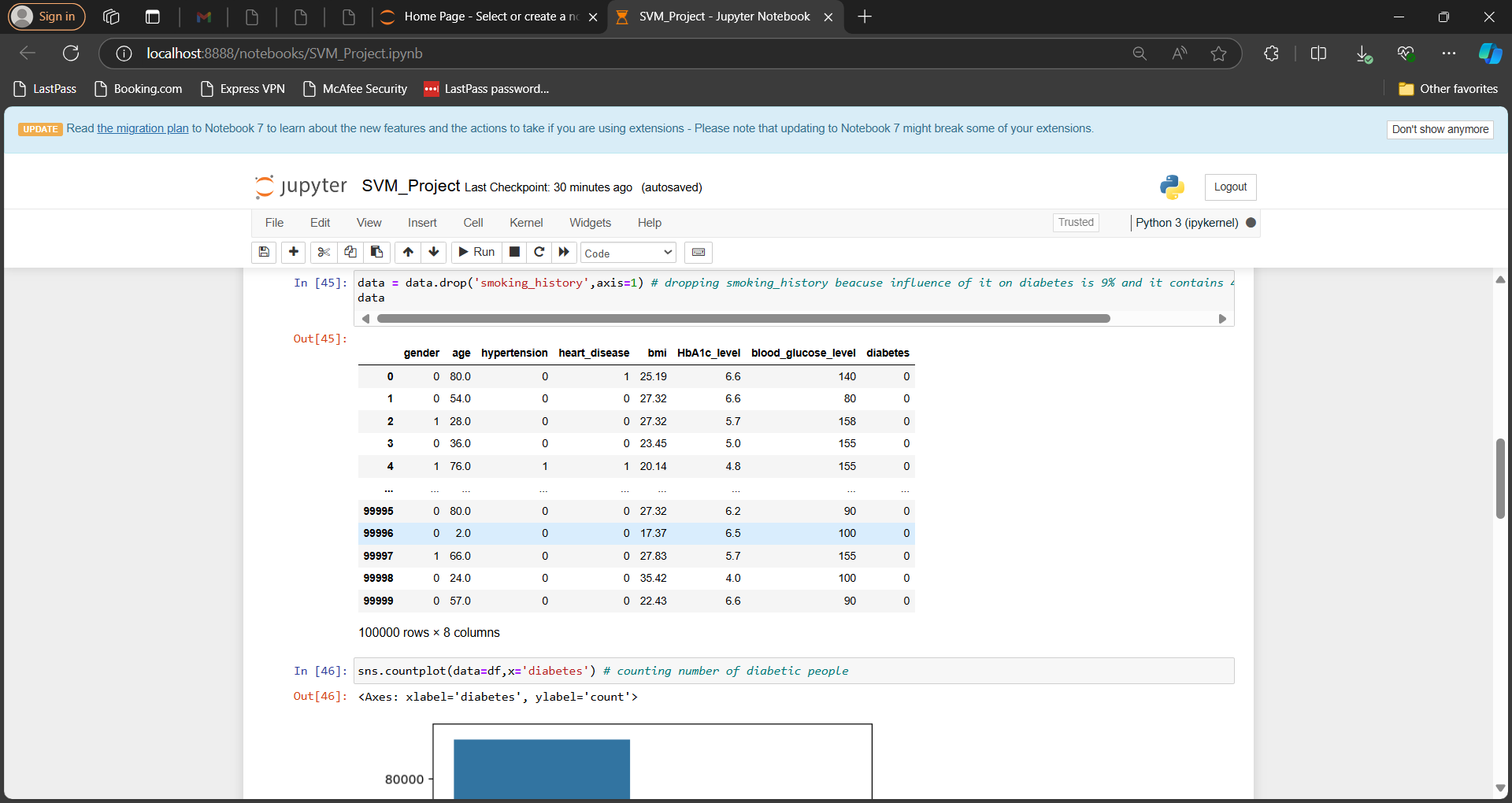
**AIM** : Predicting whether a person has diabetes or not

ALGORITHM : SVM

DATASET : diabetes\_prediction\_dataset ( Kaggle)

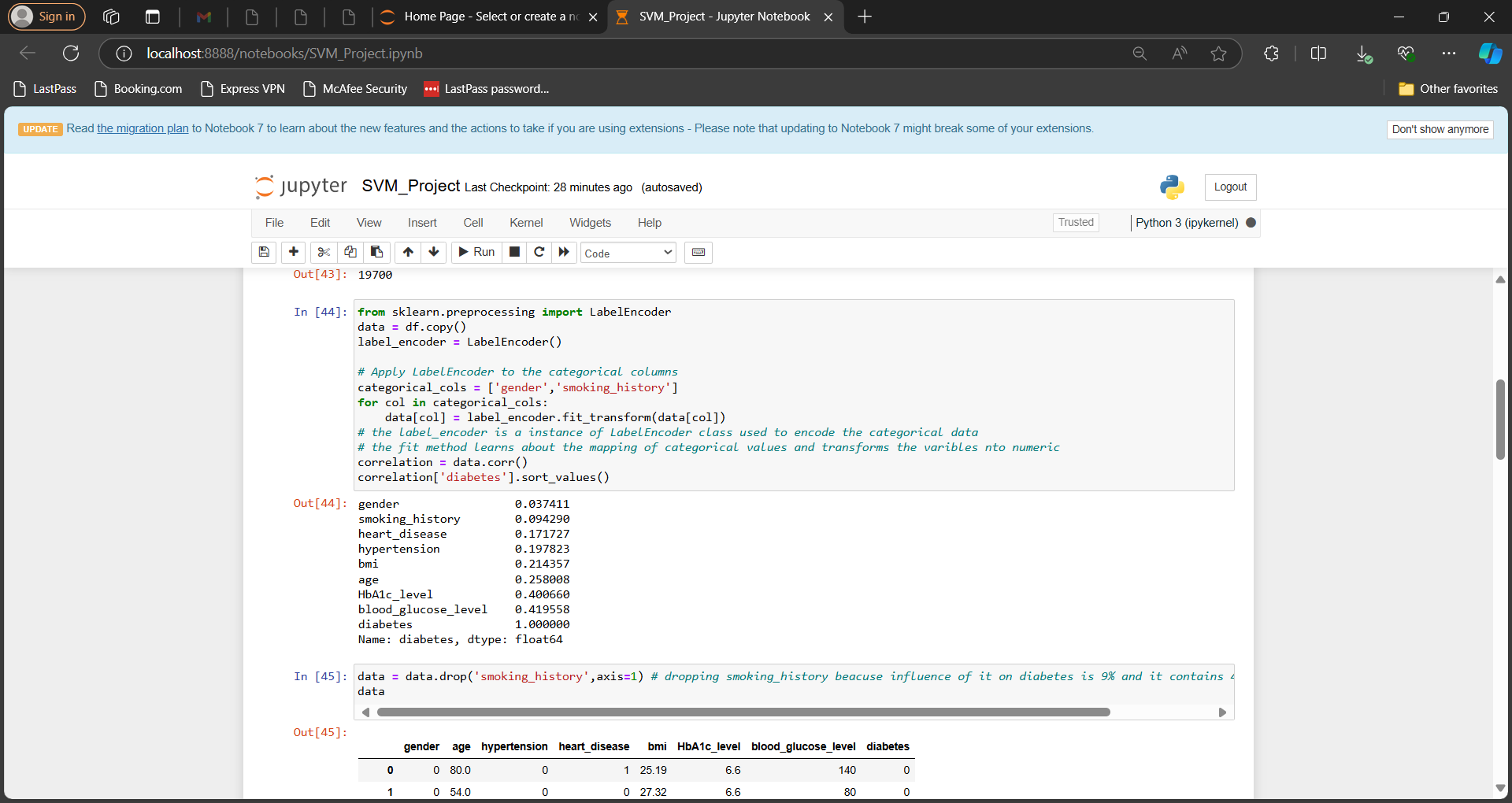
**STEP 1**

Cleaning the data. Drop the columns which have less corelation with the target column and which has more irrelavent values like No Info. If it contains NaN we can use interpolation method to replace that values.



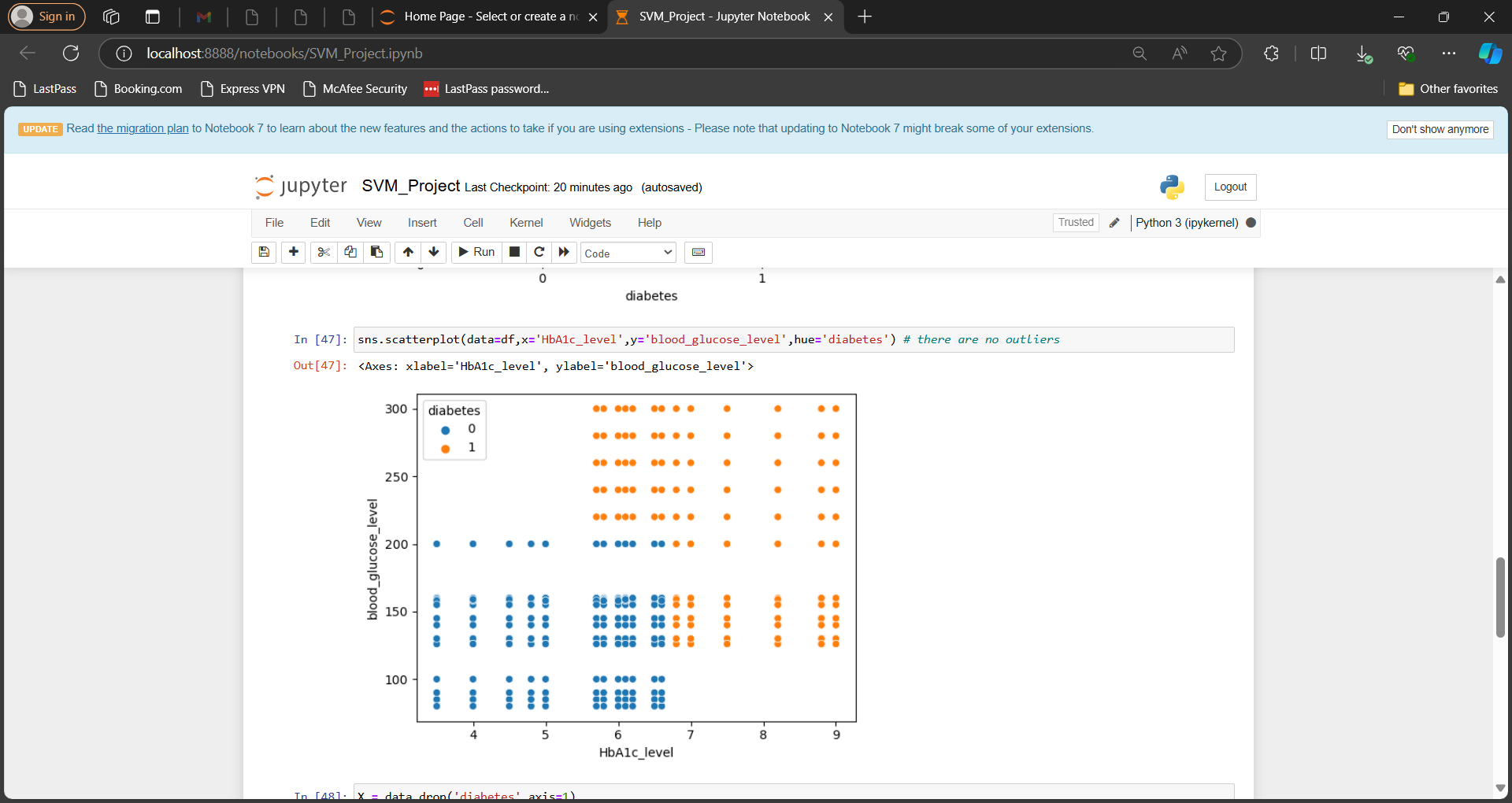
**STEP 2**

Hot encoding is used using the label encoder class for converting the categorical data into numerical format. The correlation method shows the relationship between the different features and the target values.



**STEP 3**

The scatter plot which depicts there is no outliers.



**STEP 4**

The SVM model is used to predict the target value using various parameters like C-> 1, kernel-> linear. Using GridSearchCV we select the best parameters among n number of trial parameters.

**STEP 5**

Different accuracy measuring methods such as the accuracy score, classification report and the confusion matrix is used.

