

The Crime Landscape of Salt Lake City

Basic Information

Members:

- Archit Rathore - u1144416 (archit@cs.utah.edu)
- Yash Gangrade - u1143811 (yashgangrade09@gmail.com)
- Rebeka Mukherjee - u1141112 (rebeka.mukherjee@utah.edu)

Github Repository: <https://github.com/yashgangrade09/dataviscourse-pr-crime-landscape.git>

Motivation

The general consensus between us team members was to work with data that pertained to some societal factors. Another thing that we wanted out of our project was to be able to derive insights on a much finer geospatial resolution (state/city). Finally, we also wanted our visualization to be relatable and not dealing with data of a technical nature. This led us to explore the Utah Open Data catalog (<https://opendata.utah.gov/>) and we finally zeroed in on the police cases dataset that lists all reported crimes in the Salt Lake county.

This dataset checks off all our boxes, namely:

- Pertains to societal factors
- Is localized
- Is relatable and may be of equal interest to both a layman and a specialist
- Has the potential to inform policy decisions

Data

We are using the [Police cases dataset](#) from 2008-2016 (tentatively) from Utah Open Data catalog. Each record in the data corresponds to an instance of reported crime at one of the police stations in Salt Lake City. It contains information about the type, time and date of occurrence and reporting, and location of the crime. There are about 50K+ records per year. The data is open access.

The geospatial data will be pulled from either Google Maps or OpenStreetMaps (more inclined to use OpenStreetMaps). Showing crime statistics in context of local information such as businesses presents would be a 'nice-to-have' addition.

Data Processing

The Open Data catalog provides an export API to easily download data in a variety of formats. We will be exporting the data as CSV through the web pages for each year. The data is mostly clean with very few missing values ($<0.1\%$) for which we intend to drop the row itself.

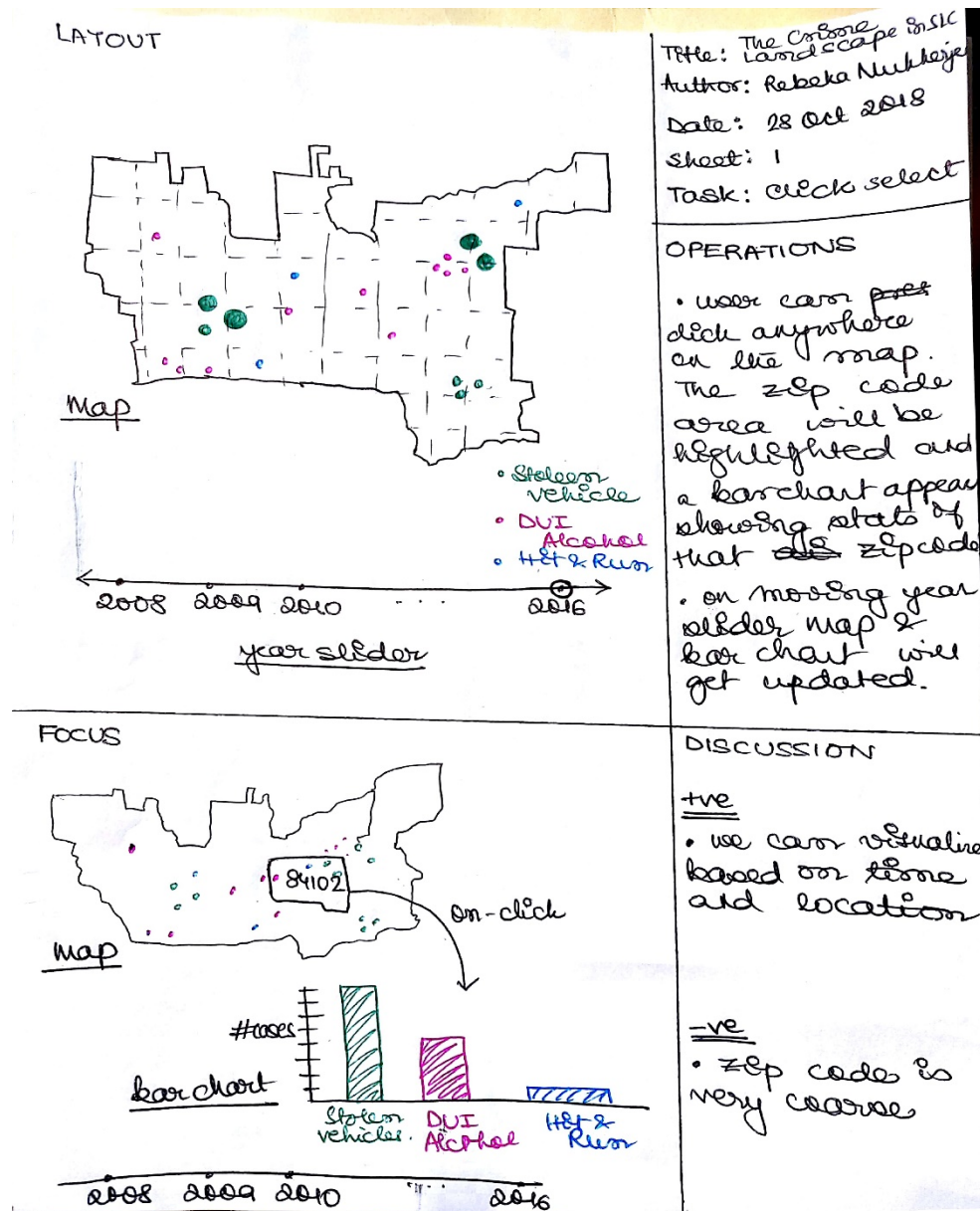
We will perform data aggregation based on time of occurrence, location, type of crime etc. This will be done using python/Excel and stored as separate files that are loaded on demand (instead of aggregation in browser).

The maps data should not require much processing as it will be fetched as is from the API.

Visualization Design

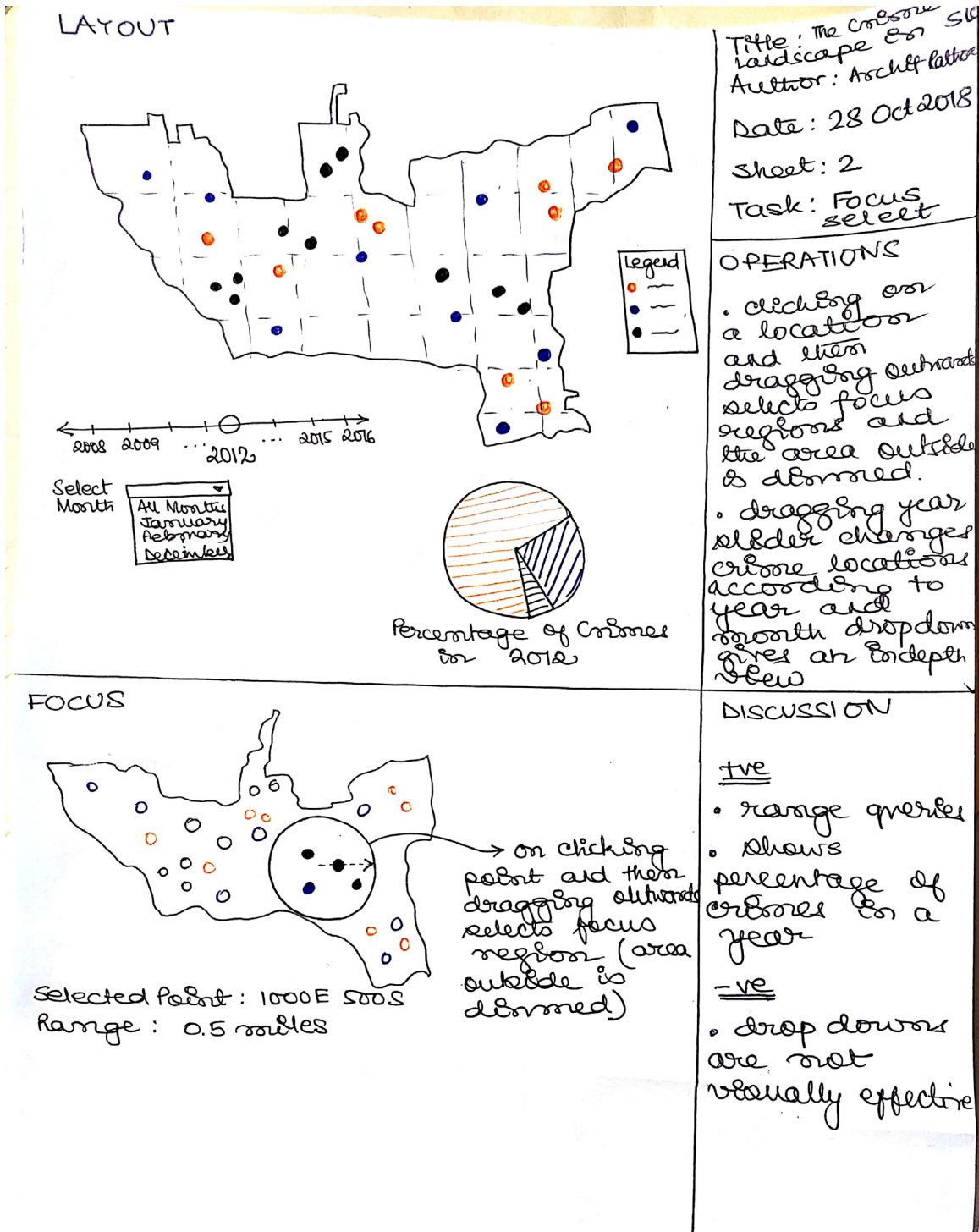
Prototype 1

In this prototype design, we are using a map view to show a map of Salt Lake county. This map will have a year slider and a month slider corresponding to the selected year. It will have certain points highlighted with different colors. These points represent different types of crimes happened at different locations in the county. A user can select a zip code and that entire area will be highlighted and a bar chart of different crimes will be created for that zip area. For every year and month selected, there will be a single bar chart.



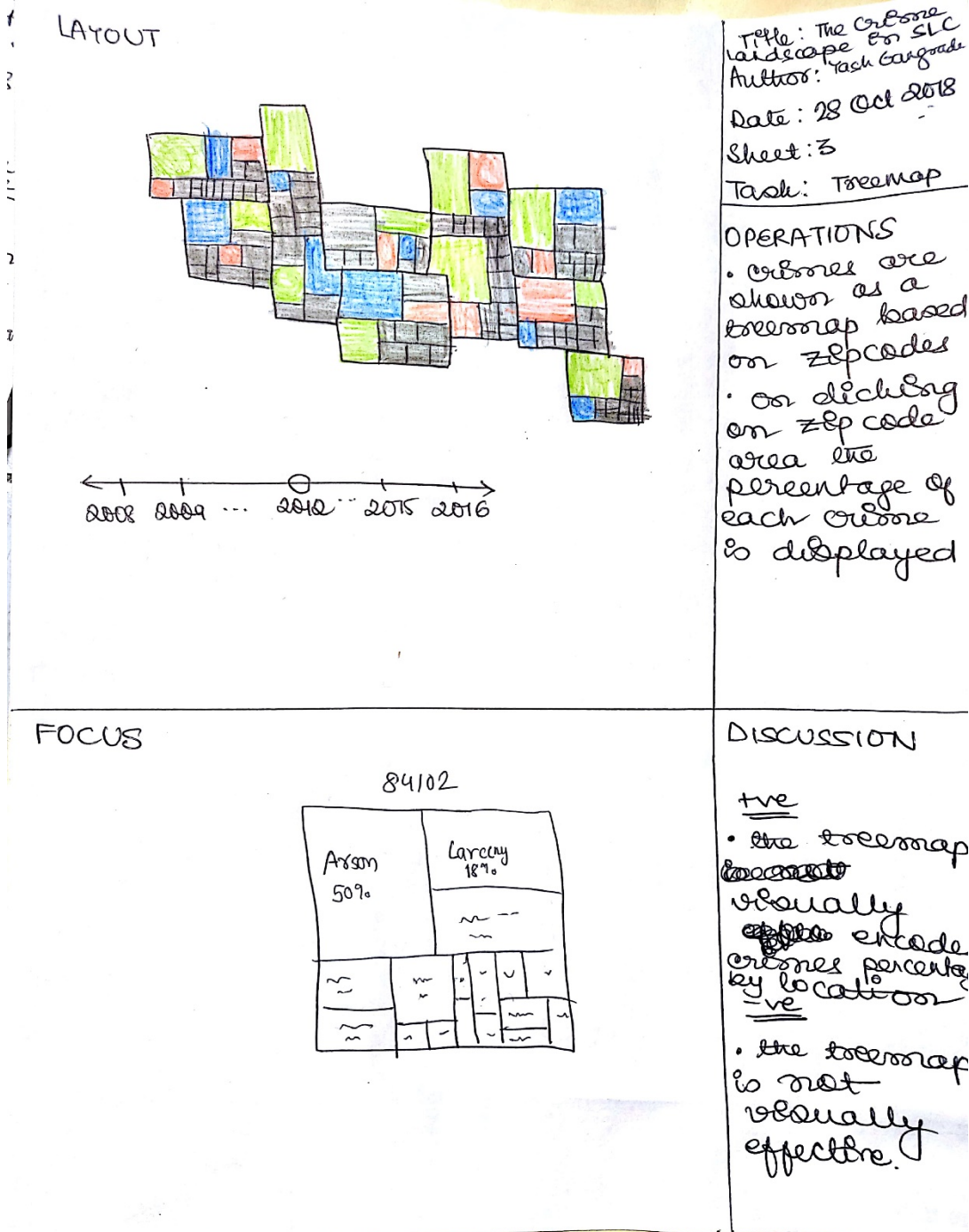
Prototype 2

We have a similar map view as prototype 1. A pie chart is added to represent distribution of crimes in the county for a selected year and month. However, here instead of clicking the user can click on a point and drag outwards to trace a circle which will then populate the information about that region below the map.



Prototype 3

We replaced the geological map with a tile map where each tile corresponds to one zip code. Each tile houses a treemap visualization of crimes for that zip. Clicking on any of the tiles expands the treemap in the tile and shows a detailed view of the distribution of crime.

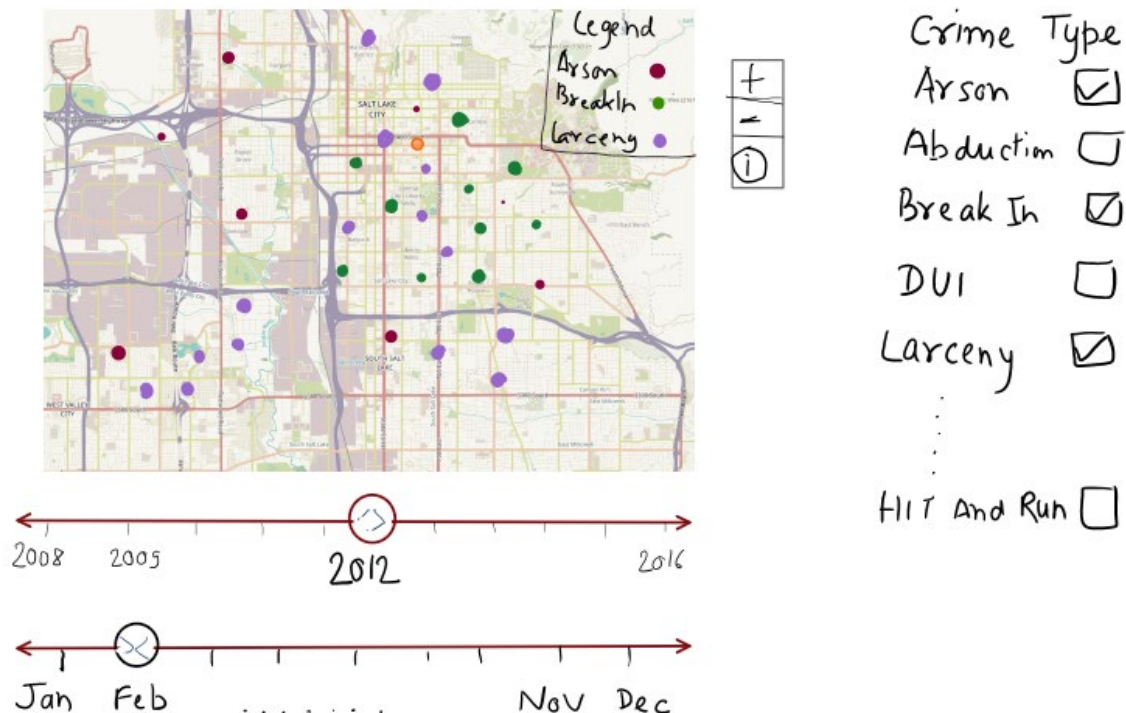


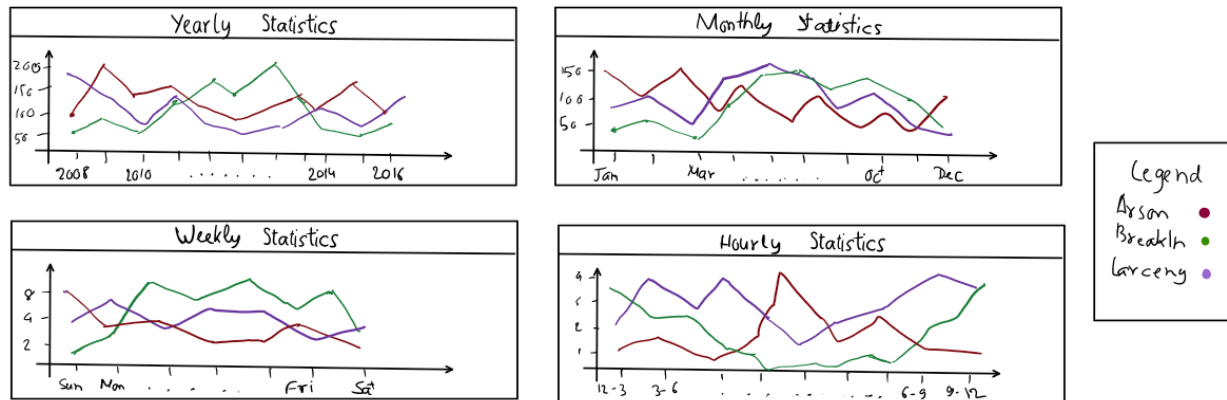
Final Design

In the final design, we intend to show a map view of Salt Lake County, and a statistics view to show the yearly, monthly, weekly and hourly statistics.

- The map view shows the location points of crimes selected in selection view and time selected by the time slider. The mark used is points and channels are color (type of crime) and position (location of crime).
- A time slider enables the user to see the trend of crimes over the years. On clicking the year slider it expands to show the months for the selected year. This enables the user to visualize the crime trends on a finer scale. It collapses to show only the years when clicked again.
- A selection view enables users to select the crimes they wish to visualize through checkboxes.
- The map view has a semantic zooming feature with zip code level aggregation at the topmost level and blocks at the lowest level.
- The statistics view shows yearly, monthly, weekly and hourly statistics of aggregated data about selected crimes as line charts.

The Crime Landscape of Salt lake City





Must-Have Features

- Functional map of Salt Lake City with an interactive overlay of the crime data.
- Interacting with individual data points shows details through tooltip.
- Show data over a span of years (through a year slider).
- “Semantic zooming” - when zoomed out the interface shows aggregated statistics, zooming in progressively refines the scale of aggregation until only individual datapoints remain. This would allow user to control the level of aggregation.

Optional Features

- Incorporate other data from the Map API such as number of businesses in the region, distance from the nearest police station etc.
- Toggle between detailed and summary view.
- Radial range selector.

Project Schedule

Week 1: October 29 - November 5

- Oct 30 - Peer Feedback
- Work to do: Pre-process Datasets

Week 2: November 6 - November 13

- Nov 9 - Project Milestone due
- Work to do: Create and populate the map view

Week 3: November 14 - November 21

- Meeting with Mentor to finalize things
- Work to do: Implement Statistics view

Week 4: November 22 - November 30

- Nov 30 - Final Report due
- Work to do: Wrapping up things, attempt optional features, creating screencast.