What is your “unit of analysis” and more importantly why?

New tools and methods for teaching undergraduate social science students to think about data

Parvaneh Abbaspour & E.J. Carter
Aubrey R. Watzek Library • Lewis & Clark College
Portland, Oregon
Phenomena \rightarrow Data
Phenomena $\rightarrow$ Data

Human Condition $\rightarrow$ Numerical Values

Social world $\rightarrow$ Numerical Values
Ultimately, our attempts to describe the world through data are not processes of merely finding what is really there, but an active epistemological project of description and making the world knowable.

All data creation processes, even if unknowingly, make assumptions about the world and what exists as a unique unit that can be analyzed.

- J. Joque

...the categories through which data are identified, collected, aggregated and mediated are social and ideological constructs... we must ultimately account for them and their effects, while also recognizing their constructed nature. All of the data work that social scientists, scientists, marketers and others do inevitability rests on a very complicated process of defining and recognizing categories of things in the world.

- J. Joque

Undergraduate Student Awareness of Data Creation Processes

Science vs. Social Science
Undergraduate Social Science Student Awareness of Data Creation Processes

- Logistical limitations
- Lack of experience as data creators
- Limited life experience & mental tolerance for the reality of an entirely grey world
Data Availability

- An embarrassment of riches

Tends to compound abstraction from data generation.
Data Availability

- An embarrassment of riches

Tends to compound abstraction from data generation.
Common Approaches to helping students Find Data
1. Data Repositories
1. Data Repositories

- ICPSR
1. Data Repositories

- ICPSR

- Indexes curated datasets – embedding another layer of assumptions into data product students not naturally disposed to recognize or account for.
2. Unit of Analysis Worksheet

**FINDING DATA: WORKSHEET**

### Define your topic

<table>
<thead>
<tr>
<th>Specific yet flexible</th>
<th>What is your research question or notion? What are the related questions?</th>
</tr>
</thead>
</table>

### Define your units of analysis

<table>
<thead>
<tr>
<th>Who or What?</th>
<th><strong>Social Unit</strong>: This is the population that you want to study.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• People (individuals, couples, households)</td>
</tr>
<tr>
<td></td>
<td>• Organizations and Institutions (companies, political parties, nation states)</td>
</tr>
<tr>
<td></td>
<td>• Commodities and Things (crops, automobiles, arrests)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>When?</th>
<th><strong>Time</strong>: This is the period of time you want to study.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Cross-sectional (one specific point in time) vs. Longitudinal/Panel study (study over time)</td>
</tr>
<tr>
<td></td>
<td>• Current information vs. historical information (each scenario has its own challenges)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Where?</th>
<th><strong>Space</strong>: There are two main types of geographic classifications.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Political boundaries (nation, state, county, school district, etc.)</td>
</tr>
<tr>
<td></td>
<td>• Statistical/census geography (metropolitan statistical areas, tracts, block groups, etc.)</td>
</tr>
</tbody>
</table>
2. Unit of Analysis Worksheet

Data Menu or Wishlist
Lessons from Year One
Lessons from Year One

- Economics students generally good at math - often savvy statistics.
Lessons from Year One

- Economics students generally good at math - often savvy statistics.

- Generally lack experience with data creation and are blissfully unaware of the complexities and assumptions embedded in translating human world into numerical values.
Lessons from Year One

• Economics students generally good at math - often savvy statistics.

• *Generally* lack experience with data creation and are blissfully unaware of the complexities and assumptions embedded in translating human world into numerical values.

• *Not naturally disposed to question clean, clear answers they so seek.*
3. Consult the 2° Literature
3. Consult the 2° Literature

Integrate data conceptualization into first step of research assignment –

*conducting a literature review & preparing an annotated bibliography.*
Data Annotation

Level 1. – Basics
[Data Sources + Research Outcome]

1. What question is being asked & answered?
   2. What is the conclusion reached?

3. Note the data source(s) used to answer the question.

4. Does the paper include a citation or reference to the data set? Are there any other clues as to data access?
Data Annotation

Level 1. – Basics
[Data Sources + Research Outcome]

1. What question is being asked & answered?
2. What is the conclusion reached?
3. Note the data source(s) used to answer the question.
4. Does the paper include a citation or reference to the data set? Are there any other clues as to data access?
Data Annotation: Determinants

Level 2.

List the determinants the researchers have defined.

Note the data source for each.
Data Annotation: Determinants

Level 2+

List the determinants the researchers have defined.

What *justification* have they provided for each?

Note the data source for each.
Data Annotation: Unit of Analysis

Level 3.

What is the *unit of analysis*?
What is the *unit of analysis*?

*Why* have they chosen this unit?
Data Annotation: Example

The Impact of Gaming on the Indian Nations in New Mexico*

Thaddieus W. Conner, University of Oklahoma
William A. Taggart, New Mexico State University

Objective. This study examines the economic and social impact of Indian gaming on the residents of the 22 pueblos and tribes in New Mexico. Method. We employ a naturally occurring quasi-experimental design that classifies each of the Indian Nations into one of two groups, gaming and nongaming, depending on the continuous operation of a “Las Vegas” style casino for multiple years in the 1990s. For these two groups we compare aggregate, primarily U.S. Census, data spanning 25 indicators in both 1990 and 2000. Results. Although improvements were evident for both groups, nine of the 12 economic measures and six of the 13 social measures revealed a growing disparity favoring gaming nations during the 1990s, while six other measures suggested declining but continuing differences. These findings persisted in light of controls for population and urbanization, though many of the economic differences disappeared for the rural nations. Conclusion. Gaming has had a positive economic and social impact on the gaming pueblos and tribes in New Mexico, especially for the more urbanized nations. The gaming nations are enjoying higher incomes, lower levels of poverty, and improvements in selected social areas compared to those nations opting not to pursue casino gaming in the 1990s.

The Impact of Gaming on the Indian Nations in New Mexico*

Thaddieus W. Conner, University of Oklahoma
William A. Taggart, New Mexico State University

Objective. This study examines the economic and social impact of Indian gaming on the residents of the 22 pueblos and tribes in New Mexico. Method. We employ a naturally occurring quasi-experimental design that classifies each of the Indian Nations into one of two groups, gaming and nongaming, depending on the continuous operation of a “Las Vegas” style casino for multiple years in the 1990s. For these two groups we compare aggregate, primarily U.S. Census, data spanning 25 indicators in both 1990 and 2000. Results. Although improvements were evident for both groups, nine of the 12 economic measures and six of the 13 social measures revealed a growing disparity favoring gaming nations during the 1990s, while six other measures suggested declining but continuing differences. These findings persisted in light of controls for population and urbanization, though many of the economic differences disappeared for the rural nations. Conclusion. Gaming has had a positive economic and social impact on the gaming pueblos and tribes in New Mexico, especially for the more urbanized nations. The gaming nations are enjoying higher incomes, lower levels of poverty, and improvements in selected social areas compared to those nations opting not to pursue casino gaming in the 1990s.


**Annotation:** …

**Research question:** What is the economic & social impact of gaming … on the pueblos & tribes of New Mexico.

**Conclusion:** Gaming has had a positive social and economic impact … on the gaming pueblos and tribes of New Mexico.

**DATA Annotation:**

**Data source:** U.S. Census Data (primarily)…

**Data set citation:** Census data compiled by the Harvard Project on American Indian Economic Development (Taylor and Kalt, 2005).
The Impact of Gaming on the Indian Nations in New Mexico*

Thaddieus W. Conner, University of Oklahoma
William A. Taggart, New Mexico State University

Objective. This study examines the economic and social impact of Indian gaming on the residents of the 22 pueblos and tribes in New Mexico. Method. We employ a naturally occurring quasi-experimental design that classifies each of the Indian Nations into one of two groups, gaming and nongaming, depending on the continuous operation of a “Las Vegas” style casino for multiple years in the 1990s. For these two groups we compare aggregate, primarily U.S. Census, data spanning 25 indicators in both 1990 and 2000. Results. Although improvements were evident for both groups, nine of the 12 economic measures and six of the 13 social measures revealed a growing disparity favoring gaming nations during the 1990s, while six other measures suggested declining but continuing differences. These findings persisted in light of controls for population and urbanization, though many of the economic differences disappeared for the rural nations. Conclusion. Gaming has had a positive economic and social impact on the gaming pueblos and tribes in New Mexico, especially for the more urbanized nations. The gaming nations are enjoying higher incomes, lower levels of poverty, and improvements in selected social areas compared to those nations opting not to pursue casino gaming in the 1990s.
Determinant List

**Economic Impacts (12 measures)**

- Per-capita income
- Median household income

--3 indicators of poverty
- % of families living below poverty line
- % of children living in poverty
- Deep poverty rate - % of individuals living at less than 75% of poverty level
- % of individual receiving SSI
- Unemployment rate
- Location quotient

**Social Impacts (13 measures)**

--4 measures of housing
- Individuals living in crowded homes
- % of homes lacking complete plumbing
- % of homes lacking a complete kitchen
- % of structures built between 1990-94 & % of structures built between 1995-2000

--2 measures of health
- % of mothers reporting diabetes
- % of mothers receiving little prenatal care

--3 interconnected measures of education
- % of individuals 25+ having:
  - Earned an associates degree or higher
  - Earned a high school diploma
  - Completed less than ninth grade
Social measures - definitions

1. Individuals living in crowded homes – a house with > 1 person per room
   Rationale:
   Source:

2. % of homes lacking complete plumbing – a house lacking hot & cold piped water or a toilet.
   Rationale:
   Source: US Census

3. % of homes lacking a complete kitchen – a house lacking either a stove or refrigerator

5. % of mothers reporting diabetes
   Rationale: Historically, Native Americans have reported a higher level of incidents of diabetes.

6. % of mothers receiving little prenatal care
   Rationale: Historically, Native Americans have received less prenatal care

**Annotation:** …

**DATA Annotation:**

**Research question:** What is the economic & social impact of gaming …on the pueblos & tribes of New Mexico.

**Conclusion:** Gaming has had a positive social and economic impact … on the gaming pueblos and tribes of New Mexico.

**Data source:** U.S. Census Data (primarily)…

**Data set citation:** Census data compiled by the Harvard Project on American Indian Economic Development (Taylor and Kalt, 2005).

**Unit of Analysis:** Pueblos & Tribes of New Mexico two time points 1990 & 2000

**Rationale:** …
## Data Review: Unit of Analysis

<table>
<thead>
<tr>
<th>Possibilities</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. States:</td>
<td>+ States serves as primary regulator of gaming activities</td>
</tr>
<tr>
<td>Data source: …</td>
<td></td>
</tr>
<tr>
<td>Frequency: Annual</td>
<td></td>
</tr>
<tr>
<td>2. Counties:</td>
<td>+ Data readily available on of social and economic characteristics.</td>
</tr>
<tr>
<td></td>
<td>+ In greater proximity to casino operations.</td>
</tr>
<tr>
<td>Data Source: …</td>
<td></td>
</tr>
<tr>
<td>Frequency: Annual</td>
<td></td>
</tr>
<tr>
<td>3. Indian Tribes:</td>
<td>+ Data is greatest proximity to casino operations.</td>
</tr>
<tr>
<td>&amp; Pueblos</td>
<td></td>
</tr>
<tr>
<td>Data Source: US Census</td>
<td></td>
</tr>
<tr>
<td>Frequency: Decadal</td>
<td></td>
</tr>
</tbody>
</table>
Data Review: Determinant Inventory

Economic Impacts

Measures of Income
- Per-capita income
- Median household income

Indicators of Poverty
- % of families living below poverty line
- % of children living in poverty
- Deep poverty rate

Measures of employment
- % of individual receiving SSI
- Unemployment rate
- Location quotient

Miscellaneous

DATA SOURCE

Measures of Income
- Census, County Reports,…

Indicators of Poverty
- Census, County Reports,…

Measures of employment
- Census, County Reports,…

Miscellaneous
Why does it matter?

The Impact of Gaming on the Indian Nations in New Mexico*

Thaddieus W. Conner, *University of Oklahoma*

William A. Taggart, *New Mexico State University*

*Objective.* This study examines the economic and social impact of Indian gaming on the residents of the 22 pueblos and tribes in New Mexico. *Method.* We employ a naturally occurring quasi-experimental design that classifies each of the Indian Nations into one of two groups, gaming and nongaming, depending on the continuous operation of a “Las Vegas” style casino for multiple years in the 1990s. For these two groups we compare aggregate, primarily U.S. Census, data spanning 25 indicators in both 1990 and 2000. *Results.* Although improvements were evident for both groups, nine of the 12 economic measures and six of the 13 social measures revealed a growing disparity favoring gaming nations during the 1990s, while six other measures suggested declining but continuing differences. These findings persisted in light of controls for population and urbanization, though many of the economic differences disappeared for the rural nations. *Conclusion.* Gaming has had a positive economic and social impact on the gaming pueblos and tribes in New Mexico, especially for the more urbanized nations. The gaming nations are enjoying higher incomes, lower levels of poverty, and improvements in selected social areas compared to those nations opting not to pursue casino gaming in the 1990s.

Revised Worksheet to aid in Finding Data

What phenomena will you explore?

_____x____ on ___y___ , ___z___ , ______

Define the **type of determinant** you’d like to find and provide a justification for each?

<table>
<thead>
<tr>
<th>Determinant Type</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td></td>
</tr>
</tbody>
</table>

As specific measures are found: record exactly how they are measured & tabulated.

What possible **units of analysis** could you have? What are the advantages to each? Who might collect each type of data?

1. 
2. 
Questions?
Questions?

Thank you!