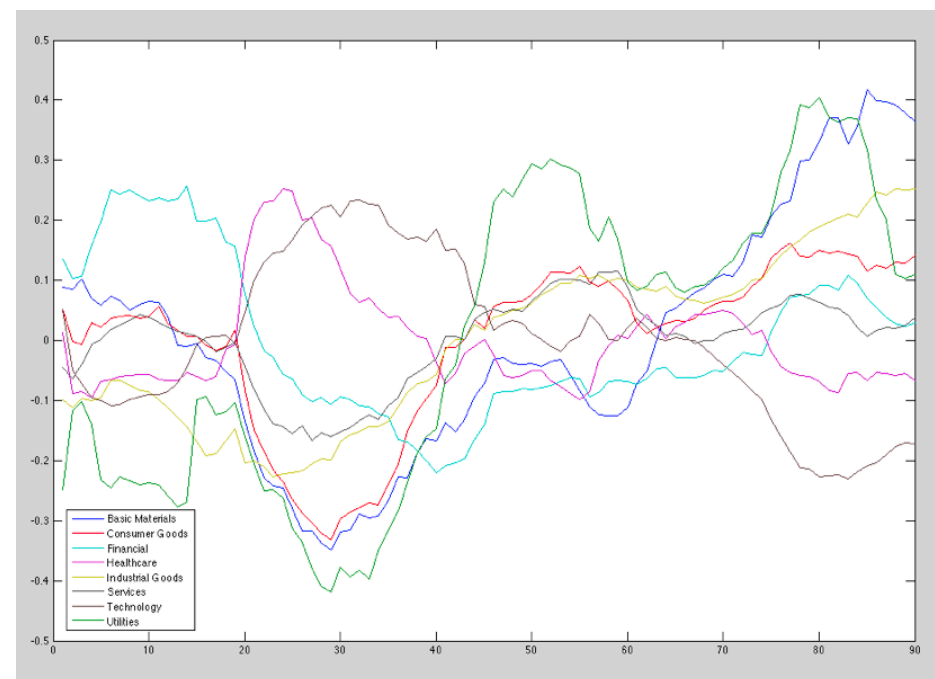


Unsupervised Learning in Complex Systems

Part D: An Introduction to the Vulcan Economy

CSSS 2008



Gregory Leibon
Memento Security & Dartmouth College

What we do...

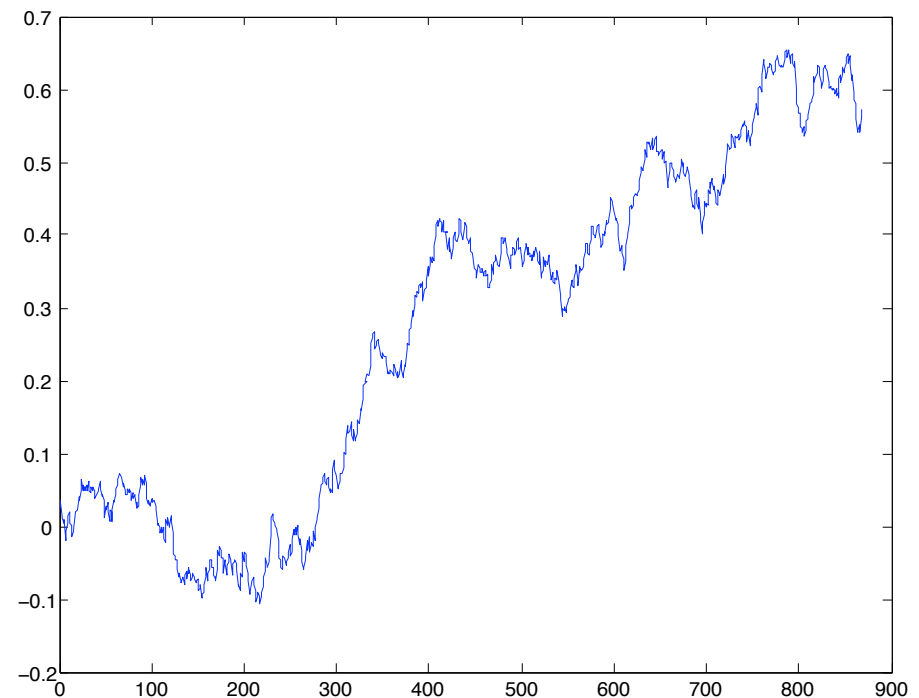
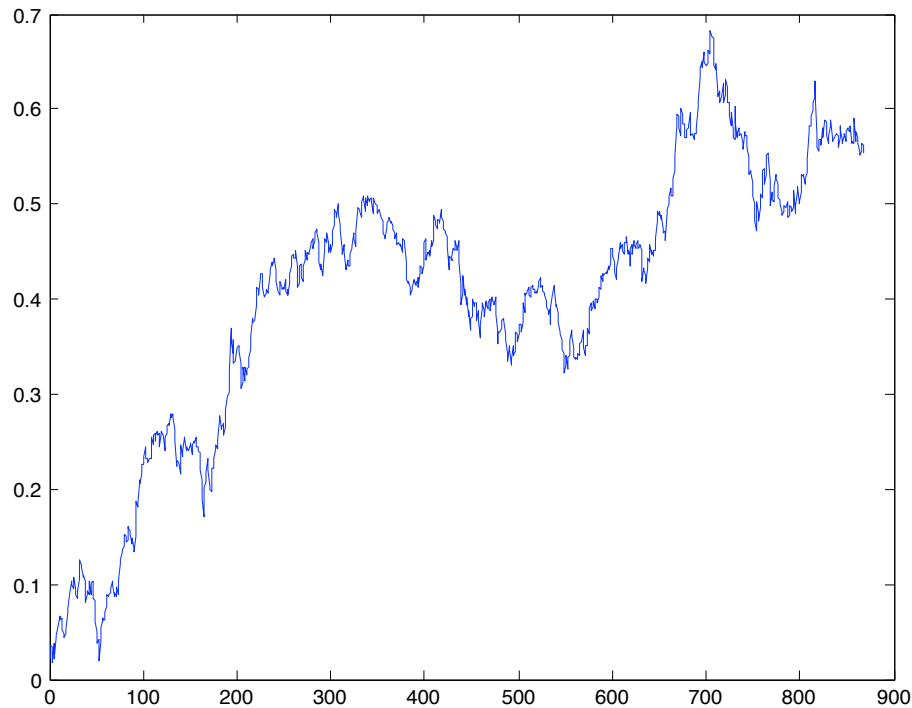
- Explore the use of simulation by examining the Vulcan economy.
- Describe the Partition Decoupling Method.
- Examine the model in the Equities Market and the 109th congress.

Vulcan Economy



Vulcan Equities

$$X_t = \frac{S_t - S_{t-1}}{S_{t-1}} \approx d(\ln(S_t))$$



$$d(\ln(S_t)) = \sigma(t)dB_t + c(t)dt$$

Correlation Metric

Normalize: $\hat{X} = \frac{X - \langle X \rangle}{\sqrt{\langle (X - \langle X \rangle)^2 \rangle}}$

Correlation: $\rho(X, Y) = \hat{X} \cdot \hat{Y}$

Metric: $d(X, Y) = 2 \sin(\theta/2) = \sqrt{2(1 - \rho(X, Y))}$

Vulcan
Examples:

N=500

L=1,000

Daily Ticks

Can you hear the shape of the Vulcan market?

$$\hat{X} = \frac{X - \langle X \rangle}{\sqrt{\langle (X - \langle X \rangle)^2 \rangle}}$$

$$\rho(X, Y) = \hat{X} \cdot \hat{Y}$$

$$d(X, Y) = 2 \sin(\theta/2) = \sqrt{2 (1 - \rho(X, Y))}$$

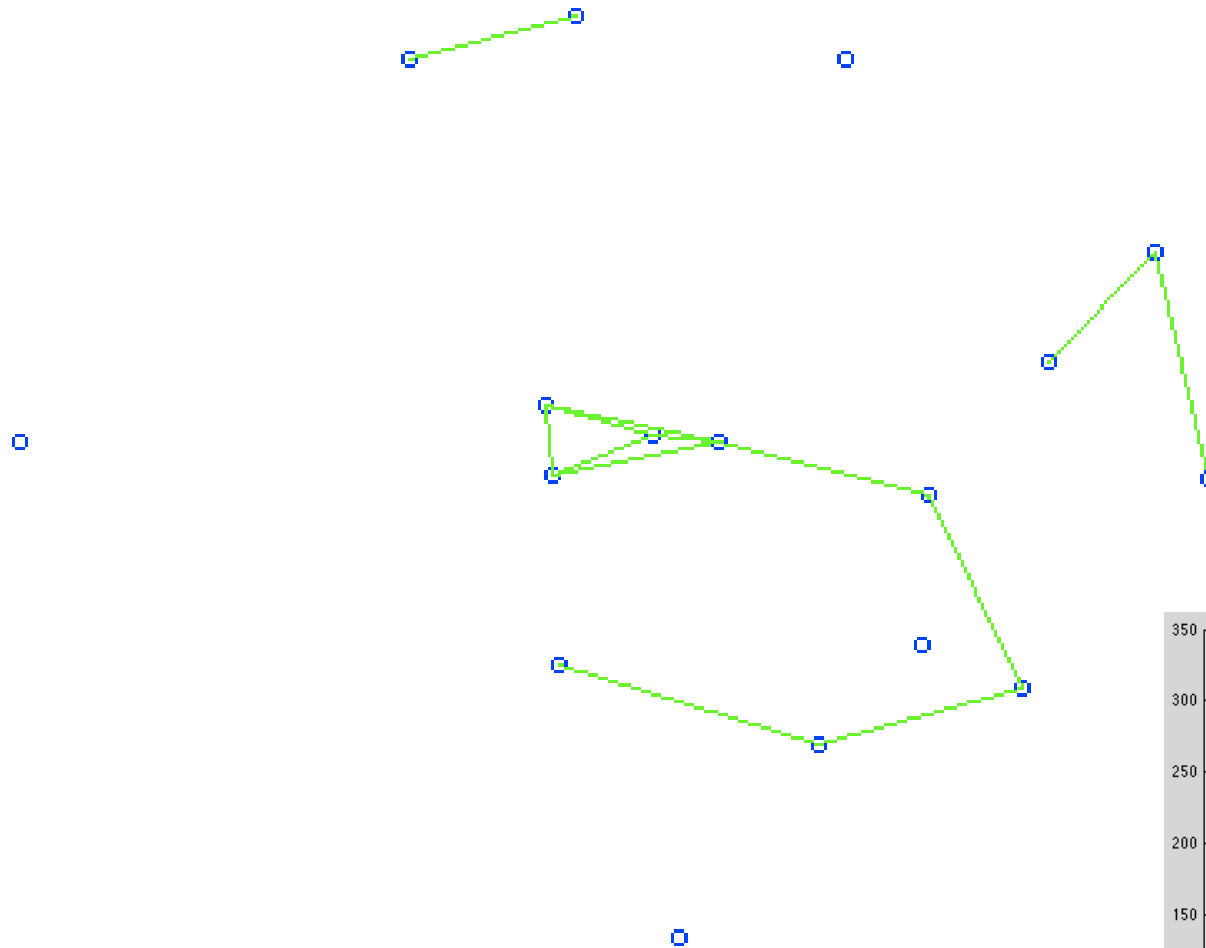
Candidate
Weights

$$\text{Conductance}_s = e^{\frac{-d(X, Y)^2}{\sigma^2}}$$

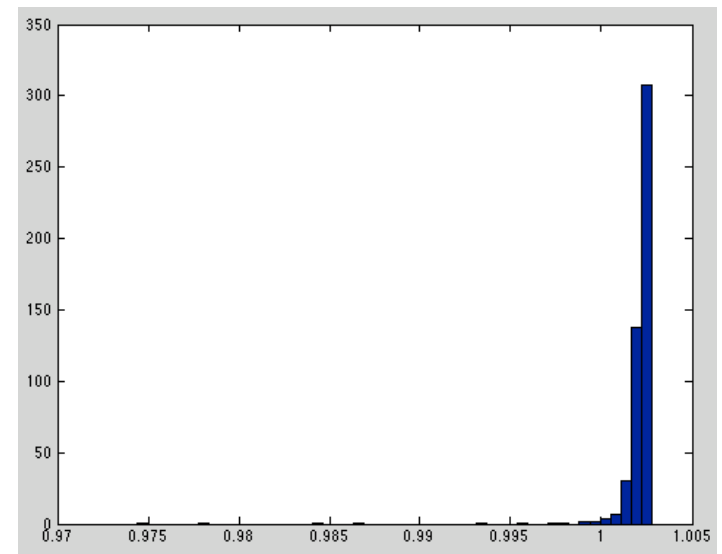
$$\text{Conductance}_e = \frac{1}{d(X, Y)}$$

...and zero on
diagonal,
force a move.

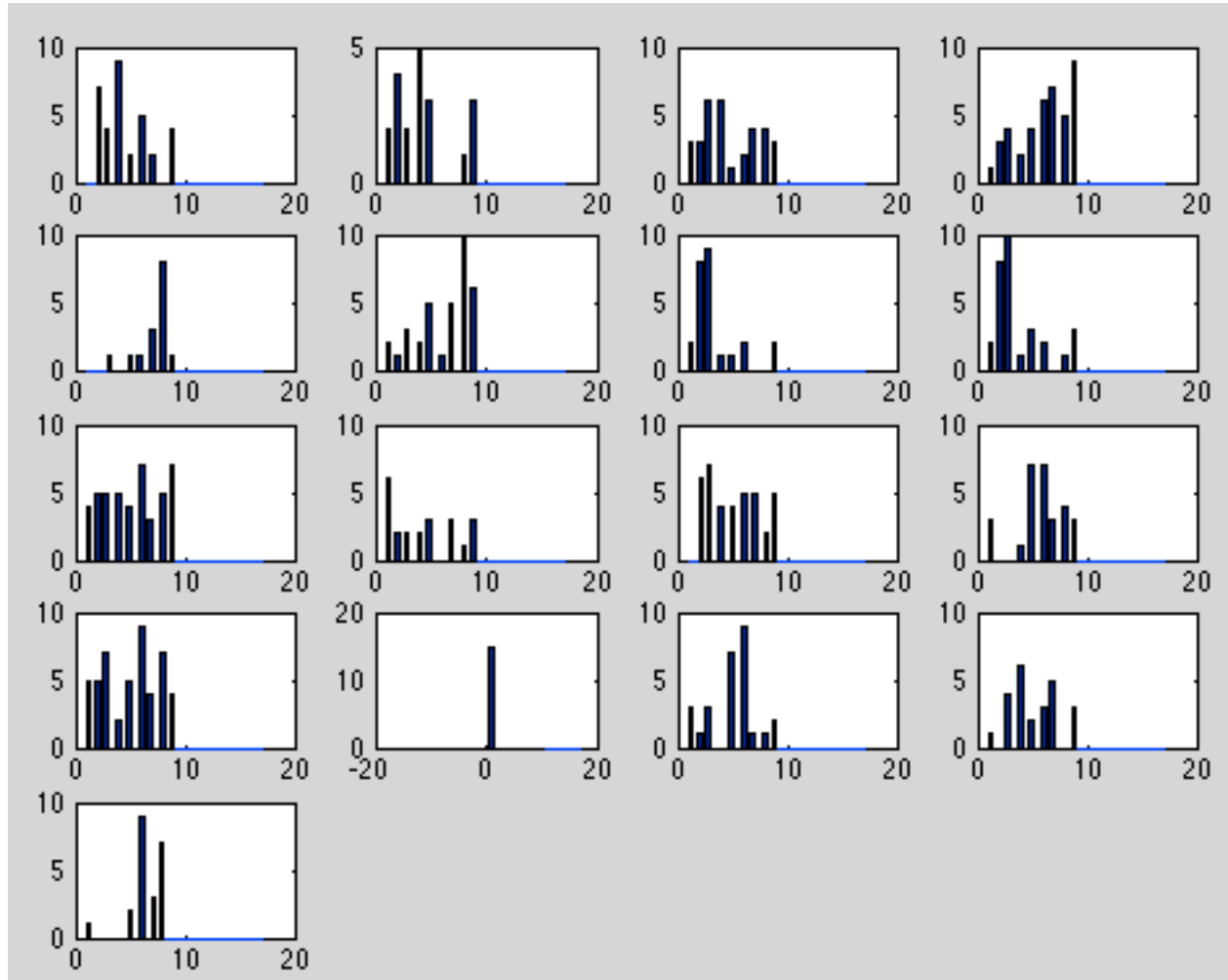
17 Clusters



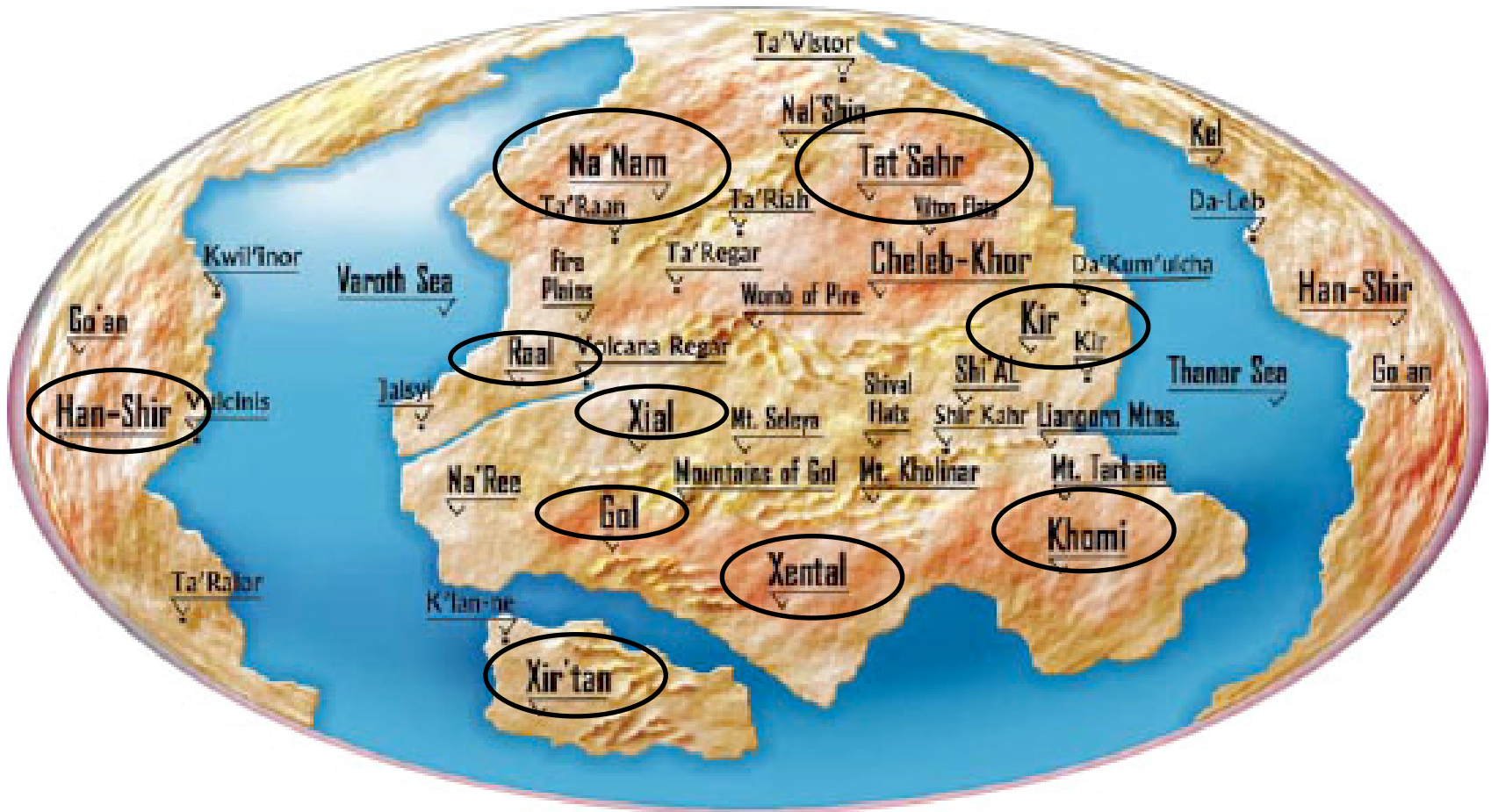
Original
Distribution



Not Sectors....

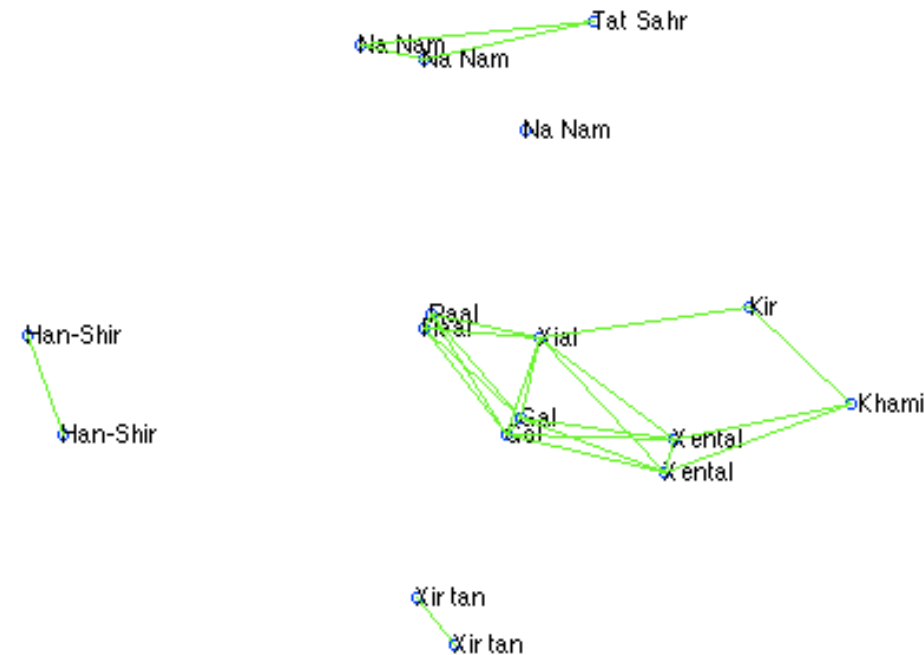
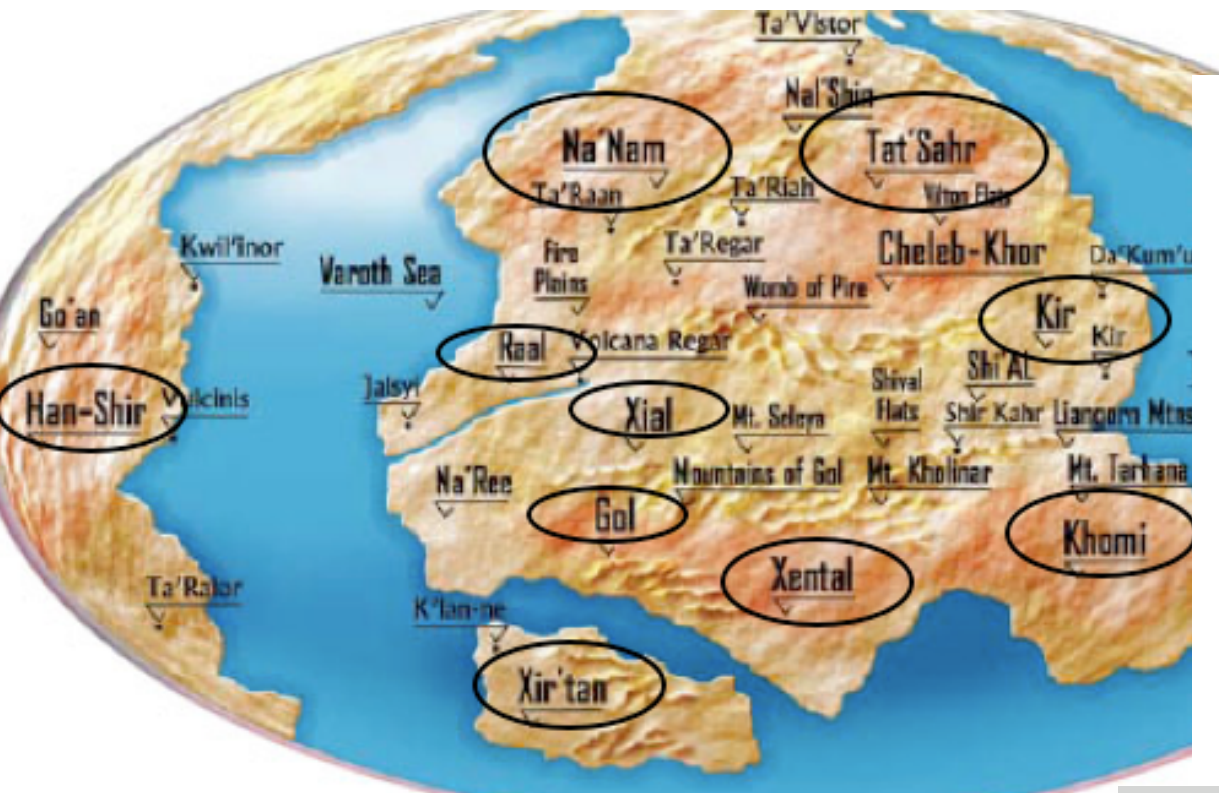


Geographic Isolation

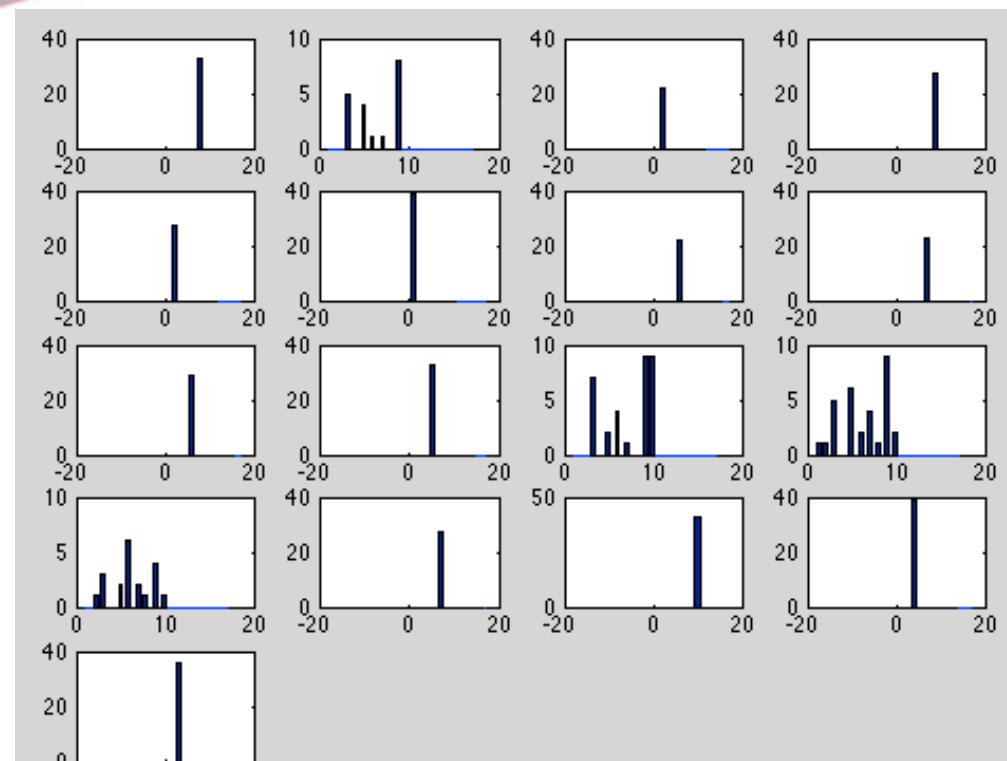


Vulcan Economic Hubs

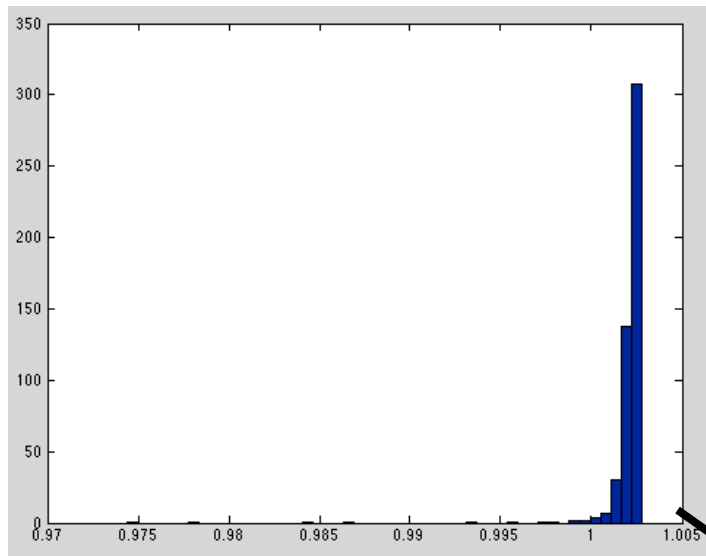
“Seems logical to me captain”



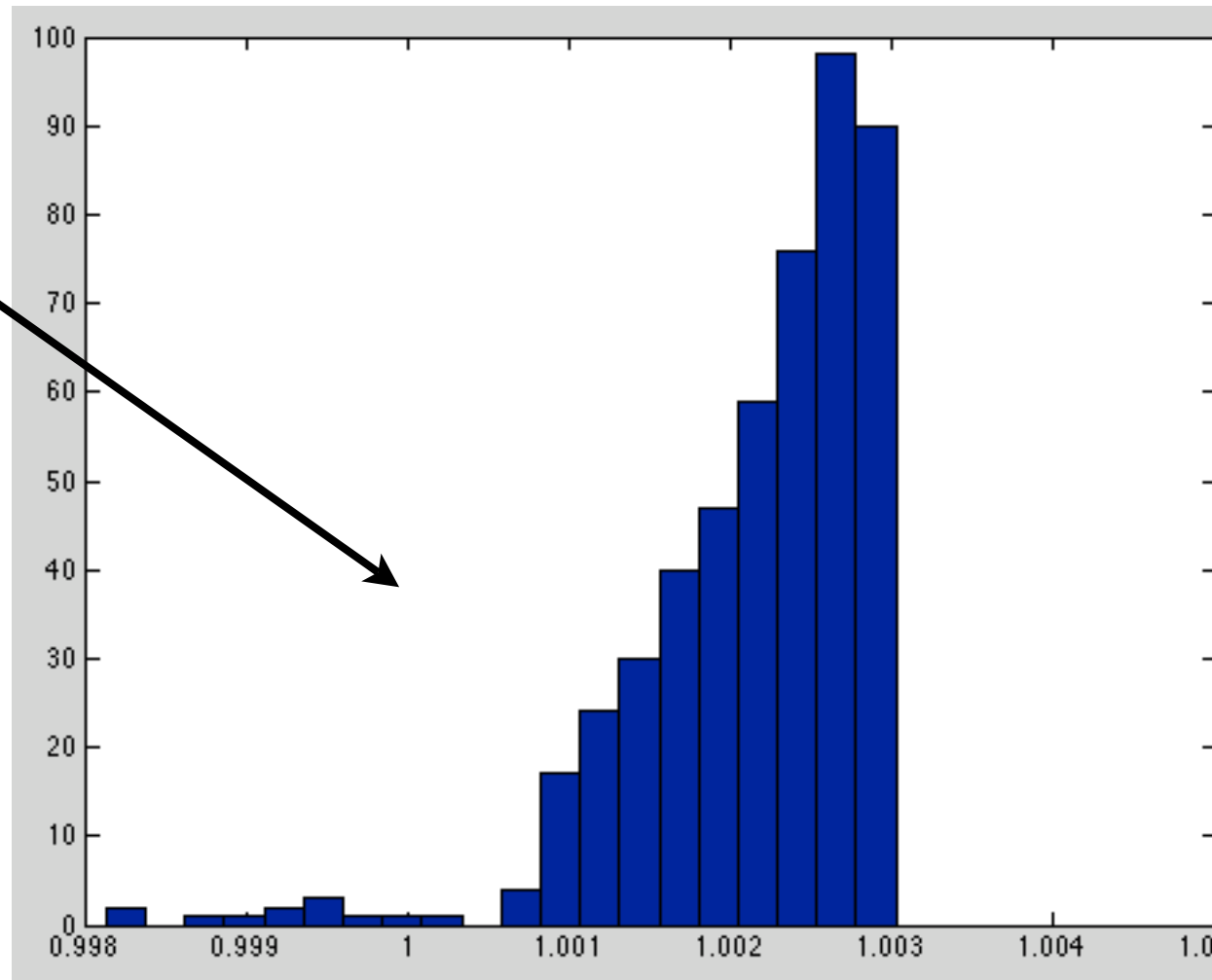
Cluster Break Down By Region



Remove Regions



What remains?

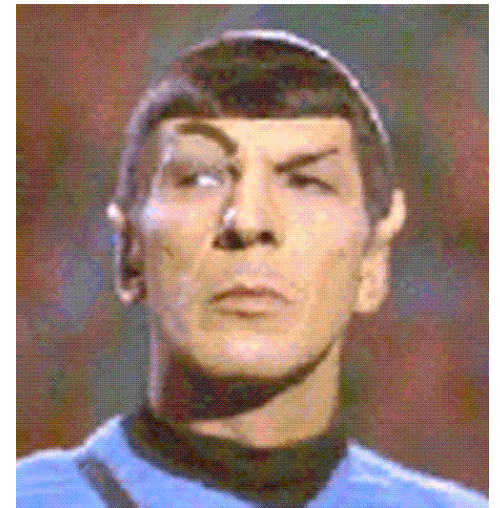


Fun Project!

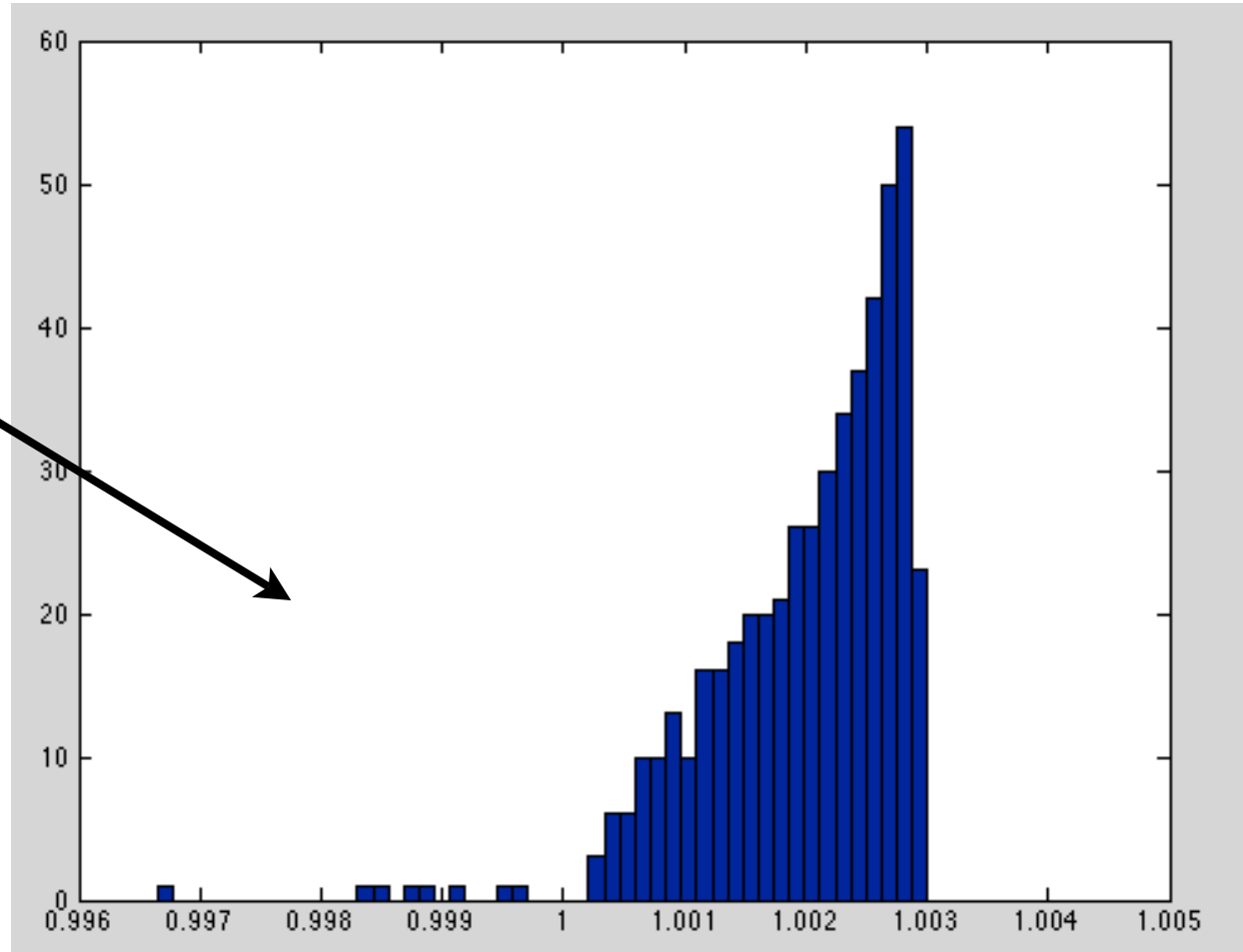
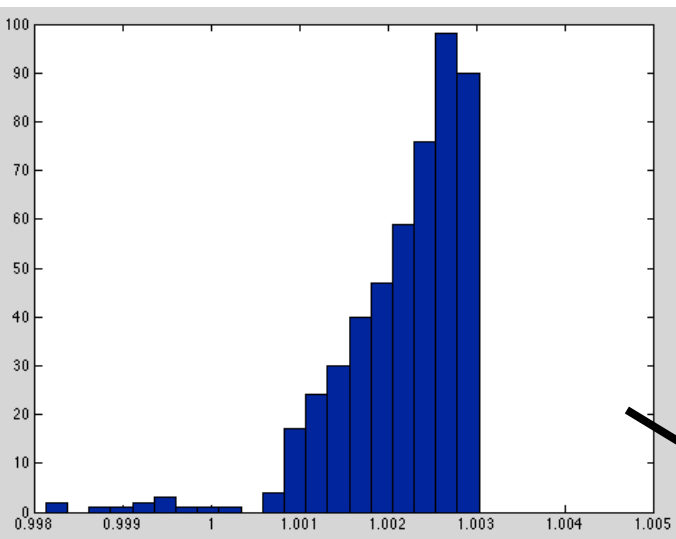
- Explain the next scale's “partition”.
- Devise ways to study discretely over lapping clusters.

(Taken from CSSS08 Lectures)

Partition Decoupling Method (Basically) Cluster, Scub, Iterate.



What were the
Clusters?



Vulcan Sector Rotation

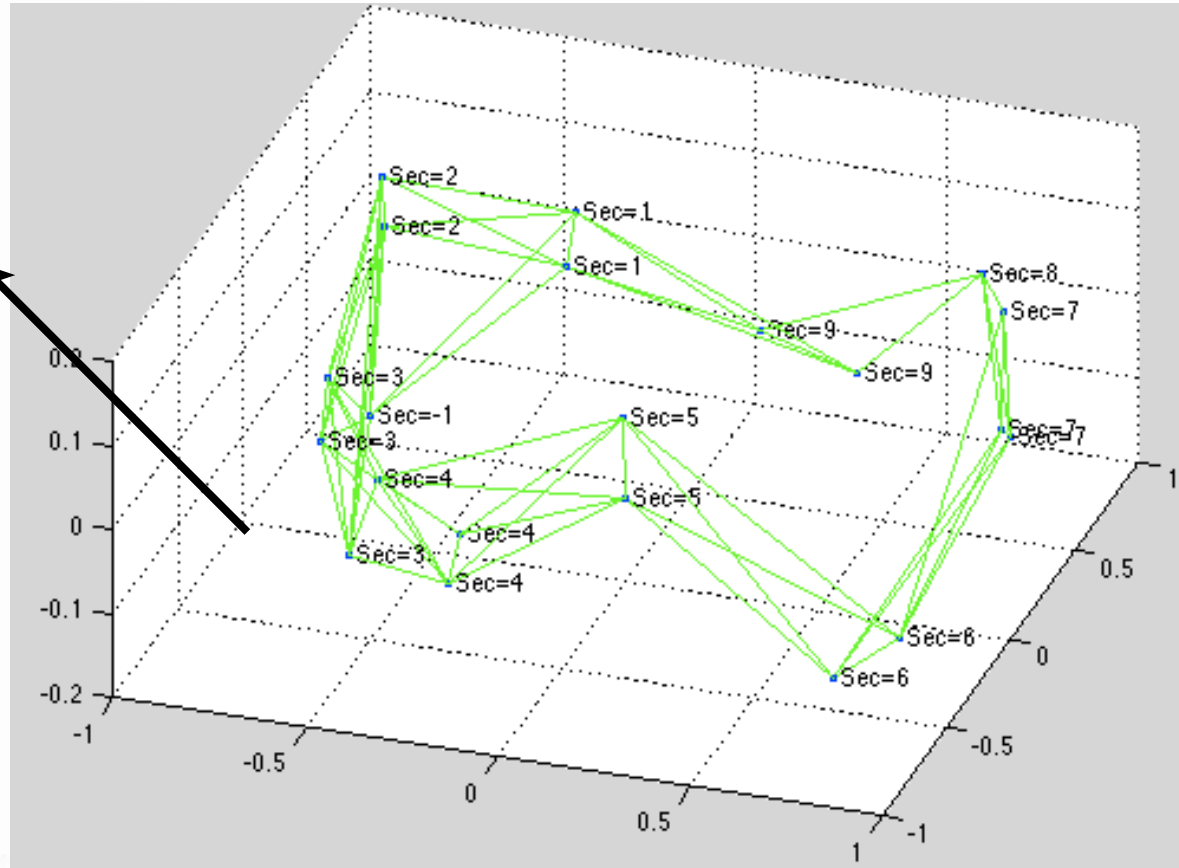
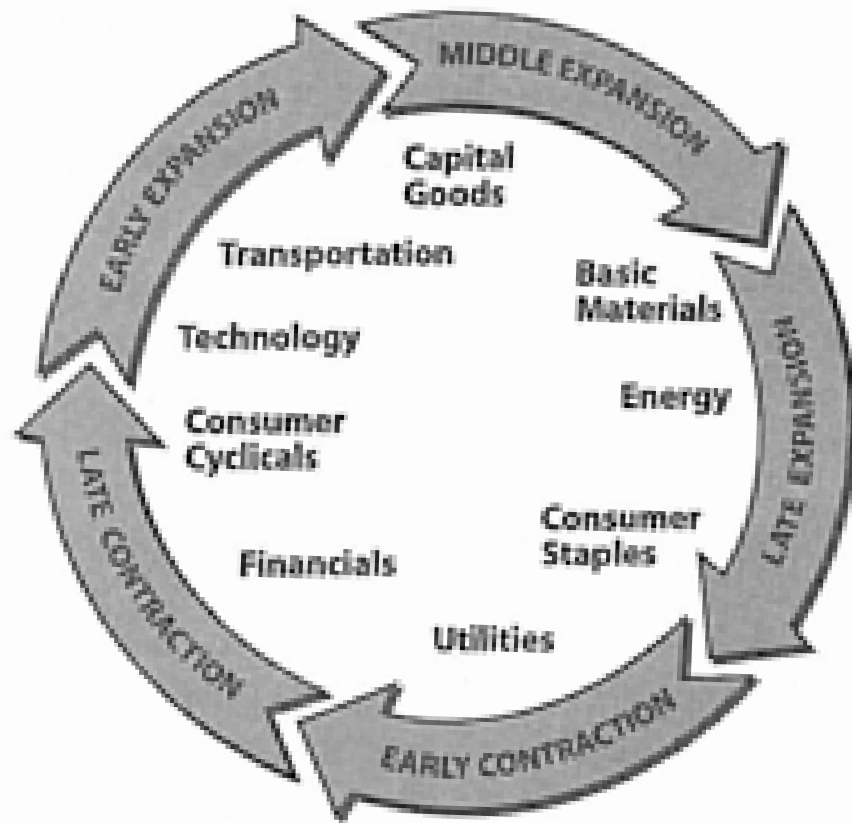
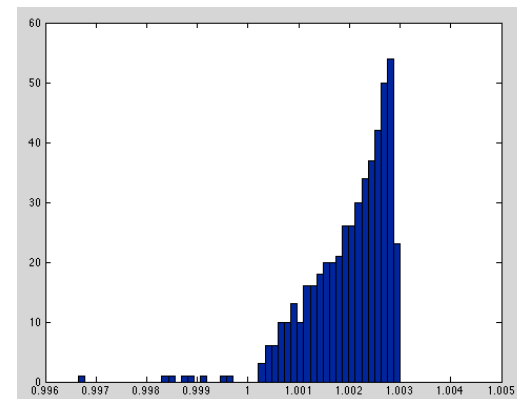
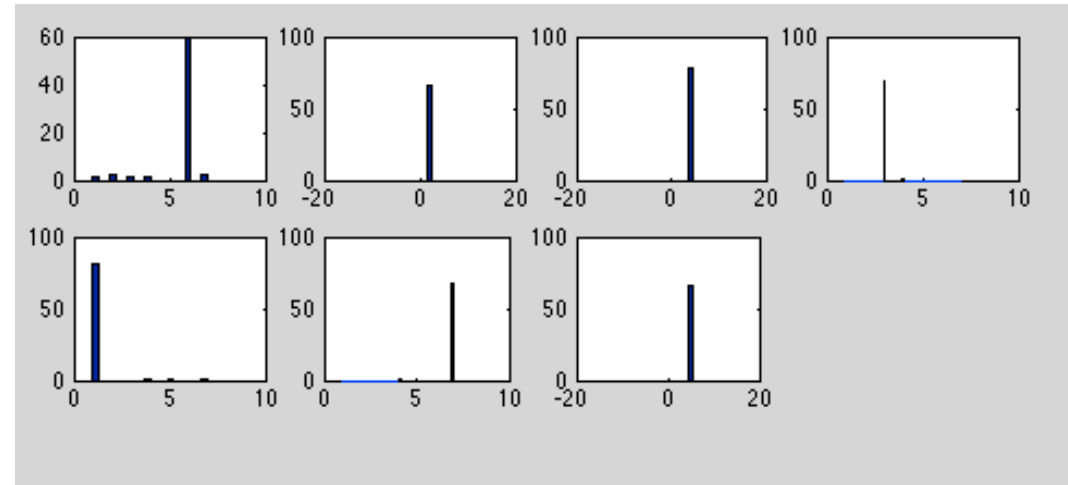


FIGURE 13.1 Technology and transportation leadership during 2003 fits Early Expansion phase.

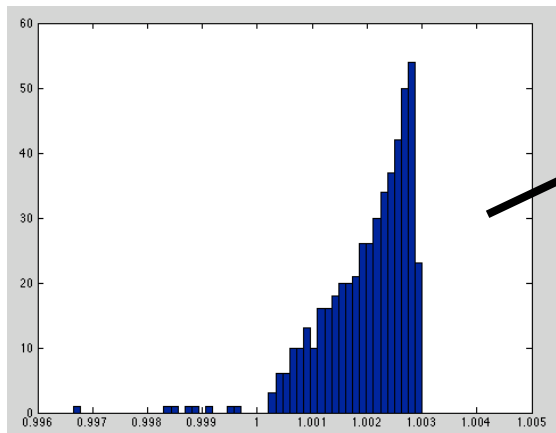
What could be left?



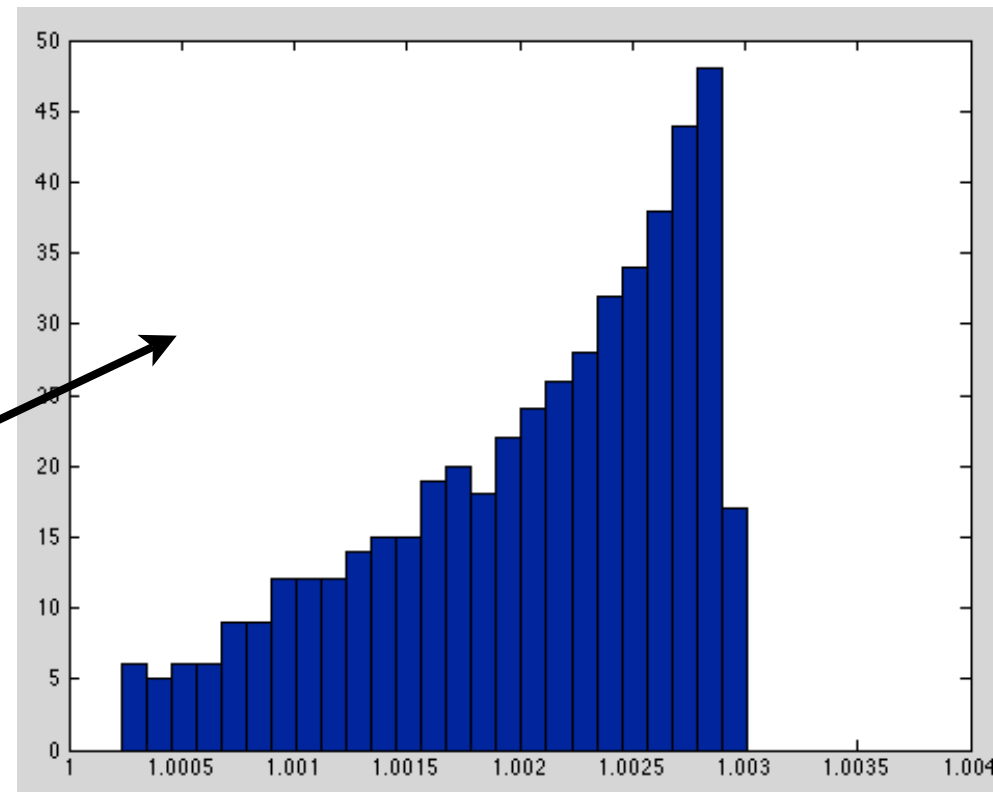
Have we found the null model yet?



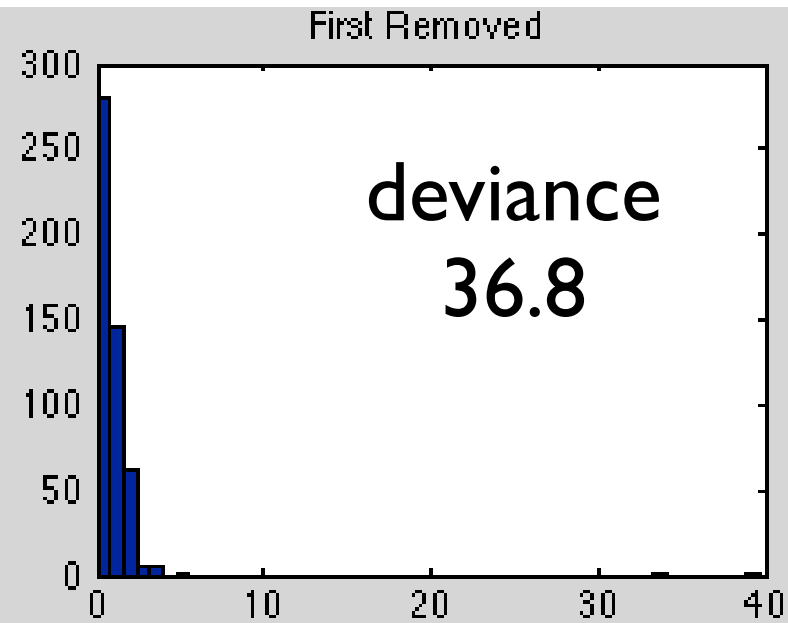
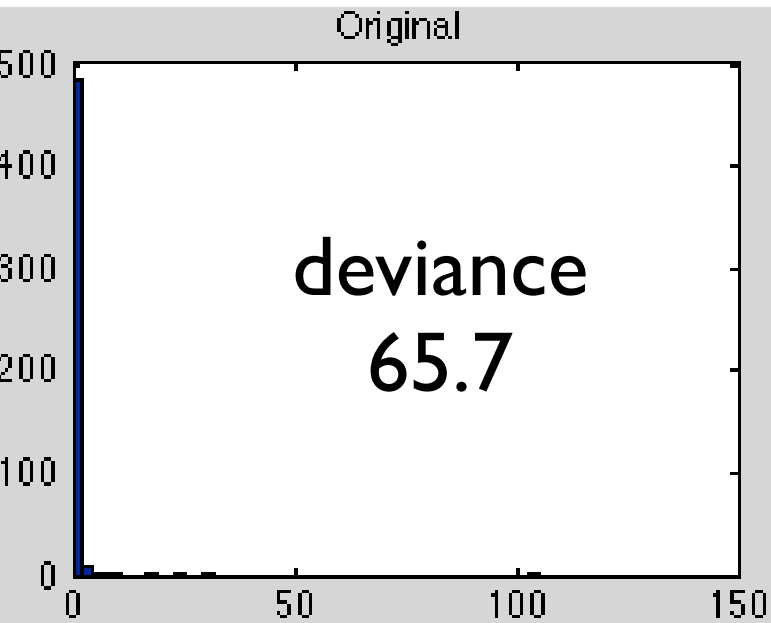
So it would seem.



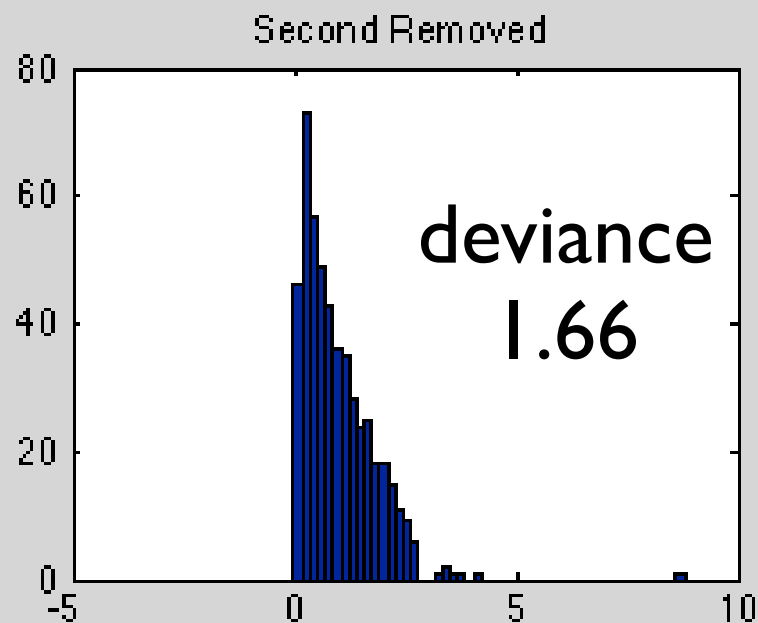
Scrub 7
Vulcan
Clans



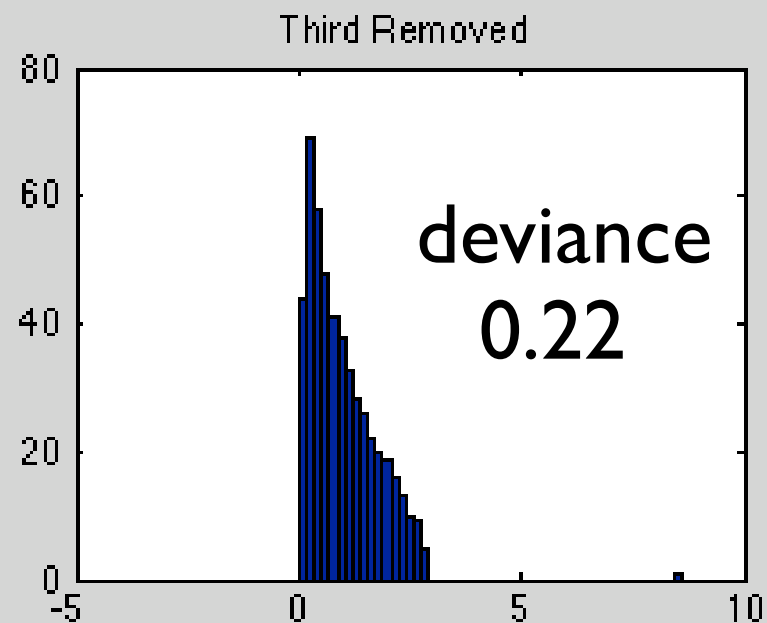
Scrubbing



scrub 10



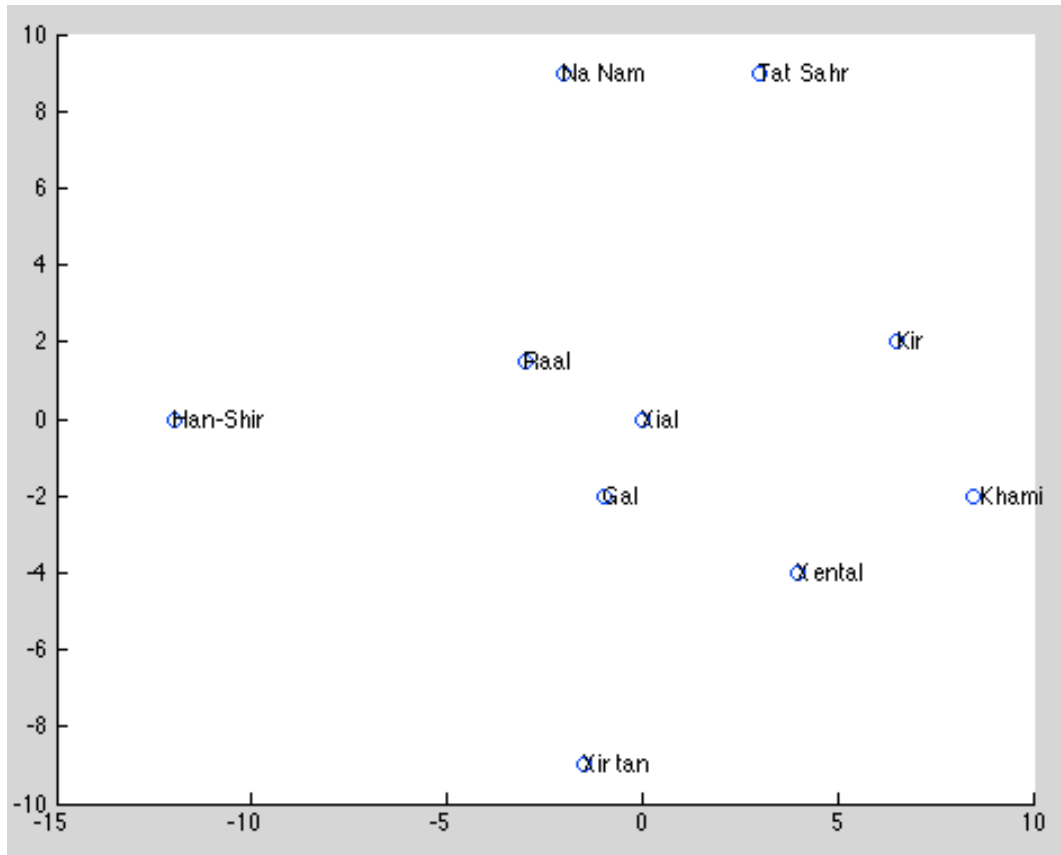
scrub 9



scrub 7

Making the simulation...

Names={'Unknown','Unknown','Han-Shir','Kir','Khami','Tat Sahr','Xental','Na Nam','Xial','Raal','Gal','Xir tan'};



P =
[-12.0 0
6.5 2.0
8.5 -2.0
3.0 9.0
4.0 -4.0
-2.0 9.0
0 0
-3.0 1.5
-1.0 -2.0
-1.5 -9.0]

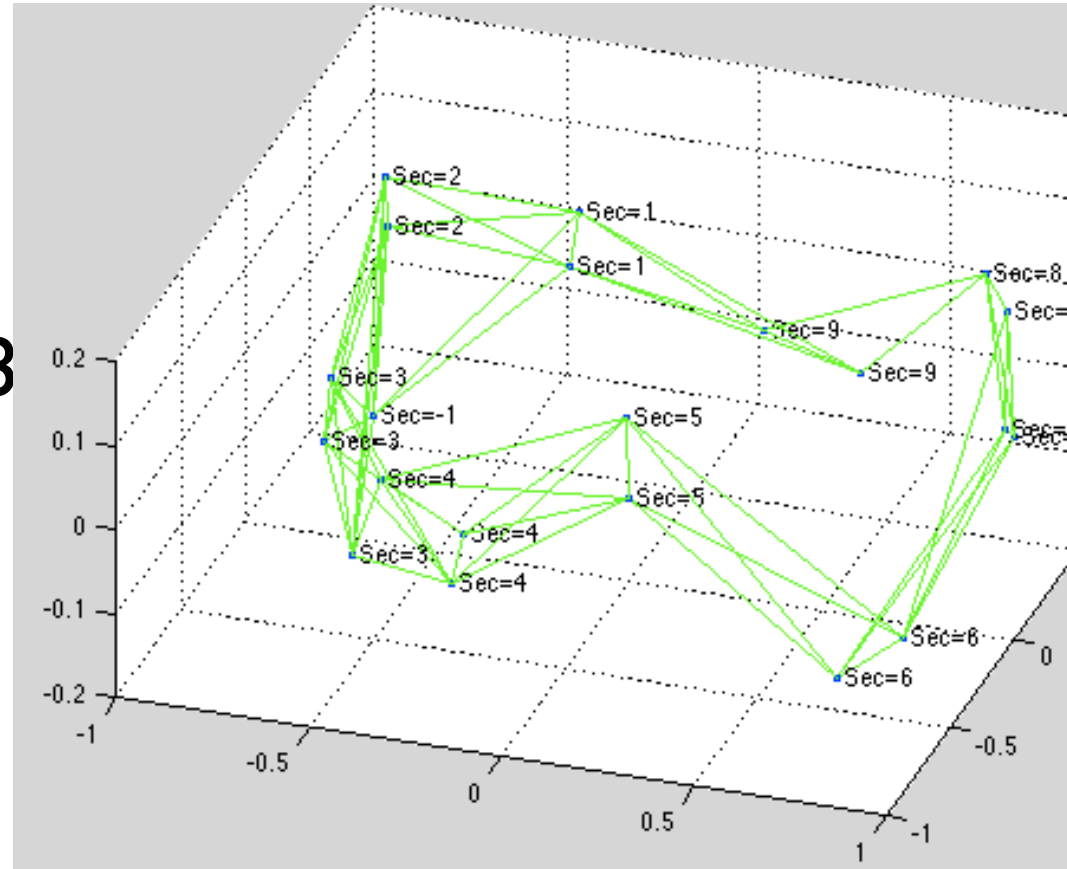
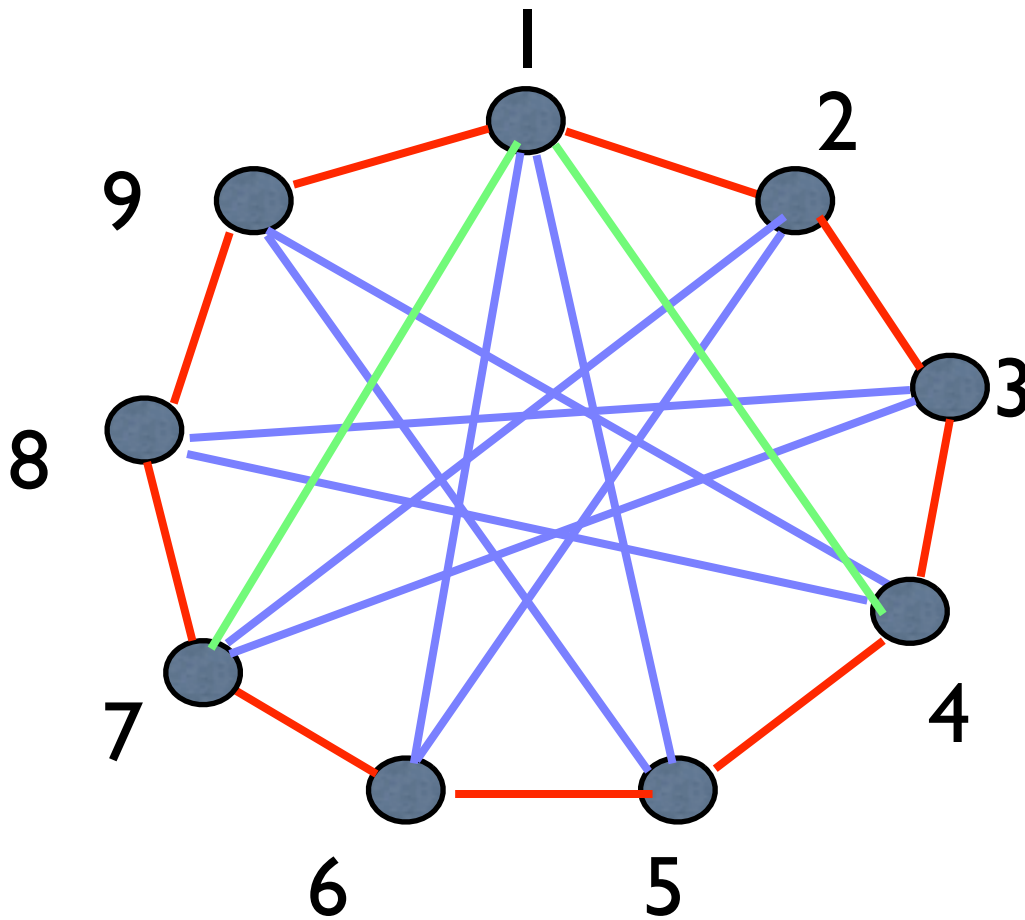
Fun
Fact!

```
SQ=squareform(pdist(P));
COR=(1-SQ/C).^2
[O E]=eig(COR);
%COR=O*E*O'
V=O*sqrt(E')*O';
%V*V'=COR
```

```
G=[];
for i=1:(length(Names)-2)
    Fake= random('Normal',0,1,d,1);
    Fake=Fake-mean(Fake);
    G=[G Fake];
end
```

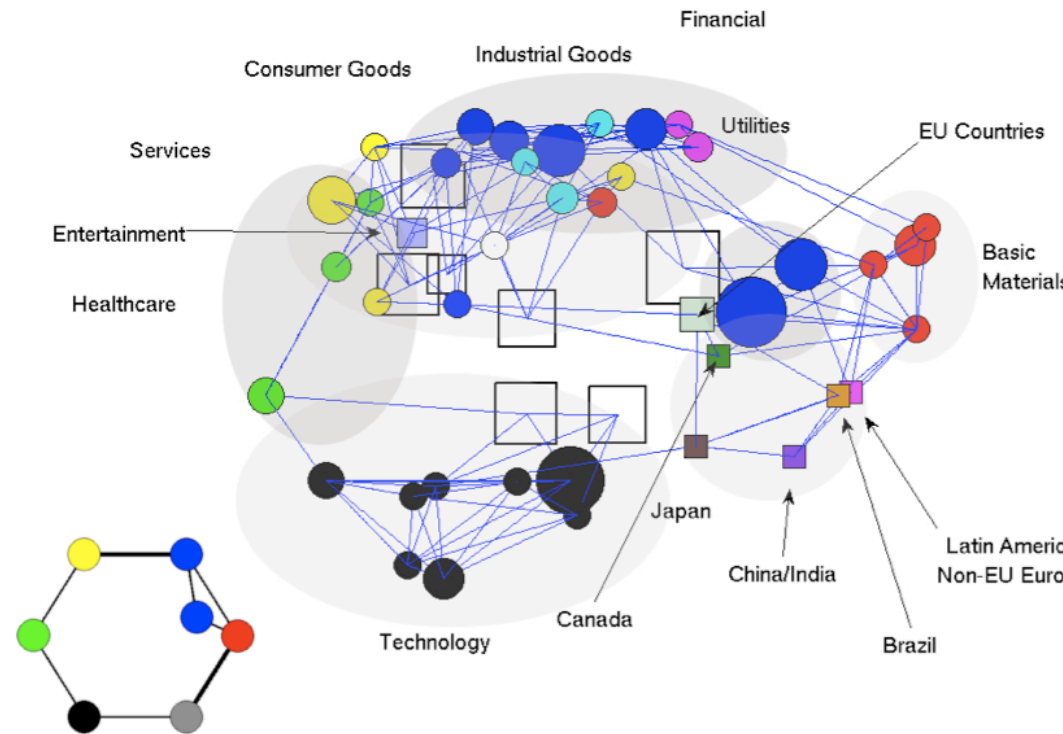
```
R=[];
for i=1:(length(Names)-2)
    r=0;
    for j=1:(length(Names)-2)
        r=r+V(i,j)*G(:,j);
    end
    R=[R r];
end
```

Simulating Sector

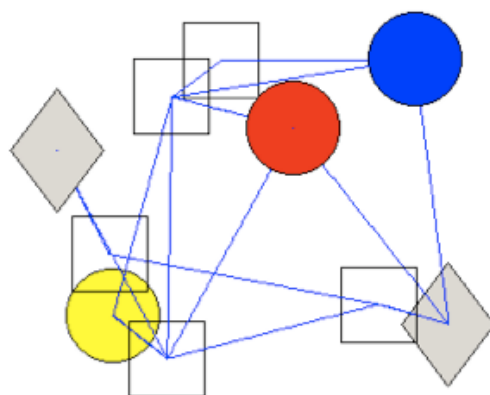


Religions we random and independent.

The Earth Market!



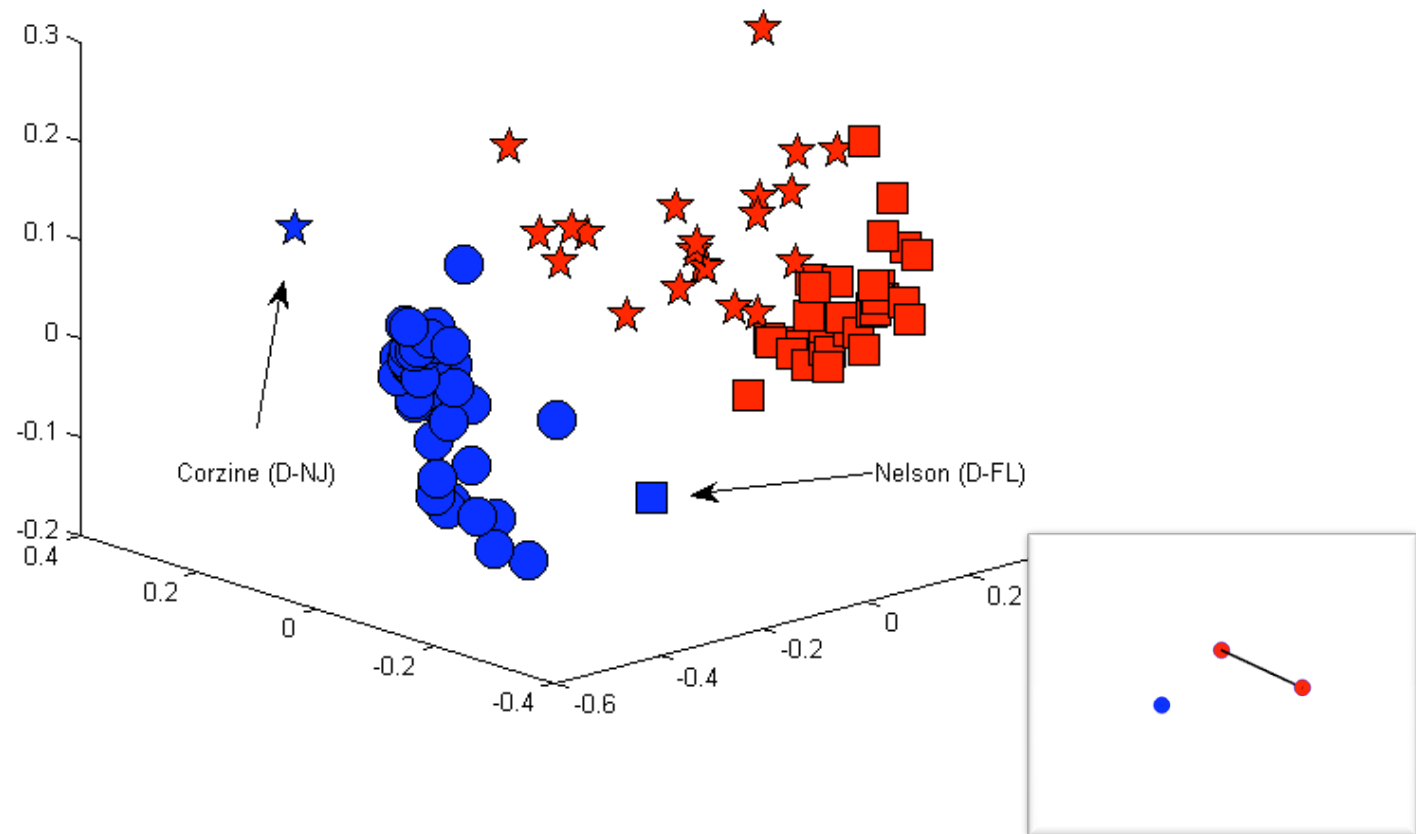
actions. For example, the diamond shaped clusters contain a mixture of multiple sectors. The first is predominantly Consumer Goods, Industrial Goods and Services, while the second is predominantly Financial, Healthcare, Services and Technology. However, both clusters contain significant commonalities. In the first, the equities in the Service sector are almost all related to the shipping industry, which obviously serves to distribute Consumer and Industrial Goods. In the second, the equities in the Financial, Services and Technology sectors are related to companies that either provide services or do business with healthcare companies (e.g. health insurance companies, drug companies, management services, healthcare based REITs, etc.). Equities in both of these clusters are drawn



109th Congress

Partition 1

Initial Clustering of 109th Senate



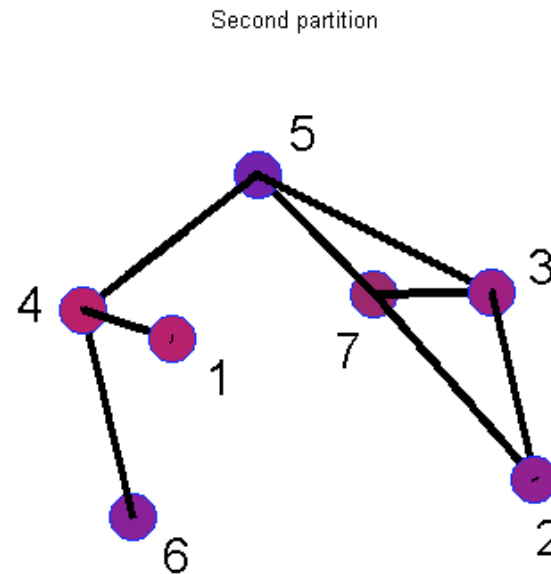
Partition 2

- Clean initial clusters via projection and renormalization of the residual data series
- Construct dissimilarity matrix using Euclidean distance
- Use spectral clustering: too much residual structure to use a random model, so use multidimensional scaling of the dissimilarity matrix and “elbowology” – 7 clusters

Partition 2

- Party effects removed, the color of the nodes is a proportional mixture of blue (Democrats) and red (Republicans)
- Use a measure of information strength, we isolate the votes which have the most impact on the formation of each cluster
- These votes indicate dominant issues resulting in cluster formation

1. Restraint on spending
2. Pro-business, “small government”
3. Immigration/Homeland Security
4. Foreign Policy
5. Poison Pills
6. Energy
7. Immigration



Partition 2

Cluster 6: Energy

1. H.R. 6, Energy Independence and Security Act of 2007, Amendment 783, To strike the section providing for a comprehensive inventory of Outer Continental Shelf oil and natural gas resources
2. S. 2320, A bill to make available funds included in the Deficit Reduction Act of 2005 for the Low Income Home Energy Assistance Program for fiscal year 2006 (waive CBA motion and cloture)
3. S. 3711, cloture, Gulf of Mexico Energy Security Act