



Análisis inteligente de datos en la Gerencia de Créditos y Cobranzas de una empresa comercial

2024

*Juárez Daiana Rocio
daianajuarz@gmail.com
Tutor: Medina Galván Marcelo*



ÍNDICE

Resumen	1
Palabras Claves	1
Introducción	2
Situación Problemática	3
Preguntas de Investigación	4
Objetivo General	5
Objetivos Específicos	5
Marco Metodológico	5
Marco Teórico	6
Datos - La pirámide del conocimiento	6
Primera Etapa - Situación actual de la Gerencia	15
Segunda Etapa - Análisis	24
Análisis Exploratorio de Datos (AED)	24
Árbol de Decisión	28
Cluster	30
Recomendaciones	39
Conclusiones	41
Anexo	42
Referencias	43



Resumen

En las organizaciones "data-centric", los datos son activos clave que generan valor para colaboradores y clientes, a diferencia de las "data-driven", donde su uso se limita a aplicaciones específicas sin integración total. Este desafío está presente en Comfy S.A., una empresa con 56 años de experiencia en la venta minorista y mayorista de artículos para el hogar en Tucumán, Salta y Catamarca. En su Gerencia de Créditos y Cobranzas, aunque se manejan grandes volúmenes de datos, no se transforman en información estratégica para la toma de decisiones. Además, la falta de integración con el área comercial dificulta la creación de estrategias para atraer y retener clientes hacia los créditos personales.

El objetivo es proponer un análisis inteligente de datos para optimizar la toma de decisiones gerenciales e idear estrategias de crédito basadas en datos. El marco teórico incluye conceptos como la pirámide del conocimiento, análisis de datos, organizaciones basadas en datos, toma de decisiones y machine learning, con especial énfasis en la preparación y limpieza de datos.

El estudio utiliza un enfoque mixto y diseño exploratorio secuencial (DEXPLOS). Primero, se analizan datos cualitativos mediante observación directa y entrevistas con colaboradores para comprender el funcionamiento de la gerencia. Luego, se realiza un análisis cuantitativo con herramientas como Excel, Google Colab y Python, identificando preferencias de clientes, rangos etarios predominantes y segmentos con mayor potencial. Esto permite proponer estrategias para mejorar la retención y captación de clientes.

Palabras Claves Análisis inteligente, Gestión de datos, Toma de decisiones, Créditos y Cobranzas



Introducción

En las organizaciones centradas en datos también llamadas "data-centric", estructuran su funcionamiento en torno a los datos, convirtiéndolos en el núcleo de su estrategia y operación, esto implica que los datos dejan de ser meros recursos y se transforman en activos claves que permiten la generación de valor no solo para los colaboradores internos, sino también para los clientes.

Esto contrasta con las organizaciones "data-driven", donde los datos son utilizados principalmente en aplicaciones específicas, pero sin una integración plena en todos los niveles organizacionales. Sin embargo, muchas organizaciones aún enfrentan desafíos significativos en la adopción de una cultura verdaderamente basada en datos, lo que limita su capacidad de ser competitivas en un entorno dinámico.

En este último, se ubica Comfy SA, una empresa familiar con 56 años de trayectoria en el mercado, dedicada a la comercialización minorista y mayorista de artículos para el hogar y electrodomésticos en la provincia de Tucumán. Con una sede central en San Miguel de Tucumán y quince sucursales distribuidas en Tucumán, Salta y Catamarca, la empresa ha consolidado su posición en el mercado regional.

La gerencia de Créditos y Cobranzas está centralizada en la sede central y está conformado por un equipo de 11 personas. Este equipo incluye al gerente de Créditos y Cobranzas, dos analistas de créditos, dos auditores internos, dos administrativos, tres operadores de call center y un responsable del archivo. Aunque cada miembro tiene claras sus funciones, a menudo asumen tareas adicionales de sus compañeros o realizan actividades operativas cuando es necesario.

El equipo se distingue por la antigüedad de sus integrantes en la empresa, muchos de los cuales han pasado por diferentes roles y ascendiendo en la empresa, como vendedores, administrativos, auditores, e incluso han trabajado en empresas competidoras del mismo rubro y



sector. Esta experiencia acumulada les permite tomar decisiones basadas tanto en su conocimiento como en su intuición.

Situación Problemática

La venta a crédito personal es fundamental para Comfy S.A., lo que hace clave la labor de la Gerencia de Créditos y Cobranzas. El proceso de crédito abarca desde que el cliente decide financiar una compra hasta el pago de la última cuota y la entrega del pagaré o libre deuda.

El otorgamiento de créditos es manejado por los administrativos, quienes una vez que pasaron el filtro por los vendedores donde estos consultan previamente en el motor si el cliente puede comprar a crédito, pasan luego hacia los administrativos quienes deben cargar manualmente los datos del cliente en el sistema. Esta tarea presenta problemas debido a errores de tipeo que a menudo resultan en el rechazo de créditos. Para nuevos clientes, se realiza una verificación exhaustiva de la información mediante el motor de decisión que recopila y compara información de las siguientes fuentes: BCRA, SIISA, Riesgonet y Veraz y, en algunos casos, Anses. Una vez aprobado, el cliente debe firmar un pagaré y un contrato de mutuo, y su información se registra en el sistema como "Cuenta Apertura" si se trata de un cliente nuevo.

Para clientes frecuentes, la decisión de otorgar crédito se basa en el análisis del margen disponible y el historial de pagos previos. Aunque el sistema proporciona un análisis preliminar, la intervención de analistas o del gerente es necesaria para casos que requieren un examen más detallado.

En la gestión de cobranzas, los operadores de call centers reciben planillas mensuales con datos de clientes en mora, registran observaciones y actividades, y solicitan intervención de



visitadores si es necesario. Las cuentas incobrables se derivan a juicio y, finalmente, se envían a un estudio jurídico para su gestión.

La gerencia de créditos y cobranzas de la empresa enfrenta desafíos significativos debido a la necesidad de un esfuerzo constante de sus colaboradores para transformar los datos que recopila el sistema en información útil para la toma de decisiones y creación de estrategias de crédito. A pesar de los avances logrados en la formalización de procesos de crédito y en el control de la morosidad, la dependencia de planillas de cálculo y herramientas manuales como Excel ralentiza la obtención de información relevante, lo que impide aprovechar el potencial de los datos disponibles. Esta situación, sumada a la alta participación del gerente en tareas operativas, limita su capacidad para centrarse en cuestiones estratégicas, además la falta de integración entre el área comercial y la gerencia bajo estudio, la ausencia de un plan de acción claro, objetivos definidos y una presentación de resultados adecuada, ha llevado a la ejecución de estrategias sin resultados.

Preguntas de Investigación

A partir del problema identificado se plantearon las siguientes preguntas de investigación:

1. ¿De qué manera trabaja actualmente la gerencia con los datos?
2. ¿Cuáles son las necesidades actuales de la gerencia en relación con los planes estratégicos de créditos? ¿Qué planes sería conveniente proponer?
3. ¿Qué herramientas de análisis de datos podrían recomendarse para optimizar la gestión de créditos y cobranzas?



Objetivo General

El objetivo general de este trabajo es proponer herramientas de análisis de datos inteligentes que proporcionen información útil para la toma de decisiones y la creación de planes de acción para la Gerencia de Créditos y Cobranzas.

Objetivos Específicos

- Evaluar las fortalezas y debilidades en los procesos de análisis de datos de la gerencia de créditos y cobranzas.
- Descubrir las necesidades comerciales y de crédito de la empresa a partir del análisis de los datos.
- Explorar herramientas de análisis de datos que podrían ser recomendadas para optimizar la gestión de créditos y cobranzas.

Marco Metodológico

Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada (meta inferencias) y lograr un mayor entendimiento del fenómeno bajo estudio. (Sampieri, 2018)

Para el desarrollo del trabajo se utilizó un enfoque mixto con diseño exploratorio secuencial (DEXPLOS), el cual implica una fase inicial de recolección y análisis de datos cualitativos, donde se exploraron las características generales de la Gerencia Créditos y Cobranzas. Posteriormente se recabaron y analizaron datos cuantitativos con el fin de proponer



indicadores que contribuyan a mejorar la toma de decisiones.

Los instrumentos de recolección de datos utilizados son:

- Observación participante: realizada por el investigador durante su estadía en el lugar de trabajo.
- Entrevistas con expertos: entrevista en profundidad con la Gerente del Departamento y colaboradores que forman parte del mismo.
- Base de datos: brindada por los sistemas informáticos “Puma”

Para el análisis de datos se utilizan herramientas de análisis de datos, entre ellas Excel, Google Colab y Python.

Marco Teórico

Datos - La pirámide del conocimiento

Cuando se habla de conocimientos, se debe hacer una distinción entre datos (hechos discretos, registros), información (datos organizados, con significado) y conocimiento. Davenport y Prusak definen el conocimiento como “una mezcla fluida de experiencia estructurada, valores, información contextual e internalización experta que proporciona un marco para la evaluación e incorporación de nuevas experiencias e información”. El conocimiento “se origina y es aplicado en la mente de los conocedores”, y en la organización “queda arraigado en rutinas, procesos, prácticas y normas institucionales”.

Lo importante del conocimiento, señalan estos autores, es que está cerca de la acción y que por ello permite la toma de decisiones. Se compone de experiencia, “verdad práctica”, complejidad,

criterios, reglas implícitas (intuición), así como de valores y creencias. Cada uno de estos elementos tiene sus propias características.

Datos: Los datos son la mínima unidad semántica, y se corresponden con elementos primarios de información que por sí solos son irrelevantes como apoyo a la toma de decisiones.

Información: La información se puede definir como un conjunto de datos procesados y que tienen un significado (relevancia, propósito y contexto), y que por lo tanto son de utilidad para quién debe tomar decisiones, al disminuir su incertidumbre.

Conocimiento: El conocimiento es una mezcla de experiencia, valores, información y know-how que sirve como marco para la incorporación de nuevas experiencias e información, y es útil para la acción.

Sabiduría: En este último nivel Cleveland describió la sabiduría simplemente como "conocimiento integrado: información que se vuelve súper útil". Otros autores han caracterizado la sabiduría como "saber qué hacer correctamente". La sabiduría implica utilizar el conocimiento para el bien mayor. Por eso, la sabiduría es más profunda y más exclusivamente humana. Requiere un sentido del bien y del mal, lo correcto y lo incorrecto, lo ético y lo no ético.

Título: Pirámide del Conocimiento



Fuente: elaboración propia



Análisis de Datos

El análisis de datos es el proceso de examinar, limpiar, transformar y modelar un conjunto de datos con el objetivo de descubrir información útil, extraer conocimientos y tomar decisiones informadas. Implica la aplicación de técnicas y herramientas estadísticas, matemáticas y de visualización para identificar patrones, tendencias y relaciones en conjuntos de datos. El análisis de datos permite revelar insights, responder preguntas y resolver problemas, ayudando a las organizaciones y personas a comprender mejor el mundo que les rodea, optimizar procesos y tomar acciones basadas en evidencia. (Innovación Digital 360, s.f.).

Por otro lado, el análisis inteligente de datos se refiere a la aplicación de técnicas avanzadas de análisis, como el machine learning, la inteligencia artificial (IA) y el big data, para identificar patrones y extraer conocimientos significativos a partir de grandes volúmenes de datos. Este enfoque va más allá de los análisis descriptivos tradicionales, permitiendo a las organizaciones predecir comportamientos futuros y tomar decisiones proactivas.

Según Provost y Fawcett (2013), el análisis inteligente de datos es esencial para la toma de decisiones basada en datos (data-driven decision making), ya que permite descubrir relaciones ocultas en los datos y optimizar las operaciones empresariales. Además, el uso de herramientas como dashboards interactivos facilita la visualización y comprensión de los datos, lo que mejora la capacidad de la gerencia para tomar decisiones informadas y oportunas (Few, 2012).

Organizaciones basadas en datos

En un mundo altamente interconectado, los datos se han convertido en un recurso esencial para la transformación social. La abundancia de datos generados impone desafíos que requieren un tratamiento cuidadoso. Por lo tanto, la cultura organizacional basada en datos es crucial, ya que



facilita la toma de decisiones informadas, fomenta conversaciones enriquecedoras y permite cuestionar el conocimiento establecido, creando un entorno seguro y propicio para el cambio organizacional.

“Una cultura basada en datos es aquella que basa sus decisiones en hechos derivados de datos, investigación, experiencia y aportes de fuentes creíbles, en lugar de en suposiciones y “corazonadas”. En una cultura basada en datos, éstos están disponibles, son accesibles, fiables y se confían a todos los miembros de la organización para que informen sus decisiones y acciones” (Educause, 2022, p.8).

María del Pilar Villamil Giraldo (2024), Profesora asociada Departamento Ingeniería de Sistemas y coordinadora de la maestría de Ingeniería de Información Universidad de Los Andes en una revista académica menciona: “En una organización basada en datos, éstos dejan de ser recursos y se transforman en activos estratégicos de la organización que representan la manera como la organización genera valor no solo para sus propios colaboradores, sino para sus clientes. Es un ejercicio recursivo, donde los aportes y creación de experiencias diferenciales internas traducidas en agilidad, eficiencia y estandarización, se traducen en nuevas oportunidades y capacidades que la organización puede generar y transferir a sus clientes. De esta forma los datos se convierten en un pilar fundamental para movilizar y asegurar la promesa de valor de empresa y su evolución de la mano de las expectativas de sus consumidores”.

Lo primero es aclarar el término de organizaciones basadas en datos, que se puede llevar a organizaciones centradas en datos (data centric) o guiadas por datos (data-driven). En las primeras, la toma de decisiones se basa en los datos y toda la estructura organizacional y funcional gira en torno a ellos. En contraposición con las empresas data-driven que están centradas en aplicaciones particulares, y no se entiende la utilidad de los datos a nivel de apertura al cambio y oportunidad



que se tiene de ser competitivos apalancando sus acciones y decisiones en los datos. Es por ello, que se debe empezar por resolver la pregunta ¿para qué sirven? y principalmente cómo aportan en la construcción del “negocio”, tanto en la parte operativa como en la táctica y la estratégica.

Según la consultora Gartner, el concepto de organizaciones basadas en datos se refiere a aquellas que centran sus decisiones estratégicas y operativas en el análisis exhaustivo de datos. Estas organizaciones utilizan técnicas avanzadas de ciencia de datos para transformar datos brutos en información útil, que luego es utilizada para optimizar procesos, mejorar productos y servicios, y prever tendencias futuras. (Gartner, 2023).

Toma de decisiones gerenciales

Se puede decir que la toma de decisiones gerenciales implica identificar un problema y seguir un proceso que incluye la obtención de información, la toma de decisiones, la acción y la evaluación del desempeño de la empresa. No obstante, en la práctica, este proceso enfrenta obstáculos como la incertidumbre, que impacta en la toma de decisiones, y la falta de información sobre las alternativas y sus consecuencias, lo que dificulta la interpretación y la decisión. (Choo, 1991, como se citó en Jarrahi, 2018, citado en Diestra Quinto et al., 2023).

Machine Learning

Machine Learning que en español significa aprendizaje automático, es una parte de la inteligencia artificial, que busca construir programas por computadora que aprendan automáticamente en base a experiencia adquirida. Los programas creados con Machine Learning, no necesitan ser programados con reglas explícitas que definan las tareas que debe seguir para el cumplimiento de estas, por el contrario, esta lógica mejora automáticamente en base al uso de



algoritmos que analicen los datos, buscando reconocer patrones y tendencias para el entendimiento de los datos. (Bagnato, 2020, como se citó en Muñoz Cavero & Luyo Pérez, 2022, p. 28).

Según explica Sandoval (2018), los algoritmos de Machine Learning se dividen en dos categorías principales: aprendizaje supervisado y no supervisado.

Aprendizaje supervisado: Es cuando entrenamos un algoritmo de Machine Learning dándole las preguntas (características) y las respuestas (etiquetas). Así en un futuro el algoritmo pueda hacer una predicción conociendo las características. Entre ellos encontramos los siguientes modelos:

- Modelos lineales: Estos tratan de encontrar una línea que se “ajuste” bien a la nube de puntos que se disponen. Aquí destacan desde modelos muy conocidos y usados como la regresión lineal (también conocida como la regresión de mínimos cuadrados), la logística (adaptación de la lineal a problemas de clasificación -cuando son variables discretas o categóricas-).
- Redes neuronales: Las redes artificiales de neuronas tratan, en cierto modo, de replicar el comportamiento del cerebro, donde tenemos millones de neuronas que se interconectan en red para enviarse mensajes unas a otras. Esta réplica del funcionamiento del cerebro humano es uno de los “modelos de moda” por las habilidades cognitivas de razonamiento que adquieren.
- Modelos de árbol: Son modelos precisos, estables y más sencillos de interpretar básicamente porque construyen unas reglas de decisión que se pueden representar como un árbol. A diferencia de los modelos lineales, pueden representar relaciones no lineales para resolver problemas.



- Random Forest: Bosques Aleatorios es un tipo de clasificador que agrupa un conjunto de árboles estructurados distribuidos de manera idéntica que arrojan un voto unitario para la clase más popular su uso es tanto para tareas de clasificación como regresión usando voto mayoritario y ponderación respectivamente. La combinación de dichos arboles claro está que bajo ciertas condiciones proporciona un mejor resultado dando como resultado un método más preciso, estable, dinámico que busca el equilibrio entre el sesgo y la varianza del bosque. (Patiño Pérez et al., 2020, p. 105)

Los árboles de decisión, sus partes claves son las siguientes:

- Nodos Raíz: El primer nodo del árbol, donde comienza la división del conjunto de datos.
- Nodos Internos: Nodos que representan decisiones basadas en atributos. Cada nodo interno tiene ramas que conducen a otros nodos internos o a nodos hoja.
- Ramas: Las conexiones entre nodos que representan los resultados de las decisiones.
- Nodos Hoja: Los nodos finales que representan la salida del modelo: etiquetas de clase en la clasificación o valores numéricos en la regresión.

Aprendizaje no supervisado: Aquí solo le damos las características al algoritmo, nunca las etiquetas. Queremos que nos agrupe los datos que le dimos según sus características. El algoritmo solo sabe que como los datos comparten ciertas características, de esa forma asume que pueda que pertenezcan al mismo grupo. Entre ellos encontramos:

- Cluster: se encarga de formar grupos diferentes dentro de los datos. Al usar algoritmos de agrupamiento se encuentra la estructura en los datos, de manera que los elementos del mismo clúster (o grupo) sean más similares entre sí que con los otros grupos generados. (Román, citado en Ramírez Mendoza, 2022).



El desarrollo de estos modelos sigue dos fases: la *fase de entrenamiento*, en la que el algoritmo aprende patrones a partir de un conjunto de datos, y la *fase de prueba*, donde se evalúa la precisión de sus predicciones. Un alto grado de precisión indica que el algoritmo ha adquirido un buen nivel de aprendizaje.

Preparación de los datos

Para la aplicación del modelo, es necesario seguir una serie de pasos de preparación que aseguren la calidad y relevancia de los datos utilizados. Según lo indicado en el documento Análisis Inteligente de Datos 2024 de la Universidad Austral, este proceso incluye varias fases, desde la selección y limpieza de los datos hasta su transformación, con el objetivo de obtener un conjunto de datos estructurado y listo para el modelado. A continuación, se detallan estas fases esenciales en la preparación de datos:

Fase de Selección de datos: Implica elegir las fuentes y los datos específicos que serán utilizados en el análisis. Esta fase es crucial para asegurar que se esté trabajando con la información relevante y de calidad.

- Determinar las bases de datos, archivos, APIs, u otras fuentes donde se encuentra la información.
- Recolección de datos: Extraer los datos necesarios desde las fuentes identificadas.
- Relevancia de los datos: Evaluar si los datos recolectados son pertinentes y útiles para el análisis objetivo.
- Volumen de datos: Asegurar que la cantidad de datos es suficiente para obtener resultados significativos.



Fase de Limpieza de datos: La limpieza de datos es el proceso de corregir o eliminar datos inexactos, incompletos o irrelevantes. Esta fase es esencial para asegurar la calidad y la integridad de los datos.

- **Eliminación de duplicados:** Identificar y remover registros duplicados que puedan sesgar el análisis.
- **Manejo de valores faltantes:** Decidir cómo tratar los datos ausentes, ya sea imputándolos, eliminándolos o dejándolos como están según la situación.
- **Corrección de errores:** Corregir datos erróneos o inexactos (errores tipográficos, valores fuera de rango, etc.).
- **Normalización:** Asegurar que los datos estén en un formato consistente y estandarizado (por ejemplo, fechas en el mismo formato, unidades de medida coherentes).

Fase de Exploración: Implica analizar los datos para entender sus características principales y encontrar patrones preliminares. Esta fase es clave para formular hipótesis y guiar el análisis posterior.

- **Análisis descriptivo:** Utilizar estadísticas descriptivas (media, mediana, moda, desviación estándar, etc.) para resumir las características principales de los datos.
- **Visualización de datos:** Crear gráficos y visualizaciones (histogramas, diagramas de dispersión, box plots, etc.) para identificar patrones y relaciones en los datos.
- **Detección de anomalías:** Identificar valores atípicos o anómalos que puedan necesitar una atención especial o una revisión adicional.
- **Análisis de correlación:** Evaluar las relaciones entre diferentes variables para entender cómo se influyen mutuamente.



Fase de Transformación de datos: Implica modificar los datos para que se ajusten mejor a las necesidades del análisis o del modelo que se va a utilizar. Esta fase prepara los datos para el análisis final.

- Creación de nuevas variables: Generar nuevas variables derivadas de las existentes para capturar mejor la información relevante (por ejemplo, agregar variables calculadas, como ratios o índices).
- Escalado de datos: Ajustar la escala de los datos (normalización, estandarización) para asegurar que todas las variables contribuyan de manera equitativa al análisis.
- Codificación de variables categóricas: Transformar variables categóricas en un formato adecuado para el análisis.
- Reducción de dimensionalidad: Aplicar técnicas como PCA (Análisis de Componentes Principales) para reducir el número de variables manteniendo la mayor cantidad posible de información.

Primera Etapa - Situación actual de la Gerencia

A partir de la inmersión en la organización, la observación directa de los procesos y las entrevistas a la gerente de Créditos y Cobranzas y al encargado de Control de Gestión, se identificó cómo se desarrolla el proceso de otorgamiento y cobranza del crédito. Este proceso abarca desde la atención inicial al cliente en el local hasta la cancelación total del crédito y la entrega de la correspondiente documentación de libre deuda o pagaré.

A continuación, se describen las principales etapas del proceso:

1) Otorgamiento del crédito

1.1 Atención al cliente: El proceso inicia cuando el cliente ingresa al local y selecciona el producto a adquirir. El vendedor toma sus datos y realiza una consulta en el sistema a través

del motor de decisión para determinar si el cliente es apto para acceder al crédito. El sistema puede arrojar tres resultados: *aprobar* (el cliente puede acceder al crédito), *revisar* (hubo un error en la carga de datos) o *rechazar* (el cliente no cumple con los requisitos). En caso de aprobación, se emite un presupuesto y se ofrecen distintos planes de pago. Si el cliente acepta, pasa a la siguiente etapa.

1.2 Instrumentación del crédito y facturación: Un administrativo solicita la documentación necesaria (DNI, boleta de servicio y recibo de sueldo) y completa los datos del cliente en el sistema. Se verifica la identidad del cliente, luego se emiten y firman los documentos correspondientes, entre ellos la solicitud de crédito, el pagaré, contrato de mutuo y fotocopia de DNI. Finalmente, se constituye el legajo de crédito y posteriormente son enviados al Sector de Archivo para auditoría y resguardo. Posteriormente se genera y emite la factura por duplicado.

1.3 Entrega del producto: Una vez procesada la venta, el cliente se dirige con la factura original hacia expedición para retirar el producto en el salón o solicitar el envío a domicilio.

Gráfico N°1: Fragmentación del proceso Otorgamiento del crédito



Fuente: elaboración propia.

2) Cobranza al cliente

Etapas de mora

La gestión de cobranzas se organiza en distintas etapas, según el nivel de mora del cliente y el grado de contacto logrado con este. Cada una de estas etapas tiene como



objetivo garantizar la recuperación de los créditos otorgados. Las principales fases son las siguientes:

A. 11 a 70 días de mora: La gestión de cobranzas se activa automáticamente cuando un cliente tiene 11 días de atraso. Si no se logra contacto mediante llamadas o mensajes se deriva la cuenta a un visitador que entrega una notificación en el domicilio. El visitador debe presentar el acuse de recibo para certificar el contacto con el cliente.

B. 71 días de mora en adelante: Tras haber intentado contactarse por vía telefónica y haber enviado un visitador (como se menciona en el punto 1), además de haber remitido cartas informando sobre la posible derivación a un estudio jurídico para iniciar un embargo, se realiza una última gestión por parte de la empresa. En esta instancia final, se busca contactar al cliente para ofrecerle una oportunidad de regularizar su deuda, brindándole la opción de refinanciarla.

C. Envío a estudio jurídico: El estudio jurídico realiza un tratamiento extrajudicial para recuperar la deuda. Si no se logra el pago, se considera "cuenta agotada" y se inicia el juicio ejecutivo para el embargo.

D. Cliente incobrable: Si tras el juicio no se logra el cobro, el cliente se clasifica como incobrable.

2.1 Gestión de cobranza: El proceso comienza con la asignación automática de las cuentas en mora, que se consideran tales a partir de los 11 días de atraso. Los operadores del call center gestionan estas cuentas mediante llamadas para notificar al cliente sobre las cuotas vencidas y acordar el pago. Todas las interacciones, como mensajes, llamadas, acuerdos de pago o refinanciaciones, se registran en el sistema para crear el historial del cliente. La gestión

abarca desde los 11 hasta los 70 días de mora. Durante este período, se facilitan los datos bancarios para transferencias o los horarios de atención para pagos presenciales en sucursal.

2.2 Verificación del depósito: Si el cliente paga por transferencia bancaria, el administrativo de cobranzas se encarga de verificar que el dinero haya ingresado correctamente a la cuenta de la empresa.

2.3 Registración de la cobranza: Tras verificar el pago, el administrativo de cobranzas registra en el sistema y se envía el recibo al cliente, ya sea por correo electrónico o WhatsApp.

2.4 Cobranza en sucursal: Si el pago se realiza presencialmente, el sistema registra el cobro de forma instantánea y se emite el recibo correspondiente.

Gráfico N°2: Fragmentación del proceso Cobranza del Crédito mediante transferencia bancaria



Fuente: elaboración propia.

Gráfico N°3: Fragmentación del proceso Cobranza del Crédito en sucursal



Fuente: elaboración propia.

3) Entrega de Pagaré o Libre Deuda

3.1 Solicitud del Cliente: Al abonar la última cuota, el cliente puede solicitar el pagaré o libre deuda. Esta solicitud se realiza ante el administrativo de cobranzas si el pago fue por transferencia, o ante el cajero si el pago fue presencial.

3.2 Gestión en el Archivo: El administrativo o cajero envía la solicitud al sector de archivo por correo electrónico, incluyendo el nombre de la cuenta y los datos necesarios del cliente para facilitar la ubicación del legajo de crédito que contiene el pagaré.

3.3 Extracción y Registro: El encargado del archivo extrae el pagaré, registra la extracción en el sistema para dejar constancia y lo entrega al administrativo o cajero, quien se lo proporciona al cliente.

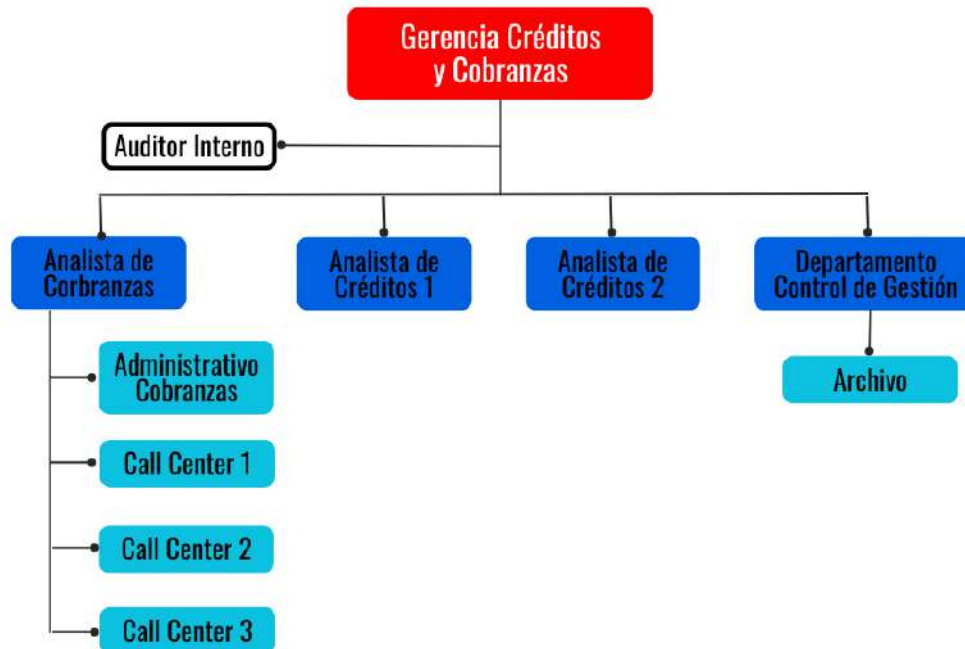
Grafico N° 4: Entrega de Pagaré o Libre Deuda



Fuente: Elaboración propia

A continuación, se presenta el organigrama propuesto para la Gerencia de Créditos y Cobranzas, en el cual se observa la estructura jerárquica y las relaciones funcionales entre los distintos roles y áreas clave:

Gráfico N°5: Organigrama Gerencia de Créditos y Cobranzas



Fuente: Elaboración propia

A lo largo del proceso descrito, el sistema PUMA recopila y registra datos en dos módulos: uno correspondiente a los créditos y otro a la cobranza. Estos datos surgen de cada venta a crédito, consultas en el motor de decisiones, y todas las actividades relacionadas con la gestión de cobranzas.

A partir de las entrevistas realizadas, se obtuvo una visión clara sobre cómo la empresa maneja actualmente estos datos. A continuación, se presentan los principales hallazgos de dichas entrevistas, junto con una nube de palabras que los resume:

- ✓ La gerente realiza tareas diversas, desde enviar mensajes a través de Teleprom (mensajes masivos a clientes en mora), participar en reuniones con el directorio, evaluar resultados y proponer mejoras, creación de procedimientos internos, el diseño de campañas de crédito, supervisión de analistas y del área jurídica para asegurar el cumplimiento de los



indicadores, capacitaciones, responde a consultas de los colaboradores de varias sucursales, controla facturas comerciales, autoriza márgenes de crédito, y asiste a reuniones con asesores.

Debido a la variedad de actividades, muchas de ellas operativas, y a la necesidad de resolver problemas ajenos a su rol, la gerente enfrenta dificultades para dedicarse a temas estratégicos.

- ✓ La gerencia ha comenzado a implementar análisis de datos desde junio de 2023 con la contratación de un asesor financiero con experiencia en empresas del rubro. Previo a esto, las decisiones se tomaban en base a intuiciones, experiencia y consultas en el motor de decisión, sin un análisis exhaustivo de los datos. Esto ocasionó decisiones que afectaron negativamente el rendimiento de la gerencia. Un ejemplo de esto ocurrió en la sucursal de los Valles, caracterizada por atender a un alto volumen de clientes informales (sin recibo de sueldo que respalde sus ingresos). Al conceder créditos a estos clientes, se asumió un riesgo significativo que resultó en una alta morosidad, perjudicando el rendimiento general de la sucursal.
- ✓ Inicialmente, se creía que el problema estaba en la gestión de cobranzas. Sin embargo, se concluyó que el verdadero desafío estaba en el proceso de otorgamiento de crédito. Al mejorar la selección de clientes y conceder créditos a los que presentaban un menor riesgo, la empresa identificó que las cobranzas se facilitarían, ya que los mejores clientes tienden a ser también los mejores pagadores.
- ✓ Se intentó utilizar herramientas como Power BI, pero se encontraron con dificultades para integrar bases de datos y la computadora se ralentiza excesivamente, lo que limita el uso efectivo de la herramienta e impide su aplicación para análisis avanzados.



- ✓ Actualmente, el análisis de datos se realiza mediante reportes generados en Excel y procesados en un dashboard que emplea tablas dinámicas y fórmulas. Sin embargo, se ha observado que la base de datos a veces contiene celdas vacías y datos desactualizados o sin una correcta parametrización. Este sistema ofrece una visión básica de la situación de la empresa, pero presenta limitaciones: el uso intensivo de fórmulas genera errores, especialmente cuando hay cambios en las celdas o el orden de las columnas. Además, el dashboard es extenso y difícil de interpretar rápidamente, lo que dificulta la toma de decisiones en reuniones gerenciales o con el directorio, debido a la falta de gráficos que resuman la información de manera clara.
- ✓ Se señaló la falta de integración entre el área comercial y la Gerencia de Créditos y Cobranzas, lo cual impide trabajar en conjunto para desarrollar estrategias de atracción y retención de clientes. Además, el área comercial ha ejecutado estrategias sin objetivos claros, metodologías definidas ni reportes adecuados, demostrando desconocimiento de las necesidades tanto de la gerencia como de cada sucursal y cliente.
- ✓ Los entrevistados coincidieron en que el análisis de datos es clave para mejorar la toma de decisiones y elaborar mejores planes estratégicos para captar y retener clientes. No obstante, señalaron que se enfrentan a limitaciones debido al desconocimiento sobre herramientas analíticas avanzadas, aunque expresaron interés en implementarlas en el futuro.
- ✓ Finalmente, los indicadores clave para la gerencia incluyen montos otorgados, cantidad de clientes nuevos, renovaciones, cancelaciones, mora general y por sucursal, asignaciones a gestión telefónica y estudios, y refinanciaciones

A partir de observación directa de la sede central y canales de difusión se reconoció lo siguiente:

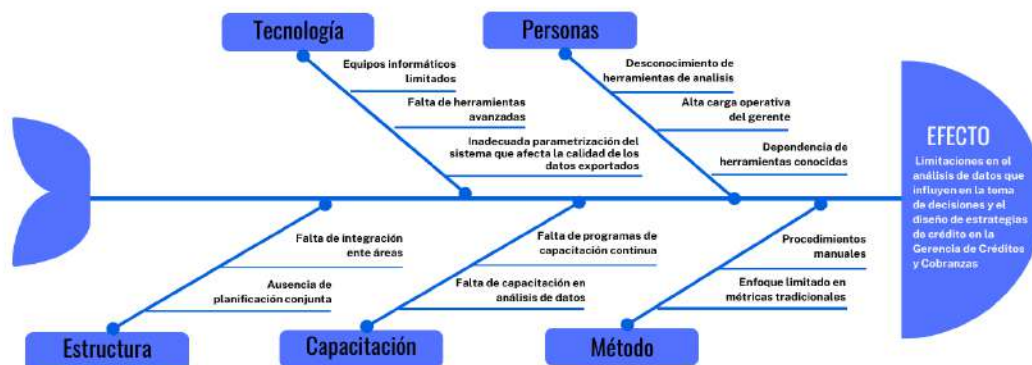
- ✓ La presencia en redes sociales, incluyendo una cuenta activa en Instagram y la explotación de reels, no menciona la promoción de consultas sobre créditos personales. Además, las historias están desactualizadas, mostrando solo dos planes activos como "Cliente Preferencial 30 a 60 días" con sus respectivas condiciones. Tampoco se cuenta con cartelera atractiva que invite a los clientes a consultar sobre créditos personales.

Para identificar y analizar los factores que influyen en la Gerencia de Créditos y Cobranzas en Comfy S.A., a partir de las entrevistas se elaboró un diagrama de cola de pez. Esta herramienta permite desglosar las principales causas que limitan el uso de herramientas de análisis de datos adicionales, una toma de decisiones oportuna y la planificación estratégica de crédito

El diagrama se organiza en varias categorías clave que representan los factores involucrados: Personas, Métodos, Tecnología, Estructura Y Capacitación. Cada categoría muestra sub-causas específicas que contribuyen al problema central.

A continuación, se presenta el diagrama de cola de pez para ilustrar visualmente este análisis de causas raíz:

Gráfico N°6: Diagrama Cola de Pez



Fuente: Elaboración propia

A continuación, se presenta una nube de palabras que resume los términos más repetidos durante las entrevistas

Nube de palabras N°1: Factores Clave en la Gestión de Créditos y Cobranzas



Fuente: elaboración propia

Segunda Etapa - Análisis

Análisis Exploratorio de Datos (AED)

En esta segunda etapa se continuo con el relevamiento de las bases de datos a partir del sistema PUMA. Con base en las fases de procesamiento de datos expuestas en el marco teórico.

Imagen N°1: Base de datos de Créditos otorgados en el mes

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Lim.
1	Sucursal	Fecha	Hora	Factura	Nº Cuenta	Documento	Edad	Telefonos	Validación	T	Ocupacion	Deuda Actua	Importe	Capital	Crédito	Plan	
2	Rivadavia	01/12/2023	17:59:00		17925	GOMEZ, RUP		74	Validado	- 01	Jubilados - P	0	161100	98009	161100		6
3	Rivadavia	01/12/2023	19:16:00		D5608	MORENO, PA		85	Validado	- 01	Privados - R	139290	557160	232779	557160		12
4	Rivadavia	01/12/2023	19:26:00		A2113	ACOSTA, AN		87	Validado	- 01	Jubilados - P	69480	427920	265247	277920		12
5	Rivadavia	01/12/2023	19:33:00		70755	ALVAREZ, CL		54	Validado	- 01	Pensionados	0	51240	28097	51240		7
6	Rivadavia	01/12/2023	19:51:00		A2298	TABARCACHI		61	Validado	- 01	Jubilados - P	0	84280	79501	44280		4
7	Monteros	01/12/2023	10:59:00		10262	RODRIGUEZ,		54	Validado	- 21	Pensionados	0	474436	438777	119236		4
8	Monteros	01/12/2023	11:12:00		Z3701	HERRERA, RA		78	Validado	- 11	Jubilados - P	0	230310	178200	182260		4
9	Monteros	01/12/2023	12:21:00		32868	RODRIGUEZ,		67	Validado	- 01	Jubilados - P	0	156960	95515	156960		6
10	Monteros	01/12/2023	20:40:00		D5625	TOLEDO, DAI		37	Validado	- 01	Adm. Públicos	137912.75	164400	100048	164400		6
11	Monteros	01/12/2023	21:06:00		BC376	ARIAS, NORR		0	Validado	- 01	Adm. Públicos	52009	435498	343774.44	182496		6
12	Bella Vista	01/12/2023	10:56:00		80840	SANCOR COC		61	Manual	- 24/	Sociedad Col	10098608	305000	305000	305000		1
13	Bella Vista	01/12/2023	09:34:00		C6424	BARRIONUEV		69	Validado	- 01	Jubilados - P	68868	275760	115235	275760		12
14	Bella Vista	01/12/2023	10:18:00		AQ789	DECIMA, AN		64	Validado	- 01	Jubilados - P	0	178720	92748	178720		8
15	Bella Vista	01/12/2023	11:56:00		8554	SISTERNA, IH		74	Validado	- 01	Jubilados - P	0	248520	151199	248520		6
16	Bella Vista	01/12/2023	19:19:00		18685	OSORES, GR		53	Validado	- 01	Adm. Públicos	0	129600	79839	129600		6
17	Bella Vista	01/12/2023	19:20:00		18685	OSORES, GR		53	Validado	- 01	Adm. Públicos	0	61320	45983	61320		3
18	Bella Vista	01/12/2023	20:56:00		79152	LOBO, DANIE		31	Validado	- 01	Privados - Cc	79007	366076	182067	316088		12
19	Bella Vista	01/12/2023	21:14:00		68820	GONZALEZ, F		45	Validado	- 01	Privados - Cc	70396	235800	98539	235800		12
20	Trancas	01/12/2023	10:52:00		C3819	FLORES, CLAI		89	Validado	- 31	Informales	0	24077.4	24077.4	18057.4		3
21	Trancas	01/12/2023	10:55:00		31316	CENTENO, GI		67	Validado	- 15	Pensionados	0	74760	45869	74760		6

Fuente: Elaboración propia



La base de datos contiene 47 variables que incluyen información del cliente que recibió el crédito, el tipo de producto seleccionado, detalles de la financiación y datos del motor de decisión SISA. A partir de esta base, se busca conocer las características de los clientes, sus preferencias de crédito, y cómo las sucursales utilizan el margen disponible, con el objetivo de diseñar planes de acción que estimulen un mayor otorgamiento de créditos.

Con el objetivo de comprender las variables y sus características antes de aplicar los modelos, se llevó a cabo un análisis exploratorio de datos. Del conjunto total de variables, se seleccionaron las más relevantes para el análisis. El dataset contiene las siguientes columnas: Sucursal, Edad, Ocupación, Deuda Actual, Importe, Capital, Crédito, Plan, Límite de Crédito, Límite de Cuotas, Proporción Margen, Coeficiente, Cliente Nuevo, Mora Máxima e Ingresos Veraz.

Clasificación de variables

Nombre de Variable	Definición	Tipo
Sucursal	Lugar donde se realizó el crédito.	Cualitativa, nominal.
Edad	Indica la edad del cliente en un rango de 19 a 80 años.	Cuantitativa, continua.
Ocupación	Describe la actividad laboral o situación ocupacional del cliente entre ellos Informales, Jubilados, Monotributistas, Pensionados, Empleados Públicos y Privados.	Cualitativa, nominal.
Deuda Actual	Valor que indica si el cliente posee deuda al momento del otorgamiento del crédito.	Cuantitativa, continua.
Importe	Valor total del producto adquirido, que incluye tanto el precio de contado como el costo de la financiación.	Cuantitativa, continua.
Capital	Valor precio de contado del artículo.	Cuantitativa, continua.
Crédito	Diferencia entre Importe y Tipo (anticipo).	Cuantitativa, continua.
Plan	Plan de cuotas elegido por el cliente para financiar el producto.	Cuantitativa, ordinal



Lim. Crédito	Valor que indica el margen disponible asignado al cliente para gastar.	Cuantitativa, continua.
Lim. Cuota	Valor que indica el límite disponible para el gasto por cada cuota.	Cuantitativa, continua.
Proporción Margen Usado	Cociente entre el importe y el límite de crédito.	Cuantitativa, continua.
Coeficiente	Tipo de plan de pago seleccionado, con o sin anticipo.	Cuantitativa, continua.
ClienteNuevo	Indica si se trata de un cliente nuevo o ya existente	Cualitativa, binaria.
MoraMax	Días de mora máxima del cliente.	Cuantitativa, discreta.
IngresoVeraz	Estimación de los ingresos del cliente, según datos del sistema Veraz.	Cuantitativa, continua.
Familia	Agrupación de productos según categorías comerciales a la cual pertenece el producto adquirido por el cliente.	Cualitativa, nominal.

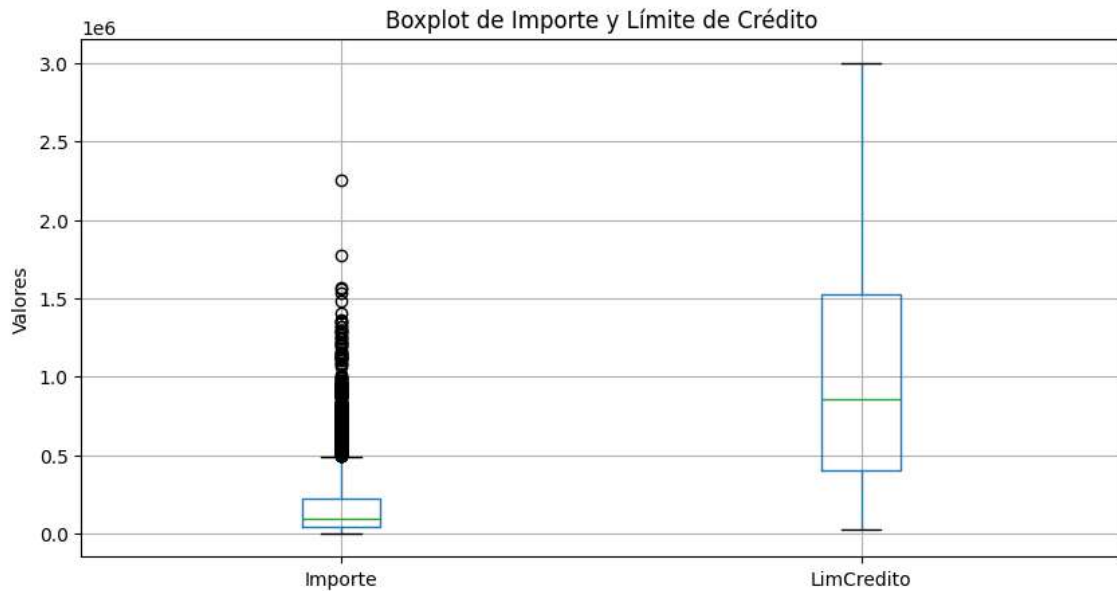
Cantidad de Registros y Columnas

De las variables analizadas, 13 no contenían valores nulos, mientras que la variable "Coeficiente" presentaba 3956 registros con valores faltantes. Además, se identificaron valores atípicos en las variables "Importe" y "Límite de Crédito".

Para el tratamiento de estos problemas, se eliminaron los registros con valores faltantes en la variable "Coeficiente", ya que impedían la correcta ejecución del modelo. Asimismo, se identificaron valores atípicos en las variables "Límite de Crédito" e "Importe" que podían distorsionar el análisis. Por ello, se decidió eliminar estos valores extremos.

Previamente a la aplicación del modelo, se analizó la distribución de las variables numéricas del dataset, tales como el importe del crédito y el límite de crédito otorgado, mediante gráficos boxplot. Estos gráficos ayudaron a identificar los valores atípicos, que luego fueron eliminados con el objetivo de obtener una representación más precisa y estable del comportamiento de los clientes. A continuación, se presenta el gráfico que representa valores extremos:

Gráfico N°7: Gráfico de Caja Importe y Limite de Crédito



Fuente: Elaboración propia

Frecuencias para Variables Cualitativas

En el análisis de las variables cualitativas se identificaron las siguientes características:

Sucursal: las sucursales con mayor actividad son Trancas (1367 registros), Cerrillos (992) y San Fernando (773).

Ocupación: los clientes pertenecen principalmente a los segmentos Público (1912), Privados (1801) y Jubilados (1576).

Coficiente (Tipo de Crédito): el tipo de crédito más utilizado es CRÉDITO SIN ANTICIPO (3232), seguido por CRÉDITO 1 PAGO 30 DÍAS (287).

Primera Operación: la mayoría de los clientes no es nueva, con 5513 registros, mientras que 1498 corresponden a clientes en su primera operación.



Árbol de Decisión

El objetivo es identificar los factores que influyen en la elección del tipo de pago por parte del cliente. Para ello, se aplicaron los modelos Random Forest y árbol de decisión con el fin de determinar las características más relevantes.

Selección de variables: Se inició explorando la base de datos para tener un entendimiento de las variables y su relación. Posteriormente, se seleccionaron las variables independientes que podrían explicar la variable dependiente y resultar útiles para el modelo, entre las cuales se incluyeron el importe, el crédito, el plan de pagos y el tipo de pago.

X	Variable	Definición
	Importe	Valor correspondiente al precio de contado más la financiación.
	Crédito	Diferencia entre Importe y Tipo (anticipo).
	Plan	Plan de cuotas elegido por el cliente para financiar el producto.
	ImporteCuota	Cociente entre Importe y Plan.
Y	Tipo	Tipo de pago elegido (Crédito con anticipo, sin anticipo, entrega de efectivo o un pago en 30 días).

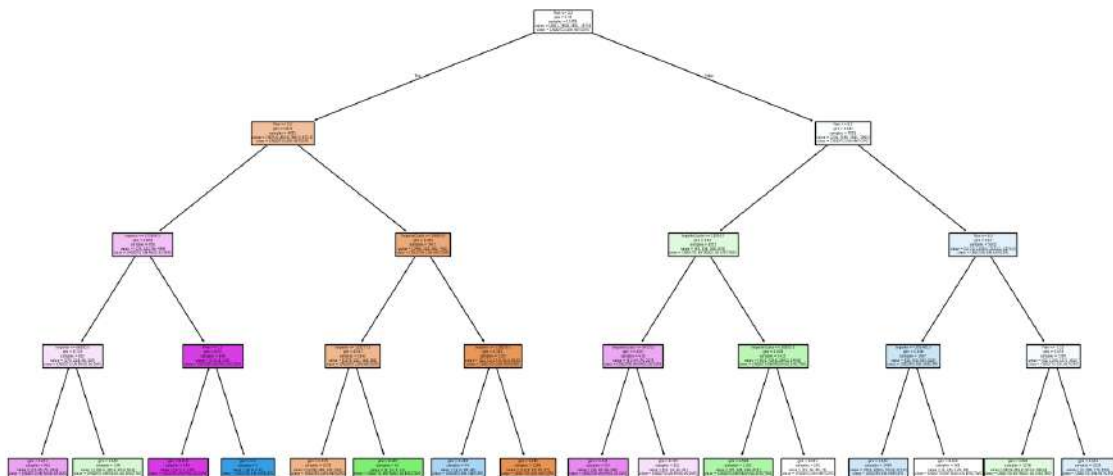
Aplicación del modelo: Se aplicó el modelo Random Forest. La principal característica que impacta en la decisión del tipo de pago son: ImporteCuota e Importe fueron las características más influyentes (aproximadamente un 39% cada una), Plan de Cuotas también fue importante, aunque menos (22.5%), por último crédito fue la menos influyente.

- **ImporteCuota:** esta es la característica más importante. Esto indica que el valor de la cuota tiene un peso considerable para determinar el tipo de pago. Los clientes probablemente eligen el tipo de financiamiento en función de cuánto tienen que pagar en cada cuota.
- **Importe:** es la segunda característica más importante, esto sugiere que el monto total de la compra es crucial para la elección del tipo de pago. Los clientes pueden decidir el tipo de pago según el valor total que deben financiar.

- **Plan:** el número de cuotas no es el factor principal, pero sigue siendo relevante para la decisión final.

Luego se decidió realizar un árbol de decisión a partir del modelo de árbol forest para poder visualizar la información. A continuación, se presenta el árbol de decisión:

Grafico N°8: Árbol de decisión Preferencia de tipo de pago



Fuente: Elaboración propia con Google Colab

Nivel 1 y 2: la primera división se basó en el Plan de Cuotas. Esto indica que la cantidad de cuotas que el cliente elige es un factor determinante en la decisión sobre el tipo de pago. A menor número de cuotas generalmente se asocia con tipos de pago más directos o con menos interés (como un pago a 30 días o pago en efectivo).

Nivel 3: el árbol dividió según el Importe del Crédito, esto sugiere que el valor del producto también influye en la decisión. A mayor importe, los clientes parecen preferir créditos con anticipo para minimizar el costo total. (como crédito con anticipo o pago en efectivo)

Nivel 4: se hicieron divisiones en base al importe de cada cuota, lo cual refleja que los clientes tienen en cuenta cuánto tendrán que pagar mensualmente. Un importe de cuota más bajo



podría hacer más atractivas ciertas opciones de financiamiento, un importe más alto tiende a ser más atractivo no realizar ningún anticipo.

- ✓ Se observa que a medida que disminuye la cantidad de cuotas y el importe de cada una, aumenta la preferencia por realizar pagos parciales o pagos a 30 días. De acuerdo con la experiencia, este tipo de pago suele aplicarse cuando el valor del producto financiado es bajo. Por otro lado, un mayor número de cuotas implica una financiación más prolongada y un costo por cuota más elevado, siendo esta la situación más común.

Cluster

Cluster de Clientes según sus preferencias de compra

El objetivo de este análisis es explorar patrones y segmentar a los clientes en grupos que compartan características comunes. Para ello utilizamos el método de clustering de K-means sobre los datos proporcionados, analizando variables como edad, ocupación, importe de compra, plan de cuotas, si se trata de un cliente nuevo o existente y familia de producto que elige.

Metodología

Selección de variables: Se seleccionaron las variables a utilizar entre ellas: edad, ocupación, importe, plan, cliente nuevo y familia. Estas variables fueron preparadas y organizadas previamente en una base de datos en formato Excel. La variable “Familia” está compuesta por diferentes tipos de productos: Audio, Bazar, Bebes, Blanco, Calzado, Camping y Jardín, Climatización, Colchones, Equipamientos Comerciales, Herramientas, Hogar, Indumentaria, Informática, Juguetes, Muebles, Muebles de Oficina, Pequeños Electrodomésticos, Rodados, TV, Telefonía y Video Juegos. Dado que



“Familia” puede tomar diferentes valores, se decidió agrupar los productos y la ocupación en categorías que facilitaran la comprensión y el análisis del modelo. A continuación, se presenta la clasificación utilizada para agrupar las variables:

Agrupación variable “Familia”:

- **Grandes Electrodomésticos:** se incluyeron los productos relacionados con Climatización, TV, Hogar y Equipamientos Comerciales.
- **Tecnología:** se incluyeron los productos relacionados con Audio, Informática, Telefonía y Pequeños Electrodomésticos.
- **MueblesHogar:** se incluyeron los productos relacionados con Muebles, Muebles de Oficina, Colchones, Blanco y Bazar.
- **Moda:** se incluyeron los productos relacionados con Calzado e indumentaria
- **Infantes:** se incluyeron los productos relacionados con Bebés, Video Juegos y Juguetería.
- **Exterior:** se incluyeron los productos relacionados con Camping y Jardín, Herramientas y Equipos de gimnasia.
- **Rodados:** se mantuvo como Rodados.

Imagen N° 2: Preparación de base de datos

	A	B	C	D	E	F	G
1	Edad	Ocupacion	Importe	Plan	ClienteNuevo	IngresoVeraz	Familia
2	74	Jubilados	161100	6	NO	188562	Tecnologia
3	65	Privados	557160	12	SI	188562	Grandes Electrodomesticos
4	67	Jubilados	427920	12	NO	107810	Muebles
5	54	Pensionados	51240	7	NO	107810	Tecnologia
6	61	Jubilados	84280	4	NO	107810	Tecnologia
7	54	Pensionados	474436	4	NO	107810	Tecnologia
8	78	Jubilados	230310	4	NO	280724	Muebles
9	37	Publica	164400	6	SI	107810	Grandes Electrodomesticos
10	62	Publica	435498	6	NO	107810	Tecnologia
11	69	Jubilados	275760	12	NO	188562	Muebles
12	64	Jubilados	178720	8	NO	107810	Grandes Electrodomesticos
13	74	Jubilados	248520	6	NO	107810	Grandes Electrodomesticos
14	53	Publica	129600	6	SI	464368	Grandes Electrodomesticos
15	53	Publica	61320	3	SI	464368	Hogar
16	31	Privados	366076	12	NO	188562	Tecnologia
17	46	Privados	235800	12	NO	188562	Tecnologia
18	40	Informales	24077.4	3	NO	33000	Moda
19	67	Pensionados	74760	6	NO	33000	Moda
20	56	Pensionados	101398	5	NO	280724	Moda
21	56	Publica	59177.8	3	NO	188562	Moda

Fuente: Elaboración propia

La base de datos fue cargada en Google Drive para su análisis en Google Colab. Se realizó la normalización de las variables numéricas para garantizar escalas uniformes. Para aplicar el modelo de clustering, fue necesario transformar las variables categóricas en valores numéricos mediante codificación. Se asignó un número único a cada categoría de las variables ClienteNuevo, Familia y Ocupacion. La codificación utilizada fue la siguiente:

Imagen N°3: Codificación Variables Categóricas

```
Codificación para la variable 'ClienteNuevo':  
NO: 0  
SI: 1  
  
Codificación para la variable 'Familia':  
Exterior: 0  
GrandesElectrodomesticos: 1  
Infantes: 2  
Moda: 3  
MueblesHogar: 4  
Rodados: 5  
Tecnologia: 6  
  
Codificación para la variable 'Ocupacion':  
Informales: 0  
Jubilados: 1  
Monotributista: 2  
Pensionados: 3  
Privados: 4  
Publico: 5
```

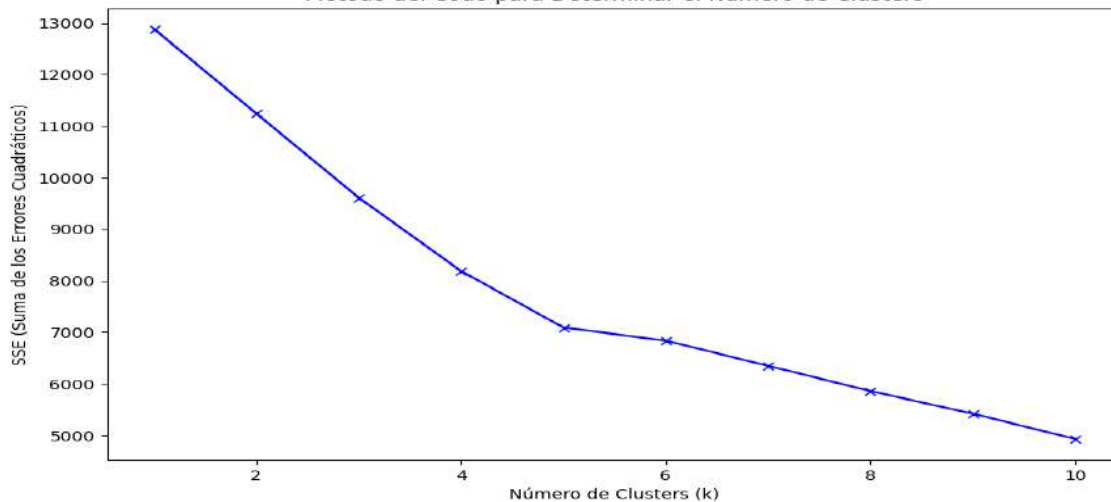
Fuente: Elaboración propia con Google Colab

Esta codificación se utilizó para garantizar la correcta aplicación del modelo, ya que este solo opera con variables numéricas.

Selección del Número de Clusters: Se utilizó el método del codo para determinar el número óptimo de clusters. En el gráfico del codo, se observa que el ángulo más pronunciado se daba en el valor 4, por lo que se optó por trabajar con cuatro clusters.



Gráfico N°8: Gráfico de Codo
 Método del Codo para Determinar el Número de Clusters



Fuente: Elaboración propia con Google Colab

Aplicación de K-means: se empleó el algoritmo K-Means con el propósito de segmentar a los clientes en grupos homogéneos según sus características.

Análisis de los Clusters: se obtuvieron estadísticas descriptivas para cada cluster, lo que nos permitió identificar algunas diferencias entre los grupos. Los resultados obtenidos muestran cuatro clusters principales, los cuales fueron interpretados en términos de edad, nivel de ingresos, preferencia por planes de cuotas (planes cortos, medios, largos), ocupación y preferencias de productos.

Tabla N°1: Estadísticas Descriptivas

Cluster	Edad	Importe	Plan	ClienteNuevo	IngresosVeraz	Informales	Jubilados	Monotributista	Pensionados	Privados	Publico	Exterior	GrandesElectrodomesticos	Infantes	Moda	MueblesHogar	Rodados	Tecnologia
0	0.35	0.02	0.17	0.11	0.06	0.46	0	0	0.09	0.17	0.28	0.01	0.02	0.02	0.85	0.03	0.01	0.07
1	0.81	0.06	0.36	0.22	0.12	0	1	0	0	0	0	0.06	0.31	0.01	0.12	0.14	0.03	0.34
2	0.44	0.07	0.29	0.26	0.11	0	0	0	0.17	0.43	0.4	0.05	0.28	0.04	0	0.15	0.06	0.43
3	0.42	0.06	0.22	0.27	0.16	0	0	1	0	0	0	0.01	0.23	0.03	0.33	0.16	0.04	0.2

Fuente: Elaboración propia

Se procedió a obtener información sobre la proporción de clientes nuevos y existentes, así como la preferencia de planes de financiamiento, dado que la tabla de estadísticas descriptivas no



permite analizar variables categóricas. Las estadísticas descriptivas se enfocan en valores promedio y generales, mientras que para identificar proporciones y preferencias fue necesario usar tablas de frecuencia, que permiten un análisis más detallado de las categorías dentro de cada cluster.

Tabla N°2: Proporción de clientes nuevos y existentes en porcentaje

Proporción de clientes nuevos y existentes en cada cluster (en %):		
ClienteNuevo	0	1
Cluster		
0	88.552559	11.447441
1	77.961783	22.038217
2	73.940435	26.059565
3	72.649573	27.350427

Fuente: Elaboración propia con Google Colab

*Referencia: Clientes existentes=0; Clientes nuevos=1

Tabla N°3: Preferencia de planes de financiamiento en cada cluster

Preferencia de planes de financiamiento en cada cluster (valores originales):		
Cluster	Plan más Frecuente (original)	Porcentaje que Prefiere este Plan (%)
0	11.052632	49.257017
1	12.894737	27.006369
2	11.315789	25.916380
3	11.052632	25.641026

Fuente: Elaboración propia con Google Colab

Clusters identificados

- Cluster 0: este cluster está integrado por clientes jóvenes con ingresos. Una gran proporción de estos clientes son trabajadores informales. Su interés principal se centra en productos de Moda entre ellos indumentaria y calzados. Tienen una gran preferencia por planes de aproximadamente 11 cuotas y está formado por clientes existentes.
- Cluster 1: está compuesto por clientes adultos mayores por jubilados, quienes presentan un nivel de ingresos moderado notablemente superior al Cluster 0 y prefieren optar por planes de hasta 13 cuotas. Este grupo está formado en su gran mayoría por clientes existentes y tiene una fuerte preferencia por productos de Grandes Electrodomésticos y Tecnología.



- Cluster 2: este grupo está formado por clientes de edad media con ingresos moderados. Estos clientes prefieren utilizar planes medios de hasta 12 cuotas. Está compuesto por clientes existentes y que trabajan en los sectores privado y público. Los productos más comprados por este cluster son Tecnología, Grandes Electrodomésticos y Muebles para el hogar.
 - Cluster 3: los clientes de este grupo tienen una edad intermedia e ingresos altos. Tienen preferencia por planes de hasta 11 cuotas. Su ocupación predominante son los monotributistas, son clientes que ya operaron alguna vez en la empresa y tienen una preferencia por la categoría Moda y Grandes Electrodomésticos.
- ✓ Teniendo en cuenta que esta base de datos corresponde a los últimos meses del año, se considera que los resultados obtenidos son coherentes con la situación de la empresa. Por ejemplo, se observa que, mes a mes, las personas de edad media tienden a mostrar una preferencia por artículos para equipar su hogar. En el caso de los jubilados, se evidencia una inclinación hacia productos tecnológicos, ya que frecuentemente prestan su cuenta para adquirir productos en nombre de familiares o realizan compras de teléfonos celulares para sus nietos. Respecto a los jóvenes, suelen tener menor estabilidad económica debido a la naturaleza de su ocupación, que en muchos casos es informal, tienden a optar por planes un año de duración o más. En cuanto a los clientes que trabajan en el sector público y privado, presentan cierta estabilidad financiera debido a su ocupación, lo cual les permite optar por planes de financiamiento cortos y medios.
- ✓ Otro aspecto relevante observado a partir de los datos es que la empresa, en general, se maneja principalmente con clientes existentes, lo que indica una dificultad para atraer

nuevos clientes. Esto representa una oportunidad potencial para implementar estrategias de atracción de clientes en todos los clusters analizados.

- ✓ Por otro lado, en los cuatro grupos analizados se observa una preferencia significativa hacia los planes de 11 a 12 cuotas, principalmente los del grupo formado por clientes informales.

Cluster de Sucursal según proporción de margen de crédito utilizado

El propósito de este análisis es examinar la proporción de utilización del margen disponible considerando distintas variables, tanto numéricas como categóricas.

Metodología

Selección de variables: Se seleccionaron las variables a utilizar, tales como sucursal, ocupación, deuda actual, importe, límite de crédito, ingreso estimado y proporción.

Imagen N° 4: Preparación de base de datos

1	Sucursal	Ocupacion	Deuda Actual	Importe	LimCredito	IngresosVeraz	Proporcion
2	Rivadavia	Publico	0	41160	3000000	107810	0.014
3	Rivadavia	Privados	0	96300	295000	280724	0.326
4	Rivadavia	Privados	0	255240	3000000	863396	0.085
5	Monteros	Jubilados	0	105550	3000000	464368	0.035
6	Monteros	Privados	0	142350	297000	188562	0.479
7	Monteros	Publico	108760	405200	601902	33000	0.673
8	Monteros	Publico	0	38220	250000	33000	0.153
9	Trancas	Pensionados	207498	13311	225000	33000	0.059
10	Trancas	Publico	0	20516	390000	107810	0.053
11	Trancas	Pensionados	381832	254520	644805	33000	0.395
12	Trancas	Privados	0	42341.6	52677.98	33000	0.804
13	Trancas	Informales	0	43536	102000	33000	0.427
14	Trancas	Informales	104370	19170	160000	107810	0.120
15	Trancas	Publico	0	210720	1137525	188562	0.185
16	Trancas	Pensionados	243848	17540.6	154000	280724	0.114
17	Trancas	Informales	0	21082	128000	107810	0.165
18	Trancas	Publico	0	41130	540000	280724	0.076
19	Cerrillos	Publico	1271160	45935	2390317.5	280724	0.019
20	Cerrillos	Pensionados	75960	29490	510000	107810	0.058
21	Cerrillos	Jubilados	0	40966.6	1340572.5	464368	0.031

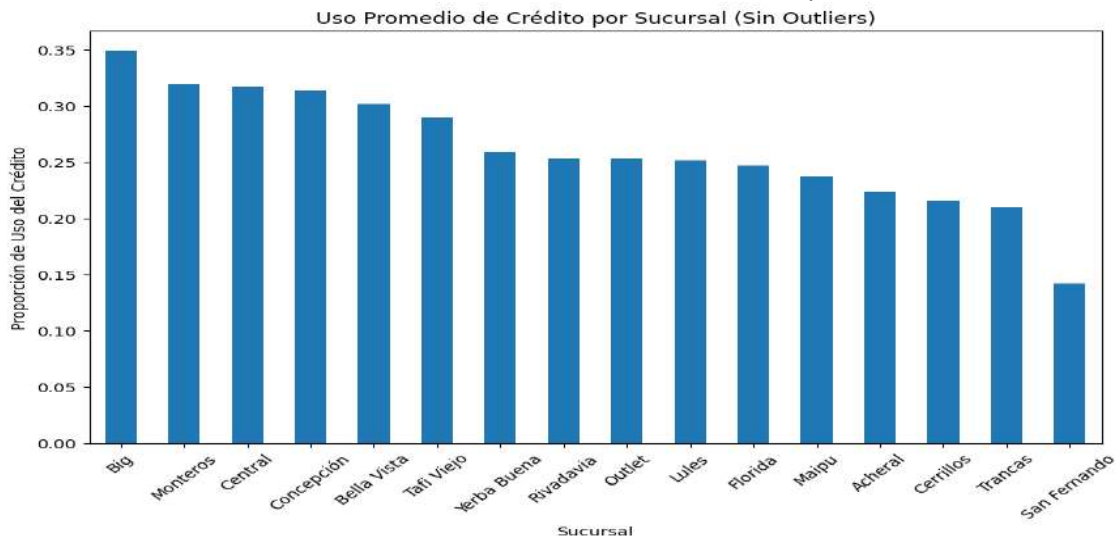
Fuente: Elaboración propia

Posteriormente, se cargó la base de datos en Google Drive para poder utilizar Google Colab.



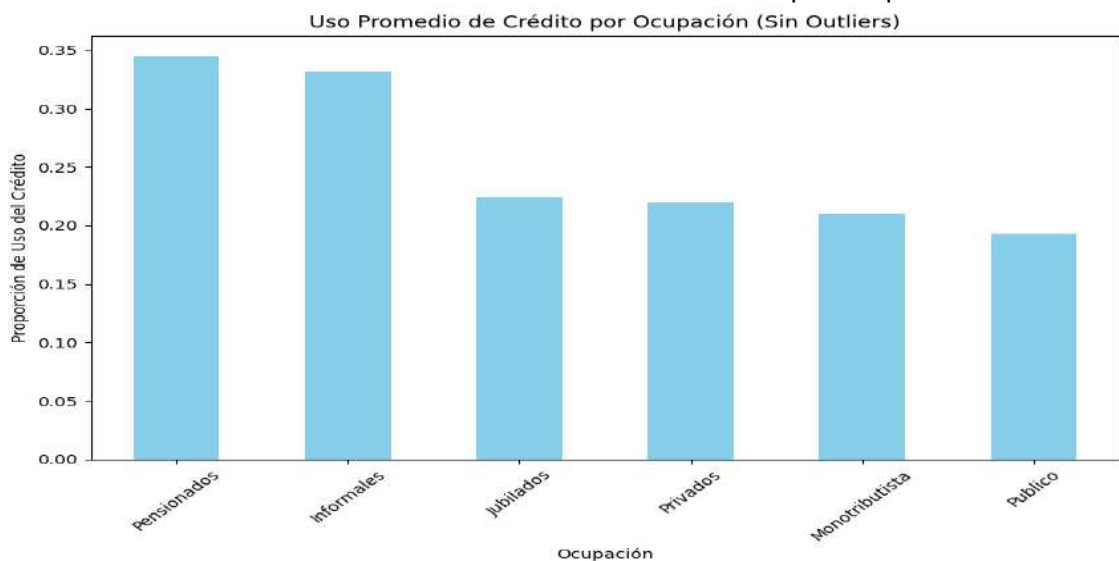
Luego, se analizó el uso promedio del crédito, tanto por sucursal como por ocupación. Para ello, se utilizó la proporción del crédito utilizado y se realizó una agrupación por sucursal y ocupación. De esta forma, se identificaron las sucursales y ocupaciones con un mayor promedio de uso del crédito, lo que permitió reconocer diferencias significativas en el comportamiento de los clientes según su ubicación y perfil laboral. A continuación, se presentan los gráficos:

Gráfico N°10: Gráfico de barras Uso Promedio del Crédito por Sucursal



Fuente: Elaboración propia con Google Colab

Gráfico N°11: Gráfico de barras Uso Promedio de Crédito por Ocupación



Fuente: Elaboración propia con Google Colab



- ✓ Se observa que algunas sucursales, como "Big" y "Monteros", presentaban un uso más intensivo del crédito, mientras que otras, como "San Fernando", mostraban un menor uso. En cuanto a las ocupaciones, se encontró que los pensionados e informales tendían a utilizar una mayor proporción de su crédito disponible.

Aplicación de Clustering: Para identificar grupos de clientes con comportamientos similares, se aplicó el algoritmo de clustering K-means utilizando únicamente variables numéricas relacionadas con características financieras. Los clientes se agruparon en tres clusters, cada uno representando un grupo con características similares en cuanto al uso del crédito. Posteriormente, se relacionaron estos grupos con las variables categóricas "Sucursal" y "Ocupacion".

Tabla N°4: Estadísticas Descriptivas

Cluster	Importe	LimCredito	IngresoVera	Proporcion
0	144935	2224381	625744	0,079
1	481673	1372035	264298	0,43
2	99167	715284	143668	0,24

Fuente: Elaboración propia

- Cluster 0: Representado por clientes con un límite de crédito alto e ingresos elevados, pero que utilizan una proporción pequeña de su crédito (7.9%). Las sucursales como San Fernando y Cerrillos parecen tener una mayor concentración de este tipo de clientes, lo que indica un comportamiento conservador en el uso de crédito.
- Cluster 1: Este grupo tiene un límite de crédito menor que el Cluster 0, con ingresos medios. Sin embargo, la proporción de crédito utilizado es la más alta (43.6%), lo cual indica que estos clientes tienden a aprovechar más su crédito disponible. Este cluster representa clientes que aprovechan más su límite de crédito en comparación con otros grupos. Las sucursales como Bella Vista y Lules tienen clientes que tienden a utilizar más el crédito disponible



- Cluster 2: Los clientes en este cluster tienen ingresos menores y un límite de crédito moderado. La proporción de crédito utilizado también es moderada, aproximadamente un 23.5%. La gran cantidad de clientes en Trancas indica que esta sucursal maneja principalmente a este tipo de clientes, que no usan tanto su crédito.
- ✓ Se observa que las ocupaciones más predominantes según el modelo son Jubilados, Público y Privados, las cuales se destacan en los tres grupos identificados. En el cluster 2, se destaca la presencia además de clientes informales y pensionados, lo que podría indicar que estos clientes no cuentan con ingresos tan estables como los de otros clusters, lo cual representa una oportunidad para incentivar un mayor uso del crédito otorgado.
- ✓ Respecto al grupo 1, los clientes de este segmento son los que hacen un mayor uso del crédito disponible en comparación con otros grupos, lo que podría indicar una mayor dependencia del financiamiento. Dado que son clientes compradores se podría ofrecer productos de mayor costo o recompensas que premien la lealtad de estos clientes.
- ✓ En cuanto al grupo 0, se podría implementar estrategias para incentivar el uso del crédito a través de promociones o beneficios específicos, con el objetivo de aumentar la proporción de utilización del crédito disponible y captar su atención.

Recomendaciones

El análisis de datos constituye una herramienta esencial para la toma de decisiones en la organización, ya que permite interpretar información clave sobre el comportamiento de los clientes, identificar patrones y explorar oportunidades de crecimiento. La capacidad de comprender las preferencias y características de los clientes actuales y potenciales permite a la



empresa orientar sus esfuerzos hacia estrategias efectivas de fidelización y captación, logrando así una ventaja competitiva en el mercado.

A partir de los resultados obtenidos, se formulan las siguientes recomendaciones para fortalecer la relación con los clientes actuales y atraer nuevos segmentos:

- ✓ Los clientes actuales entre ellos adultos mayores y personas con solvencia económica representan la base más sólida de la empresa, ya que una proporción significativa renueva sus créditos de manera recurrente. Implementar un sistema de puntos acumulativos fortalecerá esta relación, incentivando compras futuras con beneficios como descuentos, productos exclusivos o servicios adicionales. Esto no solo refuerza la fidelidad, sino que también genera una percepción positiva hacia la empresa al brindar un valor tangible por su lealtad.
- ✓ Los datos reflejan que los jóvenes constituyen un mercado potencial. La implementación de planes de crédito adaptados, con cuotas flexibles (planes medios de en promedio 11 cuotas) y descuentos iniciales, facilitará su acceso a los productos. Paralelamente, fortalecer la comunicación visual y digital, mediante campañas atractivas en redes sociales y contenido dinámico en la página web es esencial destacar en estas plataformas los beneficios de los créditos personales, integrándolos en el entorno digital que estos clientes frecuentan.
- ✓ Modernizar los puntos de venta con pantallas LED y material visual enfocado en los beneficios de los créditos personales generará un impacto directo en la experiencia del cliente. Este enfoque visual conecta especialmente con los jóvenes, mientras que folletos físicos o digitales, adaptados para los adultos mayores, aseguran que cada segmento reciba la información de manera efectiva.



- ✓ La centralidad del cliente debe ser el eje del trabajo conjunto entre la Gerencia de Créditos y Cobranzas y el área comercial. A través de un enfoque colaborativo, se puede lograr una personalización que potencie tanto la captación como la retención de clientes.
- ✓ Incorporar un sistema de gestión de relaciones con el cliente (CRM) permitirá gestionar y analizar datos de manera más eficiente, mejorando la segmentación y personalización de las ofertas.
- ✓ Los monotributistas, identificados como un segmento con ingresos significativos y menos frecuente en los créditos representan una oportunidad para diversificar la base de clientes. Ofrecerles beneficios exclusivos y condiciones ajustadas a sus necesidades posicionará a la empresa como una opción confiable y atractiva para este grupo.

Conclusiones

Este trabajo representa un primer paso hacia la comprensión del valor que aporta el análisis de datos como herramienta estratégica dentro de la Gerencia de Créditos y Cobranzas. Se ha identificado que el uso de estas herramientas no solo permite conocer en profundidad el comportamiento y las características de los clientes, sino que también favorece la toma de decisiones informadas y la formulación de estrategias orientadas a mejorar la competitividad de la empresa.

El interés de la gerencia por avanzar en la adopción de un enfoque centrado en el análisis de datos destaca como un factor clave. Sin embargo, se han identificado brechas significativas en la colaboración entre áreas, lo que limita el aprovechamiento pleno del conocimiento generado. La ausencia de sinergia y trabajo en equipo puede convertirse en una barrera para la ejecución efectiva de estrategias centradas en el cliente quien es el pilar fundamental de la empresa.



Por otro lado, se reconoce el compromiso de su personal, que apuesta a la mejora continua y al aprendizaje constante, así como en la existencia de sistemas robustos capaces de generar datos valiosos. Estos datos, si son correctamente gestionados, tienen la capacidad de transformarse en información útil y, posteriormente, en conocimiento estratégico que impulse el desarrollo de la organización.

Anexo

<p>Guía de entrevista Entrevistado: Gerente de Créditos y Cobranzas</p>	<p>Objetivo: Conocer en profundidad el funcionamiento del departamento, los procesos de manejo de datos, la toma de decisiones y los problemas observados.</p>
<p>1. Sobre el funcionamiento actual del departamento:</p> <ul style="list-style-type: none"> • ¿Podrías describir cómo funciona el departamento de créditos y cobranzas en su día a día? • ¿Cómo se distribuyen las tareas y responsabilidades dentro del equipo? • ¿Cuáles son los principales procesos que sigue el departamento para gestionar los créditos y cobranzas? <p>2. Sobre la gestión de datos actual:</p> <ul style="list-style-type: none"> • ¿Qué tipo de datos son los más relevantes para el departamento en su operativa diaria? • ¿Cómo se recolectan, organizan y analizan actualmente los datos en el departamento? • ¿Qué dificultades enfrentas actualmente al gestionar grandes volúmenes de datos? <p>3. Toma de decisiones:</p> <ul style="list-style-type: none"> • Cuando se debe tomar una decisión ¿En qué te basas para tomarla? ¿Cómo utilizas los datos para respaldar esas decisiones? • ¿Qué tipo de análisis o informes son necesarios para que puedas tomar decisiones más informadas? • ¿Sientes que tienes acceso a la información necesaria en el momento adecuado para tomar decisiones? <p>4. Problemas y desafíos observados:</p> <ul style="list-style-type: none"> • ¿Cuáles son los mayores desafíos o problemas que observas actualmente en el proceso de gestión de créditos y cobranzas? • ¿Hay alguna fase del proceso donde identifiques más dificultades o cuellos de botella? • ¿Qué cambios crees que podrían mejorar la eficiencia y efectividad del departamento? <p>5. Herramientas útiles y necesidades:</p> <ul style="list-style-type: none"> • ¿Consideras que el departamento cuenta con las herramientas de análisis adecuadas para gestionar los créditos y cobranzas? ¿Por qué? 	



- Si pudieras implementar nuevas herramientas de análisis ¿cuáles crees que serían más útiles para optimizar el trabajo?
- ¿Qué características crees que debería tener una nueva herramienta de análisis para mejorar la toma de decisiones en el departamento?

6. Conclusión:

- ¿Hay algún tema que no hayamos tratado y que consideres relevante para entender mejor el funcionamiento del departamento?

<p>Guía de entrevista Entrevistado: Encargado de Control de Gestión</p>	<p>Objetivo: Entender el proceso de análisis de datos en el departamento de créditos y cobranzas, las técnicas utilizadas y explorar oportunidades de mejora en la gestión de datos.</p>
<p>1. Sobre el análisis de datos:</p> <ul style="list-style-type: none"> • ¿Qué herramientas y métodos utilizas para analizar los datos? ¿Qué tipo de análisis realizas con los datos del departamento? (Ej. análisis de tendencias, segmentación de clientes, ratios de morosidad, etc.) • ¿Cuáles han sido los hallazgos más significativos que has logrado a partir del análisis de datos? • ¿Cómo decides qué datos o variables son los más relevantes para cada informe? <p>2. Desafíos y oportunidades:</p> <ul style="list-style-type: none"> • ¿Cuáles son los mayores desafíos que enfrentas al momento de analizar los datos o generar los informes? • ¿Consideras que el análisis de datos podría ser más eficiente con alguna herramienta o metodología adicional? ¿Cuáles serían? <p>3. Sobre los informes mensuales:</p> <ul style="list-style-type: none"> • ¿Qué información sueles incluir en los informes mensuales del departamento? ¿Qué indicadores clave (KPI) consideras? • ¿Sientes que los informes actuales reflejan de manera completa el estado del departamento? ¿Qué aspectos crees que podrían mejorarse o incluirse? <p>4. Toma de decisiones basada en datos:</p> <ul style="list-style-type: none"> • ¿Qué tipo de decisiones importantes crees que se toman en el departamento a partir de los informes y análisis de datos que realizas? • ¿Sientes que los datos actuales permiten tomar decisiones informadas? <p>5. Conclusión:</p> <ul style="list-style-type: none"> • ¿Hay algún aspecto del análisis de datos o los informes mensuales que no hayamos mencionado y que consideres relevante mencionar? 	

Referencias

Asistic. (2024). *Organizaciones basadas en datos – Conocimientos de las TIC aplicados a la*

Sociedad del Siglo XXI. <https://sistemas.acis.org.co/index.php/sistemas/issue/view/31/29>



Cultura Organizacional Basadas en Datos. (2024). *Understanding and developing a data-informed culture*. Educause.

Davenport, T. H., & Prusak, L. (1998). *Conocimiento en acción: Cómo las organizaciones gestionan lo que saben*. Harvard Business Press.

Diestra Quinto, N. M., Cordova Villodas, A. J., Caruajulca Montero, C. P., Esquivel Cueva, D. L., & Nina Vera, S. A. (2023). *La inteligencia artificial y la toma de decisiones gerenciales [Artificial intelligence and managerial decision making]*. *Revista Iberoamericana de Ventas Aplicadas*, 11(3), 12-22. https://riva.upeu.edu.pe/index.php/ri_va/article/view/1631/1914

Few, S. (2012). *Muéstrame los números: Diseñando tablas y gráficos para iluminar* (2da ed.). Analytics Press.

Gartner. (2023). *Cuadrante Mágico para Plataformas de Ciencia de Datos y Aprendizaje Automático*.

Hernández Sampieri, R. (2018). *Metodología de la investigación: Las rutas cualitativas, cuantitativas y mixtas*. McGraw-Hill.

Innovación Digital 360. (s.f.). *Análisis de datos: Técnicas y metodologías para la aplicación de analytics*. <https://www.innovaciondigital360.com/big-data/analisis-de-datos-tecnicas-y-metodologias-para-la-aplicacion-de-analytics/>



Muñoz Cavero, J. A., & Luyo Pérez, A. E. (2022). *Sistema de gestión de evaluación crediticia en el sector cooperativo de ahorro y crédito con análisis predictivo usando machine learning. Universidad Peruana de Ciencias Aplicadas (UPC).*

Patiño Pérez, D., Silva Bustillos, R., Munive Mora, C., & Botto Tobar, M. (2023). *Predicción de Covid-19 con el uso del Algoritmo Random Forest y Redes Neuronales Artificiales.*

Provost, F., & Fawcett, T. (2013). *Ciencia de datos para los negocios: Lo que necesitas saber sobre minería de datos y pensamiento analítico de datos. O'Reilly Media.*

Ramírez Mendoza, D. Y. (2022). *Métodos de machine learning con algoritmos de clúster no supervisados: Una alternativa de segmentación de las pymes colombianas para plantear estrategias de acuerdo con sus condiciones económicas. Repositorio de la Universidad EAFIT.*
<https://repository.eafit.edu.co/server/api/core/bitstreams/ff833e45-7eaa-4c8f-be8e-fa2245e08bf2/content>

Sandoval, L. J. (2018). *Algoritmos de aprendizaje automático para análisis y predicción de datos. Escuela de Ingeniería en Computación.*
<https://d1wqtxts1xzle7.cloudfront.net/88252094/162564840-libre.pdf>