Hood Hunter: A House Hunter's Guide to Narrowing Neighbourhoods

Abiramy Kuganesan, Ivan Song, and Lufei Liu

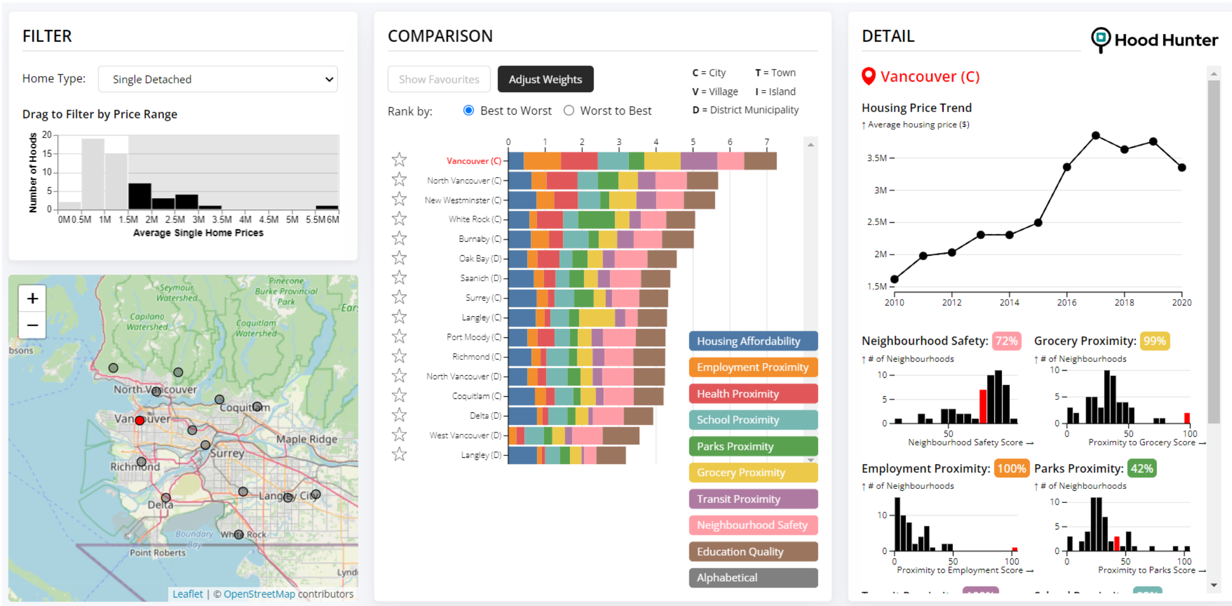

Fig. 1. The Hood Hunter exploration tool is accessible as a web application. To learn more, visit https://ivansong3.github.io/hoodhunter/.

Abstract— This work provides an overview of Hood Hunter, an explorer visualization tool to support prospective home buyers in narrowing neighbourhoods for a potential home purchase. In order to validate the need for such a tool, a survey was conducted. The main objective was to understand the core criteria that BC residents evaluate when picking the perfect neighbourhood for their residential housing purchase. Grounded by these considerations, and prospective home buyers' current task flow, the iterative design process behind this explorer visualization tool is presented. Upon analysis of the data and task abstractions presented in this work as well as synthesis of the pros and cons of preliminary proposed solutions, a wireframe of the final solution was established. This prototype was implemented using D3 Observable, HTML, and CSS. It is presented as a web application for users to explore. Early stage evaluation of the visualization tool was performed by 5 participants. Feedback from participants validate the usefulness of such a tool and potential extensions for future work.

Index Terms—house affordability, neighbourhood search

1 Introduction

In recent years, several urban centres across Canada have faced challenges with affordable housing. These trends have been exacerbated over the course of the pandemic with home buyers looking to move out of densely populated regions to nearby sub-urban neighbourhoods to support newly emerging remote work cultures.

Several factors have contributed to the unafforadable conditions of the Canadian residential real-estate market. Notably, the presence of foreign buyers in the Toronto and Vancouver real estate markets has contributed significantly to the current housing crisis incurred by local residents [12].

Although the house search process is unique to every buyer, certain recurring classes of requirements have been shown to be persistent across buyers and varying demographics. Olanrewaju and Wong conducted a study to determine key criteria that home buyers considered when purchasing a home [21]. Through an exploratory factor analysis, they were able to identify and rank clusters of requirements according to their impact on buying decision. These clusters include price configuration, accessibility, transportation, and sustainability.

Prospective home buyers spend a large amount of time and resources researching neighbourhoods for proximity to city centres, low crime rates, quality schools, and more. While each home buyer's criteria and ranking of criterion importance are different, identifying neighbourhoods that meet their criteria, while also falling within their budget, can be a cumbersome task.

The primary objective of this project is to support prospective home buyers in narrowing down potential neighbourhoods for their residential property purchase. To narrow the scope of this project, we focus on the residential real-estate market in British Columbia. Seeing that the Toronto and Vancouver housing crises are coupled [12], this project aims to initially address the region where the housing crisis has had the most impact. As such, we propose an explorer tool to support home buyers in identifying neighbourhoods of interest that meet their

- *All authors contributed equally to this work.*
- *Abiramy Kuganesan. E-mail: akuganes@cs.ubc.ca.*
- *Ivan Song. E-mail: ivansong3@alumni.ubc.ca.*
- *Lufei Liu. E-mail: liulufei@student.ubc.ca.*

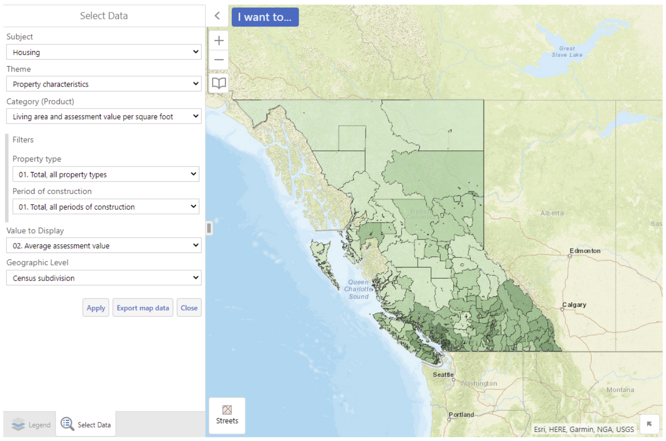

Fig. 2. The Canadian Statistical Geo-spatial Explorer [28].

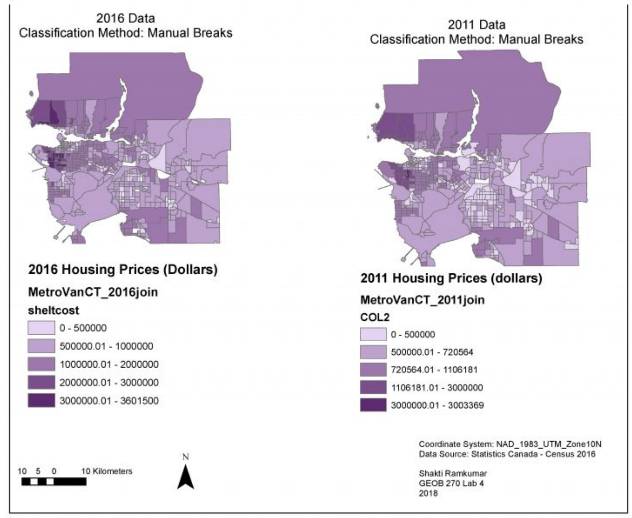

Fig. 3. A static visualization of housing affordability in Metro Vancouver by Ramkumar [22].

specific search criteria in accordance to personal importance. This work outlines the design process that supports the synthesis of this tool. An ideal task following an interaction with the proposed tool would be for the user to begin searching for homes in the neighbourhood of interest.

2 Related Work

Several housing affordability visualization tools already exist for house buyers to explore. Many are provided by government agencies such as *BC Assessment Maps* [4] and the *Canadian Statistical Geo-spatial Explorer* [28] (Figure 2). Both of these tools show regional residential housing prices. There are also non-government visualizations, such as a visualizer in a UBC blog [22] (shown in Figure 3), which do not allow for interaction. Other visualizations for neighbourhood characteristics such as safety and education are also provided separately by local governments such as *VPD GeoDash* [31]. These tools locate recent crimes and the Foundation Skills Assessment scores (FSA) for each school district in the province of British Columbia [5]. Another attribute of interest relates to transportation options, such as those included in *WalkScore* [32]. However, each of these tools focus on a specific attribute and do not consider overall neighbourhood livability. Other existing tools weigh a combination of these attributes to help potential home buyers research neighbourhoods in Metro Vancouver, including *Find Your Neighbourhood* by Vancouver Magazine [19] (shown in Figure 4) and *Find a Hood* [9]. Both of these tools require a user survey and offer less room for exploration.

In academic literature, Rinner describes a pilot study of the usefulness of geographical visualization in urban quality of life evaluations based in Toronto [24]. Liu et al. [15] and Balsas [3] show example visualizations and considerations for regional livability. Shabanzadeh et al. visualized livability in Tehran's metropolitan districts using several choropleth maps [25]. Other works such as a Malaysian study on neighbourhood evaluation [21] clearly highlight requirements that impact home-buying decisions. These considerations contribute to our proposed visualization design.

Geographical visualizations are outlined in works such as those by Cartwright et. al [8] and MacEachren et. al [18]. We rely on past works such as map visualizations of spatial and spatiotemporal data [26], cartograms [20], and Hotmaps [11] to explore trade-offs between various types of maps, task abstraction taxonomy, and color, respectively. Work by van Kreveld et al. helps us understand implications of diagram placement on maps [30] and we leverage ideas proposed in Lineup [14], necklace maps [27], and data stripes [23] for our potential solutions. In our final solution, we mimic the use of stacked bars for ranking neighbourhoods from Lineup [14] and distribution based filtering from Crossfilter [10] to filter our data. We also observe insights from Latif et al. which explore the relationship between text and geographical visualizations in data-driven stories and their influence on the reader's understanding [16].

3 Data and Task Abstractions

We first identify creiteria of importance in Section 3.1, then elaborate on our data and task abstractions.

3.1 Identifying Criteria of Importance

The grounding premise of this project relies on potential home buyers finding dwelling type, budget range, community safety, quality of education, proximity to amenities, and commuter friendliness to be the core criteria when searching for a neighbourhood. Although these criteria are validated and derived from a Malaysian study [21], we verify that this trend applies to a North American context through the distribution of the survey presented in Figures 15, 16, and 17. This survey was distributed through Reddit in the following threads:

- r/britishcolumbia
- r/vancouver
- r/CanadaHousing
- r/RealEstate
- r/SampleSize

A total of 56 respondents between the ages of 18-54 contributed to the following findings. 25 respondents were British Columbia residents and 31 residents were North Americans that reside outside of British Columbia. Responses from British Columbia residents directly inform our project while the responses from the secondary participants, belonging to the broader North American demographic, inform the scalability of our assumptions. Figure 5 shows respondents' mean rankings for criteria in accordance to their importance. Here, a ranking of 6 would indicate the most important criterion while a ranking of 1 would indicate the least important criterion. Cost of purchase price takes precedence over any other criterion. Neighbourhood safety was second most important criterion, followed by access to parks and recreation. Additionally, survey respondents indicated that proximity to employment, grocery stores, and schools were also important considerations.

3.2 Dataset Information

We obtain our data from multiple sources in order to consider different attributes that relate to our analysis. These datasets include the Canadian Census Criminal Code Violations from 2020 [28], BC Foundational Skills Assessment data (FSA) from the 2020 to 2021 school year [5], average home purchase prices according to home type from 2010 to 2020 [7], and proximity measure data provided by Statistics Canada [29]. These datasets are intended to support the following dimensions of neighbourhood search:

Dataset Source	Attribute Name	Attribute Description	Attribute Type	Attribute Details
Census (735 items)	Census subdivision	Geographical region	C	735 categories
	All Criminal Code Violations Excluding Traffic	Safety of the region, normalized per 100,000 capita	Q	Sequential; Min: 2,159.72; Max: 24,793.39;
Quality of School Scores (360 items)	District Name	Name of school district	C	60 categories
	Grade	Level of study in educational year of schooling	C	2 categories (4, 7)
	FSA Skill Code	Type of assessment issued	C	3 categories (writing, reading, numeracy)
	Score	Average assessment score for assessment issued	Q	Sequential; Min: 1; Max: 612.4;
Canadian Mortgage and Housing Absorbed Homeowner and Condominium Units (520 items)	Census subdivision	Geographical region	C	735 categories
	Dwelling type	Type of home	C	4 categories
	Year	Year the sales data was aggregated for	O	10 years
	Average home price	Average home price per dwelling type within a specific region for specified year.	Q	Sequential; Min: 364,746; Max: 6,531,910;
Proximity Measures Database (35,345 item)	Census subdivision	Geographical region	C	735 categories
	Longitude and latitude	Coordinates of a location	O	35,345 points
	Proximity scores	Closeness to source of employment, grocery stores, health care, primary/secondary education, public transit, and neighbourhood parks.	Q	Sequential; Min: 0; Max: 1;

Table 1. Summary of data attributes. C - Categorical; Q - Quantitative; O - Ordinal;

Attribute Name	Attribute Description	Attribute Type	Ordering Direction	Range / Categories
Census subdivision	Geographical region	C	–	52 categories
Dwelling type	Type of home	C	–	4 categories
Normalized safety score	Safety of the region, normalized per 100,000 capita	Q	Sequential	Min: 0, Max: 1
Normalized education score	Average combined FSA score, normalized to highest	Q	Sequential	Min: 0, Max: 1
Normalized housing affordability	Most recent average home price, normalized to highest	Q	Sequential	Min: 0, Max: 1
Average home price (2010-2020)	Average home prices for each year	Q	Sequential	Min: 364,746 Max: 6,531,910
Normalized proximity scores	Closeness to source of employment, grocery stores, health care, primary/secondary education, public transit, and neighbourhood parks.	Q	Sequential	Min: 0 Max: 1

Table 2. Data abstraction of updated combined dataset. C - Categorical; Q - Quantitative;

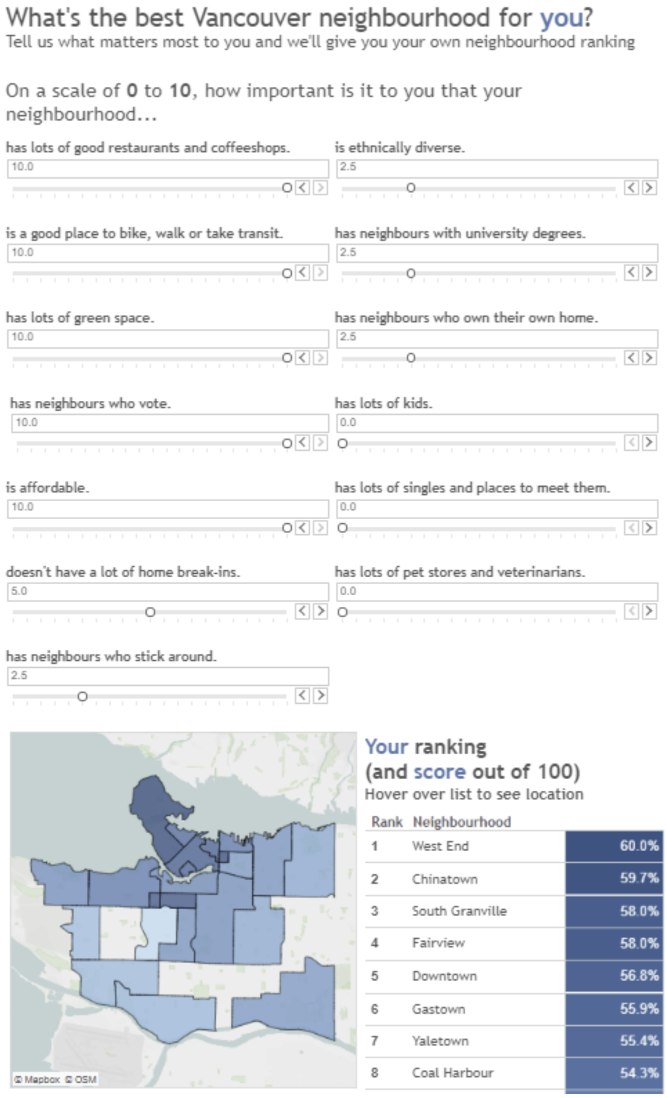

Fig. 4. Find Your Neighbourhood: a survey based interactive visualization tool by Vancouver Magazine [19].

3.2.1 House Preference and Affordability

The Canadian Mortgage and Housing Corporation (CMHC) offers a dataset of averages for absorbed homeowner and condominium units in Canada, by dwelling type and municipality for urban centres with more than 50,000 residents [6]. These dwelling types include single detached homes, semi-detached homes, row homes or townhouses, and apartments. We use annual data from 2010 to 2020 to provide the most current averages with trend information from the recent decade.

3.2.2 Safety

Statistics Canada provides a dataset on incident based crime statistics across Canada [28]. The most recent report was generated from all incidents in 2020. For relevance, we select the crime rates normalized by 100,000 population for all criminal code violations excluding traffic violations to use in our tool. The geographic level of analysis chosen was police service and detachment for the richest data available on this subject.

3.2.3 Quality of Education

The BC Foundational Skills Assessment from 2020 to 2021 school year provides an overview of literacy (reading and writing) and numeracy

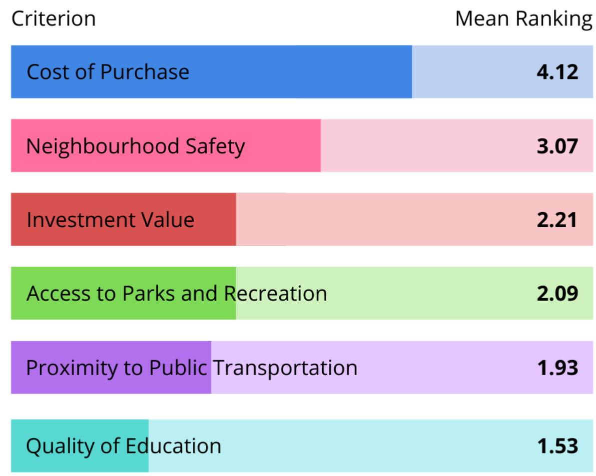

Fig. 5. Survey respondents ranked the importance of cost of purchase, quality of nearby schools, neighbourhood safety, proximity to public transportation, investment value, and access to parks. The mean rankings for each criterion are presented here. A higher mean ranking indicates higher importance for that criterion. Figure 22 represents additional criteria that respondents reported considering when engaging in this neighbourhood search task.

in grades 4 and 7 students by school district. This dataset contains 60 unique districts with 360 items [5].

3.2.4 Proximity to Amenities

According to our survey described in Section 3.1, home buyers often prioritize proximity to amenities such as neighbourhood parks or transportation. Statistics Canada offers a dataset of proximity measurements [29] for several coordinates in each census subdivision. Proximity measures are based on the distance between a reference coordinate within a census subdivision and a service. Statistics Canada uses a simple gravity model similar to the two-step floating catchment area method [17] to convert this information to a meaningful index, which considers the size of the service in addition to the distance between a reference geographic block and all the nearby blocks in which the service is located. Multiple reference coordinates are provided per census subdivision. There are 10 proximity measures and each is included as a normalized index value.

3.3 Data Abstraction

All of our datasets are organized by census subdivisions, such as cities, villages, towns, and others, with the exception of the FSA dataset, which is organized by school districts. We convert the FSA school districts into equivalent census subdivisions using administrative boundaries data from the government of British Columbia [13] and combine our dataset into a single table as the source of our visualization. Since census subdivisions also include regions, which encompasses other subdivisions such as cities, we extract only non-overlapping subdivisions from the dataset. These datasets also have a temporal element, each associated to a year between 2016 to 2021. We select the most recent data available to us, some stemming from the 2016 census and others acquired more recently. Therefore, we believe each attribute is the best representation of the present state and do not adjust for time differences between attributes. For house preference and affordability, we optionally consider a separate representation to communicate any time series data available to us.

Table 1 outlines the detailed data abstraction of these datasets, evaluated based on the raw data. These tables exclude columns that we do not plan to include in our visualization. Our final dataset combines these attributes into a single table, organized by census subdivisions. Since average home price data is only available for regions with greater than 50,000 population, we reduce our final dataset to match this subset of census subdivisions. Our final dataset has 52 items, reflecting the 52 census subdivisions we consider in our visualization. These locations are more relevant since they are more popular to live in and the reduced set still covers the majority of the province. Also, we use an average of scores for all types of FSAs for each district as the normalized education score for each census subdivision, which covers all literacy and numeracy for grade 4 and 7. The data abstraction for this updated dataset is included as Table 2.

3.4 Task Abstraction

Our task abstraction highlights the intended users of our tool and their expected behaviour.

3.4.1 Who

The intended users of our visualization tool consist of potential home buyers who are trying to pinpoint which neighbourhood to purchase residential housing in. The visualization tool should help users answer questions of where to look for housing prior to house hunting. Secondary users may also include investors, realtors, and renters. For these users, certain attributes may be less relevant than others. For instance, proximity scores may be more relevant to renters than mean purchase price. This specific iteration of the Hood Hunter tool will target home buyers interested in residential housing within British Columbia. One additional question we hope to help users answer is when to look for housing in a particular neighbourhood. We will optionally include this feature in our visualization based on time and resource constraints.

3.4.2 Actions

From our survey responses, 3 primary actions were identified as common across many prospective users' current neighbourhood search task flows.

1. Filter neighbourhoods that meet certain criteria such as budget range for a specific type of home.

2. Compare a narrow list of neighbourhoods according to user-specific dimensions of interest.

3. Dive deeper into a specific neighbourhood to understand specific assets and limitations.

At the *search* abstraction level, we expect the users to perform either *lookup* or *browse*. In the most common case, users will have a budget and a few criteria in mind when searching for places to live in. This behaviour would then fall into the category of browsing. In other cases, a user might already have a neighbourhood in mind and want to look up certain attributes regarding that neighbourhood. *Lookup* could also follow after *browse* as users start to gain a better understanding of their neighbourhoods of interest.

For *query*, we expect users to *identify*, *compare*, and *summarize*. Users can use our visualization to identify attributes regarding a particular neighbourhood after *lookup* or identify a particular neighbourhood with certain attributes after browsing. One example is to identify the neighbourhood with the cheapest housing. Another use of the visualization is help users make comparisons between multiple neighbourhoods to find the most suitable one. Lastly, the ability to summarize is not only useful for users to have a high-level overview of the attributes for all neighbourhoods in BC, it also provides them with opportunities to identify trends in certain attributes for particular neighbourhoods.

In terms of the *analyze* abstraction level, the main use for our visualization tool is to consume existing data regarding neighbourhoods in BC. In particular, users will use our visualization to discover new insights regarding which neighbourhoods are more suitable for their needs. Another use of the visualization tool is to produce new information about neighbourhoods. In some cases, we will be transforming raw quantitative data to ordered ranged data to improve the legibility of the information presented to them.

3.4.3 Targets

The main high-level targets are *trends* and *distributions* which provide users with rich insights about the various attributes across neighbourhoods for filtering, comparison, and detailed inspection. Seeing the distribution of the data according to a specific criteria when filtering can be more informative than specifying values without much context. An example of how *trends* might be useful for detailed inspections is for a user who cares about the investment value of housing. They can compare how fast the housing price within a particular neighbourhood is growing compared to another. For the specific attribute of housing price, *distribution* and *extremes* are also targets of interest. The distribution of the price for houses sold in a neighbourhood will give a more accurate representation of pricing than solely average price. Looking for extreme minima and maxima in pricing data over time can also enable users to identify when the most appropriate time is to purchase houses in a particular neighbourhood. A secondary target of interest might be *outliers* when users are trying to find the most suitable neighbourhood for specific criteria. One example is a user who prioritizes safety; they might want to find the neighbourhood with the lowest crime rate in BC.

4 Solution

This section details our initial ideas for the Hood Hunter tool in Section 4.1, which then influence our idiom and design choices for our final design.

4.1 Proposed Solutions

This section outlines three potential scenarios and the possible solutions associated with them.

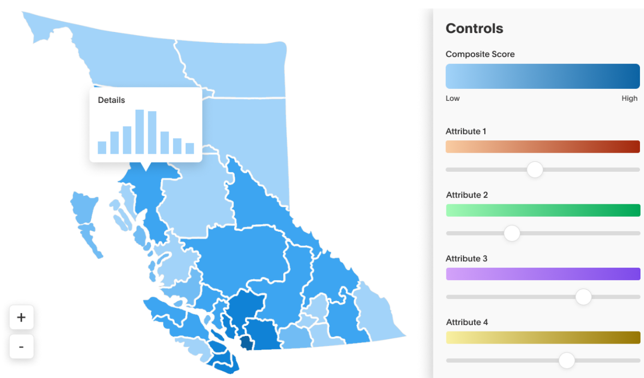

Fig. 6. Solution A mockup

4.1.1 Solution A & Scenario

Figure 6 shows a possible visualization design. In this example, a user may be interested in finding a potential neighbourhood in British Columbia that is most suitable for them to live in. They are in an exploration phase and are open to the possibility of living anywhere in the province. The map provides the user with a generalized view. The main map is divided into Census subdivisions and the luminance channel encodes a composite score based on the user's priorities. These priorities are defined by the position of the slider inputs. The hue channel encodes different regional characteristics and the luminance of each of these indicators encode their value. To explore different attributes, the user can also select a specific attribute to colour the map in replacement of the composite score.

Once the user has selected a suitable region, they may zoom into the map further to view smaller Census subdivisions on the map. When they have narrowed down their search to specific subdivisions, the user can select the subdivision and view further details. For instance, a histogram encoding the distribution of housing prices in the subdivision with corresponding lengths.

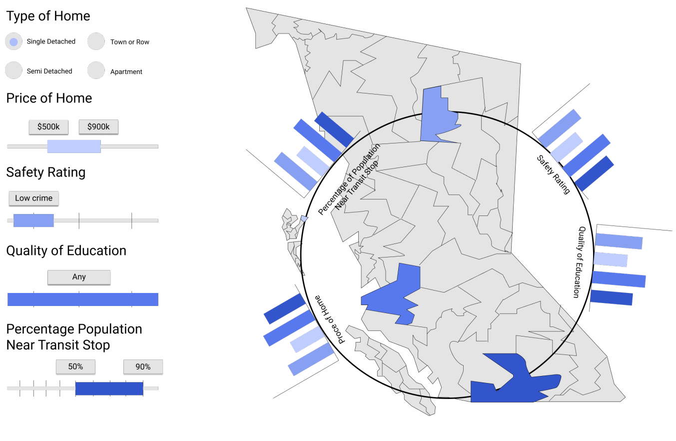

Fig. 7. Solution B mockup

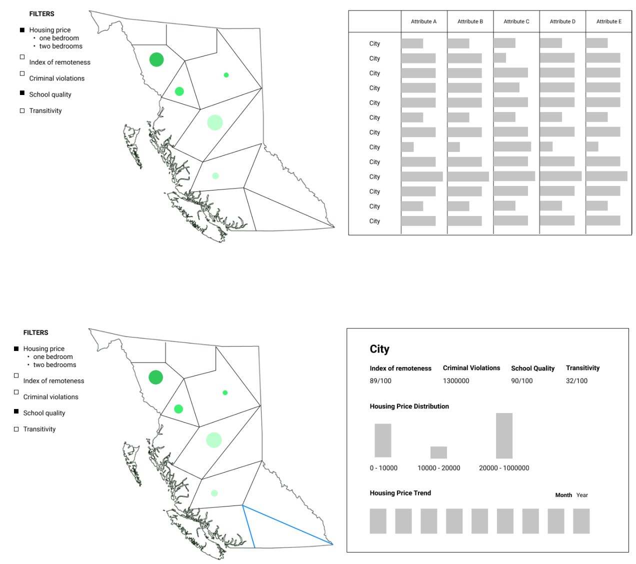

Fig. 8. Solution C mockup

4.1.2 Solution B & Scenario

Loosely inspired by the concept of Necklace Maps [27], Figure 7 presents an interactive explorer tool to support neighbourhood search.

The mark of type containment encodes census subdivisions. The fit of neighbourhood is a cumulative score of each attribute. The visual channels saturation and luminance redundantly encode a unique neighbourhood and the fit of the neighbourhood. The top four results are shown to assist the user in the task of filtering. Each neighbourhood's attribute value is represented by a mark of type line. The hue and luminance of the mark on the histogram denote the neighbourhood on the map. Additional labelling may be used to display neighbourhood names.

In order to manipulate this tool, the user is required to input their desired criteria to filter the neighbourhoods for best fit. These inputs will be user defined ranges of tolerance for each attribute. The user may drag the line mark and expand it to indicate a wider tolerance and shrink the length to indicate a narrower tolerance. The user is asked to input type of home preference through a multiple choice selector. As the user toggles these controls, a cumulative score will be computed according to the ratings inputted by the user. The top four neighbourhoods will be assigned saturation and luminance values indicating best fit. Darker and more saturated regions will indicate better fit than lighter and less saturated regions. The histograms along the necklace may be used to easily understand and compare each attribute for the top four neighbourhood recommendations.

4.1.3 Solution C & Scenario

Figure 8 showcases another possible visualization design which allows users to visualize particular attributes of interest to support the task of neighbourhood comparison. On the left hand side, a user can filter which attribute to visualize on the map. Each neighbourhood is represented with a point mark where size and saturation channels are used to encode two attributes. A user can choose up to two attributes to visualize with the default attribute being housing price encoded by the size channel. As this attribute's assumed significance is high, it would be justified to encode it using a channel with relatively higher effectiveness. Users can then choose which other attribute they want to encode as well on the map. On the right, there is a table with embedded bars showcasing attributes for every neighbourhood. Users can use this table to easily rank and compare neighbourhoods by attributes of interest.

The lower mockup on Figure 8 shows how the display changes once a neighbourhood is selected. The selected area will be highlighted as the table with embedded bars is replaced by a summary of key information regarding the selected neighbourhood. This feature will support the task of lookup, offering an overview of a particular neighbourhood. The overview will showcase the values for attributes such as index of remoteness, criminal violations, school quality, and transitivity. It will also showcase the neighbourhood's housing price distribution and housing price trend overtime. This feature can be particular useful if users want to learn more about a neighbourhood they are further interested in.

4.2 Solution

4.2.1 Interim Solution

Through synthesis of these 3 solutions, we were able to identify pros and cons that contributed to the design iteration described in this section. Figure 9 presents our interim iteration of the visualization tool. As previously mentioned, from the survey, we identified respondents' common neighbourhood search task flows. Similarly, we design our final solution to support these core tasks of filtering, comparing, and looking up details.

On the left pane of the mockup is the filter view which allows users to filter neighbourhoods in BC by attribute values. Each attribute is paired with a a histogram which simultaneously shows users the attribute's distribution and allows users to select their range of interest. The filtered neighbourhoods are then highlighted on the map on the right. The histogram is meant to give users more context on what is considered a good or bad attribute value. This plot is intended to help clarify more ambiguous attributes such as crime score. A composite score generated by this filtering task highlights appropriate census subdivisions on the map view.

On the right pane of the mockup is the ranking view which facilitates comparison between filtered neighbourhoods. Taking inspiration from Line-up, the attribute values are represented by stacked horizontal bars with color encoding each attribute value. The filtered neighbourhoods may be ranked based on the selected attributes. Users can also choose to change the weight of each attributes which effects the multi-attribute ranking of the neighbourhoods. Lastly, users can star the neighbourhoods of interest which bring them to the top of the bar graph for easy comparison.

At the bottom of the mockup is the detail view which displays detailed information of a single selected neighbourhood triggered by a mouse click. The detail view will first show one short text description of the neighbourhood. Then, the average housing price in the last 10 years in the neighbourhood is encoded by a line graph allowing users to estimate the neighbourhood's investment value. The other attributes are then shown with the value on top and the distribution encoded by a histogram below. The bin that the attribute value falls under will be highlighted.

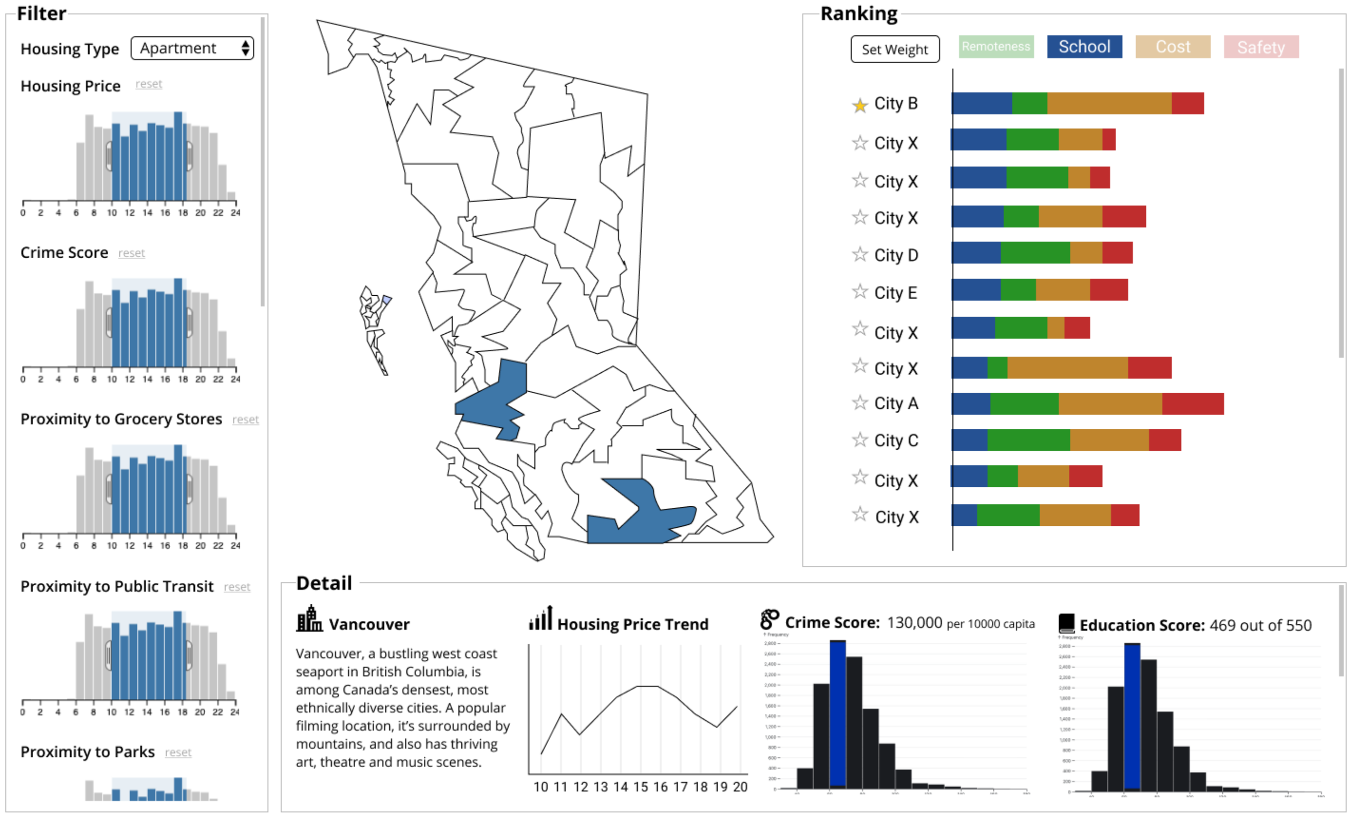

Fig. 9. Interim solution mockup

All three views are juxtaposed on the screen to better facilitate interactions between them without requiring users to recall from memory. For example, after a user is interested in a neighbourhood after checking its details, they can directly star the neighbourhood in ranking view and make comparisons. Clicking on the census subdivision on the map will bring up the detail pane for that census subdivision. Clicking the census subdivision title on the ranking view has the same effect.

4.2.2 Idiom Choices

We encode data described in Table 2 using the *length* channel to show magnitude and *colour* channel to distinguish between attributes. We choose a stacked bar chart so that bars are aligned and easier to judge relative values. We also include resorting by specific attributes with the associated bar aligned for better comparison as suggested by Weber's Law – similar to Lineup [14]. This effect is implemented as a *jump cut* for simplicity.

Our combined dashboard features *juxtaposed* views that are *linked* and *multiform*, sharing the same colours for attributes. These views share data with an *overview-detail* scheme and one view is dedicated as *detail-on-demand*. Our map view has *linked navigation* and zooms according to selections from the ranking view in an *animated transition*. The map also includes a *superimposed static layer* encoded with circle *marks* to indicated locations with available data.

We allow *item filtering* to *reduce* and better manage complexities. We implement a *scented widget* modelled after Crossfilter [10] to provide user with more context of the overall dataset. Then, selected cities are highlighted both on the map and in the detailed view with *linked highlighting* using the *colour* channel – specifically with red to draw attention.

4.2.3 Design Choices for Final Iteration

After further inspection of our mock up and cross reference with our survey results, we realized that the most relevant crossfilter was the one associated with house pricing. To reduce the visual overload on the screen, we decided to reduce the filtering pane to just this. For the price range crossfilter, we experimented with completely discarding bars that are not within a specific budget range and dynamically changing y axis to focus on the disparities between bars within the focus region. This design does not provide much visibility of the neighbourhoods outside of the selection and so, we choose to stick with the classic crossfilter design implementation. We acknowledge that having bars on both the min and max region of the crossfilter selection would be more intuitive to the drag action required to manipulate it. Due to limited time, we were unable to implement this idea.

Initial iterations of the solution utilize a choropleth map for the map view. We choose to filter out overlapping census subdivisions and focused on five specific census subdivision types with the highest level of granularity. These types include cities, towns, villages, island municipalities, and district municipalities. These census subdivisions are much too small to code with colour. As such, we decide to use a symbol map to indicate the location of census subdivisions. Since these census subdivisions are somewhat clustered within the lower half of British Columbia, we focus the map on this region by default. We encode filtered neighbourhoods with a semi-transparent colour channel corresponding to the hue of the filtered bars on the crossfilter and a solid outline to make it easier to distinguish the number of neighbourhoods within a cluster. The use of colour aims to form the notion that the neighbourhoods in black on the map are within the indicated price range.

Since length comparison is the strongest along a normalized line, we implement a way to compare individual attribute values on the stacked bar chart in the comparison view. By clicking on the legend for a specific attribute value, users are able to see that value for each neighbourhood aligned at the left vertical axis. This allows for easier visual comparison of this attribute across all the filtered neighbourhoods. Legend placement on the comparison view continues to be a point of concern. Due to limited screen real estate and the assumption that most users will choose to view neighbourhoods from best to worst, we place the legend on the bottom right of the panel to avoid as much occlusion as possible. We also considered dynamically placing the legend on the bottom right or top right depending on whether whether rank by best to worst or worst to best is selected. Due to time limitations, we were unable to implement this idea.

From the survey collected, investment value was ranked as one of the top criteria BC residents care about when choosing neighbourhoods (Figure 5). As such, we choose to put housing price trend as the largest component at the top of detail view. We use line chart to visualize price trend due to its effectiveness in showcasing trends and use dots and dotted lines to visualize missing data or data gaps. Each score's distribution is visualized using histogram due its wide usage for representing distribution which ensures greater familiarity with the intended users.

As percentiles might not be familiar to all user types, we experimented with presenting the label for neighbourhood performance for each attribute as a percentage or fraction out of one. We ran a user study with five individuals to assess which of the format is more effective at conveying the message. The objective is to convey that the neighbourhood ranked in the Nth percentile of best performance for a specific attribute when compared to other neighbourhoods. We presented half of the attribute results as percentages and half of the attribute results as fractions for their analysis. Four out of five participants found the percentage representation to be most effective for this purpose.

Our solution's initial implementation had all the different views combined in a similar layout without borders or labels. The initial intent was to give as much space to each view as possible. After collecting an initial round of feedback on the interface layout, the two participants reported the interface to be disorganized and messy. This information led us to implement larger white space coupled with drop shadows, borders, and a background colors to make each view visually distinct. Labels were also added to better communicate each view's affordances.

4.3 Tools

For pre-processing, we use Python and QGIS Geographic Information System. We implement the individual components of the exploration tool using Observable D3. In particular, we leverage the Vegalite and Leaflet [1] packages for development. Our loading page and main web page use HTML and CSS. The specific use of these tools is discussed further in Section 5.

5 Implementation

Our final implementation can be separated into three separate views and their interactions, which creates the full user interface.

5.1 Map and Filter View

QGIS Geographic Information System [2] is used to convert Census Subdivision shapefiles to GeoJSON format and to transform the coordinate system of these shapefiles to latitude and longitude values. The Canadian Census Subdivisions were filtered for entries containing British Columbia as the province name. A Python pre-processing script is then used to identify the central coordinate of the census subdivision. The Leaflet.js library [1] creates the map component of the proposed visualization tool. Circle marks are placed in the position identified by the pre-processing script for each census subdivision. This implementation is pictured in the bottom half of the filter pane in Figure 1.

Each neighbourhood can be filtered by price range using the crossfilter pictured on the top half of Figure 1. This crossfilter is implemented using the Vega Crossfilter implementation in Observable D3. Due to limitations in data, not all neighbourhoods offer pricing data for each dwelling type. Modifications are made to this implementation for formatting the histogram bins dynamically based on the amount of data available per house type. We implement the house type selector using Observable D3. It influences the histogram bins shown on the crossfilter and the corresponding neighbourhoods.

Implementation and integration of the map component with the crossfilter component proved to be a difficult task. We tried the original crossfilter package interfaced with a Vega map and a Leaflet map prior to switching to a Vega Crossfilter implementation. We initially interfaced with a Vega map but due to difficulties implementing our desired interactions, we use a Leaflet map which relies on Open Street Maps.

5.2 Ranking / Comparison View

For our comparison view, we generate a horizontal stacked bar chart in our Observable notebook using D3 and its SVG tools. This plot presents a composite score and normalized components from each attribute from the data as presented in Figure 1 (middle). The colour legend doubles as a user selection which redraws the plot sorted by the selected attribute, ranging either from best to worst or worst to best. In addition, we add hover effects to highlight the neighbourhood of interest so users can clearly distinguish the association between a bar and its neighbourhood. The weights of each attribute can also be adjusted through a modal with slider inputs, implemented as a combination of the Observable notebook and our HTML/CSS dashboard. Finally, we add SVG stars that allow users to select *favourite neighbourhoods*, effectively filtering neighbourhoods to only those they would like to consider and easier comparison.

5.3 Detail View

Figure 1 (right) demonstrates our implementation the detail view. Based on the city selected, we show detailed information such as housing price trend and the attribute scores of the selected neighbourhood coupled with a histogram for each score distribution. In the histogram, we also highlight the bin where the neighbourhood falls into. We implement our details view using the native Observable Plot and the Line Chart library to create the appropriate histograms and line chart. Highlighting and hover tool-tip effects are added by our team. The line chart is also customized to plot points and dotted lines to clearly visualize missing data. All of these were done by editing the source functions using D3. We also add code to organize the charts, display scores, and display city name using HTML and CSS. Helper functions are also implemented in JavaScript to adjust the data dynamically to the housing type and the neighbourhood chosen.

5.4 Interaction

The three views described above are linked together in our final Hood Hunter dashboard (Figure 1). They are linked by shared global variables, which include `selected_csd` and `hovered_csd`, as well as a shared dataset that is filtered by a crossfilter. A detailed data flow is described in Figure 10. We combine the views in an Observable notebook and embed them in our combined dashboard.

5.5 User Interface

The combined interface is implemented with HTML/CSS/JavaScript and hosted as a web app through Github Pages. The individual views are embedded as JavaScript through D3 Observable Notebook's cell embedding feature which utilizes their Runtime API.

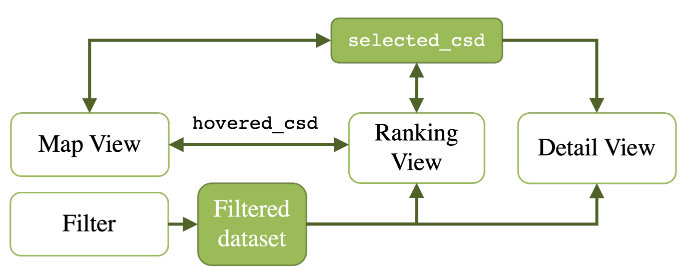

Fig. 10. Data flow between views in final combined dashboard.

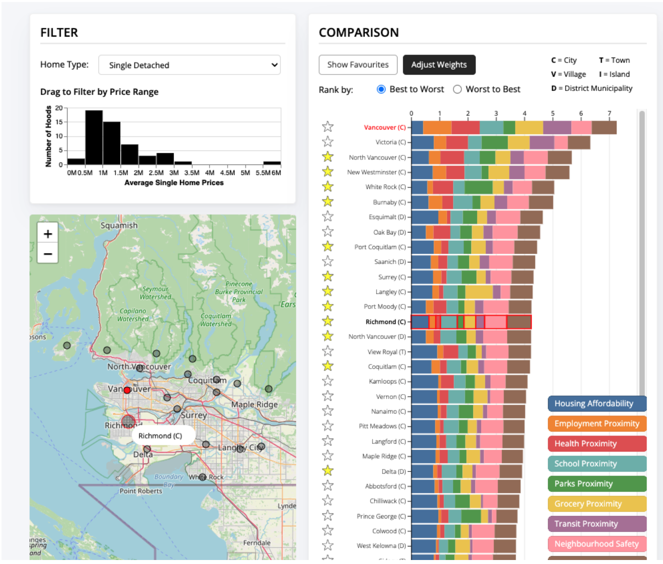

Fig. 11. Map view of Vancouver with neighbourhood hover interaction.

6 Milestones

Table 3 outlines our milestones, their estimated deadlines, actual completion times, and the team members assigned to the task. The total amount of hours estimated is 80 hours per group member. The actual total amount of hours was 95.75 hours per group member.

7 Results

We present two scenarios of use that demonstrate how our solution Hood Hunter facilities the tasks of our intended users. Overall, we feel that Hood Hunter.... We also discuss the user studies that were conducted in evaluating our solution which involved some informal feedback on early implementations, A/B testings, and questionnaires.

7.1 Scenario 1: Moving out of the city

Sam lives in Vancouver, is in his mid 30s, and is looking to start a family with his partner. He is looking for a house which is suitable for a family. However, he has noticed that housing price in Vancouver is too expensive for him to afford the type of house he wants. Since his work has changed to be remote after the Pandemic, he is opened to exploring options outside of Vancouver in nearby cities or towns as they most likely have cheaper houses. Sam sees Hood Hunter on the web and decides to give it a try. He first chooses single detached to be his home type. He then chooses the price range of 0 to 2 million dollars which is within the range of his budget. He finds Vancouver on the map and starts exploring around it. He has heard about most of these neighbourhoods around Vancouver, but does not really know the specifics about them. He notices that hovering over the neighbourhoods on the map also highlight them on the comparison view (Figure 11).

After shifting his focus on the comparison view, he realizes that the neighbourhoods are ranked based on the combined attribute score. He then adjusts the attribute weights based on his personal needs, increasing the weights of housing affordability, education quality, and decreasing weights of employment proximity and grocery proximity (Figure 12). He finds this feature really helpful as the ranking is now personalized. He wants to only compare the neighbourhoods around Vancouver so he finds them on the comparison view and favorites them to only show those. This task takes some time and is a bit frustrating. From the top of the ranking, he clicks on the neighbourhoods he is interested in to see their housing price trends and attribute scores to gain a better understanding of their strengths and weaknesses. Any neighbourhood that he isn't interested in, he takes out by unstarring it. At the end, Sam has a better understanding of which neighbourhoods around Vancouver that are more suitable for his needs.

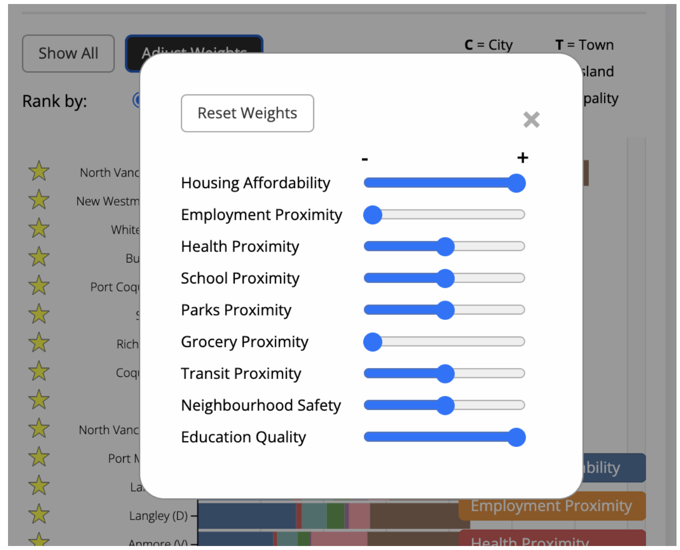

Fig. 12. Weight adjustments for individual attributes on the comparison view's stacked bar chart. Adjusting the weight for an attribute modifies how much that attribute contributes to the composite score which is encoded by the total length of a bar.

7.2 Scenario 2: Moving to BC

Lizzy is in her mid 20s, lives in Calgary, but is looking to move to BC after graduation. She has a few places in mind after hearing friends saying nice things about Vancouver and Victoria. Nonetheless, she wants to learn more about the neighbourhoods in BC to help her make the decision of where to move to. Since she is just starting her career, she wants to looks for somewhere to rent. After opening Hood Hunter, she first notices that the dashboard seems to be more tailored toward house buyers with the pricing information. Nonetheless, she finds the other attributes helpful. She also assumes that there will be some correlation between housing affordability and rental prices. Since she is really curious about how the neighbourhoods would rank base on a few attributes she is particularly interested in, she mainly uses the comparison view ranked based on a single attribute selected. She finds that the bars corresponding to the attribute selected being in same position makes it easy to compare between neighbourhoods. She also likes how the geographic location of the neighbourhoods are highlighted as she hovers over the bar going down the ranking list. Using this method, Lizzie is able to discover some interesting insights. For example, ranking by housing affordability, Lizzy discovers a few cities close to Victoria such as Langford which is significantly cheaper than Victoria (Figure 13). After clicking on Langford, she also sees in the detail view that for a few attributes she cares about such as neighbourhood safety and parks proximity, Langford scores relatively high compared to the other neighborhoods (Figure 14). She favorites Langford and dives into Google to learn more about it.

7.3 User Study

We conducted an user study with seven participants, referred to as P1-P7. All of the participants are in the age range of 22 - 27 years old. Six of the participants self identify as male, and one participant self identify as female. Three of the participants are students in the fields of Computer Engineer, Mechanical Engineer, and Biology. The other participants are working professionals who work as a Product Designer, Computer Engineer, Computer Scientist, and Data Engineer. All participants are not current students of CPSC 547.

Participant were first asked to use Hood Hunter to complete the given task of finding a neighbourhood that they would want to buy a house at in BC. This survey was done remotely without observation from researchers. After participants completed the tasks, participants were asked four questions about what steps they took to accomplish the

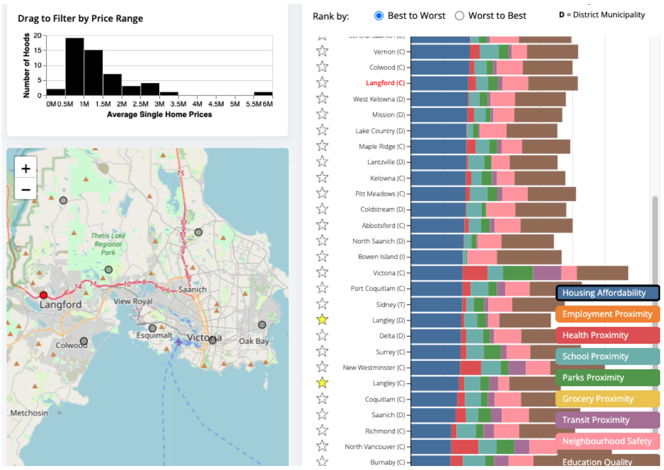

Fig. 13. Ranking by housing affordability.

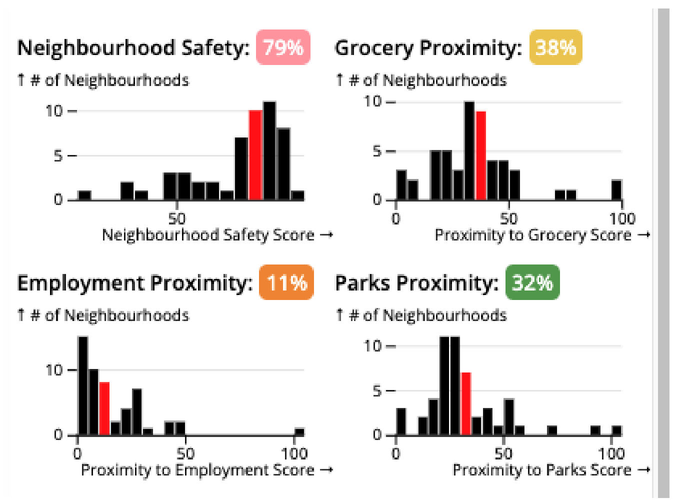

Fig. 14. Histograms in detail view showcasing score distribution with the selected neighbourhood's bin highlighted.

task, what they liked and disliked about their experience with the tool, and whether they would use this tool to look for a neighbourhood to live in. Overall, participants found the tool to be intuitive, useful, and informative. They did, nonetheless, found a number of limitations and issues that would be important to address to increase the tool's utility and effectiveness.

7.3.1 User Flow

As expected, most participants followed the intended order of filter, compare, then detail. This result might also be large influenced by the layout of the interface which structures the views in a left to right order. This order, however, was not the case with all participants. For example, P7 "first looked at the map to see what options were listed, after that [they] started looking through the top several neighbourhoods in each neighbourhood attribute tab". Interestingly, the map was an important view for many participants to use as the primary method to navigate through neighbourhoods. This effect is most likely due to the strong relation between neighbourhoods and their geographic locations. Another notable observation is that both P2 and P3 used the stars not as a way to favorite the neighbourhoods they like after comparing, but as a way to filter neighbourhoods for comparison. P2, for example, mentions that they first "restrict to Greater Vancouver Regional District using favorites" then adjusted weights and made comparisons.

7.3.2 Strengths of the Tool

Most participants mentioned that they liked the look and layout of the interface which "is really clean and each widget provides useful and distinct information"(P7). P1, P3, and P7 particular mentioned how they liked the comparison view which made it "relatively easy to make a final decision from several options"(P7).

7.3.3 Difficulties with the Tool

Two participants had concerns over how the data is generated. P6, for example, mentioned that the "scores [that] were calculated in the details were very black-box, not sure how to interpret the scores". This concern can greatly effects how much users trust the tool and limits its effectiveness. Two other participants also mentioned that the full dashboard was a bit overwhelming at first with "too many things presented at first"(P5). P7 suggested to ease users into the full dashboard and mentioned how "a short introduction into each widgets capabilities/functions might be useful". Although, initially overwhelming, the the two participants did find the interface to be quite intuitive and useful after the initial learning period.

7.3.4 Real Life Usage

Six out of seven participants indicated that they would use the tool to look for a neighbourhood to live in. P1 and P3 both thought the tool would be useful as part of their process to "use really early on for information at first"(P3), "as it provides good high-level data comparing all neighbourhoods at a glance"(P1). This also matches our intended use of the tool which is meant to make the early process of house hunting better and more convenient. There were, however, a few limitations that were mentioned by participants. First, P7 noted that the tool "doesn't really provide a sense of what a neighbourhood you're interested is like" where they suggest to "include some qualitative aspects (pictures, nearby landmarks/attractions, etc) into the details section". Both P2 and P7 also brought up the issue of neighbourhoods being too broad where they would be "more compelled to use this tool if it provided more specific information within each neighbourhood, to the specificity of actually neighbourhoods, rather than entire cities/towns".

8 Discussion and Future Work

This section describes the strengths and weaknesses of our solution, what we learned from this project, and possible future directions for our tool.

8.1 Strengths and Limitations

Unlike previous housing affordability visualization tools, Hood Hunter visualizes several key attributes on a single screen. The data is compiled from multiple sources and helps the user explore various neighbourhoods without having to navigate away from the application to search for these attributes. We also ensure this data is intuitive to users by using common visualization idioms such as stacked bar charts, histograms, and line charts that are generally familiar and obvious to understand.

Our solution communicates data to the user in a highly transparent manner. By using a crossfilter for our price range selection, users have full visibility to all available prices and the number of potential neighbourhoods in each. The same idea is present in our detail view, where histograms show the entire range of attribute scores and the proportion of neighbourhoods that fall into each bin. For our comparison view, the Lineup-style ranking system shows precisely how the composite score is calculated and helps us visualize multiple attributes as opposed to similar visualization tools that only show a single attribute. Adjusting attribute weights is also directly visible in our dashboard. However, despite our best efforts, our user study indicates many users still consider the tool like a "black-box", which motivates for more transparency of data sources. Due to time limitations, we do not add additional descriptions to score metrics.

We include several views in our final solution, which offers user a broad range of exploration options to help identify their ideal neighbourhood. The juxtaposed views are also easily digestible by our use of linked highlighting. However, we find that fitting all the views on a single screen is difficult and can be overwhelming to a new user. Given more time, our solution can include an introduction tutorial that appears for first-time users and provides a walk through of the tool. We can also include a *help* modal to guide users, which could also address the previous concern of understanding score metrics. This idea was also proposed by users during our user study.

Another consequence from the limited screen space is our smaller font sizes. Although they are legible to a general audience, the text may be too small for older audiences or those that are visually impaired. There is also no mobile-friendly version.

While our Hood Hunter dashboard attempts to meet the requirements set out in Section 3.4, we are largely limited by the data available to us. For example, we do not have detailed pricing data to help users determine the best time to purchase a home and we are missing data points for several dwelling types in many cities. We are also limited by static data in our implementation, so these data gaps cannot be filled automatically and future data cannot be easily incorporated into the visualization.

We are also limited by our time constraints for this project. Our crossfilter implementation increments in half million dollars for filtering by price range, which may be too broad. Alternatively, we can include text-inputs for finer granularity in selections. Section 8.3 describes other possible additions to our solution given more time and resources, such as additional bilateral linking between views.

8.2 Lessons Learned

Through our experience in building Hood Hunter, we find that stakeholder input is very helpful in guiding the final solution. Our survey results highlighted the importance of proximity measures, which we would not have considered ourselves. However, we also note that searching for publicly available datasets to meet a specific set of requirements can be very difficult. There is generally additional post-processing involved in these situations. For example, we converted the geographical regions in our FSA dataset, CMHC dataset, and proximity measures dataset to census subdivisions in order to match all attributes.

We also better understand motivation behind visualization best practices. For example, our original choropleth map idea clearly demonstrated its weakness of smaller regions being less visible, which was especially disadvantageous in our tool since large cities often have high population density in a small region. Also, overloading the colour channel is difficult to avoid since we use colour to encode our attributes but also require a highlight colour that does not interfere.

We also explore Observable notebooks and libraries such as Leaflet and Crossfilter in building our tool. These frameworks require more time to learn than we originally anticipate, with multiple approaches to achieve the same result and no obvious debugging mechanism. However, we now understand these tools better, which can help us in future visualizations. Lastly, our timeline fails to consider efforts required to style the final dashboard, which is an important step to ensure the multiple views are in sync and organized so they are not confusing to users.

8.3 Future Work

While we strongly believe that our tool provides value to this solution space, there is still work to be done to make it intuitive and integrate well into the entire house and neighbourhood search pipeline. We detail a few extensions to improve our current solution.

8.3.1 Functional and Interactive Features

Census boundaries are currently not shown within the tool. By transforming the red opaque circle mark into a transparent red boundary upon neighbourhood selection could be more meaningful to users. This feature would allow them to explore within the neighbourhood to learn more information on the map view.

Browsing is an important action during this neighbourhood search process. It is most similarly paralleled by the hover interaction type. Hovering over neighbourhoods in the detailed view's individual histogram bars per attribute could link to neighbourhoods on the map being highlighted with the blue colour channel. This effect could also carry across to the comparison view where the stacked bar graph can be highlighted with a blue border and the neighbourhood name text can be shown in blue for those select neighbourhoods. This feature would be helpful for users to understand what other neighbourhoods might also be in a similar bin.

Geographic area filter would also be a helpful addition allowing users to drag over the map to filter neighbourhoods within the selection area. Since the user study revealed that many users like to filter the neighbourhoods before comparing, usually by geographic region, area filter could address the inconvenience of filtering by favorites and allow for multiple selections filter in the user's designed region.

8.3.2 User Experience Enhancement

During our user feedback study, many users indicated not knowing the correct order of interactions or being overwhelmed by the amount of information making it difficult to know what each component does.To address this concern, interactions with the exploration tool may be guided at first with a walk through that requests users to perform an action according to their criteria while visually blocking other components out. This guided tutorial would eventually lead to the display of the full dashboard at which point the user will be familiar with each of the components and the impact that interactions may have on other components.

8.3.3 Smart Extensions

Beyond this tool, survey participants mentioned that next steps included a deep dive into the neighbourhood, often driving around to perform an assessment. With the coordinate level information about the proximity data that we currently have, we can provide more granularity as to which streets within a neighbourhood may be more desirable than others. Integrating with the Open Street Maps API could allow for users to search for how potential homes of interest in their desired neighbourhoods compare within our tool. Users may drop a pin or enter an address at which point the proximity values can be shown in the detail view as a percentile for performance in comparison to other homes within a specific neighbourhood.

Price projections for the housing market is a hot topic. These projections can be extremely helpful for buying decisions and understanding investment value over time. Integrating pricing projections into the line chart in our detail view with a blue colour channel and dotted line connector for a specified month or range of years could be beneficial for this task.

9 Conclusions

In this work, we presented Hood Hunter, an explorer visualization tool to help potential home buyers narrow down neighbourhoods of interest. A survey was first conducted on potential users to better understand their process, needs, pain points, and criteria of interest when choosing the neighbourhood to live in. Based on the survey results, we iterated on multiple designs to finalize on an interactive dashboard with 3 core components which reflect the user tasks of filter, comparison, and detail. We implemented the solution with D3 observable notebook embedded in a web application using HTML/CSS/Javascript for easy accessibility for the intended users. A user study with 7 participants revealed that House Hunter was intuitive to use and was useful in supporting the task of choosing neighbourhoods. It did, however, reveal a number of shortcomings and limitations of the tool which we hope to address with future work.

References

[1] Leaflet: An open-source javascript library for mobile-friendly interactive maps.

[2] QGIS: A free and open source geographic information system.

[3] C. J. Balsas. Measuring the livability of an urban centre: an exploratory study of key performance indicators. *Planning, Practice & Research*, 19(1):101–110, 2004.

[4] BC Assessment. Bc assessment maps, 2021.

[5] BC Data Catalogue. Foundation skills assessment 2017/18-2020/21, 2021.

[6] Canada Mortgage and Housing Corporation. Absorbed units by price range, 2021.

[7] Canada Mortgage and Housing Corporation. Price quartiles and averages: Absorbed homeowner and condo units, 2021.
[8] W. Cartwright, S. Miller, and C. Pettit. Geographical visualization: past, present and future development. *Journal of Spatial Science*, 49(1):25–36, 2004.
[9] Clockworkmice Ltd. Find your perfect neighbourhood, 2021.
[10] J. Davies and T. Carden. Crossfilter: Fast multidimensional filtering for coordinated views, 2021.
[11] D. Fisher. Hotmap: Looking at geographic attention. *IEEE trans. on visualization and computer graphics*, pp. 1184–1191, 2007.
[12] J. Gordon. Reconnecting the housing market to the labour market: Foreign ownership and housing affordability in urban canada. *Canadian Public Policy*, 46, 2020. doi: 10.3138/cpp.2019-009
[13] Government of British Columbia. Administrative boundaries, 2021.
[14] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. Lineup: Visual analysis of multi-attribute rankings. *IEEE trans. on visualization and computer graphics*, pp. 2277–2286, 2013.
[15] L. Jianxiao, B. Han, and M. Wang. Using multi-source data to assess livability in hong kong at the community-based level: A combined subjective-objective approach. *Geography and Sustainability*, 1(4):284–294, 2020.
[16] S. Latif, S. Chen, and F. Beck. A deeper understanding of visualization-text interplay in geographic data-driven stories. In *Computer Graphics Forum*, vol. 40, pp. 311–322, 2021.
[17] W. Luo and Y. Qi. An enhanced two-step floating catchment area (e2sfca) method for measuring spatial accessibility to primary care physicians. *Health & place*, 15(4):1100–1107, 2009.
[18] A. M. MacEachren, M. Wachowicz, R. Edsall, D. Haug, and R. Masters. Constructing knowledge from multivariate spatiotemporal data: integrating geographical visualization with knowledge discovery in database methods. *Int'l Journal of Geographical Information Science*, 13(4):311–334, 1999.
[19] S. McLachlan. Find your neighbourhood: The vanmag neighbourhood ranking tool, 2018.
[20] S. Nusrat and S. Kobourov. Task taxonomy for cartograms. 2015.
[21] A. Olanrewaju and C. Wong. Evaluation of the requirements of first time buyers in the purchase of affordable housing in malaysia. *Journal of Housing and the Built Environment*, 35:309–333, 2020. doi: 10.1007/s10901-019-09681-z
[22] S. Ramkumar. Visualizing housing affordability in metro vancouver, 2018.
[23] R. Rao and S. K. Card. The table lens: merging graphical and symbolic representations in an interactive focus+ context visualization for tabular information. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 318–322, 1994.
[24] C. Rinner. A geographic visualization approach to multi-criteria evaluation of urban quality of life. *International Journal of Geographical Information Science*, 21(8):907–919, 2007.
[25] R. Shabanzadeh Namini, M. Loda, A. Meshkini, and A. Roknedineftekhari. Comparative evaluation of livability indicators of the metropolitan tehran's districts. *Int'l Journal of Urban Sustainable Development*, 11(1):48–67, 2019.
[26] M. Shaito and R. Elmasri. Map visualization using spatial and spatio-temporal data: Application to covid-19 data. In *PErvasive Technologies Related to Assistive Environments (PETRA) Conf.*, pp. 284–291, 2021.
[27] B. Speckmann and K. Verbeek. Necklace maps. *IEEE trans. on visualization and computer graphics*, pp. 881–889, 2010.
[28] Statistics Canada. Canadian statistical geospatial explorer, 2021.
[29] Statistics Canada. Proximity measures database – early release, 2021.
[30] M. van Kreveld, É. Schramm, and A. Wolff. Algorithms for the placement of diagrams on maps. In *ACM Int'l workshop on Geographic information systems*, pp. 222–231, 2004.
[31] Vancouver Police Department. Geodash crime map, 2021.
[32] Walk Score. Cities & neighbourhoods, 2020.

10 Appendix

10.1 Attribute of Interest Survey

The following survey was used to understand the most important criteria to prospective home buyers when searching for a neighbourhood in which to purchase residential property. This survey focuses on BC home buyers specifically, but includes the wider North American channel as a secondary target. Synchrony between results gathered from both types of respondents indicate that a solution developed to visualize these attributes is scalable to a country-wide visualization tool.

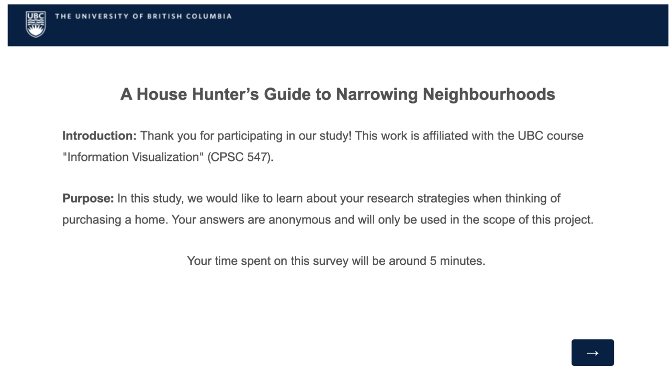

Fig. 15. General description of survey objective

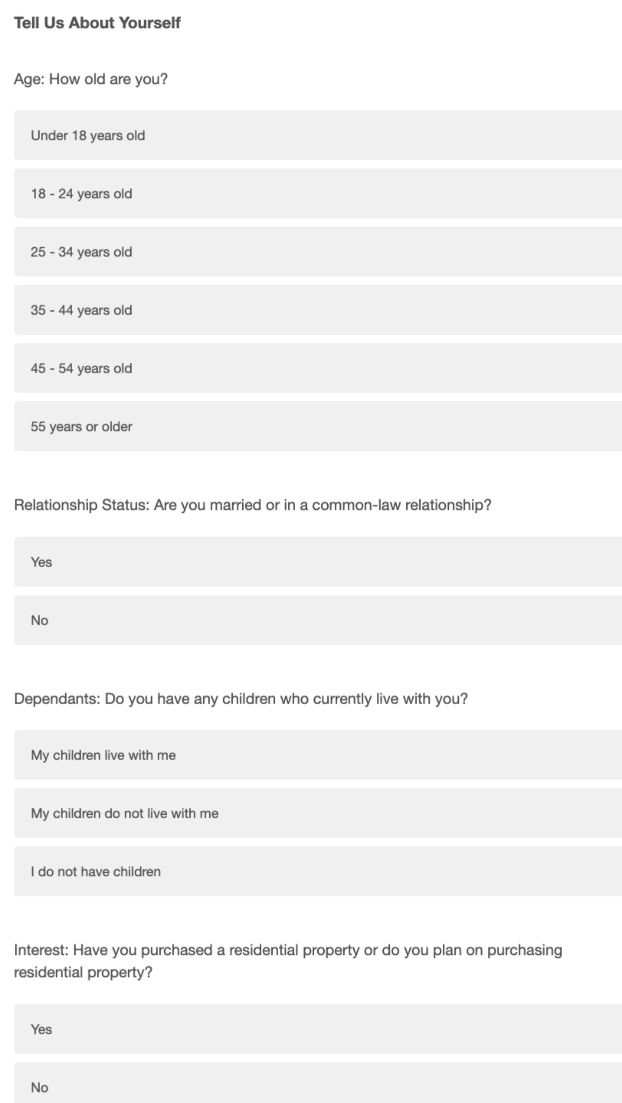

Fig. 16. Demographic information about survey respondent

10.1.1 Survey Results

Figure 18 represents the age distribution of survey respondents. To gauge whether respondents possessed any interest in purchasing residential property, respondents were asked if they had any interest in

Fig. 17. House search preferences of survey respondent

purchasing residential property or had purchased residential property

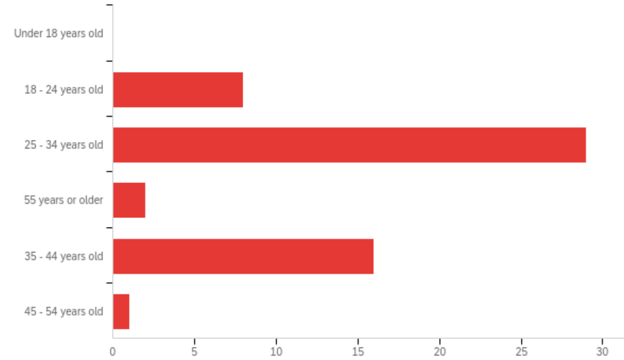

Fig. 18. Age distribution of survey respondents.

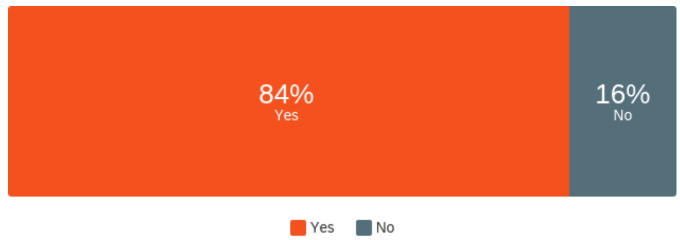

Fig. 19. Assessment of past, present or future intent to purchase residential property.

already. Figure 19 shows that 84% of survey respondents fit our primary target demographic.

Respondents were asked to rank the following criteria from 1 to 6 in descending order of their importance: cost of purchase, quality of schools, neighbourhood safety, proximity to public transportation, investment value, and access to parks and recreation. The results of this task are shown in Figure 20. As indicated in Figure 21, it is noteworthy that 73% of respondents did not have children. This demographic may explain the low importance ranking for the quality of schools criterion. To ensure that our criteria classes were not solely biased by prior works in the space, survey respondents were asked to share any other criteria that were relevant to the residential property search. Figure 22 depicts a word cloud encompassing these responses. As proximity to work and grocery stores were identified as recurring themes, we amended our original data to include a dataset that encompasses proximity to grocery stores, employment, health services, and more.

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Cost of purchase	1.00	6.00	1.88	1.21	1.47	56
2	Quality of nearby schools	1.00	6.00	4.43	1.55	2.39	56
3	Neighbourhood safety	1.00	6.00	2.93	1.47	2.17	56
4	Proximity to public transportation	1.00	6.00	4.07	1.64	2.67	56
5	Investment value	1.00	6.00	3.79	1.68	2.81	56
6	Access to parks and recreation	1.00	6.00	3.91	1.26	1.58	56

Fig. 20. Respondents were asked to rank the importance of each criterion when home hunting.

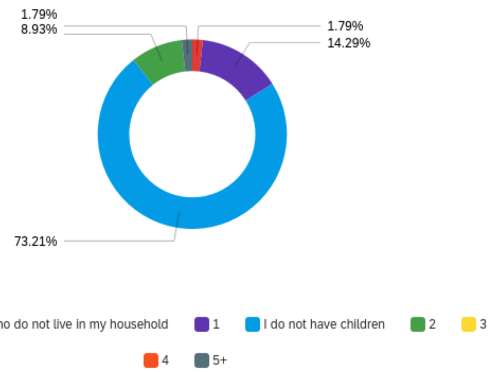

Fig. 21. Demographic data about respondents' number of children.

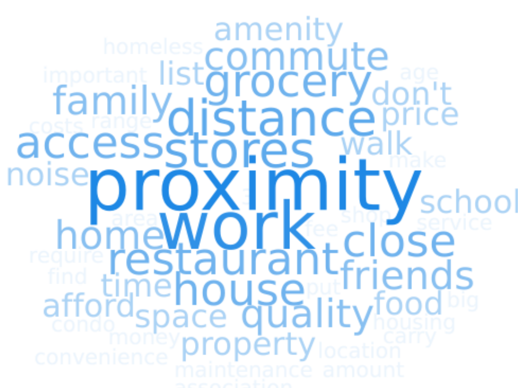

Fig. 22. Word cloud of responses from survey respondents about criteria not mentioned in Figure 5 that are important to them when choosing where to live.

Table 3: Milestone timeline

Milestone	Task	Estimaed Hours Per Person	Actual Hours Per Person	Estimated Deadline	Actual Deadline	Asignee
Pitch	Individual pitches	2h	2h	Sept 29	Sept 29	Everyone
Proposal	Idea selection meeting	1h	1h	Oct 7	Oct 7	Everyone
	Survey development and existing solution research	1.5h	1.5h	Oct 12	Oct 12	Ivan
	Pre-proposal report writing	1.5h	1.5h	Oct 12th	Oct 12th	Abi, Lucy
	Pre-proposal Meeting	3h	3h	Oct 13	Oct 13	Everyone
	Collect datasets	1h	1h	Oct 14	Oct 14	Everyone
	Define data and task abstraction	2h	2h	Oct 14	Oct 14	Everyone
	Proposal Write-up	3h	3.5h	Oct 21	Oct 21	Everyone
Update Report	Proposal Feedback Meeting	0.5h	1h	Nov 2	Nov 2	Everyone
	Analyze survey results	1h	0.5h	Oct 26	Nov 4	Everyone
	Evaluate pros and cons of proposed solutions and decide on an unified design	1h	3h	Oct 23	Nov 4	Everyone
	Design mock-ups for individual views	-	2h	-	Nov 8	Everyone
	Data cleaning, filtering, and normalization	4h	5h	Nov 8	Nov 11	Everyone
	Tool familiarization	10h	6h	Oct 26	Nov 11	Everyone
	Create initial implementation for individual views	10h	10h	Oct 26	Nov 16	Everyone
	Update writeup	2h	3h	Nov 16	Nov 16	Everyone
Implementation Deadline	Implement filter view with map	14h	15h	Nov 21	Nov 21	Abi
	Implement ranking view	14h	15h	Nov 21	Nov 21	Lucy
	Implement detail view	14h	15h	Nov 21		Ivan
	Integrate views and implement interactions	10h	13.5h Abi and Lucy; 3h Ivan	Dec 3	Dec 12	Everyone
	UI Implementation	-	13h	-	Dec 12	Ivan
	Bug fixing and clean-up	-	4h	-	Dec 13	Abi and Lucy
	Improvements to address feedback from Tamara	-	2h	-	Dec 13	Everyone
Feedback and Iteration	Meeting with Tamara for feedback on solution and presentation	-	1h	-	Dec 13	Everyone
Final Presentation	Finish presentation	3h	5h	Dec 14	Dec 14	Everyone
	Record presentation	-	2.5h	-	Dec 14	Everyone
	Record demo and edit presentation	-	2h	-	Dec 14	Ivan
Validate Implementation	Create evaluation plan	1h	0.5h	Dec 3	Dec 14	Everyone
	Recruit participants and conduct evaluation	2h	1.5h	Dec 6	Dec 15	Lucy and Abi
	Summarize results	2h	0.25	Dec 6	Dec 16	Everyone
Final Report	Finalize report	6h	6h	Dec 17	Dec 17	Everyone
Total	-	80h	95.75h	-	-	-