Method

1. Participants

Eighteen students (15 females and 3 males) at Purdue University participated in the usability evaluation. Due to software error, one participant's record was not saved and thus was dropped from data analysis. Two participants are master's students and the rest (15 participants) are undergraduate students. Participants' age ranged from 18 to 33 (M = 21.1 and SD = 3.8). Participants reported extensive experience with Google Maps or similar products (e.g., Bing Maps, MapQuest, etc.) but very little experience with GIS applications alike the ones tested in this study.

2. Response Measures and Evaluation design

Response measures of the usability evaluation included: (1) score of the successfulness of each task (0 – completed with ease, 1 – completed with difficulty, and 2 – failed to complete); (2) time to complete each task; (3) number of times help or prompts were needed for each task; (4) number of times errors occurred for each task; and (5) participants' usability ratings of each application, measured by the System Usability Scale (SUS; Brooke 1996). We recorded participants' comments regarding each application during the tasks. We also took notes about participant's behavior when completing the tasks, such as visually searching for particular functions on the interface, facial expressions, and body language.

Due to the time constraint for each evaluation session, a randomized incomplete block design was used, in which each participant completed evaluations of three of the six applications. The order of the applications was balanced across participants to minimize the carryover effect. The detailed evaluation arrangement for participants is shown in Table 1. Note that Participant 6 was dropped from the data analysis.

Table 1. Evaluation arrangement for participants.

Participant	**First Application**	**Second Application**	**Third Application**
1	ESRI Business Analyst Online	Proquest Statistics Datasets	Reference USA
2	Social Explorer	SimplyMap	PolicyMap
3	Reference USA	Proquest Statistics Datasets	ESRI Business Analyst Online
4	Proquest Statistics Datasets	Social Explorer	ESRI Business Analyst Online
5	ESRI Business Analyst Online	Reference USA	Proquest Statistics Datasets
*6**	*PolicyMap*	*SimplyMap*	*Social Explorer*
7	Proquest Statistics Datasets	ESRI Business Analyst Online	Social Explorer
8	Social Explorer	PolicyMap	SimplyMap

9	SimplyMap	PolicyMap	Reference USA
10	Reference USA	ESRI Business Analyst Online	Proquest Statistics Datasets
11	SimplyMap	Reference USA	PolicyMap
12	PolicyMap	Social Explorer	SimplyMap
13	PolicyMap	SimplyMap	Social Explorer
14	SimplyMap	Reference USA	PolicyMap
15	ESRI Business Analyst Online	Reference USA	Proquest Statistics Datasets
16	Social Explorer	PolicyMap	SimplyMap
17	Proquest Statistics Datasets	ESRI Business Analyst Online	Social Explorer
18	Reference USA	Proquest Statistics Datasets	ESRI Business Analyst Online

3. Tasks and Procedure

The tasks participants performed in the usability evaluation included: (1) create a customized map about business-related information (e.g., unemployment rate, number of business establishments, family income, etc.) in the Indianapolis, IN area; (2) change the map unit so that the information is displayed by zip code; (3) change the colors of information areas on the map; (4) change the data ranges corresponding to the different colors; (5) search for another location (Provo, UT) on the map; and (6) export the map and save it to local computer. These tasks represent the typical workflow of using a web-based mapping application for users without GIS background and experience. Therefore, participants' performance of these tasks and their subjective feedback are critical measures of the usability evaluation.

Participants first read and signed a consent form and completed a simple demographic survey. They were given an introduction regarding the basics of web-based mapping applications, including the idea of choosing and displaying spatial data on maps, map units (the basic individual map area for spatial data), map color and legend, and zooming and panning the map. After the introduction, participants completed the tasks with the first application. They were instructed to voice their expectations, difficulties, and general comments about the application during the tasks. Participant activities on the computer screen and voice were recorded using TechSmith Morae software. A researcher sat next to participants, answered questions and provided prompts when participants explicitly requested, and made observation notes about participant behavior. Participants completed the SUS questionnaire and answered open questions about their overall experience of each application. At the end of the evaluation, participants ranked their preferences of the applications they used. Each evaluation session lasted approximately one hour.

Results

1. Descriptive statistics

The descriptive statistics for response measures for each task and application are shown in Tables 2-5. The range of task score is 0 (task completed with ease), 1 (completed with difficulty, and 2 (failed to complete); the time to complete a task is measured in seconds; and the number of helps/prompts for each task and number of errors are simple counts.

Table 2. Descriptive statistics for task score.

Application	Task	Mean	SD	Min.	Max.
ESRI Business Analyst Online	1	0.63	0.74	0	2
	2	0	0	0	0
	3	0	0	0	0
	4	0.50	0.55	0	1
	5	0.17	0.41	0	1
	6	0.57	0.79	0	2
PolicyMap	1	0.57	0.53	0	1
	2	0	0	0	0
	3	0.33	0.82	0	2
	4	0.20	0.45	0	1
	5	0	0	0	0
	6	1.00	0.82	0	2
Proquest Statistical Datasets	1	0.13	0.35	0	1
	2	0	0	0	0
	3	0.33	0.82	0	2
	4	2.00	0	2	2
	5	0	0	0	0
	6	0	0	0	0
Reference USA	1	0.20	0.45	0	1
	2	2	0	2	2
	3	2.00	0	2	2
	4	2	0	2	2
	5	0	0	0	0
	6	0.33	0.71	0	2
SimplyMap	1	0.14	0.38	0	1
	2	0.29	0.49	0	1
	3	0.67	1.03	0	2
	4	0	0	0	0
	5	0.50	0.55	0	1
	6	0.43	0.79	0	2
Social Explorer	1	0.29	0.49	0	1
	2	0.33	0.52	0	1
	3	0	0	0	0
	4	0.17	0.41	0	1
	5	0	0	0	0
	6	0.14	0.38	0	1

Table 3. Descriptive statistics for task times.

Application	Task	Mean	SD	Min.	Max.
ESRI Business Analyst Online	1	126.59	90.69	36.63	305.83
	2	20.31	12.52	8.15	34.44
	3	7.32	2.70	5.25	12.60
	4	63.73	36.96	23.67	125.50
	5	28.73	25.95	8.50	78.33
	6	70.23	51.87	22.87	177.77
PolicyMap	1	80.23	52.49	22.50	165.90
	2	15.41	11.31	5.27	30.79
	3	17.16	21.84	3.54	61.07
	4	37.34	30.61	13.84	85.14
	5	17.96	12.41	7.72	38.66
	6	96.75	69.71	31.05	186.23
Proquest Statistical Datasets	1	41.55	39.41	9.66	127.02
	2	15.50	14.38	7.17	32.11
	3	32.41	16.86	3.72	51.62
	4	49.53	28.98	31.01	82.92
	5	12.76	5.54	9.18	19.15
	6	29.85	7.18	17.38	38.84
Reference USA	1	73.31	48.21	22.17	133.68
	2	11.20	0.06	11.15	11.24
	3	27.92	7.02	22.95	32.88
	4	25.09	1.06	21.11	29.54
	5	17.37	2.97	14.12	20.65
	6	45.65	17.84	20.47	71.48
SimplyMap	1	60.67	27.63	34.52	113.04
	2	29.92	32.61	6.47	80.85
	3	19.73	13.68	6.90	35.63
	4	26.98	17.34	13.56	50.32
	5	72.47	34.17	31.92	114.98
	6	61.94	35.51	27.98	128.10
Social Explorer	1	117.92	96.96	24.98	297.07
	2	21.70	11.69	6.05	36.82
	3	10.52	7.71	4.30	22.08
	4	47.16	64.15	12.36	176.51
	5	14.94	3.01	9.85	17.74
	6	39.03	18.57	21.37	68.20

Table 4. Descriptive statistics for number of help/prompts.

Application	Task	Mean	SD	Min.	Max.
ESRI Business Analyst Online	1	0.38	0.52	0	1
	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0

	5	0.17	0.41	0	1
	6	0	0	0	0
PolicyMap	1	0	0	0	0
	2	0	0	0	0
	3	0	0	0	0
	4	0.20	0.45	0	1
	5	0	0	0	0
	6	0.43	0.79	0	2
Proquest Statistical Datasets	1	0	0	0	0
	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0
	5	0	0	0	0
	6	0	0	0	0
Reference USA	1	0.20	0.45	0	1
	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0
	5	0	0	0	0
	6	0	0	0	0
SimplyMap	1	0	0	0	0
	2	0.14	0.38	0	1
	3	0	0	0	0
	4	0	0	0	0
	5	0.50	0.55	0	1
	6	0.14	0.38	0	1
Social Explorer	1	0.29	0.49	0	1
	2	0.17	0.41	0	1
	3	0	0	0	0
	4	0	0	0	0
	5	0	0	0	0
	6	0	0	0	0

Table 5. Descriptive statistics for number of errors.

Application	Task	Mean	SD	Min.	Max.
ESRI Business Analyst Online	1	1.75	1.39	0	4
	2	1.50	1.52	0	4
	3	0.33	0.82	0	2
	4	1.33	0.82	0	2
	5	1.17	1.17	0	3
	6	1.71	1.38	0	4
PolicyMap	1	2.14	1.35	0	4
	2	0	0	0	0
	3	0.50	0.84	0	2
	4	0.80	1.30	0	3

	5	1.00	1.10	0	3
	6	3.43	2.82	0	8
Proquest Statistical Datasets	1	1.25	1.16	0	3
	2	1.00	1.00	0	2
	3	1.50	1.22	0	3
	4	0.33	0.58	0	1
	5	1.00	1.00	0	2
	6	1.14	1.07	0	2
Reference USA	1	1.20	0.84	0	2
	2	0.50	0.71	0	1
	3	0.50	0.71	0	1
	4	2.00	0	2	2
	5	0.50	1.00	0	2
	6	1.33	1.12	0	3
SimplyMap	1	1.43	1.27	0	4
	2	1.14	1.21	0	3
	3	0.83	0.98	0	2
	4	1.00	0.82	0	2
	5	2.17	1.33	1	4
	6	1.43	1.13	0	3
Social Explorer	1	3.86	2.91	0	8
	2	0.83	1.17	0	3
	3	0.43	0.79	0	2
	4	0.67	1.21	0	3
	5	0.20	0.45	0	1
	6	0.71	1.50	0	4

The descriptive statistics for SUS total scores are shown in Table 6 where higher score means higher overall usability rating. The range of rating for each statement of the SUS questionnaire is 1 (*strongly disagree*) to 5 (*strongly agree*). Participant ratings for statements 2, 4, 6, 8 and 10 were reversed (i.e., changing original rating of 1 to 5 to make the direction of rating consistent with other statements) so as to calculate the SUS total score. The range of the SUS total score is from 10 to 50.

Table 6. Descriptive statistics for SUS total scores.

Application	Mean	SD	Min.	Max.
ESRI Business Analyst Online	31.00	8.16	19	42
PolicyMap	33.43	7.68	24	43
Proquest Statistical Datasets	35.29	4.35	29	42
Reference USA	37.13	7.02	24	46
SimplyMap	40.14	4.81	30	45
Social Explorer	37.00	5.29	32	45

2. Usability issues

We reviewed participant comments and behavior notes and cross-referenced them with the quantitative response measures. In order to identify potential usability issues from the evaluation results, we applied the usability principles proposed by Dix et al. (2003). They proposed a set of principles that should support usability in human-computer interaction, including: (1) learnability, the ease with which new users can begin effective interaction and achieve maximal performance; (2) flexibility, the multiplicity of ways the user and system exchange information; and (3) robustness, the level of support provided to the user in determining successful achievement and assessment of goal-directed behavior. Each of these high level principles incorporates various factors. We list the factors relevant to this study in Table 7. Our analysis of the qualitative results focused on violations of these factors and their impact on participants' task performance. In this section, we first discuss specific usability issues of each application, and then we summarize design characteristics of web-based mapping applications to support usability.

Table 7. Principles for identifying usability issues.

Principle	**Factor**	**Definition**
Learnability	Predictability	Support for the user to determine the effect of future action based on past interactions.
	Synthesizability	Support for the user to assess the effect of past operations on the current state.
	Familiarity	The extent to which pre-existing knowledge and experience can be applied to the new system.
	Consistency	Conformity in input-output behavior between similar situations or task objectives.
Flexibility	Customizability	Modifiability of the interface and information shown on the interface by the user or the system.
Robustness	Recoverability	Ability of the user to take corrective action after making changes to the system.
	Responsiveness	User perception about the rate of communication with the system.
	Task conformance	The degree to which the system supports all the tasks the user wishes to perform and in the way that the user understands them.

2.1. ESRI Business Analyst Online

Participants seemed to have difficulty finding data or beginning to create a customized map, which caused the highest average task score (0.63) and longest average task time (126.59 seconds) for ESRI among the six applications. The correct link for creating a customized map is the "Create Color-Coded Map" in the toolbar (Figure 1), but its location does not provide a strong cue to participants that they should start there. Participants spent time checking other areas of the interface and asked for more help for this task than other applications (average number of help/prompts is 0.38).

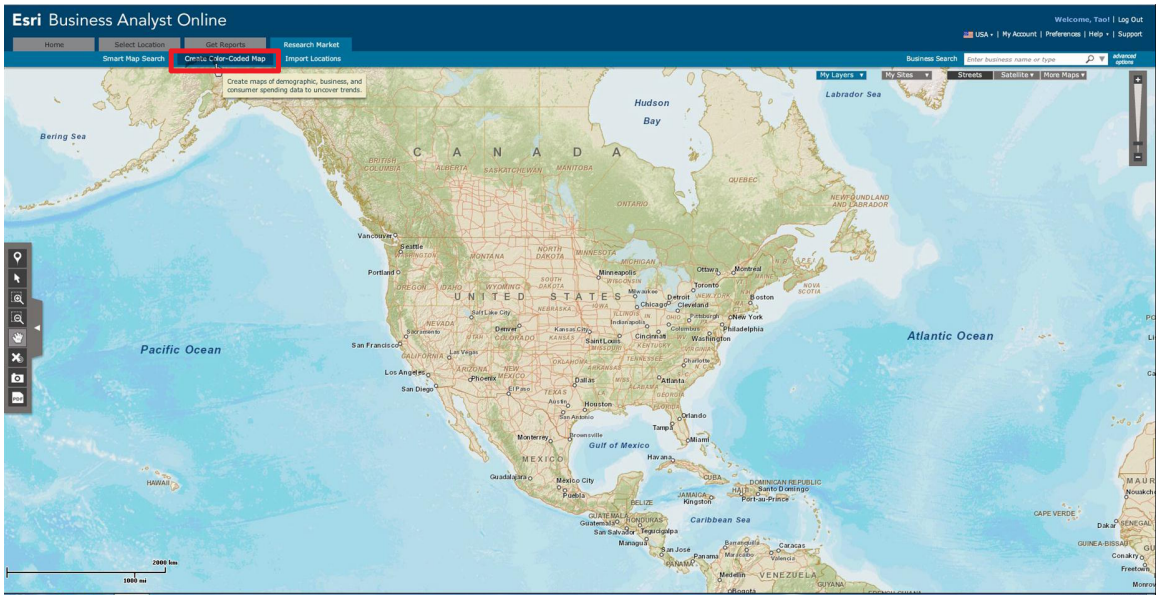

Figure 1. ESRI interface lacks an obvious starting point. Its 'Create Color-Coded Map' is too subtle.

Participants had the highest average task scores and longest average task time for editing the data range in ESRI. This is likely to be caused by the dialog design for editing data range (Figure 2). Participants are supposed to drag and drop the cut-off points or enter the value in the textboxes shown in the diagram to adjust the data ranges for different color shades on the map. The diagram covers the complete range of data and the maximal value of the data is large, but the cut-off points are presumably based on percentiles of the data. Therefore, those cut-off points and textboxes are often shown very close to each other and at the very top or bottom of the dialog, making the selection of the points and data entering difficult.

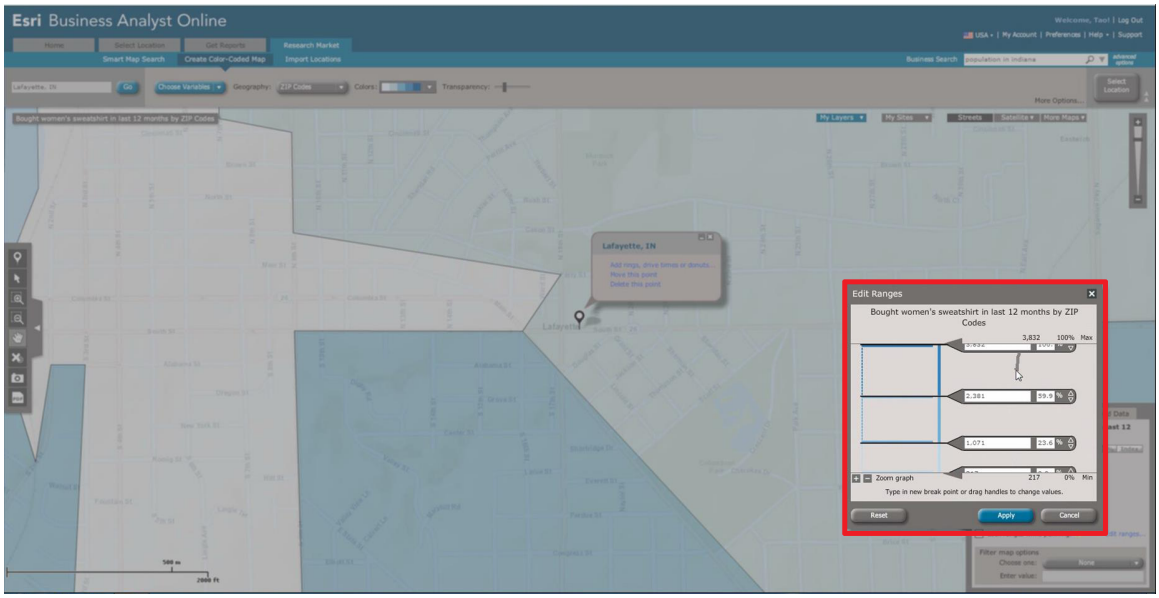

Figure 2. ESRI's interface uses drag-and-drop mechanisms to edit data range. Note the top and bottom cut-off points obscured by the boundary of diagram.

When asked to export the map, the majority of the participants chose to export the map as a PDF file. It took a considerable amount of time for the ESRI system to generate the PDF file. When the PDF file is generated, the system shows a dialog with a download link for the PDF (Figure 3). Participants tended to ignore the download link; they simply clicked on the OK button and closed the dialog. It was difficult for participants to find the PDF in the system if they closed the dialog.

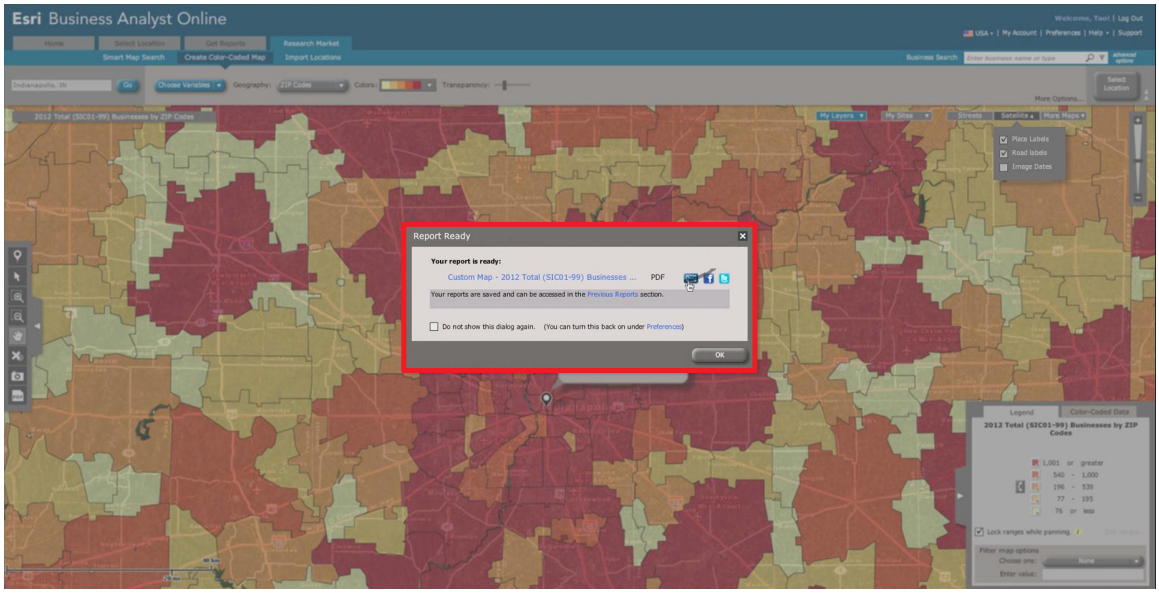

Figure 3. ESRI's export is slow, intervenes work and requires too many steps from the user.

There are some interface design issues that also affected the system's usability. For example, the color shades are overlaid on top of the map, making it difficult to read text labels of the map such as city and street names. The system has an option to change the transparency of the color shades, but it is often difficult for participants to achieve a balance between the colors and text labels. Another case is the select location button on the toolbar appears more prominent than the location search box (Figure 4), which drew participants' first attention when they were asked to find a location on the map. When the user submitted a location search for cities (e.g., searching for Provo, UT), the map was zoomed in too much and participants had to zoom out to see the city area. Regarding the map itself, one participant commented that it was difficult to tell the state's boundary on the map, especially when the map is showing data by counties.

Figure 4. ESRI's interface hides location search in subsections, while offers 'Select Location' as an obvious tab and button.

2.2. PolicyMap

The map display can show both data layers (thematic maps with values aggregated to various shaded areas or geographies) and sites (point datasets, like addresses), with interface controls on the top (for data layers) and left (for sites) of the map display (Figure 5). Two participants said having two menu bars (top and left) caused confusions to them when creating a customized map. Several participants commented that the data layers shown on top of the map display contain too much information. Some data categories have very long dropdown menus, which are difficult for visual search (Figure 6).

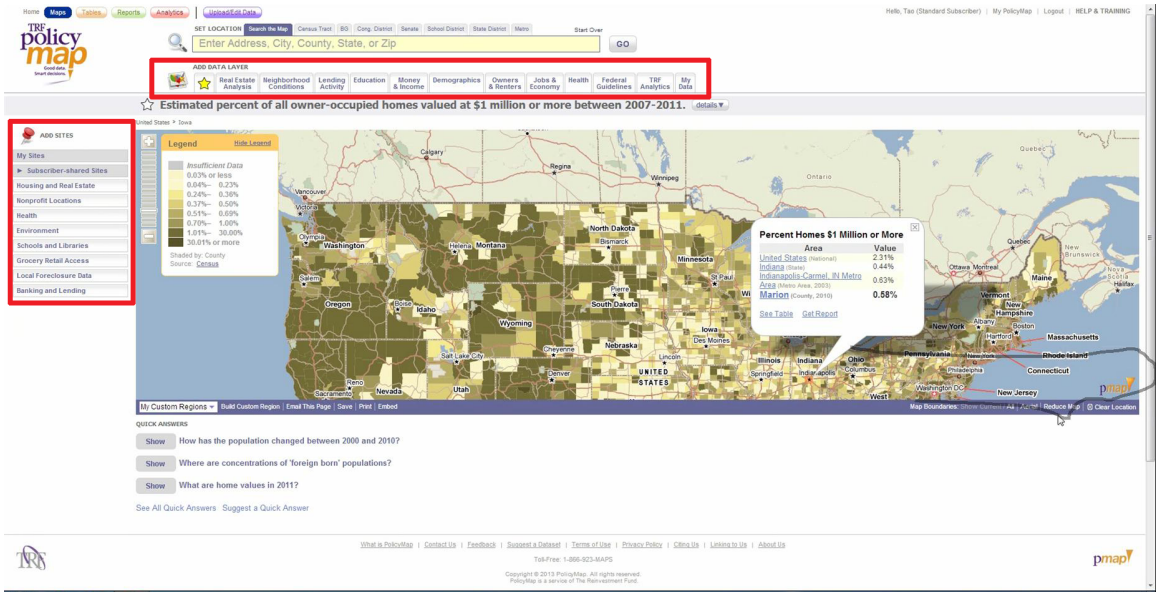

Figure 5. PolicyMap interface offers two menus (for data and location sites), which can confuse users.

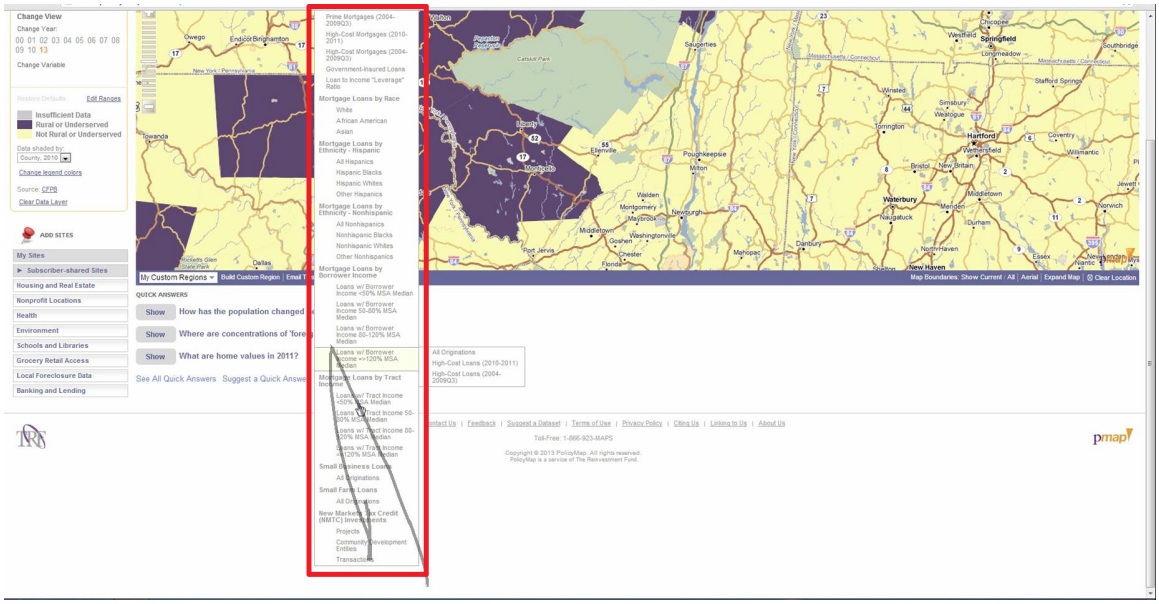

Figure 6. PolicyMap's drop down menus are so long that users have to scroll them.

There is a "Save" link below the map display area and most of the participants first noticed it when exporting the map. However, this link is for saving the current map in the system, not actually exporting the map as a downloadable file (Figure 7). The option for exporting map as PDF file is included in the "Print" link. It takes some time for the system to generate the PDF file. When the PDF file is ready, the system shows a notification at the upper left corner of the map display (Figure 8). A number of participants did not notice that notification because of its subtleness. These issues led to the highest average task score (1) and longest average task time (96.75 seconds) for exporting map among all applications.

Figure 7. PolicyMap confuses users with two export options: a) 'Save' (to the system) and b) 'Print' (as PDF/on paper).

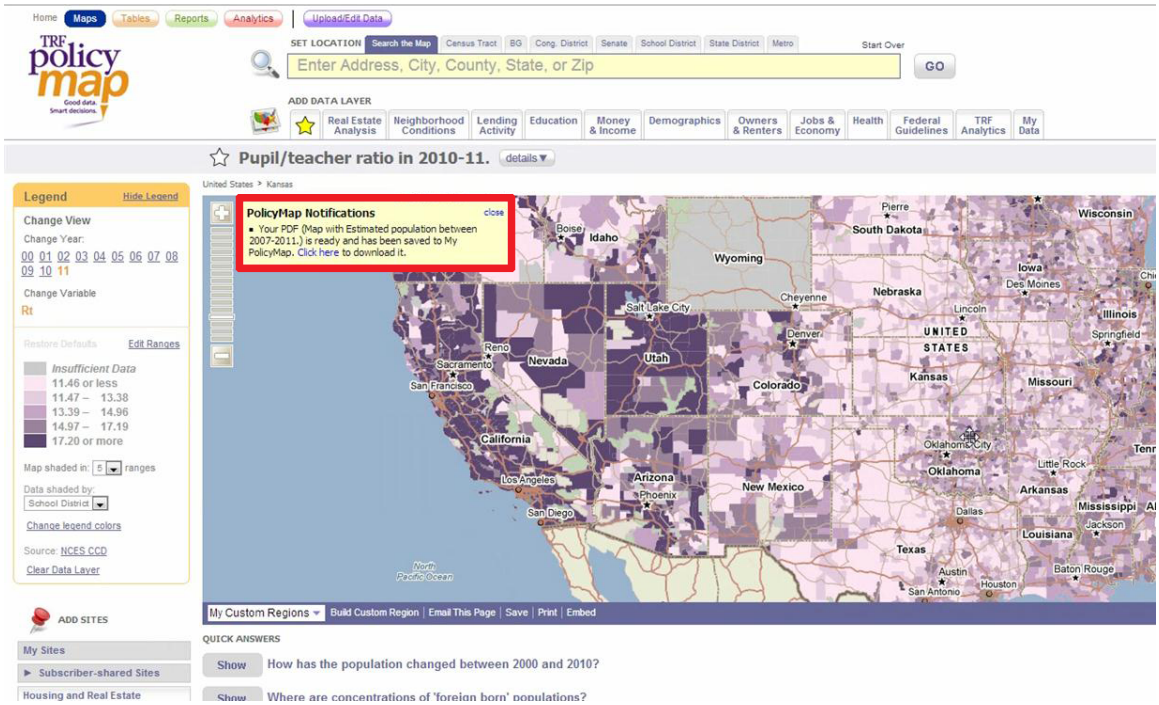

Figure 8. PolicyMap's subtle feedback about exported map image being ready for download is often unnoticed.

Overall, participants thought the text on the interface is small and thus it was not easy to find the options for different tasks. A number of participants mentioned that the map display is relatively small compared to other applications. The system has an option for expanding the map display area; however, with the expanded map participants could not see other useful components of the interface such as data layers, color, and data range (Figure 9).

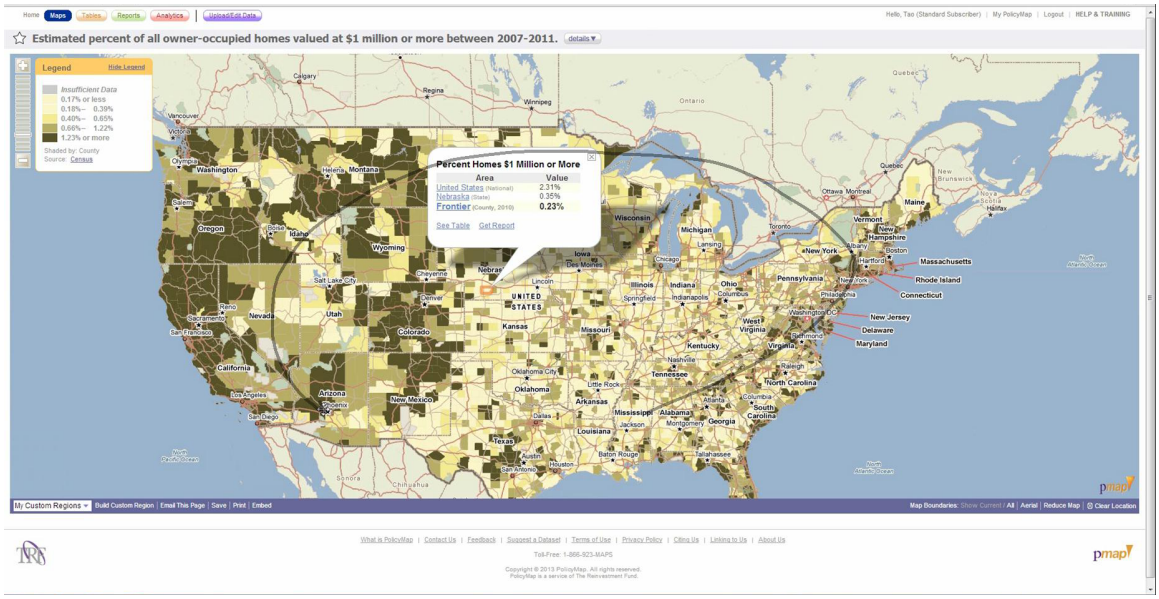

Figure 9. PolicyMap's enlarged map view eliminates useful customizability options from user's view and access.

2.3. Proquest Statistical Datasets

By default the system shows the data in a chart form. Participants had difficulty locating the button to show the data in a map (Figure 10). Participants also noticed that not all datasets available in the system could be displayed on the map. The system does not provide explicit indication of whether a dataset can be displayed on the map. Some datasets are limited at the state level, making the map viewable only at the state level.

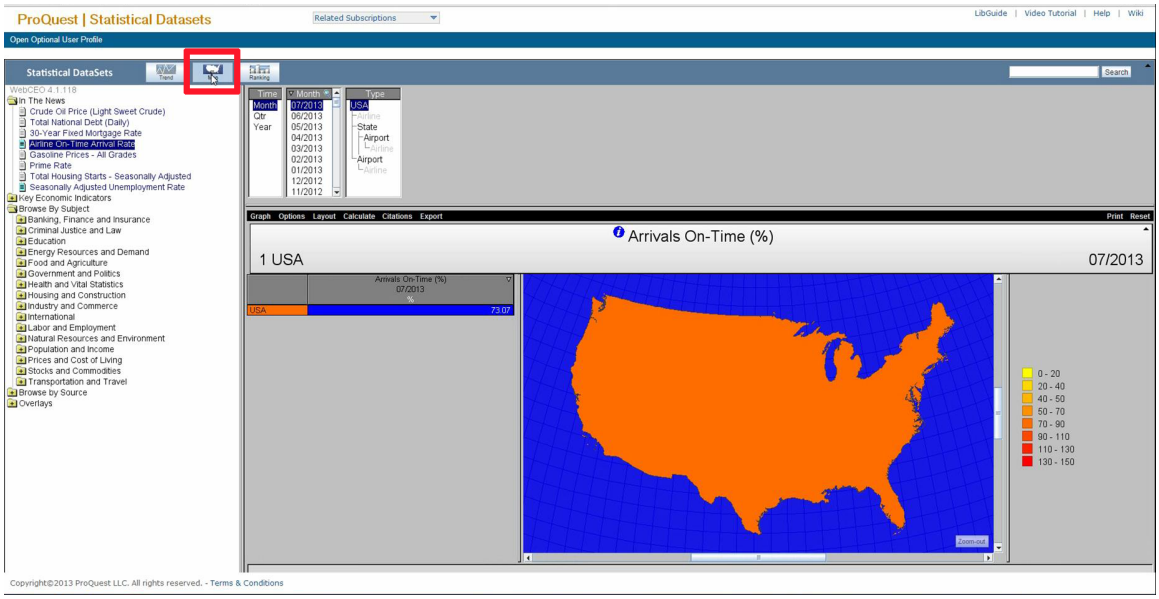

Figure 10. Proquest interface focuses on data, offering the map view as an additional option.

Contrary to other applications export option was clearly displayed above the map in Proquest interface. The exported PDF file of the map includes the data table, but it does not show the complete list of states [Figure 11]. The system supports copying the map as an image to other applications like Microsoft Word, but it does not allow participants to directly save the map as image files.

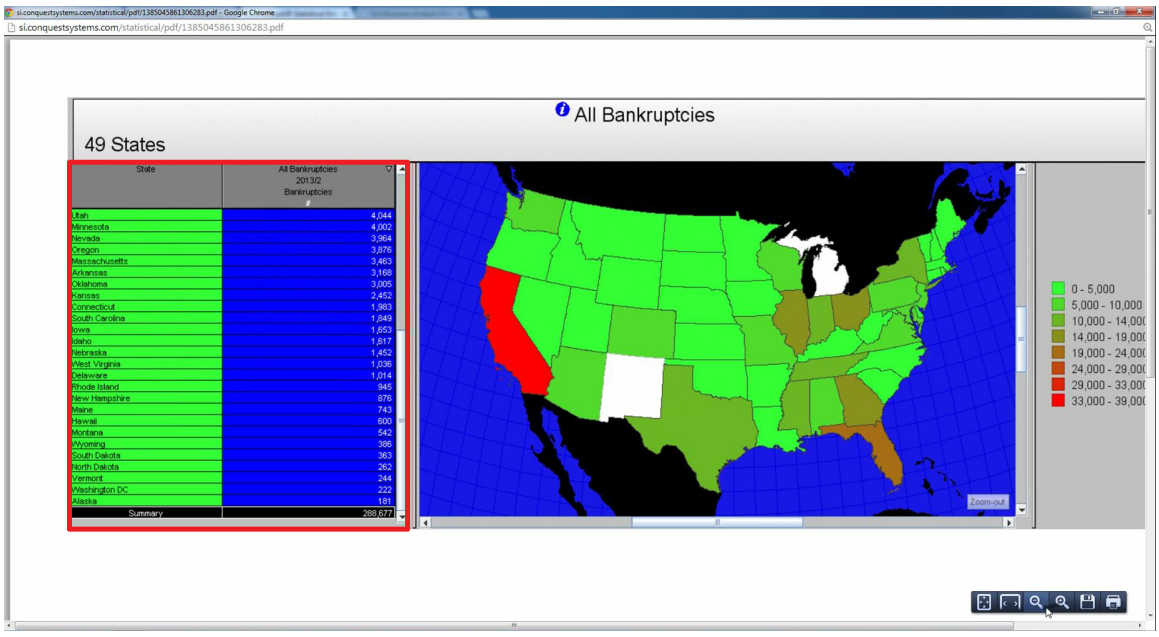

Figure 11. Proquest's exported image unnaturally cuts the appreciated data list.

A lot of participants commented that the map display is very limited in terms of the map quality and functions. The map shows the border of the states but does not show other common geographical information like state and city names, highways, and streets. If participants select a state, the map only shows the counties of that state. Unlike other applications, the map cannot show the whole country by counties. The map display allows zooming in to multiple levels; however, when participants tried to zoom out, the map display directly went back to the whole country level. The default colors of shades are too close to differentiate, especially when they are shown on adjacent areas on the map. Participants also mentioned that the response of the map display is noticeably slow.

Other usability issues include that the system does not allow changing data ranges for the color shades (average task score is 2). One participant commented that the options of the map display are in the middle of the interface and difficult to find. He expected those options to be at the top of the interface like in other applications.

2.4. Reference USA

Participants commented that the map display of Reference USA looks like Google Maps. They felt that Reference USA is very limited compared to other applications. It can only display one type of data (US business in this evaluation) on the map. The distribution of data points (location, number of employees, or sales volume) is shown as a heat map (Figure 12a) when the zoom level is high (e.g., the whole US). But if the zoom level is at city or below, the map shows individual locations (Figure 12b). Several participants did not understand this transition from heat map to individual locations. There is no option to change the map unit or the color (average task score for both tasks is 2). The map is very slow when loading data and changing zoom levels. It uses contrasting colors (green to red) to show the intensity of data, but the green

color is very close to the map background. Other applications use monochrome but distinctive colors for the shades on the map, which is more pleasing to view. Furthermore, the map area is small. Sometimes the mouse's scroll wheel can be out of the map's input focus and thus scrolls the page.

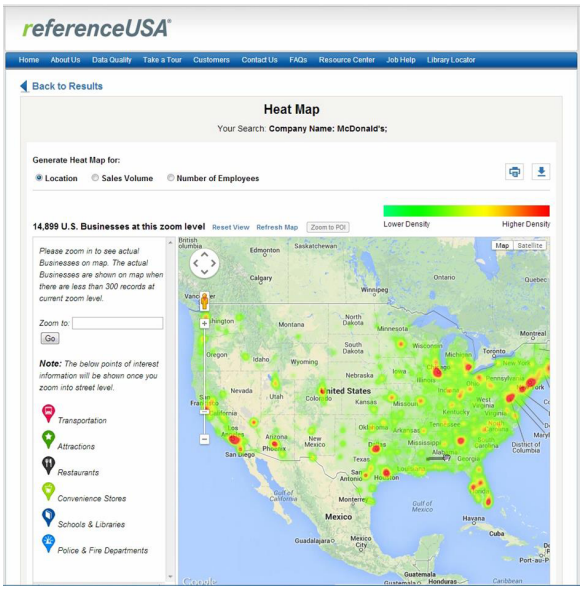

Figure 12a. Reference USA interface shows the heat map only when zooming out.

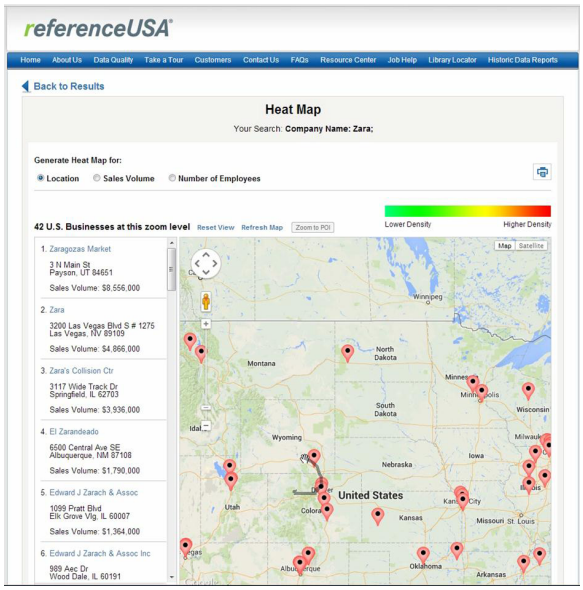

Figure 12b. It shows locations pins only when zooming in.

There is no dedicated export function in Reference USA. When asked to export the map, most participants selected the icon for printing the page (Figure 13). The exported PDF displayed the entire interface. Participants found the map clear to read and they appreciated the list of companies on the left hand side when saving a zoomed in crop of the map (Figure 12b).

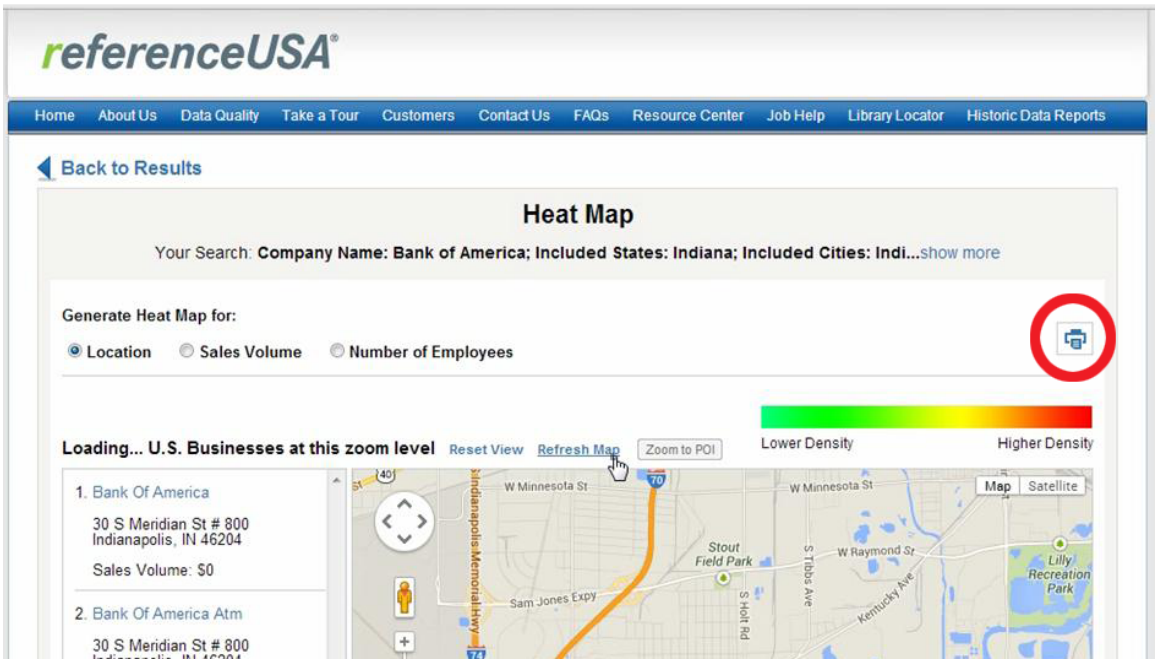

Figure 13. Reference USA interface offers only the 'Print' icon as the export option.

2.5. SimplyMap

The biggest issue for participants was that the interface did not provide a good cue for moving forward after they selected a dataset (or "variable" in SimplyMap). As shown in Figure 14, the only way for participants to go to the next step is to close the dialog of variables. The system shows a notification at the lower left corner of the dialog when the participant selected a variable. But this notification is very subtle and few participants noticed it. This was the major reason that participants took relatively long time (average task time is 60.67 seconds) to create a customized map. The seam issue happened when participants were asked to selection a location on the map (average task time is 72.47 seconds). Participants expected that after they selected the variable, the dialog would automatically close and the map would load the data.

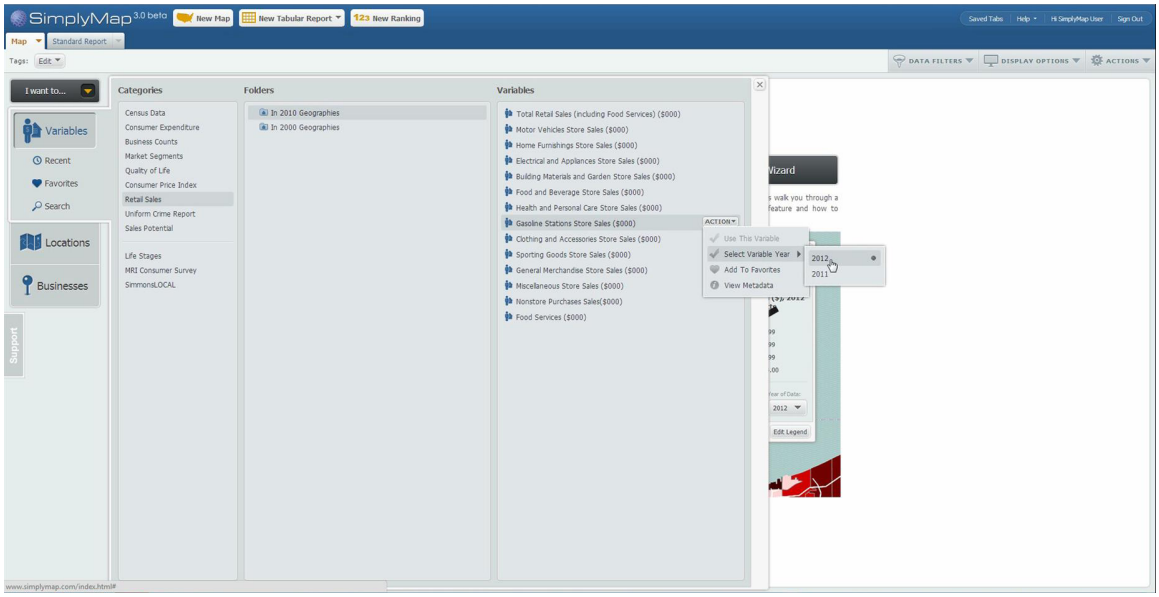

Figure 14. SimplyMap's interface lacked instructions on how to load the selected data on the map.

As a minor issue of the interface design, the option of changing map color is under the link of "Edit Legend" (Figure 15), which was confusing for several participants.

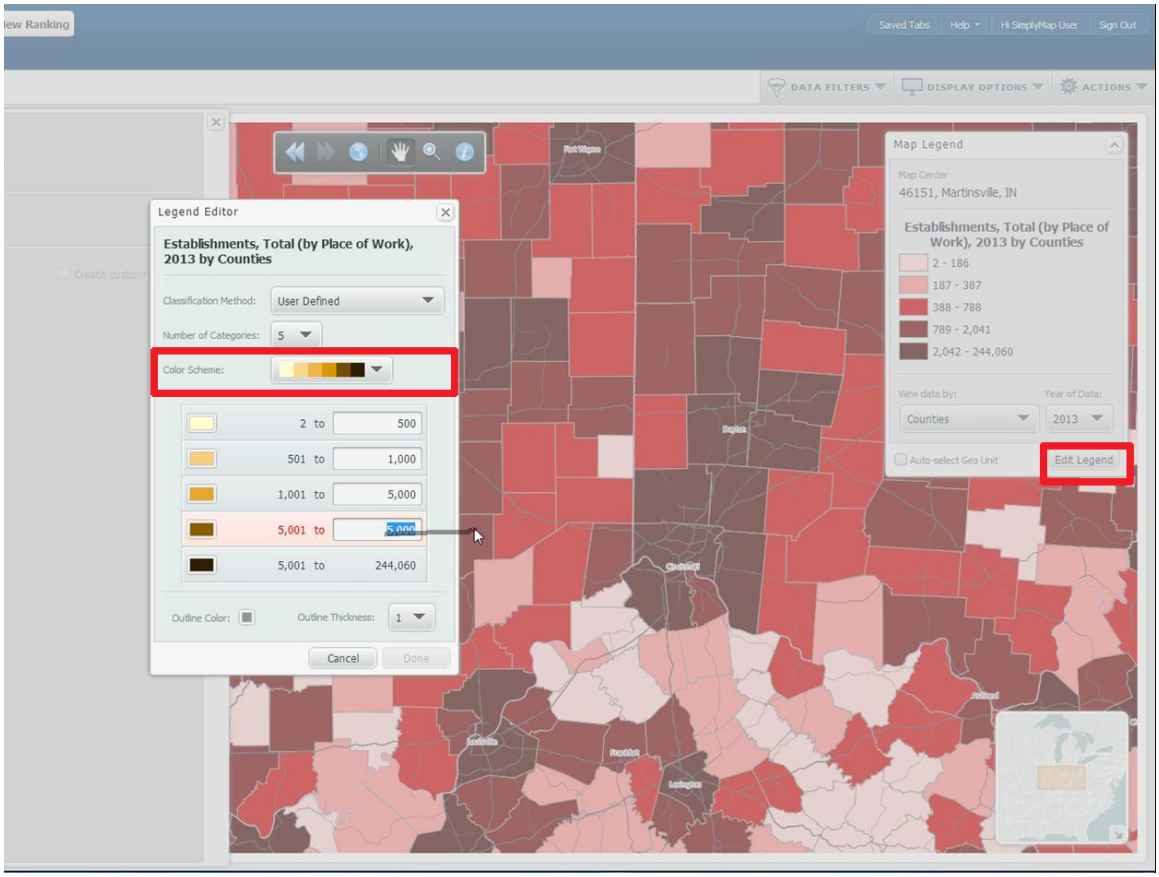

Figure 15. SimplyMap interface presented the option to change maps' colors under 'Edit Legend'.

Unlike other applications, SimplyMap does not have a search box for locations. Participants had to use the locations dialog (Figure 16) to find the location. In the dialog, they had to first choose state and then choose cities within the selected state. Few participants knew they could also type the first letter of state and city names to speed up the process.

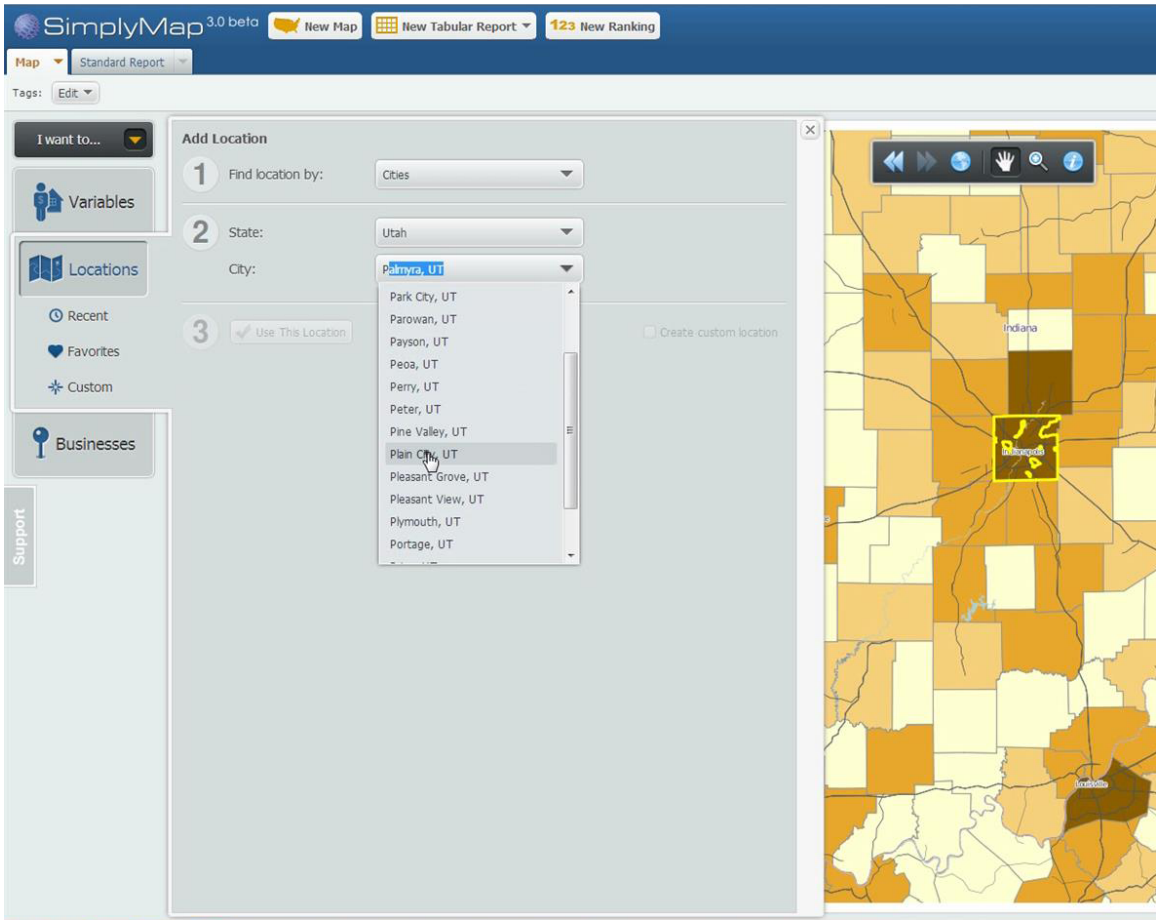

Figure 16. SimplyMap's interface has no location search, only a 'Location' dialog.

When exporting the map, the interface did not ask participants where to save the exported file. The average task time for exporting map is 61.94 seconds, part of which was spent on finding where the export file was saved on the computer.

2.6. Social Explorer

The average task time for creating a customized map with Social Explorer is 117.92 seconds, which is relatively long. Our further analysis showed that a few participants spent their time exploring the four different types of data visualization available in Social Explorer (dot density, shaded area, bubble, and histogram), which caused the longer task time. This should not be considered as a usability issue, but rather an indication that participants may need a tutorial of those data visualizations.

The interface has an “Export” button above the map display for exporting map. When participants click on that button, the interface shows a dialog for options like exported file size and format. After participants click on the “Prepare for download” button of the dialog (Figure 17a), the button changes to “Download” when the exported file is ready (Figure 17b). Most participants did not notice that change of button label. They expected to click the button of the export dialog and the system would ask where to save the exported file. They did not realize that they had to click the “Download” button again.

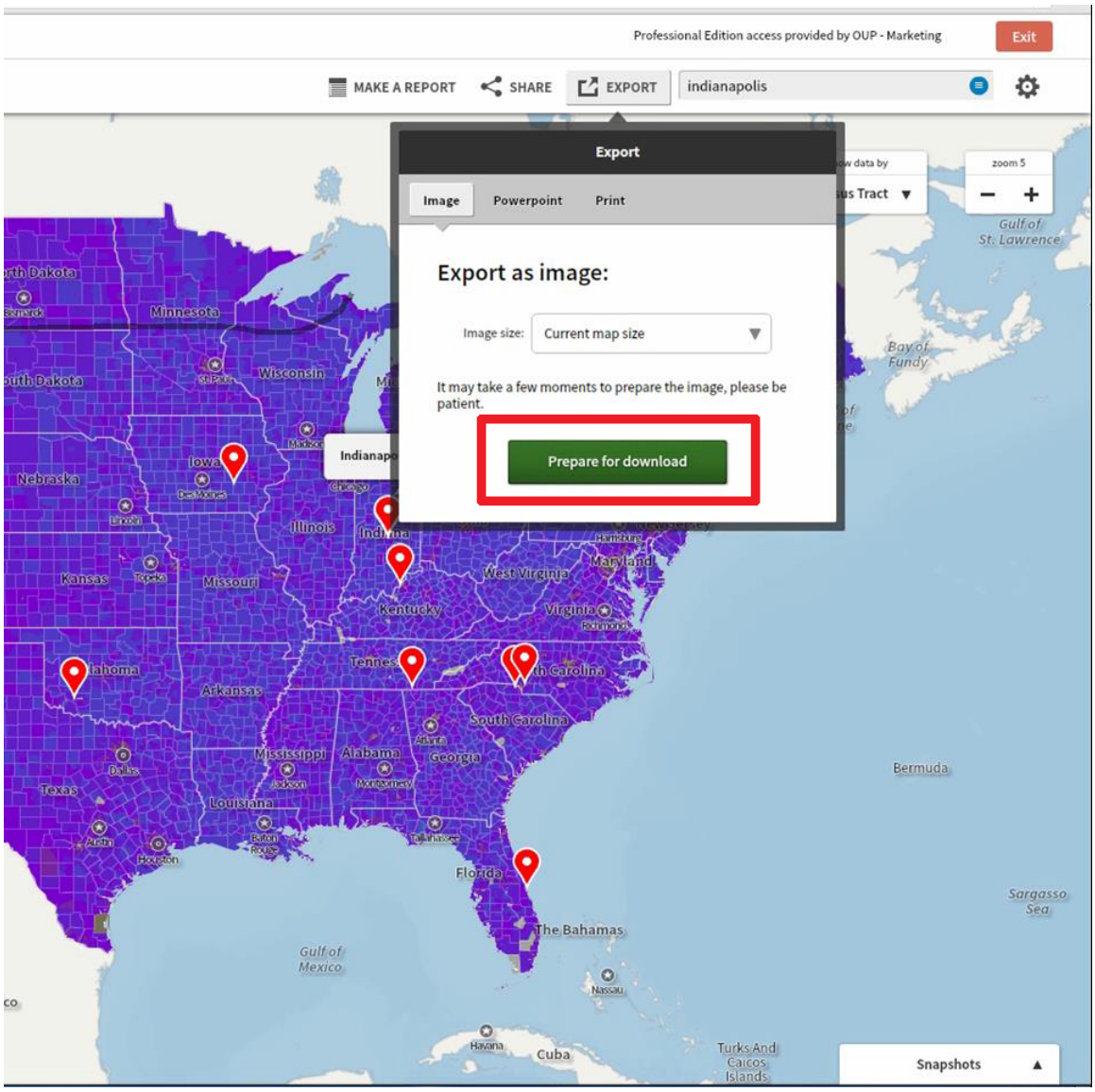

Figure 17a. The export dialog of Social Explorer with the “Prepare for download” button.

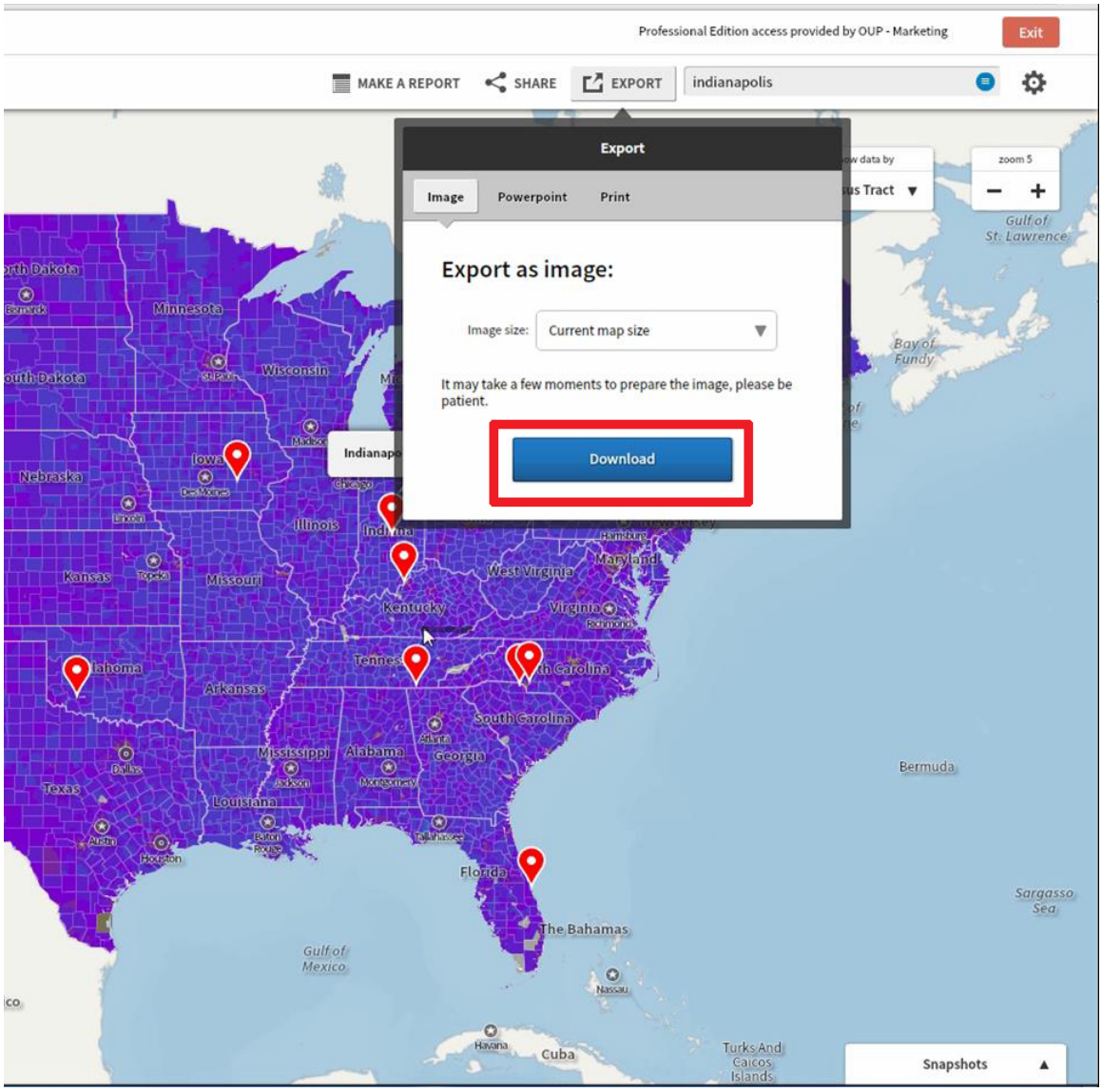

Figure 17b. The export dialog of Social Explorer with the “Download” button.

For changing the data range of color shades, Social Explorer’s interface (Figure 18) allows drag and drop of each cut-off points. A couple of participants found it harder to enter exact values than simply typing the numbers on other applications.

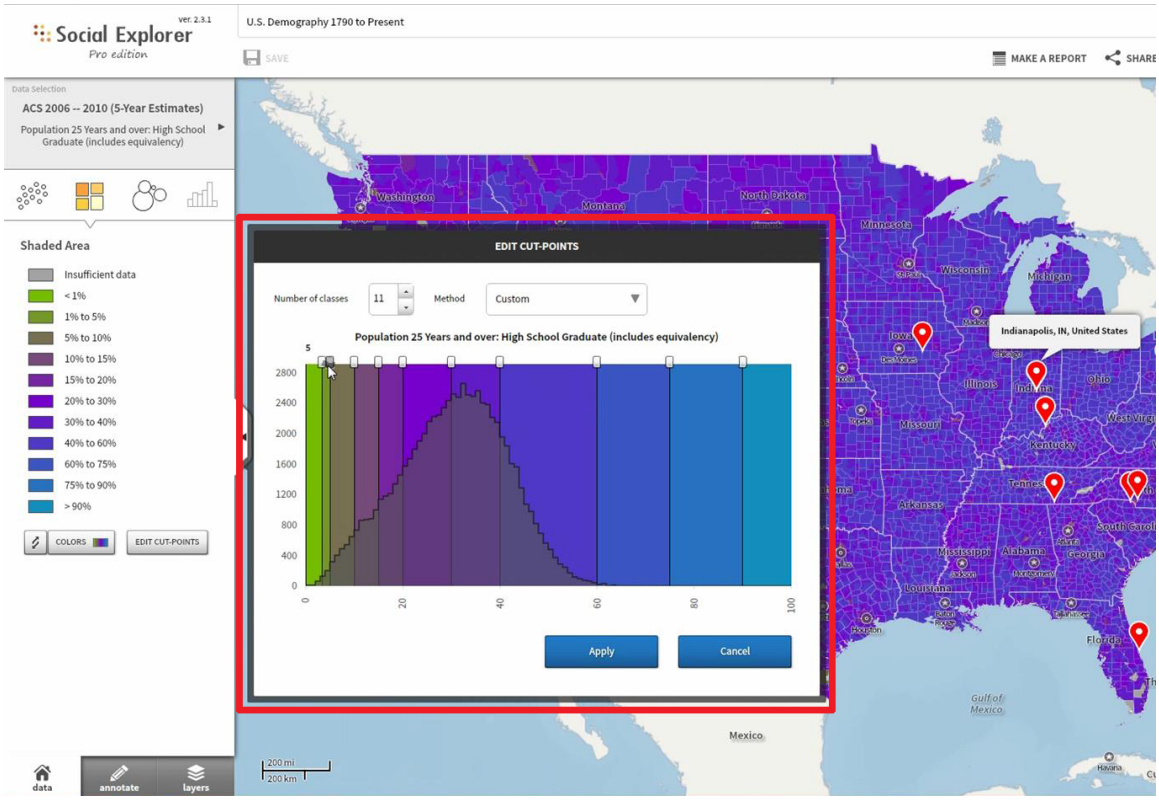

Figure 18. Social Explorer's drag and drop mechanism makes it hard for users to control the values to edit data range.

3. General evaluation

We asked participants which application they would prefer the most at the end of the evaluation sessions. SimplyMap and Social Explorer were mentioned six times each. The average SUS score of SimplyMap and Social Explorer was 40.14 (highest among the six applications) and 37.00 (the third), respectively. Reference USA has the second highest SUS score, but its data and map functions are very limited. Positive comments regarding SimplyMap include: distinguishable map colors; the variables, locations, and business tabs on the interface clearly show the action sequence for creating a map; the color shades are visually clear; the map is easy to export; and the system allows multiple tabs for different maps. Positive comments about Social Explorer include: very clear data categories; nice timeline for datasets; the map is clearly labeled and its interaction reminds of experience with Google Maps; different options easy to find on the interface; and the map layers have a lot of customizable options.

Participants offered positive comments to other applications as well. For ESRI, they liked the large, beautiful map, the ease of changing the map's colors, and the ability to change base map to other types (e.g. streets, satellite, and topographic). For PolicyMap, some participants liked the data layers directly laid out for easier visual search. PolicyMap shows city outline by default, which could be very useful in some situations. Participants liked the hierarchical structure of datasets in Proquest, but at the same time they felt it would take longer time to search for a

particular dataset than in other applications. Table 8 presents a summary of participants' comments on various aspects of the six applications.

< Insert Table 8 ("Summary table.docx") about here >

Discussion

In general, our evaluation has shown that participant task performance measures (task score and task time) are good indicators of potential usability issues. The high task scores for creating a customized map on ESRI revealed that participants had difficulty finding the right place to start. The export map task with PolicyMap has average task score of 1.0, which is caused by the confusing "Save" link (saving map to "My PolicyMap" but not exporting map). Two other high average task scores showed the limitations of applications: Proquest does not support changing data range; and Reference USA does not support changing map unit. Similarly, the average task time of 70.23 seconds for exporting map in ESRI revealed that most participants ignored the download link in the report ready dialog and they had to try exporting the map again. The average task time of 72.47 seconds for the changing location on the map task with SimplyMap suggested that its location search is inefficient. SimplyMap uses a drop down list for location selection instead of a more flexible location search box.

The number of errors made by participants provided additional evidence of the usability issues revealed by task score and task time measures. For example, the average 3.43 errors when exporting map with PolicyMap and 2.17 errors when changing location with SimplyMap were caused by the "Save" link in PolicyMap and no search box in SimplyMap, respectively. The number of errors also showed additional information about user behavior. Participants made in average 2.14 errors when creating a customized map with PolicyMap. A possible reason for this high number of errors is that PolicyMap has two toolbars (left and top). A number of participants were not clear whether they needed to start from data layers on top of the map or sites layer on the left. The average number of errors when creating a customized map with Social Explorer is also high (3.86), which is probably caused by participants exploring the different types of data visualizations on the map. For dot density and bubble visualizations, participants had to adjust the size of dots or bubbles to make the map legible. This process may involve additional trials and errors. Compared to the number of errors, the number of help or prompts requested by participants is less indicative of usability issues, as the data were relatively rare across tasks and applications.

In general, the response measures and our observations showed that participants encountered most difficulties and confusion during the first 'creating a customized map' task and the last 'exporting map' task. An important reason for this is the lack of appropriate guidance for new users in the applications (learnability is not well supported). The appropriate action sequence for creating a map was not communicated well to participants, and they were not able to relate their experiences with common websites (including Google Maps) to these applications. SimplyMap does have a prominent "Start Here" button to guide users, but it lacks a clear cue for

participants after they selected a dataset. Participants could have spent more time and efforts in order to understand the interface based on initial interactions. A common comment from participants was that given some time of learning and trials, they should be able to use the applications relatively well for the tasks we gave them. For exporting maps, some applications require time to prepare the files and thus their interfaces need to show notifications to the user when the exported file is ready. ESRI, PolicyMap, and Social Explorer have different notification mechanism and interface designs, but participants either failed to perceive important information (ESRI) or did not notice the signal when the exported map is ready (PolicyMap and Social Explorer).

Another important source of usability issues was the different interface layouts for map unit, color, legend and data range, and location search functions. Some applications like Social Explorer and SimplyMap have relatively good groupings of those options, while participants spent more time finding those options in other applications. Participants had difficulty understanding some of the terminology used in the applications, such as layer, color shade, legend, census tract, and census group. There is also a need to provide brief, understandable description of datasets as well as search functions for datasets and locations.

We rephrased the identified usability issues into prospective design suggestions: what a new GIS mapping system (1) must follow, (2) should follow, if possible, (3) could follow, for better user experience, and (4) must avoid at all times (to avoid usability issues). We used the usability principles identified in Dix et al. (2003): learnability, flexibility and robustness, to evaluate the level of each identified usability issue. An additional category, aesthetics, was added to include participant comments about visual features, such as map size, perceived beauty and usefulness, and legibility (text and color contrast). The results are presented in Table 9.

Table 9. Design characteristics for web-based mapping applications.

	System Must Have	**System Should Have**	**System Could Have**	**System Must Avoid**
Predictability	• Tooltips upon hover • Obvious start point • Obvious search box • Clear map zoom & drag	• System overview • One menu bar • Obvious help link • Obvious export	• Map immediately visible	× Barriers to entry × Cluttering the interface × Scattering options around the interface × Bury options too deep
Synthesizability	• Expanding data dialog • Distinguishable colors • Area boundaries on the map	• Group data into consistent tabs • List search results • Focus map to search results	• Responsive site (avoid scrolling leaks)	× Intrusive dialog boxes × Too subtle notifications × Hard to see data visualizations

Familiarity	• Understandable data categories • Big buttons for frequently used functions • One search box with filters • Commonly understood area boundaries, e.g. states, countries (recognition)	• Black text for active tabs • Timeline for selecting year • Right click to save map image • List results alphabetically • Google Map-like navigation	• Icons for tabs and data categories • Additional options bottom-eight	× System unusable without help × Unclear icons & labels × Adding steps (e.g. naming search) × Too much feedback (confuses user from correct action) × Unexpected functions, (e.g. saving to the system)
Consistency	• Clean & structured interface • Hover: additional information (e.g. tooltip, area info, zoom level)	• Zoom based mapping • Option for colorblind people		× Greens-reds for Colorblind
Flexibility/ Customizability	• One search option with filters • Change map colors • Edit data (e.g. year) • Precise data range controls (type in)	• Range of data • Grey out unavailable options • Clear labels for map areas (e.g. state, city, street) • Create own annotations • Customizable map legend • Many export formats (jpg, pdf, print, screen capture, share)	• Present multivariate data • Multiple tabs (multitasking) • Filter results after search • Edit export image (e.g. add data charts, notes • Many ways to visualize data on the map • Edit colors for data ranges	× Offering non-executable options (e.g. cannot map data) × Too many options (confusing)
Recoverability	• Undo & Redo buttons	• Auto-Save • Easily re-frame map to any location • Real-time editing	• Tick/Un-tick v options	× Spelling sensitive search × Inaccurate drag scales × Text based help
Responsiveness	• Delay feedback • Grey unusable options	• Ask where to export • Zoom to results after search • List results to select from •	• Agile zooming	× Scroll leaks from map to web site
Task Conformance	• Quality export map that is immediately presentable (clear) • Location pins visible at all zoom levels	• Option to export to the computer • Up to date data (2013)	• Link to external sources (e.g. LexisNexis)	× Data overlaying map details × Constrained data range editing × Capturing the entire interface when exporting × Illegible export map

Aesthetics	• Large map (legible) • Useful map based visualization	• More map than numbers (visual over numeric) • Monochrome (pleasing) • Clear data location pins (findable)	• Beautiful • Quality export • An experience	× Intimidating interface × Unappealing interface × Too small elements (text, labels) × Too many contrasting colors on the map (overwhelming) × Too little color contrast (hard to differentiate)

Most important, especially for new users, is system learnability, the degree to which a system is easily understandable and usable for a new novice user. A system design must offer instantaneous usability to the user. Having a clear and structured interface with obvious elements, such as a clear 'start here', 'export map' and 'edit' options, are vital for a user to recognize what they can do in a system. Offering obvious mechanisms that the user is already familiar with in the online domain, like 'search box with filters' and 'site help', is equally important for the user to know what they can accomplish and how. For system specific functions, offering quick help in the form of tooltips can benefit the user by allowing him to problem solve on the go and accomplish their task. Familiar icons and labels that the user can understand based on their past web experience can effectively reduce the system learning curve. It is important to avoid entry barriers, which may lead to the user to leave the system without accomplishing their task.

Good systems guide their users through its structure. Offering one clear menu (instead of two or more) is an obvious way to reduce options for the user. Good guidance involves visualizing user steps, e.g. expandable dialogs, allowing the user to gain an overview of the system and how the options have been structured. Guiding the user to appropriate action sequence must take as few steps as possible. When a system is slow, feedback must be provided, however, intrusive dialogs that can intervene work, or which involve too much feedback about the system progress (e.g. prepare report changing to report ready) must be avoided. When designing help function, avoid taking more than 30 seconds of users' time. Step-by-step video (or other visual formats) is most effective for learning. The help should offer an overview of the system's options, introduce the search function and indicate how the system gives feedback (e.g. how does a notification look like).

Visual cues are important for guiding the user in a subtle manner. Frequently used functions must be presented big and clear, while less used functions should be grouped into consistent tabs. A system should always seek to adopt generally understood mechanism such as timeline for year selection, filters for a search box, alphabetical listing, black for active and grey for inactive options, etc. For instance, right clicking on the map to export it as an image, is likely to be an intuitive online behavior due to other platforms using the same mechanic (e.g. Google Images). It is important to avoid cluttering the interface with too many options, scattering the options around the interface and hiding functions too deep into sub-folders. This makes things hard to find and would exhaust the user with too many steps.

Customizability has important implications for users' perception of a system's usefulness. The system design must account for different users' goals and support their varying needs. To this end, it must allow the user to start from different points of interest, e.g. location, business, data. The user must be able to change the map's colors and edit data parameters accurately but flexibly. The system should offer access to a wide range of data, but grey out any options that are not executable. Ideally, it should have a customizable map legend, allow user to add personal annotations to the map and offer varying export options. Exporting the visualized map is one of the key goals for a GIS system user. The system should facilitate export options beyond pdf export and printing, e.g. saving the map as an image, allowing the user to insert it into other documents; a screen shot has a small file size, and sharing the map online could allow easy peer-to-peer sharing. The system could also allow the user to pick and choose what details will be on the exported image (e.g. labels, colors, additional data charts). While the system could be very flexible by allowing the user to open many tabs for different maps, change how data is visualized on the map or allow multiple data sets to be visualized concurrently, it is important to avoid giving too many options.

System robustness (recoverability, responsiveness and task conformance) determines how effective user's experience with it is. Any software must allow user to recover from mistakes (e.g. undo, redo), offer feedback when processing user's requests, eliminate any options that cannot be executed (to avoid task failure) and allow the user to accomplish their task as close to their terms as possible. Ideally, a system would save user's work upon crashing and allow them to resume to their work where they left off (with minimal damage). In similar vein, adjusting data range, framing the location on a map and ticking/un-ticking options to display on the map, should be equally available options for users to effectively accomplish their tasks. A responsive system design would prompt users when they expect it. For instance, when exporting a map the system should ask the user where to save it and do it automatically once the system is ready with generating it. To prevent fragility in usability, avoid spelling sensitive search, inaccurate drag and drop editing scale and illegible maps.

Aesthetically, a good GIS system must have a big legible map where data is visualized in monochrome but easily distinguishable colors. It should be a pleasant visual experience with natural (familiar) navigation that requires no learning. Zooming should be smooth and dragging agile. To make the map immediately legible the system should offer clear contour lines and labels for defined areas such as countries, states, cities and streets. For instance, state outlines are very familiar shapes for Americans and therefore clearly understood, while city outlines are generally less understood, but may be important for understanding the data visualized on the map. Any additional information about an area should be initially hidden to avoid cluttering the map. It should be made easily available upon hovering the area via a small pop up window. The map could be custom designed for aesthetic beauty and offer interactive elements that ignite wonder in users, encouraging them to explore more. All aesthetic decisions made must avoid jeopardizing legibility. Avoid small pale text and too many colors.

Conclusion

In order to effectively evaluate usability of web-based mapping applications and ensure user performance and ease of use in critical tasks, there is a need for a structured assessment approach and usability metrics. We have previously identified a user-centered design process for developing research repositories (Zhang et al. 2013). Here, we propose a set of metrics for assessing usability and identifying potential usability issues for web-based mapping or similar applications. As shown in Table 9, this metrics include testing tasks, task performance measures, participants' comments, and observation notes.

The usability testing tasks should have at least three types of tasks: create customized maps, adjust map options, and export map. Table 10 lists major steps for observing each type of task. Based on our analysis, task successfulness, task time and number of errors are effective measures of task performance. As usability issues generally have impact on task performance, any unusual patterns of task performance should be further examined with the guidance of usability principles. We also list aspects of participants' comments relevant to usability assessment. Together, the task performance measures, participants' comments and observation notes should provide strong evidence for identifying usability issues and suggesting design improvements.

Table 10. Usability metrics for web-based mapping applications.

		Observation notes
Testing tasks	Create customized maps	Getting started Select dataset Select location
	Adjust map options	Map unit Color Data range Map label
	Export map	Select export options (file format and size) Notification mechanism
Task performance measures	Task successfulness (score) Task time Number of errors	
Participant's comments	Initial impression Ease of use Options, toolbar, menu Zoom Map legibility Export Help Aesthetics	

References

Brooke, J., 1996. SUS-A Quick and Dirty Usability Scale. In P. W. Jordan et al., eds. *Usability Evaluation in Industry*. London: Taylor & Francis.

Dix, A. et al., 2003. *Human-Computer Interaction* 3rd ed., Prentice Hall.

Zhang, T., Maron, D.J. & Charles, C.C., 2013. Usability Evaluation of a Research Repository and Collaboration Web Site. *Journal of Web Librarianship*, 7(1), pp.58–82. Available at: http://dx.doi.org/10.1080/19322909.2013.739041 [Accessed September 17, 2013].