Forming Peer Groups: Strategic Operations, the Comparative Process, and a Spreadsheet Model

Dr. Richard Howard, Director of Institutional Research, University of Minnesota

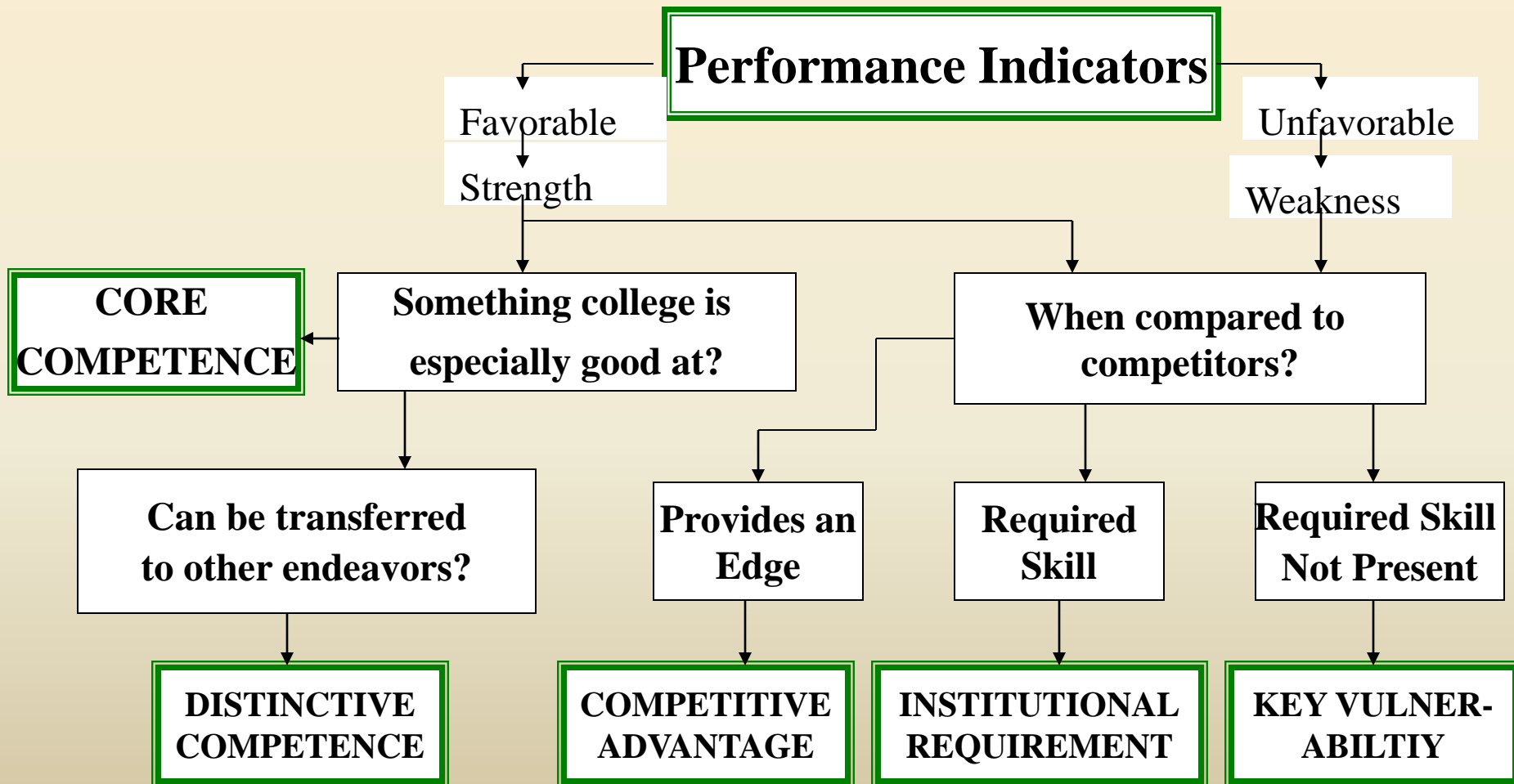*Dr. Josetta McLaughlin, Associate Professor of Management, Roosevelt University*

Dr. Gerald McLaughlin, Associate Vice President, DePaul University

AIR
Seattle, WA
May, 2008

Strategic Operations

- **Strategic operations** are those activities that help a college develop a competitive advantage and neutralize key vulnerabilities. These operations use indicators to identify and monitor the key activities.
- The status of the institution can then be informed based on knowing and comparing these indicators to other similar institutions.
- This workshop will identify the key concepts in this process and will provide hands-on exercises as participants go through those key steps. It will include use of a spreadsheet model to form the comparative group.

COMPETENCE & COMPETITIVE ADVANTAGE: HOW THEY FIT TOGETHER

Performance Indicators

Favorable

Strength

Unfavorable

Weakness

CORE COMPETENCE

Something college is especially good at?

When compared to competitors?

Can be transferred to other endeavors?

Provides an Edge

Required Skill

Required Skill Not Present

DISTINCTIVE COMPETENCE

COMPETITIVE ADVANTAGE

INSTITUTIONAL REQUIREMENT

KEY VULNER-ABILTIY

WHY CORE COMPETENCIES ARE IMPORTANT

CORE COMPETENCIES are the key to building a *COMPETITIVE ADVANTAGE*

CHARACTERISTICS

Core competencies may be a function of

- **Tangible Resources**
- **Intangible Resources**
- **Institutional Capabilities**

Core competencies may have their roots in qualities that are hard to quantify

- **Institutional History/Tradition**
- **Institutional Culture**
- **Institutional Location/Reputation**

CRITERIA FOR CORE COMPETENCIES

- **Valuable**
  - **Students are willing to pay.**

- **Rare**
  - **A capability (or combination of capabilities) that is possessed by few of the competitors.**

- **Costly to Imitate**
  - **The competency may be linked, for example, to a socially complex network extending beyond graduation.**

DATA FOR STRATEGIC PLANNING

- **Public Domain Data: Systematic data that is in the public domain and made available for use and discussion, e.g., IPEDS, U.S.News & World Report.**
  - **Assessment**
  - **Accountability**
  - **Advocacy**
- **Private Domain Data**

WHAT DO PUBLIC DOMAIN DATA TELL US?

How we compare with competitors on

- **Financial Health (e.g., money)**
- **Physical Well-Being (e.g. buildings)**
- **Information Capital (e.g., decision support)**
- **Human Capital (e.g., students & faculty)**

Requires the formation of the appropriate group of institutions (Peer, Aspiration, Competitor)

WHY THESE DATA ARE IMPORTANT

They provide:

- A road map for decision makers…
- A basis for **comparison**…
- A guide to understanding **vulnerabilities**…
- A guide for identifying **core competencies**…

CLOSING THOUGHTS ON APPEARING STRATEGIC

- Institutionalize the Strategic Management process through continual monitoring

- Use institutional and other data to support strategic decisions and implementation

- Bring data into the context of decisions that are being made

- Obtain data that are sufficient, relevant, timely, and reliable

- Summarize data to create information and intelligence

The Comparative Process

BENCHMARKING

What is Benchmarking?

- **"A benchmark, or <u>standard</u> by which something can be <u>measured</u>, is used to compare peers to improve results."** (source: NACUBO,http://www.nacubo.org/x8322.xml)

- **Benchmarking in higher education is the use or development of standards against which best practices can be identified and performance can be compared.**

- **Factoid: "Googling"** *<u>Benchmarking College University</u>* **will yield more than 86,700 hits.**

Why Is Use of Benchmarking Important?

- **Effective planning, decision making, and oversight require that you understand**
  - **the <u>context</u> of the institution and**
  - **how you <u>compare</u> to peers.**
- **Use of benchmarking and comparative data in development of key indicators can**
  - **lead to better <u>understanding</u> of institutional context**
  - **facilitate efforts to <u>monitor</u> progress and <u>support</u> strategic decision making.**

How Can Benchmarking Be Used for Decision making?

Benchmarking can be used to examine the Institution's external and internal:

- **Evaluate the competition.**
- **Formulate reasonable objectives.**
- **Identify areas of weaknesses.**
- **Guide policy development.**
- **Provide judgments for:**
  - **Budget requests**
  - **Salary adjustments**
  - **Teaching loads**
  - **Setting tuition**

Why Do Institutions Form Groups For Benchmarking?

- **Norms**

- **Nature of the Higher Education Sector**
  - **Institutions in Higher Education >7,000**
  - **Different Sectors in Higher Education**
    - **Public**
    - **Private Not-for-Profit**
    - **Private For-Profit**
  - **Sizes of Institutions within Sectors Vary**
  - **Data are Available for Forming Groups**
  - **Complex Organizations Require Strategic Planning and Monitoring**

Why Do an Increasing Number of Colleges Form Groups?

- **Competitive situation**
  - **Students and housing**
  - **Faculty and staff**
  - **Money and donors**
  - **Sales and services**
- **Complex environment**
  - **Accountability**
  - **Transparency**
 - **Diversity**
  - **Inclusiveness**
  - **Flexibility**
  - **Coordination**

Why Do an Increasing Number of Colleges Form Groups? (cont.)

- **Tough decisions need data for conflicting pressures:**
    - **More options, less time**
    - **More costs, less money**
    - **More skills, less foundation**
    - **More convenience, less brand**
- **Technology has increased expectations and abilities:**
    - **e-learning**
    - **Portals**
    - **Automated services**
    - **i-pods**

What Are the Steps in Benchmarking?

1. **Audit your Situation**
2. **Select Measures**
3. **Collect Data**
4. **Form Groups**
5. **Monitor Outcomes**
6. **Use Results**
7. **Adjust Process**
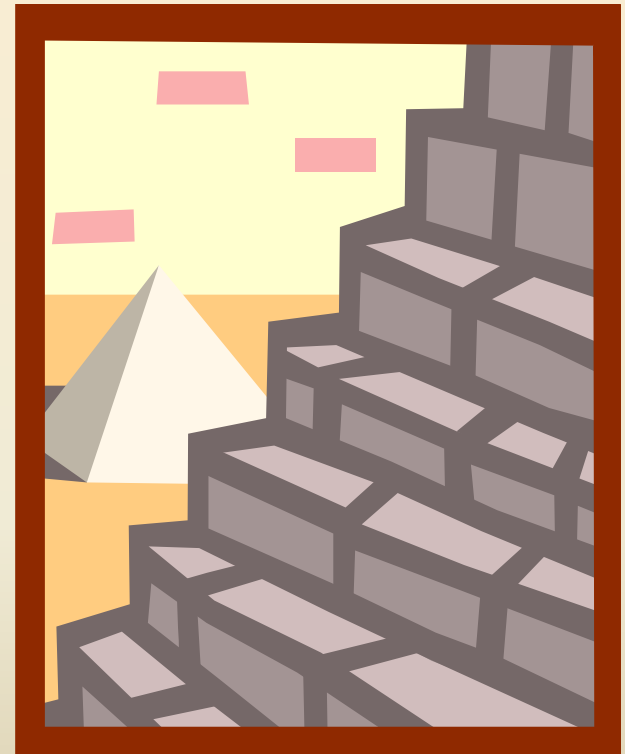
See Appendix for additional information

Checking the Outcomes…

- Techniques – Graphic, Statistical, (measure, Ratio, Index)
- Tools and Data
  - AGB
  - NCES-Executive Peer Tool
  - NCES-PAS
  - NCES-Data dump
  - CSRDE, CUPA-HR, Delaware, Moody, etc.
  - Guidestar, Ed Trust
  - Carnegie 2007

Comparison Groups

Uses of Comparison Groups

- **Evaluate Competition**
- **Provide Benchmarks**
- **Identify Areas of Weaknesses**
- **Guide Policy Development**
- **Provide Justification For:**
  - **Budget Requests**
  - **Salary Adjustments**
  - **Teaching Loads**
  - **Setting Tuition**

Process for Selecting Comparison Groups

- **Identify key strategic issues facing your institution**
- **Consider domains where there are indicators of success**
- **Review measures of success within domains**
- **Specify relevance of each measure**
- **Select subset with screens for issue areas**
- **Identify preliminary group**
- **Adjust for qualitative reasons**
- **Use, evaluate, and adjust**

Creating Meaningful Groups

- **Identify key strategic issues**
- **Consider areas where**

   success is essential
- **Review success measures**
- **Specify relevance of each**
- **Select subset with screens**
- **Identify preliminary group**
- **Adjust for qualitative reasons**
- **Use, evaluate, and adjust**

Different Types of Groups

- **"Traditional" Reference Group**
  - Preformed Groups – Athletic
  - Regional Groups - Geographical
- **Competitors**
  - for specific resources
  - HS Graduates, Research Funds, Grants
- **"Aspiration" Peers - Benchmarking**
  - What we want to be
  - More Money, Better Students, More Endowment
- **Comparable Groups (Peers?)-**
  - US News, Large Catholic
  - Our standing in a representative set (Carnegie)
  - A lot like us in some ways.
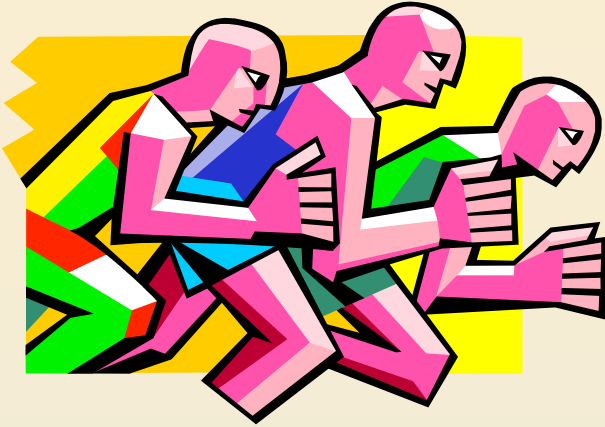
Types of Comparison Groups
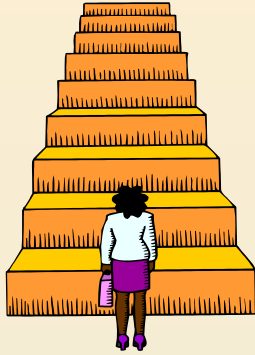
Aspirational
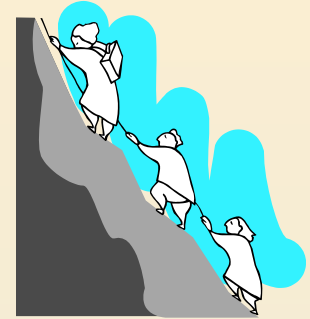
- **Dissimilar but reflect characteristics that the home institution desires to have.**

- **If presented as a peer group, you risk your credibility with the intended audience.**

Types of Comparison Groups

Competitor

- **Typically these are competing with your institution for students, faculty, and/or financial resources.**

- **May not be similar in mission, size, or complexity.**

- **Similarity may or may not be an issue.**
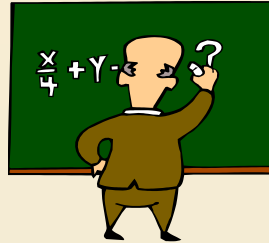
Types of Comparison Groups

Peer

- **Similar in role, scope, or mission**
- **Similar, not identical – key distinction**
- When variables such size, program content or research dollars are used to select the institutions, an acceptable degree of similarity needs to be defined.

Procedures for Developing Peer Groups

Data & Statistics   **Proximity Analysis**

Judgment   **Panel Review**
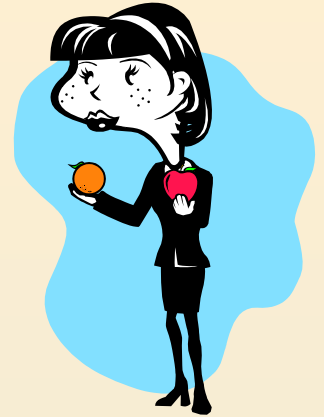
Data & Judgment   **Threshold Approach**

Data, Statistics, & Judgment   **Hybrid Approach**

How Are Group Members Selected?

- **Primary Factors**

  Finance, Carnegie Category, Mission, Size, Student Body Characteristics, Selectivity, Control, Carnegie Basic Category

- **Secondary Factors**

  Research Dollars, Region, Proximity, Residence Halls, Signature Programs, Service, Religious Affiliation, Athletic Conference, Historic Comparison Group, Faculty's Recommended Salary Group, Politics, Priorities/Weighted Factors, Judgment

Tentative variables for selecting peers

- – Local appropriations as % of State + Local
- – Total Fall headcount enrollment
- – Percent of faculty that are full-time
- – Student intensity measure (Fall full-time to part-time ratio times fall headcount to full-year unduplicated headcount ratio)
- – Percent minority in fall headcount
- – Percent of all awards that are allied health
- – Percent of all awards that are career-technical
- – Comparison variables, between colleges
- – A variety of fall headcount measures
  - • Percent minority, Percent full-time, etc

Data and Statistics

- **Nearest Neighbor**
  - **Select variables (can weight or not),**
  - **Standardize (or not)**
  - **Compute Distance (Multivariate Euclidian or correlation)**
  - **Rank in distance, Cut in vicinity with Scree Test**
- **Clusters**
  - **Like the nearest Neighbor but compute all distances**
  - **Use Cluster Analysis or MDS to form Groups.**
  - **Select preferred solution based on JUDGMENT**

Cluster Analysis

A Cluster is a group of like entities that are more alike than they are with entities in other clusters.

Caveats in Cluster Analysis

- **There is no one best way to do a cluster analysis**
- **There are many methods and most lack rigorous statistical reasoning or proofs**
- **Cluster analysis is used in different disciplines, which favor different techniques for:**
  - **Measuring the similarity or distance among subjects relative to the variables**
  - **and the clustering algorithm used**
- **Different clustering techniques can produce different cluster solutions**
- **Cluster analysis is supposed to be "cluster - seeking", but in fact it is "cluster - imposing"**

Other Approaches

- **Judgment**
  - **Panel of experts**
  - **Identify the rules and institutions**
  - **Iterate process (Delphi, Group Technique, etc.)**
- **Threshold Approach (NCHEMS)**
  - **Select the variables (Size, Money, Staffing)**
  - **Set the range (of ten set at 1/2X < X < 2X)**
  - **Iterate the process to get appropriate solution.**
- **Hybrid**
  - **Usually start with Statistics**
  - **Go to Judgment often based on Threshold**

Technical Issues

- **What size(s) and how many groups?**
- **What are the critical areas of concern and what critical indicators exist?**
- **What objective measures should be used? Magnitude, Performance Ratios, or Proportional Profiles?**
- **Should weighting be by domain area or by variable?**
- **What is the appropriate role for reasoned review?**
- **What is the appropriate improvement process? Success measures, refinement process, and strategic use?**

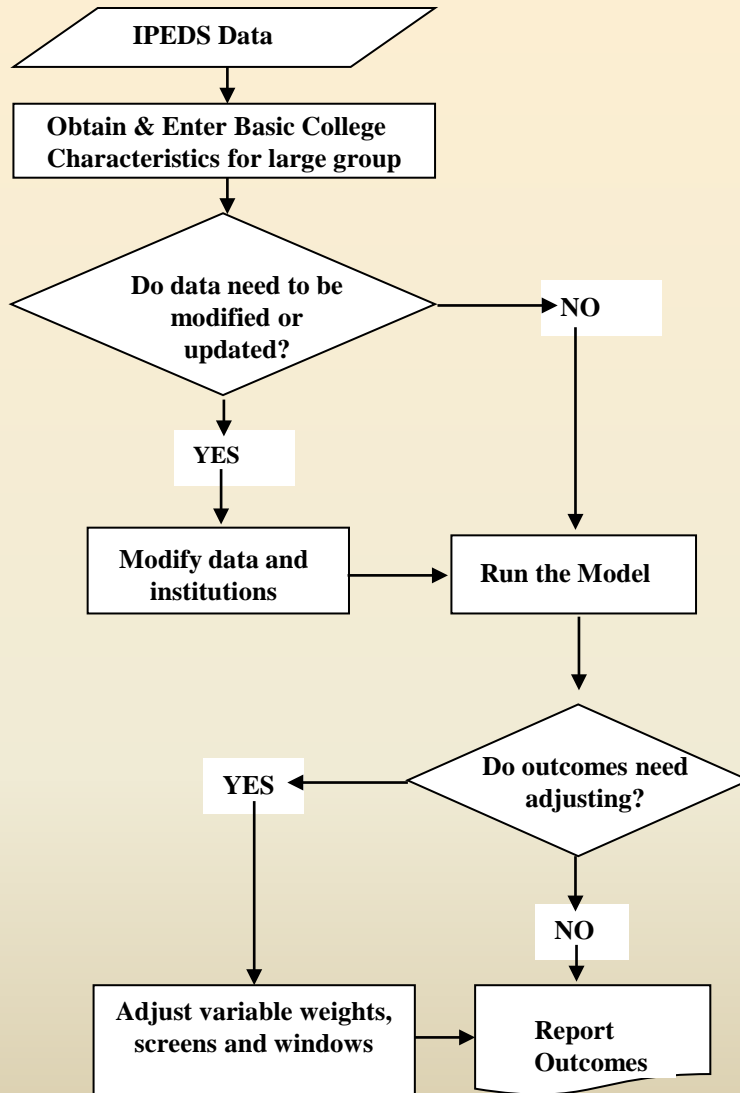- **Where you start determines where you end.**

Selecting Comparison Institutions Requires

- **An understanding of overt and hidden political agendas**

- **An awareness of different types of comparison groups that can be developed**

- **Understand that at some level the methodology used to select the comparison group will reflect the politics surrounding the issue.**

COMPUTER MODELING

- **Modeling and computer simulation are powerful tools for understanding.**

- **Examples of use:**
  - **Queues in heavy traffic**
  - **Manufacturing**
  - **Marketing**
  - **Education**

Reference Group Model

```
┌─────────────────────────────┐
      IPEDS Data
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│ Obtain & Enter Basic College │
│ Characteristics for large group │
└─────────────────────────────┘
              │
              ▼
         ◇ Do data need to be        ──────►  NO
           modified or                          │
           updated? ◇                           │
              │                                  │
             YES                                 │
              │                                  │
              ▼                                  ▼
┌──────────────────┐        ┌──────────────────┐
│ Modify data and  │──────► │   Run the Model   │
│  institutions    │        └──────────────────┘
└──────────────────┘                 │
                                     ▼
   YES ◄──────  ◇ Do outcomes need
    │             adjusting? ◇
    │                 │
    │                NO
    │                 │
    ▼                 ▼
┌──────────────────────┐   ┌──────────────────┐
│ Adjust variable weights,│ │     Report        │
│ screens and windows   │──►│    Outcomes       │
└──────────────────────┘   └──────────────────┘
```

Reference Group Model

Purpose:

- Produce a list of most similar colleges.

- Consider major components of college.

- Develop from the IPEDS data.

- Include students, faculty, revenues, academic programs, and expenditures.

- Allow varying importance of factors.

Model Components

Inputs:

- **Basic institutional characteristics**
- **Enrollments**
- **Staffing**
- **Finance**
- Degrees by program area

Model Components

Processes:

- **Form Balanced Scorecard Measures**
  - **Basic Characteristics(11)**
  - **Students(5), Faculty(6), Program(7)**
  - **Finance(6), Outcomes(9)**
- **Define <u>Same</u>, <u>Similar</u>, <u>Different</u>**
- **Weight variables**
- **Sum differences**
- **Rank neighbors**

Model Components

- **Output:**
  - **–Distribution of proximities**
  - **–Color coding of distances**
  - **–Scores and means of measures**
- **Iteration(s)**
  - **–Set Screens**
  - **–Adjust weights**

Model Process
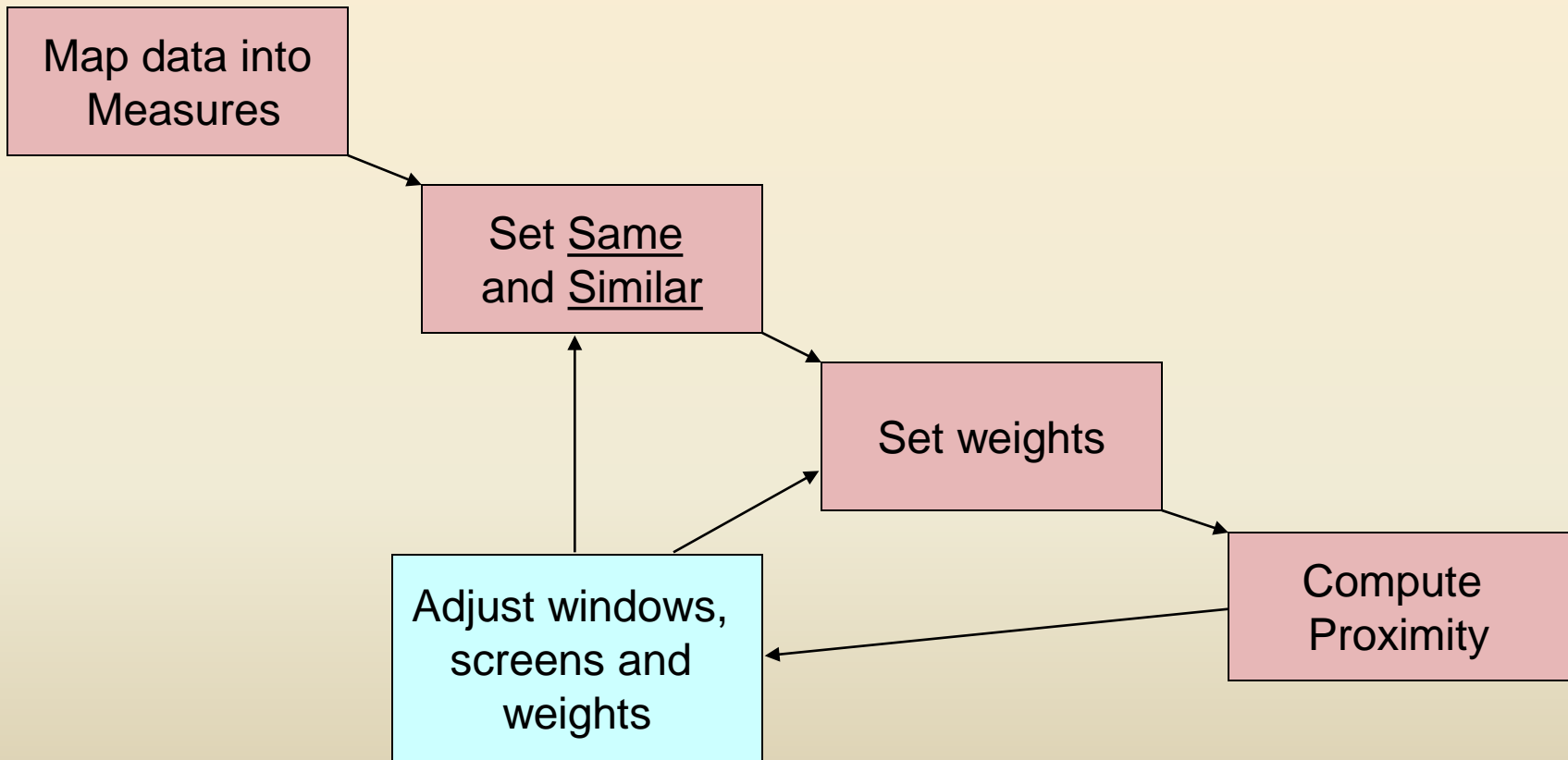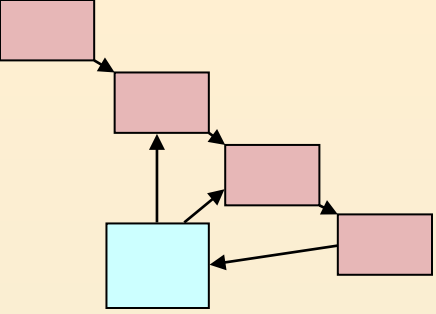
Developing the Group

- **Identify key characteristics of the focus institution. This is typically done to include size, academic degree level, and primary focus/mission.**

- **Select normative group with those characteristics**

- **Use the .mvl scripts to extract data from IPEDS on that group of institutions**

- **Download each of the three data sets to .cvs files. Sort in UNITID order.**
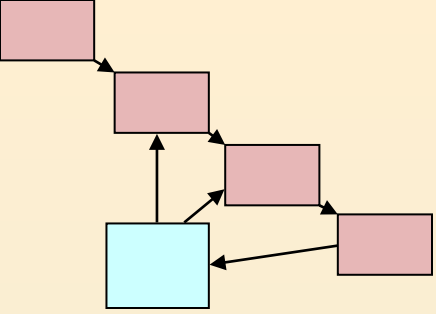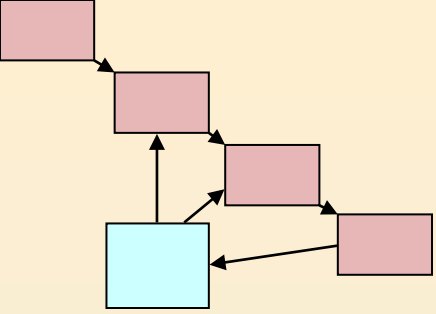
Developing the Group

- **Copy – Paste into the Normative model.**
- **Select the first row of the Public-Private Finance and Drag the formulas in it for the full set of institutions**
- **Select the first row of the BSC Measures and Drag the formulas for the full set of institutions. Delete institutions with major data problems. Compute missing values for those with minor problems.**
- **Compute the windows for SAME and SIMILAR**

Developing the Group

- **Continuous variables can have the windows set with Mean and Standard Deviation. Categories are selected for the nominal variables based on the definitions.**

- **The Focus institution is placed in the top row. Weights are assigned to the individual measures and the distances.**

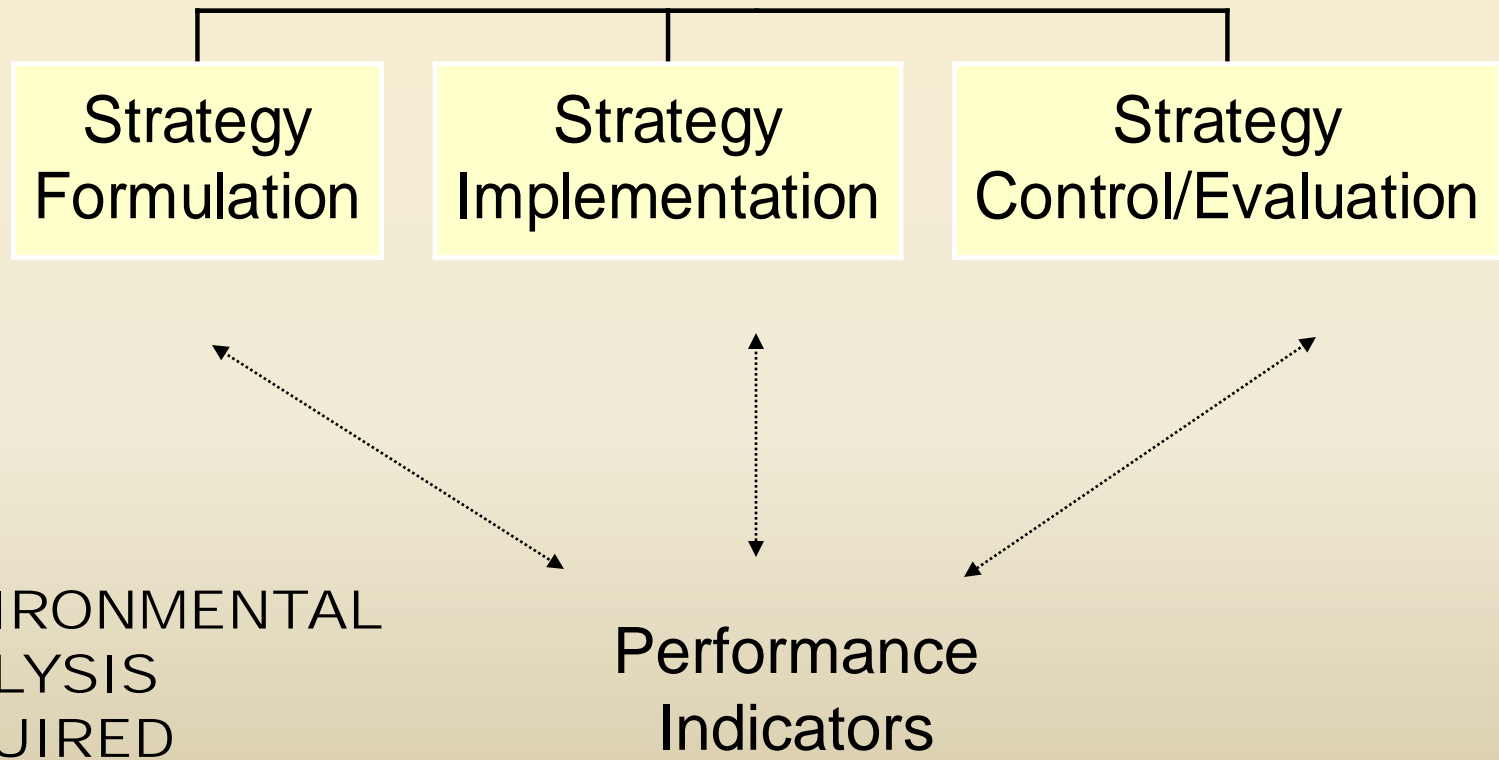- **The equations defining SAME and SIMILAR are Copied from the first row to other rows.**

Developing the Group

- **The integer scores are summed and the sum is divided by the sum of the weights forming a proximity measure.**

- **The data from BSC Measures and the proximity measure from Windows are copied to Distances.**

- **The data from Windows is copied to Neighbors. These two sheets must be sorted separately as they have common sources but are not connected through formulae.**

Appendices

MODELS

STRATEGIC MANAGEMENT MODEL

Strategy Formulation

Strategy Implementation

Strategy Control/Evaluation

ENVIRONMENTAL ANALYSIS REQUIRED

Performance Indicators

THE GENERIC STRATEGIES

	Lower Costs	Higher Costs
Broad Market	**Overall Cost Strategy**	**Differentiation Strategy**
Narrow Market	**Focused Overall Cost Strategy**	**Focused Differentiation Strategy**

SUMMARY TOOLS

STRATEGIC AUDITING

- Assessment of Current Performance
- Assessment of External Environment
  - Mega Environment (Political, Economic, Social, and Technological Trends
  - **Competitor Environment (Peer Groups)**
- **Assessment of Internal Environment**
- Identification of Strategic Alternatives
- Recommendations
- Action Plans

AUDIT:
Evaluate and Anticipate Key External Events

- **Political**
  - Financial Aid
  - Future Support of k-12
- **Economic**
  - Interest Rates
  - Economy for industries
- **Social**
  - Demographics
  - Career Demand
- **Technical**
  - Distance Learning
  - Knowledge Management

AUDIT:
Institutional Success Factors

Can you – Do you?:

- Provide quality education

- Support Student Success

- Offer Viable Programs

- Enhance the Mission
  - teaching, research, service
- Maintain institutional viability
  - applicants, donors, bankers
  - business, government
  - Compliance, risk management

- ## **<u>Compete effectively</u>**

THINKING ABOUT CORE COMPETENCIES, COMPETITIVE ADVANTAGE & VULNERABILITIES

	Institution's Core Competencies	Institution's Vulnerabilities
External Opportunities (from outside the institution)	**Do core competencies enable the institution to take advantage of opportunities?**	**Can the institution address vulnerabilities by taking advantage of some external opportunity?**
External Threats (from outside the institution)	**Do core competencies enable it to neutralize threats?**	**Can the institution prevent vulnerabilities and threats combining to destroy institutional effectiveness?**

CAVEATS: CORE COMPETENCIES & COMPETITIVE ADVANTAGE

- Core competencies are "something you are good at," but core competencies **change** over time.

- Core competencies can become core rigidities in the absence of ongoing **strategic management**.

- **Flexibility** is a must in a dynamic environment.

CORE COMPETENCIES & COMPETITIVE ADVANTAGE

Having a core competency does not automatically translate into having a competitive advantage.

Decisions can only lead to a competitive advantage if they result in outcomes that are **better** than those of the competitor.

Why Is Auditing the Situation Important?

**S
T
E
P**

1

- External: Anticipate Key Future Events
  - **Political**
    - Financial Aid
    - Future Support of K-12
    - Higher Education Reauthorization Act
  - **Economic**
    - Interest Rates
    - Economy for industries
  - **Social**
    - Demographics
    - Career Demand
  - **Technical**
    - Distance Learning
    - Knowledge Management

- Internal: Identify Institutional Success Factors
  - **Quality education**
    - Student Success
    - Viable Programs
  - **Mission Enhancement**
    - Teaching, research, service
    - Accessibility
  - **Institutional Effectiveness**
    - Applicants,
    - Donors
    - Government
  - **Operational Effectiveness and Efficiency**

How Are Measures Selected?

S
T
E
P

2

Start with the problems or issues!

- What is the problem/opportunity?
  - External Audit
  - Internal Audit
  - Strategic Plan
  - TOWS
  - Portfolio Matrix Analysis
  - Red Flags
- What kinds of data and information are needed to answer the questions?

When Are the Data Appropriate for Supporting Decisions?

S
T
E
P

3

- **Relevance** – Appropriate to discussion
- **Utility** – Usefulness to multiple groups
- **Applicability** – Meets needs of multiple groups
- **Interpretability** – Correctly understood
- **Credibility** – Believable
- **Fairness** – Balanced perspective
- **Scope** – Sufficient to major aspects
- **Availability** – Timely and accessible
- **Measurability** – Captured in a number
- **Cost** – Reasonable financial expenditure

What Decisions Need to Be Made to Form Groups?

S
T
E
P

4

- Types of groups that are appropriate
- How to form the groups
- Types of tools used to organize the data and information (Shown in Comparative Data)
  - Strategic audits
  - KPI's
  - Balanced scorecards
  - Program portfolios
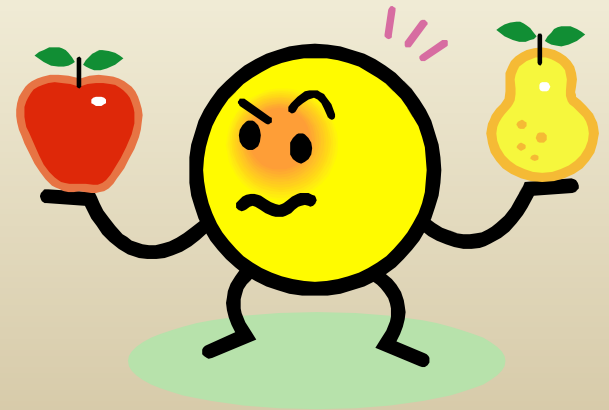
How Can Benchmarking Be Used to Monitor Outcomes?

S
T
E
P

5

Form a Basis for Judgments Based on:

- **Pre-stated Goal**
  - Mission-related
  - Conceptual Model/Belief
- **Trends Over Time**
  - Increase [Good goes up]
  - Decrease [Bad goes down]
  - Stable [Sometimes stability desired]
- **Comparisons with others**
  - Direct Competitors (1 – 4 institutions)
  - Peers (2-7 institutions)
  - Similar Group (10 – 25 institutions)
  - Large Normative Group (30+ institutions)

How Are the Results Used?

S
T
E
P

6

- **To Describe Core Competencies**
  - Capabilities, Resources, Location
  - History, Culture, Reputation
  - Valuable, Rare, Costly to imitate

- **To Create a Competitive Advantage**
  - Strength + Opportunity
  - When compared to competitor
  - Can not be easily duplicated

- **To Remove Key Vulnerability**
  - Weakness + Threat
  - When compared to competitor
  - Can not be easily changed
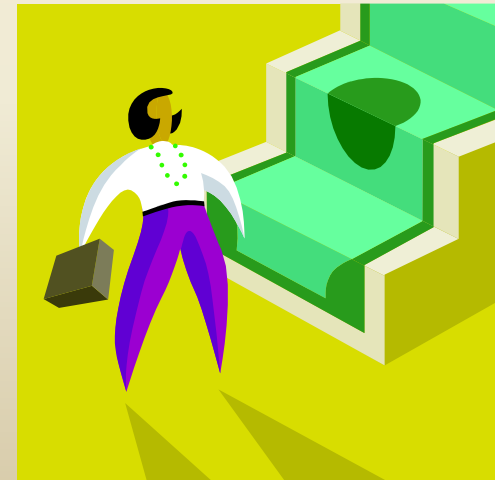
What Types of Adjustments Might Be Needed?

S
T
E
P

7

- Group size

- Number of groups

- Type of methodology used

- Clarification of "how much difference makes a difference"

- Priorities for various criteria

- Identification of competitors

- Reevaluation of importance of specific types of change over time

Predetermined

- **_Natural_** – **Institution belongs to a highly visible grouping – the nature of the specific comparable characteristic needs to evaluated.**

- **_Traditional_** – **group membership is based on historical relationships – tends to be a familiar association and maybe widely accepted.**

Types of Comparison Groups

Predetermined (cont.)

- *<u>Classification-based</u>* – a grouping used for national or regional reporting – e.g. Carnegie Classification

- Tend to have credibility and are usually recognized.

- Typically based on one or two characteristics resulting in wide variability on other institutional dimensions.

DATA

EXTERNAL COMPARISONS
PERHAPS YOU SHOULD BE CONCERNED IF:

- Your numbers are out of line when compared to those of your competitors. For example:
  - Your expenditures for non-instructional activities increases more or is higher than those of competitors.
  - You have higher discount rates than do the competitors.
 - You have substantially more programs than do the competitors.

Tentative variables for selecting peers

– **Local appropriations as % of State + Local**
– **Total Fall headcount enrollment**
– **Percent of faculty that are full-time**
– **Student intensity measure (Fall full-time to part-time ratio times fall headcount to full-year unduplicated headcount ratio)**
– **Percent minority in fall headcount**
– **Percent of all awards that are allied health**
– **Percent of all awards that are career-technical**
– **Comparison variables, between colleges**
– **A variety of fall headcount measures**
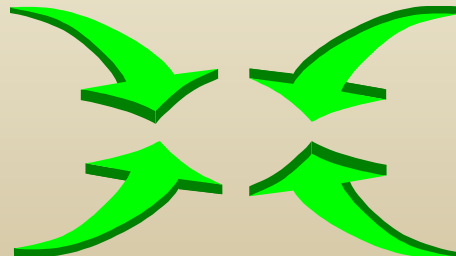 • **Percent minority, Percent full-time, etc**

Domains and Variables - Examples

Domain

Tuition and Financial Aid
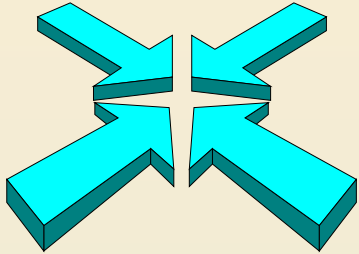
Variables

Tuition and fees / FT students
Financial aid / FTE students
Institutional aid as % of total aid
External aid as a % of total aid
% of students on work study
Inst. aid as a % of tuition + fee revenue

Students

Full time enrollment
First time applicants
Acceptances as a % of applicants
Enrolled as a % of acceptances
Enrollment by ethnic status
Ratio of the number of seniors to freshman
Ratio of graduate students to undergraduate

Finance Measures

(Typically the state focus)

- Total E & G Expenditures
- Total instructional expenditures
- Instructional expenditures on salaries only
- Instructional expenditures as a percentage of total E & G
- Other academic related expenditures
- Total academic related expenditures
- Total academic related expenditures as a percentage of total E & G
- Institutional support expenditures

Finance measures (cont.)

- Institutional support as a percentage of total E & G
- Public service expenditures
- Public service as a percentage of total E & G
- Student services expenditures
- Student services as a percentage of total E & G
- Total revenues
- Total tuition & fees revenues
- Tuition and fees revenue as a percentage of total revenue
- Total state revenues
- State revenues as a percentage of total revenues

Adjust Procedures

- What size group
- How many groups
- Which methodology
- How much difference makes a difference
- Priorities of various criteria
- Identification of competitors
- Importance of change over time

Predetermined (cont.)

- *Jurisdictional* – **Institutions that share the same political or legal jurisdiction.**

(Often the jurisdictional definition is the state line.)

- Typically this type of comparison group is used in politically driven situations even though the institutions may have little else in common.

Aspiration Groups

Relative Importance of Aspiration Reasons
Colleges of Business
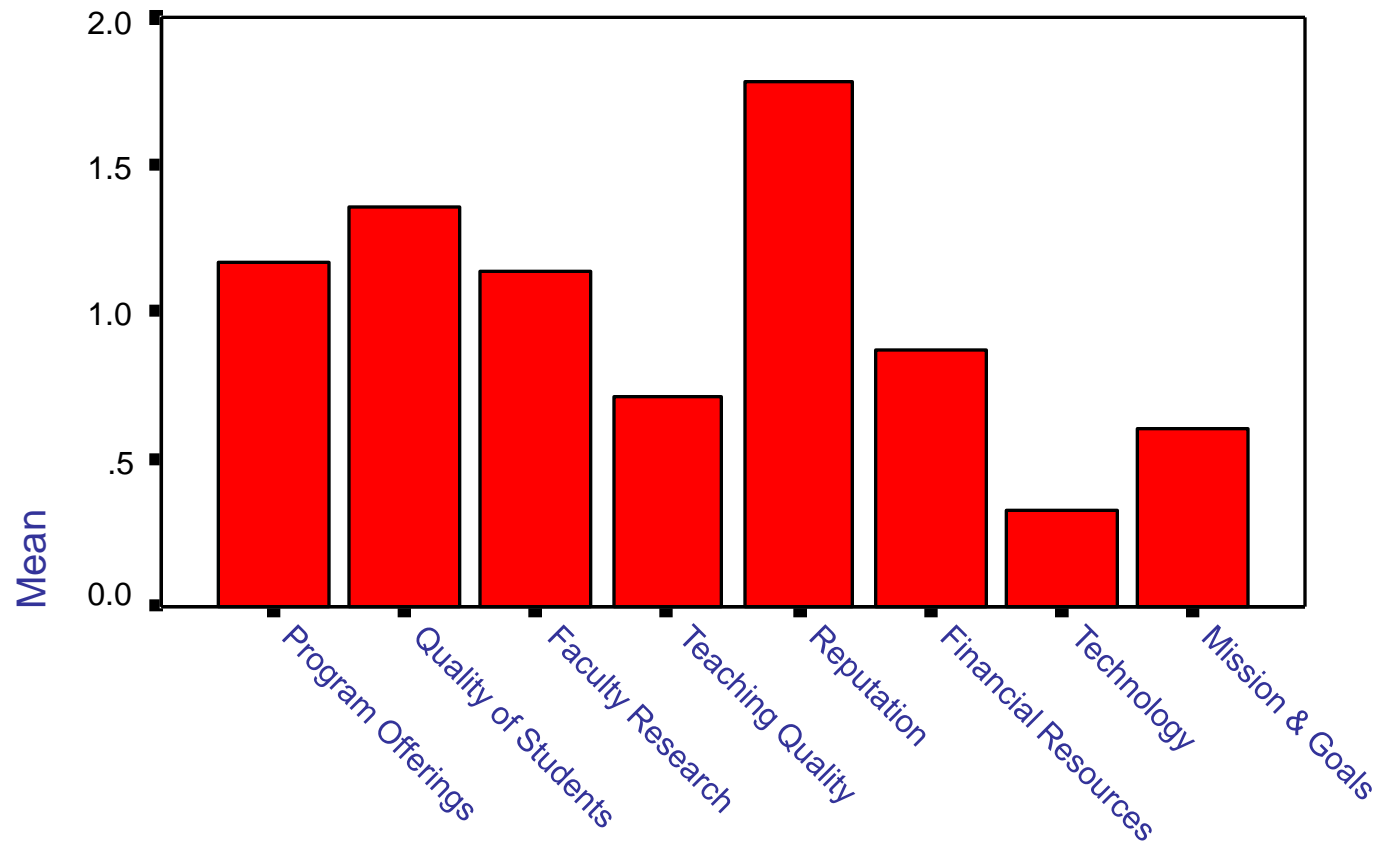
Compete for Students on:

- **Price –** Cost, Opportunity
- **Time and effort** required
- **Value added** – social, economic, intellectual, personal
- **Convenience** – attendance
- **Brand loyalty** – reputation, based on what you promise, communicate, and deliver

http://www.universitybusiness.com/page.cfm?p=65

NCES Categories

- **Public (Enrollment size)**
  - **Community Development & Career Institutions (<2k)**
  - **Community Connector Institutions (2k – 10k)**
  - **Community Mega-Connector Institutions (>10k)**
- **Private Not-For-Profit Institutions**
  - **Allied Health Institutions (100% Allied Health)**
  - **Connector Institutions**
- **Private For-Profit Institutions**
  - **Career Connector Institutions**
  - **Certificate Institutions (100% Certificate)**

NCES Classification System

Some Variables

- **Urbanicity**
- **12 Month Unduplicated HC**
- **% full-time first-time**
- **% Minority Students**
- **% Older Students**
- **% part-time Faculty**
- **% Awards as Certificates**
- **% Awards in Allied Health**
- **% Awards in occupational specialties**
- **% revenues from state and local support**
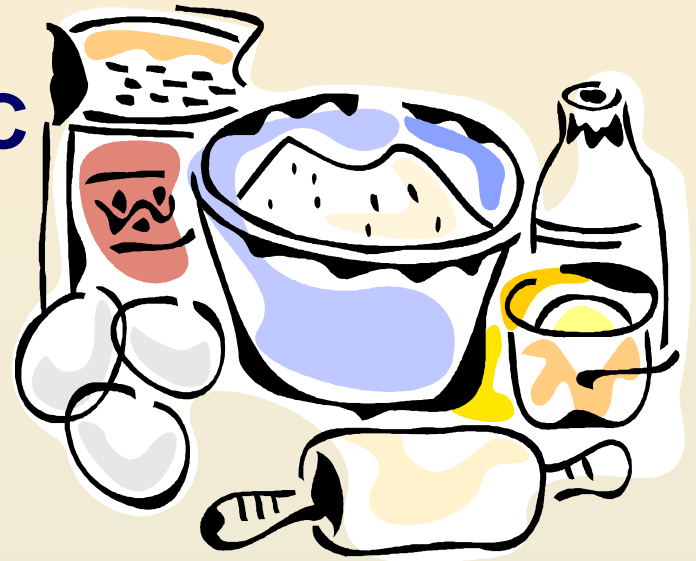
Process for Selecting Comparison Groups

- **Identify key strategic issues facing your institution**
- **Consider domains where there are indicators of success about strategic issues**
- **Review measures of success within domains**
- **Specify relevance of each measure**
- **Select subset with screens for issue areas or compute distance statistic**
- **Identify preliminary group**
- **Adjust for qualitative reasons**
- **Use, evaluate, and adjust**

Issues of Indicators

- **Need group of measures for similarities**
  - Often <u>BIG</u> items like control, size, region, type.
  - May need to refine with secondary items
    - Look at the Inputs – Money, students, faculty
    - Look at the processes – Business model, "productivity"

  OR
    - Look at the Outputs – Degrees, research, service

  THEN
  - Select the comparator measures.

Complete the sentence:

- **For institutions like us in terms of**
  - _____,
  - _____, and
  - _____ -

- **How do we do on**
  - _____,
  - _____, and
  - _____ ?

Steps in Cluster Analysis

1.   **Variable selection, construction of data base, testing assumptions**
2.   **Selecting measure of similarity or distance**
3.   **Selecting clustering algorithm**
4.   **Determining number of clusters**
5.   **Profile clusters**
6.   **Validation**

http://www.shsu.edu/~icc_cmf/cj_742/stats10.doc

Similarity/Distance Measures

- **Squared Euclidean Distance**
- **Euclidean Distance**
- **Pearson Correlation Coefficient**
- **Mahalanobis D^2**
- **Cosine of Vector Variables**
- **Minkowski Metric**
- **City Block or Manhattan Distances**
- **Jaccard's Coefficient**
- **Chebychev Distance Metric**
- **Gower's Coefficient**
- **Distances in the Absolute Power Metric**
- **Simple Matching Coefficient**

Clustering Techniques

- **Hierarchical** – Cases are formed into groups and the groups are themselves formed into groups at different levels to form a tree.

- **Optimization-partitioning** techniques – Clusters are formed by optimizing some 'clustering criterion' with mutually exclusive clusters.

- **Density or mode-seeking** – clusters are formed by searching for regions containing relatively dense concentration of cases.

- **Clumping** Techniques in which classes or clumps can overlap.

- **Other** – where the technique does not totally fit in one of the cases above.

STRATEGIC OPERATIONS

DATA AS TOOLS

Hierarchical – Cases Form a Tree

- **Agglomerative** – series of successive fusions of elements into groups. Can be based on the closest neighbor (single linkage) or to minimize the furthest distance between a pair of elements if a group is formed (furthest neighbor or complete linkage)

- **Divisive** – partition the set of n elements into finer partitions.

- **Issue -** When a step is performed it is irreversible.

- **Key Decision** - When to stop - e.g. How many groups are there?

Hierarchical Methods - Agglomerative Methods

- **Single average/linkage (nearest neighbor to cluster member)**

- **Complete average/linkage (furthest neighbor – minimize maximum for pairs)**

- **Average linkage (minimize)**

- **Ward's error sum of squares (minimize within cluster SS)**

- **Centroid method (minimize distance to centroid)**

- **Median clustering (new centroid median of two merged centroids)**

Cluster Analysis: Charles M. Friel Ph.D., Criminal Justice Center, Sam Houston State University

Hierarchical Methods -Divisive Methods

- **Trace methods**
  - **Minimize Trace of Within Cluster Sum of Squares**
  - **Minimize Determinant of Within Cluster Sum of Squares**
  - **Minimize Determinant of WCSS**
  - **Maximize Trace of BW^{-1}**
- **Automatic Interaction Detection (AID/CHAID) – Answer Tree: Split into groups based on the split that maximizes the resulting F statistic.**

Other Clustering Techniques

- **Optimization-partitioning techniques**
  1. **Initiate clusters**
  2. **Allocate entities**
  3. **Reallocate based on some algorithm**
  - **K-means Clustering (Randomly generate K groups and assign each element to nearest centroid – recompute centroids and iterate assignments. Stop when get same assignments**
  - **Splinter-Average Distance method (Move entity to splinter group when closer that to original group. Form splinter group when can't move any other individuals)**

- **Density or mode-seeking – Built on assumptions of the distribution and look for areas of low density. Relative low density indicates the presence of another cluster. Also includes Fuzzy-Set methods.**