**Team members:**

|  |  |  |
| --- | --- | --- |
| **First Name** | **Last Name** | **ASU ID** |
| Arun Raj | Deva | 1218410609 |
| Chandrahaas | Chintalaboguda | 1218686443 |
| Manikanta | Vankayala | 1218816326 |
| Pavan Kalyan Reddy | Thota | 1218436466 |
| Sai Teja Vishal | Jangala | 1218332037 |
| Vamsi Krishna | Kanagala | 1218608781 |

**Summarize:**

Our project analyzes the pandemic outbreak across several cities in 2009. As part of the project we want to extract important information like how fast the outbreak is spreading, analyze the common symptoms across all patients, mortality rate and other characteristics of the pandemic spread. We are designing and implementing few visualizations for better understanding of the pandemic outbreak.

**Domain abstraction:**

MC2 2010 VAST mini challenge is about a major outbreak that occurred in 2009 across the world. Health officials have provided us with information which will help them characterize the spread of the disease in these countries. Analysing this information given to us by the health officials now, will help us understand the seriousness of this or any future pandemic and extract statistics. Here we are given spatio-temporal data, and studying this data after its preprocessing, we visually create data flows that will make the realization of data easy for future study.

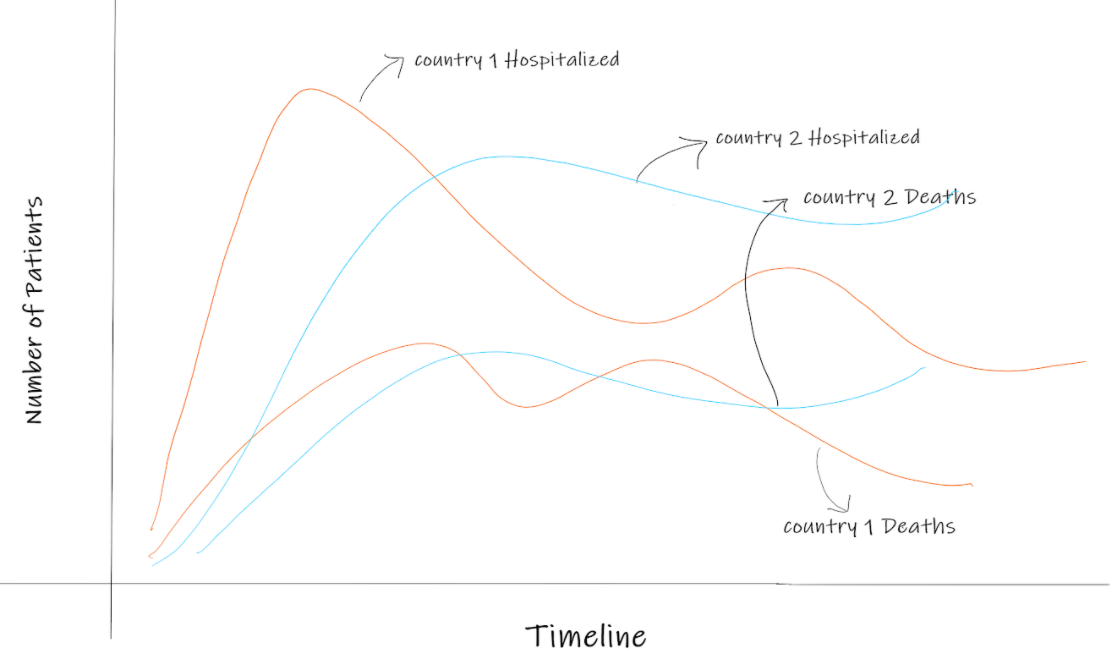
**Data/Task Abstraction**:

Health officials provided us with two excel sheets containing patients data for each location across 11 countries. These files contained several data points representing patient information and attributes with categorical, quantitative and sequential data. One of the excel sheets consists of patients id, age, gender, syndrome and date of admittance while the other one consists of date of death of patients with their id’s.

**Idiom abstraction (visualizations + interactions):**

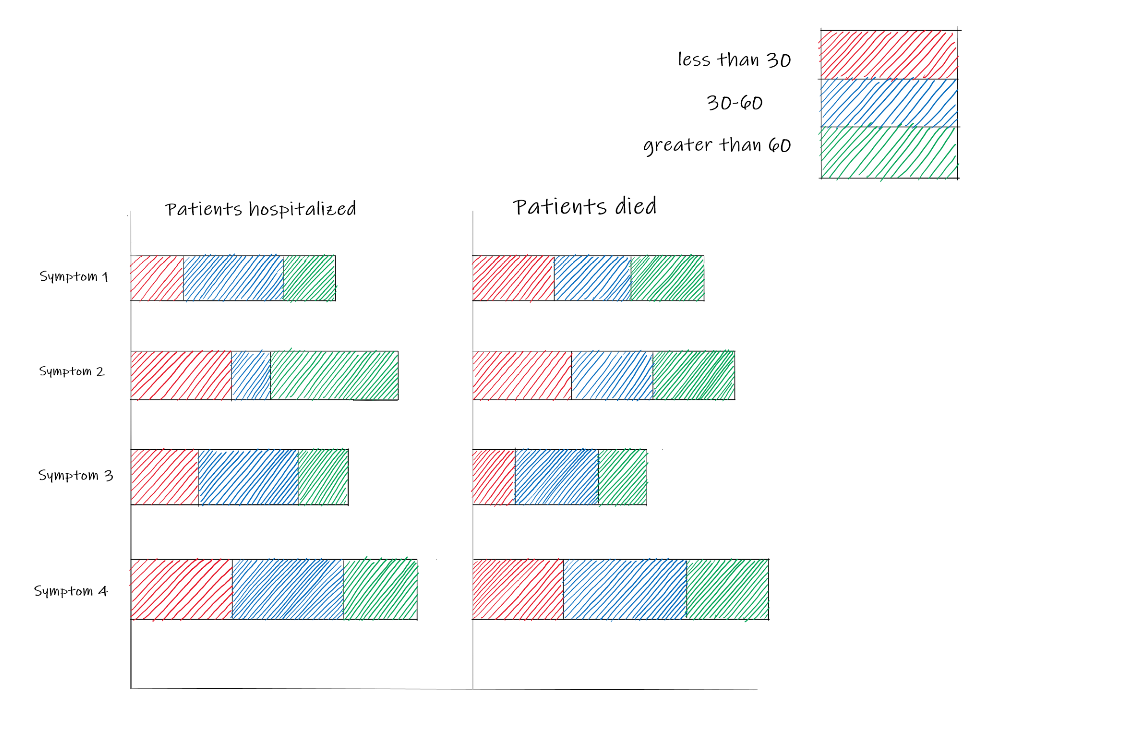
Using the data after the preprocessing, we will implement 5 visualizations to address the pandemic situation.

1. Country wise pandemic analysis.



The main aim of this visualization is to analyse the pandemic spread across all given countries. Each country has 2 lines, one line to show the total patients admitted across the given timeline and another line to show the deaths for the same country to compare the number of deaths to the number of affected people. In the same way we will plot the lines for all the countries where we can hover on one line to highlight the graph and show a tooltip with important information for a particular country and contrast the graph for other countries. In this way we will also have a scope to find any anomalies across countries.

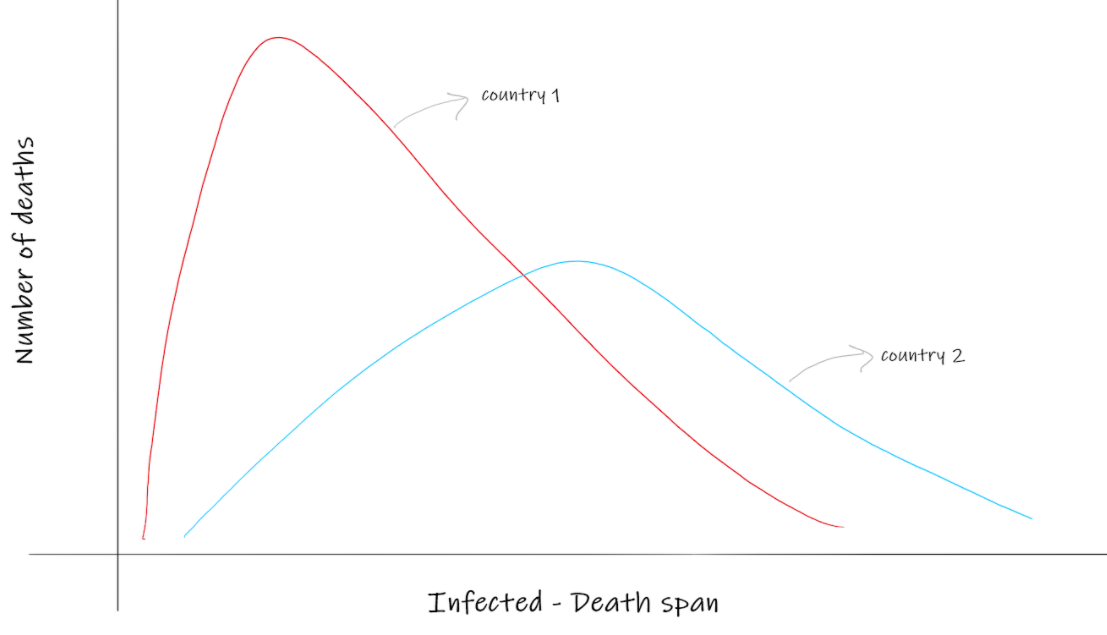
1. Symptoms analysis based on patients hospitalized and patients death with their age.



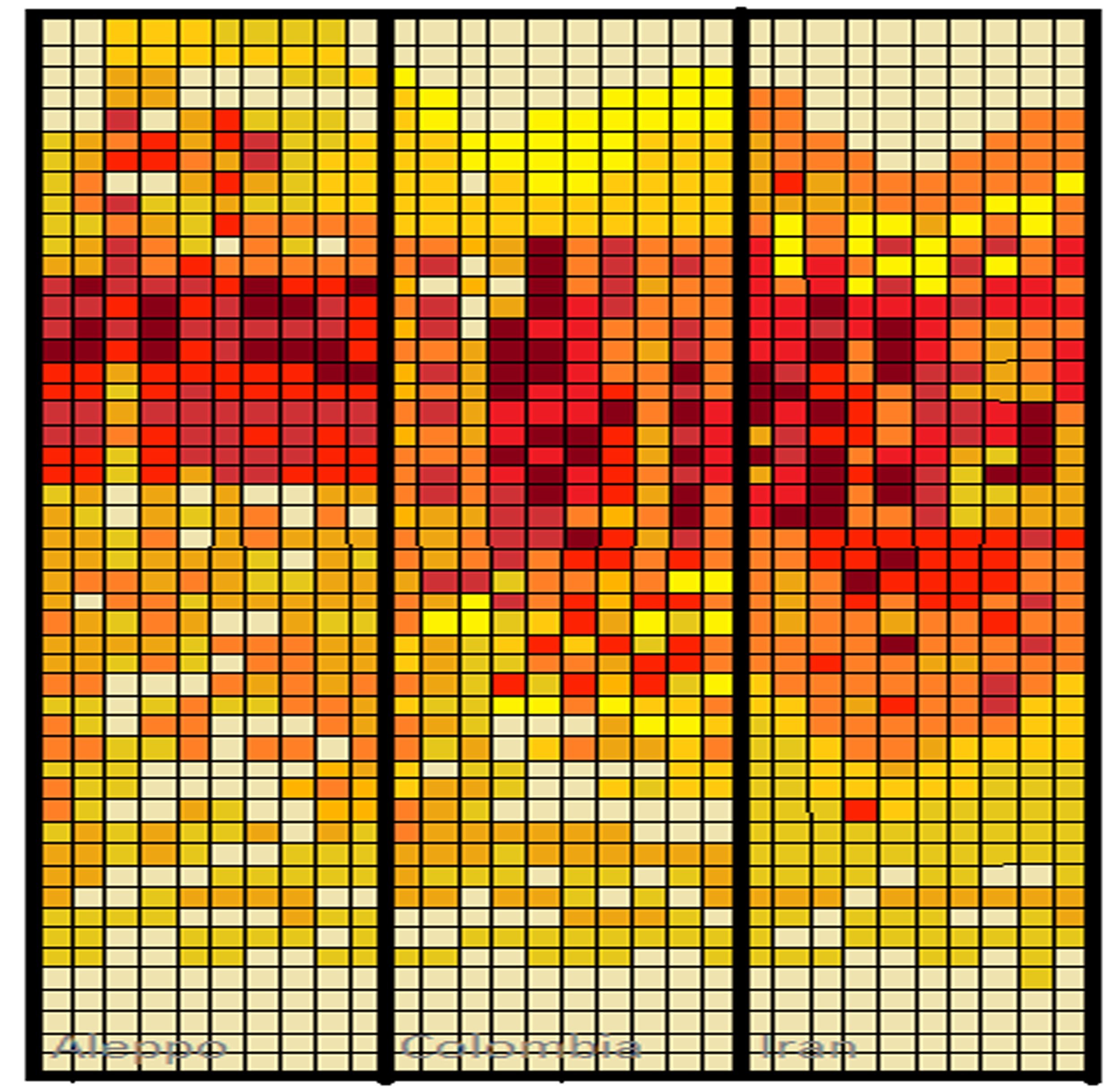
This visualization shows the cardinality importance of the symptoms analysed on different age groups. The above visualization is based on the percentage of the patients hospitalized. Also, the bar size for each symptom will vary based on the number of patients infected with the symptom. Example1: For symptom 1 the number of patients infected for age group 30 - 60 is higher compared to the other age groups, but the patients died because of symptom 1 in age group less than 30 is higher. From this visualization we can draw the cardinality importance of symptoms for each age group. Example2: From the size of the bars from the patients hospitalized graph we can compare which symptom is important compared to another symptom. Symptom 2 is highly affected compared to Symptom 1. This visualization can give important symptoms for each group and also the common symptom over all age groups.

1. Analysing the death span for every country.

The below visualization used here is a line graph with color hues for each country. The line represents the death span and number of deaths for every country. For example, as in the below graph, country 1 has a peak death span of say 5 days, meaning that patients in country 1 are dying fast compared to country 2. This also gives us to analyse the death patterns in different countries and find any anomalies. This graph also can tell the life expectancy of the newly admitted patients.

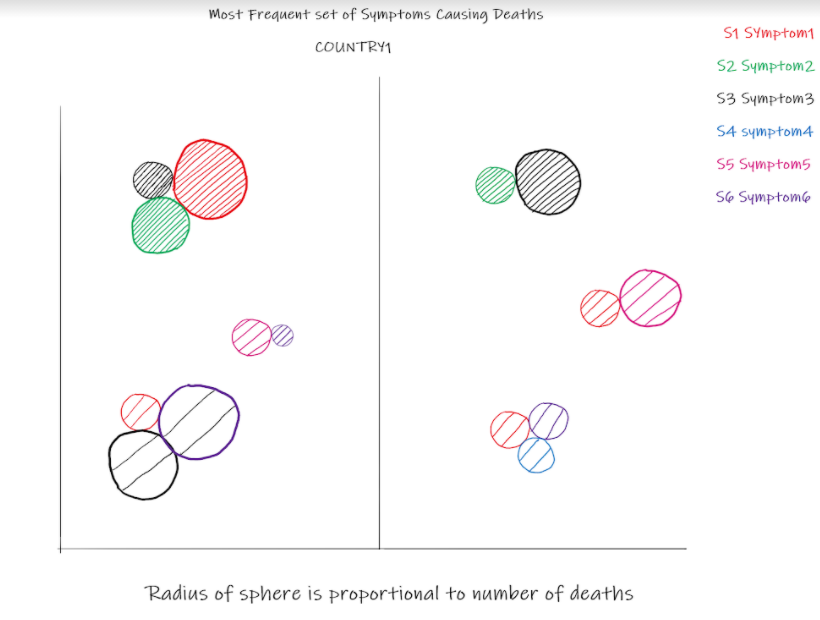


1. Tree map to analyse the mortality rates across all countries.



Above treemap visualization gives the information about the change of mortality rates over time (April 20-June 13) for each syndrome: Every section represents a country. Within each country, Time is mapped from top to bottom. Syndromes are mapped from left to right. As the mortality rate increases, the color luminance increases. Suppose if we observe column one in the above visualization, Middle part of it is much darker when compared to the other part of the column that means the mortality rate during that period is high and the people who died with each syndrome is higher.

1. **Innovative**: Analyzing the frequent set of syndromes



The Aim of this Visualization is to show the frequent set of syndromes causing deaths in the male and female. Every circle in the cluster is always together representing a frequent symptom set. The radius of the circle is proportional to the death counts and their color luminance is proportional to their frequency. We are going to find these syndrome patterns using the Apriori Algorithm. This visualization gives an insight about the high risk syndromes in male and female groups.

**Algorithm abstraction:**

We are using python for data pre-processing. The given raw data needs to be processed before implementing visualizations. Initially, we will analyse the “SYNDROME” column from the given data and then need to assign a common alias for the same type of syndrome. For example we found a few syndromes such as rash, rash on body, rash on hand which mean the same so, we want to bring all the similar syndromes under one common category before implementing the visualizations.

**Implementation plan:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Task no** | **Task** | **Week** | **Task assigned** |
| 1. | Pre-process the given raw data. | 1 and 2 | Arun, Manikanta, Pavan |
| 2. | Design and build the application structure. | 1 and 2 | Chandrahaas, Vamsi, Vishal |
| 3. | Develop the 3rd visualization. | 3 | Arun, Pavan |
| 4. | Implement 2nd visualization. | 3 | Vamsi, Vishal |
| 5. | Implement 5th visualization. | 4 and 5 | Arun, Pavan, Vamsi |
| 6. | Develop 4th visualization. | 4 and 5 | Chandrahaas, Manikanta, Vishal |
| 7. | Implement 1st visualization. | 6 | Chandrahaas, Manikanta |
| 8. | Test the developed visualizations and rework on any issues found. | 6 | Arun, Pavan, Chandrahaas, Vishal |
| 9. | Analyze to draw conclusions from the visualizations and work on the report. | 7 | Everyone |
| 10. | Presentation of project. | 8 | Everyone |