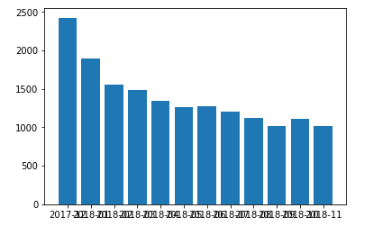
**Problem Statement 1: Plot new users acquired every month on a bar chart.**

The new users acquired every month are:

|  |  |
| --- | --- |
| Month | New Users |
| December 2017 | 2424 |
| January 2018 | 1892 |
| February 2018 | 1549 |
| March 2018 | 1490 |
| April 2018 | 1346 |
| May 2018 | 1264 |
| June 2018 | 1273 |
| July 2018 | 1208 |
| August 2018 | 1120 |
| September 2018 | 1020 |
| October 2018 | 1110 |
| November 2018 | 1015 |

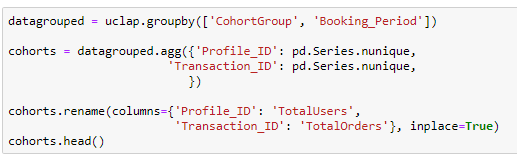
Bar Plot:

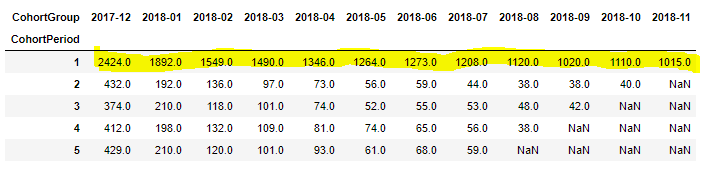


Interpretation:

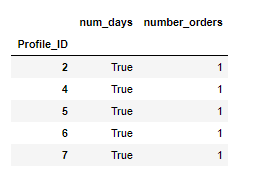
It can be seen that the number of new users is decreasing rapidly with the increase in time. Effective marketing and advertisement channels can be put in place so as to increase the number of new subscriptions in a month.

Code:





**Problem Statement 5: Distribution of users by frequency of their 90-day repeat.**

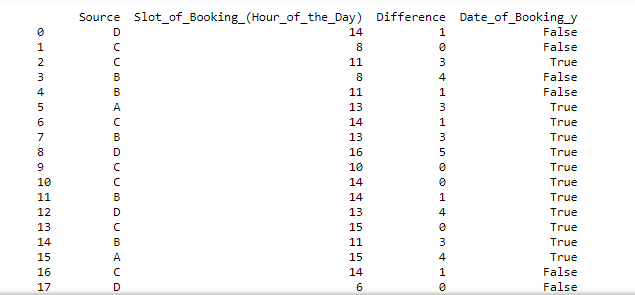


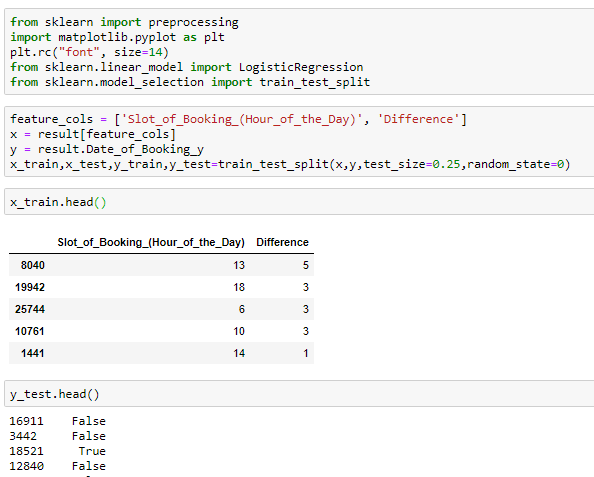
Here, Profile\_ID is the unique identity value of each customer

Num\_days is the frequency of order placing within 90 days (True Condition)

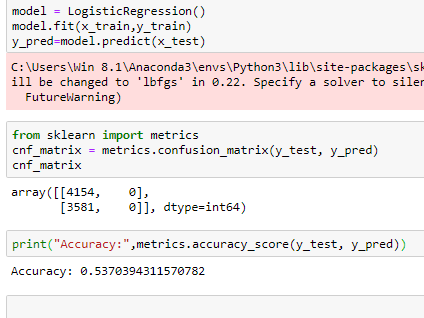
Number\_orders is the frequency value i.e number of transactions done within 90 days

**Problem Statement 4: Logistic Regression Model**

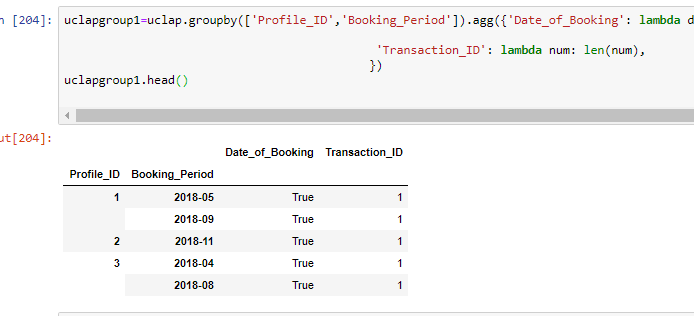


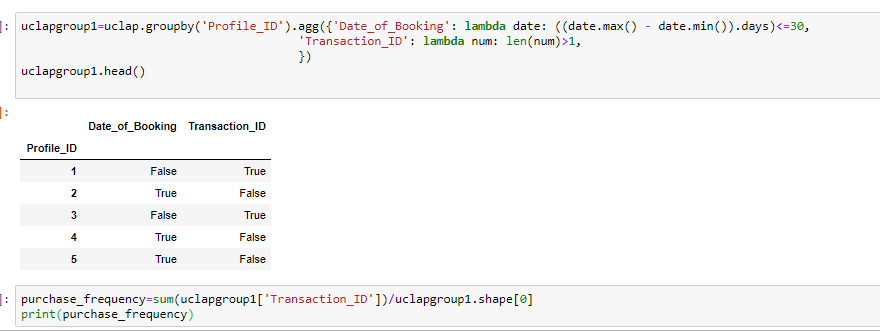


The above Logistic Regression Model gives an Accuracy of 53.7% in predicting the 90 day repeat rate of customers for all the months.

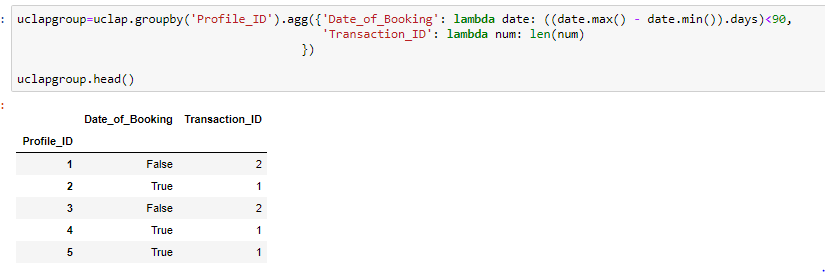


**Problem Statement 2: 30 day repeat rate for users in the month of December 2017.**





**Problem Statement 3: 90 day repeat rate for users in the month of January, February, March.**



True/False corresponds to number of repeat days of order being less than 90 (True) and greater than 90 (False).