venn

Case Study:

Pass Rate Fluctuations

Hibah Naviwala

Executive Summary

What is driving the drop in Pass Rate?

- **Declining Image integrity**, driven by poor image quality especially pertaining to “coloured_picture” records

- **Data quality errors**, with an increasing number of compromised records

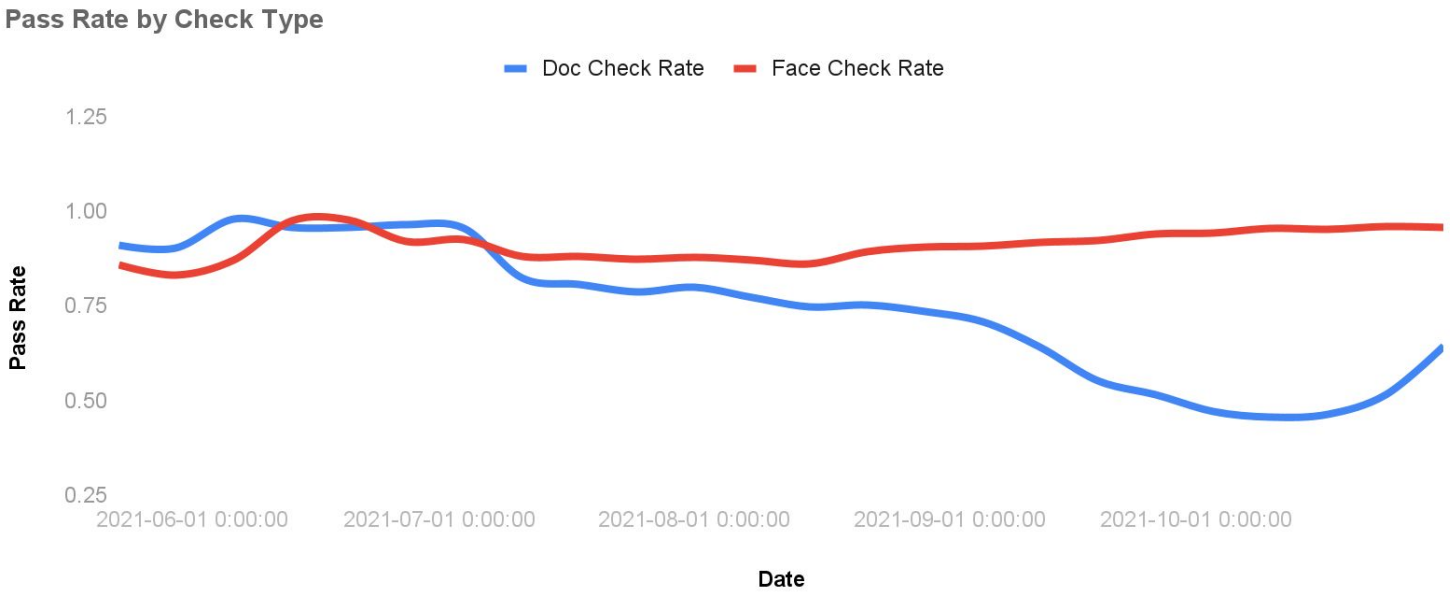
Recommendations:

1. Liaise with IDV Verify to ensure that “Coloured_Picture” algorithm is accurate

2. Only flagging a record as “rejected” after 2 completed KYC attempts

3. Human in the loop support from IDV Verify to cross check customers with their database of compromised records

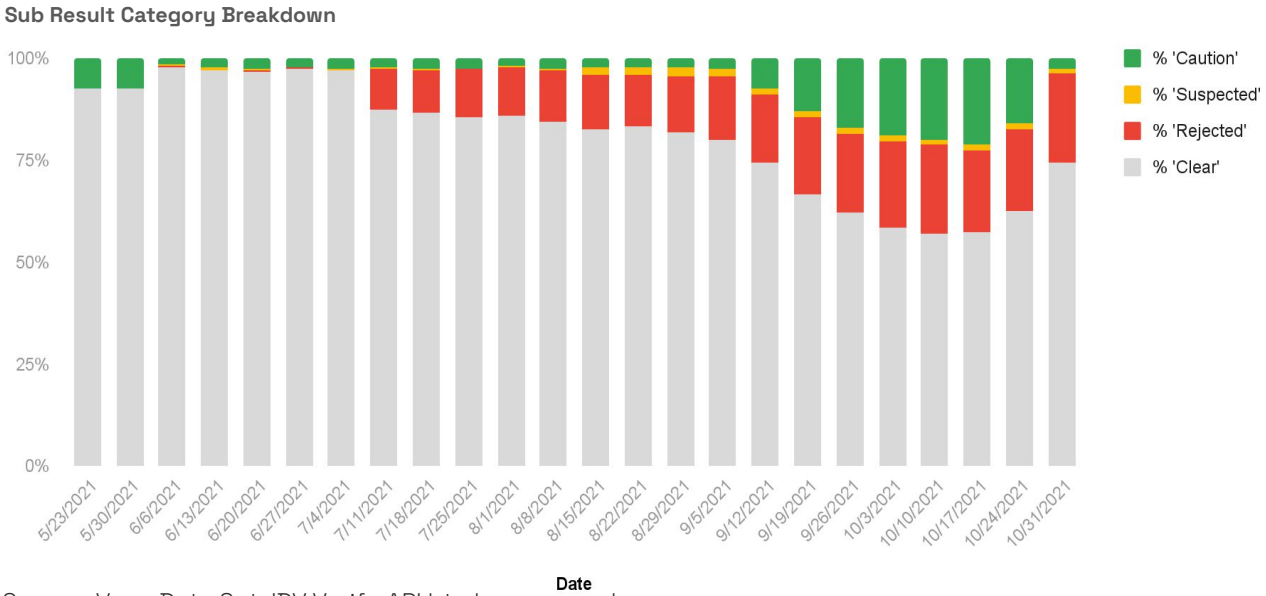
The Pass Rate is made up of Facial Similarity Check & Document Check, with the latter driving the drop in Rates



Source: Venn Data Set; IDV Verify API html source code

Further investigation into Doc Check reveals two driving factors for declining rates:

decreasing image integrity & more non-fraudulent failures



Sub Result Category Breakdown:

Caution: Other verifications failed, but don't point to fraud

Suspected: Document thought to be fraudulent

Rejected: The check can't be processed further due to reasons such as poor image integrity

Clear: All underlying verifications pass

Source: Venn Data Set; IDV Verify API html source code

Let’s take a deep dive into each lever:

Decreasing Image Integrity

Represented by “Rejected” records from previous slide

Increasing Non-Fraudulent Failures

Represented by “Caution” records from previous slide

Let’s take a deep dive into each lever:

Decreasing Image Integrity

Represented by “Rejected” records from previous slide

Increasing Non-Fraudulent Failures

Represented by “Caution” records from previous slide

Decreasing Image Integrity

Image integrity, the primary check pertaining to “Rejected” records, is made up of the following sub checks or “inputs”:

- **Supported_document:** is the doctype recognized?
- **Image_quality:** is the image clear enough?
- **Colour_picture:** is it a colour image?
- **Conclusive_document_quality:** whether the quality was high enough to perform fraud inspection

Decreasing Image Integrity

Assumption:

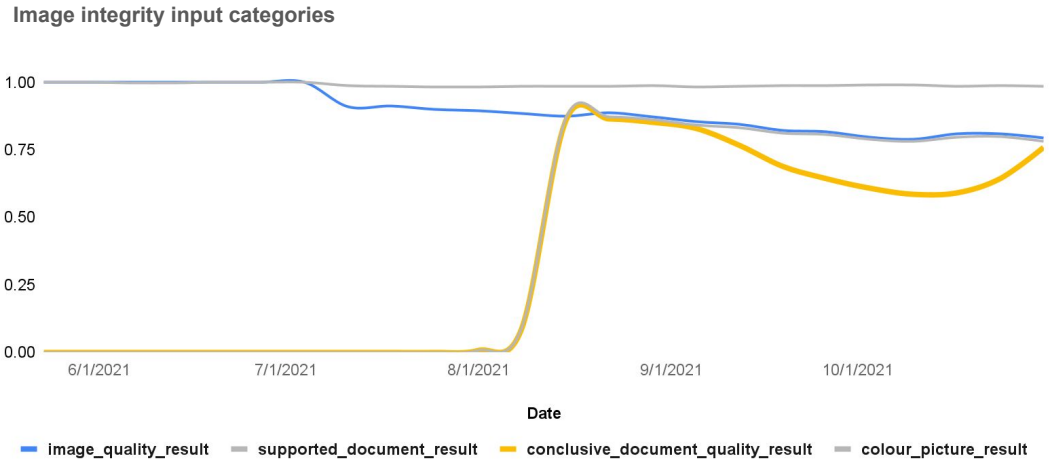
The blanks in the data are not due to net new checks/variables being introduced, but simply missing data.

We have decided NOT to impute missing values, so as not to introduce bias into this data.

Assumption:

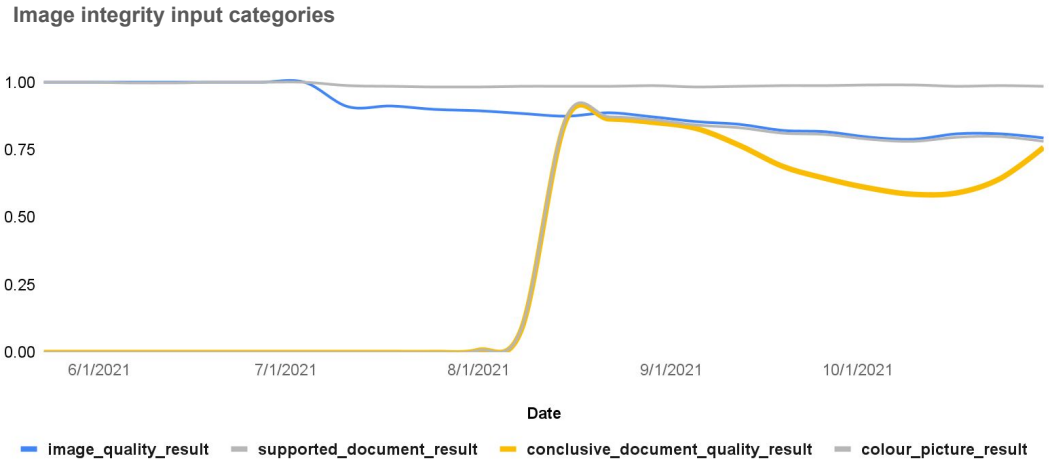
For the purposes of this case, we are assuming that the IDV Verify API was NOT deprecated any point during which this data was collected (it is now deprecated).

Image integrity inputs reveal drop in overall image quality

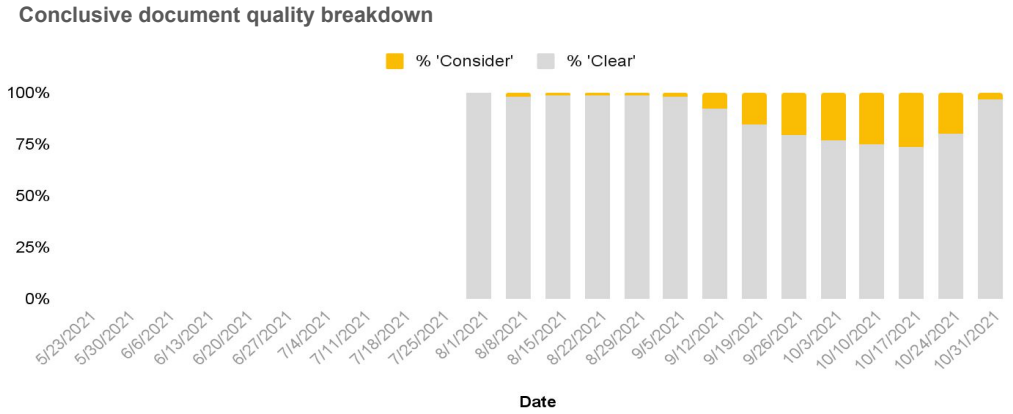


Source: Venn Data Set; IDV Verify API html source code

Image integrity inputs reveal drop in overall image quality



And a decrease in the number of records fit for fraud inspection



Source: Venn Data Set; IDV Verify API html source code

Decreasing Image Integrity

ROOT CAUSE: Coloured Picture Checks

The drop in image integrity, one of the main factors driving declining pass rates, seems to be due to the quality of uploaded images growing worse, so much so that an increasing amount of records cannot even be evaluated for fraudulent activity.

This is likely due to the **“coloured_picture”** check - the overall drop in image integrity contrasted with the minimal changes in the other variables suggests that “coloured_picture” is the root cause.

Decreasing Image Integrity

PROPOSED SOLUTIONS:

Based on the identified root cause, we recommend that Venn:

1. coordinates with IDV Verify to ensure that the “coloured_picture” check is accurate

2. There are many customers who have not completed 2 checks. Moving forward, we recommend that Venn only flag records as “Rejected” AFTER 2 complete attempts for each customer.

Decreasing
Image Integrity

Represented by “Rejected” records from previous slide

Increasing Non-Fraudulent Failures

Represented by “Caution” records from previous slide

Increasing Non-Fraudulent Failures

Checks pertaining to non-fraudulent characteristics:

- **Data Validation:** Asserts whether algorithmically-validatable elements are correct.
- **Data Consistency:** Verifies if data appearing in multiple places on the document is consistent
- **Data Comparison:** Compares data extracted from the document to the data provided by the applicant through the API or applicant form
- **Police Record:** Asserts whether the document has been identified as lost, stolen, or otherwise compromised in police databases
- **Compromised Document:** Asserts whether the image of the document has been found in IDV Verify's internal database of known compromised documents.

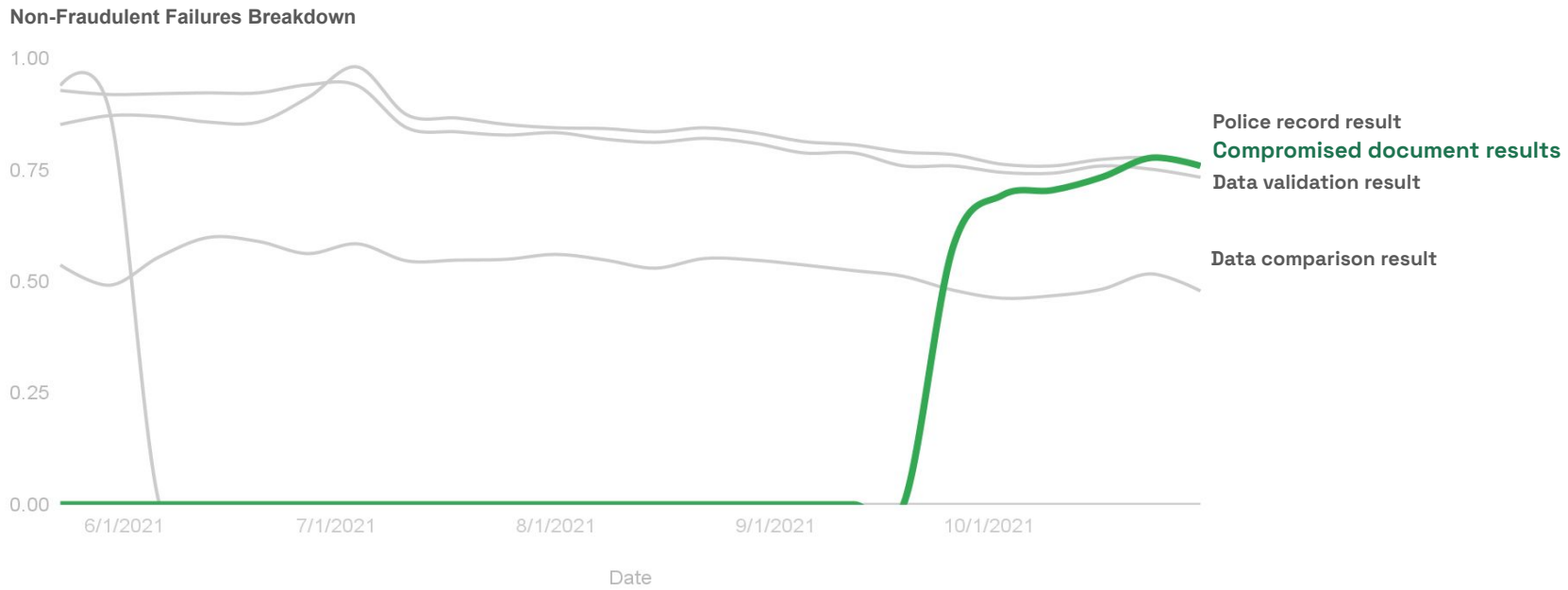
Increasing Non-Fraudulent Failures

Assumption:

The blanks in the data are not due to net new checks/variables being introduced, but simply missing data.

We have decided NOT to impute missing data, so as not to introduce bias into this data.

With only minimal drops in data validation & consistency metrics, it's likely that the increase in "Caution" records is driven by compromised documents



Source: Venn Data Set; IDV Verify API html source code

Increasing Non-Fraudulent Failures

ROOT CAUSE: Compromised Document Results

The uptick in the number of records being flagged as “Caution” contrasted with the minimal drops in “data validation”, “police record results”, and “data consistency”, it seems likely the root cause here is more **“compromised document results”.**

Increasing Non-Fraudulent Failures

PROPOSED SOLUTIONS:

To minimize errors resulting from data quality checks, we recommend that Venn introduce a new process for “Caution” records:

- Such records should be flagged to an IDV Verify (human) employee, who is tasked with comparing these records to their existing database of known compromised documents. Each record may need to be assessed on a case by case basis.

venn

Thank You!