

Unsuppressed Labor Market Data Tells a Complete Story

FEBRUARY 21, 2019 BY JAMES HOWARD

Suppressed Data — The Pros and Cons

Data are a collection of numbers that, when put in the right hands, can be translated into stories. These stories can go on to guide local developers and researchers as they plan community and economic development, align education offerings, help companies develop talent strategies, and more.

On the other hand, when placed in the wrong hands, data can be used to provide details on companies that could leave them exposed to competitive threats. Consequently, government entities such as the Bureau of Labor Statistics or American Community Survey suppress data that could be linked to an individual person or a particular firm. The anonymizing process, called suppressions or non-disclosures, takes place when an employer has fewer than 10 employees or an industry has fewer than three firms in a given region. When this is the case, the state or government entities that collect the data replace the numbers with N/A or *.

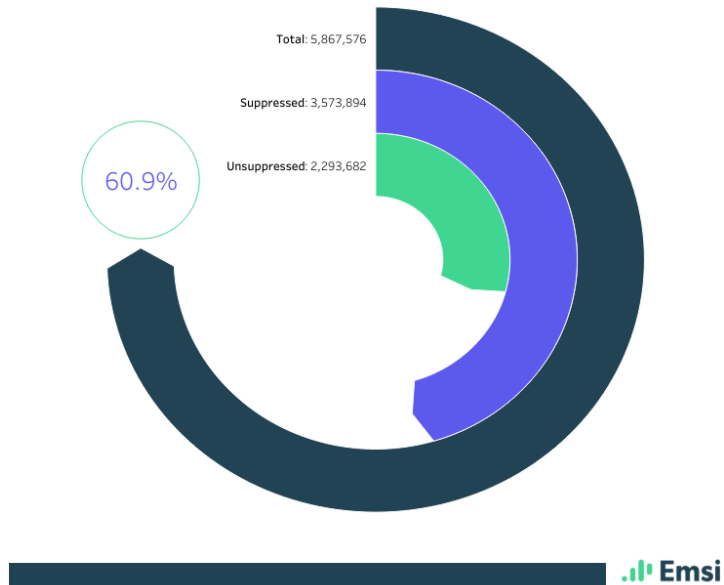
Suppressed data takes place in order to protect the privacy of both individuals and firms. In a world where privacy is extremely rare and data mining is being done by search engines, retailers, and anyone else trying to sell you something, there is a comfort in knowing the government has limits on the granularity of the data it is willing to release.

Now, for most people, this is good news. Why wouldn't it be? Well, there are researchers and economic developers that would like nothing more than to take all data, in its most granular state, and apply it to their research. In fact, there is a great amount of research that can be done with all the data that are suppressed. This research can lead to increased economic prosperity and a better understanding of local economies, as well as give insight into the lives and challenges of the everyday workforce.

The reason suppressed data becomes an issue for developers is because a large majority has gaps of unknown information, which can lead to inaccurate decision making. [Here](#) is an example when wrong assumptions were made with distorted year-over-year job change which led to faulty decision making. In these situations, Emsi's ability to unsuppress data becomes invaluable. In the 2018.4 data run, Emsi unsuppressed over 3.5 million of the nearly 5.9 million QCEW data points it received from the BLS. By those numbers, 61% of the data from QCEW alone was initially suppressed.

Unsuppressed Labor Market Data Tells a Complete Story

FEBRUARY 21, 2019 BY JAMES HOWARD

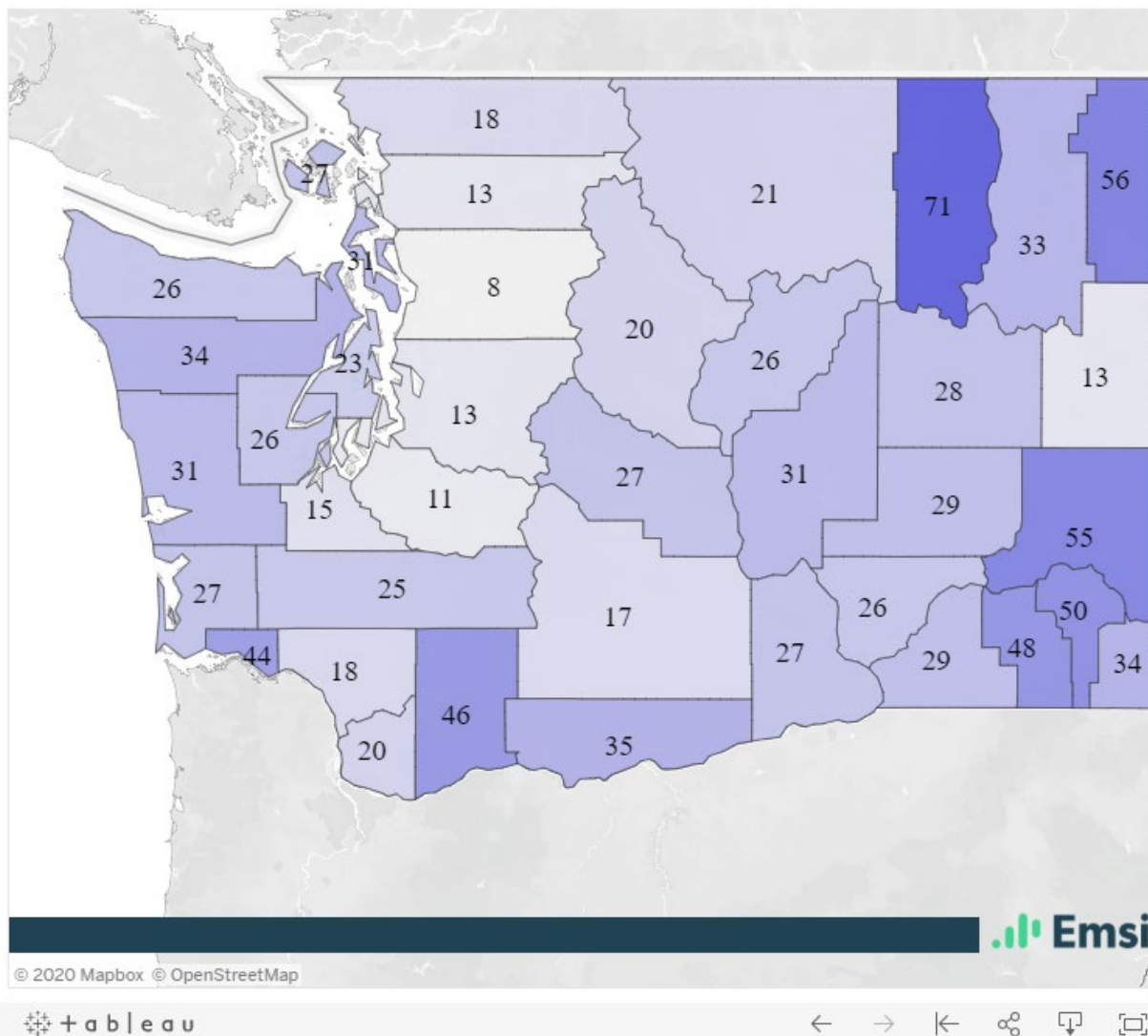


How do suppressions affect a region? Well, if we look at Washington, we can break down how many industries are being suppressed by county. The map shown below, illustrates the number of industries suppressed at the 3-digit NAICS (North American Industry Classification System) level. This an important figure because the NAICS system goes as deep as six digit levels. Therefore, the number of industries suppressed can be quantified even further, should a researcher choose to go more granular.

Unsuppressed Labor Market Data Tells a Complete Story

FEBRUARY 21, 2019 BY JAMES HOWARD

NUMBER OF INDUSTRIES SUPPRESSED AT THE 3-DIGIT NAICS LEVEL



NUMBER OF INDUSTRIES SUPPRESSED AT THE 3-DIGIT NAICS LEVEL

If we were to stop here, one would assume that suppression is only problematic in rural areas. This is not the case. There is a large amount of data that is suppressed at the city level. Usually starting at the 4-digit NAICS level, it is common to have the employment numbers and wages for entire industries suppressed. To illustrate this, let's zoom in on the Seattle-Tacoma-Bellevue, Washington, MSA where we see 86 of 310, or nearly 28%, 4-digit industries experiencing some sort of data suppression, mostly surrounding employment or wages. This goes to show that suppression is not just a rural issue, but urban as well.

Unsuppressed Labor Market Data Tells a Complete Story

FEBRUARY 21, 2019 BY JAMES HOWARD

The Important Uses of Unsuppressed Data

Using a proprietary algorithm, Emsi has unsuppressed QCEW and other census estimates down to the 6-digit NAICS level for over 15 years. And to take things even further, Emsi has applied an anti-suppression method at the ZIP code level, meaning data, when present, are available at the most granular industry and geographic level. With this level of granularity researchers are making decisions with data that they can feel confident in.

Emsi believes that there is a story to be told with data, and if the right data can be put into the right hands, growth and prosperity will result. To understand why Emsi unsuppresses data, one only needs to think of being able to tell a complete story. Stories are built on details, and to tell an accurate story, it is necessary to have access to the most detailed, reliable data available to make better data-driven decisions for policy makers.

To learn more about suppressions or to see Emsi's unsuppressed data for your own local industry and geographic level, contact James Howard at james.howard@economicmodeling.com. As a member of Emsi's consulting team, James can answer any data-related questions you might have.