Università degli Studi di Milano Bicocca
Master Degree in Data Science
Master Degree in Theories and Technologies of Communication

Visualizing Road Accidents in Great Britain (2005-2015)

Project report for the Data Visualization course

Danilo Lanzilotti (887967)
Luca Iarocci (894066)
Marco Lariccia (887508)
Samir Doghmi (897358)

academic year 2022/2023

CONTENTS

1. Introduction 2
1.1 Great Britain: a Sense of Scale 3

2. Data 4
2.1 Accidents 4
2.2 Casualties 4
2.3 Vehicles 5

3. Infographics 6
3.1 Understanding Road Accidents in Great Britain 6
Figure 3.1 Understanding road accidents in Great Britain 6
3.2 Distribution of Road Accidents over Month, Day of the Week, and Hour 9
3.2.1 Absolute terms 9
3.2.2 Relative terms 11
3.3 The Role of Gender and Age in Road Accidents 11
3.4 The Role of Vehicle Type and Gender in Road Casualties 13
3.5 Insight into the Age Group of Casualties 14

4. Assessment 17
4.1 Users 17
4.2 Heuristic evaluation 17
4.2.1 Our tasks 17
4.3 Psychometric questionnaire 18
4.4 Results 19
4.4.1 Heuristic test 19
4.4.2 Heuristic evaluation 20
4.4.3 Psychometric questionnaire 23

5. conclusion 26

1. Introduction

Every death and serious injury on the road is a preventable tragedy and yet, on average, in the last decade five people have died every day on the road in the UK and 84 were seriously injured.[1]

While overall road accidents have decreased significantly over the past two decades, the rate of decline of road casualties has significantly slowed down since 2010[2]. Studying the causes of road accidents as well as the circumstances in which they occur is crucial when it comes to further reducing the number of related injuries and death. Improving road safety is an ongoing effort as demonstrated by the numerous actions taken by the government and the constant monitoring of the phenomena through an in depth data collection regarding these events by the Department of Transport. In July 2019, the government released the Road Safety Declaration, which enunciates a two-year action plan containing 74 action measures. The statement emphasizes a shift in thinking toward a safe system approach. Kevin Marshall, director of TWM Traffic Control Systems, also said - "We want to develop systems that increase road safety, for both drivers and pedestrians: our new hazard warning sign was designed with motorcyclists' safety in mind, and we hope local authorities across the UK will review signage at their known accident blackspots, especially now that the darker, wetter winter months are coming, with roads therefore even more dangerous."

Furthermore the UK emerges as one of the most frightening states for tourists struggling with car driving. Driving in the UK, of course, is not impossible, but it’s important to always follow rules, limits and advice carefully. Moreover, car rental companies put stickers in cars reminding the driver about the key thing: keep to the left. One of the main risks, in fact, for those about to drive in the UK is precisely having to use left-hand drive. However, there are risks from which having to pay attention becomes an obligation to safeguard one's own life and the lives of others, and these stem from the likelihood of a traffic accident.

Through this project, our goal is to visualize the data regarding accidents to understand what are the general characteristics of these occurrences, focusing primarily on the role of age, gender and means of transportation. Our meta-objective is configured, instead, to try to raise awareness of this issue, try to understand how important it is to take into account a number of factors to avoid, as much as possible, these unfortunate events.

[1] https://www.brake.org.uk/get-involved/take-action/mybrake/knowledge-centre/uk-road-safety
[2] https://www.gov.uk/government/collections/road-accidents-and-safety-statistics

1.1 Great Britain: a Sense of Scale

Here are reported a series of information regarding Great Britain that may help contextualize the numbers of this phenomenon. According to the Department of Transport, as of 2021:

- Population: 67.1 million
- Road network: 424,480 km
    - 37% urban roads
    - 62% rural roads
    - 1% motorways
- Registered motor vehicles: 39.8 million
    - 82% cars
    - 12% goods vehicles
    - 3% motorcycles
    - 3% others

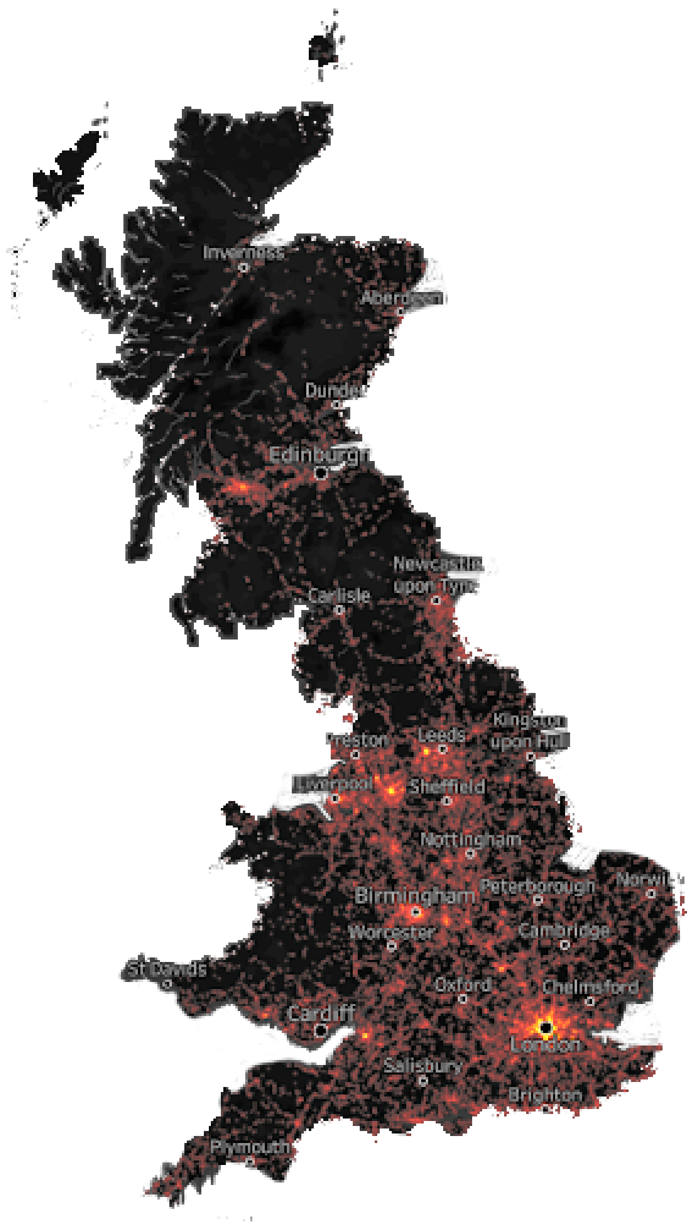


Figure 1.1 Geographical distribution of fatal accidents over the 2005-2015 period

2. Data

The dataset[3] used in this project is a collection of datasets published by the UK Department of Transport[4] containing several informations about all the reported road accidents during the period 2005-2015. The Data is recorded and reported by the traffic police on-scene who conducted assessments and provided feedback of the dynamics of the scenario to the Department for Transport by filling the STAS19 form[5]. The data is organized in three tables, each containing information about different aspects of the accidents: Accidents, Casualties, and Vehicles.

In this context a road accident is defined as a road accident resulting in human death or personal injury that have been notified to the police within 30 days of occurrence, and in which one or more vehicles are involved.

To perform the visualization of these data on Tableau, we combined the three data sets by Incident Index, resulting in a set of 70 columns and 4,287,593 observations. Due to the extreme level of detail in which each event is described a selection of the feature to visualize was needed. Following a brief description of the aforementioned tables covering only the features used in this project.

2.1 Accidents

The first dataset "Accidents.csv" contains the following information:

- Accident Index: The incident reference is a unique code that contains a total of 7, or fewer, numbers and/or alphabetic characters. It allows the police to identify the incident in the event of a query;

- Accident Severity: can be fatal, serious or slight. In the event that multiple injuries are reported per accident, the severity of the accident takes the value of the highest severity degree of of the affected individual;

- Date: date on which the accident occurred;

- Time: time on which the accident occurred.

2.2 Casualties

The second dataset "*Casualties.csv*" contains the following information:

[3] https://www.kaggle.com/datasets/silicon99/dft-accident-data
[4] https://www.gov.uk/government/collections/road-accidents-and-safety-statistics
[5] https://www.gov.uk/government/publications/stats19-forms-and-guidance

- Age Band of Casualty: age group of casualties involved in a road accident;

- Age of Casualty: exact or estimate age of wounded parties implicated;

- Casualty Class: can assume values driver/rider, vehicle/pillion passenger or pedestrian;

- Casualty Severity: could be fatal, serious or slight.

    - 'Fatal' injury includes only those cases where death occurs in less than 30 days as a result of the accident. 'Fatal' does not include death from natural causes or suicide.
    - Examples of ‘Serious ‘ injuries are: Broken neck, concussion, fracture, deep cuts and burns.
    - Examples of 'Slight' injuries are: Whiplash, Bruising, shallow cuts and slight shock requiring roadside attention;

- Sex of Casualty: may take value male or female. If the casualty is a driver, then the sex entered here must be the same as that in Sex of Driver.

2.3 Vehicles

The third dataset Vehicles.csv contain the following information:

- Age Band of Driver: age band of driver involved in a road accident;

- Age of Driver: exact age or estimate. If a driver is injured, then the age entered here must be repeated in Age of Casualty for the driver casualty record for this vehicle;

- Sex of Driver: may take value male or female;

- Vehicle type: Vehicles (other than cars or mini-buses used as taxis or private hire vehicles) are coded according to their construction and not according to their use at the time of the accident. For example, a van which is being used for the carriage of passengers should nevertheless be coded as a goods vehicle

3. Infographics

We realized a series of interactive visualizations organized in a Tableau story. Each page of our story proposes a different perspective on the matter, aiming to give the viewer an overall understanding of the characteristics of this phenomenon. Our work revolves around different aspects of an accident, namely the time and date in which it occurred, the age and sex of the people involved, the type of vehicle involved, and the severity of the accidents as well as the casualties. Through the use of interactive filters we aim to make the viewer aware and at the same time curious about the phenomenon leaving room for an autonomous exploration of the data based on the user inputs; while text is used to help and guide the user in noticing peculiarities or relevant trends in the data. Following is a detailed analysis of the realized visualizations composing our story. Follow this link to see and interact with the final product of our work.

3.1 Understanding Road Accidents in Great Britain

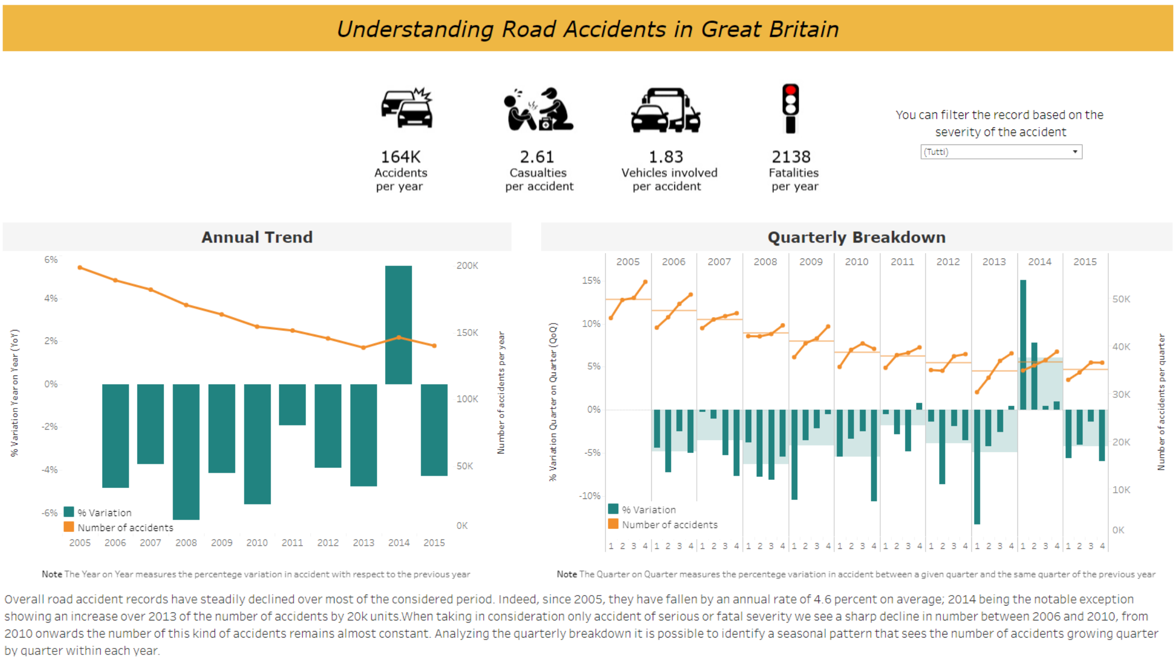


Figure 3.1 Understanding road accidents in Great Britain

In this page (Figure 3.1) we aim to give a general understanding of the scale and trend of the phenomenon over the period in analysis. This is achieved through three main visual elements: a didascalic visualization (Figure 3.2) reporting some relevant statistics regarding the matter, an annual trend analysis (Figure 3.3), and a quarter trend analysis (Figure 3.4).

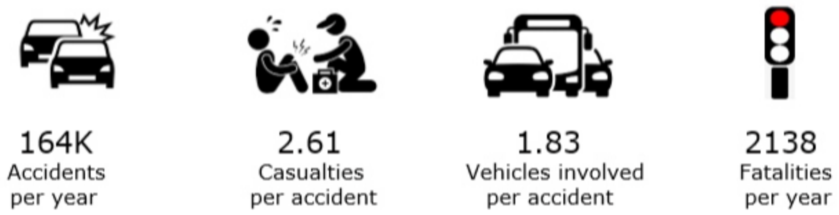


Figure 3.2 A didascalic visualization of some statistic about road accidents

A visual representation of the annual trend (Figure 3.3) is achieved with a dual axis graph in which the absolute number of accidents per year and the relative percentage variation, compared to the previous year, are shown simultaneously. Although the overall trend and scale of the phenomenon can be deduced by only looking at the line plot of the number of accidents, the percentage variation year on year stresses the improvement or worsening of the situation. Colour palette aside, to better separate the information shown, we decided to use a bar chart for the percentage variation while maintaining a line chart style for the representation of the number of accidents.

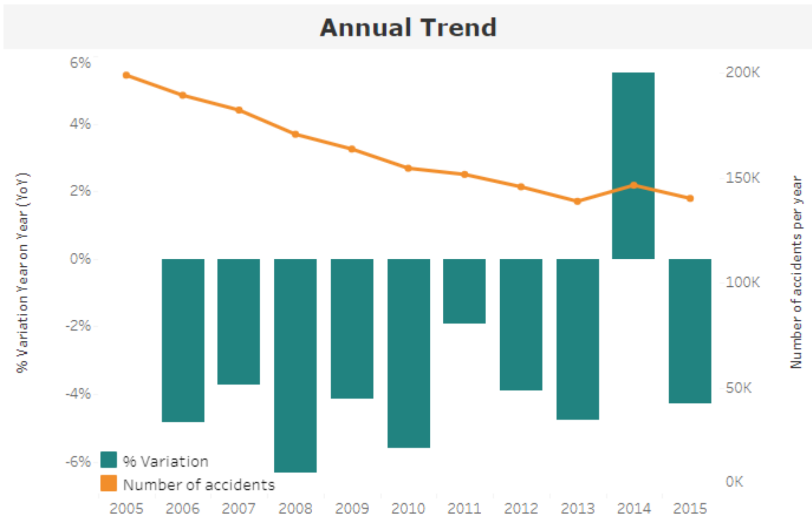


Figure 3.3 Annual trend of accidents in Great Britain over the period 2005-2015

Beside showing the annual trend, we decided to provide a quarterly breakdown of the situation (Figure 3.4) with the same graphic style shown before, that is to say a dual axis graph in which the absolute number of accidents per quarter and the relative percentage variation, compared this time to the same quarter of the previous year, are shown simultaneously. This graphic, paired with the previous one, gives a more in-depth understanding of the phenomenon, highlighting a quarter seasonal trend, with the numbers of accidents growing throughout the year quarter by quarter, and the periods characterized by the most drastic variations, both positive and negative. In addition, the annual average of both the number of accidents and its percentual variation are recalled as a line and area respectively, the color of which is a desaturated version of the color associated with the dimension of which represents the average.

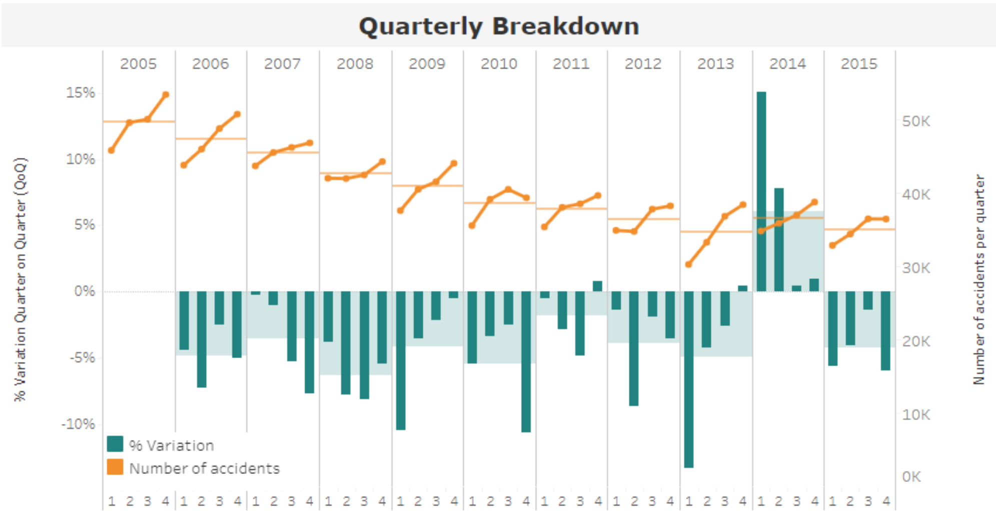


Figure 3.4 Quarter trend of accidents in Great Britain over the period 2005-2015

Users can interact with the graphics through a filter, available on the top right section of the page, giving the possibility to visualize only a subgroup of the data based on the severity of the accidents. The annual and quarterly trend graphics change dynamically based on the selection leading to a further understanding of the trend and scale of the matter by class of severity. An example is provided in Figure 3.5, in which a comparison between the long period trend of all accidents and fatal accidents is proposed.

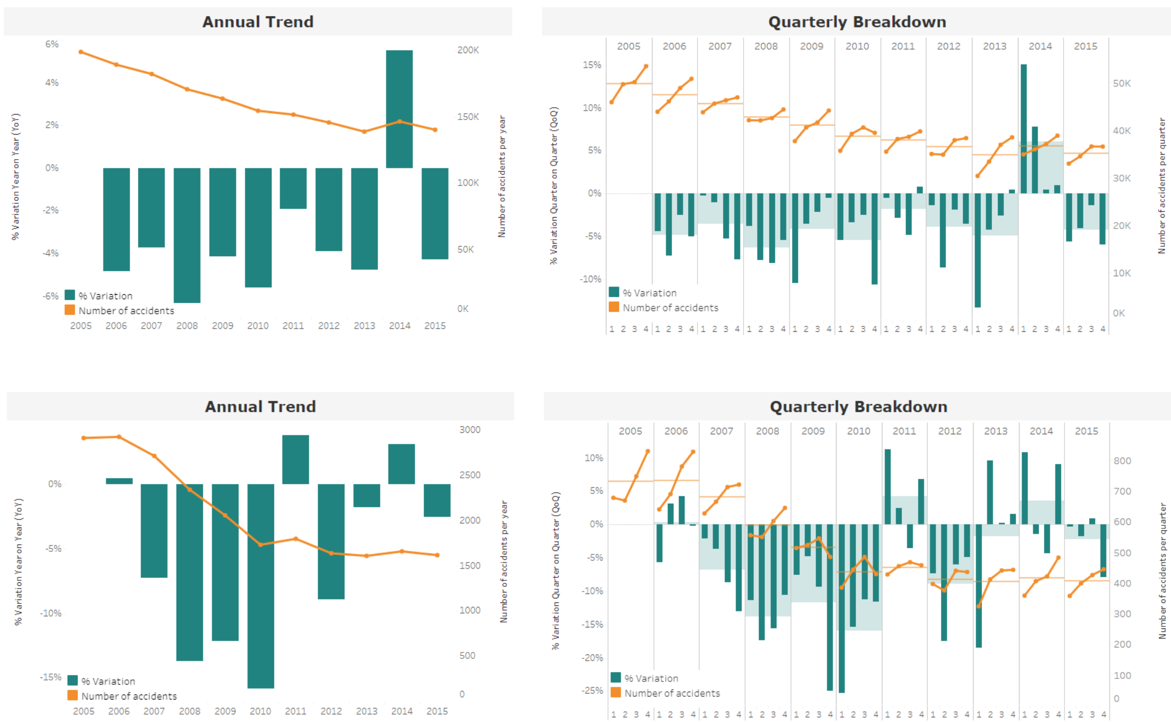


Figure 3.5 The effect of the filter on the annual trend and the quarterly breakdown; all severity accidents (top image) compared to fatal accidents (bottom image)

Filters are a fundamental part not only of this first page, but of the whole story. The user interaction through filters is a key aspect when it comes to exploring and understanding of the complex phenomenon of road accidents. We decided to introduce this aspect from the beginning with the intent of getting the user acquainted with this tool and showing its potentially drastic effects even on a simple graph as the ones shown in this page.

3.2 Distribution of Road Accidents over Month, Day of the Week, and Hour

The second point of our analysis of the road accidents phenomenon revolves around the study of its distribution over month, day of the week, and hour of the day with the intent of temporally characterizing the phenomenon and highlighting its seasonality relative to different time scales. The subject is explored in two pages of identical structure, presenting respectively the distribution of road accidents in absolute and relative terms.

3.2.1 Absolute terms

In this first page (Figure 3.6) on the distribution of road accidents over month, day of the week, and hour of the day we focus on the raw numbers of the phenomenon, answering questions like: When - in terms of hour of the day, day of the week or month - we have registered peaks in the number of accidents? With this goal in mind, we realized a visualization composed of four graphics aggregating the data by month, day of the week, hour of the day, and both day of the week and time of the day. The aggregated data is visualized through the use of three bar charts and a heatmap. A brief textual comment points out some interesting patterns and may help the user in reading and exploring the graphics.

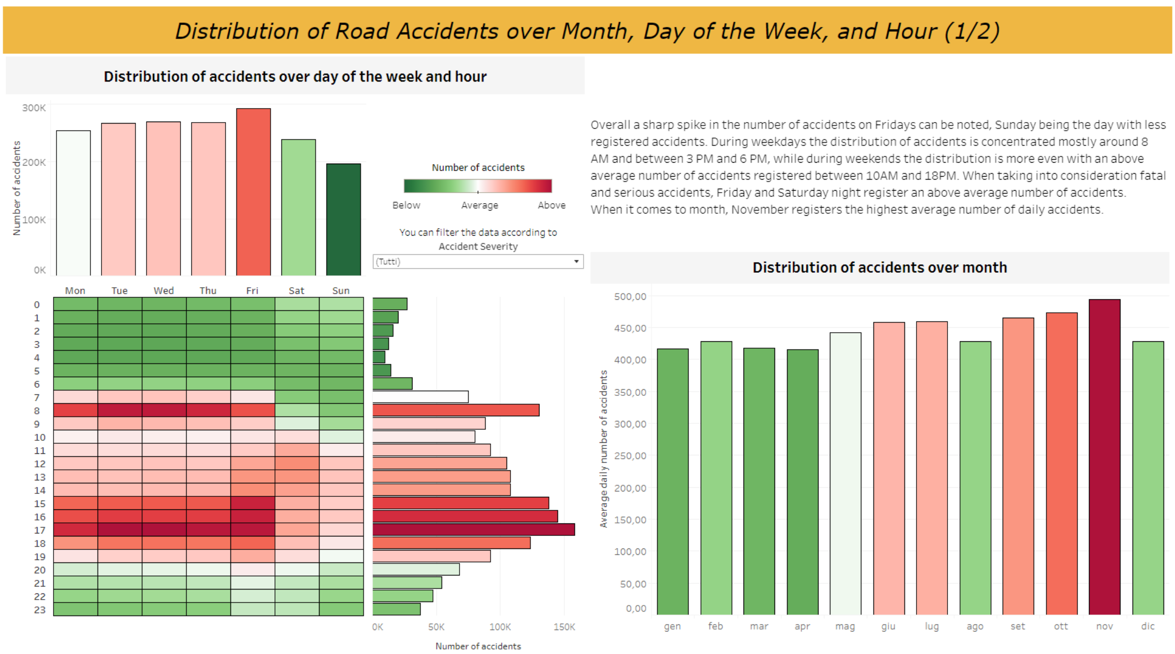


Figure 3.6 Distribution of Road Accidents over Month, Day of the Week, and Hour (1/2)

The distribution over day and hour and the relative combined distribution are bundled together in a single visual group composed of a heatmap, representing the combined distribution over day and hour, and two bar charts, representing the marginal distribution over day and hour of the number of accidents. The three graphs share the same color scheme, a red-green-white diverging palette centered on the average of the respective graph. The white corresponds to the average, while intuitively the reds are assigned to those points in time which have registered more accidents than the average, and the greens are assigned to those that have registered less accidents than the average. While the color is an essential part of the heatmap, in the case of the bar charts is an example of visual redundancy conveying, apart from the relative position relative to the average of the distribution, the same information of the length of the bar.

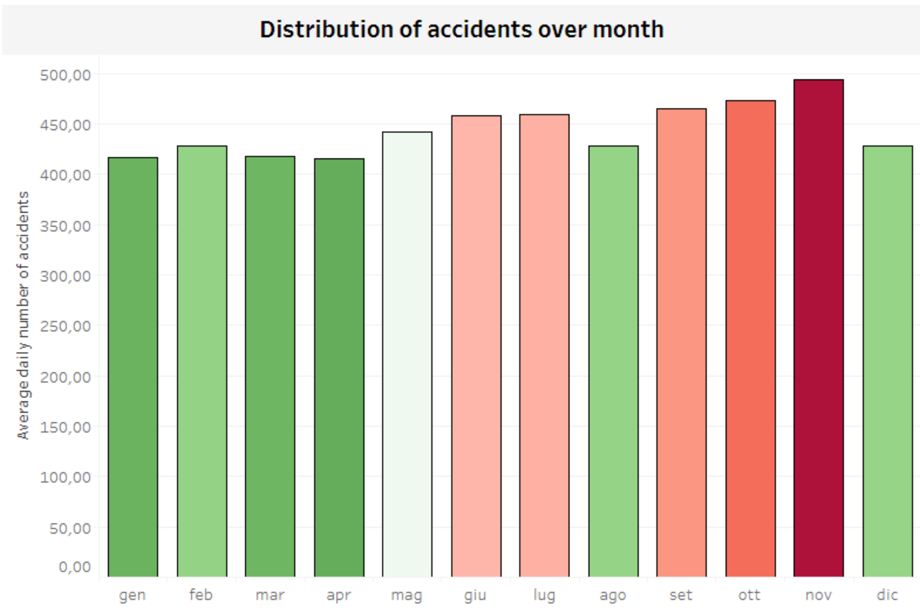


Figure 3.7 Distribution of accidents over month

The analysis of the distribution of accidents over month (Figure 3.7) is normalized by the number of days in a month, thus the quantity represented in this graph is the average daily number of accidents by month. Normalizing the data when it comes to aggregations over months is enìxtrimely important, indeed without this step we assist to a distortion of the underlying data resulting from the difference in the number of days each month has. The graph shares the same color palette as the other graph of the page.

Similar to what is found in the prior page, users can interact with the graphics through a filter, positioned under the color legend, giving the possibility to visualize only a subgroup of the data based on the severity of the accidents. All four graphics change dynamically making possible comparisons between the distributions of accidents over time by class of severity.

3.2.2 Relative terms

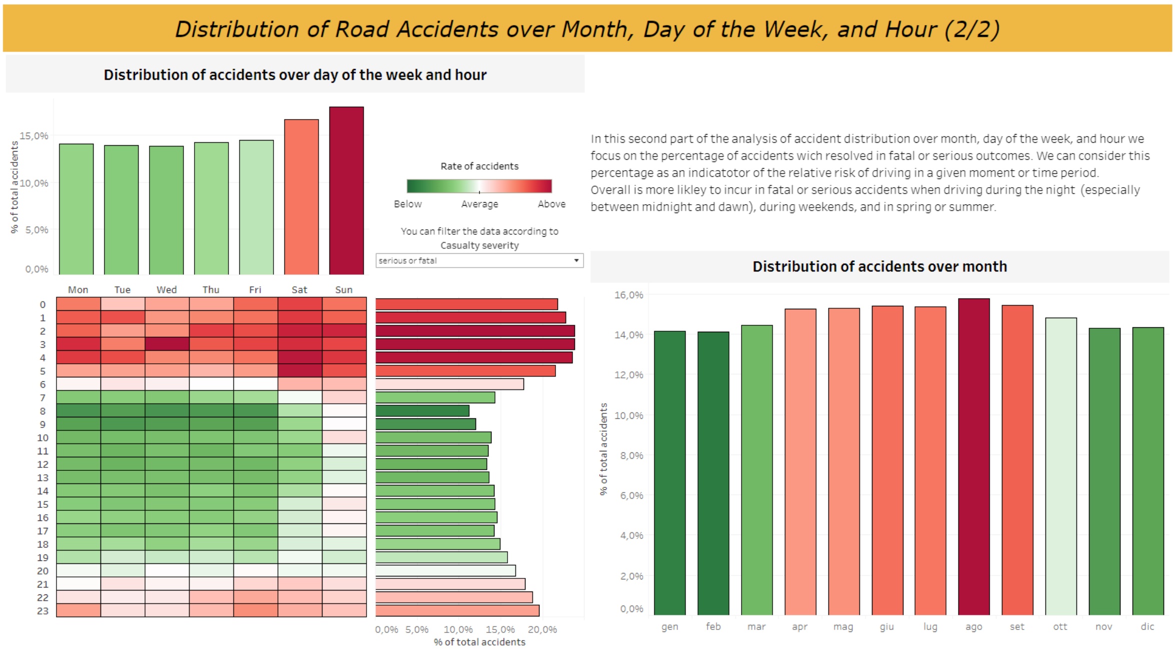


Figure 3.8 Distribution of Road Accidents over Month, Day of the Week, and Hour (2/2)

In this second page (Figure 3.8) on the distribution of road accidents over month, day of the week, and hour of the day we focus on the percentage of the total accidents registered in a given time interval resolved in severe or fatal outcomes. This approach shows when it is relatively more risky to be involved in a road accident and emphasizes the statistical probability of incurring serious or fatal injuries if involved in an accident in a given time interval. The visualization shares the same visual structure and color palette as the previous page, the change of perspective on the subject being the only difference between the two.

Also in this page it is possible to interact with the graph through a filter selecting between “fatal” and “serious or fatal”. The four graphics composing the visualization change dynamically showing which percentage of the total number of accidents registered in a given time interval resolved in either “fatal” or “serious or fatal” outcomes.

3.3 The Role of Gender and Age in Road Accidents

In this third dashboard (Figure 3.9), our objective is to give a visual representation of the distribution of age groups, sex and casualty severity for the number of drivers and casualties injured in a road accident. The topic is explored in graphs of the same layout, in order to facilitate an immediate side-by-side confrontation between the modules. This is achieved by means of two main elements: two butterfly charts and two sets of pie charts that enable to identify the role gender and age plays in road accidents.

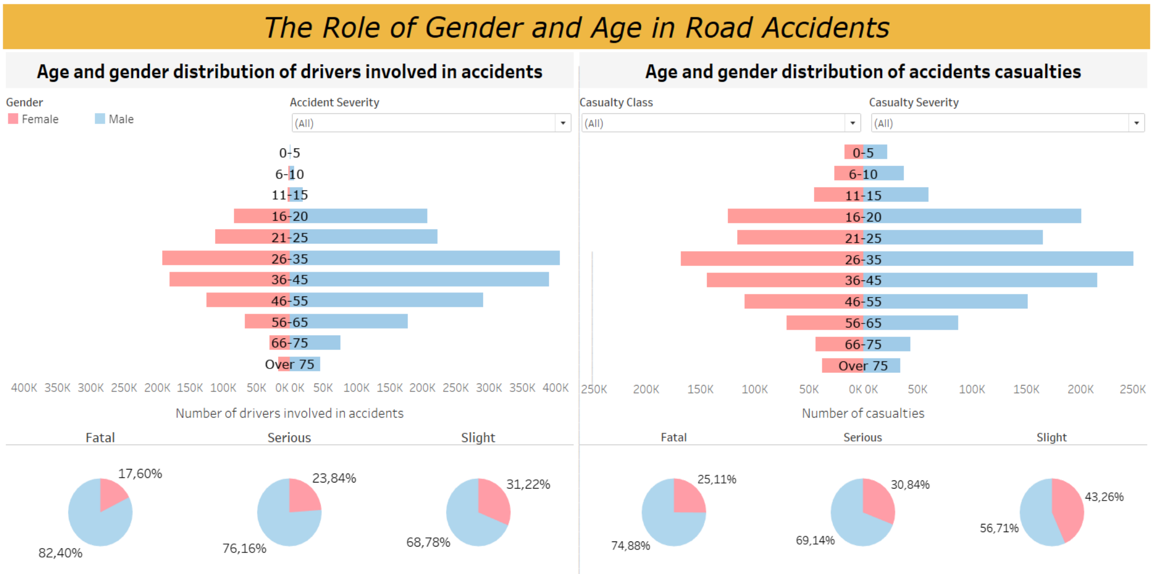


Figure 3.9 The role of gender and age in road accidents

In order to separate female and male gender, pink and blue were employed as colors. The filtering of data by accident severity is also provided in both graphs. Moreover, the same attribute is expanded, at the lower half of the dashboard, into 3 pie charts showing the proportions of men and women in fatal, serious and slight accidents severity.

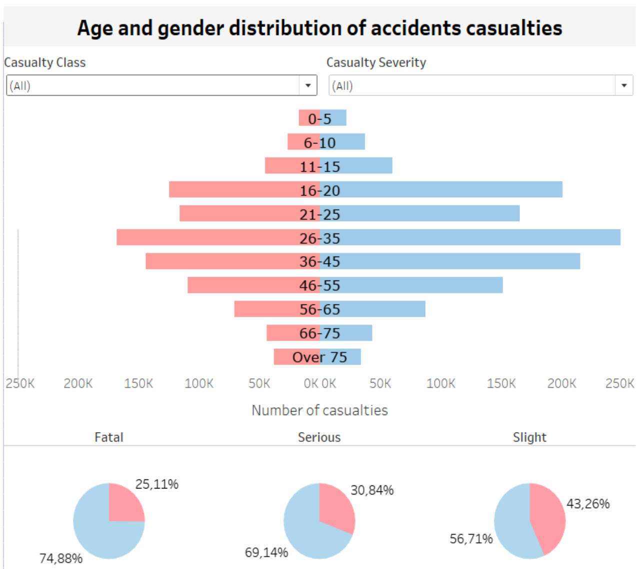


Figure 3.10 Age and gender distribution of accident casualties

In view of the implementation and application of more in-depth analyses of the data, in the second half of the story, the age and gender distribution of accident casualties (Figure 3.10), can be interacted by a second parameter by selecting between “All”, “driver/rider”, “pedestrian”, and "veh./pillion passenger”.

3.4 The Role of Vehicle Type and Gender in Road Casualties

On this page (Figure 3.11) we focus on the percentage of total road accidents by vehicle type.This angle shows when it is comparatively riskier to be personally injured in a traffic accident and highlights in absolute numbers how many drivers and casualties suffered slight, serious, or fatal severity when involved in an accident for a given vehicle. The visualization shares the same color palette as the previous dashboard, with the only difference being a more detailed focus on the topic.

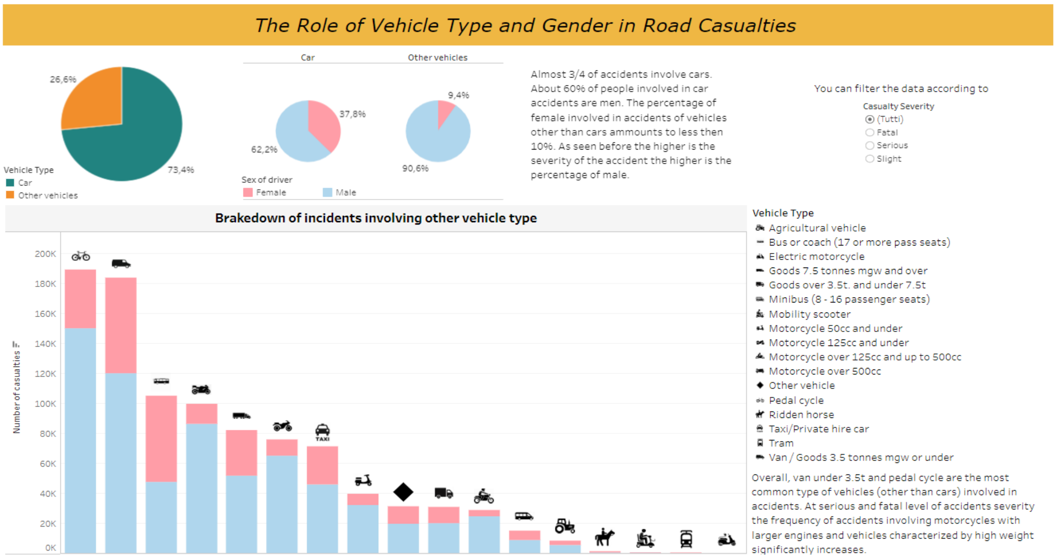


Figure 3.11 Age and gender distribution of accident casualties and drivers by vehicle type

Starting from the right, the focus was first on the type of vehicle driven by the drivers reported. Since there was a net skew toward car accidents, as they are implicated in almost 3/4 of the records, it was chosen to deploy a pie chart and group the rest of the vehicles into the "other vehicles" variable. The percentage of females involved in accidents of vehicles other than cars amounts to less then 1/10. As seen above, the higher the severity of the accident, the greater the percentage of men involved.

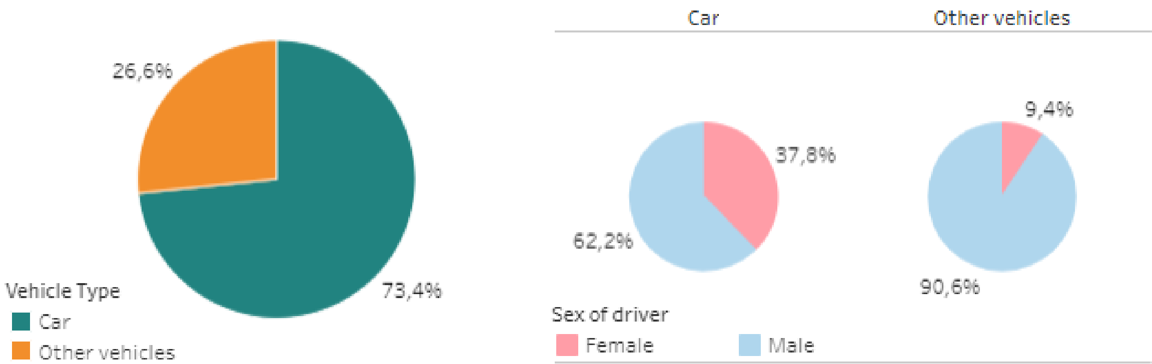


Figure 3.12 Breakdown of the typology of vehicles concerned with subsequent insights into the weight distribution of gender

In this section, it has been decided to show, in a histogram, the number of accidents casualties by "other type of vehicle" attribute and by gender of the affected individuals. In an effort to simplify the readability of the chart, some icon visualizations have been placed on top of the matching bins. These have been sorted in descending order and explained with a legend alongside. Again, users can interact with the graphs through a filter, available in the top right-hand section of the page, which gives the option of viewing only a subset of data according to the casualty severity.

In general, a consistent pattern in the displayed items is discernible: across all vehicles the majority of victims are male.While referring to vehicular category, vans under 3.5t and pedal bicycles are the most common vehicle among those injured in accidents. In terms of the degree of gravity of serious or fatal accidents, vulnerable road users (motorcyclists and cyclists) are more likely to be involved, than would be expected. Moreover, The frequency of accidents with the highest number of fatal victims are motorbikes with an over 500 cylinder capacity.

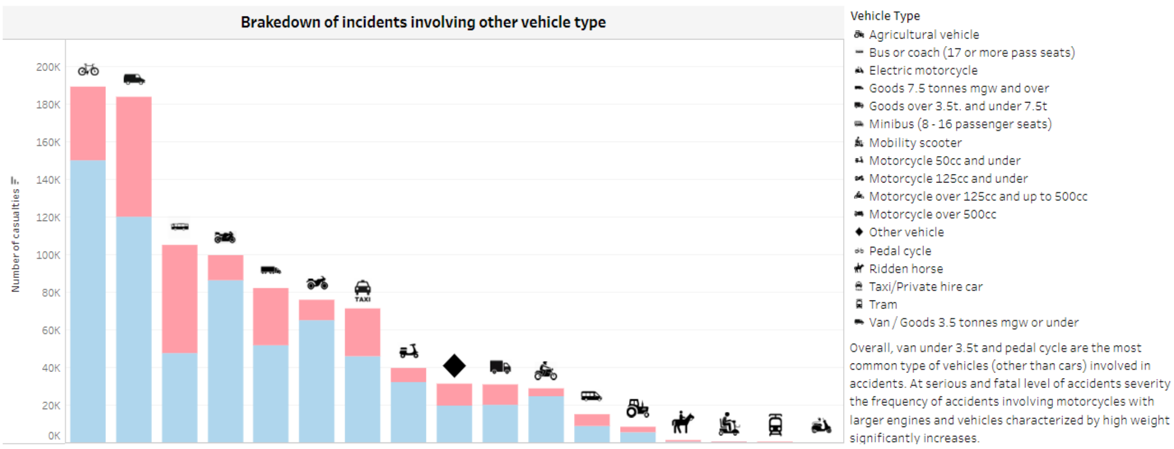


Figure 3.13 Age and gender distribution of accident casualties

3.5 Insight into the Age Group of Casualties

Once having taken a closer look at the gender of drivers and injured people in the police records, in this final dashboard (Figure 3.14) of our story we analyze the breakdown of accidents by day of the week and by hour with reference to age groups. The purpose of the legend and filters is to highlight the several trends over the span of the day and the week, and enable the user to customize and 'play around' with the class Casualty and casualty severity filters, already introduced on the previous pages. The identical pattern of colors have been used to emphasize a given age group, making it easier to understand the charts.

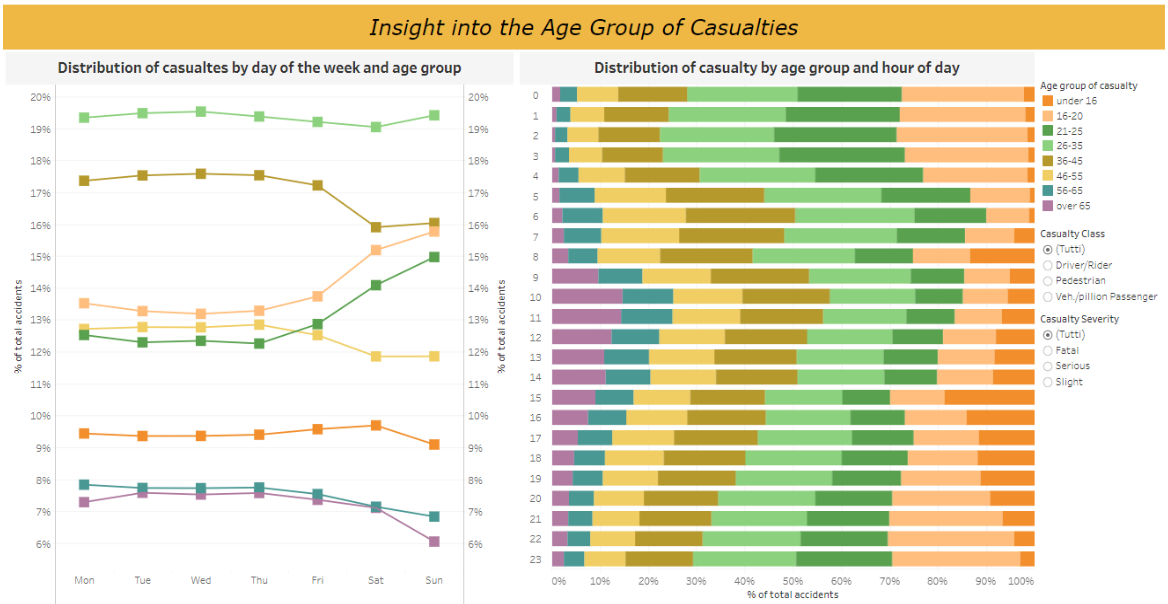


Figure 3.14 Insight into the age group of casualties

A visual representation of the trend over the reference period 2005-2015, is obtained by means of a double-axis graph showing the percentage of accidents for each age band of casualties involved, in parallel on the x-axis are shown the days of the week. By linking each of the marks of the same color, the figure is more effective and useful for discerning various hints.

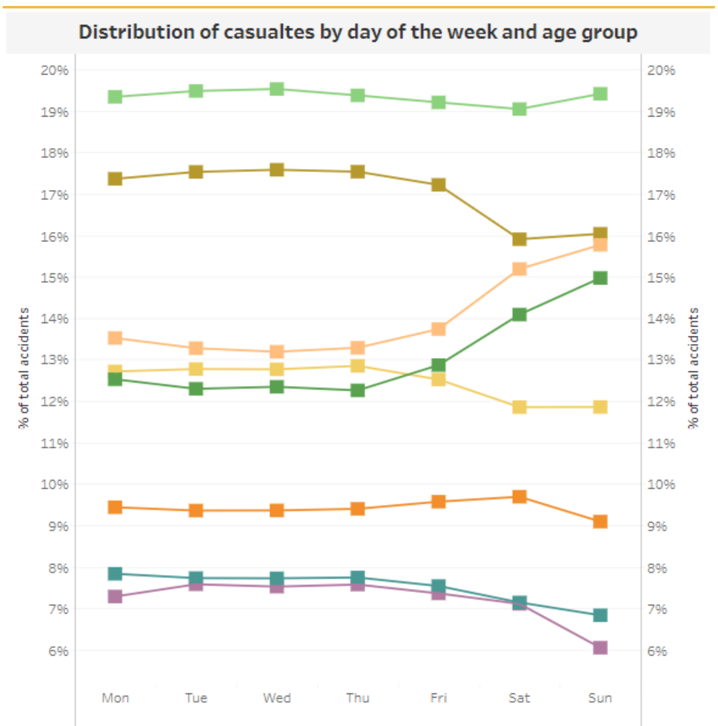


Figure 3.15 Distribution of casualties by day of the week and age group

It is noteworthy that the risk of getting involved in a road accident is proportionally related to the lifestyle level of each age class: For instance, young people 16-20 are more likely to be injured on Friday through Sunday; while the trend seems more stable during the rest of the week. In turn, the 55-64 age band shows an opposite swing.

In Figure (3.16), a stacked chart is portrayed which establishes a relationship between the time and the percentage of accidents for each age group involved during the day. The graph is intended to show the differences between the various classes.

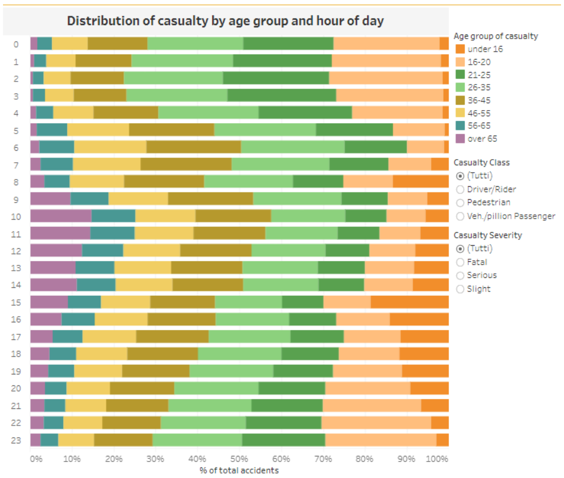


Figure 3.16 Distribution of casualty by age group and hour of day

What follows is an example of how the data could be filtered, and the changes from the baseline setting are remarkable. It can be seen how older pedestrians are over-represented in Fatal severity. A major cause of the high mortality rate of pedestrians over 65 may be related to the physical vulnerability of the elderly. Since their bones are more fragile and their soft tissues less elastic, they are at greater risk of serious injury, even if the impact forces are the same. At the same time, the elderly are more likely to be injured in an accident because loco-motor functions deteriorate with advancing years.

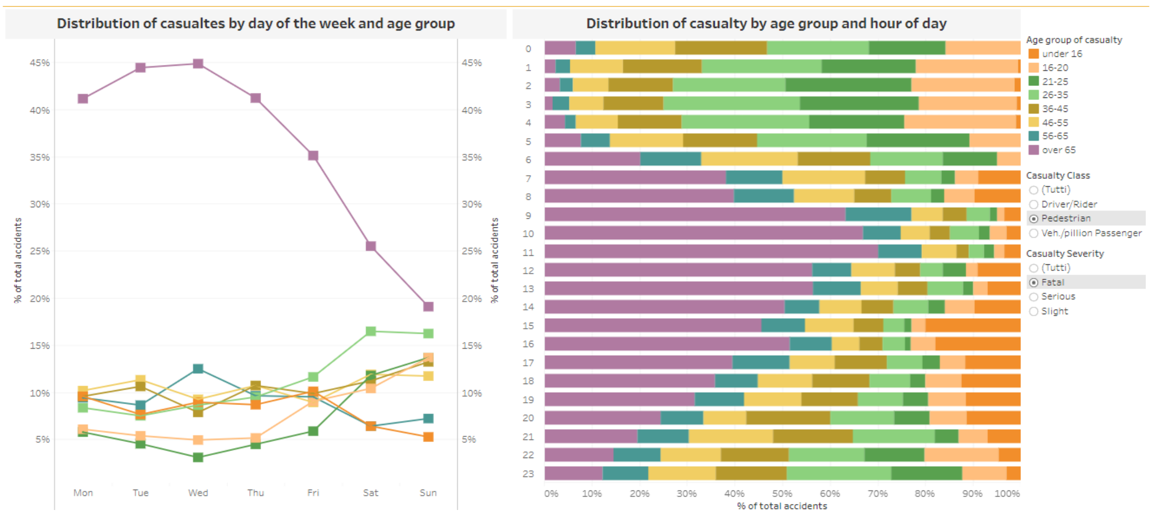


Figure 3.17 The effect of the filter on both graphs; fatal casualty with pedestrian as casualty class (top image) is compared to fatal accidents with drivers as casualty class (bottom image)

4. Assessment

In this paragraph we focused our attention on the evaluation of the infographics previously presented. There are different ways to evaluate an infographic: qualitative-quantitative, absolute-comparative, and formative-summative.

In our case, we conducted the qualitative-quantitative evaluation, which is characterized by two different kinds of methods: the heuristic evaluation, which aims to highlight the problems of usability and readability of the infographic with a psychometric questionnaire, which is used to evaluate some quality dimensions of the interaction of the users; the user test, is the other method which has the goal of observing user interactions with the infographic through specially designed tasks.

4.1 Users

The infographic must be consistent with the user's expectations, it has to put him at ease with simple but also complete elements. It must not contain elements of confusion or that lead to some kinds of mistakes. It also aims to lead the user to ask himself questions and to solve them through it.

We decided to involve for our evaluation of the infographic a number of 24 users. All the users will get involved in several activities as later described.

4.2 Heuristic evaluation

The heuristic evaluation is the process in which the users evaluate the infographic and respond to specific created tasks to identify problems and doubts about it. In this way that allowed us to identify the problems related to our data visualization that need to be settled out in the final infographic.

The heuristic evaluation that we have created consists of five different tasks that cover all the different macro areas of the project. Users will be asked to answer simple questions regarding the infographics but at the same time also to speak and think aloud in order to express doubts and questions, allowing us to understand even the smallest problems related to the infographic in question. During this phase all the errors made by the same users in responding to the tasks will therefore be noted and the execution time for each task will be taken into account.

4.2.1 Our tasks

Five tasks have been identified to carry out the heuristic analysis. It was asked to our users to answer in detail about the following questions:

Task-1

Take into consideration the graph on page 1 (Understanding road accidents in Great Britain), showing the number of road accidents in the Uk between 2005-2015. In which quarter was there the largest percentage change in the number of accidents compared to the previous year? Also try to report the numerical value.

Task-2

Take into consideration the graph on page 2 (Distribution of Road Accidents over Month, Day of the Week, and Hour), on what day of the week and at what time did the most serious accidents occur in the period 2005-2015?

Task-3

Take into consideration the graph on page 4 (The Role of Gender and Age in Road Accidents), considering the class casualty “pedestrian”, how many women aged 16-20 have been involved in serious accidents?

Task-4

Take into consideration the graph on page 5 (The Role of Vehicle Type and Gender in Road Casualties), for which type of vehicle (other than cars) have the most women been involved in serious accidents?

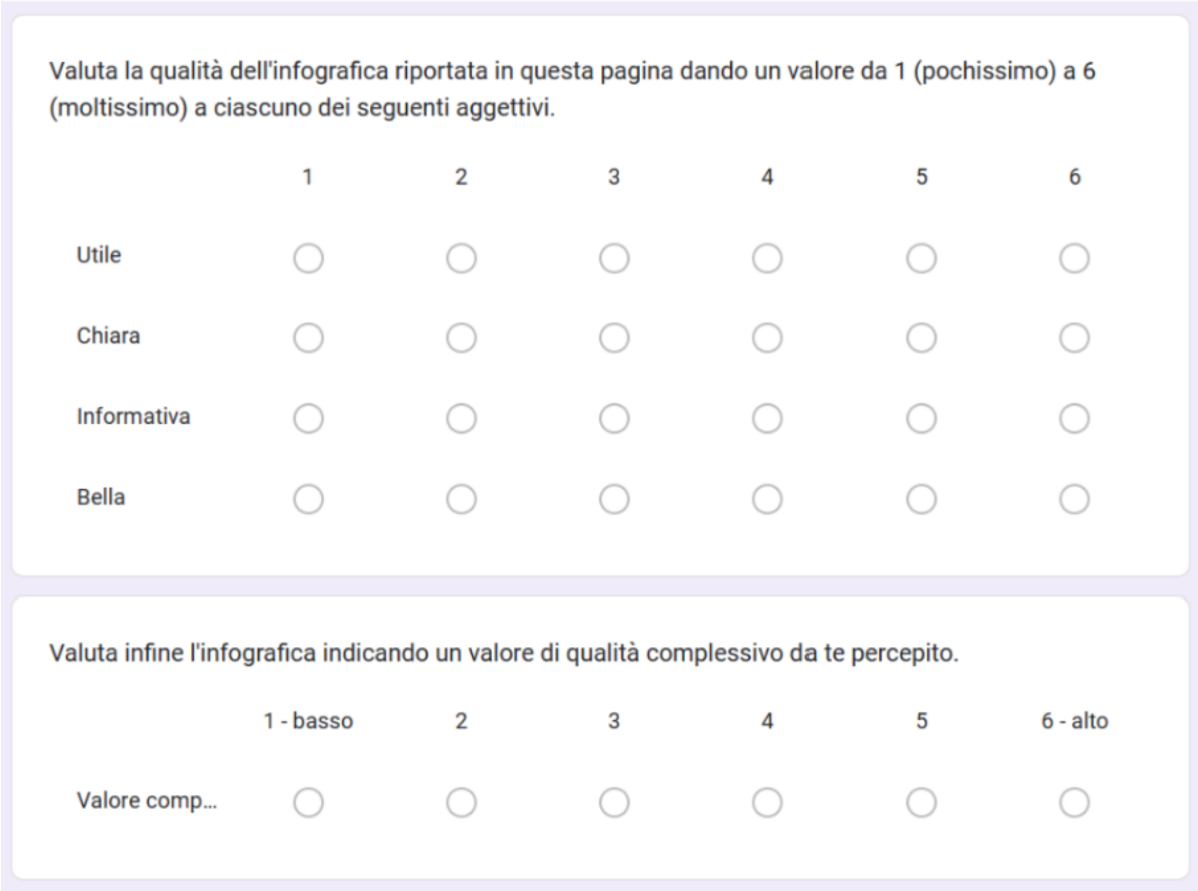
Task-5

Take into consideration the graph on page 6 (Insight into the Age Group of Casualties), what day and at what time are the 'Over 65' age group least likely to be involved in a fatal accident?

4.3 Psychometric questionnaire

For the realization of the psychometric questionnaire we decided to adopt the Cabitza-Locoro scale; Figure 4.1 shows an example of a questionnaire created in this way. This scale allows the assessment of the quality of the infographic on a scale of 1 to 6 for the following five fields:

- Utility
- Clarity
- Informativeness
- Beauty
- Overall score



Valuta la qualità dell'infografica riportata in questa pagina dando un valore da 1 (pochissimo) a 6 (moltissimo) a ciascuno dei seguenti aggettivi.

	1	2	3	4	5	6
Utile	○	○	○	○	○	○
Chiara	○	○	○	○	○	○
Informativa	○	○	○	○	○	○
Bella	○	○	○	○	○	○

Valuta infine l'infografica indicando un valore di qualità complessivo da te percepito.

	1 - basso	2	3	4	5	6 - alto
Valore comp...	○	○	○	○	○	○

Figure 4.1 Psychometric questionnaire by Cabitza-Locoro

4.4 Results

The results obtained from the evaluations described above will now be reported and analyzed.

4.4.1 Heuristic test

The purpose of this test is to evaluate the work that we have done with a "think aloud" protocol, taking into account all the problems encountered by the users throughout the navigation, so that we have the possibility to figure out how to improve our work.

We involved 3 different users, who found themselves to deal with these three main problems:

USER 1	During the interaction of dashboard 4 it emerged that the distributions of casualty severity by gender were not easy to understand through filters alone.

USER 2	Does not understand the possibility of being able to click on the filters and select from the available attributes, those that are more interested in exploring.
USER 3	Regarding the dashboard on pages 2 and 3, the user stated that he did not notice a clear difference between the colors referring to the scale.

After having collected and analyzed all the problems encountered, we made some little changes to the data visualization.

1. To remedy this misunderstanding, we have created pie charts that provide an overview of the gender percentage by severity of the accident, both for the driver and for the injured person.

2. We have added above each filter a short description that suggests that the filters are interactive.

3. We have changed the reference color scale, using white to indicate the average, red to indicate values above the average and green to indicate values below the average. Also, we have defined a black grid to divide each cell of the heat maps.

4.4.2 Heuristic evaluation

We decided to dispense the heuristic evaluation we have already presented to six people, who had to carry out the five tasks we provided. For each user, we asked them to provide us with the following personal data: gender, age (to select from age groups), level of familiarity with the topic (to choose between three options; beginner, average expert and expert) and education level (higher diploma, bachelor degree, master degree).

ID	SEX	EXPERIENCE LEVEL	AGE	EDUCATION LEVEL
USER 1	F	Average expert	18-25	Bachelor degree
USER 2	F	Average expert	18-25	Bachelor degree
USER 3	M	Average expert	18-25	Higher diploma
USER 4	F	Beginner	18-25	Higher diploma
USER 5	F	Expert	25-34	Master degree
USER 6	F	Average expert	18-25	Bachelor degree

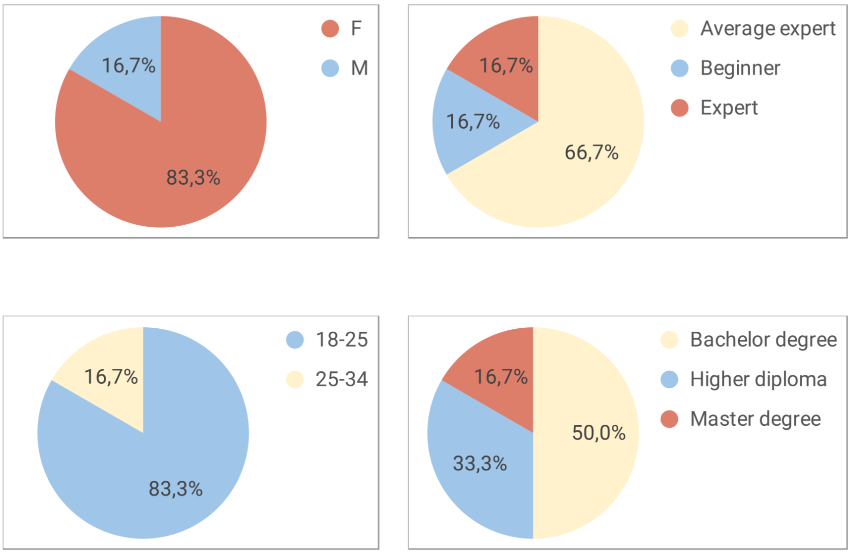


Figure 4.2 Pie charts about our users' general information

We now report the data collected regarding the time taken to solve each task and any mistakes made.

TASK	ID	EXECUTION TIME	EXECUTION MODALITY
TASK 1	USER 1	203s	C-A
	USER 2	105s	C-S
	USER 3	61s	C-S
	USER 4	67s	C-S
	USER 5	87s	C-S
	USER 6	150s	C-S
TASK 2	USER 1	112s	C-S
	USER 2	110s	C-S
	USER 3	54s	C-S
	USER 4	90s	C-A
	USER 5	88s	C-S
	USER 6	148s	C-S
TASK 3	USER 1	90s	C-S
	USER 2	58s	C-S
	USER 3	50s	C-S
	USER 4	68s	C-S
	USER 5	77s	C-S
	USER 6	134s	C-S
TASK 4	USER 1	110s	C-S
	USER 2	96s	C-S
	USER 3	38s	C-S
	USER 4	31s	C-S
	USER 5	73s	C-S
	USER 6	165s	C-S
TASK 5	USER 1	126s	C-A
	USER 2	115s	C-A
	USER 3	47s	C-S
	USER 4	30s	C-S
	USER 5	81s	C-S
	USER 6	147s	C-A

For what concerns the execution modality, we used C that stands for “corretta” (correct) and E for “errata” (wrong), S stands for “solitaria” (alone) and A for “assistita” (assisted).

The task resolutions registered as “assisted” are the ones characterized by any kind of help or explanation. As it can be seen from the table, no users made errors in reading the graphs. However, in 3 different tasks, users were assisted in their response. While in tasks 3 and 5 no request for help has even been made.

TASK 1	Specifically in the first task, the user stated that he was confused in viewing the negative columns. The interviewer explained why.
TASK 2	In the second task, a user had difficulty in locating the filter. In this case the interviewer helped him to find it.
TASK 3	In the fifth task, the error of the filter is confirmed, specifically one user states that he would have placed the filter closer to the graph because this way he does not have the possibility to visualize what it refers to.

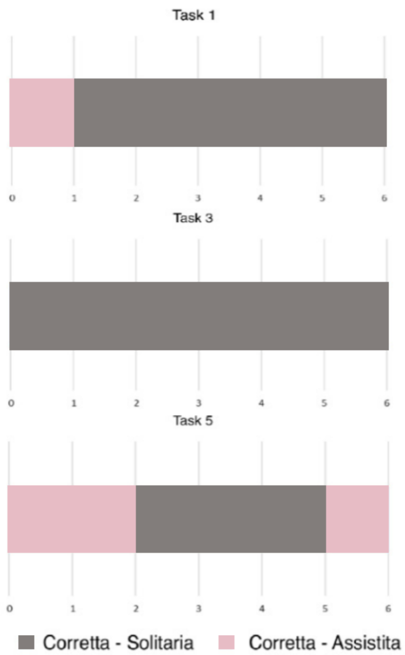


Figure 4.3 User’s performances

After the analysis of the user’s performances, we proceeded by taking in consideration the execution times for each task. To do so, we created a violin plot (Figure 4.4), aiming to immediately clear the variation of the execution times in carrying out the tasks.

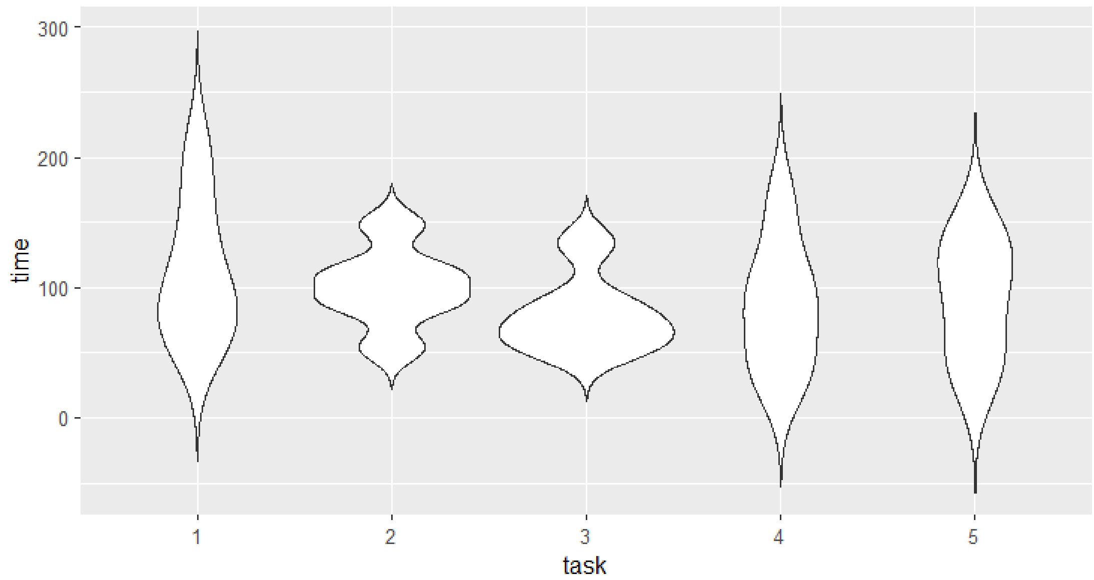


Figure 4.4 Violin plot of the time of execution of the proposed tasks

Looking at the violin plot it is possible to notice how task 1 and task 5 were characterized by a longer execution time, perhaps because both responses were assisted. The answer two and three, are instead those with less execution time (even though one user has been assisted in the second task).

4.4.3 Psychometric questionnaire

Finally, we provided the psychometric questionnaire with the Cabitza-Locoro scale to 24 users. We now report with the bar charts below (Figure 4.4) the analysis regarding the data we collected.

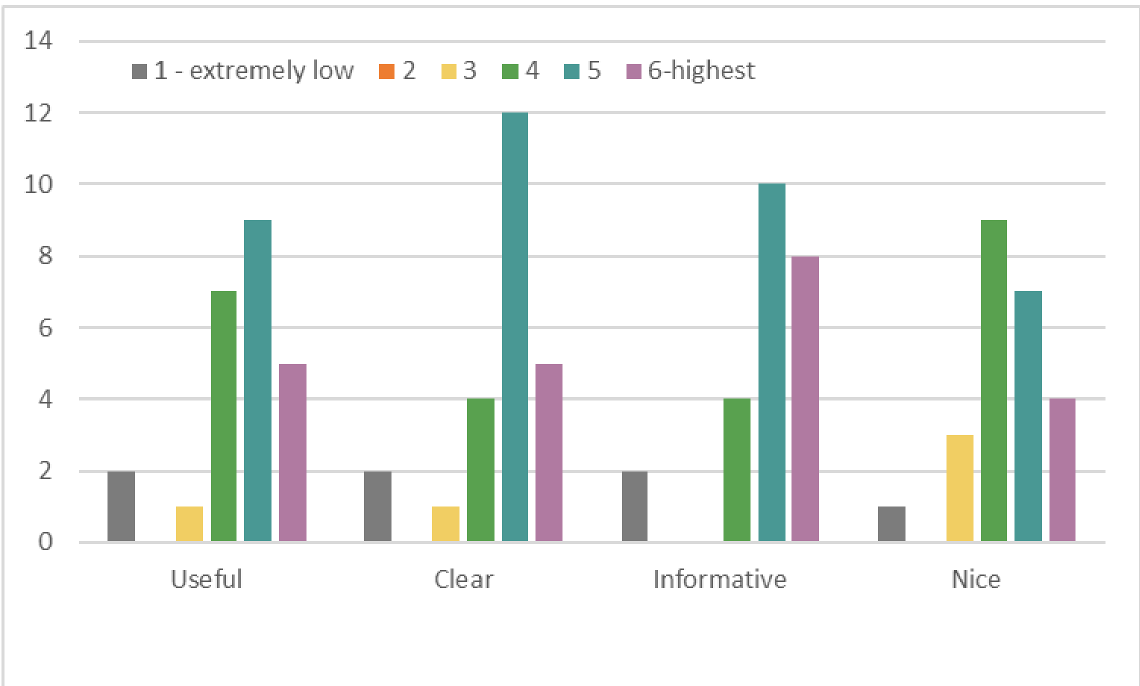


Figure 4.4 Results of the questionnaire visualized through bar chats

Regarding the results of our questionnaire, we have received quite positive feedback. Starting from the graphs we can say that as far as usability is concerned, our users evaluated it positively, the most chosen value was 5, as many as 9 users out of 24 chose it (38%). As far as clarity is concerned, here too our users judged it positively, attributing it a value of 5, as many as 12 out of 24 users (50%) chose this value. While as far as information is concerned, users have mainly chosen the values 5 and 6, the sum of these two values constitutes almost all 18 users out of 24 (75%) have chosen one of these two values. Finally, in regards on how interesting the dashboard was evaluated rather positively, the most chosen values was 4.

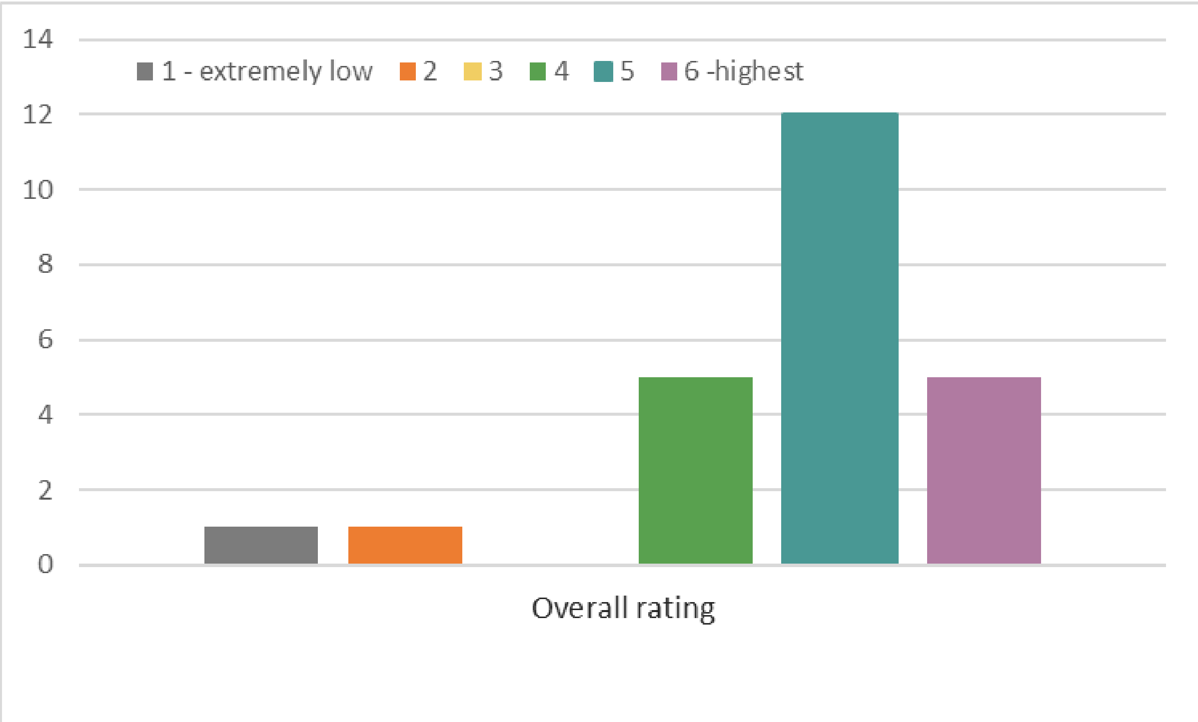


Figure 4.5 The overall rating using the Cabitza-Locoro scale given by the 24 users

As we can see the infographic has been appreciated by most users, with the value of 5 for overall rating, 12 users on 24 total the 50%, for us is a good result, instead we believe that are some aspects to fix and improve.

5. Conclusion

The theme covered in our work has raised a lot of positive feedback, as reflected in the results of the questionnaire submitted. Despite the long execution time, we have an interesting high rate of correct responses. Particularly we are glad that we have received a positive response about clarity and informative aspects of our dashboards.

Through the interactive filters, we have tried to give as much flexibility as possible to the viewer to explore the dashboards and then attempt to input some information about their personal profiling in order to investigate under what conditions they might face an higher risk of being involved in a traffic accident. In this sense, we designed the heuristic test tasks to not be immediate, but to stimulate the user to challenge himself and more carefully read the chart. This is a very recurring theme in newspapers and TV-news, that has raised a lot of questions and concerns. Hence, from the starting point, our goal through the analysis of road accidents in Great Britain, was to give more awareness of accident dynamics, identify potential trends and in our limited capacity, give some pin-points to the readers. We would have liked to analyze it over a longer time period and have more data on the background of the casualties in an effort to do more in-depth analysis.

We can consider ourselves very satisfied with our work, since we are a group with people from different backgrounds, and none of us have ever had the chance to work with Tableau before. We believe that Tableau is a very helpful tool to represent data in a graphical way, making it easier to understand the topics. We also think that tool could be very useful throughout our future working career. In the end, we started learning how to choose the best graph in all circumstances, aware that the first idea is not always the best to represent our information.