By the Numbers:
Finding and Using Data from the Internet

Julie N. Zimmerman, Professor of Rural Sociology and Cameron McAlister, Graduate Assistant.

Finding data is only a click away.

This publication provides tips to help in finding and using data online.

Part of having 21st century skills for a 21st century world means being able to find and use data that are available online. Since secondary data are so easily available, knowing how to find and use these data are now an expected part of our work.

How do I begin?

With so much data available online, it is easy to quickly feel overwhelmed. Oftentimes, when we go looking for data, it is for a particular project, program, or purpose in mind. When it comes to finding data, **let that purpose be your guide** by asking yourself:

What do I want to know?

Why do I want to know it?

Answering these two questions can help you identify numbers that are not just interesting, but are useful for your specific needs. Another way of focusing your search is to ask:

How will knowing this number help in making a decision?

Answering these two questions can help you distinguish between numbers that are useful and numbers that are interesting but may not be the most useful ones for your particular needs.

When it comes to finding data online, resist the urge to let your search engine find it for you. It can produce a dizzying long list and it might not be clear how old the data are or where the numbers came from. Instead, use your search engine to find the website, not the number.

In general, you will encounter two different types of data websites. I call them original source websites and compilation websites.

Original Source Websites

Original source websites are websites that are generated by the organization or agency that **produced the data in the first place**. These can contain a specific data source, or they can be data hubs where you can access several data sources that are all produced by the same agency. A good example of this is the Census Bureau's website https://data.census.gov.

The good news about original source data websites is that they will contain the **most up-to-date data** and they have all of the geographies available for that data (states, counties, etc).

Compilation Websites

The second kind of website you will encounter are compilation websites. These websites contain data from **multiple data sources** that are compiled **in one place**. Most are topic-specific websites like the national Kids Count Data Center (http://datacenter.kidscount.org). The good news about topic specific compilation websites is that you can find data from many different sources in one place.

Kentucky: By The Numbers
http://kybtn.ca.uky.edu

Updated: Sept. 2020

When it comes to using compilation websites remember that it takes time to pull all of the data together. So, compilations or profiles may **not** have the most up-to-date data available.

(If you think there may be more recent data available, check the list of sources they used and then go to the original source website.)

> Links to commonly used websites
> with county level data are available on the
> Kentucky: By The Numbers website
>
> **https://kybtn.ca.uky.edu/**
>
> Just click on "Links to Additional Data Sites."

Be careful where you "click"

With websites and web technology constantly changing, data sites have become a lot more complicated. Before, the main focus was providing access to the data. Now, in addition to data, those same websites can now offer a variety of online options such as data visualization, maps, or even analysis tools. In addition, no two websites work in the same way.

Here are a few tips to help in navigating data websites.

Do not rely on the "back button" on your web browser.

Once you are inside a data website, if it offers you a "back" button as part of the website, be sure to use it. On some websites if you use your browser's back button to go back a step, it will actually kick you out of the website entirely. If this happens, it means starting all over again.

What looks like open space might be a hidden link.

Links to other features of a website are not always clear. Online tools like mapping can be useful but if you don't have good internet access, they can take a while to load. Because links can be hidden, even if something looks like dead space, it may not be and you can suddenly find yourself in a part of the website you did not intend to use.

Sometimes you still have to click "go."

It may sound silly, but on some websites, after you make a selection, it will automatically take you to that page. On others, the website won't do anything until you actually click "go" (or "search"). So, if the website you are using doesn't seem to accept your selection, it may be that you just need to click "go."

When you start to use a data website for the first time, take a look around. It can save you time in the end.

I found my data, now what do I do?

Locating data for your county is just the beginning and knowing how much you have of something is only one piece of information. For example, finding out that in 2018 your county had 3,235 teenagers may be informative, it does not tell you if that is a large part of all the total population in your county.

With just a few calculations, you can learn even more. In case it's been a while, here are some helpful reminders.

Distributions

Oftentimes we may want to know what *the percent of something* is. To use the number of teenagers in Calloway County as an example, the most recent data come from the Census Bureau's population estimates. For Calloway County in 2018, there were 3,235 teens ages 15-19 years old.

To calculate a distribution, we would compare this number to the total population. For our example of Calloway County, the 2018 estimate for the total population was 39,135 people.

From here, the math is simple. To calculate a percent, you take the number of teens ages 15-19 years old and divide by the total population. It looks like this:

> The number ages 15-19 in your county
>
> **divided by**
>
> the total number of people your county
>
> **multiply the result by 100**
>
> equals the % of the total population in your county who are 15-19 years old

In our Calloway County example, according to the 2018 population estimates, 8.3 % of all people were teenagers ages 15-19 years old.

Aggregate or Numerical Change

If you want to know if a number has changed, there are two ways of doing this: aggregate change and percent change.

Figuring out the *aggregate or numerical change* is straightforward. It simply means the total amount that

something has increased or decreased between two points in time.

Let's look at our example again.

According to the 2018 population estimates, we know that there were 3,235 teenagers ages 15-19 years old in Calloway County. But, at the 2010 Census, there were 3,307 teenagers in that age group.

To find the aggregate change, all you do is take the number of teenagers in 2018 and subtract the number of teenagers in the earlier year of 2010. In this case, the number of teenagers in Calloway County ages 15-19 decreased by 72 during the 8-year period.

Percent Change

While it may be useful to know that the number of teenagers ages 15-19 years old decreased by 72, you may want to know more. How quickly is the change happening? How does change in our county compare with another county, region, or the state? To do this, you will want to figure out the rate or *percent change*.

To find the percent change for the number of youth ages 15-19 for our example county, it would look like this:

> the number ages 15-19 in **2018**
> *minus*
> the number ages 15-19 in **2010**
> *then divide by*
> the number ages 15-19 in **2018**

An easy way of remembering this is:

> **(New-Old) / Old**

The result should be a number that is **less that one**. This is the **rate** of change and it can be positive or negative.

To turn this into the percent change, simply **multiply by 100**. In the case of Calloway County, between 2010 and 2018 there was a 2.2 % decrease in the number of teenagers ages 15-19 years old.

Be aware that when it comes to understanding change over time, the dates you pick can affect your results. If one or both of the points in time are unusually high or

unusually low, the result can make it seem like there has been more (or less) change than is usually the case.

What to watch out for…

Since we have gotten used to having so much data so easily available online, we can have high expectations that we will find exactly what we need. But that might not be the case.

Here are some things to keep in mind.

- All data are collected with a particular purpose in mind. That original purpose can impact what you find. So, regardless of how we might want to use the data, what is counted and how it is counted depends on why the data were collected in the first place.

- When you are selecting data, always make sure you understand what the data are counting and how they are being counting. Just because data sources might use the same words that we do in our everyday language, it does not mean that they are always referring to the same thing.

- Even if two different sources use the same terms, they might not be using the same definitions. In both cases, the numbers are accurate. The numbers are different because the definitions, or criteria, were different.

- Finally, there can still be times when no data are available. It could that the data were not disclosed because the person or responding unit was potentially identifiable. It could also be that the data just simply do not exist.

> *To learn more about how data are counted, see:*
> *What Counts?*
> *Understanding Data on Jobs, Income, and Unemployment*
> *A copy is available is on the*
> *Kentucky: By The Numbers website*
>
> **https://kybtn.ca.uky.edu/**

Online publicly available data are an important source of data. However, they are **only one kind of data**.
Depending upon your needs, other types of data (such as conducting a survey or focus groups) or other sources of data (such as local organizations) may be more useful.

Useful Websites with County Level Data
(Updated: September 2020)

Kentucky: By The Numbers
"Kentucky: By the Numbers Data Series" as well as publications and resources for finding and using data.

https://kybtn.ca.uky.edu/

Building Strong Families County Profiles
Themes vary by year. Produced by Family and Consumer Sciences Extension. Recent themes include housing, health, and food.

http://fcs-hes.ca.uky.edu/content/bsf

Kentucky Center for Statistics (KY STATS)
Kentucky data include career/college readiness, graduation rates, early childhood profiles and post-secondary attendance.

https://kcews.ky.gov/

Labor Market Profiles
KYLMI – Kentucky Labor Market Information. Labor force data, wage data, industry data, unemployment rates, etc.

https://kystats.ky.gov/KYLMI

Kentucky Cabinet for Economic Development
Community data, industrial site and building information for cities and counties.

http://selectkentucky.com/

Kentucky Health Facts
County data profiles contain many health and social indicators.

http://www.kentuckyhealthfacts.org/

County Health Rankings
Provides maps, data, and rankings on health outcomes, factors, and additional health related measures.

https://www.countyhealthrankings.org/

CDC Diabetes Atlas
Contains data on diabetes and related risk factors.

https://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html

Kentucky State Data Center
2010 Census, American Community Survey, population projections, and data for Area Development Districts.

http://ksdc.louisville.edu/

data.census.gov
Access to Census Bureau data including the American Community Survey, annual Population Estimates, County Business Patterns, Economic Census and Non-Employer Statistics.

https://data.census.gov/

Kids Count Data Center
Data on indicators of youth well-being including economic, education, and social data.

http://datacenter.kidscount.org/

Official Unemployment Rates
The Bureau of Labor Statistics produces monthly and annual unemployment rates.

http://stats.bls.gov/lau/

Official Poverty Estimates
County estimates of individuals living in poverty, children in poverty, and median household income.

https://www.census.gov/programs-surveys/saipe.html

Census of Agriculture
Click on "Find current data by…" to access pdf files of tables.

https://www.nass.usda.gov/AgCensus/

Food Environment Atlas
Mapping website with county data including grocery store proximity, food and nutrition assistance programs, and community characteristics.

http://www.ers.usda.gov/data-products/food-environment-atlas

Food Access Research Atlas
Mapping website with food access indicators for low-income and other census tracts using different measures of supermarket accessibility.

https://www.ers.usda.gov/data-products/food-access-research-atlas/

Community and Economic Development Initiative of Kentucky (University of Kentucky)
Economic development data by theme.

http://cedik.ca.uky.edu/CountyDataProfiles

Employment and Income
Data from the U.S. Bureau of Economic Analysis.

https://www.bea.gov/data/economic-accounts/regional

Kentucky
By The Numbers

http://kybtn.ca.uky.edu