

Photo by Andrew "Donovan" Valdivia on Unsplash

Design of an interactive data analysis tool

Visualization Practical Work

Universidad Politécnica de Madrid | Big Data - Data Visualization | January 2021

Camarena Torres, Lucas

Marín Estañ, Ana

Rodrigo Lacave, Blanca

# Content

[Problem Characterization in the Application Comain 1](#_Toc62149766)

[Data and Task Abstractions 1](#_Toc62149767)

[Data Abstractions 1](#_Toc62149768)

[Task Abstractions 1](#_Toc62149769)

[Interaction and visual encoding 1](#_Toc62149770)

[Algorithmic implementation 1](#_Toc62149771)

[Results 1](#_Toc62149772)

[Conclusion 1](#_Toc62149773)

[References 2](#_Toc62149774)

[Appendix 1. How to run the code 3](#_Toc62149775)

# Problem Characterization in the Application Comain

Violent encounters with police represent a significant cause of morbidity and mortality in the USA [1]. And, over the past five years there has been no reduction in the racial disparity in fatal police shooting victims despite increased use of body cameras and closer media scrutiny, according to a new report by researchers at Yale and the University of Pennsylvania [2].

In 2015, *The Washington Post* began to log every fatal shooting by an on-duty police officer in the United States. In five years, more than 5,000 thousand shootings recorded [3]. The data used in this project was obtained from Kaggle [4]. However, the selected dataset is a preprocessed version to the *The Washington Post*'s dataset [3].

# Data and Task Abstractions

## Data Abstractions

## Task Abstractions

For each visualization that is presented in the application, it is necessary to make a description of the task abstraction, including its actions and targets.

### Geographical exploration

For these three visualizations, new information is produced: shooting frequencies, grouped by states or cities are obtained.

#### Shooting city distribution

In which urban areas are police shootings more frequent?

**Actions**: Besides the shooting frequency for the cities, new information is produced: the cities' location (longitude and altitude) are added. The aim of this question is to explore identify US urban areas in which police shootings are more frequent

**Targets**: these actions are applied over the -geographical- distribution of the shooting frequency.

#### Victim's profile

Does the state in which the shooting occurred influence the profile of the victim?

**Actions**: This task involves the shooting frequency in each state as well as the victims' demographic features. With these two, users can explore the proportion of victims of a particular profile across the states. Therefore, new knowledge is produced.

**Targets**: these actions are applied over the -geographical- distribution of the shooting frequency as well as the attributes corresponding to the victims' demographics.

#### Seasonality

Is there a seasonal pattern in shooting across states?

**Actions**: The aim of this task is to discover any seasonal patterns depending on the state or area (e.g. Are there more shooting in coastal states in summer?). From the shooting date, new information is created: the season, the quarter and the day of the week.

**Targets**: these actions are applied over the -geographical- distribution of the shooting frequency as well as the shooting date.

# Interaction and visual encoding

### Geographical exploration

#### Shooting city distribution

In which urban areas are police shootings more frequent?

**Visual encoding**:

For this idiom, the data used are an ordinal, quantitative attribute and geometry data (a map). These data are represented through a bubble map where the position of each bubble is the location of the city, and its diameter corresponds to the number of shootings.

**Interaction:**

The users can filter the year interval of the shootings that is represented. When the user hovers a bubble, the city name, state and number of shootings can be consulted. Moreover the map can be zoomed in an out, and can be downloaded as PNG.

#### Victim's profile

Does the state in which the shooting occurred influence the profile of the victim?

**Visual encoding**:

This idiom also involves several attributes regarding the victim's demographics (race and gender, categorical; age, quantitative) used to filter the number of shootings. It also involves geometry data. In this case, state location.

Therefore, the selected representation is a choropleth, where the color intensity that fills each state corresponds to the relative frequency of shootings with that victim profile:

**Interaction:**

This idiom allows users to filter the victims' demographic features (gender, race, age range) to filter the results.

#### Seasonality

Is there a seasonal pattern in shooting across states?

**Visual encoding**:

Like in the previous idiom, the different -ordinal, cyclic- attributes related to time (created from the date) are used to filter the number of shootings. And again, geometry data (state location) is involved.

The selected representation for this is a choropleth, where the color intensity that fills each state corresponds to the number of shootings that occurred at that period of time.

**Interaction:**

User interaction is given by the filter options, which in this case are nested: the selected seasonality will lead to a particular filtering option (e.g. if "Month" is selected for Seasonality, the filtering options will be "January, "February"...).

# Algorithmic implementation

For the creation of new information (like season attributes and locations) *lubridate* and *maps* packages were required. Moreover, some were needed for data wrangling libraries, like *tidyverse* or *plyr*.

Regarding the geographic exploration idiom implementation, the packages *choroplethr* and *plotly* are used.

# Results

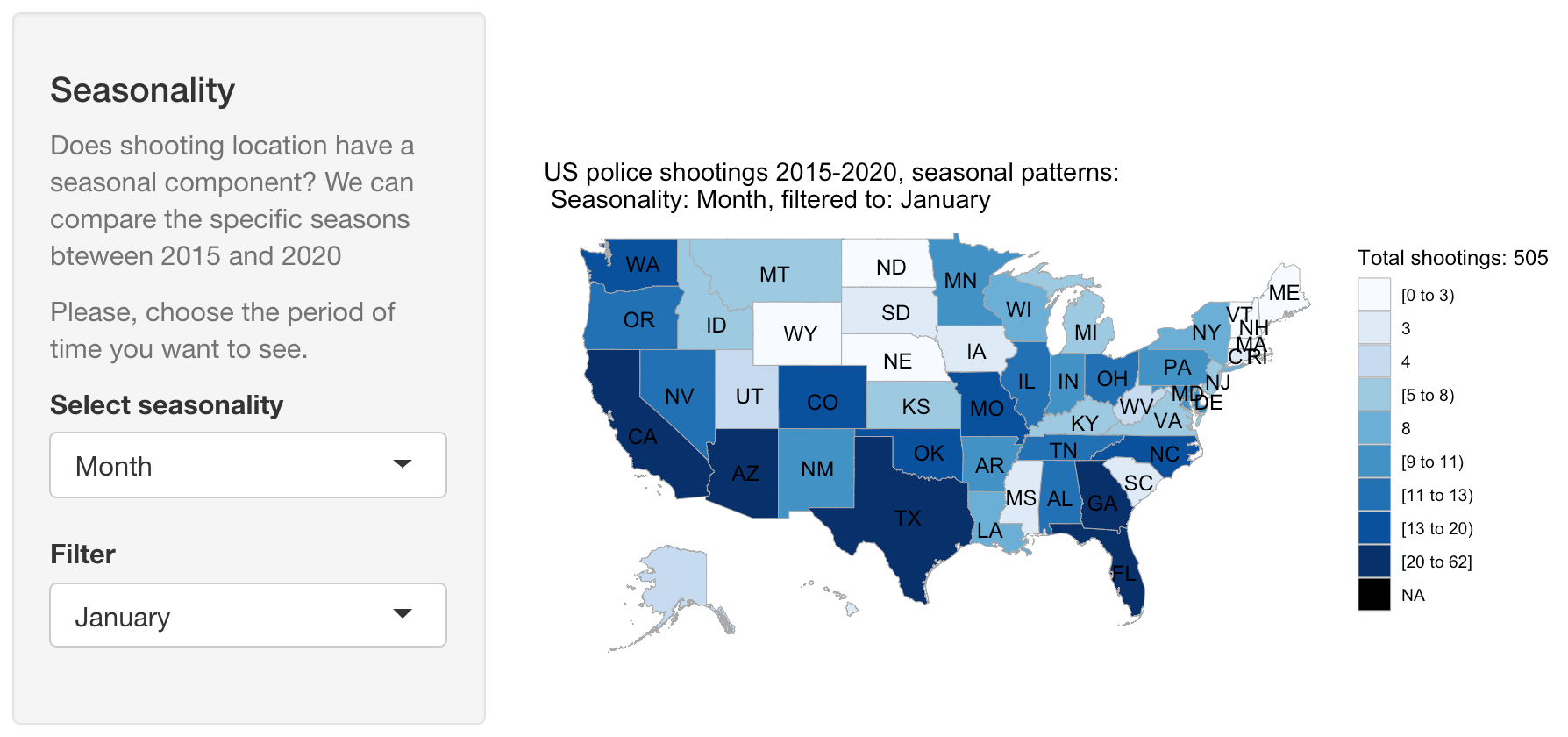
### Geographical exploration

#### Shooting city distribution

Gráfico

Descripción generada automáticamente

#### Victim's profile



#### Seasonality

Diagrama

Descripción generada automáticamente con confianza baja

# Conclusion

# References

[1] F. Edwards, H. Lee, and M. Esposito, “Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 34, pp. 16793–16798, Aug. 2019.

[2] E. Lett, E. N. Asabor, T. Corbin, and D. Boatright, “Racial inequity in fatal US police shootings, 2015-2020,” *J. Epidemiol. Community Health*, Oct. 2020.

[3] “Police shootings database 2015-2021 - Washington Post.” [Online]. Available: https://www.washingtonpost.com/graphics/investigations/police-shootings-database/. [Accessed: 21-Jan-2021].

[4] “US Police Shootings | Kaggle.” [Online]. Available: https://www.kaggle.com/ahsen1330/us-police-shootings/notebooks. [Accessed: 21-Jan-2021].

# Appendix 1. How to run the code