# DEVELOPING AN INTERACTIVE DASHBOARD ON DISABILITY POPULATION IN THE U.S. WITH TRENDS ANALYSIS

EXECUTIVE SUMMARY BY **JIAWEI WANG** | AUGUST 2020

INTRODUCTION

Facts supported by data are very powerful. When people want to explain the importance of supporting people with disabilities (PWDs), data can strongly illustrate this importance. For example, the American Community Survey, in 2018, estimated the amount of civilian non-institutionalized population who reported having a disability in the United States to be 40,637,764, which is 12.6% of the civilian non-institutionalized US population. For comparison, according to the United States Census Bureau, the estimation of PWD population in the U.S. is larger than that of the entire population of California for the same year; this comparison reaffirms the notion that the needs of PWDs cannot be ignored as they make up a very significant percent of the U.S. population. Besides the large population number, there exist other facts from data that illustrate the importance to support PWDs.

However, data may not be so easy to find or utilize. First, searching for facts with data has an unavoidable limitation due to the searching words and existing data tables. The process of searching a fact with data on the internet is actually using keywords to match with the titles or descriptions of summarized data tables. In order to find the desired facts with data, people must have great searching techniques and patience to look over websites. Meanwhile, the information that people can get from the searching process is limited up to the content of existing summarized tables. Second, for people with knowledge and the ability to transfer data into summarized tables, high-quality, unsummarized datasets are not always available and usable. Even with the condition to get good datasets, making data into facts and information requires a lot of time.

Considering potential users, who need data of summarized tables but more flexible to find out desired facts with data, this research project is established to build an interactive data dashboard about PWDs in the United States. "Interactive" could allow users to explore data based on their needs. "Dashboard" offers maps, graphs, and tables that give straightforward, intuitive, and professional information from data.

TASK AND GOAL

The task of this research project was to build an interactive data dashboard about the disability population in the United States. With the "interactive" feature, the data dashboard includes filters of demographic

characteristics and other independent filters. With the "dashboard" feature, the data dashboard updates data in maps, tables, and graphs synchronically, to show the data at the same time frame.

The goal of this research project was to build a basic data tool for anyone who is interested in utilizing data as facts and needs convenient, non-coding, and instant operations.

METHODOLOGY

To build the interactive data dashboard, a project outline with three stages was initialized:

- Data Collection: Find professional, accurate, and unsummarized raw data.
- Data Process and Estimates: Process raw data and make data estimates from datasets.
- Data Visualization: Visualize processed data and make adjustments at any time.

*1. Data Collection*

Because the data dashboard is about the disability population, and the property of original datasets should be unsummarized, this research project mainly considered population surveys as the main source of data. Meanwhile, the ideal datasets should adopt the same definition of disability consistently through time, to avoid the break of analysis to the same disability population.

According to the Centers for Disease Control and Prevention(2019), the U.S. Department of Health and Human Services considers a set of six disability questions in the American Community Survey as "the data standard for survey questions on disability…[It] represents a minimum standard, and the questions and answer categories should not be changed." The ACS categorizes disability into 6 different types: (United States Census Bureau, 2017):

- Hearing difficulty: deaf or having serious difficulty hearing.
- Vision difficulty: blind or having serious difficulty seeing, even when wearing glasses.
- Cognitive difficulty: Because of a physical, mental, or emotional problem, having difficulty remembering, concentrating, or making decisions.
- Ambulatory difficulty: Having serious difficulty walking or climbing stairs.
- Self-care difficulty: Having difficulty bathing or dressing.
- Independent living difficulty: Because of a physical, mental, or emotional problem, having difficulty doing errands alone such as visiting a doctor's office or shopping.

The goal of data collection was to achieve  accurate, consistent, large-volume, public, and unsummarized dataset. The 1-Year Estimates dataset, named Public Use Microdata Sample (PUMS), published by the ACS, is reliable and satisfies the goal of data collection. PUMS has adopted the standard six disability questions
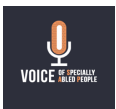
from 2008 to 2018 and  surveys 1% of the American population each year. While anyone can access their data, individual data records cannot be used to identify respondents. Most importantly, this dataset contains original answer records, which means the data is not summarized and contains all the basic demographic information that the interactive data dashboard needs.

For each year between 2008 and 2018, CSV data files were downloaded from the ACS. Household records and personal records in each year can be linked through a serial number. The total number of variables for each year is around 400 to 500. The total number of records for each year was around 400 million.

*2. Data Process and Estimates*

In this stage, Python was used to process original datasets. The variables analyzed included  gender, age, race, population types, disability status, disability types, education level, employment status, etc. After processing, this research project used methods of estimation found in ACS 1-year PUMS Accuracy documents to estimate the number and percent of the population with different characteristics with the corresponding margin of error with 90% confidence. The estimated results that appeared in ACS 1-year PUMS Estimates were also compared with the estimates generated by this research to ensure that there was no difference and correct estimating methods were used. Due to the sampling property of the PUMS dataset from the actual ACS dataset, the estimating result using the PUMS dataset had a little difference from the result using the actual ACS dataset. But the accuracy of estimates generated from PUMS was verified by the ACS and the error is in a normal range.

*3. Data Visualization*

After the data was processed in Python, it was integrated into Tableau to make the dashboard. Three main dashboards were designed with topics of population, employment, and personal income. These three main pages all have selected basic demographic filters, like Population Type, Year, Gender, Age, Race, Education Level, and Employment Status. Maps and tables contain value format selectors, like percent or number, and location selectors, like any state in the US or the whole nation. There are four more pages of trend analysis, including the trend of the disability population with different disability types, ages, genders, races, and their combinations from 2008 to 2018.

In this step, estimates from the last stage were not simply summarized in tables and put into Tableau. Instead, this research used the "calculation field" feature in Tableau to use the same estimating method verified in the last stage. After initializing the dashboard, the result shown in Tableau was compared with the result generated by Python in Stage 2 with proper adjustments until no difference in the comparison.
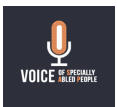
CONCLUSION

This research project has built an interactive data dashboard about the PWD population in the United States from 2008 to 2018 by using raw data of population surveys. The "interactive" feature is implemented by adding selected filters, and the "dashboard" feature is fulfilled by the synchronous maps, graphs, and tables. This interactive data dashboard can be considered as a basic tool for people who are interested in finding facts with data of PWDs. Users can find out facts of a specific group of PWDs through using filters of Population Type, Year, Gender, Age, Race, Disability Status and Types, Education Level, and Employment Status. The existence of this data dashboard of Besides the current content of the data dashboard, more elements and different topics of trend analysis are expected to be added in the future, including distribution of education level, trend analysis of employment, and etc.

—————————————

[1] "2018 ACS 1-year PUMS estimates [CSV]." United States Census Bureau. Accessed by August 11, 2020.

https://www2.census.gov/programs-surveys/acs/tech.docs/pums/estimates/pums.estimates.18.csv?#

"Disability Characteristics." United States Census Bureau. Accessed August 11, 2020.

https://data.census.gov/cedsci/table?q=disability&tid=ACSST1Y2018.S1810&t=Disability.

[2] "Disability Datasets." Centers for Disease Control and Prevention. November 26, 2019.

https://www.cdc.gov/ncbddd/disabilityandhealth/datasets.html.

[3] "How Disability Data are Collected from The American Community Survey." United States Census Bureau. October 17, 2017.

[4]https://www.census.gov/topics/health/disability/guidance/data-collection-acs.html#:~:text=All%20three%20surveys%20ask%20about,considered%20to%20have%20a%20disability.

[5] "Introduction to the Public Use Microdata Sample (PUMS) File." United States Census Bureau. April 30, 2018.

https://www.census.gov/programs-surveys/acs/guidance/training-presentations/acs-intro-pums.html.

[6] "Public Use Microdata Sample (PUMS) Accuracy of the Data (2018)." United States Census Bureau. Accessed by August 11, 2020.
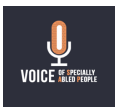
https://www2.census.gov/programs-surveys/acs/tech.docs/pums/accuracy/2018AccuracyPUMS.pdf?#

[7] "QuickFacts California." United States Census Bureau. Accessed August 11, 2020.

https://www.census.gov/quickfacts/CA.

[8] Sirius Fuller, "Using American Community Survey Estimates and Margins of Error," United States Census Bureau, April 18, 2018,

https://www.census.gov/content/dam/Census/programs-surveys/acs/guidance/training-presentations/20180418.MOE.pdf

## ABOUT THE AUTHORS

**Jiawei Wang** is a master student at the University of Southern California studying Applied Mathematics. Before that, Jiawei received a bachelor's degree in Mathematics from University of Washington in 2019.

Jiawei hopes to practice quantitative research, data analysis, and statistical information. He focuses on developing problem-solving ability and mathematical modeling, including theoretical mathematics study and applicable statistical skills. In free time, he is interested in reading, game culture, and social media operations.

## WHY THIS MATTERS TO VOSAP

Over 400 MN records are processed in this project, using data from American Community Survey (ACS), US census bureau. Like most of the modern technological world, VOSAP believes in the power of data and more specifically, how it can be used to make a strong advocacy point, drive change in the disability sector. Through the development of an interactive dashboard of US data on disability from 2008 to 2018, VOSAP is able to expand its efforts in educating policy makto mainstream SAP (PwD) in our society in the US, providing trend analysis, graphic representation and deep insights.

## ABOUT VOICE OF SPECIALLY ABLED PEOPLE INC.

The Voice of Specially Abled People (VOSAP) is a global advocacy organization built on the principles of Empowerment of Specially Abled People. In Special Consultative Status with UN ECOSOC, VOSAP is working to create an Inclusive and Accessible world by accelerating implementation of UN Sustainable Development Goals (SDGs) and goals of UN CRPD (Convention on the Rights of Persons with Disabilities) treaty. The organization has created the VOSAP Mobile App through which volunteers can take a pledge to volunteer and rate the accessibility of public places, creating a crowdsource platform to aggregate demand for accessibility.