



Maestría en Tecnología de Bases de Datos

Documento Final de Proyecto de Investigación Aplicada 2

Un Data Warehouse para la Gestión de Riesgo de Crédito

Rivera Elizondo Juan Carlos

Febrero 2024

TRIBUNAL EXAMINADOR

Este proyecto fue aprobado por el Tribunal Examinador de la carrera: **Maestría en Tecnología de Bases de Datos**, requisito para optar por el título de grado de **Maestría**, para el estudiante: **Rivera Elizondo Juan Carlos**.

JOSE ALBERTO CABEZAS
JAIKEL (FIRMA)
JAIKEL (FIRMA)

Digitally signed by JOSE ALBERTO CABEZAS
JAIKEL (FIRMA)
Date: 2024.03.06
10:38:43 -06'00'

MBD José Cabezas Jaikel
Tutor

MARISOL NUÑEZ VASQUEZ
(FIRMA)

Firmado digitalmente por MARISOL NUÑEZ VASQUEZ (FIRMA)
Fecha: 2024.03.15
15:52:34 -06'00'

MBD. Marisol Núñez Vásquez
Lector 1

JASON ULLOA HERNANDEZ
(FIRMA)

Firmado digitalmente por JASON ULLOA HERNANDEZ (FIRMA)
Fecha: 2024.03.18
11:43:23 -06'00'

M.Sc. Jason Ulloa Hernández
Lector 2



San José, Costa Rica, 29 de febrero de 2024

Un Data Warehouse para la Gestión de Riesgo de Crédito

Primero Juan Carlos Rivera Elizondo, Segundo José Alberto Cabezas Jaikel

Resumen—Este documento contiene una investigación aplicada, en la que se ofrece una solución a las instituciones del sector financiero que se dedican al otorgamiento de crédito, para que puedan gestionar apropiadamente su riesgo de crédito. La solución se plantea por medio de depósitos de datos (Data Warehouses) y cubos OLAP.

Abstract—This document contains applied research, which aims to offer a solution to financial sector institutions that are dedicated to granting credit, so that they can appropriately manage their credit risk. The solution is proposed through Data Warehouses and OLAP cubes.

Palabras claves—Riesgo crediticio, morosidad, matriz de transición, tasa de recuperación, probabilidad de incumplimiento, depósitos de datos, cubos OLAP, ETL.

I. INTRODUCCIÓN

En la actualidad, las entidades financieras que otorgan crédito necesitan medir el riesgo asociado a participar en este negocio. Entre estas, las de medianas a grandes dimensiones suelen ostentar una división de riesgo financiero que se encarga de gestionar el riesgo de crédito. En contraste, las entidades de pequeña envergadura, por las dimensiones de sus recursos, pueden carecer de una división propiamente dicha dentro de su estructura, pero sí delegan sobre alguno de sus funcionarios con conocimiento en riesgo financiero, la función de gestionar el riesgo de crédito.

El riesgo de crédito se delimita como la probabilidad de pérdida sufrida por el acreedor ante el incumplimiento de pago por parte del deudor. Tal incumplimiento puede ser parcial, cuando el deudor liquida solo una fracción de la deuda, o total, en el caso de que el deudor no realice pago alguno.

La gestión de riesgo implica evaluar dicha probabilidad y concebir estrategias que amortigüen las eventuales pérdidas.

Gestionar conlleva medir; medir para controlar, medir para diseñar estrategias mitigadoras del impacto del riesgo sobre la cartera crediticia. Medir implica recolectar datos, y en riesgo de crédito muchas veces se requiere que estos datos sean históricos, a veces de hasta 5 años o más. Tal necesidad reclama habilidades y herramientas para la gestión eficiente de volúmenes considerables de datos.

Actualmente, y desde tiempos inmemoriales, las

instituciones dedicadas a la concesión de créditos enfrentan la imperativa tarea de gestionar el riesgo inherente a su actividad. Para gestionarlo, necesitan procesar información histórica cuyas muestras hayan sido tomadas con una frecuencia mínima de un mes, y cuyo rango de análisis puede superar los cinco años. Esto representa grandes cantidades de información a procesar, pero existen aún instituciones que no usan herramientas adecuadas para procesar este volumen de datos, que les permitan obtener los resultados con la prontitud, precisión y eficiencia que exige la actividad empresarial.

La gestión del riesgo crediticio demanda procesar ingentes cantidades de datos. Por lo general, estos datos son suministrados en archivos de texto, mensualmente, y el acumulado de cinco años o más puede sobrepasar fácilmente el millón y medio de registros o los 2 GB de espacio de almacenamiento. Tal realidad invalida herramientas como Excel y Access para la gestión del riesgo crediticio.

Algunas instituciones, al desconocer herramientas apropiadas, insisten en usar estas, lo que conlleva consumir notorias cantidades de tiempo en la obtención de resultados, mediante procesos excesivamente manuales, propensos a errores humanos. Esta situación incide negativamente en una gestión ineficaz del riesgo crediticio, al carecer de indicadores oportunos, o tomar decisiones basadas en indicadores erróneos alejados de la realidad.

II. MATERIALES Y MÉTODOS

La automatización del almacenamiento y procesamiento de registros históricos, con el fin de gestionar el riesgo crediticio, se concreta eficientemente mediante un Data Warehouse. Esta herramienta, implementable a través de Microsoft SQL Server u otros motores de bases de datos, da pie al desarrollo de cubos OLAP¹. Dichos cubos posibilitan la generación de indicadores cruciales, tales como: curvas de morosidad, análisis de cosechas, matrices de transición, tasas de recuperación y probabilidades de impago, inherentes a la cartera de crédito, los cuales constituyen indicadores esenciales para la gestión de riesgo crediticio.

La Fig. 1 desglosa el flujo de datos en una solución de esta índole. Los datos ingresan mediante archivos de texto provenientes de las bases de datos transaccionales de la institución. Estos archivos se cargan temporalmente en el Área

¹ OLAP: Procesamiento Analítico En Línea por sus siglas en inglés (On Line Analytical Processing).

de Ensayo (*Staging Area*), donde son procesados mediante ETLs² para su almacenamiento permanente en el Data Warehouse. Este último es consumido por cubos OLAP, implementados en SSAS³. Los cubos son manipulables con fines de análisis y presentación mediante tablas dinámicas en Excel.



Fig. 1. Flujo de datos desde la entrada hasta la salida.

Para el desarrollo de esta investigación se obtuvo la cartera de crédito de una institución costarricense anónima, entre los años 2013 y 2017. La identidad de las operaciones y de los clientes fueron protegidas por el derecho que le otorga la ley 8968 (Ley de Protección de la Persona frente al tratamiento de sus datos personales).

El Data Warehouse que se diseñó e implementó para almacenar los datos de esta cartera de crédito se muestra en la Fig. 2.

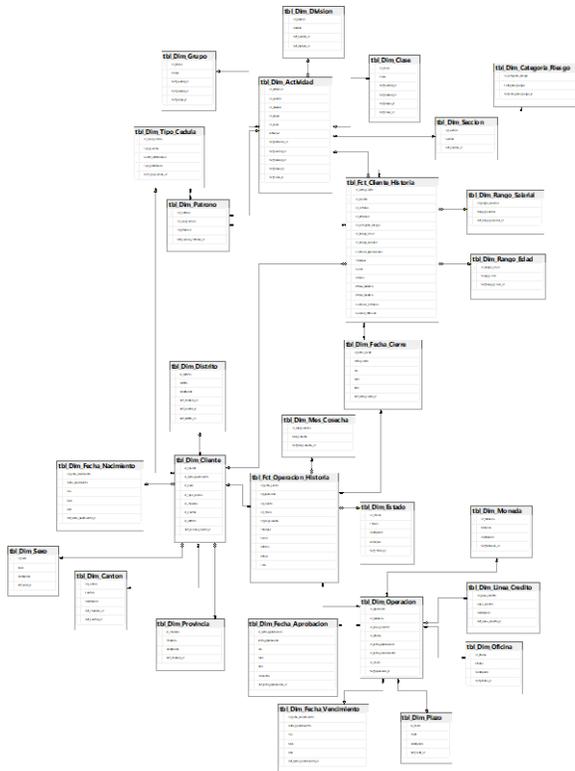


Fig. 2. Diagrama físico del Data Warehouse db_RC_DW.

Este está compuesto por 2 tablas de hechos y 26 tablas de dimensiones. Algunas tablas de dimensiones se comparten

² ETL: Extracción, Transformación y Carga por sus siglas en inglés (Extract-Transform-Load).

entre las tablas de hechos. El Data Warehouse usa un esquema en copo de nieve.

La Fig. 3 ilustra los objetos que componen el Área de Ensayo (Staging Area) db_RC_SA. La tabla tbl_Bureau_Credito recibe datos del archivo de texto correspondiente al bureau de crédito. La tabla tbl_Cartera_Credito acoge datos relativos a la cartera de crédito. Las dos tablas restantes funcionan como tablas intermedias para respaldar la transformación de los datos. El procedimiento almacenado encargado de extraer, transformar y cargar los datos en el Data Warehouse es sp_Importar_Datos. Los otros tres procedimientos son invocados por sp_Importar_Datos y representan etapas modulares del proceso.

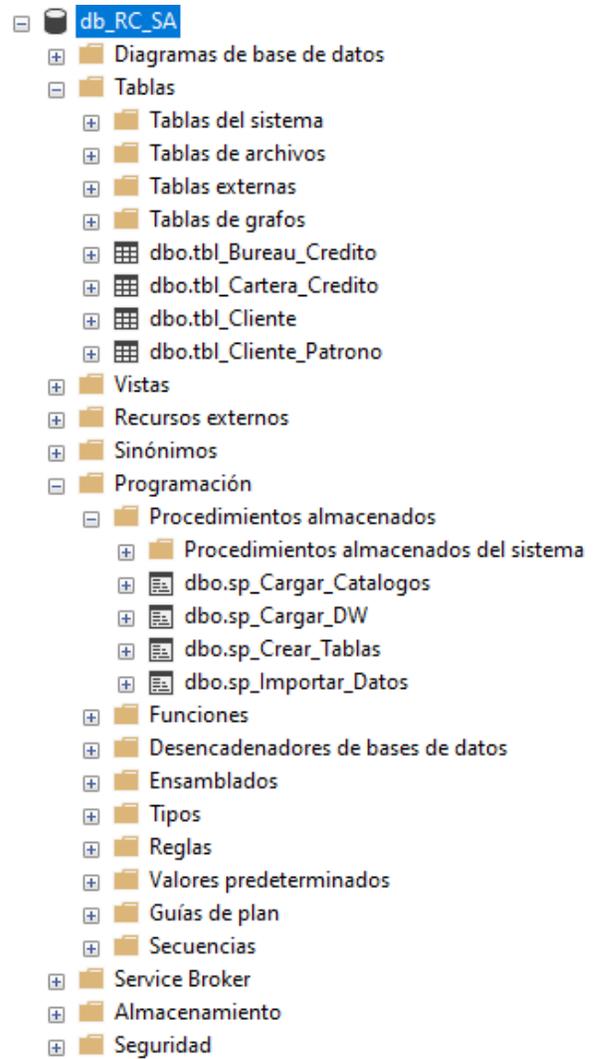


Fig. 3. Objetos que componen el staging area db_RC_SA.

La Fig. 4 presenta la estructura del cubo cbo_Mora, destinado al análisis de la morosidad en la cartera de crédito. Adopta el esquema estrella y cuenta con dimensiones como: la fecha de cierre de la cartera (Dim_Fecha_Cierre), el sexo del cliente

³ SSAS: SQL Server Analysis Services. Herramienta que ofrece Microsoft para procesamiento analítico en línea y minería de datos.

(Dim_Sexo), la moneda de la operación (Dim_Moneda), la provincia del cliente cuando entró a cartera (Dim_Provincia), el sector económico al que pertenece el cliente (Dim_Seccion), la fecha de nacimiento del cliente (Dim_Fecha_Nacimiento), el rango de edad al que pertenece el cliente (Dim_Rango_Edad), el rango salarial del cliente (Dim_Rango_Salarial) y el tipo de cédula del cliente (Dim_Tipo_Cedula). Entre las medidas se encuentran: cantidad de operaciones (Operacion), monto del principal (Principal), monto del saldo (Saldo), monto de intereses (Interes), días de atraso (Atraso), tasa de interés (Tasa), cantidad de operaciones en mora mayor a 90 (Operación_MM90) y saldo en mora mayor a 90 (Saldo_MM90).

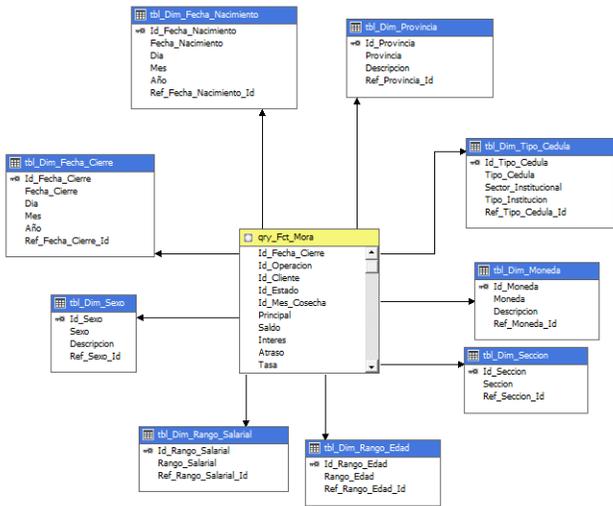


Fig. 4. Diagrama físico del cubo cbo_Mora.

En la Fig. 5 se muestra el esquema estrella con el que se construyó el cubo cbo_Cosecha. Este se utiliza para analizar las cosechas de la cartera de crédito. Cosecha se refiere al comportamiento de la mora mayor a 90 a lo largo del tiempo, de las operaciones aprobadas en un trimestre determinado del año. Independientemente del trimestre en que se aprobaron, todas se comparan desde el mes cero. Este cubo cuenta con las dimensiones: mes de cosecha (Dim_Mes_Cosecha), fecha de aprobación (Dim_Fecha_Aprobacion), el estado de la operación (Dim_Estado), el rango de edad del cliente (Dim_Rango_Edad), el rango salarial del cliente (Dim_Rango_Salarial), el sexo del cliente (Dim_Sexo) y el tipo de cédula del cliente (Dim_Tipo_Cedula). También cuenta con las medidas: cantidad de operaciones (Operacion) y el monto del saldo (Saldo).

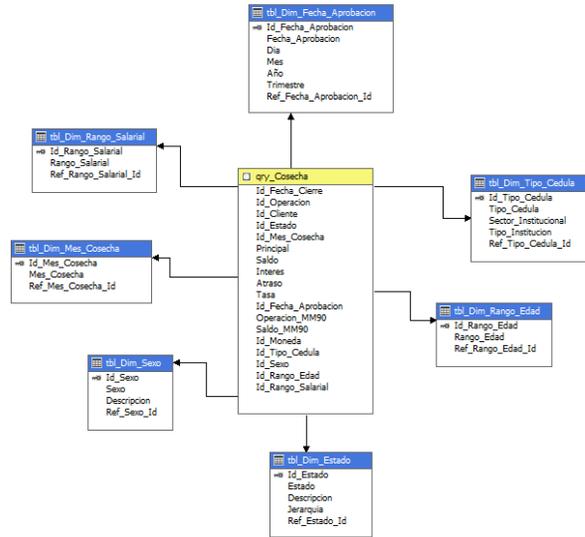


Fig. 5. Diagrama físico del cubo cbo_Cosecha.

Las matrices de transición permiten medir los rodamientos en una cartera. Estos corresponden a las operaciones que pasan de un estado a otro, o de una categoría a otra. El cubo que permite calcular esto es el cbo_MT, este tiene un esquema en estrella y se puede ver en la Fig. 6. Cuenta con las dimensiones: estado de la operación (Dim_Estado), categoría de riesgo (Dim_Categoria_Riesgo), sexo del cliente (Dim_Sexo), moneda de la operación (Dim_Moneda), provincia del cliente cuando entró a cartera (Dim_Provincia), sector económico al que pertenece el cliente (Dim_Seccion), rango de edad del cliente (Dim_Rango_Edad), tipo de cédula del cliente (Dim_Tipo_Cedula) y rango salarial del cliente (Dim_Rango_Salarial). Además, cuenta con las medidas: cantidad de operaciones (Operacion) y el monto del saldo (Saldo).

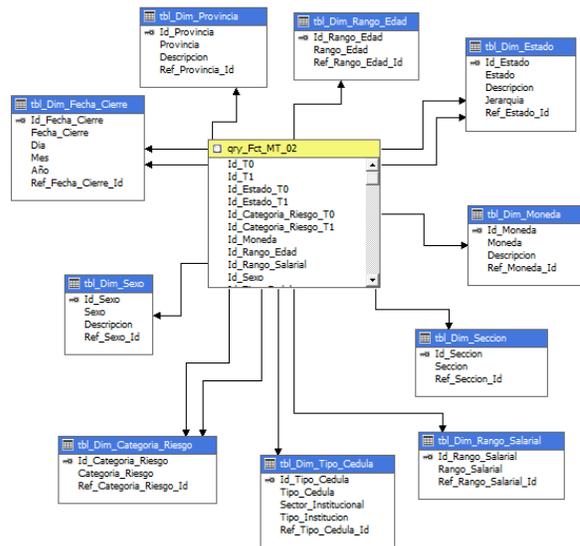


Fig. 6. Diagrama físico del cubo cbo_MT.

Las tasas de recuperación y las probabilidades de impago son complementarias, por esa razón ambas pueden calcularse en un mismo cubo. Las tasas de recuperación se refieren a cuántas operaciones se recuperaron de mora mayor a 90 a mora menor o igual a 90 en un plazo de 12 meses. Las probabilidades de incumplimiento se refieren a cuántas operaciones se deterioraron de mora menor o igual a 90 a mora mayor a 90 en un plazo de 12 meses. Ambos indicadores se miden con el cubo cbo_TRPD y su estructura se puede observar en la Fig. 7, esta corresponde a un esquema estrella, y cuenta con las dimensiones: fecha de cierre (Fecha_Cierre), moneda de la operación (Dim_Moneda), provincia del cliente cuando entró a cartera (Dim_Provincia), rango de edad del cliente (Dim_Rango_Edad), rango salarial del cliente (Dim_Rango_Salarial), sector económico al que pertenece el cliente (Dim_Seccion), sexo del cliente (Dim_Sexo), tipo de cédula del cliente (Dim_Tipo_Cedula) y estado de la operación (Dim_Estado). Asimismo, cuenta con las medidas: cantidad de operaciones deterioradas (Es_Deteriorada), cantidad de operaciones recuperadas (Es_Recuperada), monto del saldo deteriorado (Saldo_Deteriorado), monto del saldo recuperado (Saldo_Recuperado), cantidad de operaciones en mora menor o igual a 90 (Es_Normal), cantidad de operaciones en mora mayor a 90 (Es_Morosa), saldo en mora menor o igual a 90 (Saldo_Normal), saldo en mora mayor a 90 (Saldo_Deteriorado), cantidad de operaciones (Operacion) y monto del saldo (Saldo).

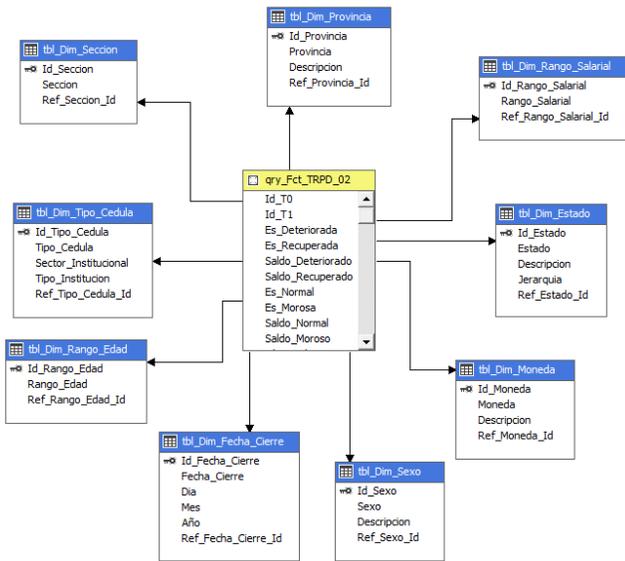


Fig. 7. Diagrama físico del cubo cbo_TRPD.

III. RESULTADOS

La Fig. 8 muestra el comportamiento de la mora mayor a 90 días en el tiempo, según la moneda. Se puede observar que, para la cartera estudiada, en el caso de colones presenta un comportamiento estable y tendiente a la baja con una pequeña

pendiente, mientras que en el caso de los dólares presenta una tasa mayor y tendiente al alza con una pendiente más grande.

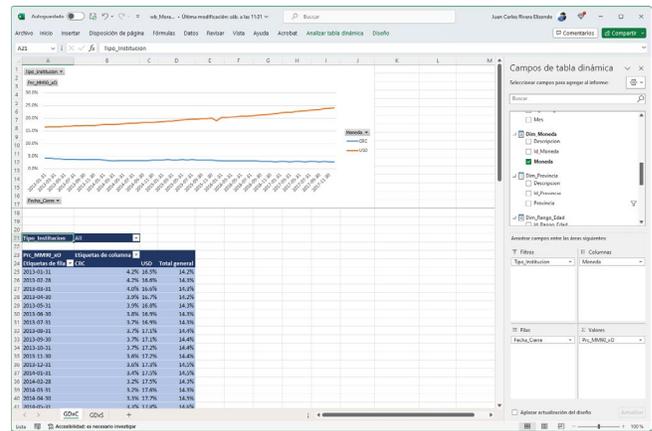


Fig. 8. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según la moneda. Con referencia a la cantidad de operaciones.

En la Fig. 9 se puede apreciar el mismo análisis según el sexo. Hombres y mujeres presentan el mismo comportamiento creciente en el tiempo, aunque la mora de los hombres es mayor a la de las mujeres. Los sexos no definidos presentan un comportamiento descendente en el tiempo.

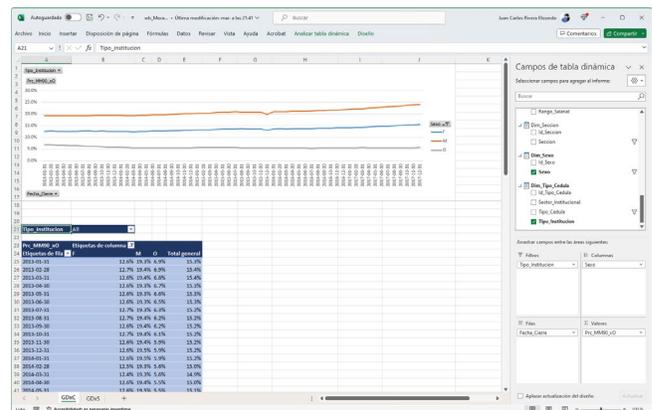


Fig. 9. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el sexo. Con referencia a la cantidad de operaciones.

La Fig. 10 muestra la mora mayor a 90 en el tiempo según el rango de edad del cliente. Todas las curvas presentan un comportamiento relativamente constante en el tiempo. Las moras más altas se encuentran para los clientes de mayor edad, y las más bajas para los clientes de menor edad.

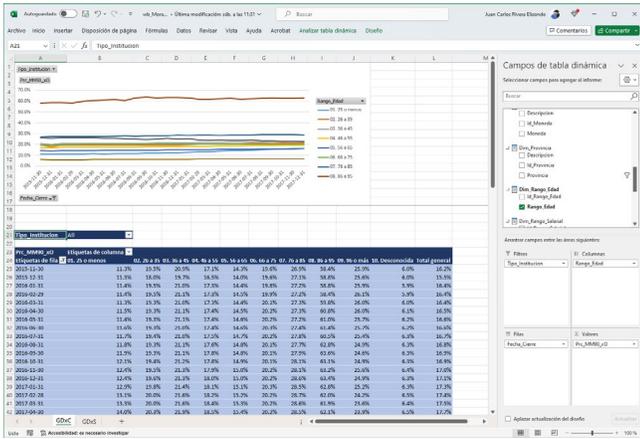


Fig. 10. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el rango de edad. Con referencia a la cantidad de operaciones.

Al analizar el comportamiento de la mora mayor a 90 en el tiempo por provincia, se puede observar un comportamiento ascendente para la provincia P00, pero relativamente constante para las demás. P04 y P06 fueron excluidas por presentar un comportamiento anormal. Cabe destacar también que la provincia P00 sobrepasa el 15 % de mora, mientras que las otras están por debajo del 8 %.

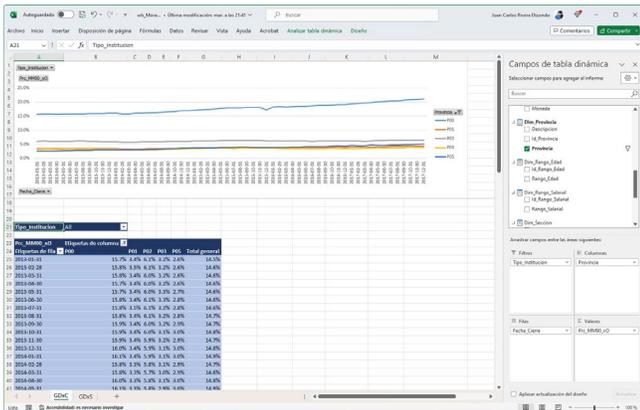


Fig. 11. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según la provincia. Con referencia a la cantidad de operaciones.

En la Fig. 12 se muestra la mora mayor a 90 en el tiempo según el rango salarial del cliente. Para las tres curvas con mayor mora se observa un comportamiento ascendente en el tiempo, mientras que para las demás se observa un comportamiento relativamente constante. Se puede notar también que las curvas de mayor morosidad se corresponden con los rangos de menor ingreso salarial, mientras que las de menor morosidad se corresponden con los rangos de mayor ingreso salarial.

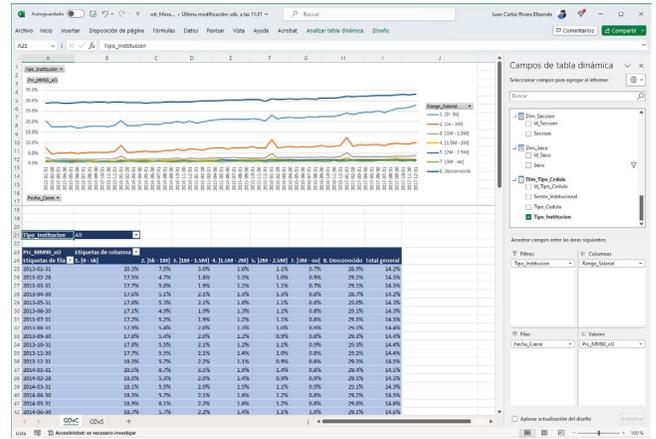


Fig. 12. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el rango salarial. Con referencia a la cantidad de operaciones.

Una variable que tiene mucha importancia para los analistas de riesgo es el sector económico. Este puede verse con la clasificación en secciones según el CIU (Clasificación Industrial Internacional Uniforme). En la Fig. 13 se puede observar el comportamiento de la mora mayor a 90 en el tiempo de cuatro secciones muy importantes para la economía. Se puede observar que, de las cuatro, la categoría Agricultura, Ganadería, Silvicultura y Pesca tiene la mayor mora, mientras que Enseñanza es la que tiene la menor. De las cuatro, tres de ellas presentan un comportamiento creciente en el tiempo, mientras que Enseñanza presenta un comportamiento relativamente constante.

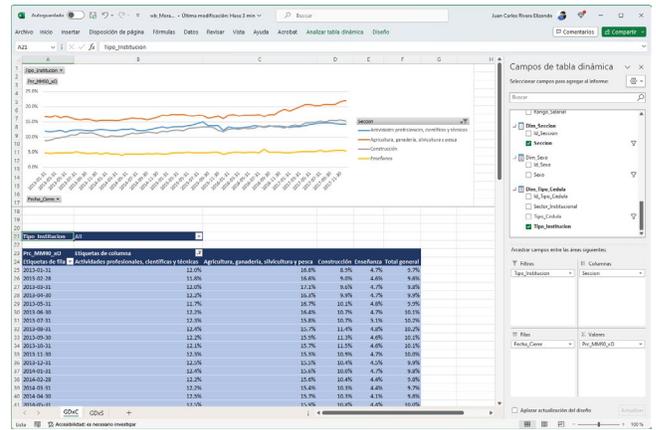


Fig. 13. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el sector económico. Con referencia a la cantidad de operaciones.

Al clasificar la cartera por tipo de institución se puede observar en la Fig. 14 que las organizaciones son las que presentan la mayor morosidad, mientras que las instituciones presentaban la menor morosidad hasta julio del 2014, después de eso ronda la morosidad de las personas físicas. Todas presentan un crecimiento en el tiempo en el último año.

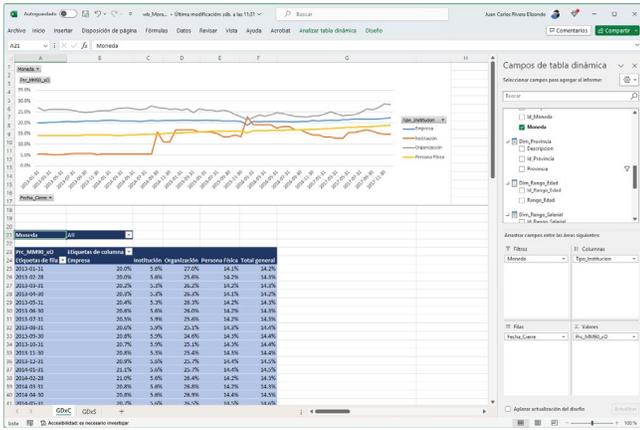


Fig. 14. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el tipo de institución. Con referencia a la cantidad de operaciones.

El comportamiento de la mora mayor a 90 en el tiempo es muy similar entre tipos de cédula. Las personas físicas extranjeras presentan el mayor nivel de morosidad, mientras que las físicas nacionales presentan el menor nivel de morosidad. El nivel de morosidad de las jurídicas privadas está entre las dos anteriores, y todas son crecientes en el tiempo. Esto puede observarse en la Fig. 15.

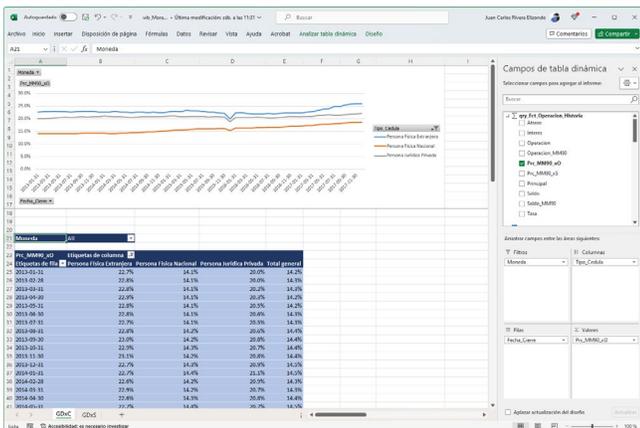


Fig. 15. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el tipo de cédula. Con referencia a la cantidad de operaciones.

La morosidad también puede analizarse en función del saldo en lugar de la cantidad de operaciones. Esto se muestra en la Fig. 16. Puede observarse que el comportamiento puede cambiar al hacerlo por saldo, tal y como ocurre para las personas físicas extranjeras y las personas jurídicas privadas, que la morosidad se vuelve decreciente en los últimos dos años.

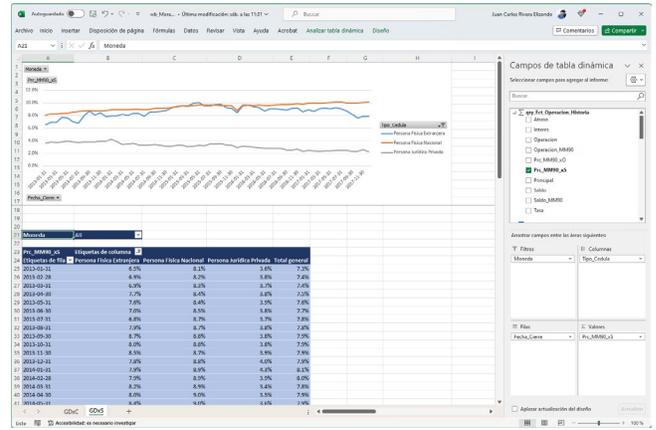


Fig. 16. Cubo cbo_Mora. Comportamiento de la mora mayor a 90 en el tiempo según el tipo de cédula. Con referencia al saldo de las operaciones.

Las cosechas se usan para comparar la mora mayor a 90, de operaciones que se aprobaron en diferentes trimestres. Todas se comparan desde el mes cero de cosecha. El mes cero corresponde al mismo mes en que se aprobó la operación. Esto permite distinguir cuál cartera se está comportando mejor. Cuando se analizan cosechas se suele considerar solamente la cartera activa, es decir, se excluyen aquellas operaciones que se encuentran en estado incobrable (I) o adjudicada (X). En algunas carteras las readevaluaciones impiden que la curva inicie en 0 %.

La Fig. 17 muestra que, en el caso de las personas de sexo masculino, la cartera que presenta la mejor morosidad es la del 2013-04, mientras que la que presenta la peor es la del 2013-02.

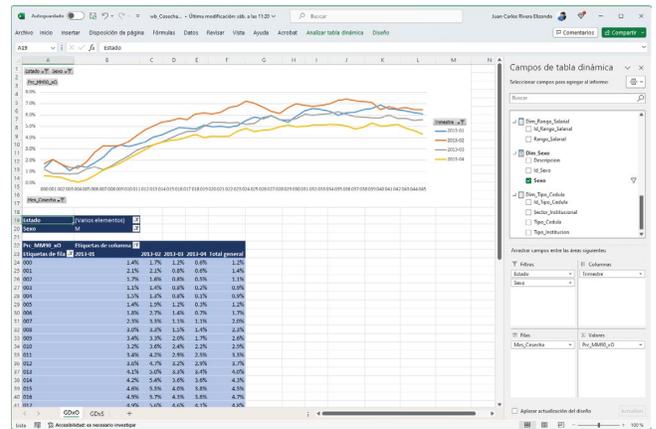


Fig. 17. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes de sexo masculino. En cuanto a la cantidad de operaciones.

Para el caso del sexo femenino, la cartera que presenta la mejor morosidad es la del 2013-04 y la que presenta la peor morosidad es la del 2013-02. Esto se puede ver en la Fig. 18.

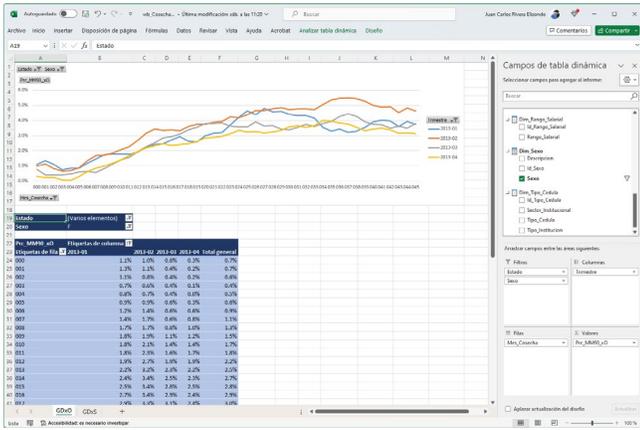


Fig. 18. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes de sexo femenino. Con referencia a la cantidad de operaciones.

Al tomar los clientes con edad menor o igual a 25 años y aquellos con la edad desconocida se puede observar que la cartera de mayor morosidad es la 2013-02 y la de menor morosidad es la del 2013-04, pero después del mes 32 es superada en rendimiento por la del 2013-01. Esto puede verse en la Fig. 19.

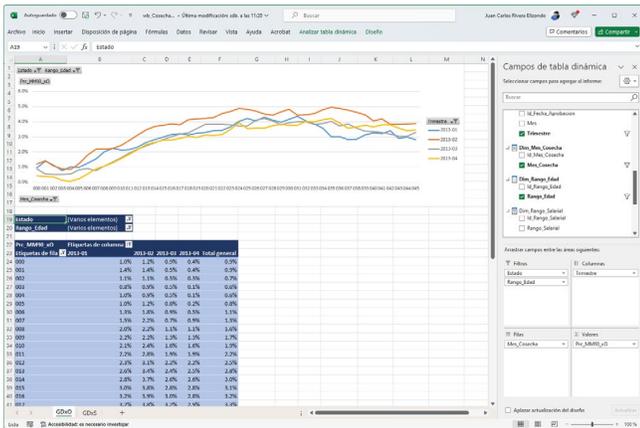


Fig. 19. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes con 25 años o menos y clientes con edad desconocida. Con referencia a la cantidad de operaciones.

Si el mismo ejercicio se hace con los clientes entre 26 y 35 años y los clientes con edad desconocida, se obtiene que la cartera con mayor morosidad es la del 2013-02, y la de menor morosidad es la del 2013-04. Además, se observa una importante caída en la morosidad para las cuatro curvas a partir del mes 33. Todo esto puede observarse en la Fig. 20.

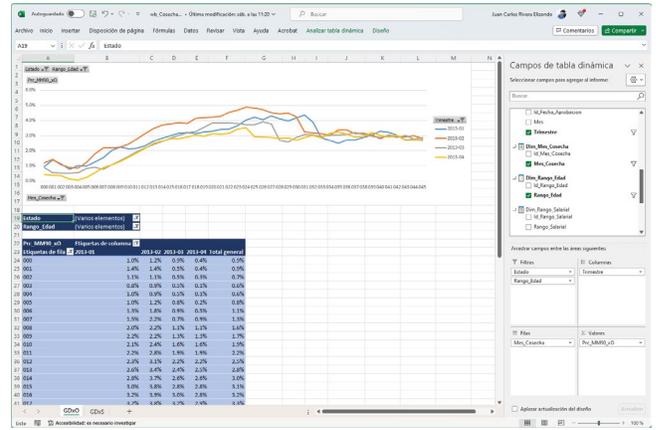


Fig. 20. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes entre 26 y 35, y clientes con edad desconocida. Con referencia a la cantidad de operaciones.

Repitiendo el ejercicio, esta vez para los clientes que ganan quinientos mil colonos o menos, se obtiene que la mayor morosidad la presenta la cartera 2013-01 y la menor morosidad en los últimos 6 meses la presenta la cartera del 2013-02. Esto puede observarse en la figura Fig. 21.

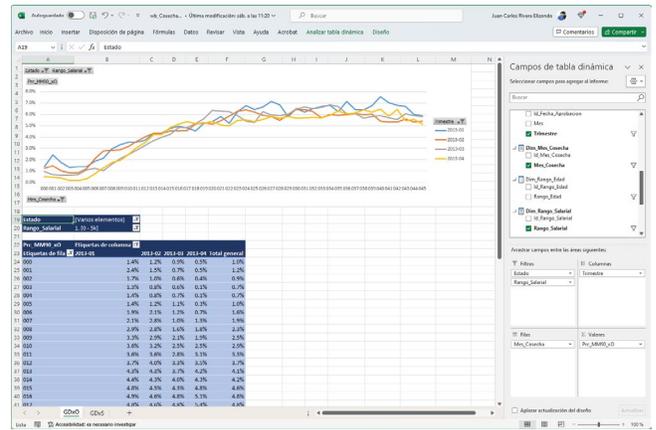


Fig. 21. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que ganan quinientos mil colonos o menos. Con referencia a la cantidad de operaciones.

Para los clientes que ganan entre quinientos mil y 1 millón de colonos, la peor morosidad la presenta la cartera del 2013-02 y la mejor morosidad la presenta la cartera del 2013-04. Esto se puede ver en la Fig. 22.

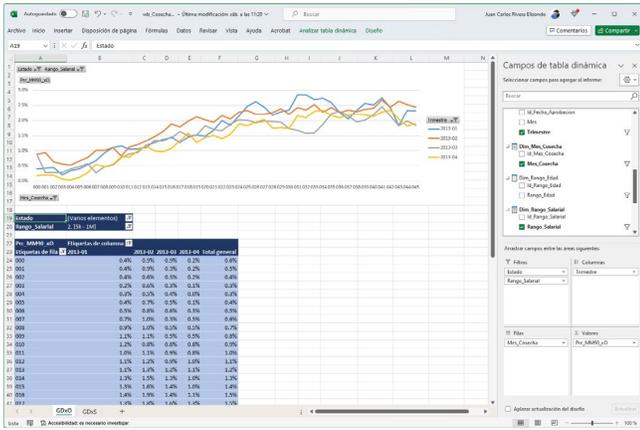


Fig. 22. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que ganan entre quinientos mil y 1 millón de colones. Con referencia a la cantidad de operaciones.

Al hacer el análisis tomando en cuenta solamente las personas físicas, se obtiene que la peor morosidad es la de la cartera del 2013-02 y la mejor morosidad es la de la cartera del 2013-04. Las demás se encuentran entre ellas dos. Véase Fig. 23.

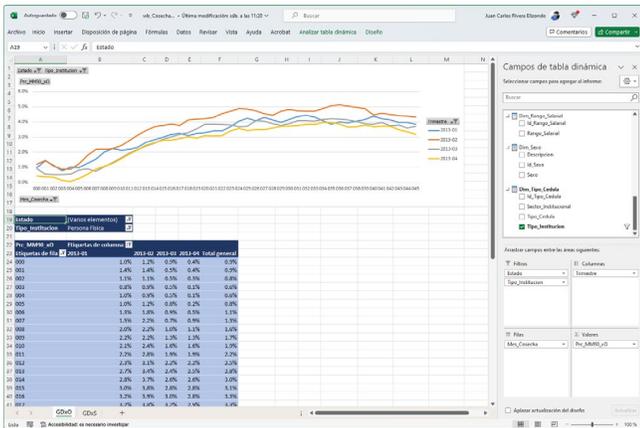


Fig. 23. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que corresponden a personas físicas. Con referencia a la cantidad de operaciones.

En la Fig. 24 se puede observar que para el caso de las empresas el peor comportamiento de morosidad lo presenta la cartera del 2013-02 y el mejor lo presenta la cartera del 2013-01.

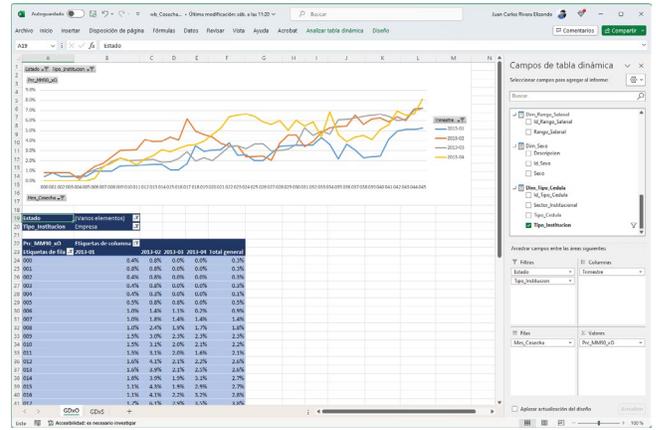


Fig. 24. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que corresponde a empresas. Con referencia a la cantidad de operaciones.

El comportamiento de las personas físicas extranjeras se puede ver en la Fig. 25. La peor cosecha la presenta la cartera del 2013-02 y la mejor la presenta la cartera del 2013-01.

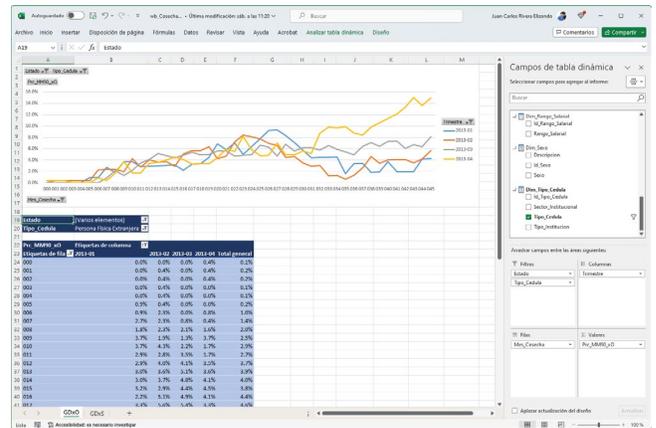


Fig. 25. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que corresponden a personas físicas extranjeras. Con referencia a la cantidad de operaciones.

Para las personas físicas nacionales, la peor cosecha corresponde a la cartera del 2013-02, y la mejor cosecha corresponde a la cartera del 2013-04. Esto puede observarse en la Fig. 26.

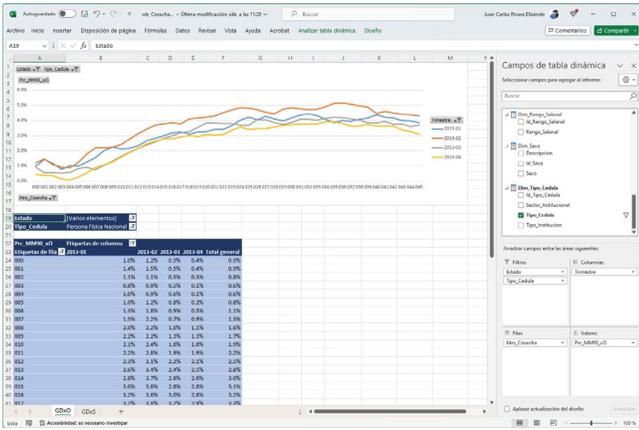


Fig. 26. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que corresponden a personas físicas nacionales. Con referencia a la cantidad de operaciones.

Las cosechas pueden analizarse también desde el punto de vista del saldo en lugar de la cantidad de operaciones. Esto puede verse en la Fig. 27. Puede notarse que el comportamiento puede cambiar al verlo desde el saldo. En este caso la cartera del 2013-01 pasa a ser la peor cosecha y la cartera del 2013-04 sigue siendo la mejor cosecha.

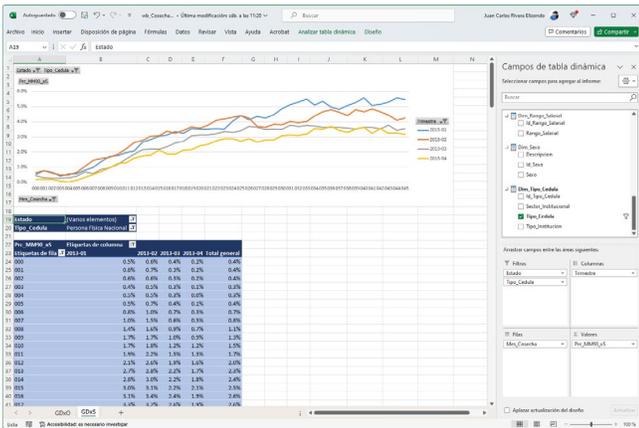


Fig. 27. Cubo cbo_Cosecha. Comportamiento de las cosechas de las operaciones nacidas en cada trimestre del 2013 para los clientes que corresponden a personas físicas nacionales. Con referencia al saldo de las operaciones.

En la Fig. 28 se puede ver la transición de operaciones entre categorías de riesgo. La diagonal marcada en color amarillo corresponde a las operaciones que se mantienen en la misma categoría. Por debajo de la diagonal están las operaciones que mejoraron de categoría y por encima de la diagonal están las operaciones que deterioraron su categoría. Las operaciones en la columna Unknown corresponden a las operaciones canceladas en el plazo de transición, que en este caso es de un mes.

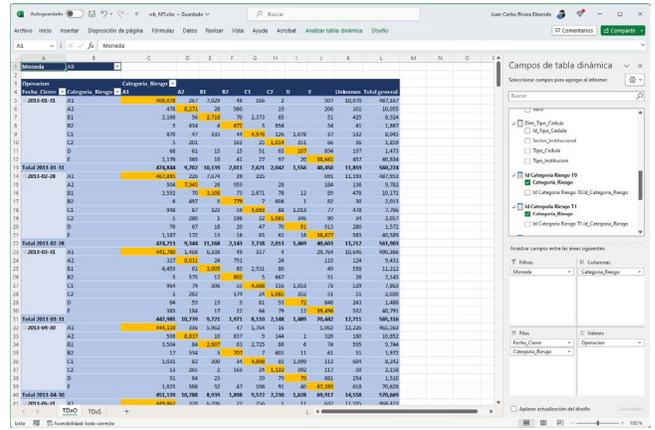


Fig. 28. Cubo cbo_MT. Matriz de transición de las operaciones, entre categorías de riesgo, en el intervalo de un mes.

La Fig. 29 muestra la transición de operaciones entre estados de operación, y la Fig. 30 muestra los mismos, pero en este caso corresponde a la transición de los saldos de las operaciones.

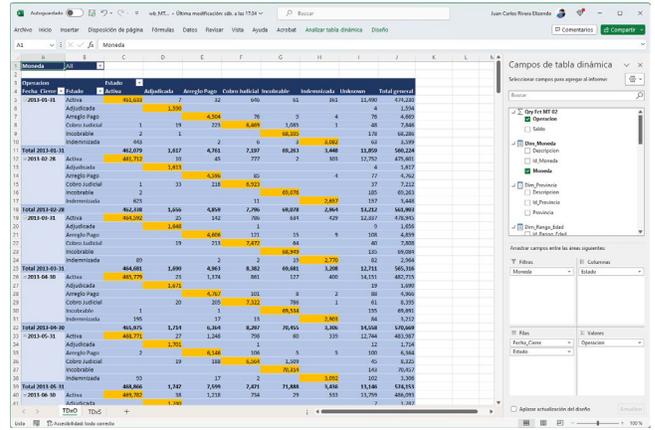


Fig. 29. Cubo cbo_MT. Matriz de transición de las operaciones, entre estados, en el intervalo de un mes.

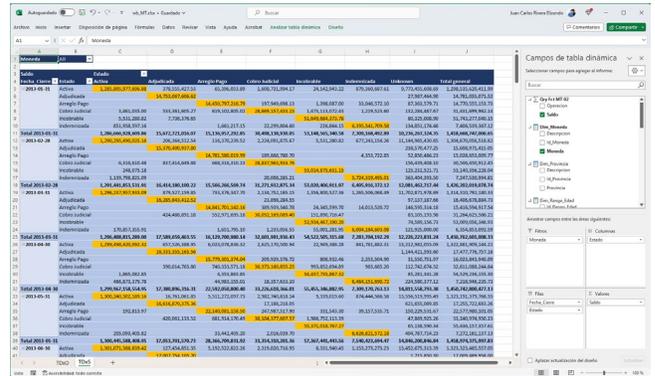


Fig. 30. Cubo cbo_MT. Matriz de transición de los saldos de las operaciones, entre estados, en el intervalo de un mes.

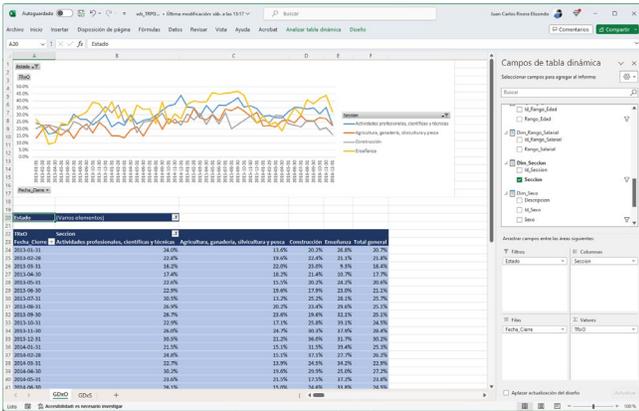


Fig. 35. Cubo cbo_TRPD. Comportamiento de las tasas de recuperación en el tiempo según el sector económico. Con referencia a la cantidad de operaciones.

Igualmente, al clasificar las tasas de recuperación por provincia se observa un comportamiento muy fluctuante y volátil en el tiempo, lo que no permite definir desde el gráfico cuál de ellas tiene el mejor y el peor comportamiento. Esto se puede apreciar en la Fig. 36.

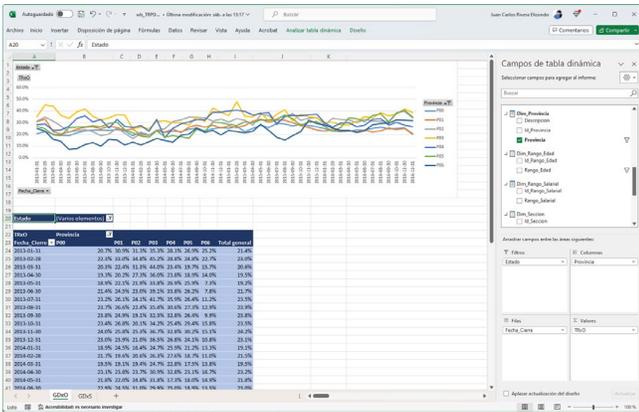


Fig. 36. Cubo cbo_TRPD. Comportamiento de las tasas de recuperación en el tiempo según la provincia. Con referencia a la cantidad de operaciones.

En la Fig. 37 se hace una comparación entre las tasas de recuperación de las personas físicas y las empresas. Se puede notar que ambas presentan un comportamiento muy fluctuante, pero se destaca que las tasas de las personas físicas son mejores que las de las empresas.

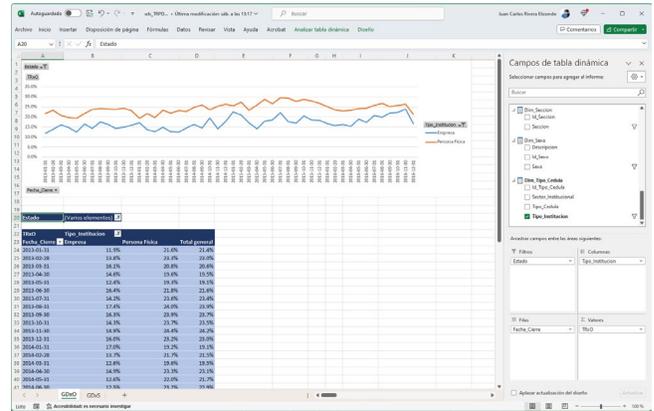


Fig. 37. Cubo cbo_TRPD. Comportamiento de las tasas de recuperación en el tiempo según el tipo de institución. Con referencia a la cantidad de operaciones.

Si se hace la clasificación por el tipo de cédula. Las tasas de recuperación de las operaciones de las personas físicas nacionales predominan sobre las tasas de las personas físicas extranjeras y las personas jurídicas privadas. Tal y como se puede observar en la Fig. 38.

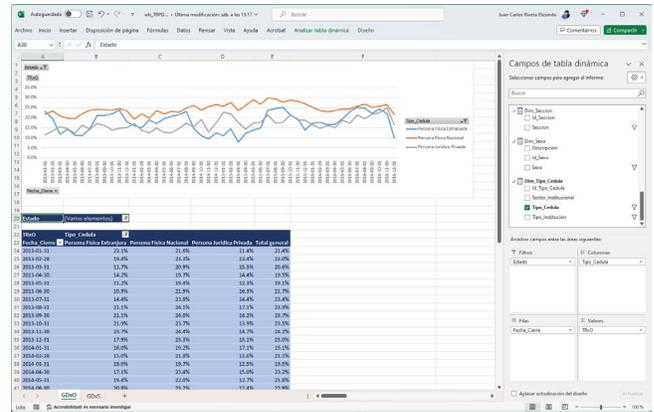


Fig. 38. Cubo cbo_TRPD. Comportamiento de las tasas de recuperación en el tiempo según el tipo de cédula. Con referencia a la cantidad de operaciones.

Es crucial señalar que este análisis puede replicarse empleando el saldo de las operaciones en vez de la cantidad. El comportamiento puede ser diferente, tal y como se muestra en la Fig. 39 para la clasificación por tipo de cédula.

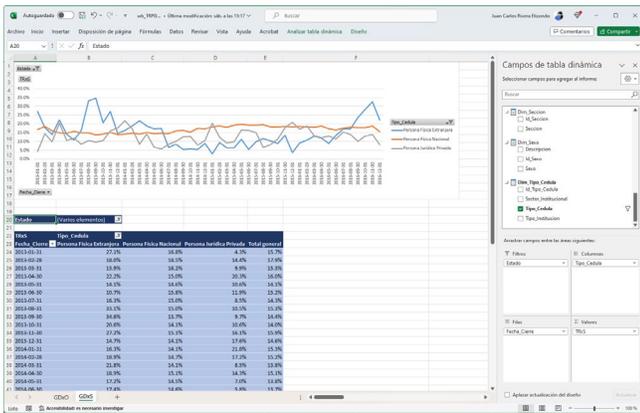


Fig. 39. Cubo cbo_TRPD. Comportamiento de las tasas de recuperación en el tiempo según el tipo de cédula. Con referencia al saldo de las operaciones.

Otro indicador muy importante para el analista de riesgo es la probabilidad de impago, también conocida como probabilidad de incumplimiento. Las siglas PD vienen del inglés Probability of Default. Este indica que tan probable es que el cliente incumpla con el pago del préstamo según el comportamiento de pago que ha tenido en el tiempo.

En la Fig. 40 se observa que la probabilidad de impago en colones es mayor a la probabilidad de impago en dólares, además que presenta un comportamiento creciente en el tiempo, mientras en dólares se mantiene relativamente constante.

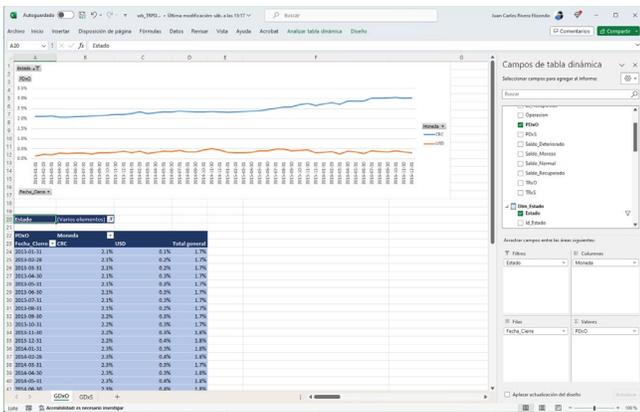


Fig. 40. Cubo cbo_TRPD. Comportamiento de las probabilidades de impago en el tiempo según la moneda. Con referencia a la cantidad de operaciones.

Esto mismo se puede analizar usando los saldos de las operaciones en lugar de la cantidad de las operaciones, tal y como se puede observar en la Fig. 41. El comportamiento mostrado puede ser diferente. En este caso se observa para ambas un comportamiento decreciente en el tiempo y además hubo periodos en los que la probabilidad de impago de dólares superó a la de colones.

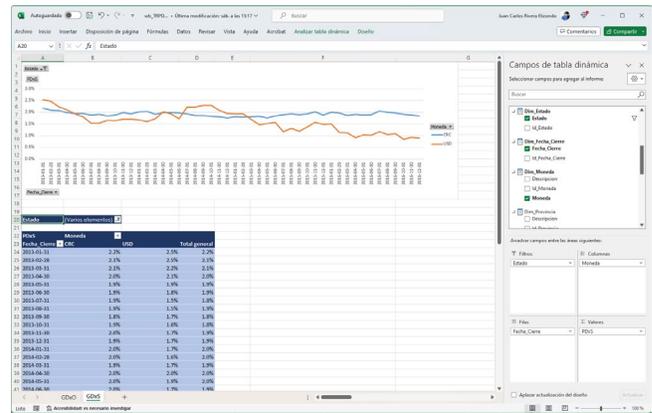


Fig. 41. Cubo cbo_TRPD. Comportamiento de las probabilidades de impago en el tiempo según la moneda. Con referencia al saldo de las operaciones.

IV. DISCUSIÓN

Para la cartera estudiada, las observaciones realizadas en la Fig. 8 muestran un mejor comportamiento de mora mayor a 90 en las operaciones en colones que en las operaciones en dólares. Mientras que la Fig. 31 muestra mejores tasas de recuperación en dólares que en colones. La Fig. 40 presenta una probabilidad de impago mayor en colones que en dólares.

En la Fig. 9 se puede apreciar una mora mayor a 90 mejor para las mujeres que para los hombres. Mientras que en cuanto a la recuperación en la Fig. 32 se puede observar que la recuperación entre hombres y mujeres es muy similar.

En la Fig. 10 se observa una mora mayor a 90, menor para los clientes de menor edad que para los clientes de mayor edad. Por otra parte, la Fig. 33 muestra tasas de recuperación muy similares, entre rangos de edad.

La Fig. 11 muestra una mora mayor a 90, mayor en la provincia P00 en comparación con las demás provincias. Al observar las tasas de recuperación en la Fig. 36 se ven tasas muy similares, fluctuantes y volátiles.

La Fig. 12 muestra mayor morosidad a menor ingreso salarial y viceversa. Mientras que la Fig. 34 muestra recuperaciones similares entre los rangos salariales, pero la tasa de los clientes de menor ingreso destaca por ser la peor.

En la Fig. 13 se puede apreciar menor morosidad en las operaciones prestadas al sector Enseñanza y mayor morosidad en las prestadas al sector Agricultura, Ganadería, Silvicultura y Pesca. La Fig. 35 muestra que las tasas de recuperación entre ellas son muy similares, fluctuantes y volátiles.

Se puede apreciar en la Fig. 14 que la mora mayor a 90 de las personas físicas es la menor de todas. Igualmente, la Fig. 37 muestra tasas de recuperación mayores para las personas físicas. Lo cual es positivo en ambos casos.

La Fig. 15 muestra menor morosidad para las personas físicas nacionales y la Fig. 38 muestra para estas mismas personas, tasas de recuperación mayores.

V. CONCLUSIONES

- Se logró diseñar e implementar un Data Warehouse para la gestión de riesgo de crédito, usando como ejemplo la cartera de crédito aportada por una institución anónima del mercado costarricense.
- Se diseñaron e implementaron ETLs que permiten cargar el Data Warehouse desde archivos de texto.
- Se diseñaron e implementaron cuatro cubos OLAP que permiten medir cinco indicadores de riesgo de crédito. Estos indicadores son: la morosidad, las cosechas, los rodamientos, las tasas de recuperación y las probabilidades de impago.
- La cartera estudiada tiene la menor morosidad en colones, pero las mejores tasas de recuperación y las probabilidades de impago más bajas las tiene en dólares.
- La cartera estudiada tiene menor morosidad en mujeres que en hombres, y tasas de recuperación muy similares entre ambos sexos.
- La cartera estudiada tiene menor morosidad a menor edad del cliente y viceversa. Las tasas de recuperación son similares entre rangos de edad.
- La cartera estudiada tiene la mayor morosidad en la provincia P00 y las tasas de recuperación son muy similares entre todas las provincias.
- La cartera estudiada tiene mayor morosidad, a menor ingreso salarial y viceversa. Las tasas de recuperación son similares entre los rangos salariales, excepto para el menor rango salarial, que sí destaca por ser peor que los demás.
- Entre los cuatro sectores económicos analizados en la cartera estudiada destaca Enseñanza por tener la menor morosidad, y el sector Agricultura, Ganadería, Silvicultura y Pesca por tener la peor. Las tasas de recuperación son muy similares entre los cuatro sectores.
- En la cartera estudiada, las personas físicas presentan las menores morosidades y las mayores tasas de recuperación.
- En la cartera estudiada, entre las personas físicas, son las físicas nacionales las que presentan la menor morosidad y las mayores tasas de recuperación.

REFERENCIAS

- [1] Banco Central de Nicaragua (Mayo, 2020). Estabilidad Financiera: Riesgo de Crédito y Operacional. Recuperado desde https://www.bcn.gob.ni/sites/default/files/cursos_econ_finan/2020/estabilidad%20financiera/Sesion%205_Estabilidad%20financiera.pdf
- [2] Evaluando Software (2022). Cubos OLAP de información para la toma de decisiones. Recuperado desde <https://www.evaluandosoftware.com/cubos-olap-informacion-la-toma-decisiones/>
- [3] García, P. (s.f). El Riesgo de Crédito. Recuperado desde <https://slideplayer.es/slide/1073256/>

- [4] Mesalles J. Luis (2008, Agosto 22). Décimo Cuarto Informe Estado de la Nación en Desarrollo Humano Sostenible. Análisis de la Evolución del Crédito en Costa Rica Durante el 2007, y sus Implicaciones sobre la Economía. Recuperado desde <https://repositorio.cepal.org/handle/11362/25849>
- [5] Microsoft Learn (2023). Extracción, transformación y carga de datos (ETL). Recuperado desde <https://learn.microsoft.com/es-es/azure/architecture/data-guide/relational-data/etl>
- [6] Orfin (2021). ¿Ha aumentado la morosidad bancaria?. Recuperado desde <https://orfin.es/facts/ha-aumentado-la-morosidad-bancaria/>
- [7] Powerdata.es (2023). Data Warehouse: todo lo que necesitas saber sobre almacenamiento de datos. Recuperado desde <https://www.powerdata.es/datawarehouse#:~:text=Un%20Data%20Warehouse%20es%20un,recuperar%20y%20f%C3%A1cil%20de%20administrar.>
- [8] SAS (2023). Riesgo del crédito qué es y porqué importa. Recuperado desde https://www.sas.com/es_pe/insights/risk-management/credit-risk-management.html
- [9] SUGEF. 2021 Bottom Up Stress Test (2022, Abril). Recuperado desde https://www.sugef.fi.cr/informacion_relevante/manuales/manual_sicveca/Prueba_de_%20Tension-BUST/2021%20BUST%20Guia%20Metodologica.pdf



Juan Carlos Rivera Elizondo, licenciado en ingeniería electrónica con once años de experiencia en riesgo de crédito, tanto en el sector público como en el sector privado. Aspirante al grado de maestría en tecnología de bases de datos por la Universidad Cenfofec.



José Alberto Cabezas Jaikel, máster en tecnología de bases de datos de la Universidad Cenfofec, con más de veinte años de experiencia en el desarrollo de soluciones analíticas en plataformas de Microsoft y Oracle.