Map Marker Accuracy:
Evaluating Spatial Data Quality in a Simplified Mapping Survey

by

Christopher Sean Ghormley

A GIS practicum report submitted in partial fulfillment of the requirements for the degree of

Master of Science
in
Geographic Information Science

MS in GIS Committee:
Jiunn-Der (Geoffrey) Duh, Co-Chair
David Banis, Co-Chair
Christopher Grant
Martin Swobodzinski

Portland State University
2025

Abstract

Collecting spatial information using maps is a recurring need in research and policy (Evers, 2020; Gottwald et al., 2016; Hitchins, 2018; Tang & Liu, 2016). Survey participants frequently abandon map-based surveys at the first mapping task (Poplin, 2015). In his survey of wildfire risk management personnel, Evers (2020) applied a simplified web-mapping interface that uses circles as the sole input geometry for all spatial representations. The current project examines an updated version of this web map. An anonymous survey gathered data from participants who completed six mapping exercises within Portland, Oregon, and provided feedback on the survey experience. A working hypothesis proposed that greater experience with web maps and/or GIS, higher familiarity with the target locations in the survey, or time lived in the region, would improve spatial accuracy. Submissions were compared to target geometries using three metrics: centroid distance, area coverage, and shape similarity. Participants successfully marked point and area locations with 90% and 95% positional accuracy, respectively. The effects of hypothesized factors on accuracy were either nonexistent or statistically insignificant. Linear and logit regression analysis showed that map display resolution was significantly associated with data quality ($p<0.001$), explaining 32.6% of variance in centroid distance when a Box-Cox transform was applied ($\lambda=-0.0202$, $p<2e-16$). Participant-specified regions exceeded target size on average by a factor of 1.5 when boundaries were visible. The effect of input device on positional accuracy approached statistical significance ($p=0.056$), with phone users exhibiting higher error rates (12.5% vs. 7.2% for laptops and 1.9% for desktops). These findings validate the circle-based approach for collecting reliable spatial data with minimal training, while suggesting specific interface improvements such as enforcing minimum zoom levels. Results have immediate application for wildfire risk management data collection and broader implications for PPGIS applications at similar map scales.

Dedication

When embarking on a project you need other people to keep you sane. I owe a huge debt of gratitude to the family and loved ones who helped me along the way.

To friends and colleagues who offered support and encouragement.

To my parents, brother, and sisters, who were there when I needed them.

To Veronica, for everything.

My profound thanks to you all.

Acknowledgements

Completing this project has been a journey that was only possible because of the generous efforts of faculty and staff of the Geography, and Environmental Science and Management departments at Portland State University. I am indebted to the following individuals.

To David Banis, for discussions on public-participation GIS projects that provided context for this one, and office hours beyond the call of duty.

To Dr. Cody Evers, for coaching me on survey strategy and analysis ideas, and patiently explaining the work this project was meant to support.

To Dr. Geoffrey Duh, for thoughtful and timely editorial guidance that exceeded my expectations, and infectious enthusiasm for geography and GIS specifically.

To Jackson Voelkel, for giving me a running start on spatial data analysis in R.

Thanks also to Dr. Max Nielsen-Pincus, Dr. Christopher Grant, Joann Ng, Dr. Andrés Holz, and Dr. Martin Swobodzinski.

I would particularly like to thank the anonymous fellow graduate student who assured me that if I kept working on the project, I would eventually finish it.

While many individuals helped me along the way, any remaining errors are mine alone.

Support for this work was provided through a Graduate Research Assistantship funded by the 2021 Bipartisan Infrastructure Law, titled "Fueling Adaptation: Leveraging Community Capacity to Reduce Wildfire Risk" (BIL Project Number: POC2).

Table of Contents

List of Tables

List of Figures

Introduction

In October 2023, I was approached by David Banis and Cody Evers to support an upcoming survey project involving spatial data collection on the web. The US Forest Service was gathering information about wildfire risk reduction work, at the Wildland-Urban Interface (WUI)[1] as a part of the Bipartisan Infrastructure Law-funded "Fueling Adaptation" project Wildfire Governance Survey (WGS). A key instrument in this project would be an online survey of personnel responsible for making land management decisions or performing educational outreach to landowners. Survey data, including work scope, project geography, and professional connections, would help the research team better understand wildfire management capacity and coverage while simultaneously mapping a professional social network. To support the WGS, I adapted an existing web map implementation (Evers, 2020) to a new data collection site that embeds in the selected online survey platform. The combination of an online survey with a spatial data entry tool may be considered a form of Public-Participation GIS (PPGIS).

The current project applied the same spatial survey workflow used by the WGS. A better understanding of the accuracy limitations that pertain to these data may help guide the selection of this tool for future survey campaigns. Involving individuals who are not expert Geographic Information System users or trained geographers in information gathering or decision-making processes activates a broader knowledge base and ideally includes those whose experience would not otherwise be recorded (Gottwald et al., 2016). Enabling wider access implies ease of use and low training overhead (Gieseking, 2018; Gottwald et al., 2016). The user interface design of the mapping tool and indeed the survey in which it was embedded were thus focused on minimizing barriers to entry. The survey required minimal training and was accessible on virtually any graphical user interface with internet access.

While this report was being prepared, a historic burning of several communities around Los Angeles, California was underway. The largest fire, known as the Palisades Fire, started on January 7, 2025, and burned 23,448 acres before being recorded as fully contained on February 10, 2025. It destroyed 6,833 structures and killed 12 people ("Palisades Fire," 2025). The urgent need to understand the scope and effectiveness of wildfire risk reduction activity is beyond doubt. One can certainly imagine the utility of incorporating PPGIS processes in disaster preparation, response, and recovery.

Lands vulnerable to wildfire may fall under any type of ownership: municipal, county, regional, state, tribal, federal. The departmental mission may vary significantly between different landowners/stewards. And yet one fire may affect several jurisdictions simultaneously. Funding is most needed where the risk is high and existing capacity to mitigate that risk is least. Understanding organizational connectivity on the ground (project adjacency) as well as in professional relationships and collaborations, is valuable when steering funding. The resulting survey data may also help land managers to understand the effectiveness of different kinds of risk mitigations, such as direct fuel reduction programs or "defensible space" promotion campaigns at the WUI (Anderson et al., 2024).

[1]"The WUI is the zone of transition between unoccupied land and human development. It is the line, area or zone where structures and other human development meet or intermingle with undeveloped wildland or vegetative fuels." (*What Is the WUI?*, n.d.)

The project described in this paper took shape in February and March 2024. Before the survey arm of the study (Campbell et al., 2024) went live, the team expressed interest in independent validation of the survey method.

In collaboration with Cody Evers and David Banis, I assessed the suitability of an interactive web map with a simplified interface for collecting spatial data within a web survey environment. The results provide some confidence that a web map survey can produce consistent, accurate spatial data with minimal participant training, while also identifying systematic weaknesses for future improvement.

Contributions to Geographical Information Science

This work contributes to the practice of geographical information science in two ways. First, by refining a simple-to-use interactive web map for collecting spatial data, as for a PPGIS project. Second, by characterizing the quality of the spatial data collected through the interactive map. The data collected show that most users who are familiar with locations (e.g., a neighborhood park or a building) can accurately mark them on a map with minimal training, whereas marking the extent of an area (e.g., the perimeter of a neighborhood, park, or land parcel boundary) is more challenging. The instrument used to collect data could affect the precision and accuracy of the data. To facilitate appropriate use, data should be reported with the best available information about their accuracy. This project describes an approach to estimating the accuracy of user-submitted data on an interactive web map. It also suggests some approaches to reporting the spatial uncertainty of PPGIS data.

Background

Location data play a critical role in web research surveys, enabling precise event or object documentation, project planning, and targeted interventions. Embedding a map in a survey is not novel, but it is far from routine. Even in advanced survey platforms like Qualtrics, integrated geospatial data collection remains challenging. To bridge the gap, Qualtrics introduced an ArcGIS integration, but it requires an Esri backend to operate (Qualtrics, 2024). A more flexible (e.g., open source) option provides some leeway for user interface experimentation while avoiding additional licensing fees. The high accessibility and low cost of such a platform makes an attractive tool for a subset of PPGIS applications. It can also answer a specific call for inexpensive, "good enough" GIS software (Gieseking, 2018).

Debates over the utility of PPGIS data have gone on for decades (Schuurman, 2000). It may be best to simply yield to the argument that data production in a GIS is inherently biased and proceed with caution. Perspectives on how PPGIS data production fits into GIS and geographic understanding more generally range from the conceptual (Schuurman, 2006) to the practical (Gottwald et al., 2016). As a practical matter, GIS professionals are bound by a code of professional ethics that does not apply to every volunteer. What are appropriate limits of trust that should be extended to these data?

Like all user-entered data, PPGIS data are susceptible to errors. Higher precision should not be mistaken for higher accuracy or reliability. Previous work (Tang and Liu, 2016) developed a framework for rating several PPGIS applications' "strengths, weaknesses, opportunities, and threats", using language developed in strategic planning literature since the 1960s, further described and reviewed in Benzaghta et al. (2021). Tang and Liu point out that applications benefited from simplicity but necessarily suffer data quality issues. These can range from software operation errors due to inadequate training or inattention, to judgement errors, and as they state "[e]ven specialists will make mistakes". They conclude that the highest quality inputs to PPGIS projects are likely to come from experienced contributors. In a situation where every contribution is from a new participant, some level of error is inevitable and probably impossible to filter out.

In this vein, previous work has investigated the usability characteristics of PPGIS tools (Gottwald et al., 2016; Hitchins, 2018; Poplin, 2015). Gottwald et al. (2016) performed a usability assessment on one tool, applying the conclusions of a pilot study to improve software accessibility. Hitchens (2018) measured user interface interactions with an online survey collecting recreational uses of public lands in Central Oregon. User interface design choices influence submitted data accuracy and precision, and it is important to ensure that a data collection tool fits the intended purpose.

Hitchins analyzed how map users produce GIS data in the context of a specific survey tool[2]. The project analyzed user behavior, such as zooming and panning the map, that influenced data accuracy and completeness. He emphasized a need to "calibrate" PPGIS data collection tools, much as any instrument needs to be calibrated. This is an important reference for the current project, as Hitchins' conclusion provides actionable guidance for reducing error rates and improving participant retention in future projects.

A key problem Hitchins identified was the inflexibility of the mapping data entry interface, that is, the inability to edit or delete created data and annotations. The total length of the survey activity was also a potential problem, since over a quarter of participants stated that it took longer than expected, and the final question yielded the lowest participation rate. To some extent this can be seen as an expectation management problem, but respect for volunteers' time dictates removing unnecessary survey steps. Finally, datasets from the mapping and textual portions of the survey were linked through a manually entered unique ID.

Every survey project requires numerous tradeoffs to match available time and resources to required data quantity, quality, and volume. A typical PPGIS project, such as one discussed in Poplin (2015), is an unattended website with little or no supervision of data entry. Once participants arrive at the site and view the interface, the design must effectively lead them through the process of data entry. It is vital to balance the motivation to collect more information from individual participants with the risk of losing data when participants abandon a survey task because of question difficulty, interface complexity, or excessive explanatory text. Indeed, 43% of the participants in an early experiment reported in Poplin (2015) stopped at the first mapping task. The same study found that about half of a cohort of students with recent GIS experience had difficulty creating line features in an unfamiliar web GIS interface.

The experience gap, seen with a substantial fraction of survey participants in a typical unsupervised data collection effort, prompts optimization for ease of use. A typical GIS editing interface supports the creation of points, lines, and polygons by drawing individual vertices. The interaction model for creating each feature type is different, and there are many ways to do it wrong, as this author has learned from experience.

Considering the lessons learned from previous research, some adjustments are worthwhile. The web map used in this project, simply called the "point it out" tool (PIO), enables point deletion. PIO also eliminates the user ID entry step as a potential point of failure. The map layer drawn by PIO is embedded in a Qualtrics survey, creating a seamless interface and facilitating subsequent GIS

[2]"Connections with Forests and Grasslands of Central Oregon Survey" (CFGCOS)

analysis alongside non-spatial data analysis. Since the source code for the map layer was open source, the data entry features were tailored specifically to minimize participant training burden and the time to complete.

An alternative used for the implementation studied here, as an alternative to points, lines, and polygons, is to draw only circles. Zooming the map in and out can change the scale of the circle, to indicate greater precision. The circle's centroid can stand in for a point, and a pattern of circles can be employed to draw a line or fill in a polygon, with less training. These shapes can operate at different scales, whether to indicate a particular building or event site, a section of a river, the extent of a park or a stand of trees, a neighborhood – or parcel of land being considered for a prescribed burn.

PSU researcher Cody Evers applied a version of the software evaluated here to wildfire management research (Evers, 2020). That configuration presented multiple map layers and provided the option to annotate points with text or by selecting from a list of descriptors. The survey collected detailed background information from each survey participant, project location data with annotations, and referrals to additional survey participants.

Because this project recruits from a pool of volunteers with a variety of experience levels and focuses on specific usability and accuracy questions, it uses an abbreviated survey and fewer user interface elements (as compared to Evers, 2020). The configuration used for this project uses only a basemap, with no switchable layers, and no text or checkbox annotation. The FA survey uses the same basemap, with a few minor configuration changes detailed in the Methods section.

This project investigates the following questions:

1. What quantifiable spatial errors occur when volunteers mark locations, and what methods can be used to evaluate data entry deviations?
2. Using the proposed evaluation rubric, what is the empirical error rate, and what does this reveal about data source reliability?
3. How do systematic data patterns inform approaches to mitigating geospatial data entry errors?

Methods

Interface Design

The web mapping tool used in the project was adapted from previous work (Evers, 2020) and had already been selected for the Fueling Adaptation Wildfire Governance Survey. Its development was guided by the principle of balancing data precision with user accessibility. While individuals in the target demographic of potential volunteers typically possessed basic familiarity with web maps, they would not be expected to have substantial experience with map-based data entry systems. This required the development of an interface that would function seamlessly across multiple platforms, accommodating both traditional desktop computing environments and touch-enabled mobile devices, while requiring minimal training.

These design goals drove a compromise. Though the underlying software architecture supports any GIS geometry—points, lines, and polygons—this tool uses only circular geometry. This decision was driven by several factors: reduction of cognitive load during the learning process, minimization of training requirements, and mitigation of potential user errors during unsupervised data entry sessions.

In practice, the circular data entry tool proved versatile. Individual circle centers effectively represented point locations, while sequences of circles could be employed to approximate linear features. Additionally, overlapping circles proved adequate to draw polygons, thus fulfilling all requirements within a simplified paradigm. In this implementation, zoom level adjusts the area covered by the circle and provides explicit context to the analyst about the map scale displayed at the time of data entry. The single-shape approach effectively eliminated confusion related to tool switching, while maintaining sufficient functionality for the research objectives.

Technical Architecture

The "Point It Out" (PIO) tool used in this project, adapted from Evers (2020), was reimplemented as a database-based web application with Docker Compose (Merkel, 2014), leveraging the Django web framework (Django Software Foundation, 2024) with a PostgreSQL database backend (PostgreSQL contributors, 2024). The system's cartographic functionality was built on the OpenLayers JavaScript library (*OpenLayers*, n.d.), allowing for integration with any compatible XYZ tile service, including ArcGIS Online for basemap provision (Esri, 2024). The implementation included a facility for additional vector layers and incorporated configurable parameters for map extent and zoom limitations, to ensure appropriate spatial constraints for specific research contexts.

The Django framework simplifies database-backed application development by translating the application data model between Python code and SQL and a built-in administrative website backend for editing configurations and downloading data.

The site was designed for embedding in an HTML inline frame to provide a seamless survey interface to the user. A short JavaScript for each question in the Qualtrics survey using map input passed required metadata into the mapping application through a URL template.

For a more complete listing of software used in the map, see Appendix C – Software.

User Interface Implementation

The interface design prioritized operational simplicity. Navigation functionality adhered to established conventions, implementing scroll-wheel and pinch-gesture zoom controls, alongside click-and-drag or touch-and-drag panning mechanisms. As the user zooms into the map view, the detail level increases, and the area represented by the circle on the map also increases (while the pixel radius holds steady). Desktop and laptop users received hover-based circle previews, while mobile users were presented with explicit real-world diameter measurements via text display. Placed markers were rendered as transparent circles to maintain visibility of underlying cartographic features. Markers were removed from the map by activating a placed marker and confirming the deletion.

The map used for this project centered on a portion of central Portland, Oregon, where survey recruitment took place. The Esri National Geographic basemap (Esri, 2024) was selected for its balanced level of detail across zoom levels. The interface was bound to an extent, range of zoom levels linked to circle sizes, and vector overlays linked to a specific URL. The same server could thus display and record data for multiple map configurations – from different surveys/survey questions – simultaneously.

Data Collection Framework

The system's data collection architecture captured a set of spatial and temporal parameters for each user interaction. The system creates a record when a user taps or clicks on the map, including the point as longitude and latitude, radius in meters, displayed map resolution in meters per pixel, the creation timestamp, a response identifier linking the point to a user interaction, and an activity identifier associating the point with a survey question. The system preserved data regarding deleted points, facilitating subsequent analysis of user behavior and task complexity.

Survey Design

The survey collected relevant user characteristics and led participants through a series of exercises, followed by questions about the user's experience with the map application itself. Survey mapping tasks include identifying buildings or points of interest and marking regions on the map. The intent was to evaluate how successfully participants would locate and mark described points of interest that may be familiar to participants who live in the area. Site selection was important because the relevant question was mapping usability, not site familiarity. A secondary goal was to characterize the range of inputs that could be expected from typical map users. Centroid distance, area ratio, and shape similarity were calculated for each submission and target. Mapping results were joined with volunteered participant information to identify trends.

The tool's efficacy was evaluated through a structured survey comprising six exercises. These were divided equally between point-marking tasks, focusing on specific buildings and parks, and area-definition tasks targeting geographic regions. This exercise framework was designed to assess user performance across a spectrum of scenarios, ranging from unambiguously marked locations to areas requiring local knowledge or external reference consultation.

The survey presented six map data entry exercises to each participant. There were three point-marking and three area definition tasks. For the point exercises, the task was to place a point on the described target, which was either a building or a park. The specific instruction was, "**Click or tap** the center of the feature on the map to create a circle." In the context of the Wildfire Governance Survey this project was supporting, the points selected for this survey were analogous to individual buildings being considered for defensible space evaluation. The instructions specified an ideal circle size, implying that the map should have been zoomed in to a high detail level. These three locations were Powell's City of Books, the Millar Library at Portland State University, and Council Crest Park. For the area filling exercises, the task was to approximate a polygon drawn with individual vertices. These areas were selected for expected familiarity among target participants and would have been analogous to treatment project areas in the Wildfire Governance Survey.

Area exercises for this project were selected to provide a variety of challenges: one clearly delineated area on the basemap (Mt. Tabor), one that required matching map detail to textual information in the survey question (Willamette River), and one without any textual or basemap reference (Pearl District). Wildfire Governance Survey responses could have resembled any of the above scenarios. A region within a National Forest might have been clearly marked on the selected basemap, while marking a private land parcel might have required outside reference. It was most important to examine the case where there was the least guidance for data input. This was where the survey team would depend on the participant's existing knowledge and resourcefulness the most.

Zoom in or out to set the circle diameter you want, then click or tap to create overlapping circles to fill the area. Use as much or as little detail as you like. It's OK to not completely fill the area. Just cover most of it.

Since the instructions invite low-detail submissions, the only criteria used to select valid submissions was centroid distance.

Complete survey contents are listed in Appendix A - Survey Questions.

Recruitment

The recruitment process was executed in two distinct phases. The initial phase utilized institutional channels, specifically PSU class email distributions, during March and April 2024. This was followed by a social media-based recruitment effort via Facebook in June 2024, which proved particularly effective, generating approximately fifty percent of total survey responses within a three-day period.

Data Processing

Data from the survey were downloaded as a comma-separated values (CSV) file from the Qualtrics website, using the "text" option. Data collected by the web map application were also downloaded as CSV from the Django admin backend site. A raw data record from the mapping application contains the fields listed in Table 1.

Table 1. Spatial database fields.

Field Name	Type	Purpose
ResponseID	Alphanumeric	ID field from Qualtrics
ProjectID	Integer	Exercise number
Radius	Float	Radius of the circle
Geom	Point (SRID 4326[3])	Location of cursor when clicked
Resolution	Float	Map display resolution meters/pixel

For point exercises, the input should be a single row in the database. Area exercises may contain an arbitrary number of rows.

These tables were imported into RStudio and joined by the Qualtrics-generated response ID, merging the data into one table containing all the submitted points and all the survey responses.

The survey contained an ungraded tutorial activity, three point-selection activities, and three area-filling activities. Area responses (consisting of individual circles) were joined into single polygons using the *st_union* operation from the sf library in R (Pebesma, 2018). This dataset was filtered and summarized in R, using the libraries listed in Appendix C – Software to generate all the tables, maps, and charts reported below.

[3] The spatial reference identifier 4326, denoting a longitude/latitude pair, is also known as WGS 84. Before performing any calculations, all points were projected into the Oregon Lambert spatial reference system, using metric units. This projection has an estimated accuracy of ≤1.0 meters for distances and ≤0.01% for areas ("Oregon Coordinate Systems," n.d.; "NAD83 / Oregon LCC (m)," 2021; *Oregon Lambert Error Statistics*, 1996).

Evaluating Responses

All exercises have an ideal target polygon representing the named building, park, geographical feature, or neighborhood, digitized in QGIS (QGIS.org, 2024) from an OpenStreetMap base layer (OpenStreetMap contributors, 2024), or drawn from public records (City of Portland, n.d.). Individual exercises were assigned to "valid" and "invalid" groups by the criteria described below.

Point Placement

For point exercises, the input should be a single circle input recorded as longitude, latitude and a radius. The recorded radius is determined by the zoom level of the map when the point was created.

Points were considered valid if they fell within the target (the block, building outline, or area boundary found on the map), determined by the *st_intersects* operation from the sf library.

The Cartesian distance (st_distance) between the point and the target centroid[4] allows comparison between participants. The Cartesian distance divided by the equal-area circle radius ((1) gives a normalized relative centroid distance ((2), used to compare performance within and between exercises. A valid submission's relative distance is less than one.

$$r_{equal-area} = \sqrt{\frac{A_{target}}{\pi}} \tag{1}$$

$$Relative\ Distance = d_{relative} = \frac{d_{centroid}}{r_{equal-area}} \tag{2}$$

The relative radius is the submitted radius divided by the radius specified in the instructions, used to determine how well participants followed instructions ((3).

$$Relative\ Radius = r_{relative} = \frac{r_{user}}{r_{exercise}} \tag{3}$$

Area Filling

Area exercises present more challenges than point evaluation. Comparing geospatial polygon shape and size must be factored in along with location when matching volunteered geographical information to authoritative data (Fan et al., 2016), or when grouping global navigation system traces together into common routes (Kim et al., 2024). Kim et al. use the centroid distance between polygons, the shape similarity calculated by the ratio of the area and perimeter, and similarity of area to group similar polygons together.

The objective of ranking survey submissions is different in this case, but the methods used are similar. For the purposes of this project, the three measurements considered were (1) distance

[4] All spatial operations described operate on "geometry" objects projected to the Oregon LCC metric projection using the *sf* library ("NAD83 / Oregon LCC (m)," 2021; Pebesma, 2018). The spatial extent of objects and distances in this study is small enough to make any errors from these approximations negligible.

between submission and target centroid, (2) area similarity, and (3) shape similarity. These three measurements provide a concise characterization of the submitted polygon.

Relative centroid distance is calculated as in the point evaluation above. Valid submissions are those within the equal-area circle (relative distance less than or equal to one). An alternative approach would accept any submitted polygon with a centroid falling inside the target polygon. This works poorly when the target shape has a long aspect ratio, where the submission could have little of its area falling inside the target. A submitted polygon with a centroid falling inside the target but outside the equal-area circle is unlikely to approximate the target. In the end, the submitted polygons that match the target best have similar centroids, similar areas, and similar shape indices.

Two ways of comparing areas between survey submissions were considered. The relative area, defined in (4, shows the degree to which submissions under- or over-estimated the size of a target. A second metric, the area score (defined in (5), gives a unidirectional deviation from the target area.

$$Relative\ Area = A_{relative} = \frac{A_{user}}{A_{target}} \tag{4}$$

$$Area\ Score = \left| \frac{A_{user}}{A_{target}} - 1 \right| \tag{5}$$

The shape index ((6)[5] can be interpreted as "circularity" or "roundness". The result of the shape index formula is exactly one for circles, and less than one for non-circular shapes. It is about 0.785 for squares, and 0.260 for a rectangle with an aspect ratio of 10:1. This is a simple but useful measure to rank the similarity of submitted shapes to the target. Of the target shapes used in this project, the Mt. Tabor target has an index value of 0.624, while the Willamette River shape is much longer, yielding a value of 0.296. The Pearl District polygon has an index of 0.693.

$$Shape\ Index = I_s = \frac{4\pi * Area}{Perimeter^2} \tag{6}$$

The shape defect formula ((7) helps indicate shape similarity, but its value can be counterintuitive. Consider that the defect between a square and a circle is 0.215; between a square and a rectangle with an aspect ratio of 3:1 it is 0.196. Since the circle input method is inherently imprecise, a significant defect should be expected.

$$Shape\ Defect = D_s = |I_{s0} - I_{s1}| \tag{7}$$

While the acceptance criteria for areas are based solely on centroid distances, relative areas and shape defect values are reported to show how they differ from the targets. High quality submissions have favorable values on all three indices: low centroid distance, relative area close to unity, and low shape defect.

[5] The shape index used here is also used to evaluate the degree of legislative district gerrymandering (Polsby & Popper, 1991), or the isoperimetric ratio. For a discussion and links to other sources see also user25201 et al. (2014, February 7)

Results

Dataset Metrics

During the survey collection period from March to June 2024, 74 survey responses met selection criteria. Of those, 10 participants did not interact with the map at all. It is unknown whether that was due to lack of time or interest, or because of a technical issue. There were seven activities, including a tutorial and six exercises. There were 57 responses to the first exercise and 51 responses to the sixth exercise, marking a slight drop in completion rate over the course of the survey. The survey software recorded 53 "finished" surveys (final submit button clicked). There is nothing recorded in the survey software on mapping question completion, so some discrepancy between response rates is expected.

A pilot version of the survey was distributed in early March 2024. The purpose of the pilot was to test the survey and map integration, refine questions, and estimate time to complete the survey. Pilot survey responses can be identified by comparing the data types and contents of the responses. Non-anonymous responses were also created in the process of testing the survey.

The data obtained from survey responses not marked as "anonymous" or from a prerelease version of the survey (i.e., test responses) were removed. After these responses were removed, 57 records remained, with declining response rate across six activities. The final mapping exercise collected 51 responses.

The median time to complete the survey was about 12 minutes (Figure 1), with a range of 4 to 32 minutes. This excludes surveys that recorded much longer durations, which were assumed to be mostly idle time.

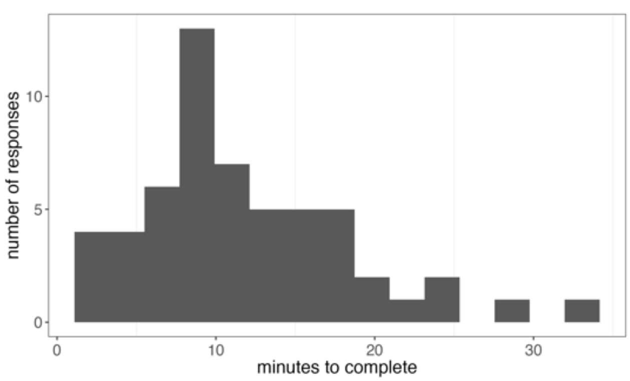

Figure 1. Histogram of time to complete survey.

The largest self-reported age group was 45-54 years old, at 27 participants. The next largest group was 25–34-year-olds, with 8 participants. The majority of the 45-54's who completed the survey (18/27, 67%) did so in June 2024. This is not surprising, since the author shares this age group with much of the social network from which many of the participants were recruited (Table 2).

Table 2. Survey submissions by age group.

Age group	n	percent
Under 18	0	0.00%

Table 2. Survey submissions by age group.

Age group	n	percent
18-24 years old	1	1.90%
25-34 years old	8	15.40%
35-44 years old	6	11.50%
45-54 years old	27	51.90%
55-64 years old	6	11.50%
65+ years old	4	7.70%
Total	52	100.00%

The device used for data entry varied. A substantial minority of completed responses (15) were submitted on phones, which highlights the importance of designing for and testing on small touchscreens (Table 3).

Table 3. Survey submissions by user device.

User device	n	percent
Desktop	9	17.30%
Laptop	26	50.00%
Phone	15	28.80%
Tablet	2	3.80%
Total	52	100.00%

Most participants reported that the survey took about as much time as they expected (Table 4).

Table 4. Survey submissions by perceived time to complete.

Perceived time to complete	n	percent
Quite a bit less time	1	1.90%
Somewhat less time	8	15.40%
About as much time as expected	35	67.30%
Somewhat more time	7	13.50%
Quite a bit more time	0	0.00%
NA	1	1.90%
Total	52	100.00%

Differences between users might be better understood by analysis of mouse movement, map panning, and clicking behavior, zoom levels, and idle time between interactions. These data were collected with a "user behavior" script running on the map website but not analyzed as part of the project.

Survey Participant Background

The survey attempted to qualify the user by level of experience with web maps, GIS or map editing tools, and familiarity with the geographic area covered by the survey mapping exercises. The survey avoided asking personally identifying questions such as name or email, and did not provide a field for race/ethnicity or sex since this information was not considered relevant. There was a field for ZIP code, but due to the small number of responses this field is not useful (33 unique ZIP codes are represented, three outside the Portland Metro area). Qualtrics provides an IP-based geolocation, but it does not allow us to infer anything useful. Survey participants were asked the following questions prior to any mapping activity:

1. How familiar are you with mapping and way finding apps, such as Google Maps, Apple Maps, AllTrails, TrailForks, etc.?
2. Have you used any GIS or map editing tools, such as ArcGIS Pro, ArcGIS Online, MapBox, Carto, Felt, etc.?
3. How long have you lived in the Portland area?

The first question, "How familiar are you with mapping and way finding apps, such as Google Maps, Apple Maps, AllTrails, TrailForks, etc.?" returned 52 "moderately" or "extremely" familiar responses, two "slightly" and one "never used" response. The user base largely knows how to operate a web map (Table 5).

Table 5. Survey submissions by web map experience.

Web map experience	n	percent
Never used	0	0.00%
Not familiar at all	1	1.90%
Slightly familiar	2	3.80%
Moderately familiar	26	50.00%
Extremely familiar	23	44.20%
Total	52	100.00%

The second question, "Have you used any GIS or map editing tools?" returned 35 users with some experience (23 extensive) and 22 users who had never edited a map or used GIS before (Table 6).

Table 6. Reported GIS experience.

GIS experience	n	percent
Never used	21	40.40%
Used casually	10	19.20%
Have extensive experience/formal training	21	40.40%
Total	52	100.00%

A clear majority (32 of 52 completed responses) reported living in Portland over 15 years. Only six had lived in Portland for less than a year (or not at all) (Table 7).

Table 7. Reported time lived in Portland.

Time lived in area	n	percent
Never	3	5.80%
Less than a year	3	5.80%
1-5 years	4	7.70%
5-10 years	6	11.50%
10-15 years	4	7.70%
More than 15 years	32	61.50%
Total	52	100.00%

Mapping Experience

As the survey concludes, a few questions evaluated the user experience.

To the question, "How familiar are you with the places named in this survey?", about half (25/52) reported that they were extremely familiar with the places mentioned, and 22 said they knew most of them. Five participants were only slightly or completely unfamiliar with the places in the survey (Table 8).

Table 8. Reported familiarity with survey locations.

How familiar were survey locations?	n	percent
Never visited/unsure	0	0.00%
Not too familiar (one or two)	1	1.90%
Slightly familiar (some of them)	4	7.70%
Moderately familiar (most of them)	22	42.30%
Extremely familiar (all of them)	25	48.10%
Total	52	100.00%

Regarding map usability, 43/52 found it either extremely or moderately easy to read details on the map. Three participants found reading the map "extremely" or "moderately" difficult, and seven gave a neutral response (Table 9).

Table 9. Reported ease or difficulty of reading map details.

Easy to read details	n	Percent
Extremely difficult	0	0.00%
Moderately difficult	2	3.80%
Neither easy nor difficult	7	13.50%
Moderately easy	28	53.80%
Extremely easy	15	28.80%
Total	52	100.00%

On the question of map controls (zoom, pan, marking points), 48/53 participants found the tools easy to use, and 5 gave a neutral or negative answer (Table 10).

Table 10. Reported ease of manipulating map.

Easy to zoom/pan/mark	n	percent
Extremely difficult	1	1.90%
Moderately difficult	1	1.90%
Neither easy nor difficult	2	3.80%
Moderately easy	28	53.80%
Extremely easy	20	38.50%
Total	52	100.00%

When asked whether additional map elements would improve usability, nine participants called for additional layers or map details (Table 11). Among these, one common observation was that street labels were inconsistent. Three suggested that a map legend would help. Two requested "online help, tutorial, or other documentation". Most (31) reported "none of these" or checked none of the options.

Sixteen participants responded to the question "Was there anything else about the map that was particularly helpful? Difficult? Anything else you noticed?". These fall into a few themes. The full set of comments is found in Appendix B – User comments.

Table 11. Summary of comments from survey.

Feature request summary	n	Comment
Difficulty zooming	5	Suggestion: Enter the scale numerically
Map details missing (street labels)	6	Common issue among tile-based web maps, hard to fix.
Layer missing	5	Suggestion: add neighborhood boundary layer to help with exercise 6/Pearl District
Missing feature: draw rectangle	1	
Difficulty with touch interface	4	Expressed difficulty zooming or placing points on phone or tablet
Interface challenge: make point creation easier	2	Users expressed frustration, but it was unclear how to do this in practice.
Unclear instructions about how to fill in areas	2	

The survey design was a compromise between maximizing data collection per record and reducing survey duration to prevent participants from dropping out. There were fewer mapping exercises and less explanatory text than originally planned. As expected, some of the comments reflect inattention to (or lack of clarity in) the instructions.

At least one lost opportunity resulted from this instructional brevity. There was no guidance on the use of outside reference to complete mapping exercises, and no follow-up question to determine which references had been used, if any. Controlling this variation could have added important context.

Regarding exercise 6, the Pearl District, one participant comment stated, "neighborhoods here have wonky and unclear borders, so what exactly are the district limits...and who cares." Considering this comment along with the other participants who complained about the lack of a neighborhood boundary layer, the exercise appeared to cause genuine confusion. In retrospect, the concept of a neighborhood boundary can be approached from a functional, sociological, or political angle, and the way the question is asked affects the answer (Deng, 2016; Chaskin, 1997). In short, asking participants to describe the boundaries of a neighborhood is a deeper question than originally intended.

The Pearl District exercise represented an attempt to see how survey participants dealt with a greater degree of uncertainty. The crowd-sourced estimate of the neighborhood's size and shape conforms closely to the official outline, though the data for this exercise are much more variable than for the other exercises. Median centroid distance across all attempts was 254 meters, while the median centroid distance for all other exercises combined was just 22 meters. Individual opinions about the boundaries of a neighborhood can differ from official or commercial definitions, and the survey instructions failed to specify which was wanted.

The next section reviews participant performance on the mapping activities, then attempts to place them in context of the survey responses.

Mapping Performance Metrics

The primary goal was to measure the accuracy with which participants would mark known points of interest or fill in known areas.

Point Placement

Table 12 contains a summary of the submitted points. For point exercises, the relative distance is as defined in ; relative distance values under 1.0 are considered valid. Relative radius is defined in .

Table 12. Valid point summary for point exercises.

			Relative Distance		Relative Radius	
Activity (valid point)	**n**	**%**	**Median**	**IQR (Q1..Q3)**	**Median**	**IQR (Q1..Q3)**
01 Powell's Books (yes)	49	86	0.20	0.17 (0.15..0.32)	0.99	0.08 (0.95..1.03)
01 Powell's Books (no)	8	14	3.12	7.8 (2.22..10.02)	2.77	3.9 (1.38..5.28)
02 Millar Library (yes)	52	93	0.24	0.24 (0.13..0.37)	0.98	0.09 (0.93..1.02)
02 Millar Library (no)	4	7	14.49	10.52 (9.25..19.76)	2.87	2.34 (1.86..4.2)
03 Council Crest Park (yes)	51	93	0.09	0.07 (0.07..0.14)	1.02	0.14 (0.98..1.12)
03 Council Crest Park (no)	4	7	10.12	4.97 (8.33..13.29)	2.21	2.27 (1.31..3.58)
Total (yes)	152	90	0.16	0.21 (0.09..0.29)	0.99	0.08 (0.96..1.05)
Total (no)	16	10	8.32	14.54 (2.98..17.51)	2.77	3.87 (1.41..5.28)

Over three sequential point exercises shown in Figure 2, the total number of completed activities drops slightly, while accuracy improves. Accepted points are grouped close to the target centroid, with at least 75% of them within one quarter of the allowed distance. Failing points tend to be significantly outside the target, and median relative distance for failing points increases over the course of the three exercises. Point placement on Council Crest Park is less precise in absolute terms, but this is not surprising since the area of the park is over five times larger than the two city-block locations.

Powell's Books is the first exercise after a brief tutorial, and 86% of participants correctly locate it on a map (Figure 3).

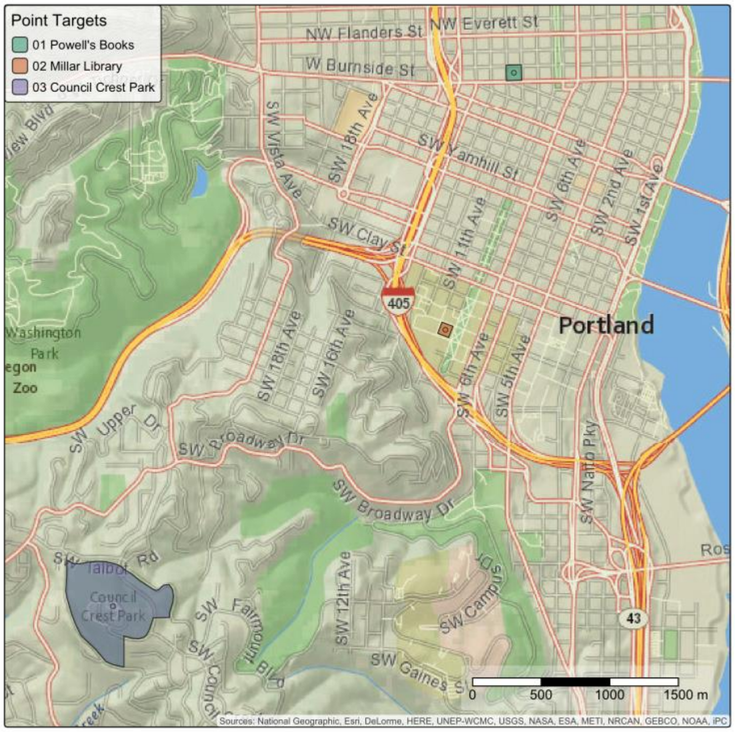

Figure 2. Point exercise targets.

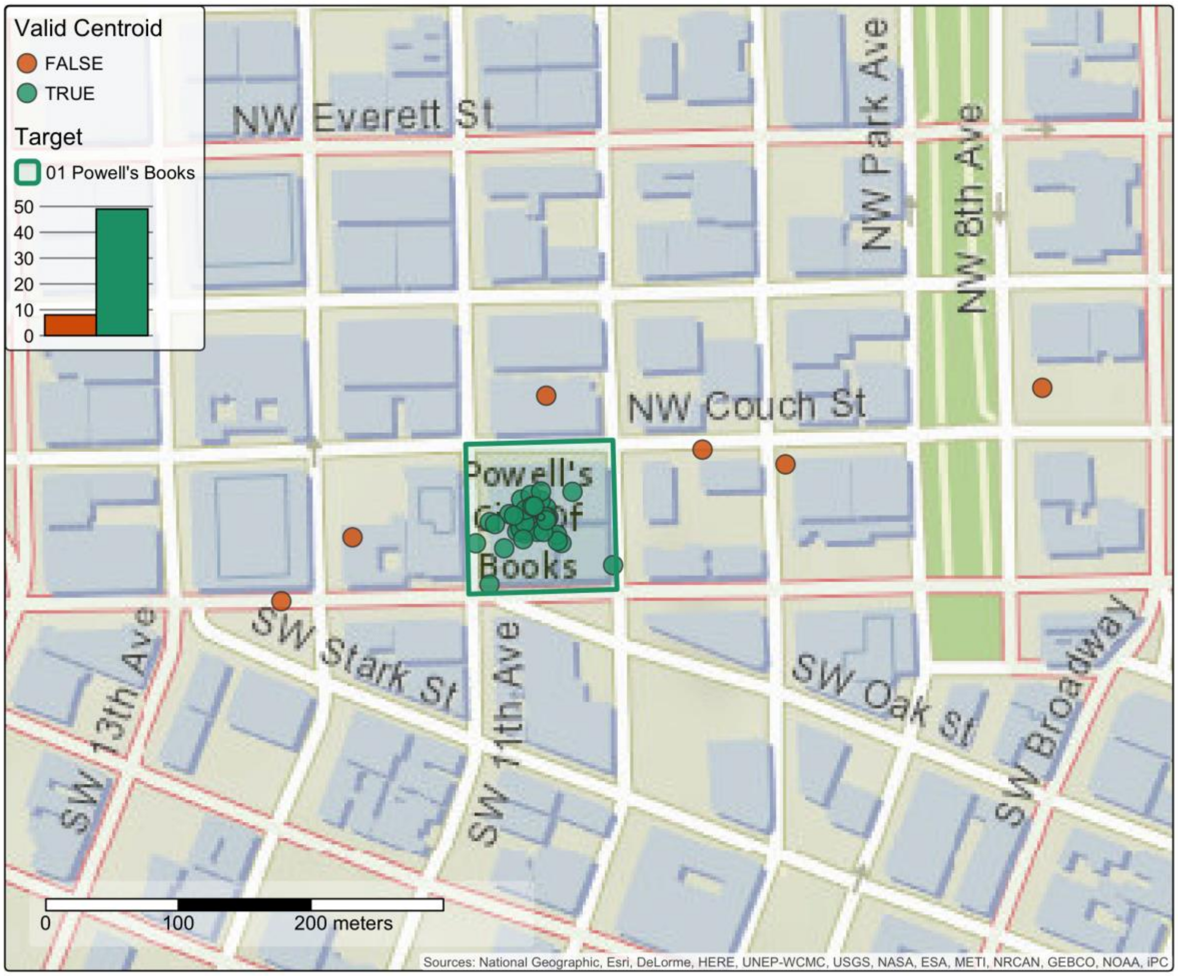

Figure 3. Powell's Books point performance.

Accepted points (based on distance) also tend to follow the survey instructions, with the median relative radius no more than 2% different from specified and a small interquartile range. Some amount of variation can be explained by the lack of precision in the map zoom controls. For failing points, the median relative radius is 1.9 to 2.77 times that specified. The circle radius in this software configuration is locked to zoom level (higher zoom leads to smaller circle radius). There is some evidence that zoom level is related to success with these exercises. What this might mean for future work is discussed in detail below.

The Portland State University Millar Library may be less familiar to Portlanders than Powell's Books, but survey participants were able to locate it on a map (Figure 4). Less than 10% of submissions missed the mark.

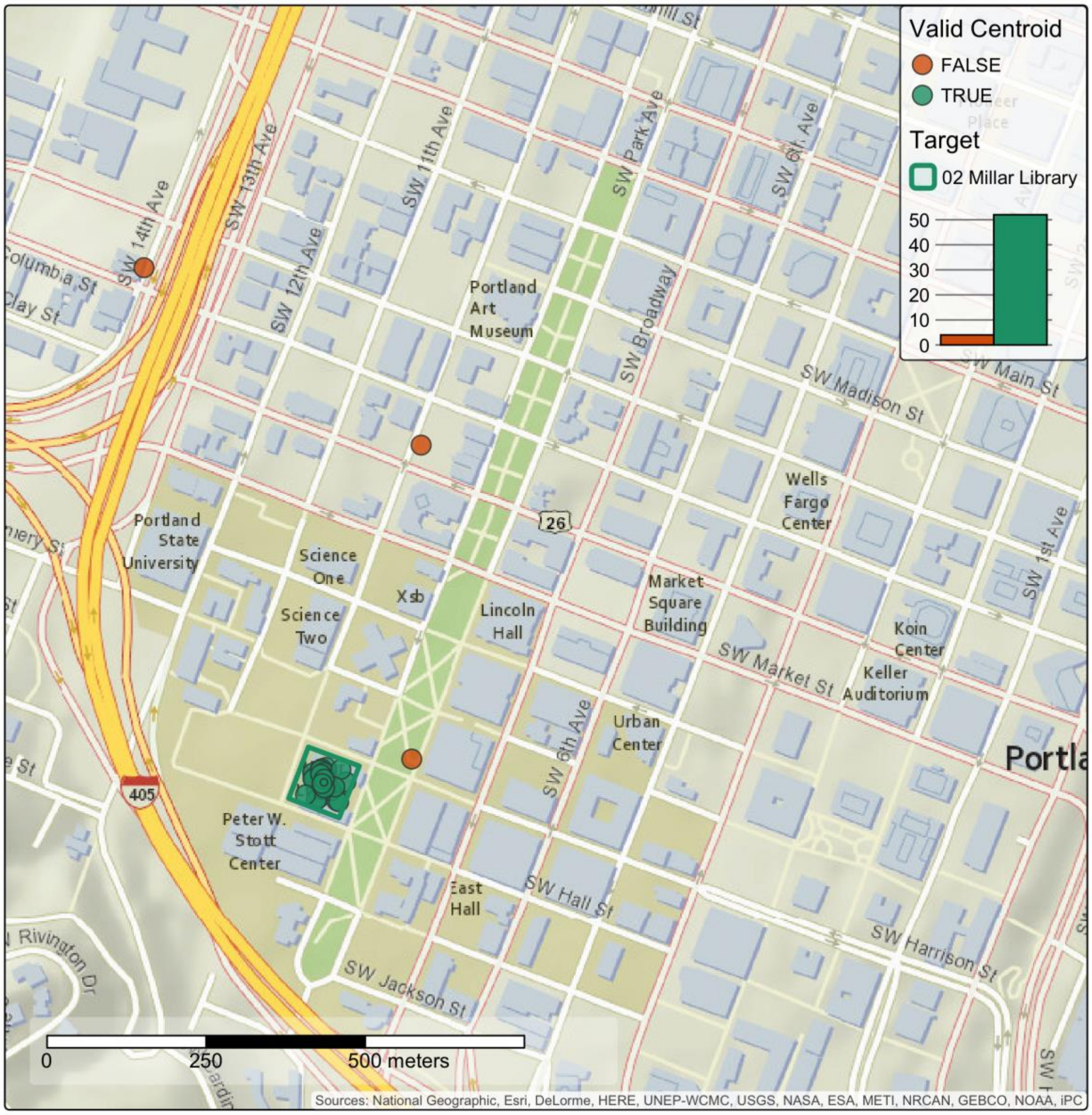

Figure 4. Millar Library point performance.

Survey participants had about the same level of success locating Council Crest Park (Figure 5), a popular destination according to average rating and review count in Google Maps listings. Submissions that failed landed a median of 14 times the allowed distance from the target.

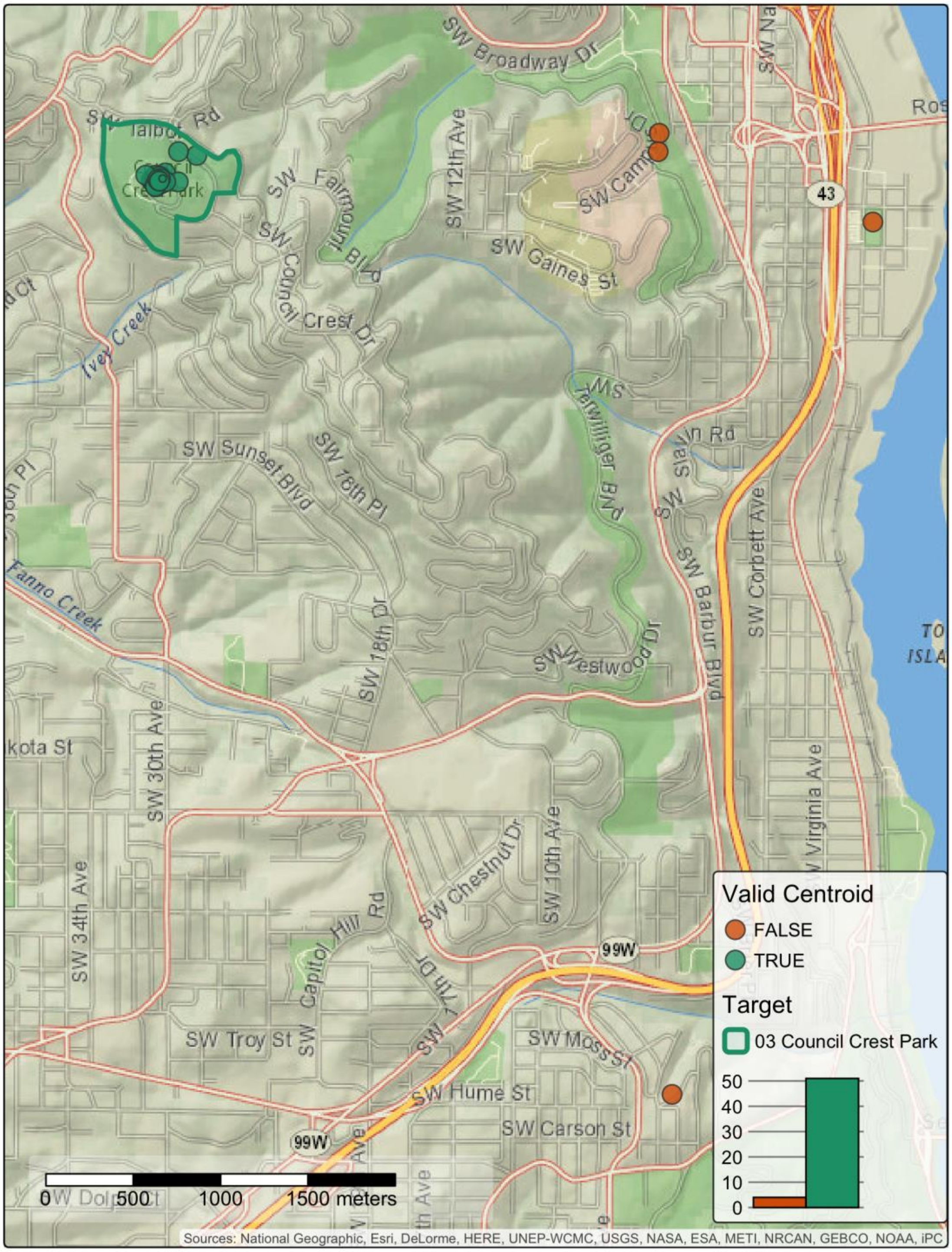

Figure 5. Council Crest Park point performance.

Area Filling

For area exercises, valid centroids are determined by distance between the submitted centroid and the target centroid. The maximum allowed centroid distance is the target polygon's equal-area circle, as calculated by (see Figure 6 for a visualization of valid centroid zones). Table 13 gives a summary of the relative distance and area statistics for the area filling exercises.

Table 13. Valid centroids summary for area exercises.

			Relative Distance		Relative Area	
Activity (valid centroid)	**n**	**%**	**Median**	**IQR (Q1..Q3)**	**Median**	**IQR (Q1..Q3)**
04 Mount Tabor Park (yes)	54	100	0.06	0.07 (0.04..0.11)	1.49	0.31 (1.37..1.69)
05 Willamette River (yes)	52	96	0.16	0.16 (0.07..0.22)	1.5	0.71 (1.34..2.05)
05 Willamette River (no)	2	4	1.84	0.17 (1.75..1.92)	2.94	2.28 (1.8..4.08)
06 Pearl District (yes)	45	88	0.43	0.38 (0.2..0.57)	1.04	1 (0.69..1.68)
06 Pearl District (no)	6	12	1.26	0.15 (1.16..1.31)	1.12	1.02 (0.59..1.62)
Total (yes)	151	95	0.13	0.25 (0.06..0.31)	1.45	0.64 (1.14..1.78)
Total (no)	8	5	1.31	0.35 (1.2..1.56)	1.12	1.08 (0.63..1.71)

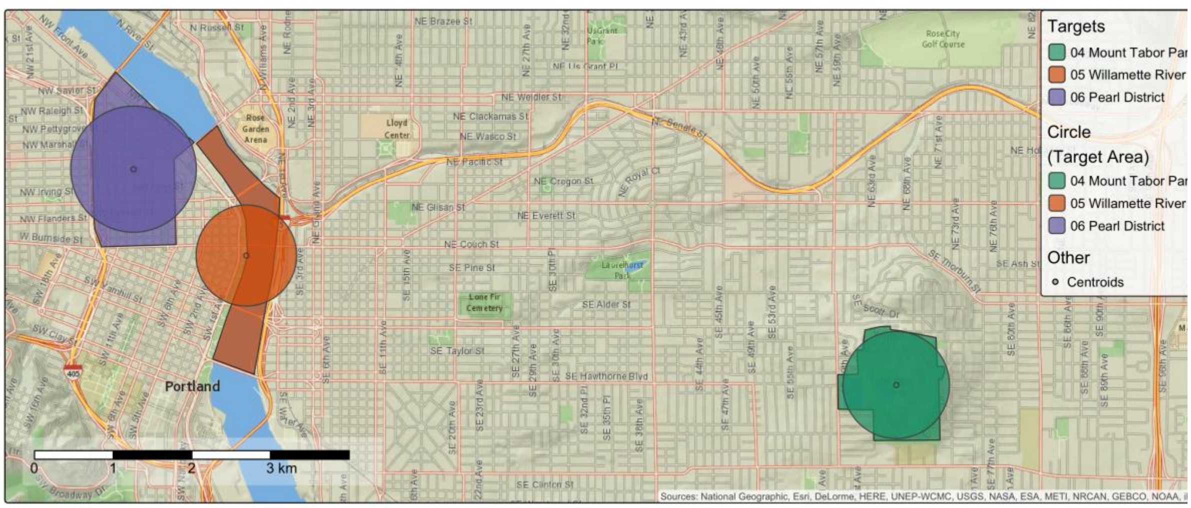

Figure 6. Study target areas, with valid centroid circles.

The basemap for the Mt. Tabor exercise (Figure 7) shows a high contrast boundary, and the label is legible at minimum zoom level, making the target easy to locate even at minimum zoom.

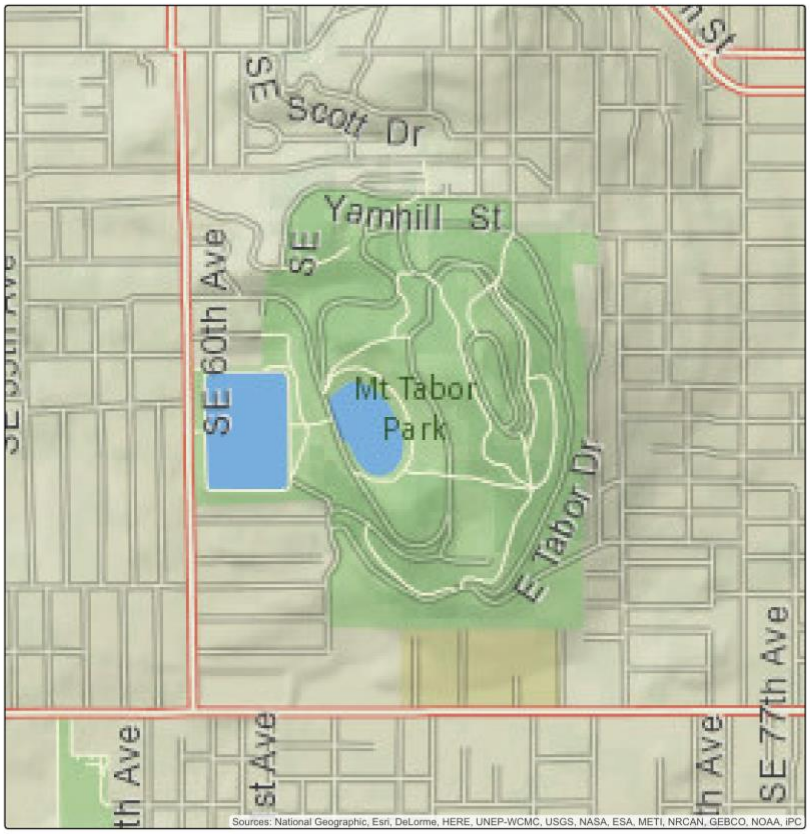

Figure 7. Basemap for Mt. Tabor exercise.

All 54 submissions on the Mt. Tabor area exercise passed the relative distance test. Another way to look at quality is to select the submissions that fall in the closest 90% of centroid distance, polygon area, and greatest shape similarity – an intentionally liberal filter. There are 44 submissions (81%) that meet all three criteria (Figure 8). Applying this filter results in a rough approximation of the target area for Mt. Tabor.

These 44 shapes approximate the target shape but still tend to overfill it. To rank areas strictly by deviation from the target, the "area score" is the absolute value of the submitted area subtracted from the target area, divided by the target area, minus 1 (). For a submission that perfectly fills the target, this value is zero.

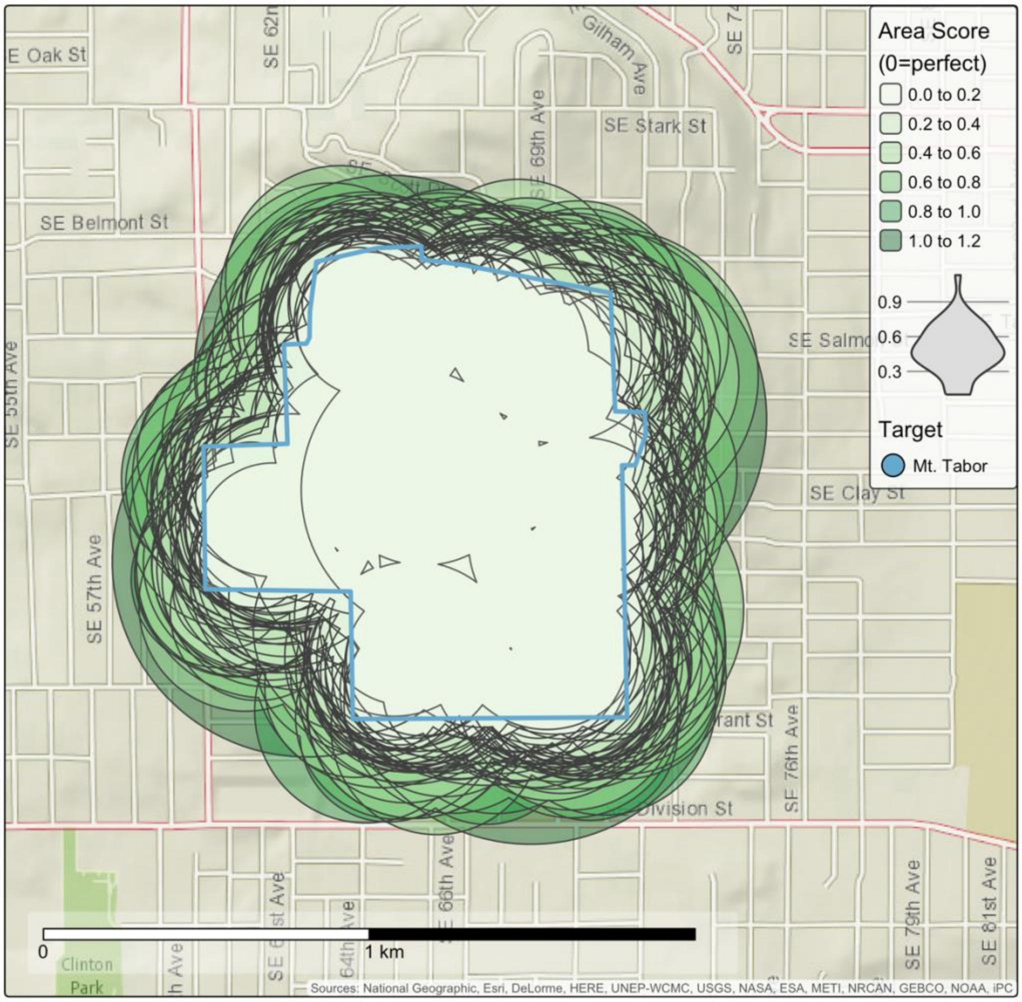

Figure 8. Mt. Tabor, all submissions falling in the first 90% of distance, area, and shape.

Participants were shown an example of filling an area, instructed to use as much or as little detail as they liked, and told "It's OK to not completely fill the area. Just cover most of it." (see Appendix A - Survey Questions). Still, most of these top-rated responses overfilled the area by about 50%. The violin plot in the legend panel in Figure 8 shows the shape of this distribution. Only four participants entered a polygon smaller than the target. While the responses with centroids furthest from the target were well inside the acceptance criteria, the quality of these submissions' shape and area is also lower (Figure 9).

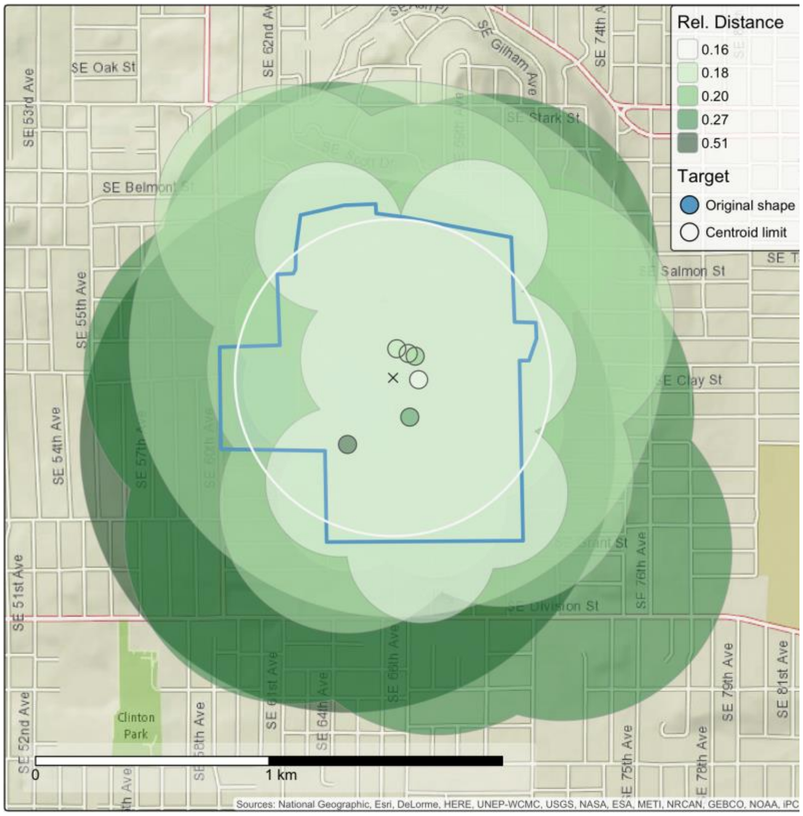

Figure 9. Mt. Tabor submissions, greatest centroid distance.

For the Willamette River exercise, the participant was directed to "[c]enter the map on the Willamette River, between the Broadway and Hawthorne Bridges (roughly between NW Lovejoy and SE Hawthorne Streets). Fill this portion of the river with circles." The basemap is shown in Figure 10.

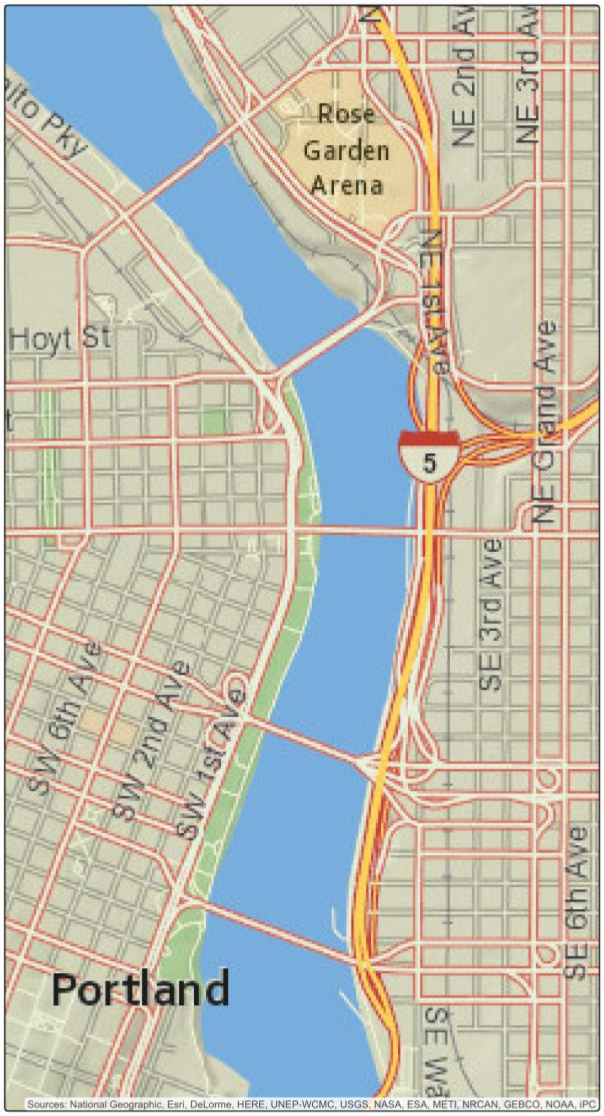

Figure 10. Basemap for Willamette River exercise.

The Willamette River exercise was more challenging than Mt. Tabor. The shape to be filled was less regular, and two bridges had to be identified by zooming in. Two of 56 submissions failed the distance test for the Willamette River exercise. The top 90% of responses are shown in Figure 11. As with the Mt. Tabor exercise, the 90% response map provides a recognizable estimate of the Willamette River exercise target, and submitted polygons overfill by about 50%. Most of the 10% most distant centroids fall within the cutoff circle, but there also appears to be a pattern of poor conformance to the target areas and shapes (Figure 12). Some participants filled space between different bridges or failed to zoom in enough to draw a precise shape.

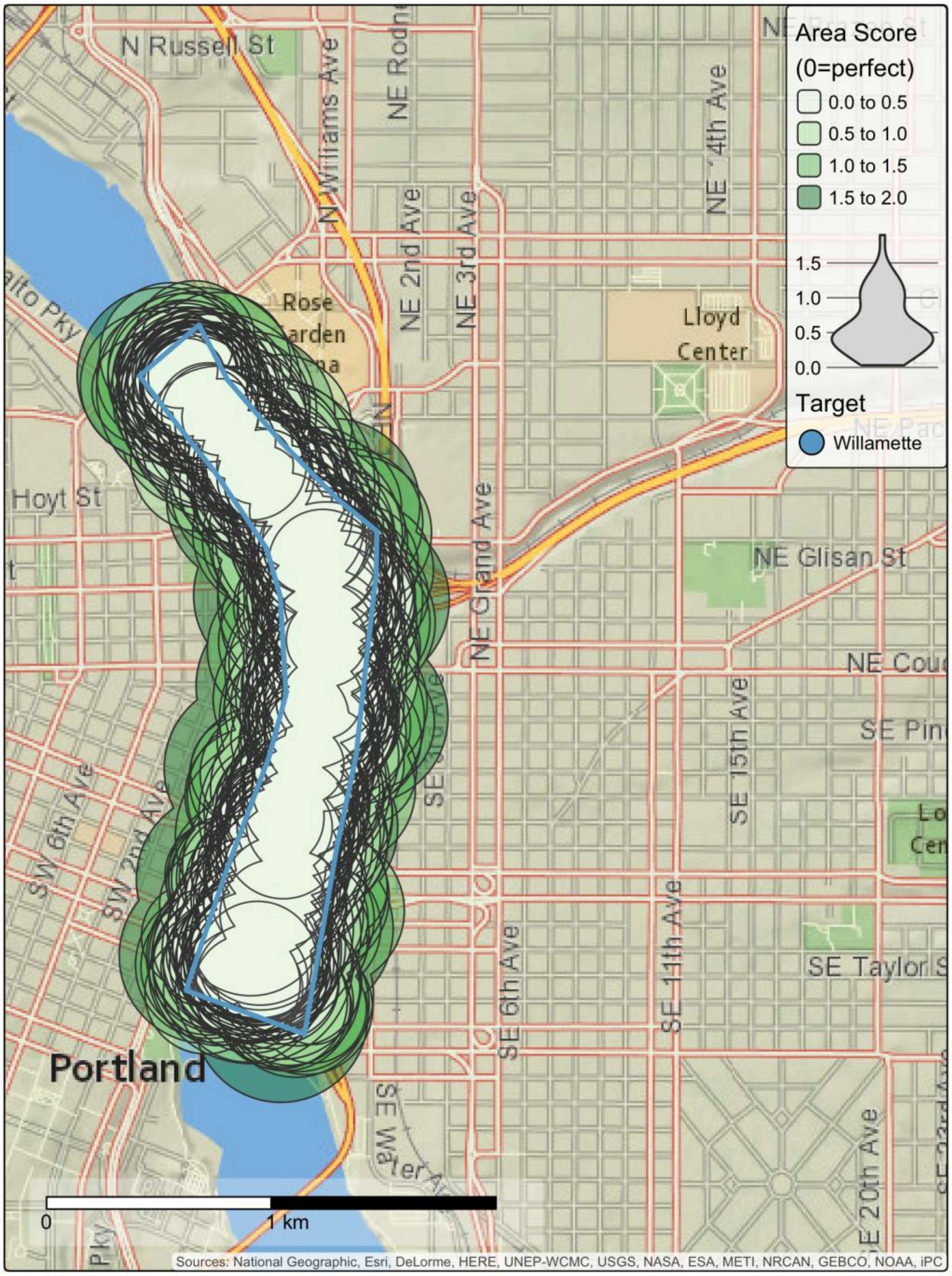

Figure 11. Willamette River, all submissions falling in the first 90% of distance, area, and shape.

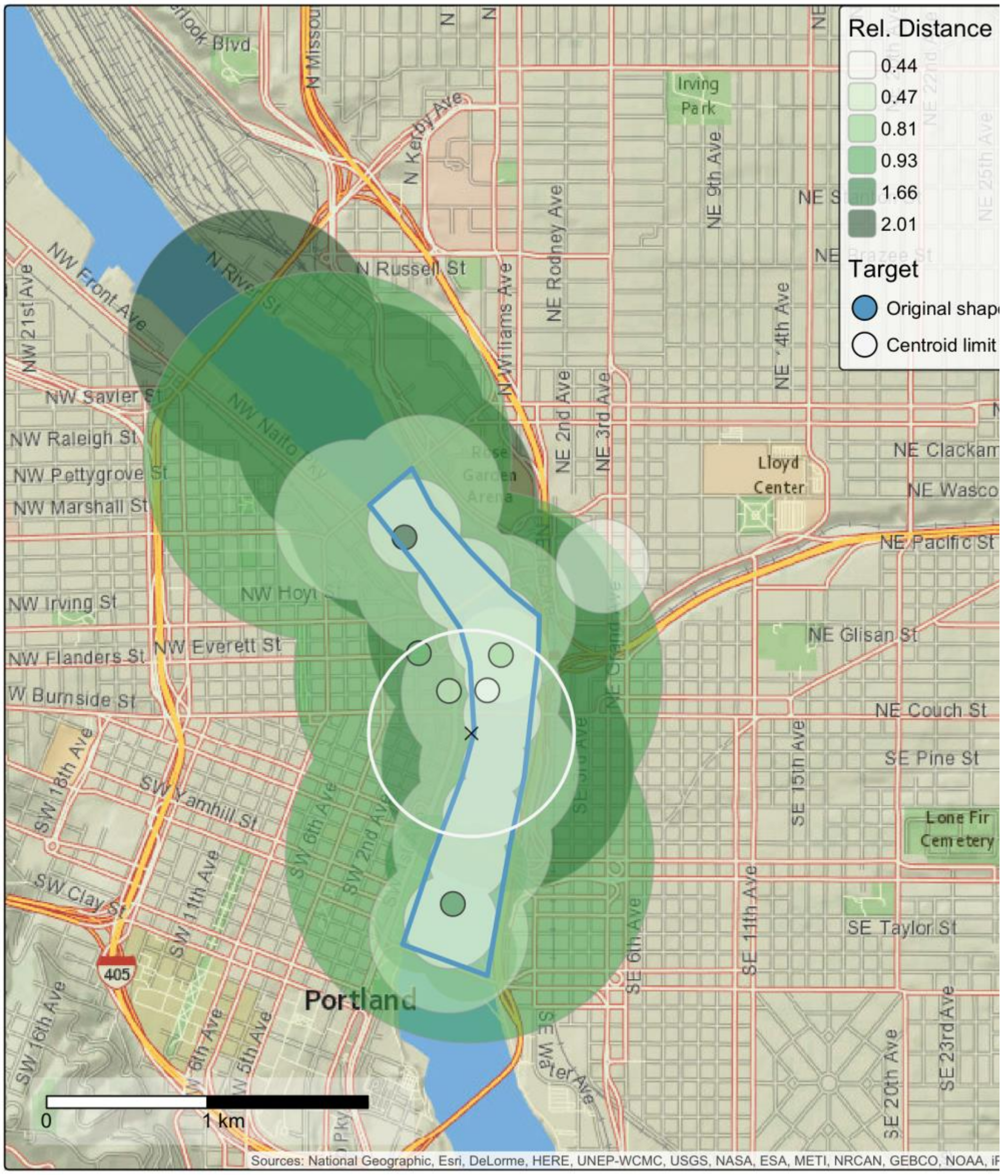

Figure 12. Willamette River, top 10% (worst) centroid distance.

The Pearl District exercise results is qualitatively different. The official neighborhood map according to the City of Portland defines the neighborhood boundaries as Burnside Street to the south, NW 14th Avenue to the west, the Willamette River to the north, NW Broadway to the east. No neighborhood boundary appears on the basemap presented for the exercise (Figure 13).

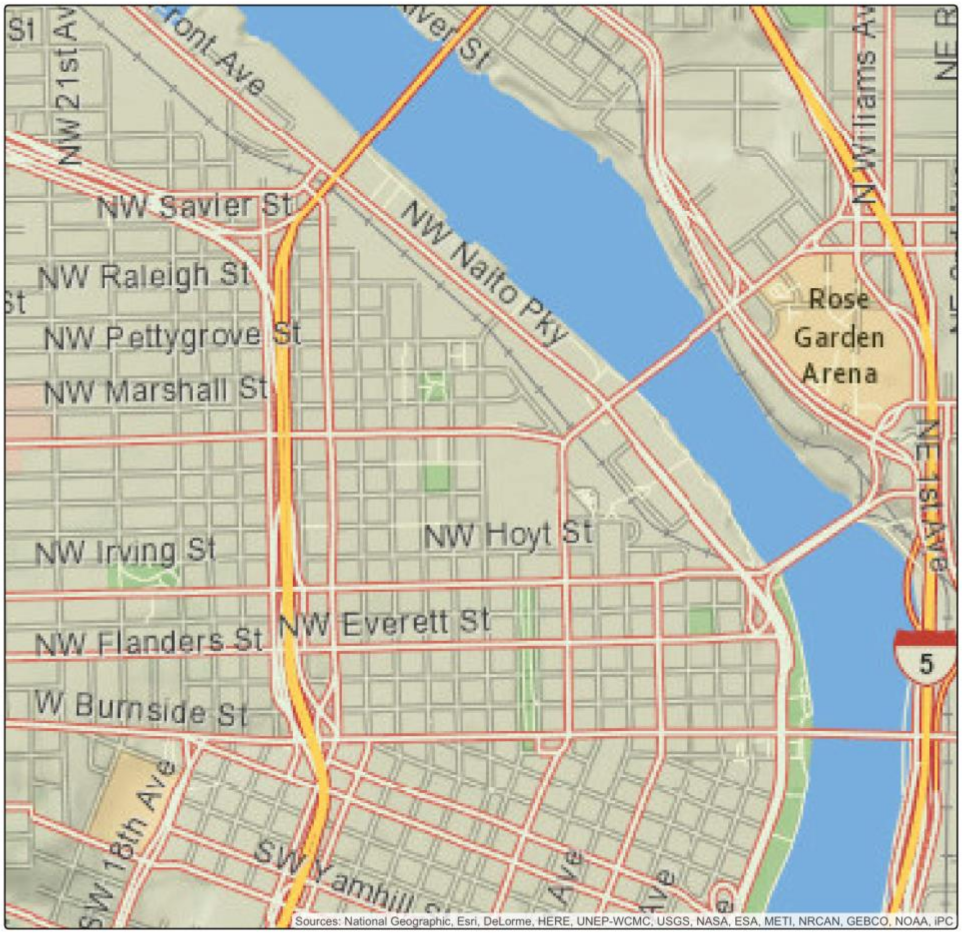

Figure 13. Basemap for Pearl District exercise.

The instructions stated, "center the map on the Pearl District, in inner NW Portland. Fill in the neighborhood with circles." Several respondents commented on this exercise, saying it would be easier with a neighborhood boundary layer. When designing the survey, the author did not appreciate the sheer number of ways that neighborhood boundaries can be conceived (Deng, 2016). Failure to place a "valid" polygon for this exercise could indicate disagreement with the official city boundary, ignorance, or failure to locate identifying features on the basemap, making interpretation even more difficult.

While six of 51 submissions failed the distance test for the Pearl District exercise, the most distant centroid was only 52% past the cutoff. These submissions make a poor representation of the target polygon (Figure 14). In contrast, there is notable clustering around the target among the submissions falling in the top 90% group (Figure 15). Despite the clustering, unlike the results for the Mt. Tabor and Willamette River exercises, there is no clear consensus in the top 90% submissions for the Pearl District. The lack of clarity could be explained by participants using (or not using) outside references to inform their survey responses. Without a displayed boundary, participants also appear to overfill less, as shown in the violin chart.

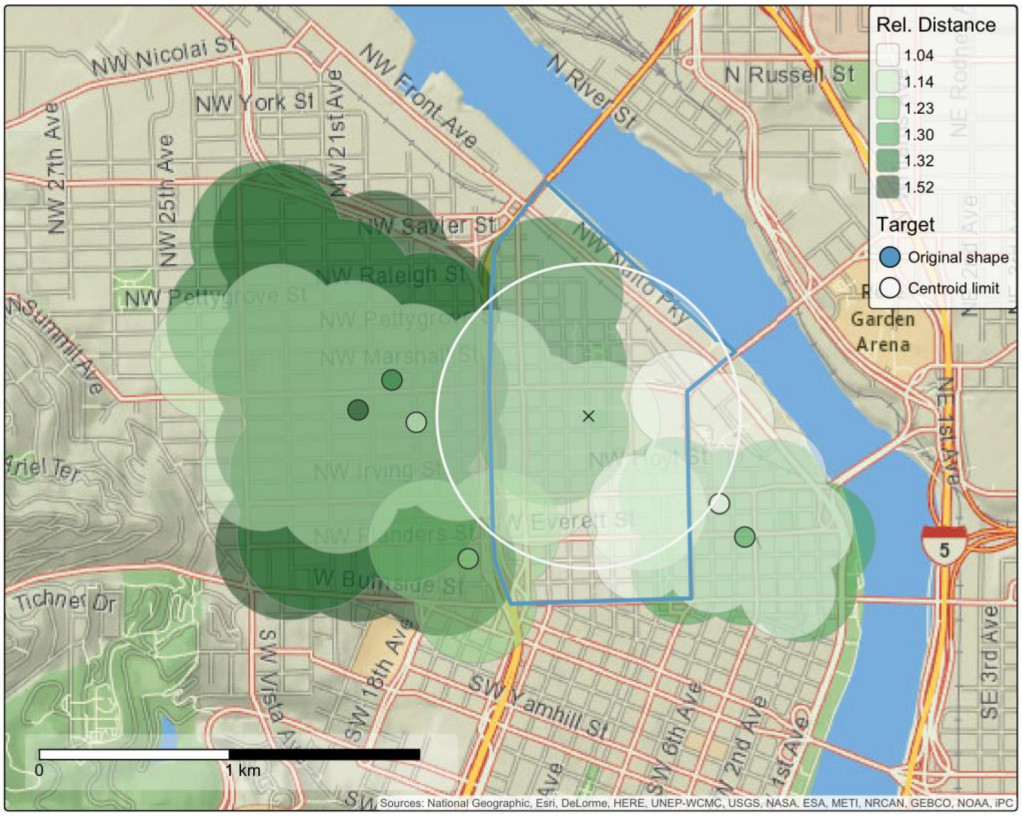

Figure 14. Pearl District, top 10% (worst) centroid distance.

The point here is not that there are some responses that poorly represent the target area, but that the responses come close enough to easily differentiate them from surrounding areas of the same type. One can imagine a future set of exercises that called upon participants to locate a set of adjacent neighborhoods. Based on the evidence shown in Figure 15, the consensus boundaries of a target area take shape.

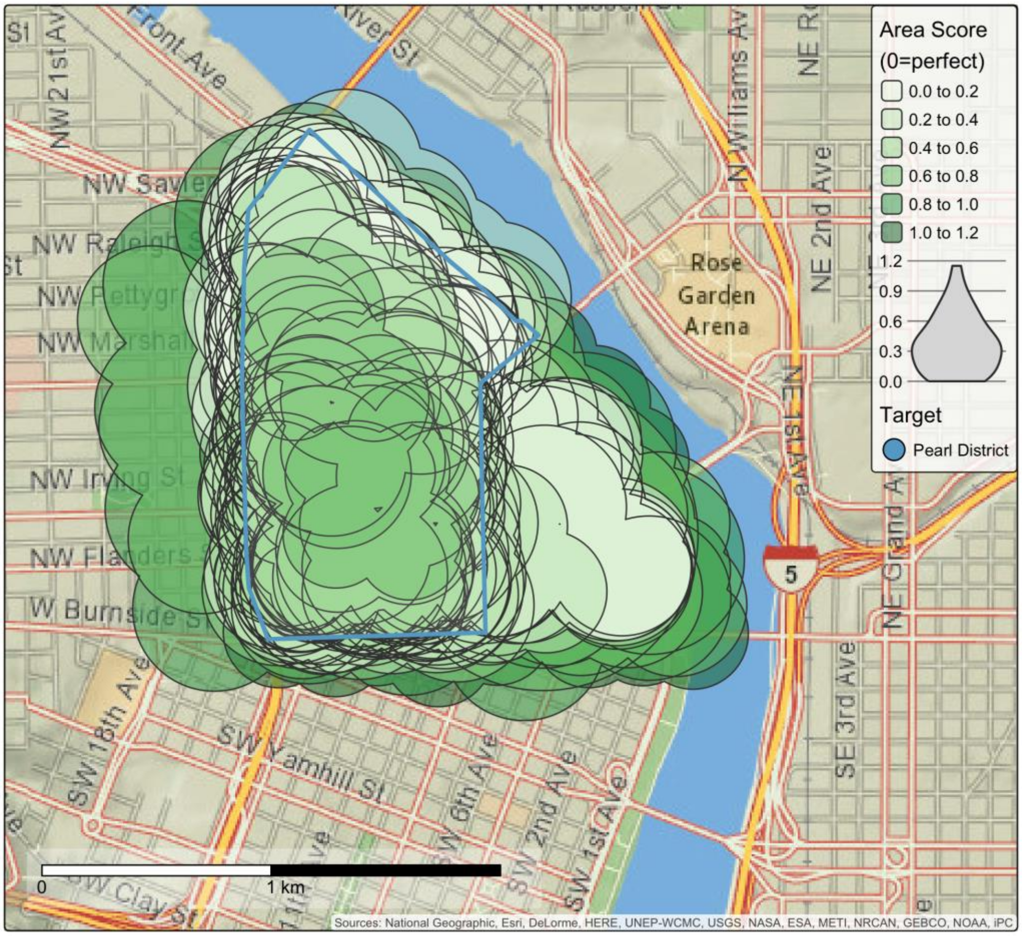

Figure 15. Pearl District, all submissions falling in the first 90% of distance, area, and shape.

Selecting the submissions that fall in the middle of the range for location, area, and shape yields a better estimate of the target boundary (Figure 16). To draw this map, an average centroid was determined by taking the centroid of the convex hull including all the submitted centroids. The distance from this center point to each polygon's centroid was recorded. Then the submitted polygons for the Pearl District exercise were filtered:

1. Distance to average center less than the median.
2. Area in the interquartile range (median±25%)
3. Shape index in the interquartile range

Most of the submissions meeting the above filter criteria closely matched the target area despite the missing basemap reference. This suggests participants likely consulted external references (like Google Maps) during the survey. Since participants weren't instructed to avoid using outside references, this isn't necessarily wrong—it just makes it difficult to determine whether accuracy came from participants' prior knowledge or from consulting other sources during the survey.

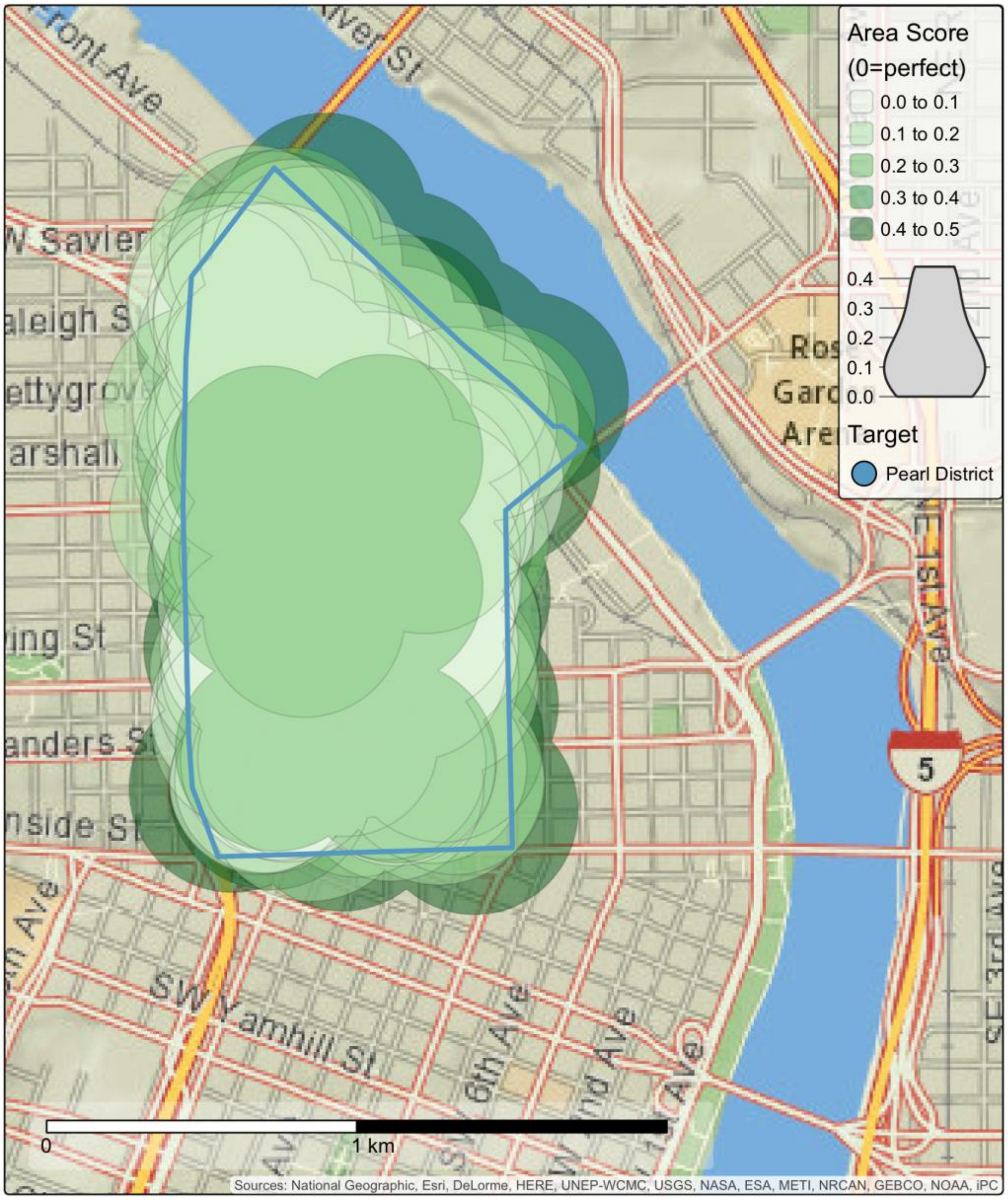

Figure 16. Submitted polygons in the IQR for area and shape, closest to mean center.

Potential Sources of Inaccuracy

Probably the most important goal of this work was to develop an understanding of the variables that might lead to lower accuracy point placement. Related, but harder to quantify, is a sense of how easy the tool is to use and what might be done to improve any shortcomings.

One source of errors may be individuals who are rushed, lack basic computer skills, or are inattentive to details. Of the 24 recorded errors (of 327 attempts on all exercises), most of them are repeats within a survey response – 12 of them occur in 5 survey responses (Table 14). A larger dataset or different experimental design would be needed to show if there is a real tendency for some individuals to be prone to error.

Table 14. Errors recorded, grouped by survey participant.

Date	Response ID	# Failed	User Device
3/24/24	PBF05	4	Phone
3/20/24	13MPS	2	Laptop
4/26/24	ASSR1	2	Phone
6/4/24	O93T0	2	Laptop
6/5/24	ZGP9C	2	Phone
3/24/24	IPU7C	1	Laptop
4/4/24	YWQBB	1	Laptop
4/4/24	AC3G6	1	Laptop
4/4/24	VCYDP	1	Laptop
4/5/24	YMRBG	1	Desktop
6/3/24	CDTGD	1	Laptop
6/4/24	P1SLS	1	Phone
6/4/24	HCKTV	1	Phone
6/4/24	PIPSQ	1	Laptop
6/4/24	DJVDH	1	Phone
6/4/24	CYSMD	1	Tablet
6/4/24	PBS89	1	Laptop

Several potential relationships between survey metadata and point placement accuracy were examined. These include age, device used for input, stated knowledge or skills, technical issues or expressed difficulty, and survey duration (or perceived duration). None of these factors had a statistically significant effect on mapping exercise success ($p \geq 0.05$, see Appendix D – Regression for output of these regression runs).

Among these, the device used for data entry stood out, with a p-value of 0.056 for phones. Data entered on phones had an error rate of 12.5%, compared to 1.9% of desktop computer data and 7.2% of laptop-entered data. The error rate for tablet users was 1 in 12 exercise submissions, with only two survey responses coming from tablet users.

Zoom Level vs. Centroid Distance

There is a statistically significant relationship between the zoom level and success with an exercise, as shown in Table 15. Zoom level is reported as the number of displayed map pixels that would represent one kilometer on the ground, px/km[6].

Table 15. Activity success vs. zoom level.

	Highest zoom level (pixels/km)		
Activity (valid centroid)	**Min**	**Median**	**Max**
01 Powell's Books (yes)	319	474	598
01 Powell's Books (no)	50	170	465
02 Millar Library (yes)	325	474	597
02 Millar Library (no)	55	249	492
03 Council Crest Park (yes)	81	457	597
03 Council Crest Park (no)	33	243	460
04 Mount Tabor Park (yes)	33	312	597
04 Mount Tabor Park (no)	N/A	N/A	N/A
05 Willamette River (yes)	55	312	597
05 Willamette River (no)	129	237	345
06 Pearl District (yes)	55	312	502
06 Pearl District (no)	255	315	399

The median zoom level for successful submissions is higher in all exercises that had any failures, a relationship visualized in Figure 17. In this figure, centroid distances are normalized by the scoring radius and zoom level decreases from left to right. The highest number of valid points (below the horizontal line) is at the far left.

There appears to be some association between high zoom level and centroid accuracy, but there is not a strict causal relationship. High zoom level is not required for accuracy: observe the successful submissions at low zoom levels in the Mount Tabor Park exercise. And high zoom levels are not a guarantee of success, as seen in the first three point-exercises.

---

[6] OpenLayers reports resolution as meters per pixel. To make this easier to understand conceptually, this has been scaled as 1/resolution * 1000 to yield the number of pixels representing 1000 meters on the ground (px/km). The amount of physical area covered by those pixels may differ from screen to screen.

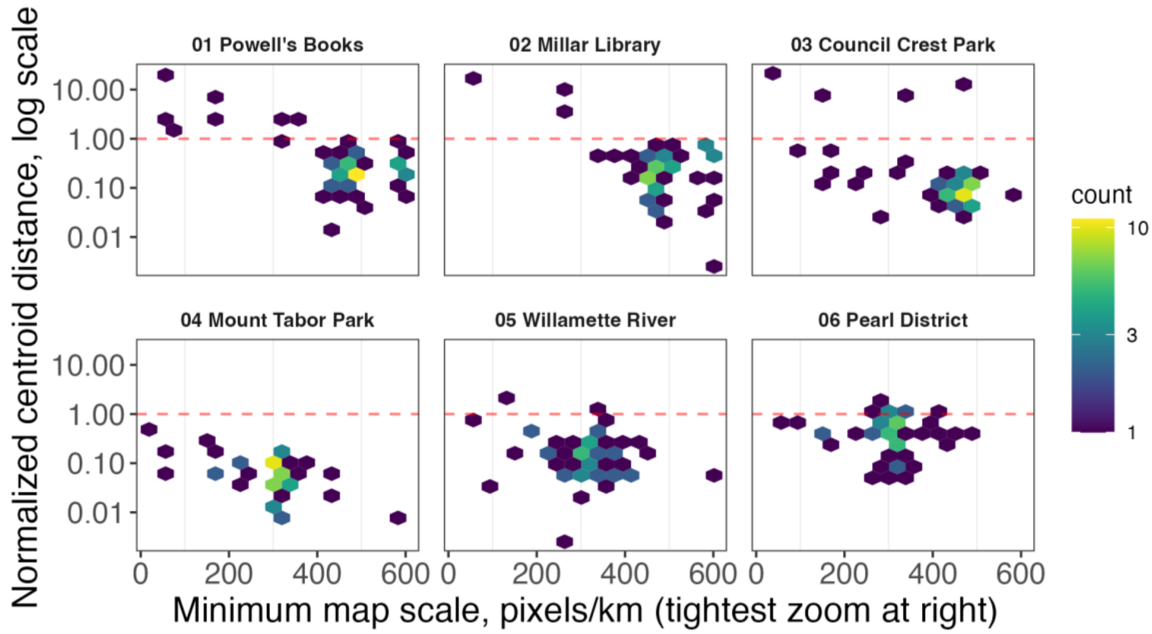

Figure 17. Submission relative centroid distance, vs. map scale in pixels/km.

The relationship between displayed map scale and exercise success may also be visualized as a boxplot, as in Figure 18. Errors are associated with smaller map scales in most cases. The exception is the Pearl District exercise, where there was no reference on the basemap and thus more uncertainty about its placement and boundaries.

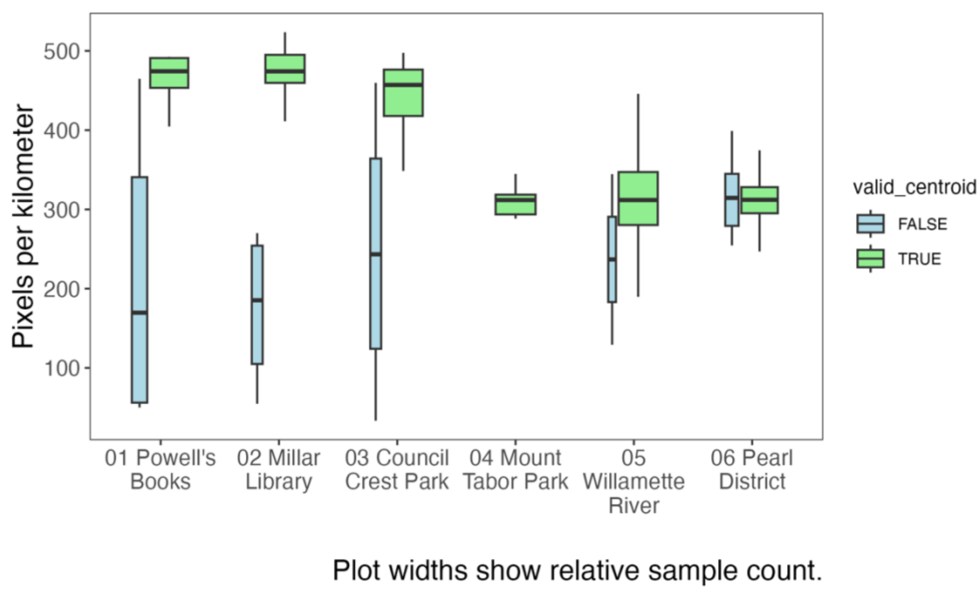

Figure 18. Box plot of map display resolution, grouped by centroid validity.

The map is initially displayed at a zoom level of approximately 18 meters per pixel; about 55 pixels represent one kilometer. Since each polygon submitted for area exercises may be an aggregate union of many rows, there is a range of resolution values as the user zooms in and out to locate the target. Most of the users who attempted the exercise zoomed in. Those who also created a circle close to the specified radius necessarily zoomed into a tighter band, since the radius was tied to zoom level.

There are several ways to look at this relationship with regression models. One could perform logistic regression to model valid centroid status, or linear regression to predict centroid distance. The centroid distance was also normalized by scoring radius to obtain a relative distance – used to compare results between exercises (see Figure 17). Stepping through these options may be instructive.

Logit regression on the relationship between map scale and "valid centroid" shows a significant effect (p=1.1e-7). Residual deviance is less than null deviance, providing confidence that this model explains at least some of the variation. Because this is a logit model, R-squared is not available to provide an easy (if potentially misleading) evaluation of model quality. For details, see > # *GLM for max px/km (*valid_centroid), in Appendix D – Regression.

The coefficient in this model predicts that for every 100-unit increase in resolution, the probability of success increases by about 2.5 times (e^(0.0093 * 100)). To more comprehensively evaluate the model, a simulation using the "Residual Diagnostics for HierARchical Models" (DHARMa) package (Hartig, 2024) was run. The plots and metrics generated by DHARMa help assess this model's suitability (Figure 19, Figure 20). Residuals closely following the diagonal line in the Quantile-Quantile (QQ) plot indicate a good fit. This logit model passes all the diagnostic tests run by DHARMa.

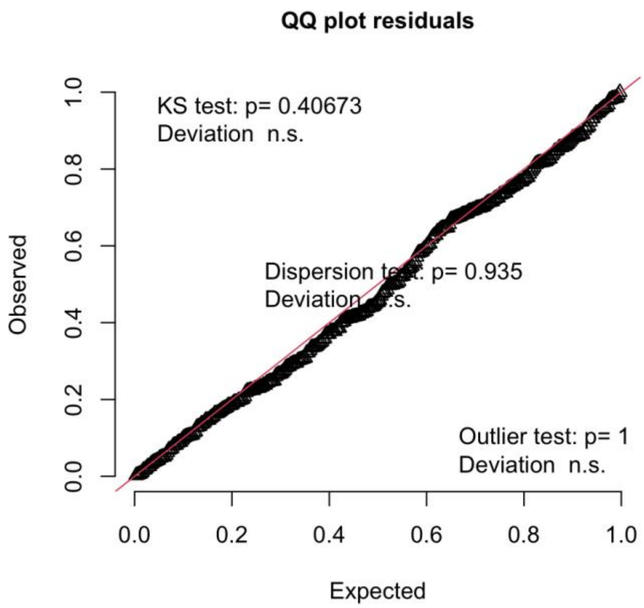

Figure 19. QQ plot, simulated vs. observed values for map scale to valid centroids model.

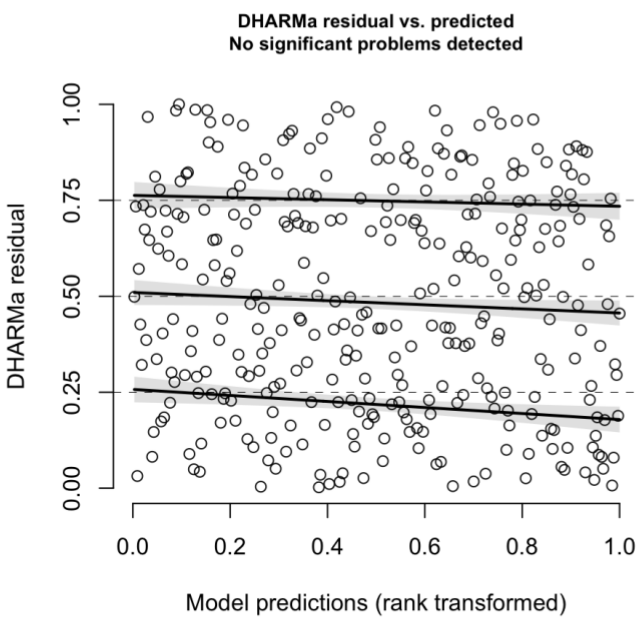

Figure 20. DHARMa residual plot for map scale to valid centroids model.

To quantify the size of the effect observed in the logit model, linear regression was applied to the relationship between map scale and centroid distance. The linear regression model predicts that for each one-unit increase in map resolution (pixels per kilometer), the centroid distance decreases by 0.865 meters ($p < 0.001$), though the model explains only about 8.8% of the variance. For details of the model results, see > # *LM for max px/km (*centroid distance), in Appendix D – Regression.

For a 100-pixel increase in resolution, the model predicts an 87-meter decrease in centroid distance. However, residuals are large. DHARMa simulation results illustrate nonlinearity in the model, as shown in Figure 21 and Figure 22 (note the S-curve in the QQ plot in Figure 21).

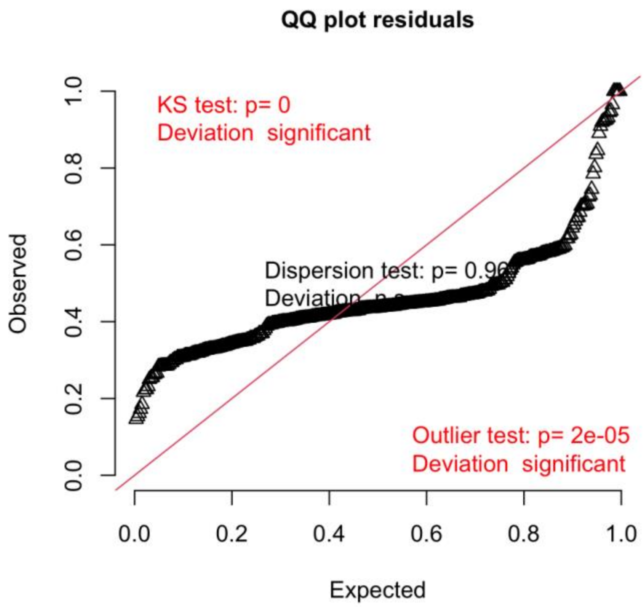

Figure 21. QQ plot, simulated vs. observed values for map scale to centroid distance model.

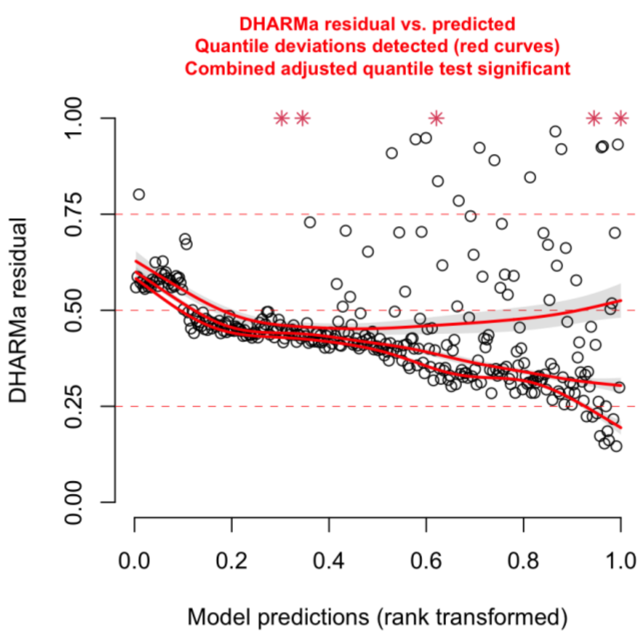

Figure 22. DHARMa residual plot for map scale to centroid distance model.

These diagnostics raise questions about direct predictions from the linear model. While there may be a relationship between map scale and point placement success, there are likely to be limits to this relationship at both ends of the scale range. At small scales, an error of a few pixels can result in a large deviation. Increasing resolution in this range will have a disproportionately large effect. At large scales, increasing resolution at the margin has little effect, as additional details are not present in the map and accuracy reaches its upper limit.

As an exercise, a Box-Cox transformation (Bobbitt, 2020; Box & Cox, 1964; Venables & Ripley, 2002) was applied to the centroid distance to obtain a better-fitting model. The Box-Cox algorithm finds a value for lambda (λ) of -0.0202 that best approximates the normal distribution when the dependent variable is transformed according to (8) (see > *# LM for max px/km (box-cox transform on* centroid distance), in Appendix D – Regression, below, for more detail on how the algorithm was applied).

$$y(\lambda) = \frac{\left(y^{\lambda} - 1\right)}{\lambda}, y \neq 0$$
$$y(\lambda) = \log(y), y = 0 \quad (8)$$

The transformed model has much better fit (adjusted R-squared 0.326, p<2e-16). A DHARMa simulation on this model also returns a linear QQ plot, with no significant deviation (Figure 23**,** Figure 24).

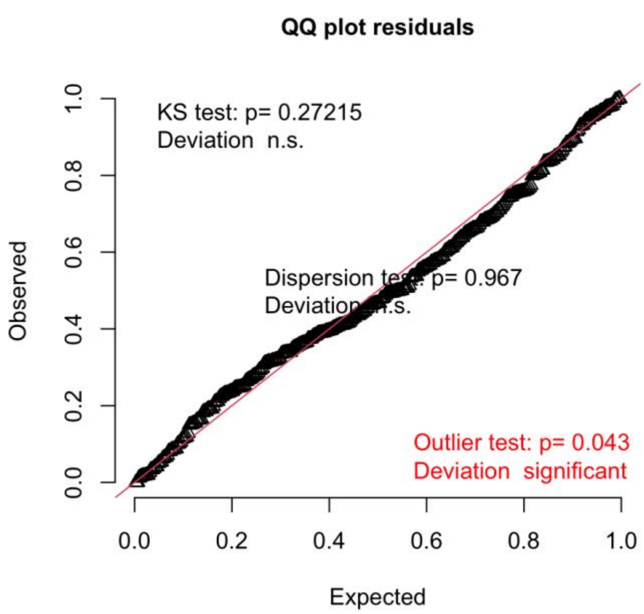

Figure 23. QQ plot, simulated vs. observed residuals for map scale to Box-Cox transformed centroid distance model.

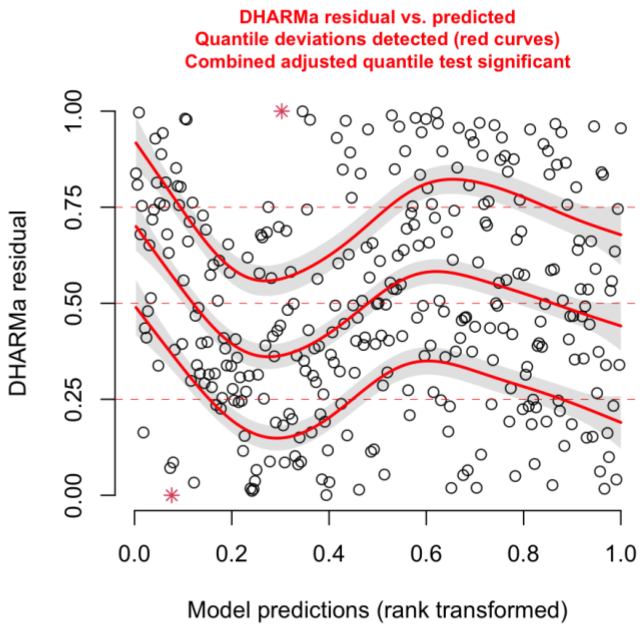

Figure 24. DHARMa residual plot for map scale to Box-Cox transformed centroid distance model.

It can be challenging to intuitively understand the predictions of a model based on a Box-Cox transform, because the results need to be transformed back to the original units. Figure 25 shows how centroid distance values predicted by the Box-Cox transform model line up with experimental values. The standard deviation error bars in Figure 25(b) show how variability limits the strength of the model.

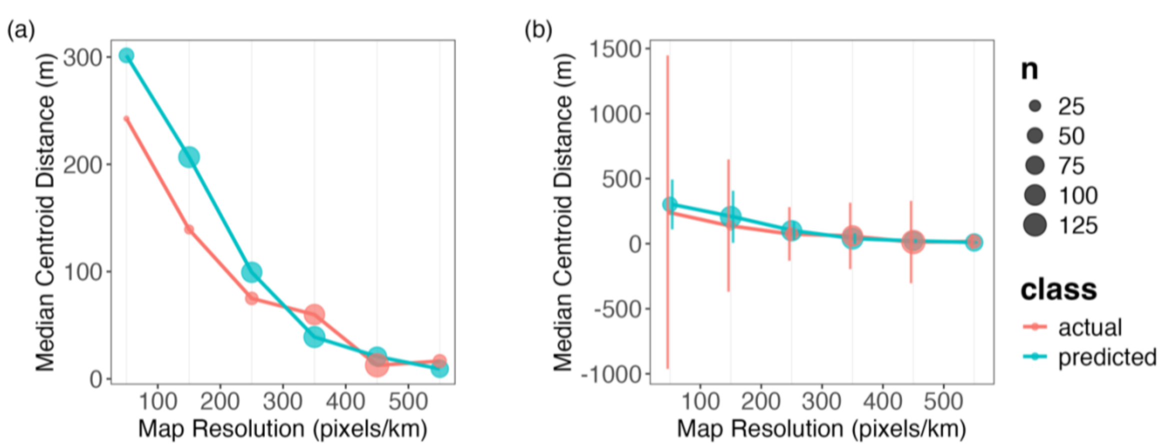

Figure 25. (a) Median centroid distance binned by map resolution, compared to Box-Cox model prediction. (b) with error bars representing standard deviation.

Zoom Level vs. Relative Area

The relative area metric also bears a relationship to map scale, as shown in Figure 26. Since submitted area tends to be about 1.5 times the target area, the "valid area" metric may not be particularly useful as a clear dividing line. For reference, proposed relative area cutoffs of 0.75 and 1.75 are shown as dotted red lines on the plot. These cutoff levels are used to sort submissions into valid and invalid areas.

Figure 27 shows the distribution of zoom level with respect to submissions with "valid" areas. Together, these plots show a trend towards higher accuracy with higher map detail.

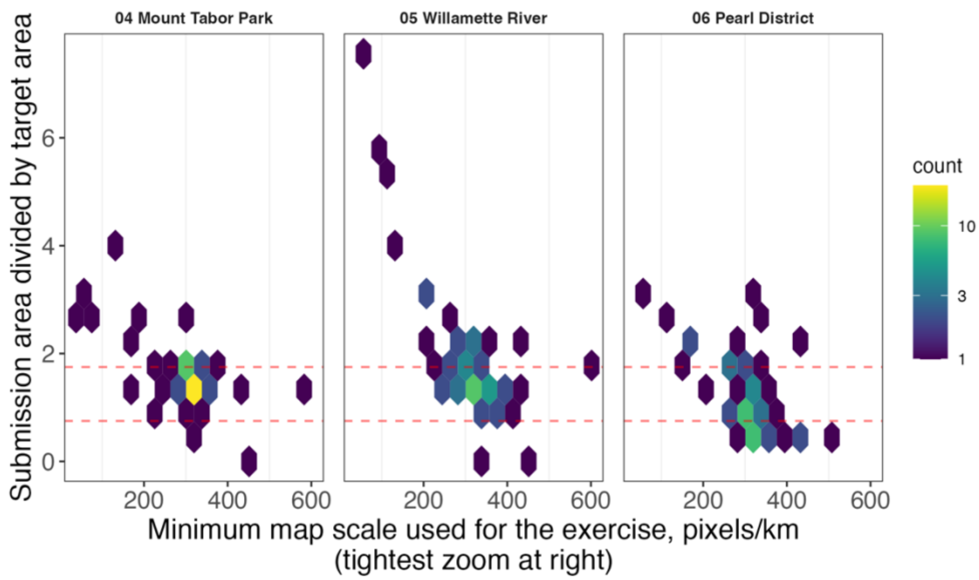

Figure 26. Submission relative area, by map display resolution.

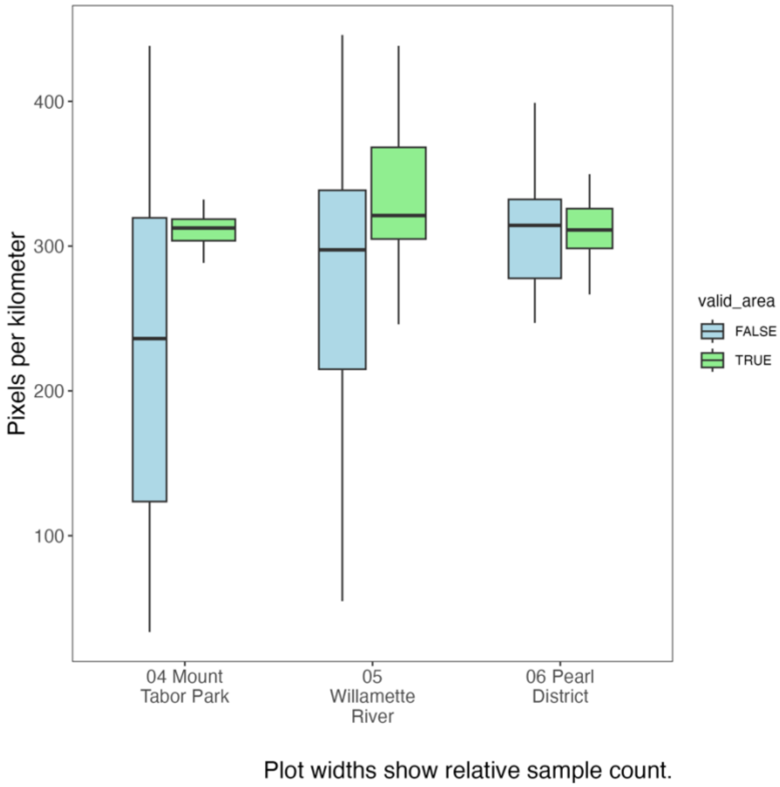

Figure 27. Box plot of map display resolution, grouped by area validity.

Zoom Level vs. Shape Defect

With the shape defect metric, over 80% of submissions have a value under 0.2 (Figure 28). This is the arbitrary cutoff used to collect "valid" shapes. As shown with the area metric, valid shapes are tightly clustered at higher resolutions (Figure 29).

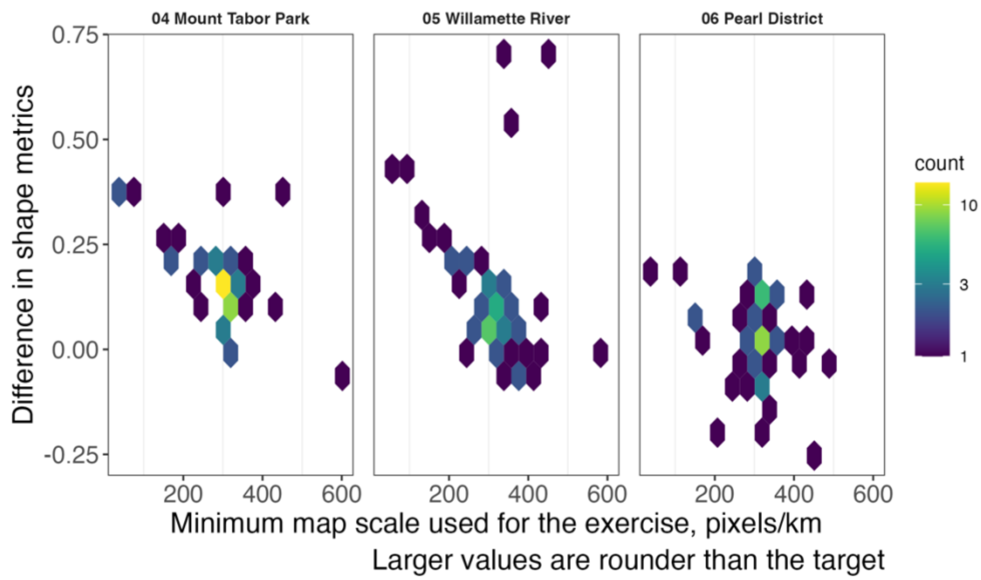

Figure 28. Distribution of shape defect, by map display resolution.

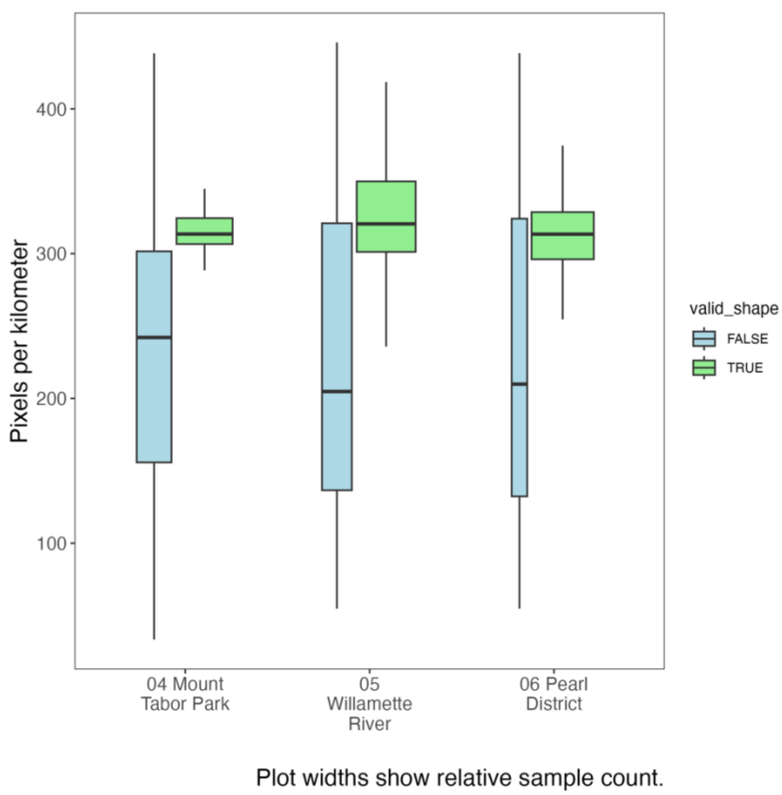

Figure 29. Distribution of map display resolution, grouped by shape validity.

Applying the same modeling approach used for centroid distance to the relative area metric on exercises 4-6 also resulted in models with a significant dependency on map scale (adjusted R-squared 0.32, p<5e-15). For details see > # *LM for max px/km (*rel_area), in Appendix D – Regression.

The model fails several DHARMa quality tests (Figure 30, Figure 31), but not as many or to the same extent as found in the centroid distance model. At 0.792, the residual standard error is larger than the interquartile range for this measure. Applying Box-Cox transformation does not produce any improvement, with most diagnostics either remaining the same or declining in quality. There is therefore evidence of a significant trend but the model has little predictive power.

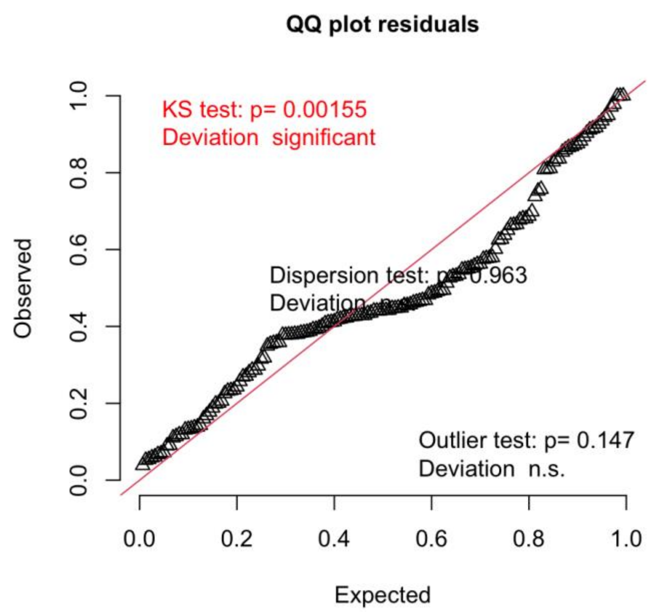

Figure 30. QQ plot, simulated vs. observed residuals for map scale to relative area model.

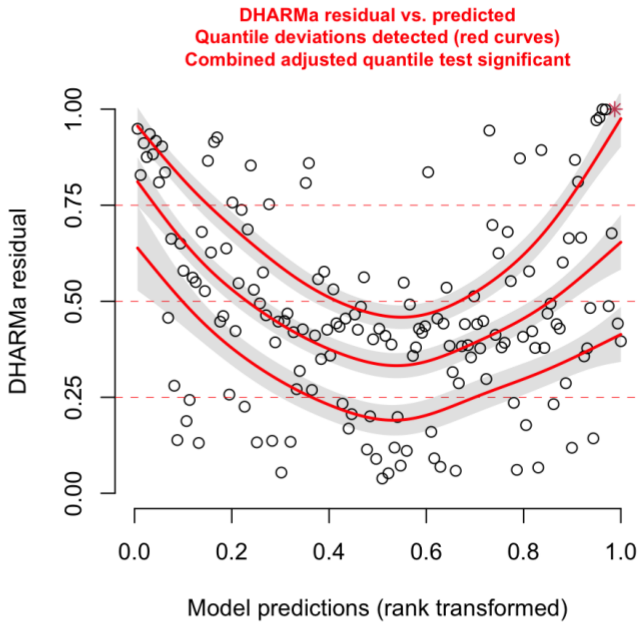

Figure 31. DHARMa residual plot for map scale to relative area model.

The shape defect model also suggests a significant connection between map scale and shape defect (adjusted R-squared 0.129, p=2e-6). See > # *LM for max px/km (*shape_defect), in Appendix D – Regression, for details.

The median value for shape defect over three exercises, 0.11, is identical to the residual standard error of the model, and about the same as the interquartile range of 0.12. There are a few problems highlighted by the DHARMa evaluation as well (Figure 32, Figure 33). With a low R-squared value and relatively high RSE, this model does not offer much predictive power, but it supports the trend.

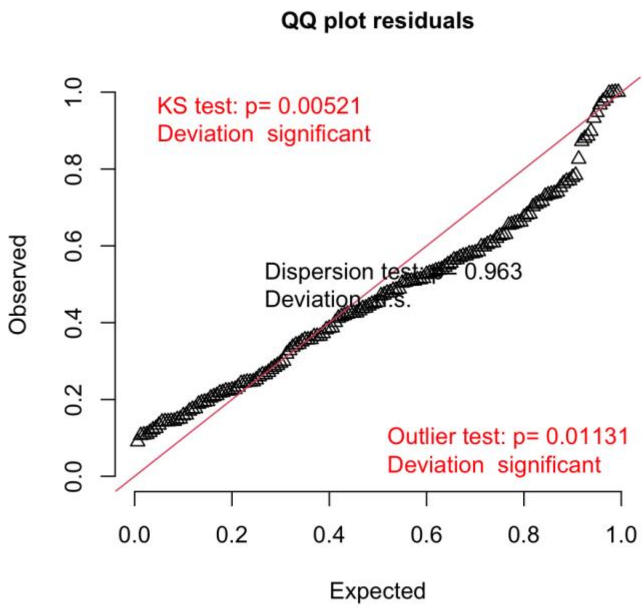

Figure 32. QQ plot, simulated vs. observed residuals for map scale to shape defect model.

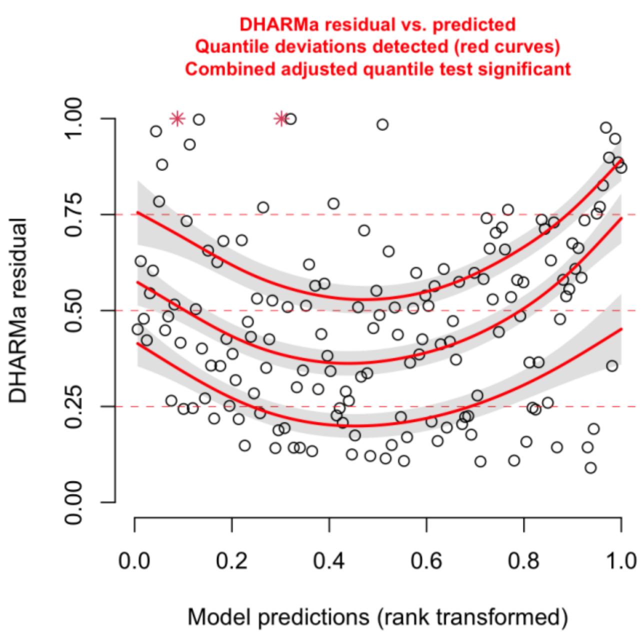

Figure 33. DHARMa residual plot for map scale to shape defect model.

Table 16 contains a list of median and interquartile range values for each of the evaluation metrics calculated for all exercises.

Table 16. Area metrics summary.

	Relative Distance		Relative Area		Shape Defect	
Activity	**Median**	**IQR**	**Median**	**IQR**	**Median**	**IQR**
04 Mount Tabor Park	0.06	0.07	1.49	0.31	0.16	0.09
05 Willamette River	0.16	0.20	1.50	0.77	0.08	0.11
06 Pearl District	0.45	0.43	1.04	1.00	0.08	0.09

Summary

The key objective of this project was to quantify the accuracy with which survey participants marked locations on this map. Several hypothesized explanations for variations in data quality were examined and rejected, including levels of application skill, life experience, time spent entering data, and measures of application ease of use. The two variables that appear to have some effect on data quality are map scale (highly significant) and input device (approaching statistical significance).

Discussion

This project defined a successful point location as falling within the target footprint. Similarly, an area submission was counted as a success if its centroid fell close to the centroid of the target, with the critical distance depending on the area of the target. Based on this rubric, participants had little difficulty locating points on the map, with a success rate of 90%. The improved 95% success rate for areas may be related to gained experience or may result from the averaging effect of multiple data points.

Users tended to overfill the space when performing area exercises: the median submitted area is 19% larger than the target overall. About 65% of all submissions filled an area between 75% and 175% of the target.

For the well-defined areas, Mount Tabor and the Willamette River between two bridges, the median area filled is about 1.5 times the actual area. There is much less certainty about the boundaries of the Pearl District among participants: the crowdsourced area estimate is very similar to the target, but the interquartile range is larger than it is for Mount Tabor or Willamette River submissions.

Most submitted polygons (84%) are rounder than the target areas. This should not be surprising, since the unit of data entry is a circle. While shapes visually approximate the targets (see Figure 8, Figure 11, Figure 15), only 42% of submitted shapes are as similar as 2:1 rectangles and squares – not similar in absolute terms. Perhaps a more nuanced understanding of the shape metric used in the analysis could be achieved by comparing the shape metrics of a set of polygons generalized by methods typically used for cartography (Wang & Müller, 1998).

If this data entry tool is applied to future quantitative work, it will be important to perform a calibration exercise to take these tendencies into account. The relationship between displayed map scale and accuracy described at length in the Results section suggests that limiting data entry to higher zoom levels would improve data quality.

Directions for Future Work

Results from this project as well as unforeseen difficulties with the design of the survey used to collect the data provide some opportunities for improvement and unanswered questions. For example, areas entered for exercises where a boundary was shown on the basemap tended to be 1.5 times the area on the map. Only one exercise, the Pearl District, lacked a map boundary. In this case, the median submitted area was much closer to the true target area. Is this repeatable? One way to ask this question would be to present the same exercise with two basemaps (to random participants): one showing the boundary and one not. The map and survey software used for this project would support this A/B experiment without modification.

All the points used in the test survey likewise have a distinct building outline or boundary, plus a clear text label. A follow-up survey using points without this kind of hinting may produce different results.

A practical application of this tool is the Wildfire Governance Survey for which data collection closed in Fall 2024. The basemap used here, Esri National Geographic, a relief map with a balance of natural features and human-built structures, was also used for that survey effort. But the map scale of that survey is quite different from the tests reported here. For example, the minimum circle size for that map configuration is 0.5 mi (2640 ft). The minimum circle size for the test survey was 800 ft. The test survey reported here was focused on urban areas that were likely to be familiar to the participants. It could have focused on larger natural features such as mountains and lakes, or major recreation sites. Testing the tool at a similar scale to its application would make for more comparable results and expand understanding of the accuracy of public participation GIS data.

A follow-up study could also improve the interpretability of the data by focusing on repeated exercises and direct comparisons. The future study could include exercises that would fill several areas that are both marked (lakes, rivers) and unmarked on the map (neighborhoods, districts, regions). Effective A/B testing with different basemaps could show the effect of boundaries and labeling. Setting zoom level thresholds for data entry could improve quality by limiting unintentional data points and ensuring that a higher level of detail is displayed when the map is open for input.

Limitations

Changes to the survey might have controlled for some variables and allowed stronger conclusions. The sample count is also smaller than hoped. A participant sample that matches the target population for similar surveys could expand the applicability of results. An alternative survey could have also provided a larger menu of locations and allowed participants to choose the ones with which they were most familiar.

Some participants may have used outside references when performing mapping exercises. Gathering more detailed feedback from participants, such as whether a participant looked at outside references, would be valuable for interpreting results. Outside references significantly change the scope of the claims one can make, particularly about the Pearl District exercise.

Conclusion

This project evaluated volunteers' ability to mark features on a web map. While additional research could refine these findings, over 90% of users successfully marked points and areas as instructed.

The collected data support the use of this software and methodology for gathering geospatial data from the public, provided sufficient validation is in place to avoid overinterpreting potentially erroneous data.

The project addressed several key questions:

What quantifiable spatial errors occur when volunteers mark locations, and what methodological framework can effectively characterize data entry deviations?

The project developed multiple metrics to measure how participant-submitted points and polygons deviated from target features. For points, the evaluation considered intersection with target polygons and distance from target centroids. For polygons, the analysis examined centroid distance, area, and a shape metric. Valid polygons were defined as those whose centroids fell within a target's equal-area circle, limiting deviation from the target polygon's center (see Figure 6). This approach accepted submissions near irregular target centers (as demonstrated in the Willamette River exercise, Figure 10) while rejecting significantly off-center submissions.

Using the proposed evaluation rubric, what is the empirical error rate, and what does this reveal about data source reliability?

The distance-based rubric showed a 92% success rate across all attempts. However, when examining area and shape metrics, most submissions exceeded target size (median 1.45, interquartile range 1.13-1.78). While the specific causes warrant further investigation, this likely stemmed from limitations in the data entry method. Users with relevant geographic knowledge but no prior experience with the mapping tool could still accurately mark features.

A notable concern was the 5-10% of submissions that deviated significantly from targets. Error patterns suggested that accuracy depended more on map context than participant experience or background. While clearly marked features were generally marked accurately, a significant error rate persisted that training and software modifications seem unlikely to eliminate.

The percent error rate remained low even at small target size. The Millar Library exercise target was about 60 meters square, and all successful submissions were less than 27 meters from its centroid. The web map may work well to point out features this small, if there is sufficient detail on the map.

How do systematic data patterns inform approaches to mitigating geospatial data entry errors?

For feature creation, the best strategy for avoiding errors would be to limit data entry at small map scales. As this is already a configurable feature in the map software, it should be applied on future projects.

Projects soliciting volunteered location data needing greater precision could benefit from methods for multiple independent sampling, but this will rarely be feasible. Fortunately, much PPGIS data collection demands little precision, including the current project for which the tool is being used. For example, the immediate application of this data entry tool is locating up to 25,000-acre project sites within landscapes divided into districts of about 250,000 acres (Anderson et al., 2024).

Individual user-submitted polygon quality varies widely, but submissions closest to the mean centroid, median area, and median shape provide an improved approximation of the target. Input from multiple participants can be aggregated to improve confidence, as seen above in Figure 16.

Expanding training materials has the potential to improve performance. However, any such effort requires balancing against volunteer time commitment and potential participant dropout. Some PPGIS users never read the instructions at all (Poplin, 2015).

Only a simple training exercise was included in the project survey. More training material, while remaining sensitive to volunteer time commitment, could also give survey participants time to become familiar with the basemap symbology. Additional training would yield no benefit for those who do not read instructions and may increase the participant dropout rate.

The mapping software at the center of this project is meant to be used for practical data collection. Minimally trained participants in this project accurately located the targets, but this cannot necessarily be taken for granted for all geographies and all basemaps. When designing a project, it may be best practice to include sentinel exercises on known features relevant to the project. The data set would thus provide some built-in assurance of a participant's diligence and the resulting data quality.

References

Anderson, R. M., Hixson, K., Johnston, K., Mockrin, M. H., & Mowery, M. (2024). *Adapting to Wildfire: A Review of Ongoing Wildfire Planning and Mitigation Efforts Across 4 Ten Landscapes in the Western United States* (General Technical Report No. PNW-GTR-XXX [in press]). U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Benzaghta, M. A., Elwalda, A., Mousa, M., Erkan, I., & Rahman, M. (2021). SWOT analysis applications: An integrative literature review. *Journal of Global Business Insights*, *6*(1), 55–73. https://doi.org/10.5038/2640-6489.6.1.1148

Bobbitt, Z. (2020, October 13). How to Perform a Box-Cox Transformation in R (With Examples). *Statology*. https://www.statology.org/box-cox-transformation-in-r/

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society*, *26*(2), 211–252.

Campbell, Lindsay, Evers, Cody, & Hickson, Kathy. (2024, August). *Fueling Adaptation: Mapping organizations and networks addressing wildfire risk in the Wildfire Crisis Strategy's Colorado Front Range Landscape*. https://research.fs.usda.gov/sites/default/files/2024-08/pnw-fuelingadaptationonepager-coloradofrontrange.pdf

Canonical & contributors. (2024). *Ubuntu Linux* (Version 22.04.5) [Python]. Canonical. https://ubuntu.com/

Chaskin, R. J. (1997). Perspectives on Neighborhood and Community: A Review of the Literature. *Social Service Review*, *71*(4), 521–547. https://doi.org/10.1086/604277

City of Portland. (n.d.). *Pearl District*. City of Portland Official Website. Retrieved November 14, 2024, from https://www.portland.gov/neighborhoods/pearl-district

Deng, Y. (2016). Challenges and complications in neighborhood mapping: From neighborhood concept to operationalization. *Journal of Geographical Systems*, *18*(3), 229–248. https://doi.org/10.1007/s10109-016-0232-z

Django Software Foundation. (2024). *Django* (Version 5.1) [Python, Javascript, HTML]. Django Software Foundation. https://www.djangoproject.com/

Encode OSS Ltd. (2024). *Django REST Framework* (Version 5.1) [Python, Javascript, HTML]. Encode OSS Ltd. https://www.django-rest-framework.org/

Esri. (2024). *National Geographic World Map* [Basemap tile layer]. https://www.arcgis.com/home/item.html?id=b9b1b422198944fbbd5250b3241691b6

Evers, C. (2020). *Tending the Fire: Wildfire Risk Management at the Interface* [Dissertations and Theses, Portland State University]. https://doi.org/10.15760/etd.7289

Fan, H., Yang, B., Zipf, A., & Rousell, A. (2016). A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. *International Journal of Geographical Information Science*, *30*(4), 748–764. https://doi.org/10.1080/13658816.2015.1100732

GDAL/OGR contributors. (2024). *GDAL/OGR Geospatial Data Abstraction Software Library* [Computer software]. Open-Source Geospatial Foundation. https://doi.org/10.5281/zenodo.5884351

Gieseking, J. J. (2018). Operating anew: Queering GIS with good enough software: Operating anew. *The Canadian Geographer / Le Géographe Canadien*, *62*(1), 55–66. https://doi.org/10.1111/cag.12397

Google. (n.d.). *Council Crest Park*. Google Maps. Retrieved November 22, 2024, from https://maps.app.goo.gl/ysR9oGXQccQRUCZ67

Gottwald, S., Laatikainen, T. E., & Kyttä, M. (2016). Exploring the usability of PPGIS among older adults: Challenges and opportunities. *International Journal of Geographical Information Science*, *30*(12), 2321–2338. https://doi.org/10.1080/13658816.2016.1170837

Gregory Oschwald. (2024). *GeoIP2* (Version 4.8) [Python]. MaxMind. https://www.maxmind.com/en/geoip-databases

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal*, *40*(1), 27–37. https://doi.org/10.1179/000870403235002042

Hartig, F. (2019, November 19). CrossValidated: How to interpret "quantile residuals" [StackExchange]. *CrossValidated*. https://stats.stackexchange.com/questions/436481/how-to-interpret-quantile-residuals

Hartig, F. (2024). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models* [Computer software]. https://CRAN.R-project.org/package=DHARMa

Hitchins, T. (2018). *Assessing PPGIS Usability and its Relationship to Spatial Data Production: A Case Study* [Dissertations and Theses, Portland State University]. https://doi.org/10.15760/etd.6405

jQuery. (n.d.). [Project home page and documentation site]. jQuery. Retrieved December 1, 2024, from https://jquery.com/

Kim, H.-C., Son, W.-J., Lee, J.-S., & Cho, I.-S. (2024). Identification of Maritime Areas With High Vessel Traffic Based on Polygon Shape Similarity. *IEEE Access*, *12*, 92253–92267. https://doi.org/10.1109/ACCESS.2024.3422398

Merkel, D. (2014). Docker: Lightweight linux containers for consistent development and deployment. *Linux Journal*, *2014*(239), 2.

NAD83 / Oregon LCC (m). (2021, September 10). *Epsg.Io*. https://epsg.io/2991

OpenLayers. (n.d.). [Project home page and documentation site]. OpenLayers. Retrieved November 30, 2024, from https://openlayers.org/

OpenStreetMap contributors. (2024). *OpenStreetMap*. https://tile.openstreetmap.org/

Oregon Coordinate Systems. (n.d.). [State government]. *Geospatial Enterprise Operations*. Retrieved March 6, 2025, from https://www.oregon.gov/eis/geo/Pages/coordinate-system.aspx

Oregon Lambert Error Statistics. (1996). State of Oregon.

Palisades Fire. (2025, February 10). [Government Agency]. *The Department of Forestry and Fire Protection, Current Emergency Incidents*. https://www.fire.ca.gov/incidents/2025/1/7/palisades-fire

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439. https://doi.org/10.32614/RJ-2018-009

Polsby, D. D., & Popper, R. (1991). The Third Criterion: Compactness as a Procedural Safeguard Against Partisan Gerrymandering. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2936284

Poplin, A. (2015). How user-friendly are online interactive maps? Survey based on experiments with heterogeneous users. *Cartography and Geographic Information Science*, *42*(4), 358–376. https://doi.org/10.1080/15230406.2014.991427

Posit team. (2024). *RStudio: Integrated Development Environment for R* [Computer software]. Posit Software, PBC. http://www.posit.co/

PostGIS Project Steering Committee & others. (2024). *PostGIS, spatial and geographic objects for postgreSQL* [Computer software]. https://postgis.net

PostgreSQL contributors. (2024). *PostgreSQL* [Computer software]. https://www.postgresql.org/

QGIS.org. (2024). *QGIS Geographic Information System* [Computer software]. QGIS Association. http://www.qgis.org

Qualtrics. (n.d.). *ArcGIS Extension Basic Overview*. Qualtrics. Retrieved November 14, 2024, from https://www.qualtrics.com/support/integrations/arcgis-extension-basic-overview/

Qualtrics. (2024, June 20). *Qualtrics Weekly Product Release Notes—June 20, 2024*. Qualtrics. https://community.qualtrics.com/product-release-notes-96/weekly-product-release-notes-june-20-2024-28345

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schuurman, N. (2000). Trouble in the heartland: GIS and its critics in the 1990s. *Progress in Human Geography*, *24*(4), 569–590.

Schuurman, N. (2006). Formalization Matters: Critical GIS and Ontology Research. *Annals of the Association of American Geographers*, *96*(4), 726–739. https://doi.org/10.1111/j.1467-8306.2006.00513.x

Tang, Z., & Liu, T. (2016). Evaluating Internet-based public participation GIS (PPGIS) and volunteered geographic information (VGI) in environmental planning and management. *Journal of Environmental Planning and Management*, *59*(6), 1073–1090. https://doi.org/10.1080/09640568.2015.1054477

Tennekes, M. (2018). **tmap**: Thematic Maps in R. *Journal of Statistical Software*, *84*(6). https://doi.org/10.18637/jss.v084.i06

Tennekes, M., & Puts, M. (2023). cols4all: A Color Palette Analysis Tool. *EuroVis 2023 - Short Papers*, 37–41. https://doi.org/10.2312/EVS.20231040

USDA US Forest Service Pacific Northwest Research Station. (n.d.). *Fueling Adaptation: Leveraging Community Capacity to Reduce Wildfire Risk* [Government Agency]. US Forest Service Research. Retrieved November 25, 2024, from https://research.fs.usda.gov/pnw/projects/bil-fueling-adaptation-socioecological-assessment

user25201, radouxju, Taylor H., & Vince. (2014, February 7). Calculating roundness/compactness of polygon? [Expert discussion forum]. *GIS StackExchange*. https://gis.stackexchange.com/questions/85812/calculating-roundness-compactness-of-polygon

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed). Springer.

Wang, Z., & Müller, J.-C. (1998). Line Generalization Based on Analysis of Shape Characteristics. *Cartography and Geographic Information Systems*, *25*(1), 3–15. https://doi.org/10.1559/152304098782441750

What is the WUI? (n.d.). U.S. Fire Administration. Retrieved January 14, 2025, from https://www.usfa.fema.gov/wui/what-is-the-wui.html

Wickham, H. (2011). Ggplot2. *WIREs Computational Statistics*, *3*(2), 180–185. https://doi.org/10.1002/wics.147

Appendix A - Survey Questions

Following is a rendering of the survey text and question flow as presented to participants. Items in [brackets] are explanatory or placeholders and did not literally appear on the survey.

[introduction/consent/instructions]

Help us improve a map-based survey tool!
We want to improve our understanding of how online map users locate and mark points on a map. This survey will lead you through a few simple exercises. There is no right or wrong answer. The survey is anonymous. We do request a few pieces of personal information, such as your age and zip code, and how long you have lived in the Portland area. The survey software also uses your IP address to estimate your computer's approximate location. We record metadata about your interaction with the maps such as zooming, panning and clicking. **There is no way to identify you from this information.**
If there are any questions you don't want to answer, you can skip to the next question.
Should you take this survey? You can complete this survey whether you are familiar with web-based mapping or not. It should take between 5 and 10 minutes to complete.
Are there any risks? We don't know of any risks to you from being in the study that are greater than the risks you encounter in everyday life.
Any amount of data is helpful. Please complete the survey to the best of your ability. You can skip to the end and submit a partial survey if you run out of time.
Thank you for your contribution to this research!

   a. I want to take the survey
   b. I would like to stop now [logic ends survey here]

Section 1: Getting started

1. How familiar are you with mapping and way finding apps, such as Google Maps, Apple Maps, AllTrails, TrailForks, etc.?
   a. Never used
   b. Not familiar at all
   c. Slightly familiar
   d. Moderately familiar
   e. Extremely familiar
2. Have you used any GIS or map editing tools, such as ArcGIS Pro, ArcGIS Online, MapBox, Carto, Felt, etc.?
   a. Never used
   b. Used casually
   c. Have extensive experience/formal training
3. How long have you lived in the Portland area?
   a. Never
   b. Less than a year
   c. 1-5 years
   d. 5-10 years
   e. 10-15 years
   f. More than 15 years

Section 2: Map Tutorial

Spend a moment becoming familiar with the map below.
On desktop browsers, you can zoom in and out with an up and down scrolling action and pan the view around by clicking and dragging.
If using a touchscreen, pinch to zoom and drag to move the view around.

Zooming adjusts the circle size! The bottom-right corner of the map shows the circle diameter. **Clicking or tapping** on the map **marks a circle** at that location. On laptops or desktops, the mouse pointer also shows the size of the circle.
If you create a point by mistake, click or tap on it and select "Delete" to remove it.

First Task
1. On the map below, zoom into central Portland and notice how the circle diameter changes.
2. Click to create a few points.
3. Then, delete at least one point you created.

[include interactive map of central Portland, Oregon - Figure 34]
Section 3: Locating Features
The following questions ask you to mark some points on the map.

Zoom in or out to set the circle diameter to about 1000 ft. **Click or tap** the center of the feature on the map to create a circle.
Remember that if you make a mistake, you can click and choose "Delete".
If you are unable to find the location on the map, you can skip to the next question.
Note: All of the locations for this survey are inside the red box.

1. Locating Features (1/3)
   Place a 1000-ft circle on Powell's City of Books, NW Burnside at 10th Ave.

[include interactive map of central Portland, Oregon - Figure 34]
]

2. Locating Features (2/3)
   Place a 1000-ft circle on the Portland State University Library (Millar Library), about 2 blocks west of Broadway between Hall and Harrison.

[include interactive map of central Portland, Oregon - Figure 34]

3. Locating Features (3/3)
   Place a 1000-ft circle on Council Crest Park, in Southwest Portland, about a mile west of the Ross Island bridge/US-26.

[include interactive map of central Portland, Oregon - Figure 34]

Section 4: Filling areas
The following questions ask you to fill in specific areas with circles. The purpose of this exercise is to try an alternate method to mark areas on maps (polygons).
Zoom in or out to set the circle diameter you want, then click or tap to create overlapping circles to fill the area.
Use as much or as little detail as you like.

Note: new circles always start outside existing ones.
Figure 35 is a completed example of Lone Fir Cemetery in SE Portland.

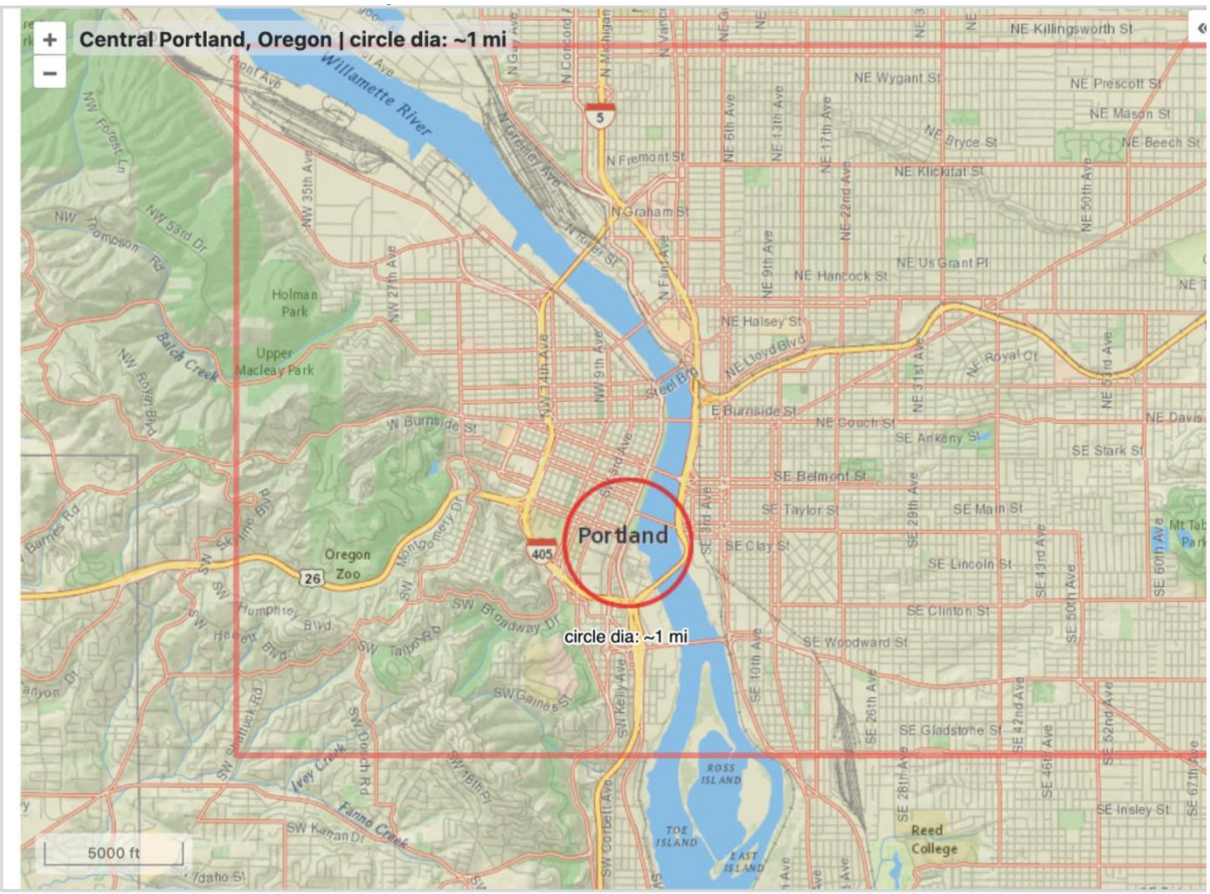

Figure 34. Interactive map displayed for survey.

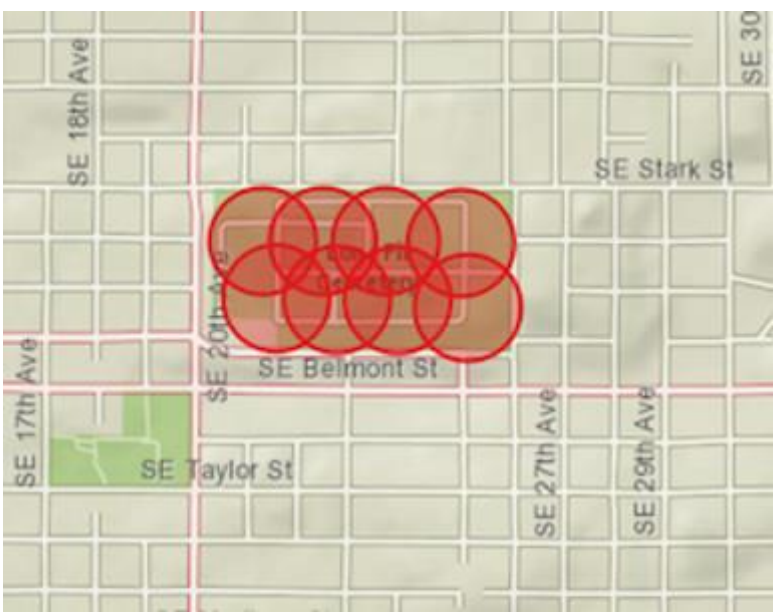

Figure 35. Example of filling an area with circles.

It's OK to not completely fill the area. Just cover most of it.

If you are unable to find the region on the map, you can skip to the next question.

1. Filling Areas (1/3)
   Locate Mt. Tabor Park, in SE Portland, north of Division Street and east of 60^{th} Avenue.
   Fill in the park with circles.
   Recommended circle diameter is about 1500 ft.

[include interactive map of central Portland, Oregon - Figure 34]

2. Filling Areas (2/3)
   Center the map on the Willamette River, between the Broadway and Hawthorne Bridges (roughly between NW Lovejoy and SE Hawthorne Streets).
   Fill this portion of the river with circles.
   Recommended diameter is 1000-1500 ft.

[include interactive map of central Portland, Oregon - Figure 34]

3. Filling Areas (3/3)
   Now, center the map on the Pearl District, in inner NW Portland.
   Fill in the neighborhood with circles.
   Recommended circle diameter: about 1500 ft.

[include interactive map of central Portland, Oregon - Figure 34]

Section 5: Feedback and finishing up

1. How familiar are you with the places named in the mapping portion of this survey? Did you know about where to find them on the map?
   a. Never visited/unsure
   b. Not too familiar (one or two)
   c. Slightly familiar (some of them)
   d. Moderately familiar (most of them)
   e. Extremely familiar (all of them)
2. Overall, how easy or difficult was it to read details on the map?
   a. Extremely difficult
   b. Moderately difficult
   c. Neither easy nor difficult
   d. Moderately easy
   e. Extremely easy
3. How easy or difficult was it to zoom, pan, and mark items?
   a. Extremely difficult
   b. Moderately difficult
   c. Neither easy nor difficult
   d. Moderately easy
   e. Extremely easy
4. Did the mapping activity take more or less time than you expected?
   a. Quite a bit less time
   b. Somewhat less time
   c. About as much time as expected

    d. Somewhat more time
    e. Quite a bit more time
5. Were there any elements that were missing from the map that would have helped you, such as:
    a. Map legend
    b. Additional map layers/details
    c. Online help, tutorial, or other documentation
    d. Something else (please mention below)
    e. None of these
6. Was there anything else about the map that was particularly helpful? Difficult? Anything else you noticed? [fill in the blank]
7. How old are you?
    a. Under 18
    b. 18-24 years old
    c. 25-34 years old
    d. 35-44 years old
    e. 45-54 years old
    f. 55-64 years old
    g. 65+ years old
8. What kind of device are you using to complete this survey?
    a. Phone
    b. Laptop
    c. Desktop
    d. Tablet
9. What is your US Zip Code? [fill in the blank]

Click the submit button below to submit your survey.
We thank you for your time spent taking this survey.
Your response has been recorded.

[end of survey]

Appendix B – User comments on survey

User device	comments
Desktop	I liked how it shows what the circle diameter is. Some streets were missing that would have been helpful.
Desktop	I liked that a click was all needed to mark, delete, or drag/pan.
Desktop	Re previous question: neighborhoods here have wonky and unclear borders, so what exactly are the district limits...and who cares.
Desktop	The instruction to use a certain diameter circle and the identification of that zoom diameter on the map was very useful. Finding a way to highlight that zoom diameter on map (call attention to it) would be useful. It took me a few seconds to find it.
Desktop	The interface won't allow the placing of overlapped circles. It seems that if the cursor is inside a placed circle, then the only option is to delete the circle, not placing a new one. I used a fairly typical size monitor and had difficulty to see the instructions and the map or the upper part of the map and the circle size text. It would be helpful to tell the users that the map questions would appear on the next page (after they click Next) given the fact that the first question has the map on the same page.
Desktop	The street names showing up was helpful, but I do not know where the pearl district is, so having neighborhoods named would have been nice.
Laptop	A dynamic scale bar would be helpful. Filling in areas was a little clunky and inexact... When marking points, a crosshair in the circle might be better than the mouse cursor.
Laptop	A layer with the district labels would have been helpful.
Laptop	Fonts were small, moved it to a bigger monitor. Thought question 4 wasn't working as I expected the first map on that actual page.
Laptop	I couldn't get the circles the right recommended size. Most were too big or too small.
Laptop	I'm used to seeing neighborhoods overlaid in Google Maps. My mental map of the Pearl is pretty nebulous.
Laptop	If it had neighborhood designations, I missed them.
Laptop	It might be nice to be able to search text on the map itself using ctrl-f.
Laptop	labels for the places would have been helpful when zoomed out, street labels were sparse when zoomed in all the way, made it difficult to navigate to the library
Laptop	More buttons would be nice. Right clicking should do more. I'm using a laptop without a mouse, so I couldn't zoom to 1000 ft diameter circle (800 or 1200).
Laptop	Neighborhood indications would have made it easier for me to make sure I was accurately marking out the Pearl District
Laptop	Once or twice I used google maps to locate where a feature was so I could mark in with circles. It's not clear if it was more important to fill an area completely with circles and go outside the boundaries mentioned or use fewer circles not cover an area completely.

User device	comments
Laptop	Other shapes for the highlighting area would be nice, also if you could click and drag to highlight an entire area rather than using the circles to encapsulate the entire area.
Laptop	Panning in and out made street names easier to read. Took a couple of attempts to "get the hang" of it.
Laptop	The scale in the lower left I mistook for circle diameter and didn't see the actual circle diameter value next to the circle until a later task.
Laptop	trackpad zoom was slow when using the pinch gesture, but the scroll gesture worked way better
Laptop	Two things: One is about terminology. In your introductory instructions, the diameter of the circle did not change with zooming in and out. What changed was the distance REPRESENTED BY the diameter. The SCREEN diameter remained the same. There were plenty of visual cues to the change in diameter to guide me through, though. Second, the jumps in scales often made it impossible to settle on the diameter suggested by the exercise.
Phone	Hard to see circle size on mobile device
Phone	I had to zoom in a lot to get bridge names, and some street names (for example, Lovejoy) never appeared. I was familiar with the way streets in NW Portland worked, so I knew where Lovejoy was, but I never saw the street name on the map. (As for the bridges, I mostly knew the ones in the northern part of the city, because that's where I lived.)
Phone	I'm on an iphone and it's hard to anticipate where the circle will be. So I wound up with gaps or overlapping circles. Getting the zoom is right is always a challenge so you can see the details but don't lose context.
Phone	It was easier to place circles accurately when I realized I could move them if I didn't lift my finger.
Phone	Like I sort of know where the pearl is
Phone	On the filling exercises, wasn't sure if every inch of target had to be covered? I aimed for rough estimate
Phone	Some street names disappeared if you assumed in too close
Phone	Zooming to the recommended circle diameter scale was a bit tedious. I would love to be able to punch in a scale numerically.
Tablet	I didn't notice that I could adjust the diameter of the circle until close to the end, but then again, I'm not the sharpest tool in the shed.

Appendix C – Software

Map/data collection

The survey data for this paper was collected using Qualtrics software, June 2024 version of Qualtrics. Copyright © 2024 Qualtrics, Provo, UT, USA. https://www.qualtrics.com

The geospatial data for this paper was collected using a stack of open-source software including the following projects. Code available at https://github.com/csghormley/point-it-out.

> Django, with GeoDjango extension (Django Software Foundation, 2024)
> Django REST framework (Encode OSS Ltd., 2024)
> GeoIP2 (Gregory Oschwald, 2024)
> GDAL (GDAL/OGR contributors, 2024)
>
> OpenLayers 9 (*OpenLayers*, n.d.)
> jQuery 3.7.1 (*jQuery*, n.d.)
>
> PostgreSQL 16, with postGIS extension (PostgreSQL contributors, 2024), (PostGIS Project Steering Committee & others, 2024)
>
> Docker (Merkel, 2014)
> Ubuntu Linux 22.04 (Canonical & contributors, 2024)

Analysis

Data analysis for this project was prepared using *R* (R Core Team, 2024) and *RStudio* (Posit team, 2024), making use of the *Tidyverse* suite of data engineering tools (Wickham et al., 2019). *DHARMa* (Hartig, 2024) was particularly helpful for evaluating regression models.

Maps were generated using *tmap* (Tennekes, 2018). Other visualizations made use of *ggplot2* (Wickham, 2011). Color palettes for visualizations provided by *ColorBrewer* (Harrower & Brewer, 2003) and *Cols4All* (Tennekes & Puts, 2023).

Polygons used in the analysis were drawn in QGIS (QGIS.org, 2024).

Maps displayed during data collection and included in this paper contain the Esri National Geographic World Map tile layer as a basemap (Esri, 2024).

Appendix D – Regression

```
> require(dplyr, MASS)
> HOME<-"~"
> graded_resp <- read_rds(glue("{HOME}/graded_resp.RDS"))
> area_resp <- graded_resp %>% filter(projectid>3)
```

> # GLM with time related variables

```
> centroid_glm <- glm(valid_centroid ~ time_more_less + duration + is.na(comments),
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ time_more_less + duration + is.na(comments),
    family = binomial(link = "logit"), data = graded_resp)

Coefficients:
                                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                                  1.66e+01   9.80e+02    0.02     0.99
time_more_lessSomewhat less time            -1.46e+01   9.80e+02   -0.01     0.99
time_more_lessAbout as much time as expected -1.38e+01   9.80e+02   -0.01     0.99
time_more_lessSomewhat more time            -1.35e+01   9.80e+02   -0.01     0.99
duration                                     8.42e-06   2.94e-05    0.29     0.77
is.na(comments)TRUE                         -5.28e-01   4.45e-01   -1.19     0.24

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 167.26  on 299  degrees of freedom
Residual deviance: 161.50  on 294  degrees of freedom
  (27 observations deleted due to missingness)
AIC: 173.5

Number of Fisher Scoring iterations: 15
```

> # GLM for GIS experience

```
> centroid_glm <- glm(valid_centroid ~ used_gis,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ used_gis, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                                  Estimate Std. Error z value Pr(>|z|)
(Intercept)                          2.260      0.303    7.45  9.3e-14 ***
used_gisUsed casually                0.719      0.665    1.08     0.28
used_gisHave extensive experience    0.305      0.460    0.66     0.51
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 169.70  on 314  degrees of freedom
Residual deviance: 168.32  on 312  degrees of freedom
  (12 observations deleted due to missingness)
AIC: 174.3

Number of Fisher Scoring iterations: 5
```

> # GLM for mapping experience

```
> centroid_glm <- glm(valid_centroid ~ used_mapping,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)
```

```
Call:
glm(formula = valid_centroid ~ used_mapping, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                                Estimate Std. Error z value Pr(>|z|)
(Intercept)                        0.693      1.225    0.57     0.57
used_mappingSlightly familiar      1.253      1.439    0.87     0.38
used_mappingModerately familiar    1.771      1.261    1.40     0.16
used_mappingExtremely familiar     2.051      1.272    1.61     0.11

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 170.64  on 320  degrees of freedom
Residual deviance: 167.97  on 317  degrees of freedom
  (6 observations deleted due to missingness)
AIC: 176

Number of Fisher Scoring iterations: 5
```

> # GLM for time in city

```
> centroid_glm <- glm(valid_centroid ~ lived_portland,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ lived_portland, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                                 Estimate Std. Error z value Pr(>|z|)
(Intercept)                        2.7726     1.0308    2.69   0.0071 **
lived_portlandLess than a year     0.0606     1.4565    0.04   0.9668
lived_portland1-5 years           -0.6931     1.1989   -0.58   0.5632
lived_portland5-10 years           0.1452     1.2607    0.12   0.9083
lived_portland10-15 years         -1.3863     1.1456   -1.21   0.2263
lived_portlandMore than 15 years  -0.0480     1.0730   -0.04   0.9643
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 170.64  on 320  degrees of freedom
Residual deviance: 165.13  on 315  degrees of freedom
  (6 observations deleted due to missingness)
AIC: 177.1

Number of Fisher Scoring iterations: 5
```

> # GLM for age group

```
> centroid_glm <- glm(valid_centroid ~ age_group,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ age_group, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)                 1.609      1.095    1.47     0.14
age_group25-34 years old    1.459      1.313    1.11     0.27
age_group35-44 years old    1.224      1.315    0.93     0.35
age_group45-54 years old    0.809      1.133    0.71     0.48
age_group55-64 years old    0.215      1.197    0.18     0.86
age_group65+ years old      1.526      1.498    1.02     0.31
```

```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 168.25  on 305  degrees of freedom
Residual deviance: 164.55  on 300  degrees of freedom
  (21 observations deleted due to missingness)
AIC: 176.6

Number of Fisher Scoring iterations: 5
```

> # GLM for ease of use (1)

```
> centroid_glm <- glm(valid_centroid ~ easy_zoom_pan_mark + easy_read_details,
+                      data=graded_resp, family = binomial(link="logit"))
Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ easy_zoom_pan_mark + easy_read_details,
    family = binomial(link = "logit"), data = graded_resp)

Coefficients:
                                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                                     15.65    1715.42    0.01     0.99
easy_zoom_pan_markModerately difficult          18.37    2662.86    0.01     0.99
easy_zoom_pan_markNeither easy nor difficult    17.49    1715.42    0.01     0.99
easy_zoom_pan_markModerately easy                2.04       1.43    1.42     0.15
easy_zoom_pan_markExtremely easy                 2.19       1.46    1.49     0.14
easy_read_detailsNeither easy nor difficult    -14.96    1715.42   -0.01     0.99
easy_read_detailsModerately easy               -15.45    1715.42   -0.01     0.99
easy_read_detailsExtremely easy                -15.19    1715.42   -0.01     0.99

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 169.22  on 311  degrees of freedom
Residual deviance: 163.08  on 304  degrees of freedom
  (15 observations deleted due to missingness)
AIC: 179.1

Number of Fisher Scoring iterations: 17
```

> # GLM for ease of use (2)

```
> centroid_glm <- glm(valid_centroid ~ easy_zoom_pan_mark,
+                      data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ easy_zoom_pan_mark, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                                     0.693      1.225    0.57     0.57
easy_zoom_pan_markModerately difficult         16.873   1615.104    0.01     0.99
easy_zoom_pan_markNeither easy nor difficult   16.873   1142.052    0.01     0.99
easy_zoom_pan_markModerately easy               1.743      1.256    1.39     0.17
easy_zoom_pan_markExtremely easy                1.792      1.273    1.41     0.16

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 169.22  on 311  degrees of freedom
Residual deviance: 164.68  on 307  degrees of freedom
  (15 observations deleted due to missingness)
AIC: 174.7

Number of Fisher Scoring iterations: 16
```

> # GLM for ease of use (3)

```
> centroid_glm <- glm(valid_centroid ~ easy_read_details,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ easy_read_details, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                                     17.6     1142.1    0.02     0.99
easy_read_detailsNeither easy nor difficult    -15.1     1142.1   -0.01     0.99
easy_read_detailsModerately easy               -15.2     1142.1   -0.01     0.99
easy_read_detailsExtremely easy                -15.0     1142.1   -0.01     0.99

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 169.22  on 311  degrees of freedom
Residual deviance: 167.16  on 308  degrees of freedom
  (15 observations deleted due to missingness)
AIC: 175.2

Number of Fisher Scoring iterations: 16
```

> # GLM for place familiarity

```
> centroid_glm <- glm(valid_centroid ~ familiar_places,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ familiar_places, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                                                  Estimate Std. Error z value Pr(>|z|)
(Intercept)                                           16.6     1073.1    0.02     0.99
familiar_placesSlightly familiar (some of them)      -13.7     1073.1   -0.01     0.99
familiar_placesModerately familiar (most of them)    -14.2     1073.1   -0.01     0.99
familiar_placesExtremely familiar (all of them)      -14.1     1073.1   -0.01     0.99

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 169.22  on 311  degrees of freedom
Residual deviance: 168.17  on 308  degrees of freedom
  (15 observations deleted due to missingness)
AIC: 176.2

Number of Fisher Scoring iterations: 15
```

> # GLM for device used

```
> centroid_glm <- glm(valid_centroid ~ user_device,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ user_device, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)           3.97       1.01    3.93 0.000084 ***
user_deviceLaptop    -1.42       1.06   -1.34    0.179
user_devicePhone     -2.02       1.06   -1.91    0.056 .
user_deviceTablet    -1.57       1.45   -1.08    0.279
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 168.25  on 305  degrees of freedom
Residual deviance: 162.11  on 302  degrees of freedom
  (21 observations deleted due to missingness)
AIC: 170.1

Number of Fisher Scoring iterations: 6
```

> # LM for device used

```
> centroid_lm <- lm(centroid_dist ~ user_device,
+                   data=graded_resp)
> summary(centroid_lm)

Call:
lm(formula = centroid_dist ~ user_device, data = graded_resp)

Residuals:
   Min     1Q Median     3Q    Max
  -195   -118    -94    -34   4005

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)           75.2       51.8    1.45    0.147
user_deviceLaptop     48.8       60.3    0.81    0.418
user_devicePhone     121.2       65.8    1.84    0.066 .
user_deviceTablet     31.1      121.4    0.26    0.798
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 380 on 302 degrees of freedom
  (21 observations deleted due to missingness)
Multiple R-squared:  0.0125,     Adjusted R-squared:  0.00269
F-statistic: 1.27 on 3 and 302 DF,  p-value: 0.283>
```

> # LM for relative distance

```
> rel_lm <- lm(rel_dist ~ user_device,
+                   data=graded_resp)
> summary(rel_lm)

Call:
lm(formula = rel_dist ~ user_device, data = graded_resp)

Residuals:
   Min     1Q Median     3Q    Max
 -1.24  -0.88  -0.75  -0.15  51.24

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.210      0.539    0.39     0.70
user_deviceLaptop    0.723      0.628    1.15     0.25
user_devicePhone     1.032      0.685    1.51     0.13
user_deviceTablet    0.381      1.265    0.30     0.76

Residual standard error: 3.96 on 302 degrees of freedom
  (21 observations deleted due to missingness)
Multiple R-squared:  0.00778,    Adjusted R-squared:  -0.00207
F-statistic: 0.79 on 3 and 302 DF,  p-value: 0.5

> # Transform resolution (meters per pixel) to something easier to conceptualize:
> # pixels per kilometer (1 / resolution * 1000)
> # as px/km increases, detail increases and so should success rate
> # for point exercises min/max are the same, but areas may capture different
> # zoom levels
```

```
> # GLM for min px/km
> centroid_glm <- glm(valid_centroid ~ min_px_km,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ min_px_km, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.28418    0.51995   -0.55     0.58
min_px_km    0.00896    0.00178    5.02    5e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 171.57  on 326  degrees of freedom
Residual deviance: 142.79  on 325  degrees of freedom
AIC: 146.8

Number of Fisher Scoring iterations: 6

> # GLM for max px/km (valid_centroid)
> centroid_glm <- glm(valid_centroid ~ max_px_km,
+                     data=graded_resp, family = binomial(link="logit"))
> summary(centroid_glm)

Call:
glm(formula = valid_centroid ~ max_px_km, family = binomial(link = "logit"),
    data = graded_resp)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.46627    0.52053   -0.90     0.37
max_px_km    0.00930    0.00175    5.32  1.1e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 171.57  on 326  degrees of freedom
Residual deviance: 138.81  on 325  degrees of freedom
AIC: 142.8

Number of Fisher Scoring iterations: 6

> # evaluate max px/km model (valid centroid)
> sr <- simulateResiduals(fittedModel = centroid_glm, n=2000)
>
> plotQQunif(sr)
> plotResiduals(sr)

> # LM for max px/km (centroid distance)
> centroid_lm <- lm(centroid_dist ~ max_px_km, data=graded_resp)
> summary(centroid_lm)

Call:
lm(formula = centroid_dist ~ max_px_km, data = graded_resp)

Residuals:
   Min     1Q Median     3Q    Max
  -384   -122    -55      0   3765

Coefficients:
```

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  464.611     62.800    7.40  1.2e-12 ***
max_px_km     -0.865      0.154   -5.61  4.4e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 354 on 325 degrees of freedom
Multiple R-squared:  0.0882,     Adjusted R-squared:  0.0854
F-statistic: 31.5 on 1 and 325 DF,  p-value: 4.38e-08

> # evaluate max px/km model (centroid distance)
> sr <- simulateResiduals(fittedModel = centroid_lm, n=2000)
>
> plotQQunif(sr)
> plotResiduals(sr)
qu = 0.25, log(sigma) = -3.423734 : outer Newton did not converge fully.
```

> # LM for max px/km (relative distance)

```
> rel_lm <- lm(rel_dist ~ max_px_km, data=graded_resp)
> summary(rel_lm)

Call:
lm(formula = rel_dist ~ max_px_km, data = graded_resp)

Residuals:
   Min     1Q Median     3Q    Max
 -2.69  -1.07  -0.37  -0.08  51.79

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.08225    0.66915    4.61  5.9e-06 ***
max_px_km   -0.00580    0.00164   -3.53  0.00048 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.77 on 325 degrees of freedom
Multiple R-squared:  0.0369,     Adjusted R-squared:  0.0339
F-statistic: 12.5 on 1 and 325 DF,  p-value: 0.000478

>
> # evaluate max px/km model (relative distance)
> sr <- simulateResiduals(fittedModel = rel_lm, n=2000)
>
> plotQQunif(sr)
> plotResiduals(sr)
qu = 0.75, log(sigma) = -3.598596 : outer Newton did not converge fully.
```

> # LM for max px/km (box-cox transform on centroid distance)

```
> # box-cox transformation on centroid_dist ~ max_px_km
> # (only valid for continuous, not boolean, data)
> bc <- boxcox(graded_resp$centroid_dist ~ graded_resp$max_px_km)
> lambda <- bc$x[which.max(bc$y)]
>
> # create a subset including the transformed dependent variable
> bc_lm_graded_resp <- graded_resp %>%
+   select(rel_dist, centroid_dist, max_px_km, user_device) %>%
+   mutate(L_centroid_dist = (centroid_dist^lambda-1)/lambda) %>%
+   mutate(L_rel_dist = (rel_dist^lambda2-1)/lambda2)

> bc_centroid_lm <- lm(L_centroid_dist ~ max_px_km, data=bc_lm_graded_resp)
>
> summary(bc_centroid_lm)

Call:
lm(formula = L_centroid_dist ~ max_px_km, data = bc_lm_graded_resp)
```

```
Residuals:
   Min     1Q Median     3Q    Max 
-4.386 -0.682 -0.077  0.675  4.484 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.924556   0.214971    27.6   <2e-16 ***
max_px_km   -0.006649   0.000528   -12.6   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.21 on 325 degrees of freedom
Multiple R-squared:  0.328,	Adjusted R-squared:  0.326 
F-statistic:  159 on 1 and 325 DF,  p-value: <2e-16

> sr <- simulateResiduals(fittedModel = bc_centroid_lm, n=2000)
> plotQQunif(sr)
> plotResiduals(sr)
> plotResiduals(sr, bc_lm_graded_resp$user_device, xlab="user device")
```

> # LM for max px/km (box-cox transform on relative distance)

```
> bc_rel_lm <- lm(L_rel_dist ~ max_px_km, data=bc_lm_graded_resp)
> 
> summary(bc_rel_lm)

Call:
lm(formula = L_rel_dist ~ max_px_km, data = bc_lm_graded_resp)

Residuals:
   Min     1Q Median     3Q    Max 
-5.759 -0.927  0.026  0.904  5.480 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.71803    0.26870   -2.67   0.0079 ** 
max_px_km   -0.00323    0.00066   -4.90  1.5e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.51 on 325 degrees of freedom
Multiple R-squared:  0.0687,	Adjusted R-squared:  0.0658 
F-statistic:   24 on 1 and 325 DF,  p-value: 1.55e-06

> sr <- simulateResiduals(fittedModel = bc_rel_lm, n=2000)
> 
> plotQQunif(sr)
> plotResiduals(sr)
> plotResiduals(sr, bc_lm_graded_resp$user_device, xlab="user device")
```

> # AREA AND SHAPE MODELS

```
> area_resp <- graded_resp %>% filter(projectid>3)
> 
```

> # LM for max px/km (rel_area)

```
> area_lm <- lm(rel_area ~ max_px_km, data=area_resp)
> summary(area_lm)

Call:
lm(formula = rel_area ~ max_px_km, data = area_resp)

Residuals:
   Min     1Q Median     3Q    Max 
-1.423 -0.373 -0.112  0.277  4.829 
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.26950    0.20398   16.03   <2e-16 ***
max_px_km   -0.00512    0.00059   -8.67    5e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.792 on 157 degrees of freedom
Multiple R-squared:  0.324,	Adjusted R-squared:  0.32
F-statistic: 75.2 on 1 and 157 DF,  p-value: 4.96e-15

>
> # evaluate max px/km model (rel_area)
> sr <- simulateResiduals(fittedModel = area_lm, n=2000)
>
> plotQQunif(sr)
> plotResiduals(sr)
> plotResiduals(sr, area_resp$user_device, xlab="user device")
> plotResiduals(sr, area_resp$projectid)
>
> # box-cox transformation on rel_area ~ max_px_km
> bc <- boxcox(area_resp$rel_area ~ area_resp$max_px_km)
> lambda <- bc$x[which.max(bc$y)]
>
> # box-cox transformation on shape_defect ~ max_px_km
> bc2 <- boxcox(area_resp$shape_defect ~ area_resp$max_px_km)
> lambda2 <- bc2$x[which.max(bc2$y)]
>
> # create a subset including the transformed dependent variable
> bc_lm_area_resp <- area_resp %>%
+   select(projectid, rel_dist, centroid_dist, rel_area, shape_defect, max_px_km,
user_device) %>%
+   mutate(L_rel_area = (rel_area^lambda-1)/lambda) %>%
+   mutate(L_shape_defect = (shape_defect^lambda2-1)/lambda2)
>
```

> # LM for max px/km (box-cox transform on rel_area)

```
> area_lm <- lm(L_rel_area ~ max_px_km, data=bc_lm_area_resp)
> summary(area_lm)

Call:
lm(formula = L_rel_area ~ max_px_km, data = bc_lm_area_resp)

Residuals:
   Min     1Q Median     3Q    Max
-1.551 -0.178  0.015  0.269  1.968

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.496854   0.144091   10.39  < 2e-16 ***
max_px_km   -0.003322   0.000417   -7.97  3.1e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.56 on 157 degrees of freedom
Multiple R-squared:  0.288,	Adjusted R-squared:  0.283
F-statistic: 63.5 on 1 and 157 DF,  p-value: 3.1e-13

>
> # evaluate max px/km model (box-cox transform on rel_area)
> sr <- simulateResiduals(fittedModel = area_lm, n=2000)
>
> plotQQunif(sr)
> plotResiduals(sr)
> plotResiduals(sr, bc_lm_area_resp$user_device, xlab="user device")
> plotResiduals(sr, bc_lm_area_resp$projectid)
>
```

> # LM for max px/km (shape_defect)

```
> shape_lm <- lm(shape_defect ~ max_px_km, data=area_resp)
> summary(shape_lm)

Call:
lm(formula = shape_defect ~ max_px_km, data = area_resp)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.1517 -0.0707 -0.0080  0.0379  0.6201 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.2640301  0.0283190    9.32   <2e-16 ***
max_px_km   -0.0004045  0.0000819   -4.94    2e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.11 on 157 degrees of freedom
Multiple R-squared:  0.134,	Adjusted R-squared:  0.129 
F-statistic: 24.4 on 1 and 157 DF,  p-value: 2.01e-06

> 
> # evaluate max px/km model (shape_defect)
> sr <- simulateResiduals(fittedModel = shape_lm, n=2000)
> 
> plotQQunif(sr)
> plotResiduals(sr)
> plotResiduals(sr, area_resp$user_device, xlab="user device")
> plotResiduals(sr, area_resp$projectid)
> 
```

> # LM for max px/km (box-cox transform on shape_defect)

```
> shape_lm <- lm(L_shape_defect ~ max_px_km, data=bc_lm_area_resp)
> summary(shape_lm)

Call:
lm(formula = L_shape_defect ~ max_px_km, data = bc_lm_area_resp)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.9458 -0.2335  0.0584  0.1979  1.3744 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1.016877   0.096223  -10.57  < 2e-16 ***
max_px_km   -0.001539   0.000278   -5.53  1.3e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.374 on 157 degrees of freedom
Multiple R-squared:  0.163,	Adjusted R-squared:  0.158 
F-statistic: 30.6 on 1 and 157 DF,  p-value: 1.32e-07

> 
> # evaluate max px/km model (shape_defect)
> sr <- simulateResiduals(fittedModel = shape_lm, n=2000)
> 
> plotQQunif(sr)
> plotResiduals(sr)
> plotResiduals(sr, area_resp$user_device, xlab="user device")
> plotResiduals(sr, area_resp$projectid)
```