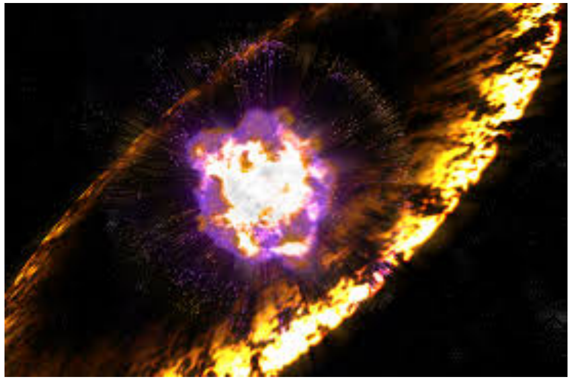
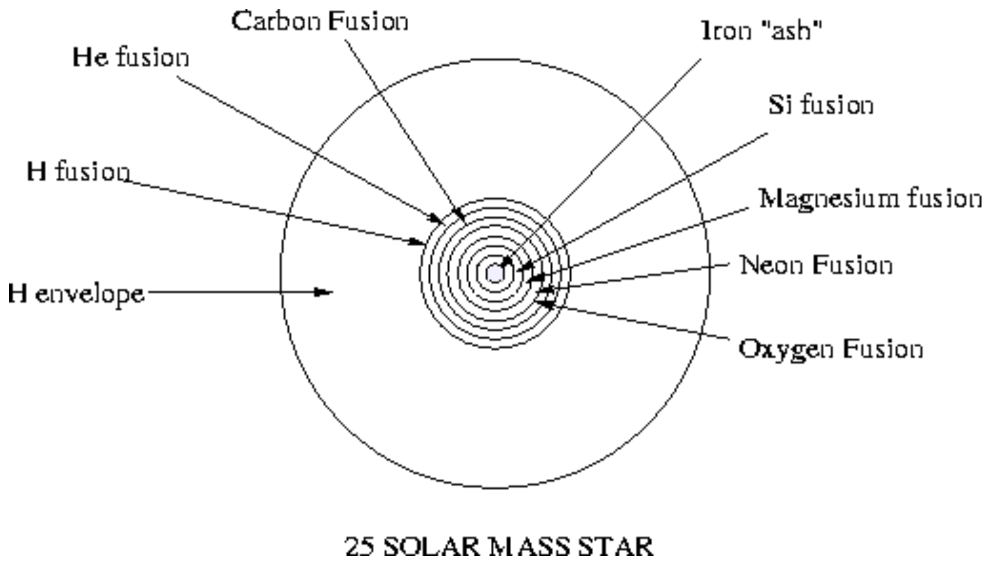
Supernova Light Curve Properties Clustering Analysis

Emir Karamehmetoglu

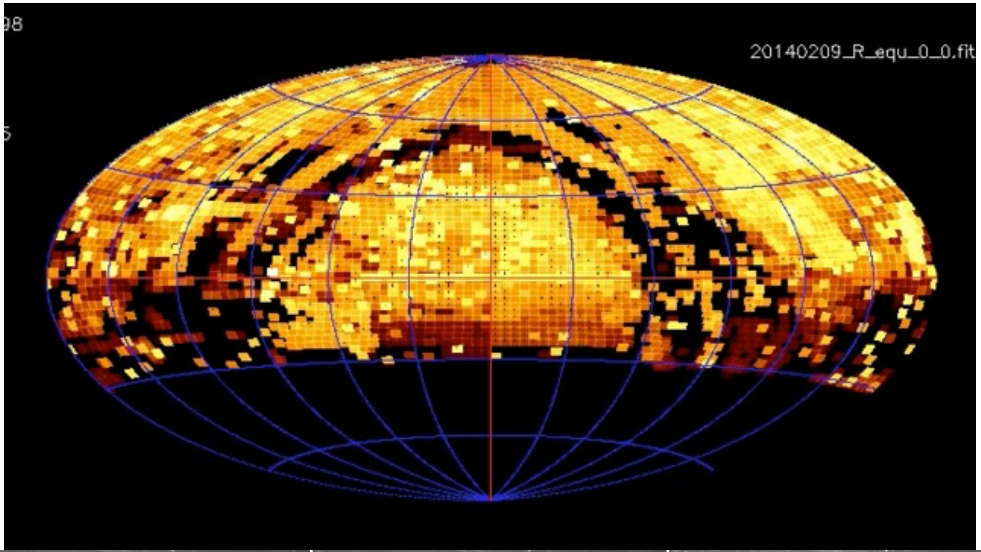


Stripped-Envelope Supernova

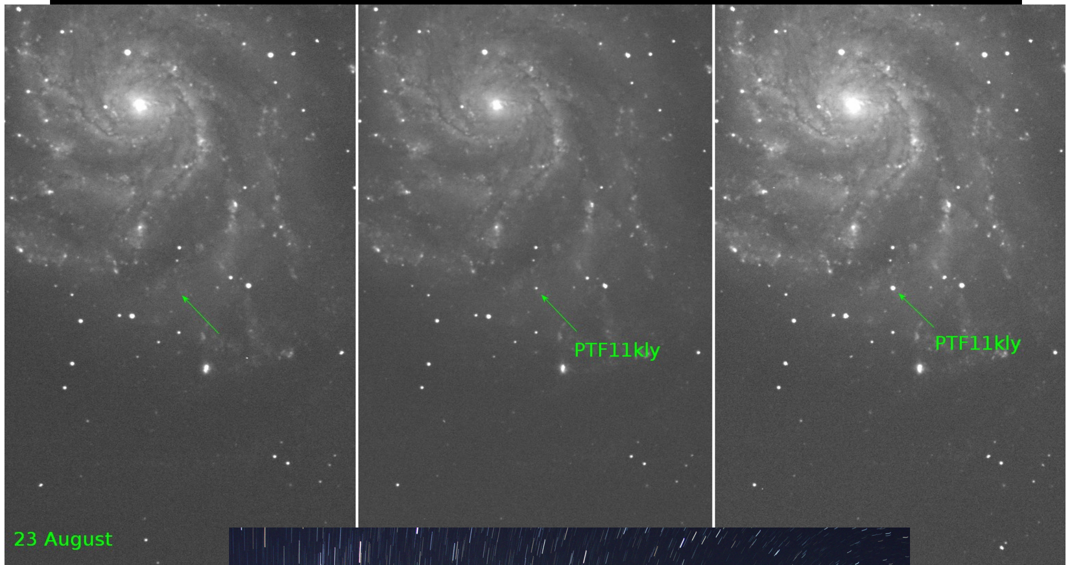
- Massive stars lose their outer envelopes, explode as "stripped-envelope" SNe.
- Spectroscopically they show little or no Hydrogen.
- Progenitors thought to lose their outer layers by winds or in a binary.



iPTF Stripped Envelope SN Dataset



Massive dataset of 1000's of SNe, a small part (About ~200) of which are stripped envelope (SE).

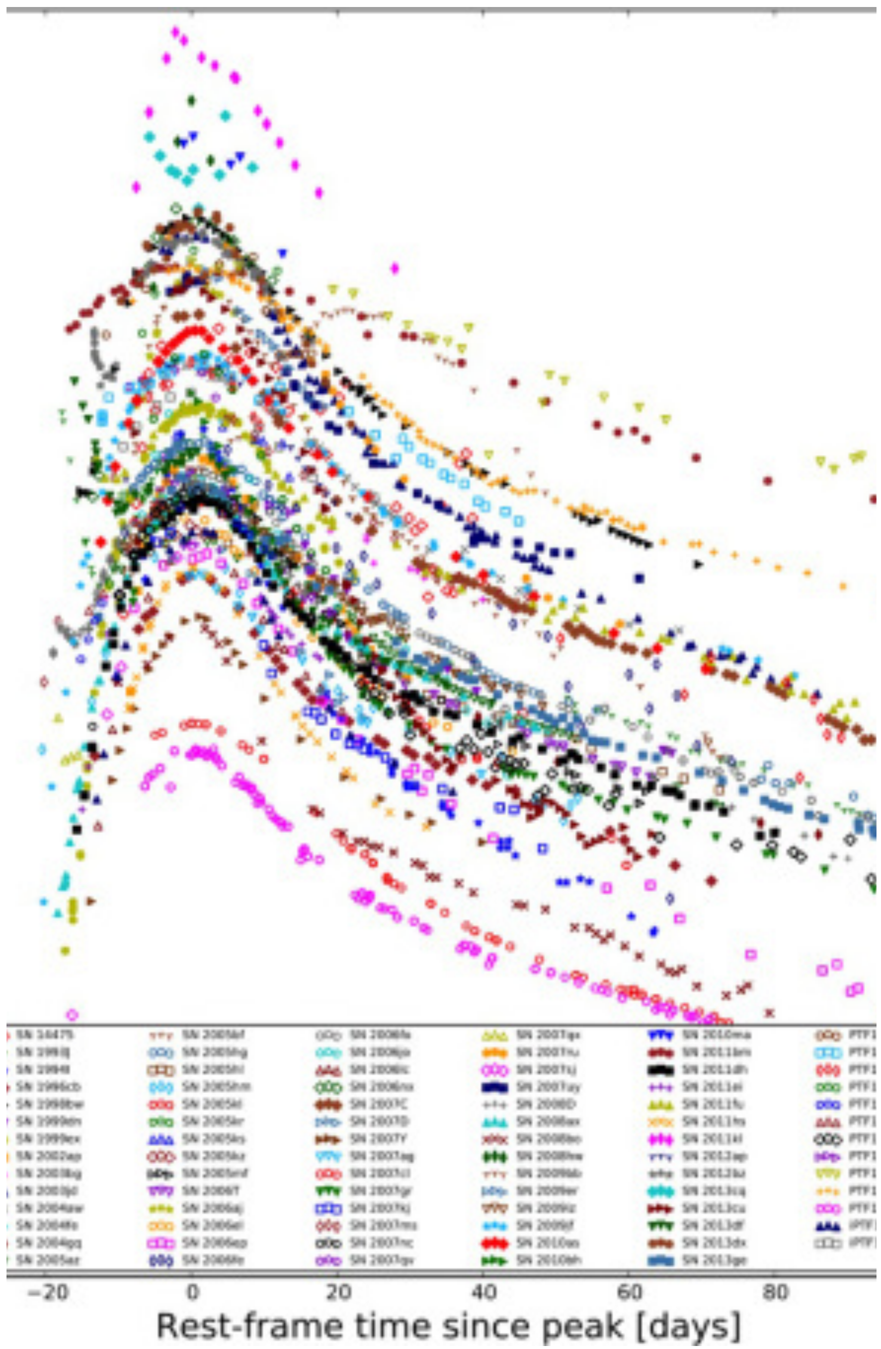


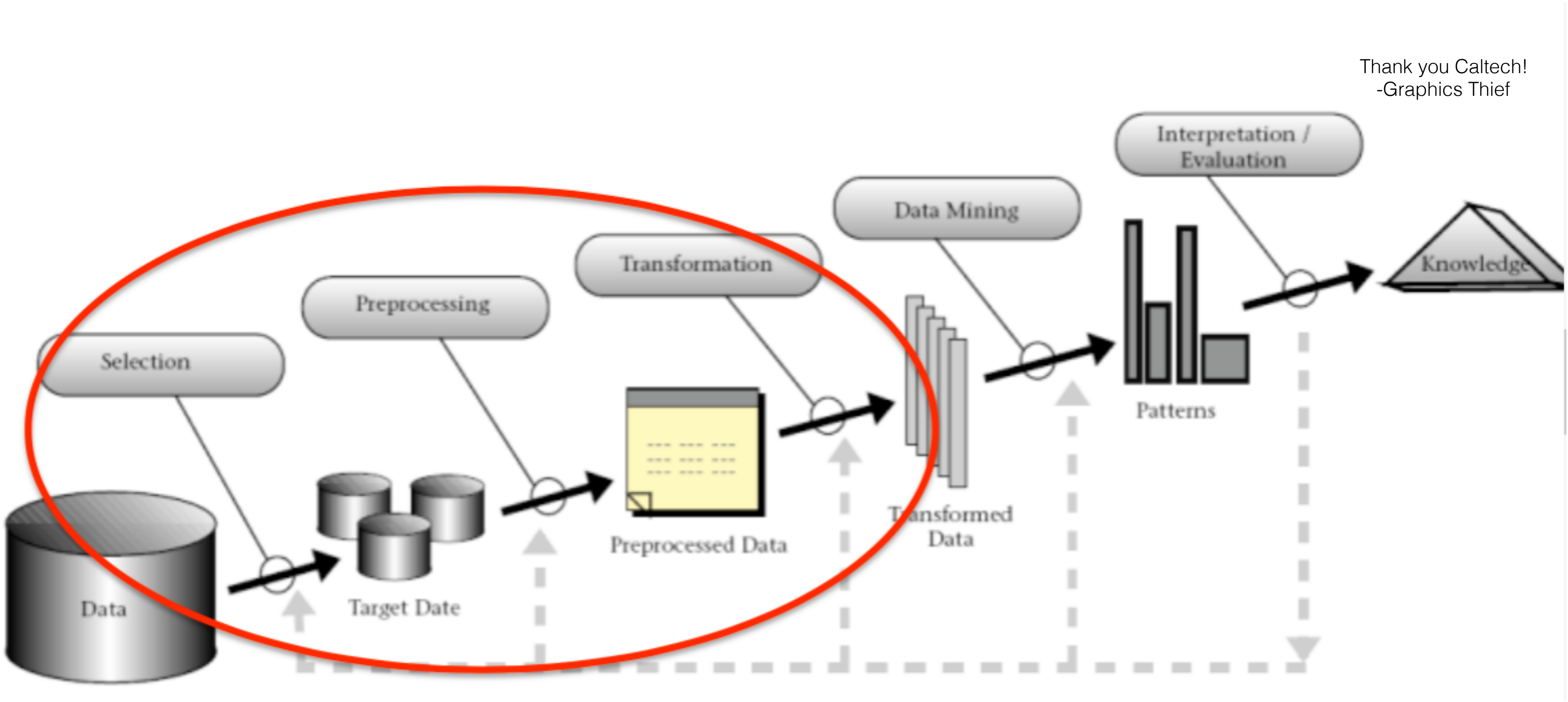
Lot's of light curves, and some spectra that allows us to characterize them as stripped envelope.

But if they come from similar progenitors and belong to the same spectral class, do they also have similar light curves?

Physical Motivation

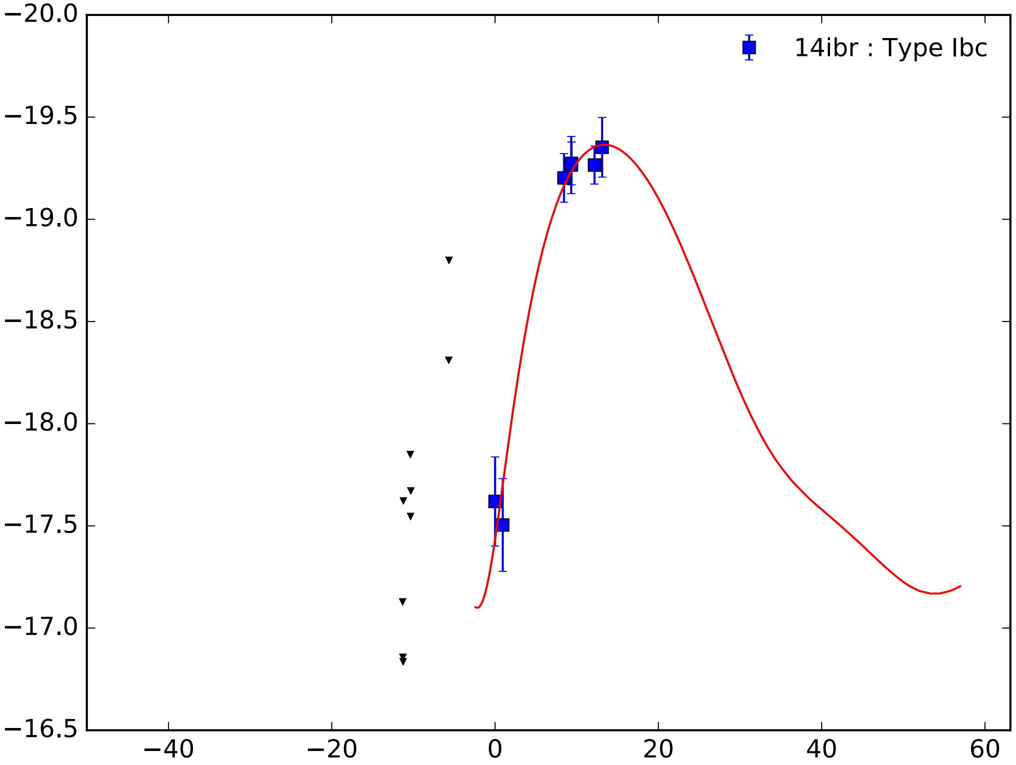
- SE Supernova (SN) Light Curves are powered radioactive decay of Ni56.
- Differences in light curves could be due to amount of Ni56 synthesized, amount of material ejected, distribution of Ni56, or due to alternate powering mechanisms.
- Finding those differences might allow us to pinpoint exactly which type of massive stars produce which SE SN.

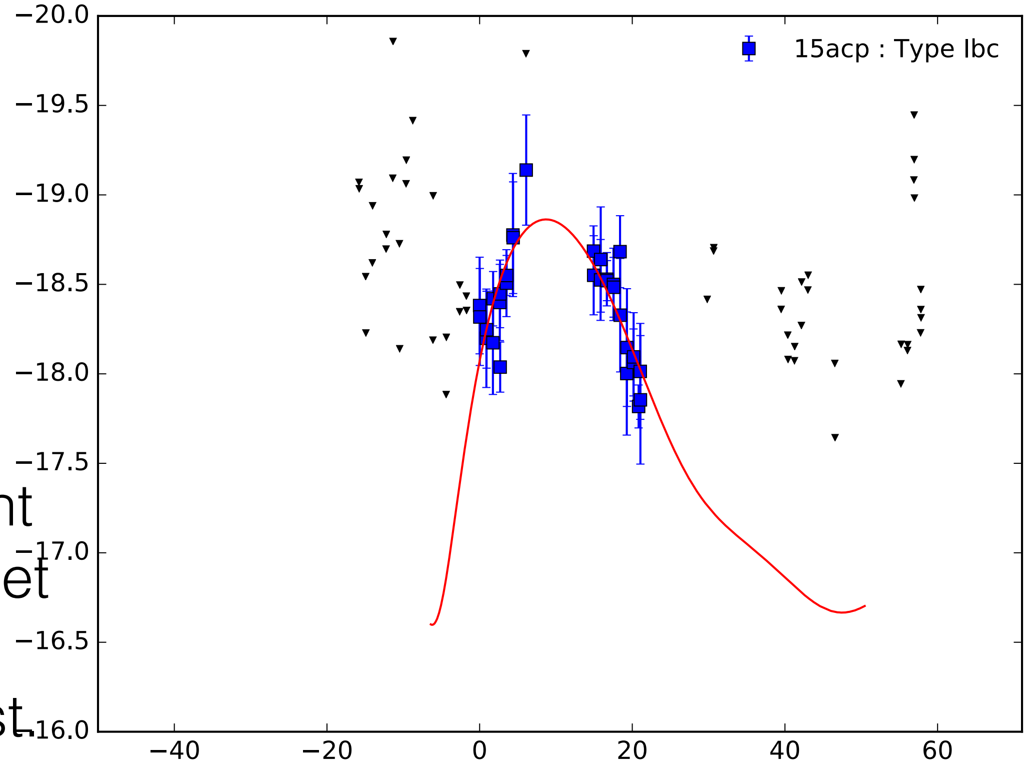




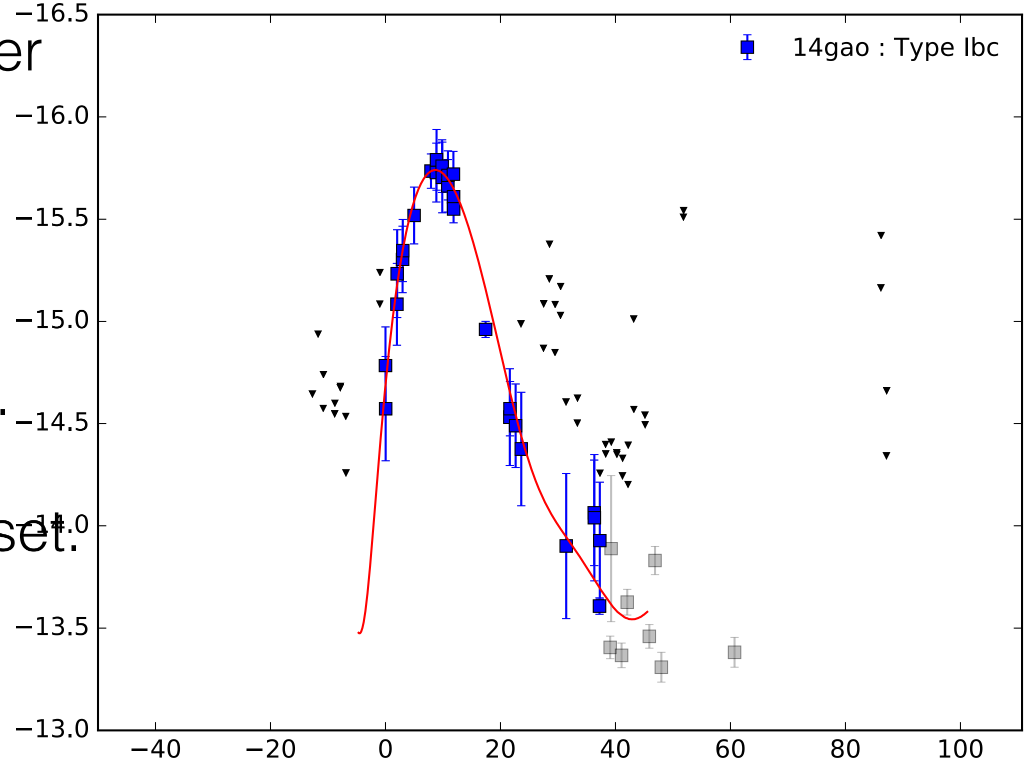
What Next?

"Knowledge Discovery in Databases"

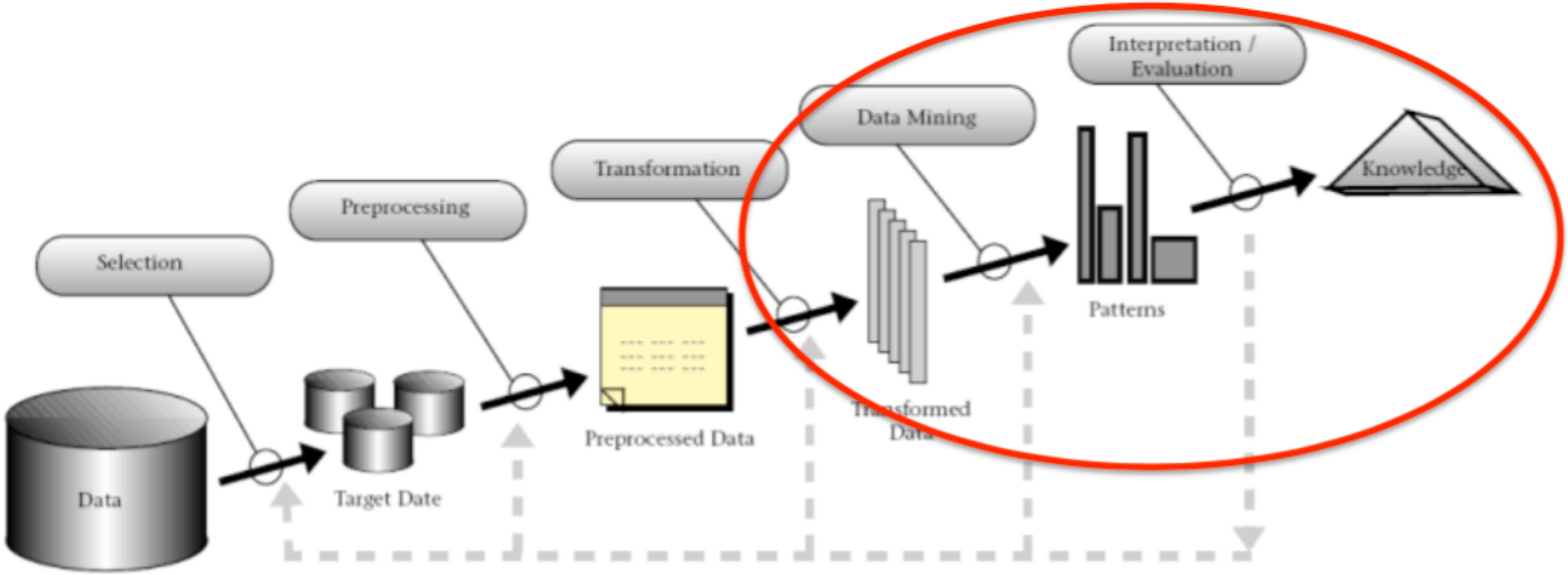




- However poor data called for another solution..
- Model fitting.
- This is a simple regression problem.
- Reduces dimensionality of the dataset.



Thanks again Caltech! Keep it up!
-Graphics Thief

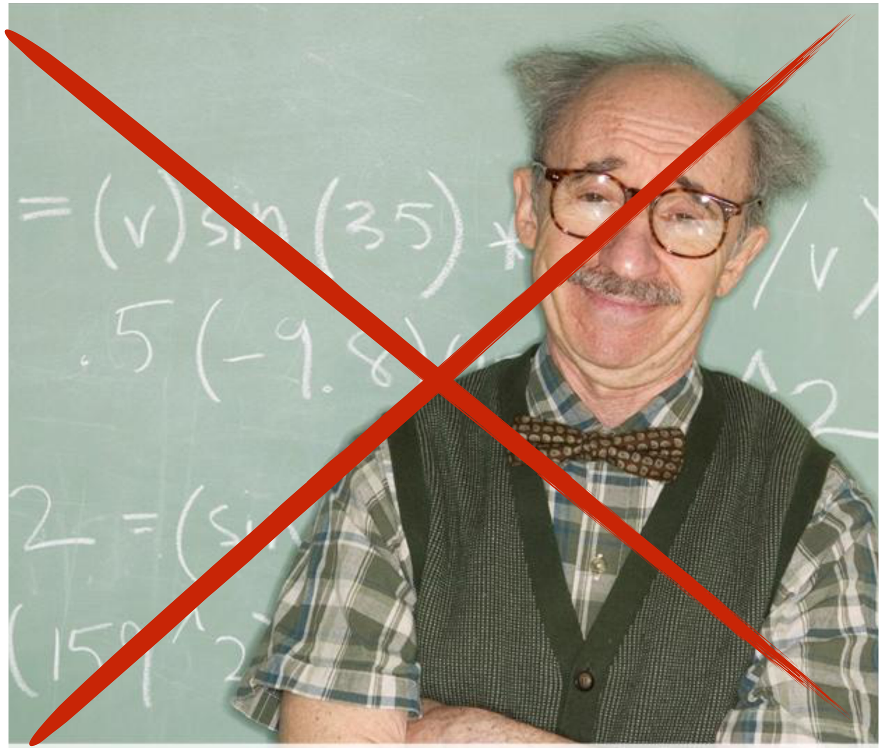


"Data mining"

Clustering/Classification/Regression/Visualization

Supervised or Unsupervised

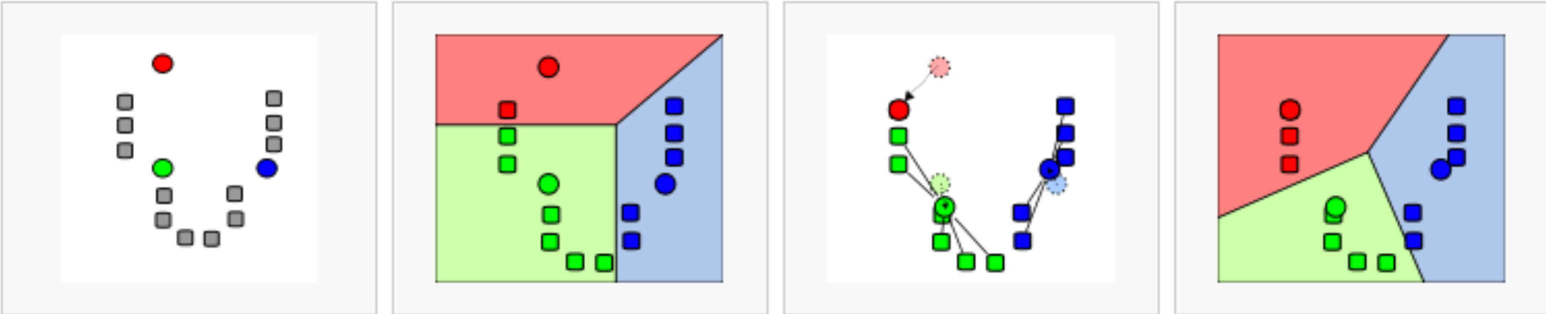
- Classification (Supervised):
  - Look for things that we want to find
  - One-SVM.
- Clustering (Unsupervised):
  - Find things and see if we even want them
  - Connectivity based, Centroid-based, Distribution based, Density based.
  - K-means and DBSCAN



$$\underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2$$

K-Means(++)

Demonstration of the standard algorithm



1. *k* initial "means" (in this case *k*=3) are randomly generated within the data domain (shown in color).

2. *k* clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.

3. The centroid of each of the *k* clusters becomes the new mean.

4. Steps 2 and 3 are repeated until convergence has been reached.

Thanks Wikipedia,
You're great!
-Graphics Thief

Why?

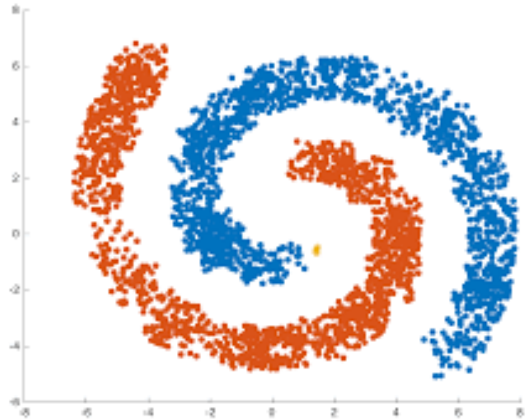
- Naturally expect to find at least one cluster of "normal" SESN.

- Also want to find relatively cleanly separated clusters, to find "real" different SN, not just clustering due to sampling bias.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN)

Ester at al. 1996

```
DBSCAN(eps=0.9,min_samples=5).fit(X)
```



- Grows from initial cores.

- Minimizes distance between nearby points.

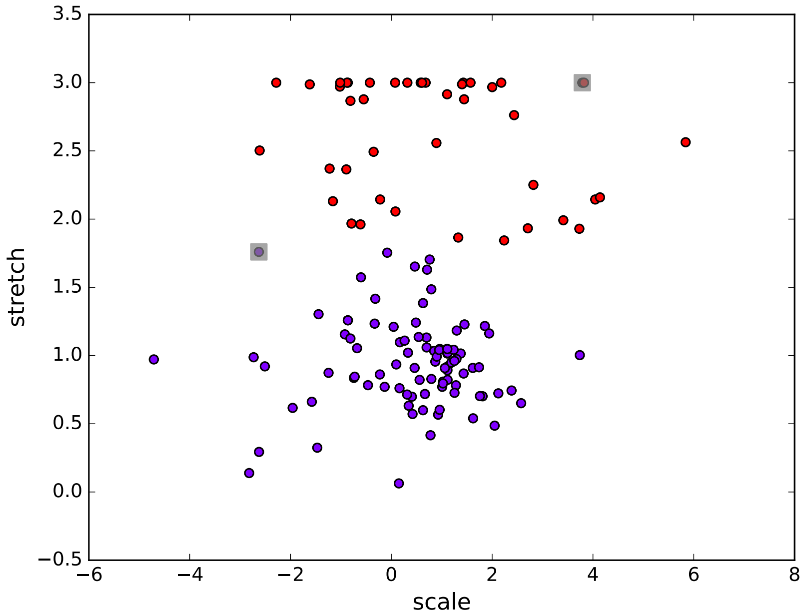
Why?

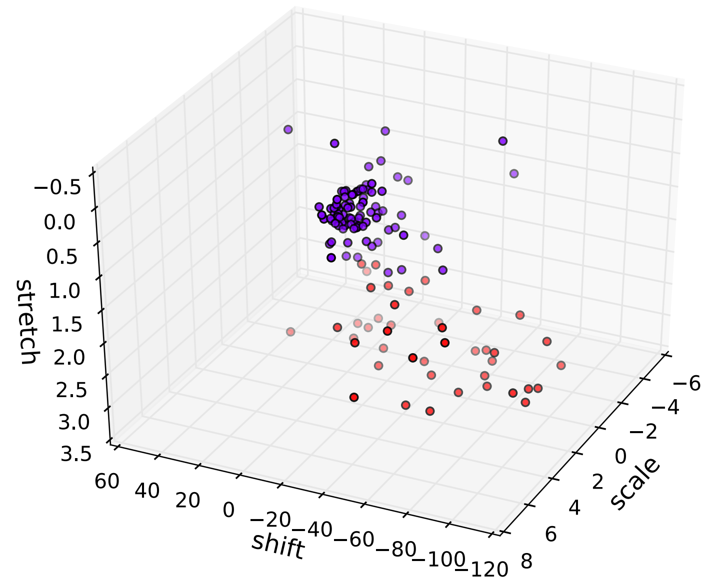
- Hierarchical Algorithm

- We do have noise in our data (Sampling bias/misfitting)

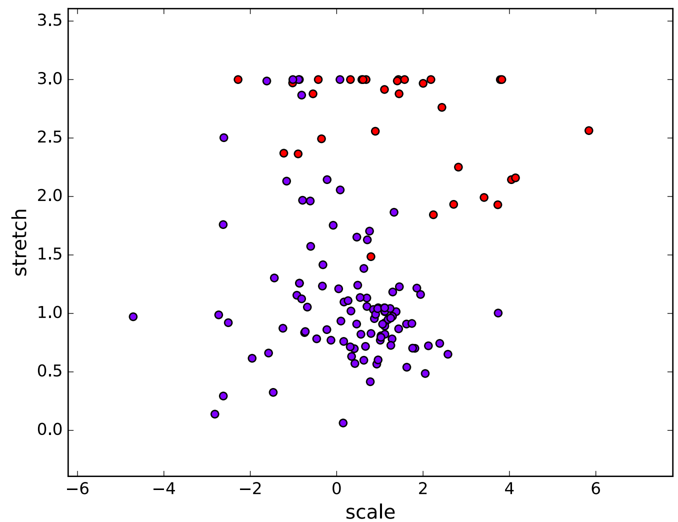
- Want to find dense regions representing possible SN sub-classes.

Some Results (See Demo)





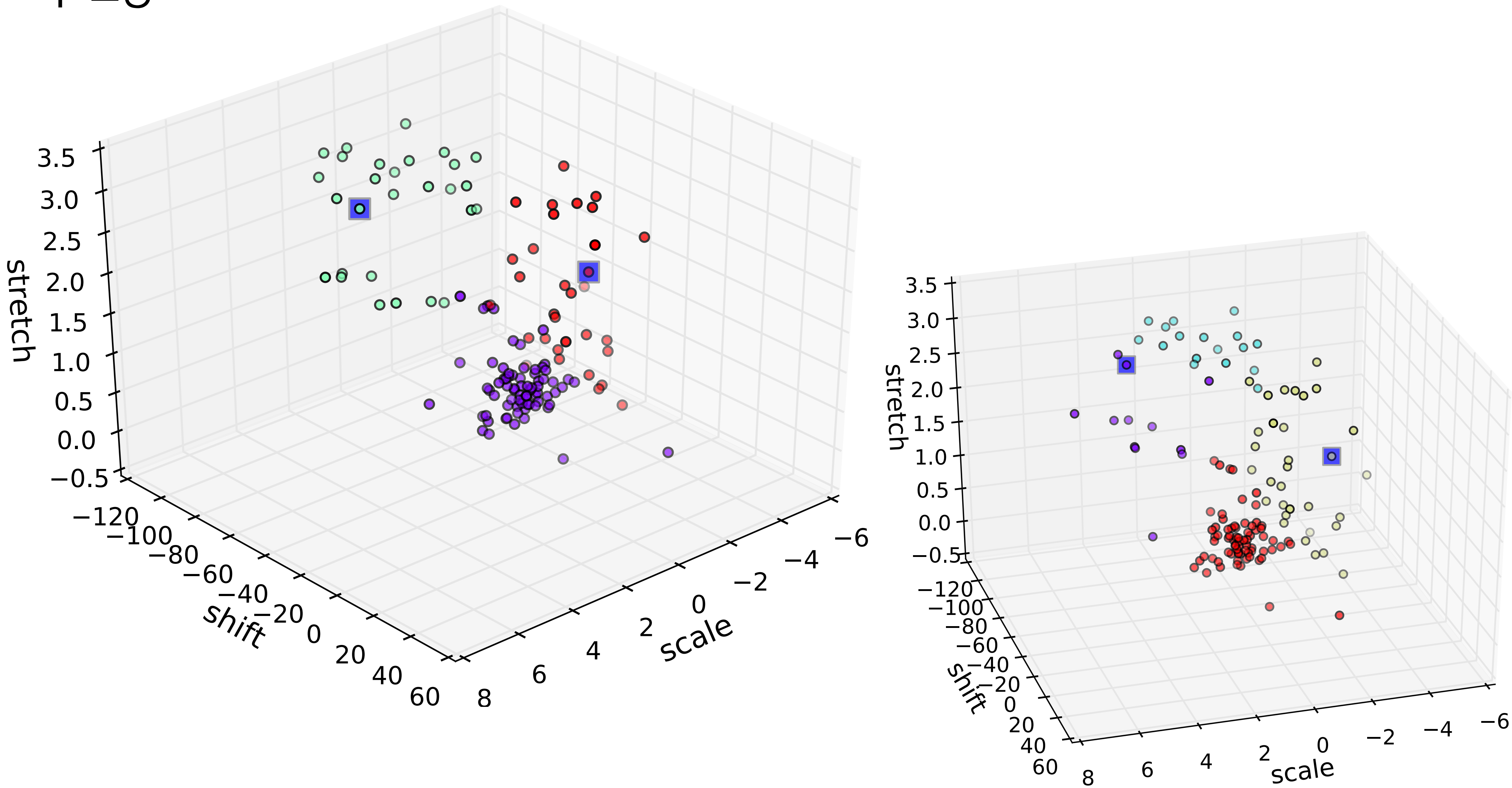
K=2
P=2



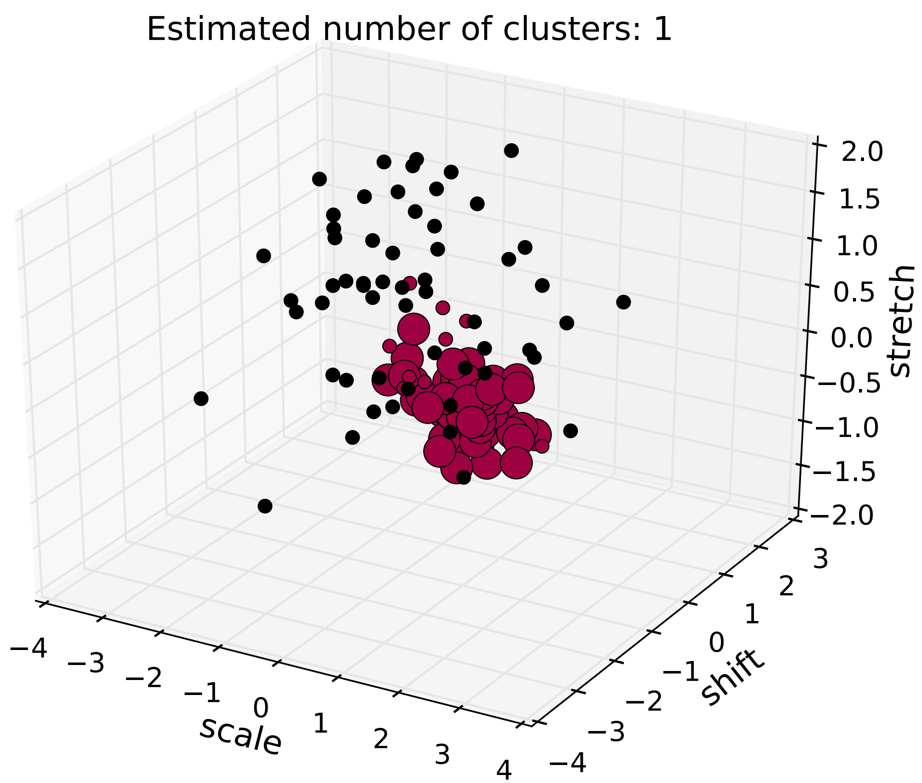
K=2
P=3

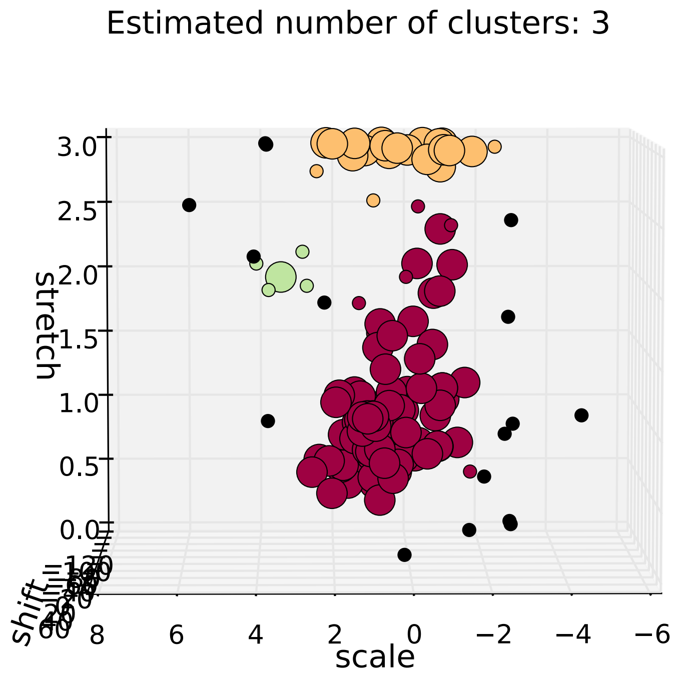
Some Results (See Demo)

K=3
P=3



Some Results (See Demo)



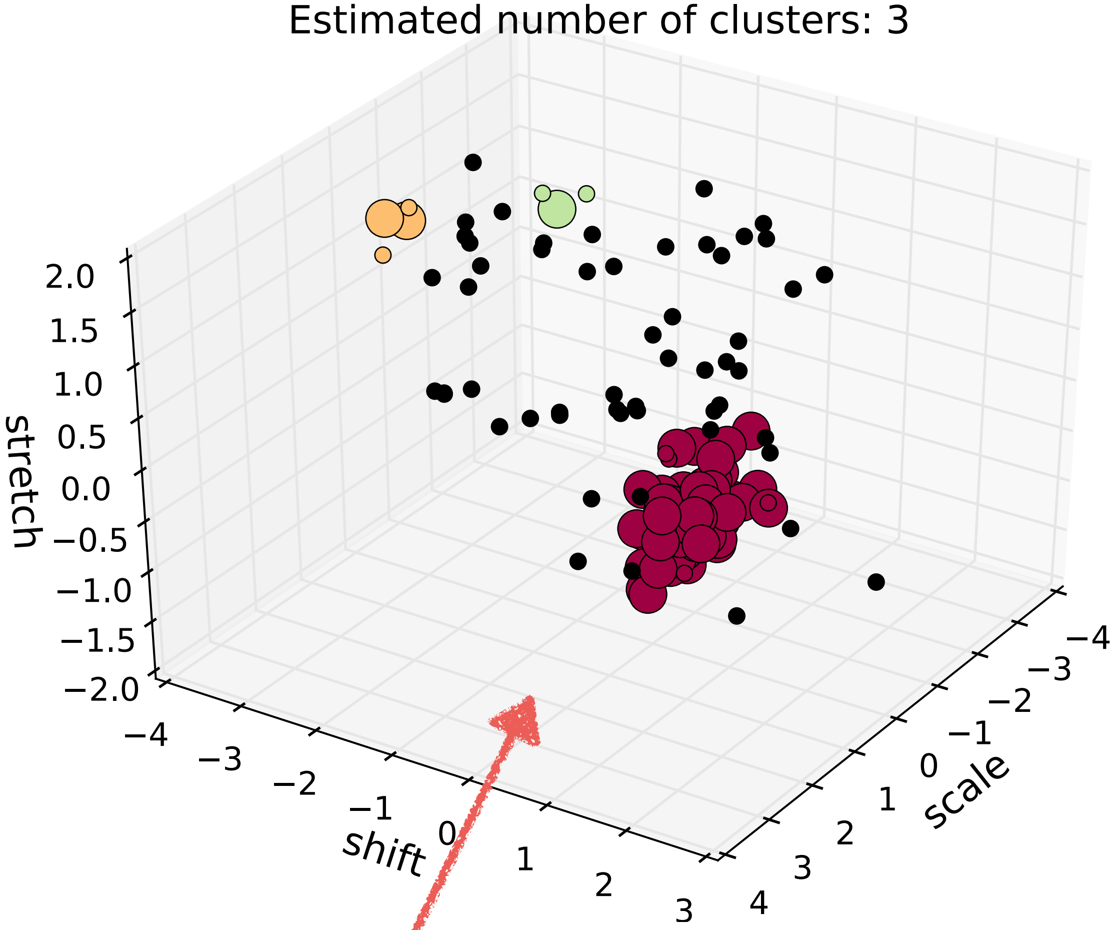


N_samples=High
E=low
P=3

(Strict)

N_samples=High
E=low
P=2

Some Results (See Demo)



N_samples=Default/Low
E=low
P=3

N_samples=Default/Low
E=High
P=3