

# Education and Traffic fatalities, what can we expect from “more educated” states when it comes to driving?

## Introduction

The following visualisation project is based in a Shiny App for interactive visualizations with R using the Shiny Library.

This visualisation project will aim to give a general understanding of the current situation of the traffic fatalities with a focus in three main aspects:

- The most dangerous highways across the US with the specific location that present the highest records of traffic fatalities.
- The relationship between education and traffic fatalities per state (focus on TX, CA, FL)
- The type of intersection and hour of the day that is more dangerous across the US.

## User Guide

The following User Guide provides the instructions for exploring the narrative visualisation.

### 1. Visualisation Name

*Traffic Fatalities in the US 2015.*

### 2. Intended Use

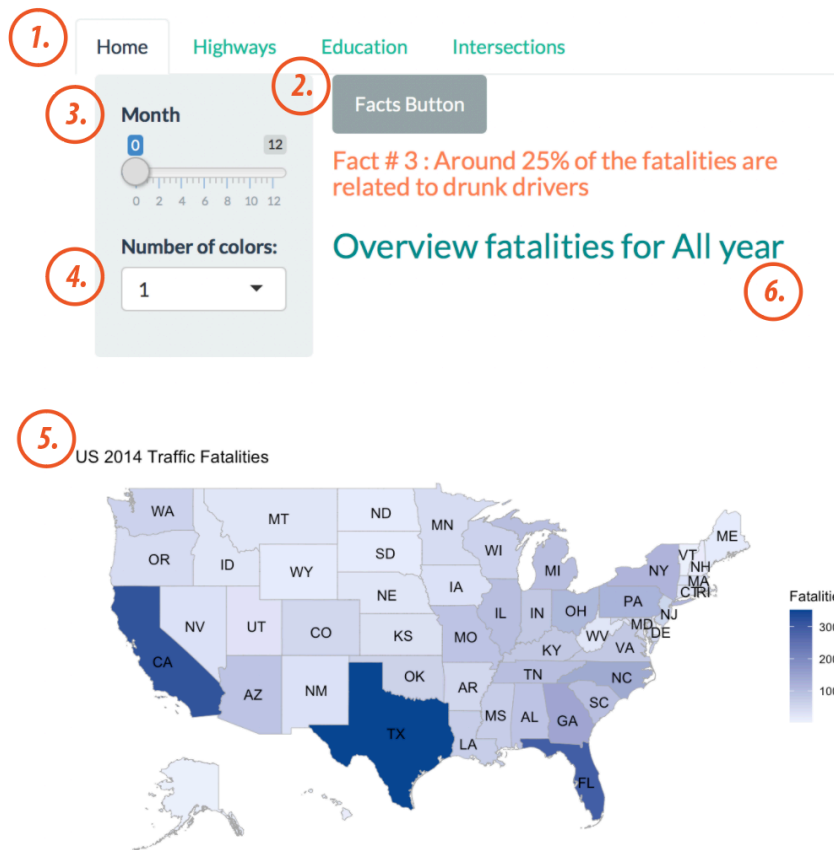
The visualisation described in the following user guide is oriented towards the government and aims to provide a big picture of the traffic fatalities in the US for 2015. The main purpose of the visualisation is to exhibit which states are the most problematic in the rates of driving fatalities and supply additional information to implement regulations and policies that allow to improve the safety in the most dangerous spots for these states.

### 3. UI Description

The *Traffic Fatalities in the US 2015* visualisation is divided in 4 sections as follows:

- a. **Home Section:** In this tab the user will find a choropleth map of the US that provides a general overview of the record of traffic fatalities for 2015. This page contains 6 points that are explain below.

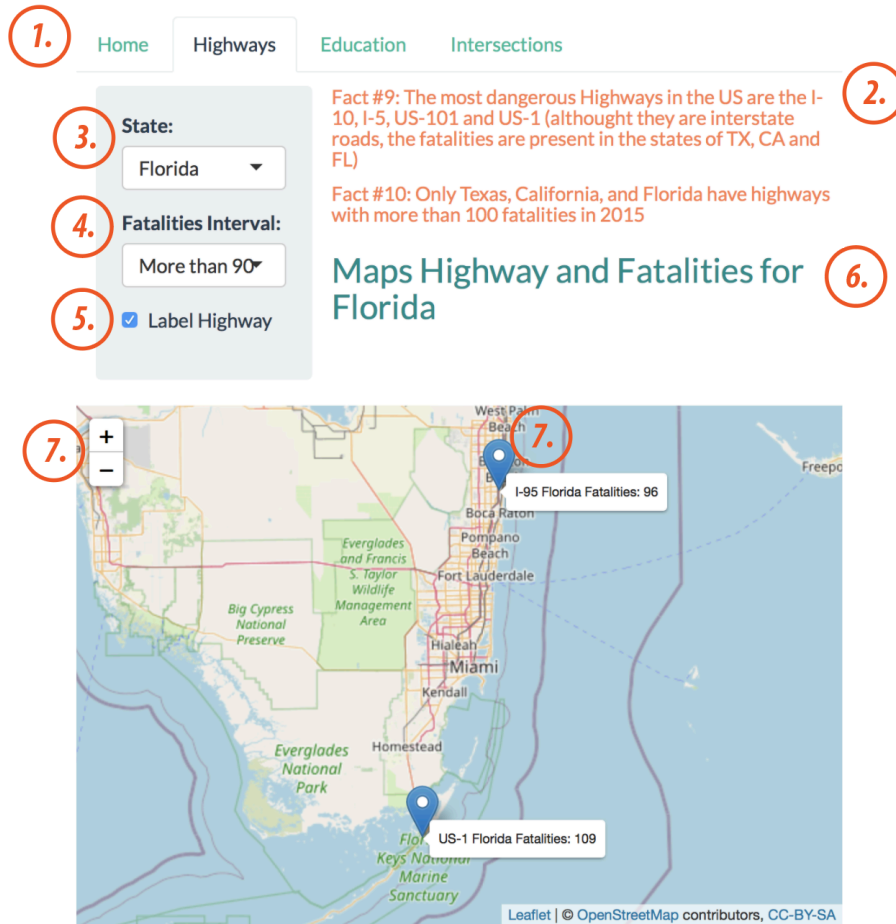
## Traffic Fatalities in the US 2015



1. Tabset panel with access to other tabs for different levels of detail
2. Facts Button: Provides different facts in order to engage the user with some of the findings of the study.
3. Slider bar to filter by the month or whole year by selecting the option 0
4. Number of colors for the choropleth map. This gives the user the flexibility to see several scales in order to group more states under different ranges for the amount of driving fatalities
5. The Choropleth map adapts to the selection of the user in the steps mentioned above.
6. Interactive title for the section showing the selected month or year.

b. **Highways Section:** In this tab the user will find all the information related to the specific highways that were identified as the most dangerous in the US. This interface provides the user with several interactive and informative components that are described below the image.

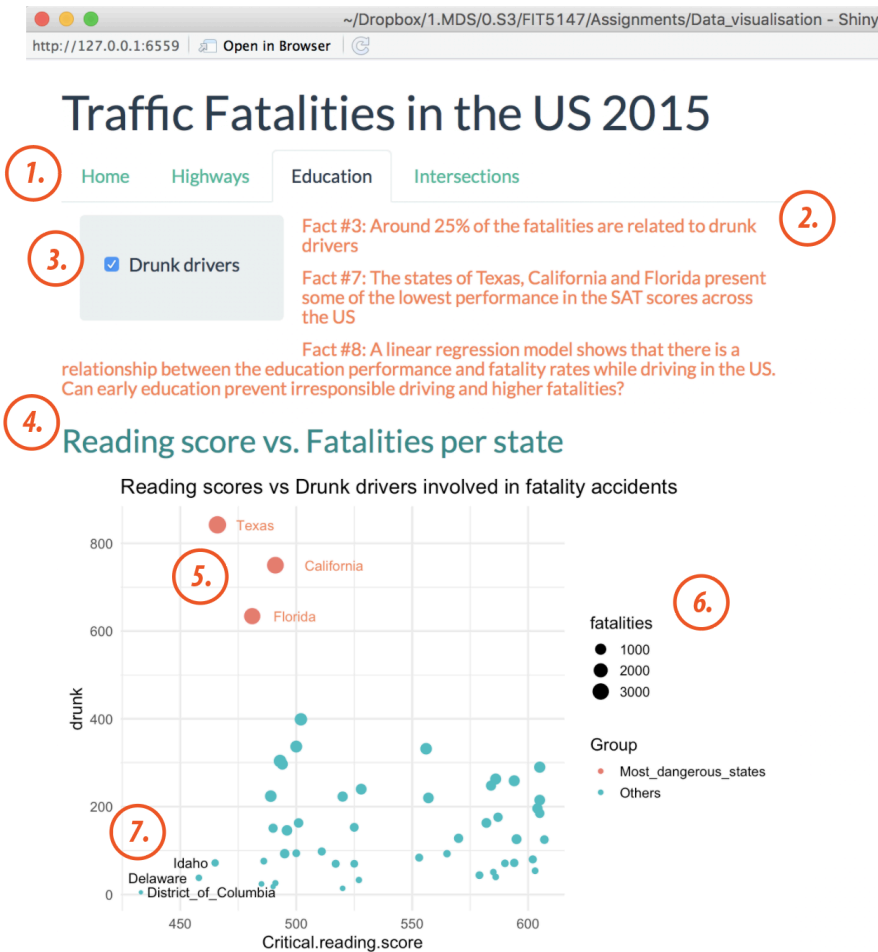
# Traffic Fatalities in the US 2015



1. Tabset panel again to navigate across the different sections
2. List of the related facts for the Highways section
3. Dropdown menu for filtering the state (or all country)
4. Dropdown menu for selecting different intervals in order to provide more or less detail on the number of fatalities per road
5. Checkbox to display the labels of the markers in the map with additional information
6. Interactive title updated according to the selected state or whole country
7. Leaflet OpenStreetMap with zoom buttons and hover option to show tags of the markers with details

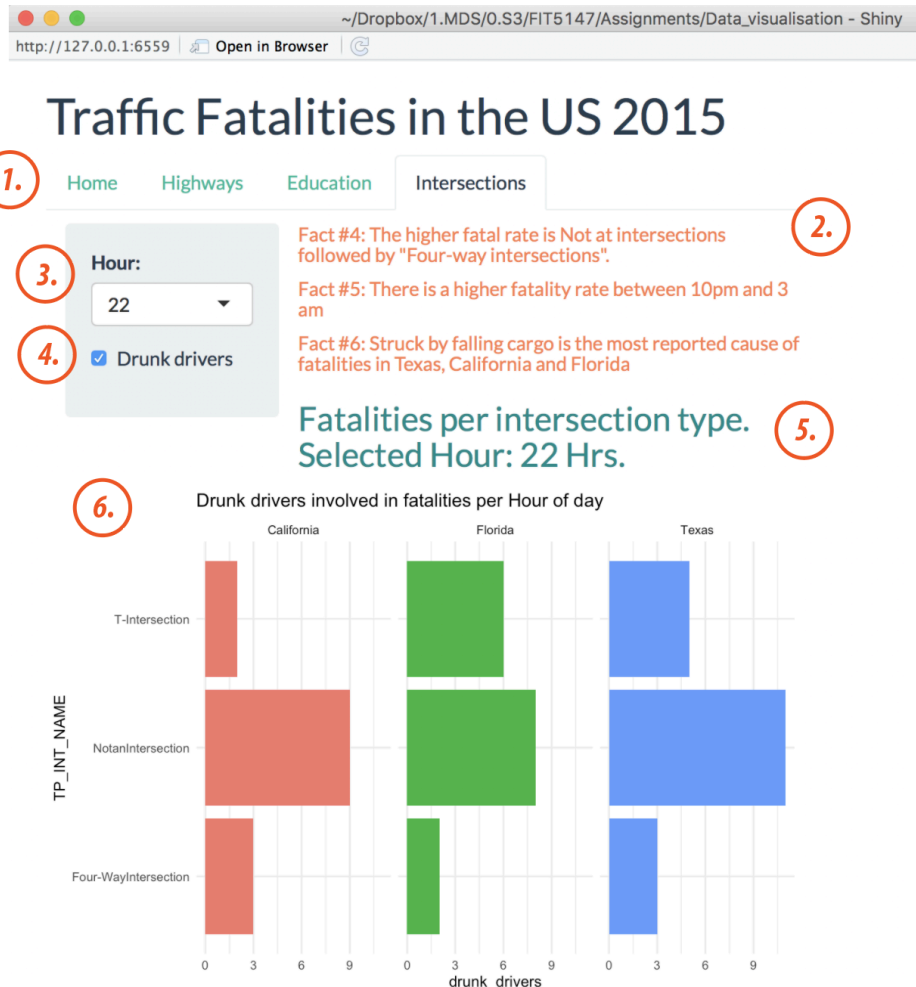
- c. **Education Section:** This tab provides the main insights of the relationship between education and traffic fatalities. The main purpose of this tab is to show to the user how the states of Texas, California and Florida appear to be '*clustered*' in a problematic zone of the graph where the academic performance is low and the drunk driver related

fatalities are very high. The main components of the user interface are numerated in the image below and further explained after the image.



1. Tabset panel again to navigate across the different sections
2. List of the related facts for the Highways section
3. Checkbox to include the Drunk drivers effect in the visualisation
  - a. Without check: shows the reading score vs. fatalities
  - b. With check: drunk effect is included as y-label and fatalities now becomes an extra dimension (size of the points). Also the labels for states in risk because of low academic performance are displayed
4. Title of the section
5. Cluster of the three most dangerous states with their corresponding label and different colour
6. Legend with the information for colours and and size
7. As mentioned in 3.b, the label for the states at risk for their low academic performance.

- d. **Intersections Section:** The final tab of the visualisation provides the user with specific information related to the intersection type where the fatalities took place and is grouped for the mentioned states (TX, CA, FL). In a similar manner, the following image shows the components of the tab and the explanation is provided below.



1. Tabset panel to navigate across the different sections
2. List of the related facts for the Highways section
3. Dropdown menu for filtering by hour (or full day)
4. Checkbox to display the effect of Drunk drivers only
5. Interactive title updated according to the selected hour of the day
6. Barchart displaying the total fatalities (or drunk related fatalities) per intersection type for each of the 3 problematic states

## References

- Wagner, I. (2020). Licensed drivers in the U.S. - total number by state 2018. Retrieved from <https://www.statista.com/statistics/198029/total-number-of-us-licensed-drivers-by-state/>
- Golden Oak Research Group LLC, "U.S. Income Database Kaggle". Publication: 5, August 2017. Accessed, 01, April 2020
- Panagiotis D. (2020). Five Design Sheets - A framework for prototyping information visualization interfaces through sketching. Retrieved from <http://pdritsos.com/projects/FDS/>
- The Data visualisation Catalogue. (2020). Retrieved from <https://datavizcatalogue.com/>

## Appendix - Description of data sources

1. Tabular data: 32K rows x 52 columns with spatial and temporal attributes. Several attributes contain codes that need to be searched in the user's manual. (<https://data.world/transportation/2015-traffic-fatalities>)
2. Tabular data in xls with info for each state in 56 rows and 19 columns. The tables contain multiple line heading so a pre-processing task for reshaping must be done. The data contains some additional text at the end. ([https://nces.ed.gov/programs/digest/d16/tables/dt16\\_226.40.asp](https://nces.ed.gov/programs/digest/d16/tables/dt16_226.40.asp))
3. Tabular data with 32K rows and 19 columns. Geo locations, spatial data. ([https://www.kaggle.com/goldenoakresearch/us-household-income-stats-geo-locations#kaggle\\_income.csv](https://www.kaggle.com/goldenoakresearch/us-household-income-stats-geo-locations#kaggle_income.csv))