Secrets of the Super Searchers II – Statistics – Transcript of audio

Good afternoon, everyone. Joe here. Doing one last sound check. Will be getting ready to start in about two minutes.

Good afternoon, everyone. Welcome to the Academy. We have another terrific webinar for you today. My name is Joe. I am with my colleague Ashley. She is tech support. Our presentation today is secrets of the super searchers theories. Here are the statistics. With us today is our presenter. Suzanne. The depository coordinator at the New Mexico state library. Until fairly recently she was the government information library and a North Dakota State University. A co-region for that state. Previously to that she was the FPL regional at the University of Montana Missoula. And then the state library. She was the editor and author of government essentials which was published in 2018. She told me she has two cats and one spouse. With that introduction I will handed over to Susanne Caro. She was taken from here.

I just realized I muted myself. That's a great way to get started. Hi, everyone. Welcome. What I will be covering today is based on a number of questions that people submitted. Most of which it up being focused around statistics. That will be the focus of today. First, just some general questions that came in. What is the most go to reference source to answer most questions? Since this is more of a statistical event. I would have to say the census. Followed by the Congressional research service. I love their publication because it gets a good starting point plus some helpful background information. For ordinal older statistics. Historical statistics of the United States colonial times. And then another place is to go to the statistical abstract. If you want a free voice resource. Find out what agencies have the information you are looking for. Statistical programs of the united states government is a great source. Of course the more direct replacement would be the ProQuest statistical aspect. Another question is involving academic statistics. Where do you go to find those? Would you do when people are asking you for the statistics that you provide them but they have no idea how to use that data?

I would say if someone is trying to access the data or say they got a sheet full of numbers to try to determine what they are actually been looking at then look at the reports that came with the statistics. Those options provide a lot of information. That will help someone to understand it. Also if you are in a academic situation then you might have a data library or a subject specialist whose more familiar with the content that can be helpful. I would also say to ask of the agency that provided the information to begin with. They like it when they use their data. They want people to use the data. Frequently they can also help. I don't think we all need to be specialist in every type of data. If there is someone who is like that then I would love to meet them. How is the data collected? What statistical programs are used to analyze it? How do you handle the skewed data?

How the data is collected is depending on the agency and information collect it. It sometimes may be mandatory or particular agencies to receive data. From states and other departments for example. Statistical programs could be something like Python. Those require some training to use but they are with statistical programs that you can use. They can go through all of the data and get some good information that you are looking for. How you handle skewed data?

I think that is up to the researcher. It is, of course, going to depend. If you go back to the report of the company information that you can see what was included in what was not included. Frequently they will say that they took out the iPlayer outlier to have something more consistent. As long as you also recognize and acknowledge that the data is not going to be perfect.

Are the quality control measures in place? Yes, and they vary. Are there any measures for federal entities to decide or implement what data gets rolled up from state and locals? That is frequently related to legal requirements. If you are in a particular program or sticking the federal funds then there are often reports requiring that go along with those. Those are going to vary as well. So you where you find different types of data? What do these different places actually have? Where should you send people who are looking for these types of data? We are going to look at some of these resources. What kind of data they are providing and how it differs. What statistical program or reports will be mandatory and guaranteed by law to the budget-cutting. I would say that it is honestly 1. I have this chart to give you an idea of what this looks like. I'd say that the one area where they are probably not going to be discontinued is going to be this census because that is actually in the U.S. Constitution. Everything else has to have its budget cut. Everything else can be easily changed. The laws and standards are going to have differences depending on what kind of data they are getting and how it is being handled. There is a lot of variation. For the national Institute of health then I choose variance. In some cases there will be similarities. For example, the Bureau of economic analysis is going to have somewhat different information from the Bureau of Labour statistics. They have more international data in BEA. Then the Bureau of statistics will have more national data. The census will have the demographic, education, income and housing. They are the broadest of all of these. Part because they have additional questions that get asked as a part of a more specific census. Such as that they have a questionnaire that is going out regarding effects of COVID. It's a broad area. At times you will see census of agriculture. They have very different reasons for collecting data. There are data sets and databases with the federal scientific statistic. One centralized place would possibly be data.gov. It does have a very large number of searchable data sets. Often it is just the data and you do really want to go and get those reports. From this I just did a quick search to see what agencies were having the larger number. The U.S. Census is there and it is followed by the USGS. That is based on the previous chart. This is how much they are producing. This is how much you can find on data.gov. They have started this from the secrets of the super searchers. You can go back if you have not seen that yet. Go to the Academy and put the URL in the side. You can also find other trainings and webinars. This is idea of how many agencies are also providing data sites. Data.gov is great but it is also very large. If you know that someone is looking specifically for condensed data then he can be faster to go directly to that source. What is my strategy for doing it? This is my strategy but one of my first questions is what is a level of information that is needed? Does someone need 50 years of statistics or granular data? Do they need broader data? Are they hoping to find something general? Are they looking for something that is specific?

What agency will be most likely to have the information. Does the website have a report or data section? I want to know how far back I need to go because that can really change where I look. I might have to go to the stat or I may have to look at multiple agencies because they changed over time. I want to know if there is a statistical abstract because I am not going to know exactly where all of these things are right at the top of my head. I want to find something that will help me get into the right direction. I could also be the Congressional research specialist. Can I contact the agency if I think I should have a input? In my finding tables that will close but not quite what I'm looking for. I will often contact the agency to say that I'm looking for it but I am finding only this. I can either confirm that they don't have it or they will provide me with a specific table that I needed. They will put them into direct contact with the person was asking the question so that they can interact. Are they in access? Demographic data et cetera. Let's talk about the census. The education. This is where we will try to get a little bit experimental with this. I'm going to ask you all to do a bit of a quiz. Technically there is no wrong answer. If you are looking for how many school districts in the U.S. require some form of school uniform as compared to the dress code then where would you start? This is asking where you would start on

this. Giving a minute to think about it. I see someone is looking at the national center of statistics. At least one person with the education of the state. Okay. We can go ahead into where I look first. That is the and CES. That is the national education system. I didn't search just for school uniforms. I don't know him very much about it this is a good place to start. I found that they had how many schools needed to require uniforms. There is a response between 2017 2018 where 20% is given. What's really important that it is also giving you the source information. I went ahead and use that arrow because why not? If I can go to the source information than I can find additional content. Going to the school survey on safety. I also found information about this. Specifically the schools that are excluded. This can be very helpful to have an idea of what you are actually looking at when you are looking at the statistics. They will not include all of these different areas.

You do have to pay attention to the dates. Slightly older information here. Here's another question for you. Where to find data on adults with some college but no degree. Again, this is an education question. We have the census for the statistics. The census microdata as well. I am seeing a lot of document coming up. Remember, too, there are no wrong answers because all of these sources will give you content. Some may not be exactly what you think.

If you go to data the census of -- I will do a little bit more of a walk-through on this. If you do a advanced search then you can select your geography and your geography can be national or state. It can be Metro. There's a number number of options. I always go for America community surveys of five years. It's going to lead to more accurate data with that. They've come find five years worth of data and it will be more accurate. If you're trying to find something more recent such as the information that they are gathering on long-term effects of COVID then you do not want the five-year but the one year. If you look at the top of the education that will give you a number of tables. Table S1501 with the educational entertainment to give you some college and no degree. It's also giving you margin of error. Is not the full census. There is also something called DataLabs from the national Center for education statistics. You can do a search for no degree. It will give you options like the census does. Credits earned over time. No degree Amnon enrolled. Looking at this too -- here's the source. A study with 2009. If I search for educational attainment than I can get the statistics which is honestly one of my favorite things to have found. It's absolutely fantastic with this statistic. One of those that they have included with table 501.20 employment population ration. Highest level of education attainment. That information was coming from the census. Also when you're looking at these, pay attention to the table numbers. Often they are fairly consistent. You can look from one year to the next. DataLabs has a very nice layout. For example, you select your variables. You arrange them and you go ahead to do your search. Here we have postsecondary students. You would go in and choose your various items. In this case, we have calm variables by stop. I came across this with what the heck is this all about? Once you have what variables you want you can go ahead to hit run. I wanted to find out the stop out. They have a codebook that has definitions for all of these. In this case it applies to respondents who have not yet completed the degree or certificate/enrolled in/plan to enroll. These are individuals who have taken courses but have not actually received the degree yet. Sometimes one of the big challenges is learning the different terminologies for the agencies. In this case -- I have estimates only. I've never stopped out for various years. One of the issues is that some of their data sets are very dated. Again, this only goes to 2014. That might be all that you need or you might need additional information. Census microdata. I love census microdata. If you have not jumped in and played with the data then you really should. It's actually very simplistic. In this case, I have my selected geographies. Specifically going for the most region versus estate. Also years. This is the number that I have for some college but no degree. My variables were education and educational attainment. I also have the code here. Those are pretty consistent. Again. I selected some college. This will let you bookmark your table. You can bookmark it and go back. If you

have not played with the microdata while this is basically where you start. You can select the data site and it. You can select whatever year you want and what data set. It's five years. And you select your topic. A large number in variables. If you know then you can go ahead and filter by that. In this case I've got educational attainment. Then I slipped my geography. This time I went for New Mexico. You can also limit your content. This can help you narrow down what they will look like. For school, these are all of the burials. It goes preschool all the way to doctoral. The information I want is just right here. You just select those two. You can also select ages as well. In a similar manner. If you are always looking for a specific number of the ages that this is where you go. Otherwise you will have to look at the census and main tables. They have a very set number. If you want everyone over the age of 18 and that is where you do it. It will give you a preview of what your table will look like. You can always move from around your goals and columns. Sometimes it won't let you depending on your selection. Here is our table. Now I can see the total for New Mexico. Some college of less than one year and one year of college but no degree. You can always go back and change some of the variables in your data set or change the year. This is a lot and extremely helpful.

You can also use this for the percentage of people in their 30s or older to hold their higher education. This is a very similar process to what I just showed. You've got your survey and your variables. You can change that and custom groups. Changing it to 30 and 99 years. I did have to move education to columns. I removed the needed variables. I don't want to know about people who are --. I can focus on the doctoral and masters et cetera.

There is another question. Social dynamics of the coastal communities. Are they growing economically? Are there increase public and private partnerships? A very easy question. Where would you start with this. The Federal Reserve or fax? The economic analysis. Will are very fond. If we start with census microdata. Then you can get the five-year estimate. You can go for whatever geography you want. In this case a microdata. For variables on this that I use residential mobility. There is also a migration option. In this case I chose several state that are coastal. That I also selected a number of counties in California. Just so you can see how detailed that information is. The question is specifically when someone moved. That can give you an idea of health long someone has been here all the way up to 30 years. That does give you little bit of an idea of the people that are in. If we can go to this then we can go for migration and geographic mobility. One of the tables that will come up is this geographic mobility I selected part. Again this will give me information on various things. Sometimes lower levels as well. Giving me that demographic information. If you have and use the census advanced search than it hasn't changed all that much from when it was the American community survey search. Again, you have your geographies. You can choose multiple locations. I really like. You don't always have that option. Your topics. In this case is population with people and mobility. We have our tables. That is just that. It also has some other inclinations. You can look at a number of these. You can see which one will be there. Fred is a great resource. It does have some of the broader information as far as industries go. You can actually take to improve your ability and familiarity with the materials. The categories angiography again. I think that is 1998. Up to the present. It's a gray area. Also COVID is located there. Giving some context. It's pretty straightforward. Going for the category. Taking a look at all series. Selected state before I did that. About 76,000. 17,000 and then 342. We have gross domestic product with all industries. Total in California. That's the chart. Also a lot information on consumer spending and employment income. Some reason with outdoor recreation. It is in there though. They also have a lot of industry facts. This is what it looks like. You can select a state. This is where it tends to only have one as an option. This is my economic profile. Going to the count. Going over into the more detailed additional options. If I look at the gross domestic product then it will give me some additional ones. I find that the quality are going to have a majority of information that I want. Looking at multiple industries. I go for

California and the table. They will ask for your time. Going with more than one. I have the data for agriculture. We can see if some industries are emerging or maybe some are losing interest. If you decide you want to change this then you can modify it. You can go directly to the personal income. There's a lot going on here is a fun place to explore. When is the next census data coming out? That depends on the project. It could be released if it is along the lines of the 1950s information. There are so many products that they are coming out with. The American housing survey and they will be coming out in 2023 and 2024. They have this project calendar here. That will give you a list of a lot of the different products that are going to be released. I think it takes a little bit longer than we would like for some of these to be available. Let's take a quick look at the environment. Here's a question. Used by decade and state in the U.S. Where should we start. EPA, USGS, or USDA? Everyone will have a different perspective when it comes to pesticide use. This is going to be interesting. A lot of people are saying USDA. We got a couple USGS. That is the U.S. geological survey. Of you for the EPA. Again. No wrong answers. I started with the USGS. Part of that is because they are interested in water. It turns out they are also working with the national oceanic and atmospheric administration. Here on this quality assessment. They have a program looking pesticide use. The statistics that they have expanded from 1992 to 2019. I am hoping they will up date that a little bit. It is also a little bit fun. For a lot of this information it's actually specific to compounds. This is what their site looks like. It will be estimated with the pesticide use. If you look at the county levels. You can get quite a few information you can only do one type of pesticide at a time. If you have someone who really needs to know about one particular pesticide being used then that is a great place to. This is what it looks like. This is specifically for the use of them. It gives you this nice map of the United States. Also gives you the use by year. You can get an idea of what plants are being exposed to specific pesticides. They have a different interest. They are not interested in water quality. They are trying to help people grow things. They have the national agricultural statistics service that has a lot of fantastic data. They also have the survey of agricultural chemical use for the program. This is a program that will help with a lot of information about management and chemicals. Some of the information is specifically 1990 2000 or are available in the quick stats. It's a great report in the U.S. agriculture. 21 select crops, nine teen 62 2008. That is a great place to go. This is a quick snap of what it looks like. When I was using this I got very excited because look at all of these different things. You can only pick one with the euro. There is an area to show you. This is going to change all of your options in the other fields. In this case I went environmental, field crops, and then you have to choose one or the other. You can't do multiples in any of these. Corn, corn pest management. Again, showing the program with the state. It's going to be somewhat limited. This is what it looks like. The location in the commodity. They are using the pesticide. A lot of what they are interested in the economics with how much it is costing. We are only getting a few states to be a part of the program. That information protects us. If we go into that report of the select crops then I get this nice graphic. It shows other pesticides. If you just wanted to see how much of a pesticide has been used over time the this can be good. Is there any information on how much fertilizer is used? Can it be broken down by state? Is there international information? Unfortunately, I did not get to the international portion but you can use Quickstart for this. In this case for the sector then you use economics. That will change some of your options. You can specifically go for extensive and fertilizer totals. You can go for state and amount. You can come from the U.S. economic research service. There is also a publication. Price tables in a single workbook. You can see from here that it is giving multiple states. This is not going to give you a ton of fertilizer but it will give you some data. How about hunting statistics by type of animal, region and across time? Fish and wildlife service. We have national fishing, hunting and wildlife associated recreation. They have been coming out soon 1955. Then officially hunting. Looking at the pesticide. I just wanted to see the percentage. Let's talk about our health. Where do I find statistics on the rate of births and deaths of the United States? We have the CDC? They have a lot of different products. I know there's a lot of people coming up with it right now. I'm guessing that is the hint of the vital statistics system. The

2022 report. The number that is increasing or declining. How it has changed from year-to-year. There's also the vital statistics of the portal that you can go to. That will give you a bit more freedom in what you are selecting. And then the totality United States report. This one is from 2021. That will give you information about life expectancy. Unfortunately it is also killing everyone. They have a lot of great and if a graph charts. They have life tables. In the case of the life table of 1866 to the present. There is a general map of the vitals statistic report. Some of these go back to 1890. There is a lot of content here. How many people in the U.S. broken down by racial categories have received the initial two COVID-19 doses? Broken down by political party categories. This question has a tiny bit of --. I cannot imagine the backlash that would happen if the CDC or any other government agency tried to break it down by political party. That would cause a bit of problems. The CDC does have the COVID-19 vaccination in the United States. The vaccination distribution and covers report. They will have information about the demographic. In the location. What you can do is get the voting. Going into the format that you want. It might take a little bit more work. There's an article recently published on this. I am very curious about it. Budgets and defense. Looking for this question. One of them is U.S. spending.gov. This has infographics. It has government spending. All the different agencies. What you do is start off with the agency and then narrow it down. You can have the Department of Defense and select a section. It will get smaller and smaller. If we are looking for how much for the Army. We can get the piece of information. This will also equally link you to the source content. These federal agencies have some sort of area. They have the defense.gov. They have a lot of different reports. The acquisition cost by the weapon system. There is also the national budget estimate known as the green book. The budget request are very consistent with URLs. If you find one of these and it has the year in and you can change the year to pull up the different report. Where can you find information about the U.S. budget? There is also of info.gov. Then they have the budget the United States government and other things. The office of management and budget. The office of the treasury also has a lot of content. They have financial guide to give you a lot of content as well. That is for this particular graphic. Interested in helping people understand how the money is being spent. One place you can start is the CRS. They will always have something updated on the U.S. budget. I found the great figure here. It is telling me the sources for the historic tables. I can follow that. Go to the Whitehouse.gov and find the budget section. Finding all of these tables. Going back to 1789. A nice source of information. What percentage of the United States Marine Corps pilots leave the U.S. Marine Corps annually? I had no idea about this. I went to one of my other go to places which is the Government accountability populace. They have so many reports that are available online. They are easy to search. They have something for the topic. In this case I found the military personnel and evaluating the pilot workforce requirements. Also led me to the term of retention. Also a technical as a report. Once you have found the terminology that he gets a lot easier. Labor and income. What is the average pay difference for men and women in various professions? The labor statistics have more information of labor. They have a section of subjects. Going by demographics. They have a table support. Trying to find this content. They have a wage statistic. They also have special warts for the issue. If you want to search for something like the earnings that you can earn this in 2020 the also have a report for median usual weekly earnings of the full-time wage and salary workers. Detailed by 2020 and over. We have the year and the number of workers. Then we have all of these occupations. A quick look at Justice. Can one find U.S. government statistics on the number of arrests that never go to trial? So, this is one of the reasons why I have screenshots. This great site is currently unavailable. I have the same thing happened to me so screenshots. There is the crime data Explorer. This is the database you can use to find this information. One thing it found was 9290 for present our results through the plea deal. You can also go to the Bureau of Justice statistics. I know I'm going through this a bit quickly. This is one of those areas where it is not a problem to ask. You can see them to judge the cases. A total. I was wondering if I could take the number of jury cases and the number of judge hold cases. Subtract that from the total cases. Will I get basically a number of plea deals? I reached out and, yes, the person at the

court statistics project was able to confirm that for me. Sometimes that can be very hope for researchers. Saying this is what they are able to find. This is what we advise to do in order to get that content. One of the more challenging questions is that I need a lot of data. A lot of data over time. There is the growing question about residents of Texas in 1900 whose mother tongue was German and set. Who were natives of Germany and where were they located? The U.S. Census had special volumes that were for each state. That will give you a lot of information for the state. Some of these have specifically what language people were speaking and where they were coming from. I would say in the 1990s or 1900s. They were all interested in immigration. They would have those types of questions. He was the 1920 census. Composition characteristic. Then German. Going for the entire state. We are seeing various cities. Going to it somewhat different. Going with a stem degree. Going around 2001. Taking some evaluation as we find it. Collecting the information. The terminology that they were using. The reports and were they consistent? Does the information exists in that. In this case, there's the National Science Foundation. They have a survey of college graduates, scientists and engineers statistical data. That goes back to about 1993. Going with the education statistics which is a rockstar. He has a lot of content and it goes to 1990. They have the number of post secondary institution. It's going to be a more specific title then stem. In fact, one of the tables is conferred by the institution to the field of study. They do have some of this information sometimes it is not going to be using the same terminology. This is the section from the 1950s. It is specifically engineering a non-engineering related. You might have to combine and go through a lot of different tables to find what you are looking for. Ordinary and down the question. If you just want to go by education and that is a lot easier to do. You want to know which agencies are handling this. Office of education and under Department of Interior. Office of education security agencies. On and on to the Department of Education. This will give you an idea of who was caught that information and when. Here is that digested educational statistic. I love them. They put this information together. Going back to 1869 and 70. It is not super detailed. It is a doctoral degree. It is still a great deal. How does one find income to poverty ratios for rural New York State counties? You can go to the census and look at income and poverty. Then County for that state. When you do that you should get some income levels. Work with the definition to decide what levels are going to be there. We have this county here. You can choose as many counties as you would like. I don't know how amazing the data into a visualization is but I put it together to basically mimic the expensive database. It is all Biddle Joy's Asian visualization. There is also search strategies. The U.S. government start as well. They are also pretty good. I know that we just covered a lot. I think we have a few minutes for questions.

Thank you, Susanne. A fantastic webinar. I really loved it. Any questions for Susanne? We are short on time but we do have time for questions . Please put them in there. Going into the chat.

There is a question about why did the census discontinue the statistical abstract? The official was budget. They decided that they wanted to refocus and this was also 2012 with the last one. There's a lot of budget cuts going on. That was something that did get cut. I think it was about the information that people had a bit of outrage. We really relied on that.

Any more questions for Susanne?

Remember, too, the slides will be available. There is a handout that charts what I put together. You can always come back and take a look at some of the screenshots. Remind yourself where you can go. If you have any questions as well then I'd be more than happy to answer those. At least to the best of my ability.

Please fill out the webinar satisfaction survey. I think you may put in their links to our training repository. This webinar and the others are there. Here's one. Kendra, what is a good source for active transportation like bicycling use?

There could be a couple. The department of presentation should have some information about. Also there is a specific question with the ACS that asked how people can commute. I think bicycling is one of those options. You might be able to see how much commuting won't necessarily be recreational.

Any more questions for Susanne? We have a bit of time.

I would say to take a look for any national bicycling associations. They will be the ones that are definitely collecting statistics. They may also have a break down of the number of communities that are increasing by planes. Trying to improve the bicycling culture in that location. Such a great question.

Best bet. International statistics generally.

I did not go into the international. They go to part three. I believe they go to the U.N.

Any more questions? Making some good comments. Check that out if you would. Keep those questions coming and we will make the best use of our last four minutes. I do have to end up three. We do have a bit of time.

I would say that if anyone in the audience thinks they are a super searcher then let me know. I will be happy to talk about resenting that. Putting my email in the chat. Send me a message if you would like to do one of the super searcher presentations. If you have expertise in some subsection. I think I will line up for the wall. Let me know. Any more questions for Susanne? We have a few more minutes. We do have three more West webinars today. We have one on the 23rd with the consensus 101. It ties in with this webinar it should be a great one. Questions. Also in our repository is the fall and spring conferences. All the presentations are available to check out. Give that a look. Getting all bunch of shout outs. Thank you. Okay. Three more minutes. I don't want to shortchange anyone. This slide deck will be used greatly in the community. I'm sure the audience members will tell others about it. It will --

There is a great question about the effects of bank loans. That's a good question. There is a civil rights commission. I suspect that they probably had hearings on redlining. That could be a good place to start. They will have testimony from people who have been researching this. They probably would have to have information from HUD and a number of other organizations. That is also tied into banking. One of institutions regarding banking. I have a feeling that is also something that people don't like to talk about. I would definitely take a look at the civil rights commission hearings to see if there's anything there. Possibly taking a look at search redlining in info.gov.

Thank you, Sudan. They are asking about whether we will have presentations around the fall. I'm assuming the summertime. We will issue an announcement on that. There is a link for the redlining information. Any last questions? I have one more minute. I probably should close out.

Fantastic, webinar. Let me closeout. Thank you so much, Susanne. A terrific webinar. Anyone else? Come back to follow. I volunteer. Thank you, Ashley for the tech support. Thank you audience. Come back for another one. Come back for the census 101 on Tuesday. Thank you and have a good day. [Event Concluded]