Data Analysis & Automation : An Overview

Archit Srivastava, Veena K Divya and Arvind Bhagwath

July 14, 2021

Data Analysis and Automation - An Overview

Archit Srivastava, Prof. Veena Divya K., and Mr. Arvind Bhagwath

Abstract. Automation has helped humanity grow and achieve many things.This is the core ideology behind this review chapter. The intends to reduce the manual work required in making the daily Dashboard Report by using the knowledge of IT Automation, specifically data handling and analysis process. The initial process was to work on the front end to develop the required visuals and make them user friendly while making progress in the back end development which is the core aspect when comes to the Automation aspect of the project. This paves way to give us an overview about what the actual dashboard automation steps look like. The software stack used to in this review chapter are majorly related to data handling and visualization. Tools like *MS SQL Server Management Studio* and *MS Power BI* are a perfect amalgamation required to create professional looking dashboard with cutting edge back end to drive the automation. The chapter elegantly gives the overview about *SSIS packages* and how to connect them to *Power BI* via an SQL data source connection for regular data refresh, thus reducing the manual work required and simultaneously helping the clients access a fully automated and interactive dashboard for their daily needs. This review chapter was inspired from the core subjects of *Industrial Automation* and *Virtual Instrumentation* and extending the knowledge of the same to the field of Data Analysis. This chapter is an elegant amalgamation of vastly different fields of Data Science and Analysis and Industrial Instrumentation and takes into account the knowledge from both fields to develop the knowledge about IT Automation.

Keywords: SQL, SSIS, Power BI, Automation

1 Introduction

Data Analysis and Engineering has been one of the most anticipated areas of scientific development which is seen by humanity. With increasing demand and supply of goods world wide, thus increasing the status of living, there has been an ever increasing demand for Data Related positions across the globe. Data Analysis forms a major role when it comes to connecting the worlds of complex mathematical analysis and common people. [1] Effective Data Analysis practices can help users effectively develop an insight about what is seen from the study. Companies like Hewlett Packard Enterprise are heavily driven by data and have many in depth studies being created based on the insights developed. Daily

report generation is the backbone of any firm that needs to stay ahead in the market. These reports help the Executive authorities to skim through the daily Business made and then help them to settle upon major business decisions, which drives the economy of the Company.[1]   Many reports that are generated on a regular basis, are responsible to tell the end users, about vital information like, which products were not sold and have been kept on Hold for what tenure and what is the valuation of them. The delay in selling of products may be because of the following points and thus this effects the supply chain as a whole:

– Delayed Supply Chain Response
– Natural Calamities at the shipping area
– Delay in production
– Reduction in consumption of the end product

These types of parameters in any organization, are a crucial part to assess where the actual Business Unit is incurring losses and then eventually reducing it to involve less loss attained.[2] These daily reports form the backbone of the organization and it helps such ventures to rise and improve on their business on a daily basis. The project intends to look into the depth of the processes of automation which were inculcated from the core subjects in the curriculum. The subject of Industrial Automation helped develop important insight about what Hardware and Software Automation referred to and the project intensively researches about the same aspect of Software automation and it's collective effect over the whole Supply Chain. The project is inspired from the area of **Industrial IT Automation**, an amalgamation of Industrial and IT Automation. The majority of work done in the project is based on the IT Automation aspect dealing with the areas of Data Extraction, Transformation and Loading process and then automating the process that will have an end effect over the Industrial Automation aspect, thus optimizing the Manufacturing of goods and services at the producer level.

2   Industrial Automation

In the present processing plants and shop floors, modern computerization is all over the place and it is hard to envision a creation line without mechanization. Mechanical automation utilizes control frameworks and hardware, like PC programming and robots, to perform errands that were generally done physically. These frameworks work mechanical hardware consequently, fundamentally decreasing the degree of administrator inclusion and oversight required. Mechanization frameworks commonly comprise of input circles and tactile projects that naturally change working conditions to meet the ideal qualities dependent on continuous information. The advances utilized in modern computerization have progressed fundamentally as of late because of digitalization of assembling, particularly with the Mechanical Web of Things (IIoT) – opening up an ever increasing number of chances for organizations to exploit mechanization arrangements.

3   IT Automation

3.1   Introduction

IT automation is the way toward making programming and frameworks to supplant repeatable cycles and diminish manual mediation. It speeds up the conveyance of IT framework and applications via robotizing manual cycles that recently required a human touch. With IT automation, programming is utilized to set up and rehash directions, cycles, or strategies that save time and let loose IT staff for more essential work. With the ascent of virtualized organizations and cloud benefits that require quick, complex provisioning, automation is an irreplaceable system for aiding IT groups convey administrations with improved speed, consistency, and security.

IT automation is an amazing asset that can scale a business, give huge expense reserve funds, and permit IT staff to zero in on vital instead of managerial work. A wide scope of server farm and cloud tasks can be mechanized, bringing about quicker activities. On account of automation, IT conditions can scale all the more rapidly with less mistakes and are more receptive to business needs. A completely mechanized climate can diminish the opportunity to conveyance for creation prepared assets from weeks to not exactly a day.

3.2   Working of IT Automation

IT automation programming can play out a scope of IT assignments and cycles, from easy to complex. For instance, automation can be utilized to make systems administration or security formats and outlines and to design applications and arrangement creation prepared framework.

Late IT automation patterns incorporate the utilization of man-made brainpower and AI (two diverse however related innovations) to make more brilliant cycles that arrangement with more erratic circumstances. These advancements are as yet in their beginning phases, yet they could permit robotized cycles to learn and improve as they go. Automation apparatuses themselves are additionally getting all the more impressive, permitting IT staff to fabricate work processes all the more rapidly.

3.3   Need of IT Automation

IT automation is valuable to supplant tedious undertakings and permit IT staff to stay aware of the expanding scale and intricacy of IT activities and cloud foundation. In a cutting edge IT climate, the speed and size of administrations are a lot for even a huge and committed group to oversee. IT automation permits groups to work in a setting where it's normal to have to (for instance) set up and arrange a huge number of workers.

4   Data Integration and Analysis

Data analysis is a cycle of examining, purifying, transforming, and demonstrating data fully intent on finding helpful data, illuminating ends, and supporting dynamic. Data analysis has numerous aspects and approaches, incorporating assorted strategies under an assortment of names, and is utilized in various business, science, and sociology spaces. In the present business world, data analysis assumes a part in settling on choices more logical and assisting organizations with working viably. [14]

Data combination is a forerunner to data analysis, and data analysis is firmly connected to data perception and data scattering.ETL is a kind of data combination measure alluding to three unmistakable yet interrelated advances (Extract, Transform and Load) and is utilized to blend data from numerous sources commonly to construct a Data Distribution center, Data Hub, or Data Lake.[15]The most widely recognized misstep and confusion made when planning and building an ETL arrangement is hopping into purchasing new instruments and composing code prior to having a thorough comprehension of business necessities/needs.

There are some principal things that ought to be remembered prior to pushing ahead with carrying out an ETL arrangement and stream.

Need of ETL It is fundamental to appropriately arrange and get ready data to load it in your preferred data stockpiling arrangement.[13] The triple mix of ETL gives urgent capacities that are ordinarily joined into a solitary application or set-up of devices that assistance in the accompanying regions:

— Offers profound verifiable setting for business.
— Upgrades Business Insight answers for dynamic.
— Empowers setting and data accumulations with the goal that business can produce higher income or potentially set aside cash.
— Empowers a typical data archive.
— Permits check of data transformation, accumulation and estimations rules.
— Permits test data correlation among source and target framework.
— Assists with improving efficiency as it systematizes and reuses without extra specialized abilities.

An essential ETL cycle can be arranged in the underneath stages:

— Data Extraction
— Data Cleaning
— Transformation
— Load

5   Data Automation Tools

5.1   Power BI

Microsoft Power BI is a business intelligence stage that furnishes nontechnical business clients with devices for conglomerating, dissecting, envisioning and

sharing data.[1] Power BI's UI is genuinely instinctive for clients acquainted with Dominate and its profound incorporation with other Microsoft items makes it an extremely flexible self-administration device that requires minimal forthright preparing.

A free form of Power BI is proposed for little to average size business proprietors; an expert variant called Power BI In addition to is accessible for a month to month membership charge. Clients can download an application for Windows 10, called Power BI Desktop, and local mobile applications for Windows, Android and iOS gadgets. There is likewise Power BI Report Worker for organizations that should keep up their data and reports on premises.[5] That form of Power BI requires an uncommon variant of the desktop application - suitably called Power BI Desktop for Power BI Report Worker.

Microsoft Power BI is utilized to discover bits of knowledge inside an association's data. Power BI can help associate unique data sets, transform and clean the data into a data demonstrate and make outlines or charts to give visuals of the data. The entirety of this can be imparted to other Power BI clients inside the association.

The data models made from Power BI can be utilized in a few different ways for associations, including recounting stories through graphs and data representations and inspecting "consider the possibility that" situations inside the data. Power BI reports can likewise respond to inquiries progressively and assist with anticipating to ensure divisions meet business measurements. [5]

Power BI can likewise give leader dashboards to heads or administrators, giving administration more understanding into how divisions are getting along.

5.2   SQL Server Managemet Studio

Microsoft SQL Worker Management is a high level improvement climate that empowers us to design, oversee and administrate SQL Worker database motors.[6] SSMS is extremely mainstream and broadly utilized by the database engineers and directors on account of the accompanying benefits:

5.3   Microsoft Visual Studio

The Visual Studio coordinated improvement climate is an inventive take off platform that you can use to alter, troubleshoot, and assemble code, and afterward distribute an application. A coordinated advancement climate (IDE) is an element rich program that can be utilized for some parts of programming improvement. Far beyond the standard supervisor and debugger that most IDEs give, Visual Studio incorporates compilers, code consummation apparatuses, graphical creators, and a lot more highlights to facilitate the product advancement measure.

Visual Studio is an Incorporated Advancement Environment(IDE) created by Microsoft to foster GUI(Graphical UI), console, Web applications, web applications, mobile applications, cloud, and web administrations, and so on With the

assistance of this IDE, you can make oversaw code just as local code.[7] It utilizes the different foundation of Microsoft programming advancement programming like Windows store, Microsoft Silverlight, and Windows Programming interface, and so forth. [8] It's anything but a language-explicit IDE as you can utilize this to compose code in C#, C++, VB(Visual Essential), Python, JavaScript, and a lot more dialects. It offers help for a day and a half programming dialects. It is accessible for Windows just as for macOS.

Advancement of Visual Studio: The main adaptation of VS(Visual Studio) was delivered in 1997, named as Visual Studio 97 having variant number 5.0. The most recent variant of Visual Studio is 15.0 which was delivered on Walk 7, 2017. It is additionally named as Visual Studio 2017.[14] The upheld .Net System Adaptations in most recent Visual Studio is 3.5 to 4.7. Java was upheld in old variants of Visual Studio yet in the furthest down the line adaptation doesn't offer any help for Java language.

6   Frontend Development

We got an indepth knowledge about the tools used during the entirety of project, namely: *Power BI, SSMS and Visual Studio.* The Project Development was specifically consisting of 2 parts:
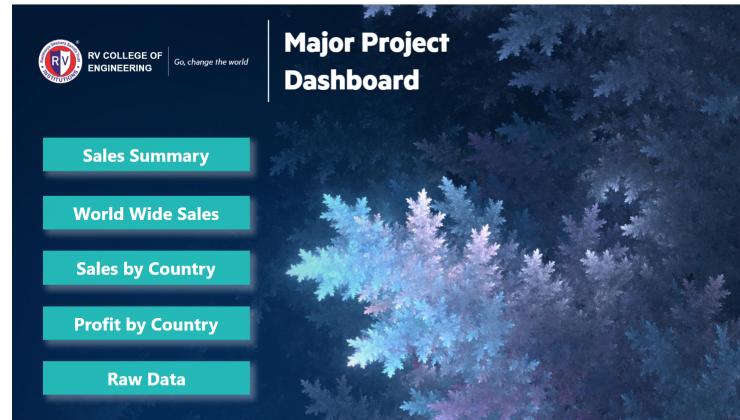
1. Frontend Development
2. Backend Development

This chapter will give an insight on what the frontend visuals look like in the project and what is inferred from the charts and different views inculcated in the Dashboard.

The frontend development, which is the major part of the project was carried out extensively in Microsoft Power BI. [20] The frontend development was crucial as it directly interacted with the end user and gave insight about what and how the data interprets to them. Frontend dashboards as previously mentioned are interactive and change values in accordance with the data being interacted with and the filters applied.

6.1   Landing Page
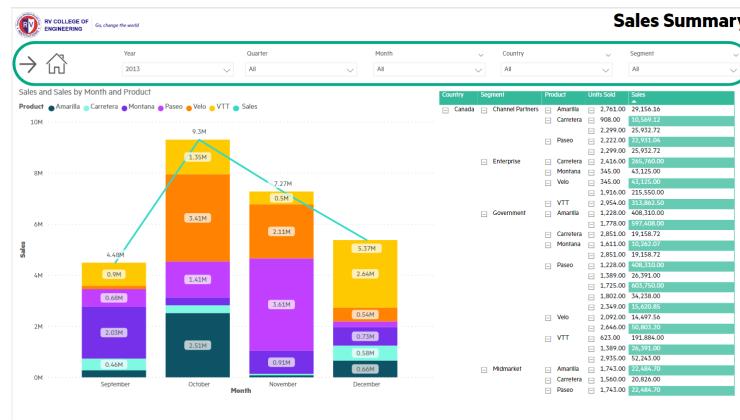
The landing page of the dashboard makes navigation easier for end user. This is the most attractive looking visual of the dashboard to attract the customer to interact more with the dashboard. Every dashboard requires a landing page to give vital information about the dashboard and the visuals embedded inside it.[17] It has buttons for navigation to different visuals for giving an ease of access to the user to see the visualization. It needs to be attractive for customer appeal to draw the attention and make it more interactive and user friendly.

Fig. 1. Landing Page

6.2 Sales Summary Page

The sales summary page is one of the major page which describes about how the sales have been in the particular duration. This page provides the summary view along with a clustered bar graph and table with the following column:

1. Country
2. Segment
3. Product
4. Units Sold
5. Sales

Fig. 2. Sales Summary Page

The bar graph gives us an insight about how the sales of the particular product has been performing throughout the year selected. The Summary View is Important when it comes to getting a quick insight about the data being monitored.

6.3   World Wide Sales

This visual is a necessary visual when it comes to executive level customer viewing the dashboard. The World Wide sales has a Pie Chart which depicts the overview of the data according to per country and tells what part of the pie is each country corresponding to. The second chart is a clustered bar graph which shows sales of each Holds by country and product.
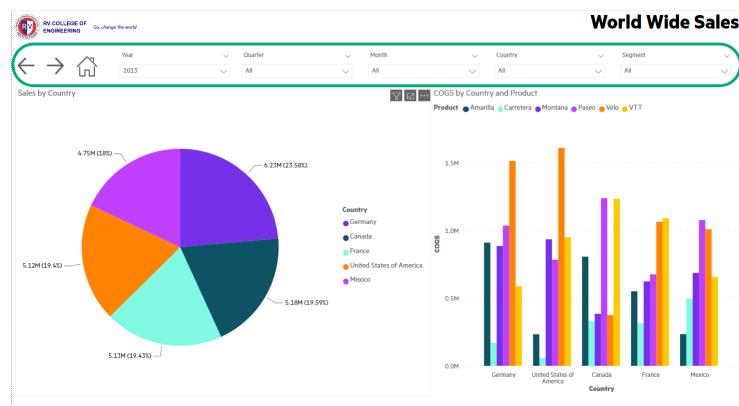
Fig. 3. World Wide Sales

Getting the world Wide picture about the business helps getting useful insights about what sales has performed and how it will make an impact on the decisions that will influence all the business decisions.

6.4   Sales by Country

The Sales by Country visual is an essential visual when the end user wants to get a more deeper insight about how the sales and profits have performed for each country. The values thus been shown in the stacked area chart gives us the sales per month per country in the year 2014.

The stacked column chart tells us about the profits made by each country per segment. These vital in formations help the customer decide which country should be kept at priority to increase sales and improve the logistics to reduce the holds.
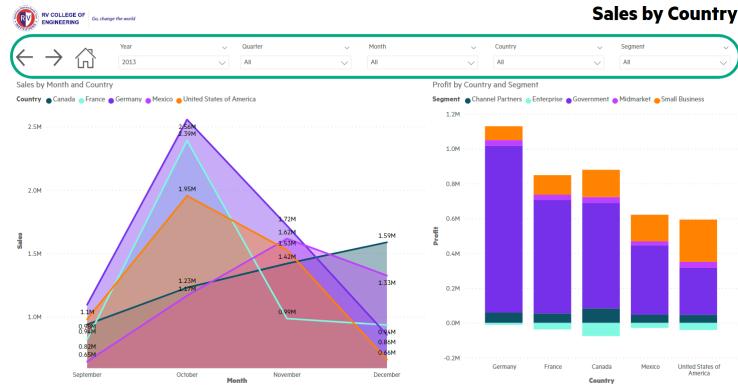
Fig. 4. Sales by Country

6.5   Porfit by Country

This visual tells us that the country with the maximum profit will have the least holds because it has performed well in the years. The clustered column chart on the left tells us which country made the profit in the year of 2014, and the tables on the right tells us about the profits made per country and product in a tabular format in accordance to the years 2013 and 2014 respectively.This visual is deeper layer visualization of the Sales by Country Page and helps to get insights specifically focused on profits made.
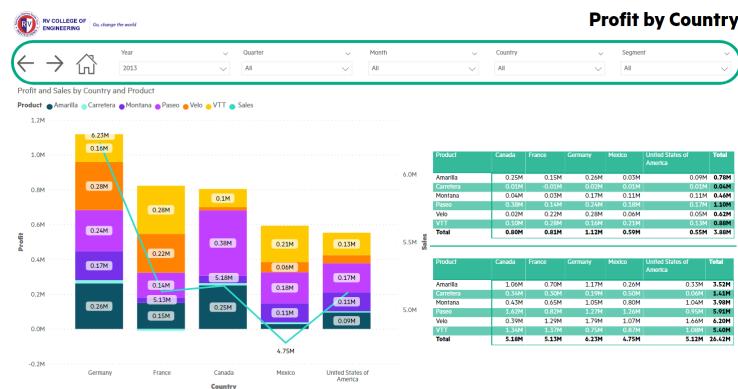
Fig. 5. Profit by Country

6.6   Raw Data

This view is the last view in almost all the dashboards when they go into development. This view helps to get information about how much data and what type of it, the dashboard is consisting off. This view helps to get the customer familiar with the data and raise few changes in the dashboard for further insights which they find out statistically. This view is helpful for developers to get an insight

Fig. 6. Raw Data

about the actual raw data and help them mimic or develop new insights on for the dashboard.

The chapter concludes with the explanation of all the visuals in the front end dashboard developed and paving a way for the backend development of the project which will be covered in the next chapter.

7   Backend Development

The backend development of the project specifically focused with the work being done in two specific environments.

1. SQL Server Management Studio
2. Visual Code

The above two softwares are exclusive Microsoft software and are used to develop and integrate the SSIS package required for the process of automation. The details about the methods used and performed in the backend endgineering process is discussed below.

7.1 Microsoft SQL Server Management Studio

The SSMS is a Miscrosoft SQL Service provided to create and maintain data warehouses for easy data ETL. [12] For the project the data warehouse was created in SSMS and the Staging Table, the Main Table and all the stored procedures were mapped in this software.
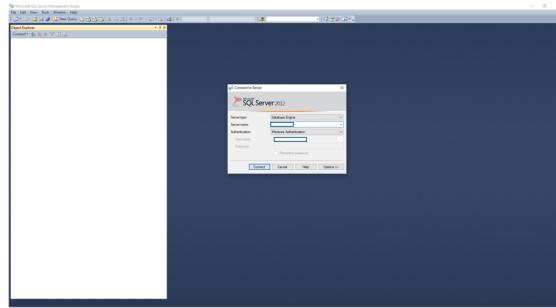
Fig. 7. SSMS Startup Menu

The visual above shows us the window when we open the SSMS.[11] It takes in the values of Server Name and the login credential to connect to the network to check the data.
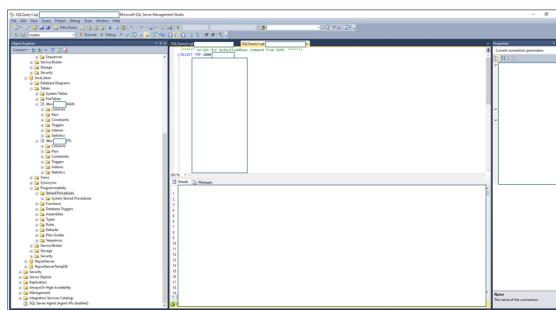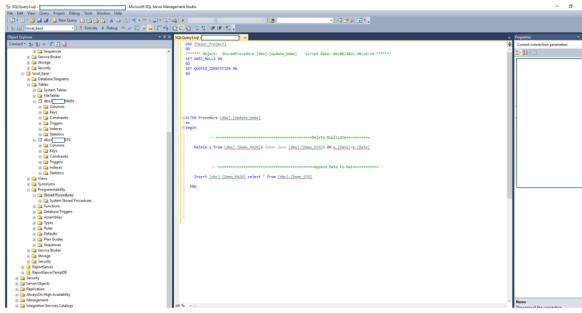
Fig. 8. Top 1000 rows

The visual described above gives us the view of the top 1000 data values and columns of the table that will be finally injected in the dashboard.
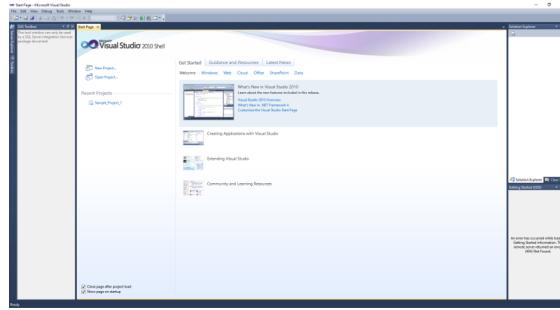
The third visual above tells us description about the stored procedure created to clean the data and the table specifically deleting any kind of duplicates.

The steps and snaps explained above were the major steps described briefly about the data Transformation and Loading process

Fig. 9. Stored Procedure

7.2   Visual Studio

Microsoft Visual Studio is a Graphical Interface based programming environment and is similar to *NI LabView* used for Virtual Instrumentation. The programming environment helps to create SSIS packages which are used to automate various tasks related to the development of the dashboards and their daily refreshes required. For the development of the daily dashboard, the following steps were taken.

Fig. 10. Visual Studio Startup Menu

The visual above tells us about how the visual studio pane looks like when it starts up. Talking about ETL, there are a lot of ETL devices that give valuable highlights to help in all means of data move and ETL. Microsoft additionally delivered SQL Worker Coordination Administrations as a piece of the Microsoft SQL Worker item. Mix Administrations assists with diminishing the measure of work utilizing segments and a planned design for data move.[8] At present, Combination Administrations is one of the main ETL apparatuses in the market due to its power and ability to perform data move ETL represents Extract, Transform, and Load. Extract is the way toward getting data from diverse data sources, Transform is the progression where we apply transformation to data,

such as conglomeration, get together with other datasets, applying rules, part or union outcome set, etc.
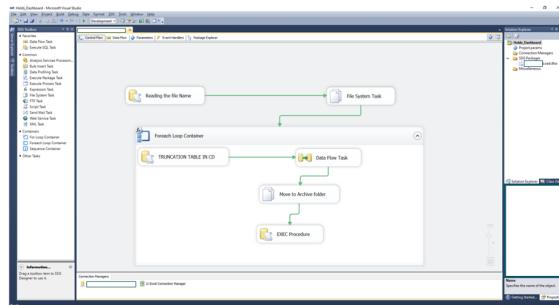
Fig. 11. The SSIS Package

The visual above shows us the actual SSIS package developed to automate the dashboarding process and the daily refresh involved.
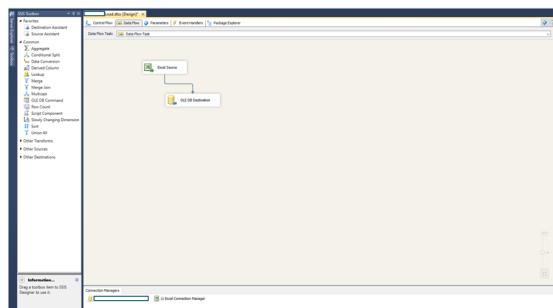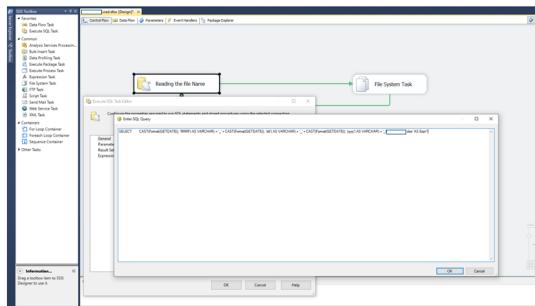
Fig. 12. The Data Flow Connection Model

This snap shows the data flow connection made in the data flow model of the SSIS package. This is the module that connects various different sources of data like Excel in this case to one common OLE DB Destination so that data integrity is maintained. SSIS packages have a lot to offer and provide necessary automation to the ETL steps. SSIS has a motor that is introduced as a piece of the Microsoft SQL Worker establishment media (in the event that you pick the combination administrations during establishment). [7]

Fig. 13. SQL Query Written

This visual shows us the SQL query written to automatically tell us about the date of extraction of the data in a specified format.

Fig. 14. Raw Data File

There are a lot of types of Raw Data available, and they have columns in them to describe their nature. The above screenshot gives us an overview about a sample dataset created an the fields included in the excel are:

1. **Region:** This is the region of operation in which the data is being used
2. **Country:** This is the country in which the sales have been made.
3. **Party:** This is the major distributor in the country which procures the product.
4. **Customer:** This is the secondary distributor to which the Party sells the product.

5. **Shipped To:** This is the final end user which procures the product.
6. **Number ID:** This is the ID provided to each transaction made by party.
7. **Shipped ID:** This is the ID provided when the product is shipped.
8. **Sales:** This is the amount of sales made in the on the date.
9. **Profit:** This is the profit made on the sales.
10. **Holds:** The amount of data on holds in the warehouses.
11. **Date:** The current date of report generation.
12. **Date Received:** This is the date when the order is recieved by the party.

The SSIS package development was necessarily the most important step of the project as it automates the manual data entry process to the dashboard and the refreshes of the data in the dashboard. [15]

8 Results

After getting insights about the processes involved and how the integration is done to the frontend with the backend. The results can be seen below:

8.1 Frontend Result

The main observation that we make in frontend is in the dates of the data loaded. Below we can see the Data initially is there of 2013. We see that in 2013 from the summary page is that the sales made in October 2013 was significantly more than September and over the period of the next two months the sales of the products collectively dropped but was above sales made in September.
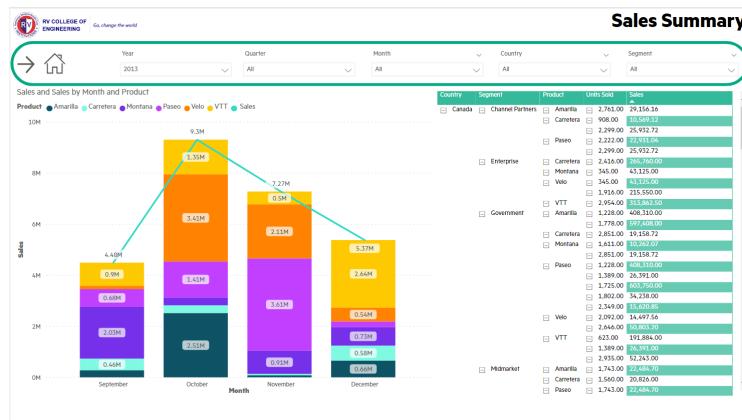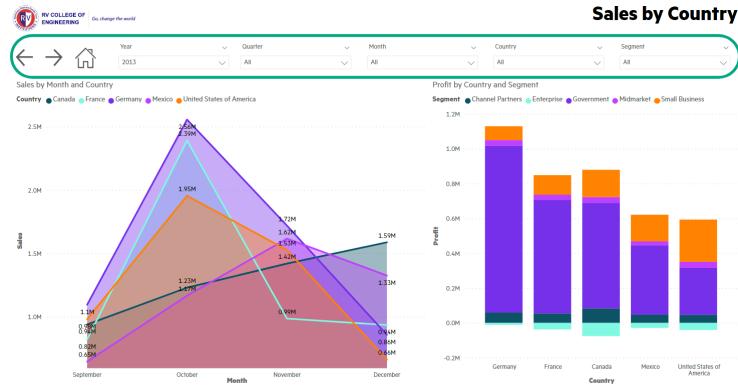
Fig. 15. Sales Summary in 2013

The results that we see above are the data that have been loaded from the file containing records of 2013. The table on the right gives us an insight about how the sales of the organization have performed over the period.

Fig. 16. Sales by Country in 2013

The visualization above tells us about how the sales in each country have performed over the period of September to December in the named 5 countries, Canada, France, Germany. Mexico and USA. The specific above visual tells us that during the mentioned period of 2013, Germany was the country with the most sales made, and this has made the most impact on business when compared to the other four big nations. This insight helps to take decisions regarding the business to be made in the other four country or USA to be specific to increase sales in the nation.

Fig. 17. World Wide Sales in 2013

This third visual tells us more about the sales made in 2013 across the globe. This visual gives an in-depth knowledge about what was the sales made across the globe. We can see that the sales distribution of the firm was almost equal in

all the 5 countries and neared about in the range of 18% - 23.8%. This was almost equal distribution all over the countries in 2013. The visual tells us details about COGS per product per country. We see that the product Velo performed the most prominently in USA and Velo remained the top or the second performer in almost all countries of operation.
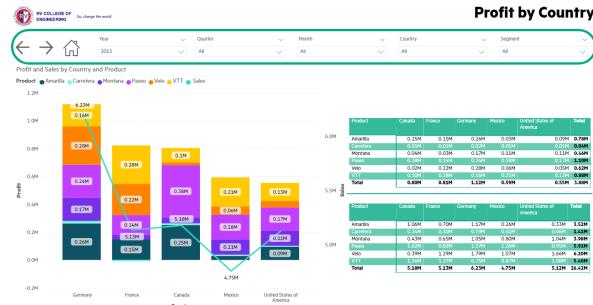
Fig. 18. Profit by Country in 2013

Profits are the most important factor for any company made and this is the most primary thing that is required. The visualization as explained in the previous chapters, this page tells us more about the profits made by per country over the period of 4 months of September, October, November and December of 2013. We infer from this stacked graph is that Germany was the major player in total profits made by the company. Germany made a profit of $1.12M which was almost 50% more than USA and Mexico and almost 30% more profit from France and Canada.
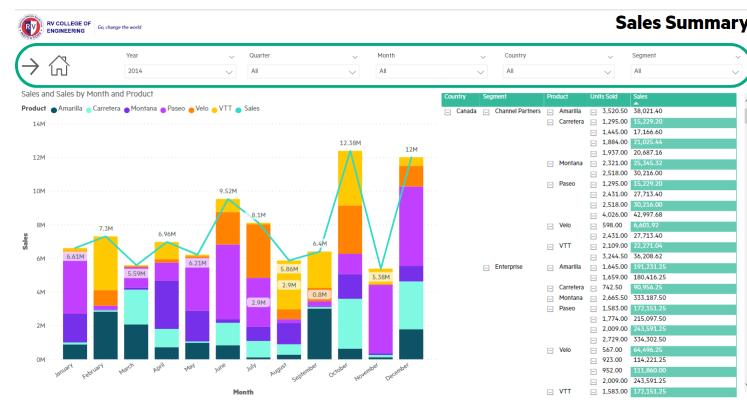
Fig. 19. Sales Summary in 2014 after data refresh

18   Archit Srivastava, Prof. Veena Divya K., and Mr. Arvind Bhagwath

The inflow of huge data in such organization leads to a lot of problems. The automation performed in the project helps to remove such disadvantages in the data handling process. Now when the SSIS package is executed, the data of 2014 is loaded in dashboard and the visualizations get updated based on the same. This is the automation that is performed in the project. The step of ETL related to the project is automated and below we see the results after SSIS execution.

Now when the SSIS package is executed, we see that the year changes from 2013 to 2014 and the corresponding data is changed in the visualization. The above visual when we compare it to the Sales Summary Page in 2013, in 2014 we find that maximum sales were made in October 2014 and the sales across the year was almost stagnant.

This again suggests that when the sales were compared from 2013 in the same duration, it has significantly increased over the period. The sales made by the company in October 2013 was $9.3M and increased to $21.6M in 2014, a significant 232% increase in sales made.

This was only possible due to the insights made by the executives based on the dashboard and the business decision. We see that there was a significant increase in sales overall and shows a growth seen overall over the year.
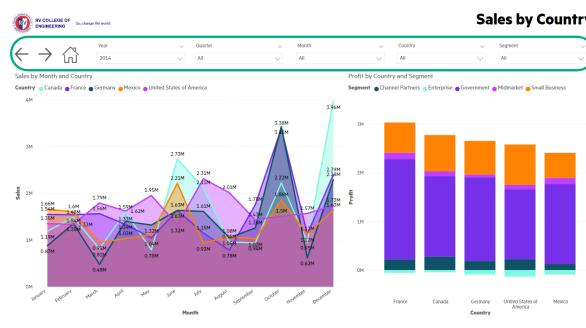
Fig. 20. Sales by Country in 2014 after data refresh

The page above which shows us the updated results after the Automated ETL procedure was executed, it tells us about the sales made by each country inn 2014. The graph in the page shows us the rise in sales that was made in 2014 with respect to 2013. We can infer from the graph that in 2014, France was the country with majority of sales made. The sales increase in France from $1.1M in 2013 to $3.8 M in 2014.

This suggests that the insights developed from the visualization of 2013 dashboard, helped the executives to increase the sales in France by nearly 150% over the next 3 Quarters. This means that the executives were benefitted by the insights made and made the proper decisions related to the sales in the country.

This help significantly increase the sales in France. With the other countries, we can see that sales in USA also increased in 2014 when compared to sales made

in 2013 which kept USA at the 4th position in 2014 than 5th position in 2013. This directly infers that the executive board was able to make proper decisions related to the sales made in USA over 2014 which significantly increased the overall revenue of the products sold.

Fig. 21. Worldwide Sales in 2014 after data refresh

Getting insight about the sales made across the globe is the most important aspect to any business. The visual above shows us that almost there was equal distribution which was like previous year. The sales distribution in 2014 ranged from $16M - $20M. The COGS distribution across the country shows us that Pasos was the most popular product being sold in 2014. This infers that in Mexico, the COGS for Pasos reached a global maximum across the organization with the total valuation of $5.7M.


Fig. 22. Profits made per country in 2014 after data refresh

Profits, the most important parameter which is of most importance for any company. We see in the visual above we infer that France was the most significant player in the financial year. The profits made by France was $0.6M more than the sales made. In 2013 the profits made by France was $0.8M and it increased by almost 375% to $3M in 2014. This was the result because of the insights and business decisions made over because of the frontend designed for the data insight development for the executive level of people.

8.2   Backend Result

The result of backend includes the successful execution of the SSIS package and the copying of the Excel file from the Input Folder to the Archive Folder. Below we can see the successful execution of the SSIS package, which leads to the copy of the target file from the input folder to the Archive folder.
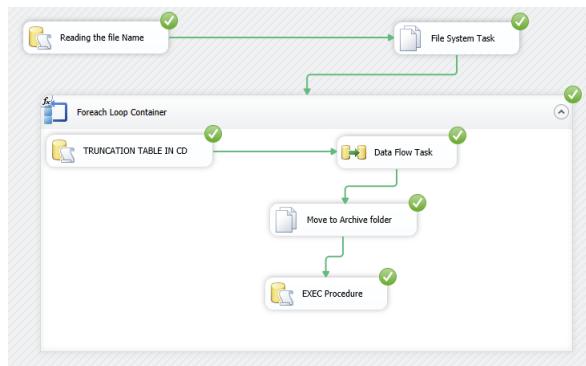
Fig. 23. Successful Execution of the SSIS package

The SSIS package developed for the automation step performs the following steps:

1. Reading the File Name: Reading the file name from the input file location in a particular format.
2. File System Task: This step tells the SSIS package to copy the file read, and then intake the contents of the excel file.
3. Truncation Table in CD: This is a step specifically done in SSMS. The contents of the Staging Table are cleared by this step.
4. Data Flow Task: This data flow task is responsible for the contents of the excel data being taken from step 2 into the now empty Staging table.
5. Move to Archive: This steps moves the file from the input folder into an archive folder, thus creating a folder which contains all the archive data of the input excel data formats for the dashboard.

6. EXEC Procedure: After the staging table gets populated from the excel data, the data from the staging table is then copied and pasted into a Main Table which acts as an archive to the historic data thus keeping in the data from different days.

We can see below the results before execution of the SSIS package where the file is there in the Input Folder.
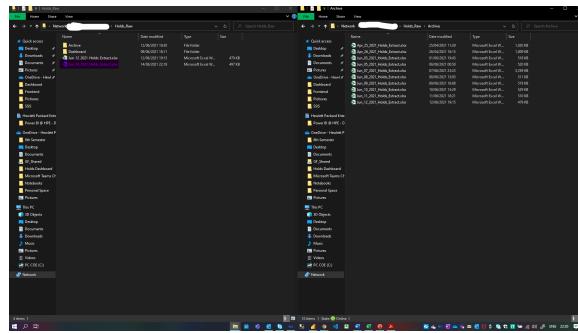
Fig. 24. File in the Input Folder before Execution

Next when we execute the SSIS package the EXCEL file then moves to the Archive folder and all its data is pushed to the table in the dashboard. This is the execution step that we see on surface happening. The movement of the data file from input folder to the archive folder ensures that the SSIS package was executed Successfully. This also inferred that the data in the excel file was taken as an input from the data file and then populated into the staging table. The SSIS package then pushes the data in staging table to the Main Table there in the data base created in SSMS. This then is used to directly populate the frontend dashboard to display data.
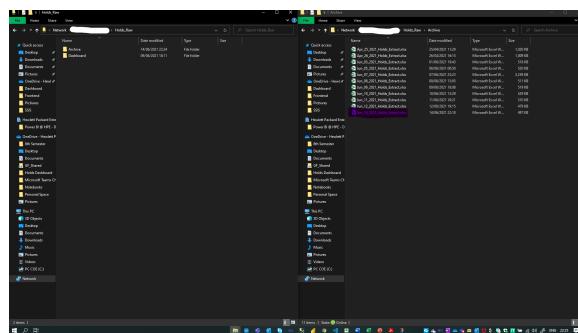
Fig. 25. File in the Archive Folder after execution

The snap below shows us the result visible when the execution of the SSIS package was complete. We can see the Data file being moved from the input folder into the Archive folder.

Thus, we can conclude that we successfully Automated the Holds Dashboard using Power BI, SSMS and Visual Studio. This helped to increase knowledge about IT Automation and increased the knowledge about Visual Programming which was part of the curriculum in the subject of Virtual Instrumentation.

Dashboards have always been the major source of information consumption when it comes to insights about data. Dashboard development and automation being the most important departments in data focused firms like Hewlett Packard Enterprise, Gartner and PhonePe. Thus the successful completion of the Holds Dashboard Automation project has incurred the skills required to create aesthetically pleasing, effective and efficient dashboards for executive level usage.

References

1. Microsoft Power BI, Amrapali Bansal, A. K. Upadhyay, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-7 Issue-3, July 2017
2. Research Data Analysis with Power BI, Vijay Krishnan S Bharanidharan G Krishnamoorth, 11th International CALIBER-2017 Anna University,Chennai, Tamil Nadu 02-04 August, 2017
3. Analysis and Design of Visualization of Educational Institution Database using Power BI Tool, Mandava Geetha Bhargava, K. Tara Phani Surya Kiran & Duvvada Rajeswara Rao, Global Journal of Computer Science and Technology: Software & Data Engineering, August 2018
4. W. Hu,H. Xie,M. Nakas, W. Shi and M. Wang, "Power BI for Impacts Analysis on Cost of Living Caused by Industry Prevalence in Smart Cities,"3rd International Conference on Smart Grid and Smart Cities, 2019
5. Study on Business Intelligence Tools for Enterprise Dashboard Development, K.Gowthami1, M.R. Pavan Kumar2 , International Research Journal of Engineering and Technology (IRJET), 04, Apr -2017
6. ETL tools for Data Warehousing: An empirical study of Open Source Talend Studio versus Microsoft SSIS, Ranjith Katragadda, Sreenivas Sremath Tirumala, David Nandigam, ICWISCE'2015 International Conference on Web Information System and Computing Education, January 2015
7. Leonard, Andy. (2021). Refactoring the SSIS Package Hierarchy. 10.1007/978-1-4842-6482-9_17, January 2021
8. Root, Randal & Mason, Caryn. (2012). Concluding the ETL Process with SSIS. 10.1007/978-1-4302-3489-0_8. , January 2012
9. Leonard, Andy. (2018). SSIS Catalog Compare. 10.1007/978-1-4842-3276-7_7. , November 2018
10. Fouché, Guy & Langit, Lynn. (2011). Intermediate SSIS. 10.1007/978-1-4302-3325-1_8. , January 2011
11. Silva, Yasin & Almeida, Isadora & Queiroz, Michell. (2016). SQL: From Traditional Databases to Big Data. 413-418. 10.1145/2839509.2844560.

12. Shetty, Sucharitha & Rao, B & Prabhu, Srikanth. (2021). Growth of relational model: Interdependence and complementary to big data. International Journal of Electrical and Computer Engineering (IJECE). 11. 1780. 10.11591/ijece.v11i2.pp1780-1795.
13. Ding, Bailu & Chaudhuri, Surajit & Narasayya, Vivek. (2020). Bitvector-aware Query Optimization for Decision Support Queries. 2011-2026. 10.1145/3318464.3389769.
14. Song, Suk-Kyu. (2009). A Genetic Algorithm for Minimizing Query Processing Time in Distributed Database Design: Total Time Versus Response Time. The Kips Transactions:partd. 16D. 295-306. 10.3745/KIPSTD.2009.16-D.3.295.
15. Mercurio, Ralph & Merrill, Brian. (2021). Power BI. 10.1007/978-1-4842-6936-7-12.
16. Aspin, Adam. (2020). Discovering and Loading Data with Power BI Desktop. 10.1007/978-1-4842-5763-01.
17. Clark, Dan. (2017). Advanced Topics in Power BI. 10.1007/978-1-4842-2577-6-14.
18. Rad, Reza. (2018). PowerPoint Integration with Power BI: Sharing, Security, and Deployment Options for Microsoft Power BI Solutions. 10.1007/978-1-4842-4015-1_24.
19. Matthias, Daniel & Managwu, Chidozie. (2021). Data analytics and visualization using Power BI and MS Excel for COVID-19 (Coronavirus). 10.13140/RG.2.2.25204.48001.
20. Aspin, Adam. (2016). Discovering and Loading Data with Power BI Desktop. 10.1007/978-1-4842-1805-1_2.
21. Silva[1], Geraldo & Silva[1], Allan & Coelho[1], Priscila & Pimenta[1], Igor & Rocha, Vladimir Alexei & Souza, Flávio & de Carvalho, Luiz. (2021). Power BI and Time Series in Budget Management of Business Support Services for a Large Brazilian Company.
22. Dobreva, Maria & Pavlov, Nikolay & Rahnev, Asen. (2019). INTEGRATE POWER BI WITH WPF DESKTOP APPLICATIONS.
23. Clark, Dan. (2020). Publishing Reports and Creating Dashboards in the Power BI Portal. 10.1007/978-1-4842-5620-6_9.