1. What is the distribution of the customer ages?

Ans :- I have done the age distribution in the gap of 10 years and it is starting from age 18. It is like:

18 - 24
25 - 34
35 - 44
45 - 54
65 - 74
75 - 84
85 - 120

In this way I have done the age distribution.

2. What is the relationship between customer age and subscription?

Ans :- Customers having the age above 55 are having greater chances of buying the subscription because the ratio of people contacted above age of 55 and their buying subscription is greater than those of having age less than 55 yrs.

We can visualize it by below image.
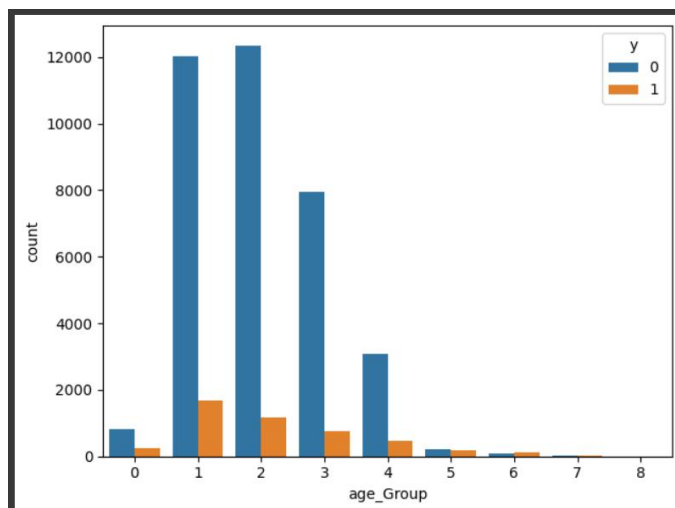
If we take the ratio of 0 and 1 of age group after 4 or 5 it is greater than those of age group less than 4.

3. Are there any other factors that are coorelated with subscription?
Ans :- The 'duration' factor is highly coorelated with the buying of subscription.

4. What is the accuray, precision, recall and f1 score of logistic regression model?
Ans :-

```
[97] model = LogisticRegression(max_iter=1000)
     model.fit(X_train, Y_train)
     predicted = model.predict(X_test)
     report = classification_report(Y_test,predicted)
     print(report)

                   precision    recall  f1-score   support

               0       0.91      0.98      0.94     12030
               1       0.61      0.23      0.34      1563

        accuracy                           0.90     13593
       macro avg       0.76      0.61      0.64     13593
    weighted avg       0.87      0.90      0.87     13593
```

5. What is the most importatnt features for logistic regression model?
Ans :- Age and Duration are the most important factors for logistic regression model.

6. How can you improve the performance of logistic regression model?
Ans :- We can improve the performance of logistic regression model in following ways:
- Hyperparameter Tuning
- Using Regularization techniques like Lasso and Ridge
- More effective feature selection

7. What are limitations of the logistic regression model?
Ans :- The limitations of our logistic regression model are:
i. It is quite sensitive to outliers.
ii. Might not handle imbalance dataset well
iii. It is quite limited expressiveness, means it may not capture complex interations between independent features

Submitted By : Priyanshu Kumar Patel
Email Id : priyanshukrpatel333@gmail.com
Contact No. :- +91 8873996561