Density heat map visualization of spatially and temporally localized events

Pierre Couy Hugo Fauvet Louis Manhès

1 Introduction

As population density in cities rises, it is becoming increasingly difficult for people to move from one point in the city to another. We propose a method for visualizing and analyzing punctual events localized in time and space, and use it with Uber pickups in New-York City from April to September 2014. This data was released following a Freedom of Information Law request submitted by the NYC Taxi & Limousine Commission in July 2015.

2 Related work

2.1 About the choice of a color scheme

A wide variety of heat maps are used to display all kinds of data. These heat maps often use a color map ranging from blue or green for lower values to red for higher values. However, this is not always the best solution : depending on the type of data displayed, other color schemes may provide a more intuitive visualization.

In [5], Kenneth Moreland deals with the issue of rainbow color maps. According to him, and numerous other researchers in the field, this kind of color map is inefficient at meaningfully displaying scalar fields, while for instance giving the illusion of a gradient where there is none, or highlighting low interest regions of the visualization. To address this problem, the author discusses the use of diverging color maps.

Another point of view is to avoid as much as possible the use of colors in heat maps. In [1], the authors advise the use of gray-scale for heat maps for several reasons :

- Color maps lack the natural and intuitive ordering of gray scale
- Uncontrolled changes in luminance cause a loss of information when converting to gray-scale or printing
- Color maps can lead to the perception of a non-existent gradient

2.2 Grouping individual events depending on zoom level

Most of the available tools for visualizing geographical data on a map allow the user to zoom in and out. When visualization includes markers, showing them as individual markers can lead to a map that is difficult to interpret. In order to avoid this phenomenon, it is possible to make clusters of markers, and display each cluster as its own marker and eventually adapting their sizes to the number of elements in the cluster. This feature is for instance used in *Google Maps* and is described in [4].

3 Technical Design

3.1 Visualization

The design of the project is a one-page web application displaying in the main area a density heat map [3] overlaid on a geographical map of the desired location and on the right side a control panel that allows the users to tune their search parameters. The user can choose a predefined dataset or load his own in order to create the heat map.

In the control panel, the user is able to precisely filter the events he wants to display on the heat map. In order to tune parameters, several type of fields will be offered depending on the parameter :

Checkbox lists : A list of checkboxes, one for each possible value, is displayed. All of them are checked by default. Data matching the checked checkboxes will be included in the heat map.

Double sliders : A time line is displayed, with two sliders on it. Data matching the time window between the two sliders will be included in the heat map

The user will be able to tune the following parameters :

Year(s) : Checkbox list

Month(s) : Checkbox list or double slider ranging across a year

Day(s) of week : Checkbox list

Time of day : Double slider ranging across a day

The goal of this fine-tuning is to give the user the ability to visualize meaningful data. For instance, the user may want to separate data for work days and weekends, but still be able to filter by a specific time of the day, such as early in the morning. This way, it is possible to display data about the usage of a transportation service during morning commutes relative to its average usage. Another use case would be to compare data from winter months with data from summer months.

We made a first mock-up sketch which is displayed in figure 1. The user is asked to setup a few parameters relatives to his own preferences. For instance, the user can pick a date and time in the future (e.g his departure for office) and visualize the affluence of one kind of transportation for similar times in the past (same week day, same hour, same period of year) through the use of the density heat map.

Moreover it is also possible to combine multiple datasets in order to have a quick comparison between different types of transport, or between different time filters. The comparison is made through the use of one dataset as a reference for the other. The resulting heat map is then displayed using a diverging color map (cf section 3.2) and gives an overview of the density in one dataset relative to the other. We are exploring the possibility to draw the correlation between two datasets, even though we are not fully decided on how this feature should work.

Finally another specific feature will be to give the user the ability to set a duration, then make a time window of this duration slide across the entire temporal frame of the data. As the window slides, the map is updated live to display an animated heat map.

3.2 Color Scheme

The choice of a specific color map should be taken seriously into consideration and it should not be based on one's preferences. Indeed, a color map carries informations that could lead to some misinterpretations if not used correctly.

Color maps may be split in three categories :

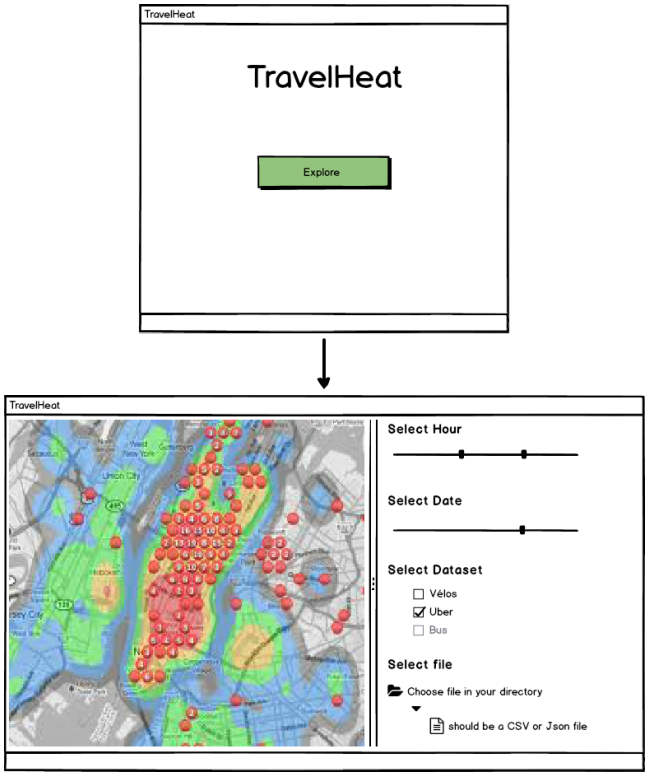

Figure 1: Mockup of the graphical UI

Diverging color maps, for which colors vary in two opposite directions from a central color

Sequential color maps which may consist of variations in luminance for a single hue, but not necessarily

Qualitative color maps are made of uncorrelated colors which are not meant to be ordered

Examples for these categories are displayed in figure 2. Choosing a diverging map for plotting density related data can lead to some confusions because they are for data with extremes at both sides of a neutral value, which is not the case for a density of transport usage. Therefore we will use the sequential color map for this task and we shall privilege the use of the diverging color map for the heat map generated by the difference between two kinds of datasets.

3.3 Dataset

In order to have some predefined transportation data to make our heat map visualization, we downloaded the Uber pickups in New York City Dataset from the web plateform Kaggle. This directory contains data from over 4.5 million Uber pickups in New York City from April to September 2014, and 14.3 million more Uber pickups from January to June 2015. The files are separated by month and each has the following columns :

- `Date/Time` : The date and time of the Uber pickup
- `Lat` : The latitude of the Uber pickup
- `Lon` : The longitude of the Uber pickup
- `Base` : The TLC base company code affiliated with the Uber pickup

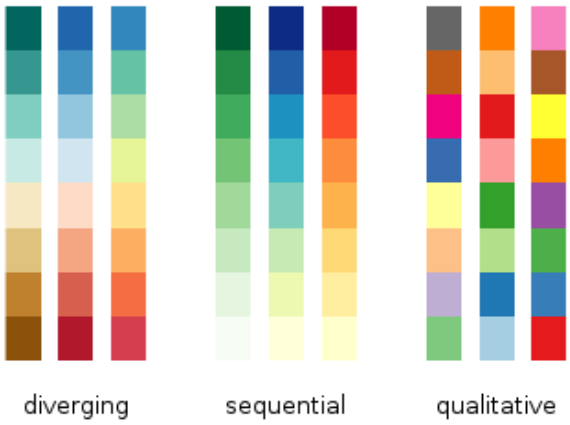

Figure 2: Examples of color maps

We planned to dismiss the Base attribute because such an information is not useful for the purpose of our work.

We discuss here the *Uber pickups* dataset, but a core aspect of our project is its ability to display any data provided in the CSV format with at least 3 columns : one for date and time, and two for latitude and longitude. However, even temporally and spatially localized data can be unfit for this visualization. Here are some other suggestions of datasets that may produce meaningful visualizations with this project :

- Stamping of transport tickets
- Bike share withdrawals/deposits
- Car parkings entries/exits

Other datasets not related to transportations may be meaningfully displayed by our project, but this topic is not discussed here.

3.4 Tools

Among the available technologies, we found that a combination of Google Maps, Leaflet and the framework D3.js would match perfectly for a web-based application.

Indeed, we found a well described visualization on a french website [2] with the use of advanced mapping techniques.

4 Conclusion

In this work, we propose the first steps for a web application that generate some heatmap based visualization in order to give the user a comprehensive exploration of the transportation in his current location.

We demonstrate through the use of a mockup design the possibility of integrating such visualizations in a web-based application.

References

[1] D. Borland and R. Taylor. Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications*, 2007. doi: 10.1109/MCG.2007.323435

[2] E. Frigot. DATAVIS tutoriaux javascript pour la visualisation de donnes, 2016.

[3] N. Gehlenborg and B. Wong. Points of view: Heat maps. *Nature Methods 9, 213*, 2012. doi: nmeth.1902

[4] I. Google. Utilitaire de regroupement de marqueurs google maps android.

[5] K. Moreland. Diverging Color Maps for Scientific Visualization. *Proceedings of the 5th International Symposium on Visual Computing*, 2009. doi: 10.1007/978-3-642-10520-3_9