Meeting notes

Shu SHANG

26/04/2017

1 Main points

- User in-the-loop *vs.* automatic way
  - We choose to start from the automatic way bacause it is the basic problem of this project.
  - User in-the-loop can be extended to it after we handle the automatic way.
- What make a visualization interesting?
  - An interesting report *Les fractures françaises: le logement, les raisons de la crise*
    - We are interested in visualizations having **time series** as one of its axis.
      - How? We can **Group by time** (Especially, **by year** should be interesting as year has few distinct values, spaning over a large range. We can also think about **group by month...**)
      - Put several subsets of data together. For instance, in the case of *Fractures française*, we are interested in 4 subsets of tenant: 1. **locataire parc social** 2. **locataire secteur libre** 3. **accédants à la propriété** 4. **propriétaires san prêt en cours**. The question is: How much of their annual revenue did they use to cover their loyer/loan during the past ten years? → We can get 4 visualizations for each subset of the data, each visualization has year as X-axis and the percentage of revenue as Y-axis. → We can then compare the results of 4 subsets. (**Scale** problems? How can we automatically analyze the 4 subsets are comparable? What kinds of metric should we propose to analyze the differences of the 4 visualizations?)
      - In the example above, one or more results may stand out compared with others, the visualization can thus become more interesting. Here, we may have a particular subset of people we concern most, other groups can viewed as reference data (Recall the seedb example). In some cases, visualization shows interesting insights compared with some reference data (For instance, between **prix de logement** vs. **time** and **revenu** vs. **time**), while others originally show interesting insight (For instance, trend of house price over time).
      - Think about proposing new metrics (For instance, SEEDB proposed a new metric *Utility* based on deviation...) to compare the differences of visualizations and find interesting ones.
    - We are intereted in some quantile features

- For instance, in the example of *Fractures française*, 4-quantiles: les 25% les plus modestes, deuxième quartile, troisième quartile, les 24% les plus aisés. (quartiles based on revenue)

2 Next

- Transformation from bar chart to trend line
- Think about the ideas proposed in the meeting, visualizations having time series as axis, visualizations having quantile features...