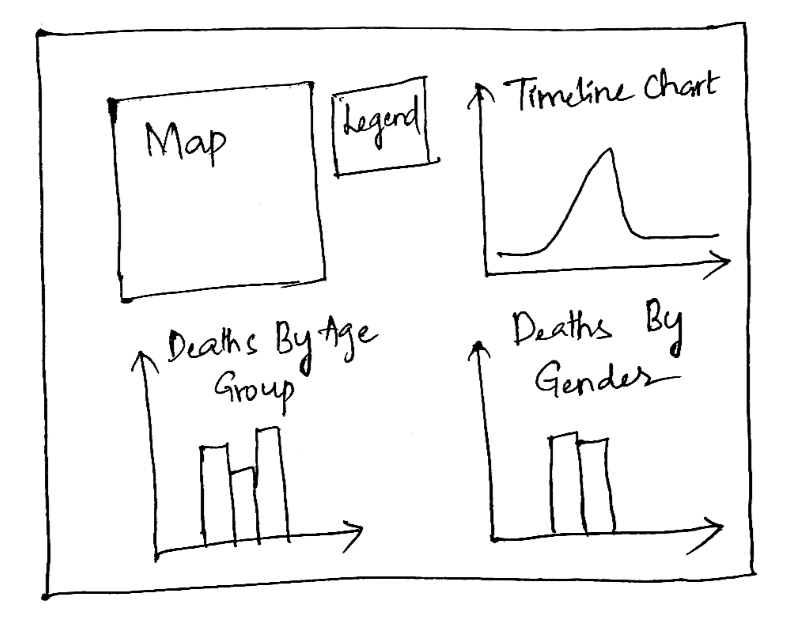
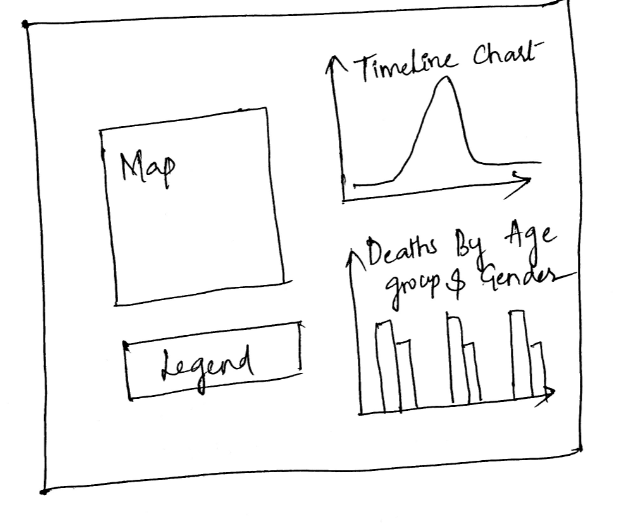
**Project 1 – John Snow Cholera Map using D3 - Documentation**

**Design Process:**

After going through the requirements, it is clear that we need 4 charts over all – Map, timeline chart and 2 bar graphs. So, I have initially started with this design.



The initial design consists of map with streets and pumps locations along with locations of deaths and a legend for the map on its right. To the right of it, the timeline chart. Below the map, a bar graph showing the number of deaths by age group. Below the timeline chart, another bar graph showing the number of deaths by gender. After implementing this design, I have noticed that, user have to scroll up and down to observe the deaths on a particular day in each of these charts, which can be selected by clicking on the data points on the timeline chart. To avoid this issue, I had to slightly modified the design as shown below.



I have slightly centered the map to the left and placed the legend below the map. To the right I have added the timeline chart and a grouped bar chart with number of deaths by age group and gender. I have found that this design is more effective than the previous one, so that user can select the data points on the timeline chart and both map and the grouped bar chart will be updated and it will be easier to observe the changes on both the charts, without having to scroll up and down. I found this design to be perfect to include zoom feature for the map, which otherwise would overlap with the legend on zoom. Later, I have also included a secondary axis below the timeline chart to select a time period to reflect on the timeline chart.

**Rationale of design choices:**

Map: For drawing the map, I have used data from streets.json file, which contains locations of the streets, pumps.csv file, which contains the locations of the pump. For drawing the streets, I have used SVG path elements. I have chosen circles to represent pumps and colored it in pink. I have picked the colors from the color brewer website, which are colorblind safe. I have tried and tested different colors and found that this color made it easier to locate the pumps and stands out from the other data points like the locations of deaths, which are also represented using circles, but with a smaller radius, in different colors. Locations of the deaths have been mapped using data from the file deaths\_age\_sex.csv. I have included two buttons, deaths by age and deaths by gender, which enables user to view the deaths on the map, where the circles representing locations of the deaths will be colored based on either age group or gender, depending on the selection. For color coding based on gender, I have selected two diverging colors in orange and purple tones from the color brewer website, which makes it easier for the user to identify two genders easily. For color coding based on age group, I have selected 6 diverging colors for each age group ranging from red to blue tones. I have also included names of the major streets and approximately mapped workhouse and brewery locations. I have also color-coded workhouse and brewery to make it easy for the user to locate them on the map. I have also included tooltips for circles representing locations of deaths, displaying age and gender information, which will be displayed on hovering over the data point. I have also included zoom and drag features for the map, which enables the user to drag or zoom in and out the map.

Timeline Chart – Date Vs Number of Deaths: To represent timeline chart, I have created a line graph using SVG path elements, with Date on the X axis and number of deaths on the Y axis. I have also added circles to represent each data point, so that it will be easier for the user to click on that data point to view deaths related to that particular day on the map and the grouped bar chart. The color coding on the map to represent deaths on a particular day is based on gender, to keep it simpler. User can always use tooltips to know the age group. I have also included tooltips showing date and number of deaths on that day, which will be displayed on hovering over the data point. I have also included a secondary x axis to select the time frame for the brush feature, which updates the timeline chart for that particular selection period.

Grouped Bar Chart – Deaths by Age and Gender: For the grouped bar chart, I have included Age group on the x axis, number of deaths on y axis and bars with 2 groups color coded, to represent gender. For color coding the bars, I have used the same colors which I have used in the map to represent gender in order to maintain uniformity. This grouped bar chart will be updated based on the user’s selection of data points on the timeline chart. This can be reset by clicking either deaths by age or deaths by gender button.

**Observations from the Visualizations:**

From the map, it is obvious that majority of the deaths have been recorded near the pump on the broad street and brewery locations. From the time line chart, we can observe that there is an increasing trend in the number of deaths till September 1st and a gradual decrease after that. Also, majority of the deaths have been recorded during the end of August and first week of September. From the grouped bar chart, we can observe that young children below 10 years and older age group people above 80 years were the most effected. Also, both genders are almost equally effected due to this epidemic.