

Data Literacy for Rural Economies

Andrew J. Van Leuven

December 8, 2022

Today's Objectives

Part One

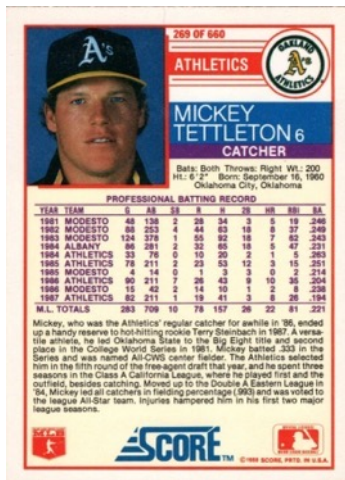
- Discuss some data basics: what is it, how to find it, how to work with it
- Become familiar with a variety of data sources pertaining to rural economic and community development

Part Two

- Discuss how to use that data to help community leaders and decision makers in your county with their economic goals.
- Discuss what types of data (or data-related topics) are worth covering in a possible follow-up in-service training

Data Basics

What is Data?



269 OF 660

ATHLETICS

MICKEY TETTLETON 6
CATCHER

Bats: Both Throws: Right Wt.: 200
Ht.: 6' 2" Born: September 16, 1960
Oklahoma City, Oklahoma

PROFESSIONAL BATTING RECORD

YEAR	TEAM	G	AB	R	H	2B	HR	RBI	BA	
1981	MODESTO	46	138	2	28	34	3	5	19	.246
1982	MODESTO	88	253	4	44	83	18	6	37	.249
1983	MODESTO	124	378	1	55	92	18	7	62	.243
1984	ALBANY	86	281	2	32	65	18	5	47	.231
1984	ATHLETICS	33	76	0	10	20	2	1	5	.263
1985	ATHLETICS	78	211	2	23	53	12	3	16	.251
1985	MODESTO	4	14	0	1	3	3	0	2	.214
1986	ATHLETICS	90	211	7	26	43	9	10	35	.204
1986	MODESTO	15	42	2	14	10	1	2	8	.236
1987	ATHLETICS	82	211	1	19	41	3	6	28	.194
M.L. TOTALS		283	709	10	78	157	26	22	81	.221

Mickey, who was the Athletics' regular catcher for awhile in '85, ended up a handy reserve to hot-hitting rookie Terry Steinbach in 1987. A versatile athlete, he led Oklahoma State to the Big Eight title and second place in the College World Series in 1981. Mickey batted .333 in the Series and was named All-CWS center fielder. The Athletics selected him in the fifth round of the free-agent draft that year, and he spent three seasons in the Class A California League, where he played first and the outfield, besides catching. Moved up to the Double A Eastern League in '84, Mickey led all catchers in fielding percentage (.993) and was voted to the league All-Star team. Injuries hampered him in his first two major league seasons.

SCORE™ © 1987 SCORE, PHOTO U.S.A.

- Data is plural for *datum*, which refers to **a piece of information**
- We use data to answer questions or to have a better idea of what is going on
- Data are rectangular:
 - Rows are **observations** (horizontal)
 - Columns are **variables** (vertical)

Data: An Example

Example Data: Hotels & Motels in Shawnee

Company Name	Year Founded	Physical Address	Physical City	Physical County	Sales	Employees
SAC & FOX CASINO	2004	42008 WESTECH RD	SHAWNEE	POTTAWATOMIE	\$6,461,694	22
BLACK HAWK CASINO	2013	42008 WESTECH RD	SHAWNEE	POTTAWATOMIE	\$5,173,580	19
GRAND CASINO HOTEL RESORT	2011	777 GRAND CASINO BLVD	SHAWNEE	POTTAWATOMIE	\$2,743,873	57
KICKAPOO CASINO	2013	38900 W MACARTHUR ST	SHAWNEE	POTTAWATOMIE	\$1,246,710	29
MIDWEST HERITAGE INN OF SHAWNEE	2014	4851 N KICKAPOO AVE	SHAWNEE	POTTAWATOMIE	\$1,032,519	14
LKV ENTERPRISES INC	1991	4900 N HARRISON ST	SHAWNEE	POTTAWATOMIE	\$1,000,000	18
SHIV PROPERTIES, LLC	2007	4900 N HARRISON ST	SHAWNEE	POTTAWATOMIE	\$789,940	19
LAXMI INN, INC.	2007	5401 ENTERPRISE CT	SHAWNEE	POTTAWATOMIE	\$439,557	17
HOLIDOG INN, LLC	2016	14285 ACME RD	SHAWNEE	POTTAWATOMIE	\$437,089	10
TRAVIS ENTERPRISES INC	1984	KICKATOOSPUR	SHAWNEE	POTTAWATOMIE	\$407,533	12
SHAWNEE HOTELS, INC.	2002	4909 N UNION AVE	SHAWNEE	POTTAWATOMIE	\$356,782	4
LAXMI INC	1989	5107 N HARRISON ST	SHAWNEE	POTTAWATOMIE	\$291,979	53
THUNDERBIRD CASINO	2018	2051 GORDON COOPER DR	SHAWNEE	POTTAWATOMIE	\$193,672	14
KICKAPOO MOTEL	1999	4118 N MARKET AVE	SHAWNEE	POTTAWATOMIE	\$173,626	2
COLONIAL INN	1986	4700 N HARRISON ST	SHAWNEE	POTTAWATOMIE	\$172,560	3
COMFORT INN	2010	5400 ENTERPRISE CT	SHAWNEE	POTTAWATOMIE	\$114,864	10
BUDGET INN	2010	14204 HIGHWAY 177	SHAWNEE	POTTAWATOMIE	\$103,280	12
SUPER 8	2005	4900 N HARRISON ST	SHAWNEE	POTTAWATOMIE	\$102,538	1
SHAWNEE LODGING, LLC	2003	623 KICKAPOO SPUR ST	SHAWNEE	POTTAWATOMIE	\$94,347	14

Source: Mergent, Inc. (2022)

- What are the observations?
- What are the variables?

Types of Data

- **Quantitative Data:** can be measured using numbers; answers questions like “how many?” and “how often?”
 - **Continuous:** fractions make sense (1.5 hot dogs, 3.14 slices of pie)
 - **Discrete:** fractions don't make sense (1.5 dogs, 3.14 employees working at the pie stand)
- **Qualitative Data:** describes the attributes or properties that an object possesses
 - **Nominal:** uses names or categories (public or private, male or female)
 - **Ordinal:** uses a ranking or ordering system (A+/B/C- or always/sometimes/never)

The Source Matters!

- **Primary Data:** data you collected yourself (usually never gathered before)
- **Secondary Data:** data you got from someone else

If you have to ask, it's secondary data!



Data Sources

The US Census Bureau

- The US Census Bureau is responsible for producing data about the American people and economy
- Original creator of administrative sub-units used to this day: “delineation of small geographic entities based on population, topography, and housing characteristics”
- Chiefly responsible for conducting the decennial census; also has other, more frequent operations under its umbrella (next two slides)



American Community Survey

- Sent to approximately 295,000 addresses monthly (3.5 million per year)
- Gathers information previously contained only in the long form of the decennial census, such as: ancestry, educational attainment, income, language proficiency, migration, disability, employment, and housing characteristics.
- **1-year estimates** are available for areas with a population of at least 65,000 people.
- **5-year estimates** are available for areas down to the block group scale, on the order of 600 to 3000 people
- ACS uses **sampling** methods to collect data; in smaller geographies, the error bands can be much larger

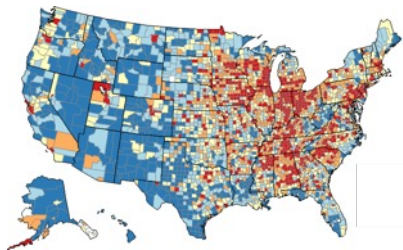


AMERICAN
COMMUNITY
SURVEY

U.S. CENSUS BUREAU

County Business Patterns

- An annual series that provides sub-national economic data by industry
- Includes the number of establishments, employment, first quarter payroll, and annual payroll.
- **Researchers** use the data for studying the economic activity of small areas and for analyzing economic changes over time.
- **Businesses** use the data for analyzing market potential, measuring the effectiveness of sales and advertising programs, setting sales quotas, and developing budgets.
- **Government** agencies use the data for administration and planning.



The Bureau of Labor Statistics (BLS)

- The principal fact-finding agency for the U.S. government in the broad field of labor economics and statistics
- “Collects, processes, analyzes, and disseminates essential statistical data to the American public, the U.S. Congress, other Federal agencies, State and local governments, business, and labor representatives.”
- Produces the following (among others):
 - U.S. Consumer Price Index
 - Current Population Survey
 - The American Time Use Survey
 - Current Employment Statistics
 - Quarterly Census of Employment and Wages (QCEW)
 - National Compensation Survey



The Bureau of Economic Analysis (BEA)

- Government agency that provides official macroeconomic and industry statistics
- Most notably reports about the gross domestic product (GDP) of the United States and its various units—states, cities, counties and MSAs.
- Provides information about personal income, corporate profits, and government spending
- The **industry economic accounts**—presented both in an input-output framework and as annual output by each industry—provide a detailed view of the interrelationships between U.S. producers and users and the contribution to production across industries
- The **regional economic accounts** provide information about the geographic distribution of U.S. economic activity and growth.



- Short for “Federal Reserve Economic Data”
- Contains more than 500,000 economic time series from 87 sources; widely reported in the media and play a key role in financial markets
- FRED covers: banking, business/fiscal, consumer price indexes, employment and population, exchange rates, gross domestic product, interest rates, monetary aggregates, producer price indexes, reserves and monetary base, U.S. trade and international transactions, and U.S. financial data.



Data from Other Sources

- Depending on what you are looking for, data can come from a wide variety of sources
- It is important to be aware of:
 - **what** the data point is measuring,
 - **how** data were collected (sampling strategies, etc.) and
 - **where** (i.e., the level & type of geography) the data are from
- It is equally important to be aware of **why** data were collected in the first place (i.e., motive of data-collecting agency or interest group)



Working with Data

- **How can we use data to answer a specific question about the local economy?**
 - First thing: understand the tool we're working with
 - Second thing: use context to look for patterns or trends

Understanding the Data

Consideration #1: TIME

Important questions to consider:

- When was this data collected?
- What time period does this data correspond with?
- Do we expect the data to meaningfully vary over time? How *much* time?
- When will this data be collected (and publicly released) next?

Understanding the Data

Consideration #2: GEOGRAPHY

Important questions to consider:

- What places does this data correspond with?
- Was this data only collected in one place, or is it available everywhere?
- Do we expect the data to meaningfully vary across places?
- Was everyone in the “geographic unit” accounted for?
- Are all places comparable? (example: Tulsa County and Roger Mills County)

Geography

Units of Geography



The Soda Fountain Eatery Restaurant

- **Address:** 108 W Broadway St, Anadarko, OK 73005
- **Coordinates:** 35.07474998749038, -98.244082643422
- **County FIPS:** 40015 – Caddo County
- **Full FIPS:** 400151621024040 – Census Tract 1621.02, Block 4040
- **Wider Context:** Oklahoma, USA, North America, Earth
- What **other geographic units** might this parcel of land belong to?

Geography is Sometimes Inconsistent!



BEA Regions



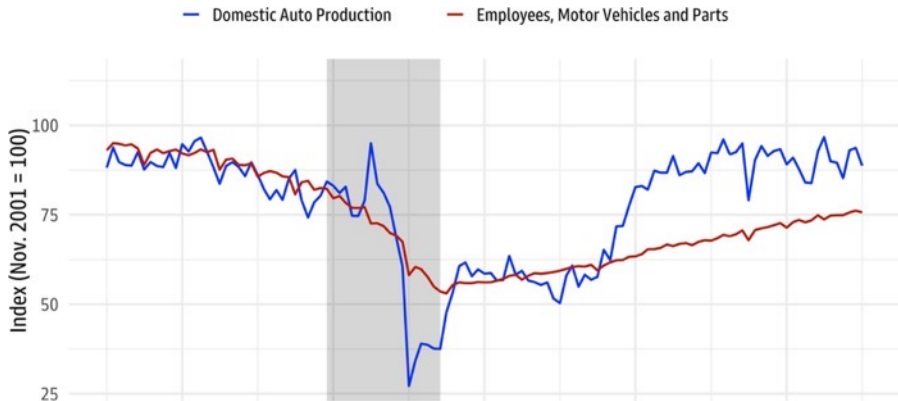
Census Regions & Divisions

Actually Using Data

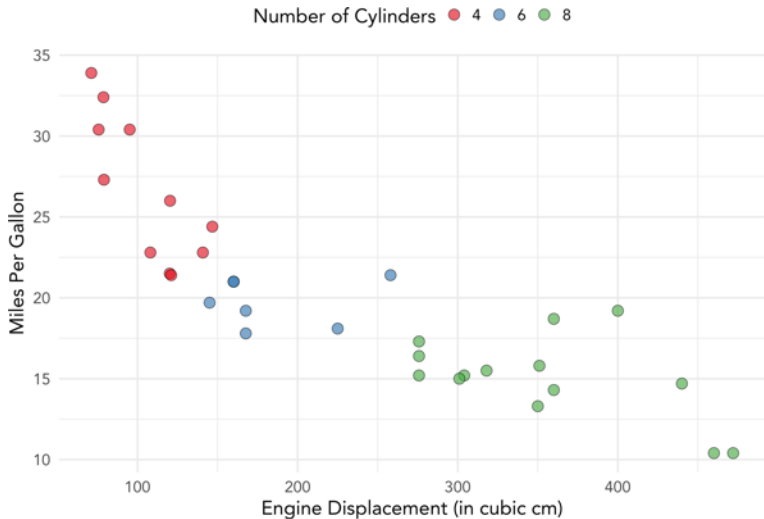
What Do We Actually Do With Data?

- There are many additional considerations beyond time and geography, but this is not a statistics class! Let's just stick with those two.
- Once we are aware of the timeframe and geography represented, how can we use context to look for patterns or trends?
 - Data about **your place** can give us specific information about observations that apply to our county.
 - Data about **all places** can give us an idea of how our county stacks up with others.
 - Data that includes **many time periods** can give us a sense of the direction things are going
- **Visualizing** data is often how we actually tell the story.

Visualization #1: Comparison



Visualization #2: Relationships



Putting Data in its Proper Context

Data Isn't Just Numbers and Spreadsheets

- My role is to help folks understand data, but I often cannot interpret it!
- You might have learned some new terms today, but most of you are already experts at using “data” to tell a story about your county.
- **Your data sources:** experience + social embeddedness
- As county educators, you likely possess more data about your county than anyone else alive!
- The trick is to mesh your “anec-data” with existing secondary data: use your inside knowledge to support what the chart/table/map says.

Recap

- Data refers to **pieces of information** that inform us about the world and help us answer questions.
- There are many ways to categorize data. **Understanding those categories** helps you to know your data.
- **Know your data!** Primary data collection is hard & time-intensive, so it helps to become familiar with secondary data.
- Geography plays a huge role in how data is collected and what story the data tells. **Be aware of geography** when using data.
- A single data point is often useless if we cannot put it in context with other data points!

End of Part 1

Questions?

Part 2: Demonstration and Discussion

A Quick Demonstration

- BLS Local Area Unemployment Statistics – [LINK](#)
- Census Search Bar *"median family income in McCurtain County Oklahoma in 2010"* – [LINK](#)
- BEA employment by county – [LINK](#)
- Excel demo: dispensaries and crime/poverty

Discussion Questions

- How much does the “**searching**” process scare/overwhelm you? How much does the “**manipulation/visualization**” process scare/overwhelm you?
- What **kinds of questions** do you find yourselves wishing you had better data to answer? What kinds of questions do you think we are already well equipped to answer? What questions are unanswerable, but we try anyway?
- Who is your “first phone call” when you need questions answered with data?

Discussion Questions, continued

- Do you feel like the people in your county who do community/economic development **full-time** (econ development authority, chamber, jobs board) are appropriately using data in their position?
- What have been your experiences with **primary data** collection? Have you felt confident about your analysis (or whatever you did with the data)?
- How much do you **trust** governmental administrative data (Census, BEA, BLS, etc.) versus your “gut” when thinking about your local economy? Do we ever put too much of an emphasis on using quantitative data?

Questions and Feedback

- Specific questions about something we covered but was unclear?
- Specific questions about something you *thought* we'd be covering but didn't?
- General questions about using data for rural community development?
- Suggestions for a follow-up or continuation from this conversation?

Thank You!

Andrew J. Van Leuven

andrew.vanleuven@okstate.edu



OKLAHOMA STATE
UNIVERSITY