


# *SYRIATEL CUSTOMER CHURN*



*Beatrice Kirui*



# *OVERVIEW*

- The Syria Churn dataset is a collection of data that focuses on customer churn in a telecommunications company operating in Syria.
  - The objective of this project is to analyze the dataset and develop a predictive model to identify factors that contribute to customer churn.
  - By understanding these factors, the company can take proactive measures to reduce churn and improve customer retention.
- 

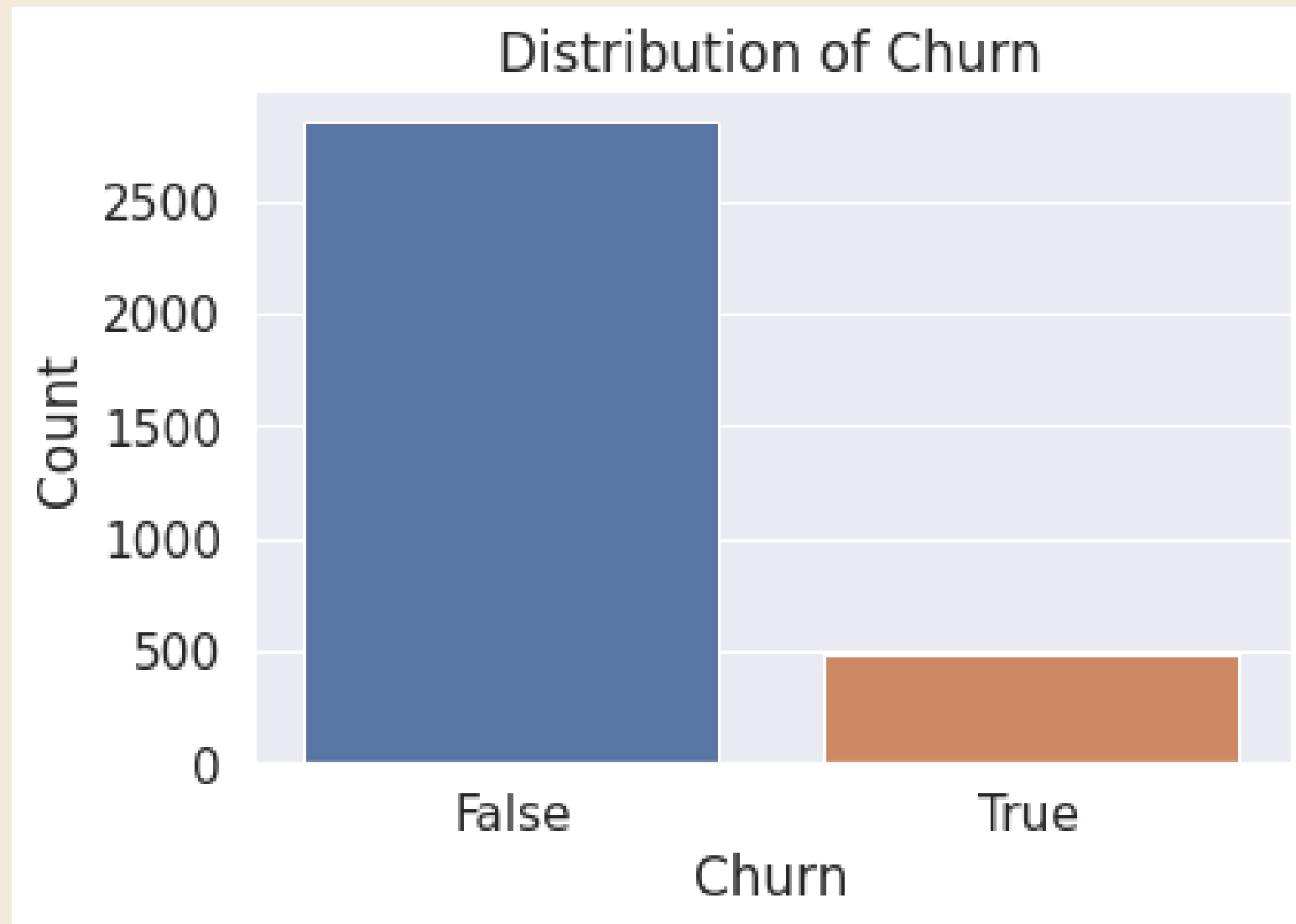


# *Business Understanding*

- Syria-Tel is a telecommunications company in Syria. They have realised that some of their customers have started to churn, discontinue their service.
- 
- Customer churn refers to the phenomenon of customers leaving or discontinuing their services with the company. It is a critical issue for businesses as it affects revenue, customer satisfaction, and overall business performance.

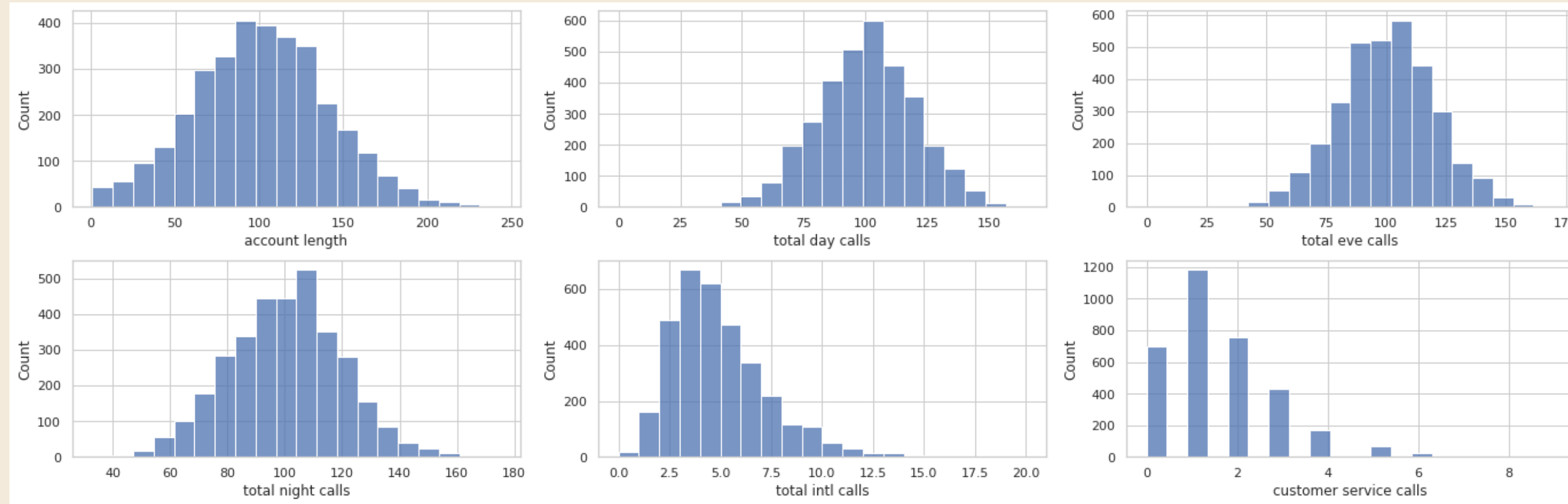


# *Distribution of Churn*



- Out of 3,333 customers in the dataset ,483 have churned whereas 2850 have not churned.
- The distribution shows class imbalance which we will address before modelling.

# *Distribution Plot for Numeric Features,*



- All of the plots displayed above except 'customer service calls' plot show a normal distribution.
- Total international calls is skewed to the right but still is normally distributed.



# *Confusion Matrix*

## *For Log*

## *Regression.*

Actual	Negative	Positive
	Negative	Positive
Negative	573	150
Positive	25	86

- True negatives (TN): In this case, there are 573 instances where the model correctly predicted non-churned customers.
- False positives (FP): There are 150 instances where the model incorrectly predicted customers as churned, but they were actually non-churned.
- False negatives (FN): There are 25 instances where the model incorrectly predicted customers as non-churned, but they were actually churned.
- True positives (TP): There are 86 instances where the model correctly predicted churned customers.

# *Evaluation of Our Final Model*

- The final model, a Random Forest classifier - accuracy of 96.64%. This indicates that the model correctly classified the majority of instances in the test set.
- Precision score of 1.0, meaning that all instances predicted as positive were indeed true positives. However, there is room for improvement in terms of recall, as the model captured only 74.77% of the actual positive instances.
- The F1-score- 0.8556701030927836. This suggests a balanced performance between the two, although recall appears to be the weaker aspect of the model's performance.
- 
- Nonetheless, the weighted average F1-score of 0.96 demonstrates a high level of overall model performance.


# *Conclusion*

- In conclusion, the Random Forest model exhibited strong accuracy and precision, making it a reliable classifier. However, the model's lower recall indicates that it may not identify all positive instances accurately. It is crucial to consider the specific requirements and priorities of the problem at hand when interpreting these evaluation metrics and deciding on the final model. Further refinements could focus on improving the model's recall without sacrificing its high accuracy and precision.



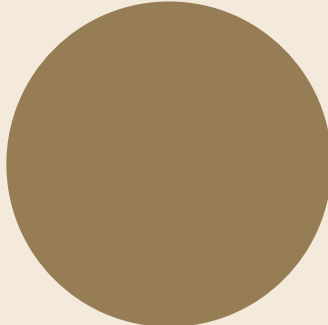


# *Reccomendations*

1. Improve Customer Service
  2. Personalize Retention Strategies
  3. Enhance Communication Channels
  4. Offer Incentives for Long-Term Commitment
  5. Monitor Customer Satisfaction
  6. Invest in Network Infrastructure
  7. Retrain and Empower Customer Service Representatives
- 



# *Next-Steps*

1. Track Key Metrics.
  2. Refine Retention Strategies.
  3. Predictive Analytics.
  4. Customer Segmentation.
  5. Enhance Data Collection and Integration.
  6. Competitive Analysis.
  7. Proactive Customer Outreach.
- 



*Thank You*

