Process Book

Background and Motivation

Medical data is often visualized using charts and graphs. We looked around a number of bio-medical websites and we could only find static images with no or very little interactions. For the physicians/ researchers this is a bottleneck as they are not able to modify the graphs/data for context based analysis. Also the users don't usually have the knowledge of databases or even the datasets which are used in visualization. The objective of this project is to provide an interactive visualization system for the users using which they can visualize and change datasets, compare different datasets, export datasets to their reports. The application will also provide simple user interface to import new data sets.

The primary motivation for doing this project is that we didn't find any good visualization online for bio-medical domains. We also discussed with the users about their expectations from a visualization application. The general feedback we got from this exercise is that the existing applications are heavy and involves a learning curve for which non-technical users are reluctant. Our application will be a light weight visualization system with nominal learning curve which enables non technical users to visualize and personalize their data. We will also incorporate basic utilities on data like export, transform, import and personalize in our application with simple user interface.

Project Objectives

Most of the visualization tools in the medical domains lack ease of use, interaction and intuitiveness and personalization. The main objective along with learning and applying programming languages is to understand and improve the aesthetic and perceptive properties of the visualization. We chose to do this exercise with non-technical users (medical domain) which give us a better insight of how the visualization tools should be.

We classified our objectives into 2 groups; Data based objectives and Aesthetic objectives.

Data objectives:

1. Context based data selection: We will provide utilities to make context based data selection. The users will be allowed to select the filters of their choice.
2. Comparison: Data comparison views will be available for the user.
3. Personalized views (extended objective): Users will be allowed to personalize their views and select the view of their choice.

Aesthetic objectives:

1. Bring intuitiveness in the visualization: We will research over existing methods used in other visualization tools to bring self explanations. We will do our own analysis over the effectiveness of visualization and also consider publically available feedbacks on designs.
2. Ease of use: As stated above our application will have a nominal learning curve. We will ensure that the user interface is easy to learn and use.
3. Personalize cosmetic properties: We will provide flexibility to the users to change the properties like fill and text of the visualization. The tool will recommend the user to use certain combinations of color sets to personalize visualization along with the graphical views.

We hope to cover not just the technical but cognitive aspects of how the visualization tool should be like. We will learn multiple designs while researching on our design and also the user interactions. We also plan deploy it and get feedback from target users. This will be a good learning exercise and will give us more insights about an effective visualization tool.

Data

For this project we are visualizing flu trends of United States for a period of 2003-2004 to 2014-2015. The data is available on 'Google Flu Trends'.

Data URL : https://www.google.org/flutrends/about/data/flu/us/data.txt

Data Processing

The unwanted text from the data is removed manually.
Visualization data is converted into a master CSV file : 'flu_trends_data.csv' which contains data for all the dates , states and regions.
We created different data structures to save the datasets for different views

- `seasonsData` : stores season wise values for all the states for the selected year.
- `statesData` : stores data for individual states and it's major cities.
- `regionsData` : stores data for given regions
- `yearsData` : stores year wise data for all the states.

Exploratory Data Analysis

There are 2 main aspects of the data that we wanted to show, comparison and multi dimension data. We thought of several designs to show comparison. We tried multi series graph for showing comparison, it gave a very clear intuition of the design. But as the number states grew it became crowded. We decided to go with pie-charts and limited the number of state selections to 10.

Design Evolution

The following images describes the design evolution process.

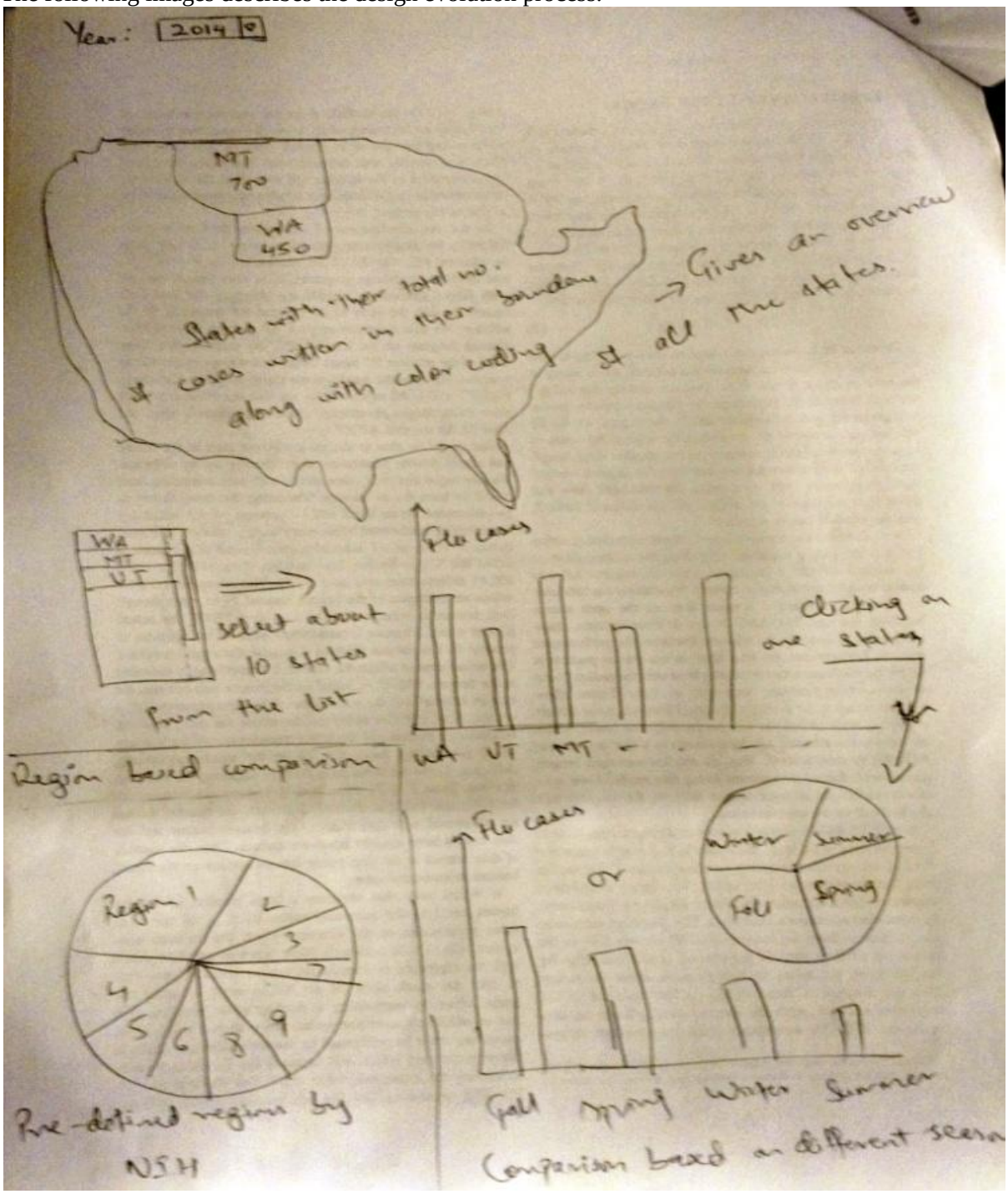


Figure 1: First Version of Design

Figure 1 shows the first version of the design. We started with basic blocks of visualization and kept on modifying keeping ourselves in place of the user. The following 2 figures shows how we landed up with the final design for this visualization.

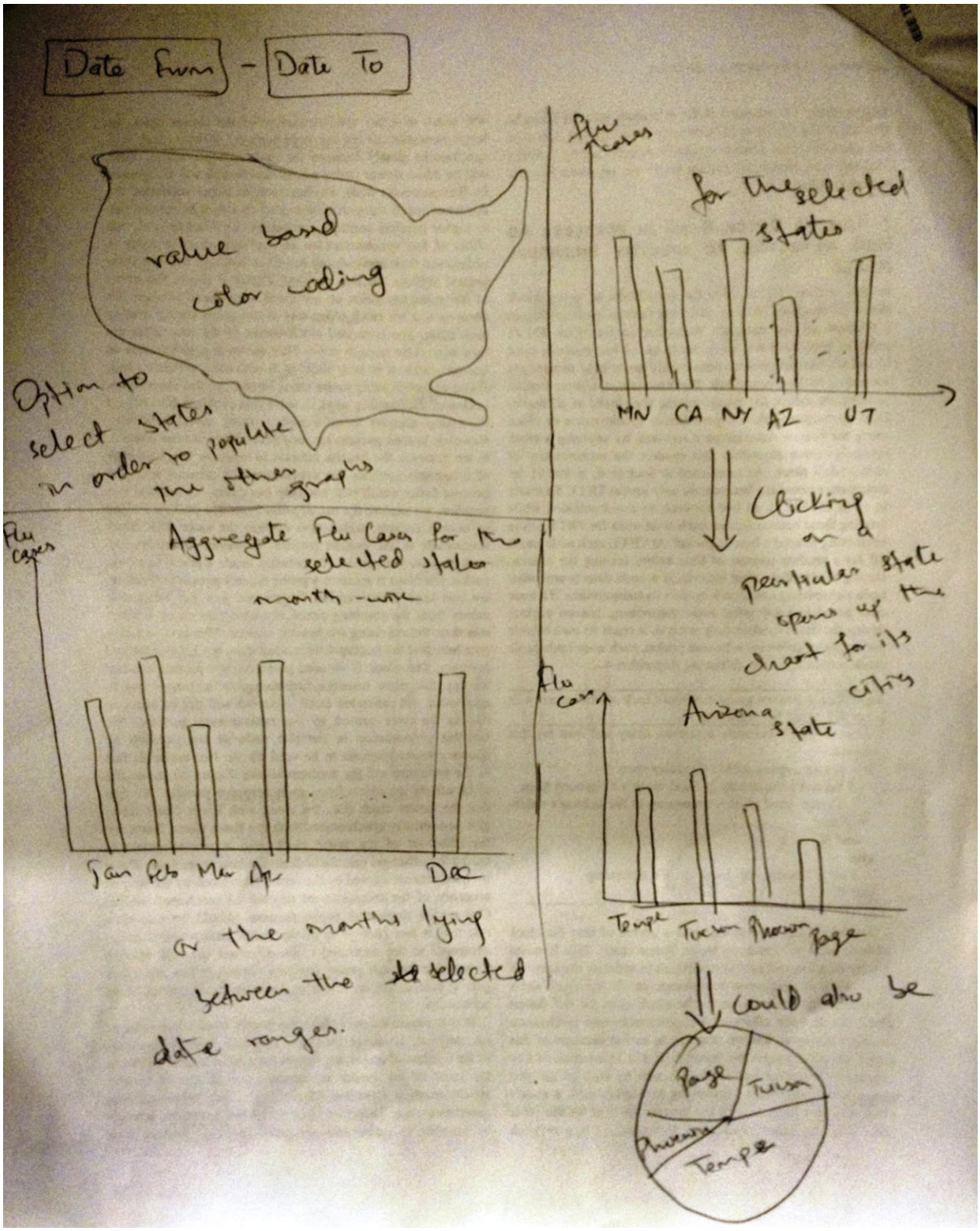


Figure 2: Second Version of Design

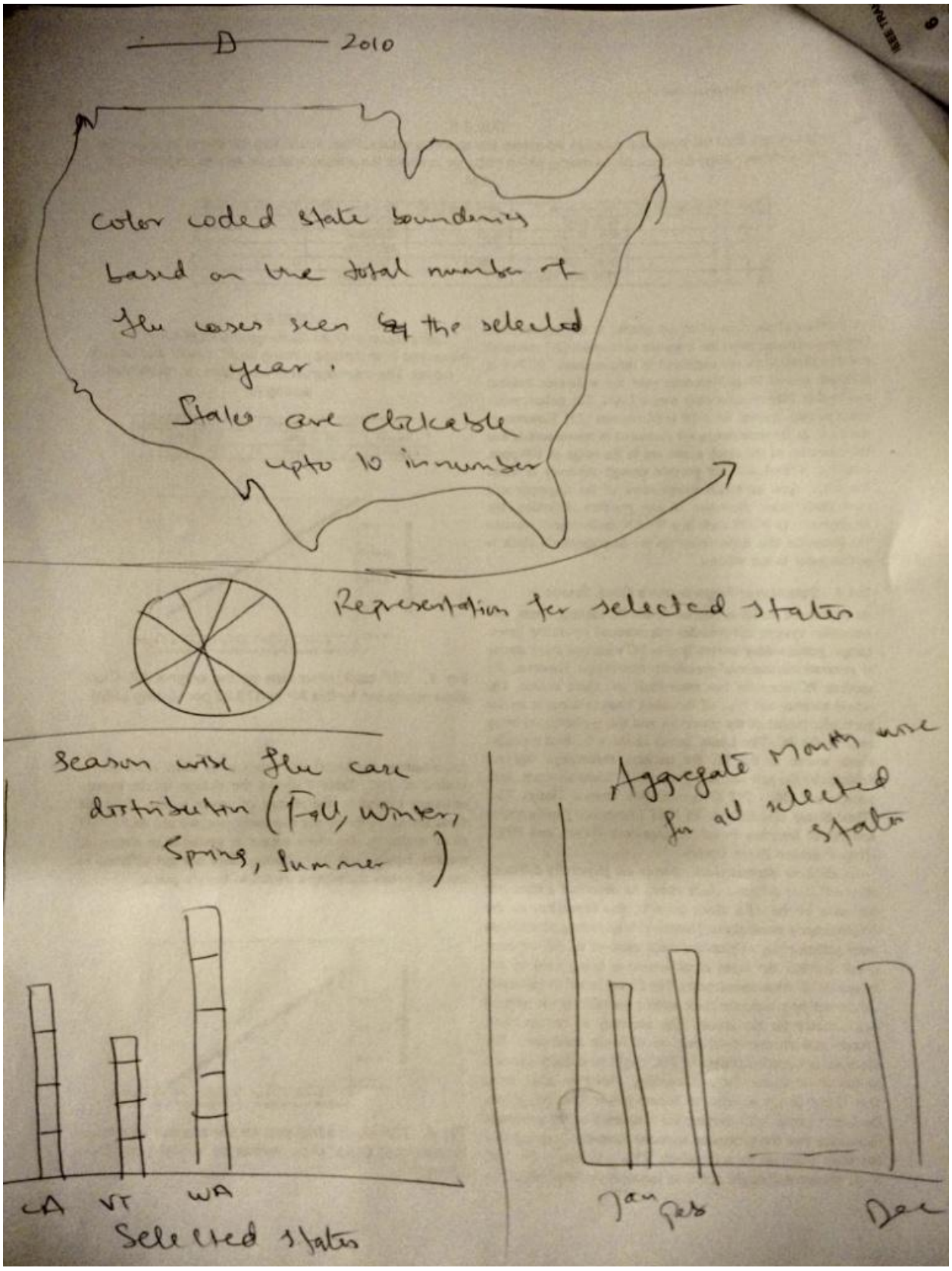
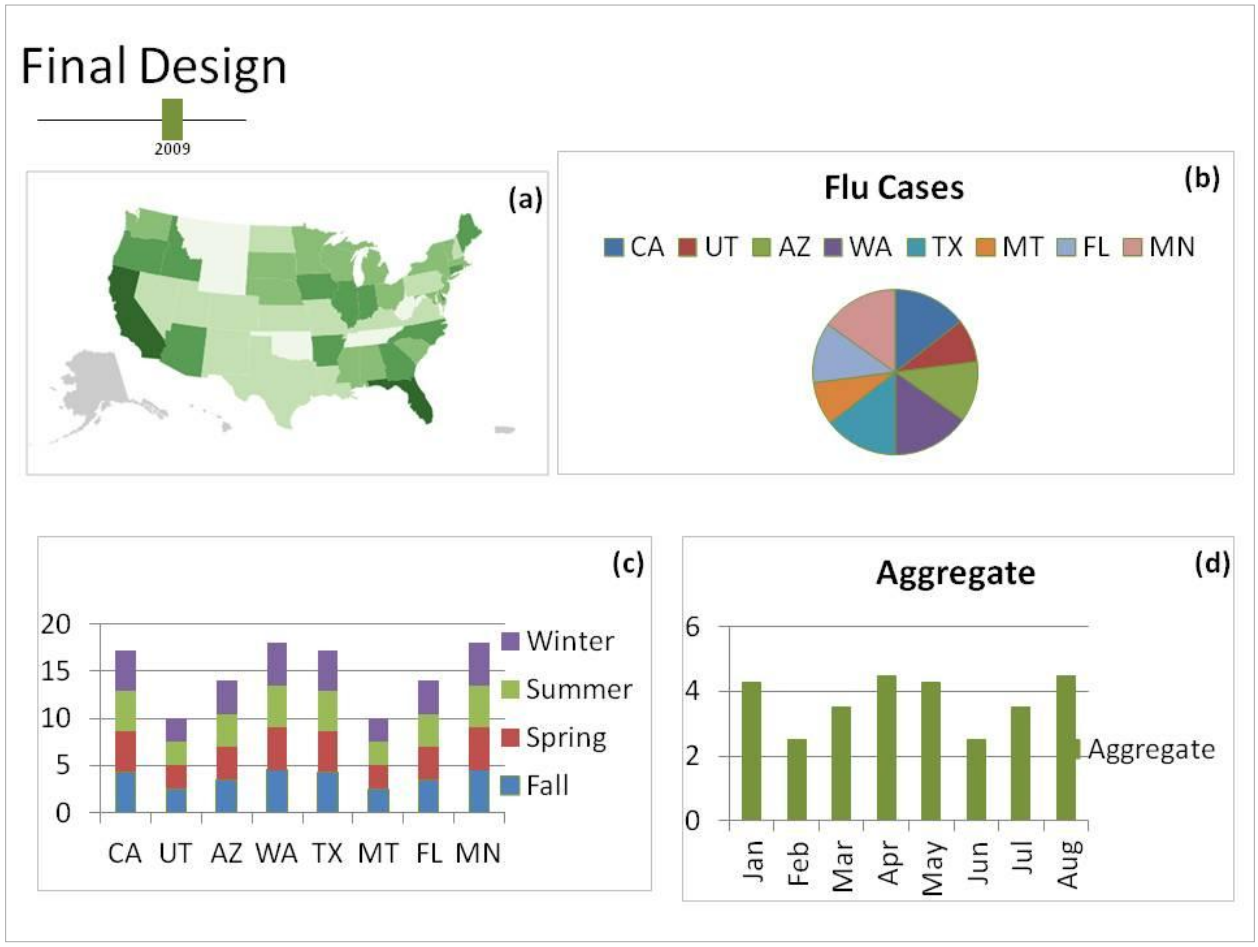


Figure 3: Final Design rough sketch

Selected Design



Design description

a) Map view with states ordered on the flu density for the selected year (default: 2003). User can select multiple states from the map.
b) This view shows a pie chart which better compares the flu cases in the selected states for the year selected using the slider.
c) This view shows the aggregated flu activity during seasons for the selected states. The Stacked graph shows the comparison of number of cases.
d) This view explores the data for the months of selected year. The flu activity of all the selected states is aggregated on the month and displayed in a bar graph.

Implementation

We designed the skeleton of our project.

Codes

- Structure of main page using HTML5 , CSS3.
- Created place holders for different views.
- Data processing and populating data structures
- Enabled slider interaction and updating the Monthly Reports View based on that.

Evaluation

We went through a number of options to design our tool. The scale of data is too large and we wanted to avoid any tabs or multiple visualization designs. In our data set there are multiple dimensions in which the data can be viewed. Years, months, seasons, regions, cities, states and their comparisons, if we had to design a visualization for showing all this information it would have been very difficult to keep it in one page. We prioritized our dimensions and designed based on the more generalized information which the user might find useful. We also limited the number of selections to be made for the states.

There are still improvements which could be made in our design. We are thinking of including as many dimensions in our visualizations yet keeping the tool simple.