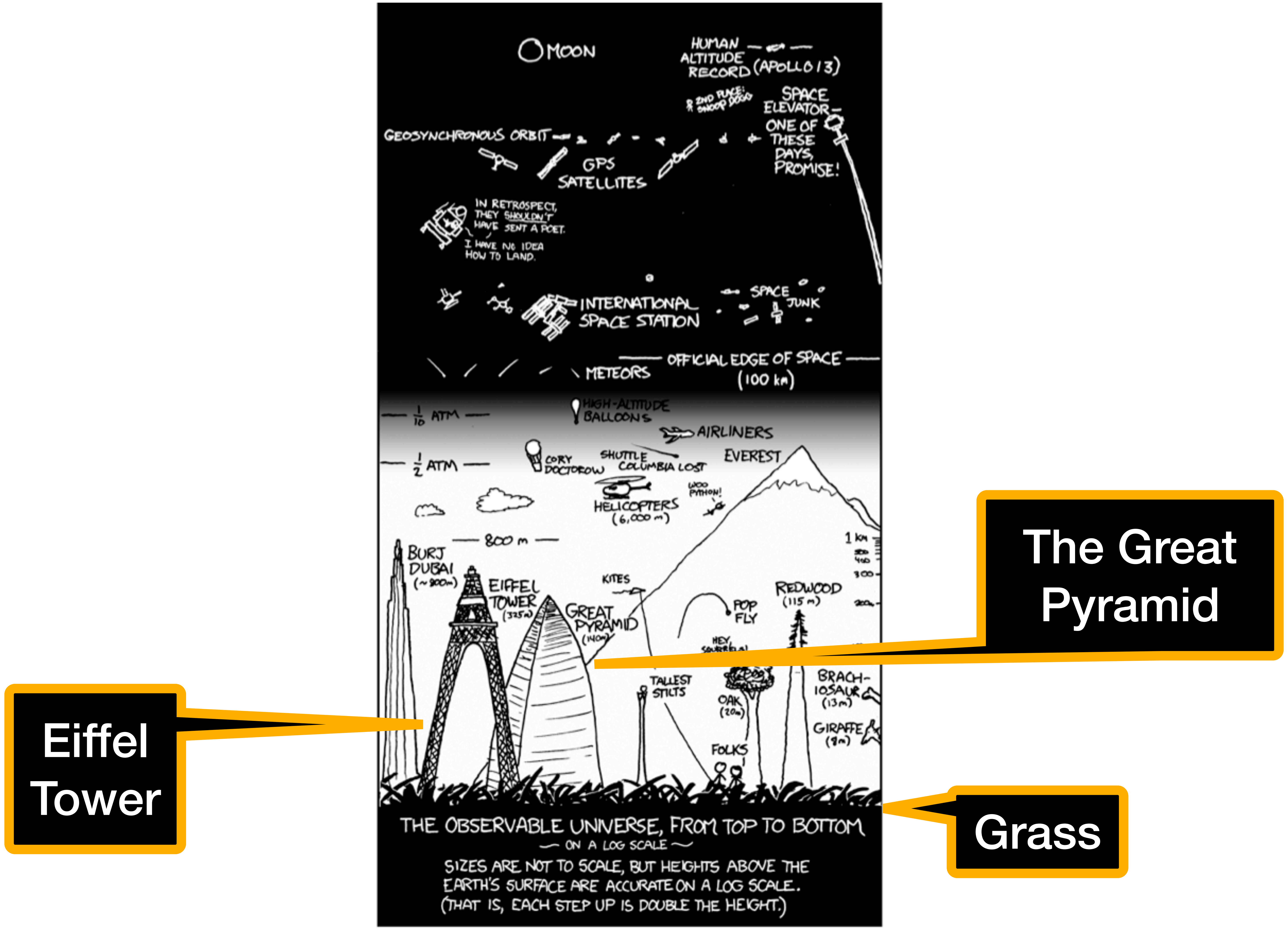
$$435\ yottameters = 46\ billion\ light - years$$

Top of observable universe

Display data
over a very wide range of
values

Height above earth's surface





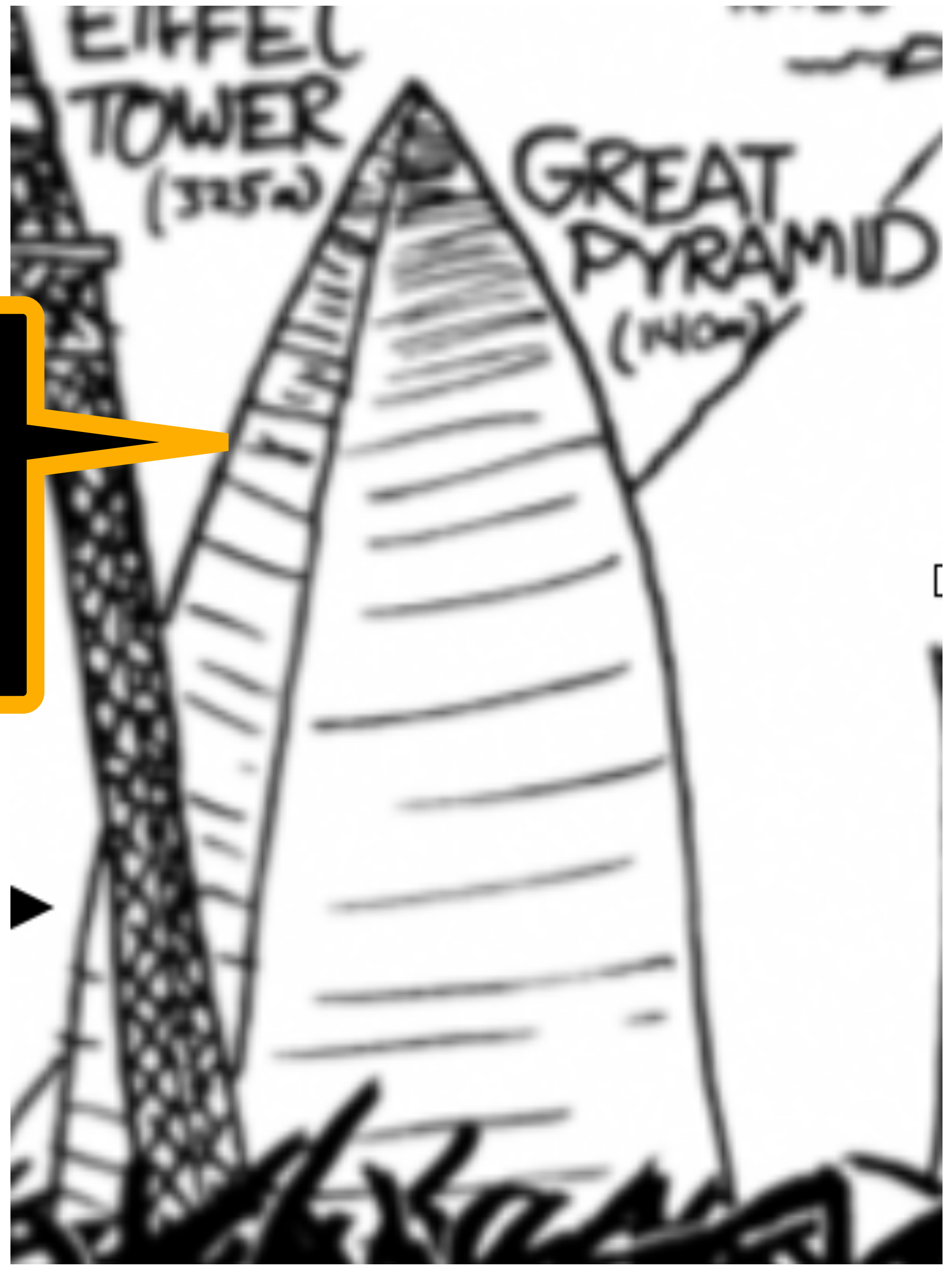
https://xkcd.com/482: Height, CC BY-NC 2.5
https://www.explainxkcd.com/wiki/index.php/482:_Height

The Great Pyramid: linear vs log scale

Linear scale:
Pyramid

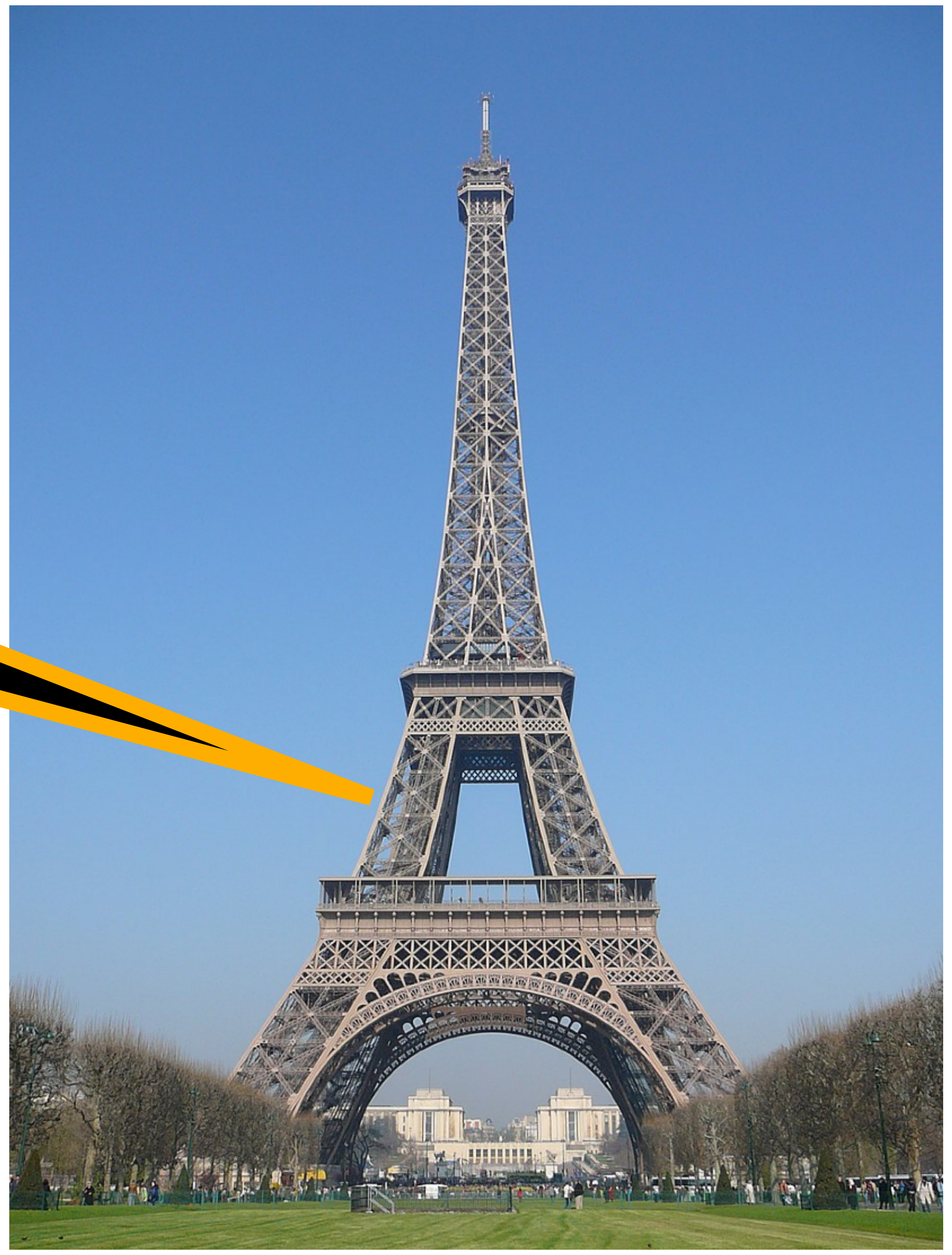
Daguerreotype of the Great Pyramid taken by Frédéric Goupil-Fesquet
on 20 November 1839

Log scale:
Stretched up



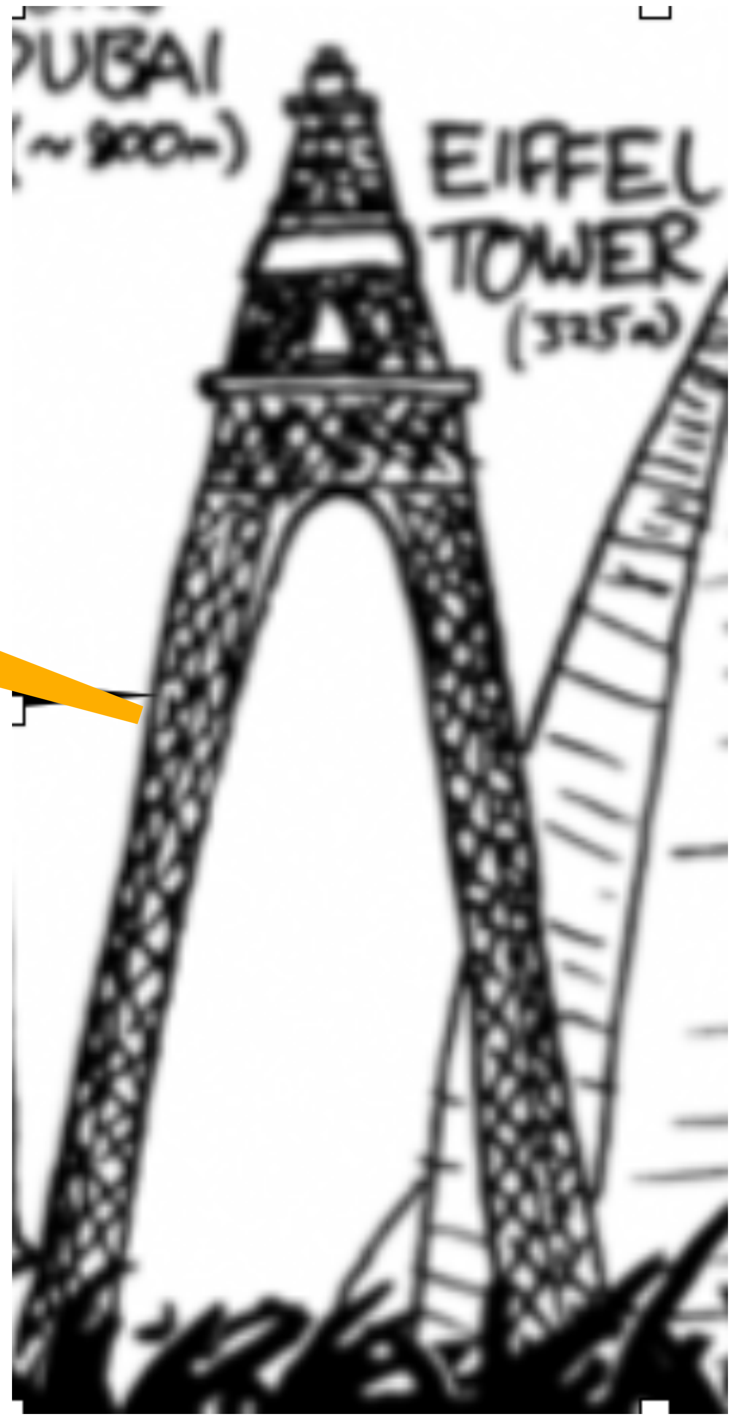
The Eiffel Tower: linear vs log scale

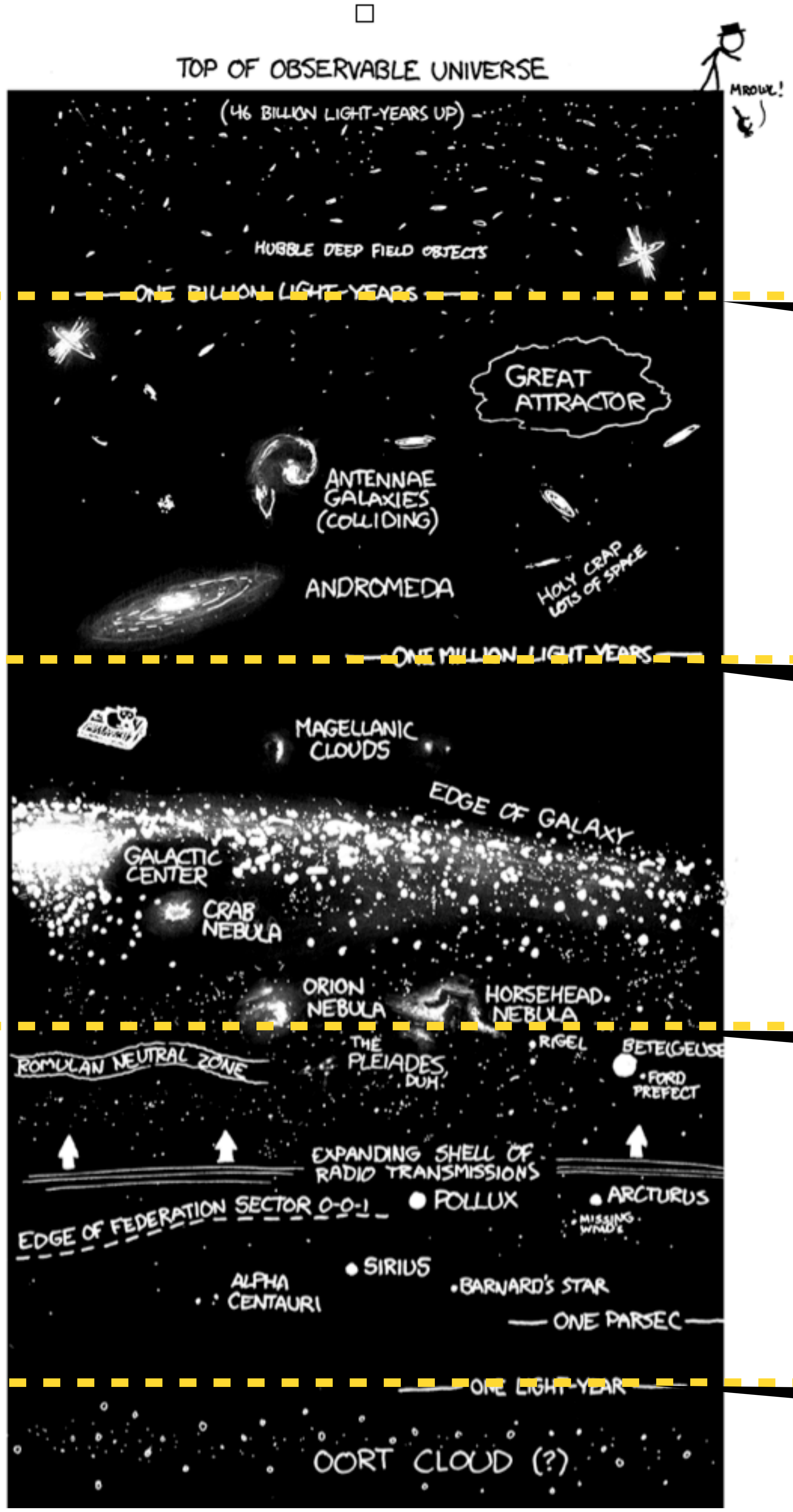
Eiffel Tower



Linear scale: Exponential Curve

Log scale: Straight line





1 Billion Light-Years
$9.46 * 10^{24} m$

10^3

1 Million Light-Years
$9.46 * 10^{21} m$

10^3

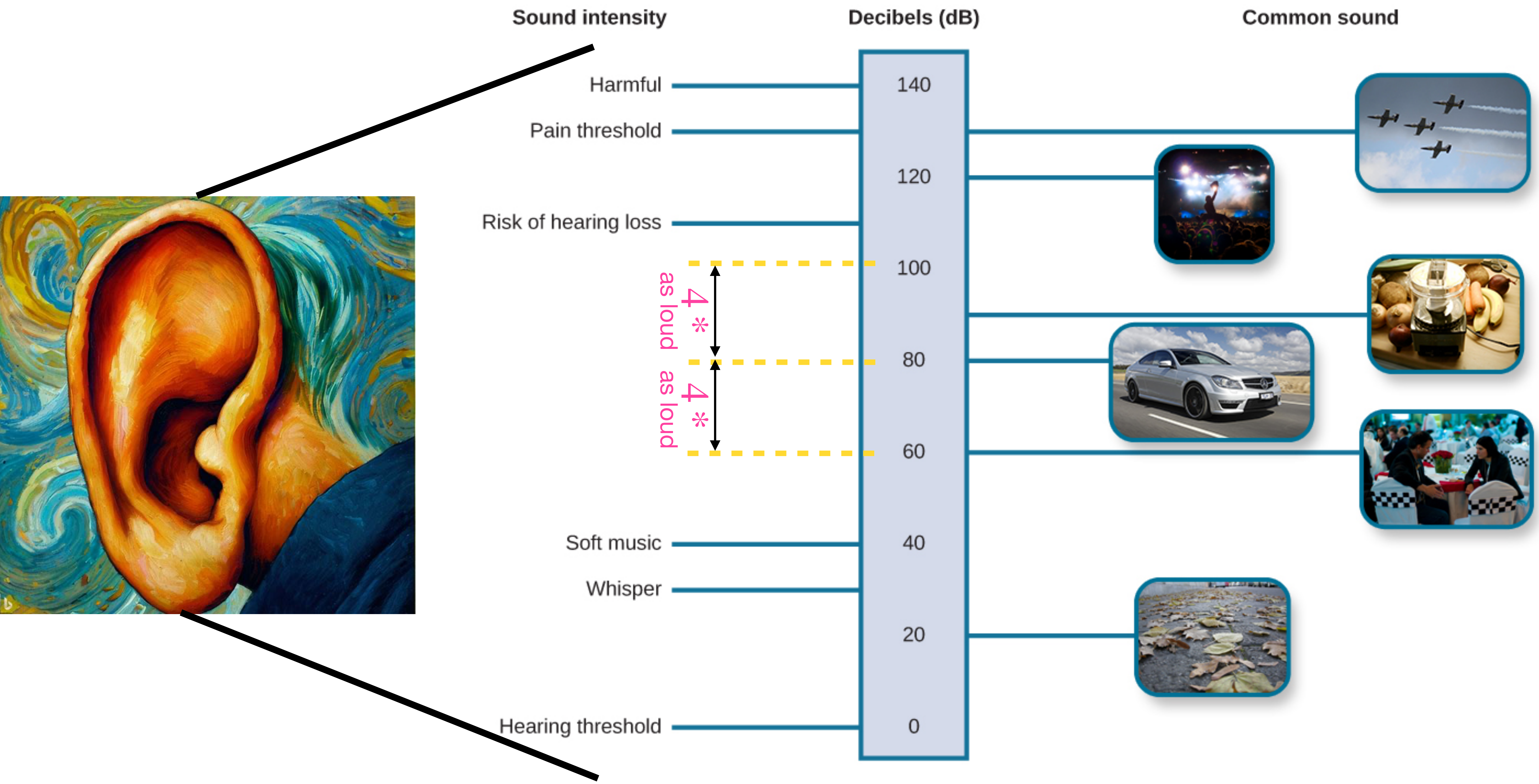
1000 Light-Years
$9.46 * 10^{18} m$

10^3

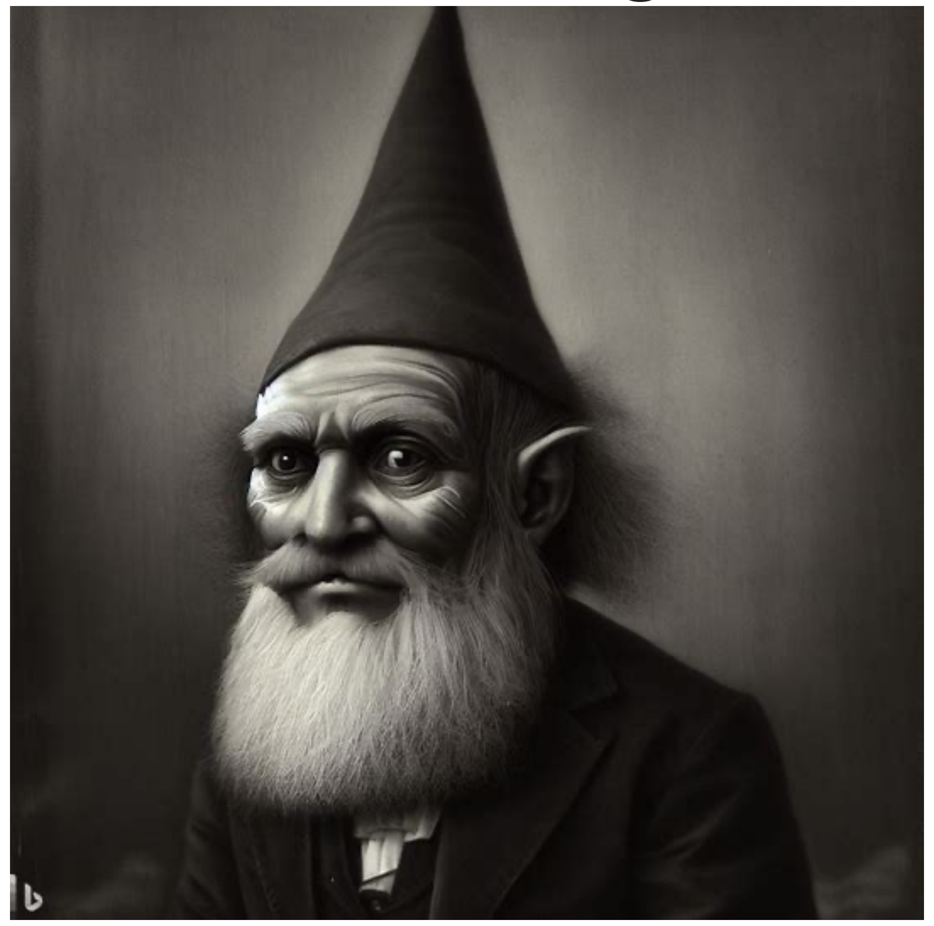
1 Light-Year
$9.46 * 10^{15} m$

https://xkcd.com/482: Height, CC BY-NC 2.5
https://www.explainxkcd.com/wiki/index.php/482:_Height

Ears perceive sounds at log scale



Gustav, as a gnome

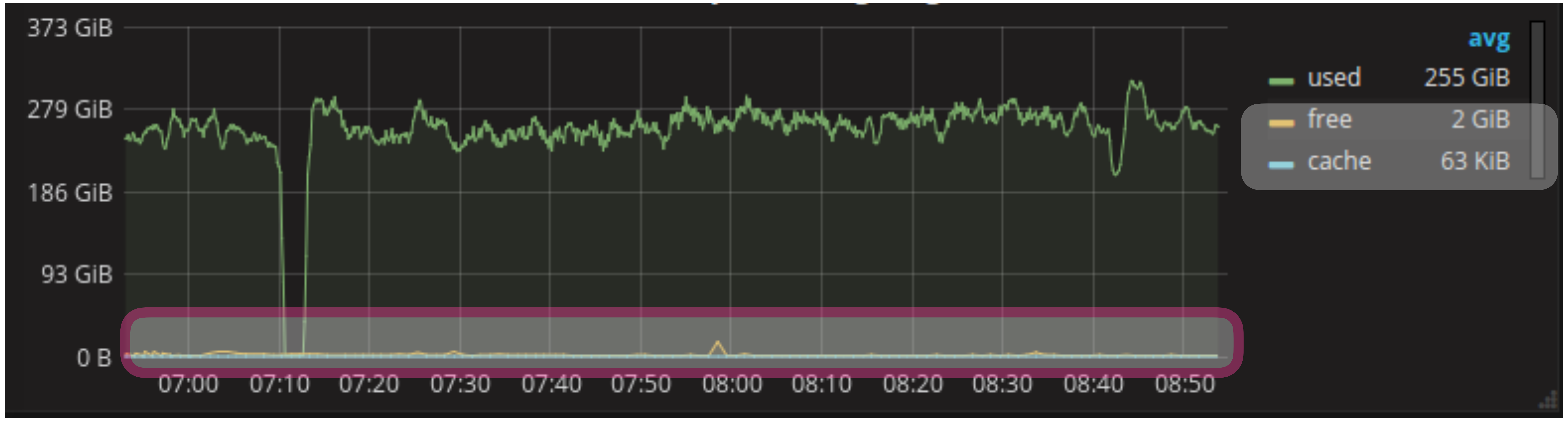


Gustav Fechner
(1801-1887)

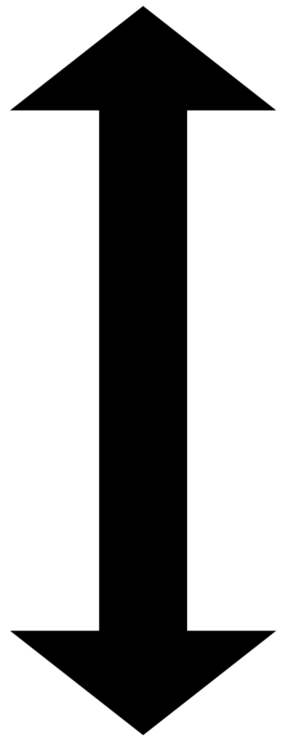
Lumen Learning, https://courses.lumenlearning.com/wsu-sandbox/chapter/waves-and-wavelengths/, (CC BY-SA 4.0)

https://commons.wikimedia.org/wiki/File:Gustav_Fechner.jpg

Linear

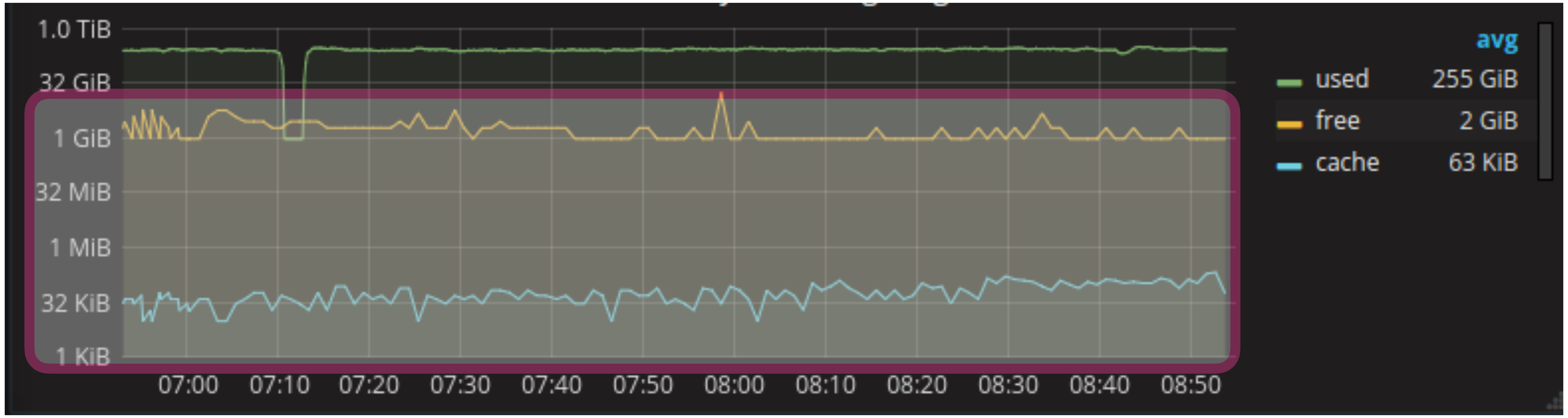


$255GiB = 2.67 * 10^8 KiB$



63 KiB

Log Scale



https://community.grafana.com/t/advanced-graphing-part3-right-y-axis-log-scales/400

Origins of logarithm

Large calculations required for:

- Astronomy
- Celestial navigation of ships

Mathematicians pain points in 1600s

"…**nothing** is **more tediou**s, ..**than** the **great delays suffered** in the tedium of **lengthy multiplications and divisions**,

.. I have found an **amazing way of shortening** .. **multiplications**, **and divisions** .. **by** .. **addition, subtraction**, …"

Description of the Wonderful Law of Logarithms, 1614, John Napier

John, as a gnome

John Napier
(1550-1617)

The wonderful laws of logarithms

Turn **multiplications into additions**

$$\log(x * y) = \log(x) + \log(y)$$

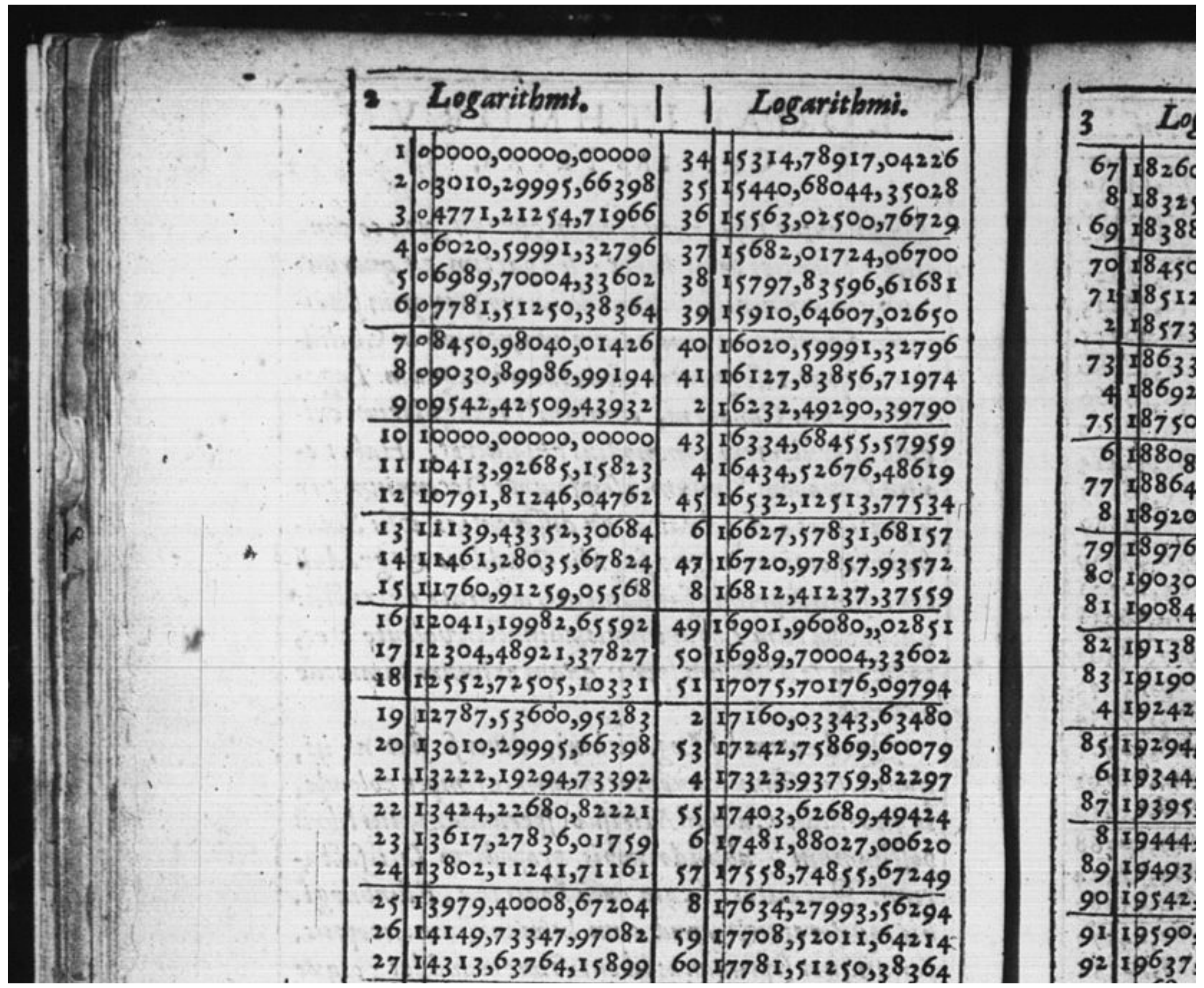
$$\log(10^2 * 10^3) = \log 10^2 + \log 10^3$$

Turn **divisions into subtractions**

$$\log(\frac{x}{y}) = \log(x) - log(y)$$

$$\log(10^3/10^2) = \log 10^3 - \log 10^2$$

Henry Briggs' first table of common logarithms, Logarithmorum Chilias, 1617
Prima, from 1617.
This page tabulates the base-10 logarithms of the numbers 0 to 67 to fourteen decimal places.



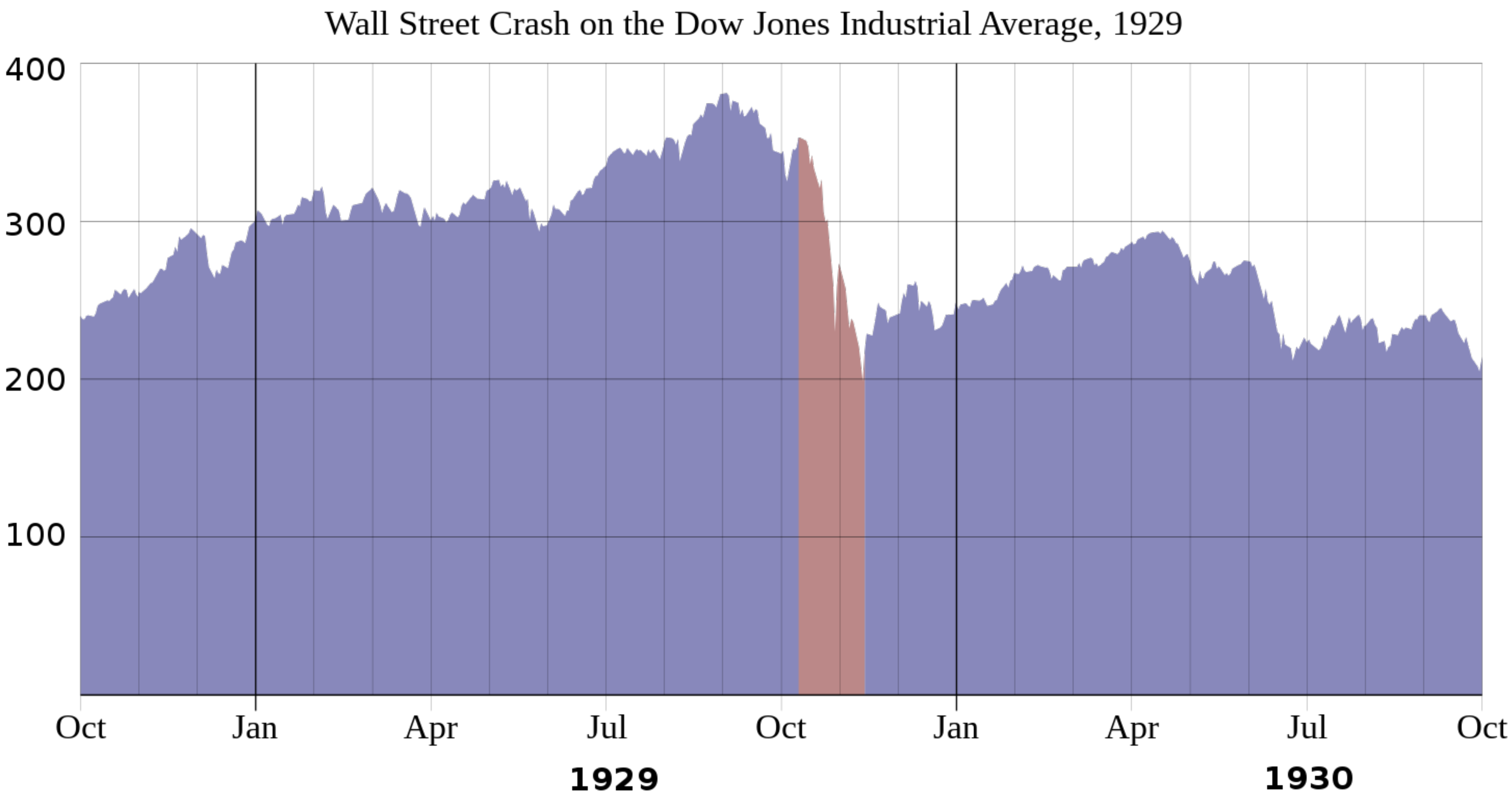
2	Logarithmi.			Logarithmi.	
1	0	0000,00000,00000	34	1	5314,78917,04226
2	0	3010,29995,66398	35	1	5440,68044,35028
3	0	4771,21254,71966	36	1	5563,02500,76729
4	0	6020,59991,32796	37	1	5682,01724,06700
5	0	6989,70004,33602	38	1	5797,83596,61681
6	0	7781,51250,38364	39	1	5910,64607,02650
7	0	8450,98040,01426	40	1	6020,59991,32796
8	0	9030,89986,99194	41	1	6127,83856,71974
9	0	9542,42509,43932	2	1	6232,49290,39790
10	1	0000,00000,00000	43	1	6334,68455,57959
11	1	0413,92685,15823	4	1	6434,52676,48619
12	1	0791,81246,04762	45	1	6532,12513,77534
13	1	1139,43352,30684	6	1	6627,57831,68157
14	1	1461,28035,67824	47	1	6720,97857,93572
15	1	1760,91259,05568	8	1	6812,41237,37559
16	1	2041,19982,65592	49	1	6901,96080,,02851
17	1	2304,48921,37827	50	1	6989,70004,33602
18	1	2552,72505,10331	51	1	7075,70176,09794
19	1	2787,53600,95283	2	1	7160,03343,63480
20	1	3010,29995,66398	53	1	7242,75869,60079
21	1	3222,19294,73392	4	1	7323,93759,82297
22	1	3424,22680,82221	55	1	7403,62689,49424
23	1	3617,27836,01759	6	1	7481,88027,00620
24	1	3802,11241,71161	57	1	7558,74855,67249
25	1	3979,40008,67204	8	1	7634,27993,56294
26	1	4149,73347,97082	59	1	7708,52011,64214
27	1	4313,63764,15899	60	1	7781,51250,38364

Image from microfilm of copy in the British Museum.
https://commons.wikimedia.org/wiki/File:Logarithmorum_Chilias_Prima_page_0-67.jpg

Why logarithmic scale?

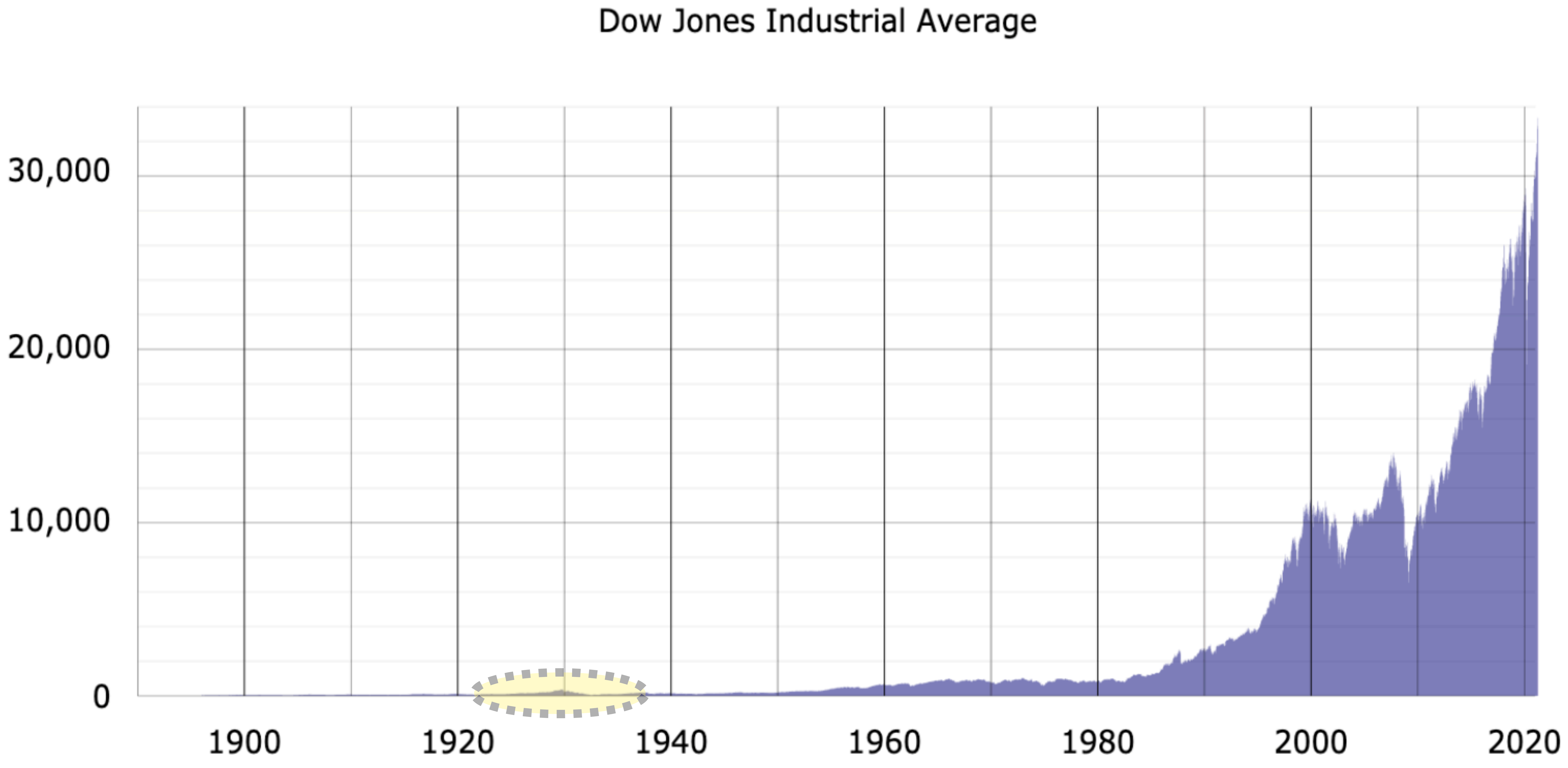
- Display data over a very wide range of values in a compact way
- Display relative change instead of absolute quantity of change

1929 Wall Street Crash

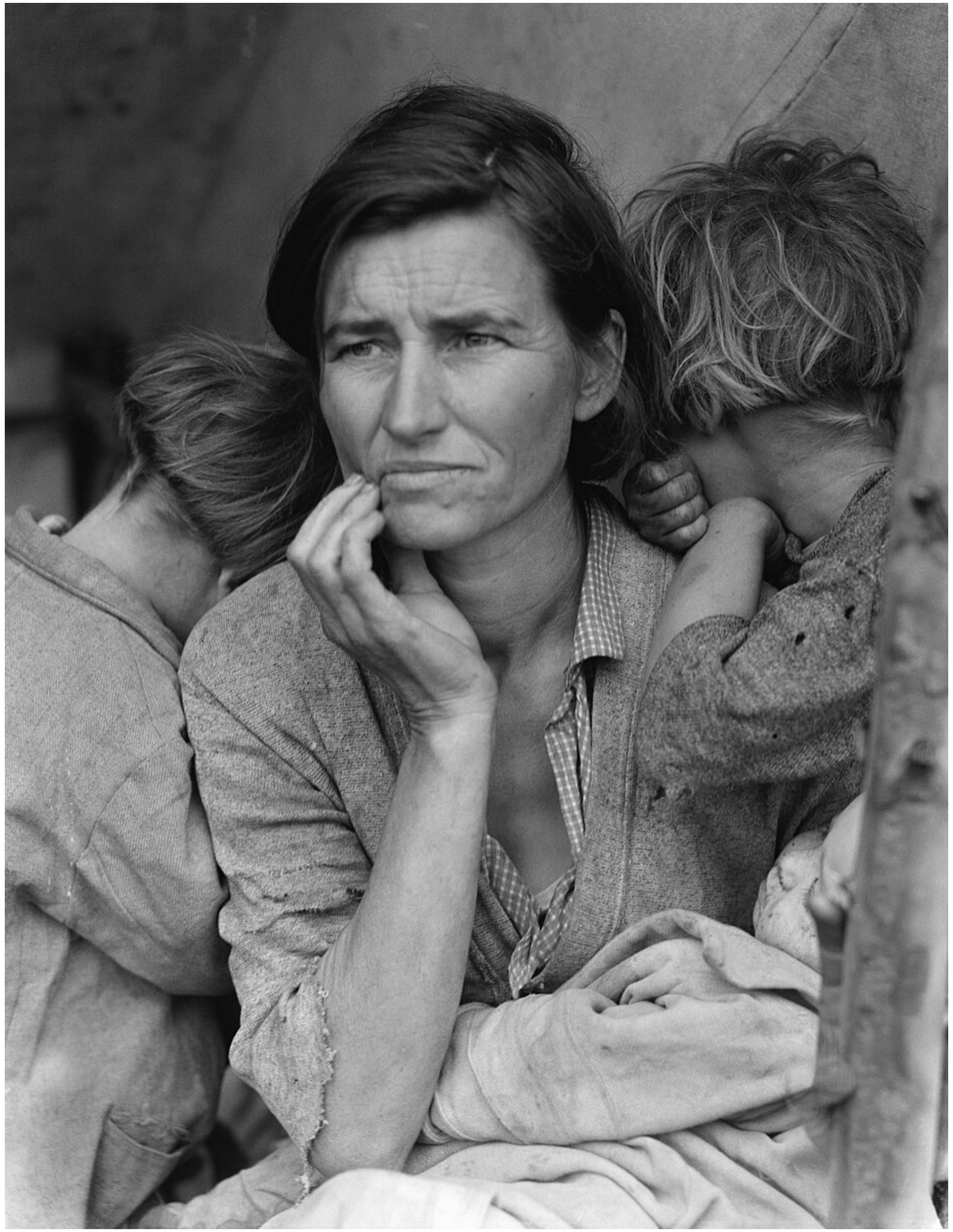


https://commons.wikimedia.org/wiki/File:1929_wall_street_crash_graph.svg

Dow Jones Industrial Average



https://vanderbei.princeton.edu/Economy/



Migrant mother, Nipomo, California by Dorothea Lange
https://en.wikipedia.org/wiki/File:Lange-MigrantMother02.jpg

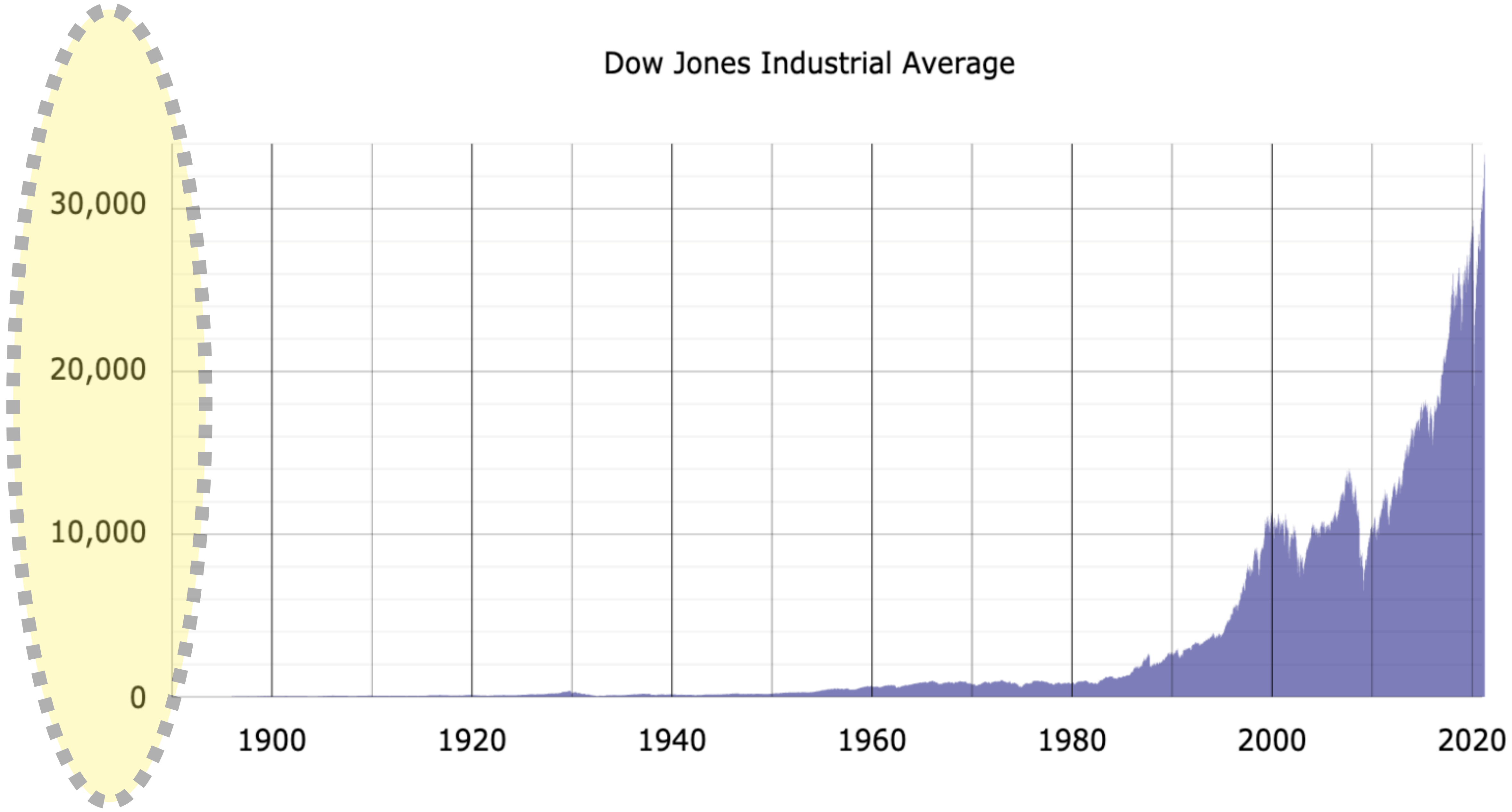
Dorothea Lange (1895-1965),

[Resettlement Administration](#) [photographer](#) with a 4×5 camera

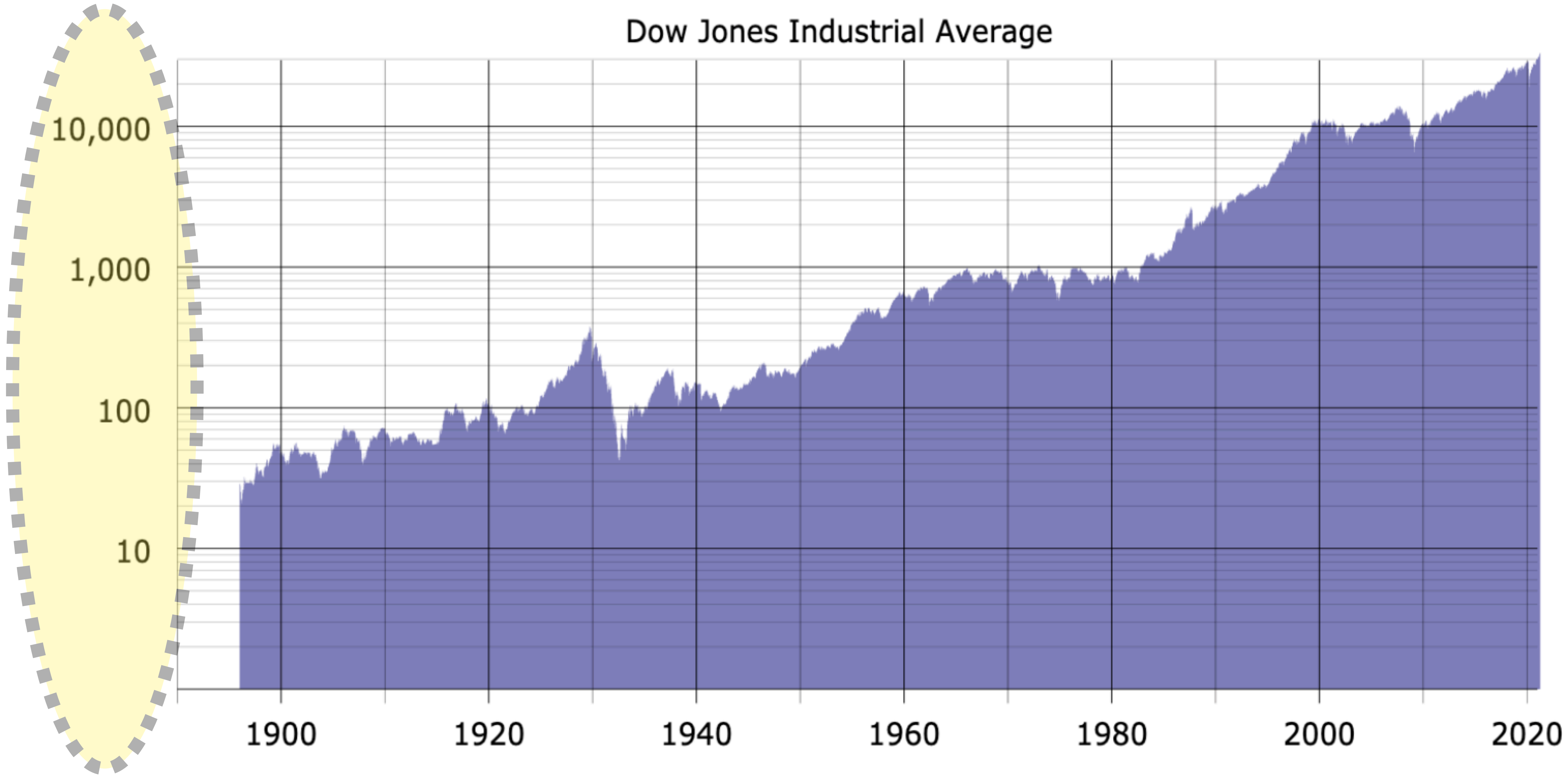
Photographer: Rondal Partridge

https://en.wikipedia.org/wiki/Dorothea_Lange#/media/File:Dorothea_Lange_atop_automobile_in_California_(restored).jpg

Spot the difference



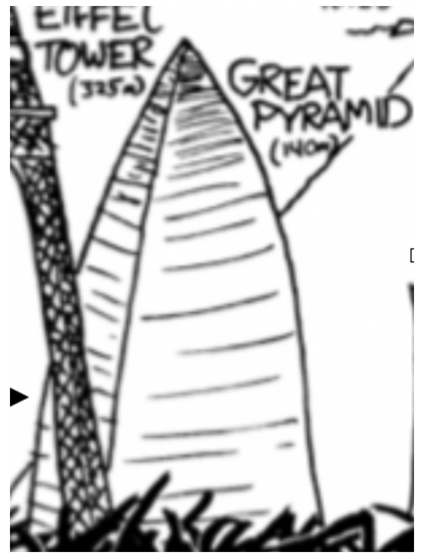
Linear

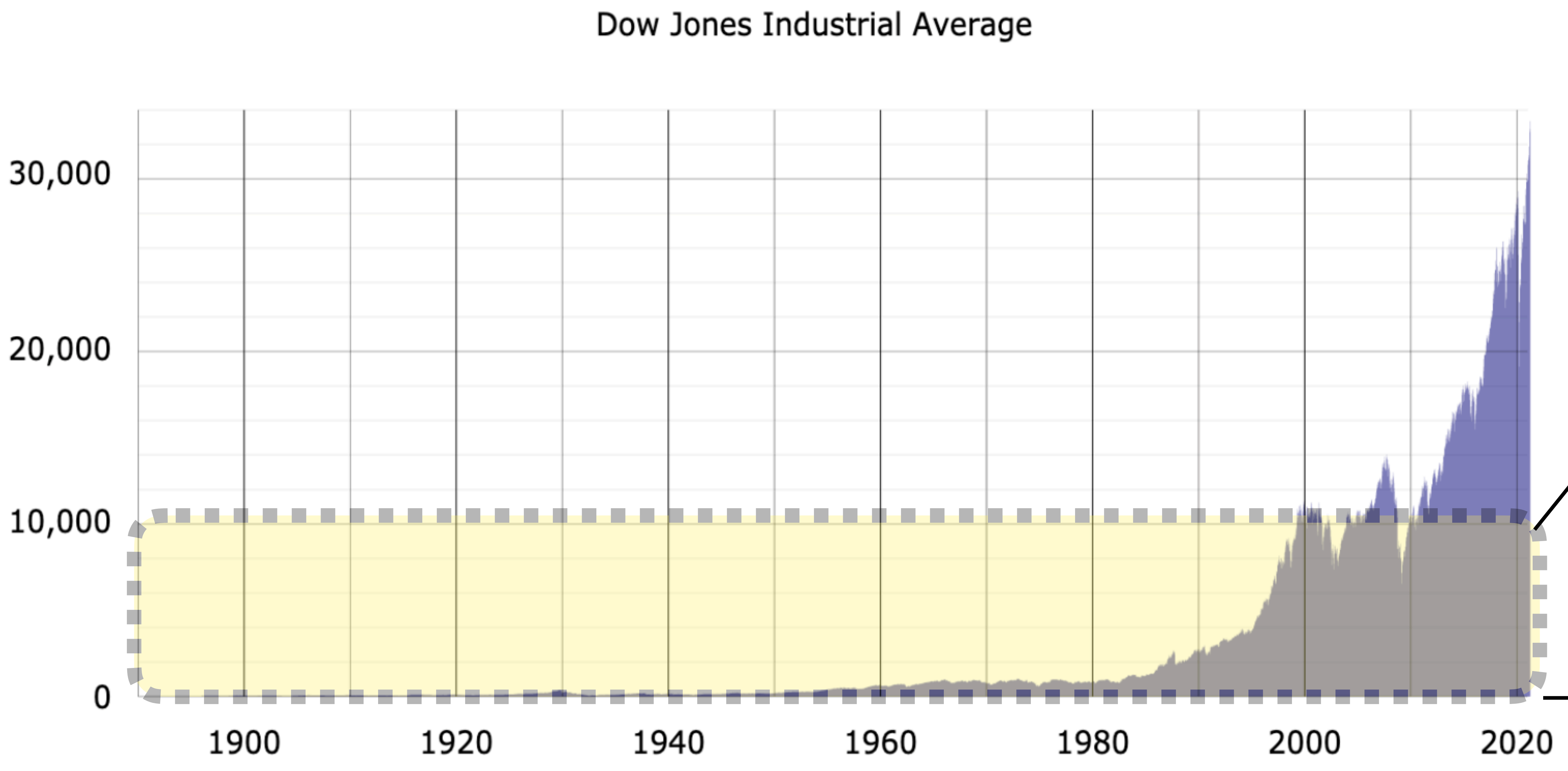


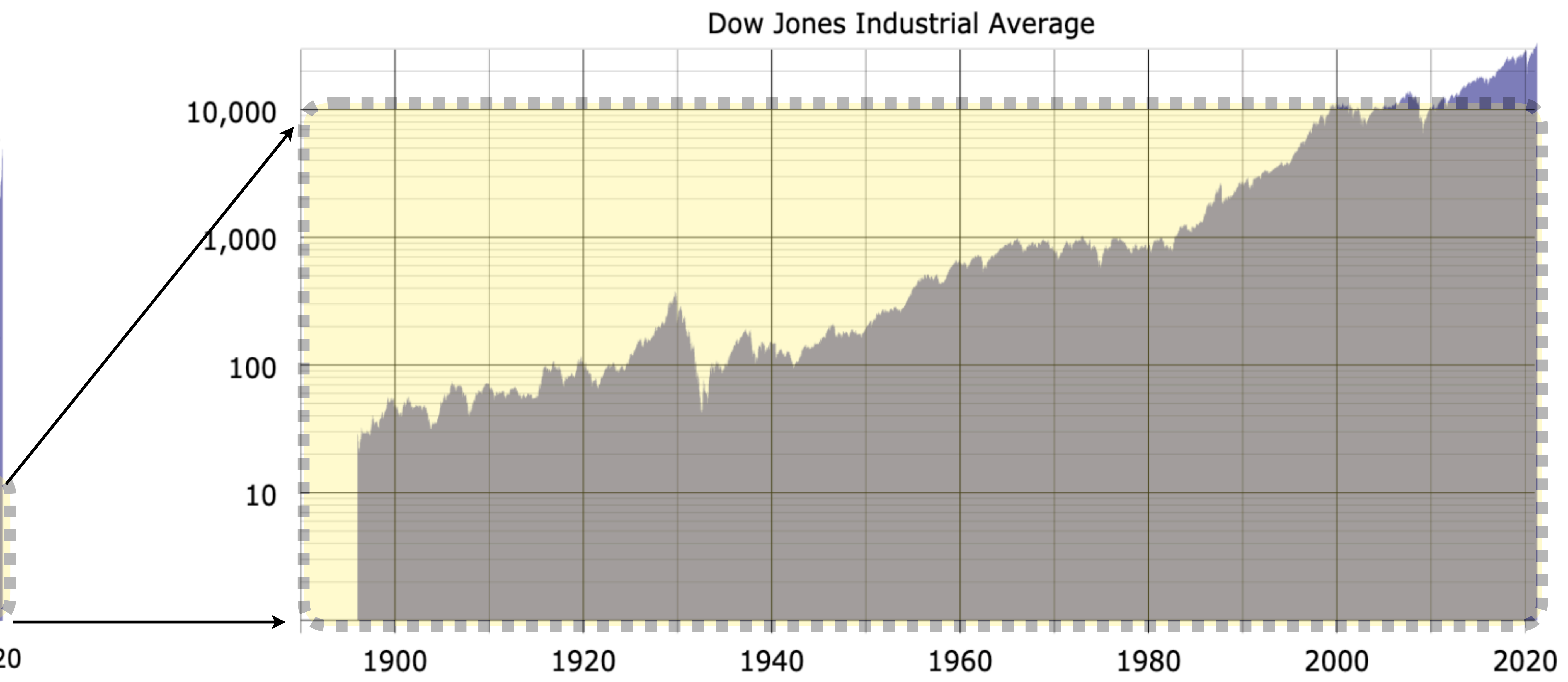
Log scale

https://vanderbei.princeton.edu/Economy/

Displaying relative change

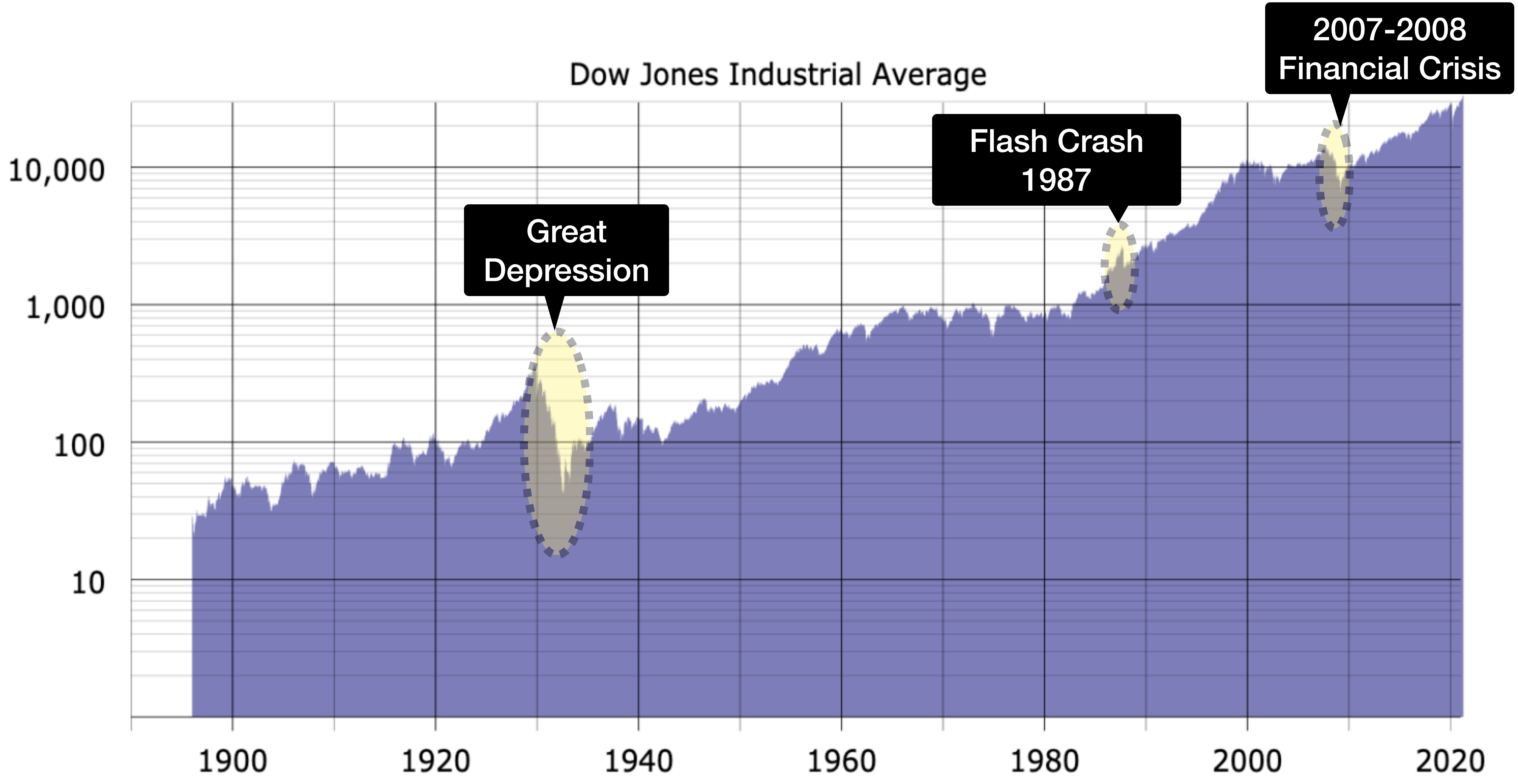






https://vanderbei.princeton.edu/Economy/

Dow Jones Industrial Average 1900-2021

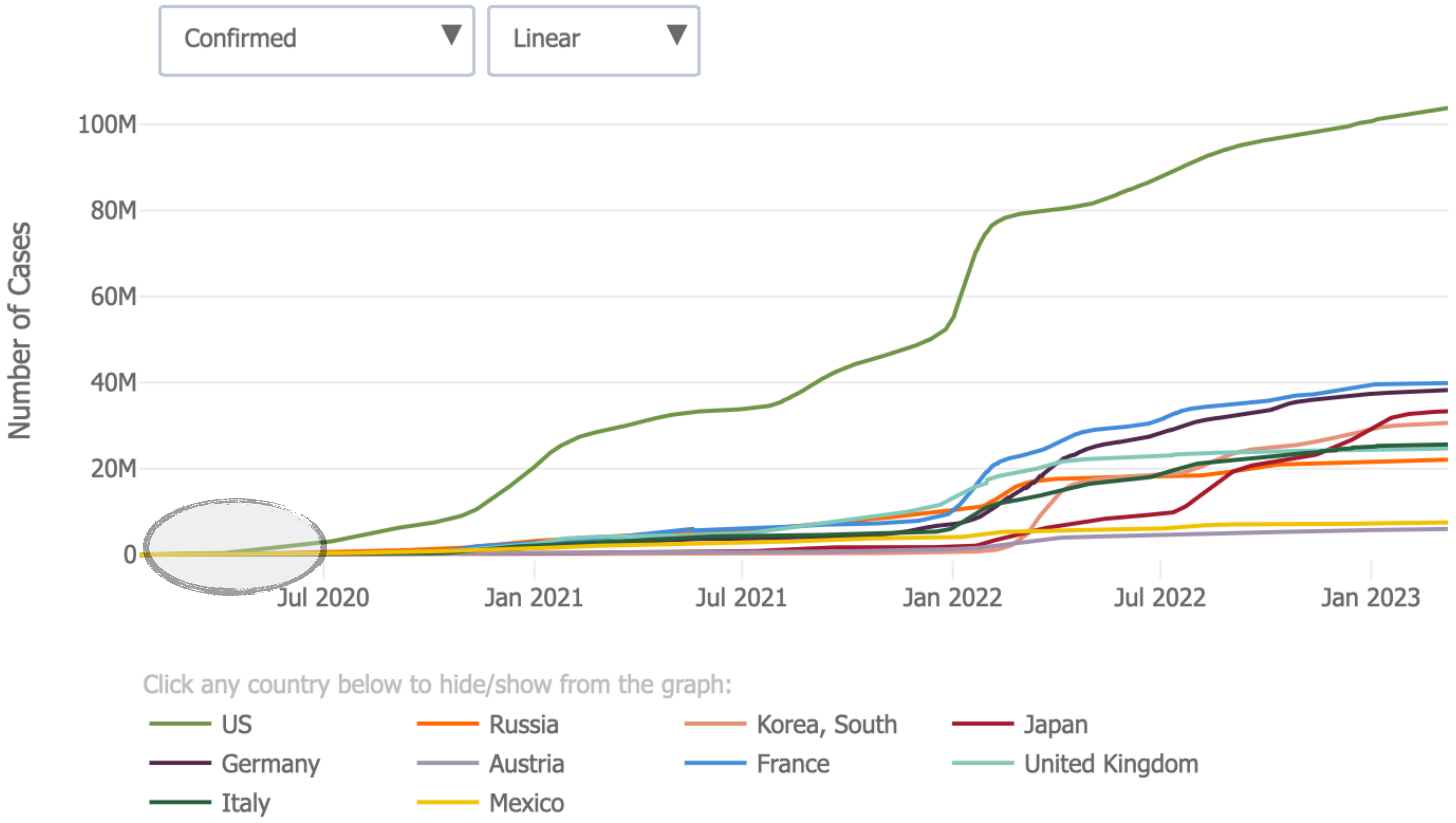


https://vanderbei.princeton.edu/Economy/

Why logarithmic scale?

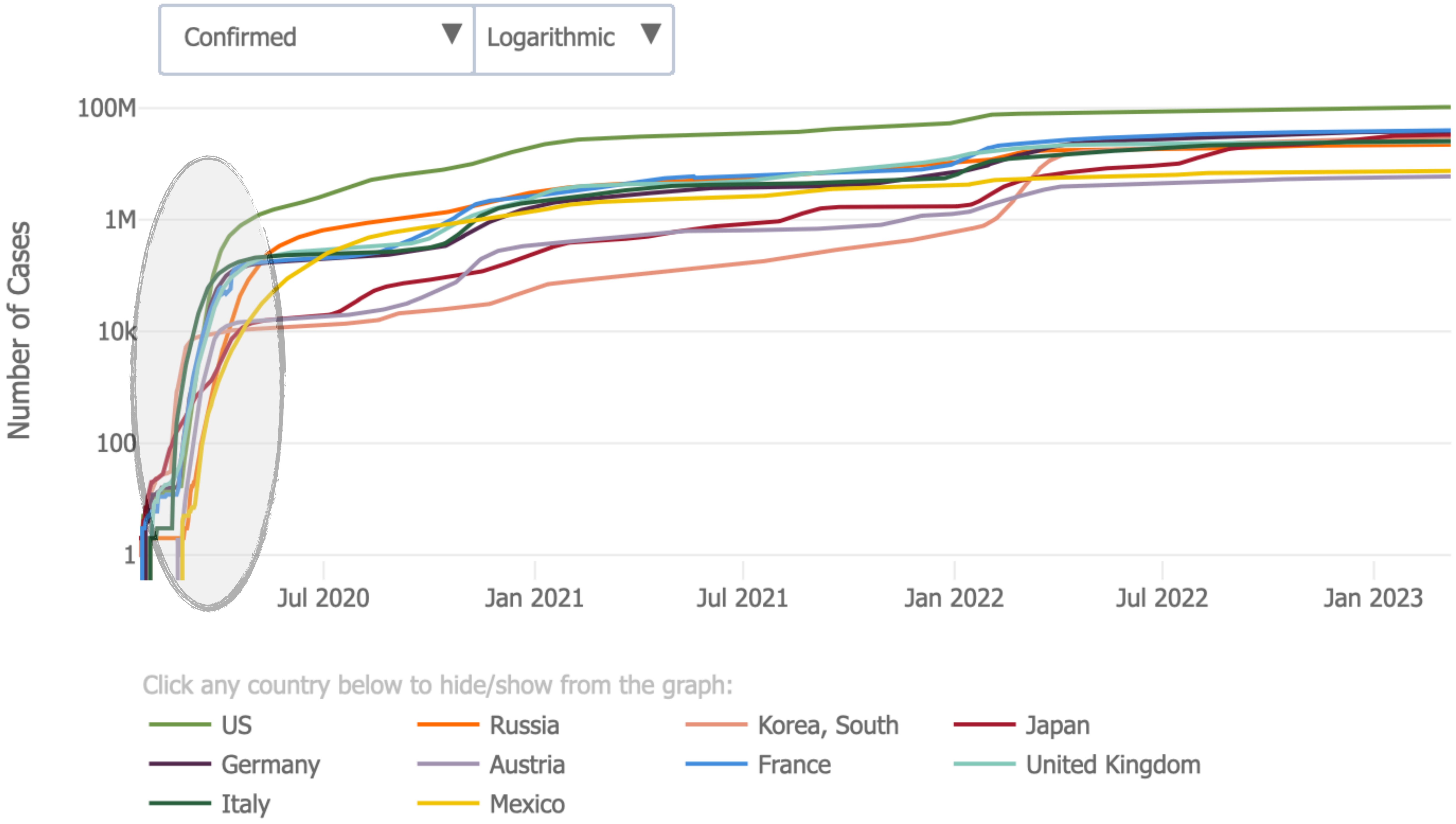
- Display data over a very wide range of values in a compact way
- Display relative change instead of absolute quantity of change
- Easier to identify trends like exponential growth

Confirmed Covid-19 cases



Johns Hopkins Coronovirus Research Center, https://coronavirus.jhu.edu/data/cumulative-cases

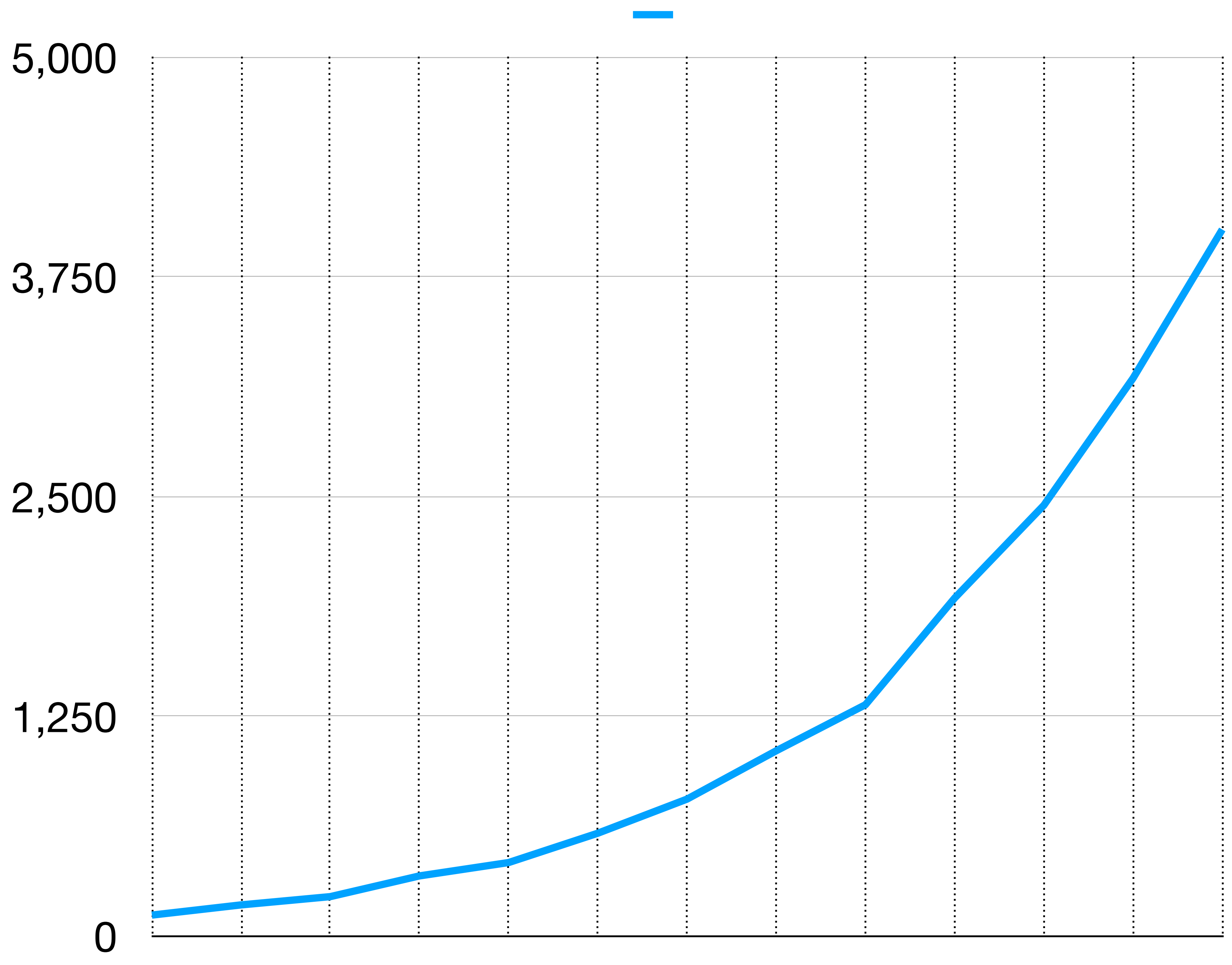
Confirmed Covid-19 cases: logarithmic



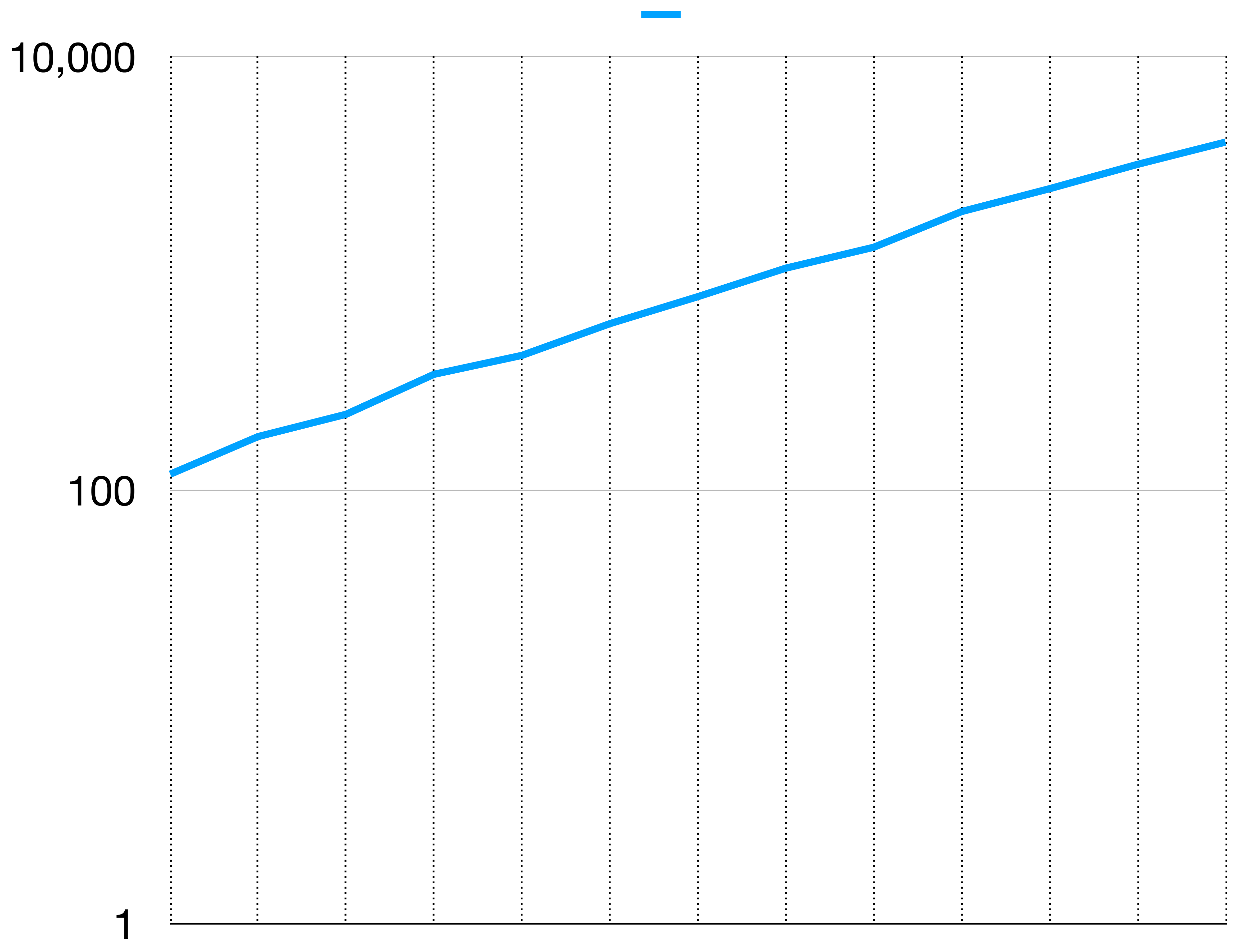
Johns Hopkins Coronovirus Research Center, https://coronavirus.jhu.edu/data/cumulative-cases

Cases in US: 2020-03-04 to 2020-03-16

Linear



Logarithmic



Date	Cases
2020-03-04	118
2020-03-05	176
2020-03-06	223
2020-03-07	341
2020-03-08	417
2020-03-09	584
2020-03-10	778
2020-03-11	1053
2020-03-12	1315
2020-03-13	1922
2020-03-14	2450
2020-03-15	3173
2020-03-16	4019

https://leancrew.com/all-this/2020/03/exponential-growth-and-log-scales/,

Latency (logarithmic heatmap)

Display data over a wide range of values in a compact way

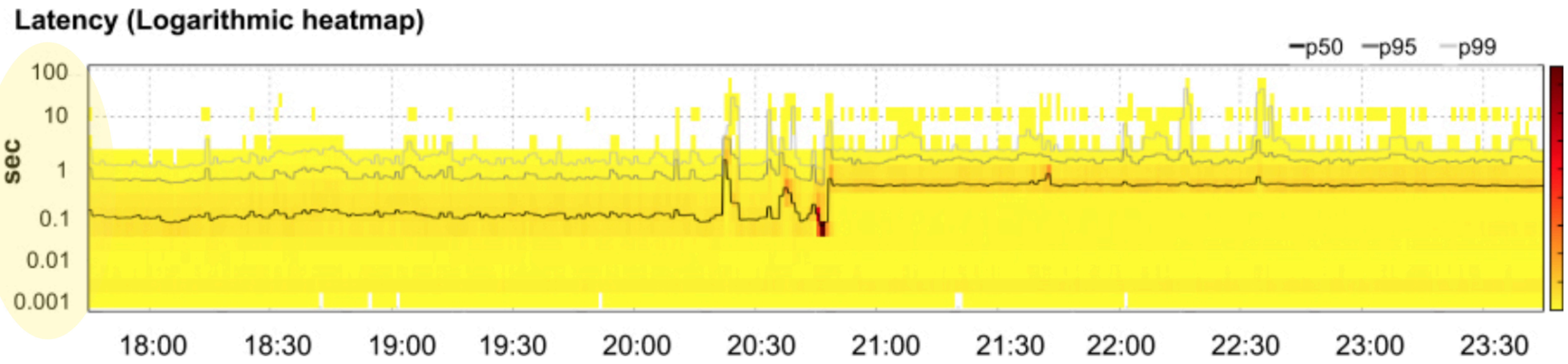


Figure 12-4. Application's latency, showing 50th, 95th, and 99th percentiles (lines) with a heatmap showing how many requests fell into a given latency bucket at any point in time (shade)

How do we describe the shape of data?

- Center of data

Describe the center of the data

Indices of central tendencies

1. Mean (average)

Mean (average)

$$\$80,000 + \$65,000 + \$135,000 + \$120,000 = \$400,000$$

Sum

count

$$\frac{\$400,000}{4} = \$100,000$$

Mean

Fulcrum by weight of gnomes




$80,000	$65,000	$135,000	$120,000

Mean (average)





$80,000	$65,000	$135,000	$120,000	$7,000,000,000

$$\frac{(\$400{,}000 + \$7{,}000{,}000{,}000)}{5} = \$1{,}400{,}080{,}000$$

Mean

“Then there is the man who drowned crossing a stream with an average depth of six inches.”

- W. I. E. Gates

Average limitations

- Averages hide the outliers.
- Outliers **skew** averages.
- Averages are not enough alone to describe the data!

https://orangematter.solarwinds.com/2016/11/18/why-percentiles-dont-work-the-way-you-think/

Describing the center of the data

Indices of central tendencies

1. Mean
2. Median

Median

Sort the data

Median

$65,000	$80,000	$120,000	$135,000	**$7,000,000,000**

Outlier

Fulcrum by number of gnomes

$65,000	$80,000	$135,000	**$7,000,000,000**

Median = ($80,000 + $135,000) / 2 = $107,500

NFL players salaries

Mean: $2,000,000

Median: $860,0000

https://work.chron.com/much-money-nfl-player-make-year-2377.html

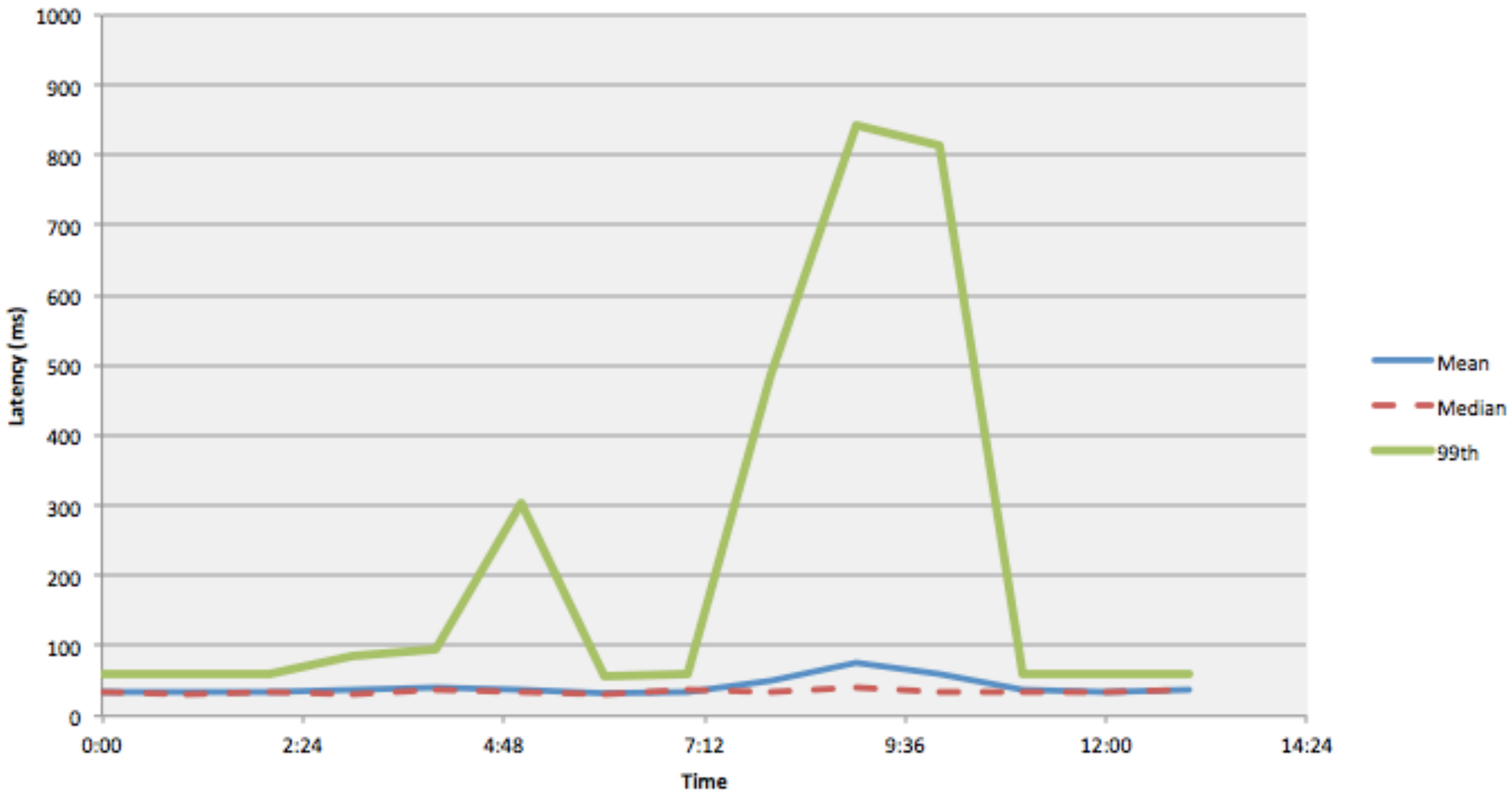
Summarizing the center of the data

Indices of central tendencies

1. Mean or average
2. Median
3. Mode - when it’s unique, this is the value that appears the most often in a data set

https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch11/mode/5214873-eng.htm#

Mean, median latency



https://www.elastic.co/blog/averages-can-dangerous-use-percentile

Average CPU usage

Cluster	Provider	Nodes	Average CPU usage past 6h	Max CPU usage (6h)	Average memory usage past 6h	Max memory usage ...
se-demo-cluster	gce	9	47%	49.65%	27%	27.82%
opentelemetry-sedemo	gce	3	8%	32.50%	28%	32.03%
opentelemetry-demo	gce	3	6%	7.27%	20%	21.79%

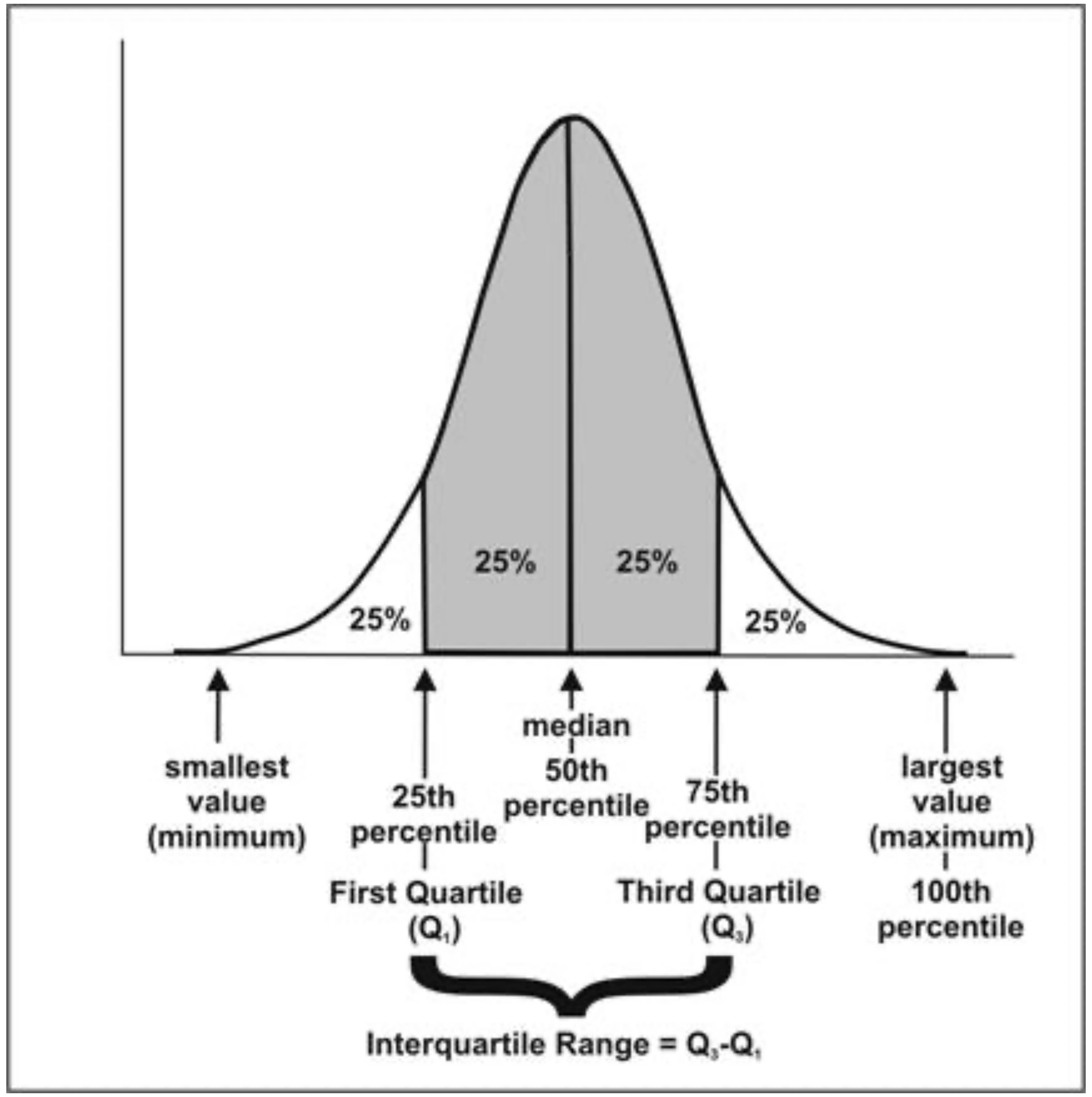
https://grafana.com/blog/2023/03/03/how-to-optimize-resource-utilization-with-kubernetes-monitoring-for-grafana-cloud/

How do we measure the spread in data?

https://www.statcan.gc.ca/en/wtc/data-literacy/catalogue/892000062020003

How do we measure the spread in data?

1. Range
   - max - min
2. Inter-quartile range (IQR)
   - Distance between 25th and 75th percentiles
3. Standard deviation



https://www.cdc.gov/csels/dsepd/ss1978/lesson2/section7.html#ALT27

Standard deviation

How concentrated the values are around the mean.

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

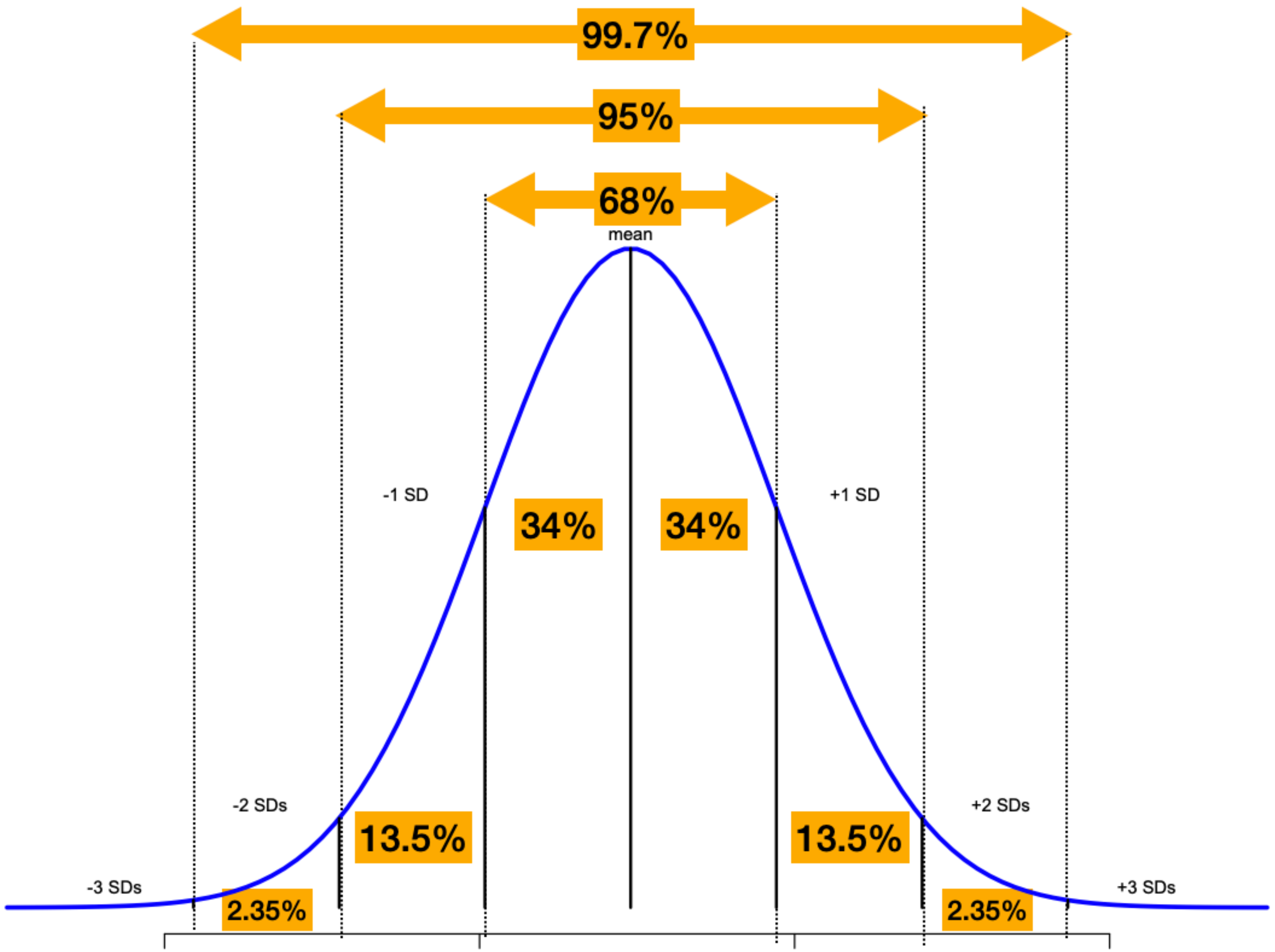
$\bar{x}$: mean

n: number of values

Variance is the square of standard deviation => s^2

Analogies from The Art of Statistics, David Spiegelhalter
Statistics for Dummies, Deborah J Rumsey, PhD

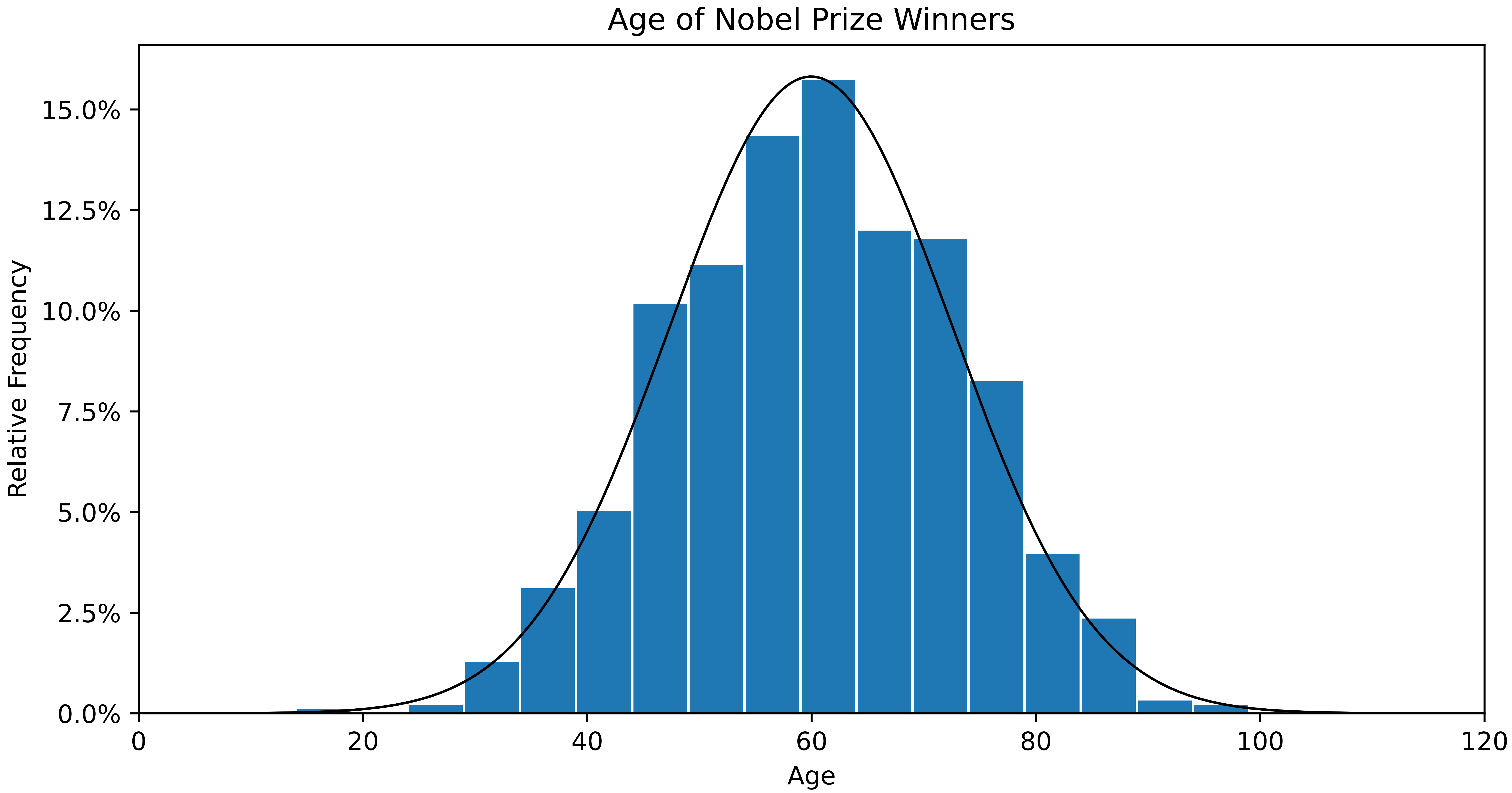
Empirical rule



Carl, as a gnome

Carl Friedrich Gauss
(1777 - 1855)

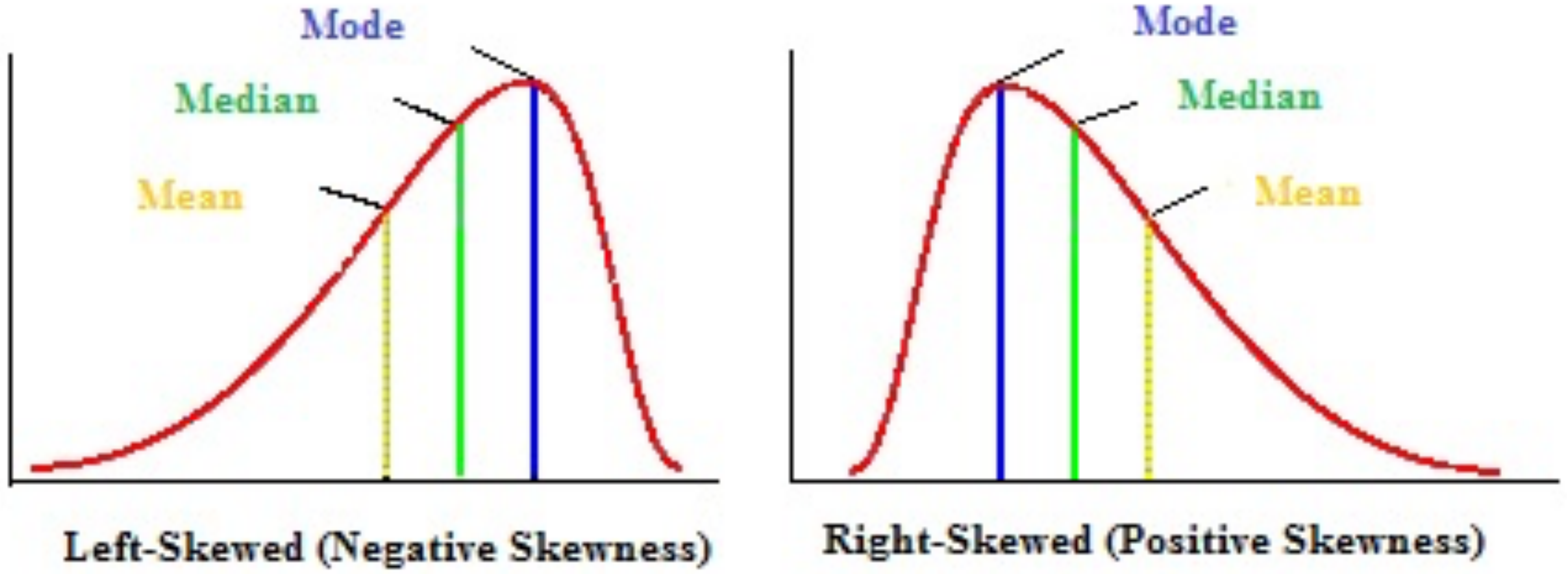
Age of Nobel Prize Winners

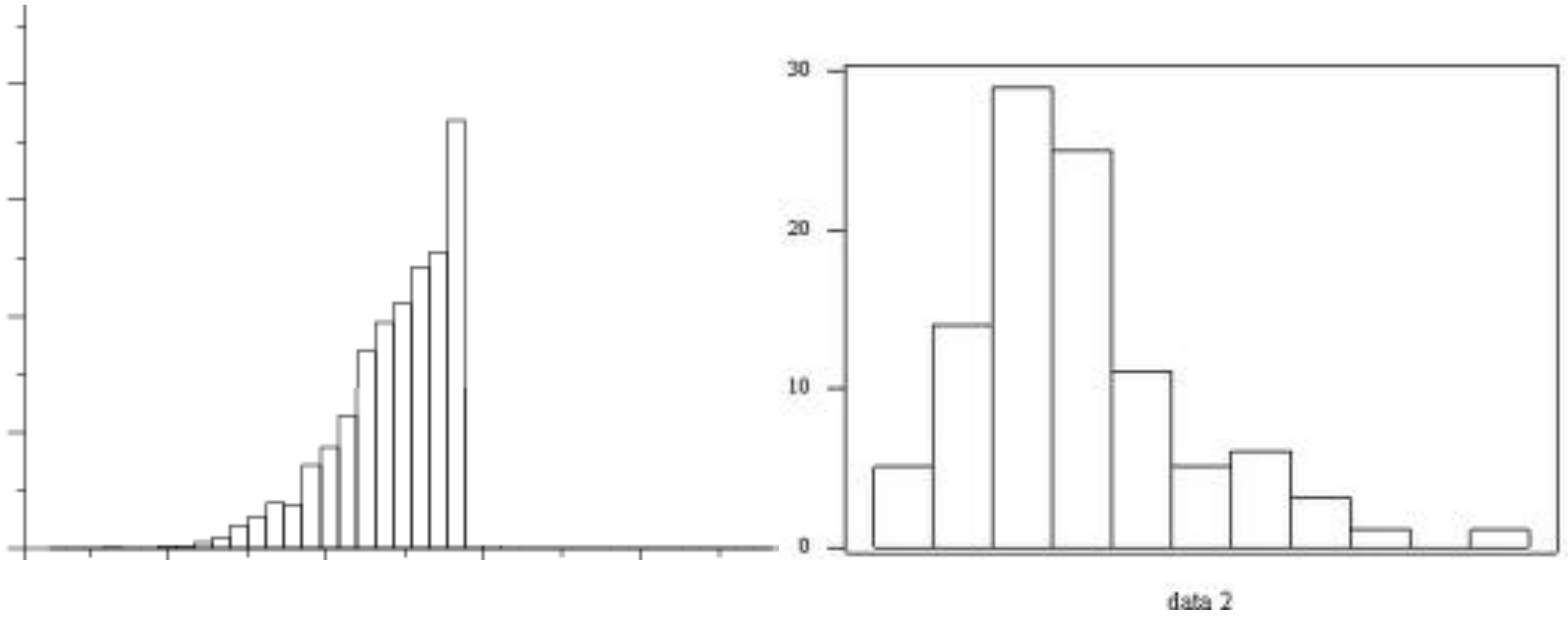


Histogram of the age of Nobel Prize winners when they won the prize and normal distribution fitted to the data.

https://www.w3schools.com/statistics/statistics_normal_distribution.php

Skewed distributions





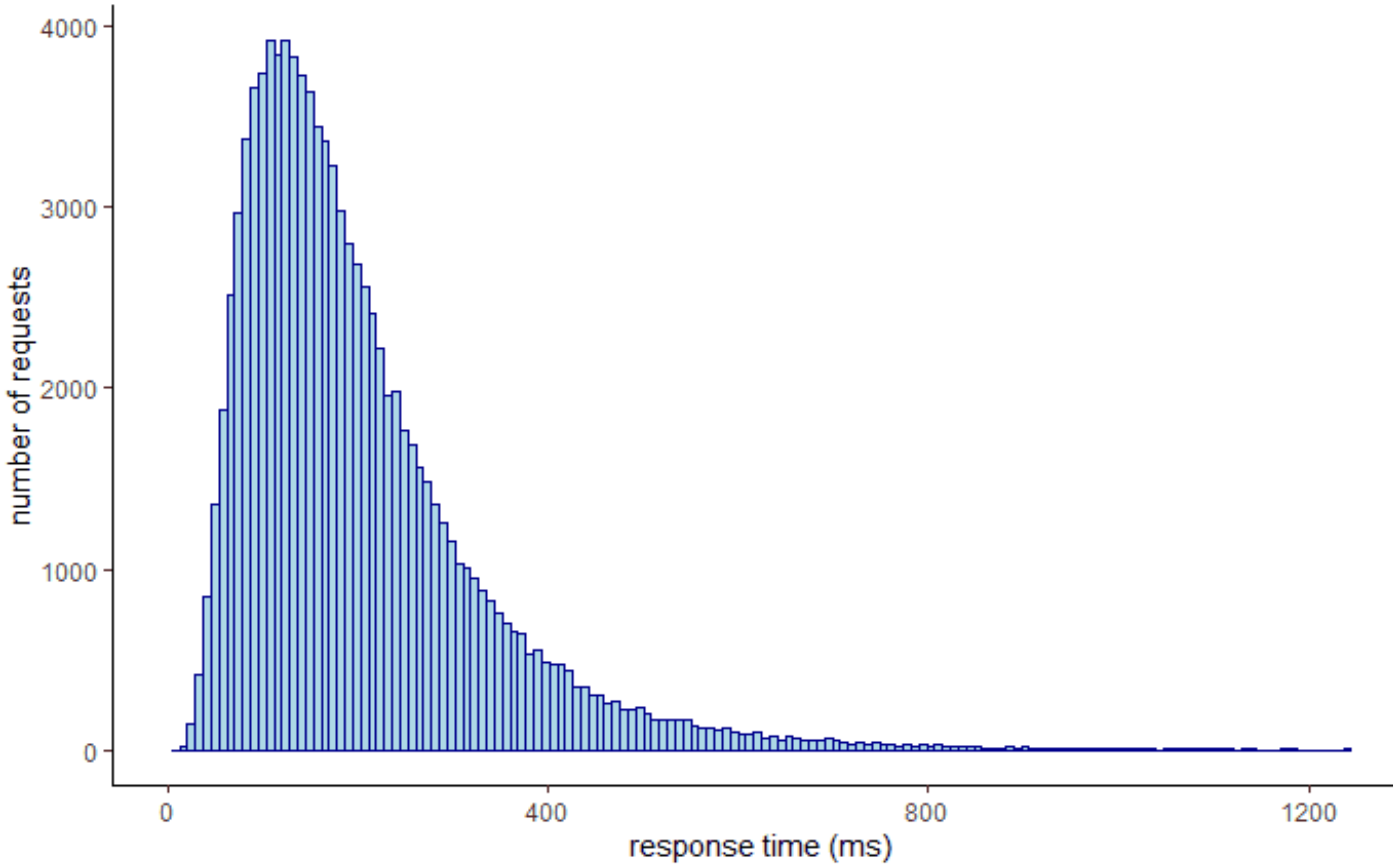
Stephanie Glen. "Skewed Distribution: Definition, Examples" From **StatisticsHowTo.com**: Elementary Statistics for the rest of us! https://www.statisticshowto.com/probability-and-statistics/skewed-distribution/

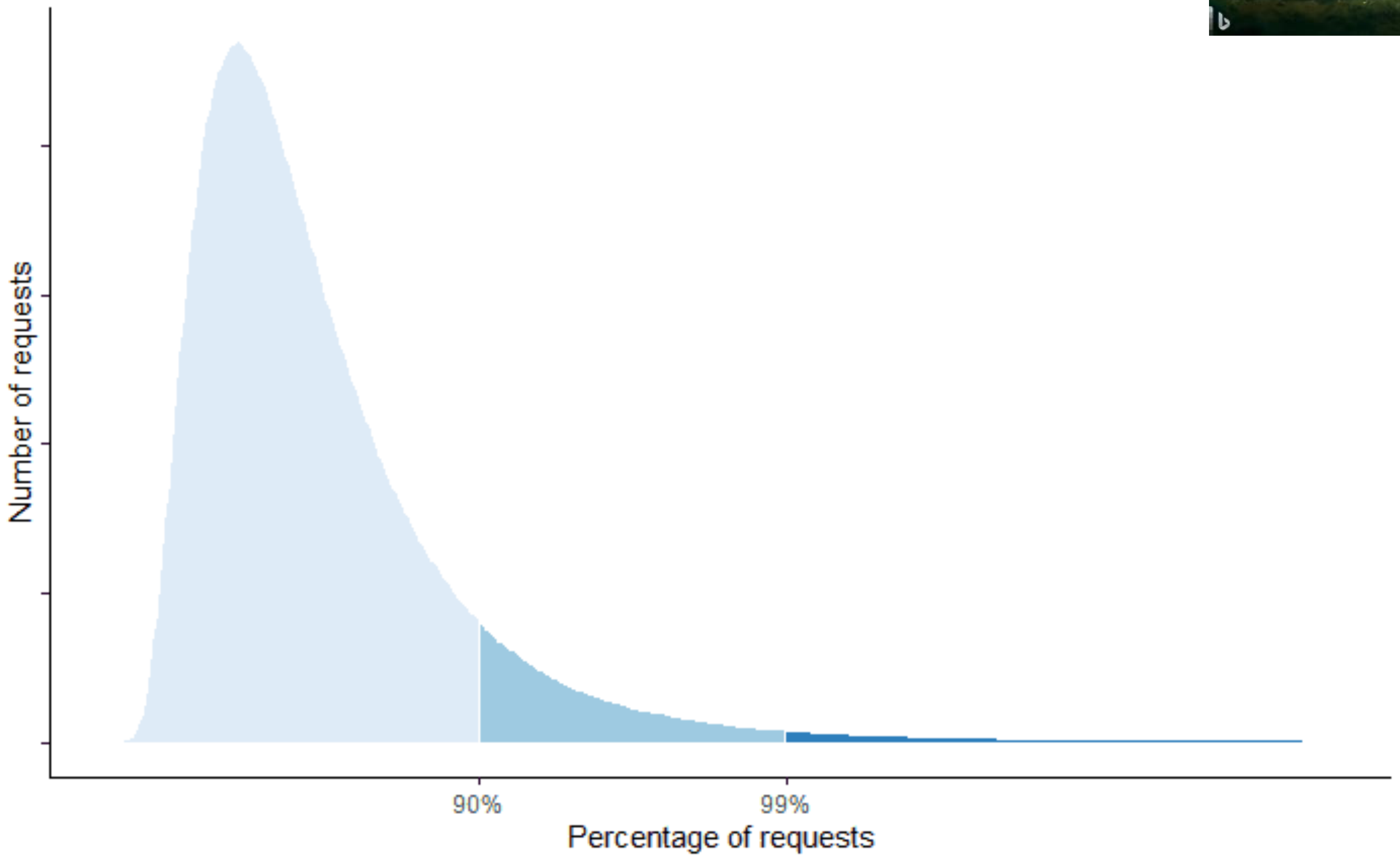
"We generally prefer to **work with percentiles rather than the mean** (arithmetic average) of a set of values.

Doing so makes it possible to **consider the long tail of data points**, which **often** have significantly **different** (and more interesting) **characteristics than** the **average**."

https://sre.google/sre-book/service-level-objectives/

Long tail of latency





https://robertovitillo.com/why-you-should-measure-tail-latencies/

How do we measure relative standing

Exam scores => Low mean (hard exam) vs high mean (easy exam)

Percentile analogy

- In a race of 100 people, p99 is the “time” for the 99th person to cross the finish line.

Vishwesh [redacted] · Aug 26, 2021

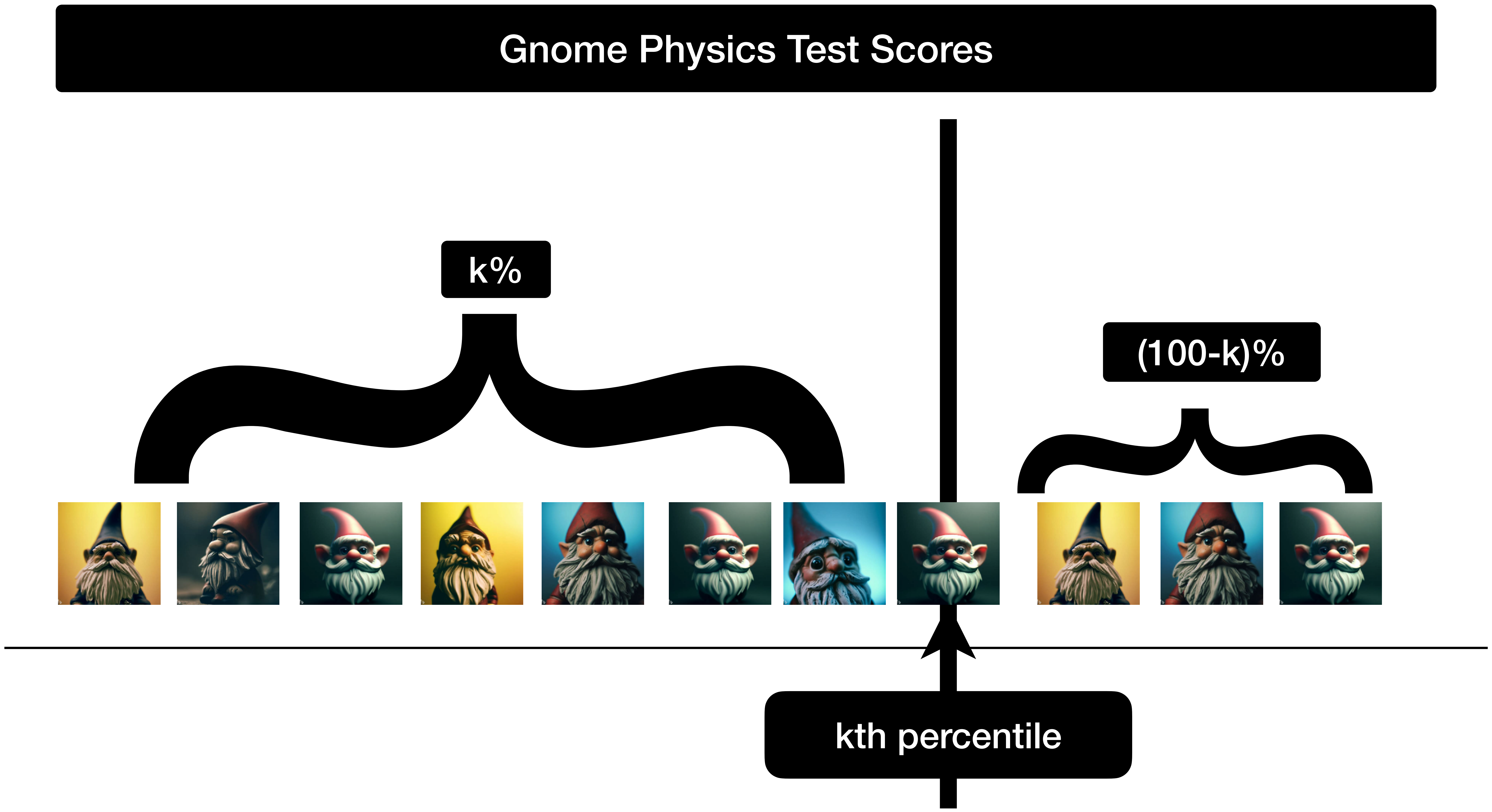
p99 = if 100 people running in a race, its time taken for 99th person to cross the finish line.

Apply the same thing to all other percentiles!

4

https://last9.io/blog/your-percentiles-are-incorrect-p99-of-the-times/

Percentile



Quantile

- 2-quantile: median (p50)
- 4-quantiles: quartiles
    - interquartile range, **midspread** or **middle fifty** → IQR = $Q_3 - Q_1$.
- 100-quantiles: percentiles

https://en.wikipedia.org/wiki/Quantile
Stephanie Glen. "Quantile: Definition and How to Find Them in Easy Steps" From **StatisticsHowTo.com**: Elementary Statistics for the rest of us!
https://www.statisticshowto.com/quantile-definition-find-easy-steps/

ϕ-quantile

Gnome Physics Test Scores

ϕ*100%

ϕ-quantile at
ϕ*Nth value

(100-ϕ)*100%

https://en.wikipedia.org/wiki/Quantile

How to calculate percentiles?

Naive implementation to find p90

$0.9 * 13 = 11.7$ => 12

1	2	3	4	5	6	7	8	9	10	11	12	13
20	34	34	35	40	40	41	54	55	71	73	77	79

10% of scores are higher than 77.

P90
(0.9-quantile)

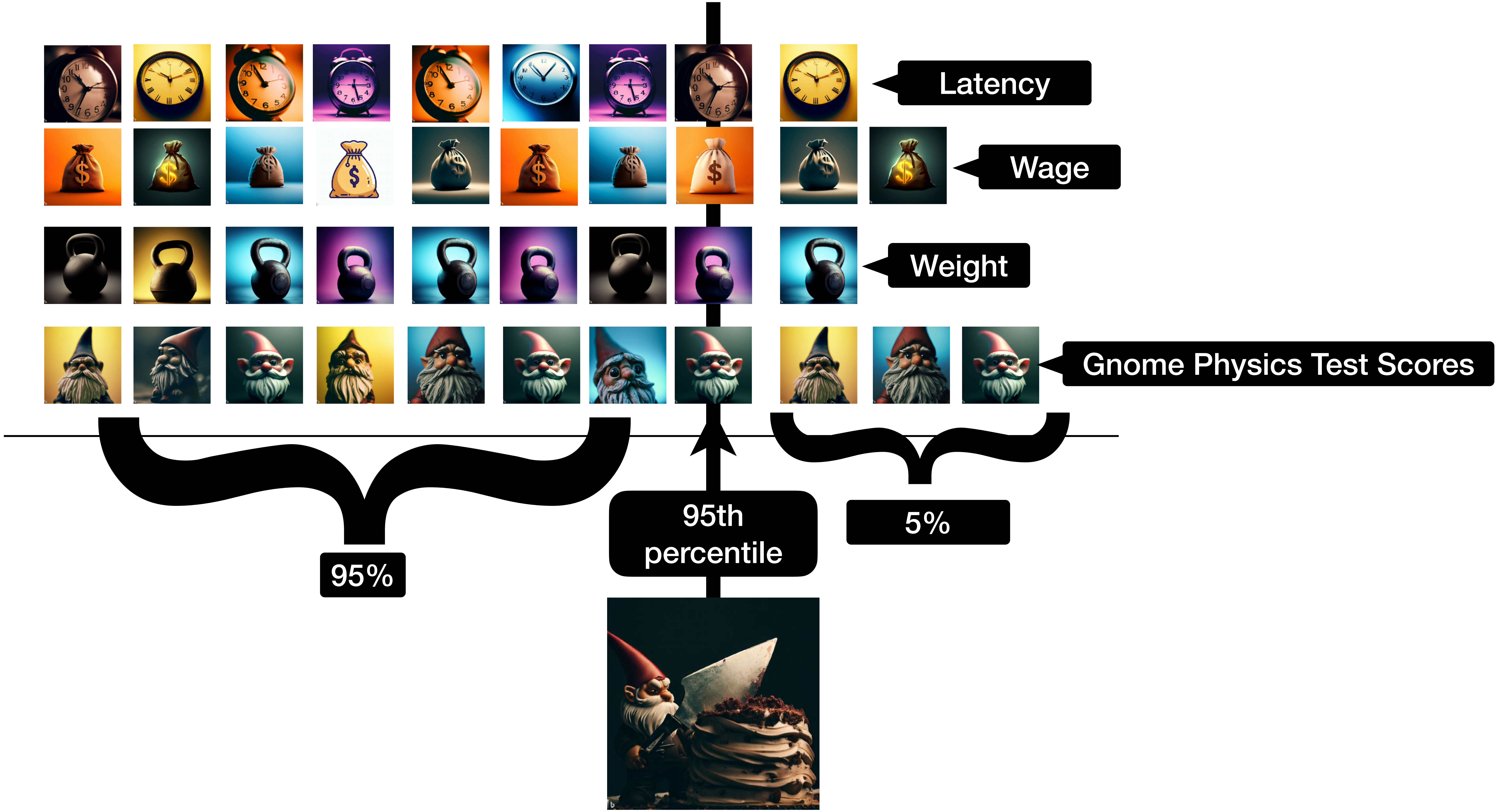
90% of scores are less than 77.

U.S. household income dispersion

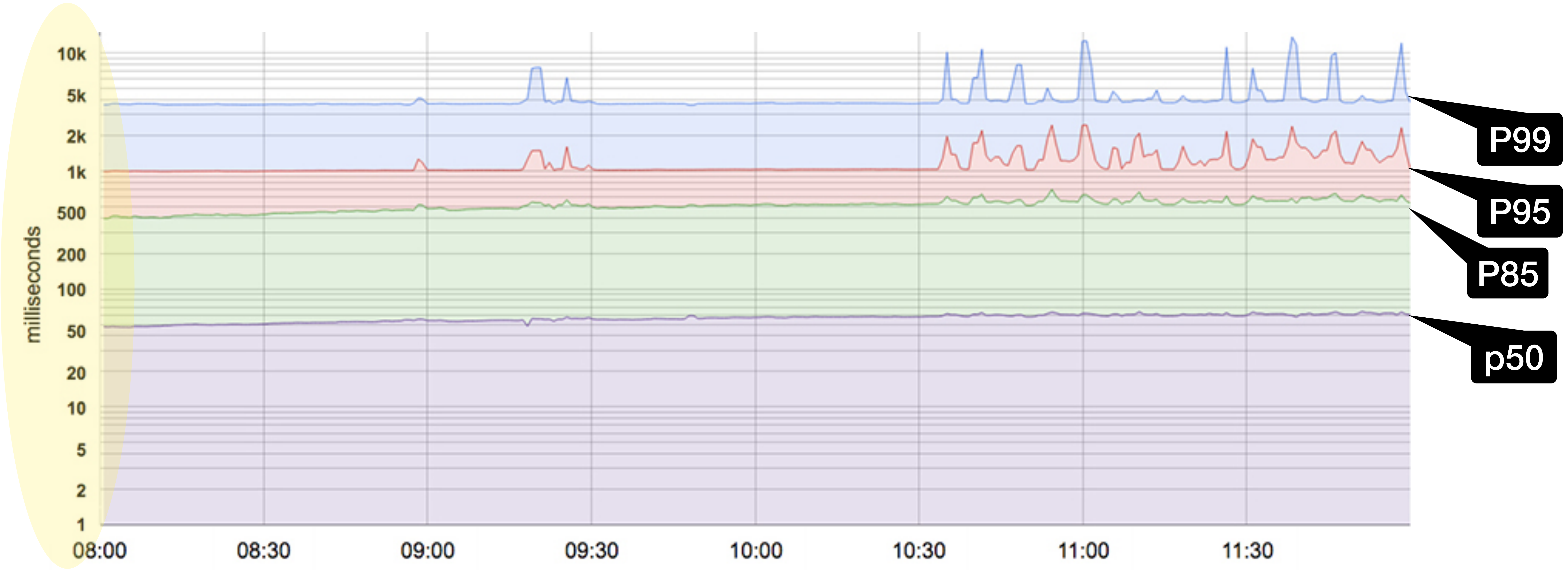
Year	Measures of income dispersion												
	Household income at selected percentiles										Household income ratios at selected percentiles		
	10th percentile limit	20th percentile limit	30th percentile limit	40th percentile limit	50th (median)	60th percentile limit	70th percentile limit	80th percentile limit	90th percentile limit	95th percentile limit	90th/10th	90th/50th	50th/10th
2021	15,660	28,007	40,524	55,000	70,784	89,744	113,210	149,131	211,956	286,304	13.53	2.99	4.52
2020[1]	16,386	28,544	41,763	55,044	71,186	89,534	113,519	148,620	211,438	287,841	12.90	2.97	4.34
2019	16,984	29,762	42,815	56,700	72,808	91,657	116,267	151,017	213,171	286,138	12.55	2.93	4.29
2018	15,784	27,621	39,924	53,948	68,168	85,823	108,071	140,265	198,844	268,368	12.60	2.92	4.32

Table A-4a at https://www.census.gov/content/dam/Census/library/publications/2022/demo/p60-276.pdf

Percentiles have universal interpretation



p50, p85, p95, p99 latency



50th, 85th, 95th, and 99th percentile latencies https://sre.google/sre-book/service-level-objectives/, ([CC BY-NC-ND 4.0]())

How do we calculate percentiles?

To calculate percentiles across potentially billions of values

approximate percentiles are calculated

https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-metrics-percentile-aggregation.html#search-aggregations-metrics-percentile-aggregation-approximation
https://www.cs.wustl.edu/~jain/papers/ftp/psqr.pdf

Approximate percentiles from a stream

Compressed data structures:
t-digest, Greenwald-Khanna,
HDRHistogram, KLL

Approximate
kth percentile
at time t

Key take aways

- Logarithmic scale graphs display wide range of data in a compact way
- Observe the long tail of latency with percentiles
- Averages get skewed by outliers
- Percentiles are most of the time approximated

Thank You

Lerna Ekmekcioglu
https://www.linkedin.com/in/lerna/

References

- Statistics for Dummies, 2nd Edition, Deborah J Rumsey, PhD
- The Art of Statistics, How to Learn from Data, David Spiegelhalter
- Frequency Trails: Modes and Modality, Brendan Gregg
- Tail Latency Might Matter More Than You Think, Marc Brooker
- The t-digest: Efficient Estimates of Distributions, Ted Dunning