# Operation Challenge Report

Created by: Mu Yuhong
04-Apr-2019

## Table of Contents

## Executive Summary

This report aims to understand the possible causes of KYC pass rate drop. The report consists of three sections. The first section explains what happened and the significance of the issue. The second section illustrates the steps in identifying the possible causes. The third section covers recommendation to address this issue.

The investigation results suggest that the pass rate drop is strongly correlated with Image Integrity result in document check process and may not be related to the actual quality of documents uploaded by perspective customers. A breakdown investigation on Image Integrity result further reveals that there are two features associated with. From 11-Jul-2017 onwards, Image Quality is correlated with Image Integrity pass rate, while from 18-Oct-2017 onwards, the Conclusive Document Quality result comes into play and is strongly correlated to Image Integrity result. Moreover, The Conclusive Document Quality result is a new feature introduced since 2-Aug-2017.

To summarise, the drop is probably caused by technical issues from KYC vendor, Vertitas. Moreover, since the study is only based on quantitative analysis, a further investigation with Veritas is highly recommended to confirm root causes.

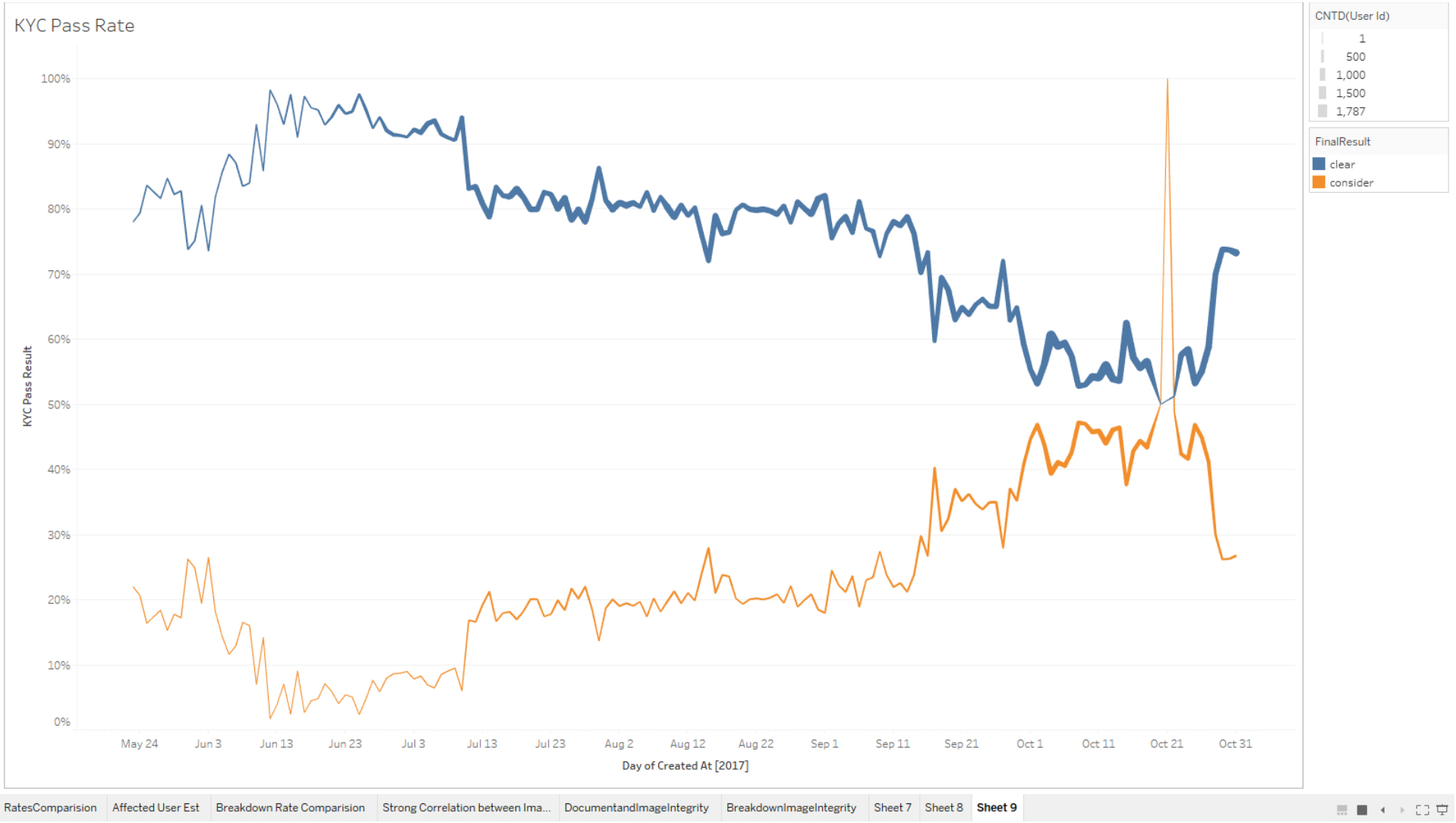
To prevent future failures, I recommend a further study with Veritas to confirm root cause, to provide workaround to mitigate impact in short-term, and to implement feasible long-term solution such as rigorous API performance test and partial group AB test before fully rollout for API upgrade, etc. Meanwhile, I recommend Client to evaluate Veritas' capability and to explore other service providers if necessary.

## Background

In an age of increasing financial regulation and a high risk of identity fraud, Client requires prospective customers to go through a Know Your Customer (KYC) process as part of customer onboarding experience to verify their identity before allowing them to make transactions with Client accounts. To ensure regulation compliance and a smooth onboarding experience, prospective customers need to pass both identity document check and facial similarity check. Client relies on Vertitas, a KYC solution provider, to perform these checks.

One critical success matrix of KYC process is checking pass rate. As prospective customer volume keeps increasing, a consistent and reliable pass rate become extremely important for the company's growth.

However, after examining the data from both identity document check and facial similarity check, it's obvious that pass rate has been dropped substantially since July-2017 and reached its bottom from 20-Oct-2017 to 22-Oct-2017, during which the KYC service was nearly down as shown in Figure 1.

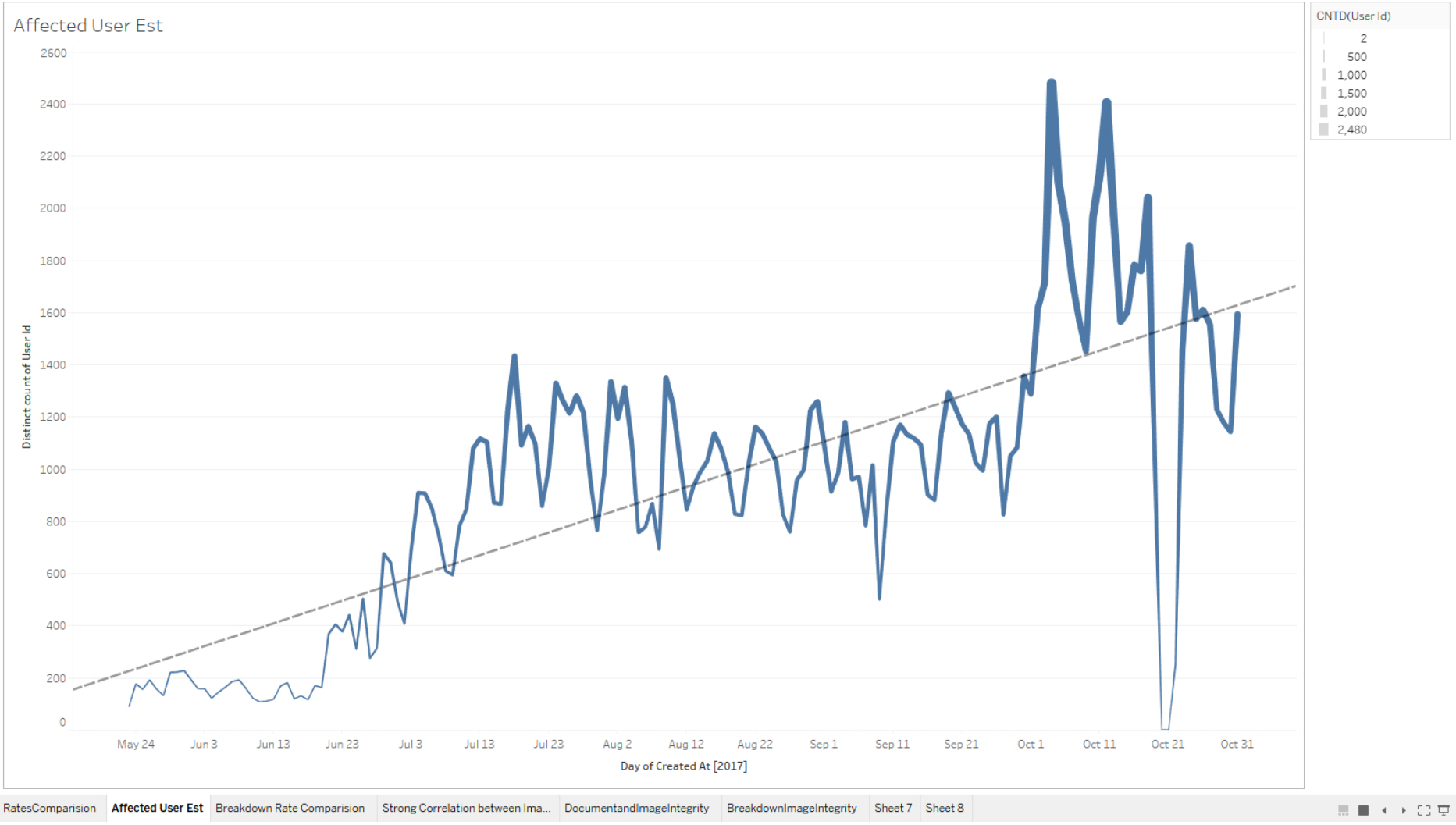



*(Figure 1: Pass rate trend of KYC)*

If the data is accurate, this service disruption undoubtedly frustrated our prospective customers and result in high unnecessary expenses. Based on history trends, it is estimated that at least 1,400 users each day were unable to register on Client due to the drastic decrease of the pass rate.

Consequently, these registration failures might result in high volume of service tickets and burden the customer service. Additional tickets may also incur high cost if Client is using third party ticketing service such as Zendesk etc.

The KYC service needs to be restored as soon as possible, and the root cause needs to be identified and resolved quickly.




*(Figure 2: Number of affected users due to the pass rate drop)*

## Identifying Root Cause and Contributing Factors

The first step of the diagnose is to identify all the factors that contributed to the pass rate change. There are two major types of contributing factors here:

**Internal**. These are actions that the company has taken which resulted in changes of the verification process, within the verification procedure. Examples include tighten of the verification policy, etc.

**External**. These are things that happen to Client, regardless if the company wants them to or not. Examples include third party API service upgrade or disruption, sudden increase of fraudulent users, competitor anonymous attack, etc.

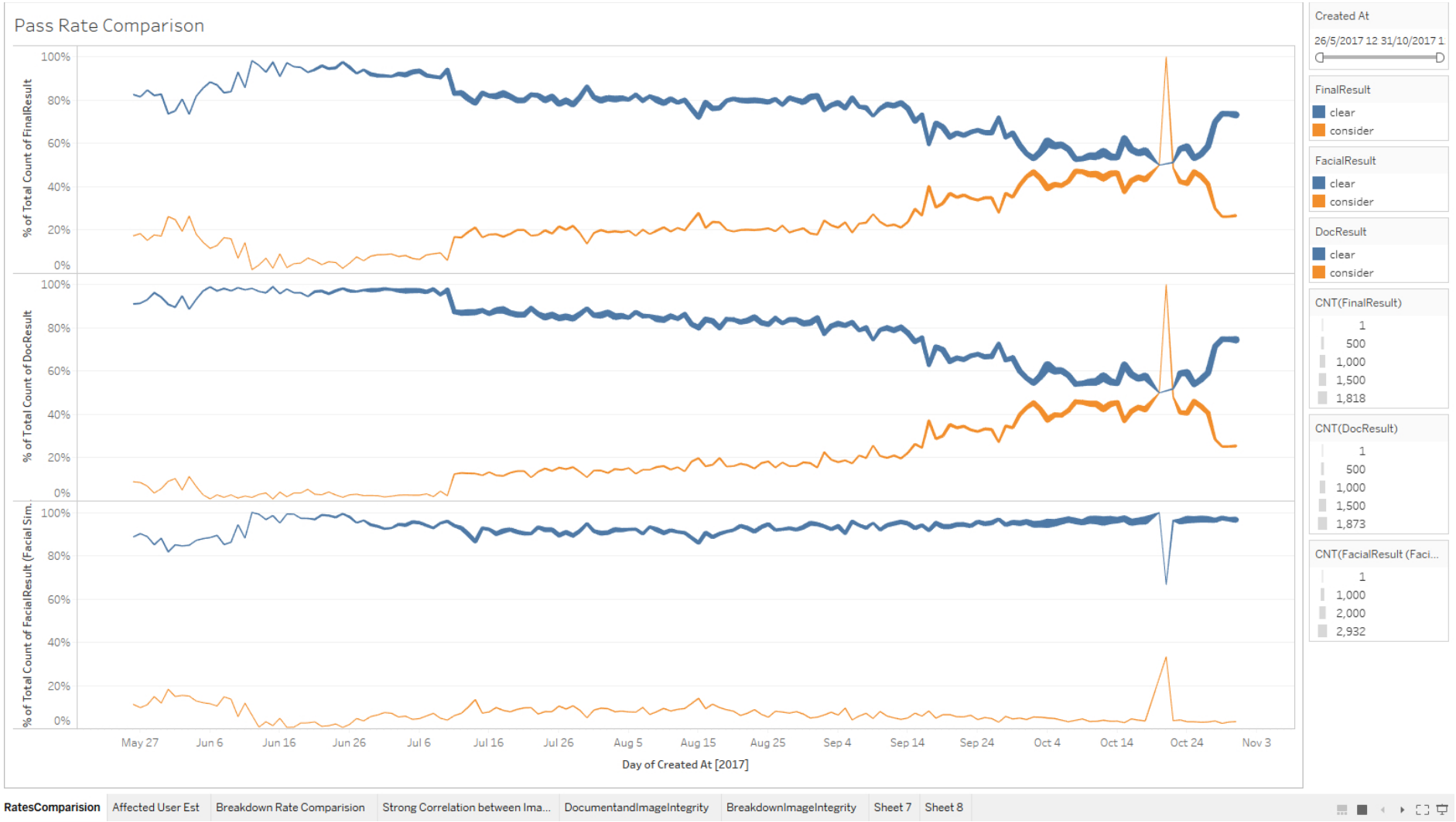
In this case, the investigation will be focused on external factors since there is no internal process changes implemented within the investigation period. The first step of investigation is to derive KYC result by combining both facial similarity check results and identity document check results. Those who pass both checks will be considered Pass.

How do these three rates correlate? Could both checks positively correlated to the drop of final pass rate?

To answer the question, transformation and visualization of these three factors are required. Instead of looking into each single factor, a combined visualization is created. The detailed steps and findings are listed below.

### Step 1: Data Preparation and Visualization

1) Derive results by combining facial similarity check results and document check results.
2) Visualize data by plotting pass rate of below three factors by attempt date.
    a. Facial similarity check pass rate
    b. Document check pass rate
    c. Final KYC pass rate
3) The chart is show in Figure 3.




*(Figure 3: Comparison of KYC pass rate, facial similarity pass rate, and document check pass rate)*

The visualization gives intriguing insights. The document check pass rate is dropped along with the final pass rate, while facial similarity check pass rate remains relatively constant during the same period and recovered quickly from the disruption on 22-Oct-2017. The plot suggests a strong, positive association between the final pass rate and that of document report check result.

According to the API document[1], the document report can be further broken down into seven components, Visual authenticity, Image integrity, Data validation, Data consistency, Data comparison, Police record, and Compromised document. Each component represents different features that used by Veritas to predict whether the report result should be pass or consider.

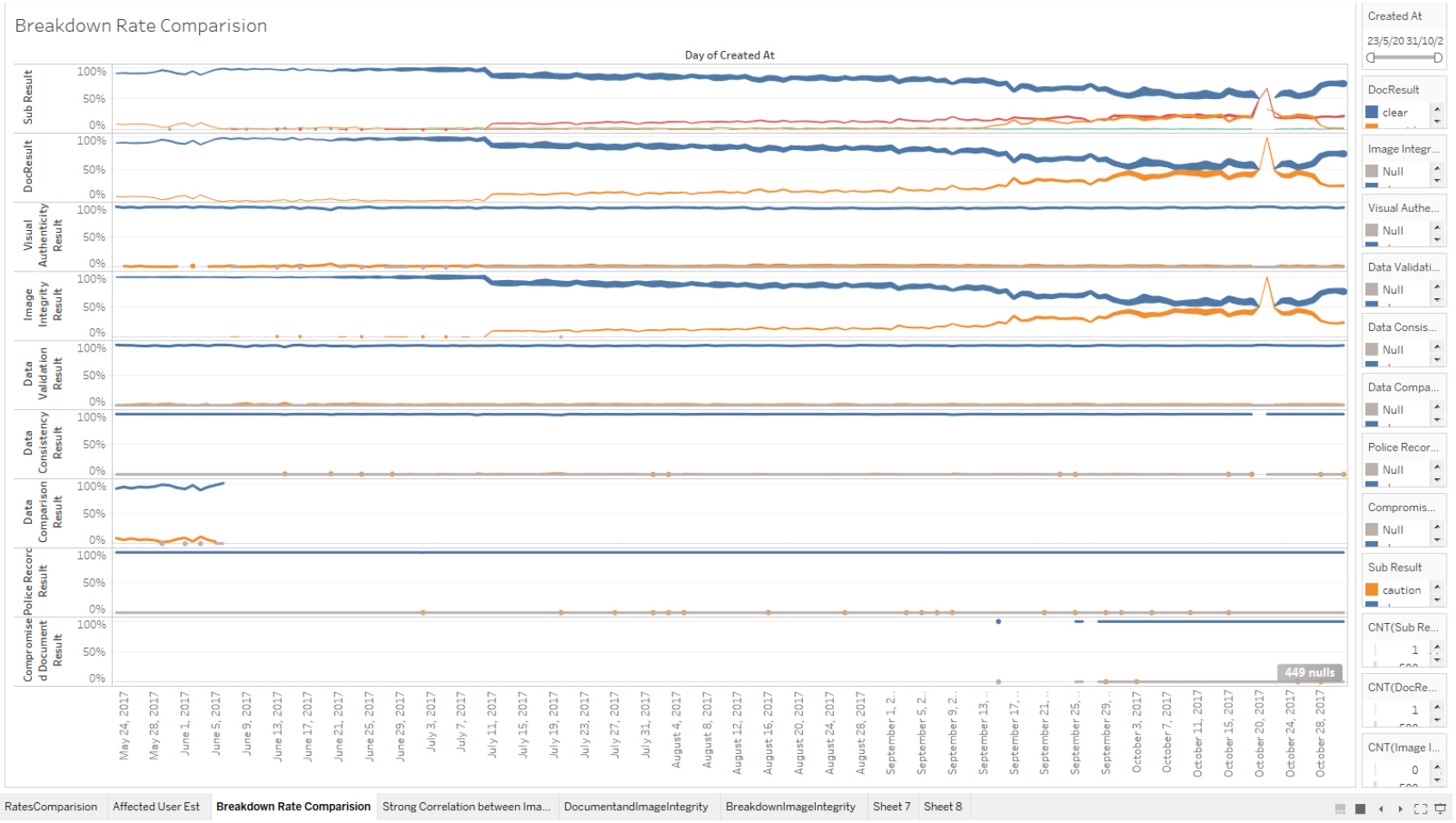
Hence, a deeper analysis is required to investigate the correlation between the seven components and the document report result.

## Step 2: Breakdown Document Result and Examine Pass Rate of its Components

1) Visualize data by plotting pass rate of below factors and document report result by attempt date.
   a) Visual authenticity
   b) Image integrity
   c) Data validation
   d) Data consistency
   e) Data comparison
   f) Police record
   g) Compromised document

2) The chart is show in Figure 4.

[1]Refer to Vertitas whitepaper section: Report Types> Identity Verification> Document Report




*(Figure 4: Comparison of document report rate and its seven components)*

The document pass rate is strongly correlated with Image Integrity pass rate, suggesting that the document uploaded by user was of insufficient quality to verify according to the API. While the pass rates of other components remain relatively constant, suggesting weak correlation with document checking failure.

The result of Data Compassion component is not captured since 07-Jun-2017, however, according to Veritas specification, the field value should be captured. Whether it is caused by technical issue? or other reasons? A further study with technical team is required. Without further study, we can't exclude the component out of cause investigation process.

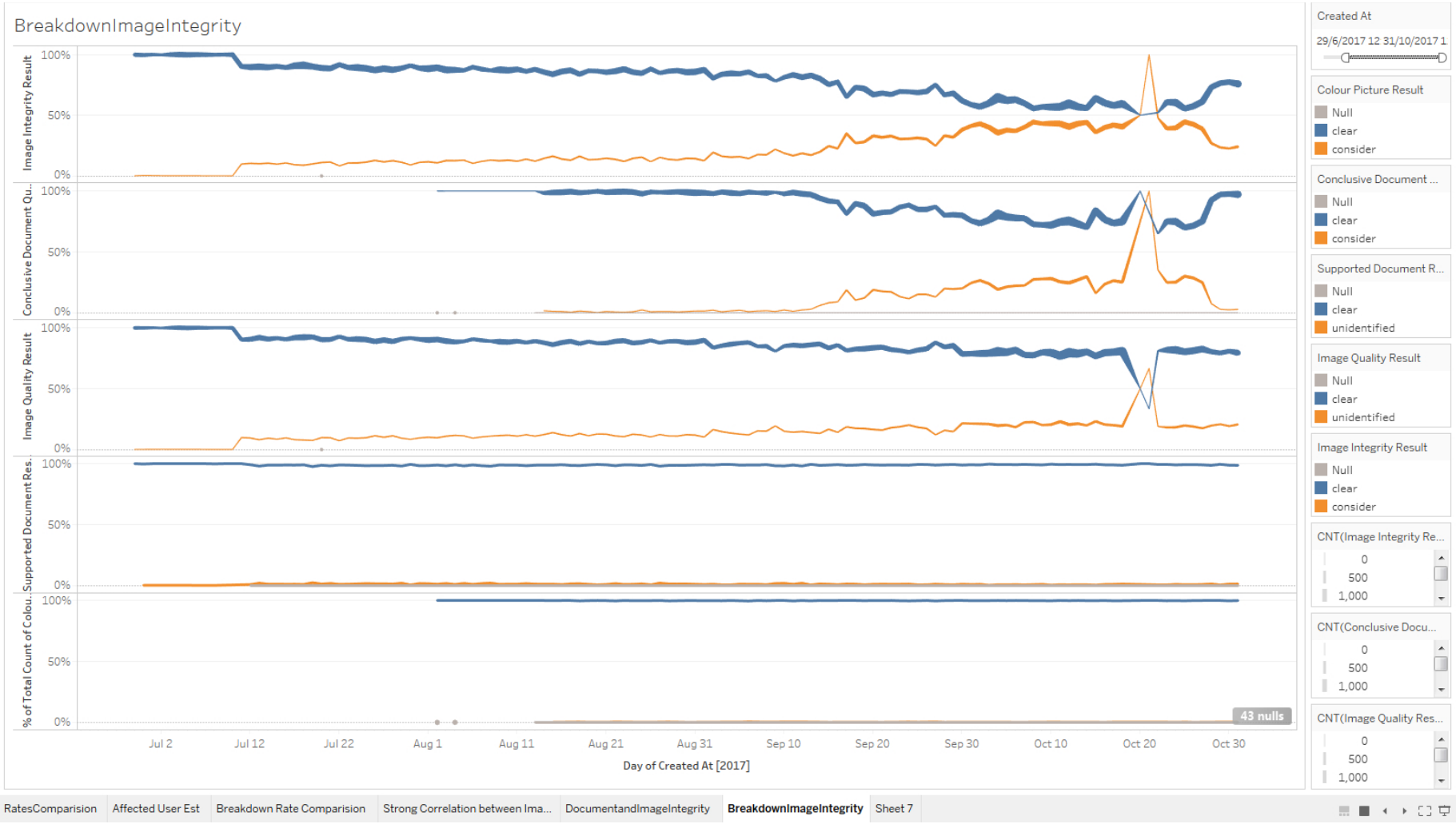
According to the API document[2], the Image Integrity can be further broken down into 4 components - Supported document, Image quality, Colour picture, and Conclusive document quality. Each component represents different features that used by Veritas to predict whether the Image Integrity result should be pass or consider.

As the investigation continues, the Image Integrity is further broken down, and a deeper analysis is required to investigate the correlation between the four components and the Image Integrity result.

### Step 3: Break Down Image Integrity Result and Examine Pass Rate of its Components

1) Break down Image Integrity checking into 4 components3, and compare Image Integrity pass rate with that of below components.
   a) Supported document
   b) Image quality
   c) Colour picture
   d) Conclusive document quality
2) The chart is show in figure 5.

[2]Refer to Vertitas whitepaper section: Report Types> Identity Verification> Document Report
[3] Refer to Vertitas whitepaper section: Report Types> Identity Verification> Document Report




*(Figure 5: Comparison of Image Integrity rate and its 4 components)*

The Image Integrity pass rate dropped along with image quality pass rate from 10-Jul-2019 onwards, while from 13-Sep-2017 onwards the drop of Image Integrity pass rate is also correlated with conclusive document quality pass rate.

Further, the Image Quality pass result dropped drastically from 18-Oct-2017 to 22-Oct-2017 and recovered quickly afterwards. The Image Quality drop could be the closely related with service disruption.

Moreover, it is noticed that the conclusive document quality is a new feature introduced on 2-Aug-2017, and data is only captured from 2-Aug-2017 onwards. A further investigation with Vertitas is required to check if the new feature indeed caused issue.

These observations suggest that the textual elements of documents user uploaded are of inferior quality for Vertitas to perform a fraud inspection. For example, some parts of the documents except for non-textual part are not visible for algorithm to detect.

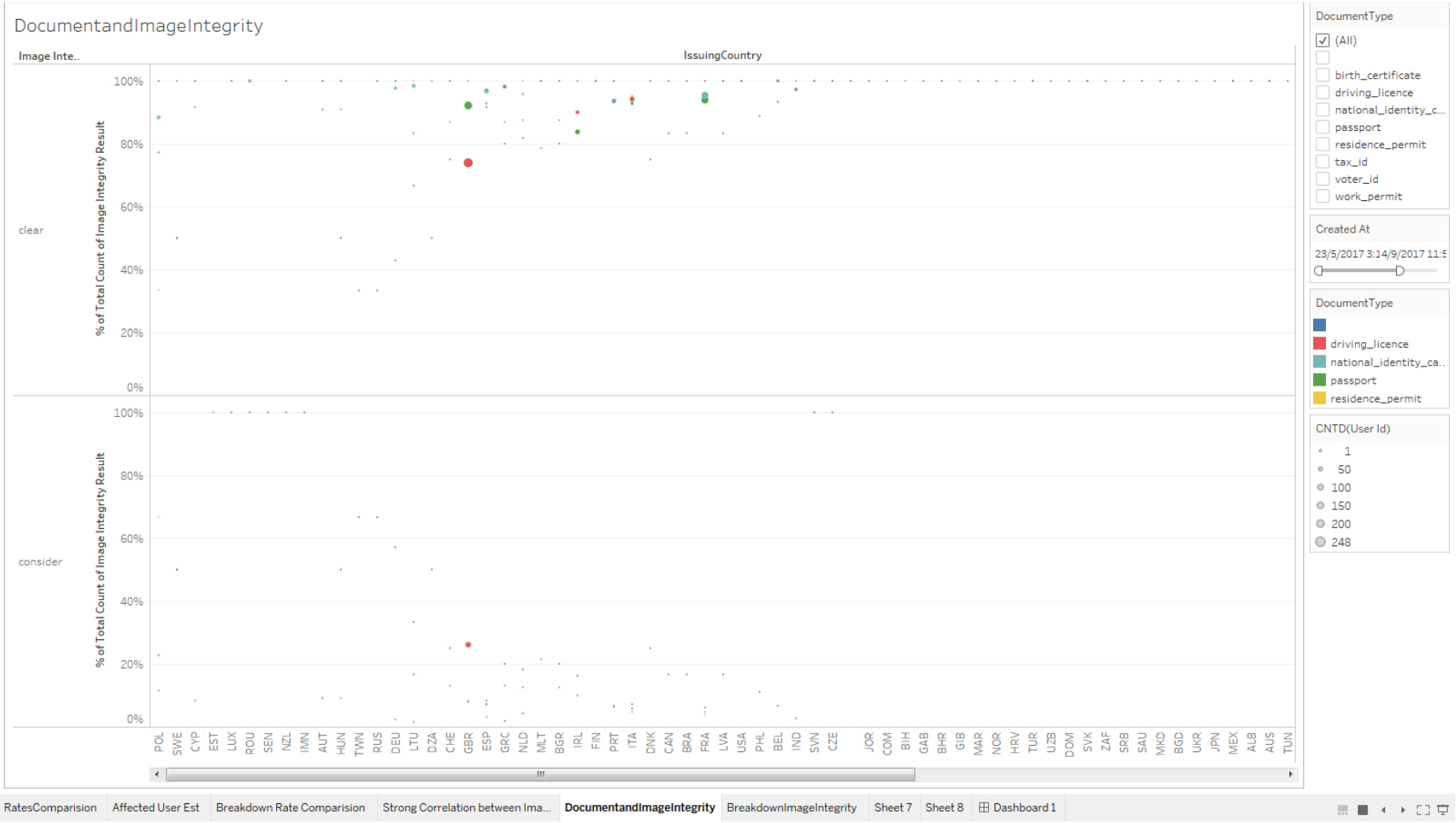
Could Image integrity is correlated with Image Quality? How these two features are defined and assessed? Since the white paper did not reveal these technical details, it is recommended to raise the issue to Vertitas team to further investigate.

Why can't identity documents check detect textual elements? Could it be caused by technical changes from Vertitas? Or a sudden change of identity documents made the algorithm unable to detect textual elements on documents? To answer this question, I further explored the correlation among issuing country, document type, and pass rate of before and after incidence period.

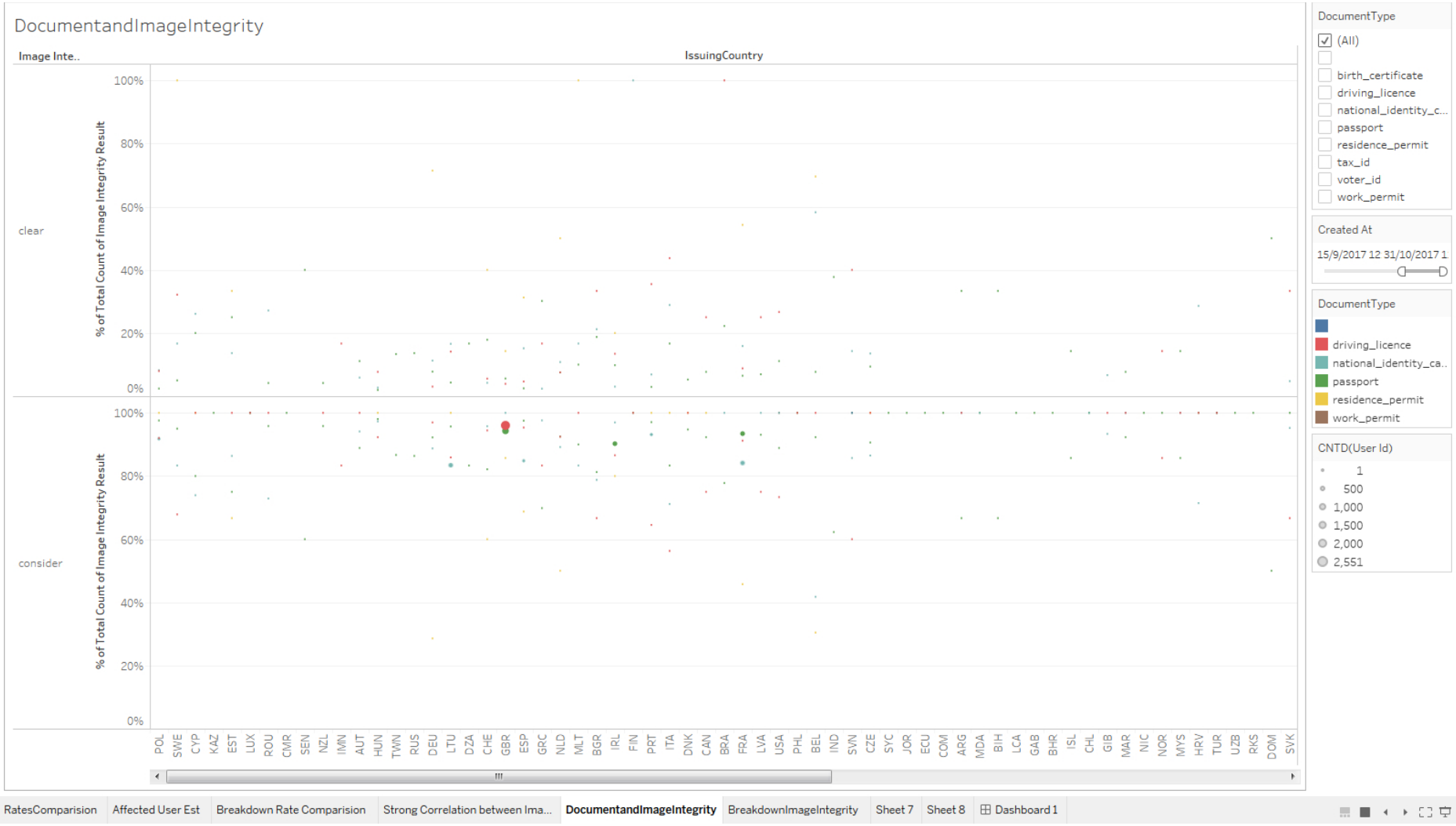
## Step 4: Explore Document Type and Image Integrity Correlation

1) Extract Document Type and Issuing Country from Property 4field.
2) Plot number of attempts, Image Integrity pass rate and Image Integrity consider rate by issuing country, and dates.
3) Slice 2017 in two periods, before period and after period. Before period refers to dates before 15-Sep-2017, after period refers to dates after 15-Sep-2017
4) The charts are generated for before and after and shown in figure 6 and figure 7 respectively.

[4] Refer to Vertitas White Paper: any data extracted from the document through OCR is returned in the properties attribute.




*(Figure 6: Correlation between Document Type and Image Integrity Correlation)*



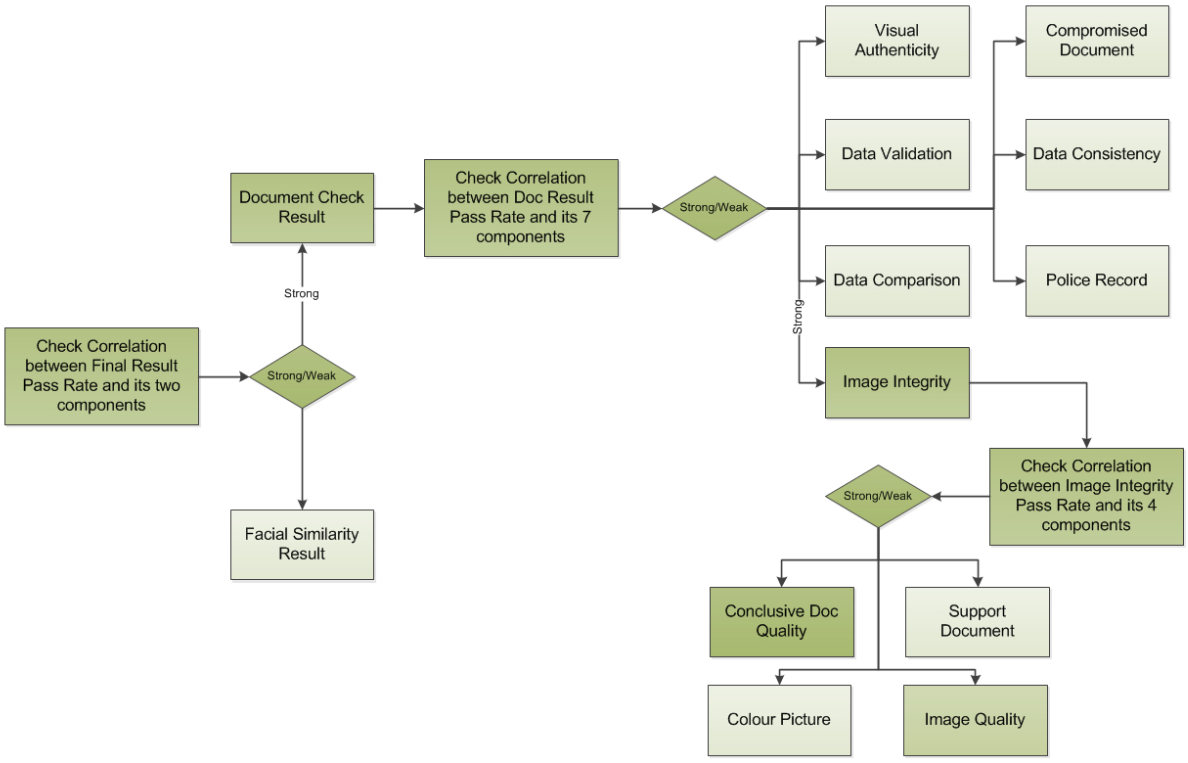

*(Figure 7: Correlation between Document Type and Image Integrity Correlation)*

There is no obvious pattern observed for a specific issuing country or document type. Taking GBR for example, where most of Client users come from, majority customers use driving license for KYC verification. Before 15-Sep, the pass rate is more than 60%. However, after 15-Sep, the pass rate dropped to less than 10%.

With document type, issuing country unchanged, a clear pattern is surfaced after comparing before and after, the pass rate dropped substantially in after period irrespective of issuing country and document type.

In conclusion, the pass rate drop is strongly correlated with technical issues in document check process and is not related to the quality of documents uploaded by perspective customers.

The overall investigation process is shown in figure 8.




*(Figure 8: Overview of Investigation Process)*

## Conclusion

Given above findings from data analysis, it can be positively concluded that the fall of KYC pass rate is not related to documents uploaded by perspective customers. It could be caused by technical issues from Veritas. Moreover, since the study is only based on quantitative analysis, a further investigation with Veritas is highly recommended to confirm root causes.

## Recommendation

To further confirm investigation results, we need to raise the issue and share our investigation findings to Vertitas. The investigation can be used to facilitate troubleshooting, but a though checking by Vertitas is required to confirm the root cause. Meanwhile, a workaround should be implemented immediately to mitigate impact. If the newly introduced feature -conclusive document quality feature - indeed caused the issue, we can explore by using AB testing to see if it is possible to turn off the feature while ensuring KYC proceed smoothly.

Internally, we need to identify perspective customers who were affected by the service disruption, apologize, and invite them to try again once service is restored.

Moreover, we need to communicate with Vertitas that if there are any feature changes in future, we should be notified as well. Prevention is better than cure.

Lastly, we can take this opportunity to evaluate the performance of Veritas and explore other service partners if possible. The perspective customer number in Client kept growing. Our service provider needs to upgrade themselves as well to get prepared as well.

## Appendix:

i. The investigation process overview, steps, and findings can be found here: Steps.docx
*ii.* The data analytics is complete in Tableau and Excel. The tableau workbook can be accessed from here: Tableau workbook
*(Note: You need to use Tableau Desktop or Tableau Reader to view the file)*
iii. The figures in this report can be accessed from here: Charts