"INCENTIVE COMPENSATION AND SALES CREDITING MANAGER"

Project Report submitted in partial fulfillment of the requirements for the Degree of

BACHELORS OF TECHNOLOGY

By

YOGESH KHURANA

Enrollment No: 151289

UNDER THE GUIDANCE OF

Gaurav Ahuja

(Senior Business Technology Analyst)

Arvind Pothula

(Solution Delivery Consultant)

Department of Computer Science and Engineering

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

May-2019

DECLARATION

*I hereby declare that the work reported in this report entitled **"INCENTIVE COMPENSATION AND SALES CREDITING MANAGER"** in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** submitted in the department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology Waknaghat is an authentic record of my own work carried out over a period from Feb 2019 to May 2019 under the supervision of **(Arvind Pothula)** (Solution Delivery Consultant, ZS Associates) and **(Gaurav Ahuja)** (Senior Business Technology Analyst, ZS Associates).*

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Yogesh Khurana

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

Arvind Pothula
Solution Delivery Consultant
ZS Associates India Pvt. Ltd.
World Trade Center, Tower 3, Kharadi, Pune 411014, Maharashtra, India
T | +91 20 6739 9427   M | +91 91 6863 5760
www.zs.com

ZS Impact where it matters.

Acknowledgement

I would like to express our special thanks of gratitude to my mentor **Mr. Arvind Pothula** *who gave us the golden opportunity to do this project on the topic* **INCENTIVE COMPENSATION AND SALES CREDITING MANAGER**, *which also helped us in doing a lot of Research and we came to know about so many new things.*

I am really thankful to him.

Secondly, we would also like to thank Lab assistants who helped us a lot in finalizing this project within the limited time frame.

YOGESH KHURANA(151289)

Table of Contents

LIST OF FIGURES

LIST OF ACRONYMS & ABBREVIATIONS

EC2	Elastic Compute Cloud
CEC	Centers of competencies
ETL	Extract Transform Load
HCP	Health Care Practitioner
IC	Incentive Compensation
IT	Information Technology
JQD	Javelin Quota Designer
MBO	Management by Objectives
JIM	Javelin Incentive Manager
JSCR	Javelin Sales Compensation Reports
SCM	Sales Crediting Manager
CDP	Customer Data Platform
AWS	Amazon Web Services
SQL	Structured Query Language
QC	Quality Checks
DQM	Data Quality Module
FTP	File Transfer Protocol
SBE	Sales Bearing Entity

ABSTRACT

My five month internship project, Sales Crediting Manager part of Incentive Compensation, was with ZS Associates, Pune Office starting from the February 2019.The report is prepared to cover the basic details of the problems and the solutions employed. This main target area of the project were sales representatives (also known as reps) of pharma companies, to figure out a way to provide incentives and compensations to the representatives based on their performance. The main aim of the project was to find a way to handle such large sales data and manage it accordingly so that it can be further utilized in the correct format.

The major problem faced by the client was to track the performance of the sales representatives, taking into account all the difficulties the sales representatives are going through, develop methodology to calculate their incentives and compensations. The methodology is elaborated with details properly in upcoming chapters but in general the incentive plan depends on the product they are promoting and the territories assigned to them. The Plans were designed by respective Business units Leads and were verified with the help client.

After the completion of all the phases, the expected results were a digital scorecard for all the sales representatives available on their mobile phones for both android and iOS. The information available on the scorecards includes, their earnings and goals of the month. These scorecards are then used by the reps to understand what exactly they will be getting by the end of their plan period and where to put their efforts to earn more.

This Incentive Compensation is accomplished using a ZS Software Tool Suite named Javelin whereas Sales Crediting is done using AWS Services like RedShift, EC2 and S3. This Report gives an explanation of said tools. Javelin is based on SQL, Unix and javelin proprietary tools. The finals reports are formatted according to the pre decided XML format.

Chapter – 1

Introduction

1.1 General Introduction

Incentive Compensation is a process of rewarding the sales representatives on the basis of the sales of the product and the profit made by the company. There can be multiple plans on which the sales representatives are rewarded. Cost of the product, profit, budget, revenue, market share are some of the factors on which generally the incentive compensation plan depends. Historical sales data and future predictions are taken into account while developing an incentive compensation plan.

1.2 Organization

ZS Associates (ZS) is a management consulting firm which majorly works on the sales and marketing problems of the pharma industry. Prabhakant Sinha and Andris Zoltners, professors of marketing at the Kellogg School of Management at Northwestern University founded ZS in 1983. The firm employs more than 5,000 employees in 23 offices in the Asia, Europe and America and is headquartered in Evanston. ZS's CEC centers are present in Pune, New Delhi and Bangalore, India. 49 of the 50 largest drug-manufacturers and 17 of the 20 largest medical device makers are the clients of ZS, apart from offering services like financial services, consumer products, telecommunications, industrial products and logistics industries & Transportation.

1.3 Motivation

ZS Associates is Management Consulting Firm specializing in the pharma industry. Our main focus is to solve the sales and marketing problems of the clients. The problem statement posed here is Incentive compensation involves various calculations and

computation of goals for the reps to achieve in any given fiscal quarter. The computation of goals involves reducing complexity of plans and maintaining other dependent factors like fair treatment across all the reps, potential to increase market share and achieve market dominance. This involves consideration and analyzing of data form various sources to drive methods like regression analysis to compute fair and simple goals and create an impact on the market. The incentives are captured accordingly from data generated in the quarter and compensated fairly to various reps across the business units of an organization.

1.4 Objective(s)

The objective of the project is to design end to end IC systems to generate incentive compensation by using various ETL methods. The design of the system includes transformation of data from various sources, implementation of various IC plans and robust testing of ETL systems to compute accurate incentives to large sets of data. The objectives of the project include design of IC and goal setting systems. Goal setting systems implement plans analyzing various market trends, previous market performance and various factors are considered to compute market potential and quotas(goals) for reps.

1.5 Target Specification

Our client requires us to deliver the 3 reports to the IC team which were basically the HCP level sales report, Territory level sales report as well the IC report. HCP level sales report were used to calculate the HCP level sales for the hcp level analysis and to calculate the hcp level incentives and compensations. Territory level sales reports were sales aggregated to the territory level. We also include the district, region and nation level sales. Territory level sales reports are used to calculate the payouts of the territory sales representatives, territory level sales managers, district sales managers, regional sales managers and the nation sales manager. IC report is used for the analysis of the product how it is performing as compared to the market. To identify the rise or drop in sales as compared to the previous month or quarter sales. As well IC report

are also used to identify the goals of the sales representative for the next month.

1.6 Scope of Work

The IC systems are designed to compute incentives and generate reports on proprietary systems of ZS on quarterly basis for reps across various business units of an organization. This involves analyzing various data sources, ETL, data validation with pre-designed system featured to be modular, scalable and robust for long term efficiency. The systems are designed to be easily replicated and simple to be redesigned in the event of requirement from clients

Chapter – 2

BACKGROUND THEORY

2.1 Introduction

Most of the ZS are from the pharma industries, whenever a product cycle changes pharma companies get in touch with the ZS for consultation. As pharma companies spend lot of money on research and development of the product, the price of the product is generally high. To launch the product in the market, the pharma industries are requires patent which is generally of 6 to 7 years. After the patent expires the competitors develop the same kind of product with cheaper price which reduces the market share of the original product which eventually leads to the saturation level of the product. Therefore different planning is required to be done for different product at different levels.

Generally pharma companies hire hundreds of sales representatives which are assigned to a specific territory, and their main motive is to promote the product in the territory assigned to them. They work under the territory managers.

In USA, pharmacists are required the to store the data regarding whatever product they sell which is further acquired and aggregated by the government organization named IMS. Consulting and and analytics firm acquire this data to perform analysis on it. All of this data and the data provided by the client is used to develop the incentive compensation plans for the client.

ZS takes a holistic approach to transform information capabilities, account factors such as existing technology infrastructure, information-sharing patterns, and interactions between IT and business teams. ZS blends keen insights, business acuity and a collaborative approach to help gain a competitive advantage from business technology

> ➢ Functional requirements; Secure server, User friendly access, Sequenced transaction

- Product features; Incentive compensation design, Quota design, Sales compensation reports.

- User characteristics: Sales representatives across a pharmaceutical organization within various hierarchy levels.

- Assumptions and Dependencies: --ND--

- Domain Requirements; ZS internal Network, License for access to the administrators and end users

- User Requirements: Tablet, Personal Computer, Android or IOS base mobile phone.

The overall system is divided into two phases Sales Compensation and Sales Planning. The output of both of these are combined to produce the quality reports.

2.2 Sales Planning

Sales Planning phase deals with the planning of sales which can be further divided into three steps names, call planning, roster management and territory management.

2.2.1 Call Planning Team

Call planning can be defined as the process of assigning the list of HCPs which a sales representative must target in his territory. Call planning team is responsible for the designing of call plan for various sales representatives on the basis of their skills, product importance, years of experience etc. ZS provides Javelin Call Plan manager which is used by the sales representatives to track their call plan progress.

Fig. 2.1. Call Planning tool: Javelin Call Plan Manager

2.2.2 Roster Management Team

Roster Management team is responsible for managing the roster for the sales representatives. They work in alignment with the call planning team and provide us with the roster details like HCP ids, representatives ids, their eligibility and area of working.

2.2.3 Territory Management Team

The territory management team is responsible for dividing the nation into regions, regions into districts, districts into territories and territories into zips. They generally divide on the basis of number of HCPs present, and they are responsible so that no territory is highly reporting and other one is low on sales. Territory management team is also responsible for the alignment which is generally zip territory alignment which is used by the sales crediting and sales compensation teams.

Fig. 2.2. Territory management tool: Javelin Territory Designer.

Fig. 2.3. Territory Management tool: Javelin Alignment Manager

2.3 Sales Compensation

Sales compensation is responsible for the final implementation of the sales plans developed by the sales planning teams. Sales compensation teams includes goal setting team and sales crediting team.

2.3.1 Goal Setting Team

Goal setting team is responsible for defining the territory level goals, which have to be accomplished by the sales representatives in order to earn incentives. They generally work on Javelin Incentive manager, Javelin Quota designer and Javelin Quota refiner.

Javelin Incentive Manager is a powerful platform enabling the implementation of Incentive Compensation Plans by its support for even the most complicated business rules. With a dedicated Plans section supporting a plethora of pre-configured plans apart from having a customizable framework to implement custom IC Plans based on the manually configured business rules driven by user implemented procedures and specially configured input files. Additionally, a robust quality check framework is available which helps in detecting and correcting the erroneous data fast and easy.
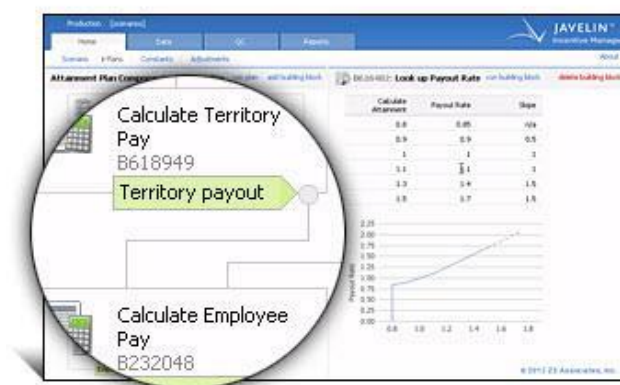

Fig. 2.4. Sales Compensation tool: Javelin Incentive Manager.

JQD is Web-based software that establishes fair, motivating and effective sales goals at every level across the sales organization. With JQD, you can streamline the goal-setting process, test multiple inputs and quickly see the impact of àny changes. The tool is customized to define various data analysis methods like regression to derive accurate predictions with effective quotas (goals).
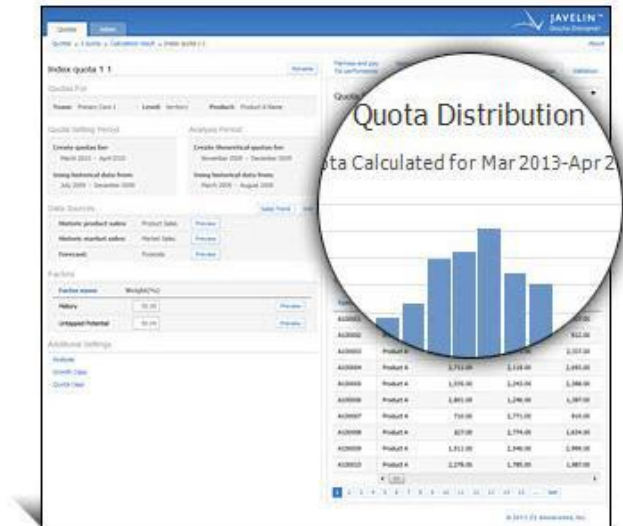

Fig. 2.5. Goal setting tool: Javelin Quota Designer

To improve sales effectiveness, reps and managers require visibility into current and future payout potential. JSCR enables seamless and easy configuration of web-based scorecards and compensation reports via the data-output from JIM making it an important reporting platform within the Javelin ecosystem.
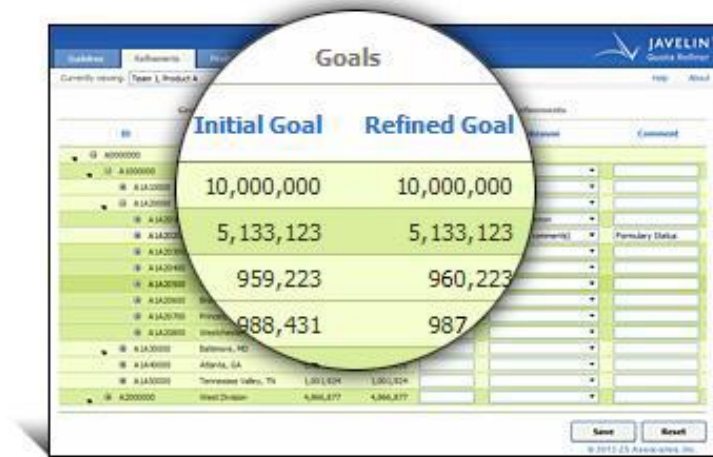
Fig. 2.6. Goal Setting tool: Javelin Quota Refiner

2.3.2 Sales Crediting Team

Sales Crediting team is responsible for collecting the raw sales data from INS, transforming it into the correct format, and applying all the business rules. It uses UNIX (EC2 Machine 4 8.x nodes) to retrieve and sort the data. AWS Redshift is used as the data warehouse and S3 as the bucket storage. Sales crediting team is responsible for analysis and data cleaning of data collected. The sales are collected, business rules are applied, and territory level sales are generated. The territory level sales are rolled up to district level sales and the district level sales are rolled up to give nation level sales.

Fig. 2.7. Amazon REDSHIFT logo

Fig. 2.8. Amazon EC2 logo

2.3.3 Reporting and MBO team

Reporting and MBO team are responsible for the development of the final reports and payouts of the sales representatives. Reporting and MBO works on JSCR which the final web utility seen by the sales representative and the client. It is a highly customizable and advanced tool featured to report to various clients and their potential clients for continued realization of data flow and reporting. With several added functionalities such as easy iPad accessibility, a dedicated publish option and a customized look and feel, JSCR gains a definite edge over the traditional excel reports thus improving the overall user experience.

Fig. 2.9. MBO team tool: Javelin MBO Utility

JSCR enables seamless and easy configuration of web-based scorecards and compensation reports via the data-output from JIM making it an important reporting platform within the Javelin ecosystem.

Fig. 2.10. (a) Reporting Team tool: Javelin Sales Compensation Reports (view 1)
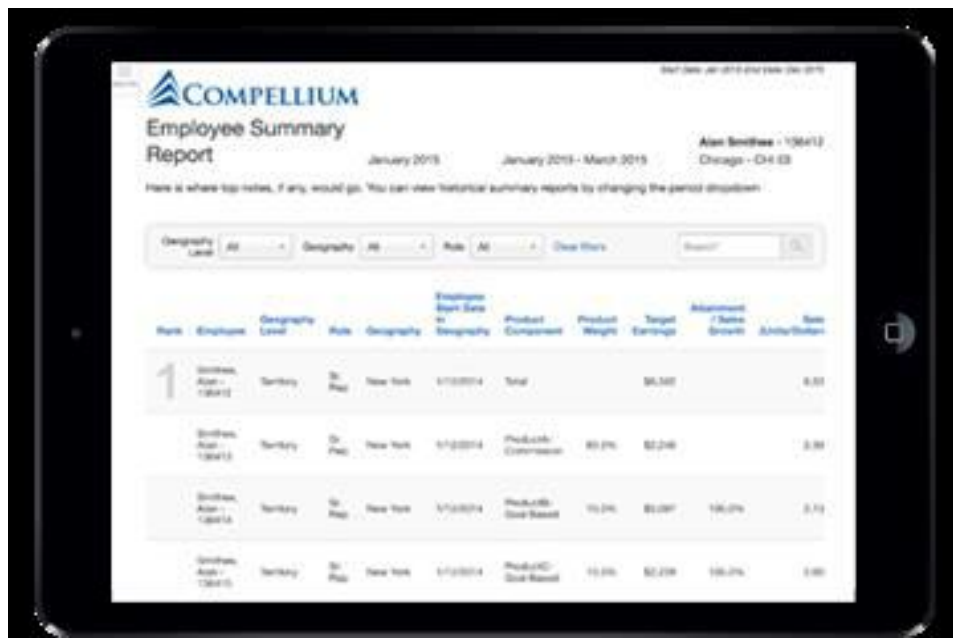

Fig. 2.10. (b) Reporting Team tool: Javelin Sales Compensation Reports (view 2)

JSCR offers insight into an individual's key performance metrics. As a result, sales force can easily track essential metrics against current plans and recognition programs.

Fig. 2.11. (a) Reporting Team tool: Javelin Sales Compensation Analytics (view 1)

Fig. 2.11. (b) Reporting Team tool: Javelin Sales Compensation Analytics (view 2)

Chapter – 3

SALES CREDITING SYSTEM DESIGN AND OVERVIEW

3.1 Overview:

Sales Crediting is a process of determining "who gets credit" and "how much credit" for a transaction or event.

Fig. 3.1. Sales Crediting Overview

3.1.1 Data Sources:

Data sources for Sales Crediting Manager includes data from client, third party data sources and ZS internal data.

> ➢ **CLIENT DATA: -** Client provide us the details of business rules and market definitions which are required to be applied to the raw sales for the process of incentive compensation. Business rules includes rules stating prescribes and outlets to include, factoring to be applied, call plan alignment, zip territory alignment, manual sales adjustments

and many more. Market definitions are required to identify the product and market which are included in the final incentives calculation as well for the comparison with the competitors.

➢ **Third-Party data sources:** Third party data sources companies like IQVIA, IMS etc. provide us the raw sales data on which we apply the business rules and market definitions.

➢ **ZS Internal data:** Sometimes we are required to apply some manual adjustments for some quarter which specifically is not a business rule, we need the feed the system with internal data with specific format.

3.1.2 Sales Crediting key activities: Below are the key activities performed by the sales crediting manager.

➢ **Alignment:** Generally, the raw sales comprises of HCP level sales and with the help of SCM we are responsible for aligning the HCP level sales with the zips and territories. Alignment is generally defined by the two-step process:

➢ **Call Plan alignment:** In call plan alignment, we align the reps with the retail sales, as reps are responsible for the retail sales.

➢ **Zip Territory alignment:** We do Zip territory alignment for the non-retail sales, non-retail sales includes, mass sales by the pharmacies, hospitals etc.

➢ **Adjustments:** Sales adjustments comprises of manual sales adjustments and factoring which we need to apply on the raw sales data. Zip delta sales is also the part of manual adjustments.

➢ **Exclusions/Inclusions:** These are the business rules we have to apply specific to inclusion and exclusion of some territories and HCP ids.

- **Rollups:** Rollups includes aggregating zip level sales to territory level sales. Territory level sales to district level sales. District level sales are aggregated to region level sales and finally region level sales to the nation-wide sales. It is important to do roll up as the provide us different level of comparison as well help us to calculate the incentives of district and region level managers.

3.1.3 Output reports: Output of SCM is different level of reports including, HCP level sales, territory level sales, IC Comp deed, and delta reports. The reports are further used by the IC team for incentive compensation calculation, analytics team for analysis and by the sales planning team for goal setting and call planning.

3.2 Features of Sales Crediting Manager: Below are the some of the features of the sales crediting system manager.

- **Quick System Processing:** Apply Multiple Business Rules on large amounts of data quickly.

- **Ease of loading Data:** Load variety of data sources without extensive configuration.

- **End to End DQM System:** Configure Validations on Inputs and Business Rules at every step.

- **Data Processor System:** Configurable building blocks to implement business logic such as exclusion, roll-ups.

- **Easy Configuration tool:** Configuration Tool to quickly and easily setup runs from scratch with a variety of Business Rules.

- **Javelin Integration:** Integration with Javelin Systems to allow automatic transfer of outputs to JIM, SCM triggering via JIM UI etc.

- **Scalable and Secure:** The SCM system can be easily scaled up as per

data and processing requirement.

> Data Encryption, VPC, VPN Tunnels etc. are in place to provide security.

> **Integration with Reporting tools:** Reporting tools like Tableau, Microstrategy can easily be implemented onto SCM to generate reports based on live data which is not limited to just inputs and outputs.

3.3 Sales Crediting Process Generic: Quality checks are done on input files and then the files are standardized according to the requirements. Then pre-alignment business rules are applied to these input. Reps are assigned to a territory known as alignment. After alignment post-alignment business rules are applied. And then we rollup and aggregate the data.
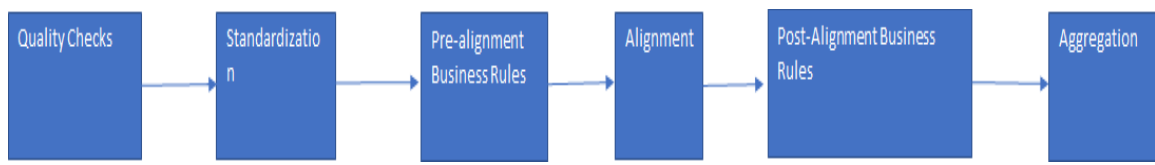

| Quality Checks | Standardizatio n | Pre-alignment Business Rules | Alignment | Post-Alignment Business Rules | Aggregation |

Fig. 3.2. Sales Crediting Process

> **Quality Checks:** Quality checks include scenarios where we have to remove not null datas, duplicate datas (present when we perform joins), inconsistent data (when the datas in two corresponding tables are inconsistent with each other.)

> **Standardization:** Standardization includes data granularity, market defination.

> **Data granularity** : We may get input input in the form of daily, weekly or monthly basis and the final reporting is always done on monthly basis .

> **Market Defination**: Brand product refers to sales of all dosage of a particular drug is considered as a single product.

- ➢ **Pre-alignment Business Rules:** Apply factoring, adjustments, inclusions and exclusions.

  - ➢ **Factoring:**   1) Trx=1/2 Nrx

    2) If a customer A reports product sales as 1 Unit for a sale of 100 samples of 1 product, while another customer B reports 1 Unit for sale of 10 samples of the same product. Then factoring needs to be applied. A=10*B

    - ➢ **Adjustments:** Sales adjustments comprises of manual sales adjustments and factoring which we need to apply on the raw sales data. Zip delta sales is also the part of manual adjustments.

    - ➢ **Exclusions/Inclusions:** These are the business rules we have to apply specific to inclusion and exclusion of some territories and HCP ids.

    - ➢ **Alignment:** Generally, the raw sales comprises of HCP level sales and with the help of SCM we are responsible for aligning the HCP level sales with the zips and territories. Alignment is generally defined by the two-step process:

      - ➢ **Call Plan alignment:** In call plan alignment, we align the reps with the retail sales, as reps are responsible for the retail sales.

      - ➢ **Zip Territory alignment:** We do Zip territory alignment for the non-retail sales, non-retail sales includes, mass sales by the pharmacies, hospitals etc.

- ➢ **Post-Alignment Business Rules:** Includes the same steps as the pre-alignment business rules. But in the case of post alignment, customer + territory level rules are applied and in the case of pre-alignment, customer level rules are applied.

- ➢ **Aggregation:** Rollups includes aggregating zip level sales to territory level sales. Territory level sales to district level sales. District level sales are

aggregated to region level sales and finally region level sales to the nation-wide sales. It is important to do roll up as the provide us different level of comparison as well help us to calculate the incentives of district and region level managers.

3.4 Sales Crediting Architecture: The input level sales are received by the CDP system which is the FTP server. Then the required files are downloaded from the M-Box to our land (Unix server). Then the files are stored in the Amazon S3. The files stored on S3 is accessible by Amazon Redshift.

Fig. 3.3. Sales Crediting Architecture

➤ **CDP(Customer Data Platform):** We receive input from CDP sysytem. We receive files on MBOX from CDP, which is an common layer for sharing of files between CDP and ZS. M-Box is an FTP server where we can provide the user names and password to acces the files.

➤ **Landing:** The required files are downloaded from the M-Box. The files are unzipped. Awking of the files are done.

➤ Awking maybe done for multiple purposes:

> ➤ Column swapping

> ➤ Dos to unix conversion

> ➤ To exclude the header

➤ **S3:** S3 is a storage location on cloud.

➤ **Redshift:** Redshift is the datawarehouse service where actual data processing takes place (business rules are applied). We can load data from S3 to Redshift

and vice versa.

3.5 Process Overview:
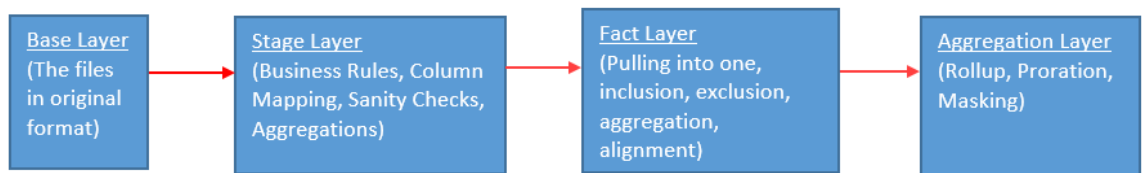
Fig. 3.4. Process Overview

➢ **BASE LAYER:** Base layer contains anything and everything that we are loading to our system.The input to the base layer are the awked files and the business input files. These files are directly loaded into tables without applying filters. Business input files contain the business rules and their ids that tell us about the rules that we have to apply at the further layers.

The base layer contains the files in the original format. We require this to compare with tables that we get at the further layers.

➢ **STAGE LAYER:** At the stage layer the basic business rules are applied. We also filter the data according to the requirements.

   ➢ **Ex:** The base layer table contains data for 35 months and we require data for 24 months. So we would filter out the 24 months data.

   ➢ Column Mapping is also a part of stage layer.

      ➢ **Ex:** In one file the column name is product_id whereas in the other file the name is prod_id, therefore we need to map the product_id column to prod_id column.

   ➢ We also apply null, duplicate data, inconsistent data checks at this layer.

   ➢ Aggregation refers to summing up of daily or weekly data to convert

into monthly data.

- **FACT LAYER:** Pulling the data into one table from multiple tables happens at this layer.

    - Inclusion and Exclusion Rules are applied here.

    - **Example :** The rules that we have to apply have an id. We have a flag column which tells us about the rule that has been applied. Let us first include all the rules, then we could exclude some of the rules and then we could include some exceptional cases.

    - Aggregation could be performed at this stage too if it hasn't been performed at the previous stage.

    - Pivoting Up Data is converting the data from row level to column level.

    - Alignment is assigning customers to territories. This gives us the data at the territory level.

    - The final table that we obtain at this stage is compared with the one at the base level to check that the validations and rules have been correctly applied.

    - Adjustments is performing adjusmtents for a particular customer for a particular time period.

AGGREGATION LAYER: Previously we have aligned customer level data to territory level data, now we need to align territory level sales to the district level sales.

➢ Example:

Customer	Territory	Sales
J1	A	4
J2	A	3
J1	B	5
J2	B	5
J3	C	6

Table 3.1 Raw sales

Territory	District
A	X
B	X
C	Y

Territory	Sales
A	7
B	10

Table 3.2 Territory District Mapping   **Table 3.3** Territory Level Sales

So, the district level sales for district X become 17.

Similarly, we can aggregate the district level sales to get the region level sales and the region level sales to get the nation level sales.

➢ Prorating data means reducing the baseline period data to get equivalent data.

Example : Let baseline period be from Jan'18 to Jun'18 and current period be from Jan'19 to Jun'19. But currently we have the data just for

Jan'19-Feb'19. We need to compare this with the baseline trend. Therefore, we multiply the baseline data with 8/26.

Masking refers to hiding some of the details which we can't reveal to the reps.

Chapter – 4

IMPLEMENTATION

4.1 IC Manager Implementation:

The main aim was to develop an IC system which can act as a single place for all the output reports.

There are various steps involved in the IC system setup. The first step involves collecting the required input files and uploading them into the landing layer of the IC system.[Fig 4.1].

Fig. 4.1. All the Necessary files uploaded

Many of the files required to be reformatted according to the system fixed format. All the formatting and the manual changes along with the uploading of data is done in the data preparation tab of javelin system [Fig 4.2]. In addition with uploading files, validation checks are also done in the data preparation tab. There are different types of validation and quality checks which are done on the input files in the data preparation tab.
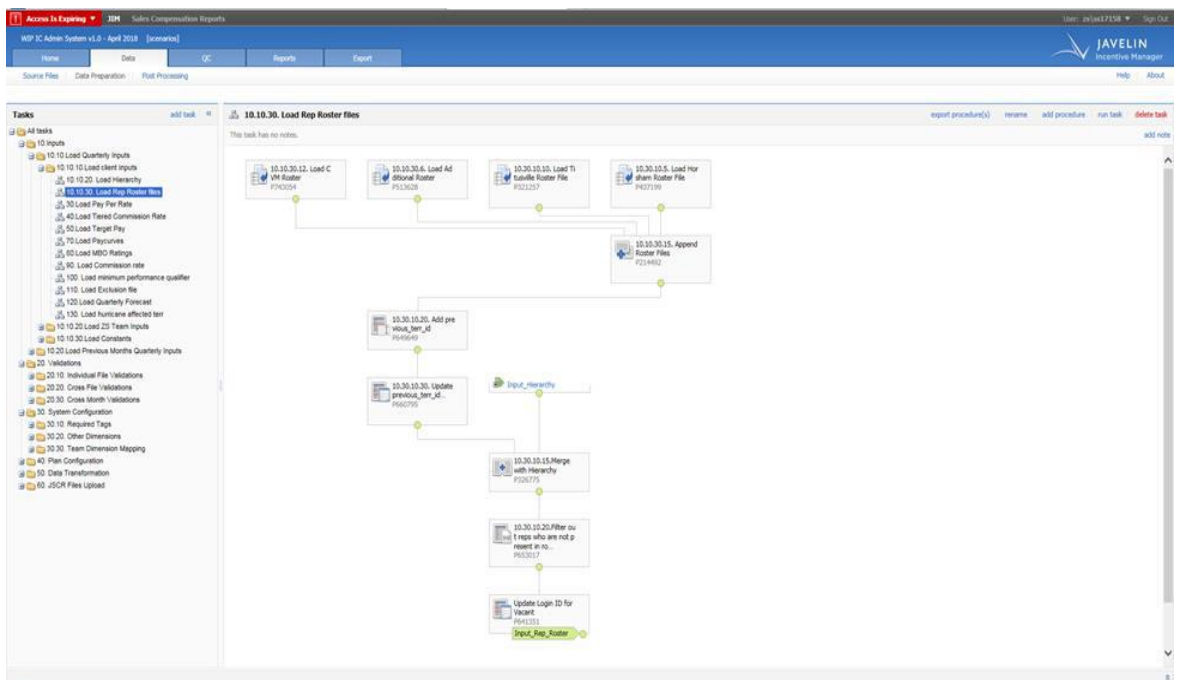
Fig. 4.2. Data Preparation Tab

After the uploading and preparing data, we create the required IC plans. IC Plans [Fig 4.3] can be defined as the set of instructions which are performed on the input data to filter out the required data on which we perform the required calculations. Generally there are multiple IC plans which depends on Business rules, BUs, products, and requirements etc.

Fig. 3.3. Plan Tab

According to the requirements and need preparation of various IC plans were done with their set procedures. All of the data rows are identified with the help of data tags. Thus to ease the process of transferring the data from one phase to another we created various data tags in the IC system[Fig 4.4]. The plans differ across business units, therefore they are identified according to the tag[Fig 4.5]
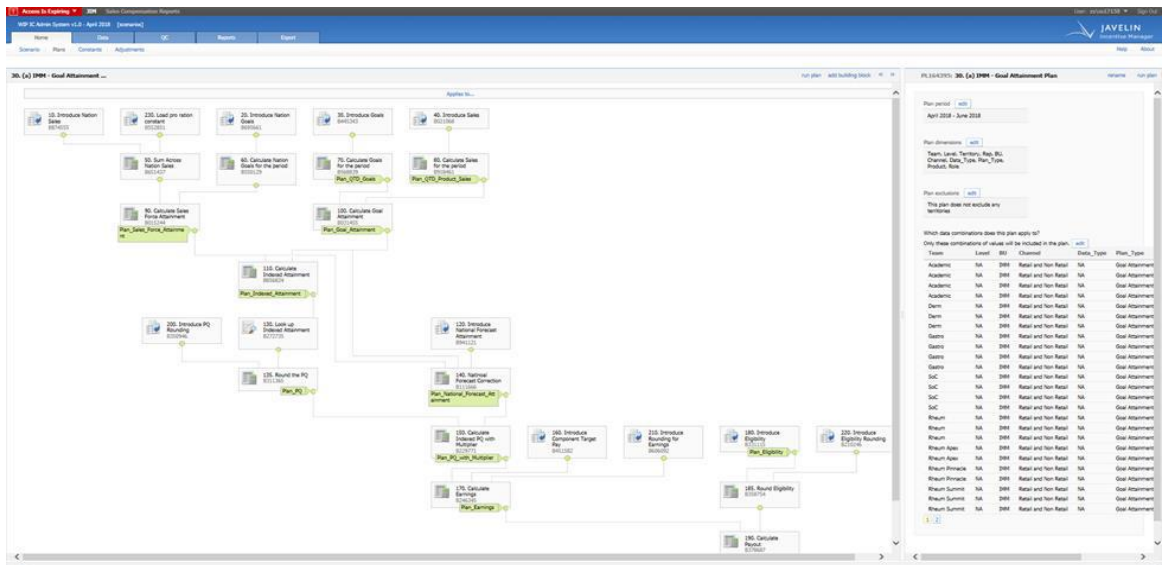
Fig. 4.4. Flow of Plans
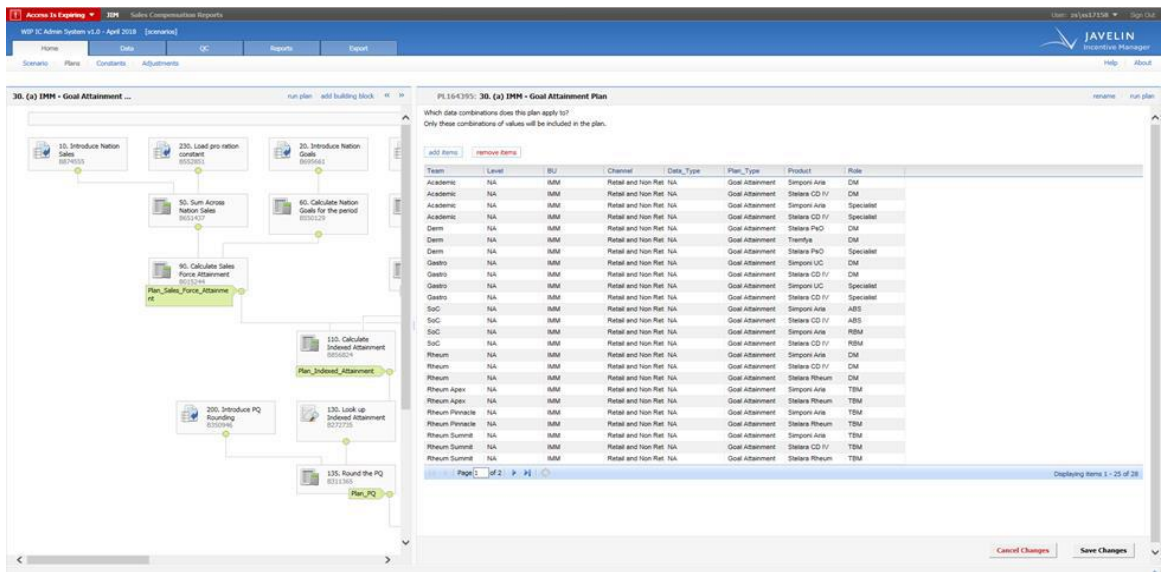
Fig. 4.5. Filter of Plans

Once the plans are created and all the calculations are done data is transferred to the post processing phase [Fig 4.6]. The final phase where reports are generated is this phase. At this step we have already acquired all the data required to prepare reports. Now in this step a master data base is created through which we can extract the required information for all the reports. JSCR require files to be created as reports and report files.

JSCR tag is finally attached to the report files which proves it's authenticity. From the master database all the required information is extracted, required for the reports.
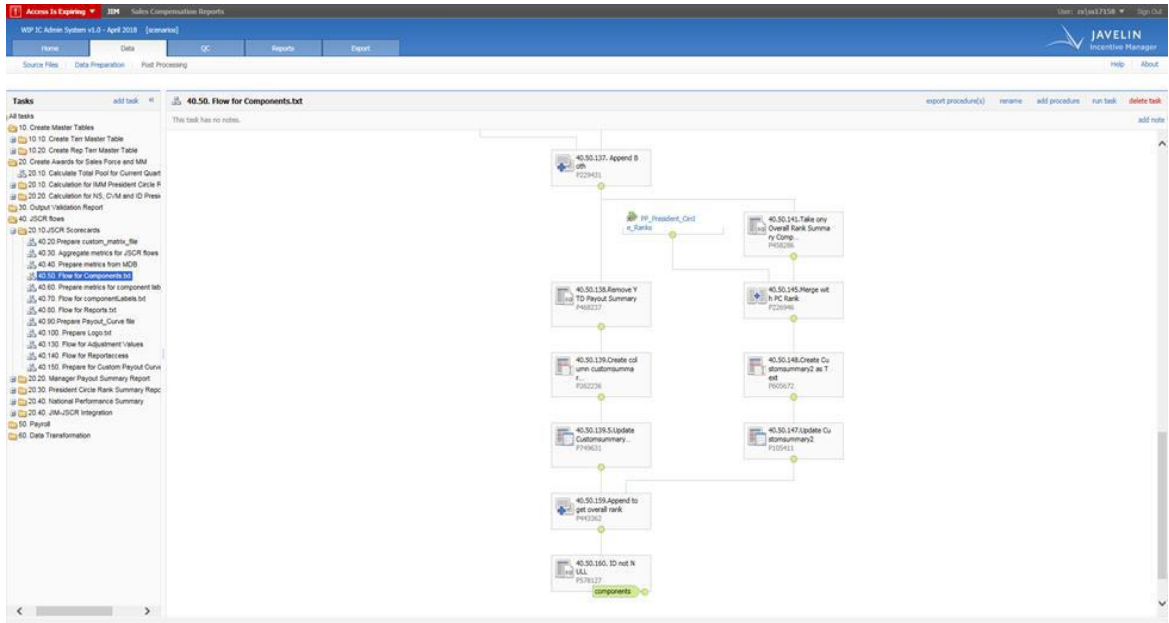
Fig. 4.6. Post Processing Tab

In the last step the reports are reformatted to the required XML format designed to publish the JSCR reports [Fig 3.7]. The client is provided with the temporary access to the reports for the approval. Once we get the validation from the client, we publish the sales reports so that sales representatives can view the reports in their respective javelin accounts.
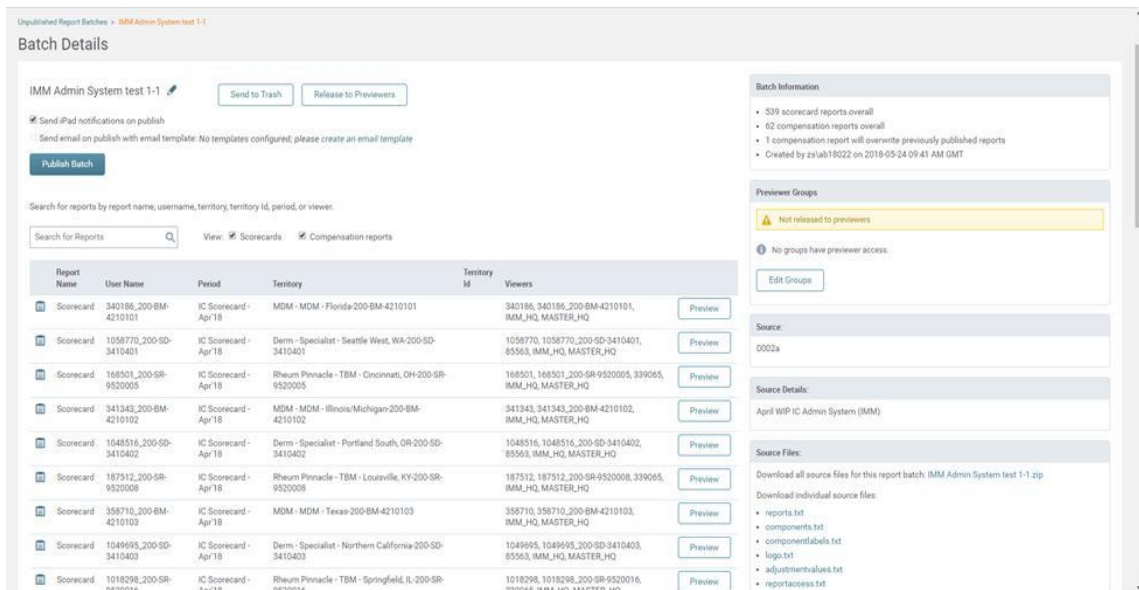
Fig. 4.7. Batch of all the reports produced

4.2 Sales Crediting Manager Implementation

The **SCM** sales crediting tool offers process setup by choosing and sequencing pre-developed, customizable codebase.

4.2.1 Hardware and Software requirements

SCM framework will run on a AWS EC2 machine compatible with the following specifications:

➢ Minimum ram of 2GB (For optimal system performance).

➢ A complete installation from the compressed installation file should take about 20MB of disk space.

➢ As installation process needs to communicate with the AWS infrastructure elements such as S3, Redshift a working internet connection on client machine is required.

The following AWS infrastructure elements are pre-requisite for system installation:

- Amazon EC2 (*RedHat Linux 3.10.0*): Orchestration Engine

- AWS Redshift: Relational Database

- Amazon S3 bucket space: File hold disk space

The following access/tools will be needed to communicate with AWS elements:

- ZSSERVICES Remote Desktop Access.

- PUTTY <Version 0.65>: For EC2 UNIX Shell access

- Filezilla <Version 3.20>: For transferring files from Client to AWS EC2 Machine

- SQL Workbench < Version 121>: For interacting with AWS Redshift.

4.2.2 SYSTEM CONFIGURATION

4.2.2.1 Sales Crediting Interface (SCI)

Configuration in the SCM Tool is carried out using the Excel SCI Tool which connects to the database in order to pull details, performs changes locally and then pushes them to the Database. The SCI tool can be used for the first time system setup & ongoing changes.

- The Sales Crediting Interface – consists the following tabs:

  - **Cover** – Cover page of the Tool. Used to configure new run / modify an existing one

  - **Admin Config** – Contains the database details as well as the Team, Channel, Market etc.

  - **SalsA Component Repository** – Contains the list of steps & sub steps

in the process and their ID's

- ➢ **QC Def** – Contains the quality checks for each functional area. Can be used to automatically fill the CTL_QC_PROF tab when a new file is added

- ➢ **Process Control** – Used to add new steps to the process/ Delete existing steps

- ➢ **Business Rule Query** – Used to configure the new steps which are being added (including source, target etc.)

- ➢ **Process Steps** - Contains the list of steps in the current process

- ➢ **Dataset Profile** – Contains the dataset name and associated parameters

- ➢ **Business Rules** - Contains the list of SQL queries along with source and target

- ➢ **QC Steps** – Contains the list of DQM checks to be applied from Base to Stage

- ➢ **Process QC Steps** – Contains the list of Process QC Checks to be applied after any sequence/subsequence

- ➢ **CTL_DQA_KEY_STORE** – Contains the DQA Key at which Log record will be kept of the data for each step

4.2.2.2 Process Control Table

This is the most important table to be filled out when configuring SCM as it acts as the master control table, controlling all the process steps to be executed. It consists of steps divided into 5 different categories, namely – Environment Prep, Data Loading, Applying Business Rules, Output and Cleanup.

It connects to two separate tables for controlling specific tasks such as fetching of input datasets and performing data sanity checks on the base tables. This table contains the details about each process to be executed such as the name of the executable, order in which the processes need to be run, team, channel, market information using which the process of Sales Crediting can be layered etc.

4.2.2.3 Dataset Profile Table

This table contains the details about each input dataset which needs to be loaded for processing. This drives the automated fetching of datasets from the remote location to the SCM environment and subsequent loading onto the SCM tables.

4.2.2.4 Business Rule Query Table

Once all business rules have been added to the process control in the necessary sequence, this table provides specific information about the data on which the rule needs to be applied, the supporting inputs the rule needs, the operation to be performed on sales in the form of queries etc.

4.2.2.5 DQA Key Store Control table

For each business rule applied, sales summary (i.e., record count, summed up NRx sales and summed up TRx sales) is logged after the application of each business rule to trace change in sales post application of the rule. The level at which this information is captured is defined in this table.

4.2.2.6. Process QC

To QC all the running processes in the SCM system, we can use the Process QC Framework (DQM Module Section) The list of quality checks is exhaustive for each of the table and a user has to use an active flag column to activate or deactivate any quality check.

4.2.2.7 DQM Module

This table consists of all the quality checks that have been built into SCM for validating input data while moving from LAND layer to STG layer and can be applied on the dimension inputs like channel profile and hierarchies, product profile and hierarchies or sales inputs. The list of quality checks is exhaustive for each of the table and a user has to use an active flag column to activate or deactivate any quality check.

4.2.2.8 Auto Trigger

The Auto Trigger module is an added functionality for the SCM tool which allows the user to effectively utilize the DEAD-Time for system. Due to time zone differences it has been noticed that for most teams, there is a delay of about ~12 hrs between when the raw data files arrive vs when the team start processing it.

The auto trigger module aims to utilize this dead time to trigger the process automatically as soon as the files arrive and pass all the QC Steps configured in the auto trigger module.

The auto trigger module is divided into three sub modules which can be triggered independently as per the use case. Below sections details all the Sub-modules.

4.2.2.9 Auto Push

This module is designed to allow the SCM System to automatically transfer Sales data to a JIM Scenario. This is intended to augment the existing SD enabled JIM-SCM Integration which allows transfer of files manually.

Features:

> Ability to transfer SCM Output Files automatically to JIM Scenario without any intervention from Users

> ➢ The Auto-push Functionality can be configured as an extra step in Process Control Table of SCM

> ➢ Detailed logging maintained to track process execution

> ➢ Email Notifications to notify users after the transfer. Detailed information on each file which was configured for transfer.

4.2.2.10 Archival

The archival module is built into SCM for auto archiving the Redshift data onto S3 files. This is done based on user defined queries. The result of the queries is stored into flat files on S3 bucket.

4.2.3 Cycle Creation

The SCM engine runs on the concept of *"cycle time id and scenario id"*. For each run, the users need to configure the processing control tables based one run cycle time id and scenario id. These are two 6 and 2 digit integers logically describing a system run.

4.2.4 System Trigger

There are two ways to trigger a SCM process.

> ➢ We can trigger the SCM process directly by executing the SCM Wrapper script. This requires all the control tables configured on the system and all the dataset files present on the FTP location.

> ➢ We can configure the control tables for auto trigger module (Refer to section 12 of this document) and the dataset profile table will be configured as and when the sales files are received.

4.2.5 SCM Processing Logs

SCM processing logs are maintained in at two levels. The logs are maintained for each individual script that is run and for the wrapper/nohup scripts executed during processing.

The wrapper scripts have their individual logging files where all the important script parameters and outputs. However, the SCM engine calls several of the wrappers in nohup mode. When a script is called in nohup mode, all the terminal output statements are logged into separate files.

4.2.6 SCM Governance Logs

The governance module creates several logs that are general and do not depend on SCM processing. All the logs are file based and stored in tab delimited TEXT files. Logs are created if any of the below event is triggered using the SCM Governance module:

- ➢ Cluster Start

- ➢ Cluster Stop

- ➢ Snapshot Create

- ➢ Snapshot Delete

4.2.7 S3 and redshift Transfer Logs

In addition to all the script level logging, SCM also logs all the transfer data between your S3 bucket and Redshift done by SCM processing. This may include loading of files from S3 to Redshift during the normal processing run or unloading the files when performing system archiving activities.

4.2.8 Process Logging

For each process executed, a log table records the error code and message if any, no. of

records in input, no. of records in output and so on.

4.2.9 QC Logging

For each QC performed on any dataset id, logging is done at both summary and detail level. The summary log captures information such as ID of the QC performed, no. of records that failed the QC etc. The detailed QC log table stores all the erroneous records.

4.2.10 DQA Summary Log

For each business rule applied, summary is logged at a defined level (product, time period and data type) to identify the number of records that were modified. The fact table also contains two additional columns, business rule log and business rule value which identify the rule applied and the value of the changed entity before application of the rule respectively, thus maintain a record level log.

4.2.11 Sales Data

Sales data captures the sales recorded over a time period, invoicing transactions involving the key entities listed below. Sales data can be stored as normalized or de-normalized sales buckets.

➢ **Sales Bearing Entity (SBE)**: Any entity for which the system receives sales data is defined as SBE. Few examples of this type of entity are: prescribers, hospitals, accounts, zips, zip and specialty etc.

   ➢ All SBEs are identified by unique id

   ➢ SBEs can have a hierarchy. For ex: a group of doctors affiliated to a hospital or a group of small hospitals affiliated to a hospital chain etc.

➢ **Product**: Product can be defined as one molecule or a portfolio/ market a) It can have a hierarchy

   ➢ It can be a non-standard pharmaceutical product such as samples,

brochures as well etc.

- ➤ Product can be aligned to a SBE via Position.

- ➤ **Position**: Position can be defined as a role with a unique identifier. It is expressed in terms of a team, product, geography and a SBE type. It is similar to the concept of commonly used term 'territory' a) It can have a hierarchy

- ➤ **Payor**: Payor is the insurance plan that pays for a particular prescription. In cases where a particular prescription is not covered by any payor (private or federal), payor field of a transaction will be populated by the value 'CASH'

- ➤ **Channel**: Channel provides information about sales channel used to generate a particular drug sale. It can have a hierarchy but it is primarily used to apply specific business rules and enable selective processing of sales data

- ➤ **Market**: Market provides information about the market to which the product belongs. One product may belong to more than one market.

- ➤ **Data type**: Data type provides information about the metric in which sales figures are recorded e.g., Dollars, Scripts, Units, Bottles etc.

4.2.12 BRL Entity Master Data

While the fact table only stores the identifiers of each of the entities listed above, separate tables such as, profile tables are used to capture more information about each of these entities and hierarchy tables are used to store any hierarchical information about each of these entities, if it exists.

- ➤ **Sales Bearing Entity (SBE) Table:**

  - ➤ **SBE Profile Table**: This table stores all data related to a SBE such as specialty if SBE is a prescriber or sub category code if SBE is an account, address of the SBE and so on.

- ➢ **Product Tables:**

  - ➢ **Product profile Table**: This table stores all data related to a product such as its name, form, strength, manufacturer, price etc. As products have various formulation strengths under a particular brand, it can be considered as a hierarchy and this table records profile information about products at form strength level as well as at a brand level and so on.

  - ➢ **Product hierarchy Table**: This table stores a child to parent relationship assigning a lower level number to the child and a higher level number to the parent e.g., a product 'P1' of brand, 'BRAND 1', here the product with specific form strength would have a level 10 and the brand(parent) would have a level 20. This table records child to parent relationships such as various formulation strengths under a specific brand, various brands under a specific market, various formulation strengths under a specific market etc. This hierarchy information can be very useful when aggregation is required from the child to parent level.

- ➢ **Position Tables:**

  - ➢ **Position profile Table**: This table stores all data related to a position such as position name, team to which the position belongs etc. As position has a hierarchy, this table records profile information about all positions e.g., at territory level, at district level, at region level etc.

  - ➢ **Position hierarchy Table**: This table stores a child to parent relationship assigning a lower level number to the child and a higher level number to the parent e.g., a territory, 'Terr 1' in the district, 'Dist 1', here the territory would have a level 10 and the district (parent) would have a level 20. This table records child to parent relationships such as territories in a district, territories in a region, districts in a region etc. All territories belonging to a team also states a child to parent relationship. This hierarchy information can be very useful when aggregation is required

from the child to parent level.

- **Channel Tables:**

  - **Channel profile Table**: This table stores all data related to a channel such as channel name, description etc. As channel has a hierarchy, this table records profile information about all channels e.g., at physician retail level, at zip level, at retail level etc.

  - **Channel hierarchy Table**: This table stores a child to parent relationship assigning a lower level number to the child and a higher level number to the parent e.g., the channel Physician Retail (PHY-RTL) belongs to the Retail (RTL) channel, here the 'PHY-RTL' channel would have a level 10 and the 'RTL' channel would have a higher level 50. This hierarchy information can be very useful when aggregation is required from the child to parent level.

- **Payor tables:**

  - **Payor profile Table**: This table stores all data related to a payor such as payor name, plan information etc. As payor can have a hierarchy, this table records profile information about all channels e.g., at physician retail level, at zip level, at retail level etc.

  - **Payor hierarchy Table**: This table stores a child to parent relationship assigning a lower level number to the child and a higher level number to the parent e.g., the payor, 'Pay 1' may belong to the payor, 'Pay all' at the national level, here the 'Pay 1' payor would have a level 10 and the 'Pay all' payor would have a higher level 50. This hierarchy information can be very useful when aggregation is required from the child to parent level.

4.2.13 File Quality Check Module

This module is used to perform quality checks on flat files present in the landing layer of SCM. File QC Module supports 2 types of checks as below:

➢ File Level Checks: The checks related to format, file extension, file names, file headers etc.

➢ File Data level checks: The checks related to file data such as unique check, data range check, domain check, length check, null check etc.

4.2.14 Base to Stage QC Module

This module is used to perform QC checks while moving data from Base to Stage. Records which are flagged are inserted in the LOG_QC_DTL Table and filtered out while moving data to stage.

4.2.15 Process QC

This framework is used to perform pre-defined process quality checks after each step is completed.

4.2.16 Fetch Dataset Module

SCM can handle both file type and database type data sources. Both could be in the same environment as your SCM machine or could be a remote FTP location or a different database. Fetch dataset module is used to fetch file/database type source to your own environment.

4.2.17 AWK Module

In case the source is a file, it may not be necessary that the source is always in the desired format. It may contain irrelevant columns, improper field ordering as compared to out processing tables.

4.2.18 Copy to Source

Once the files are on the landing layer on EC2 machine, they need to be uploaded onto the S3 bucket in order to process them and move them into the Redshift database. To achieve this, SCM has a source directory where the files are kept. The Copy to source module copies the file from land, moves them to EC2 Source folder with Standardized dataset name, and based on the split flag splits the file. In case the has_header flag is Y, the file header is removed. Once done, the sales file (Split or single) is/are moved to the corresponding source location on the S3 bucket.

4.2.19 Load Base

The file load process now can transform any input structure to any desired table structure.

Based on the number of columns in the Awked file, this module will create a temporary table (suppose Temp_Table_1) and load the data from input file to the table directly. Since this table was created in the exact same structure as the input file, the file load process is direct.

4.2.20 Adjustment Building Block

This building block adjusts sales values in terms of arithmetic operations (addition or subtraction).

For example, the values in sales data have to be adjusted for all products sold by the hospital with ID as '90573456' in the last month by decreasing the recorded sales by $100 due to returned sales.

4.2.21 Aggregation Building Block

The aggregation building block rolls up the sales data as required by the business rule.

For example, rolling up the sales data across customers to territory level.

4.2.22 Alignment Building Block

The alignment building block aligns the customer level sales information with the geography information to assign 'geography' values to each sales record and also splits sales for a customer among geographies as needed.

For example, T1 and T3 are two geographies which get 67% and 33% credit for sales from customer C1.

4.2.23 Select Insert Building Block

The exclusion building block excludes the sales records which are required based on some entities like product, subcategory etc. It flags the records which have to be excluded but doesn't remove them physically in order to increase traceability. The excluded records are however discarded just before creation of final credited sales.

For example, exclude sales from accounts, with subcategory H3, selling product, P1.

4.2.24 Fact Derivation Block

The fact derivation building block derives new facts from existing facts. It is applied row by row and the derived data is appended to the input data.

For example, conversion of sales from units to dollars or derivation of market share of each product etc.

4.2.25 Factoring Building Block

The factoring building block factors the sales figures by multiplying them by a certain number.

For example, discount sales from account with subcategory A1 by 20%.

Chapter – 5

Data Quality Assurance and Analysis

5.1 Quality Check Attributes

To perform quality checks on the data, we have special DQM (Data Quality Module) which is configured to perform various quality checks on the data. In DQM there are various tables which need to be configured for quality check.

5.1.1 File Quality Check Module

This module is used to perform quality checks on flat files present in the landing layer of SCM. File QC Module supports 2 types of checks as below:

➢ File Level Checks: The checks related to format, file extension, file names, file headers etc.

➢ File Data level checks: The checks related to file data such as unique check, data range check, domain check, length check, null check etc.

Types of Checks

➢ File Level Checks o File Name Check: Check whether the file name is as expected or not

➢ File Extension Check: Check whether the file is present with the expected extension as csv or txt or not

➢ File Delimiter Check: Check whether the file delimiter is as expected or not

➢ File Record Count Check: Check whether record count in the file matches with manifest or not

> File Header Check: Check whether file headers are as expected or not

> File Data Level Checks o Primary Check: Check whether file records are unique on expected attributes

> Data Range Checks: Check whether the values for any attribute is within the specified data range or not

> Data Domain Check: Check whether the values of any attribute are among the expected domain values or not. Domain can be defined for a maximum of 3 values

> Data Length Check: Check whether the values of any attribute are of expected length or not

> Data Null Check: Check whether the values of any attribute are not null

5.1.2 Base to Stage QC Module

This module is used to perform QC checks while moving data from Base to Stage. Records which are flagged are inserted in the LOG Table and filtered out while moving data to stage

For details of Logging Refer to section **QC Logging** above.

Types of Checks

> NULL check: Check if any entity has null value. No Parameter.

> Length check: Check if the records of a specified field are of specified length

- ➢ Type check: Check if the entity is of a required data type

- ➢ Domain check: Check if the entity value is a desired value

- ➢ Range check: Check if the Column value is between specified values

- ➢ Referential Integrity check: Check if the records of a file overlap with other. If records not found in reference Table then error

- ➢ Referential Integrity Distinct check: Same as above but LOG_QC_DTL not at row level

- ➢ Uniqueness check: Check if a field or combination of fields has got unique records. No Parameter.

5.1.3 Process QC

This framework is used to perform pre-defined process quality checks after each step is completed.

The QC Parameters are separated from each other by colon.

- ➢ Domain Check - Enter list of valid values expected

- ➢ Range Check – Min and max values expected for the column

- ➢ Type Check – Valid values are Decimal, Integer, Alphabetic, Alphanumeric

- ➢ Length Check – Min and max length of the values in the column

- ➢ Integrity Check – Reference Table and Reference Column name

Chapter – 6

RESULT ANALYSIS

4.1 Introduction

As I was the part of the sales crediting team, I majorly worked for the handling of raw data sales, configuring the AWS and the sales crediting manager tables, applying alignment, market definitions, various business rules and manual adjustments. I was also responsible for the final sales reports which we give to the IC team for IC processing.

4.2 Result Analysis

I used to deliver the 3 reports to the IC team which were basically the HCP level sales report, Territory level sales report as well the IC report.

HCP level sales report were used to calculate the HCP level sales for the hcp level analysis and to calculate the hcp level incentives and compensations.

Territory level sales reports were sales aggregated to the territory level. We also include the district, region and nation level sales. Territory level sales reports are used to calculate the payouts of the territory sales representatives, territory level sales managers, district sales managers, regional sales managers and the nation sales manager.

IC report is used for the analysis of the product how it is performing as compared to the market. To identify the rise or drop in sales as compared to the previous month or quarter sales. As well IC report are also used to identify the goals of the sales representative for the next month.
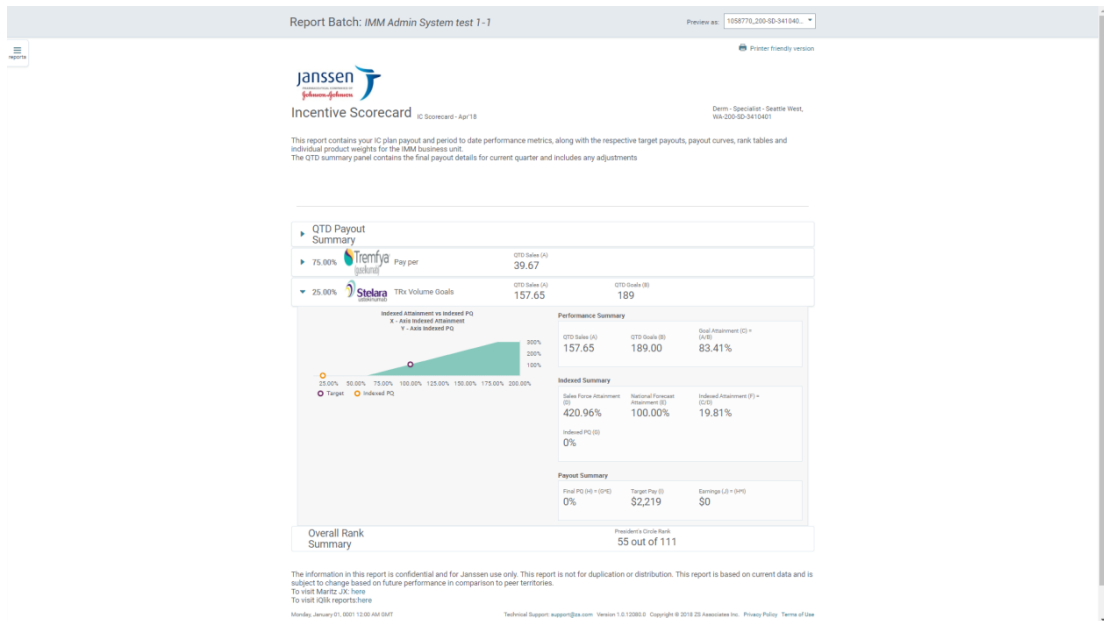
Fig. 4.1. Scorecard of reps

Chapter – 7

CONCLUSION AND FUTURE WORK

7.1 Brief summary of work

➢ Benchmark the design and performance against key industry standards (e.g. plan type, metrics, leverage, mix)

➢ Assess the program's alignment with strategic objectives through stakeholder interviews and sales force surveys

➢ Analyze the program for fairness, pay for performance, biases and fiscal responsibility

➢ Synthesize results to create actionable insights to enhance sales compensation program performance

7.2 Conclusion

The objective of the project was to prepare an error free quality reports so that there are no errors in incentives calculation and payouts and sales representatives are delivered with what they deserve. This is an on going project on month to month basis and we would be using the current established system until the client changes the business requirements.

7.3 Future scope of work

After the establishment of the sales crediting and Incentive Compensation system, we are working on the data analysis so that we can provide various insights to the client using machine learning and forecasting techniques.

REFERENCES

[1] Wikipedia – http://www.wikipedia.org/
[2] ZS Associates – https://www.zs.com/