1. Data Input (Manufactured)

The **SubdivisionPersons.csv** file is the source data file. Which will be used for subsequent analytics and Train / Test sets.

It has **17318 rows** that are randomly generated. These will be divided as 75% train set, 15% test set and 10% validation set.

It has following columns:

1. DistCode
2. District: 23 districts in West Bengal
3. Subdivision: 70 subdivisions
4. PersonName: random character set / placeholder
5. Gender
6. Age
7. WorkerType
8. RegisteredYesNo: as of 2018
9. Registered2017
10. Registered2016
11. Analytics Applied

**SR002: Choropleth for percentage of beneficiaries:**

3 choropleth maps are generated to indicate the percentage of beneficiaries over the years 2016, 2017 and 2018 across districts of West Bengal. These show the % of beneficiaries as

number of existing SSY beneficiaries in each district X 100

total beneficiaries as per eligible population of that district

This has been achieved by grouping the total beneficiaries and the registered beneficiaries in each year as per districts from the SubdivisionPersons.csv file. This is done inside pandas data frames in python script file GenerateAnalytics.py.

Accordingly, 3 choropleth maps of West Bengal, one each for years 2016 to 2018 are generated as html files. These show the percentage of beneficiaries over the years 2016 to 2018 across districts of West Bengal.

**SR003: Balloon popup details:**

Every district will have a balloon pointer that can be clicked to see the details. When a user clicks a balloon, it shows the name of the district as a hyperlink. When the hyperlink is clicked, the details analytics page for the district will be shown. Currently in the demo, the detail is shown only for the district Dakshin Dinajpur. For rest of the districts, the detail is short circuited to SSY or company home page.

Following 2 types of analytics is shown for the district Dakshin Dinajpur as part of this demo in the details page.

1. **Details diagram 1:** Histogram of age of registered users in the year 2018 in Birbhum district overlaid by distribution of population over age in the entire state for the same year. The bins for the above distributions are chosen from ages 18 to 65 years, every 10 years apart.
2. **Details diagram 2:** Bar diagram showing male Vs female registration percentages in SSY scheme over the years 2016 to 2018 in the district of Dakshin Dinajpur. This is overlaid by a line diagram showing the overall registration percentage for both genders in the entire state.

Artificial Intelligence Based Prediction

1. Demo System Features

User should he able to register to the demo website.

User should be able to login to the website.

User should be able to access account page using the website.