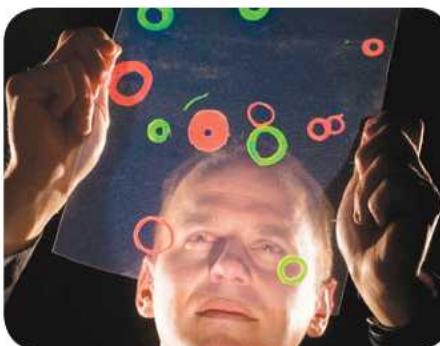


Exceptional service in the national interest



Development and Exploration of Cultural Archetypes Using JMP

Presented by Thor D. Osborn

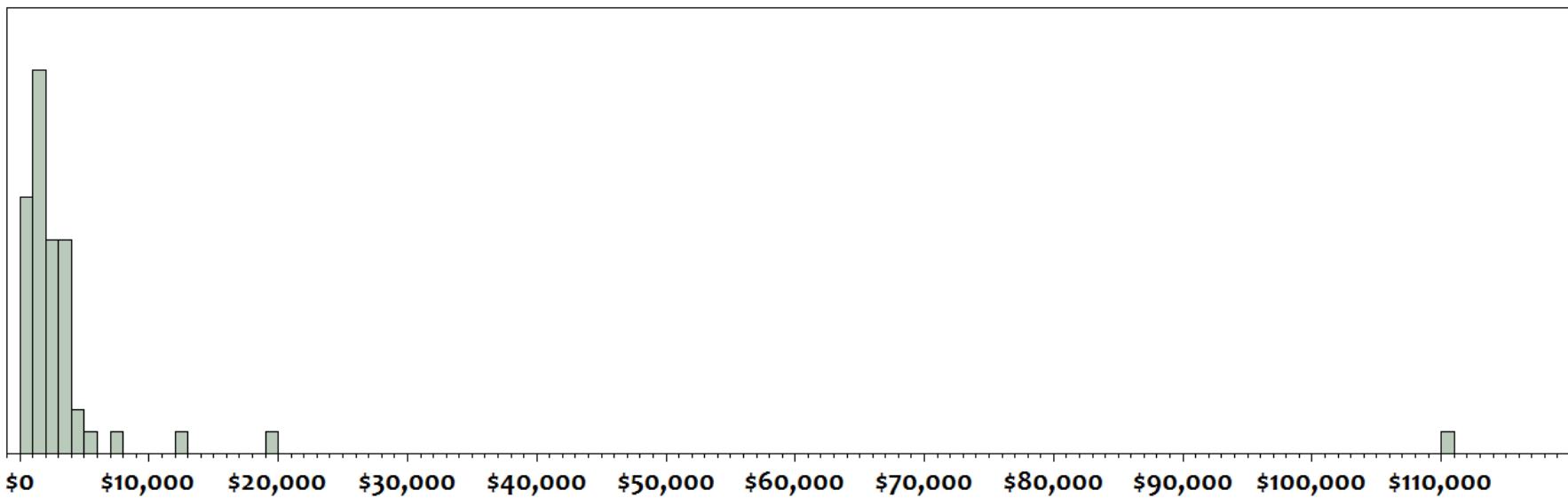
September 12, 2012

Overview

- ▶ Societies are often said to have cultural similarities and differences
 - Can cultural archetypes having distinct tendencies be systematically identified?
 - How can archetype membership be determined?
 - ◆ Without inserting substantial personal bias?
 - ◆ Without a Ph.D. in anthropology?
 - JMP enables exploration
 - ◆ Fluid extraction of results from one platform as inputs to the next
 - ◆ Graphical design of equations for data pre-processing (similar to Mathcad)
 - ◆ Interactive graphical display of results for ease of presentation
- ▶ Onward – exploring inward-flowing greenfield FDI

The Outcome Data

- ▶ Foreign Direct Investment (FDI) is important to economic development
- ▶ Inward-flowing “greenfield” FDI (IGFDI) is especially important to less developed economies
- ▶ There are vast differences between winners and losers – why?



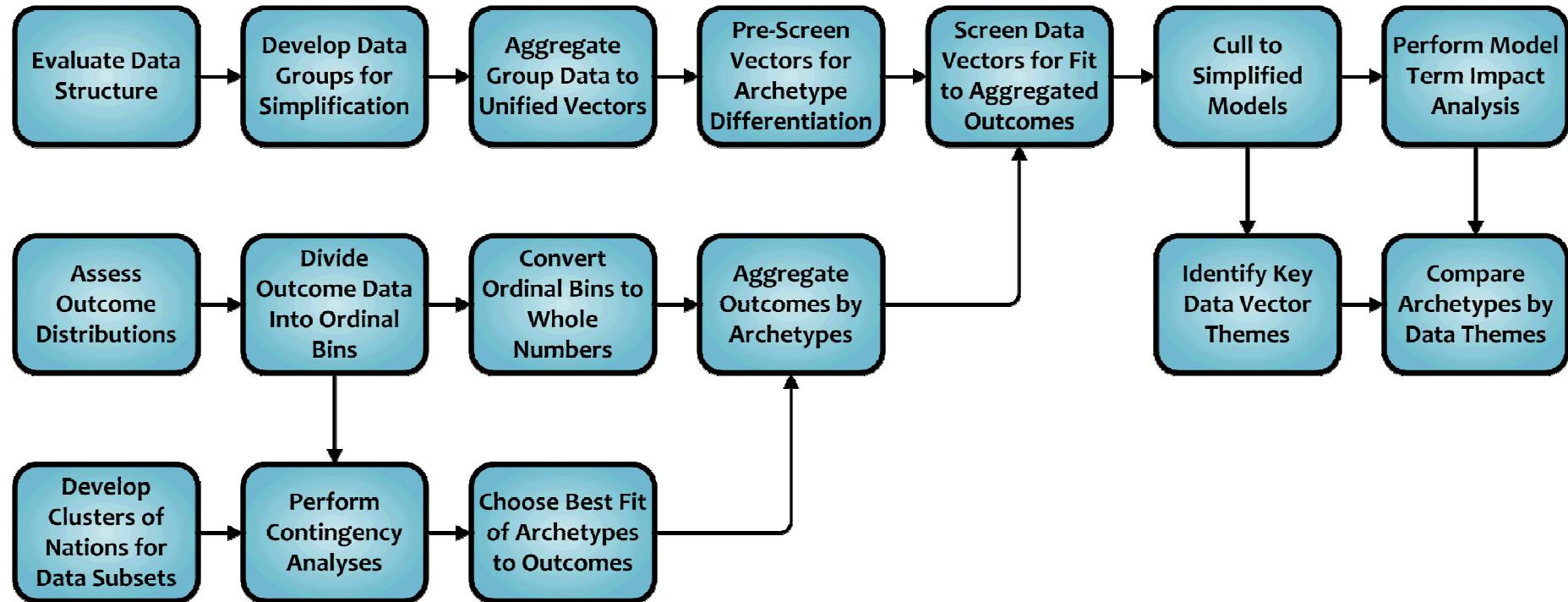
The Comparison Data

- ▶ All of the data used in this analysis are from published materials – most can be readily downloaded from online sources
- ▶ GLOBE study of leadership traits and expectations around the world – 57 societies used in this analysis
- ▶ Societal traits from Culturally Endorsed Implicit Leadership Theory
- ▶ Socioeconomic data and characterizations from a broad range of sources
 - ▶ World Bank
 - ▶ World Economic Forum
 - ▶ Heritage Foundation
 - ▶ Transparency International
 - ▶ Freedom House
 - ▶ Vision for Humanity
 - ▶ International Food Policy Research Institute
 - ▶ United Nations

The data chosen to represent the societies necessarily impacts the contextually-based societal archetype structure:

different data → different results

The Model Development Process

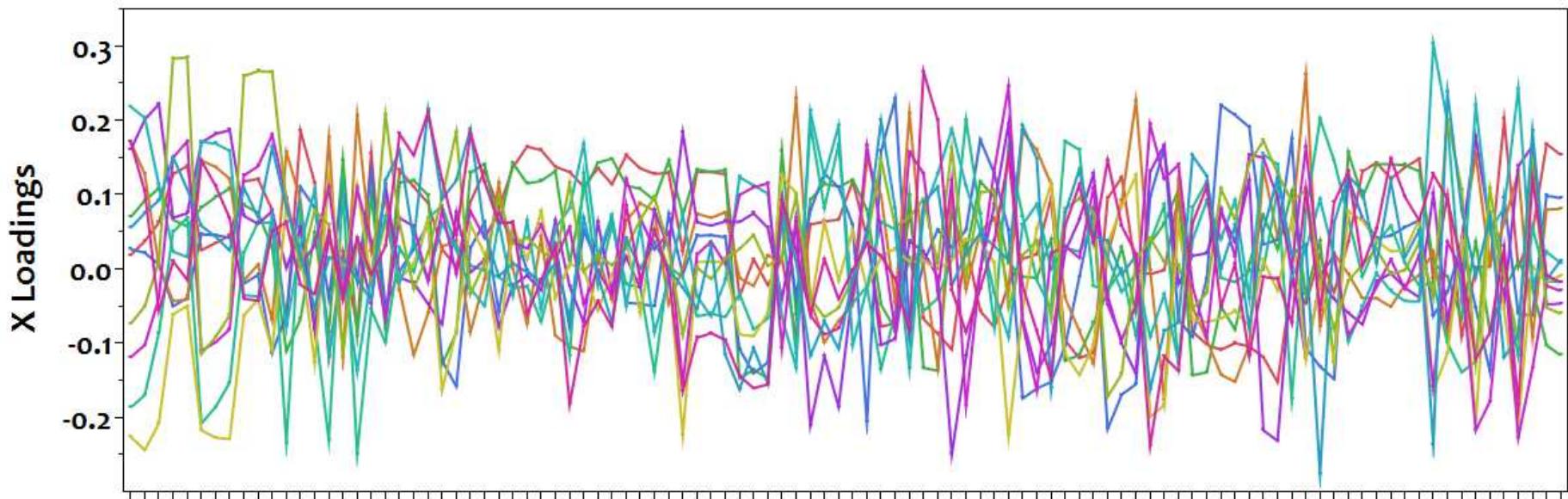


**Data consolidation and pattern-matching
for intuitive comprehension and hypothesis
generation – *not* hypothesis testing**

One Alternative - PLS

► Loadings plot

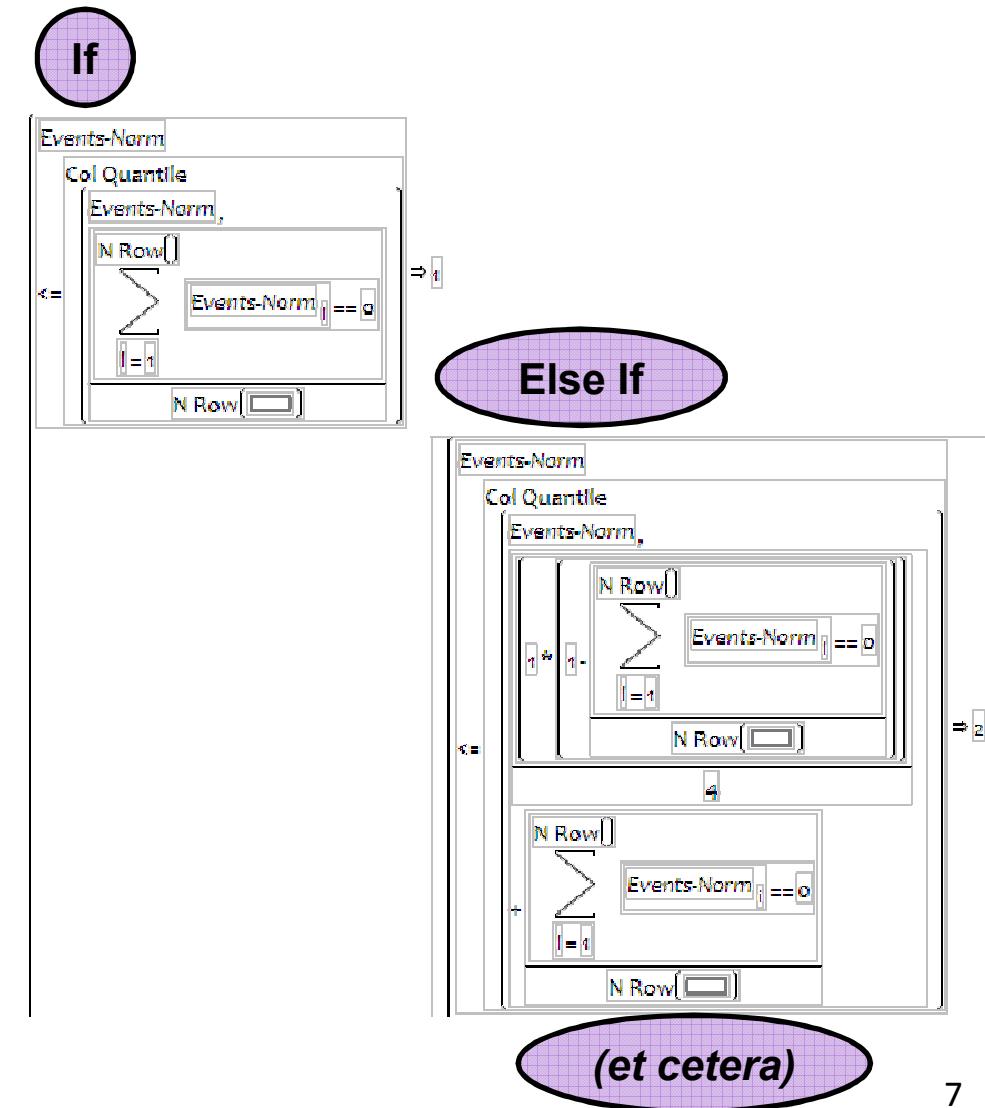
- 12 latent variables, each is composite of original 90 inputs (not shown)
- Model explains 99.3% of variation in outputs



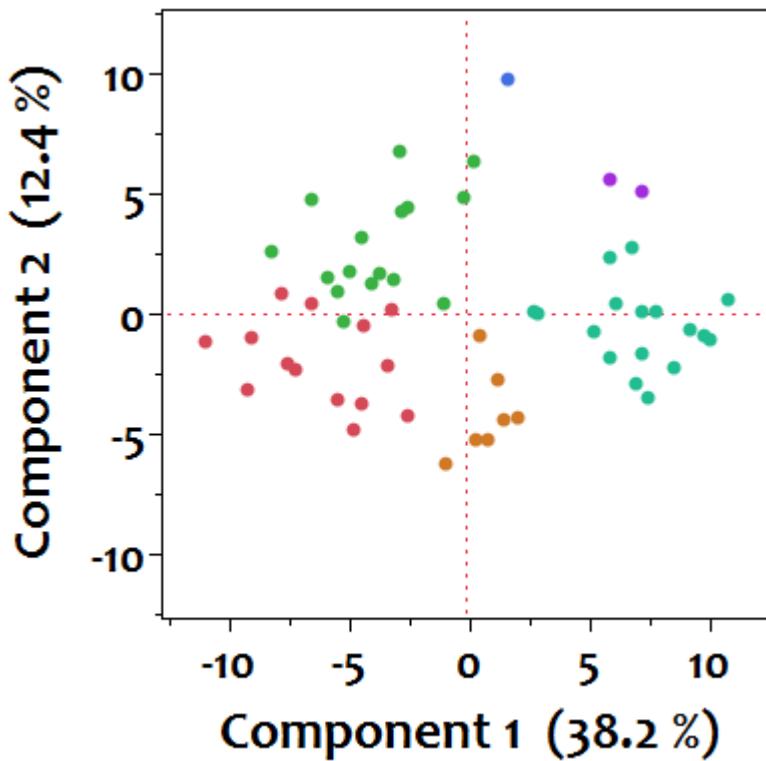
Mathematically, PLS can “explain” nearly all of the variation – intuitively, somewhat less

Convert Outcome Data to Ordinals

- ▶ Generate new column and use Formula Editor
- ▶ Note traditional graphic mathematical notation
- ▶ Regular quantiles generated for IGFDI data, but JMP enables alternatives
 - ▶ The formula at right creates modified “quintiles” that place all zeroes in the first level instead of arbitrarily assigning equal values to separate levels



Develop Clusters of Societies – 1/2

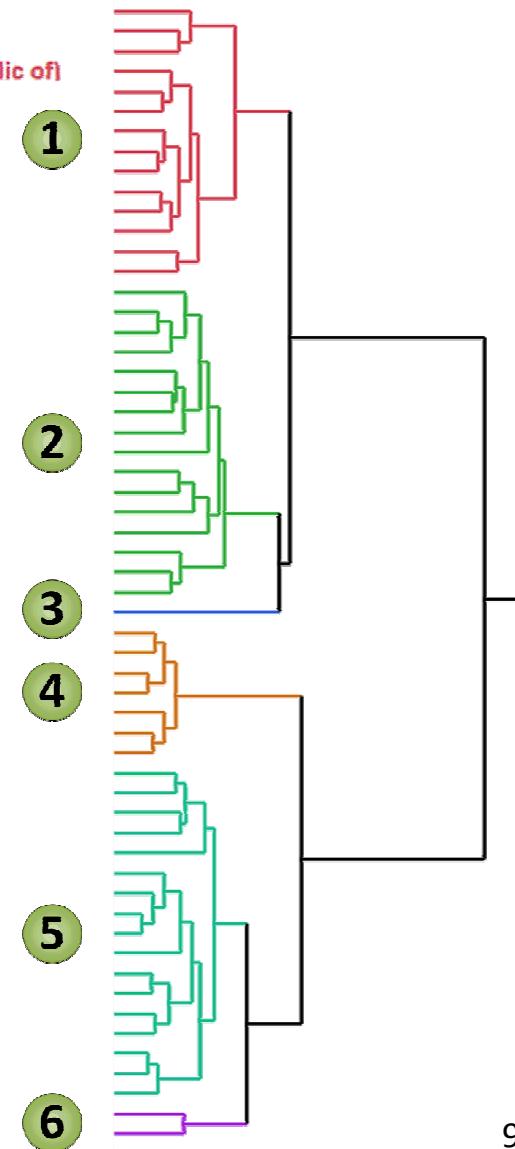


- ▶ Objective is to cluster societies by co-similarity
- ▶ Use principal components analysis to find first principal component of data subset
- ▶ Analyze → Multivariate Methods → Principal Components
 - ▶ Select all input data columns, hit “OK”
 - ▶ Save first principal component to data table

Develop Clusters of Societies – 2/2

- ▶ Analyze → Multivariate Methods → Cluster
 - Select all data in subset for Ys
 - Select Country for Label
 - Select Prin 1 for Order
 - Hit “OK”
 - Red triangle menu
 - Set number of clusters, N
 - Save clusters to data table
 - Repeat for all N
- ▶ Post-processing
 - Image extracted using SnagIt
 - Identifiers superimposed using MS Visio
 - Cluster labeling: future enhancement for JMP?

-Zimbabwe
-Nigeria
-Zambia
-Venezuela (Bolivarian Republic of)
-Ecuador
-Argentina
-Guatemala
-El Salvador
-Colombia
-Turkey
-Mexico
-Brazil
-Bolivia
-Namibia
-Iran (Islamic Republic of)
-Philippines
-Indonesia
-India
-Thailand
-Costa Rica
-Malaysia
-Kuwait
-China
-Egypt
-Albania
-South Africa
-Morocco
-Russian Federation
-Georgia
-Kazakhstan
-Qatar
-Greece
-Hungary
-Italy
-Spain
-Poland
-Slovenia
-Portugal
-Korea (Republic of)
-Israel
-France
-Japan
-United States
-Ireland
-United Kingdom
-Australia
-Canada
-New Zealand
-Finland
-Sweden
-Netherlands
-Denmark
-Austria
-Germany
-Switzerland
-Hong Kong, China (SAR)
-Singapore



Perform Contingency Analysis

- ▶ Analyze → Fit Y by X
 - Select cluster columns for X
 - Select normalized quantile columns for Y
 - JMP automatically engages the Contingency “personality” of the Fit Y by X platform because X is categorical and Y is ordinal
 - Hit “OK”
 - Statistical test outcomes appear below mosaic plot
 - Save output as text file for easy searching

Cast Selected Columns into Roles

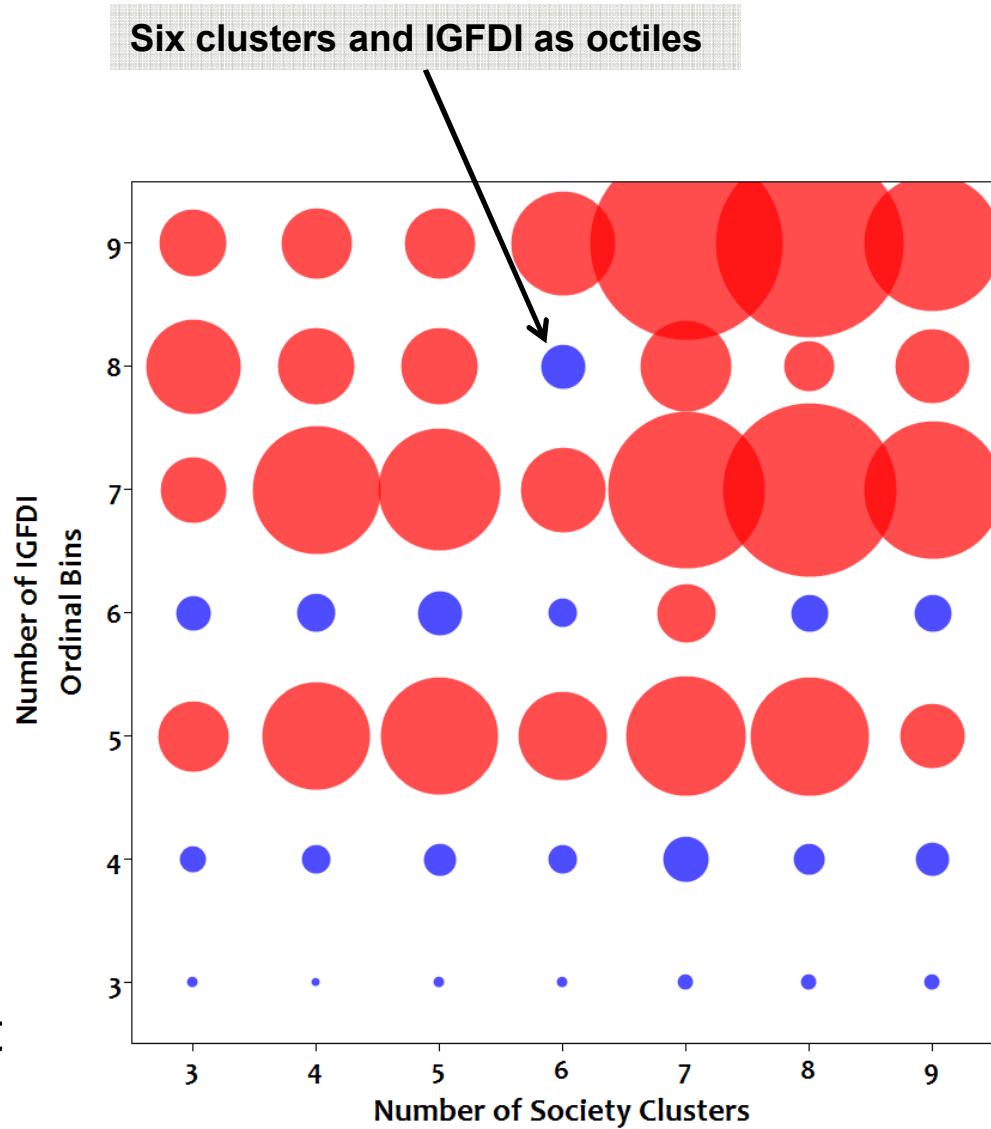
Y, Response	<ul style="list-style-type: none"> FDI Inward Greenfield Per Capita Tertile FDI Inward Greenfield Per Capita Quartile FDI Inward Greenfield Per Capita Quintile FDI Inward Greenfield Per Capita Sextile FDI Inward Greenfield Per Capita Septile FDI Inward Greenfield Per Capita Octile FDI Inward Greenfield Per Capita Nonile
X, Factor	<ul style="list-style-type: none"> Cluster C3 Cluster C4 Cluster C5 Cluster C6 Cluster C7 Cluster C8 Cluster C9
Block	optional
Weight	optional numeric
Freq	optional numeric
By	optional

Action

OK
Cancel
Remove
Recall
Help

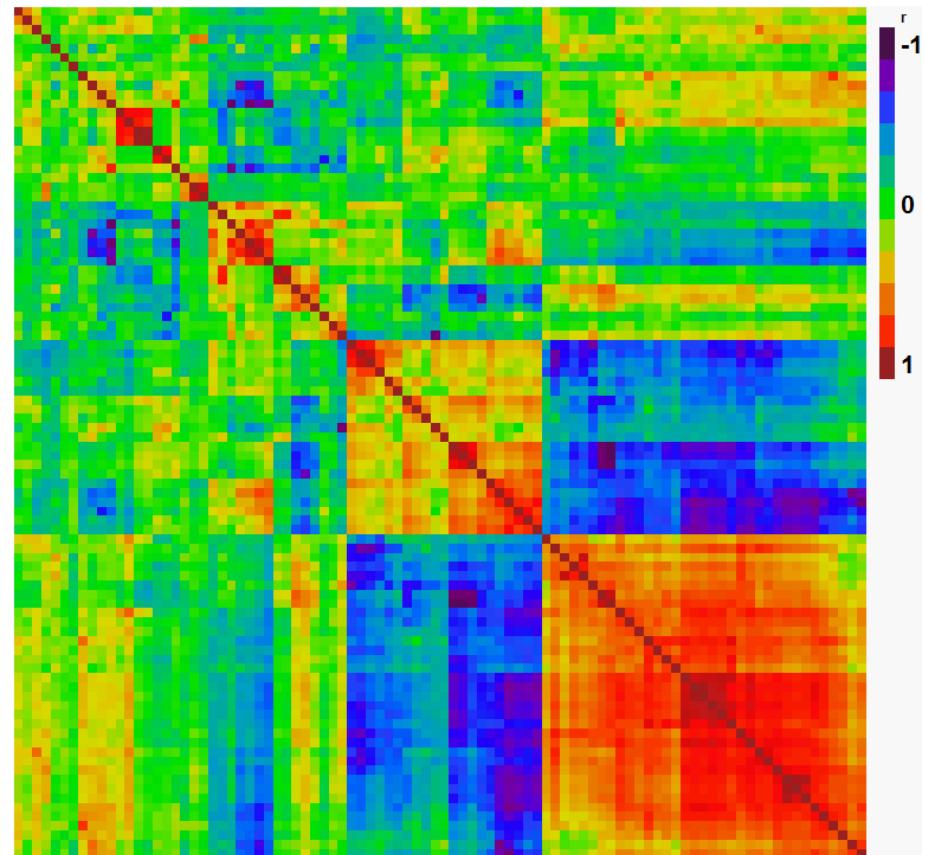
Identify Best Match

- ▶ Use table of contingency analysis results
 - Pearson Chi-Square Probabilities
 - Use Formula Editor to make test column – check limit and whether Chi-Square suspect
- ▶ Graph → Bubble Plot
 - Ordinal bins as Y
 - Cluster size for X
 - Pearson Chi-Square probability for Data
 - Color by test column
- ▶ Use judgment to find most suitable combination



Examine Data Structure

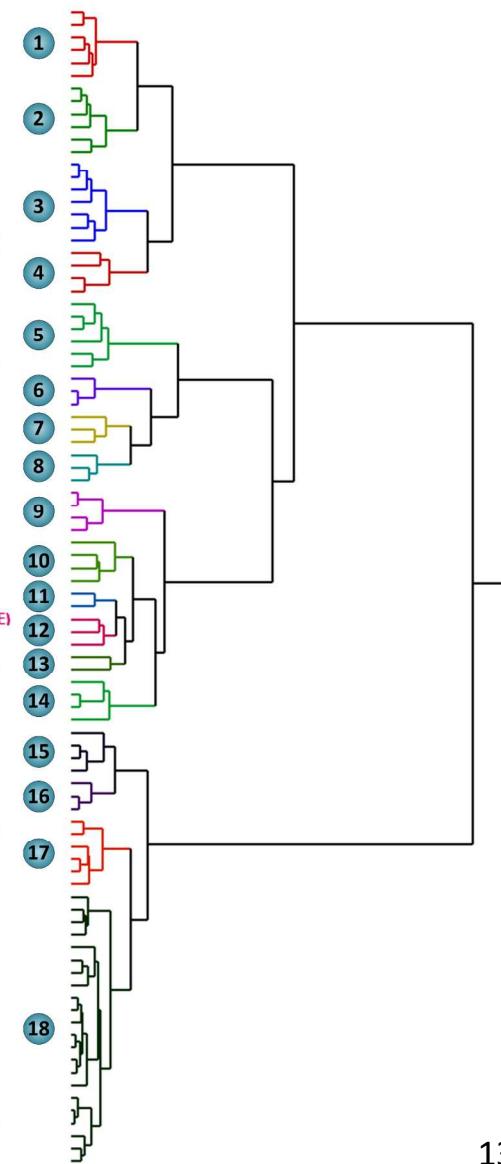
- ▶ Use full data table
- ▶ Analysis → Multivariate Methods → Multivariate
 - ▶ All input data for Ys
 - ▶ Hit “OK”
 - ▶ Red triangle menu: Color Maps → Cluster the Correlations
 - ▶ Visual indication that groupings of similar data exist in data set
 - ▶ Right-click over correlation table to save as a new table



Simplify Using Data Groups

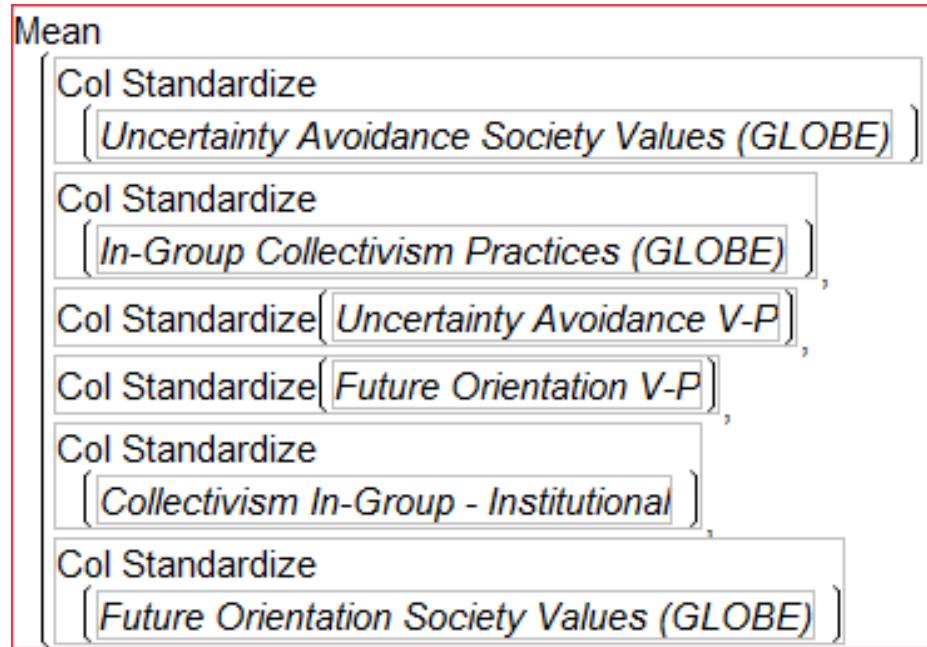
- ▶ Use table generated from correlation matrix
- ▶ Perform PCA to obtain first principal component
- ▶ Cluster as before using Ward's method and Prin1
- ▶ Choose N=18 as a compromise between simplification and granularity
- ▶ Save Clusters to data table

- Uncertainty Avoidance Society Values (GLOBE)
- In-Group Collectivism Practices (GLOBE)
- Uncertainty Avoidance V.P
- Future Orientation V.P
- Collectivism In-Group - Institutional
- Future Orientation Society Values (GLOBE)
- Child Dependency Ratio 2010
- Total Fertility Rate 2005-2010
- Rate of Population Natural Increase in Pct
- Gini Household Index Score
- Composite Gini Index
- Homicides Per 100,000
- FH Press Freedom Rating
- FH Civil Liberties Rating
- FH Political Rights Rating
- CLT - Self-Protective
- Fiscal Freedom Index
- Government Spending Index
- Energy Consumption Rate of Increase 05 to 09 (%)
- CLT - Humanes Orientation
- Humanes Orientation Society Practices (GLOBE)
- Assertiveness Society Values (GLOBE)
- Assertiveness V.P
- Power Distance Society Practices (GLOBE)
- Inconsistency V.P (GLOBE)
- Inconsistency (GLOBE)
- Inconsistency P.V (GLOBE)
- Performance Orientation V.P
- Performance Orientation Society Values (GLOBE)
- CLT - Team Oriented
- CLT - Charismatic or Value Based
- Visionary Leadership Society Scores (GLOBE)
- Assertiveness Society Practices (GLOBE)
- Humanes Orientation Society Values (GLOBE)
- Humanes Orientation V.P
- Gender Egalitarianism V.P
- Gender Egalitarianism Society Values (GLOBE)
- CLT - Participative
- Net Energy Surplus Per Capita 2009 (MBTU)
- Energy Production Per Capita 2009 (MBTU)
- Net International Migration Rate
- Energy Consumption Per Capita 2009 (MBTU)
- Energy Production Rate of Increase 05 to 09 (%)
- Female Prisoners (%)
- Population Density
- WEF Macroeconomic Stability (Bas)
- Power Distance Society Values (GLOBE)
- Power Distance V.P
- Performance Orientation Society Practices (GLOBE)
- Institutional Collectivism Practices (GLOBE)
- Labor Freedom Index
- CLT - Autonomous
- Gender Egalitarianism Society Practices (GLOBE)
- Prison Pop Per 100K
- Primary Energy Production 2009 (QBTU)
- Primary Energy Consumption 2009 (QBTU)
- WEF Market Size (Eff)
- Urban Population
- Adult Literacy Rate
- Education Index
- Combined Gross Enrolment Ratio
- Old Age Dependency Ratio 2010
- VH Democracy Index
- WBGI - Voice and Accountability
- Future Orientation Society Practices (GLOBE)
- Uncertainty Avoidance Society Practices (GLOBE)
- WEF Group Innovation and Sophistication (Eff)
- WEF Goods Market Efficiency (Eff)
- WEF Institutions (Bas)
- WEF Labor Market Efficiency (Eff)
- Trade Freedom Index
- Financial Freedom Index
- Business Freedom Index
- Economic Freedom Index Overall Score
- WEF Health and Primary Education (Bas)
- WEF Economic Freedom (Eff)
- UNPD Simulate Mean Age at Marriage - Female
- GDP Per Capita 2007 \$US
- WBGI - Regulatory Quality
- Property Rights Index
- WEF Technological Readiness (Eff)
- WBGI - Control of Corruption
- Corruption Perceptions Index
- WBGI - Rule of Law
- WBGI - Government Effectiveness
- WEF Higher Education and Training (Eff)
- WEF Business Sophistication (Inn)
- WEF Innovation (Inn)
- WEF Group Innovation and Sophistication Factors
- WEF Group Basic Requirements
- WEF Infrastructure (Bas)
- WEF Group Efficiency Enhancers



Make Unified Group Vectors

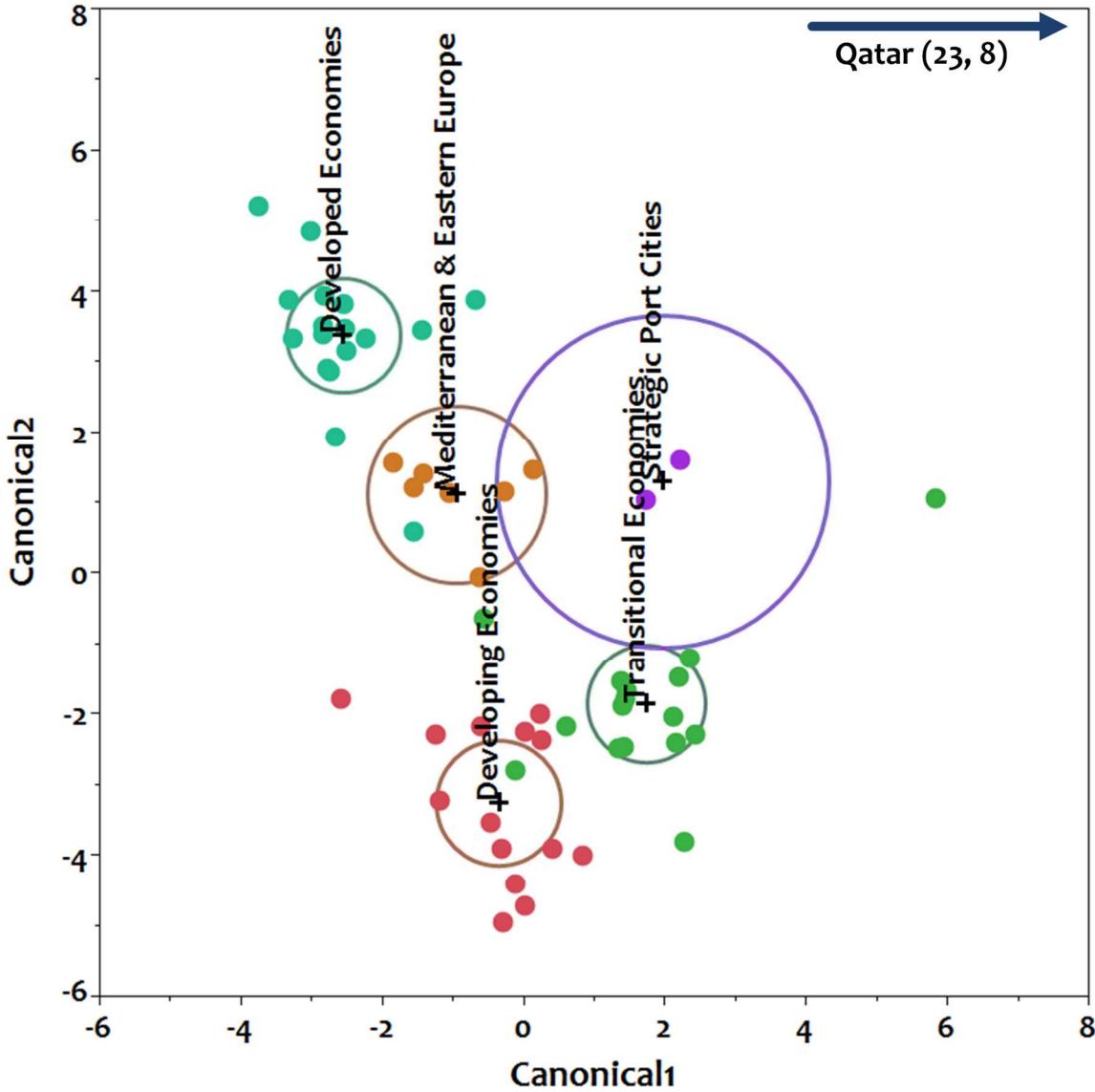
- ▶ Add new columns to main data table
- ▶ Use Formula editor to generate means of column normalized cluster members
 - Example: Data Group 1



The screenshot shows a list of six 'Col Standardize' formulas, each enclosed in a box with a red border. The formulas are:

- Col Standardize [*Uncertainty Avoidance Society Values (GLOBE)*],
- Col Standardize [*In-Group Collectivism Practices (GLOBE)*],
- Col Standardize [*Uncertainty Avoidance V-P*],
- Col Standardize [*Future Orientation V-P*],
- Col Standardize [*Collectivism In-Group - Institutional*],
- Col Standardize [*Future Orientation Society Values (GLOBE)*]

Pre-Screen to Differentiators

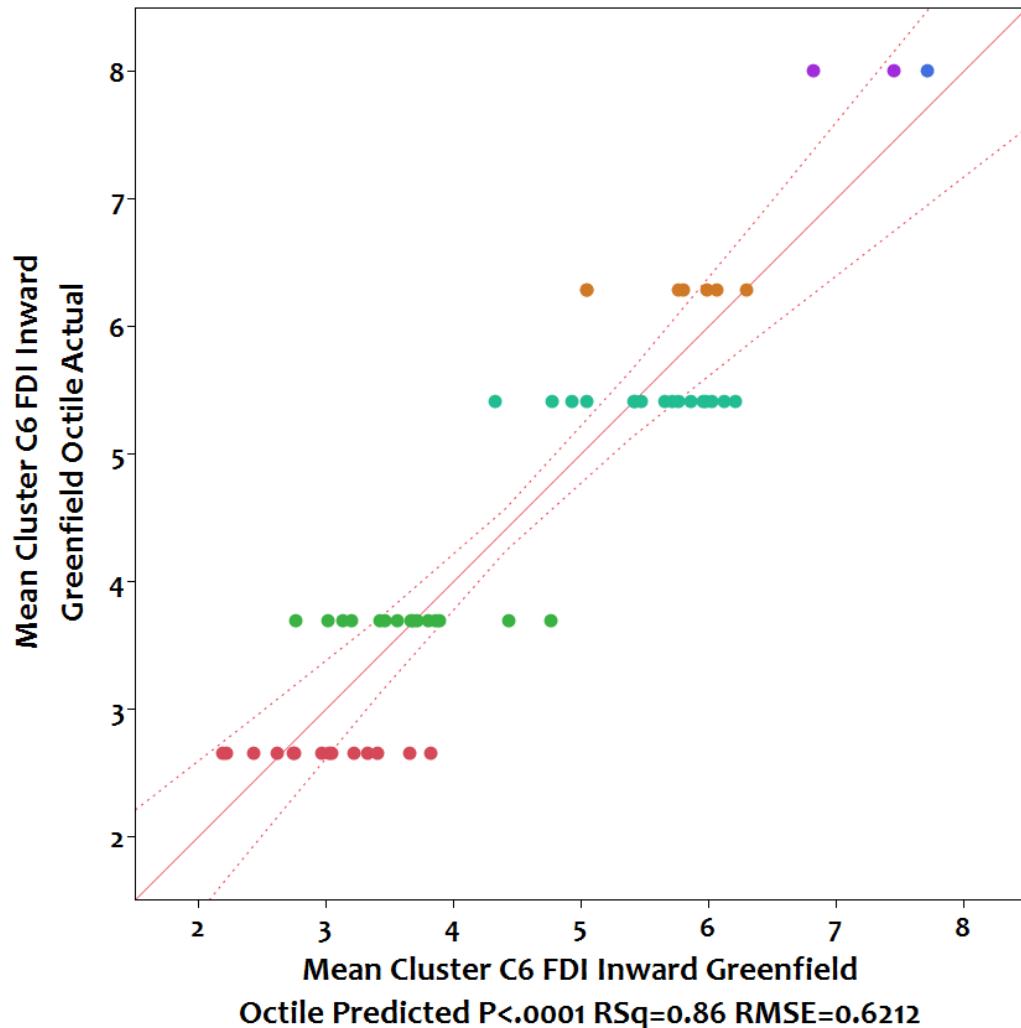


- ▶ Use Discriminant platform
- ▶ Stepwise variable selection
- ▶ Add all, remove one at a time until $p=0.1$

Out of 18 vectors, seven are non-differentiating with respect to cluster identification – these can be removed for final model determination

Model Fit to Vectors

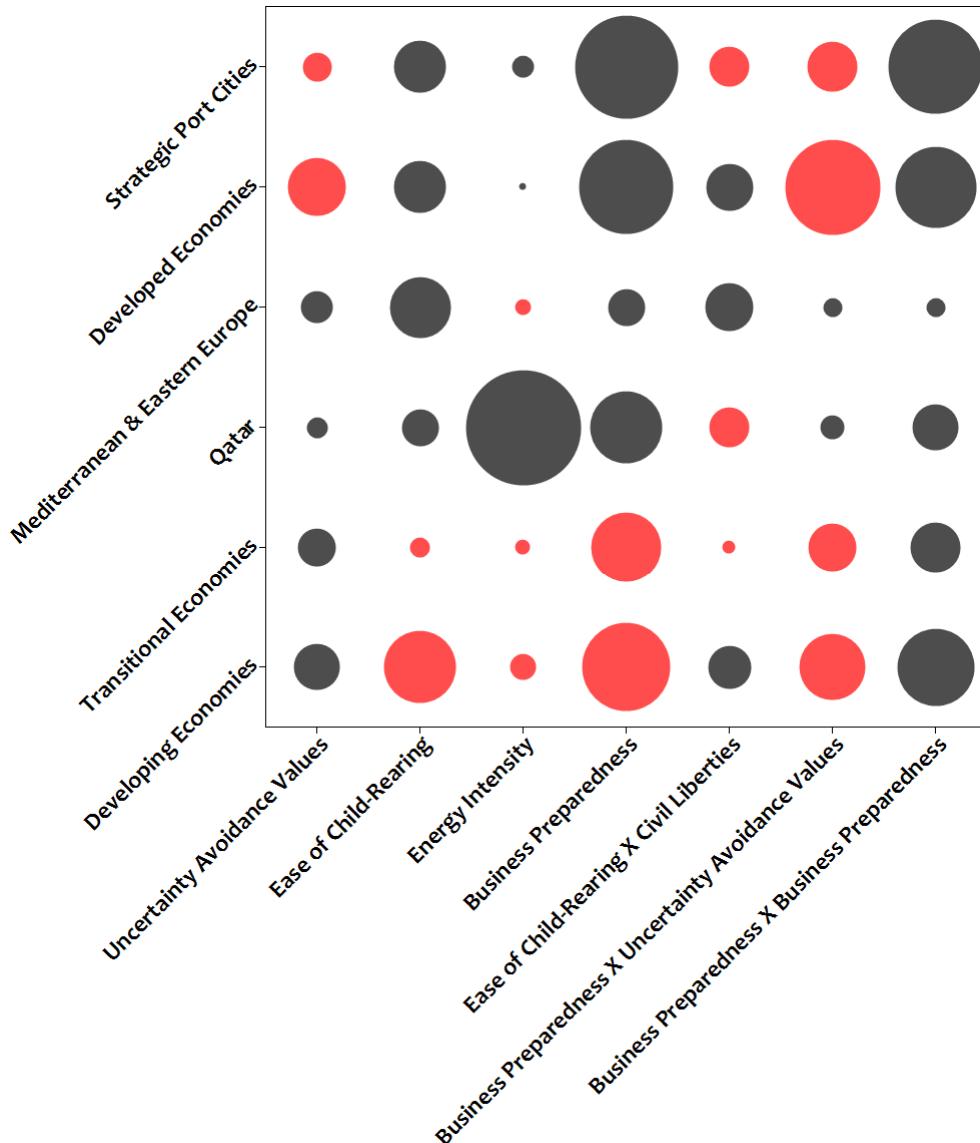
- ▶ Only five input data vectors with cohesive themes
- ▶ Clear separation of societal clusters aka cultural archetypes
- ▶ Useful for qualitative understanding



Semi-Quantitative Model Drivers

Group	Group Theme	Representative Variable	Correlation
1	Uncertainty Avoidance Values	Uncertainty Avoidance V-P (Values – Practices)	0.9376
2	Ease of Child-Rearing	Child Dependency Ratio	0.9691
3	Civil Liberties	FH Civil Liberties Rating	0.8850
9	Energy Intensity	Energy Production Per Capita	0.9818
18	Business Preparedness	Government Effectiveness	0.9836

Characterize Societies by Impact



- ▶ Model terms multiplied by society cluster mean input values
- ▶ Illustrates societal differences and similarities
- ▶ Identifies factors with limited relevance – see Energy Intensity

Conclusions

- ▶ Relatively stable cultural factors correlated with IGFDI
 - ▶ Inference: IGFDI trends unlikely to change rapidly regardless of immediate changes in economic marker data
- ▶ Method Has Limitations
 - ▶ Hierarchical clustering still dependent on data order
 - ▶ Influence of personal biases curtailed but not eliminated
 - ▶ Results are path-dependent and must be assessed for reasonableness
- ▶ Method Has Benefits
 - ▶ Useful cultural groupings can be developed from openly available data
 - ▶ Approach offers qualitative heuristic benefit for comprehension