

Data Visualization of Missing and Unidentified Persons in the United States

Bandr AlSwyan

Ricky Severino

American University

4400 Massachusetts Ave, NW

CSC-493-001

September 15, 2018

Setting/Context

The environment in which our capstone will run on is a browser based application that the user will be able to interact with.

Capstone Problem and Rationale

The purpose of this capstone project is to give the general public the information and tool to identify and potentially assist locating missing individuals in the United States. The need for this technology arise in its simplicity as the tools that are currently available are either complex or not modern enough to provide the complete potential and use to reach different people. The use of a simple tool that does not require prior knowledge is the goal that we want to achieve. We will apply and support our data with statistical analyses that will reflect better understanding of the data collected.

The data that will be integrated into the application will be retrieved from the National Institute of Justice's National Missing and Unidentified Persons System (NamUs) website and categorized as public information. According to the privacy agreement that is in correspondence with the Department of Justice (DOJ), third party applications can use the data to help professionals reach the purpose in which this data was created and collected, and therefore meeting the demand of our project, giving us the permission to use the data.

The application will consist of an HTML, CSS, and Javascript website, and the data will be represented through R-Studio. A database will be used for the retrieval of information using Structured Query Language (SQL). The website provided will display missing persons, represented by a point on the map. When the user clicks on this point, information regarding the individuals will popup to the side of the page. In this popup tab, the user will be able to see information regarding name, gender, ethnicity, age, etc. On another side of the page, a search filter will be provided so the user can make specific queries about the information displayed. Among other options, the search filter will have the option to display either points representing people, or a heat map representing the density of missing persons in that area. The page will also provide a link to another page which will show statistics about the data mentioned above. These statistics will show a plethora of correlations and relationships between different variables provided by the data source.

Objectives

The objective of this project is to provide the general public with better understanding of the missing people in the United States in an interactive and simple manner. The response to the needs mentioned above is to design a web tool that can be easily accessed and used through an interface that the user can draw conclusions with and from.

Deliverables

- Website
 - Data mapping and representation
 - Points and heatmap
 - Queries: allowing users to filter the data shown
 - Statistical analyses (creating new data)
 - Resources to authorities

Project Description

Timeline:

1. Finding and getting to know the data
 - a. Exploring the data to know what we will work with and to check if it's legal/free to use
 - b. Figuring out the format of the data and how it will be collected
2. Collecting data
 - a. Figuring out how to collect and input the data into our application
 - b. Group the data into workable format
3. Processing data
 - a. Converting and organizing data so that it meets the requirement of the tools that we will work with
 - b. Outputting new information driven from the collected data through statistical transformation
4. Representing data
 - a. Create an interactive template of the website that will show the data
 - b. Display the outputted statistical data
5. Enhancement and optimization of the application
 - a. Putting the website into function and creating the necessary direction for the user
 - b. Debugging
6. Presentation

Resources

Hardware	Computers
Software	Browsers, RStudio, HTML/CSS, Javascript, MapBox GL API, MySQL
Data Formats	Json, GeoJson, CSV, SQL, R
Thoughts	Hope and Perseverance