**Documentación: Cálculo e Interpretación de la Pendiente en la Evolución de Provisiones**

**1. Objetivo del Cálculo de la Pendiente**

El propósito principal de calcular la pendiente en este contexto es medir la **tendencia** en el comportamiento de la provisión de un cliente a lo largo del tiempo. Específicamente, la pendiente nos indica:

* **Si la provisión está aumentando:** Una pendiente positiva significa que, en promedio, el valor de la provisión crece conforme pasan los meses.
* **Si la provisión está disminuyendo:** Una pendiente negativa indica una tendencia decreciente en el valor de la provisión.
* **Si la provisión se mantiene estable:** Una pendiente cercana a cero sugiere poca o ninguna variación a lo largo del tiempo.

**2. Metodología: Uso de la Regresión Lineal**

Para cuantificar la tendencia, se utiliza una **regresión lineal simple**, que es una técnica estadística que modela la relación entre dos variables. En este caso, las variables son:

* **Variable independiente (x):** El tiempo, representado en este caso por la numeración de los meses en los que se observa la provisión. Debido a que algunos clientes pueden tener datos en cero en determinados meses, únicamente se consideran aquellos meses en los que la provisión es mayor a cero. Se asigna a cada mes válido un número (por ejemplo, 1 para el primer mes con datos, 2 para el segundo, etc.).
* **Variable dependiente (y):** El valor de la provisión en cada uno de los meses válidos.

El modelo de regresión lineal se expresa de la siguiente forma:

y=β0+β1⋅x+εy = \beta\_0 + \beta\_1 \cdot x + \varepsilony=β0​+β1​⋅x+ε

donde:

* β0\beta\_0β0​ es el intercepto, es decir, el valor de la provisión cuando x=0x=0x=0 (aunque en este caso el primer dato se mapea a x=1x=1x=1).
* β1\beta\_1β1​ es la pendiente, que representa el cambio promedio en la provisión por cada unidad de tiempo (mes).
* ε\varepsilonε es el término de error o residuo.

**3. Lógica para el Cálculo de la Pendiente**

La función implementada sigue estos pasos:

1. **Extracción de Datos Válidos:**  
   Dado que cada cliente tiene 12 columnas (una para cada mes), se extraen los valores y se filtran aquellos meses en que la provisión es mayor a cero. Esto es esencial para:
   * Evitar que valores en cero (que podrían representar la ausencia de actividad o datos no disponibles) distorsionen la tendencia real.
   * Asegurar que la regresión se realice únicamente con información útil.
2. **Normalización del Tiempo:**  
   Los meses en que se cuenta con datos válidos se numeran de forma secuencial (1, 2, 3, …). Esto es necesario porque, si bien cada fila tiene 12 columnas, no todos los meses aportan información. Al renumerar se genera una escala de tiempo coherente para cada cliente.
3. **Aplicación de la Regresión Lineal:**  
   Con los datos de tiempo (meses válidos) y las provisiones correspondientes, se utiliza la función linregress de la librería scipy.stats para calcular:
   * La **pendiente (β1\beta\_1β1​)**.
   * El intercepto (aunque no se utiliza directamente en la interpretación principal).
   * El coeficiente de correlación rrr (del cual se obtiene R2R^2R2 para evaluar el ajuste).
   * Otros parámetros estadísticos como el error estándar, el valor p, etc.
4. **Obtención de la Pendiente:**  
   La pendiente calculada indica el **cambio promedio en la provisión por cada mes**. Por ejemplo, una pendiente de 100 significa que, en promedio, la provisión aumenta en 100 unidades por mes.

**4. Interpretación de la Pendiente**

La interpretación de la pendiente se basa en su valor y en el contexto del negocio:

* **Pendiente Positiva (β1>0\beta\_1 > 0β1​>0):**
  + **Significado:** La provisión está aumentando con el tiempo.
  + **Interpretación:** Esto podría indicar un crecimiento en el riesgo asumido o en la asignación de recursos para provisiones, dependiendo del contexto operativo de la compañía.
  + **Ejemplo:** Si la pendiente es +150, en promedio la provisión aumenta en 150 unidades cada mes.
* **Pendiente Negativa (β1<0\beta\_1 < 0β1​<0):**
  + **Significado:** La provisión está disminuyendo a lo largo del tiempo.
  + **Interpretación:** Esto podría reflejar una disminución en el riesgo, una mejora en la eficiencia o una reducción en la exposición, según la situación del cliente.
  + **Ejemplo:** Una pendiente de -75 indica una disminución promedio de 75 unidades mensuales en la provisión.
* **Pendiente Cero o Cercana a Cero (β1≈0\beta\_1 \approx 0β1​≈0):**
  + **Significado:** No se observa una tendencia definida de crecimiento o decrecimiento.
  + **Interpretación:** La provisión se mantiene relativamente estable a lo largo del tiempo. Es importante corroborar este hallazgo con otros indicadores o revisando la dispersión de los datos (por ejemplo, mediante el R2R^2R2).

**5. Complemento: Validación con R2R^2R2**

Aunque la pendiente es la medida principal para determinar la tendencia, es crucial acompañarla del coeficiente de determinación R2R^2R2. Este valor ayuda a evaluar **qué tan bien el modelo lineal explica la variabilidad en los datos**:

* Un R2R^2R2 cercano a 1 indica que la línea de tendencia es una buena representación del comportamiento real de la provisión.
* Un R2R^2R2 bajo sugiere que, a pesar de tener una pendiente calculada, los datos presentan alta variabilidad y la tendencia lineal puede no ser representativa.

Esta validación es importante para asegurarse de que la interpretación de la pendiente sea confiable.

**6. Conclusión**

La lógica detrás del cálculo de la pendiente se fundamenta en la aplicación de un modelo de regresión lineal sobre datos previamente filtrados y normalizados. La pendiente nos permite comprender de manera sencilla la dirección y magnitud del cambio en la provisión a lo largo del tiempo. Al combinar esta métrica con el R2R^2R2 y otros indicadores (como el cambio total), se obtiene una visión robusta y completa del comportamiento de la provisión, lo que es esencial para la toma de decisiones y el análisis del riesgo en el entorno de la compañía.