Quality Control in Manual Data Collection

Osman Seker

2022 Laurea

Laurea University of Applied Sciences

Quality Control in Manual Data Collection

Osman Seker

Bachelor of Business Administration

Thesis

November, 2022

Laurea University of Applied Sciences **Abstract**

Degree Programme in Service Business Management

Bachelor of Business Administration

Osman Seker

Quality Control in Manual Data Collection

Year 2022 Number of pages 35

The purpose of the thesis project was to design and implement a quality control process for manual data collection. The client company for this thesis was Nordea Bank. After the implementation, the collected data as part of quality control will be analyzed to gain an understanding of the quality level in the unit. The unit will be provided with improvement ideas or mitigating actions based on the analysis findings.

A quantitative research method was used to support the analysis by collecting data for three months in order to have a sufficient amount. Literature regarding quality control can be found for physical products made in a factory environment, but it was difficult to find supporting material for this thesis. Therefore, KPIs and other internal agreements were a way to measure success.

In the unit of Data Collection, no manual quality control had been established before this thesis. The only form of quality control was done with automatic quality control, which was not capable of checking for errors in the data.

The process of quality control was fully implemented in four different countries during December 2021, and the data collection for the analysis begun in January 2022. Data was collected for three months and in April of 2022, the analysis begun. The results were alarming as the results showed that the actual error rate was three times higher than expected.

Analysis continued narrowing down to specific areas in a case, which are most prone to errors. Together with the Team Leads of the Data Collection teams, the author was able to provide the teams with mitigating actions to reduce the error rates.

The process was not implemented only for the purpose of this thesis, and is still continuing as an important part of Manual Data Collection. KPIs were implemented and are reported on continuously for stakeholders and senior management.

Keywords: Quality Control, Data Collection, Process Design

Contents

1 Introduction

This thesis has four objectives: Design- and implement a Quality Control process, analyze the results to develop an understanding of the quality levels in Credit Data Collection and provide improvement ideas of mitigating actions to the teams based on the analysis results. Prior to December 2021, Nordea's Credit Data Collection unit has not had any structured Quality Control process implemented. The process was done in collaboration with several units within Nordea: Credit Services, Operations Change Execution, Risk Models and Group Internal Audit.

Process of setting up the control begun in July 2021 with the goal of having a well-working control fully implemented by April 2022.

1.1 Company description

Nordea Bank is the biggest financial services group in the Nordics. It is formed by acquisitions and mergers of the Finnish, Swedish, Danish and Norwegian banks that happened between the years of 1997 and 2001. Nordea employs approximately 26 thousand people of which five thousand are working in the company's headquarters in Helsinki.

Nordea is serving almost ten million private customers and over half a million corporate customers of which 2650 are large corporates and institutes. According to Nordea's annual report (2021) in December of 2021, assets under management (AUM) were reported to be €411 billion.

After finalizing their Baltic operations in 2019, the Nordic countries are considered to be Nordea's home market.

The organization consists of five business areas: Personal Banking, Business Banking, Large Corporates & Institutions, Asset & Wealth Management, and Group Business Support (Nordea 2022).

1.1.1 Purpose and values

"We enable dreams and aspirations for a greater good" can be seen in Nordea's marketing in television or on the internet. The company claim that their purpose and values combine the daily endeavors help customers reach their hopes and dreams.

The company values can be seen in their culture, where people serve customers with joy, work in collaboration across the organizational borders, hold genuine accountability of their work, raise discussion topics boldly and challenge each other.

The four values at Nordea are:

- Collaboration – For the common good
- Ownership – It starts with me
- Passion – To serve our customers
- Courage – To do what is right

1.2 Thesis description

This thesis seeks to understand the level of quality in the collected data in Nordea's Data Collection unit by designing and implementing a Quality Control process. The thesis will design the process in collaboration with a Team Lead in Credit Data Collection, establish Service Level Targets and Key Performance Indicators for quality with the stakeholders, train the employees to match the requirements of the process, create mitigating actions with Team Leads if the Service Level Targets are not met, and create a reporting structure that allows daily result monitoring for the Senior Management.

Thesis requires collaboration between multiple units within Nordea and multiple process designs and documents need to be approved by the Senior Management.

1.3 Research objective

The purpose of the thesis is to understand the level of quality in Credit Data Collection by establishing Quality Control process for error screening on a sample of cases. When the understanding has been gained, root-cause analysis needs to be performed in order to understand the errors better. This thesis aims to provide beneficial data for the data collection teams to understand their quality level as well as providing improvement ideas or mitigating actions for more accurate data collection.

Quality Control framework is formed by using theoretical literature to support the requirements from Group Internal Audit and Senior Management. Data is collected between January 2022 and March 2022 in order to have sufficient amount of data for any further analysis.

1.4 Thesis structure

This thesis consists of an introduction to the company, the unit of Credit Data Collection and its stakeholders. It will describe the purpose of the unit and the core tasks and the importance from a high-level perspective. This thesis will largely focus on the quality aspect as the objective is to provide more understanding on the level of quality, by introducing a brand new process. Quality Control process is designed and implemented fully by the author.

This thesis will gather sufficient amount of data to perform an analysis, which aims to highlight pain points in the data collection process. Highlighted areas will be investigated further and a root cause analysis will be performed together with the Team Leads using a Fishbone diagram.

The thesis ends with providing improvement ideas or mitigating actions depending on the analysis results according to the thesis objective.

2 Credit Data Collection in Nordea

Credit Data Collection is an unit under Group Business Support consisting of approximately sixty employees. Its primary purpose is to collect customer data from bank's internal systems according to European Central Bank's requirements. Data is collected manually from both retail and corporate customers, who are in default.

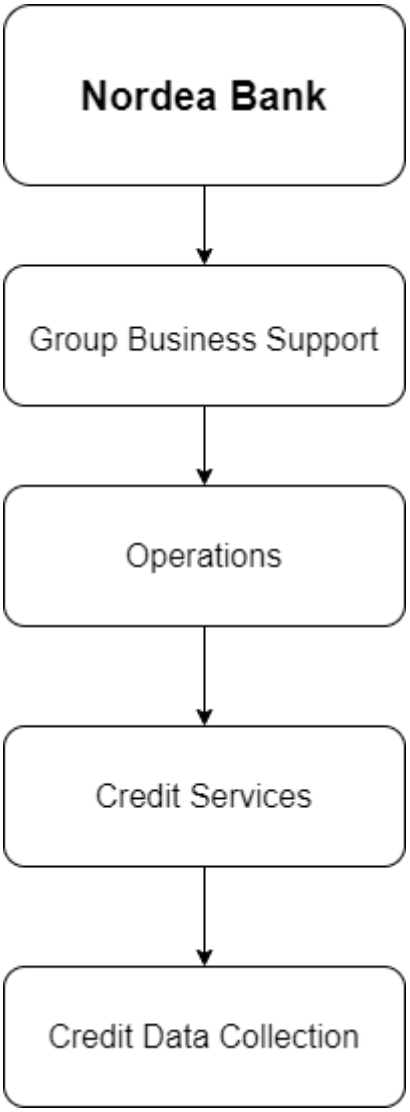

Figure 1: Organization structure

Credit Data Collection consists of one Head of Unit, three Team Leads, four Business Developers and approximately fifty Data Collectors. The unit is divided into five teams: Finnish-, Swedish-, Norwegian-, Danish-, and Business Development team. All teams collect data from their country's customers, except for the Business Development team as their task is to develop the whole data collection process as well as drive other tasks such as reporting.

Most of the employees are working in the Tallinn office as only seven people are living outside of Estonia in Finland and Denmark.

2.1 What is Credit Data Collection?

The first team in Credit Data Collection was introduced in 2018 as the European Central Bank started requiring financial institutions to perform data collection from defaulted customers. Data collection is to be done in all large financial institutions in the EU. For the first year, the

data collection was done fully manually with a team of external consultants, but as the requirements grew, Nordea realized the need for an actual team. That team was then later in 2019 created in Tallinn. Nowadays all Nordic countries have their own teams which are responsible for collecting data from their customers.

Credit Data Collection in Nordea gathers data from various banks' systems and external sources to develop an overall picture of the status of a defaulted customer. A default customer is a person or a business who's bank rating has dropped below 1 on a scale of 6 to -2. The rating is formed by many factors such as:

- Amount of products and services
- Meeting the loan repayment deadlines
- Unusual financial behavior etc.

The most common reason for a rating drop is missing a repayment deadline. A customer becomes defaulted once they have missed a repayment three times (three months) in a row. After those missed repayments, the loan is then sent to an external debt collection agency. Data collection in Nordea starts from the day that the customer has defaulted.

Until April of 2022, the data was collected to an Excel template which had 87 data fields unevenly distributed between the five sheets, that could be filled. The template was divided into five sheets:

- Customer information – Information on the customer such as: the default date, name, age, customer ID etc.
- Collateral information – Information on the possible collateral such as: Collateral ID, the type of collateral, valuation etc.
- Exposure information – Information on all products that the customer has such as: exposure ID, Amount of exposure at default, accrued fees at default etc.
- Transaction information – All transactions from those exposures are documented since the default date.
- Recovery process status updates – All status updates in the recovery process are documented since the default date.

By May of 2022, all four teams have moved on from using Excels to filling in the collected data into a low-code platform called PEGA. The same five sheets of data collection are still in place as well as all data fields.

2.2 Stakeholder introduction

Once the data is collected on a defaulted customer, it is sent to another unit in Nordea called Risk Models. They are responsible of creating models based on the data that has been provided to them.

When designing the Quality Control framework, in addition to receiving help from a colleague, Operations Change Execution provided assistance. Implementing a brand new process requires project management knowledge and they were able to provide that.

The whole process of implementing Quality Control lasted for nine months. During those months, the process required multiple acceptances from the senior management. Those people were the Head of Credit Data Collection, Head of Credit Services, multiple Analysts in Risk Models, Risk Manager of Credit Services and Auditors from Group Internal Audit.

The main stakeholder in this project was Risk Models as they are the end-user of the collected data.

2.2.1 Risk Models

In the organization chart, Risk Models gets separated from Credit Data Collection after Operations. They belong to Operations, but are not part of Credit Services.

Risk Models develops, maintains and ensures the performance of Nordea's internal risk models for credit, market, counterparty credit risk, and collective impairments (Charter for Risk Models 2022).

3 What is Quality Control?

"Quality control is the set of measures and procedures to follow in order to ensure that the quality of a product is maintained and improved against a set of benchmarks and that any errors encountered are either eliminated or reduced" (Techopedia, para. 1). Quality Control is constant testing of the product to determine that the product manufacturing is consistent and in line with the customer requirements.

The use of clearly defined controls is one of the features of quality control. In most organizations, quality control is provided by a department (either internal- or third-party team) that gives a set of standards to be followed for the products. The process relies on testing and inspection of the product, a clear picture can be gained of the quality.

According to Simplilearn (2021), quality control has two main goals:

1. To ensure that the products are as uniform as possible
2. To minimize errors and inconsistencies within them

Quality control itself does not have a uniform or an universal process, because quality is relative. As an example, in food and drug products, errors can put people at risk and therefore create significant liability for the company. These types of industries must rely heavily on scientific measures, while others may require a more qualitative methods.

There are multiple benefits for an organization to have a quality control process in place. Investments in quality control measures protect the reputation of the company, prevent products from being unreliable and therefore increase the trust of the consumers. The process ensures that a company can look at evidence-based data and research to ensure that the products are living up to their standards, making the quality control necessary (Simplilearn 2021).

Quality control is often interchangeably with quality assurance. Although these terms are similar, there are distinct differences between these two concepts. Quality assurance and quality control are both aspects of quality- or risk management. Typically, quality assurance activities and responsibilities cover virtually all of the quality system in one fashion or another, while quality control is a subset of the quality assurance activities. Quality assurance is providing confidence that quality requirements will be fulfilled, while quality control is focused on fulfilling quality requirements. Quality assurance relates to have a process is performed and quality control is more the inspection aspect of quality management.

3.1 Quality Control design in Credit Data Collection

Quality control is a continuous process in Credit Data Collection in which case files are manually screened for any errors in the data collection phase. It’s purpose is to improve the quality of the data that the Credit Data Collection produces for Risk Models. Manual Quality Control screens for errors by comparing data in the case templates to the data in the source systems. It helps to reduce the errors by identifying the root causes and organizing feedback sessions. Quality Control also monitors and escalates the data quality by producing reports for teams and stakeholders.

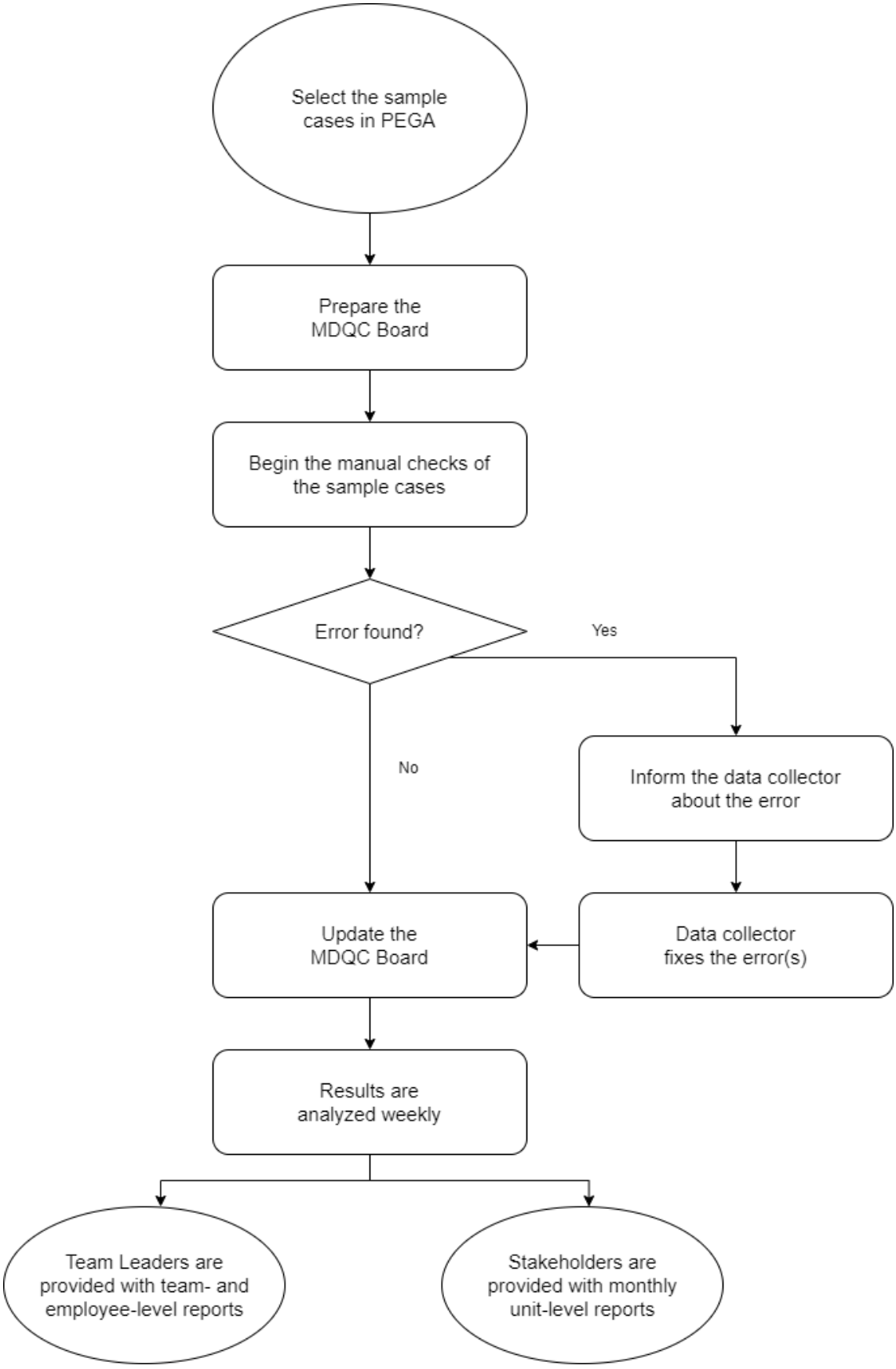

Figure 2: Quality Control flowchart

4 Project plan

Before developing the process any further, a project plan had to be created that states clearly when, how and by who connected tasks are completed. Project plan needed to be done very well as it can be an important defense against the common pitfalls that cause projects to fail. LaPrad (2020) states that, there are large amounts of benefits in having a project plan, such as:

- Stress minimization: Looking at the final product without seeing the steps in between can cause stress. Project plan can be thought of as an instruction manual that guides the team through the steps to success.
- Confidence: Having a well thought of plan with realistic goals can be a confidence booster. Deadlines are set which eliminates the need to wonder about the progress and what tasks should be tackled next.
- Communication: Communication is important in project management and dealing with a team and stakeholders, there's a need to communicate the progress constantly. Project plan can be visible for all, which makes it a very effective communication tool to have.
- Unity of the team: Having a common idea of the final product motivates the team as they are not simply completing tasks without knowing why. All tasks are laid out to the plan which makes the priorities clear and expectations aligned.

At Nordea, all employees have an access to a standard project plan template which was used in this project too. It is primarily used in all projects in the company and can be acquired from Nordea's intra pages. It has four standard slots that need to be filled:

I. Deliveries: This column is for what is intended to be created. These can be thought of as mid-goals on a way to finishing the project.
II. Activities: It is used to describe the activities that are needed in order to complete the deliveries. Larger deliveries often have more than one activity, while smaller ones can be completed with just one activity.
III. Responsible: Each activity that has been listed, need an assigned responsible that is either working on the task or the contact person for it. The role can also be assigned for multiple people to ensure it's progress in case of any issues.
IV. Timeline: Each action needs a timeline of when the action can be expected to start and finish. In a big project, plans are hardly ever waterproof and timelines need to be adjusted from time to time. That is why any deviations need to be marked, so any deadlines or milestones can be adjusted accordingly.

The project plan was done in collaboration with a team leader from Swedish team and a Senior Project Manager from the Operations Change Execution team. The plan was presented to the Head of Unit at Credit Data Collection and the Head of Credit Services and approved afterwards.

The project plan

Delivery	Activity	Responsible	Weeks
Design the Manual Quality Control framework	Create a Control Framework	Osman Seker	31-37
Design a reporting framework to monitor the effectiveness of Quality Control	Creation of the Quality Control Board-file	Osman Seker	31-39
Develop a macro to perform data validation	Define requirements for the macro	Osman Seker & Macro developer	35-41
Document Manual Quality Control framework	Standard Operating Procedure and Framework documents approved	Osman Seker	35-45
Develop a macro to perform data validation	Macro development	Macro developer	37-44
Training all Credit Data Collection teams	Developing a pilot plan and material for training	Osman Seker	39-40
Training all Credit Data Collection teams	Training the pilot staff	Osman Seker	40-41
Training all Credit Data Collection teams	Process testing (pilot) phase	Swedish Team	41-44

Training all Credit Data Collection teams	Gathering feedback from the pilot team	Osman Seker	41-44
Training all Credit Data Collection teams	Process improvements based on the feedback	Osman Seker	43-45
Training all Credit Data Collection teams	Present pilot findings for the senior management	Osman Seker	45
Training all Credit Data Collection teams	Training the remaining teams	Osman Seker	45
Implement the updated Manual Quality Control framework	Quality Control fully implemented in Finland and Norway	Osman Seker	47-50
Implement the updated Manual Quality Control framework	Quality Control fully implemented in Sweden and Denmark	Osman Seker	47-50
Establish a reporting structure	Present the first reporting for all countries	Osman Seker	1
Error Analysis	Data collection for analysis	All teams	1-13
Error Analysis	Data analysis on collected data	Osman Seker	13-15

Table 1: Project plan

4.1 Quality Control framework

Designing a framework for the process was the first delivery in the project plan. Control framework is a formal high-level document which describes the Manual Quality Control process, roles and responsibilities, reporting and mitigating actions. It is an ongoing living document, and it is reviewed quarterly as the process evolves.

The document starts off by describing the purpose of the process and what it aims to do. It then gives an introduction to the case templates and expresses the scope of the process by telling what cases will be selected as part of the sample, by whom and when. After that, all roles and their responsibilities are listed complimented by a RACI chart. The last part of the document is about reporting and mitigating actions.

To create an effective and beneficial reporting, team leaders and stakeholders had to be interviewed. It is important that the reports that will be created are used, and the information gained is valuable for all parties. Once the reporting need was mapped, it was then documented into the framework.

Timeline for this task was from 1st of August until 24th of September. The document had to be fully written and approved by the Head of Credit Services during this timeframe. It ended up being a 10 page Word-document, which was modified multiple times before submitted for approval on the 20th of September.

4.2 Development of an VBA macro in Excel

Second delivery on the plan was to develop an Excel macro to perform data validation on a case. According to Corporate Finance Institute (2022) VBA macros use the Visual Basic Application in Excel to create custom user-generated functions and speed up manual tasks by creating automated processes. Instead of going through each data field in a case, the Data Collectors would be able to click a button and the macro would check the fields and provide feedback in a matter of seconds.

The macro needed to be created so that it could be downloaded by each Data Collector to be used after they have completed a case. The purpose of the macro is to perform automatic checks on the data that has been inserted into the case and providing feedback on the accuracy, so the case can be corrected immediately if errors are found. The goal with creating the macro was to reduce the amount of errors the Data Collectors make.

The process starts with planning as always. Before the coding starts, it is important to gain an understanding of what the macro should do and how. The ultimate goal was that the macro would be able to compare the data in a case with the data in the source systems to verify that it is used correctly. This goal was unfortunately unreachable as there are too many systems in use in different formats. Together with the VBA macro developer, we disregarded the option of building a macro of this kind very quickly.

Instead, we decided to focus on format and logic issues within a case. This means that the macro would check each data field and look for any format or logic related issues (e.g. field for date cannot have any text in it). As mentioned earlier, a case has 87 data fields and the

macro would go through all of them. Data collectors would need to open the macro while having the case file open. Clicking a button on the macro, it would go through the case file in a matter of seconds and provide feedback based on the findings.

The macro was released in a common SharePoint which is used by all teams in Credit Data Collection. I organized a training session for the staff where I presented the macro and gave instructions on how it should be used. The next day, the macro was taken to use in all four teams.

4.3 Quality Control Board

Quality Control Board is an Excel file, which purpose is to collect all the errors found on the Quality Control process. It includes information of all cases that were selected as part of the sample. The board is used as the source of all reporting for Quality Control. The board is filled with data everyday as the manual checks are ongoing. Each team has selected one of two employees to be the Controllers, who are responsible for checking the cases, communicating with the data collectors and filling in the errors to the board. The board will be the source of the data for the quality analysis in the later stages of the process.

The board follows a standardized structure, which is used in all reporting in Credit Data Collection. There are 18 fields that need to be filled per each case, with the first 11 being standard fields:

1. Date – Specific date when the case was checked (e.g. 01.01.2022)
2. Week – Week number
3. Month – Name of the month
4. Controller ID – Instead of using names which are recognizable, we use employee identification ID numbers. This field is for the ID of the person who has done the manual check on the case
5. User – ID of the person who completed the case
6. Country – In which team/country was the case completed in
7. PEGA ID – Cases are completed in the PEGA system and all cases have a PEGA ID
8. Common ID – All cases have a Common ID, which is connected to a specific customer
9. Segment – Is the case Retail/Corporate
10. Secured flg – Does the case include collateral information
11. Exposure type – What types of exposures does the case include
12. Error found? – Does the case have an error found or not
13. Error – Sheet – In which of the five sheet was the error found in
14. Error – Field – In what specific data field was the error found in
15. Error description – Free text description of the error, not mandatory to fill
16. Error value – Insert the value which was found to be incorrect

17. Correct value – Insert the value which is correct
18. Fixed? – Is the case fixed with the correct data or not

4.4 Documentation to support the process

After designing the process, it was required that two crucial documents are prepared:

I. Quality Control Framework
II. Quality Control Standard Operating Procedure (SOP)

Framework was completed earlier while the process was being designed. It was approved by the Senior Management and it only needed to be updated quarterly if any changes were to happen on the process.

The Standard Operating Procedure is a document which describes the step-by-step process that must be taken to perform a routine activity. SOPs have to be followed the exact same way every time to guarantee that the organization remains consistent and in compliance with industry regulations and business standards. Brush (2021) states that SOPs can benefit a business by reducing errors, increasing efficiencies and profitability, creating a safe work environment and producing guidelines for how to resolve issues and overcome obstacles.

Questions that should be answered in the SOP are:

- Who performs what?
- What roles are there and what do they perform?
- What goals do each role have?
- What actions for each role need to happen to reach the goals?

In this project, the importance of a good SOP is crucial. Almost all of the employees in the unit are working from home (or outside the office), and in different countries, which limits the amount and the quality of communication between the people. Therefor, it is vital that the steps and responsibilities are clearly described and presented.

4.5 Staff training

According to the Quality Control framework, each of the four teams must appoint 1-2 Controllers, why are responsible for the manual checking. These people, together with the Team Leads were invited to a training session during week 45 of 2021. Due to the Covid situation at the time, the training could not be held live and it took place in MS Teams.

The training session lasted for 1,5 hours and it was recorded for later use. Training session included topics such as: Introduction to the process, rules for checking, KPIs, instructions for Quality Control board and reporting.

All participants were given a task to give their teams' an introduction to the process, so the Data Collectors would know their roles.

4.6 Creating a reporting structure

In a data driven business environment, it is crucial to have an effective reporting structure established. Before starting to build a structure, it is crucial to understand what areas are going to be followed upon in the future, what areas are valuable in terms of quality and what should be reported for the stakeholders and teams in weekly/monthly/quarterly meetings. After gathering ideas from the Team Leads and the Head of Unit, three main things stood out that need to be reported on:

1. Failure rate – Failure rate expresses how many cases that have been through the Manual Quality Control process have one or more errors in the data. An example: Ten cases were checked and one of the cases had one or more errors. The failure rate would then be 10%.
2. Sample size – Sample size tells how many cases have been selected as part of the sample for Manual Quality Control. According to the Control Framework, each team must select 10% of the completed cases as part of sample and that number needs to be monitored. If the sample size is too small, it cannot represent the whole team. If the sample size is too big, the team would be wasting resources in checking too many cases as 10% is considered to be a representative size.
3. Error distribution – Error distribution highlights the areas in the case, which have the most errors and therefor prone to errors in the future. Finding out this information is extremely important as the teams can have an understanding of where their Standard Operating Procedures are lacking information. It also sheds light on what areas are difficult in the whole data collection process, which can be focused by the Business Developers to seek improvements.

After establishing the reporting needs, it was needed to ensure that all data was available in the Quality Control Board. Reporting sheet was opened on the same file, and the needed charts were built based on the existing data.

Failure rate chart needs seven different data points: Case type, Month, Quarter, Week, Country, Cases checked and cases with errors. Cases checked and cases with errors are shown in the chart, and the rest are used as filters. Here is an example of the results from weeks 15-17 in Credit Data Collection from all teams:

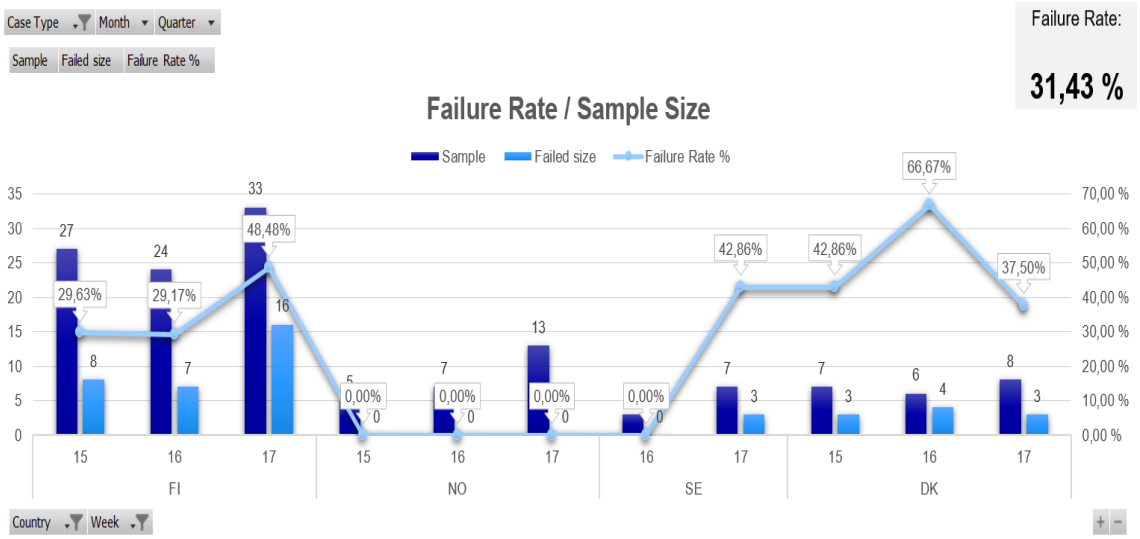

Figure 3: Failure rate example

As the first chart already covers the first two requirements for reporting, there's need for only for more additional chart type to be built. Here's an example of the Finnish teams' error distribution in the different sheets:

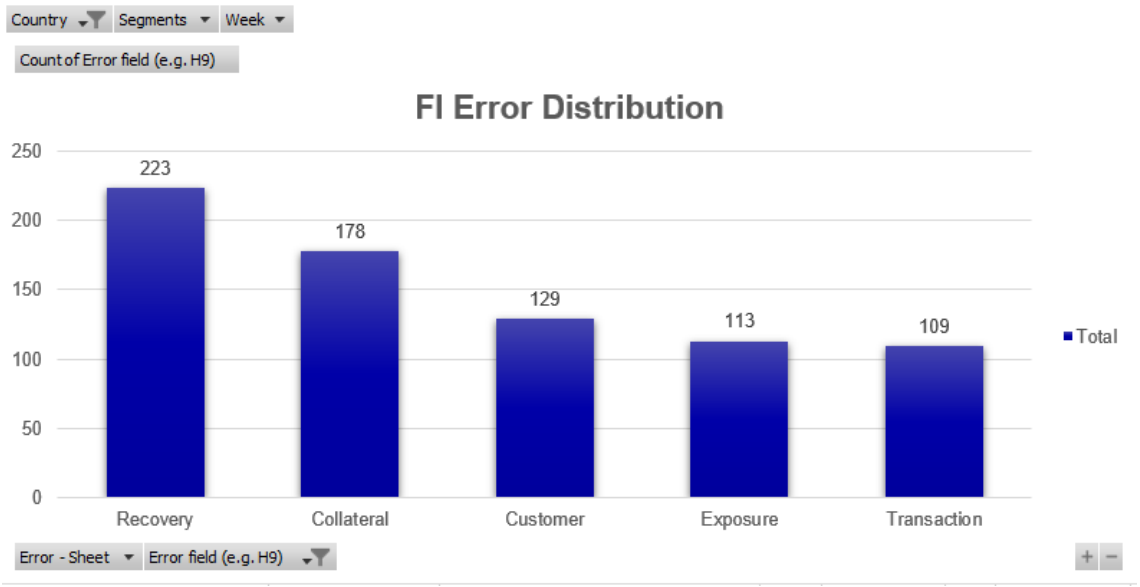

Figure 4: Error distribution in sheets example

5 Quality Analysis

After gathering data for three months and looking at the first results, it was clear that a deeper analysis had to be made in order to gain an understanding of the quality level in

Credit Data Collection. In a quarterly meeting between Credit Data Collection and Risk Models, a KPI was set to monitor the level of data quality:

I. Green – Failure rate is below 10%. This is an accepted rate and no mitigating actions are needed.
II. Yellow – Failure rate is between 11-15%. If failure rate is yellow, the importance quality needs to be emphasized in daily team meetings and Data Collectors must give extra focus on the data accuracy.
III. Red – Failure rate is above 15%. When the failure rate is red, there are certain mitigating actions to be taken. Teams must find out the individual employees who's failure rate is higher. Those employees are given a week to improve their results, before a development meeting will take place organized by the Team Lead. In this meeting, the Team Lead and the Data Collector will fill out a development form, which describes the content of the meeting. In this meeting, they must together decide what actions are to be taken in order to improve the results. Usually the actions are extra training or working with a pair for a fixed amount of time.

The first three months' results were quite alarming already, as the failure rates are as follows:

- January 31,6%
- February 36,2%
- March 23,1%

Each week, all teams were on red and there was a genuine need of understanding the poor performance in more detail.

5.1 General error analysis

For the error analysis, the data that was used was from the beginning of January until the end of March. This means that the cases that were selected as part of sample and manually checked during January-March, were used as the data in the analysis. Data was taken from the Manual Quality Control Board, which is constantly updated as part of the process. The goals of the analysis are:

I. Understand the types of errors that are made in the data collection process
II. Gain ability to pinpoint specific data fields that are more prone to errors
III. By performing a root cause analysis, have explanations why errors are made in specific data fields

During this time, 827 individual cases went through the Manual Quality Control process according to the framework and the standard operating procedure. Out of those cases, 307

included one or more errors within the case data. Total amount of errors found in these "failed" cases was 1084. This means that on average, a case that fails has 3,5 errors in it. Out of 827 cases checked, 520 passed the checks as no errors could be found. Failure rate during this timeframe was 37%.

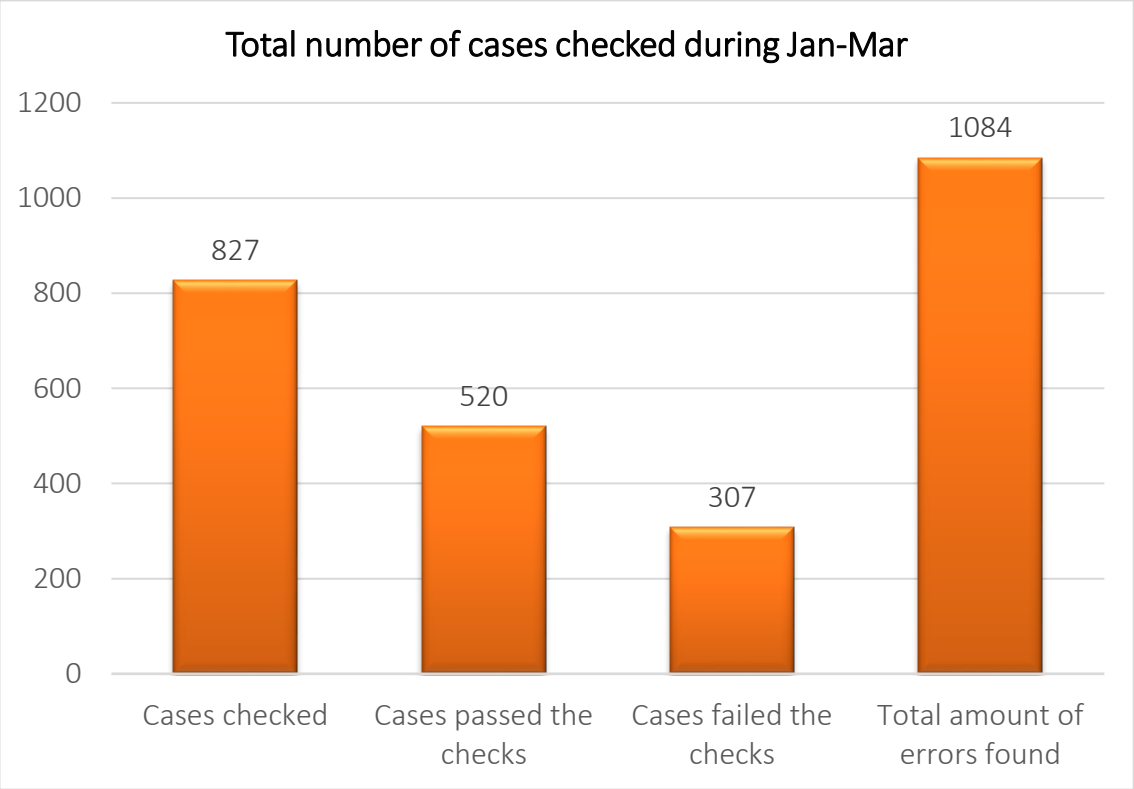

Chart 1: Total number of cases checked

5.2 Distribution between high and low errors

As these failure rates were communicated to the stakeholders and the need for a deeper analysis was established, Credit Data Collection asked Risk Models for a classification of errors. Risk Models provided a classification model, which places errors into high or low classification based on the data fields. This means that different data fields in a case are classified differently and some errors are more critical than others. This is based on which fields have a bigger impact on the models that they are building according to the data provided by Credit Data Collection. All 87 data fields within the cases got a classification and were placed on either high or low importance.

After receiving this model, the 1084 errors found in previous section needed to be classified, in order to understand, if there is a possibility that the errors made are not having a big impact on Risk Models.

By classifying the errors, it was found out that a majority (57%) of the errors were of high importance.

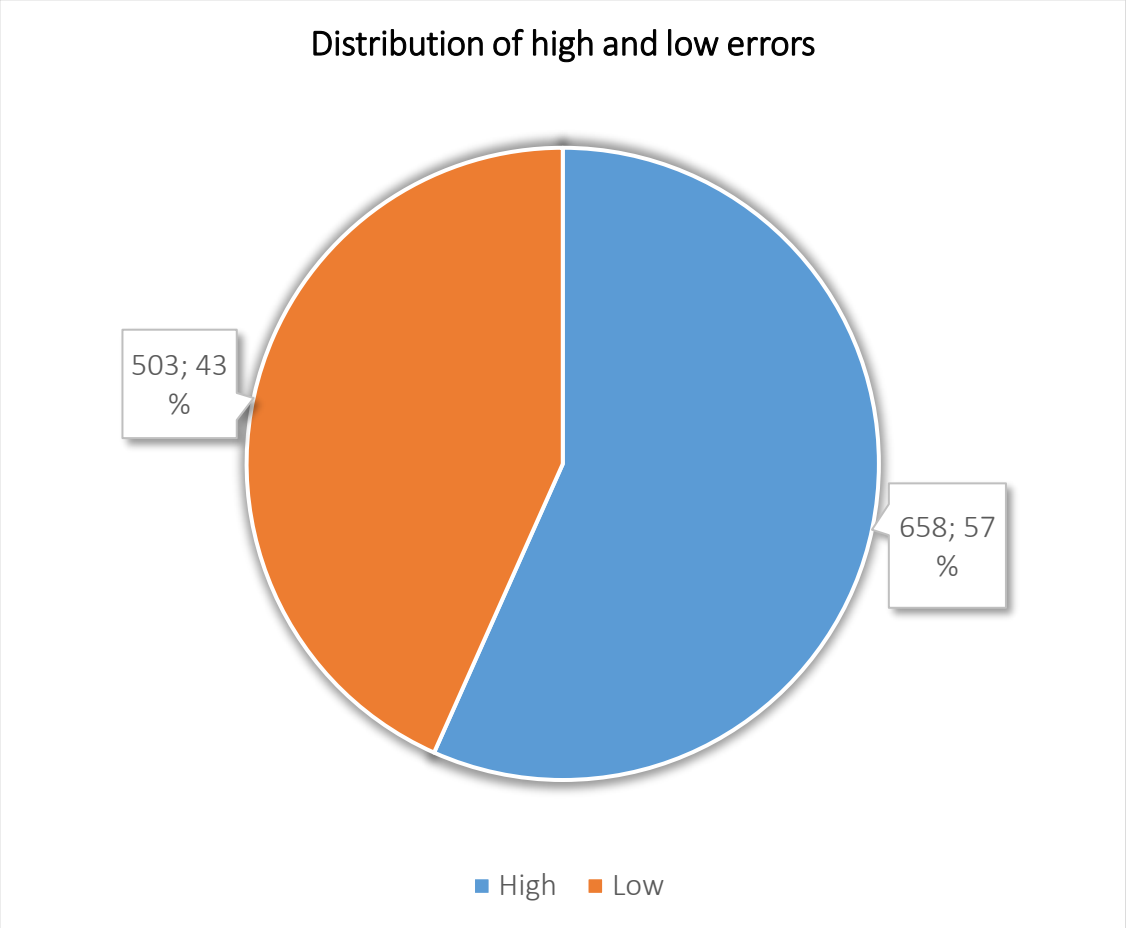

Chart 2: Distribution of high and low errors

Looking at the teams' results separately showed no real difference and a clear trend was to be seen:

- Denmark: 55%
- Finland: 58%
- Sweden: 58%
- Norway: 54%

Majority of the errors that are made in Data Collection are of high importance. These data fields would need to be prioritized in the later stages, when introducing mitigating actions.

5.3 Error distribution between the sheets

After the previous finding of 658 high classified errors, it is needed to understand where in a case can these errors by mostly found. High errors can be made in any of the five sheets, but generally speaking, the more data fields there are in a sheet, the more prone to errors it is.

By this logic, the most amount of high errors should be found from either the Collateral- or the Transaction-sheet as those have most data fields to be filled.

After locating the errors, the results were surprising:

1. Recovery information: 33%
2. Transaction information: 20%
3. Customer information: 19%
4. Collateral information: 18%
5. Exposure information: 10%

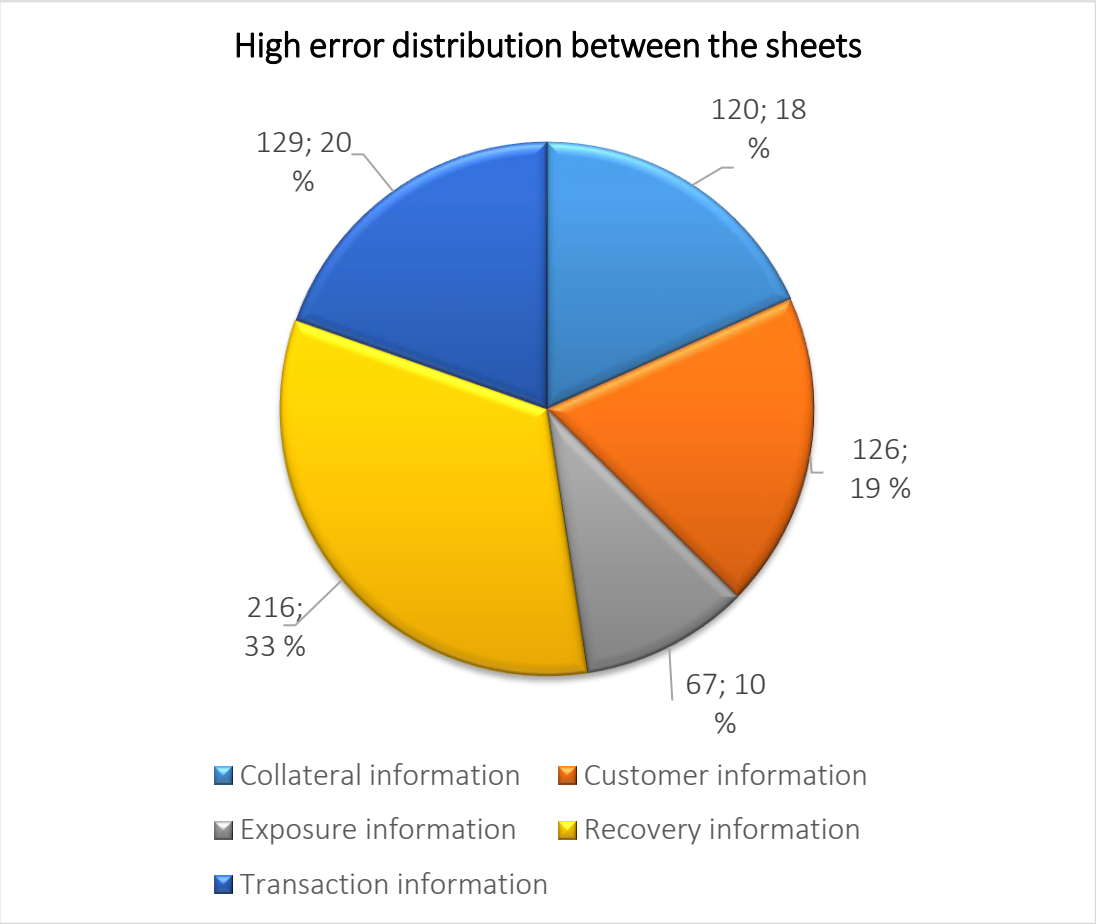

Chart 3: High error distribution in sheets

Third (216 out of 658 cases) of all high errors are in the Recovery sheet. What makes this finding surprising is that the sheet can be filled entirely with the information picked up earlier in the case. No external data sources are needed. 20% of the high errors were in the Transaction sheet, which is also surprising. Even though the sheet has a lot of data fields to be filled, it is mostly filled automatically with various Excel VBA Macros that are in use.

These findings can indicate two things:

1. Standard Operating Procedure for Recovery-sheet is not updated with the latest case rules or the content is inaccurate
2. VBA Macros that are used in collecting transaction data are not working properly or the instructions are unclear

Low classified errors were located as well as part of this analysis and two things stood out. Most of the low errors were made in the Collateral-sheet as 37% of the low errors were there. Only 2% of the low errors happened in the Transaction-sheet. This could be an indication that most of the data fields in the Transaction-sheet are classified as high, considering the amount of high errors found in the sheet.

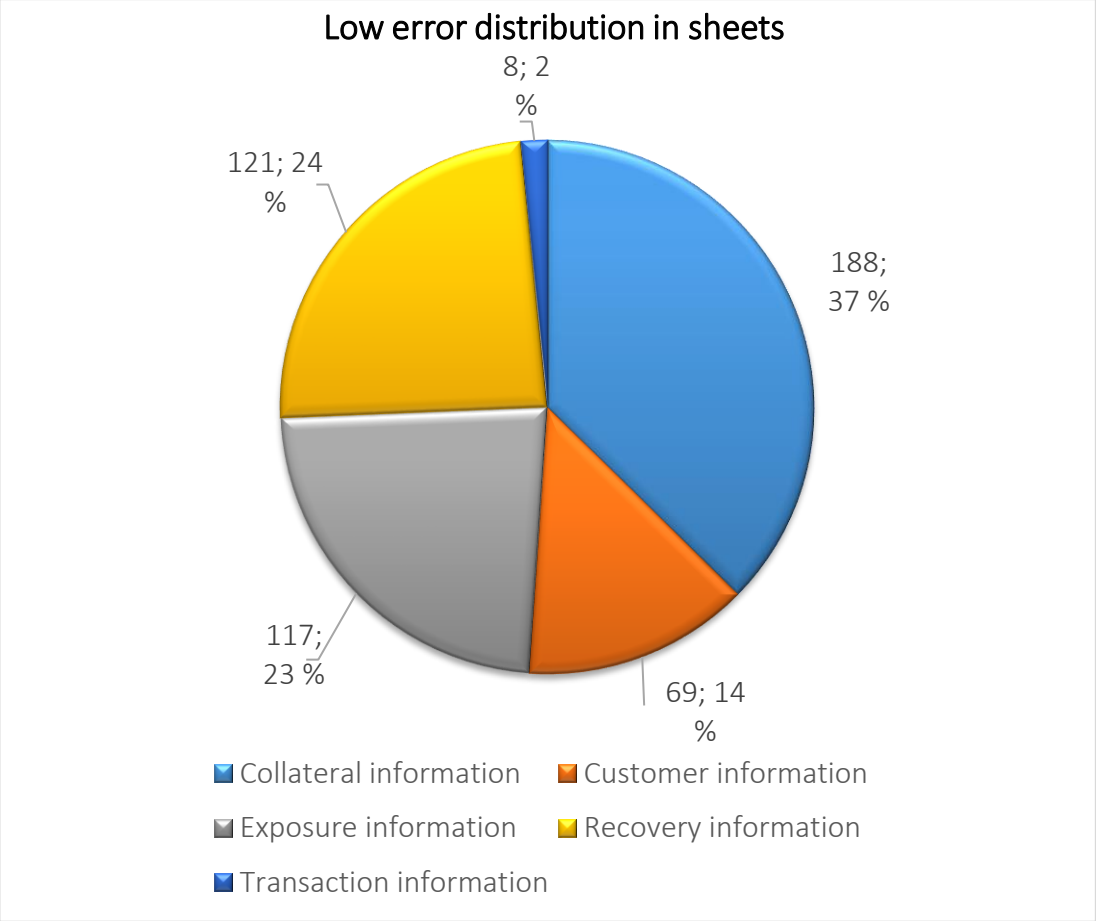

Chart 4: Low errors in sheets

5.4 Errors in specific data fields

As majority of the errors made were classified as high, it was clear that the analysis should focus on the high errors for their impact on Risk Models. In the previous chapter, it was found that third of the high errors were located in Recovery-sheet, and the next step was to go a little deeper.

As the analysis aims to pinpoint areas and scenarios where the data collection is prone to errors, it was crucial to find out the specific data fields, which had the most high errors.

All 87 data fields in a case have a name. The chart below presents the data fields which were classified as high and had the most amount of errors:

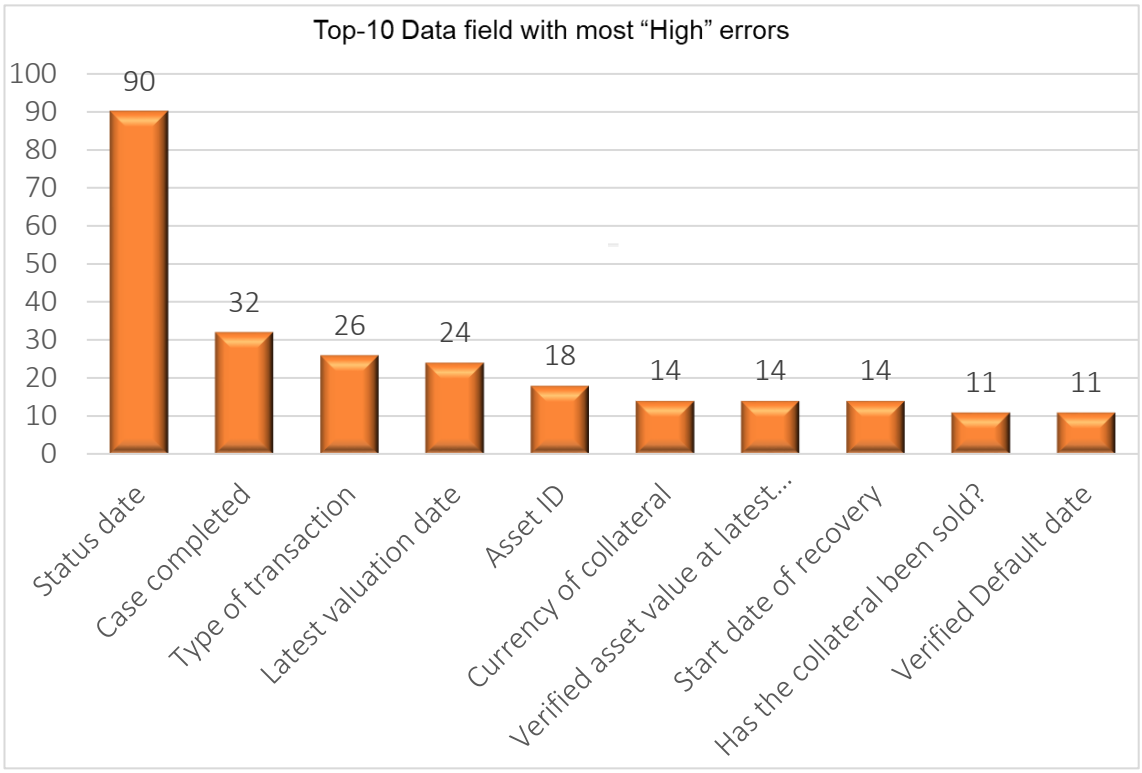

Chart 5: Data fields with most high errors

Out of these ten data fields, two are located in the Recovery-sheet:

I. Status date – Date when a status was changed for an exposure (product) in recovery process
II. Start date of recovery – Date when a customer has entered recovery process

Even if only two data fields from the Recovery-sheet are in the top 10, the amount of errors made in Status date was extremely high as 13,6% of all high errors were made in that specific field. This is a clear indicator that everything connected to this data field needs to be examined in more detail.

The other eight fields appearing on the chart are located as follows:

Customer-sheet – Case completed, Verified default date

Collateral-sheet – Latest valuation date, Currency of collateral, Verified asset value at latest valuation, Has the collateral been sold

Transaction-sheet – Type of transaction, Asset ID

5.5 Ishikawa Analysis

After analyzing the data gathered between January and March, it is clear that multiple issues that need to be tackled were found. The results overall were quite alarming, as the teams were aiming to be under 10% failure rate as stated in the KPIs. In order to do a concrete root cause analysis, it is reasonable to focus on the areas were most of the high errors were made. That way this analysis can have a real impact on the quality in Credit Data Collection.

The areas that will be investigated are:

1. Why is the Status Date-field so prone to errors compared to others?
2. Why is the Recovery-sheet so prone to errors compared to others?
3. Reasons why the teams are making a lot of errors in the top 3 fields

In order to understand the disappointing results, the teams' thoughts needed to be heard as they are the ones producing the data and the errors in data collection. Ishikawa or the Fishbone diagram is a tool for understanding, which factors are leading to an existing problem. It is a widely used quality control tool in many businesses (Sabater 2021). As listed above, the analysis introduced the three problems existing in Data Collection. Together with the Team Leads, we discussed these problems and I acted as a facilitator by helping them group the potential causes according to their level of importance.

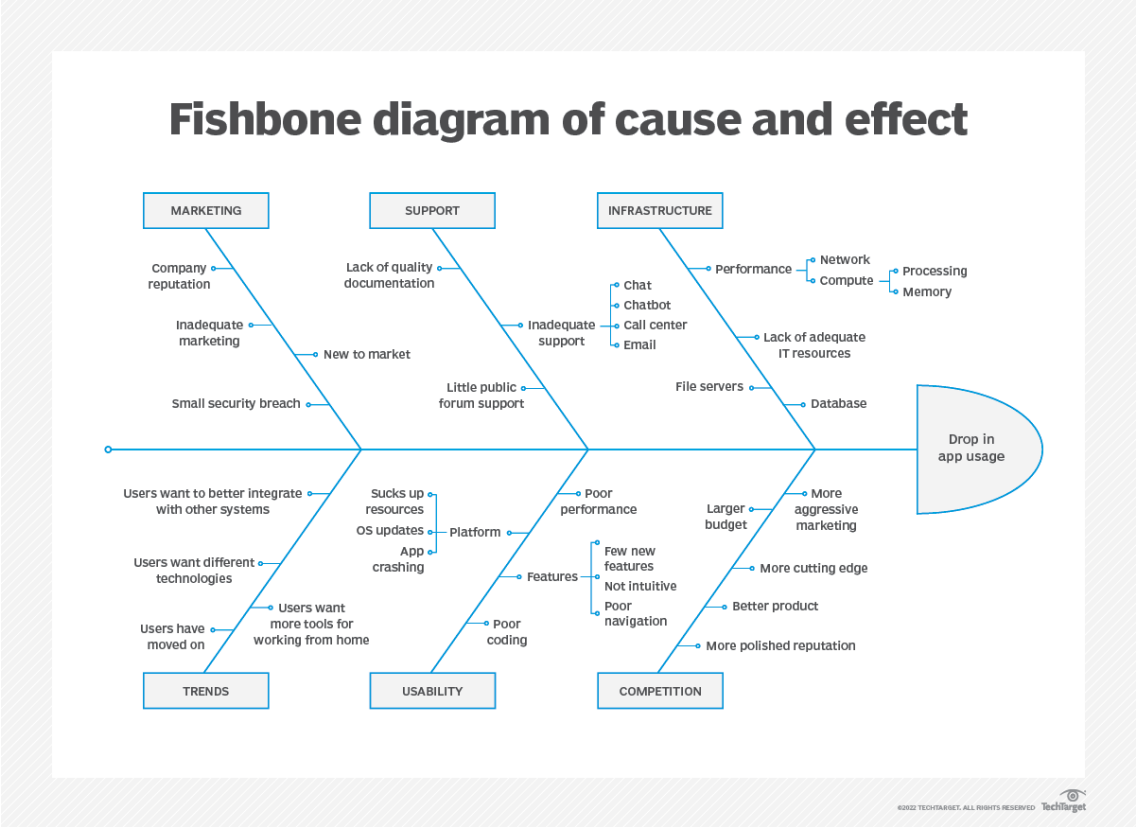

Figure 5: Example of a fishbone diagram (Lewis 2020)

Multiple possible causes for the results were found during the meeting, with the main ones being:

I. Employee turnover – Credit Data Collection has been struggling with keeping experienced employees as they have received better offers from elsewhere. This has caused the balance of experienced-unexperienced to shift, and the teams are in need of experienced workers. Teams have received a lot of new employees, who are not as skilled yet and are therefor making more mistakes in data collection.

II. Rule changes – There are a lot of rules for data collection set by Risk Models to ensure the unit collects relevant data in a correct manner. Risk Models has instructions from European Central Bank on what data should be used in the models they create. As these requirements change, the effect can be seen in data collection as the data now needs to be collected in a different way. Recently, the unit has been under a lot of rule changes, and the employees have not had the time to get used to the new way. This then results in data collected in a wrong way and therefor more errors are found.

III. Exceptional cases – Just like ordinary peoples' finances are different to each other, so are the cases. Some cases are very easy to complete, while some of them might require lot more work. During January and March, the teams were handling extraordinary amount of exceptional cases. These cases often times include more data to be collected or are in situations that have no set of rules established yet.

IV. Attentiveness – Data collection can at times be monotonous, which can make the employee feel bored or less interested. It is largely repetitive as on average an employee completes 25 cases per day. Just like many other jobs, employees in Credit Data Collection are working mainly from home due to the Covid-pandemic. Based on many calculations, the production did not suffer from this switch at all, but there has not been any investigation on what is the impact on attentiveness.

5.6 Mitigating actions

After discovering these areas where improvement would have a massive impact on the failure rate and the overall quality level in Credit Data Collection, it was time to think about the actions, so the improvement would happen. As the root cause discussion was so productive with the Team Leads, it made sense to continue the discussions with them around the topic of mitigating actions. A meeting was held and four actions to be taken were listed:

- Teams will be introducing weekly training/workshop sessions for all Data Collectors. These sessions will be organized whenever the failure rate goes above 10%. In these sessions, the teams will take a closer look at the errors found during the week and discuss the errors with the goal of everyone understanding how to solve the issues in the future. These meetings will be led by a facilitator from each team. Session materials are saved and stored as evidence, that action has been taken whenever the quality level is on red.
- Team Leads will begin to use of individual development plans for under-performing employees. The plan will be used if an employee's failure rate has been above 15% for more than three weeks. The document will be filled by the Team Lead and the Data Collector and it describes the reason for the use, actions for improvement and short-term targets. The documents will be saved and stored as evidence.
- Standard Operating Procedures are going to moved to another platform. Until this point, the documents have always been in a Word format, which has made it inconvenient to report missing or unclear information to the responsible people. These documents will be moved to Confluence which allows a quicker reaction time when it needs to be updated. It is also easier and faster to go through when a Data Collector is searching for an answer on a specific issue.
- Data Collection will be moved from Excel to a new platform called PEGA. Credit Data Collection will be able to implement automatic quality checks on the data collection sites, which will reduce errors caused by attentiveness. These errors are generally small mistakes such as format issues, which can easily be avoided with PEGA.

These actions were presented to the Senior Management and Risk Models in a meeting during May of 2022.

6 Conclusion

The project initially started in autumn of 2021 as Group Internal Audit performed a check on different units within Operations. After their check on Credit Data Collection, they released a report which stated seven different issues in the process. One of them was the lack of a quality control process. By April of 2022, the process should be fully implemented and evidence of the process should be gathered and stored for validation.

Creating a design of the process was fairly simple and straightforward: A sample of cases should be screened for errors in the accuracy manually (or semi-manually). The process and the found errors needed to be well documented. Reporting structure must be established and created in a way that it requires minimal work to update and that it can be viewed at any point. As the unit had no previous quality control set up, all KPIs had to be based on expectations and discussions with stakeholders.

Once the design was documented and approved by the Senior Management and Risk Models, a project plan was created. The project plan had to include everything and it was updated once a week. Other Group Internal Audit issues were handed out the other Business Developers, and a check in meeting was held weekly. In those meetings, the status of all projects was discussed and the meeting was a forum for raising any concerns and issues that had possibly come along.

Once the Standard Operating Procedure was written and approved, a very quick pilot-phase was done with the Swedish team. They were selected as the team that would run the process for a week, to see if there are any issues that needed to be fixed before the full launch with all teams. Nothing came up during the pilot, and the trainings could be organized with all teams.

Just before the trainings took place, a VBA macro was developed together with a developer. The macro allowed the Data Collectors to run a format- and logic-checks on a case after it was completed. The macro would give an instant feedback if it found any errors.

Quality Control in Credit Data Collection begun in December of 2021. All four teams were had been trained and 10% of all completed cases during the week were selected to be part of the sample. These cases were checked by 1-2 Controllers and any errors found were recorded in a Quality Control Board file.

Data was collected for three months between January and March of 2022 to have a sufficient amount for an analysis. During this time, the teams were able to manually check over eight hundred individual cases. This amount of data was enough to start analysing and drawing conclusions.

Based on the data available in the Quality Control Board, an analysis took place with the goal of understanding the types of errors that the teams are making and providing improvement ideas or mitigating actions based on the results. It was clear from the very beginning of the analysis, that the teams' failure rates were a lot higher than expected. The failure rate during January and March was almost forty percent, when the expected rate was below ten percent. Total amount of errors was surprising as well as 307 failed cases resulted in over a thousand errors found.

As the analysis progressed, it was found that most of the errors made were of high importance. 57% of the errors were located in data fields that were rated as high importance by Risk Models.

The analysis was able identify key data fields in a case that are the most prone to errors. These data fields were then brought up in a meeting with the Team Leads and a session was held where the objective was to find reasons why the Data Collectors are making errors in these fields. Several reasons were found and the information was collected and forwarded to the teams and Senior Management.

After this, the teams were able to correct any mistakes in their data collection SOPs, introduce weekly workshops to reduce their errors and a new employee development plans were introduced.

These mitigating actions have resulted in much better quality in Manual Data Collection. During the summer of 2022, the teams were able to reduce their failure rates down to 10% and in many weeks, no errors were found at all. Failure rates are still reported monthly as one of the KPIs and the average failure rate in a team is approximately 13% for the whole year.

References

Electronic

Brush, K. 2021. Standard Operating Procedure (SOP). Accessed 16 November 2022. https://www.techtarget.com/searchbusinessanalytics/definition/standard-operating-procedure-SOP

Corporate Finance Institute. 2022. VBA Macros: How to set up macros in VBA. Accessed 16 November 2022. https://corporatefinanceinstitute.com/resources/excel/study/excel-vba-macros/

LaPrad, L. 2020. 10 Reasons why planning is important in project management. Accessed 16 November 2022. https://www.teamgantt.com/blog/why-planning-is-important-in-project-management

Lewis, S. 2020. Fishbone diagram (Ishikawa cause and effect). Accessed 16 November 2022. https://www.techtarget.com/whatis/definition/fishbone-diagram

Nordea. 2022. Organization. Accessed 16 November 2022. https://www.nordea.com/fi/tietoa-meista/keita-olemme/organisaatio

Nordea. 2021. Annual Report. Accessed 16 November 2022. https://www.nordea.com/en/doc/annual-report-nordea-bank-abp-2021.pdf

Sabater, V. 2021. Ishikawan diagrammi ja ongelmanratkaisu. Accessed 16 November 2022. https://mielenihmeet.fi/ishikawan-diagrammi-ja-ongelmanratkaisu/

Simplilearn. 2022. What is Quality Control: Definition, Benefits, Examples, and Top Techniques Explained. Accessed 16 November 2022. https://www.simplilearn.com/what-is-quality-control-article

Techopedia. What Does Quality Control (QC) Mean? Accessed 23 November 2022. https://www.techopedia.com/definition/12191/quality-control-qc

Figures

Charts

Tables