How to handle complexity: 1 previous strategy + 3 more

• derive new data to show within view
• change view over time
• facet across multiple views
• reduce items/attributes within single view

Reminder
This table has 100 million items
• pair up, discuss how to have scalable approach, create sketch to illustrate
– [8 min]
– socrative: true when done

Filter & Aggregate

Idiom: FilmFinder
• dynamic queries/filters for items
– tightly coupled interaction and visual encoding idioms, so user can immediately see results of action

Idiom: Crossfilter
• item filtering
– coordinate axes/controls combined
– all scanned histogram builders update when any ranges change

Idiom: Small multiples
• encoding same
data: none shared
– idiom: crossfilter
different items
different condition keys,
same graph keys), same attributes
expression values for node colors
– (same network layout for nodes/games)
navigation: shared
Idiom: histogram

- static item aggregation
- task: find distribution
- data: table
- derived data
  - new table: keys are bins, values are counts
  - pattern can change dramatically depending on discretization
- bin size crucial
- opportunity for interaction: control bin size on the fly

Idiom: scented widgets

- augmented widgets show information scent
- better cues for information foraging: show whether
- in drilling down further vs. looking elsewhere
- concise use of space: histogram on slider
Distortion costs and benefits
- benefits
  - combine focus and context information in single view
- costs
  - length comparisons impaired
  - network/tree topology
    - comparisons unfeasible
    - connections, constraints
  - effects of distortion unclear
    - original structure unclear
    - object constancy/trackability maybe impaired

Capturing & using material reflectance
- reflectance measurement: interaction of light with real materials (spheres)
  - result: 104 high-res images of material
  - each image 4MP pixels
  - goal: image synthesis
    - simulate completely new materials
  - need for more concise model
    - 104 images  4MP pixels  = 400M dims
    - want concise model with meaningful limbs
    - how shiny/greasy/metallic
    - DR to the rescue!

Linear DR
- first try: PCA (linear)
  - result: error falls off sharply after ~45 dimensions
  - scree plot: error vs number of dimensions in lowD projection
  - problem: physically impossible intermediate points when simulating new materials
    - specular highlights cannot have holes!

Nonlinear DR
- second try: charting (nonlinear DR technique)
  - scree plot suggests 10-15 dims
  - note: dim estimate depends on technique used!

Finding semantics for synthetic dimensions
- look for meaning in scatterplots
  - synthetic dims created by algorithm but named by human analysis
  - points represent real-world images (spheres)
  - people inspect images corresponding to points to decide if set could have meaningful name
  - cross-check meaning
    - arrows show simulated images (teapots) made from model
    - check if those match dimension semantics

Understanding synthetic dimensions
- plot: Specular-Metallic vs Diffuseness-Glossiness
  - meaningful name
  - specular highlights cannot have holes!

Capturing & using material reflectance
- reflectance measurement: interaction of light with real materials (spheres)
  - result: 104 high-res images of material
  - each image 4MP pixels
  - goal: image synthesis
    - simulate completely new materials
  - need for more concise model
    - 104 images  4MP pixels  = 400M dims
    - want concise model with meaningful limbs
    - how shiny/greasy/metallic
    - DR to the rescue!

Linear DR
- first try: PCA (linear)
  - result: error falls off sharply after ~45 dimensions
  - scree plot: error vs number of dimensions in lowD projection
  - problem: physically impossible intermediate points when simulating new materials
    - specular highlights cannot have holes!

Nonlinear DR
- second try: charting (nonlinear DR technique)
  - scree plot suggests 10-15 dims
  - note: dim estimate depends on technique used!

Finding semantics for synthetic dimensions
- look for meaning in scatterplots
  - synthetic dims created by algorithm but named by human analysis
  - points represent real-world images (spheres)
  - people inspect images corresponding to points to decide if set could have meaningful name
  - cross-check meaning
    - arrows show simulated images (teapots) made from model
    - check if those match dimension semantics

Understanding synthetic dimensions
- plot: Specular-Metallic vs Diffuseness-Glossiness
  - meaningful name
  - specular highlights cannot have holes!

Capturing & using material reflectance
- reflectance measurement: interaction of light with real materials (spheres)
  - result: 104 high-res images of material
  - each image 4MP pixels
  - goal: image synthesis
    - simulate completely new materials
  - need for more concise model
    - 104 images  4MP pixels  = 400M dims
    - want concise model with meaningful limbs
    - how shiny/greasy/metallic
    - DR to the rescue!

Linear DR
- first try: PCA (linear)
  - result: error falls off sharply after ~45 dimensions
  - scree plot: error vs number of dimensions in lowD projection
  - problem: physically impossible intermediate points when simulating new materials
    - specular highlights cannot have holes!

Nonlinear DR
- second try: charting (nonlinear DR technique)
  - scree plot suggests 10-15 dims
  - note: dim estimate depends on technique used!

Finding semantics for synthetic dimensions
- look for meaning in scatterplots
  - synthetic dims created by algorithm but named by human analysis
  - points represent real-world images (spheres)
  - people inspect images corresponding to points to decide if set could have meaningful name
  - cross-check meaning
    - arrows show simulated images (teapots) made from model
    - check if those match dimension semantics

Understanding synthetic dimensions
- plot: Specular-Metallic vs Diffuseness-Glossiness
  - meaningful name
  - specular highlights cannot have holes!

Capturing & using material reflectance
- reflectance measurement: interaction of light with real materials (spheres)
  - result: 104 high-res images of material
  - each image 4MP pixels
  - goal: image synthesis
    - simulate completely new materials
  - need for more concise model
    - 104 images  4MP pixels  = 400M dims
    - want concise model with meaningful limbs
    - how shiny/greasy/metallic
    - DR to the rescue!

Linear DR
- first try: PCA (linear)
  - result: error falls off sharply after ~45 dimensions
  - scree plot: error vs number of dimensions in lowD projection
  - problem: physically impossible intermediate points when simulating new materials
    - specular highlights cannot have holes!

Nonlinear DR
- second try: charting (nonlinear DR technique)
  - scree plot suggests 10-15 dims
  - note: dim estimate depends on technique used!