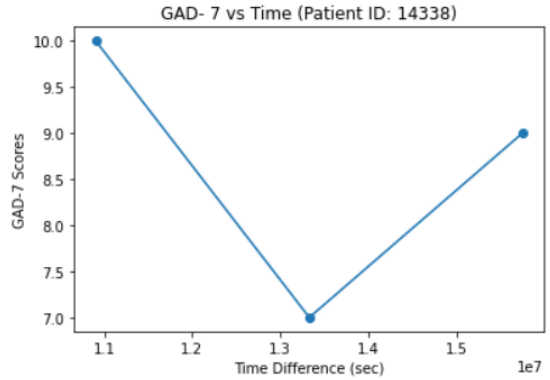
Part 1:

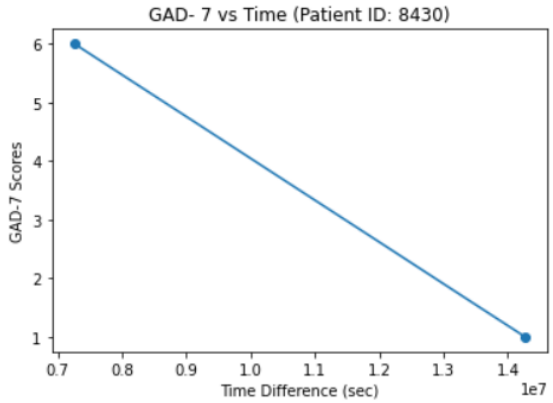
The given dataset consists of 53689 readings of GAD-7 scores (a measure for generalized anxiety disorder) measured at different time intervals for 15502 patient IDs. The dataset can be visualized individually as below for each patient record to track their progress over time. This would help the clinicians and mental health providers to generate diagnostics for their patients through this assessment.

For each patient, a line graph was plotted with the GAD-7 scores in the y-axis and the date of the test taken in the x-axis. The user is required to enter the specific patient ID to generate the corresponding plot.

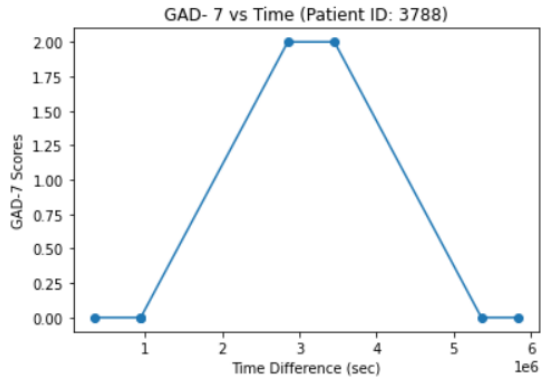
```
Enter patient_id: 14338
```



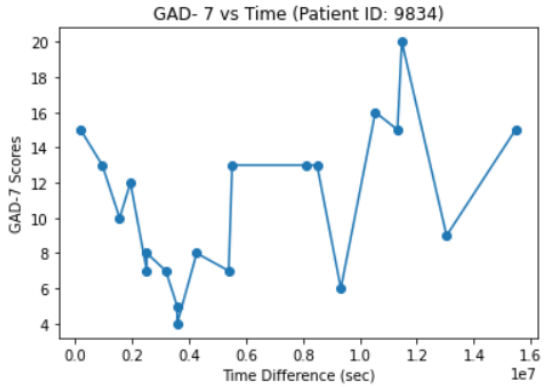

```
Enter patient_id: 8430
```



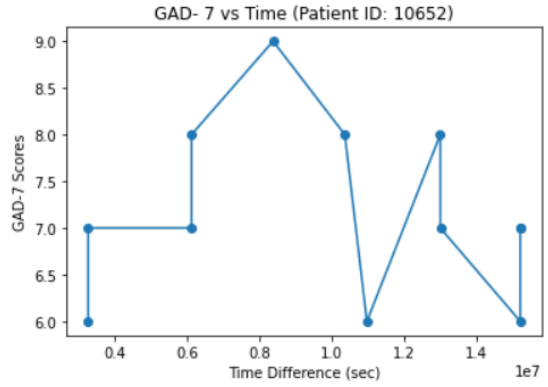

```
Enter patient_id: 3788
```



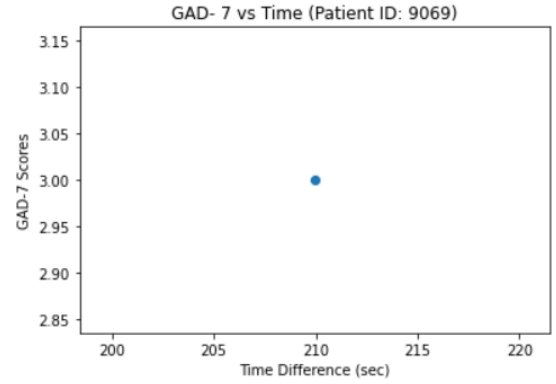
```
Enter patient_id: 9834
```



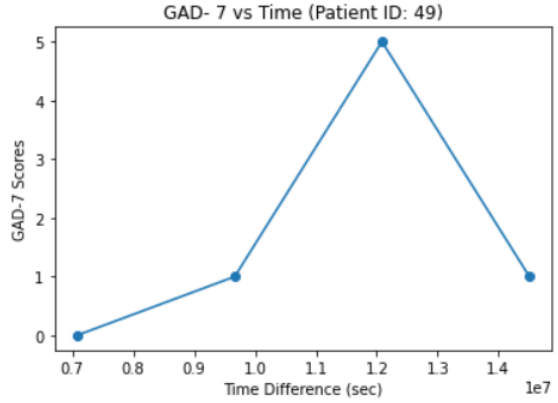

```
Enter patient_id: 10652
```




```
Enter patient_id: 9069
```




```
Enter patient_id: 49
```



From the plots (and plots obtained for other patients apart from the above), one can infer that each patient has a different history of testing. Some patients have been taking the test for quite sometime and the other are beginners of the GAD-7 test (some patients like patient_id = 9069 have just started the test). Also, the pattern of the plots is distinct for each patient (the number of peaks and troughs and their order in the graphs differ).

One can assume that external factors are needed to analyse these test patterns. Some of the factors that would be useful to study the results are:

- type of therapy the patient is undertaking
- age
- BMI
- health conditions/history of medical illness (hypoglycemia, stroke, hyperthyroidism)
- alcohol/drug consumption
- recent trauma
- environmental factors (stress due to work, academics, family, friends)

Also, with the external factors, machine learning models can be trained as an additional application to predict scores by receiving patient details as inputs.