# Business Intelligence Development Project

## Closing Report

Lin Hellsten

13.12.2019

## Table of contents

# 1 Introduction

This report will introduce the Business Intelligence Development Project which developed the BI solutions for Cronus International Ltd according to the user stories (referring to the append). It will discuss the project solution through the following aspects:

- Data Structure and visualization
- Refresh Setting
- Project Delivery
- Conclusion

This report will only discuss the solutions for the 1st and 3rd user stories, which is Customer Analysis and Purchaser Workload analysis. The rest solutions will be discussed by Suvash in his report.

## 1.1 Update History

### 1.1.1 Iteration One

In solution of Customer Analysis, the user originally requested a solution for analysing the customer discount by sales amount. However, the user changed their requirement after some time that they would like to understand the Customers by RFM Model and the current discount.

### 1.1.2 Iteration Two

Base on the teacher's comment, the following update has been made, which is highlighted:

- Master data --- Customer table and Vendor table were imported into the solution
- Table naming practice has been improved. Please refer to Chapter 2.
- Added analysis for total purchase number per Vendor in visualization
- Validated the solution data with NAV. Please refer to Chapter 4.

# 2 Data Structure

In this session, the data structure will be introduced, including:

- Server, Database and Tables
- Data Cleansing
- Primary Key Determination
- Measure creation

## 2.1 Server, Database and Tables

| | |
|---|---|
| Server Name | *NAV2015* |
| Database Name | *BIG4TF023_Rehn_2022* |
| Company Name | *Cronus 01* |

### 2.1.1 Tables used in Customer RFM and Discount Analysis

| | |
|---|---|
| Table Name | *01 Cronus Int_W1$ Sales Header Merged (Refer to 2.1.3)* |
| | *01 Cronus Int_W1$ Sales Line Merged (Refer to 2.1.3)* |
| | *01 Cronus Int_W1$ Sales Line Discount* |
| | *01 Cronus Int_W1$ Exchange Currency Rate* |
| | *01 Cronus Int_W1$ Customer* |
| | *RFM Analysis* |

### 2.1.2 Tables used in Purchaser Workload Analysis

| | |
|---|---|
| Table Name | *01 Cronus Int_W1$ Purchase Header Merged (Refer to 2.1.3)* |
| | *01 Cronus Int_W1$ Purchase Line Merged (Refer to 2.1.3)* |
| | *01 Cronus Int_W1$ Salesperson_Purchaser* |
| | *01 Cronus Int_W1$ Vendor* |

### 2.1.3 Merged Tables

| Tables to be merged | Merged Table |
|---|---|
| *01 Cronus Int_W1$ Sales Header* | *01 Cronus Int_W1$ Sales Header Merged* |
| *01 Cronus Int_W1$ Sales Header Archive* | |
| *01 Cronus Int_W1$ Sales Line* | *01 Cronus Int_W1$ Sales Line Merged* |
| *01 Cronus Int_W1$ Sales Line Archive* | |
| *01 Cronus Int_W1$ Purchase Header* | *01 Cronus Int_W1$ Purchase Header Merged* |
| *01 Cronus Int_W1$ Purchase Header Archive* | |
| *01 Cronus Int_W1$ Purchase Line* | *01 Cronus Int_W1$ Purchase Line Merged* |
| *01 Cronus Int_W1$ Purchase Line Archive* | |

## 2.2 Data Cleansing

Since the data was extracted from Microsoft NAV, the data were already well structured. However, there were some blank columns which, per my research, were mis-opened orders and lines. Therefore, the blank columns were excluded during the Data Transformation phase.

## 2.3 Primary Key Determination

Since the data in this solution will include the archived data, the primary key will combine the fields of Document Type and Document Number with Version No. The detailed steps are as follow:

- Filter the rows in which version No equals to null or 1
- Replace null as 0 in fields of version No
- Combine Document Type, Document No and Version No as primary key

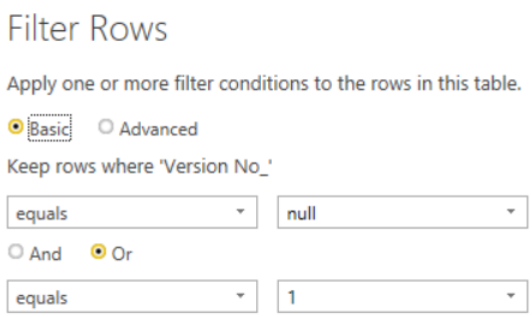



*Figure: Filter rows in which the Version No_ equals to null or 1*

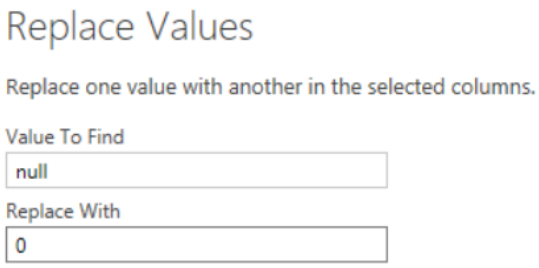



*Figure: Replace null as 0*

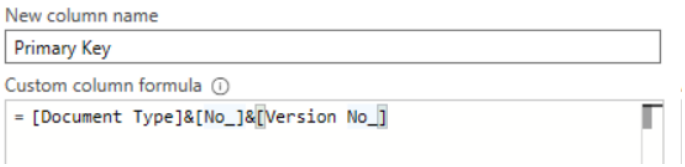



*Figure: Create the primary key*

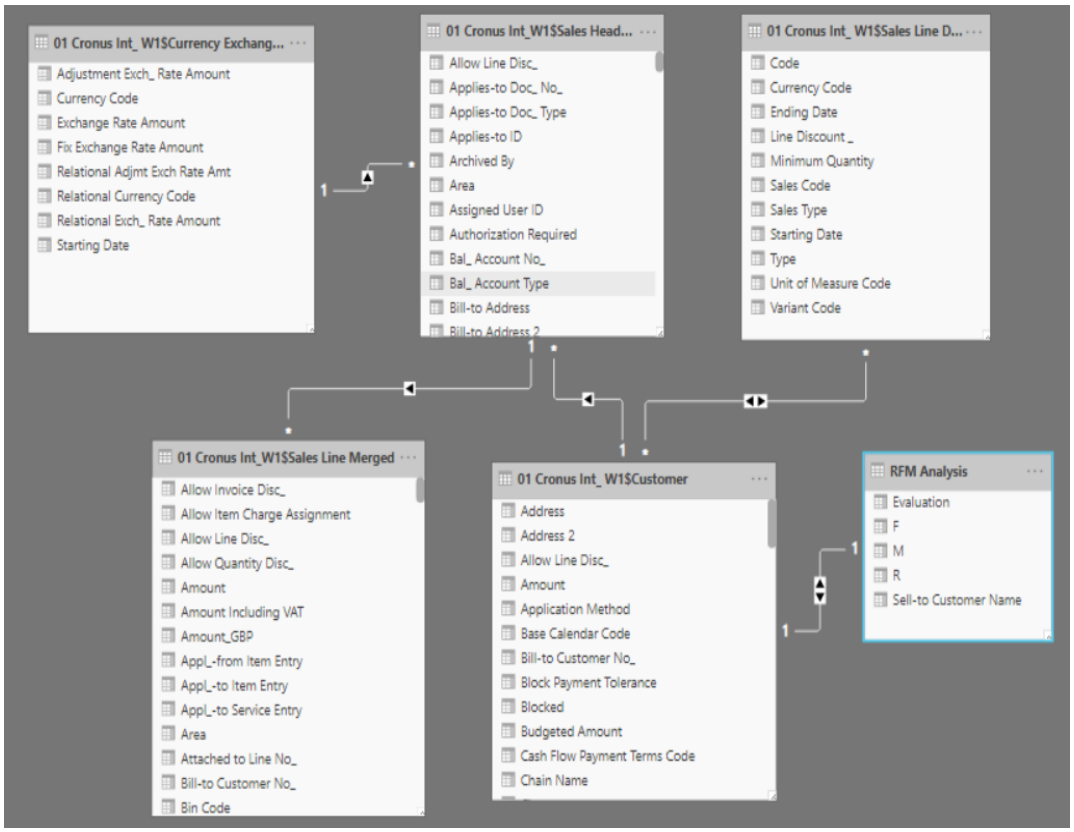



*Figure: Data model for Customer RFM analysis*

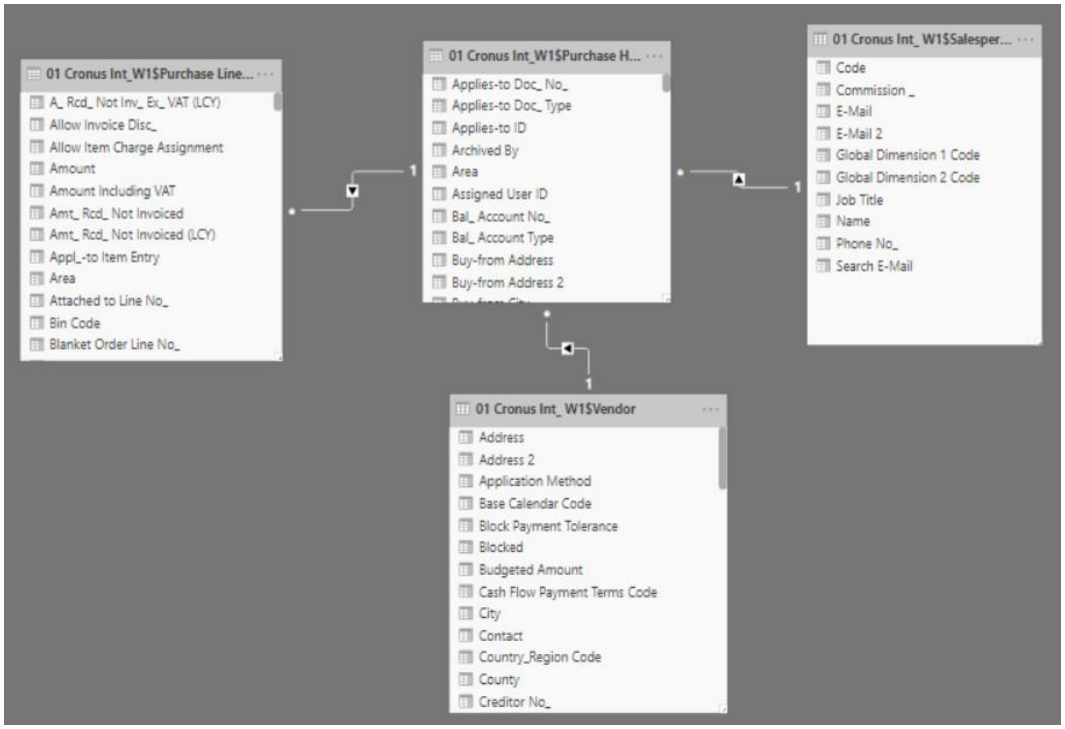



*Figure: Data model for Purchaser workload analysis*

***Remark:*** *To clarify the data model for each solution, additional pages were created base on the solutions.*

## 2.4 Measures

In this session, the measures and detailed steps will be presented. No measure needs to be created for Purchaser Workload Analysis. There are two new columns, three measures and one new table will be needed for Customer RFM Analysis, which are:

- The new column **"Time Period"** to calculate how many months since the sales order was created till current date:

  ***= DATEDIFF('RFM_Sales Header'[Document Date];TODAY();MONTH)***

- The measure **"R"** to calculate the minimum time period between current date and sales order created date.

  ***= MIN('RFM_Sales Header'[Time Period])***

- The measure **"F"** to calculate how many sales orders were placed

  ***= COUNTA('RFM_Sales Header'[Primary Key])***

- The measure **"M"** to calculate the total sales amount.

  ***= COUNTA('RFM_Sales Header'[Primary Key])***

- After the measures are created, a new table **"RFM Analysis"** will be created by using the following fuction:

  ***= SUMMARIZE('RFM_Sales Header';'RFM_Sales Header'[Sell-to Customer Name];"R";'RFM_Sales Header'[R];"F";[F];"M";[M])***

- A new column is created to group the customers by the result of “R”,”F”,”M”:

***= IF('RFM Analysis'[R]<AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]>AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M]> AVERAGE('RFM Analysis'[M]), "Best Customer",***

***IF('RFM Analysis'[R]<AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]>AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] < AVERAGE('RFM Analysis'[M]), "Potential Best Customer",***

***IF('RFM Analysis'[R]<AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]<AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] > AVERAGE('RFM Analysis'[M]), "Potential Best Customer",***

***IF('RFM Analysis'[R]<AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]<AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] < AVERAGE('RFM Analysis'[M]), "Others",***

***IF('RFM Analysis'[R]>AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]>AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] > AVERAGE('RFM Analysis'[M]), "Potential Customer",***

***IF('RFM Analysis'[R]>AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]>AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] < AVERAGE('RFM Analysis'[M]), "At risk",***

***IF('RFM Analysis'[R]>AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]<AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] > AVERAGE('RFM Analysis'[M]), "Potential Best Customer",***

***IF('RFM Analysis'[R]>AVERAGE('RFM Analysis'[R])&& 'RFM Analysis'[F]<AVERAGE('RFM Analysis'[F]) && 'RFM Analysis'[M] < AVERAGE('RFM Analysis'[M]), "At risk", "Others")))))))***

- To convert all sales amount into GBP. A new column is created by using below function:

***= if('RFM_Sales Line'[Currency Code]="",'RFM_Sales Line'[Amount],'RFM_Sales Line'[Amount]*RELATED('Currency Exchange Rate'[Relational Exch_ Rate Amount])/RELATED('Currency Exchange Rate'[Exchange Rate Amount]))***

### 2.5 Visualization

Please refer to the Power BI Desktop file or Power BI Service Report.

## 3 Refresh Settings

In Power BI Service, the Gateway is connected to NAV2015 of Haaga-Helia University of Applied Sciences. The Scheduled Refresh can be set according to user's requirement.

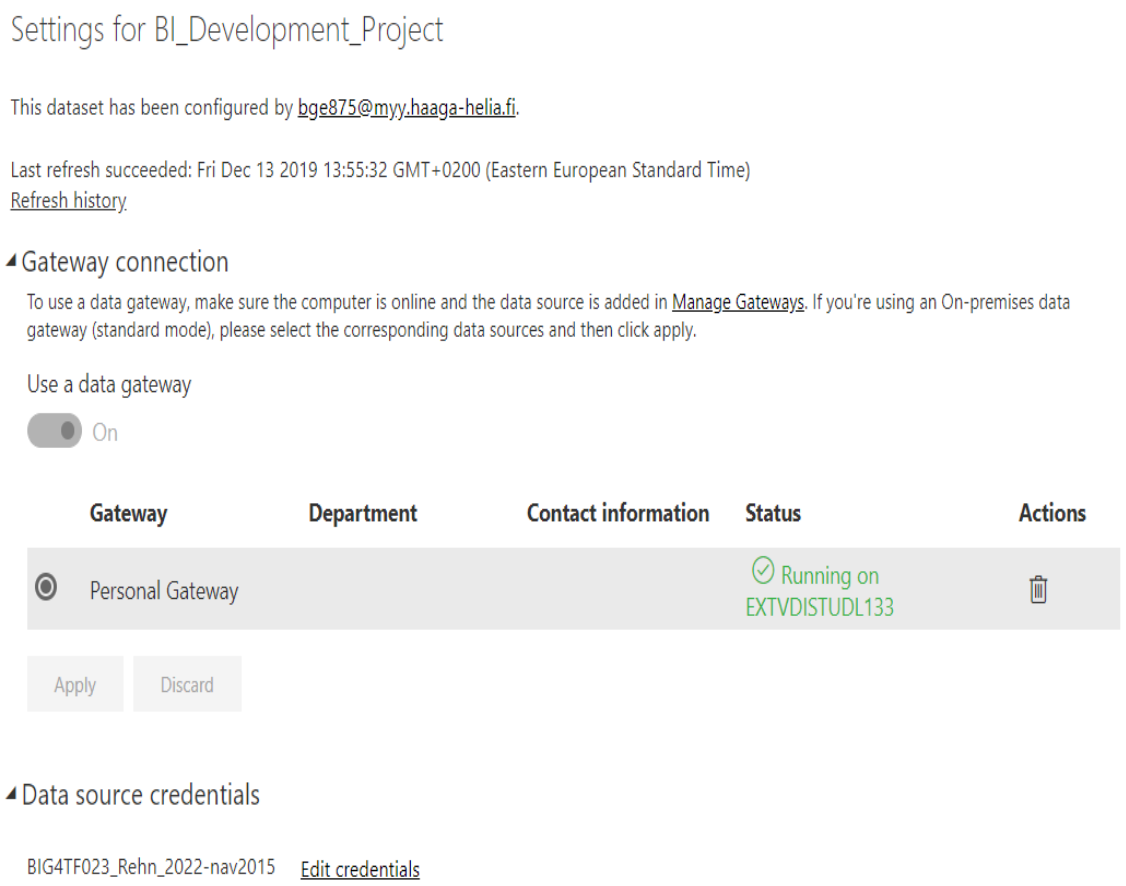



*Figure: Gateway connection*

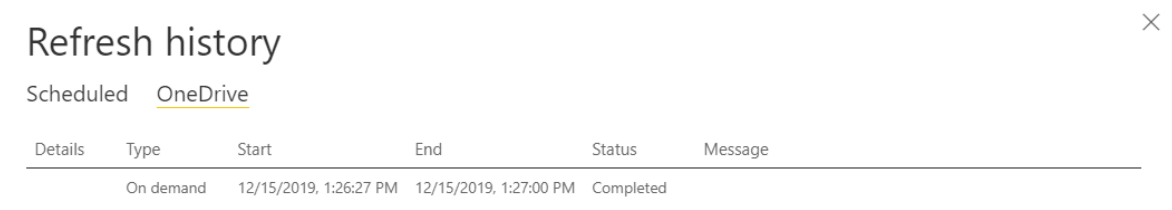



*Figure: Refresh tested*

# 4 Data Validation

In this part, the data in the solution will be validated with NAV.

## 4.1 RFM Analysis Validation

To validate the customer discount, the following steps were followed:

| | |
|---|---|
| Step One | Select Top 3 Maximum Line Discount with their Customer Name in the visual |
| Step Two | Select Line Discount equals to 100 and 50 with Document No and Customer No in Table '01 Cronus_Int_W1$ Sales Line Merged' |
| Step Three | Compare the data with NAV by Document No and Customer No |
| Step Four | Select one of the best customer 'London Light Company' as to validate the RMF calculation accuracy. |
| Step Five | Compare the RFM result with NAV |

**Customer Discount Analysis**

| Customer | Disc_Group | Line_Disc (Min) | Line_Disc (Max) | Lin_Disc(Avg) |
|---|---|---|---|---|
| Blanemark Hifi Shop | | 0,00 | 100,00 | 16,93 |
| Beef House | | 0,00 | 50,00 | 8,75 |
| Karoo Supermarkets | RETAIL | 10,00 | 50,00 | 16,00 |

*Figure: Top 3 Maximum Line Discount from Customer Discount Analysis*

| Document Type | Document No_ | Line No_ | Sell-to Customer No_ | Type | No_ | Location Code | Posting Group |
|---|---|---|---|---|---|---|---|
| 1 | 1077 | *20000* | 49525252 | *2* | 70200 | GREEN | RAW MAT |
| 0 | 1076 | *30000* | 60000 | *2* | LS-MAN-10 | WHITE | RESALE |
| 0 | 1075 | *40000* | 60000 | *2* | LS-MAN-10 | WHITE | RESALE |
| 1 | 1142 | *40000* | 60000 | *2* | LS-MAN-10 | WHITE | RESALE |
| 1 | 1141 | *30000* | 60000 | *2* | LS-MAN-10 | WHITE | RESALE |
| 1 | 1100 | *10000* | 27321782 | *2* | 1000 | YELLOW | FINISHED |
| 1 | 1098 | *20000* | 44756404 | *2* | 1700 | BLUE | FINISHED |

*Figure: Document No and Customer No with Line Discount equals to 100 and 50 in Table '01 Cronus_Int_W1$ Sales Line Merged'*

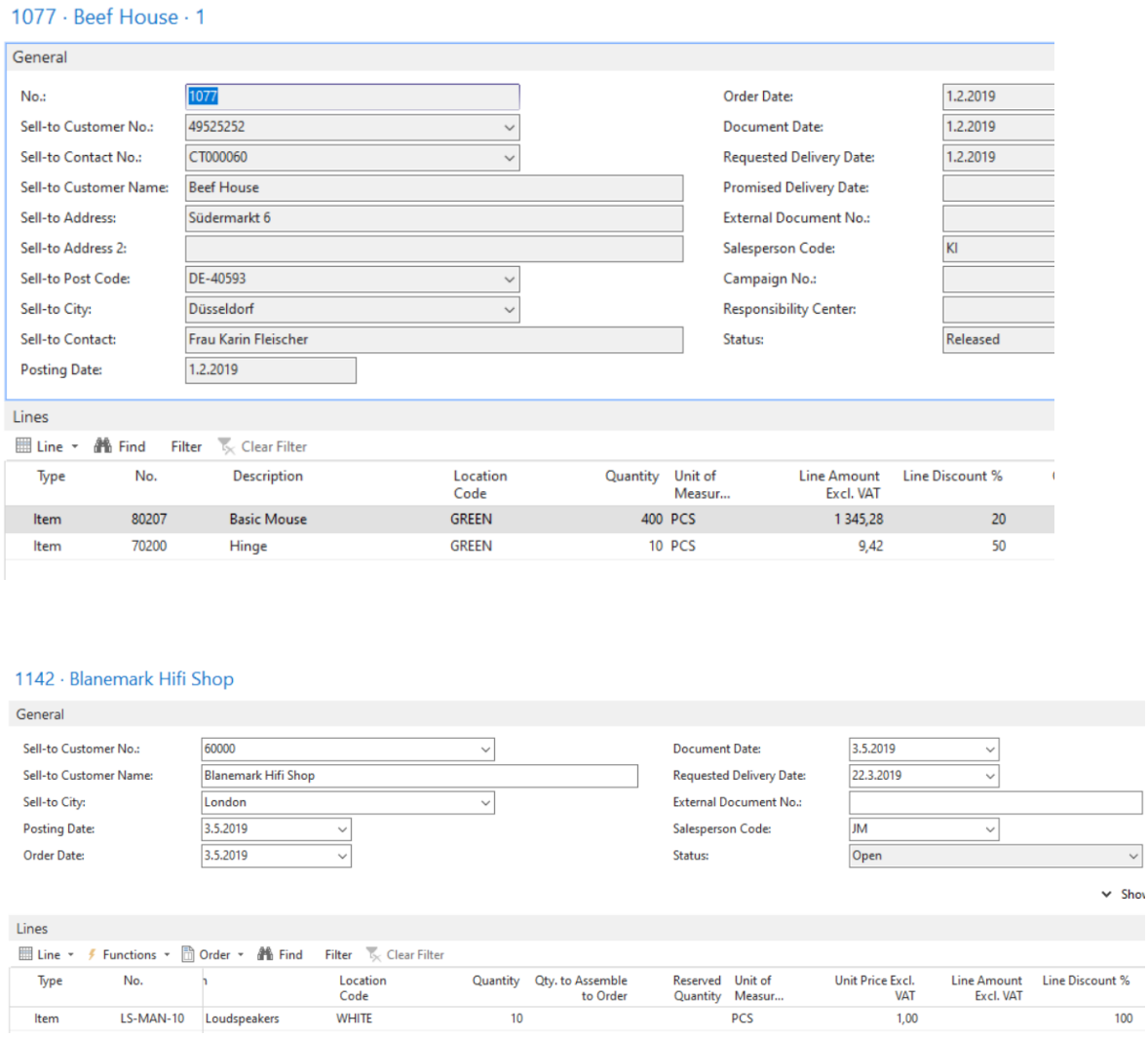
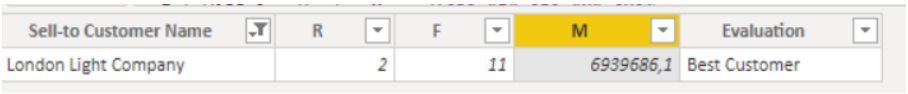



*Figure: Archived and non-archived data were examined in NAV*



| Sell-to Customer Name | R | F | M | Evaluation |
|---|---|---|---|---|
| London Light Company | 2 | 11 | 6939686,1 | Best Customer |

*Figure: RMF calculation result of 'London Light Company', which indicate the latest purchase date was two month ago.*

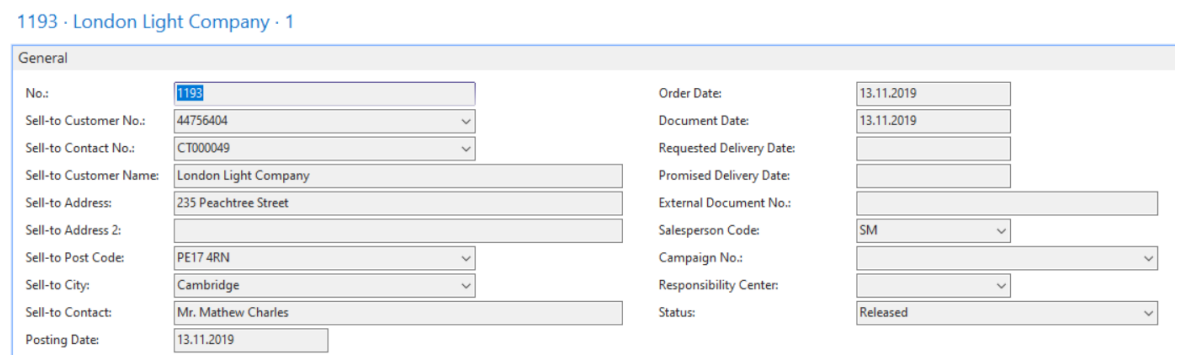



*Figure: NAV showed the latest document data of London Light Company was 13.11.2019, which is match with its RFM result.*

## 4.2 Nr of Purchase Validation

The steps to validate the Nr of purchase are as follow:

Step One Select one of the purchasers to examine the data

Step Two Compare the result with the data in NAV

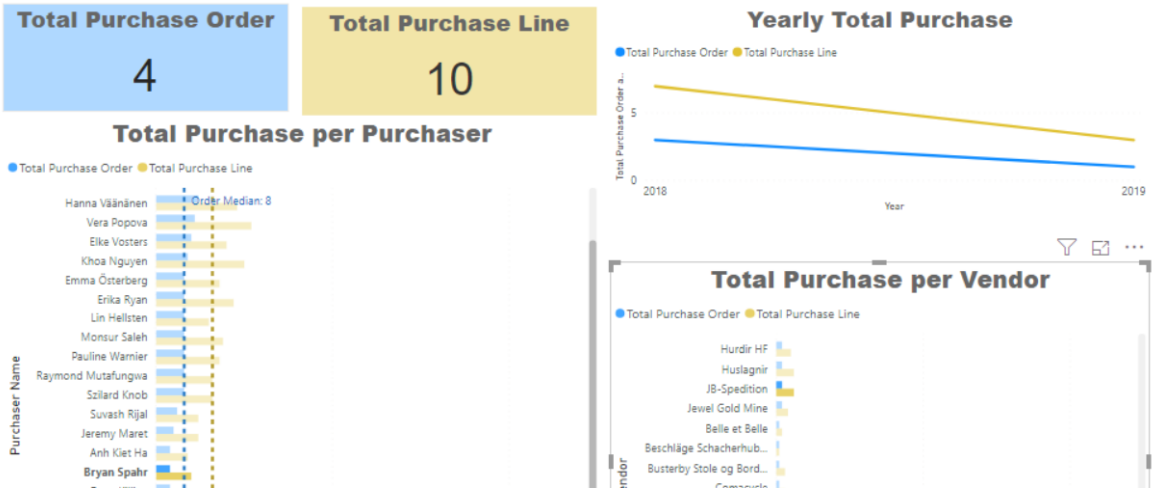
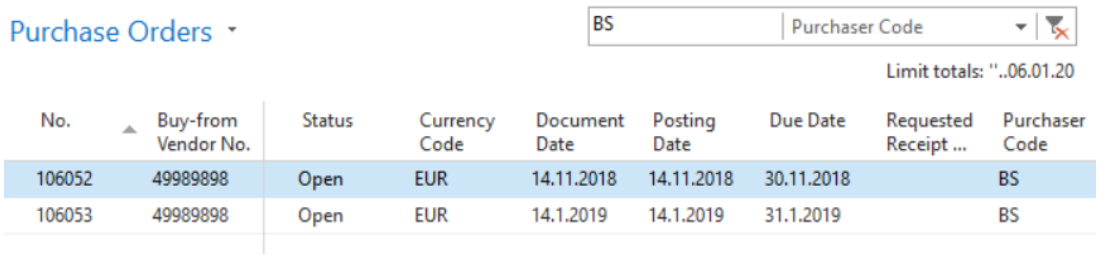


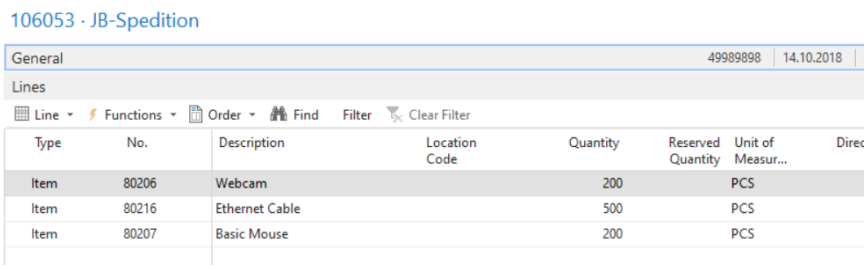

*Figure: The solution showed 4 purchase orders with 8 lines belong to Bryan Spahr. Orders and lines were created in 2018 and 2019. One of the Vendor was JB-Spedition.*




| No. | Buy-from Vendor No. | Status | Currency Code | Document Date | Posting Date | Due Date | Requested Receipt ... | Purchaser Code |
|---|---|---|---|---|---|---|---|---|
| 106052 | 49989898 | Open | EUR | 14.11.2018 | 14.11.2018 | 30.11.2018 | | BS |
| 106053 | 49989898 | Open | EUR | 14.1.2019 | 14.1.2019 | 31.1.2019 | | BS |

*Figure: 2 open purchase orders created by Bryan Spahr. The orders were created in 2018 and 2019.*

| | | | | |
|---|---|---|---|---|
| Item | 80002 | Computer III 600 MHz | 100 | PCS |
| Item | 80103 | 19" M009 Monitor | 100 | PCS |
| Item | 80211 | Quietkey Keyboard | 200 | PCS |




| Type | No. | Description | Location Code | Quantity | Reserved Quantity | Unit of Measur... | Direc |
|---|---|---|---|---|---|---|---|
| Item | 80206 | Webcam | | 200 | | PCS | |
| Item | 80216 | Ethernet Cable | | 500 | | PCS | |
| Item | 80207 | Basic Mouse | | 200 | | PCS | |

*Figures: Detailed purchase lines in each purchase order. There were 6 lines in total.*

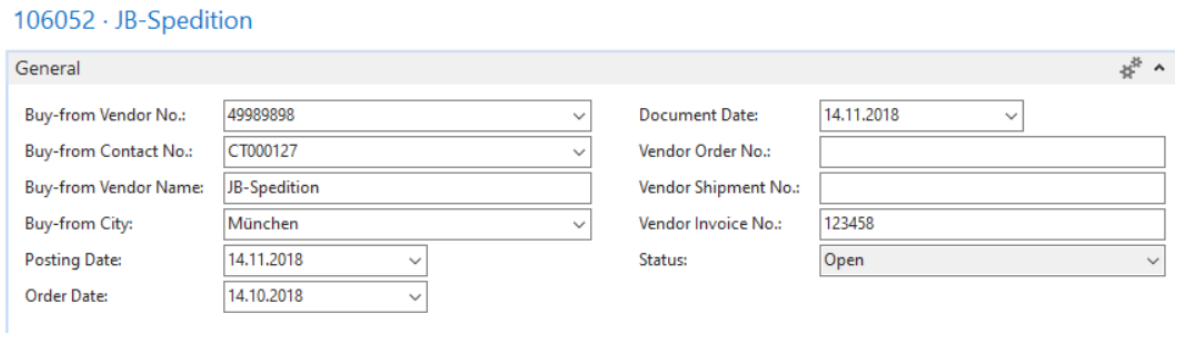
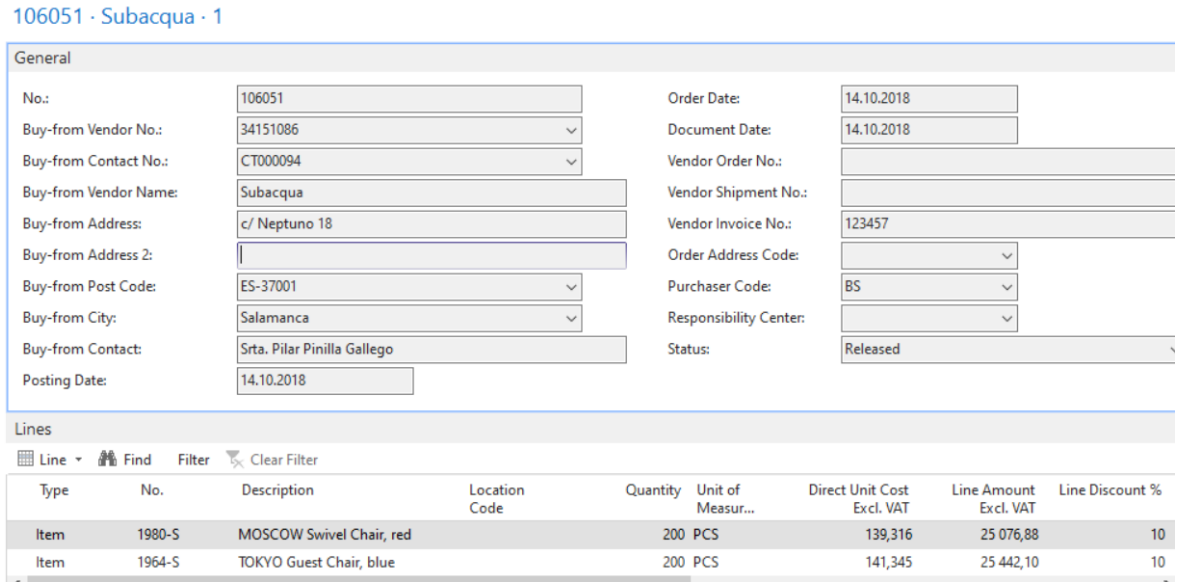
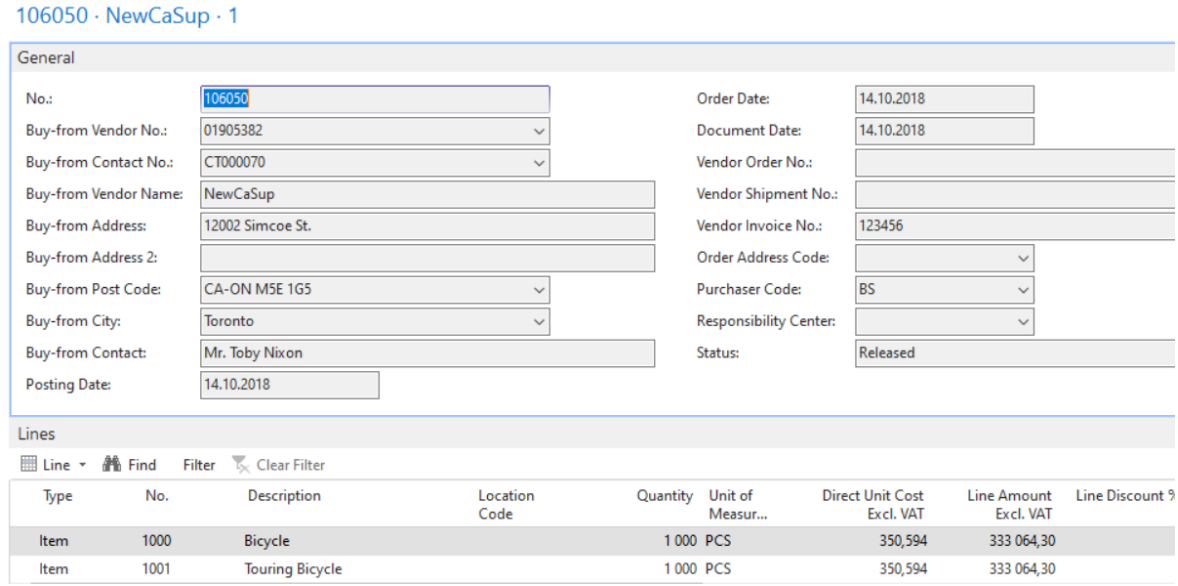



*Figure: The vendor of purchase order 106052 is JB-Spedition*




*Figures: 2 archived purchase order with detailed purchase lines created by Bryan Spahr. There were 4 purchases in total.*

# 5 Project Delivery

The two parts solutions of this project are combined and delivered to users through Power BI Cloud platform --Power BI Service Pro as a whole solution. Users can view the reports, add comments and edit the reports according to their granted right.

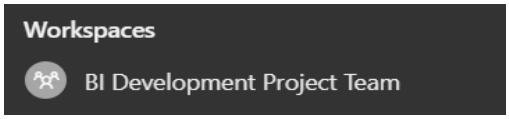



*Figure: A specific workspace was created for the project*

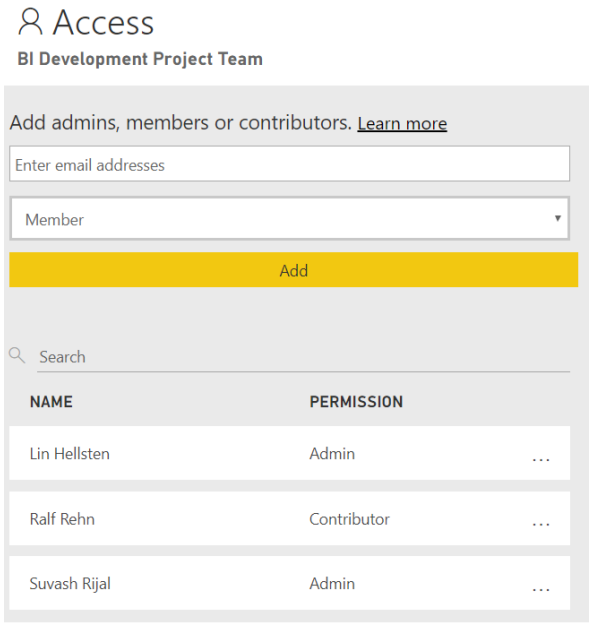



*Figure: Different right can be granted to team members and users*


*Figure: Access solutions through the specific workspace in Power BI Service*

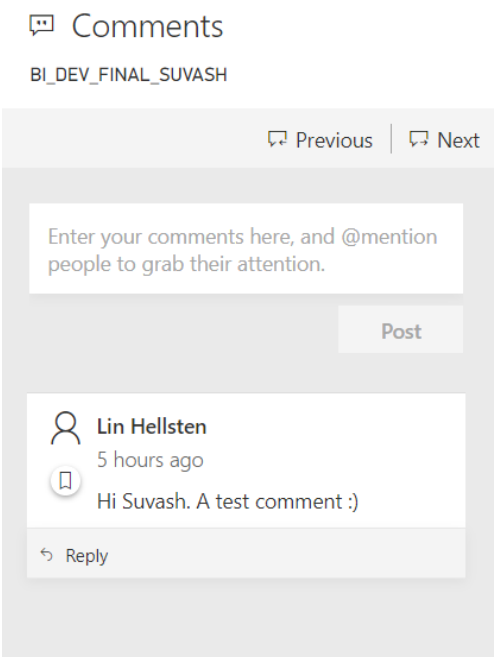



*Figure: Comments can be added through Power BI Service*

# 6 Conclusion

In this session, the achievement and challenge of team collaboration and individual work will be discussed.

## 6.1 Achievement

Through this project, we experienced the real-life environment on agile analytics development. At the project initiation phase, we had a team meeting to brainstorm, prioritize user stories, define the features, ect. After we defined the requirements and features, we started to develop own solutions respectively. However, according to Agile, we communicated and shared our issues, findings ect positively during the entire project development. For example, we shared the cleansed and transformed datasets for sales through Power BI Service and Suvash tested it in his own Power BI Desktop environment since he decided to use the same tables for his solutions, which improves team's productivity and reduce the repetitive work. Meanwhile we also discussed and tested a lot about the measures that used in the solutions (Will discuss more in next chapter).

Though the whole project took longer time than what we expected, the methodology was well followed and went smooth and the BI solutions were developed as designed.

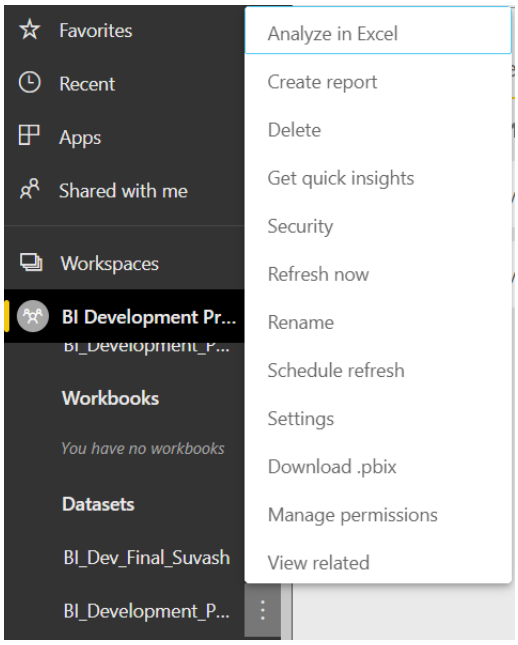



*Figure: Datasets can be downloaded as pbix file and used in Power BI Desktop*

## 6.2 Challenge

As mentioned above, we encountered some issues on the measures. After Suvash shared the teacher's comment about currency, I realized that it was not taken into account in my solution either. Therefore, we started to work on it. At the beginning, I thought about take the dynamic currency data through the Web. Meanwhile, Suvash considered to use "Unit Price" and "Unit Cost (LCY)" to calculate the currency rate. Later I noted that NAV has its own Currency Tables. However, the metadata of the Exchange Currency Rate confused me. After some research, I understood that the field "Exchange Rate Amount" represents the Local Currency Amount while "Relational Exch_Rate Amount" represents the Foreign Currency Amount.

It happened when I worked on the primary key. Take Sales Data as the example. I appended the active and archived data as new tables named Sales Header and Sales Line. After that, I tried to create the data model between Sales Header and Sales Line using the combination key of Document Type and Document No. It created a many to many relationships, which was not useful for the solution. It took me quite long time to understand that NAV archives data more than once. Therefore, I created the relationships for Sales Header and Sales Line by filtering the active data (which is not archived) and the first-time archived data (whose version No is 1).

In my perspective, the unfamiliar with NAV system is the most challenge and hinder my progress on the BI solution development.

### 6.3 Conclusion

Though the project is a bit more complex than what I thought, I found myself very interested to it and acquired much more than what I expected on both NAV system and Power BI as well as the process of analytics project development.

## Append

Project Backlog.xlsx