# Comparative visualization of different factors involved in road transport accidents in France

Elisa Denier* M2 Bioinformatique
Lucía Castro García† M2 Bioinformatique
Lorraine Soudade‡ M2 Bioinformatique

**ABSTRACT**

Road transport accidents, its causes and consequences are widely recongnized as relevant, and several sensibilisation campaigns are launched all over the world for this matter. Nevertheless, a lot of them consist only on plain text or unsightly graphics of difficult comprehension, which fail in giving accessible information for readers. In this article we present a very complete comparative visualization of different factors involved in road transport accidents in France in 2016.
These visualization is made using the D3 javascript library, and it contains multiple types of charts chosen carefully to best explain each factor, such as a calendar chart, an animated donut chart, animated line and pie charts, an innovative impact chart, a stacked chart and a French map. We have achieved to represent interesting factors such as the days and hours where most accidents take place, the vehicle's parts which are the most susceptible to collide, the gravity of accidents, the sex ratio and predominant age groupe of actors involved, and finally the geographic localisation of each accident. All in all, we have achieved a very complete visualisation designed to be understood by all types of public and allows readers to easily draw conclusions by themselves. For instance, we concluded that in 2016 men where involved in more accidents than women, that the age range of the majority of people involved was between 25 and 45 years old, that in France most accidents occured at 8am and at 6pm, and the list goes on. You can visit our project github page here, containing our visualization.

**Index Terms:** Visualization—javascript D3 library—Road transport accidents—2016; sensibilisation—preventive information—dynamic charts

## 1 INTRODUCTION

Road transport is the favorite and most frequent way for people all over the world to travel. There are more than one billion vehicles in the whole planet, and this number never stops to rise even though transport vehicles' counts have dramatically drop since the Transport Revolution in the 20th century [4]. In France alone, until January the 1st 2015, more than 38 million of vehicles travel the roads of this country, and studies have proved that 83% of the total passenger traffic count is ensured by road transport, as well as 80% of freight transport (INSEE, 08/2007).

There are obviously many positive aspects in the fact that road transport is available for everyone. But there is also a darker side, which is the increase of road accidents. In 2016, there have been around 3500 people killed on the road, and mortality statistics don't seem to go down fast enough.
According to the French road transport accidents assessment for 2016 [8], the number of people killed in a road transport accident has stabilized after two years of increasement. Nevertheless, other accident indicators have raised: the number of injured people (+2,6%) and the number of injured people being hospitalized (+1,6%). What's more, this assessment informs us about the age group of people involved in the majority of accidents. The most vulnerable age group are the young people from 18 to 24 years old, with 597 of them killed in 2016. We find also information about the sex ratio of people involved in accidents, being men the most vulnerable group as for accident victims. We can also find information related to the type of vehicle involved, different accident factors such as alcohol or drugs, and other interesting statistics. The big problem with this assessment is obviously not the truthfulness of the information provided or its pertinence, which are redoubtable, but its format. It's 100% plain text.

Actually, most people are visual. We are interested in colors, forms, big fonts and images, whether we talk about a child or and adult. Thus, the only way to get people interested in a delicate and complicated matter such as road transport prevention measures, it's through well designed informative visualizations which speak for themselves.

Therefore, we have decided to focus our project on this subject. We consider that road safety is an essential element in our society, but unfortunately it doesn't seem that people are aware enough on this matter. In order to decrease the number of accidents, people need better and more visual ways to realize the dangers. Hence, our goal is to create interactive and interesting visualizations which allow people to easily come across information about accident-related factors. For example, in which French departments accident counts are higher, what is the most dangerous type of vehicle, what situations are to avoid before driving, and so on.
Apart from preventing purposes, this project has two extra motives: to complete the work that already exists on the internet but that we consider is unfinished, and to put to the test some society stereotypes, for instance the fact that women are involved in more road accidents than men.

## 2 RELATED WORK

Before embarking on the data research for our project and its actual development, we searched for existing visualizations based on our same subject in order to get a global idea about what has already been done and the different approches developers have taken regarding the road accidents problem. This step was important because we had to take into consideration that we wanted to do something useful and innovative and mostly, we wanted to do something that was lacking in the already existing resources. We decided to search for and to inspire ourselves from infography works, as we consider this type of data visualization as very easily understandable and accessible for everyone.

We choose to present the two most interesting visualizations we found based on accident rate. These cover an specific matter about our subject, specially the men/women ratio for this first

*github: ElisaDnr
†github: lcastrogarcia
‡github: lsoudade

visualization [6] that we explain in detail in the first bullet point section, and the motorcycle accident rate specifically for the second visualization [2], that we address in the second bullet point section.

- **Men/women ratio in accident rate approach**
  This visualization particularly covers the fact that most fatal accidents involve men rather than women, whether it is victims being mostly men or fatality actors being mostly men. The men/women rate graphic is very clear and precise, and it convinced us of the fact that the best way to represent basic data is with a simple and concise graphic, much more efficient than a complicated one overloaded with information. Hence, we decided to inspire our work on this visualization to represent similar data, since it is crucial that this kind of graphic catches the attention of the reader from the very first glimpse to make sure that the information is transmitted quickly and efficiently.
  This infography covers also some accident factors, for instance alcohol, drugs and not having a driving licence. These factors where a cause in 82% of accidents, which is a very useful information when trying to prevent people and to persuade them to think twice before driving.

- **How to avoid a road accident in a motorcycle**
  This visualization covers exclusively accidents involving a motorcycle. We have trully appreciate this visualization because it includes some innovative and original graphics, which are also entertaining and pleasant-looking. With the first look the reader understands immediately what the visualization is about and the essence of the information is instantly transmitted. The graphics about helmet use and about the types of collision are particularly striking, which is why we have inspired ourselves on this last visualization to create our "impact chart" graphic in our own visualization.

## 3 Available Data

After choosing the subject for our project and after searching for existing data visualizations based on road accidents, we have searched for open data online and we have found a data base containing information about road transport accidents in France [5].The data that we used for our project is available for download in the "Ressources" section.
Our data consists in multiple files (CARACTERISTIQUES, USAGERS, LIEUX, VEHICULES), each of which contain different kinds of information.

- Information went from people involved in accidents (USAGERS, containing data about the vehicle id, the affected part of the vehicle, the gravity of the accident, the sex of the person involved, the status of the person involved whether driver or victim, the year of birth of the people involved, the type of ride whether the accident was produced when the person was goind to work, or was heading back home or to the grocery store, the type of security available in the vehicle, etc).

- The characteristics of the accident (CARACTERISTIQUES, containing data about the date, hour, light, department, weather conditions, the address, a GPS code, latitude and longitude).

- The location where the accident took place (LIEUX, containing information about the type of road whether a National Road, a parking lot, etc, the road number, the type of circulation allowed, etc).

- And finally the vehicles involved (VEHICULES, containing data about the type of vehicle, the obstacles that could cause the accident, the impact zone, etc).

A useful description of this data is available online as well [7]. Several parameters are available for public exploitation, which allows us to come up with many different possible representations.

## 4 Project Description

For our visualization design, we took into consideration two main important points:

- Show in a simple and playful way the accidents having taken place in France in 2016. To use multiple types of charts in order to show as much information as possible in a reduced space.

- Doing this in a sensibilization purpose. There are still an enormous quantity of accidents and the cause of this is in part due to the fact that people are not well prevented because they would not bother to read some large paragraphs containing complicated information. Our visualization has to be designed in a way that information is easily and quickly understood.

### 4.0.1 Distribution of accidents per day

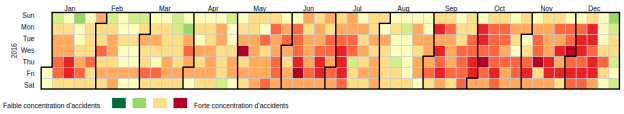


Figure 1: Calendar chart: A visualization of the distribution of accidents per day. Color code: days in red are days with the highest concentration of accidents, whereas days in green are days with the least concentration of accidents.

This first visualization shows the quantity of accidents per day in a 2016 calendar. Each day is represented by a square, and depending on the density of accidents that day, the square is colored with different shades, going from dark green to dark red. Days in dark green represent days with the least density of accidents, and the green goes lighter and turns into red as the accident counts increase per day. This representation in very visual and easy to understand, readers can quickly recognize what dates are the most susceptible to accidents since they are in bright red, for instance June the 6th, a date where a lot of students and workers start holidays so vehicle density in cities increases. This visualisation has also easy-to-understand legends and tooltips to indicate the date when we move the mouse on a square.

#### 4.0.2 Distribution of accidents per hour, detail on type of ride made

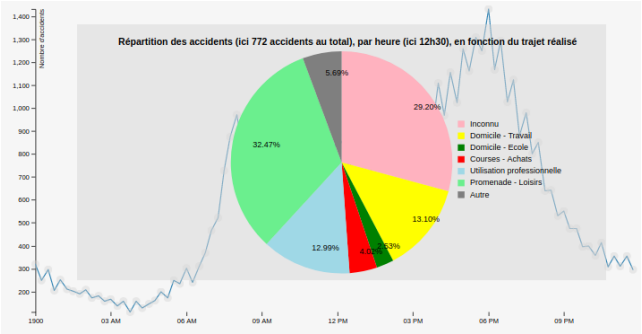


Figure 2: Line chart with ride type pie chart per quarter of an hour: A visualization of the distribution of accidents per hour. X axis: hours, Y axis: number of accidents. Each point represents a quarter of an hour, all of which contain a pie chart representing the percentage of ride types that people made during the accident: going to work, going back home, going to the grocery store, etc.

We present here an innovative visualization, because there is no existing graphic of this kind on the internet based on accident rate representation. This visualization aims to inform people about what hours of the day we have to be more careful in order to avoid an accident, because vehicle density in streets is not the same throughout the day. Readers can easily notice two peaks, one at 8am and another at 6pm, corresponding to the hours where people go and come back from work. Notice that each quarter of hour is represented by a dot in the line chart, and that each dot is interactive. By clicking on it, readers can acces to a pie chart giving detailed information about the type of ride that people involved in the accident where doing, it could be going to the grocery store, going back home from work, etc, in that quarter of hour. The pie chart has bright colors to easily distinguish from each type of ride. This visualization is very user-friendly as it only takes to click on a dot to see the pie chart and to click elsewhere to make it disappear, and click on another dot to display the pie chart corresponding to another quarter of hour.

#### 4.0.3 Distribution of accidents in terms of vehicle impact zone

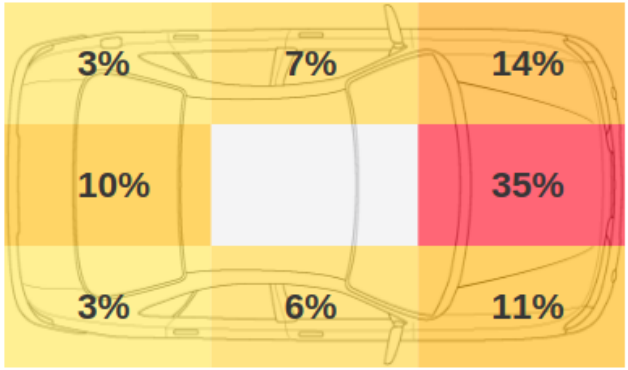


Figure 3: Impact Chart. A visualization of the distribution of accidents in terms of vehicle impact zone. The probability of impact for each vehicle zone is represented with shades of red, going from white to bright red. Read squares represent the vehicle zone the most susceptible to collide with another vehicle, whereas light yellow represents the zones the least susceptible to collide with another vehicle.

This is another of our innovative graphics, as we haven't seen a graphic of this kind on the internet. This visualisation aims to prevent people about the dangers of getting injured in an accident depending on the part of the car they are occupying. For instance, according to this visualization readers can easily come to the conclusion that the most vulnerable part of the car is the front seats, mostly the middle in bright red, as well as the sides in bright orange. Thanks to this graphic they will know that the danger of getting injured is higher when occupying this places. On each square we represented the percentage of times that another vehicle collided with these specific parts of the car during 2016 accidents.

#### 4.0.4 Distribution of accidents per sex


Figure 4: Stacked chart. Visualization of the distribution of accidents per sex. Color code: blue represents the percentage of accidents where the actor was a man, and in red the percentage of accidents where a woman was the accident actor.

This visualization is simple but precise and it efficiently transmits the piece of information that we wanted to transmit to the reader. According to this graphic, in 2016 the vast majority of accidents where caused by men. This is an important detail to know, because road security services could start making their accident awareness visualizations addressed specifically to men to make them more efficient.

#### 4.0.5 Distribution of accidents per age groupe and sex

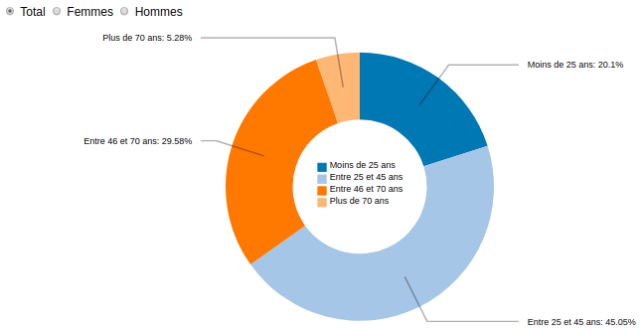


Figure 5: Dynamic donut charts. A visualization of the distribution of accidents per age groupe and sex. This visualization contains 3 representations, depending on if we take into account the totality of accidents, accidents where only women where involved or only men where involved. Dynamic radio buttons are available on top of the visualization, in order to change from one sight to another.

This animated and interactive visualization aims to alert people about the most vulnerable age groups in terms of accident victims. Depending on which type of representation the user wants to display (total, women or men), a different donut chart is going to be generated. If we click on "total", the donut chart is going to represent the percentage of accidents where different age groups where involved in the totality of accidents. If we click on "women", the same percentages are going to be calculated but this time on the basis of accidents where only women where involved, and if we click on "men", the calculations are going to be made on the basis

of accidents where only men where involved. Users can easily realise that in the three cases, the most vulverable groupe is between 25 and 45 years old. This is a very user-friendly visualization, it only takes a click to change from one representation to another, and dynamic tooltips and legends allow better understanding of information provided. The colors are produced automatically with a D3 color palette. This representation was inspired by this work [3].

#### 4.0.6 Distribution of accidents in terms of gravity

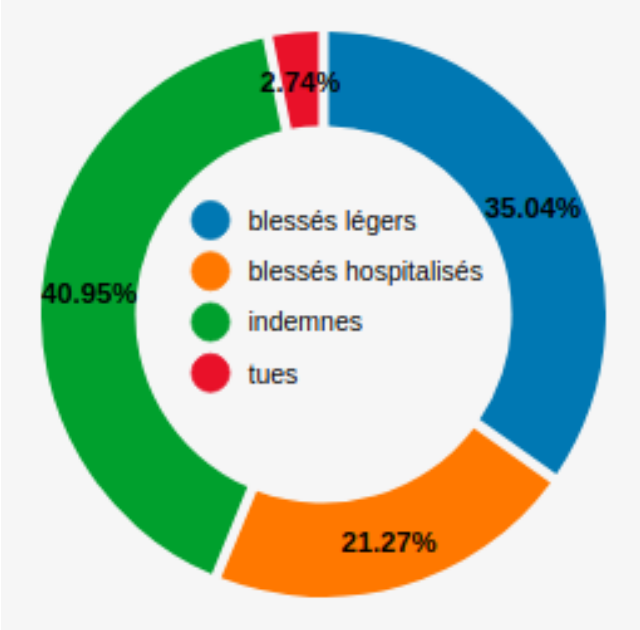


Figure 6: Donut chart. A visualization of the distribution of accidents by gravity. Color code: fatal victims in red, hospitalized injured people in orange, mildly injured people in blue and uninjured people in green.

This simple visualization informs us about the gravity of most accidents. Each part of the donut chart represents the percentage of fatal, mild, hospitalized and uninjured accident victimes in 2016. Users can easily realise that most of accident victims result uninjured or mildly injured.

#### 4.0.7 Distribution of accidents per geographic localisation

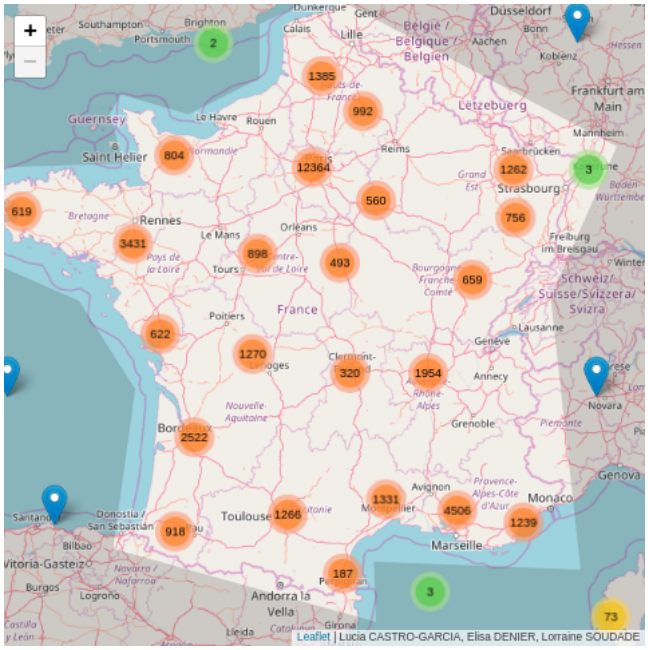


Figure 7: Map chart. A visualization of the distribution of accidents per geographic localisation. The French map is generated with the leaflet library [1]. Accidents are clustered by region in colored bubbles all over the map, indicating the density of accidents in each region, going from green (low density) to orange (high density).

This final visualization aims to prevent users about the French regions and departments the most vulnerable in terms of accident rate. This could raise awareness in people when travelling of they choose a destination with a higher rate of accidents. This is a very user-friendly and interactive representation, allowing the user to zoom in and zoom out on the French map to see the accident rate in not only each region, but each department and even each city. The colored bubbles will automatically regenerate when zooming in or out, according to the new calculation of accident rate. If you zoom in enough, you will be able to get the detail of the exact street name where a certain accident took place.

## 5 Discussion

### 5.0.1 Strong points

- We have a very complet visualization: each graphic provides useful information about accident factors in France in 2016. They are user-friendly and readers can easily draw conclusions by themselves about accident facts and factors. They are not overloaded with complicated information, and even if they are very complete they are not difficult to understand.

- Our visualization covers many aspects of accident factors and also the context of the accident, which is an information that people forget to represent most of the time. For istance, there is no existing visualization about the type of ride that people where doing in the moment of the accident.

- We have created some innovative and original visualizations, in particular the vehicle impact chart and the hour line chart, with details in ride type represented by multiple pie charts generated interactively. These graphics did not exist on the internet, so we could say that a really positive aspect of our

work is to bring new information that could possibly interest more people and raise accident awareness more efficiently.

- All of our visualizations are accessible for everybody, no matter the age group, and its quick understanding could possibly increase the efficiency of the information transmission. They are interactive, easy to use. Each graphic uses a pertinent color code, with different shades representing increasing or decreasing densities, from green to red, from yellow to red, and our color code was mostly universal. Most of the time, red colors defined a high density of accidents, for example in the calendar chart.

- Our web interface is also user-friendly and pleasant-looking. For a better layout we have used boostrap4 [1], an HTML, CSS, and JS framework. For an easier navigation, we have generated a navigation bar on top of the web page for the user to only click on each label to be able to reach each graphic quickly.

#### 5.0.2 Weak points

- Our wisualization certainly is complete, but we have many independent graphics which are not mutually connected, so most of the visualizations are at best 2D, but we have not reached a visualization where multiple dimensions are represented at once.

- Universalize the color palette. We have used pertinent colors in ours visualizations and we have tried to keep the same colors for extreme values (the lowest or the highest), but it would be interesting to keep a fixed color palette for all visualizations in order to be sure that the reader does not get disturbed by the colors changes from one visualization to another.

- With our visualization it is not possible to represent for instance, the mean gravity of accidents per region in France. It would have been interesting to plot different D3 graphics (donut charts or line charts or other) on top of the French map, and that they could be automatically generated when we change regions or zoom in and zoom out.

## 6 Conclusion

Our initial goal was to respond to a lack of global visualizations on the subject of accident rate in France, always keeping in mind to satisfy the unexperienced reader. This goal has led us to produce visualizations which are adapted to the public we are aiming for, that is to say, the widest possible audience, since the entire population is concerned by road security issues.
For this matter, we have tried to be really careful in ensuring the simplicity and understanding of our representations, with emphasis on color palettes that are meaningful and intuitive for the reader. In addition, we cared to give a meaningful name to our graphics so that the user captures the representation message at a first glance. The captions accompanying the graphics, as well as the tooltips and legends are present to accompany the reader in his understanding of graphics. Furthermore, we have ensured at least a little interactivity in our graphics for those who wish to go deeper into the details of certain representations, without however complicating the visualization and compromising its quick understanding.
The different angles shown cover a large part of the subject, and keep the interest of the reader over the visualization. Finally, while keeping in mind that many improvements can be made, whether at the interactive level and in detail, this project is a good starting point in the field, which until then was strikingly lacking global and precise representations on this subject.

## References


[1] V. Agafonkin. Leaflet, an open-source javascript library for mobile-friendly interactive maps, 2017. Available at http://leafletjs.com/.

[2] F. Casse. Survivre en moto: les chiffres à connaître, May 2013. Available at https://www.francecasse.fr/blog/wp-content/uploads/2013/05/Accident

[3] J. Cruz-Benito. Animated donut chart with labels, legend and tooltips.

[4] C. Magdelaine. La place de l'automobile dans nos sociétés : faits et chiffres, November 2016. Available at https://www.notre-planete.info/ecologie/transport/placeauto.php.

[5] N. Mergitur. Base de données accidents corporels de la circulation, September 2017. Available at https://www.data.gouv.fr/fr/datasets/base-de-donnees-accidents-corp

[6] O. national interministériel de la sécurité routière. L'accidentologie routière sous l'angle hommes/femmes, May 2016. Available at https://www.permisapoints.fr/assets/uploads/2016/06/infographie-fem

[7] O. national interministériel de sécurité routière. Description des bases de données annuelles des accidents corporels de la circulation routière, September 2017. Available at https://www.data.gouv.fr/s/resources/base-de-donnees-accidents-corp

[8] Securite-routiere.gouv.fr. Bilan définitif de l'accidentalité routière 2016, June 2017. Available at http://www.securite-routiere.gouv.fr/medias/espace-presse/publicati