

NWX-US DEPT OF COMMERCE

Moderator: Deborah Rivera Nieves
August 28, 2019
1:00 pm CT

Coordinator: Thank you for standing by. At this time, all participants are in a listen-only mode. During the question-and-answer session if you would like to ask a question, you may press star 1 on your phone. Today's conference is being recorded. If you have any objections you may disconnect at this time. Now I would like to turn the meeting over to Mr. Chase Sawyer. You may begin.

Chase Sawyer: Thank you. Hello everyone and welcome to our webinar today. We're going to be discussing the Census Bureau's Planning Database and the ROAM Application. And this is going to help you be able to plan for surveys and for census - and the upcoming census.

I'm joined today by Kathleen Kephart and Suzanne McArdle. And they're going to help walk us through these applications and the Planning Database.

So to get started I just wanted to go ahead and give an overview of what we're going to be discussing. First, I'm going to give some basics about the American Community Survey and what we do with - for the ACS. We're also going to talk about the 2020 Census and how it relates to the ACS. We'll then go over the Planning Database and what a low response score is. And then

finally we'll have a live demo of the ROAM Application and let you know what you can expect going forward with the Planning Database and the newest - the next version of ROAM.

And so for - to give you some background on the American Community Survey the American Community Survey is the leading edge of survey design and continuous improvement and quality data.

We help to provide local statistics on a number of topics such as age, children, veteran status, income and commuting statistics among many other topics. And we do this by gathering survey responses from 3.5 million addresses each year.

And this helps us to inform federal agencies of \$675 billion worth of federal funding. Some examples of government agencies that use the ACS for funding include USDA, the Highway Association, among others.

And we have three annual releases for our major products. These releases are for our one year estimates which are for large populations, our one year supplemental estimates that are for smaller populations, and our five year estimates, which can be used to gather data for very small populations.

And so coming up here in the next month we will begin to release our one year estimates on September 26, 2019.

And then our one year supplemental estimates are usually released in October but this year we're actually going to have them released in February.

And then finally the five year estimates will begin being released on December 19th of 2019.

For more information about these releases as well as the release of public use micro data sample files you can go to the URL below to learn more about the data release schedule.

And so now I'm going to go ahead and turn the time over to Suzanne and she's going to talk to us a little bit about the 2020 Census.

Suzanne McArdle: Hi. Good afternoon. My name is Suzanne McArdle. And I have spent 11 years of my census career in the Geography Division of the U.S. Census Bureau. And I'm really excited to be here to discuss how the American Community Survey is actually informing some work for the 2020 Census so this is a really neat overlap and cross of different parts of the Census Bureau working together right here.

So with that being said, the goal of the 2020 Census is to count everyone once, only once and in the right place.

And at this point in the decade if we were sitting in a room full of Census Bureau employees you would've heard everybody join in to recite that. The mission of counting everyone once, only once and in the right place seems like an incredibly simple endeavor even though we know that the act of doing this is not actually simple. It's incredibly important and we take this very seriously.

And we have a number of different data sets and pools that are helping us reach this goal. You may have heard the phrase that the Decennial Census is the largest peace time mobilization undertaken by our federal government.

To throw numbers out there to speak to the sheer scale of this, and to put it into a little bit better perspective, the Census Bureau is actually estimating that by 2020 there will be more than 330 million people living in about 140 million housing units.

So the Census Bureau has been using the data set that Kathleen is going to be highlighting today, the Planning Database. And then ROAM, the Response Outreach Area Mapper which is the tool that I'll walk you through during a scary live demo in a little bit after Kathleen highlights the Planning Database for a number of different 2020 Census planning efforts.

The Census Bureau had been using this Planning Database and our new mapping tool to do a number of different things. These have been used to help recruit and hire staff. Think about where and with what language skills do we need people out in the field as well as informing the level of field resources and staff necessary to actually do the work.

Another quick example before we get into the weeds here is how we've been identifying hard to count areas using a metric called the low response score that you'll learn more about in just a moment to figure out where we're recruiting and hiring folks that can actually serve as trusted messengers with the 2020 Census and again making sure that they have the language skills that are necessary to operate within specific communities.

A number of trusted messengers are actually hired directly by the Census Bureau. They're called the partnership specialists. And as we sit here talking about this, a number of those partnership specialists are actually operating across the entire country to engage, educate and empower the communities in different ways. They're using these tools, the Planning Database and ROAM

to inform their outreach and communication plans within the communities that they're serving.

And because we're talking about a data set and an application that are both publicly available, they're able to take these tools to show tribal, state and local officials as well as community organizations and other partners and show them how to make use of these tools in their local communities.

The Census Bureau is really lucky in that we have a number of local people and groups operating in their own communities to help folks understand that being counted means being represented in our government. And it also means federal dollars. Both of those things impact all of us, people who work here at the Census Bureau and your neighbors.

So the main reason for all of this work is to try to minimize the nonresponse follow-up efforts. That's when someone doesn't respond to the 2020 Census and we need to re-contact them until they do. That's one of the costliest parts of the 2020 Census.

And so we all can make use of these data driven decision making power tools to help boost self-response to avoid as much of that costly follow-up activity as possible.

So I'm going to let Kathleen run away with helping you understand a little bit more about the American Community Survey and how that actually feeds into this Planning Database that I've been referring to.

Kathleen Kephart: All right, thank you Suzanne. My name is Kathleen. And I've been with the U.S. Census Bureau for ten years now. And I've kind of worked all over. So

right now I'm working in the Center for Behavioral Science Methods which actually produces the Planning Database.

I want to give you a little background and kind of add to what Chase said about the American Community Survey. The ACS is the premier source for detailed population and housing information as Chase mentioned.

And kind of historically it started in 2005. It became a critical element of our Reengineered Decennial Program. So we switched from gathering - instead of decennial, every ten year data, we switched to annual data on a small sample or subset of the U.S. population.

Now the sample is large enough that we can aggregate it over five years. We can produce estimates at very small levels of geography including block group and tract which I'm going to talk about today because that's what the Planning Database is at. Because we were collecting more information about households, language spoken, income on the ACS we were able to make the 2010 Census and the 2020 Census be a relatively short form with just a few questions such as name, sex, age, date of birth, race, ethnicity, relationship and tenure (rent versus own) for each U.S. household.

Okay. So the Planning Database and this is not - I don't mean to imply this is the only levels of ACS data of geography available. But the Planning Database is available at the tract and the block group level. And the ROAM is available at the tract level.

We sample - we take the ACS sample and we aggregate it over five years through these estimates at the smaller geographies.

Just a little background, a census tract is a subset of a county. It's contained within counties. There's more than 74,000 tracts in the U.S. And these generally have a population between 1200 and 8000 with the optimum size being about 4000 people in each tract.

A block group is a further subset of a census tract so there's multiple, typically multiple block groups in a tract. There are more than 220,000 block groups in the U.S. And they generally have a population between 600 and 3000.

So a little more about where you can find ACS estimates as well as the Planning Data.

Okay, so you may be wondering if we have these ACS estimates, why do we have a Planning Database?

And historically we've used the Planning Database internally to plan for field resources, to create tailored communication and partnership campaigns. So we work with nonprofits, churches, local community governments to engage people and help them understand why they need to respond to the 2020 Census.

1990 we had a hard to count score which was on an early internal Planning Database that wasn't publicly available. And households in each tract were assigned a score of how likely we thought they were to self-respond or we'd need to send out field interview in the nonresponse follow-up; the higher the score the harder to count. We then used this information to decide where to hire the most field representatives for nonresponse follow-up and for the resources for advertising, etcetera.

In 2010 we continued hiring and resource decisions using the Planning Database. And additional staff across the country used the score to identify areas that needed extra outreach, engagement and education. The 2020 Census hard to count score has been upgraded to what we call the low response score which I'll give you a little background later in the presentation. This is now free and publicly available to both internal and external stakeholders.

So what's in the Planning Database?

I like to call it the greatest hits of the ACS five year estimates and 2010 Census variables. So this is your gender, age, racial information, household income etcetera. It's kind of the most popular variables and the variables that we and field use to plan for census and survey operations.

So at a population person level you have gender, age, education, poverty level status; at a household level language spoken in the households, the relationship if it's a related family or a nonrelated household roommate situation, and the median income of the household.

At a housing unit level we have the ten year so the percentage or the number of rented versus owned units in a tract or block group and the number of units, housing units in a tract or block group.

For census operational variables on the file we have if it was a mail back area which means in 2010 we mailed a census form to them and they were expected to mail it back. There are other areas where we canvas in person. And those kind of - the characteristics tend to be different so this information can be useful for survey practitioners.

Also if it was a bilingual tract so in tracts where we predicted a high percentage of Spanish speakers we would mail a bilingual census form.

Another big thing is we calculate the percentage variables for each estimate. So the official ACS tract and block group estimates are count, you know we - there is 200 men. There's, you know, the actual count. Whereas on the Planning Database we calculate as a percentage of the total population this tract has 30% males versus females.

And we've actually calculated all of those for you for all the variables on the Planning Database, right on the Planning Database.

Like I said it's available at a tract and block group level. And it's available as a comma delimited file that you can just download and open right in Excel or through a census API. And the API, the 2019 Planning Database is currently available as a CSV. It will shortly be available as an API.

Okay, this data has a broad scope of potentially useful. It's useful for identifying areas with likely low survey response that you're going to need to target additional effort to get people to respond. Stratifying small areas on different characteristics, if you want to do age or income, any kind of variable that's available you can stratify.

Creating thematic map which is what the ROAM does for you. It takes the Planning Database data set and makes beautiful tract level maps as Suzanne will show you.

And you can enhance your reports with population metrics. So if you're doing a report on DC you can calculate different characteristics for DC as I will do later in the presentation.

Okay why use the Planning Database?

It's easier to download than the full ACS Summary files. I don't know if anyone's tried to download the files. But even on a high speed Internet connection it can take seven to ten minutes.

And then you really need to know what you're doing and there's a little bit more post-processing that needs to happen whereas the Planning Database you can click, download and open in Excel and within two minutes, well depending on your computer processor, you have all the variables right at your fingertips.

We've already kind of combed through all the potential variables on the ACS for the most popular and the ones people typically use and request. And we've mapped them to their 2010 counterparts for each tract and block group as well as creating the low response score.

We update the Planning Database and we only put the latest ACS data. And the ROAM Application creates tract level geographic maps that you can more easily access to figure out where these different areas are.

So the structure of the file, if you open it in Excel or any program, it's your choice, the first thing you'll notice is the geographic identifiers. On the block group file it's the block group ID. On the tract, the tract ID.

Then we place the demographics, socioeconomic and household level data. And we always try and keep the order of the variables consistent. We put the census estimate for that geography first followed by the ACS estimate and the

ACS margin of error which is very important to pay attention to when you're working with point estimates.

So for example you would see male Census 2010. That would be the count of males for that tract or block group followed by the point estimate from the latest ACS data followed by the margin of error around that point estimate for ACS.

Then we put the operational data including the mail return rate and the low response score followed by the percentages and margin of error for the percentages. And we try and keep these listed in the same order as the respective point estimates.

So now you'll see the percent males in the Census 2010, the percent males ACS in that block group or tract and then the margin of error for that percentage.

So here's some examples of cool ways to use the Planning Database. Let's say you wanted to know some information about Washington, D.C. You could take and aggregate the tract or block group estimates. You could also go to the ACS Web site but I'm using the Planning Database for this example. Kind of and we can see that as of the 2013-17 ACS five year estimate there's about 672,000 people living in DC. The male to female ratio is pretty on par with the U.S. national average. The population under 5 years-old is also pretty similar to the national. The population that identifies as Hispanic is a little lower in DC than the U.S. national level and the population that was not born in the U.S. is also on par with the U.S. as a whole. So that's kind of an example, some interesting data you can pull quickly.

All right, so in the 2016 Census Test the Planning Database was used to pull some information for Harris County, Texas which is where the test site was. And we found there are about .5 million people in 292 block groups. And as you could see in Harris County, 14.8% of the households no one over the age of 14 speaks English very well so these are non-English speaking households, about 3 times higher rate than the national average. The population of 18 to 24, not very interesting, it's about the same as the national. Renter occupied almost twice the national average suggesting that this population is more mobile and may move more frequently. And also the population was lower education is a little higher than the national.

All right, so here's a map that we created of Harris County, Texas. If you'll look remember the previous slide, the national average for households where no one speaks or not, sorry, the national. The - for this area it's about 14.8%. However if you look at where that's distributed those dark purple areas, 50% to 80% of the households no one in the household speaks English.

So it's really important when we're hiring field staff that we send people who speak this language and can help people respond.

So what if we want to identify these areas that needs support for a language other than English?

We find the block groups in that area that have a high percentage of housing units where no one over the age of 14 speaks English very well. And then we look at what language is spoken in these tracts.

So looking at these block groups actually we see that not surprisingly Spanish is the most common language other than English that's being spoken. One thing I wanted to highlight this estimate of other that's 3.2%. Look at in the

margin there you see its margin of error is 4.8%. If the margin of error is ever bigger than the point estimate the point estimate should not be used. It's not reliable at all. I just kind of wanted to highlight that.

So looking at this data we can tell that we need to send Spanish staff out to conduct these interviews.

Okay, so as promised I'm going to get into the low response score. This is our kind of updated hard-to-count score from 2010. Well, yes, pre-2010. This is the predicted low of census self-nonresponse at the tract or block group. It's important to highlight nonresponse so the higher the score the more households we expect will not self-respond.

We're using an Ordinary Least Squares regression on 2 dozen metrics from the Planning Database that are using the Census 2010 mail return rate as their dependent variable. The values at each tract or block group range from zero to 100%. And if you see a household or sorry, a tract or block group with an LRS at 25% that means we're estimating that 25% of households in the best tract will not self-respond to the 2020 Census.

The LRS is updated yearly as the new ACS five year estimates become available.

Here's some examples – or rather, these are the variables that we're using to predict 2020 self-response, age, married family households, renter occupied units. So you can kind of see that the higher the percentage of married family households is tracked as negatively associated. So nonresponse goes down if there's more married households. However if there's more renter occupied units the percentage predicted nonresponse goes up. And that's kind of what we would expect.

Okay so here's an example just using Excel with a Planning Database where I've hidden all the additional variables except for the geo locator for the block group ID and the low response score. And I have filtered so I'm only looking at Weld County, Colorado.

Okay. And this is how you can use Excel to kind of define what is a median for a particular county or tract something you're interested in.

Okay, so using the quartile command in Excel and selecting all the cells we want to include in the county, we can see that the smallest low response score is 8.3% so in Weld County, Colorado the tract at the high or lowest nonresponse we're predicting will be about 8.4% of households. For the first quartile we're predicting 17.0% and for the median we're predicting about 20% so about half of households are going to have the low 20% nonresponse which is pretty decent.

Third quartile we're predicting 23.7% and the greatest we're predicting that one tract in Weld County, 36% of households will not respond to the 2020 Census. Here's kind of the tracts laid out for you from 8.3% to 36.4%.

Some important things to pay attention to with the low response score, we're only - 2010 only had a mail self-response option. So we're only considering mail self-response when we created the low response score.

The 2020 Census is going to offer Internet, phone and mail as response options. Some census tracts have a small number of housing units in the mail back universe so their estimates aren't as reliable, for instance Indian reservations and very rural areas.

The LRS cannot be generated for Puerto Rico or select census tracts that their boundaries have changed since 2010 because they don't match up with the ACS estimates.

And always if the LRS is extremely high take a closer look. Make sure you're not dealing with a rural tract or something else that's kind of off about that tract before you make decisions about it.

And like I said unfortunately they're not currently margins of error for the LRS so it's just a point estimate. It's hard to say which - what is statistically significant high or low score at this time.

If you'd like to see more discussion of the LRS methodology including that list of predictor variables we included, please refer to this paper.

All right, so people often ask us what is a high LRS?

Generally there's not a hard and fast rule. But we say, you know, if you've got greater than 30% predicted nonresponse for a tract or block group it's probably high. However you really need to think locally. Almost everyone we talk to, you know, cares about a certain census region or state or a county. Really look at what the distribution of low response scores for that area is, not just the national.

So, you know, if you're concerned about one of these particular local areas extract like the example I gave where you just look at that area and produce a custom distribution. What is a median or the 75th percentile for that area or place?

What score is one or two standard deviations above the average?

You know and then you can identify those tracts that match your cutoff for the area you're interested in.

Okay. So you might be wondering, okay, you mentioned this ROAM. What is the ROAM?

This is a web mapping application that provides a map and data table interface for users to identify and learn about hard to count tracts. It's accessible through a web browser. You don't have to download anything. You do have to download the Planning Database and it's hosted by the Census Bureau. And with that I'm going to hand off to the expert who actually stated this.

Suzanne McArdle: All right so this is Suzanne again. And we're going to dive right into a live demo of the application itself. And I think looking at the maps themselves lend this - lend a little bit more - help make this make more sense in terms of the low response score and some other things that are available in here.

So the foundation of the Response Outreach Area Mapper or ROAM as you can see in - as we enter into the application is a national map of census tracts. So we're looking at census tracts that are shaded by the low response score.

So as Kathleen mentioned that the low response score, the LRS, is simply the predicted mail nonresponse rate. So this legend over here on the right side of the map and I do just want to say right now that I'm going to be going to the actual Web site that this is publicly available to show you how to navigate to the application itself and to show you how to navigate to some of the resources that we've already created and put onto the web for folks to refer to.

So if we look at the legend on the right hand side we can see that the higher the low response score is the darker that area is shaded on the map.

And so this darkest blue color here are census tracts that we consider our hardest to count.

So this hardest to count set of census tracts is 30% to 48.4%. So you can think of this as this area in Northeast Nevada. These census tracts we're predicting that 30% to 48.4% of the households in these tracts are not going to respond to the 2020 Census.

And so while I'm here I did just want to click on a census tract to show you that all of that other amazing ACS information that Kathleen highlighted is accessible right through this simple web map viewer. You don't have to download the Planning Database to do this. You don't have to join it to a geo spatial file. This is something that you can use sort of right out of the browser.

If we scroll through this popup box here, we can see that this particular census tract is in Elko County, Nevada. And we can see that 32.3% of the households here are likely to not self-respond so even though we can get the actual low response score for each individual census tract in the country by using this tool.

We can see that there are 2900 people who live in this census tract. And we can get some other information about the people in households here. So the median household income is about \$30,000. The median age 34.8. We have some other age metrics here. One of the historically hard to count populations are children under the age of 5.

So we've selected a number of the ACS variables that lend themselves to identifying hard to count populations so that folks who work here at the Census Bureau and people working out in their local communities can use this information to tailor and, you know, really hone in on the message that they're sending with their outreach.

So you can see that there are a number of other variables here. Kathleen mentioned a number of them earlier, educational attainment, race, language spoken at home. We have additional variables that speak to the mobility of the population. That's another hard to count population. People who have moved recently so we can see that we have the, you know, the percent of people who have moved from another residence within the last year.

And we have some other housing unit metrics. So is there vacant housing in this census tract? Are there multi-unit structures in the census tract?

With the latest ACS five year release we're actually able to get access to a number of different Internet availability and Internet accessibility estimates. And so we've also put those into the application so that we can see that, you know, 61, almost 62% of the population here has access to broadband Internet and a computing device. This is incredibly helpful information since, you know, for this decade we'll be introducing safe and secure Internet self-response as Kathleen mentioned earlier as well.

So right off the bat this tool is very informative just by zooming into an area and clicking on an area to learn more about it.

I did kind of want to highlight some of the other types of functionality and things you can do. And then we'll dive deeper into some more specialized examples of things that you can do with this tool.

So we have a number of Navigation buttons in the upper left of the tool itself. We have this Home button here that'll take us back to the full extent of the contiguous U.S. And you can see that we do have Alaska and Hawaii in here as well. We also have some bookmarks that are created in the upper left here, if you'd like to navigate to a particular area on the fly.

So we do have - as Kathleen mentioned, the low response score is not calculated for Puerto Rico. But we do have all of the ACS estimates and the ACS data that is disseminated for Puerto Rico.

And so you'll notice that some of the census tracts on the map are in this grey category. That just simply means that low response score is not calculated.

One of the things that you can do if you need to remember in instances where particular census tracts do not have - where we can easily reference the scenarios in which the low response score is not calculated for a particular census tract.

And so that does include where the geography has changed. There we go. And where mail out and mail back areas or excuse me, and which areas in 2010 that weren't mail out or mail back.

So let's go back to the contiguous U.S. And you can see that we can zoom in to a particular area of interest using those buttons. We can zoom out from that area. We can go back to a previous extent and we can even set our own bookmarks.

So let's say that we want to come back to the view of Atlanta here. So if we go back to the contiguous U.S. Then we can come back to Atlanta.

One of the things I do want to mention is that every time you do access the ROAM Application you'll be starting from scratch. So if you have set any bookmarks and you've reset your web browser history or your cookie settings you may not have bookmarks that you've set before.

So you should just operate in a manner that you know that any time you come back to the web browser you may be starting from scratch. So that's just something that I wanted to point out so that people don't necessarily spend a ton of time configuring one version of their web browser on that.

So we do have a couple of other ways to navigate across the map. And I'd like to go ahead and explore some of those now.

And so this search box here has a number of different types of ways that we can navigate to an area.

So you could actually go to an area on the map by address. You can just zoom to a particular state, a particular county that you're interested in. If you are interested even if you know exactly what census tract you want to go to, we can navigate by that by something called Regional Census Center and Area Census Office. Excuse me.

These are geographies at which the Census Bureau internally manages our field areas. And so we've, you know, gone ahead and created these layers and put them into the public version of the map so that we can be very transparent with the way that we're operating in the field.

You can search, you know, navigate to a particular tribal area, Alaska Native Regional Corporation, Minor Civil Division, school district, you can navigate

to a particular congressional district, state legislative district, zip code tabulation area. It's known as ZCTAs). So this is the best spatial representation of a zip code that we have available. And finally you can search by a particular place.

And so right now just to show you how this search bar is going to work I'd like to navigate to Langley Park, Maryland. All right, and you can see that the App automatically takes us to the full extent of the place that's called Langley Park, Maryland so that this shaded area in orange is Langley Park.

Now for anybody who knows where Langley Park, Maryland is this is a good enough reference for us to be able to do - for us to be able to understand where we are. But a number of you on the phone I imagine you're not familiar with where Langley Park, Maryland is.

So I wanted to step outside of this and just highlight some of the other functionality that we have. So one of the things that we can do is actually change the base map so the underlying data behind the low response score data layer we can get a sense from looking at satellite imagery and that low response score what type of area we're in.

So we can see that we're not out in rural America here. It looks like we're in an area of industrial mixed use area along a route mixed industry here. We have a number of different residential areas, single family homes, looks like we have some townhomes here. So these different underlying data layers can really help us understand what's going on in the outside world.

If I scroll it back to the original base map we can turn off the low response score to reveal the underlying street network to get a little bit more reference on where we are.

And so one of the other things that we can do is change the transparency of this layer as well if you need a different effect going on with that.

So again, while we're in this particular area I wanted to bring up the data table and show you how we may make use of some querying power from the ROAM Application.

So one of the things I'd like to do is just shift. So for anybody who doesn't know where we are still, I'm sorry about that, we are just outside of Northeast Washington, D.C.

And so what I'd like to do is query the data window. So any census tracts that's within this map window I'd like to go ahead and query it to see if we can find census tracts that have a high percentage of Spanish speakers at home.

So I have the Filter by Map Extent button on this data table. And I access the data table by this simple little tab down in the bottom here.

So I'd like to go ahead and use the options and filter tool and add a filter expression to this map. So again I'm accessing the data for all of the census tracts that are currently in my map window.

And so I'd like to go ahead and find our language variable; we do have a data dictionary available on census.gov. And I'll go ahead and show that to you. All of the data variables that are in this application are named exactly the same as they are in the Planning Database. So for any of you who are familiar with, already using the Planning Database, we wanted to make this a seamless transformation.

So let's find census tracts where more than a quarter of the population speaks Spanish in the home. So at the bottom left of this data table now I have 23 features.

So one of the things that we can do here is select this; it's not really a button but select the record header at the left of each record. And we can shift, click all of these census tracts and highlight them directly on the map.

So you can see to the Northeast of Washington, D.C. we just highlighted an entire pocket of census tracts where more than a quarter of the population here speaks primarily Spanish in their home. This is incredibly telling of the messaging that we need to use in this area, the obvious language that we need to use that messaging with.

If we are a partnership specialist that actually worked for the Census Bureau you'd also be working to identify trusted messengers within this area.

And so what do I mean by that?

I mean that there are field staff around the country working to identify individual establishments within particular areas to visit and build rapport and a relationship with.

And so one of the other things that we can do to make this map even more useful is adding local data to the map so can search our online category for something like a library. And add different libraries onto this map.

And if I were a partnership specialist I would be using this Maryland layer here. We'll let that load. And we can see that there are a couple of libraries.

This guide to - in the center of the map and one to the east of that entire, you know mostly Spanish speaking area. And we can plan to have an actual event there and have a bilingual partnership specialist host that event and share the message of the census and what it means when you respond.

Another thing that I wanted to show you while we're here is that you can actually export a subset of the census tracts to a CSV file.

So let's say that I'd like to keep a hold of these 23 census tracts here. I can go back to my Options menu and export selected to CSV.

And now I've gone ahead and I've saved down this entire data table that has all of the variables that were also included in that popup we talked about at the beginning.

And now you have your own little subset of the Planning Database at your fingertips to be able to, you know, bring that into a third party application if you'd like. You didn't need to download the full Planning Database. And you were able to use this geographical interface to be able to identify areas or different populations that you wanted to collect and download data for.

One of the other things that I want to go ahead and show here, I'm going to unselect all of these things, and so every type of geography. I'm going to remove those libraries for a moment too. So the different levels of geography that we have in that search box we also have them available as reference layers in the application.

So if we go to this additional layer we can see. Let's go ahead and turn on this school district. Excuse me, zip code. Go ahead and do zip codes. It's

something that everybody recognizes and lives within and many neighborhoods are known by their zip code.

So now we can click on the map. And we can see the full extent of this 20783 zip code. And we can see that to the southwest of this zip code is where the hardest to count areas are. So now we can take this map of the low response score and identify, you know, within a particular type of geography where my hard to count areas are.

One of the other things that partnership specialists are doing on a daily basis is meeting with local legislators. And so we can turn on this legislative area here and we can see that we have our congressional district boundaries and our state legislative boundaries.

So we're going to go ahead and let's keep on the same area that we focused on in - earlier in Langley Park. We can click on a census tract and also click to see the full extent of that actual congressional district.

And we can say where this Congressperson would want to be spending most of their outreach resources to conduct the 2020 Census. And that's in these darker areas that actually, you know, most of the part of Prince George's County, Maryland that borders the District of Columbia.

So no matter why an area is hard to count our 2020 Census outreach consists of installing voices directly in the community. The communities that we've helped with to show this tool to are using this tool and seeing that ROAM is simply reinforcing what they already kind of have a gut instinct about in their different communities.

So this is really serving as a mechanism to allow tribal, state and local areas to also tailor their own outreach efforts and direct to their own outreach funding that's coming from various levels out in the field.

And so this is the first decade that the Census Bureau is using a tool like this. And it's also the first decade that we're making this available directly to external stakeholders to help inform their outreach as well.

Before I pass it back to Kathleen to wrap up here, one of the additional layers that is brand new that a lot of people are not aware of that has been added to ROAM is something called the 2020 Census Audience Segmentation Data.

And so this was the result of a lot of research. There's a paper that was just published. I'm going to go back to the legend here so we can understand what we're looking at now.

But in addition to the low response score where you can, you know, identify areas that are hard to count and then learn more about them, this particular data layer has classified each of the census tracts in the country into one of eight different types of census tracts.

So these are ranked in order of responsiveness so responsive suburbia is a census tract that's more inclined to respond to the 2020 Census.

And as we go down this list the response goes down as well. So I'm going to hand it back over to Kathleen.

But I just wanted to make you aware that this is a data layer that's available in ROAM.

And if we now click on a census tract we get, you know, this particular census tract is considered a multicultural, mosaic census tract.

And if we click on that link right here from within the popup we get an actual summary of information that shows us who is there and how likely we think they'll respond and what type of response we think they'll give whether that's Internet or otherwise.

So I'm going to go ahead and turn it back over to Kathleen just to wrap up. And then I'm excited to answer some of these questions that we've seen coming through the chat here.

So let's dive right back in here for Kathleen.

Kathleen Kephart: Great. Thanks Suzanne. So some of the new things in 2019 we've got the ACS five year Internet and technology variables are now available at the tract level, not the block group but on the tract file. This includes count of households with broadband access, the number of households with smartphone only access so that means in that household the only way to access the Internet is a web or a mobile device which is important to know when you're planning survey response.

And then some other characteristics about the number of people with access, etcetera that are available on the file.

We're adding several variables that have existed for several years to the Planning Database at the request of the National Academy of Sciences. These include 3 to 4 year-old school enrollments, households with a child under 5 present. This is to help target areas that are more likely to have high

populations of children and the undercount of children efforts which is a serious problem that many surveys and 2020 Census faces.

We're also now including the ACS median age for each block group and tract and as well as Suzanne kind of mentioned this, 2020 Census Audience Segmentation. Because it's not enough to just know that there's going to be low response. We kind of need to know how to tailor the message and how we can reach these populations.

Okay, this is really exciting, the ACS five year census tract self-response rate. This data product is the actual ACS self-response rate at a tract level. And it is not publicly available anywhere except the Planning Database.

And in the ROAM you can't get these color maps with the self-response rate. But it does come up in the popup. I don't know if you noticed.

So it's really neat that this is now available so you can compare. You know there - I don't know if there is another national survey besides the actual Decennial Census that you can look at the self-response rates.

So the self-response rate I should add here that you're seeing includes Internet, phone and mail so it's non-interviewer basically. It's all combined.

Okay so this tract segmentation, the goals of our segmentation are to kind of provide an overarching framework for understanding the country because it's not enough to just know like they're going to have high nonresponse. We need to know more about them to reach them, to motivate them.

What we do with these segments is we use geography to bring together behavioral, demographic, attitudinal and media usage data for campaign

planning. What's the best way? If we're going to do an advertising campaign, what's the best way to reach this tract? Is it through online ads, on Facebook, etcetera? Is it through newspaper or radio ads?

This is important information for the partnership campaign to have and the segmentation data is bringing it to them.

So we're kind of simplifying a lot of complex data and information about the tract by identifying the key shared characteristics in the tract.

So this is going to inform like I said our overall program strategy for reaching out, partnership and advertising. Communications, creative development and campaign optimization and tracking during enumeration.

And I just wanted to add the ROAM won't contain live 2020 response rate information but that data will be available through census.gov. We're going to have kind of a live response rate to live tract for people to know okay we need to do something to increase self-response.

Okay, all right, so what are some of the tract segments?

So this is kind of that cool multicolor map that Suzanne just showed us. This is the breakdown of those segments that we've created. Your Responsive Suburbia tends to be, you know, these tracts that overall have more single family homes, married couples, they're highly educated. And we also expect that, you know, we're going to have 71% self-response from these people. So we're not as concerned about them as maybe say the multicultural and mosaic group that I'll dive into in a second.

So this is just kind of an example of these different segments we've created that kind of combine a lot of variable and data sources to explain what the majority of the population is like in each tract. Main street middle, country roads, downtown dynamics, student and military communities, sparse bases, multicultural and mosaic which I'll get into and then rural, delta, and urban enclaves.

So knowing this information can help us kind of figure out what kind of partnerships we need with nonprofits or churches? What's the best way to reach these groups? And also how to plan advertising as far as billboards, buses, etcetera.

Okay the multicultural and mosaics, so the biggest thing, most of these tracts we find are in California Central Valley, parts of New Mexico, Texas and Florida with a couple additional concentrations in urban areas.

Who are these people?

The tracts in this segment have relatively high concentration of foreign-borne residents as well as low percentages of college educated residents and a majority of the people are - in this segment are Hispanic.

How will they respond? On average we're predicting this segment is going to respond at a low rate with a below average percentage of response coming to us online.

Let's see, so you can kind of see here on these bar charts, the grey is the national average. And the purple is in this segment the percentage of households that are Hispanic. You've got almost 60% versus the national average of about 18%.

So you can also see the owner versus renter. About 71% have Internet access at home. That's a little lower than the national average of 77%, etcetera.

All right, okay. So we can kind of - we've got some more demographic data about these, the multicultural and mosaic and all of this data. Is this on the ROAM Web site?

Suzanne McArdle: This is - yes, this is available on the ROAM Web site.

Kathleen Kephart: Okay.

Suzanne McArdle: So if you click on a particular census tract from the ROAM map of the Audience Segment and click on the URL it actually takes you to each one of these individual key map summaries.

Kathleen Kephart: Okay so you can kind of get a more in-depth profile of each of that big chart I showed you, responsive suburbia, multicultural, mosaic, et cetera.

So kind of diving in, you know, you see like I said 34% foreign-born, about 45% married couple households, 28% more higher than the national average have a child under 6. The median household income is below the national average at about \$45,000. And only 15%, about half the national average are college educated.

So like we thought it's a little lower educated population and also much higher than the national average non-English speaking, 18% of households.

This other cool thing that we have is the mindset composition. And this is kind of these profiles we've created of where people stand on responding. On

the left hand side you have your eager engagers. We predict about 19% of the national population, that grey bar are eager engagers. These are the people, the letter comes in their mailbox. They pull it open. They go online. They respond. They're done. They're excited. They're engaged.

We're predicting only about 11% of these segments, the people in these segments are eager engagers who are just going to respond right away with no kind of encouragement. Most people kind of fall in the middle, your fence sitters. They're not actively opposed to responding. They probably wouldn't take much encouragement. But we need some kind of outreach or buy-in to get them to respond.

Then on the far right we have the disconnected doubters. These people have no interest. They have actively chosen to not engage and they may be harder to encourage. So any kind of marketing, we need to really target.

You know in the middle we also have the confidentiality minded. This is a good example. So in this mosaic or in this segment, sorry, multicultural, mosaic, 33% of people are concerned about their confidentiality and what the government is going to do with the data from the census.

So maybe the best way to reach them is to engage with partners they trust. And develop advertising to reassure them about what - that their data will be protected and how it will be used.

All right, so how do we reach them?

If we wanted to do say, you know, advertising they're about on par with the national average as far as the number of newspapers they typically read, the

number of magazines. So if you're doing ads in newspapers and magazines it makes sense in this area.

They drive a little bit less than the national average in a seven day period so maybe billboard, not as much for what we call main driving routes. They listen to a similar number of radio so maybe radio ads would work, same with TV. And they use the Internet kind of on par so all that your regular - it makes sense to advertise in all your regular outlets.

Okay. So we hope you've kind of been engaged. And I'm going to hand off to Chase.

Chase Sawyer: Great, thank you Kathleen and Suzanne. To go ahead and finish up I just want to go ahead and let you know some ways that you can keep connected with the ACS and the U.S. Census Bureau. One way that you can do this is through our Gov Delivery System. This may be the way that most of you became aware of the webinar.

And it's actually a way that we'll go ahead and make sure that you know that the slides are available once we have those posted in Census Academy.

So if you're not signed up and would like to receive those slides or just would like to receive updates from the U.S. Census Bureau I encourage you to go ahead and sign up to our email list.

You can also find out more about the American Community Survey at [census.gov/acs](https://www.census.gov/acs). We also have a Customer Service Center line. And I've also included our email for user support as well as questions for the Planning Database.

Finally I just encourage you to source us or the American - or the Planning Database to let people know where you found that data so that they can use it for themselves.

And so with that Operator I think we're going to get ready to start queuing questions.

And then we can also go ahead and look at getting a few questions from the chat line.

Coordinator: All right. And of course at this time if - on the line if you would like to ask a question you may press star 1 on your phone and record your name clearly when prompted. To withdraw your question you can press star 2.

And please be advised that you are allowed to ask a question as well as a follow-up question and then you will be removed from the queue at that time. So we'll just allow a few minutes for some calls to come into queue.

Chase Sawyer: Great, thank you Operator. And so for some of the first questions that we're getting on the chat, one of them I thought was, what are some of the limitations of ROAM in rural areas?

And I think this is something we've went over if we want to just touch on it again.

Suzanne McArdle: Yes. So most of the country does have a low response score calculated for the census tracts so obviously if - actually let me go back to the map just to help support this.

So if an area was in a 2010 mail out/mail back area it likely has a census tract, has a low response score calculated for that given census tract.

So let's just take a look back at the contiguous U.S. real quick. And you can see that most of the country does have a value associated with a low response score. So is either yellow to blue. You should take note of how many people are in a particular census tract when you're taking into account the low response score for that area.

But in terms of limitations, the data is available for most rural areas.

One of the other questions Chase that I just saw come in is related to American Indian areas.

And so if anybody is familiar with Navajo here, you'll see that it does not have a low response score calculated for most of the land there because that area was not in a mail out/mail back area in 2010 so there was not a mail return rate associated with it. And we were not able to calculate the low response score. That doesn't mean that this tool is not helpful or not available for those particular areas. We still have the American Community Survey estimates that are included here.

So if we click onto the map, we can see that we do have the actual American Community Survey data and information from the Planning Database included, you know Internet accessibility and other things included as well.

One of the other things that I did notice is somebody asked if there is a Planning Database file to download for the entire nation for all census tracts and for all census block groups.

And so that's a really good point and we do have the link directly to the Planning Database to share with you. So the Planning Database for census tracts is one national comma delimited file. So you're downloading the entire Planning Database at the census tract level.

Kathleen Kephart: So tract and block group, so there's two files. But yes they don't match.

Suzanne McArdle: Right, right. So you'll get one Planning Database file for the census tracts and then one for the census blocks and each one of those files contains all of those geographies for that particular level.

So one of the other things that I mentioned that I was going to provide and I never went back to it is how to actually access ROAM directly from the web.

So if you go to www.census.gov/roam. That's R-O-A-M. You'll stumble upon this landing page for ROAM. We have a number of the different links to the Planning Database right here. The methodology that Kathleen pointed out earlier and a number of different resources including what's new from last time since we have published this application a couple of different times. A User Guide, a data dictionary, FAQs, a number of things to help you get started using this.

So census.gov/R-O-A-M is the best resource to go to.

Kathleen Kephart: Okay.

Chase Sawyer: Great. Thank you so much Suzanne. Operator, can we go ahead and get questions from the line now?

Coordinator: Yes. Our first question comes from (Chris Anderson). Your line is now open.

(Chris Anderson): Yes. I was just wanting to find out when might have this again because I've been unable to access any of the video, only been able to hear it on the phone.

Chase Sawyer: All right, well yes. I'm sorry to hear that that happened. As I mentioned earlier the gov deliveries, we'll go ahead and we'll send out the slides. I should've mentioned earlier though that we'll also include a link to the YouTube video of this webinar.

So you can feel free to review that and then yes, that's another way to keep in touch about when we'll be having other webinars like this.

Suzanne McArdle: And (Chris).

(Chris Anderson): Thank you so much.

Suzanne McArdle: Just one more point here in the short term before that recording is ready if you do go to [census.gov/roam](https://www.census.gov/roam) there's actually a previous version of this recorded webinar that we've provided for folks. It doesn't have any of that 2020 Census Audience Segmentation Data in there but that might be enough to get you started in the meantime.

(Chris Anderson): Appreciate it. Thank you.

Suzanne McArdle: No problem. Thank you.

Coordinator: All right. And our next question comes from (Michael Sullivan). Your line is now open.

(Michael Sullivan): Hello. I wonder if you have any guidance about sort of when we might want to rely on a low response score, the mailed response from 2010 and the ACS five year nonresponse rate.

And is the low response score based sort of taking the 2010 mailed response and then applying demographic variables to generate investment of future risk or does it also take into account the ACS five year nonresponse rate?

Kathleen Kephart: Okay. So first the low response score, it's using the 2010 mail as its dependent variable. But it is using updated ACS to...

(Michael Sullivan): Right.

Kathleen Kephart: Data as the independent variables. It does not currently use the ACS self-response data. But it correlates. We know we've done correlations between the predicted low response score and the ACS self-response rate and they're very highly correlated. So it's doing a really good job predicting ACS self-response even though the current model and score doesn't use the self-response.

I just wanted to add, because you said ACS nonresponse. So the ACS response is actually the self-response rate. It's not nonresponse. So the higher it is the better unlike the low response score which is nonresponse. The higher that is that's bad. Self-response we want, you know, 70% of households self-responding or more.

(Michael Sullivan): Okay so I guess the rationale for using the low response score is that it takes sort of the known response score from 2010 but add these more current demographic variables.

Kathleen Kephart: Yes, exactly.

(Michael Sullivan): To assess (unintelligible) essentially. Okay.

Kathleen Kephart: Exactly. Yes. It actually - it does a really good job predicting self-response rate with the latest ACS self-response rate data. Exact - yes, it's been updated.

(Michael Sullivan): Okay.

Coordinator: All right, and our next question comes from (Angel Kristanus). Your line is now open.

(Angel Kristanus): Hi there. I was wondering if there was any data that kind of touched upon new buildings or new construction over the last couple years. And also like if this will be updated when the 2018 five year will be released.

Kathleen Kephart: So new construction that ACS does provide estimates on that. But that is not on the Planning Database.

(Angel Kristanus): Okay. So it doesn't pull in anything from the census like new building permits or anything like that.

Kathleen Kephart: No. It does not. No.

(Angel Kristanus): Okay.

Kathleen Kephart: Census has a whole area that specializes in new construction if you go to census.gov. But the Planning Database doesn't use any of that.

(Angel Kristanus): Okay, thank you.

Coordinator: All right. And as of now there are no further questions in the queue.

Suzanne McArdle: Can you scroll towards -

Chase Sawyer: Thanks Operator. We're going to go ahead and then take just quick look at the chat real quick. Okay so we did have another question in the chat that we're able to go ahead and answer. It was we had a user ask that they wanted to use the Planning Database or ROAM to calculate census, the 2000 or 2010 undercounts for specific race or ethnic groups. And we're going to have Kathleen answer that for us.

Kathleen Kephart: Okay. Not currently. The low response score doesn't do that at this time. We do take into account the percentage of different racial groups as part of the predictors. But we don't produce low response scores by race or ethnicity at this time. But data users can use the data however they want. But I thought we can answer, like if you scroll up, they asked about -

Suzanne McArdle: Right, so you're going to find it there.

Chase Sawyer: All right. And we have another question that we're going to go ahead and answer about naval bases.

So the question was - sorry about that. So the question was that a data user had noticed that there are high low response scores for navy base and university. And they wanted to know is that because of the transitory nature, group quarters count methodology or something else.

Kathleen Kephart: So it's kind of a combination of all of the above. The low response score only uses households that were in the mail back universe and a lot of this population because they're in group quarters enumeration they're not in.

So we're basing a score, even if the area gets a score it's not on all the housing that's in group quarters. It's only those households that were in the mail out/mail back universe.

So it might look high but it's just - it's inaccurate because it's based on, you know, only 5 out of 100 households or something.

And also the transitory nature plays into it as well around colleges and naval bases.

Suzanne McArdle: And to that point if you use ROAM and actually click on the individual census tracts that you're noticing are on a naval base or college or university you can basically ask that question of the popup. Like, you know, what are the indicators that are sort of showing? Is there a transitory population? Are there large percentage of the people that have moved off that area as opposed to one year ago? What's the median age? How many percentage are under the age of 5 or between 18 and 24? That type of thing.

So a lot of these questions are something that you can actually ask directly of the application and the data themselves. So a lot of people are interested in different parts of the population or different geographic areas. And so this tool is really opening up a mapping gateway into this data set in ways like we haven't been able to provide before. Sure.

Chase Sawyer: All right, well we are past the hour. So I just want to thank everybody for calling in today and learning about the Planning Database and ROAM.

If you do have any additional questions please feel free to reach out to us at either the Customer Support Line or either of the email addresses provided. Thank you everyone. Have a great day.

Coordinator: And that does conclude today's conference call. Thank you for your participation. You may disconnect at this time.

END