**Data Literacy Instruction for SparkMap Users**

Authors: Shaye Morrison and Sarah Pauley

# **I. Background and Statement of the Problem**

SparkMap is a subscription-based service where customers (typically public health professionals, local government employees, and nonprofit organizational members) can access cleaned, analyzed, and benchmarked data to create customized maps and community assessments. The goal of SparkMap products is to use data-based maps and assessments to show trends in a community and areas for improvement. However, we consistently find a pain point for users on SparkMap is not understanding how to read and interpret the data. For example, they have a hard time understanding the source of the data, which datasets can be used together and compared, and how to understand the differences between geographies that data sets cover. The goal of this project, then, is to improve data literacy among SparkMap users so that they have a more consistent and confident experience interpreting and applying SparkMap data. To do this, we will use Census 2020 data and the Map Room as an example to guide learners in understanding data literacy, interpretation, and visualization.

# **II. Needs Analysis**

A. Overview

In order to create an effective data literacy training for SparkMap users, we needed to learn about the characteristics of current SparkMap users, the usability of the SparkMap platform/ data, current challenges experienced by SparkMap users, and SparkMap’s goals for users. Additionally, we needed to explore broad definitions of data literacy and how data literacy instruction can be implemented in various settings.

Data literacy in general is a normative need. Data literacy can be defined multiple ways, but there are certain standards that individuals should meet to be considered data literate. When individuals do not meet those standards, there is a normative need. “Data literacy” can encompass a wide variety of definitions, from the very broad to the very specific. In the tutorial “What is Data? What is Data Literacy?” librarian Meryl Brodsky defines data literacy as “the ability to think critically about and use data.” In a lecture, APAC Technology Evangelist Mac Bryla (2018) argues that the definition of data literacy should be expanded beyond simply the ability to read a chart to include critical thinking and questioning skills. Bryla recognizes seven components of data literacy, defining it as the ability to 1) derive meaningful information from data and argue with data, 2) question whether data is useful, 3) question whether data is correct, 4) understand how aggregations were calculated, 5) understand basic statistics, 6) use facts to communicate complex ideas – without misleading the audience, and 7) the ability to use all the facts, including missing data.

As Bryla argues, statistical literacy is one component of data literacy. Gal (2002) defines statistical literacy as “the ability to understand and critically evaluate statistical results that permeate daily life, coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions” (p. 2). Gal points out that, while formal instruction (e.g. high school or college math and statistics classes) tends to focus on pure statistics, adult consumers are most likely to encounter statistics in the context of larger articles or reports; reading contexts include TV or newspaper reports, advertisements, handouts at community or political events, and data included in workplace reports. The fact that data is integrated into these reading contexts means that people must learn how to synthesize information from multiple numerical and statistical representations with written information. Gal breaks “statistical literacy” into five components: literacy skills, statistical knowledge, mathematical knowledge, context knowledge, and critical questions. These components overlap: literacy skills may refer to understanding data presented within a paragraph, while statistical knowledge may refer to ability to read charts and graphs. To a large extent, both types of knowledge require users to be familiar with key vocabulary: terms associated with percent (percent change, percent chance, percent above or below, percent gain/ loss, times more/ less likely), terms associated with data collection (random, representative, reliable, margin or error, sampling error, biases), measures of central tendency (mean, median, mode), and other numerical/ statistical terms (fractions, part-whole relations, odds, ratios, skewed, etc). To be statistically literate, adults need to be aware of the ways in which data can be presented or interpreted differently: median may be a more accurate measure of central tendency when the mean is skewed by extreme outliers; reporters may use modifiers before percentages (e.g. “only 10% vs. “a full 10%”) to emphasize or de-emphasize data. Finally, users must be willing to think critically about data and to “adopt, without external cues, a questioning attitude towards quantitative messages that may be misleading, one-sided, biased, or incomplete in some way” (p. 18).

Among SparkMap users, data literacy can be considered both a felt need and an expressed need. SparkMap’s help requests show that users need assistance reading and interpreting the data. Because users are requesting help in these areas, there is demand for a training program.

B. Data Methods and Analysis

We collected data through five main processes. First, we reviewed past SparkMap customer surveys to interrogate consistent roadblocks with understanding SparkMap data. Secondly, we used historical support inquiries to develop an understanding of what kinds of data questions customers have and where points of confusion lie. Third, we conducted interviews with SMEs to understand more about the data within the product, as well as the types of data literacy questions customers routinely ask. Fourth, we conducted an immersion exercise to explore any challenges or questions that arise for novice SparkMap users. Lastly, we conducted a literacy review to understand more about Census 2020 data, data literacy in general, and data literacy education, to inform our instructional development.

|  |  |  |
| --- | --- | --- |
| **Data Source** | **Analysis Informed** | **Sample/Data Size | Sampling Method** |
| Client Surveys | Learner analysis  Orienting Context  Performance Context | -sample size: We need to go back and review this information  –method: Shaye reviewed SparkMap’s client surveys, which were previously obtained by SparkMap |
| Historical Support Inquiries | Gap analysis  Task Analysis | –sample size: We need to go back and review this information  -Method: Shaye reviewed SparkMap’s historical support inquiries |
| SME Interviews | Task Analysis  Performance Context | –Shaye interviewed two individuals who work at SparkMap |
| Novice Immersion/ Observation Exercise | Task Analysis  Gap Analysis | – sample size: 5 (Sarah + 4 volunteers)  – all volunteers work in health, mental health, or education-related fields; close to SparkMap’s customer base  – informal observation and interviews with volunteers  – novice users read/ interpreted a sample SparkMap map  – limitation: small sample size, informal and somewhat inconsistent observation/ interview methods |
| Literature Review | Orienting Context  Task Analysis  Instructional Context | – sample size: 9 articles, videos, or instructional presentations on adult data literacy; 8 articles or webpages from the US Census Bureau |

C. Gap Analysis

*Data literacy – including reading and interpreting data drawn from maps can be a challenge, even for educated professionals. While basic statistics is often a required college course, many people may go through college without gaining much experience with data visualization or maps. Thus, when users subscribe to a service such as SparkMap, they face a learning curve.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Current State** | **Desired State** | **Root Cause** | **Evidence of Root Cause** | **Strategies to Address** |
| Learners do not understand the background information of Census 2020 data sets (e.g. survey type, data source, date of collection, geography of data, etc). | Learners will be able to identify background information (e.g., survey type, data source, date of collection, geography of data) on a Census 2020 data set needed for interpretation | Learners do not know the definitions of terms such as ZIP, ZCTA. These definitions are not currently provided directly on SparkMap tools. | Historical support tickets | –provide definitions via job aid or video tutorial  – give learners a self-paced knowledge check to review knowledge |
| Learners are uncertain how to read SparkMap maps. They do not know how to use to key to translate the visualization into words. | Learners will be able to use the map key and search tools on a SparkMap map to extract necessary information. | Learners are just learning how to use SparkMap’s tools. They may be unfamiliar with the SparkMap’s layout or with reading choropleths (maps) in general. | Historical support tickets | –provide a tutorial that visually and verbally walks learners through the process of accessing SparkMap tools and extracting information  –provide a knowledge check that requires learners to read a map to answer factual questions |
| Learners are uncertain how to draw conclusions from maps or tend to draw incorrect conclusions. | Learners will be able to interpret maps to draw and state conclusions about a given population (e.g. compare data from multiple counties to draw a conclusion about broader trends in a state). | Learners may have limited experience with using maps to draw conclusions. They may be unsure what questions to ask to analyze data.  Some maps visualize information in a way that may be unfamiliar to many users or that requires users to think about concepts in a different way. | Historical support tickets  Novice immersion strategy | –expository approach; provide rules of accurate and inaccurate data interpretation and then offering examples and non-examples  –open-ended questions and a self-check; questions require learners to use a map to draw conclusions; learners check their answers against the correct interpretation |

i. Implications of Gap Analysis for Training Design

We have combined learner needs into three categories that are based on Bloom’s taxonomy. First, there is a gap in lower-order thinking skills, such as identifying and understanding basic terminology from US Census data sets. This gap can be addressed through a brief training and job aids. Secondly, there is the challenge of correctly reading maps. Before they can use SparkMap data for advocacy or decision-making, users need to be confident that they are accurately reading our maps. Users need to a) be shown how to find necessary information and translate visual information into sentences. They also need to practice answering fact-level questions. This need matches the application level of Bloom’s taxonomy. By following along with a tutorial and checking their answers to fact-level questions, users will gain the experience and confidence needed to start drawing conclusions with the data. Finally, users need to be able to draw accurate conclusions in order to make full use of their SparkMap subscription. Users need to be able to analyze and interpret information that relates to their field and their community. Our training will provide practice by having learners look at data sets that relate to one or more common fields (healthcare, local government, etc). We will write questions that require users to draw conclusions about a data set. We will also provide the correct answers. This will allow learners to come up with their own answer and check their work against the provided answer. There are also many pitfalls novices face when interpreting data sets, so we will also provide examples of incorrect interpretations. By following along with multiple examples and nonexamples, learners will gain understanding of how to apply their knowledge in a variety of situations. Our goal is for learners to fully utilize SparkMap’s tools for their sectors and communities. We believe that this training will result in fewer support requests over basic data literacy concepts and more independent data interpretation.

D. Learner Analysis

We used SparkMap’s user data and surveys to obtain information about current users.This information was obtained prior to the start of this project and was current as of February 2022.

|  |  |  |
| --- | --- | --- |
| **Category** | **Data Sources** | **Learner Characteristics** |
| Entry behaviors | We still need to review this information. | We still need to review this information. |
| Prior Knowledge of Topic | Historical SparkMap data?  User surveys? | – Majority of users have a baseline understanding of what data is |
| Attitudes toward content | Historical SparkMap data?  User surveys? | – SparkMap is a subscription-based service. We believe that users are subscribing to SparkMap for a purpose and find inherent value in the tool and likely in data.  –Support requests do indicate that users may find data interpretation overwhelming and complex. |
| Attitudes toward delivery system [technology, training format] | Historical SparkMap data  User surveys | – SparkMap is an online platform. Users work in a variety of fields in a variety of geographic locations. Therefore, a web-based training is an appropriate training format. |
| Motivation for Instruction | Historical SparkMap data  User surveys | – Majority of users use SparkMap’s services as part of their job; therefore, it is likely that they need to be able to interpret this data in order to make decisions or present information in their jobs. |
| Demographics [specifics vary by project- age, gender, years of experience in job, etc] | Historical SparkMap data  User surveys | – over 1000 users since January 1, 2022  – 68% female; 32% male  – tend to skew younger (40% aged 18-24; 21% aged 25-34; 17% aged 35-44; 12% 55-64; 3% aged 65+)  –vast majority live in the US and speak English  – most users work in public health, local government, nonprofit, or education-related fields |

i. Implications of Learner Analysis for Training Design

SparkMap’s users are individual subscribers, based in a variety of organizations and geographic locations. They tend to be professionals working in health, local government, nonprofit, or education-related fields. They have chosen, either independently or with encouragement from superiors, to subscribe to SparkMap in order to visualize data that is relevant to their job. These factors mean that an asynchronous, self-paced, online tutorial approach is the best way to deliver our training. This way, both current and future SparkMap users can access this training whenever and wherever they need it.

**E. Contextual Analysis**

Our training will be an asynchronous, self-paced, online set of tutorials and knowledge checks. The target demographic will be professionals working in health, local government, nonprofit, and education-related fields. For the orienting context, we analyzed how data visualization is used in these fields. For the instructional context, we focused on data literacy needs among American adults and how to best deliver content in an online, self-paced context. The performance context is closely related to the orienting context in that learners will be expected to apply their knowledge to create reports in their separate fields.

a. Orienting Context

As shown through the learner analysis, SparkMap’s users are a diverse group, not centered in one organization or even one geographic location. Rather, individuals from various organizations have chosen to subscribe to SparkMap in order to better interpret and visualize data for their jobs. SparkMap users do tend to represent four main sectors: healthcare, nonprofits, local governments, and education. Users in these sectors might use different types of data for different purposes. Historically, health organizations have used data and maps to pinpoint the epicenter and stop the spread of infectious disease. Examples span the 1850s, when a physician used maps to determine the cause of a cholera outbreak, to the 2020s, with the now-ubiquitous Covid-19 tracking maps. Healthcare providers may also use data to explore the relationship geographic area and risk factors for or prevalence of a certain disease. Local governments, meanwhile, are more likely to focus on issues that can be addressed by local governments: traffic, housing, roads, policing, etc. Governments need to use data to understand pinpoint areas most in need of improvement in order to make decisions regarding funding and staffing. All sectors may use data visualization for advocacy.

These different sectors also overlap in what data they use and for what purposes. A recent article on the SparkMap blog discusses the health, environmental, and economic impacts of commuting (Krohn). Healthcare providers might explore this data to understand the needs of patients in their area. Meanwhile, local governments might focus on the economic impact of commuting, while nonprofit organizations might use the data to target community services that would improve the quality of life for commuters. Another blog article, “Heart Disease – Designing Community Interventions” demonstrates how the goals of the health and nonprofit sectors (and possibly local governments and education as well) map overlap. This article uses several data sets to explore the physical risk factors of heart disease, as well as social determinants, such as food deserts. Food deserts are an actionable area for community intervention, where multiple sectors might work together on a solution.

In general, the underlying need of SparkMap users to understand terminology, data sources, and visualization is the same for all users, regardless of specific sector.

Current user data indicates that SparkMap’s current subscribers tend to be individual employees, rather than entire organizations. This means that our data literacy training needs to have buy-in from these individual users, but not necessarily from managers, at least not at this initial step. Our learner analysis has shown that there is a demand for self-paced trainings on interpreting and using data, so the buy-in already exists.

b. Instructional Context

Because user data shows that SparkMap tends to be used by individuals at organizations rather than by organizations as a whole, asynchronous, self-paced online instruction is the best delivery format for this instruction. Online asynchronous trainings may include a combination of video explanation and examples followed by practice questions with immediate feedback. Two tutorials created by business librarian Meryl Brodsky for Eastern Michigan University Library provide excellent examples of this kind of instruction. One tutorial, entitled “How to Read a Chart” provides examples of different kinds of charts along with explanations of how charts are used and direct instruction on how to extract meaning from a chart. For example, on a bar chart, readers should look at the title, identify what information each axis represents, identify the range of values covered by each axis, look for patterns or trends (such as overall upward trends, looking for exceptions or outliers), and finally read the data on the chart; Brodsky also points out that users should take note of the data’s source. In the tutorial entitled “What is Data? What is Data Literacy?” Brodsky provides examples showing how a user’s understanding can be manipulated by the way data is presented. For instance, it displays two charts depicting average global temperature from 1880 to 2014. In one chart, the y-axis goes from 0 to 100, which means that the reader can’t see much variation over time. However, if the y-axis range is from 55 to 60, it is clear that the trend has steadily increased over the past century. Both videos include multiple choice practice questions at the end. “What is Data Literacy?” tests learners over key vocabulary, while “How to Read a Chart” asks questions about sample charts. The United States Census Bureau also provides brief tutorials explaining various census-related concepts (e.g. census-designated places, census tracts, and census blocks) using videos and drawings. However, these videos are intended for brief, informal instruction and are focused on defining vocabulary rather than helping learners practice or apply skills.

Data literacy instruction can also be presented in writing alone. For example, Emerson’s pamphlet “Visual Information for Advocacy”illustrates the diverse types of charts and visual representations non-governmental organizations can use to present a message to their target audience. The pamphlet includes some contemporary and historical examples of maps being used to drive policy decisions. For example, in 1859, a map showing the location of water pumps and cholera deaths in a London neighborhood led to a pump being shut down – and thus to the end of the cholera outbreak (p. 13). One challenge with this text-only approach is that users have to make more inferences than they would when watching a screencast or video that actually highlights specific parts of a map or chart during the explanation. Written instruction can be more effective when it breaks a process down into steps. For instance, in this blog post entitled “Data Literacy: Top Five Tips for Interpreting Maps,” Oldrey lays out five simple, yet crucial tips for interpreting data maps: read the legend, pay attention to scale, don’t enter with assumptions, enter with focus, and filter with caution. Oldrey mentions key vocabulary and anticipates common user assumptions, which is helpful. However, overall, articles and pamphlets do not provide learners with the same level of multimodal and interactive elements as video tutorials do.

c. Performance Context

In terms of performance, the majority of SparkMap users would implement the instruction in the form of reports to improve their communities. Whether that be nonprofits, local government, health consultants, public health departments, or other types of community leaders, often data from SparkMap is used as evidence for grant funding packages, to demonstrate areas of community need, or to increase education on community deficiencies and areas of excellence. Because there is such a variety of users it is challenging to articulate one specific context in which the instruction would be applied. Generally, however, improved data literacy could be applied in a few specific ways. First, data literacy instruction would result in easier navigation of the SparkMap Map Room tool. With an understanding of how to read a data set and ascertain information about how data was collected and what it specifically focuses on, users will better understand maps of selected data layers. Second, the instruction will help learners integrate data into reports and grant applications. After going through the training, learners will be armed with information on how to draw conclusions from maps and formulate them in a way that supports the problem they want to address. Overall, then, the main performance context of interest will be creating a SparkMap map and distilling the information into an accurate and clear argument.

ii. Implications of Contextual Analysis for Training Design

The primary implication of the contextual analysis is that our training needs to use data sets that are relevant to the performance contexts of our learners. That is, we need to use data sets related to health, education, and community needs or services. The training design may need to use a framing device that puts the learner in the shoes of a person in a job similar to their own (e.g. a community health expert). At the same time, it is important that our training is accessible to learners in all fields; in order to focus on data comprehension and interpretation, we need to avoid any field-specific jargon in our training.

# **III. Task Analysis**

[Describe methods for collecting task analysis]

A. Facts/Concepts

Data Literacy Background

SparkMap’s data is presented in a type of map called a choropleth, which is a map where each administrative area (e.g. nation, state, or county) is filled with a different color to show the aggregated statistics of that area. In an article geared toward helping cartographers create effective, trustworthy COVID-19 choropleths, Juergens (2020) explains three variables shown on thematic maps: map project, scale or level of detail, and classification method. Juergens points out that data can look different depending on the granularity of the administrative area: a larger area (e.g. state-level) can show the bigger picture, while smaller units (e.g. county-level) can better reveal regional differentiation; in the case of COVID data, smaller units better display hotspots. A well-designed map maintains the same level of granularity throughout the map; for example, a world map should not represent some nations as single units while dividing others into provides or states. Classification method encompasses the number of thematic classes, the way classes are defined, and the color scheme applied to distinguish between classes. Ideally, choropleth maps use between four and six classes. These classes should be defined in a way that allows users to make comparisons. Absolute numbers should be normalized against the population (for example, converted into cases per 100,000 inhabitants). Knowing the different factors that go into creating effective maps can help users better interpret those maps.

The U.S. Census is a major data source for SparkMap; thus, it is important for users to understand census terminology. The U.S. Census Bureau’s Census Academy provides videos explaining key vocabulary and demonstrating how to access different data sets. While SparkMap eliminates the need for users to access data sets directly from the census website, Census Academy provides helpful background on census-related vocabulary. For example, it is useful for SparkMap users to understand the various administrative units used in census data: CDP (census designated place), census track, census block, and ZCTA (ZIP code tabulation area). These terms are explained in videos listed in the bibliography and defined in the list below.

* **US Census Bureau:** The US Census Bureau is one of many government agencies dedicated to collecting data on the economy, demographics, and geographical dispersion of Americans (United States Census Bureau, 2017). In fact, this is the largest statistical agency in the United States. The goal of the Bureau is to provide thorough and accurate data about America and its people.
* **Decennial Census:** an account of population and housing in the United States that occurs once every 10 years (United States Census Bureau, 2017). In this survey, people from all states and US territories are counted. This census is incredibly important because it “determine(s) the number of seats for each state in the U.S. House of Representatives and are used to draw congressional and state legislative districts and to distribute hundreds of billions of dollars in federal funds each year” (United States Census Bureau, 2017, para. 6).
* **Demographic surveys:** those which measure information regarding socioeconomic status, race, poverty, education, housing, and other subjects (United States Census Bureau, 2017).
* **Economic surveys:** surveys of specific parts of the national economy. These occur on monthly, quarterly, and annual bases (United States Census Bureau, 2017).
* **Sponsored surveys:** demographic and economic surveys collected for other governmental agencies (United States Census Bureau, 2017).
* **Category:** the type of data within a data set. Examples include demographic, economic, education, housing, etc.
* **Indicator:** the content being measured. For example, total population.
* **Data source:** The agency, company, and specific timeframe of the data. This would indicate results from a specific survey.
* **Source geography:** the geographic boundaries a data set is available at. Examples include census tract, county, Zip code, ZCTA, state, and address (defined below).
* **Census tract:** subsets of a county or other geographic area that is updated based on the decennial census (United States Census Bureau, Glossary).
* **Zone Improvement Plan (ZIP) code:** areas of land grouped together based on US Postal Service Delivery routes. These are determined by the US Postal Service.
* **ZIP Code Tabulation Area (ZCTA):** geographical areas based on US Postal Service delivery routes. They may or may not be the same as Zip Codes, as they are determined by the United States Postal Service (United States Census Bureau, 2018, July 2).
* **State:** A geographical area that falls within the boundaries of a particular state.
* **Address:** the particular place where someone lives or an organization is situated (Oxford Languages, 2022).
* **State benchmark:** information that compares state level data to national level data.

B. Procedures

For the purpose of this training, we are going to learn how to interpret data using the SparkMap Map Room and using an example data set from the US Census Bureau 2020 Decennial Census.

1. In your web browser, go to [www.sparkmap.org](http://www.sparkmap.org)
2. Select the button in the top toolbar that says “Get Started”
3. Scroll down and select the button that says “Make a map”
4. In the “Search Data” pop up box, type in “Census 2020”
5. Select the “Total Population, US Census 2020” data set. Once you click on it and the box is checked, select the blue “Add to Map” button at the bottom of the screen. You should then see a map populate with predominant race data.
6. In the “Map Layers” box, you will see the key to interpreting the map. On the top of that box, there are 3 dots, “...” Click on these dots and select the “data info” button. This will pop up a box with information about the data set.
7. Gain information about the data set.
   1. The description at the top tells you what the data layer displays. For example, in the Predominant Race layer it describes the racial categories included in the 2020 Census.
   2. Release date: This tells you the month and year that the data set was released by the source. This is important because it helps you grasp how recent the data is. Since this is the decennial census, we know the data released in 2021 is the most recent.
   3. Release cycle: If you do not know how often a data set is released, this will tell you, which helps to determine the recency of the data you’re using.
   4. Data geographical unit: This tells you the geography of the data set you are looking at.
   5. Data source: This is the organization that collected the data as well as which specific survey the data was sourced from.
   6. Data source description: This tells you about the data source. For the example data set, it provides information on the U.S. Census bureau and what the decennial census determines.
   7. This information is important to understand because it provides context of your data set. Understanding background information on a data source, how often the data is collected, and which geographies it is collected at provides you with the knowledge to decide what kinds of data you can compare and what conclusions can be drawn from the data.
8. Once you have gained information about the data set, exit out and click on your location of interest on the map or search it in the toolbar on the right that says “Enter a location.” For the purposes of this demonstration, we will use St. Louis, Missouri.
9. Once you select your location, in the map layers box, you will see the geography defaults to county. If you click on St. Louis City, you will see that the population in the 2020 Census was 301,578 people. However, if you change the geography, you will find that the total population increases. Within the total population data set, you have the ability to select state, county, census tract, or block group geographies.
10. Change your geography from county to tract. Once you do this, you will see the boundaries in St. Louis change dramatically. If you click on St. Louis City on the map, you will see that the total population is 3,729 in Census Tract 1273. If you move to Census tract 1193 in St Louis city, however, you see the population is 6,100 people.
11. Interpret the data. Based on the map, you can draw a few conclusions. First, you can determine that the total population of St. Louis City in the 2020 Census was 301,578 people. By investigating further, you are able to determine the most and least populated census tracts within this county. Therefore, you can make determinations on which areas might be most or least impacted by your community interventions.
12. To add a greater level of understanding, add additional data layers. To do this, select the blue “add data” button in the top left corner.
13. Again, search “Census 2020” in the search data box.
14. Select the “Population Living in Group Quarters - Correctional, US Census 2020” layer and add to map.
15. When comparing information from two data layers a few considerations are important. First, look to see if the data source is the same. Since we used Census 2020 data for both, ours are the same. Next, Be sure you are examining the same geography. Set both geographies to county. We already know the population of St. Louis City is 301,578. If we switch our map layer, we see that 1,213 of those residents were living in correctional facilities at the time of the 2020 Decennial Census. Therefore, we can draw the conclusion that at the time of the Census, 0.4% of the population of St. Louis City was living in correctional facilities. We can then compare with surrounding counties. For example, St. Louis County has a total population of 1,004,125 people. At the time of the Census, 1930, or 0.19% of those citizens were living in correctional facilities.
16. Draw conclusions. Why does this matter? Let’s imagine you are working on a local campaign to reform prisoners. Using this data, you could propose to target efforts in St Louis City instead of St Louis County. Although some might argue that because St. Louis County has a larger population it should be the target, accurate interpretation of this data will allow you to accurately state that a larger percentage of residents in St. Louis City are incarcerated and therefore would benefit more from the intervention.

C. Attitudes

It will be important for us to consider that learners could have negative attitudes toward the content of the training. Data is something that individuals can find overwhelming and complex. Therefore, it will be important that we make the instructional materials as engaging, straightforward, and applicable as possible to keep learners interested.

We do believe that a benefit of these learners in particular is that data interpretation is part of their job or project at hand. Because SparkMap is a subscription-based service, users are coming here for a purpose and find inherent value in the tool, and likely in data. As such, it is possible they believe the training will be value added to their SparkMap experience and make accomplishing their goal easier.

# **IV. Training Design**

A. Broad Goals

* Improving data literacy–understanding data set origins, collections methods, and what types of data sets can be compared to draw conclusions.
* Improving data visualization and interpretation–knowing how to read a map
* Improving conclusions made with data–learners should understand how to take the information provided by data and be able to put it into reports to distill what it means for their community.

B. Overview of Training Design

* Because of the type of learners (e.g., from all different industries across the entire nation) the training will be conducted online in an asynchronous manner.
* Learners will walk through several different modules (as outlined in below in D) that take them through the steps of identifying background information on data, using the key of a map to extract necessary information, interpret maps to draw conclusions from the data, and write conclusions in their own words.
* Utilizing a storyboard format, learners will engage in watching videos, interacting with maps on their own, and completing interactive quizzes and knowledge checks for objective and subjective mastery.

C. Learning Objectives

* Learners will be able to identify background information (e.g., survey type, data source, date of collection, geography of data) on a Census 2020 data set needed for interpretation
* Learners will be able to use the map key and search tools on a SparkMap map to extract necessary information.
* Learners will be able to interpret maps to draw and state conclusions about a given population (e.g. compare data from multiple counties to draw a conclusion about broader trends in a state).

D. Learning Experiences

|  |  |  |  |
| --- | --- | --- | --- |
| Learning Objective | Instructional Strategy | Assessment Strategy | Time |
| Learners will be able to identify background information (e.g., survey type, data source, date of collection, geography of data) on a Census 2020 data set needed for interpretation | Presentation strategy: Provide definitions with examples and highlight similarities/differences between key terms (Table 9.7, p. 234)  Generative strategy: Provide scaffolding to help learners locate, assess, and validate information by utilizing a knowledge check (Table 9.6, p. 229) | After presenting the content, learners will complete a knowledge check using a multiple choice format to assess their ability to correctly identify background information about a data set. Questions will focus on identifying the correct geography level, survey type, data source, etc. Remediate wrong choices as needed (Table 9.4; p. 228). | 10-30 minutes, depending on user need for remediation following knowledge check. |
| Learners will be able to use the map key and search tools on a SparkMap map to extract necessary information. | Presentation strategy: Tutorial with both verbal and visual walkthrough. Provide a job aid with the steps to follow for successful map interpretation.  Generative strategy: Have learners reflect on a map provided and practice extracting information by answering questions | Learners will be presented with an interactive map and questions about it. These questions will be fill in the blank, rather than multiple choice to help learners apply information in accordance with Bloom’s Taxonomy. Examples of questions include having learners find the percentage of individuals below the poverty line in a county, or the rate of unemployment in a state. Remediate wrong choices as needed (Table 9.4; p. 228). | 30-45 minutes depending on need for remediation with knowledge check. 15 minutes of this estimated time is accounted for the tutorial and job aid. |
| Learners will be able to interpret maps to draw and state conclusions about a given population (e.g. compare data from multiple counties to draw a conclusion about broader trends in a state). | Presentation strategy: Utilize the expository approach (p. 221) by providing rules of accurate and inaccurate data interpretation and then offering examples and non-examples.  Generative strategy: Knowledge check | For the assessment there will be two parts, a knowledge check and self assessment. First, learners will be presented with examples of maps and data interpretation write ups, learners will indicate if the write up is a correct or incorrect interpretation of the data. Remediate wrong choices as needed (Table 9.4; p. 228).  Second, learners will engage in a self check where we will present them with a multilayer map and ask them to draw conclusions (e.g., compare the unemployment rate and average educational attainment in Missouri with neighboring states of Iowa, Illinois, and Kansas to conclude the state with highest economic potential). After answering, learners will be presented with our correct interpretation to check their answer against. | 20-30 minutes |

E. Assessment Plan

* For formative assessment, learners will engage in self-assessment, quiz, and debriefing exercises as outlined above. Since this is an asynchronous online activity, some of the more interactive formative assessments are challenging to conduct. However, we believe since these are adult learners who use data to some extent in their professions, performing self-assessment, quizzes, and debriefing exercises will be sufficient.

F. Implementation Plan

Logistics: This is an online asynchronous training

1. Technology needs: a computer with access to the internet; speakers to listen to audio components

2. Room layout needs: N/A

3. Resources and other materials needs: power adapter for computer to ensure you have access to the entire training

Timing considerations:

1. Prior to Activity: Prior to the activity, learners should enroll in a SparkMap subscription to have access to the tools discussed

2. During Activity: During the activity, learners should devote complete attention to the training in a distraction free environment

3. After Activity: After training, learners should utilize the job aid to create an interpret a map with data specific to their own needs. They should revisit the trainings and contact our team if any lingering data interpretation questions arise.

# **V. Evaluation Plan: 15 points**

A. Formative Evaluation

* For formative evaluation, we will engage in expert review and one-to-one exercises. In the expert review, we will speak with our SME’s and knowledge experts on the SparkMap team to ensure the product is meeting their goals and that all content included is technically accurate. In the one-to-one portion Sarah and I will each recruit someone with a basic understanding of data to walk through the training supervised. By doing so, we will gain clarification on points of the training that are confusing or ineffective to change before deploying the course.

B. Summative Evaluation

* As summative evaluation, we will engage in levels 1 and 2 of Kirkpatrick’s levels.
* For level 1, we will gauge learner reaction by including a satisfaction survey at the end of the training. In this survey we will ask learners to indicate rate with a Likert-scale how helpful the training was and how satisfied the learner was. We will also include two open ended responses. One open ended response will allow learners to include any other comments or suggestions on the training. The second open ended response will prompt learners to leave any lingering questions following the training which will be responded to by a SparkMap team member within 48 hours.
* For level 2, we will provide answers and full explanations on the knowledge checks throughout the training. Showing the correct answers will allow learners to understand which concepts they have mastered and which areas they should spend more time on.

C. Confirmative Evaluation

* To assess the effectiveness of the training, we will track SparkMap support questions regarding data interpretation as well as data citing. Since one goal of the training from the organizational perspective is to reduce customer inquiries about data interpretation, a trend in decreasing support questions about data interpretation would indicate success of the training. Further, we will also track the number of sources SparkMap maps get. If people are using and citing the data in the wild, it is indicative that the training is worth the cost.

**Instructional Approach**

The Data Literacy Instruction for SparkMap Users training course will utilize a fully remote instructional model in which users will engage with a variety of materials. Learners will complete three modules (two of which are depicted in this document) in the Canvas LMS to complete the training. Within each module, learners will engage in generative and summative assessments. The training is scaffolded such that mastery of each module will be applied in subsequent models. The goal is the current and future SparkMap users, dispersed across the country in a variety of professions, will walk away with an understanding of how to interpret data within SparkMap’s Map Room and synthesize it for reports and materials of their own.

**Estimated Session Timing & Delivery Method**

|  |  |  |
| --- | --- | --- |
| Module 1 | 30 minutes | Remote & Independent |
| Module 2 | 45 minutes | Remote & Independent |
| Module 3 | 45 minutes | Remote & Independent |
| TOTAL TIME: ~ 2 hours | | |

**Content and Activities**

The content and instructional activities are outlined in the tables below. You will find one table for each module. Within each table, you will find information about the learning objectives, instructional approach, assessment strategy, estimated time, and materials needed.

**Module 1: Identifying Background Information on Data and SparkMap**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Learning Objective** | **Instructional Approach** | **Assessment Strategy** | **Estimated Time** | **Materials Needed** |
| 1a. Learners will be able to recognize and define terms related to US Census geographic units | Presentation strategy: In a voicethread presentation, provide definitions with examples and highlight similarities/ differences between key terms.  Generative strategy: Give learners an opportunity to check their ability to recognize and apply terms and definitions via a knowledge check | Knowledge Check:  true/ false, multiple choice, and/ or short-answer questions that require learners to understand and apply the definitions of various terms. Learners will be able to submit their answers for immediate feedback (including explanations) via Google Forms | -presentation: 10 minutes  – knowledge check: 5 minutes | – handout of terms and definitions  –presentation (Google Slides, uploaded to Voicethread)  – knowledge check questions (on Google Forms)  \***Appendix A** – handout draft  \***Appendix B** - knowledge check questions  \*attached Google slides - presentation draft [**https://bit.ly/3kmzr7G**](https://bit.ly/3kmzr7G) |
| 1b. Learners will be able to recognize and define terms related to US Census data sources. | Presentation strategy: In a voicethread presentation, provide definitions with examples and highlight similarities/ differences between key terms.  Generative strategy: Give learners an opportunity to check their ability to recognize and apply terms and definitions via a knowledge check | Knowledge Check:  true/ false, multiple choice, and/ or short-answer questions that require learners to understand and apply the definitions of various terms. Learners will be able to submit their answers for immediate feedback (including explanations) via Google Forms | Presentation – 5 minutes  Knowledge check – 5 minutes | -handout of terms and definitions  –presentation (Google Slides, uploaded to Voicethread)  – knowledge check questions (on Google Forms)  **\*Appendix C** – handout draft  \*attached Google slides - presentation draft [**https://bit.ly/censusterms**](https://bit.ly/censusterms) |
| 1d. Learners will be able to find examples of various geographic units and data sources and validate whether an example matches a given term or definition | Presentation strategy: Learners may rewatch the presentations and use the handouts from objectives 1a and 1b as needed  Generative strategy: Give learners an opportunity to practice locating, assessing, and validating information via a knowledge check | Knowledge check: Given a sample data set and true/ false, multiple choice and/or short-answer questions, learners will be able to identify or check examples and nonexamples of various terms | –knowledge check: 10-15 minutes | –knowledge check  –sample data set (embedded, or with links or screenshots provided)  \***Appendix D** - sample knowledge check |

**Module 2: Extracting Information from a SparkMap Map**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Learning Objective** | **Instructional Approach** | **Assessment Strategy** | **Estimated Time** | **Materials Needed** |
| 2A: Learners will be able to create a map on SparkMap. | At the beginning of the module, learners will be directed to make a map using Sparkmap. They will follow step by step instructions articulated in **Appendix E**. | A self-assessment will be completed after this step, where learners will compare their map to the example map uploaded in the module to ensure they followed all steps correctly. | 5 minutes | Personal computer, access to the internet, access to [www.sparkmap.org](http://www.sparkmap.org/), access to Canvas training module. |
| 2B: Learners will be able to use and interpret the map key. | Once learners have made their example map, they will watch a VoiceThread tutorial on how to read and navigate the map key. The script of the tutorial is articulated in **Appendix F**.  \*Note, this will be visual and audio, so the walkthrough will have a cursor viewers can follow. | Knowledge Check: Following the tutorial, learners will use information learned, to complete a knowledge check. The knowledge check questions are a variety of true/false, multiple choice, and fill in the blank. The full knowledge check questions can be found in **Appendix G.**  Learners will be able to see which answers were wrong, but will not be shown the correct answers.    Learners will not be able to move forward in the training until they get 100% on the knowledge test. | 20 minutes | Personal computer, access to the internet, access to [www.sparkmap.org](http://www.sparkmap.org/), speakers/headphones to hear audio, and access to Canvas training module. |
| 2C: Learners will be able to find information about a location of interest on a SparkMap map.  2D: Learners will be able to extract information from a SparkMap map. | Once learners have completed the first knowledge check, they will be directed to another tutorial where they will learn how to find data regarding specific locations and interpret that data. Tutorial script can be found in **Appendix H.** | Knowledge Check: Following the tutorial, learners will use information learned, to complete a knowledge check. The knowledge check questions all free response, to meet higher order levels of Bloom’s Taxonomy.. The full knowledge check questions can be found in **Appendix I.**  Learners will be able to see which answers were wrong, but will not be shown the correct answers.  Learners will not be able to move forward in the training until they get 100% on the knowledge test. | 20 minutes | Personal computer, access to the internet, access to [www.sparkmap.org](http://www.sparkmap.org/), access to Canvas training module. |

**Module 3: Interpret SparkMap Maps to Write Conclusions**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Learning Objective** | **Instructional Approach** | **Assessment Strategy** | **Estimated Time** | **Materials Needed** |
| Learners will be apply their map interpretation knowledge to draw and state conclusions about a given population (e.g. compare data from multiple counties to draw a conclusion about broader trends in a state). | Utilize the expository approach (p. 221) by providing rules of accurate and inaccurate data interpretation and then offering examples and non-examples. To provide examples and non-examples, we will provide users with case studies and a map that they will practice drawing conclusions from.  A sample is indicated in **Appendix J.** | Knowledge check & self check.  For the assessment there will be two parts, a knowledge check and self assessment. First, learners will be presented with examples of maps and data interpretation write ups, learners will indicate if the write up is a correct or incorrect interpretation of the data. Remediate wrong choices as needed (Table 9.4; p. 228).  Second, learners will engage in a self check where we will present them with a multilayer map and ask them to draw conclusions (e.g., compare the unemployment rate and average educational attainment in Missouri with neighboring states of Iowa, Illinois, and Kansas to conclude the state with highest economic potential). After answering, learners will be presented with our correct interpretation to check their answer against. | 45 minutes | Personal computer, access to the internet, access to [www.sparkmap.org](http://www.sparkmap.org/), access to Canvas training module. |

**Appendix A:**

**Objective 1A: US Census Geographic Unit Terms**

**Common Geographic Terms:**

* **Address:** the particular place where someone lives or an organization is situated (Oxford Languages, 2022).
* **Borough:** an administrative unit that can have a different meaning in different locations:
  + Alaska: county equivalent
  + New York: minor civil division
  + Connecticut, New Jersey, and Pennsylvania: incorporated place
* **Congressional District (CD):** an area from which one representative is elected to the U. S. House of Representatives
  + All congressional districts in a given state contain roughly the same number of residents
  + The District of Columbia, Puerto Rico, and each Island Area (U.S. Virgin Islands, Guam, American Samoa, and Commonwealth of the Northern Mariana Islands) each contain a single congressional district.
  + Congressional districts are redrawn based on each decennial census. This means that the geographic boundaries of a congressional district may vary by decade. Boundaries may also be changed more than once during a decade.
  + For data on historic congressional districts, visit American Factfinder:
    - 106th Congress, seated in 1999 and 2000 - Census 2000
    - 109th Congress, seated in 2005 and 2006 - ACS
    - 110th Congress, seated in 2007 and 2008 - Census 2000, ACS
    - 111th Congress, seated in 2009 and 2010 - 2010 Census, ACS
    - 112th Congress, seated in 2011 and 2012 - same as 111th Congress
* **County:** a political and administrative division of a state, providing certain local governmental services
* **Parish:** county equivalent in Louisiana
* **State:** A geographical area that falls within the boundaries of a particular state.
* **State Legislative District:** an area from which members are elected to state legislatures
  + Most state legislatures have an upper chamber (senate) and lower chamber (house). Thus, most states have two sets of legislative districts: upper and lower.
  + Nebraska and the District of Columbia each have only one legislative chamber. These are represented as upper chamber legislative entities.
  + Like congressional districts, state legislative districts are redrawn periodically based on census data.

**Census-Specific Geographic Units**

* **Block:** the smallest geographic unit for Census data.
  + Boundaries may include:
    - Visible features (streets, roads, streams, railroad tracks)
    - Nonvisible boundaries (selected property lines; city, township, school district, or county boundaries).
  + In towns and cities, blocks often correspond to individual city blocks bounded by streets
  + In rural areas, blocks may include many square miles and may some boundaries that are not streets
  + Blocks may change from census to census
* **Block Group:** A statistical subdivision of a census tract,
  + generally contains between 600 and 3,000 people or 240 and 1,200 housing units
  + the smallest geographic unit for which the Census Bureau tabulates sample data
* **Census Area:** county equivalent in parts of Alaska that are not organized into boroughs
* **Census designated place (CDP):** statistical equivalent of incorporated places. These are places that do not have legal powers or functions, but do contain settled concentrations of population.
* **Census tract:** A small, relatively permanent statistical subdivision of a county
  + Boundaries normally follow visible features (e.g. streets, streams), but may also follow legal geographic boundaries (e.g. county lines) or non-visible features (e.g. school districts).
  + Ideally contain about 4,000 people or 1,600 housing units
  + Census tract boundaries do not change from census to census because the goal is for users to be able to compare statistics for the same census tract over multiple decades. If the population in a tract grows considerably larger than 4000, it may be subdivided; however, the original boundaries remain intact.
* **Incorporated Place:** A place with legal powers and functions
  + Usually includes: cities, towns, boroughs, and villages
  + The following are NOT treated as incorporated places:
    - Boroughs in Alaska (treated as statistical equivalents of counties).
    - Towns in the New England states, New York, and Wisconsin (treated as minor civil divisions (MCDs).
    - Boroughs in New York (treated as MCDs).
  + Hawaii and Guam do not contain any incorporated places
* **Public Use Microdata Area (PUMA):** A statistical area defined to contain a population of 100,000 or greater for which the Census Bureau tabulates public use microdata sample (PUMS) data. American Community Survey and decennial census population and housing microdata are disseminated using these defined areas. The American Community Survey also publishes one year estimate data for PUMAs
* **Zone Improvement Plan (ZIP) code:** areas of land, grouped together based on US Postal Service Delivery routes. These are determined by the US Postal Service.
  + May be assigned to a section of a street, a collection of streets, an establishment, structure, or group of post office boxes, for the delivery of mail.
* **ZIP Code Tabulation Area (ZCTA):** geographical areas based on US Postal Service delivery routes.

**Appendix B:**

Objective 1A Knowledge Check

True/ False questions

1. True or False: Census tracts are subdivisions of counties
2. True or False: A census tract can be subdivided if the population drastically increases (true – if the population grows to substantially over 4000, the tract can be subdivided)
3. True or False: Census tracts can be combined when the population in an area drastically decreases (false – tracts remain the same over time)
4. True or False: A ZCTA is a subdivision of a county (false – ZCTAs are based on US Postal Service delivery routes. ZCTAs may contain multiple counties.
5. True or False: A block group is a subdivision of a tract
6. True or False: Ideally, census tracts are roughly equal in land area. (false – census tracks are ideally rough equal in population. Ideally, each tract encompasses about 4000 residents. However, actual numbers may vary by hundreds or even thousands)
7. The boundaries of which geographic unit are LEAST likely to change over time?
   1. Congressional district
   2. State legislative district
   3. Census tract
   4. Census block
8. Arrow Rock, Missouri is a small community with a population of 56 and a municipal government that includes a mayor and a board of trustees. Arrow Rock is an example of a(n)...
   1. Incorporated place
   2. Census-designated place
   3. Census Area
   4. Census Community
9. Edinburgh, Missouri is a small community with a population of 82 and no municipal government. Edinburgh is an example of a(n)...
   1. Incorporated place
   2. Census-designated place
   3. Census Area
   4. Census Community
10. In Southwest Missouri, the counties of Jasper, Newton, McDonald, Barry, Lawrence, Polk, Greene, Christian, Stone, Taney, and part of webster are all represented by the same member of the US House of Representatives. This area is an example of a(n)...
    1. State legislative district
    2. Congressional district
    3. County group
    4. Senate district

**Appendix C:**

Objective 1A Handout: US Census Data Source Terms

* **Category:** the type of data within a data set. Examples include demographic, economic, education, housing, etc.
* **Data source:** The agency, company, and specific timeframe of the data. This would indicate results from a specific survey.
* **Data type**: different statistics that have been tabulated for a data set. Common data types include percent, total, and rate (e.g. cases per 1000)
* **Decennial Census:** an account of population and housing in the United States that occurs once every 10 years (United States Census Bureau, 2017). In this survey, people from all states and US territories are counted. This census is incredibly important because it “determine(s) the number of seats for each state in the U.S. House of Representatives and are used to draw congressional and state legislative districts and to distribute hundreds of billions of dollars in federal funds each year” (United States Census Bureau, 2017, para. 6).
* **Demographic surveys:** those which measure information regarding socioeconomic status, race, poverty, education, housing, and other subjects (United States Census Bureau, 2017).
* **Economic census:** Provides a detailed portrait of the United States' economy once every five years, from the national to the local level. It covers most of the U.S. economy in its basic collection of establishment statistics.
  + Collective name for the censuses of construction, manufactures, minerals, minority- and women-owned businesses, retail trade, service industries, transportation, and wholesale trade, conducted by the Census Bureau every five years, in years ending in 2 and 7
* **Economic surveys:** surveys of specific parts of the national economy. These occur on monthly, quarterly, and annual bases (United States Census Bureau, 2017).
* **Indicator:** the content being measured. For example, total population.
* **Release Cycle:** how often a given type of data set is released. For example, demographic data from decennial census is released decadally, or every ten years
* **Release Date:** when a data set is released to the public. This may be a different year than the one in which the data was collected.
* **Source geography:** the geographic boundaries a data set is available at. Examples include census tract, county, Zip code, ZCTA, state, and address (defined below).
* **Sponsored surveys:** demographic and economic surveys collected for other governmental agencies (United States Census Bureau, 2017).
* **State benchmark:** information that compares state level data to national level data.
* **US Census Bureau:** The US Census Bureau is one of many government agencies dedicated to collecting data on the economy, demographics, and geographical dispersion of Americans (United States Census Bureau, 2017). In fact, this is the largest statistical agency in the United States. The goal of the Bureau is to provide thorough and accurate data about America and its people.

**Appendix D**

Objective 1D: Applying definitions

This knowledge check uses the following data set: Women that Gave Birth, ACS 2012-16.

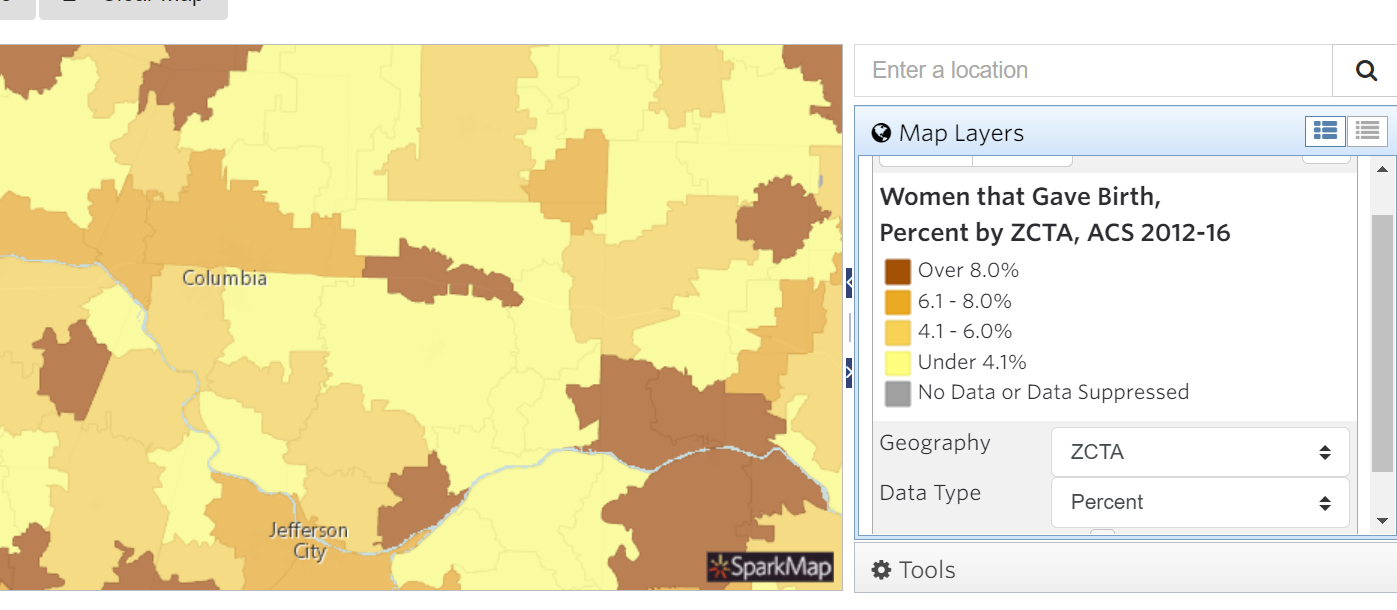
Data Set Information

|  |
| --- |
| Data Set: Women that Gave Birth, ACS 2012-16  **Data Type**  Percent, Total  **Data Geographic Unit**  State, County, Place, Tract, US Congressional District, State Senate District, State House District, School District (Elem.), School District (Sec.), ZCTA, PUMA  **Data Source**  US Census Bureau, American Community Survey: 2012-16  **Data Source Description**  The American Community Survey (ACS) is a nationwide survey designed to provide communities with reliable and timely social, economic, housing, and demographic data every year. The ACS has an annual sample size of about 3.5 million addresses, with survey information collected nearly every day of the year. Data are pooled across a calendar year to produce estimates for that year. As a result, ACS estimates reflect data that have been collected over a period of time rather than for a single point in time as in the decennial census, which is conducted every 10 years and provides population counts as of April 1. The Census Bureau combines 5 consecutive years of ACS data to produce estimates for geographic areas with fewer than 65,000 residents. These 5-year estimates represent data collected over a period of 60 months. Because the ACS is based on a sample, rather than all housing units and people, ACS estimates have a degree of uncertainty associated with them, called sampling error. In general, the larger the sample, the smaller the level of sampling error. Data users should be careful in drawing conclusions about small differences between two ACS estimates because they may not be statistically different.  Citation: Citation: U.S. Census Bureau: UNDERSTANDING AND USING AMERICAN COMMUNITY SURVEY DATA: WHAT ALL DATA USERS NEED TO KNOW (2018).  For more information about this source, including data collection methodology and definitions, refer to the American Community Survey data users website. |

Sample Questions:

1. Which of the following is NOT a geographic unit for which this data set has been tabulated?
   1. County
   2. Place
   3. Tract
   4. Block Group
   5. ZCTA
2. What is the data source for this data set?
   1. The decennial US Census
   2. The US Census Bureau’s American Community Survey
   3. A sponsored survey for the Department of Health and Human Services
   4. The US Department of Agriculture
3. What data types are available for this data set? Type all terms into the box (percent, total)
4. How many women gave birth in Columbia, Missouri between 2012 and 2016? (Hint: geography = place, type = total)
   1. Over 500
   2. 101-500
   3. 1-100
   4. 0
   5. No data or data suppressed

For questions #5 and 6, use the following figure:



1. What unit of geography is being shown here? (ZCTA)
2. What data type is being shown here? (percent)
3. In how many Missouri Congressional Districts did under 4.1% of women give birth between 2012 and 2016?
   1. None of them
   2. 1 of them
   3. About half of them
   4. All of them
4. Find the tract that includes Centralia, Missouri (this is slightly north of Columbia, at the top of Boone County). How many women in this tract gave birth between 2012 and 2016? (type in the exact number) (answer=11)
5. Find the ZCTA that includes Hallsville, Missouri (this is just north of Columbia). How many women in this ZCTA gave birth? (type in the exact number) (answer=32)
6. Find the State Senate district that include Columbia, Missouri. What percent of women in this state senate district gave birth?
   1. Over 8.0%
   2. 6.1-8.0%
   3. 4.1-6.0%
   4. Under 4.1%
   5. No data or data suppressed
7. Which place had a higher percent of women who gave birth?
   1. Columbia, MO
   2. Jefferson City, MO

**Appendix E**

Objective 2A: Creating a Map in SparkMap

Step-by-Step instructions to make a map in SparkMap. This will be used as a job aid.

1. In your web browser, visit [www.sparkmap.org](http://www.sparkmap.org/)

2. Select the “Get Started” button on the top toolbar.

3. Scroll down to the Map Room and click the green “Make a Map” button.

4. In the “Search Data” box, type “Population Census 2020” we will use this map throughout this module as an example.

5. Select the “Predominant Race/Ethnicity by County, US Census 2020” data set and it will automatically be added to your map.

6. You have now created your first SparkMap, map!

7. Compare your map to the map uploaded as “Example Map” in the module. Be sure the correct data set is shown and the map layers box looks the same.

8. If your map looks different, revisit steps 2-7.

**Appendix F**

Objective 2B: Use & Interpret a SparkMap Map Key

Script for Video Tutorial

*Hello and welcome to your first SparkMap tutorial! We’re excited for your interest in our map room. Let’s get started by walking through our map key, otherwise known as map layers.*

*For this example, we will be walking through a map using the “Predominant Race/Ethnicity by County, US Census 2020” data layer. This should be the same map you have created. We’d love for you to walkthrough with us! If this isn’t the map you’ve created, please visit the first document in this module.*

*First, when interpreting a map in SparkMap it is important to notice the tools embedded in the map. On the left hand side, you will be able to zoom in with the “+” button, zoom out with the “-“ button, and reset the map to its original view with the house button. Further, you can change the appearance of the map using the button that looks like four squares right here.*

*Next, on the far right side of the window, you will see a “Map Layers” box. In this box, it shows you your data set, as well as what each of the colors represents. For example, in this data set dark blue means the predominant population of a county is over 90% White, whereas bright pink indicates the primary population of a county is less than 50% Native Hawaiian/Pacific Islander.*

*On the right side of the map layers box, you will notice three dots. If you click on those, you can select “data info” to learn more background information about the data set, which was covered in module one, search within the data using the “query data” tool or remove the layer from the map.*

*Lastly, at the bottom of the map layers box, you can change the geography of the map for example from county to block group, and also the transparency of the map layers. This is especially helpful when you have more than one layer on the map.*

*Explore your map and key a little bit to become even more familiar.*

*When you’re ready, visit the next step in the module to test your knowledge on building a map and using the key.*

**Appendix G**

Objective 2B: Interpreting a SparkMap Map Key

Knowledge Check Questions

1. The data set used in the map comes from the US Department of Defense (2016)?
   1. True
   2. False
2. Which of the following geographies can you change this map to?
   1. City
   2. Census Tract
   3. Block Group
   4. Country
3. Using the “data info” tool, what is the release cycle of this data set?
   1. Decadal
4. Which color indicates that the predominant race/ethnicity of the area if 70%-90% Asian?
   1. Purple
   2. Green
   3. Orange
   4. Grey
5. Based only on geographic area, what is the predominant race/ethnicity of Alaska?
   1. Hispanic
   2. Black
   3. Some other race
   4. Native American/Alaska Native

**Appendix G**

Objectives 2C and 2D: Find Location of Interest & Extract Information

Video Tutorial Script

*Welcome back to your SparkMap data interpretation tutorial. In this VoiceThread, we will move on to the next step of data interpretation with SparkMap: finding location-specific information. There are two ways you can find information out about a specific location, let’s go back to the map you created and explore them together!*

*First, navigate to the “enter a location” search bar above the “map layers” box. Type in “Chicago, IL.” When you hit enter, the map will automatically zoom you in to Chicago. Here, you can see three different colors, all shades of blue. Using your skills from the last walkthrough, you can tell that part of Chicago has a predominant race of <50% White, a second part has a predominant race of 50-70% White, and a third portion has a predominant race of 70%-90% White. While this is interesting, it isn’t really helpful if understanding the specific county populations.*

*So, the second way to find out location-specific information is by clicking on the map.*

*First, click on the lightest blue area. When you do that, a whole box of information comes up. You can see that this specific area is Cook County in Chicago. You can also see that 40.47% of the population identifies as White, and see the breakdown of number of residents and percentage of residents identifying with each race/ethnicity group.*

*Repeat these steps for the other two colors in this section of the map. As you click around this area, you should find information for DuPage, Will, Lake, and Porter Counties.*

*We suggest you practice this process with other areas throughout the US. Remember, you can search an area in the toolbar above “map layers” and zoom out to click around the map.*

*When you’re ready, visit the next step in the module to test your knowledge on searching for location-specific data in the map.*

**Appendix I**

Objectives 2C and 2D: Find Location of Interest & Extract Information

1. Find Atlanta, Georgia on the map. Locate DeKalb County. What specific percentage of the county is the predominant race/identity?

**Answer: 50.29**

1. Also in DeKalb County, how many participants indicated their predominant race as Asian (Non-Hispanic)?

**Answer: 50,076**

1. Find Boone County, Missouri on the map. What percentage of the population indicated their primary race as Two or More Races?

**Answer: 5.76**

1. Switch the geography to block group. Go to Denver, Colorado. One block group has a predominant race/ethnicity of <50% Asian. Which block group is it?

**Answer: Block group 3**

1. What is the primary percentage and race/ethnicity in Quitman County, Mississippi?

**Answer: 73.61% Black**

**Appendix J**

Objective 3: Drawing Conclusions

Sample Case Study

Laura and Bruce are local lawmakers in Spokane, Washington. Recently, they have been charged with targeting and distributing funds from the American Rescue Plan Act. To start, they are interested in focusing on the accessibility to broadband internet in their area. Using SparkMap, they have created a map of internet providers and fiber internet access. Based on their map, they draw the conclusion that **all tracts in Spokane have access to at least 2 internet providers and 50% of the county has access to Fiber Internet.**

Using their map, is this a correct interpretation?

Map access: <https://cares.page.link/zsQe>

**References**

Brodsky, M. How to read a chart. Easterrn Michigan University Library. <https://www.emich.edu/library/help/tutorials/assets/read_chart/story_html5.html>

Brodsky, M. *What is data? What is data literacy?* Eastern Michigan University Library. <https://www.emich.edu/library/help/tutorials/assets/data_intro/story_html5.html>

Data literacy: Skills of critical thinkers. (2018, October 25). Tableau Software. Youtube. <https://www.youtube.com/watch?v=jJHkKPShqZs>

Emerson, J. (2008). Visualizing information for advocacy: An introduction to information design. Tactical Technology Collective. <https://www.opensocietyfoundations.org/publications/visualizing-information-advocacy-introduction-information-design>

Gal, I. (2002). Adults’ statistical literacy: Meanings, components, responsibilities. International Statistical Review 80.1, 1-51.

Glossary. United States Census Bureau. <https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13>

Juergens, C. (2020, October 23). Trustworthy COVID-19 mapping: Geo-spatial data literacy aspects of chloropleth maps. Journal of Cartography and Geographic Information 70: 155-161. <https://link.springer.com/content/pdf/10.1007/s42489-020-00057-w.pdf>

Krohn, J. (2022, March 11). The health, environmental, and economic impacts of commuting. *SparkMap.* <https://sparkmap.org/blog/the-health-environmental-and-economic-impacts-of-commuting/>

Oldrey, K. (2017, December 1). Data literacy: Top five tips for interpreting maps. Champaign County Regional Planning Commission.

<https://ccrpc.org/news/2017/top-five-tips-for-interpreting-maps/>

Rogers, L. (2017, September 19). Different measures of demographic change and the stories they tell. United States Census Bureau. <https://www.census.gov/newsroom/blogs/random-samplings/2017/09/population-change.html>

Survey and Topic One-Pagers. United States Census Bureau. <https://www.census.gov/data/academy/resources/one-pagers.html>

U.S. Census Bureau at a glance. (2017, October 24). United States Census Bureau. <https://www.census.gov/about/what/census-at-a-glance.html>

What are census blocks? (2021, August 3). United States Census Bureau. <https://www.census.gov/data/academy/data-gems/2021/what-are-census-blocks.html>

What is the definition of that variable? Making sense of census terminology. (2018, July 2). United States Census Bureau. <https://www.census.gov/data/academy/data-gems/2018/definitions.html>

What is a CDP? Making sense of census geography. (2018, May 21). United States Census Bureau. <https://www.census.gov/data/academy/data-gems/2018/cdp.html>

What is a census tract? Making sense of census geography. (2018, July 2). United States Census Bureau. <https://www.census.gov/data/academy/data-gems/2018/tract.html>

What is a ZCTA? Making sense of census geography. (2018, July 2). United State Census Bureau. <https://www.census.gov/data/academy/data-gems/2018/zcta.html>