Multimedia: Visualisation of student residences in Leuven

Simon Chuptys, Niels Claeys, Jeroen de Coninck, Stijn de Haes
KU Leuven
Belgium
[simon.chuptys;niels.claeys;jeroen.d3.coninck;stijn.de.haes]@gmail.com

ABSTRACT
In this paper, the different stages in the development of an information visualization application will be explained. The application presents information about the student housing situation in Leuven. To start with, the different users that can benefit from a student housing application are identified by means of use case scenarios. From this starting point the evolution of the visualization is documented, together with design decisions and motivations. The relation to related work and a motivation for the chosen colors and for picking the size of the circles to illustrate the amount of residences are discussed. Lastly, an evaluation involving test users is discussed which concludes that the visualization is useful but there is still room for improvements such as showing more data or showing more characteristics of a residence.

Author Keywords
HCI; information visualization; data analysis;

1. BASIC IDEA

The visualisation described in this paper presents information about the residences of the students in the city of Leuven. The reason for choosing this visualisation is that, being students ourselves, we can relate to the topic: most of us had to search for a student residence and therefore it would be interesting to work on this visualisation. Furthermore, we are convinced that finding a suitable residence is a common concern amongst our fellow students, who will therefore be interested in the visualization.

The purpose of the visualisation is not to show individual houses and their prices but rather to aid the student to identify the best locations to start his search for a new residence. We offer them an interactive visualisation through which they can find patterns in the data, gain insights and, most importantly, find a residence. Possible patterns could be: only residences around the Oude Markt fit my budget range, all my fellow students have houses near their campus, all engineering students have houses outside the center of Leuven . . .

For every residence, the price, surface area, length of the contract, overall satisfaction, address, distance to the campus and the field of study is used (see section 'Housing Data' for more info). The student housing data will be shown on a map of Leuven. In order to aid the students in exploring the data, they can pick a preferred price and area range. Only corresponding residences will be shown on the map. This way, the user can see where most houses fitting his criteria are located. In order to further narrow the search, it is necessary that the user interacts closely with the visualization.

The visualisation can be used by different users, all with their own vision on the data, as explained below.

1.1 New student

The first (and main) usage scenario is for a new student who is looking for a residence in Leuven. This person doesn't know Leuven nor the typical attributes of student residences or locations of residences in Leuven. This user wants to get an overview of following attributes:

1. The **prizes** and **surface areas** of student residences.
2. Which **locations** are good for a student residence and which places are to be avoided (for instance, too expensive).
3. Where students from a similar **field of study** are housed.
4. Overall **satisfaction score**: what current residents think about their home (Are they satisfied with it? Would they rather have another residence?).
5. How long the **travel time** from the residence to the campus is.

We used these attributes because we found them most important when selecting a residence. In the evaluation section, we also present the opinions of other students on their preferred attributes.

The purpose of this visualisation is to help the student answering questions related to the attributes listed above and to guide him towards a suitable location for his new residence.

1.2 Settled student

This student already has a residence and he is curious to see if he pays a correct price. Maybe there are cheaper residences in the same area? This person will mainly focus on comparing his current residence with other residences. The student might want to give a review (or more data) of his current residence.

For settled students, the visualization will probably be used in one of the following ways: Either the student wants to focus on a very limited amount of residences or he wants to get a broad overview of all the residences in Leuven. Both ways of data exploration should be supported by the visualization.

1.3 Residence provider

This person is interested in the general distribution of the residence characteristics in Leuven. He wants to know which price he should ask for the residence in order to stay competitive, while not asking a price that is too low. The following information would be useful to him:

1. **Prices** of residences close by and more generally in Leuven.
2. **Surface areas** of residences close by / general for Leuven.
3. Other **accommodations** offered for these prices.

Based on the different use cases, we developed a visualization that tries to fulfil as much as possible the needs of the different users. Our main focus lies with the students because they will probably use the visualization the most.

2. HOUSING DATA

In order to be able to answer the questions discussed in the previous section, the visualization needs a lot of different data. This section discusses how this data is gathered and motivates the data selection.

2.1 Gathering the data

One of the difficult things was the gathering of the necessary information. We looked at a few different options:

Firstly, there exist many real estate offices that rent hundreds of student houses in Leuven. After the first contact, it was obvious that they did not want to help with this visualisation because they did not trust us with their data. There didn't exist a prototype so we could not gain their trust. This was a shame because they could also benefit from our visualization.

Next, there is the university of Leuven which has a lot of information about the students and their residences. They were a lot more cooperative and it was possible to get data from them. The only problem was that this data was not available during the first month of implementation. Therefore another approach was needed to gather some initial data and start our first visualisation.

The last option to gather the necessary data was crowdsourcing. The data was gathered by means of a form which was distributed trough social media to reach a big audience. This approach worked very well: within a week our data almost reached 200 entries which was more than enough to get started with.

2.2 The Data

In the end two sets of data were used: The KU Leuven data on one side which contains the following information: 1) Address 2) Price (monthly) 3) Contract duration 4) Surface area 5) Field of study of the current inhabitant. On the other side there was the crowdsourcing data which has some additional residence properties, being the mean travel duration to classes and an overall satisfaction rating (on a scale from 1 to 5). The KU Leuven data was restructured to have one row for each location with the corresponding residences and fields of study associated with it. There are approximately 500 locations which leads to information about approximately 4800 students and 1800 residences. Important to understand is that the data does not give a direct relation between the field of study and the residence which also explains the difference in size between them.

Our crowdsourcing data is stored in a Google Docs spreadsheet and is imported with MISO into our visualisation (see section 5). The KU Leuven data was received in 2 Microsoft Excel spreadsheets. These files are processed into an array of JSON objects which can be loaded directly in the visualisation. More information on the storage of our data can be found in section 5.0.3.

2.3 Why these data?

Since we envisioned our visualization from the beginning to utilize a map, gathering the addresses of the different residences was necessary. The main goal of the visualization was to help students in finding suitable residences. Conventional methods (most often used by real estate offices) to compare different residences typically rely on the price and surface area, hence this data was also included in our visualization.

Furthermore, the visualization also needed to provide guidance to new students who typically have little knowledge about the distribution of residences in Leuven. To aid them the fields of study was introduced in our visualization together with the travel distance to class. This gives new students the possibility to look where their classmates are housed, which can be used as a starting point.

The overall rating of a residence seemed interesting to us since it can both help students in selecting a residence and residence owners in improving their facilities (this data was only included in the crowdsourcing data and was therefore not used in our visualization).

3. DESIGN OF VISUALISATION & INTERACTION

The following paragraphs discuss the different stages in the development of the visualisation. During the process, weekly feedback was provided by means of a blog where people could experiment with and comment on the latest version of the visualization. Furthermore, co-students could evaluate the visualization during weekly presentations that were held in the multimedia courses.

3.1 Mock-up

The starting point is a static mock-up of our visualization (see figure 1). Our initial ideas were inspired by a similar visualisation about the housing prices in London [1]. From the start, it was clear that the map of Leuven should play a major role. Therefore, this map is the most dominant visual element in our mock-up. The user should be able to zoom and pan the map, allowing to switch between a more global or detailed view. The location of student residences is indicated by small circles on the map. Three buttons on the map allow the user to change the color of the circles to denote either price, surface area or price per square meter. The circle radius is proportional to the number of residences that are available at the same location. To allow the user to filter the displayed residences, sliders are added that control the price range and surface area range. Furthermore, the travel distance from a specific location to the residences can be investigated. This option will create an overlay mask on the map, indicating which residences lie within the specified travel distance. Initially, the travel distance was delivered by the crowdsourcing data, afterwards Google maps API provided more accurately results (as discussed in 6.2).

Our main concern at this point is what should happen when a large number of residences are displayed on the map: Will the view be too cluttered? Is it needed to aggregate residences that are close to each other? What color would this aggregation have...? An implemented prototype version which uses actual data will give us more insight in these matters.

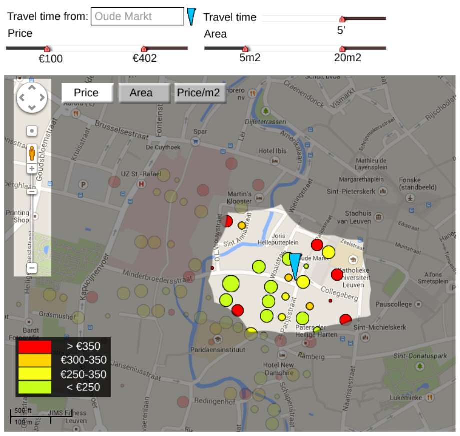

Figure 1. The static mockup of the visualisation.

3.2 V0.1

Based on the mock-up, an initial version was implemented. Figure 2 shows this visualization based on the data gathered from our crowd sourcing efforts. In this first version, we focussed on the map display, the color coding of residences and the filtering using sliders. Based on this version our initial concerns were unnecessary: the individual residences remain distinguishable, even when showing the full map of Leuven. A problem that arose was the scale, that was used for the color coding. The colors scale linearly in the range specified by the sliders. Because our data contains a few outliers, they make sure that almost all residences have the same color. One could argue that, if the user wants a more interesting color coding, they should change the sliders to exclude the outliers. However, we would like to present the user with an overview that still provides a useful color coding. Therefore some alternatives to linear scaling were tried out in future versions. The comments of our co-students on this issue confirmed our concerns.

3.3 V0.2

In our second version, shown in figure 3, a couple of subtle adjustments were added. First of all, pop-ups are introduced to give extra information about a location when hovering over it. This way the user can see the average price of the residences at the selected location together with the surface area and the number of residences. Furthermore, the color legend is now incorporated with the rest of the GUI, so that it displays the range of values corresponding to each color.

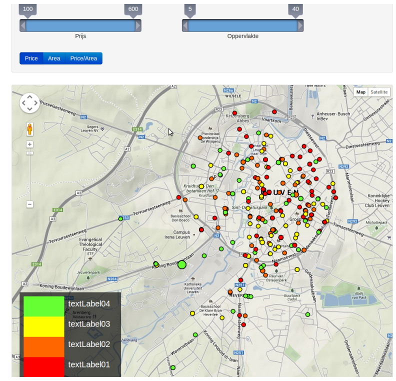

Figure 2. Version 0.1

We also experimented with the color-scale: Instead of using a linear scale, a quantile scaling was implemented, meaning that each color contains about the same number of residences. Two different opinions exist on the results:

- On one hand, the ranges on the color legend now give a better understanding of the distribution of residences in the selected price and area ranges.
- On the other hand, the map now always shows an equal number of circles of every color, losing this same insight in the distribution on the map. This is an aspect we found problematic.

Because we found that the linear scaling represented the data in a better way, it was chosen for the visualization. A related issue that arose during a discussion with our co-students was that, when dragging the sliders, suddenly a lot of residences appeared or disappeared because of the distribution of the residences. The user should be presented with a way to tell if a lot of residences lie in the specified ranges. This issue is addressed in the next version.

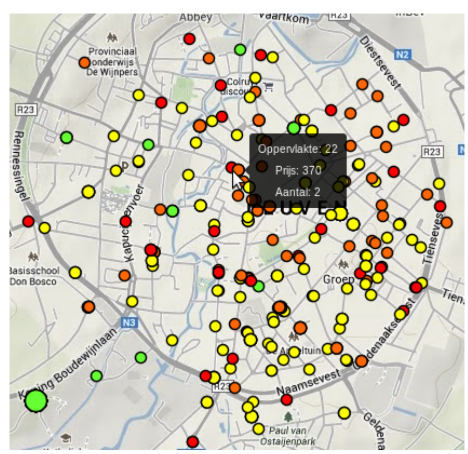

Figure 3. Version 0.2 (cropped)

3.4 V0.3

In the third iteration, visible in figure 4, extra functionality is added: bar charts and distance calculation. The bar charts above the range sliders show the distribution of the prices and surface areas. With this information, the user gets a better notion of how many residences exist with the specific price/surface area, solving the issues that arose during the previous version. More information on the distance calculation can be found in section 6.2.

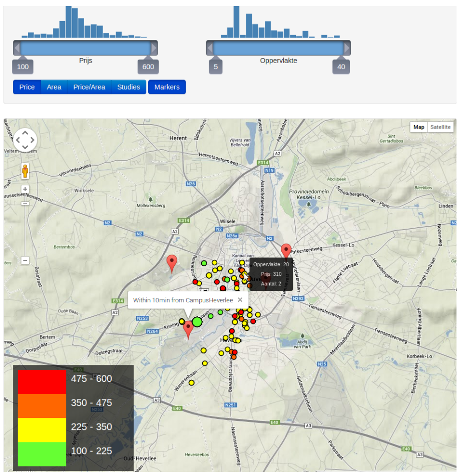

Figure 4. Version 1.0

3.5 V1.0

In the final iteration, shown in figure 5, the data received from the KU Leuven was integrated into our visualisation. This means that a lot more data is shown, requiring us to use JSON (more information in section 5.0.3) instead of a google spreadsheet to store the data. Another improvement made in this iteration is to show information of individual residences at one location (shown in the upper right corner). This shows 2 bargraphs, representing the price, area and a stacked graph illustrating the distribution of the fields of study when the user hovers over a circle. This was a necessary improvement because in the KU Leuven data, one location often represents multiple residences. The map only shows the averages over all residences on a specific location, but the extra view also shows the individual characteristics of each residence.

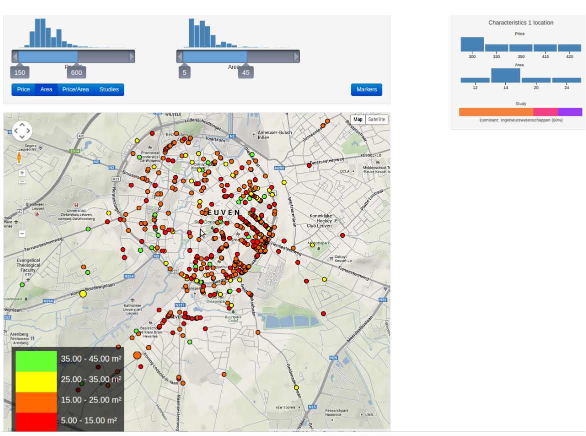

Figure 5. Version 0.3

4. THEORETICAL BACKGROUND

In order to create a good visualization, there has been done a lot of research on different visualization aspects. In this section some of this research will be highlighted and applied to our visualization.

4.1 Information aesthetics

When developing an application that visualizes information, it is important to understand what the focus of the visualization should be. A visualization can be situated at any point in the quadrant created by *data focus* and a *mapping focus*. [7] defines and discusses these different focus groups (see figure 6).

- The horizontal axis identifies the methods used to map the data into visual form. Direct mappings give the user the ability to infer the data from the visual representation. Interpretive mappings often involve stylistic influences.

- The data focus (on the vertical axis) wants to facilitate the communication of information. An intrinsic focus provides insight into the data. Extrinsic focusses on the communication of the meaning related to the underlying data.

Based on the techniques that are used, different kind of visualizations will be created. In this framework, our visualization is situated in the bottom left corner. This can be inferred by the goal of our visualization. The idea is to help users to get insights into our data and to display this data as useful as possible. This defines the positioning of our visualization.

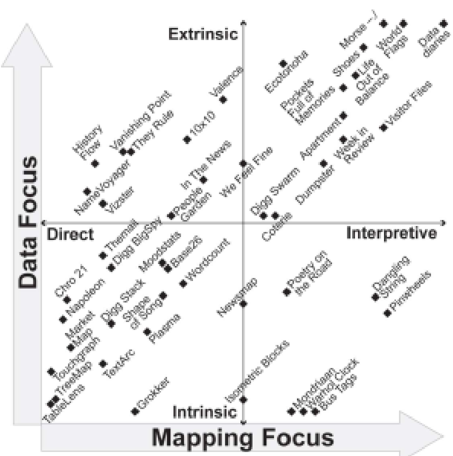

Figure 6. This illustrates information aesthetics of different visualizations.

4.1.1 Relevance to our visualization

Data Focus Our visualization has primarily an intrinsic data focus because it wants to help people find an interesting region to live. This way the visualization wants to provide insight into the underlying data.

Direct mapping We mainly use a direct mapping approach because the underlying data can be inferred by the color of the houses. This gives the user the necessary tools to find an interesting region where the housing is to his liking.

Extra Additionally, our visualization promotes interactivity. This interactivity wants to help users get more insight into the data which enables them to narrow down the interesting regions.

4.2 Human perception

When designing a good information visualization, one needs to understand how humans perceive what they see. For example, our vision is sensitive to changes in color and is trained to identify patterns in visual data. [5] discusses these characteristics and we will apply them to our visualization. The paper argues that a good visualization should:

1. indicate the relation between objects: All the circles represent different houses. Based on the highlighted button users can see which attribute is shown,

2. represent quantities accurately. The colors show an approximation of the values but hovering over the circles displays the exact values,

3. make it easy to compare quantities. The different colors make it easy to compare the values,

4. make it easy to rank the values. The ranking of values can also be easily performed based on the colors. It is intuitive to use red circles for expensive houses and green for cheap ones which makes the visualization easy to understand,

5. make it clear how the information should be used. The sliders, buttons and the ability to zoom make it obvious that the main idea is to play with the visualization. Only by using it, can the patterns be seen.

4.3 Inspiring visualizations

One of the most inspiring visualizations was [4], which describes the liveability of Melbourne. This is because the visualization uses similar data and because it illustrates the data in an interactive manner which leads to a better understanding for the user. One remark on the visualization is that it is difficult to understand the scoring scheme which is something that we must avoid. It was not clear to us how the scores were defined which leads to a less effective visualization because the user is not sure what a color represents.

Another inspiring visualization shows the Irish population in the US[3]. This visualization is nice due to the simplicity of only using one color. It is also a great example of a visualization that displays data on a large scale. One minor issue of the visualization is that it only uses one attribute which is easier than showing multiple attributes on the same map.

The next visualization that inspired us displays information about meteorites[9]. The good things about this visualization are the way in which the circles change when zooming in and the presentation of the bar chart. The bar chart describes nicely when a meteorite was found or hit the ground.

Lastly our visualization was also inspired by a visualization of the housing prices in Slovenia [2]. The data used is similar to ours. Furthermore zooming in is nicely handled in the visualization. A minor remark is that the visualization uses to many colors for specifying the prices of the different residences.

4.4 Theory on creating a visualization

In this section the article discusses how well our visualization follows the theory behind visualizing data.

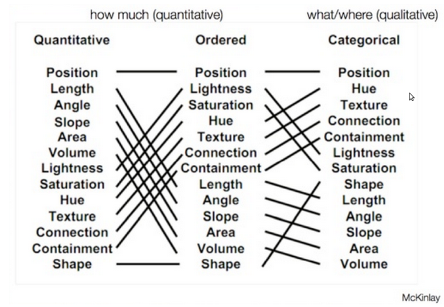

Figure 7. Techniques to display attributes in a visualization.

Figure 7 shows what techniques work best for displaying different sorts of data. In our visualization both quantitative data (prices, areas, number of residences) as well as categorical data (field of studies) are used. For the number of residences we use area. Based on the figure this is not the best choice but it's not too bad a choice because position was already taken. The different characteristics: area, price, price over area and studies, are represented as categorical data (for example: price ranges). These categories are represented by hue which is an excellent choice given the represented figure.

Another guideline in information visualization is:

"Overview first, zoom and filter, then details-on-demand."

which was stated by Ben Shneiderman [8]. Our visualization follows this principle perfectly. First the users are presented with an overview of all the residences in Leuven. Afterwards they can filter the residences in which they are interested in and hovering over a particular residence, gives them in-depth details.

Lastly there also exist the Gestalt principles which describe some patterns that are good to use in a visualization. Our visualization shows primarily the law of similarity and the law of isomorphism. Similarity is illustrated by the different circles, that represent the different houses. Isomorphism is used for the color coding, for example when talking about prices red is expensive and green is cheap.

5. SOFTWARE DESIGN

The software that displays our visualization and makes it interactive is written in JavaScript. With the help of a couple of different libraries (D3.js, jquery), our code is restricted to 2 files: gui.js and map.js. The first file handles our controls (buttons and sliders) that are able to update our visualization together with the bar charts. The code for interacting with the map resides only in map.js. There are two versions of the visualisation: one for the crowdsourcing data and one for the KU Leuven data. The crowdsourcing version loads its data straight from the Google Docs spreadsheet where it is stored and accumulates data for residences in the same location on the fly. This was infeasible for the KU Leuven version because this dataset is much larger and because there is no direct relation between the field of study attribute and the price and area attributes. The KU Leuven data was preprocessed instead (see further down).

5.0.1 Map.js
The code starts off by creating the map with the legend on it and afterwards the dataset is imported. When the data is ready, the callback function is executed. This function creates the DOM elements on the map and ties the data to them. It also sets up the function that keeps their position in sync with the map. Furthermore this file contains the functions to color the earlier created DOM elements according to their attributes, to place the markers for important locations, and in the case of the crowdsourcing data also to accumulate the data for residences in the same location.

5.0.2 Gui.js
In this file the interface elements are tied to JavaScript code. Using the functions defined in Map.js these can then interact with the map and the DOM elements on it. This file also contains the function to filter out residences by specifying a price range, an area range and a maximum distance to one of the important locations. The code for creating the various barcharts in our visualisation can be found at the end of the file.

5.0.3 KU Leuven data
The KU Leuven data was delivered to us as two Microsoft Excel spreadsheets. We merged these into one and accumulated the residences in the same locations by converting all addresses to coordinates and using these to match residences. We exported the spreadsheets to comma-separated value text files and have written some parsers to get the matching and formatting work on these files. The end result is a JavaScript file containing an array of JSON objects. There is one object for every geographical location which has an array holding price-area pairs and an array holding the amount of students for each field of study at that location. The distance to the important locations and the name of the dominant field of study are attributes as well.

5.1 Tools used
1) **Data importing** is done through Miso[1]. This library is used because it allows us to import our crowd-sourced data directly from Google Docs with just a few lines of code. When using the data from the KULeuven, the data was moved to a JSON file. It was possible to keep using MISO for reading this JSON file.

2) **Data handling** is done with D3.js. D3.js ties our data to DOM elements and updates them. This framework was a logical choice since every member of the team has familiarized himself with D3 at the beginning of the course.

3) As a **geospatial API**, the Google Maps API is used. Google Maps has a well documented and extensive geospatial API. Thanks to Google maps it is easy to render a map with zoom & pan controls, and draw markers on it. Furthermore Google maps also provides support for the calculation of travel distances and the conversion between addresses and coordinates.

4) JavaScript is simplified using **jQuery**. jQuerys tagline is "Write less, do more" which is certainly correct from our point of view. To write even less, the visualization uses two jQuery plugins: Tipsy for the tooltips and JQRangeSlider for our slider controls

5.2 Opinion on d3.js
After working with d3.js for a semester, we have a good idea about the benefits and drawbacks of d3. D3 is useful and convenient to work with once you overcome the the initial learning curve. Certainly for people who are new to Javascript it is a bit of a struggle. The power of d3 lies primarily in the easy visualisation of data, certainly when this data changes.

In the beginning, d3 seemed a bit like a burden because we only wanted to display the circles on a map and this functionality is also offered by Google. The placement of the circles was not so easy in the beginning, also interacting with the map and the circles was cumbersome at first.

In a later stage, we found out that d3 has a lot of advantages, especially when implementing the different bar charts. The bar charts that change with hovering over different locations can be easily represented with d3. The necessary thing to know is that d3 makes a distinction between entering data, leaving data and changed data.

Another important advantages of d3 is that it has a good user community, a lot of examples for every problem can be found here.

6. IMPLEMENTATION ISSUES

6.1 Housing department data
One of the issues was incorporating the data received from the housing department of the KULeuven into our visualization. They provided us with two files, respectively containing addresses and studies or prices. The mapping of these two files was done on latitude and longitude coordinates because this is relatively simple. This conversion was done through a tool which supports the retrieval of thousands of address locations [2]. The problem here was that the coordinates were sometimes not accurate enough, which meant that a lot of the initial data was lost. Because the most significant loss of data was in the file that linked the studies to the addresses and this data is less critical to our visualisation we chose not to revisit the matching. It would however be a good thing to do when there is some extra time. Around 1800 residences are left, which is enough for providing a good representation of the residences in Leuven.

6.2 Distance calculation
For distance calculation, the first option was to let the user place markers on the map and enable him to specify the maximum travel time. Based on this information the residences

[1] `http://misoproject.com/dataset/`

[2] http://smartystreets.com/products/liveaddress-lists

that lie further away will be filtered out. However, the response times are only acceptable for small datasets. Furthermore, the Google maps API also limits the number of simultaneous distance calculation requests allowed, so our initial idea became infeasible.

We therefore opted for a different approach. A number of locations that are important for students living in Leuven (like the campuses of the different fields of study) are selected and the travel time to each of these location is pre-calculated. For this visualization, 4 important locations were selected: the train station, campus Heverlee, Gasthuisberg and GroupT (all are locations frequently visited by students). These locations are displayed by markers on the map. The user can select one of these markers and the residences outside a 10 minute travel time are filtered out. A button to reset/disable/enable the markers is also included.

6.3 Working in group

During the implementation, there were also some issues with assigning different jobs to everyone. Because only 2 files are used in our representation of the data, it was difficult in the beginning to work on the visualization independently. Initially we only had 2 tasks: one group worked on the visualization and the other on the user interface. Later, when more functions were added the assignment of different tasks to people became less of a problem. This was mostly because almost all functions were independent and only required the data to be available, making it easier to work independently from each other.

7. EVALUATION

7.1 Empirical studies in information visualization

Paper [6] proposes 7 scenario's to evaluate a visualization, only the ones related to our visualization will be discussed. The focus for our visualization lies on getting additional insights and not on user experience or usability.

7.1.1 Evaluating Visual Data Analysis and Reasoning

This scenario assesses the visualization's ability to support visual analysis and reasoning about the data. The following questions should be answered:

1) Data exploration: How does it support processes aimed at seeking information, searching, filtering, reading and extracting information? Our visualization provides sliders to set price and area ranges. Also, the user can choose to either use color coding for price, area or price/area. While the sliders have the purpose of narrowing down the dataset, the colors are meant to quickly get an idea about the selected parameter. The actual data (the numbers) can be accessed on a per-residence basis by hovering over the residence.

2) Knowledge discovery: How does it support the schematization of information or the (re-)analysis of theories? In our case, theories like residences in noisy areas are cheaper can easily be verified by having a look at the color-coded map (provided the user knows which areas are noisy).

3) Hypothesis generation: How does it support hypothesis generation and interactive examination? By examining the map of all residences, one might find patterns in the colors that are displayed to form hypotheses.

4) Decision making: How does it support the communication and application of analysis results? Residences that do not specify the users needs are filtered from the map by means of the sliders. Which of the remaining residences are to be inspected in more detail is up to the user, who can hover over them to acquire specific data.

7.1.2 Evaluating communication through visualization

this scenario tries to reveal whether the visualisation helps users in understanding the data. The main question that should be answered by our visualization is the following:

Do people learn better/faster using the visualization tool? Our tool will help students with the hardest part of finding a residence, which is getting to know what the options are. Using our tool, students will get an idea about typical price and area ranges. Furthermore, they get an overview of which residences are worth a visit, which is important to know since the time span that residences are open to public is typically very limited. An evaluation of our visualization with actual test persons will assess if our visualization succeeds in the goals we have in mind. When evaluating the visualization, our main focus lies on the *added value* for the intended group of users. However, at the time the evaluation took place (December) , it was difficult to find students that are looking for a new residence. Therefore, during the evaluation we asked test subjects to impersonate a person searching for a new residence. Furthermore, our evaluation does not assess the *usability* of the developed application (although this is an important aspect to evaluate), because it is not the focus of this project.

7.2 Test setup

Two tests were performed to evaluate our visualization: a face-to-face evaluation and an anonymous questionnaire. The face-to-face evaluation was performed with 9 test subjects: 2 students with computer science background, 4 students from other fields of study and 3 parents of students. The test subjects were both students and parents since we expect that these are also the type of people who will primarily use our visualization. The anonymous questionnaire was filled in by the same people, while also being shared online.

During the face-to-face evaluation, test users were introduced to our visualization by just stating that it shows student housing in Leuven. Then, the following questions were asked:

1. (main question) Find one (or a limited number) of residences that suit(s) your needs, using the visualization.
2. Where are the most expensive residences located?
3. Which area has the largest concentration of law students?
4. Give two characteristics of residences close to gasthuisberg.

The first question gives us an understanding in how users search for a new residence. The other questions encourage the test person to explore different functionalities of our visualization, so that we can assess if these functionalities indeed help the user in exploring the data. During the test, users were asked to think aloud, enabling us to better understand their actions.

Afterwards, the test subjects were asked to fill in a questionnaire, anonymously and without interference from the test coordinators. The questionnaire tries to assess the added value of our visualization through the following questions:

1. The visualization helped me to choose a suitable student housing.
2. The visualization provides insight into the characteristics of the student housings in Leuven.
3. The visualization helps me to verify hypothesis about the student housings in Leuven.
4. I think it is fun to use this visualization and to learn more about the student housings in Leuven.
5. I would be willing to pay the following amount to use this visualization if the complete list of student residences were available.
6. Choose 3 parameters that are most important for you when searching for a student residence: price, the other residents, number of other residents, area, distance to class, price over area, accommodations, fellow students of your discipline.

7.2.1 Test results

The face-to-face evaluation (involving 9 test subjects) enabled us to draw the following conclusions:

1. When trying to find a suitable residence through using our visualization, a lot of different behaviours were observed:
    - Two persons tried all the different functionalities that our visualization offered and repeatedly switched between them to make a decision.
    - Two persons started their search based on the location of residences (either close to the station or close to the campus). Other functionalities were used to refine the selection. (Remark 1: neither of them found the 'filter by travel time' functionality on its own) (Remark 2: two of the test persons tested with a version where this functionality was not present).
    - One person didn't use any of the filtering or colouring functionalities, only the positions of the residences on the map.
    - The four remaining test persons used some of the functionalities, based on what they thought were important selection criteria.
2. When answering questions that required the test persons to utilize specific functionalities of the visualization, they performed rather well: they were able to determine which functionalities would lead them to the answers and they had little trouble making a conclusion based on what the visualization showed them.
3. Additional comments from the test persons:
    (a) Four persons mentioned they enjoyed playing with the functionalities to explore the data.
    (b) Two persons reflected on their own (current) residence (eg 'quite expensive compared to what I see here')
    (c) Three persons mentioned the 10min-travel time filter would be more useful if it was limited to 5min. (Three out of the seven persons that tested this functionality)

The anonymous questionnaire gave following results (the numbering corresponds to the the numbering of the questions discussed in the previous section):

(0 = completely disagree, 5 = completely agree)

1. Choosing a residence: range 3 to 5, average 4
2. Understanding characteristics: range 4 to 5, average 4.6
3. Verifying hypotheses: range 3 to 5, average 3.9
4. Fun to use: range 2 to 5, average 4.2
5. Average price willing to pay: 1.2 Euro
6. Most important characteristics: 1)accommodations 2)price 3)price per surface area

We interpret these results as follows: Our visualization does a good job giving people more insight in the data. Also, if people were to find a new residence, they would benefit from using this visualization. When choosing a residence, it turns out that the field of study is not considered important (this characteristic was not chosen by any test person as important). However, we still think it a good idea to include it in the visualization, since our face-to-face testing revealed that it is one of the characteristics that is most used when exploring the data (for fun or out of interest). The characteristic that the questionnaire shows to be most useful is 'offered accommodations'. The reason our visualization doesn't have the functionality to filter characteristics is that this data is hard to come by. Also, it is non trivial to visualize this data, since a lot of variation exists and it is not standardized, which makes is difficult to compare on a large scale.

8. CONCLUSION

This paper described the creation of a visualization that shows student housing in Leuven, using D3.js and Google maps. The relevance of the visualization towards the different users was illustrated. The evolution of the visualization is described through different stages where adjustments were made based on the feedback received by different students. This paper also discussed the rationale for the chosen colors and why the area of the circles represent the number of residences at a given location among other design decisions. The evaluation showed that the visualization helps in the search of a new residence: it accommodates different methods of data exploration and offers relevant informations on student housing. However, the evaluation also showed that there is still room for improvement: a major missing feature is information about student housing accommodations, a non trivial but interesting addition that could be developed in the future.

9. REFERENCES

1. Housing prices in london. `http://www.mysociety.org/2007/more-travel-maps/morehousing`. Accessed: 2013-12-10.
2. Housing prices in slovenia. `http://visual.ly/map-real-estate-prices-slovenia-m2?view=true`. Accessed: 2013-12-10.
3. Where irish-americans live. `http://trends.truliablog.com/vis/irish-america/`. Accessed: 2013-12-10.
4. Datalicious. Melbourne liveability. `http://www.theage.com.au/national/greater-melbourne/our-liveable-city/map`. Accessed: 2013-12-10.
5. Few, S. *Data Visualization for Human Perception*. The Interaction Design Foundation, Aarhus, Denmark, 2013.
6. Lam, H., Bertini, E., Isenberg, P., Plaisant, C., and Carpendale, S. Empirical studies in information visualization: Seven scenarios. *Visualization and Computer Graphics, IEEE Transactions on 18*, 9 (2012), 1520–1536.
7. Lau, A., and Vande Moere, A. Towards a model of information aesthetics in information visualization. In *Information Visualization, 2007. IV '07. 11th International Conference* (2007), 87–92.
8. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, VL '96, IEEE Computer Society (Washington, DC, USA, 1996), 336–.
9. Torre, R. Macrometeorites. `http://visual.ly/macrometeorites?view=true`. Accessed: 2013-12-10.