

Mental Health Activities from the Household Pulse Survey (experimental data)

Overview

To support the U.S. national recovery, the Census Bureau together with other designated Federal Agency issued the “Household Pulse Survey” a 20-minutes online survey that inquires how the coronavirus pandemic is impacting the U.S. States from a social and economic perspective. The survey’s question is about the following areas: childcare, education, employment, energy use, food security, health, housing, household spending, Child Tax Credit payments, and intention to receive a COVID-19 vaccination.

More details: <https://www.census.gov/programs-surveys/household-pulse-survey.html>

Objective

From the public available tables, the analysis will be focused on “Table 4. Mental Health Activities in the last 4 weeks, by selected Characteristics” for phase 3.1 (Period: April 14th to July 5th, 2021) available here:

<https://www.census.gov/data/tables/2021/demo/hhp/hhp33.html>

The task is to perform an initial data and exploratory analysis of some of the data and derive insights or correlations between the selected characteristics and perform a better segmentation based on possible correlations.

Key questions

- Which age range adhered the most to mental health activities¹
- Does the family status or household size have an impact on conducted mental health activities? (i.e., do single participate more often in health activities?)
- The survey categorizes different income ranges. Do they play a major role in participating in mental health activities?
- Are there any patterns that can be identified across the analyzed Metropolitan Areas²?

Data

For this case study, data will be from the census.gov publicly available resources:

<https://www.census.gov>

and especially the datasets “Table 4. Mental Health Activities in the last 4 weeks, by selected Characteristics” of the weeks 28th till 33rd of the project.

¹ Participation to mental health activities will be defined in the data dictionary of this project brief

² Metropolitan areas will be defined in the data dictionary of this project brief

Table 4 will be customized to focus on the Metropolitan Areas and the demographic and social aspects of the case study.

Note: *These data are experimental. Users should take caution using estimates based on subpopulations of the data – sample sizes may be small and the standard errors may be large.***

Data Dictionary & Abbreviation meaning

Question	Q
Received counseling or therapy from a mental health professional such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker.	Q1
Needed counseling or therapy from a mental health professional but did not get it for any reason.	Q2
Took prescription medication to help with any emotions or with your concentration, behavior, or mental health.	Q3
Yes	yes
No	no
Did not report	na
Period	W
Race	
<i>Hispanic or Latino (may be of any race)</i>	Hispanic/Latin
<i>White alone, not Hispanic</i>	White
<i>Black alone, not Hispanic</i>	Black
<i>Asian alone, not Hispanic</i>	Asian
<i>Two or more races + Other races, not Hispanic</i>	Other races
Marital Status	
<i>Marital Status – did not report</i>	Marital status na
Household size	
<i>1 person in the household</i>	HH1
<i>2 people in the household</i>	HH2
<i>3 people in the household</i>	HH3
<i>4 people in the household</i>	HH4
<i>5 people in the household</i>	HH5
<i>6 people in the household</i>	HH6
<i>7 or more people in the household</i>	HH7
Presence of children under 18 years old	
<i>Children in household</i>	child<18y_yes
<i>No children</i>	child<18y_no

Respondent or household member experienced loss of employment income	
<i>Yes</i>	loss of empl. income_yes
<i>No</i>	loss of empl. income_no
<i>Did not report</i>	loss of empl. income_na
Respondent currently employed	
<i>Yes</i>	current. emp_yes
<i>No</i>	current. emp_no
<i>Did not report</i>	current. emp_na
Household income	
<i>Income</i>	inc
Cities Metro Area	
<i>New York Metro Area</i>	New York-Newark-Jersey City, NY-NJ-PA Metro Area
<i>Los Angeles Metro Area</i>	Los Angeles-Long Beach-Anaheim, CA Metro Area
<i>Chicago Metro Area</i>	Chicago-Naperville-Elgin, IL-IN-WI Metro Area
<i>Dallas Metro Area</i>	Dallas-Fort Worth-Arlington, TX Metro Area
<i>Houston Metro Area</i>	Houston-The Woodlands-Sugar Land, TX Metro Area
<i>Washington DC Metro Area</i>	Washington-Arlington-Alexandria, DC-VA-MD-WV Metro Area
<i>Miami Metro Area</i>	Miami-Fort Lauderdale-Pompano Beach, FL Metro Area
<i>Philadelphia Metro Area</i>	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD Metro Area
<i>Atlanta Metro Area</i>	Atlanta-Sandy Springs-Alpharetta, GA Metro Area
<i>Phoenix Metro Area</i>	Phoenix-Mesa-Chandler, AZ Metro Area
<i>Boston Metro Area</i>	Boston-Cambridge-Newton, MA-NH Metro Area
<i>San Francisco Metro Area</i>	San Francisco-Oakland-Berkeley, CA Metro Area
<i>Riverside Metro Area</i>	Riverside-San Bernardino-Ontario, CA Metro Area
<i>Detroit Metro Area</i>	Detroit-Warren-Dearborn, MI Metro Area
<i>Seattle Metro Area</i>	Seattle-Tacoma-Bellevue, WA Metro Area

Analysis Criteria

- The project folder follows industry standards in terms of structure and naming conventions
- Analysis has been conducted using Jupyter notebooks and the Anaconda libraries manager

- Analysis has been conducted using Python and relevant libraries (pandas, numpy, os, matplotlib, scipy, seaborn and geopandas)
- Data has been cleaned. Duplicate data, missing data, and mixed-type columns have been checked and addressed.
- Data ethics have been kept in mind when dealing with data, especially customer information.
- The final report includes evidence of analysis methodology, clear answers to the questions in this brief, insights, and visualizations.