



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**Propuesta de modelo de riesgo para prevenir pérdidas en  
crédito al consumo: Caso de una sociedad financiera de  
objeto múltiple (SOFOM)**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**ACTUARIA**

**P R E S E N T A:**

**GABRIELA MANZANERA HERNÁNDEZ**



**DIRECTOR DE TESIS:  
ACT. JOSÉ FERNANDO SORIANO FLORES  
2013**



## **Jurado Asignado**

Presidente: Act. Alberto Cadena Martínez

Vocal: Act. Jorge Otilio Avendaño Estrada

Secretario Tutor: Act. José Fernando Soriano Flores

Suplente: Dra. Ruth Selene Fuentes García

Suplente: Act. Ricardo Villegas Azcorra.



## *Agradecimientos*

*A Dios por la vida.*

*A mis padres por su amor absolutamente incondicional.*

*A la Universidad Nacional Autónoma de México, mi segunda casa.*

*A mis profesores, con admiración y respeto.*

*Un agradecimiento especial a mi tutor, Act. José Fernando Soriano Flores, por su orientación en esta etapa especial de mi vida.*



*Vayan mi trabajo y los esfuerzos de mi vida a la memoria de mi  
hermano Alonso.*

*Hoy me quedo un tiempo más en este mundo, viviendo doble vez,  
una por ti y otra por mí.*

*Te quiero.*





# Índice general

<b>Introducción</b>	<b>1</b>
<b>1. Administración de riesgo en crédito al consumo</b>	<b>11</b>
1.1. El riesgo . . . . .	11
1.2. Tipos de crédito . . . . .	13
1.3. Ciclo de la administración del riesgo . . . . .	14
1.4. Modelos preventivos de pérdidas . . . . .	15
<b>2. Ciclo del riesgo</b>	<b>17</b>
2.1. Origen del crédito . . . . .	18
2.2. Administración de portafolio . . . . .	18
2.3. Gestión de cobranza . . . . .	19
2.4. Modelo de riesgo . . . . .	21
<b>3. Bases técnicas actuariales. Tópicos estadísticos a utilizar</b>	<b>25</b>
3.1. Base de datos . . . . .	25
3.2. Estadística . . . . .	27
3.2.1. Estadística paramétrica . . . . .	28
3.2.2. Estadística no paramétrica . . . . .	35

---

<b>4. Construcción del modelo de prevención de pérdidas</b>	<b>41</b>
4.1. Tipos y formas de construir modelos . . . . .	41
4.1.1. Consideraciones del modelo . . . . .	44
4.1.2. Variables empleadas . . . . .	45
4.1.3. Análisis bivariado . . . . .	51
4.1.4. Insumo para regresión logística . . . . .	60
4.1.5. Metodología empleada . . . . .	63
4.2. Pruebas de Bondad de Ajuste . . . . .	66
4.3. Metodología empleada para construir el Prevent Score . . . . .	68
4.3.1. Validación del modelo . . . . .	73
<b>Conclusiones</b>	<b>79</b>
<b>Anexo</b>	<b>81</b>

# Introducción

*Crédito es, sin duda, confianza.*

El origen etimológico de la palabra “crédito” se remonta al latín *creditus*, vocablo resultado de la sustantivación del verbo “creyer” traducido al castellano como “creer”. Como su origen lo indica, el crédito hace alusión a confianza, certidumbre y seguridad. El filósofo, político y economista inglés John Stuart Mill<sup>1</sup> (1806-1873) aporta su propio enfoque al concepto definiéndolo como “el permiso para usar el capital de otro”.

El crédito yace en la confianza depositada al deudor por parte del acreedor, quien le proporciona un capital pactado por ambos. Dicho capital no será únicamente de índole financiero: puede consistir de igual manera en bienes o servicios. Se puede deducir que el crédito abarca cualquier forma de pago diferido en la que el prestamista deposita su confianza en el deudor ya que el prestatario no cubrirá inmediatamente su deuda. El deudor goza de los derechos de disponer de “lo prestado”, por lo tanto, está comprometido a devolver los recursos solicitados al prestatario en el plazo previa-

---

<sup>1</sup>John Stuart Mill (1806-1873). Economista, lógico y filósofo británico. Hijo del también economista James Mill, fue educado de forma exclusiva por éste según los estrictos principios del *Emilio* de Rousseau. Mill es uno de los claros ejemplos de la economía clásica y de la corriente teórica del utilitarismo antes de ceder paso a un incipiente socialismo.

Formado en el pensamiento de Adam Smith y David Ricardo, Mill se abrió al positivismo y otras nuevas ideas, también muy influenciado por su mujer Harriet Taylor en cuanto a sus teorías feministas.

Mill escribió ensayos sobre diversas materias, desde ética y lógica hasta economía y psicología. Su obra más conocida es “Sobre la libertad” (1859).

mente definido y según las condiciones establecidas para dicho préstamo, incluyendo los intereses devengados, seguros y costos asociados si ése fuera el caso.

El otorgamiento de un crédito implica un contrato, el cual por si mismo genera obligaciones y derechos para ambas partes. En su acepción jurídica, el “crédito” es una “promesa de pago”, estableciendo un vínculo jurídico prestamista-prestatario. El deudor tendrá la obligación de pagar; asimismo, el acreedor tendrá el derecho de reclamar el respectivo pago. La confianza es la base del crédito; sin embargo, implica un riesgo. La posible pérdida que asume un agente económico sería consecuencia del incumplimiento de las obligaciones contractuales que incumben a las contrapartes con las cuales se relaciona. Este riesgo de crédito es compensado por el “interés”, dependiendo de la probabilidad de incumplimiento del deudor. Riesgo-Interés son dos razones directamente proporcionales, ante un mayor riesgo (enfrentado por el acreedor) mayor será la tasa de interés que tendrá que cubrir el deudor.

### **Origen del crédito**

*Tan antiguo como la propia civilización.*

Precedente a la Era Cristiana, ya podemos encontrar en la antigua Roma los primeros indicios del desarrollo crediticio. Hoy en día existen constancias de leyes y decretos que establecían penas corporales al deudor moroso o que no cumplía con los acuerdos pactados previamente con el acreedor. Asimismo, se han encontrado manuscritos históricos en los cuales se establecían penas al deudor que incumplía con algún convenio estipulado, como la confiscación de sus bienes, el arresto, la pena de muerte e incluso su venta como esclavo.

En un inicio, el préstamo se realizaba en especie, y no fue sino hasta la aparición y el uso de la moneda cuando hallamos los incipientes principios crediticios. El banquero como el prestamista, empezaron a cobrar un papel de suma importancia, ya que su

labor era de cambista<sup>2</sup> y mercader<sup>3</sup> de metales preciosos. Cobraba los réditos a sus clientes cuando los deudores radicaban en el extranjero, y tenía la obligación de pagar las deudas de los clientes locales a los acreedores que residían en otros lugares. En el siglo XII nos encontramos con los primeros bancos.

Los babilonios dejaron escritos de órdenes de pago en tablillas de barro. Los griegos y los romanos, en cambio, utilizaban la letra de cambio<sup>4</sup>.

Cuando el mercader viajaba depositaba los fondos con el banquero de su ciudad, quien le extendía un título que respaldaba el monto depositado, el cual se cobraría en el sitio de destino. El proceso evolutivo del crédito vivió una notable mejoría con el comercio marítimo. Se requerían recursos económicos para el movimiento de importación y exportación, el cual era desarrollado como sociedad mercantil, en donde los prestamistas y comerciantes eran dueños de la mercancía; si la embarcación naufragaba, el deudor quedaba libre de la obligación de cubrir su deuda.

### **Importancia del crédito y su función social**

*Cuando el acceso al financiamiento es limitado, también se restringen las posibilidades de crecimiento de una economía.*

Dentro del sistema en que vivimos, el dinero ha sido el eje de las actividades económicas. Su uso es fundamental, ya que a partir de este podemos proveerle de un valor a nuestros bienes. Han evolucionado las diversas formas de intercambio de

---

<sup>2</sup>Cambista: se refiere a la persona que cambia dinero en los mercados; en establecimientos de cambio.

<sup>3</sup>Mercader: Persona que trata o comercia con géneros vendibles, es decir, mercaderías o mercancías. Se le aplican diferentes calificativos en función de la mercancía con la que trabaja: mercader de hierro, mercader de telas, mercader de finanzas, etc.

<sup>4</sup>Letra de cambio: Es el título de crédito o de valor formal y completo que contiene una orden incondicionada y abstracta de hacer pagar a su vencimiento, ya sea al tomador o a su orden, una suma de dinero en un lugar determinado, vinculando solidariamente a todos los que en ella intervienen.

productos y servicios, e irremediablemente se requiere un instrumento de intercambio. Hoy en día, para facilitar el intercambio comercial, el uso del dinero es combinado con un sistema de crédito.

La fuente principal de ingreso económico la constituye el trabajo remunerado. La remuneración más baja obtenida por una jornada de trabajo (salario mínimo) está diseñada para proporcionar a una familia los medios necesarios para obtener los elementos requeridos para su sostenimiento; sin embargo, este ideal político y económico es evidentemente utópico en nuestro país. Para cubrir sus necesidades básicas, las familias recurren a diferentes alternativas para obtener ingresos. Principalmente, recurren a la operación crediticia más común: el préstamo.

El crédito repercute en la reactivación del sistema productivo de un país de la siguiente manera:

1. Cuando el crédito genera un mayor consumo (público o privado) se promueve la demanda presente de bienes y servicios, fomentando la producción de las empresas.
2. Cuando el crédito aumenta la inversión privada, se eleva la capacidad productiva de las empresas, beneficiando la productividad del sector económico que genera mayores fuentes de empleo.
3. Cuando el crédito aumenta en la inversión pública, se generan beneficios sociales en diversos ámbitos (salud, educación, conectividad, etcétera).

### **Crédito al consumo**

El crédito al consumo ocupa nuestra atención. El “crédito al consumo” o “crédito de adquisición de bienes de consumo duradero” es el producto otorgado para comprar un bien o servicio de uso personal en plazos determinados. El crédito al consumo es el contrato por el que un empresario otorga o se compromete a otorgar a su cliente crédito bajo la forma de pago diferido, préstamo, apertura de crédito, o cualquier otro medio equivalente de financiamiento para satisfacer necesidades personales, pudiendo

ser el empresario cualquier persona física o moral que en el desempeño de su actividad, profesión u oficio, concede créditos. En cambio, el consumidor es la persona física que actúa con un propósito ajeno a su actividad o profesión en estas relaciones contractuales.

Existe un fuerte vínculo entre el “crédito” y el “consumo”, pese a ser dos conceptos independientes. Ambos interfieren en la reactivación económica de un país: el acto del consumo se ve favorecido por el aplazamiento del precio, sin embargo, dicho aplazamiento conlleva obligaciones al deudor. Actualmente, el crédito al consumo abarca diversos convenios, tales como venta a plazos, préstamo personal, apertura de crédito, tarjetas de crédito, etcétera.

Existen dos tipos de créditos al consumo:

- Crédito personal. Se basa en la persona natural; dicha persona cuenta con excedentes de capital, ya sea por bienes o por sus ingresos, lo que le permite solventar el préstamo y realizarlo.
- Crédito de tarjeta. El banco da una tarjeta al cliente para hacer uso de consumo. La tarjeta tiene un determinado monto para gastar (artículos personales u otros).

### **Importancia del crédito al consumo**

Para la economía, este tipo de crédito permite apoyar su crecimiento fomentando el consumo. Al permitir distribuir el financiamiento de un bien durante su tiempo de vida, el crédito de consumo permite optimizar la gestión de la tesorería familiar (caja familiar), beneficiando evidentemente la economía de un país.

*El sistema capitalista no podría sostenerse ni reproducirse si fallara el sistema de crédito.*

Los créditos al consumo incentivan el consumo de las personas, activando el sistema productivo de un país. Además, juegan un papel importante para el Gobierno, ya

que le permiten llevar a cabo inversión social en infraestructura y en otros propósitos para mantener la actividad económica de un país en niveles aceptables o mejor aún, favorables. Sin embargo, el nivel de endeudamiento (los créditos) debe mantenerse a niveles aceptables, verificando que no se afecte la estabilidad económica de los mexicanos y, en consecuencia, de nuestro país.

El otorgamiento del crédito al consumo implica un riesgo; por lo tanto, resulta de vital importancia prevenir las pérdidas una vez ya otorgado el crédito. Esto es, establecer estrategias que permitan aminorar los riesgos en los cuales incurre el acreedor y así prevenir el deterioro futuro de su portafolio de inversión. Para el cliente, el uso responsable del crédito facilita realizar gastos de consumo e inversión por encima de lo que permiten los ingresos corrientes.

En nuestro país la cartera de crédito al consumo y el saldo de las tarjetas crediticias registraron en noviembre del 2011 pasado el incremento mensual más alto en tres años, lo que confirma la tendencia de recuperación en estos segmentos de financiamiento. Los resultados se dan luego de que ambas carteras mostraron una importante contracción entre 2008 y 2009, derivada del incremento en la morosidad, principalmente en los plásticos crediticios. Así, de acuerdo con la Comisión Nacional Bancaria y de Valores (CNBV), el saldo de esta cartera respecto a octubre pasó de 392 mil 329.7 millones a 400 mil 127.6 millones de pesos, lo que significó un incremento en el mes de 1.99 por ciento, el mayor aumento porcentual y nominal con 7,798 millones de pesos en los últimos tres años. Esta mejoría responde al crecimiento de 2.0 por ciento o 7,470 millones de pesos en la cartera vigente, y de 1.76 por ciento o 328 millones de pesos en el saldo vencido. De esta manera, el índice de Morosidad (IMOR) del crédito al consumo se mantuvo prácticamente sin cambio, al pasar de 4.77 a 4.76 por ciento, al igual que el índice de Cobertura de Cartera Vencida con Estimaciones Preventivas (ICOR), que varió de 190 a 189 por ciento. En tanto, por cuarto mes consecutivo la cartera de tarjeta de crédito observó una variación positiva, pero esta vez el incremento mensual fue el mayor de los últimos tres años, tanto en términos nominales como porcentuales. Este circuito de crédito reportó aumento en su saldo en cuatro mil 364 millones de pesos, esto es 2.07 por ciento por encima



del observado al cierre de octubre previo, con lo que pasó de 211 mil 048.2 millones a 215 mil 411.7 millones de pesos. En este mes, también destacó la continuidad en la disminución de la morosidad, al pasar de 5.82 a 5.67 por ciento.

Por otro lado, la cartera de créditos personales mostró variación positiva de 2.04 por ciento, esto es, 2,279 millones de pesos respecto a octubre pasado; dicho crecimiento estuvo acompañado de un aumento en el IMOR de 4.03 a 4.20 por ciento. La CNBV precisó que, respecto a los circuitos de crédito que integran al consumo, los principales otorgantes de financiamiento automotriz (considerando el número de créditos originados en octubre) fueron BBVA Bancomer, Banorte, Scotiabank y Volkswagen Bank, con tasas de interés promedio entre 13.31 y 14.11 por ciento. Sin embargo, el número de créditos totales otorgados por estos cuatro bancos disminuyó de 13 mil 588 en septiembre a 12 mil 378 en octubre, lo cual implicó una contracción en monto financiado en el mes de 12.39 por ciento. En el segmento de créditos personales y otros créditos al consumo, Banco compartamos, S.A. se mantuvo como el principal otorgante con 4,020 millones de pesos originados en octubre, a través de 563 mil 882 créditos y con una tasa anual promedio de 77.38 por ciento. El crédito promedio otorgado por este banco bajó de siete mil 146 a siete mil 129 pesos, estos créditos están orientados al sector de microfinanzas.

En lo referente a los créditos de nómina, BBVA Bancomer se mantuvo como el principal referente, con una originación en octubre de 61 mil 150 créditos (14.2 por ciento menos que los otorgados el mes anterior) que acumularon un monto de mil 514 millones de pesos y con tasa promedio de 33.68 por ciento, 1.24 puntos porcentuales por debajo de la registrada en septiembre. Mientras, en el segmento de créditos para la adquisición de bienes muebles, Banco Azteca es la única institución con originación relevante, con un millón 036 mil 025 créditos y un monto otorgado de mil 299 millones de pesos.

Además, el crédito promedio de este banco se incrementó de mil 196 a mil 254 pesos. El organismo regulador y supervisor también informó que, al 30 de noviembre de 2010, la banca múltiple y las sociedades financieras de objeto limitado y múltiple reguladas presentaron un saldo de cartera total de dos billones 112 mil millones de

pesos, es decir, 1.36 por ciento más con respecto al cierre del mes anterior.

El índice de morosidad mide el porcentaje de los créditos concedidos por los bancos y cajas de ahorro cuyo pago tiene un retraso de más de tres meses. La tasa de morosidad es un factor sensible a los ciclos económicos de recesión y expansión de la economía. Una inflación inesperada redistribuye la riqueza de los acreedores a favor de los deudores, favoreciendo a los prestatarios de dinero y perjudicando a los que lo han prestado. Las empresas serán en este caso entidades de crédito, otorgando aplazamientos de pago a sus clientes, lo que las lleva a enfrentar riesgos crediticios muy elevados y, en el peor de los casos, a una situación irremediable de desventaja competitiva. El peligro que la existencia de morosidad representa para las empresas es una cuestión que debe ser tratada y corregida eficientemente. Las empresas requieren de una adecuada documentación para cada una de las operaciones comerciales que se realizan, con el fin de proteger eficazmente el riesgo de impago en el que incurren. Es necesario garantizar el cobro mediante un adecuado aseguramiento jurídico de todas aquellas operaciones realizadas con los deudores. La documentación legal necesaria le brindará a la empresa judicialmente el derecho de crédito, lo cual la beneficiará ahorrándole costos innecesarios.

### **La crisis del 2007 y la crisis europea**

El origen de la crisis del 2007 a nivel mundial tuvo su origen en los Estados Unidos. En el verano de 2007 se pone de manifiesto que el nivel de defaults (morosos) del sector hipotecario estaba muy por encima de las medias estadísticas históricas. Debido a una mala gestión del crédito se suscitó dicho problema. A causa de ello un gran número de cédulas hipotecarias titulizadas sobre estos préstamos perdieron su cotización y, de forma repentina, se volvió ilíquido un mercado que hasta entonces procuraba financiamiento a una parte importante del sector inmobiliario de los Estados Unidos. A partir de ahí, y dado el entramado de productos estructurados alrededor de estos activos titulizados se genera una creciente desconfianza en el mercado financiero. La crisis que se suscitó en el año 2007 a nivel mundial originada en los

Estados Unidos a causa de la venta de materia prima a precios altos, la sobrevaloración del producto, entre otras causas, provocó que se vieran afectados diversos países, entre ellos nuestro país con quien tiene relaciones comerciales originando desempleo y aumento de precios de diferentes productos de consumo básico.

Dicha situación originó una gran incertidumbre financiera mundial y fuertes necesidades de liquidez de diversas instituciones para poder cumplir con sus compromisos financieros. En el año 2008 México se ubicó en la posición número 11 en las economías más grandes del planeta, según el Fondo Monetario Internacional, por encima de España y debajo de los Estados Unidos, China, Japón, India, Alemania, Reino Unido, Francia, Rusia, Italia y Brasil.

La incertidumbre financiera mundial llevó a que grandes inversionistas mundiales busquen deshacerse de sus activos en otros países y traten de convertir esos recursos de moneda local a dólares. Los inversionistas extranjeros sacaron de México recursos por US\$ 22,190,000 depositados en la Bolsa Mexicana de Valores o en bonos de deuda del gobierno federal.

Por otro lado la crisis europea es una serie de acontecimientos que han venido afectando negativamente desde principios del año 2010 a los 16 estados miembros de la UE que conforman la Eurozona o Zona Euro, esto es, que han adoptado el euro como moneda única.

Durante este periodo los estados de la Eurozona han venido padeciendo una crisis de confianza sin precedentes, con ataques especulativos sobre los bonos públicos de varios de sus miembros, turbulencias en sus mercados financieros y bursátiles, y una caída del valor cambiario de la moneda única en un contexto de incertidumbre y dificultad por alcanzar un acuerdo colectivo que todavía persiste.

Es así que se propone un modelo de riesgo para prevenir pérdidas en crédito al consumo que pronostica la probabilidad de que un cliente vigente comience a deteriorar su comportamiento de pago en su fecha límite, recurriendo a una de las herramientas estadísticas con mayor capacidad para el análisis de datos: la regresión logística. El objetivo esencial que tiene esta técnica es modelar cuánto influye en la probabilidad la aparición de un suceso dicotómico, en este caso un cliente tendrá o no

un buen comportamiento en un futuro. Basándose en datos reales de una SOFOM y empleando una regresión logística se pronosticó el comportamiento de los clientes. La intención es, una vez estimada la probabilidad de incumplimiento de pago aprovechar el poder de predictibilidad del modelo para que la institución desarrolle estrategias preventivas de cobranza: “Mayor intensidad de cobro para clientes con mayor probabilidad de incumplimiento de pago”. De esta manera el resultado repercutirá en la frecuencia en la que los clientes caen en impagos, y se pretende que dicha frecuencia sea menor para pérdidas futuras.

El presente trabajo se encuentra estructurado como sigue:

En el primer capítulo se presentan los principales conceptos de la administración de riesgo en crédito al consumo. Posteriormente se describen detalladamente los diferentes tipos de mercados de consumo y finalmente se expone el ciclo de la administración del riesgo .

En el segundo capítulo se describen los tópicos estadísticos a los cuales se recurrió para la realización del modelo de riesgo propuesto.

En el tercer capítulo se expone el ciclo del riesgo y las 3 etapas que lo componen:

1. Origen del crédito
2. Administración de portafolio
3. Recuperación de pérdidas por crédito y origen

En el cuarto capítulo se presenta la metodología para la realización del modelo de riesgo propuesto para la prevención de pérdidas en crédito al consumo.

# Capítulo 1

## Administración de riesgo en crédito al consumo

### 1.1. El riesgo

El término riesgo tiene varias acepciones dependiendo del área de estudio de que se trate, pero desde el punto de vista actuarial está referido al producto de la probabilidad de que un evento (generalmente no deseable) ocurra, y al daño esperado debido a la ocurrencia de dicho evento (pérdidas económicas). El riesgo crediticio es la posibilidad de que un prestatario o contraparte no pueda cumplir con sus obligaciones de acuerdo con los términos previamente acordados por ambas partes.

Todas las instituciones de crédito se encuentran expuestas a riesgos en diversos ámbitos como resultado de su propia operación y para hacer frente a dichos riesgos deberán ser económicamente solventes para enfrentar sus obligaciones y de igual manera tener un rendimiento sobre su capital.

Como principio básico de la administración de riesgos, a mayor riesgo debe tenerse un mayor monto de recursos reservados, para así poder hacer frente al monto de pérdidas que podrían suscitarse.

Existen diversos tipos de riesgos financieros:

- **Riesgo de mercado.** Pérdida potencial por cambios en los factores de riesgo que inciden sobre la valuación o sobre los resultados esperados de las operaciones activas, pasivas o causantes de pasivo contingente, tales como tasa de interés, tipos de cambio e índices de precios, etcétera.
- **Riesgo de liquidez.** Se define como la pérdida potencial por la imposibilidad o dificultad de renovar pasivos o de contratar otros en condiciones normales por la venta anticipada o forzosa de activos a descuentos inusuales para hacer frente a las obligaciones, o bien, por el hecho de que una posición no pueda ser oportunamente enajenada, adquirida o cubierta mediante el establecimiento de una posición contraria equivalente.
- **Riesgo operativo.** Se define como la pérdida potencial por fallas o deficiencias en los controles internos, por errores en el procesamiento y almacenamiento de las operaciones o en la transmisión de información, así como por resoluciones administrativas y judiciales adversas, fraudes o robos, y comprende, entre otros, al riesgo tecnológico y al riesgo legal. En cumplimiento con las disposiciones emitidas por la CNBV (Comisión Nacional Bancaria y de Valores), para la gestión del riesgo operativo, incluyendo el tecnológico y legal, se definieron categorías de riesgos las cuales están sustentadas sobre los criterios previstos en los controles internos.
- **Riesgo de Crédito.** Pérdida potencial por falta de pago de un acreditado o contraparte en las operaciones que efectúan las instituciones, incluyendo las garantías reales o personales que les otorguen, así como cualquier otro mecanismo de mitigación utilizado por las instituciones.

El riesgo de crédito asociado a un acreditado o contraparte se da en función de los siguientes aspectos:

1. El monto de la exposición de riesgo de la contraparte.

2. La probabilidad de incumplimiento.
3. La potencial recuperación de las garantías pactadas al inicio de la transacción.

Es este último tipo de riesgo precisamente el cual se aborda en el presente trabajo.

## 1.2. Tipos de crédito

Definición de los tipos de crédito existentes desde el punto de vista de riesgo crediticio:

- Crédito hipotecario. Contrato por el que una entidad financiera concede una línea de financiamiento o préstamo hipotecario a un titular, de la que éste puede ir disponiendo cantidades, según sus necesidades. Es dinero que entrega el banco o financiera para adquirir una propiedad ya construida, un terreno, la construcción de viviendas, oficinas y otros bienes raíces, con la garantía de la hipoteca sobre el bien adquirido o construido; normalmente es pactado para ser pagado en el mediano o largo plazo (8 a 40 años, aunque lo habitual son 20 años).
- Créditos revolventes. Préstamo a corto plazo establecido por una cantidad fija y que tiene una vigencia que puede ir de uno a tres años, del cual se puede disponer en efectivo en el momento en que el deudor lo requiera. Está diseñado para optimizar la disponibilidad del capital. Esta deuda puede ser pagada por el cliente de manera diferida en el tiempo. Una tarjeta de crédito es un buen ejemplo de un crédito revolvente, ya que cuando pagas tu deuda vuelves a tener saldo disponible en la tarjeta.
- Créditos de servicios. Préstamo a corto plazo para el pago de servicios: agua, luz, telefonía, etcétera.

- Crédito personal. Es aquel que se otorga para un uso libre del dinero solicitado; por ejemplo para un viaje, comprar una televisión, cancelar deudas pendientes, etcétera. El banco o entidad financiera por lo general solicita menos requisitos para gestionarlo; inclusive existen créditos personales preacordados en donde un cliente bancario puede disponer de un determinado dinero cuando lo desee, sin presentar ningún tipo de documentación al momento de realizar la operación. Este tipo de créditos suele tener un costo financiero total más elevado que los créditos prendarios o créditos hipotecarios y el plazo de cancelación no supera las 60 cuotas (5 años).

### **1.3. Ciclo de la administración del riesgo en crédito al consumo**

El objetivo de la administración del riesgo en crédito al consumo es maximizar la tasa de rendimiento ajustada por el riesgo en el cual incurre el acreedor, manteniendo la exposición al riesgo de crédito dentro de los límites aceptables. Una eficaz administración del riesgo de crédito resulta de vital importancia para el éxito a largo plazo de cualquier organización que otorgue el mismo. Evidentemente, dicha organización tendrá que administrar el riesgo de crédito inherente en su cartera y el riesgo en créditos otorgados. Asimismo, deberá tener presente la necesidad de identificar, medir, monitorear y controlar el crédito, determinando si su capital es adecuado para hacer frente a estos riesgos. Mediante la obtención de la probabilidad de incumplimiento y del monto esperado se obtiene la pérdida esperada de un crédito. Finalmente, la administración del riesgo en toda área de riesgo de crédito se basa en el “ciclo de riesgo”, el cual es detallado en el capítulo 3.

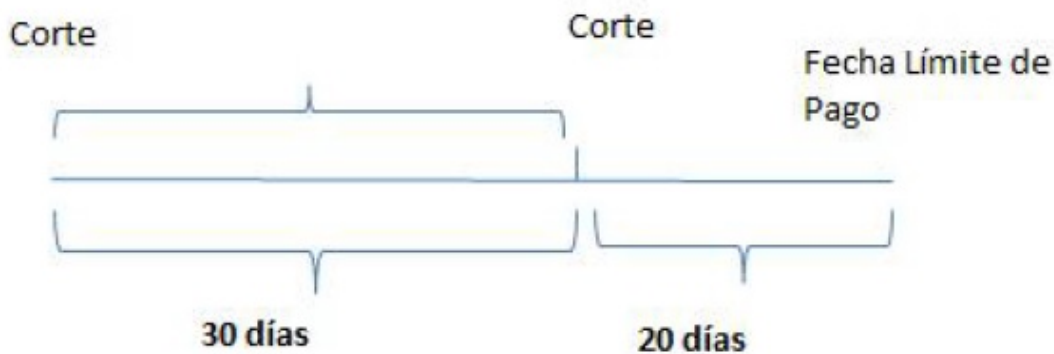


## 1.4. Modelos preventivos de pérdidas

Se puede definir un modelo de prevención de pérdidas en este contexto, como una herramienta que ayuda a pronosticar la probabilidad de que un cliente, estando totalmente vigente en sus pagos, al momento de ser evaluados los mismos, llegada su fecha límite de pago incumpla con sus obligaciones previamente pactadas en el contrato de crédito.

Es importante identificar este tipo de clientes, pues de manera temprana se detecta su comportamiento de pago que en un futuro puede causar pérdidas a la institución. Como un ejemplo de estos modelos mencionaremos el caso de la operación de una tarjeta de crédito. En una tarjeta de crédito existen dos fechas importantes:

- Fecha límite de pago. Fecha en la cual un cliente tiene que pagar los saldos generados al corte.
- Fecha de corte. Fecha en la que se cierra la evaluación del estado de cuenta. Generalmente la fecha límite de pago es 20 días después de la fecha de corte.



Es así que el modelo preventivo de pérdidas calcula la probabilidad de que un cliente que acaba de cortar, llegada su fecha límite de pago, no realice el mismo, mediante un modelo que denominaremos como "Prevent Score".



# Capítulo 2

## Ciclo del riesgo

El ciclo de riesgo es la forma en la cual se puede resumir la administración de riesgo en crédito al consumo, pues resume perfectamente el conjunto de acciones que hay que seguir para administrar pérdidas por crédito. Es así que puede aplicarse a la administración de cualquier tipo de crédito enfocado al consumo. Finalmente el ciclo de riesgo consta de 3 partes fundamentales:

1. Originación del crédito.
2. Administración de portafolio.
3. Recuperación de pérdida por crédito y origen.



Veamos más a detalle cada componente.

## 2.1. Origen del crédito

En esta parte del ciclo de riesgo nos preguntamos a qué cliente le daremos crédito, dado que nunca ha sido nuestro cliente, y para ello hacemos una selección de los “mejores clientes”, de acuerdo a sus características demográficas y de buró de crédito, que se terminan convirtiendo en una probabilidad de incumplimiento de pago. El cliente deberá cumplir con una serie de requisitos para que le pueda ser otorgado el crédito. La institución que otorgue créditos deberá poseer su propia política de otorgamiento de créditos con base en los estatutos de la Comisión Nacional Bancaria y de Valores. Las características demográficas del cliente hacen referencia a los datos generales del cliente (nombre, edad, estado civil, domicilio particular, dirección de su empleo, ingresos, antigüedad, tanto en su domicilio como en su trabajo, referencias de familiares, amigos o vecinos, etcétera). Dependiendo de la institución, los requisitos podrán extenderse, como la necesidad de un aval, muestra de algún tipo de garantía, etcétera. La consulta del buró de crédito del cliente permite conocer la solvencia moral de las personas, lo cual promoverá la minimización del riesgo crediticio.

## 2.2. Administración de portafolio

En esta parte del ciclo de riesgo nos preguntamos a qué clientes se les va a dar más crédito, dado que ya forman parte de nuestra cartera de crédito; es decir, ya que tienen un comportamiento de pago con la institución. Una vez seleccionados los clientes, éstos ya forman parte de nuestro portafolio. Ahora, el siguiente paso es “la Administración del Portafolio”, proceso de análisis del cliente que lleva a tomar dos principales medidas:

1. Otorgar un mayor crédito.
2. Disminuir el crédito.

Este proceso está basado en la premisa: A los clientes con un buen comportamiento crediticio habrá que darles más crédito y a los clientes con un comportamiento no satisfactorio habrá que disminuirles o cancelarles el mismo. La recuperación de pérdidas derivadas del otorgamiento de créditos es una labor de vital importancia para el rescate de cuentas pendientes de pago. Toda institución que otorgue créditos está sujeta al riesgo de que algunos clientes se nieguen a cumplir con sus obligaciones crediticias, por lo cual la cobranza es un proceso elemental, tanto para prevenir pérdidas, como para recuperar las cuentas que no hayan pagado. Para garantizar un buen trabajo de cobranza y un adecuado control de las cuentas debe llevarse un estricto registro de aquellas que han permanecido insolutas, excediendo las condiciones normales de venta y que ya están vencidas. La antigüedad de cuentas significa su separación en diferentes categorías.

### **2.3. Gestión de cobranza**

En esta parte del ciclo de riesgo nos preguntamos qué cliente nos va a pagar, dado que actualmente no está pagando. Una de las actividades vitales en la vida de una empresa, sea cual sea su giro es, sin duda, la cobranza. Representa el corazón que bombea los recursos financieros para la subsistencia de la misma, por tal razón, debe ejercerse con profesionalismo.

La cobranza representa el final del ciclo de venta del crédito; obteniendo el pago del cliente se logra su culminación.

#### **Tipos principales de cobranza**

- Cobranza educativa. Como su nombre lo indica, su objetivo principal es ir educando al cliente sobre el adecuado y correcto manejo de su crédito. Se usa en clientes nuevos sin atraso o con pocos días de mora. Una técnica recurrente utilizada aquí es hacer llamadas de bienvenida y servicio a los clientes, recordándoles su fecha de pago y los estímulos y castigos derivados de su comportamiento crediticio.

- Cobranza preventiva. Una gran cantidad de clientes que no pagan a tiempo, dejan de hacerlo no porque no lo recordaron o no le dieron importancia al adeudo; sino por que no fueron recordados del mismo en el momento correcto.
- Cobranza correctiva o extrajudicial. Cuando un cliente incurre en mora (retraso en sus pagos), es imprescindible tomar acciones de manera inmediata para corregir el rumbo, pues mientras más atraso tenga, menor será el porcentaje de recuperación de la cartera (los clientes se mudan, pagan otros adeudos y se quedan sin bienes para garantizar el adeudo, etcétera). Tener una estrategia de cobranza que involucre múltiples vías de contacto incluyendo llamadas telefónicas, cartas, visitas, es indispensable para que los deudores paguen a tiempo.
- Cobranza judicial. En el caso de que el cliente decida no pagar en los tiempos acordados (previa negociación extrajudicial), lo mejor es proceder a la cobranza judicial mediante un juicio mercantil. La decisión debe tomarse pronto, pues los pagarés y otros documentos del cliente tienen una vigencia determinada y si llegan a vencer, difícilmente podrán utilizarse en juicio.

### **Principales causas que originan la morosidad:**

- Falta de cobranza preventiva (anticiparse al vencimiento).
  1. Créditos mal soportados en cuanto a garantías.
  2. Rebasar capacidad de pago en el otorgamiento.
  3. Otorgamiento de créditos inadecuados (mal investigados y no verificados).
- Créditos otorgados muy alejados de la sucursal, que limiten el desplazamiento de clientes para pagar o para su cobranza.
- Eventos de crisis económica o de la naturaleza.
- Derivados del cliente: Mala organización, olvido, enfermedad, mala cultura de pago, etcétera.

- Gestiones de cobranza deficientes: Falta de seguimiento como: maltrato, comunicación inadecuada, etcétera.

En este trabajo me enfocaré a la cobranza preventiva, pues sólo se requiere ver qué clientes, estando vigentes en sus pagos van a “empezar” a incumplir, llegada su fecha límite de pago.

Este ciclo se dice que es de mejora continua, pues constantemente una institución financiera está originando clientes, administrando su portafolio de clientes y, por tanto, cobrando. La idea es que mediante modelos de riesgo gestionemos mejor el riesgo en todo el ciclo.

En todos los casos la respuesta será dicotómica, es por ello que se eligió emplear una regresión logística, ya que se ajusta adecuadamente al análisis.

## 2.4. Modelo de riesgo

Para cada parte del ciclo de riesgo existen, en términos generales, diferentes modelos de riesgo:

1. **Originación.** El score de originación también conocido como Acquisition Score, o Modelo de riesgo de originación, es una estimación de la conveniencia general de la institución que otorgue el crédito sobre la base de la información que posee acerca del comportamiento que han tenido sus clientes, y en base a las características de los mismos va a determinar cómo seleccionar a sus nuevos clientes. Generalmente se consideran variables demográficas y de buró de crédito. Para cada cliente, la suma de cada puntaje asignado a cada variable analizada por la institución será el puntaje final del mismo, y en base a los puntos de corte determinados, se decidirá si se le da el crédito o no.

Con mayor información sobre el acreditado se puede tener una medida más precisa de la probabilidad de incumplimiento. Si se establece qué créditos con ciertas características tienen un determinado puntaje, se puede construir la

tabla de puntaje para todas las combinaciones posibles y así obtener el Score Card.

2. Administración del Portafolio. El Score de Comportamiento, también conocido como Behavior Score o Modelo de riesgo de comportamiento, predice la probabilidad de que un producto ya otorgado se convierta en parte de la cartera morosa de la institución. Este score asigna un nivel de riesgo en el proceso de seguimiento del crédito en base a la información crediticia contenida en las bases de datos de la institución. Sus beneficios son que permite precisar estrategias y controlarlas, aumenta la satisfacción del cliente, facilita las decisiones para realizar ventas cruzadas y reduce los gastos de cobranzas y de autorizaciones.
3. Cobranza. El Score de Cobranza, también conocido como Collection Score o Modelo de riesgo de cobranza. El Score de Cobranza permite identificar qué clientes tienen mayor probabilidad de pagar sus deudas. Gracias a este Score se pueden implementar estrategias distintas para los distintos segmentos de clientes morosos, además de optimizar la inversión en cobranza y aumentar evidentemente la tasa de recuperabilidad. El Score de Cobranza constituye una oportunidad de mejora para la labor que realizan los departamentos de gestión de cobranza de las instituciones; ya que su objetivo es encaminar las estrategias persuasivas de cobranza a segmentos de cartera caracterizados por tener un riesgo de pago, por lo tanto, repercute positivamente en la eficacia de la recuperación. Esta herramienta cuantitativa (puntuación o Score de clientes) se basa en el comportamiento histórico de pagos, variables demográficas y socioeconómicas de los clientes, proporcionando lineamientos y prioridades respecto de determinaciones de cobro e incidiendo positivamente en la eficacia del recaudo.

En cuanto a los modelos de riesgo de cobranza, podemos encontrar sólo “modelos” que predicen diferentes niveles de morosidad:

- Modelos de Prevención de Pérdidas (Prevent Score). Pronostica si un cliente vigente en su fecha límite de pago caerá en mora.



- Modelos de cobranza temprana. Pronostica la probabilidad de que el estado de mora de un cliente pase de mora temprana (entre 1 y 29 días), a un rango de 30 a 59 días de retraso en el pago.
- Modelos de cobranza avanzada (30 a 120 días de mora).

En el presente trabajo me enfocaré en el score preventivo de pérdidas financieras.



# Capítulo 3

## Bases técnicas actuariales. Tópicos estadísticos a utilizar

En este capítulo se abordarán los tópicos estadísticos requeridos en la elaboración de un modelo estadístico para pronosticar la probabilidad de que un cliente vigente reporte un deterioro en su comportamiento de pago en fecha límite.

### 3.1. Base de datos

Una base de datos es un conjunto de información almacenada en memoria auxiliar que permite acceso directo y un conjunto de programas que manipulan dichos datos. Hoy en día, el uso de bases de datos permite la manipulación de información de manera eficaz y ágil.

Una base de datos contiene entidades de información que están relacionadas vía organización y asociación. La arquitectura lógica de una base de datos se define mediante un esquema que representa las definiciones de las relaciones entre las entidades de información. La arquitectura física de una base de datos depende de la configuración del hardware residente. Sin embargo, tanto el esquema (descripción lógica) como la organización (descripción física) deben adecuarse para satisfacer los

requerimientos funcionales y de comportamiento para el acceso al análisis y creación de informes.

Las ventajas de las bases de datos son:

- Coherencia de los resultados.
- Mayor disponibilidad de los datos para el conjunto de los usuarios.
- Mayor valor informativo (puesto que se captan relaciones propias del mundo real entre los datos, tiene más información que la suma individual de los datos por separado).
- Mayor eficiencia en la recopilación, validación e introducción de los datos en el sistema.
- Reducción del espacio de almacenamiento.
- Limpieza de una base de datos. La limpieza de una base de datos es fundamental para el correcto análisis de los datos a estudiar. Se realiza para eliminar los registros duplicados, corregir los datos erróneos y, de esta manera, agilizar el análisis. Limpiar la base de datos a utilizar para la creación de un modelo estadístico es de vital importancia para que el pronóstico sea lo más certero posible.

Para el desarrollo del modelo propuesto en el presente trabajo se empleó fundamentalmente el manejo de bases de datos de una institución financiera.

Para el manejo de la información hubiera resultado imposible recurrir a la aplicación Excel, por su limitada capacidad para el almacenamiento de datos, ya que el volumen empleado fue muy extenso. Evidentemente, antes de la creación del modelo estadístico se limpió la base de datos proporcionada por dicha institución y, para ello, se empleó la aplicación informática SAS (Statistical Analysis System), paquete especialmente diseñado para resolver problemas en el área de la estadística, misma herramienta empleada en todo el desarrollo del modelo propuesto.

Para este trabajo se emplearon dos fuentes de información de la institución analizada:

- **Originación:** La base de originación constituye una recopilación de todas las solicitudes del crédito. Contiene toda la información que se requiere respecto al cliente para otorgarle un crédito: características demográficas, de comportamiento y de buró de crédito. La base de originación está conformada por solicitudes aprobadas y no aprobadas. Para efectos de este trabajo se emplearon únicamente los créditos dados de alta, porque se requirió observar su comportamiento crediticio.
- **Comportamiento de pago.** La base del comportamiento de pago nos muestra la información del saldo de los montos entregados a los acreditados, más los intereses devengados no cobrados, además de la mora al día por cada cliente. Para efectos de este trabajo, esta base permite conocer el comportamiento crediticio con la institución.

## 3.2. Estadística

*La estadística es una ciencia formal que estudia la recolección, análisis e interpretación de datos de una muestra representativa, ya sea para ayudar en la toma de decisiones o para explicar condiciones regulares o irregulares de algún fenómeno o estudio aplicado, de ocurrencia en forma aleatoria o condicional.*

La estadística se ocupa de los métodos científicos para recolectar, organizar, resumir, presentar y analizar datos, así como de obtener conclusiones válidas y tomar decisiones con base en el análisis de datos. La estadística consta de las fases de obtención de información, de análisis y de presentación e interpretación de los resultados y elaboración de métodos.

Al recolectar datos que determinan las características de un grupo de individuos u objetos, por ejemplo, refiriéndonos al tema central del presente trabajo; “Prevenir

pérdidas en crédito” al consumo; los datos serán: la edad del cliente, la fecha de apertura del crédito, la periodicidad de sus pagos, el monto solicitado, el monto otorgado, el número de recibos vencidos del cliente, etcétera; es decir, todas las características del cliente que pueden ser útiles para analizar su comportamiento. Resultaría impráctico analizar todos los clientes, especialmente si el número de créditos otorgados por una organización es grande. En lugar de examinar todo este grupo de clientes se examinará una parte de éstos, es decir, una muestra representativa del grupo. Cuando una muestra es representativa, el análisis puede dar conclusiones importantes acerca de la población. Para analizar los datos que se van a requerir para la elaboración del modelo propuesto se requiere del dominio de los siguientes tópicos.

### 3.2.1. Estadística paramétrica

La estadística paramétrica es una rama de la estadística que comprende los procedimientos estadísticos y de decisión, que están basados en las distribuciones de los datos reales, las cuales son determinadas usando un número finito de parámetros. Cuando se desconoce totalmente qué distribución siguen los datos, deberá aplicarse primero un test no paramétrico, que permita conocer la distribución.

La mayoría de los procedimientos paramétricos requiere conocer la forma de distribución para las mediciones resultantes de la población estudiada.

La estadística paramétrica clásica plantea tres tipos de problemas:

1. Estimación puntual: en la cual se pretende dar un valor al parámetro a estimar.
2. Estimación por intervalos: en la cual se busca un intervalo de confianza.
3. Contrastes de hipótesis: donde se busca contrastar información acerca del parámetro.

### Análisis de regresión

El análisis de regresión es parte de la estadística paramétrica y busca encontrar modelos a partir de las relaciones causales entre una variable dependiente y una o más

variables independientes, que permita pronosticar un resultado conociendo el valor estimado de una variable independiente. Una de las dos variables que llamaremos  $x$  (variable dependiente), puede ser medida sin error apreciable; mientras que la otra variable,  $Y$  (variable independiente), es una variable aleatoria. Para el análisis será necesario enfocarnos en la dependencia de  $Y$  en términos de  $x$ .

Se considerará el experimento de seleccionar en primera instancia  $n$  valores  $x_1, \dots, x_n$  de  $x$ , y luego observar los valores de  $Y$  que correspondan a dichos valores  $x$ , de tal manera que podamos obtener una muestra de la forma:

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

En el análisis de regresión, se supone que la media  $\mu$  de  $Y$  depende de  $x$ , es decir, que es una función  $\mu = \mu(x)$ . La curva  $\mu(x)$  se llama curva de regresión de  $Y$  con base en  $x$ . En su caso más sencillo, podemos expresarla como sigue:  $\mu(x) = \alpha + \beta(x)$ .

Le llamamos recta de regresión de  $Y$  con base en  $x$ , y a la pendiente  $\beta$  se le llama coeficiente de regresión.

Cualquier función  $\mu(x)$  se puede aproximar con suficiente exactitud mediante una función lineal si  $x$  varía en un intervalo corto. La muestra puede emplearse para estimar los parámetros  $\alpha$  y  $\beta$ , mediante el método de mínimos cuadrados.

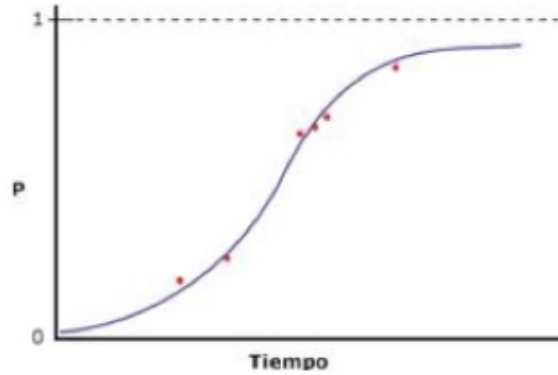
Partamos de las siguientes suposiciones:

- (A<sub>1</sub>) Para cada  $x$  fija la variable  $Y$  es normal con media  $\alpha + \beta(x)$  y varianza  $\sigma^2$  (no dependiente de  $x$ ).
- (A<sub>2</sub>) Las  $n$  ejecuciones del experimento que proporciona la muestra son independientes.

### **Análisis de regresión logística**

Los modelos de regresión logística son modelos de regresión que permiten estudiar si una variable binomial depende, o no, de otra u otras variables (no necesariamente

binomiales). Si una variable binomial de parámetro  $p$  es independiente de otra variable  $X$ , se cumple  $p = p(X)$ , por consiguiente, un modelo de regresión es una función de  $p$  en  $X$  que a través del coeficiente de  $X$  permite investigar la relación anterior.



En estadística, la regresión logística es un modelo de regresión para variables dependientes o de respuesta binomialmente distribuidas. Es útil para modelar la probabilidad de un evento ocurriendo como función de otros factores. La regresión logística es usada extensamente en las ciencias médicas y sociales.

### Etapas del modelo de regresión logística

1. Formulación del modelo de regresión:

$$y = \beta_0 + \beta_i x + u$$

2. Estimación de parámetros. Se estiman los parámetros mediante el método de mínimos cuadrados:

- Calculamos la esperanza en  $y = \beta_0 + \beta_i x + u$  para  $x = x_i$ :

$$E[y|x_i \beta_0 + \beta_i x_i]$$

- Llamamos  $p_i$  a la probabilidad de que  $y$  tome el valor 1 cuando  $x = x_i$ :

$$p_i = P(y = 1|x_i)$$



- la esperanza de  $y$  es:

$$E[y|x_i]P(y = 1|x_i) * 1 + P(y = 0|x_i) * 0 = p_i$$

- $p_i = \frac{1}{1+e^{-\beta_0-\beta_i x_i}}$

Entonces la predicción  $y_i$  estima la probabilidad de que un individuo con características definidas por  $x = x_i$  pertenezca a la población correspondiente a  $y = 1$ .  $\beta_0 + \beta_i x_i$  garantiza esta restricción, ya que el modelo podría arrojar probabilidades mayores a 1; sin embargo, maximiza la separación entre los grupos, independientemente de la distribución de los datos.

Si nuestro objetivo es estimar directamente la probabilidad de pertenecer a cada uno de los grupos, debemos transformar la variable respuesta de manera que se garantice un resultado de cero o uno.

Como:

$$p_i = \frac{1}{1 + e^{-\beta_0 - \beta_i x_i}}$$

es una función continua, entonces:

$$\begin{aligned} g_i &= \log\left(\frac{p_i}{1-p_i}\right) = \log\left(\frac{\frac{1}{1+e^{-\beta_0-\beta_i x_i}}}{1 - \frac{1}{1+e^{-\beta_0-\beta_i x_i}}}\right) \\ &= \log\left(\frac{\frac{1}{1+e^{-\beta_0-\beta_i x_i}}}{\frac{e^{-\beta_0-\beta_i x_i}}{1+e^{-\beta_0-\beta_i x_i}}}\right) = \log\left(\frac{1}{e^{-\beta_0-\beta_i x_i}}\right) \\ &= -\log(e^{-\beta_0-\beta_i x_i}) = -(-\beta_0 - \beta_i x_i) = \beta_0 + \beta_i x_i. \end{aligned}$$

Por lo que al hacer la transformación obtendremos un modelo lineal. La variable  $g$  representa, en una escala logarítmica, la diferencia entre las probabilidades de pertenecer a ambas poblaciones y al ser una función lineal de las variables explicativas, facilita la estimación y la interpretación del modelo.

En el caso de la regresión logística, que es una situación en la que la variable de respuesta sólo tiene dos resultados posibles, llamados en forma genérica, “éxito” y

“fracaso” representados por 0 y 1, la respuesta es en esencia cualitativa, porque los calificativos “éxito” y “fracaso” son totalmente arbitrarios. Por ejemplo, la respuesta podría ser el resultado de una prueba de funcionamiento eléctrico para un dispositivo semiconductor, que da como resultado un “éxito” que indica que el dispositivo trabaja bien, o un “fracaso” que podría deberse a un corto, un circuito abierto u otro problema de funcionamiento.

Supóngase que el modelo tiene la forma:

$$Y_i = x'_i \beta + \varepsilon_i.$$

En donde  $x'_i = (1, x_{i_1}, \dots, x_{i_k})$ ,  $\beta = (\beta_0, \dots, \beta_k)$  y la variable de respuesta  $y_i$  toma los valores 0 ó 1. Se supondrá que la variable de respuesta  $y_i$  es una variable aleatoria de Bernoulli. Sea  $P(y_1 = 1) = p_i$  y  $P(y_1 = 0) = 1 - p_i$ . Ahora bien, como  $E(\varepsilon_i) = 0$ , el valor esperado de la variable de respuesta es:

$$E(y_i) = 1(p_i) + 0(1 - p_i) = p_i.$$

Esto implica que:

$$E(y_i) = x'_i \beta = p_i.$$

De donde la respuesta esperada, determinada con la función de respuesta  $E(y_i) = x'_i \beta$ , no es más que la probabilidad de que la variable de respuesta tenga el valor 1.

### Estimación de parámetros en un modelo de regresión logística

La forma general del modelo de regresión logística es:

$$Y_i = E(y_i) + \varepsilon_i.$$

Donde las observaciones  $y_i$  son variables aleatorias independientes de Bernoulli, cuyos valores esperados son:

$$E(y_i) = p_i = \frac{e^{x'_i \beta}}{1 + e^{x'_i \beta}}.$$

Se usará el método de mínimos cuadrados como se refirió anteriormente para estimar los parámetros del predictor lineal  $x'_i \beta$ .

Cada observación de la muestra sigue la distribución de Bernoulli, por lo que la distribución de probabilidades de cada observación es:

$$f_i(y_i) = p_i^{y_i} (1 - p_i)^{1-y_i}$$

con  $i = 1, \dots, n$  y cada observación  $y_i$  toma el valor 0 ó 1. Como las observaciones son independientes, la función verosimilitud no es más que:

$$L(y_i, y_1, \dots, y_n, \beta) = \prod_{i=1}^n f_i(y_i) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i}.$$

Aplicando logaritmo tenemos:

$$\ln(L(y_i, y_1, \dots, y_n, \beta)) = \ln\left(\prod_{i=1}^n f_i(y_i)\right) = \sum_{i=1}^n y_i \ln\left(\frac{p_i}{1-p_i}\right) + \sum_{i=1}^n \ln(1-p_i).$$

Como  $1 - p_i = (1 + e^{x'_i \beta})^{-1}$  y  $\eta_i = \ln\left(\frac{p_i}{1-p_i}\right) = x'_i \beta$ , el logaritmo de la verosimilitud se puede expresar como sigue:

$$\ln(L(y, \beta)) = \sum_{i=1}^n y_i \eta_i - \sum_{i=1}^n \ln(1 + e^{\eta_i}).$$

Con frecuencia en los modelos de regresión logística se tienen observaciones o intentos repetidos en cada nivel de las variables  $x$ , esto sucede mucho en los experimentos diseñados. Sea  $y_i$  la cantidad de 1 observado en  $i$  y  $\eta_i$  la cantidad de intentos en cada observación. Entonces, el logaritmo de la verosimilitud se transforma en:

$$\ln(L(y, \beta)) = \sum_{i=1}^n y_i p_i + \sum_{i=1}^n \eta_i \ln(1 - p_i) - \sum_{i=1}^n y_i \ln(1 - p_i).$$

Se podrían usar métodos numéricos de búsqueda para calcular los estimados  $\hat{\beta}$  por máxima verosimilitud. Sea  $\hat{\beta}$  el estimado final de los parámetros del modelo. Entonces se puede demostrar que  $E(\hat{\beta}) = \beta$  y  $Var(\hat{\beta}) = (X'V^{-1}X)^{-1}$ .

El valor estimado del predictor lineal es  $\hat{\eta}_i = x'_i \hat{\beta}$  y el valor esperado del modelo de regresión logística se escribe como sigue:

$$\hat{Y}_i = \hat{p}_i = \frac{e^{\hat{\eta}_i}}{1 + e^{\hat{\eta}_i}} = \frac{1}{1 + e^{-x'_i \hat{\beta}}}.$$

Ahora bien, la prueba de hipótesis en la regresión logística se basa en pruebas de cociente de máxima verosimilitud, que es un procedimiento para muestras grandes, por lo que los procedimientos de prueba se basan en la teoría asintótica. El método de la razón de verosimilitud conduce a un estadístico llamado devianza.

La desviación del modelo compara el logaritmo de la verosimilitud del modelo ajustado con el logaritmo de la verosimilitud de un modelo saturado, que es un modelo que tiene exactamente  $n$  parámetros y se ajusta perfectamente a los datos de la muestra. Para el modelo de regresión logística eso significa que las probabilidades  $p_i$  son totalmente irrestrictas, lo que al igualar  $p_i = y_i$  con  $y_i = 0, 1$  se maximizará la verosimilitud. El valor máximo de la función logaritmo de verosimilitud es cero.

### Selección de parámetros

Existen varios métodos para construir el modelo de regresión, es decir, para seleccionar de entre todas las variables que introducimos en el modelo cuáles son las que se requieren para explicarlo. El modelo de regresión se puede construir utilizando las siguientes técnicas:

- Técnica de pasos hacia adelante (Forward): Consiste en ir introduciendo las variables en el modelo, únicamente si cumplen una serie de condiciones, hasta que no se pueda introducir ninguna más debido a que ninguna cumpla la condición impuesta.
- Técnica de pasos hacia atrás (Backward): Se introducen en el modelo todas las variables y se van suprimiendo si cumplen una serie de condiciones definidas a priori, hasta que no se puedan eliminar más, es decir, hasta que ninguna variable cumpla la condición impuesta.
- Técnica por pasos (Stepwise): Combina los dos métodos anteriores, adelante y atrás, introduciendo o eliminando variables del modelo si cumplen una serie de condiciones definidas a priori, hasta que ninguna variable satisfaga ninguna de las condiciones expuestas de entrada o salida del modelo.

- Técnica de introducir todas las variables obligatoriamente (Enter): Esta técnica de selección de variables para construir el modelo de regresión produce que el proceso de selección de las variables sea manual, partiendo de un modelo inicial en el que se obliga a que entren todas las variables seleccionadas se va evaluando qué variable es la que menos participa en él y se elimina, volviéndose a construir un nuevo modelo de regresión aplicando la misma técnica, pero excluyendo la variable seleccionada y aplicando el mismo proceso de selección. Este proceso se repite reiteradamente, hasta que se considere que el modelo obtenido es el que mejor se ajusta a las condiciones impuestas y que no se puede eliminar ninguna variable más de las que lo componen.

### **3.2.2. Estadística no paramétrica**

La estadística no paramétrica es una rama de la estadística que estudia las pruebas y modelos estadísticos cuya distribución subyacente no se ajusta a los llamados criterios paramétricos. Su distribución no puede ser definida a priori, pues son los datos observados los que la determinan. La utilización de estos métodos es recomendable cuando no se puede asumir que los datos se ajusten a una distribución conocida, cuando el nivel de medida empleado no sea, como mínimo, de intervalo.

A continuación se presenta la prueba de Kolgomorov Smirnov, ya que más adelante se empleará para mostrar una aplicación práctica del modelo, la recalibración de un score y simultáneamente el impacto del modelo en la sociedad financiera de objeto múltiples SOFOM, para incrementar la diferencia entre las dos poblaciones Goods y Bads.

"Goods": Clientes con buen comportamiento crediticio.

"Bads": Clientes con mal comportamiento crediticio.

Dos conceptos ya previamente establecidos por la institución con base en el score de los clientes.

- Kolgomorov Smirnov. Prueba no paramétrica empleada para determinar la bondad de ajuste de dos distribuciones de probabilidad entre sí. La prueba de

Kolmogorov-Smirnov es apropiada únicamente para distribuciones continuas. La hipótesis a probar es que cierta función  $F(x)$  es la función de distribución de una población de la que se ha tomado una muestra  $x_1, \dots, x_n$ .

Los pasos de la prueba de Kolmogorov Smirnov se muestran a continuación:

1. Calcular los valores de la función de distribución  $F(x)$  de la muestra  $x_1, \dots, x_n$ .
2. Determinar la desviación máxima  $\alpha$ .
3. Escoger un nivel de significancia  $\alpha$  (5 %, 1 % o uno semejante).
4. Determinar la solución  $c$  de la ecuación:

$$P(a \leq -c) = \alpha.$$

Si  $\alpha \leq -c$ , no se rechaza la hipótesis. Si  $a > c$ , se rechaza la hipótesis.

### Ejemplo

La siguiente tabla muestra la prueba de Kolmogorov-Smirnov. En primera instancia aparecen las diferencias más extremas (absoluta, negativa y positiva) entre las funciones de distribución de ambas muestras. Y a continuación se muestra el resultado de la tipificación de la diferencia más extrema en valor absoluto ( $Z$  de Kolmogorov- Smirnov = 2.134 ) junto con el nivel crítico bilateral (significación asintótica bilateral = 0). Puesto que el nivel crítico es menor que 0.05, podemos rechazar la hipótesis de desigualdad de distribuciones y concluir que los dos grupos comparados difieren significativamente en salario inicial.

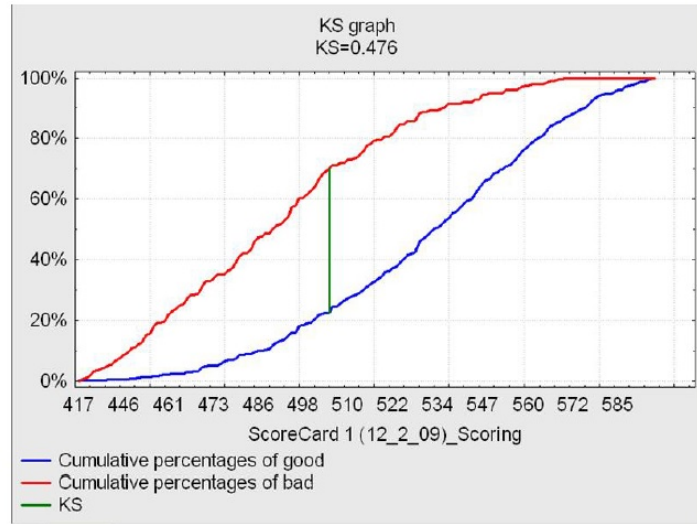
## Prueba de Kolmogorov-Smirnov

		Salario inicial
Diferencias más extremas	Absoluta	,237
	Positiva	,000
	Negativa	-,237
Z de Kolmogorov-Smirnov		2,134
Sig. asintót. (bilateral)		,000

La siguiente tabla recoge el resultado de la prueba de rachas. La tabla ofrece el número mínimo y máximo del número de rachas, el valor del estadístico  $Z$  en cada supuesto y el nivel crítico unilateral (significación asintótica unilateral). La presencia de 25 empates hace que el resultado obtenido en ambas situaciones sea muy diferente. En estos casos puede recurrirse a otras pruebas para tomar una decisión.

Salario inicial			
	Número de rachas	Z	Sig. asintót. (unilateral)
Mínimo posible	40 <sup>a</sup>	-16,576	,000
Máximo posible	200 <sup>a</sup>	4,923	1,000

a. Hay 25 empates inter-grupos que implican 348 casos.



## Divergencia

*La divergencia de la información es una medida no simétrica de la similitud o diferencia entre dos funciones de distribución de probabilidad  $P$  y  $Q$ .*

Prueba no paramétrica que mide la efectividad de cómo el scorecard asignado a una serie de variables de una población analizada separa dicha población en dos grupos. Para efectos de este trabajo consideremos estos dos grupos; clientes “buenos” y clientes “malos”, entendiéndose como cliente “BUENO”, aquel que tiene un buen comportamiento crediticio, correspondiéndole un determinado puntaje a cada variable analizada en el scorecard, y como cliente “MALO”, aquel que tiene un mal comportamiento crediticio, correspondiéndole un determinado puntaje a cada variable analizada en el scorecard.

En cuanto mayor sea la divergencia, mayor será la separación entre la media de la población de clientes malos y clientes buenos. A continuación se muestra cómo calcular la divergencia  $D$ :

$$D = \frac{2(\mu_g - \mu_b)^2}{\zeta_g^2 - \zeta_b^2}.$$



Donde:

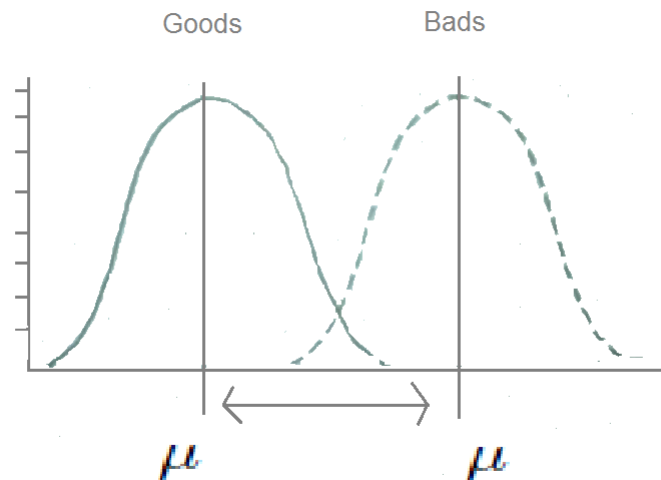
$\mu_g$  = Media de los clientes buenos.

$\mu_b$  = Media de los clientes malos.

$\zeta_g$  = Desviación estándar de los clientes buenos.

$\zeta_b$  = Desviación estándar de los clientes malos.

Gráficamente, la divergencia entre dos poblaciones (“BUENOS” y “MALOS”), donde  $\mu$  representa la media de la distribución se ve de la siguiente forma:





# Capítulo 4

## Construcción del modelo de prevención de pérdidas

### 4.1. Tipos y formas de construir modelos

Como ya vimos, existe un modelo de riesgo para cada fase del ciclo de riesgo. Se pueden identificar tres formas de construir los modelos:

1. Árboles de decisión: Como su nombre lo indica, son una “herramienta de decisión” basada en la “llave de decisión” y clasificación de la población, que a su vez forman “nodos de ramas”. Un nodo se puede definir como la intersección de una llave de decisión y la clasificación de la población para esa llave. Una rama es el resultado final. Las premisas de un árbol de decisión son que al menos tiene que haber una llave de decisión y la clasificación de la población debe superar el 5% de la población.

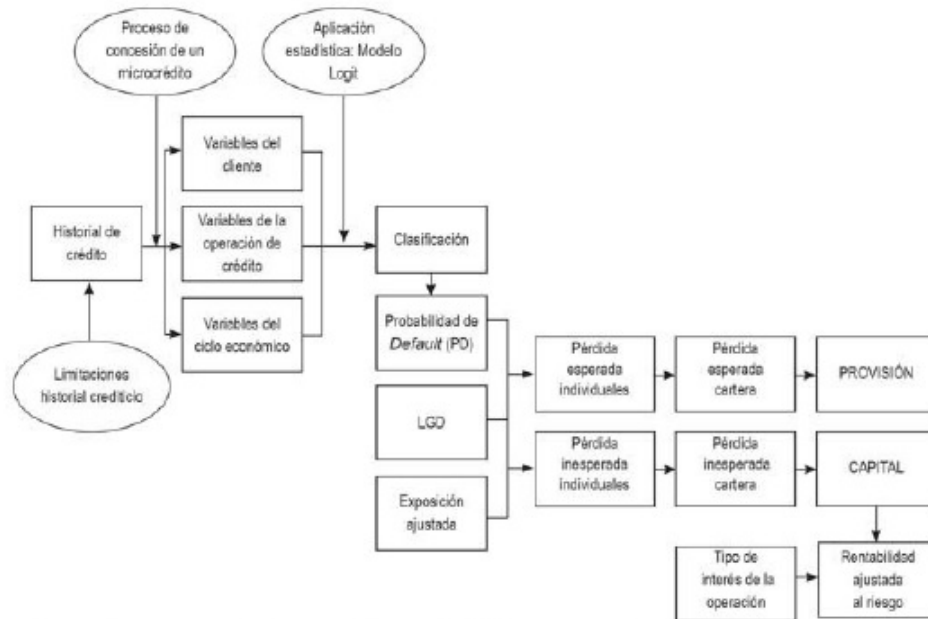


Figura 2. Esquema general de la metodología para el análisis del riesgo de microcrédito

2. Redes neuronales: El objetivo de este análisis es encontrar algunas relaciones para grupos determinados de la población, de acuerdo con sus características particulares. Se cruzan variables de cada cliente, del crédito y del comportamiento, contra la variable default (ocurre o no ocurre) que establece un procedimiento de clasificación y permite determinar las ponderaciones necesarias y establecer la probabilidad de ocurrencia. Las redes neuronales emulan la estructura y comportamiento del cerebro humano. Son un conjunto de algoritmos matemáticos que encuentran las relaciones no lineales entre conjuntos de datos; se emplean como herramientas para la predicción de tendencias.

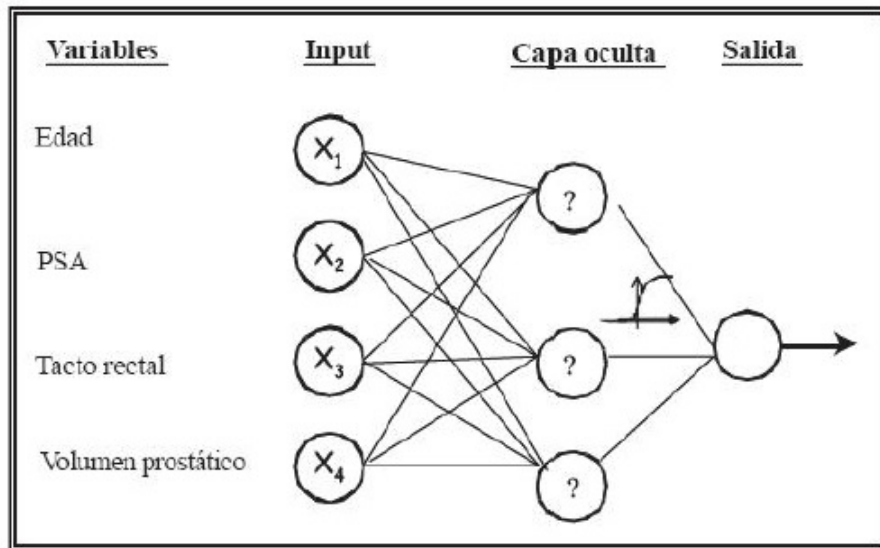
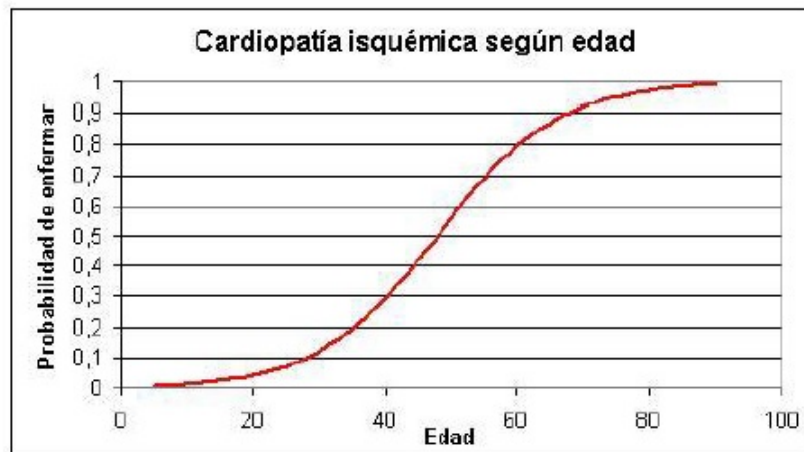


FIGURA 1: Estructura de la Red Neuronal Artificial construida para predecir cáncer de próstata, ajustando por diversas covariables. Set de derivación.

3. Regresión Logística: Es una herramienta estadística de análisis multivariado de uso predictivo o explicativo. La regresión estadística analiza datos cuya distribución es binomial.



Para cualquier forma que se elija para construir un modelo de riesgo, se pueden identificar dos tipos de construcción:

1. Experto.

Los modelos son construidos con información de otras instituciones, no así con la información de la institución de interés. Estos modelos son hechos cuando la institución no cuenta con información de su portafolio, como instituciones de reciente creación.

2. Estadístico.

Los modelos son construidos con información de la misma institución, y por lo tanto son los modelos que mejor funcionan.

Este último es el método que se empleará en el presente trabajo. La obtención de información para la elaboración del modelo de riesgo propuesto proviene de una SOFOM, cuyo objetivo es el otorgamiento de créditos al consumo y se propone una estrategia preventiva de cobranza para dicha SOFOM. La manipulación de bases de datos obtenida es llevada a cabo en el software SAS. Como el suceso a estudiar es dicotómico (el cliente pagará o no pagará), se empleó una regresión logística; la cual permite modelar cómo influye en la probabilidad de aparición de un suceso la presencia o no de diversos factores y su valor o nivel.

#### 4.1.1. Consideraciones del modelo

Las consideraciones del modelo son las siguientes:

1. Base principal: En primera instancia, se emplearon las exigibilidades de pago del mes de mayo del año 2011 (68,636 datos).
2. Colocación de bandera de cliente GOOD/BAD. Para la construcción de la bandera se partió de los siguientes supuestos:

- Cliente GOOD. Se empleó la base de datos con las exigibilidades del mes de mayo del año 2011 para la institución financiera, determinando que el cliente que con cero días de mora en su fecha de vencimiento y que dos días posteriores a dicha fecha siguiera presentando cero días de mora fuera un cliente GOOD. Es decir, todo cliente que dos días después de su fecha límite de pago estuviera al corriente, se consideró como bueno, pues cumplió con su pago.
- Cliente BAD. Se empleó la base de datos con las exigibilidades del mes de mayo del año 2011 para la institución financiera, determinando que el cliente que con cero días de mora en su fecha de vencimiento y que dos días posteriores a dicha fecha presentara más de cero días de mora fuera un cliente BAD. Pues se infiere que no realizaron pago alguno o bien un pago suficiente.

Después de dividir la base en estos dos grupos los datos se redujeron a 51,893. El nivel de morosidad de los clientes se obtuvo de la base del portafolio de la SOFOM.

#### **4.1.2. Variables empleadas**

Se emplearon 13 variables demográficas, 3 de buró de crédito y 10 variables de comportamiento del cliente. Tanto las variables demográficas como las de buró se obtuvieron de la base de origen, que almacena todas las solicitudes de los clientes. Es importante mencionar que se puede emplear un número considerable de variables, en este caso, se escogieron las de mayor importancia para fines de la realización del modelo de riesgo propuesto.

##### **Variables demográficas:**

- Fecha de nacimiento del cliente.
- Fecha de alta del crédito.

- Sexo del cliente.
- Estado donde reside el cliente.
- Nivel de estudios del cliente.
- Automóvil (Si ha dado en garantía un automóvil o no).
- Contrato laboral del cliente.
- Edad del cliente. Se refiere, como su nombre lo indica, a la edad del cliente y cómo influye dicha variable en su comportamiento crediticio.
- Número de dependientes del cliente. Se refiere al número de personas que dependen económicamente del cliente, lo cual proporcionará una noción de los gastos del cliente.
- Score (Calificación otorgada por la SOFOM al cliente). Esta variable es una calificación otorgada por la SOFOM al cliente. Las variables consideradas, así como sus respectivos puntajes, ya se encontraban establecidas previamente por la institución.
- ATP (Calificación otorgada por la SOFOM al cliente). El ATP, por sus siglas en inglés, significa: Ability to Pay. Esta variable nos indica la capacidad de pago del cliente, lo cual nos puede ayudar a predecir cómo será el comportamiento crediticio del cliente.
- Estado civil del cliente.
- Tipo de vivienda del cliente. Esta variable se encuentra dividida como sigue:
  1. Propia
  2. Familiares
  3. Rentada
  4. Pagándola



Resultaría lógico que un cliente con vivienda propia tuviera una mejor posición económica y, por consecuencia, tendría un mejor comportamiento crediticio.

- Sucursal. La SOFOM cuenta con 60 sucursales, en su mayoría en el Distrito Federal y Estado de México, y las restantes en el interior de la república.
- Código postal del cliente. Éste nos indica la zona donde reside.
- Capital. Es la línea de crédito que le fue otorgada al cliente por parte de la SOFOM
- Plazo. Es el tiempo que fue determinado para realizar el pago.
- Tipo de crédito. La SOFOM divide esta variable en dos grupos: consumo y microcrédito. Los créditos para el consumo como su nombre lo indica son para el consumo personal del cliente, por otro lado el microcrédito es un crédito para el negocio del cliente.
- Tipo de producto. SOFOM divide esta variable en dos grupos: preferencial y regular.
- Ubicación del microcrédito (de ser el caso). Esta variable aplica para los clientes con microcréditos, y está referida a la ubicación física del negocio del cliente.
- Tipo de microcrédito (de ser el caso). Esta variable aplica para los clientes con microcréditos, y está referida al tipo de negocio del cliente; es decir, si es propio o no.
- Lugar del microcrédito. Esta variable aplica para los clientes con microcréditos, y nos indica si el negocio del cliente es fijo o móvil.
- Horario laboral del cliente. Como su nombre lo indica, es el horario de la jornada laboral del cliente.

- Tipo de solicitud del cliente. Esta variable nos indica el tipo de solicitud del cliente establecida por la SOFOM, y se encuentra dividida como sigue:
  1. Former
  2. Nuevo
  3. Reestructura
  4. Renovación
- Periodicidad del crédito. Las periodicidades de pago que maneja la SOFOM son mensuales, quincenales y semanales.
- Ingreso total del cliente. Esta variable nos indica los ingresos brutos mensuales del cliente.
- Pago mensual por el crédito. Esta variable nos indica el monto que tiene que cubrir el cliente mensualmente para amortizar su deuda con la SOFOM.
- Gastos totales del cliente. Esta variable nos indica los gastos totales mensuales del cliente.
- Fuentes de ingreso variables.
- Cómo se enteró el cliente de la SOFOM. Esta variable la estableció la SOFOM para conocer la forma en la que el cliente se había enterado de la existencia de la de la institución, ya sea por algún tipo de publicidad, por otro cliente, o por alguna renovación o reestructura.
- Número de teléfonos que presenta el cliente. Es el número de teléfonos que proporcionó el cliente en el registro de su crédito para poder localizarlo.
- Comprobante de ingresos del cliente. Esta variable nos indica el tipo de solicitud del cliente establecida por la SOFOM, y se encuentra dividida como sigue:
  1. Cuenta de haberes

2. Notas de compra
  3. Estados de cuenta
  4. Facturas
  5. Nómina
  6. Recibos de pensión
- Frecuencia de percepción de ingresos del cliente. La frecuencia de percepción de ingresos es empleada por la SOFOM para determinar la periodicidad de los pagos del cliente. Esta frecuencia puede ser diaria, mensual, quincenal o variable, en ciertos casos.
  - Número de semanas en libros del cliente (WOB). Esta variable nos indica el número de semanas que el cliente ha estado en la cartera crediticia de la SOFOM. WOB por sus siglas en inglés significa Weeks on books.
  - Sistema de nacimiento. La SOFOM maneja dos sistemas llamados sistema 1 y sistema 2, este último es relativamente nuevo y esta variable nos indica bajo qué sistema fue otorgado el crédito.

**Variables de buró de crédito:**

- Score otorgado por el buró de crédito para el cliente. Esta variable nos va a dar una noción del futuro comportamiento crediticio de un nuevo cliente.
- Bandera: clientes con historial y sin historial crediticio. La SOFOM divide a sus clientes en dos grupos: clientes con historial y sin historial crediticio.
- Gastos en el buró de crédito del cliente. Como su nombre lo indica, y con base en la información proporcionada por el buró de crédito, esta variable proporciona información respecto a los gastos del cliente.

**Variables de comportamiento:**

- Máximo de días de morosidad (Obtenido del portafolio de la SOFOM). Esta variable se obtiene del portafolio de crédito de la SOFOM, e indica el número máximo de días de mora en que ha incurrido el cliente antes o igual a su próxima exigibilidad.
- Promedio de días que el cliente se tarda en pagar la exigibilidad. (Obtenido del portafolio de la SOFOM). Esta variable se obtiene del portafolio de crédito de la SOFOM e indica en promedio (porcentaje) el número de días que el cliente tarda en pagar sus exigibilidades.
- Porcentaje de capital ya pagado (Obtenido del portafolio de la SOFOM). Esta variable se obtiene del portafolio de crédito de la SOFOM e indica en porcentaje la cantidad que el cliente ya ha abonado a capital.
- Número de impagos en los últimos  $n$  meses ( $N = 5$ ) (Obtenido de la base de pagos de la SOFOM). Esta variable se obtiene de la base de pagos de la SOFOM e indica el número de impagos del cliente en los últimos 5 meses.
- Número de adelantos en los últimos  $n$  meses ( $N = 5$ ) (Obtenido de la base de pagos de la SOFOM). Esta variable se obtiene de la base de pagos de la SOFOM e indica el número de adelantos del cliente en los últimos 5 meses.
- Máximo número de recibos vencidos (Obtenido del portafolio de la SOFOM). Esta variable se obtiene del portafolio de crédito de la SOFOM e indica el máximo número de recibos vencidos del cliente, antes o igual a su próxima exigibilidad.
- Número  $k$  recibos vencidos con  $k = 1, \dots, 5$  (Obtenido del portafolio de la SOFOM). Esta variable se obtiene del portafolio de crédito de la SOFOM e indica el número de recibos vencidos de clientes, antes o igual a su próxima exigibilidad.

- Cuotas ya pagadas (Obtenido de la base de pagos de la SOFOM). Esta variable se obtiene de la base de pagos de la SOFOM e indica el número de cuotas ya pagadas del cliente, antes o igual a su próxima exigibilidad.
- Número de condonaciones (Obtenido de la base de condonaciones de la SOFOM). La variable se obtiene de la base de condonaciones de la SOFOM, e indica el número de condonaciones que ha tenido el cliente, antes o igual a su próxima exigibilidad.
- Número de renovaciones (Obtenido de la base de renovaciones de la SOFOM). La variable se obtiene de la base de renovaciones de la SOFOM, e indica el número de renovaciones que ha tenido el cliente, antes o igual a su próxima exigibilidad.

### **4.1.3. Análisis bivariado**

El análisis bivariado es el tratamiento estadístico de dos variables de manera simultánea. Este análisis permite estudiar la relación entre dos variables, ya que mide el cambio entre ambas variables al mismo tiempo, diseñando tablas con tabulaciones cruzadas, es decir, las categorías de una variable se cruzan con las categorías de una segunda variable, implicando el análisis comparativo de dos variables, una de las cuales modifica a la otra. Este análisis permitirá saber si existe o no una asociación, la fuerza de esta asociación, o si hay una diferencia entre las variables. Existen varias formas de realizarlo, y se muestran a continuación:

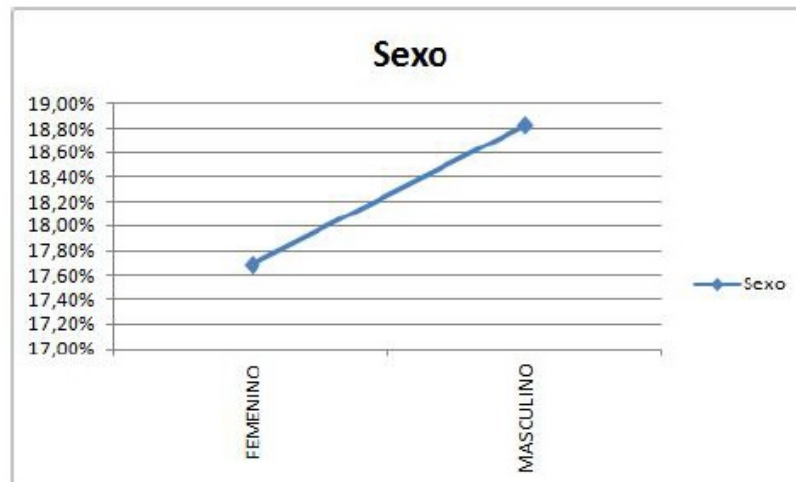
1. Contar frecuencias dobles: Consiste en contar dos características a la vez en un individuo. Estas frecuencias dobles se usan en tablas anidadas o de doble entrada, para el caso de variables nominales u ordinales. Las variables nominales únicamente permiten la clasificación cualitativa, es decir, se pueden medir sólo en términos de si los elementos individuales pertenecen a ciertas categorías distintas. Debido a la falta de orden o intervalos iguales, no podemos realizar

operaciones aritméticas u operaciones lógicas en los datos nominales. En cambio, las variables ordinales son variables ordinales discretas, pero sus estados se ordenan en una secuencia significativa. Las operaciones aritméticas resultan imposibles pero se pueden realizar operaciones lógicas con ellas.

2. Correlación simple: Procedimiento para variables cuantitativas. Consiste en computar el producto de las desviaciones respecto de la media para dos variables, evaluadas sobre un mismo individuo. El análisis bivariado da paso al análisis multivariado. Para efectos del modelo propuesto se analiza el comportamiento de cada variable respecto al nivel de morosidad que toma.

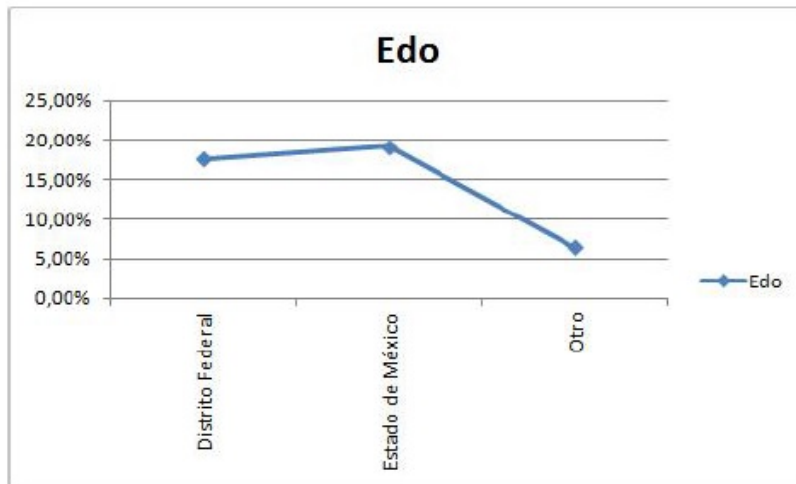
A continuación se muestran los resultados del análisis bivariado que se realizó. La columna ID fue asignada de acuerdo al nivel de mora de cada variable. Algunas variables fueron agrupadas en cuanto a sus valores cualitativos, como el nivel de estudios. La razón fue lograr su representatividad para la regresión logística. Este agrupamiento fue en base al nivel de mora.

ID SEXO	ALTAS	MOROSOS	%MOROSOS
1 FEMENINO	30792	5446	17.69%
2 MASCULINO	22096	4161	18.83%



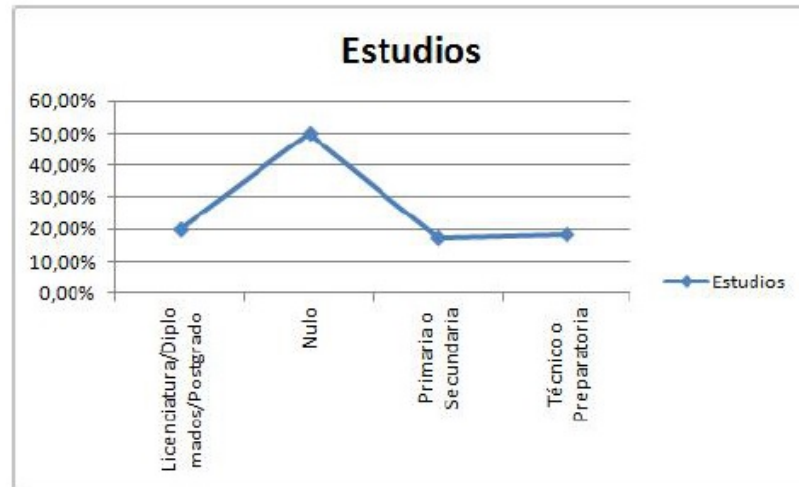
Se aprecia claramente que el comportamiento crediticio, tanto de las mujeres como de los hombres son similares, aunque las mujeres tuvieron un mejor comportamiento crediticio que los hombres en la ventana de tiempo analizada.

ID	EDO	ALTAS	MOROSOS	%MOROSOS
2	DISTRITO FEDERAL	16738	2944	17.59%
3	ESTADO DE MEXICO	33992	6524	19.19%
1	OTRO	2158	139	6.44%



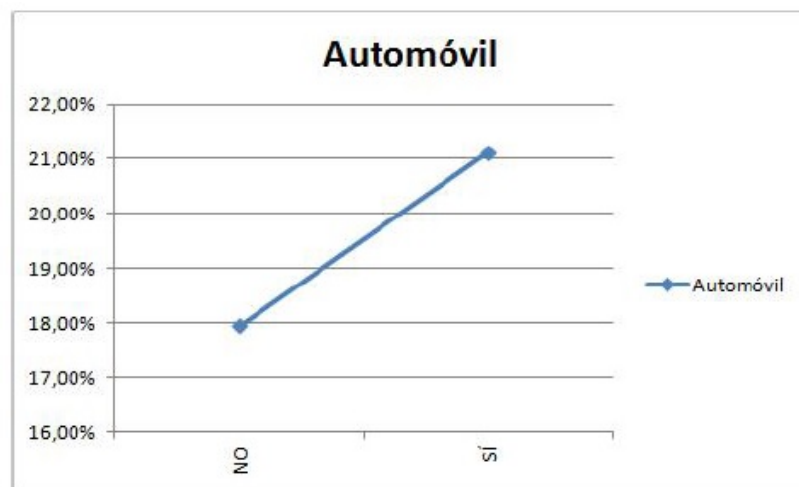
Se puede observar que clientes domiciliados, tanto en el Estado de México como en el Distrito Federal, tuvieron un comportamiento crediticio similar. Aquellos clientes domiciliados en otro estado fueron mucho menos morosos.

ID	ESTUDIOS	ALTAS	MOROSOS	%MOROSOS
3	Licenciatura/Diplomados/Postgrado	4592	920	19.97%
4	Nulo	4	2	50.00%
1	Primaria o Secundaria	23767	4164	17.52%
2	Técnico o Preparatoria	24525	4521	18.43%



Es interesante observar que los clientes con un mayor nivel de estudios (Licenciatura/Diplomados/Postgrado) presentaron un mayor nivel de mora que aquellos clientes con menor preparación (Primaria o Secundaria y Técnico o Preparatoria)

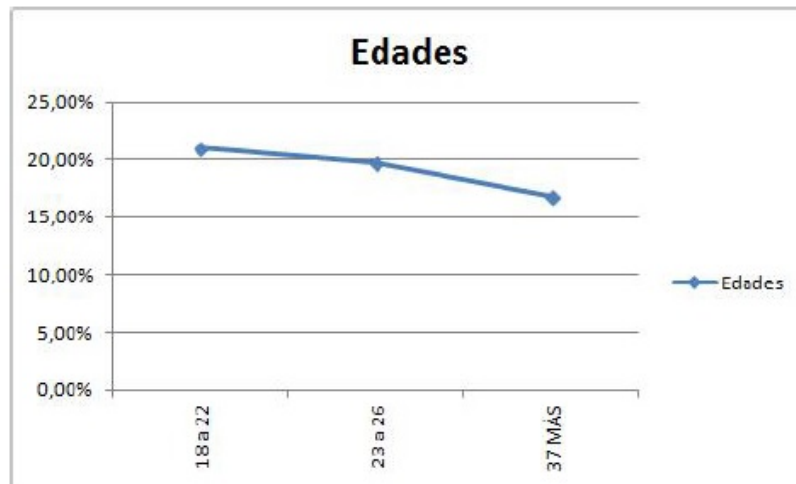
ID	AUTOMÓVIL	ALTAS	MOROSOS	%MOROSOS
1	NO	49282	8845	17,95%
2	SI	3606	762	21,13%



Aquellos clientes que presentaron como garantía su automóvil presentaron mayor nivel de mora que los que no presentaron dicha garantía.

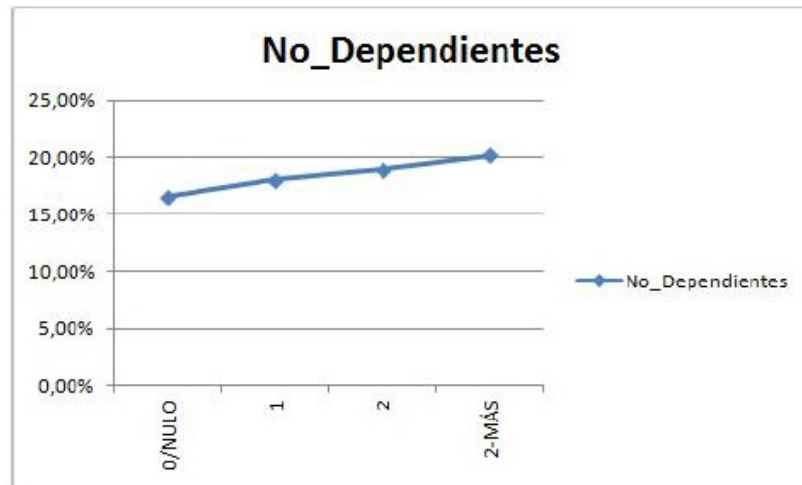


ID	EDADES	ALTAS	MOROSOS	%MOROSOS
3	1 a 22	3577	750	20.97%
2	23 a 36	19679	3876	19.70%
1	37 MAS	29632	4981	16.81%



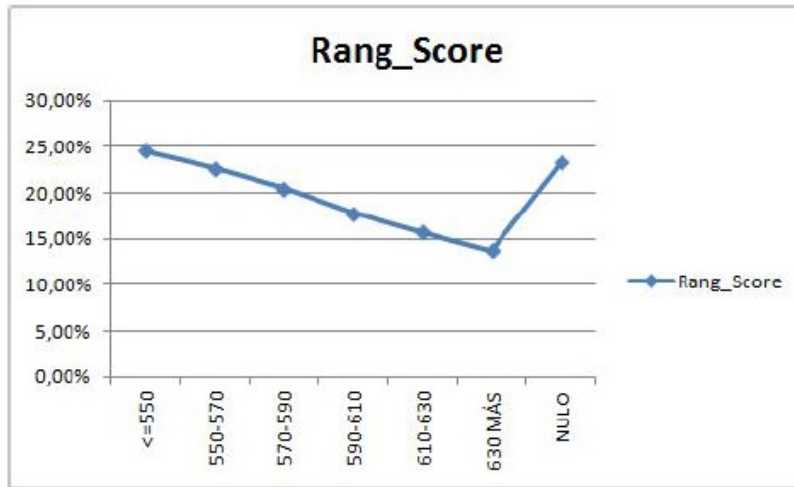
Aquellos clientes menores a los 36 años de edad fueron marcadamente más morosos que los clientes de mayor edad, lo cual desde mi punto de vista resultaría lógico: A mayor edad generalmente existe una mayor estabilidad económica, además de un mayor sentido de la responsabilidad.

ID	NO DEPENDIENTES	ALTAS	MOROSOS	%MOROSOS
1	0/NULLO	7077	1175	16.60%
2	1	31728	5723	18.04%
3	2	10873	2061	18.98%
4	2-MAS	3210	648	20.19%



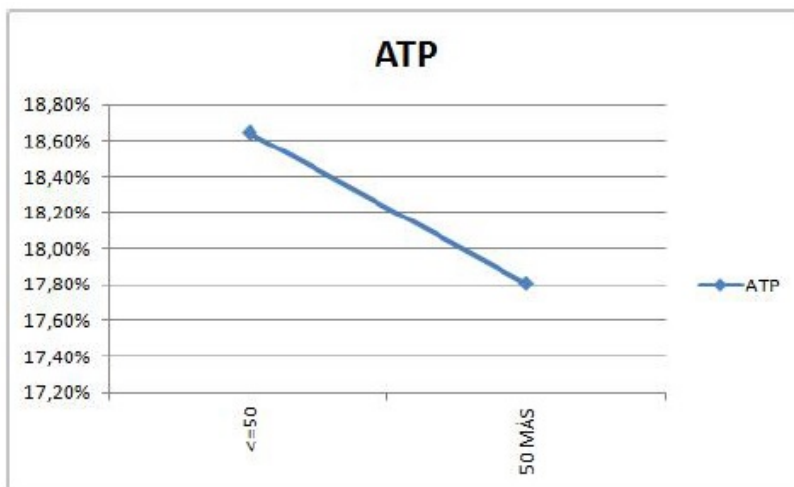
Se puede observar claramente que a mayor número de dependientes, mayor mora. Desde mi punto de vista este resultado tiene mucho sentido, un hijo representa una gran responsabilidad económica, lo cual en ocasiones no le permite al padre cubrir sus deudas económicas.

ID	RANG	SCORE	ALTAS	MOROSOS	%MOROSOS
7	<=550		1546	380	24.58%
5	550 - 570		3359	762	22.69%
4	570 - 590		12297	2526	20.54%
3	590 - 610		18503	3289	17.78%
2	610 - 630		13808	2173	15.74%
1	630 MAS		3332	467	13.74%
6	NULL		43	10	23.26%



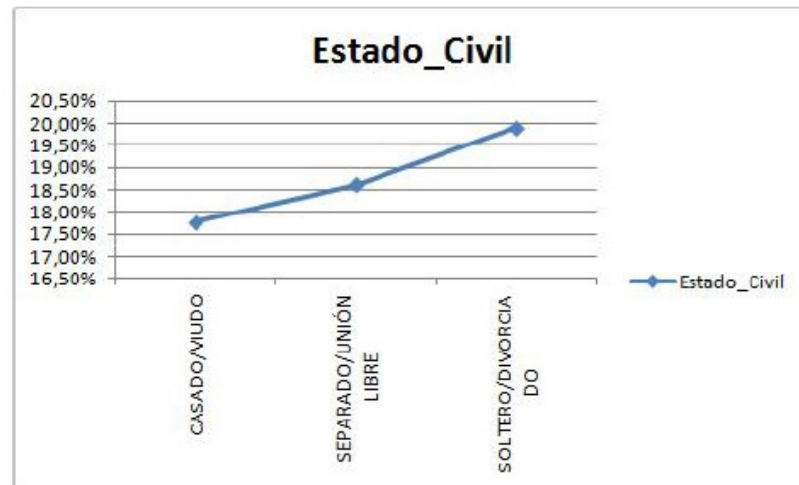
Se consideró importante analizar el nivel de mora de los clientes respecto del puntaje adquirido al momento de adquirir el crédito con la institución, ya que dicho Score fue construido con base en varias variables demográficas y de comportamiento del cliente. El resultado mostró la tendencia esperada, a mayor score menor nivel de mora.

ID	ATP	ALTAS	MOROSOS	%MOROSOS
2	<=50	22178	4138	18,66%
1	50 MAS	30710	5469	17,81%



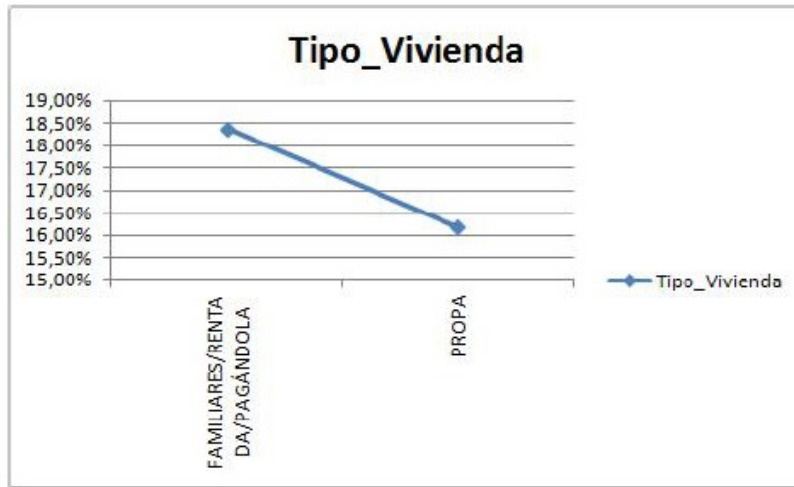
Se consideró importante analizar el nivel de mora de los clientes respecto del puntaje. El resultado mostró la tendencia esperada, a mayor puntaje en el ATP (Ability to Pay ó Capacidad de pago) menor nivel de mora.

ID	ESTADO CIVIL	ALTAS	MOROSOS	%MOROSOS
1	CASADO/VIUDO	40002	7110	17,77%
2	SEPARADO/UNIÓN LIBRE	5376	1001	18,62%
3	SOLTERO/DIVORCIADO	7510	1495	19,92%



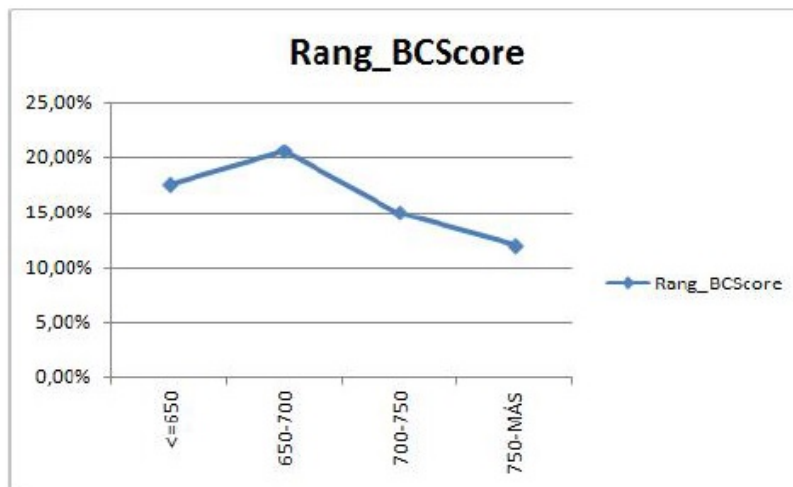
Resulta interesante observar el comportamiento crediticio del cliente en relación con su estado civil. Los clientes casados o viudos son los clientes que mejor pagaron y los clientes solteros o divorciados son los que peor pagaron, lo cual, desde mi punto de vista, tendría dos vertientes: por un lado, el matrimonio implica una mayor responsabilidad económica, por lo tanto, se esperaría que se incrementaran las probabilidades de impago del cliente y, por otro lado, el matrimonio implica una responsabilidad moral y alguien comprometido en cuanto a su vida personal sería comprometido en todos los aspectos de su vida, de manera fundamental el económico.

ID	TIPOVIVIENDA	ALTAS	MOROSOS	%MOROSOS
2	FAMILIARES/RENTADA/PAGANDOLA	47448	8727	18,39%
1	PROPIA	5440	880	16,18%



Aquellos clientes con vivienda propia fueron menos morosos que los clientes que no poseen una vivienda propia, aunque la diferencia resulta mínima.

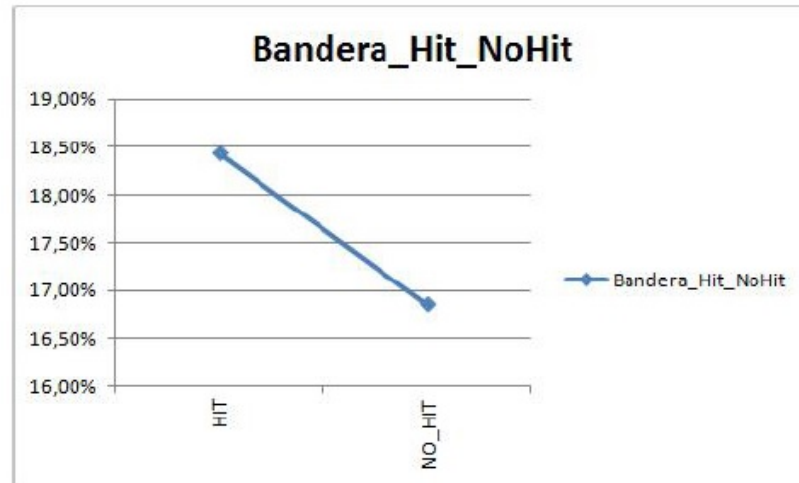
ID	RANG BCSCORE	ALTAS	MOROSOS	%MOROSOS
3	<=650	29350	5623	17,59%
4	650a700	14773	3040	20,58%
2	700a750	12475	1868	14,97%
1	750 MAS	748	90	12,03%



El resultado del análisis del Behavior Score o Score de comportamiento mostró, a grandes rasgos, la tendencia esperada: a mayor score (mejor comportamiento

crediticio en el buró de crédito) menor nivel de mora.

ID	BANDERA	HIT	NOHIT	ALTAS	MOROSOS	%MOROSOS
2	HIT			43530	8029	18.44%
1	NOHIT			9358	1578	16.86%



Aquellos clientes con historial crediticio (HIT) mostraron un peor comportamiento crediticio con respecto a los clientes sin historial crediticio (NO HIT) lo cual, desde mi punto de vista no resultaría lógico, pero el hecho es explicable si se observa el volumen de la población de clientes HIT (marcadamente superior) y la población no NO HIT.

#### 4.1.4. Insumo para regresión logística

Se utilizó la técnica de regresión logística para construir el modelo de riesgo, y se tuvo que construir un insumo con variables cuantitativas y por rangos. El ID que se otorgó en el análisis bivariado fue el valor que tomó cada variable. El insumo de la regresión se construyó en función al análisis bivariado, dejando todas las variables cuantitativas. El ID que se otorgó a cada variable está en función de la morosidad encontrada por parte del cliente, y sigue la misma coherencia para cada variable analizada.

## Regresión logística

La regresión logística es una herramienta estadística de análisis multivariado, puede tener un uso explicativo o predictivo. El objetivo del análisis es predecir la ocurrencia de cierto evento, determinar qué variables pesan más para aumentar o disminuir la probabilidad de que suceda el evento en cuestión. La determinación de probabilidad de ocurrencia del evento a cierto individuo, así como la determinación del peso de cada una de las variables dependientes para esta probabilidad, se basa en las características que presentan los individuos.

La regresión logística analiza datos distribuidos binomialmente de la forma:

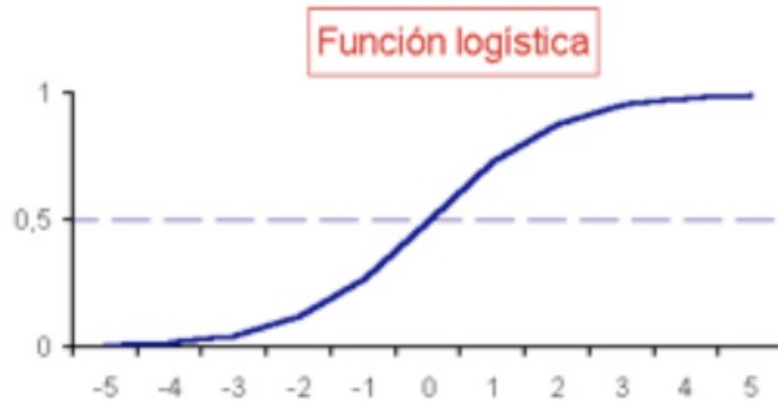
$$Y_i \sim B(p_i, n_i)$$

para  $i = 1, \dots, m$ ; donde los números de ensayos Bernoulli no son conocidos y las probabilidades de éxito  $p_i$  son desconocidas.

El modelo se obtiene con lo que cada ensayo (valor de  $i$ ) y el conjunto de variables explicativas/independientes puedan informar acerca de la probabilidad final. La regresión logística es un modelo de regresión donde se desea conocer la relación entre una variable dependiente cualitativa dicotómica y la relación entre una o más variables explicativas independientes o covariables, ya sean cualitativas o cuantitativas. El modelo de regresión logística parte de un modelo exponencial de la siguiente forma:

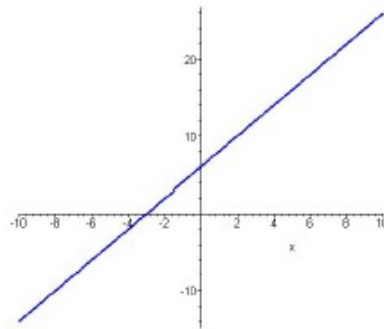
$$P(y = 1|x) = \frac{e^{b_0 + \sum_{i=1}^m b_i x_i}}{1 + e^{b_0 + \sum_{i=1}^m b_i x_i}}$$

cuya representación gráfica es:



Sabemos que mediante una transformación logarítmica (logit), logramos obtener una función lineal, como se muestra a continuación:

$$\log \left( \frac{P(y = 1|x)}{1 - P(y = 1|x)} \right) = b_0 + \sum_{i=1}^m b_i x_i$$



Donde  $y$  es la variable dependiente, en este caso, la bandera de cliente bueno o malo, y  $x$  nuestras variables independientes (todas las variables incluidas en el modelo (demográficas, de buró y de comportamiento). También cabe mencionar que el proceso para encontrar las  $b_i$ , es mediante métodos numéricos, en función de la significancia de las variables que explican la predicción.



#### **4.1.5. Metodología empleada**

Cómo se mencionó anteriormente, se utilizó una de las herramientas estadísticas con mayor capacidad para el análisis de datos, la regresión logística, y para ello se empleó el software SAS.

1. Se empleó la tabla insumo, construida con base en valores cuantitativos.
2. La variable dependiente fue la variable que se asignó como cliente bueno/malo; es decir, cliente con buen comportamiento crediticio o cliente moroso.
3. Se eligió pronosticar a las clientes morosos, es decir los clientes con valor 1 (ya que así se le asignó previamente: cliente bueno= 0, cliente malo= 1).
4. Se sometieron las variables cualitativas a la prueba de significancia.
5. Se eligió que el modelo tuviera intercepto u ordenada al origen (constante, modelo con puntuación inicial).
6. Se seleccionaron las variables significativas con la opción forward selection, que consiste en incluir variables de una en una.
7. El nivel de significancia fue del 5%.
8. Se obtuvo la calibración del modelo.
9. Se calculó el score en función de la regresión logística.
10. Se realizó un análisis de ranqueo del modelo total y por periodicidad.

#### **Resultados, análisis de la regresión logística y escalamiento del modelo**

De las variables incluidas en el modelo, no todas resultaron significativas. Las siguientes variables fueron las variables significativas:

- Sexo del cliente.

- 
- Nivel de estudios del cliente.
  - Edad del cliente.
  - Número de dependientes del cliente.
  - Score otorgado por el buró de crédito para el cliente.
  - Plazo (Para realizar el pago).
  - Tipo de crédito.
  - Lugar del microcrédito.
  - Tipo de solicitud del cliente.
  - Periodicidad del crédito.
  - Ingreso total del cliente.
  - Pago mensual por el crédito.
  - Número de semanas en libros del cliente.
  - Sistema de nacimiento (la SOFOM maneja dos sistemas).
  - Máximo de días de morosidad.
  - Promedio de días que el cliente se tarda en pagar la exigibilidad.
  - Porcentaje de capital ya pagado.
  - Cuotas ya pagadas.
  - Número de condonaciones.
  - Número de impagos en los últimos  $n$  meses.
  - Número de veces en 1 recibo vencido.

- Número de veces en 2 recibos vencidos.

Las variables restantes no fueron consideradas significativas. En particular se infiere, por su comportamiento, que dichas variables muestran una marcada tendencia de comportamiento respecto a su morosidad.

<b>Análisis del estimador de máxima verosimilitud</b>					
<b>VARIABLE</b>	<b>Chi-cuadrado de Wald</b>	<b>Parámetro Pr&gt;ChiSq Intercept</b>	<b>DF</b>	<b>Estimador</b>	<b>Error Estándar</b>
INTERCEPTO	149,282	<.0001	1	-3,0833	0,2524
SEXO	3,5585	0,0592	1	0,0532	0,0282
ESTUDIOS	2,1303	0,1444	1	-0,0411	0,0282
EDADES	12,9354	0,0003	1	0,0802	0,0223
NO_DEPENDIENTES	5,4823	0,0192	1	0,0438	0,0187
RANG_BCSCORE	20,9648	<.0001	1	0,083	0,0181
T_PLAZO	10,6936	0,0011	1	0,1625	0,0497
TIPO_CRÉDITO	12,6066	0,0004	1	-0,2405	0,0677
LUGAR_MICROCRÉDITO	3,2848	0,0699	1	0,067	0,037
TIPO_SOLICITUD	12,046	0,0005	1	-0,0869	0,025
TIPO_PERIODICIDAD	60,9894	<.0001	1	0,2168	0,0287
INGRESTO_TOTAL	4,5946	0,0321	1	0,0748	0,0349
PAGO_MENSUAL	2,5183	0,1125	1	0,0372	0,0234
WOB	17,8863	<.0001	1	-0,0597	0,0141
NAC-SIST	5,4978	0,019	1	-0,177	0,0755
MAXIMO_DM	60,3892	<.0001	1	0,261	0,0336
PROMEDIO_DM	75,7516	<.0001	1	0,4183	0,0481
PCT_CAPPAG	141,1125	<.0001	1	-0,1268	0,0107
CTASPAGADAS	534,9414	<.0001	1	-0,2195	0,00949
CONDONACIONES	3,5299	0,0603	1	0,2311	0,123
NUM_IMPAGOS	2464,6925	<.0001	1	0,556	0,012
V_R1	13,563	0,0002	1	-0,0877	0,0238
V_R4	40,0358	<.0001	1	-0,2823	0,0446
RENOVACIONES	59,2926	<.0001	1	-0,284	0,0369

## 4.2. Pruebas de Bondad de Ajuste

Las llamadas pruebas de bondad de ajuste son procedimientos de pruebas de hipótesis que comparan los resultados de una muestra con los que se espera obtener cuando la hipótesis nula es verdadera. La conclusión de esta prueba de hipótesis se basa en lo cercano que están los resultados de la muestra, con respecto a los resultados esperados.

A continuación se expone la definición de las medidas globales de la bondad de ajuste:

1. Devianza del modelo: Es una medida del grado de diferencia entre las frecuencias observadas y predichas por el modelo de la variable dependiente, de forma que, a mayor devianza, peor es el modelo. Su cálculo es  $-2$  veces el logaritmo neperiano de la verosimilitud del modelo. La devianza orienta durante la etapa de selección del modelo final.
2. Razón de verosimilitud: El estadístico que se usa es  $G$ , que es  $-2$  veces el logaritmo neperiano del cociente entre la verosimilitud del modelo con el conjunto de  $p$  covariables introducidas en el mismo y la del modelo sólo con la constante. Este estadístico sigue una distribución  $\chi^2$  con  $p$  grados de libertad. Si este estadístico alcanza significación estadística indica un buen ajuste, quiere decir que uno o más de los coeficientes de las covariables introducidas en el modelo es distinto de 0.
3.  $\chi^2$  residual de Pearson y devianza de residuos: Se trata de dos pruebas basadas en los residuos de un modelo que siguen una distribución  $\chi^2$  con  $m - (p + 1)$  grados de libertad. La ausencia de significación indica que el ajuste del modelo es bueno.
4. Prueba de Hosmer-Lemeshow: Cuando los patrones de covariables siguen una distribución  $n$ -asintótica, la  $\chi^2$  residual de Pearson y la devianza de residuos no se ajustan bien a una distribución  $\chi^2$ . Por ello, en general será más apropiada

la prueba de Hosmer-Lemeshow, que agrupa los  $n$  sujetos en  $m$  patrones según criterios estadísticos. Los sujetos se agrupan según los 9 deciles de las probabilidades esperadas; a partir de aquí puede construirse una tabla de contingencia de  $10 \times 2$  de la que puede construirse un estadístico que seguirá una distribución  $\chi^2$  con 8 grados de libertad. La ausencia de significación indica un buen ajuste del modelo.

5. Área bajo la curva ROC: Una curva ROC enfrenta en un sistema de ejes la sensibilidad (en el eje  $y$ ), al complementario de la especificidad (en el eje  $x$ ). El procedimiento consiste en determinar las correspondientes tablas de clasificación de puntos de corte de  $P(Y = 1|X)$  crecientes (0.1, 0.2, ..., 0.8, 0.9, 1), y determinar a partir de ellas las correspondientes sensibilidades y especificidades.

Para obtener el punto de corte óptimo para predecir  $Y$ , es decir, el punto de corte que ofreciera mayor sensibilidad y especificidad, recurriríamos al análisis gráfico de la curva, seleccionando como punto de corte aquél que correspondiera con el punto de inflexión de la curva ROC. Otra forma sería analizar una gráfica en la que se representaran para cada punto de corte (en el eje  $x$ ) su sensibilidad y especificidad (en el eje  $y$ ); el punto de corte óptimo coincidiría con aquél en el que se cruzaran las dos curvas. La comparación entre modelos respecto a la capacidad predictiva de los mismos puede hacerse comparando la forma de las curvas y el área bajo las mismas; las mejores curvas serán aquellas con área más próxima a la unidad. Como regla general, un área de 0.5 implica ausencia de discriminación; entre 0.7 y 0.79 es una discriminación aceptable; entre 0.8 y 0.89 es excelente; 0.9 ó superior es una discriminación excepcional.

El área bajo la curva se calcula de la siguiente forma:

$$\frac{UMX}{n_0 \times n_1}$$

donde  $UMW$  es el estadístico  $U$  de Mann-Whitney y  $n_0$  y  $n_1$  son el número de sujetos con  $y = 0$  y  $y = 1$ .

Cabe aclarar que el modelo de riesgo propuesto no contempla un procedimiento de prueba de bondad de ajuste, ya que las bases empleadas para la realización de la regresión logística son de uso exclusivo de la SOFOM.

### 4.3. Metodología empleada para construir el Prevent Score

Una vez que obtuvimos los coeficientes de la regresión logística, recordemos que estos valores nos darán una probabilidad entre cero y uno correspondiente a la probabilidad de que un cliente estando al corriente con sus pagos llegada su fecha límite de pago no realice el pago, por lo que usaremos dichos coeficientes para utilizar otra escala en función a un puntaje que denominaremos Prevent Score:

$$Score = \left( \text{puntaje deseado} + \frac{PDO}{Ln2} \right) * \left( B_0 + \sum_{i=1}^n B_i * X_i \right)$$

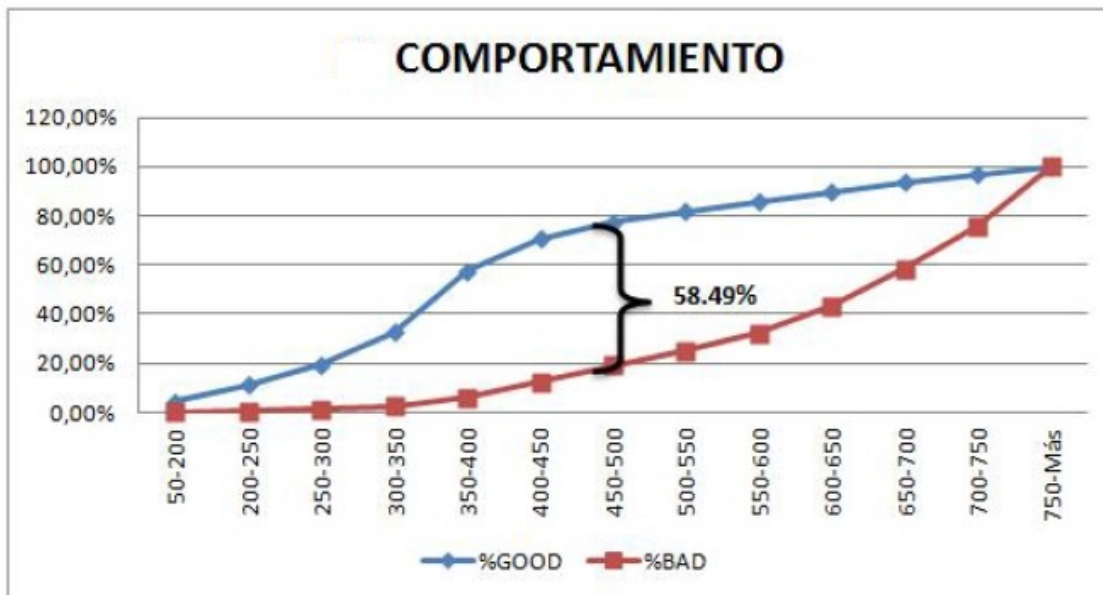
Donde:

El Puntaje deseado corresponde al puntaje de Score donde deseamos que este el promedio de Score, en este caso utilizamos 100 puntos.

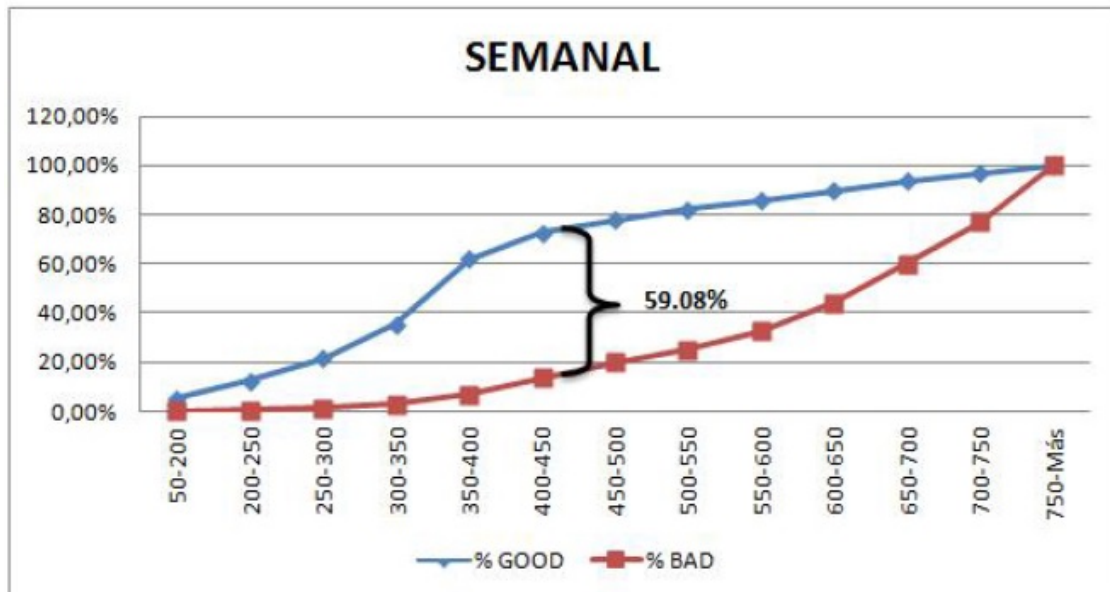
PDO mejor conocido como Points to Double Odd's que traducido quiere decir: Puntos de Score deseados que queremos que existan para que suplique la probabilidad de pago, en este caso usamos 750 Puntos.

Una vez explicado esto podemos medir el desempeño del modelo.

COMPORTAMIENTO									
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA
50-200	2161	2161	8	8	2169	0,37%	4,53%	0,08%	4,45%
200-250	3205	5366	22	30	3227	0,68%	11,24%	0,32%	10,92%
250-300	3970	9336	80	110	4050	1,98%	19,56%	1,16%	18,40%
300-350	6216	15552	144	254	6360	2,26%	32,58%	2,67%	29,91%
350-400	11982	27534	323	577	12305	2,62%	57,69%	6,07%	51,62%
400-450	6264	33798	594	1171	6858	8,66%	70,81%	12,32%	58,49%
450-500	3161	36959	649	1820	3810	17,03%	77,44%	19,15%	58,29%
500-550	2121	39080	553	2373	2674	20,58%	81,88%	24,97%	56,91%
550-600	1832	40912	687	3060	2519	27,27%	85,72%	32,19%	53,53%
600-650	1946	42858	1049	4109	2995	35,05%	89,79%	43,23%	46,56%
650-700	1872	44730	1450	5559	3322	43,65%	93,72%	58,49%	35,23%
700-750	1545	46275	1643	7202	3188	51,54%	96,95%	75,77%	21,18%
750-Más	1454	47729	2303	9505	3757	61,30%	100,00%	100,00%	0,00%
								KS	58,49%

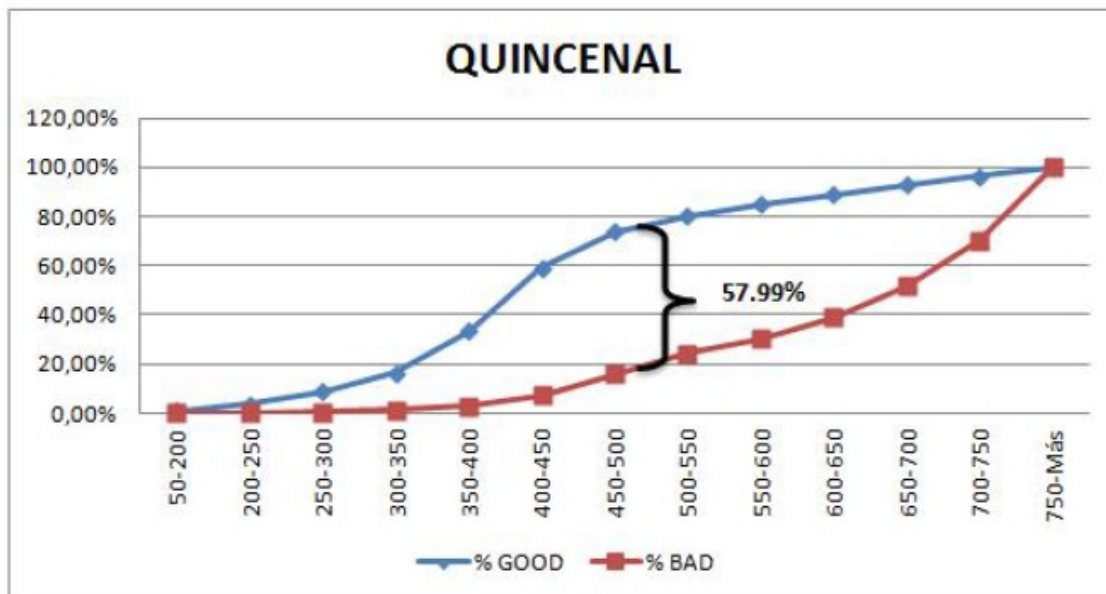


SEMANAL									
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA
50-200	2110	2110	8	8	2118	0,38%	5,14%	0,11%	5,03%
200-250	3015	5125	22	30	3037	0,72%	12,48%	0,40%	12,08%
250-300	3649	8774	68	98	377717	1,83%	21,37%	1,32%	20,05%
300-350	5730	14504	130	228	5860	2,22%	35,33%	3,08%	32,25%
350-400	10870	25374	290	518	11160	2,60%	61,80%	6,99%	54,81%
400-450	4513	29887	498	1016	5011	9,94%	72,79%	13,71%	59,08%
450-500	2138	32025	472	1488	2610	18,08%	78,00%	20,08%	57,92%
500-550	1677	33702	368	1856	2045	18,00%	82,09%	25,04%	57,05%
550-600	1512	35214	550	2406	2062	26,67%	85,77%	32,46%	53,31%
600-650	1687	36901	861	3267	2548	33,79%	89,88%	44,08%	45,80%
650-700	1624	38525	1187	4454	2811	42,23%	93,83%	60,09%	33,74%
700-750	1317	39842	1259	5713	2576	48,87%	97,04%	77,08%	19,96%
750-Más	1215	41057	1699	7412	2914	58,30%	100,00%	100,00%	0,00%
								<b>KS</b>	<b>59,08%</b>

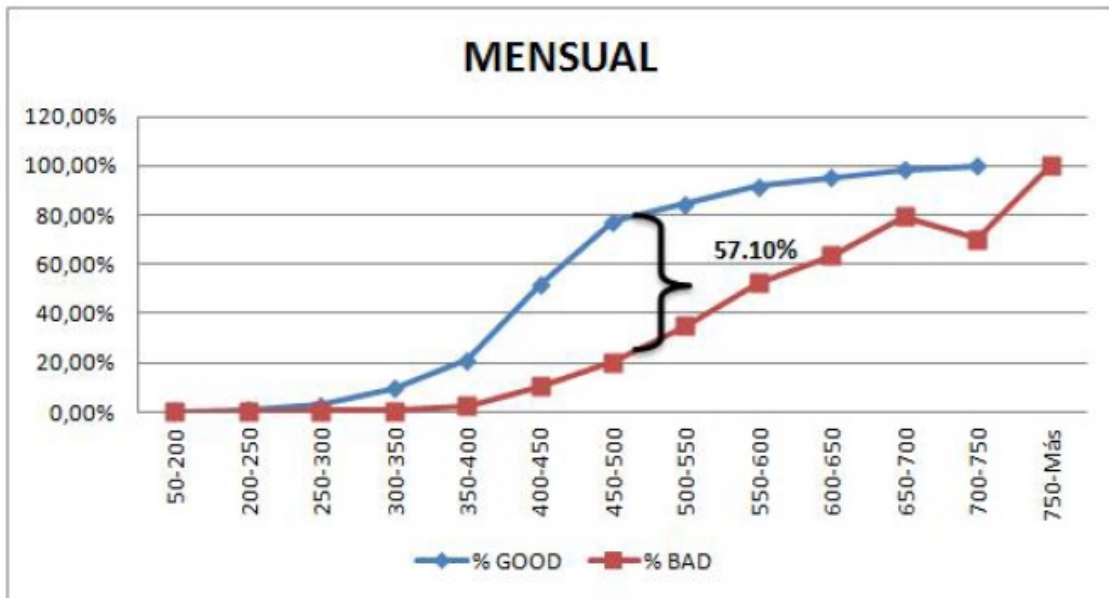




QUINCENAL									
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA
50-200	50	50	0	0	50	0,00%	0,81%	0,00%	0,81%
200-250	187	237	0	0	187	0,00%	3,82%	0,00%	3,82%
250-300	312	549	11	11	323	3,41%	8,86%	0,58%	8,28%
300-350	455	1004	14	25	469	2,99%	16,20%	1,31%	14,89%
350-400	1058	2062	30	55	1088	2,76%	33,27%	2,89%	30,38%
400-450	1608	3670	84	139	1692	4,96%	59,21%	7,29%	51,92%
450-500	903	4573	162	301	1065	15,21%	73,78%	15,79%	57,99%
500-550	411	4984	163	464	574	28,40%	80,41%	24,34%	56,07%
550-600	288	5272	111	575	399	27,82%	85,06%	30,17%	54,89%
600-650	242	5514	171	746	413	41,00%	88,96%	39,14%	49,82%
650-700	234	5748	239	985	473	50,53%	92,74%	51,68%	41,06%
700-750	220	5968	353	1338	573	61,61%	96,29%	70,20%	25,09%
750-Más	230	6198	568	1906	798	71,18%	100,00%	100,00%	0,00%
								<b>KS</b>	<b>57,99%</b>



MENSUAL										
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA	
50-200	1	1	0	0	1	0,00%	0,22%	0,00%	0,22%	
200-250	3	4	0	0	3	0,00%	0,86%	0,00%	0,86%	
250-300	9	13	1	1	10	10,00%	2,80%	0,66%	2,14%	
300-350	31	44	0	1	31	0,00%	9,46%	0,66%	8,80%	
350-400	54	98	3	4	57	5,26%	21,08%	2,65%	18,43%	
400-450	143	241	12	16	155	7,74%	51,83%	10,60%	41,23%	
450-500	120	361	15	31	135	11,11%	77,63%	20,53%	57,10%	
500-550	33	394	22	53	55	40,00%	84,73%	35,10%	49,63%	
550-600	32	426	26	79	58	44,83%	91,61%	52,32%	39,29%	
600-650	17	443	17	96	34	50,00%	95,27%	63,58%	31,69%	
650-700	14	457	24	120	38	63,16%	98,28%	79,47%	18,81%	
700-750	8	465	31	151	39	79,49%	100,00%	70,20%	29,80%	
750-Más			36		45			100,00%	-100,00%	
									KS	57,10%



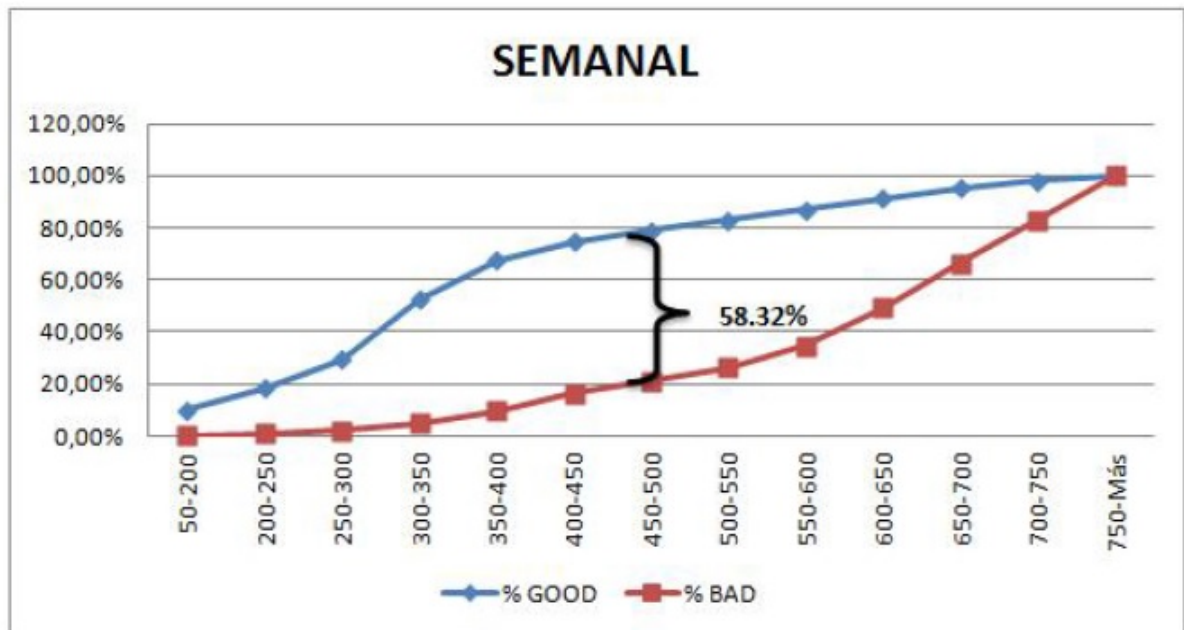
### 4.3.1. Validación del modelo

Con el fin de comprobar la efectividad del modelo se obtuvo el cálculo del Kolgomorov-Smirnov del mes de junio y se muestra continuación el total semanal, quincenal y mensual. Los resultados evidencian que el modelo efectivamente funciona.

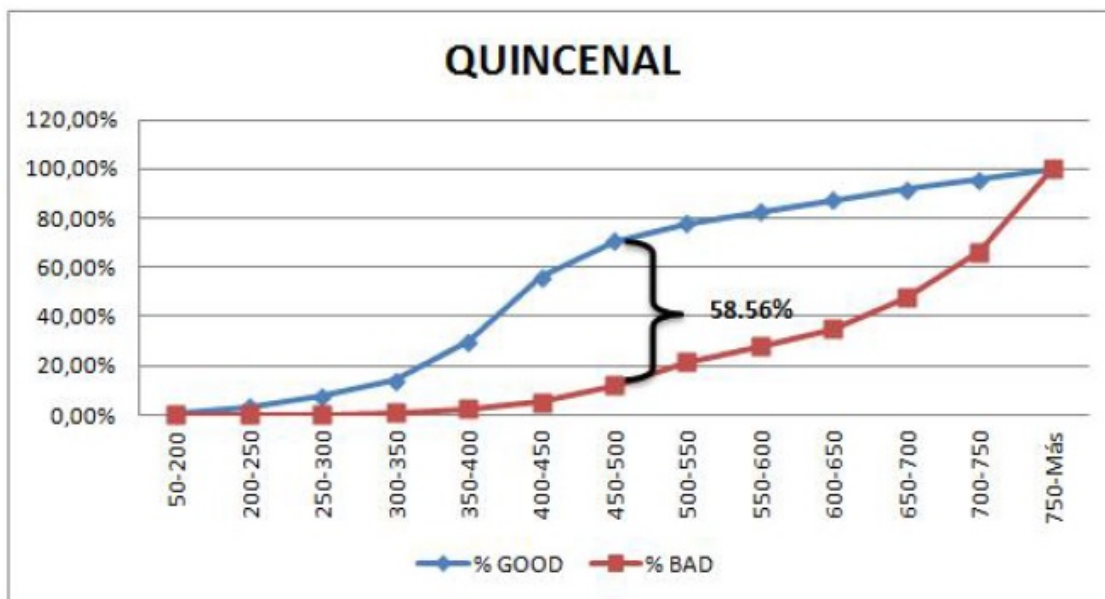
COMPORTAMIENTO										
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA	
50-200	4198	4198	19	19	4217	0,45%	8,80%	0,20%	8,60%	
200-250	3459	7657	44	63	3503	1,26%	16,04%	0,66%	15,38%	
250-300	4892	12549	104	167	4996	2,08%	26,29%	1,76%	24,53%	
300-350	9894	22443	201	368	10095	1,99%	47,02%	3,87%	43,15%	
350-400	7246	29689	401	769	7647	5,24%	62,20%	8,09%	54,11%	
400-450	4635	34324	555	1324	5190	10,69%	71,91%	13,93%	57,98%	
450-500	2904	37228	513	1837	3417	15,01%	78,00%	19,33%	58,67%	
500-550	2054	39282	573	2410	2627	21,81%	82,30%	25,36%	56,94%	
550-600	1898	41180	759	3169	2657	28,57%	86,28%	33,34%	52,94%	
600-650	2210	43390	1233	4402	3443	35,81%	90,91%	46,31%	44,60%	
650-700	1868	45258	1525	5927	3393	44,95%	94,82%	62,36%	32,46%	
700-750	1387	46645	1637	7564	3024	54,13%	97,73%	79,58%	18,15%	
750-Más	1084	47729	1941	9505	3025	64,17%	100,00%	100,00%	0,00%	
					57234					
								KS		58,67%



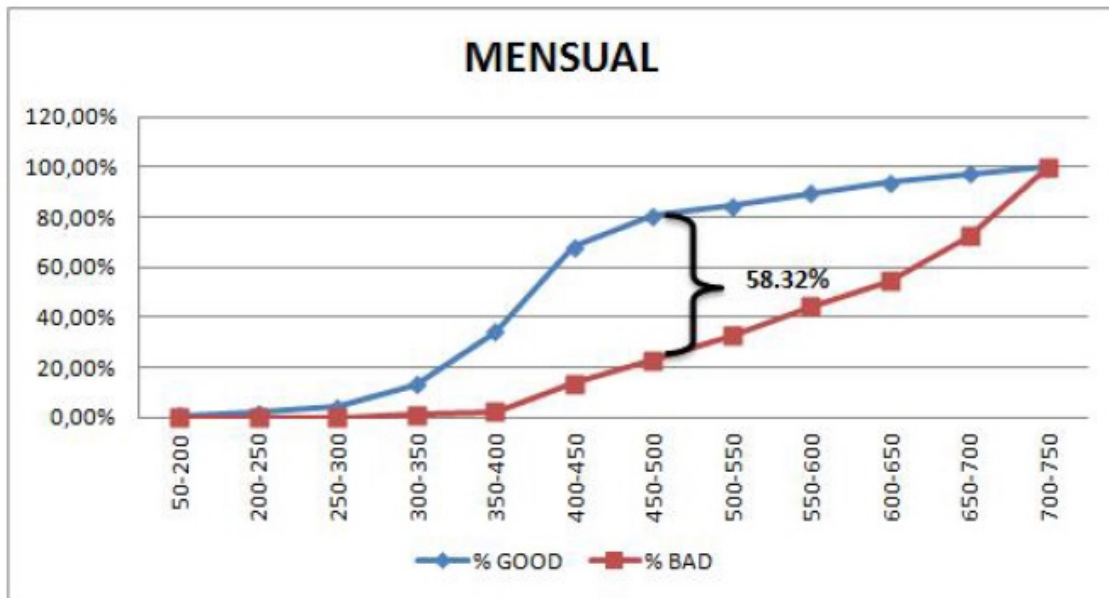
SEMANAL									
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA
50-200	4178	4178	19	19	4197	0,45%	10,18%	0,26%	9,92%
200-250	3284	7462	43	62	3327	1,29%	18,17%	0,84%	17,33%
250-300	4616	12078	99	161	4715	2,10%	29,42%	2,17%	27,25%
300-350	9490	21568	186	347	9676	1,92%	52,53%	4,86%	47,67%
350-400	6217	27785	372	719	6589	5,65%	67,67%	9,70%	57,97%
400-450	2901	30686	498	1217	3399	14,65%	74,74%	16,42%	58,32%
450-500	1836	32522	363	1580	2199	16,51%	79,21%	21,32%	57,89%
500-550	1557	34079	376	1956	1933	19,45%	83,00%	26,39%	56,61%
550-600	1573	35652	618	2574	2191	28,21%	86,84%	34,73%	52,11%
600-650	1894	37546	1075	3649	2969	36,21%	91,45%	49,23%	42,22%
650-700	1571	39117	1265	4914	2836	44,61%	95,27%	66,30%	28,97%
700-750	1132	40249	1252	6166	2384	52,52%	98,03%	83,19%	14,84%
750-Más	808	41057	1246	7412	2054	60,66%	100,00%	100,00%	0,00%
							KS		58,32%



QUINCENAL									
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA
50-200	20	20	0	0	20	0,00%	0,32%	0,00%	0,32%
200-250	172	192	1	1	173	0,58%	3,10%	0,05%	3,05%
250-300	271	463	5	6	276	1,81%	7,47%	0,31%	7,16%
300-350	393	856	15	21	408	3,68%	13,81%	1,10%	12,71%
350-400	986	1842	27	48	1013	2,67%	29,72%	2,52%	27,20%
400-450	1634	3476	55	103	1689	3,26%	56,08%	5,40%	50,68%
450-500	908	4384	129	232	1037	12,44%	70,73%	12,17%	58,56%
500-550	438	4822	179	411	617	29,01%	77,80%	21,56%	56,24%
550-600	306	5128	123	534	429	28,67%	82,74%	28,02%	54,72%
600-650	291	5419	136	670	427	31,85%	87,43%	35,15%	52,28%
650-700	278	5697	241	911	519	46,44%	91,92%	47,80%	44,12%
700-750	237	5934	352	1263	589	59,76%	95,74%	66,26%	29,48%
750-Más	264	6198	643	1906	907	70,89%	100,00%	100,00%	0,00%
								KS	58,56%



MENSUAL										
Score	GOOD	GOOD ACUM	BAD	BAD ACUM	CLIENTES	%BAD's	%GOOD	%BAD	DELTA	
200-250	3	3	0	0	3	0,00%	0,63%	0,00%	0,63%	
250-300	5	8	0	0	5	0,00%	1,69%	0,00%	1,69%	
300-350	11	19	0	0	11	0,00%	4,01%	0,00%	4,01%	
350-400	43	62	2	2	45	4,44%	13,06%	1,07%	11,99%	
400-450	100	162	2	4	102	1,96%	34,18%	2,14%	32,04%	
450-500	160	322	21	25	181	11,60%	67,93%	13,37%	54,56%	
500-550	59	381	18	43	77	23,38%	80,38%	22,99%	57,39%	
550-600	19	400	18	61	37	48,65%	84,39%	32,62%	51,77%	
600-650	25	425	22	83	47	46,81%	89,66%	44,39%	45,27%	
650-700	19	444	19	102	38	50,00%	93,67%	54,55%	39,12%	
700-750	18	462	33	135	51	64,71%	97,47%	72,19%	25,28%	
750-Más	12	474	52	187	64	81,25%	100,00%	100,00%	0,00%	
									KS	57,39%



A continuación, un comparativo del K-S de mayo contra del mes de junio, ambos de 2011.

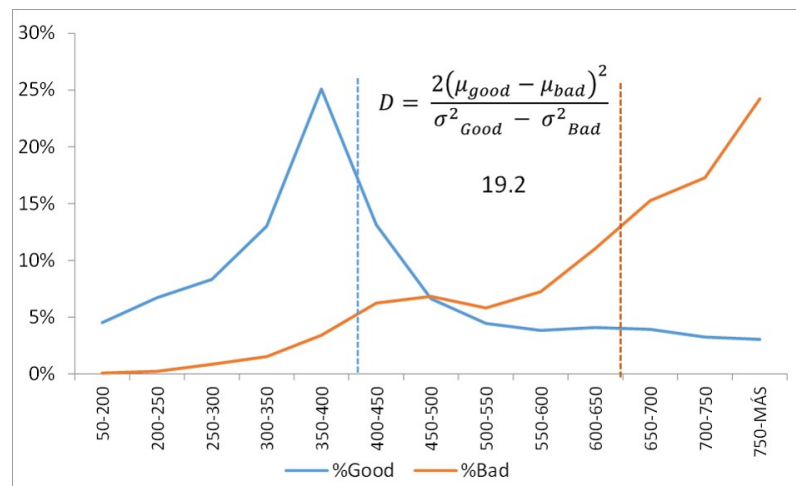
	KS JUNIO	KS MAYO	Diferencia (K-S)
<b>TOTAL</b>	58.67%	58.49%	0.18%
<b>SEMANAL</b>	58.32%	59.09%	0.77%
<b>QUINCENAL</b>	58.56%	57.99%	0.57%
<b>MENSUAL</b>	57.39%	57.10%	0.28%

La certeza del modelo es importante: el Kolgomorov-Smirnov del mes de mayo para toda la SOFOM, que fue el mes analizado, es del 58.49 %, es decir, de cada 10 clientes cerca de 6 pagarán, para los restantes por lo tanto, habría que tomar una medida preventiva, como lo es una llamada de recordatorio. En cuanto a la periodicidad semanal, el Kolgomorov-Smirnov es del 59.09 %, lo cual quiere decir que de cada 6 clientes cerca de 6 pagarán. En cuanto a los restantes, de igual manera habría que aplicar una medida preventiva para evitar pérdidas económicas para la SOFOM.

En el caso quincenal, se aprecia un Kolgomorov-Smirnov del 57.99 %. Aquí el porcentaje de predicción baja un poco, respecto al caso semanal, lo cual tendría lógica, ya que la mayoría de los créditos que se otorgan tienen una periodicidad semanal, que nos permite tener más información y, por ende, realizar un pronóstico más certero, aunque un Kolgomorov-Smirnov del 57.99 % es un porcentaje igualmente importante. Le sigue el caso de los clientes con periodicidad mensual, cuyo Kolgomorov-Smirnov de 57.10 % baja, con respecto ahora, incluso, a los clientes con periodicidad quincenal, y esto es porque el monto menor de clientes de la SOFOM es mensual. Nunca se daría el caso idóneo de tener un KS del 100 %; sin embargo, el objetivo que se persigue es obtener el Kolgomorov-Smirnov mayor.

A continuación se muestra el cálculo de la divergencia de las dos poblaciones estudiadas (% Goods y % Bads)

Calculo de estadísticos de prueba para validar predictibilidad del Modelo Prevent										
Prevent Score	CLIENTES	GOOD	Dist g	BAD	dist bad	Clase	M <sub>good</sub>	M <sub>bad</sub>	$[\sigma_{\wedge 2}]_{(Good)}$	$[\sigma_{\wedge 2}]_{(Bad)}$
50-200	2,169	2,161	5%	8	0%	125	270,125	1,000	177,727,334	2,095,118
200-250	3,227	3,205	7%	22	0%	225	721,125	4,950	111,812,780	3,729,868
250-300	4,050	3,970	8%	80	1%	275	1,091,750	22,000	74,274,434	10,469,144
300-350	6,360	6,216	13%	144	2%	325	2,020,200	46,800	46,811,875	13,995,234
350-400	12,305	11,982	25%	323	3%	375	4,493,250	121,125	16,209,377	22,130,008
400-450	6,858	6,264	13%	594	6%	425	2,662,200	252,450	1,094,654	26,634,239
450-500	3,810	3,161	7%	649	7%	475	1,501,475	308,275	12,633,556	16,980,187
500-550	2,674	2,121	4%	553	6%	525	1,113,525	290,325	27,188,332	6,906,110
550-600	2,519	1,832	4%	687	7%	575	1,053,400	395,025	48,805,545	2,619,719
600-650	2,995	1,946	4%	1,049	11%	625	1,216,250	655,625	88,470,071	144,870
650-700	3,322	1,872	4%	1,450	15%	675	1,263,600	978,750	129,700,521	2,121,251
700-750	3,188	1,545	3%	1,643	17%	725	1,120,125	1,191,175	151,574,402	12,795,291
750-MÁS	3,757	1,454	3%	2,303	24%	775	1,126,850	1,784,825	191,823,823	44,016,294
<b>Total</b>	<b>57,234</b>	<b>47,729</b>	<b>100%</b>	<b>9,505</b>	<b>100%</b>	<b>Medias =</b>	<b>412</b>	<b>637</b>	<b>22,589</b>	<b>17,321</b>
									<b>D =</b>	<b>19.2</b>





# Conclusiones

Una de las mejores herramientas estadísticas para construir los modelos de riesgo es la regresión logística. Es una técnica que predice la ocurrencia de determinados eventos, basándose en las características de los individuos de dichos eventos, siendo éstos en la mayoría de los casos un hecho dicotómico.

En muchas ocasiones de nuestra vida diaria deseamos predecir la conducta humana, y ésta depende de un comportamiento previo (ventana de tiempo) y de innumerales variables cualitativas que caracterizan a cada individuo. La regresión logística resulta una probabilidad condicional, nuestra variable dependiente es nuestra predicción (se comportará de esta manera o no) y nuestras variables independientes son aquellas variables cualitativas que posteriormente convertimos en cuantitativas (como ya vimos anteriormente en el modelo). El software SAS nos permitió obtener los coeficientes de la regresión logística, para posteriormente concluir gracias a un K-S, empleando la ventana de tiempo considerado con anterioridad, que nuestro punto de corte sería de 750 puntos de SCORE con una efectividad de 64 % del K-S. El modelo que se propuso puede aplicarse a cualquier empresa, mientras que se quiera predecir un evento dicotómico.

El prevent score efectivamente predice el comportamiento de pago en la fecha límite de pago para todas las frecuencias de pago ya que la validación no mostró diferencias significativas de KS. Una vez que podemos predecir que clientes no realizaran su pago en la FLP (fecha límite de pago) podemos generar estrategias de cobranza preventiva para que el cliente no llegue a niveles de mora avanzada que le provoquen pérdidas por crédito a la institución.



# Anexo

```
/**VALIDACIÓN DE LA PREVENTIVA CON UN MES POSTERIOR A MAYO  
Gabriela Manzanera***/
```

```
LIBNAME DIARIOS "D:\Bases\DIARIOS";  
LIBNAME ORIGINA "D:\Bases\Originacion";
```

```
/**Importar exigibilidades de Mayo***/
```

```
data DIARIOS.EXIGIBILIDADES_mayo2011 ;  
  %let _EFIERR_ = 0; /* set the ERROR detection macro variable */  
  infile 'D:\Bases\DIARIOS\Exigibilidades_mayo2011.txt'  
delimitter='09'x MISSOVER DSD  
  lrecl=32767 firstobs=2 ;  
  informat N__de_Contrato $13. ;  
  informat Fecha_De_Vto_ date9. ;  
  format N__de_Contrato $13. ;  
  format Fecha_De_Vto_ date9. ;  
input  
      N__de_Contrato $  
      Fecha_De_Vto_  
;  
  if _ERROR_ then call symput('_EFIERR_',1); /* set ERROR  
detection macro variable */  
run;
```

```
/**Conciliar Exigibilidades de Mayo 2011***/
```

```
PROC SQL; CREATE TABLE DIARIOS.Exigibilidades_mayo2011 AS  
SELECT A.*  
      , CASE WHEN C.Acct_num IS NULL THEN A.N__de_Contrato ELSE  
C.Acct_num_nuevo END AS CONTRATO  
FROM DIARIOS.Exigibilidades_mayo2011 AS A  
LEFT JOIN DIARIOS.CONCILIACIONb AS C  
ON A.N__de_Contrato = C.Acct_num  
; QUIT;
```

```
/**Colocar la bandera a las Exigibilidades***/
```

```
PROC SQL; CREATE TABLE DIARIOS.Exigibilidades_mayo2011 AS  
SELECT E.Contrato  
      ,E.Fecha_De_Vto_  
      ,R.DiasMoroso  
from DIARIOS.Exigibilidades_mayo2011 as E  
LEFT JOIN DIARIOS.TBL_REPCARTERA_V8 AS R  
ON E.Contrato=R.CONTRATO  
WHERE E.Fecha_De_Vto_= R.FecSeguimiento  
AND R.DiasMoroso=0  
;quit;  
data DIARIOS.Exigibilidades_mayo2011;set  
DIARIOS.Exigibilidades_mayo2011;rename  
DiasMoroso=Morosidad_Exigibilidad;run;
```

```
PROC SQL; CREATE TABLE DIARIOS.Exigibilidades_mayo2011 AS  
SELECT C.*  
      ,R.DiasMoroso  
      ,Case when R.DiasMoroso>0 then "BAD" ELSE "GOOD" END AS  
FLAG_GOOD_BAD  
from DIARIOS.Exigibilidades_mayo2011 as C
```

```

LEFT JOIN DIARIOS.TBL_REPCARTERA_V8 AS R
ON C.Contrato=R.CONTRATO
where C.Fecha_De_Vto_+2=R.FecSeguimiento
;quit;

```

```

/****LEFT JOIN CON ORIGINACION****/

```

```

PROC SQL; CREATE TABLE RESUMEN AS
SELECT p.Contrato
      , p.diasmoroso
      , p.FLAG_GOOD_BAD
      , P.Fecha_De_Vto_
      , O.Fecha_Nacimiento
      , O.Fecha_Alta
      , O.Sexo
      , O.Estado
      , O.Nivel_Estudios
      , O.Automovil
      , O.Contrato_laboral
      , FLOOR( (P.Fecha_De_Vto_ - O.Fecha_Nacimiento)/365 ) AS EDAD
      , O.NoDependientes
      , O.AntigDom
      , O.AntigTrabajo
      , O.SCORE_BUENO
      , O.ATP
      , O.EdoCiv
      , O.TipoVivienda
      , O.Sucursal
      , O.CP
      , O.BC_Score
      , CASE WHEN O.BC_SCORE_FINAL=-9 THEN'NO_HIT' ELSE 'HIT' END
AS FLAG_HIT_NOHIT
      , O.Capital
      , O.Plazo
      , O.Tipo_Cr_dito
      , O.Producto
      , O.Micro_cr_dito_Ubicaci_n
      , O.Micro_cr_dito_Tipo
      , O.Micro_cr_dito_Lugar
      , O.Horario_laboral
      , O.TipoSolicitud
      , O.Periodicidad
      , FLOOR(O.Ingreso_Fijo + O.Ingreso_variable) AS IngresoTotal
      , O.Gastos_Bur__Cr_dito
      , O.Pago_mensual_pr_stamo
      , FLOOR(O.Renta + O.Servicios + O.Otros_Gastos) AS
GastosTotales
      , O.Fuentes_Ingreso_variables
      , O.Como_se_Entero
      ,Case when O.Tel_Casa is not null then 1 else 0 end as Tel_1
      ,Case when O.Tel_Movil is not null then 1 else 0 end as Tel_
2
      ,CASE WHEN O.Tel_Casa is not null then 1 else 0 end as Tel_3
      ,CASE WHEN O.Tel_Movil is not null then 1 else 0 end as Tel_
4
      ,CASE WHEN O.Tel_Arrendador is not null then 1 else 0 end as
Tel_5
      ,CASE WHEN O.Tel1_Empleo is not null then 1 else 0 end as
Tel_6
      ,CASE WHEN O.Tel2_Empleo is not null then 1 else 0 end as
Tel_7
      ,CASE WHEN O.Telefono_Cosolicitante is not null then 1 else
0 end as Tel_8

```

```

,CASE WHEN O.Tel_Ref_Fam1 is not null then 1 else 0 end as
Tel_9
,CASE WHEN O.Tel_Ref_Fam2 is not null then 1 else 0 end as
Tel_10
,CASE WHEN O.Tel_Ref_Psn11 is not null then 1 else 0 end
as Tel_11
,CASE WHEN O.Tel_Ref_Psn12 is not null then 1 else 0 end
as Tel_12
,O.TipoComprobanteIngresos
,O.FrecuenciaPrecepcionIngresos
,FLOOR((P.Fecha_De_Vto_ - O.Fecha_alta)/7) as wob
,CASE WHEN C.acct_num_nuevo IS NOT NULL THEN 'ANFEXI'
WHEN LENGTH(P.CONTRATO) = 9 THEN 'ANFEXI' ELSE
'FINAY' END AS SISTEMA

```

```

FROM DIARIOS.Exigibilidades_mayo2011 AS P
LEFT JOIN ORIGINA.Originacion18a AS O
ON p.Contrato = O.CONTRATO
LEFT JOIN DIARIOS.CONCILIACIONB AS C
ON P.CONTRATO = C.acct_num_nuevo

```

```
; QUIT;
```

```

PROC SQL; CREATE TABLE RESUMEN AS
SELECT O.*
, FLOOR(O.Tel_1+ O.Tel_2+ O.Tel_3+ O.Tel_4 + O.Tel_5+ O.Tel_6
+O.Tel_7+O.Tel_8+O.Tel_9+O.Tel_10+O.Tel_11+O.Tel_12) as TELS
FROM RESUMEN AS O
;QUIT;

```

```
/**PARA INCLUIR LA VARIABLE SISTEMA_DE_NACIMIENTO**/
```

```

/*VARIABLES COMPORTAMIENTO*/
/*

```

```

1.MAXIMO DE DÍAS DE MOROSIDAD
2.PROMEDIO DE DÍAS QUE EL CLIENTE SE TARDA EN PAGAR LA
EXIGIBILIDAD
3.PORCENTAJE DE CAPITAL YA PAGADO
4.IMPAGOS ULTIMOS N MESES (N=5)
5.MAX NÚM DE RECIBOS VENCIDOS
6.NUMERO DE VECES EN 1 RV , 2 RV, 3 RV, ETC
7.CUOTAS YA PAGADAS
8.NÚMERO DE CONDONACIONES
9.NÚMERO DE RENOVACIONES
*/

```

```
/* 1.MAXIMO DE DÍAS DE MOROSIDAD*/
```

```
PROC SQL; CREATE TABLE Maxdiasmoroso AS /**MÁXIMO DÍAS QUE TARDA EN
PAGAR EL CLIENTE**/ /***/
```

```

SELECT T.CONTRATO
, E.Fecha_De_Vto_
, MAX(T.DIASMOROSO) AS MAX_DM
, MAX(recibosvencidos) AS MAX_RV
, AVG(T.DIASMOROSO) AS AVG_DM

```

```

FROM DIARIOS.TBL_REPCARTERA_V8 AS T
INNER JOIN DIARIOS.EXIGIBILIDADES_mayo2011 AS E
ON T.CONTRATO = E.Contrato
AND T.FECSEGUIMIENTO <= E.Fecha_De_Vto_
GROUP BY T.CONTRATO, E.Fecha_De_Vto_ ;QUIT;

```

```

proc sql; create table Resumen as
Select R.*
      , M.MAX_DM
      , M.MAX_RV
      , M.AVG_DM
from Resumen as R
inner join Maxdiasmoroso as M
      on R.Contrato=M.CONTRATO
      and R.Fecha_De_Vto_= M.Fecha_De_Vto_
;quit;

/*3.PORCENTAJE DE CAPITAL YA PAGADO */

/*****

proc sql; create table Resumen as
select R.*

((T.CapitalVigente+T.CapitalTramite+T.capvencexigible+T.capvencnoexig
ible)/T.Total)*100 as pcappag

FROM Resumen AS R
      left join DIARIOS.TBL_REPCARTERA_V8 as T
      on R.Contrato=T.Contrato
      AND T.FECSEGUIMIENTO = R.Fecha_De_Vto_
      ;QUIT;

/*4.IMPAGOS ULTIMOS N MESES*/

/**Número de Impagos los últimos n meses***/
/** Con n=5**/
PROC IMPORT OUT= WORK.pagosa
      DATAFILE= "D:\Bases\DIARIOS\vencimiento.txt"
      DBMS=TAB REPLACE;
      GETNAMES=YES;
      DATAROW=2;
RUN;

PROC SQL; CREATE TABLE Pagosa AS
      select P.*
      , case when c.Acct_num_nuevo is null then P.n_de_contrato
      else c.Acct_num_nuevo end as N__de_Contrato
      from pagosa as p
      left join DIARIOS.conciliaciona as c
      on P.n_de_contrato= c.Acct_num
; quit;
data pagosa; set pagosa; rename fecha_de_pago=fechapago
fechaVencimiento=Fecha_Vencimiento_Cubierto
Forma_de_Pago_Cliente=Forma_de_Pago;run;
proc sql; create table pagosa as
select P.N__de_Contrato
      ,P.fechapago
      ,P.Fecha_Vencimiento_Cubierto
      ,P.Forma_de_Pago
      from pagosa as p
      ;quit;

```

```

PROC IMPORT OUT= WORK.pagosf
      DATAFILE= "D:\Bases\DIARIOS\Pagos FINAYv2.txt"
      DBMS=TAB REPLACE;
      GETNAMES=YES;
      DATAROW=2;
RUN;
proc sql; create table pagosf as
select P.N__de_Contrato
      ,P.fechapago
      ,P.Fecha_Vencimiento_Cubierto
      ,P.Forma_de_Pago
      from pagosf as p
      ;quit;

PROC SQL; CREATE TABLE Pagos AS
SELECT * FROM pagosa UNION ALL
SELECT * FROM pagosf ; QUIT;

proc sql; create table pagos1 as
select N__de_contrato, Fecha_De_Vto_ ,count(IMPAGOS) AS NUM_IMPAGOS
      FROM
      (
select t.*
      , R.Fecha_De_Vto_
      , case when t.Fechapago>t.Fecha_Vencimiento_Cubierto then
'Impago'
              when t.Fechapago=t.Fecha_Vencimiento_Cubierto then "pago"
              when t.fechapago<t.Fecha_Vencimiento_Cubierto then
"adelanto" else "nuncapago" end as impagos
      from pagos as t
      INNER JOIN RESUMEN AS R
              ON T.N__de_Contrato = R.CONTRATO
              AND t.Fecha_Vencimiento_Cubierto <= R.Fecha_De_Vto_
      )z
      WHERE Z.IMPAGOS='Impago'
              AND Z.Fechapago >="01JAN11"D
              AND Z.Fechapago <="31MAY11"d

GROUP BY N__de_contrato,Fecha_De_Vto_
;quit;

PROC SQL; CREATE TABLE Resumen AS
SELECT r.*
      ,case when P.num_IMPAGOS is null then 0 else P.num_IMPAGOS end
as Impagos
FROM RESUMEN AS R
LEFT JOIN PAGOS1 AS P
ON R.CONTRATO=p.N__de_Contrato
and r.Fecha_De_Vto_=p.Fecha_De_Vto_
;quit;

proc sql; create table adelanto as
select N__de_contrato, Fecha_De_Vto_ ,count(IMPAGOS) AS
num_adelantos
      FROM
      (
select t.*
      , R.Fecha_De_Vto_
      , case when t.Fechapago>t.Fecha_Vencimiento_Cubierto then
'Impago'
              when t.Fechapago=t.Fecha_Vencimiento_Cubierto then "pago"
              when t.fechapago<t.Fecha_Vencimiento_Cubierto then
"adelanto" else "nuncapago" end as impagos

```



```

from pagos as t
INNER JOIN RESUMEN AS R
      ON T.N__de_Contrato = R.CONTRATO
      AND t.Fecha_Vencimiento_Cubierto <= R.Fecha_De_Vto_
)z
WHERE Z.IMPAGOS="adelanto"
      AND Z.Fechapago >="01JAN11"D
      AND Z.Fechapago <="31MAY11"d

GROUP BY N__de_contrato, Fecha_De_Vto_
;quit;

PROC SQL; CREATE TABLE Resumen AS
SELECT r.*
      , case when P.num_adelantos is null then 0 else P.num_adelantos
end as adelantos
FROM Resumen AS R
LEFT JOIN adelanto AS P
ON R.CONTRATO=p.N__de_Contrato
and r.Fecha_De_Vto_=p.Fecha_De_Vto_
;quit;

proc sql; create table coutas_pagadas as
select p.N__de_contrato, r.Fecha_De_Vto_, count
(p.Fecha_Vencimiento_Cubierto) as ctaspagadas
from pagos as p
inner join resumen as r
      on p.N__de_contrato = r.contrato
      and p.fechapago <= r.Fecha_De_Vto_
group by 1, 2 ; quit;

PROC SQL; CREATE TABLE Resumen as
SELECT R.*
      ,CASE WHEN C.ctaspagadas IS NULL THEN 0 ELSE C.ctaspagadas END
AS ctas_pagadas_final
FROM RESUMEN AS R
LEFT JOIN COUTAS_PAGADAS AS C
ON R.contrato=c.n__de_contrato
and r.Fecha_De_Vto_=c.Fecha_De_Vto_

;QUIT;

PROC SQL; CREATE TABLE VECES_WK AS
SELECT CONTRATO
      , SUM(V1) AS V_1RV
      , SUM(V2) AS V_2RV
      , SUM(V3) AS V_3RV
      , SUM(V4) AS V_4RV
      , SUM(V5) AS V_5RV
      , SUM(V6) AS V_6RV
      , SUM(V7) AS V_7RV
      , SUM(V8) AS V_8RV
FROM
(
SELECT FECSEGUIMIENTO
      , CONTRATO
      , CASE WHEN recibosvencidos = 1 THEN 1 ELSE 0 END AS V1

```

```

, CASE WHEN recibosvencidos = 2 THEN 1 ELSE 0 END AS V2
, CASE WHEN recibosvencidos = 3 THEN 1 ELSE 0 END AS V3
, CASE WHEN recibosvencidos = 4 THEN 1 ELSE 0 END AS V4
, CASE WHEN recibosvencidos = 5 THEN 1 ELSE 0 END AS V5
, CASE WHEN recibosvencidos = 6 THEN 1 ELSE 0 END AS V6
, CASE WHEN recibosvencidos = 7 THEN 1 ELSE 0 END AS V7
, CASE WHEN recibosvencidos = 8 THEN 1 ELSE 0 END AS V8
, CASE WHEN recibosvencidos = 9 THEN 1 ELSE 0 END AS V9
, CASE WHEN recibosvencidos >= 10 THEN 1 ELSE 0 END AS V10
_MAS
FROM DIARIOS.TBL_REPCARTERA_V8 /*** CAMBIAR AQUI***/
WHERE FECSEGUIMIENTO <"01MAY11."D
AND weekday(FECSEGUIMIENTO) = 2
AND PERIODICIDAD IN ('SEMANAL', 'Semanal')
AND diasmoroso <= 30

```

```
) Z GROUP BY CONTRATO; QUIT;
```

```

PROC SQL;
CREATE TABLE DIARIOS.DIAS_QUINCENA AS

SELECT FECSEGUIMIENTO
FROM
(
SELECT ANO
, MES
, MIN(DIF_1_D) AS MIN_DIF_1_D
, MIN(DIF_16_D) AS MIN_DIF_16_D
, FECSEGUIMIENTO
FROM
(
SELECT FECSEGUIMIENTO
, YEAR(FECSEGUIMIENTO) AS ANO
, MONTH(FECSEGUIMIENTO) AS MES
, DAY(FECSEGUIMIENTO) AS DIA
, DAY(FECSEGUIMIENTO)- 1 AS DIF_1_D
, CASE WHEN DAY(FECSEGUIMIENTO) < 16 THEN 99 ELSE ABS(DAY
(FECSEGUIMIENTO)- 16) END AS DIF_16_D
FROM
(
SELECT FECSEGUIMIENTO, COUNT(CONTRATO) AS CLIENTES
FROM DIARIOS.TBL_REPCARTERA_V8 /*OJO CON LA VERSION */
GROUP BY FECSEGUIMIENTO
)Z
)Y
GROUP BY ANO, MES

)X
WHERE ( DAY(FECSEGUIMIENTO) = MIN_DIF_1_D+1 OR DAY(FECSEGUIMIENTO) =
MIN_DIF_16_D+16 )
ORDER BY FECSEGUIMIENTO
; QUIT;

```

```

PROC SQL; CREATE TABLE VECES_QN AS
SELECT CONTRATO
, SUM(V1) AS V_1RV
, SUM(V2) AS V_2RV
, SUM(V3) AS V_3RV
, SUM(V4) AS V_4RV
, SUM(V5) AS V_5RV
, SUM(V6) AS V_6RV

```

```

        , SUM(V7) AS V_7RV
        , SUM(V8) AS V_8RV
FROM
(
    SELECT FECSEGUIMIENTO
        , CONTRATO
        , CASE WHEN recibosvencidos = 1 THEN 1 ELSE 0 END AS V1
        , CASE WHEN recibosvencidos = 2 THEN 1 ELSE 0 END AS V2
        , CASE WHEN recibosvencidos = 3 THEN 1 ELSE 0 END AS V3
        , CASE WHEN recibosvencidos = 4 THEN 1 ELSE 0 END AS V4
        , CASE WHEN recibosvencidos = 5 THEN 1 ELSE 0 END AS V5
        , CASE WHEN recibosvencidos = 6 THEN 1 ELSE 0 END AS V6
        , CASE WHEN recibosvencidos = 7 THEN 1 ELSE 0 END AS V7
        , CASE WHEN recibosvencidos = 8 THEN 1 ELSE 0 END AS V8
        , CASE WHEN recibosvencidos = 9 THEN 1 ELSE 0 END AS V9
        , CASE WHEN recibosvencidos >= 10 THEN 1 ELSE 0 END AS V10
_MAS
    FROM DIARIOS.TBL_REPCARTERA_V8
    WHERE FECSEGUIMIENTO <="01MAY11"D
    AND FECSEGUIMIENTO IN (SELECT * FROM DIARIOS.DIAS_QUINCENA)
    AND PERIODICIDAD IN ('QUINCENAL', 'QUINCENAL', 'Quincenal')
    AND diasmoroso <= 30

) Z GROUP BY CONTRATO; QUIT;

```

```

PROC SQL;
CREATE TABLE DIARIOS.DIAS_MES AS

SELECT FECSEGUIMIENTO
    FROM
(
    SELECT ANO
        , MES
        , MIN(DIF_1_D) AS MIN_DIF_1_D
        , MIN(DIF_16_D) AS MIN_DIF_16_D
        , FECSEGUIMIENTO

FROM
(
    SELECT FECSEGUIMIENTO
        , YEAR(FECSEGUIMIENTO) AS ANO
        , MONTH(FECSEGUIMIENTO) AS MES
        , DAY(FECSEGUIMIENTO) AS DIA
        , DAY(FECSEGUIMIENTO)- 1 AS DIF_1_D
        , CASE WHEN DAY(FECSEGUIMIENTO) < 16 THEN 99 ELSE ABS(DAY
(FECSEGUIMIENTO)- 16) END AS DIF_16_D
FROM
(
    SELECT FECSEGUIMIENTO, COUNT(CONTRATO) AS CLIENTES
        FROM DIARIOS.TBL_REPCARTERA_V8 /*OJO CON LA VERSION*/
        GROUP BY FECSEGUIMIENTO
    )Z
    )Y
GROUP BY ANO, MES

)X
WHERE DAY(FECSEGUIMIENTO) = MIN_DIF_1_D+1
ORDER BY FECSEGUIMIENTO
; QUIT;

```

```

PROC SQL; CREATE TABLE VECES_MN AS
SELECT CONTRATO
    , SUM(V1) AS V_1RV

```

```

, SUM(V2) AS V_2RV
, SUM(V3) AS V_3RV
, SUM(V4) AS V_4RV
, SUM(V5) AS V_5RV
, SUM(V6) AS V_6RV
, SUM(V7) AS V_7RV
, SUM(V8) AS V_8RV
FROM
(
  SELECT FECSEGUIMIENTO
    , CONTRATO
    , CASE WHEN recibosvencidos = 1 THEN 1 ELSE 0 END AS V1
    , CASE WHEN recibosvencidos = 2 THEN 1 ELSE 0 END AS V2
    , CASE WHEN recibosvencidos = 3 THEN 1 ELSE 0 END AS V3
    , CASE WHEN recibosvencidos = 4 THEN 1 ELSE 0 END AS V4
    , CASE WHEN recibosvencidos = 5 THEN 1 ELSE 0 END AS V5
    , CASE WHEN recibosvencidos = 6 THEN 1 ELSE 0 END AS V6
    , CASE WHEN recibosvencidos = 7 THEN 1 ELSE 0 END AS V7
    , CASE WHEN recibosvencidos = 8 THEN 1 ELSE 0 END AS V8
    , CASE WHEN recibosvencidos = 9 THEN 1 ELSE 0 END AS V9
    , CASE WHEN recibosvencidos >= 10 THEN 1 ELSE 0 END AS V10
  _MAS
  FROM DIARIOS.TBL_REPCARTERA_V8 /*OJO CON LA VERSION */
  WHERE FECSEGUIMIENTO <="01MAY11"D
  AND FECSEGUIMIENTO IN (SELECT * FROM DIARIOS.DIAS_MES)
  AND PERIODICIDAD IN ('MENSUAL', 'Mensual')
  AND diasmoroso <= 30

) Z GROUP BY CONTRATO; QUIT;

PROC SQL; CREATE TABLE VECES AS
SELECT * FROM VECES_WK UNION ALL
SELECT * FROM VECES_QN UNION ALL
SELECT * FROM VECES_MN ; QUIT;

proc sql; create table resumen as
select R.*
,case when v.V_1RV is null then 0 else v.V_1RV end as V1RV
,case when v.V_2RV is null then 0 else v.V_2RV end as V2RV
,case when v.V_3RV is null then 0 else v.V_3RV end as V3RV
,case when v.V_4RV is null then 0 else v.V_4RV end as V4RV
,case when v.V_5RV is null then 0 else v.V_5RV end as V5RV
,case when v.V_6RV is null then 0 else v.V_6RV end as V6RV
,case when v.V_7RV is null then 0 else v.V_7RV end as V7RV
,case when v.V_8RV is null then 0 else v.V_8RV end as V8RV

  from resumen as r
  left join veces as v
  on r.Contrato=v.Contrato

;quit;

/****Conciliar las condonaciones****/

data DIARIOS.CONDONACIONES
;
%let _EFIERR_ = 0; /* set the ERROR detection macro variable */
infile 'C:\Documents and Settings\m914089\Escritorio
\CONDONACIONES.txt' delimiter='09'x
lrecl=32767 firstobs=2 ;
informat Contrato $12. ;
informat MES date9. ;

format Contrato $12. ;

```

```

format MES date9. ;
INPUT
      CONTRATO $
      MES
/* set ERROR detection macro variable */
; run;

```

```

PROC SQL; CREATE TABLE CONDONACIONES AS
SELECT A.contrato
      ,A.mes
      , CASE WHEN C.Acct_num IS NULL THEN A.CONTRATO ELSE
C.Acct_num_nuevo END AS CONTRATO5
FROM DIARIOS.condonaciones AS A
LEFT JOIN DIARIOS.CONCILIACIONb AS C
      ON A.CONTRATO = C.Acct_num
; QUIT;

```

```

proc sql; create table condonaciones as
select r.CONTRATO,r.Fecha_De_Vto_,count(p.contrato5) as condonaciones
from CONDONACIONES as p
inner join Resumen as r
      ON p.contrato5= r.contrato
      And p.mes <= r.Fecha_De_Vto_
group by r.CONTRATO, r.Fecha_De_Vto_

;quit;

```

```

proc sql; create table Resumen as
select t.*
      ,CASE WHEN C.CONDONACIONES IS NULL THEN 0 ELSE C.CONDONACIONES
END AS NUM_CONDONACIONES
from resumen as t
left join CONDONACIONES as c
on T.CONTRATO=C.CONTRATO
AND T.Fecha_De_Vto_=C.Fecha_De_Vto_

;QUIT;

```

/\*\*CONCILIAR RENOVACIONES\*\*\*/

```

PROC SQL; CREATE TABLE RENOVACIONES AS
SELECT A.*
      , CASE WHEN C.Acct_num IS NULL THEN A.Origin ELSE
C.Acct_num_nuevo END AS CONTRATO_CONCILIADO
FROM RENOVACIONES AS A
LEFT JOIN DIARIOS.CONCILIACIONb AS C
      ON A.Origin = C.Acct_num
; QUIT;
PROC SQL; CREATE TABLE Resumen AS
SELECT R.*
      , CASE WHEN C.Acct_num_nuevo IS NULL THEN R.Contrato ELSE
C.Acct_num_nuevo END AS CONTRATO_VIEJO
FROM Resumen AS R
LEFT JOIN DIARIOS.CONCILIACIONb AS C
      ON R.Contrato = C.Acct_num_nuevo
; QUIT;

```

```

PROC SQL; CREATE TABLE RESUMEN AS
SELECT R.*
      ,CASE WHEN G.A IS NOT NULL THEN 1

```

```

        WHEN G.B IS NOT NULL THEN 2
          WHEN G.C IS NOT NULL THEN 3
            WHEN G.D IS NOT NULL THEN 4
              WHEN G.E IS NOT NULL THEN 5
                WHEN G.F IS NOT NULL THEN 6
                  WHEN G.G IS NOT NULL THEN 7
                ELSE 0 END AS NUM_RENOVACIONES
FROM resumen AS R
LEFT JOIN RENOVACIONES AS G
on r.CONTRATO_VIEJO=g.a
AND g.Fecha_renovacion<=R.Fecha_De_Vto_

;QUIT;

proc sql; create table Insumo_ler_tabla as
select CONTRATO
      , FLAG_GOOD_BAD
      , CASE WHEN Sexo IN ('FEMENINO') THEN 'FEMENINO'
              WHEN Sexo in ('MASCULINO') THEN 'MASCULINO' ELSE
"NULO" END AS SEXO_
      , case when Estado in ('DISTRITO FEDERAL','DISTRITO FEDERAL
D.F') then 'DISTRITO FEDERAL'
              WHEN Estado in ('ESTADO DE MEXICO','ESTADO DE MEXICO
MEX') THEN 'ESTADO DE MEXICO' ELSE 'OTRO' END AS EDO
      , CASE WHEN Nivel_Estudios IN
('DIPLOMADOS','Licenciatura','UNIVERSIDAD','POSTGRADO
(MAESTRIA)','Postgrado') THEN 'LICENCIATURA/DIPLOMADOS/POSTGRADO'
              WHEN Nivel_Estudios IN ('PREPARATORIA','Técnico o
Preparatoria') THEN 'Técnico o Preparatoria'
              WHEN Nivel_Estudios IN
('PRIMARIA','Primaria','SECUNDARIA','Secundaria') THEN 'Primaria o
Secundaria' ELSE 'NULO'END AS ESTUDIOS

      , CASE WHEN AUTOMOVIL IN ('SI') THEN 'SI'
              WHEN AUTOMOVIL IS NULL THEN "NO" ELSE 'NULO'END AS
AUTOMOVIL_
      , case when edad>=1 AND edad<=22 THEN "1 a 22"
              WHEN edad>=23 AND edad<=36 THEN "23 a 36"
              WHEN edad>=37 THEN "37 MAS" ELSE "NULO" END AS EDADES

      , CASE WHEN NoDependientes=0 THEN "0"
              when nodependientes is null then "0"
              WHEN NoDependientes=1 THEN "1"
              WHEN NoDependientes=2 then "2 "
              WHEN NoDependientes>2 THEN "2-MAS" END AS
NO_DEPENDIENTES

      ,CASE WHEN SCORE_BUENO<=550 THEN "<=550"
              WHEN SCORE_BUENO> 550 AND SCORE_BUENO< 570 THEN '550a570'
              WHEN SCORE_BUENO>= 570 AND SCORE_BUENO< 590 THEN
'570a590'
              WHEN SCORE_BUENO>= 590 AND SCORE_BUENO< 610 THEN
'590a610'
              WHEN SCORE_BUENO>= 610 AND SCORE_BUENO< 630 THEN
'610a630'
              WHEN SCORE_BUENO> 630 THEN
'630-MAS' ELSE 'NULO' END AS RANG_SCORE

      ,CASE WHEN ATP<=50 THEN "<=50"
              WHEN ATP>50 THEN '50MAS'
              ELSE "NULO" END AS ATP_

      , CASE WHEN EDOCIV IN ('CASADA/O BIENES S', 'CASADO/A BIENES

```

```

M', 'Casado','VIUDO/A', 'Viudo' ) THEN 'CASADO/VIUDO'
      WHEN EDOCIV IN ('DIVORCIADO/A', 'Divorciado',
'SOLTERO/A', 'Soltero') THEN 'SOLTERO/DIVORCIADO'
      WHEN EDOCIV IN ('SEPARADO/A', 'Separado','UNION
LIBRE', 'Unión Libre') THEN 'SEPARADO/UNIÓN LIBRE' ELSE 'NULO' END AS
ESTADO_CIVIL
      , CASE WHEN TIPOVIVIENDA IN
('FAMILIAR','Familiares','RENTADA','Rentada','PAGANDO','Pagandola')
THEN 'FAMILIARES/RENTADA/PAGÁNDOLA'
      WHEN TIPOVIVIENDA IN ( 'PROPIA','Propia') THEN
'PROPIA' ELSE 'NULO' END AS TIPO_VIVIENDA

      , CASE WHEN BC_SCORE<=650 THEN "<=650"
      WHEN BC_SCORE>650 AND BC_SCORE<=700 THEN '6500 a 700'
      WHEN BC_SCORE>700 AND BC_SCORE<=750 THEN '700 a
750'
      WHEN BC_SCORE>750
      THEN '750 mas' ELSE 'NULO' END AS RANG_BCSCORE

      , CASE WHEN CAPITAL<=5000 THEN "<=5000"
      WHEN CAPITAL>5000 THEN ">5000" ELSE 'NULO' END AS
MONTO

      , CASE WHEN PLAZO<=50 THEN "<=50"
      WHEN PLAZO>50 THEN ">50"
      ELSE 'NULO' END AS T_PLAZO

      , CASE WHEN Tipo_Cr_dito IN ('CONSUMO') THEN 'CONSUMO'
      WHEN Tipo_Cr_dito IN ('MICROCREC','MICRONEGOCIO') THEN
'MICRONEGOCIO' ELSE 'NULO' END AS TIPO_CREDITO

      , CASE WHEN Producto IN
('PREFERENCI','Prefere','Preferencial') THEN 'Preferencial'
      WHEN Producto IN ('REGULAR', 'Regular') THEN
'REGULAR' ELSE 'NULO' END AS TIPO_PRODUCTO

      , CASE WHEN Micro_cr_dito_Ubicaci_n IN('DE CASA EN
CASA','CASA','CASA FAMILIAR','DESDE CASA','DESDE SU CASA
DOMICILIO','EN CASA','EN SU CASA','EN EL HOGAR','EN EL LUGAR
DONDE RE','PATIO CASA','VENTA DESDE SU CASA',' VENTA EN SU
CASA','HOGAR',' CASA FAMILIAR','FAMILIAR' 'TRABA EN EL HOGAR')
THEN 'DOMICILIO'
      WHEN Micro_cr_dito_Ubicaci_n LIKE ('CASA') THEN
'DOMICILIO'
      WHEN Micro_cr_dito_Ubicaci_n IS NULL THEN 'NULO'
      WHEN Micro_cr_dito_Ubicaci_n IN('LOCAL') THEN
'LOCAL' ELSE 'OTRO' END AS UBICACION_MICROCREDITO

      , CASE WHEN Micro_cr_dito_Tipo IN ("PROPIO","CASA","EN SU
CASA","EN SU MISMO DOMICILI","FAMILIAR","FAMILIAR UBICADO EN
HOGAR") THEN 'PROPIO'
      WHEN Micro_cr_dito_Tipo IS NULL THEN 'NULO' ELSE 'NO
PROPIO' END AS TIPO_MICROCREDITO

      , CASE WHEN Micro_cr_dito_Lugar IN
('FIJO','LOCAL','SEMIFIJO','PUESTO') THEN 'FIJO'
      WHEN Micro_cr_dito_Lugar IN ('MOVIL','OTROS') THEN
'MOVIL' ELSE 'NULO' END AS Tipo_MicrocrEdito

      , CASE WHEN TipoSolicitud IN ('FORMER') THEN 'FORMER'
      WHEN TipoSolicitud IN('NORMAL','NUEVO') THEN 'NUEVO'
      WHEN TipoSolicitud IN('REESTR','REESTRUCTURA',) THEN
'REESTRUCTURA'
      WHEN TipoSolicitud IN('RENOVA','RENOVADA','RENOVADO')
THEN 'RENOVACIÓN 'ELSE 'NULO' END AS TIPO_SOLICITUD

      , CASE WHEN Periodicidad in ('MENSUAL','Mensual') then
'MENSUAL'
      WHEN Periodicidad in ('QUINCENAL','Quincenal')

```

```

THEN 'QUINCENAL'
      WHEN Periodicidad in ('SEMANAL','Semanal') THEN
'SEMANAL' ELSE 'NULO' END AS TIPO_PERIODICIDAD

      , CASE WHEN IngresoTotal<=10000 THEN '<=10000'
      WHEN IngresoTotal>10000 THEN '>10000'
      ELSE 'NULO' END AS INGRESO_TOTAL

      , CASE WHEN Gastos_Bur__Cr_dito>=0 AND Gastos_Bur__Cr_dito<=
500 THEN '0 a 500'
      WHEN Gastos_Bur__Cr_dito>500
      THEN '500 mas' ELSE 'NULO' END AS GTOS_BURO_CREDITO

      , CASE WHEN Pago_mensual_pr_stamo<=300 THEN "<=300"
      WHEN Pago_mensual_pr_stamo>300 AND
Pago_mensual_pr_stamo<=1000 THEN '300 a 1000'
      WHEN Pago_mensual_pr_stamo>1000 AND Pago_mensual_pr_stamo
<=3000 THEN '1000 a 3000'
      WHEN Pago_mensual_pr_stamo>3000 then '3000 mas'
      ELSE 'NULO' END AS PAGO_MENSUAL

      , CASE WHEN GastosTotales<=500 THEN "<=500"
      WHEN GastosTotales>500 AND GastosTotales<=899 THEN
'500-899'
      WHEN GastosTotales>899
      THEN '>899' ELSE 'NULO' END AS GTOS_TOT

      , CASE WHEN Como_se_Enterо IS NULL THEN 'NULO'
      WHEN Como_se_Enterо IN ("ANUNCIO AZOTEA", "ANUNCIO
CALLE/MANTA", "APASCO", "CAMBACEO",
"COMISIONISTA", "INTERNET", "LETRERO", "LLAMADA TELEFONICA"
,"MARQUESINA", "MODULO", "PASANDO POR
SUCURSAL", "PERIODICO", "POSTER / MICRONEG", "POSTER
MICRONEGOCIO", "REVISTA", "VOCEO", "VOLANTE",
"NINGUNO", "AMISTAD / CONOCIDO",
"RECOMENDACION", "FAMILIAR") THEN 'PUBLICIDAD'
      WHEN Como_se_Enterо IN ("RENOVACION", "REESTRUCTURA"
,"FORMER") THEN "RENOVACION/REESTRUCTURA/FORMER" ELSE 'NULO'
END AS CENTERO

      ,CASE WHEN TipoComprobanteIngresos IN ("CTA HABERES", "Carta
de haberes", "NOTA COMPRA", "Notas de compra", 'Carta laboral a mano',
'Carta laboral membretada') then 'CARTA LABORAL'
      WHEN TipoComprobanteIngresos IN ("EDO DE CTA", "Estado
de Cuenta", "Estado de cuenta", "Estados de
Cuenta", "FACTURAS", "Facturas") then 'FACTURAS'
      WHEN TipoComprobanteIngresos IN ("REC PENSION"
"Recibos de pensión" "NOMINA" "Nómina" "DECLARACION"
"Declaración de Impuestos") then 'NÓMINA'
      when TipoComprobanteIngresos is null then 'NULO' ELSE
'OTRO' END AS TIP_CPTE_INGR

      , CASE WHEN FrecuenciaPrecepcionIngresos IN ("DIARIO",
"Diario") then "DIARIO"
      WHEN FrecuenciaPrecepcionIngresos IN ("MENSUAL",
"Mensual") then "MENSUAL"
      WHEN FrecuenciaPrecepcionIngresos IN ("QUINCENAL",
"Quincenal", "VARIABLE", "Variable") then "QUINCENAL/VARIABLE"
      WHEN FrecuenciaPrecepcionIngresos IN ("SEMANAL",
"Semanal") then "SEMANAL"
      WHEN FrecuenciaPrecepcionIngresos is null then
'NULO' ELSE 'OTRO' END AS FREC_PREC_INGR

      , CASE WHEN WOB>=0 AND WOB<=5 THEN '0 a 5'
      WHEN WOB>5 AND WOB <=10 THEN '5 a 10'
      WHEN WOB>10 AND WOB <=20 THEN '10 a 20'

```



```

        WHEN WOB>20 AND WOB <=40 THEN '20 a 40'
        WHEN WOB>40 THEN '40 mas' ELSE 'NULO'
        END AS WOB_

, CASE WHEN SISTEMA='ANFEXI' THEN 'ANFEXI'
        WHEN SISTEMA='FINAY' THEN 'FINAY' ELSE 'NULO' END AS
NAC_SIST

, CASE WHEN TELS<=4 THEN '<=4'
        WHEN TELS>4 THEN '>4' ELSE 'NULO' END AS NUM_TELS

, CASE WHEN MAX_DM = 0 THEN '0'
        WHEN MAX_DM >= 1 AND MAX_DM <= 2 THEN '1-2'
        WHEN MAX_DM >= 3 AND MAX_DM <= 6 THEN '3-6' /***/
        WHEN MAX_DM >= 7 THEN '7 mas'
        ELSE 'NULO' END AS MAXIMO_DM

,CASE WHEN MAX_RV=0 then '0'
        WHEN MAX_RV>=1 AND MAX_RV<=4 THEN '1 a 4'
        WHEN MAX_RV>5 THEN '5mas/nulo' ELSE '5mas/nulo' END
AS MAXIMO_RV

, case when avg_dm = 0 then '0'
        WHEN avg_dm > 0 AND avg_dm <= 0.318 THEN '0-0.318'
        WHEN avg_dm > 0.318 THEN '0.318-+'
        ELSE 'NULO' END AS PROMEDIO_DM

, CASE WHEN pcappag>=0 AND pcappag<=10 THEN '0 a 10'
        WHEN pcappag>10 AND pcappag<=30 THEN '10 a 30'
        WHEN pcappag>30 AND pcappag<=50 THEN '30 a 50'
        WHEN pcappag>50 AND pcappag<=70 THEN '50 a 70'
        WHEN pcappag>70 and pcappag<=80 then '70 a 80'
        WHEN pcappag>80 and pcappag<=90 then '80 a 90'
        when pcappag>90 then '90mas'
        else 'NULO' end as pct_cappag

, CASE WHEN ctas_pagadas_final>=0 AND ctas_pagadas_final<=5
then '0 a 5'
        WHEN ctas_pagadas_final>5 AND ctas_pagadas_final<=10 THEN
'5 a 10'
        WHEN ctas_pagadas_final>10 AND ctas_pagadas_final<=15
THEN '10 a 15'
        WHEN ctas_pagadas_final>15 AND ctas_pagadas_final<=20
THEN '15 a 20'
        WHEN ctas_pagadas_final>20 AND ctas_pagadas_final<=40
THEN '20 a 40'
        WHEN ctas_pagadas_final>40 AND ctas_pagadas_final<=60
THEN '40 a 60'
        WHEN ctas_pagadas_final>60 AND ctas_pagadas_final<=80
THEN '60 a 80'
        WHEN ctas_pagadas_final>80 AND ctas_pagadas_final<=
100 THEN '80 a 100'
        WHEN ctas_pagadas_final>100 THEN '100 mas' ELSE
'NULO' END AS CTASPAGADAS

, CASE WHEN NUM_CONDONACIONES=0 THEN '0'
        WHEN NUM_CONDONACIONES>1 THEN 'MAS QUE CERO'
        ELSE 'NULO' END AS CONDONACIONES

, CASE WHEN Impagos= 0 then '0'

        when Impagos=1 then '1'
        WHEN Impagos>=2 AND Impagos<=3 THEN '2 a 3'
        WHEN Impagos>3 AND Impagos<=4 THEN '3 a 4'
        WHEN Impagos>4 AND Impagos<=5 THEN '4 a 5'

```

```

        WHEN Impagos>5 AND Impagos<=10 THEN '5 a 10'
        WHEN Impagos>10 AND Impagos<=15 THEN '10 a 15'
        WHEN Impagos>15

        THEN '15 mas' ELSE 'NULO' END AS NUM_IMPAGOS

, CASE when adelantos=0 then '0'
      when adelantos=1 then '1'
      WHEN Adelantos>1 AND Adelantos<=2 THEN '1 a 2'
      WHEN Adelantos>2 AND Adelantos<=3 THEN '2 a 3'
      WHEN Adelantos>3 AND Adelantos<=4 THEN '3 a 4'
      WHEN Adelantos>4 AND Adelantos<=5 THEN '4 a 5'
      WHEN Adelantos>5 AND Adelantos<=10 THEN '5 a 10'
      WHEN Adelantos>10 AND Adelantos<=15 THEN '10 a 15'
      WHEN Adelantos>15 AND Adelantos<=20 THEN '15 a 20'
      WHEN Adelantos>20
      then '20mas' else 'NULO' END AS NUM_ADELANTOS

, CASE WHEN V1RV=0 THEN '0'
      WHEN V1RV>1
      THEN '1 mas' ELSE 'NULO' END AS V_R1

, CASE WHEN V2RV=0 THEN '0'
      when v2rv>1 then '1MAS'
      ELSE 'NULO' END AS V_R2

, CASE WHEN V3RV=0 THEN '0'
      when v3rv>1 THEN 'lmas' ELSE 'NULO' END AS V_R3

, CASE WHEN V4RV=0 THEN '0 '
      when v4rv>1 THEN 'lmas' ELSE 'NULO' END AS V_R4

, CASE WHEN V5RV=0 then '0'
      when v5rv>1 THEN '1MAS' ELSE 'NULO' END AS V_R5

, CASE WHEN V6RV=0 then '0'
      when v6rv>1 then'lmas' else 'NULO' END AS v_r6

, CASE WHEN NUM_RENOVACIONES=0 THEN '0'
      WHEN NUM_RENOVACIONES>0 THEN 'MAS DE CERO' ELSE
'NULO' END AS RENOVACIONES

FROM RESUMEN
;QUIT;

```

```

PROC SQL; CREATE TABLE num_impagos AS
SELECT impagos , COUNT(CONTRATO) AS CLIENTES, SUM(BADS) AS BAD
FROM
(
SELECT num_impagos
      , CONTRATO
      , CASE WHEN FLAG_GOOD_BAD = 'BAD' THEN 1 ELSE 0 END AS BADS
FROM Insumo_1er_tabla
)
GROUP BY 1
;QUIT;

```

```

/*****/

```

```

proc sql; create table Insumo_2da_tabla as

```

```

select CONTRATO
,FLAG_GOOD_bad
,CASE WHEN FLAG_GOOD_BAD='GOOD' THEN 0
      WHEN FLAG_GOOD_BAD='BAD' THEN 1 ELSE 0 END AS BANDERA
,CASE WHEN Sexo IN ('FEMENINO') THEN 1
      WHEN Sexo in ('MASCULINO') THEN 2 ELSE 2 END AS SEXO_
,case when Estado in ('DISTRITO FEDERAL','DISTRITO FEDERAL
D.F') then 2
      when Estado is null then 3
      WHEN Estado in ('ESTADO DE MEXICO','ESTADO DE MEXICO
MEX') THEN 3 else 1 END AS EDO
,CASE WHEN Nivel_Estudios IN
('DIPLOMADOS','Licenciatura','UNIVERSIDAD''POSTGRADO
(MAESTRIA)','Postgrado') THEN 3
      WHEN Nivel_Estudios IN ('PREPARATORIA','Técnico o
Preparatoria') THEN 2
      WHEN Nivel_Estudios IN
('PRIMARIA','Primaria','SECUNDARIA','Secundaria') THEN 1 ELSE 4 END
AS ESTUDIOS
,CASE WHEN AUTOMOVIL IN ('SI') THEN 2
      WHEN AUTOMOVIL IS NULL THEN 1 ELSE 2 END AS AUTOMOVIL_
,case when edad>=1 AND edad<=22 THEN 3
      WHEN edad>=23 AND edad<=36 THEN 2
      WHEN edad>=37 THEN 1 ELSE 3 END AS EDADES
,CASE WHEN NoDependientes=0 THEN 1
      when nodependientes is null then 1
      WHEN NoDependientes=1 THEN 2
      WHEN NoDependientes=2 then 3
      WHEN NoDependientes>2 THEN 4 END AS NO_DEPENDIENTES
,CASE WHEN SCORE_BUENO<=550 THEN 6
      WHEN SCORE_BUENO> 550 AND SCORE_BUENO< 570 THEN 5
      WHEN SCORE_BUENO>= 570 AND SCORE_BUENO< 590 THEN 4
      WHEN SCORE_BUENO>= 590 AND SCORE_BUENO< 610 THEN 3
      WHEN SCORE_BUENO>= 610 AND SCORE_BUENO< 630 THEN 2
      WHEN SCORE_BUENO> 630 THEN 1
ELSE 6 END AS RANG_SCORE
,CASE WHEN ATP<=50 THEN 2
      WHEN ATP>50 THEN 1
      ELSE 2 END AS ATP_
,CASE WHEN EDOCIV IN ('CASADA/O BIENES S','CASADO/A BIENES
M','Casado','VIUDO/A','Viudo') THEN 1
      WHEN EDOCIV IN ('DIVORCIADO/A','Divorciado',
'SOLTERO/A','Soltero') THEN 3
      WHEN EDOCIV IN ('SEPARADO/A','Separado','UNION
LIBRE','Unión Libre') THEN 2 ELSE 3 END AS ESTADO_CIVIL
,CASE WHEN TIPOVIVIENDA IN
('FAMILIAR','Familiares','RENTADA','Rentada','PAGANDO','Pagandola')
THEN 2
      WHEN TIPOVIVIENDA IN ('PROPIA','Propia') THEN 1
ELSE 2 END AS TIPO_VIVIENDA
,CASE WHEN BC_SCORE<=650 THEN 3
      WHEN BC_SCORE>650 AND BC_SCORE<=700 THEN 4
      WHEN BC_SCORE>700 AND BC_SCORE<=750 THEN 2
      WHEN BC_SCORE>750
      THEN 1 ELSE 4 END AS RANG_BCSCORE
,CASE WHEN FLAG_HIT_NOHIT IN ("HIT") THEN 2
      WHEN FLAG_HIT_NOHIT IN ("NOHIT") THEN 1 ELSE 2 END
AS BANDERA_HIT_NOHIT
,CASE WHEN CAPITAL<=5000 THEN 1
      WHEN CAPITAL>5000 THEN 2 ELSE 2 END AS MONTO
,CASE WHEN PLAZO<=50 THEN 2
      WHEN PLAZO>50 THEN 1
      ELSE 2 END AS T_PLAZO
,CASE WHEN Tipo_Cr_dito IN ('CONSUMO') THEN 2
      WHEN Tipo_Cr_dito IN ('MICROCRED','MICRONEGOCIO')
THEN 1 ELSE 3 END AS TIPO_CREDITO

```

```

,CASE WHEN Producto IN
('PREFERENCI','Prefere','Preferencial') THEN 2
      WHEN Producto IN ('REGULAR','Regular') THEN 1 ELSE
3 END AS TIPO_PRODUCTO
      ,CASE WHEN Micro_cr_dito_Ubicaci_n IN('DE CASA EN
CASA','CASA','CASA FAMILIAR','DESDE CASA','DESDE SU CASA
DOMICILIO','EN CASA','EN SU CASA','EN EL HOGAR','EN EL LUGAR
DONDE RE','PATIO CASA','VENTA DESDE SU CASA',' VENTA EN SU
CASA','HOGAR',' CASA FAMILIAR','FAMILIAR' 'TRABA EN EL HOGAR')
THEN 2
      WHEN Micro_cr_dito_Ubicaci_n LIKE ('CASA') THEN 2
      WHEN Micro_cr_dito_Ubicaci_n IS NULL THEN 1
      WHEN Micro_cr_dito_Ubicaci_n IN('LOCAL') THEN 4 ELSE
3 END AS ubicacion_microcred
      ,CASE WHEN Micro_cr_dito_Tipo IN ("PROPIO","CASA","EN SU
CASA","EN SU MISMO DOMICILI","FAMILIAR","FAMILIAR UBICADO EN
HOGAR") THEN 2
      WHEN Micro_cr_dito_Tipo IS NULL THEN 1 ELSE 3 END AS
tipo_microcred
      ,CASE WHEN Micro_cr_dito_Lugar IN
('FIJO','LOCAL','SEMIFIJO','PUESTO') THEN 1
      WHEN Micro_cr_dito_Lugar IN ('MOVIL','OTROS') THEN 2
ELSE 3 END AS LUGAR_MICROCREDITO
      ,CASE WHEN TipoSolicitud IN ('FORMER') THEN 1
      WHEN TipoSolicitud IN('NORMAL','NUEVO') THEN 2
      WHEN TipoSolicitud IN('REESTR','REESTRUCTURA',) THEN
4
      WHEN TipoSolicitud IN('RENOVA','RENOVADA','RENOVADO')
THEN 3 ELSE 5 END AS TIPO_SOLICITUD
      ,CASE WHEN Periodicidad in ('MENSUAL','Mensual') then 4
      WHEN Periodicidad in ('QUINCENAL','Quincenal') THEN 3
      WHEN Periodicidad in ('SEMANAL','Semanal') THEN 1 ELSE 2
END AS TIPO_PERIODICIDAD
      ,CASE WHEN IngresoTotal<=10000 THEN 2
      WHEN IngresoTotal>10000
      THEN 1 ELSE 2 END AS INGRESO_TOTAL
      ,CASE WHEN Gastos_Bur__Cr_dito>=0 AND Gastos_Bur__Cr_dito<=500
THEN 1
      WHEN Gastos_Bur__Cr_dito>500
      THEN 2 ELSE 2 END AS GTOS_BURO_CREDITO
      ,CASE WHEN Pago_mensual_pr_stamo<=300 THEN 4
      WHEN Pago_mensual_pr_stamo>300 AND
Pago_mensual_pr_stamo<=1000 THEN 2
      WHEN Pago_mensual_pr_stamo>1000 AND
Pago_mensual_pr_stamo<=3000 THEN 3
      WHEN Pago_mensual_pr_stamo>3000
      THEN 1 ELSE 4 END AS PAGO_MENSUAL
      ,CASE WHEN GastosTotales<=500 THEN 1
      WHEN GastosTotales>500 AND GastosTotales<=899 THEN 2
      WHEN GastosTotales>899
      THEN 3 ELSE 3 END AS GTOS_TOT
      ,CASE WHEN Como_se_Enterо IS NULL THEN 3
      WHEN Como_se_Enterо IN ("ANUNCIO AZOTEA", "ANUNCIO
CALLE/MANTA", "APASCO", "CAMBACEO",
"COMISIONISTA", "INTERNET", "LETRERO", "LLAMADA TELEFONICA",
"MARQUESINA", "MODULO", "PASANDO POR
SUCURSAL", "PERIODICO", "POSTER / MICRONEG", "POSTER
MICRONEGOCIO", "REVISTA", "VOCEO", "VOLANTE",
"NINGUNO", "AMISTAD / CONOCIDO",
"RECOMENDACION", "FAMILIAR") THEN 1
      WHEN Como_se_Enterо IN ("RENOVACION", "REESTRUCTURA",
"FORMER") THEN 2 ELSE 1 END AS CENTERO
      ,CASE WHEN TipoComprobanteIngresos IN ("CTA HABERES", "Carta
de haberes", "NOTA COMPRA","Notas de compra", 'Carta laboral a mano',
'Carta laboral membretada') then 2
      WHEN TipoComprobanteIngresos IN ("EDO DE CTA", "Estado

```

```

de Cuenta", "Estado de cuenta"      ,"Estados de
Cuenta"      ,"FACTURAS",      "Facturas") then 4
      WHEN TipoComprobanteIngresos IN ("REC PENSION"
"Recibos de pensión"      "NOMINA"      "Nómina"      "DECLARACION"
"Declaración de Impuestos") then 3
      when TipoComprobanteIngresos is null then 1 ELSE 1 END
AS TIP_CPTE_INGR
      ,CASE WHEN FrecuenciaPrecepcionIngresos IN ("DIARIO",
"Diario") then 5
      WHEN FrecuenciaPrecepcionIngresos IN ("MENSUAL",
"Mensual") then 2
      WHEN FrecuenciaPrecepcionIngresos IN ("QUINCENAL",
"Quincenal", " VARIABLE", "Variable") then 6
      WHEN FrecuenciaPrecepcionIngresos IN ("SEMANAL",
"Semanal") then 3
      WHEN FrecuenciaPrecepcionIngresos is null then 4
ELSE 1 END AS FREC_PREC_INGR
      ,CASE WHEN WOB>=0 AND WOB<=5 THEN 1
      WHEN WOB>5 AND WOB <=10 THEN 2
      WHEN WOB>10 AND WOB <=20 THEN 3
      WHEN WOB>20 AND WOB <=40 THEN 4
      WHEN WOB>40 THEN 1 ELSE 1 END AS WOB_
      ,CASE WHEN SISTEMA='ANFEXI' THEN 2
      WHEN SISTEMA='FINAY' THEN 1 ELSE 2 END AS NAC_SIST
      ,CASE WHEN TELS<=4 THEN 2
      WHEN TELS>4 THEN 1 ELSE 2 END AS NUM_TELS
      , CASE WHEN MAX_DM = 0 THEN 1
      WHEN MAX_DM >= 1 AND MAX_DM <= 2 THEN 2
      WHEN MAX_DM >= 3 AND MAX_DM <= 6 THEN 3
      WHEN MAX_DM >= 7 THEN 4
      ELSE 4 END AS MAXIMO_DM
      ,CASE WHEN MAX_RV=0 then 1
      WHEN MAX_RV>=1 AND MAX_RV<=4 THEN 2
      WHEN MAX_RV>5 THEN 3 ELSE 3 END AS MAXIMO_RV

      , case when avg_dm = 0 then 1
      WHEN avg_dm > 0 AND avg_dm <= 0.318 THEN 2
      WHEN avg_dm > 0.318 THEN 3
      ELSE 3 END AS PROMEDIO_DM

      ,CASE WHEN pcappag>=0 AND pcappag<=10 THEN 7
      WHEN pcappag>10 AND pcappag<=30 THEN 6
      WHEN pcappag>30 AND pcappag<=50 THEN 5
      WHEN pcappag>50 AND pcappag<=70 THEN 4
      WHEN pcappag>70 and pcappag<=80 then 2
      WHEN pcappag>80 and pcappag<=90 then 3
      when pcappag>90 then 1 else 7 end as pct_cappag
      ,CASE WHEN ctas_pagadas_final>=0 AND ctas_pagadas_final<=5
then 3
      WHEN ctas_pagadas_final>5 AND ctas_pagadas_final<=10
THEN 4
      WHEN ctas_pagadas_final>10 AND ctas_pagadas_final<=
15 THEN 6
      WHEN ctas_pagadas_final>15 AND ctas_pagadas_final<=
20 THEN 9
      WHEN ctas_pagadas_final>20 AND ctas_pagadas_final<=
40 THEN 8
      WHEN ctas_pagadas_final>40 AND ctas_pagadas_final<=
60 THEN 7
      WHEN ctas_pagadas_final>60 AND ctas_pagadas_final<=
80 THEN 5
      WHEN ctas_pagadas_final>80 AND ctas_pagadas_final<=
100 THEN 2
      WHEN ctas_pagadas_final>100 THEN 1 ELSE 9 END AS
CTASPAGADAS

```

```

,CASE WHEN NUM_CONDONACIONES=0 THEN 1
      WHEN NUM_CONDONACIONES>1 THEN 2
      ELSE 2 END AS CONDONACIONES
, CASE WHEN Impagos= 0 then 1

      when Impagos=1 then 2
      WHEN Impagos>=2 AND Impagos<=3 THEN 3
      WHEN Impagos>3 AND Impagos<=4 THEN 4
      WHEN Impagos>4 AND Impagos<=5 THEN 5
      WHEN Impagos>5 AND Impagos<=10 THEN 6
      WHEN Impagos>10 AND Impagos<=15 THEN 7
      WHEN Impagos>15

      THEN 8 ELSE 8 END AS NUM_IMPAGOS

, CASE when adelantos=0 then 10
      when adelantos=1 then 9
      WHEN Adelantos>1 AND Adelantos<=2 THEN 8
      WHEN Adelantos>2 AND Adelantos<=3 THEN 7
      WHEN Adelantos>3 AND Adelantos<=4 THEN 6
      WHEN Adelantos>4 AND Adelantos<=5 THEN 5
      WHEN Adelantos>5 AND Adelantos<=10 THEN 4
      WHEN Adelantos>10 AND Adelantos<=15 THEN 3
      WHEN Adelantos>15 AND Adelantos<=20 THEN 2
      WHEN Adelantos>20
      then 1 else 10 END AS NUM_ADELANTOS

, CASE WHEN V1RV=0 THEN 1
      WHEN V1RV>1
      THEN 3 ELSE 2 END AS V_R1

, CASE WHEN V2RV=0 THEN 1
      when v2rv>1 then 3
      ELSE 2 END AS V_R2

, CASE WHEN V3RV=0 THEN 1
      when v3rv>1 THEN 2 ELSE 3 END AS V_R3

, CASE WHEN V4RV=0 THEN 1
      when v4rv>1 THEN 3 ELSE 2 END AS V_R4

, CASE WHEN V5RV=0 then 1
      when v5rv>1 THEN 3 ELSE 2 END AS V_R5

, CASE WHEN NUM_RENOVACIONES=0 THEN 1
      WHEN NUM_RENOVACIONES>0 THEN 2 ELSE 2 END AS
RENOVACIONES

FROM RESUMEN ;quit;

libname insumo 'D:\Bases\Insumo' ;

data insumo.Insumo_2da_tabla ; set Insumo_2da_tabla; run;

/*
                                Análisis del estimador de máxima
verosimilitud

                                Error
Chi-cuadrado                                Pr > ChiSq
de Wald      Parámetro                      DF      Estimador      estándar
149.2820     Intercept                       1        -3.0833      0.2524
              <.0001

```

3.5585	SEXO_ 0.0592	1	0.0532	0.0282
2.1303	ESTUDIOS 0.1444	1	-0.0411	0.0282
12.9354	EDADES 0.0003	1	0.0802	0.0223
5.4823	NO_DEPENDIENTES 0.0192	1	0.0438	0.0187
20.9648	RANG_BCSCORE <.0001	1	0.0830	0.0181
10.6936	T_PLAZO 0.0011	1	0.1625	0.0497
12.6066	TIPO_CREDITO 0.0004	1	-0.2405	0.0677
3.2848	LUGAR_MICROCREDITO 0.0699	1	0.0670	0.0370
12.0460	TIPO_SOLICITUD 0.0005	1	-0.0869	0.0250

1

1

0:57 Saturday, August 29, 2011 18

Procedimiento LOGISTIC

Análisis del estimador de máxima

verosimilitud

Chi-cuadrado de Wald	Parámetro Pr > ChiSq	DF	Estimador	Error estándar
60.9894	TIPO_PERIODICIDAD <.0001	1	0.2168	0.0278
4.5946	INGRESO_TOTAL 0.0321	1	0.0748	0.0349
2.5183	PAGO_MENSUAL 0.1125	1	0.0372	0.0234
17.8863	WOB_ <.0001	1	-0.0597	0.0141
5.4978	NAC_SIST 0.0190	1	-0.1770	0.0755
60.3892	MAXIMO_DM <.0001	1	0.2610	0.0336
75.7516	PROMEDIO_DM <.0001	1	0.4183	0.0481
141.1125	pct_cappag <.0001	1	-0.1268	0.0107
534.9414	CTASPAGADAS <.0001	1	-0.2195	0.00949
3.5299	CONDONACIONES 0.0603	1	0.2311	0.1230
2164.6925	NUM_IMPAGOS <.0001	1	0.5560	0.0120
13.5630	V_R1 0.0002	1	-0.0877	0.0238
40.0358	V_R4 <.0001	1	-0.2823	0.0446
59.2926	RENOVACIONES <.0001	1	-0.2840	0.0369

\*/

```
PROC SQL; CREATE TABLE INS_PREVENT_SCORE AS
SELECT z.*
      , 700+(80/LOG(2))*(-3.0833+
                    0.0532*SEXO_
```

```

-0.0411*ESTUDIOS
+0.0802* EDADES
+0.0438*NO_DEPENDIENTES
+0.0830*RANG_BCSCORE
+0.1625*T_PLAZO
-0.2405*TIPO_CREDITO
+0.0670*LUGAR_MICROCREDITO
-0.0869 *TIPO_SOLICITUD
+0.2168*TIPO_PERIODICIDAD
+0.0748*INGRESO_TOTAL
+0.0372*PAGO_MENSUAL
-0.0597 * WOE_
-0.1770*NAC_SIST
+0.2610*MAXIMO_DM
+0.4183*PROMEDIO_DM
-0.1268*pct_cappag
-0.2195* CTASPAGADAS
+0.2311* CONDONACIONES
+0.5560* NUM_IMPAGOS
-0.0877 * V_R1
-0.2823*V_R4
-0.2840*RENOVACIONES
)
AS PSCORE

FROM INSUMO.INSUMO_2DA_TABLA AS z
;QUIT;

PROC SQL; CREATE TABLE VALIDA00 AS SELECT MAX(PSCORE), MIN(PSCORE)
FROM INS_PREVENT_SCORE ; QUIT;

PROC SQL; CREATE TABLE RESUMEN_INS_PREVENT_SCORE AS
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS

FROM (
SELECT
CONTRATO
/*, CASE WHEN TIPO_PERIODICIDAD in (2,1) THEN 'SEMANAL'
WHEN tipo_PERIODICIDAD=4 THEN 'MENSUAL'
WHEN TIPO_PERIODICIDAD=3 THEN 'QUINCENAL' END AS PERIODO*/
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'
WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'
WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'
WHEN PSCORE > 750 THEN '750-MAS'
END AS P_SCORE
, BANDERA
FROM INS_PREVENT_SCORE
)Z
GROUP BY 1
; QUIT;

/*PROC SQL; CREATE TABLE RESUMEN_INS_PREVENT_SCORE AS
SELECT PERIODO, P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA)
AS BADS

FROM (

```



```

SELECT
CONTRATO
, CASE WHEN TIPO_PERIODICIDAD in (2,1) THEN 'SEMANAL'
      WHEN tipo_PERIODICIDAD=4 THEN 'MENSUAL'
      WHEN TIPO_PERIODICIDAD=3 THEN 'QUINCENAL' END AS PERIODO
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
      WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
      WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
      WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
      WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
      WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
      WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
      WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
      WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
      WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'
      WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'
      WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'
      WHEN PSCORE > 750 THEN '750-MAS'
      END AS P_SCORE
, BANDERA
FROM INS_PREVENT_SCORE
)Z
GROUP BY 1, 2
;      QUIT;*/

PROC SQL; CREATE TABLE SEMANAL AS
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS
from(select *
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
      WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
      WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
      WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
      WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
      WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
      WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
      WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
      WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
      WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'
      WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'
      WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'
      WHEN PSCORE > 750 THEN '750-MAS'
      END AS P_SCORE
      FROM INS_PREVENT_SCORE

)Z

WHERE TIPO_PERIODICIDAD in (2,1)

group by 1
; QUIT;

```

```

PROC SQL; CREATE TABLE MENSUAL AS
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS
from(select *
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
      WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
      WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
      WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
      WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
      WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
      WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
      WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
      WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
      WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'

```

```
    WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'  
    WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'  
    WHEN PSCORE > 750 THEN '750-MAS'  
    END AS P_SCORE  
    FROM INS_PREVENT_SCORE
```

```
)Z
```

```
WHERE tipo_PERIODICIDAD=4  
group by 1  
; QUIT;
```

```
PROC SQL; CREATE TABLE QUINCENAL AS  
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS  
from(select *  
    , CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'  
    WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'  
    WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'  
    WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'  
    WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'  
    WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'  
    WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'  
    WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'  
    WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'  
    WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'  
    WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'  
    WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'  
    WHEN PSCORE > 750 THEN '750-MAS'  
    END AS P_SCORE  
    FROM INS_PREVENT_SCORE
```

```
)Z
```

```
WHERE TIPO_PERIODICIDAD=3  
group by 1  
; QUIT;
```

```
/**VALIDACIÓN DE LA PREVENTIVA CON UN MES POSTERIOR A junio  
Gabriela Manzanera***/
```

```
LIBNAME DIARIOS "D:\Bases\DIARIOS";  
LIBNAME ORIGINA "D:\Bases\Originacion";
```

```
/**Importar exigibilidades de junio***/
```

```
data DIARIOS.EXIGIBILIDADES_junio2011 ;  
  %let _EFIERR_ = 0; /* set the ERROR detection macro variable */  
  infile 'D:\Bases\DIARIOS\Exigibilidades_junio2011.txt'  
delimitter='09'x MISSOVER DSD  
  lrecl=32767 firstobs=2 ;  
  informat N__de_Contrato $13. ;  
  informat Fecha_De_Vto_ date9. ;  
  format N__de_Contrato $13. ;  
  format Fecha_De_Vto_ date9. ;  
input  
      N__de_Contrato $  
      Fecha_De_Vto_  
;  
  if _ERROR_ then call symput('_EFIERR_',1); /* set ERROR  
detection macro variable */  
run;
```

```
/**Conciliar Exigibilidades de junio 2011***/
```

```
PROC SQL; CREATE TABLE DIARIOS.Exigibilidades_junio2011 AS  
SELECT A.*  
      , CASE WHEN C.Acct_num IS NULL THEN A.N__de_Contrato ELSE  
C.Acct_num_nuevo END AS CONTRATO  
FROM DIARIOS.Exigibilidades_junio2011 AS A  
LEFT JOIN DIARIOS.CONCILIACIONb AS C  
ON A.N__de_Contrato = C.Acct_num  
; QUIT;
```

```
/**Colocar la bandera a las Exigibilidades***/
```

```
PROC SQL; CREATE TABLE DIARIOS.Exigibilidades_junio2011 AS  
SELECT E.Contrato  
      ,E.Fecha_De_Vto_  
      ,R.DiasMoroso  
from DIARIOS.Exigibilidades_junio2011 as E  
LEFT JOIN DIARIOS.TBL_REPCARTERA_V8 AS R  
ON E.Contrato=R.CONTRATO  
WHERE E.Fecha_De_Vto_= R.FecSeguimiento  
AND R.DiasMoroso=0  
;quit;  
data DIARIOS.Exigibilidades_junio2011;set  
DIARIOS.Exigibilidades_junio2011;rename  
DiasMoroso=Morosidad_Exigibilidad;run;
```

```
PROC SQL; CREATE TABLE DIARIOS.Exigibilidades_junio2011 AS  
SELECT C.*  
      ,R.DiasMoroso  
      ,Case when R.DiasMoroso>0 then "BAD" ELSE "GOOD" END AS  
FLAG_GOOD_BAD  
from DIARIOS.Exigibilidades_junio2011 as C
```

```

LEFT JOIN DIARIOS.TBL_REPCARTERA_V8 AS R
ON C.Contrato=R.CONTRATO
where C.Fecha_De_Vto_+2=R.FecSeguimiento
;quit;

```

```

/****LEFT JOIN CON ORIGINACION****/

```

```

PROC SQL; CREATE TABLE RESUMEN AS
SELECT p.Contrato
      , p.diasmoroso
      , p.FLAG_GOOD_BAD
      , P.Fecha_De_Vto_
      , O.Fecha_Nacimiento
      , O.Fecha_Alta
      , O.Sexo
      , O.Estado
      , O.Nivel_Estudios
      , O.Automovil
      , O.Contrato_laboral
      , FLOOR( (P.Fecha_De_Vto_ - O.Fecha_Nacimiento)/365 ) AS EDAD
      , O.NoDependientes
      , O.AntigDom
      , O.AntigTrabajo
      , O.SCORE_BUENO
      , O.ATP
      , O.EdoCiv
      , O.TipoVivienda
      , O.Sucursal
      , O.CP
      , O.BC_Score
      , CASE WHEN O.BC_SCORE_FINAL=-9 THEN'NO_HIT' ELSE 'HIT' END
AS FLAG_HIT_NOHIT
      , O.Capital
      , O.Plazo
      , O.Tipo_Cr_dito
      , O.Producto
      , O.Micro_cr_dito_Ubicaci_n
      , O.Micro_cr_dito_Tipo
      , O.Micro_cr_dito_Lugar
      , O.Horario_laboral
      , O.TipoSolicitud
      , O.Periodicidad
      , FLOOR(O.Ingreso_Fijo + O.Ingreso_variable) AS IngresoTotal
      , O.Gastos_Bur__Cr_dito
      , O.Pago_mensual_pr_stamo
      , FLOOR(O.Renta + O.Servicios + O.Otros_Gastos) AS
GastosTotales
      , O.Fuentes_Ingreso_variables
      , O.Como_se_Entero
      ,Case when O.Tel_Casa is not null then 1 else 0 end as Tel_1
      ,Case when O.Tel_Movil is not null then 1 else 0 end as Tel_
2
      ,CASE WHEN O.Tel_Casa is not null then 1 else 0 end as Tel_3
      ,CASE WHEN O.Tel_Movil is not null then 1 else 0 end as Tel_
4
      ,CASE WHEN O.Tel_Arrendador is not null then 1 else 0 end as
Tel_5
      ,CASE WHEN O.Tel1_Empleo is not null then 1 else 0 end as
Tel_6
      ,CASE WHEN O.Tel2_Empleo is not null then 1 else 0 end as
Tel_7
      ,CASE WHEN O.Telefono_Cosolicitante is not null then 1 else
0 end as Tel_8

```

```

,CASE WHEN O.Tel_Ref_Fam1 is not null then 1 else 0 end as
Tel_9
,CASE WHEN O.Tel_Ref_Fam2 is not null then 1 else 0 end as
Tel_10
,CASE WHEN O.Tel_Ref_Psn11 is not null then 1 else 0 end
as Tel_11
,CASE WHEN O.Tel_Ref_Psn12 is not null then 1 else 0 end
as Tel_12
, O.TipoComprobanteIngresos
, O.FrecuenciaPrecepcionIngresos
, FLOOR((P.Fecha_De_Vto_ - O.Fecha_alta)/7) as wob
, CASE WHEN C.acct_num_nuevo IS NOT NULL THEN 'ANFEXI'
WHEN LENGTH(P.CONTRATO) = 9 THEN 'ANFEXI' ELSE
'FINAY' END AS SISTEMA

```

```

FROM DIARIOS.Exigibilidades_junio2011 AS P
LEFT JOIN ORIGINA.Originacion18a AS O
ON p.Contrato = O.CONTRATO
LEFT JOIN DIARIOS.CONCILIACIONB AS C
ON P.CONTRATO = C.acct_num_nuevo

```

```
; QUIT;
```

```

PROC SQL; CREATE TABLE RESUMEN AS
SELECT O.*
, FLOOR(O.Tel_1+ O.Tel_2+ O.Tel_3+ O.Tel_4 + O.Tel_5+ O.Tel_6
+O.Tel_7+O.Tel_8+O.Tel_9+O.Tel_10+O.Tel_11+O.Tel_12) as TELS
FROM RESUMEN AS O
;QUIT;

```

```
/**PARA INCLUIR LA VARIABLE SISTEMA_DE_NACIMIENTO**/
```

```
/*VARIABLES COMPORTAMIENTO*/
/*
```

```

1.MAXIMO DE DÍAS DE MOROSIDAD
2.PROMEDIO DE DÍAS QUE EL CLIENTE SE TARDA EN PAGAR LA
EXIGIBILIDAD
3.PORCENTAJE DE CAPITAL YA PAGADO
4.IMPAGOS ULTIMOS N MESES (N=5)
5.MAX NÚM DE RECIBOS VENCIDOS
6.NUMERO DE VECES EN 1 RV , 2 RV, 3 RV, ETC
7.CUOTAS YA PAGADAS
8.NÚMERO DE CONDONACIONES
9.NÚMERO DE RENOVACIONES
*/

```

```
/* 1.MAXIMO DE DÍAS DE MOROSIDAD*/
```

```
PROC SQL; CREATE TABLE Maxdiasmoroso AS /**MÁXIMO DÍAS QUE TARDA EN
PAGAR EL CLIENTE**/ /***/
```

```

SELECT T.CONTRATO
, E.Fecha_De_Vto_
, MAX(T.DIASMOROSO) AS MAX_DM
, MAX(recibosvencidos) AS MAX_RV
, AVG(T.DIASMOROSO) AS AVG_DM

```

```

FROM DIARIOS.TBL_REPCARTERA_V8 AS T
INNER JOIN DIARIOS.EXIGIBILIDADES_junio2011 AS E
ON T.CONTRATO = E.Contrato
AND T.FECSEGUIMIENTO <= E.Fecha_De_Vto_
GROUP BY T.CONTRATO, E.Fecha_De_Vto_ ;QUIT;

```

```

proc sql; create table Resumen as
Select R.*
      , M.MAX_DM
      , M.MAX_RV
      , M.AVG_DM
from Resumen as R
inner join Maxdiasmoroso as M
      on R.Contrato=M.CONTRATO
      and R.Fecha_De_Vto_= M.Fecha_De_Vto_
;quit;

/*3.PORCENTAJE DE CAPITAL YA PAGADO */

/*****

proc sql; create table Resumen as
select R.*

((T.CapitalVigente+T.CapitalTramite+T.capvencexigible+T.capvencnoexig
ible)/T.Total)*100 as pcappag

FROM Resumen AS R
      left join DIARIOS.TBL_REPCARTERA_V8 as T
      on R.Contrato=T.Contrato
      AND T.FECSEGUIMIENTO = R.Fecha_De_Vto_
      ;QUIT;

/*4.IMPAGOS ULTIMOS N MESES*/

/**Número de Impagos los últimos n meses***/
/** Con n=5**/
PROC IMPORT OUT= WORK.pagosa
      DATAFILE= "D:\Bases\DIARIOS\vencimiento.txt"
      DBMS=TAB REPLACE;
      GETNAMES=YES;
      DATAROW=2;
RUN;

PROC SQL; CREATE TABLE Pagosa AS
      select P.*
      , case when c.Acct_num_nuevo is null then P.n_de_contrato
      else c.Acct_num_nuevo end as N__de_Contrato
      from pagosa as p
      left join DIARIOS.conciliaciona as c
      on P.n_de_contrato= c.Acct_num
; quit;
data pagosa; set pagosa; rename fecha_de_pago=fechapago
fechaVencimiento=Fecha_Vencimiento_Cubierto
Forma_de_Pago_Cliente=Forma_de_Pago;run;
proc sql; create table pagosa as
select P.N__de_Contrato
      ,P.fechapago
      ,P.Fecha_Vencimiento_Cubierto
      ,P.Forma_de_Pago
      from pagosa as p
      ;quit;

```

```

PROC IMPORT OUT= WORK.pagosf
          DATAFILE= "D:\Bases\DIARIOS\Pagos FINAYv2.txt"
          DBMS=TAB REPLACE;
          GETNAMES=YES;
          DATAROW=2;
RUN;
proc sql; create table pagosf as
select P.N__de_Contrato
      ,P.fechapago
      ,P.Fecha_Vencimiento_Cubierto
      ,P.Forma_de_Pago
      from pagosf as p
      ;quit;

PROC SQL; CREATE TABLE Pagos AS
SELECT * FROM pagosa UNION ALL
SELECT * FROM pagosf ; QUIT;

proc sql; create table pagos1 as
select N__de_contrato, Fecha_De_Vto_ ,count(IMPAGOS) AS NUM_IMPAGOS
      FROM
      (
select t.*
      , R.Fecha_De_Vto_
      , case when t.Fechapago>t.Fecha_Vencimiento_Cubierto then
'Impago'
              when t.Fechapago=t.Fecha_Vencimiento_Cubierto then "pago"
              when t.fechapago<t.Fecha_Vencimiento_Cubierto then
"adelanto" else "nuncapago" end as impagos
      from pagos as t
      INNER JOIN RESUMEN AS R
              ON T.N__de_Contrato = R.CONTRATO
              AND t.Fecha_Vencimiento_Cubierto <= R.Fecha_De_Vto_
      )z
      WHERE Z.IMPAGOS='Impago'
              AND Z.Fechapago >="01FEB11"D
              AND Z.Fechapago <="30JUN11"d

GROUP BY N__de_contrato,Fecha_De_Vto_
;quit;

PROC SQL; CREATE TABLE Resumen AS
SELECT r.*
      ,case when P.num_IMPAGOS is null then 0 else P.num_IMPAGOS end
      as Impagos
      FROM RESUMEN AS R
      LEFT JOIN PAGOS1 AS P
      ON R.CONTRATO=p.N__de_Contrato
      and r.Fecha_De_Vto_=p.Fecha_De_Vto_
      ;quit;

proc sql; create table adelanto as
select N__de_contrato, Fecha_De_Vto_ ,count(IMPAGOS) AS
num_adelantos
      FROM
      (
select t.*
      , R.Fecha_De_Vto_
      , case when t.Fechapago>t.Fecha_Vencimiento_Cubierto then
'Impago'
              when t.Fechapago=t.Fecha_Vencimiento_Cubierto then "pago"
              when t.fechapago<t.Fecha_Vencimiento_Cubierto then
"adelanto" else "nuncapago" end as impagos

```

```

from pagos as t
INNER JOIN RESUMEN AS R
      ON T.N__de_Contrato = R.CONTRATO
      AND t.Fecha_Vencimiento_Cubierto <= R.Fecha_De_Vto_
)z
WHERE Z.IMPAGOS="adelanto"
      AND Z.Fechapago >="01FEB11"D
      AND Z.Fechapago <="30JUN11"d

GROUP BY N__de_contrato, Fecha_De_Vto_
;quit;

PROC SQL; CREATE TABLE Resumen AS
SELECT r.*
      , case when P.num_adelantos is null then 0 else P.num_adelantos
end as adelantos
FROM Resumen AS R
LEFT JOIN adelanto AS P
ON R.CONTRATO=p.N__de_Contrato
and r.Fecha_De_Vto_=p.Fecha_De_Vto_
;quit;

proc sql; create table coutas_pagadas as
select p.N__de_contrato, r.Fecha_De_Vto_, count
(p.Fecha_Vencimiento_Cubierto) as ctaspagadas
from pagos as p
inner join resumen as r
      on p.N__de_contrato = r.contrato
      and p.fechapago <= r.Fecha_De_Vto_
group by 1, 2 ; quit;

PROC SQL; CREATE TABLE Resumen as
SELECT R.*
      ,CASE WHEN C.ctaspagadas IS NULL THEN 0 ELSE C.ctaspagadas END
AS ctas_pagadas_final
FROM RESUMEN AS R
LEFT JOIN COUTAS_PAGADAS AS C
ON R.contrato=c.n__de_contrato
and r.Fecha_De_Vto_=c.Fecha_De_Vto_

;QUIT;

PROC SQL; CREATE TABLE VECES_WK AS
SELECT CONTRATO
      , SUM(V1) AS V_1RV
      , SUM(V2) AS V_2RV
      , SUM(V3) AS V_3RV
      , SUM(V4) AS V_4RV
      , SUM(V5) AS V_5RV
      , SUM(V6) AS V_6RV
      , SUM(V7) AS V_7RV
      , SUM(V8) AS V_8RV
FROM
(
SELECT FECSEGUIMIENTO
      , CONTRATO
      , CASE WHEN recibosvencidos = 1 THEN 1 ELSE 0 END AS V1

```



```

        , CASE WHEN recibosvencidos = 2 THEN 1 ELSE 0 END AS V2
        , CASE WHEN recibosvencidos = 3 THEN 1 ELSE 0 END AS V3
        , CASE WHEN recibosvencidos = 4 THEN 1 ELSE 0 END AS V4
        , CASE WHEN recibosvencidos = 5 THEN 1 ELSE 0 END AS V5
        , CASE WHEN recibosvencidos = 6 THEN 1 ELSE 0 END AS V6
        , CASE WHEN recibosvencidos = 7 THEN 1 ELSE 0 END AS V7
        , CASE WHEN recibosvencidos = 8 THEN 1 ELSE 0 END AS V8
        , CASE WHEN recibosvencidos = 9 THEN 1 ELSE 0 END AS V9
        , CASE WHEN recibosvencidos >= 10 THEN 1 ELSE 0 END AS V10
_MAS
    FROM DIARIOS.TBL_REPCARTERA_V8
    WHERE FECSEGUIMIENTO <"01JUN11."D
    AND weekday(FECSEGUIMIENTO) = 2
    AND PERIODICIDAD IN ('SEMANAL', 'Semanal')
    AND diasmoroso <= 30

```

```
) Z GROUP BY CONTRATO; QUIT;
```

```

PROC SQL;
CREATE TABLE DIARIOS.DIAS_QUINCENA AS

SELECT FECSEGUIMIENTO
    FROM
    (
    SELECT ANO
        , MES
        , MIN(DIF_1_D) AS MIN_DIF_1_D
        , MIN(DIF_16_D) AS MIN_DIF_16_D
        , FECSEGUIMIENTO
    FROM
    (
    SELECT FECSEGUIMIENTO
        , YEAR(FECSEGUIMIENTO) AS ANO
        , MONTH(FECSEGUIMIENTO) AS MES
        , DAY(FECSEGUIMIENTO) AS DIA
        , DAY(FECSEGUIMIENTO)- 1 AS DIF_1_D
        , CASE WHEN DAY(FECSEGUIMIENTO) < 16 THEN 99 ELSE ABS(DAY
(FECSEGUIMIENTO)- 16) END AS DIF_16_D
    FROM
    (
    SELECT FECSEGUIMIENTO, COUNT(CONTRATO) AS CLIENTES
        FROM DIARIOS.TBL_REPCARTERA_V8 /*OJO CON LA VERSION */
        GROUP BY FECSEGUIMIENTO
    )Z
    )Y
    GROUP BY ANO, MES

    )X
WHERE ( DAY(FECSEGUIMIENTO) = MIN_DIF_1_D+1 OR DAY(FECSEGUIMIENTO) =
MIN_DIF_16_D+16 )
ORDER BY FECSEGUIMIENTO
; QUIT;

```

```

PROC SQL; CREATE TABLE VECES_QN AS
SELECT CONTRATO
    , SUM(V1) AS V_1RV
    , SUM(V2) AS V_2RV
    , SUM(V3) AS V_3RV
    , SUM(V4) AS V_4RV
    , SUM(V5) AS V_5RV
    , SUM(V6) AS V_6RV

```

```

        , SUM(V7) AS V_7RV
        , SUM(V8) AS V_8RV
FROM
(
    SELECT FECSEGUIMIENTO
        , CONTRATO
        , CASE WHEN recibosvencidos = 1 THEN 1 ELSE 0 END AS V1
        , CASE WHEN recibosvencidos = 2 THEN 1 ELSE 0 END AS V2
        , CASE WHEN recibosvencidos = 3 THEN 1 ELSE 0 END AS V3
        , CASE WHEN recibosvencidos = 4 THEN 1 ELSE 0 END AS V4
        , CASE WHEN recibosvencidos = 5 THEN 1 ELSE 0 END AS V5
        , CASE WHEN recibosvencidos = 6 THEN 1 ELSE 0 END AS V6
        , CASE WHEN recibosvencidos = 7 THEN 1 ELSE 0 END AS V7
        , CASE WHEN recibosvencidos = 8 THEN 1 ELSE 0 END AS V8
        , CASE WHEN recibosvencidos = 9 THEN 1 ELSE 0 END AS V9
        , CASE WHEN recibosvencidos >= 10 THEN 1 ELSE 0 END AS V10
    _MAS
    FROM DIARIOS.TBL_REPCARTERA_V8
    WHERE FECSEGUIMIENTO <="01JUN11"D
    AND FECSEGUIMIENTO IN (SELECT * FROM DIARIOS.DIAS_QUINCENA)
    AND PERIODICIDAD IN ('QUINCENAL', 'QUINCENAL', 'Quincenal')
    AND diasmoroso <= 30

) Z GROUP BY CONTRATO; QUIT;

```

```

PROC SQL;
CREATE TABLE DIARIOS.DIAS_MES AS

SELECT FECSEGUIMIENTO
    FROM
(
    SELECT ANO
        , MES
        , MIN(DIF_1_D) AS MIN_DIF_1_D
        , MIN(DIF_16_D) AS MIN_DIF_16_D
        , FECSEGUIMIENTO
    FROM
(
    SELECT FECSEGUIMIENTO
        , YEAR(FECSEGUIMIENTO) AS ANO
        , MONTH(FECSEGUIMIENTO) AS MES
        , DAY(FECSEGUIMIENTO) AS DIA
        , DAY(FECSEGUIMIENTO)- 1 AS DIF_1_D
        , CASE WHEN DAY(FECSEGUIMIENTO) < 16 THEN 99 ELSE ABS(DAY
(FECSEGUIMIENTO)- 16) END AS DIF_16_D
    FROM
(
    SELECT FECSEGUIMIENTO, COUNT(CONTRATO) AS CLIENTES
        FROM DIARIOS.TBL_REPCARTERA_V8
        GROUP BY FECSEGUIMIENTO
    )Z
    )Y
    GROUP BY ANO, MES

)X
WHERE DAY(FECSEGUIMIENTO) = MIN_DIF_1_D+1
ORDER BY FECSEGUIMIENTO
; QUIT;

```

```

PROC SQL; CREATE TABLE VECES_MN AS
SELECT CONTRATO
    , SUM(V1) AS V_1RV

```

```

, SUM(V2) AS V_2RV
, SUM(V3) AS V_3RV
, SUM(V4) AS V_4RV
, SUM(V5) AS V_5RV
, SUM(V6) AS V_6RV
, SUM(V7) AS V_7RV
, SUM(V8) AS V_8RV
FROM
(
  SELECT FECSEGUIMIENTO
    , CONTRATO
    , CASE WHEN recibosvencidos = 1 THEN 1 ELSE 0 END AS V1
    , CASE WHEN recibosvencidos = 2 THEN 1 ELSE 0 END AS V2
    , CASE WHEN recibosvencidos = 3 THEN 1 ELSE 0 END AS V3
    , CASE WHEN recibosvencidos = 4 THEN 1 ELSE 0 END AS V4
    , CASE WHEN recibosvencidos = 5 THEN 1 ELSE 0 END AS V5
    , CASE WHEN recibosvencidos = 6 THEN 1 ELSE 0 END AS V6
    , CASE WHEN recibosvencidos = 7 THEN 1 ELSE 0 END AS V7
    , CASE WHEN recibosvencidos = 8 THEN 1 ELSE 0 END AS V8
    , CASE WHEN recibosvencidos = 9 THEN 1 ELSE 0 END AS V9
    , CASE WHEN recibosvencidos >= 10 THEN 1 ELSE 0 END AS V10
  _MAS
  FROM DIARIOS.TBL_REPCARTERA_V8 /*OJO CON LA VERSION */
  WHERE FECSEGUIMIENTO <="01JUN11"D
  AND FECSEGUIMIENTO IN (SELECT * FROM DIARIOS.DIAS_MES)
  AND PERIODICIDAD IN ('MENSUAL', 'Mensual')
  AND diasmoroso <= 30

) Z GROUP BY CONTRATO; QUIT;

PROC SQL; CREATE TABLE VECES AS
SELECT * FROM VECES_WK UNION ALL
SELECT * FROM VECES_QN UNION ALL
SELECT * FROM VECES_MN ; QUIT;

proc sql; create table resumen as
select R.*
,case when v.V_1RV is null then 0 else v.V_1RV end as V1RV
,case when v.V_2RV is null then 0 else v.V_2RV end as V2RV
,case when v.V_3RV is null then 0 else v.V_3RV end as V3RV
,case when v.V_4RV is null then 0 else v.V_4RV end as V4RV
,case when v.V_5RV is null then 0 else v.V_5RV end as V5RV
,case when v.V_6RV is null then 0 else v.V_6RV end as V6RV
,case when v.V_7RV is null then 0 else v.V_7RV end as V7RV
,case when v.V_8RV is null then 0 else v.V_8RV end as V8RV

  from resumen as r
  left join veces as v
  on r.Contrato=v.Contrato

;quit;

/****Conciliar las condonaciones****/

data DIARIOS.CONDONACIONES ;
  %let _EFIERR_ = 0; /* set the ERROR detection macro variable */
infile 'C:\Documents and Settings\m914089\Escritorio
\CONDONACIONES.txt' delimiter='09'x
lrecl=32767 firstobs=2 ;
informat Contrato $12. ;
informat MES date9. ;

format Contrato $12. ;

```

```

format MES date9. ;
INPUT
      CONTRATO $
      MES
/* set ERROR detection macro variable */
; run;

```

```

PROC SQL; CREATE TABLE CONDONACIONES AS
SELECT A.contrato
      ,A.mes
      , CASE WHEN C.Acct_num IS NULL THEN A.CONTRATO ELSE
C.Acct_num_nuevo END AS CONTRATO5
FROM DIARIOS.condonaciones AS A
LEFT JOIN DIARIOS.CONCILIACIONb AS C
      ON A.CONTRATO = C.Acct_num
; QUIT;

```

```

proc sql; create table condonaciones as
select r.CONTRATO,r.Fecha_De_Vto_,count(p.contrato5) as condonaciones
from CONDONACIONES as p
inner join Resumen as r
      ON p.contrato5= r.contrato
      And p.mes <= r.Fecha_De_Vto_
group by r.CONTRATO, r.Fecha_De_Vto_

;quit;

```

```

proc sql; create table Resumen as
select t.*
      ,CASE WHEN C.CONDONACIONES IS NULL THEN 0 ELSE C.CONDONACIONES
END AS NUM_CONDONACIONES
from resumen as t
left join CONDONACIONES as c
on T.CONTRATO=C.CONTRATO
AND T.Fecha_De_Vto_=C.Fecha_De_Vto_

;QUIT;

```

/\*\*CONCILIAR RENOVACIONES\*\*\*/

```

PROC SQL; CREATE TABLE RENOVACIONES AS
SELECT A.*
      , CASE WHEN C.Acct_num IS NULL THEN A.Origin ELSE
C.Acct_num_nuevo END AS CONTRATO_CONCILIADO
FROM RENOVACIONES AS A
LEFT JOIN DIARIOS.CONCILIACIONb AS C
      ON A.Origin = C.Acct_num
; QUIT;
PROC SQL; CREATE TABLE Resumen AS
SELECT R.*
      , CASE WHEN C.Acct_num_nuevo IS NULL THEN R.Contrato ELSE
C.Acct_num_nuevo END AS CONTRATO_VIEJO
FROM Resumen AS R
LEFT JOIN DIARIOS.CONCILIACIONb AS C
      ON R.Contrato = C.Acct_num_nuevo
; QUIT;

```

```

PROC SQL; CREATE TABLE RESUMEN AS
SELECT R.*
      ,CASE WHEN G.A IS NOT NULL THEN 1

```

```

        WHEN G.B IS NOT NULL THEN 2
          WHEN G.C IS NOT NULL THEN 3
            WHEN G.D IS NOT NULL THEN 4
              WHEN G.E IS NOT NULL THEN 5
                WHEN G.F IS NOT NULL THEN 6
                  WHEN G.G IS NOT NULL THEN 7
                ELSE 0 END AS NUM_RENOVACIONES
FROM resumen AS R
LEFT JOIN RENOVACIONES AS G
on r.CONTRATO_VIEJO=g.a
AND g.Fecha_renovacion<=R.Fecha_De_Vto_

;QUIT;

proc sql; create table Insumo_ler_tabla as
select CONTRATO
      , FLAG_GOOD_BAD
      , CASE WHEN Sexo IN ('FEMENINO') THEN 'FEMENINO'
              WHEN Sexo in ('MASCULINO') THEN 'MASCULINO' ELSE
"NULO" END AS SEXO_
      , case when Estado in ('DISTRITO FEDERAL','DISTRITO FEDERAL
D.F') then 'DISTRITO FEDERAL'
              WHEN Estado in ('ESTADO DE MEXICO','ESTADO DE MEXICO
MEX') THEN 'ESTADO DE MEXICO' ELSE 'OTRO' END AS EDO
      , CASE WHEN Nivel_Estudios IN
('DIPLOMADOS','Licenciatura','UNIVERSIDAD','POSTGRADO
(MAESTRIA)','Postgrado') THEN 'LICENCIATURA/DIPLOMADOS/POSTGRADO'
              WHEN Nivel_Estudios IN ('PREPARATORIA','Técnico o
Preparatoria') THEN 'Técnico o Preparatoria'
              WHEN Nivel_Estudios IN
('PRIMARIA','Primaria','SECUNDARIA','Secundaria') THEN 'Primaria o
Secundaria' ELSE 'NULO'END AS ESTUDIOS

      , CASE WHEN AUTOMOVIL IN ('SI') THEN 'SI'
              WHEN AUTOMOVIL IS NULL THEN "NO" ELSE 'NULO'END AS
AUTOMOVIL_
      , case when edad>=1 AND edad<=22 THEN "1 a 22"
              WHEN edad>=23 AND edad<=36 THEN "23 a 36"
              WHEN edad>=37 THEN "37 MAS" ELSE "NULO" END AS EDADES

      , CASE WHEN NoDependientes=0 THEN "0"
              when nodependientes is null then "0"
              WHEN NoDependientes=1 THEN "1"
              WHEN NoDependientes=2 then "2 "
              WHEN NoDependientes>2 THEN "2-MAS" END AS
NO_DEPENDIENTES

      ,CASE WHEN SCORE_BUENO<=550 THEN "<=550"
              WHEN SCORE_BUENO> 550 AND SCORE_BUENO< 570 THEN '550a570'
              WHEN SCORE_BUENO>= 570 AND SCORE_BUENO< 590 THEN
'570a590'
              WHEN SCORE_BUENO>= 590 AND SCORE_BUENO< 610 THEN
'590a610'
              WHEN SCORE_BUENO>= 610 AND SCORE_BUENO< 630 THEN
'610a630'
              WHEN SCORE_BUENO> 630 THEN
'630-MAS' ELSE 'NULO' END AS RANG_SCORE

      ,CASE WHEN ATP<=50 THEN "<=50"
              WHEN ATP>50 THEN '50MAS'
              ELSE "NULO" END AS ATP_

      , CASE WHEN EDOCIV IN ('CASADA/O BIENES S', 'CASADO/A BIENES

```

```

M', 'Casado','VIUDO/A', 'Viudo' ) THEN 'CASADO/VIUDO'
      WHEN EDOCIV IN ('DIVORCIADO/A', 'Divorciado',
'SOLTERO/A', 'Soltero') THEN 'SOLTERO/DIVORCIADO'
      WHEN EDOCIV IN ('SEPARADO/A', 'Separado','UNION
LIBRE', 'Unión Libre') THEN 'SEPARADO/UNIÓN LIBRE' ELSE 'NULO' END AS
ESTADO_CIVIL
      , CASE WHEN TIPOVIVIENDA IN
('FAMILIAR','Familiares','RENTADA','Rentada','PAGANDO','Pagandola')
THEN 'FAMILIARES/RENTADA/PAGÁNDOLA'
      WHEN TIPOVIVIENDA IN ( 'PROPIA','Propia') THEN
'PROPIA' ELSE 'NULO' END AS TIPO_VIVIENDA

      , CASE WHEN BC_SCORE<=650 THEN "<=650"
      WHEN BC_SCORE>650 AND BC_SCORE<=700 THEN '6500 a 700'
      WHEN BC_SCORE>700 AND BC_SCORE<=750 THEN '700 a
750'
      WHEN BC_SCORE>750
      THEN '750 mas' ELSE 'NULO' END AS RANG_BCSCORE

      , CASE WHEN CAPITAL<=5000 THEN "<=5000"
      WHEN CAPITAL>5000 THEN ">5000" ELSE 'NULO' END AS
MONTO

      , CASE WHEN PLAZO<=50 THEN "<=50"
      WHEN PLAZO>50 THEN ">50"
      ELSE 'NULO' END AS T_PLAZO

      , CASE WHEN Tipo_Cr_dito IN ('CONSUMO') THEN 'CONSUMO'
      WHEN Tipo_Cr_dito IN ('MICROCREC','MICRONEGOCIO') THEN
'MICRONEGOCIO' ELSE 'NULO' END AS TIPO_CREDITO

      , CASE WHEN Producto IN
('PREFERENCI','Prefere','Preferencial') THEN 'Preferencial'
      WHEN Producto IN ('REGULAR', 'Regular') THEN
'REGULAR' ELSE 'NULO' END AS TIPO_PRODUCTO

      , CASE WHEN Micro_cr_dito_Ubicaci_n IN('DE CASA EN
CASA','CASA','CASA FAMILIAR','DESDE CASA','DESDE SU CASA
DOMICILIO','EN CASA','EN SU CASA','EN EL HOGAR','EN EL LUGAR
DONDE RE','PATIO CASA','VENTA DESDE SU CASA',' VENTA EN SU
CASA','HOGAR',' CASA FAMILIAR','FAMILIAR' 'TRABA EN EL HOGAR')
THEN 'DOMICILIO'
      WHEN Micro_cr_dito_Ubicaci_n LIKE ('CASA') THEN
'DOMICILIO'
      WHEN Micro_cr_dito_Ubicaci_n IS NULL THEN 'NULO'
      WHEN Micro_cr_dito_Ubicaci_n IN('LOCAL') THEN
'LOCAL' ELSE 'OTRO' END AS UBICACION_MICROCRECITO

      , CASE WHEN Micro_cr_dito_Tipo IN ("PROPIO","CASA","EN SU
CASA","EN SU MISMO DOMICILI","FAMILIAR","FAMILIAR UBICADO EN
HOGAR") THEN 'PROPIO'
      WHEN Micro_cr_dito_Tipo IS NULL THEN 'NULO' ELSE 'NO
PROPIO' END AS TIPO_MICROCRECITO

      , CASE WHEN Micro_cr_dito_Lugar IN
('FIJO','LOCAL','SEMIFIJO','PUESTO') THEN 'FIJO'
      WHEN Micro_cr_dito_Lugar IN ('MOVIL','OTROS') THEN
'MOVIL' ELSE 'NULO' END AS Tipo_MicrocrEdito

      , CASE WHEN TipoSolicitud IN ('FORMER') THEN 'FORMER'
      WHEN TipoSolicitud IN('NORMAL','NUEVO') THEN 'NUEVO'
      WHEN TipoSolicitud IN('REESTR','REESTRUCTURA',) THEN
'REESTRUCTURA'
      WHEN TipoSolicitud IN('RENOVA','RENOVADA','RENOVADO')
THEN 'RENOVACIÓN 'ELSE 'NULO' END AS TIPO_SOLICITUD

      , CASE WHEN Periodicidad in ('MENSUAL','Mensual') then
'MENSUAL'
      WHEN Periodicidad in ('QUINCENAL','Quincenal')

```

```

THEN 'QUINCENAL'
      WHEN Periodicidad in ('SEMANAL','Semanal') THEN
'SEMANAL' ELSE 'NULO' END AS TIPO_PERIODICIDAD

      , CASE WHEN IngresoTotal<=10000 THEN '<=10000'
      WHEN IngresoTotal>10000 THEN '>10000'
      ELSE 'NULO' END AS INGRESO_TOTAL

      , CASE WHEN Gastos_Bur__Cr_dito>=0 AND Gastos_Bur__Cr_dito<=
500 THEN '0 a 500'
      WHEN Gastos_Bur__Cr_dito>500
      THEN '500 mas' ELSE 'NULO' END AS GTOS_BURO_CREDITO

      , CASE WHEN Pago_mensual_pr_stamo<=300 THEN "<=300"
      WHEN Pago_mensual_pr_stamo>300 AND
Pago_mensual_pr_stamo<=1000 THEN '300 a 1000'
      WHEN Pago_mensual_pr_stamo>1000 AND Pago_mensual_pr_stamo
<=3000 THEN '1000 a 3000'
      WHEN Pago_mensual_pr_stamo>3000 then '3000 mas'
      ELSE 'NULO' END AS PAGO_MENSUAL

      , CASE WHEN GastosTotales<=500 THEN "<=500"
      WHEN GastosTotales>500 AND GastosTotales<=899 THEN
'500-899'
      WHEN GastosTotales>899
      THEN '>899' ELSE 'NULO' END AS GTOS_TOT

      , CASE WHEN Como_se_Enterо IS NULL THEN 'NULO'
      WHEN Como_se_Enterо IN ("ANUNCIO AZOTEA", "ANUNCIO
CALLE/MANTA", "APASCO", "CAMBACEO",
"COMISIONISTA", "INTERNET", "LETRERO", "LLAMADA TELEFONICA"
,"MARQUESINA", "MODULO", "PASANDO POR
SUCURSAL", "PERIODICO", "POSTER / MICRONEG", "POSTER
MICRONEGOCIO", "REVISTA", "VOCEO", "VOLANTE",
"NINGUNO", "AMISTAD / CONOCIDO",
"RECOMENDACION", "FAMILIAR") THEN 'PUBLICIDAD'
      WHEN Como_se_Enterо IN ("RENOVACION", "REESTRUCTURA"
,"FORMER") THEN "RENOVACION/REESTRUCTURA/FORMER" ELSE 'NULO'
END AS CENTERO

      ,CASE WHEN TipoComprobanteIngresos IN ("CTA HABERES", "Carta
de haberes", "NOTA COMPRA", "Notas de compra", 'Carta laboral a mano',
'Carta laboral membretada') then 'CARTA LABORAL'
      WHEN TipoComprobanteIngresos IN ("EDO DE CTA", "Estado
de Cuenta", "Estado de cuenta", "Estados de
Cuenta", "FACTURAS", "Facturas") then 'FACTURAS'
      WHEN TipoComprobanteIngresos IN ("REC PENSION"
"Recibos de pensión" "NOMINA" "Nómina" "DECLARACION"
"Declaración de Impuestos") then 'NÓMINA'
      when TipoComprobanteIngresos is null then 'NULO' ELSE
'OTRO' END AS TIP_CPTE INGR

      , CASE WHEN FrecuenciaPrecepcionIngresos IN ("DIARIO",
"Diario") then "DIARIO"
      WHEN FrecuenciaPrecepcionIngresos IN ("MENSUAL",
"Mensual") then "MENSUAL"
      WHEN FrecuenciaPrecepcionIngresos IN ("QUINCENAL",
"Quincenal", "VARIABLE", "Variable") then "QUINCENAL/VARIABLE"
      WHEN FrecuenciaPrecepcionIngresos IN ("SEMANAL",
"Semanal") then "SEMANAL"
      WHEN FrecuenciaPrecepcionIngresos is null then
'NULO' ELSE 'OTRO' END AS FREC_PREC INGR

      , CASE WHEN WOB>=0 AND WOB<=5 THEN '0 a 5'
      WHEN WOB>5 AND WOB <=10 THEN '5 a 10'
      WHEN WOB>10 AND WOB <=20 THEN '10 a 20'

```

```

        WHEN WOB>20 AND WOB <=40 THEN '20 a 40'
        WHEN WOB>40 THEN '40 mas' ELSE 'NULO'
        END AS WOB_

, CASE WHEN SISTEMA='ANFEXI' THEN 'ANFEXI'
        WHEN SISTEMA='FINAY' THEN 'FINAY' ELSE 'NULO' END AS
NAC_SIST

, CASE WHEN TELS<=4 THEN '<=4'
        WHEN TELS>4 THEN '>4' ELSE 'NULO' END AS NUM_TELS

, CASE WHEN MAX_DM = 0 THEN '0'
        WHEN MAX_DM >= 1 AND MAX_DM <= 2 THEN '1-2'
        WHEN MAX_DM >= 3 AND MAX_DM <= 6 THEN '3-6' /***/
        WHEN MAX_DM >= 7 THEN '7 mas'
        ELSE 'NULO' END AS MAXIMO_DM

,CASE WHEN MAX_RV=0 then '0'
        WHEN MAX_RV>=1 AND MAX_RV<=4 THEN '1 a 4'
        WHEN MAX_RV>5 THEN '5mas/nulo' ELSE '5mas/nulo' END
AS MAXIMO_RV

, case when avg_dm = 0 then '0'
        WHEN avg_dm > 0 AND avg_dm <= 0.318 THEN '0-0.318'
        WHEN avg_dm > 0.318 THEN '0.318-+'
        ELSE 'NULO' END AS PROMEDIO_DM

, CASE WHEN pcappag>=0 AND pcappag<=10 THEN '0 a 10'
        WHEN pcappag>10 AND pcappag<=30 THEN '10 a 30'
        WHEN pcappag>30 AND pcappag<=50 THEN '30 a 50'
        WHEN pcappag>50 AND pcappag<=70 THEN '50 a 70'
        WHEN pcappag>70 and pcappag<=80 then '70 a 80'
        WHEN pcappag>80 and pcappag<=90 then '80 a 90'
        when pcappag>90 then '90mas'
        else 'NULO' end as pct_cappag

, CASE WHEN ctas_pagadas_final>=0 AND ctas_pagadas_final<=5
then '0 a 5'
        WHEN ctas_pagadas_final>5 AND ctas_pagadas_final<=10 THEN
'5 a 10'
        WHEN ctas_pagadas_final>10 AND ctas_pagadas_final<=15
THEN '10 a 15'
        WHEN ctas_pagadas_final>15 AND ctas_pagadas_final<=20
THEN '15 a 20'
        WHEN ctas_pagadas_final>20 AND ctas_pagadas_final<=40
THEN '20 a 40'
        WHEN ctas_pagadas_final>40 AND ctas_pagadas_final<=60
THEN '40 a 60'
        WHEN ctas_pagadas_final>60 AND ctas_pagadas_final<=80
THEN '60 a 80'
        WHEN ctas_pagadas_final>80 AND ctas_pagadas_final<=
100 THEN '80 a 100'
        WHEN ctas_pagadas_final>100 THEN '100 mas' ELSE
'NULO' END AS CTASPAGADAS

, CASE WHEN NUM_CONDONACIONES=0 THEN '0'
        WHEN NUM_CONDONACIONES>1 THEN 'MAS QUE CERO'
        ELSE 'NULO' END AS CONDONACIONES

, CASE WHEN Impagos= 0 then '0'

        when Impagos=1 then '1'
        WHEN Impagos>=2 AND Impagos<=3 THEN '2 a 3'
        WHEN Impagos>3 AND Impagos<=4 THEN '3 a 4'
        WHEN Impagos>4 AND Impagos<=5 THEN '4 a 5'

```



```

        WHEN Impagos>5 AND Impagos<=10 THEN '5 a 10'
        WHEN Impagos>10 AND Impagos<=15 THEN '10 a 15'
        WHEN Impagos>15

        THEN '15 mas' ELSE 'NULO' END AS NUM_IMPAGOS

, CASE when adelantos=0 then '0'
      when adelantos=1 then '1'
      WHEN Adelantos>1 AND Adelantos<=2 THEN '1 a 2'
      WHEN Adelantos>2 AND Adelantos<=3 THEN '2 a 3'
      WHEN Adelantos>3 AND Adelantos<=4 THEN '3 a 4'
      WHEN Adelantos>4 AND Adelantos<=5 THEN '4 a 5'
      WHEN Adelantos>5 AND Adelantos<=10 THEN '5 a 10'
      WHEN Adelantos>10 AND Adelantos<=15 THEN '10 a 15'
      WHEN Adelantos>15 AND Adelantos<=20 THEN '15 a 20'
      WHEN Adelantos>20
      then '20mas' else 'NULO' END AS NUM_ADELANTOS

, CASE WHEN V1RV=0 THEN '0'
      WHEN V1RV>1
      THEN '1 mas' ELSE 'NULO' END AS V_R1

, CASE WHEN V2RV=0 THEN '0'
      when v2rv>1 then '1MAS'
      ELSE 'NULO' END AS V_R2

, CASE WHEN V3RV=0 THEN '0'
      when v3rv>1 THEN 'lmas' ELSE 'NULO' END AS V_R3

, CASE WHEN V4RV=0 THEN '0 '
      when v4rv>1 THEN 'lmas' ELSE 'NULO' END AS V_R4

, CASE WHEN V5RV=0 then '0'
      when v5rv>1 THEN '1MAS' ELSE 'NULO' END AS V_R5

, CASE WHEN V6RV=0 then '0'
      when v6rv>1 then'lmas' else 'NULO' END AS v_r6

, CASE WHEN NUM_RENOVACIONES=0 THEN '0'
      WHEN NUM_RENOVACIONES>0 THEN 'MAS DE CERO' ELSE
'NULO' END AS RENOVACIONES

FROM RESUMEN
;QUIT;

```

```

PROC SQL; CREATE TABLE num_impagos AS
SELECT impagos , COUNT(CONTRATO) AS CLIENTES, SUM(BADS) AS BAD
FROM
(
SELECT num_impagos
      , CONTRATO
      , CASE WHEN FLAG_GOOD_BAD = 'BAD' THEN 1 ELSE 0 END AS BADS
FROM Insumo_1er_tabla
)
GROUP BY 1
;QUIT;

```

```

/*****/

```

```

proc sql; create table Insumo_2da_tabla as

```

```

select CONTRATO
,FLAG_GOOD_bad
,CASE WHEN FLAG_GOOD_BAD='GOOD' THEN 0
      WHEN FLAG_GOOD_BAD='BAD' THEN 1 ELSE 0 END AS BANDERA
,CASE WHEN Sexo IN ('FEMENINO') THEN 1
      WHEN Sexo in ('MASCULINO') THEN 2 ELSE 2 END AS SEXO_
,case when Estado in ('DISTRITO FEDERAL','DISTRITO FEDERAL
D.F') then 2
      when Estado is null then 3
      WHEN Estado in ('ESTADO DE MEXICO','ESTADO DE MEXICO
MEX') THEN 3 else 1 END AS EDO
,CASE WHEN Nivel_Estudios IN
('DIPLOMADOS','Licenciatura','UNIVERSIDAD','POSTGRADO
(MAESTRIA)','Postgrado') THEN 3
      WHEN Nivel_Estudios IN ('PREPARATORIA','Técnico o
Preparatoria') THEN 2
      WHEN Nivel_Estudios IN
('PRIMARIA','Primaria','SECUNDARIA','Secundaria') THEN 1 ELSE 4 END
AS ESTUDIOS
,CASE WHEN AUTOMOVIL IN ('SI') THEN 2
      WHEN AUTOMOVIL IS NULL THEN 1 ELSE 2 END AS AUTOMOVIL_
,case when edad>=1 AND edad<=22 THEN 3
      WHEN edad>=23 AND edad<=36 THEN 2
      WHEN edad>=37 THEN 1 ELSE 3 END AS EDADES
,CASE WHEN NoDependientes=0 THEN 1
      when nodependientes is null then 1
      WHEN NoDependientes=1 THEN 2
      WHEN NoDependientes=2 then 3
      WHEN NoDependientes>2 THEN 4 END AS NO_DEPENDIENTES
,CASE WHEN SCORE_BUENO<=550 THEN 6
      WHEN SCORE_BUENO> 550 AND SCORE_BUENO< 570 THEN 5
      WHEN SCORE_BUENO>= 570 AND SCORE_BUENO< 590 THEN 4
      WHEN SCORE_BUENO>= 590 AND SCORE_BUENO< 610 THEN 3
      WHEN SCORE_BUENO>= 610 AND SCORE_BUENO< 630 THEN 2
      WHEN SCORE_BUENO> 630 THEN 1
ELSE 6 END AS RANG_SCORE
,CASE WHEN ATP<=50 THEN 2
      WHEN ATP>50 THEN 1
      ELSE 2 END AS ATP_
,CASE WHEN EDOCIV IN ('CASADA/O BIENES S','CASADO/A BIENES
M','Casado','VIUDO/A','Viudo') THEN 1
      WHEN EDOCIV IN ('DIVORCIADO/A','Divorciado',
'SOLTERO/A','Soltero') THEN 3
      WHEN EDOCIV IN ('SEPARADO/A','Separado','UNION
LIBRE','Unión Libre') THEN 2 ELSE 3 END AS ESTADO_CIVIL
,CASE WHEN TIPOVIVIENDA IN
('FAMILIAR','Familiares','RENTADA','Rentada','PAGANDO','Pagandola')
THEN 2
      WHEN TIPOVIVIENDA IN ('PROPIA','Propia') THEN 1
ELSE 2 END AS TIPO_VIVIENDA
,CASE WHEN BC_SCORE<=650 THEN 3
      WHEN BC_SCORE>650 AND BC_SCORE<=700 THEN 4
      WHEN BC_SCORE>700 AND BC_SCORE<=750 THEN 2
      WHEN BC_SCORE>750
      THEN 1 ELSE 4 END AS RANG_BCSCORE
,CASE WHEN FLAG_HIT_NOHIT IN ("HIT") THEN 2
      WHEN FLAG_HIT_NOHIT IN ("NOHIT") THEN 1 ELSE 2 END
AS BANDERA_HIT_NOHIT
,CASE WHEN CAPITAL<=5000 THEN 1
      WHEN CAPITAL>5000 THEN 2 ELSE 2 END AS MONTO
,CASE WHEN PLAZO<=50 THEN 2
      WHEN PLAZO>50 THEN 1
      ELSE 2 END AS T_PLAZO
,CASE WHEN Tipo_Cr_dito IN ('CONSUMO') THEN 2
      WHEN Tipo_Cr_dito IN ('MICROCRED','MICRONEGOCIO')
THEN 1 ELSE 3 END AS TIPO_CREDITO

```

```

,CASE WHEN Producto IN
('PREFERENCI','Prefere','Preferencial') THEN 2
      WHEN Producto IN ('REGULAR', 'Regular') THEN 1 ELSE
3 END AS TIPO_PRODUCTO
      ,CASE WHEN Micro_cr_dito_Ubicaci_n IN('DE CASA EN
CASA','CASA','CASA FAMILIAR','DESDE CASA','DESDE SU CASA
DOMICILIO','EN CASA','EN SU CASA','EN EL HOGAR','EN EL LUGAR
DONDE RE','PATIO CASA','VENTA DESDE SU CASA',' VENTA EN SU
CASA','HOGAR', 'CASA FAMILIAR','FAMILIAR' 'TRABA EN EL HOGAR')
THEN 2
      WHEN Micro_cr_dito_Ubicaci_n LIKE ('CASA') THEN 2
      WHEN Micro_cr_dito_Ubicaci_n IS NULL THEN 1
      WHEN Micro_cr_dito_Ubicaci_n IN('LOCAL') THEN 4 ELSE
3 END AS ubicacion_microcred
      ,CASE WHEN Micro_cr_dito_Tipo IN ("PROPIO","CASA","EN SU
CASA","EN SU MISMO DOMICILI","FAMILIAR","FAMILIAR UBICADO EN
HOGAR") THEN 2
      WHEN Micro_cr_dito_Tipo IS NULL THEN 1 ELSE 3 END AS
tipo_microcred
      ,CASE WHEN Micro_cr_dito_Lugar IN
('FIJO','LOCAL','SEMIFIJO','PUESTO') THEN 1
      WHEN Micro_cr_dito_Lugar IN ('MOVIL','OTROS') THEN 2
ELSE 3 END AS LUGAR_MICROCREDITO
      ,CASE WHEN TipoSolicitud IN ('FORMER') THEN 1
      WHEN TipoSolicitud IN('NORMAL','NUEVO') THEN 2
      WHEN TipoSolicitud IN('REESTR','REESTRUCTURA',) THEN
4
      WHEN TipoSolicitud IN('RENOVA','RENOVADA','RENOVADO')
THEN 3 ELSE 5 END AS TIPO_SOLICITUD
      ,CASE WHEN Periodicidad in ('MENSUAL','Mensual') then 4
      WHEN Periodicidad in ('QUINCENAL','Quincenal') THEN 3
      WHEN Periodicidad in ('SEMANAL','Semanal') THEN 1 ELSE 2
END AS TIPO_PERIODICIDAD
      ,CASE WHEN IngresoTotal<=10000 THEN 2
      WHEN IngresoTotal>10000
      THEN 1 ELSE 2 END AS INGRESO_TOTAL
      ,CASE WHEN Gastos_Bur__Cr_dito>=0 AND Gastos_Bur__Cr_dito<=500
THEN 1
      WHEN Gastos_Bur__Cr_dito>500
      THEN 2 ELSE 2 END AS GTOS_BURO_CREDITO
      ,CASE WHEN Pago_mensual_pr_stamo<=300 THEN 4
      WHEN Pago_mensual_pr_stamo>300 AND
Pago_mensual_pr_stamo<=1000 THEN 2
      WHEN Pago_mensual_pr_stamo>1000 AND
Pago_mensual_pr_stamo<=3000 THEN 3
      WHEN Pago_mensual_pr_stamo>3000
      THEN 1 ELSE 4 END AS PAGO_MENSUAL
      ,CASE WHEN GastosTotales<=500 THEN 1
      WHEN GastosTotales>500 AND GastosTotales<=899 THEN 2
      WHEN GastosTotales>899
      THEN 3 ELSE 3 END AS GTOS_TOT
      ,CASE WHEN Como_se_Enterо IS NULL THEN 3
      WHEN Como_se_Enterо IN ("ANUNCIO AZOTEA", "ANUNCIO
CALLE/MANTA", "APASCO", "CAMBACEO",
"COMISIONISTA", "INTERNET", "LETRERO", "LLAMADA TELEFONICA",
"MARQUESINA", "MODULO", "PASANDO POR
SUCURSAL", "PERIODICO", "POSTER / MICRONEG", "POSTER
MICRONEGOCIO", "REVISTA", "VOCEO", "VOLANTE",
"NINGUNO", "AMISTAD / CONOCIDO",
"RECOMENDACION", "FAMILIAR") THEN 1
      WHEN Como_se_Enterо IN ("RENOVACION", "REESTRUCTURA",
"FORMER") THEN 2 ELSE 1 END AS CENTERO
      ,CASE WHEN TipoComprobanteIngresos IN ("CTA HABERES", "Carta
de haberes", "NOTA COMPRA","Notas de compra", 'Carta laboral a mano',
'Carta laboral membretada') then 2
      WHEN TipoComprobanteIngresos IN ("EDO DE CTA", "Estado

```

```

de Cuenta", "Estado de cuenta"      ,"Estados de
Cuenta"      ,"FACTURAS",      "Facturas") then 4
      WHEN TipoComprobanteIngresos IN ("REC PENSION"
"Recibos de pensión"      "NOMINA"      "Nómina"      "DECLARACION"
"Declaración de Impuestos") then 3
      when TipoComprobanteIngresos is null then 1 ELSE 1 END
AS TIP_CPTE_INGR
      ,CASE WHEN FrecuenciaPrecepcionIngresos IN ("DIARIO",
"Diario") then 5
      WHEN FrecuenciaPrecepcionIngresos IN ("MENSUAL",
"Mensual") then 2
      WHEN FrecuenciaPrecepcionIngresos IN ("QUINCENAL",
"Quincenal", " VARIABLE", "Variable") then 6
      WHEN FrecuenciaPrecepcionIngresos IN ("SEMANAL",
"Semanal") then 3
      WHEN FrecuenciaPrecepcionIngresos is null then 4
ELSE 1 END AS FREC_PREC_INGR
      ,CASE WHEN WOB>=0 AND WOB<=5 THEN 1
      WHEN WOB>5 AND WOB <=10 THEN 2
      WHEN WOB>10 AND WOB <=20 THEN 3
      WHEN WOB>20 AND WOB <=40 THEN 4
      WHEN WOB>40 THEN 1 ELSE 1 END AS WOB_
      ,CASE WHEN SISTEMA='ANFEXI' THEN 2
      WHEN SISTEMA='FINAY' THEN 1 ELSE 2 END AS NAC_SIST
,CASE WHEN TELS<=4 THEN 2
      WHEN TELS>4 THEN 1 ELSE 2 END AS NUM_TELS
, CASE WHEN MAX_DM = 0 THEN 1
      WHEN MAX_DM >= 1 AND MAX_DM <= 2 THEN 2
      WHEN MAX_DM >= 3 AND MAX_DM <= 6 THEN 3
      WHEN MAX_DM >= 7 THEN 4
      ELSE 4 END AS MAXIMO_DM
,CASE WHEN MAX_RV=0 then 1
      WHEN MAX_RV>=1 AND MAX_RV<=4 THEN 2
      WHEN MAX_RV>5 THEN 3 ELSE 3 END AS MAXIMO_RV

, case when avg_dm = 0 then 1
      WHEN avg_dm > 0 AND avg_dm <= 0.318 THEN 2
      WHEN avg_dm > 0.318 THEN 3
      ELSE 3 END AS PROMEDIO_DM

,CASE WHEN pcappag>=0 AND pcappag<=10 THEN 7
      WHEN pcappag>10 AND pcappag<=30 THEN 6
      WHEN pcappag>30 AND pcappag<=50 THEN 5
      WHEN pcappag>50 AND pcappag<=70 THEN 4
      WHEN pcappag>70 and pcappag<=80 then 2
      WHEN pcappag>80 and pcappag<=90 then 3
      when pcappag>90 then 1 else 7 end as pct_cappag
,CASE WHEN ctas_pagadas_final>=0 AND ctas_pagadas_final<=5
then 3
      WHEN ctas_pagadas_final>5 AND ctas_pagadas_final<=10
THEN 4
      WHEN ctas_pagadas_final>10 AND ctas_pagadas_final<=
15 THEN 6
      WHEN ctas_pagadas_final>15 AND ctas_pagadas_final<=
20 THEN 9
      WHEN ctas_pagadas_final>20 AND ctas_pagadas_final<=
40 THEN 8
      WHEN ctas_pagadas_final>40 AND ctas_pagadas_final<=
60 THEN 7
      WHEN ctas_pagadas_final>60 AND ctas_pagadas_final<=
80 THEN 5
      WHEN ctas_pagadas_final>80 AND ctas_pagadas_final<=
100 THEN 2
      WHEN ctas_pagadas_final>100 THEN 1 ELSE 9 END AS
CTASPAGADAS

```

```

,CASE WHEN NUM_CONDONACIONES=0 THEN 1
      WHEN NUM_CONDONACIONES>1 THEN 2
      ELSE 2 END AS CONDONACIONES
, CASE WHEN Impagos= 0 then 1

      when Impagos=1 then 2
      WHEN Impagos>=2 AND Impagos<=3 THEN 3
      WHEN Impagos>3 AND Impagos<=4 THEN 4
      WHEN Impagos>4 AND Impagos<=5 THEN 5
      WHEN Impagos>5 AND Impagos<=10 THEN 6
      WHEN Impagos>10 AND Impagos<=15 THEN 7
      WHEN Impagos>15

      THEN 8 ELSE 8 END AS NUM_IMPAGOS

, CASE when adelantos=0 then 10
      when adelantos=1 then 9
      WHEN Adelantos>1 AND Adelantos<=2 THEN 8
      WHEN Adelantos>2 AND Adelantos<=3 THEN 7
      WHEN Adelantos>3 AND Adelantos<=4 THEN 6
      WHEN Adelantos>4 AND Adelantos<=5 THEN 5
      WHEN Adelantos>5 AND Adelantos<=10 THEN 4
      WHEN Adelantos>10 AND Adelantos<=15 THEN 3
      WHEN Adelantos>15 AND Adelantos<=20 THEN 2
      WHEN Adelantos>20
      then 1 else 10 END AS NUM_ADELANTOS

, CASE WHEN V1RV=0 THEN 1
      WHEN V1RV>1
      THEN 3 ELSE 2 END AS V_R1

, CASE WHEN V2RV=0 THEN 1
      when v2rv>1 then 3
      ELSE 2 END AS V_R2

, CASE WHEN V3RV=0 THEN 1
      when v3rv>1 THEN 2 ELSE 3 END AS V_R3

, CASE WHEN V4RV=0 THEN 1
      when v4rv>1 THEN 3 ELSE 2 END AS V_R4

, CASE WHEN V5RV=0 then 1
      when v5rv>1 THEN 3 ELSE 2 END AS V_R5

, CASE WHEN NUM_RENOVACIONES=0 THEN 1
      WHEN NUM_RENOVACIONES>0 THEN 2 ELSE 2 END AS
RENOVACIONES

FROM RESUMEN ;quit;

libname insumo 'D:\Bases\Insumo' ;

data insumo.Insumo_2da_tabla ; set Insumo_2da_tabla; run;

/*
                                Análisis del estimador de máxima
verosimilitud

                                Error
Chi-cuadrado                                Pr > ChiSq
de Wald      Parámetro                    DF      Estimador      estándar
149.2820      Intercept                    1        -3.0833      0.2524
              <.0001

```

3.5585	SEXO_ 0.0592	1	0.0532	0.0282
2.1303	ESTUDIOS 0.1444	1	-0.0411	0.0282
12.9354	EDADES 0.0003	1	0.0802	0.0223
5.4823	NO_DEPENDIENTES 0.0192	1	0.0438	0.0187
20.9648	RANG_BCSCORE <.0001	1	0.0830	0.0181
10.6936	T_PLAZO 0.0011	1	0.1625	0.0497
12.6066	TIPO_CREDITO 0.0004	1	-0.2405	0.0677
3.2848	LUGAR_MICROCREDITO 0.0699	1	0.0670	0.0370
12.0460	TIPO_SOLICITUD 0.0005	1	-0.0869	0.0250

1  
0:57 Saturday, August 29, 2011 18

1

Procedimiento LOGISTIC

Análisis del estimador de máxima

verosimilitud

Chi-cuadrado de Wald	Parámetro Pr > ChiSq	DF	Estimador	Error estándar
60.9894	TIPO_PERIODICIDAD <.0001	1	0.2168	0.0278
4.5946	INGRESO_TOTAL 0.0321	1	0.0748	0.0349
2.5183	PAGO_MENSUAL 0.1125	1	0.0372	0.0234
17.8863	WOB_ <.0001	1	-0.0597	0.0141
5.4978	NAC_SIST 0.0190	1	-0.1770	0.0755
60.3892	MAXIMO_DM <.0001	1	0.2610	0.0336
75.7516	PROMEDIO_DM <.0001	1	0.4183	0.0481
141.1125	pct_cappag <.0001	1	-0.1268	0.0107
534.9414	CTASPAGADAS <.0001	1	-0.2195	0.00949
3.5299	CONDONACIONES 0.0603	1	0.2311	0.1230
2164.6925	NUM_IMPAGOS <.0001	1	0.5560	0.0120
13.5630	V_R1 0.0002	1	-0.0877	0.0238
40.0358	V_R4 <.0001	1	-0.2823	0.0446
59.2926	RENOVACIONES <.0001	1	-0.2840	0.0369

```

PROC SQL; CREATE TABLE INS_PREVENT_SCORE AS
SELECT z.*
      , 700+(80/LOG(2))*(-3.0833+
                        0.0532*SEXO_

```

```

-0.0411*ESTUDIOS
+0.0802* EDADES
+0.0438*NO_DEPENDIENTES
+0.0830*RANG_BCSCORE
+0.1625*T_PLAZO
-0.2405*TIPO_CREDITO
+0.0670*LUGAR_MICROCREDITO
-0.0869 *TIPO_SOLICITUD
+0.2168*TIPO_PERIODICIDAD
+0.0748*INGRESO_TOTAL
+0.0372*PAGO_MENSUAL
-0.0597 * WOE_
-0.1770*NAC_SIST
+0.2610*MAXIMO_DM
+0.4183*PROMEDIO_DM
-0.1268*pct_cappag
-0.2195* CTASPAGADAS
+0.2311* CONDONACIONES
+0.5560* NUM_IMPAGOS
-0.0877 * V_R1
-0.2823*V_R4
-0.2840*RENOVACIONES
)
AS PSCORE

FROM INSUMO.INSUMO_2DA_TABLA AS z
;QUIT;

PROC SQL; CREATE TABLE VALIDA00 AS SELECT MAX(PSCORE), MIN(PSCORE)
FROM INS_PREVENT_SCORE ; QUIT;

PROC SQL; CREATE TABLE RESUMEN_INS_PREVENT_SCORE AS
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS

FROM (
SELECT
CONTRATO
/*, CASE WHEN TIPO_PERIODICIDAD in (2,1) THEN 'SEMANAL'
WHEN tipo_PERIODICIDAD=4 THEN 'MENSUAL'
WHEN TIPO_PERIODICIDAD=3 THEN 'QUINCENAL' END AS PERIODO*/
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'
WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'
WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'
WHEN PSCORE > 750 THEN '750-MAS'
END AS P_SCORE
, BANDERA
FROM INS_PREVENT_SCORE
)Z
GROUP BY 1
; QUIT;

/*PROC SQL; CREATE TABLE RESUMEN_INS_PREVENT_SCORE AS
SELECT PERIODO, P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA)
AS BADS

FROM (

```

```

SELECT
CONTRATO
, CASE WHEN TIPO_PERIODICIDAD in (2,1) THEN 'SEMANAL'
      WHEN tipo_PERIODICIDAD=4 THEN 'MENSUAL'
      WHEN TIPO_PERIODICIDAD=3 THEN 'QUINCENAL' END AS PERIODO
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
      WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
      WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
      WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
      WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
      WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
      WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
      WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
      WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
      WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'
      WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'
      WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'
      WHEN PSCORE > 750 THEN '750-MAS'
      END AS P_SCORE
, BANDERA
FROM INS_PREVENT_SCORE
)Z
GROUP BY 1, 2
;      QUIT;*/

PROC SQL; CREATE TABLE SEMANAL AS
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS
from(select *
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
      WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
      WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
      WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
      WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
      WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
      WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
      WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
      WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
      WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'
      WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'
      WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'
      WHEN PSCORE > 750 THEN '750-MAS'
      END AS P_SCORE
      FROM INS_PREVENT_SCORE

)Z

WHERE TIPO_PERIODICIDAD in (2,1)

group by 1
; QUIT;

```

```

PROC SQL; CREATE TABLE MENSUAL AS
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS
from(select *
, CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'
      WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'
      WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'
      WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'
      WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'
      WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'
      WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'
      WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'
      WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'
      WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'

```



```
    WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'  
    WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'  
    WHEN PSCORE > 750 THEN '750-MAS'  
END AS P_SCORE  
FROM INS_PREVENT_SCORE
```

```
)Z
```

```
WHERE tipo_PERIODICIDAD=4  
group by 1  
; QUIT;
```

```
PROC SQL; CREATE TABLE QUINCENAL AS  
SELECT P_SCORE , COUNT(CONTRATO) AS CLIENTES, SUM(BANDERA) AS BADS  
from(select *  
    , CASE WHEN PSCORE > 50 AND PSCORE <= 200 THEN '50-200'  
    WHEN PSCORE > 200 AND PSCORE <= 250 THEN '200-250'  
    WHEN PSCORE > 250 AND PSCORE <= 300 THEN '250-300'  
    WHEN PSCORE > 300 AND PSCORE <= 350 THEN '300-350'  
    WHEN PSCORE > 350 AND PSCORE <= 400 THEN '350-400'  
    WHEN PSCORE > 400 AND PSCORE <= 450 THEN '400-450'  
    WHEN PSCORE > 450 AND PSCORE <= 500 THEN '450-500'  
    WHEN PSCORE > 500 AND PSCORE <= 550 THEN '500-550'  
    WHEN PSCORE > 550 AND PSCORE <= 600 THEN '550-600'  
    WHEN PSCORE > 600 AND PSCORE <= 650 THEN '600-650'  
    WHEN PSCORE > 650 AND PSCORE <= 700 THEN '650-700'  
    WHEN PSCORE > 700 AND PSCORE <= 750 THEN '700-750'  
    WHEN PSCORE > 750 THEN '750-MAS'  
END AS P_SCORE  
FROM INS_PREVENT_SCORE
```

```
)Z
```

```
WHERE TIPO_PERIODICIDAD=3  
group by 1  
; QUIT;
```



# Bibliografía

- [1] *Circular Única de Bancos (DISPOSICIONES DE CARACTER GENERAL APLICABLES A LAS INSTITUCIONES DE CREDITO)*, Diario Oficial de la Federación 2011, <http://www.cnbv.gob.mx/Prensa/Presentaciones%20Seminario%20Corresponsales/i.%20Circular%20%20C3%20%9Anica%20de%20Bancos.pdf>.
- [2] Lipschutz S.y Lipson M., *Probabilidad*, Editorial McGraw Hill 2a. edición.
- [3] Mendenhall W. y Terry S., *Un segundo curso de Estadística: Análisis de regresión*. Editorial Prentice Hall 7a. edición.
- [4] Newbold Paul, *Estadística para la Administración y Economía*. Editorial PRENTICE HALL. 6<sup>a</sup> Edición, 2008.
- [5] Pagano Robert R., *Estadística para las ciencias del comportamiento*. Editorial CENGAGE LEARNING. 9<sup>a</sup> Edición, 2011.
- [6] Pérez U., *El costo de la canasta básica se incrementó 70 por ciento en 10 años, informa la CNC*, La Jornada, <http://www.jornada.unam.mx/2011/02/08/index.php?section=politica&article=017n1pol>.