New Era for Construction Industry

Data to Knowledge Made Possible by Visualization
(Construction Change Order Visualization Analysis)

CPSC 533C Projects Updates By Chao-Ying Chiu
Construction industry is plagued by:
- Conditions difficult to predict
- Too many conditions influence each other
  So a construction project is difficult to control

Solutions?
- Gain more knowledge about all aspects of it through analyzing data pertinent to projects

However two bottlenecks…
- Data collection costly and incomplete
- Data collected hard to interpret
- The less you gain from data, the less you endeavor to pursue them---Vicious cycle!
Project domain, data, and tasks

- **Domain-- Change orders**
  - Frequent change orders disturb everyone’s plans
  - Little is known about them—ineffective control
  - Every one is unhappy when they are out of control—>
    everyone does not like to spend more....

- **Data**
  - Data collected pertinent to change orders process

- **Tasks**
  - Prove that data is interpretable and useful if visualized
  - Analyze principles of using information visualization in the construction industry domain
Project Approach

- **Selection of information visualization**
  - Analyze characteristics of domain, data, task
    - Literature review of domain
    - Clean up and organize data at hands
    - Scope tasks
  - **Analyze visualization requirement**
    - Literature review of related visualization techniques
  - **Identify available visualization tools**
- **Exploratory data analysis**
  - Iterate steps in phase 1 if necessary
- **Critiques selected solutions VS current practice**
What have been found so far (1)…

- **In Terms of Domain**
  - Lots of “what to do”, but few about why
  - Complex interactions in the domain remain unknown
  - Practitioners might need to explore questions before they can ask questions

- **In Terms of Data**
  - Seems to have mysterious “Time” property of data:
    - “Timing” roles played in cause effect
    - Correlations between “Time” attributes and other attributes
  - Lots of “Texts” data need pre-process
  - But basically is safe to organize them into relations
<table>
<thead>
<tr>
<th>General Contractor</th>
<th>Site Instruction or Request of Information Number</th>
<th>Change Order Description</th>
<th>Date Issued</th>
<th>Projected Cost</th>
<th>Date Approved</th>
<th>Approved Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>nominal</strong></td>
<td><strong>nominal</strong></td>
<td><strong>“text”</strong></td>
<td><strong>date</strong></td>
<td><strong>quantity</strong></td>
<td><strong>date</strong></td>
<td><strong>quantity</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>nominal</strong></td>
</tr>
<tr>
<td>Request for Information</td>
<td>Site Instruction</td>
<td>Date Sent</td>
<td>Date Due</td>
<td>Date Replied</td>
<td>Comments</td>
<td>Site Instruction Number</td>
</tr>
<tr>
<td><strong>nominal</strong></td>
<td><strong>“text”</strong></td>
<td><strong>date</strong></td>
<td><strong>date</strong></td>
<td><strong>date</strong></td>
<td><strong>nominal</strong></td>
<td><strong>nominal</strong></td>
</tr>
</tbody>
</table>
In Terms of Tasks
- Better to focus on finding unknown facts and their corresponding visualization scenario
- And then summarize good practices and principles

In Terms of Visualization Requirement
- Let users quickly “try” data—>Retrieve and show data simultaneously
- Increase possibility of seeing something—>Different ways of showing the same data
- Innovative ways of showing “Time” property of data
What have been found so far (3)…

**In Terms of Tools Selections**

- **Tableau**
  - Visually query databases
  - Surprising images appear out of your expectations, but flexible for you to try another way
  - However, difficult to stuff many dimensions of data in a single table

- **Advizor**
  - Use claimed exhaustive 15 charts to automatically retrieve and show data for users
  - Different charts can be juxtaposed for comparisons
  - However, is it really exhaustive and effective?

- **TimeSearcher or Lifelines (not decided yet)**
Advizor's 15 charts

<table>
<thead>
<tr>
<th>Region</th>
<th>Revenue(sum)</th>
<th>Margin % (avg)</th>
<th># Customers(sum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South</td>
<td>87,188,143.00</td>
<td>10.34</td>
<td>1,742</td>
</tr>
<tr>
<td>West</td>
<td>153,660,754.00</td>
<td>9.90</td>
<td>2,037</td>
</tr>
<tr>
<td>East</td>
<td>35,051,186.00</td>
<td>11.42</td>
<td>1,108</td>
</tr>
<tr>
<td>Midwest</td>
<td>51,544,817.00</td>
<td>11.16</td>
<td>1,152</td>
</tr>
<tr>
<td>OVERALL</td>
<td>327,444,900.00</td>
<td>10.33</td>
<td>6,039</td>
</tr>
</tbody>
</table>

Region Filter
- [ ] (All)
- [ ] (None)
- [ ] East
- [ ] Midwest
- [ ] South
- [ ] West

Statistics:
- Age: [Graphs showing age distribution]
- Brand Loyalty: [Graphs showing brand loyalty]
- Income: [Graphs showing income distribution]
- OilCo: [Graphs showing oil company distribution]
What have been changed (1)…

- Exploratory data analysis phase
  - Focus on understanding the exploration scenario of finding facts:
    - Step1: A domain expert identify what data at hands
    - Step2: Use his knowledge to visually select data
    - Step3: Inspired by images of data shown and facts revealed
    - Step4: Iterate step2 to step3
  - Rather than answering pre-defined questions
What have been changed (2)…

- **Critique phase**
  - Focus on identifying which features and images more suitable for domain information needs
  - Rather than identifying pros and cons (it depends on applications)
  - The facts that the tools identified are affordable and apparently outwit Excel. No need to do cost/benefit analysis and to compare with Excel.