

By the Numbers:

Finding and Using Data from the Internet



Resources for accessing data in real time
<http://www.ca.uky.edu/snarl>

Julie N. Zimmerman, Professor, Rural Sociology. (Revised: Feb 2012, Updated June 2013.)

It used to be that finding data meant going to the library.

Today, many agencies and organizations post data directly to the web.

Because of this, accessing data on your county is now only a click away.

While the focus in this publication is on secondary data, it is also important to remember that they are only one kind of data available.

Other kinds of information (such as conducting a survey or focus groups) or other sources of data (such as local organizations) may be more useful for your particular purposes.

This publication provides help in finding data online, help in figuring out what to do with it, and tips on what to watch out for.

Updated: June 2013

How do I begin?

With so much data available online, it is easy to quickly feel overwhelmed. Often times, when we go looking for data, it is for a particular project, group, or purpose in mind.

When it comes to finding data, the first step is to let that purpose be your guide and ask yourself 2 key questions:

What do I want to know?

Why do I want to know it?

Answering these two questions can help you distinguish between numbers that are useful and numbers that may be interesting, but are not necessarily useful ones for your particular needs. Another way of focusing your search is to ask:

How does knowing this number help us make a decision?

Different types of websites

When it comes to finding data online, you will generally encounter two different types of websites: 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. Most often, we think of these in terms of government agencies (such as the U.S. Census Bureau) because they produce a lot of the data we commonly use. But, other agencies and groups may also produce their own data.

The good news about original source data websites is that they tend to contain a lot of data, they have the most up-to-date data, and they can have the most levels of geography (states, nation, counties, cities, etc.).

The bad news is that because they contain so much data, they can be more challenging to navigate. Still, once you get the hang of it, original source websites are a great resource.

Compilation Websites

The second kind of website you will encounter are *compilation websites*. These websites contain data from many different sources that the organization compiled together in one place and usually for a specific purpose.

The Kids Count Data Center website is a good example.

The good news about compilation websites is that you can often find topic-specific data from many different sources in one place.

Compilation websites can also be easier to navigate. Some are even designed as static websites where you click directly onto a page with your data.

The bad news is that since compilations must be gathered by someone - and that takes time - these websites **may not have** the most up-to-date data.

(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.)

At the end of this publication is a list of websites with commonly used data.

The URLs are hyperlinked on the "Kentucky: By The Numbers" website (<http://www.ca.uky.edu/snarl>).

Just click on "Links to Additional Data Sites."

Useful Websites

To find data for Kentucky, you can either go to websites that focus on Kentucky or national websites that have data for the entire nation.

The "[Kentucky: By the Numbers Data Series](#)" provides access to commonly used county-level data. Developed through Cooperative Extension to support local programming and decision making, issues contain county level data on a number of social, demographic, and economic themes. Some issues also have easy to print maps. The data series is produced on an occasional basis as part of the Kentucky: By the Numbers program. You can find it on the "Kentucky: By the Numbers" website.

<http://www.ca.uky.edu/snarl>

Another useful website for Kentucky data is the [Kentucky State Data Center](#) at the University of Louisville. (Every state has its own State Data Center.) Here you can find easy access to data from the 2010 Census, the American Community Survey, and other data for Kentucky.

<http://ksdc.louisville.edu/>

The U.S. Census Bureau provides many data that we commonly use. To access data including the American Community Survey, the 2010 or 2000 Census, or the most recent Annual Population Estimates, go to the [American Factfinder](#) website:

<http://factfinder2.census.gov>

Two national resources that have compiled data from different sources are [Quickfacts](#)

<http://quickfacts.census.gov/qfd/>

and the [Atlas of Rural and Small Town America](#)

<http://www.ers.usda.gov/data-products/atlas-of-rural-and-small-town-america.aspx>

These are all good one-stop places for some of the commonly used data for states and counties.

There are many more websites located on the last page of this publication. If you are still not sure where to find the data you are looking for, a good place to check is called [Fedstats](#). This website is intended to help people locate data produced by the federal government.

<http://www.fedstats.gov>

To help you get started, there are some easy to use graphic instructions for some of the commonly used websites.

You can find these on the "Kentucky: By The Numbers" website <http://www.ca.uky.edu/snarl>.

I found the data, now what do I do?

Locating data for your county is only just the beginning. With so much data available, it is easy to get lost amongst the numbers. For example, finding out that in 2011 your county had 2,603 jobs in retail may be informative, it does not tell you if that is a large part of all the jobs in your county.¹

Distributions

One way to deal with this is to figure out the *distribution*. In the case of our county, it would be the *percent of all jobs* that are in the retail sector. To do this you just need two numbers for the county; the number of jobs in retail and the total number of jobs. Both of these need to be for the same point in time or year.

To calculate the distribution, you take the number of jobs in retail and divide by the total number in jobs. It looks like this:

the number of jobs in retail in your county
divided by
 the total number of jobs in your county
equals
 the rate of retail jobs in your county

The result should be a number less than one. To make this a percent, simply take the rate and multiply by 100. That way you will know the percent of all jobs that are in retail.

¹ The data used in the examples are for Calloway county, Kentucky

For example, in 2011 our county had a total of 22,336 jobs and 2,603 of these were in retail. This means that 11.7% of all the jobs in our county are in retail.

Another way to look at data is to look at **change over time**. There are two ways of doing this: aggregate change and percent change.

Aggregate or Numerical Change

Figuring out the *aggregate or numerical change* simply means the total amount that something increased or decreased between two points in time. Let's look at our retail example again.

In our county, in 2011 there were 2,603 jobs in retail in your county but in 2009, there were 2,530 jobs in retail. To find the aggregate change, all you do is take the number of retail jobs in 2011 and subtract the number of retail jobs in the earlier year of 2009. In this case, our county has gained 73 jobs in retail over the 2 year period.

Percent Change

Another way to look at change over time is to look at the *percent change*. While it may be useful to know that 73 jobs in retail were gained in our county's economy, you may also want to know if that is a large change or not, especially if you want to compare it with another county, region, or the state. To find the percent change in retail jobs for our example it would look like this:

the number of jobs in 2011
minus
the number of jobs in 2009
then divide by
the number of jobs in 2009

Another way of remembering this is:

(New-Old) / Old

The result should be a number that is less than 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 our retail

example, our county had a 2.9% increase in the number of jobs in retail.

When it comes to understanding change over time, it is possible that the dates you select can affect your results. If one or both of the points in time you selected were unusually high or unusually low, this would make it seem that there has been more (or less) change than what was more typically the case.

What to watch out for...

Since we have gotten used to having so much data easily available online, we can have high expectations that we will find just what we need. But more often than you might expect, it still might not exactly be what you were looking for.

Here are some key questions to ask yourself before you use the data you have found.

Who collected it?

All data are collected with a particular purpose in mind. Regardless of the source, you should always be aware of who collected the data.

Why was it collected?

Since data are collected with a particular purpose in mind, this can also impact what you find. In other words, 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.

How is it measured?

When you are selecting data, always make sure you understand what the data are counting and how they are being counted. The numbers you find can be different depending upon what definition was used. All of the numbers are accurate, they are different because the definitions, or criteria, were different.

What can I live with?

Even with all of the data that are now available online, there will be times when the data we want don't exist, are not as up-to-date as we would like, or are not as reliable as we would like. There will also be times when data are undisclosed if the person or responding unit is identifiable.

Useful County-Level Data Websites

(Last Updated: June 2013)

Kentucky Sources

National Sources

Kentucky: By The Numbers

Info available: In addition to the "Kentucky: By the Numbers Data Series," you can also find links to useful publications and resources for using data from the American Community Survey.

<http://www.ca.uky.edu/snarl>

Kentucky State Data Center

Info available includes: 2010 Census, American Community Survey, population projections, and data for Area Development Districts.

<http://ksdc.louisville.edu/>

Workforce Kentucky

Info available: Labor force data, wage data, industry data, unemployment rates, etc.

<http://www.workforcekentucky.ky.gov/>

Kentucky Cabinet for Economic Development

Info available: Community data, industrial site and building information for cities and counties. Under "Data Center," click on "Community Profiles."

<http://www.thinkkentucky.com/>

County Health Profiles

Info available: County data profiles contain many health and social indicators.

<http://chfs.ky.gov/dph/epi/cohealthprofiles.htm>

or

<http://www.kentuckyhealthfacts.org/>

Community and Economic Development Initiative of Kentucky (University of Kentucky)

Info available: Economic development data by theme.

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

Building Strong Families County Profiles

Info available: Themes vary by year. Recent themes include health, economy, and food.

<http://www.ca.uky.edu/hes/?p=437>

American FactFinder

Info available includes: American Community Survey, 2010 Census, annual population estimates.

<http://factfinder2.census.gov>

Kids Count Data Center

Kentucky Kids Count Databook:

<http://kyyouth.org/kentucky-kids-count/>

Info available: Data profiles on indicators of child well-being including economic, education, and social data.

<http://datacenter.kidscount.org/>

Bureau of Economic Analysis

Info available: Data on employment and income by economic sector. To find county data, click on "State and Local Area Personal Income and Employment."

<http://www.bea.gov/regional/index.htm>

National Agricultural Statistics Service

Info available: For "Quick Stats," click on "Data and Statistics." For the current and previous Census of Agriculture, click on "Census."

<http://www.nass.usda.gov/>

Official Poverty Estimates

Info available: County estimates of individuals living in poverty, children in poverty, and median household income.

<http://www.census.gov//did/www/saipe/>

County Business Patterns

Info available: Data on businesses by type for states, counties, and zip codes.

<http://www.census.gov/econ/cbp/index.html>

Bureau of Labor Statistics

Info available: Monthly and annual unemployment rates.

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



Resources for accessing data in real time
<http://www.ca.uky.edu/snarl>

Source: "By the Numbers: Finding and Using Data from the Internet" June 2013. Julie N. Zimmerman. University of Kentucky.

<http://www.ca.uky.edu/snarl/KentuckyByTheNumbers/AdditDataSites.htm>