How to run the program:

The entire code for the program has been written using python 3.9.6. The codes have been broken into modules namely:

- APK: This module creates a GUI and likewise interacts with other modules to produce the charts
- preprocessing module: Cleans the datasets and extracts additional features from the datasets
- Visualization module: This module contains functions that generate visual charts from the dataset.

The entire program run**s fully** from the **APK module.**

 To run the entire program, run the APK module either from the command prompt terminal (by double clicking the APK module) or from the python terminal. The charts will Open in a new window, close each figure to generate the next figure.

Dependency Information:

The dependency library used are: pandas, matplotlib and pywaffle. These libraries can be installed on your local system by using the pip command in the command prompt terminal

How to run the unit tests:

The Jupyter Notebook in the root directory of the Source folder contains the unit testing codes for the Visualization functions.

Details of charts produced:

1. The Bar graphs below shows the Top 5 areas with the highest number and top 3 areas with lowest number of cases alongside age distribution:
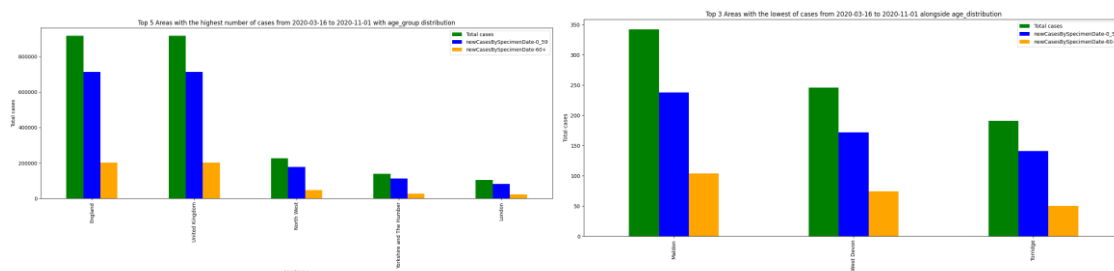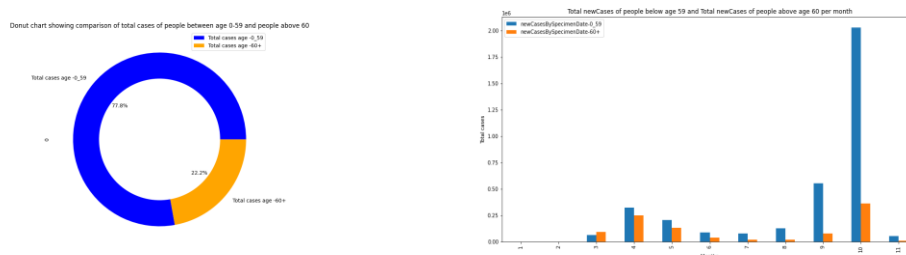

- ✓ X-axis: AreaName [Names of the areas in the datasets]
- ✓ Y-axis: Total Cases [Total cases observed in the respective area through the dataset]
- ✓ Description: The graphs shows that England and Uk recorded the highest cases while Torridge the lowest and people of age 0-59 contributed significantly in all areas.

2. A Pie chart showing the comparison of total cases by month and a bar chart showing the comparison of total cases through the month of october, the month with the highest recorded case:

✓ X-axis: Months/ Total Cases [The Months and Total Cases recorded through the dataset]
✓ Y-axis: Area
✓ Description: The graphs shows that the month with the highest number of cases is 10(October); with more than 50% of the total cases recorded in that month and that the top 5 areas with highest cases that month also contributed significantly (> 50%) of total cases that month.
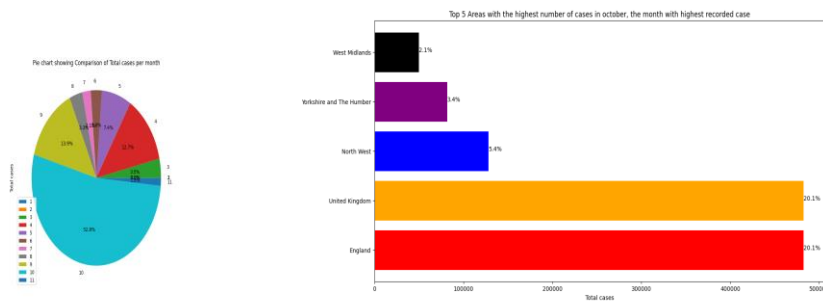

3. A Donut chart and a Grouped bar chart showing the distribution of the cases by unique age range (those below 59 and those above 60):

✓ X-axis: Months [The Months cases were recorded through the dataset]
✓ Y-axis: Total Cases [Total cases observed in the respective Months through the dataset]
✓ Description: While the pie chart shows that majority of the total cases recorded were from those between age 0-59, The grouped bar graph shows that distribution across each month.
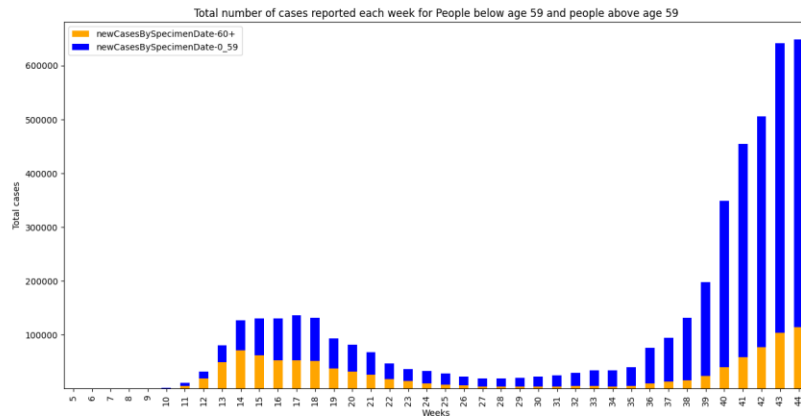
4. Graphs below shows the distribution of cases per Day and Week:
✓ X-axis: Dates/weeks [The Days/Weeks cases were recorded through the datasets]
✓ Y-axis: Total Cases [Total cases observed in the respective days/weeks through the dataset]
✓ Description: The graph shows that a first spike occurred between 22-04 and 2020-06 and a
    second spike much greater than the first occurred from 2020-10 through 2020-11.
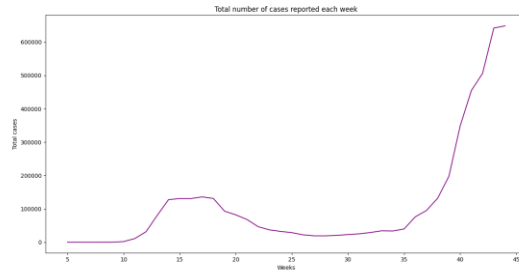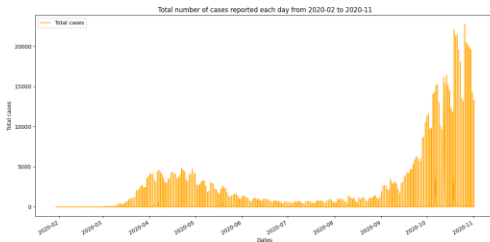

5. Stacked Bar-Graph showing the Total case distribution of age groups below and above 59 across
   the weeks.

    Y-axis: Weeks,   X-axis: Total Cases [Total cases observed in each week]

Total number of cases reported each week for People below age 59 and people above age 59

6. Bar graph showing total number of cases recorded for each age groups:
✓ X-axis: Age-Groups [The age-range of people as grouped in the dataset]
✓ Y-axis: Total Cases [Total cases observed for each age-range through the dataset]
✓ Description: The graph shows that age-group 20-24 contributed the most to total recorded cases while ages 5-9 contributed the least.

Total Number of cases recorded for each age group
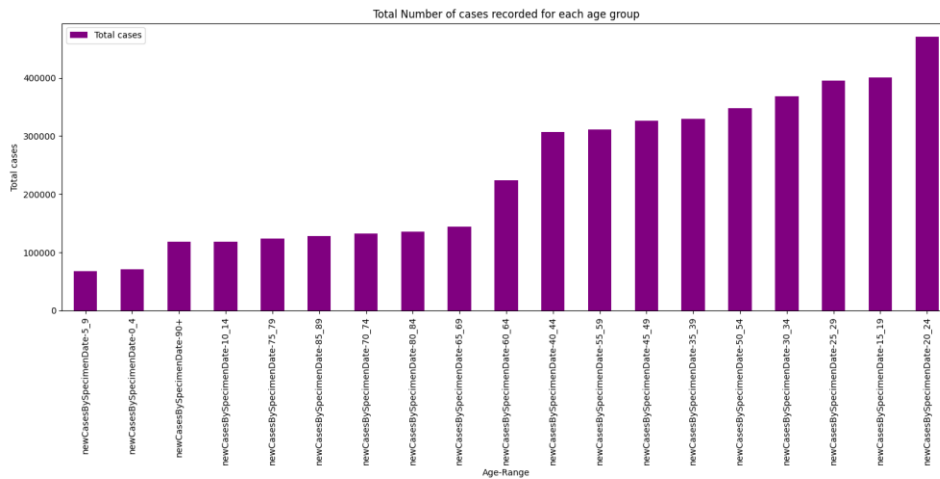
7. Waffle chart showing proportion of each age-group contribution on the 10th Month; the month with highest recorded cases. The chart reveals that age-group 20-24 and 15-19 contributed most significantly to the recorded cases that month.
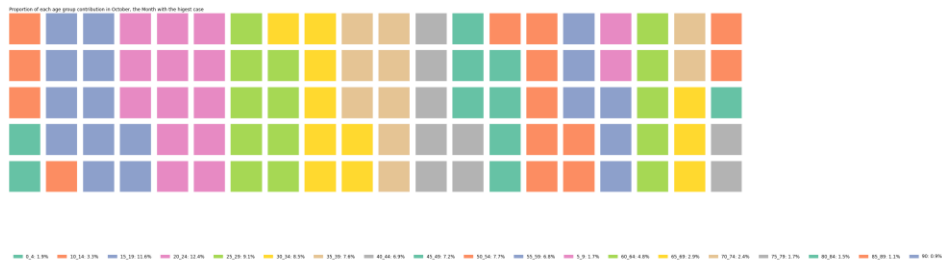
8. Graphs showing total number of cumulative cases by rolling rates and sums through the days: