CSE 494/591
Data Visualization Projects
Dr. Sharon Hsiao
2015/01/21
Final Project Deliverables

- **Intelligent** Interactive visualization
  - must be accessible online
  - submit all source codes or executable files as a zip.
- 6-10(min-max) pages paper, unlimited extra pages for references.
  - Introduction, Motivation, Visualization Design (implementation), Methodology (Clearly state why&how can your data visualization be used to solve the research questions), Evaluation Plan, Discussions & Future Work, References.
- Presentation slides should also be submitted. You will have to present this work (demo/explain it) in class.
Evaluation

- **30%** collecting data, cleaning data, conducting data analysis;
- **45%** prototyping visualization, implementing visualization (clarity, consistency, aesthetic, originality);
- **25%** demonstrating the implementation in report and presentation (technically sound? appropriate and sufficient references?)

(if it's completed as a group I +-(I – P) * 20% of the average group peer review)

i.e.

Group score = 90 (A-)
90 + (98-90) x 0.2 = 91.6
90 – (90-70) x 0.2 = 86
● Total 8-10 Projects
● Email TA to sign a project by forming a team of 2-3. First come first serve.
● Project sign up due: 2/4 Wednesday noon
(same as assignment 1 due, individual project proposal due)

● Project alternatives: individual project (2 pages proposal is required: dataset descriptions, research questions, motivation)
What do I do?

- **Computer Science Education**
  - programming
  - personalized (adaptive) tools
    - visual analytics
    - visual recommenders
- **CSI (Computing Systems & Informatics)**

Datasets

1. Stackoverflow Dataset: selected topics in Java
2. Stackoverflow Dataset: stack exchange API
3. Yelp Academic Dataset: (Phoenix)
4. Yelp Academic Dataset: all other available cities

http://www.yelp.com/dataset_challenge
Stackoverflow

- Java:
  ```
  {type, title, content, code, user_id, time, vote, reputation, accept_rate, tags}
  ```

- Unbounded:
  ```
  { (all of above), badges, featured, no-answered, upvote, flags, favorite, etc.} https://api.stackexchange.com/docs
  ```
Yelp

- **Business Objects**
  ```json
  {  'type': 'business',
      'business_id': (a unique identifier for this business),
      'name': (the full business name),
      'neighborhoods': (a list of neighborhood names, might be empty),
      'full_address': (localized address),
      'city': (city),
      'state': (state),
      'latitude': (latitude),
      'longitude': (longitude),
      'stars': (star rating, rounded to half-stars),
      'review_count': (review count),
      'photo_url': (photo url),
      'categories': [(localized category names)]
  }
  ```

- **Review Objects**
  ```json
  {  'type': 'review',
      'business_id': (the identifier of the reviewed business),
      'user_id': (the identifier of the authoring user),
      'stars': (star rating, integer 1-5),
      'text': (review text),
      'date': (date, formatted like '2011-04-19'),
      'votes': {
          'useful': (count of useful votes),
          'funny': (count of funny votes),
          'cool': (count of cool votes)
      }
  }
  ```

- **User Objects**
  ```json
  ...  ```
Project Categories:

1. Visual Analytics
2. Visual Recommenders
1. Stackoverflow Java
2. Stackoverflow unbounded
3. Yelp: Phoenix
4. Yelp: unbounded

Intelligent Interactive visualization

A. Visual Analytics
B. Visual Recommender
A Visual Analytics! But NOT Intelligent!
example of a simple **intelligent** visual analytics [http://twitter.github.io/interactive/sotu2014/#p1](http://twitter.github.io/interactive/sotu2014/#p1)

After all, that’s the spirit that has always moved this nation forward. It’s the spirit of citizenship – the recognition that through hard work and responsibility, we can pursue our individual dreams, but still come together as one American family to make sure the next generation can pursue its dreams as well.

Citizenship means standing up for everyone’s right to vote. Last year, part of the Voting Rights Act was weakened. But conservative Republicans and liberal Democrats are working together to strengthen it; and the bipartisan commission I appointed last year has offered reforms so that no one has to wait more than a half hour to vote. Let’s support these efforts. It should be the power of our vote, not the size of our bank account, that drives our democracy.

Citizenship means standing up for the lives that gun violence steals from us each day. I have seen the courage of parents, students, pastors, and police officers all over this country who say “we are not afraid,” and I intend to keep trying, with or without Congress, to help stop more tragedies from visiting innocent Americans in our movie theaters, shopping malls, or schools like Sandy Hook.

Citizenship demands a sense of common cause; participation in the hard work of self-government; an obligation to serve to our communities. And I know this chamber agrees that few Americans give more to their country than the men and women of the military, who put their lives on the line every day for our freedom.
A recommender! But NOT Visual!
examples of a visual recommender


Figure 1: Screenshot of Tasteweights illustrating the main interaction features: (a) changing the weight of an item (b) restoring the default value of an item (c) removing an item (d) changing the weight of a context source (e) changing the visible portion of a context source (f) navigating context sources.
another visual recommender example

different domains, bounded or unbounded parameters

1.A. & 3.A. stackoverflow:Java & Yelp:Phoenix intelligent visual analytics
1.B. & 2.B.
2.A. & 4.A. stackoverflow:All & Yelp:All intelligent visual analytics
3.B. & 4.B.

for example:
- bounded domain: a semantic code visual analytics;
- unbounded domain: innovative exploratory visual analytics
Break traditional *list-style* of recommendation!

1.A. & 3.A.
1.B. & 2.B. stackoverflow:Java & stackoverflow:All visual recommender
2.A. & 4.A.
3.B. & 4.B. Yelp:Phoenix & Yelp:All visual recommender

for example:
- bounded domain: designing a code snippet visual recommender; local cuisine visual recommender
- unbounded domain: multi-modal visual recommender; geolocation visual recommender
Project #9: Da Vinci project: Analytical Art
Task: crawl cubism collections, utilize image processing algorithm analyze the collections, and produce an intelligent visual analytics. It can be used to detect counterfeits, analyze & understand art, facilitate art education.

Pablo Piccaso (1907) Les Demoiselles d'Avignon
How Google "Translates" Pictures into Words Using Vector Space Mathematics

Google engineers have trained a machine-learning algorithm to write picture captions using the same techniques it developed for language translation.

http://www.technologyreview.com/view/532886/how-google-translates-pictures-into-words-using-vector-space-mathematics/
Approach:

1. explore data sets
2. explore existing interactive visualizations
3. finalize problems to solve
4. prototyping
5. start coding, start writing