How High?

The Cost of Recreational Drugs

Group Members

- Chalenge Masekera (chalenge@ischool.berkeley.edu)
- Marjan Ghahremani (marjan@ischool.berkeley.edu)
- Pablo Arvizu (parvizu@ischool.berkeley.edu)

Instructor: Professor Marti Hearst

URL: http://people.ischool.berkeley.edu/~parvizu/infoviz/howhigh/

Project Goals

Our initial goals for this project was to show the effects of recreational drug consumption in the United States has in the rest of the World by the way of crimes such as homicide, trafficking, extortion and kidnapping. Since the United States is largest illicit drug market in the world, Mexico is the main country from where drugs are transported to the U.S. We decided to focus on these two countries and build an interactive infographic to show the "cost" of different types of recreational drugs in terms of the footprint they left behind. We wanted to show how this problem not only affects Mexico but the rest of the world too.

We wanted the interface to allow the user to understand the context of the drug problem in the US in a guided way. At the same time, we wanted to enable the user to explore the different information and statistics that we present and let them come up with their own insights. As a whole we wanted our project to be a narrative visualization with exploratory elements.

Related Work

The United Nations Office of Drug and Crime has a website with several interactive maps that attempt to display statistics about drug prevalence in the population over 12 years old in several countries. Even though the information is relevant we concluded that the visualization had a lot of shortcomings and was not able to relay insights to the user.

We found several infographics that helped us learn more about the topic, but even though the overall subject was drugs, the emphasis was not similar to what we had in mind. Some of them focused solely on the economic cost of the war on drugs from the United States perspective, others focused on the benefits of legalizing marijuana in the United States. None of these really touched the subject of how recreational drugs consumption affects other countries.

We found a few others that were helpful in defining our scope, particularly the following:

- **Invasion of the Drug Cartels -** A very well done and interesting infographic by the National Post about the drug cartels in Mexico and how they have been "invading" the United States (picture below). These infographic was as useful as it was inspiring to continue with our project.

- **Global Drug Trafficking -** An infographic done by RIA in Russia about drug trafficking focusing on the main countries producing illicit drugs and the routes they take to the main consumers.

- **Illegal Drugs in the Americas -** This infographic by NGP Pharma focused on the production of illicit drugs through the Americas as well as drug seizures by regions. Even though this infographic is not easy to understand, it helped us realize some of the data that can be found as well as the relation between production and the consumption in the U.S.
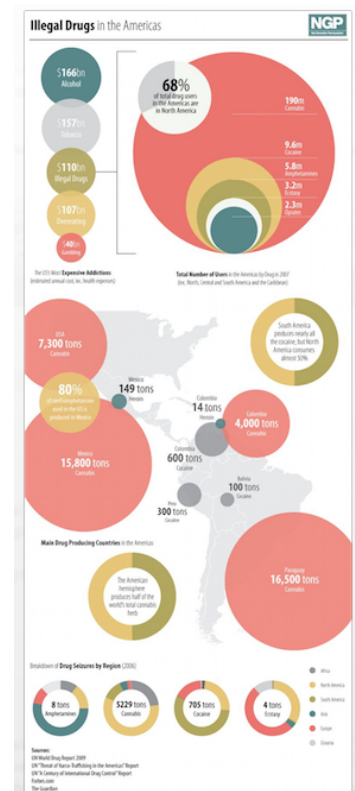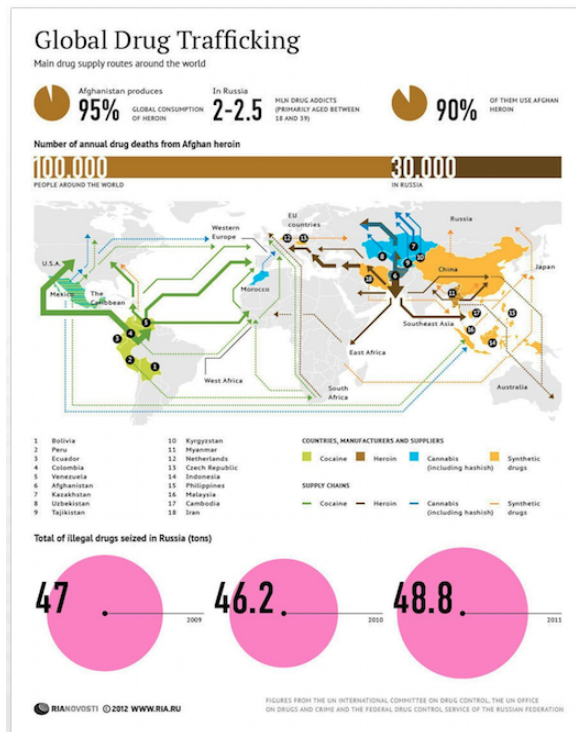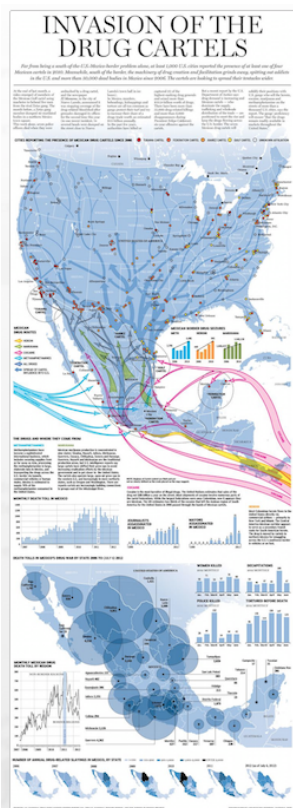
Overall, infographics were a great resource into exploring and understanding different ways to visualize the type of data that we would be dealing with. We could see some visualizations that really did not work, like using donuts for distributions, but then we noticed some of them really worked, like the choropleth by state.

Description

Our visualization consists of three different screens and within each screen there are multiple sections. Some of these sections work together through linking and brushing, while some of them only complement each other on the content level.

**First Screen: Recreational Drugs consumption in the United States**
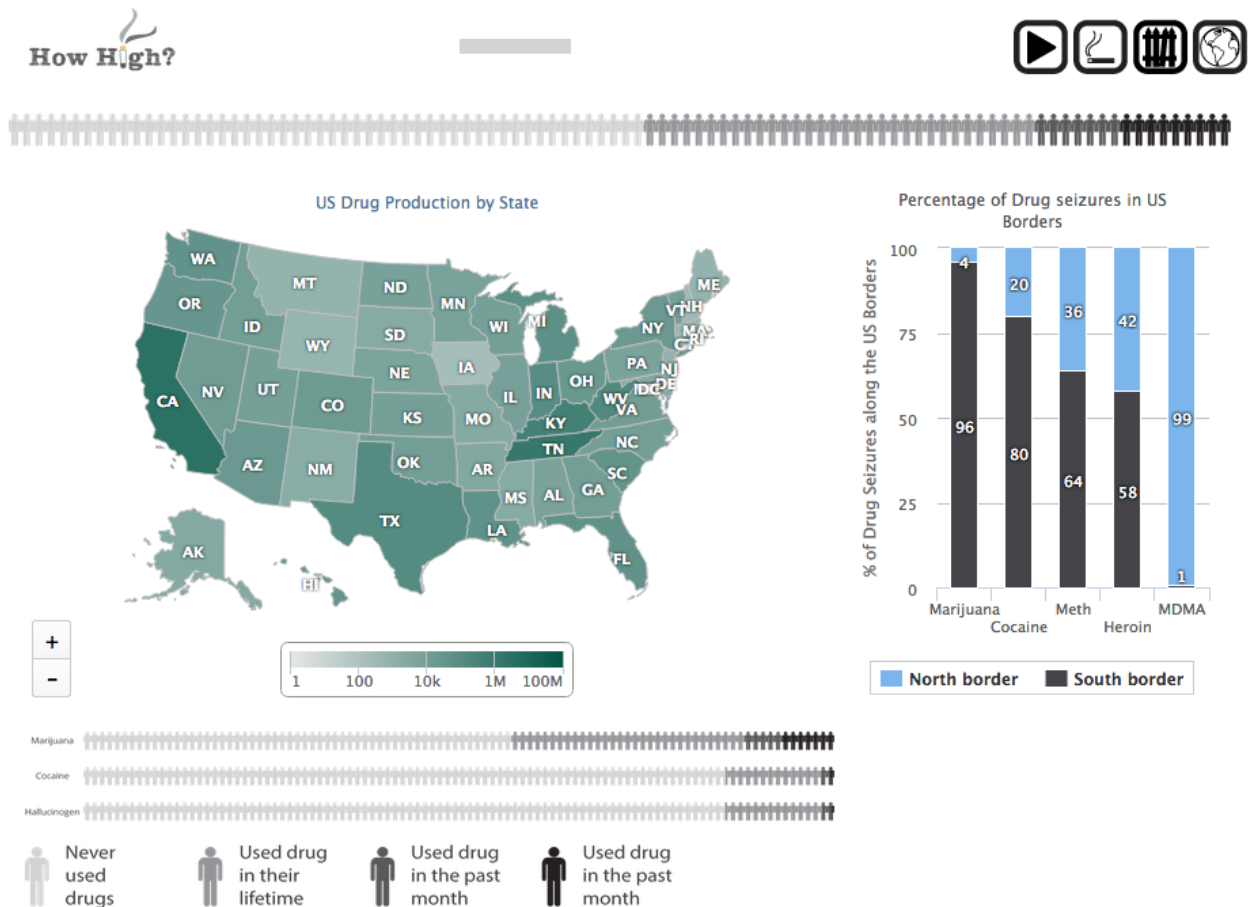
The first screen of the visualization is focused on the United States. It provides statistics about drug production within each state, and general drug use patterns among the American people, as well as drug use patterns for the three most commonly used drugs. It also provides an overview about the routes through which illicit drugs are transported to the United States.


This simple infographic/chart shows the overall condition of drug use in the U.S. for people aged 12 or older. It gives the user a general idea of the percentage of U.S population who have tried illicit drugs at some point their lives, as well as those who have used drugs in the past year and past month.
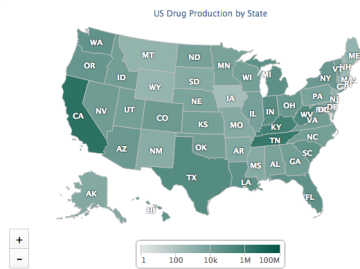
The choropleth map of the United States, which appears as the main component of the first screen, shows marijuana production by state. When the user hovers over a state, the state is highlighted and the tooltip shows the statistics for that state.
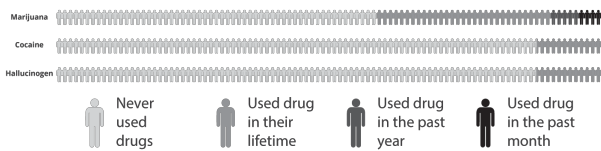

This section follows the same format as the introductory infographic. It shows the drug use in the U.S. for people aged 12 or older for three of the most common illicit drugs: Marijuana, Cocaine, and Hallucinogens.
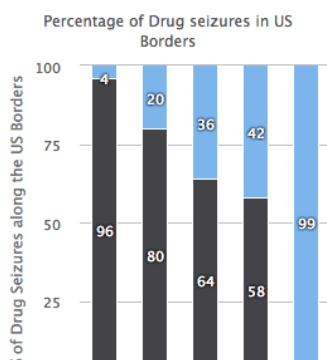
This chart shows the percentage of drug seizures along the north and south borders of the U.S. We used columns instead of horizontal bars to increase the match between the system and the real world, so the users can intuitively tell the north and south apart.

Second Screen Illustration: Violent Crimes in Mexico

This visualization shows drug related crimes in Mexico. Having established that most drugs that flow into the US come from Mexico, we wanted to explore whether with increased drug usage in the US, drug crimes are also on the rise in Mexico.
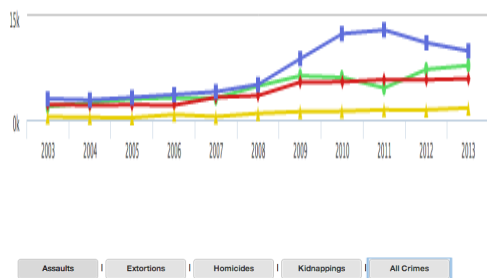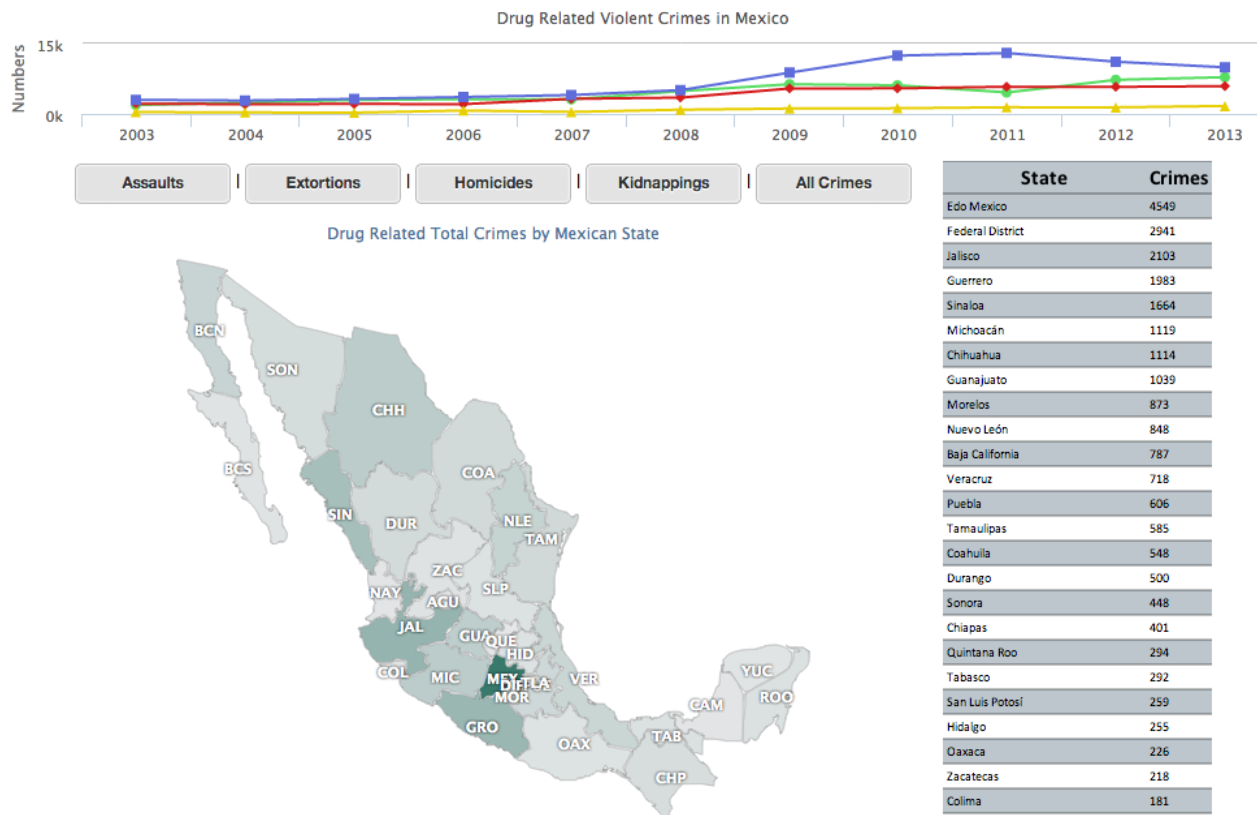
The line chart shows the trend of particular types of violent crimes for the past 10 years. Each line represents a type of crime. The chart is linked to all the other maps such that selecting a type of crime on the map cascades to the other visualizations on the screen

The five buttons represent categories of violent crimes committed in Mexico as labeled. Selecting one filters all the charts on the visualization.
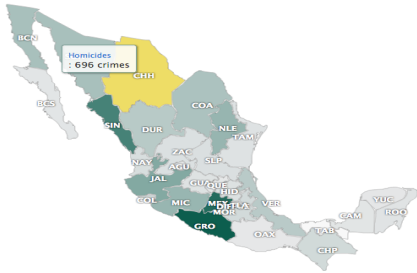
The choropleth map crime statistics by a type of crime in each state in Mexico. Darker colors represent more incidents of crime. Hovering over the state brings up a tooltip which shows the actual number of crimes committed.


The table which is to the right of the map ranks the states according to selected type of crime. The table is linked to the map such that when a user hover over a state on the map, the appropriate result in the table is highlighted so that it shows where exactly the state ranks for the crime.

Third Screen Illustration: World Drug Seizures and Drug Related Crimes

The third screen in our project focused on the worldwide statistics about drug seizures and drug-related crimes. In that sense we created two views within the same screen, one for seizures (top image) and one for the crimes (bottom image).
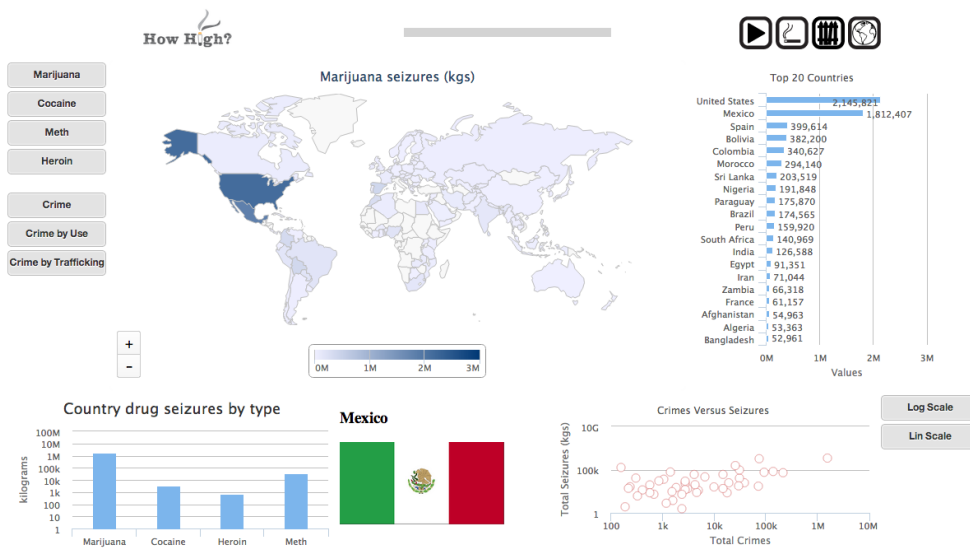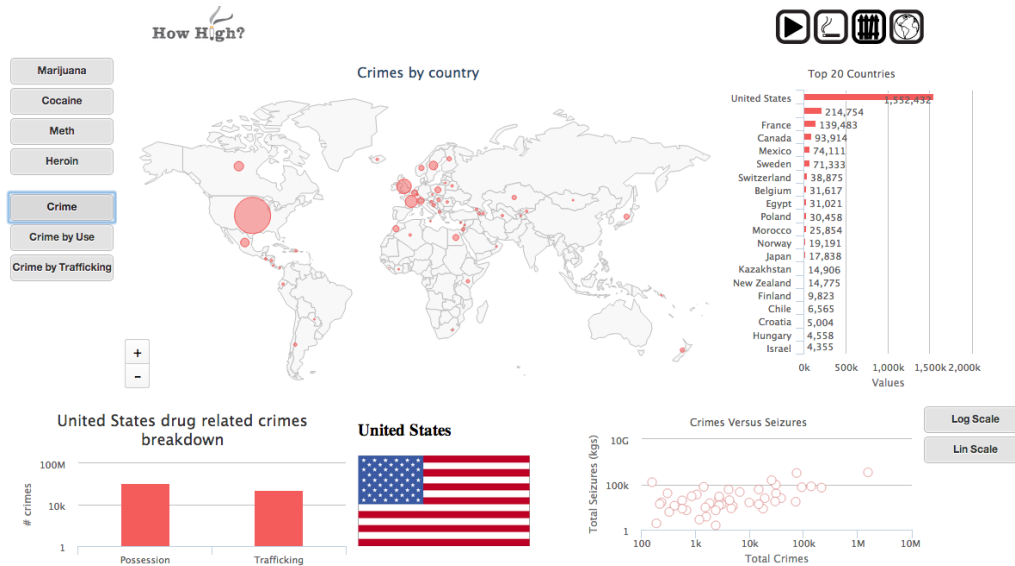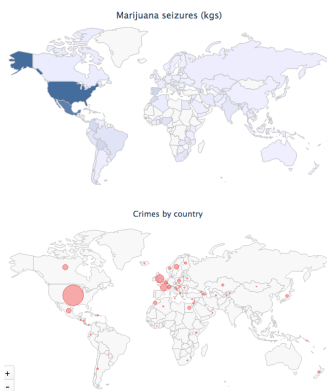
Both screens have the same structure and use the same charts with slight changes in the data being presented. Since this screen has multiple elements they are explained in detail in the following table.
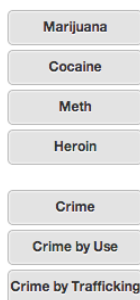
	On the left side of the screen we placed buttons to allow the user to select the different either the type of drugs or the types of drug-related crimes they want to visualize in the multiple charts of this screen.
	Selecting either one will dynamically change the information in the map and the top 20 ranking by countries.
	The map shows the as a blue choropleth the seizures of the selected drug by country in kilograms or can show the number of drug related crimes by country displayed with red bubbles of different sizes.
	The map can be explored with the mouse which by hovering over the different countries will show the number of seizures of crimes for that country. Also, the map allows the user to select a specific country to show a breakdown of the seizures by drug type or the crimes by type in the column graph of the bottom.
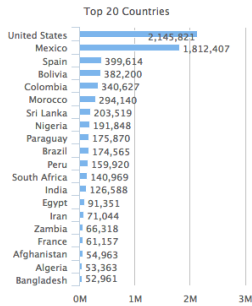
The chart on the right side show the Top 20 Countries of the selected statistic in a descending order. This allows the user to have an overall idea of which countries either affect or are affected the most by the selected statistic.

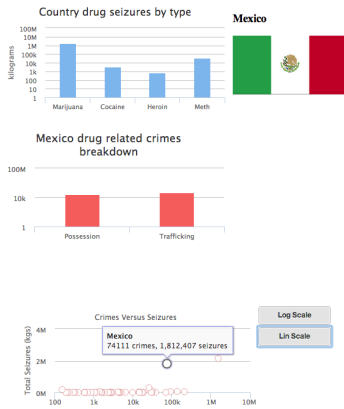The bottom left graph show a breakdown of a specific country by the selected category. For drug seizures, shows the overall amount of kgs seized for each drug and for the crimes the total number of crimes reported by type. This chart is plotted in a logarithmic scale to avoid confusion with small amounts reported for some countries. The side flag allows the user to have a clear view of which country is selected.

Finally, on the bottom right side we plotted a scatter graph of drug seizures (y axis) and drug related crimes (x axis) in an attempt to show the overall relation of these two statistics. We added the option to show it in a linear or a logarithmic scale, which allowed us to find a couple of outliers.

Data

The data we used was our own compilation from various sources but mainly from the Drugs Enforcement Agency (DEA), the United Nations Office of Drug and Crime (UNODC), and the US Department of State websites. We were able to mine other data from the UNODC public database. For the statistics about Mexico's crimes we used information gathered manually from the Mexican Government's Executive Secretariat Delinquency annual report.

- Past drug use by drug type in the U.S. in population over 12 years old
- Marijuana production in the US by state for 2012
- Violent Crime statistics in Mexico by state and by type for the past 10 years
- Percentage of Drug Seizures along the borders by type
- Drug Seizures by country in kilograms and by type

- Drug Related Crimes by country and by type

We used a number of tools focused on different parts of the project as highlighted below:

Data cleaning and preparation: Whilst we manually prepared most of the data we used from a number of various source, we used online CSV to JSON file converters as well as Microsoft Excel to create JSON data format we required for the visualization.

Data Analysis: We used Tableau for exploratory data analysis. Since we had not found compelling data which tied with our initial project goals we hoped data exploration would give us trends on drug usage that we tie to our story as well as help us uncover unique insights.

Visualization: To implement the actual visualization we used JavaScript and CSS3 for data manipulation and for the overall look and feel. All the charts and maps were implemented using the Highcharts.js library. We also used Adobe Illustrator to design the isotypes used to show drug usage patterns.

Testing: Testing was done manually since we did not have time to implement automatic testing throughout the development process.

Definition of Project Goal and Objectives

This involved all the team members and we discussed the project goals. We created an outline of the three 'stages' of visualizations that we thought we would tell the story as shown:
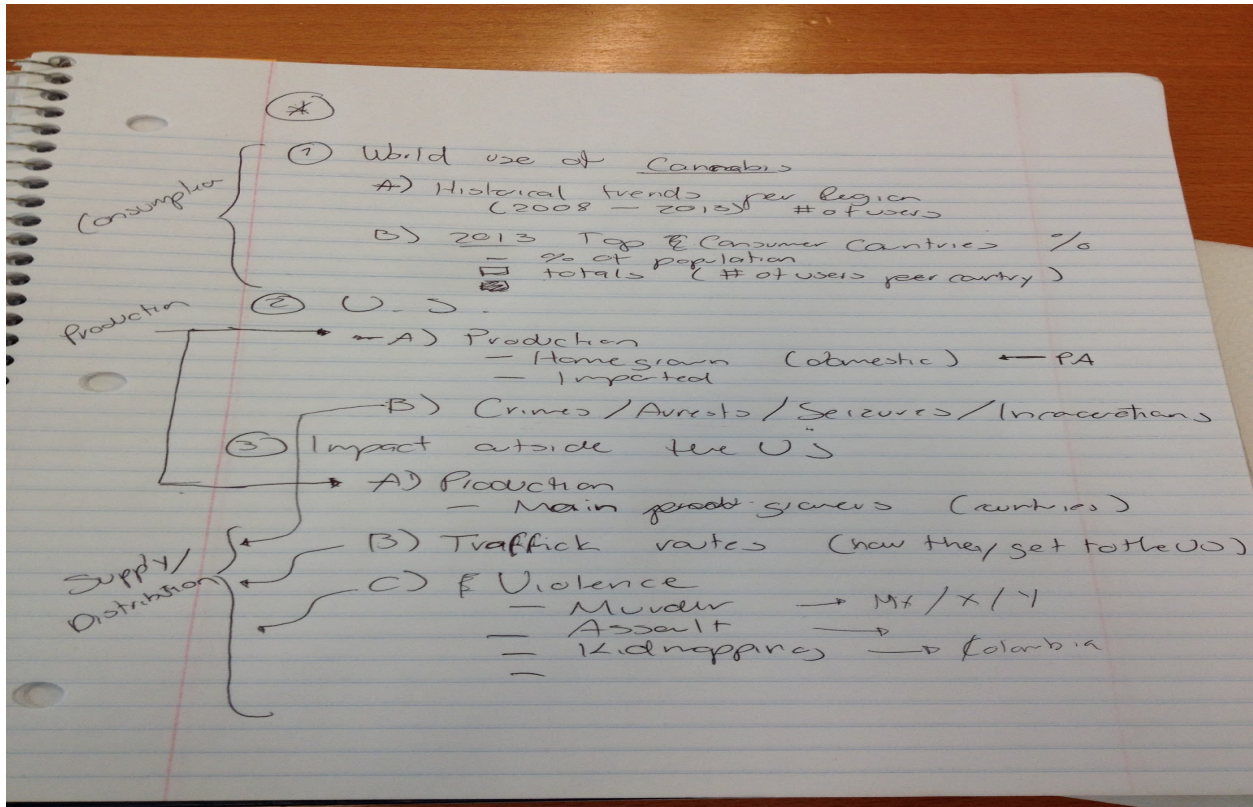
We identified the information we wanted to show and tasked each of the project members to gather appropriate data for the identified project parts. Since we did not know of any data set that readily had the data we decided to first gather the data then make design decisions with a firm understand of the data.

Data Understanding, Discovery and Collection

Contrary to what we assumed, finding data was the most challenging part of our project. We were not able to find a single dataset that provided all the information we needed for either the U.S. and Mexico or the world. We initially started by searching for detailed information that would be able to show each of the sections we had identified in our project goals. This resulted in seemingly endless data search across many different reports up until we gave up and went back to redefinition of project goals.

Redefinition of Project

Whilst trying to understand the data we had managed to collect lots of incoherent and often incomplete data with a wide range of topics. We then looked at the available data and analyzed it to discover what story we could narrate that was closely related to our initial goals. We identified the following topics as relevant to our narrative visualization:

- Drug Usage and production patterns in the US
- Drug flows into the US and drug seizures along the borders
- Drug related violence statistics in Mexico
- Drug production and trafficking patterns across the world.
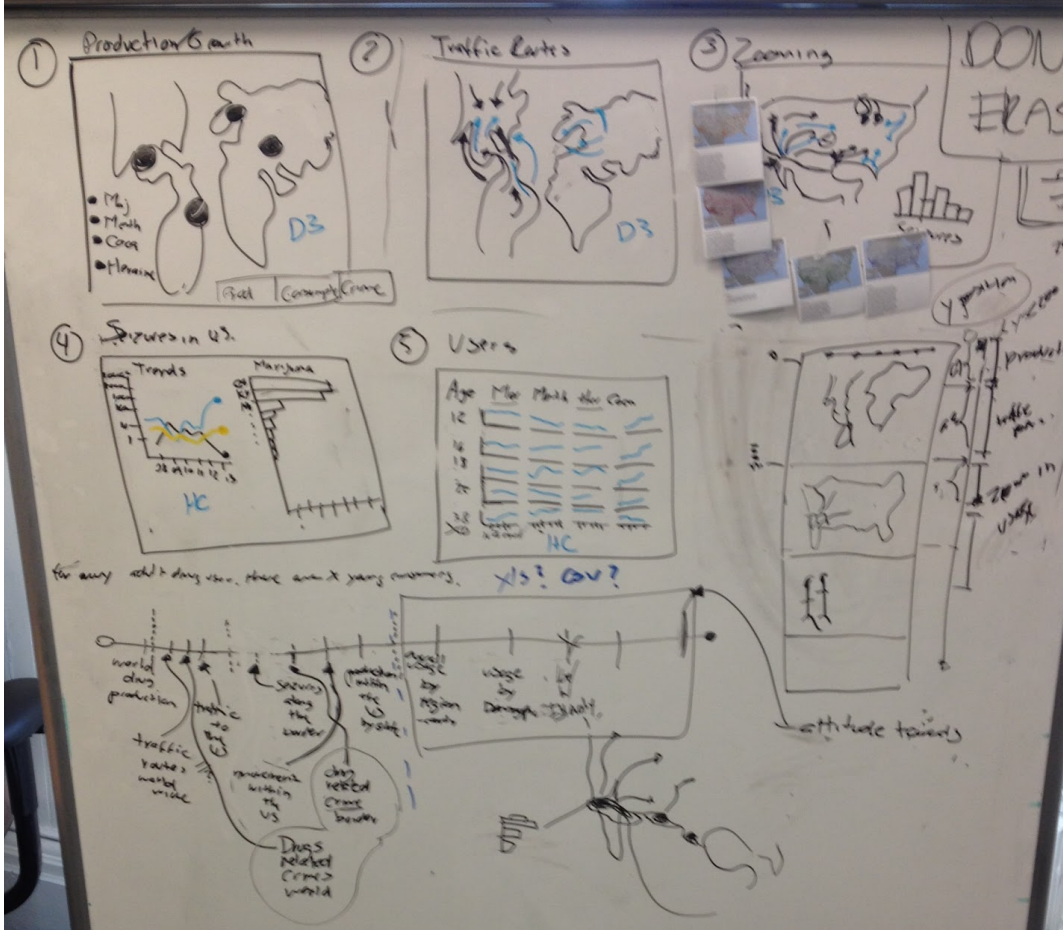
Initial Design Formulation

With the 'new' objectives we mocked up our first design which was based on a lot on the initial ideas of showing where the drugs are produced, how they flow into the main market (US) and lastly show the usage trends. This initial concept is illustrated below:
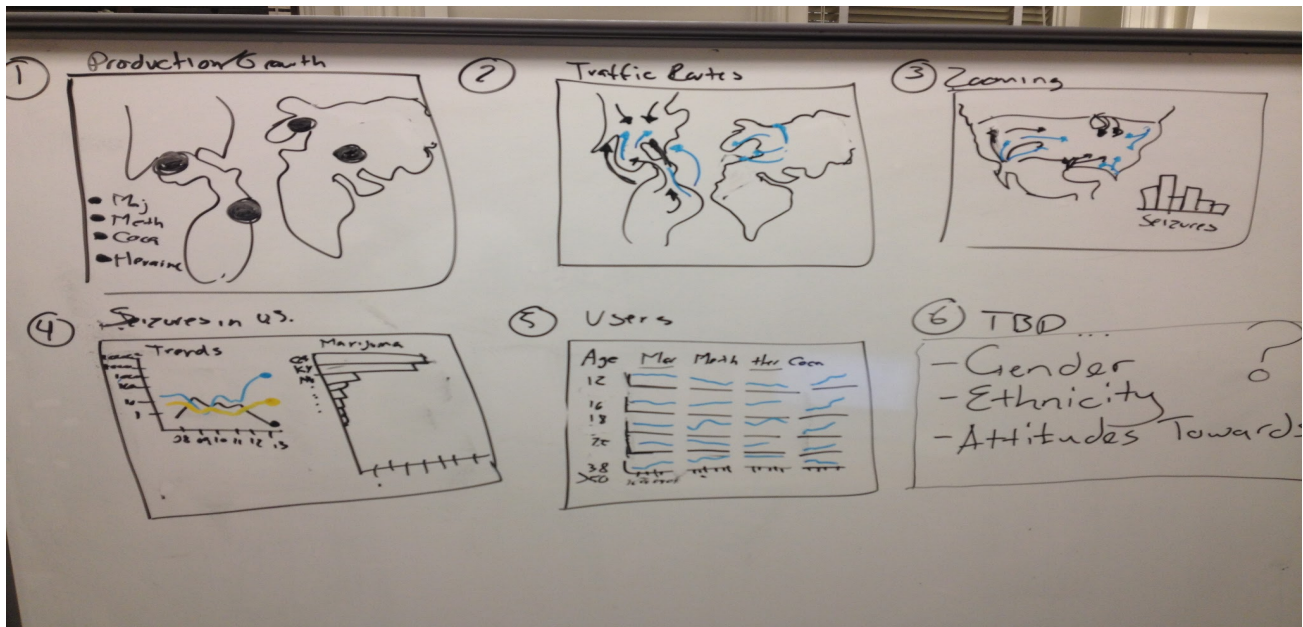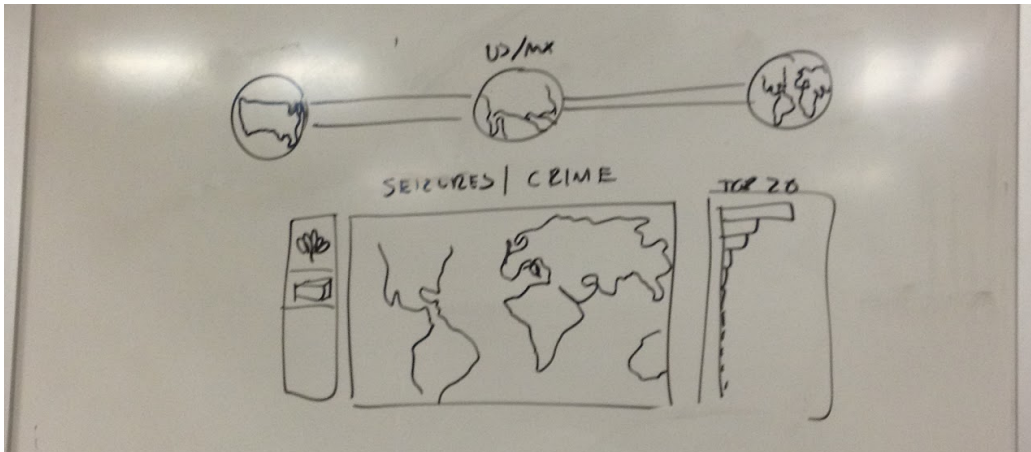

[First Iteration]

We envisioned a narrative visualization which would narrate as follows using the order numbered on the storyboard using a single screen. Tooltip text would be implemented that would highlight useful insights that the user should be drawn to.

1) First show the production patterns of various drugs across the world.

2) The visualization would then transition to show drug flows from production sources to market where the drugs are consumed.

3) The third stage would involve the visualization and map zooming into the US/Mexican border and show drug seizures along the border as well as other drug related crimes in the border states.

4) The fourth visualization would show trends in drug production and crimes in all the other states in the US.

5) The last visualization would then transition to show drug usage patterns within the US according to various dimensions such as age, gender and ethnicity. We intended this to highlight how drug behavior patterns are changing and affecting minors among other factors.

[Second Iteration]

[Third Iteration]

[Fourth Iteration]
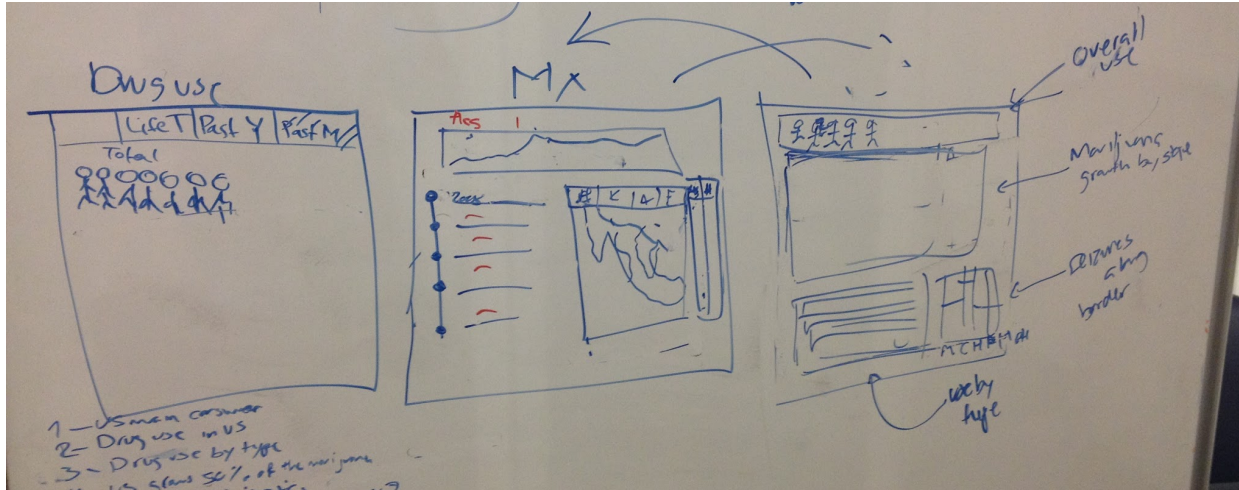
[Final Iteration]

We went through multiple iteration until we landed with our final interface which included 3 main screens:

1) The initial screen would show the drug past usage of the population in the US above 12 years of age, but would do it with the help of an animated infographic and a bubble as a guide. We would also add the marijuana production by state in the U.S.

2) Would focus on the crime statistics in Mexico by state. Allowing the user to select each type of crime and see the results in a map and the table.

3) The world view which would give the user more freedom to explore statistics about crime and seizures while comparing rankings and breakdowns of such statistics. This would be the most information intensive screen of all.

Data Preparation

Since we were unable to find a 'complete' data set we pieced together data from various sources in csv format for easy analysis and exploration. After tying the data we converted all the json format as it seemed much easier to work with in Highcharts. We ended up with a lot of files so as not to cram all the data in one file. An example of the data structure we used is shown below:

```
|[{"code":"US","value":1552432,"ranking":1,"name":"United States"},
{"code":"UK","value":214754,"ranking":2,"name":"United Kingdom"},
{"code":"FR","value":139483,"ranking":3,"name":"France"},
{"code":"CA","value":93914,"ranking":4,"name":"Canada"},
{"code":"MX","value":74111,"ranking":5,"name":"Mexico"},
{"code":"SE","value":71333,"ranking":6,"name":"Sweden"},
{"code":"CH","value":38875,"ranking":7,"name":"Switzerland"},
{"code":"BE","value":31617,"ranking":8,"name":"Belgium"},
{"code":"EG","value":31021,"ranking":9,"name":"Egypt"},
{"code":"PL","value":30458,"ranking":10,"name":"Poland"},
{"code":"MA","value":25854,"ranking":11,"name":"Morocco"},
{"code":"NO","value":19191,"ranking":12,"name":"Norway"},
{"code":"JP","value":17838,"ranking":13,"name":"Japan"},
{"code":"KZ","value":14906,"ranking":14,"name":"Kazakhstan"},
{"code":"NZ","value":14775,"ranking":15,"name":"New Zealand"},
{"code":"FI","value":9823,"ranking":16,"name":"Finland"},
{"code":"CL","value":6565,"ranking":17,"name":"Chile"},
{"code":"HR","value":5004,"ranking":18,"name":"Croatia"},
{"code":"HU","value":4558,"ranking":19,"name":"Hungary"},
{"code":"IL","value":4355,"ranking":20,"name":"Israel"},
{"code":"KE","value":4146,"ranking":21,"name":"Kenya"},
{"code":"MU","value":4056,"ranking":22,"name":"Mauritius"},
{"code":"SI","value":3063,"ranking":23,"name":"Slovenia"},
{"code":"KG","value":2785,"ranking":24,"name":"Kyrgyz Republic"},
{"code":"AZ","value":2784,"ranking":25,"name":"Azerbaijan"},
{"code":"EC","value":2773,"ranking":26,"name":"Ecuador"},
{"code":"MV","value":2418,"ranking":27,"name":"Maldives"},
{"code":"BY","value":2383,"ranking":28,"name":"Belarus"},
{"code":"DO","value":1862,"ranking":29,"name":"Dominican Republic"},
{"code":"NI","value":1796,"ranking":30,"name":"Nicaragua"},
{"code":"GE","value":1616,"ranking":31,"name":"Georgia"},
```

Each of the json files we created consisted mostly of a dictionary with four variables:

(i) Code - This could be either the country code or state code

(ii) Name - The name of the state or country

(iii) Value - Stored a continuous variable that was the statistic being measured for example number of violent crimes or amount of seized drugs in kgs.

Design Implementation

Most of the feedback we received after an initial design implementation was of the effect that we needed to focus a lot on the flow of the visualization and 'finding threads that can link the specific data' together. We then changed the order of our narrative from starting broadly looking at the world to looking specifically at the patterns in the US, into the effects in Mexico where most drugs come from, then lastly generalize to give an overall picture of the whole world.

Results

All the users really liked the informational bubble that we used to guide them through the first section of the visualization, but they demanded for more control. They wanted to be able to pause, go back, and control the speed of the tour. The team was also aware of this shortcoming, but we could not implement this functionality due to the time limitation.

Users also liked that they could explore the data based on various segments, such as drug type and crime, and within each segment they could explore the details for each country. Overall, users found the visualization very informative specifically about the United States and Mexico. Some of them found the numbers about the U.S. quite surprising. One user said:

> *"I knew that the drug problem in the U.S. is larger than many other countries, but I didn't know that the discrepancy is so significant."*

Another one which we'll leave as anonymous:

> *"Looks great, but now I feel guilty about some things I've done in the past. I am gonna go now."*

We also got good feedback about improvements in the interaction as well as the presentation of the information.

- One user mentioned that we should remove the animation that redraws the maps every time you change a stats (type of crime or drug) because then it is difficult to compare the stat with the previous one in the same map.
- Another important visualization improvement would be to now show multiple lines in the very first visualization during the introduction video because it is confusing as it seems that they are different statistics. The suggestions is to do the fade in but on the same isotype visualization for the breakdown of past usage.
- Some of them preferred to see 'per capita' numbers instead of raw numbers.
- Another suggestions was a normalization of these numbers to 100,000 inhabitants per country.

- Another suggestion for the border seizures graph was to label the series differently. Just use Mexico and Canada instead of South and North which can be confusing.
- For the bubble, maybe another color would have been better since it might get confusing with some of the visualizations.

Key Challenges

Unavailability of Data: We spent most of the time gathering data and aggregating it in a bid to ensure the data tells our story and not the other way around. This however proved tricky and we ended up changing some objectives. We learned that because of the nature of the subject, there are not many official and accurate statistics about drug consumption and production in the US or other countries of the world. Most of this numbers are estimates based on seizures or polls. Since consumption and production and illegal in most of the world there is no way of knowing the exact data for this activities.

Different data code: Some of the data we got had different codes/ resource descriptions for the same entities, for example country codes varied a bit across data sources. We thus had to manually do cross overs of the names to match them.

Beta version of Highmaps.js: The maps API of Highcharts.js (Highmaps) is still in beta version and most functionality is not documented. We thus had to learn through trial and error and looking at examples which sometimes functioned unpredictability. Also, we came up with a bug in the Highmaps library which didn't include the functions to create column charts which is part of the original Highcharts library. Because of this we had to include both libraries but making slight modifications in the way we referenced them in our javascript code to avoid conflicts.

Highchart limitation: Whilst highcharts are quite useful for basic graphs, it becomes tricky when you try to customize some functionality.

Narrative: Since our project had multiple visualizations, statistics, and charts we figured that it would be hard for the users to focus on important insights and might be lost with the

exploratory nature of the implementation, particularly in the first screen which introduces the context and the nature of the whole project. In order to fulfill our narrative objective along with the interactive nature of our project we developed a "bubble" element for the first screen of the visualization. This bubble would assist the users in introducing the whole subject guiding them through our animated infographic of the first screen (some screenshots shown below).

This bubble was implemented entirely with javascript and jQuery. It works with a single function that receives the target coordinates, the dimensions, the text/html for the content, and the duration of the pause after the transition. While the transition function was simple, the complex part was figuring out a way in which we could call linked or ordered instructions for this bubble to be executed in the moments we wanted to without having to hardcode them all. Our solution was to define a JSON file with the requirements for each step and a create a function that would execute such instructions by calling itself recursively while there were elements left loaded from the JSON file.
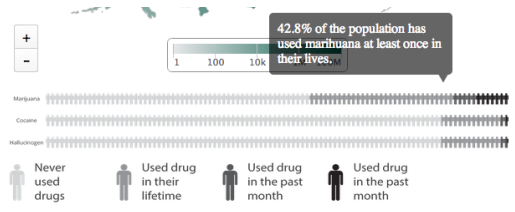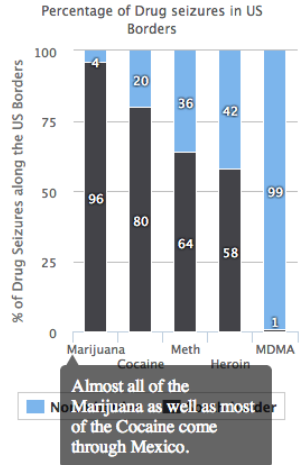
Task Delegation

Activity	Chalenge	Pablo	Marjan
Structure	25%	50%	25%
Data Collection	50%	25%	25%
Design	25%	25%	50%
Maps	75%	25%	0%
Graphs	25%	50%	25%
Wireframing	50%	50%	0%
Interactions	50%	50%	0%
User Testing	25%	0%	75%
Bubble	0%	100%	0%
Narrative	0%	50%	50%
Graphics	25%	25%	50%

References

http://en.wikipedia.org/wiki/Mexican_Drug_War#Effects_in_Mexico

http://visual.ly/super-charged-story-marijuana

http://www.unodc.org/wdr/

http://www.unodc.org/wdr/en/maps-and-graphs.html

https://stats.unodc.org/#state:0

http://www.justice.gov/dea/resource-center/statistics.shtml

http://www.justice.gov/dea/pr/multimedia-library/publications/drug_of_abuse.pdf#page=68

http://www.cnn.com/2013/09/02/world/americas/mexico-drug-war-fast-facts/

http://www.state.gov/j/inl/rls/nrcrpt/2013/vol1/204050.htm

http://www.secretariadoejecutivosnsp.gob.mx/es/SecretariadoEjecutivo/Incidencia_Delictiva_Nacional_fuero_comun