Individual Report 3: Solid Ground

Niko Antoun

Winter 2024

1 Description of Data

The Excel file provided by Solid Ground contains data on Outcome Services within the years 2022 and 2023. There are two kinds of statistics in this data set, where the data under the National Performance Indicator (NPI) is indicated under the participants served and outcomes achieved, as well as individual services under raw counts. The given data set is divided into several categories, as separate spreadsheets: Broadview, Children's Services, Community Initiative, Community Food Education (CFE), Regional Access Point (RAP), Housing Stability Project (HSP), JourneyHome Rapid Re-Housing (JHRRH), Residential Services, Solid Ground Transportation (SGT), Supportive Services, Tenant Services, and Volunteer Services. At the beginning of the Excel file, there is an additional spreadsheet, called 'Summary', which provides a view of the categories on a single page. There also exists a handful of separate sheets that have a small amount of data, and some others have cells with no data, such as the entire Community Initiative sheet. I was told of the meaning of no-provided data is that Solid Ground did not collect data from a certain category. Before adjusting, collectively across all categories, there were 39 different unique fields of data and 106 total fields of data across all categories. After adjusting my cleaned data to ignore zeroed fields, there are 19 unique fields and 69 collective fields of data.

There is also a secondary Excel sheet that displays a list of 50 miscellaneous services with two fields: service name and participant count. Out of every single listed service, there are 25 unique services that can be broken down into six main categories: Community, Educational Support, Food Accessibility, Housing, Transportation, and Other. This sheet of data has significantly less complexity involved, and similar processes taken as for the main Excel sheet will be used to visualize the data.

2 Methodology

2.1 Preliminaries: Data Cleanup

Initially working through Excel, the data's native environment, normalizing all of the data into one, collective table with a 'Category' was the first move. Normalizing the data by moving it onto the master file was another process that involved splitting the headers, removing unnecessary colors, and other basic Excel operations. All numerical data within the table was originally set with a general datatype—a step to help future visualization. Assigning the appropriate datatypes to numbers entailed declaring numbers as numerical (double) and percentages. There also existed conditional formatting for certain percentages as a pre-set rule in the original spreadsheets, such that the condition a percentage is in between specific values. A significant benefit of normalizing the data in the master spreadsheet is unifying the cell rules, since each independent spreadsheet has set similar or equivalent rules. Now that all of the data is unified on a single table, potential rule changes can be made and applied with a single click. Under the master spreadsheet, the format is now set between 1 and 1000, thus not running out of constraints. The final step made was creating

pivot tables, a major asset to visualizing this data. The benefit of pivot tables is that they allow me to filter and combine categories, providing a direct route to creating visualizations.

The next phase entailed generating visualizations, which proved to be a tricky task due to the format of my Excel tables. While I was integrating my data into Tableau, the system did not process my Excel files as initially desired, but I found that I was not alone with my data format struggles after communicating my position in a class meeting. As a solution, I made a copy file to restructure my data into a similar Tableau-friendly format she had created. Functional refinements involved removing any extra labels, categories, and colors, where convenience refinements included appending label names. I therefore decided to follow a similar process as the team when it comes to considering data platforms.

2.2 Preliminaries: Visualizations

Through Python, choosing the appropriate method for visualization was a key decision to ensure the provided data is represented in a deliverable manner. Considering the nature of the statistics, utilizing bar charts to visualize each Outcome Service appeared to be the best approach, where the Turnout Target, Actual Turnout, and Successful Outcomes are the x-axis; and each y-axis represents the number of participants, alongside the Target Outcomes represented as a dotted line. Each bar chart has Python-generated labels for each variable, and additionally, Adobe Photoshop was used to manually label the actual numbers in a way to better implement good design principles. Figure 1 is the bar chart generated for the Outcome Service 'Individuals Who Obtained Safe and Affordable Housing.' For the Misc. Service Count data, since the data requires a more trivial route for visualization. Figure 2 maps all 25 of the unique fields and category counts onto bar charts.

Figure 1: Bar Chart Displaying Example Data

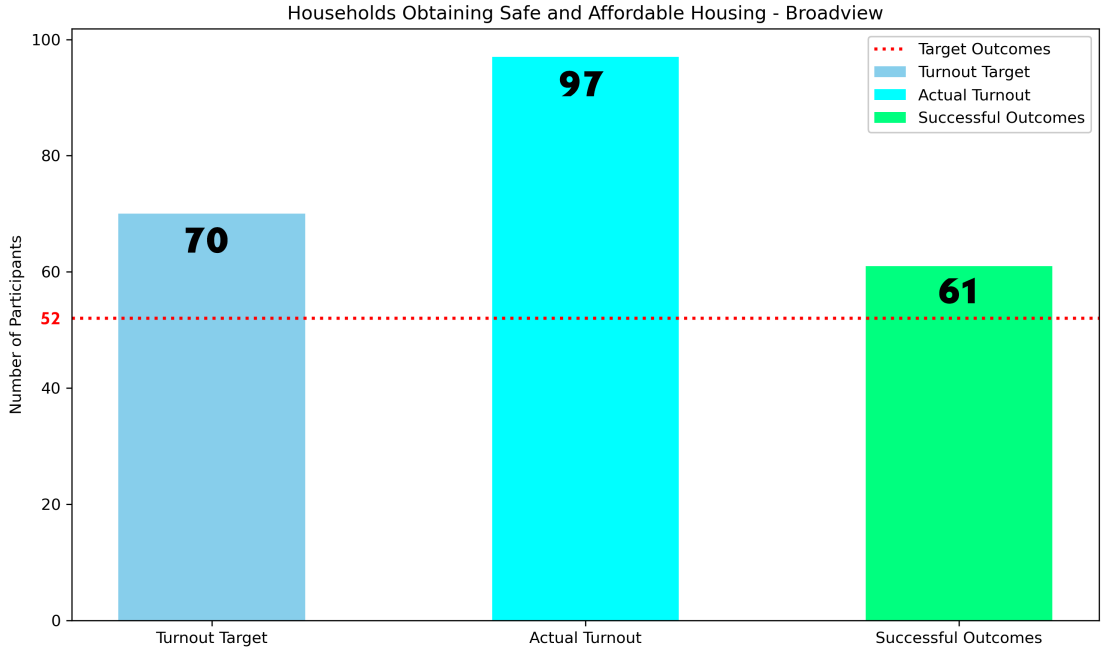
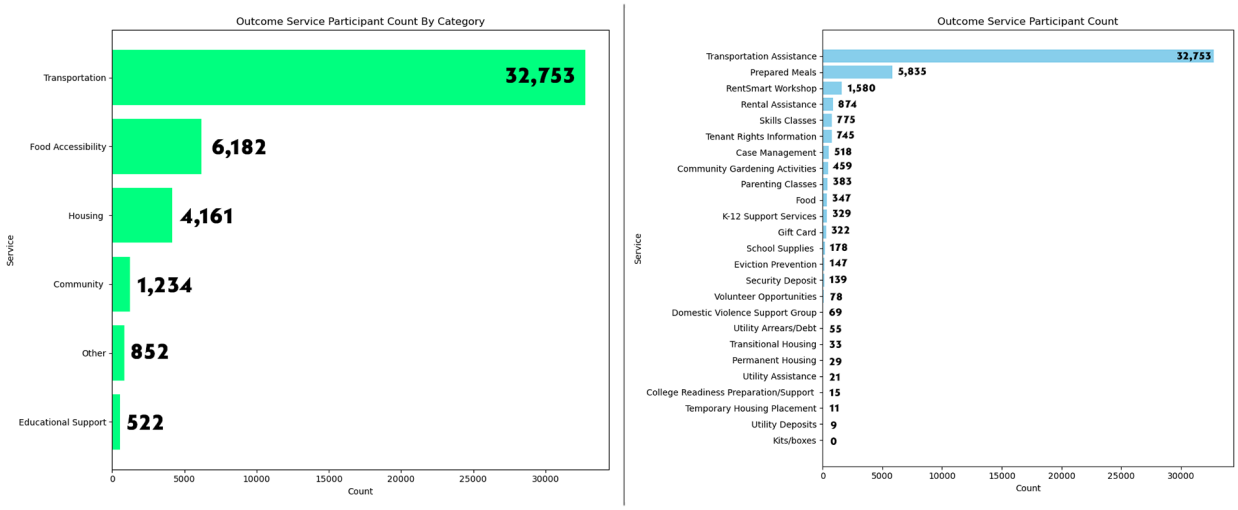


Figure 2: Bar Chart Displaying Miscellaneous Service Counts



2.3 Analysis and Measurements

A dimension to the NPI Outcome Services data that should be considered is implementing a relative target the number of successful participants, which appears to be set as a fixed number. In the context of the Outcome Services data, the "Target Outcomes" field is constructed with respect to the initial target participants, rather than the service's actual turnout. This can cause a misunderstanding on determining a program's success and strength, especially for the vast majority of programs that overperformed the target participants goal, and thus makes fulfilling a fixed goal easier. Considering certain NPI rows where there is a greater participant turnout than the initial "Target Participants" field, the bar for the initial "Target Outcomes" is inadvertently lowered due to a higher volume of participants. For instance, analyzing the Broadview category's"Number of Households That Obtained Safe and Affordable Housing" service, the target number of people to serve was set at 70, but 97 people were actually served. The successful outcome target is 52, i.e., approximately a 74% success outcome rate, i.e., a target of 73 successes. However, the actual successful outcome was 61 people, exceeding the fixed target by nine participants and missing the relative target by 12. In other words, the fixed target was met, but the true relative target was not met.

An appropriate measurement for the relative target outcome would be considering the rate of successes, allowing for an account of a higher-than-expected participant number. With 97 people served, 27 over the target, and 61 achieving a successful outcome, this translates to a success rate of 53.6%, not meeting the analyzed target rate of 74%. Therefore, a rate-based approach is more relevant, as it factors out the possibility of misleading statistics. Considering a success rate, such as 74%, is more indicative of actual performance than a fixed outcome goal. This method underscores the benefit of considering a success rate over a fixed number, providing a more accurate reflection

of performance called Relative Outcome Success Target Rate (ROSTR):

$$\text{ROSTR} = \frac{\text{Outcome Taret}}{\text{Target Participants}}$$

which yields a crucial key in determining the Relative Outcome Success Target Number (ROSTN):

$$\text{ROSTN} = \text{Actual Participants} \cdot \text{ROSTR}$$

and we can confirm these formulas work since all of the numbers of the fixed target of successful outcomes is always be less than the actual turnout of participants, and less than or equal to the target number of participants. After calculating each service's ROST rate and number, I was finally able to see the full picture of each service's true impact on its respective cause.

2.4 Visualizations

Applying the relative outcome target to the bar chart gives us clearer visualization on an Outcome Service's effectiveness. Figure 3 yields an updated version of Figure 1 with the addition of the relative target rate, denoted as a solid green line. Notice that, in Figure 3, that the number of successful participants is not sufficient to meet the relative success goal, even though it met the fixed success outcome; the number of successful participants is nine more than the fixed target but 12 less than the relative target. In other words, the program's fixed target was met, but the relative success target was not met, thus telling us this program was not effective in successfully serving the ratio with respect to the 97 participants who participated. As for the miscellaneous service count data, primarily due to the reason of significantly different counts, Figure 4 has an adjusted the x-axis to display on a logarithmic rather than a linear format. Both versions of the miscellaneous service count and category visualizations will be useful for Solid Ground to use as a reference, and can used interchangeably.

Figure 3: Bar Chart Displaying Outcome Service with Relative Success Target

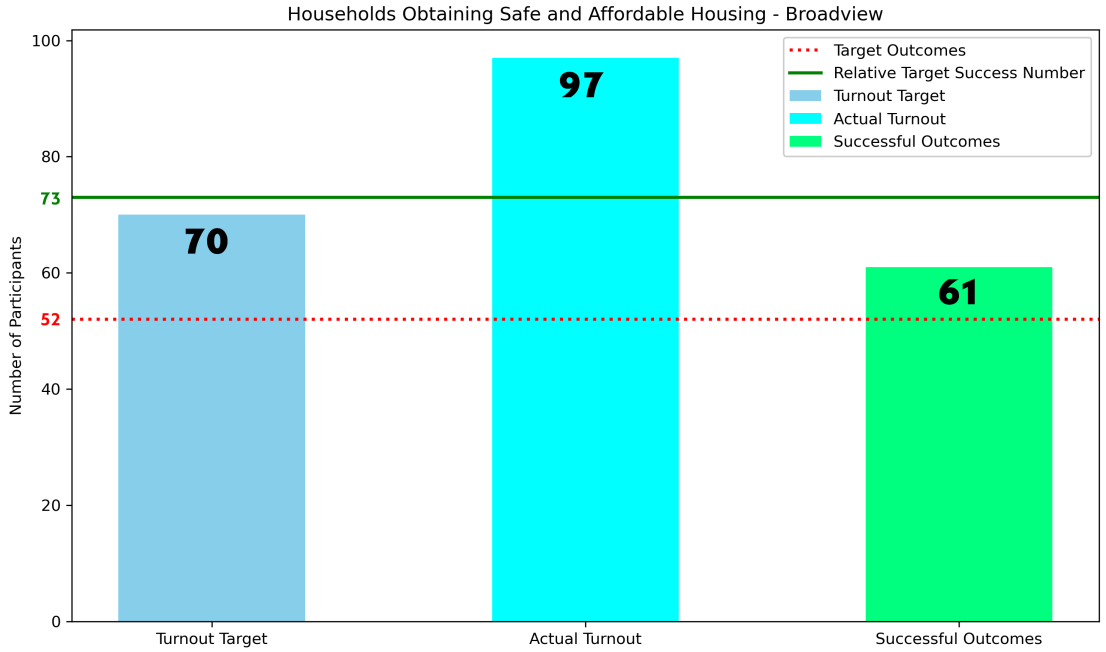
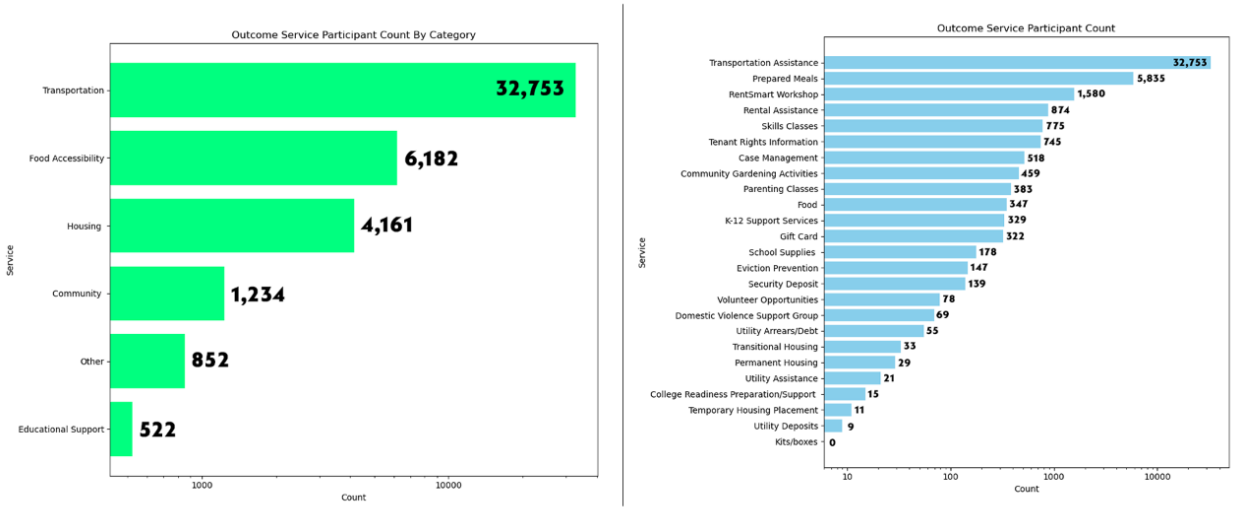


Figure 4: Bar Chart Displaying Miscellaneous Service Counts on Logarithmic Scale



3 Results

After analyzing and generating visualizations for each Outcome Service, we saw that a majority of services still met the relative success goal. Out of all the outcome services that met the fixed goal, only five did not meet the relative success target. 16 service programs that met both the fixed

Table 1: NPI Outcome Services Results

Category	Num. of Services	Fixed Successes	Relative Successes	Close call?
Broadview	4	3	2	0
Children's Services	1	0	1	1
CFE	1	1	1	0
HSP	4	4	4	4
JHRRH	2	2	0	1
Residential Services	4	4	4	2
Tenant Services	2	2	2	2
Volunteer Services	1	1	0	0

and relative targets, and within that number, every service within Children's Services, CFE, HSP, Residential Services, and Tenant Services was found successful. Another exception found in my analysis was the Children's Service's 'Children and Youth Who Improved Academic Progress," where the relative target was met but the fixed target was not. The reason for this exception was the factor of underperformance in target participant turnout. There also exist scenarios where the factor of a service's success and effectiveness was a close call than compared to the fixed target, where 10 services comfortably that made the fixed successful outcome target are within about 10% of the relative success target. All of the results of the NPI Outcome Services are summarized in Table 1.

4 Conclusion

Based on the statistics Solid Ground has provided, their programs are clearly effective in making change amongst individuals, families, and communities. While there will always be room for growth in evolving programs over time, the Housing Stability Program is the strongest among Solid

Ground's services. Meanwhile, the category of services that has the most growth appears to be between the Broadview category and the JourneyHome Rapid Re-Housing service, which depends on considering the fixed and relative goals. Regardless if Solid Ground prefers to function on setting fixed goals or chooses to consider relative goals in the future, the nonprofit organization needs to overestimate its participant turnout targets for its services since almost every NPI Outcome Service overperformed its participant turnout target. For the Miscellaneous Outcome Service Participant Count, the largest volume of clients served is unanimously in the transportation category, seeing over 32,000 people. Whereas for the least-participated category of service, Educational Support totals 522 participants. Considering the contrast in participant turnouts, Solid Ground should allocate more resources into Transportation Services in order to keep up with demand, whereas on the other hand, the organization should focus on enrolling more participants in Educational Support.

5 The Team and Future Progression

In my team, including myself, there are four members working on the Solid Ground data project. Since four Excel spreadsheets were provided to our group, we decided to divide the work by file among our team, where one person focuses on one collection of data.

5.1 Other Team Members' Progress

The group is developing interactive visuals using age, ethnicity, and gender factors from the Client Characteristics file, allowing community partners to track service utilization across demographics. The interactive feature enables users to compare individual factors and services, such as selecting a race and service in the race visualization. In addition, the team intends to provide a tool that allows

users to see the percentage of genders served, with the goal of providing a more accurate portrayal of demographic services.

In our Community Needs Assessment file, we focused on determining the importance and accessibility of various services in different zip codes. My group member's initial research focused on the importance of housing aid, which was visualized using a map chart of the Seattle urban area. Furthermore, my classmate designed a visualization that connects the number of responses to the average inequality value for each zip code, revealing the individual zip code when passed over with a pointer.

For the Customer Satisfaction Survey, we looked at the levels of satisfaction to see whether there were any relationships between client demographics and services received. The data suggests a consistent pattern across the board: if a customer scores one component of their experience highly, this pleasure is likely to be reiterated throughout other areas of the survey, implying consistency in the client experience regardless of the service aspect. The mapping assists in determining which locations have reported the highest and lowest levels of satisfaction with Solid Ground services.

5.2 Future Work

The team is looking to finish the process of organizing and understanding everyone's assigned provided data. Connecting our material will help us to achieve our ultimate goal of presenting a single narrative is the final crucial step. However, before we can fully connect our data, we must finish a few individual file activities, such as finalizing any addition visualizations, determining the conclusive narrative of our own respectively assigned files, and thus ultimately, connecting each other's narratives together to compose a cohesive conclusion.