Understanding Revolut's drop in KYC pass rate

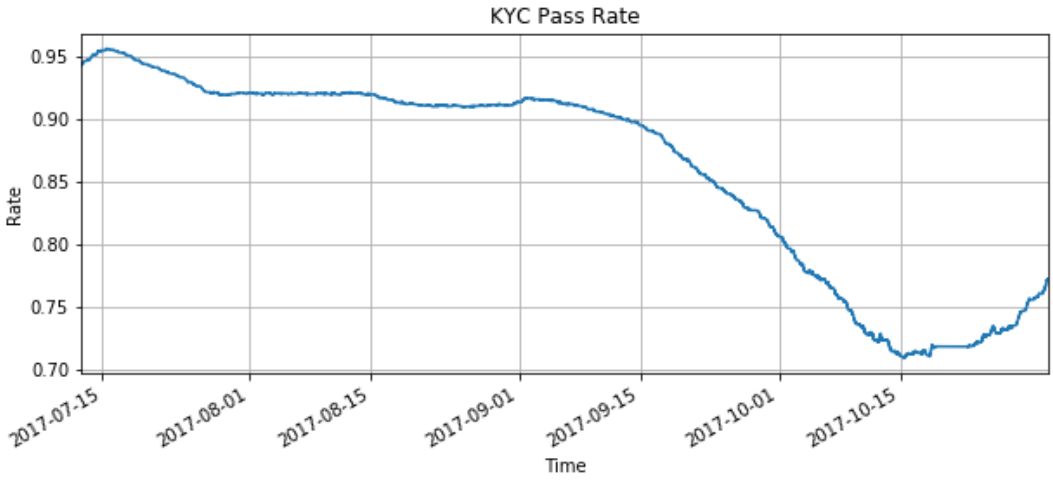
Introduction & Background:

Each prospective customer has to go through a Know Your Customer (KYC) process by submitting a government-issued photo ID and a facial picture of themselves to an external partner Veritas. Veritas then performs 2 checks:

- Document check: To verify that the photo ID is valid and authentic;
- Facial Similarity check: To verify that the face in the picture is the same with that on the submitted ID.

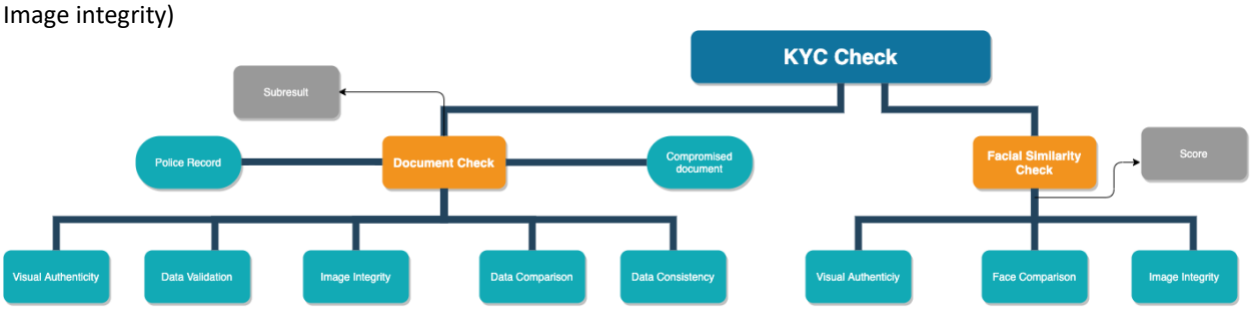
The customer will 'pass' the KYC process and get onboarded if the results of both Document and Facial Similarity checks are 'clear'. If the result of any check is not 'clear', the customer has to submit all the photos again. Each customer has up to 2 attempts.

From May 2017 to October 2017, the pass rate has substantially decreased as seen in the figure below. To understand the root cause of this drop, we have analyzed Veritas API data (document & facial similarity check's data).



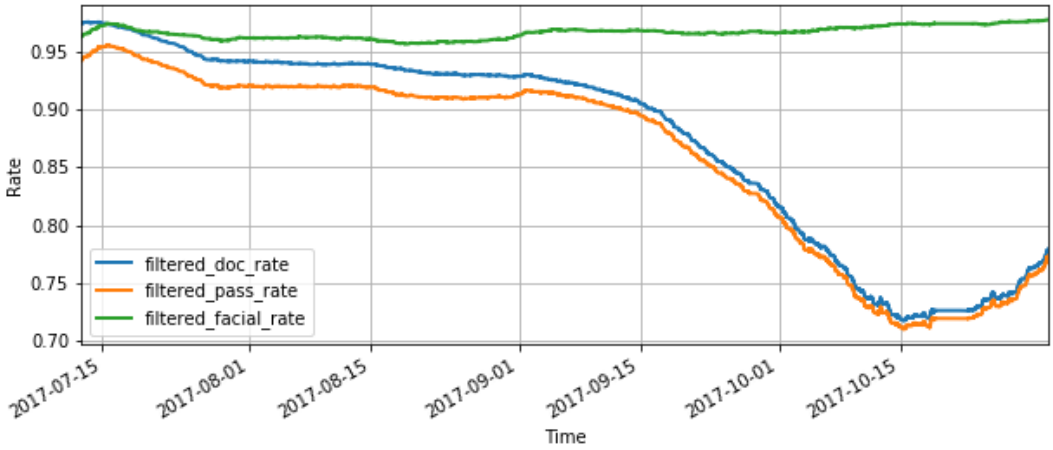
Below is a diagram of the overall structure of Veritas Process. The tree can be read in the following way: one node corresponds to one check or condition that is cleared if all its leaf nodes (that also correspond to checks or sub-checks) are cleared (or not reported in some cases). The graph below contains the two first layers only and the clearance of some of the leaf nodes displayed also depends on sub-checks in the KYC Process.

The node sub-result provides a more detailed information if the document check is failed. The node Score contains a facial similarity check score calculated using the facial similarity sub-checks (Visual authenticity, Face comparison, Image integrity)



Analytics:

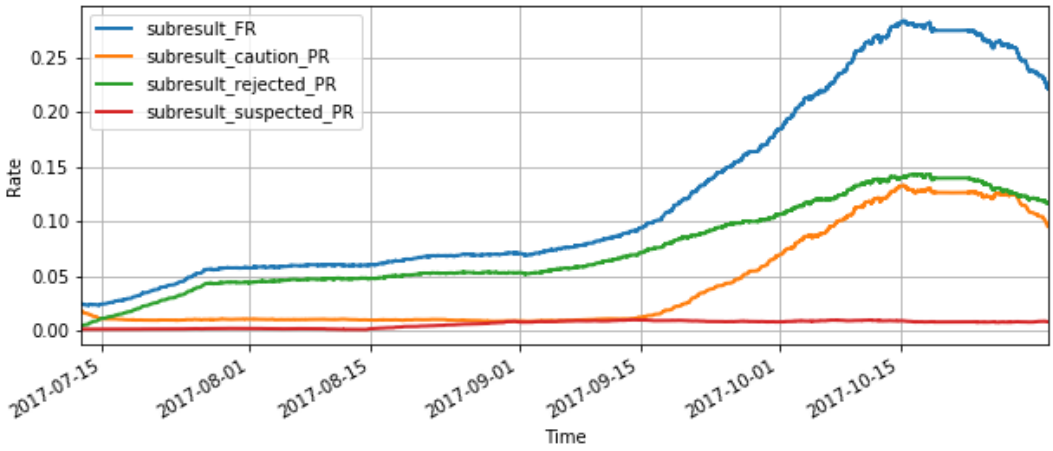
We use a top down approach to track down the root cause of the drop in KYC. We first investigate whether the drop is due to a decline in Facial Comparison check or Document check. The figure below shows that the facial comparison pass rate remained flat (it actually slightly increased) and suggests that the drop in KYC pass rate (here labelled filtered_pass_rate) is entirely due to a decline in document check (there is a 0.99 correlation between the two series).



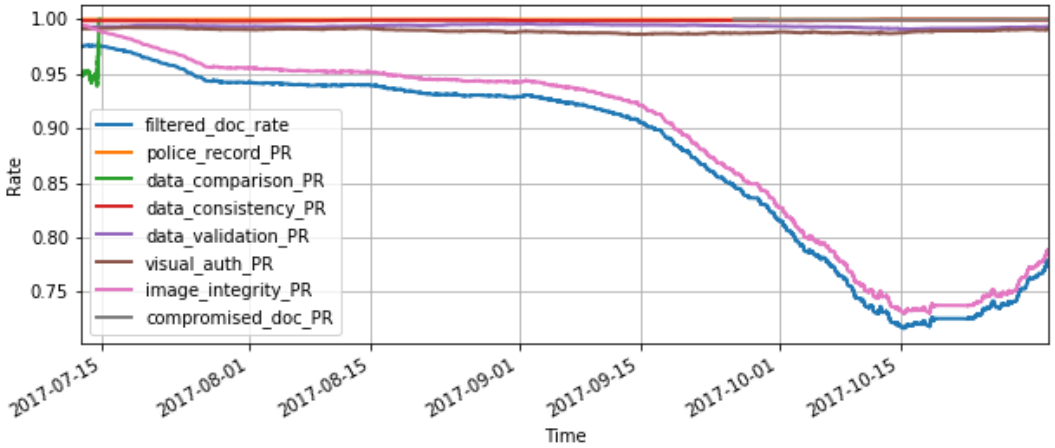
As stated earlier, if the document check is failed, the subresult feature allows us to have a more precise understanding of the underlying reason.

- Rejected: If the report has returned information where the check cannot be processed further (poor quality image or an unsupported document).
- Suspected: If the document that is analyzed is suspected to be fraudulent.
- Caution: If any other underlying verifications fail but they don't necessarily point to a fraudulent document (such as the name provided by the applicant doesn't match the one on the document)

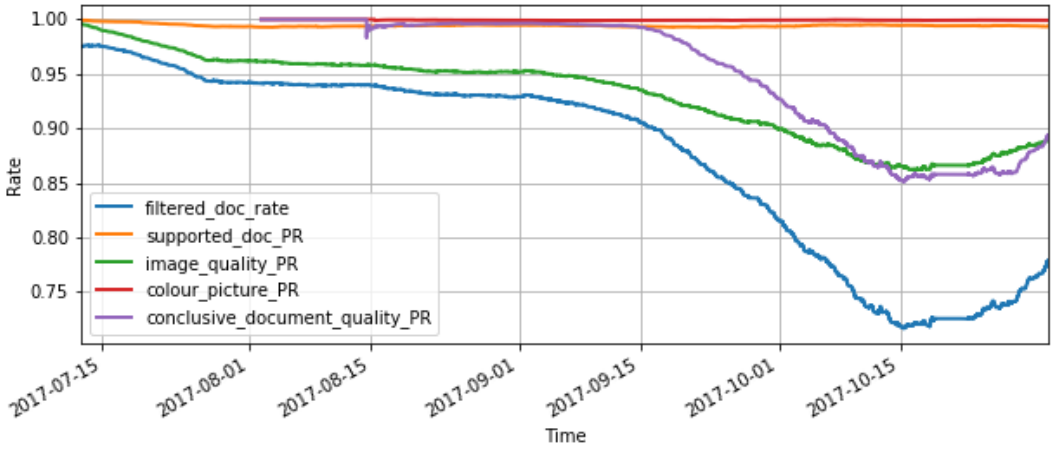
The plot below shows that the increase in failure rate (labelled as subresult_FR) is first due to an increase in 'rejected' in end of July and then due to an increase 'caution' around mid-September.
The fraud thesis can be safely rejected at this point i.e the decrease in pass rate is not due to an increase in fraudulent attempts.



Again, we investigate which sub-check involved in 'Document Check' and understand that the drop is entirely due to the Image Integrity check that is not cleared, hence implying that the documents weren't of sufficient quality to verify or that Veritas had a systemic for verifying the documents.
In the figure below, filtered_doc_rate corresponds to the clearance rate of the documents check.

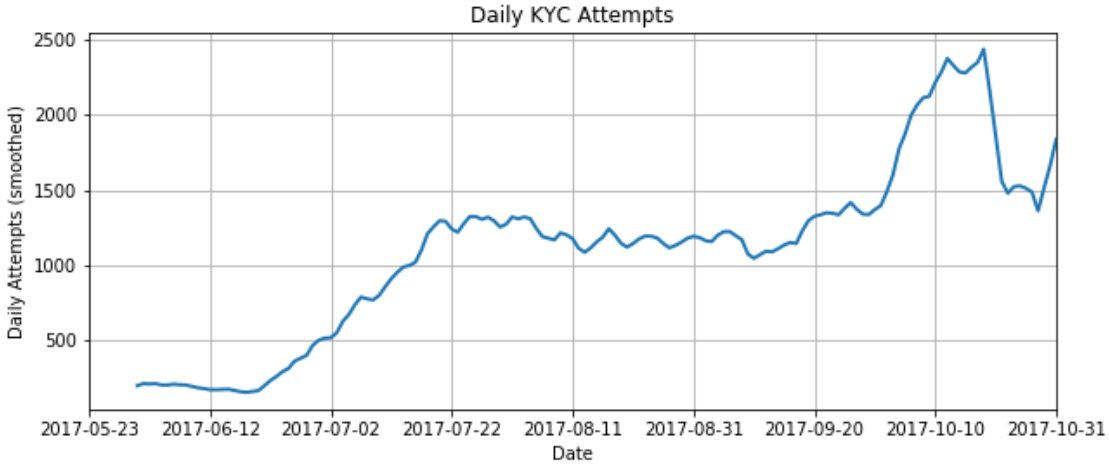


The image integrity check is comprised of a supported document check (asserts whether the document presented is valid), an image quality check, a color picture check and a conclusive document quality check (asserts whether the document is of enough quality to be able to perform a fraud inspection). We first notice that 'color picture' and 'conclusive document quality' only started to be reported in August. We also understand that the first drop in clearence rate is due to the image_quality pass rate decreasing and that the second is mainly due to a decrease in the conclusive_document_quality pass rate which was not recorded before August (hence implying that the pass rate before august could have been biased and that the clearance methodology has changed over time)

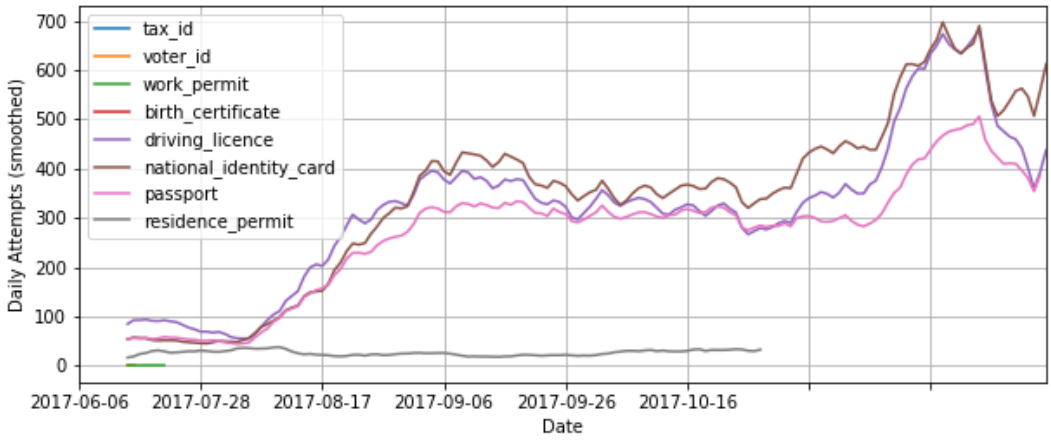


By plotting the daily attempts, we clearly see that there is a correlation between the two jumps in failure rate and the two jumps in daily attempts. We can infer that it is due to a hyper growth in the KYC attempts, meaning that the Veritas API has difficulties scaling and that the problem is due to an overflow.
We can see that the daily KYC attempts time series and image quality Pass rate time series are negatively correlated, meaning that each time there is a positive jump in the daily KYC attempts we notice a drop in the Image Quality Pass rate: the image quality check doesn't seem to be able to scale at more than 500 attempts/day.

We also notice that the conclusive document quality pass rate is negatively correlated with the daily KYC attempts. In this case, even if data is only reported starting from August, we can infer that this check can scale to up to 1500 daily attempts but brutally drops when exceeding this threshold. We can also infer that the increase in 'Caution' in subresult is due to a decrease in conclusive_document_quality clearance rate.



Finally, we check if there is a correlation between the failure rate and the attempts with a certain type of document (ex Passport). Other features like gender, date of birth, date of expiry, nationality etc. are irrelevant for our problem. The graph below reject the above thesis: we would have noticed a spike (or two, one for each jump) for one specific document if it was the case while we see that passport, ID card and driving license follow the same trend over time.



Summary:

- The drop in KYC pass rate is due to a decline in document clearance rate.
- We notice a first drop (5-10%) in beginning of July and a second drop (10-15%) in mid-September.
- The decrease in document clearance rate is entirely due a decrease in Image Integrity
- The drop in KYC pass rate does not seem to be due to an increase in fraudulent attempts or a change in user's behaviors
- The first drop in image integrity is due to a decrease in Image quality. The second drop is due to a decrease in Conclusive document quality (which only starts to be recorded in August)
- Veritas seems to have changed its process and have added additional Image integrity sub-checks in August (Conclusive document quality, Color Picture), the pass rate before August could hence be biased.
- There is a direct correlation between the (positive) jumps in the Daily KYC attempts (which is increasing) and the (negative) jumps in KYC pass rate, implying that the Veritas API has difficulties scaling.
- We can infer that the image quality check fails at scaling above 500 daily attempts, the conclusive document quality check fails at scaling above 1500 daily attempts.

Solutions:

1) **Work with Veritas to improve their product:**

    - Make the process more transparent and more user friendly
        - Some additional conditions have been added recently, making the process more constrained.
        - The process is clearly not user friendly as only 50% of those who failed at their first attempt tried again.
    - Work on scalability (Assuming the inferences were right i.e, Image quality and Conclusive document quality drop is due to a failure at scaling)
        - Some conditions are directly altered by the growth in attempts (Image Quality, Conclusive Document Quality).
    - If the issue is not scalability, understand why image quality and conclusive document quality rate decreased

2) **Find a competitor with a technology adequate to the next growth phase of Revolut (i.e able to scale, user friendly)**