ETL Developer at Nassprotech Consulting

Submitted By

Muhammad Sajid Osman

ID: 191-35-2660

Department of Software Engineering

Daffodil International University

Supervised By

Khalid Been Md. Badruzzaman Biplob

Senior Lecturer

Department of Software Engineering

Daffodil International University

The internship report has been submitted in fulfillment of the conditions for the degree of Bachelor of Science in Software Engineering

Fall-2022

© All right Reserved by Daffodil International University

APPROVAL

This thesis titled on "**Jr. Software Engineer at Nassprotech Consulting**", submitted by **Muhammad Sajid Osman (ID: 191-35-2660)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Chairman

Dr. Imran Mahmud
Head and Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1

Tapushe Rabaya Toma
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2

Khalid Been Badruzzaman Biplob
Lecturer (Senior)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

External Examiner

Md. Tanvir Quader
Senior Software Engineer
Technology Team
a2i Programme

© Daffodil International University

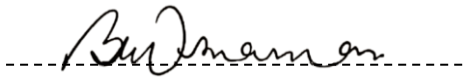
DECLARATION

I hereby declare that this internship report "ETL Developer at Nassprotech Consulting" is grounded on the activities of my four months internship at Nassprotech Consulting Company. This is an original report prepared by me and supervised by “Khalid Been Md. Badruzzaman Biplob, Senior Lecturer, Department of Software Engineering".

This report is unique and has not been preliminarily used for any academic purpose. All the data and information contained in this report have been collected by myself. I shall be solely responsible for any miscue in the said information and data and if this data is used away for misleading purposes.

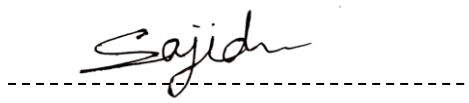
Eventually, this report has been submitted to the “Department of Software Engineering, Daffodil International University”, in partial fulfillment of the conditions for the Bachelor of Science degree.

Supervised By:



Khalid Been Md. Badruzzaman Biplob
Senior Lecturer
Department of Software Engineering
Daffodil International University

Submitted By:



Muhammad Sajid Osman
ID: 191-35-2660
Department of Software Engineering
Daffodil International University

© Daffodil International University

ACKNOWLEDGEMENT

Foremost, I am thanking the Almighty Allah for giving me the opportunity to complete my internship properly.

I would admire my deep respect to my supervisor "**Khalid Been Md. Badruzzaman Biplob, Senior Lecturer, Department of SWE**", due to his generous and amicable guideline. Without whose support this internship would not have been possible for me. I am eternally grateful to sir for his valuable time towards me and patiently solving all the issues I faced during my internship period.

Also I would heartily elicit my earnest gratitude to my organization's supervisor "**Md Sabbir Choudhury, CEO & Project Manager, Nassprotech Consulting**" for offering me the scope to work in his organization. His inspiration motivated me to work full time. For his constant support, I have learned a lot and able to apply it in my work.

I would like to show my sincere reverence to my mentor "**Asif Khan Shakir, Senior Data Analyst at Save the Children International**", Sir, who has always supported me with his valuable suggestions. I am forever indebted to Sir.

I elicit my earnest thankfulness to "**Dr. Imran Mahmud, Associate Professor & Head In-Charge, Department of SWE**", and respected faculty members of the "Software Engineering Department" for their help in completing my internship.

 © Daffodil International University

EXECUTIVE SUMMARY

To complete the degree of BSc in Software Engineering I had to do a thesis in my final year or an internship from any reputed company. In that case, I opted for an internship and served as an intern for four months at Nassprotech Consulting Company. Based on my experience and learning from the internship. This report "ETL Developer at Nassprotech Consulting" has been prepared.

Nassprotech is one of America's leading software development companies providing backend support for various organizations and software. The journey of Nassprotech started in 2016 by Md Sabbir Choudhury. Nassprotech usually deals with warehouse, database, and ETL.

In this report, I have discussed the type of work I have done in the organization. My working experience, motivation, and feelings are presented in this report. Here I tried to explain the type of work I did, how I completed it, and what I learned from it. My problems, scope, and various limitations while working are also included here.

Most of the time during the internship I had to work with database and ETL tools. Here I have always had to work following the software development life cycle. I always had to focus on bringing out a better quality product through analysis. Also, my various soft skills have also improved a lot by working in this company. How to communicate within the organization, and how to behave and talk with superiors are all that I learned by working here. I have thoroughly enjoyed my internship program at Nassprotech and the internship program has benefited me a lot in my future career and academic studies.

 © Daffodil International University

TABLE OF CONTENTS

DECLARATION i
ACKNOWLEDGEMENT iii
EXECUTIVE SUMMARY iv
CHAPTER 1: INTRODUCTION 1
1.1 Overview 1
1.2 Background 1
1.3 Origin of the Report 1
1.4 Motivation 1
1.5 Objectives 2
1.6 Scope 2
1.7 Limitations 2
CHAPTER 2: COMPANY OVERVIEW 3
2.1 About Nassprotech 3
2.2 Services of Nassprotech 3
2.3 Mission & Vision 3
2.4 Management 4
2.5 Chart of Management 4
2.6 SWOT Analysis of Nassprotech 5
2.6.1 Strengths 5
2.6.2 Weaknesses 5
2.6.3 Opportunities 5
2.6.4 Threats 5
2.7 Contact Details 6
2.8 Mentor Details 6
CHAPTER 3: COMPANY CULTURE 7
3.1 Team I Work With 7
3.2 Working Environment & Protocols 7
3.2.1 Rules & Regulations 7
3.2.2 Facilities 8
3.3 Comparative Analysis of Office Culture 8
3.3.1 Mixed Up With Office Culture 8
3.3.2 Entertainments & Refreshments 8
3.4 Internee Life Cycle 8
3.4.1 Getting Started 8

© Daffodil International University

3.4.2 Recruiting Policies....8
3.5 Professional Environment....9
3.6 Personal Expectations....9
3.7 First Day at Office....9
CHAPTER 4: TECHNOLOGY EMPLOYING....10
4.1 Fundamental Technologies....10
4.2 Supportive Technologies....10
4.3 Technology in Use for the Projects....11
CHAPTER 5: COMPANY ROLES & RESPONSIBILITY....12
5.1 Training....12
5.2 Project Name....12
5.3 Project Plot & Requirements....12
5.4 Solution Providing....12
5.5 Project Goal....13
5.6 Assigned Responsibilities....13
5.7 My Contributions....13
5.7.1 Pervasive Map Design....13
5.7.1.1 Source Connection....14
5.7.1.2 Target Connection....14
5.7.1.3 Map Fields....14
5.7.2 Pervasive Process Design....14
5.7.2.1 RIFL Scripting....14
5.7.2.2 Database....14
5.7.2.3 Invoker....15
5.7.3 Pentaho Transformation Design....15
5.7.4 Pentaho Job Design....15
5.7.5 Technical Process Documentation....16
5.7.6 SQL Query Design....16
5.7.7 Create SQL Store Procedure....16
5.7.8 Tickets Solving....16
5.7.9 Unit Testing to FICO Debt Manager....17
5.8 Challenges....17
CHAPTER 6: INTERNSHIP OUTCOMES....18
6.1 Overcome Problems....18
6.2 Working Practices....18
6.3 Technological Enhancement....18

 © Daffodil International University

6.4 Non-Technical Growth 18
6.5 Achievement 19
CHAPTER 7: CONCLUSION & FUTURE SCOPE 20
7.1 Conclusion 20
7.2 Future Scope 20
APPENDICES 21
Appendix A: Pervasive Source Connection 21
Appendix B: Pervasive Target Connection 21
Appendix C: Pervasive Map Design 22
Appendix D: Pervasive Process Design 23
Appendix E: Pervasive Scripting 24
Appendix F: Pervasive Database Connection 24
Appendix G: Pervasive Invoker Setup 25
Appendix H: Pentaho Job Design 25
Appendix I: Pentaho Transformation Design 27
Appendix J: ETL Process Diagram Design 29
Appendix K: ETL Process Documentation 30
Appendix L: SQL Stored Procedure 32
Appendix M: FICO Debt Manager 33
REFERENCES 34

 © Daffodil International University

CHAPTER 1: INTRODUCTION

1.1 Overview

The internship is a program through which we can work in real-life industry and gain experience from there. Through this, we can learn the discipline of working in a company. As an undergraduate student, I had to complete an internship at a reputed company to complete my graduation. Where I had to work following the software development process and learn the etiquette and discipline of working in a company. In view of this, I started my internship at Nassprotech Consulting Company. Where I worked as an intern for four months.

Through the internship, I was able to gain real-life working experience and learn about the work environment which was very important for my course studies and also my personal career. Most of the time I had to work with the database there. I faced many problems during the internship and my teammates helped me to overcome them. I am grateful to the Nassprotech family for giving me the opportunity to do such a wonderful job.

1.2 Background

I joined Nassprotech Consulting Company on July 15, 2022, as an ETL Developer. I had to train there for the first month, and after that, they added me to their main developer team. I had to work with structured databases and various ETL tools there. Also, I had to write documentation and draw diagrams at various times for different processes. There I also participated in various team meetings and programs.

1.3 Origin of the Report

This report represents four months of my working period at Nassprotech Consulting. How I worked in Nassprotech, what were my responsibilities in the organization, and how I fulfilled them, all those things are mentioned in this report. This report also includes my achievements from how I have contributed to the organization.

The report was supervised by “Khalid Been Md Badruzzaman Biplob, a respected Senior Lecturer, in the Department of Software Engineering”, at Daffodil International University.

1.4 Motivation

The internship program gave me a great opportunity to enhance my skills and experience. The internship connects me with the software industry. I always wanted to work in a reputed

© Daffodil International University

software company, the internship gave me that opportunity. The experience and knowledge gained from this will help me in the future. By working in this company I got a thorough understanding of the software industry. Which will help me to become a complete asset to the country and the workplace. And I am determined to give my best service for this.

1.5 Objectives

The main objective of this report is to know in detail about Nassprotech Consulting Company and present the activities which I performed there by following the software development life cycle. Some special objectives of this report are:

- To describe the activity and founding history of the organization.
- To explain the responsibility I have performed in the organization.
- To describe the team I worked with.
- To describe the environment of the organization.
- To explain the contribution I have made to the organization.
- To explain the experience which I have gathered from the organization.

1.6 Scope

As an intern, I got the opportunity to see the working process of the company very closely. I had the opportunity to work with real-life databases and data of hospitals and banks there. Have seen and worked on large ETL processes from scratch. Which will enable me to work in multinational database companies in future. I got to know business patterns and product development processes by working there. Also, I got experience working with multiple ETL tools working there. Which will help me to become an efficient ETL developer in future.

1.7 Limitations

As Nassprotech works with databases of different companies and they have a lot of sensitive information about different companies, which they have to maintain confidentiality. To maintain this confidentiality and as per my agreement with the company I cannot disclose much information outside. Moreover, due to a lack of sufficient time and resources, I had to face many difficulties.

© Daffodil International University

CHAPTER 2: COMPANY OVERVIEW

2.1 About Nassprotech

Nassprotech is a leading business process development and consulting company in the USA. They are well known for providing business process automation and complete business solutions. They basically create ETL processes to manage databases and provide backend support to clients. They provide software development, database maintenance, and data warehouse-related services.

2.2 Services of Nassprotech

Nassprotech Consulting offers the following services to their clients:

- Business Process Management.
- Business Consulting Services.
- Applications & Scripting.
- Consulting & Custom Development.
- Web Applications & API's Development.
- Financial Services Consulting.

2.3 Mission & Vision

Their main aim is to make Nassprotech one of the leading software companies in the data warehouse sector and spread Nassprotech across countries. Nassprotech is working towards providing maximum support to their clients with minimum effort. Their mission is to deliver quality products to their clients. Their main missions are:

- Providing quality services and products to clients.
- Becoming a leading company in this field.
- Spreading the company to different countries.
- Creating strategic partners in the market field.
- Keeping existing customers happy.

© Daffodil International University

2.4 Management

Nassprotech Consulting has a well-organized management system. Every decision of the company is taken through team meetings and discussions. The management hierarchy is been followed while assigning work to everyone through meetings. Customer handling and support system is also very good here. The superior officers are also very friendly with their employees. There is a very good relationship with everyone in the office.

2.5 Chart of Management

A chart of management represents the hierarchy of management in a company. The chart of management of Nassprotech Consulting is presented below:

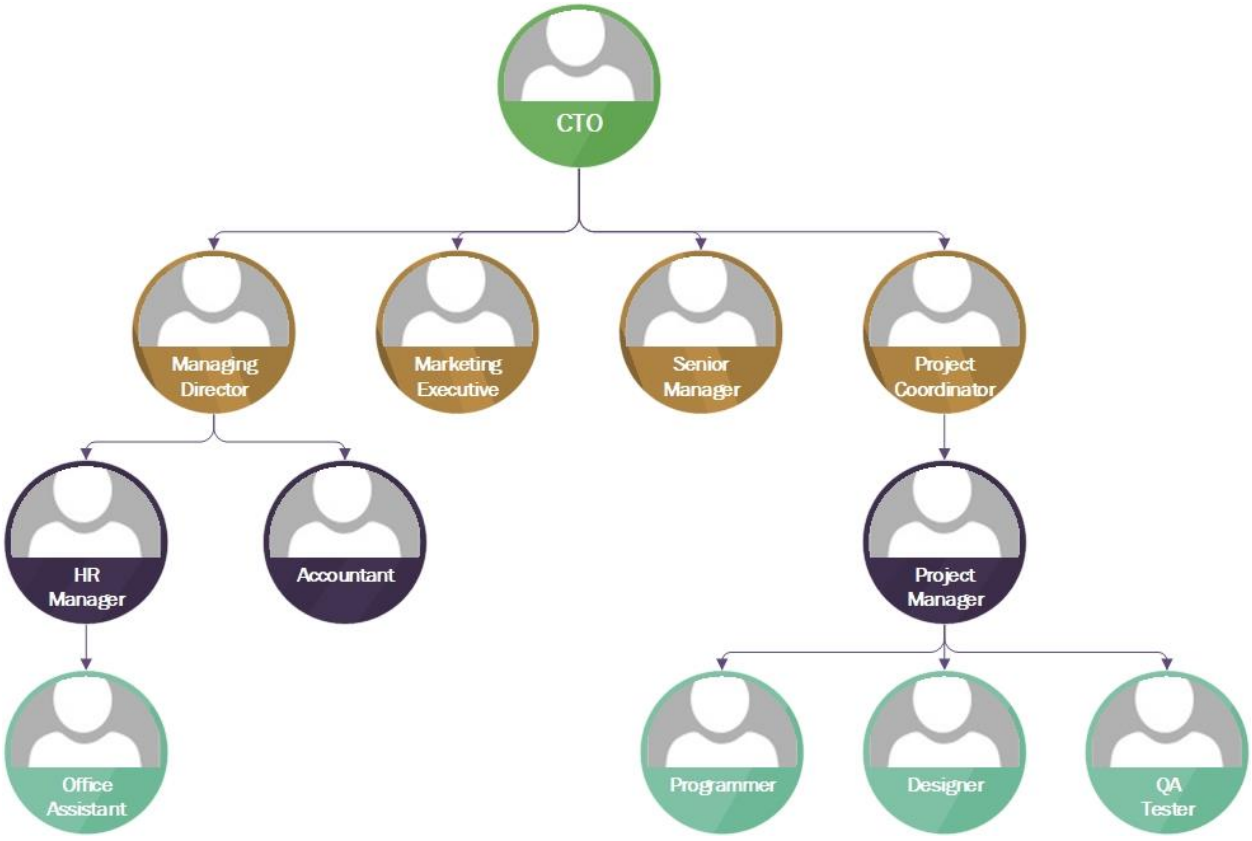


Figure: 2.1 Chart of Management

© Daffodil International University

2.6 SWOT Analysis of Nassprotech

SWOT analysis is a technique that helps to find out a company's strengths, weaknesses, future opportunities, and threats. SWOT analysis of Nassprotech Consulting is presented below:

2.6.1 Strengths

- Nassprotech has skilled manpower, who can easily do many complex tasks.
- They have advanced data analysis capabilities.
- They have good infrastructure to work with software development and data warehouse.
- They have many supporting partners in the marketplace.
- The company has a nice working environment.
- They can complete many big projects in a very short time.

2.6.2 Weaknesses

- Most of the time they have to maintain a high level of privacy while developing products otherwise they have to pay a high level of price.
- They don't have a good website.
- The company faces difficulties in spreading across multiple countries due to fewer manager.
- There is conflict over the company's successor.

2.6.3 Opportunities

- Nassprotech is connected with a good community from where they can get support.
- The demand for data warehouse is increasing day by day.
- The government also investing in this sector.

2.6.4 Threats

- The number of competitors is increasing day by day.
- There will be a big disaster if the data gets leaked.
- Due to not having a good website, it will also be difficult to get new projects in the future.
- There may be future threats to the company's successor.

© Daffodil International University

2.7 Contact Details

Company name: Nassprotech Consulting.

Website: https://www.nassprotech.com/

Email: info@nassprotech.com

Address: Minnesota City, MN, USA.

2.8 Mentor Details

Name: Md. Sabbir Choudhury (CEO & CTO)

Email: sabbirchoudhury@nassprotech.com

© Daffodil International University

CHAPTER 3: COMPANY CULTURE

The best thing about Nassprotech is that they want to build a family, not a team. Here the team members stay together and work together, the bonds between them go beyond simple co-worker relationships. There are certain rules and regulations of the company which all the employees of the company follow properly. This chapter describes the company's rules, regulations, culture, and environment.

3.1 Team I Work With

Here I am working with the ETL developer team. Our team consists of five members, two of whom develop ETL and the rest support their work. At the beginning of each day, our project manager assigns work to us and after that, we complete the work. Everyone in our team has a harmonious relationship and we all support each other in our work. I got a lot of support from everyone on the team. They helped me a lot to understand and complete my work.

Two key team members contributed to creating the SQL stored procedure and the rest helped to create the data mapping and process steps. Thus we create a large ETL process. Our project manager also encouraged us to work and sometimes helped us a lot.

3.2 Working Environment & Protocols

Every employee of Nassprotech has to follow certain rules and regulations. The rules and regulations provided by Nassprotech are given below.

3.2.1 Rules & Regulations

- Every employee has to work in the office from 7 pm to 4 am.
- Communication and exchange of information between employees must be done through company-provided email.
- All decisions must be discussed with management before approval.
- Sensitive information of the company can't be disclosed outside.
- Must attend every meeting of the company.
- Daily work report should be submitted to everyone supervisor.
- All communication between employees must be in English.

© Daffodil International University

3.2.2 Facilities

- They provided us with strong internet connection.
- Granted us 2 days holiday per week.
- The company provided us with high configuration laptops.
- Granted us bonuses at various occasions.

3.3 Comparative Analysis of Office Culture

Nassprotech has a conducive working environment. Here all the teams are composed of fresher's and experienced people. This allows fresher's to gain experience and learn on the job from seniors.

3.3.1 Mixed Up With Office Culture

Everyone in the company is properly familiar with the company culture and everyone adheres to it properly. I am also familiar with the company culture and am very happy to work here. They gave me complete freedom to work from home. I am now fully engaged with the company culture.

3.3.2 Entertainments & Refreshments

Everyone can work comfortably in the company, no force or pressure is created to work. Everyone is given ample time for refreshments. Nassprotech doesn't provide us with any entertainment but still, we don't feel bored to work because we find our joy in gossiping.

3.4 Internee Life Cycle

All interns are required to work as an intern for four months after joining the company. Those who perform well get a permanent job in the company. Every company follows some rules for hiring interns. Nassprotech follows certain policies in hiring interns.

3.4.1 Getting Started

I started my internship at Nassprotech Consulting Company on 15th July 2022. Before starting the internship I had to face an interview with Nassprotech authority. Where various questions are asked about my goals and purpose for working in the company.

3.4.2 Recruiting Policies

At the start, I sent my CV to Nassprotech Company through one of my references. As there are some senior brothers from our department in the company, I get an easy opportunity to give an

 © Daffodil International University

interview. After completing the interview they gave me some work to do. As I completed them successfully, I got the opportunity to join the company as an intern. After joining the company they gave me on-the-job training for some time.

3.5 Professional Environment

As I said earlier, the working environment here is very good. All kinds of facilities are provided to the employees here. Moreover, the employees of the company make the environment of the company better. I was able to adapt to the company environment very easily. My teammates were very friendly with me.

3.6 Personal Expectations

I had some personal goals and expectations that I wanted to achieve from this internship.

- Learning to work in a company following the software development life cycle.
- Gaining detailed knowledge about ETL and database.
- Gaining an understanding of the environment of a multinational company.
- To gain a detailed understanding of SQL queries and stored procedures.
- Learning to communicate and take responsibility in the company.
- Gain advanced data analysis knowledge.
- Creating my own identity in the industry.

3.7 First Day at Office

I was very nervous on my first day of joining the office. I was very worried about the company culture, work rules, and regulations. But everything went well. I was asked to attend by 7 pm, but I arrived before that. Everyone welcomed me as a new recruit. That day they gave me a company email and access to a personal computer. Also, teach me all the rules and regulations of the company. After some time the CEO of the company arrived there and welcomed me. He discussed my study area and my various responsibilities. After that, he explained to me my duties and handed me over to my senior colleagues. All in all I had a great first day.

© Daffodil International University

CHAPTER 4: TECHNOLOGY EMPLOYING

4.1 Fundamental Technologies

Since I worked as a data warehouse and ETL developer at Nassprotech, I had to use multiple tools related to databases and ETL. Also, I had to use some programming language to create the ETL process.

The programming languages I had to use were:

1. RIFL (Rapid Integration Flow Language)
2. JavaScript
3. Java

I used these programming languages to set variables, and apply various conditions within the ETL process.

I also had to write SQL queries to interact with the database in the ETL process. I had to use MS SQL Server tools for this task.

The ETL tools I had to use were:

1. Action Data Integrator.
2. Pentaho.

Also, I had to use other tools for creating documentation and working with the database, they are:

1. Microsoft Visio.
2. Microsoft SQL Server.

4.2 Supportive Technologies

Our team mainly provides background support to a software named Fico Debt Manager. We use Fico Debt Manager Software to validate our newly developed processes and queries. This software includes many operations like client management, portfolio management, account management, etc. which we use to create ETL processes and SQL-stored procedures. Also we use MS Visual Studio for report generation and Jira for testing.

 © Daffodil International University

4.3 Technology in Use for the Projects

The technologies we are using in the current project are:

- Pentaho.
- Microsoft SQL Server.
- JavaScript

We are currently converting the previously built ETL processes in Action Data Integrator to Pentaho processes. That's why we are using Pentaho ETL developer software. Also, we are using MS SQL Server software to manage the database and create SQL procedures. We use JavaScript programming language for programming within processes.

© Daffodil International University

CHAPTER 5: COMPANY ROLES & RESPONSIBILITY

5.1 Training

Working as an intern in the company, I had to fulfill many responsibilities. For that, I had to take training from seniors for the first few days. I took training there in ETL tools, databases, and many more. They teach me how to work with these tools. The tools and technique I have trained at Nassprotech are given below:

- My SQL
- SQL Stored Procedure
- Action Data Integrator.
- Pentaho
- Excel
- FICO Debt Manager
- MS SQL Server

5.2 Project Name

The project I am working on its name is **New Business**. It is basically an ETL process of a hospital.

5.3 Project Plot & Requirements

Fico Debt Manager is an ERP system of a hospital. The ERP system has many ETL processes, one of them being the New Business Process. The requirements of this project are data validation, verification, and data processing as per company policy. Moreover, creating different types of reports from the data is also a requirement of this project. From time to time the requirements of the clients also change.

5.4 Solution Providing

We have developed ETL processes to automate this process of data validation and data management. Within the ETL process, we have integrated the system, data, and the databases. Within the ETL process, we have created a data map for data validation, verification, and business rule implementation. We have created many SQL queries and SQL-stored procedures for connecting ETL and database. Along with that, we have done programming to interact with the system. Above all, we have provided a complete solution to automate the processes.

 © Daffodil International University

5.5 Project Goal

The goal of this project is to create an ETL process to automate business processes. A process that collects data from various sources then transforms the data as per requirements and finally saves the data to the database. Also, creating various SQL stored procedures to generate reports.

5.6 Assigned Responsibilities

I worked as an ETL developer at Nassprotech. There I created the ETL process through various ETL tools. I also had to write SQL queries. The following are the responsibilities I performed there.

- Pervasive Mapping and Process Design.
- Write RIFL Scripting Language.
- Develop ETL in Pentaho.
- Draw diagrams and Documentations.
- Writing SQL Queries.
- Create Store Procedures.
- Solving Tickets.
- Unit Testing to FICO Debt Manager.

5.7 My Contributions

I have performed the duties assigned to me properly. I designed the map using the Action data integrator. At the same time, I have designed the complete process with the combination of many maps. Created the script using RIFL scripting language. I also created the entire ETL process using the Pentaho tool. While creating the ETL process I also had to create its diagrams and documentation. I also contributed to the development of SQL Stored Procedures. I also had to do unit testing on the FICO Debt Manager software.

5.7.1 Pervasive Map Design

Pervasive Map Design is a tool of Action Data Integrator through which data-related work is done. Through this, we implement data cleansing, data validation, and verification, and apply business rules to data. Here first we have to provide a data source in the source connection, then we have to give the address of where the final data will be stored and we have to set what to change in the data. Finally, if we run the map, it will automatically do all the work related to the data it was supposed to do. I have created several such maps in Nassprotech.

 © Daffodil International University

5.7.1.1 Source Connection

Source Connections is a part of Pervasive Map Designer. Where the source of the data has to be mentioned, that means where the data will come from. Here we can provide data in Excel files, text files, and CSV files in various formats as source. We can also use databases as sources. Here we can also connect to the database table. It is also possible to fetch data from multiple database tables. **Appendix A** has some pictures of these.

5.7.1.2 Target Connection

Target Connections is also a part of Pervasive Map Designer. Where the processed data is to be stored, the address has to be given there. Like Source Connection it also accepts Excel files, CSV files, TXT files, and as well as database. Data can also be stored in multiple database tables if desired. **Appendix B** has some pictures of these.

5.7.1.3 Map Fields

After setting the source connection and target connection, the map field is used for this data manipulation. Here first we have to determine which data of the source connection will go to which target. After that, we have to determine how to change which data through programming. For this, we can create different events in the map, do data filtering, and also scripting. **Appendix C** has some pictures of these.

5.7.2 Pervasive Process Design

At Pervasive Process Design, we create a complete ETL process. A process consists of multiple maps, conditions, and scripts. Through this process, we connect the database and data to the system. We had to create such an ETL process. **Appendix D** has some pictures of these.

5.7.2.1 RIFL Scripting

RIFL is a scripting language through which we program in the process. Where we can access files from the system by writing code. Can create folders and files on the system. Here we can do data manipulation using various conditions, variables, and codes. We had to write such a scripting language while creating the process. **Appendix E** has some pictures of these.

5.7.2.2 Database

The database is another step of the Pervasive process through which we execute database queries inside the ETL process. Through this, we connect to ETL and database tables. Reading data from the database, updating data, deleting and storing data in the database are all done

 © Daffodil International University

through it. Through this, we also execute the SQL Store Procedures. For ETL we create and execute SQL queries through it. **Appendix F** has some pictures of these.

5.7.2.3 Invoker

Through Invoker we can send emails to the administrator. After validating and verifying the data, we attach the problematic data to an excel file and send it to the administrator through Invoker. I have also contributed to the creation of these Invokers. **Appendix G** has some pictures of these.

5.7.3 Pentaho Transformation Design

Pentaho is another ETL developing software like Pervasive. ETL process can be created using Pentaho software. I had to convert ETL processes created in Pervasive to Pentaho process. Pentaho Transformation is a tool in Pentaho Process Designer, through which we work with data. What we used to do with Pervasive Map Designer can be done with Pentaho Transformation. It is easier to work with than Pervasive. I had to do data validation, and verification by creating a Pentaho transformation at Nassprotech.

Pentaho transformation also has many steps like Get Variable, Set Variable, Database Connection, Scripting, validation, verification, and many more. Through Set Variable and Get Variable we declare the variable and get the value of the variable. We connect to the database through a database connection step. Where we need to do programming, we use the JavaScript step. By filtering we can filter the data, and also we validate the data using various validation steps. We have to create a transformation using various steps like these. **Appendix I** has some pictures of these.

5.7.4 Pentaho Job Design

Pentaho Job is a complete ETL process. Within the Pentaho job, we create multiple transformations. Through Pentaho jobs, we interact with our systems and warehouses. Inside the Job, we determine which transformation will run after which transformation. Any operation with the system should be mentioned in all the jobs. Within a job, we have to perform various tasks including file management, SQL query execution, script writing, condition checks, etc.

We do scripting inside a job using JavaScript language. We need to work with variables so that the whole system is not static, we use a variable step to declare variables. ETL processes can be run for different conditions using condition checks. Using these things we create a complete job or ETL process. **Appendix H** has some pictures of these.

 © Daffodil International University

5.7.5 Technical Process Documentation

I have also been involved in the preparation of documentation and reports. We had to write documentation for all the ETL processes we created. We had to draw a separate diagram for each transformation or map and write its full description. We had to mention in the document how each step will work, how their flow will be in the system. I also had to draw a control flow diagram for each step. In this way, I have created a completely functional design document. **Appendix J, K** has some pictures of these.

5.7.6 SQL Query Design

Since I have been dealing with warehouses and databases, I have had to write SQL queries for that. Since our company works with relational database systems, I had to write the query for a relational database. Creating ETL processes requires writing SQL queries at various times, I had to design SQL queries for that. Clients send us their requirements for generating various types of reports. To generate all those reports we had to create SQL queries. Most of the time I had to write such SQL queries.

5.7.7 Create SQL Store Procedure

A SQL stored procedure is a bunch of SQL statements that we can store as one query inside the database. So that those queries can be reused multiple times later on. One of the advantages of SQL stored procedure is that we can use conditions, loops, functions, and variables in it. Which in short can be called programming to run SQL statements. The client sends us the requirement document to create the SQL stored procedure. Then we create the client's stored procedure according to that document. These stored procedures are then linked to the ETL so that the stored procedures are also run with the ETL process. Clients mainly use SQL-stored procedures for various report generation, database servicing, and management tasks. **Appendix L** has some pictures of these.

5.7.8 Tickets Solving

Our work is not limited to creating any stored procedure or ETL process. After creating a process we have to send it to the testing team. The testing team tests it and sends tickets to us then we have to solve those tickets. I also had to resolve such tickets after creating a new SQL stored procedure or map. After creating a new SQL stored procedure I had to spend the next few weeks solving tickets. The process had to be modified and redesigned for different data, different conditions.

© Daffodil International University

5.7.9 Unit Testing to FICO Debt Manager

FICO Debt Manager is an ERP system. All the ETL processes, stored procedures, and maps we create are for Fico Debt Manager Software. We mainly provide backend support for this software. When we create a stored procedure or ETL process, we test it on the Fico Debt Manager testing server. By running the process in this software, we can understand whether our process is working properly or not. Fico Debt Manager Software has many other sections including account management, work management, portfolio management. We test our system by providing demo data in all sections. Verify that the process we created is working properly. If there is any error, I correct the process accordingly. **Appendix M** has some pictures of these.

5.8 Challenges

My main challenge working here was to create processes as per the client's requirements. Because sometimes a query returns correct results, but it doesn't work properly for some exceptional data. It was a big challenge for me to make the process work properly for all types of data. Moreover, understanding the client's requirements correctly was a big task for me, because if the requirements were not understood correctly, the process creation would not be correct.

© Daffodil International University

CHAPTER 6: INTERNSHIP OUTCOMES

6.1 Overcome Problems

I faced many problems during the internship, at first I was very worried about it but slowly I learned to overcome my fear. I learned to deal with problems instead of running away from them. I learned to conquer my fear. I gradually overcome these problems of speaking English, understanding clients' requirements by reading documents, and making processes according to requirements. After overcoming all kinds of adversities, I am now capable of working in any company.

6.2 Working Practices

I have had good working experience by working at Nassprotech. I have developed a routine of working which will benefit me a lot in future. I am more motivated to work by working in a beautiful environment. I have learned all the rules and regulations of working in a company. All in all, I developed a good work habit.

6.3 Technological Enhancement

While working at Nassprotech I have been introduced to many new tools and software. There were many databases and ETL tools that I had never worked with before. Working with these new tools has greatly improved my technological skills. I have also gained a lot of knowledge in the database and warehouse sector. I have acquired a lot of skills in using new technologies, and I have learned how to adapt to new technologies easily. This knowledge of technology will help me a lot in the future workplace.

6.4 Non-Technical Growth

Along with technical growth, I have also had a lot of non-technical growth. I have also improved my personal skills a lot. I have improved my English speaking skills since most of the times I had to communicate in English in the office. Now I can speak English fluently. I have also improved my communication skills a lot. Before, I felt a lot of inertia in communicating, but now I can communicate fluently. I have improved some of my own in terms of leading. The attitude of giving leadership in me through working here has created. The attitude of giving leadership in me through working here has created.

© Daffodil International University

6.5 Achievement

The main purpose of joining my internship was to achieve the experience of working in the industry, which I was able to achieve in full. At the same time, I have gained a good knowledge of the data warehouse and database sector. I have also improved my personality and behavior a lot. All these experiences and achievements will help me a lot in the future.

© Daffodil International University

CHAPTER 7: CONCLUSION & FUTURE SCOPE

7.1 Conclusion

My working experience at Nassprotech was a dream come true. I never thought that I can work in such a good company. The experience of working in such a critical sector like a warehouse and database was also invaluable. I feel fortunate to have this opportunity. Working here I have been able to improve my technical and non-technical skills. Learned how a complete ETL process works. I got to know about the working environment in the company. All my colleagues in the company were very friendly to me, I got to know and learn a lot from them. The experience I have gained through working at Nassprotech in the past few months, I believe it will help me to create a bright future.

7.2 Future Scope

In today's software industry, the demand for data warehouses is increasing day by day. Because the amount of our data is increasing day by day, it is becoming very important to manage this data. Being a highly demanding job, my experience in the data warehouse sector will open up new horizons for me. Nassprotech Company wants to train and relocate us to America so that they can increase their human resources. In that case, it is a great opportunity for me to perform well here and move abroad.

© Daffodil International University

APPENDICES

Appendix A: Pervasive Source Connection

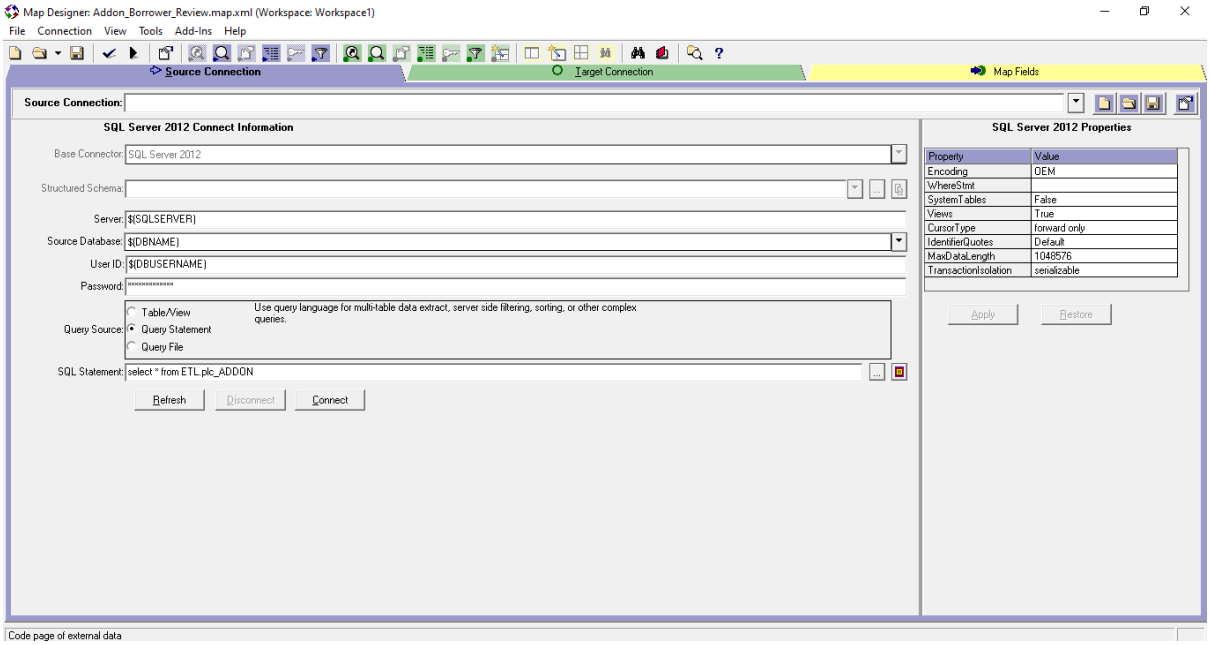


Fig A.1: Pervasive Source Connection

Appendix B: Pervasive Target Connection

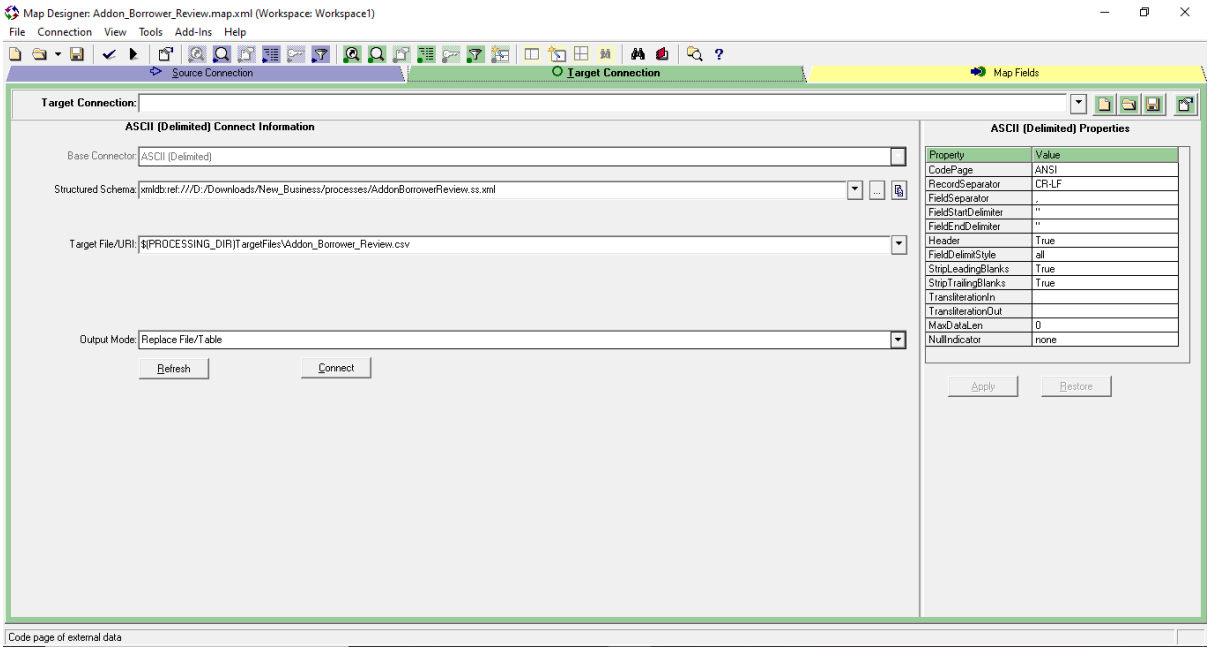


Fig B.1: Pervasive Target Connection

© Daffodil International University

Appendix C: Pervasive Map Design

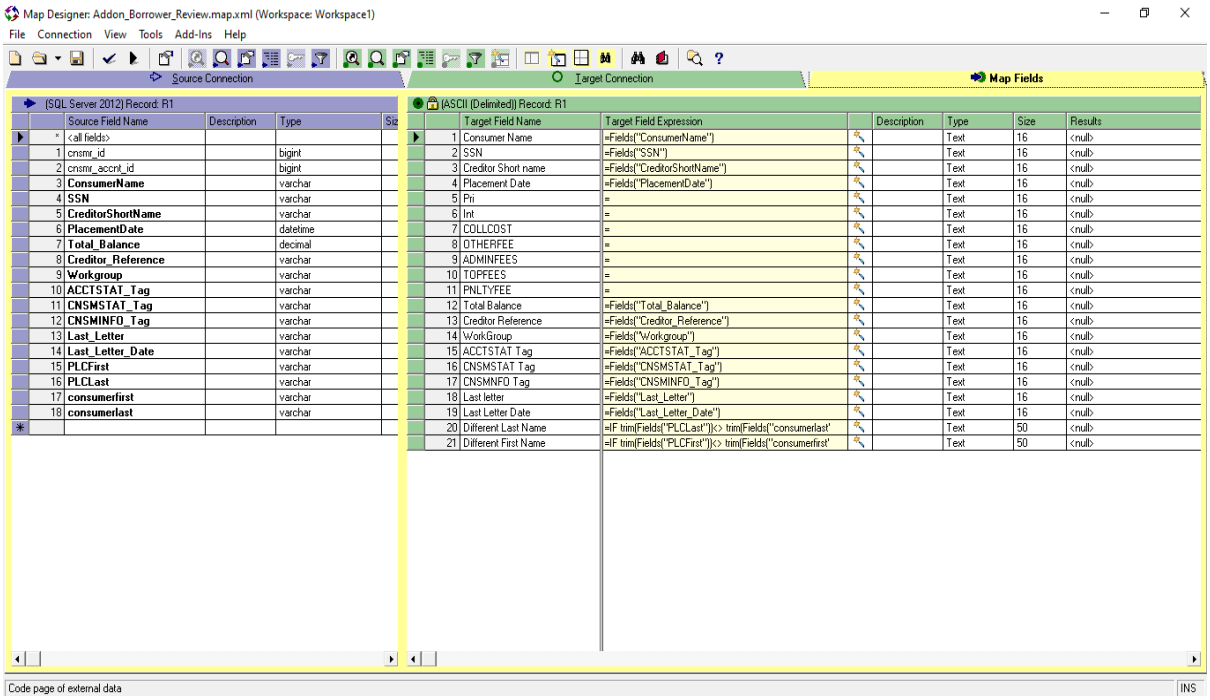


Fig C.1: Pervasive Data Mapping

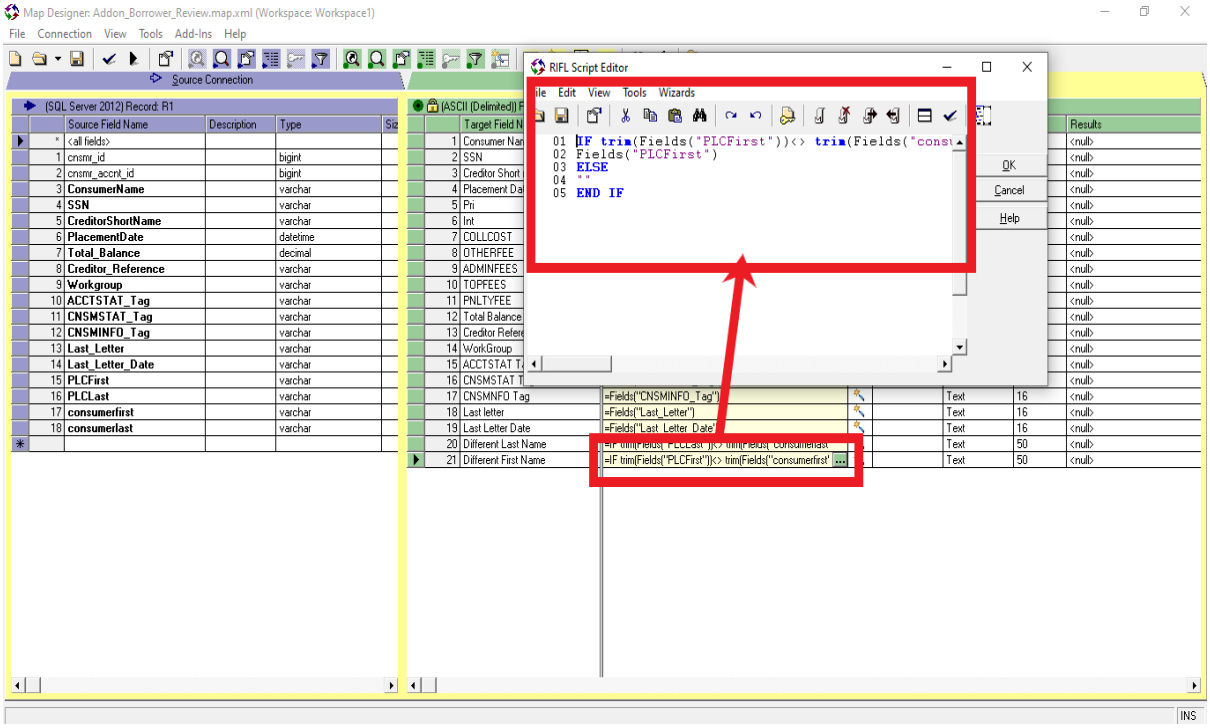


Fig C.2: Pervasive Data Validation

© Daffodil International University

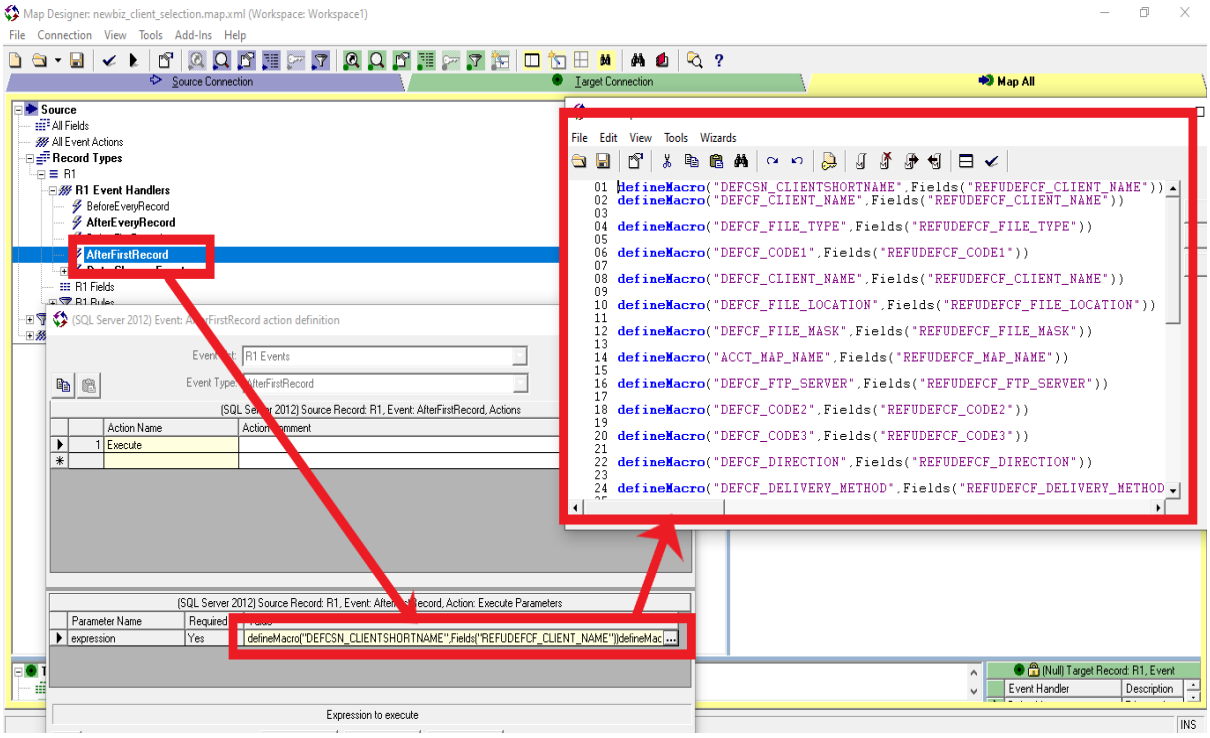


Fig C.3: Pervasive Setting Events

Appendix D: Pervasive Process Design

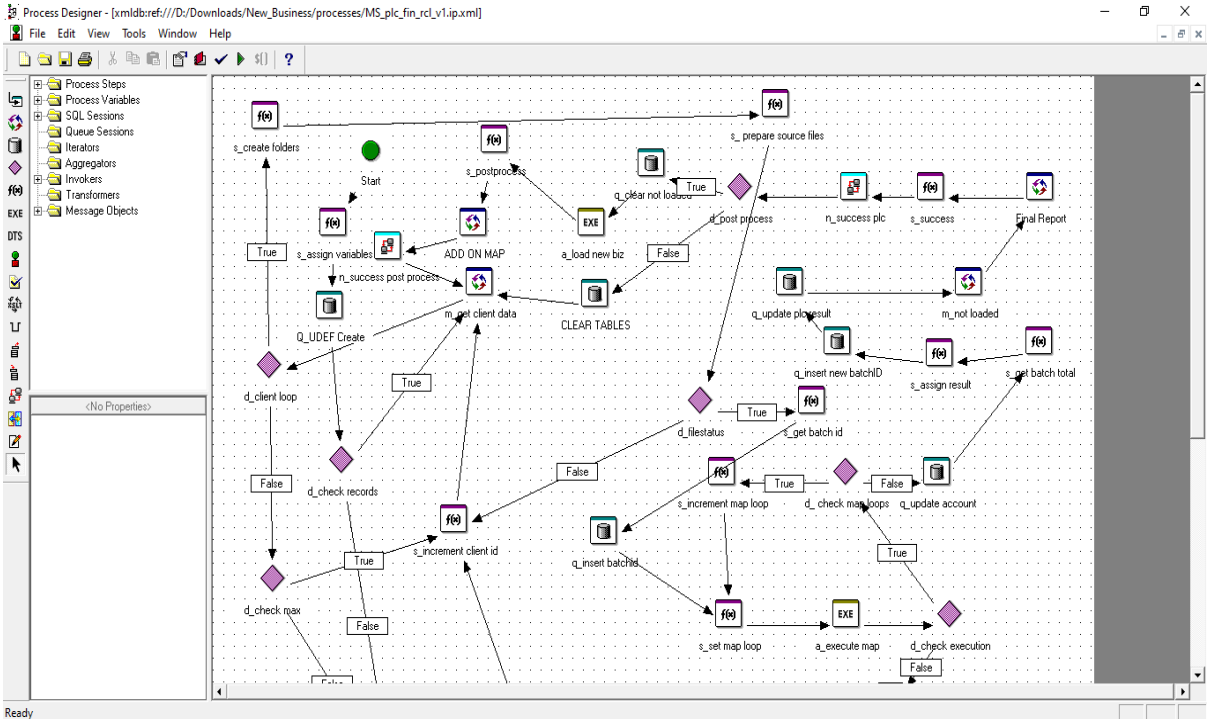


Fig D.1: Pervasive Process Design

© Daffodil International University

Appendix E: Pervasive Scripting

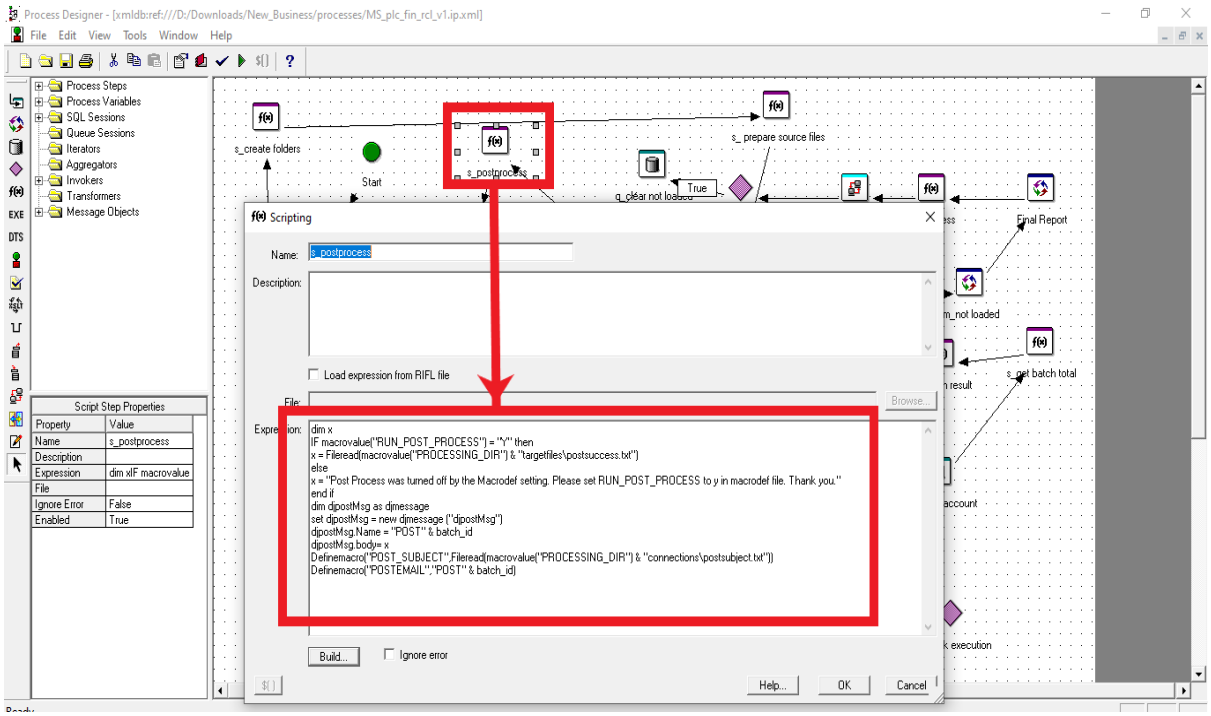


Fig E.1: Pervasive Scripting

Appendix F: Pervasive Database Connection

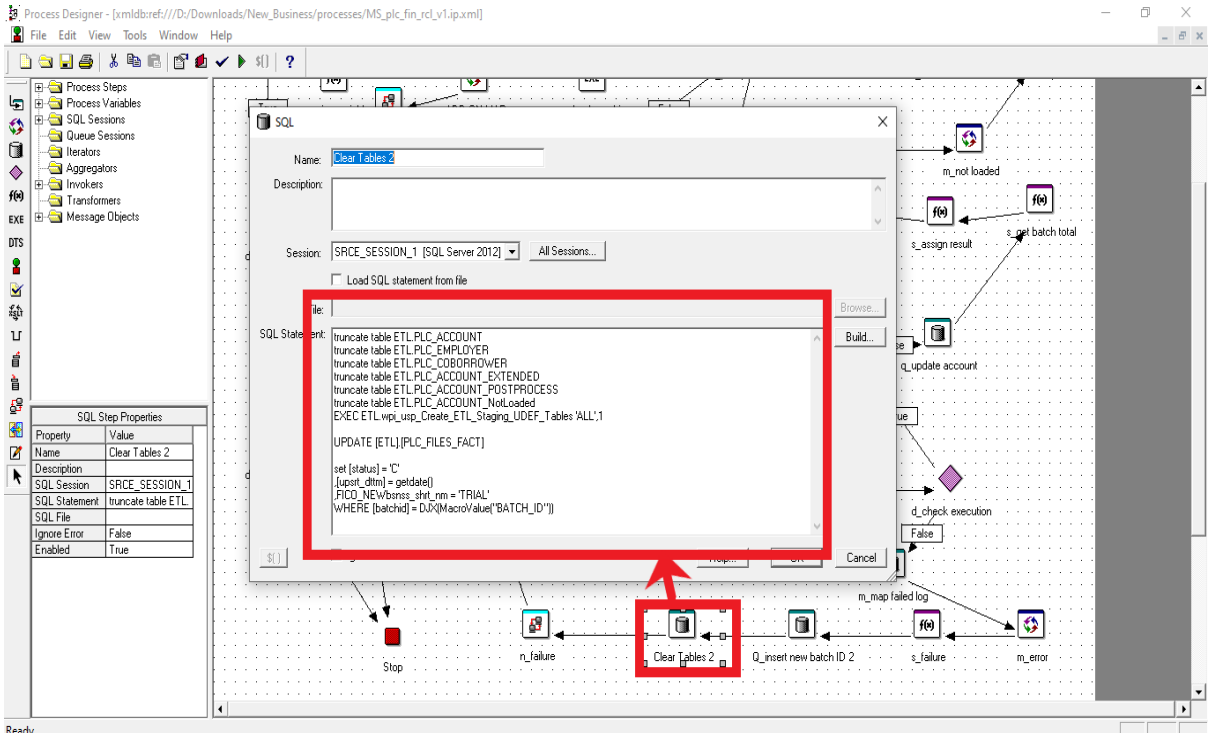


Fig F.1: Pervasive Database Connection

© Daffodil International University

Appendix G: Pervasive Invoker Setup

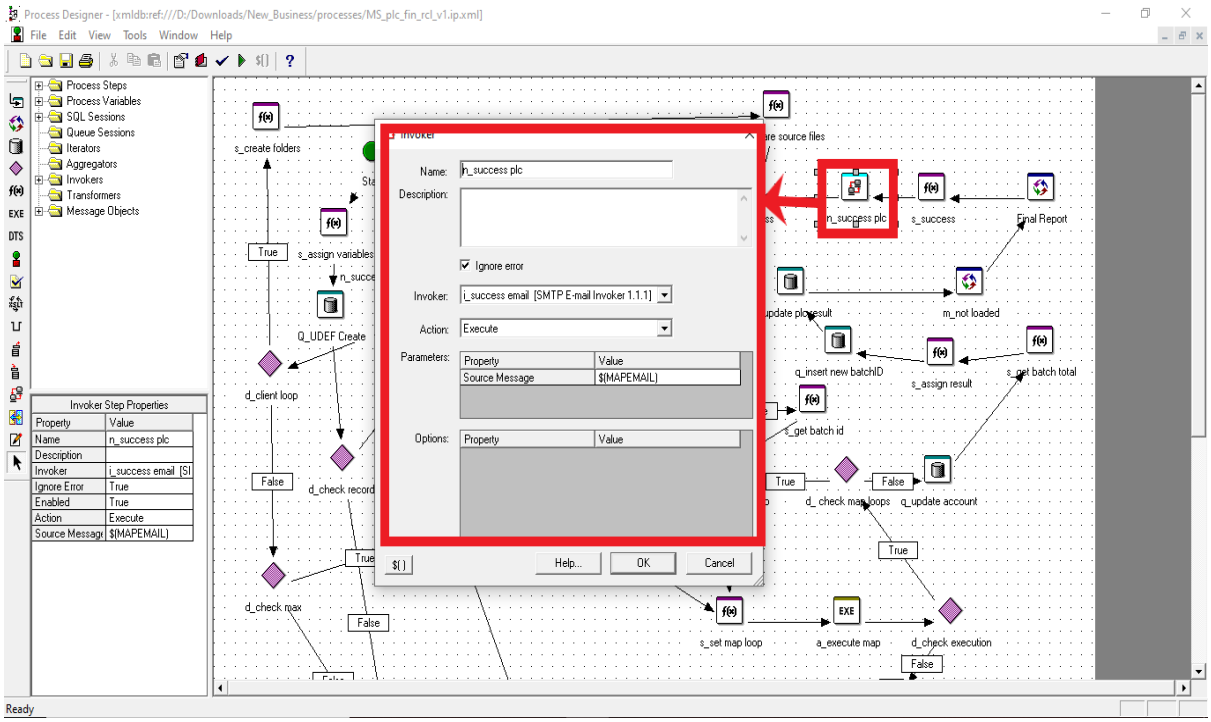


Fig G.1: Pervasive Invoker Setup

Appendix H: Pentaho Job Design

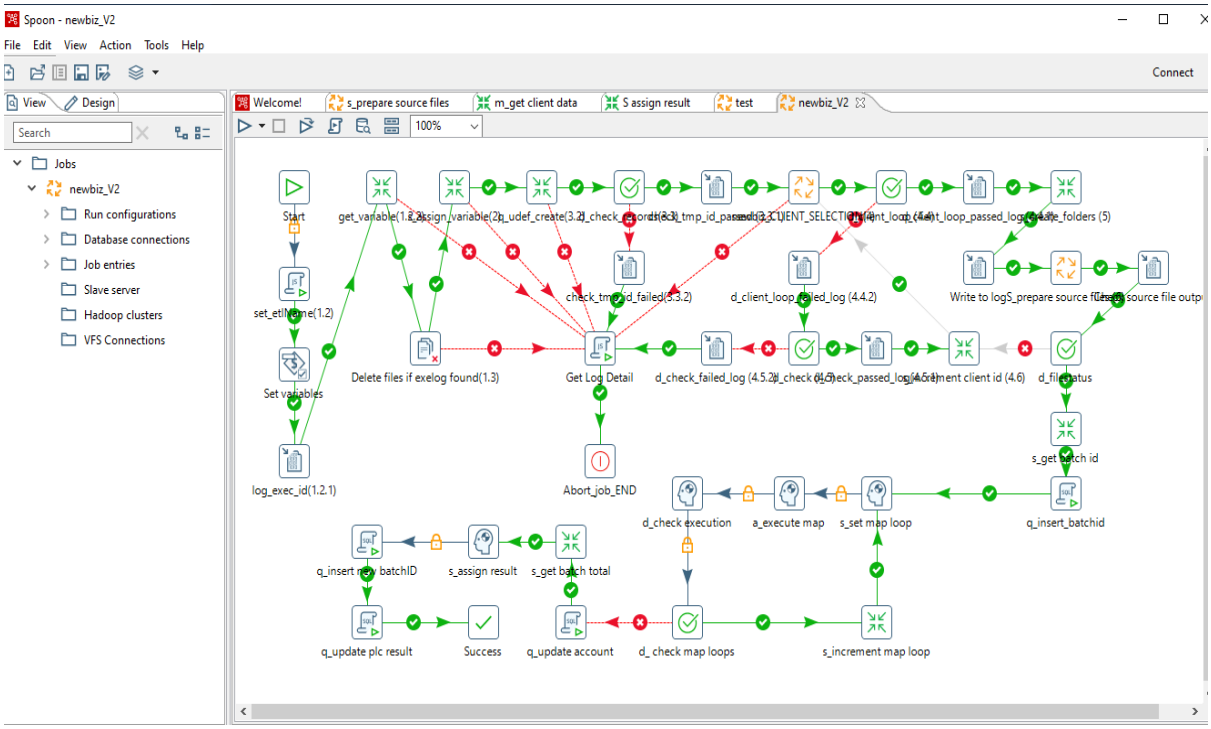


Fig H.1: Pentaho Job Design

© Daffodil International University

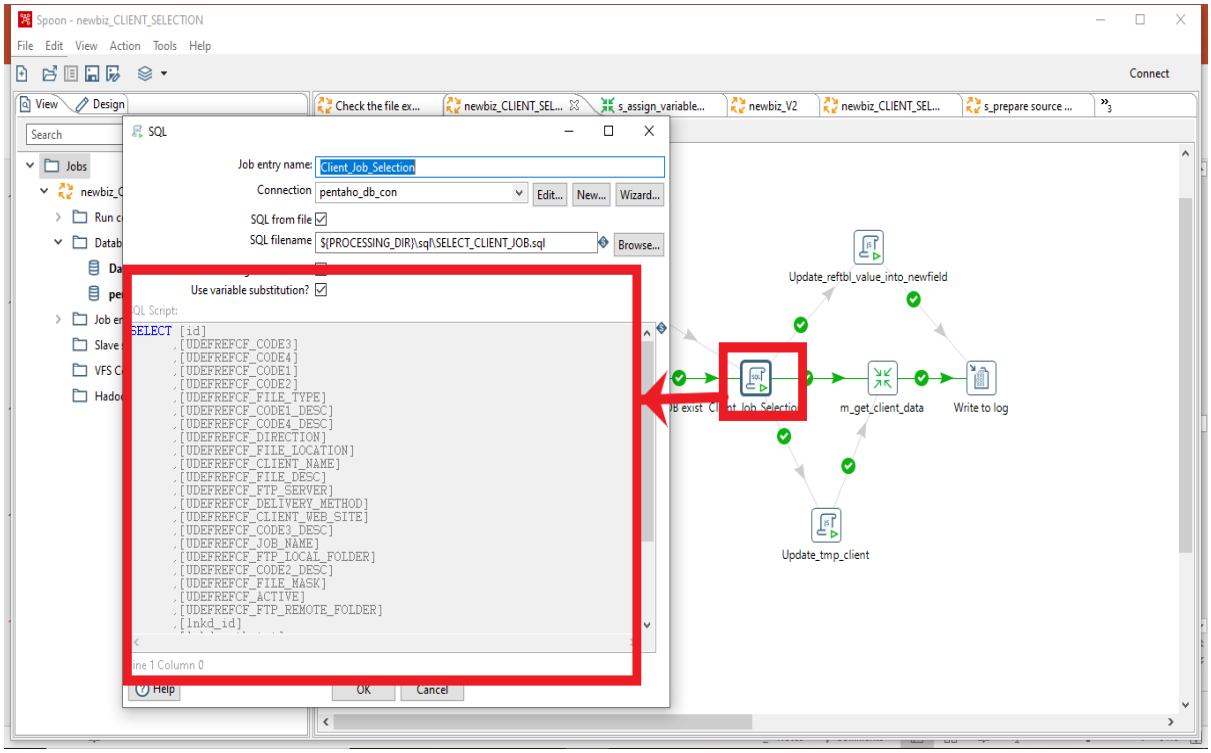


Fig H.2: Pentaho Query Writing

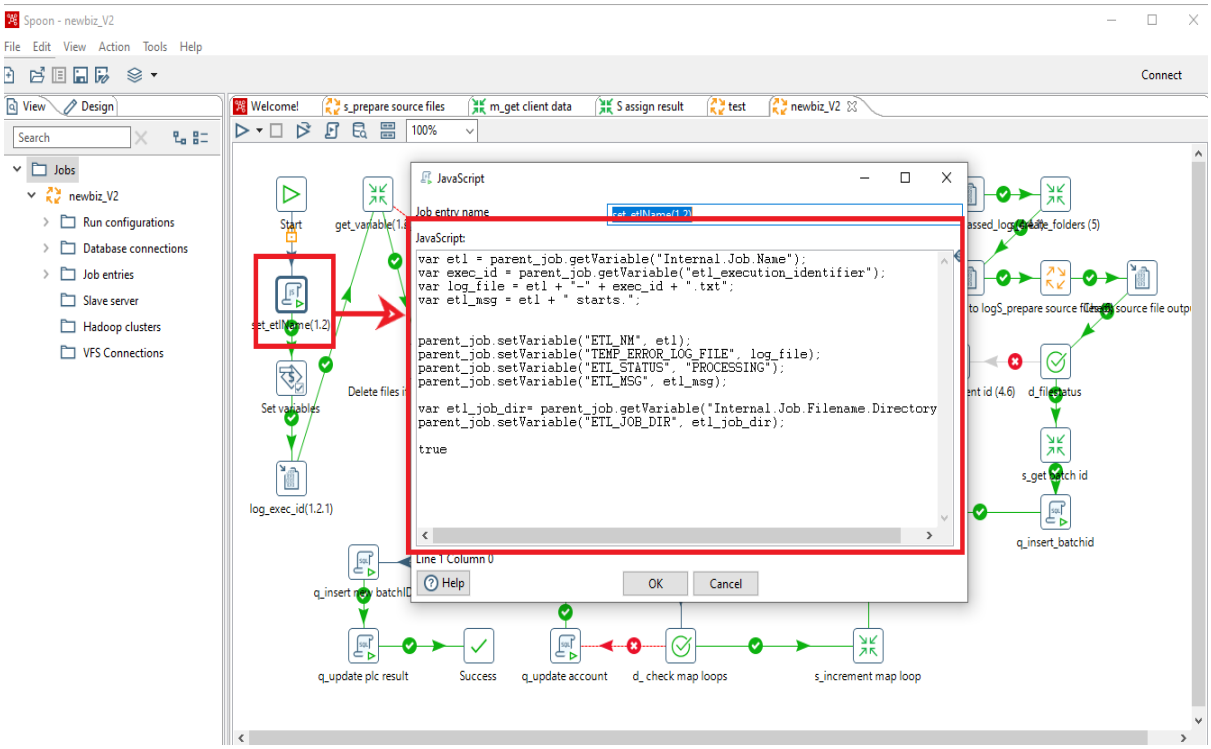


Fig H.3: Pentaho Scripting

© Daffodil International University

Appendix I: Pentaho Transformation Design

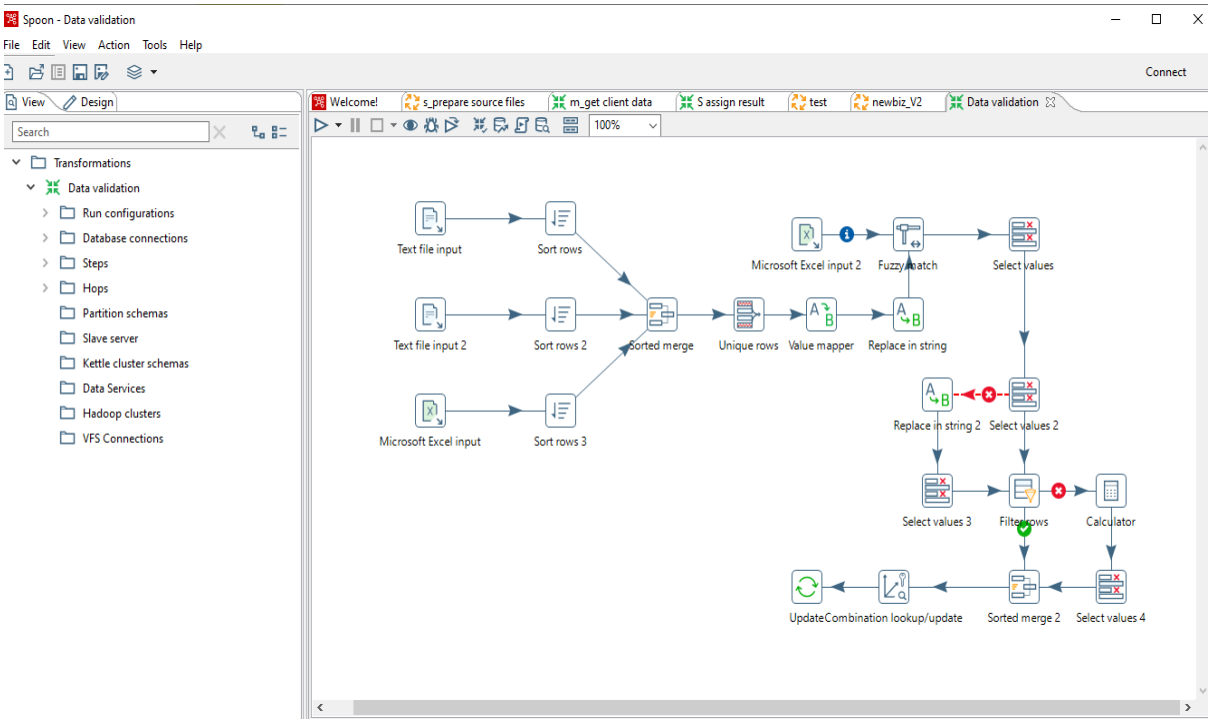


Fig I.1: Pentaho Transformation

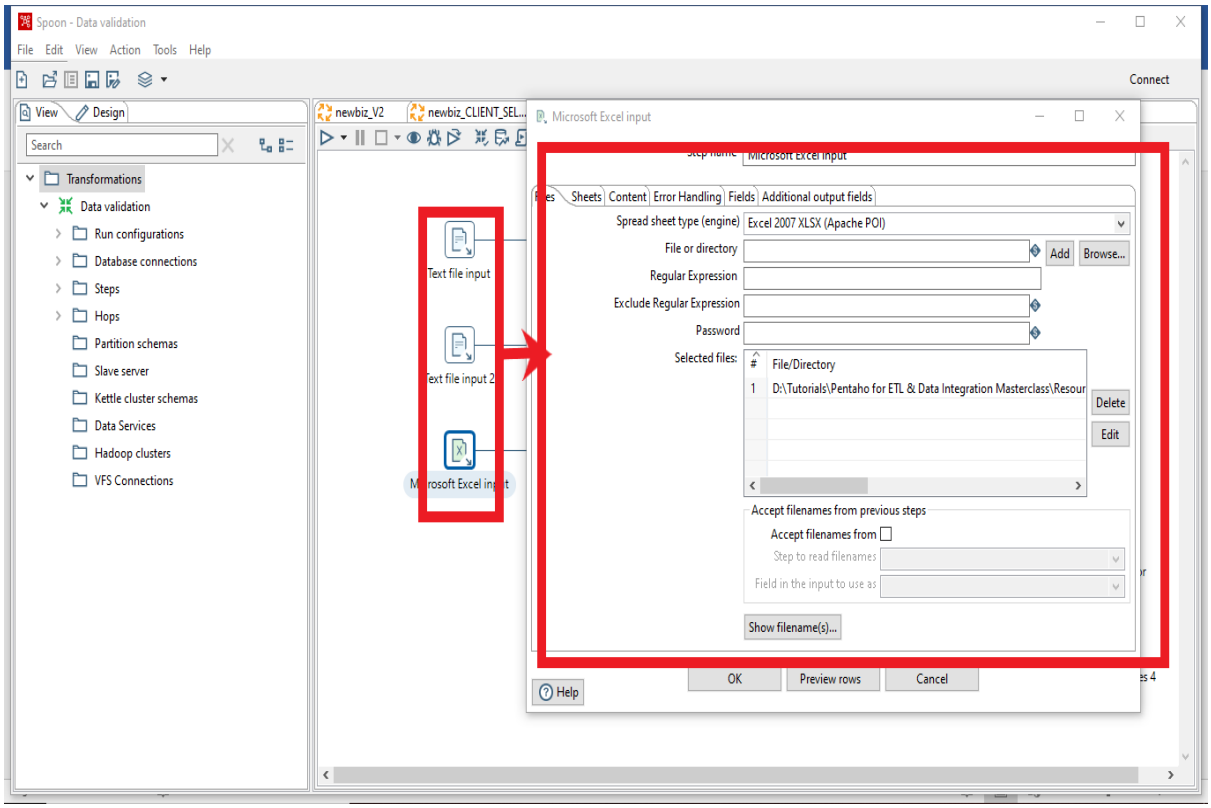
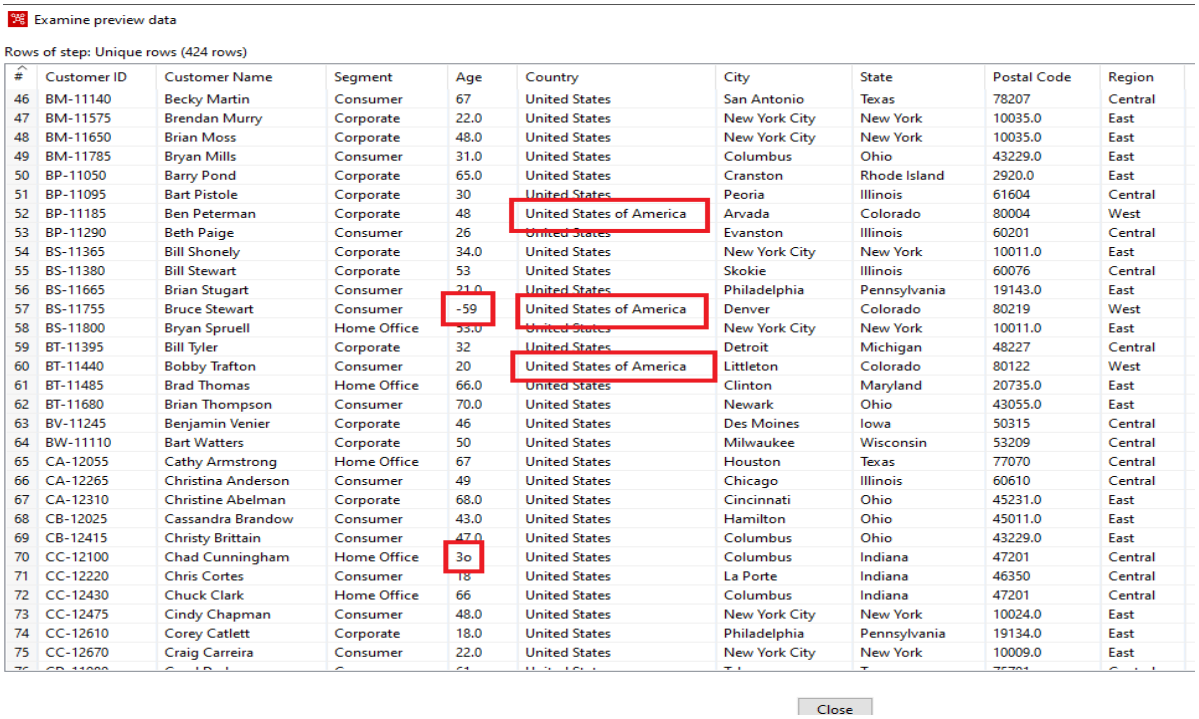


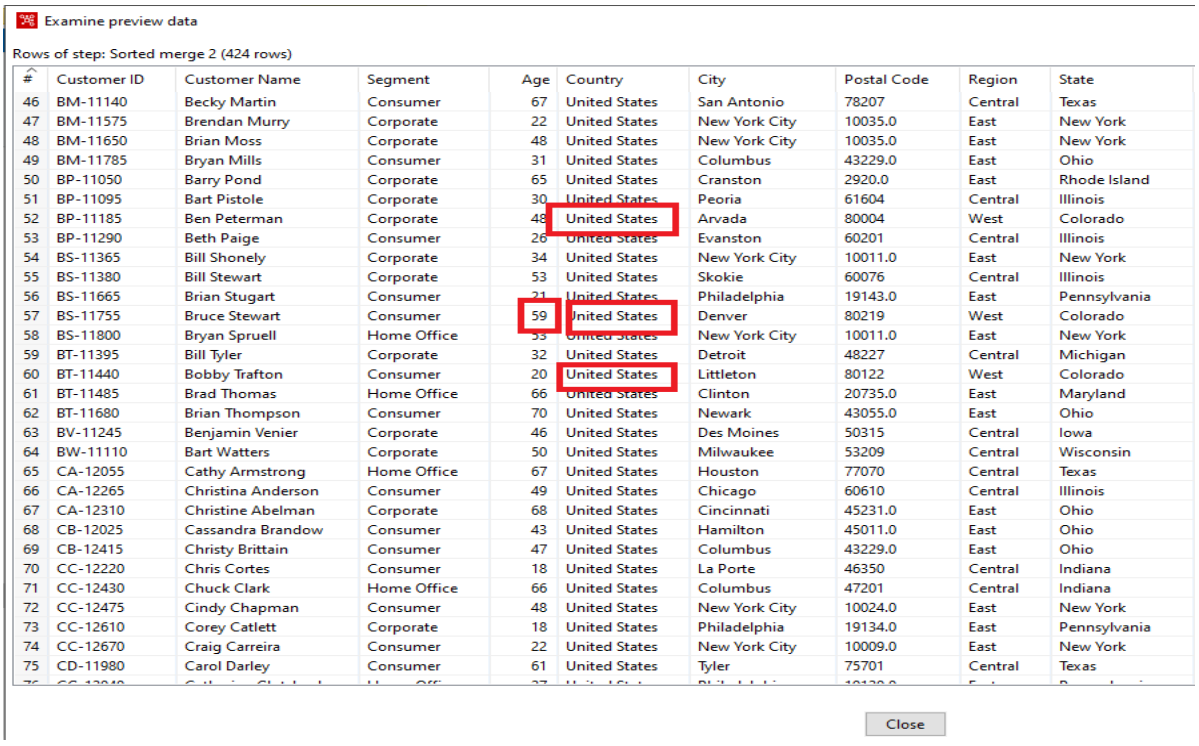
Fig I.2: Pentaho Loading Data from Different Source

© Daffodil International University



#	Customer ID	Customer Name	Segment	Age	Country	City	State	Postal Code	Region
46	BM-11140	Becky Martin	Consumer	67	United States	San Antonio	Texas	78207	Central
47	BM-11575	Brendan Murry	Corporate	22.0	United States	New York City	New York	10035.0	East
48	BM-11650	Brian Moss	Corporate	48.0	United States	New York City	New York	10035.0	East
49	BM-11785	Bryan Mills	Consumer	31.0	United States	Columbus	Ohio	43229.0	East
50	BP-11050	Barry Pond	Corporate	65.0	United States	Cranston	Rhode Island	2920.0	East
51	BP-11095	Bart Pistole	Corporate	30	United States	Peoria	Illinois	61604	Central
52	BP-11185	Ben Peterman	Corporate	48	United States of America	Arvada	Colorado	80004	West
53	BP-11290	Beth Paige	Consumer	26	United States	Evanston	Illinois	60201	Central
54	BS-11365	Bill Shonely	Corporate	34.0	United States	New York City	New York	10011.0	East
55	BS-11380	Bill Stewart	Corporate	53	United States	Skokie	Illinois	60076	Central
56	BS-11665	Brian Stugart	Consumer	21.0	United States	Philadelphia	Pennsylvania	19143.0	East
57	BS-11755	Bruce Stewart	Consumer	-59	United States of America	Denver	Colorado	80219	West
58	BS-11800	Bryan Spruell	Home Office	33.0	United States	New York City	New York	10011.0	East
59	BT-11395	Bill Tyler	Corporate	32	United States	Detroit	Michigan	48227	Central
60	BT-11440	Bobby Trafton	Consumer	20	United States of America	Littleton	Colorado	80122	West
61	BT-11485	Brad Thomas	Home Office	66.0	United States	Clinton	Maryland	20735.0	East
62	BT-11680	Brian Thompson	Consumer	70.0	United States	Newark	Ohio	43055.0	East
63	BV-11245	Benjamin Venier	Corporate	46	United States	Des Moines	Iowa	50315	Central
64	BW-11110	Bart Watters	Corporate	50	United States	Milwaukee	Wisconsin	53209	Central
65	CA-12055	Cathy Armstrong	Home Office	67	United States	Houston	Texas	77070	Central
66	CA-12265	Christina Anderson	Consumer	49	United States	Chicago	Illinois	60610	Central
67	CA-12310	Christine Abelman	Corporate	68.0	United States	Cincinnati	Ohio	45231.0	East
68	CB-12025	Cassandra Brandow	Consumer	43.0	United States	Hamilton	Ohio	45011.0	East
69	CB-12415	Christy Brittain	Consumer	47.0	United States	Columbus	Ohio	43229.0	East
70	CC-12100	Chad Cunningham	Home Office	3o	United States	Columbus	Indiana	47201	Central
71	CC-12220	Chris Cortes	Consumer	18	United States	La Porte	Indiana	46350	Central
72	CC-12430	Chuck Clark	Home Office	66	United States	Columbus	Indiana	47201	Central
73	CC-12475	Cindy Chapman	Consumer	48.0	United States	New York City	New York	10024.0	East
74	CC-12610	Corey Catlett	Corporate	18.0	United States	Philadelphia	Pennsylvania	19134.0	East
75	CC-12670	Craig Carreira	Consumer	22.0	United States	New York City	New York	10009.0	East

Fig I.3: Data before Transformation



#	Customer ID	Customer Name	Segment	Age	Country	City	Postal Code	Region	State
46	BM-11140	Becky Martin	Consumer	67	United States	San Antonio	78207	Central	Texas
47	BM-11575	Brendan Murry	Corporate	22	United States	New York City	10035.0	East	New York
48	BM-11650	Brian Moss	Corporate	48	United States	New York City	10035.0	East	New York
49	BM-11785	Bryan Mills	Consumer	31	United States	Columbus	43229.0	East	Ohio
50	BP-11050	Barry Pond	Corporate	65	United States	Cranston	2920.0	East	Rhode Island
51	BP-11095	Bart Pistole	Corporate	30	United States	Peoria	61604	Central	Illinois
52	BP-11185	Ben Peterman	Corporate	48	United States	Arvada	80004	West	Colorado
53	BP-11290	Beth Paige	Consumer	26	United States	Evanston	60201	Central	Illinois
54	BS-11365	Bill Shonely	Corporate	34	United States	New York City	10011.0	East	New York
55	BS-11380	Bill Stewart	Corporate	53	United States	Skokie	60076	Central	Illinois
56	BS-11665	Brian Stugart	Consumer	21	United States	Philadelphia	19143.0	East	Pennsylvania
57	BS-11755	Bruce Stewart	Consumer	59	United States	Denver	80219	West	Colorado
58	BS-11800	Bryan Spruell	Home Office	33	United States	New York City	10011.0	East	New York
59	BT-11395	Bill Tyler	Corporate	32	United States	Detroit	48227	Central	Michigan
60	BT-11440	Bobby Trafton	Consumer	20	United States	Littleton	80122	West	Colorado
61	BT-11485	Brad Thomas	Home Office	66	United States	Clinton	20735.0	East	Maryland
62	BT-11680	Brian Thompson	Consumer	70	United States	Newark	43055.0	East	Ohio
63	BV-11245	Benjamin Venier	Corporate	46	United States	Des Moines	50315	Central	Iowa
64	BW-11110	Bart Watters	Corporate	50	United States	Milwaukee	53209	Central	Wisconsin
65	CA-12055	Cathy Armstrong	Home Office	67	United States	Houston	77070	Central	Texas
66	CA-12265	Christina Anderson	Consumer	49	United States	Chicago	60610	Central	Illinois
67	CA-12310	Christine Abelman	Corporate	68	United States	Cincinnati	45231.0	East	Ohio
68	CB-12025	Cassandra Brandow	Consumer	43	United States	Hamilton	45011.0	East	Ohio
69	CB-12415	Christy Brittain	Consumer	47	United States	Columbus	43229.0	East	Ohio
70	CC-12220	Chris Cortes	Consumer	18	United States	La Porte	46350	Central	Indiana
71	CC-12430	Chuck Clark	Home Office	66	United States	Columbus	47201	Central	Indiana
72	CC-12475	Cindy Chapman	Consumer	48	United States	New York City	10024.0	East	New York
73	CC-12610	Corey Catlett	Corporate	18	United States	Philadelphia	19134.0	East	Pennsylvania
74	CC-12670	Craig Carreira	Consumer	22	United States	New York City	10009.0	East	New York
75	CD-11980	Carol Darley	Consumer	61	United States	Tyler	75701	Central	Texas

Fig I.4: Data after Transformation

© Daffodil International University

Appendix J: ETL Process Diagram Design

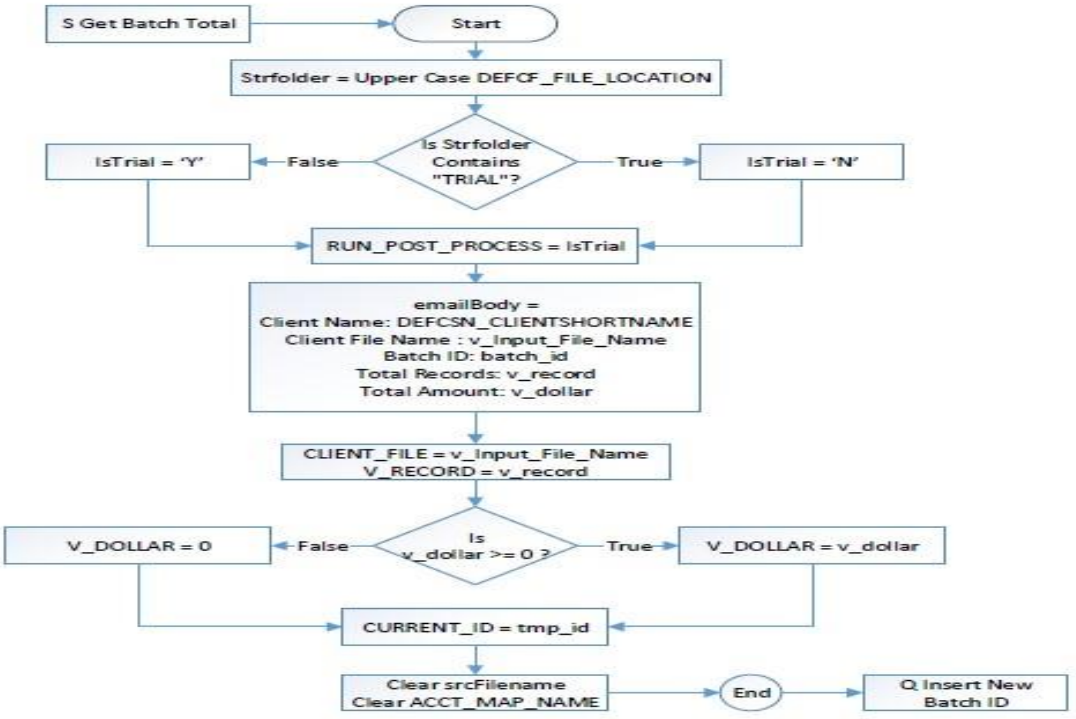


Fig J.1: Process Diagram

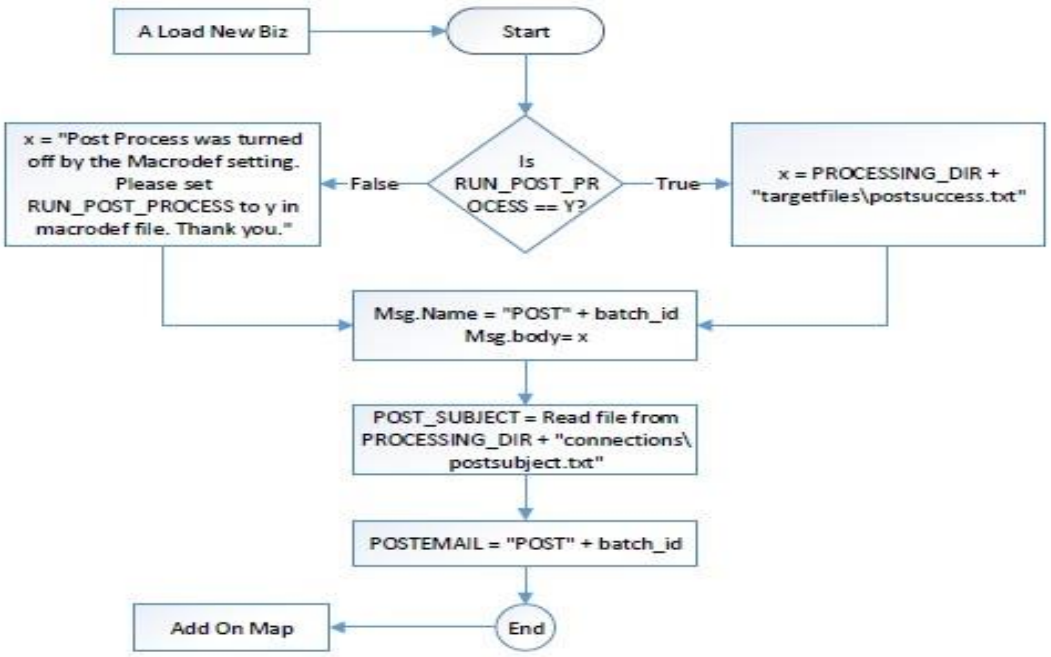


Fig J.2: Process Diagram

© Daffodil International University

Appendix K: ETL Process Documentation

Document Information

Document Title	Pentaho New Business Functional Design Document
Document ID	New_Biz_FDS
Version	1.0
Owner	MD
Author	Nassprotech

Amendment History

Document Version#	Submission Date	Modified By	Modifications
0.1	03/08/2022	Sajid	Initial requirements documentation

Fig K.1: ETL Process Documentation

TABLE OF CONTENTS

New Business Project Diagram 4
1. Start 5
2. S Assign Variables 5
3. Q UDEF Create 6
4. M Get Client Data 7
5. S Increment Client Id 8
6. S Create Folders 9
7. S Prepare Source Files 9
8. S Get Batch Id 10
9. Q Insert Batch Id 11
12. S Increment Map Loop 11
13. Q Update Account 12
14. S Get Batch Total 12
16. Q Insert New Batch Id 13
17. Q Update PLC Result 13

Fig K.2: ETL Process Documentation

 © Daffodil International University

1. Start

For starting step, we created a job named newbiz. In the starting step, we have to declare several process variables and also initialize their value. We have taken a JavaScript file. On that file, we declare several process variables. Then we have taken a log file to print the value of the variables.

We have to check whether the exelog file exists or not. For that, we used a delete file step. Where we check if the exelog file exists or not. If exist then we will delete the file. Then we proceed to the next step.

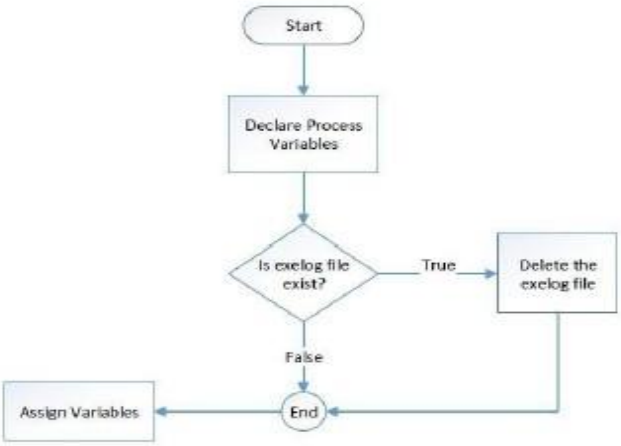


Fig K.3: ETL Process Documentation

16. Q Insert New Batch Id

In this step, we execute an SQL query from a SQL file to UPDATE ETL.PTH_BATCHID_LKP Table and also insert data into ETL.PTH_BATCHID_LKP table. For this task, we used a SQL step inside the main job and on that step, we select the SQL file for executing.

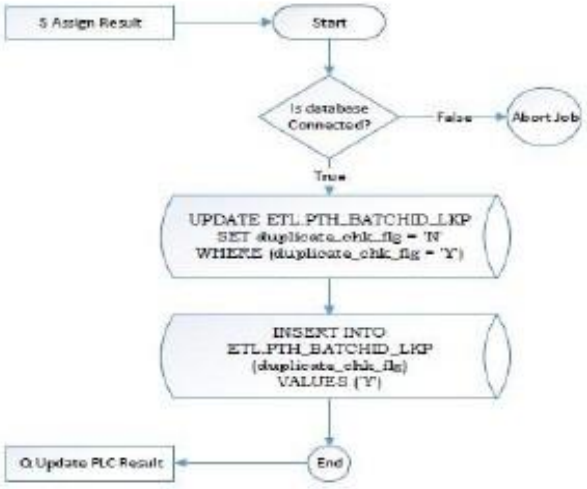


Fig K.3: ETL Process Documentation

© Daffodil International University

Appendix L: SQL Stored Procedure

Windham PROFESSIONALS

Overview

Objectives

- Identify consumers with a IVRPAY consumer tag that have no accounts with an IVR CLIENT PORTAL PYMNTS UDP in a Submitted to Client Status

Scope

Requirements

Identify consumers that have an IVRPAY consumer tag

Review ALL accounts under the consumer and identify all IVR CLIENT PORTAL PYMTS UDP'S

If there is no IVR CLIENT PORTAL PYMTS UDP with a status of Submitted to Client-remove the IVRPAY consumer tag

- If any of the UDP'S has a submitted to client status-DO NOT remove the consumer tag

Fig L.1: Client Requirement Document

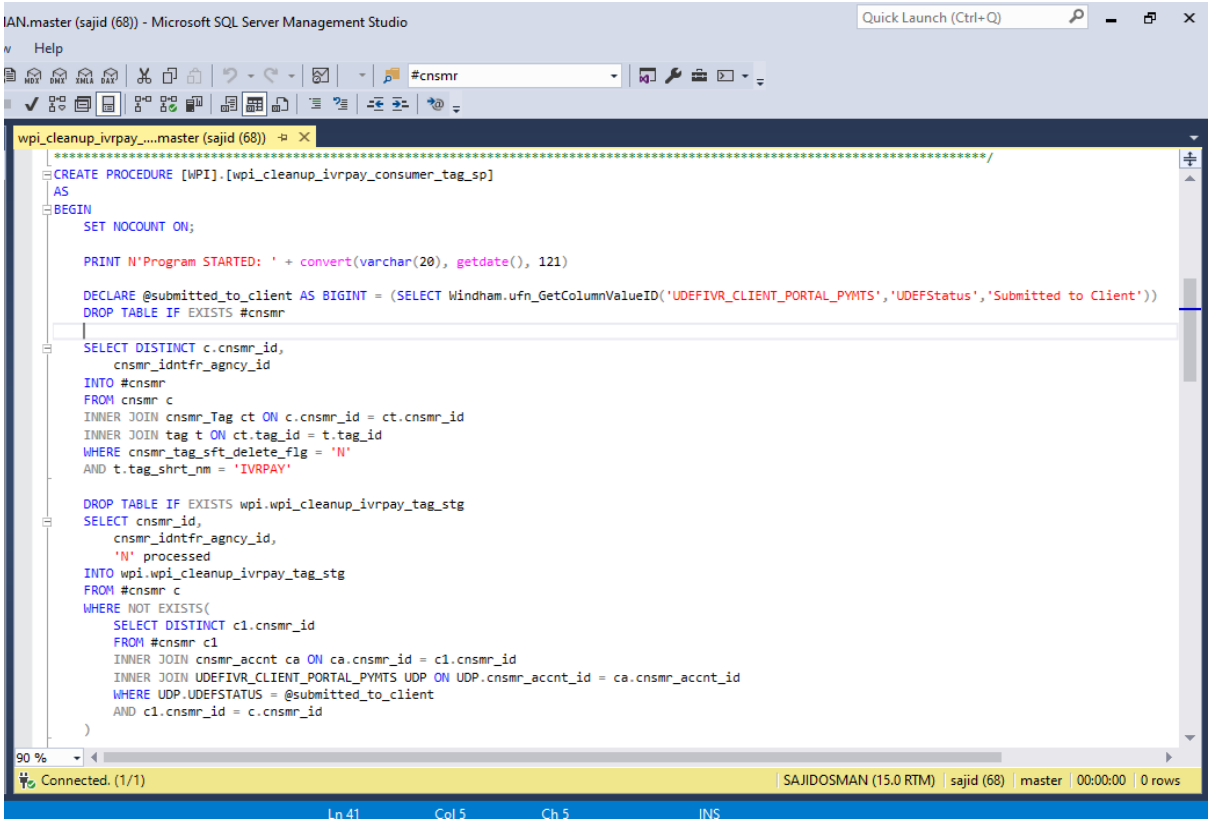


Fig L.2: SQL Stored Procedure

© Daffodil International University

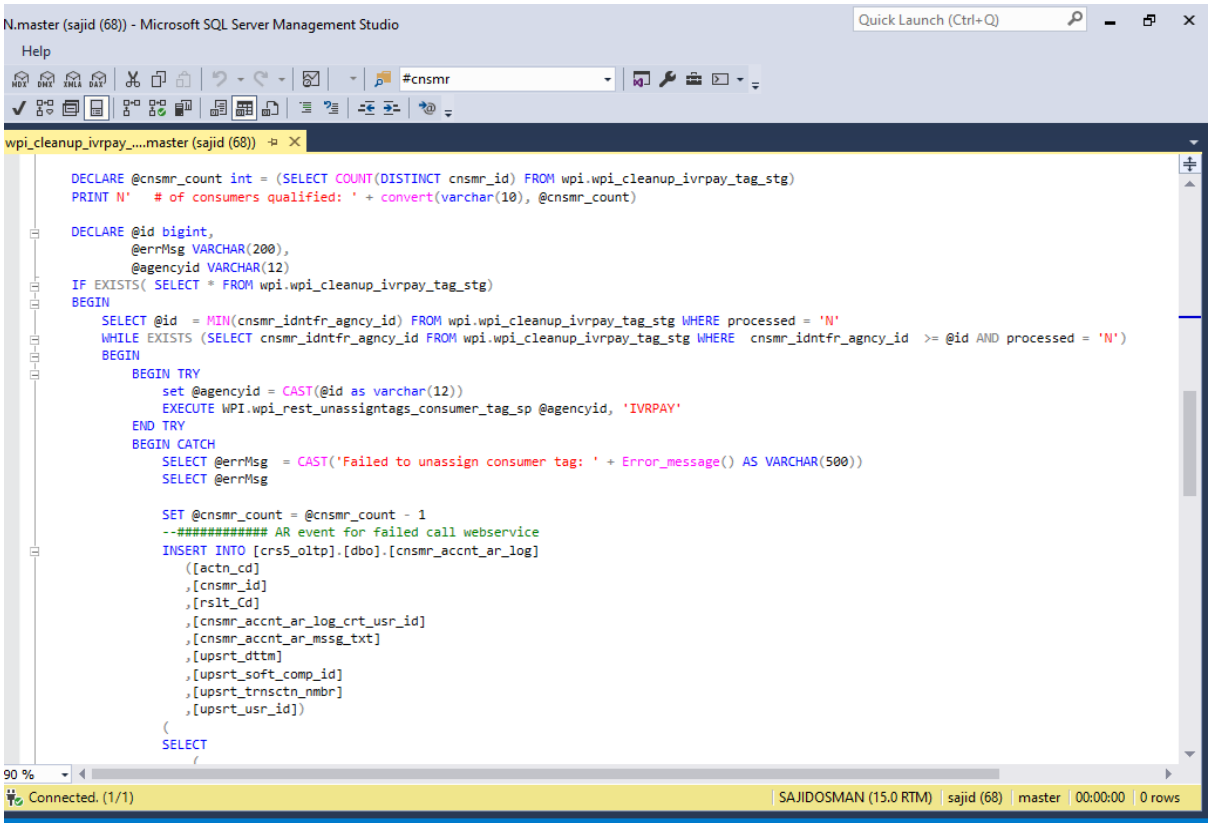


Fig L.3: SQL Stored Procedure

Appendix M: FICO Debt Manager

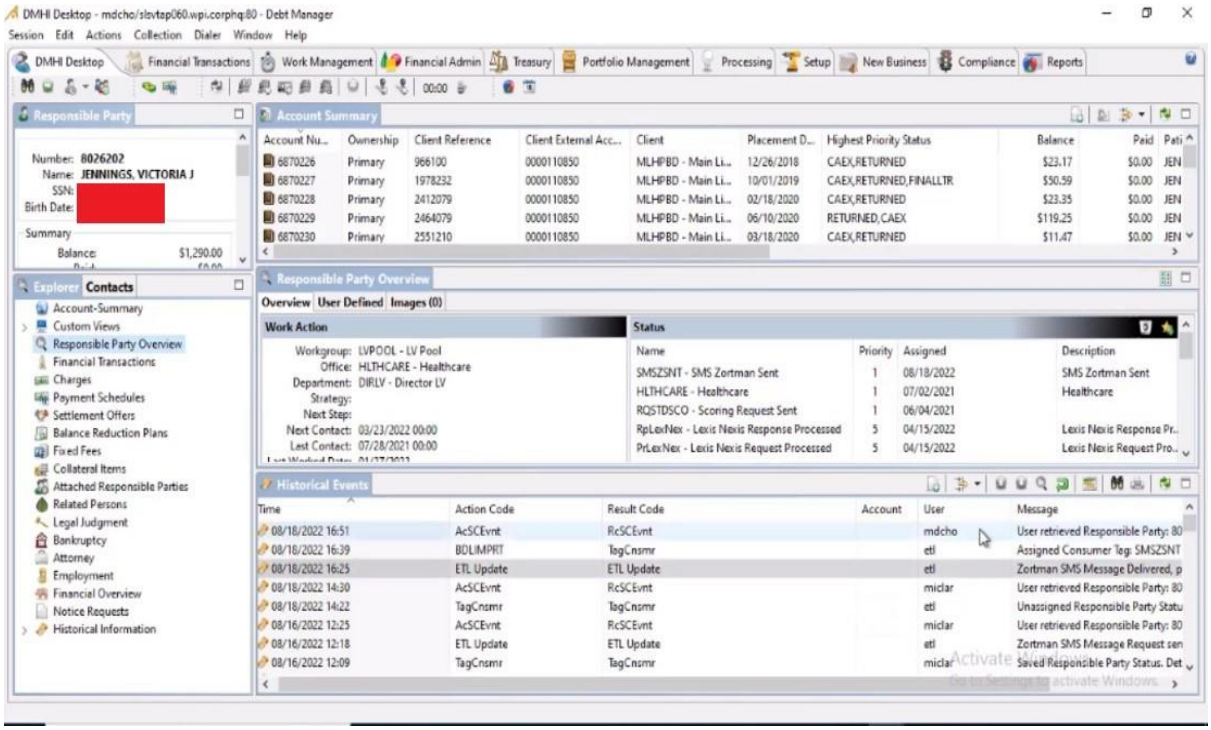


Fig M.1: Fico Debt Manager

© Daffodil International University

REFERENCES

1) Documentation of Pervasive ETL software: https://www.bitool.net/software/pervasive-etl.html
2) Documentation of Pentaho Data Integration tool: https://help.hitachivantara.com/Documentation/Pentaho/8.2/Products/Data_Integration
3) Fico Debt Manager Software's About Us Page: https://www.fico.com/en/about-us
4) Facebook page of Nassprotech Consulting: https://m.facebook.com/nassprotech/posts
5) Nassprotech Consulting Company details: https://www.nassprotech.com/about/
6) The working style of Nassprotech Consulting: https://www.nassprotech.com/how-we-work/
7) Services of Nassprotech Consulting: https://www.nassprotech.com/business-consulting-services/

© Daffodil International University