



**FACULTAD DE POSTGRADO
TRABAJO FINAL DE GRADUACIÓN**

**PERFILAMIENTO DE CLIENTES EMPLEANDO
HERRAMIENTAS DE ANALITICA DE NEGOCIOS PARA 5
MIPYMES EN LA CIUDAD DE SAN PEDRO SULA, HONDURAS**

SUSTENTADO POR:

**SAMARIA POLETH HERNANDEZ VASQUEZ
NELSON ANTONIO CABRERA BONILLA**

PREVIA INVESTIDURA AL TÍTULO DE

**MÁSTER EN
ANALITICA DE NEGOCIOS**

SAN PEDRO SULA, CORTES, HONDURAS, C.A.

SEPTIEMBRE, 2023

UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA

UNITEC

FACULTAD DE POSTGRADO

AUTORIDADES UNIVERSITARIAS

RECTORA

ROSALPINA RODRÍGUEZ GUEVARA

PRORECTOR/ SECRETARIO GENERAL

ROGER MARTÍNEZ MIRALDA

VICERRECTOR ACADÉMICO

JAVIER ABRAHAM SALGADO LEZAMA

DIRECTORA NACIONAL DE POSTGRADO

ANA DEL CARMEN RETTALLY

**PERFILAMIENTO DE CLIENTES EMPLEANDO
HERRAMIENTAS DE ANALITICA DE NEGOCIOS PARA
5 MIPYMES EN LA CIUDAD DE SAN PEDRO SULA,
HONDURAS**

**TRABAJO PRESENTADO EN CUMPLIMIENTO DE LOS
REQUISITOS EXIGIDOS PARA OPTAR AL TÍTULO DE
MÁSTER EN**

ANALITICA DE NEGOCIOS

ASESOR METODOLÓGICO

MsC. ALBA GABRIELA GARAY ROMERO

ASESOR TEMATICO

MAE. KEVIN EDUARDO FÚNEZ FÚNEZ

MIEMBROS DE LA TERNA:

MAE. CARLOS AMADOR

MAE. JOSUE MEGIA

MAE. GERARDO LUJANO

DERECHOS DE AUTOR

© Copyright 2023
Samaria Poleth Hernandez Vásquez
Nelson Antonio Cabrera Bonilla

Todos los derechos son reservados.



FACULTAD DE POSTGRADO

PERFILAMIENTO DE CLIENTES EMPLEANDO HERRAMIENTAS DE ANALITICA DE NEGOCIOS PARA 5 MIPYMES EN LA CIUDAD DE SAN PEDRO SULA, HONDURAS

**Samaria Poleth Hernandez Vásquez
Nelson Antonio Cabrera Bonilla**

Resumen

El presente informe se centra en el perfilamiento de clientes y su aplicación en cinco MYPIMES de diversos rubros en la ciudad de San Pedro Sula. El objetivo principal es analizar mediante la técnica RFM de la herramienta de analítica KNIME el perfilamiento de clientes para comprender mejor las necesidades y preferencias de sus clientes, y cómo esta información se traduce en poder implementar estrategias de marketing más efectivas que ayuden a incrementar sus ingresos. El estudio comienza con una investigación sobre el perfilamiento de clientes y sus ventajas en la toma de decisiones empresariales, a través de ello se investiga sobre las herramientas de la analítica de datos existente en el mercado y se toma como una de ellas KNIME que es la herramienta para utilizar para el perfilamiento, Luego, se describe la metodología utilizada para recopilar datos de las cinco empresas participantes, que representan sectores como ventas en línea, distribución de productos alimenticios, distribuidora de baterías para autos y cafetería. A través del análisis de datos con la técnica de RFM se muestran los resultados de la investigación proporcionando información valiosa sobre cómo el perfilamiento de clientes puede adaptarse y beneficiar a las 5 MYPIMES clasificando sus clientes y a través de esto empezar a tomar decisiones estratégicas.

Palabras claves: (Perfilamiento, Segmentación, RFM, Analítica, KNIME)



GRADUATE SCHOOL

CUSTOMER PROFILING USING BUSINESS ANALYTICS TOOLS FOR 5 SMES IN THE CITY OF SAN PEDRO SULA, HONDURAS

**Samaria Poleth Hernandez Vásquez
Nelson Antonio Cabrera Bonilla**

Abstract

This report focuses on customer profiling and its application in five MSMEs of various areas in the city of San Pedro Sula. The main objective is to analyze through the RFM technique of the KNIME analytical tool the profiling of customers to better understand the needs and preferences of their customers, and how this information translates into being able to implement more effective marketing strategies that help increase their income. The study begins with an investigation on customer profiling and its advantages in business decision making, through it is investigated on the tools of data analytics existing in the market and is taken as one of them KNIME which is the tool to use for profiling, Then, It describes the methodology used to collect data from the five participating companies, representing sectors such as online sales, distribution of food products, distributor of batteries for cars and cafeteria. Through data analysis with the RFM technique, the results of the research are shown, providing valuable information on how customer profiling can adapt and benefit the 5 SMES by classifying their customers and through this start making strategic decisions.

Palabras claves: (Profiling, Segmentation, RFM, Analytics, KNIME)

DEDICATORIA

A nuestros esposos;

Por ser nuestro apoyo inquebrantable y fuente constante de inspiración a lo largo de este arduo camino de investigación. Su amor y paciencia han sido nuestro refugio en los momentos de desafío, y su inquebrantable fe en nosotros han sido nuestro motor. Gracias por estar a nuestro lado y ser el pilar de nuestra familia.

A nuestros hijos;

Cada día, ustedes nos recuerdan que la búsqueda del conocimiento es un legado que debemos transmitir de generación en generación.

Esperamos que esta tesis sea un testimonio de la importancia de perseguir nuestros sueños y nunca dejar de aprender.

Con todo nuestro amor,

Samaria Hernández

Nelson Cabrera

AGRADECIMIENTO

En el transcurso de este viaje académico que ha sido nuestra tesis, deseamos expresar nuestro más sincero agradecimiento a quienes han sido fundamentales en este proceso.

Primero y, ante todo, a Jehová, por su guía y fortaleza inquebrantable. Su amor y apoyo incondicional han sido nuestro refugio en los momentos de duda y fuente de esperanza en cada paso del camino.

A nuestros esposos y queridos hijos, quiénes han compartido este viaje. Sus sacrificios, comprensión y amor constante han sido nuestra mayor motivación y razón para esforzarnos aún más.

A nuestros valiosos compañeros de maestría, quienes compartieron risas, desafíos y éxitos a lo largo de este viaje. Sus aportaciones y amistad han enriquecido nuestra experiencia académica de manera inigualable.

A nuestro asesor metodológico KEVIN FUNES, cuya experiencia y orientación fueron fundamentales en la formulación y ejecución de esta investigación. Su paciencia y compromiso con nuestro crecimiento académico han sido invaluable.

Cada uno de ustedes ha desempeñado un papel esencial en nuestro camino hacia la culminación de esta tesis. Estamos profundamente agradecidos por su apoyo constante y su presencia en nuestra vida.

Con gratitud sincera,

Samaria Hernández

Nelson Cabrera

ÍNDICE DE CONTENIDO

DEDICATORIA	1
AGRADECIMIENTO	2
ÍNDICE DE CONTENIDO	3
CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN	6
1.1 INTRODUCCIÓN	6
1.2 ANTECEDENTES DEL PROBLEMA	6
1.3 DEFINICIÓN DEL PROBLEMA	7
1.4 OBJETIVOS DEL PROYECTO.....	8
1.4.1 OBJETIVO GENERAL	8
1.4.2 OBJETIVOS ESPECÍFICOS.....	8
1.5 JUSTIFICACIÓN.....	8
CAPÍTULO II. MARCO TEÓRICO	10
2.1 ANÁLISIS DE LA SITUACIÓN ACTUAL.	10
2.2 CONCEPTUALIZACIÓN	12
2.3 TEORÍAS DE SUSTENTO	14
2.3.1 BASES TEÓRICAS	14
2.4 MARCO LEGAL	15
CAPÍTULO III. METODOLOGÍA	17
3.1 CONGRUENCIA METODOLÓGICA.....	17
3.1.1 MATRIZ METODOLÓGICA.....	17
3.1.2 ESQUEMA DE VARIABLES DE ESTUDIO.....	18
3.2 ENFOQUE Y MÉTODOS	19
3.3 DISEÑO DE LA INVESTIGACIÓN	19
3.3.1 POBLACIÓN	19
3.3.2 MUESTRA	19
3.3.3 TÉCNICAS DE MUESTREO.....	19
3.4 TÉCNICAS, INSTRUMENTOS Y PROCEDIMIENTOS APLICADOS	20
3.4.1 TÉCNICAS	20
3.4.2 INSTRUMENTOS ELABORADOS	20
3.4.3 PROCEDIMIENTOS	20

3.5	FUENTES DE INFORMACIÓN.....	21
3.5.1	FUENTES PRIMARIAS	21
3.5.2	FUENTES SECUNDARIAS.....	21
CAPÍTULO IV. RESULTADOS Y ANÁLISIS		22
4.1	INFORME DE PROCESO DE RECOLECCIÓN DE DATOS	22
4.1.1	INTRODUCCIÓN AL ANÁLISIS	22
4.1.2	PERFILAMIENTO DE CLIENTES	22
4.1.3	RECOLECCIÓN DE DATOS	23
4.2	RESULTADOS Y ANÁLISIS DE LAS TÉCNICAS APLICADAS	24
4.2.1	PROCESO DE SEGMENTACION DE CLIENTES	24
4.2.2	CRITERIOS UTILIZADOS PARA CADA EMPRESA	25
4.3	ANÁLISIS INFERENCIAL Y MODELOS APLICADOS.....	32
4.3.1	FLUJO DE TRABAJO EN KNIME	33
4.3.2	GENERACIÓN DE RESULTADOS.....	34
4.3.3	COMPARACION DE RESULTADOS	38
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES.....		41
5.1	CONCLUSIONES	41
5.2	RECOMENDACIONES	42
CAPÍTULO VI. APLICABILIDAD.....		43
6.1	NOMBRE DE LA PROPUESTA	43
6.2	JUSTIFICACIÓN DE LA PROPUESTA	43
6.3	ALCANCE DE LA PROPUESTA	43
6.4	DESCRIPCIÓN Y DESARROLLO	43
6.4.1	PREPARACION DE DATOS.....	43
6.4.2	PROCESAMIENTO DE DATOS	43
6.4.3	VISUALIZACION Y DESPLIEGUE	44
6.5	MEDIDAS DE CONTROL	44
6.6	CRONOGRAMA DE IMPLEMENTACIÓN Y PRESUPUESTO	46
6.6.1	CRONOGRAMA DE IMPLEMENTACION.....	46
6.6.2	PRESUPUESTO	47
REFERENCIAS BIBLIOGRÁFICAS.....		48

ANEXOS	50
Anexo 1 Documentación flujo de trabajo de KNIME.....	50
Anexo 2 Fichas de observación para análisis de campos críticos.	54
Anexo 3 Notas de Aceptación uso de los datos.	55

CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1 INTRODUCCIÓN

A pesar de la tremenda irrupción, en la última década, que ha tenido el análisis de datos como instrumento de apoyo para la mejora continua de los negocios, no es esta una opción viable para los pequeños emprendimientos que les permita encontrar en la información que generan oportunidades de mejora y crecimiento.

La analítica de negocios es hoy por hoy una herramienta poderosísima que puede llegar a impulsar a estos pequeños negocios a aumentar la eficiencia de los procesos y los costos, impulsar la estrategia y el cambio, supervisar y mejorar el rendimiento financiero, por medio de una efectiva interpretación de los datos. Sin embargo, aunque las opciones de programas y herramientas digitales que le permite hacer esta labor son múltiples y variadas, en la mayoría de los casos son costosas y requieren de una inversión bastante onerosa por parte de dichas empresas.

Es así como surge la idea de demostrar que con el uso de herramientas de analítica de datos de libre acceso y con la mínima inversión, y conocimientos básicos en el uso de tecnología los dueños de estos pequeños negocios pueden exponer sus datos a ser analizado por algoritmos de perfilamiento de clientes y con la información resultante puedan tomar decisiones oportunas y relevantes para el crecimiento de sus negocios.

1.2 ANTECEDENTES DEL PROBLEMA

Aunque se carece de información actualizada, los informes disponibles nos dicen que las MIPYMES son la columna de la economía hondureña y en la actualidad son un eje importante en el país. Según las cifras de CID / Gallup (2000), en su Estudio de micro y pequeña empresa no agrícola en Honduras, se contabilizan 257,422 micros y pequeñas empresas que se caracterizan como unidades productivas de hasta cincuenta empleados, cuyo aporte porcentual es significativo relativo al Producto Interno Bruto (PIB), siguiendo en orden de importancia los establecimientos manufactureros (incluyendo la agroindustria) con un 25 %. En cuanto al tamaño de la planta, se advierte un claro predominio de los establecimientos de autoempleo o con un empleado, con el 74.4 % del total. (Amaya, O. A./ Macías, M.M., 2013).

Con la llegada de la pandemia del COVID-19, los retos para estos pequeños negocios por mantener su vigencia han sido enormes, y han surgido necesidades diferentes, sobre todo las

relacionadas con los aspectos digitales. Esfuerzos para promover la digitalización de los negocios, como El Plan de Digitalización MIPYME, impulsado por el Servicio Nacional de Emprendimiento y Pequeños Negocios (SENPRENDE) de Honduras, la Organización de Estados Americanos (OEA) y la empresa de ventas en línea Kolau Marketing, el cual otorgaría 15 mil licencias para que los emprendedores tengan una página digital y accedan a los beneficios del comercio electrónico (Navarro, D., 2021) , son mínimos con respecto al potencial y alcance que los pequeños negocios pudieran acceder a partir de mover sus negocios a plataformas digitales.

En este sentido en la actualidad, producto de los procesos de adaptabilidad por los que tuvieron que pasar los pequeños emprendimiento, durante y post pandemia del COVID-19, en el año 2020, en los que en su mayoría migraron sus negocios a plataformas virtuales y en consecuencia encontraron en estas un recurso de almacenamiento y generación de datos, acumulando un recurso de suma importancia para las empresa y del cual no se han obtenido mayores beneficios producto del análisis procesamiento con el propósito de obtener información que les permita tomar decisiones oportunas, como lo puede ser el que tengan un perfilamiento adecuado de su cartera de clientes.

1.3 DEFINICIÓN DEL PROBLEMA

Las micro, medianas y pequeñas empresas en San Pedro Sula cuentan con información producto de sus transacciones de negocios la cual no están aprovechando para mejorar la toma de decisiones empresariales, incluyendo la posibilidad de perfilar adecuadamente sus carteras de clientes.

¿Tienen potencial los datos generados por estos pequeños emprendimientos de generar información útil para toma de decisiones?, ¿Pueden ayudar estas decisiones tomadas a partir de esta información hacer crecer a estos pequeños negocios?, ¿Existen herramientas de analítica de negocios en el mercado de uso libre y que se adecuen a las necesidades actuales de estos pequeños negocios?, ¿Es posible la información de ventas generada por estos pequeños emprendimientos nos permita generar un perfilamiento adecuado de su cartera de clientes?.

1.4 OBJETIVOS DEL PROYECTO

1.4.1 OBJETIVO GENERAL

Desarrollar el perfilamiento de clientes con la ayuda de modelos de analítica de negocio para 5 MIPYMES hondureñas, con la finalidad de mejorar las estrategias de rentabilidad y marketing.

1.4.2 OBJETIVOS ESPECÍFICOS

- Realizar un análisis exhaustivo de los datos generados por las 5 MIPYME seleccionadas con el fin de identificar las diferentes variables clave de ventas que puedan proporcionar información útil para la toma de decisiones fundamentadas en los negocios.
- Emplear KNIME como la herramienta de analítica de datos que permita mediante el análisis de los datos de ventas generar perfiles para todos los clientes de las MIPYMES en los últimos 4 años.
- Generar recomendaciones y acciones en base a los perfiles generados para mejorar las estrategias de mercadeo.

1.5 JUSTIFICACIÓN

La incidencia del análisis de datos en la actualidad es fundamental para que las personas tomen decisiones correctas. En el pasado, las decisiones eran empíricas, espontáneas y/o reactivas, ya que no se tomaba en cuenta ningún tipo de análisis de registros a través de los datos o información disponible. En muchas ocasiones, estas decisiones podrían tener consecuencias o adversidades negativas debido a una mala elección. Hoy en día, todavía existe la posibilidad de tomar decisiones incorrectas, pero también existen medidas de contingencia anticipadas que se pueden llevar a cabo si llegara a ocurrir una mala decisión. Los mismos datos se encargan de mostrar qué hacer o no hacer en el futuro. Por esta razón, es tan importante contar con herramientas de análisis de datos que nos ayuden a tomar decisiones acertadas.

El presente estudio está enfocado en pequeños emprendedores que, en términos de tamaño de la empresa, determinado por el número de empleados, se puede decir que existe un dominio de las empresas de tipo autoempleo o con un empleado, representando un 74.4% del total. Ampliando

este segmento con las microempresas (hasta 10 empleados), se trata de un total de 98%. Por lo tanto, la pequeña empresa con 11 hasta 50 trabajadores tiene una presencia en términos numéricos muy limitada con apenas un 2%. (CID/Gallup, 2000). A partir de algoritmos de clasificación aplicados a sus datos de ventas con herramientas de analítica de datos de uso libre, estos pequeños emprendedores tendrían la oportunidad de realizar tomas de decisiones oportunas y acciones necesarias para abordar cada segmento de sus clientes de la manera que mejor corresponda a sus estrategias de mercado.

Estas herramientas pueden proporcionar datos importantes que llevan a los dueños de negocios a tomar decisiones determinantes en función del comportamiento de sus clientes en cuanto a recencia, regularidad y montos monetarios les consumen y de esta forma tomar acciones que les permitan hacer crecer sus negocios.

CAPÍTULO II. MARCO TEÓRICO

2.1 ANÁLISIS DE LA SITUACIÓN ACTUAL.

La República de Honduras está localizada en el corazón de Centroamérica, siendo sus principales actividades económicas las relacionadas con el sector agropecuario, además del comercio, manufacturas, finanzas, y servicios públicos. Una de las estrategias con las cuales Honduras espera impulsar el emprendimiento a nivel nacional es la generación de un ecosistema emprendedor favorable que conceda a toda la población inclusive a la más desfavorecida mayores oportunidades de desarrollo y crecimiento.

De cualquier forma, es imperativo abordar este tema con énfasis especial, ya que Honduras es el tercer país con la más baja calificación en el coeficiente de GINI entre los países de Latinoamérica y el Caribe después de Haití (PNUD, 2012).

En cuanto a la actualidad del emprendimiento en Honduras, cabe mencionar que el país cuenta con grandes retos que se deben abordar con estrategias contundentes en diferentes temas que generan amenazas al ecosistema emprendedor. Éstas residen en factores como la inseguridad, el crimen, la violencia y la burocracia, entre otros. Situación que con seguridad desincentiva la creación de empresa y la inversión en negocios físicos y optan por las ventas de productos y servicios en línea mayormente; a raíz de la pandemia de COVID19 el auge del emprendimiento hondureño fue visto a plenitud en el año 2020 donde la utilización de las redes sociales fue la mejor herramienta para generar ingresos a través de las ventas en línea.

Datos notables de emprendimiento en Honduras se presentan a continuación para contextualizar el país a nivel internacional a partir de informes reconocidos: En el informe Doing Business el subíndice mejor calificado es la facilidad para conseguir crédito ubicándolo en la posición 13 de 189 economías analizadas, una herramienta fundamental en el momento de convertir una idea de negocios en una realidad.

Honduras se encuentra en una posición ventajosa en el otorgamiento de créditos en comparación con el promedio regional, en esta medida se debe asegurar a los empresarios las condiciones adecuadas para que los procedimientos, costos y tiempo necesarios para importación y exportación no se conviertan en un obstáculo.

De acuerdo con el Índice de Emprendimiento y Desarrollo Global (GEDI), las fortalezas en las actitudes emprendedoras de los hondureños residen en la aceptación del riesgo, la percepción de habilidades, los competidores, el reconocimiento de oportunidad y los nuevos productos.

En el índice GEDI Honduras se ubica en la posición 101 para el reporte 2014. Como se puede observar en la ilustración siguiente se puede ver que el país se destaca en las variables individuales del subíndice Actitudes Emprendedoras, siendo los factores mejor calificados para el país la aceptación del riesgo, la percepción de habilidades, los competidores, el reconocimiento de oportunidad y los nuevos productos. Mientras tanto las falencias que más sobresalen en Honduras y le dan una posición relativamente baja en comparación con los otros 120 países calificados son según este índice: la inversión por parte del gobierno en investigación y desarrollo, la innovación de procesos, la inversión informal, el capital de riesgo y el alto crecimiento de los emprendimientos.

Los esfuerzos que se han identificado en Honduras para el fomento de la cultura emprendedora son muy pocos, y se caracterizan por ser aisladas, no coordinados entre sí y desarrollados por iniciativas privadas, muy incipientemente en algunas universidades, y en ONGs, sin participación alguna de las asociaciones del sector privado, como por ejemplo, el Consejo Hondureño de la Empresa Privada (COHEP), Asociación de Medianas y Pequeñas Industrias de Honduras (ANMPIH) Cámaras de Comercio y de instituciones del Estado.

Esta misma afirmación es válida para el área educativa, cultural y financiera, donde se encuentran acciones institucionales de educación, formación, capacitación y financiamiento que están concentradas en acciones exclusivas para micro, pequeñas y medianas empresas (MIPYME) sin considerar acciones específicas y especializadas para emprendimiento. Destacando además la ausencia de financiamiento no tradicional, el tema de género y lo relacionado a migrantes retornados.

La Red de Emprendimiento de Honduras está conformada por Instituciones Públicas-Privadas- Academia y Cooperantes que desarrollan actividades de apoyo para el fomento del emprendimiento en el país, denominado como un espacio de consenso, para socializar y aportar ideas hacia la construcción de la Estrategia País. Este trabajo conjunto ha permitido desarrollar una agenda de prioridades, celebrando una serie de talleres con la participación de directores de Programas, Municipalidades, Organizaciones No Gubernamentales, Asesores nacionales de

cooperación externa, asesores técnicos y Organizaciones Gremiales.

A partir de los resultados obtenidos en la identificación de los programas de emprendimiento, se han invitado a formar parte de este espacio actores clave considerando criterios como: Liderazgo Institucional en el fomento de programas de emprendimiento a nivel nacional y territorial, rol en el mapa de emprendimiento nacional, Capacidad Instalada y recursos disponibles.

2.2 CONCEPTUALIZACIÓN

- **Perfilamiento de clientes:** Es el proceso de analizar y segmentar a los clientes en grupos homogéneos con características y comportamientos similares, con el objetivo de comprender mejor sus necesidades, preferencias y comportamientos de compra. (Hair, J. F., Celsi, M., Ortinau, D. J., & Bush, R. P. (2019). Investigación de mercados. McGraw-Hill Education).
- **Segmentación de mercado:** Es el proceso de dividir el mercado en grupos más pequeños y homogéneos de consumidores con características similares, como demográficas, geográficas, psicográficas o comportamentales. (Kotler, P., & Armstrong, G. (2017). Principios de marketing. Pearson).
- **Análisis de comportamiento del cliente:** Es el estudio y evaluación de los patrones y tendencias en los comportamientos de compra, interacciones y preferencias de los clientes a lo largo del tiempo, con el fin de identificar oportunidades y personalizar estrategias de marketing. (Wedel, M., & Kannan, P. K. (2016). Análisis de marketing para entornos ricos en datos. Journal of Marketing, 80(6), 97-121).
- **Herramientas de análisis de datos:** Son software o plataformas que permiten recopilar, organizar, visualizar y analizar grandes volúmenes de datos con el objetivo de obtener información valiosa para la toma de decisiones. (Provost, F., & Fawcett, T. (2013). Ciencia de datos para empresas: lo que necesita saber sobre la minería de datos y el pensamiento analítico de datos. O'Reilly Media).
- **Minería de datos:** Es una técnica que consiste en extraer información valiosa y útil a partir

de grandes conjuntos de datos, utilizando técnicas estadísticas y de aprendizaje automático. (Han, J., Kamber, M., & Pei, J. (2011). Minería de datos: conceptos y técnicas. Elsevier).

- **Aprendizaje automático:** Es una rama de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender a partir de los datos, sin necesidad de ser programadas explícitamente. (Alpaydin, E. (2010). Introducción al aprendizaje automático. MIT press).
- **Modelo predictivo:** Es un modelo matemático o estadístico que se utiliza para predecir el comportamiento futuro de un fenómeno o evento, a partir de los datos históricos. (Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introducción al análisis de regresión lineal. John Wiley & Sons).
- **Visualización de datos:** Es la representación gráfica de los datos con el objetivo de hacerlos más comprensibles y accesibles para el usuario final. (Few, S. (2012). Muéstrame los números: Diseñando tablas y gráficos para iluminar. Analytics Press).
- **El Análisis RFM:** Es una técnica de segmentación de clientes que se basa en tres factores: Recencia, Frecuencia y Monetización. Estos factores se utilizan para evaluar el comportamiento de compra de los clientes y segmentarlos en grupos con características similares. El análisis RFM permite a las empresas identificar a los clientes más valiosos y establecer estrategias de marketing personalizadas para cada grupo. (Kotler, P., & Armstrong, G. (2017). Fundamentos de marketing. Pearson Educación).
- **La Recencia:** en términos de RFM (Recencia, Frecuencia, Monetización) se refiere al tiempo transcurrido desde la última interacción del cliente con la empresa. La recencia es un factor importante en el análisis de la cartera de clientes, ya que los clientes más recientes tienden a ser más valiosos para la empresa y a tener un mayor potencial de compra en el futuro. El análisis de la recencia permite segmentar a los clientes según su comportamiento de compra y establecer estrategias de marketing personalizadas para cada segmento. (Kotler, P., & Armstrong, G. (2017). Fundamentos de marketing. Pearson Educación).
- **La Frecuencia:** En términos de RFM (Recencia, Frecuencia, Monetización) se refiere a la cantidad de veces que un cliente ha interactuado con la empresa durante un período de tiempo determinado. La frecuencia es un factor importante en el análisis de la cartera de

clientes, ya que los clientes que interactúan con la empresa con mayor frecuencia tienden a ser más valiosos y a tener un mayor potencial de compra en el futuro. El análisis de la frecuencia permite segmentar a los clientes según su comportamiento de compra y establecer estrategias de marketing personalizadas para cada segmento. (Kotler, P., & Armstrong, G. (2017). Fundamentos de marketing. Pearson Educación).

- **La Monetización:** En términos de RFM (Recencia, Frecuencia, Monetización) se refiere al valor monetario total que un cliente ha gastado en la empresa durante un período de tiempo determinado. La monetización es un factor importante en el análisis de la cartera de clientes, ya que los clientes que han gastado más dinero en la empresa tienden a ser más valiosos y a tener un mayor potencial de compra en el futuro. El análisis de la monetización permite segmentar a los clientes según su comportamiento de compra y establecer estrategias de marketing personalizadas para cada segmento. (Kotler, P., & Armstrong, G. (2017). Fundamentos de marketing. Pearson Educación).

2.3 TEORÍAS DE SUSTENTO

2.3.1 BASES TEÓRICAS

El perfilamiento de clientes es un proceso fundamental en el campo del marketing que permite a las empresas comprender mejor las necesidades y preferencias de sus clientes para ofrecer productos y servicios más personalizados. Para lograr esto, es necesario contar con herramientas y técnicas de análisis de datos que permitan recopilar, procesar y analizar grandes volúmenes de información. En esta tesis, se explorarán algunas teorías fundamentales que respaldan el uso de herramientas de análisis de datos para el perfilamiento de clientes en empresas PYMES.

La teoría de la segmentación de mercado es un enfoque clave para el perfilamiento de clientes, ya que permite dividir el mercado en grupos más pequeños y homogéneos de consumidores con características similares. Al segmentar el mercado, las empresas pueden identificar las necesidades y preferencias únicas de cada grupo, lo que les permite adaptar sus estrategias de marketing y ofrecer productos o servicios más personalizados. La teoría del comportamiento del consumidor se enfoca en comprender los patrones y tendencias en los

comportamientos de compra, interacciones y preferencias de los clientes a lo largo del tiempo. Al analizar el comportamiento del consumidor, las empresas pueden adaptar sus estrategias de marketing para satisfacer esas necesidades y preferencias específicas de cada grupo de clientes. La teoría del aprendizaje automático es fundamental para el análisis de datos y el perfilamiento de clientes, ya que permite identificar patrones y tendencias en los datos que no son evidentes a simple vista.

En conjunto, estas teorías respaldan la importancia del uso de herramientas de análisis de datos para el perfilamiento de clientes en empresas MIPYMES. En esta tesis, se explorará cómo estas teorías pueden aplicarse en la práctica para mejorar la comprensión del cliente y poder tomar las decisiones necesarias en función del perfil en particular de cada uno.

2.4 MARCO LEGAL

En Honduras, el marco legal relacionado con el uso de herramientas de análisis de datos se encuentra principalmente en las siguientes leyes y regulaciones:

- **Ley de Protección de Datos Personales:** La Ley de Protección de Datos Personales establece los principios y requisitos para el tratamiento de datos personales en Honduras. Esta ley garantiza la protección de la privacidad y los derechos de las personas en relación con el uso, recolección, almacenamiento y transmisión de sus datos personales.

Desde abril de 2018, el Congreso Nacional de la República de Honduras ha mantenido en suspenso el tercer y último debate para la aprobación del proyecto de Ley de Protección de Datos Personales. Este proyecto es innovador en Honduras, ya que actualmente no existe una legislación que proteja los datos personales. Sin embargo, se han identificado ciertos conceptos que podrían prestarse para malinterpretaciones o carecen de un buen desarrollo conceptual, lo que podría abrir la puerta a posibles arbitrariedades.

A falta de una legislación especial, los datos personales en Honduras cuentan con al menos una protección reconocida en la Ley del Instituto de Acceso a la Información Pública, Decreto Legislativo No. 170-2006. En los artículos 24 al 26 de esta ley se reconoce el derecho al Hábeas Data, la protección de los datos personales y se presenta la figura del Comisionado Nacional de Derechos Humanos como una oficina facultada para incoar acciones para la protección de datos personales. Además, se establece una prohibición en

la cual ninguna persona puede solicitar datos personales que puedan generar algún tipo de discriminación o poner en riesgo los derechos morales y patrimoniales de ese individuo.

- **Ley de Comercio Electrónico:** La Ley de Comercio Electrónico, publicado en el diario Oficial La Gaceta el 27 de abril de 2015 bajo el Decreto No. 149-2014, regula las transacciones comerciales realizadas a través de medios electrónicos en Honduras. Esta ley puede ser relevante para el uso de herramientas de análisis de datos, ya que aborda aspectos como la protección de datos, la seguridad de la información y la privacidad en el entorno digital.

CAPÍTULO III. METODOLOGÍA

3.1 CONGRUENCIA METODOLÓGICA

3.1.1 MATRIZ METODOLÓGICA

Título: “PERFILAMIENTO DE CLIENTES EMPLEANDO HERRAMIENTAS DE ANALITICA DE NEGOCIOS PARA 5 EMPRESAS MIPYMES EN SAN PEDRO SULA”					
Formulación del Problema	Objetivo General	Objetivos específicos	Preguntas de Investigación	Variable Dependiente	Variables Independientes
<p>Las micro, medianas y pequeñas empresas en San Pedro Sula cuentan con información producto de sus transacciones de negocios la cual no están aprovechando para mejorar la toma de decisiones empresariales, incluyendo la posibilidad de perfilar adecuadamente sus carteras de clientes.</p>	<p>Desarrollar el perfilamiento de clientes con la ayuda de modelos de analítica de negocio para 5 MIPYMES hondureñas, con la finalidad de mejorar las estrategias de rentabilidad y marketing.</p>	<p>Realizar un análisis exhaustivo de los datos generados por las 5 MIPYME seleccionadas con el fin de identificar las diferentes variables clave de ventas que puedan proporcionar información útil para la toma de decisiones fundamentadas en los negocios.</p>	<p>¿Tienen potencial los datos generados por estos pequeños emprendimientos de generar información útil para toma de decisiones?</p>	<p>Perfilamiento de Clientes</p>	<p>Análisis RFM</p>
		<p>Emplear KNIME como la herramienta de analítica de datos que permita mediante el análisis de los datos de ventas generar perfiles para todos los clientes de las MIPYMES en los últimos 4 años.</p>	<p>¿Pueden ayudar estas decisiones tomadas a partir de esta información hacer crecer a estos pequeños negocios?</p>		<p>Recencia</p>
		<p>Generar recomendaciones y acciones en base a los perfiles generados para mejorar las estrategias de mercadeo.</p>	<p>¿Existen herramientas de analítica de negocios en el mercado de uso libre y que se adecuen a las necesidades actuales de estos pequeños negocios?</p>		<p>Frecuencia</p>
		<p>Generar recomendaciones y acciones en base a los perfiles generados para mejorar las estrategias de mercadeo.</p>	<p>¿Es posible la información de ventas generada por estos pequeños emprendimientos nos permita generar un perfilamiento adecuado de su cartera de clientes?</p>		<p>Monetización</p>

Tabla 1 - Matriz Metodológica

Fuente: Inspiración Propia.

3.1.2 ESQUEMA DE VARIABLES DE ESTUDIO

Variable	Definición		Dimensiones	Indicador
	Conceptual	Operacional		
Análisis RFM	El análisis RFM (Recency, Frequency, Monetary) es una técnica de segmentación de clientes que se utiliza para clasificar a los clientes en función de su comportamiento de compra. Esta técnica se basa en tres variables clave: la recencia (la última vez que un cliente realizó una compra), la frecuencia (con qué frecuencia realiza compras un cliente) y el valor monetario (cuánto dinero gasta un cliente en promedio en cada compra). (Peppers, D., & Rogers, M. (2004). Gestión de las relaciones con los clientes: un marco estratégico. John Wiley & Sons.)	Técnica de perfilamiento utilizada en un flujo de trabajo de KNIME con el propósito de segmentar los clientes en base a la data de ventas proporcionada por las empresas.	Recencia	Fecha Compra
				Fecha Análisis
			Frecuencia	Transacciones
				Productos
			Monetización	Precios
				Cantidades

Tabla 2 - Esquema de variables de estudio

Fuente: Inspiración Propia.

3.2 ENFOQUE Y MÉTODOS

El presente estudio tiene un enfoque **DESCRIPTIVO** con un método **CUANTITATIVO** con un método de muestreo **NO PROBABILÍSTICO**.

3.3 DISEÑO DE LA INVESTIGACIÓN

3.3.1 POBLACIÓN

Población, Es el conjunto de personas u objetos de los que se desea conocer algo en una investigación. "El universo o población puede estar constituido por personas, animales, registros médicos, los nacimientos, las muestras de laboratorio, los accidentes viales entre otros". (PINEDA et al 1994:108). En el caso de esta investigación la población objeto de estudio serán los clientes que han realizado transacciones de ventas, datos proporcionados por las diferentes MYPIMES con una antigüedad que van desde 5 meses a 5 años conteniendo en ella la información necesaria para el testeo de escenarios utilizando el algoritmo de RFM según el modelado de datos ya identificado utilizando la herramienta de KNIME.

3.3.2 MUESTRA

La Muestra es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que se le llama población, Las muestras pueden ser probabilísticas o no probabilísticas. (H. Sampieri, citado por Balestrini 2001 Pág. 141). Existen varias clases de muestras no probabilísticas, para Hernández, Fernández y Baptista en su Metodología de la Investigación (1991) se hallan clasificadas de la manera siguiente: Muestras de sujetos voluntarios, Muestras de expertos, Sujetos tipos o Stakeholders y Muestras por cuota. La muestra de la población objeto de estudio en esta investigación será no probabilística ya que se tomará seleccionando 5 tipos de rubros donde a las MYPIMES se les pedirá compartan su histórico de ventas de los últimos 5 meses a 5 años respetando los campos críticos necesarios para los testeos de los diferentes escenarios de la investigación.

3.3.3 TÉCNICAS DE MUESTREO

Muestreo, es el método utilizado para seleccionar a los componentes de la muestra del total de la población. "Consiste en un conjunto de reglas, procedimientos y criterios mediante los cuales

se selecciona un conjunto de elementos de una población que representan lo que sucede en toda esa población". (MATA et al, 1997:19). En esta investigación se utilizará una técnica de muestreo NO PROBABILÍSTICA ya que se ha decidido seleccionar 5 rubros de diferentes empresas, únicamente para que puedan brindar el objeto de estudio que será preciso ya que cada MYPIME compartirá la información con los campos que se le pidan brindar.

3.4 TÉCNICAS, INSTRUMENTOS Y PROCEDIMIENTOS APLICADOS

3.4.1 TÉCNICAS

Las técnicas utilizadas para esta investigación es la Observación Cuantitativa, ya que se analizará a través del algoritmo de RFM los datos históricos de ventas de los últimos 5 meses a 5 años recopilados de los clientes seleccionados.

3.4.2 INSTRUMENTOS ELABORADOS

Ficha de Observación, se elaborará una plantilla de archivo con los campos críticos necesarios con los que cada cliente brindará la información de su histórico de ventas de los últimos 5 meses a 5 años. Este es el que permitirá realizar el análisis de algoritmos de perfilamiento de clientes y del cual se obtendrán los datos que se le mostrarán al cliente a través de un tablero de mando para que pueda interactuar con las diferentes recomendaciones y pueda el mismo cliente diseñar sus estrategias de mercado para poder llegar a esos clientes potenciales. Puede verse en el Anexo 2.

3.4.3 PROCEDIMIENTOS

Se utilizará el instrumento de una ficha de observación que será la plantilla de los campos críticos correlacionados necesarios para realizar los testeos de perfilamiento a través del RFM, los necesarios para poder determinar las variables que involucran el análisis como ser; Recencia, la cual requiere fecha de ultima compra, Frecuencia, que requiere un conteo de las compras que realiza el cliente y Monetización que requiere de una suma de valores de compra obtenida de las cantidades y precios que detallen las facturas de venta.

3.5 FUENTES DE INFORMACIÓN

3.5.1 FUENTES PRIMARIAS

Se obtendrá un reporte histórico de ventas de los últimos 5 años donde los clientes brindaran los datos de sus ventas y datos demográficos de sus clientes según la plantilla que contiene los campos críticos a analizar

3.5.2 FUENTES SECUNDARIAS

No aplican para el presente trabajo de investigación.

CAPÍTULO IV. RESULTADOS Y ANÁLISIS

4.1 INFORME DE PROCESO DE RECOLECCIÓN DE DATOS

4.1.1 INTRODUCCIÓN AL ANÁLISIS

Dado que el objetivo principal de esta tesis es establecer un perfilamiento de clientes a través del análisis de datos; que ayude a las 5 MIPYMES objetos de estudio en San Pedro Sula a aumentar su volumen de ventas y tomar decisiones estratégicas de mercado, hemos procedido a recopilación de los datos de ventas de al menos 5 clientes dedicados a diferentes rubros que han decidido participar en esta investigación; quienes han proporcionado de forma directa la data solicitada con el formato y estructura especificados.

Estos datos recopilados contienen información de las ventas realizadas por 5 MIPYMES pertenecientes a los rubros de: Venta de tiendas por departamento, distribución de alimentos, supermercados, ventas y distribución de batería para autos y cafetería todos ubicados en San Pedro Sula con contenido que datan de 6 meses hasta 5 años de histórico de ventas.

4.1.2 PERFILAMIENTO DE CLIENTES

La herramienta de analítica de datos que se ha decidido utilizar en esta investigación para llevar a cabo el perfilamiento de las MYPIMES que participaran del proceso es KNIME.

KNIME (Konstanz Information Miner) es una plataforma de código abierto que se utiliza para el análisis de datos y la minería de datos. Ofrece una interfaz gráfica que permite a los usuarios crear flujos de trabajo de análisis de datos sin necesidad de escribir código.

Características Clave:

- **Flujos de Trabajo Visuales:** KNIME permite a los usuarios arrastrar y soltar componentes para crear flujos de trabajo visuales, lo que facilita la construcción y personalización de análisis de datos.
- **Amplia Integración:** Ofrece una amplia gama de extensiones y conectores para integrar con bases de datos, herramientas de análisis, y más.
- **Procesamiento de Datos:** Permite la manipulación y transformación de datos, incluyendo la limpieza, filtrado, agregación y unión de conjuntos de datos.

- **Modelado Predictivo:** Facilita la creación de modelos de aprendizaje automático y análisis predictivo.
- **Visualización:** Proporciona herramientas para la visualización de datos y resultados.

Aplicación en el Perfilamiento de Clientes:

KNIME es útil en el perfilamiento de clientes porque permite la preparación de datos, la segmentación de clientes, la creación de modelos predictivos para identificar comportamientos de compra y preferencias, y la generación de informes visuales para la toma de decisiones. El flujo de trabajo aplicado desde KNIME utilizará un algoritmo de clasificación basado en la metodología RFM.

4.1.3 RECOLECCIÓN DE DATOS

Como parte de la recopilación de datos hemos tenido el acceso a los datos de venta de 5 diferentes empresas con diferentes rubros, los cuales han proporcionado los datos detallados de venta que datan de los 5 meses hasta los 5 años de operaciones que tienen algunos de ellos; estos datos contienen en su estructura los campos requeridos para poder desarrollar un algoritmo de perfilamiento RFM. La tabla 3 muestra la estructura de estos campos.

NOMBRE DE CAMPO	TIPO	TAMAÑO	DESCRIPCION
NUMERO_FACTURA	Int	-	Número de Factura Generada
PRODUCTO	Char	25	Código del ITEM Facturado
DESCRIPCION_PROD	Char	60	Nombre que describe el Producto
CANTIDAD_VENDIDA	Numeric	18,8	Cantidad del ITEM Facturado
FECHA_DE_FACTURA	Datetime	-	Fecha de la facturación
PRECIO_AFFECTADO	Numeric	18,8	Precio del ITEM Facturado
CODIGO_DE_CLIENTE	Int	-	Código del Cliente de la Factura
NOMBRE_CLI_A_FACTU	Char	60	Nombre del Cliente de la Factura

Tabla 3 – Estructura de campos para recolección de datos.

Fuente: Inspiración Propia.

4.2 RESULTADOS Y ANÁLISIS DE LAS TÉCNICAS APLICADAS

4.2.1 PROCESO DE SEGMENTACION DE CLIENTES

El perfilamiento de clientes mediante la metodología RFM es una técnica ampliamente utilizada en marketing y análisis de negocios para clasificar y segmentar a los clientes en función de su historial de compras y su comportamiento. RFM es un acrónimo que representa tres dimensiones clave:

- **Recencia (Recency):** Esta dimensión se refiere a la "recencia" de la última compra de un cliente. Cuanto más reciente haya sido la compra, mayor será la puntuación asignada. Por ejemplo, un cliente que realizó una compra en los últimos 30 días recibiría una puntuación alta en la dimensión de recencia.
- **Frecuencia (Frequency):** La dimensión de "frecuencia" mide con qué frecuencia un cliente realiza compras. Los clientes que compran con mayor frecuencia obtienen una puntuación más alta en esta dimensión. Por ejemplo, un cliente que compra cada semana recibiría una puntuación alta en frecuencia.
- **Monetización (Monetary):** La dimensión de "monetización" se relaciona con cuánto dinero gasta un cliente en sus compras. Los clientes que gastan más en general reciben una puntuación más alta en esta dimensión. Por ejemplo, un cliente que realiza compras grandes y costosas obtendría una puntuación alta en monetización.

Para llevar a cabo el perfilamiento de clientes mediante RFM, se asignan puntuaciones a cada cliente en estas tres dimensiones. Por lo general, se utiliza una escala numérica, como 1-5 o 1-10, donde 1 representa la menor recencia, frecuencia o gasto, y el número más alto representa la mayor recencia, frecuencia o gasto.

Luego, se pueden combinar estas puntuaciones para crear segmentos de clientes. Por ejemplo, un cliente con puntuaciones altas en recencia, frecuencia y monetización se consideraría un cliente "VIP" o de alto valor. En cambio, un cliente con puntuaciones bajas en todas las dimensiones podría considerarse como un cliente "inactivo" o de bajo valor.

El perfilamiento de clientes RFM es valioso porque permite identificar grupos de

clientes con características similares, lo que facilita la personalización de estrategias de marketing, la retención de clientes y la toma de decisiones basadas en dato.

4.2.2 CRITERIOS UTILIZADOS PARA CADA EMPRESA

Los criterios que fueron utilizados para todas y cada una de la empresa de la muestra son exactamente los mismos, tomando la data y adecuarla a la plantilla y estructura que lee Knime al cargar la data para su procesamiento, en formato Excel (Xlsx), y procesada como a continuación explicamos con el objetivo de obtener cada elemento que componen al RFM como método de clasificación;

- **Recencia:** Basados en los días que han pasado desde que se hizo la compra, ocupamos el campo **FECHA_DE_FACTURA** y obtuvimos la diferencia a la fecha que se ejecute el flujo de trabajo en KNIME.
- **Frecuencia:** Usamos agrupamiento por cliente para contar la cantidad de compras que se realiza por cliente, usamos el campo **NUMERO_DE_FACTURA**.
- **Monetización:** Sumamos el total de las ventas, obteniendo el monto de venta al multiplica el **PRECIO_AFECTADO** por la **CANTIDAD_VENDIDA** de la data proporcionada por las empresas.

En este sentido, para el flujo de trabajo usado en KNIME, se ha aplicado el nodo **Auto-Binner**, este nodo permite agrupar datos numéricos en intervalos, llamados Bins. utiliza los cuantiles para determinar los rangos de valores que caen en cada Bin. Esto permite discretizar una variable numérica en intervalos categóricos, lo cual puede facilitar su análisis y visualización.

Los cuantiles son segmentaciones que dividen los datos en partes iguales. Por ejemplo, el percentil 25 divide los datos en dos partes, donde el 25% de los valores cae por debajo de ese percentil y el 75% restante cae por encima. Para el caso de nuestra investigación asignamos

Para nuestro flujo de trabajo aplicado en Knime, para propósitos de definición de los Bins, se decidió configurar el nodo Auto-Binner con cuantiles de ejemplo, especificando los siguientes valores: 0.0, 0.25, 0.75, 1.0, los valores de cuantiles de ejemplo se interpretan de

la siguiente manera:

- El cuantil 0.0 representa el valor mínimo de la distribución. Es el valor por debajo del cual se encuentra el 0% de los datos.
- El cuantil 0.25 representa el primer cuartil de la distribución. Es el valor que divide a la distribución en cuatro partes iguales y por debajo del cual se encuentra el 25% de los datos.
- El cuantil 0.75 representa el tercer cuartil de la distribución. Es el valor que divide a la distribución en cuatro partes iguales y por debajo del cual se encuentra el 75% de

los datos.

- El cuantil 1.0 representa el valor máximo de la distribución. Es el valor por debajo del cual se encuentra el 100% de los datos.

Ejecutado el nodo de **Auto-Binner** los resultados para cada una de las datas de las diferentes empresas y los rangos calculados por los cuantiles, los mostramos en las siguientes tablas;

REGENCY [Binned]	FRECUENCY [Binned]	MONETARY [Binned]	RANGE REGENCY	RANGE FRECUENCY	RANGE MONETARY
Bin 1	Bin 1	Bin 1	[13,53]	[1,5]	[34.5,2,007.42]
Bin 2	Bin 1	Bin 1	(53,1,012]	[1,5]	[34.5,2,007.42]
Bin 3	Bin 1	Bin 1	(1,012,1,715]	[1,5]	[34.5,2,007.42]
Bin 1	Bin 1	Bin 2	[13,53]	[1,5]	(2,007.42,58,464.463]
Bin 2	Bin 1	Bin 2	(53,1,012]	[1,5]	(2,007.42,58,464.463]
Bin 3	Bin 1	Bin 2	(1,012,1,715]	[1,5]	(2,007.42,58,464.463]
Bin 2	Bin 1	Bin 3	(53,1,012]	[1,5]	(58,464.463,359,851,879.745]
Bin 3	Bin 1	Bin 3	(1,012,1,715]	[1,5]	(58,464.463,359,851,879.745]
Bin 1	Bin 2	Bin 1	[13,53]	(5,72]	[34.5,2,007.42]
Bin 2	Bin 2	Bin 1	(53,1,012]	(5,72]	[34.5,2,007.42]
Bin 3	Bin 2	Bin 1	(1,012,1,715]	(5,72]	[34.5,2,007.42]
Bin 1	Bin 2	Bin 2	[13,53]	(5,72]	(2,007.42,58,464.463]
Bin 2	Bin 2	Bin 2	(53,1,012]	(5,72]	(2,007.42,58,464.463]
Bin 3	Bin 2	Bin 2	(1,012,1,715]	(5,72]	(2,007.42,58,464.463]
Bin 1	Bin 2	Bin 3	[13,53]	(5,72]	(58,464.463,359,851,879.745]
Bin 2	Bin 2	Bin 3	(53,1,012]	(5,72]	(58,464.463,359,851,879.745]
Bin 3	Bin 2	Bin 3	(1,012,1,715]	(5,72]	(58,464.463,359,851,879.745]
Bin 1	Bin 3	Bin 2	[13,53]	(72,7,331]	(2,007.42,58,464.463]
Bin 2	Bin 3	Bin 2	(53,1,012]	(72,7,331]	(2,007.42,58,464.463]

Bin 3	Bin 3	Bin 2	(1,012,1,715]	(72,7,331]	(2,007.42,58,464.463]
Bin 1	Bin 3	Bin 3	[13,53]	(72,7,331]	(58,464.463,359,851,879.745]
Bin 2	Bin 3	Bin 3	(53,1,012]	(72,7,331]	(58,464.463,359,851,879.745]
Bin 3	Bin 3	Bin 3	(1,012,1,715]	(72,7,331]	(58,464.463,359,851,879.745]

Tabla 4 – Resultados Auto-Binner para la data de la MIPYME Distribuidora de alimentos.

Fuente: Inspiración Propia.

RECENCY [Binned]	FRECUENCY [Binned]	MONETARY [Binned]	RANGE RECENCY	RANGE FRECUENCY	RANGE MONETARY
Bin 1	Bin 1	Bin 1	[13,18]	[2,8]	[1,900,4,423.75]
Bin 2	Bin 1	Bin 1	(18,33.25]	[2,8]	[1,900,4,423.75]
Bin 3	Bin 1	Bin 1	(33.25,104]	[2,8]	[1,900,4,423.75]
Bin 1	Bin 1	Bin 2	[13,18]	[2,8]	(4,423.75,7,860]
Bin 2	Bin 1	Bin 2	(18,33.25]	[2,8]	(4,423.75,7,860]
Bin 3	Bin 1	Bin 2	(33.25,104]	[2,8]	(4,423.75,7,860]
Bin 2	Bin 2	Bin 1	(18,33.25]	(8,12]	[1,900,4,423.75]
Bin 1	Bin 2	Bin 2	[13,18]	(8,12]	(4,423.75,7,860]
Bin 2	Bin 2	Bin 2	(18,33.25]	(8,12]	(4,423.75,7,860]
Bin 3	Bin 2	Bin 2	(33.25,104]	(8,12]	(4,423.75,7,860]
Bin 1	Bin 2	Bin 3	[13,18]	(8,12]	(7,860,11,200]
Bin 2	Bin 2	Bin 3	(18,33.25]	(8,12]	(7,860,11,200]
Bin 3	Bin 2	Bin 3	(33.25,104]	(8,12]	(7,860,11,200]

Bin 1	Bin 3	Bin 2	[13,18]	(12,18]	(4,423.75,7,860]
Bin 2	Bin 3	Bin 2	(18,33.25]	(12,18]	(4,423.75,7,860]
Bin 1	Bin 3	Bin 3	[13,18]	(12,18]	(7,860,11,200]
Bin 2	Bin 3	Bin 3	(18,33.25]	(12,18]	(7,860,11,200]
Bin 3	Bin 3	Bin 3	(33.25,104]	(12,18]	(7,860,11,200]

Tabla 5 – Resultados Auto-Binner para la data de la MIPYME Tienda por Departamentos.

Fuente: Inspiración Propia.

RECENCY [Binned]	FRECUENCY [Binned]	MONETARY [Binned]	RANGO RECENCY	RANGO FRECUENCY	RANGO MONETARY
Bin 1	Bin 1	Bin 1	[13,284]	[1,1]	[5,1,850]
Bin 2	Bin 1	Bin 1	(284,1,075.25]	[1,1]	[5,1,850]
Bin 3	Bin 1	Bin 1	(1,075.25,1,657]	[1,1]	[5,1,850]
Bin 1	Bin 1	Bin 2	[13,284]	[1,1]	(1,850,3,650]
Bin 2	Bin 1	Bin 2	(284,1,075.25]	[1,1]	(1,850,3,650]
Bin 3	Bin 1	Bin 2	(1,075.25,1,657]	[1,1]	(1,850,3,650]
Bin 1	Bin 1	Bin 3	[13,284]	[1,1]	(3,650,373,820]
Bin 2	Bin 1	Bin 3	(284,1,075.25]	[1,1]	(3,650,373,820]
Bin 3	Bin 1	Bin 3	(1,075.25,1,657]	[1,1]	(3,650,373,820]
Bin 1	Bin 2	Bin 1	[13,284]	(1,2]	[5,1,850]
Bin 2	Bin 2	Bin 1	(284,1,075.25]	(1,2]	[5,1,850]
Bin 3	Bin 2	Bin 1	(1,075.25,1,657]	(1,2]	[5,1,850]

Bin 1	Bin 2	Bin 2	[13,284]	(1,2]	(1,850,3,650]
Bin 2	Bin 2	Bin 2	(284,1,075.25]	(1,2]	(1,850,3,650]
Bin 3	Bin 2	Bin 2	(1,075.25,1,657]	(1,2]	(1,850,3,650]
Bin 1	Bin 2	Bin 3	[13,284]	(1,2]	(3,650,373,820]
Bin 2	Bin 2	Bin 3	(284,1,075.25]	(1,2]	(3,650,373,820]
Bin 3	Bin 2	Bin 3	(1,075.25,1,657]	(1,2]	(3,650,373,820]
Bin 1	Bin 3	Bin 1	[13,284]	(2,2,177]	[5,1,850]
Bin 2	Bin 3	Bin 1	(284,1,075.25]	(2,2,177]	[5,1,850]
Bin 3	Bin 3	Bin 1	(1,075.25,1,657]	(2,2,177]	[5,1,850]
Bin 1	Bin 3	Bin 2	[13,284]	(2,2,177]	(1,850,3,650]
Bin 2	Bin 3	Bin 2	(284,1,075.25]	(2,2,177]	(1,850,3,650]
Bin 3	Bin 3	Bin 2	(1,075.25,1,657]	(2,2,177]	(1,850,3,650]
Bin 1	Bin 3	Bin 3	[13,284]	(2,2,177]	(3,650,373,820]
Bin 2	Bin 3	Bin 3	(284,1,075.25]	(2,2,177]	(3,650,373,820]
Bin 3	Bin 3	Bin 3	(1,075.25,1,657]	(2,2,177]	(3,650,373,820]

Tabla 6 – Resultados Auto-Binner para la data de la MIPYME Distribuidora de Baterías para Autos.

Fuente: Inspiración Propia.

RECENCY [Binned]	FRECUENCY [Binned]	MONETARY [Binned]	RANGO RECENCY	RANGO FRECUENCY	RANGO MONETARY
Bin 1	Bin 1	Bin 1	[5,5]	[1,3]	[30,127.826]
Bin 2	Bin 1	Bin 1	(5,8]	[1,3]	[30,127.826]

Bin 3	Bin 1	Bin 1	(8,22]	[1,3]	[30,127.826]
Bin 1	Bin 1	Bin 2	[5,5]	[1,3]	(127.826,562.174]
Bin 2	Bin 1	Bin 2	(5,8]	[1,3]	(127.826,562.174]
Bin 3	Bin 1	Bin 2	(8,22]	[1,3]	(127.826,562.174]
Bin 2	Bin 2	Bin 1	(5,8]	(3,12]	[30,127.826]
Bin 3	Bin 2	Bin 1	(8,22]	(3,12]	[30,127.826]
Bin 1	Bin 2	Bin 2	[5,5]	(3,12]	(127.826,562.174]
Bin 2	Bin 2	Bin 2	(5,8]	(3,12]	(127.826,562.174]
Bin 3	Bin 2	Bin 2	(8,22]	(3,12]	(127.826,562.174]
Bin 1	Bin 2	Bin 3	[5,5]	(3,12]	(562.174,448,188.002]
Bin 1	Bin 3	Bin 2	[5,5]	(12,11,260]	(127.826,562.174]
Bin 1	Bin 3	Bin 3	[5,5]	(12,11,260]	(562.174,448,188.002]
Bin 2	Bin 3	Bin 3	(5,8]	(12,11,260]	(562.174,448,188.002]

Tabla 7 – Resultados Auto-Binner para la data de la MIPYME Cafetería.

Fuente: Inspiración Propia.

4.3 ANÁLISIS INFERENCIAL Y MODELOS APLICADOS

El proceso continuó creando una tabla donde juntáramos todas las posibles combinaciones que se pudieran dar (27 en total), en un diccionario donde sustituimos los nombres de los Bins por una letra del alfabeto (A-Bin 3, B-Bin 2, C-Bin 1) esto para el caso de los factores Monetización y Frecuencia, para el factor Recencia la combinación fue (A-Bin 1, B-Bin 2, C-Bin 3), lo que dispusimos para obtener un campo (**CUSTOMER_TYPE**) que contendría el resultado de la segmentación según la combinación de factores para cada cliente. Dicho diccionario mostramos a continuación;

COMBINACION	SEGMENTACION
'A'_B'_A'	Cliente Top
'A'_B'_B'	Cliente Top
'A'_B'_C'	Cliente Leal
'A'_C'_A'	Cliente Leal
'A'_C'_B'	Cliente Leal
'A'_C'_C'	Cliente Perdido
'B'_A'_A'	Cliente Top
'B'_A'_B'	Cliente Top
'B'_A'_C'	Cliente Leal
'B'_B'_A'	Cliente Top
'B'_B'_B'	Cliente Leal
'B'_B'_C'	Cliente Leal
'B'_C'_A'	Cliente Leal
'B'_C'_B'	Cliente Leal
'B'_C'_C'	Cliente Perdido
'C'_A'_A'	Cliente por Abandonar
'C'_A'_B'	Cliente por Abandonar
'C'_A'_C'	Cliente por Abandonar
'C'_B'_A'	Cliente por Abandonar
'C'_B'_B'	Cliente por Abandonar

'C'_B'_C'	Cliente por Abandonar
'C'_C'_A'	Cliente Perdido
'C'_C'_B'	Cliente Perdido
'C'_C'_C'	Cliente Perdido

Tabla 7 – Diccionario de Segmentación de Clientes.

Fuente: Inspiración Propia.

4.3.1 FLUJO DE TRABAJO EN KNIME

Una vez establecido el criterio procesamos la data con el flujo de trabajo de KNIME cuyo objetivo es procesar los datos proporcionados y generar como salida un archivo de Excel con los datos segmentados para propósito de análisis y despliegue a criterio del usuario final, y que se muestra a continuación;

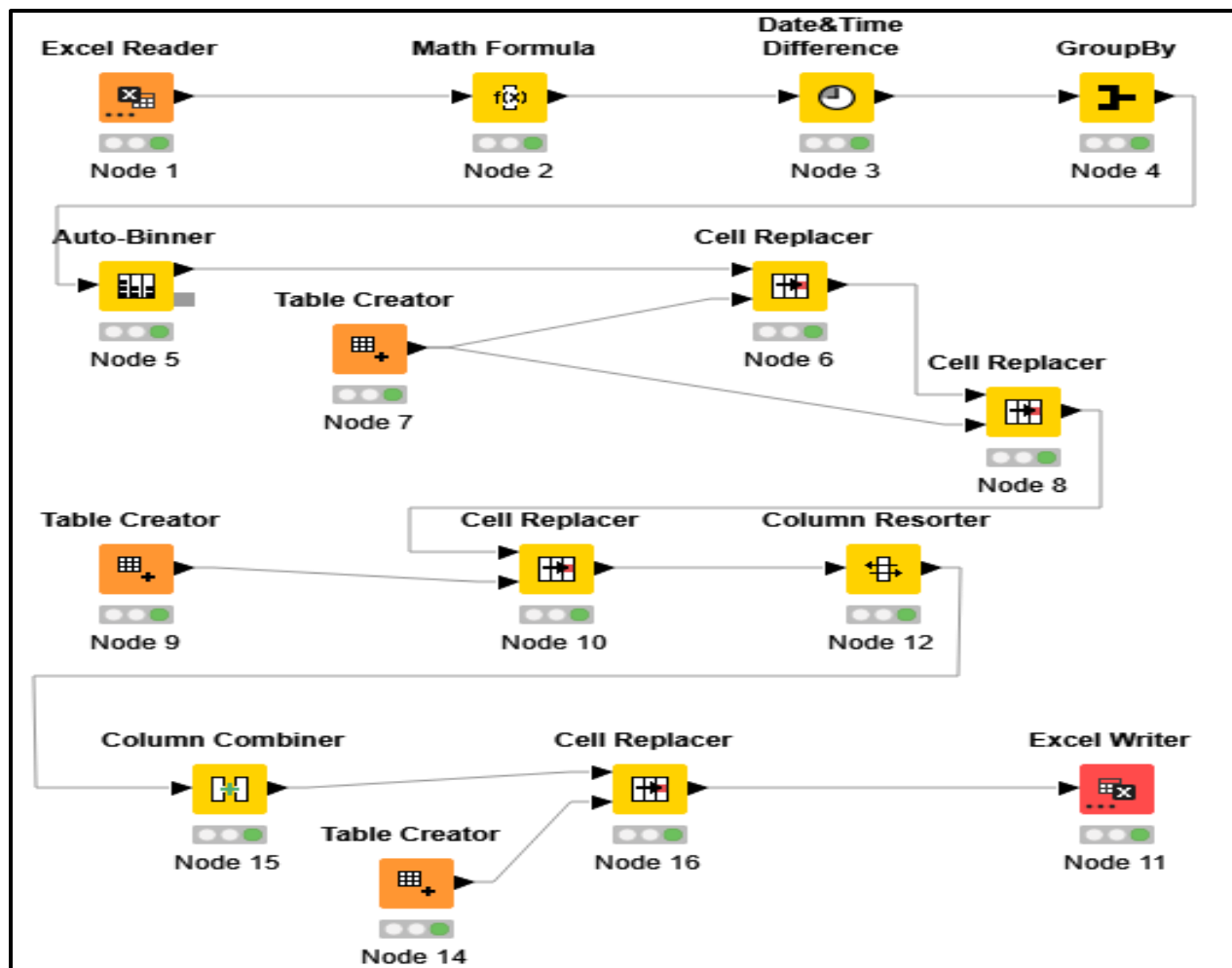


Ilustración 2 – Flujo de trabajo en KNIME para realizar perfilamiento de clientes.

Fuente: Inspiración Propia.

4.3.2 GENERACIÓN DE RESULTADOS

Con la data proceda pro el flujo de trabajo en KNIME, obtuvimos un archivo de Excel con los resultados del perfilamiento una vez aplicados los conceptos RFM y procesado el algoritmo del nodo **Auto-Binner**, dichos datos fueron cargados usando una herramienta de despliegue de uso libre como los es **POWER BI DESKTOP**, de Microsoft, y los resultados para cada empresa procesada mostramos a continuación;

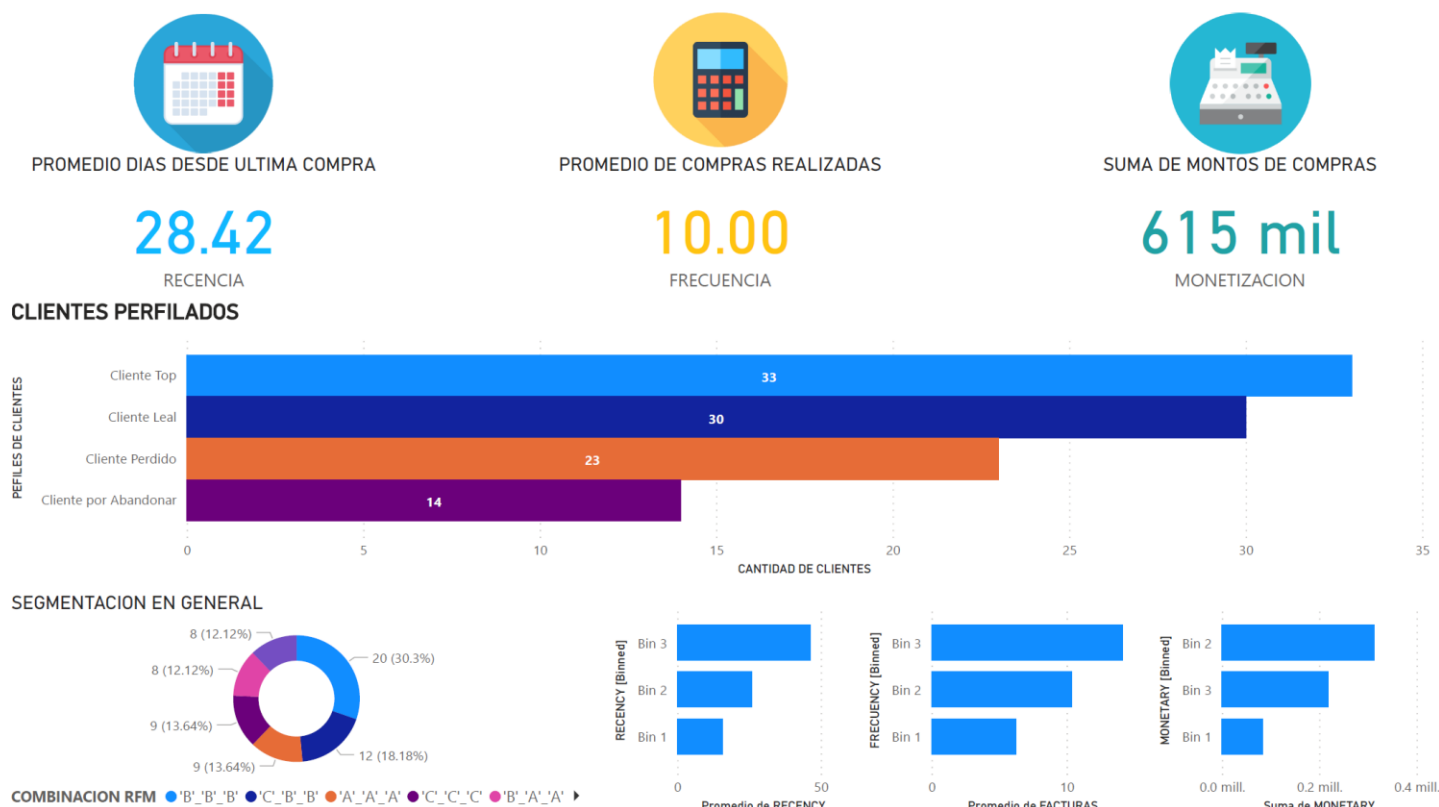


Ilustración 3 - Tablero dinámico PBI, se muestran los resultados de la data procesadas para la MIPYME Ventas en Línea.

Fuente: Inspiración Propia.



PROMEDIO DIAS DESDE ULTIMA COMPRA

697.37

RECENCIA



PROMEDIO DE COMPRAS REALIZADAS

2.07

FRECUENCIA

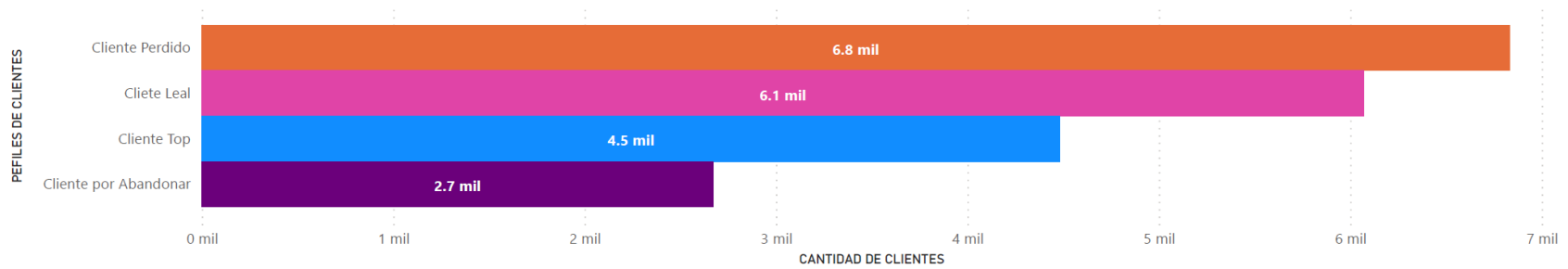


SUMA DE MONTOS DE COMPRAS

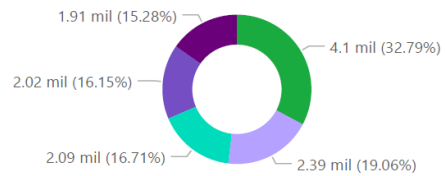
74 mill.

MONETIZACION

CLIENTES PERFILADOS



SEGMENTACION EN GENERAL



COMBINACION RFM ● 'B'_'C'_'B' ● 'C'_'C'_'B' ● 'B'_'C'_'A' ● 'B'_'C'_'C' ● 'C'_'C'_'C'

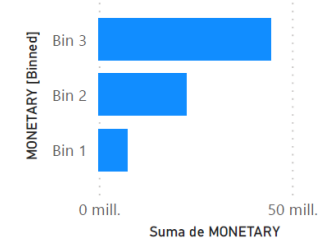
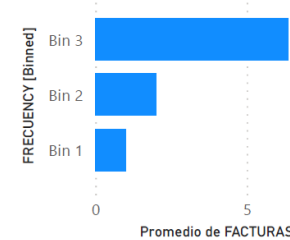
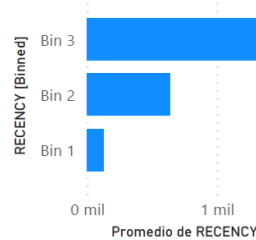


Ilustración 4 - Tablero dinámico PBI, se muestran los resultados de la data procesada para la MIPYME Baterías para Autos.

Fuente: Inspiración Propia.



PROMEDIO DIAS DESDE ULTIMA COMPRA

593.92

RECENCIA



PROMEDIO DE COMPRAS REALIZADAS

140.27

FRECUENCIA

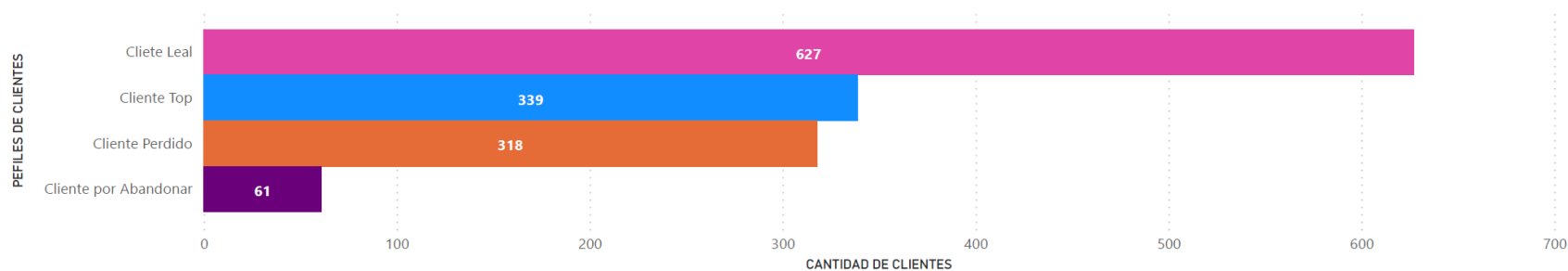


SUMA DE MONTOS DE COMPRAS

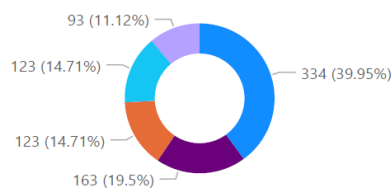
3 mil M

MONETIZACION

CLIENTES PERFILADOS



SEGMENTACION EN GENERAL



COMBINACION RFM ● 'B'_B'_B' ● 'C'_C'_C' ● 'A'_A'_A' ● 'B'_B'_A' ● 'C'_C'_B'

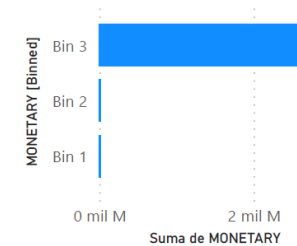
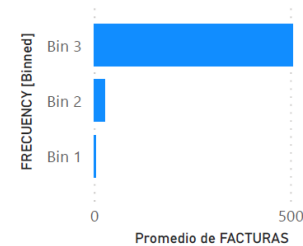
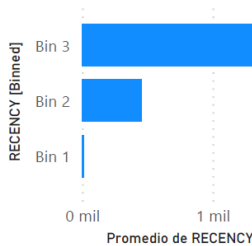


Ilustración 5 - Tablero dinámico PBI, se muestran los resultados de la data procesada para la MIPYME Distribuidora.

Fuente: Inspiración Propia.



PROMEDIO DIAS DESDE ULTIMA COMPRA

23.88

RECENCIA



PROMEDIO DE COMPRAS REALIZADAS

237.71

FRECUENCIA

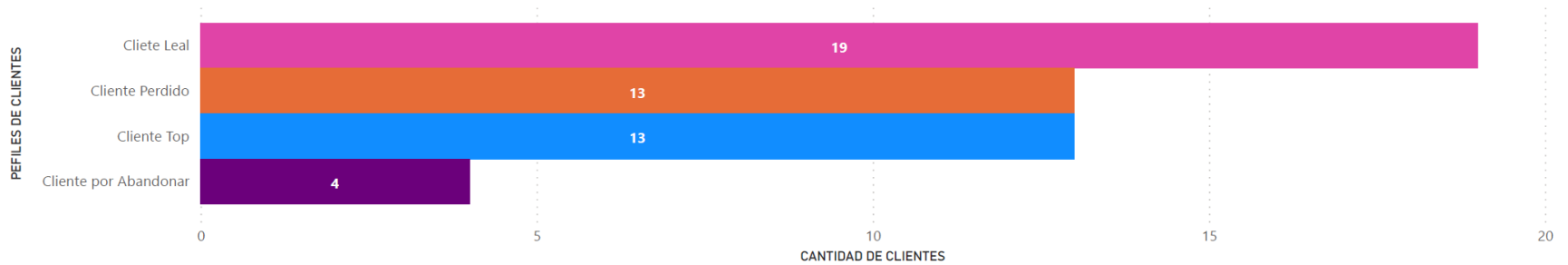


SUMA DE MONTOS DE COMPRAS

465 mil

MONETIZACION

CLIENTES PERFILADOS



SEGMENTACION EN GENERAL

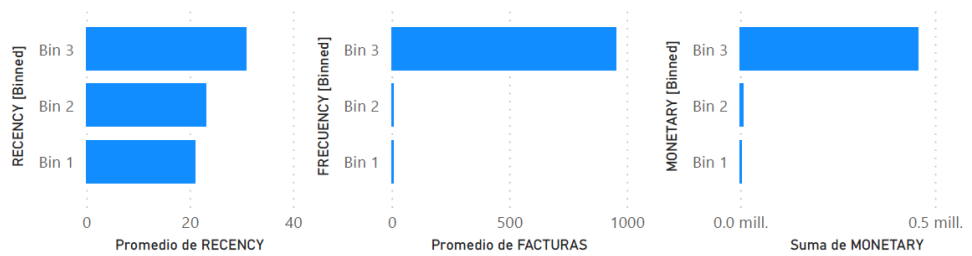
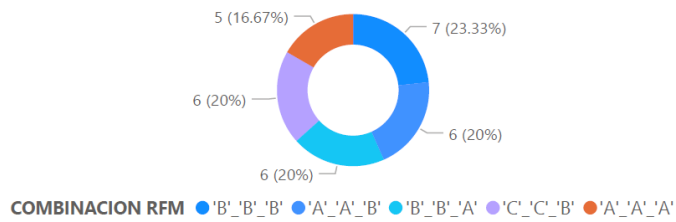


Ilustración 6 - Tablero dinámico PBI, se muestran los resultados de la data procesada para la MIPYME Cafetería.

Fuente: Inspiración Propia.

4.3.3 COMPARACION DE RESULTADOS

De los diferentes orígenes de datos procesados el resultado de la segmentación es homogéneo a pesar del hecho que las datas corresponden a empresas de diferentes rubros y con diferentes antigüedades, el algoritmo de clasificación siempre arrojo resultados muy de acorde a las realidades del estado actual de estas empresas, segmentándolos en su mayoría correctamente según esta realidad. Es obvio que al diccionario y distribución de segmentación deberán de realizarles ajustes según lo requiera la política interna e idiosincrasia de cada empresa en particular de modo que el mismo se ajuste a la necesidad de cada cual.

Mostramos como ejemplo los resultados del perfilamiento de clientes para una de las empresas, SAMAC que se dedica a la venta y distribución de a modo de tienda por departamento y producto en línea, la dueña y coparticipe de esta este trabajo de tesis, pudo constatar de primera mano que la información generada por el flujo de trabajo de KNIME con respecto a la segmentación de clientes que ellos atienden como empresa es bastante apegada a su realidad, y al tener los resultados pudo identificar acciones a seguir a partir de la información brindada y la cual mostramos a continuación;

PEFILES DE CLIENTES	CANTIDAD DE CLIENTES	NOMBRE_CLI_A_FACTUAR	PRODUCTO [Binned]	MONETARY [Binned]	REGENCY [Binned]	MONETARY_ABC	FREQUENCY_ABC	REGENCY_ABC	OVERALL_SEGME
Cliente por Abandonar	0101-1997-01611	Karina Villalvir	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0408-1983-00021	Julia Melgar	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0501-1987-06946	Ingrid Rivera	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0502-1995-00399	Cinthia Fugon	Bin 3	Bin 3	Bin 3	A	A	C	'C_'A_'A'
Cliente por Abandonar	0511-1992-01331	Katherine Cruz	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0511-1998-01986	Sarahi Rodríguez	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0601-1985-05814	Bessy Serrato	Bin 2	Bin 3	Bin 3	A	B	C	'C_'B_'A'
Cliente por Abandonar	0601-1994-00794	Olga Hernandez	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0703-1986-05066	Keila Garcia	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0801-1979-12704	Merlyn Gonzalez	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0801-1984-01990	Vanessa Vallecillo	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	0801-1989-18854	Diana Dominguez	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	1804-1995-02750	Oscar Ventura	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'
Cliente por Abandonar	1806-1980-01005	Elsy Escobar	Bin 2	Bin 2	Bin 3	B	B	C	'C_'B_'B'

Ilustración7 – Clientes por abandonar segmentación aplicado para la MIPYME SAMAC.

Fuente: Inspiración Propia.



PROMEDIO DIAS DESDE ULTIMA COMPRA

43.79

RECENCIA



PROMEDIO DE COMPRAS REALIZADAS

10.36

FRECUENCIA

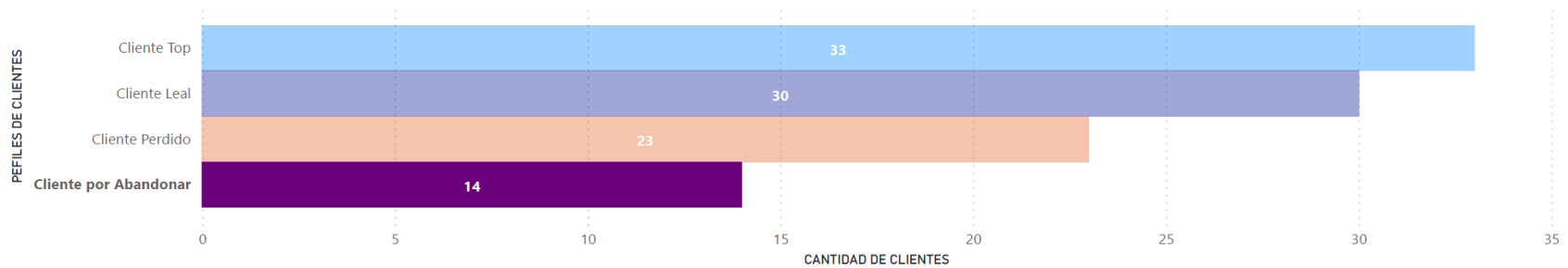


SUMA DE MONTOS DE COMPRAS

90 mil

MONETIZACION

CLIENTES PERFILADOS



SEGMENTACION EN GENERAL

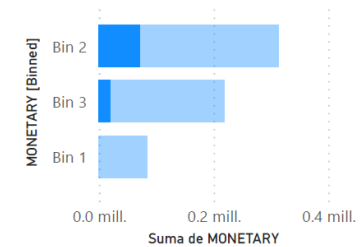
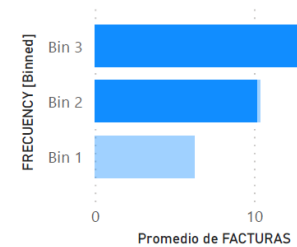
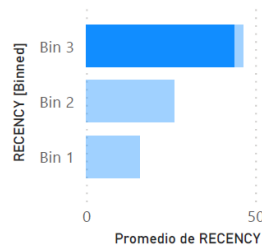
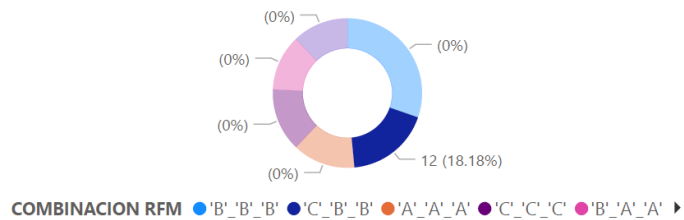


Ilustración 8 - tablero dinámico PBI Desktop, se muestran los resultados de la data procedas para la MIPYME SAMAC.

Fuente: Inspiración Propia.

Como se observa en la imagen anterior podemos tener de primera mano información valiosa relacionada con los segmentos de clientes después de procesar la data de la empresa, para el caso nos dice que tiene al menos **14 de clientes** que según combinación de criterios del RFM para esta **empresa se encuentran en un punto que corren el riesgo de retirarse** como clientes, es decir, clientes que no han comprado recientemente, su frecuencia o número de compras es poca y que los montos pueden ser buenos en algunos casos pero que son menores al de otras segmentaciones.

En este caso el tablero interactivo muestra que para estos 14 clientes el promedio de días que han pasado desde la última compra, es decir **la Recencia, es de 43.79 días**, el promedio de compras realizadas, **la Frecuencia, es de 10.36** y la suma de montos de compras, o sea la **Monetización, alcanza los 90,000 lempiras**. Solo ese último dato debe de alertar a los dueños de la empresa a hacer algo por retener a estos clientes que el mismo **representa el 14% del total de los ingresos** por venta que muestra la data proceda para esta empresa.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

Dado que en la actualidad de San Pedro sula existen MIPYMES que procesan miles de datos todos los días a través de las diferentes transacciones de ventas que realizan y debido a la necesidad de interpretación y aprovechamiento de esta información se decidió llevar a cabo esta investigación de perfilamiento de clientes a través de los datos recopilados y obtenidos de las empresas participantes donde voluntariamente nos dieron el acceso a sus datos de ventas a través de una solicitud de campos requeridos para llevar a cabo los escenarios.

Los escenarios de perfilamiento de clientes de las 5 MIPYMES participantes se lograron desarrollar utilizando la técnica del RFM en KNIME, Python y R y dado los resultados podemos mencionar que a implementación de la metodología RFM en KNIME ha demostrado ser una estrategia eficaz para la segmentación de clientes basada en su comportamiento de compra en cada una de las datas procesadas, La recencia, frecuencia y valor monetario de las compras han sido las métricas sólidas para evaluar la importancia y el compromiso de los clientes para cada una de las empresas evaluadas.

Gracias a que KNIME proporciona una plataforma versátil y accesible para procesar grandes conjuntos de datos y calcular de manera eficiente las métricas RFM la segmentación de clientes ha permitido identificar grupos de clientes de alto valor que pueden ser objetivo de estrategias de retención y fidelización.

Estos resultados de la segmentación a través de la técnica de RFM son fácilmente interpretables y nos ayudan proporcionar información valiosa para la toma de decisiones que harán cada una de las empresas participando.

La privacidad de los datos de los clientes debe ser una prioridad en todo momento, asegurándose de cumplir con todas las regulaciones aplicables.

5.2 RECOMENDACIONES

Se recomienda realizar un análisis exhaustivo de los parámetros RFM para adaptarlos a las necesidades y objetivos específicos de cada empresa, considerando, por ejemplo, la naturaleza de los productos o servicios que ofrecen ya que los mejores resultados se obtendrán manteniendo actualizada la información de compra y aplicar regularmente la segmentación RFM a medida que se acumulan nuevos datos.

Se sugiere explorar la posibilidad de combinar la segmentación RFM con otras técnicas de análisis de datos, como el aprendizaje automático, para obtener una comprensión más profunda de los patrones de comportamiento del cliente si en todo caso la empresa quisiera saber a más detalle alguno otro dato en específico que le ayude a tomar mejores decisiones o a desarrollar mejores estrategias.

Con los resultados obtenidos es crucial que se desarrollen estrategias de marketing y retención de clientes identificando claramente los distintos segmentos de clientes mediante RFM, para personalizar las ofertas y comunicaciones porque de esta forma será más efectiva la retención.

Se recomienda llevar a cabo un seguimiento constante de la efectividad de las estrategias implementadas después de la segmentación RFM y ajustarlas según sea necesario.

Continuar investigando y manteniéndose actualizado sobre las mejores prácticas en el uso de RFM y KNIME para maximizar los beneficios de esta estrategia de segmentación de clientes.

CAPÍTULO VI. APLICABILIDAD

6.1 NOMBRE DE LA PROPUESTA

Segmentación de Cliente para fidelización

La fidelización de clientes se refiere al proceso de mantener a los clientes existentes de una empresa satisfechos y comprometidos con la marca, con el objetivo de asegurar su lealtad a largo plazo. Esto implica ofrecer productos o servicios de alta calidad, una atención al cliente excepcional y experiencias personalizadas que satisfagan las necesidades y expectativas de los clientes. (Reichheld, F. F. (1996). El efecto lealtad: la fuerza oculta detrás del crecimiento, las ganancias y el valor duradero. Harvard Business Press).

6.2 JUSTIFICACIÓN DE LA PROPUESTA

La fidelización de clientes es importante porque los clientes leales son más propensos a repetir compras, recomendar la marca a otros y defenderla en caso de críticas o problemas. El perfilamiento de clientes mediante la metodología RFM es una técnica ampliamente utilizada en marketing y análisis de negocios para clasificar y segmentar a los clientes en función de su historial de compras y su comportamiento.

6.3 ALCANCE DE LA PROPUESTA

Se propone implementar el uso de herramienta KNIME, con el propósito de ejecutar un flujo de trabajo que disponga de un algoritmo de clasificación, usando la metodología de perfilamiento RFM, para poder procesar los datos de ventas y obtener la información de los clientes agrupados en segmentos que nos permita identificar el nivel de fidelidad de su cartera.

6.4 DESCRIPCIÓN Y DESARROLLO

6.4.1 PREPARACION DE DATOS

Se requiere preparar datos de histórico de venta con el fin de obtener las variables necesarias requeridas para la metodología de perfilamiento RFM.

6.4.2 PROCESAMIENTO DE DATOS

Procesar los datos del histórico de venta con el flujo de trabajo de la herramienta KNIME, con el fin de que nos genere un archivo en formato EXCEL (.xlsx), con el detalle

de la segmentación resultante.

6.4.3 VISUALIZACION Y DESPLIEGUE

Cargar el archivo generado haciendo uso de POWER BI Desktop de Microsoft, para poder indicar el resultado de la clasificación y los indicadores necesarios para que se puedan tomar decisiones para fidelización de sus clientes en base la información resultante.

6.5 MEDIDAS DE CONTROL

Es esencial implementar medidas de control efectivas para garantizar el éxito y la continuidad del programa. A continuación, presentamos algunas medidas que las MIPYMES que implementen el proyecto deben considerar:

- **Establecimiento de Indicadores Clave de Desempeño (KPIs):** Deberá definir métricas específicas que le permitan evaluar el rendimiento del programa de perfilamiento de clientes. Esto podría incluir tasas de retención de clientes, aumento en las ventas por cliente, satisfacción del cliente, entre otros.
- **Seguimiento Continuo:** Deberá establecer un sistema de seguimiento y monitoreo constante de la información de los clientes. Esto implica actualizar y enriquecer los perfiles de los clientes a medida que se obtienen nuevos datos y comportamientos.
- **Análisis de Resultados:** Realizar análisis periódicos de los datos recopilados para identificar patrones, tendencias y oportunidades de mejora. Esto le ayudará a adaptar su estrategia de fidelización según los hallazgos.
- **Automatización de Procesos:** Utilizar herramientas y sistemas de gestión de relaciones con el cliente (CRM) para automatizar parte del proceso de perfilamiento y seguimiento, lo que garantiza la consistencia y la eficiencia en la gestión de la información.
- **Retroalimentación del Cliente:** Establecer canales para recibir retroalimentación directa de los clientes sobre sus experiencias y preferencias. Esto le ayudará a ajustar tu enfoque y servicios según sus

necesidades cambiantes.

- **Capacitación del Personal:** Asegurarse que el personal esté capacitado en la utilización de los datos de perfilamiento y en la implementación de estrategias de fidelización. Esto garantizará un enfoque coherente en todos los puntos de contacto con el cliente.
- **Mantenimiento de la Privacidad de los Datos:** Cumple con las regulaciones de protección de datos y garantiza que la información de los clientes se almacene y maneje de manera segura y ética.

Al establecer estas medidas de control, se podrá maximizar los beneficios del perfilamiento de clientes para la fidelización, garantizando una estrategia sólida y sostenible a lo largo del tiempo.

6.6 CRONOGRAMA DE IMPLEMENTACIÓN Y PRESUPUESTO

6.6.1 CRONOGRAMA DE IMPLEMENTACION

Tiempo		Actividad	Detalle de Actividades
Semana	1 y 2	Fase de Planificación	1. Definición de objetivos y metas del perfilamiento de clientes. 2. Selección de herramientas y tecnologías de perfilamiento. 3. Formación del equipo de proyecto.
Semana	3 y 4	Fase de Recopilación de Datos	1. Diseño de cuestionarios y formularios de perfilamiento. 2. Implementación de sistemas de recopilación de datos. 3. Recopilación de datos iniciales de los clientes.
Semana	5 y 6	Fase de Análisis y Segmentación	1. Procesamiento y análisis de los datos recopilados. 2. Segmentación de la base de clientes en grupos o categorías. 3. identificación de patrones de comportamiento.
Semana	7 y 8	Fase de Estrategia de Fidelización	1. Desarrollo de estrategias de fidelización personalizadas para cada segmento. 2. Implementación de programas de recompensas o incentivos. 3. Diseño de campañas de marketing específicas para cada segmento.
Semana	9 y 10	Fase de Implementación	1. Puesta en marcha de las estrategias de fidelización. 2. Capacitación del personal en la nueva estrategia. 3. Monitorización inicial de resultados.
Semana	11 y 12	Fase de Evaluación y Optimización Continua	1. Evaluación de los resultados obtenidos. 2. Ajustes y mejoras en las estrategias según los resultados. 3. Mantenimiento y actualización continua de la base de datos de clientes.

Tabla 8 – Cronograma de Implementación.

Fuente: Inspiración Propia.

6.6.2 PRESUPUESTO

Tarea	Descripción	Valor
Tecnología y Software	Costos relacionados con la adquisición de software de CRM, herramientas de análisis de datos y sistemas de recopilación de datos.	Lps. 15,000.00 Laptop *Software de uso Libre
Recurso Humano	Salarios y capacitación del equipo involucrado en el proyecto, que incluye analistas de datos, personal de marketing y otros profesionales, incluye contrataciones de asesores externos.	Lps. 70,000.00/Mensual
Marketing y Promoción	Fondos destinados a campañas de marketing, publicidad, y programas de fidelización, como descuentos o recompensas para los clientes.	Lps. 100,000.00/Mensual
Capacitación	Gastos relacionados con la formación del personal en las nuevas estrategias y herramientas.	Lps. 50,000.00/Anual

Tabla 9 – Presupuesto.
Fuente: Inspiración Propia.




REFERENCIAS BIBLIOGRÁFICAS




- Amaya, O. A./ Macías, M.M. (2015). Situación de las MIPYMES en Honduras, retos y oportunidades. Portal De La Ciencia, 5, 87-97.
- Navarro, D. (2021). <https://dplnews.com/honduras-apoyara-a-15-mil-emprendedores-para-digitalizar-a-las-mipymes/>, Recuperado en Agosto, 2023.
- La Gaceta (2022) Acuerdo ejecutivo STSS-308-2022 https://www.tsc.gob.hn/web/leyes/Acuerdo_Ejecutivo_STSS-308-2022.pdf/, Recuperado en Agosto, 2023.
- <https://www.tsc.gob.hn/web/leyes/Reglamento%20del%20Servicio%20de%20Internet%20o%20Acceso%20a%20Redes%20Inform%C3%A1ticas.pdf>, Recuperado en Agosto, 2023.
- Hub – UNITEC <https://hub.unitec.edu/tecnologia/disenyo-y-desarrollo-web-una-profesion-innovadora-orientada-al-emprendimiento-digital/>
- Gastón Becerra y Juan Pablo López Alurralde, Universidad de Buenos Aires Argentina (2017), Big data y Data mining. Un análisis crítico acerca de su significación. https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=teorias+de+sustento+para+el+ analisis+de+datos+big+data&btnG=/, Recuperado en Agosto, 2023.
- Universidad Politécnica de Valencia, Walid Ghoobar, E. (2017). Un sistema de recomendación basado en perfiles generados por agrupamiento y asociaciones. <http://hdl.handle.net/10251/94049/>, Recuperado en Agosto, 2023.
- PINEDA, Beatriz; DE ALVARADO, Eva Luz; DE CANALES, Francisca 1994 Metodología de la investigación, manual para el desarrollo de personal de salud, Segunda edición. Organización Panamericana de la Salud. Washington.
- MATA, María Cristina; MACASSI, Sandro; 1997 Cómo elaborar muestras para los sondeos de audiencias. Cuadernos de investigación No 5. ALER, Quito.
- Hernández, Fernández y Baptista en su Metodología de la Investigación (1991)



- Hair, J. F., Celsi, M., Ortinau, D. J., & Bush, R. P. (2019). Investigación de mercados. McGraw-Hill Education.
- Wedel, M., & Kannan, P. K. (2016). Análisis de marketing para entornos ricos en datos. *Journal of Marketing*, 80(6), 97-121.
- Provost, F., & Fawcett, T. (2013). Ciencia de datos para empresas: lo que necesita saber sobre la minería de datos y el pensamiento analítico de datos. O'Reilly Media.
- Han, J., Kamber, M., & Pei, J. (2011). Minería de datos: conceptos y técnicas. Elsevier.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introducción al análisis de regresión lineal. John Wiley & Sons.
- Kotler, P., & Armstrong, G. (2017). Fundamentos de marketing. Pearson Educación.
- Reichheld, F. F. (1996). El efecto lealtad: la fuerza oculta detrás del crecimiento, las ganancias y el valor duradero. Harvard Business Press.
- Peppers, D., & Rogers, M. (2004). Gestión de las relaciones con los clientes: un marco estratégico. John Wiley & Sons.)
- H. Sampieri, citado por Balestrini 2001 Pág. 141). Existen varias clases de muestras no probabilísticas, para Hernández, Fernández y Baptista en su Metodología de la Investigación (1991)
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (1991). Metodología de la investigación. McGraw-Hill.



ANEXOS

Anexo 1 Documentación flujo de trabajo de KNIME

Nodo	Descripción	Uso en Flujo de Trabajo
	<p>Excel Reader: Este nodo lee archivos de Excel (formatos xlsx, xlsxm, xlsb y xls). Puede leer uno o varios archivos al mismo tiempo, pero solo lee una hoja por archivo.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos, para cargar la data proporcionada por las empresas en formato Excel (xlsx).</p>
	<p>Math Formula: Este nodo evalúa una expresión matemática basada en los valores de una fila. Los resultados calculados pueden ser añadidos como una nueva columna o utilizados para reemplazar una columna de entrada.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para crear la columna MONETARY, como resultado de multiplicar el PRECIO_AFECTADO por la CANTIDAD_VENDIDA de la data proporcionada por las empresas.</p>
	<p>Date & Time Difference: Calcula las diferencias entre dos celdas de fecha y hora y agrega una nueva columna. La salida puede ser una duración o una granularidad seleccionada (años, meses, semanas, días, horas, minutos, segundos, milisegundos, nanosegundos).</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para crear la columna RECENCY, basándonos en los días que han pasado desde que se hizo la compra, ocupamos el campo FECHA_DE_FACTURA y obtuvimos la diferencia a la fecha que se ejecute el flujo de trabajo</p>

	<p>GroupBy: Agrupa las filas de una tabla por los valores únicos en las columnas de grupo seleccionadas. Se crea una fila para cada conjunto único de valores de la columna de grupo seleccionada. Las columnas restantes se agregan en función de la configuración de agregación especificada.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para crear la columna FREQUENCY, agrupando la data por cliente para contar la cantidad de compras que se realiza a cada cliente, hicimos un conteo del campo NUMERO_DE_FACTURA.</p>
	<p>Auto-Binner: Este nodo permite agrupar datos numéricos en intervalos, llamados bins. Hay dos opciones de nomenclatura para los bins y dos métodos que definen el número y el rango de valores que caen en un bin.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para establecer los diferentes segmentos de clientes basados en las columnas creadas: RECENCY, FREQUENCY, MONETARY. Catalogando a cada una con un Bin, de acuerdo con el rango en que los valores de estas se encontraban según lo determinaba el rango de cuantiles especificados en la configuración del nodo.</p>
	<p>Table Creator: Permite la creación manual de una tabla de datos. Los datos se pueden ingresar en una tabla similar a una hoja de cálculo.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para crear tablas con valores, diccionario, de categorías A,B o C para los Bins asignados por el algoritmo a los campos FREQUENCY, RECENCY y MONETARY, también se usó para darle nombre a las categorías finales de los clientes de acuerdo a las combinaciones de Bins que se dieron de los tres</p>

		factores que componen el RFM.
	Cell Replacer: Reemplaza las celdas en una columna según la tabla de diccionario (segunda entrada). El nodo tiene dos entradas: la primera entrada contiene una columna de destino cuyos valores deben ser reemplazados usando la tabla de diccionario (segunda entrada).	Usado en el flujo de trabajo de perfilamiento que utilizamos para crear columnas FREQUENCY_ABC, RECENCY_ABC y MONETARY_ABC, que contendrán los valores resultantes del diccionario de datos para cada Bins arrojado por el algoritmo. En otra instancia se usa para crear columna de CUSTOMER_TYPE que contendrá la categoría final que tendrá el cliente de acuerdo con la combinación de Bins asignados por el algoritmo.
	Column Resorter: Este nodo cambia el orden de las columnas de entrada, en función de la configuración definida por el usuario. Las columnas se pueden desplazar en pasos simples hacia la izquierda o hacia la derecha, o completamente al principio o al final de la tabla de entrada. Además, las columnas también se pueden ordenar según su nombre. La tabla reordenada se proporciona en el puerto de salida.	Usado en el flujo de trabajo de perfilamiento que utilizamos para definir el orden en que queremos que queden ordenadas las columnas FREQUENCY_ABC, RECENCY_ABC y MONETARY_ABC, de modo que cumplan el criterio RFM; RECENCIA Primero, luego FRECUENCIA y por último MONETIZACION.

	<p>Column Combiner: Combina el contenido de un conjunto de columnas y agrega la cadena concatenada como una columna separada a la tabla de entrada. El usuario debe especificar en el diálogo las columnas de interés y algunas otras propiedades, como el delimitador para separar los diferentes contenidos de las celdas y las opciones de citado.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para combinar las columnas, previamente ordenadas, FREQUENCY_ABC, RECENCY_ABC y MONETARY_ABC, con el fin de obtener una denominación única para cada cliente combinando cada criterio de RFM y posteriormente asignarle un tipo de clientes de acuerdo con el criterio establecido para cada una de las veintisiete (27) combinaciones posibles.</p>
	<p>Excel Writer: Este nodo escribe la tabla de datos de entrada en una hoja de cálculo de un archivo de Excel, que luego puede ser leído con otras aplicaciones como Microsoft Excel. El nodo puede crear archivos completamente nuevos o agregar datos a un archivo de Excel existente. Al agregar, los datos de entrada se pueden agregar como una nueva hoja de cálculo o después de la última fila de una hoja de cálculo existente.</p>	<p>Usado en el flujo de trabajo de perfilamiento que utilizamos para generar como salida un archivo de Excel con los datos segmentados para propósito de análisis y despliegue a criterio del usuario final.</p>

Anexo 2 Fichas de observación para análisis de campos críticos.

ClientID	Gender	EducationLevel	MaritalStatus	Age	Territory	HasChildren	JoinDate	SalesIn2022	ChurnIn2022
C0000806173	Female	Undergraduate	Single	22	Centro	N	08/06/2020	\$ 5,726.00	0
C0000578289	Female	High School	Married	30	Sur	Y	07/03/2020	\$ 5,368.00	1
C0000696548	Male	Undergraduate	Married	55	Norte	Y	02/11/2020	\$ 3,477.00	0
C0000968182	Male	Elementary	Single	28	Norte	N	09/20/2019	\$ 1,713.00	1
C0000806399	Female	Undergraduate	Single	30	Occidente	N	01/12/2019	\$ 48,726.00	0
C0000768694	Female	High School	Single	24	Oriente	N	11/12/2019	\$ 44,644.00	0
C0000189952	Female	Elementary	Single	40	Norte	N	01/28/2020	\$ 9,624.00	0
C0000227245	Female	Undergraduate	Married	45	Centro	N	08/22/2020	\$ 8,538.00	0
C0000312953	Female	Undergraduate	Married	42	Occidente	N	06/12/2021	\$ 10,041.00	1

Atributo	Descripcion
ClientID	Identificador Unico del Cliente
Gender	Genero del cliente
EducationLevel	Nivel educativo mas alto alcanzado
Age	Indica la edad del cliente
Territory	Indica la region origen de cliente
MaritalStatus	Estado Civil
HasChildren	Tiene hijos. Y = Si, N = No
JoinDate	Fecha en que realizo su primera compra.
SalesIn2022	Ventas atribuibles al cliente durante 2022.
ChurnIn2022	Indica si el cliente cancelo su cuenta en 2022. 0 = El cliente todavia posee una cuenta Activa, 1 = El cancelo su cuenta.

ClientID	DateOfSales	Invoice	IdProduct	Descriptions	Quantity	Price	Subtotal	Taxes	TotalAmount	PayMode
C0000926218	03/15/2021	45551	PR-0001	Leche	1	25	25	3.75	28.75	TC
C0000301017	03/18/2021	52536	PR-0002	Cereal	5	25	125	0	125	EF
C0000536864	03/21/2021	59521	PR-0003	Detergente Liquido	20	30	600	90	690	TF
C0000933062	03/24/2021	66506	PR-0004	Detergente en Polvo	2	32	64	9.6	73.6	EF
C0000554246	03/27/2021	73491	PR-0005	Acondicionador	1	100	100	0	100	EF
C0000589182	03/30/2021	80476	PR-0006	Shampoo	5	150	750	112.5	862.5	TC
C0000182633	04/02/2021	87461	PR-0007	Harina	3	55	165	24.75	189.75	DE
C0000287491	04/05/2021	94446	PR-0008	Manteca	5	55	275	0	275	DE
C0000944667	04/08/2021	101431	PR-0009	Frijoles	4	22	88	13.2	101.2	TF

Atributo	Descripcion
ClientID	Identificador Unico del Cliente
DateOfSales	Fecha de Venta
Invoice	Numero de docuemnto
IdProduct	Código de Producto
Descriptions	Describe el Producto
Quantity	Cantidad de Producto
Price	Precio del Producto
Subtotal	Subtotal de la venta
Taxes	Impuesto Sobre Venta
TotalAmount	Total de la Venta
PayMode	Forma de Pagos

Anexo 3 Notas de Aceptación uso de los datos.

samac

Nota de Aceptación de Uso de Datos

Por medio de la presente, otorgo mi consentimiento para el uso de los datos de ventas de mi empresa, **NOVEDADES SAMAC** por parte de Maestranteros del post grado de analítica de negocios, Nelson Cabrera y Samaria Hernández, con fines específicos y bajo las condiciones previamente acordadas.

Entiendo y acepto que los datos de ventas de mi empresa serán utilizados exclusivamente para los siguientes fines: ***Análisis y Estadísticas:** * Los datos de ventas se utilizarán para realizar análisis y estadísticas relacionadas con el desempeño de **SAMAC** con el objetivo de mejorar nuestras estrategias y operaciones.

Acepto que los maestranteros Nelson Cabrera y Samaria Hernández se comprometen a tratar mis datos de ventas con confidencialidad y a cumplir con todas las leyes y regulaciones de protección de datos aplicables.

Este consentimiento es válido desde la fecha de mi aceptación y podrá ser revocado por escrito en cualquier momento si así lo requiero. Quedo a disposición para cualquier aclaración o consulta adicional sobre el uso de los datos de ventas de SAMAC

Atentamente,



Samaria Poleth Hernandez Vasquez

Representante Legal

SAMAC

01 de agosto 2023



+504 9810-7614 ■ Samac_hn

✉ Samacsps@hotmail.com

@ Samac HN