Information Visualization 2015 - Project Report

Xiaoli Li (10874968)
Floris de Bruin (5772583)
Minh Ngo Le (10897402)
Sebastian Dröppelmann (5783453)

March 30, 2015

Contents

1 Concept

1.1 Topic

The aim of this project is to help people find a spot to buy a house within the municipality of Amsterdam. To achieve that we visualize the leisure infrastructure that exists in Amsterdam, since since they are most likely interested in buying a house in regions with leisure places of their desire. A user should be able to interactively explore the different regions of Amsterdam and gain information about the distributions of different types of leisure that exist in these regions. The representation should also be visually appealing and intuitive for the user. Also the user should be provided with the possibility to compare different areas. to easily spot the more interesting areas.

1.2 What other Work has

A representation of the basis of the dataset can be seen on the Opendata website. [12]. This however is no representation of leisure places over a city, just another representation for the data we used. Another representation of leisureplaces can be found on a map of the city of Vienna. The representation there simply shows the different places of different types of leisure. [8]

1.3 Where this Project differs

This project differs from the other parties in several ways. Instead of directly showing the different places of leisure as different icons, this projects first represents the areas by score for the different types of leisure. This leads the user directly to the interesting areas per type. Also the user is offered the ability to compare the different regions that he finds interesting. The representation of different types of leisure places is also shown as POI's on the map, but also can be reduced to the regions that the user finds interesting.

2 The Dataset

2.1 Data Gathering

For the purpose of finding leisure places in Amsterdam a dataset containing all sport grounds like public Basketball courts and Football fields was used. Furthermore, a list of nonresidential functions of buildings in Amsterdam was used to find different groups of leisure buildings over the city. The last dataset used for this task was a representation of all parks and public green areas in Amsterdam. All three sets are provided by Open Geodata [13].

Two more data sources where used to fill the regional information. The first supplied the region polygons to draw the postcode regions on the map. [5] The second provided information about the average price per m^2, the average surface area of a living unit and the average price per living unit. [9]

At last one more dataset was used to bind the information of the street names and housenumbers to the correct postcode. This data was downloaded from postcodedata.nl [3]

2.2 Data Processing

From the datasets mentioned above, two final datasets where built up. One containing the information about the different buildings of the different leisure types and a second with information about the regions such as the polygon to draw the region, the weighted score for each type of leisure or the average price per m^2.

2.3 Preparation

2.3.1 Grouping

First the data needed to be grouped to different categories of leisure so that it could be represented in a meaningful way without being to rough or too fine grained. After analyzing the data the following groups where chosen.

- cultural (museum/monuments/historical)
- green (parks/lakes/gardens)
- culinary (restaurants/coffees)
- sport (gyms/sport fields)
- relaxation (spas/saunas/massage salons)
- spiritual (churches/temples/shrines)
- activity (clubs/bars/dance halls/cinemas)
- nonleisure (everything that does not contribute to leisure)

2.3.2 Building Information Dataset

The Data about the pars, buildings and sport places where aggregated into one dataset to represent the different places on the map and to gather statistics about the different types of leisure within the different regions. The preprocessing included at first to gather the important pieces of information out of the building dataset. The information collected from the dataset contained:

- street name
- housenumber
- name of building
- categorycode
- categorydescription

Since the dataset did not contain the postcode directly, the street name and housenumber information was used to derive the postcode through usage of the postcode dataset. Since the original dataset just contained x and y coordinates of some map, but no geodatapoints, the latitude and longitude of the part of the street in the postcode where used. This was also supplied by the postcode dataset. The categories supplied per building where also mapped to the selected leisure categories. For a full view of the mapping look at Appendix I. During data processing the entries of the dataset that had no information on the street and could therefore not be mapped correctly to the regions where left out. Also some buildings that had street names but no housenumber where mapped to the street, since except for a few streets in Amsterdam they mostly reside in one postcode area. This could present a minor distortion but not more than a few buildings. From the total 29000 buildings roughly 1400 got pruned because of missing information.

For the sports dataset the name, category description and geocoordinates could be gathered directly from the dataset. The category was set to sport for all 350 places in the dataset. Finally the geocoordinates where used to find the polygon of the region they resided in, and then where given the postcode of that region.

During parsing the park data the category-description was set to park and the category was set to green. the name and geocoordinates could directly be extracted. Also the size of the park was kept for calculation of the score, which is described in the next section. The park dataset contained 87 entries.

The final dataset contained roughly 29000 entries with following information:

- 4 digit postcode
- name of building
- leisure category
- category description
- latitude
- longitude

2.3.3 Region Information Dataset

For the regional information the dataset containing the average price, size and price per m^2 could be read directly from the dataset. Also the region size and polygon information could be extracted directly from the mentioned dataset. Finally the scoring of the categories had to be calculated. For the building data and the sport places, the scoring was done by summing up all locations of a certain type in the region and then dividing that by the size of the region. This way the average count of leisure buildings per m^2 was calculated.

For the parks scoring was calculated by taking the total amount of park area in a region and dividing it by the size of the region.

After these scores where calculated they where rescaled to values from 0 to 100 where 100 was the highest value of all regions and 0 was 0.

This dataset contained the following data:

- 4 digit postcode (categorical)
- score per category (interval)
- total surface per region (interval)
- average livingunit price (interval)
- average livingunit size (interval)
- average price per m^2 (interval)
- geometrical data for drawing polygons (spacial)

2.4 Integration of the datasets

All the required data is made available via a REST webinterface that is called from javascript. This way js is only busy updating data that is processed on the server and can be called and loaded on demand. For communication JSON is used.

2.5 Data Quality

The dataset relates to the requirements described in Stephen Few's *Now you see it* [7] in following form:

- **High volume** In total around 29000 locations where incorporated into the final dataset, with the information about their categories, sub-category descriptions, names and locations. Since the original dataset contains almost every non-residential building in Amsterdam this can be considered high-volume.

- **Historical** the main dataset ignores the historical aspect, since there is just one point of measurement in time of each location in the dataset.
- **Consistent** Since the dataset is collected over a longer period, the date of collection had to be ignored to achieve a certain amount of consistency in the dataset. Also data that could not be placed spatially, due to missing address information had to be ignored during the build up of the final dataset. With leaving out all incomplete information it is ensured that the final dataset contains all necessary information for each entry. Therefore it is consistent.
- **Multivariate** the variety of data is limited but due to the use of two different types of datasets, one on regional levelm and one more fine-grained, the variety of different data is high.
- **Atomic** Since the smallest level of information that can be reached are certain shops on locations, this information is the base of the dataset. The data is first presented as aggregate data of the counts of each category, but can then be explored till the level of each building individually.
- **Clean** Unnecessary data from the original datasets was removed, and it was ensured that only complete and uniform data exists in the final dataset. This way the dataset can be considered clean.
- **Clear** the labeling of the data is easily understandable as it contains latitude and longitude as geocoordinates, the main categories as English words, the sub categories as complete descriptions in dutch (which should have been translated), the postcode as the normal dutch four digit version without the letters and the name of each place as text. This way the compete dataset is easily understandable and interpretable.
- **Segmented** The data is segmented in different ways. one segment is the spatial segment into different postcodes, the other segmentation is the one in several types of leisure. The types of leisure also have sub categories, which are presented on showing information of specific buildings but which could also be used to group the data differently and more fine grained. The locations can also be grouped by the geocoordinates, provided per location.
- **Known pedigree** The data comes from the local government of Amsterdam or sources that work together with the local government. There is information available on when the data of the different locations was collected.

2.6 The use of the data

There are two different use types for the data, one is the representation of the data on a regional level on the map, the other is the representation of the different buildings of certain types.

First the polygon data is drawn on the map. The scoring of the different categories is accessible in different ways. First as choropleth map-layer where we color the regions in the different scores for the chosen leisure-categories. Secondly as the different axes of the star plot and the star combination plot. The score can also be seen upon hovering over a axis. The other regional information is accessible through clicking on the star plot of the region of which more information is desired. Finally the building data is shown on the location of the geocoordinates. The category-description and the name of the building are shown upon hovering over the POI marker on the map. This way all the information in the two datasets is accessible, and can be viewed according to the interest of the user.

The final product can be seen on root.org.ua [2].

3 Story Telling

In this section we describe our design in a story telling, with screenshots from the demo implemented. In a short explanation beforehand we used all of the different story telling models to a certain amount.

The slideshow can be found in the way that we have an actual slide show as introduction, but also in the way that you would first have the help slideshow, which can be seen as the first slide. then the user can explore and compare as a second slide, and then at last go to a place where he can buy a home as final slide.

The second story telling method would be the drill down method which can mainly be found in the fact that the user can start with a general comparison of the areas with the help of the leisure filters over all regions. Then compare the data as a drill down step on all levels of the leisure categories and finally explore single POI's on the map mode.

The last storytelling method, but also the one that our project has most similarities with is the martini glass model. First the user is guided through the usage of the project, and then set loose to explore the different regions, compare them and play around with them until he finally picks the cherry and goes further and buys a house.

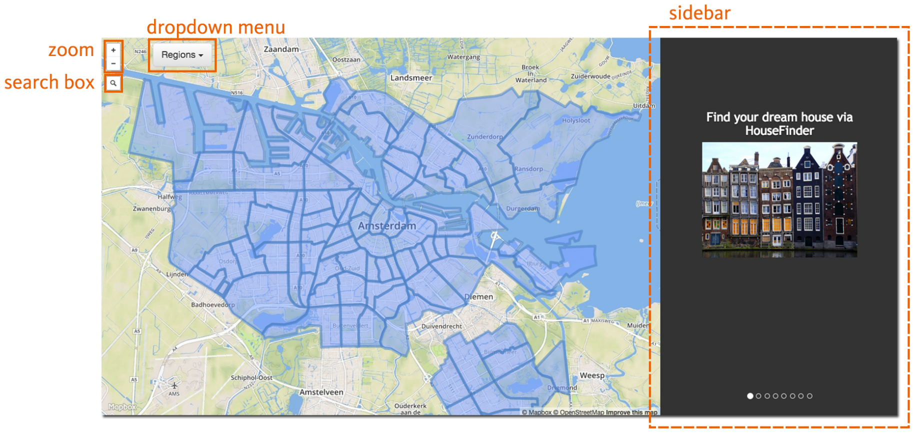

Figure 1: Initial view

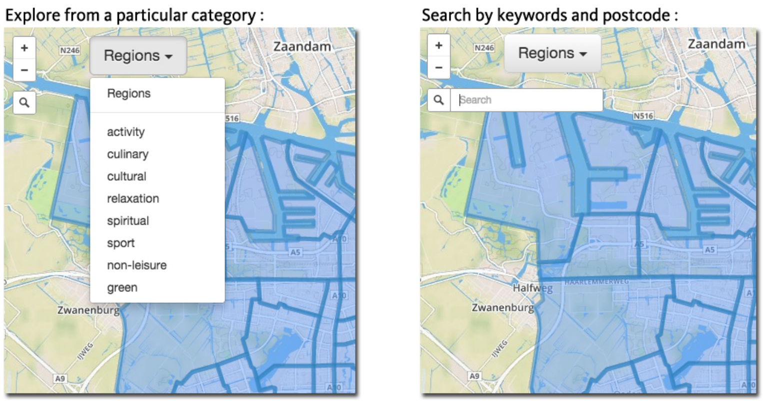

Figure 2: Drop-down menu and search box as starting points

Multi-view As can be seen in Figure 1, the interface consist of two views. The main view is built on the top of a map and it illustrates geographical information of regions in Amsterdam. On the right there is a sidebar. Initially the sidebar gives a brief tutorial, while during the use of the program, it shows individual star plots of the chosen regions, as well as a combined bigger star plots diagram for users' convenience comparing characters of the chosen regions.

Multiple starting points of exploration At the beginning of exploration, users can either select from the drop-down menu for a certain category of leisure places or input the keywords/postcode to find a region directly (Figure 2).

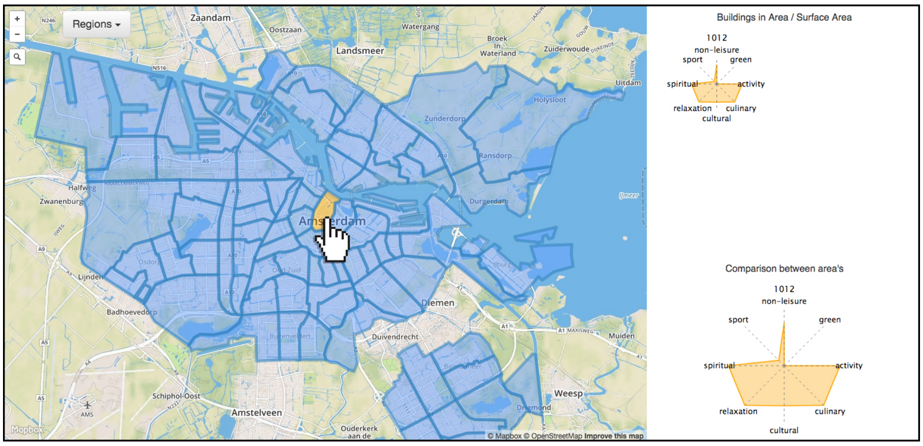

Figure 3: Choose a region

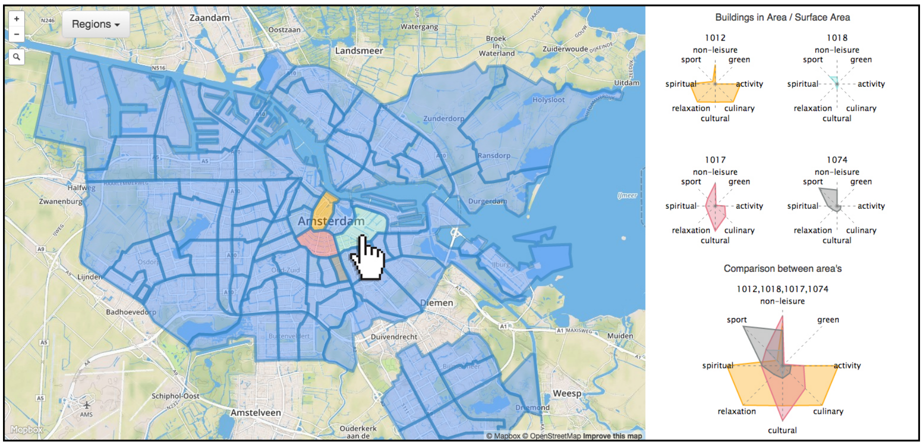

Figure 4: Choose multiple regions

Interaction with the map When clicking on a region (Figure 3), a star plots that describes the characters of the region will appear on the sidebar. It is possible to choose at most four regions the same time (Figure 4). Meanwhile the user can compare the chosen regions via the combined star plots. By clicking the chosen region once more, the user can deselect it.

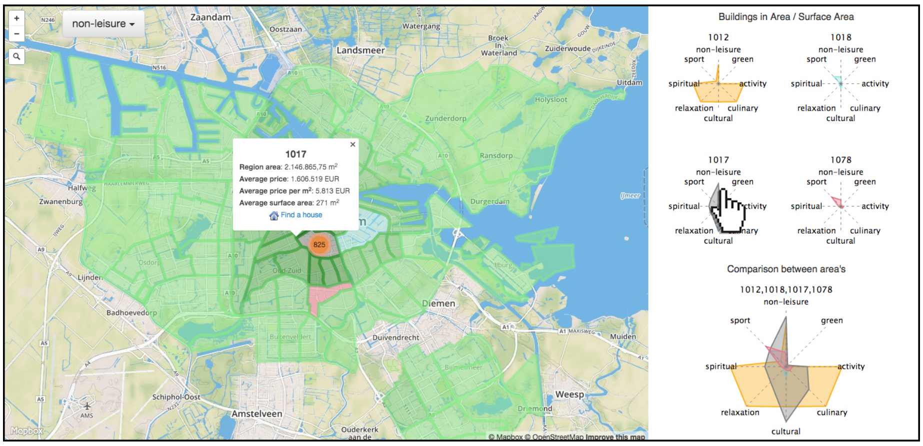

Figure 5: Click on the axes of the small star plots, checking the category "non-leisure" in one particular region

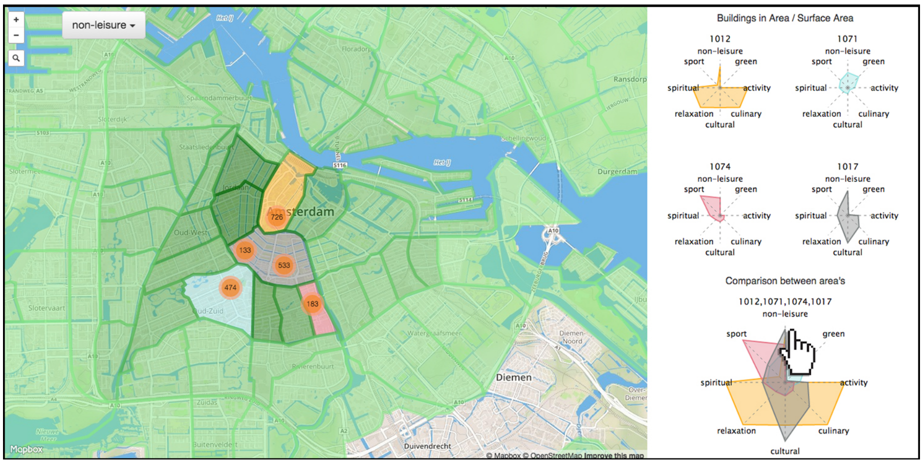

Figure 6: Click on the axes of the big star plots, checking the category "non-leisure" in one particular region

Interaction with star plots When clicking on the axes of a individual star plots, the system shows the detailed information of the chosen region. For example in Figure 5, the "non-leisure" axis of the region with the postcode 1017 is clicked, a pop-up penal opens and the number of its non-leisure places is shown near the region. Similarly, when clicking on the "non-leisure" axis of the combined star plots (Figure 6), figures of all chosen regions appear on the map.

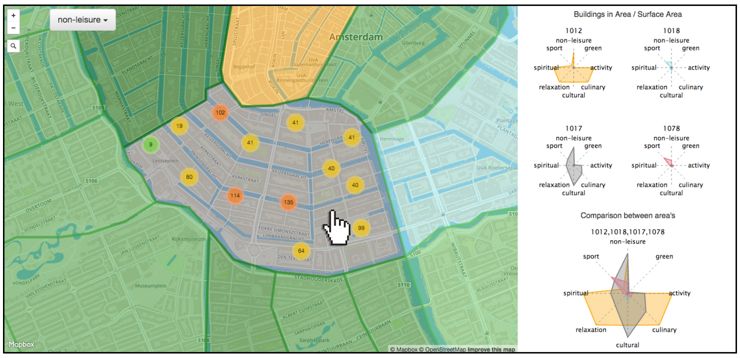

Figure 7: Zoom in/out - an in-between level

Figure 8: Zoom in/out - the most detailed level

Figure 9: Click on one of the “non-leisure places”

Scalable data representation By zooming in and out the map, users gain different levels of information representation, from an overview down to details of individual data points, with many levels in-between. The figures shown are also changed accordingly, so that the user can see how many non-leisure places near specific streets exist (Figure 7, Figure 8). In addition, the detailed information of a particular place can be found by clicking its POI icon. (Figure 9).

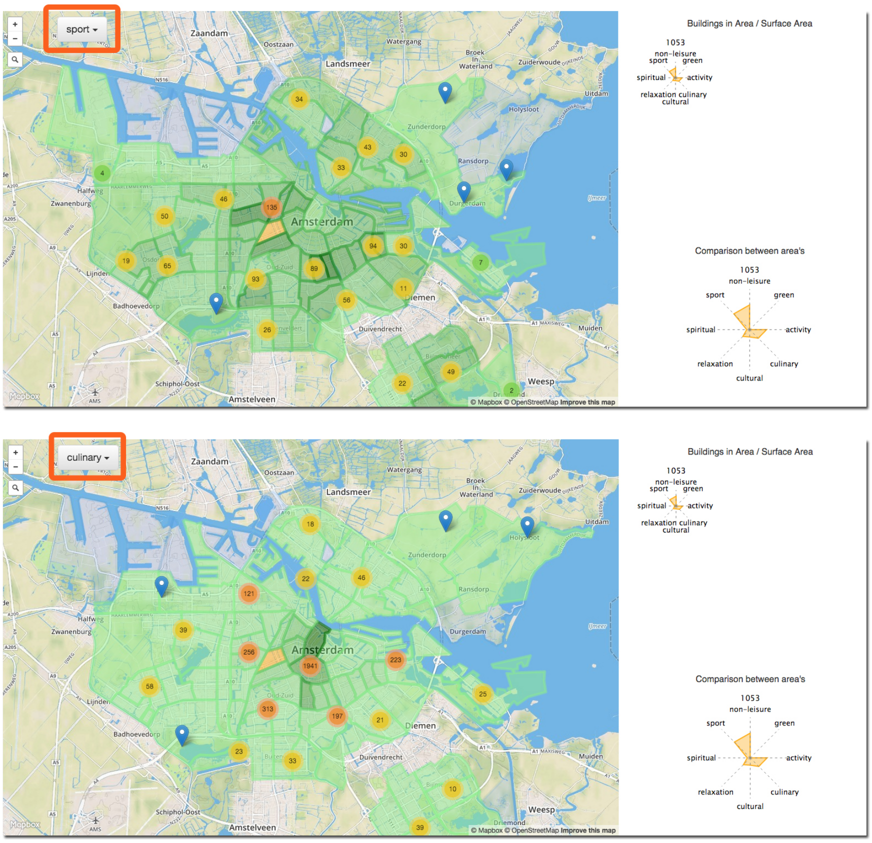

Figure 10: Explore different categories of leisure places

Look into different categories As mentioned before, users can start their exploration by choosing particular categories of leisure places. Figure 10 is an example of filtering with category. The density of buildings or places in the chosen category is shown in different colors (colors of the region shapes and the background of figures).

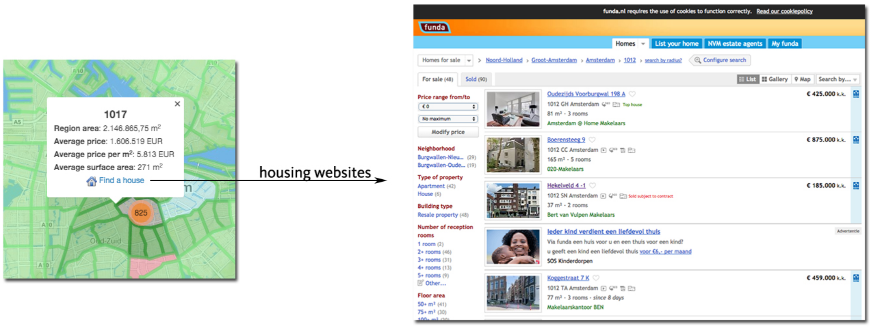

Figure 11: Redirect to housing websites

Redirect to housing websites When an user finds the region he is most interested in, she can click the "Find a house" link to find houses currently for sale in this region on funda.nl.

4 Interaction Design

Based on the notion of user intent, in this section we describe our interaction design according to the seven categories of interaction in Information Visualization [18]. For each category we provide the corresponding use case(s). A UML diagram can be found as following (12).

Select The user is provided with the ability to select regions of interest on the map. When clicking a region, its detailed information (in different categories of leisure places) can by found on the right sidebar.

Explore The system enables users to examine a different subset of data cases. When clicking on the axes of a star plot, the map will highlight the numbers of buildings that are belonging to the selected leisure category.

Reconfigure. The main visualization is based on spatial information of the different types of buildings. Besides, the comparing functionality of star plots is not influenced by the sorting of the item. Therefore the user is not allowed to rearrange the elements on the interface such as axes and the order of small star plots.

Encode For visualizing region information, two display mechanisms where developed. First, a scoring mechanism for different leisure categories is shown on the map, using variation on the depth of colors, in addition the number of buildings/places directly on the shape of the region is shown. Second, users can gain insight from star plots, which not only indicate the amount of certain leisure places in a region but also provide opportunity for users to compare several regions.

Abstract/Elaborate the user is provided with the ability to adjust the

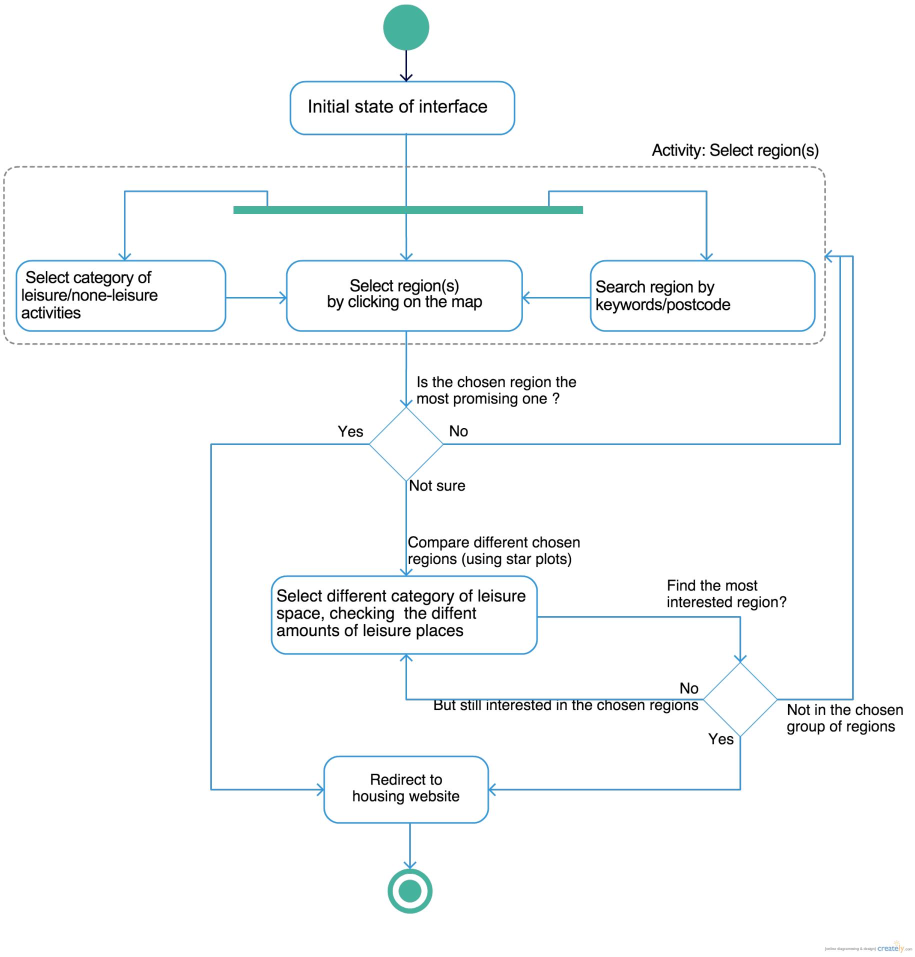

Figure 12: UML diagram

level of representation by zooming in and out, allowing her to alter the representation from an overview down to details of individual cases, and many levels in-between. On the wider view, the map shows shapes of regions and the amount of buildings in one particular category, while when zooming in, the user can find more detailed information such as how many leisure places near a street and what exactly the places are.

Filter Three basic filtering functions are provided to the user. First, the user can choose particular categories to filter the data by different types of leisure activities, which are shown in varies color on the map when selecting. Further, they can search by keywords or postcode if they already have an idea about their interesting regions. Besides, the user can click freely on the map, filtering the regions manually.

Connect When a user selects a certain leisure category, the map will then highlight all regions that have leisure places in of the chosen type.

5 Visual thinking design

5.1 Conforming to the visual thinking design criteria

Previous studies[16] have described core functional cognitive components for visual query and outlined the visual thinking process. Here we exam our design choices using the criteria by Colin Ware.

First a visual queue was used to direct the user to the point where the star plot is going to open through the use of a sliding animation when the first region is clicked. This directs the users attention to the place where something is going to happen and information will show up.

Also the amount of the POI's on the map is kept small due to the fact that if too many points of interest would be shown, the amount is combined so that there are no more than a few points to keep track of. This makes the points easily findable due to the fact that a human can only keep 3-5 items in visual memory and roughly 9 items in the egocentric map. Through the reduction of POI markers the user an easily remember the positions and amount of POI's in the different areas. As a further improvement the combined POI markers are color coded to make the amount easily comparable through the medium of color which again is a good visual query for pattern recognition, so that the amounts can easily be compared.

Through using the same color for the star plots and the regions on the map we reduce the time and cognitive load, compared to if we where using labels or other means, since this triggers a search for a pattern can be finished in one step, namely comparing color, instead of several different pattern recognition steps when showing a label of the region where one would have to compare different patterns for each number in the four-digit postcode or even go deeper into semantic coding for the recognition.

5.2 illuminating the path

In Illuminating the path [15], we learned that the following four means can enhance user recognition. Hereby we explain how we applied them in this project:

Recognition instead of recall Recognizing information generated by a visualization is easier than recalling that information by the user. Therefore in our design, users need not to remember data from previous steps. For instance, when selecting different regions and comparing their leisure places, the user can still see the location of the regions on the map while checking the figures on the right side.

Abstraction and aggregation According to [Card et al., 1991 [14]; Resnikoff, 1989 [1]], visualizations simplify and organize information, supplying higher centers with aggregated forms of information through abstraction and selective omission. In our system, we are careful about the level of representation. When zooming in and out, the user can see different information on the map, from the total number of one kind of leisure places to the name of one single park.

Visual schemata for organization Our dataset is large, therefore we need a rather clearer structure of organization. In the database we have all regions in Amsterdam, and visually every region is presented as a shape on the map according to its location information. The relation between regions is intuitive hence easy to recognize by users. The star plots on the sidebar is abstract but since their appearance are related to the clicking of the regions, user can still understand the relationship between two views and meanings of the abstract graphs.

Value, relationship, trend Visualizations can be constructed to enhance patterns at all three levels: value, relationship, and trend. In our case, firstly we use an intuitive map to show the regions, which follows the pattern for user to recognize a geographical item. Then we use different colors to show different values – the smaller the figure is, the lighter the color be.

6 Evaluation Design

For evaluating, we will conduct an informal usability test. In particular, we focus on evaluating how well the display, selection and interaction can support users in exploring information about leisure infrastructure in Amsterdam.

The research questions for this evaluation are:

- Does the system provide useful information for given tasks ?
- Does user perceive the system easy to use ?

6.1 Participants

According to studies[11, 10], five users can help us to identify most of the usability problems. Participants in this evaluation are not necessarily real users, they only need to be fairly familiar with websites browsing. Besides, the participants should not have seen our interface in advance. The user recruitment can be based on email invitations and open invitations on campus in University of Amsterdam.

6.2 Set up

The experiments should take place at the participants' home or places where they feel relax. Every participants will be asked to complete simulated tasks and following, answer a list of questions corresponding to those tasks. Such questions include their thoughts about our prototype, its usefulness and the interaction flow.

The experiment have two parts:

6.2.1 Introduction of the usability test

Participants will be asked demographic questions and for informed consent. Afterwards, they will be shown a piece of paper with a simple scenario and testing tasks on it. The following paragraph is a sample scenario:

"You are 34 years old, living in a rental house in Utrecht. You have found a new job in Amsterdam and decided to move to Amsterdam with your partner and 2 children of 5 and 8. You consider a good neighborhood for the children very important and are willing to live just outside the city center, besides you hope your children can choose from many kinds of sports. You and your partner prefer to do exercise in a park, but not gyms.

With this scenario in mind, the participants need to perform a series of tasks. Each new task will be an addition to the task before and will be in the end result in the completion of a primary goal: find a suitable area/region/community for family to live, based on the information provided on our platform.

The tasks were as follows: (1) find the most leisure area(s) in Amsterdam; (2) considering your situation, find possible regions for your family; (3) compare the pros and cons of several regions and (4) check different sport facilities surrounding your chosen area.

6.2.2 Experiment Session

While performing the task, the participants should be strongly encouraged to think-aloud[17] in order to gather as much qualitative data as possible. This think-aloud method as an analysis protocol has been used in many usability studies and is a popular qualitative and informal evaluation method, especially for exploratory studies such as the one presented in this paper. In addition to the standard think-aloud method, we also will carry out a short survey after each performed task to instigate a dialog with the participant about the performed task. The question of the survey includes how they perceive the ease of use of the interface and general impressions about our design.

7 Scientific Excellence

This project uses several design patterns from the field of information visualization to help people find a new place to live. To find interesting areas for the people the criterion of leisure has been chosen as a basis, since leisure is an important factor in the life of humans as it is correlated to fun and relaxation [6]. To achieve the ordering of leisure places this project was based on the idea of Yuan, Zheng and Xie [19] and used POI's which are related to different types of leisure to grade regions within Amsterdam on these different types. The idea of the POI grading was combined with a POI representation which was enhanced through the use of Colin Ware's visual queue criteria's. [16]

8 Reflection

8.1 Pros and cons of the project

Initially our project has been faced with a problem of redundancy of man powers. There are 6 person in the team but this resource has not been used effectively during the fact that most of time was spent for transferring knowledge from one person to another.

We have confirmed in practice the assumption introduced by Frederick Brooks [4] about the efficiency of team working in small group of people by finishing the project in time with decreasing amount of team members.

Another good thing that we have realized that our project has solved specific problems that can be used in practice and it's focused on concrete type of customers that are going to seek a house.

The bad thing is that provided time range was not enough to finish all minor features. For example, with a large amount of map data our system still needs some big delay to load and cluster them (this problem has been solved now by some visualization "cheat" to visualize data gradually). Another example of existent problems is usability of analyzing tools (like our star plots) that can still be improved in the future.

8.2 Team functioning

8.2.1 Background of team members

A major part of the team has a Computer Science, Artificial Intelligence background, a minor (but not less important) - Information Science. The good thing is that every team member has already some pre-knowledge and understanding in technologies and concepts that we have used in our project that leaded to constructive and operative discussion and effective problem solving in prospect.

In addition, a major part of the team are practicing programmers that are familiar with concepts of product management and development.

8.2.2 Responsibility of team members

During the fact that our project consists of two major parts (frontend and backend) it was decided to have one person fully working on the backend part (Sebastian Dröppelmann), and one person fully responsible on the frontend part (Floris de Bruin). Xiaoli Li, as a person that is most familiar with usability and design, was responsible for this thing in the project. Minh Ngo was working as a "free artist" and was not responsible for anything, but mostly his contribution includes the project architecture, REST API for data requesting and map visualization.

Despite of the fact described above, there aren't any restrictions about possible improvements into the project by contributors. Every implemented feature has been discussed before being implemented that means that the project involves efforts of each team member equally at each stage.

8.3 Team members self-evaluation

We are the team and should be evaluated as a team (One for all, all for one!). In spite of this fact, we would like to say some kind of regret because of the loss of two our initial team members during our short-term battle against vicious enemies.

A Appendix I

The mapping of the category labels to the chosen leisure categories.:

D1.01, activity
D1.02, activity
D1.03, activity
D1.04, activity
D1.05, activity
D1.06, activity
D1.07, activity
D1.08, activity
D1.09, activity
D1.10, activity
D1.14, activity
D1.15, activity
D1.16, activity
H2.06, activity
H2.07, activity
H2.08, activity
H2.09, activity
H2.10, activity
H2.11, activity
V3.01, activity
V3.02, activity
V3.03, activity
V3.04, activity
V3.05, activity
V3.07, activity
V3.08, activity
V3.09, activity
V4.01, activity
V4.02, activity
V4.03, activity
V4.04, activity
V4.05, activity
V4.06, activity
V4.07, activity
V5.01, activity
V5.02, activity
V5.03, activity
H2.01, culinary
H2.02, culinary
H2.03, culinary
H2.04, culinary
H2.05, culinary
V2.01, cultural
V2.02, cultural
V2.03, cultural

V2.04, cultural
V2.05, cultural
V2.06, cultural
V2.07, cultural
V2.08, cultural
V2.09, cultural
D2.05, relaxation
D2.06, relaxation
D2.08, relaxation
D2.09, relaxation
H1.01, relaxation
H1.02, relaxation
M3.01, relaxation
M3.02, relaxation
M3.03, relaxation
M3.04, relaxation
V3.06, relaxation
V3.10, relaxation
M5.01, spiritual
M5.02, spiritual
M5.03, spiritual
M5.04, spiritual
V1.01, sport
V1.02, sport
V1.03, sport
V1.04, sport
V1.05, sport
V1.06, sport
V1.07, sport
V1.08, sport
V1.09, sport
V1.10, sport
V1.11, sport
V1.12, sport
V1.13, sport
V1.14, sport

References

[1]

[2] Team 10. Leisure map of amsterdam. `root.org.ua/citymap`, 2015. [Online; accessed 30-March-2015].

[3] Instant Amsterdam. Postal codes in the netherlands. `http://instantamsterdam.com/postcode.htm`, 2015. [Online; accessed 25-February-2015].

[4] Frederick Brooks. *The Mythical Man-Month.* Addison-Wesley, 1975.

[5] CBS. Allochtonen in amsterdam. `https://www.google.com/fusiontables/DataSource?docid=1mjJ3ixFpeQKQ6vBybfKXTqQlzyV67Ycj3ckpN6QA#rows:id=1`, 2015. [Online; accessed 25-February-2015].

[6] J Dattilo and Dennis Howard. The complex and dynamic nature of leisure experience. *Journal of Leisure research*, 26(3):195–211, 1994.

[7] Stephen Few. *Now you see it: simple visualization techniques for quantitative analysis.* Analytics Press, 2009.

[8] Vienna Government. Citymap. `http://www.wien.gv.at/stadtplan/en/#t`, 2015. [Online; accessed 30-March-2015].

[9] Gemiddelte huizenprijs. Gemiddelte huizenprijs. `http://www.gemiddeldehuizenprijs.nl/gemiddeldehuizenprijs/index.php`, 2015. [Online; accessed 30-March-2015].

[10] Jakob Nielsen. Why you only need to test with 5 users, 2000.

[11] Jakob Nielsen and Thomas K Landauer. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 206–213. ACM, 1993.

[12] City of Amsterdam. Non-residential functions. `http://maps.amsterdam.nl/functiekaart/`, 2015. [Online; accessed 25-February-2015].

[13] City of Amsterdam. Open geo data. `http://maps.amsterdam.nl/open\_geodata/`, 2015. [Online; accessed 25-February-2015].

[14] G.G. Robertson S.K. Card and J.D. Mackinlay. The information visualizer, an information workspace. in proceedings of the sigchi conference on human factors in computing systems. *CHI'91*, pages 181–188, 1991.

[15] James J Thomas. *Illuminating the path:[the research and development agenda for visual analytics].* IEEE Computer Society, 2005.

[16] Colin Ware. Visual queries: The foundation of visual thinking. In *Knowledge and information visualization*, pages 27–35. Springer, 2005.

[17] Peter C Wright and Andrew F Monk. The use of think-aloud evaluation methods in design. *ACM SIGCHI Bulletin*, 23(1):55–57, 1991.

[18] Ji Soo Yi, Youn ah Kang, John T Stasko, and Julie A Jacko. Toward a deeper understanding of the role of interaction in information visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1224–1231, 2007.

[19] Jing Yuan, Yu Zheng, and Xing Xie. Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 186–194. ACM, 2012.