Day 34-40

Customer Churn Analysis

Project:

https://colab.research.google.com/drive/1EzoEtCR8-CVQ-b0gzJumfpBxNgxCVUHY?usp=sharing

About the data:

A marketing agency has many customers that use their service to produce ads for the client/customer websites

Name : Name of the latest contact at Company

Age: Customer Age

Total_Purchase: Total Ads Purchased

Account_Manager: Binary 0=No manager, 1= Account manager assigned

Years: Totaly Years as a customer

Num_sites: Number of websites that use the service.

Onboard_date: Date that the name of the latest contact was onboarded

Location: Client HQ Address

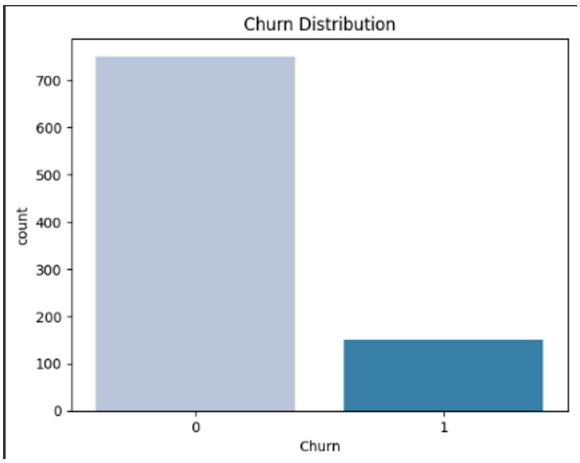
Company: Name of Client Company

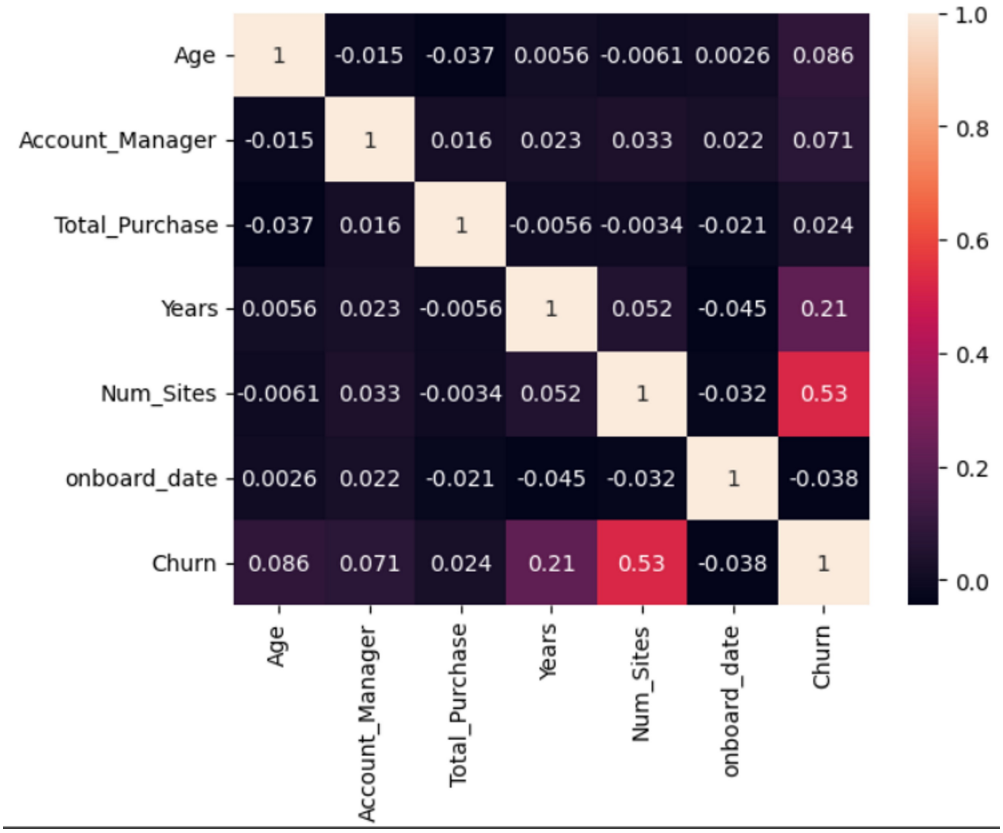
Technology used:

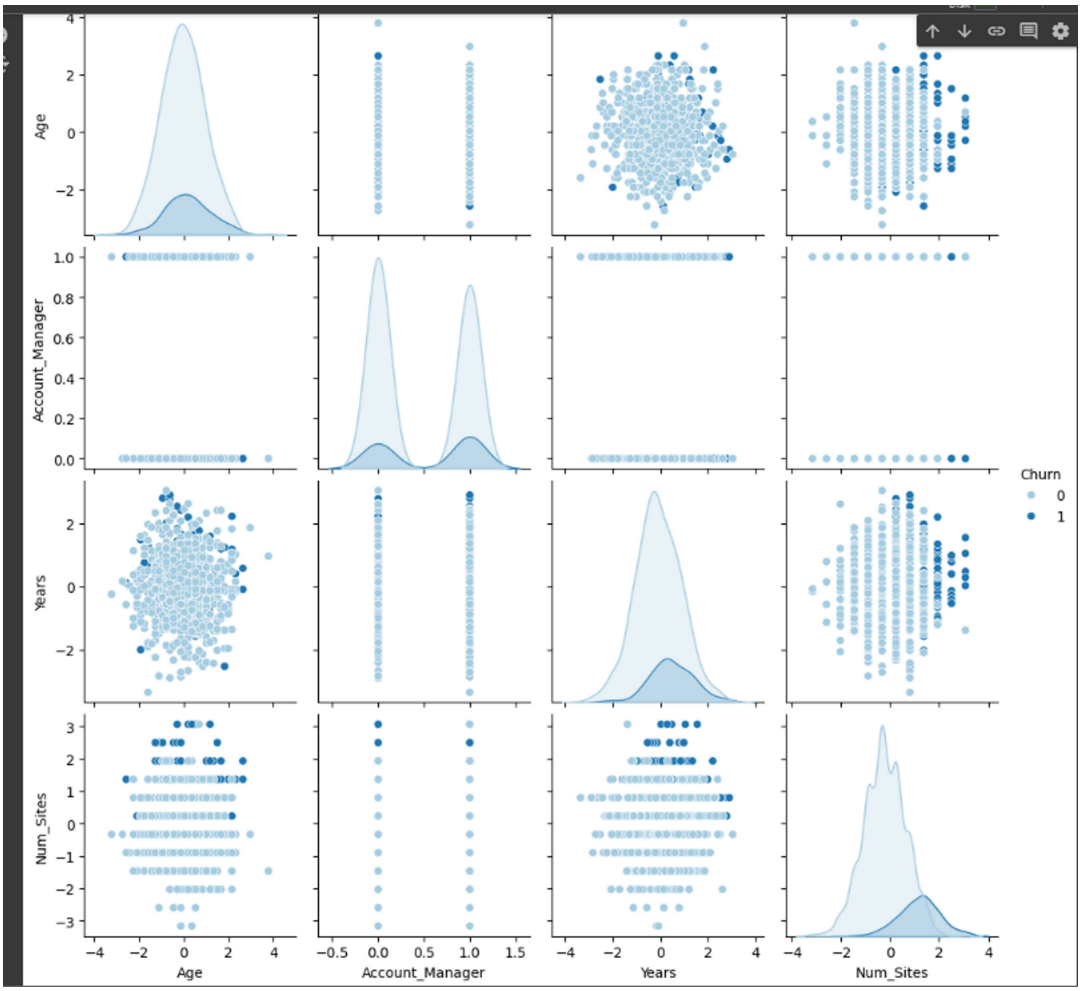
Feature Engineering, Exploratory Data Analysis (seaborn), Statistics (t-test, chi-square test)

Results:

EDA:


Churn Distribution
700
600
500
400
300
200
100
0
count
0
1
Churn


Age
Account_Manager
Total_Purchase
Years
Num_Sites
onboard_date
Churn
1 -0.015 -0.037 0.0056 -0.0061 0.0026 0.086
-0.015 1 0.016 0.023 0.033 0.022 0.071
-0.037 0.016 1 -0.0056 -0.0034 -0.021 0.024
0.0056 0.023 -0.0056 1 0.052 -0.045 0.21
-0.0061 0.033 -0.0034 0.052 1 -0.032 0.53
0.0026 0.022 -0.021 -0.045 -0.032 1 -0.038
0.086 0.071 0.024 0.21 0.53 -0.038 1
1.0
0.8
0.6
0.4
0.2
0.0



Statistics:

(H0 - Null Hypothesis, H1 - Alternate Hypothesis)

1. Age VS Churn:

   H0 : Age affects Churn, H1: Age do not affect Churn

   P value = 0.009 < 0.05

   Since, 'Age' is statistically significant at 0.9%. Hence, reject null hypothesis in favor of alternate hypothesis

2. Total Purchase VS Churn:

   H0 : Total Purchase affects Churn, H1: Total Purchase do not affect Churn

   P value = 0.4 > 0.05

   Since, 'Total Purchase' is not statistically significant at 40%, but with reference to the correlation matrix is 0.024, implies weak positive.

Thus, we may fail to reject the null hypothesis, but Total purchase and Churn may not be related strongly

3. Years VS Churn

   H0 : Years affects Churn, H1: Years do not affect Churn

   P value = 8.2 > 0.05

   Since, 'Years' is not statistically significant at 82%.. Hence, do not reject null hypothesis.

4. Num_Sites VS Churn

   H0 : Num_Sites affects Churn, H1: Num_Sites do not affect Churn

   P value = 3.34 > 0.05

   Since, 'Num_Sites' is not statistically significant at 34%. Hence, do not reject null hypothesis

5. Account_Manager VS Churn:

   H0 : Account Manager affects Churn, H1: Account Manager do not affect Churn

   P value = 0.04 < 0.05

   Since, 'Account_Manager' is statistically significant at 4%. Hence, reject null hypothesis in favor of alternate hypothesis

Conclusion:

Hence, from the above, we can say that Number of Sites and Years affect the Churn rate.

Thus, it is advisable to increase the number of ads on websites. Additionally, it is advisable to retain the old customers. Since, Account Manager has a negligible role, maybe by managing the old customers via an account manager can give them insights. This is a way promoted personalization.