

Term Project Milestone #1: Choosing the Topic and Creating the Initial Dataset

Michelle Helfman

Bellevue University

DSC540-T301 Data Preparation (2233-1)

Catherine Williams

April 9, 2023

Choosing the Topic and Creating the Initial Dataset

Project Subject Area: Moving to a new city is a life-changing event, but most people do not know where to start the decision process because the information is scattered all over the internet. My project creates a "Moving Starter Kit" of basic demographic, economic, education, and additional location-based information.

Data Sources:

- Flat File:
 - Moving Starter Kit Flat File – Information at the Metropolitan/Micropolitan, State, and City level of basic demographic, economic, education, and additional location-based information. See Include Codebook for column details.
 - See Moving Starter Kit Flat File.xlsx
- API:
 - Statistical weather data for each “anchor city” listed in the Moving Starter Kit Flat File to provide climate information and historical weather conditions.
 - Open Weather - <http://api.openweathermap.org>
- Website:
 - The 4 regions of the United States: South, West, Northeast, and Midwest. The regions will allow for grouping information in visualizations.
 - The SpreadsheetGuru, List of the 50 States, Abbreviations, & Capitals - <https://www.thespreadsheetguru.com/blog/list-united-states-capitals-abbreviations>

Relationships: Everything starts with the Moving Starter Kit Flat File. The Weather API is accessed using the longitude and latitude of the anchor_city, state, and a hard-coded US for

country code. For the United States geographic regions, either state name or state code can be used.

America Is On The Move: Information To Get The Search Started

Whether it is "pulling up stakes," "getting out of town," or just changing locations, people in America are on the move. People move for various reasons, political, religious, or cultural. Others move for better education opportunities for their children, to be closer to loved ones, or to find somewhere with less crime. Retirees seek better health care and stretch their retirement savings and social security. With more opportunities to work remotely, internet availability and speed interest people.

After deciding to focus on information to facilitate moving to a new city, I went out looking for data, nothing controversial or polarizing, like politics or religion, just readily available information from mainstream websites. After searching the internet, I found demographic, economic, and internet information on one website, retirement information on another, and information on crime on a third website. In all, five different websites were used to create 11 datasets. The datasets were then uploaded into a SQL Server database, created a cross-reference table, and wrote a stored procedure to consolidate the information into one final flat file. Next, the Moving Starter Kit file will be transformed into a more consistent and readable, and the missing information will be addressed. Weather API and regional data will be identified and added to the Moving Starter Kit. Finally, visualizations of the Moving Starter Kit and a report will be written to complete the process. The biggest challenge for this project will be my need for experience in web scraping and using an API, but I am excited about the opportunity to learn new skills.

This project is data-driven with facts about different metropolises in the United States. There are no opinions, editorials, or personal beliefs, just facts so someone can start looking for a new home.

References

DP05: ACS DEMOGRAPHIC AND HOUSING ESTIMATES

U.S. Census Bureau, 2021 American Community Survey 1-Year Estimates Retrieved April 2, 2023

[https://data.census.gov/table?q=United+States+demographics&g=010XX00US,\\$31000M1&tid=ACSDP1Y2021.DP05&tp=true](https://data.census.gov/table?q=United+States+demographics&g=010XX00US,$31000M1&tid=ACSDP1Y2021.DP05&tp=true)

S1902: MEAN INCOME IN THE PAST 12 MONTHS (IN 2021 INFLATION-ADJUSTED DOLLARS) U.S. Census Bureau, 2021 American Community Survey 1-Year Estimates Retrieved April 2, 2023

<https://data.census.gov/table?q=income&tid=ACSST1Y2021.S1902&moe=false&tp=true>

S2301: EMPLOYMENT STATUS

U.S. Census Bureau, 2021 American Community Survey 1-Year Estimates Retrieved April 2, 2023

[https://data.census.gov/table?q=employment+acs&g=010XX00US\\$3100000&tid=ACSS1Y2021.S2301&moe=false&tp=true](https://data.census.gov/table?q=employment+acs&g=010XX00US$3100000&tid=ACSS1Y2021.S2301&moe=false&tp=true)

S2801: TYPES OF COMPUTERS AND INTERNET SUBSCRIPTIONS

U.S. Census Bureau, 2021 American Community Survey 1-Year Estimates Retrieved April 2, 2023

[https://data.census.gov/table?q=internet&g=010XX00US\\$3100000&tid=ACSST1Y2021.S2801&moe=false&tp=true](https://data.census.gov/table?q=internet&g=010XX00US$3100000&tid=ACSST1Y2021.S2801&moe=false&tp=true)

City and Town Population Totals: 2010-2019

U.S. Census Bureau, 2021 Retrieved March 28, 2023

<https://www.census.gov/data/tables/time-series/demo/popest/2010s-total-cities-and-towns.html#ds>

Income Tax Rates By State

Tax-Rates.org — The 2022-2023 Tax Resource Retrieved March 31, 2023

<https://www.tax-rates.org/taxtables/income-tax-by-state>

United States Airports, by Name

Airport Codes Retrieved March 31, 2023

<https://www.airportcodes.us/us-airports-by-name.htm>

Offenses Known to Law Enforcement by State by City, 2019

FBI - Federal Bureau of Investigation, 2019 Crime in the United States Retrieved April 7, 2023

<https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/tables/table-8/table-8.xls/view>

Most Educated States

McCann, Adam (February 13, 2023) *2023's Most & Least Educated States in America*

<https://wallethub.com/edu/e/most-educated-states/31075>

Best & Worst States to Retire

McCann, Adam (January 23, 2023) *2023's Best States to Retire*

<https://wallethub.com/edu/best-and-worst-states-to-retire/18592>

Cross Reference Table – Joined together by FIPS Codes.

PRINCIPAL CITIES OF METROPOLITAN AND MICROPOLITAN STATISTICAL AREAS,

MARCH 2020 – 1270 Rows

U.S. Census Bureau, 2021 Retrieved April 1, 2023

<https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html>

American National Standards Institute (ANSI) and Federal Information Processing Series (FIPS)

Codes

U.S. Census Bureau, Retrieved April 7, 2023

<https://www.census.gov/library/reference/code-lists/ansi.html#state>