

# Slicing multi-dimensional spaces

Thomas Torsney-Weir

# About me

---

- Born: Allentown, PA



# About me

---

- Born: Allentown, PA
- Undergrad: Georgetown



# About me

---



- Born: Allentown, PA
- Undergrad: Georgetown
- Finance: NYC



# About me

---



- Born: Allentown, PA
- Undergrad: Georgetown
- Finance: NYC
- Master's: Simon Fraser



# About me

---



- Born: Allentown, PA
- Undergrad: Georgetown
- Finance: NYC
- Master's: Simon Fraser
- PhD/Postdoc: University of Vienna



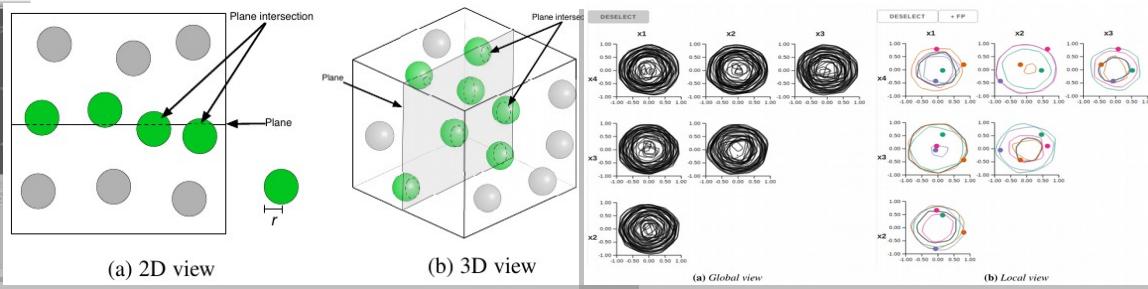
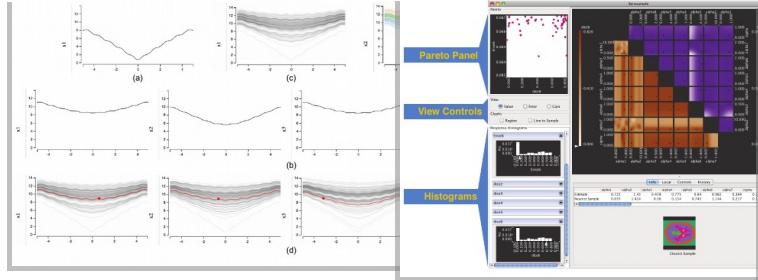
# About me

---

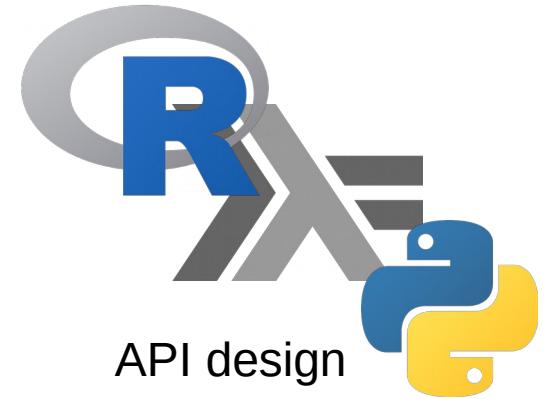


- Born: Allentown, PA
- Undergrad: Georgetown
- Finance: NYC
- Master's: Simon Fraser
- PhD/Postdoc: University of Vienna
- Next: Swansea University!

# My research

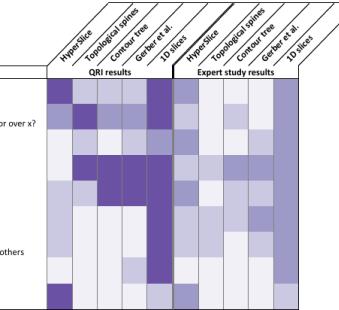


Multi-dimensional spaces

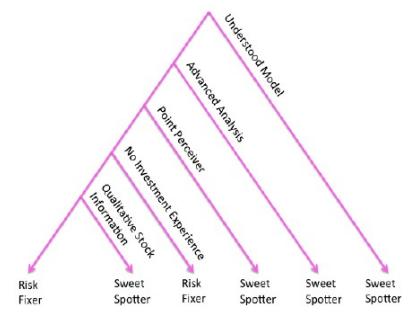


API design

Task	Task description for discrete data items from [AES05]	Our adaption to continuous scalar functions
Retrieve value	"Given a set of specific cases, find attributes of those cases"	Given an $x$ , what is the function value?
Filter	"Given some concrete conditions attribute values, find data cases satisfying those conditions."	For what parameter values is the function equal or over $x$ ?
Compute derived value	"Given a set of data cases, compute an aggregate numeric representation of the data cases"	Summary statistics: variance, mean, SA
Find extremum	"Find data cases possessing an extreme value of an attribute over its range within the data set"	Find local/global min/max
Determine range	"Given a set of data cases and an attribute of interest, find the span of values within the set"	What is the range of possible outputs?
Characterize distribution	"Given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute's values over the set"	What types of shapes do the manifolds have
Find anomalies	"Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers"	Do areas of the manifold have shapes unlike any others
Cluster	"Given a set of data cases, find clusters of similar attribute values"	Areas of the manifold have similar shapes
Correlate	"Given a set of data cases and two attributes, determine useful relationships between the values of those attributes"	1D vs 2D relationships

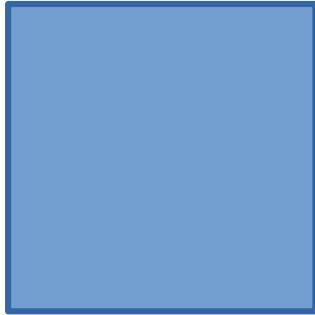


Users/tasks

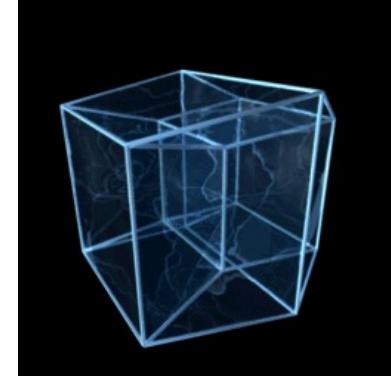


# Viewing these things is hard

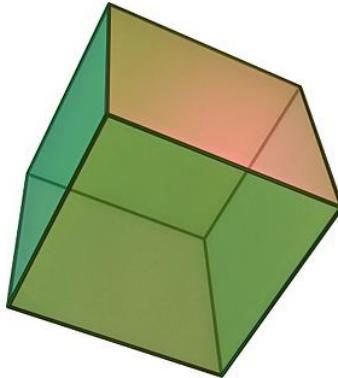
2D



4D



3D



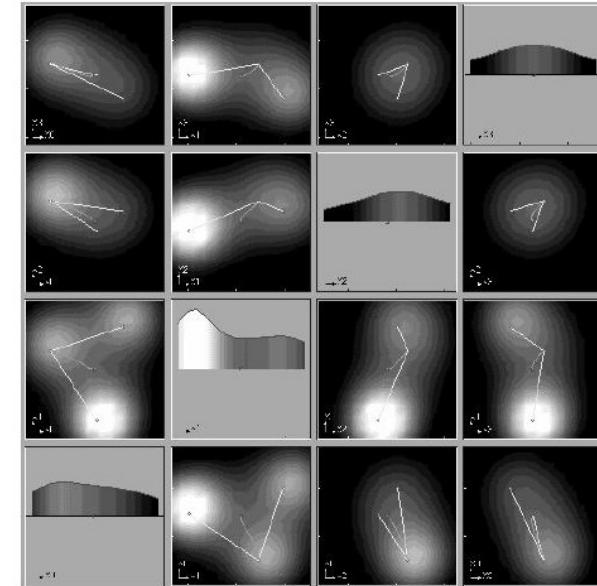
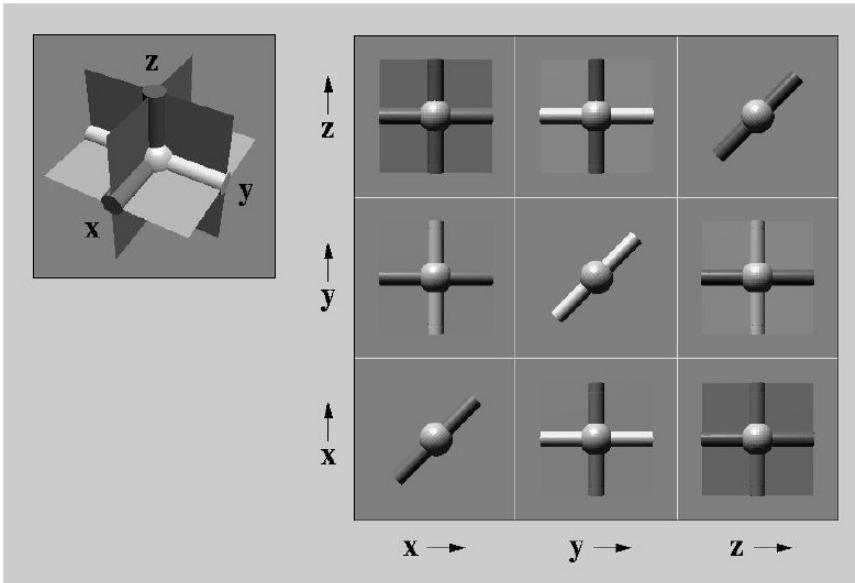
>4D



# Slicing

---

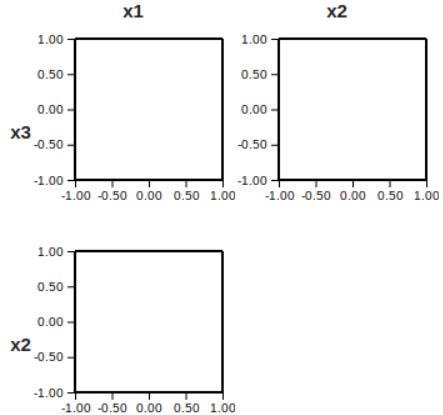
- General view
- Easy to understand metaphor



# Cube

---

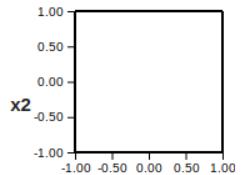
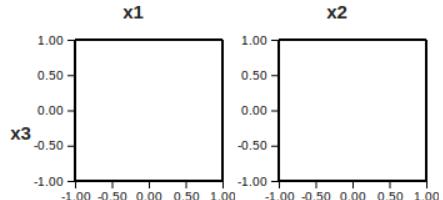
3D



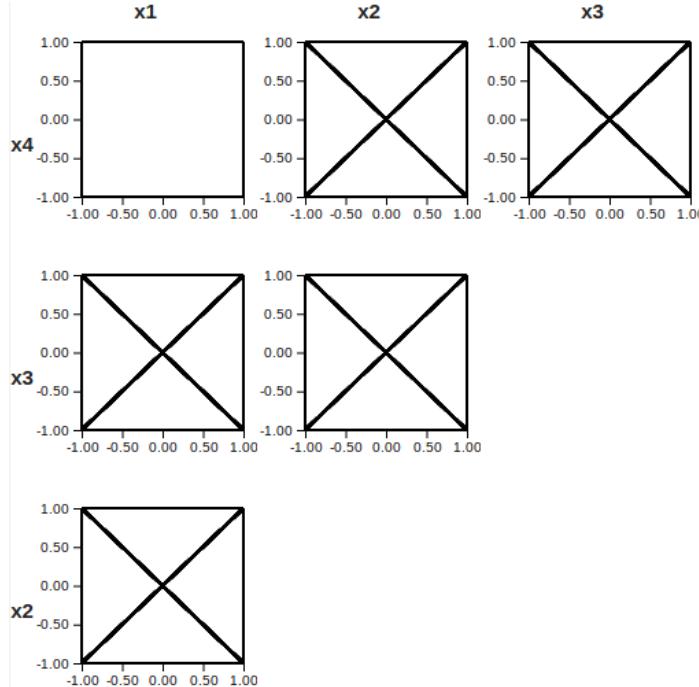
# Cube

---

3D



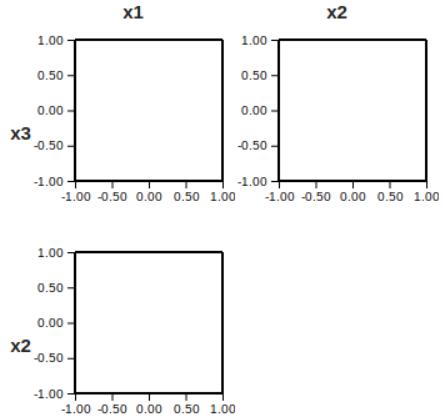
4D



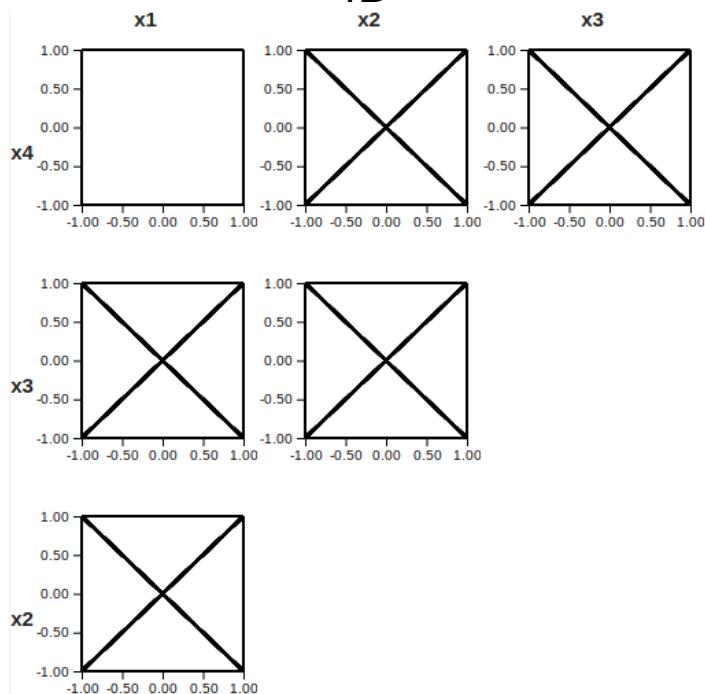
# Cube

---

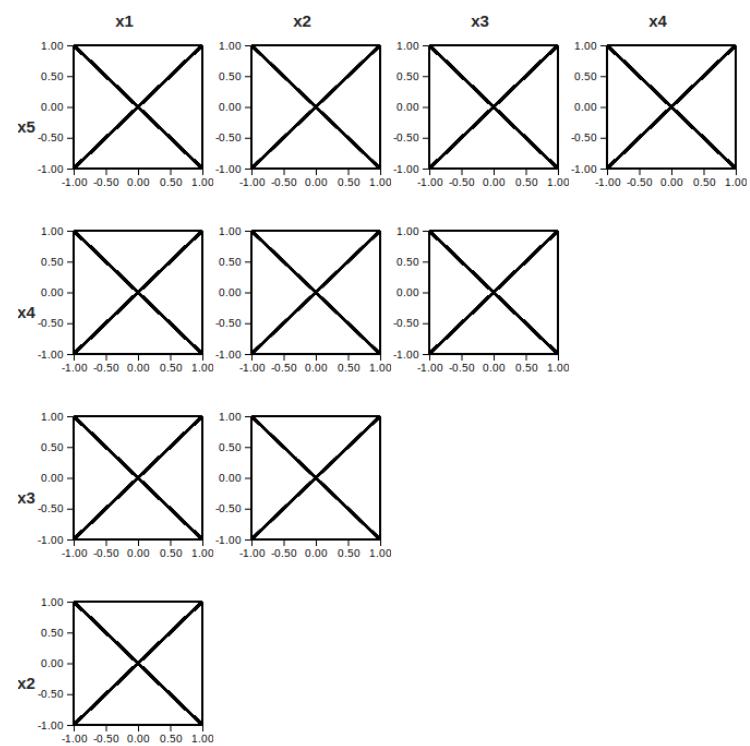
3D



4D



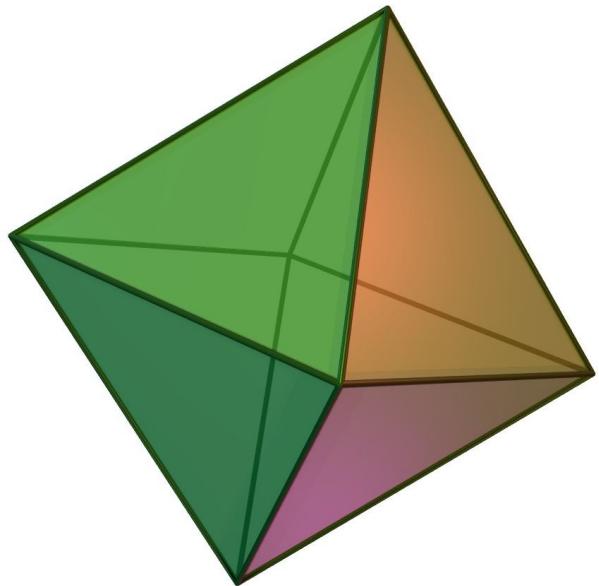
5D



# Orthoplex

---

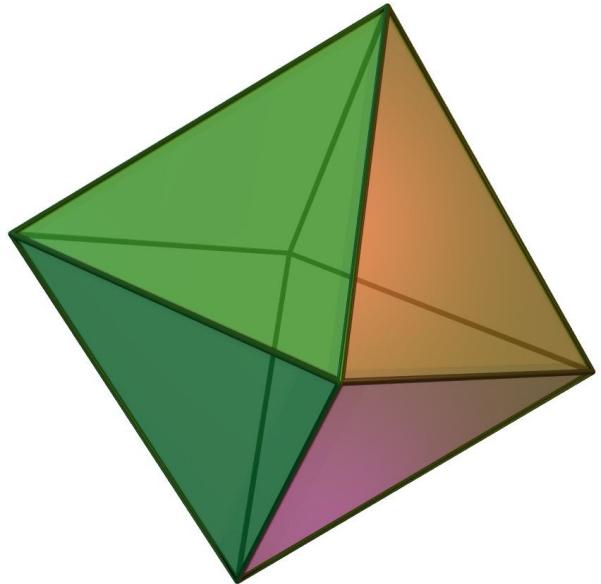
3D



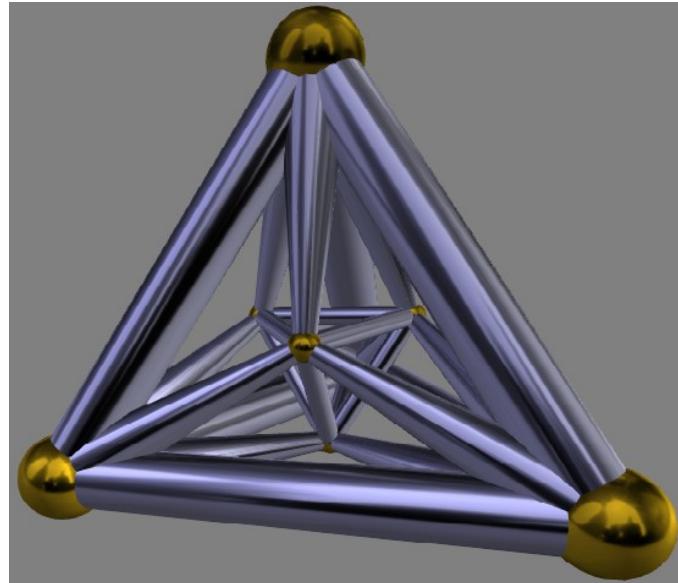
# Orthoplex

---

3D



4D

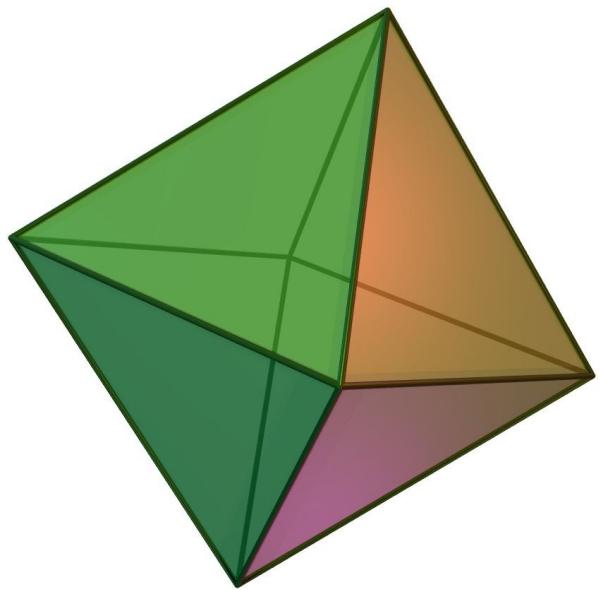


Sommerville, 1929

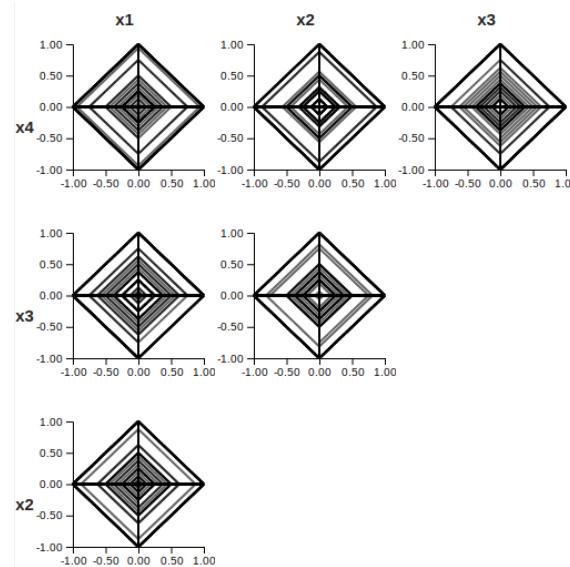
# Orthoplex

---

3D



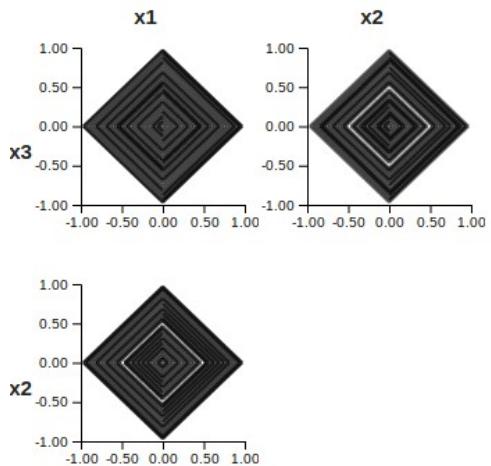
4D



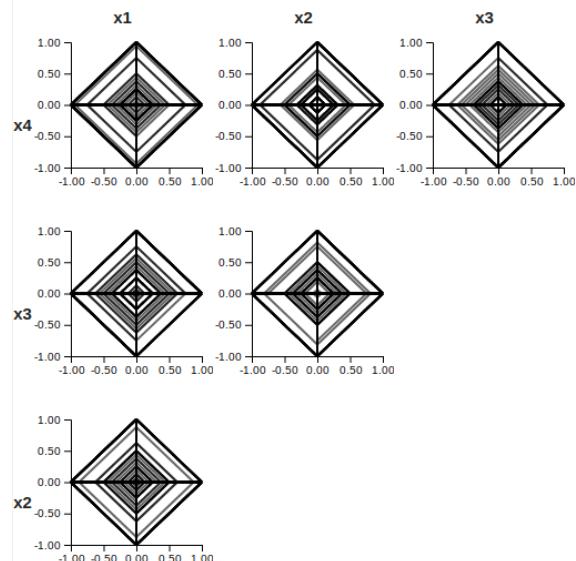
# Orthoplex

---

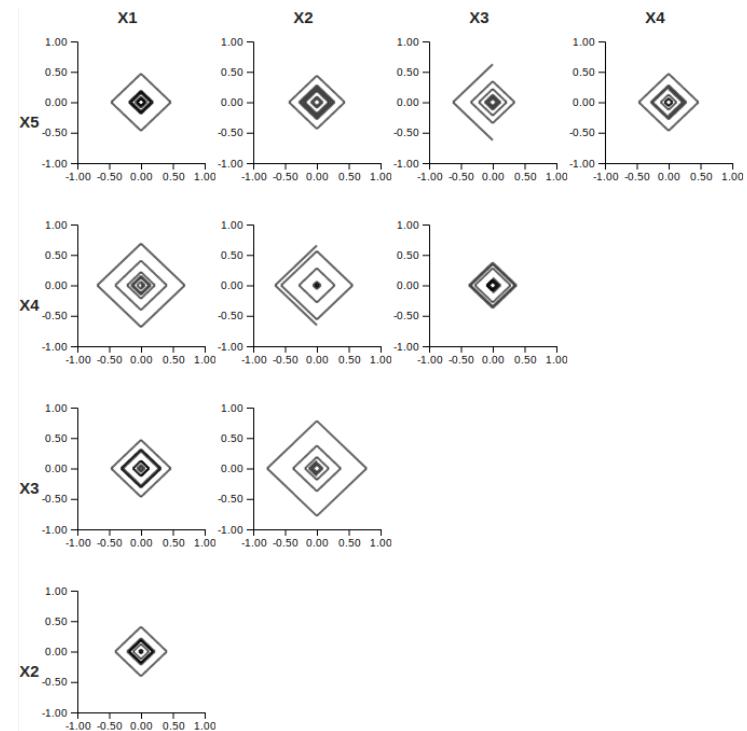
3D



4D

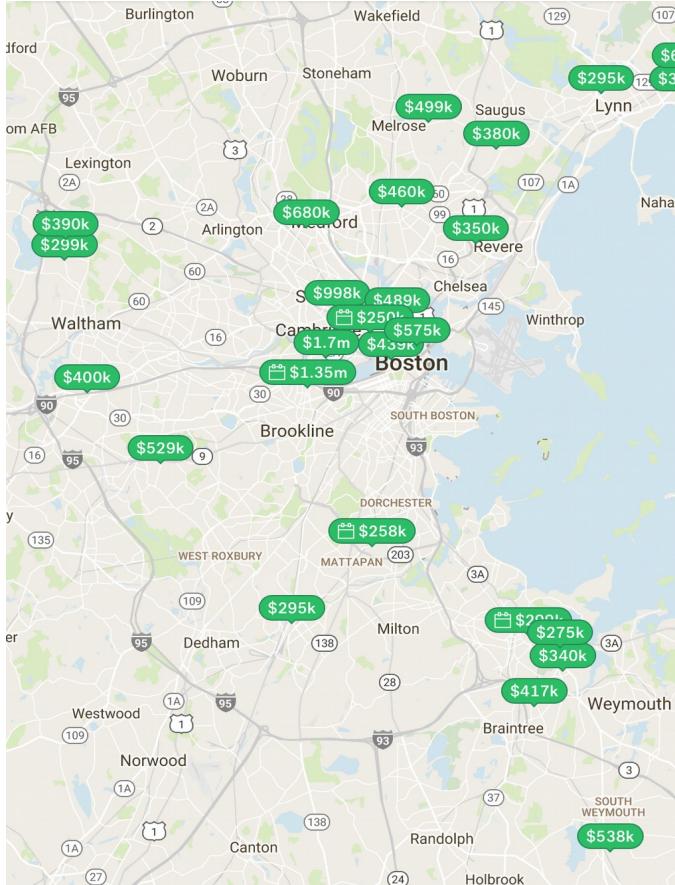


5D



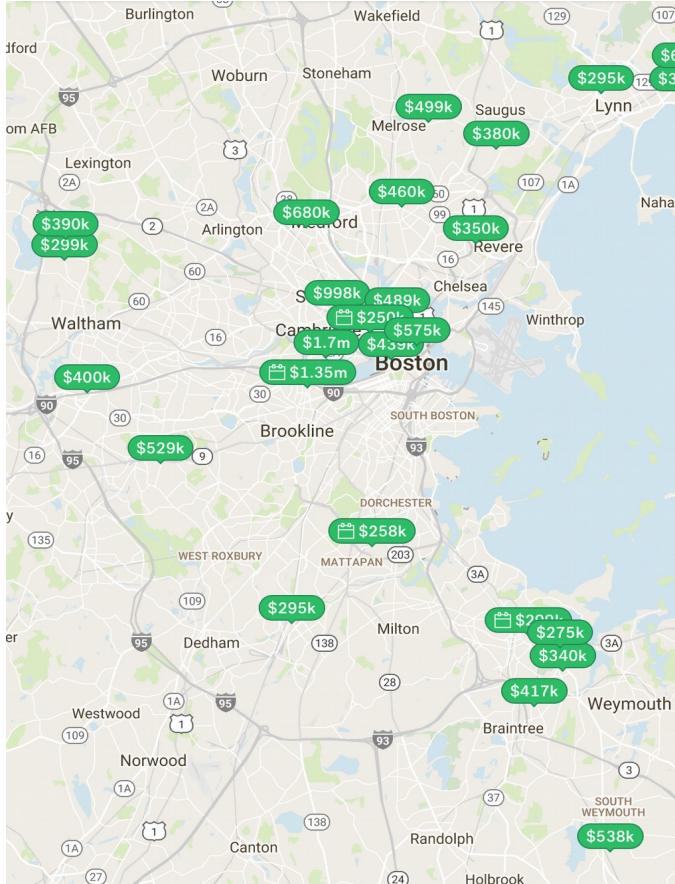
# Housing prices in Boston

---



# Housing prices in Boston

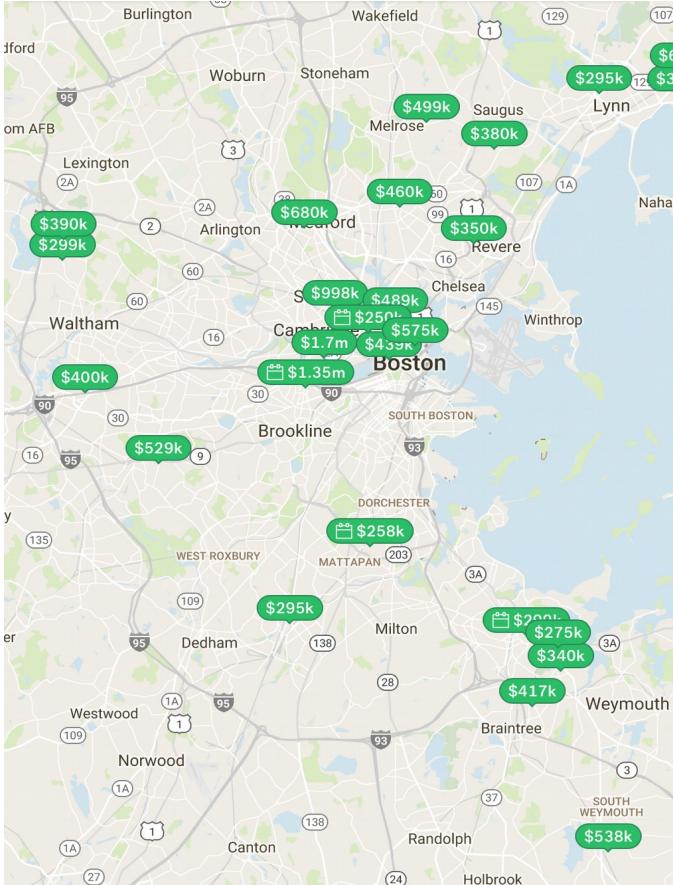
---



- 1) What's the most expensive house?
- 2) What factors contribute the most to changes in price?
- 3) How much does house price change with safety?
- 4) Is the relationship linear/logarithmic/etc?

# Housing prices in Boston

---



- 1) What's the most expensive house?
- 2) What factors contribute the most to changes in price?
- 3) How much does house price change with safety?
- 4) Is the relationship linear/logarithmic/etc?

Continuous model

# UCI housing dataset

---

13 factors → Median home price

## Key factors

CRIM      Crime rate

LSTAT      % lower income status

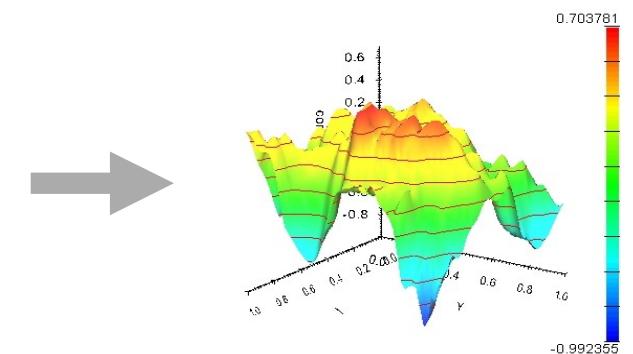
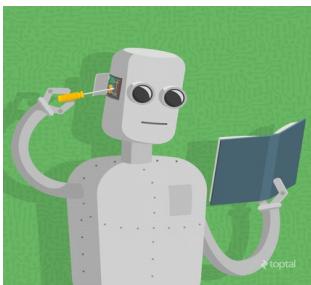
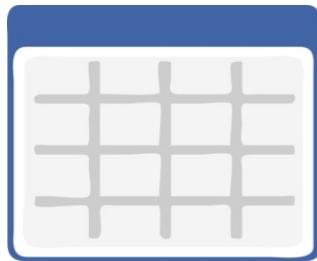
NOX      Nitric oxides concentration

RM      Average rooms per dwelling

# Building a model

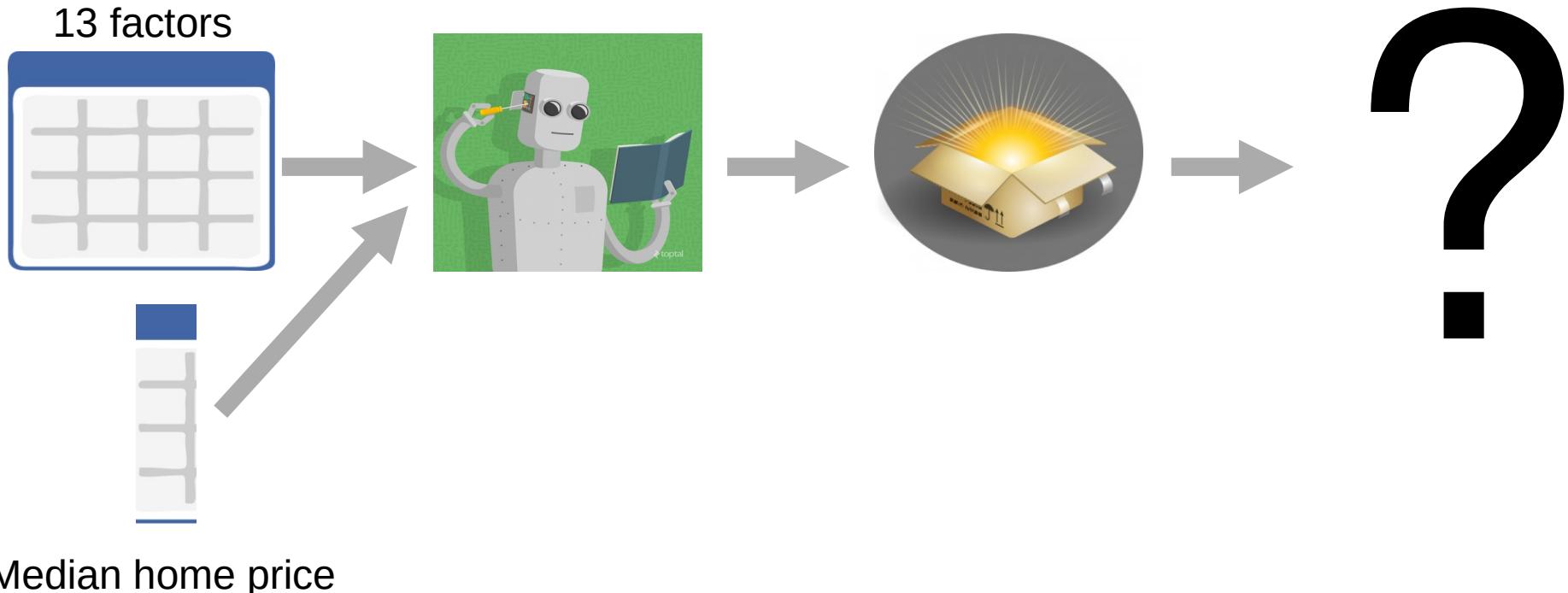
---

13 factors



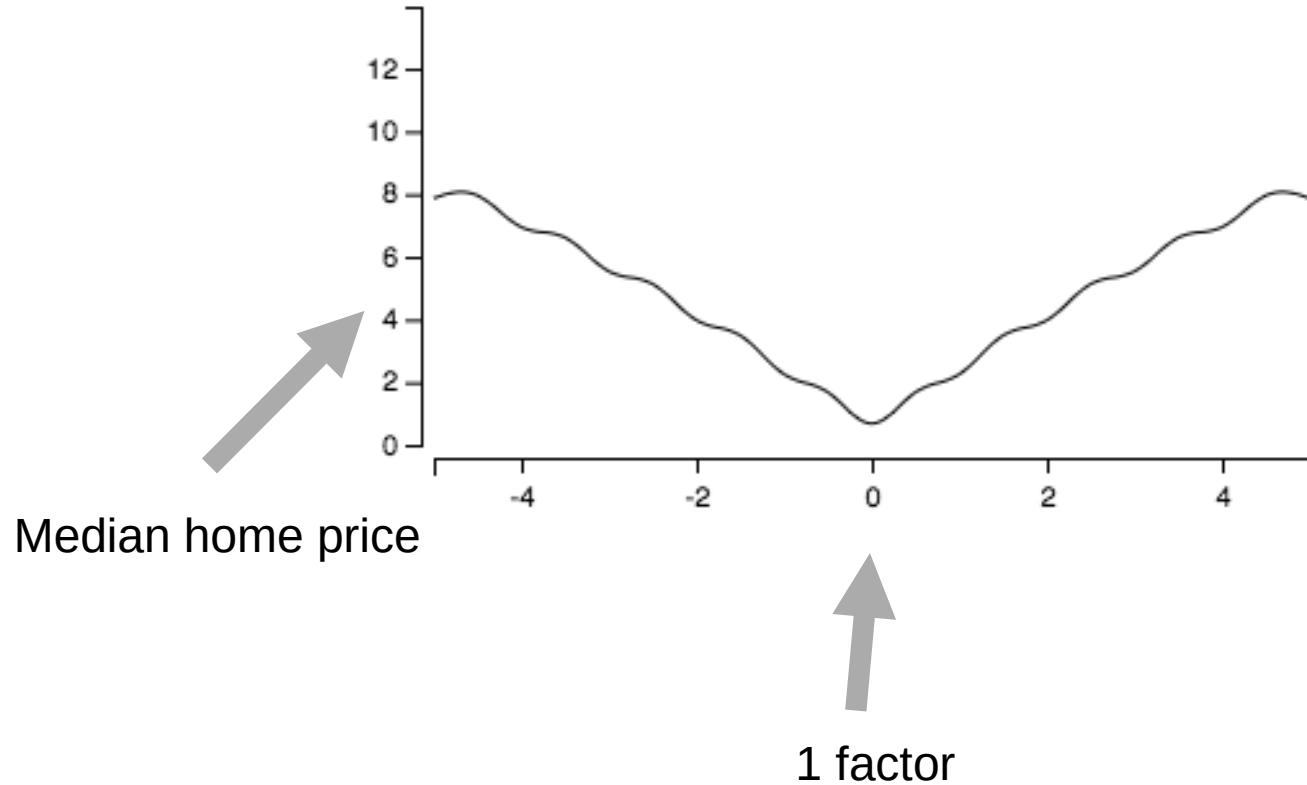
Median home price

# Building a model



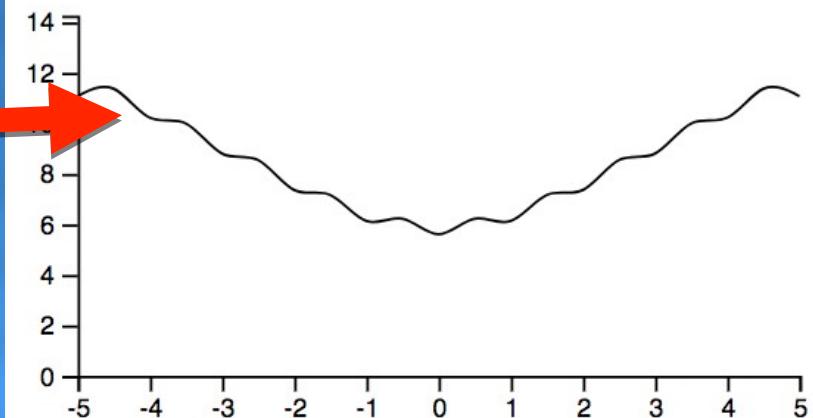
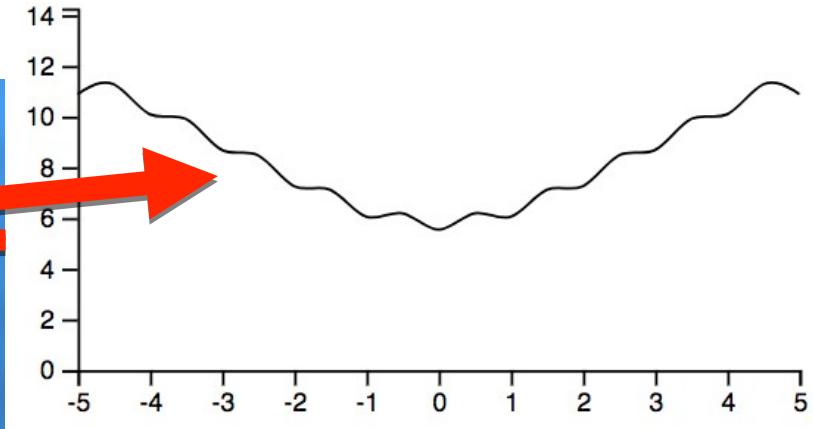
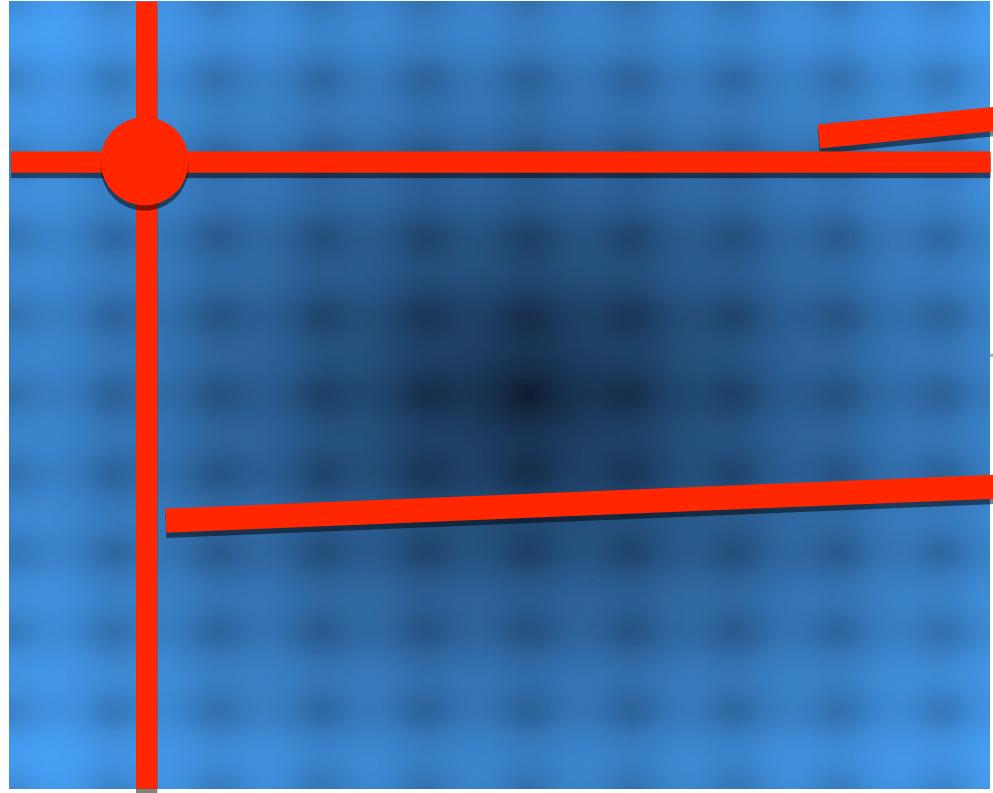
# Visualization

---



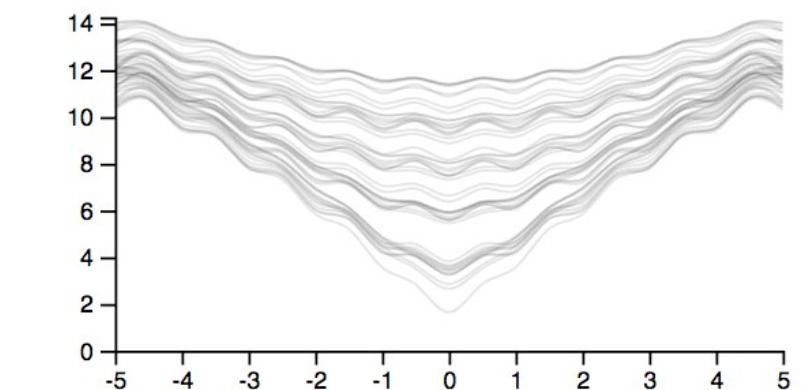
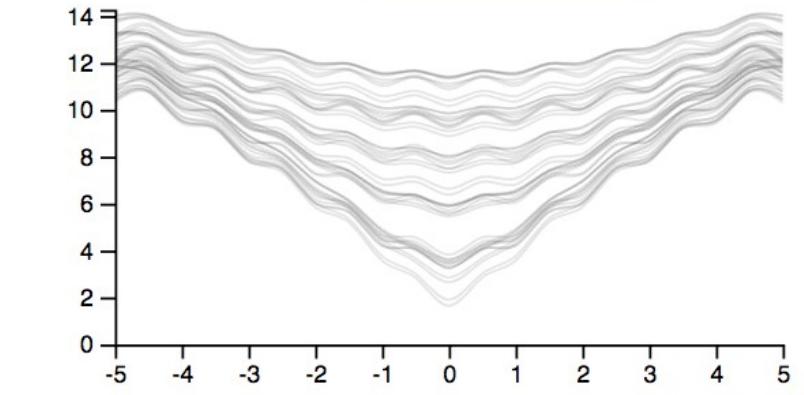
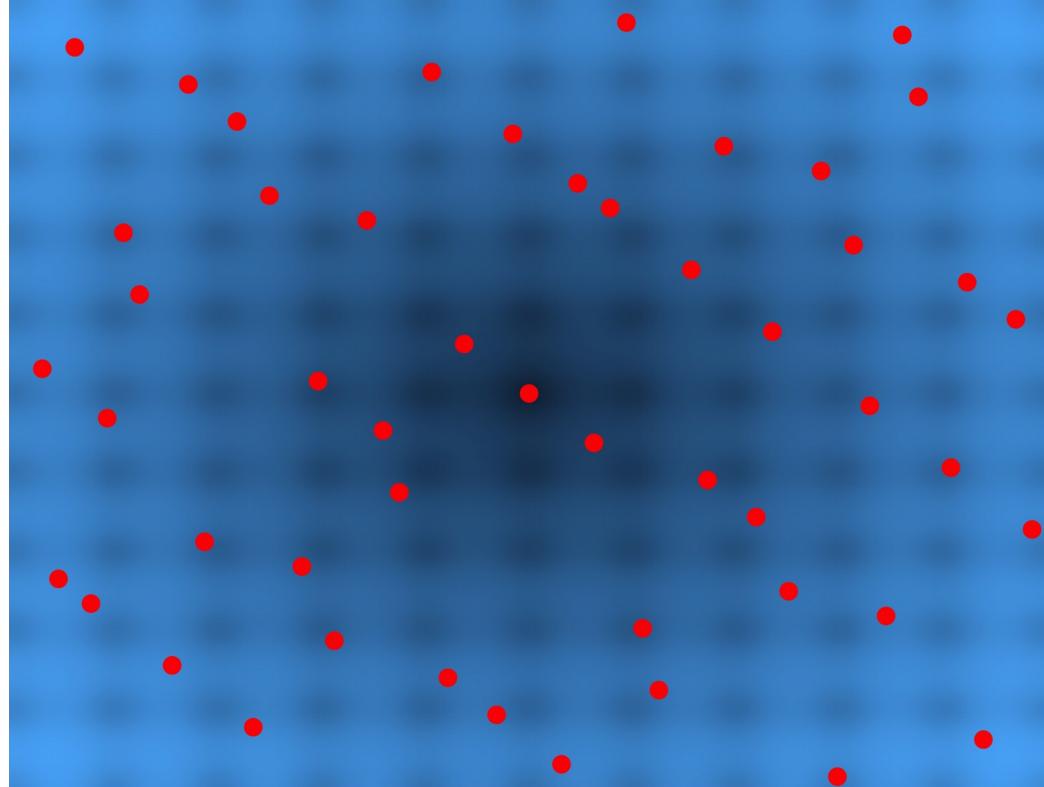
# Slicing

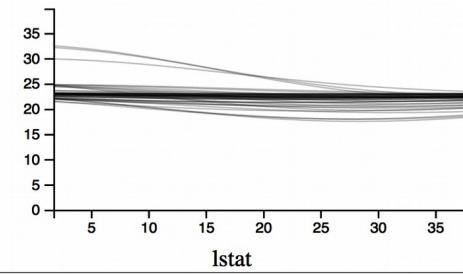
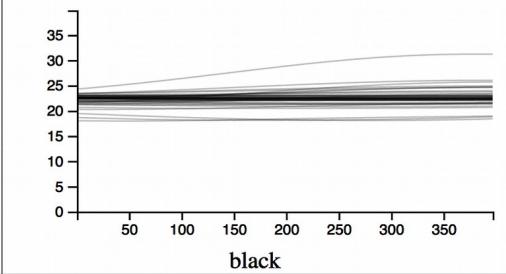
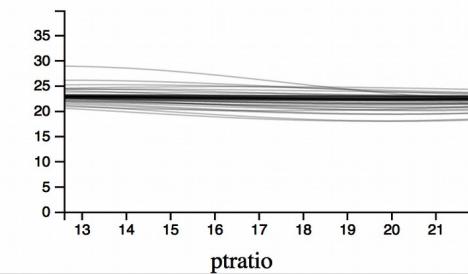
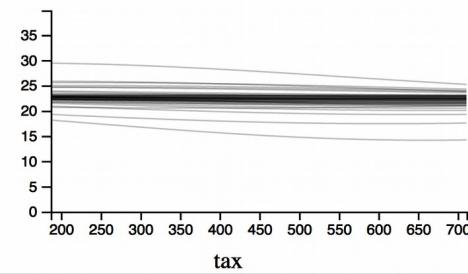
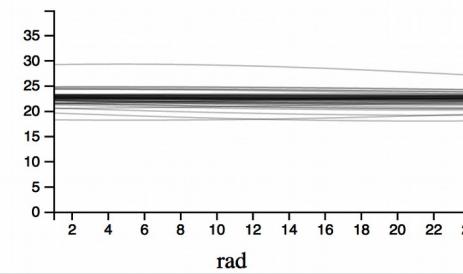
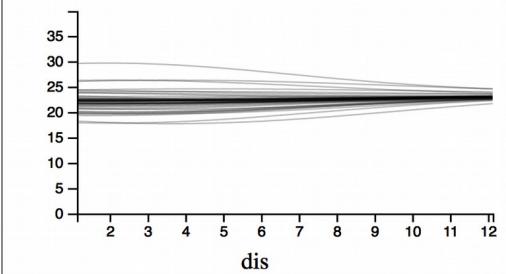
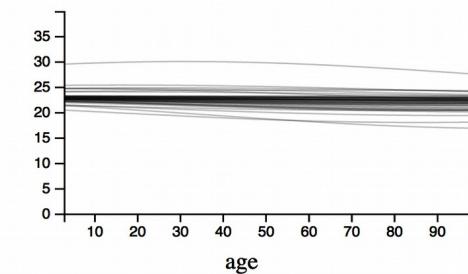
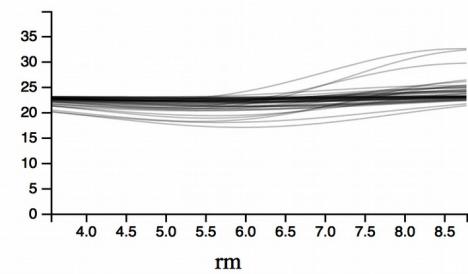
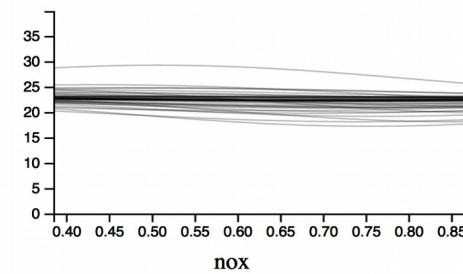
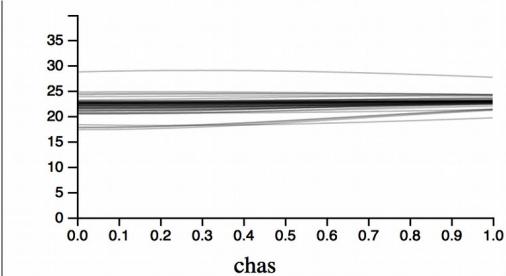
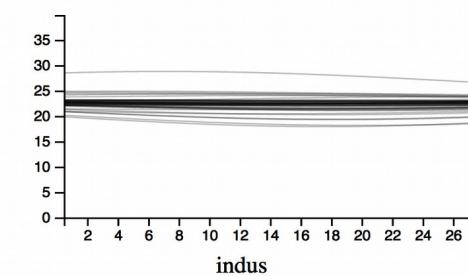
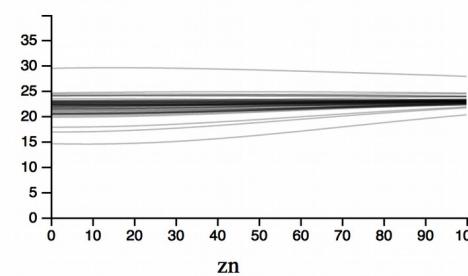
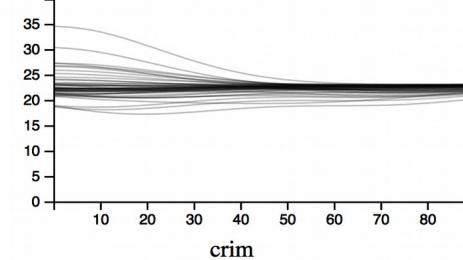
---

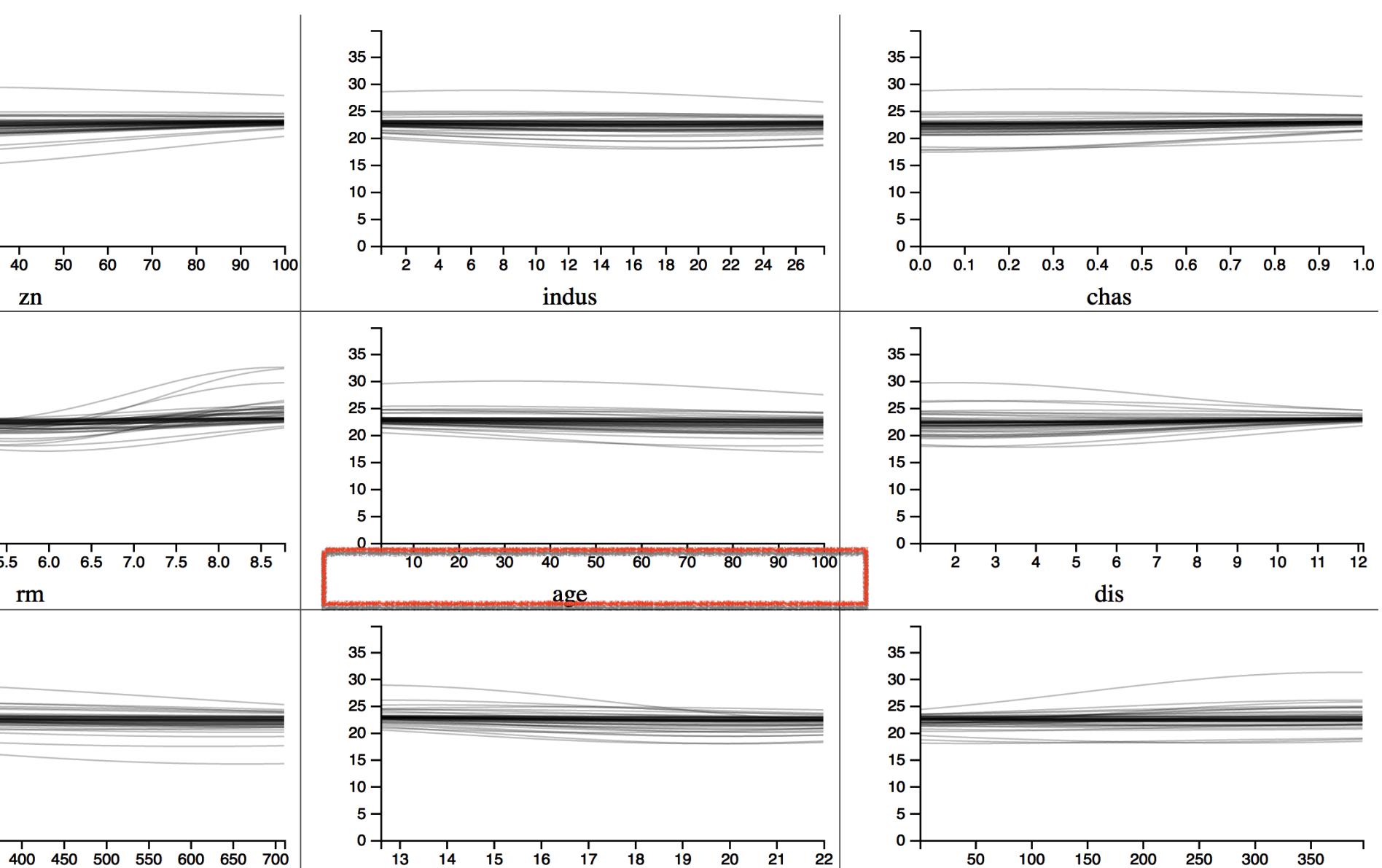


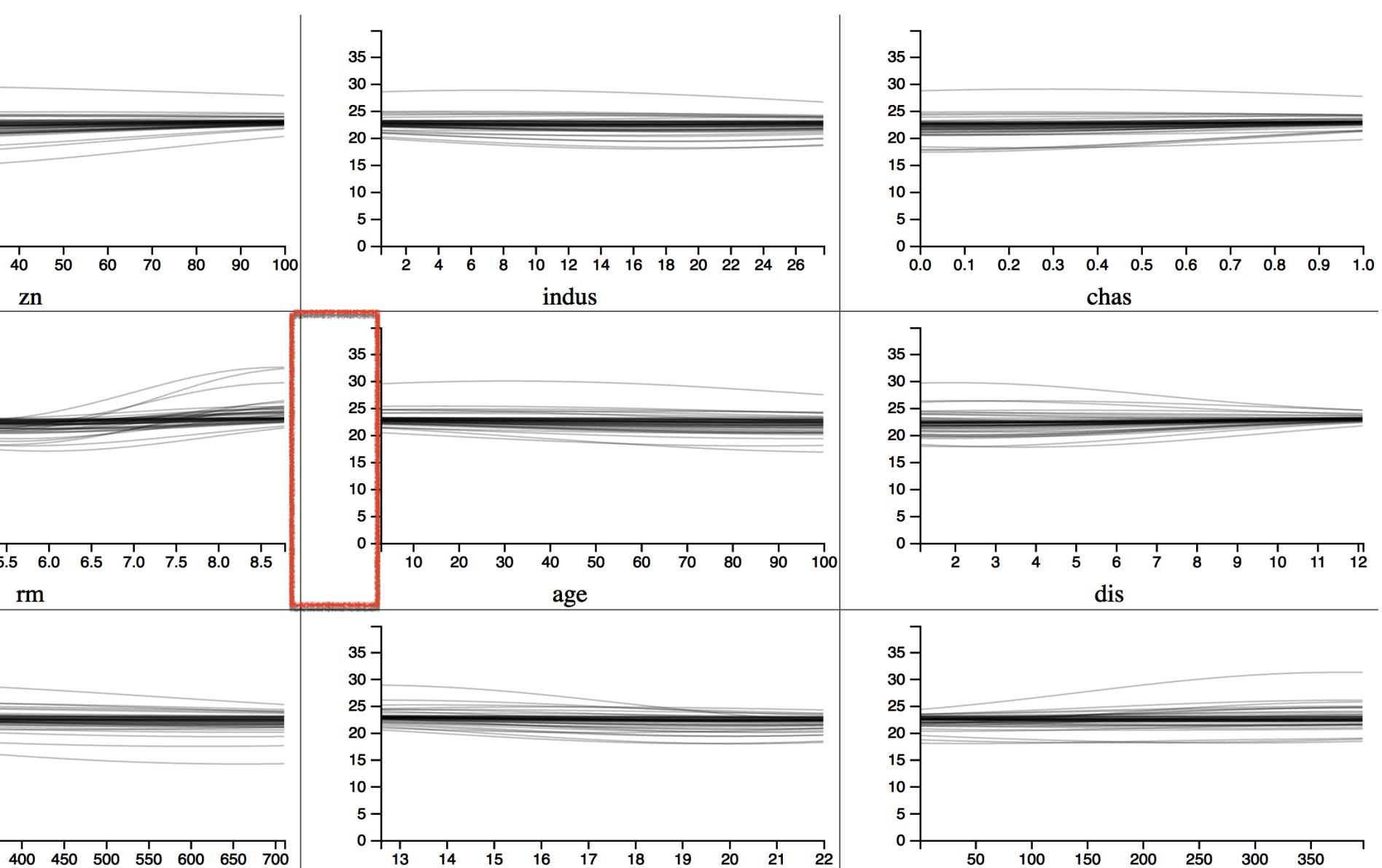
# Slicing

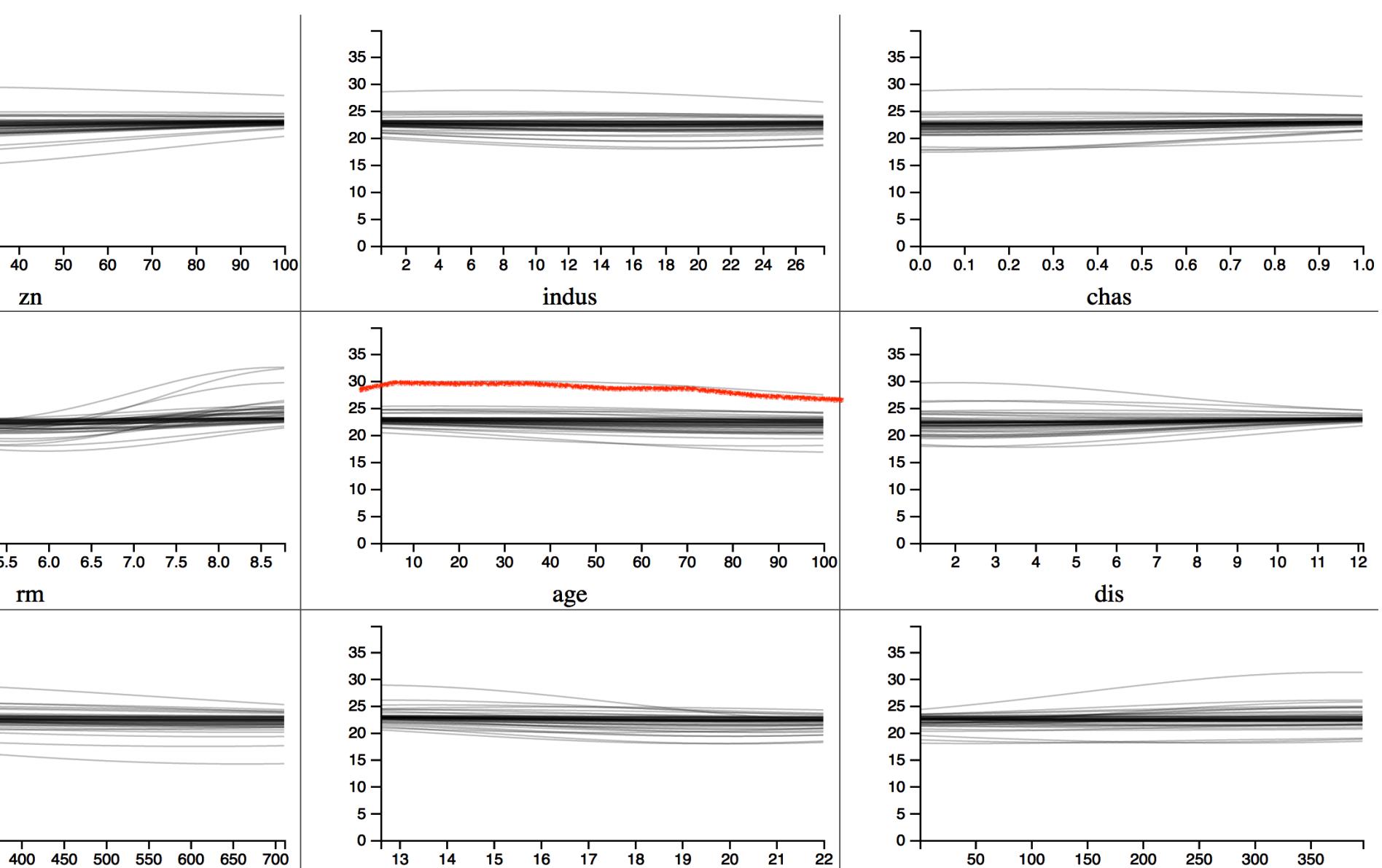
---



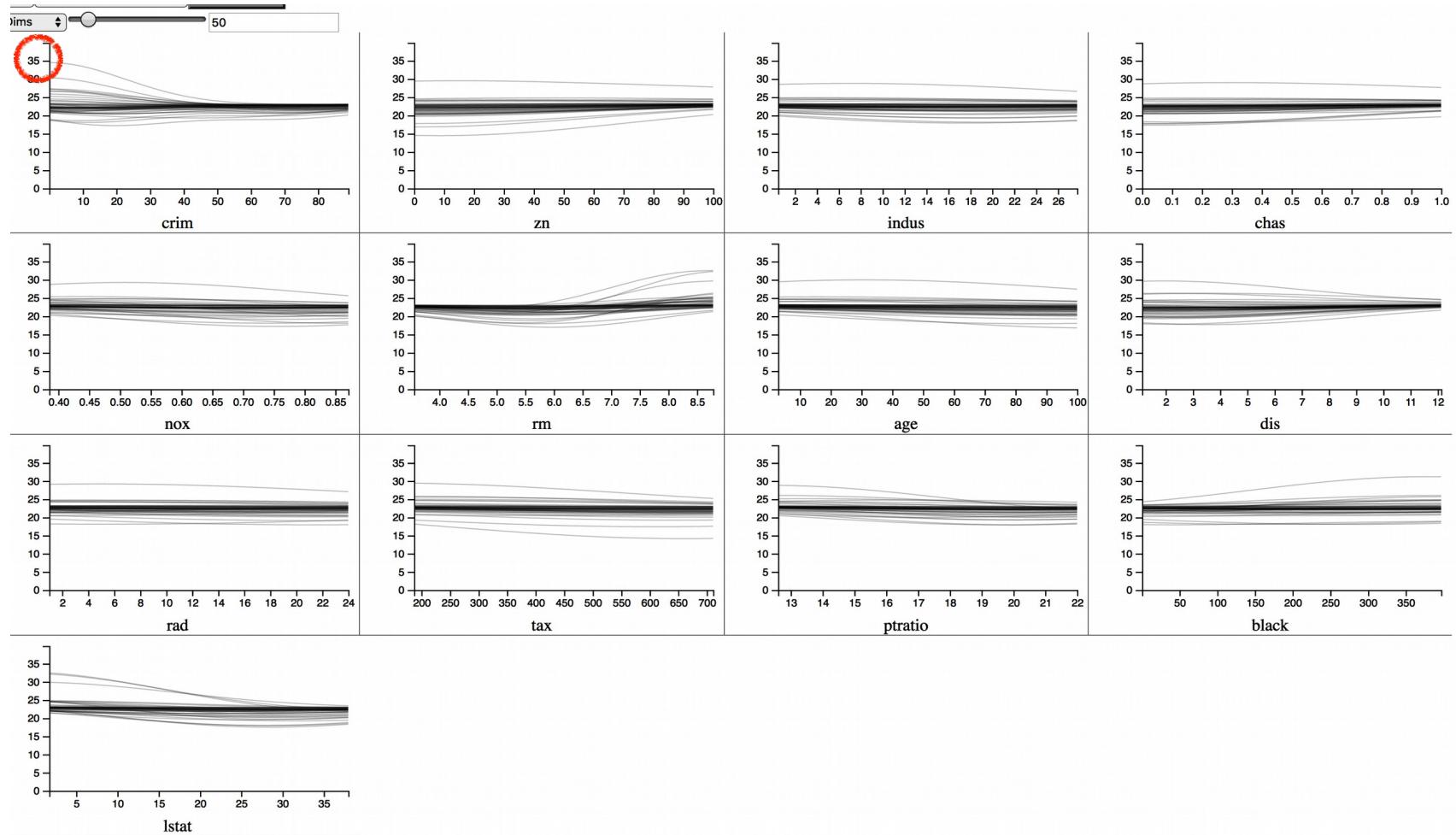






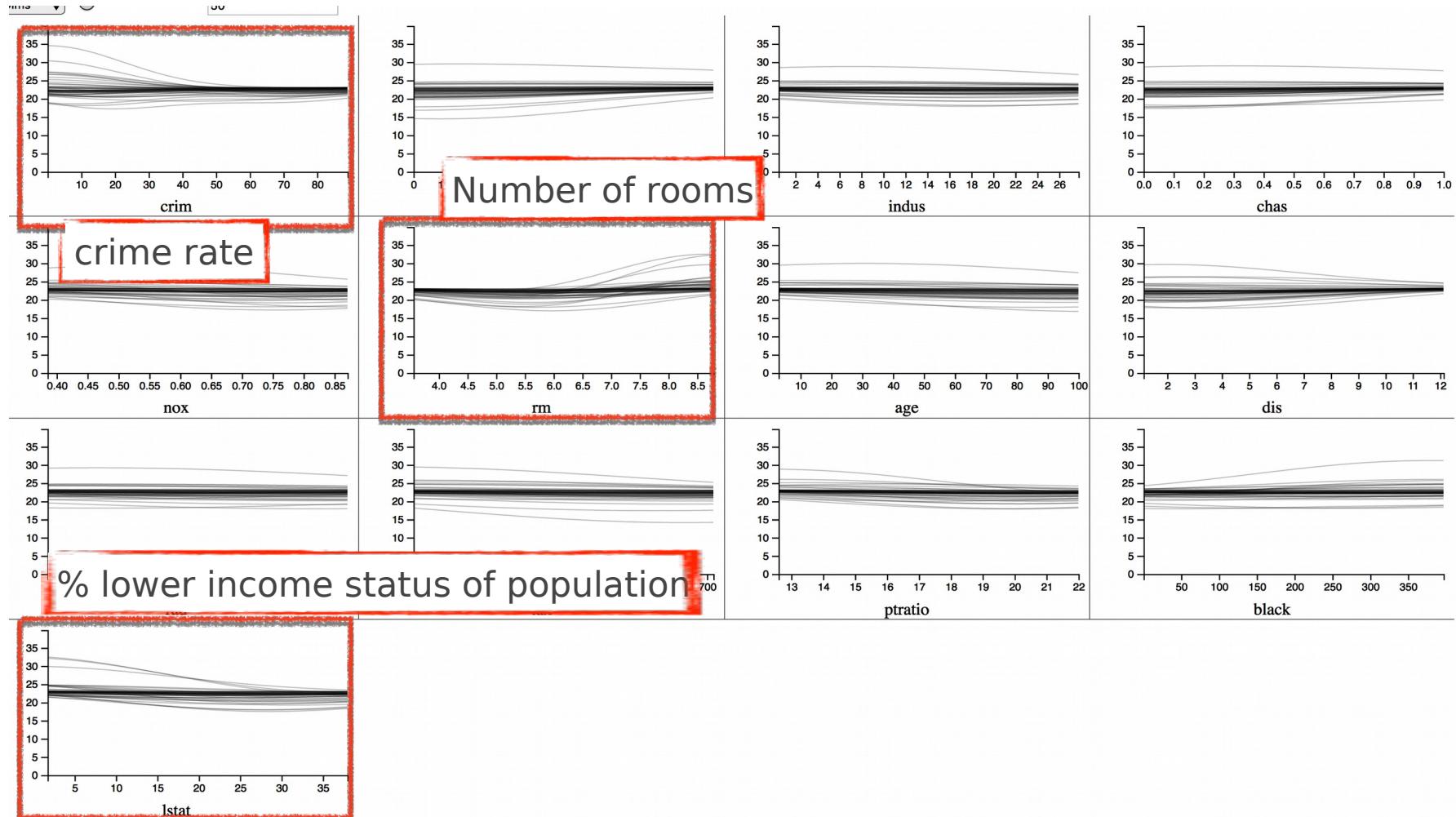


# Q1: Most expensive house

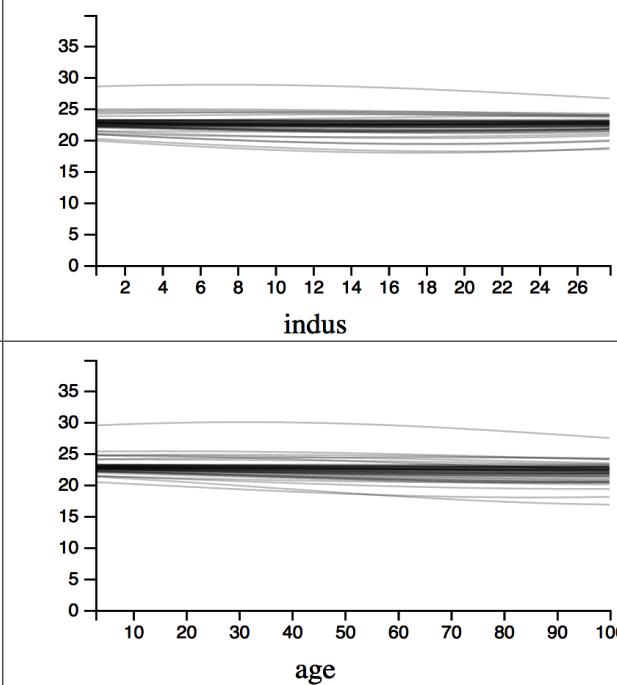
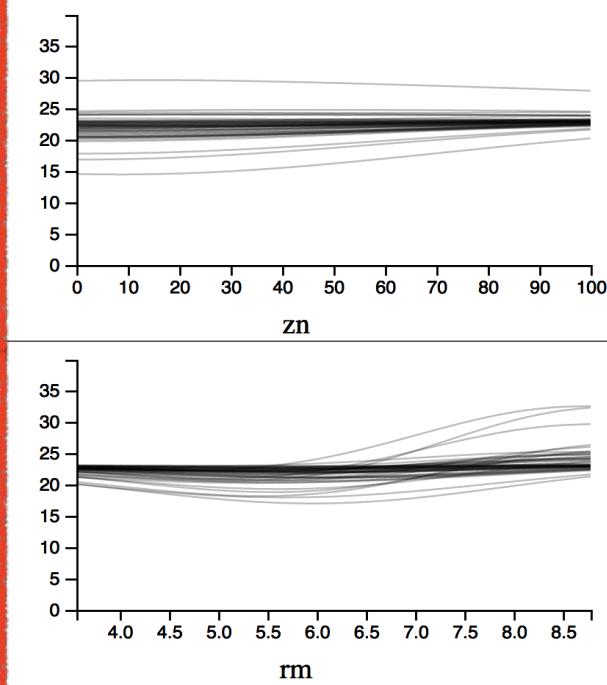
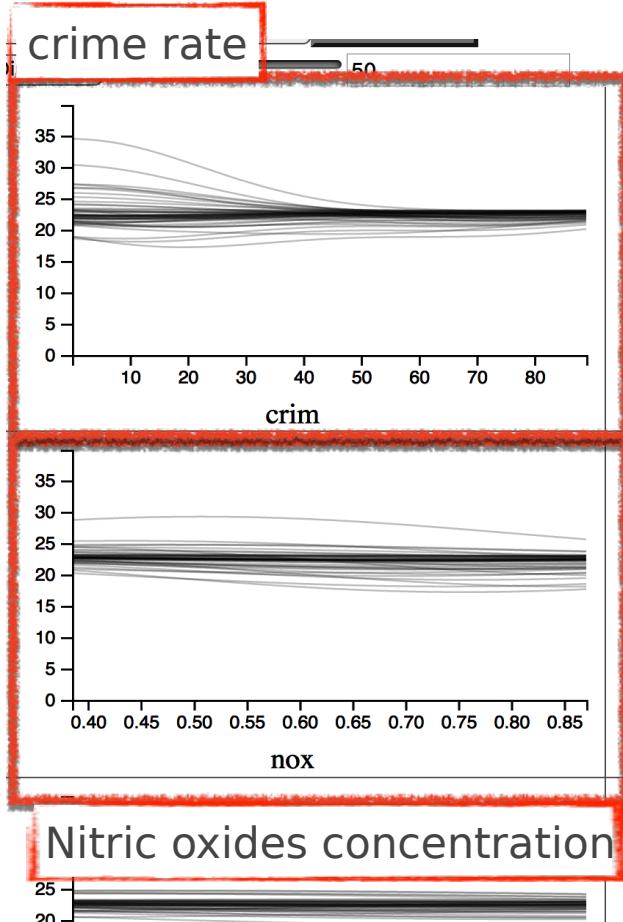


# Q2:

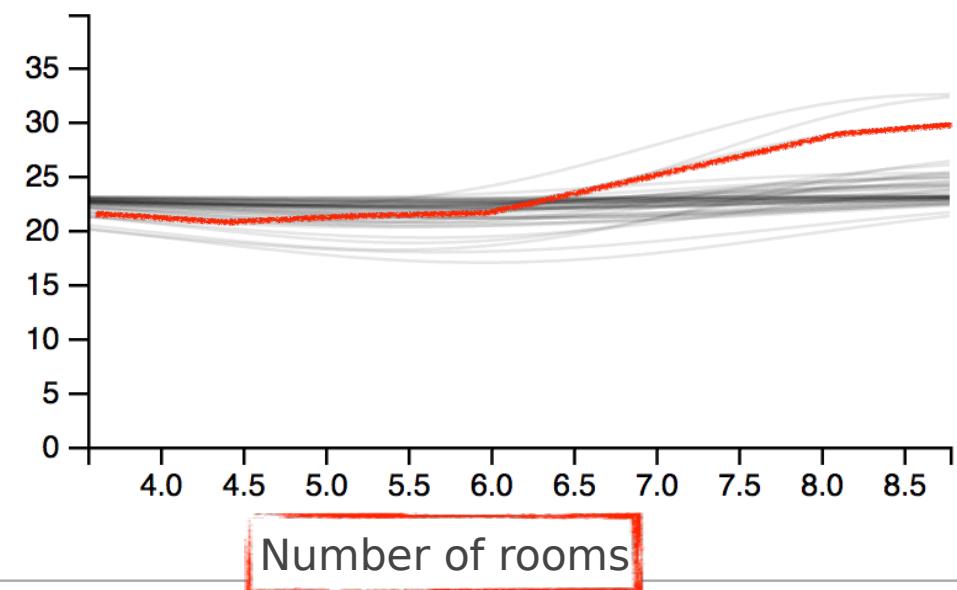
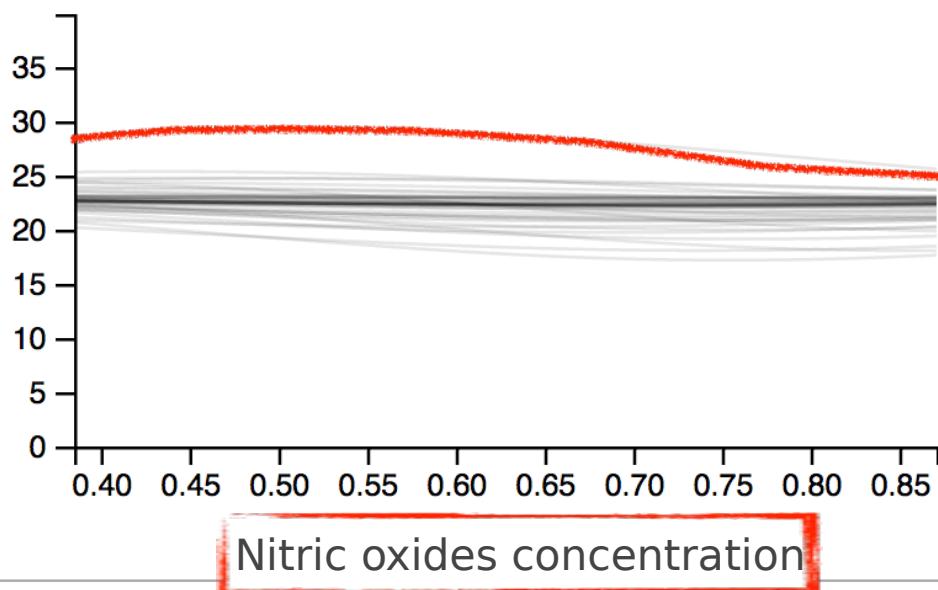
---



# Q3: Influence of safety?

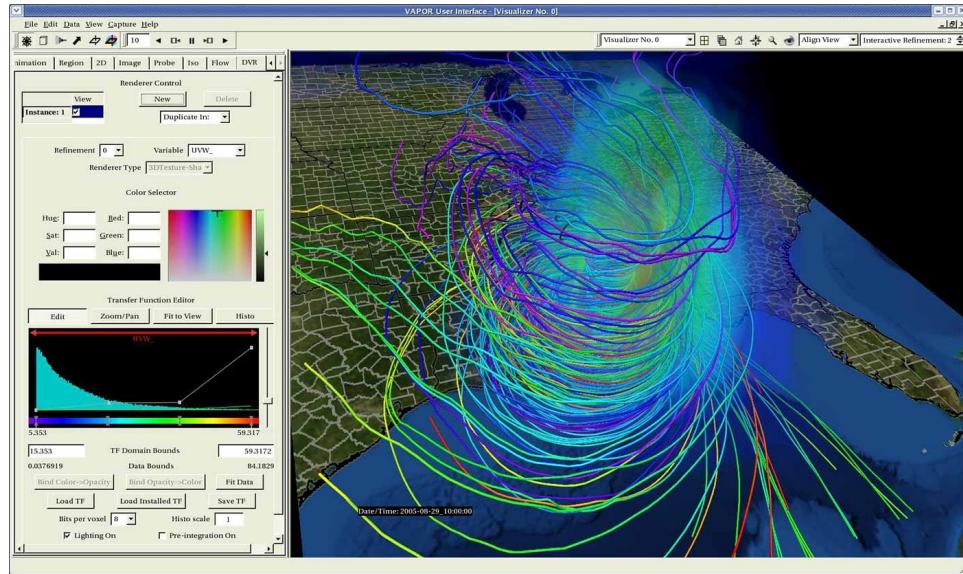


# Q4: Type of relationship?

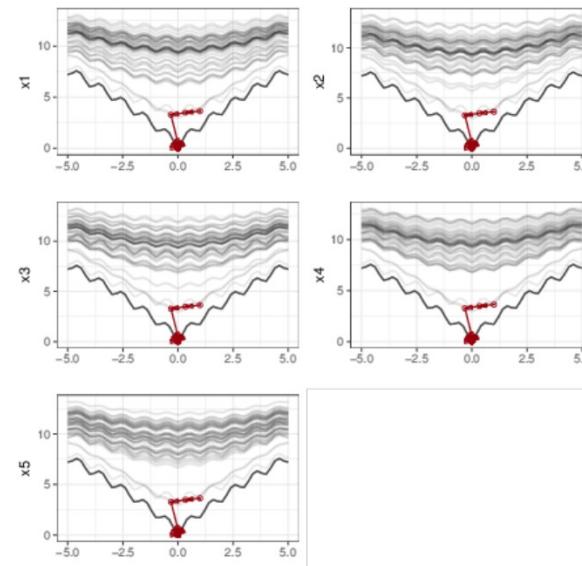


# Applications

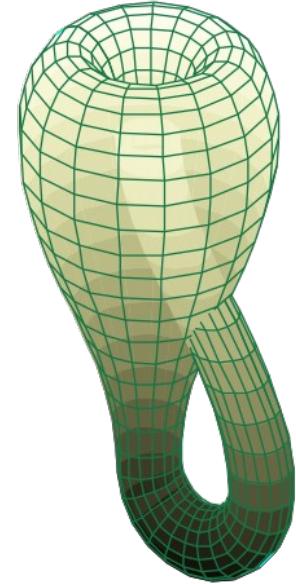
---



Simulations



Optimization



Polytopes

# Tasks

---

- Started with Amar, Eagan, and Stasko, 2005
- Meaning when studying input/output relationships
- “Sort” had no continuous analog
- 9/10 others had analogues

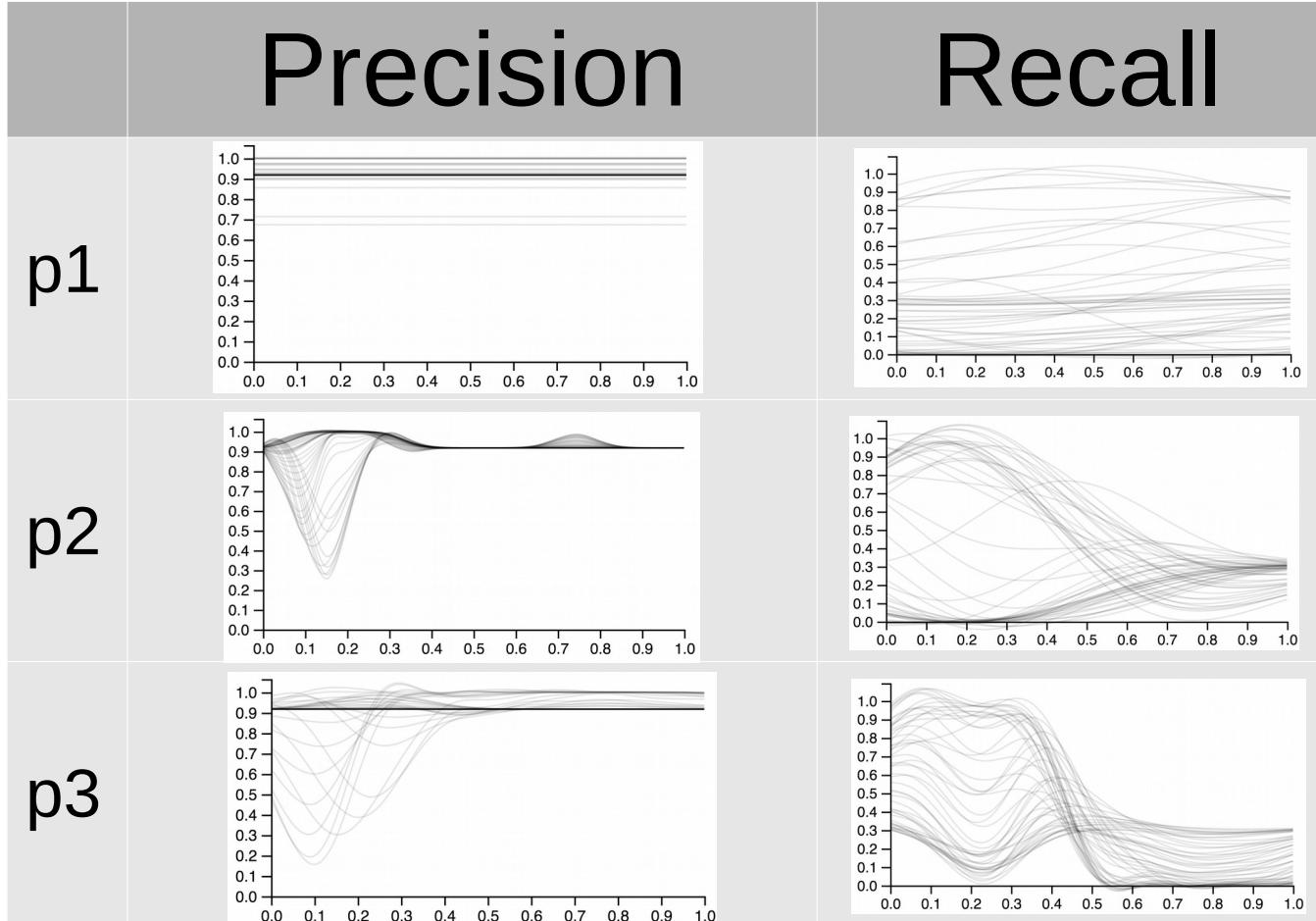
# Tasks

---

Task	Task description for discrete data items from [AES05]	Our adaption to continuous scalar functions
Retrieve value	"Given a set of specific cases, find attributes of those cases"	Given an x, what is the function value?
Filter	"Given some concrete conditions attribute values, find data cases satisfying those conditions."	For what parameter values is the function equal or over x?
Compute derived value	"Given a set of data cases, compute an aggregate numeric representation of those data cases"	Summary statistics: variance, mean, SA
Find extremum	"Find data cases possessing an extreme value of an attribute over its range within the data set"	Find local/global min/max
Determine range	"Given a set of data cases and an attribute of interest, find the span of values within the set"	What is the range of possible outputs?
Characterize distribution	"Given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute's values over the set"	What types of shapes do the manifolds have
Find anomalies	"Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers"	Do areas of the manifold have shapes unlike any others
Cluster	"Given a set of data cases, find clusters of similar attribute values"	Areas of the manifold have similar shapes
Correlate	"Given a set of data cases and two attributes, determine useful relationships between the values of those attributes"	1D vs 2D relationships

# Characterize distribution

---



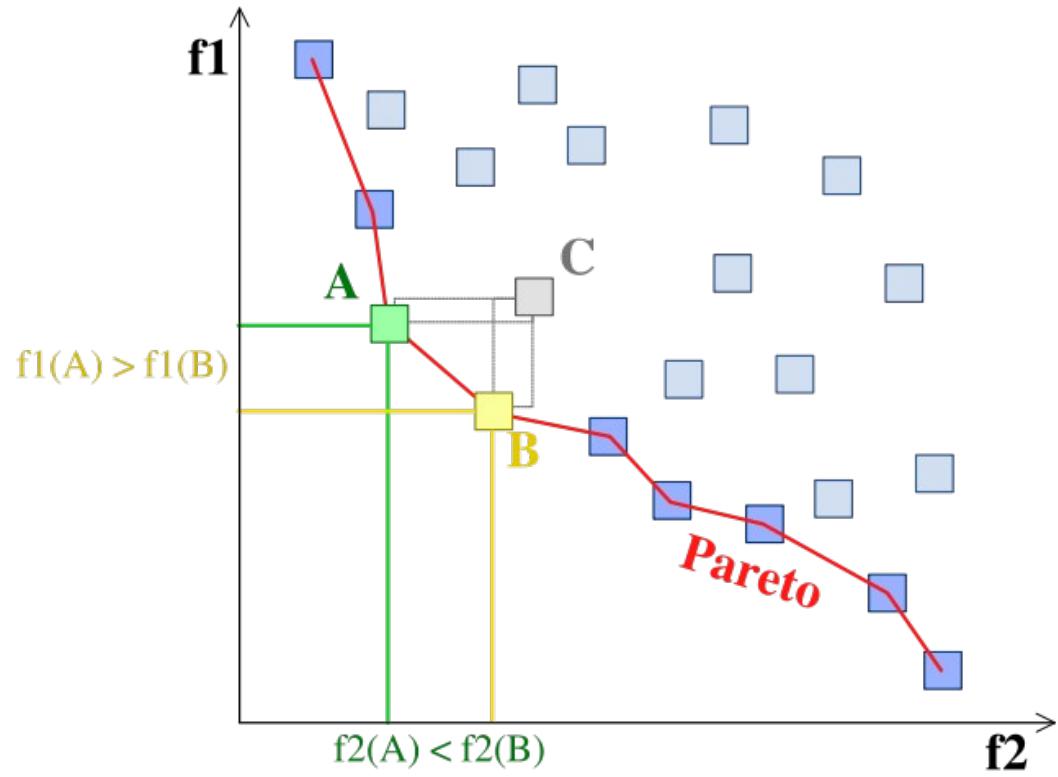
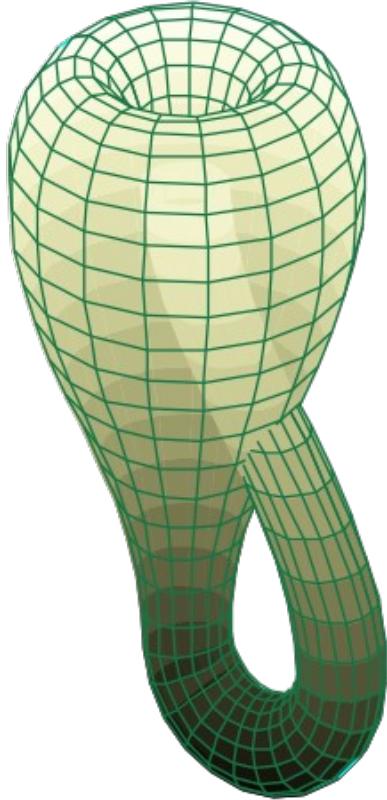
# Tasks

# Tasks

Our adaption to continuous scalar functions	QRI results					Expert study results			
	HyperSlice	Topological spines	Contour tree	Gerber et al.	1D slices	HyperSlice	Topological spines	Contour tree	Gerber et al.
Given an x, what is the function value?	Medium	Low	Low	High	Medium	Low	Medium	Low	High
For what parameter values is the function equal or over x?	Medium	High	Medium	Medium	Medium	Low	Medium	Medium	High
Summary statistics: variance, mean, SA	Low	Medium	High	High	Medium	Low	Medium	Medium	High
Find local/global min/max	Low	High	High	High	Medium	Low	Medium	Medium	High
What is the range of possible outputs?	Medium	Low	Low	Low	Medium	Low	Medium	Low	High
What types of shapes do the manifolds have	Medium	Low	Low	Low	Medium	Low	Medium	Low	High
Do areas of the manifold have shapes unlike any others	Low	Low	Medium	Low	Medium	Low	Low	Low	High
Are areas of the manifold similar to other types	Medium	Low	Low	Low	Medium	Low	Low	Low	High
1D vs 2D relationships	High	Low	Low	Low	Medium	Low	Low	Low	High

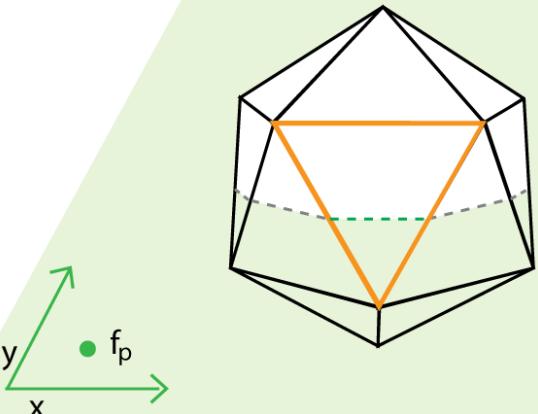
# Polytopes

---

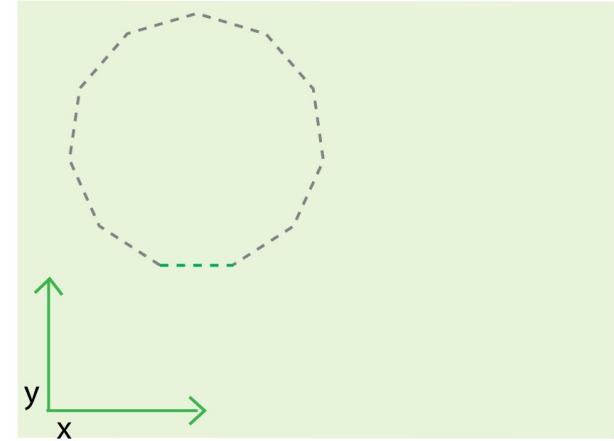
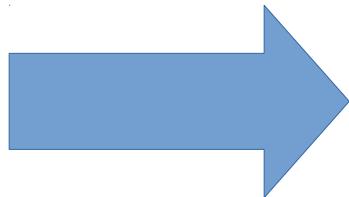


# Polytopes

---



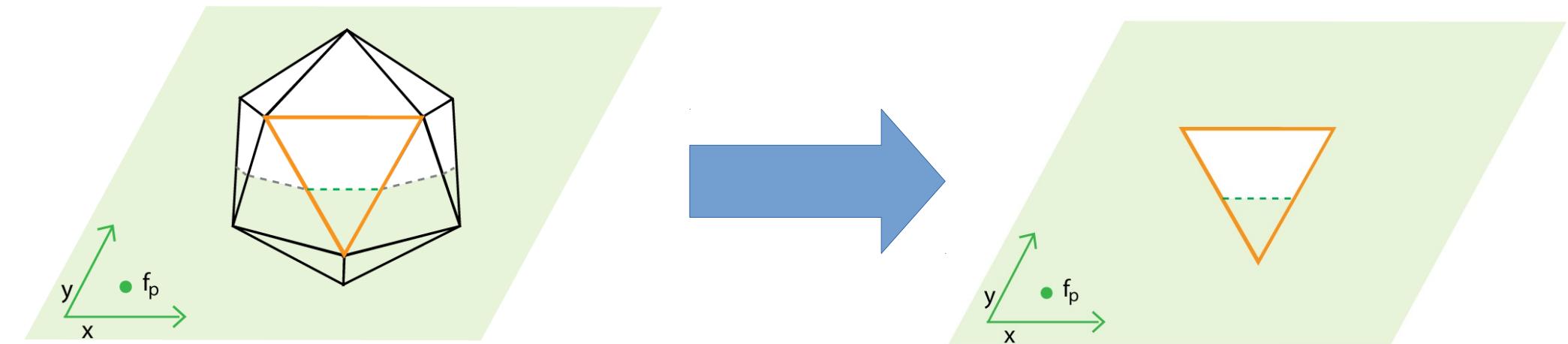
?



# Polytopes

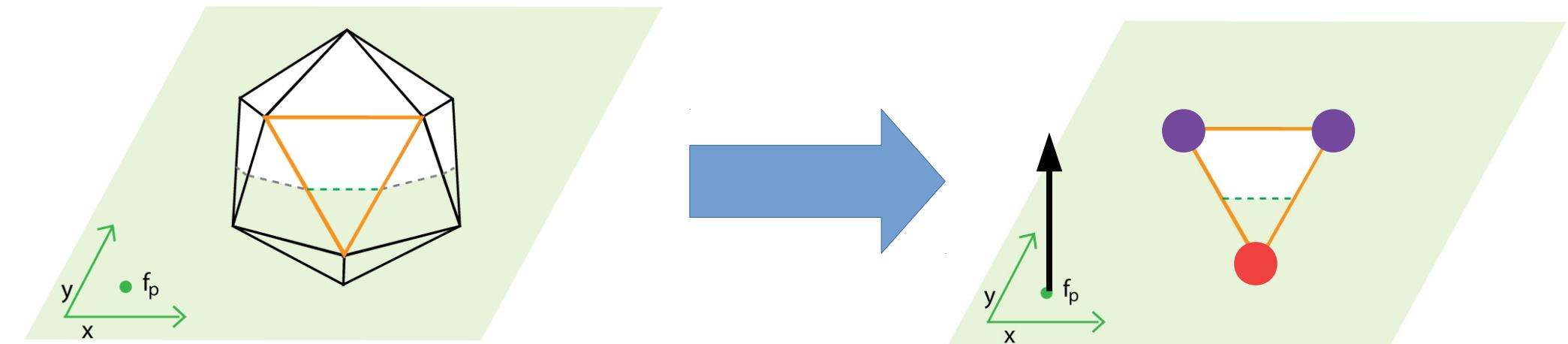
---

?



# Polytopes

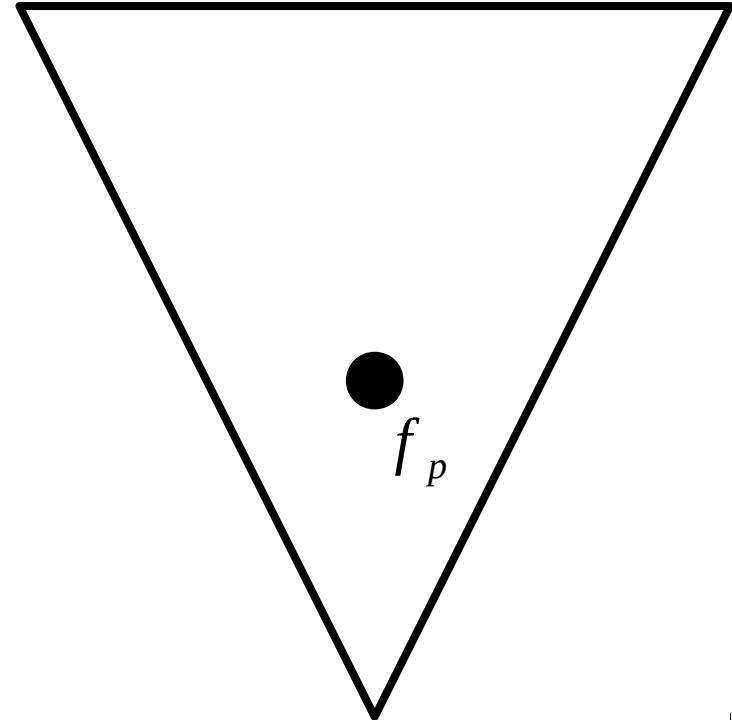
---



# Hypersliceplorer

---

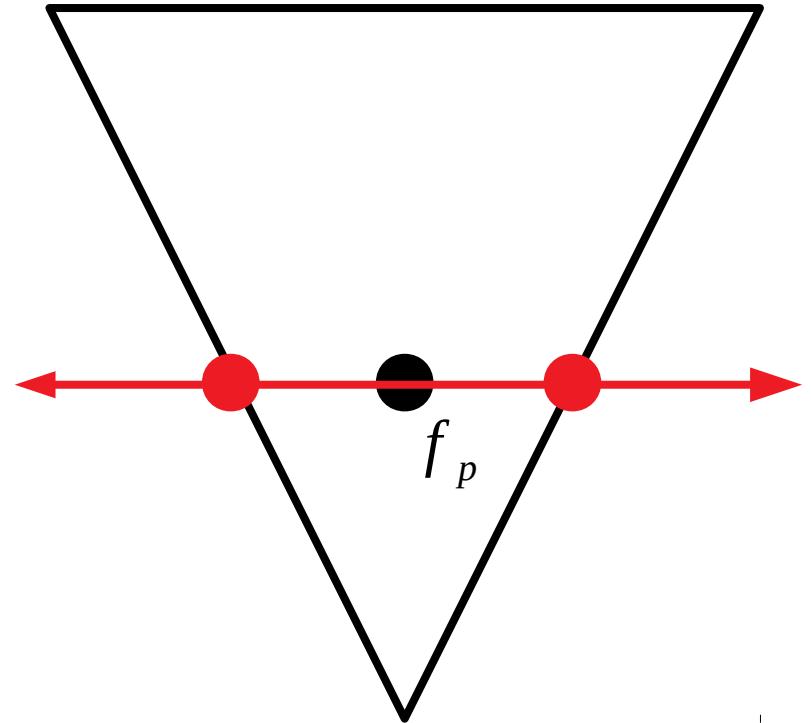
$$f_p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_d \end{bmatrix}$$



# Hypersliceplorer

---

$$f_p = \begin{bmatrix} x \\ p_2 \\ p_3 \\ \vdots \\ p_d \end{bmatrix}$$

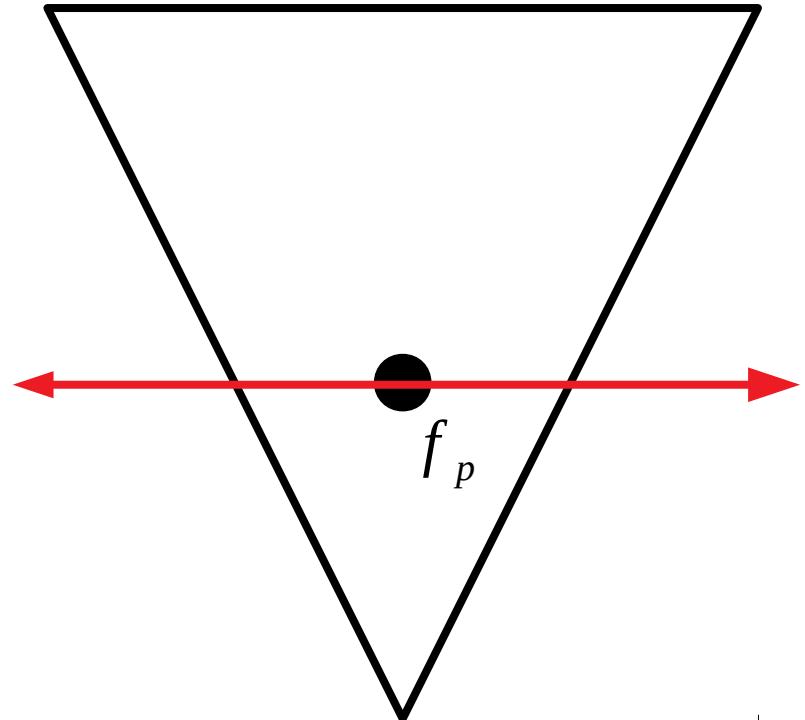


# Hypersliceplorer

---

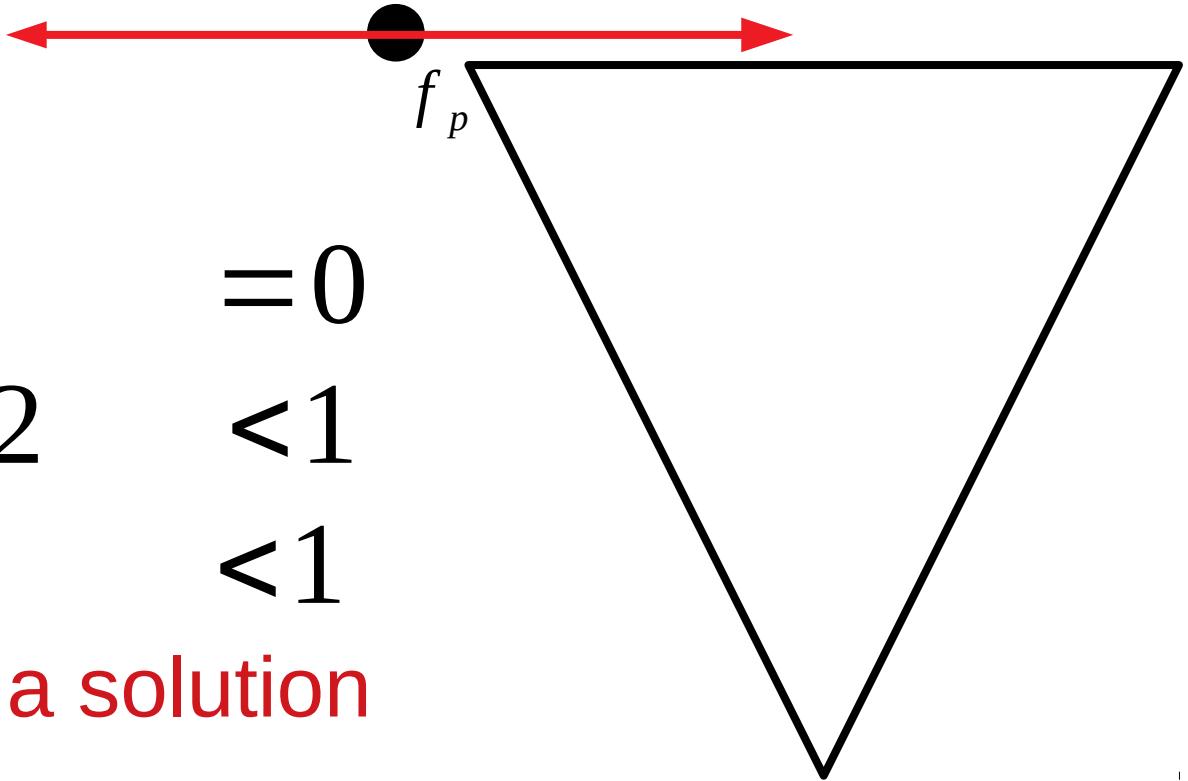
$$\begin{array}{ll} 0.24x + 0.27 & = 0 \\ -0.43x + 0.12 & < 1 \\ 0.78x - 0.2 & < 1 \end{array}$$

There will be a solution



# Hypersliceplorer

---

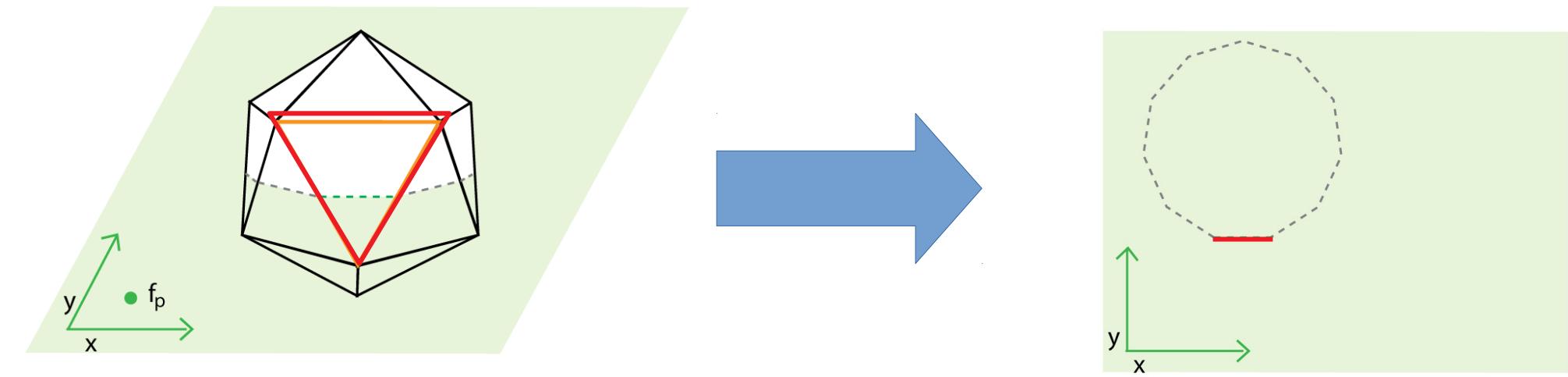


$$\begin{array}{ll} 0.24x + 0.27 & = 0 \\ -0.43x + 0.12 & < 1 \\ 0.78x - 0.2 & < 1 \end{array}$$

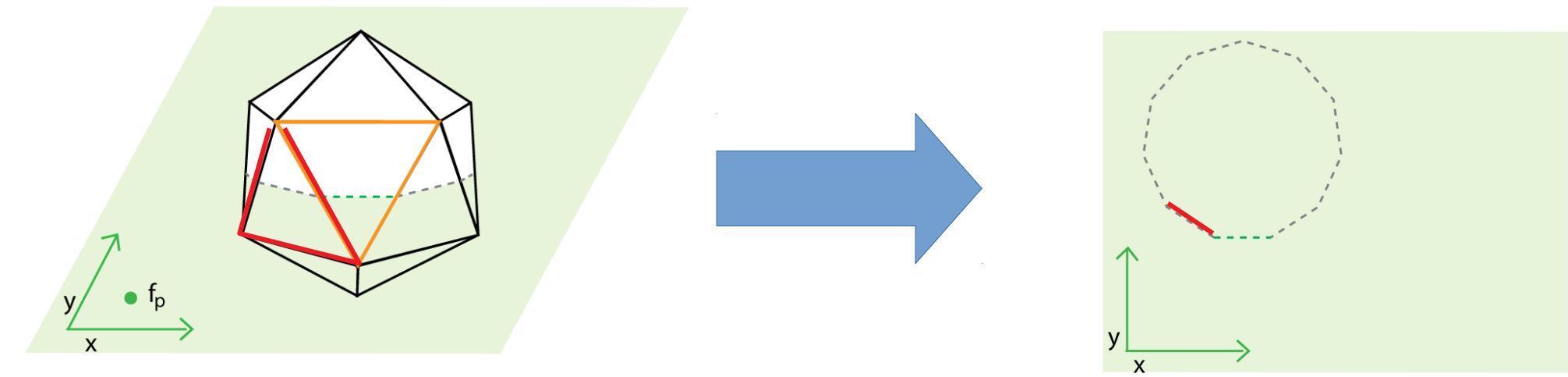
There will not be a solution

# Polytopes

---



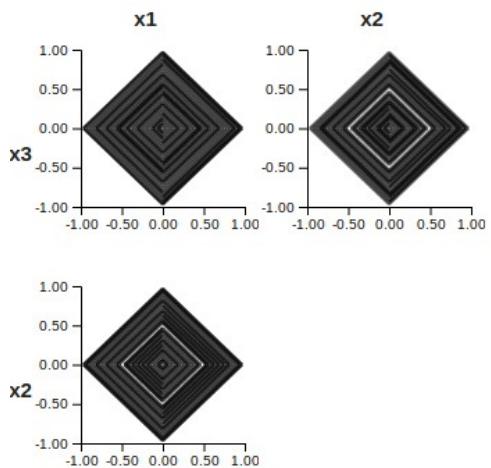
# Polytopes



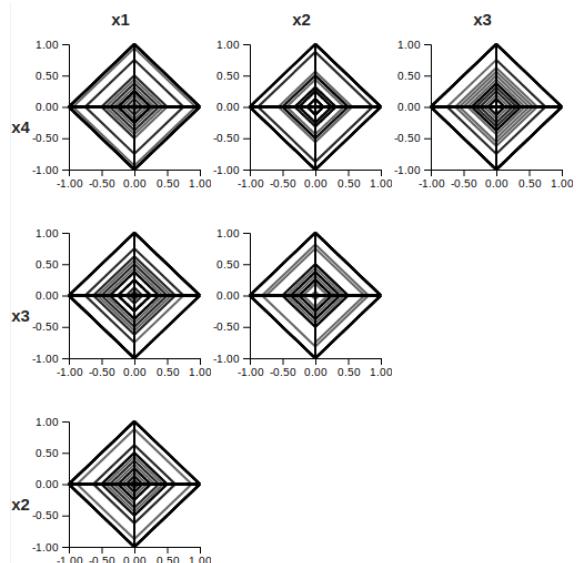
# Orthoplex

---

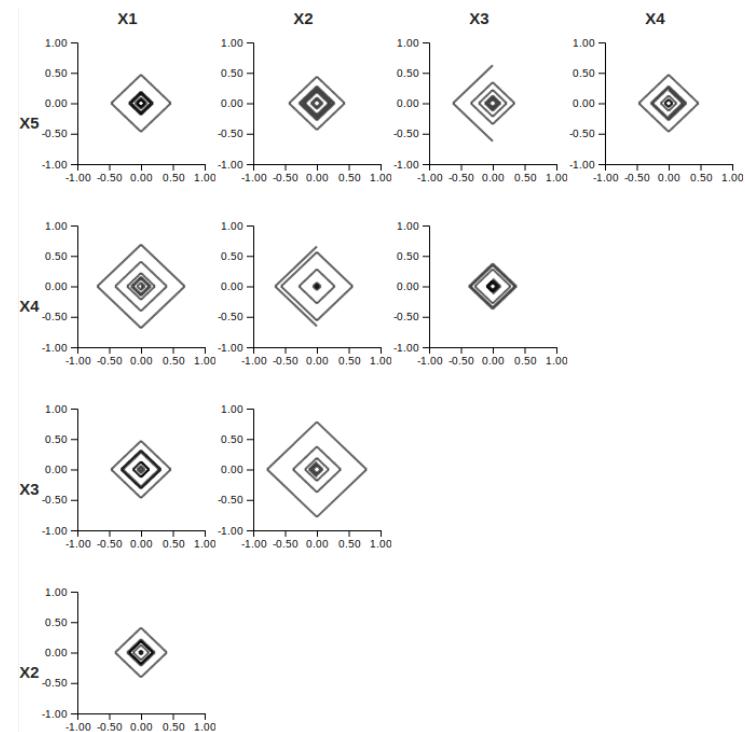
3D



4D

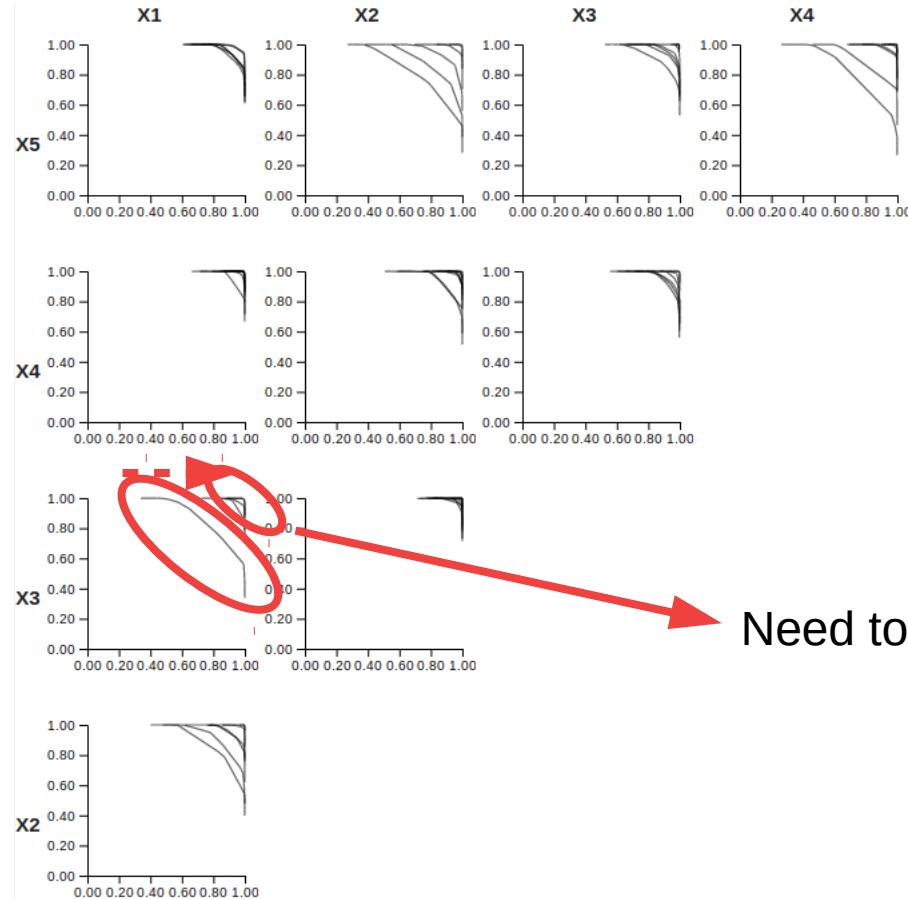


5D



# Pareto fronts

---



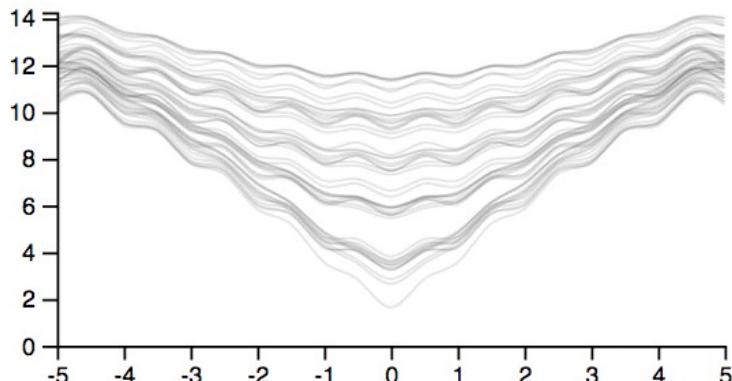
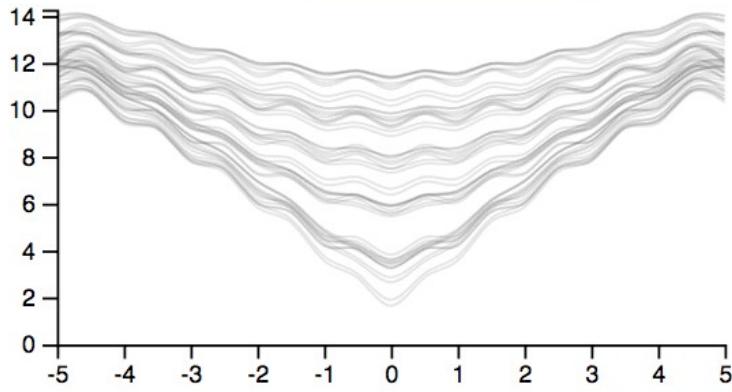
Need to lower other objective values

# Conclusion

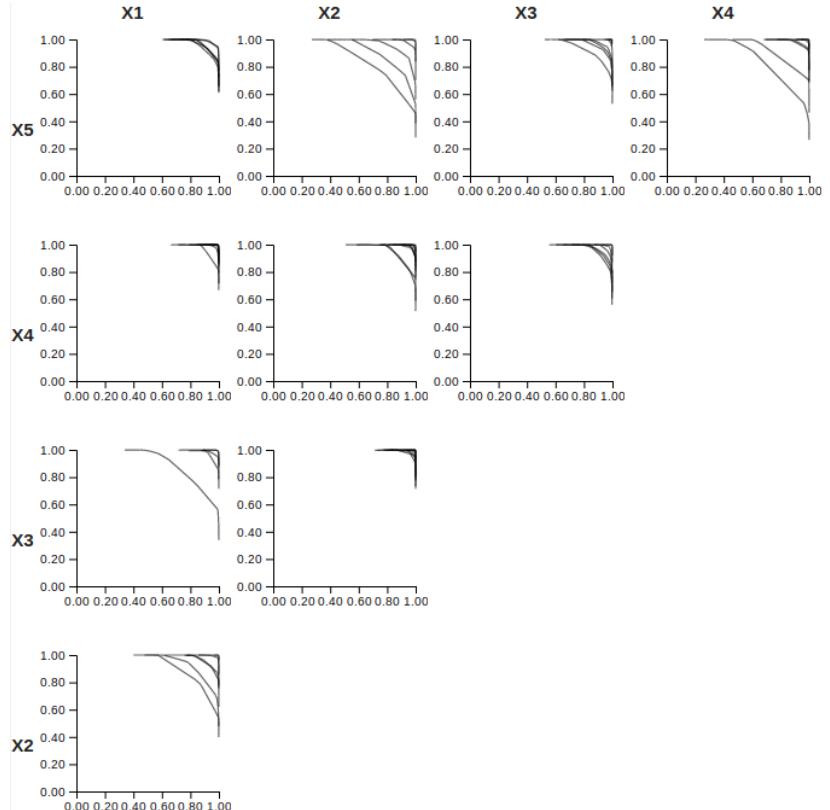
---

Considering data/tasks is vital!

1D: functions



2D: polytopes



# Future

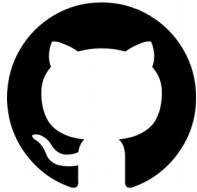
---

- R/Python library
- Detailed task performance
- Specific applications

# Thanks!

[www.tomtorsneyweir.com](http://www.tomtorsneyweir.com)

[thomas.torsney-weir@univie.ac.at](mailto:thomas.torsney-weir@univie.ac.at)



@gabysbrain

