Visualization of Accidentology in Paris, risk factors and evolution of the traffic accident

Bourhis Kilian p1609961 | Lechat Nicolas p1205627 | Nguyen Ngoc-Bien p1715271

Abstract

This paper present the project we work on during the course of Data Visualization from November 2017 to January 2018. We choose to work on the topic of the accidentology as it is a nowadays problem especially in the biggest cities. Hence we recover the accidentology data about the city of Paris for the period of 2012-2013, the most recent and public data we found.

Our objective was to create a visualization simple to understand and enough intuitive for the user with the aim to offer him the tools to understand the traffic risks of the city and found some interesting correlations. To manage this, we represent the different points of interest on a map of Paris on which the user is able to interact directly by zooming or hovering an area to discover its informations. We also provide multiple filtering options to help the user to focus on what he is interesting or proceed to comparisons.

The result of this project is a demo available at this address `https://garet005.github.io/m2-2018-DataViz-KSN/`[8] which already offer the possibility of detect some of the simplest risk factors like the most risked boroughs or periods. A base for future complex statistics comparisons is also present when the user seek for informations about an area.

1 Introduction

Nowadays, traffic security is one of the most important issue in our cities. A lot of data are collected every day in order to identify the most risked sections, improve the road network or prevent future accidents. The goal of our study is to offer a simple visualization tool to the citizen of Paris who may want to know which route is more risky or have an idea of the number of accident in its agglomeration.

Mapping for big cities has been very well done in the past. These maps contributed to help people gathering a lot of informations, making it easier to find their way and facilitate their traveling more than ever. Never before the mapping tools have been as important and effective as it is today, and it will increasingly play a more important role in the lives of individuals in modern societies. The idea of building a map of traffic accidents is an important part of our work, as we want to help people to understand the risks of their city traffic as well as discover the main factors that are correlated with these risks.

As the visibility and the simplicity of our visualization is important to us we decided to work with informations dynamically generated by the user. Hence we have started to work on a filtering tool which is not mandatory to discover the data but can greatly help the user restricting the area of it research by displaying only the informations, areas he wants. Nevertheless we also wanted to keep the capacity for someone which may not know what he is looking for, to start from a general point of view and then study more in depth the traffic risks thanks to another view by which the user could found what can be interesting for him. To do so, we virtually split our map in two levels of visualization on which you can swap by a simple zoom on a targeted area on the map.

Another effort we make was to attract people with a visualization on which they can recognize a street or a district in which they would know some traffic informations. A main advantage of our topic is the abundance of relevant documents. In fact, there are a lot of maps of different cities that have been built, which have given us different ideas to build our own maps. In doing so we hope users will be concern by our visualization as we tried to offer a visualization on which our informations are inlaid.

In order to explain our visualization, where we take our ideas and how we build it we will start by a quick state of the art about traffic and city visualization. Then we will explain how we have done this project, our choices of design and the ideas we did not manage to implement during the project. Eventually we will discuss about the limits of our visualization and how we could improve it.

2 Related Work

We had a lot of examples of different maps with different types of areas and different cities for references [7]. We found many different sources like on `http://blockbuilder.org`.

As example for our subject, we used the website `http://itisaasta.com/nycs/nycs.html`. As we can see on this website, there is a lot of possibilities and informations. This website is also simple and expressive. On a the metric's criteria "School", each school is shown by a blue dot, and this point was very interesting for our different ideas for some reasons. When we hover a point, we have informations about the school, a graph showing the statistics related to the quality of the school and the relative comparison between that school and other schools. There are also many

icons available to choose from, such as geographic area selection, school type selection, program selection, level selection, and more. This is a well-designed, well-elaborated map that makes it easy to find, search and access information.

Figure 1: Visualization of Schools in New-York at http://itisaasta.com/nycs/ .

Another map which helped us a lot is the traffic map of London, at `http://vis.oobrien.com/trafficcounts/`.

This map is used to describe informations about vehicles participating in traffic over different time periods. The author uses concentric circles with different sizes and colors to convey different details. There are also a lot of icons that allow users to display various types of vehicles, in different time periods.

When the user hover over a circle, a detailed info sheet will be displayed to help the user get the most detailed information. This is also a well-elaborated and comprehensive map that gives users an insight into the London traffic situation for many years.

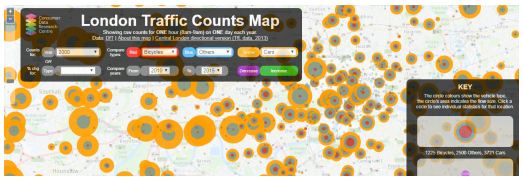

Figure 2: Visualization of traffics in London at http://vis.oobrien.com/trafficcounts/ .

We also refer to the Paris's map created by Vincent Garreau at `http://vincentgarreau.com/dataparis/`. There is many areas related to the lives of people (left of the website) and the areas of different districts of the city, which depicts various statistics for metro stop areas, residential or commercial centers, schools, hospitals, cultural centers.

Here, the author gives us a variety of approaches to information such as average income, annual taxes, unemployment rates, political characteristics, student ratios, high education ratios, marital status ratios, retirement age rations, average age, and so on. Also, users can choose to display their own metro stops, to zoom in or zoom out the map. Each metro stop will be represented by a circle with different colors representing different statistical results. When the user hovers on a circle, a small table will display the detail statistics about the field that the user has selected. This is an elaborated map which provides a lot of useful indications. However the zoom is very limited in this visualization, making it very general.

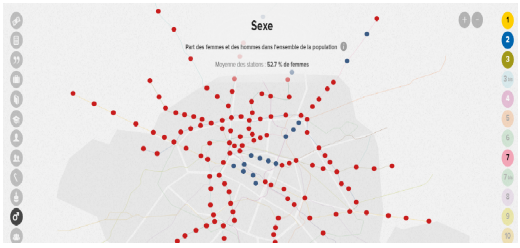

Figure 3: Visualization of many statistics in Paris at http://vincentgarreau.com/dataparis/ .

3 Techniques and Choice of Design

As our main target are the regular citizen of Paris, our interface will be as simple as possible to help them highlight their main interests and inform them about the traffic security in their city.

To manage our visualization we use a form for the choice of filtering. The visualization itself is built on a single view on which we distinct two sub-views which will determine the level of accuracy available. The main view whose goal will be to give the general situation of Paris with its major locations of interest. The second view will be more interactive with the user if he want to get more details on one of the point of interest. In order to display the informations the user is looking for we use a window on the left of the main view.

3.1 Dataset

We choose to use a dataset produce by the city of Paris that we found on the open data of Paris website. The data are available in different kind of format but we choose to use the GeoJSON file which will be useful to display data according to their coordinate on the Paris map. This sample of data offer us a large amount of details about the condition of the accident such as the type of vehicles implies, the role of the different person, a precise localization (latitude and longitude), date and time.[6]

3.2 Form and Filtering Algorithm

The way we proceed to manage the filtering options is not the best but according to the difficulties we encounter on the visualization itself we decided to stay with this simple but effective form. If we had more time, our plan was to move the form on the right of the visualization in order to see the whole page without need to scroll down to the page. We would also use slider for time and date intervals as well as a calendar to select the days, months and/or years with a multiple date picker system. We think that these ideas would

have bring a more compact and modern form.

Our form offer the possibility of selecting the borough, an interval of time and the choice of dates. For the dates we offer two possibilities for the user : use a simple interval or pick the days, months or years in which he is interested. This second possibility make possible for instance to study the number of accident during holidays seasons, week end etc. Of course none of these options are required in order to obtain a visualization. By default, if the user do not choose any filtering options all the traffic accident will be display.

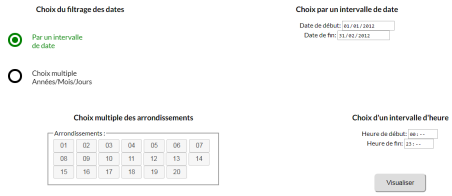

Figure 4: Table to choose different parameter to display.

The form stay accessible all the time and can use in any level of zoom to add or remove options of filtering.

Once the user updated the form and click to see the result we call our filtering algorithm to create or update two objects we will use for the visualization. The first one is the list of the filtered accident, each accident is composed by its location (latitude, longitude), the id of the point of interest in which it is affected and all the details of the accident itself. The second object represent all the the points of interest with its center (latitude, longitude) which is in fact the position of one of the accident inside the area. It also has an id, a marker layer and the statistics of all of its accidents. Thereby, we just have to go through the list of points of interest to draw our circles and the list of accidents for the markers.

3.3 Main View

This will be the home visualization, representing the map of Paris with what we call the points of interest. These points represent regions where accident occurred during a selected period of time. As we explain it before we decided to represent a point of interest by a circle. The idea was to have the points of interest displayed on the main view and then the precise area draw dynamically when zooming for the focused view. Nevertheless, generating the different points of interest according to the filtering of the user bring the problem of affect the accidents to their closest area with the difficulty that this area may not be created yet. We manage it by defining a radius for the perimeter of all our area in which an accident must be affect in. So for each accident, in our filtered list, we look for the closest area in which it can be affect (comparison between the distance to the center of these areas). If our accident can't be affect to an area (to far or just no area created yet) we create a new area whose center are the coordinate of the accident. Despite this method, problems will appear when we will look at the focused view and will be discuss at this moment.

To display the point of interest on the map we use the d3 svg which allow us to easily draw our circles [5] and a mercator projection [3] to transform our latitude longitude into X,Y pixels coordinate relative to our svg div. The circle are updated each time the user moves or zoom on the map to maintain a correct position. We also add a progressive increase to the size of the areas to follow the zoom of the user. To overlaid our points of interest in our map we used the leaflet.js [1] [2] and d3.js libraries.

A gradient color (7 colors) on the points of interest and a legend is also provide to help the user to spot the areas where the number of accident is very important. Unfortunately, our legend is not dynamic according to the filtering because of a lack of time. Our legend is more suited when their are no filtering options of time or date (because the number of accident will strongly decrease) as our legend is scale when all the accidents are display in an area. We also lack of time to add an other visual aid which would consist of highlighting the areas where the gravity of the injuries are the most important.

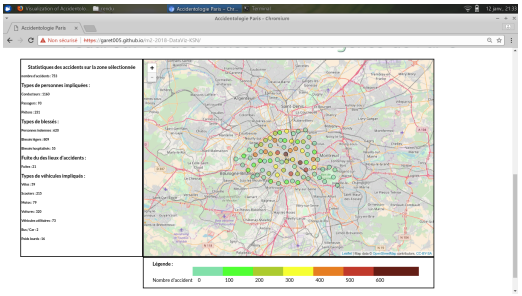

Figure 5: Visualization of Accidentology in Paris by main view.

Moreover, the user has the possibility of display statistics about a point of interest by hovering the area. These statistics are displayed in the left window and categorized in 5 main fields : the number of accident, the type of person and vehicle involved and the level of the injuries. An interesting functionality would be to allow the user to select two areas or boroughs and generate a comparative graphic on the different field.

3.4 Focused View

The focused view appear at a certain level of zoom on the map, it is marked by a lower opacity of the areas. In this "view" the user is now able to display all the accident in an area by clicking on it (the border of the area become blue). Each accident will be display by a marker and will remain until the user re-click on the area to hide the accidents. It seems that when multiple accidents are very close the markers are just superposed (this is why some area have less markers than indicated). We encounter an other problem when we tried to display pop-ups by clicking on a marker. Indeed, each area has it own layer where the markers are stored, but these layers seems to be behind the leaflet map, hence they are not selectable and we did not manage to

find a solution.

To add our markers we added a marker layer to each of our area [4]. We then go through our list of filtered accident to add them to the corresponding layer depending their area id.

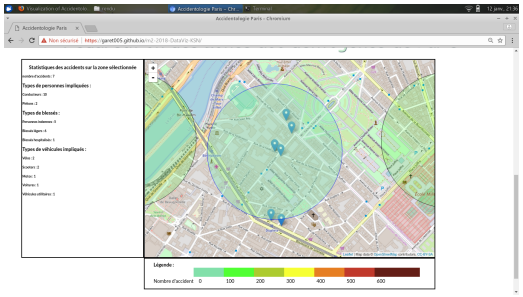

Figure 6: Visualization of accidentology in Paris by focused view.

As it can be seen in this view, the form of our area are not representative of the geographic position of their accidents. To solved this problem, discovered during the development of the project, we had the idea to change the form of our areas in the focused view by replacing the circles by polygons. This idea has been tried but it was technically very difficult to create these polygons. These polygons would have solved the correctness of the geographic representation of our area and by the way bring a new information to the user by indicated the area where there is no accident (all the spaces outside our colored areas).

4 Discussion

In the end of our project, we noticed some troubles, some limits and some ways to improve our demo.

4.1 Problems Encounter

Because of lack of time, some bugs remains.

In first, circles don't encompass all accidents linked to it ; in the beginning of the project, our idea was to have two separated javascript files, one for the global view, and one for the focused view. On the global view, we wanted to represent with this circles the different points of interest - areas with accidents - but during the project, we change our minds with the idea to use only one javascript file, so gather the global view and the focused view. We wanted to change the visualization with the zoom. Presently we have, with the zoom, the changing from the global view to the focused view, but we failed to change circles with other polygonal shape to cover every points linked to the focused area.

In a second point, on the focused view, when we click on a circle to reveal coordinates of each accident then we decide to change with using some new parameter without hiding coordinate points, this ones remain on the map because we don't clear all of the different markers layers.

We have a problem with the legend that scale on all datas and not on selected datas.

In a last point, we have the mixed-content, some external script that we are using are not in https.

4.2 Limits

One of the first limit we have is that we only have one dataset of the traffic of Paris. Our visualization would greatly benefits to have multiple dataset which could be link to the road accident in order to find some advance correlations.

Another import limit is that we do not offer the choice of the type of data which can be visualized on the map. Indeed, we only display the area where accident happened and the location of these accident. It would have been more interesting to add other type of informations to display by using an other filter. For instance we could display only the gravity of the accident by coloring the map with a color gradient, the accident where a vehicule involved runaway with a specific icon or only the accident where a pedestrian was involved.

4.3 Improvement ways

To improve our project, we have many ideas which have not been implemented. Some of them have been mentioned during the description of the different part of our project and we will expose a few more.

Despite the few bugs we would have to correct, a main improvement would be to replace the left window multiple graphics below the map to display by a more visual way. In case of multiple areas selected the graphics could do a comparison (gravity of the injuries, time of the accident etc). We could also add below the map other graphics but this time on the overall data. This could highlight which type of vehicle or borough is the more risky, a calendar view with the number of accident to see immediately which period are the most dangerous etc.

We also have the idea to add some other csv or GeoJSON files and cross the results, like the weather during the same period to verify if accidents append during some specific climates, or like alcohol or drugs consumption to verify if there is a possibility that some substances can be responsible of some accidents.

5 Conclusion

To conclude, we have a simple and intuitive visualization that represents Paris and its accidentology which covers the period 2012/2013 and every accident referenced over this duration.

User can select some specific dates or periods and choose precisely one or more districts. They also can select multiple time slot to be more specific on their research.

Thanks to this, each user who will use our website will realize that accidents and road troubles are not harmless and that can happen to everyone, a kind of road prevention. Users also can consider areas where accidents are frequents and change his path or be more attentive whether on foot, by bike or by car.

Some programmers can use our work which is commented and change the map and the dataset, or improve this one to enhance the road prevention, the understood of the traffic risks and its factors.

References

[1] V. Agafonkin. Leaflet, an open-source javascript library for mobile-friendly interactive maps.

[2] M. Bostock. D3 + leaflet.

[3] M. Bostock. Geographies (d3-geo).
[4] d3noob. Adding a marker to a leaflet.js map.
[5] d3noob. Map using leaflet.js and d3,js overlaid.
[6] M. de Paris/Direction de la Voirie et des Déplacements. Open data paris - accidentologie.
[7] M. Dörk and S. Meier. Visualizing cities.
[8] B. Kilian, L. Nicolas, and N. Ngoc-Bien. Accidentologie paris.