Revolut KYC Operations Challenge

Problem:

The KYC pass rate has decreased substantially in the recent period: what are the drivers of this decrease, and what solutions can be presented?

Assumptions:

- For both the Docs Report and the Facial Sim Report, a client has passed if the overall result is both "Clear."
- All data collected is "as is" (i.e. if a value is not populated it is because the information is not available for that particular data point); Due to this, cleansed data maintains the integrity of the source files provided.
- If both steps' results are "Consider" then it is not a pass, it is also not a pass if one item is "Clear" and the other "Consider."

Conclusion:

Four main problems have been identified as a cause for the decrease in the KYC pass rate which are discussed in further detail in the follow pages:

1. Too many attempts made by the same user
2. Loss of potential customers after one try
3. Facial similarity unscored data
4. Document step pass rate by document type and country

Overall, each problem may have minimal impact on the pass rate individually, but when combined together, can reasonably explain the decrease based on the data presented.

I would also recommend **adjusting the calculation of the KYC pass rate** by requiring an order for completion before the next step can be taken (i.e. must pass the Document Check prior to moving on to the Facial Sim check). There were a combined 44,107 attempts or 25% of total attempts where one step was "Clear" and the other "Consider." By requiring one step to be completed before moving on, the number of attempts used in the formula will be significantly lower and not skew the data. The below table shows a matrix of the combined results of both steps.

	Facial Sim Result	
Doc Result	clear	consider
clear	126,889	5,511
consider	38,596	5,406

High-level Observations/ Caveats:

Limitations of Problems Identified: *These problems highlighted are not a complete assessment of the drivers of the decrease in the pass rate. It is an assessment based on the available information at the time and without access to customer feedback or other value drivers. In addition, the solutions identified are not an exhaustive list of what can be done to address the problems.*

***Lack of Data Completeness**: Overall with the facial similarity results, there are inconsistencies regarding the completeness of data with much of the data have null values. It should be ensured that all fields must be populated to some degree (even with N/A) to be able to capture the full details of each attempt and understanding the root of the cause.*

Conclusions Explained and Solution:

1. Too many attempts

Using the data provided, we filtered how many attempts were made per user based on the User ID. Based on this analysis, there were 176,404 attempts made by 142,742 users made.

The below chart shows a breakdown of the number of attempts per user:

Attempts per User	# of Users	Attempts
1	110,374	110,374
2	31,118	62,236
3	1,145	3,435
4	78	312
5	7	35
6	2	12
Total	**142,724**	**176,404**

Based on the information provided, it is clear that there are significant instances (approx. 77%) of users only making one attempt. However, there are only two attempts allowed in the KYC process, and the data shows that there are 1,232 users that made 3 or more attempts. This represents **2.15% of the total attempts**, and while not large, may be a contributing factor to the decline we are seeing in the pass rate measure at it increases the denominator of the factor.

In order to solve this, we need to create a control in the system that prohibits users from making more than 2 attempts and ensures that one user isn't able to make multiple attempts. This could mean gathering more personal information from the user to make sure there are not duplicates.

2. Loss of potential customers after one try

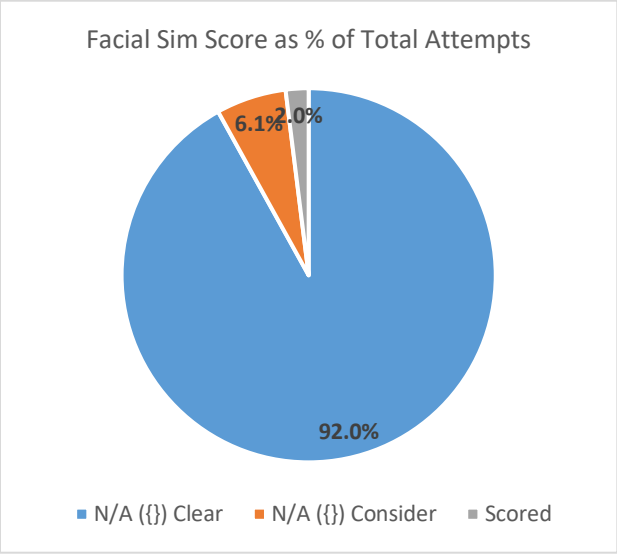
In order to test how many users failed after one try and were lost after not passing, we calculated how many users attempted the process 1 time and were marked as "Consider" for both steps. If the user receives both "Consider" then it is assumed they did not pass. There were **3,264 users that made** one (1) attempt and did not pass. *Because of this, I would consider this as a potential lost customer as they might have passed after the second try.*

Overall, these 3,000+ users **represent 2.3% of total users** and 1.9% of total attempts with all of these attempts contributing negatively to the pass rate. In addition, these one-time try users represent **60.4% of attempts where both steps were "Consider."**

We should implement an initiative to encourage potential clients to always try for a second time. This could be in the form of **multiple e-mail follow ups** or a prompt for a **customer service window** that pops up to chat with the client and provide assistance if needed. The goal of this solution is to reduce this metric to zero as there is a possibility that the user could be converted to a customer in the second attempt. If the user has changed their mind about

becoming a customer, then we should add a field to provide that information so those attempts can be excluded from the KYC pass rate.

3. Facial Similarity Unscored Data



There were **172,920 total instances** where facial similarity was not scored ({} in dataset and changed to N/A for presentation purposes). Of these 172,920 unscored attempts, 162,217 attempts were marked as "Clear" whereas only 10,702 were marked as "Consider."

The solution to this issue is to ensure that all the data is scored appropriately since everything under a score of 0.5 is not cleared. We need to ensure that there are controls in place that score each attempt and mark it appropriately. In theory, these unscored attempts can greatly skew our in our understanding of pass rate especially if they are not marked correctly.

4.1 Docs Report Success by Document Type

An analysis was performed to understand the nature of the documents that were used in each attempt for the Document Check step and the associated designation. The following table shows the results.

	clear	consider			
Document Type	clear	caution	rejected	suspected	**Grand Total**
National Identity Card	49,745	4,099	1	768	54,613
Driving License	44,865	5,418	7	202	50,492
Passport	35,134	6,206	1	887	42,228
Blank	3	22	26,081	-	26,106
Residence Permit	2,620	246	-	63	2,929
Work Permit	16	1	-	-	17
Voter ID	11	-	-	-	11
Tax ID	7	-	-	-	7
Birth Certificate	1	-	-	-	1
Grand Total	**132,402**	**15,992**	**26,090**	**1,920**	**176,404**

Overall, the National Identity Card received the highest marks of "Clear" results representing over 37% (49,745 of 132,402) of the total "Clear" attempts. Following is the Driving License and Passport which collectively with ID cards make up 97.99% of the "Clear" attempts. Voter ID, Tax ID and Birth Certificates were sparsely used, but did not yield "Consider" results. Both Driving

License and National Identity Cards were the two forms of documents that contributed to the most "Clear" results for Facial Similarity.

Unsurprisingly, blank forms of documents were rejected representing 26,081 of the total 176,404 attempts. *I believe these rejections should be removed from the attempts as they just contribute to total attempts used in the pass rate calculation and skew the data as no document was provided.*

4.2 Docs Report "Consider" Results by Country

The below shows an excerpt of the highest attempts marked "Consider" based on Document type and Country.

	Driving License	National Identity Card	Passport	Residence Permit	Grand Total
GBR	3,877	3	2,083	70	6,033
FRA	154	1,126	1,154	90	2,524
LTU	273	834	270	10	1,387
IRL	262	34	1,056	8	1,360
ESP	58	574	194	32	858
POL	144	444	184	11	783

The data shows that certain forms of documents are preferred in certain countries. For example, the National Identity Card should be utilized in GBR compared to Driving License or Passport whereas a Driving License is preferred in FRA compared to the others as they have the lowest number of "Consider" marks compared to the other forms of documents.

I would recommend limiting the types of documents that can be used for this verification. In general I would recommend promoting the use of Driving License and National Identity Card documents with Passport following, but then tailoring the recommended forms of documents based on the country where the user is based.

I would also recommend ensuring that all data points be filled out prior to submitting the document for a check as the data provided did not have records for every data point. If a client had failed to input some information or if the information did not populate via the photo, then it is highly likely that the document would not pass this step.