

crediclub

CÁLCULO DE DISPERSIÓN GRUPAL DE ASESORES Y SUCURSALES

Matilde Simental Herrera

Trainee Datos. Programa Be Bold

Julio 2025

An abstract graphic consisting of numerous thin, overlapping, wavy lines that create a sense of movement and depth. The lines are light gray and flow from the bottom left towards the top right, with some lines curving back to form a loop-like structure on the right side.

Índice

1. Descripción del proyecto	2
2. Metodología para medida de dispersión	2
2.1. Área de polígono cóncavo	2
2.2. Área de polígono convexo	2
2.3. Desviación de grupos a un centroide	3
3. Aplicación a Asesores	4
3.1. Fuentes y descripción de datos	4
3.2. Resultados y Análisis	5
4. Aplicación a Sucursales	6
4.1. Resultados y Análisis	6
4.2. Conclusión	9

1. Descripción del proyecto

En este proyecto se implementó una metodología para analizar y comparar las áreas de cobertura o dispersión de los grupos asignados a los asesores de CrediClub, con el objetivo de identificar si existía una relación entre dicha dispersión y la productividad de los asesores. Posteriormente, esta misma metodología fue adaptada para evaluar la dispersión de los grupos a nivel sucursal y explorar su posible relación con el desempeño de cada sucursal definido por los indicadores Cartera, % de Atraso 4 a 30 días, % de Liquidación, y % de Renovación.

2. Metodología para medida de dispersión

Se propusieron tres maneras de obtener una medida comparable de la dispersión de los asesores:

- Área de polígono cóncavo.
- Área de polígono convexo.
- Desviación estándar de las distancias de los grupos a un centroide.

2.1. Área de polígono cóncavo

Esta opción consistía en identificar las coordenadas geográficas de los grupos asignados a cada asesor y calcular el área resultante al unir los puntos más externos (outer points) en el espacio (Ver Fig. 1a). Esta unión se realizaría utilizando la distancia mínima entre dichos puntos. Sin embargo, esta alternativa no fue desarrollada a profundidad, ya que las áreas generadas tendían a ser excesivamente amplias y no representaban con precisión la cobertura real del asesor.

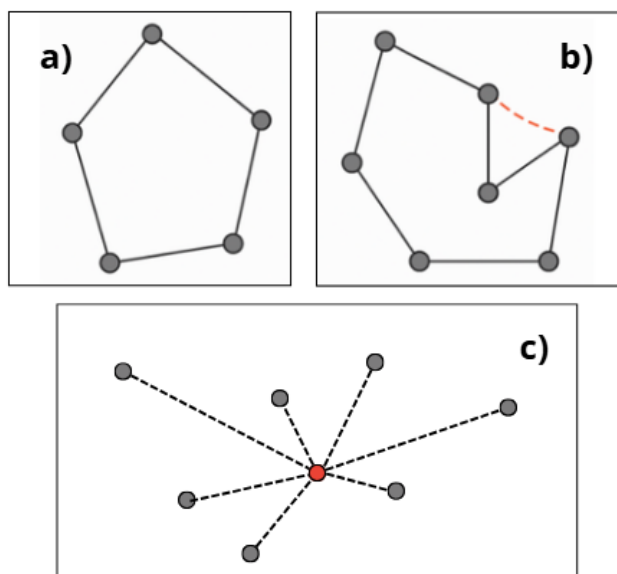


Figura 1: Visualización de las tres metodologías.

2.2. Área de polígono convexo

Este enfoque consistía en tomar las coordenadas de todos los grupos asignados a un asesor y generar un contorno que los envolviera completamente, sin dejar espacios hacia dentro, es decir, un polígono sin concavidades (Ver Fig. 1b). Este tipo de figura se conoce como *envolvente convexa* o *convex hull*.

Si bien un polígono convexo es más preciso que un polígono cóncavo al evitar irregularidades y asegurar un contorno uniforme, este enfoque presenta una limitación importante: no es fácilmente comparable entre asesores. Para construir el polígono convexo ajustado a los grupos de cada asesor, es necesario definir un parámetro de ajuste que determine qué tan ceñido debe estar el contorno a los puntos.

El problema radica en que este parámetro no puede ser el mismo para todos los casos, ya que la distribución geográfica de los grupos varía entre asesores. Por lo tanto, el parámetro de ajuste tendría que definirse de manera personalizada para cada conjunto de grupos, lo que impide establecer una base de comparación uniforme y objetiva entre todos los asesores.

Debido a esta falta de estandarización y comparabilidad, se optó por no utilizar este enfoque en el análisis final.

2.3. Desviación de grupos a un centroide

Esta metodología consiste en calcular el centroide del conjunto de grupos de cada asesor, es decir, el punto promedio de sus coordenadas geográficas (latitud y longitud). Una vez determinado este punto central, se mide qué tanto se alejan los grupos respecto a él. La desviación se obtiene a partir de la distancia individual de cada grupo al centroide, y refleja qué tan concentrados o dispersos están los grupos alrededor de su punto medio.

Se obtiene entonces la distancia de cada grupo al centroide por medio de la fórmula de Haversine (Ver ecuación 2.3), la cual permite calcular la distancia entre dos puntos en la superficie de la Tierra considerando su curvatura. Esta fórmula requiere las coordenadas geográficas (latitud y longitud) de los puntos de interés y devuelve la distancia en kilómetros.

$$a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2\left(\frac{\Delta\lambda}{2}\right)$$

$$c = 2 \cdot \arctan 2\left(\sqrt{a}, \sqrt{1-a}\right)$$

$$\text{distancia} = R \cdot c$$

Una vez calculada la distancia de cada grupo al centroide correspondiente, se obtiene la distancia promedio de los grupos respecto al centroide. A partir de estas distancias individuales, se calcula la desviación estándar (en km), la cual refleja qué tanto varían los grupos con respecto a esa distancia media.

Un valor bajo de desviación indica que los grupos se encuentran poco dispersos, mientras que un valor alto sugiere una mayor dispersión geográfica en la distribución de los grupos del asesor.

Es importante señalar que este método también presenta ciertas limitaciones. Por ejemplo, si un asesor tiene pocos grupos pero todos se encuentran alejados entre sí de manera uniforme, la desviación estándar podría seguir siendo baja, a pesar de que en la práctica exista una alta dispersión geográfica. Esto puede generar una subestimación de la verdadera cobertura del asesor.

Aun con esta debilidad, la desviación estándar resulta ser una métrica útil en términos generales. Como se puede observar en el gráfico 2, existe una relación clara entre la distancia promedio al centroide y la desviación estándar: a mayor distancia promedio, generalmente también se incrementa la dispersión. Aunque se presentan algunos outliers que no siguen esta tendencia, en general se identifica una relación aproximadamente lineal.

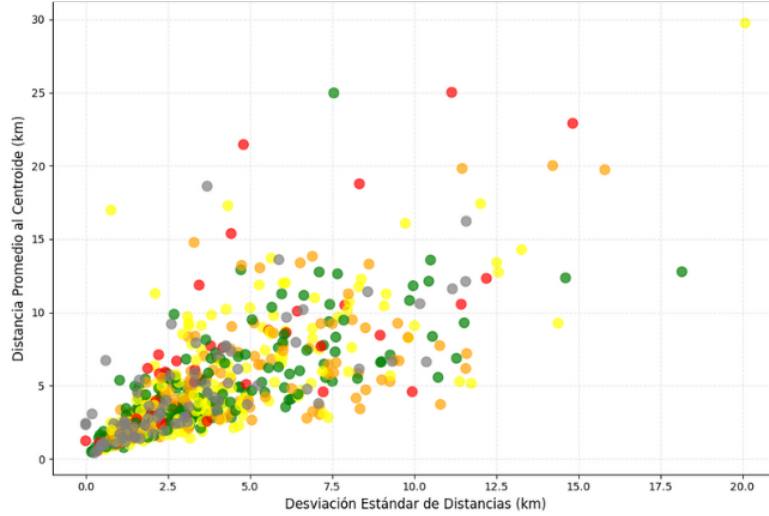


Figura 2: Relación entre la desviación estándar de las distancias y la distancia promedio de los grupos al centroide por asesor.

3. Aplicación a Asesores

El propósito del análisis es ver si la dispersión de los grupos del asesor está relacionado con su productividad.

3.1. Fuentes y descripción de datos

Se obtuvieron los asesores de CrediClub y sus grupos de la tabla `datasets.ma.carteragrupalvigente` a fecha corte 2025-05-14. El integrante del grupo del cual es la casa sede se obtuvo de la tabla `lakehouse.silver.srraorigruci` y las coordenadas de la casa sede se obtuvieron de `lakehouse.silver.srraoriprospe`. La productividad del asesor se obtuvo de `incentivos.analiticarh.calificaciones`.

También se decidió implementar la distancia promedio de la casa del asesor a sus grupos, esta distancia fue calculada al igual con la fórmula de Haversine [2.3](#).

La productividad del asesor es categórica y se divide en 4 categorías: Alta, Media Alta, Media Baja y Baja. Para compararlas de manera categórica también se categorizó la desviación de grupos de cada asesor como se muestra en la tabla

Métrica	Categorización
Distancia promedio al centroide	Baja: 0 a 4 km Media: Más de 4 a 11 km Alta: Más de 11 km
Desviación estándar al centroide	Baja: 0 a 3 km Media: Más de 3 a 8.5 km Alta: Más de 8.5 km
Distancia promedio asesor - grupos	Baja: 0 a 5 km Media: Más de 5 a 18 km Alta: Más de 18 km

Cuadro 1: Categorización de métricas de dispersión

3.2. Resultados y Análisis

El resultado final es una tabla (Tabla 3) que contiene, para cada asesor, su ID y nombre, el número de grupos asignados, la distancia promedio de los grupos al centroide (en kilómetros), la desviación estándar respecto al centroide, y la distancia promedio entre el asesor y sus grupos. Además, se incluye la categorización correspondiente de cada una de estas métricas, así como el nivel de productividad del asesor.

Asesor	NomAsesor	LatCentroide	LonCentroide	PromDistCentroide	DesvDistCentroide	Productividad_Asesor
17996	NORA GUADALUPE ARANDA GUERRERO	25.744429	-100.356041	2.552544	1.786399	Alta
18013	MARIA ISABEL GUZMAN VALLES	25.689060	-100.492050	1.946581	1.812686	Media alta
18314	CARMEN JULIA VALDEZ ARREOZOLA	25.867910	-100.234077	9.024249	3.475336	Media alta
18472	LEONOR ALEJANDRA TALAVERA VELAZQUEZ	25.520514	-103.387817	1.893380	1.090428	Alta
18804	NORMA ALICIA MEDELLIN TERAN	23.727362	-99.156212	2.314193	1.384012	Media baja

Figura 3: Enter Caption

Ya que se buscaba saber si se relaciona o no la dispersión del asesor con su productividad. Se generó una gráfica en la que se observa la proporción de asesores que tienen baja, media y alta dispersión grupal, que caen dentro de las distintas categorías de productividad, así podremos ver si en dispersiones altas predominan los asesores de productividad baja, y en dispersiones bajas predominan los que tienen alta productividad.

Chi cuadrado de independencia

Para evaluar si existe una relación significativa entre la dispersión geográfica de los grupos asignados a un asesor y su nivel de productividad, se aplicó un análisis de Chi cuadrada de independencia (χ^2). Que permite determinar si dos variables categóricas están asociadas entre sí o si son independientes. Si el valor resultante supera cierto umbral crítico, se rechaza la hipótesis nula, concluyendo que sí existe una asociación entre la dispersión y la productividad.

El valor calculado fue de $\chi = 6,8$, mientras que el valor crítico correspondiente al nivel de significancia elegido fue de $\chi = 15,5$. Dado que el valor calculado es menor que el valor crítico, esto indica que no se encontró evidencia suficiente para afirmar que existe una relación estadísticamente significativa entre la dispersión y la productividad de los asesores.

En conclusión, aunque visualmente se aprecia en la Figura 4 una tendencia en la que disminuye la proporción de asesores con alta productividad conforme aumenta la dispersión de sus grupos, y, a su vez, incrementa la proporción de asesores con baja productividad en niveles altos de dispersión, **esta relación no fue estadísticamente significativa** según el análisis realizado.

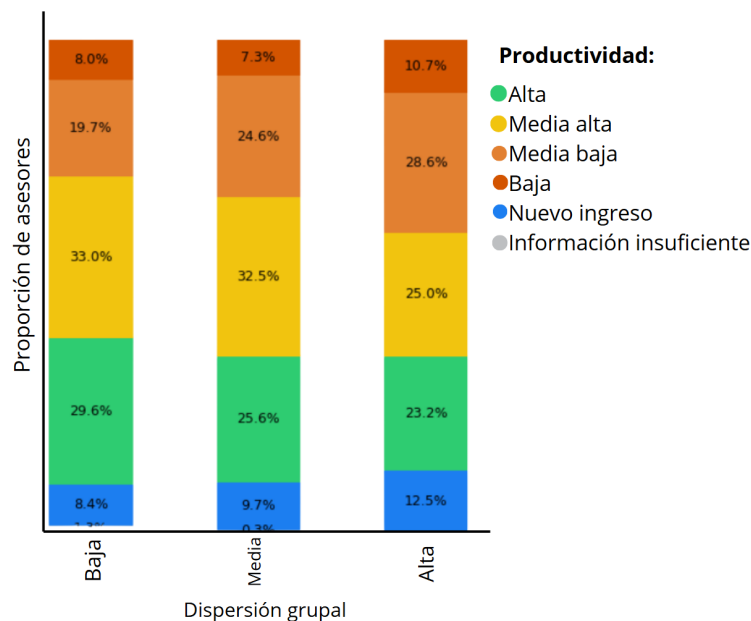


Figura 4: Enter caption

Finalmente, para facilidad de visualización, se realizó un tablero en Power BI en el que se puede buscar al asesor, ver la dispersion de sus grupos geograficamente y la categoría de dispersion asignada y la productividad.

4. Aplicación a Sucursales

Se implementó la metodología detallada en la Sección 2 en los grupos pero ahora a nivel sucursal.

Los grupos y la sucursal a la que pertenecen se extrajeron de la tabla `datasets.ma.carteragrupalvigente` a `FechaCorte` 2025-06-04. Se analizaron 173 sucursales. El objetivo es saber si existe una relación entre indicadores de la situación de las sucursales y la dispersión de sus grupos. Los indicadores a comparar con la dispersión grupal son:

- Cartera. ¹
- Porcentaje de Atraso de 4-30 días. ²
- Porcentaje de Renovación.
- Porcentaje de Liquidación.

Estos indicadores fueron extraídos del tablero KPI General MA, de la tabla `Situacion Sucursales`.

4.1. Resultados y Análisis

Se analizaron esas tablas en Python, el código utilizado se puede encontrar en bajo el **nombre**. El resultado es un listado de Sucursales, con su respectiva desviación estándar que es la medida de dispersión grupal, los indicadores, y la categorización de cada uno de estos. Esta categorización se muestra en la Tabla 2.

En el código se puede descargar la tabla final.

¹Saldo pendiente de los créditos

²Monto en atraso de 4 a 30 días

Columna	Baja	Media Baja	Media Alta	Alta
Cartera	$\leq 8.4M$	$(8.4M - 15.4M]$	$(15.4M - 22.9M]$	$> 22.9M$
% Atraso (4-30)	≤ 0.01	$(0.01 - 0.03]$	$(0.03 - 0.05]$	> 0.05
% Renovación	≤ 0.50	$(0.50 - 0.70]$	$(0.70 - 0.85]$	> 0.85
% Liquidación	≤ 0.65	$(0.65 - 0.80]$	$(0.80 - 0.95]$	> 0.95
Desv. Centroide (km)	≤ 2.98	$(2.98 - 5.00]$	$(5.00 - 7.60]$	> 7.60
Dist. Promedio Centroide (km)	≤ 4.00	$(4.00 - 6.27]$	$(6.27 - 10.00]$	> 10.00

Cuadro 2: Categorías numéricas

Análisis

En la Fig. 5 se puede observar la proporción de sucursales que cae dentro de cada categoría de dispersión, respecto a cada indicador de desempeño. Visualmente, se observa una correlación en la mayoría de los casos.

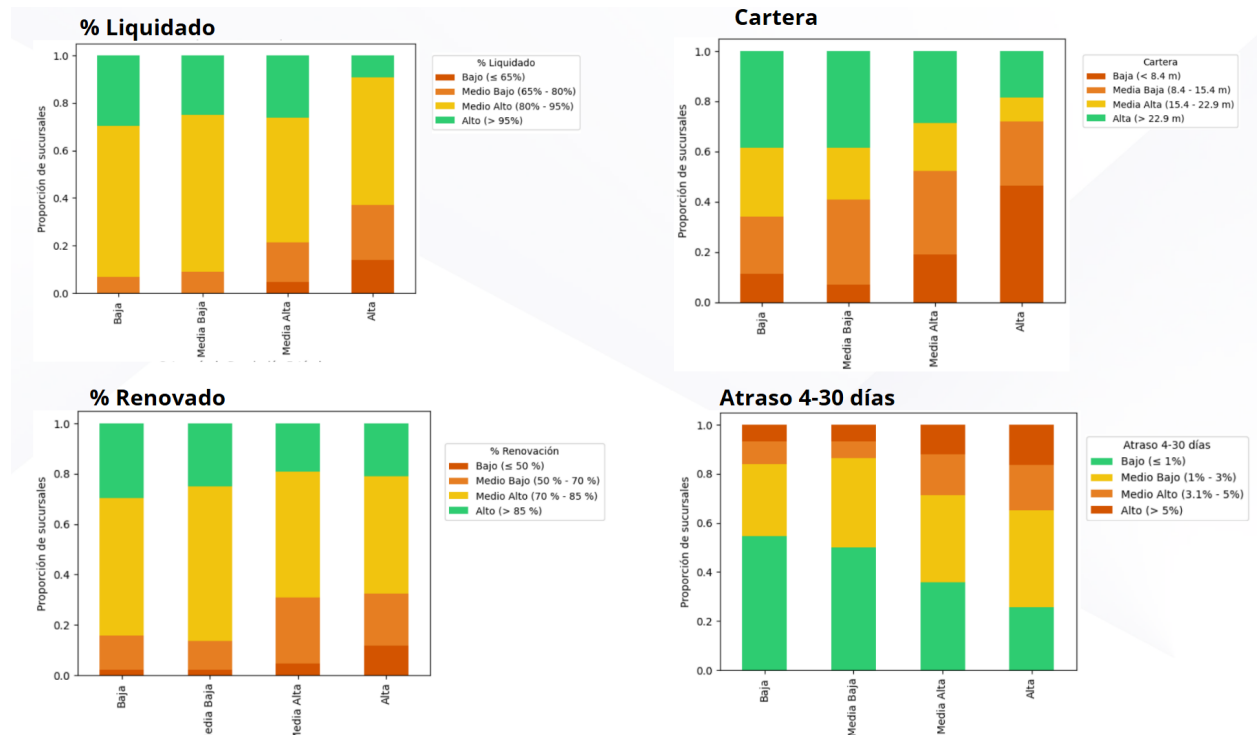


Figura 5: Indicadores de desempeño de sucursales vs dispersión de grupos.

Por ejemplo, en el caso de la **cartera** (Fig. 5a), la participación de carteras bajas es mayor en las sucursales con alta dispersión, mientras que en aquellas con baja dispersión predominan, en proporción, las carteras altas.

Para el **porcentaje de liquidación**, se identifica una disminución progresiva en la proporción de sucursales con alta liquidez a medida que aumenta la dispersión grupal, y, de forma inversa, un aumento en la proporción de sucursales con baja liquidez en niveles altos de dispersión.

En cuanto al **porcentaje de renovación**, la tendencia es menos marcada, pero se mantiene el mismo patrón general: en las sucursales con baja dispersión tiende a concentrarse la mayor proporción de sucursales con alta renovación, y en las de alta dispersión, aunque con menor diferencia, se incrementa la participación de aquellas con renovaciones bajas.

Finalmente, en el **atrasoFIG** se observa una mayor proporción de sucursales con bajo nivel de atraso en contextos de dispersión baja, y una mayor presencia de sucursales con alto nivel de atraso en aquellas con mayor dispersión grupal.

Surge ahora la pregunta, de si existe un abasto suficiente de asesores correspondiente a la dispersión y cantidad de grupos de la sucursal. A fin de evaluar esto, se establecieron las siguientes categorías de grupos y asesores por sucursal:

- **Grupos:**
 - Pocos 25 o menos
 - Medios De 26 a 42
 - Muchos Más de 43
- **Asesores:**
 - Pocos Menos de 4
 - Medios De 4 a 7
 - Muchos Más de 8

A partir de esta categorización, en el gráfico 6 se muestra la distribución proporcional de sucursales según la cantidad de grupos que manejan (pocos, medios o muchos), en función del número de asesores asignados a la sucursal.

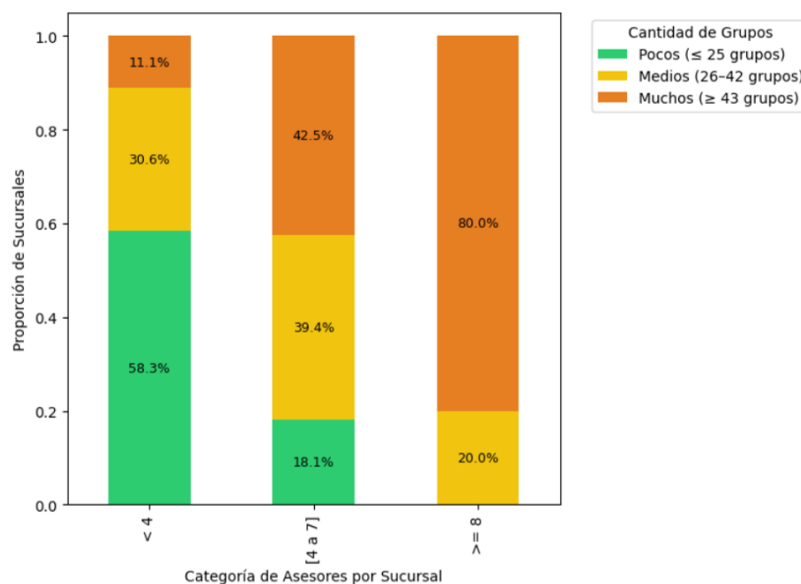


Figura 6: Asesores y cantidad de grupos por sucursal

- El 58 % de las sucursales con **menos de 4 asesores** atienden **pocos grupos**, y el 80 % de las sucursales con 8 o más asesores tienen muchos grupos, lo cual resulta razonable.
- Sin embargo, el 11 % de las sucursales con pocos asesores atienden **más de 43 grupos**, lo cual podría representar una sobrecarga.

Se identificaron las sucursales que pertenecen a la categoría de tener pocos asesores y una alta cantidad de grupos asignados, correspondiente al 11 % que se observa en el primer bin del gráfico. En estos casos, se analizaron los indicadores de desempeño así como la dispersión de sus grupos para detectar patrones relevantes (Fig. 7). Por ejemplo:

Sucursal	Asesores	Grupos	CarteraCat	AtrasoCat	%RenovacionCat	%LiquidacionCat	DesviacionCat
ENSENADA	3	43.0	Media Alta	Medio Alto	Medio Bajo	Medio Bajo	Media Alta
CUAJIMALPA	3	49.0	Media Baja	Bajo	Alto	Alto	Media Baja
MONTERREY PTE	3	54.0	Alta	Medio Alto	Medio Alto	Medio Bajo	Media Baja
MONTERREY SUR	2	55.0	Alta	Bajo	Alto	Alto	Baja

Figura 7: Enter Caption

Sucursal	Asesores	Grupos	CarteraCat	AtrasoCat	%RenovacionCat	%LiquidacionCat	DesviacionCat
MAZATLAN-MAZATLAN 2	11	61.0	Alta	Bajo	Alto	Medio Alto	Alta
APIZACO	10	61.0	Alta	Medio Bajo	Medio Alto	Medio Alto	Alta
HERMOSILLO-MIGUEL ALEMAN	8	43.0	Media Alta	Bajo	Medio Alto	Medio Bajo	Alta
SAN LUIS RIO COLORADO	7	51.0	Alta	Medio Alto	Medio Alto	Medio Bajo	Alta
MOCHIS	6	67.0	Alta	Medio Alto	Medio Alto	Medio Bajo	Alta
CUAUTLA	6	49.0	Media Baja	Medio Bajo	Medio Bajo	Bajo	Alta
VERACRUZ NORTE	6	68.0	Alta	Medio Alto	Medio Alto	Medio Bajo	Alta

Figura 8: Sucursales con muchos grupos y alta desviación.

Se observa un mejor desempeño en sucursales con baja dispersión geográfica, como es el caso de Monterrey Sur, que a pesar de contar únicamente con dos asesores, sus indicadores son todos buenos. Esto sugiere que, incluso cuando el número de grupos es elevado y la cantidad de asesores poca, una baja dispersión facilita la gestión, permitiendo atender eficazmente a los grupos con menos asesores.

Surge ahora la pregunta de si en sucursales muy dispersas (alta desviación) y con más de 43 grupos, una mayor cantidad de asesores mejora el desempeño de esta. Para responder esto, se compararon las sucursales con:

- Más de 43 grupos.
- Alta desviación de los grupos.

Estas sucursales se ordenaron de acuerdo al número de asesores que tienen, y se evaluaron cinco indicadores de desempeño: *cartera*, *atraso*, *% renovación*, *% liquidación* y *desviación centroe*.

En esta muestra, las sucursales que cuentan con un mayor número de asesores para atender una alta cantidad de grupos (más de 43) —es decir, con una mejor distribución de grupos por asesor— tienden a presentar mejores indicadores de desempeño, lo cual sugiere una carga de trabajo más equilibrada.

4.2. Conclusión

Referencias