HAST-E Executive Summary

Christopher Bromley, Ryan Shannon, Christopher Smith, Sam Fritz-Schreck, Austin Mcgahan

Military families face frequent Permanent Change of Station (PCS) moves, causing significant stress. Each year, over 400,000 military families PCS. Unfortunately, it is challenging for military families to find adequate information about the areas they're moving to, which can add to the already stressful time. To tackle this issue, our team developed an intuitive decision support tool called the Housing Area Support Tool - Expeditionary (HAST-E). This tool helps military families search for their ideal housing location using a data-informed approach, which reduces the stress of finding the right place to live.

HAST-E takes user input preferences across a range of housing considerations, such as affordability, crime, school quality, walkability, and recreation. It then creates a customized, simple, and easy-to-visualize dashboard with recommended areas to live. To achieve our project objectives, we used a three-step method: data curation from numerous open sources, computation, and visualization.

We compiled data from 26 files and eight unique sources, including U.S. Census Keyhole Markup Language (KMLs), Basic Allowance for Housing (BAH) Primer/Rates, Housing & Urban Development (HUD) Fair Market Rent(FMR), the Federal Bureau of Investigation (FBI) Uniformed Crime Report, the Great Schools API, the Environmental Protection Agency's Walkability Index sourced from the Environmental Systems Research Institute (ESRI) API, and the OpenBrewery database API. We merged the data into multiple GeoDataFrames built off the census mapping KML files using GeoPandas. Furthermore, we conducted appropriate computations to normalize the metrics for final computation into an overall commutated score. This final commutation includes weighted averages of each area we identified for appropriate housing selection.

Using the centroid of each census block for input search locations, we conducted a brute force comparison of the top locations within 18 nautical miles of the duty station location. Leveraging the Param package to handle data inputs and user preferences, in combination with the ability to dynamically refresh through the Folium map, we created an interactive dashboard.
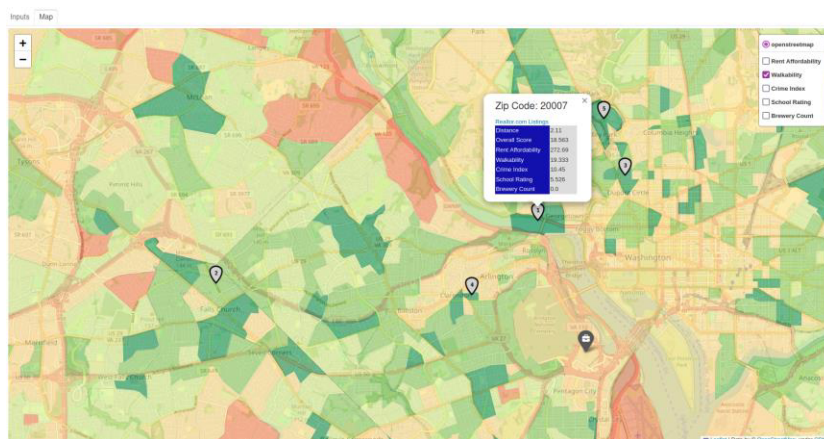

Figure 1: Snapshot of HAST-E results

We chose Folium due to its ease of use and robust sub-functions. With Folium, you can easily create a wide range of map layers, including choropleths and custom markers, and even include a layer control to allow users to toggle different layers on and off. However, it's important to note that Folium relies on one-way rendering, which means that to change the display of data, the map needs to be re-rendered. To accomplish this, we leveraged the Param package. This allows for user-modifiable parameters, arguments, and attributes. By creating a map refresh function dependency on a Param action button, Folium maps can be quickly reloaded with new information, making it a flexible tool for a variety of mapping needs.

However, there are some shortfalls of HAST-E. The biggest one is its computation time, which takes around one minute following user input. This inefficiency is likely due to for loops, and any computational gains would provide the end-user with a more usable dashboard. Secondly, the team leveraged GitHub as a collaboration tool, but its usefulness was limited due to the lack of familiarity amongst many group members. Nonetheless, GitHub allowed our team to maintain strict version control which prevented unexpected errors.

Throughout the data curation, computation, and visualization, we identified several shortfalls, such as data accuracy and overall data applicability. For example, through the Brewer database API pull we obtained over 15,000 unique breweries. We expected this large return of data to yield significant positive results for our program output. Unfortunately, only 58 brewers were applicable because of our focus on the National Capital Region (NCR). This fact was not recognized until the visualization phase of the program. Once mapped, it is easy to identify a few clusters of data with large areas without any valid returns. Using common sense, we were easily able to determine the large shortcoming of our data set. At times, we used diluted data to provide a valid return, such as a weighted average of the national crime index over a specific population when no crime was reported in a particular zip code. This weighted average is a "best guess" and may not be indicative of the actual crime in that area. Additionally, inconsistencies in reporting locations between data sets created gaps in our computation or created a dilated data value for that area.

Despite the identified shortfalls, there are vast future research opportunities or expanded usefulness for this tool, such as expanding the tool outside the National Capital Region (NCR) and integrating real-time housing options via external sources like realtor.com or zillow.com. Using tools like Nominatim would increase the ease of use. Overall, HAST-E provides military families with a valuable tool to reduce the stress of finding their ideal housing location during a PCS move.