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

Background

- **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 pursuit 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

Change Order Registry									
General Contractor Change Order Number	Site Instruction or Request of Information Number	Change Order Description	Date Issued	Projected Cost	Date Approved	Approved Amount	Trade Change Order Issued		
							Trade Name	Revision Number	Issued Amount
nominal	*nominal*	**text**	***date***	*quantity*	***date***	*quantity*	*nominal*	*nominal*	*quantity*

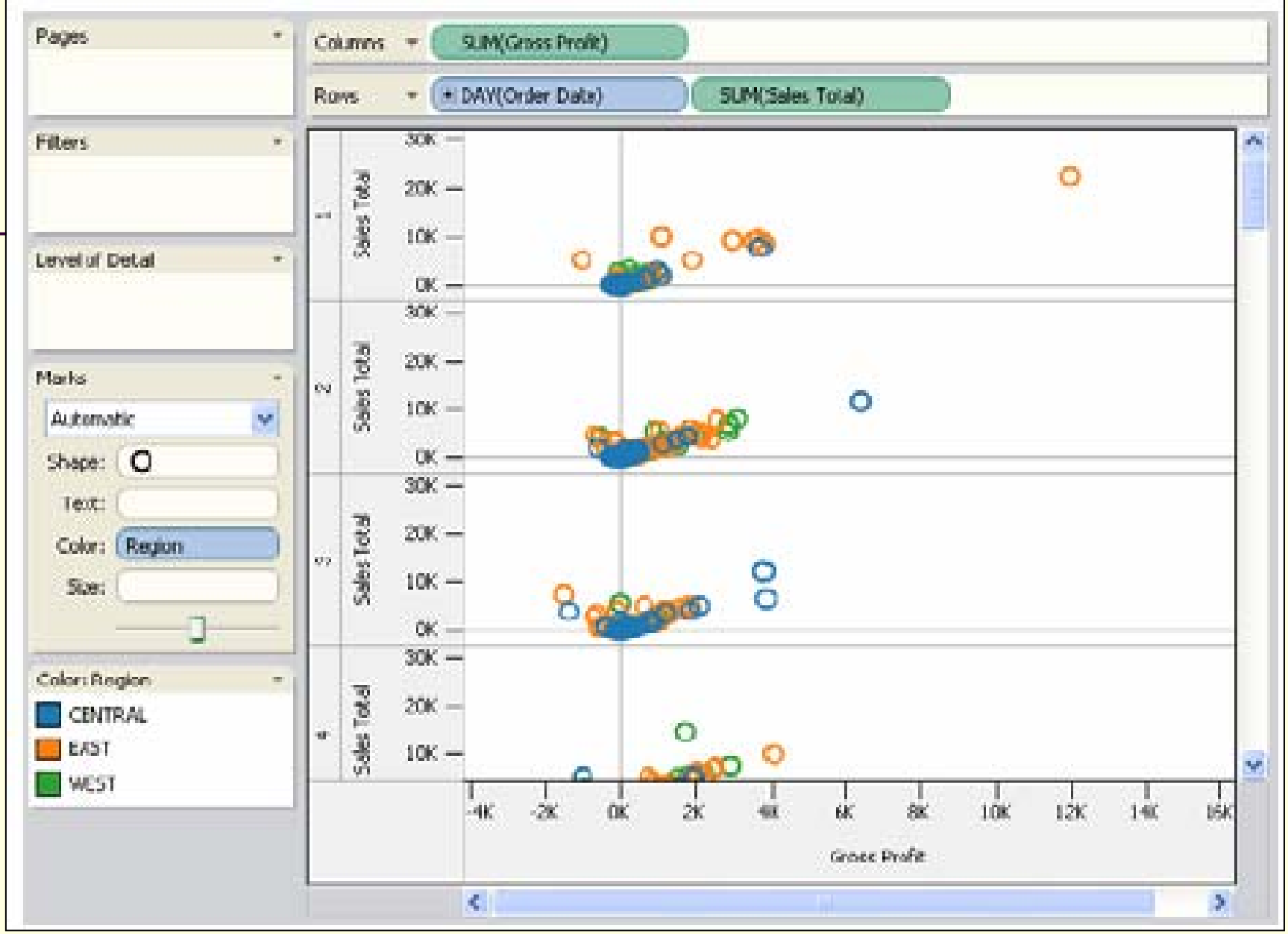
Request for Information						Site Instruction				
Request for Information Number	Request for Information Description	Date Sent	Date Due	Date Replied	Comments	Site Instruction Number	Issued By	Site Instruction Description	Attachement	Date Issued
nominal	**text**	***date***	***date***	***date***	*nominal*	*nominal*	*nominal*	**text**	*nominal*	***date***

What have been found so far (2)…

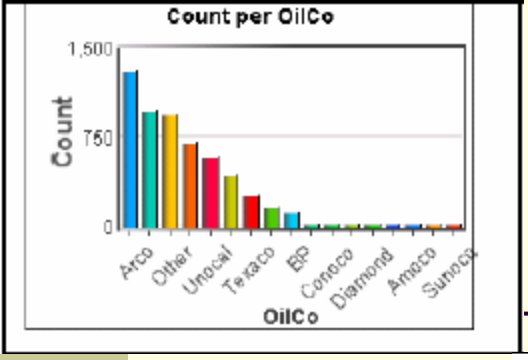
- **In Terms of Tasks**
  - Better to focus on finding unknown facts and their corresponding visualization scenario
  - And then summarize good practices and principles
- **In Terms of Visualization Requirement**
  - Let users quickly “try” data->Retrieve and show data simultaneously
  - Increase possibility of seeing something->Different ways of showing the same data
  - Innovative ways of showing “Time” property of data

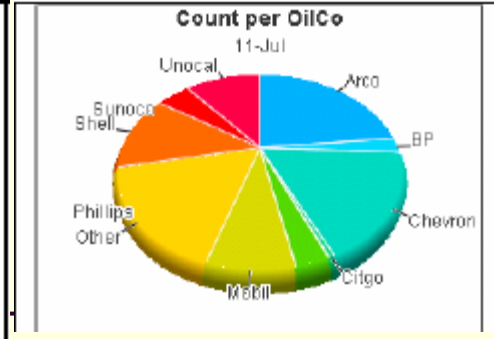
What have been found so far (3)…

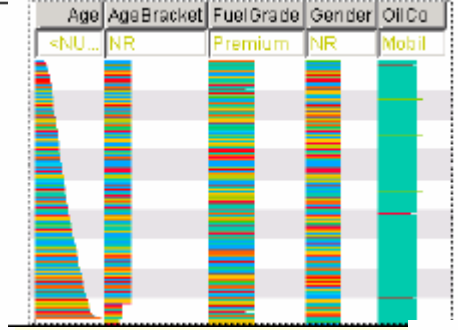
- **In Terms of Tools Selections**
  - Tableau
    - Visually query databases
    - Surprising images appear out of your expectations, but flexible for you to try another way
    - However, difficult to stuff many dimensions of data in a single table
  - Advizor
    - Use claimed exhaustive 15 charts 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)

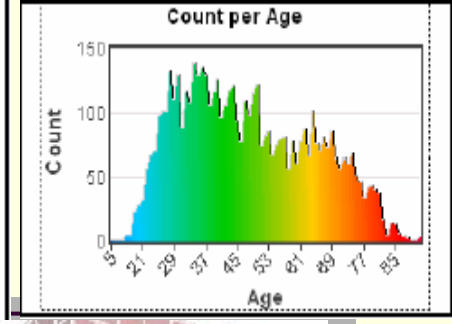

Pages
Filters
Level of Detail
Marks
Automatic
Shape:
Text:
Color:
Region
Size:
Color: Region
CENTRAL
EAST
WEST
Columns
SUM(Gross Profit)
Rows
DAY(Order Date)
SUM(Sales Total)
Sales Total
30K
20K
10K
0K
-4K
-2K
0K
2K
4K
6K
8K
10K
12K
14K
16K
Gross Profit

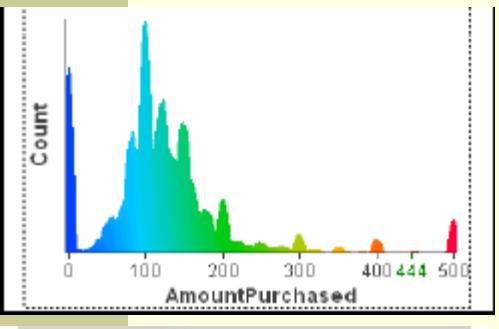
Tableau

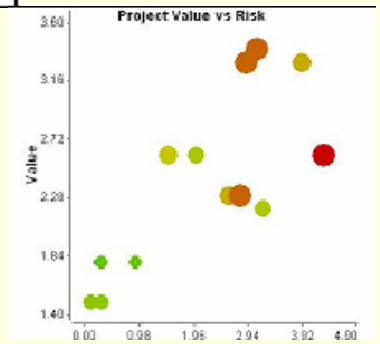


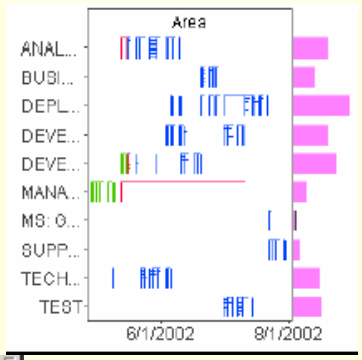


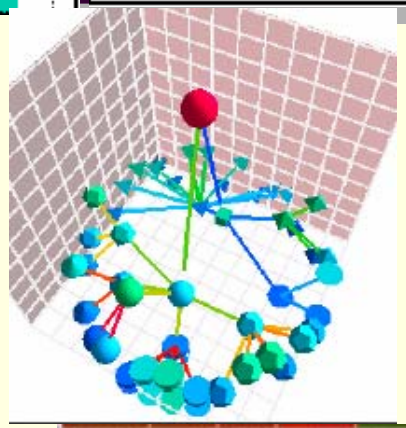


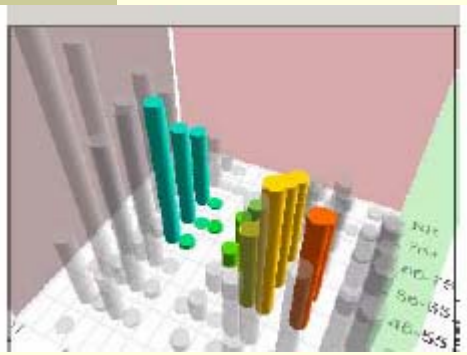


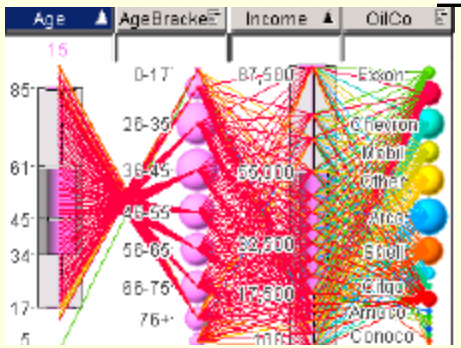


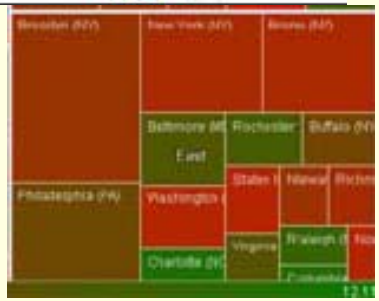












Region	Revenue(sum)	Margin %(avg)	# Customers(sum)
South	87,188,143.00	10.34	1,742
West	153,660,754.00	9.90	2,037
East	35,051,186.00	11.42	1,108
Midwest	51,544,817.00	11.16	1,152
OVERALL	327,444,900.00	10.55	6,039

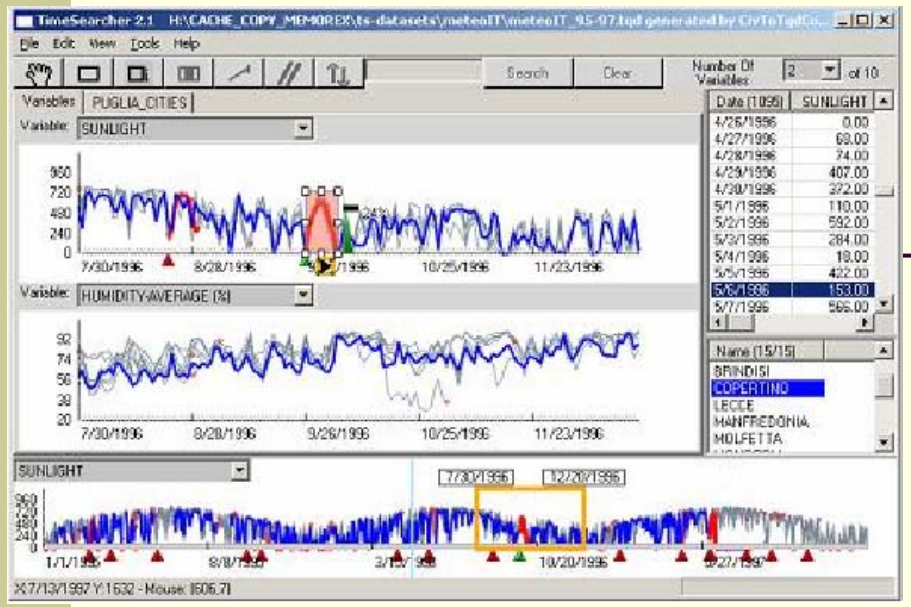
Region Filter

☐ (All) ☐ (None)

☐ East
☑ Midwest
☐ South
☑ West

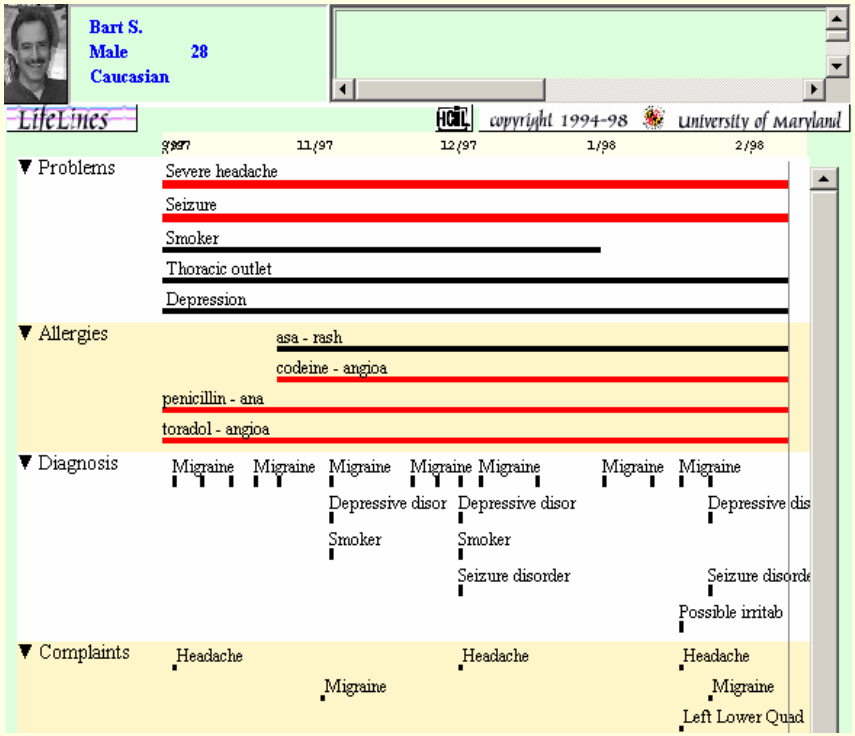
Statistic	Age	BrandLoyalty	Income	OilCo
	Inte...	String	Integer	String
Count	5,612	5,612	5,612	5,612
Selected	5,612	5,612	5,612	5,612
Excluded				
Unique	79	4	16	17
Uniq Sel	79	4	16	17
Mode	34	TwoOrThree	55,000	Arco
Mode Sel	34	TwoOrThree	55,000	Arco
Mean	48		37,564	
Mean Sel	48		37,564	
Std Dev	17		26,211	

Advizor's 15 charts



TimeSearcher 2

LifeLines



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.