**PROJECT 1**

**Hello, Stranger!**

**Introduction**

We spend majority of our time with strangers – in bus, in college, in the streets etc. How many of us take time to interact or show a pleasing gesture with a stranger we meet in public? Even a smile!

My project is an analysis of responses given by people towards strangers. This data collection has helped me analyze the main reasons as to why people tend to disregard fellow beings during various situations.

I collected the data, as a part of my daily routine, while travelling through different places and observing the behavior of random people towards me. I categorized them based on Age Group, Gender, Time and Location where I met them.

**Data Collection**

The data collection process took 6 days. My daily routine includes a visit to the park, university campus, office building and the library. I also go to church very often. I meet a lot of strangers at all these places every day. I have assumed few information like Gender and Age group of the people who were analysed.

The individual data points are recorded in the ‘Project\_Data.csv’ file and the ‘data.csv’ file contains the summarized data.

**Age group is divided into three categories:**

0 – 30 years

31 – 60 years

61 years and above.

**Time of day is divided into three categories:**

‘Morning’ – 07:00am – 12:00pm

‘After Noon’ – 12:01pm – 5:00pm

‘Evening’ – 5:01pm – 10:00pm

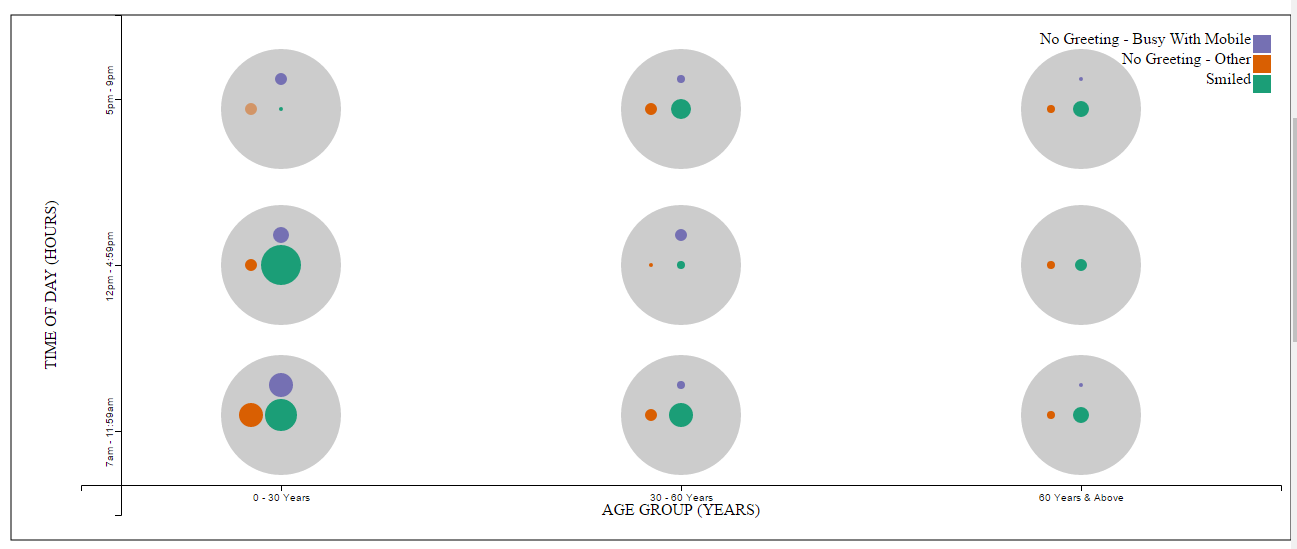
**Data Visualization**

Two different data visualizations have been created.

1. Variation of Scatter Plot incorporated with various other features including a donut chart
2. Donut charts displaying the overall count of people who did not respond to me as they were busy with Mobile Phones. This data has been categorized based on three conditions:
3. Age
4. Time
5. Gender

**Visualization 1:**

The below visualization shows the count of people who ‘Smiled’, ’Did not Smile due to use of Mobile Phones’ and ‘Did not smile due to other reasons’ based on 2 categories – the ‘Time of day’ when I met them and the ‘Age Group’.



Each bubble represents the number of each category (‘Smiled,’NotSmiledMobilePhone’,’NotSmiledOthers’) with the radius of the bubble being the total count of people in that category.

For eg:

Let us consider the 3 bubbles on the lower left corner of the visualization.

The green circle represents the people who smiled and it has a larger radius as the people who smiled during ‘7am – 11:59am’ and who are in the age group of 0 – 30 years was greater than the people in the same category who did not smile and the ones who did not smile as they were using mobile phones.

The data for this is loaded from the data.csv file which is then parsed in the code. This csv file contains a summary of the entire data that was collected as this shows the final total counts under each category.

**Visualization techniques implemented:**

1. The data has been visualized by incorporating features of scatter plot, bubble chart and donut chart.
2. On mouse hover, the bubbles show the details.
3. On clicking the outer circle (grey circle) which contains the 3 circles, a donut chart appears which displays the actual count of each category.
4. When this donut is displayed, the other details in that circle are not displayed.
5. On clicking inside the donut, the old data is displayed again as 3 circles.

**Visualization 2 : Final Analysis of Data Collected**

The second visualization in the form of Donut chart shows the number of people who were busy with mobile phones, based on various categories like Age, Time and Gender. The data has been provided to display the data from the ‘Project\_data.csv’ file. This file has the original data points that were collected.

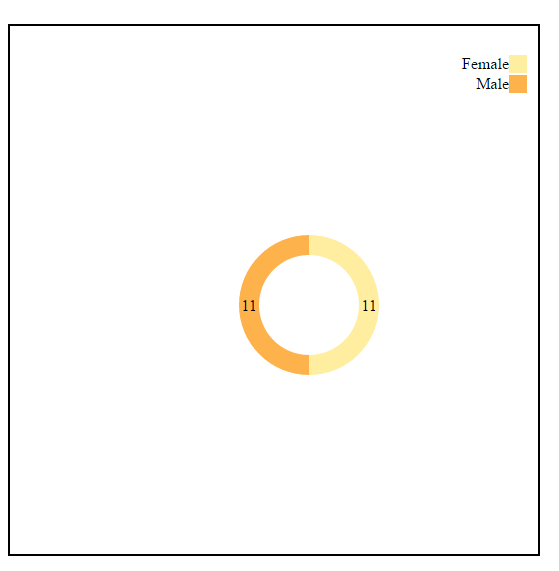
From the data that has been collected, it can be assumed that the people in the age group of 0 – 30 years tend to be more held up by mobile phones (this is an analysis which is just based on this data that was collected).

**Visualization techniques implemented:**

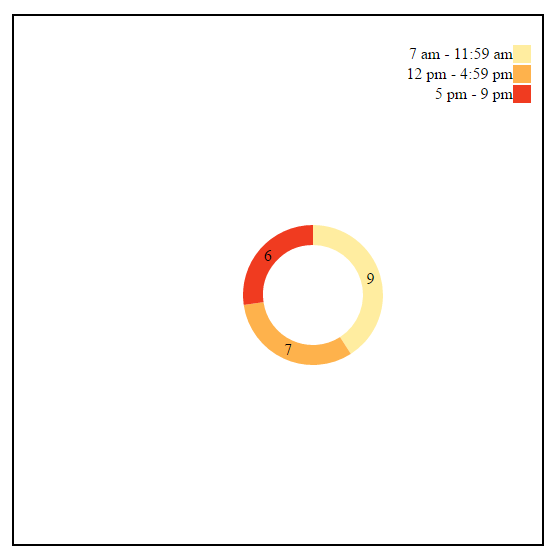
1. Donut chart with transitions and animations

The following screenshots display a summary of the charts:

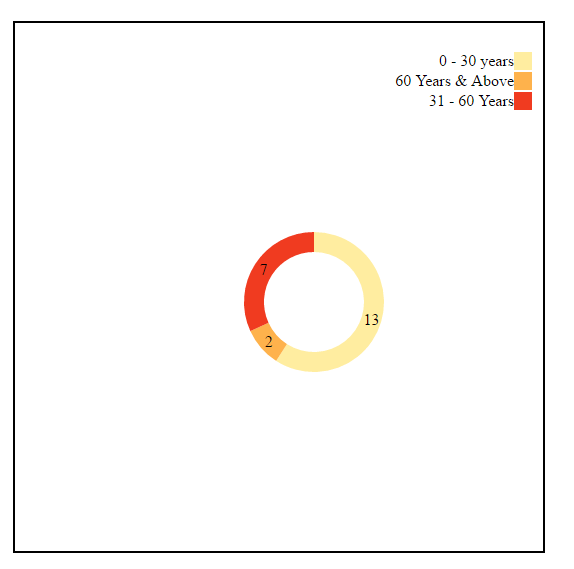
**Based on Gender:**



**Based on Time of Day:**



**Based on Age Group:**



The three donut charts appear as an animation and thus only one would be visible at a time.

Modification to Existing Techniques:

1. Normally in a scatter plot, the axis is numerical or just one axis is categorical. But in this case, both the axis are categorical which makes the visualization even more complex.
2. The idea of bubble chart has been incorporated with Scatter Plot.

References:

1. Scatter Plot : <http://bl.ocks.org/weiglemc/6185069>
2. For incorporationg various features, various links from **bl.ocks.org/mbostock** have been referred.