

The Decision Boundary Map

An Interactive Visual Interface for Making Informed Decisions
and Selections in the Presence of Tradeoffs

Klaus Mueller and Shenghui Cheng

Visual Analytics and Imaging Lab
Computer Science Department
Stony Brook University

Prologue

Decisions are often rendered in the presence of **tradeoffs**

- however, tradeoffs are often difficult to recognize and balance
- especially when there are many factors playing a role

Decisions are often rendered in the presence of **many factors**

- there's high potential to overwhelm the human decision maker
- they might render decisions that they are not fully sure about
- only to recognize later that better choices could have been made

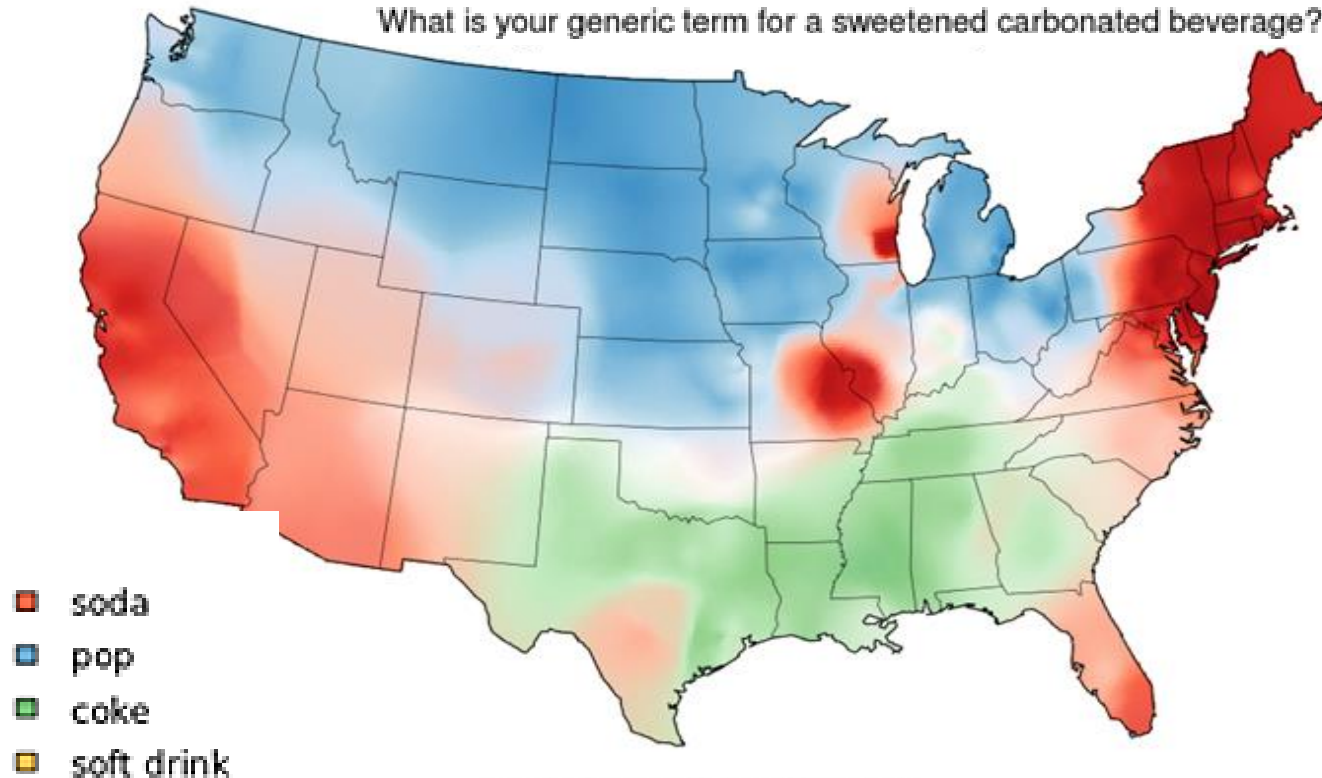
We propose a **visual interface** to ease this pain

The Data Deluge

We Need Something Intuitive



How About Something Like A Map?

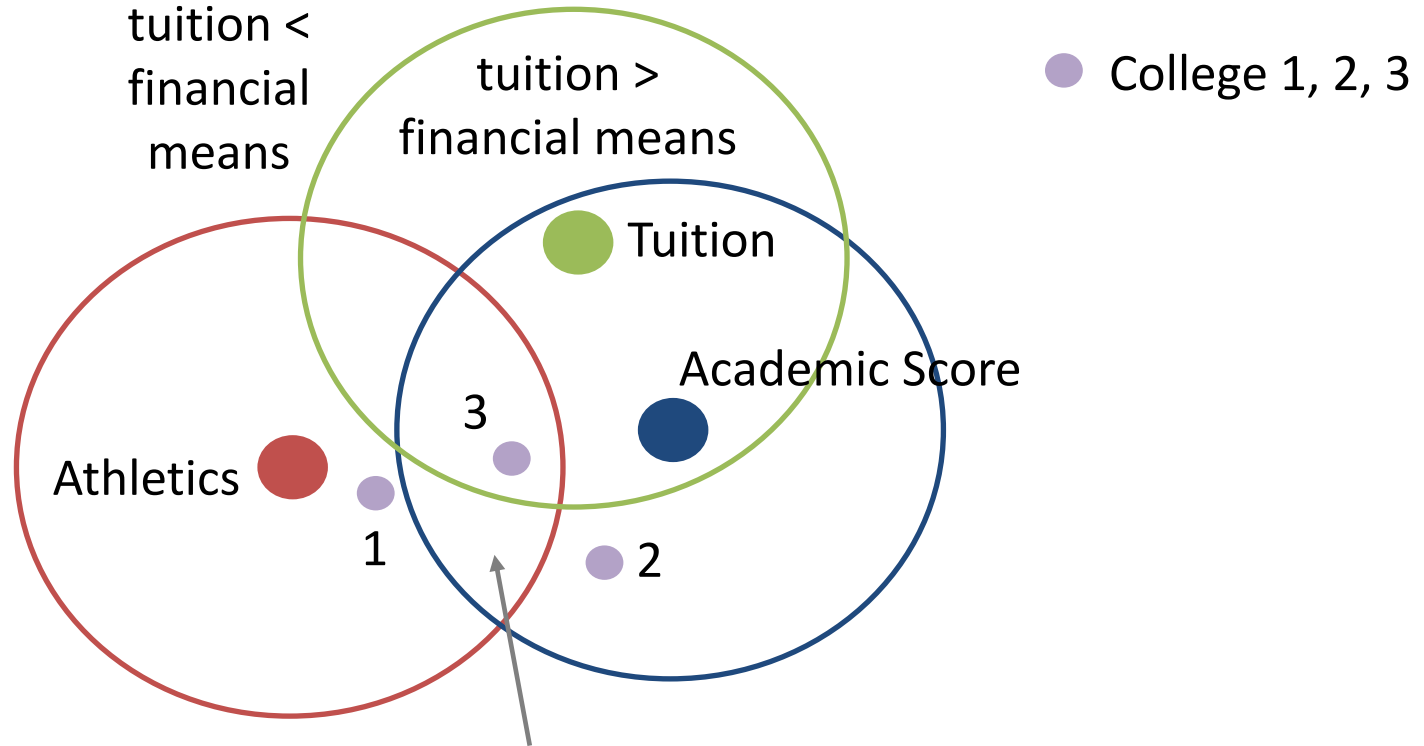


Example: Decide on a College

Many factors (and possible tradeoffs):

- athletics
- academics
- students per faculty
- social life
- tuition
- safety
- weather
- size
- and many more...

Concept: Visualize Using Decision Circles



no dream school here: good athletics, low tuition, high academic score

Interlude: The Data Matrix

Each college is an N -dimensional vector of attributes

- the data matrix has M rows of colleges each with N attributes
- the M colleges reside in an N -dimensional space of attributes

Goal is to flatten this N -D space into a 2D map with colleges

- this is called *embedding*
- embedding is essentially an optimization problem
- preserve the N -dimensional distance relations in the 2D layout

Interlude: Space Embedding

General idea:

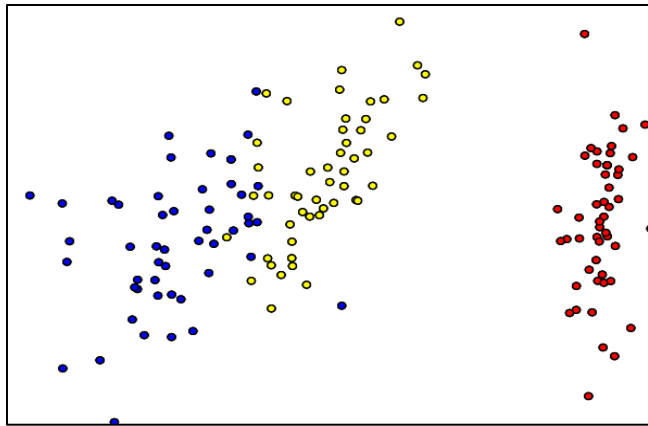
- preserve N-D space distances δ_{ij} in 2-D space d_{ij}

- minimize

$$stress = \sqrt{\frac{\sum_{ij} (d_{ij} - \delta_{ij})^2}{\sum_{ij} \delta_{ij}^2}}$$

- Multi-Dimensional Scaling (MDS)
- similar data map to similar places
→ Similarity Map

- Universities
- Ivy League
- Regional



But...



...are these schools so different?

Space Embedding: Shortcomings

Embedding the colleges is only half the information

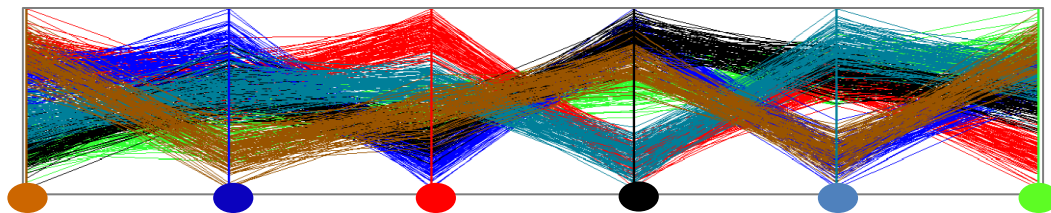
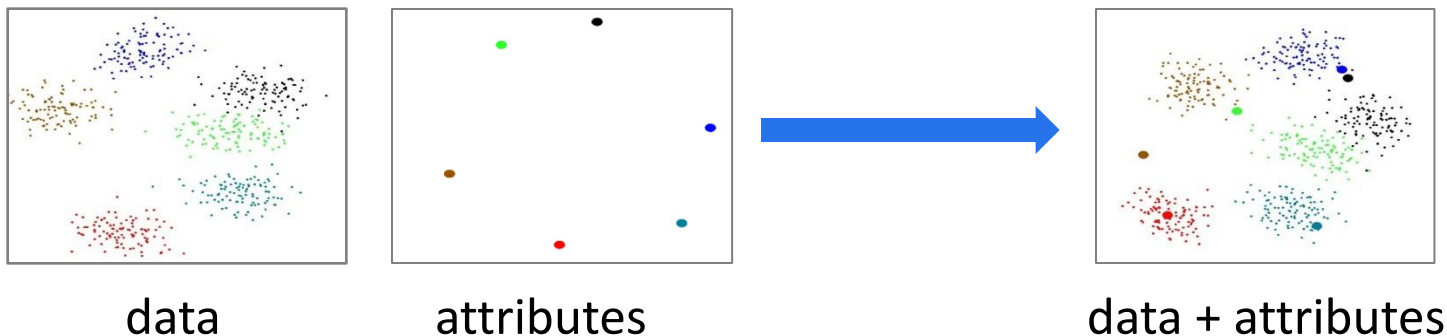
- we need to embed the attributes as well
- then we can tell what makes the colleges different
- it gives the decision context
- this requires an extension/augmentation to the data matrix

Introducing: the joint data-attribute matrix

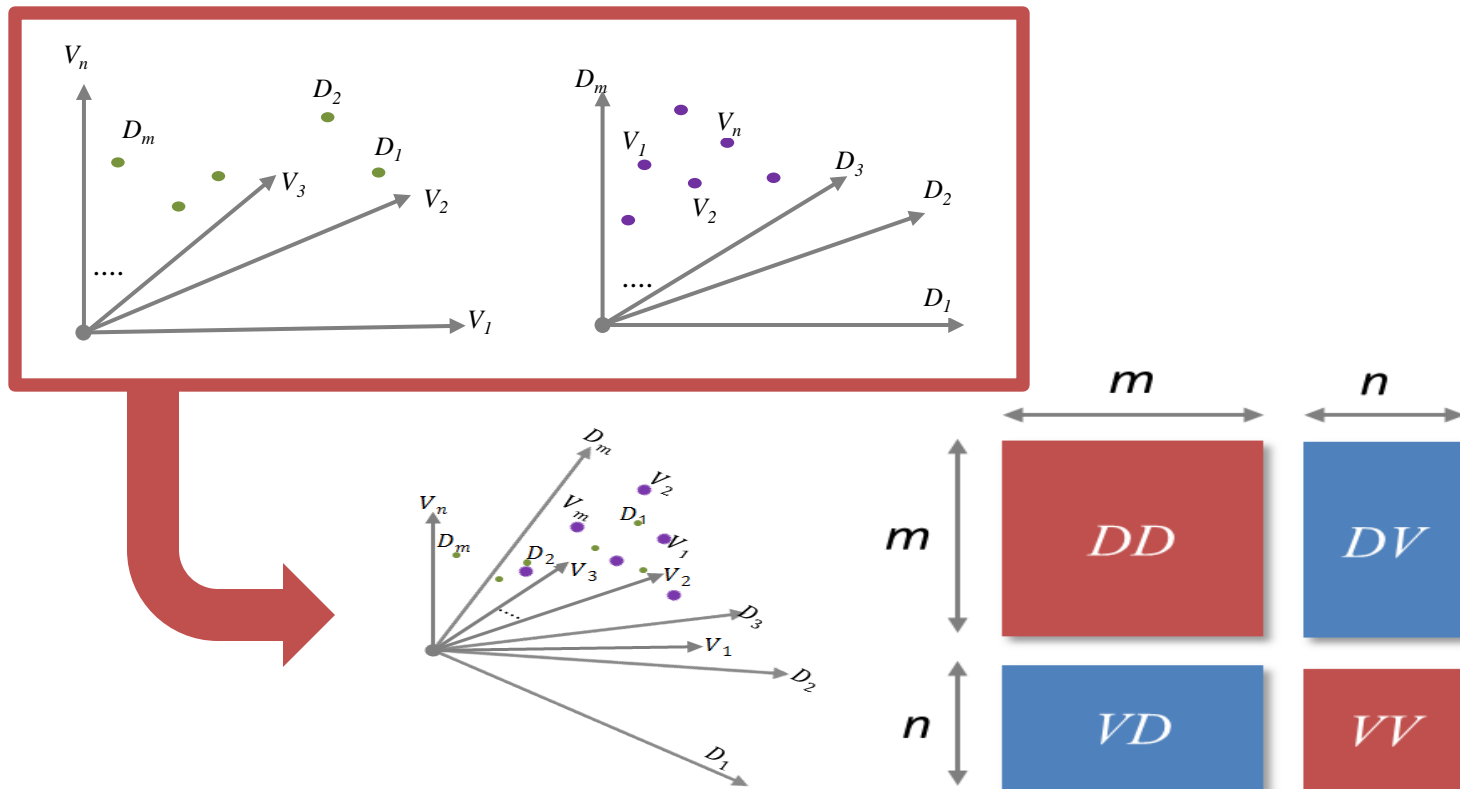
The Joint Data-Attribute Matrix

Best of both worlds

- similarity map of the data is based on vector similarity
- similarity map of the attributes is based on pairwise correlation



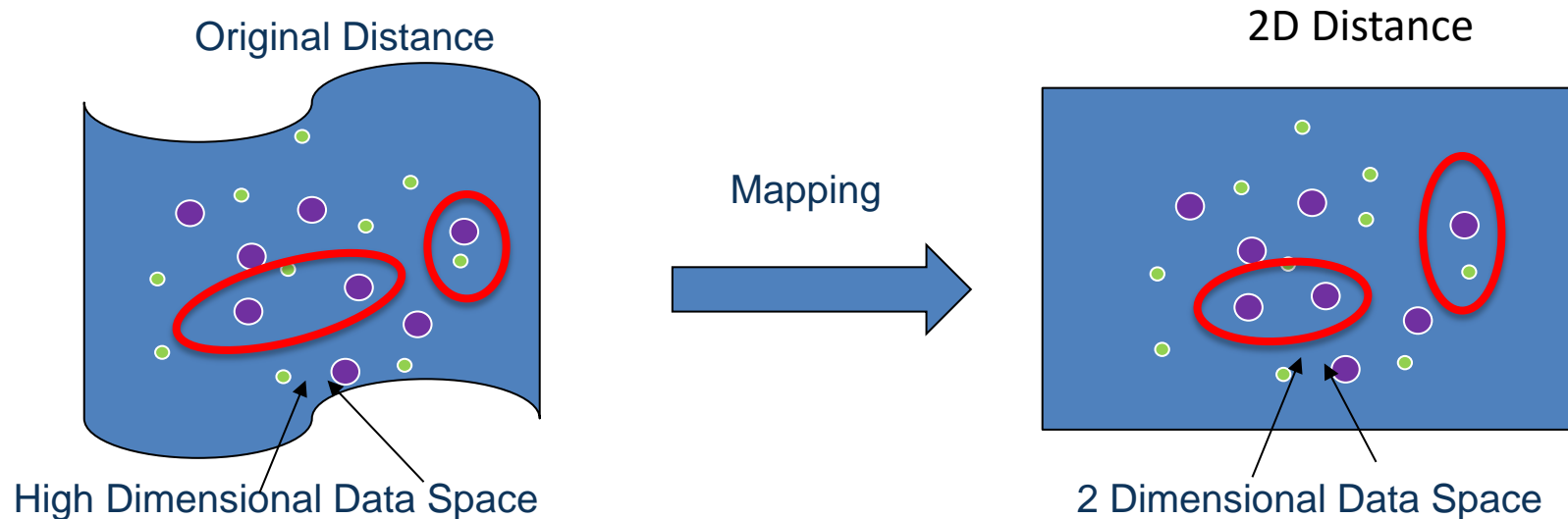
Achieved by Joint Matrix Optimization



Flexible MDS Layout Schedules

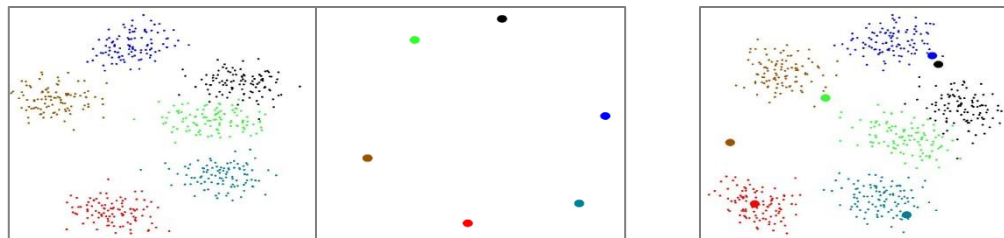
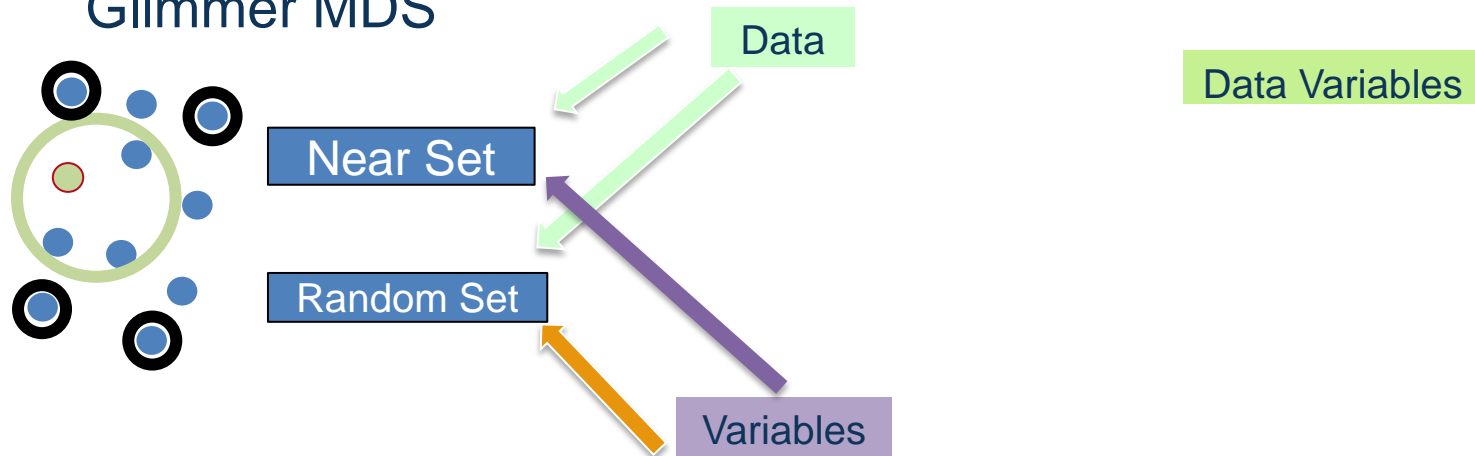
Optimization schedules for data and attribute layout can vary

- allows us to steer accuracy of the layout to preserve data relationships of attribute relationships or achieve a balance



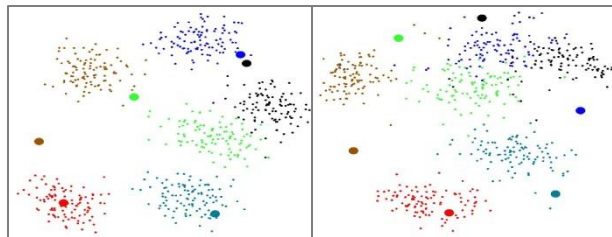
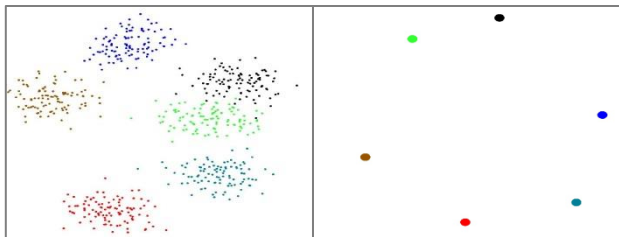
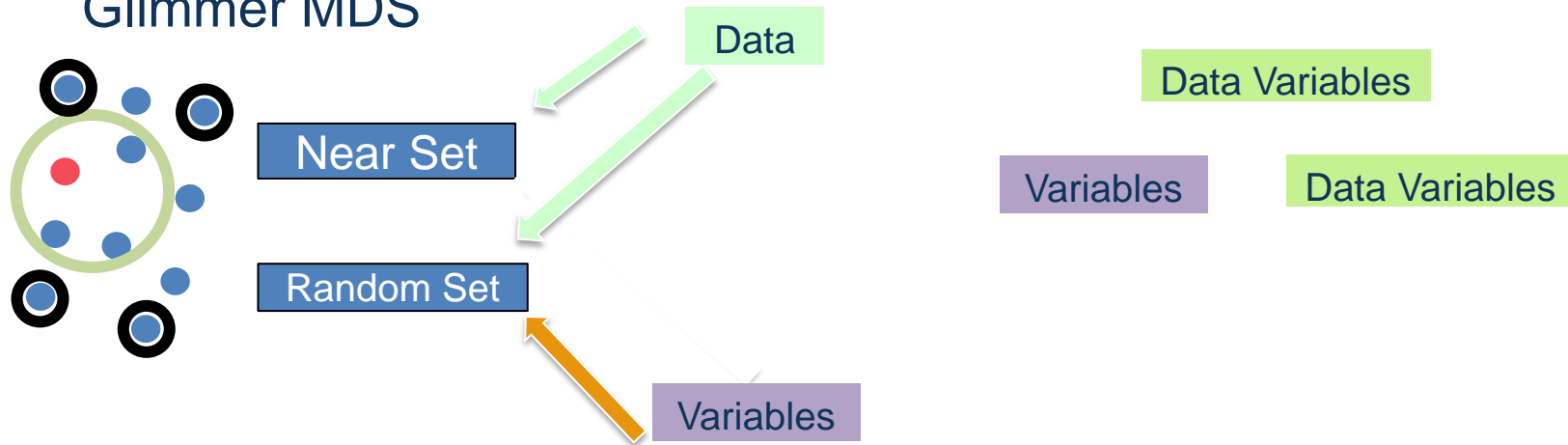
MDS Layout Schedules

Glimmer MDS



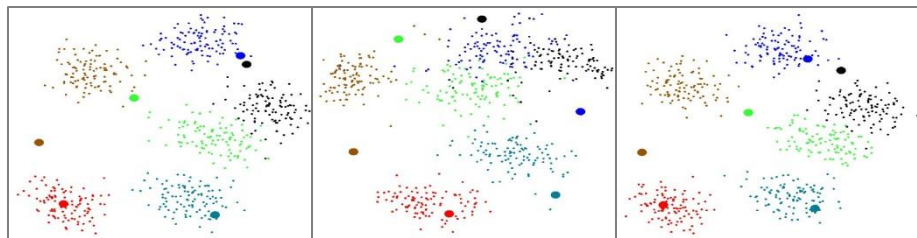
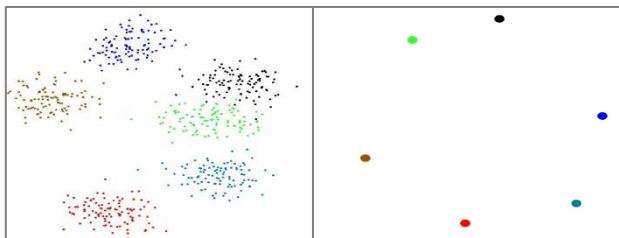
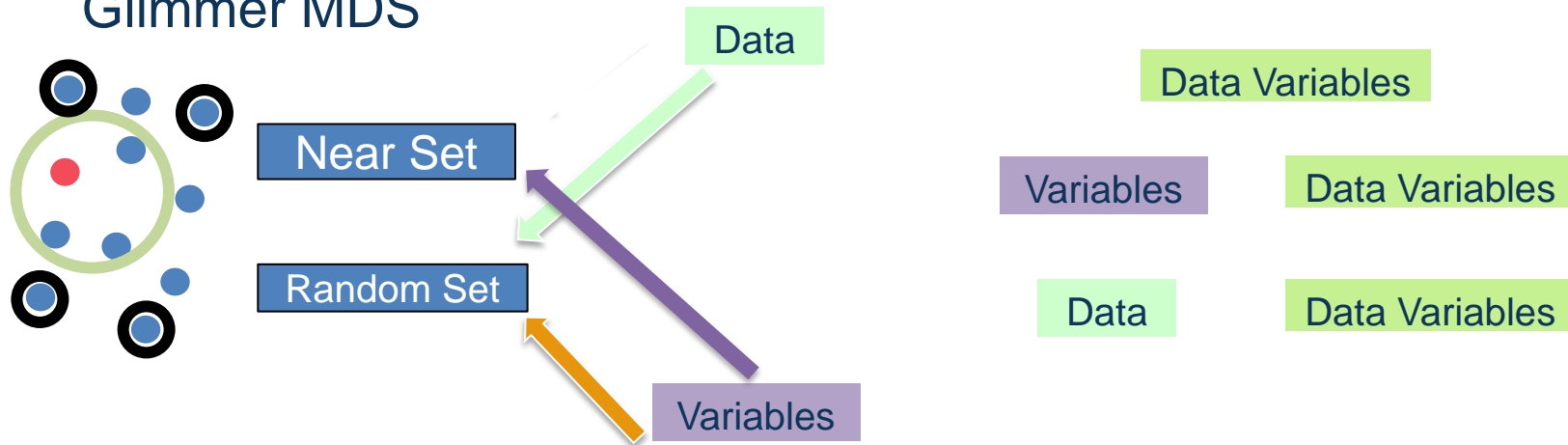
MDS Layout Schedules

Glimmer MDS



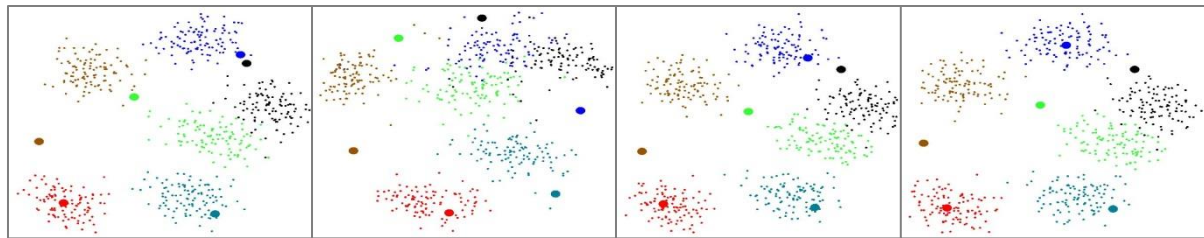
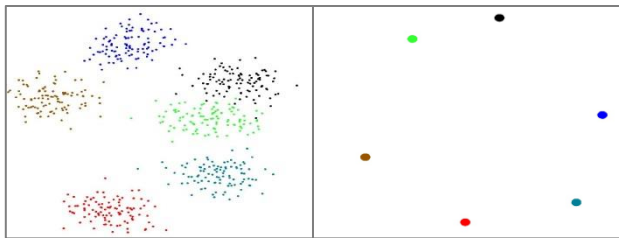
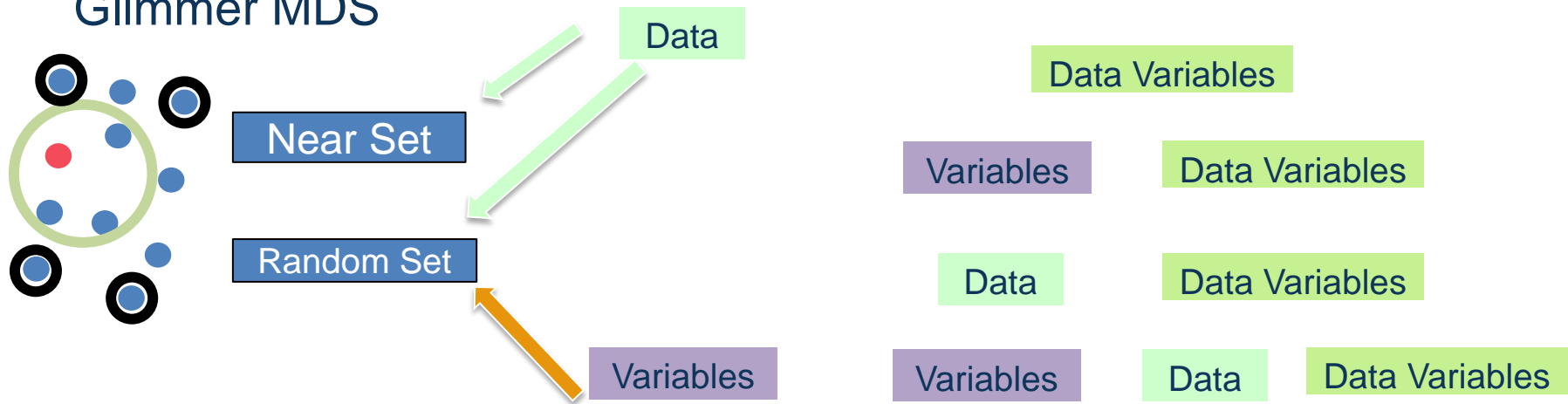
MDS Layout Schedules

Glimmer MDS



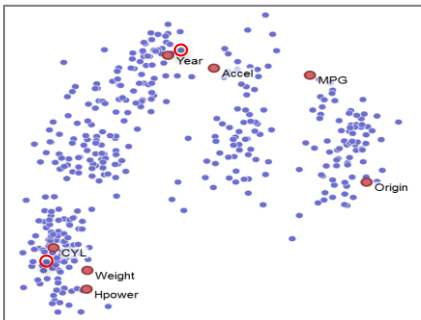
MDS Layout Schedules

Glimmer MDS



Constructing The Map

Chevrolet Cavalier
MPG: 34
CYL: 4
Hpower: 88
Weight: 2395
Accel: 18
Year: 82
Origin: 1



Pontiac Catalina
MPG: 16
CYL: 8
Hpower: 170
Weight: 4668
Accel: 11.5
Year: 75
Origin: 1

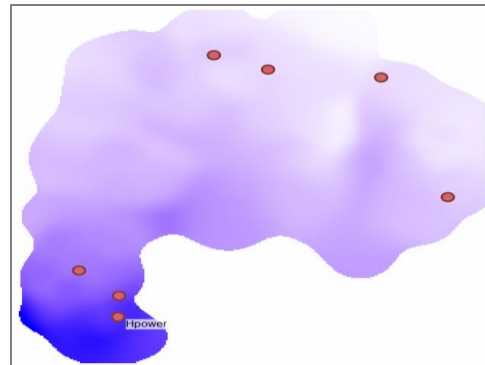
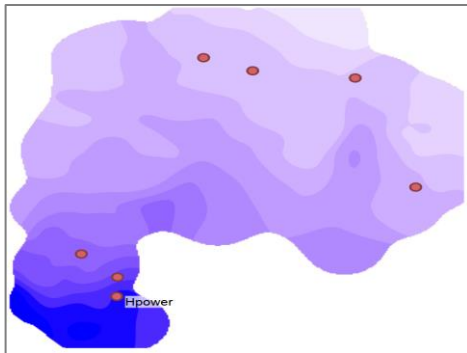
Adaptive Kernel
Density Estimation



Nadaraya-Watson
kernel regression



Contour boundary



Online Deployment

Our application runs interactively in modern web browsers

- data processing on back-end server, interactive graphics on client
- can be accessed anywhere and anytime, free of charge (for now)
- users can simply upload a spreadsheet with the data matrix
- application for a domain name .net and .com is in process
- a rudimentary version is already online and will have been fully developed at the time of the workshop



Use Cases

General public (based on data we have already visualized):

- potential car buyers selecting cars fitting their preferences
- college-age kids and parents selecting the best-fitting college
- wine shoppers navigating the confusing landscape of wines
- investors picking stocks in light of specific investment profiles

Researchers (based on inquiries we have received):

- consulting firms for foundations and nonprofits to inform clients
- disease researchers to gain insight into the interaction of traits

Use Case: Deciding on the Best College

Data are a fusion of two datasets:

- College rankings from US News and World Report
- College Prowler website

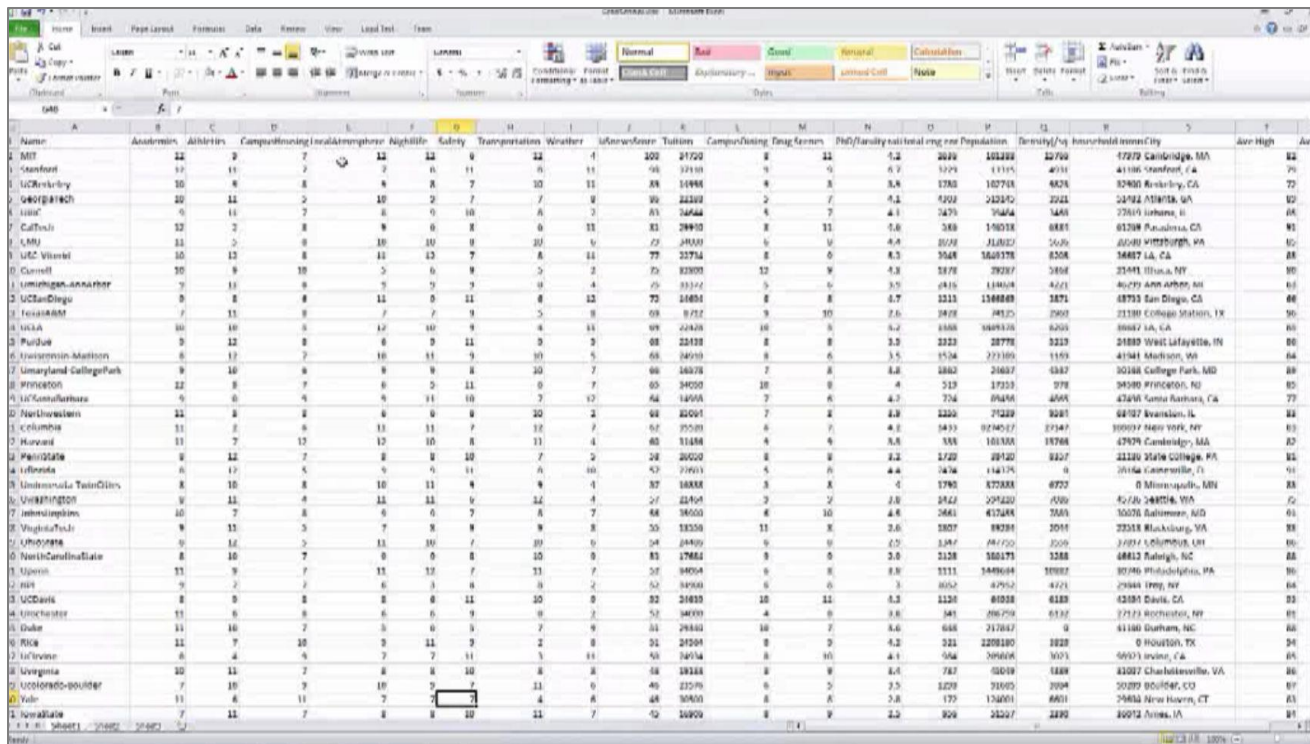
A video is online [here](#)

- some annotated stills are presented in the following

The Following Slides are Stills of the Video

Spreadsheet of University Data

Overwhelming and difficult to rank and weigh tradeoffs

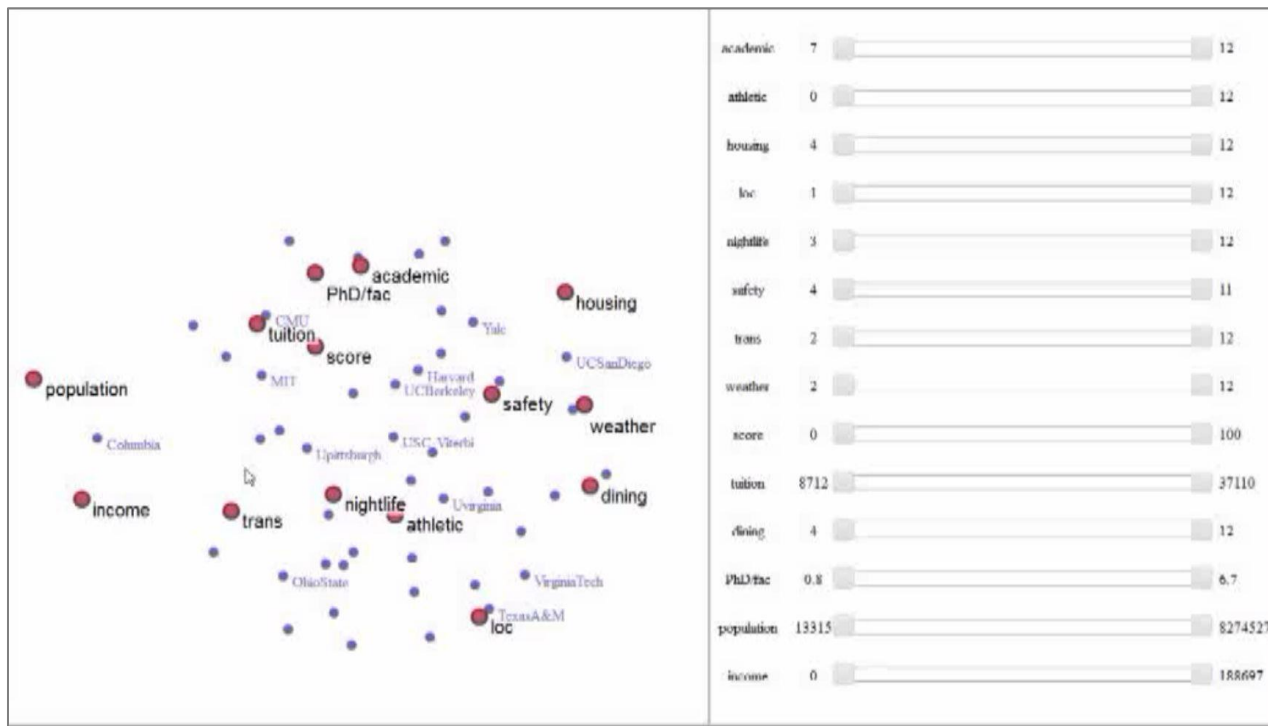


Name	Academic	Athletics	Campus	Housing	Local	Atmosphere	Nightlife	Safety	Transportation	Weather	Life	Tuition	Campus	Dining	Drug	Gender	PhD/Graduate	National	eng	pop	Population	Density/sq	hours	school	mean	City	Ave	High	Ave
MIT	12	9	7	12	12	9	12	9	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
Stanford	12	11	7	7	7	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
UC Berkeley	10	8	8	8	8	7	10	11	10	11	11	10	11	10	11	11	10	11	11	11	11	11	11	11	11	11	11	11	11
Harvard	10	8	11	9	10	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Yale	9	11	7	8	8	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Caltech	12	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
CMU	11	9	8	8	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
UCLA	10	13	8	11	13	7	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
Georgetown	10	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
University of Michigan	7	11	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
UC San Diego	8	8	8	11	8	11	8	11	8	12	7	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
UConn	7	11	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
UVA	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Purdue	9	12	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Wisconsin	8	12	7	10	11	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
University of Maryland	9	10	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
Princeton	12	8	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of California	9	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
Northwestern	11	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
Columbia	11	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
Rutgers	11	7	12	10	8	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
Penn State	8	12	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
UFlorida	8	12	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Texas	8	10	8	10	11	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
University of Washington	9	11	8	11	11	9	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
University of Virginia	10	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of North Carolina	9	11	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Oregon	9	12	8	11	10	9	11	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
University of North Carolina	8	10	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Wisconsin	11	9	7	11	12	7	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
University of Michigan	9	2	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of California	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Texas	11	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Florida	11	10	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Wisconsin	11	7	10	9	11	9	11	9	11	9	11	9	11	9	11	9	11	9	11	9	11	9	11	9	11	9	11	9	11
University of Virginia	8	4	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of North Carolina	10	11	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
University of Colorado	7	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9
University of Utah	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11	8	11
University of Arizona	7	11	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8

spreadsheet
with
colleges as
rows and
attributes as
columns

The Colleges Shown in Attribute Context

Red nodes: attributes (decision factors); blue nodes: colleges

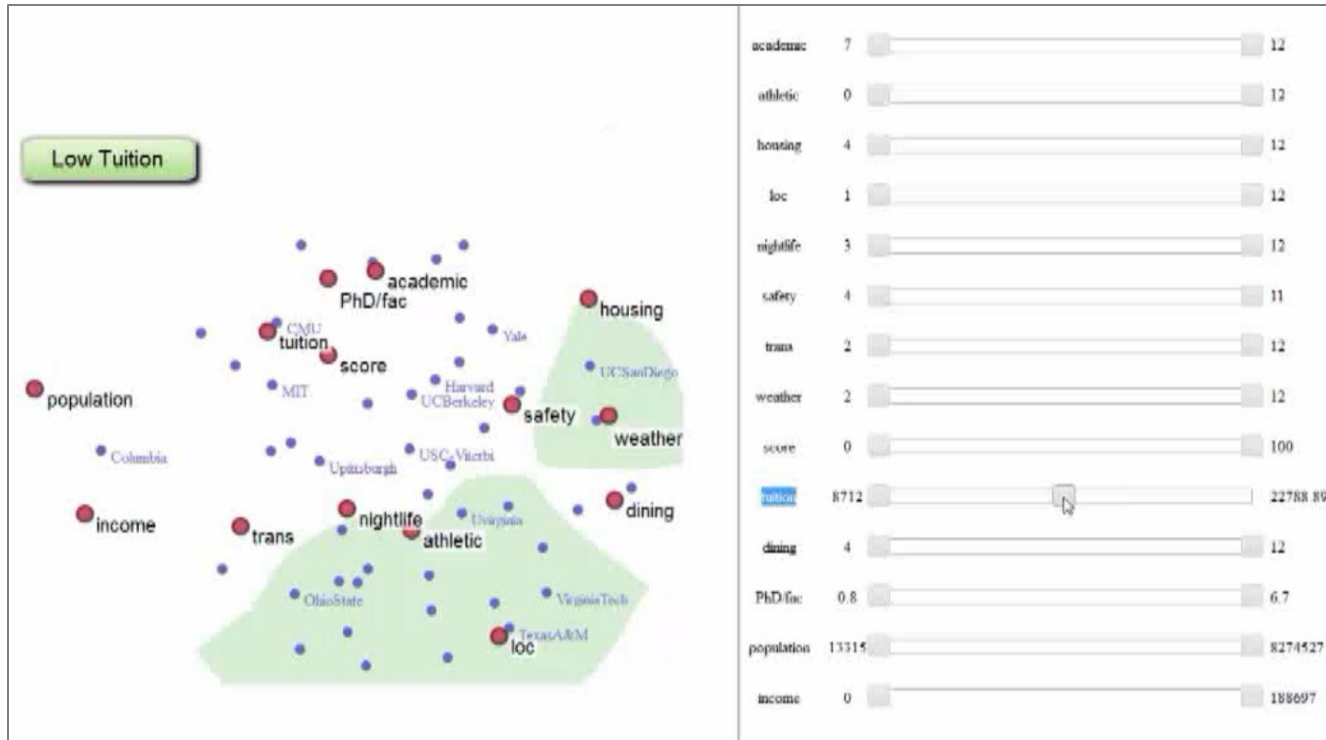


similar
colleges /
attributes
map closely

colleges
with higher
attribute
values map
close to that
attribute
node

Decision Region for Tuition < \$27,000

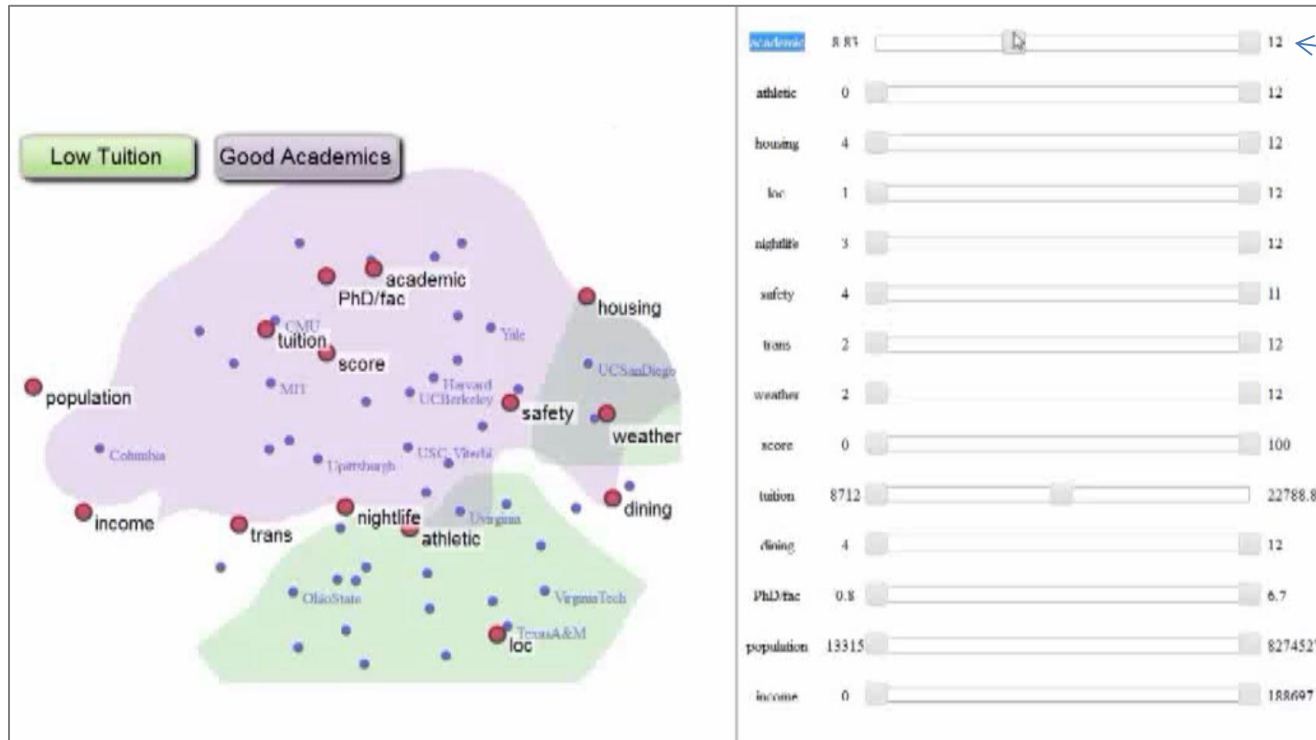
The green regions contain all colleges that fit that condition



tuition
slider
controls
range and
shape of
the green
tuition
region
(fully
accurate)

Add Decision Region for Academic Score > 8.3

The purple region contains all colleges that fit that condition



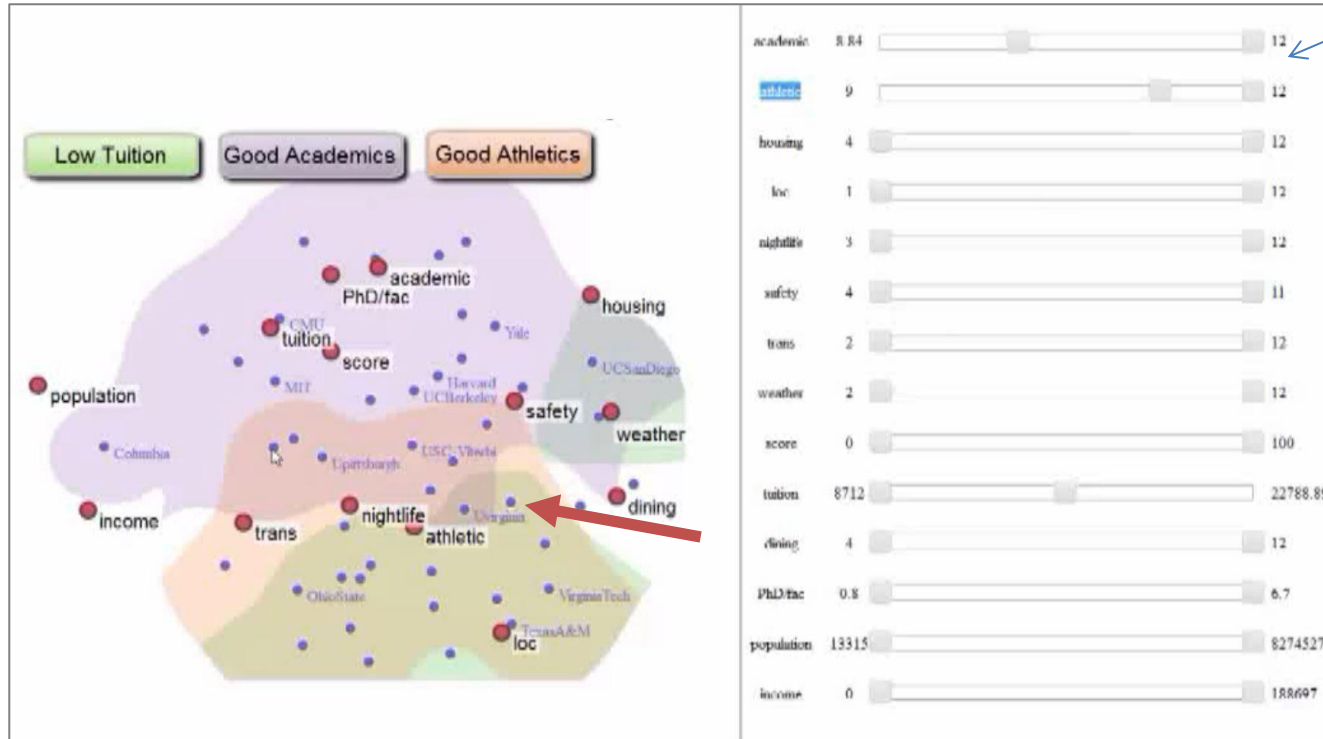
academics slider controls the shape of the academics region

some colleges fall into both decision regions, tuition and academics : these are the perfect colleges

some colleges are close by: these are the trade-off colleges

Add Decision Region for Athletics Score > 9

The purple region contains all colleges that fit that condition



athletics slider
controls the shape of
the athletics region

only one college falls
into all three decision
regions: this is the
perfect college for that
user (U Virginia)

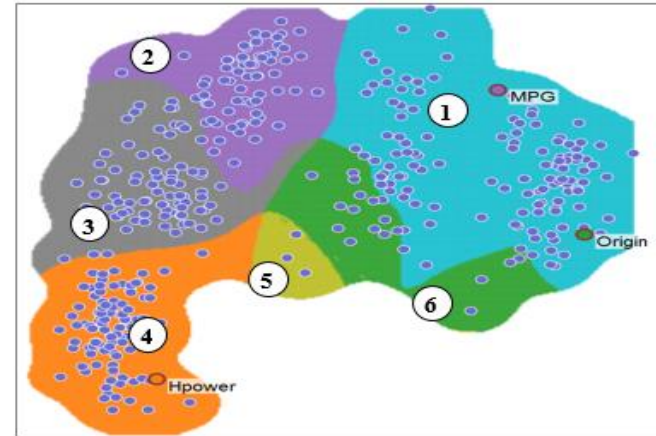
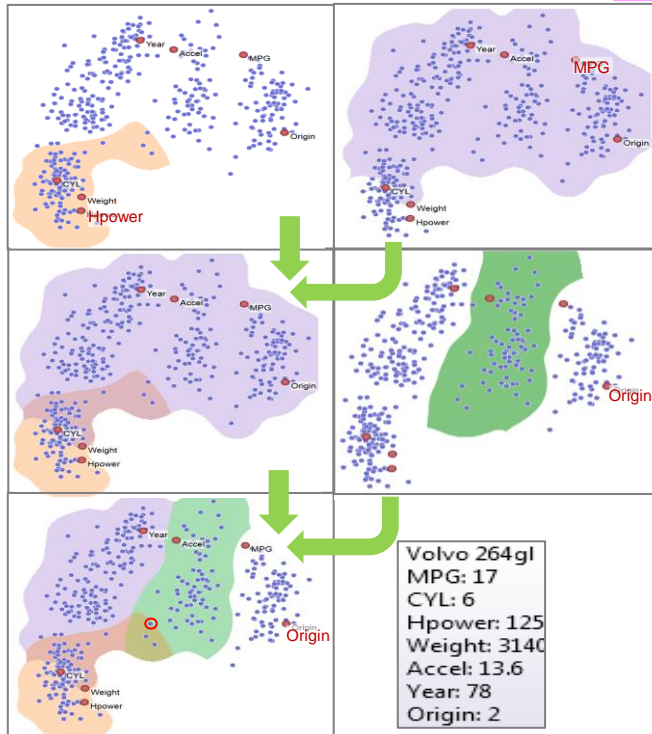
some colleges are
close by: these are the
trade-off colleges

Use Case: Overviewing Types of Cars

Horsepower: 116~192

MPG: 16~30

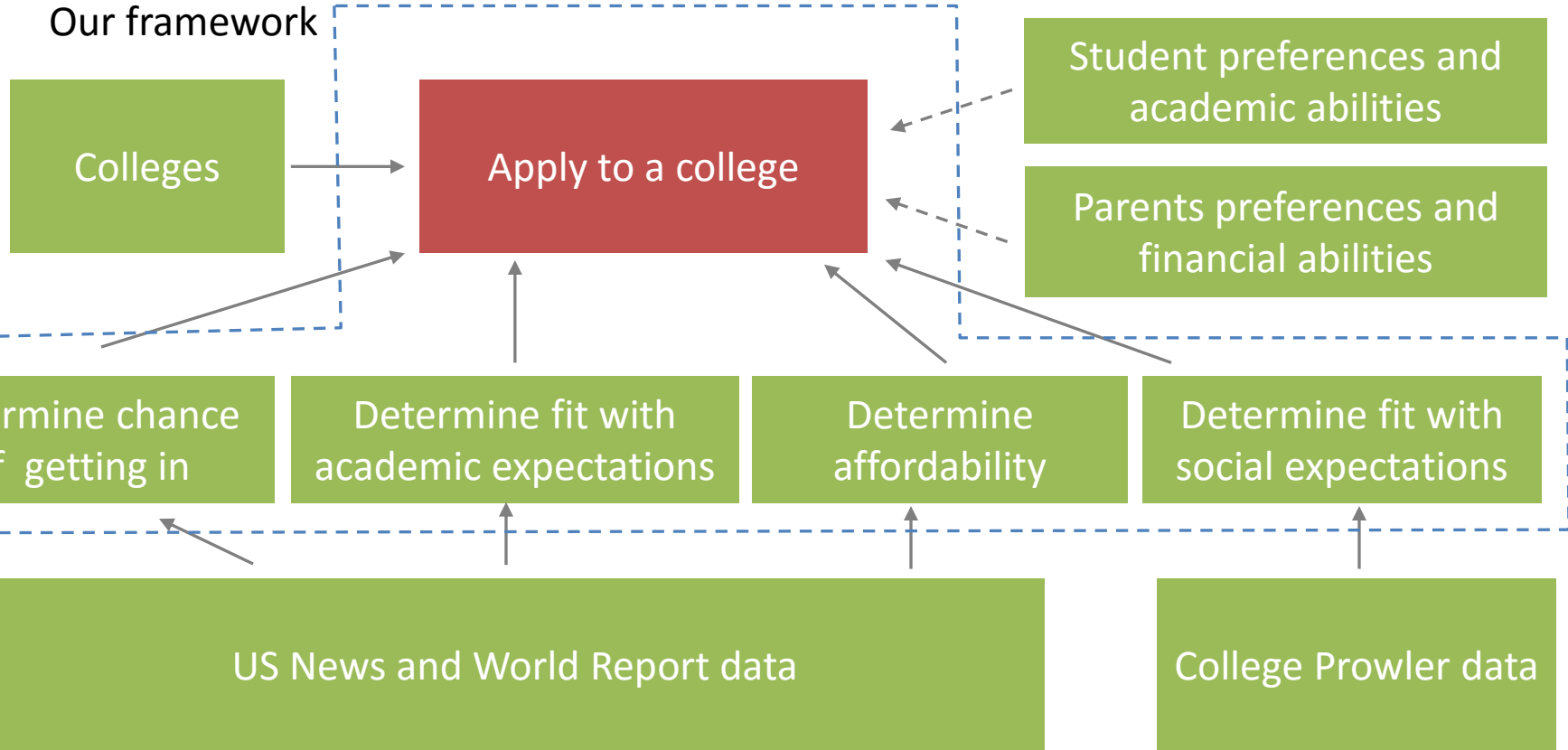
European Car



- 1 Euro-Japanese efficient compact cars
- 2 US efficient compact cars
- 3 US semi-efficient medium-power cars
- 4 US big block gas guzzlers
- 5 Euro-Japanese gas guzzlers
- 6 Euro-Japanese semi-efficient medium-power cars

DMN for College Selection Decision

Our framework



Conclusions and Future Work

User studies suggest that the framework is an intuitive environment to make decisions in selection tasks with many attributes (factors)

Some current work focuses on reducing the number of sliders

- give users options on the attributes they wish to control
- determine a set of good attributes automatically

