

Escuela Superior Politécnica del Litoral

Facultad de Ciencias Sociales y Humanísticas

Desarrollo de un Modelo de Regresión para la Gestión de Cobranza en la
Contratista COINSER de Interagua: Aplicación de Clústeres y Probabilidades
Estadísticas en la Recuperación de Cartera Vencida (ADMI-1150)

Proyecto Integrador

Previo la obtención del Título de:

Nombre de la titulación

Licenciado en Auditoría y Control de Gestión

Presentado por:

Fernanda Fiorella Cunalema Arana

Doménica Stefania Narváez Bustamante

Guayaquil - Ecuador

Año: 2025

Dedicatoria

Con profunda gratitud y un inmenso amor, dedico este trabajo, que es el fruto de mi esfuerzo y sacrificio, a mis más grandes amores: mis padres, quienes, desde que era pequeña, han sido mis pilares. Gracias por brindarme siempre su fortaleza, amor y dedicación incondicional. Ustedes han formado cada paso de mi vida académica y me han enseñado a no rendirme, a luchar por mis sueños sin importar las dificultades. A mis abuelos, que, con su sabiduría y ternura, me inspiran a ser una mujer exitosa. A mis mejores amigas, Gene y Nai, por su presencia, consejos y acompañamiento en mi vida. A mi bebé conejo, mi sobrino, que este año llegó para alegrarme. Y, a mis mascotas (Coco, Chester y Coddy), por su presencia en las largas noches de desvelo.

Fernanda Cunalema Arana.

Dedicatoria

Dedico este proyecto a mis motores más grandes, mis hijas, Samantha y Joselyn. Gracias a su amor profundo, cada día realizamos juntos un sinnúmero de esfuerzos y sacrificios para que pueda avanzar en mis estudios. Este logro es para ellas y para mí, con la esperanza de alcanzar una vida mejor y de brindarles un futuro lleno de oportunidades. A mi mamá, Kennya Bustamante, quien siempre me ha dado su apoyo incondicional y ha creído en mí en todo momento. Ella ha sido mi pilar, acompañándome en cada paso hacia la realización de mis metas. También agradezco a mi papá y a mi hermano, quienes han hecho todo lo posible para ayudarme en este proceso. Finalmente, a mi familia y amigos, quienes me han acompañado con su apoyo y compañía, siendo una fuente constante de motivación y aliento.

Doménica Narváez Bustamante

Agradecimiento

Mi más sincero agradecimiento a Dios, por ser mi guía principal en este proceso. A mis profesores Christian y Daniel por su acompañamiento con sabiduría y paciencia. A mis amistades por darme paz en los momentos de angustias, especialmente a Cristhel S. por su generosidad al brindarme su tiempo y energía cuando más necesita. A todas las personas que, de una u otra forma, contribuyeron, les agradezco de todo corazón.

Fernanda Cunalema Arana.

Agradecimiento

Gracias a Dios por darme la fuerza y constancia necesarias para alcanzar este objetivo que tanto anhelé. Mi agradecimiento también va para mi mamá, quien siempre ha estado a mi lado, brindándome su amor y apoyo en cada paso de mi carrera. A mis hijas, les doy las gracias por su paciencia y comprensión durante aquellos días en los que no pude brindarles toda la atención que necesitaban y merecían.

A mis tutores, Christian y Daniel, les agradezco profundamente por su paciencia y por compartir generosamente su conocimiento. Finalmente, gracias a todas las personas que me sirvieron de inspiración para ser mejor cada día.

Doménica Narvárez Bustamante

Declaración Expresa

Nosotros Cunalema Arana Fernanda Fiorella y Narváez Bustamante Doménica Stefania acordamos y reconocemos que:

La titularidad de los derechos patrimoniales de autor (derechos de autor) del proyecto de graduación corresponderá al autor o autores, sin perjuicio de lo cual la ESPOL recibe en este acto una licencia gratuita de plazo indefinido para el uso no comercial y comercial de la obra con facultad de sublicenciar, incluyendo la autorización para su divulgación, así como para la creación y uso de obras derivadas. En el caso de usos comerciales se respetará el porcentaje de participación en beneficios que corresponda a favor del autor o autores.

La titularidad total y exclusiva sobre los derechos patrimoniales de patente de invención, modelo de utilidad, diseño industrial, secreto industrial, software o información no divulgada que corresponda o pueda corresponder respecto de cualquier investigación, desarrollo tecnológico o invención realizada por mí/nosotros durante el desarrollo del proyecto de graduación, pertenecerán de forma total, exclusiva e indivisible a la ESPOL, sin perjuicio del porcentaje que me/nos corresponda de los beneficios económicos que la ESPOL reciba por la explotación de mi/nuestra innovación, de ser el caso.

En los casos donde la Oficina de Transferencia de Resultados de Investigación (OTRI) de la ESPOL comunique los autores que existe una innovación potencialmente patentable sobre los resultados del proyecto de graduación, no se realizará publicación o divulgación alguna, sin la autorización expresa y previa de la ESPOL.

Guayaquil, 11 de octubre del 2024.



Cunalema Arana
Fernanda Fiorella



Narváez Bustamante
Doménica Stefania

Evaluadores

Vera Alcivar Christian

Profesor de Materia

Contreras Moscol Daniel Jossepp

Tutor de proyecto

Resumen

Este proyecto integrador busca generar impacto positivo para las contratistas, el cual tiene como objetivo mejorar la gestión de cobranza de la cartera morosa de Interagua que es administrada por la contratista Coinser S.A.S, incrementando un modelo predilecto de regresión logística y técnicas de conglomerado por clústeres para un rendimiento óptimo de la recuperación de cartera morosa. La idea sustentada es entregar un modelo predictivo que puede identificar a los usuarios con mayor probabilidad de regularizar la deuda de los morosos, permitiendo enmarcar los esfuerzos de cobranza de manera más eficiente. Se justifica la investigación por la creciente morosidad en el sector de servicios básicos, la cual afecta la viabilidad financiera y la credibilidad de la contratista donde su capacidad para reactivar el servicio a los usuarios morosos es compleja. Para el desarrollo del proyecto, se utilizó la base de datos de los usuarios de Interagua del último trimestre del 2024, donde se aplicaron técnicas estadísticas para la ejecución de la regresión logística, utilizando herramientas como Excel, Minitab y Power BI. El resultado logró predecir de manera efectiva a los usuarios con índice de probabilidad de pago, es decir, en base a este proyecto se realizó una proyección para el primer trimestre del 2025 dando buena aprobación para la recuperación de la cartera morosa y beneficios económicos futuros para la Contratista. En conclusión, el modelo es eficaz y realizable, ya que ofrece una solución que optimiza la gestión constante de las cuadrillas, incrementando las reconexiones mensuales y los resultados positivos para Interagua.

Palabras Clave: Viabilidad financiera, Interagua, Morosidad, Modelo predilecto.

Abstract

This integrative project seeks to generate positive impact for contractors, which aims to improve the collection management of Interagua's delinquent portfolio that is managed by the contractor Coinser S.A.S., increasing a preferred logistic regression model and clustering techniques for optimal performance in the recovery of delinquent portfolio. The idea is to deliver a predictive model that can identify the users most likely to regularize the debt of the delinquent ones, allowing to frame the collection efforts in a more efficient way. The research is justified by the growing delinquency in the basic services sector, which affects the financial viability and credibility of the contractor where its ability to reactivate the service to delinquent users is complex. For the development of the project, the database of Interagua users of the last quarter of 2024 was used, where statistical techniques were applied for the execution of logistic regression, using tools such as Excel, Minitab and Power BI. The result was able to effectively predict users with a probability of payment index, i.e., based on this project, a projection was made for the first quarter of 2025, giving good approval for the recovery of the delinquent portfolio and future economic benefits for the Contractor. In conclusion, the model is effective and feasible, since it offers a solution that optimizes the constant management of the crews, increasing monthly reconnections and positive results for Interagua.

Keywords: Financial viability, Interagua, Delinquency, Preferred model.

Índice general

| | |
|---|----|
| Evaluadores..... | 7 |
| Resumen | 8 |
| Abstract | 9 |
| Índice general..... | 10 |
| Índice de figuras..... | 12 |
| Índice de tablas | 13 |
| Capítulo 1 | 14 |
| 1 Introducción..... | 15 |
| 1.1 Descripción del Problema | 17 |
| 1.2 Justificación del Problema | 19 |
| 1.2.1 Población en estudio..... | 20 |
| 1.3 Objetivos | 22 |
| 1.3.1 Objetivo general..... | 22 |
| 1.3.2 Objetivos específicos..... | 23 |
| 1.4 Marco teórico..... | 23 |
| 1.4.1 Marco conceptual..... | 23 |
| Capítulo 2 | 44 |
| 2 Metodología..... | 45 |
| 2.1 Alternativas de Solución | 45 |
| 2.2 Desarrollo de un modelo predilecto: | 46 |
| 2.2.1 Diseño Conceptual y Metodología Adoptada | 46 |
| 2.2.2 Recolección de datos | 47 |
| 2.2.3 Herramientas que se utilizaron..... | 47 |

| | | |
|-------------|---|----|
| 2.3 | Población y muestra | 51 |
| Capítulo 3 | | 56 |
| 3 | Resultados y análisis | 57 |
| 3.1 | Análisis de las variables | 57 |
| 3.2 | Resultados por Minitab | 68 |
| 3.3 | Análisis de Clústeres | 77 |
| 3.4 | Presentación de Resultados Power Bi | 83 |
| 3.5 | Análisis de Costos | 87 |
| Capítulo 4 | | 89 |
| 4 | Conclusiones y recomendaciones..... | 90 |
| 4.1 | Conclusiones..... | 90 |
| 4.2 | Recomendaciones..... | 91 |
| Referencias | | 95 |

Índice de figuras

| | |
|---|----|
| Figura 1 <i>Segmentación de ciclos dentro de guayaquil</i> | 39 |
| Figura 2 <i>Reconocimiento de los sectores peligrosos</i> | 40 |
| Figura 3 <i>Detalle del Método</i> | 69 |
| Figura 4 <i>Eventos relacionados a la Variable dependiente</i> | 69 |
| Figura 5 <i>Ecuación de regresión logística ajustada</i> | 69 |
| Figura 6 <i>Análisis de los coeficientes</i> | 70 |
| Figura 7 <i>Relaciones de probabilidad</i> | 73 |
| Figura 8 <i>Resumen del modelo</i> | 74 |
| Figura 9 <i>Pruebas de bondad de ajuste</i> | 75 |
| Figura 10 <i>Análisis de la varianza</i> | 76 |
| Figura 11 <i>Dispersión a partir de CLUSTERES con un $K=3$</i> | 78 |
| Figura 12 <i>Centroides de grupo</i> | 79 |
| Figura 13 <i>Partición final</i> | 79 |
| Figura 14 <i>Gráfica de dispersión de Rango vs. Saldo</i> | 82 |
| Figura 15 <i>Reconexiones 4to Trimestre Periodo 2024</i> | 84 |
| Figura 16 <i>Reconexiones Proyectadas al 1er Trimestre Periodo 2025</i> | 85 |

Índice de tablas

| | |
|--|----|
| Tabla 1 <i>Segmentación de ciclos por productos y deuda de los usuarios.</i> | 21 |
| Tabla 2 <i>Cantidades de planillas vencidas por número de factura</i> | 22 |
| Tabla 3 <i>Codificación Dummy Reconexión</i> | 57 |
| Tabla 4 <i>Codificación de Facturas</i> | 58 |
| Tabla 5 <i>Codificación \$Adeudado</i> | 59 |
| Tabla 6 <i>Codificación del Rango</i> | 60 |
| Tabla 7 <i>Codificación de los ciclos</i> | 61 |
| Tabla 8 <i>Codificación Corte</i> | 62 |
| Tabla 9 <i>Codificación Predio</i> | 63 |
| Tabla 10 <i>Codificación Contribuyente</i> | 64 |
| Tabla 11 <i>Codificación Call center</i> | 65 |
| Tabla 12 <i>Codificación Visita</i> | 66 |
| Tabla 13 <i>Codificación Corte</i> | 66 |
| Tabla 14 <i>Codificación Contravención</i> | 67 |
| Tabla 15 <i>Presentación de los diagramas Análisis 2024 vs 2025</i> | 86 |
| Tabla 16 <i>Análisis conclusión</i> | 87 |

Capítulo 1

1 Introducción

En Ecuador, una de las ciudades más habitada es Guayaquil, pero esta no se aleja de los retos financieros que atraviesa el país (Chávez Páez, 2021). En la escala de la posición económica alrededor de 96% de ecuatorianos aun contando con trabajo les está perjudicando la presente circunstancia del país a causa de los componentes de incremento de la inseguridad, emergencia energética y fenómenos climáticos lo que genera como resultado un déficit en los abonos de los pagos mensuales que tiene el beneficiario (INEC, 2024).

La morosidad en los pagos, un problema que se ha perpetuado en los últimos años, representa un reto considerable para las empresas locales. En el primer semestre del 2024 el Ecuador atravesó un periodo de recesión fuerte en la economía de todos los ecuatorianos; datos liberados por el Banco Central del Ecuador, la tasa de morosidad crediticia crece y está siendo evaluada como la más alta en el mercado, la cifra tiene un incremento de 0,8 puntos porcentuales en comparación con el 2023, esto trae como impacto la reducción de la actividad económica, menor inversión y producción. El crédito que tiene una alta tasa de incumplimiento es el microcrédito, ya que tiene componentes como: gastos, formación y domicilio. Lo que se comprende que hay un período de un declive financiero que demuestra una reducción en la comercialización, inactividad laboral y una menor financiación de creadores de negocios en el país, inclusive problemas para tener acceso a préstamos recientes disminuyendo la capacidad de compra para núcleos familiares en Ecuador (Banco Central del Ecuador, Información Económica, 2024).

Según la información liberada por el Instituto Nacional de Estadísticas y Censos, INEC organización superior y designada de ofrecer datos estadísticos de manera

funcional para el progreso y evaluación de la condición del país. Se realiza una comparación entre el 2023 y el 2024, en donde se demuestra que la falta de trabajo tiene una modificación de 0,2 puntos porcentual de incremento distinto a el 2023 donde se ve con exactitud que el país está excluyendo a los ecuatorianos a no tener trabajo, se puede entender que en el 2024 hubo un aumento de 21.390 personas sin trabajo en el Ecuador, teniendo en cuenta a los adultos y jóvenes. Dando alcance a nuestro punto fijo que es Guayaquil, existe una diferencia en base a las demás provincias; en el Guayas representando el 3,1% de la población referencial estadística a nivel nacional; el desempleo bajó -0,5 puntos porcentual en referencia al anterior año, pero como tal la población en edad para trabajar PET es 3,2 millones de guayasenses en perfil para seguir laborando, pero la población económicamente activa PEA es de 1,9 millones aquellos que cuentan con un empleo y la población económicamente inactiva PEI 1,3 millones sin conseguir algún empleo digno. Lo que significa que, el contexto profesional en Guayaquil, con un índice de inactividad laboral que, si bien ha reducido levemente, continúa siendo alta y con una diferencia relevantes entre la PET y la PEA, tiene un efecto inmediato e intenso en el impago de crédito y, por lo tanto, en la economía regional y nacional (INEC, 2024).

Las organizaciones proveedoras en Ecuador trabajan en un ambiente complicado, en el cual la restauración de portafolio se ve perjudicada por aspectos económicos y sociales, como el estado económico de los usuarios y las circunstancias de peligro en algunas zonas o sectores (Banco Central del Ecuador, Información Económica, 2024). En la actualidad, COINSER es responsable de designar convenios para los métodos de pago de los clientes. El aumento irregular laboral, tiene un alto índice de interés, restringe la

destreza de pago de los ecuatorianos, la inseguridad en ciudades como Guayaquil, pone a los colaboradores de gestión de cobros a desafíos en barrios peligrosos (Consejo de Participación Ciudadana , 2024).

El fraude también tiene un rol importante en la administración de cobros, ya que puede provocar procedimientos incorrectos y perjudicar la claridad en las operaciones. (Raimundo Soto, 2003) En este contexto, es importante establecer instrumentos o recursos y estructuras que faciliten perfeccionar el manejo de deudas, disminuyendo los peligros e incrementando la restitución de cuentas.

Frente a esta situación, el presente plan unificador establece la ejecución de un esquema de análisis que, a través de la práctica de estrategias de clustering y cálculo probabilístico, posibilite clasificar a los clientes con deudas y anticipar su posibilidad de cumplimiento, es decir de pago. Este trabajo busca la recaudación efectiva en COINSER, al identificar a los clientes con incumplimiento en pagos y establecer propuestas de gestión individualizada de cobros. También que este trabajo sea útil para próximas investigaciones en la administración de riesgos financieros en empresas.

1.1 Descripción del Problema

La empresa INTERAGUA es responsable de suministrar el servicio de agua potable al 90% de los habitantes del cantón guayaquil (Interagua, Veolia, 2022). Por lo tanto, al tener más de 2,62 millones de clientes por la prestación del servicio potable; la empresa requiere proveedores para que sean encargados del proceso de cobranza, por este motivo, acude a trabajar juntamente con contratistas, empresas que laboran al 100% para la recuperación de la cartera vencida.

El presente proyecto se enfoca en las actividades de la Contratista COINSER SAS la cual gestiona la recuperación de la cartera morosa a partir de los 5 meses de deuda. Es importante recalcar que la fuente de ingresos de la contratista proviene de cada reconexión del servicio ejecutada, esto implica que previamente se tuvo que haber realizado toda la gestión necesaria para que el usuario regularice su deuda de tal forma se genera la reconexión de su servicio, finalmente una cuadrilla de la contratista asiste al predio y deja el servicio en estado activo. Actualmente se evidencia que se realizan gestiones que no tienen ningún efecto en la recuperación de cartera, como el de call center, como tal al principio estuvo considerado como una excelente herramienta para poder contactarse con los clientes específicamente se referenciaba a los barrios que no se puede entrar debido a la inseguridad, pero es importante mencionar que no se cuenta con datos informativos del usuario y esto impide abordar la misión, por lo tanto, se desperdician los recursos disponibles y al ser un volumen elevado de usuarios implica que no se podrá realizar la gestión de cobranza para todos los usuario o predios que se requiere durante el periodo mensual, es por este motivo que es necesario crear herramientas que permitan focalizar las operaciones en aquellos usuarios que tengan mayor probabilidades de realizar un convenio de pago o saldar la deuda completa.

La investigación busca proponer soluciones más eficientes y basadas en datos históricos haciendo uso de modelos matemáticos y estadísticos para la correcta planificación de las operaciones, las cuales pueden ser: gestión de call center, gestión de visita al predio, cortes y cortes “duros” al momento en que se evidencia una reconexión fraudulenta o clandestina por parte del usuario. Todas estas actividades involucran recursos humanos por ende monetarios, es así como nace la necesidad de optimizar y

focalizar las actividades con una correcta planificación a fin de generar la mayor cantidad de reconexiones que son importante para la contratista rendir objetivamente a la empresa Interagua.

1.2 Justificación del Problema

Los proveedores como COINSER SAS buscan optimizar el rendimiento de la gestión de pagos, se dedica en la restitución de cuentas morosas y la maximización para ejecutar tareas que se salen de control a organizaciones como Interagua. Se ha percibido que los trámites ejecutados a veces carecen de eficiencia, ocasiona una gestión ineficiente de los recursos accesibles y, como resultado, reduce el potencial de atender correctamente a todos los clientes que necesiten cuidado. Es relevante por el alto número de personas y el requerimiento de potenciar el impacto de cada participación reducida de parte del outsourcing.

La investigación tiene como objetivo desarrollar herramientas y estrategias que permitan focalizar las operaciones de cobranza en aquellos usuarios con mayor probabilidad de regularizar su situación financiera. Al hacer uso de modelos matemáticos y estadísticos, se busca optimizar la planificación de actividades como la gestión de call center's, visitas a predios y acciones de corte, especialmente en casos de reconexiones fraudulentas. Este ajuste no solo facilitará una óptima distribución de recursos humanos y financieros, sino que también impulsará a incrementar el número de reanudaciones funcionales y, por lo tanto, a optimizar la viabilidad financiera de INTERAGUA.

Este estudio reside en su capacidad para modificar y renovar un procedimiento, hoy en día ineficiente en uno más eficiente y sustentado en datos. Esto no solo favorecerá

a INTERAGUA y a sus contratistas, sino también a los clientes y consumidores, quienes obtendrán una asistencia más productiva y equitativa en la administración de sus deudas.

1.2.1 Población en estudio

Este análisis se enfocará en la ciudad de Guayaquil, Ecuador, con una perspectiva en los consumidores del servicio de agua potable suministrado por INTERAGUA, particularmente aquellos con cuentas por cobrar vencidas superiores a 6 meses administrada por el proveedor COINSER SAS.

Guayaquil se encuentra segmentada en ciclos, como se muestra en la siguiente tabla:

Tabla 1*Segmentación de ciclos por productos y deuda de los usuarios.*

| Etiquetas de fila | Cuenta de PRODUCTO | Suma de PRODUCTO2 | Suma de ADEUDADO | Suma de ADEUDADO2 |
|-------------------|--------------------|-------------------|------------------------|-------------------|
| I | 76 | 1,25% | \$ 151.085,24 | 2,03% |
| N | 63 | 0,66% | \$ 208.857,72 | 2,80% |
| Y | 49 | 0,97% | \$ 163.171,96 | 2,19% |
| X | 118 | 2,83% | \$ 127.960,51 | 1,72% |
| R | 163 | 3,30% | \$ 223.806,27 | 3,00% |
| T3 | 212 | 8,39% | \$ 157.788,56 | 2,11% |
| S | 177 | 5,00% | \$ 200.105,85 | 2,68% |
| H | 117 | 1,96% | \$ 220.203,95 | 2,95% |
| O | 19 | 0,34% | \$ 41.347,59 | 0,55% |
| K | 245 | 5,15% | \$ 339.615,13 | 4,55% |
| T | 465 | 15,64% | \$ 487.197,97 | 6,53% |
| E | 100 | 1,54% | \$ 301.522,08 | 4,04% |
| V1 | 87 | 3,68% | \$ 72.771,15 | 0,98% |
| B | 291 | 4,74% | \$ 401.104,06 | 5,38% |
| R2 | 230 | 4,43% | \$ 587.750,43 | 7,88% |
| M | 63 | 1,31% | \$ 116.917,33 | 1,57% |
| Q | 346 | 9,74% | \$ 500.310,70 | 6,71% |
| P | 106 | 2,60% | \$ 282.439,59 | 3,79% |
| U | 461 | 10,07% | \$ 673.125,20 | 9,02% |
| G | 26 | 0,41% | \$ 20.078,86 | 0,27% |
| T2 | 71 | 2,33% | \$ 21.458,32 | 0,29% |
| J | 101 | 3,04% | \$ 245.623,69 | 3,29% |
| V | 19 | 0,74% | \$ 10.856,52 | 0,15% |
| F | 89 | 1,56% | \$ 493.081,34 | 6,61% |
| D | 314 | 5,56% | \$ 789.345,83 | 10,58% |
| A | 75 | 1,10% | \$ 253.379,16 | 3,40% |
| C | 184 | 1,67% | \$ 369.558,60 | 4,95% |
| Total, general | 4267 | 100,00% | \$ 7.460.463,61 | 100,00% |

*Fuente: Contratista COINSER SAS**Elaborada por: Autores de este proyecto*

La base proporcionada por INTERAGUA a la contratista COINSER SAS es de un total de deuda por \$ 6.885.844,25 la cual se encuentra segmentada en 27 ciclos con un total de 3753 productos, es decir, medidores con deuda mayores a 6 meses.

Cada medidor tiene una cantidad de planillas adeudas que se encuentran distribuidas en rangos de la siguiente manera:

Tabla 2

Cantidades de planillas vencidas por número de factura

| Rangos | Cuenta de PRODUCTO |
|----------------------|---------------------------|
| >= 81 fact | 63 |
| 11 a 15 fact | 30 |
| 16 a 24 fact | 82 |
| 25 a 40 fact | 79 |
| 41 a 50 fact | 60 |
| 51 a 60 fact | 3852 |
| 6 a 10 fact | 69 |
| 61 a 80 fact | 32 |
| Total general | 4267 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

Nota: Del total de 4267 medidores con deuda gestionados por la contratista COINSER, 63 tienen una cantidad de planillas vencidas mayores a 81 facturas.

Este estudio está restringida al área de servicios básicos, específicamente en la administración de recuperación de deuda de los servicios de agua potable. La perspectiva se centra en el estudio de enfoques implementadas por COINSER SAS para restaurar cuantas vencidas.

Esta investigación tiene como objetivo recuperar cuentas vencidas, incrementar los restablecimientos y brindar recursos a organizaciones del sector, crear un efecto favorable en la estabilidad económica de COINSER SAS.

1.3 Objetivos

1.3.1 Objetivo general

Optimizar la gestión de cobranza trimestral de la cartera vencida en la contratista COINSER de Interagua, mediante la aplicación de técnicas de clústeres y análisis de probabilidad, con el fin de mejorar la efectividad en la recuperación de deudas según el perfil de los usuarios en el cantón de Guayaquil.

1.3.2 Objetivos específicos

1. Analizar los factores que inciden en la cartera de clientes mediante estadística descriptiva para visualizar la situación actual y posterior.
2. Realizar un modelo matemático de regresión para analizar las variables más influyentes para la generación de la reconexión del servicio.
3. Evaluar el impacto de las técnicas aplicadas en la efectividad de la recuperación de deudas.
4. Establecer los clústeres según la probabilidad de pago de los usuarios para optimizar la gestión de cobranza.

1.4 Marco teórico

1.4.1 Marco conceptual

El marco teórico del presente estudio abarca los principales factores relacionados con la gestión de la morosidad en INTERAGUA, la cual se enfrenta a desafíos significativos debido a la actual situación económica, social y operativa de Ecuador, con especial énfasis en el cantón Guayaquil. Este estudio tiene variables económicas, sociales, operativas y geográficas en un esquema completo apoyado por estudios locales y globales, y se enmarca en la normativa actual en el país.

La **morosidad** es la tardanza o infracción en pagos, es decir falta de compromiso económico en períodos estipulados, perjudicando a los clientes y a las organizaciones suministradoras del servicio. En el contexto de servicios públicos, el agua apta para el consumo, el impago no solo implica una problemática económica para las empresas contratistas, sino un problema social, ya que restringe la entrada de los usuarios a servicios primordiales y perjudica el nivel de vida (Palmer, 2013).

La morosidad representa eventos negativos que perjudican de manera directa la viabilidad financiera de una organización, específicamente en el sector de servicios públicos. Según el reporte del (Banco Central del Ecuador, Gestión de la Liquidez del Sistema Financiero No. 227, 2024), el aumento de las deudas atrasadas, que en febrero de 2024 alcanzó USD 3.262,8 millones, demuestra las dificultades que afrontan las organizaciones para solventar sus gastos administrativos y responsabilidades monetarias. Bajo estas circunstancias, las empresas de servicios públicos deben destinar mensualmente fondos o bienes relevantes para asegurar la permanencia de sus procesos, lo que dificulta aún más su habilidad o potencial de flexibilidad operativa frente al incumplimiento de pagos puntuales por parte de los clientes. El incumplimiento en los pagos afecta la solvencia rápida y aumentan los gastos de funcionamiento vinculados con el manejo de morosidad, dañando la viabilidad financiera a largo plazo y la posibilidad de estas empresas para financiar y optimizar sus servicios.

Según la (Asobanca, 2023) , al terminar diciembre de 2023, el sistema bancario ecuatoriano cerró en 3,2%, un aumento de 1 punto porcentual el año pasado, perjudicando especialmente a los sectores de consumo y microfinanzas con tasas de 4,8% y 6,6%. Lo que refleja la capacidad en pagos de empresas y hogares, perjudicando la estabilidad financiera. En sectores productivos y de servicios de las organizaciones, la creciente cartera improductiva alcanzó los USD 1.349 millones disminuyendo su acceso a financiamiento y aumentó los costos asociados a provisiones para la gestión de riesgos. Este fenómeno impacta la capacidad de inversión y crecimiento empresarial, agravando la incertidumbre económica en un Ecuador que enfrenta retos de liquidez y recuperación

financiera. Pues de tal manera, es importante conocer los factores que le dan flujo a la morosidad: económico, social y operativo.

Factores económicos

La recesión, el desempleo y la inflación actualmente afectan en gran significancia a la económica de los ecuatorianos, lo cual tiene como consecuencia que la población no pueda reconocer sus obligaciones de pago en el tiempo determinado por la empresa prestadora del servicio. La economía tiene un impacto en los niveles de empleo, refleja el flujo del mercado laboral. Durante periodos de crisis económica, las organizaciones enfrentan una reducción en sus ingresos, pocas ventas y una contracción de los mercados, las obliga a aplicar medidas de reducción de costos, despidos o congelaciones en la contratación, perseverando su viabilidad financiera. Por lo que, durante esos periodos, el incremento en la demanda de bienes y servicios impulsa la generación de empleo, ya que las empresas requieren mayor personal para atender las necesidades del mercado. Estos ciclos económicos muestran el índice de empleo y desempleo, posicionando a las fluctuaciones económicas como factor clave en los cambios de empleos (Generation Spain, 2024).

Según (Banco Central del Ecuador, Información Económica, 2024) en Ecuador, la morosidad tuvo un nivel alto en 2024, 5.3%, porcentaje mayor en los últimos cinco años. Este equivalente a un alza de 0.8 puntos porcentuales respecto a 2023, relacionado a aspectos económicos y sociales. El desempleo incrementa la morosidad, las personas sin ingresos estables tienen necesidades como la alimentación diaria y vivienda sobre el pago de servicios.

(INEC, 2024) menciona que el índice de desempleo se ubicó en 4% en agosto de 2024, aumento respecto al año anterior, refleja que la pobreza ha incrementado. Con 21,390 personas adicionales sin empleo en comparación con el año anterior, refleja un mercado laboral debilitado donde los ingresos irregulares se convierten en la norma, especialmente para los microempresarios y trabajadores informales. En Guayaquil, que concentra el 3.1% de la población económicamente activa (PEA) nacional, la situación es paradójica: aunque el desempleo disminuyó en un 0.5%, la alta informalidad laboral (57%) limita el acceso de las familias a créditos formales y dificulta el pago regular de servicios básicos como agua potable.

Factores sociales

En Ecuador, específicamente en Guayaquil donde el contexto socioeconómico es desafiante, la inseguridad y la falta de empleo son dos factores sociales clave que afectan los métodos de cobranza. La inseguridad ha alcanzado niveles alarmantes, actos delictivos que dificultan las operaciones en ciertas zonas (Andrea Jaramillo, 2024). Los cobradores se enfrentan a riesgos al ejecutar sus actividades en barrios peligrosos, por lo que, empresas como COINSER SAS, contratista de INTERAGUA, se ven obligadas a modificar sus operaciones, una de ellas puede ser brindar el servicio de call centers para no exponer a su personal. Pero puede que no siempre sea efectivo, ya que no permite una interacción directa con los clientes y limita las estrategias o métodos de negociación. La inseguridad genera costos adicionales para las empresas al tener que invertir en medidas de protección para sus trabajadores, lo que reduce sus márgenes operativos y compromete la estabilidad del servicio (Coinser SAS, 2024).

Según la Organización Internacional del Trabajo (OIT), la falta de trabajo es una problemática en América Latina y el Caribe. La estrategia FORLAC 2 por la OIT indica que las personas sin empleo cuentan con ingresos inestables, poca protección social y acceso limitado a beneficios financieros, por ende, no pueden tener créditos o cumplir con sus obligaciones de pagos. Sin un ingreso estable, las instituciones financieras y proveedoras de servicios básicos tienen a los trabajadores informales como un segmento de alto riesgo. Esta falta de crédito y redes de seguridad laboral perpetúa la vulnerabilidad económica de los individuos y del sistema financiero. La OIT destaca la necesidad de políticas públicas y estrategias integrales para la formalización laboral como un mecanismo clave para mejorar la viabilidad financiera y fortalecer la inclusión crediticia en la región (OIT, 2024).

En Ecuador, la inseguridad es una problemática alarmante, Guayaquil es una de las ciudades que se posiciona como un lugar afectado por la violencia y el crimen organizado. Según el Observatorio de Seguridad Ciudadana, en 2024 se reportaron más de 1,200 incidentes violentos en zonas de servicio de INTERAGUA, perjudicando a las organizaciones a realizar actividades como la cobranza en terreno. Estas áreas, obligan a los cobradores que tengan un acceso limitado a ciertos sectores, aplicando medidas de seguridad estrictas, como escoltas o la dependencia de los call centers, y a la larga generará costos operativos. Este entorno compromete la eficiencia de las estrategias de cobranza y repercute en la sostenibilidad financiera de las empresas proveedoras de servicios básicos (OECD, 2024).

Según (INEC, 2024), el 57% de la población guayaquileña trabaja en el sector considerado informal, lo que trae al país una ausencia de contratos estables, ingresos

regulares y acceso a beneficios laborales. Generando un nivel alto de precariedad financiera, limitando a las familias a planificar sus gastos y cumplir puntualmente con el pago de servicios básicos como el agua potable. Además, saldar deudas acumuladas y la morosidad como tal, afecta a los hogares y empresas concesionarias. La inseguridad e informalidad laboral en Guayaquil es una problemática en la economía local. Por lo que, se debe establecer estrategias integrales, que deben incluir políticas públicas enfocadas en el trabajo, creación de redes y medidas para proteger al personal y las operaciones empresariales en áreas de riesgo.

El sector de los servicios públicos y la recuperación de su cartera vencida se busca consigo estar vinculada al estado de suspensión del servicio, ya sea parcial o total, esto se evidencia en la normativa 142 de 1994 de Colombia, El estado de suspensión, es un factor crítico en el comportamiento de pago de los clientes. Considerando suspensión parcial, pues se restringe ciertos aspectos del servicio, puede servir como un recordatorio menos intrusivo, incentivando a los usuarios a regularizar su situación. Mientras que, la suspensión total, implica la interrupción completa del servicio, pues se ejerce una presión mayor y tiende a ser más efectiva en la recuperación de pagos atrasados, dando que se puede generar tensiones sociales y operativas; la clasificación de suspensión permite segmentar a los usuarios según su nivel de riesgo y diseñar estrategias diferenciadas de cobranza. Las alternativas correspondientes para la gestión eficiente se comprenderán por los usuarios morosos que tienen su suspensión del servicio de dos formas: “parcial” indicados para ofrecerle una negociación para regularizar su deuda o gestión de convenio siempre cuando sean firmados en base a contratos. En el caso de que sea “total” se podría presentar medidas más dedicadas al usuario y estrictas. (Funcionpublica, 2024).

El Municipio de Quevedo por medio de la **Empresa Pública de Agua Potable y Alcantarillado (EPMAPAQ)** presenta dificultades con la cartera vencida, viabilidad financiera y calidad del servicio. Hace dos años, se libera un informe basado en procedimientos técnicos y análisis del sistema municipal, en donde establecieron un prototipo de mejora, incluso buscaron estrategias para contabilizar el agua no considerada en la ciudad, a partir de esto, se buscó optimar la infraestructura y los sistemas de operación afianzados al entorno medioambiental y las afectaciones financieras que trae consigo (Ministerio del Ambiente y Agua EPMAPAQ, 2021).

En el cantón Durán, la **Empresa Municipal de Agua Potable y Alcantarillado (EMAPAD-EP)** tiene un elevado índice en el incumplimiento de planillas, ocasionando pérdidas millonarias como el mantenimiento de los sistemas de distribución de agua potable, esto compromete la viabilidad financiera de la empresa y su habilidad para satisfacer las necesidades básicas de las personas. La EMAPAD-EP, a través de la **Resolución Administrativa No. -EMAPAD-EP-GG-2024-002**, aprobó el **Plan Anual de Contrataciones (PAC) 2024**, prioriza propuestas para mejorar la gestión financiera y operativa. En estas estrategias está la planificación de contrataciones para obras de infraestructura, adquisición de tecnología para el control y monitoreo de la red de distribución, y el fortalecimiento de procesos administrativos mediante sistemas digitales para mejorar el método de cobros (EMAPAD-EP, 2024).

Adicionalmente, la empresa ha establecido medidas para la **reprogramación de deudas** para los clientes en mora, campañas de sensibilización y la mejora en la transparencia y comunicación con los usuarios. Estas buscan reducir la cartera vencida y fomentar una cultura de responsabilidad entre los usuarios. Por lo que, se ha enfocado en

la sostenibilidad del servicio mediante la planificación y la priorización de proyectos para el sistema. De esta manera, la EMAPAD-EP se esfuerza por asegurar que los recursos sean aplicados de manera estratégica y mantener la continuidad y calidad del suministro de agua potable en Durán, siguiendo principios de eficiencia y normativas nacionales (EMAPAD-EP, 2024).

Del mismo modo en otra empresa pública, Interagua en Guayaquil, representa un desafío significativo para su estabilidad financiera y operativa. Al cierre de 2020, Interagua, con aproximadamente 56.000 usuarios, registró una facturación de 163,8 millones de dólares, de los cuales 16 millones correspondían a cartera vencida (Interagua, Home page, 2013).

Para abordar este desafío, Interagua aplicó estrategias de cobranza, como:

Convenios de pago: regularizaciones en sus deudas a los usuarios, mediante acuerdos según sus capacidades económicas.

Canales de atención en línea: aplicación móvil y Suite Virtual, los clientes podrán consultar saldos, facturas y ejecutar pagos, agilizará el método de cobranza y comunicación.

Gestión comunitaria: Interagua ofrece facilidades de pago adaptadas a las necesidades socioeconómicas de cada comunidad (Interagua, Home page, 2013).

La Corporación Nacional de Electricidad (CNEL EP) en la ciudad de Guayaquil, necesita desarrollar un **plan de mejora de crédito y cobranza** para optimizar la calidad del servicio eléctrico. En los sectores marginales de la ciudad de guayaquil: Socio Vivienda, Nueva Prosperina y Guerreros del Fortín, existe una alta demanda de morosidad debido a que es imposible que las cuadrillas encargadas de cobranza no

pueden entrar a estos barrios. El plan de medidas como el uso de tecnologías para el seguimiento de deudas, capacitaciones continuas al personal, procesos de corte y reconexión del servicio eléctrico en casos de incumplimiento. La propuesta espera optimizar la percepción del servicio. Al implementar estas estrategias, CNEL EP pretende reducir índices de morosidad y garantizar una cultura de pago responsable para la parte operativa y financiera de la empresa, alineándose a los principios y calidad del servicio (Brachfield, 2008).

(Mogollón Gómez, J, 2021) , indica que hay una correlación entre una gestión de cobranza y resultados financieros de las organizaciones. Este informe mostró procesos incorrectos en la gestión de cobranza, como la falta de seguimiento a cuentas vencidas y la aplicación de políticas de cobro. Se utilizó el coeficiente de correlación de Spearman para demostrar que la gestión de cobranza impacta en la parte financiera al mejorar el capital de trabajo y recuperar cuentas por cobrar. Conocer también, las estrategias sé que verificaran en aplicaciones de avisos al vencimiento, negociaciones con los clientes, uso de tecnologías, diseño de políticas de cobranza y la capacitación continua del personal para evitar deudas morosas.

Según (Brachfield, 2008), considera que las herramientas más eficientes de cobranza son:

Llamadas telefónicas. – es el acto de mantener un diálogo directo, facilita resolver dudas de las personas, negociar plazos e indicar al cliente en cumplir con sus obligaciones.

Mensajes de texto: notificación a los clientes por SMS y se considera que es una manera económica de cómo se notificar a los usuarios sobre sus deudas o temas de pagos, asegurar que la información llegue al destinatario de manera clara.

Correo electrónico: aplicado para enviar notificaciones más detalladas a los usuarios como facturas o términos de refinanciamiento, tal vez es menos inmediato que los mensajes de texto, pero es valiosa en temas formales.

Visitas personales: útiles para localizar a clientes que se esconden o no disponen de medios tecnológicos, permiten establecer un contacto directo que puede fortalecer la relación entre la empresa y el cliente.

Cartas: documento que debe estar redactado de manera clara y persuasiva a los sobre ponerse al día con sus obligaciones financieras, manteniendo un tono para cuidar la buena relación con el usuario.

Se enfatiza en cómo se caracteriza el perfil de un usuario moroso considerando siempre su naturaleza deudora y el ambiente social. Es decir, en barrios con acceso limitado a tecnología, las cartas pueden ser más efectivas, por lo que, en áreas urbanas con mayor conectividad, los mensajes de texto y las llamadas telefónicas son preferibles. Este enfoque optimiza carteras vencidas y una buena comunicación con los clientes, ayudándoles con sus estados de cuenta (Brachfield, 2008).

La aplicación de **call centers** en el desarrollo de cobranza es un elemento para la recuperación de deudas y las operaciones de las organizaciones. El estudio "Análisis de la Productividad en un Call Center de Cobranza", las tecnologías como el **marcador predictivo** son aquellas que van a dar facilidad para reducir el tiempo de inactividad de los agentes al priorizar las llamadas con éxito. Esta tecnología aumenta la productividad

de los agentes hasta un 75%, se traduce en una mejora del 48% en la operativa general del call center (Ruiz Fúnez, 2014).

El marcador predictivo en sectores como el financiero, la recuperación de cartera vencida es crítica para mantener la rentabilidad. Un plan estratégico correcto, basado en el análisis de datos históricos y la segmentación de clientes, es crucial para los beneficios de estas herramientas tecnológicas (Ruiz Fúnez, 2014).

Las **visitas presenciales** son importantes cuando las llamadas telefónicas o correos electrónicos no han sido efectivas o incluso el contacto registrado no es el correcto, también permite al cobrador obtener una visión clara sobre la situación del usuario, sus condiciones reales en su entorno económico. Se evalúa que tan bueno es realizar visitas a los deudores, esta gestión se conoce como una oportunidad para negociar de forma personalizada, fortalecer relaciones comerciales y evaluar la capacidad de pago del deudor. Dentro de esta página web en donde se publica artículos relacionados al tema, se conocer que existe un modelo que es más costoso en comparación con herramientas digitales, del cual resultan indispensable en casos complicados, ya que facilita la detección de posible es insolvencias y permite diseñar soluciones adaptadas a las necesidades del cliente. Un recurso importante para llevar una gestión de cobranza efectiva es dar la visita a los usuarios en donde ellos toman comunicación con el encargo de la empresa, en donde le da a conocer los métodos que tiene la compañía para que se regularice o incluso adquiera convenios en base a las políticas, siendo este una clave para la sostenibilidad económica de cualquier empresa (Campos, R., 2017).

La **suspensión o corte del servicio de agua potable** es una estrategia clave dentro de la gestión de cobranza, para los escenarios de incumplimiento prolongado por

parte de los usuarios. Se dice que en el Estudio Previo de la Empresa Pública Aguas de Manta (EPAM), esta medida no solo busca mitigar las pérdidas financieras derivadas de la morosidad, sino también fomentar una cultura de pago responsable entre los clientes. Los principios éticos y operativos dan validez a esta implementación donde se evidencia una conversación pertinente con el usuario dando descuentos en la facilidad de pago, siempre cuando se cumpla con las políticas aplicadas por le empresa (Municipio de Manta, 2021).

Para que la gestión de cobranza sea exitosa depende de un enfoque equilibrado y adaptativo. Se debe utilizar el corte del servicio como última medida tras agotar alternativas como gestión de centro de atención telefónica o call center para convenios de pago, visitas al predio y recordatorios. Por eso las empresas se aseguran de que el procedimiento cumpla con las regulaciones legales y sea comunicado de manera ética para preservar la relación con el cliente (Calle, Tania, 2021).

El historial del pago es considerado uno de los factores más contundentes para los prestamistas, de tal forma por medio de esto evalúan la solvencia crediticia del usuario y la capacidad económica analizando el aérea laboral; es importante que toda prestamista pone sus políticas donde el usuario da fe de su colaboración a tiempo del pago, en algunos casos empresas prestadoras de servicios brindan con el compromiso de cancelar su deuda a mes vencido, en el caso de que no se cumpla existe una tarifa por mora, reconexión o tarifa de interés. Por tal motivo, los pagos atrasados pueden reducir el puntaje crediticio y estos aparecen en el historial hasta siete años. El informe crediticio es un documento donde se muestra el historial incluyendo cuentas de: saldos, pagos,

consultas y registros públicos (FasterCapital, Historial de pagos como afecta su calificación de capacidad de pago y como mejorarlo, 2024).

El registro de los usuarios en el Servicio de Rentas Internas (SRI) es una herramienta clave en la gestión de recuperación de cartera morosa, ya que permite segmentar a los clientes según su naturaleza jurídica (persona natural o jurídica). El estudio contenido en el repositorio de la (Espinosa Jarrin, Kevin Enrique, 2020) presenta conocimientos y evalúa a cada perfil tributario de los usuarios facilita el diseño de estrategias personalizadas para abordar la morosidad, debido a que las empresas pueden identificar patrones de pago y detectar riesgos asociados con actividades económicas específicas. Dentro de la clasificación según el RUC en el SRI existen personas naturales: usuarios que manejan sus propios ingresos y son considerados bajos, incluso conocen cada movimiento que debe tener su sueldo para el pago de sus deudas; sociedades: empresas relacionadas al ámbito económico y dan soporte al financiamiento de la economía de país, en base a los pagos se consideran instituciones que necesitan siempre regularizar sus deudas en el tiempo determinado; y los que no están registrados bajo un RUC: personas de libre comercio, no tienen ligado un establecimiento de comercio, siendo estos los más desafiantes para cancelar deudas pendientes; esta categorización es debido es para que no exista una interrelación entre los estados financieros de los usuarios (Espinosa Jarrin, Kevin Enrique, 2020).

(Calvo, O., & Aurelio, C. D., 2002), dentro de este artículo de la revista española, nos menciona que la regresión logística viene del concepto estadístico, siendo utilizada para conocer la relación entre una variable dependiente binaria y múltiples variables independientes (estas siendo cuantitativas o cualitativas). El artículo destaca que el

tamaño de la muestra es un factor crítico en la precisión y validez de los modelos de regresión logística, pues un tamaño insuficiente puede llevar a estimaciones sesgadas y pérdida de poder estadístico. De tal manera, se aborda la importancia de garantizar una relación adecuada entre los "eventos por variable independiente" para evitar problemas de sobreajuste y maximizar la capacidad predictiva del modelo.

Según (Knüsel, L. , 2007) , el documento explica que el cálculo del tamaño mínimo de muestra para estimar parámetros binomiales con márgenes de error y niveles de confianza precisos ha sido tradicionalmente considerado complejo debido a las infinitas posibilidades de evaluación de probabilidad. Sin embargo, se demuestra que para determinar exactamente el tamaño de muestra se requiere métodos computacionales modernos. Las técnicas como la evaluación de probabilidades discretas en lugar de aproximaciones clásicas basadas en el teorema del límite central. Los parámetros para establecer el tamaño de muestra incluyen el margen de error absoluto o relativo, donde se evalúa la cobertura de probabilidades en función del intervalo binomial. Pues se garantiza una mayor exactitud y menor conservadurismo, más eficiente en experimentos estadísticos, como los aplicados en estudios de regresión logística para la toma de decisiones informadas.

La **segmentación conductual en la cobranza**, en CR Software exponen que, por este medio se permite identificar patrones específicos en base al comportamiento de los clientes para desarrollar planificaciones de recuperación más efectivas. Se utilizan datos históricos y conductuales para clasificar a los usuarios en actitudes hacia el cumplimiento de pagos, como su frecuencia de morosidad, reacción a recordatorios y disposición a negociar. Esta técnica es fundamental reconocer la eficacia de la cobranza al personalizar

los métodos en base a las necesidades de cada grupo, sino que también fortalece la relación con el cliente al implementar soluciones más empáticas y adaptadas a cada usuario. La segmentación es esencial para entender cómo los factores como la ubicación geográfica, la inseguridad y la informalidad laboral influyen en el comportamiento de pago, permitiendo de forma correcta priorizar acciones y maximizar la recuperación de cartera vencida en un entorno de alta complejidad social y económica (Paneda, B., 2023).

La agrupación de riesgo crediticio utilizando algoritmos como **K-Means** y **agrupación jerárquica** es una metodología avanzada que permite distribuir a los clientes en grupos homogéneos con características compartidas relacionadas con el riesgo crediticio. **K-Means es algoritmo**, el método más recomendado para el razonamiento de datos, inicia con la selección del número de grupos (k) que se desea identificar. Es decir, cada cliente es representado por un conjunto de variables, que son relevantes en el modelo predilecto, en donde se formaran K conglomerados. Los centroides son generados aleatoriamente, en donde se evidencia que sirven como puntos de referencia para cada clúster dando referencia la representación por figuras de cada uno. Después se conoce que cada usuario tendrá una asignación de clúster en base a los centroides (siendo el índice referencial de cada conglomerado) y se encuentra más cercano según una métrica de distancia, como la euclidiana. En función que el algoritmo avanza, recalcula los centroides basándose en los promedios de los puntos asignados a cada clúster. Se muestra finalmente, un conjunto de conglomerados que representan la separación o dispersión en la que se encuentran los datos y por tanto se lo considera como un claro ejemplo del resultado de riesgo crediticio. Es decir, los clústeres en el caso del proceso a realizar por el modelo de resultado se podrían clasificar usuarios con valor alto de

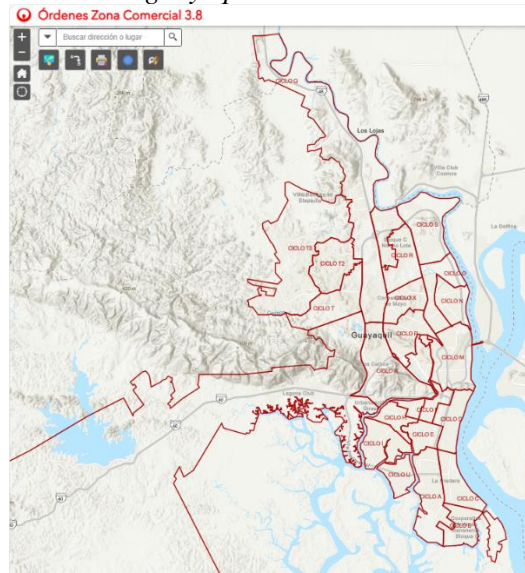
cancelar, riesgo del sector o número de facturas adeudadas. Tal acontecimiento, ayuda a las empresas a la toma de decisiones en base a factores claves existentes en la población y desarrollar estrategias para cada grupo de conglomerado. Es importante mencionar que, la **agrupación jerárquica** complementa este análisis al proporcionar una estructura más visual y escalonada. Es decir, fijar un número de clústeres al inicio, este método crea un dagrama que ilustra cómo los datos se agrupan progresivamente desde los puntos individuales hasta un solo clúster que abarca toda la población. Se considera especialmente útil para analizar relaciones entre grupos y ajustar estrategias basándose en niveles específicos de riesgo (FasterCapital, Agrupacion de riesgo crediticio como agrupar el riesgo crediticio utilizando K Means y agrupacion jerarquica, 2024).

Las **variables geográficas** dan un impacto significativo en la gestión de la cobranza y en la incidencia de la morosidad en los servicios básicos, en particular en una ciudad como Guayaquil, cuya estructura territorial y socioeconómica es diversa y desigual. Las variables permiten analizar y segmentar a los usuarios según su ubicación, características del entorno y acceso a servicios, ofreciendo información clave para diseñar estrategias de cobranza más efectivas (Marcela Guachamín, 2024).

Es importante mostrar la parte segmentada-distribuida de Interagua en guayaquil, debido a que se enfatizara esta distribución por cada ciclo con la probabilidad de pago.

Figura 1

Segmentación de ciclos dentro de guayaquil



Fuente: Ordenes Zona Comercial 3.8

ArcGIS Web Application

Elaborado por: Interagua

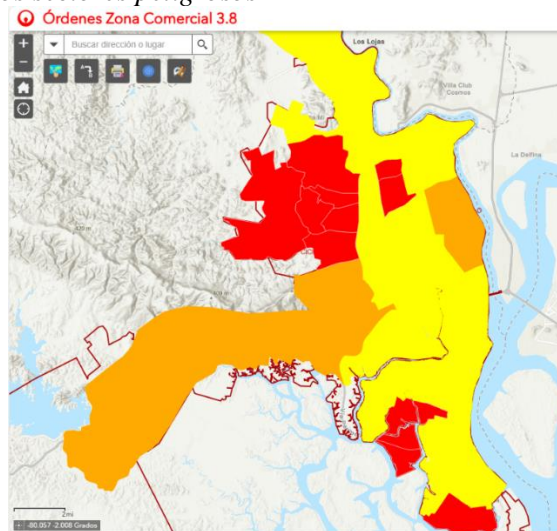
Geográficamente esta división como se muestra en la imagen, el alcance que tiene el servicio brindado por Interagua, considerando que es una estrategia para analizar y ejecutar mejores desafíos para la gestión de servicios básicos, en el contexto de la recuperación de cartera vencida. Guayaquil tiene gran parte territorial significativa, con zonas urbanas muy pobladas, áreas verdes con poco acceso y sectores rurales como Tenguel y Posorja. Las diferencias territoriales impactan de forma directa en la capacidad de las empresas, como INTERAGUA y su contratista COINSER SAS, las estrategias de cobranza efectivas, especialmente en un entorno caracterizado por inseguridad y desigualdad socioeconómica (Interagua, Home page, 2013).

Conociendo mejor la parte de distribución, podemos afirmar, que los ciclos K y M, cuentan con condiciones socioeconómicas buenas, mayor infraestructura y menor riesgo de inseguridad. Facilitando que las cuadrillas encargadas puedan dar visita al predio con total normalidad sin tener el factor radical de la inseguridad y realizar sus

actividades de cobranza, lo cual reduce los niveles de morosidad. También consideramos, las zonas verdes o campestres como los ciclos T, T2 o S enfrentan desafíos territoriales, debido a la propagación geográfica, el difícil acceso a las zonas, y los altos niveles de inseguridad que dificultan la labor del equipo encargado de gestionar las rutas de visita, lo cual perjudica llegar a un alcance a sus usuarios y que ellos conozcan su irregularidad con el servicio de agua potable (Interagua, Home page, 2013).

En las zonas más vulnerables, prevalece los trabajadores informales, quienes representan el 57% de la población económicamente activa en Guayaquil; este ciclo atraviesa un déficit en sus ingresos inestables y una limitada capacidad de planificación financiera, lo que incrementa el riesgo de morosidad en los pagos. Es decir que, los altos índices de inseguridad en estas áreas obligan a las empresas a depender de estrategias de cobranza remota, como los call centers, las cuales no siempre resultan efectivas (Interagua, Home page, 2013).

Figura 2
Reconocimiento de los sectores peligrosos



Fuente: Ordenes Zona Comercial 3.8
ArcGIS Web Application
Elaborada por: Interagua

La imagen presentada, se visualiza la distribución que tiene Interagua en base a cada ciclo y clasifica según un color característico en representando el nivel de inseguridad en los ciclos comerciales de Guayaquil, los colores con su significado son: amarillo para zonas de baja peligrosidad, naranja para peligrosidad media, y rojo para alta peligrosidad; existe una relación entrelazada entre los colores que caracteriza la inseguridad de las zonas con los desafíos en la gestión de regularizar a los usuarios deudores del servicio básico. Los sitios marcados en rojo, que abarcan barrios críticos tanto en las afueras de la ciudad como en áreas urbanas, representan un desafío considerable para las empresas como INTERAGUA y su contratista COINSER SAS. Esta zona se dice que, los cobradores enfrentan altos riesgos al realizar visitas presenciales, lo que limita el acceso directo a los usuarios y obliga a depender de métodos de cobranza remota, como los call centers, que en muchas ocasiones son insuficientes para reducir la morosidad (Interagua, Home page, 2013).

La falta de seguridad en las zonas de cobranza de Guayaquil, especialmente en aquellas marcadas como de alta peligrosidad, se plantea un desafío crítico para las empresas concesionarias de servicios básicos como Interagua y la contratista Coinser SAS. Por lo cual, se requieren estrategias operativas diferenciadas, como la inversión en seguridad para el personal, la implementación de herramientas digitales y tecnologías predictivas que prioricen a los usuarios con mayor probabilidad de pago, y la dependencia de métodos remotos, como los call centers, para mitigar los riesgos asociados a la inseguridad. Por lo tanto, es indispensable conocer la normativa, pues el Código Orgánico de Organización Territorial, Autonomía y Descentralización, expone en su artículo 55 la responsabilidad de los Gobiernos Autónomos Descentralizados (GAD)

de respaldar la provisión equitativa de servicios básicos, incluso en contextos de alta vulnerabilidad, mientras que su artículo 137 permite diseñar políticas locales para reestructurar deudas o establecer tarifas diferenciadas para usuarios en condiciones críticas (COOTAD, 2019). Adicionalmente, la Ley Orgánica de Recursos Hídricos, Uso y Aprovechamiento del Agua (2014) refuerza este principio en su artículo 3, reconociendo el acceso al agua como un derecho humano importante y se exige que los servicios estén en zonas afectadas por inseguridad y desigualdad (LEY ORGANICA DEL AGUA, 2014).

Superintendencia de Servicios Públicos (SSP), es clave para ser regulador y supervisor especializado en garantizar la prestación de servicios básicos, la entidad busca equilibrar los intereses de los usuarios y que la empresa siga brindando servicios de calidad y avance. El (COOTAD, 2019) en su artículo 55, toma competencias exclusivas en base a los gobiernos autónomos y da veracidad sobre la clasificación existente de predios y servicios públicos, el literal d hace referencia a la prestación de servicios públicos de agua potable y demás servicios; finalmente se crea tarifarios diferenciado para cada predio y a partir de esto se genera que por cada usuario la probabilidad de su pago es distinta y clave en función al modelo de regresión.

La regresión logística permite clasificar a los deudores según la probabilidad de pagar, en donde los gestores van a priorizar esfuerzos en aquellos casos con mayor posibilidad de éxito. Estos modelos logísticos permiten que los gestores de cobranza dirijan sus recursos hacia las cuentas con mayor probabilidad de ser recuperadas, es decir que se podrá ir mejorando la eficiencia y reduciendo costos operativos (Thomas, L. C., Crook, J. N., & Edelman, 2002).

Estos modelos predilectos ayudan a segmentar la cartera de clientes en categorías de riesgo, lo que permite diseñar estrategias específicas según el perfil de cada deudor. La regresión logística es una herramienta fundamental en la ejecución de estrategias de cobranza, al identificar diferentes segmentos de clientes según sus comportamientos de pago y características individuales (Anderson, R., 2007), "The Credit Scoring Toolkit".

Del mismo modo, permite evaluar la influencia de variables clave (ingresos, historial de pagos, antigüedad de la deuda, entre otros) en el comportamiento del cliente. El impacto que se evidencia en las regresiones se da por múltiples factores simultáneamente es crucial en el análisis de cobranzas, permitiendo a los gestores entender qué variables afectan más el cumplimiento. (Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X., 2013), "Applied Logistic Regression".

Finalmente, después de una investigación exhaustiva, la regresión logística es una herramienta poderosa en la gestión de cobranzas, no solo por su capacidad predictiva, sino también por su contribución a la eficiencia operativa y la personalización de estrategias. La **regresión logística** ha influenciado significativamente en la gestión de cobranzas de desarrollo para la capacidad para modelar probabilidades de eventos binarios, como el cumplimiento o incumplimiento de pagos, optimizando estrategias de recuperación de deudas; y su implementación permite a las empresas de cobranza maximizar los recursos y recuperar mayores volúmenes de deuda de manera más efectiva (Dolores, & Rodríguez Pérez, J. C., 2000).

Capítulo 2

2 Metodología

Este capítulo se podrá describir en detalle la metodología utilizada para abordar el problema de la morosidad en la gestión de cobranzas de INTERAGUA, la cual es operada en parte por la contratista COINSER SAS. Se desarrolla basado en técnicas estadísticas avanzadas, que incluyó el diseño de un modelo predictivo de regresión logística y una estrategia de segmentación por clústeres para optimizar la recuperación de la cartera vencida. Se presenta las alternativas de solución que este caso son los clústeres, de los cuales se selecciona la más adecuada de acuerdo con los datos, la selección de la mejor opción, el diseño conceptual y el detalle del diseño implementado, así como las especificaciones técnicas del modelo donde se evidenció la distribución de los ciclos estratificados en el cantón Guayaquil. Además, se exponen las que las convalidaciones del resultado sean éticas y legales necesarias para garantizar la validez y aplicabilidad de los resultados.

2.1 Alternativas de Solución

Existen varios problemas identificados por la contratista COINSER SAS, en lo cual se evaluaron tres alternativas basadas en atacar de raíz este inconveniente, basadas en la factibilidad técnica, operativa y económica.

1. *Incremento del personal de cobranza:*

➤ Se consideró aumentar el personal de call center y visitas domiciliarias con el fin de abarcar la gestión de cobranza a todos los ciclos distribuidos dentro de Guayaquil. Aunque la alternativa es viable representa altos costos operativos, logísticos e incluso riesgos relacionados con la inseguridad en ciertas zonas-barrios de Guayaquil y es de conocimiento que actualmente en algunos sectores es imposible entrar debido a las organizaciones delictivas.

2. *Implementación de un sistema de clasificación simple:*

➤ Este sistema tuvo un enfoque en donde se utilizaba criterios básicos, como el monto de la deuda y el tiempo de vencimiento con que por medio de estos se clasificara para a esos usuarios con mayor deuda se pueda dar prioridad en la gestión de cobranza. Aunque esta alternativa sea más económica, como tal no integraba un análisis multifactorial ni consideraba variables contextuales que influyen en la probabilidad de pago.

2.2 Desarrollo de un modelo predilecto:

➤ Alternativa que involucraba la construcción de un modelo de regresión logística con la finalidad de predecir la probabilidad del pago de los usuarios con deuda e implementaron técnicas de clustering para segmentar a los usuarios según su perfil de morosidad en donde se tuvo prioridad según el desarrollo del modelo.

Al conocer cada alternativa descrita, se terminó por aceptar que la más eficaz para la problemática es la tercera debido a la capacidad de poder integrar múltiples variables y proporcionar predicciones mejoradas las cuales ayudaron en la nueva gestión de cobranza. Por otro lado, se consideró importante esta alternativa debida que esta solución era escalable y podía replicarse en otros sectores de servicios básicos.

2.2.1 *Diseño Conceptual y Metodología Adoptada*

Por medio de un modelo predilecto de regresión logística se fue optimizando **la gestión de la cartera vencida** de INTERAGUA, gestionada por COINSER SAS. La contratista necesita priorizar a los usuarios con mayor probabilidad de cumplir con sus pagos, lo que permitirá diseñar estrategias de cobranza más efectivas y personalizadas.

2.2.2 Recolección de datos

Coinser S.A.S, nos entregó la data correspondiente al último trimestre del 2024, en donde por medio de esta se fue considerando el planteamiento de las variables representativas y que tenga peso para generar el modelo, como tal dentro de datos se encontraros información de buen alcance el nivel de interpretación fue un poca compleja, pero con las investigaciones previas se fue conociendo el modelo de base de datos brindado por el cliente, se gestionó un análisis para considerar factores altos y bajos para el abordamiento de los objetivos planteados. Inicialmente, se utilizó la base de datos proporcionada por INTERAGUA y gestionada por COINSER SAS del **cuarto trimestre del 2024**; la cual se evidenciaron registros históricos de los usuarios morosos. Las variables analizadas fueron:

- Características sociodemográficas (zona geográfica, tipo de predio).
- Expediente de pagos (# de facturas adeudadas, valor total por pagar).
- Términos contractuales (tipo de predio).
- Data demográfica (distribución de cada ciclo, dando énfasis al nivel de inseguridad).
- Gestión de cobro (centro de atención telefónica, visita al usuario, corte del servicio, contravención, reconexión).

2.2.3 Herramientas que se utilizaron

- *Excel*: utilizó esta herramienta para integrar y consolidar la base de datos que combina la información de clientes con todas las gestiones realizadas por la contratista, con el fin de obtener una visión completa y detallada de las acciones

ejecutadas y su peso sobre cómo se comporta el pago de los usuarios. La contratista es encargada de:

- **Gestiones de Centro De Atención Telefónica:** denominada también “call center”, es el monitoreo telefónico para recordatorios de pago o negociaciones.
- **Visitas al predio:** la realización de verificar de forma presencial el estado del servicio y conocer si está en cumplimiento del pago.
- **Cortes del servicio de agua:**
 - **Cortes sencillos:** Suspensiones parcial del servicio por mora, en donde se paga un fee por tal acontecimiento.
 - **Cortes duros con sanción:** Suspensión total, incluso se detectan fraudes y generan multas a los usuarios. Las (VCO), se registran cuando el usuario ha reconectado ilegalmente el servicio después de un corte.
- **Lecturas del consumo:** se basa en que, si existe diferencias entre las lecturas del mes anterior y del mes actual, con el fin de identificar diferencias significativas que puedan indicar reconexiones ilegales, consideradas como fraude.
- **Reconexiones del servicio:** predios que fueron reconectados, lo cual refleja que alguna gestión tuvo efecto en que el usuario realice un convenio o regularice su deuda. También se incluyeron casos de reconexión por otros motivos que se busca entender en el análisis.

En conclusión, se unieron las bases de datos de clientes y gestiones realizadas en un formato estructurado y formulado en Excel, adhiriendo información como: (call center, visitas al predio, cortes sencillos, cortes por contravención -VCO-, y lecturas de consumo), los resultados de las columnas son dadas en fechas, es decir cuando se realizó

la diligencia por la cuadrilla se registra a fecha como evidencia de la gestión. El análisis específico en este apartado es con el fin de conocer las variables que se interrelacionan directamente y el comportamiento que tiene cada usuario en base a aspectos fijos como la zona en que viven o monto adeudado; conocer la naturaleza de cada individuo va a permitir dar una respuesta positiva o negativa con la renegociación.

- *Excel Power Query*: herramienta necesitada para limpiar y organizar la base de datos proporcionada por INTERAGUA y gestionada por COINSER SAS. La base de datos se fue eliminando datos aberrantes de la siguiente forma: se aplicaron transformaciones como la eliminación de duplicados, normalización de valores, detección de inconsistencias, y estructuración de las variables clave como la ubicación geográfica, el número de facturas vencidas y el monto de deuda. Dando como conjunto datos uniformes u homogéneos, a evaluar para los siguientes procesos, hasta llegar al modelo predilecto.

Se transformó variables cualitativas en numéricas, facilitando así su inclusión en el modelo estadístico. Esta transformación consistió en asignar números únicos a las categorías específicas de las variables cualitativas, como el tipo de cliente, la región o el estado del caso de cobranza, asegurando una codificación adecuada para el análisis. El objetivo principal de esta codificación fue identificar y medir el impacto de dichas variables cualitativas en la probabilidad de recuperación de la cartera. Por ejemplo, se cuantifico variables como el "tipo de deudor" o el "estado de la deuda", el modelo logró evaluar con mayor precisión cuáles factores tenían mayor influencia en el comportamiento de pago de los clientes.

- *Power BI*: herramienta en donde se subió la base ya depurada y se hizo gráficas para la presentación de los datos procesados, se visualizó de forma clara y dinámica las tendencias de morosidad, los ciclos geográficos críticos y las categorizaciones de usuarios. Además, el proceso utilizado fue reconocer valores atípicos o inconsistencias que no fueron resueltas en la etapa de limpieza; por medio de estas técnicas empleadas se facilitaron la identificación de áreas problemáticas y ayudaron a priorizar las estrategias operativas, haciendo que las decisiones fueran basadas en información comprensible y bien estructurada. Después, se integró bases de datos del Servicio de Rentas Internas (SRI) Datasets - intersri - Servicio de Rentas Internas y del Instituto Nacional de Estadística y Censos (INEC), fueron descargadas en un repositorio de todas las provincias considerando que dentro de guayaquil existen habitantes que tienen un contrato con Interagua, pero su lugar de nacimiento fue en provincia. Para poder unir las bases de datos, simplemente al número de cedula registrada en la base se le añadió los numerales “001”, dando validez al formato entregado inicial. La principal importancia de esta unión de base de datos es conocer si el usuario esta registrado bajo una sociedad, persona natural o no tiene RUC; dentro de esta base de datos existen variables con denominación de tipo de contribuyente, # de establecimientos, actividad económica que realiza con la posibilidad de conocer si con aquella información tributaria el modelo toma una eficiencia; permitiendo identificar factores que influyen en su capacidad para pagar deudas.

- *Minitab*: es un software con múltiples desarrollos y en este caso, fue fundamental para el modelo predictivo enfocado en regresión logística, enfatizado a nuestros objetivos. Se identificaron las variables más relacionadas en la probabilidad de

pago de los usuarios. Permitiendo evaluar la significancia por el valor p-value, que basándonos en la teoría no debe ser mayor a 0,05; a partir de generar el modelo se dieron predicciones confiables sobre los factores clave al comportamiento de los deudores. Por lo tanto, se lograron los resultados previstos que sirvieron como base para segmentar a los usuarios en clústeres y diseñar estrategias de cobranza adaptadas a cada perfil.

- Finalmente, para que la base de datos pueda generar una regresión logística, en base a la variable dependiente y variables independientes se le fue dando la codificación dummy; es decir la variable dependiente binaria tomó dígito de 0 y 1; mientras que las variables independientes si son categóricas fueron codificadas en numerales convirtiendo columnas binarias (dummy) donde se dio por cada columna una categoría, la cual será explicada en el capítulo 3.

2.3 Población y muestra

El cálculo de la muestra se realizó aplicando criterios estadísticos fundamentales que garantizaron la representatividad de los datos y la precisión de los resultados obtenidos. A continuación, se detallan los criterios utilizados:

Población

La población del estudio estuvo conformada por la totalidad de usuarios correspondientes a los ciclos comerciales del 4to trimestre del año 2024 de la ciudad de Guayaquil gestionados por INTERAGUA a través de su contratista COINSER SAS. Esta población incluyó a todos los usuarios con cartera vencida identificados en los 27 ciclos mapeados, de los cuales presentaban diversas características en cuanto a niveles de inseguridad, condiciones económicas y geográficas. En total, los usuarios de Interagua

registrados al 2024 son más de 51 mil clientes dentro del alcance de la empresa. La base de datos del trimestre fue de 4.267 usuarios, por lo que hizo un barrido de información con el SRI donde se obtuvo una nueva data, permitiendo un análisis integral que abarcó variables como el monto de la deuda, número de facturas vencidas, ubicación geográfica, nivel de inseguridad en las zonas y las condiciones categóricas como el tipo de contribuyente.

Nivel de confianza

Se estableció un nivel de confianza del **95%**, lo que significó que los resultados obtenidos a partir de la muestra seleccionada reflejarían con un 95% de certeza los patrones presentes en la población total. Este nivel de confianza se seleccionó como estándar en estudios estadísticos, asegurando conclusiones sólidas y generalizables al resto de la población.

El nivel de confianza del **95%** fue seleccionado para el cálculo del tamaño de la muestra debido a que representa un equilibrio ideal entre precisión estadística y viabilidad práctica en la investigación. Este nivel de confianza asegura que, en promedio, 95 de cada 100 muestras tomadas incluirán el parámetro poblacional verdadero, lo que proporciona resultados suficientemente representativos. Es ampliamente utilizado en estudios sociales, económicos y operativos debido a su aceptación como estándar estadístico, ofreciendo un balance adecuado entre la rigurosidad científica y la eficiencia operativa. Optar por niveles más altos, como el 99%, implicaría un aumento significativo en el tamaño de la muestra y los costos asociados, mientras que un nivel más bajo, como el 90%, podría comprometer la precisión y confiabilidad de los resultados. La elección del 95% respalda la validez del modelo propuesto en este estudio, asegurando

conclusiones estadísticamente significativas y alineadas con las mejores prácticas metodológicas (Hernández Sampieri, Fernández Collado, Baptista, 2014).

Margen de error

Se configuró un margen de error del 5%, lo cual estableció el rango de incertidumbre aceptable entre los resultados obtenidos de la muestra y la realidad de la población total. Un margen de error del 5% permitió que las estimaciones derivadas del análisis variaran en $\pm 5\%$, lo que garantizó precisión sin requerir un tamaño de muestra excesivo.

Proporción esperada (P)

Se asumió una **proporción esperada del 50%**, reflejó la frecuencia con la que se espera que un evento ocurra dentro de la población estudiada. Es un parámetro crítico para el cálculo de la muestra, ya que su valor influyó directamente en la precisión del tamaño calculado. En una distribución binomial, el complemento de P, denotó como $q = 1 - p$, representa la probabilidad de que el evento no ocurra.

Cálculo del tamaño de la Muestra

El cálculo del tamaño de la muestra se realizó bajo el supuesto de que la población seguía una **distribución binomial**, que es apropiada para variables categóricas donde solo existen dos posibles resultados, como el cumplimiento o incumplimiento en el pago de las deudas. Este enfoque estadístico permitió determinar el número mínimo de observaciones necesarias para garantizar una representatividad adecuada, considerando los parámetros clave de confianza y error.

Fórmula utilizada

La fórmula para el cálculo del tamaño de la muestra basada en una distribución binomial es:

$$n = \frac{Z^2 \times p \times q}{e^2 + (N - 1)}$$

Donde:

e) n : *Tamaño de la muestra*

e) Z : *Valor correspondiente al nivel de confianza; para un nivel de confianza del 95%,*

$$Z = 1.96$$

e) p : *Proporción esperada de éxito en la población (en este caso, se asumió $p = 0.5$, al desconocerse su distribución exacta).*

e) q : *Complemento de p , definido como $q = 1 - p$ (también $q = 0.5$).*

e) e : *Margen de error aceptado, que se fijó en $e = 0.5\%$.*

Aplicación de la fórmula

Sustitución de valores:

1. $Z = 1.96$

2. $p = 0.5$

3. $q = 0.5$

4. $e = 0.05$

Cálculos intermedios:

1. $Z^2 = (1.96)^2 = 3.8416$

2. $p \cdot q = 0.5 \cdot 0.5 = 0.25$

3. $e^2 = (0.05)^2 = 0.0025$

Sustitución en la fórmula:

$$n = \frac{3.8416 \times 0.25}{0.0025}$$

Resultado:

$$n = \frac{0.9604}{0.0025 + (4267 - 1)} = 354$$

Por lo tanto, el tamaño mínimo de la muestra requerido para representar a la población fue de 354 usuarios.

El cálculo del tamaño de la muestra determinó que, con al menos **354 usuarios**, se podía garantizar una representatividad estadística adecuada para una población total de **4,267 usuarios**, con un nivel de confianza del **95%** y un margen de error del **5%**.

Finalmente, se usará la totalidad de la muestra registrada en la base de datos del trimestre, debido a que al correr el modelo con el tamaño de muestra sale que el modelo no se puede predecir debió a la falta de información. Esta muestra más grande fortaleció la precisión del análisis y la validez de los resultados, asegurando que todas las variables relevantes, como las características categorías, geográficas y operativas, fueran incluidas. Además, permitió generar inferencias robustas y confiables para la población total, lo que respaldó el diseño de estrategias de cobranza más efectivas y ajustadas al contexto específico de los usuarios con deudas.

Capítulo 3

3 Resultados y análisis

3.1 Análisis de las variables

En esta sección se presentaron los resultados obtenidos del análisis de las variables relacionadas en la ejecución del modelo predictivo en regresión logística con distribución binomial. La variable dependiente fue la probabilidad de que exista una reconexión, dando como finalidad la regularización del pago; definida como la respuesta que se tuvo por los usuarios en base a las gestiones de cobranzas realizadas por la contratista Coinser S.A.S, la variable fue establecida de manera categorizada y dicotómica en base al comportamiento teniendo relación con la fecha de reconexión de cada usuario:

- **Cumplimiento (valor = 1)**, son los usuarios que normalizaron su deuda y se verifica que el servicio se reconectó. Por tal motivo, si el usuario tiene una fecha establecida dentro de la base de datos el valor dummy que se estableció fue 1.
- **Incumplimiento (valor = 0)**, aquellos que no mostraron interés suficiente en regularizar su estado de morosidad. Por tal motivo, dentro de la base de datos no se reflejó que existía una fecha de reconexión y el valor dummy que se estableció fue 0.

Tabla 3
Codificación Dummy Reconexión

| VARIABLE DEPENDIENTE | |
|----------------------|-------------|
| RECONEXIÓN | VALOR DUMMY |
| 10/10/2024 | 1 |
| N/A | 0 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

La variable dependiente *reconexión* se aprobó como un análisis clave para el desarrollo eficiente del modelo predilecto de regresión logística, la validez de la variable se fundamenta en su naturaleza binaria, alineándose con los requisitos estadísticos, la cual

hace modelar la probabilidad del evento específico basado en múltiples factores entrelazados; a partir de esta variable se proporcionó una base sólida, consistente para ejecutar y evaluar un análisis con el impacto de las variables independientes, pues que será probado y aplicable en escenarios reales de cobranza.

Las variables independientes se especificaron en aspectos categorizados y numéricos que se reflejaron de la siguiente manera:

- Número de facturas adeudadas: se consideró una variable crítica-cuantitativa para representar la cantidad de facturas vencidas ya sea que fueron acumuladas por los usuarios durante varios meses. De tal forma, esta métrica proporciona un índice directo del nivel de deuda exacto por cada cliente moroso, dando paso a una de las primeras señales de incumplimiento financiero. Operativamente se indicó que a un mayor número de facturas adeudadas se tiene una mayor complejidad de pago, pues esto implica que el usuario tuvo un compromiso financiero más significativo dando valor a la variable dependiente del modelo al cumplimiento lento.

Tabla 4
Codificación de Facturas

| |
|---------------------------|
| VARIABLE INDEPENDIENTE |
| #FACTURAS |
| 6 |
| 334 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

La tabla muestra la cantidad de facturas mínima siendo 6 y el máximo 334, las cuales se van a reflejar en toda la matriz mapeada.

- Saldo pendiente: correspondió al monto total que el usuario adeuda, se consideró una variable económica más relevante para predecir el comportamiento de

pago en el modelo, pues el indicador reflejó la magnitud del compromiso financiero que el usuario tuvo con la empresa, dando como conclusión que se vio significativamente en la capacidad y disposición para regularizar las deudas dentro del tiempo pactado. De tal manera, se espera que los usuarios con saldos por pagar más bajos sean los que llegaron a cumplir más rápido, dando como lógica que a menor carga financiera más manejable en el pago de la deuda. Esta variable dentro del modelo predilecto permitió no solo enfocarse en el tiempo de pago, sino diseñar políticas acercadas al maximizar la recuperación de cartera y minimizar el riesgo financiero para la empresa.

Tabla 5
Codificación \$Adeudado

| VARIABLE INDEPENDIENTE | |
|------------------------|--------------|
| \$ ADEUDADO | LN(ADEUDADO) |
| 15,77 | 2,76 |
| 157.472,58 | 11,97 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

En la tabla se muestra el valor adeudado por usuario en el cual utilizó la herramienta de minitab para hacer el logaritmo natural (LN) de la variable saldo pendiente debido a que la distribución de los saldos mostraba una asimetría positiva significativa, esto quiere decir, había una gran cantidad de valores pequeños y una menor proporción de valores muy altos. Esta herramienta nos ayudó en reducir la asimetría de los datos, logrando una distribución más cercana a la normalidad. La capacidad del modelo para generar las relaciones entre las variables y permitió una interpretación más lineal y proporcional de los efectos del saldo pendiente sobre la probabilidad de regularización de la deuda.

- Rango de adeudo: se clasificó en 8 rangos en intervalos definidos por la cantidad de factura por pagar de cada cliente refiriéndose a que el producto (medidor) se

encuentra en un rango, esta variable fue considerada con un enfoque que permitió segmentar a los usuarios en grupos homogéneo con un análisis de comportamiento de pago en base al nivel de deuda. De tal forma, al incluir esta variable en el modelo predilecto se buscó optimizar las estrategias de cobranza y se mejora la capacidad de predicción.

Tabla 6
Codificación del Rango

| VARIABLE INDEPENDIENTE | |
|------------------------|-----------|
| RANGO | RANGO NUM |
| 6 a 10 fact | 1 |
| 11 a 15 fact | 2 |
| 16 a 24 fact | 3 |
| 25 a 40 fact | 4 |
| 41 a 50 fact | 5 |
| 51 a 60 fact | 6 |
| 61 a 80 fact | 7 |
| >= 81 fact | 8 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

En la tabla se muestra el rango facturado y la categoría relacionada que se fue dando valores para la base de datos en minitab.

- Ciclos: se identificó la ubicación geográfica de los usuarios en base a cada ciclo de facturación y fue relacionada con las condiciones socioeconómicas y de inseguridad por sector en todo guayaquil, vía la costa y parte de Posorja. Se consideró que esta variable es crucial en el análisis debido a que cada ciclo tiene una característica particular (en este caso se la definió en base al nivel de peligrosidad). La categorización se dio en base al nivel siendo este (Alto, Medio y Bajo) según el ciclo ubicado en el territorio mostrado en la imagen 2.

Tabla 7
Codificación de los ciclos

| VARIABLE INDEPENDIENTE | |
|------------------------|--------------------------------|
| CICLOS | CATEGORÍA NIVEL DE INSEGURIDAD |
| R | 3 |
| X | 2 |
| T3 | 3 |
| R2 | 3 |
| E | 3 |
| Q | 2 |
| U | 3 |
| K | 2 |
| M | 1 |
| S | 1 |
| T | 3 |
| H | 2 |
| V1 | 2 |
| B | 3 |
| D | 2 |
| O | 1 |
| A | 2 |
| G | 2 |
| T2 | 3 |
| J | 1 |
| C | 2 |
| P | 3 |
| I | 3 |
| F | 2 |
| N | 1 |
| V | 2 |
| Y | 2 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

La tabla muestra que existen los 27 ciclos antes mencionados y que esta categorizado por el nivel de inseguridad considerado en el marco teórico en la imagen 2, donde se identifica en un mapa de guayaquil: nivel alto 3, nivel medio 2 y nivel bajo 1.

- Estado de corte: esta variable reflejó la condición del servicio de agua potable, en donde se categorizará si está suspendido o suspendido total; se consideró que

esta variable ES estrechamente relacionada con la urgencia de la regularización del servicio. En el caso de la suspensión del servicio es un factor determinante que hace que los usuarios tengan motivación en priorizar el pago de esta deuda para que su servicio se reconecte. Normalmente, los usuarios en estado de Suspendido Total muestran un mayor cumplimiento rápido en comparación con aquellos en estado de Suspendido, pues es importante mencionar que el servicio es indispensable para sus actividades cotidianas o comerciales.

Tabla 8
Codificación Corte

| VARIABLE INDEPENDIENTE | |
|-------------------------------|--------------------|
| CORTE | VALOR DUMMY |
| SuspendidoTotal | 1 |
| Suspendido | 2 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

En la tabla se muestra la variable independiente Corte la cual en la base de datos se mostró según el estado del usuario con su codificación para gestión del modelo “Suspendido Total” =1 o “Suspendido” =2.

- Tipo de predio: la distribución de los usuarios se dio en base a las características físicas y funcionales de sus propiedades, casas, departamentos, terrenos baldíos, locales comerciales y entre otros. Además, se puede incluir que al reflejar características clave del usuario y su relación con la propiedad, se convierte en un elemento esencial para comprender y predecir el comportamiento de pago dentro del modelo de regresión logística.

Tabla 9
Codificación Predio

| VARIABLE INDEPENDIENTE | |
|--|------------------|
| PREDIO | CATEGORÍA PREDIO |
| CASA, VILLA, CHALÉS Y DEPARTAMENTO | 1 |
| LOCAL, EDIFICIO, OFICINA, GALPONES QUIOSCO, GARAJE Y URBANIZACIONES | 2 |
| SOLARES VACÍOS, ÁREAS VERDES Y OTROS | 3 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

En la tabla se evidencia que distribución del predio en base a categorías para el análisis respectivo para la creación del modelo, esta clasificación se dio considerando que:

1. Residenciales: Propiedades como casas, villas, chalés, departamento; que tuvieron un comportamiento más consistente en cuanto a cumplimiento.
2. Comerciales: Locales, edificio, oficinas, garaje, urbanizaciones, quiosco y galpones; los cuales estuvieron sujetos a fluctuaciones económicas más amplias, afectando su capacidad de pago.
3. Predios no ocupados: Solares vacíos, áreas verdes y otros; considerado un segmento con mayor probabilidad de incumplimiento debido a su baja utilización y prioridad.
 - Tipo de contribuyente: se consideró clave para influir en el comportamiento de pago. Se fue categorizando con el fin de identificar las diferencias estructurales y patrones distintos que tienen estos tipos de usuario.

Tabla 10*Codificación Contribuyente*

| VARIABLE INDEPENDIENTE | |
|------------------------|----------------------------|
| CONTRIBUYENTE | CATEGORÍA CONTRIBUYENTE |
| PERSONA NATURAL | 1 |
| SOCIEDAD | 2 |
| N/A | 3 |

*Fuente: Contratista COINSER SAS**Elaborado por: Autores de este documento*

Se visualizó la clasificación existente de los contribuyentes, donde cada uno toma una categoría para el análisis del modelo; se podrá analizar de la siguiente manera:

- ✓ Persona natural (valor= 1): usuarios que pueden presentar mayor variabilidad en el momento de regularizar su deuda, esto se debe a que la actividad económica que realizan en ocasiones es muy baja en ingresos económicos.
- ✓ Sociedad (valor=2): empresas o sociedades que por su nivel de apertura en el mercado de negocios tienen más apego comercial, teniendo más estabilidad económica.
- ✓ No aplica el usuario (valor= 3): aquellos que no están ingresados en el sistema como contribuyentes, en este caso es más complejo debido a que no se conoce ni siquiera la actividad económica que realizan, afectando a la gestión de cobranza.
 - Call center: gestión realizada por la contratista con el fin de comunicarse con el usuario para hacerle saber sobre su deuda y su propósito de pago. Si contesta el usuario: la gestión de call center llega a tener un impacto positivo en la probabilidad de reconexión, ya que actúa como un catalizador para acelerar el proceso de pago. Y cuando el cliente moroso no contestó: se consideró un desafío para las estrategias de cobranza, ya que la falta de respuesta puede deberse a factores como datos de contacto desactualizados, desinterés o incluso intenciones de no pago. Finalmente, este indicador

para entender cómo la interacción directa influye en el comportamiento de pago y, por ende, en la probabilidad de reconexión, que es la variable dependiente en el modelo de regresión logística.

Tabla 11
Codificación Call center

| VARIABLE INDEPENDIENTE | |
|------------------------|-------------|
| CALL CENTER | VALOR DUMMY |
| 10/10/2024 | 1 |
| N/A | 0 |

Fuente: Contratista COINSER SAS

Elaborado por: Autores de este documento

En la tabla se pudo visualizar que por medio de las fechas se da la gestión de call center en donde la variable explica que:

✓ **Gestión Exitosa (1 - Usuario contestó):** el usuario responde a la gestión telefónica, la empresa tiene la apertura de comunicar directamente las condiciones para regularizar su deuda, como plazos de pago, descuentos o incentivos.

✓ **Gestión Fallida (0 - Usuario no contestó):** escenarios donde el usuario no responde a la gestión telefónica, pues ende se disminuye la probabilidad de reconexión. Pues se consideró que contacto directo limita la posibilidad de negociar, lo que puede llevar a un mayor tiempo de suspensión del servicio.

○ **Gestión de visita:** por medio de esta variable se logró identificar personalmente la ubicación y al usuario referenciado, debido a que la cuadrilla encargada visita para gestionar el pago de la deuda; es muy importante esta gestión por los casos en la que los encargados del predio no responden a otros canales. El principal objetivo de esta es que al momento que el personal se encuentre el sitio el usuario decida pagar o incluse buscar otras alternativas de regularizar su deuda.

Tabla 12
Codificación Visita

| VARIABLE INDEPENDIENTE | |
|-------------------------------|--------------------|
| VISITA | VALOR DUMMY |
| 12/10/2024 | 1 |
| N/A | 0 |

Fuente: *Contratista COINSER SAS*

Elaborado por: *Autores de este documento*

Visualizando la tabla es importante reconocer que la variable se dará valor si se dio la visita al usuario en donde se llega a un acuerdo, en caso de ser imposible ubicarlo, o no se encuentran dentro del sitio pues la codificación es negativa. De esta manera se presentarán los datos para la validación del modelo:

✓ Si la visita es efectiva (valor = 1) se permitió establecer un contacto directo con el usuario, pues este acontecimiento incrementará las posibilidades de que este realice el pago y, como tal, el servicio sea reconectado.

✓ Si la visita es negativa (valor = 0), los deudores podrían no ser conscientes del estado de su deuda, disminuyendo la probabilidad de reconexión.

○ Verificación de corte: por medio de esta variable se logró evaluar si se realizó una revisión en el servicio que realmente fue suspendido, asegurando que los usuarios con deuda estén cumpliendo lo establecido, por lo cual se espera la regularización del servicio. El valor establecido por la variable se identifica de la siguiente manera:

Tabla 13
Codificación Corte

| VARIABLE INDEPENDIENTE | |
|-------------------------------|--------------------|
| CORTE | VALOR DUMMY |
| 12/12/2024 | 1 |
| N/A | 0 |

Fuente: Contratista COINSER SAS
Elaborado por: Autores de este documento

Se visualizó dentro de la tabla los valores dummy correspondientes para esta variable, por medio de esta se confirma que, si el usuario se rige a la suspensión del servicio, dando valores:

✓ Se verificó el corte (valor = 1) asegurando que la acción inicial de suspensión sea efectiva, pues se ejerció presión sobre el usuario para regularizar su deuda y obtener la reconexión del servicio.

✓ Si no se verificó el corte (valor = 0), no se logró evidenciar que el usuario mantenga el servicio de manera irregular, lo cual reduce su urgencia de pago y afecta negativamente las probabilidades de reconexión.

○ Contravención: con esta variable se pudo identificar que usuarios registraron actividades irregulares o ilegales, en caso de ser encontrada la irregularidad se registra una multa al usuario. Dentro de esta variable se puede identificar varios factores, pero el fin de esta es saber si se hizo la gestión de verificar si existe tales irregularidades.

Se mostrará como toma valor:

Tabla 14
Codificación Contravención

| VARIABLE INDEPENDIENTE | |
|------------------------|-------------|
| CONTRAVENCIÓN | VALOR DUMMY |
| 12/11/2024 | 1 |
| N/A | 0 |

Fuente: Contratista COINSER SAS
Elaborado por: Autores de este documento

Se visualizó en esta variable la existencia de un registro por fecha, esto se debe a que si se hizo la gestión de verificar el producto del usuario el valor =1, en caso de que el producto no se pueda revisar o el deudor se impone el valor= 0. Es decir, a partir de esta

variable la mayoría de los predios se regularizan debido a que se cava para la identificación y los usuarios prefieren pagar a que se dé tal acontecimiento.

Después de la explicación detallada de cada variable a utilizar para el modelo predilecto, se crea una sola base de dato en donde se evidencia la nueva matriz tomando valores binarios denominados dummy, también en ciertas variables basándose en las categorías que representan.

3.2 Resultados por Minitab

Después del proceso de dar codificación a la base de datos, fue posible el análisis por minitab, esta aplicación fue seleccionada debido a que mantiene una gran capacidad para manejar grandes conjuntos de datos. Por lo tanto, se llevaron a cabo pruebas de significancia, análisis de correlaciones, y se estimaron los coeficientes del modelo para interpretar el impacto de cada variable en la probabilidad de reconexión.

Fueron 4,267 usuarios mapeados y codificados para el respectivo análisis y ejecución del modelo predilecto de Regresión Logística. A continuación, se detallan los hallazgos liberados por minitab para el análisis del estudio:

- El modelo predilecto de regresión logística con función que fue denominada logit y codificación binaria de las variables categóricas, el cual se consideró adecuado para analizar la probabilidad de que los usuarios paguen sus deudas. En base a una muestra de 4,267 registros, el modelo permitió predecir el comportamiento de pago de los usuarios en función de las variables categóricas seleccionadas, lo que facilitó la segmentación de los usuarios con mayor probabilidad de pago y optimiza las estrategias de cobranza.

Figura 3
Detalle del Método

Método

| | |
|--|-------|
| Función de enlace | Logit |
| Codificación de predictores categóricos (1; 0) | |
| Filas utilizadas | 4267 |

Fuente: Minitab

Elaborado por: Autores de este proyecto

Nota: Evidencia la función del modelo y el total registrado para el análisis.

Figura 4
Eventos relacionados a la Variable dependiente

Información de respuesta

| Variable | Valor | Conteo |
|-------------------------|-------|--------------|
| RECONEXION DEL SERVICIO | 1 | 432 (Evento) |
| | 0 | 3835 |
| | Total | 4267 |

Fuente: Minitab

Elaborado por: Autores de este documento

Nota: Por medio de la variable Reconexión del servicio se evidencia que 432 son 1 y 3835 valor 0, es decir que existen más usuarios que no tienen registro de reconexión.

- Ecuación ajustada para predecir la probabilidad de que el evento ocurra, es decir, la probabilidad que se dio para que el usuario pague la deuda; se evaluaron las variables independientes las cuales afectaron en la probabilidad de pago.

Figura 5
Ecuación de regresión logística ajustada

Ecuación de regresión

$$P(1) = \frac{\exp(Y')}{1 + \exp(Y')}$$

$$Y' = -19 - 0,0972 \text{ LOG SALDO} + 0,00138 \text{ FACTURAS} + 0,0 \text{ CORTE}_1 + 2,615 \text{ CORTE}_2 + 0,0 \text{ RANGO}_1 - 1,690 \text{ RANGO}_2 - 2,031 \text{ RANGO}_3 - 2,563 \text{ RANGO}_4 - 3,149 \text{ RANGO}_5 - 4,352 \text{ RANGO}_6 - 3,432 \text{ RANGO}_7 - 4,785 \text{ RANGO}_8 + 0,0 \text{ PELIGROSIDAD}_1 - 0,121 \text{ PELIGROSIDAD}_2 - 0,079 \text{ PELIGROSIDAD}_3 + 0,0 \text{ PREDIO}_1 - 0,861 \text{ PREDIO}_2 - 0,341 \text{ PREDIO}_3 + 0,434 \text{ PREDIO}_4 + 0,0 \text{ CONTRIBUYENTE}_1 - 0,803 \text{ CONTRIBUYENTE}_2 - 0,080 \text{ CONTRIBUYENTE}_3 + 0,0 \text{ CALL CENTER}_0 + 20 \text{ CALL CENTER}_1 + 0,0 \text{ VISITA}_0 + 1,310 \text{ VISITA}_1 + 0,0 \text{ CORTE}_1_0 - 0,468 \text{ CORTE}_1_1 + 0,0 \text{ CONTRAVENCION}_0 + 0,709 \text{ CONTRAVENCION}_1$$

Fuente: Minitab

Elaborado por: Autores de este documento

Nota: La ecuación de regresión logística permitió calcular la probabilidad de pago de los usuarios, lo que se consideró útil para dirigir las acciones de cobranza de manera más eficiente. Se identificó las variables que tienen mayor impacto en base a la probabilidad de pago, como las gestiones de cobranza (call centers, visitas).

Figura 6
Análisis de los coeficientes

| Término | EE del | |
|---------------|---------|--------------|
| | Coef | coef. FIV |
| Constante | -19 | 303 |
| LOG SALDO | -0,0972 | 0,0708 1,85 |
| FACTURAS | 0,00138 | 0,00271 4,27 |
| CORTE | | |
| 2 | 2,615 | 0,202 1,15 |
| RANGO | | |
| 2 | -1,690 | 0,338 2,38 |
| 3 | -2,031 | 0,318 3,12 |
| 4 | -2,563 | 0,331 3,56 |
| 5 | -3,149 | 0,424 1,96 |
| 6 | -4,352 | 0,448 2,03 |
| 7 | -3,432 | 0,407 3,06 |
| 8 | -4,785 | 0,566 6,43 |
| PELIGROSIDAD | | |
| 2 | -0,121 | 0,254 2,99 |
| 3 | -0,079 | 0,274 3,37 |
| PREDIO | | |
| 2 | -0,861 | 0,341 1,06 |
| 3 | -0,341 | 0,254 1,08 |
| 4 | 0,434 | 0,494 1,11 |
| CONTRIBUYENTE | | |
| 2 | -0,803 | 0,332 1,23 |
| 3 | -0,080 | 0,156 1,13 |
| CALL CENTER | | |
| 1 | 20 | 303 1,00 |
| VISITA | | |
| 1 | 1,310 | 0,189 1,36 |
| CORTE_1 | | |
| 1 | -0,468 | 0,171 1,24 |
| CONTRAVENTION | | |
| 1 | 0,709 | 0,230 1,17 |

Fuente: Minitab

Elaborado por: Autores de este proyecto

Nota: El proceso para la regresión liberó esta tabla en donde se hizo referencia a los coeficientes.

El análisis de los coeficientes, por la regresión logística fueron:

1. **Término (Coeficientes):**

- **Constante:** El valor indicado de -19 es independiente en la ecuación del modelo. Es decir, este valor establece el punto de partida para la regresión logística.

- **Log saldo:** Se estableció como el coeficiente de -0.0972, el cual indicó que, por cada incremento en el logaritmo del saldo de la deuda, la probabilidad de pago disminuye, ya que el coeficiente es negativo.

- **Facturas:** Por medio de este coeficiente, se analizó que el 0.00138 mostró un aumento en el número de facturas vencidas tiene un pequeño pero positivo efecto sobre la probabilidad de pago. Pues, cada factura adicional aumentó la probabilidad de que el usuario pague.
- **Corte:** Los coeficientes para **Corte_1** (0) y **Corte_2** (2.615) sugieren que las acciones de corte de servicio tienen un gran impacto en la probabilidad de que el usuario pague, especialmente cuando se aplica el corte de servicio en las fases posteriores.
- **Rango:** Estos coeficientes dio valor a las distintas categorías de deuda, se mostró de la siguiente manera: rango 2 (-1.690), rango 3 (-2.031), rango 4 (-2,563), rango 5 (-3,149), rango 6 (-4,352), rango 7 (-3,432) y rango 8 (-4,785), lo cual se indicó que a medida que el rango de deuda aumentó, la probabilidad de pago disminuye.
- **Peligrosidad:** Se mostraron el rango de los coeficientes negativos para **peligrosidad 2** y **peligrosidad 3** (-0.121 y -0.079) se mencionó que los usuarios con mayor riesgo o peligrosidad tuvieron una menor probabilidad de pagar sus deudas.
- **Predio:** La codificación que se dio para esta variable **predio 2** y **predio 3** tienen coeficientes negativos, por lo tanto, se comentó que ciertos tipos de propiedades pueden estar asociados con menores probabilidades de pago. Pero existe una observación, el **predio 4** tiene un coeficiente positivo, lo que indicó que estos usuarios tienen una mayor probabilidad de pagar.
- **Contribuyente:** Se consideró negativo el coeficiente de **contribuyente 2** (-0.803) y **contribuyente 3** (-0.080), por lo que se sugirió que los contribuyentes con ciertas características tuvieron menor probabilidad de pago.

- **Call Center:** Coeficiente que es gestionado por la contratista y se dio positivo de 20 para **call center 1**, lo cual indicó que la intervención del call center tiene un efecto positivo y significativo en la probabilidad de que el usuario realice el pago.
- **Visita:** Este coeficiente de visita 1 dio como valor 1.310, por lo que se consideró, que la visita directa tiene un impacto positivo en la probabilidad de pago.
- **Corte_1 y Contravención:** En este caso los coeficientes, se dieron valores por: negativo para **corte 1** (-0.468) y positivo para **contravención 1** (0.709), donde finalmente mostraron que existió diferentes acciones afectan la probabilidad de pago, con el corte de servicio siendo menos efectivo en comparación con acciones de contravención.

2. **EE del coef.** Conocido como: (Error estándar del coeficiente)

- El error estándar midió la variabilidad de los coeficientes estimados. Y se pudo asegurar que cuanto más pequeño sea el error estándar, más confiables son los coeficientes. Por tanto, el coeficiente con menos valor es la variable factura donde se pudo corroborar que es el más fiable.

3. **FIV** conocido como: (Índice de inflación de varianza):

- Este índice midió la multicolinealidad entre las variables predictoras. Es decir, a valores mayores a 10 indican una posible colinealidad alta, lo que podría afectar la estabilidad del modelo. En base a las variables la única variable casi alta es el rango en su categoría 8 dando un valor FIV de 6,43.

Figura 7

Relaciones de probabilidad

Relaciones de probabilidades para predictores continuos

| | Relación de probabilidades | IC de 95% |
|-----------|-------------------------------|------------------|
| LOG SALDO | 0,9073 | (0,7898; 1,0424) |
| FACTURAS | 1,0014 | (0,9961; 1,0067) |

Fuente: Minitab

Elaborado por: Autores de este documento

Nota: Se analizó la relación existente entre las probabilidades predictoras de log saldo y facturas (variables independientes cuantitativas).

- **Log Saldo:** Existió una relación de probabilidades para el logaritmo del saldo pendiente es de **0,9073**. Es decir, se indicó que un aumento en el logaritmo del saldo pendiente disminuye la probabilidad de que el usuario pague su deuda. Pues, si el saldo de la deuda aumenta, la probabilidad de pago disminuye en un 9,27%, ya que la relación de probabilidades es menor que 1. Por otro lado, el intervalo de confianza al 95% para el logaritmo del saldo va de **0,7898** a **1,0424**, y se incluyó el valor 1, lo que implica que el efecto observado podría no ser estadísticamente significativo. Sabiendo que la relación de probabilidades es menor que 1, el intervalo de confianza sugiere que no podemos descartar completamente que el saldo no tenga un impacto claro sobre la probabilidad de pago en la muestra analizada.

- **Facturas:** La probabilidad para el número de facturas es de **1,0014**. En base a este valor se sugirió que, por cada factura adicional, la probabilidad de que un usuario pague aumenta en un 0,14%. Pues, la relación de probabilidades es ligeramente mayor que 1, se interpreta que el incremento en el número de facturas vencidas tiene un pequeño pero positivo efecto sobre la probabilidad de pago. Por otro lado, el intervalo de confianza al 95% para el número de facturas va de **0,9961** a **1,0067**. Este intervalo está muy cerca de 1, lo cual significó que el efecto de las facturas sobre la probabilidad de

pago es extremadamente pequeño y podría no ser relevante en términos prácticos. Sin embargo, estadísticamente, aún es posible que haya un impacto positivo.

Figura 8
Resumen del modelo

Resumen del modelo

| R-cuad. de desviación | R-cuad. (ajust) de desviación | AIC | AICc | BIC |
|-----------------------|-------------------------------|---------|---------|---------|
| 56,99% | 56,24% | 1247,28 | 1247,52 | 1387,17 |

Fuente: Minitab

Elaborado por: Autores de este documento.

Nota: Se tuvo una predicción del 56,99%, siendo un porcentaje considerable para el funcionamiento de la regresión.

Se presentó un resumen de los resultados de ajuste del modelo de regresión logística, donde se destacó varias métricas clave. Es importante conocer que el **R-cuadrado de desviación** es de **56,99%**, lo que indica que el modelo explica aproximadamente el 57% de la variabilidad en los datos, lo que sugiere un buen *nivel de ajuste*. Aunque, **R-cuadrado ajustado** es ligeramente inferior, con un **56,24%**, lo cual es una medida que ajusta el R-cuadrado cuando se tuvo en cuenta el número de variables predictoras en el modelo. Respecto a **AIC (1247,28)** y **AICc (1247,52)** son analizados muy cercanos, lo que implicó que la penalización por el número de parámetros adicionales es mínima. Además, como último detalle, el **BIC (1387,17)**, es de los que penaliza más fuertemente los modelos complejos, siendo considerado como el más alto que los otros dos, lo que sugiere que, si bien el modelo es eficiente, puede no ser el modelo más tedioso en términos de complejidad. Finalmente, estos resultados indican que el modelo tiene un ajuste razonable y es eficaz en la predicción de la variable dependiente, aunque existe una ligera complejidad en su estructura.

Figura 9

Pruebas de bondad de ajuste

Pruebas de bondad del ajuste

| Prueba | GL | Chi-cuadrada | Valor p |
|-----------------|------|--------------|---------|
| Desviación | 4245 | 1203,28 | 1,000 |
| Pearson | 4245 | 1877,82 | 1,000 |
| Hosmer-Lemeshow | 8 | 6,63 | 0,577 |

Fuente: Minitab

Elaborado por: Autores de este proyecto.

Nota: Se analizó la tabla pruebas de bondad en base a las asignaciones.

Se presentaron los resultados en base a las pruebas de bondad del ajuste, en donde se evaluó la calidad del modelo de regresión logística. La prueba realizada **de desviación**, se observa un valor de **Chi-cuadrada** de **1203,28** con un **valor p** de **1,000**, es decir, que el modelo no muestra una mala adaptación y, por lo tanto, no hay evidencia suficiente para rechazar la hipótesis nula de que el modelo se ajusta bien. Por otro lado, la **prueba de Pearson**, el valor de **Chi-cuadrada** fue **1877,82** con un **valor p** de **1,000**, lo que también indica un buen ajuste del modelo. También se tomó en cuenta, la **prueba de Hosmer-Lemeshow**, con **8 grados de libertad** y un **valor Chi-cuadrada** de **6,63** con un **valor p** de **0,577**, lo cual dio como resultado de que el valor p superior a 0,05 sugiere que no hay evidencia para rechazar la hipótesis de un buen ajuste, indicando que el modelo tiene un buen desempeño en cuanto a su capacidad de predicción.

Figura 10
Análisis de la varianza

Análisis de Varianza

| Fuente | Prueba de Wald | | |
|---------------|----------------|--------------|---------|
| | GL | Chi-cuadrada | Valor p |
| Regresión | 21 | 377,90 | 0,000 |
| LOG SALDO | 1 | 1,89 | 0,170 |
| FACTURAS | 1 | 0,26 | 0,610 |
| CORTE | 1 | 168,18 | 0,000 |
| RANGO | 7 | 113,92 | 0,000 |
| PELIGROSIDAD | 2 | 0,24 | 0,885 |
| PREDIO | 3 | 9,03 | 0,029 |
| CONTRIBUYENTE | 2 | 5,84 | 0,054 |
| CALL CENTER | 1 | 0,00 | 0,947 |
| VISITA | 1 | 48,19 | 0,000 |
| CORTE_1 | 1 | 7,47 | 0,006 |
| CONTRAVENCION | 1 | 9,50 | 0,002 |

Fuente: Minitab

Elaborado por: Autores de este proyecto.

Nota: Se analizó la tabla referente al análisis de varianza donde el valor $p \leq 0,05$.

Los resultados presentados en base al **Análisis de Varianza (ANOVA)**, enfocándose la **Prueba de Wald**, que fue utilizado para evaluar la significancia de las variables en un modelo de regresión logística. El **valor Chi-cuadrada** para la **regresión** es **377,90** con un **valor p** de **0,000**, lo que se indicó que el modelo predilecto en su conjunto es altamente significativo. Se logro analizar las variables individuales, por lo cual, se observó que algunas tenían un impacto significativo en el modelo. **Corte** (Chi-cuadrada = 168,18, valor p = 0,000) y **Rango** (Chi-cuadrada = 113,92, valor p = 0,000) se consideraron altamente significativas, por lo que se intuyó que influyen de manera sustancial en la predicción de la probabilidad de pago. Pero existió otros casos, en donde las variables como **Log saldo** (Chi-cuadrada = 1,89, valor p = 0,170) y **Facturas** (Chi-cuadrada = 0,26, valor p = 0,610) no siendo significativas, ya que sus valores p superan el umbral de 0,05, indicando que no tienen una influencia estadísticamente relevante en el

modelo. Finalmente, otras como **Contribuyente** y **Visita** también presentan valores p mayores a 0,05, lo que sugiere que su impacto es marginal.

Se concluyó que las variables **Corte** y **Rango** son las más significativas en el modelo y siendo claves para su efectividad, ya que sus valores p extremadamente bajos indican que tienen un impacto considerable en la probabilidad de pago de los usuarios. Es decir, estas variables fortalecen el modelo al ofrecer información crucial sobre los comportamientos de los usuarios. Incluso es importante mencionar que las variables como **Log Saldo** y **Facturas**, debido a su falta de significancia, podrían ser consideradas para ser excluidas del modelo en futuras iteraciones, mejorando la precisión y simplicidad del modelo al centrarse en los predictores más relevantes.

Para mostrar cómo funciona el modelo se logró calcular en dos líneas de la base de datos en donde la respuesta de la ecuación será enviada codificada en base a 0 = probabilidad negativa de que el usuario pague y 1 = probabilidad positiva de que el usuario decida regularizar su deuda. Finalmente, se logró proyectar por medio de esta ecuación el primer trimestre del 2025 identificando que usuarios se podría ir directo a gestionar y quienes se tardarán o incluso es complejo su regularización; esta analítica se podrá evidenciar en dentro del análisis realizado con Power Bi.

3.3 Análisis de Clústeres

En esta sección se presentan los resultados obtenidos en la propuesta como solución al decremento en la facturación de la Contratista de INTERAGUA, COINSER S.A.S. se planteó establecer un modelo de regresión para conocer la posibilidad de que ocurra una reconexión, la cual es el medio por el cual la empresa genera liquidez cada periodo mensual y a través de conocer esta probabilidad de ocurrencia se estableció

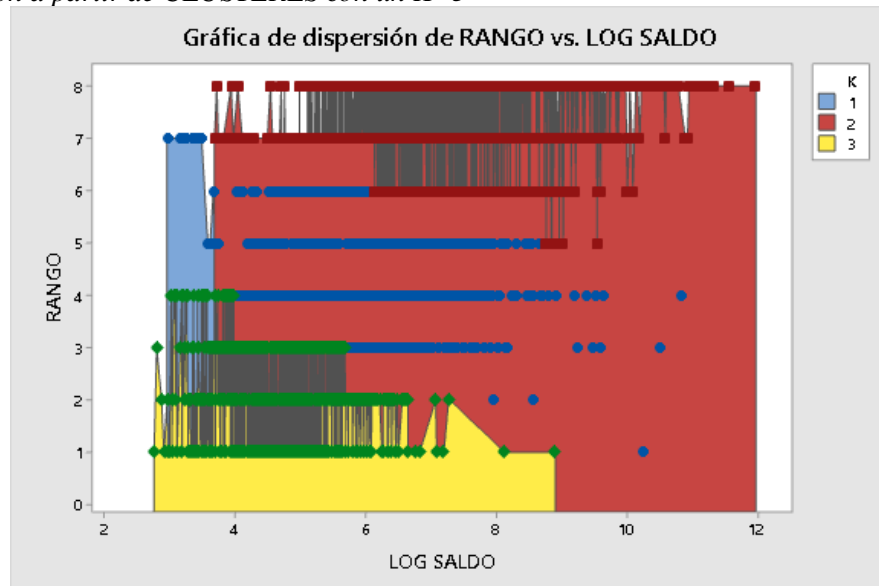
clústeres con el fin de enfocar de forma eficiente los recursos de cobranzas que posee la contratista.

En la presente se obtuvo mediante un modelo de regresión logístico la probabilidad de que ocurra una reconexión del servicio de agua potable, tal acción de ejecuta después de que el usuario realice el pago. Por ende, el modelo nos indica la probabilidad de que los usuarios realicen un convenio de pago o liquiden la deuda.

En el modelo de regresión se obtuvieron resultados que determinaron un R-cuadrado ajustado del 56.24%, lo que indica que el modelo es capaz de explicar el 56.24% de la variabilidad de la variable dependiente, es decir, la probabilidad de que se dé una reconexión y con esta nueva medición se procedió a crear los clústeres.

Para establecer los clústeres se consideró un $K=3$ y $K=4$ como se muestra a continuación:

Figura 11
Dispersión a partir de CLUSTERES con un $K=3$



Fuente: Minitab

Elaborado por: Autores de este documento

En el grafico se muestra los 3 clústeres diferenciado por colores sin embargo estos conglomerados no son adecuado ya que existe superposición en los clústeres y se requirió un ajuste en el número de clústeres para lo cual se estableció K=4 para una mayor precisión.

Después de considerar el cambio, se procedió hacerlo:

Figura 12

Centroides de grupo

Centroides de grupo

| Variable | Conglomerado1 | Conglomerado2 | Conglomerado3 | Conglomerado4 | Centroide principal |
|-----------|---------------|---------------|---------------|---------------|---------------------|
| LOG SALDO | 4,6249 | 7,0938 | 4,8042 | 6,6180 | 6,1596 |
| RANGO | 3,7896 | 7,4213 | 1,5351 | 3,9408 | 5,2011 |
| PRE RECO | 0,0087 | 0,0000 | 0,1971 | 0,0070 | 0,0314 |

Fuente: Minitab

Elaborado por: Autores de este documento.

Nota: se conocen los centroides de grupo en donde se identifica que se generaron 4

conglomerados para el análisis correspondiente.

Figura 13

Partición final

Partición final

| | Número de observaciones | Dentro de la suma de cuadrados del conglomerado | Distancia promedio desde el centroide | Distancia máxima desde centroide |
|---------------|-------------------------|---|---------------------------------------|----------------------------------|
| Conglomerado1 | 903 | 1150,483 | 0,979 | 3,619 |
| Conglomerado2 | 2008 | 4488,765 | 1,333 | 4,907 |
| Conglomerado3 | 613 | 662,476 | 0,967 | 2,454 |
| Conglomerado4 | 743 | 899,408 | 0,938 | 4,673 |

Fuente: Minitab

Elaborado por: Autores de este documento.

Nota: se presentaron la división de cada conglomerado y que cantidad de observaciones abarca cada uno.

El análisis de clústeres ha sido exitoso al clasificar a los clientes en 4 grupos distintos con comportamientos diferenciados en función de sus saldos, deudas y su disposición para regularizar su situación.

De tal manera, se considera que los clústeres se encuentran bien segmentados, dando como referencia que cada conglomerado se encuentra dentro de su posición basado en las variables seleccionada Log saldo y Rango, lo cual, permite un análisis específico del comportamiento de pago de los usuarios con nivel de morosidad.

Análisis descriptivo de los Clústeres:

- **Conglomerado 1 (Bajo saldo y deudas):**

La principal característica de este conglomerado 1 es conocer que son aquellos que su naturalidad es tener bajo saldo y monto de sus deudas no son un valor alto, sino comparado a los otros es el más bajo. Sin embargo, su baja disposición para regularizar su deuda (según el valor de PRE RECO) sugiere que, aunque son clientes con menores dificultades económicas, requieren estrategias de cobranza menos intensivas, como recordatorios o pagos a plazos más flexibles.

- **Conglomerado 2 (Altos saldos y deudas):**

Los clientes de este clúster tienen saldos elevados pendientes y deudas elevadas. A pesar de tener mayores dificultades para regularizar (según el valor bajo de PRE RECO), este grupo debe recibir estrategias de cobranza más intensivas, como visitas de cobranza y contacto telefónico constante.

- **Conglomerado 3 (Bajas deudas, pero alta disposición para reconectar):**

Los usuarios de este clúster tienen bajas deudas, pero están altamente enfocados en la reconexión. Este conglomerado se puede considerar, teniendo un distintivo debido a que son los perfiles más adecuados para regularizar la deuda por lo cual se enfoca en dar planes incentivos para que el usuario pueda dar su pago en una fecha prevista cercana.

- **Conglomerado 4 (Deudas regulares y moderada capacidad para reconectar):**

Grupo de conglomerado con aspectos regulares para una reconexión habilitándose la disposición para que los deudores pueden hacer convenios de pago por su monto total de facturas adeudadas. Se considera un planteamiento basa en cada usuario donde se analiza la situación socioeconómica y demográfica de cada moroso, para dar acciones de cobranza con incentivos para recuperar el valor vencido.

El R-cuadrado ajustado de 56.24% predice el nivel de porcentaje del modelo dando valor que es razonable y significativamente bueno; pues dando la probabilidad que se ejecutaran un porcentaje mayor de reconexión de lo antes previsto en el trimestre evaluado, pero existe otro porcentaje que es posible mejorar con más variables o datos más representativos en base al usuario. No obstante, es eficiente el modelo para tomar decisiones operativas para encaminadas al crecimiento de la contratista y veracidad en sus resultados como empresa prestadora de servicios.

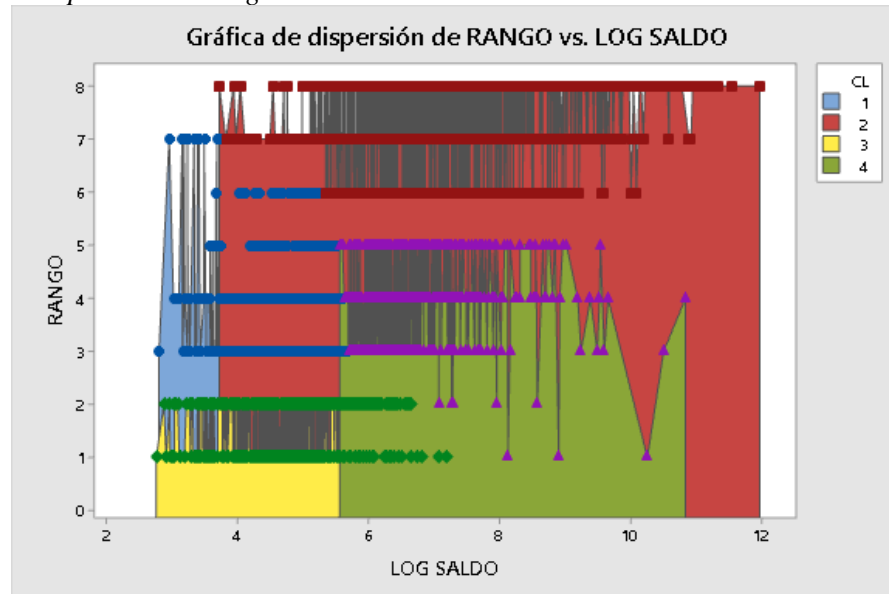
Predicción de Reconexiones:

La herramienta que se desarrollo sirve para predecir las reconexiones para el primer trimestre de 2025, logró establecer una proyección de 1,513 predios, lo que representa 3.5 veces más de lo que se ejecutó en el último del año anterior. Esto es

evidencia de que, con un enfoque correcto y un uso más eficiente de los recursos, las estrategias de cobranza pueden generar un aumento significativo en las reconexiones.

Figura 14

Gráfica de dispersión de Rango vs. Saldo



Fuente: Minitab

Elaborado por: Autores de este documento

Nota: Se presentaron la gráfica donde se evidencia el área que abarca cada conglomerado.

✓ **Conglomerado 1 (Color azul y símbolo de círculo):**

Este clúster tiene bajos valores de LOG SALDO y RANGO intermedio, lo que indica que estos usuarios tienen saldos bajos y deudas no tan altas.

Los valores bajos de PRE RECO indican que estos usuarios no están tan enfocados en regularizar su deuda.

✓ **Conglomerado 2 (Color rojo y símbolo de cuadrado):**

Este clúster tiene valores altos en LOG SALDO y RANGO. Los clientes en este clúster tienen saldos pendientes elevados y por un tiempo elevado.

El valor bajo de PRE RECO en este clúster indica que estos usuarios tienen menos probabilidades de regularizar su deuda o reconectar el servicio.

✓ **Conglomerado 3 (Color verde y símbolo de diamante):**

Este clúster tiene valores bajos de RANGO y moderados de LOG SALDO, lo que indica que estos clientes tienen bajas deudas, pero valores intermedios de saldo pendiente.

Los valores altos de PRE RECO indican que estos clientes están más enfocados en la reconexión o en regularizar su situación financiera.

✓ **Conglomerado 4 (Color morado y símbolo de triángulo):**

Este clúster tiene valores de LOG SALDO y RANGO intermedios, indicando un perfil de clientes con deuda moderada.

El valor de PRE RECO es relativamente alto, lo que sugiere que estos usuarios también están enfocados en reconectar o regularizar su servicio.

El valor $K=4$ muestra ser adecuado, ya que los clústeres están bien diferenciados según las variables LOG SALDO y RANGO. Se evidencia que las características entre los conglomerados son bastante diferenciadas por lo que se planificaría las estrategias más efectivas para cada uno de los clústeres.

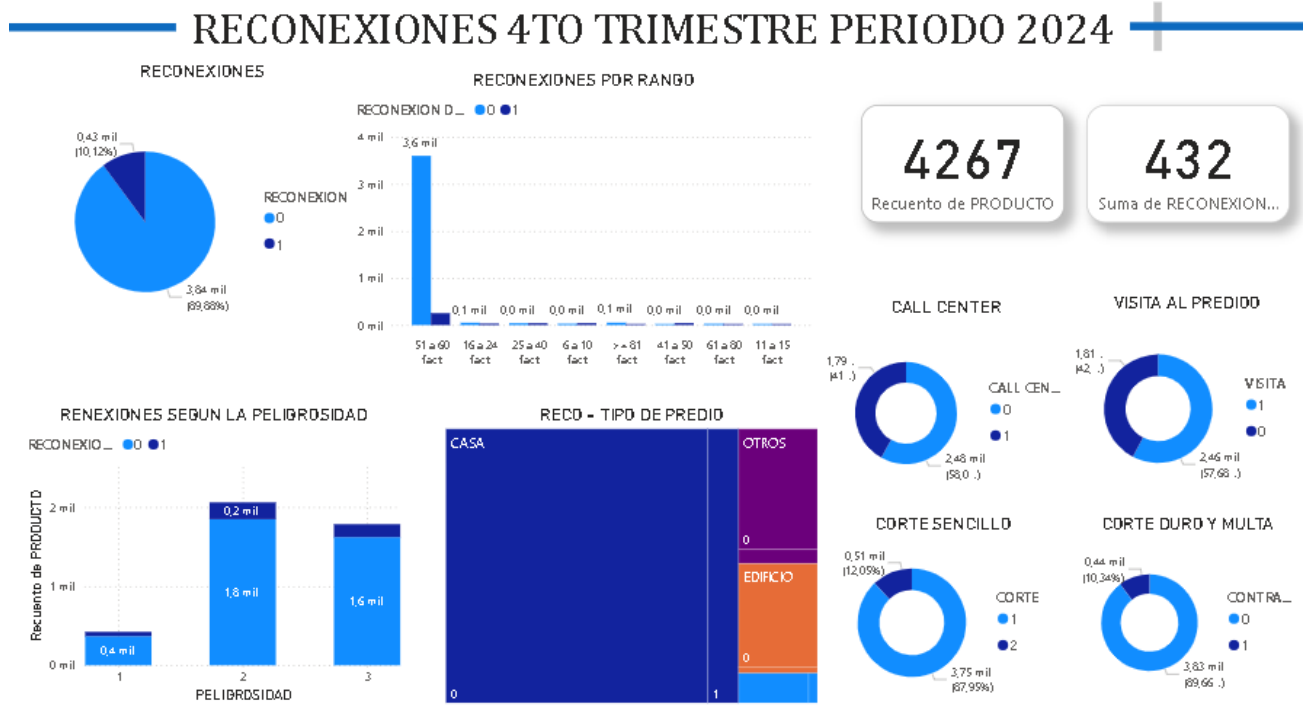
- I. Conglomerado 2 probablemente representa a clientes con mayores deudas
- II. Conglomerado 1 tiene clientes con menores dificultades.
- III. Los clústeres 3 y 4 muestran un enfoque más claro hacia la reconexión y regularización, lo que podría indicar clientes que están más enfocados en saldar sus deudas.

3.4 Presentación de Resultados Power Bi

A continuación, se presentan dos dashboard donde se visualizan los datos de las reconexiones que fueron ejecutadas en el último trimestre del año anterior y la

proyección de las reconexiones para el primer Trim 2025, que gracias a la herramienta desarrollada establece la predicción de las reconexiones a 1513 predios lo que viene a ser 3,5 veces más.

Figura 15
Reconexiones 4to Trimestre Periodo 2024



Fuente: Power Bi

Elaborado por: Autores de este documento.

Nota: Se muestran estos diagramas que fue elaborada por los datos enviados por el cliente, en donde se evidenció que el 4to trimestre del año 2024 tuvo 4,267 productos con morosidad donde es esos se dieron solo 432 reconexiones.

Figura 16
Reconexiones Proyectadas al 1er Trimestre Periodo 2025



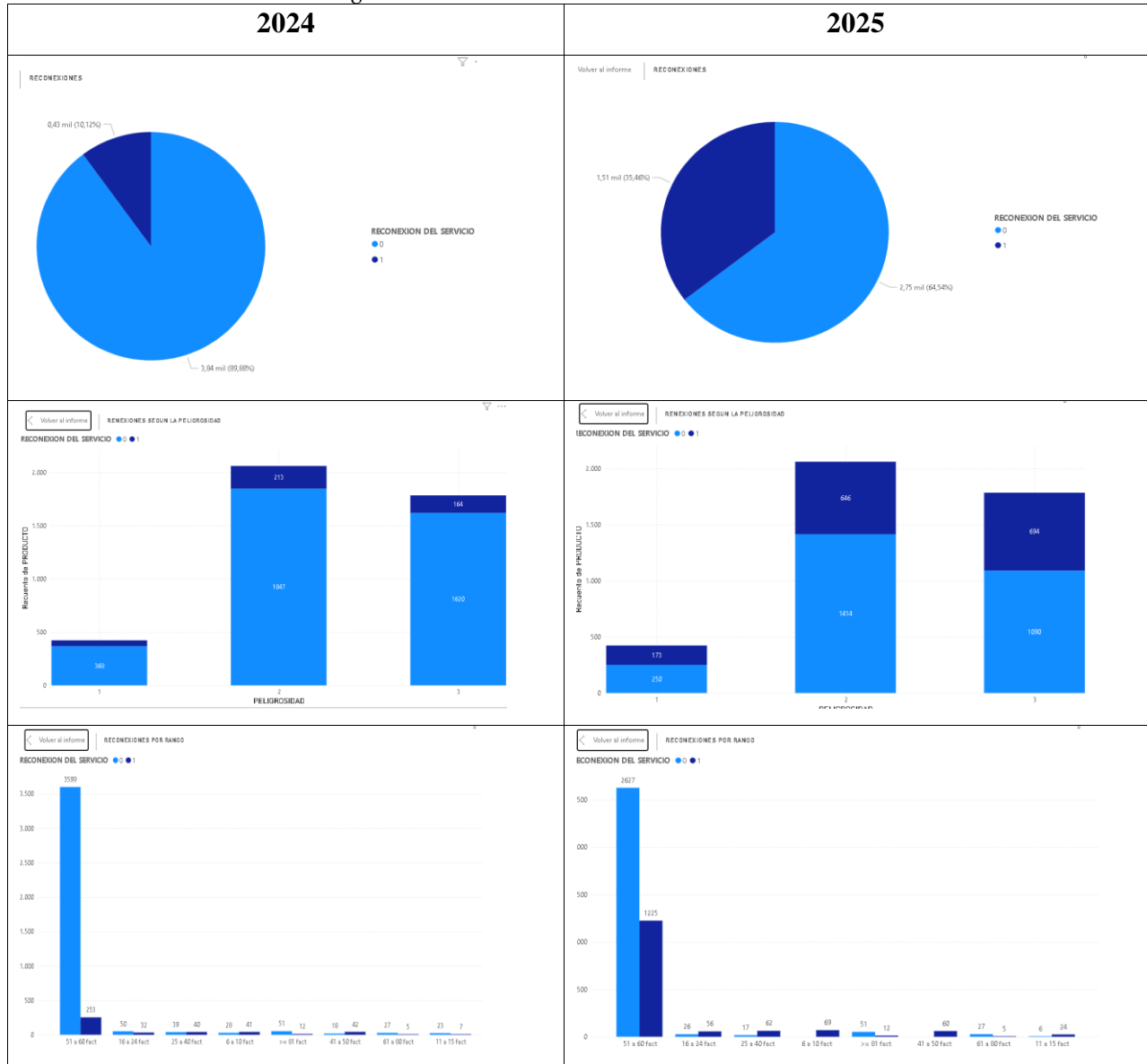
Fuente: Power Bi

Elaborado por: Autores de este documento.

Nota: Se muestran estos diagramas elaborados por medio del modelo predilecto en donde se pudo proyectar con 4,267 productos 1,513 reconexiones en beneficio para Interagua y una buena gestión realizada por la Contratista Coinser S.A.S.

A continuación, se presentaron por medio de la base de datos diagramas en donde se analizará la diferencia existente entre lo gestionado ya por el cliente en el 2024 y en base al modelo predilecto una proyección al 2025:

Tabla 15
Presentación de los diagramas Análisis 2024 vs 2025



Fuente: Power Bi

Elaborado por: Autores de este documento.

Nota: En base a estos diagramas se generará un análisis conclusión.

Tabla 16
Análisis conclusión

| Deducciones | 4TO TRIM 2024 | PROYECCIÓN 1ER TRIM 2025 |
|--|----------------------|---------------------------------|
| Reconexiones | 432 | 1513 |
| Reconexiones por zona de riesgo nivel 2 | 213 | 646 |
| Reconexiones por zona de riesgo nivel 3 | 164 | 694 |
| Reconexiones por Rango 51 a 60 fact | 253 | 1225 |

Fuente: Power Bi

Elaborado por: Autores de este documento.

Nota: Analítica en números que representan las variaciones de las reconexiones tanto por zonas de riesgo como por el rango de facturas adeudas.

Los gráficos pastel muestran la variación que existe antes y después de la aplicación de los clústeres. Los gráficos pastel evidencian que para el 4to trimestre del 2024 se realizaron 423 reconexiones representando el 10,12% del total de la cartera, y que para la proyección del primer trimestre del 2025 las reconexiones representarían el 35,46% aplicando las recomendaciones especificadas para cada clúster. Para lo cual el 10,12% representa una recuperación de la cartera para INTERAGUA por un monto de \$328.592,77 mientras que para el 2025 se proyecta el 35,46% representando una cartera de \$1.253.651,96.

3.5 Análisis de Costos

Por medio del modelo de regresión logística, se buscó optimizar la gestión de cobranzas y reconexiones, se ofrece una alternativa en base a estadística tecnológica, bajo un modelo que sea viable para la economía de la contratista dando resultado positivo a su cliente, por ese lado se complementa por un modelo que aumenta las reconexiones de manera mensual, es decir, sin necesitar costos adicionales en la operatividad post-visita que gestionaba la contratista bajo su sistema pasado. Es clave mencionar, que se tendrá que invertir en el personal dando capacitaciones en base al conocimiento de nuevas

técnicas estadísticas presentes en base al modelo, haciendo que la contratista suba tecnológicamente usando herramientas afianzadas a la recuperación de la cartera morosa. Por otro lado, para que las gestiones propuestas sean efectivas se deberá comprar bases de datos que contengan correos electrónicos y números de teléfonos actualizados para poder realizar las gestiones de cobro, esto tiene un costo mensual entre \$80 a \$200 mensuales, sin embargo, la contratista al manejar bases trimestrales puede realizar este gasto 4 veces y al mismo tiempo la contratación de un operador de call center.

Capítulo 4

4 Conclusiones y recomendaciones

4.1 Conclusiones

- El proceso dio un resultado ha demostrado ser una estrategia con una eficiencia del 57,99% al momento de revelar la probabilidad de pago, incluso que la toma de decisiones sea adecuada y planificada por el líder y su cuadrilla. Por medio de la segmentación adecuada de los usuarios con mayor probabilidad de pago, se ha optimizado el proceso de cobranza, aumentando el número de reconexiones mensuales, lo cual está alineado con el objetivo general de mejorar la eficiencia operativa de la contratista.
- Implementar un modelo predictivo basado en regresión logística es tecnológicamente factible, pues se ha integrado de manera eficiente con las herramientas y sistemas existentes en la infraestructura de la empresa. Se utilizaron varias herramientas, como Power BI y Excel, permiten una fácil visualización de los datos y la segmentación de los usuarios, lo cual facilita la toma de decisiones estratégicas para la gestión de cobranzas, donde finalmente se pudo presentar gráficamente tanto lo ocurrido en el 2024 como una proyección para el primer trimestre del 2025.
- El resultado esperado genera un impacto directo en la recuperación de cartera vencida, dando prioridad a las acciones de cobranza en usuarios con mayor probabilidad de pago, dando como resultados en términos de recuperación de deuda, cumpliendo así con los objetivos específicos de optimizar el proceso de cobranza y mejorar la rentabilidad para Interagua.
- COINSER S.A.S podrá priorizar acciones de cobranza más estratégicas en los conglomerados 3 y 4. Mejorando su viabilidad financiera y el cumplimiento de los objetivos establecidos en la gestión de la cartera vencida. La capacidad de identificar patrones y predecir comportamientos de los usuarios ha permitido diseñar estrategias de

cobranza más efectivas, alineadas con los objetivos estratégicos de la empresa, y garantizar una mayor eficiencia en los recursos utilizados.

4.2 Recomendaciones

Después del análisis y los resultados evidentes mostrados y se dio una proyección base para el primer trimestre del 2024, damos por constancia que por medio de las herramientas, el modelo matemático de regresión logística y clústeres se recomienda priorizar los recursos de cobranzas con estrategias basadas en el cada clúster:

1. Conglomerado 1 (bajas deudas y saldos):

- Se recomienda implementar estrategias menos intensivas (es decir, gestionar por call center antes de dar una visita al predio) de esa manera ahorramos costos y se presupuesta costos para alcanzar la proyección, como recordatorios por correo electrónico o envíos de mensajes ofreciendo planes de pago flexibles.
- Es importante siempre hablar con el cliente, que se envié una base de datos actualizada para un contacto eficiente con los usuarios.
- Utilizar herramientas automatizadas para reducir el costo de la cobranza como por ejemplo un chat bot.

2. Conglomerado 2 (altos saldos y deudas):

- Se recomiendan estrategias más agresivas, incluso tener a las cuadrillas más capacitadas para la gestión por visitas de cobranza o tener contacto con el cliente que nos otorgue información más específica de cada usuario (la cual debería estar en el contrato firmado por la adquisición del servicio), una parte importante y de mejora para esta operatividad es que Interagua mejore su proceso de otorgamiento de productos.
- Ofrecer acuerdos de pago escalonados o descuentos por pronto pago para incentivar a los clientes a regularizar sus deudas.

3. Conglomerado 3 (disposición alta para reconectar):

- Grupo que debe ser focalizado con rapidez con ofertas de convenios con abonos iniciales atractivos o disposiciones para condonar un porcentaje de la deuda.

- Incentivar su reconexión con promociones y ofertas especiales.

4. Conglomerado 4 (deudas intermedias y disposición moderada):

- Para este conglomerado se recomienda utilizar un enfoque personalizado, iniciando por un contacto por medio de mensajes con ofertas de pagos, convenios y condonaciones. En caso de no obtener respuesta incrementar la intensidad.

- Es importante contar con una base de datos actualizada para un contacto eficiente con los usuarios.

Optimizar el Uso de los Recursos de Cobranza:

Basado en los clústeres, las cobranzas pueden a partir de ahora asignar recursos de manera más eficiente:

Según los resultados, los clústeres con mayor probabilidad de reconexión son los conglomerados 3 y 4 los cuales deberían recibir ofertas atractivas y estrategias personalizadas, lo que resultará en un uso más eficiente de los recursos.

El clúster de alto riesgo es conglomerado 2, el cual debe recibir un enfoque más fuerte con visitas de cobranza en el cual se debe asistir con un detalle de todos convenios disponibles para la deuda, el monto exacto de abono inicial, a cuantos meses plazo y el monto mensual del convenio informando al usuario que deberá mes a mes cancelar a tiempo su factura de consumo regular más el valor del convenio, aunque a un costo mayor es más eficiente una visita para que el mensaje llegue al usuario de forma efectiva teniendo en cuenta que nadie puede subsistir sin agua a que es agua es indispensable para la vida.

Control de Resultados y Ajustes Continuos en el modelo y clústeres:

El modelo de regresión logística y el análisis de clústeres deben ser monitoreados periódicamente para ajustar las estrategias de cobranza según los resultados obtenidos. Si ciertos clústeres muestran cambios en su comportamiento de pago, se deben aplicar nuevas estrategias de cobranza según la variación que se haya detectado.

El análisis predictivo de futuras reconexiones (como el modelo que predice 1,513 reconexiones en el primer trimestre de 2025) se deberá ejecutar con el fin de ajustar las metas y planificar los recursos de cobranza.

Incentivos de Reconexión:

Es crucial en la cobranza brindar a los usuarios un panorama donde se haya regularizado, el agua potable al ser indispensable para la vida y necesario para llevar a cabo limpieza, preparación de alimentos, higiene entre otras, es muy atractivo para los usuarios brindar un panorama donde cuentan con el servicio por un costo moderado. Los incentivos de reconexión para clústeres con altas probabilidades de regularizar (Conglomerado 3 y 4). Esto puede incluir disposiciones específicas vigentes, condonaciones por un porcentaje del total de la deuda, aplazamientos de pago o reducción de cargos por reconexión.

Estos incentivos ayudarán a aumentar la tasa de reconexiones y a recuperar más ingresos sin necesidad de incrementar considerablemente el costo de la cobranza.

El modelo de regresión logística proporciona una visión clara de las variables que afectan la probabilidad de reconexión, permitiendo una segmentación más precisa.

Proyecciones de reconexiones muestran un aumento significativo en las reconexiones si se optimizan las estrategias de cobranza.

Monitoreo y ajustes continuos de las estrategias asegurarán que el enfoque de cobranza siga siendo eficiente y costos eficientes.

El objetivo principal de crear clústeres es clasificar a los usuarios de manera que el departamento de cobranza pueda aplicar estrategias más eficaces y específicas según las características de cada grupo.

En la gráfica de dispersión, se visualizan los 4 clústeres según las variables LOG SALDO (eje X) y RANGO (eje Y). Los clústeres están representados por diferentes colores y símbolos.

Referencias

- Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Obtenido de <https://archive.org/details/TheCreditScoringToolkitTheoryAndPracticeForRetailCreditRiskMgmt.DecisionAutomati?>
- Andrea Jaramillo. (Febrero de 2024). *Impacto de la inseguridad en Ecuador en el ámbito laboral*. Obtenido de <https://actuarial.com/impacto-de-la-inseguridad-en-ecuador-en-el-ambito-laboral-desafios-y-perspectivas/>
- Asobanca. (Diciembre de 2023). *Evaluación de la Banca Privada Ecuatoriana*. Obtenido de <https://asobanca.org.ec/wp-content/uploads/2024/01/Evolucion-de-la-Banca-12-2023-1.pdf>
- Banco Central del Ecuador. (Febrero de 2024). *Gestión de la Liquidez del Sistema Financiero No. 227*. Obtenido de <https://contenido.bce.fin.ec/documentos/Estadisticas/SectorMonFin/Liquidez/gli202402.pdf>
- Banco Central del Ecuador. (2024). *Información Económica*. Obtenido de <https://www.bce.fin.ec/informacioneconomica>
- Banco Central del Ecuador. (Septiembre de 2024). *La economía ecuatoriana*. Obtenido de <https://www.bce.fin.ec/boletines-de-prensa-archivo/la-economia-ecuatoriana-reporto-una-contraccion-de-2-2-en-el-segundo-trimestre-de-2024>
- Brachfield, P. (2008). *Memorias de un cazador de morosos*. Madrid: Gestión 2000. Obtenido de https://cladea.org/wp-content/uploads/2022/01/263Franklin-Lopez-Vera-and-Geovanny-Chancay-Quimis-Plan-Estrategico-Para-Recuperar-Cartera-Vencida-De-Clientes-De-Cnel-Ep-Agencia-De-Negocios-En-La-Ciudad-De-Guayaquil.pdf?utm_source
- Calderón Saldaña & Alzamora de los Godos. (2009). *Revista Salud, Sexualidad y Sociedad*. Obtenido de REGRESIÓN LOGÍSTICA APLICADA A LA EPIDEMIOLOGÍA: <https://inppares.org/magazine/Revista%20IV%202009/10-%20Regres%20Logis.pdf>
- Calvo, O., & Aurelio, C. D. (2002). *Regresión logística no condicionada y tamaño de muestra*. Obtenido de https://scielo.isciii.es/scielo.php?pid=S1135-57272002000200002&script=sci_arttext

- Campos, R. (Octubre de 2017). *COBRAR A CLIENTES MOROSOS: LAS CINCO ARMAS DEL RECOBRO. ICIRED Impagados*. Obtenido de <https://iciredimpagados.com/blog/cobrar-clientes-morosos/>
- Chávez Páez. (2021). *Informe de diagnóstico de recuperación económica urbana y resiliencia*. Obtenido de <https://repositorio.cepal.org/server/api/core/bitstreams/e7284f48-9586-45c7-a187-94659d6e18b6/content>
- Consejo de Participación Ciudadana . (2024). *SEGUNDA CONVOCATORIA A INTEGRAR OBSERVATORIO CIUDADANO*. Obtenido de Convocatorias Observatorios Ciudadanos: https://www.cpccs.gob.ec/wp-content/uploads/2024/12/2da__convocatoria_oc_seguridad_ciudadana_rbba-1.pdf
- Constitución del Ecuador. (2008). *LEY ORGANICA DE RECURSOS HIDRICOS USOS Y APROVECHAMIENTO DEL AGUA*. Obtenido de <https://educacion.gob.ec/wp-content/uploads/downloads/2012/08/Constitucion.pdf>
- COOTAD. (19 de Dic de 2019). *CODIGO ORGANICO DE ORGANIZACION TERRITORIAL*. Obtenido de <https://www.cpccs.gob.ec/wp-content/uploads/2020/01/cootad.pdf>
- Dolores, & Rodríguez Pérez, J. C. (2000). *La regresión logística: una herramienta versátil*. Obtenido de <https://www.revistanefrologia.com/es-la-regresion-logistica-una-herramienta-articulo-X0211699500035664>
- EMAPAD-EP. (Enero de 2024). *RESOLUCIÓN ADMINISTRATIVA No.- EMAPAD-EP-GG-2024-002*. Obtenido de http://www.emapad.gob.ec/app/lotaip/LOTAIP/2024/13.ADJUNTOS/RESOLUCION_No_EMAPAD-EP-GG-2024-02_Aprobacion_del_pac_PAC_2024-signed.pdf
- Espinosa Jarrin, Kevin Enrique. (2020). *Eficiencia en la recuperación de la cartera generada en los procedimientos de determinación tributaria*. Obtenido de PUCE.EDU.EC: <https://repositorio.puce.edu.ec/items/51b155eb-931c-45fe-b199-498dec2f9b0f>
- FasterCapital. (Junio de 2024). *Agrupacion de riesgo crediticio como agrupar el riesgo crediticio utilizando K Means y agrupacion jerarquica*. Obtenido de <https://fastercapital.com/es/contenido/Agrupacion-de-riesgo-crediticio--como-agrupar-el-riesgo-crediticio-utilizando-K-Means-y-agrupacion-jerarquica.html>

FasterCapital. (Junio de 2024). *Historial de pagos como afecta su calificación de capacidad de pago y como mejorarlo*. Obtenido de <https://fastercapital.com/es/contenido/Historial-de-pagos--como-afecta-su-calificacion-de-capacidad-de-pago-y-como-mejorarlo.html>

Funcionpublica. (Noviembre de 2024). *Ley 142 de 1994 - Gestor Normativo*. Obtenido de <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=2752>

Generation Spain. (Febrero de 2024). *Entendiendo el desempleo: causas, consecuencias y soluciones*. Obtenido de <https://spain.generation.org/blog/entendiendo-el-desempleo-causas-consecuencias-y-soluciones/>

Hernández Sampieri, Fernández Collado, Baptista. (2014). *Metodología de la investigación*. Obtenido de <https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y%20Baptista-Metodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf>

Hernández Sampieri, R., Fernández Collado, C., & B. (2014). *Metodología de la investigación (6ª ed.)*. McGraw Hill. Obtenido de <https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y%20Baptista-Metodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf>

Holaerp.com. (2024). *KPIs en facturación: indicadores clave de rendimiento*. Obtenido de https://doi.org/10539282/module_164386044695_Social_Sharing

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Obtenido de <https://dl.icdst.org/pdfs/files4/7751d268eb7358d3ca5bd88968d9227a.pdf?>

INEC. (2024). *Powerbi.com*. Obtenido de <https://app.powerbi.com/view?r=eyJrljoiNGUxZjQyMDUtMzg0Zi00MzI0LTk5NWEtY2JiMWUzM2YyYjdlIiwidCI6ImYxNThhMmU4LWNhZWmtNDQwNi1iMGFiLWY1ZTI1OWJkYTExMiJ9>

Interagua. (10 de Mayo de 2013). *Home page*. Obtenido de [Interagua.com.ec](https://www.interagua.com.ec/)

Interagua. (2022). *Veolia*. Obtenido de <https://www.latinoamerica.veolia.com/es/casos-estudio/interagua>

Knüsel, L. . (2007). *Exact Computation of Minimum Sample Size for Estimation of Binomial Parameters*. Obtenido de <https://arxiv.org/abs/0707.2113>

- LEY ORGANICA DEL AGUA. (2014). *LEY ORGANICA DE RECURSOS HIDRICOS USOS Y APROVECHAMIENTO DEL AGUA*. Obtenido de <https://www.regulacionagua.gob.ec/wp-content/uploads/downloads/2016/03/Ley-Org%C3%A1nica-de-Recursos-H%C3%ADricos-Usos-y-Aprovechamiento-del-Agua.pdf>
- Liliana Orellana. (2001). *¿Qué es la estadística?* Obtenido de https://www.dm.uba.ar/materias/estadistica_Q/2011/1/modulo%20descriptiva.pdf
- Marcela Guachamín. (2024). *Factores micro y macro que inciden en la morosidad*. Obtenido de <file:///C:/Users/usuario1/Downloads/466-Texto%20del%20art%C3%ADculo-2563-1-10-20240628.pdf>
- Ministerio del Ambiente y Agua EPMAPAQ. (Agosto de 2021). *INFORME TÉCNICO No.022-DAPS-SAPSRD-2021*. Obtenido de <https://www.epmapaq.gob.ec/wp-content/uploads/Archivos/RENDICION%20DE%20CUENTAS/INFORME%20TECNICO%20022-DAPS-SAPSRD-2021.pdf>
- Mogollón Gómez, J. (2021). *Gestión de Cobranza y su impacto en la Gerencia Financiera de la empresa PP S.A: Periodo 2014-2016*. Obtenido de <https://doi.org/10.37073/puriq.3.1.121>
- Municipio de Manta. (2021). *ESTUDIO PREVIO PARA LA CONTRATACIÓN DE: “SERVICIOS DE SUSPENSIÓN Y RECONEXIÓN DEL SERVICIO DE AGUA POTABLE*. Obtenido de <https://www.epam.gob.ec/wp-content/uploads/2014/08/ESTUDIO-PREVIO-PARA-CONTRATACION-SERVICIO-SUSPENSION-DE-AGUA.pdf?>
- OECD. (Septiembre de 2024). *Denuncias de contrabando, tráfico de armas, droga, etc. Ecuador*. Obtenido de <https://oeco.pdf.org/visualizador-de-datos-denuncias-de-delitos-graves/>
- OIT. (Julio de 2024). *OIT presenta Estrategia FORLAC 2.0*. Obtenido de <https://www.ilo.org/es/resource/news/oit-presenta-estrategia-forlac-2-para-abordar-informalidad>
- ONU. (Agosto de 2010). *64/292. El derecho humano al agua y el saneamiento*. Obtenido de file:///C:/Users/usuario1/Downloads/A_RES_64_292-ES.pdf
- ONU. (2010). *Agua y saneamiento*. Obtenido de <https://www.ohchr.org/es/topic/water-and-sanitation#:~:text=En%202010%2C%20la%20Asamblea%20General,de%20todos%20los%20derechos%20humanos%20C2%BB>.

- Oriol Gilabert López. . (2024). *Análisis de Segmentación de Clientes Usando K-Means*.
Obtenido de <https://medium.com/@oriolgilabertlopez/an%C3%A1lisis-de-segmentaci%C3%B3n-de-clientes-usando-k-means-una-gu%C3%ADa-pr%C3%A1ctica-72ff735ba9e7>
- Paneda, B. (2023). *La segmentación conductual en la cobranza*. Obtenido de Crsoftware.com. : <https://doi.org/1059714.fs1.hubspotusercontent-na1>
- Raimundo Soto. (2003). *LA CORRUPCIÓN DESDE UNA PERSPECTIVA ECONÓMICA*.
Obtenido de https://www.flacsoandes.edu.ec/web/imagesFTP/1275931953.raimundo_soto.pdf
- Ruiz Fúnez. (Abril de 2014). *Análisis de Productividad en un Call Center de cobranzas*.
Obtenido de <https://repositorio.unitec.edu/server/api/core/bitstreams/6cbb5f5b-7db0-4c13-9d9b-16897476cbe7/content>
- Thomas, L. C., Crook, J. N., & Edelman. (2002). *Credit Scoring and Its Applications*.
Obtenido de <https://epubs.siam.org/doi/pdf/10.1137/1.9780898718317.fm?>