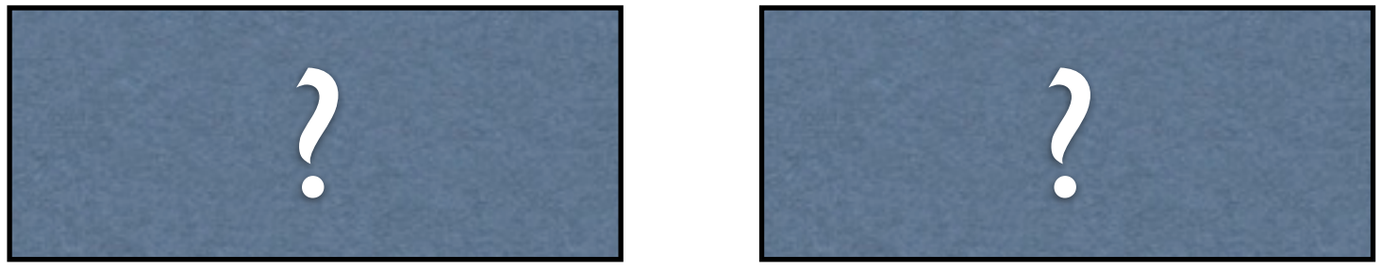
CSCI568

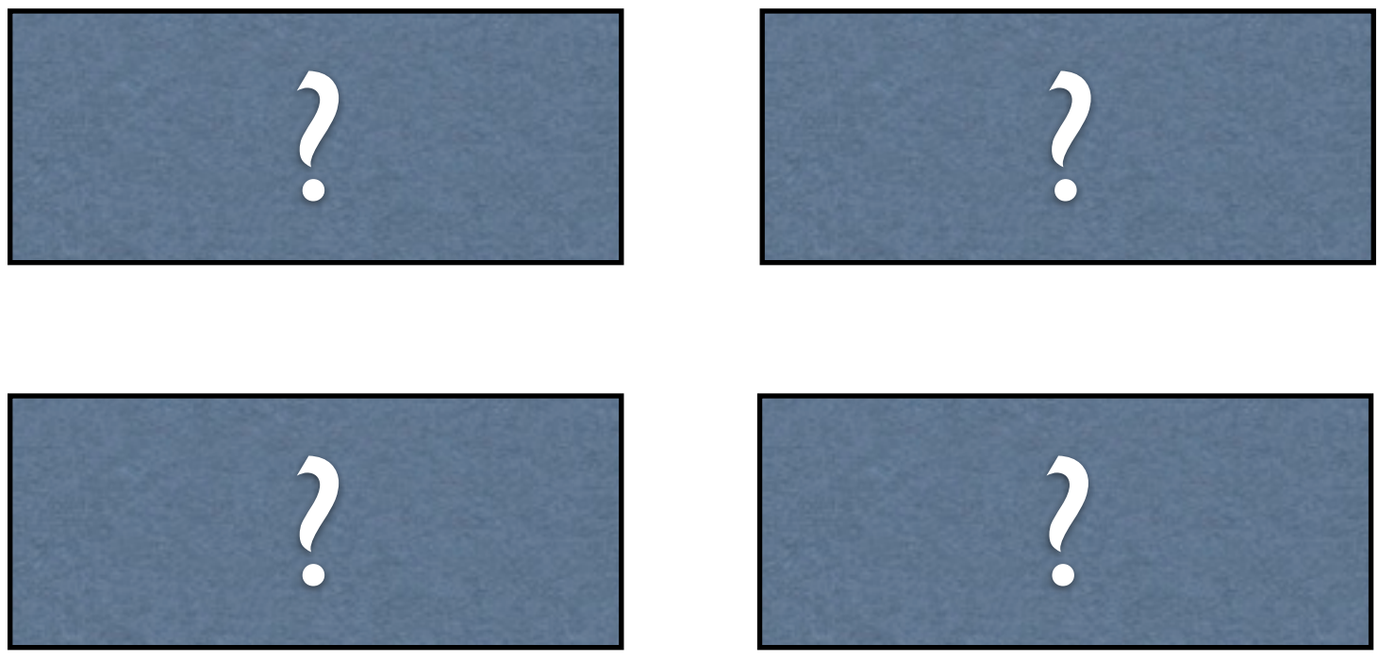
Discussion 3: Data Mining from 10,000 ft

Organize Your Mind

Two general purposes



Four families of tasks



Organize Your Mind

Two general purposes

predictive

?

Four families of tasks

?

?

?

?

Organize Your Mind

Two general purposes

predictive

descriptive

Four families of tasks

?

?

?

?

Organize Your Mind

Two general purposes

predictive

descriptive

Four families of tasks

cluster analysis

classification

association analysis

anomaly detection

Organize Your Mind

Two general purposes

predictive

descriptive

~~Four~~ families of tasks

cluster analysis

classification

exploration

association analysis

anomaly detection

visualization & interaction

Attribution: Amblin Entertainment 1985

DATA

Data Basics

- Types of data
- Quality of data
- Exploring data
  - "Summary statistics"
- Common formats in DM

Homework

- Project 2
- Reading 3