**Situation (Context): Kuukua**

For our project, we decided to focus on a sample bank Customer Churn dataset since we believed it would be somewhat relevant to Citi/financial institutions

1. **Define customer churn**: Customer churn isn’t unique to banking, and it refers to the case where customers end their relationship with a company by ceasing to use their products and services
2. **Why is customer churn important to banks:** Minimizing customer churn is a priority for banks since, in the case of a bank, retaining existing customers is cheaper than onboarding new customers. Time and money is already devoted to acquiring new customers, and if this balance is off, this can have a negative impact on a bank's long term profit. Additionally, losing customers means losing their deposits, potential loan interest, transaction fees, and other sources of income. Churn prevention allows banks to develop loyalty and trust with as many customers as possible.
   1. Customer churn can be an indicator of future growth
3. Our goal in this project was to determine which variables strongly correlate to a customer leaving this specific bank. We split these features into two categories: one category, bank relationship features, represents factors that are somewhat in the bank’s control such as balance. The other category represents external factors such as gender and age
4. Our dataset consists of 11 key features, with our target being “Exited”, a categorical variable that shows if the customer left the bank or not (1 = yes (exited), 0 = no)

**Overview/Categorical Dashboard: Lily**

Starting with the overview dashboard, it can be filtered based on each variable but we find that a customer's location plays a significant role in their decision to leave the bank. The majority of our active customers are from Germany, Spain, and France. (HOVER OVER THE COUNTRIES IN THE MAP) However, it is worth noting that the proportion of churned customers is inversely related to the population of customers in each location which suggests a potential problem in areas with fewer clients, indicating a possible lack of allocated customer service resources.

Moving on to gender distribution, our customer base consists of approximately 54% male and 45% female customers but note that due to our dataset has done been updated . Interestingly, we observe a higher churn rate among female customers, with 25% of them having churned compared to the churn rate for male customers which is 16%.

The age of our customers ranges from 18 to 92 but this wide range suggests that we need to consider the unique needs and preferences of customers across different age groups. We found that approximately 56% of customers in the age range of 50 to 70 exhibits the highest churn rate. This age group indicates the need for targeted retention strategies tailored to customers however, the distribution of customer age is right-skewed so there’s relatively fewer customers in the older age group which means we need to pay attention to retaining customers in that specific older age bracket.

**Bank Relationship : Dario**

A strong customer-bank relationship is important for lowering churn rates and retaining customers. By analyzing these factors,

A few key metrics that we derived from this data are that:

the number of exits decline as the number of products that a customer uses increases.

69% of all exits came from customers who only used one bank product.

Individuals with a balance greater than 100k displayed a 25.2% exit rate while those with less than 100 thousand balance had a 15.9% exit rate.

The bank should aim to strategize ways to combat these failing areas, possibly by improving benefits for customers in these specific target groups.

Prioritizing the customer-bank relationship positively impacts the bank's reputation and success. The bank should aim to do best by their customers. A positive byproduct of doing so is that satisfied customers are likely to recommend the bank to others, which can attract new customers and contribute to economic growth.

**External Factors: Annie**

In our analysis of the bank customer churn rate, we explored various external factors that could potentially influence customer behavior.

To begin with, we chose these factors based on their relevance to customer decision-making and their potential impact on their banking relationship. Age, for instance, could be a critical factor as it often correlates with significant life events and changing financial needs. Credit score and estimated salary, on the other hand, provide insights into a customer's financial stability and capacity to maintain banking services.

To examine the relationship between these factors and customer churn, we utilized box plots as a visualization tool.

Among these factors, age stood out as having the most significant difference within the box plots. The box plots revealed that the median age of churned customers was noticeably higher or lower compared to non-churned customers, indicating a positive correlation between age and churn rate.

To gain further insights, we explored scatter plots of age versus credit score and estimated salary. Despite not observing a direct correlation between age and these variables, we noticed that a significant concentration of data points occurred within specific age ranges for churn and non churn rates. This concentration suggested that age might play a pivotal role in influencing customer churn.

In addition to visualizations, we also examined the distribution of the data for age, credit score, and estimated salary. Understanding the distribution is crucial because it helps us assess the normality of the variables and identify any potential outliers. We discovered that age and credit score followed a normal distribution, while estimated salary exhibited an uniform distribution pattern. This finding highlighted the unique characteristics of each variable and underscored the importance of considering their impact on the churn rate separately.

Throughout the project, we heavily emphasize the use of visual graphs, such as box plots, for intuitive understanding. Box plots offer a quick and visual way to identify how certain factors, like age, can significantly impact whether a customer churns or not in a bank. By leveraging these insights, we can better understand the dynamics of customer churn and make informed decisions to mitigate it.

**Conclusion: Dario**

In order to understand which variables strongly correlate to a customer leaving this specific bank., we trained a ML model, specifically a random forest classifier, on all the churn data. Using a random forest classifier to analyze customer churn data can provide several benefits, including accurate predictions, feature importance, which can allow banks to better allocate their efforts.

Our model determined that the three most important features to making churn predictions were age, estimated salary and credit score. After this analysis we proposed different methods that the bank could implement.

A few proposed solutions we have for the bank are as follows:

1. **Employing a churn prediction model**

With a prediction model, the bank can proactively identify customers with a higher likelihood of exiting. The bank can then utilize these insights to implement tailored retention strategies that cater to individual needs and preferences of their customers.

1. **Customized Customer Engagement**

The bank can deploy personalized retention strategies, placing a particular emphasis on the 50-70 age group. This demographic has demonstrated a comparatively higher churn rate, underscoring the significance of tailoring initiatives specifically towards their needs.

1. **Establishing a Performance Monitoring System**

Maintain a continuous monitoring process to track churn rates, carefully analyze customer feedback, and assess the efficacy of retention strategies. Leveraging these data-driven insights will aid in making informed adjustments and enhancements, ensuring the optimization of customer retention efforts.

## Final Remarks

One setback in the analysis was the limited data available on gender. This was a publicly available dataset from previous years,, and the dataset only contained binary gender information. This limitation prevents the inclusion of all genders, and we hope this can be expanded upon in future analysis.

In conclusion, customer churn analysis is a vital process for banks, where customer retention is crucial for success. By understanding the factors that contribute to customer churn, and which are most impactful, banks can take proactive measures to mitigate churn and improve customer retention.

We’d like to thank all of you for attending our presentation and we hope you all enjoyed what we had to say. Feel free to ask any questions!