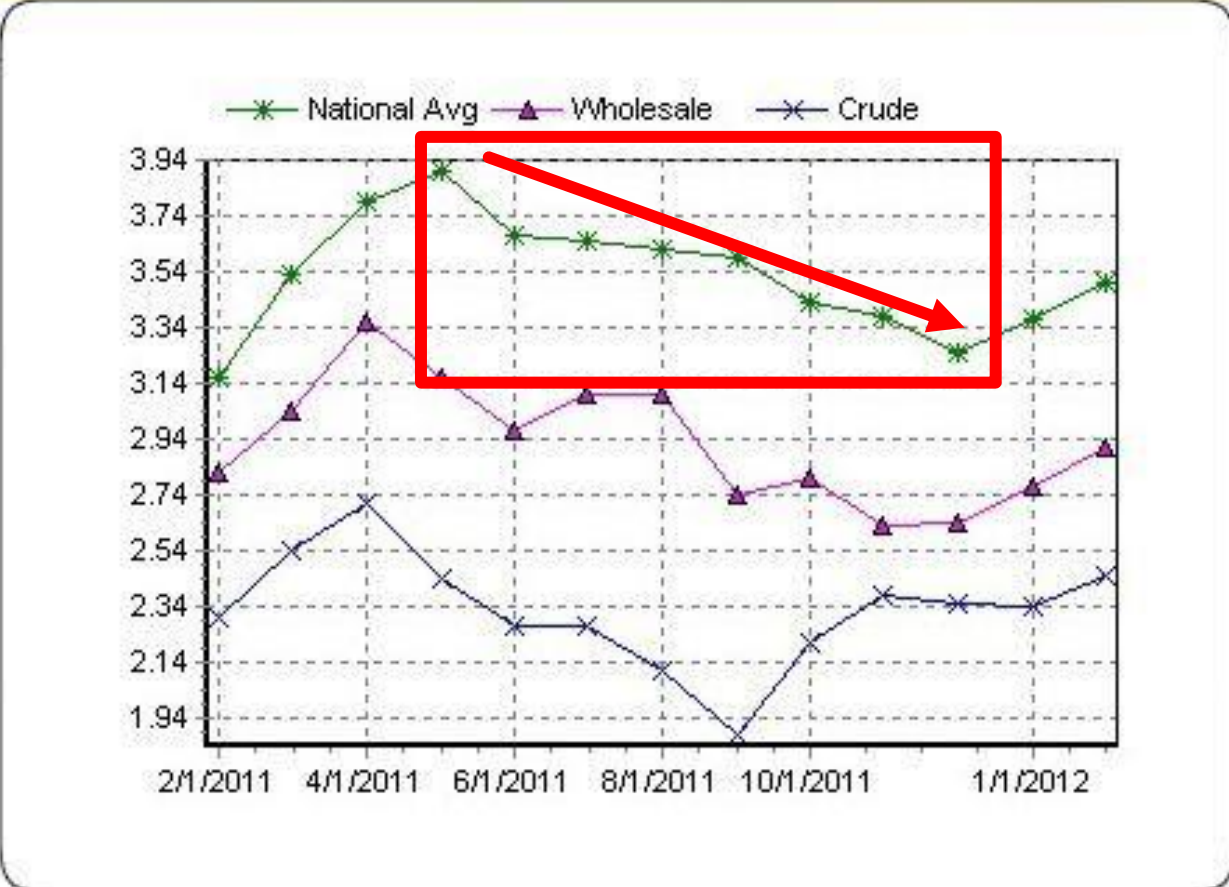

HALL-OF-FAME / HALL-OF-SHAME

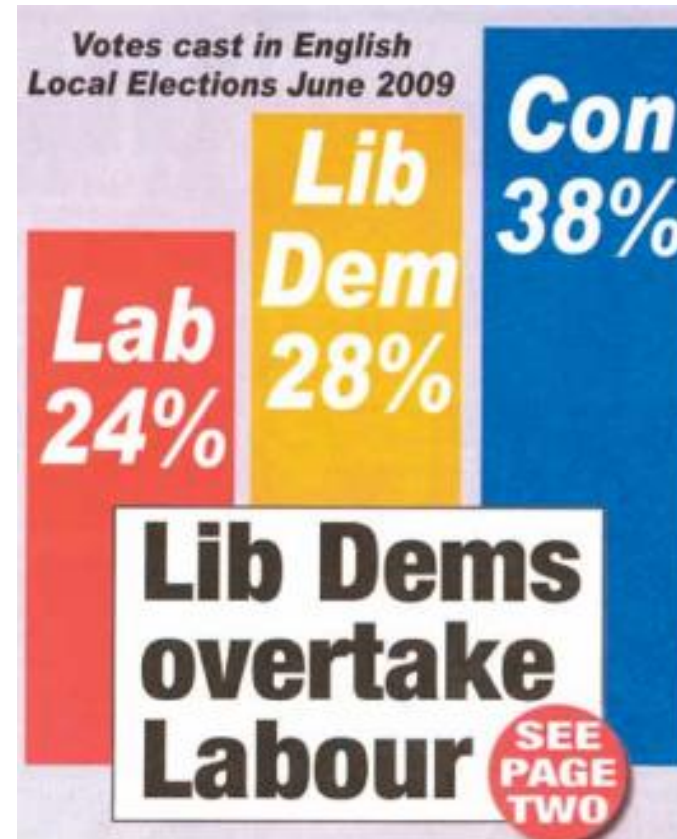
MISLEADING CHARTS



12 Month Average for Self-Serve Regular



MISLEADING CHARTS



MISLEADING CHARTS

Problems: disingenuous, selective and/or incompetent reporting

Solutions:

- Consistent scales and units of comparison
- Full time series
- No cherry picking the data range
- Cutting off -axis will exaggerate some effects
- Numbers must add up

WHAT TO WATCH FOR

Some methods yield visually striking, yet misleading, charts.

Be on the lookout for:

- tampering with axes and linear scales
- scaling effects, when representing data points as shapes or volumes
- cherry-picking by omitting certain data points

For low-dimensional datasets, a **tabular display** may provide as much information and be less likely to mislead.

YOU BE THE JUDGE

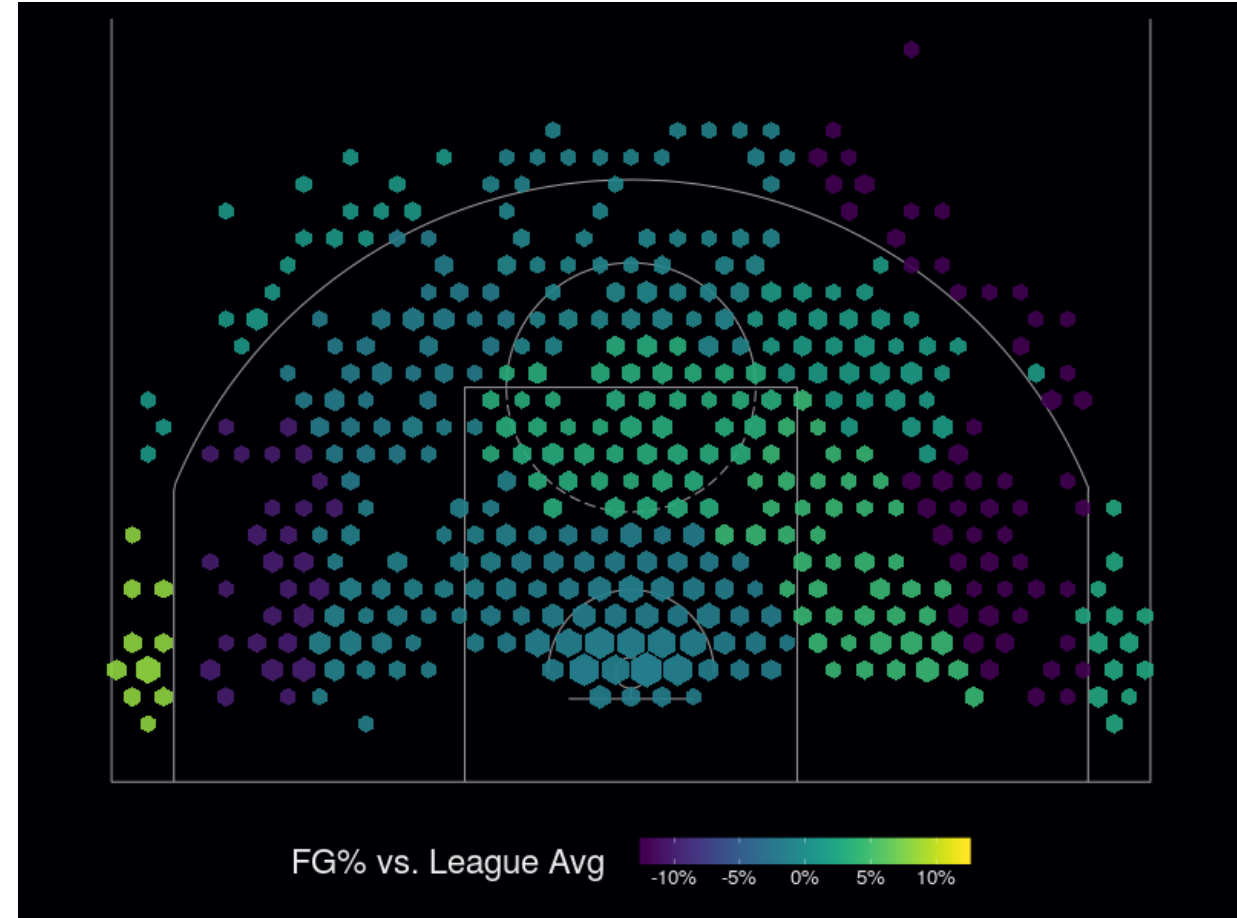
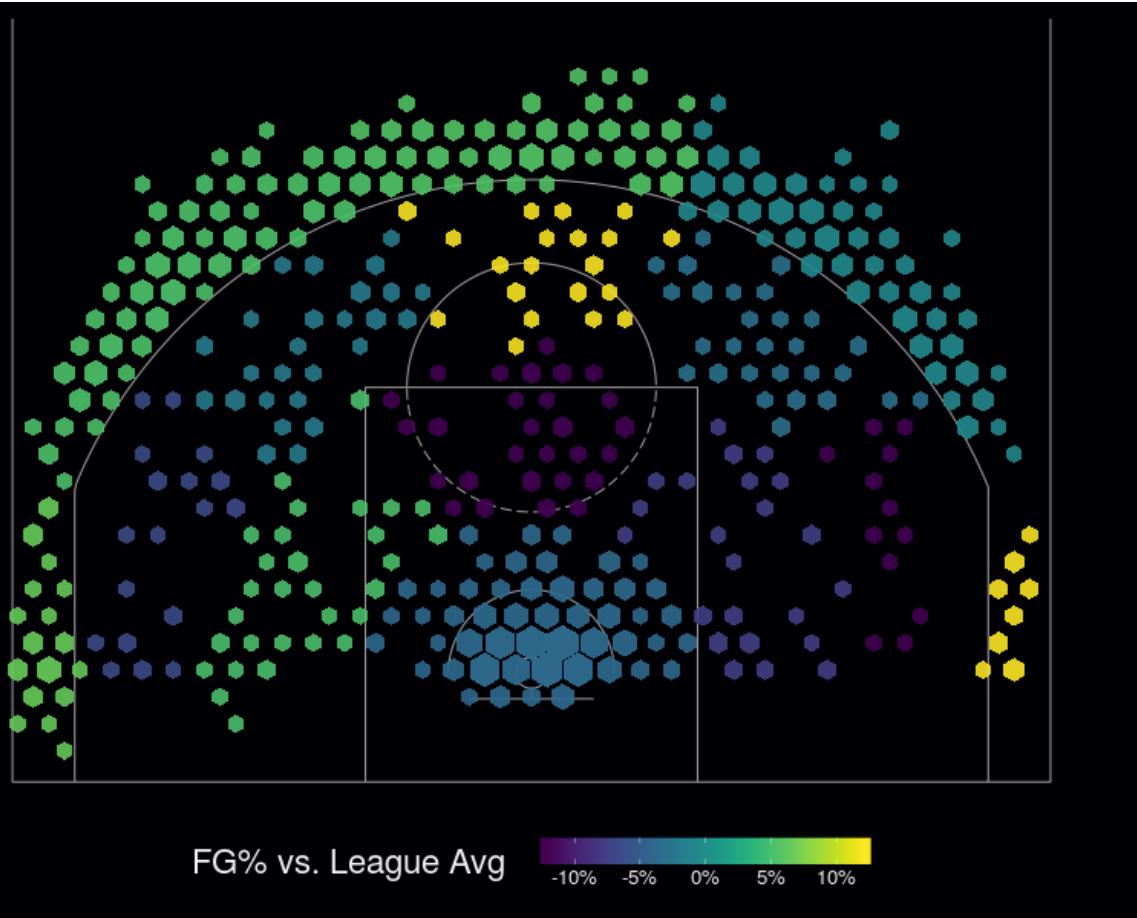
Some of the following are (arguably) good visualizations. But some are not!

Which are which? You be the judge...

NBA FG% Against League Average ('15-'16)

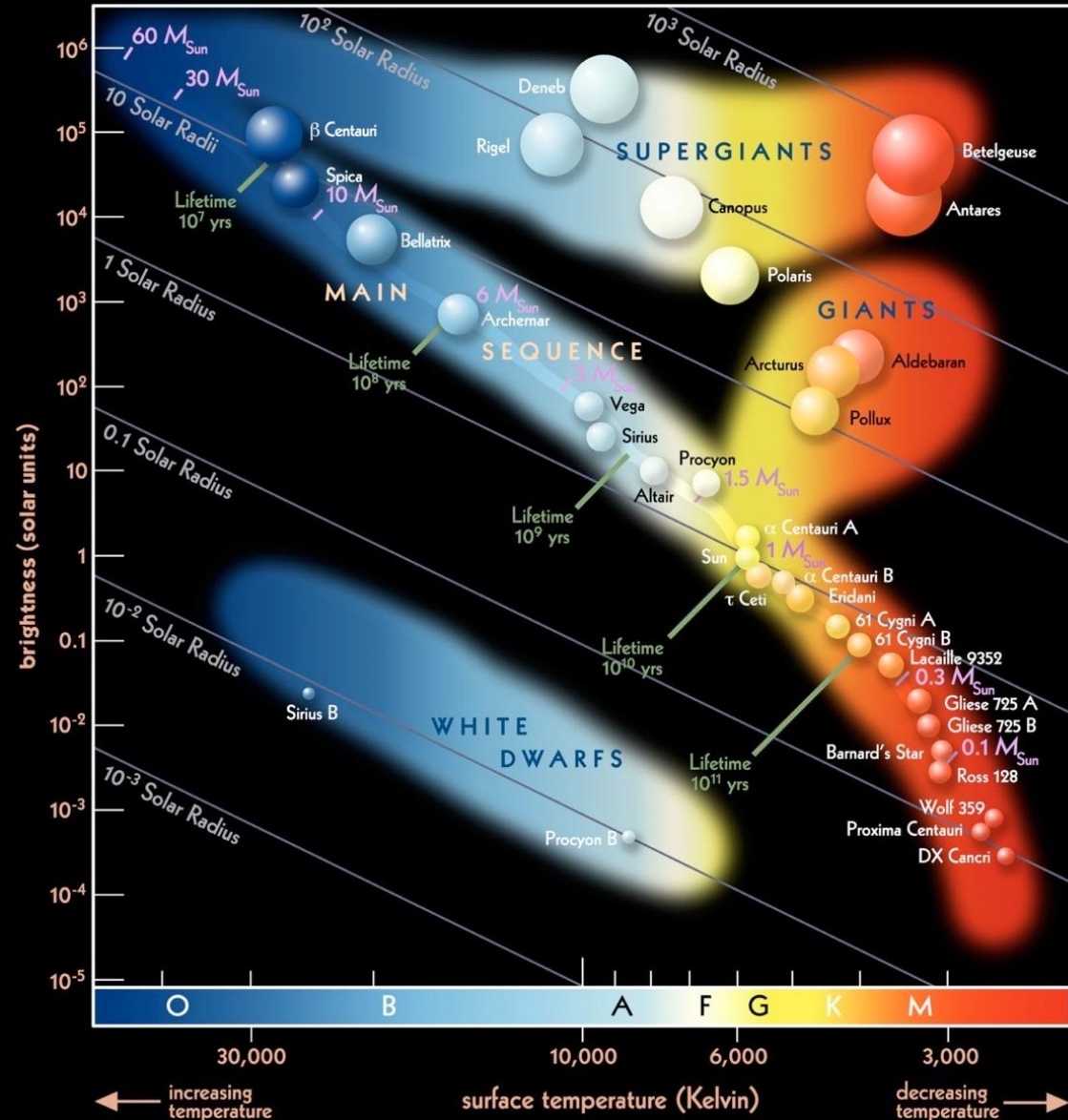
Kyle Lowry

DeMar DeRozan



What comparisons can you make? Do you understand the encoding? The context?

Hertzprung-Russell Diagram



Data Elements

- star radius (x 2)
- surface temperature (x 2)
- spectral class
- brightness
- mass
- lifetime
- name

Underlying Structure

- 4 clusters/group
- lifetime, mass and radius are related to brightness and surface temperature on the Main Sequence

Only a subset of all the stars is shown in the HR diagram.

A Model of Breast Cancer Causation

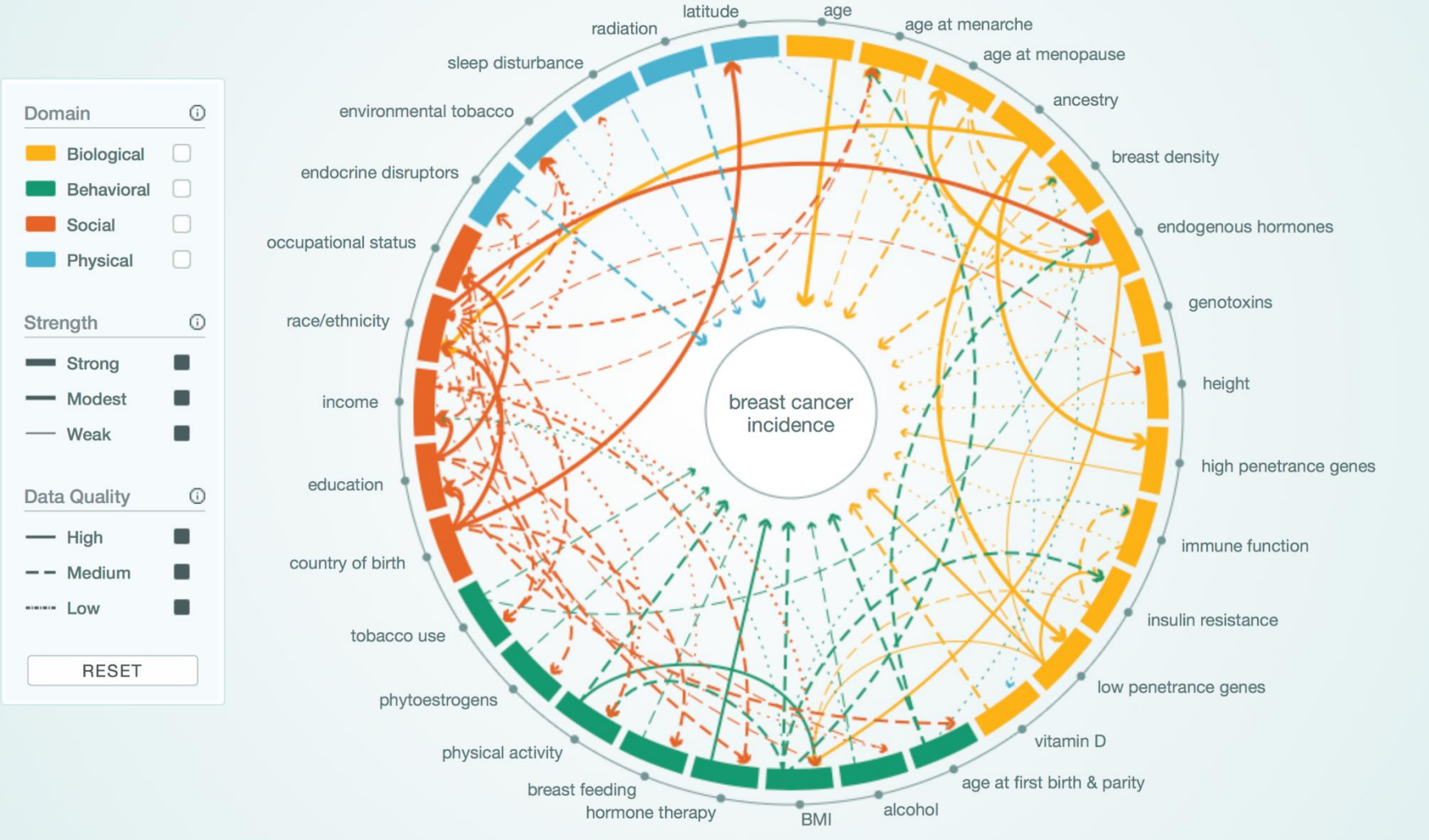
Visualizing the many factors and relationships influencing breast cancer incidence in postmenopausal women



Can you infer causality from this diagram?

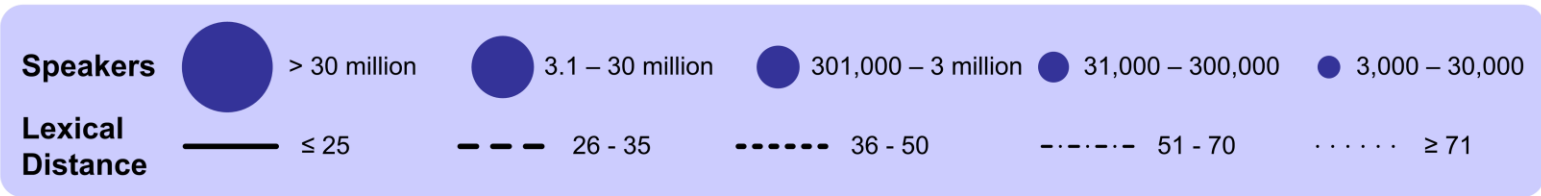
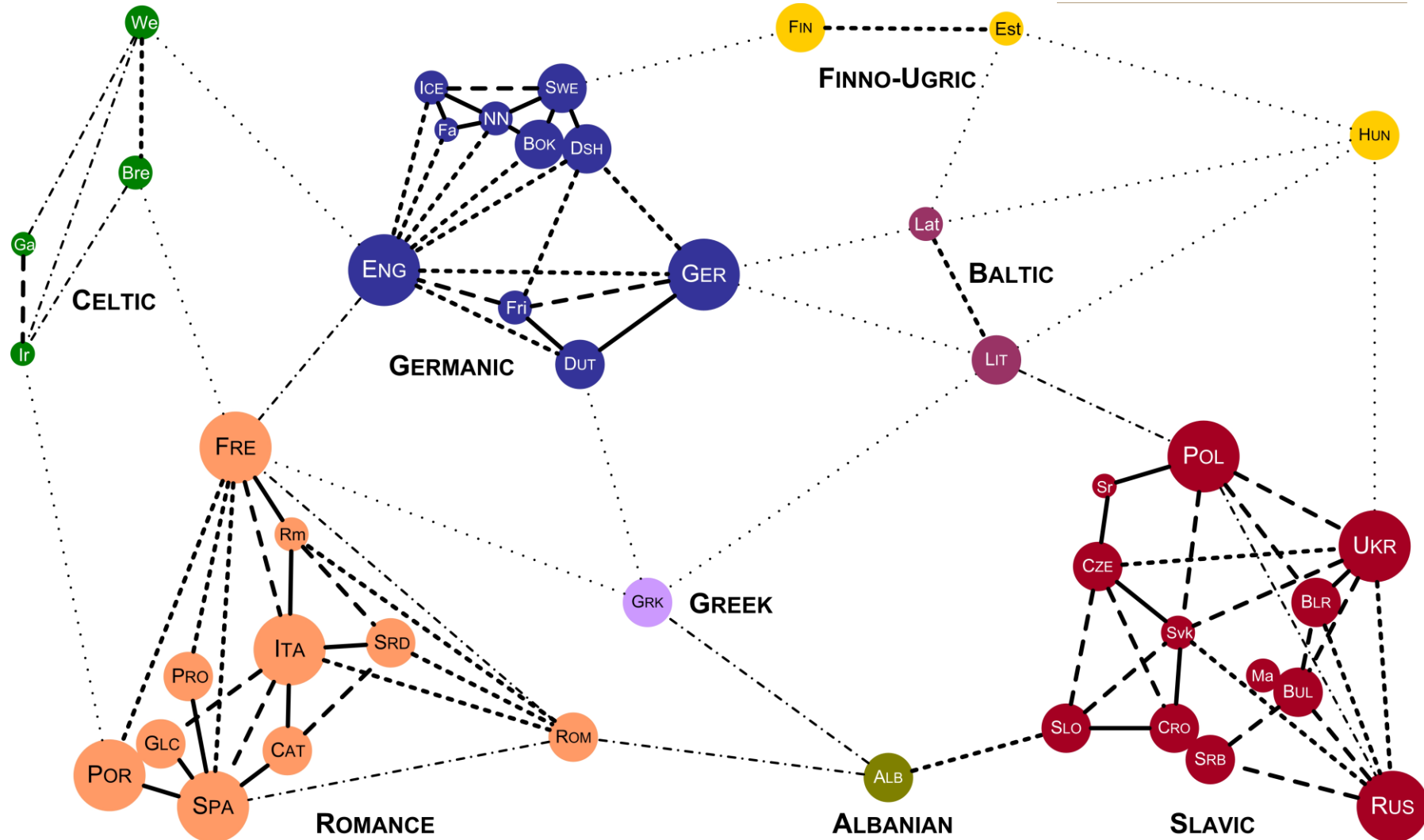
A Model of Breast Cancer Causation

Visualizing the many factors and relationships influencing breast cancer incidence in postmenopausal women

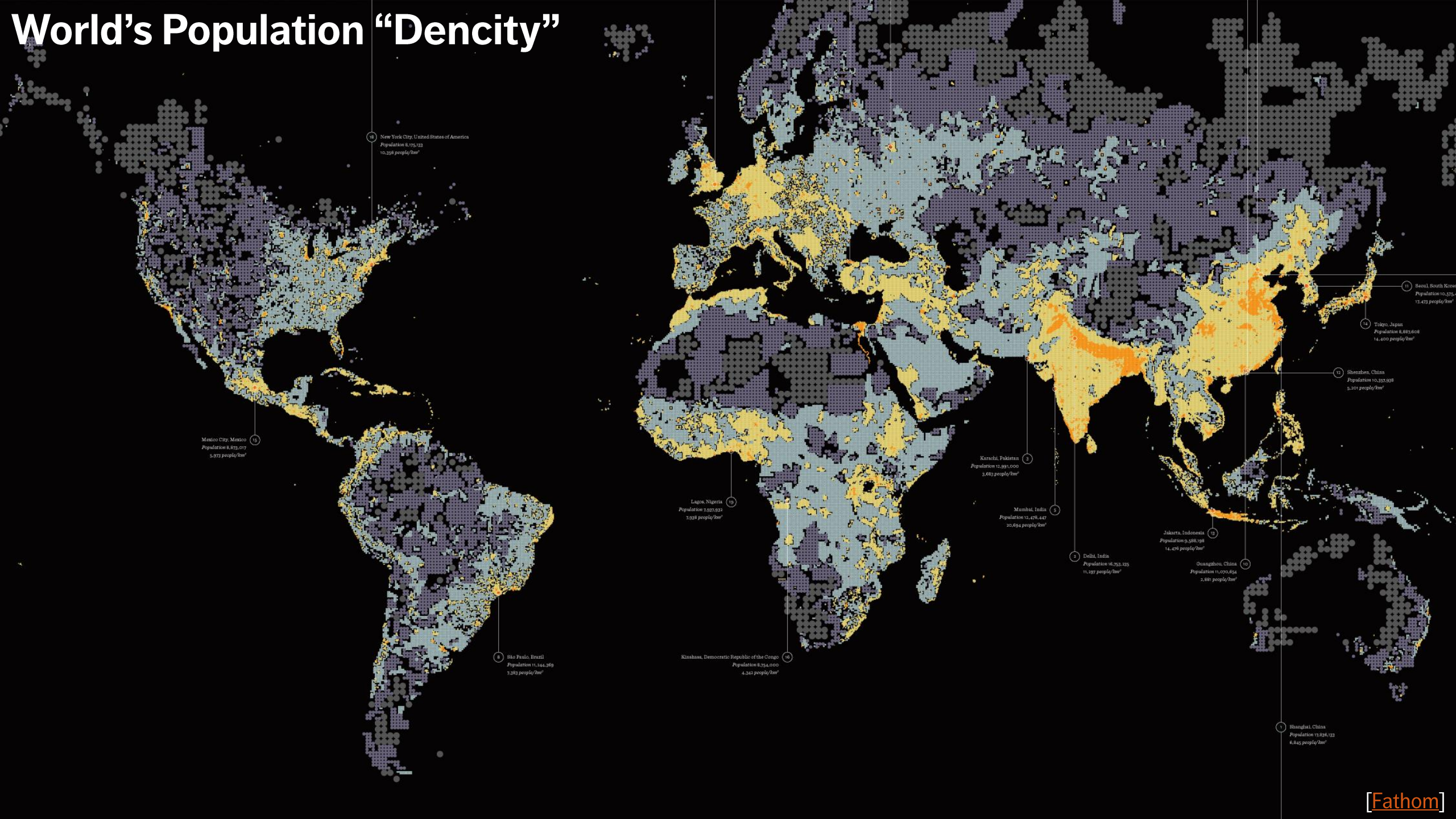


Can you infer causality from this diagram?

Lexical Distance Among the Languages of Europe



World's Population "Dencity"





MAPPING PAID PATERNITY LEAVE

HOW MUCH TIME DO OTHER COUNTRIES GUARANTEE COMPARED TO THE U.S.?

Low data density

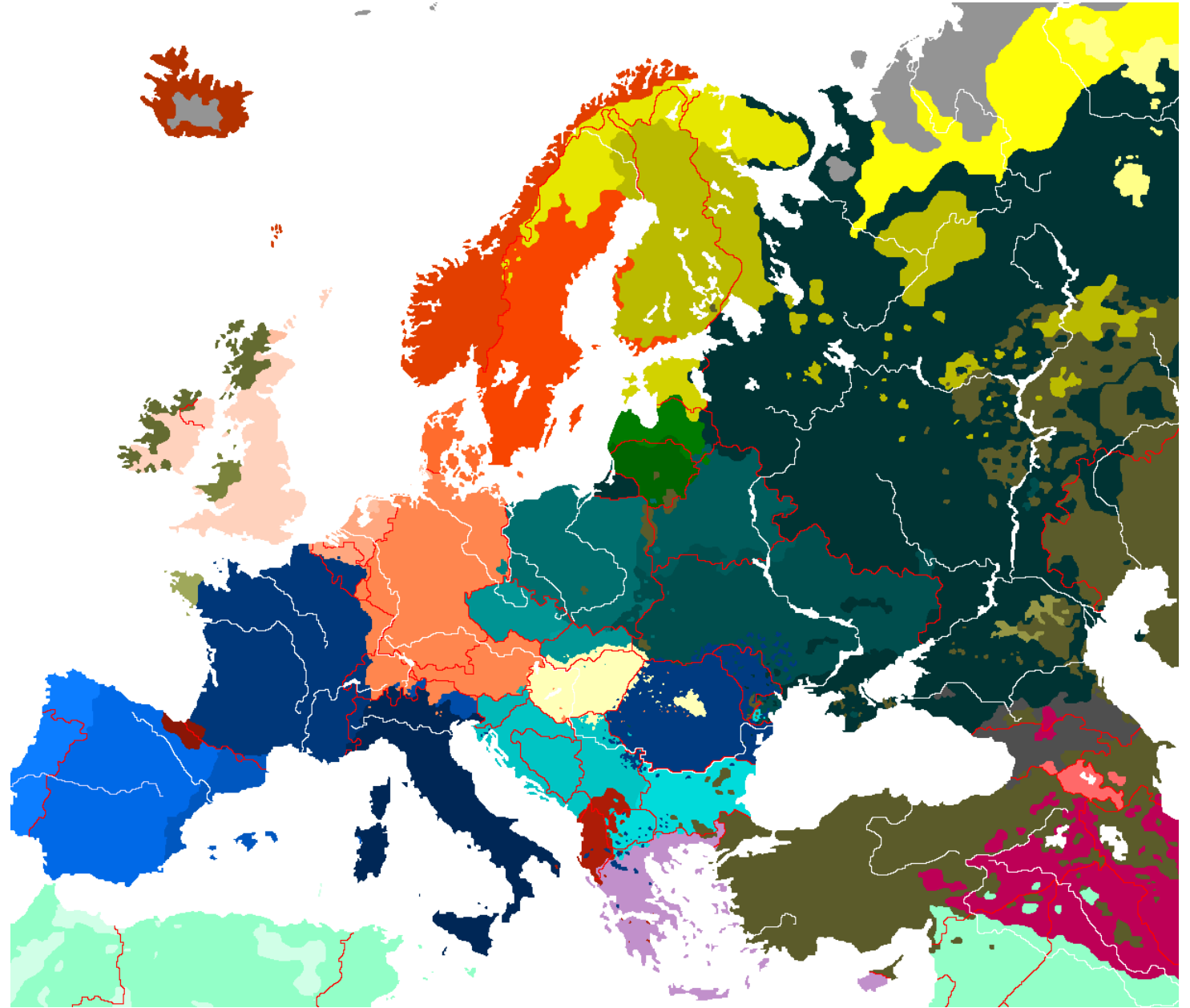
High chartjunk ratio

Scaling effects

Cherry-picking

Why not use a bar chart or a tabular display instead?





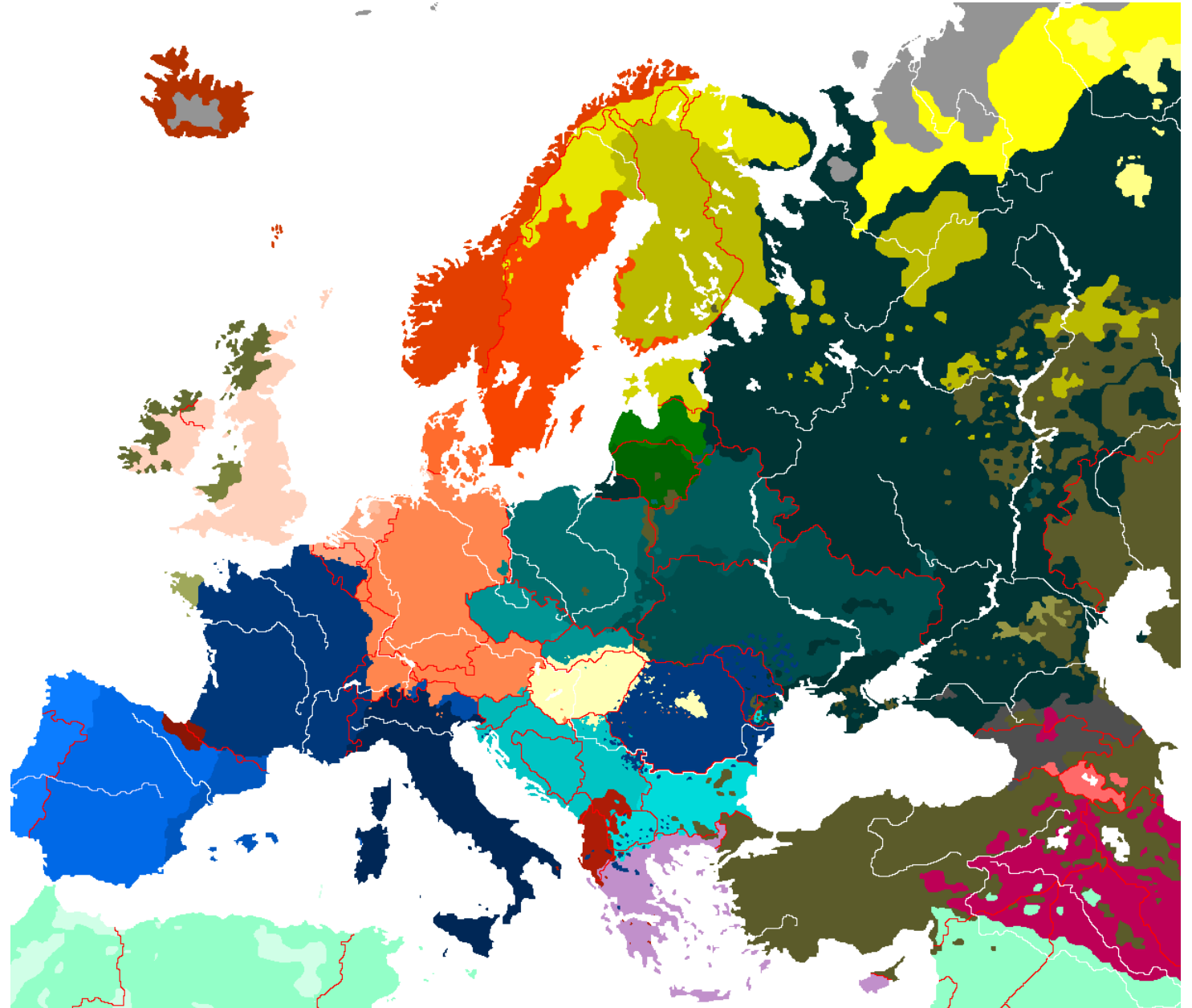
Encoding?

Population density?

Secondary languages?

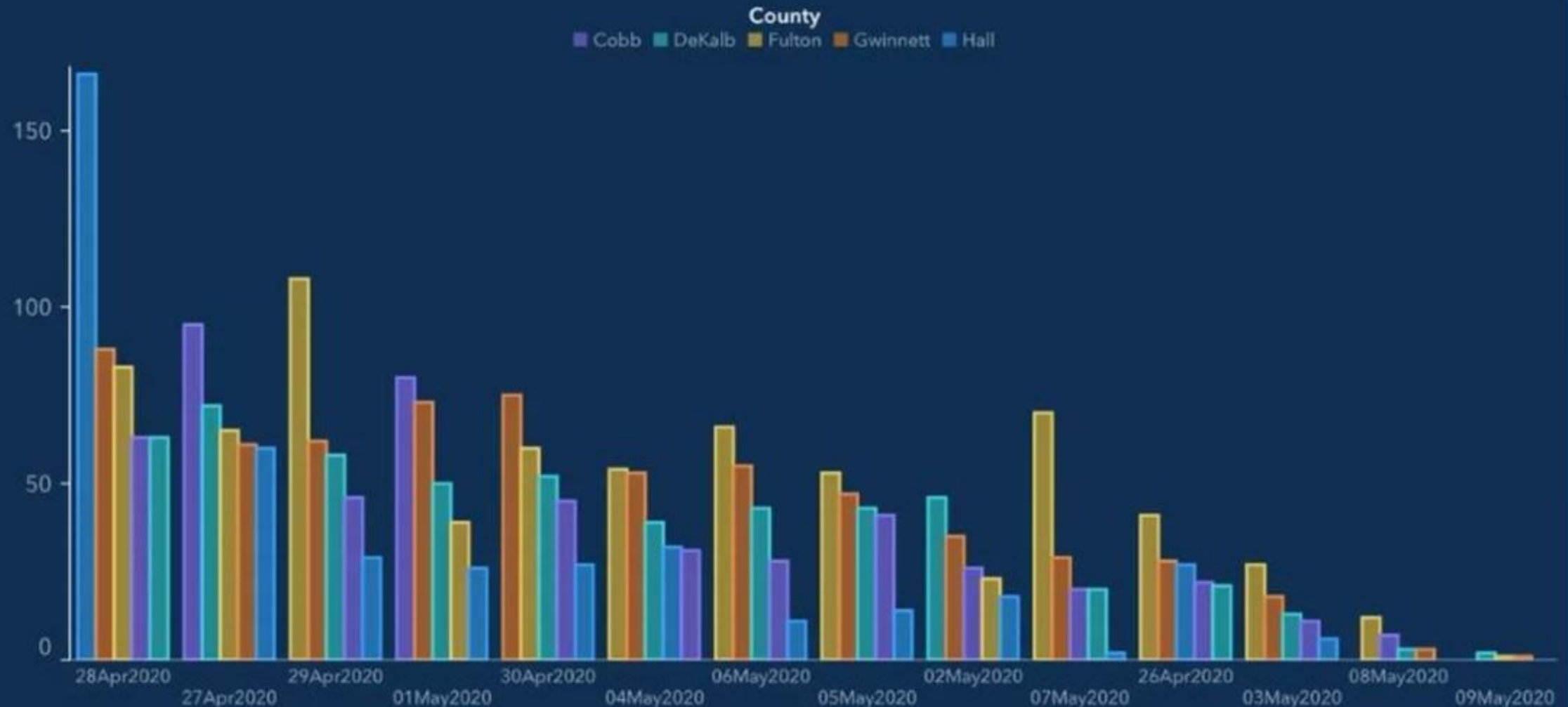
Rivers?

No data source



Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.



Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.



TAKE-AWAYS

Effective data visualizations **provide insights** and **facilitate understanding**.

The basic principles can guide your visualization design and consumption.

Be **creative**, but keep your data and your representations **honest**.

Be mindful of attempts to distort trends and conclusions with flashy visuals.

Data and code should be made available along with the displays.