Introduction to Data Exploration

Introduction to Data Exploration

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

Objective

Explain data warehouses and exploratory querying

Exploratory Search

- learning: acquiring new knowledge
  - Knowledge acquisition, comprehension and interpretation, comparison, aggregation and integration, and socialization.
- investigating: revealing new facts.
  - Accretion, analysis, synthesis, evaluation, discovery, planning and forecasting, and transformation.

- **SENSE MAKING from the REAL WORLD**

“Sense” making…What Does it Mean?

- 1st sense: from latin “sentire” or “to perceive”
  - any of the faculties, as sight, hearing, smell, taste, or touch, by which humans and animals perceive stimuli originating from outside or inside the body
- 2nd sense: “to attain awareness or understanding of…”
  - “awareness” implies vigilance in observing or alertness in drawing inferences from what one experiences
  - “understanding” is the power to make experience intelligible by applying concepts and categories

Did You Notice the Gap?

Data-Rich World

Sense Making in a Dynamic World

Data Challenges

ISQ

- I = Imprecision
- S = Sparsity
- Q = Quality/Noise

3Vs

- V = Volume
- V = Velocity
- V = Variety

HMLE

- H = High-dimensional
- M = Multi-modal
- L = Inter-**L**inked
- E = Evolving

Human Challenges: for many applications the final consumer is **HUMAN**

Data Management/Mining Techniques

Real-World Data

Data Exploration Needs to Support Both

Effective Data Manipulation

- filtering: projection, selection,
- integration: join, nearest neighbor joins
- set operations: union and intersection

Effective Data Retrieval and Analysis

- feature extraction,
- similarity search (top-k, range, skyline, nearest-neighbor)
- clustering, partitioning
- aggregation, summarization
- classification, latent analysis
- preference-driven retrieval