

CIS 4930/6930-002

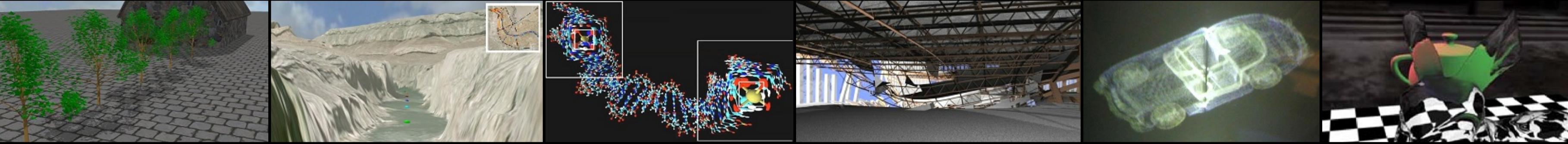
DATA VISUALIZATION



FILTERING & AGGREGATION

Paul Rosen
Assistant Professor
University of South Florida

slides credits Miriah Meyer (U of Utah)



REMINDERS

4/4/2018 – Project 6 Peer Reviews Due

4/9/2018 – Project 7 Due

4/11/2018 – Paper 4 Review Due

4/16/2018 – Project 7 Peer Review Due

4/23/2018 – Project 8 Due



Reducing Items and Attributes

④ Filter

→ Items

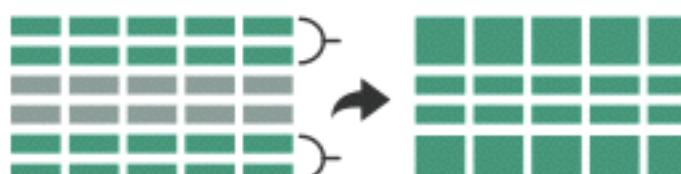


→ Attributes

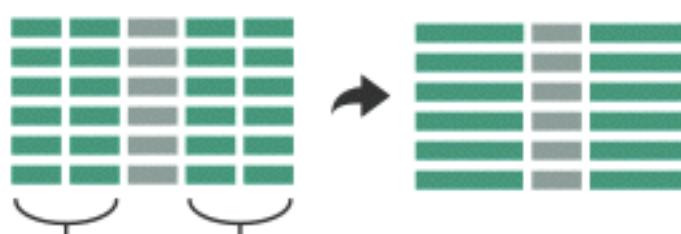


④ Aggregate

→ Items



→ Attributes



TODAY
filtering & aggregation



WHY REDUCE?



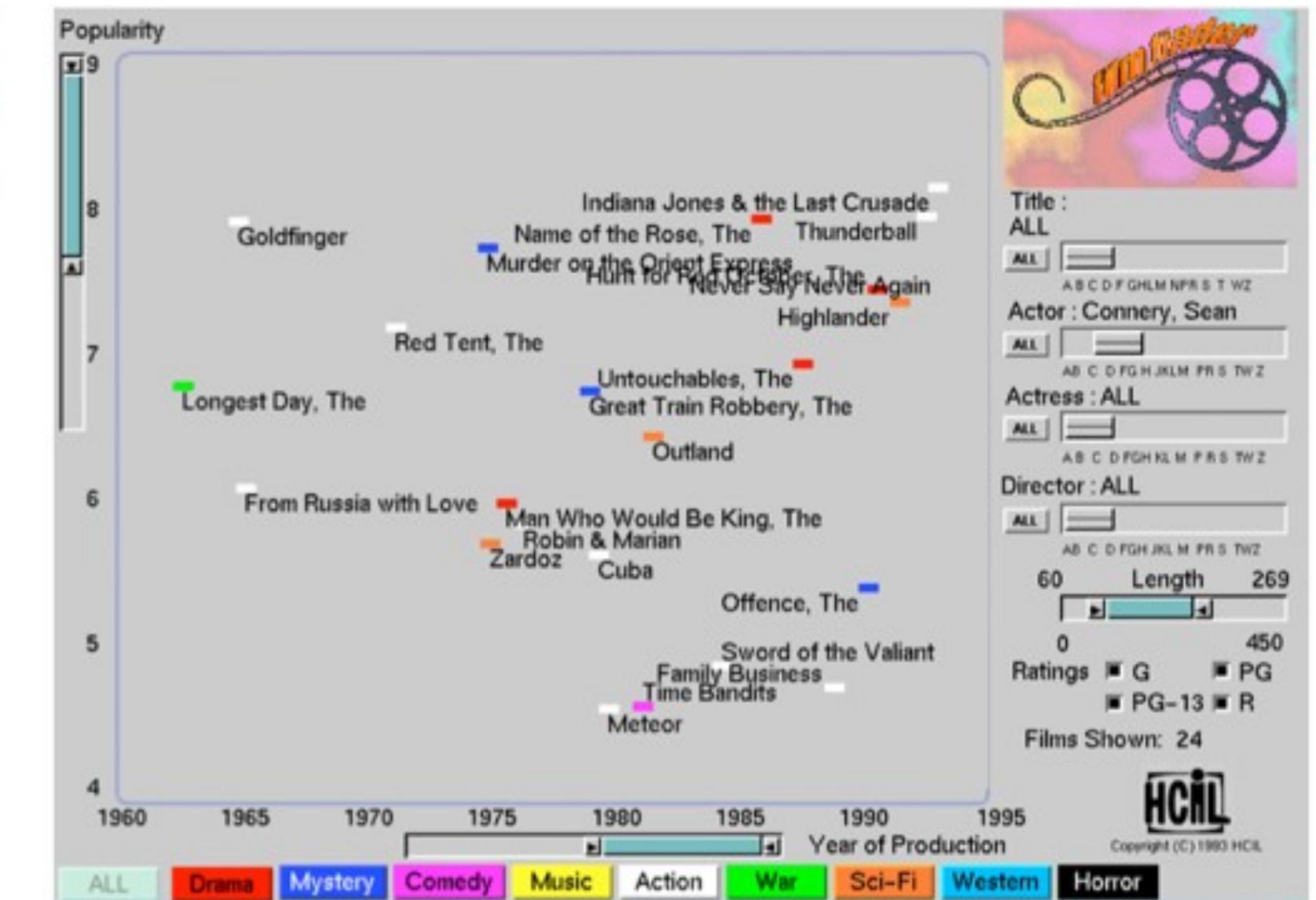
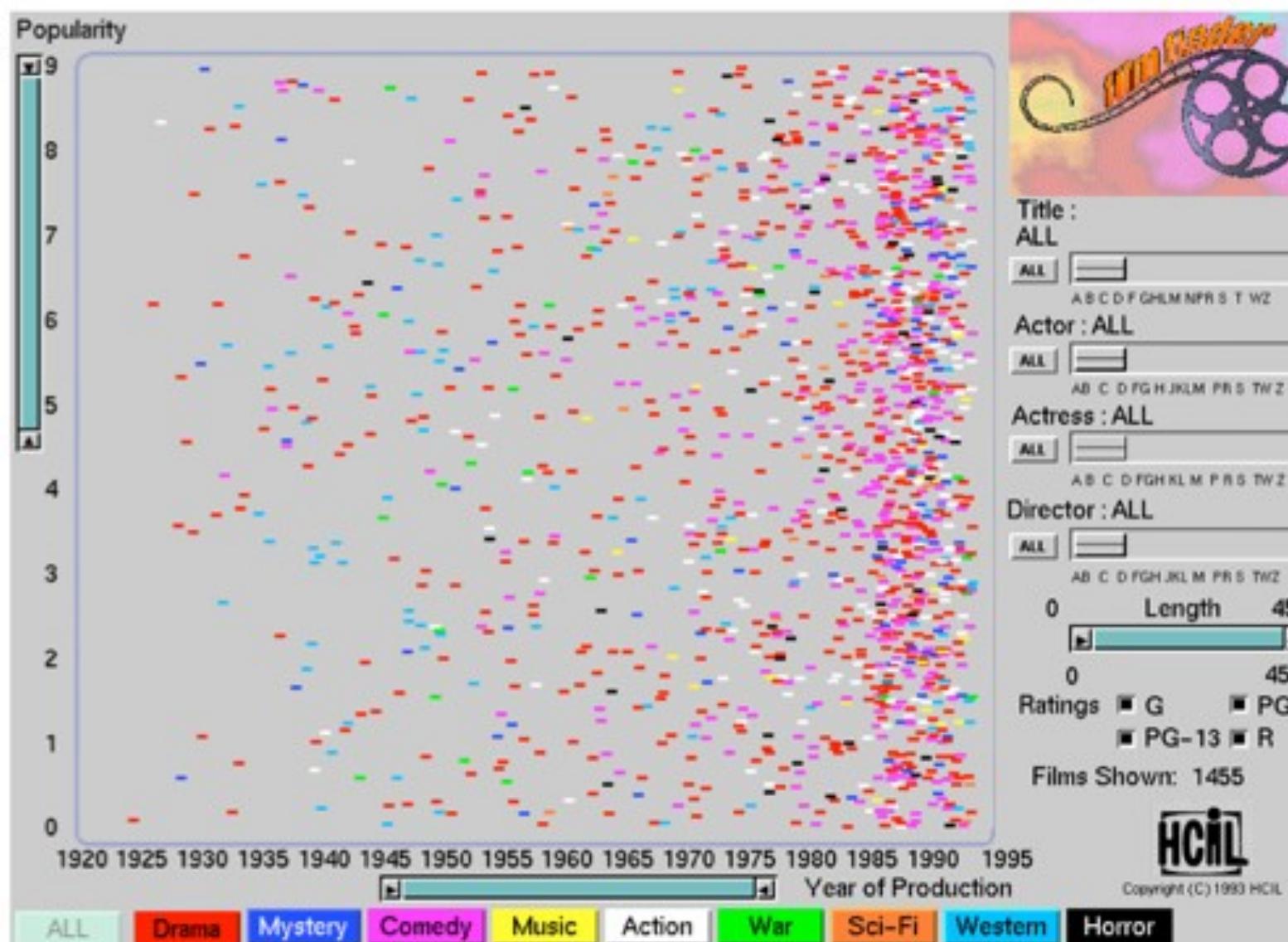
FILTER

elements are eliminated

dynamic queries
coupling between encoding and interaction so
that user can immediately see the results of an
action



ITEM FILTERING



UPDATED June 25, 2012

[RECOMMEND](#)

[TWITTER](#)

[LINKEDIN](#)

[SIGN IN TO E-MAIL](#)

[SHARE](#)

New York Health Department Restaurant Ratings Map

The New York City Department of Health and Mental Hygiene performs unannounced sanitary inspections of every restaurant at least once per year. Violation points result in a letter grade, which can be explored in the map below, along with violation descriptions. The information on this map will be updated every two weeks. For menus and reviews by New York Times critics, visit [our restaurants guide](#). [Related Article »](#)

[FIND A RESTAURANT](#)

[FIND A LOCATION](#)

[FILTER](#)



Restaurant locations are derived from the New York City Department of Health and Mental Hygiene database. Due to the limitations of the Health Department's database, some restaurants could not be placed.

By JEREMY WHITE | [Send Feedback](#)

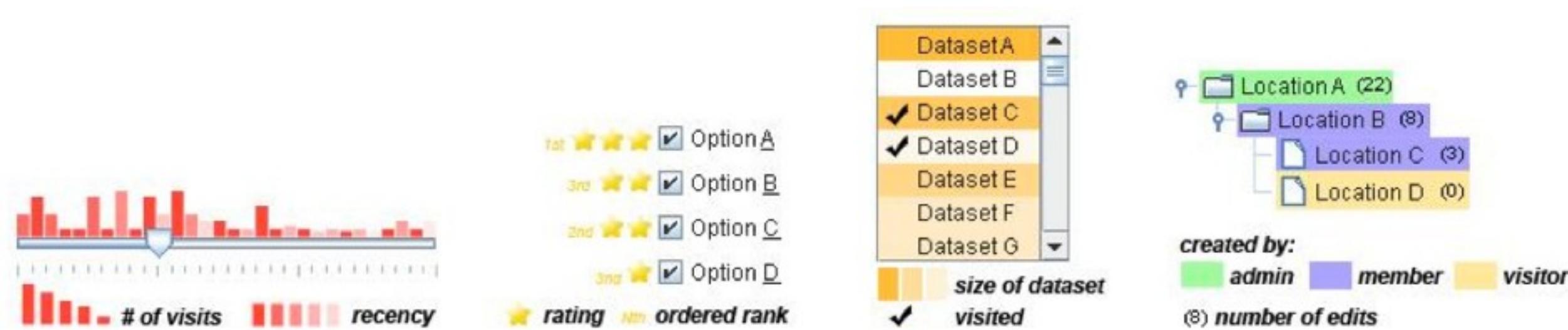
Source: New York City Department of Health and Mental Hygiene



SCENTED WIDGETS

information scent: user gets sense of data

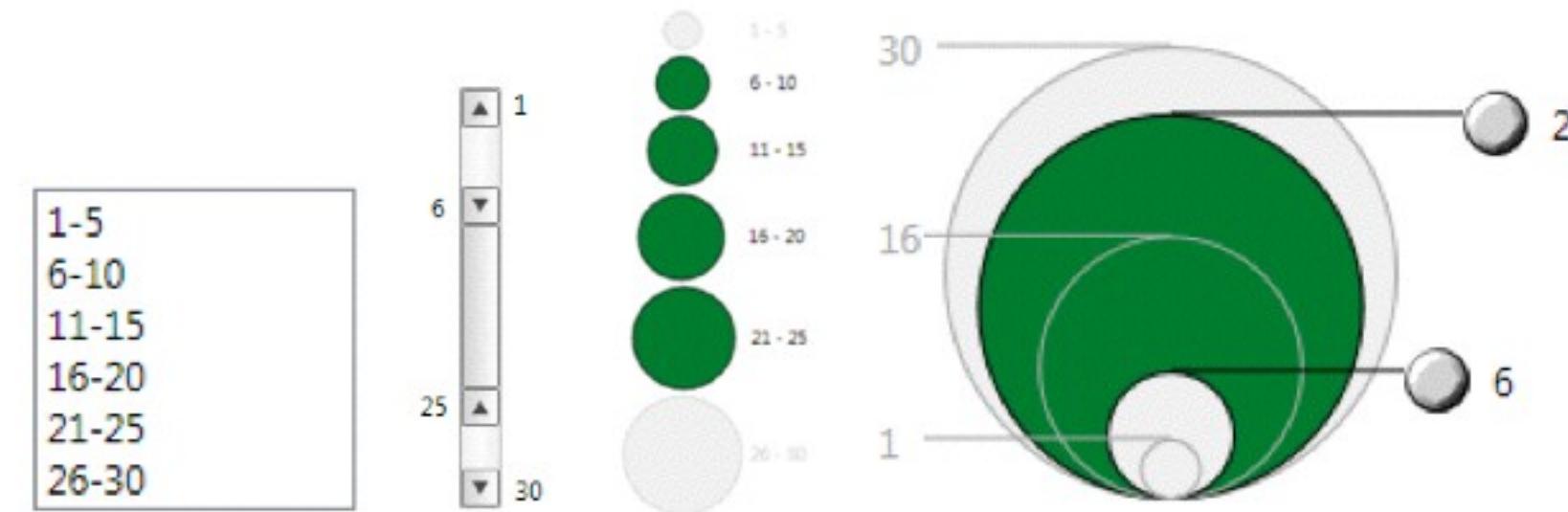
GOAL: lower the cost of information forging through better cues



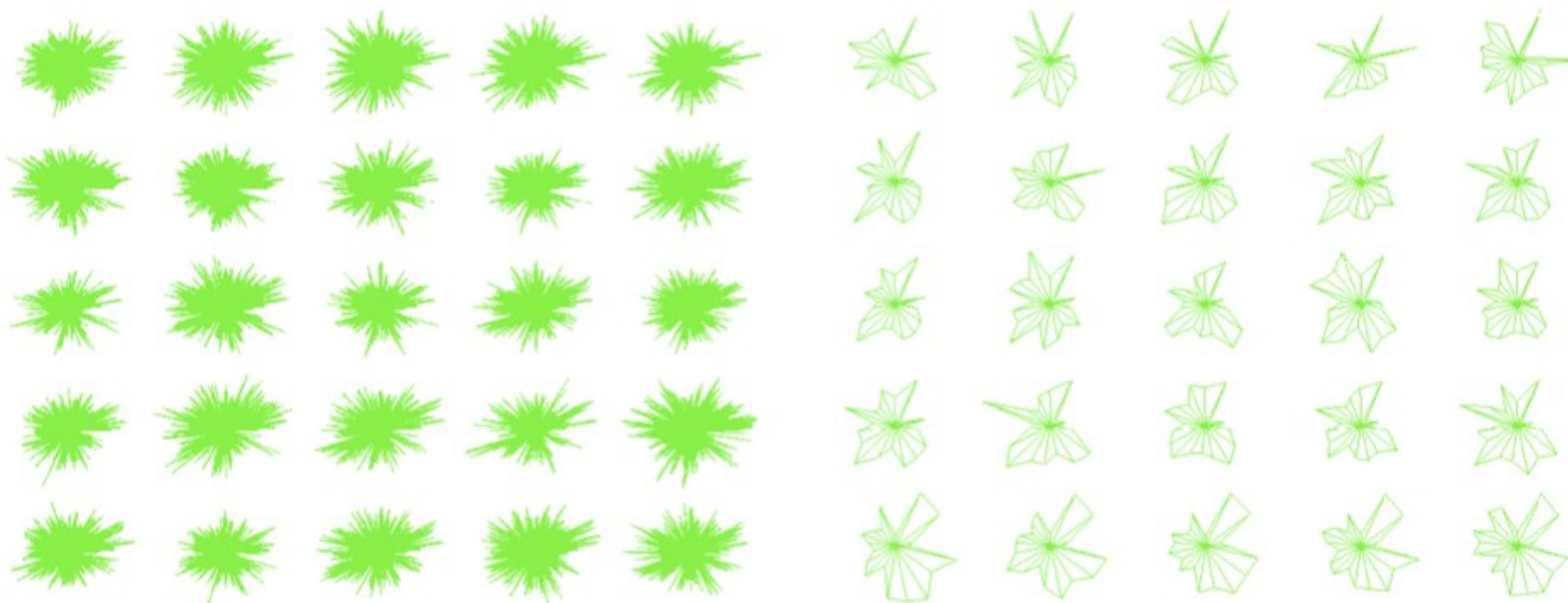
INTERACTIVE LEGENDS

controls combining the visual representation of static legends
with interaction mechanisms of widgets

define and control visual display together



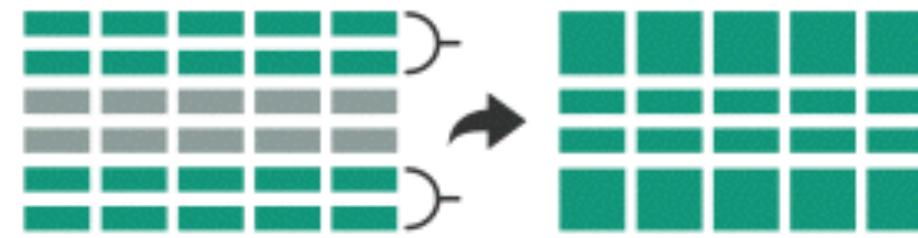
ATTRIBUTE FILTERING



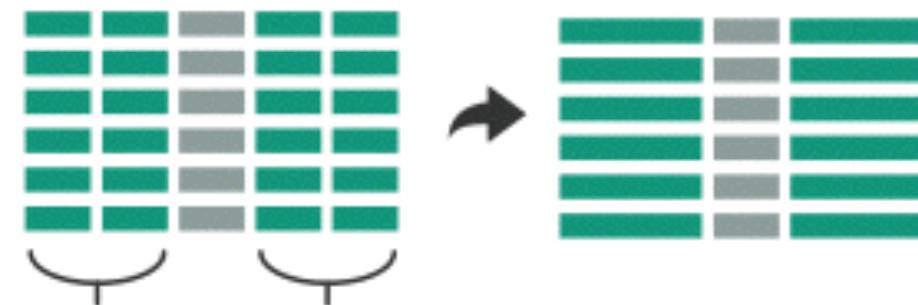
AGGREGATE

a group of elements is represented by a new derived element that stands in for the entire group

→ Items

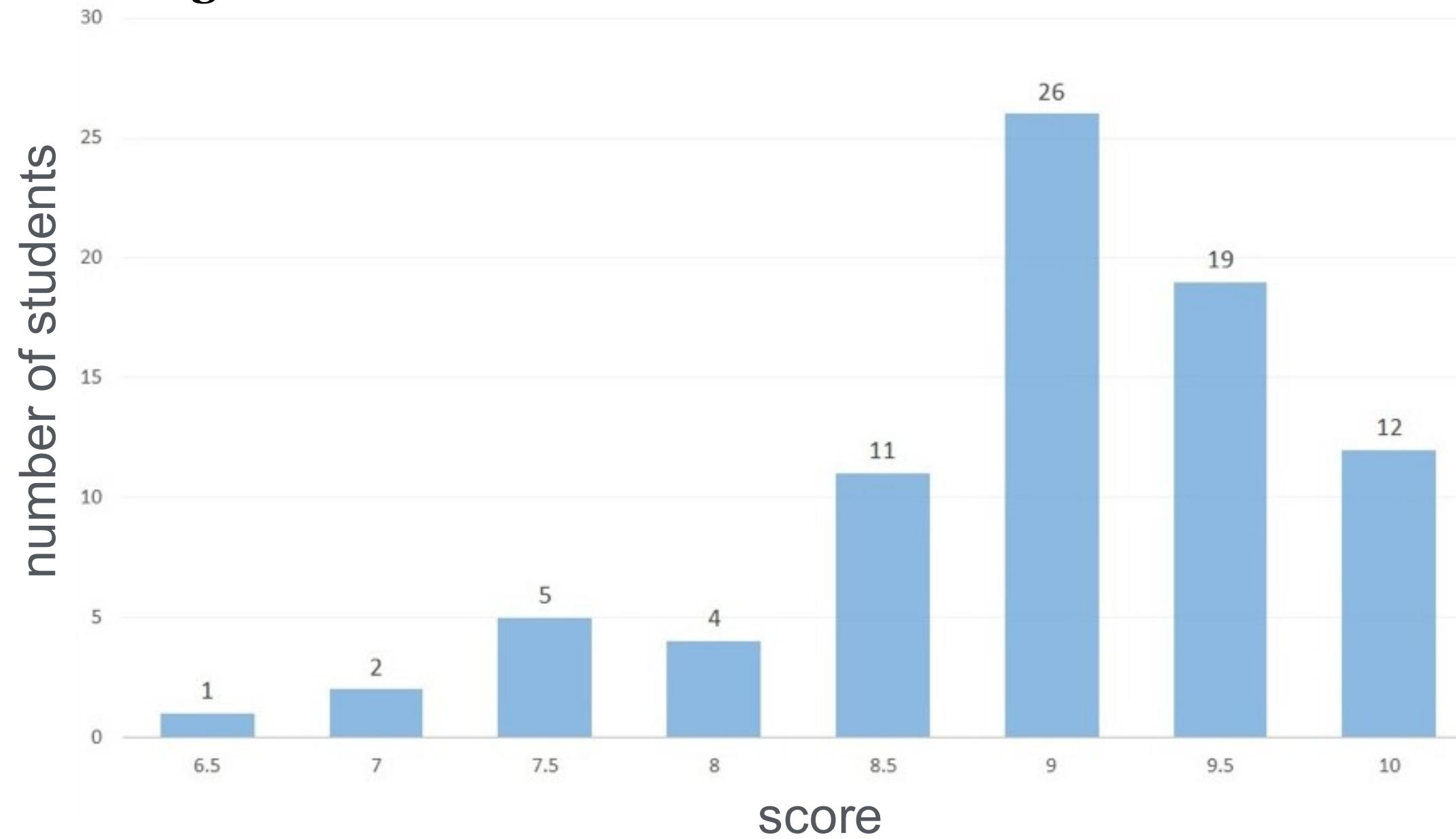


→ Attributes



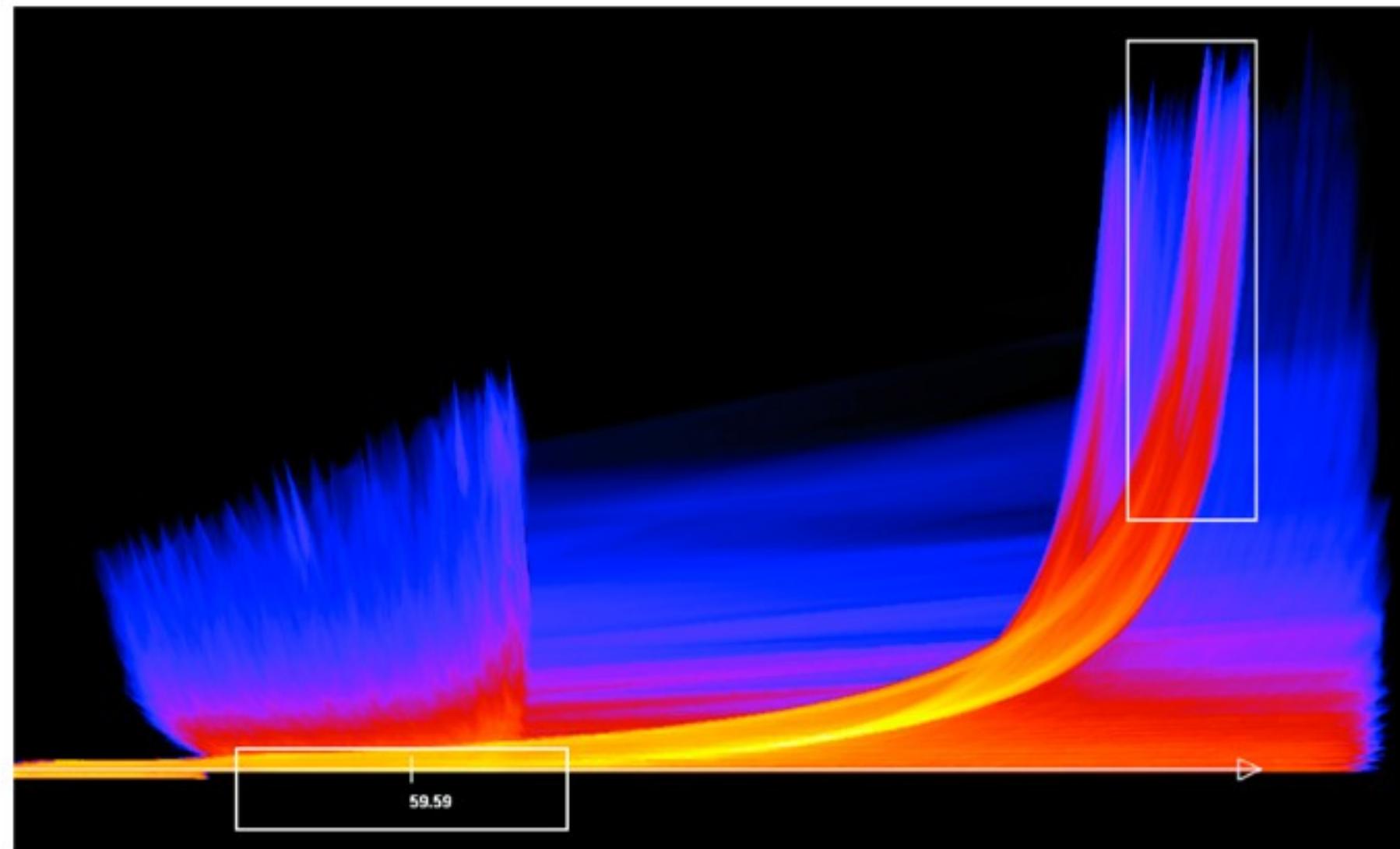
ITEM AGGREGATION

histogram



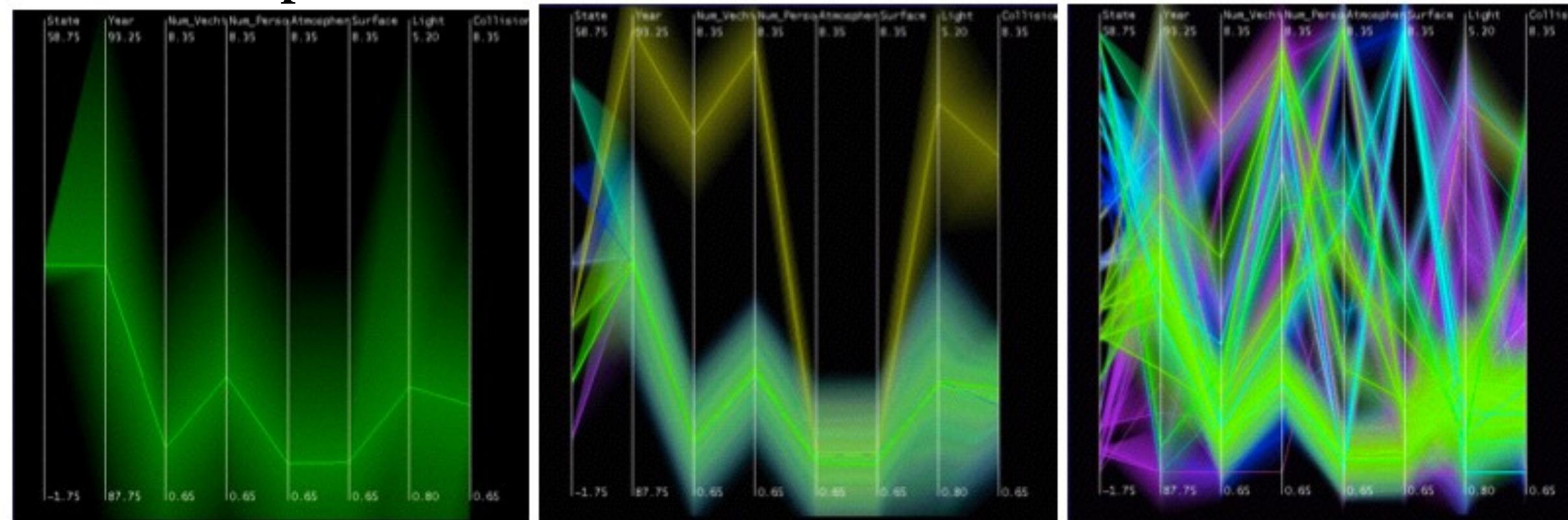
ITEM AGGREGATION

continuous scatterplot

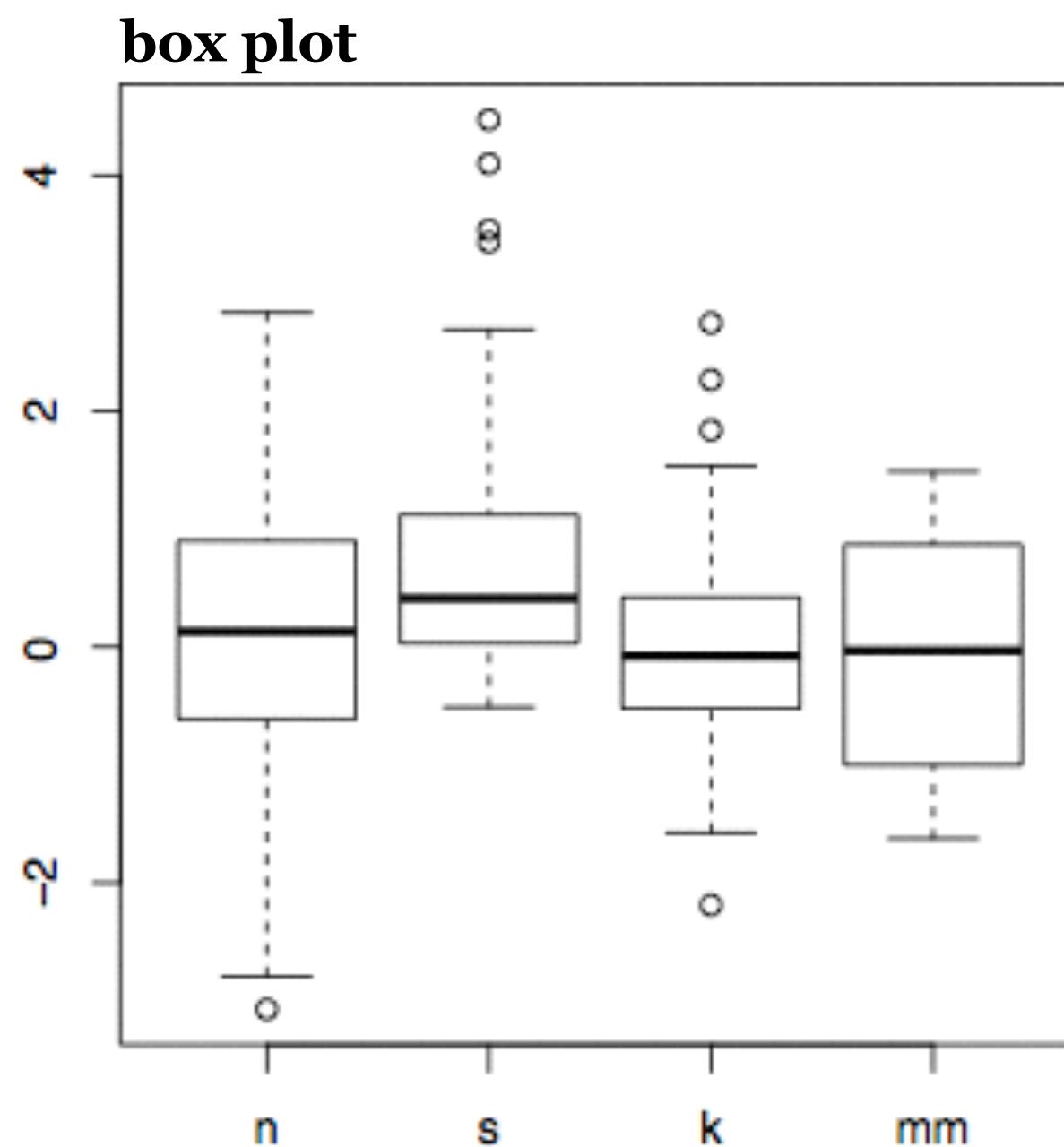


ITEM AGGREGATION

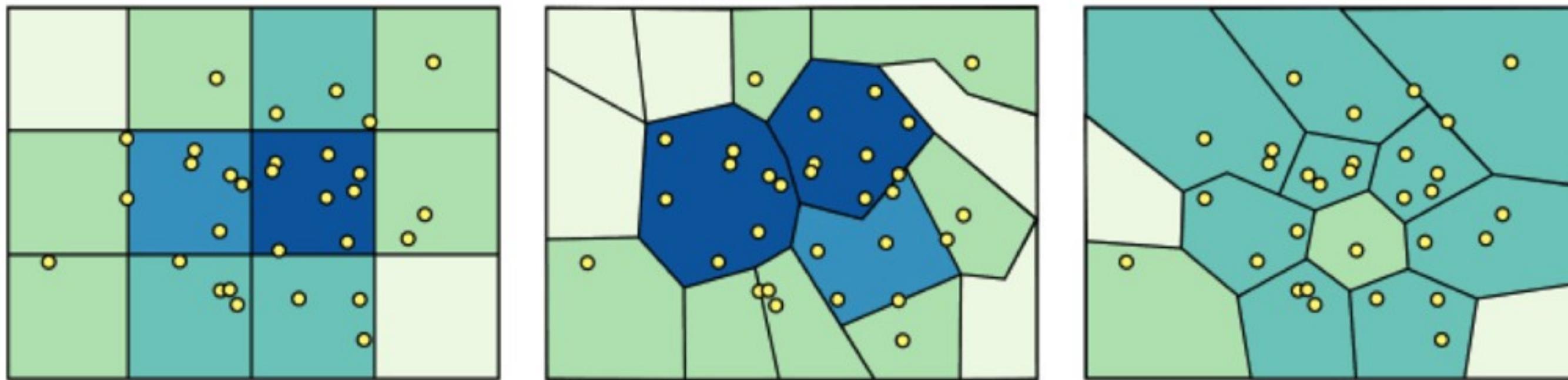
hierarchical parallel coordinates



ITEM AGGREGATION



SPATIAL AGGREGATION

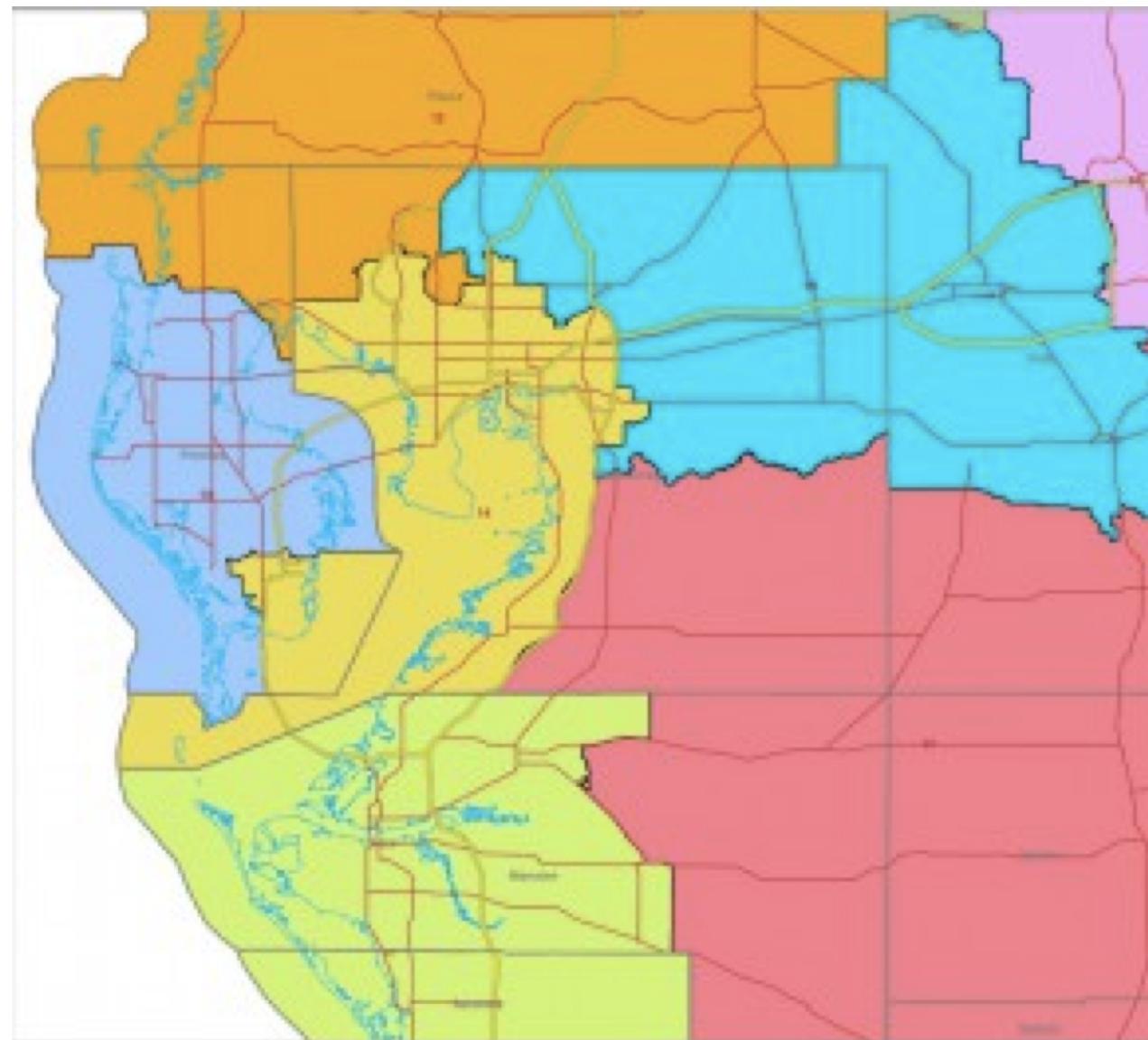


MODIFIABLE AREAL UNIT PROBLEM

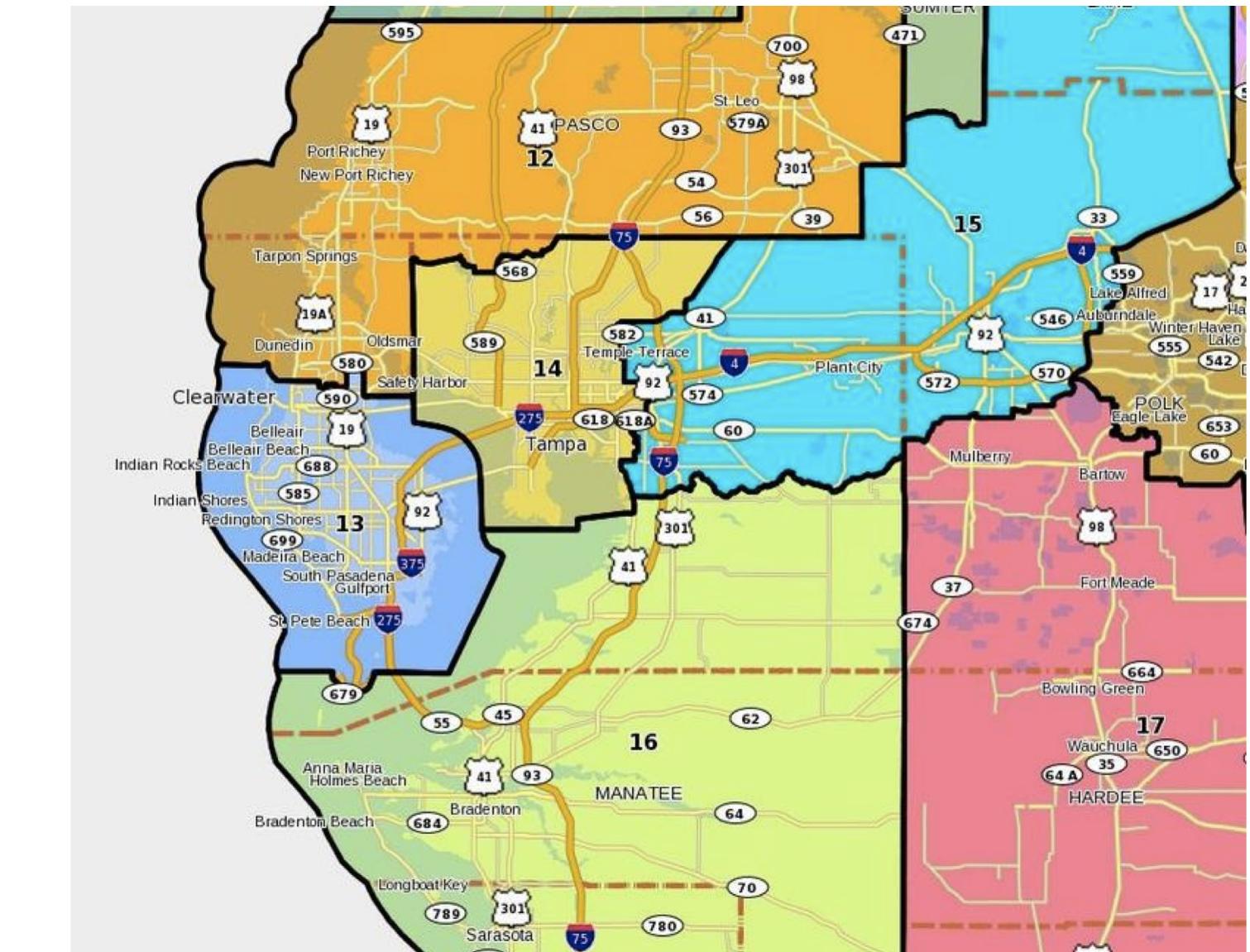
in cartography, changing the boundaries of the regions used to analyze data can yield dramatically different results



CONGRESSIONAL DISTRICTS



2014



Proposed 2015



ATTRIBUTE AGGREGATION

group attributes and compute a similarity score across the set
dimensionality reduction to preserve meaningful structure



SIMILARITY SCORES

correlation

measure of similarity between 2 or more attributes
many variants—pearson, rank, multi-way, etc.

regression

fit a model to the data
measure the quality of fit (i.e. R^2)



WHERE DO WE GO FROM HERE?

Big Data 3 V's – Variety, Velocity, Volume

Large & high-dimensional (multi-attribute) data

Solutions: statistics and probability, data mining and machine learning, and computational topology



