

INFO 290T

Human-Centered Data Management

Trust but Verify



Announcements

- Thursday's class will happen in 205
- Project Proposal Due Tonight
- “Grading” paper reviews, more soon!
- Also “grading” class participation, more soon!
- Also “grading” class presentations, more soon!

- Poll forthcoming: Class on Oct 19 and Nov



Thoughts on Paper?

- Optimistic Visualization Proposal?
- Interface?
- Evaluation?
- Writing?



Use Cases

Insights are drawn from visualizations.

For what types of insights is the “trust-but-verify” approach in the pangloss manner – as opposed to simply “wait” – a *bad* idea?



Use Cases

Insights are drawn from visualizations.

For what types of insights is the “trust-but-verify” approach in the pangloss manner – as opposed to simply “wait” – a *bad* idea?

When the subsequent decisions/takeaways are crucially dependent on some fine-grained insight, e.g.,

- Comparisons (“chairs sold more than tables, let me investigate why”)
- Trends (“Sales of chairs has been going up throughout, great! Let’s look at tables”)
- Anomalies (“no products receiving an unsafe grade, great!”)



Industry Perspective: Discussion

I'm from Tableau, considering adopting Trust but Verify – what are considerations I must think about?



Informal Industry Perspective

I'm from Tableau, considering adopting Trust but Verify – what are considerations I must think about?

- “Remembering” overload for users
- Can you actually tell what has changed?
- What if a decision to go down a path was because of an incorrect assumption?
- Looking at multiple visualizations for comparison (heatmap) – hard for people to do
- Not enough visualization types (line charts, scatterplots - not a great fit for AQP)



Informal Archaeology

- Paper this is based on: Sample + Seek
 - AQP Scheme that combines (roughly)
 - “sampling” (when there are enough tuples of a given type that a uniformly random/stratified sample suffices) and
 - “seeking” (when there are so few tuples that it’s better to use an index to look those tuples up)
- Paper based on this: Falcon!



Other Open Questions

- Was it odd that the revisiting a “remembered” visualization never caused any users to change their “insight”?
- The authors had to repeatedly prod their participants to “remember” the visualization. Good or bad?
 - Generally, what if an alternate design is to skip the verification step altogether but the system remembers the entire set of visualizations and reproduces each one?
- What if the system “auto-remembers” all the visualizations with high deviation?
- How does the system prioritize across many different simultaneous “remembered visualizations”?

