BPI Challenge 2014 – Applied data mining and process mining techniques for analyzing the impact of a change on the workload of the Service Desk and IT Operations.

Peter Van den Spiegel, Leen Dieltjens, Liese Blevi, Jan Verdickt, Paul Albertini and Tim Provinciael
KPMG Advisory, IT Advisory, Bourgetlaan 40, 1130 Brussels, Belgium
[pvandenspiegel, ldieltjens, lblevi]@kpmg.com

Abstract The service management process forms an essential part in every organization. Since businesses rely heavily on IT, each outage, issue or user service request should be dealt with as quickly as possible in order to minimize its impact on operations. For Rabobank Group ICT, we analyzed the record log files of service management system called HP Service Manager, in order to objectively verify the impact of change implementation on the workload of the Service Desk and/or IT Operations. Our analysis was performed by means of a combination of process mining and data mining techniques and tools, including MS Excel, Disco, Weka and SQL Server Analysis Services. The log files itself consisted of 147.004 Interactions, 46.606 Incidents, 466.737 Incident activities and 30.275 Changes. As part of the exercise, we investigated aspects, such as the patterns that determine if the change implementation will cause an impact on the Service Desk/IT Operation, the average period to return to a steady state, average increase/decrease of Closed Interactions once a new steady state is reached, the change in average steps to resolution, the process flow between interactions, incidents and changes per configuration item, planned versus actual timing, etc. Finally, we also made recommendations on how to further improve the current models (impact patterns) and thus increase their predictive value.

Keywords: BPI Challenge, Process Mining, Data mining, Incident management, Problem management

1 Preface

Following our participation in the previous BPI challenge, we were eager to find out what the 2014 challenge would look like. Similar to last year, the area of investigation was related to IT service management. But the depth and breadth of the required analysis was significantly more complex than last year, given the need to develop a predictive model that could be operationalized to help manage the workload of the Service Desk and/or IT operations.

All in all a great challenge, allowing us to apply several data and process mining techniques to a concrete client concern.

We would like to thank everyone in our Analytics core team who assisted and supported us during their scarce free hours. Also thanks to Prof. dr. ir. David Martens and his team who gave us additional insights for the predictive model.

2 Introduction

Due to the uncertain global economic climate companies have been faced with in past years, focus has been put on optimizing resources and cutting costs. Additionally, banking regulation has advanced notably since the 2008 financial crisis.

KPMG studies show that as a result of this increased regulation, demands by regulators have and are continuing to grow, requiring banks to disclose increasing amounts of information on their asset valuations, risk management processes and internal operations. This has also unveiled the multitude of issues that banks are facing, concerning data quality and management given the fragmented systems and processes through which this data flows. Constrained by legacy IT systems, increasingly larger budgets are being spent on improving data and risk management. [1]

Unfortunately, at the same time, increased budget limitations put additional pressure on the effectiveness of IT to deliver its services.

It is in this context that we need to view Rabobank Group ICT's problem statement: they are faced with increased software releases and a decreased time to deliver these releases. Therefore, Rabobank Group ICT would like to be able to **predict** future workloads to further optimize the resources available to Rabobank Group ICT. As such they are looking for a predictive model which can be used in a BI environment and can help reduce the impact of changes on the Service Desk and/or IT operations. [2]

In this respect, the Business Process Intelligence 2014 challenge, puts us right in the middle of the data analytics "revolution".

2 Executive Summary

Based on our analyses performed we noted a substantial correlation between the interaction/incident management and change management process at Rabobank Group ICT. At first glance, using visualizations and time series, the interaction

[1] http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/evolving-banking-regulation/Pages/Default.aspx

[2] ***Predictive data modelling*** *is the result of turning large volumes of data into valuable information. This is something a large majority of companies are faced with: How to gain more insight into their data, and derive value form it. A recent KPMG report shows that 96% of C level executives from major corporations around the world (+ 1 billion USD turnover) believe they could do more with the data they own by making better use of analytics in their organization.*

between changes and interactions/incidents appeared to be only marginal, with the exception of some specific configuration item types (e.g. software, hardware…). Overall the impact of changes on the Service Desk/ IT Operations, appeared to be limited, largely due to the substantial number of incidents and interactions on configuration items, for which no changes took place at all. However, when applying different sets of predictive data mining algorithms the correlation and potential predictability became apparent.

Below, we provide an answer on the key questions as asked by the process owner.

1 Identification of impact patterns;
   - We succeeded in identifying patterns in order to predict the impact of a change implementation the workload of the Service Desk/IT Operation, being the amount interactions, incidents, and the combination of both. By means of the Linear Regression technique we were able to predict the impact fairly well; a correlation coefficient of 0,55; 0,64; 0,59 respectively. The Decision Tree Models give even better results: they had a predictive value of 85%, 86%, 69% respectively.
   - When developing a predictive model, the necessary attention was spent on ensuring the operationalization of the model. Since Rabobank would like to apply the future predictive model as part of their day-to-day operations, it is of utmost importance that the model, although not perfect, remains easy to understand and to implement. Avoiding over fitting and a black-box-approach was key.
   - Note that, by means of process mining we noted that in many cases (67%) changes do not cause any interactions or incidents at all, for the same configuration item.

2 Parameters for every impact pattern;
   - The average workload in steady-state operation was determined by the configuration items on which incidents and/or interactions, but no changes occurred. It is approximately 697 tickets per day, which is almost half the average workload than with changes. Additional work could be put in calculating the average increase/decrease of closed interactions after a new steady state.

3 Change in average steps to resolution;
   - Only limited research was done on the average steps to resolution. We performed an analysis on individual service components, such as for the Service Component WBS000152 (within the top 10 of the most frequent service components). We noted that the change in average steps to resolution showed a downward trend. Thus for this specific Service component, a better service level is maintained over all the changes in time. Additional work could be done on analyzing the steps to resolution from a more statistical and global perspective.

4 Creativity challenge;

- Throughout our investigation, we spent considerable time on analyzing the process behind the data, as well as the quality of the data itself. Given the expectation of Rabobank Group ICT to operationalize the model, we believe it is important to notice that several signs of data quality issues have been identified, increasing the risk of a building an erroneous predictive model.

- Firstly, we noted that many of the data fields containing information about the interactions, incidents and changes have been filled in manually. Even more, several fields are not consequently filled out by the Service Desk/IT Operators, and thus contain null values. There is also a large degree of inconsistency between the different start/end date fields in the changes data table. This causes many fields to be unreliable, which off course has an impact on the outcome of the further analyses deployed.
  We also noted that the same change ID can have different attributes assigned, such as considered both as an “emergency change” and a “regular” change.

- From a process perspective, we noticed that the planning is often overestimated (30%). On average the change implementation is planned to end 20 days earlier than was requested. An unrealistic planning however causes the Service Desk/IT Operators being incapable of tackling the interactions/incidents with the needed care, as they are understaffed.

- A high level view on the provided data shows that the amount of interactions and incidents are highly correlated (0,9581). The relation with the changes on the other hand is rater low (0,6630 and 0,6992 for incidents and interactions respectively). Furthermore, as we would expect, the amount of service management related events significantly drops during the end of year period in December.

3 Approach

3.1 Methodology

The following steps have been completed in order to tackle this challenge.

Figure 1: Methodology

For this challenge we first obtained a general understanding of the process and the questions asked by the process owner. We based ourselves on the information (process descriptions, questions about the process) provided by Rabobank Group ICT and Fluxicon, ITIL best practices and our experiences in the domain of service management. We tried to identify a number of expected behaviors between the incident and change management process.

In the next step we obtained a general understanding of the data. Through data exploration techniques, we could define which data fields were suitable for further analysis and determine which analyses were possible to perform. We identified the unique ID's of the provided data tables and the possible links between the different data tables.

Based on our understanding of the data, we transformed the data to fit the format required by the tools we used. We then analyzed the data using a combination of data mining and process mining techniques. Furthermore, we applied visualization techniques to get a better insight in the analysis results and increase their readability.

3.2 Tools

Because of the nature of the challenge, we applied a combination of process mining techniques, data mining techniques and visualization techniques. To apply these techniques, we used a combination of different tools:

- Ataccama DQ Analyzer: DQ Analyzer (DQA) is a data analysis tool that combines advanced data profiling and analysis capabilities with a point-and-click interface. We used this tool during the data exploration phase to assess the data quality and to determine the suitability of the different data fields to perform further analysis.[3]

- CA ERwin Data Modeler: CA ERwin Data Modeler is a data modeling solution that offers a simple, visual interface to manage your complex data environment. It provides a collaborative data modeling environment to manage enterprise data through an intuitive, graphical interface.[4]

- Microsoft Excel 2010: Microsoft Excel is spreadsheet software that is used to create tables, calculate and analyze data. Furthermore, it has strong visualization capabilities which are very helpful in understanding the data and analysis results better. [5]

[3] http://www.ataccama.com/en/products/dq-analyzer.html
[4] http://erwin.com/products/data-modeler
[5] http://office.microsoft.com/nl-be/excel/

- Disco: Disco is a process mining tool that supports in automated process discovery, process map animation and enables to perform detailed case analysis. We used this tool during the analysis phase to identify different patterns in the data.[6]

- SQL Server Management Studio: Microsoft SQL Server 2008 Management Studio Express is an integrated environment for accessing, configuring, managing, administering, and developing all components of SQL Server, as well as combining a broad group of graphical tools and rich script editors that provide access to SQL Server to developers and administrators of all skill levels. We used this tool for data transformation purposes. [7]

- WEKA: Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.[8]

- SQL Server Analysis Services: Microsoft SQL Server Analysis Services (SSAS) delivers online analytical processing (OLAP) and data mining functionality for business intelligence applications. For data mining applications, Analysis Services lets you design, create, and visualize data mining models by using a wide variety of industry-standard data mining algorithms.[9]

4 Gain Understanding

4.1 Process understanding

In order to get a general understanding of the service management process, we investigated **ITIL's best practices** within this domain.

[6] http://fluxicon.com/disco/
[7] http://www.microsoft.com/en-us/download/details.aspx?id=7593
[8] http://www.cs.waikato.ac.nz/ml/weka/
[9] http://technet.microsoft.com/en-us/library/ms175609(v=sql.90).aspx

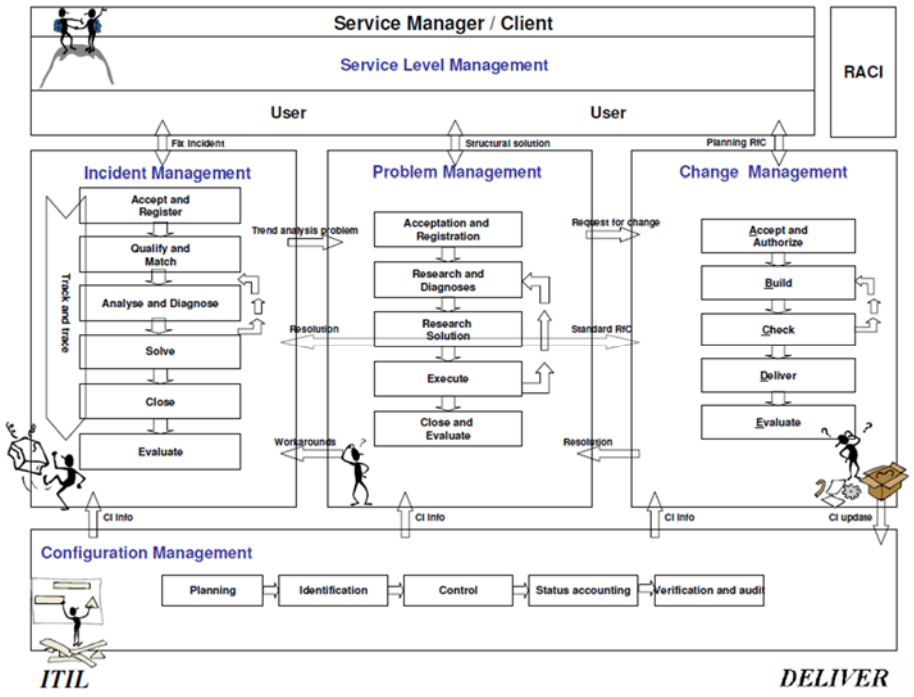


Figure 2: ITIL Service Management model

The ITIL process for Incident Management is focused on restoring the IT services for the users as soon as possible. The creation of an incident is triggered by a significant configuration item event. Once it is clear that the 1st level support is not able to resolve the incident, the 2nd (and 3rd level) support are involved. Afterwards, a problem is created in case the root cause of the incident could not be resolved. Within the Problem Management area the aim is to prevent incidents from happening and to minimize the impact of incidents that cannot be prevented. So the root cause of the problem needs to be identified and a solution must be initiated.

This could result in the initiation of a change request to “permanently” solve the incident/problem and prevent it from reoccurring. Finally within the Change Management area, multiple constructive changes are made with a minimum disruption to the IT services.

It should be mentioned that an important part of the Service Management Model, **Problem management**, was set **out of scope** by the data provider.

4.2 Problem understanding

In the current economic situation, ICT companies are faced with increased software releases in combination with a decreased time to market. For this reason, Rabobank Group ICT has implemented a change management process based on the ITIL best practices. The primary objective of this process is to enable beneficial changes to be made, with minimum disruption to IT services.

We understand that, in order to reach this objective, Rabobank Group ICT is looking for fact-based insight into the disruption to IT services caused by changes in the past enabling them to predict the workload of the Service Desk after future changes.

On the basis of this insight, Rabobank Group ICT would like to implement a predictive model to support Business Change Management in implementing software releases with less impact on the Service Desk.

In order to get the before mentioned fact-based insight, the process owner would like an answer to the following questions:

- Identification of Impact-patterns: Rabobank Group ICT is interested in the impact of a change implementation on the workload of the Service Desk. Furthermore, they are interested in the difference between the impact-patterns across the various service components. Rabobank Group ICT defines the workload of the Service Desk as the combination of Closed Interactions and Incidents.
    - What is the correlation between the implementation of a change and the number of closed interactions and incidents after the change?
    - Is there a difference in the correlation across the various service components?
    - What is decrease or increase of the number of closed interactions and incidents after the implementation of a change?
    - Are there different impact-patterns across the various service components?

- Parameters for every Impact-pattern: Rabobank Group ICT would like to define for each of the impact-patterns found a number of parameters to use in the predictive model.
    - What is the average period to return to a steady state?
    - What is the average increase/decrease of closed interactions once a new steady state is reached?

- Change in Average Steps to Resolution: Rabobank Group ICT expects form its project managers to deliver the same or better service levels after each implementation of a change and they are looking for confirmation that this challenge is indeed being across the various service components.
    - Does the number of closed interactions and incidents decreases after the implementation of a change?

- Is the decrease of closed interactions and incidents similar across the various service components?

5 Data Exploration

5.1 Data Profiling

Data profiling is a technique to examine data available in a data source and to collect statistics and information about that data. The purpose is to **evaluate the data quality** and **uncover erroneous areas** in the data organization. For this challenge, we were provided with 4 data tables:

- Detail Interaction
- Detail Incident
- Detail Incident Activity
- Detail Change

On each of these tables we applied data profiling techniques by means of the Ataccama DQ Analyzer in order to facilitate the data transformation required to perform further analysis.

5.1.1 Detail Interaction

We identified that the ‘interaction ID’ is the unique identifier for this table as all the values appear only once (Distinct column) and are unique (Unique column). Only the 'Closure Code' and 'Related incident' fields contain null values. The number of null values for 'Related incident' relates to the number of interactions that could not be solved by the first line helpdesk (36%).

Expression	*Non-null*	*Null*	*Unique*	*Distinct*
CI Name (aff)	147.004	0	2.730	4.153
CI Type (aff)	147.004	0	0	14
CI Subtype (aff)	147.004	0	10	67
Service CompWBS (aff)	147.004	0	25	289
Interaction_ID	147.004	0	147.004	147.004
Status	147.004	0	0	2
Impact	147.004	0	0	5
Urgency	147.004	0	1	6
Priority	147.004	0	0	5
Category	147.004	0	1	6

KM_number	147.004	0	286	2.36
Open Time First_Touch	147.004	0	24.685	65.848
Close Time	147.004	0	24.308	64.727
Closure Code	146.517	487	2	24
First Call Resolution	147.004	0	0	2
Handle Time (secs)	147.004	0	772	3.064
Related Incident	52.754	94.250	42.789	46.088

Figure 3: Outcome of the Attacama DQ Analyzer for the table Detail Interaction

A detailed list of findings on data quality and profiling per field, regarding the "detail interaction table", can be found in **appendix A.9**. In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

5.1.2 Detail Incident

This table can be uniquely identified with the field 'Incident ID'. There are also several fields that contain null values (Urgency, # Reassignments, Reopen time, Resolved time, Handle time (hours), Closure code, # Related Interactions, # Related incidents, # Related changes, Related Change). We understand that these fields need to be **manually filled in**, and therefore have a lower data quality.

Expression	*Non-null*	*Null*	*Unique*	*Distinct*
CI Name (aff)	46.606	0	2.062	3.019
CI_Type (aff)	46.606	0	0	13
CI Subtype (aff)	46.606	0	11	65
Service Component WBS (aff)	46.606	0	32	274
Incident ID	46.606	0	46.606	46.606
Status	46.606	0	0	2
Impact	46.606	0	0	5
Urgency	46.605	1	0	5
Priority	46.606	0	0	5
Category	46.606	0	1	4
KM number	46.606	0	383	1.825
Alert Status	46.606	0	0	1
# Reassignments	46.605	1	10	41
Open Time	46.606	0	46.067	46.336
Reopen Time	2.284	44.322	2.28	2.282
Resolved Time	44.826	1.78	44.381	44.603
Close Time	46.606	0	46.147	46.376
Handle Time (Hours)	33.734	12.872	21.726	26.592
Closure Code	46.146	460	1	14

# Related Interactions	46.492	114	18	49
Related Interaction	46.606	0	43.058	43.06
# Related Incidents	1.222	45.384	9	24
# Related Changes	560	46.046	1	4
Related Change	560	46.046	181	232
CI Name (CBy)	46.606	0	2.689	3.652
CI Type (CBy)	46.606	0	1	14
CI Subtype (CBy)	46.606	0	11	63
ServiceComp WBS (CBy)	46.606	0	34	275

Figure 4: Outcome of the Attacama DQ Analyzer for the table Detail Incident

A detailed list of findings on data quality and profiling per field, regarding the "detail incident table", can be found in **appendix A.10**. In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

5.1.3 Detail Incident Activity

We understand that the incident activities are logged automatically based on the actions performed in the incident record. Because of this automatic logging, we expect a **high data quality**. This is also evident from figure 5. None of the fields contains null values. However, the Interaction ID is in 1,21% of the cases set to '#N/B'. This can be explained by the fact that not all incidents are linked to an interaction, and therefore there is no related interaction ID to report in the incident activity record.

The IncidentActivity_Number is the primary key for this table.

Expression	*Non-null*	*Null*	*Unique*	*Distinct*
Incident ID	466.737	0	1	46.616
DateStamp	466.737	0	141.981	273.401
IncidentActivity Number	466.737	0	466.737	466.737
IncidentActivity Type	466.737	0	0	39
Assignment Group	466.737	0	4	242
KM number	466.737	0	0	1.825
Interaction ID	466.737	0	1	46.444

Figure 5: Outcome of the Attacama DQ Analyzer for the table Detail Incident Activity

A detailed list of findings on data quality and profiling per field, regarding the "detail incident activity table", can be found in **appendix A.11**. In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

5.1.4 Detail Change

In comparison with the other 3 tables, there **no single field that can serve as a unique identifier**. Especially because the Change ID is not unique. Note that several fields contain null values: Planned End, Scheduled Downtime Start, Scheduled Downtime End, Actual Start, Actual End, # Related Interactions, # Related Incidents. We understand that these fields need to be manually filled in, and therefore have a lower data quality.

From our findings, it is clear that there is a **lot of inconsistency** between the different start/end date fields in the data table. We would expect the actual end date to be in line with the close time, and thus to be between 1/10/2013 and 31/03/2013. However, the actual end date lies between 26/10/2012 and 31/02/2021. This indicates that the fields are not always correctly filled in and may impact our analysis results.

Expression	*Non-null*	*Null*	*Unique*	*Distinct*
CI Name (aff)	30.275	0	5.787	10.193
CI Type (aff)	30.275	0	0	13
CI Subtype (aff)	30.275	0	6	74
Service Component WBS (aff)	30.275	0	40	286
Change ID	30.275	0	13.974	18
Change Type	30.275	0	22	240
Risk Assessment	30.275	0	0	3
Emergency Change	30.275	0	0	2
CAB approval needed	30.275	0	0	2
Planned Start	30.275	0	7.438	11.774
Planned End	30.232	43	5.661	8.905
Scheduled Downtime Start	755	29.52	143	297
Scheduled Downtime End	744	29.531	152	307
Actual Start	27.017	3.258	8.825	13.1
Actual End	27.014	3.261	9.321	13.78
Requested End Date	30.275	0	4.147	6.642
Change record Open Time	30.275	0	9.841	15.066
Change record Close Time	30.275	0	10.498	15.549
Originated from	30.275	0	0	3
# Related Interactions	2	30.273	2	2
# Related Incidents	1.948	28.327	10	57

Figure 6: Outcome of the Attacama DQ Analyzer for the table Detail Change

A detailed list of findings on data quality and profiling per field, regarding the “detail change table”, can be found in **appendix A.12**. In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

5.1.5 Data profiling conclusion

Based on the data profiling results it can be seen that a lot of data is manually filled in. Only the Detail Incident Activity table, and the Open and Close time fields are automatically generated by the system. There appear to be different fields with null

values and not all fields are consistently filled in. Especially in the Detail Change table we have to be careful with the start and end date fields. Accordingly, the quality of most analyses is open for interpretation as it depends on the correct input of the Service Desk personnel/IT Operators. For our further analyses we will assume however that the filled out data is correct.

Based on the actual and planned timing, we see that on average 30% more time is planned than is actually needed to implement the change. In general, the change implementation also starts 49 hours later than was initially planned, although it ends about 20 hours earlier than planned. And on average the change implementation is planned to end 20 days earlier than was requested. So although the duration of a change is overestimated, it generally starts later than was planned.

5.2 Table relations

For the Detail Interaction, Detail Incident and Detail Incident Activities, the primary key is rather straight forward: Interaction ID, Incident ID and IncidentActivity Number respectively. This was also concluded in the Data Profiling phase.

For the Detail Change table however, the Change ID field is not unique. In order find a unique key, the fields CI Name (aff) and even Change record Close Time need to be added. Even though, four lines are still completely identical. It seems that if the Change record Close Time field is changed, a new line is created in the Detail Change table. In four cases however, the Change record Close Time field was then changed to exactly the same time, creating double records.

We also investigated the possible links between the four tables.

For 873 interactions (0,59%) the Related Incident field contains the value '#MULTIVALUE', and thus no pure link could be made between the Detail Interaction and Detail Incident table by means of the Related Incident ID and Incident ID fields. The Incident and Interaction table have a one-to-many relationship, so an incident can be linked with multiple interactions. Furthermore, 970 incidents have no link to an interaction in the provided Detail Interaction table.

All incidents have an incident activity, but 72 incident activities have no related incident in the incident table. The relation between incidents and incident activities is a one to many relationship.

In the Detail Incident table, the Related Change field is only filled out for 1,2% of the records, and 24 records (0,05%) contain the value '#MULTIVALUE' for this field. As the changes table contains 30275 records, no real link can be made with the incident table based on the Related Change field.

However, the Interaction, Incident and Change table can also be linked by means of the CI name (aff), CI type (aff), CI subtype (aff), and Service Component WBS (aff) fields.

In our further analysis regarding process mining, we combined the tables based on the CI name (aff) field. For the **Data Mining analysis**, we have **linked** the tables on the **Service component WBS (aff) field**.

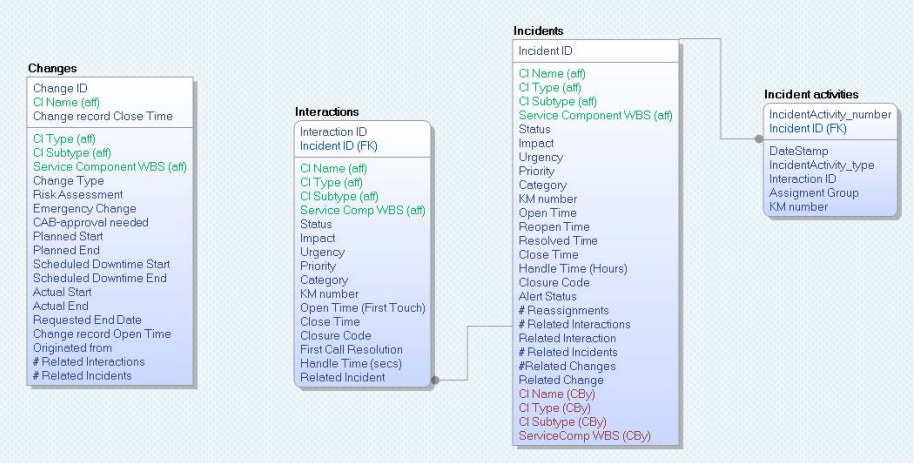


Figure 7: Output of CA ERwin Data Modeler

6 Data Insights

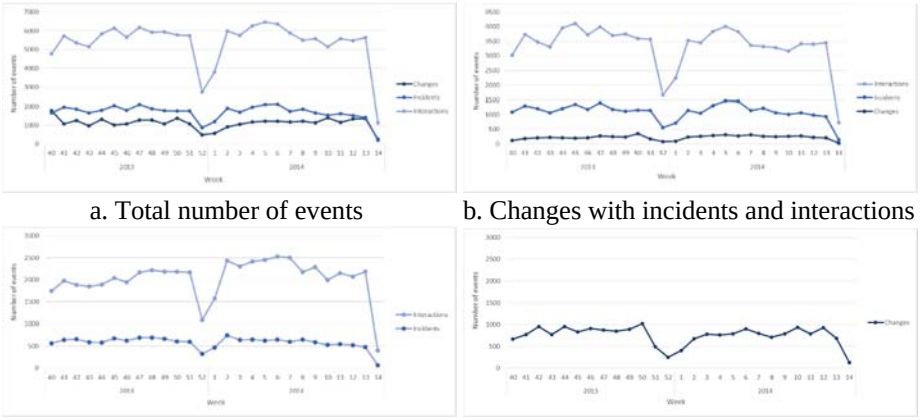
6.1 Data visualization

One of the techniques we used to get insight in the data are visualization techniques. With the help of Microsoft Excel we created several (time) graphs.

In order to get a first insight in the dispersion of events over time, we created 4 **timelines** applying different filters. We linked the events based on CI Name (aff) and only considered the events in the period 01/10/2013 - 31/03/2014. For the period filter we used Open Time for incidents and interactions, and Actual End for changes. Furthermore, we applied an additional filter on configuration item (CI Name (aff)):

a) No additional filter. The graph shows the total number of events in the period 01/10/2013 – 31/03/2014 across all configuration items. Each line represents a type of event (i.e. Change, Incident or Interaction).
b) Based on CI Name (aff) we filtered out the configuration items for which at least once occurred a change AND incident AND interaction in the period in scope. Changes, incidents and interactions that occur on the same configuration item, we consider as related events.
c) Based on CI Name (aff) we filtered out the configuration items for which NEVER occurred a change in the period in scope, only incidents and/or interactions.

d) Based on CI Name (aff) we filtered out the configuration items for which ONLY occurred changes in the period in scope, no incidents and interactions.



a. Total number of events

b. Changes with incidents and interactions

c. Incidents and interactions without changes

d. Changes without incidents and interactions

Figure 8: Events per week in the period 01/10/2013 - 31/03/2014

Across the different graphs in figure 8 the lines have a similar shape. Furthermore, when comparing the changes, incidents and interactions also a similar trend can be seen (see graph a). Especially the **incident and interaction** line show a very **high similarity**. This is also evident from the correlation ratio, which is equal to 0,9581. Changes and incidents as well as changes and interactions on the other hand are only moderately correlated, with a correlation ratio of 0,6630 respectively 0,6992.

Due to the holiday season in December, the number of changes, incidents and interactions takes a significant drop in the last week of the year.

In order to avoid bias, we created a timeline with the average number of events per week, taking into account the weekend days and holidays. From figure 9 it can be seen that this also resulted in a similar graph.

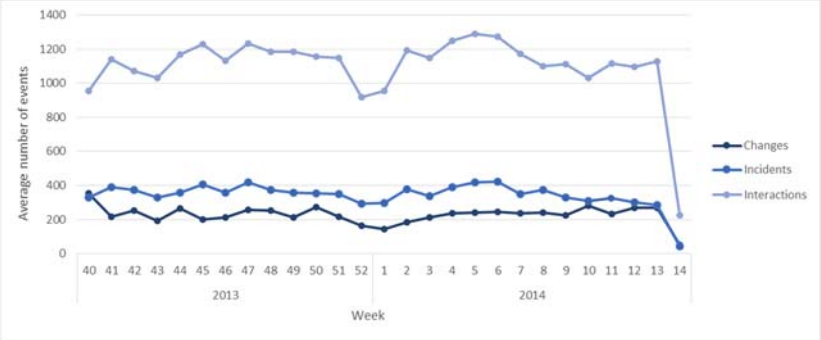


Figure 9: Average number of events in the period 01/10/2013 – 31/03/2014

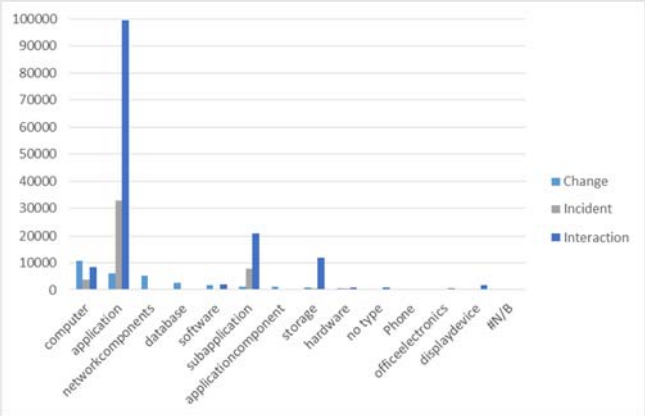


Figure 10: Number of events per configuration item type (CI Type (aff))

Another graph that helped us to get insight in the data is a Column chart presenting the number of events per configuration item (CI) type. Figure 10 shows that the number of changes, incidents and interactions are not equally spread across the different CI types. Most events take place on applications and subapplications. The third most affected CI type is computer.

Furthermore, we made a comparison between the CI types of the events over time. We made use of relative terms, enabling us to see similar trends between the different CI types. The graphs are sorted from most affected CI type to least affected CI type, being Application respectively #N/B.

The first two CI types, which cover 75% of all events in the period 01/10/2014 – 31/03/2014, show a stable pattern over time: 5% changes – 25% incidents – 70% interactions.

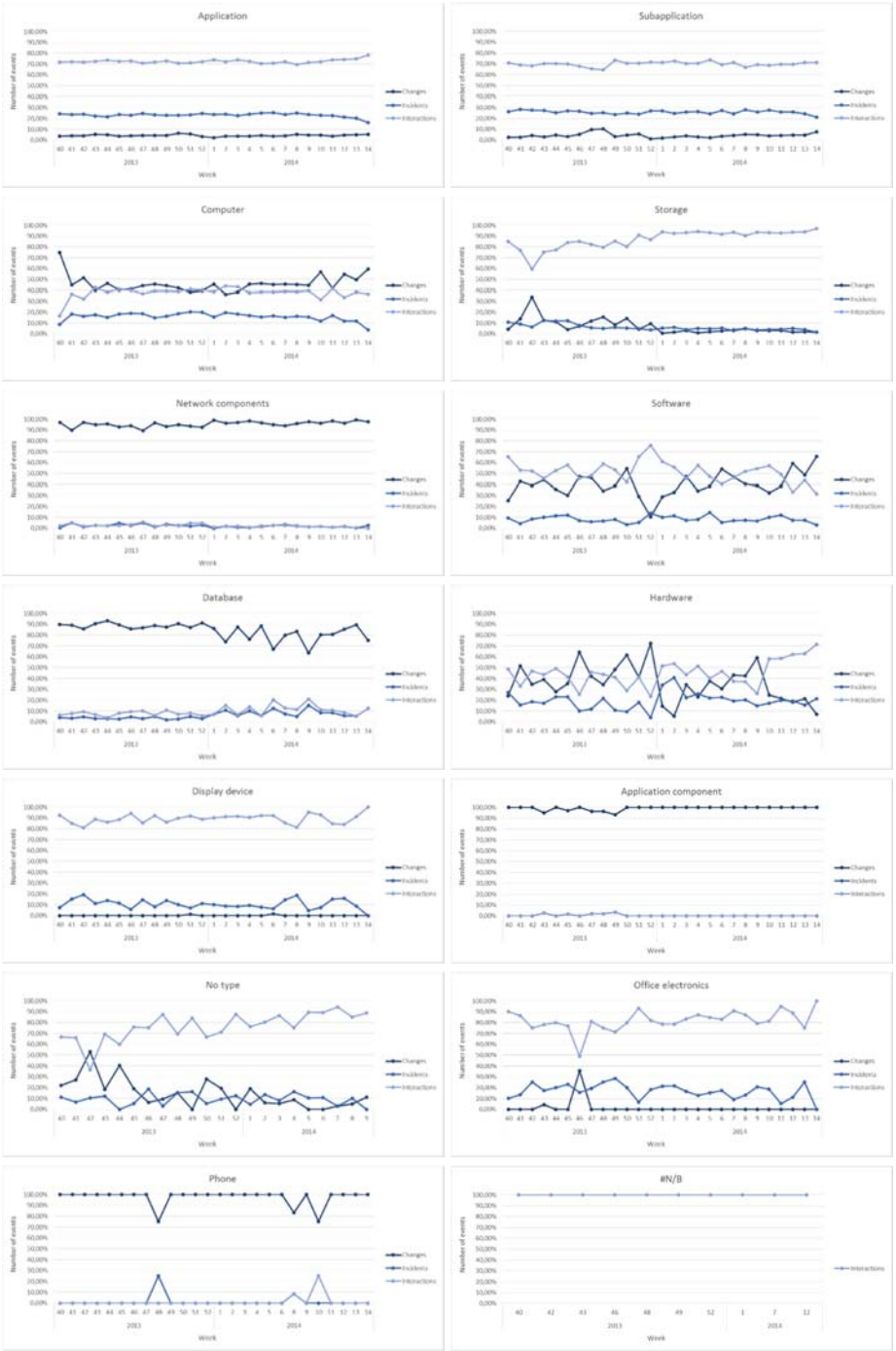


Figure 11: Comparison of the events over time between CI types in the period 01/10/2013 – 31/03/2014

6.2 Additional insights

6.2.1 Average workload of the Service Desk

We defined the average workload of the Service Desk as "the average number of tickets (including changes, interactions and incidents) that are closed per day" or expressed into a formula:

$$\text{Average workload of the Service Desk} = \frac{\text{\# Changes + \# Incidents + \# Interactions}}{\text{\# days in scope}}$$

The period in scope is from 1/10/2013 until 31/03/2014, the number of # days in scope therefore is 181 days. Following the above formula we found an average workload of the Service Desk of approximately 1237 tickets/day. In table 1 an overview can be seen per type of ticket (event):

Table 1: Overview average workload calculation

Total		Average	
Total # Changes	30275	AVG Changes/day	167,2651934
Total # Incidents	46606	AVG Incidents/day	257,4917127
Total # Interactions	147004	AVG Interactions/day	812,1767956
Total # Tickets	**223885**	AVG workload	**1236,933702**

6.2.2 Average workload of the Service Desk in steady-state operation

In our calculation of the average workload of the Service Desk in steady-state operation, we took the configuration items on which incidents and/or interactions occurred. There are 4636 configuration items (32,78%) on which incidents and interactions occurred.

We assume that the configuration items on which incidents and/or interactions, but no changes occurred represent the steady-state operation. Table 2 gives an overview of the number of configuration items impacted.

Table 2: Overview of the number of configuration items impacted

	# CI	% CI
Total CI impacted with incidents AND/OR interactions	4636	32,78%
Total CI impacted with incidents and **NO** changes (1)	2420	52,20%
Total CI impacted with interactions and **NO** changes (2)	3501	75,52%

The 2420 configuration items on which an incident occurred correspond to 15796 incidents. In order to get the total number of incidents in steady-state operation we divided this number of incidents by 52,20% (percentage of configuration item of the total number CI impacted), which resulted in approximately 30260 incidents. The average number of incidents per day could then be calculated by dividing this number by the number of days in scope, being 181 days. This results in an average of 167 incidents/day. We applied the same calculation method for interactions. The results can be found in table 2.

Table 3: Overview of the number of incidents and interactions related to the CI impacted

	# Tickets	# Tickets / % CI
# Incidents related to (1)	15796	30260,43636
# Related interactions (2)	54860	72645,23279

In order to get the total average workload, we added the average number of changes/day based on the configuration items on which occurred only changes with no related incidents or interactions.

Table 4: Average workload in steady-state operation

Average workload in steady-state operation	
AVG Changes/day in STEADY-STATE	128,281768
AVG Incidents/day in STEADY-STATE	167,1847313
AVG Interactions/day in STEADY-STATE	401,3548773
AVG workload in STEADY-STATE	696,8213765

The average workload in steady-state operation is approximately 697 tickets per day, which is almost half of the average workload with changes.

6.2.3 Planned versus actual timing

Based on the actual and planned timing, we see that on average **30% more time** is **planned** than is actually needed to implement the change. In general, the change implementation also starts 49 hours later than was initially planned, although it ends about 20 hours earlier than planned. And on average the change implementation is planned to end 20 days earlier than was requested. So although the duration of a change is overestimated, it generally starts later than was planned.

6.2.4 Evolution of the number of incident activities for a service component

The method for calculating and visualizing the evolution of the number of incident activities per Service Component is summarized below:

First a summary of the number of incident activities per incident is created by counting the number of distinct “IncidentActivity_Number IncAct” values of an inner join on IncidentID of the incident and incident activities table. To continue a specific Service Component (in this case WBS000152) is selected for which an average number of incident activities is calculated. Afterwards for each incident ID the difference between the number of related Incident activities and the average number of Incident Activities for that specific Service component (in this case 12,3876) is calculated. The difference value is set to 0 for negative values (for these incidents the total number of related incident activities is less than the average) after which the sum per week (based on the “Start Time” of the incident) of the difference values is calculated (please note that thus only the number of incident activities above the Service Component average per incident are included in the sum). After division of these weekly values by the number of incidents that occurred that week below graph of the evolution of the number of incident activities can be created. We clearly note for the selected **Service Component**

WBS000152 a downward trend, supported by the linear trend line, in the number of incident steps (incident activities) and thus an evolution to a more efficient incident resolution.

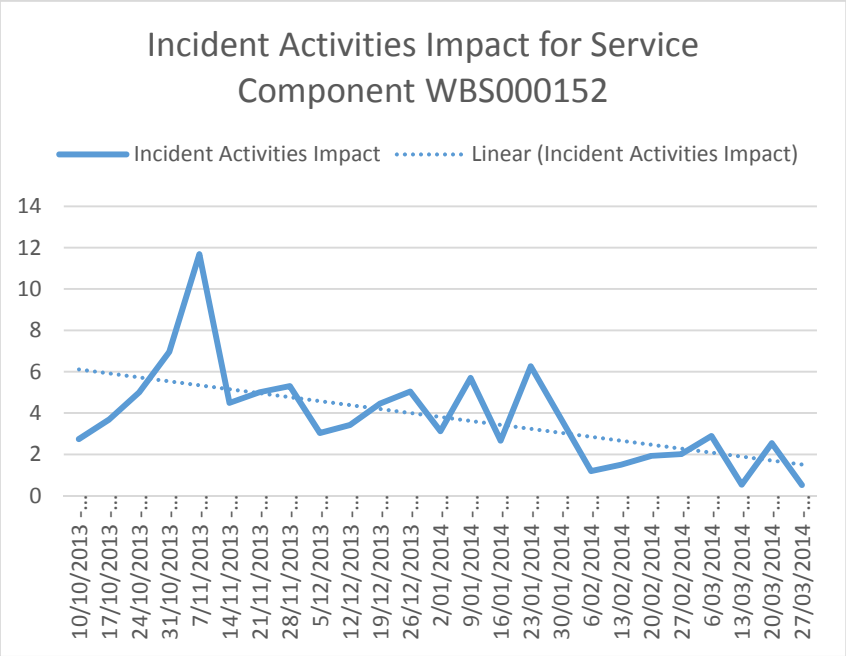


Figure 12: Incident Activities Impact for Service Component WBS000152

6.3 Process mining

6.3.1 Data transformation

In order to be able to use the process mining tool Disco we performed a number of data transformation activities. In each of the source tables Change Detail, Incident Detail and Interaction Detail, we added a column indicating the Activity parameter. We defined the activities as change, incident and interaction. We also created an additional timestamp column in which we copy pasted the start and end time that we wanted to use for the different activities. We used ‘Open Time (First Touch)’ and ‘Close Time’, ‘Open Time’ and ‘Close Time’ and ‘Actual Start’ and ‘Actual End’ for interactions, incidents respectively changes. For interactions and incidents the timestamps are automatically generated and for changes they are manually filled in. We decided not to use the automatically generated timestamps for changes because we understood from the process owner (cfr. Q&A Webinar BPI Challenge) that the abovementioned timestamps are a better indication of the actual situation. We then appended the data tables Change Detail, Incident Detail and Interaction Detail via SQL Server Management Studio (SSMS).

To import the table in Disco, we need at least the following parameters: case ID, timestamp and activity. We defined these parameters as follows:

- Case ID: CI Name (aff)
- Activity: Event (i.e. Change, Incident or Interaction)
- Timestamp: Open Time and Close Time (for changes, these are the fields Actual Start date and Actual End date)

In addition we added a number of attributes: CI Type (aff), CI Subtype (aff), Service Component WBS (aff), Change Type and Risk Assessment.

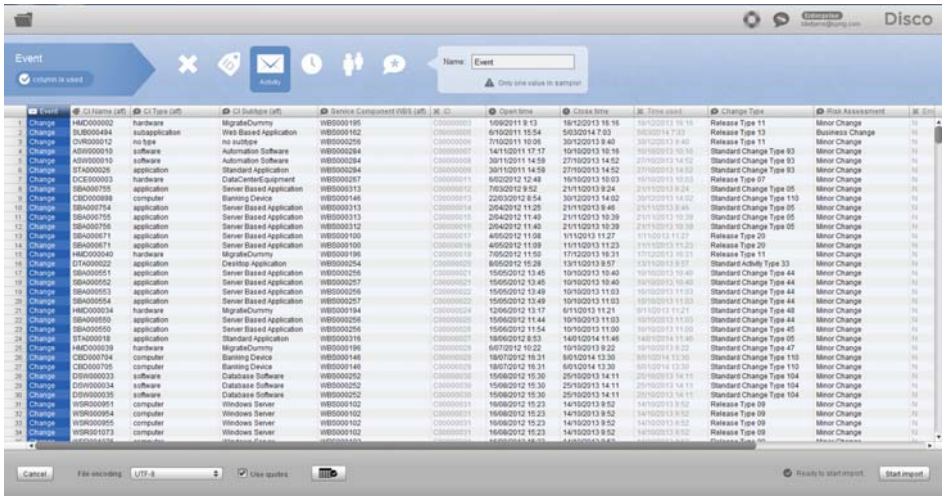


Figure 13: Import of data set in Disco

6.3.2 Process discovery

The output from the import described in section 6.3.1 is shown in figure 13. On the left site the process map with all possible paths is shown, on the right site the process map with the most followed paths. Both process maps show the case frequency per path.

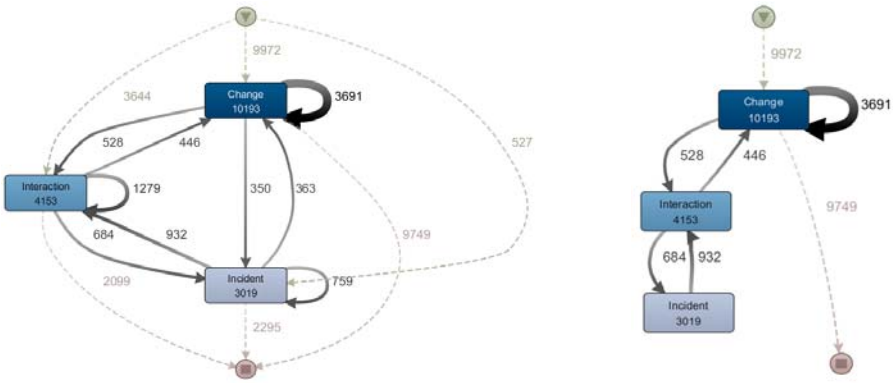


Figure 14: Output from Disco (left: all possible paths, most followed path)

When looking at the process maps, it can be seen that the **most frequent path** followed between activities is **Change -> Change**. This means that on most configuration items no incidents or interactions occur between 2 changes. This is also evident from the case statistics in Disco. Table 5 gives an overview of the top 10 most followed flows.

Top 10 most followed flows		
1	Change	39,43%
2	Change -> Change	12,48%
3	Interaction	9,16%
4	Interaction -> Incident	7,98%
5	Change -> Change -> Change	5,47%
6	Change -> Change -> Change -> Change	3,23%
7	Incident	3,02%
8	Change -> Change -> Change -> Change -> Change	1,77%
9	Interaction -> Interaction	1,13%
10	Change -> Change -> Change -> Change -> Change -> Change	1,08%
		84,75%

Table 5: Top 10 most followed flows based on the case statistics in Disco

We consider changes that do not cause any interactions or incidents to be well planned by the Change Management Department. In order to identify these changes, we used Disco's filtering options. We made use of the attribute filter to filter out all cases in which interactions and incidents occurred.

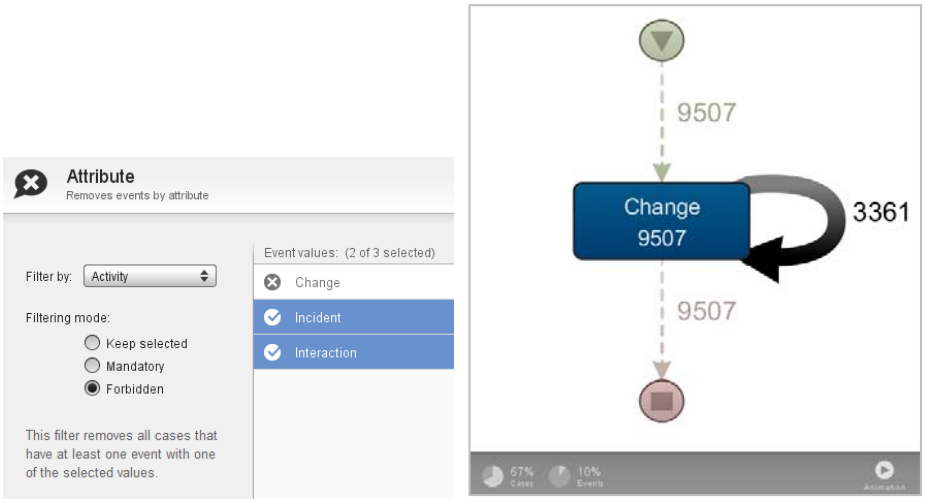


Figure 15: Configuration of attribute filter (*left*) and output (*right*)

We noted that in **67%** of the cases a change implementation **does not affect** the **Service Desk** or **IT Operations**. In 58,65% of these cases only 1 change has occurred during the period in scope. In 18,57% of the cases 2 changes occurred and in 8,14% of the cases 3 changes occurred. The maximum number of changes that has occurred on one configuration item is 303 changes. The related configuration item is DBR00114 and belongs to service component WBS000224.

In the same way, we can filter out the cases on which only interactions and incidents occurred. In figure 15 the related filter configurations can be seen. We noted that in 27% of the cases only interactions and incidents occurred. Based on the configuration item, these interactions and incidents are thus not correlated to any of the changes. In 32% of these cases only 1 interaction has occurred during the period in scope. In 28,56% of the cases an interaction and an incident occurred and in 10,81% of the cases only an incident occurred. The maximum total number of interactions and incidents occurred in a case is 10990 events and concerned configuration item SAN000182, which belongs to service component WBS000128. 96,40% of the events were interactions with a median duration of 2 minutes.

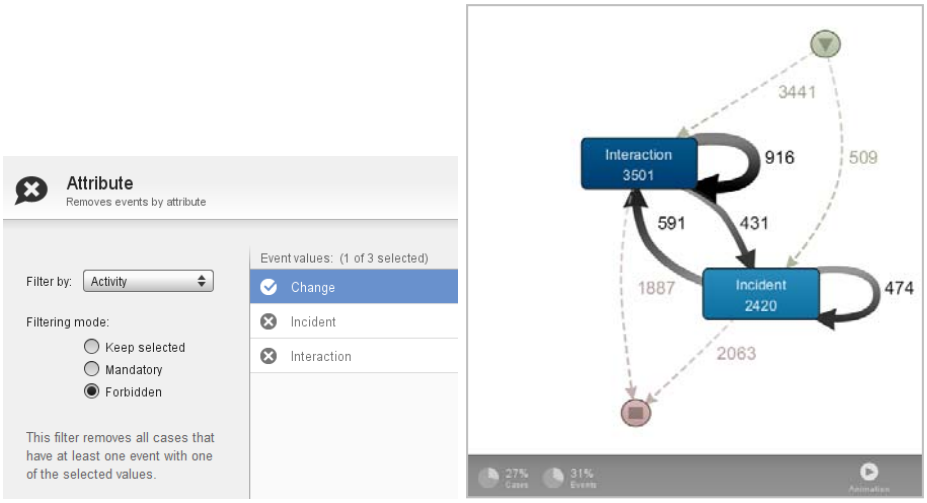


Figure 16: Configuration of attribute filter (*left*) and output (*right*)

In the remaining cases changes as well as incidents and/or interactions have occurred. The **most frequently followed path**, i.e. in 7,73% of the filtered cases, is **change-interaction-incident** as shown in figure 16. In 4 out of the top 6 possible patterns, a change is directly followed by an interaction. However, this does not always mean that there is a relation between both events. depending on the time between both events it may or not may be related to the change. If we consider 7 days to be the maximum duration for an interaction to be related to the change, we see that there only remain 2 cases in the first variant.

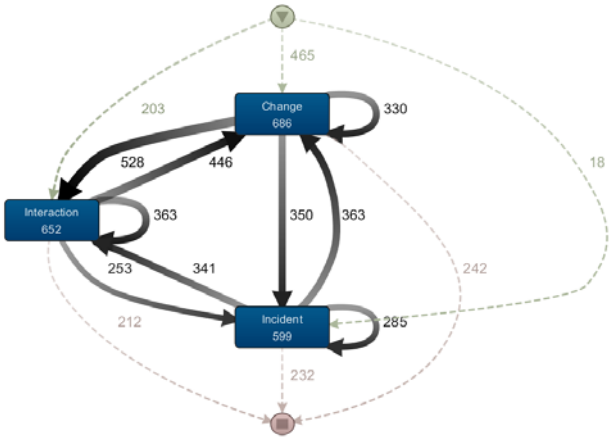


Figure 17: Process map from Disco only taking into account cases in which both changes and incidents and/or interactions occurred

6.3.3 Case analysis

We noted that on service component WBS000073 the most events (including interactions, incidents and changes) occurred during the period in scope, in total 21,15% of all events. The share of the other service components in the events decreases exponentially as can be seen from figure 17. The second most common service component is WBS000128 which includes 6,65% of all events.

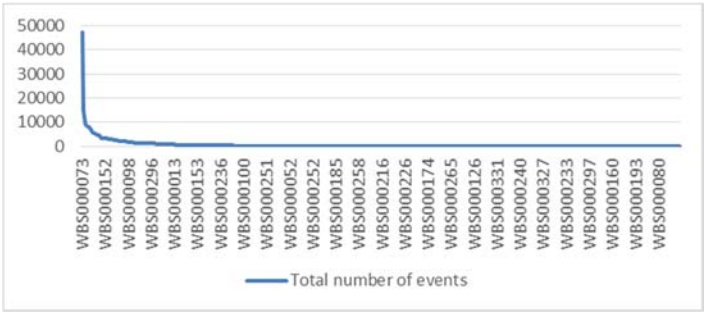


Figure 18: Distribution of events across the service components

If we have a closer look at WBS000073, we can see that 1,03% of the events concerns changes, 28,17% concerns incidents and the remaining 70,80% concerns interactions. This indicates that WBS000073 has a significant higher impact on the Service Desk and IT Operations than the other service components. Because of its big part in the workload of the Service Desk and IT Operations we decided to have a closer look at the patterns followed.

Because we want to identify any relationship between the implementation of a change and incidents and interactions, we performed further investigation of the 25 cases on which both types of events occurred. It can be seen that the number of events per case is not unequally distributed.

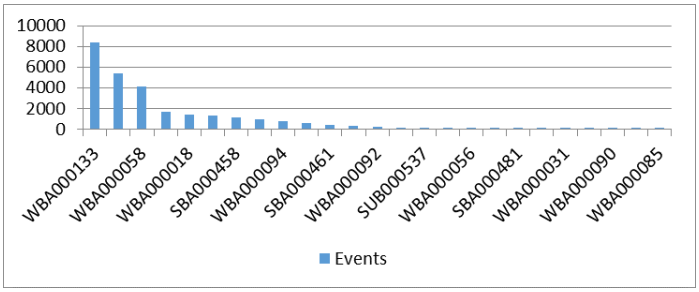


Figure 19: Distribution of events across cases (CI's) for service component WBS000073

Taking 50% of the possible paths, we can see that most cases start with a change and are followed by a combination of interactions and incidents. Interactions follow each other directly and for those that cannot be solved in first line, an incident is created.

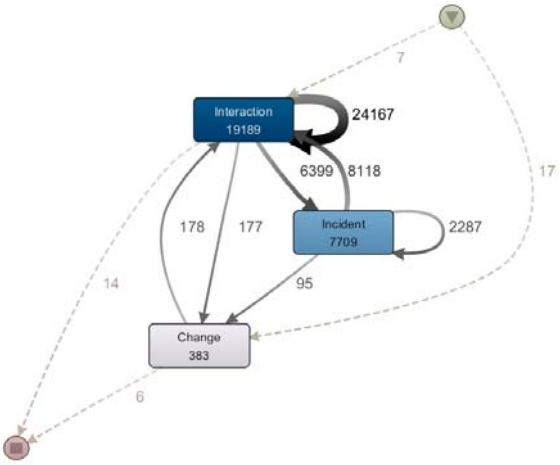


Figure 20: Process map with 50% of the possible paths for service component WBS00073

7 Model Creation & Validation

7.1 Data transformation

In order to create a dataset reflecting numerous attributes for a certain change the following a table, with conditions and content as described below, was created.

As **no unique key** was available for linking Changes to Interactions and or Incidents we opted to relate Interactions and or Incidents to changes **based on Service Component WBS (aff)** in order to create attributes.

Based on this information a combination of following fields was used as a unique key to define a certain change for which attributes were to be retrieved from the Interaction and Incident (Activities) Tables.

- **Change ID**: The unique ID of the Change
- **Service Component WBS**: As the impact of a change will be determined on Service component WBS. Please note that this field is not unique to Change ID and therefore generates another 1338 lines.
- **Actual Start:** As we interpreted the Actual Start as the moment of actual implementation the change we considered this as the latest moment before a change to predict the impact of the change based on, at that moment, current information . This value is key as it determines together with the selected time interval of an

attribute the number of Interactions and or Incident (Activities) that are related to a change. As this field is key for the calculation of our attributes and most of the attributes are calculated on a 7 day time difference basis we **filtered out all changes** with Actual Start **before the 7th of October and after the 24th of March** in order to make sure that the data did not skew at the time limits of our data set. Furthermore we assumed that when no Actual Start was filled in the Change record, no implementation occurred and therefore no change occurred. All **changes with a missing Actual Start date were filtered out** as well.
Please note that the actual start field is not unique to Change ID and Service Component WBS and therefore generates another 7 lines.

With filters included the total number of lines and thus total number of considered changes for which attributes are determined is **16295**

Amongst others, we added the following types of attributes:

- Number of interactions opened in the week after the actual start of the change (based on 'Open Time (First Touch)' from the Detail Interaction table and 'Actual Start' from the Change Detail table)
- Number of incidents opened in the week before the actual start of the change (based on 'Open Time' from the Detail Incident table and 'Actual Start' from the Change Detail table)
- Average handle time of the interactions opened during the week before the actual start date of the change (based on 'Open Time (First Touch)' and 'Close Time' from the Detail Interaction table and 'Actual Start' from the Change Detail table)
- Number of incidents during the week before the actual start date of the change with category 'request for information' (based on 'Open Time' and 'Category' from the Detail Incident table and 'Actual Start' from the Change Detail table)
- The weekday on which the change occurred (based on 'Actual Start' from the Detail Change table)
- The impact of a change on interactions as the difference of the number of interactions in the week after the change and the number of interactions in the week before the change. This was expressed as a discrete variable "IntImpact7Days" where 1 meant that the number of interactions in the week after the change was larger than the number of interactions in the week before the change. For differences equal to or smaller than 0 the discrete variable was set to 0. (based on 'Open Time (First Touch)' from the Detail Incident table and 'Actual Start' from the Change Detail table). Based on the results (table 6) for the defined 'IntImpact7Days' attribute we can for example conclude that **changes to Service component WBS000161** will have a **very limited impact** on the number of interactions in the following week as out of a total of 1653 changes only 1 change had more interactions in the following week than in the week before the change.

ServiceComponentWBS	IntImpact7Days	Count of Changes
WBS000161	0	1652

WBS000161	1	1

- **Table 6:** IntImpact7Days for changes on WBS000161

For the linking of the tables we used a **left outer join** (with above described table of changes always as the left table), meaning that, for our output, we considered all changes even if in no Interactions and Incidents affecting the Change's Service component were retrieved.

7.2 Model creation

Using the Weka tool, we executed two different data mining algorithms on our prepared data table:

- The linear regression algorithm 'LinearRegression' which was by default provided by the Weka tool itself.
- The C4.5 algorithm in order to generate a decision tree (this is the standard algorithm within the data mining domain of decision trees);

With these data mining algorithms we tried to predict whether a change implementation has a positive or negative impact on the Service Desk and/or IT Operations.

For the linear regression, the target variable was defined by taking the difference of the total amount of interactions and incidents over 7 days after the change and the total amount of these before the change. We also predicted this value for interactions and incidents separately (i.e. difference of total amount of interactions/incidents over 7 days after and before the change).

For the decision tree algorithm, the target variable was defined as follows:

- **1: positive difference** (there are more interactions and incidents after the change than before, so the change has a negative impact on the service desk/IT operations)
- **0: negative difference** (there are less or an equal amount of interactions and incidents after the change than before, so the change has a positive impact on the service desk/IT operations)

Note that we only included those variables as input for the data mining algorithms that contain information that is known before the start of a change record. In this way the generated model will be more robust.

Within the Weka tool, we used a percentage split of 66% in order to divide the data between the training set (66% of the records – to build up the model) and the test set (34% of the records – to test the model). This percentage division is also the standard setting that is defined by the tool.

7.3 Model validation

a) Linear Regression

For the linear regression technique, we used all the standard settings as defined within Weka. The prediction value of the linear regression model is assessed by the correlation coefficient.

Interactions

For the Linear Regression Model we refer to the appendix A.1. The data dictionary containing all the used variables can be found in Appendix A.7.

Model accuracy	Value
Correlation coefficient	0.5596

The created model has a predictive value as the correlation coefficient is larger than 0.5. The 5 attributes within the model that have the most impact (based on the absolute value of the coefficient) are:

1. the origin of the change being an interaction
2. the origin of the change being a problem
3. the impact of a change being a (major) business change
4. the number of incidents with a very low impact
5. the deviation of the number of incidents on the 3rd day before the change implementation from the average number of incidents spread over 21 days before the change.

Incidents

For the Linear Regression Model we refer to the appendix A.2. The data dictionary containing all the used variables can be found in Appendix A.7.

Model accuracy	Value
Correlation coefficient	0.6441

The created model has a predictive value as the correlation coefficient is larger than 0.5. The 5 attributes within the model that have the most impact (based on the absolute value of the coefficient) are:

1. the amount of incidents before the change with priority ‘3’
2. the impact of a change being a (major) business change
3. the amount of incidents before the change with priority ‘5’ (very low)
4. the amount of incidents before the change with urgency ‘3’

5. the amount of incidents before the change with impact ‘3’

Interactions and Incidents Combined

For the Linear Regression Model we refer to the appendix A.3. The data dictionary containing all the used variables can be found in Appendix A.7.

Model accuracy	Value
Correlation coefficient	0.5936

The created model has a predictive value as the correlation coefficient is larger than 0.5. The 5 attributes within the model that have the most impact (based on the absolute value of the coefficient) are:

1. the impact of a change being a (major) business change
2. the amount of incidents before the change with impact ‘5’ (very low)
3. the amount of incidents on the 4th day before the change
4. the deviation of the number of incidents on the 3rd day before the change implementation from the average number of incidents spread over 21 days before the change
5. the amount of incidents before the change with urgency ‘3’

Both the impact of a change on the interactions and on the incidents could be reasonably modeled. The model describing the impact on the combination of interactions and incidents has however the highest predictive value.

We notice that over all the three linear regression models the Risk Assessment Variable in the Change Detail table describing the impact of a change, plays an important role. Also the number of incidents before a change with an average Impact/Urgency/Priority (value ‘3’) is influential. And finally the deviation of the number of incidents on the 3rd day before the change implementation from the average number of incidents spread over 21 days before the change, has a big influence in two out of the three models.

b) Decision Tree

In order to limit the size of the tree to an acceptable and readable magnitude, we adjusted a few of the basic settings:

- confidenceFactor: The confidence factor used for pruning the tree. A confidence factor of 0.0000001 is the lowest possible value that Weka accepts in order to create a decision tree.

- minNumObj: The minimum number of instances per tree. A value of 100 is an acceptable value, and increases the accuracy of the model.

The prediction value of a decision tree is assessed by the following statistics:

- Correctly Classified Instances (%): the percentage of the records that are correctly classified with the model.
- Confusion Matrix: shows the number of instances that are misclassified for each classification.
- ROC Area: a value of 0.5 or lower indicates a random model.

Interactions

We adjusted the following parameters before running the algorithm;

Options	Value
confidenceFactor	1.0E-7
minNumObj	100

For the Decision Tree Model we refer to the appendix A.4. The data dictionary containing all the used variables can be found in Appendix A.7.

The generated decision tree has 18 leaves and has a size of 35. The first splits of the tree are made on the following attributes:

1. the deviation of the number of changes on the 7th day before the change implementation from the average number of changes spread over 21 days before the change
2. the amount of changes on the 3rd day before the change implementation
3. the amount of changes on the 6th day before the change implementation

Model accuracy	Value
Correctly classified instances (%)	85.8664 %
ROC Area	0.824

Confusion matrix		
Classification	0	1
0	4451	142
1	620	327

The decision tree model for determining the impact of a change on the number of interactions has a higher predictive value than the linear regression model.

Incidents

We adjusted the following parameters before running the algorithm;

Options	Value
confidenceFactor	1.0E-7
minNumObj	100

For the Decision Tree Model we refer to the appendix A.5. The data dictionary containing all the used variables can be found in Appendix A.7.

The generated decision tree has 18 leaves and has a size of 35. The first splits of the tree are made on the following attributes:

1. the total handle time of all interactions on the 7th day before the change implementation
2. the deviation of the number of changes on the 3rd day before the change implementation from the average number of changes spread over 21 days before the change
3. the deviation of the number of incidents on the 1st day before the change implementation from the average number of changes spread over 21 days before the change

Model accuracy	Value
Correctly classified instances (%)	86.2455 %
ROC Area	0.824

Confusion matrix		
Classification	0	1
0	4451	142
1	620	327

The model prophesying the impact of a change implementation on the number of incidents is more accurate than the Linear Regression Model, and even more than that for the Interactions Decision Tree Model.

Interactions and Incidents Combined

We adjusted the following parameters before running the algorithm;

Options	Value
confidenceFactor	1.0E-7
minNumObj	50

For the Decision Tree Model we refer to the appendix A.6. The data dictionary containing all the used variables can be found in Appendix A.7.

The generated decision tree has 36 leaves and has a size of 71. The first splits of the tree are made on the following attributes:

1. the amount of changes on the 3rd day before the change implementation
2. the amount of incidents before the change implementation with an impact of value '5' (very low) the total handle time of all interactions on the 7th day before the change implementation
3. the deviation of the number of interactions on the 7th day before the change implementation from the average number of changes spread over 21 days before the change

Model accuracy	Value
Correctly classified instances (%)	69.6029 %
ROC Area	0.649

Confusion matrix		
Classification	0	1
0	3284	137
1	1547	562

The Decision Tree Model depicting the combination of interactions and incidents has a lower predictive value that the two separate models.

In general, the Decision Tree Models are able to give better forecasts than the Linear Regression Model.

We noted that the deviation on the amount of interactions/incidents/changes (on a certain day) from the 21day average is depicted in all three models. Also the amount of interactions/incidents/changes on a certain day is influential.

7.4 SASS Microsoft Decision Tree Model

In order not to limit ourselves to only 1 tool and to support the validity of the Weka tool results above the SASS Data Mining Tool was used as well.

Through the means of SSAS a mining structure was set up in order to create predictive analyses using the different Microsoft Data Mining Algorithms.

In order to predict the value of column “IncImpact7Days” a Microsoft Decision Trees model was built based on:

- A single key column = “Change ID Ch “ (The ID of the change)
- A predictable column = “IntImpact7Days” (discrete variable to indicate if there was an increase in interactions after the change)

- Input columns = All columns as listed in de data dictionary that could be known exactly at the Actual Start of a change. (E.g.: ChaBday2, IncUrgency4, etc…)

The model's training population was set to 70% (11406 lines) of the mining model data thus leaving 30% (4888 lines) of the model data set as testing population. Based on the trained Microsoft Decision Tree model (94 nodes) the following Accuracy Chart for model 'IntIncIncAcComb7DaysImpactDT' was created to reflect the results of the predicted values for our testing population. The graph summarizes the accuracy of the predictive model compared to the ideal predictive model (being 100% of the overall population predicted 100% correct).

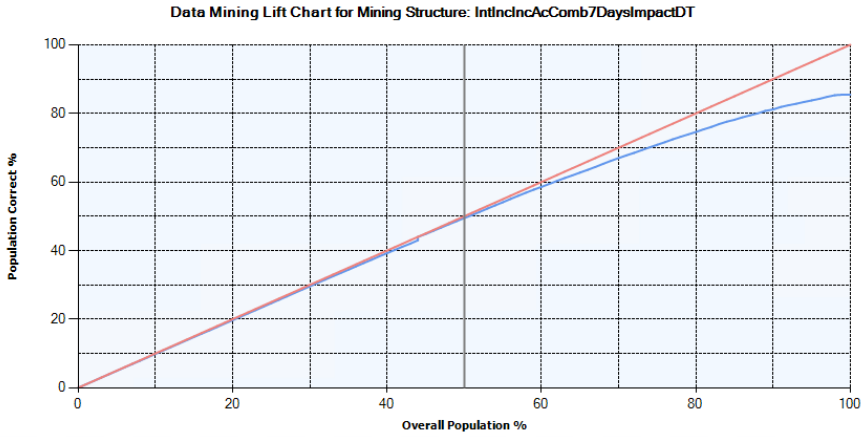


Figure 21: SSAS Lift Chart for variable IntImpact7days

We noted that the accuracy line for the predicted values is relatively close to the ideal model implying that the created model can be considered predictive. We furthermore noted in the above the accuracy chart (upper right corner) and the below model report that 86% of the "IntImpact7days" values were predicted correctly

Output Column Information for 'Int Impact7 Days'	
Content Type	Discrete
Important splits	Int Urgency5, Service Component WBSaff Ch, Cha B14 Days Average, Inc Impact3, Average Int Handle Time
Pass	86 %
Fail	14 %
Log Score	-0,281948328
Lift	0,195958138

Table 7: SSAS Data Mining Model Report for variable IntImpact7Days

More detail on the prediction results of the model is included in the Classification Matrix below.

Predicted	0 (Actual)	1 (Actual)
0	3546	399
1	161	343

Table 8: SSAS Data Mining Classification Matrix for variable IntImpact7Days

The following 5 Attributes are recognized by the model as important splits for the outcome of the prediction:

- IntUrgency5
- ServiceComponentWBS_aff_Ch
- ChaB14DaysAverage
- IncImpact3
- Average Int Handle Time

To demonstrate a Microsoft Decision Tree decision of the model that is based on one of these important splits we include the following **fragment** of the Microsoft Decision Tree Model (it is the first split of the model):

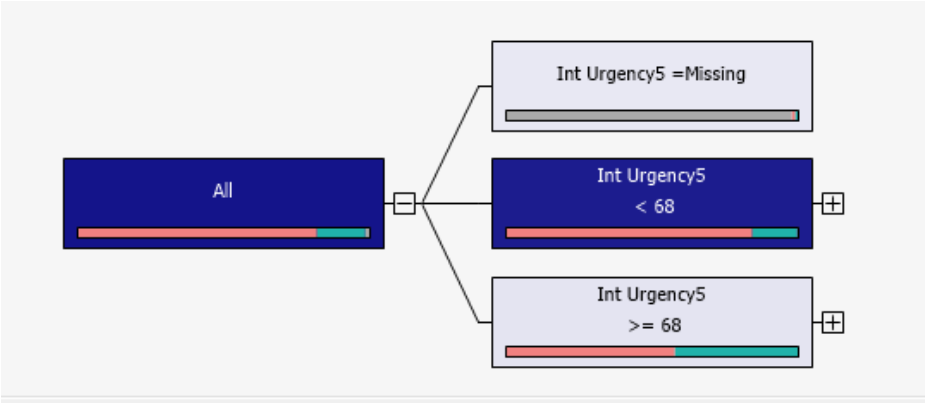


Figure 22: Microsoft Decision Tree main split on Int Urgency5

Based on the number of urgency 5 interactions in the past week the tree will divide the population. If no interaction urgency level is found (Int Urgency5 =Missing; only 160 cases) it means that for the Service Component of the change no related interactions were found in the considered time period. The below graph shows more detail of the "Int Urgency5 <68" node to clarify the node status bars.

Value	Cases	Probability	Histogram
☑ 0	8282	83,75%	
☑ 1	1531	15,49%	
☑ Missing	75	0,76%	

Figure 23: Detailed probability information of node IntUrgency5<68

From the above table we can conclude that for changes with less than 68 related “urgency=5 Interactions” in the past week it is far more likely (83,75%) that there won’t be an increase when comparing the total number of interactions after the “Actual Start” of the change with the total number of interaction before the “Actual Start” of the change. This seems logic as the higher the number of “urgency=5 Interactions” in the week before the Actual Start of the change, the more chance that the total number of interactions in the following week will increase compared to this preceding week.

An export (until level 5) of the above discussed Microsoft Decision Tree for determining column 'Int Impact7 Days' can be found in Appendix A.8.

8 Further research

The data models as have been currently created could be refined further in several ways.

In the current model the impact of weekend days is not considered. We do however incorporate the specific day of the week. Tor further improved the created models, we could also take into account the weekend days, and even more the school holidays and/or the holidays of the personnel itself. This way, unpopular vacation days will not bias the predictive model.

Furthermore, we noticed that the time range of the provided data is rather limited. All interactions/incidents/changes are closed between 1/10/2013 and 31/03/2014. In order to have a reliable base to create a model on, at least a two year data range should be provided.

Finally, most of the fields in the provided data tables have been filled out manually. With less room for manual input, the model will become more robust.

9 Appendix

- **Appendix A.1 – Linear Regression Model for interactions**

IntRealImpact7Days =

0.8423 * DayOfChange +
3.5215 * RiskAssessmentCh=BusinessChange,MajorBusinessChange +
-8.3059 * OriginatedFromCh=Interaction,Problem +
6.4639 * OriginatedFromCh=Problem +
-0.27 * ChaB7DaysAverage +
0.3361 * ChaB14DaysAverage +
-0.3065 * ChaB21DaysAverage +
0.0778 * ChaBday4 +
0.2061 * ChaBday7 +
0.2166 * ChaBD1Av21Diff +
0.124 * ChaBD6Av21Diff +
-0.1683 * ChaBD7Av21Diff +
-1.8586 * IncImpact3 +
4.5643 * IncImpact5 +
0.253 * IncUrgency3 +
-0.6681 * IncUrgency4 +
0.2063 * IncUrgency5 +
-2.005 * IncPriority5 +
-1.9851 * IncCatRFI +
-1.1602 * IncCCOther +
0.4879 * IncCCSoftware +
-0.0133 * IncB7DaysHandleTimeHrs +

0.1603 * AverageIncHandleTime +

-0.3145 * IncB21DaysAverage +

-0.1499 * IncBday1 +

1.3292 * IncBday3 +

-2.3877 * IncBday4 +

1.3724 * IncBday5 +

-0.6377 * IncBday7 +

1.3945 * IncBD1Av21Diff +

0.6477 * IncBD2Av21Diff +

2.6249 * IncBD3Av21Diff +

2.2496 * IncBD4Av21Diff +

-0.155 * IncBD5Av21Diff +

1.1245 * IncBD6Av21Diff +

1.7882 * IncBD7Av21Diff +

2.3848 * IntImpact3 +

-1.494 * IntImpact5 +

-2.4952 * IntUrgency3 +

0.1641 * IntPriority5 +

1.2712 * IntCatRFI +

-1.645 * IntB14DaysAverage +

0.4426 * IntBday1 +

0.4579 * IntBday2 +

0.2089 * IntBday3 +

0.8202 * IntBday4 +

0.5979 * IntBday5 +

0.8706 * IntBday6 +

0.1655 * IntBday7 +

-0.7673 * IntBD1Av21Diff +

-0.2366 * IntBD2Av21Diff +

-1.5131 * IntBD3Av21Diff +

0.0861 * IntBD4Av21Diff +

-1.2187 * IntBD5Av21Diff +

-0.0825 * IntBD6Av21Diff +

-0.8331 * IntBD7Av21Diff +

2.5974 * IncB14DaysAverage +

-1.577

- **Appendix A.2 – Linear Regression Model for incidents.**

IncRealImpact7Days =

0.5303 * DayOfChange +

2.8881 * RiskAssessmentCh=BusinessChange,MajorBusinessChange +

-0.5852 * OriginatedFromCh=Problem,Interaction +

-0.0008 * PlannedTimeHrs +

-0.0531 * ChaB14DaysAverage +

-0.0562 * ChaB21DaysAverage +

0.1328 * ChaBday4 +

0.1417 * ChaBday5 +

-0.0805 * ChaBday7 +

0.1301 * ChaBD1Av21Diff +

-0.0564 * ChaBD4Av21Diff +

-0.0771 * ChaBD5Av21Diff +

0.084 * ChaBD6Av21Diff +

0.1102 * ChaBD7Av21Diff +

-1.4177 * IncImpact3 +

-0.6982 * IncImpact5 +

-1.6291 * IncUrgency3 +

3.0952 * IncPriority3 +

2.1557 * IncPriority5 +

-0.471 * IncCatRFI +

-0.3832 * IncCCOther +

0.0417 * IncCCSoftware +

-0.0042 * IncB7DaysHandleTimeHrs +

0.059 * AverageIncHandleTime +

-1.2962 * IncB21DaysAverage +

-1.1319 * IncBday1 +

-0.8972 * IncBday2 +

0.0491 * IncBday3 +

-0.9649 * IncBday5 +

-0.9637 * IncBday7 +

1.0701 * IncBD1Av21Diff +

0.6086 * IncBD2Av21Diff +

0.43 * IncBD3Av21Diff +

-0.8124 * IncBD4Av21Diff +

0.584 * IncBD5Av21Diff +

-0.7655 * IncBD6Av21Diff +

0.359 * IncBD7Av21Diff +

0.5922 * IntImpact3 +

-1.0174 * IntUrgency3 +

-0.5979 * IntUrgency5 +

0.3362 * IntCatRFI +

0.2734 * IntB21DaysAverage +

0.5447 * IntBday1 +

0.389 * IntBday2 +

0.1243 * IntBday3 +

0.386 * IntBday4 +

0.1081 * IntBday5 +

0.2149 * IntBday6 +

0.1006 * IntBday7 +

-0.6988 * IntBD1Av21Diff +

-0.0787 * IntBD2Av21Diff +

-0.4135 * IntBD3Av21Diff +

0.2159 * IntBD4Av21Diff +

-0.2162 * IntBD5Av21Diff +

0.5158 * IntBD6Av21Diff +

-0.07   * IntBD7Av21Diff +

1.2994 * IncB14DaysAverage +

-1.7753

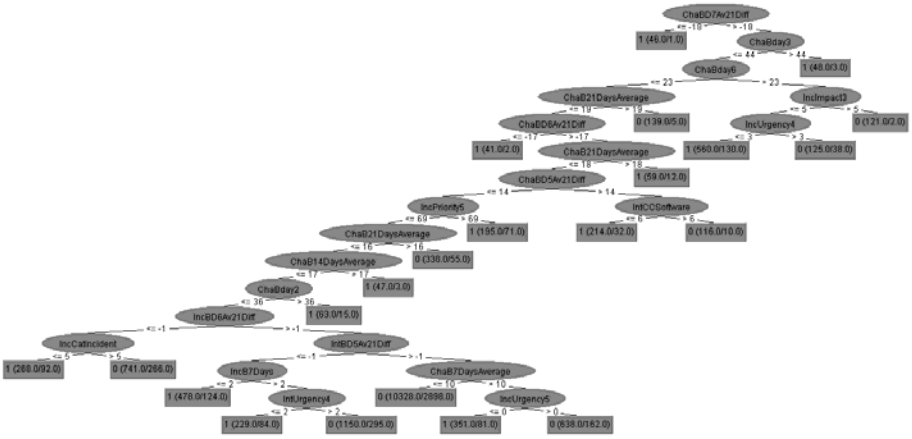
- **Appendix A.3 – Linear Regression Model for interactions and incidents combined.**

  CombinedRealImpact7Days =

  1.2064 * DayOfChange +
  6.0741 * RiskAssessmentCh=BusinessChange,MajorBusinessChange +
  -2.7425 * OriginatedFromCh=Problem,Interaction +
  0.1355 * ChaBday4 +
  0.3056 * ChaBD1Av21Diff +
  0.2064 * ChaBD6Av21Diff +
  -2.5202 * IncImpact3 +
  5.3437 * IncImpact5 +
  -1.5455 * IncUrgency3 +
  1.5557 * IncUrgency5 +
  3.2725 * IncPriority3 +
  -1.6867 * IncPriority5 +
  -2.7422 * IncCatRFI +
  -1.4788 * IncCCOther +
  0.8588 * IncCCSoftware +
  -0.0183 * IncB7DaysHandleTimeHrs +
  0.2288 * AverageIncHandleTime +
  -1.4944 * IncB21DaysAverage +
  -1.3082 * IncBday1 +
  -0.6908 * IncBday2 +
  0.3848 * IncBday3 +
  -4.291 * IncBday4 +
  0.1489 * IncBday5 +
  -1.3886 * IncBday7 +
  1.783 * IncBD1Av21Diff +
  0.388 * IncBD2Av21Diff +
  3.2413 * IncBD3Av21Diff +
  2.689 * IncBD4Av21Diff +
  -0.356 * IncBD6Av21Diff +
  1.4331 * IncBD7Av21Diff +
  2.6696 * IntImpact3 +
  -1.9962 * IntImpact5 +
  -3.1984 * IntUrgency3 +
  1.7275 * IntCatRFI +
  0.8106 * IntBday1 +
  0.8498 * IntBday2 +
  0.8869 * IntBday3 +
  0.3683 * IntBday5 +
  0.8649 * IntBday6 +

0.0919 * IntBday7 +
-1.3885 * IntBD1Av21Diff +
-0.4398 * IntBD2Av21Diff +
-2.5334 * IntBD3Av21Diff +
1.3634 * IntBD4Av21Diff +
-1.1783 * IntBD5Av21Diff +
0.5406 * IntBD6Av21Diff +
-0.9157 * IntBD7Av21Diff +
1.0855 * IncB14DaysAverage +
-2.4391

- **Appendix A.4 – Decision Tree Model for interactions.**



IntB7DaysHandleTimeHrs <= 0.133333

| ChaBD3Av21Diff <= 22: 0 (6543.0/53.0)

| ChaBD3Av21Diff > 22: 1 (42.0/16.0)

IntB7DaysHandleTimeHrs > 0.133333

| IncBD1Av21Diff <= -8: 1 (115.0/33.0)

| IncBD1Av21Diff > -8

| | ChaB14DaysAverage <= 19

| | | IntB21DaysAverage <= 1

| | | | IntB7DaysHandleTimeHrs <= 0.136944: 1 (46.0/9.0)

| | | | IntB7DaysHandleTimeHrs > 0.136944: 0 (5224.0/860.0)

| | | IntB21DaysAverage > 1

| | | | IncB7DaysHandleTimeHrs <= 0: 1 (187.0/57.0)

| | | | IncB7DaysHandleTimeHrs > 0

| | | | | IntB7DaysHandleTimeHrs <= 1.0025: 1 (72.0/17.0)

| | | | | IntB7DaysHandleTimeHrs > 1.0025

| | | | | | PlannedTimeHrs <= 1289

| | | | | | | ChaBD4Av21Diff <= -7: 0 (449.0/64.0)

| | | | | | | ChaBD4Av21Diff > -7

| | | | | | | | IncPriority4 <= 0

| | | | | | | | | IncB7DaysHandleTimeHrs <= 17: 0 (206.0/87.0)

| | | | | | | | | IncB7DaysHandleTimeHrs > 17: 1 (265.0/66.0)

| | | | | | | | IncPriority4 > 0

| | | | | | | | | ChaB14DaysAverage <= 16

| | | | | | | | | | ChaB14DaysAverage <= 10

| | | | | | | | | | | ChaBday6 <= 15

| | | | | | | | | | | | IntB7Days <= 18: 1 (156.0/58.0)

| | | | | | | | | | | | IntB7Days > 18: 0 (2014.0/634.0)

| | | | | | | | | | | ChaBday6 > 15: 1 (77.0/13.0)

| | | | | | | | | | ChaB14DaysAverage > 10: 0 (337.0/1.0)

| | | | | | | | | ChaB14DaysAverage > 16: 1 (152.0/52.0)

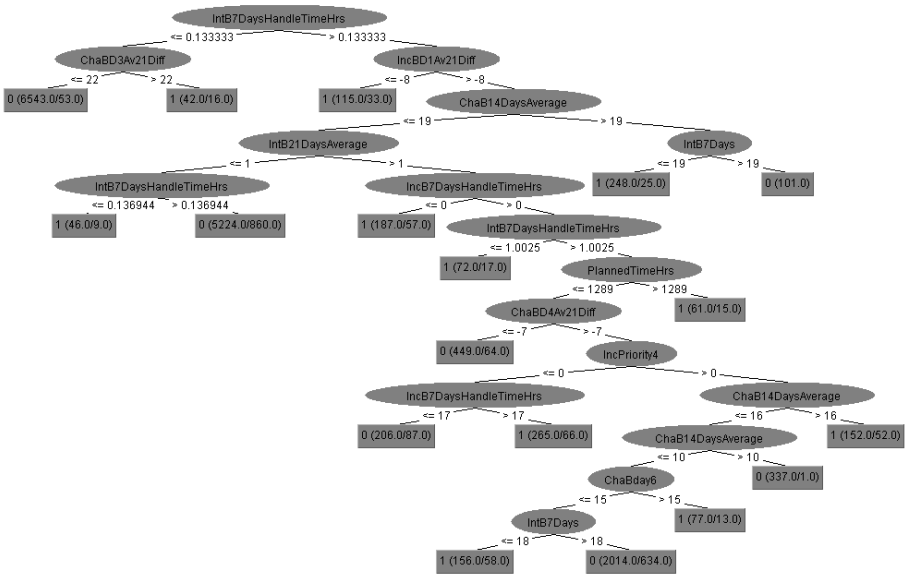
| | | | | | PlannedTimeHrs > 1289: 1 (61.0/15.0)

| | ChaB14DaysAverage > 19

| | | IntB7Days <= 19: 1 (248.0/25.0)

| | | IntB7Days > 19: 0 (101.0)

- **Appendix A.5 – Decision Tree Model for incidents.**



IntB7DaysHandleTimeHrs <= 0.133333

| ChaBD3Av21Diff <= 22: 0 (6543.0/53.0)

| ChaBD3Av21Diff > 22: 1 (42.0/16.0)

IntB7DaysHandleTimeHrs > 0.133333

| IncBD1Av21Diff <= -8: 1 (115.0/33.0)

| IncBD1Av21Diff > -8

| | ChaB14DaysAverage <= 19

| | | IntB21DaysAverage <= 1

| | | | IntB7DaysHandleTimeHrs <= 0.136944: 1 (46.0/9.0)

| | | | IntB7DaysHandleTimeHrs > 0.136944: 0 (5224.0/860.0)

| | | IntB21DaysAverage > 1

| | | | IncB7DaysHandleTimeHrs <= 0: 1 (187.0/57.0)

| | | | IncB7DaysHandleTimeHrs > 0

| | | | | IntB7DaysHandleTimeHrs <= 1.0025: 1 (72.0/17.0)

| | | | | IntB7DaysHandleTimeHrs > 1.0025

| | | | | | PlannedTimeHrs <= 1289

| | | | | | | ChaBD4Av21Diff <= -7: 0 (449.0/64.0)

| | | | | | | ChaBD4Av21Diff > -7

| | | | | | | | IncPriority4 <= 0

| | | | | | | | | IncB7DaysHandleTimeHrs <= 17: 0 (206.0/87.0)

| | | | | | | | | IncB7DaysHandleTimeHrs > 17: 1 (265.0/66.0)

| | | | | | | | IncPriority4 > 0

| | | | | | | | | ChaB14DaysAverage <= 16

| | | | | | | | | | ChaB14DaysAverage <= 10

| | | | | | | | | | | ChaBday6 <= 15

| | | | | | | | | | | | IntB7Days <= 18: 1 (156.0/58.0)

| | | | | | | | | | | | IntB7Days > 18: 0 (2014.0/634.0)

| | | | | | | | | | | ChaBday6 > 15: 1 (77.0/13.0)

| | | | | | | | | | ChaB14DaysAverage > 10: 0 (337.0/1.0)

| | | | | | | | | ChaB14DaysAverage > 16: 1 (152.0/52.0)

| | | | | | PlannedTimeHrs > 1289: 1 (61.0/15.0)

| | ChaB14DaysAverage > 19

| | | IntB7Days <= 19: 1 (248.0/25.0)

| | | IntB7Days > 19: 0 (101.0)

- **Appendix A.6 – Decision Tree Model for interactions and incidents combined.**

ChaBday3 <= 41

| IncImpact5 <= 69

| | IntBD7Av21Diff <= -1

| | | ChaBday5 <= 37

| | | | IntPriority3 <= 0: 1 (488.0/109.0)

| | | | IntPriority3 > 0

| | | | | ChaBday3 <= 37

| | | | | | ChaBday4 <= 43

| | | | | | | ChaB21DaysAverage <= 20

| | | | | | | | ChaBD2Av21Diff <= 20

| | | | | | | | | ChaB14DaysAverage <= 16

| | | | | | | | | | ChaBD3Av21Diff <= 2

| | | | | | | | | | | ChaBD3Av21Diff <= -8: 0 (111.0/21.0)

| | | | | | | | | | | ChaBD3Av21Diff > -8

| | | | | | | | | | | | IncImpact3 <= 2: 1 (404.0/142.0)

| | | | | | | | | | | | IncImpact3 > 2: 0 (276.0/81.0)

| | | | | | | | | | ChaBD3Av21Diff > 2: 0 (247.0/48.0)

| | | | | | | | | ChaB14DaysAverage > 16: 1 (227.0/35.0)

| | | | | | | | ChaBD2Av21Diff > 20: 0 (31.0/1.0)

| | | | | | | ChaB21DaysAverage > 20: 0 (45.0/10.0)

| | | | | | ChaBday4 > 43: 1 (26.0/1.0)

| | | | | ChaBday3 > 37: 0 (35.0)

| | | ChaBday5 > 37: 1 (44.0)

| | IntBD7Av21Diff > -1

| | | IncBD3Av21Diff <= -1

| | | | IncB7Days <= 5: 1 (186.0/16.0)

| | | | IncB7Days > 5

| | | | | ChaBday6 <= 29

| | | | | | ChaB7DaysAverage <= 17: 0 (806.0/268.0)

| | | | | | ChaB7DaysAverage > 17: 1 (57.0/16.0)

| | | | | ChaBday6 > 29: 1 (55.0/8.0)

| | | IncBD3Av21Diff > -1

| | | | IncBD5Av21Diff <= 1

| | | | | ChaBday5 <= 20

| | | | | | ChaBD5Av21Diff <= -17: 1 (31.0/7.0)

| | | | | | ChaBD5Av21Diff > -17

| | | | | | | ChaBday3 <= 33

| | | | | | | | ChaBD3Av21Diff <= 22

| | | | | | | | | IntBD3Av21Diff <= -2

| | | | | | | | | | IntBday5 <= 7

| | | | | | | | | | | IncBD6Av21Diff <= 0: 1 (205.0/36.0)

| | | | | | | | | | | IncBD6Av21Diff > 0: 0 (130.0/41.0)

| | | | | | | | | | IntBday5 > 7: 0 (132.0/18.0)

| | | | | | | | | IntBD3Av21Diff > -2

| | | | | | | | | | IncImpact5 <= 7

| | | | | | | | | | | IntCatRFI <= 12

| | | | | | | | | | | | IntCatRFI <= 4

| | | | | | | | | | | | | IntImpact5 <= 11

| | | | | | | | | | | | | | ChaB7DaysAverage <= 8: 0 (8595.0/2647.0)

| | | | | | | | | | | | | | ChaB7DaysAverage > 8

| | | | | | | | | | | | | | | IntB7DaysHandleTimeHrs <= 0.029444: 1 (468.0/109.0)

| | | | | | | | | | | | | | | IntB7DaysHandleTimeHrs > 0.029444: 0 (1386.0/414.0)

| | | | | | | | | | | | | IntImpact5 > 11: 1 (115.0/23.0)

| | | | | | | | | | | | IntCatRFI > 4: 0 (404.0/63.0)

| | | | | | | | | | | IntCatRFI > 12

| | | | | | | | | | | | ChaBD7Av21Diff <= 8: 0 (92.0/32.0)

| | | | | | | | | | | | ChaBD7Av21Diff > 8: 1 (50.0/5.0)

| | | | | | | | | | IncImpact5 > 7: 0 (220.0/73.0)

| | | | | | | | ChaBD3Av21Diff > 22: 1 (28.0/3.0)

| | | | | | | ChaBday3 > 33: 0 (57.0/8.0)

| | | | | ChaBday5 > 20

| | | | | | IncB7Days <= 11

| | | | | | | IncUrgency5 <= 1: 1 (320.0/62.0)

| | | | | | | IncUrgency5 > 1: 0 (50.0/10.0)

| | | | | | IncB7Days > 11: 0 (61.0)

| | | | IncBD5Av21Diff > 1: 0 (644.0/98.0)

| IncImpact5 > 69

| | IntUrgency5 <= 190: 1 (138.0/21.0)

| | IntUrgency5 > 190: 0 (71.0/7.0)

ChaBday3 > 41: 1 (60.0/6.0)

- **Appendix A.7 – Data dictionary.**

Field ***Range***	**Field Definition**	**Range**
ChangeID Ch	The unique ID of a Change-record in the Service Management tool.	Key - 18000 distinct Change ID's
ActualStartCh	Date and time the change implementation is actually started.	10/7/2013 5:00:00 AM - 3/23/2014 9:32:00 PM
ServiceComponentWBS(aff)Ch	Every CI in the CMDB is related to 1 Service Component, in order to identify which Product Manager is responsible for the CI. A Service Component is equal to a product in the Bill of Material and is part of one or more Services.	Discrete- 284 values
DayOfChange	Number to indicate the day of the week the implementation of the change based on the "Actual Start Ch" date of the Change (ID). Values assigned as follows: Monday = 1, Tuesday = 2,…Sunday = 7.	Discrete – 7 values
AffectedCITypes	Distinct number of "CI Type (aff) Ch" values related to the Change (ID)	1 - 3
AffectedCISubTypes	Distinct number of "CI SubType (aff) Ch" values related to the Change (ID)	1 - 4
AffectedCIs	Distinct number of "CI Name (aff) Ch" values related to the Change (ID)	1 - 64
RiskAssessmentCh	Impact of change: Major Business Change, Business Change or Minor Change.	Discrete – 3 values
EmergencyChangeCh	Indication if the change is an emergency fix. (Y or N)	Discrete - 0/1
CAB-approvalNeededCh	Indication is the changes needs approval by the Change Advisory Board, before implementation. (Y or N)	Discrete - Y/N
OriginatedFromCh	Indication if the change originated from for instance Problem research, or is a quick fix for an Incident.	Discrete – 3 values
ScheduledDownTimeHrs	Sheduled downtime (in hours) of the Change (ID) calculated as follows: "Scheduled Downtime End Ch" – "Scheduled Downtime Start Ch"	0-3192
PlannedStartDelayHrs	Delay (in hours) of the "Actual Start" of the Change (ID) calculated as follows: "Actual Start Ch" – "Planned Start Ch"	-8757 - 17570
RequestedEndDateDelayHrs	Delay (in hours) of the "Actual Start" of the Change (ID) calculated as follows: "Actual Start Ch" – "Requested End Date Ch"	-8757 - 17570
PlannedTimeHrs	Planned Time (in hours) of the Change (ID) calculated as follows: "Planned End Ch" – "Planned Start Ch"	0 - 13162

ChaMajorBusinessChange	Number of “Major Business Changes” (Risk Assessment Ch = Major Business Change) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaBusinessChange	Number of “Business Changes” (Risk Assessment Ch = Business Change) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaMinorChange	Number of “Minor Changes” (Risk Assessment Ch = Minor Change) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaEmergencyChangeY	Number of “Emergency Changes” (Emergency Change Ch = Y) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaEmergencyChangeN	Number of “Non-Emergency Changes” (Emergency Change Ch = N) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaCABApprovY	Number of “Required CAB-Approval changes” (CAB-Approval needed Ch = Y) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaCABApprovN	Number of “Non-Required CAB-Approval changes” (CAB-Approval needed Ch = N) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaOriginatedFIncident	Number of “from incident originated changes” (Originated from Ch = Incident) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaOriginatedFProblem	Number of “from problem originated changes” (Originated from Ch = Problem) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaB7DaysAverage	Average per day of the total number of distinct Change Id’s that affected the exact same “Service component WBS (aff)	0 - 26

	" of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	
ChaB14DaysAverage	Average per day of the total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 14 days (336 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 24
ChaB21DaysAverage	Average per day of the total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 21 days (504 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 22
ChaBday1	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the day (24 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 54
ChaBday2	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the second day (24-48 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 47
ChaBday3	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the third day (48-72 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 55
ChaBday4	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the fourth day (72-96 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 55
ChaBday5	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the fifth day (96-120 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 54
ChaBday6	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the sixth day (120-144 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 54
ChaBday7	The total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the seventh day (144-168h hours) before the "Actual Start Ch" time of the Change (ID)	0 - 54
ChaBD1Av21Diff	Difference between the "ChaBday1" field and the "ChaB21DaysAverage" field to calculate the difference between the total number of distinct Change Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the day (24 hours) before the "Actual Start Ch" time of the Change (ID) and the Average per day of the total number of distinct Change Id's that affected the	-21 - 34

	exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID)	
ChaBD2Av21Diff	Difference between the “ChaBday2” field and the “ChaB21DaysAverage” field, calculation analogue to the calculation method as explained for field “ChaBD1Av21Diff”.	-19 - 32
ChaBD3Av21Diff	Difference between the “ChaBday3” field and the “ChaB21DaysAverage” field, calculation analogue to the calculation method as explained for field “ChaBD1Av21Diff”.	-19 - 34
ChaBD4Av21Diff	Difference between the “ChaBday4” field and the “ChaB21DaysAverage” field, calculation analogue to the calculation method as explained for field “ChaBD1Av21Diff”.	-19 - 33
ChaBD5Av21Diff	Difference between the “ChaBday5” field and the “ChaB21DaysAverage” field, calculation analogue to the calculation method as explained for field “ChaBD1Av21Diff”.	-20 - 32
ChaBD6Av21Diff	Difference between the “ChaBday6” field and the “ChaB21DaysAverage” field, calculation analogue to the calculation method as explained for field “ChaBD1Av21Diff”.	-20 - 32
ChaBD7Av21Diff	Difference between the “ChaBday7” field and the “ChaB21DaysAverage” field, calculation analogue to the calculation method as explained for field “ChaBD1Av21Diff”.	-20 - 32
IncImpact3	Number of “ Impact 3 Incidents” (Impact Inc =3) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 30
IncImpact4	Number of “ Impact 4 Incidents” (Impact Inc =4) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 726
IncImpact5	Number of “ Impact 5 Incidents” (Impact Inc =5) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 127
IncUrgency3	Number of “ Urgency 3 Incidents” (Urgency Inc =3) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 30
IncUrgency4	Number of “ Urgency 4 Incidents” (Urgency Inc =4) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 722
IncUrgency5	Number of “ Urgency 5 Incidents” (Urgency Inc =5) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 127

IncPriority3	Number of “Priority 3 Incidents” (Priority Inc =3) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 30
IncPriority4	Number of “Priority 4 Incidents” (Priority Inc =4) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 722
IncPriority5	Number of “Priority 5 Incidents” (Priority Inc =5) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 127
IncCatIncident	Number of “Incident Category Incidents ” (Category Inc = Incident) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 815
IncCatRFI	Number of “Request for information Incidents ” (Category Inc = Request for Information) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 127
IncCCOther	Number of “Other Closure Code Incidents ” (Closure Code Inc = Other) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 268
IncCCSoftware	Number of “Software Closure Code Incidents ” (Closure Code Inc = Software) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 367
IncB7Days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 830
IncB7DaysHandleTimeHrs	Sum of the total handle time of all distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 11471
AverageIncHandleTime	The sum of the total handle time of all distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID) divided by the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during	0 - 621

	the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID).” IncB7DaysHandleTimeHrs”/ “IncB7Days”	
IncA7DaysAverage	Average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) after the “Actual Start Ch” time of the Change (ID)	0 - 119
IncA14DaysAverage	Average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) after the “Actual Start Ch” time of the Change (ID)	0 - 117
IncB14DaysAverage	Average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 117
IncB21DaysAverage	Average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 109
IncA21DaysAverage	Average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) after the “Actual Start Ch” time of the Change (ID)	0 - 109
IncA1day	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 229
IncA2days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 2 days (48 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 372
IncA3days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 3 days (72 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 529
IncA4days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 4 days (96 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 688
IncA5days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 5 days (120 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 817
IncA6days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change	0 - 834

	(ID) during the 6 days (144 hours) following the “Actual Start Ch” time of the Change (ID)	
IncA7days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 834
IncA14days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 1642
IncA21days	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 2290
IncBday1	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 221
IncBday2	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the second day (24-48 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 195
IncBday3	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the third day (48-72 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 244
IncBday4	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the fourth day (72-96 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 244
IncBday5	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the fifth day (96-120 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 226
IncBday6	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the sixth day (120-144 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 188
IncBday7	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the seventh day (144-168h hours) before the “Actual Start Ch” time of the Change (ID)	0 - 246
IncAday1	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change	0 - 229

	(ID) during the day (24 hours) after the “Actual Start Ch” time of the Change (ID)	
IncAday2	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the second day (24-48 hours) after the “Actual Start Ch” time of the Change (ID)	0 - 244
IncAday3	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the third day (48-72 hours) after the “Actual Start Ch” time of the Change (ID)	0 - 186
IncAday4	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the fourth day (72-96 hours) after the “Actual Start Ch” time of the Change (ID)	0 - 247
IncAday5	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the fifth day (96-120 hours) after the “Actual Start Ch” time of the Change (ID)	0 - 193
IncAday6	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the sixth day (120-144 hours) after the “Actual Start Ch” time of the Change (ID)	0 - 196
IncAday7	The total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the seventh day (144-168h hours) after the “Actual Start Ch” time of the Change (ID)	0 - 217
IncBD1Av21Diff	Difference between the “IncBday1” field and the “IncB21DaysAverage” field to calculate the difference between the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) before the “Actual Start Ch” time of the Change (ID) and the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID).	-109 - 140
IncBD2Av21Diff	Difference between the “IncBday2” field and the IncB21DaysAverage field, calculation analogue to the calculation method as explained for field “IncBD1Av21Diff”.	-109 - 111
IncBD3Av21Diff	Difference between the “IncBday3” field and the “IncB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncBD1Av21Dif”.	-108 - 157
IncBD4Av21Diff	Difference between the “IncBday4” field and the “IncB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncBD1Av21Diff”.	-107 - 153

IncBD5Av21Diff	Difference between the “IncBday5” field and the “IncB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncBD1Av21Diff”.	-105 - 134
IncBD6Av21Diff	Difference between the “IncBday6” field and the” IncB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncBD1Av21Diff”.	-109 - 111
IncBD7Av21Diff	Difference between the “IncBday7” field and the “IncB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncBD1Av21Diff”.	-101 - 155
IncAD1Av21Diff	Difference between the “IncAday1” field and the “IncA21DaysAverage” field to calculate the difference between the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) after the “Actual Start Ch” time of the Change (ID) and the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) after the “Actual Start Ch” time of the Change (ID).	-104 - 125
IncAD2Av21Diff	Difference between the “IncAday2” field and the “IncA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncAD1Av21Diff”.	-105 - 140
IncAD3Av21Diff	Difference between the “IncAday3” field and the “IncA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncAD1Av21Dif”.	-105 - 94
IncAD4Av21Diff	Difference between the “IncAday4” field and the “IncA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncAD1Av21Diff”.	-101 - 141
IncAD5Av21Diff	Difference between the “IncAday5” field and the “IncA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncAD1Av21Diff”.	-107 - 100
IncAD6Av21Diff	Difference between the “IncAday6” field and the” IncA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncAD1Av21Diff”.	-108 - 103
IncAD7Av21Diff	Difference between the “IncAday7” field and the “IncA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IncAD1Av21Diff”.	-109 - 109
IntImpact3	Number of “ Impact 3 Interactions” (Impact Int =3) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 227
IntImpact4	Number of “ Impact 4 Interactions” (Impact Int =4) that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 1727

IntImpact5	Number of " Impact 5 Interactions" (Impact Int =5) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 338
IntUrgency3	Number of " Urgency 3 Interactions" (Urgency Int =3) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 216
IntUrgency4	Number of " Urgency 4 Interactions" (Urgency Int =4) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 1720
IntUrgency5	Number of " Urgency 5 Interactions" (Urgency Int =5) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 337
IntPriority3	Number of "Priority 3 Interactions" (Priority Int =3) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 216
IntPriority4	Number of "Priority 4 Interactions" (Priority Int =4) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 1720
IntPriority5	Number of "Priority 5 Interactions" (Priority Int =5) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 337
IntCatIncident	Number of "Incident Category Interactions " (Category Int = Incident) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 1939
IntCatRFI	Number of "Request for information Interactions " (Category Int = Request for Information) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 291
IntCCOther	Number of "Other Closure Code Interactions " (Closure Code Int = Other) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168 Hours) before the "Actual Start Ch" time of the Change (ID)	0 - 644
IntCCSoftware	Number of "Software Closure Code Interactions " (Closure Code Int = Software) that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days	0 - 843

	(168 Hours) before the “Actual Start Ch” time of the Change (ID)	
IntB7Days	The total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 1988
IntB7DaysHandleTimeHrs	Sum of the total handle time of all distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 240
AverageIntHandleTime	The sum of the total handle time of all distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID) divided by the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID).” IntB7DaysHandleTimeHrs”/ “IntB7Days”	0-2
IntA7DaysAverage	Average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) after the “Actual Start Ch” time of the Change (ID)	0 - 284
IntB14DaysAverage	Average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) after the “Actual Start Ch” time of the Change (ID)	0 - 267
IntA14DaysAverage	Average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 267
IntB21DaysAverage	Average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID)	0 - 249
IntA21DaysAverage	Average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) after the “Actual Start Ch” time of the Change (ID)	0 - 249
IntA1day	The total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 543
IntA2days	The total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the	0 - 974

	Change (ID) during the 2 days (48 hours) following the "Actual Start Ch" time of the Change (ID)	
IntA3days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 3 days (72 hours) following the "Actual Start Ch" time of the Change (ID)	0 - 1336
IntA4days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 4 days (96 hours) following the "Actual Start Ch" time of the Change (ID)	0 - 1691
IntA5days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 5 days (120 hours) following the "Actual Start Ch" time of the Change (ID)	0 - 1961
IntA6days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 6 days (144 hours) following the "Actual Start Ch" time of the Change (ID)	0 - 1994
IntA7days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 7 days (168h hours) following the "Actual Start Ch" time of the Change (ID)	0 - 1994
IntA14days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 14 days (336 hours) following the "Actual Start Ch" time of the Change (ID)	0 - 3744
IntA21days	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the 21 days (504 hours) following the "Actual Start Ch" time of the Change (ID)	0 - 5233
IntBday1	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the day (24 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 506
IntBday2	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the second day (24-48 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 497
IntBday3	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the third day (48-72 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 543
IntBday4	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the	0 - 539

	Change (ID) during the fourth day (72-96 hours) before the "Actual Start Ch" time of the Change (ID)	
IntBday5	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the fifth day (96-120 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 471
IntBday6	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the sixth day (120-144 hours) before the "Actual Start Ch" time of the Change (ID)	0 - 437
IntBday7	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the seventh day (144-168h hours) before the "Actual Start Ch" time of the Change (ID)	0 - 549
IntAday1	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the day (24 hours) after the "Actual Start Ch" time of the Change (ID)	0 - 543
IntAday2	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the second day (24-48 hours) after the "Actual Start Ch" time of the Change (ID)	0 - 543
IntAday3	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the third day (48-72 hours) after the "Actual Start Ch" time of the Change (ID)	0 - 445
IntAday4	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the fourth day (72-96 hours) after the "Actual Start Ch" time of the Change (ID)	0 - 549
IntAday5	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the fifth day (96-120 hours) after the "Actual Start Ch" time of the Change (ID)	0 - 542
IntAday6	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the sixth day (120-144 hours) after the "Actual Start Ch" time of the Change (ID)	0 - 423
IntAday7	The total number of distinct Interaction Id's that affected the exact same "Service component WBS (aff) " of the Change (ID) during the seventh day (144-168h hours) after the "Actual Start Ch" time of the Change (ID)	0 - 564
IntBD1Av21Diff	Difference between the "IntBday1" field and the "IntB21DaysAverage" field to calculate the difference between the total number of distinct Incident Id's that	-248 - 309

	affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) before the “Actual Start Ch” time of the Change (ID) and the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID).	
IntBD2Av21Diff	Difference between the “IntBday2” field and the “IntB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntBD1Av21Diff”.	-249 - 290
IntBD3Av21Diff	Difference between the “IntBday3” field and the “IntB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntBD1Av21Dif”.	-249 - 335
IntBD4Av21Diff	Difference between the “IntBday4” field and the “IntB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntBD1Av21Diff”.	-247 - 324
IntBD5Av21Diff	Difference between the “IntBday5” field and the “IntB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntBD1Av21Diff”.	-246 - 255
IntBD6Av21Diff	Difference between the “IntBday6” field and the” IntB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntBD1Av21Diff”.	-248 - 240
IntBD7Av21Diff	Difference between the “IntBday7” field and the “IntB21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntBD1Av21Diff”.	-235 - 333
IntAD1Av21Diff	Difference between the “IntAday1” field and the “IntA21DaysAverage” field to calculate the difference between the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the day (24 hours) after the “Actual Start Ch” time of the Change (ID) and the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) after the “Actual Start Ch” time of the Change (ID).	-244 - 304
IntAD2Av21Diff	Difference between the “IntAday2” field and the IntA21DaysAverage field, calculation analogue to the calculation method as explained for field “IntAD1Av21Diff”.	-244 - 299
IntAD3Av21Diff	Difference between the “IntAday3” field and the “IntA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntAD1Av21Dif”.	-244 - 201
IntAD4Av21Diff	Difference between the “IntAday4” field and the “IntA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntAD1Av21Diff”.	-235 - 303

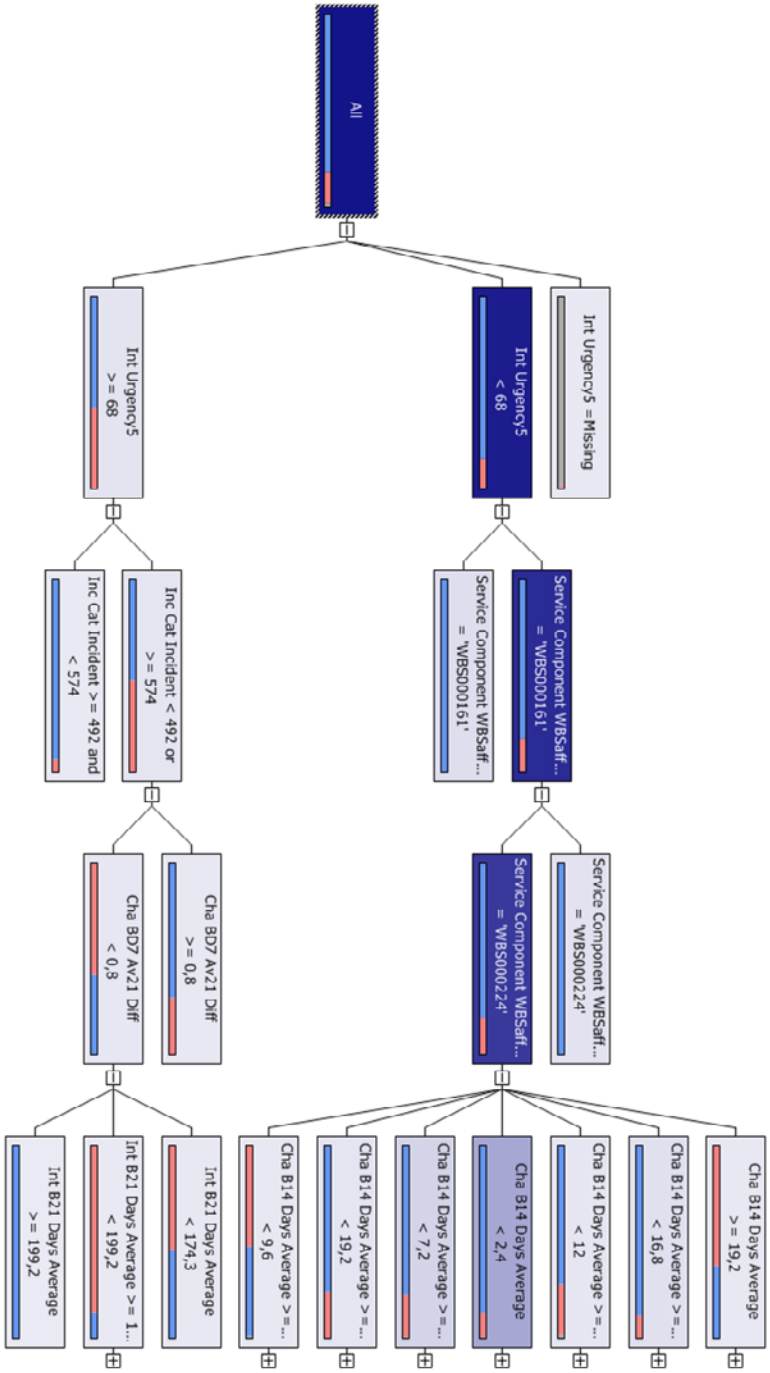
IntAD5Av21Diff	Difference between the “IntAday5” field and the “IntA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntAD1Av21Diff”.	-247 - 296
IntAD6Av21Diff	Difference between the “IntAday6” field and the” IntA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntAD1Av21Diff”.	-246 - 216
IntAD7Av21Diff	Difference between the “IntAday7” field and the “IntA21DaysAverage” field, calculation analogue to the calculation method as explained for field “IntAD1Av21Diff”.	-248 - 315
IntImpact7Days	Indicates if the average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IntImpact7Days” indicates 1.	Discrete: 1/0
IntRealImpact7Days	Difference between the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168h hours) following the “Actual Start Ch” time of the Change (ID) and the the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID) to indicate the level of increase/decrease. “IntA7Days”-“IntB7Days”	-422 - 625
IntImpact14Days	Indicates if the average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IntImpact14Days” indicates 1.	Discrete: 1/0
IntImpact21Days	Indicates if the average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of distinct Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change	Discrete: 1/0

	(ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IntImpact21Days” indicates 1.	
IncImpact7Days	Indicates if the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IncImpact7Days” indicates 1.	Discrete: 1/0
IncRealImpact7Days	Difference between the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168h hours) following the “Actual Start Ch” time of the Change (ID) and the the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID) to indicate the level of increase/decrease. “IncA7Days”-“IncB7Days”	-273 - 305
IncImpact14Days	Indicates if the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 14 days (336 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IncImpact14Days” indicates 1.	Discrete: 1/0
IncImpact21Days	Indicates if the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 21 days (504 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IncImpact21Days” indicates 1.	Discrete: 1/0
CombinedImpact7Days	Indicates if the overall number of distinct Incident and Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days	Discrete: 1/0

	(168h hours) following the “Actual Start Ch” time of the Change (ID) is larger than the total number of distinct Incident and Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “combinedImpact7Days” indicates 1.	
RealCombinedImpact7Days	Sum of the ”IntRealImpact7Days” and the “IncRealImpact7Days” fields to indicate the level of increase/decrease in the overall number of distinct Incident and Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168h hours) following the “Actual Start Ch” time of the Change (ID) and the total number of distinct Incident and Interaction Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID).	-693 - 930
IncActA7Days	The total number of Incident Activities related to distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) following the “Actual Start Ch” time of the Change (ID)	0 - 7103
IncActB7Days	The total number of Incident Activities related to distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID)	0 - 7078
IncActImpact7Days	Indicates if the average per day of the total number of Incident Activities related to distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) after the “Actual Start Ch” time of the Change (ID) is larger than the average per day of the total number of Incident Activities related to distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 Hours) before the “Actual Start Ch” time of the Change (ID). If the difference is larger than 0 , the “IncActImpact7Days” indicates 1.	Discrete: 1/0
IncActRealImpact7Days	Difference between the total number Incident activities related to distinct Incident Id’s that affected the exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168h hours) following the “Actual Start Ch” time of the Change (ID) and the the total number of Incident activities related to distinct Incident Id’s that affected the	-1948 - 1879

	exact same “Service component WBS (aff) “ of the Change (ID) during the 7 days (168 hours) before the “Actual Start Ch” time of the Change (ID) to indicate the level of increase/decrease. “IncActA7Days”-“IncActB7Days”	

- **Appendix A.8 – SSAS Decision Tree Model for interactions impact (Until level5)**



Appendix A.9 – Detailed findings on data quality & profiling for the "detail interaction table"

In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

Data field	Main findings	DM	PM
CI Name (aff)	For 16 of the 147004 interactions, the 'CI Name (aff)' field contains value '#N/B'. The most interactions occur on configuration item SAN000182 (7,26%). This configuration item belongs to type storage.		X
CI Type (aff)	For 68% of the interactions, the configuration item type where a disruption of an ICT Service is noticed, is the application. 14% of the interactions correspond to the sub-application, 8% to storage and 6% to computer. Furthermore, there are 777 interactions for which the "CI Type (aff)" field contains value 'no type'.		X
CI Subtype (aff)	For 33% of the interactions, the configuration item subtype where a disruption of an ICT Service is noticed, is the Server Based Application. 27% of the interactions correspond to the Web Based Application, 16% to Desktop Application and 8% to SAN. These subtypes belong to configuration item type application or sub-application, except for SAN, which belongs to type storage. 777 interactions are related to the CI Subtype (aff) 'no type'. These correspond with those interactions that also have a CI Type (aff) of 'no type'.		X
Service Comp WBS (aff)	Most interactions are related to service component WBS000073 (23%). The second most affected service component is WBS000128 (10%).	X	
Interaction ID	The Interaction ID is the unique identifier for data table Detail Interaction. Note that the interaction IDs do not increase incrementally.		
Status	Only 6 of the 147004 interactions are not closed, and have the status 'Open - Linked'. Furthermore, all the 'linked' interactions, could not be solved in first line, and are thus linked to an already existing incident.		
Impact	Only 2 interactions have a very high impact. They are both related to the Server Based Application CI subtype (aff), and were not solved in first line. Approximately 88% have a low or very low impact.	X	

Urgency	Only 32 interactions have a very high urgency and approximately 88% have a low or very low urgency.	X	
Priority	Only 2 interactions have a very high priority and approximately 88% have a low or very low priority. The 2 interactions with very high priority are the same interactions that are categorized as very high impact.	X	
Category	Approximately 79% of the interactions correspond to the 'incident' category, and 31% to the 'request for information' category.	X	
KM number	The most used Knowledge document is KM0002125 (4%).		
Open Time (First Touch)	All interactions are opened between 9/09/2011 and 31/03/2014. Most interactions are opened in 2013 (50%) and 2014 (49%). A very limited number is opened in 2011 and 2012.	X	X
Close Time	All interactions are closed between 1/10/2013 and 31/03/2014.	X	X
Closure Code	Approximately 40% of the interactions has an unknown disruption type (i.e. Other, Unknown, NULL, Overig, unknown). The most used closure code is 'Software'.	X	
First Call Resolution	64% of the interactions could be solved in first call.	X	
Handle Time (secs)	The handle time to resolve an interaction lies between 0 and 22530 seconds (6,3 hours). On average it takes approximately 7,5 minutes to resolve an interaction. Of all interactions, 292 have a handle time of 0 seconds. For 75% of them immediately an incident was created.	X	
Related Incident	873 interactions (0,59%) were linked to multiple incidents, which is indicated with value '#MULTIVALUE'. Approximately 64% was not linked to an incident, indicating that they were solved in first call. This is in line with our finding concerning 'First_Call_Resolution'. The incident that is most linked is IM0000220.		

Appendix A.10 – Detailed findings on data quality & profiling for the "detail incident table"

In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

Data field	Main findings	DM	PM
CI Name (aff)	For all interactions the CI Name (aff) is properly filled in. We found no NULL or #N/B values. The most incidents occur on configuration item SUB000456 (6,5%). This configuration item belongs to type sub-application.		X
CI Type (aff)	For 71% of the incidents, the configuration item type where a disruption of an ICT Service is noticed, is application. 17 % of the incidents correspond to sub-application and 8% to computer. Furthermore, there are 111 incidents for which the "CI Type (aff)" field contains value 'no type'.		X
CI Subtype (aff)	For 40% of the incidents, the configuration item subtype where a disruption of an ICT Service is noticed, is Server Based Application. 33 % of the incidents correspond to Web Based Application and 8% to Desktop Application. These subtypes belong to the CI Type (aff) application or sub-application.		X
Service Comp WBS (aff)	Most interactions are related to service component WBS000073 (29%). The second most affected service component is WBS00091 (5%).	X	X
Incident ID	Incident ID is the unique identifier for this table. Note that the incident IDs do not increase incrementally.		
Status	Only 9 of the 46606 incidents are not closed, and have the status 'Work in Progress'. All the 'Work in Progress' incidents are related to affected CI subtype 'Banking device' and service component 'WBS000146'.		
Impact	Only 3 incidents have a very high impact, and approximately 84% have a low or very low impact. The 3 very high impact interactions are related to the Web Based Application CI subtype (twice), and the Server Based Application CI subtype (once).	X	
Urgency	Only 6 incidents have a very high urgency and approximately 84% have a low and very low impact.	X	

Priority	Only 3 incidents have a very high priority and approximately 84% have a low and very low priority. The 3 incidents with very high priority are the same interactions that are categorized as very high impact.	X	
Category	Approximately 81% of the incidents is categorized as 'incident' and 19% as 'request for information'. Note that there is also a very limited number of incidents (12) that is categorized as 'complaint' or 'request for change'.	X	
KM number	The most used Knowledge document is KM0001106 (2%).		
Alert Status	All incidents have alert status 'closed'.		
# Reassignments	59% of the incidents is solved without the need to reassign the ticket to another Operator. 28% of the incidents has switched from operator once or twice.		
Open Time	All incidents have been opened between 05/02/2012 and 31/03/2014. Most interactions opened in 2013 (52%) and 2014 (47%). Note that only 21 incidents were opened in 2012.	X	X
Reopen Time	In 95% of the cases the reopen time is set to null, meaning that they were not reopened. The other 5% were reopened in the period 10/05/2013 - 31/03/2014.		
Resolved Time	All incidents have been resolved between 1/10/2013 and 31/03/2014. Note that for 1780 incidents the resolved time was blank.		
Close Time	All incidents have been resolved between 1/10/2013 and 31/03/2014. As expected, this field did not contain any null values because it is a system generated field.	X	X
Handle Time (Hours)	For 12872 records (27,62%) this field is not filled in (null value).	X	
Closure Code	Approximately 40% of the incidents has an unknown disruption type (i.e. Other, Unknown, NULL, Overig). The most used closure code is 'Software'.	X	
# Related Interactions	92% of the incidents are linked to 1 interaction. Note that 3 incidents have more than 100 related interactions.		
Related Interaction	3434 incidents (7%) were linked to multiple interactions, which is indicated with value '#MULTIVALUE'. For 114 incidents (0,2%) the value for related interaction was set to '#N/B'.		
# Related Incidents	Only 3% of the incidents is related to another incident.		

# Related Changes	Only 1% of the incidents is related to a change.		
Related Change	24 incidents (0,05%) were linked to multiple changes, which is indicated with value '#MULTIVALUE'. For approximately 99% of the changes, this field had a null value, which corresponds to the field '# Related Changes'.		
CI Name (CBy)	For 2,4% of the incidents, the field was set to '#N/B'. The configuration item that caused most incidents is SUB000456, which is the same as the most affected configuration item.		
CI Type (CBy)	Approximately 63% of the incidents is caused by configuration item type application and 17% by the sub-application.		
CI Subtype (CBy)	The configuration item subtypes that caused the most incidents are Server Based Application (36%) and Web Based Application (32%).		
ServiceComp WBS (CBy)	The service component that caused the most incidents is WBS000073 (28%). In approximately 7% of the cases, the field was set to value '#N/B'.		

A.11 – Detailed findings on data quality & profiling for the "detail incident activity table"

In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

Data field	Main findings	DM	PM
Incident ID	The same incident ID appears multiple times, once for each status change of the incident.	X	
DateStamp	The DateStamp range lies between 07/01/2013 and 02/04/2014, with a more or less equal spread over the two years (2013: 45%, 2014: 55%)		
IncidentActivity Number	The IncidentActivity Number is the unique identifier for this table. Note that this ID does not increase incrementally.	X	
IncidentActivity Type	The top 3 Incident Activity Types that occur the most are the Assignment (19%), Operator Update (12%), and Reassignment (11%).		
Assignment Group	The most used Assignment Group is TEAM0008.(18%).		
KM number	The KM0001106 Knowledge Document is used the most (4%).		
Interaction ID	For 5643 Incident Activity Records, the Interaction ID is not filled out.		

A.12 **– Detailed findings on data quality & profiling for the "detail incident activity table"**

In addition, we flagged whether the field was used for data mining (DM) and/or process mining (PM) purposes.

Data field	Main findings	DM	PM
CI Name (aff)	For all changes the field CI Name (aff) is filled out. The most affected configuration item is NET000425 (2,5%). Note that also the second and third most affected configuration items are 'NET'-configuration items, NET000217 respectively NET000426. All three of these configuration items belong to configuration item type network components.	X	X
CI Type (aff)	The most affected configuration item types are computer (35%), application (20%) and network components (17%). For 171 changes there is no configuration item type defined ('no type').	X	X
CI Subtype (aff)	The most affected configuration item subtypes are Server Based Application (15%), Windows Server (13%) and Linux Server (9%). For 171 changes there is no configuration item subtype defined ('no subtype'). These changes correspond to the 171 changes with no CI Type (aff) mentioned above.	X	X
Service Comp WBS (aff)	Most changes are related to service component is WBS000102 (14%). The second most affected service component is WBS00161 (9%).	X	X
Change ID	Change ID is not a unique identifier for this table. This indicates that the same change may occur on different configuration items.		
Change Type	The 100 most common values are all Standard Change Types, Release Types or Standard Activity types. Except for one other value, Master change that is also in the list of most common values (used in 0,22% of the cases). We summarized all the change types into the following list; Change Component, Master Change, Master Change Roadmap, Release Type 01-21, Standard Activity Type 01-54, Standard Change Type 01-163.	X	X
Risk Assessment	Most of the changes are minor changes (94%). The remaining percentage concerns business changes and major business changes. Only 113 changes are considered major (0,37%).	X	X

Emergency Change	99,70% of the changes are normal changes, the remaining 0,3% are emergency changes.	X	X
CAB approval needed	Only for 6% of the changes a CAB approval was needed.	X	X
Planned Start	The changes were planned to start between 1/06/2011 and 20/02/2021. Note that 7 changes are planned to start in the future (2015, 2017 and 2021).	X	
Planned End	The changes were planned to end between 13/10/2011 and 20/02/2021. Note that the extraction filter was set on Change record close time for the period 01/10/2013-31/03/2014, indicating that there is a big difference with the planned end for some cases. Note that 17 of the changes were expected to end in the future (2015, 2017, 2021). On average the duration of a planned change is 155,83 hours.	X	
Scheduled Downtime Start	The changes were scheduled to start between 13/10/2011 and 20/02/2021.	X	
Scheduled Downtime End	The changes were scheduled to end between 23/06/2013 and 21/03/2015. The average scheduled downtime duration is 35 hours.	X	
Actual Start	The changes were actually started between 16/10/2012 and 21/03/2021. Note that the oldest planned start time is more than a year earlier than the oldest actual start time. The same is valid for scheduled downtime start. On average, the actual start is 49 hours later than it was planned.	X	X
Actual End	The actual end date lies between 26/10/2012 and 21/03/2021. On average the duration of an actual change is 46,19 hours, which is about 30% of the average planned duration. And on average, the actual end is 20,4 hours before the planned end. So although generally the implementation of a change starts later than planned, it also finishes earlier than planned.		X
Requested End Date	The requested end date lies between 31/03/2012 and 20/02/2028. On average the end date is planned 478 hours (+/- 20 days) earlier than what was requested.	X	
Change record Open Time	This field is a system generated field and therefore is more reliable than the other start date fields. Following this field, changes were opened		

	and assumed to be started between 1/09/2011 and 31/03/2014.		
Change record Close Time	This field is a system generated field and therefore is more reliable than the other end date fields. Following this field, changes were closed and assumed to be ended between 1/10/2013 and 31/02/2014. This period was configured for the extraction.		
Originated from	Most changes are originated from a problem (64%). The remaining changes were a quickfix for an incident (36%) or an interaction (0,01%).	X	X
# Related Interactions	In more than 99% of the cases this field is equal to NULL, indicating there are no related interactions.		
# Related Incidents	For 94% of the changes, this field is not filled in. This indicates that the Operator did not see any correlation with incidents that occurred after the change implementation.		