EDA

Question 1: What is the distribution of card types among customers who close their account versus those who don't?

Answer: Based on this overview, it's evident that among customers who terminated their accounts, there was a higher proportion of blue cardholders compared to silver or gold cardholders. Nevertheless, among currently active customers, the prevalence of blue cardholders is still higher, though the contrast is less stark than in closed accounts. This implies that the type of card alone might not be a decisive factor in account closure, given the presence of active customers with different card types. Other variables likely influence decisions regarding account closure.

Question 2
Question: Do customers who were contacted more frequently last year tend to keep their accounts open?

Answer: In comparison to those who were contacted less frequently, consumers who were contacted more frequently last year were more likely to have an active account, according to this summary. For instance, among the consumers contacted twice, three times, and four times in the past year, there are more active accounts than closed accounts. It's crucial to remember that not all contact frequencies will show the same pattern, as there have been periods this year (e.g., 1, 5, and 6) when the number of closed accounts has exceeded the number of active accounts. Consequently, account closure decisions may also be influenced by other factors, even though there may be a general tendency.

Question 3
Question: What is the average number of months since the first account for customers who close their accounts compared to those who don't?

Answer: We can observe from this summary that there is a small variation between customers who close their accounts and those who don't in the average number of months after the first account. As a whole, bank customers with closed accounts have been around for a little while longer than those with open accounts. The difference, nevertheless, is not very great, indicating that account closure may not be significantly influenced by the length of the customer's connection with the bank. Decisions on account closure may be influenced more by other variables.

Question 4
Question: Do customers with a higher number of dependents tend to close their accounts more frequently?

Answer: As per the assessment, it is evident that there isn't a consistent pattern suggesting that clients with more dependents have a larger propensity to terminate their accounts. Although there are variations in the number of closed accounts between dependent counts, it doesn't seem that dependent count alone has a major role in account closure. Account closure decisions may also be influenced by other variables.

Question 5
Question: How does the utilization ratio differ between customers who close their account and those who don't?

Answer: This summary reveals that customers who closed their accounts typically exhibit a lower average utilization ratio in comparison to those with active accounts. This hints at a potential association between utilization ratio and account closure, with lower ratios potentially signaling a heightened probability of closure. However, additional analysis is necessary to validate this correlation and delve into potential underlying causes.

Logistic Regression Results:

1
roc_auc
binary
0.944

A tibble: 1 × 3
  .metric   .estimator .estimate
  <chr>   <chr>   <dbl>
1 precision binary   0.831
A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>   <chr>   <dbl>
1 recall binary   0.869
A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>   <chr>   <dbl>
1 accuracy binary   0.861
A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>   <chr>   <dbl>
1 f_meas binary   0.849
A tibble: 1 × 3
  .metric   .estimator .estimate
  <chr>   <chr>   <dbl>
1 sensitivity binary   0.869
A tibble: 1 × 3

```
  .metric    .estimator .estimate
  <chr>     <chr>        <dbl>
1 specificity binary        0.854
           Truth
Prediction     closed_account active
  closed_account          364    74
  active                  55    433
```

Decision Tree Results:

```
1
roc_auc
binary
0.9116

# A tibble: 1 × 3
  .metric    .estimator .estimate
  <chr>     <chr>        <dbl>
1 precision binary        0.880
# A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>  <chr>         <dbl>
1 recall  binary        0.862
# A tibble: 1 × 3
  .metric  .estimator .estimate
  <chr>   <chr>         <dbl>
1 accuracy binary        0.884
# A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>  <chr>         <dbl>
1 f_meas  binary        0.871
           Truth
Prediction     closed_account active
  closed_account          361    49
  active                  58    458
# A tibble: 1 × 3
  .metric    .estimator .estimate
  <chr>      <chr>        <dbl>
1 sensitivity binary        0.862
# A tibble: 1 × 3
  .metric    .estimator .estimate
  <chr>      <chr>        <dbl>
1 specificity binary        0.903
```

KNN Model Results:

1
roc_auc
binary
0.9115

A tibble: 1 × 3
  .metric   .estimator .estimate
  <chr>   <chr>   <dbl>
1 precision binary   0.806
A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>   <chr>   <dbl>
1 recall binary   0.845
A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>   <chr>   <dbl>
1 accuracy binary   0.838
A tibble: 1 × 3
  .metric .estimator .estimate
  <chr>   <chr>   <dbl>
1 f_meas binary   0.825
A tibble: 1 × 3
  .metric   .estimator .estimate
  <chr>   <chr>   <dbl>
1 sensitivity binary   0.845
A tibble: 1 × 3
  .metric   .estimator .estimate
  <chr>   <chr>   <dbl>
1 specificity binary   0.832
                Truth
Prediction   closed_account active
  closed_account   354   85
  active   65   422

Below are the main questions:

What are the factors that are associated with customers closing their credit card accounts?
Is it possible to predict whether a customer will close their account? If so, how accurate are the predictions?
How many costly errors is the model expected to produce?

Are there any actions or policies the company can implement to reduce the risk of losing their customers?