

NWX-US DEPT OF COMMERCE (US)

Moderator: John Sperry

October 25, 2018

1:00 pm CT

Coordinator: Welcome and thank you for standing by. At this time all participants will be in a listen-only mode until the question-and-answer session of today's call. At that time you may press star 1 to ask a question from the phone line.

I would like to inform all parties that today's call is being recorded. If you have any objections, you may disconnect at this time. I would now like to turn the call over to your host, Mr. John Sperry. Thank you, sir, you may begin.

John Sperry: Thank you, operator and thank you everyone for attending today. Today's Webinar Data Literacy is the seventh in the economic program Webinar series. I'm John Sperry, a Survey Statistician from the wholesale indicator branch and I'll be your host today. Today's Webinar on data literacy is the final in the series on seven Webinars on economic programs.

On the right-hand side you can see a flyer of our previous 2018 economic data Webinar series. We began these Webinars on March 15 and have been holding one of these Webinars a month focused-on key topics. At the end of

today's presentation, we'll include a question-and-answer session. Please note today's Webinar will be recorded and posted for later reference.

You'll be able to find this Webinar's materials as well as the recorded presentation and transcript at the link listed on this slide at a later date. Until this presentation is posted at that Website, in the meantime you can access all of the previous Webinars archived there.

As data literacy is such a broad topic, you may be wondering what we will cover today. For starters we'll define what data literacy is. From there we will frame the importance of becoming data literate by defining common pitfalls of Census data users. Furthermore we will frame these trappings using examples from Census datasets.

In particular I will focus-on subject areas I have working experience in meaning I'll draw examples from where I currently work in the wholesale trade indicator branch and from international trade in the area I worked-in for multiple years.

That being said you may recognize a few other datasets from previous Webinars highlighted earlier in our series but before I get too far into the presentation, I'd like to tell you a little bit about major programs here at the Census Bureau. The United States Census Bureau is the federal government's largest statistical agency.

We conduct more than 130 surveys each year including our most famous, the decennial Census which many of you know is a once in a decade population and housing count of all 50 states and U.S. territories.

The American Community Survey which is the ongoing annual survey of the nation's population and also most recently the economic Census which is the official five-year measure of American businesses.

It's like some of you listening today have already received a form for the current 2017 economic Census which will be disseminated later in 2019. As a Bureau our mission is to serve as the nation's leading provider of quality data about its people and economy. In simpler terms as I've explained to my family, we collect and report data on the economy and its people.

As I just stated, Census collects and reports data on people and the economy. While not exclusive most of Census surveys can be divided into these two areas: demographic and economic. Today's focus on data literacy will be within the context of Census economic surveys. That being said, lessons here can also be applied to demographic surveys.

A majority of the economic Census surveys would fall somewhere on the pyramid image on the top right. As with anything exceptions to this graphic could be program areas like our survey of governments, an area that was covered extensively in a previous Webinar.

In any event, the pyramid provides a distinct way of showing how a majority of our economic surveys are related. At the very top of the pyramid are monthly and quarterly surveys, home to the economic indicators which is where I've worked my entire career.

These surveys provide your most current data but the least amount of detail. Annual surveys - the green portion - have larger samples and are thereby able to provide more level of detail but are released at a slower rate than the indicators.

Finally the blue area or economic Census conducts a survey that is the least timely but measures all businesses providing the most comprehensive data available on the economy. Unlike other Webinars in our series, I'm only going to highlight a couple of critical definitions before we jump into today's material.

The reason being because of the nature of today's discussion on data literacy, it will provide greater context knowing definitions and the context of a particular data user challenge, rather than defining everything at the front end and then calling back to it 15 minutes later into the relevant example. In my experience the Webinars I've learned the most from have framed themselves this way.

With that being said for anyone joining us for the first time in this Webinar series or new data users becoming familiarized with Census economic data, perhaps the most important economic data term to understand today is NAICS, otherwise known as the North American Industry Classification System.

Throughout today's Webinar I'm going to refer to NAICS frequently because it's our primary data classification system which we publish nearly all our business data under. Every establishment in the United States is assigned a single NAICS code based on their primary activity.

For example NAICS 42 covers wholesale trade. More specific codes like NAICS 4244 cover wholesale grocery trade. The final definitions I wanted to highlight are Title 13 and 26. These titles grant Census the ability to keep survey respondents' data such as yourself or a company's protected and confidential.

In particular Title 13 is authorizing legislation and provide requirements for protecting confidentiality of respondent data whereas Title 26 provides guidance on use and protection of federal tax information.

The best way I can describe this would be Title 13 protects confidentiality for things like your name or name of a reporting company or address and Title 26 is designed to protect tax information such a company's earnings or a person's income.

Both Titles provide protections for the data so the data is used for statistical purposes only. Please keep this in mind later when we get to the question-and-answer session. Because of these title protections, we cannot answer questions specifically about any companies that may or may not be on our survey.

While we're reviewing data literacy today, if you are wondering where you can find today's data, most of the Census Bureau's data is published to the American FactFinder. That being said, you can also find the most extensive amount of international trade data in USA Trade Online and more specifically indicator-level data in the Census indicator database.

Because I don't have the time today to show you how to access USA Trade Online or the Census indicator database, if you're a new data user feel free to watch an old but still relevant Webinar on navigating Census indicator databases at the Website listed on the slide.

Finally if you're looking for data to help you start or grow a business or even understand the business landscape for a region, check-out the Census Business

Builder so now that we've laid the groundwork for today, what is data literacy and why is it so important?

I like to think of data literacy as the ability to derive meaningful information from data, in other words to be able to properly communicate the data into information. For instance, if I said the wholesale indicator stock-to-sales ratio was 1.5 for a given month, that probably wouldn't tell you very much.

But if I explained the stock-to-sales ratio represented the amount of time in months it would take a wholesaler to clear their existing inventories at the rate of current sales that may help a data user to understand why it's an important piece of information so why is data literacy important?

Let's be honest, because identifying the appropriate data for your search can be as difficult as navigating different databases. By understanding what particular data represents the data user, it will help you understand if the data are appropriate for your ends which in turn will decrease the chances of data users from misusing or misinterpreting the data.

For instance if you joined us in our Census employment statistics Webinar way back in April, you know that one weakness in our employment data is that we don't have the distinction between full and part-time positions so if you interpreted our employment figures as solely representative of full-time positions, you could be wrong.

This is especially critical when we have multiple areas at the Census Bureau, the federal government and the private sector all creating economic data with their own respective methodology. The point of data literacy is to be familiarized enough with the data and their methodologies to be able to

understand why one dataset or data point may be suitable or more suitable for your ends.

For example if a data user requested international trade data on alcohol, there would be a lot of figures they could use. Should they use seasonally-adjusted or not seasonally-adjusted data? Data under what classification system? Aggregate figures or more detailed community breakouts?

Each one of these choices would lead to a different number but reflect international trade in alcohol. Ultimately while the determination would depend on whatever analysis the data user was trying to accomplish, if a data user does not understand the differences between these options, then there is a chance they maybe misuse the data.

With that example in mind, our goal today is to cover common traps data users may fall into so that the next time they need data, they can make a more-informed decision when using Census data. That being said, the challenges we will cover today won't cover every potential pitfall users may have with selecting Census datasets.

Normally we'd be able to go into a fully-nuanced discussed of each of these. We simply just don't have enough time for that especially on topics that could warrant their own Webinars themselves. Instead, we'll be providing a high-level review of what each one of these risks are, provide examples using Census data and explain instances where a data user may prefer one option over another.

Many of these challenges cut across multiple economic survey areas and statistics in general so if you have not encountered one of these problems yet, it's likely you may encounter them in your future. Hopefully by the end of

each pitfall's discussion, you will learn something new and be able to apply this knowledge to make more-informed decisions in your future data needs.

Let's get started with perhaps one of the greatest trappings any data user could have, balancing the three mission principles of relevancy, timeliness and accuracy. By timely we mean the amount of time between the end of the reference period and when statistics are released.

For our earliest economic surveys, the economic indicators are released to come-out as early as two weeks following the reference period. Though for our most-comprehensive release the economic Census it occurs every five years in years ending in 2 and 7 and can take more than a year before it's disseminated.

By relevance we mean the ability of the data to answer the questions being asked about the economy. Issues of relevance change as the economy changes, as policy concerns evolve and as economic theory advances.

In other words if you are looking to see monthly changes in an industry, the indicator data may be more relevant to your research while the five-year economic Census would not. By accuracy we mean how close the estimates come to measuring the concepts they are designed to measure.

For any data user the estimate is accurate when it captures the thing they wanted to capture. Think of this area as covering things like non-sampling error. For instance monthly indicator surveys are voluntary and tend to have lower survey response rates than the mandatory economic Census.

Lower response rates and other forms of non-sampling error such as data process or view errors can lead to less-accurate data. As this example

highlights, generally speaking a survey will be strong on two of the three of these principles.

Another way to look at the relationships of these three principles across economic data is this. There are three types of economic surveys that the U.S. Census Bureau conducts and then provides the results.

These are economic indicators as shaded in orange, annual surveys as shaded in gray and the economic Census as shaded in teal. This graphic provides a brief portrayal of each of these surveys. I will go into more detail on each of these in the upcoming slides.

The U.S. Census Bureau's economic indicators keep the world informed by providing the first official measures of the changing U.S. economy. The economic indicators monthly and quarterly data are timely, reliable and offer comprehensive measures of the U.S. economy.

The data provide measures of current economic activity that allow analysis of performance and informed business investment in policy decisions. In our economic indicator surveys, your data items are collected then for annual surveys or the economic Census their sample sizes are generally smaller than the annual surveys and the economic Census.

Generally speaking, the data are collected at the company or company reporting unit level and statistics are published at the national level. In order to provide more timely and cost-effective data on our nation's year-to-year economic activity, the U.S. Census Bureau collects information in our annual surveys.

Our annual surveys collect fewer data items than in the economic Census. Data are reported at the company level or establishment level, are a subsample of the economic Census. The statistics are published at the national or state level and also allow for year-to-year comparisons.

Last but not least the economic Census provides the nation with the most comprehensive and detailed facts about the structure of the U.S. economy. Conducted every five years, the economic Census profiles the U.S. economy from the national to the local level. It provides official measures of output for industries and geographic areas.

The economic Census is conducted in years ending in 2 and 7 and it collects data from business establishments of companies. It is the most-detailed information on the structure of the economy are provided including data by industry, data by geography, the inputs companies consume, the outputs companies produce and the markets companies serve.

Data products from the economic Census include industry series at the national level covering mining, construction, manufacturing, retail trade, wholesale trade and services, select geographic data at the sub-national level presented by the state, metropolitan and micropolitan area, county and ZIP Code, product lines, establishment and firm size, miscellaneous subjects including topics like class of customer, type of food service, end-of-year inventory and petroleum bulk storage.

Subjects presented depend on the industry. Data from the economic Census provides the framework for measures such as the Bureau of Economic Analysis or BEA's calculation of the gross domestic product GDP and national income and product accounts.

The data are used by government and businesses for planning and economic research. The economic Census provides benchmarks and sampling frame source of data for the Census Bureau's current economic programs.

The most recent economic Census data is available from 2012. We are currently collecting data for the 2017 economic Census. Now that we know the background for all the blocks, let's go ahead and apply a case.

To bring all of our relationships together, we'll highlight a hidden gem that wasn't covered in the previous Webinar, the wholesale trade industry, represented in our indicator surveys by the monthly wholesale trade survey or MWTS, our annual surveys by the annual wholesale trade survey or AWTS and the economic Census via the Census of wholesale trade.

Altogether these surveys cover companies with employment that are primarily engaged in merchant wholesale trade in the United States. This includes merchant wholesalers that typically take title of their goods they sell such as wholesale merchants or jobbers, drop shippers, exporters and importers.

To clarify wholesales aren't companies where individuals like you or me can walk into and make an everyday purchase. Wholesalers are a business that sells to retailers, contractors or other types of businesses including farms but not to the general public or at least in any significant amount.

A majority of wholesales sales are typically conducted to other businesses; in other words just because a company may have wholesale in their name, it doesn't make them a wholesaler. It's who they sell to that matters.

As you can see here, as the releases become more timely, the sample becomes smaller becomes smaller; however, as the data becomes less timely, it may

also be more detailed for a long-term analysis. Looking at this table, depending on a data user's interest there are differences in coverage, release dates and data products which could guide their decision-making process.

Starting on the left with the MWTS it's the smallest sample, collects the least amount of data and releases the most frequently. As noted in the bottom category, MWTS does not publish data for agents and brokers and does not include data for manufacturers, sales branches and offices or MSBOs in their figures.

If a data user didn't know this or tried comparing MWTS figures to total AWTS figures that were inclusive of all wholesale, they were be comparing apples to oranges. I'll get into that a little later when I cover the pitfalls on accounting for error.

In the middle is the annual wholesale survey or AWTS. As you can see it comes-out a little bit slower than the monthly counterpart because it draws from a larger sample and collects more data. It has the ability to include industry level wholesale data that covers agents and brokers as well as MSBOs.

Lastly we have the economic Census on the right. As its name Census implies, its sample encompasses almost all wholesale employer establishments of companies, has the most items collected, subsequently leading to the most data products.

However, this extra detail comes at a cost of timeliness, in particular the 2012 economic Census started with the release disseminating starting a year after the reference period on a flow basis. The relationship between indicator,

annual survey and economic Census is played-out for almost each of our indicators in annual surveys.

Because of this, it's important to consider your needs as a data user before you select a dataset. Do you need more timely industry-relevant or more industry-accurate? As I hopefully have illustrated with the wholesale trade sector, just because something is say more timely with the wholesale indicators doesn't mean it's more industry-accurate due to its data coverage.

With that being said, why don't we take a closer look into the pitfall of accounting for error but first, what exactly do we mean by accounting for error? With any sample statistic, sampling and non-sampling statistical error exists so why is this important? How will understanding error better make you more data-literate?

Well, because not accounting for error can lead to some very serious ramifications. In fact it could change the analysis of the data. For instance, at the time of reporting the seasonally-adjusted June 2018 wholesale indicator sales estimates that was released said the sales were down 0.1% from May with a margin of error of plus or minus 0.4%.

With the margin of error at 0.4%, this tells us that the change in the value of sales could - based on this measure of sampling error - have also been unchanged at zero or even up. As a side note, even this example masks another common possible source of non-sampling error, pitfall that I will not have the time to cover today, timing of data releases in revision.

By timing, I mean that a figure may change from when it's originally reported versus its revised figure when more processing time allows for later reporters

or more reporting issues to be corrected by analysts. As you can probably tell by now, hearing a statistic is easy but understanding one takes longer.

For our purposes today, we're going to cover error in two buckets, sampling and non-sampling. Furthermore given the time allotted today, we will only be able to keep discussion to a high level. That being said, we will take a closer look at just some of the types of non-sampling error.

So what are these types of error? How are they defined? The first error is sampling error. You can think of this error as the way a statistic is different from an actual value due to the survey sample being drawn from one of many possible samples opposed to the result that would be obtained from a complete Census conducted under the same conditions.

This error occurs because characteristics differ amongst sampling units in the population and only a subset of the population is measured in this sample survey. On the flip side, we have non-sampling error.

Non-sampling error encompasses all factors other than sampling error that contribute to the total error associated with an estimate. This error may also be present in Censuses and other non-survey programs.

Non-sampling error arises from many sources: inability to obtain information on all units in the sample, response errors, differences in the interpretation of the questions, mismatches between sampling units and reporting units, requested data and data available or accessible in respondent's records or with regard to reference periods, mistakes in coding or keying the data obtained and other errors of collection, response to coverage and processing.

One example of non-sampling error is measurement error. This is when a difference between the true value of the measurement and value obtained during the measurement process occurs. Another example of non-sampling error is coverage error.

This includes both under coverage and over coverage and is the error in an estimate that results from one, failure to include all units belonging to the target population or failure to include specified units in the conduct of the survey under coverage and two, inclusion of some units erroneously either because of a defective frame or because of inclusion of unspecified units or inclusion of specified units more than once in the actual survey also known as over coverage.

Our final example of non-sampling error is non-response error. This error is observed in estimates caused by differences between respondents and non-respondents. Now that we've reviewed these types of errors, how are they reflected in Census data?

Depending on the survey at the Census Bureau, sampling error will be reflected as the margin of error, a coefficient of variation, confidence intervals or standard error.

Like with the earlier example I highlighted on monthly wholesale trades, accounting for sampling error is very important because if the error makes it possible for the estimated value to swing directions or be zero, then any interpretations of the data assuming one direction without making note of error would be flawed.

So how might non-sampling error be reflected in Census Bureau data?
Starting with measurement error, a possible representation of this error occurs

sometimes on the monthly wholesale trade survey I work on. The monthly wholesale trade survey wants sales and inventories data reported to the nearest dollar.

In other words, no cents included and to be provided in straight dollars, not thousands or millions. Measurement error then could occur when a survey respondent reports data including cents or in thousands because that's how their accounting is completed. Rest assured we have systems in place to catch this type of error if it occurs.

Next we have coverage error, borrowing yet again from my survey area of wholesale trade. An example of this would occur if the sampling frame isn't accurate due to the ongoing nature of the business cycle. In other words it can be difficult for a survey to capture companies as they start-up a business and/or cease operations.

While both births and deaths of companies can be phased onto a survey gradually, the example highlights how in the interim of samples coverage of a target population may deteriorate and thereby leave the coverage error.

Finally we have non-response error. Again using wholesale trade as a backdrop, an example of this would occur when our survey respondents don't respond and in place of their response we impute their data based on similar-sized reporting units classified in the same industry or manually adjust the data based on public 10-Ks or quarterly earnings, etcetera.

This error then could be magnified if the non-responding characteristics are different than those of respondents. What exactly do I mean by that? Think of it this way. The reason a company may not respond is because they handle their accounting themselves, it's a small mom-and-pop business.

And they just don't have the time to respond versus a business that is of similar size but may be able to afford to outsource the accounting. At the risk of sounding like a broken record, we cannot stress enough how important it is for data users such as yourself to account for error when conducting an analysis.

As such the Census Bureau recommends that individuals using these estimates factor-in this information when assessing their analysis of these data as error could affect the conclusions drawn from the estimates. With that in mind, one challenge that's related to how statistical errors are represented is the classification system the data are presented under.

As I illustrated earlier with whole sale trade indicator data not including agents or brokers or MSBO data, what's included in a statistics classification can have a major impact on its interpretation. Whole sale trade like many of the other economic data programs at Census, publish their data under the North American Industry Classification System or NAICS system.

A parallel product classification system also known as the North American Product Classification System or NAPCS provides a product classification system that is not of industry origin that can be linked to the NAICS industry structure.

While most areas use NAICS data at Census, there are exceptions. For instance, the international trade area publishes their data under multiple classification systems so why don't we take a closer look at what makes those systems different for international trade? For international trade the first classification system is the Harmonized System also known as HS.

While HS codes share the same six-digit headings for imports and exports, their differences arise over the final four digits of their 10-digit codes. These divisions create the distinction for exports under the Schedule B classification and imports under the Harmonized Tariff System.

The second classification system we have covered a bit today, North American Industry Classification System or NAICS. As the name indicates, these data are industry-based meaning this data provides a broader picture of how a particular industry is performing.

It's important to note when it comes to international trade, the NAICS data is NAICS-based data off of the Harmonized System. This is because exporters and importers do not file NAICS codes but rather report their shipments under the Harmonized System. This means we have to convert from the Harmonized System into NAICS codes.

Next we have the end use classification system. This system managed by the Bureau of Economic Analysis bases classification on the principle use of the good instead of physical characteristics. Finally the Standard International Trade Classification or SITC was created by the United Nations to enable cross-country comparisons of data.

Let's apply these systems to a good. Let's consider international trade data on red potatoes. Under the detail of the six-digit level of NAICS, red potatoes would simply be classified as potatoes. Under the more-detailed Harmonized System, you could find data on more specific potatoes such as organic, Russet or red.

Under the end use classification data red potatoes would be rolled into a code on vegetables and under the SIT system red potatoes would be classified

under a more-detailed potatoes category so if you were a data user looking for specifically red potato data, then your best bet may be the most-detailed Harmonized System.

That being said, even a survey that uses just NAICS classification, categories aren't crystal clear. Consider wholesale trade. If you weren't familiar with the industries and red data on farm products on the wholesale indicator press release, you might think that farm products is inclusive of all things that may be grown on farms.

However, that's not actually the case if you looked to see what each industry covers. For instance if you look at the abridged examples of what's covered under their listed industries, you would see that roots fall under the grocery industry and that farm supplies such as fertilizer fall under miscellaneous non-durables.

This is why it's so important for data users to review what is covered under the classification for each industry because if a data user noted that lower farm product figures were influenced by an early frost and subsequent loss on citrus crops, they would be wrong as wholesale fruits data is not included in farm products.

The same could be said if you were trying to conduct a cross-sector analysis. If a data user wanted to conduct a cross-sector analysis say on petroleum using the retail international trade, manufacturing and wholesale sectors, they might think it's okay since they share a common classification system.

However, if a data user delves into what each sector's classification system covered, they would see that the retail trade petroleum representation through gas stations may be a bit diluted to inclusion of convenience stores or other

gasoline stations such as service stations for oil changes with their gas station data.

In other words, a data user couldn't say for sure if changes in retail gas station data were due to an underlying change in petroleum sales via gasoline or from concession sales via the gas station convenience stores or even oil changes from service stations.

Needless to say this doesn't mean that the retail trade data is bad or inherently flawed, it just suggests that it might not be a good fit for a cross-sector analysis with the other sectors, the other peers to exclusively cover petroleum.

Before I wrap-up my discussion on classification systems, there's one last thing to consider. As the economy evolves and changes so do classification systems. However, sometimes classification systems may not change as fast as the economy.

Take for instance the introduction of cellphones. Twenty years ago in international trade data they may have been included under a different category before becoming a commodity with a large-enough trade to warrant their own commodity.

That being said, when classification change are made, we release notes signifying how old data are changed as a result of the new classification so now that we've covered why it's important to take the time to understand what a statistic represents through classification and to consider the error alongside a figure.

The next pitfall we will cover is only appropriate for a few data areas; when to consider the impact of valuation of dollars. By valuation of dollars we mean

nominal versus real. For those of you on the phone who may not have heard these terms, it's an entry-level economics or statistics course, here's a fast refresher.

By nominal we mean dollars that are reflecting the current price of a good it's measuring. To clarify, nominal dollars do not include adjustments to reflect factors such as seasonality or inflation. Contrasting this are real dollars which can be understood as a constant dollar level adjusted for inflation or deflation and includes adjustments to reflect seasonality.

Most data areas here at Census publish real data as seasonally-adjusted figures; however, some areas also publish real data with chain dollar. For instance one such area that uses chain dollars in their published figures is the international trade indicator area which chains dollars using a 2012 as a reference year to two of their monthly exhibits.

So why don't we see what this chain dollar data looks like graph? Here we have a graph showing seasonally-adjusted U.S. export trades in goods on a monthly basis from January 2010 to July 2018. In blue we have real chain dollars based on 2012 dollars. In orange we have seasonally-adjusted data.

As you can see in the graph, the values coverage before 2012 and start separating near the end of 2014. What this graph shows is holding constant for 2012 prices that the level of trade has risen faster on a chain basis than has in unchained seasonally-adjusted data so why might one use chain dollars?

A data user may opt to use chain-weighted figures to account for changing valuations in the dollar if the values are actually changing and/or how much of it may be attributed to inflation or deflation.

That being said, data users should be cautious when comparing chain dollar data from one survey to another. For instance the Census Bureau constant dollar series will not match the Bureau of Economic Analysis' constant dollar series because of the underlying coverage differences between current dollar and national income and product accounts or NIPA data and Census data.

And you may be thinking, okay, well what if a data user would prefer to use nominal dollars and that'd be a good question and lead us into our final pitfall, seasonal adjustment. Some surveys at the Census Bureau particularly those conducted on a monthly or quarterly basis offer data on a non-seasonally-adjusted basis as well as a seasonally-adjusted basis.

That being said, just like the aforementioned international trade example, seasonal adjustment can vary from one survey to a survey so caution needs to be given when comparing cross-survey seasonally-adjusted data against each other but let's not get too far ahead. We haven't even defined what seasonal adjustment is yet.

Again for our purposes today, we're going to keep the explanation brief. As seasonal adjustment itself, it's very nuanced and we don't have the time to go down the rabbit hole. Defined, seasonal adjustment is the process of estimating and removing seasonal effects from a time series in order to better reveal certain non-seasonal features.

Examples of seasonal effects include a July drop in automobile production as factories retool for new models and increases in heating oil production during September in anticipation of the winter heating season.

It's important to note this is why comparing cross-survey seasonally-adjusted data against each other can be difficult. A drop in manufacturing may be

accounted-for earlier in their models than say a seasonal change in the wholesale or retail automobile sectors.

When applicable we also estimate and remove trading day effects and moving holiday effects such as Easter or Labor Day during the seasonal adjustment process. Trading-day effects are reoccurring effects related to the weekday composition of the month.

The seasonal adjustment factors were developed using the X-13ARIMA-SEATS software. The X-13ARIMA-SEATS program provides summary statistics to indicate the overall effect of these seasonal adjustments. For more information on the X-13ARIMA-SEATS follow the link on the slide.

On the flip seat of seasonally-adjusted data is non-seasonally-adjusted data. As the name suggests, these values are the survey estimates for seasonal adjustment. Let's look at this in a more digestible way to understand why a data user may prefer using seasonally-adjusted data over non-seasonally-adjusted and vice versa.

For starters, a data user may prefer to use non-seasonally-adjusted data because they may want to see the seasonality in data and/or for historic analysis because many areas have longer time series, think 1980s and earlier. Was data published to a non-seasonally-adjusted manner?

Furthermore because seasonal adjustment may vary from survey to survey, using non-seasonally-adjusted data may make it easier for using multiple data sources and analysis. On the flip side a data user may prefer to use seasonally-adjusted data if they want to view percentage changes or underlying trends in data.

Or as the name suggests, they want to see a smoother data that accounts for adjustments like trading days, seasonality, inflation or deflation and finally price changes. Perhaps the best way to see the differences between seasonally and non-seasonally-adjusted data is looking at them on a graph.

Seasonal adjustments can be difficult to understand without actually seeing it in action so here we have an example of an industry's seasonal effects and we can see how seasonal adjustment removes those effects to better reveal certain non-seasonal features.

In the graph we see both seasonally-adjusted sales and not adjusted monthly wholesale indicators of beer, wine and distilled alcoholic beverages over time. Every December alcohol sales jump to high levels and then corresponding drop to low levels following in January. We can see these fluctuations pretty clearly in the non-adjusted line on the graph.

However, make no mistake, seasonal adjustment isn't exactly straightforward as we see here on the graph. For instance, if we look at the same graph but for the wholesale trade I'll call inventories, different seasonal adjustment factors would have been applied.

That's because inventories are only collected at a specific point of time versus the sale year which are an aggregate figure over a period of time. Because of this difference, inventories are adjusted only for seasonal effects whereas sales are adjusted for seasonal variations like trading days and moving holidays like Easter.

It's important to state that seasonal adjustment can even vary between industries on a same survey. For instance consider back-to-school sales. We

would want to address bumps in industries that cover school supplies and not one such as lumber, automobiles or metals.

The seasonal adjustment process removes these seasonal effects from the time series and produces the adjusted line we see in the slide which is precisely why some data users might prefer using seasonally-adjusted data.

As we wrap things up today, we hope all of you that are listening learned a little bit more about data literacy, Census data and most importantly leave with a few important lessons. The Number 1 thing is if you use Census data and are worried about falling into one of these pitfalls discussed today or ones that we didn't have the chance to cover, don't hesitate to reach-out to us.

Call us, e-mail us, tweet us, whatever works for you because we are here to help. If you have questions, let us know because our data is only as good as if the data users understand what it represents.

That being said what we are more than happy to answer questions about are methodologies or classifications, please keep in mind due to Title 13 and 26 restrictions, we cannot get into the specifics as detailed as what companies may be driving industry statistics up or down? Think of it this way.

If we did disclose that information, why would a company want to respond then? Before I move-on to our question-and-answer session, I wanted to hammer-home the importance of responding to the economic Census. If you've already replied or aren't covered by the economic Census, please take this time to prepare any questions you have for us.

One thing I can't highlight enough is that for those of you that are respondents to the economic Census is that your response matters. As we've shown today, monthly, quarterly and annual surveys are useful to you in different ways.

If you're in need of timely information, you will want to use our monthly and quarterly statistics while our annual programs provide you with the ability to do time series analysis. Currently we are in the final steps of collecting data for the 2017 economic Census. As I highlighted earlier, even if you don't use the economic Census your response matters.

We use the economic Census as the baseline and benchmark for other surveys like the wholesale trade indicator I work in, thereby reinforcing the quality of the data for you. Your response to the economic Census helps facilitate the success of many of the surveys featured today and earlier throughout our series. Your response is valuable and is always kept confidential.

Because of this trust, the quality of our data is absolutely the gold standard. You can use our economic Census statistics whenever you need highly-detailed data. The 2017 economic Census is currently being conducted. The due date was June 12th but there is still time for companies to respond.

Responses are required by law to improve the dissemination of the data and provide cost savings for taxpayers responses available only online. Best of all the results are available to you free of charge so if one of you joining us today has yet to respond, here's a quick breakdown of the information we need for you to complete the survey.

As you can tell, the categories that you see here are broad, enabling this to be our most comprehensive business data program. The information you provide is invaluable. In fact, it's dissemination to data can touch your life directly as

research or indirectly as data used by policymakers as a guide to direct policies and programs within your community.

So what's the current status of the economic Census? Well, the 2017 economic Census was launched earlier this year in May 2018 and is currently moving into the response processing stage. That being said, there are a few changes for the 2017 Census.

First, responses are only online. Next, we do have a new reporting instrument for businesses with multiple locations and finally the data collection timeframe is a bit different. While response was due June 12th, if you haven't responded yet, we are more than happy to take your response.

Should you need help or have any questions about filing, help is available online and via the telephone. Additional information and resources are available at the link provided here.

If you didn't have the chance to join us previously for one of our prior Webinars or joined us a little bit late today or want to view this Webinar at a later date, it will be made available at the link on this slide following the presentation. As this is the last Webinar in our 2018 series, we are currently fielding ideas for a new Webinar series starting next year.

Are there things you'd like to be covered? Please let us know in the chat or by calling into the question-and-answer session. On that note, I'd like to thank everyone who joined us today and all of the time and energy of my colleagues who helped produce this Webinar.

Before I field questions, I would like to remind you all because of my experience is in international trade and wholesale trade, I will be most

prepared to answer questions for those areas. That being said, I have here with me multiple subject matter experts from different program areas that can assist with questions beyond my capacity.

Should we not have anyone here that can answer your questions, we will find the appropriate person for them and on that note, operator, let's take our first question.

Coordinator: Thank you. To ask a question from the phone lines, please press star 1. You will be prompted to unmute your phone and record your name. Again it's star 1 to ask a question. Okay, and we do have one question in queue, one moment, please.

John Sperry: Okay, great. Just to give you a heads-up, if I don't respond directly it's probably because I'm thinking about how to respond so or consulting with the other people in the room with me.

Coordinator: And we do have a question coming from (Karen). (Karen) your line is open.

(Karen): Okay, I just wanted to say thank you. This is a fantastic presentation. I've not attended one of yours before so this is a great one to start with and I'm really looking forward to seeing the ones in the future because these are very helpful.

My specific question is and I thought it was really great that you listed up-front all the places that the data can be found. Do you have a notification service where I could sign-up so whenever data gets refreshed that I would get a notification?

John Sperry: Yes, we do. It's called GOV Delivery.

(Karen): Okay.

John Sperry: Yes, I think it's on the footer of the census.gov homepage and I think like if there's a specific program area you're interested in, usually there's like I think like a button or a link usually on specific program areas. It probably will let you see if I write-down so just hold-on a second.

So I'm being told by one of my other colleagues here in the room that when you sign-up for GOV Delivery, you have the option of selecting what topics you're interested in so for instance if you're like interested in economic data, you'll get updates on economic data.

Yes, so it's right at the bottom of the Website, it's just a little box that says subscribe, sign-up for updates, please enter your contact information below and I believe you will get an e-mail pretty promptly and from there it'll guide your subscriptions.

(Karen): Okay, fantastic. Thank you.

John Sperry: No problem.

Coordinator: Thank you and there are no further questions in queue at this time.

John Sperry: So we're getting a few questions coming-in via the chat. It appears we have a question, this is the first one, from an (Erin) that says which Webinar would you suggest for learning to aggregate data regarding higher education, race and poverty?

Now I specifically work on the economic side of things so I'm guessing that would be a question more suitable for our demographic counterparts so if you'd like to shoot us an e-mail to ewd.outreach@census.gov what you're seeing up there on the screen of you could send an e-mail to it's easier, pio@census.gov.

That's the public information office. They can direct you to the appropriate resource. Then we have a couple of other questions here on the chat. This one is asking considering that the surveys are done every five years, can the economic Census data be used to determine trends?

Let me think about that one, hold on just a second. (Unintelligible) so I have here with me (Bill Bradus), the Chief of the wholesale indicator data area.

(Bill Bradus): So the question of economic Census being available for trends, so the economic Census is only five years but that's where the annuals and the monthly and quarterly indicators kind of fill-in the missing years for the trends.

So every five years kind of the economic Census is done and for trends we use the more-frequent programs for those data items to fill-in the trends.

John Sperry: All right, looks like we got a couple of other questions coming-in here. This one is from (Christa), will the 2017 economic Census data include ethnicity of business owners? That is a great question. I don't know if anyone here knows that. I would say that's probably a question then for PIO which that is another question I have here.

They were asking how to spell P-I-O and that is just simply the P as Peter, I as in Ice, O as on Ostrich so P-I-O [@census.gov](mailto:pio@census.gov) for the contact there. I'm not

seeing any other questions right now via the chat. Operator did we have any others come-in over the phone?

Coordinator: No, sir, none came-in over the phone lines.

John Sperry: So just getting another question here. It says do you have any Webinar series specific for maternal and child health and at this time I do not believe so. I'm just going to ask for a quick sweep if there's any other questions via chat or I guess calling into the operator and then if not, we'll shut things down for today.

Coordinator: As a quick reminder, it's star 1 to ask a question.

John Sperry: While we're waiting for any questions if anything else comes-up later, you like think after we end today oh geez, I really had a great question, you can always e-mail me directly on the slide here. It's john.w.sperry Sperry like the shoes @census.gov.

More than happy to answer them myself but if I don't know them, I will certainly find the appropriate person for you. With that being said, I am not seeing anything on the chat. Operator, do we have anything that is coming yet?

Coordinator: No, sir, nothing from the phone lines.

John Sperry: Okay, great. Well, then once again I'd like to thank everyone for joining us today and I'd like to thank certainly all my colleagues for helping me put this Webinar together. Thank you and have a nice day.

Coordinator: Thank you. That concludes today's call. Thank you for your participation.
You may disconnect at this time.

END