

# A Knowledge-Based Selection Methodology of Peer Institutions

---

Marcel Nzeukou  
Diane Muntal

Office of Institutional Analysis & Planning  
University of Nevada, Las Vegas  
October 22, 2009

# Peerless

---

Marcel Nzeukou

Diane Muntal

Office of Institutional Analysis & Planning  
University of Nevada, Las Vegas  
October 22, 2009

# Peer Play



# A Peer for Every Season and Every Reason

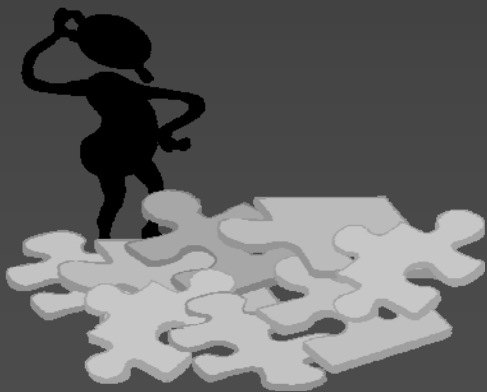
---

- UNLV Comparison Institutions:
    - The Early Years (1980-2000)
    - WICHE Universities
    - Legislative Peers – AB203
    - USN&WR 3<sup>rd</sup> and 4<sup>th</sup> tiers
    - Planning
-

# The Question

---

What are the 15 (or n) most similar institutions to UNLV?



# Choosing a Methodology

---

- Criteria:
    - Tied to strategic planning priorities
    - Methodologically legitimate
    - Easy to replicate and update
    - Provides opportunity for campus input
-

# Quantitative Methodology

---

**statistical clustering** is the most commonly used methodology

*Curry (1972) out of NCHEMS, Terenzini et al. (1980, 1983), Teeter (1984), Christal (1987) and Brinkman (1987)*

---

# Methodology: Clustering

---

- The results of a statistical clustering approach to select peers depend on:
    - Variables chosen
    - Quality of the variables chosen
    - Clustering algorithm and options
  - How do you choose the variables and an algorithm?
-



# Methodology: Cluster Algorithms

---

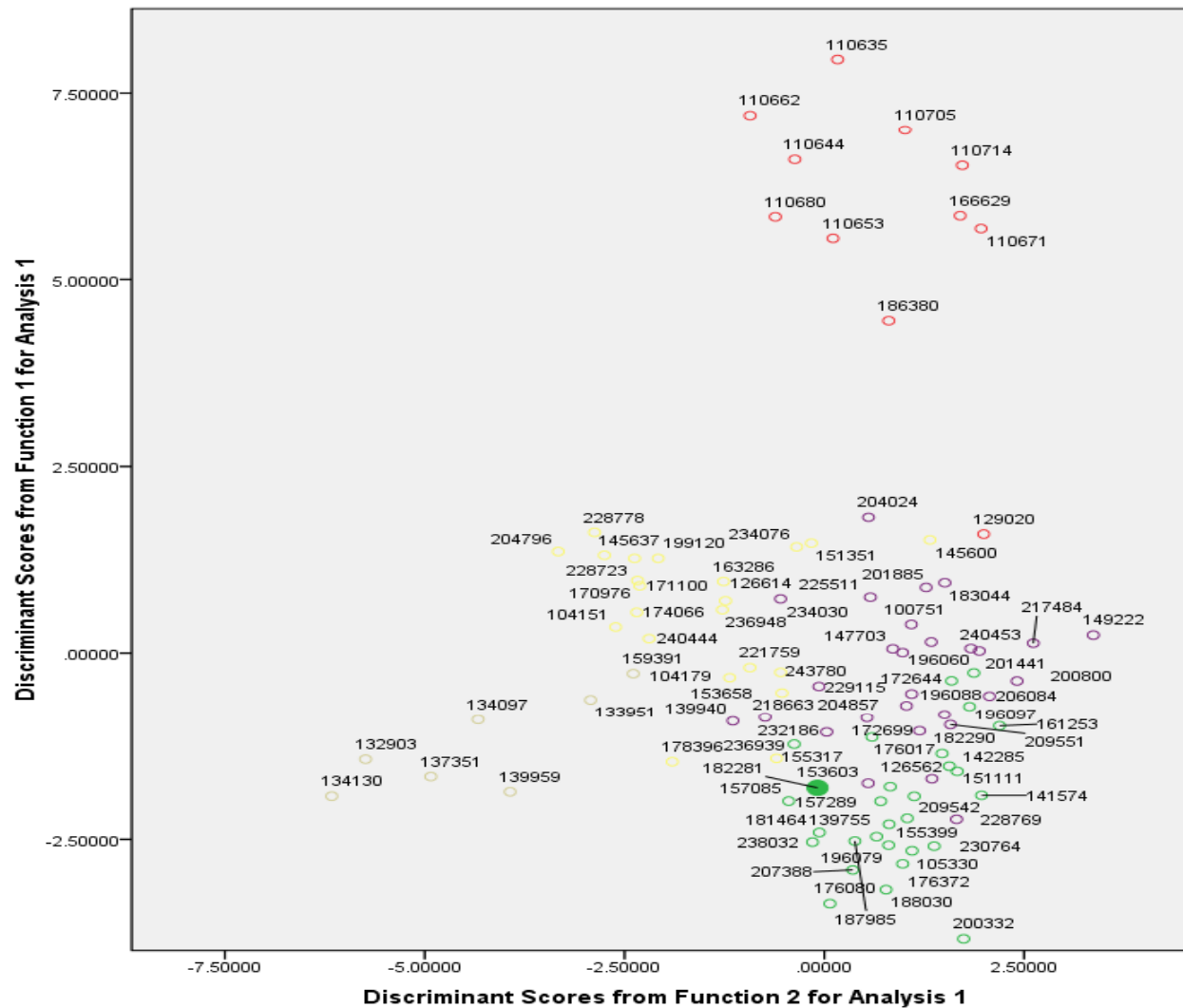
- Based on distance or similarity of measures between any two institutions
  - Distance/similarity is measured using institutional characteristics (Variables)
  - Minimize distance between institutions within clusters and maximize distance between clusters
-

# Methodology: Cluster Limitations

---

- The number of clusters is unknown
  - The number of institutions in the selected cluster may be too large or too small
  - Final peer set is a byproduct of clustering and not the focus
-

# Methodology: Cluster Limitations



# Methodology: Reference Institutions

---

- Total of over 2,800 accredited, public and private postsecondary institutions in U.S.
  - Only 102 Public, Comprehensive Doctoral Institutions including UNLV with high/very high research activity (Carnegie Classification of Institutions, 2005)
-

# Methodology: Variable Selection

---

- 54 variables grouped into 7 dimensions related to planning priorities
  - 7 dimensions are of interest to UNLV:
    - 1-Enrollment
    - 2-Student Performance
    - 3-Research
    - 4-Personnel
    - 5-Finance
    - 6-Financial Aid
    - 7-Urbanization
-

# Methodology: Variable Selection

---

- The most informative variables in each dimension are transformed into smaller number of principal components
  - Source : National Center for Education Statistics (NCES), Integrated Postsecondary Education Data System (IPEDS), 2006-07
-

# Limitations

---

- Limitations on national data elements collected
    - Reliability: variations in way national definitions are interpreted by institutions
    - Validity: how well they measure the concept
    - Missing values
-

# Variables:

## Enrollment Dimension

---

- % total enrollment for each race/ethnicity
  - % total enrollment for each gender
  - Full-time undergraduate enrollment
  - Part-time undergraduate enrollment
  - Full-time graduate enrollment
  - Part-time graduate enrollment
-



# Variables:

## Student Performance Dimension

---

- Full-time retention, part-time retention
  - Master's degrees awarded
  - Bachelor's degrees awarded
  - 4, 5, and 6-year graduate rates
  - 6-year graduation rate by gender
-

# Variables:

## Financial Aid Dimension

---

- % receiving federal grant aid
  - Average amount of federal grant aid
  - % receiving state/local & institutional grant aid
  - Average amount state/local & institutional grant aid
  - % receiving student loan aid received
  - Average amount of student loan aid
-

# Variables:

## Personnel Dimension

---

- Total FTE staff
  - FTE in instruction/research and public service
  - FTE in executive/administrative and managerial
  - FTE other employees (excluding medical schools)
-

# Variables:

## Research Dimension

---

- Total number of ranked programs
  - # Programs Ranked in Top 20
  - # Programs Ranked in Top 10
  - Doctoral degrees per tenured/tenure track faculty
  - Research expenditures as percent of total
  - Research expenditures per tenured/tenure-track faculty
-

# Variables:

## Finance Dimension

---

- Instruction as percent of total expenses
  - Public service as percent of total expenses
  - Academic support as percent of total expenses
  - Student service as percent of total expenses
  - Institutional support as percent of total expenses
  - Tuition and fees as a percent of total revenues
  - In-State Tuition and Fees
  - Out-of-State Tuition and Fees
-

# Variables:

## Finance Dimension

---

- Instruction expenses per FTE
  - Public service expenses per FTE
  - Academic support expenses per FTE
  - Student service expenses per FTE
  - Institutional support expenses per FTE
  - Average salary of full-time professors
  - Average salary of associate professors
  - Average salary of assistant professors
-

# Variables:

## Urbanization Dimension

---

- Degree of Urbanization
  - Land Grant Status
-

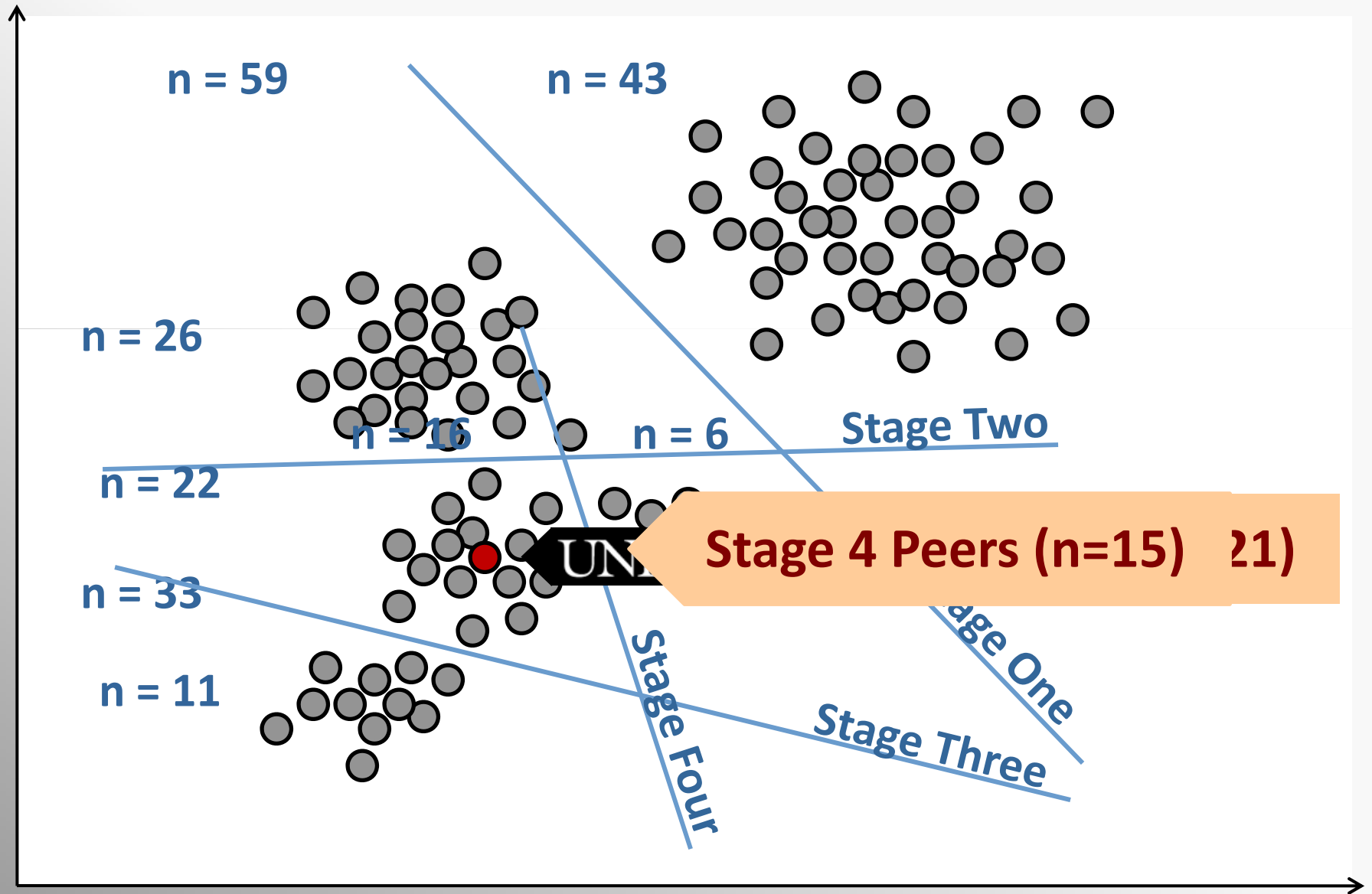
# Methodology: Sequential Clustering Elimination

---

- Two-step with sequential elimination
    - First step separates 102 comprehensive doctoral institutions into two sets: UNLV set and other institutions;
    - Eliminates institutions not in UNLV set over additional sequential stages
-



# Stage One with 90 Institutions (including UNLV)...



# Methodology: Indexing

---

- 1<sup>st</sup> stage delivers two clusters—one contains UNLV, the other drops out
  - 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> stages—at each successive stage cluster containing UNLV is split, with non-UNLV cluster dropping out
  - Each institution carries an index of 0-4, representing the number of times it clustered with UNLV, that can be used to estimate its similarity to UNLV
-

# Methodology: Advantages

---

- Allows for categorical variables (urban, rural, land grant, etc.)
  - Breaks large, complex problem into manageable steps
  - Not required to establish an ideal # of clusters
  - Identifying similar institutions becomes easier at each stage
  - Information is retained for each successive cluster using stage index
  - Stage indices provide information for external input or judgments
-

# Methodology: Disadvantages

---

- Institutions that are not selected in early stages will no longer appear in subsequent stages
  - Analysis is designed to produce peers; does not retain nearby clusters for other comparisons (e.g. aspirants)
-

# Results:

## Original Dimensions

---

- Institutions determined using 6 dimensions
  - Urbanization & land grant dimension considered separately later
  - The 10 institutions obtained form a set of “core institutions”
-

# Results:

## Stage 4 Original Dimensions

---

### **UNLV Stage 4 Peers**

### **Stage**

Florida International University	4
George Mason University	4
Georgia State University	4
The University of Texas at Arlington	4
University of Houston	4
University of North Texas	4
Virginia Commonwealth University	4

---

# Results:

## Stage 3 Original Dimensions

---

### UNLV Stage 3 Peers

### Stage

Florida State University

3

University of Central Florida

3

University of South Florida

3

---

# Results:

## Basic Dimensions Non-Peer Examples

---

University of Arizona	0
University of California-Berkeley	0
University of California-Los Angeles	0
University of Colorado at Boulder	0
University of Florida	0
University of Illinois at Urbana-Champaign	0
University of Michigan-Ann Arbor	0
University of North Carolina at Chapel Hill	0
University of Washington-Seattle Campus	0

---



# Dimension Effect Testing

---



# Dimension Effect Testing

---

- The sequential clustering algorithm applied to all dimensions is the control group
  - A treatment is obtained by applying the sequential clustering algorithm to all but one dimension
-

# Dimension Effect Summary

---

- The stage index measures the institution's proximity to UNLV associated with the presence/absence of dimension
  - The stage index computed for each set of dimensions is aggregated to measure closeness to UNLV in different dimensions
-

# Results: Dimension Effects



- Removing Finance
- Removing Research
- Adding Urbanization & Land-Grant

# Finance—Why test?

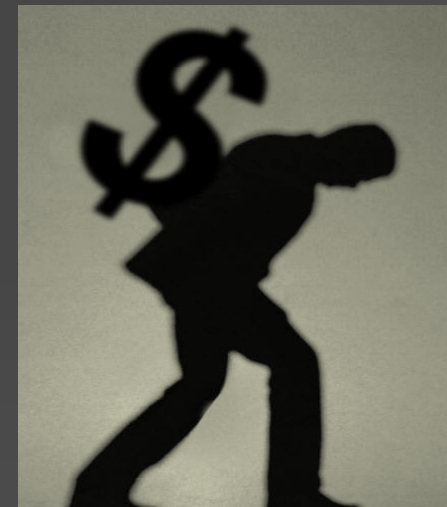
---

- To determine to what extent it constrains the composition of our peer group
  - Does it limit our peers to less well-funded institutions to the detriment of other characteristics?
-

# Impact of Removing Finance

---

- When the finance dimension is removed from the algorithm, UNLV clusters with institutions that are better funded



# Research—Why test?

- Importance of research as a strategic priority... issues similar to what we have said about finance.



# Results of Removing Research

---

- When the research dimension is removed from the algorithm, there is no change in the composition of the stage 4 and 3 clusters
  - It appears that the dimension is not bringing in additional information
-



# Results Refined: Urbanization & Land Grant

---

- Added two categorical variables to form the 7<sup>th</sup> dimension
  - Both factors help discriminate campus environment, location (urban, rural), and conditions in which it operates.
-

# Urbanization and Land Grant Effects

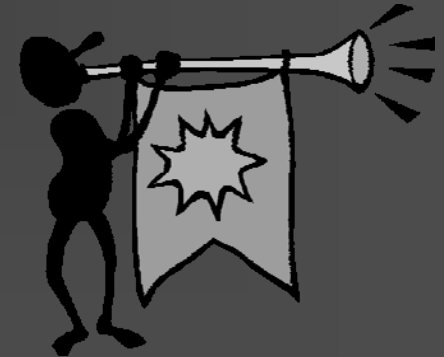
---

- Institutions that cluster more often with UNLV when Urbanization and Land Grant status are accounted for:
    - Have higher total research expenditures
    - Produce fewer doctoral degrees
    - Spend significantly more on students
-

# Final Results: Combining Peer Sets

---

- Combined peer sets to provide a more comprehensive view of candidate institutions
- Summed the stage measure for each institution to determine proximity to UNLV
- 5 new peers added to the 6 dimension results—these are highlighted in **green**.



# 15 Proposed Candidate Institutions

	All Dimension Stage	Finance Excluded Stage	Urbanization & Land Grant Stage	Total
UNLV	4	4	4	12
Florida International	4	4	4	12
Georgia State University	4	4	4	12
George Mason University	4	3	4	11
University of Houston	4	3	4	11
University of Louisville	2	4	4	10
University of New Mexico	2	4	4	10
Virginia Commonwealth	4	3	3	10

# 15 Proposed Candidate Institutions

	All Dimensions Stage	Finance Excluded Stage	Urbanization & Land Grant Stage	Total
The U. of Texas at Arlington	4	4	2	10
University of Central Florida	3	2	4	9
University of South Florida	3	2	4	9
Indiana U.-Purdue U.-Indianapolis	2	3	4	9
Temple University	2	3	4	9
University of Oklahoma-Norman	2	3	4	9
Wayne State University	2	3	4	9
University of North Texas	4	3	2	9

# Notes on Methodology

---

- Use normalized data to eliminate variable size effects (ranks, ratios)
  - By extracting only significant principal component factors:
    - Number of variables is reduced to a small set of factors
    - Number of variables in each dimension does not impact results; they are reduced to factors which contribute statistically independent information
    - Potential correlation among variables is eliminated
-

# Peer Into The Future

---

- Performance Comparisons
- Establishing Benchmarks
- Program Evaluation
- Aspirant Institutions
- Strategic Planning



# Questions?

---

- Marcel Nzeukou  
marcel.nzeukou@unlv.edu
  - Diane Muntal  
diane.muntal@unlv.edu
  - Website: <http://ir.unlv.edu/Presentations>
-