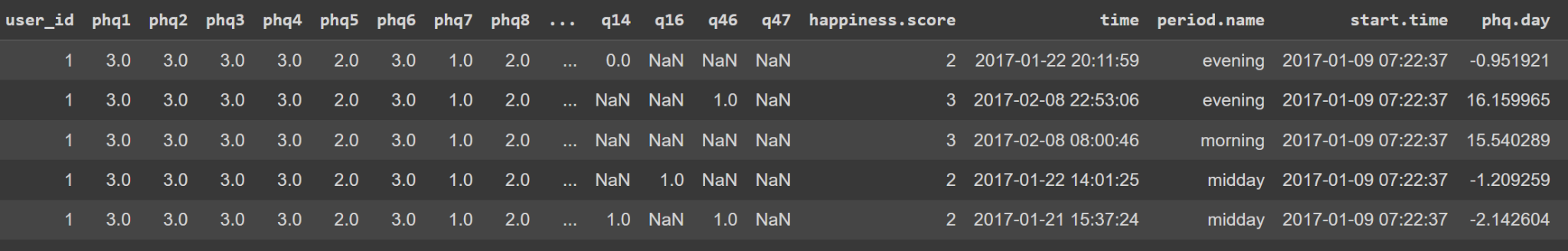
# PHQ-9 Depression Assessment

In this part we take a PHQ-9 Depression Assessment dataset- [data](https://www.kaggle.com/datasets/thedevastator/phq-9-depression-assessment)  and try to predict the happiness score based on the assessment. This dataset comprises of ambulatory assessment (AA) data taken over a span of 14 days. The dataset consists of attributes related to depression symptoms and mood ratings, as well as results from a retrospective Patient Health Questionnaire (PHQ-9) which is curated for finding out the symptoms or occurence of depression in individuals. Moreover, it also contains demographic information about the participants such as their age and gender. It contains more than 3000 datapoints.



This task can be viewed as a tabular classification task where we can leverage the tubular techniques like decision models in order to predict the happiness score of people based on the demographics and the PHQ 9 assessment. We use the label encoder to convert the string labels into integer classes so that it can be fed into the models for training. Then we move on to build our model for predictions. We make use of one of the most popular models, the Decision tree classifier model

## 

## CODE:

## Installation

Pip install pandas sklearn

## Imports

Import pandas as pd

from sklearn.preprocessing import OneHotEncoder, LabelEncoder

from sklearn.tree import DecisionTreeClassifier

from sklearn.model\_selection import train\_test\_split

## Data Loading

df = pd.read\_csv('Dataset\_14-day\_AA\_depression\_symptoms\_mood\_and\_PHQ-9.csv')

X= df.drop(columns=['time','start.time','phq.day','happiness.score'],axis=1)

y= df['happiness.score']

## Data processing

X = X.apply(LabelEncoder().fit\_transform)

le = LabelEncoder()

y = le.fit\_transform(y)

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.33, random\_state=42)

## Model Building

bst = DecisionTreeClassifier()

bst.fit(X\_train, y\_train)

bst.score(X\_test,y\_test)