CSE 578 - DATA VISUALIZATION
Project Proposal

VAST 2010 MC2: Characterization of pandemic Spread

Team members

Name	ASU ID
Sindhu Sree Aita	1220083085
Sucharith Reddy Desireddy	1219887383
Aishwarya Reddy Dwaram	1221604007
Madhura Ganga	1220012365
Vineethkrishna Vemireddy	1219499398
Avinash Senthil Kumaran	1222318630

Summary

Our task is to examine the pandemic spread in many countries during 2009. In this, we need to extract significant data like, how quick the pandemic is spreading, investigate the recurrent symptoms in all patients, death toll and different attributes of the pandemic outbreak. We are planning and executing some visualizations for preferable comprehension of the pandemic outburst.

Domain abstraction

Health department has given the data which aids in characterizing the disease's spread in specified countries. Analyzing the data provided by the health department will assist us in determining the severity of present or any other possible future pandemics and extracting statistics. We are provided the data which is spatial temporal, and after preprocessing it, then we visually design flow of data which would help the data realization simple for further research.

Data/Task Abstraction

As part of this challenge, the following tasks can be performed.

1. Examining the symptoms of patients of all age groups.
2. Death rate of the patients
3. How fast the pandemic is spreading
4. Number of cases for different age groups.
5. Visualization of the concentrations of cases and deaths.

And few other tasks can be performed based on the attributes. We are planning and executing a few perceptions for better comprehension of the pandemic episode.

Idiom abstraction

1. **Visualization of density of cases & deaths in various countries in global view**
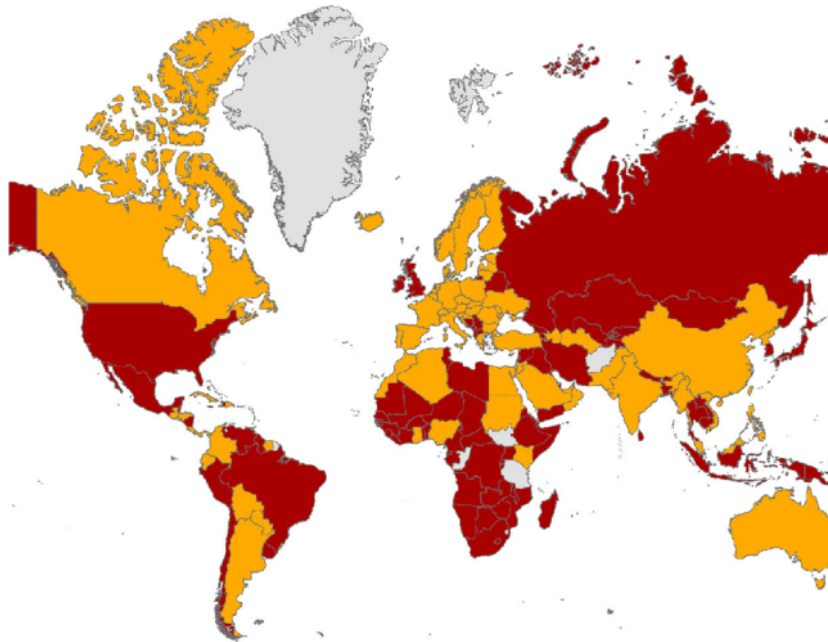

This visualization depicts the global view, which shows the density of the number of cases and death toll in different nations over time. This data will assist countries in taking the required steps to prevent a pandemic from spreading into their territory at the proper time. The mortality vs. cases comparison in various nations provides data on how effectively developed and underdeveloped countries can handle a pandemic.

**2. Analysing death rate across the countries using Tree Map**
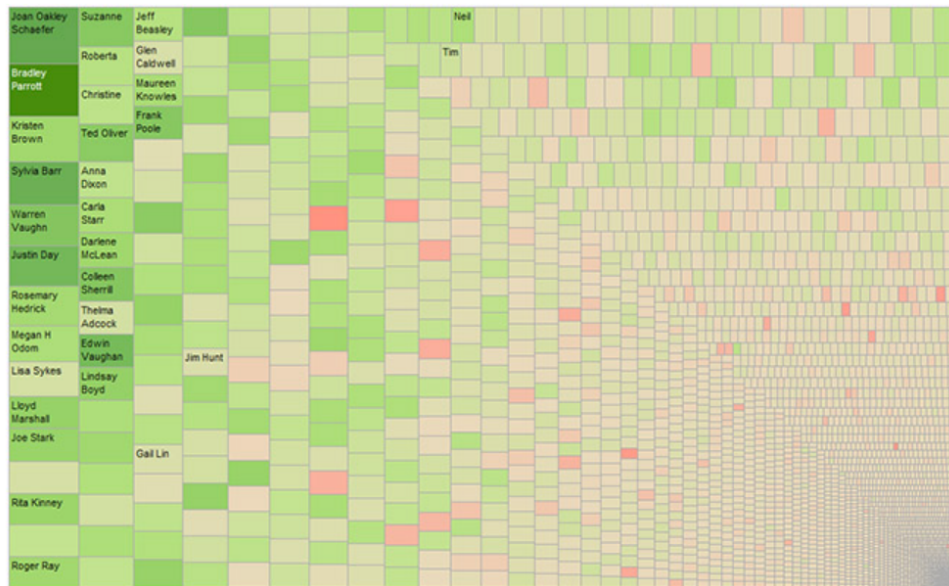
For each symptom, the above visualisation depicts the death toll across all the countries in the year of 2009. The countries are represented by the boxes. Symptoms are varied in the visualisation from left to right. Time varies from top to bottom. Death toll is represented by the luminance of the colour. Darker the colour, higher the death toll with that particular symptom.

3. Stacked bar visualization of age groups and mortality cases for certain countries
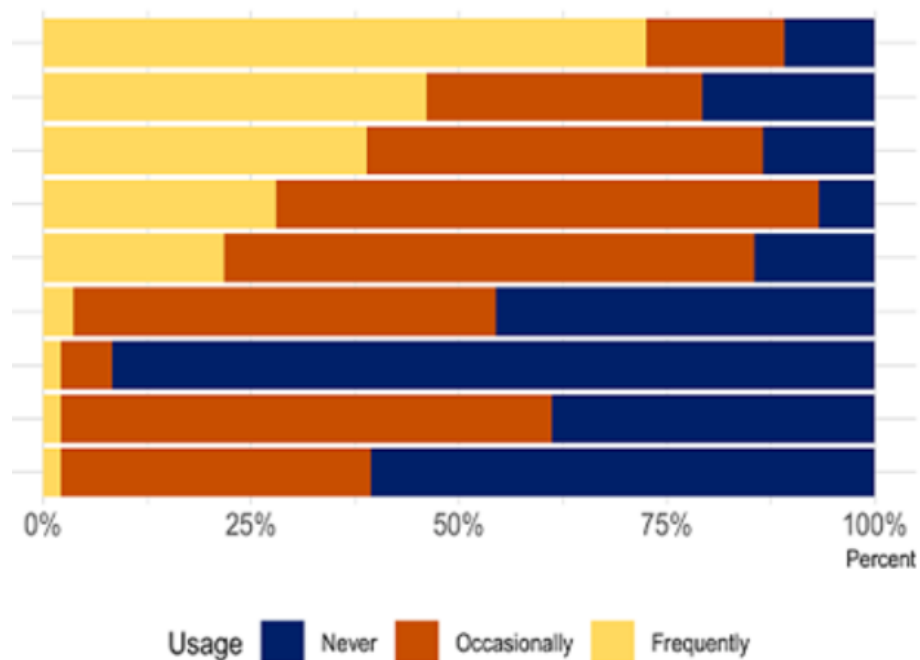
In a stacked bar visualization we would like to represent age groups and mortality cases for given countries across affected people. Each colour will represent an age group and on x-axis we have the percentage, by which we get the percentage of deaths in a particular age group. The length gives demographic data's distribution. Y-axis denotes the countries, we have a horizontal bar for each country.

4. Link chart visualization for death toll and time
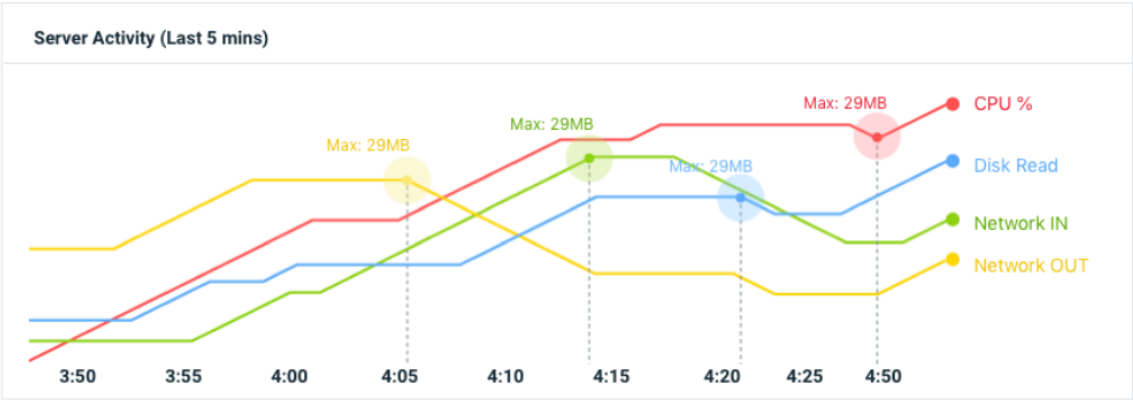
This line graph shows the temporal view. The y-axis of the line graph denotes the number of deaths, whereas the x-axis indicates the time. Each line indicates a country with its color hue. Researchers can use this line graph to analyze the death patterns during the pandemic. It also enables us to compare the effect of the pandemic between the countries. Marks are lines, and channels are x-position, y-position, and color hue.

5. Frequency chart of symptoms causing deaths in males and females (Innovative)
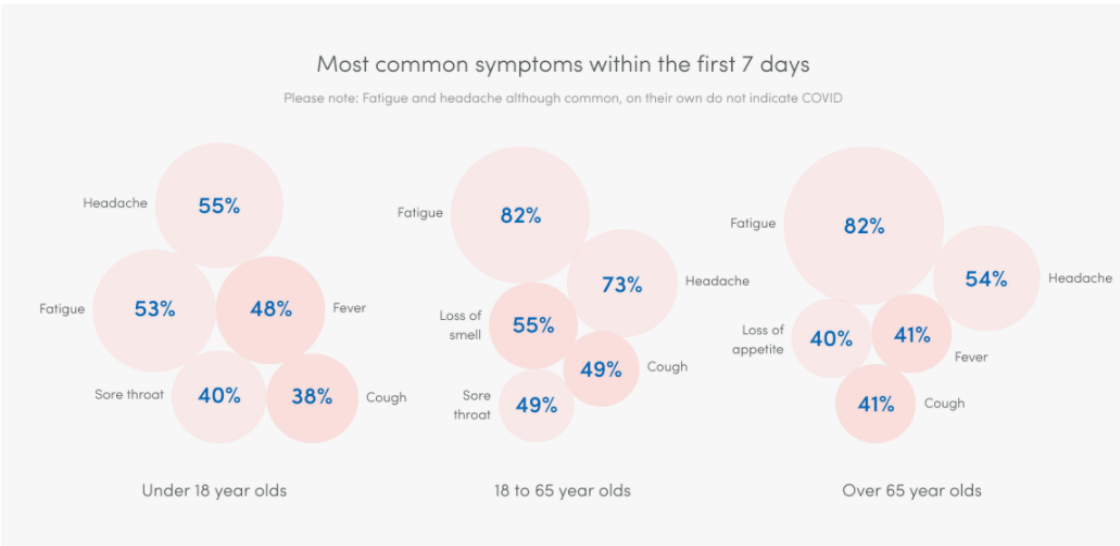
This visualization presents the common set of symptoms resulting to deaths in males and females. Each cluster is a frequent symptom set consisting of circles packed together, where each circle denotes a symptom. Clusters are created for each age group and gender. The circle area denotes the number of deaths, whereas the color luminance denotes the frequency of symptoms. This visualization gives a brief insight into the high-risk syndromes associated with gender and age group. Marks are bubbles, and channels are x-position(Age group, gender), circle area, Luminance.

Algorithm abstraction

The given raw data has to be processed first so we could use it for data visualizations. Python script will be used to derive meaningful information. For all the given 11 countries data, we will first categorize the SYNDROME column. Some syndromes mentioned in the dataset are similar, these will be assigned with familiar aliases which will make it easier to categorize or filter. Next step will be to give the proper required inputs for each of the data visualizations. For frontend we will use D3.js and HTML, CSS for web and styling.

Implementation plan

S No	Task	Week	Task allocated
1	Initial data preprocessing	1	Sindhu, Vineeth
2	Designing application structure	2	Sucharith, Aishwarya
3	Visualization - 1 & 2	3	Aishwarya, Sucharith
4	Visualization - 3 & 4	4	Madhura, Avinash
5	Visualization - 5	5	Vineeth, Sindhu
6	Testing and re-work on visualization	6	Avinash, Madhura
7	Visualization analysis, summary and report	7	Everyone
8	Project presentation	8	Everyone