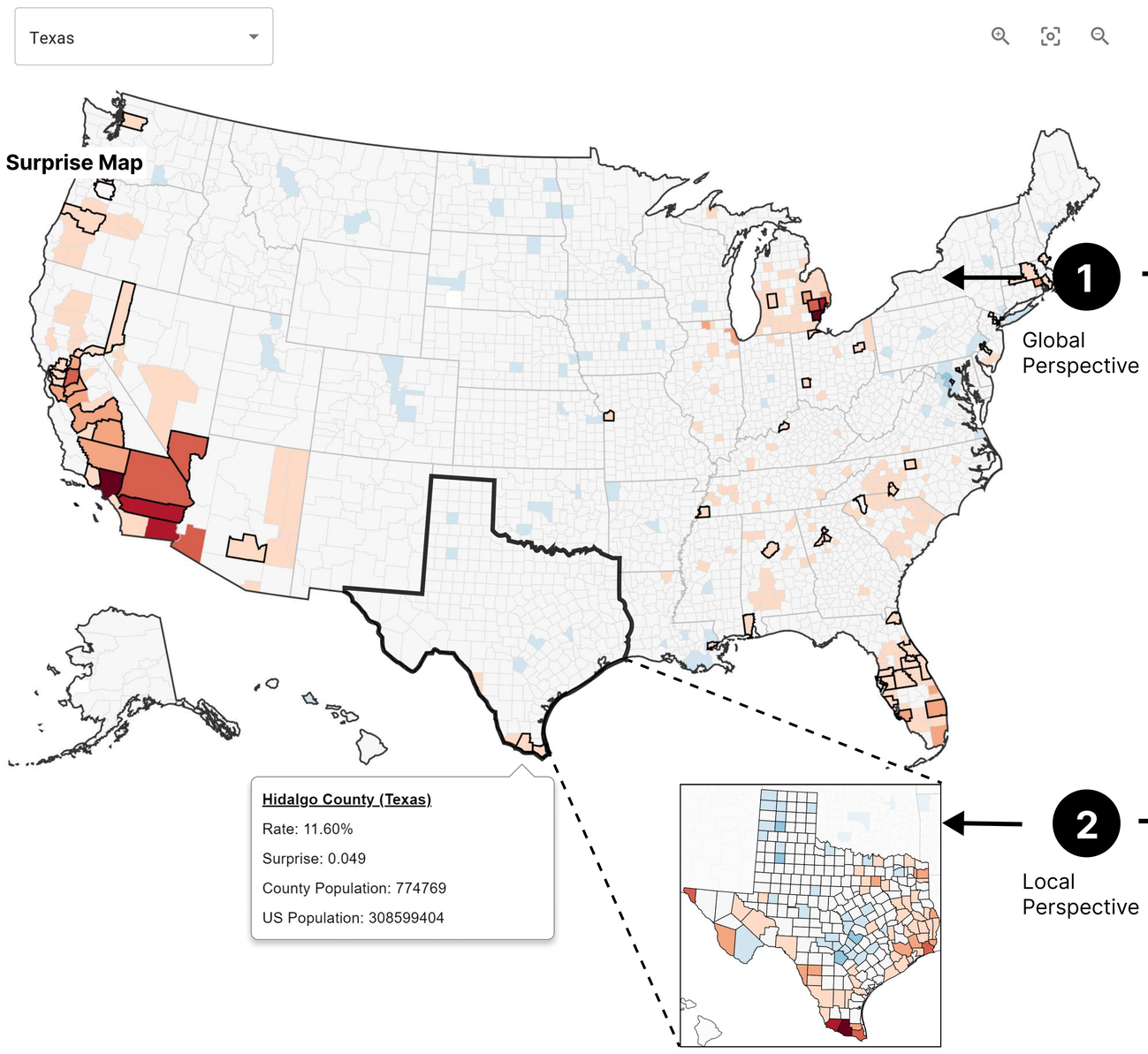
# SurpriseSync: Visual Exploration for De-biased Choropleth Maps

**Akim Ndlovu**, Hilson Shrestha, Evan M. Peck, **Lane T. Harrison**




**Which counties show significant deviation from prior beliefs based on their population?**
A user can select counties of interest (eg. high/low z-score, and or high/low population) by brushing on the funnel plot.

**Can I emphasize low surprise counties to tell a story?**
Drag color scale to positively or negatively skew outliers

**What is the relationship between county population and their associated statistical significance?**
Hover to highlight a county

**What are the counties with high unemployment rate and low population?**
Brush to select a group of counties in all 4 dimensions

**What is the relationship between rate and population with surprise/z-score for a county?**
Hover over a line to explore county metrics

People are able to generate and explore hierarchical models of belief through the use of augmenting visualizations (Surprise map, Funnel Plot and Parallel Coordinate Plot).

## Which regions on a map deviate from expected event rates?

Statistical techniques such as Bayesian surprise[1] have been shown to be useful for revealing informative regions on a map. Yet challenges remain:

## Challenges

Local — Visualizing data at a national level does not prov insights for people interested in local trends[2].

Experts — Implementing Bayesian weighted maps requires modeling and statistical expertise[1].

## Contributions

**SurpriseSync,** a system that facilitates the exploration of de-biased Choropleth maps through the use of multiple coordinated views.

An hierarchical exploration workflow that facilitates comparisons between global and local beliefs.

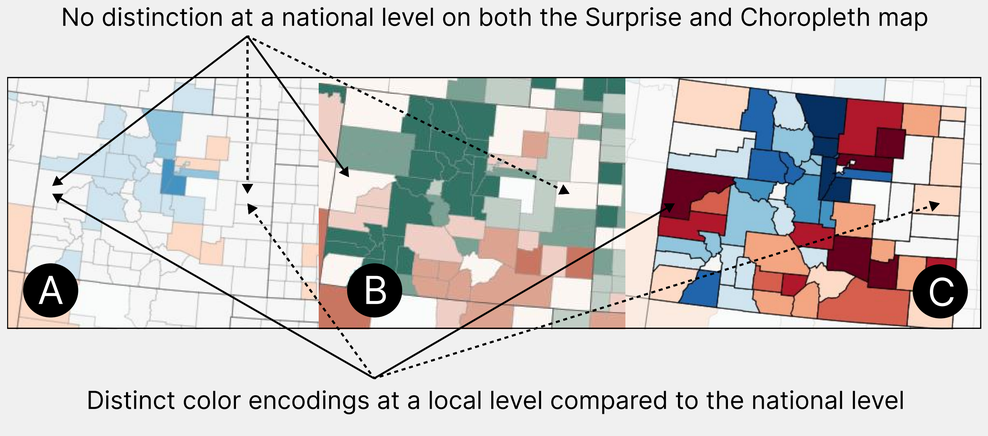
Interactive skewing of prior beliefs to induce alternative exploration insights.

## SurpriseSync Supports

1. Fine **tuning** of the **belief model** through the use of an **interactive color scale**.
2. The use of **multiple views** to explore **relationships** between counties (i.e Highlighting counties with similar metrics).
3. The **identification** of **statistically significant regions** that may be hidden in traditional maps through the use of custom interaction sequences.

## Exploration Insights

No distinction at a national level on both the Surprise and Choropleth map




Distinct color encodings at a local level compared to the national level

A visual distinction of spatial patterns for the state of Colorado, at a **A)** national ”Surprise” level **B)** national event-rates level and **C)** local ”Surprise” level. However, differences in color encoding, are more pronounced at a state level.

## Future Work

1. How does the complexity associated with adding multiple interactive views, influence peoples exploration insights?
2. Extend **SurpriseSync** to explore and formalize the notion of “multi-variable” surprise.

**REFERENCES**
[1] Correll, M., & Heer, J. (2016). Surprise! Bayesian weighting for de-biasing thematic maps. IEEE transactions on visualization and computer graphics, 23(1), 651-660.
[2] Peck, E. M., Ayuso, S. E., & El-Etr, O. (2019, May). Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-12).