Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ingeniería de Sistemas e Informática
Escuela Profesional de Ingeniería de Software

Implementación de un sistema de inteligencia de negocios para la estimación de cuotas y monitoreo del desempeño de la gerencia de ventas y producto de una empresa importadora

TRABAJO DE SUFICIENCIA PROFESIONAL

Para optar el Título Profesional de Ingeniero de Software

AUTOR

Liliana Alexandra RAMIREZ FERNÁNDEZ

ASESOR

Jorge Luis CHÁVEZ SOTO

Lima, Perú

2022

Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Ramirez, L. (2022). *Implementación de un sistema de inteligencia de negocios para la estimación de cuotas y monitoreo del desempeño de la gerencia de ventas y producto de una empresa importadora*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Escuela Profesional de Ingeniería de Software]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Liliana Alexandra Ramirez Fernández
Tipo de documento de identidad	DNI
Número de documento de identidad	72528869
URL de ORCID	https://orcid.org/0000-0001-5066-7821
Datos de asesor	
Nombres y apellidos	Jorge Luis Chávez Soto
Tipo de documento de identidad	DNI
Número de documento de identidad	08675814
URL de ORCID	https://orcid.org/0000-0002-9408-9266
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Espinoza Robles Armando David
Tipo de documento	DNI
Número de documento de identidad	08633326
Miembro del jurado 1	
Nombres y apellidos	Trujillo Trejo John Ledgard
Tipo de documento	DNI
Número de documento de identidad	06187585
Miembro del jurado 2 (Asesor)	
Nombres y apellidos	Jorge Luis Chávez Soto
Tipo de documento	DNI
Número de documento de identidad	08675814
Datos de investigación	
Línea de investigación	No aplica

Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	País: Perú Departamento: Lima Provincia: Lima Distrito: Cercado de Lima Jr. Carlos Amezaga No. 375 Universidad Nacional Mayor de San Marcos Latitud: -12.0564232 Longitud: -77.0843327
Año o rango de años en que se realizó la investigación	2022
URL de disciplinas OCDE	2.02.04 -- Ingeniería de sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
Escuela Profesional de Ingeniería de Software

Acta Virtual de Sustentación
del Trabajo de Suficiencia Profesional

Siendo las 19:12 horas del día 26 de agosto del año 2022, se reunieron virtualmente los docentes designados como Miembros de Jurado del Trabajo de Suficiencia Profesional, presidido por el Lic. Espinoza Robles Armando David (Presidente), Lic. Trujillo Trejo John Ledgard (Miembro) y el Lic. Chávez Soto Jorge Luis (Miembro Asesor), usando la plataforma Meet (https://meet.google.com/nvd-sypk-tbu), para la sustentación virtual del Trabajo de Suficiencia Profesional intitulado: ***"IMPLEMENTACIÓN DE UN SISTEMA DE INTELIGENCIA DE NEGOCIOS PARA LA ESTIMACIÓN DE CUOTAS Y MONITOREO DEL DESEMPEÑO DE LA GERENCIA DE VENTAS Y PRODUCTO DE UNA EMPRESA IMPORTADORA",*** *por la Bachiller* ***Ramirez Fernández Liliana Alexandra;*** *para obtener el Título Profesional de Ingeniera de Software.*

Acto seguido de la exposición del Trabajo de Suficiencia Profesional, el presidente invitó a la Bachiller a dar las respuestas a las preguntas establecidas por los miembros del Jurado.

La Bachiller en el curso de sus intervenciones demostró pleno dominio del tema, al responder con acierto y fluidez a las observaciones y preguntas formuladas por los señores miembros del Jurado.

Finalmente habiéndose efectuado la calificación correspondiente por los miembros del Jurado, la Bachiller obtuvo la nota de ***18 (DIECIOCHO).***

A continuación, el Presidente de Jurados el Lic. Espinoza Robles Armando David, declara a la Bachiller ***Ingeniera de Software****.*

Siendo las 20:00 horas, se levantó la sesión.

Presidente
Lic. Espinoza Robles Armando David

Miembro
Lic. Trujillo Trejo John Ledgard

Miembro Asesor
Lic. Chávez Soto Jorge Luis

Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería de Sistemas e Informática
Escuela Profesional de Ingeniería Software

INFORME DE EVALUACIÓN DE ORIGINALIDAD
Nº 024-EPISW-FISI-2022

1. *Autoridad Académica que emite el Informe de Originalidad:*	*Directora de la Escuela Profesional de Ingeniería de Software*
2. *Apellidos y Nombres de la autoridad académica:*	*Dra. Nora Bertha La Serna Palomino*
3. *Operador del programa informático de similitudes:*	*Dra. Nora Bertha La Serna Palomino*
3. *Documento evaluado:*	*Tesis para Pregrado* *Título: "**Implementación de un sistema de inteligencia de negocios para la estimación de cuotas y monitoreo del desempeño de la gerencia de ventas y producto de una empresa importadora**"*
5. *Autores del documento:*	*Ramirez Fernández, Liliana Alexandra*
6. Fecha de recepción de documento	*Recepción: 10/09/2022*
7. *Fecha de aplicación del programa detector de similitudes:*	*Revisión: 17/09/2022*
8. Software utilizado:	*Turnitin*
9. Configuración del programa detector de similitudes:	*Excluye textos entrecomillados: Sí* *Excluye biografías: Sí* *Excluye cadenas menores a 40 palabras: Sí* *Otro criterio (especificar): No*
10. Porcentaje de similitudes según programa detector de similitudes	*Ocho por ciento (8%)*
11. Fuentes originales de las similitudes encontradas	*Se adjuntan en 02 (dos) fojas al presente informe*
12. Observaciones:	*Ninguna*
13. Calificación de originalidad *i. Documento cumple criterios de originalidad, sin observaciones* *ii. Documento cumple criterios de originalidad, con observaciones* *iii. Documento no cumple criterios de originalidad.*	*Documento cumple criterio de originalidad, sin observación*
14. *Fecha del Informe:*	*14/11/2022*

Dra. Nora Bertha La Serna Palomino
Directora (e) de la EPISW

DEDICATORIA

Este trabajo está dedicado a mis padres y a mis hijos, Caleb y Leia, por brindarme su apoyo incondicional y por creer en mí cada día. Ustedes constituyen mi motivación principal para poder cumplir mis objetivos. Cada logro es para y por ustedes.

AGRADECIMIENTOS

Quiero agradecer, en primer lugar, a Dios por bendecirme con una familia maravillosa, por darnos vida y salud. Además, quiero agradecer de manera muy especial a mi familia, por impulsarme a ser una mejor persona y profesional.

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

ESCUELA PROFESIONAL DE INGENIERÍA DE SOFTWARE

Implementación de un sistema de inteligencia de negocios para la estimación de cuotas y monitoreo del desempeño de la gerencia de ventas y producto de una empresa importadora

Autor: **Ramirez Fernández, Liliana Alexandra**

Asesor: **Chavez Soto, Jorge Luis**

Título: **Trabajo de Suficiencia Profesional para optar el Título Profesional de Ingeniero de Software**

Fecha: **Julio de 2022**

RESUMEN

El informe de investigación plantea la implementación de una solución de Inteligencia de Negocios para la estimación de cuotas y el cálculo de resultados, además de realizar el monitoreo del desempeño de la gerencia de ventas y producto de una empresa distribuidora de productos de tecnología. Estos procesos eran realizados de forma manual, afectando otros procesos de la empresa, invirtiendo muchas horas hombres en trabajo operativo y sin tener los recursos para hacer un seguimiento del desempeño de los asociados. Por lo expuesto, se optó por implementar un sistema de Inteligencia de negocios como herramienta de gestión y monitoreo que le permita al área de inteligencia de negocios entregar resultados fiables y oportunos, mejorar la comunicación y confiabilidad de los asociados en una medición trasparente, reducir los reprocesos por correcciones y realizar un análisis de los sobrecumplimientos de los asociados para replantear las evaluaciones. En una primera etapa se consideró la creación de una base de datos centralizada, los procesos de carga y limpieza de datos de las diversas fuentes de información, la automatización del cálculo y una aplicación web que permita la publicación de resultados finales, así como el avance diario del cumplimiento de cada métrica. En una segunda etapa se consideró la implementación de un modelo predictivo para automatizar la determinación de las cuotas y un dashboard para dar seguimiento al performance de cada unidad de negocio.

Palabras claves: Inteligencia de Negocios, Modelo Predictivo, Sistema de Inteligencia de Negocios, Estimación de Cuotas, Dashboard, KPI de desempeño.

MAJOR NATIONAL UNIVERSITY OF SAN MARCOS
FACULTY OF SYSTEMS AND INFORMATICS ENGINEERING
PROFESSIONAL SCHOOL OF SOFTWARE ENGINEERING

Implementation of Business Intelligence System for forecast quotes and monitor the performance of the sales and product management of an importing company

Author: **Ramirez Fernández, Liliana Alexandra**

Advisor: **Chavez Soto, Jorge Luis**

Title: **Professional Sufficiency Work to opt for the Professional Title of Software Engineer**

Date: **July 2022**

ABSTRACT

The research report proposes the implementation of a Business Intelligence solution for quota estimation and calculation of results, in addition to monitoring the performance of the sales and product management of a company that distributes technology products. These processes were performed manually, affecting other processes of the company, investing many man-hours in operational work and without having the resources to track the performance of associates. Therefore, it was decided to implement a business intelligence system as a management and monitoring tool that would allow the business intelligence area to deliver reliable and timely results, increase communication and reliability of the associates in a transparent measurement, reduce reprocesses due to corrections and perform an analysis of the associates' overcompliance to reconsider the evaluations. In the first stage, we considered the creation of a centralized database, the processes of loading and cleaning data from the various sources of information, the automation of the calculation and a web application that allows the publication of results, as well as the daily progress of compliance with each metric. In a second stage, we considered the implementation of a predictive model to automate the determination of quotas and a dashboard to monitor the performance of each business unit.

Keywords: Business Intelligence, Predictive Model, Business Intelligence System, Quota Estimation, Dashboard, Performance KPI's.

TABLA DE CONTENIDO

RESUMEN vi

ABSTRACT vi

TABLA DE CONTENIDO vii

ÍNDICE DE FIGURAS ix

ÍNDICE DE TABLAS xi

INTRODUCCIÓN xi

CAPÍTULO I: TRAYECTORIA PROFESIONAL 3

1.1 PRESENTACION PROFESIONAL 3

1.2 FORMACION ACADÉMICA 3

1.3 EXPERIENCIA PROFESIONAL 4

CAPÍTULO II - CONTEXTO EN EL QUE SE DESARROLLA LA EXPERIENCIA 7

2.1. EMPRESA - ACTIVIDAD QUE REALIZA 7

2.2. VISIÓN 8

2.3. MISIÓN 8

2.4. ORGANIZACIÓN DE LA EMPRESA 8

2.5. ÁREA, CARGO Y FUNCIONES DESEMPEÑADAS 9

2.6. EXPERIENCIA PROFESIONAL REALIZADA EN LA ORGANIZACIÓN 9

CAPÍTULO III: ACTIVIDADES DESARROLLADAS 11

3.1. SITUACIÓN PROBLEMÁTICA 11

3.1.1. DEFINICIÓN DEL PROBLEMA 11

3.2. SOLUCIÓN 12

3.2.1. OBJETIVOS 13

3.2.2. ALCANCE 13
3.2.3. ETAPAS Y METODOLOGÍA 14
3.2.4. FUNDAMENTOS UTILIZADOS 16
3.2.5. IMPLEMENTACIÓN DE LAS ÁREAS, PROCESOS, SISTEMAS Y BUENAS PRÁCTICAS 29
3.3. EVALUACIÓN 69
3.3.1. EVALUACIÓN ECONÓMICA / EVALUACIÓN COSTO-BENEFICIO 69
3.3.2. INTERPRETACIÓN DEL VAN Y DEL TIR 72
CAPÍTULO IV. REFLEXIÓN CRÍTICA DE LA EXPERIENCIA 74
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES 76
5.1 CONCLUSIONES 76
5.2 RECOMENDACIONES 78
5.3 FUENTES DE INFORMACIÓN 79

ÍNDICE DE FIGURAS

Figura 1: Organigrama de la empresa ……..8
Figura 2: Fases del proyecto ……...……...……..…15
Figura 3: Proceso de un sistema de inteligencia de negocios……...……...17
Figura 4: Etapas de implementación del proyecto …….29
Figura 5: Arquitectura de la solución ……..……..38
Figura 6: Modelo de datos …….……...……..……...……..…39
Figura 7: Procesos de carga de datos ...……...……...……..…42
Figura 8: Paquete de carga de NPS …..……...……...……..…42
Figura 9: Paquete de carga de Ajustes …..……...……...……43
Figura 10: Paquete de carga de Marketing profit ...……...……..…43
Figura 11: Paquete de carga de Rebates …..……...……...……..…44
Figura 12: Paquete de carga de métricas del equipo preventa ……...……..…44
Figura 13: Store procedures desarrollados por métrica ..……...……..…45
Figura 14: Store procedures desarrollados para avance trimestral ……..…46
Figura 15: Cuadro resumen de total a pagar y resultados por mes …..……...……..…..50
Figura 16: Cuadro de resultados por métrica …..……...……..…51
Figura 17: Cuadro de criterios de evaluación …….……...……..…51
Figura 18: Dashboard de seguimiento de sobrecumplimientos recurrentes ……...…55
Figura 19: Dashboard de seguimiento de la compensación de la métrica de ventas ……..55
Figura 20: Dashboard de seguimiento de la compensación de la métrica de inventario …56
Figura 21: Análisis de variables cuantitativas …….……...……59
Figura 22: Análisis de gráficos multivariados para las variables cuantitativas ...……...…60
Figura 23: Análisis de variables cualitativas …….……...……..61
Figura 24: Análisis de componentes principales ……54
Figura 25: Resultados del Modelo de regresión múltiple ……...63
Figura 26: Resultados del Modelo Random Forest con data de entrenamiento …….64
Figura 27: Resultados del Modelo Random Forest con data de testeo ……..…65
Figura 28: Resultados del Modelo Random Forest con validación cruzada ……..…65

Figura 29: Resultados del Modelo Random Forest con validación cruzada con data de testeo ……66
Figura 30: Importancia de variables …….……..…66
Figura 31: Evaluación de modelos ……...…….…….....…67

ÍNDICE DE TABLAS

Tabla 1: Resumen de estudios superiores y grados académicos …… 3
Tabla 2: Resumen de estudios complementarios ……..……. 4
Tabla 3: Detalle de la experiencia laboral comprobada ……..……... 4
Tabla 4: Etapas de la metodología ágil empleada ……..……..... 16
Tabla 5: Roles del proyecto ……..……30
Tabla 6: Product Backlog del Proyecto ……..……...31
Tabla 7: Estimación y priorización del producto backlog del proyecto …..……..32
Tabla 8: HU-01 Identificación de las fuentes de datos …..……..……...34
Tabla 9: HU-02 Definir tratamiento de las fuentes de datos externas …..……...34
Tabla 10: HU-03 Creación de la base de datos …..……..……..……35
Tabla 11: Tabla de métricas y consideraciones …..……..……..……...35
Tabla 12: Tabla de Fuentes Externas …..……..……..……...37
Tabla 13: Resumen de la Reunión de Retroespectiva – Sprint 1 …..……..……39
Tabla 14: HU-04 Creación del ETL de carga …..……..……..……..40
Tabla 15: HU-05 Creación de sp para el cálculo de resultados por métrica …….40
Tabla 16: HU-06 Creación de sp para el cálculo de resultados totales a pagar ……41
Tabla 17: Resumen de la Reunión de Retroespectiva – Sprint 2 ……47
Tabla 18: HU-07 Configuración de perfiles y autenticación con Active Directory …….48
Tabla 19: HU-08 Cuadro de resultados por métrica ……48
Tabla 20: HU-09 Cuadro de resumen de resultados finales y total a pagar ……..49
Tabla 21: HU-10 Cuadro de criterios de evaluación ……..……..49
Tabla 22: Resumen de la Reunión de Retroespectiva – Sprint 3 ……..52
Tabla 23: HU-11 Dashboard de sobrecumplimientos recurrentes ……53
Tabla 24: HU-12 Dashboard de seguimiento del cumplimiento de ventas …….53
Tabla 25: HU-13 Dashboard de seguimiento del cumplimiento de inventario ……54
Tabla 26: Resumen de la Reunión de Retroespectiva – Sprint 4……...56
Tabla 27: HU-14 Implementación de un modelo predictivo supervisado ……57

Tabla 28: Definición de variables consideradas en el modelo …….57
Tabla 29: Resumen de la Reunión de Retroespectiva – Sprint 5 ……..58
Tabla 30: Evaluación económica del proyecto …….69
Tabla 31: Cuadro comparativo de beneficios entre soluciones ……69
Tabla 32: Costo del proyecto mensualizado …….72
Tabla 33: Beneficio del proyecto mensualizado ……...72
Tabla 34: Indicadores financieros del proyecto ……72

INTRODUCCIÓN

Cada vez son más las empresas que reconocen el valor de los datos en sus estrategias y apuestan en la implementación de soluciones de Inteligencia de Negocio (BI por sus siglas en inglés), las cuales hacen posible tener un mejor control y análisis de la información para soportar y facilitar la toma de decisiones. Las aplicaciones de Inteligencia de Negocios nos ayudan a tener un mejor conocimiento de la situación actual de la empresa y agilizar los procesos mejorando la productividad de las áreas.

En este sentido, el aprovechamiento inteligente de los datos nos permite optimizar el rendimiento de las áreas de negocio, garantizar una toma de decisiones estratégica y obtener ventajas competitivas aprovechando las capacidades del análisis predictivo. Tal es el caso del proceso del cálculo de la compensación variable de la gerencia de Ventas y Producto de una empresa distribuidora que realizaba el procesamiento de los datos de manera manual y poco efectiva. Este bajo nivel de gestión de la información afectaba la productividad del área responsable del proceso y no permitía tener una visión general del rendimiento de los asociados.

El objetivo de este informe es implementar un sistema de inteligencia de negocios que permita automatizar la estimación de cuotas y tener un mejor control del desempeño de la fuerza de ventas y producto de una empresa importadora. Con el fin de optimizar la precisión de cuotas, mejorar la productividad al automatizar tareas manuales, reducir el tiempo de espera de resultados y aumentar el rendimiento de los equipos al tener la información disponible para tomar decisiones basado en datos.

El presente documento está dividido en 5 capítulos:

En el capítulo I, expliqué mi experiencia profesional en base a todos los conocimientos adquiridos, las funciones y actividades que he desempeñado.

En el capítulo II, detallé el escenario donde se realizó el proyecto para el trabajo de suficiencia profesional, mencionando a la empresa y la experiencia realizada en la

organización, describí su organigrama, su visión, misión y las unidades de negocio de la empresa.

En el capítulo III, describí mi aporte dentro del proyecto, mencionando la situación problemática y la solución de este.

En el capítulo IV, realicé una crítica sobre la implementación del proyecto, detallando las lecciones aprendidas y comentando las oportunidades de mejora en un escenario actual.

En el capítulo V, desarrollé mis conclusiones y recomendaciones, además mencioné las fuentes de información y el glosario empleado.

CAPÍTULO I: TRAYECTORIA PROFESIONAL

1.1 PRESENTACION PROFESIONAL

Soy un profesional con grado de Bachiller en Ingeniería de Software, con especialización en Inteligencia de Negocios y Machine Learning. Cuento con más de 7 años de experiencia profesional en el análisis e implementación de soluciones analíticas de alto impacto en áreas de negocio en empresas del sector de consumo masivo, tecnología y gobierno. He trabajado en proyectos de transformación digital, analítica avanzada, innovación y mejora de procesos, enfocada en generar ventajas competitivas que permitan mejorar la rentabilidad y el crecimiento del negocio.

Cuento con capacidad para trabajar de forma creativa y analítica en un contexto de resolución de problemas, mentalizada siempre en cumplir todos mis objetivos, con alto sentido de responsabilidad y compromiso. Buen desempeño en actividades de trabajo en equipo, con la habilidad de interactuar a niveles técnicos y de negocio.

1.2 FORMACION ACADÉMICA

A continuación, detallaré mis estudios superiores y complementarios que constituyen mi formación académica.

Tabla 1

Resumen de estudios superiores y grados académicos

Educación Superior	Periodo
Grado académico de Bachiller en Ingeniería de Software Escuela Académico Profesional de Ingeniería de Software. Facultad de Ingeniería de Sistemas e Informática Universidad Nacional Mayor de San Marcos	2009 – 2015

Nota. Fuente: Propia

Tabla 2

Resumen de estudios complementarios

Especializaciones	Periodo
Certified Big Data & Machine Learning Professional BSG Institute	2021
Business Intelligence Application Developer SQL Server 2012 Cibertec	2016

Nota. Fuente: Propia

1.3 EXPERIENCIA PROFESIONAL

Tabla 3

Detalle de la experiencia laboral comprobada

Experiencia Profesional	Periodo
Cargo: Senior Business Intelligence Data Analyst Empresa: INGRAN MICRO S.A Funciones y actividades: - Liderar y desarrollar nuevos casos de uso de analítica alineados a los pilares estratégicos de la compañía. Además de optimizar los modelos existentes y monitorear los resultados. - Evaluar nuevas tecnologías y técnicas de machine learning para generar soluciones de alto impacto en las áreas de negocio. - Recolectar, procesar, asegurar la calidad de los datos y analizar grandes volúmenes de datos provenientes de diversas fuentes.	03/2019 - Actualidad

- Generar insights que soporten la toma estratégica de decisiones, identificando posibles sinergias entre diversas fuentes de información.
- Desarrollar de procesos de transformación de datos y automatización de procesos buscando mejorar la productividad y rentabilidad de las áreas operativas.

Cargo: Analista de Inteligencia de Negocios

05/2018 – 02/2019

Empresa: Genomma Lab. Perú S.A

Funciones y actividades:

- Desarrollar procesos de carga de datos para la consolidación de la información de ventas e inventarios de los clientes.
- Presentación semanal del avance de ventas por marca y cliente, detección de quiebres de inventario, oportunidades de mejora en punto de venta y tendencia por sku de los negocios de Perú y Bolivia.
- Implementación de indicadores que apoyen en la gestión comercial y gerencial.
- Optimización de los modelos existentes.
- Análisis del presupuesto base y nuevas proyecciones de ventas para el año en curso.
- Desarrollo de nuevos Dashboards y reportes que apoyen en la toma de decisiones

---

Cargo: Analista Programador BI

03/2015 – 05/2018

Empresa: ALIGNET SAC

Funciones y actividades:

- Desarrollar Paquetes de Extracción de las distintas fuentes de información (SSIS).

- Creación de base de datos multidimensionales que cumplan los requerimientos de información (SSAS).
- Desarrollar herramientas web de administración (Visual Basic) que complementen las soluciones de BI.
- Participación en la definición de planes de trabajo para la implementación de una solución de BI.
- Planificación de las nuevas aplicaciones o modificaciones actuales.

Cargo: Practicante de implementación de proyectos de TI

06/2014 – 12/2014

Empresa: BANCO CENTRAL DE RESERVA DEL PERÚ

Funciones y actividades:

- Programación de store procedure.
- Generación de reportes del Sistema de visitas del banco.
- Implementación de dashboard en QlikView.
- Documentación y mantenimiento de otras aplicaciones del banco.

Nota. Fuente: Propia

CAPÍTULO II - CONTEXTO EN EL QUE SE DESARROLLA LA EXPERIENCIA

2.1. EMPRESA - ACTIVIDAD QUE REALIZA

La empresa transnacional es uno de los distribuidores de tecnología más grande en el mundo, cuenta con más de 120 años de presencia en el mercado, tiene operaciones en 61 países y 188 centros de distribución. Esta empresa ofrece una amplia gama de servicios y productos de tecnología de las marcas más reconocidas del mercado. Este líder global en tecnología es reconocido cada año como socio estratégico de las principales marcas ocupando el número uno en premios regionales y mundiales. Tiene como principal foco la venta Business to Business (B2B), es decir su venta es hacia subdistribuidores, no realiza venta directa al cliente final. Entre sus principales proveedores se encuentra: Xiaomi, Apple, HP, Lenovo, Acer, Asus, Dell y MSI.

La subsidiaria en Perú inició sus operaciones hace 25 años. Actualmente, cuenta con oficinas y su centro de distribución en la cuidad de Lima, desde donde atiende a todo el país. A través de sus divisiones de negocio, son capaces de influir en la totalidad de la cadena de suministro de tecnología. A continuación, se detallan las divisiones de negocio:

- **Cloud**: La demanda de servicios empresariales en la nube no muestra signos de desaceleración. Con los servicios de valor que brinda la unidad, las empresas pueden adaptarse fácilmente a los cambios de la industria mientras monetizan cualquier servicio en un modelo de suscripción. Entre los servicios que se brindan se encuentran los siguientes tipos: Software as Service (SaaS), Infrastructure as a Service (IaaS) y Platform as a service (PaaS). Cuenta con un equipo de preventa y especialistas altamente capacidades que guían y asesoran al cliente en todo el proceso de compra.
- **Business and Consumer Solutions:** La empresa ayuda a los socios a mantenerse al día con las últimas tendencias comerciales, con su experiencia en productos y conocimientos del mercado ofrecen un amplio portafolio de productos como relojes inteligentes, drones, aspiradores inteligentes, cámaras de acción, así como también

otros productos de consumo masivo como computadoras gamers, portátiles, discos duros portátiles entre otros.

- **Mobility:** Esta unidad ofrece un amplio catálogo de dispositivos móviles de los principales líderes de la industria como XIAOMI y HUAWEI, ofreciendo productos de alta gama para los mercados de consumidores y empresas.

2.2. VISIÓN

Según la empresa en estudio su visión es ayudar a las empresas a realizar la promesa tecnológica.

2.3. MISIÓN

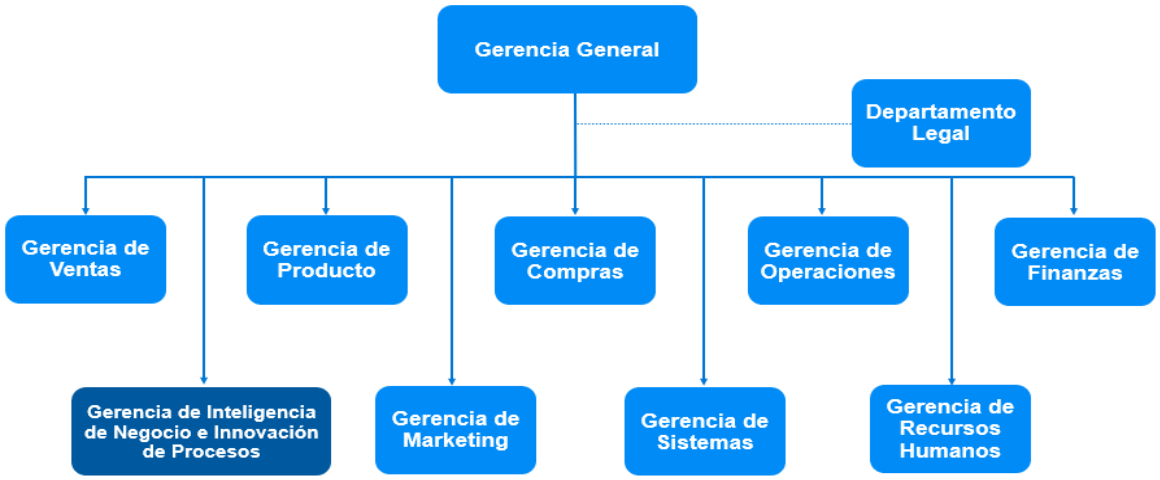
Ser el socio estratégico de nuestros clientes y proveedores, además de contribuir en su crecimiento y rentabilidad.

2.4. ORGANIZACIÓN DE LA EMPRESA

A continuación, se muestra el organigrama de la empresa mayorista, resaltando el área donde se llevó a cabo la experiencia profesional.

Figura 1

Organigrama de la Empresa



Nota. En el presente gráfico se describe la estructura organizacional y las dependencias de las gerencias de la Empresa en estudio. Elaboración Propia.

2.5. ÁREA, CARGO Y FUNCIONES DESEMPEÑADAS

Actualmente desempeño el cargo de Analista de Inteligencia de Negocios Senior en la Gerencia de Inteligencia de Negocio e Innovación de Procesos cumpliendo las siguientes funciones:

- Liderar y desarrollar nuevos casos de uso de analítica alineados a los pilares estratégicos de la compañía. Además de optimizar y calibrar los modelos que se encuentran en producción y monitorear los resultados.
- Evaluar nuevas tecnologías y técnicas de machine learning para generar soluciones de alto impacto en las áreas de negocio.
- Recolectar, procesar y asegurar la calidad de grandes volúmenes de datos provenientes de diversas fuentes.
- Generar insights que soporten la toma estratégica de decisiones, identificando posibles sinergias entre diversas fuentes de información.
- Desarrollar procesos de transformación de datos y automatización de procesos buscando mejorar la productividad y rentabilidad de las áreas operativas.
- Recibir los requerimientos de las distintas áreas y organizar junto con el área involucrada el entendimiento de su necesidad.

2.6. EXPERIENCIA PROFESIONAL REALIZADA EN LA ORGANIZACIÓN

La experiencia profesional realizada en la empresa transnacional se puede resumir en los siguientes puntos:

- Participación en las tareas de migración de servidores de base de datos y aplicaciones.
- Análisis mensual del margen comercial de la compañía, detectando desviaciones del margen y posibles errores de costos del sistema erp para asegurar que se reconozca el margen real.

- Análisis e implementación de una solución de inteligencia de negocio para el seguimiento del IGV generado cada mes con el fin de recuperar el crédito fiscal.
- Implementación de técnicas de web scraping para obtener información de fuentes externas e internas de manera automática, con el fin de enriquecer y potenciar el análisis de datos
- Análisis e innovación de procesos operativos a través de la implementación de listas de SharePoint, generación y envió de reportes automáticos, automatización de flujos de trabajo con PowerAutomate, desarrollar aplicaciones con PowerApps.
- Implementación de una página web con 2 módulos: Comisiones y Créditos. El cual permite mostrar información relevante del área de créditos y el sistema de compensación variable en base a los perfiles definidos para cada usuario.
- Implementación de una solución de analítica avanzada para determinar las proyecciones de cobranzas y pagos a proveedores, con el fin de automatizar y estimar el flujo de caja de la compañía.
- Implementación de una solución de analítica avanzada para el incremento de la línea de crédito en clientes potenciales, a través del análisis de la frecuencia de venta, comportamiento de pago y score en el sistema crediticio.
- Implementación de dashboards en PowerBI de las principales áreas del negocio que permitan dar seguimiento al avance de sus resultados versus sus objetivos y detectar desviaciones para tomar acciones estratégicas.

CAPÍTULO III: ACTIVIDADES DESARROLLADAS

3.1. SITUACIÓN PROBLEMÁTICA

3.1.1. Definición Del Problema

En la actualidad la automatización de los procesos es fundamental en todas las empresas que quieran aumentar su productividad, ya que disminuye las operaciones manuales, dando pase a que el tiempo ganado se pueda enfocar en realizar actividades que generen valor a la empresa. El proceso del cálculo de la compensación variable de la empresa en estudio era responsabilidad del área de Inteligencia de Negocios, estos procesos se realizaban de manera manual apoyados en hojas de cálculo de Excel. Las cuotas se establecían en base al promedio de ventas de los últimos cuatro trimestres poniendo un mayor peso al trimestre anterior, esto no permitía tener una proyección de cuotas precisa y confiable. La información de los resultados proviene de diferentes áreas y sistemas tanto internos como externos, el responsable de la tarea hacia el procesamiento de esta información en hojas de cálculo generando muchas horas de trabajo en la validación y limpieza de los datos. Por otro lado, los resultados de la evaluación de cada asociado se guardaban en archivos Excel considerando cortes trimestrales. Esta forma de almacenamiento impedía hacer un análisis general y comparaciones históricos de los resultados obtenidos por cada asociado, marca, unidad de negocio y gerencia.

En consecuencia, los resultados eran entregados al límite del tiempo y muchas veces con retraso al área de recursos humanos que finalmente hace la programación de pagos. En algunos casos se tuvo que postergar la fecha de pago programada de toda la planilla, por no contar con los resultados completos de los asociados con remuneración variable. Debido a que el detalle de la evaluación por métrica y total a pagar de todos los asociados se encontraba en un mismo archivo de Excel no se podía hacer público, por un tema de confidencialidad. Si un asociado quería conocer el detalle de su evaluación debía acercarse personalmente al encargado del área, esto generaba desconfianza en los asociados y muchos asociados desconocían aún las métricas que se consideraban en su medición. Además, la manera en que se almacenaban los datos hacía imposible hacer un seguimiento

del cumplimiento de los asociados, pues la información no se encontraba centralizada. Por otra parte, se generaban reprocesos por errores en los cálculos.

Por lo expuesto, con el proceso tradicional era imposible realizar el cálculo automático de resultados, realizar proyecciones de las cuotas, tener una visión general del performance de los asociados que permitiera proponer mejoras, aún menos hacer análisis y descubrir tendencias. Es fundamental contar con un sistema de inteligencia de negocios que disponga de un mecanismo adecuado para el procesamiento de información, nos permita realizar estimaciones más acertadas y hacer un análisis del cumplimiento de metas de nuestros asociados.

3.2. SOLUCIÓN

Se propone el desarrollo e implementación de un sistema de inteligencia de negocios que a través de un modelo de machine learning nos brinde estimados de ventas para la determinación de cuota de ventas. Además de realizar el cálculo de los resultados obtenidos por cada asociado de manera automática y mostrar los avances de manera diaria por cada métrica de acuerdo con el puesto del asociado. Se creó un proyecto de ETL para la consolidación y limpieza de los datos provenientes de distintas fuentes, para luego cargar la información en una base de datos centralizada. Adicionalmente, implementará una página web para mostrar los resultados alcanzados por cada métrica y el resultado final de su pago, se manejaron perfiles según el puesto de cada asociado para asegurar la confidencialidad de la información. Con el fin de hacer un seguimiento del performance de los asociados, se implementará un dashboard para analizar la información y proponer mejoras en la evaluación y métricas de los asociados.

3.2.1. OBJETIVOS

3.2.1.1 GENERAL

Implementar un sistema de inteligencia de negocios para la estimación de cuotas y monitoreo del desempeño de la gerencia de ventas y producto de una empresa importadora de productos de tecnología.

3.2.1.2 ESPECÍFICOS

- Implementar una base de datos centralizada para almacenar data histórica tanto de las cuotas como de los resultados obtenidos y totales a pagar en cada evaluación trimestral.
- Construir el proceso ETL para el procesamiento de datos e integración de las diversas fuentes de información.
- Implementar una página web como herramienta de apoyo para la publicación de resultados finales, así como el cumplimiento obtenido en cada métrica.
- Implementación de un modelo predictivo supervisado para la estimación de las cuotas.

3.2.2. ALCANCE

El proyecto comprende la implementación de una solución de Inteligencia de Negocios, para la cual se establecen los siguientes alcances:

- Diseño de los ETL para consolidar y limpiar la data proveniente de distintas fuentes de información.
- Creación de un repositorio de datos centralizado, tablas y store procedure que soporte el sistema, así como las distintas vistas que serán usadas por el aplicativo web y dashboard de seguimiento.

- Diseño y desarrollo de un dashboard para el seguimiento de resultados de los asociados en la herramienta POWER BI.
- Desarrollo de un modelo predictivo para la estimación de cuotas de venta.
- Diseño e implementación de un aplicativo web usando autenticación con el active directory para mostrar los resultados históricos, así como el cumplimiento de metas del trimestre en curso.
- Implementar pruebas para testear el flujo de carga de datos y el despliegue de la aplicación.

3.2.3. ETAPAS Y METODOLOGÍA

3.2.3.1 METODOLOGÍA

En las metodologías tradicionales, los equipos de trabajo enfocaban sus esfuerzos en la data y las tecnologías que utilizan, dejando en un segundo plano el entendimiento de las necesidades del usuario y que buscamos responder con la data disponible. En un enfoque ágil, en lugar de planificar todo por adelantado y final lanzar el sistema, permite hacer entregas pequeñas de la solución en corto plazos de manera iterativa mejorando la previsibilidad de los entregables y la retroalimentación por partes de las áreas interesadas (Tymkiw & Bournissen & Tumino, 2020).

Los entornos de inteligencia de negocios son complejos debido a los rápidos cambios de las necesidades de la información y prioridades, la existencia de muchas áreas de interés, los sobretiempos, la disponibilidad y calidad de los datos, los diferentes sistemas de origen de datos, y las tecnologías emergentes. La adopción de prácticas ágiles en un entorno de Inteligencia de Negocios es necesario para entregar productos de Inteligencia de Negocios con éxito (Overeem & Van Schaik, 2016).

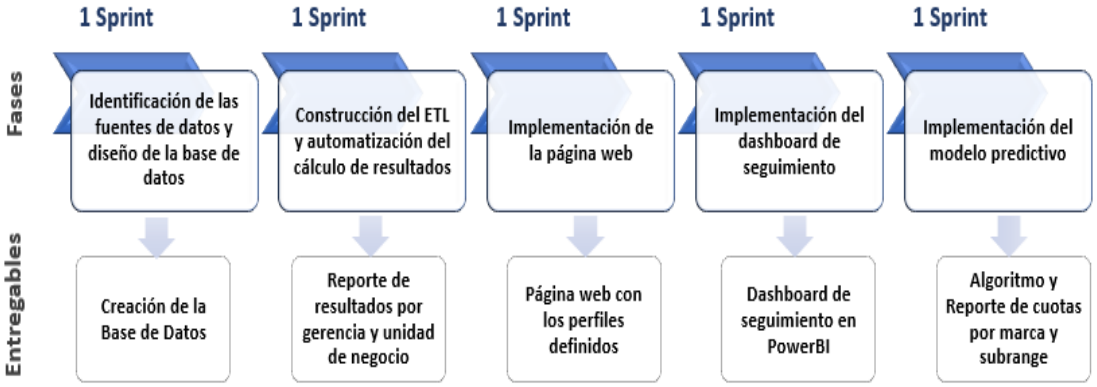
Por esta razón, para la implementación de este proyecto se ha optado por utilizar la metodología ágil Scrum, un framework que se adapta de forma eficiente a las características típicas de los proyectos inteligencia de negocios. Cabe mencionar que

Scrum no es un marco de trabajo exclusivo para proyectos de software, sino que en principio puede ser aplicado en cualquier contexto donde un equipo necesita trabajar en conjunto para lograr un objetivo en común. Para la implementación del proyecto, se tuvieron en cuenta las siguientes consideraciones:

- En la implementación de la solución se consideró 2 meses y medio.
- Se consideraron 5 sprints, cada uno duración de 15 días.
- El autor del informe de investigación estuvo presente en todas las fases de la implementación formando parte del scrum team, con el rol de Senior Business Intelligence Data Analyst.
- Se consideraron 5 fases en la implementación, que van desde la carga y procesamiento de los datos hasta la visualización de los resultados.

Figura 2

Fases del proyecto



Nota. Se muestran las fases que se consideraron en el proyecto, los sprints que comprende cada una y entregables. Elaboración Propia.

3.2.3.1 ETAPAS

Para comprender mejor como se desarrolló la solución de inteligencia de negocios, se presenta el proceso de la metodología SCRUM detallando las actividades que se consideraron en cada etapa del proyecto.

Tabla 4

Etapas de la metodología ágil empleada

Etapas	Actividades
Inicio	Entendimiento de la problemática. Definición del alcance y fases del proyecto. Definición del Scrum Team. Estimación del cronograma inicial del proyecto, así como los recursos necesarios. Creación del product backlog.
Planeamiento	Definición del objetivo del sprint y tareas consideradas en el sprint backlog. Especificación y priorización de las historias de usuario. Estimación de tiempos por cada historia de usuario. Asignación de responsables para cada historia de usuario. Definición de un plan de entrega.
Desarrollo	Implementación de las historias de usuario. Presentación de los avances en los Daily Scrum. Evaluación de los cambios y su impacto. Realización de pruebas del entregable.
Retrospectiva	Presentación de las historias implementadas al equipo. Demostración del entregable. Pruebas de aceptación.
Lanzamiento	Pase a producción. Capacitación a los usuarios.

Nota. Fuente: Propia

3.2.4. FUNDAMENTOS UTILIZADOS

3.2.4.1. SISTEMA DE INTELIGENCIA DE NEGOCIOS

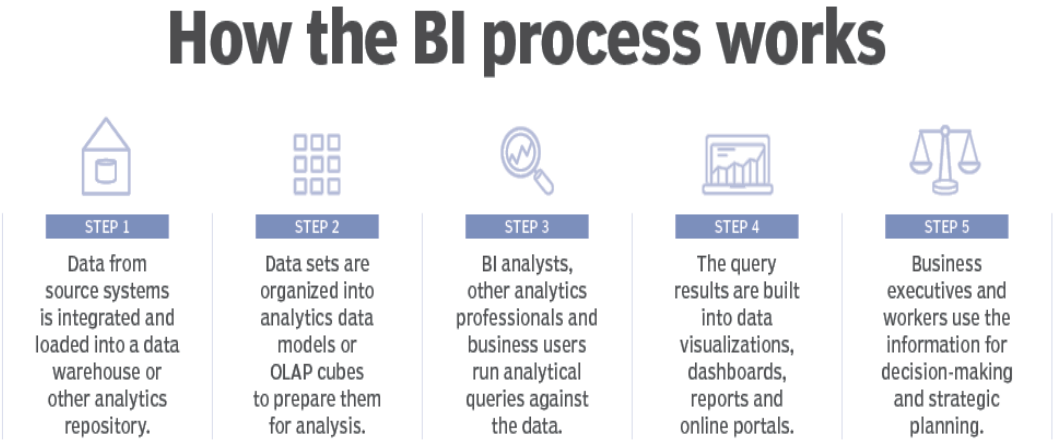
La mayoría de las empresas generan datos a cada minuto, tanto de data estructurada de sus sistemas transacciones como data no estructura como correos, comentarios en redes soluciones, facturas de proveedores, y demás. La información relevante está fluyendo en ríos virtuales de datos de manera constante y cada vez son más las empresas que toman conciencia de ello. Sin embargo, para poder sacar provecho de esos datos, se necesita contar con herramientas que permitan recopilarlos y comprenderlos. A través de

las soluciones de inteligencia de negocios, una compañía consolida, procesa y analiza los datos, y los transforma en conocimientos relevante. Estas estrategias convierten datos sin procesar en decisiones, lo cual permite que las empresas aumenten su productividad y generen ventajas competitivas. (Microsoft, 2019)

Como primer paso dentro del proceso, se deben identificar las fuentes de información con las que se disponen para el proyecto. Luego se desarrolla el flujo de trabajo para automatizar y extraer los datos desde las fuentes de origen hasta un repositorio centralizado. Los datos pueden ser información histórica o datos generados en tiempo real. Antes de presentar los datos en herramientas visuales, los datos en bruto obtenidos directamente de los sistemas transaccionales deben consolidarse y limpiarse mediante herramientas de integración de datos para asegurar la calidad de los datos, a este proceso se le denomina Extract-Transform-Load (ETL). Los datos procesados se almacenan en repositorios estructurados como datawarehouse o datamarts. Para finalmente presentar los datos en gráficos y tableros de control. (Stedman, 2020)

Figura 3

Proceso de un sistema de inteligencia de negocios



Nota. Se detallan los 5 pasos dentro del proceso de desarrollo de un sistema de inteligencia de negocios. Tomado de Stedman, Craig, Techtarget, en el informe Ultimate Guide to Business Intelligence in the Enterprise.

En la actualidad, los proyectos de Inteligencia de Negocios están incorporando nuevas herramientas de análisis avanzado, como análisis predictivo, minería de texto, análisis estadístico y análisis de big data. Un ejemplo son los modelados predictivos que permiten el análisis hipotético de diferentes escenarios comerciales. Sin embargo, cabe resaltar que los proyectos de análisis avanzado son implementados por equipos conformados por científicos de datos, ingenieros de datos, estadísticos, especialistas de machine learning y otros profesionales de análisis calificados, mientras que los equipos de BI supervisan consultas y análisis más directos de los datos comerciales.

3.2.4.2. COMPONENTES DE UN SISTEMA DE INTELIGENCIA DE NEGOCIOS

Los sistemas de inteligencia de negocios soportan todo el proceso de análisis desde la preparación de los datos hasta la explotación de estos. De acuerdo con cada solución, los componentes que se consideran pueden variar. (Stedman, 2020)

A continuación, detallaremos los componentes principales para una solución de inteligencia de negocios:

- Fuentes de datos. Estos incluyen principalmente el procesamiento de transacciones y los sistemas operativos de una organización, como ERP, CRM, finanzas, recursos humanos y cadena de suministro. Los datos de mercado, las listas de clientes externos y otros conjuntos de datos externos también se pueden recopilar.
- Almacenes de datos. Los datos de los sistemas de origen generalmente se consolidan y cargan en un almacén de datos, donde se puede acceder a ellos para usos de BI. Los datamarts más pequeños también se implementan a menudo para departamentos individuales. Además, una arquitectura de BI puede incluir un data stage operativo como área de preparación para los datos antes de que vayan al almacén de datos, así como un data lake que contiene conjuntos de grandes datos, a

menudo en formatos no estructurados o semiestructurados en lugar de los datos estructurados encontrados en el almacén.

- Las herramientas ETL (abreviatura de extraer, transformar y cargar) son el principal medio de integración y calidad de datos consolidando datos en procesos por lotes. Otros métodos de integración que se pueden emplear incluyen la integración de datos en tiempo real y la virtualización de datos, que crea vistas combinadas de conjuntos de datos sin cargarlos físicamente en un almacén de datos. La creación de perfiles y la limpieza de datos generalmente se realizan junto con la integración para identificar y corregir errores de datos.
- Tableros e informes. Los tableros en particular son una forma efectiva de distribuir información de BI a los usuarios comerciales para ayudar a impulsar la toma de decisiones y la planificación estratégica.
- BI y herramientas de análisis. Además de consultas, visualización e informes de autoservicio, las funciones de análisis que ahora se están incorporando en las plataformas de BI incluyen soporte de consultas en lenguaje natural y tecnologías impulsadas por IA que ayudan a los usuarios a encontrar y comprender datos, prepararlos para el análisis y crear visualizaciones. Las organizaciones también pueden implementar otros tipos de herramientas como software de minería de datos, sistemas de análisis en tiempo real y procesamiento analítico en línea, conocido como OLAP para abreviar, motores que almacenan datos en cubos multidimensionales para admitir análisis y cálculos sofisticados.

3.2.4.3. BENEFICIOS DE UN SISTEMA DE INTELIGENCIA DE NEGOCIOS

Un sistema de inteligencia de negocios puede ser implementada en cualquier área de una organización generando un impacto positivo. Por ejemplo, BI permite monitorear el desempeño comercial de forma continua para ejecutar acciones rápidas cuando se identifiquen desviaciones imprevistas u oportunidades. El análisis de los datos de los clientes ayuda a entender sus dolencias y enfocar los esfuerzos de marketing, ventas y servicio al cliente para mejorar la experiencia del cliente. Los cuellos de botella de la cadena de suministro y la distribución se pueden detectar antes de que ocasionen gastos

financieros. Recursos humanos puede controlar mejor el desempeño de los empleados, los costos laborales y otros datos de la fuerza laboral.

En general, los beneficios clave que las empresas pueden obtener de las aplicaciones de BI incluyen la capacidad de:

- Agilizar y mejorar la toma de decisiones.
- Optimizar los procesos comerciales internos.
- Aumentar la eficiencia operativa y la productividad.
- Detectar problemas comerciales de manera oportuna.
- Identificar tendencias comerciales y de mercado emergentes.
- Desarrollar estrategias comerciales más sólidas basadas en data.
- Impulsar mayores ventas y nuevos ingresos.
- Obtener una ventaja competitiva sobre las empresas rivales.

3.2.4.4. MACHINE LEARNING

Machine Learning es una rama de la Inteligencia de Artificial que permite programar algoritmos que aprenden por sí mismos. De igual manera, como nosotros aprendemos de nuestras experiencias y vivencias, los sistemas son capaces de aprender de los datos. Los modelos de machine learning o aprendizaje automático se entrenan con grandes cantidades de datos. Se construye un modelo a base de prueba y error, que se va ajustando a los datos y mejora su rendimiento. (Guerrero, 2019).

Los modelos de machine learning aplican métodos de análisis de datos y estadísticas para estimar resultados o futuros comportamientos. El modelado genera predicciones con un grado de probabilidad en base a las variables analizadas. En el área de marketing, por ejemplo, se han realizados proyectos para predecir el comportamiento de los consumidores y evaluar el nivel de satisfacción que se puede obtener con determinadas acciones en el mercado. (NEOATTACK, 2020)

Muchas de las empresas líderes de la actualidad, como Facebook, Google y Uber, hacen del machine learning una parte central de sus operaciones. El aprendizaje automático se ha convertido en un diferenciador competitivo importante para muchas empresas.

3.2.4.5. TIPOS DE MACHINE LEARNING

Los modelos de machine learning se clasifican según la forma en que un algoritmo aprende a ser más preciso en sus predicciones. Existen cuatro enfoques básicos: aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semisupervisado y aprendizaje por refuerzo. El tipo de algoritmo que los científicos de datos eligen usar depende del tipo de datos que quieran predecir. (Burns, 2021)

APRENDIZAJE SUPERVISADO

Los algoritmos de aprendizaje supervisado son aquellos que se utilizan para la clasificación y predicción. Debemos tener grandes cantidades de datos disponibles en los que se encuentre la variable a predecir (por ejemplo, compra o no compra). Estos datos también se denominan la variable objetivo o dependiente, ya que contienen la etiqueta (valor de resultado) para cada registro. (Burns, 2021)

Estos datos de entrenamiento son los datos a partir de los cuales el algoritmo de clasificación o predicción "aprende" o es "entrenado" sobre la relación entre las variables predictoras y la variable de objetivo. Una vez que el algoritmo ha aprendido de los datos de entrenamiento, se aplica el modelo a otra muestra de datos (los datos de validación) donde el resultado de la variable objetivo es conocido, pero no se ingresa al modelo para evaluar el sesgo entre el valor real y el valor precedido por el modelo.

Si se están probando muchos modelos diferentes, es prudente guardar una tercera muestra, que también incluye resultados conocidos (los datos de prueba) para usar con el modelo finalmente seleccionado para predecir qué tan bien funcionará. Luego, el modelo se puede usar para clasificar o predecir el resultado de interés en nuevos casos en los que se desconoce el resultado (Shmueli & Bruce & Yahav, 2018).

Los algoritmos de aprendizaje supervisado se usan en los siguientes casos:

- Clasificación binaria: Categorización de datos en dos tipos.
- Clasificación multiclase: Identificar entre más de dos tipos de categorías.
- Modelado de regresión: Estimación de valores continuos.
- Ensamblado: Combinación de las predicciones de varios modelos de aprendizaje automático para producir una predicción precisa.

APRENDIZAJE NO SUPERVISADO

En este aprendizaje automático, los algoritmos se entrenan con datos sin etiquetar. El algoritmo analiza un conjunto de datos en busca de identificar alguna relación significativa. Los datos con los que se entrenan los algoritmos, así como las predicciones o recomendaciones que generan, están predeterminados. (Burns, 2021)

Los algoritmos de aprendizaje no supervisados se pueden emplear en las siguientes situaciones:

- Agrupamiento: Selecciona los datos en grupos según la similitud de sus variables.
- Detección de anomalías: Identificación de puntos de datos inusuales en un conjunto de datos.
- Minería de asociaciones: Identificación de conjuntos de elementos en un conjunto de datos que ocurren juntos con recurrencia.
- Reducción de dimensionalidad: Reducción de la cantidad de variables.

APRENDIZAJE SEMISUPERVISADO

Este enfoque de aprendizaje automático implica una combinación de los dos tipos anteriores. Se inicia con un conjunto de datos etiquetados, es decir, se define previamente la variable objetivo a predecir. Sin embargo, el modelo es libre de explorar los datos por sí mismo y desarrollar su propia comprensión del conjunto de datos. El trabajar con un conjunto de datos etiquetado puede consumir muchos recursos y tiempo de procesamiento.

El aprendizaje semisupervisado se encuentra en un término medio entre el desempeño del aprendizaje supervisado y la eficiencia del aprendizaje no supervisado. (Burns, 2021)

Algunas situaciones donde se puede emplear el aprendizaje semisupervisado son las siguientes:

- Traducción automática: Algoritmos de enseñanza para traducir idiomas tomando como base un diccionario completo de palabras.
- Detección de fraude: Identificar casos de fraude cuando tienes unos pocos ejemplos positivos.
- Etiquetado de datos: Los algoritmos entrenados en conjuntos de datos pequeños pueden aprender a aplicar etiquetas de datos a conjuntos más grandes automáticamente.

APRENDIZAJE REFORZADO

El aprendizaje reforzado se emplea para entrenar a una máquina a completar una serie de pasos para el cual se definen reglas claras. Se definido un objetivo y un conjunto prescrito de reglas para lograr ese objetivo. Los científicos de datos también programan el algoritmo para buscar recompensas positivas, que recibe cuando realiza una acción que es beneficiosa para el objetivo final, y evitar castigos, que recibe cuando realiza una acción que lo aleja más de su objetivo final. Pero en su mayor parte, el algoritmo decide por sí mismo qué pasos tomar en el camino. (Burns, 2021) El aprendizaje por refuerzo se usa a menudo en áreas como:

- Robótica: Los robots aprenden a realizar tareas utilizando estos algoritmos.
- Videojuegos: Se ha empleado para enseñar a los bots a jugar videojuegos o para los bots de respuesta en Livechat.
- Gestión de recursos: El aprendizaje por refuerzo ayuda a las empresas a planificar cómo asignar los recursos, debido a que se cuentan con recursos finitos y es importante tener una buena gestión.

3.2.4.6. H20

H2O fue creado por la empresa H2O.ai, con el fin de combinar los principales algoritmos de Machine Learning con el Big Data. Debido a la forma de almacenamiento, procesamiento y compresión de los datos utilizando técnica de Big Data (Map/Reduce) es capaz de trabajar con grandes volúmenes de datos tanto en un único ordenador como en un cluster. H2O está desarrollado en JAVA y se puede acceder a todas sus funciones desde R, Python o Scala, así como por una interfaz web llamada Flow (Amat, 2018). Entre los algoritmos implementados en esta herramienta se destacan los siguientes:

- Kmeans
- Análisis de componentes principales
- Aggregator
- Redes Neuronales
- Random Forest
- Modelo de regresión lineal
- Modelos esamblados
- XGBoost

3.2.4.7. BENEFICIO DE H2O

Permite a los científicos de datos desarrollar proyectos de forma más rápida y eficiente, consumiendo menos tiempo en el procesamiento gracias a la manera en que comprime y almacena los datos. Además, ofrece ingeniería automática de variables, validación, ajuste, selección e implementación de modelos, interpretabilidad de aprendizaje automático, usando su propia receta, series de tiempo y generación automática de la puntuación de modelos.

Su principal ventaja frente a sus competidores es la escalabilidad. Sin embargo, sus algoritmos son igual de útiles y eficientes trabajando con volúmenes de datos grandes como pequeños.

3.2.4.8. METODOLOGÍA SCRUM

Scrum es un marco de trabajo ligero que permite a las personas, equipos de trabajo y organizaciones generar valor a través de soluciones adaptables para situaciones complejas. Se enfoca en un proceso iterativo e incremental; donde el equipo en la medida que va realizando entregas parciales, genera un valor incremental hasta culminar con el proyecto. De esta manera, permite optimizar la previsibilidad del producto y controlar el riesgo. Esta metodología consiste en los equipos scrum, sus roles, eventos periódicos y los artefactos asociados. Cada elemento tiene un propósito definido y es importante para el éxito de scrum (Schwaber & Sutherland, 2020).

ROLES

En SCRUM, se pueden identificar dos grupos: El primero se conforman por las personas comprometidas con el proyecto y proceso SCRUM, y el segundo por las personas que se afectan e interactúan con el proceso. Es importante la participación de este segundo grupo en la retroalimentación en cada revisión y en el sprint planning. (Tymkiw & Bournissen & Tumino, 2020).

Primer grupo:

- Product Owner: Es el responsable de optimizar el valor del producto, gestionar el product backlog y garantizar que los requerimientos sean comprendidos por el equipo. Debe ser consciente de las necesidades de los distintos interesados, establecer prioridades y gestionar sus expectativas. En cada sprint, el producto owner debe definir el objetivo de la iteración forma clara, y en coordinación con el scrum team.
- Scrum Master: Es el encargado de liderar el equipo para asegurar que se siga el proceso scrum. Para ello, debe brinda un constante soporte al equipo scrum, tanto de la teoría como la práctica del marco de scrum para maximizar la efectividad del equipo scrum.

- Scrum Team: Esta formado por un grupo de profesionales con diferentes habilidades y conocimientos técnicos enfocados en un objetivo en común, el objetivo del producto. El equipo scrum es el responsable de realizar todas las tareas programadas en el sprint y crear un entregable valioso en cada sprint.

Segundo grupo:

- Usuarios: son los que interactúan con el producto final.
- Stakeholders: son los que participan de las reuniones y se ven afectados por el proceso.

EVENTOS DE SCRUM

El corazón de Scrum es un sprint, que es un periodo de tiempo de un mes o menos, durante el cual se crea un incremento de producto, utilizable, y potencialmente liberable. El sprint es un contenedor para todos los eventos en Scrum. Los eventos establecidos ayudan a crear regularidad y evitar las reuniones no definidas en Scrum. A continuación, se describe cada evento:

- El Sprint Planning es una reunión que se realiza al inicio de cada sprint, donde todo el Scrum team define el objetivo del sprint. Además, se seleccionan y estiman los elementos del Product Backlog que se considerarán en el Sprint actual. El Sprint Planning tiene una duración máxima de 8 horas considerando un sprint de un mes. Para sprints más cortos, el evento suele ser más corto.
- El Daily Scrum es un tiempo de 15 minutos en donde el equipo de desarrollo presenta su avance y crea un plan para las próximas 24 horas. Estas reuniones mejoran la comunicación, permiten identificar impedimentos y ayudan a tomar decisiones rápidamente.
- Un Sprint Review es una revisión que se realiza para inspeccionar el resultado del sprint y adaptar Product Backlog si es necesario. El Scrum team presenta el resultado de su trabajo a las personas interesadas y se evalúa el cumplimiento del

objetivo. Este evento es el penúltimo dentro del sprint y podría durar máximo 4 horas.

- El Sprint Retrospective es una oportunidad para inspeccionar cómo les fue en el sprint y crear un plan de mejoras que se promulgaron durante el próximo sprint. El sprint retrospective es ultimo evento del sprint y puede durar un máximo de 3 horas.

ARTEFACTOS DE SCRUM

Los artefactos utilizados en esta metodología se describen a continuación.

- Product Backlog es una lista ordenada de todas las características y funcionalidades del producto para priorizarlas de acuerdo con las necesidades del área de negocio. De esta lista de tareas, se seleccionan historias de usuario que formarán parte del sprint backlog en el sprint planning y serán desarrolladas por el scrum team. En este punto es donde se detallan y definen aún más los elementos de trabajo pendiente del producto.
- Las historias de usuario contienen la descripción de una funcionalidad que debe incorporar al producto, y cuya implementación aporta valor al cliente.
- Sprint Backlog es el conjunto de elementos del product backlog seleccionado por el scrum team para ser completado en el sprint además realiza un plan de acción para definir como completar los objetivos seleccionados para la iteración.
- El Incremento es la suma de todos los elementos del product backlog completadas durante un sprint y todos los sprints anteriores.

3.2.4.9. SISTEMA DE GESTIÓN PARA EL CÁLCULO Y CONTROL DE LOS SERVICIOS EN CAMPO DE REDES EXTERNAS EN UNA EMPRESA DE GAS NATURAL (Velásquez, 2022)

El informe de investigación abordó la implementación de un sistema de gestión para el cálculo y monitoreo de los servicios en campo de redes externa. Puesto que el proceso inicial se realizaba de manera manual, se consideraba semanas para tener toda la

información consolidada. La información se recopilaba en campo en diferentes procesos durante todo el proceso del proyecto de redes externas, el encargado consolidaba la información. Luego de unos días, la data se enviaba a planta para su procesamiento también de manera manual en un documento Excel. En simultaneo, se generaban las evidencias del uso de los agregados en campo enviados a través del WhatsApp, las valorizaciones de igual forma se hacían de manera manual, así como el cálculo para el pago a los vendors entre otros procesos administrativos.

Este sistema ha transformado el cálculo y mejorado el control de los servicios de campo, disminuyendo el tiempo de carga de datos, cálculo de las valorizaciones, realizar los pagos a vendors, evidencias de carga, entre otros procesos que derivan de la información generada en campo en más de un 80%, mejorando los tiempos de entrega, gestionando mejor el uso de los recursos y disminuyendo los gastos administrativos en la organización.

3.2.4.10. UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA APOYAR EL ANÁLISIS DEL DESEMPEÑO DE LOS PROYECTOS EN UNA EMPRESA CONSTRUCTORA (Morante, 2021)

El informe de investigación abordó la implementación de un proyecto de inteligencia de negocios para dar soporte a la gestión de los indicadores de performance, en el departamento de gestión de proyectos, de una empresa constructora del Perú que antes emfocaba sus esfuerzo en hacer seguimiento solo s indicadores de costos, sin tomar en cuenta otros indicadores de desempeño igual de importante y significativos como tiempo (indicador del alcance de lo realizado respecto a lo estimado) y seguridad y salud (indicador de riesgo de accidentes). Los datos relacionados a estos indicadores se encontraban en diversas fuentes de datos, la consolidación, procesamiento e integración de los datos se realizaba de manera manual en hojas de cálculos. Por lo anterior, se excedía el tiempo estimado de procesamiento para la elaboración de los informes finales ocasionando dificultad para la toma decisiones de manera eficiente y rápida. Para la construcción del Datawarehouse, se consideró emplear la metodología de Ralph Kimball y la herramienta Tableau como herramienta de visualización de data. Los indicadores de Costo, Tiempo, y Seguridad y Salud fueron considerados en el análisis. Se desarrollaron 3 datamartas, los

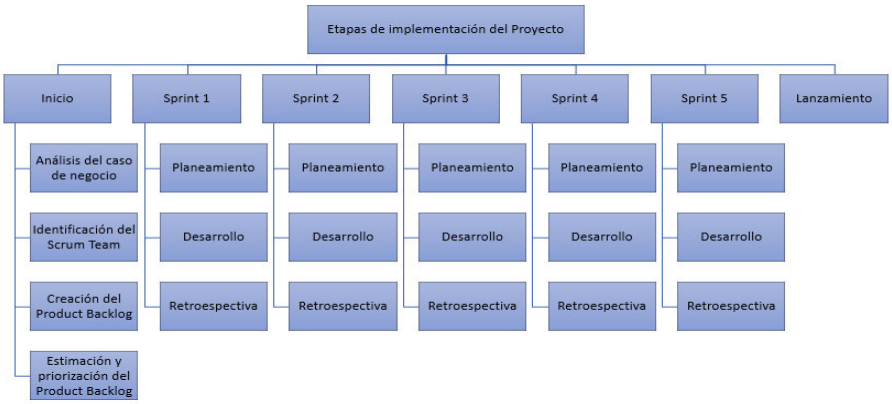
cuales contenían información consolidada y procesada. Para hacer el seguimiento de los indicadores se construyeron 3 dashboard: Gestión Operativa, Desempeño de Proyectos y Gestión de accidentes. Para la constructora era importante hacer un seguimiento de los indicadores, por lo que se proporcionó información histórica para establecer metas por cada indicador. Con el objetivo de hacer un análisis del rendimiento de nuestra propuesta, se tomaron los tiempos de procesamiento de los datos tanto del proceso anterior donde solo se contemplaba indicadores relacionados al costo del proyecto; como de la nueva propuesta donde se complementó el análisis de los indicadores de costos con los indicadores de tiempo, seguridad y salud. Del análisis anterior, se evidenció una reducción considerable e importante del tiempo del procesamiento.

3.2.5. IMPLEMENTACIÓN DE LAS ÁREAS, PROCESOS, SISTEMAS Y BUENAS PRÁCTICAS

En base a lo expuesto en la sección de etapas y metodología del presente documento, se desarrolló el proyecto bajo el enfoque de la metodología ágil Scrum en 5 fases considerando 5 sprints, los cuales se iteraron y fueron revisados con las gerencias de Producto y Ventas.

Figura 4

Etapas de Implementación del Proyecto



Nota. Se detallan las 5 etapas de metodología y los sprints que se consideraron en la implementación del proyecto. Elaboración Propia.

3.2.5.1. ETAPA INICIO

En la primera etapa del proyecto, se definió el alcance y la planificación del proyecto en base a los requerimientos de información.

3.2.5.1.1. ANÁLISIS DEL CASO DE NEGOCIO

Se realizó una revisión de las métricas empleadas en la evaluación de cada gerencia. Además de realizar entrevistas con cada responsable de las unidades de negocios, para entender el proceso anterior y recepcionar sus necesidades de información.

Como resultado del proceso anterior, se definió en coordinación con la gerencia de recursos humanos la Política de Métricas de Remuneración Variable, en donde se establecieron los lineamientos necesarios para una adecuada gestión de la compensación variable. Entre los puntos que se tomaron en cuenta en la implementación de la política se encuentran:

- Estandarizaron de las métricas consideradas por puesto, así como la definición y condiciones del cálculo de cada métrica.
- Se determinó el uso de una tabla de aceleradores para motivar las ventas, así como describir las situaciones de su no aplicación y el tratamiento de caso de especiales como protecciones.

3.2.5.1.2. IDENTIFICACIÓN DEL SCRUM TEAM

Se definió el equipo a cargo del proyecto.

Tabla 5

Roles del proyecto

Rol	Cargo
Scrum Team	Senior Business Intelligence Data Analyst
Scrum Master	Business Intelligence and Process Innovation Manager
Product Owner	Sales Manager

Nota. Fuente: Propia

El rol que desempeño la autora del proyecto fue de Senior Business Intelligence Data Analyst como parte del Scrum Team, realizando todas las tareas definidas en las historias de usuario. En la sección de etapas y metodología, se describieron las funciones y responsabilidades de cada rol.

3.2.5.1.3. CREACIÓN DEL PRODUCT BACKLOG

En base al levantamiento de requerimientos, se definieron 14 historias de usuario y se agruparon en un total de 5 iteraciones.

Tabla 6

Product Backlog del Proyecto

Fases		Historia de Usuario	Sprint
Identificación de las fuentes de datos y diseño de la base de datos	HU01	Identificación de las fuentes de datos.	Sprint 1
	HU02	Definir tratamiento de las fuentes de datos externas.	Sprint 1
	HU03	Creación de la base de datos.	Sprint 1
Construcción del ETL y automatización del cálculo de resultados	HU04	Creación del ETL de carga.	Sprint 2
	HU05	Creación de store procedures para el cálculo de resultados por métrica.	Sprint 2
	HU06	Creación de store procedures para el cálculo de resultados totales a pagar.	Sprint 2
Implementación de la página web	HU07	Como usuario necesito ingresar a la aplicación con mi usuario de red.	Sprint 3
	HU08	Como usuario necesito visualizar el resultado obtenido por cada métrica.	Sprint 3
	HU09	Como usuario necesito visualizar un cuadro resumen del total a pagar por mes.	Sprint 3
	HU10	Como usuario necesito visualizar los criterios aplicados en la evaluación.	Sprint 3
Implementación del dashboard de seguimiento	HU11	Como gerente de BI necesito identificar a los asociados con sobrecumplimiento recurrente.	Sprint 4
	HU12	Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de ventas.	Sprint 4

	HU13	Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de inventario.	Sprint 4
Implementación del modelo predictivo	HU14	Como gerente de ventas y producto necesito contar un estimado de cuotas por marcas y categoría de productos.	Sprint 5

Nota. Fuente: Propia

3.2.5.1.4. ESTIMACIÓN Y PRIORIZACIÓN DEL PRODUCT BACKLOG

Después de definir las historias de usuario, el Scrum Team realizó la priorización y estimación de horas por cada historia de usuario.

Tabla 7

Estimación y priorización del producto backlog del proyecto

	Historia de Usuario	Prioridad	Tiempo Estimado
HU01	Identificación de las fuentes de datos.	Alta	10 horas
HU02	Definir tratamiento de las fuentes de datos externas.	Alta	15 horas
HU03	Creación de la base de datos.	Alta	25 horas
HU04	Creación del ETL de carga.	Alta	25 horas
HU05	Creación de store procedures para el cálculo de resultados por métrica.	Alta	15 horas
HU06	Creación de store procedures para el cálculo de resultados totales a pagar.	Alta	10 horas
HU07	Como usuario necesito ingresar a la aplicación con mi usuario de red.	Alta	15 horas
HU08	Como usuario necesito visualizar el resultado obtenido por cada métrica.	Alta	15 horas
HU09	Como usuario necesito visualizar un cuadro resumen del total a pagar por mes.	Alta	10 horas
HU10	Como usuario necesito visualizar los	Media	10 horas

	criterios aplicados en la evaluación.		
HU11	Como gerente de BI necesito identificar a los asociados con sobrecumplimiento recurrente.	Alta	20 horas
HU12	Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de ventas.	Media	15 horas
HU13	Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de inventario.	Media	15 horas
HU14	Como gerente de ventas necesito contar un estimados de cuotas por marcas y categoría de productos.	Media	50 horas

Nota. Fuente: Propia

Para calcular el tiempo total de duración del proyecto, se debe tomar en cuenta que el recurso a cargo de la implementación dedicará 5 horas hombre de su día laboral en la implementación del proyecto, considerando que las otras 3 horas se dedican a actividades rutinarias u otra actividad solicitada. De acuerdo con lo anterior, la sumatoria de horas consumidas en cada historia de usuario da un total de 250 horas que equivalen a 50 días.

3.2.5.2. SPRINT 1

3.2.5.2.1. PLANEAMIENTO

El Sprint inicia con el Sprint Planning, reunión de planificación donde participan todos los miembros del equipo Scrum (Product Owner, Scrum Master y Scrum Team). En este evento, se define como objetivo del sprint: la identificación de las fuentes de información y la creación de la base de datos. Para esta iteración, el equipo scrum consideró del Product Backlog las siguientes historias de usuario: HU01, HU02, HU03. A continuación, se listan las tareas asociadas a las historias de usuarios de esta iteración:

- Realizar el mapeo de las fuentes de datos del proyecto.
- Realizar la clasificación de las fuentes de datos.
- Realizar reuniones con los responsables del envío de reportes de fuentes de datos externos.
- Realizar la gestión de acceso a las plataformas nuevas según sea el caso.
- Definir los plazos para el envío de datos y los formatos considerados para cada reporte.
- Crear las tablas y relaciones entre ellas del modelo de datos.

Tabla 8

HU-01 Identificación de las fuentes de datos

	HISTORIA DE USUARIO
CÓDIGO: HU01	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Identificación de las fuentes de datos	
PRIORIDAD: Alta	**ITERACIÓN**: Sprint 1
DESCRIPCIÓN: Identificar los orígenes de datos de las métricas establecidas por puestos en la Política de Remuneración Variable.	
OBSERVACIONES: Considerar los horarios de actualización de cada origen de datos para asegurar la veracidad de la información.	

Nota. Fuente: Propia

Tabla 9

HU-02 Definir tratamiento de las fuentes de datos externas

	HISTORIA DE USUARIO
CÓDIGO: HU02	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Creación de la base de datos.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 1
DESCRIPCIÓN: Identificar las métricas que provienen de fuentes externas o no formar parte de la data de BI.	
OBSERVACIONES: Gestionar los permisos a las nuevas plataformas con los gerentes y definir los formatos de envío de información.	

Nota. Fuente: Propia

Tabla 10

HU-03 Creación de la base de datos

	HISTORIA DE USUARIO
CÓDIGO: HU03	**RESPONSABLE**: Liliana Ramirez
TÍTULO: Creación de la base de datos.	
PRIORIDAD: Alta	**ITERACIÓN**: Sprint 1
DESCRIPCIÓN: Diseño y creación del modelo de datos que soporte los procesos analíticos.	
CONSIDERACIONES: Considerar que la solución de ser capaz de guardar cortes trimestrales tanto de resultados como de los parámetros considerados en cada evaluación.	

Nota. Fuente: Propia

3.2.5.2.2.DESARROLLO

Se analizaron las fuentes de datos de cada métrica establecida, se realizó el mapeo y la clasificación de cada una de ellas.

Tabla 11

Tabla de métricas y consideraciones

Métrica	Fórmula de Cálculo	Fuente
Ventas	Suma del monto vendido en el periodo (en USD)	BIPI.dbo.BO_Ventas_Consolidado
Profit	(Suma del monto vendido) – (Suma del costo neto de ventas) en el periodo (en USD)	BIPI.dbo.BO_Ventas_Consolidado
DIO (Días de Inventario)	[(valor de inventario BOH + transito) / (Costo del producto)] * número de días de venta	BIPI.dbo.BO_BalanceScoreCard
Aging 60 & 120	Valor de inventario con más de 60 y 120 días respectivamente	BIPI.dbo.BO_BalanceScoreCard
Marketing Profit (Utilidades de MKT)	Suma del monto de utilidad de marketing por eventos	Reporte enviado por correo

Rebates	Suma del monto que se obtiene por alcanzar la cuota de compra del fabricante	Reporte enviado por correo
Cobertura de Clientes	Clientes con venta en el periodo	BIPI.dbo.BO_Ventas_Consolidado
NPS	(Número de clientes Promotores / Número total de clientes encuestados) – (Número de clientes Detractores / Número total de clientes encuestados)	IM 360
Participación de clientes por la página web	(Cantidad de clientes que realizaron compras por la web en el periodo) / (Cantidad total de clientes en el periodo)	BIPI.dbo.BO_Ventas_Consolidado
Participación de ordenes por la página web	(Cantidad de facturas procesadas por la web en el periodo) / (Cantidad total de facturas procesadas en el periodo)	BIPI.dbo.BO_Ventas_Consolidado
Venta Web	Suma del monto vendido por la página web en el periodo (en USD)	BIPI.dbo.BO_Ventas_Consolidado
Certificado Vendor	Número de certificados otorgados por el vendor al asociado	Reporte enviado por correo
Nuevas Certificaciones	Número de nuevos canales certificados (habilitados) en la línea de producto.	Reporte enviado por correo
POC o/y Demos	Número de Demos o/y pruebas de conceptos (POC) realizadas en los canales.	Reporte enviado por correo
Venta Web SMB	Suma del monto vendido por la página web en el segment SMB durante el periodo (en USD)	BIPI.dbo.BO_Ventas_Consolidado
Individual	Número de cotizaciones convertidas a ordenes facturadas. Estas cotizaciones son registradas en la plataforma IM360.	IM 360

Nota. Fuente: Propia

Después de analizar y clasificar las fuentes de datos, se identificaron nuevas fuentes de datos. Por lo cual, se establecieron reuniones con las personas responsables del envió de estos resultados para validar la automatización y gestionar el acceso a estas nuevas plataformas. En algunos casos, se definió no viable la automatización de consultas de resultados de estas nuevas fuentes. Esto debido a que el proceso de las áreas se realizaba de manera manual. Se consideró definir responsables y la periodicidad del envío de información. En los otros casos, se coordinó en conjunto con las áreas involucradas la gestión de los accesos.

Tabla 12

Tabla de Fuentes Externas

Métrica	Periodo de definición	Responsable	Observaciones
Marketing Profit	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	Analista de Vendor Debit	El procesamiento de resultados por parte del área responsable es manual y no permite la automatización.
Rebates	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	Gerente de Compras	El procesamiento de resultados por parte del área responsable es manual y no permite la automatización.
NPS	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	-	Se gestionaron los accesos correctamente.
Certificado Vendor	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	Coordinador del Equipo Preventa	El procesamiento de resultados por parte del área responsable es manual y no permite la automatización.
Nuevas Certificaciones	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	Coordinador del Equipo Preventa	El procesamiento de resultados por parte del área responsable es manual y no permite la automatización.
POC o/y Demos	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	Coordinador del Equipo Preventa	El procesamiento de resultados por parte del área responsable es manual y no permite la

			automatización.
Individual	Hasta la primera semana de iniciado el siguiente periodo a evaluar.	-	Se gestionaron los accesos correctamente.

Nota. Fuente: Propia

En la siguiente figura, se muestra la arquitectura de la solución planteada considerando los tipos de fuentes de datos identificados. Tenemos los orígenes de datos que provienen del sistema transaccional SAP, esta información ya se encontraba almacenada en una base de datos SQL Server. Además de los reportes en Excel que serán compartidos por los responsables designados en la tabla anterior al cierre de cada mes y los datos provenientes de la página web de IM 360.

Figura 5

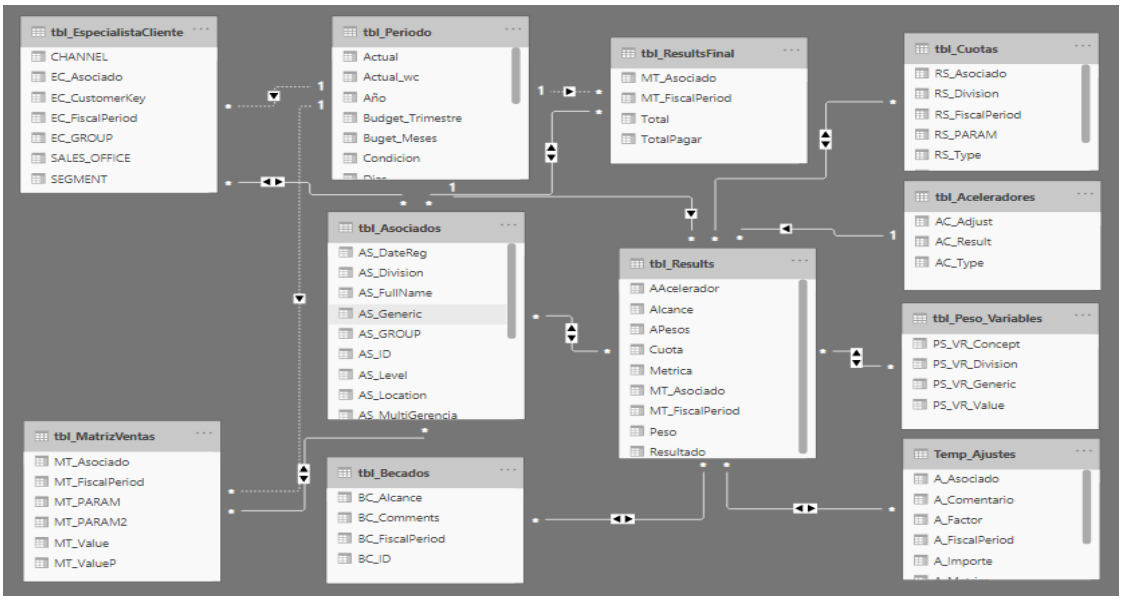
Arquitectura de la Solución

Nota. Se muestran los 4 componentes básicos de la Solución. Elaboración Propia.

Finalmente, se realizó la creación de la estructura de base de datos para soportar la solución según las consideraciones del proyecto.

Figura 6

Modelo de datos



Nota. Se muestran los tablas y relaciones consideradas en el modelo. Elaboración Propia.

3.2.5.2.3.RETROESPECTIVA

Al final de este primer sprint, se realizó la presentación de los resultados obtenidos y las tareas realizadas por el equipo de desarrollo. Se validó junto a todos los involucrados el cumplimiento de los objetivos planteados inicialmente.

Tabla 13

Resumen de la Reunión de Retroespectiva – Sprint 1

Sprint	Historia de Usuario	Comentarios
Sprint 1	Identificación de las fuentes de datos.	Todo conforme con las fuentes de datos identificadas.
Sprint 1	Definir tratamiento de las fuentes de datos externas.	Todo conforme con las consideraciones definidas para cada fuente de datos externa.
Sprint 1	Creación de la base de datos.	Se solicito asegurar que la estructura de datos permita guardar cortes trimestrales. Se solicitó considerar una tabla de ajustes para excepciones.

Nota. Fuente: Propia

Las observaciones fueron resueltas por el equipo de desarrollo antes de iniciar la siguiente iteración.

3.2.5.3. SPRINT 2

3.2.5.3.1. PLANEAMIENTO

Este segundo Sprint inició con la reunión de planificación denominada Sprint Planning. En esta reunión, participaron todos los miembros del equipo Scrum y se definió como objetivo del sprint: Construcción del proceso ETL y automatización del cálculo de resultados. Para esta iteración, el equipo scrum consideró del Product Backlog las siguientes historias de usuario: HU04, HU05 y HU06.

Tabla 14

HU-04 Creación del ETL de carga.

	HISTORIA DE USUARIO
CÓDIGO: HU04	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Creación del ETL de carga.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 2
DESCRIPCIÓN: Automatización del proceso de carga y limpieza de datos.	
CONSIDERACIONES: Considerar el formato definido para las fuentes de datos externos**.**	

Nota. Fuente: Propia

Tabla 15

HU-05 Creación de store procedures para el cálculo de resultados por métrica.

	HISTORIA DE USUARIO
CÓDIGO: HU05	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Creación de store procedures para el cálculo de resultados por métrica.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 2
DESCRIPCIÓN: Automatización del cálculo de resultados por cada métrica y	

el alcance final por métrica.
CONSIDERACIONES: Considerar la aplicación de la tabla de aceleradores según la definición de la Política de Remuneración Variable. Considerar la aplicación de ajustes en los resultados por solicitudes aprobadas por los managers respectivos.

Nota. Fuente: Propia

Tabla 16

HU-06 Creación de store procedures para el cálculo de resultados totales a pagar.

	HISTORIA DE USUARIO
CÓDIGO: HU06	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Creación de store procedures para el cálculo de resultados totales a pagar.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 2
DESCRIPCIÓN: Automatización del cálculo del total a pagar mensual y el pago del ajuste final del trimestre.	
CONSIDERACIONES: Considerar el pago al 100% de las protecciones confirmadas por recursos humanos.	

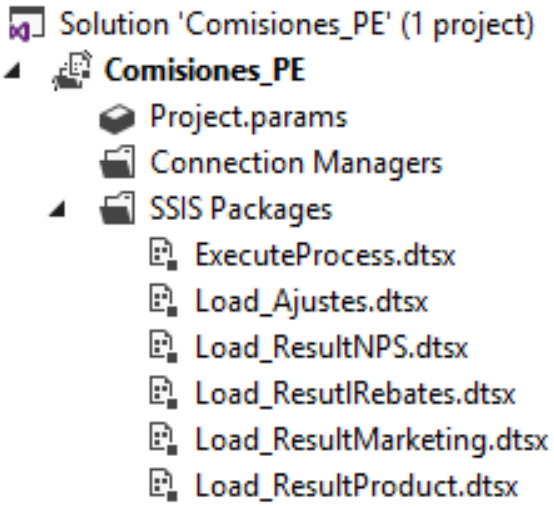
Nota. Fuente: Propia

3.2.5.3.2.DESARROLLO

La actualización de la información de este proyecto se realizó a través de un proceso ETL a través de la herramienta SQL Server Integration Services. De acuerdo con lo definido en el primer sprint, las fuentes de ventas e inventario ya se encontraban almacenadas en una base de datos propiedad del área de BI. Por lo cual, la automatización del flujo de datos se desarrolló solo para las fuentes externas pendientes y los ajustes que son excepciones validadas por los managers. Luego de implementar los paquetes de carga de datos por cada métrica pendiente, se programaron los Jobs en el SQL Agent para su ejecución automática y diaria. Se consideró como horario de ejecución las 8:30 am asegurándonos que todas las fuentes de información dependiente del proceso ya se encuentren actualizadas para ese momento.

Figura 7

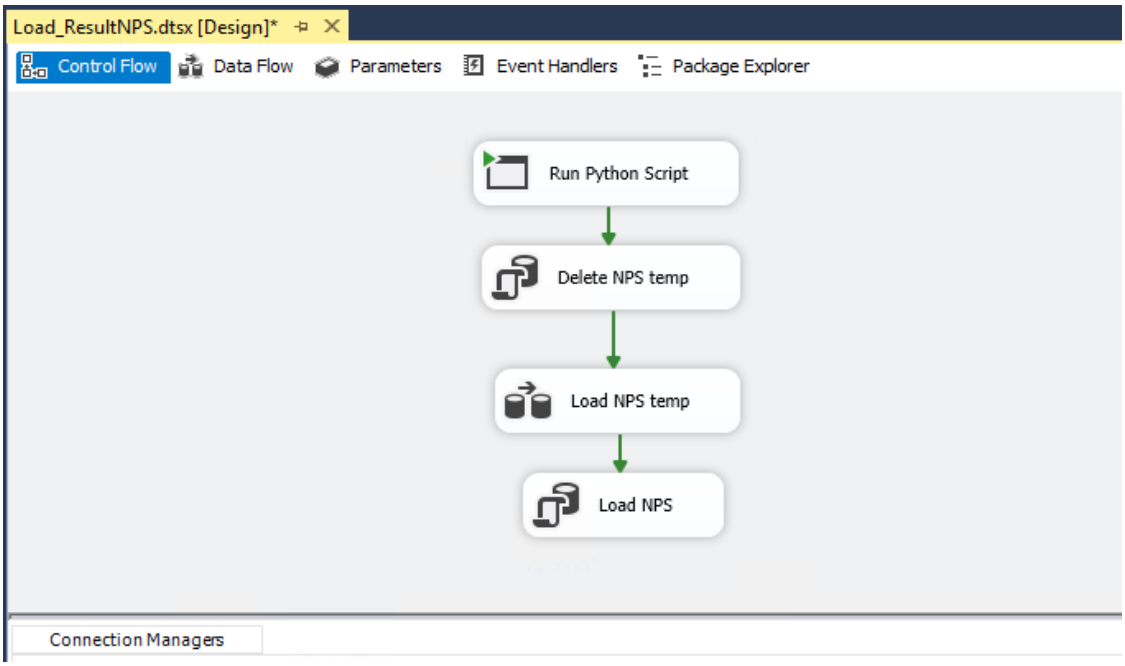
Procesos de carga de datos



Nota. Se muestran los paquetes de transformación de datos desarrollados para las fuentes externas. Elaboración Propia.

Figura 8

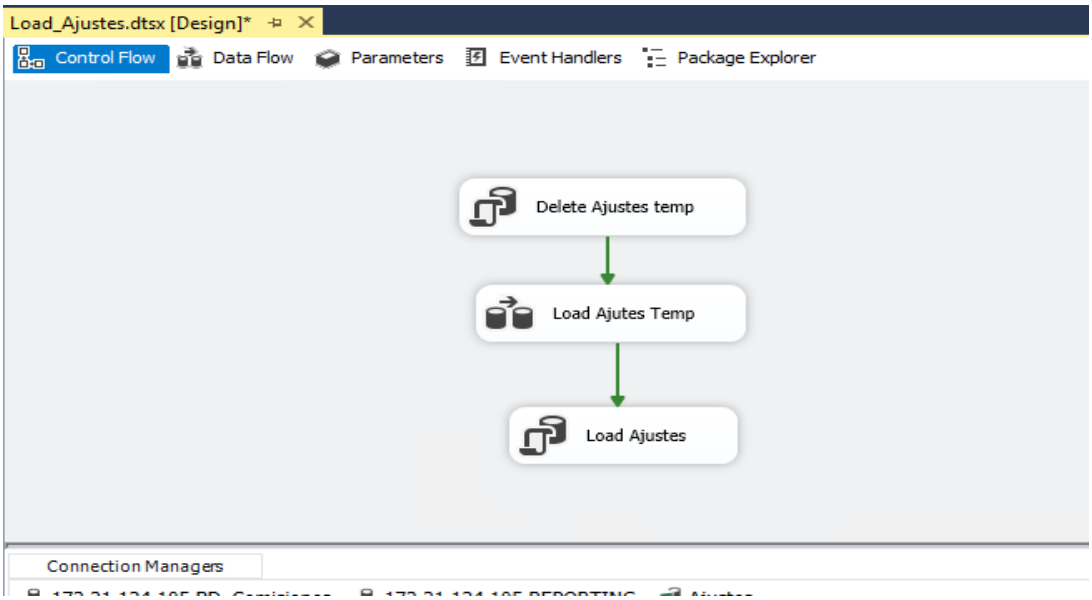
Paquete de carga de NPS



Nota. Se presenta el flujo de datos desarrollado en SSIS para cargar los resultados de NPS Elaboración Propia.

Figura 9

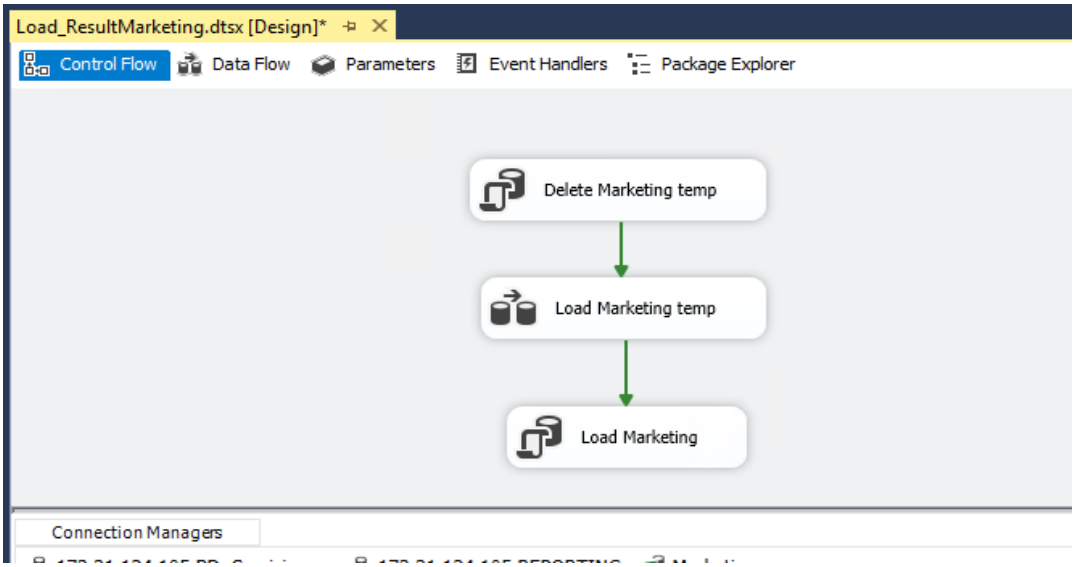
Paquete de carga de Ajustes



Nota. Se presenta el flujo de datos desarrollado en SSIS para cargar los ajustes. Elaboración Propia.

Figura 10

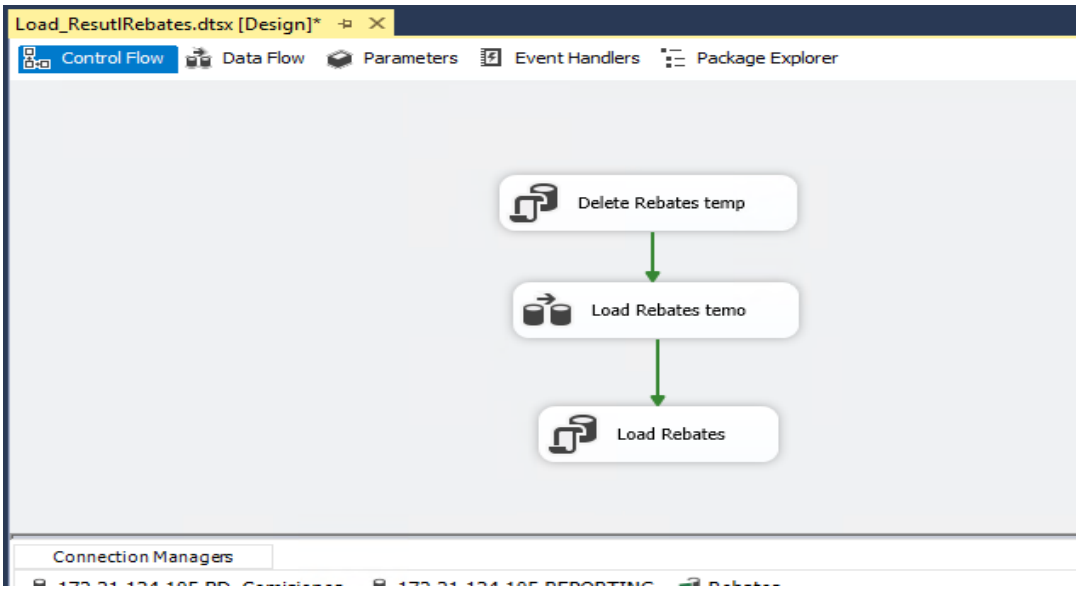
Paquete de carga de Marketing profit



Nota. Se presenta el flujo de datos desarrollado en SSIS para cargar los resultados de Marketing Profit. Elaboración Propia.

Figura 11

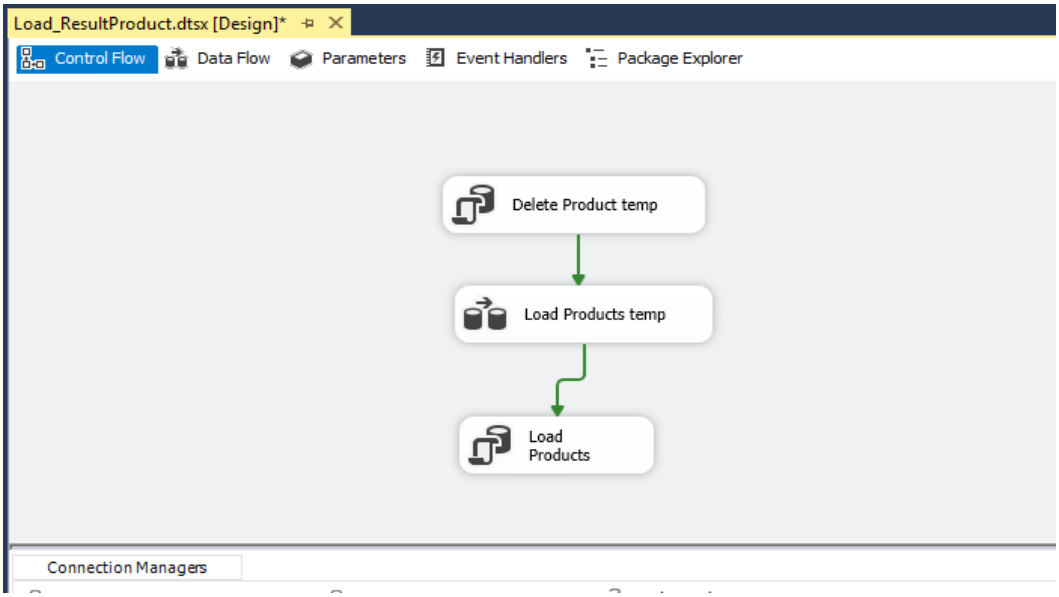
Paquete de carga de Rebates



Nota. Se presenta el flujo de datos desarrollado en SSIS para cargar los resultados de Vendors Rebates. Elaboración Propia.

Figura 12

Paquete de carga de métricas del equipo Preventa



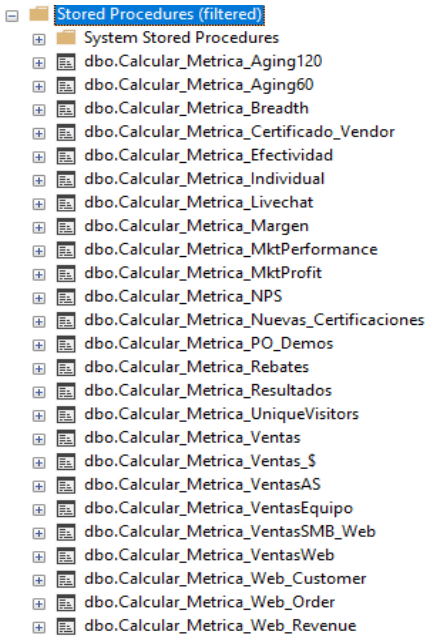
Nota. Se presenta el flujo de datos desarrollado en SSIS para cargar los resultados de Certificado Vendor, Nuevas Certificaciones, POC o/y Demos. Elaboración Propia.

Los asociados son evaluados de manera trimestral, se establecen cuotas considerando un periodo de 3 meses. Los primeros meses del trimestre se evalúa solo la venta y al final del trimestre se realiza la evaluación completa del asociado según las métricas y cuotas consideradas para su puesto. Se desarrollaron store procedures por cada métrica. Estos store procedures implementaban la lógica del cálculo según lo indicado en la ***Tabla 10***. Según lo mencionado, en los primeros meses del trimestre solo se considera la métrica de venta y se aplica unas reglas que dan como tope de pago el 90% de la variable.

Para el último mes del trimestre, se considera en el alcance por métrica la tabla de aceleradores según lo definido en la Política de Remuneración Variable. Luego de obtener el alcance considerando la tabla de aceleradores, se calcula el alcance final. Para esto, el alcance acelerado se debe multiplicar por el peso asignado a la métrica en evaluación y de esta manera obtenemos el alcance final por métrica. La suma de todos los alcances finales por métrica nos indica el resultado del trimestre.

Figura 13

Store procedures desarrollados por métrica



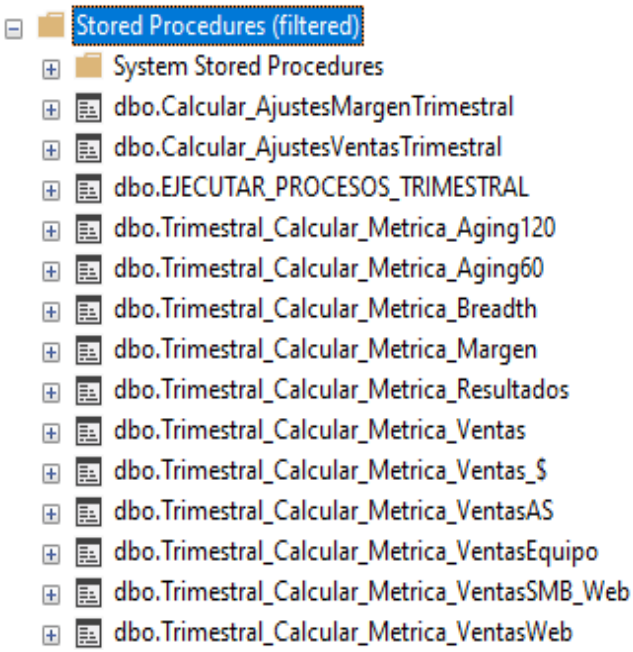
Nota. Se listan los store procedures desarrollados por cada métrica. Elaboración Propia

Los store procedures anteriores calculan el resultado del mes en curso. Es decir, en los primeros meses del trimestre donde solo se evalúa la métrica de venta, el asociado visualizaba su cuota del mes, su avance de ventas y el porcentaje de cumplimiento que representaba su avance para ese mes. De esta manera, se perdía visibilidad de saber cuánto es el porcentaje de avance considerando la cuota total del trimestre. La evaluación trimestral considerando todas las métricas solo se visualizaba hasta llegar al último mes del trimestre.

Debido a la necesidad de los interesados de visualizar la información de su evaluación proyectada, el Product Owner levantó esta observación. Se trabajó en la implementación de store procedures para los cálculos proyectados del final del trimestre. Se consideraron las métricas principales y que tenían como origen los datos previamente cargados en los repositorios de BI.

Figura 14

Store procedures desarrollados para avance trimestral



Nota. Se listan los store procedures desarrollados por el avance trimestral durante los primeros meses del trimestre. Elaboración Propia.

3.2.5.3.3.RETROESPECTIVA

Al final de este primer sprint, se realizó la presentación de los resultados obtenidos y las tareas realizadas por el equipo de desarrollo. Se validó junto a todos los involucrados el cumplimiento de los objetivos planteados inicialmente.

Tabla 17

Resumen de la Reunión de Retroespectiva – Sprint 2

Sprint	Historia de Usuario	Comentarios
Sprint 2	Creación del ETL de carga.	Todo conforme con los procesos automatizados para la carga de datos
Sprint 2	Creación de store procedures para el cálculo de resultados por métrica.	Se observó considerar resultados proyectados del final del trimestre en los primeros meses del trimestre considerando la cuota total del trimestre y el avance de resultados por métrica.
Sprint 2	Creación de store procedures para el cálculo de resultados totales a pagar.	Se observó considerar resultados proyectados del total a pagar para el fin del trimestre considerando el avance hasta la fecha.

Nota. Fuente: Propia

El equipo de desarrollo trabajó en la implementación de las observaciones y luego de ser validadas por el Scrum Team se dio pase a la siguiente iteración.

3.2.5.4. SPRINT 3

3.2.5.4.1. PLANEAMIENTO

Este tercer Sprint inició con la reunión de planificación denominada Sprint Planning. En esta reunión, participaron todos los miembros del equipo Scrum y se definió como objetivo del sprint: Implementación de la página web para la publicación de resultados y criterios de evaluación de los asociados. Para esta iteración, el equipo scrum consideró del Product Backlog las siguientes historias de usuario: HU07, HU08, HU09 y HU10.

Tabla 18

HU-07 Configuración de perfiles y autenticación con Active Directory

	HISTORIA DE USUARIO
CÓDIGO: HU07	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como usuario necesito ingresar a la aplicación con mis credenciales de red para ver la información de acuerdo con mi perfil.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 3
DESCRIPCIÓN: Implementación de una aplicación web utilizando como lenguaje de programación C# y configurar la autenticación con el Active Directory.	
CONSIDERACIONES: Considerar que la página web no será publicada, y será necesario estar conectada a la red de la empresa para acceder. Además, que se debe identificar el perfil de cada usuario para mostrar solo la información que le corresponde.	

Nota. Fuente: Propia

Tabla 19

HU-08 Cuadro de resultados por métrica.

	HISTORIA DE USUARIO
CÓDIGO: HU08	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como usuario necesito visualizar el resultado obtenido por cada métrica para poder tomar decisiones.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 3
DESCRIPCIÓN: Implementación de la vista del resultado de cada métrica: cuotas, resultado, alcance, alcance considerando la tabla de aceleradores, peso de la métrica y resultado por métrica final**.**	
CONSIDERACIONES: El resultado por métrica que se muestra en la web considera los ajustes que fueron validados por los managers. Los resultados finales por métrica son porcentajes de cumplimiento, los mismos que son sumados para obtener el resultado total del trimestre.	

Nota. Fuente: Propia

Tabla 20

HU-09 Cuadro de resumen de resultados finales y total a pagar.

	HISTORIA DE USUARIO
CÓDIGO: HU09	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como usuario necesito visualizar un cuadro resumen del total a pagar por mes.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 3
DESCRIPCIÓN: Implementación de la vista de cuadro resumen de los resultados y totales a pagar por mes.	
CONSIDERACIONES: Considerar en los filtros año y trimestre, debido a que la evaluación es trimestral y es importante ver los resultados de todos los meses en todo el trimestre de evaluación. El resultado y el total a pagar se manejan en porcentajes, debido a que solo recursos humanos tienen visibilidad del sueldo variable.	

Nota. Fuente: Propia

Tabla 21

HU-10 Cuadro de criterios de evaluación.

	HISTORIA DE USUARIO
CÓDIGO: HU10	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como usuario necesito visualizar los criterios aplicados en la evaluación.	
PRIORIDAD: Media	**ITERACIÓN:** Sprint 3
DESCRIPCIÓN: Implementación de una vista para mostrar los criterios que se aplican en la evaluación de cada asociado.	
CONSIDERACIONES: En esta vista solo se mostrarán los criterios aplicados para el trimestre en curso. No se consideró mostrar criterios históricos para evitar confusiones.	

Nota. Fuente: Propia

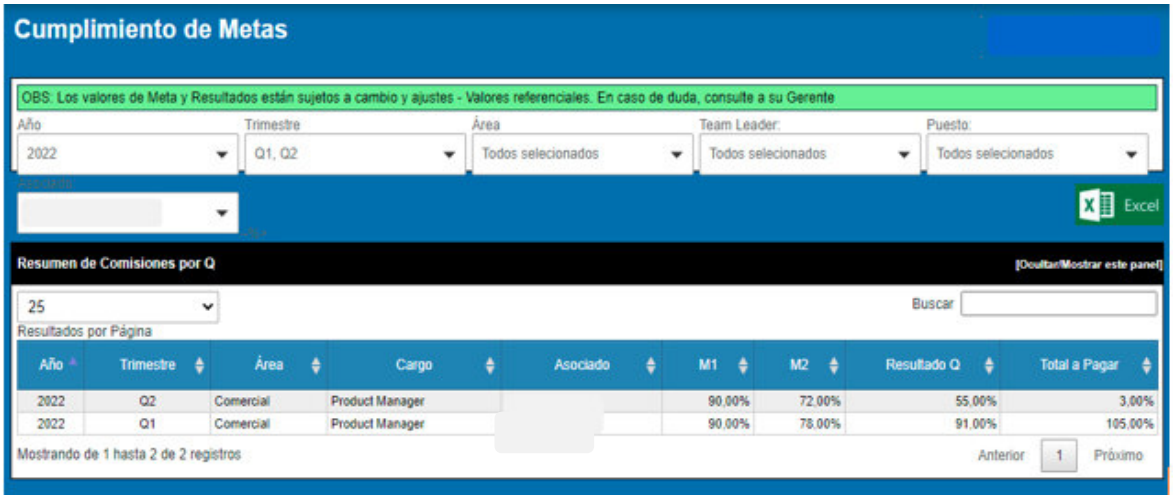
3.2.5.4.2.DESARROLLO

Como se indicó en las historias de usuario, en este sprint se realizó la implementación de página web. Se trabajó en la definición de permisos para permitir a los Team Leaders y Manager poder visualizar el resultado de los miembros de su equipo. A continuación, se listan las tareas asociadas a las historias de usuarios de esta iteración:

- Definición de perfiles de usuario e implementación de la autenticación al módulo de "Cumplimiento de Metas" utilizando las credenciales de red.
- Desarrollo de vistas en la Base de datos para ser consultadas en los cuadros de información de la aplicación.
- Integración de filtros de datos según lo definido: Año, Trimestre, Área, Team Leader, Puesto y Asociado.
- Implementación de vista de resultado final del trimestre y total a pagar. (***Figura 14***)
- Implementación de vista de resultados por métricas y el peso de esta. (***Figura 15***)
- Implementación de la vista de criterios aplicados en la evaluación. (***Figura 16***)
- Publicación de la aplicación web en el IIS en Windows Server 2012 considerando que es una intranet solo visible en la red privada de la empresa.

Figura 15

Cuadro resumen de total a pagar y resultados por mes



Nota. Se muestran los resultados obtenidos por trimestre. Elaboración Propia.

Figura 16

Cuadro de resultados por métrica

Detalle [Ocultar/Mostrar este panel]

Año	Trimestre	Mes	Asociado	Condicion	Metrica	Cuota	Resultado	Alcance	A. con Aceleradores	Peso	A. por Pesos
2022	Q2	Junio		Normal	Aging 120	46.414	104.417	44,00%	0,00%	10,00%	0,00%
2022	Q2	Junio		Normal	Aging 60	139.243	249.656	56,00%	0,00%	10,00%	0,00%
2022	Q2	Junio		Normal	Margen	$169.981	$105.641	62,00%	44,00%	40,00%	18,00%
2022	Q2	Junio		Normal	SMB	866.455	815.028	94,00%	94,00%	10,00%	9,00%
2022	Q2	Junio		Normal	Venta	$2.925.516	$2.755.011	94,00%	94,00%	30,00%	28,00%
Liliana Olivares 2022Q2 - Q											**55,00%**
2022	Q2	Mayo		Normal	Venta	$1.040.000	$748.920	72,00%	72,00%	100,00%	72,00%
Liliana Olivares 2022Q2 - M2											**72,00%**
2022	Q2	Abril		Normal	Venta	$735.516	$1.033.592	141,00%	90,00%	100,00%	90,00%
Liliana Olivares 2022Q2 - M1											**90,00%**
2022	Q1	Marzo		Normal	Aging 120	41.139	99.305	41,00%	0,00%	10,00%	0,00%
2022	Q1	Marzo		Normal	Aging 60	123.416	267.989	46,00%	0,00%	10,00%	0,00%
2022	Q1	Marzo		Normal	Margen	$173.330	$186.700	108,00%	116,00%	40,00%	46,00%
2022	Q1	Marzo		Normal	SMB	778.200	793.093	102,00%	102,00%	10,00%	10,00%
2022	Q1	Marzo		Normal	Venta	$2.570.000	$2.782.920	108,00%	116,00%	30,00%	35,00%
Liliana Olivares 2022Q1 - Q											**91,00%**
2022	Q1	Febrero		Normal	Venta	$875.000	$680.326	78,00%	78,00%	100,00%	78,00%
Liliana Olivares 2022Q1 - M2											**78,00%**
2022	Q1	Enero		Normal	Venta	$760.000	$876.160	115,00%	90,00%	100,00%	90,00%
Liliana Olivares 2022Q1 - M1											**90,00%**

PERAMI 12

Nota. Se muestran los resultados por cada métrica según el puesto. Elaboración Propia.

Figura 17

Cuadro de criterios de evaluación

Parámetros de Ventas

OBS: Los criterios que se muestran se aplican en el cáculo de comisiones en el Trimestre en curso.

Área: Todos selecionados | Team Leader: Todos selecionados | Puesto: Todos selecionados | Asociado: Todos selecionados

Excel

Parámetros de Ventas [Ocultar/Mostrar este panel]

25 Resultados por Página — Buscar

Área	Team Leader	Cargo	Asociado	Criterio	Codigo	Descripcion
BIPI		E-commerce executive		LOB	1.1-TS Comercial	1.1-TS Comercial
BIPI		E-commerce executive		LOB	1.2-TS Consumo	1.2-TS Consumo
BIPI		E-commerce executive		LOB	2-TS Advanced Solutions	2-TS Advanced Solutions
BIPI		E-commerce executive		LOB	3-TS Services	3-TS Services
BIPI		E-commerce executive		LOB	4-TS Specialty	4-TS Specialty
BIPI		E-commerce executive		LOB	5-TS Mobility	5-TS Mobility
BIPI		E-commerce executive		LOB	6-Cloud	6-Cloud
Comercial		SalesSpecialist		Subrange	459601	Lenovo Comercial Tbg Accesorios
Comercial		SalesSpecialist		Subrange	459603	Lenovo Comercial Tbg Dt
Comercial		SalesSpecialist		Subrange	459604	Lenovo Comercial Tbg Dt Ws
Comercial		SalesSpecialist		Subrange	459605	Lenovo Comercial Tbg Monitores
Comercial		SalesSpecialist		Subrange	459608	Lenovo Comercial Tbg Tablets

Nota. Se muestran los parámetros considerados en la evaluación de cada asociado. Elaboración Propia.

3.2.5.4.3.RETROESPECTIVA

Al final de este primer sprint, se realizó la presentación de los resultados obtenidos y las tareas realizadas por el equipo de desarrollo. Se validó junto a todos los involucrados el cumplimiento de los objetivos planteados inicialmente y se levantaron observaciones que fueron trabajadas antes de pasar al siguiente sprint.

Tabla 22

Resumen de la Reunión de Retroespectiva – Sprint 3

Sprint	Historia de Usuario	Comentarios
Sprint 3	Como usuario necesito ingresar a la aplicación con mi usuario de red para ver la información de acuerdo con mi perfil.	Todo conforme con esta historia de usuario.
Sprint 3	Como usuario necesito visualizar el resultado obtenido por cada métrica para poder tomar decisiones.	Se observó agregar la función de exportar datos.
Sprint 3	Como usuario necesito visualizar un cuadro resumen del total a pagar por mes.	Se observó agregar la función de exportar datos.
Sprint 3	Como usuario necesito visualizar los criterios aplicados en la evaluación.	Todo conforme con esta historia de usuario.

Nota. Fuente: Propia

El equipo de desarrollo trabajó en las observaciones y fueron validadas antes de iniciar la siguiente iteración.

3.2.5.5. SPRINT 4

3.2.5.5.1. PLANEAMIENTO

Este cuarto Sprint inició con la reunión de planificación denominada Sprint Planning. En esta reunión, participaron todos los miembros del equipo Scrum y se definió como objetivo del sprint: Implementación del dashboard de seguimiento.

Para esta iteración, el equipo scrum consideró del Product Backlog las siguientes historias de usuario: HU11, HU12 y HU13.

Tabla 23

HU-11 Dashboard de sobrecumplimientos recurrentes

	HISTORIA DE USUARIO
CÓDIGO: HU11	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como gerente de BI necesito identificar a los asociados con sobrecumplimiento recurrente.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 4
DESCRIPCIÓN: Creación de objetos visuales que ayude a navegar en los datos e identificar los asociados con sobrecumplimiento continuo.	
CONSIDERACIONES: Los resultados se actualizó de manera diaria.	

Nota. Fuente: Propia.

Tabla 24

HU-12 Dashboard de seguimiento del cumplimiento de ventas

	HISTORIA DE USUARIO
CÓDIGO: HU12	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de ventas.	
PRIORIDAD: Media	**ITERACIÓN:** Sprint 4
DESCRIPCIÓN: Creación de objetos visuales que ayude a navegar en los datos e identificar los asociados que no cumplieron la cuota de venta, pero lo compensaron con otras métricas.	
CONSIDERACIONES: Los resultados se actualizan de manera diaria.	

Nota. Fuente: Propia

Tabla 25

HU-13 Dashboard de seguimiento del cumplimiento de inventario

	HISTORIA DE USUARIO
CÓDIGO: HU13	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de inventario.	
PRIORIDAD: Media	**ITERACIÓN:** Sprint 4
DESCRIPCIÓN: Creación de objetos visuales que ayude a navegar en los datos e identificar los asociados que no cumplieron la cuota de inventario, pero lo compensaron con otras métricas.	
CONSIDERACIONES: Los resultados se actualizan de manera diaria.	

Nota. Fuente: Propia

3.2.5.5.2.DESARROLLO

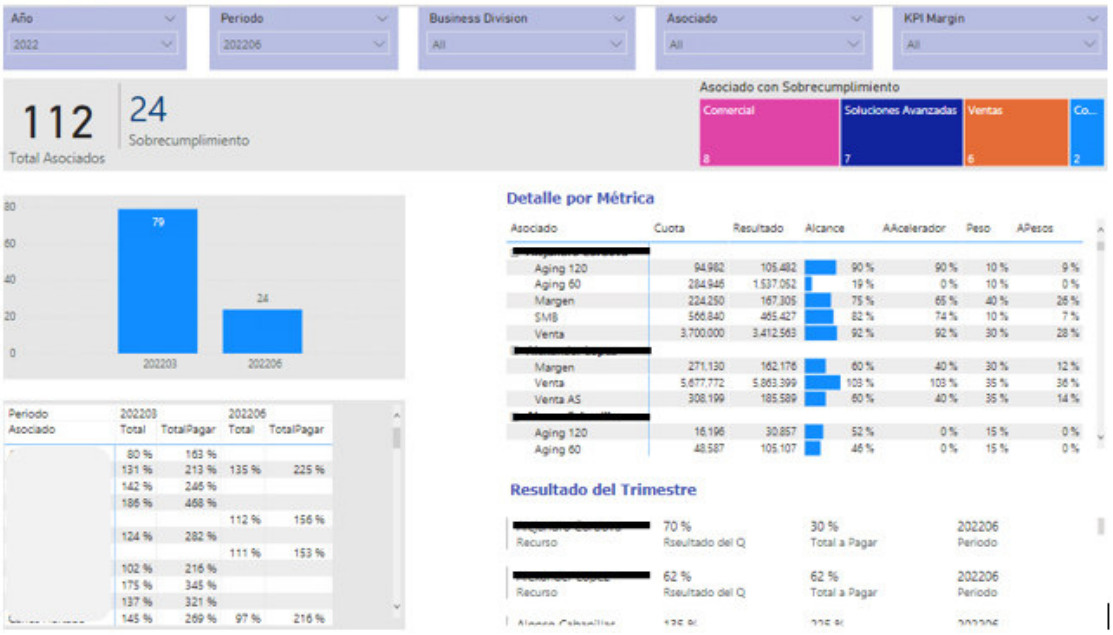
Como se indicó en las historias de usuario, en este sprint se realizó la creación del dashboard para el monitoreo del performance de los asociados. Para guiar la navegación de la información, se definieron los siguientes filtros: Año, Trimestre, División de Negocio, Asociado y KPI. En el caso del campo de KPI nos ayuda para identificar si los asociados cumplían con la condición en evaluación, es decir, presenta sobrecumplimiento, compensación de la métrica de ventas y/o compensación de la métrica de inventario.

A continuación, se listan las tareas asociadas a las historias de usuarios de esta iteración:

- Cargar los datos y definir las relaciones de las tablas en el dashboard en Power BI.
- Creación de fórmulas DAX.
- Implementación del dashboard de sobrecumplimientos recurrentes. (***Figura 18***)
- Implementación del dashboard de monitoreo de la métrica de venta. (***Figura 19***)
- Implementación del dashboard de monitoreo de la métrica de inventario. (***Figura 20***)

Figura 18

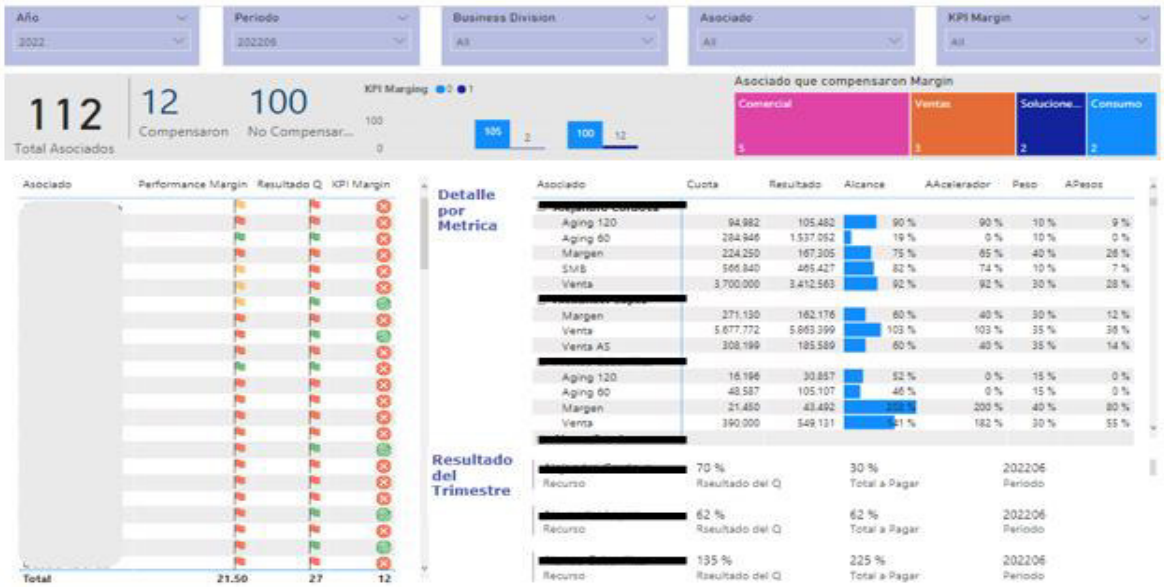
Dashboard de seguimiento de sobrecumplimientos recurrentes.



Nota. Análisis de los asociados con sobrecumplimiento continuo. Elaboración Propia.

Figura 19

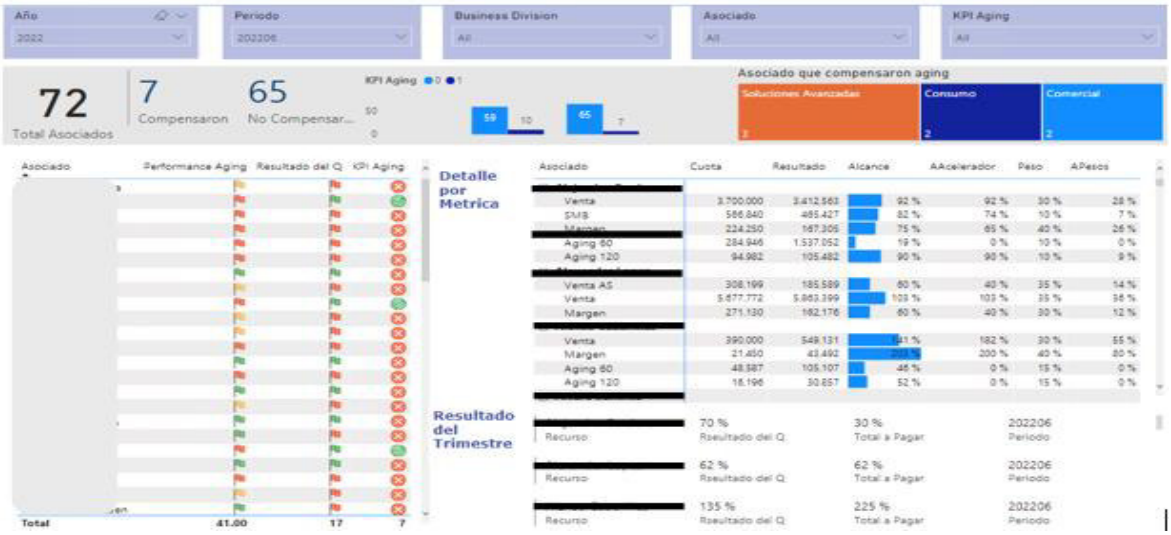
Dashboard de seguimiento de la compensación de la métrica de ventas.



Nota. Análisis del performance de asociados que no cumplieron la métrica de ventas. Elaboración Propia.

Figura 20

Dashboard de seguimiento de la compensación de la métrica de inventario.



Nota. Análisis del performance de asociados que no cumplieron la métrica de inventario. Elaboración Propia.

3.2.5.5.3.RETROESPECTIVA

Al final de este primer sprint, se realizó la presentación de los resultados obtenidos y las tareas realizadas por el equipo de desarrollo. Se validó junto a todos los involucrados el cumplimiento de los objetivos planteados inicialmente y se levantaron observaciones que fueron trabajadas antes de pasar al siguiente sprint.

Tabla 26

Resumen de la Reunión de Retroespectiva – Sprint 4

Sprint	Historia de Usuario	Comentarios
Sprint 4	Como gerente de BI necesito identificar a los asociados con sobrecumplimiento recurrente.	Todo conforme con esta historia de usuario.

Sprint 4	Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de ventas.	Todo conforme con esta historia de usuario.
Sprint 4	Como gerente de BI necesito evaluar a los asociados que compensan el no cumplimiento de métricas de inventario.	Todo conforme con esta historia de usuario.

Nota. Fuente: Propia

3.2.5.6. SPRINT 5

3.2.5.6.1. PLANEAMIENTO

Este quinto Sprint inició con la reunión de planificación denominada Sprint Planning. En esta reunión, participaron todos los miembros del equipo Scrum y se definió como objetivo del sprint: Implementación de un modelo predictivo supervisado para la estimación de cuotas de ventas. Como resultado se obtiene un reporte de cuotas por Division de Negocio y Marca. Para esta iteración, el equipo scrum consideró del Product Backlog las siguientes historias de usuario: HU14.

Tabla 27

HU-14 Implementación de un modelo predictivo supervisado

	HISTORIA DE USUARIO
CÓDIGO: HU14	**RESPONSABLE:** Liliana Ramirez
TÍTULO: Como gerente de ventas necesito contar un estimados de cuotas por marcas y categoría de productos.	
PRIORIDAD: Alta	**ITERACIÓN:** Sprint 5
DESCRIPCIÓN: Implementación de un modelo predictivo usando técnicas de Machine Learning que permitan automatizar la predicción de cuotas de ventas.	
CONSIDERACIONES: Se deben evaluar por lo menos 3 modelos para determinar cuál es que nos da mejores resultados, considerando el indicador MAE.	

Nota. Fuente: Propia

A continuación, se listan las tareas asociadas a las historias de usuarios de esta iteración:

- Entendimiento de los datos.
- Preparación de los datos.
- Modelamiento y Evaluación.

3.2.5.6.2. DESARROLLO

a) ENTENDIMIENTO DE LOS DATOS

En primer lugar, los datos son explorados en profundidad para entender sus comportamientos, identificar irregularidades o falta de data. Con el fin de asegurar la calidad de los datos y por consecuente del modelo. Se recopilan los datos de la tabla de ventas de SAP. A continuación, se describen los campos:

Tabla 28

Definición de variables consideradas en el modelo

Campo	Tipo	Esquema	Descripción
Fecha	Date	DD/MM/AAAA	Contiene la fecha de la demanda de una marca x
AÑO	Numérico de 4 dígitos	AAAA	Contiene el año de la demanda de una marca x
MES	Numérico de 2 dígitos	MM	Contiene el mes de la demanda de una marca x
BUSINESS DIVISION	Cadena	Consumo, Mobility, Soluciones Avanzado, Cloud y Comercial	División de la gerencia comercial tomando en cuenta la naturaleza de los productos ofertados.
brand	Cadena	XIAOMI, HP,	Marca de nuestros

		DELL, LENOVO, APPLE	proveedores
NET VALUE PROM 3 MESES	Numérico		Venta promedio de los 3 últimos meses posteriores
NET VALUE PROM 6 MESES	Numérico		Venta promedio de los 6 últimos meses posteriores
NET VALUE PROM 12 MESES	Numérico		Venta promedio de los 12 últimos meses posteriores
NET VALUE	Numérico		Venta en el periodo determinado

Nota. Fuente: Propia

NET VALUE es el número por predecir, este campo se define como la variable dependiente en las predicciones. Luego, haremos un análisis de las variables cuantitativas: NET VALUE PROM 3 MESES, NET VALUE PROM 6 MESES, NET VALUE PROM 12 MESES y NET VALUE.

Figura 21

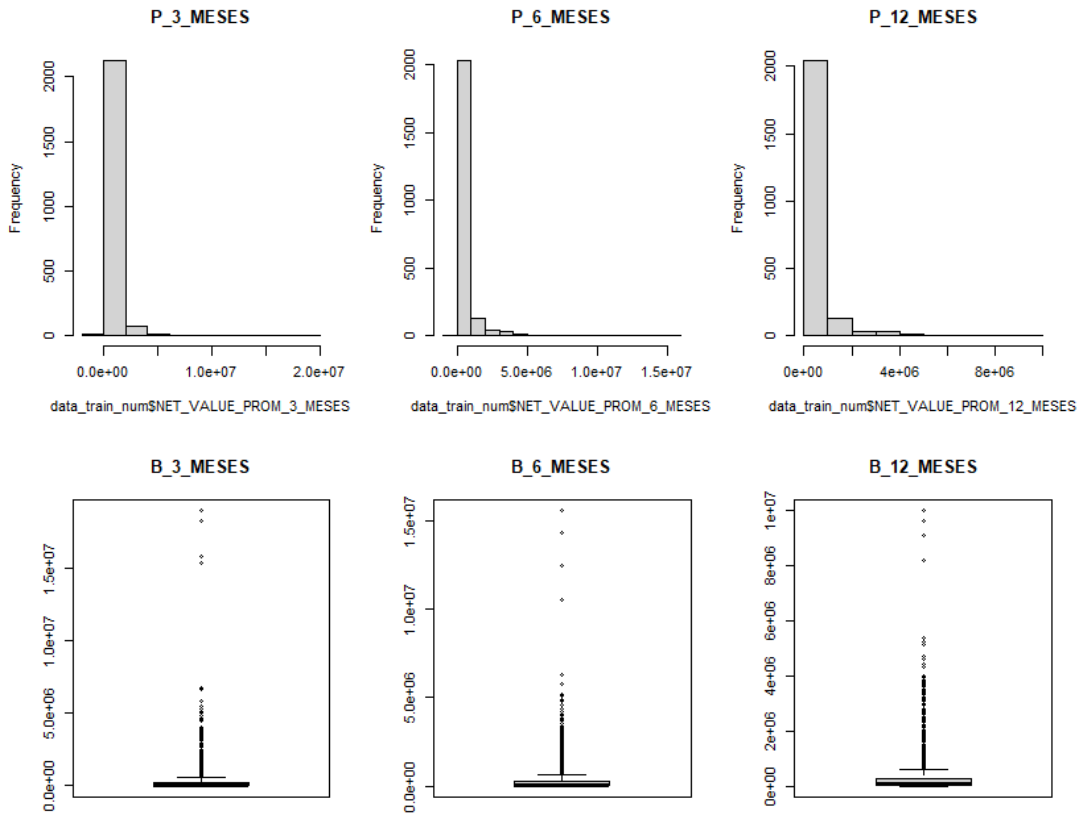
Análisis de variables cuantitativas.

```
> var(data_train_num2$NET_VALUE_PROM_3_MESES)
[1] 971944273930
> var(data_train_num2$NET_VALUE_PROM_6_MESES)
[1] 725143506780
> var(data_train_num2$NET_VALUE_PROM_12_MESES)
[1] 522680082982
> kurtosis(data_train_num2$NET_VALUE_PROM_3_MESES)
[1] 160.8201
> kurtosis(data_train_num2$NET_VALUE_PROM_6_MESES)
[1] 113.18
> kurtosis(data_train_num2$NET_VALUE_PROM_12_MESES)
[1] 53.05124
> skewness(data_train_num2$NET_VALUE_PROM_3_MESES)
[1] 10.29752
> skewness(data_train_num2$NET_VALUE_PROM_6_MESES)
[1] 8.338492
> skewness(data_train_num2$NET_VALUE_PROM_12_MESES)
[1] 5.774463
```

Nota. Se muestran los indicadores de forma, simetría y variabilidad de las variables cuantitativas. Elaboración Propia.

Figura 22

Análisis de gráficos multivariados para las variables cuantitativas.



Nota. Evaluaciones las variables predecesoras. Elaboración Propia.

Luego de analizar nuestras variables cuantitativas, destacamos la alta varianza de las variables lo cual afecta directamente en el rendimiento de nuestro modelo y debemos controlarla. Los gráficos de boxplot y el histograma nos ayudan para validar la distribución de los datos y alértanos de la presencia de outliers. De la ***figura 22***, podemos decir que el 75% de mis marcas manejan volúmenes de ventas moderado, mientras solo el 25% son las marcas con grandes volúmenes.

En el caso de las variables cualitativas, no se puede evaluar la correlación. Sin embargo, podemos utilizar en su defecto las tablas de contingencia.

Figura 23

Análisis de variables cualitativas.

```
> chisq.test(data_venta$BUSINESS_DIVISION, data_venta$BRAND)

        Pearson's Chi-squared test

data:  data_venta$BUSINESS_DIVISION and data_venta$BRAND
X-squared = 7799, df = 340, p-value < 2.2e-16
```

Nota. Se muestran los indicadores de forma, simetría y variabilidad de las variables cuantitativas. Elaboración Propia.

De la ***figura 23***, se determina que las variables si están relacionadas, lo cual puede ser perjudicial para nuestro modelo.

b) PREPACIÓN DE LOS DATOS

Inicialmente solo se consideró las siguientes variables para el modelo: Fecha, Año, Mes, Business Division, Brand y Net Value. Después, agregamos 3 variables a nuestro modelo, las cuales se generan tomando como base la variable NET_VALUE. De esta manera buscamos enriquecer nuestro modelo y mejorar la predicción. Las variables generadas fueron: NET_PROM_3_MESES, NET_PROM_6_MESES y NET_PROM_12_MESES.

Como se observa en la ***figura 21***, las variables cuantitativas presentan datos outliers. Primero realizaremos la transformación de los datos outliers a nulos. Con esta acción, tenemos un 2.5% de datos vacíos. Seguidamente de transformar los datos, procedemos a imputar los valores vacíos por números aleatorios dentro del rango 25% y 75%. Se presenta una reducción de los datos outliers. Sin embargo, no se eliminan por completo.

En el caso de las variables cualitativas, se utilizó el método dummy para transformar las variables cualitativas BUSINESS DIVISION y BRAND a variables cuantitativas, con lo cual se generan 87 variables.

c) MODELAMIENTO Y EVALUACIÓN

Para disminuir varianza de los datos aplicamos el análisis de componentes principales.

Figura 24

Análisis de componentes principales.

```
                          PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8     PC9    PC10    PC11
Standard deviation     1.82408 1.44568 1.36514 1.30540 1.15773 1.11730 1.04780 1.02052 1.01888 1.01880 1.01765
Proportion of Variance 0.03656 0.02297 0.02048 0.01873 0.01473 0.01372 0.01206 0.01144 0.01141 0.01141 0.01138
Cumulative Proportion  0.03656 0.05953 0.08001 0.09874 0.11346 0.12718 0.13925 0.15069 0.16210 0.17351 0.18489
                          PC12    PC13   PC14    PC15    PC16    PC17    PC18   PC19    PC20    PC21    PC22
Standard deviation     1.01587 1.01524 1.0142 1.01264 1.01238 1.01187 1.01036 1.0098 1.00910 1.00841 1.00788
Proportion of Variance 0.01134 0.01133 0.0113 0.01127 0.01126 0.01125 0.01122 0.0112 0.01119 0.01117 0.01116
Cumulative Proportion  0.19623 0.20755 0.2189 0.23013 0.24139 0.25264 0.26386 0.2751 0.28625 0.29743 0.30859
                          PC23    PC24    PC25    PC26    PC27    PC28    PC29    PC30    PC31    PC32    PC33
Standard deviation     1.00713 1.00710 1.00689 1.00659 1.00654 1.00633 1.00627 1.00624 1.00622 1.00622 1.00622
Proportion of Variance 0.01115 0.01115 0.01114 0.01113 0.01113 0.01113 0.01113 0.01113 0.01113 0.01113 0.01113
Cumulative Proportion  0.31974 0.33088 0.34202 0.35316 0.36429 0.37542 0.38655 0.39767 0.40880 0.41992 0.43105
                          PC34    PC35    PC36    PC37    PC38    PC39    PC40    PC41    PC42    PC43    PC44
Standard deviation     1.00622 1.00622 1.00622 1.00622 1.00622 1.00622 1.00622 1.00614 1.00612 1.00606 1.00597
Proportion of Variance 0.01113 0.01113 0.01113 0.01113 0.01113 0.01113 0.01113 0.01112 0.01112 0.01112 0.01112
Cumulative Proportion  0.44218 0.45330 0.46443 0.47556 0.48668 0.49781 0.50893 0.52006 0.53118 0.54230 0.55343
                          PC45    PC46    PC47    PC48    PC49    PC50    PC51   PC52   PC53   PC54   PC55
Standard deviation     1.00590 1.00579 1.00564 1.00562 1.00550 1.00542 1.00529 1.0052 1.0051 1.0050 1.0049
Proportion of Variance 0.01112 0.01112 0.01111 0.01111 0.01111 0.01111 0.01111 0.0111 0.0111 0.0111 0.0111
Cumulative Proportion  0.56454 0.57566 0.58677 0.59789 0.60900 0.62011 0.63121 0.6423 0.6534 0.6645 0.6756
                         PC56   PC57    PC58    PC59    PC60    PC61    PC62    PC63    PC64    PC65    PC66
Standard deviation     1.0049 1.0048 1.00446 1.00435 1.00365 1.00354 1.00327 1.00323 1.00289 1.00266 1.00243
Proportion of Variance 0.0111 0.0111 0.01109 0.01108 0.01107 0.01107 0.01106 0.01106 0.01105 0.01105 0.01104
Cumulative Proportion  0.6867 0.6978 0.70889 0.71998 0.73105 0.74211 0.75318 0.76424 0.77529 0.78634 0.79738
                          PC67    PC68    PC69    PC70    PC71    PC72    PC73    PC74    PC75    PC76    PC77
Standard deviation     1.00217 1.00204 1.00181 1.00177 1.00169 1.00155 1.00144 1.00137 1.00131 1.00114 1.00112
Proportion of Variance 0.01104 0.01103 0.01103 0.01103 0.01103 0.01102 0.01102 0.01102 0.01102 0.01101 0.01101
Cumulative Proportion  0.80842 0.81945 0.83048 0.84151 0.85253 0.86356 0.87458 0.88560 0.89661 0.90763 0.91864
                          PC78    PC79   PC80    PC81    PC82    PC83    PC84    PC85    PC86    PC87    PC88
Standard deviation     1.00103 1.00073 1.0004 1.00026 0.96380 0.82929 0.60922 0.59188 0.57502 0.51871 0.42324
Proportion of Variance 0.01101 0.01101 0.0110 0.01099 0.01021 0.00756 0.00408 0.00385 0.00363 0.00296 0.00197
Cumulative Proportion  0.92965 0.94066 0.9516 0.96265 0.97286 0.98041 0.98449 0.98834 0.99198 0.99493 0.99690
                          PC89    PC90    PC91
Standard deviation     0.40999 0.32231 0.09983
Proportion of Variance 0.00185 0.00114 0.00011
Cumulative Proportion  0.99875 0.99989 1.00000
>
```

Nota. Se muestran los resultados del análisis de los componentes principales. Elaboración Propia.

En este caso por la cantidad de las variables deberíamos tomar en cuenta hasta la componente 67 que acumula el 80% de la información, según se muestra en la ***figura 23***. Sin embargo, el análisis de componentes principales debería darnos máximo 10

componentes. Si es mayor quiere decir que la información está muy diversa, es decir, tiene mucha varianza. En lo que si ayudo el análisis de componentes principales es en reducir la correlación de las variables.

Modelo de regresión múltiple

Consideraremos probar con un modelo de regresión múltiple considerando los siguientes parámetros:

- Validación cruzada: 5
- Familia Gaussian, debido a que nuestra variable objetivo es cuantitativa
- Sin balanceo
- Hacer una búsqueda del hiperparámetro lamba
- Selección automática del solver adecuado

Figura 25

Resultados del Modelo de regresión múltiple.

```
H2ORegressionMetrics: glm
** Reported on cross-validation data. **
** 5-fold cross-validation on training data (Metrics computed for combined holdout predictions) **

MSE:  9.38861e+11
RMSE:  968948.4
MAE:  435969.7
RMSLE:  3.044072
Mean Residual Deviance :  9.38861e+11
R^2 :  -6.245759e-05
Null Deviance :2.138806e+15
Null D.o.F. :2276
Residual Deviance :2.137787e+15
Residual D.o.F. :2275
AIC :69240.03


Cross-Validation Metrics Summary:
                              mean            sd    cv_1_valid    cv_2_valid    cv_3_valid    cv_4_valid
mae                       435596.62     11528.225     431473.97      462156.6     414747.56     444039.06
mean_residual_deviance  9.3963328E11 2.77072282E11  6.4964035E11 1.67865378E12  9.6019441E11  8.2051878E11
mse                     9.3963328E11 2.77072282E11  6.4964035E11 1.67865378E12  9.6019441E11  8.2051878E11
null_deviance          4.27761294E14 1.25996028E14 2.98834597E14  7.6380651E14 4.40732029E14 3.66775879E14
r2                     -0.0011859658   7.309141E-4 -0.0011466872 -0.0030014175  -3.888136E-4  -1.445299E-5
residual_deviance      4.27755926E14  1.2599123E14 2.98834563E14  7.6378746E14 4.40729243E14  3.6677192E14
rmse                        950983.7     132784.25      806002.7     1295628.8      979895.1     905824.94
rmsle                      3.0409727    0.06944582     3.1345193     2.8528154     3.0679965      3.097775
                          cv_5_valid
mae                        425566.03
mean_residual_deviance  5.8915894E11
mse                     5.8915894E11
null_deviance          2.68657452E14
r2                     -0.0013784579
residual_deviance      2.68656479E14
rmse                        767566.9
rmsle                      3.0517573
> |
```

Nota. Se muestran el resultado del rendimiento del modelo. Elaboración Propia.

En la ***figura 25***, se muestran los indicadores de rendimiento del modelo que serán comparados con el modelo de Random Forest.

Modelo H2O Random Forest

Debido a la naturaleza de nuestros datos que presentan correlación, varianza y valores outliers, optamos por trabajar nuestro modelo ensamblado con H2O. En H20, ya no necesitamos aplicar el análisis de componentes principales. Además, los modelos ensamblados son robustos para los datos con outliers y varianza. En este caso probaremos un modelo Random Forest. Aunque estos modelos están más orientados a pronosticar variables cualitativas. También pueden ser usados para target cuantitativos. Aplicamos el primer modelo con las variables por defecto.

Figura 26

Resultados del Modelo Random Forest con data de entrenamiento.

```
Model Details:
==============

H2ORegressionModel: drf
Model ID:  primer_modelo
Model Summary:
  number_of_trees number_of_internal_trees model_size_in_bytes min_depth max_depth mean_depth min_leaves
1              50                      50              495010        20        20   20.00000        539
  max_leaves mean_leaves
1       1067   780.02000


H2ORegressionMetrics: drf
** Reported on training data. **
** Metrics reported on Out-Of-Bag training samples **

MSE:  253770321609
RMSE:  503756.2
MAE:  169867.3
RMSLE:  1.42211
Mean Residual Deviance :  253770321609
```

Nota. Se muestran el resultado del rendimiento del modelo. Elaboración Propia.

El indicador MAE nos indica la suma del sesgo, es decir, la diferencia entre el valor estimado por el modelo y el valor real. Considerando que tenemos una venta promedio mensual de 31M, tener un sesgo de 169,867 es considerar un margen de error del 0.5%, lo cual es aceptable. Probamos nuestro modelo con la data de testeo.

Figura 27

Resultados del Modelo Random Forest con data de testeo.

```
> h2o.performance(Mod_h2o_1, newdata = D_h2o_test)
H2ORegressionMetrics: drf

MSE:  166899584195
RMSE:  408533.5
MAE:  151487.9
RMSLE:  1.430116
Mean Residual Deviance :  166899584195
```

Nota. Se muestran los indicadores de rendimiento del modelo. Elaboración propia.

Los indicadores MSE y RMSE están relacionado al error de la estimación, siendo el RMSE el error cuadradito. Como podemos observar nuestros indicadores mejoran.

Modelo H2O Random Forest con Validación Cruzada

Ahora probaremos el modelo considerando la validación cruzada y configuraremos el número de árboles en 500 para mejorar el modelo.

Figura 28

Resultados del Modelo Random Forest con validación cruzada.

```
H2ORegressionMetrics: drf
** Reported on cross-validation data. **
** 5-fold cross-validation on training data (Metrics computed for combined holdout predictions) **

MSE:  2.39697e+11
RMSE:  489588.6
MAE:  163677
RMSLE:  1.450891
Mean Residual Deviance :  2.39697e+11


Cross-Validation Metrics Summary:
                              mean           sd    cv_1_valid    cv_2_valid   cv_3_valid    cv_4_valid
mae                       163789.16    13452.172     168453.12     165627.67    128728.08     169483.53
mean_residual_deviance 2.40383148E11 6.7540095E10 2.49693733E11 1.95993944E11 8.6315401E10 3.66057095E11
mse                    2.40383148E11 6.7540095E10 2.49693733E11 1.95993944E11 8.6315401E10 3.66057095E11
r2                          0.742524  0.022091733    0.68399125     0.7754823    0.7508895     0.7424666
residual_deviance      2.40383148E11 6.7540095E10 2.49693733E11 1.95993944E11 8.6315401E10 3.66057095E11
rmse                        478491.6     75594.12     499693.62     442712.03     293794.8      605026.5
rmsle                      1.4491307   0.06371163     1.3859829     1.4899349    1.4095228     1.3547045
                         cv_5_valid
mae                       186653.36
mean_residual_deviance 3.03855567E11
mse                    3.03855567E11
r2                         0.7597904
residual_deviance      3.03855567E11
rmse                       551230.94
rmsle                      1.6055084
```

Nota. Se muestran el resultado del rendimiento del modelo. Elaboración Propia.

Luego de crear nuestro modelo considerando 5 validaciones cruzadas, probamos el nuevo modelo con la data de testeo.

Figura 29

Resultados del Modelo Random Forest con validación cruzada con data de testeo.

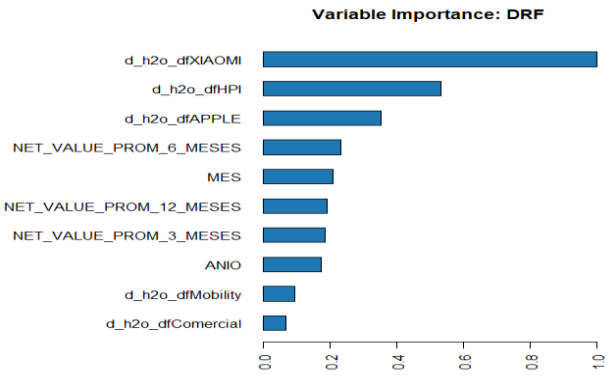
```
> h2o.performance(algo_h2o_valid)
H2ORegressionMetrics: drf
** Reported on training data. **
** Metrics reported on Out-Of-Bag training samples **

MSE:  215529346514
RMSE:  464251.4
MAE:  159054.9
RMSLE:  1.431771
Mean Residual Deviance :  215529346514

> h2o.performance(algo_h2o_valid, D_h2o_test)
H2ORegressionMetrics: drf

MSE:  161815429842
RMSE:  402262.9
MAE:  149085.5
RMSLE:  1.445395
Mean Residual Deviance :  161815429842

>
```

Nota. Se muestran el resultado del rendimiento del modelo. Elaboración Propia.

Como se aprecia en ***Figura 28***, los indicadores de rendimiento del modelo mejoran con la data de testeo. Analizamos las variables de mayor relevancia.

Figura 30

Importancia de variables.



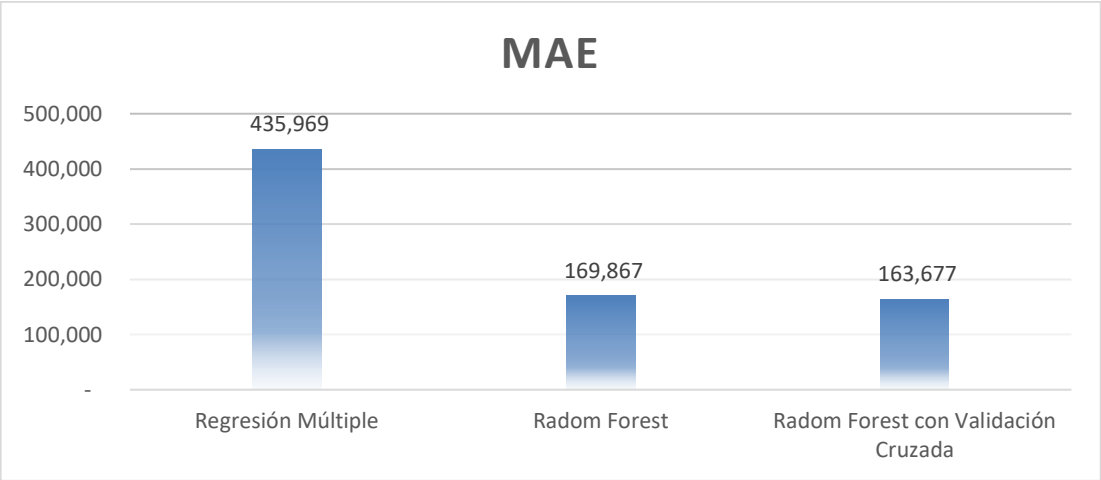
Nota. Variables según su importancia en el modelo. Elaboración Propia.

Como se muestra en la ***figura 29***, las marcas con mayor volumen de venta como XIAOMI, HPI y APPLE son las que tienen mayor influencia. Las variables NET_VALUE_PROM_3_MESES, NET_VALUE_PROM_6_MESES y NET_VALUE_PROM_12_MESES también son de influencia en el modelo. Debido a que estas variables se generaron de la variable objetivo, se espera que sean de influencia en modelo. De igual modo, al ser un modelo que tiene como variables de tiempo AÑO y MES se espera que sean variables determinantes en el modelo.

Finalmente, hacemos la evaluación de los modelos comparando el sesgo obtenido por cada modelo. Según se aprecia en la ***figura 28***, Random Forest con validación cruzada es el modelo con mejor precisión. Considerando este último como el modelo elegido.

Figura 31

Evaluación de modelos.



Nota. Se muestran los resultados de MAE obtenidos por cada modelo. Elaboración Propia.

3.2.5.6.3.RETROESPECTIVA

Al final de este último sprint, se realizó la presentación de los modelos generados y los resultados obtenidos por cada uno. Asimismo, se mencionaron las tareas realizadas por el equipo de desarrollo. Se validó junto a todos los involucrados el cumplimiento del objetivo del sprint.

Tabla 29

Resumen de la Reunión de Retroespectiva – Sprint 5

Sprint	Historia de Usuario	Comentarios
Sprint 5	Como gerente de producto necesito contar un estimados de cuotas por marcas y división de negocio.	Todo conforme con esta historia de usuario.

Nota. Fuente: Propia

3.3. EVALUACIÓN

3.3.1. EVALUACIÓN ECONÓMICA / EVALUACIÓN COSTO-BENEFICIO

El presupuesto invertido en el proyecto se calculó considerando el costo de los recursos y tiempos designados al análisis e implementación de la solución. A continuación, se muestra el detalle de cada uno:

Tabla 30

Evaluación económica del proyecto

Concepto	Recurso	Tiempo (horas)	Monto Total(S/.)
Implementación	Product Owner	48	4,091
	Scrum Mater	68	5,795
	Senior Business Intelligence Data Analyst	272	10,045
Capacitaciones	Senior Business Intelligence Data Analyst	8	200
Total		396	22,882

Nota. Fuente: Propia

En efecto, la implementación del proyecto redujo significativamente el número de horas hombres empleadas en el procesamiento de los datos, corrección por errores u omisiones en los resultados tanto por parte del área de BI como el área de Recursos Humanos que tenía que reprocesar el saldo pendiente de los asociados afectados. No obstante, existen otros beneficios importantes que aporta la solución y que contribuyeron a agilizar los procesos y mejorar la productividad y rentabilidad de las áreas involucradas. De acuerdo con lo anterior, se puede afirmar que el proyecto tiene un costo beneficio positivo para la empresa. A continuación, se detalla los beneficios que ofrece el proyecto:

Tabla 31

Cuadro comparativo de beneficios entre soluciones

Concepto a evaluar	Anterior Solución	Nueva Solución
Información centralizada	NO	SI
Información actualizada	Mensual	Diaria
Envío de Resultados Finales	2 semanas despúes del cierre	1 semana después del cierre
Horas de procesamiento y cálculo de resultados	50 horas al mes	8 horas al mes
Ánalisis de la información	NO	SI
Acceso a la información	A solicitud	24 x 7
Correcciones mensuales	15	4 ó 0

Nota. Fuente: Propia

Como una de las principales utilidades del proyecto, se encuentra la reducción de horas de trabajo y por ende la mejora en los tiempos de entrega de la información. Esto da pase a enfocar esas horas que antes se dedicaban a tareas operativas, en el análisis de la información y tomar mejores decisiones a favor del proceso.

La información consolidada en un repositorio único permite analizar los resultados históricos, identificar comportamientos y tener una visión global del proceso.

Se mejoró la experiencia del asociado. En el proceso anterior, el no tener información disponible sobre sus cuotas y forma de evaluación generaba desconcierto sobre cómo mejorar su cumplimiento y afinar sus estrategias. Con la implementación, la comunicación sobre sus resultados mejoró, teniendo un acceso a los datos 24 x 7 a través de la página web.

Desde el portal, los asociados tienen visibilidad de los criterios que se consideran en sus resultados. De esta manera, es más fácil identificar si hay algún criterio que no está actualizado y hacer la solicitud de corrección a tiempo antes de terminar la evaluación al cierre del trimestre y generar un reproceso en el ajuste.

Se impulsaron las ventas con una acertada estimación de las cuotas. De esta manera, se puedo identificar a las unidades de negocio con oportunidad de crecimiento y asignarle una cuota retadora en base a histórico.

3.3.2. INTERPRETACIÓN DEL VAN Y DEL TIR

Definiremos los ingresos y los costos de este proyecto, para luego evaluar los indicadores financieros de la inversión. El proyecto no contempló la adquisición de licencias o infraestructura adicional, se desarrolló considerando las herramientas disponibles y software libre. En el concepto de inversión inicial se consideró las horas monetizadas empleadas en la etapa de inicio y concepción del proyecto. A continuación, se detallan los costos mensuales considerados en el proyecto.

Tabla 32

Costo del proyecto mensualizado

Conceptos	Costo Mensual (S/.)
Inversión Inicial	1,659
Mes 1	7,250
Mes 2	7,250
Mes 3	3,973
Total	20,132

Nota. Fuente: Propia

En los ingresos del proyecto, se consideran la reducción de horas hombres que se consumían en procesos operativos para el cálculo de resultados y consolidación de los datos y consultas de evaluaciones individuales. Además, de la reducción de horas en reprocesos por correcciones manuales.

Tabla 33

Beneficio del proyecto mensualizado

Conceptos	Beneficio mensual (S/.)
Tareas operativas y reprocesos en el área de BI	2,364
Reprocesos en el área de Recursos Humanos	148
Total	2,511

Nota. Fuente: Propia

Para este proyecto, definiremos una tasa de descuento del 10% y periodo de evaluación de 24 meses. Luego de definir los dos puntos anteriores, obtendremos el VAN y la TIR de esta inversión.

Tabla 34

Indicadores financieros del proyecto

VAN	979
TIR	10.74%

Nota. Fuente: Propia

La solución presentada en el informe de investigación es un proyecto que por necesidad del negocio se debía implementar, debido a que era un proceso demasiado manual e impactaba de manera negativa en otros procesos de la compañía y en la experiencia del asociado. Sin embargo, como podemos observar en la **tabla 34**, este proyecto resulta también ser rentable para la compañía teniendo como resultado una tasa de beneficio (TIR) de 10.74%, superior a la tasa de descuento definida previamente.

El VAN nos ayuda a determinar la viabilidad y rentabilidad de un proyecto de inversión. En este caso el importe obtenido es positivo, lo que nos indica que el proyecto es rentable. Además, el beneficio acumulado obtenido del proyecto es de S/.35,118.

CAPÍTULO IV. REFLEXIÓN CRÍTICA DE LA EXPERIENCIA

La implementación del sistema de inteligencia de negocios generó una serie de experiencias nuevas y positivas tanto al autor de este informe como a la empresa en estudio. Entre algunas de ellas se mencionan las siguientes:

- Debido a que es la primera vez que se usa el desarrollo de modelos predictivos para la implementación de un proyecto en el área, me sentí motivada a capacitarme e incrementar mis conocimientos técnicos y habilidades analíticas.
- Se logró cumplir con el tiempo estimado del proyecto, a pesar de que hubo algunas historias de usuario que tomaron más tiempo del estimado. Lo anterior fue posible debido a que otras tareas de usuarios terminaron antes del tiempo estimado y se aprovecharon los tiempos para compensar el cronograma del proyecto. Como equipo hemos evaluado estos casos para mejorar nuestras estimaciones.
- Debido a que era la primera vez que se desarrollaba una página web como herramienta complementaria a un proyecto de inteligencia de negocios en la empresa en estudio y se encontró potencial para trabajar nuevos módulos, se ha considerado implementar más módulos para reportes de otros proyectos y dar visibilidad de los datos a usuarios que por tema de costo de licencias no pueden acceder a la plataforma de Power BI.
- La creación de un repositorio centralizado nos permitió consolidar la información de distintas fuentes y almacenar data histórica para tener una visión general de nuestro proceso e identificar oportunidades de mejora. El repositorio central fue la base para éxito de las aplicaciones que se implementaron en este proyecto, proveyendo de datos a las aplicaciones para su explotación y visualización. De esta manera, resaltamos su importancia en la preservación y acceso de los datos.
- El proyecto se desarrolló bajo la metodología SCRUM, la cual se adaptó bastante bien en este entorno de Inteligencia de Negocios aportando flexibilidad a la implementación. Debido al alto impacto de la implementación que transformó el proceso anterior, se identificaron interesados resistentes al cambio y con poca

disposición. Las entregas parciales en cada Sprint ayudaron a dar visibilidad del avance del proyecto y gestionar las expectativas de los interesados, manteniéndolos siempre informados de los avances e involucrándolos en cada reunión. De esta manera, se fortaleció el compromiso y comunicación del Scrum Team.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

El desarrollo del proyecto se realizó satisfactoriamente, logrando mejorar los tiempos de entrega de resultados y mejorando la evaluación de los asociados.

A continuación, se detalla los principales resultados obtenidos tras el desarrollo de este proyecto:

- Se automatizó el procesamiento de los datos y cálculo de los resultados a través de un proceso ETL, reduciendo las horas de trabajo operativo del área de BI de 50 a 8 horas al mes (como se muestra en la ***tabla 23***) y permitiendo mostrar el resultado del 80% de las métricas en el primer día hábil después del cierre de ventas. Por consiguiente, se redujo el tiempo de entrega de resultados finales al área de Recursos Humanos en 5 días hábiles luego del cierre de mes, a diferencia del proceso manual anterior que tomaba de 8 a 10 días hábiles obtener los resultados finales.

- Se optimizó el acceso de los datos a través de un repositorio centralizado. De esta manera se garantizó el almacenamiento de los datos históricos, con el fin de tener una visión general y hacer un correcto análisis de los datos.

- La implementación del dashboard de seguimiento del performance de los asociados permitió identificar mejoras en las evaluaciones para replantear los pesos y métricas de los asociados, dar seguimiento a los sobrecumplimientos y definir lineamientos para una correcta gestión de la compensación variable.

- Se incrementaron los resultados de ventas en 152% de la compañía versus el año anterior, la implementación de modelo predictivo para el forecast de ventas contribuyó a definir cuotas retadoras y reales que impulsara el crecimiento del negocio.

- La implementación de la página web mejoró la comunicación y el sentimiento de satisfacción de los asociados. Muchos asociados no tenían conocimiento de las métricas y pesos que se consideraban en su evaluación. La aplicación web no solo dio visibilidad de las cuotas y evaluaciones, sino además podrían ver y monitorear su avance de resultados todos los días y con esto tomar acción frente a retrasos del cumplimiento de metas. Asimismo, la página web permitía ver los criterios considerados en las evaluaciones.

- Con la mejora al acceso de información con la página web y la automatización del cálculo de resultados, se redujo un 80% el número de ajustes enviados al área de recursos humanos comparando la cantidad de ajustes enviados el año en que se implementó la solución versus el año anterior, en algunos meses no se reportaron ajustes. Estos ajustes generaban reprocesos en ambas áreas tanto en el área de Inteligencia de Negocios como Recursos Humanos.

5.2 RECOMENDACIONES

Del análisis e implementación de la solución de Inteligencia de Negocio, se plantean la revisión de las siguientes propuestas:

- En una próxima versión del aplicativo web, deberían agregarse un módulo para descargar el detalle de las registros y transacciones que se consideran en el resultado de cada métrica.

- La aplicación web es una intranet que solo funciona cuando los asociados se encuentran conectados a la VPN de la empresa. Con el fin de asegurar el principio de disponibilidad de la información se podría contemplar en un futuro una versión móvil que no requiera de la conexión a la VPN.

- La estimación de cuotas de ventas a través de un modelo predictivo se encuentra a nivel de división de negocio y marca, en una próxima etapa se puede considerar llegar a un nivel de categoría de producto. Con el fin de evaluar mejor la estacionalidad de cada categoría y por consiguiente mejorar la precisión del modelo.

- La gestión de la remuneración de variable del equipo de cobranzas se realiza de manera semiautomática. Debido a la falta de herramientas solo se considera una métrica en la evaluación de los analistas de cobranzas y no se almacena data histórica. El proceso de cobranzas presenta como oportunidades de mejora, mucho de los puntos que se abordaron con éxito en la implementación de esta solución. Como un siguiente proyecto y en base a lo aprendido, se propone la implementación de un sistema de inteligencia de negocios para la estimación de cuotas y monitoreo del desempeño del equipo de cobranzas.

5.3 FUENTES DE INFORMACIÓN

Microsoft. (Febrero del 2019). Portal Microsoft Academy. Obtenido de https://www.microsoft.com/es-mx/microsoft-365/business-insights-ideas/resources/why-data-visualization-is-a-must-for-understanding-business-intelligence

Microsoft. (Febrero del 2016). Portal Microsoft Academy. Obtenido de https://news.microsoft.com/es-xl/diferencias-entre-inteligencia-de-negocios-y-analisis-avanzados-infografia/

Neoattack. (Agosto del 2020). NEOATTACK. Obtenido de https://neoattack.com/neowiki/modelo-predictivo/

Overeem, B. & Van Schaik, S. (Julio del 2016). Scrum & Business Intelligence. Obtenido de https://www.scrum.org/resources/scrum-and-business-intelligence

Schwaber, K., & Sutherland, J. (Noviembre del 2020). La Guía Scrum. Obtenido de https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-Spanish-European.pdf

Tymkiw, N. & Bournissen, J. & Tumino, M. (Abril del 2020). SCRUM como Herramienta Metodológica para el Aprendizaje de la Programación. Obtenido de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1850-99592020000200010#ref

Guerrero, S. (Junio del 2019). Machine Learning, el futuro de la Inteligencia Artificial. Obtenido de https://www.nicmexico.mx/2019/09/04/machine-learning-el-futuro-de-la-inteligencia-artificial/

Shmueli, G. & Bruce, P. & Yahav, I. (2018). Data Mining For Business Analytics

Amat, J. (Julio del 2018). Machine Learning con H2O y R. Obtenido de https://rpubs.com/Joaquin_AR/406480

Stedman, Craig (2020). Ultimate Guide to Business Intelligence in the Enterprise. Obtenido de https://www.techtarget.com/searchbusinessanalytics/Ultimate-guide-to-business-intelligence-in-the-enterprise

Burns, Ed (2021). In-depth guide to machine learning in the enterprise. Obtenido de: https://www.techtarget.com/searchenterpriseai/In-depth-guide-to-machine-learning-in-the-enterprise

Velasquez, Waldir (2022). Sistema de gestión para el cálculo y control de los servicios en campo de redes externas en una empresa de gas natural. Obtenido de: https://cybertesis.unmsm.edu.pe/handle/20.500.12672/18066

Morante, Grace (2021). Una solución de inteligencia de negocios para apoyar el análisis del desempeño de los proyectos en una empresa constructora. Obtenido de: https://cybertesis.unmsm.edu.pe/handle/20.500.12672/17309