- **Basic Info**

  Title : US Mass Shootings (1966-2017)

  Team :

  Junior Rojas u1114542 u1114542@utah.edu

  Saman Sepehri Nejad u1113773 u1113773@utah.edu

  Github : https://github.com/samansepehri/US-Mass-Shootings

- **Background and Motivation**

In last 50 years, the United States has had over 300 mass shootings that resulted in over 1900 deaths and over 2400 injured. The death tolls and places are different in every case, and the most recent mass shooting in Las Vegas had more than 500 victims, a very tragic incident with an uncommonly high number of victims. We believe that knowing more about the details of these incidents is of interest to many people like ourselves, who are currently living in the United States.

- **Project Objectives**

We would like to get some insights about common patterns in past US mass shootings. In particular, some interesting questions we would like to answer and some quantities and patterns we would like to show include:

- Number of people killed and injured per year
- Visualize the location of the incidents on a US map
- Are these incidents more common in some states than others?
- Is there a correlation between number of people killed and injured in an incident?
- Are these incidents more common in certain months or days of the week?

- **Data**

Our main data source is www.kaggle.com, which provides a CSV file with information of US mass shootings since 1966.
URL: https://www.kaggle.com/zusmani/us-mass-shootings-last-50-years/data

- **Data Processing**

The data requires some processing to be used in our visualization. There is no “state” column in the data set, but it can be extracted from the “location”. The “date” column is a simple string, which is not appropriate for certain comparisons and analyses we want to do. We will extract the year, month and date from this column. Since we also want to analyze patterns about days of the week, we will need to extend the data with the corresponding day of the week in which each incident happened. We also need certain aggregate quantities, such as number of victims per state and total number of

victims per year, which are not directly included in the data set. We will either pre-compute some of these values or compute them on the fly with JavaScript.

- **Visualization Design**

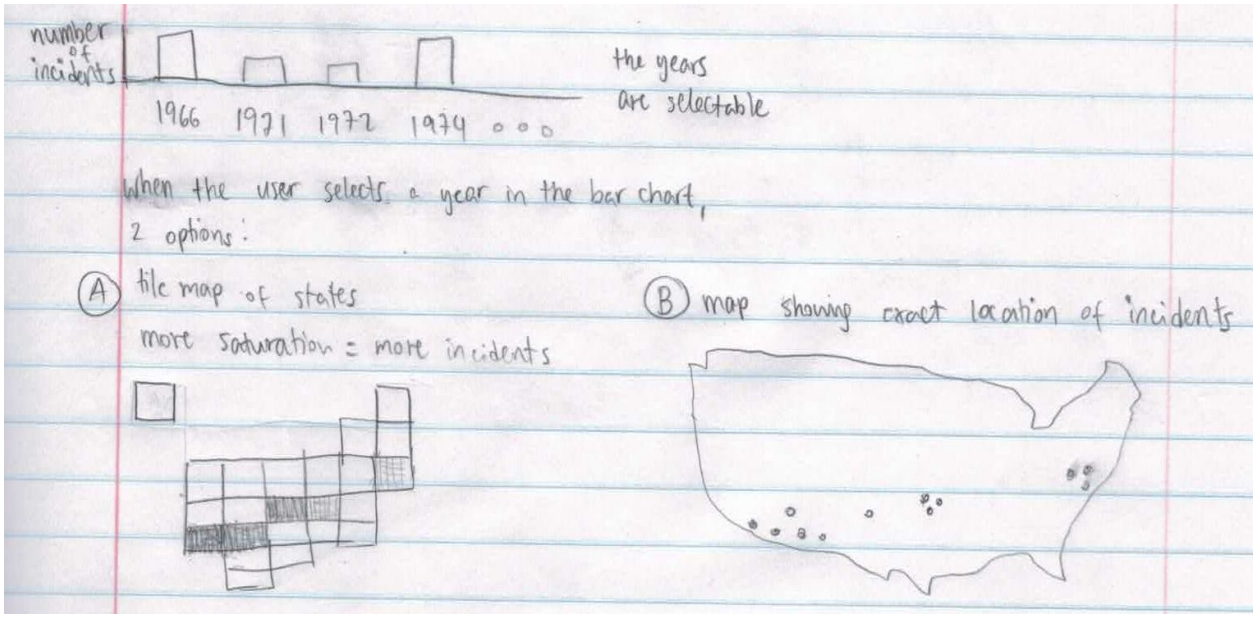
We guide our design by the general rule “overview first, zoom and filter, then details on demand”. The following sections describe some proposals to achieve an effective visualization following this idea.

Overview by year and location

We propose that the initial state of the visualization show a bar chart with the total number of incidents per year. The years are selectable. When the user selects a year from the bar chart, the visualization is updated with a map that shows more information about the selected year. Allowing the user to select more than one year might also be a useful addition.

We propose two options for the map:

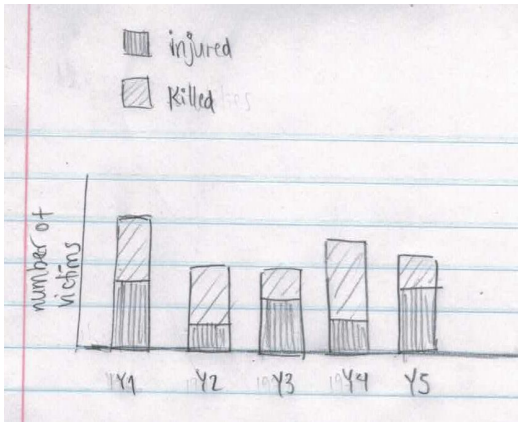
A. A tile map that shows each state and uses saturation as the channel to indicate the number of incidents.
B. A more traditional US map that shows the exact location of the incidents.



We believe both map options are valuable in different cases. Map A is useful for a quick overview of the number of incidents per state. Map B shows more detailed information of the locations and allows nice zoom and filtering, which is especially useful if most incidents happened in the same state. One important thing to note is that some incidents in our data set

do not have exact location information available, so for some years in which exact location information is scarce, Map A might be more appropriate.

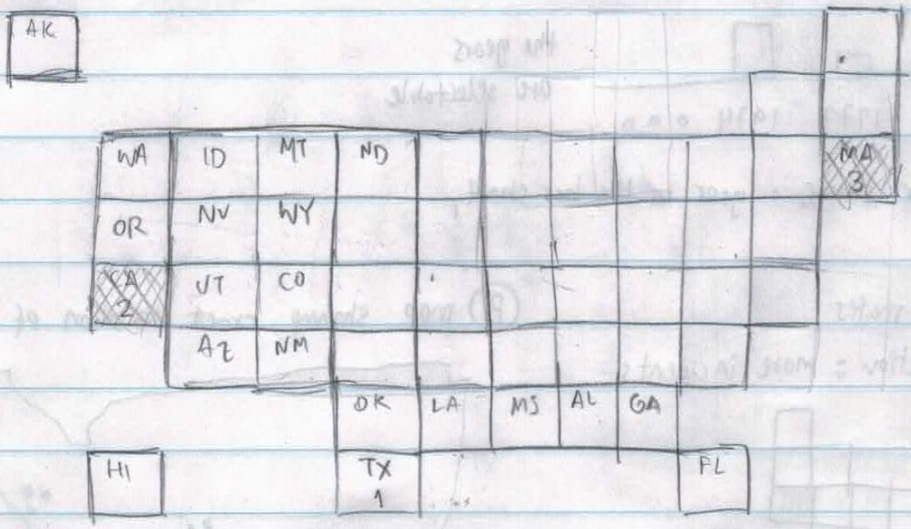
Another interesting alternative for an overview of the data per year is a bar chart of the number of victims, where each bar has two components: injured and killed, as shown in the following chart. In some cases, visualizing the total number of victims per year provides interesting insights that are absent in a chart of total number of incidents. For example, one single incident in 2017 had more than 500 victims, and the number of victims per year is often less than 50.



Filtering and details

Each map alternative proposed in the previous section provides different ways to filter the data and get more information about specific incidents.

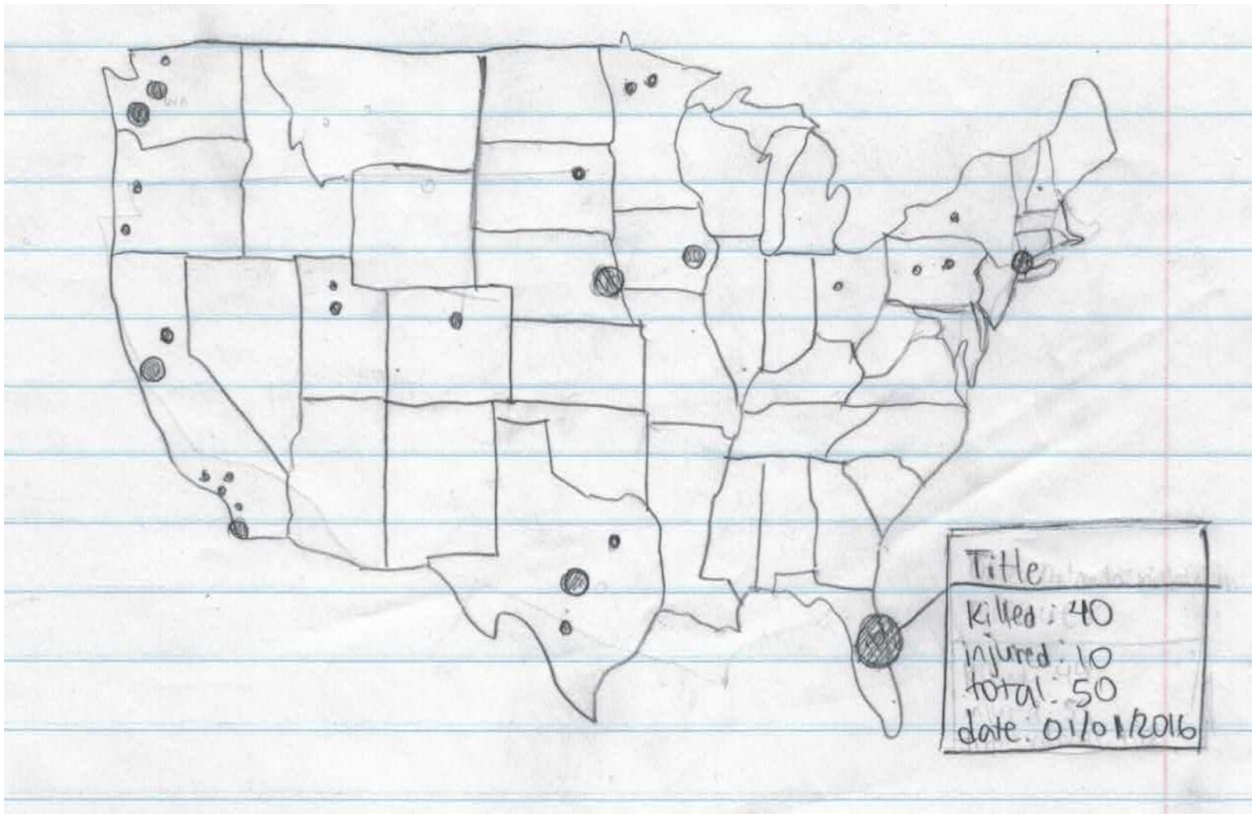
Map A: The user can select states from the tile map, which in turn displays a table of detailed information of the incidents in the selected states.



Selected States: CA, MA

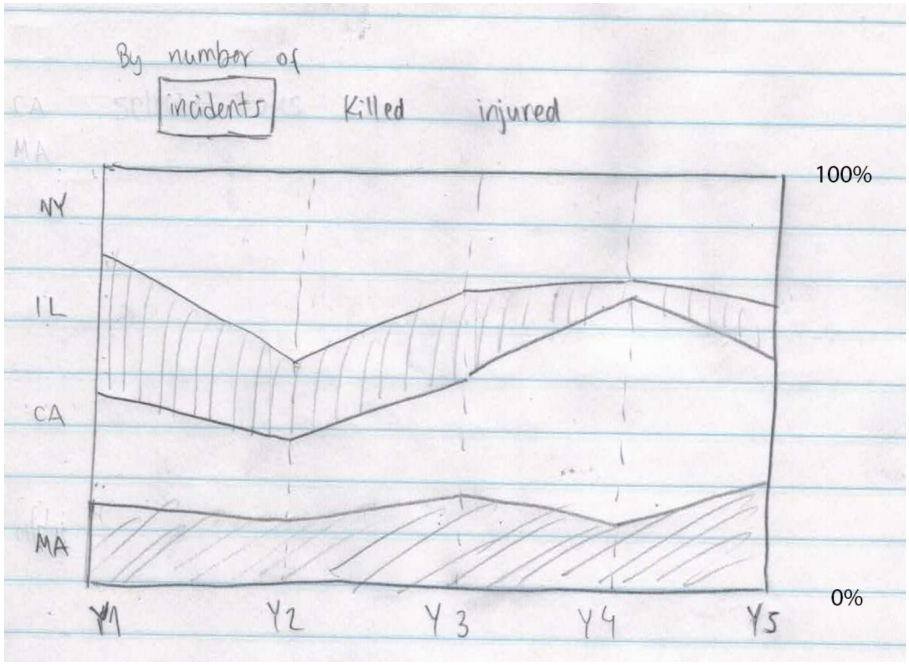
Title	Location	Date	Fatalities	Injured	...
title 1	San Bernardino, California	01/01/2016	1	5	
title 2	San Bernardino, California	02/02/2016	2	4	
title 3	Boston, Massachusetts	03/03/2016	3	3	...
title 4	Chelsea, Massachusetts	04/04/2016	4	2	
title 5	Boston, Massachusetts	05/05/2016	5	1	

Map B: This map uses circles as marks for each incident. The size of the circle can be used as a channel to encode a relevant criterion such as number of victims. When the user points to a circle, a tooltip with detailed information is shown.



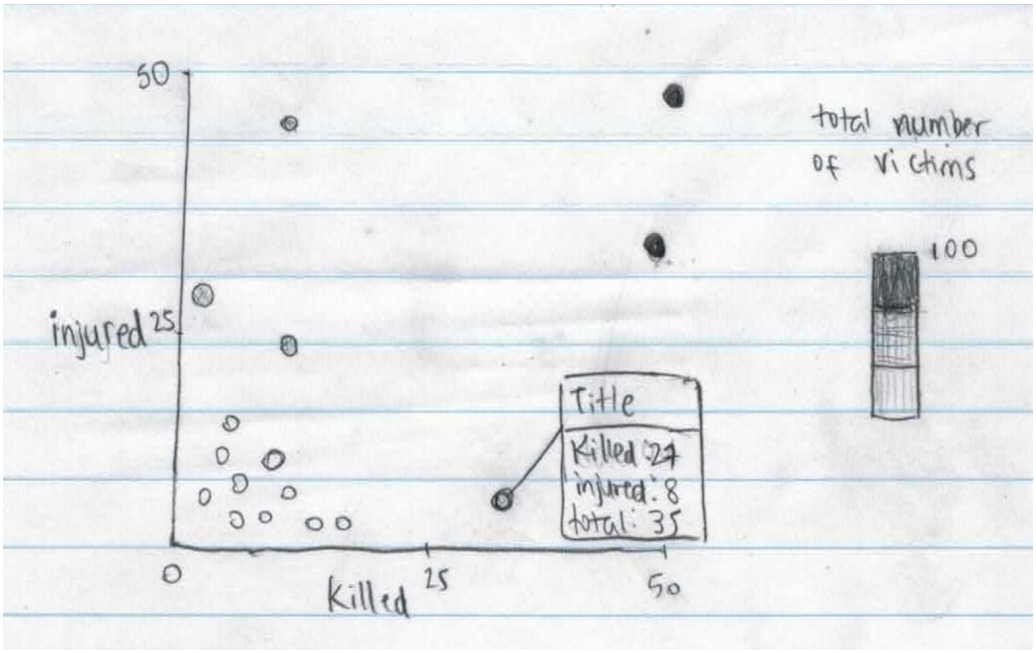
Comparison of changes over time in different states

We also want to get some insights about changes over time for different states. One alternative to achieve that is with a graph like the following, which shows the percentage of incidents in different years for different states. Other quantities of interest that could use the same type of graph are the number of people injured, killed and total victims. This type of graph is useful for our data because in some cases, plotting absolute values is not be the best choice. For example, one incident in 2017 had 585 victims, and the number of victims in one year is often less than 50. Comparing percentages is more appropriate for cases like this.

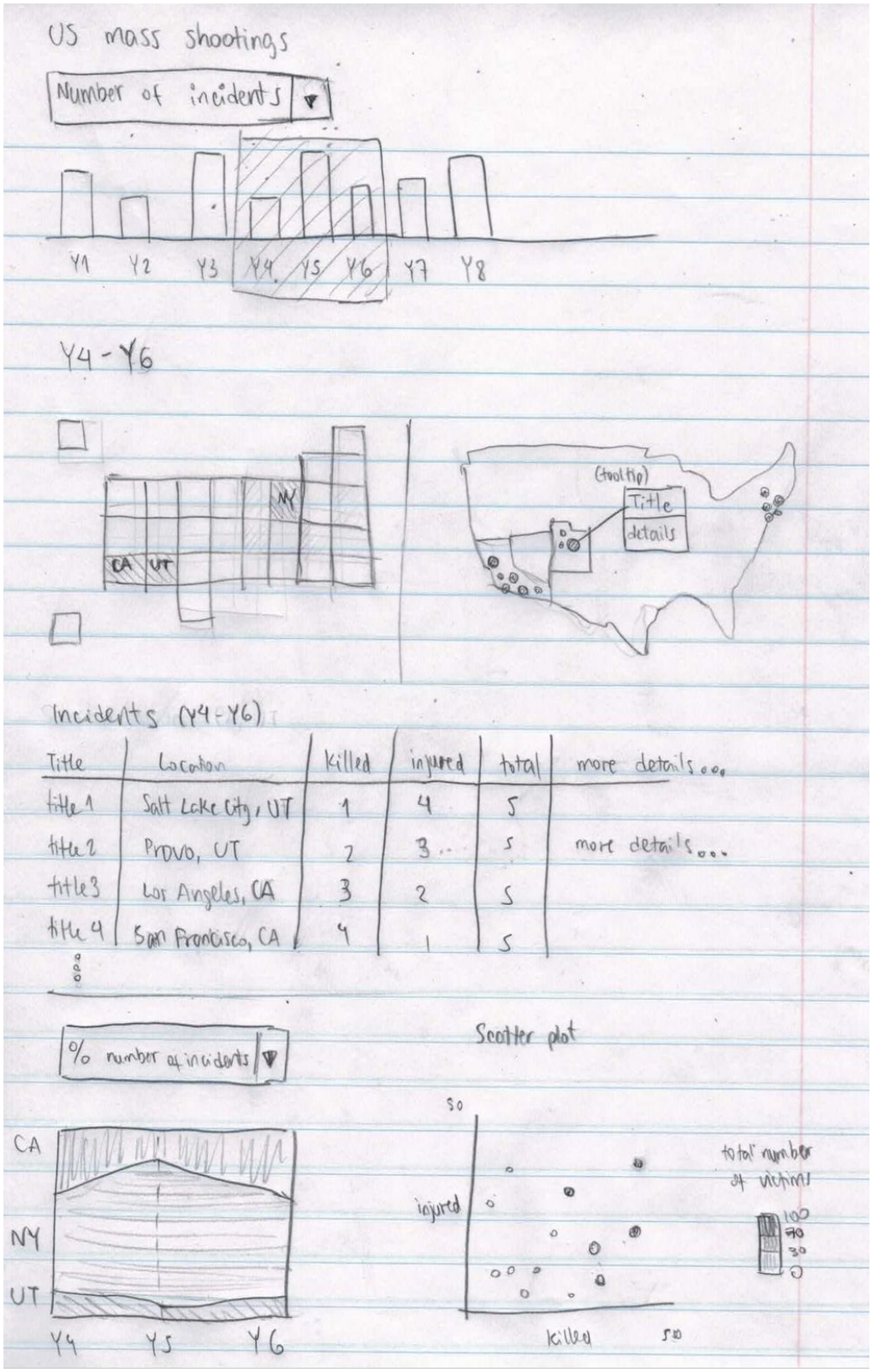


Correlation killed - injured

It would also be interesting to get some insights about correlations between number of people injured and killed in each incident, for which we propose a scatter plot. As shown in the following plot, each point corresponds to an incident and the color encodes the total number of victims. A tooltip is also available when the user points to an incident.



Proposed layout of the visualization incorporating previous designs



- **Must-have features**

  - Overview bar chart per year (either number of incidents or number of victims)
  - At least one map from the proposed options
  - The map must be updated to show information about the selected year
  - At least one graph that compares some quantity (such as incidents or victims) between states.
  - Scatter plot killed - injured
  - Table of detailed information about incidents, after filtering.

- **Optional features**

  - Allow the user switch between number of incidents and number of victims in the overview bar chart
  - Allow the user switch between different map options (tile map and geographical map) or show them side to side
  - Allow the user select more than one year
  - Allow the user switch between different comparisons (such as incidents or victims) between states
  - Allow the user rearrange the layout
  - Make the table of incidents sortable by column

- **Project Schedule**

Action Points	Oct	Nov				Dec
	Oct 28	Week 1	Week 2	Week 3	Week 4	Dec 1
Project Proposal						
Data processing						
Bar chart by year						
Table of incidents						
Map						
Project milestone						
Comparison between states						
Scatter plot						
Tweak interactivity						
Optional features						
Final submission						