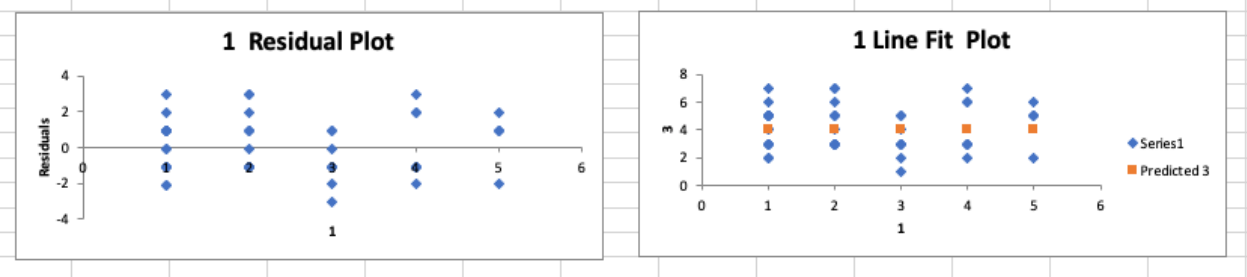
Experiment Results Explained:

The data is ordered as follows:

i. each participant is assigned a participant number, this refers to their unique visitor ID generated by hotjar. This was used in order to observe each participant's data individually, as opposed to viewing the collective data as a whole.
ii. the section that they most commonly visited was noted as well, this was determined by time spent within a section, number of clicks performed in that section, and the quantity of heat spots recorded within a section.
iii. the speed in which a participant completed the given questions was tracked by noting the time that they arrived on the web page and comparing that to when they eventually left the webpage. This data has potential to be skewed because participants may not have started right away or could have gotten distracted while performing the tasks. Since the experiment was not conducted with surveillance or within a controlled physical environment, it is difficult to tell what factors may have contributed to certain times. However, we noted that no participant spent more than 4 minutes on the tasks and no participants completed the tasks in less than a minute. While there are factors that may impact these numbers, this does give a sense of how a user may behave normally since they are not within a controlled environment — this may prove to yield more significant data. The quickest of two attempts was used to denote this variable.
iv. the average distance scrolled on the page was captured via the "AVG highest mouse scroll" variable. This was important to simply understand if participants were truly completing the tasks, this is possible because if the participant did not scroll all the way down to sections 6 and 7, it would be impossible for them to complete the tasks. Since each user was tasked to complete this twice, the average total distance traveled was used for this variable. The average scroll distance was 99% of the environment — which is expected because we would assume that participants would complete all tasks. For those (6 participants) that are less than 100%, the participant often only scrolled down about 75% of the environment in one of their tries and thus their average distance was calculated as 88% if they scrolled down 100% the other try. This variable was simply to understand if participants were completing the tasks correctly and ultimately proved that nearly all of them did.
v. the "AVG significant repeat hot spots" variable is the second most important variable that we tracked. This simply consisted of us counting the number of "hot spots" that a participants accumulated while completing the tasks. A hot spot represents a red zone — as shown in some of the screenshots. This red zone is an area that a user traveled to consistently and represents a specific x,y coordinate location in the environment. We were unable to track the exact coordinates of each hot spot, so instead we simply considered any other user hot spot within a 3EM distance of another to be similar enough to group — which is what the next variable is about. To give an idea of how much distance this is, consider the distance from the thin orange line to the bottom of the black as 3 EM. Em is a commonly used measurement for web development and is a standardized distance for doing so — we used this distance metric because our experiment's environment was built in WordPress as a website. After finding user heat maps with a similar x,y position, we grouped them together. We did this for each of the two trials and averaged the number of similar maps. For example, a value of 1.5 may represent a participant having 1 similar map in the first trail and 2 in the second; thus, the average repeat hot spots would be 1.5.

vi. the most significant variable that we observed was the “% shared most common hot spots” variable. This is because this variable takes into account the overall number of similar hotspots that a participant had. For example, if over the two trials a participant had a total of 4 similar hot spots, they would be similar to 4/56 total participants data. This was a metric that we used to determine how significant hot spots are in determining a unique user. We found many users with no other matches over the two trials, therefore producing the start of a unique user profile. Of course, factors to consider are time spent performing the tasks as well as their most commonly visited section. This is because users spending less time completing the tasks are providing less data by default and therefore their user profile is not as complete as others. Since this is an initial study, this is a great to start to understanding how this type of data can be collected at a large scale in order to determine how truly unique users are. Out of the 56 total participants, we were able to determine 19 unique user profiles with our given data — this is 34% of total participants. We expect that as the number of participants increases and the amount of data per participant increases, that a larger number of unique profiles can be constructed. This is because — as previously stated — some users spent very little time in the environment and therefore did not provide us with enough significant data to determine whether or not they have other unique characteristics. We also expect that as participants and data increase that the % of shared hot spots will decrease per participant as well — this is another factor that contributes to the number of unique user profiles.

Overall, we tracked each participant using their unique hotjar ID and assigned them new ID’s for our data (1-56) based off of the order that they completed the tasks. We tracked the most commonly visited section to see if there was any affect on similarly based off of sections most commonly visited and determined that it is not an effective way to determine if a user’s maps will be similar to others that visited the same section. We did so by running a linear regression model and found very little correlation between the two.



We found that there was a somewhat correlated relationship between the amount of time spent completing the tasks and the number of significant repeat hot spots for a participant. This makes sense because in general, we were able to grab more data for those who spent longer performing the tasks than those who didn’t. Though, it is also possible that if certain users were completing the tasks either together or one after another, then the data could be somewhat skewed. We expect this is the case for instances where two participants that are performing the tasks one after another and one user takes significantly longer than the other, but they have similar hot spots. This could potentially be the first user completing the tasks and showing the second what to do — therefore performing the tasks in a similar manner and making the hot spots obsolete. This is something that I expect with some of the less tech

savvy participants that we had. We have college students across various different majors, adults with and without college education, and users over the age of 70 who often have far less technical experience than the others. Because of this, we are getting a ton of data from various different types of users and therefore we are bound to have some conflicts along the way.

A good sign that we had was that while the number of average significant hot spots changed, the percent that the user's maps were similar to others did not correlate. This is good because that means although users may be similar to others, the amount of hot spots that they have does not make them less likely to be unique. We expected that if users were to have many hot spots, they would be more likely to be similar to others because there is a greater chance of shared spots on their maps; though, this was not true. Rather, it seems as if users with similar hotspots often did not take much time to complete the tasks and thus had less significant data to work with, making those results similar by default and by the nature of the basic tasks. We can tell that this is true by yet again running another linear regression on the time spent completing the tasks as the Y variable and the % of shared hotspots as the X variable. We observed a very small standard error of 0.03291698 and therefore believe that, again, with more data, we would expect to see even more significant correlations between the time spent and the percentage of similar hot spots. This would ultimately yield more unique users because if those who spent less time were to spend an equal amount of time completing the tasks, we would produce more accurate unique user profiles.