

26 DE NOVIEMBRE DEL 2024

Campaña comercial de retail banking

Rafael Antonio Fabio Aguilera

Itzae Abigail Flores Lara

Tarek Elías Hirmas Aboid

Felipe Santiago Nicolás Norambuena Palacios

Alon Menahem Rosenblum Keller

Ilan Enrique San Martín Kleinkopf

Profesor: Agustín Chiu - Ayudante: Lucas Salcedo

José Pedro Vargas Figueroa

Tabla de contenidos

- 01 Descripción del problema
- 02 Metodología e implementación
- 03 Resultados
- 04 Cierre

Descripción del problema

Definiciones

Retail banking:

- En este caso estaremos estudiando la campaña para créditos de consumo que hace un banco.



Definiciones

Crédito de consumo:

- Es un préstamo otorgado por el banco, cobrando una **tasa de interés** como costo del financiamiento.



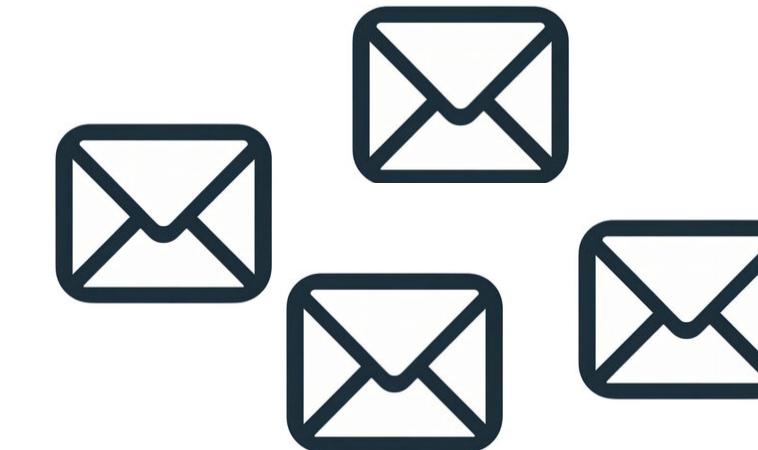
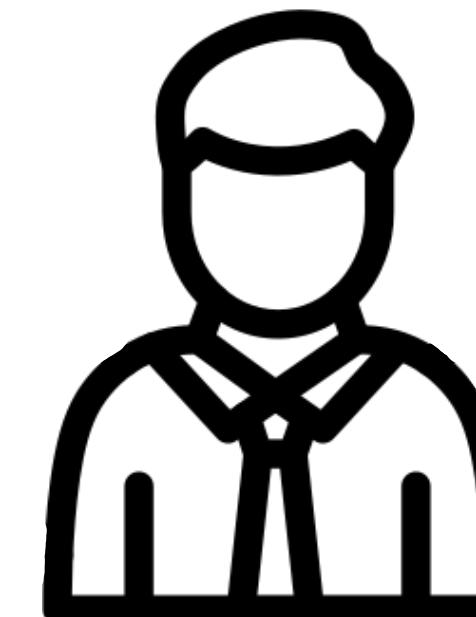
Definiciones

Tratamiento:

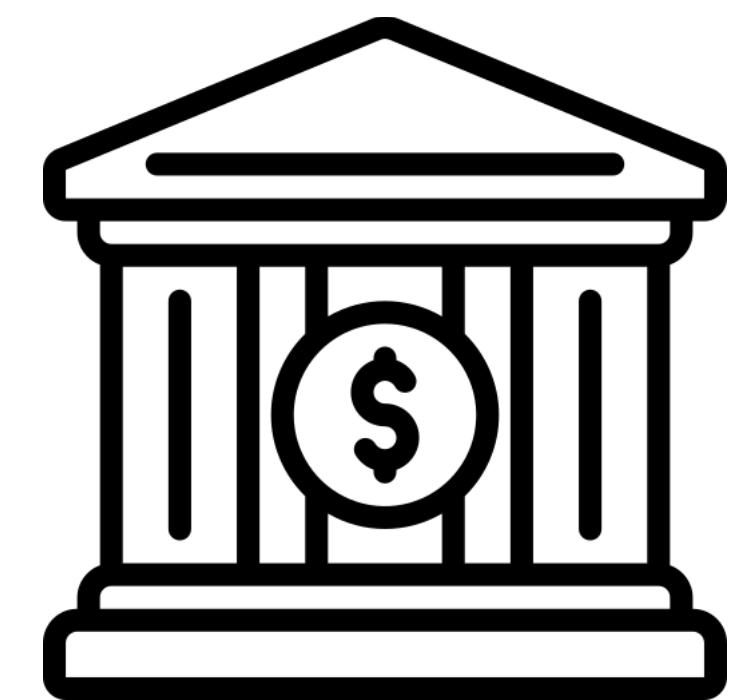
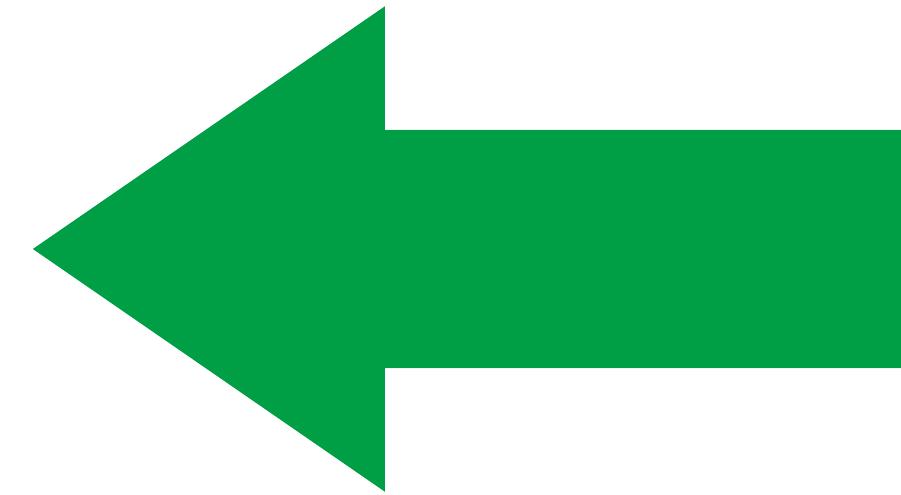
- Forma en que el banco contacta a sus clientes para ofrecer créditos de consumo.

Consiste en:

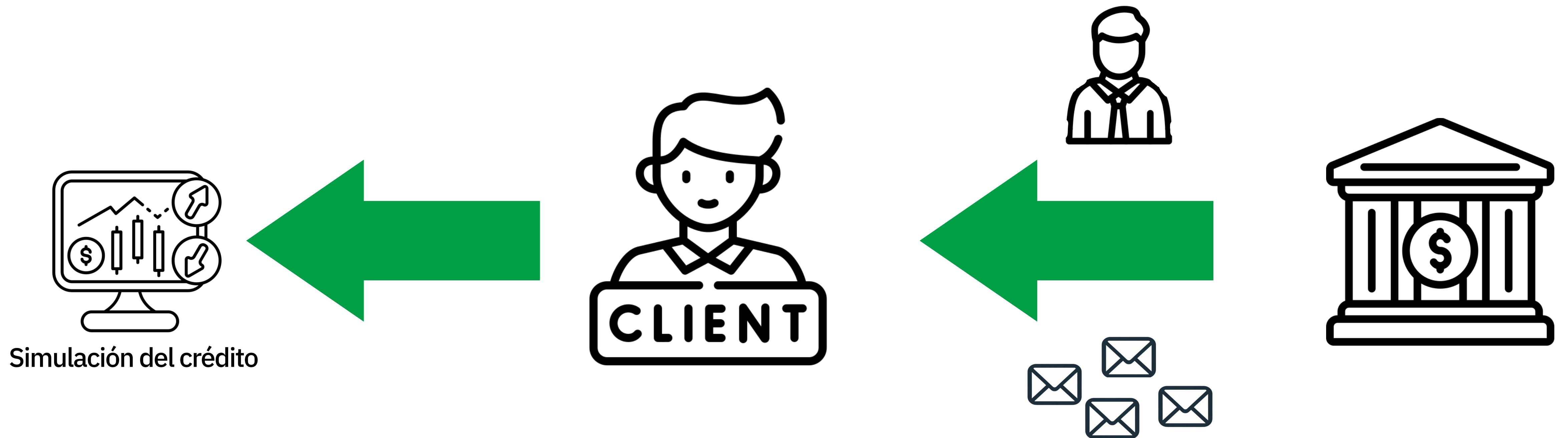
- Variable binaria de asignación de ejecutivo.
- Variable entera entre 0 y 4 de asignación de correos.



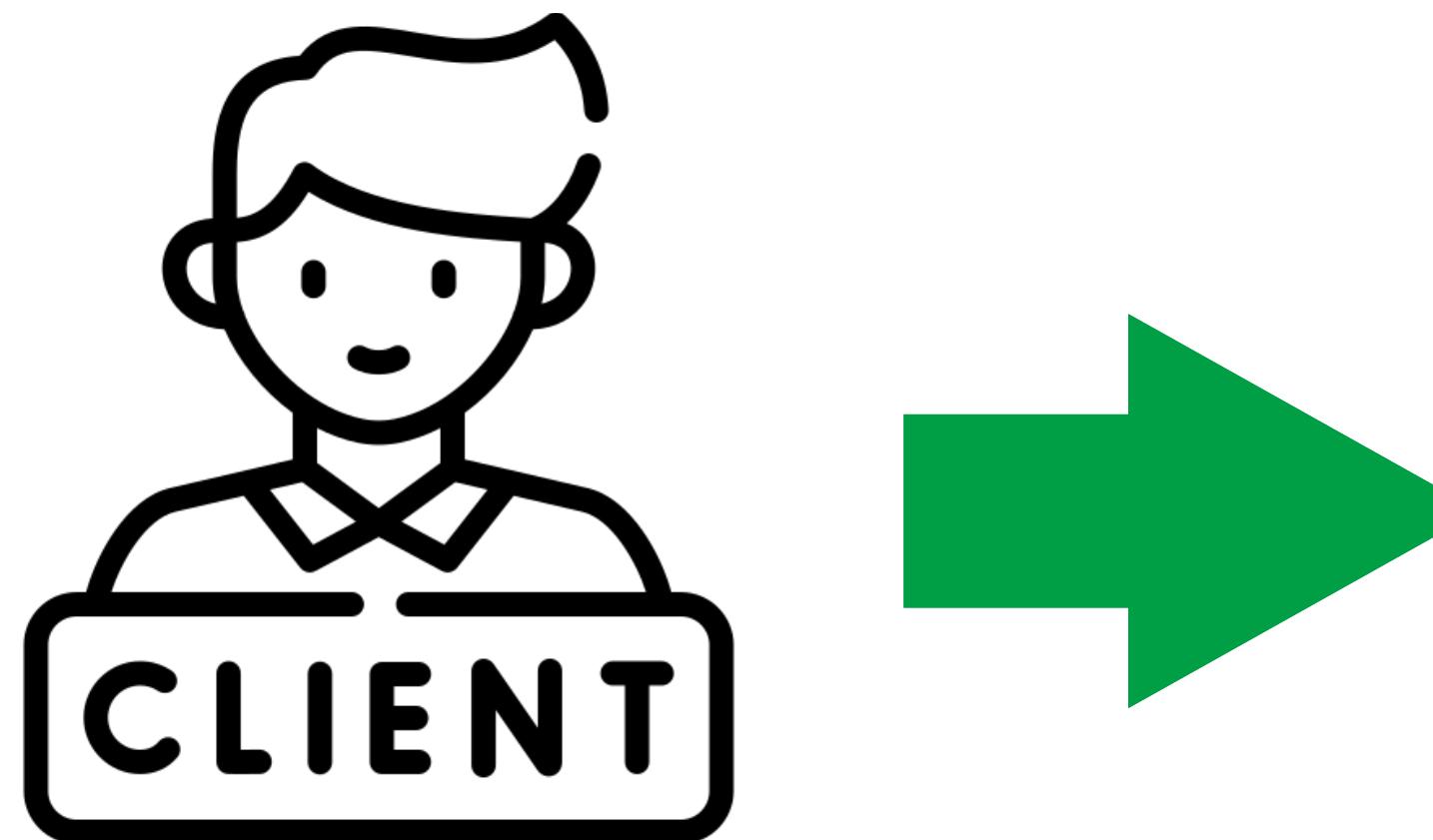
Descripción del problema



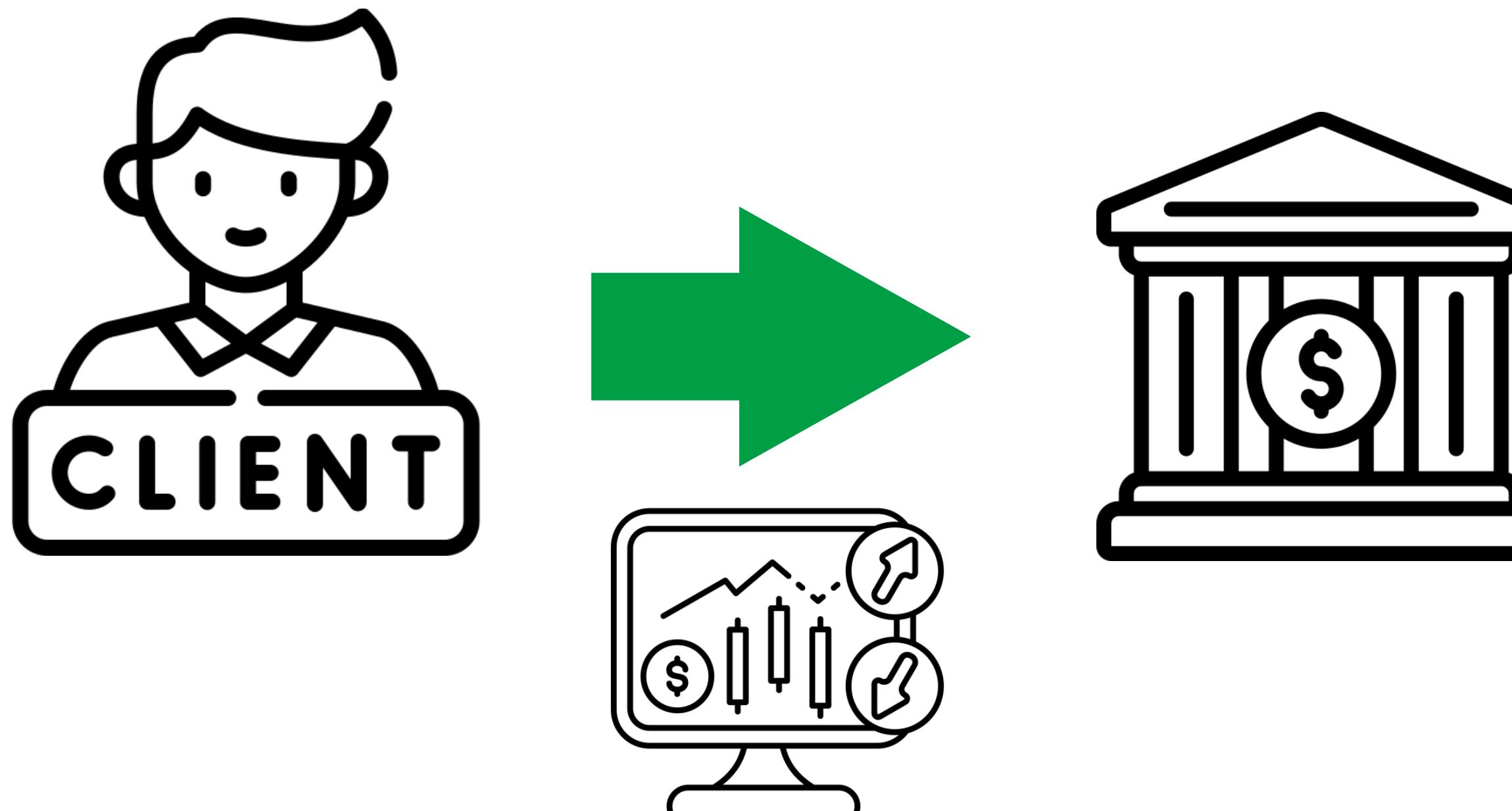
Descripción del problema



Descripción del problema

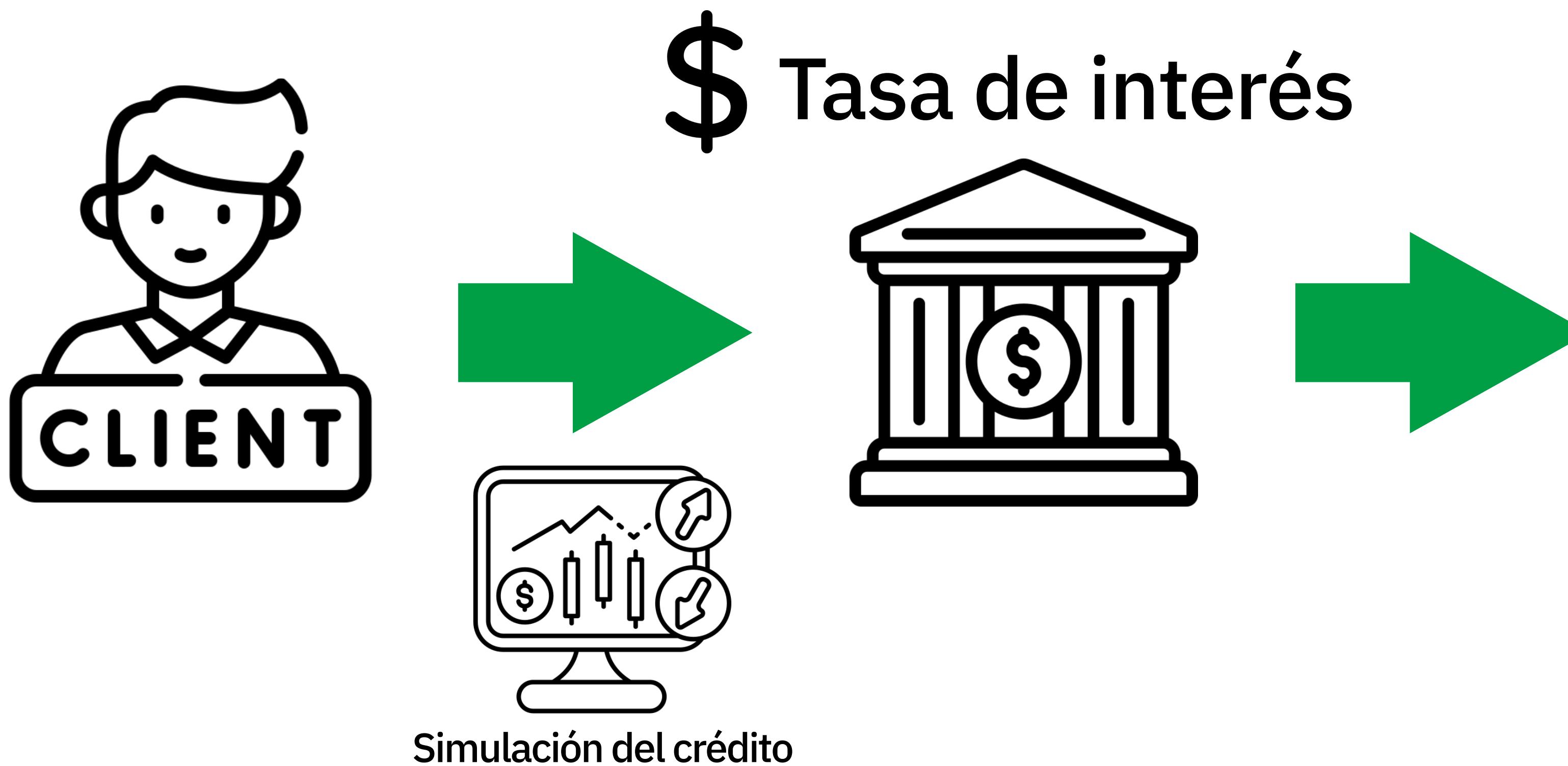


Descripción del problema

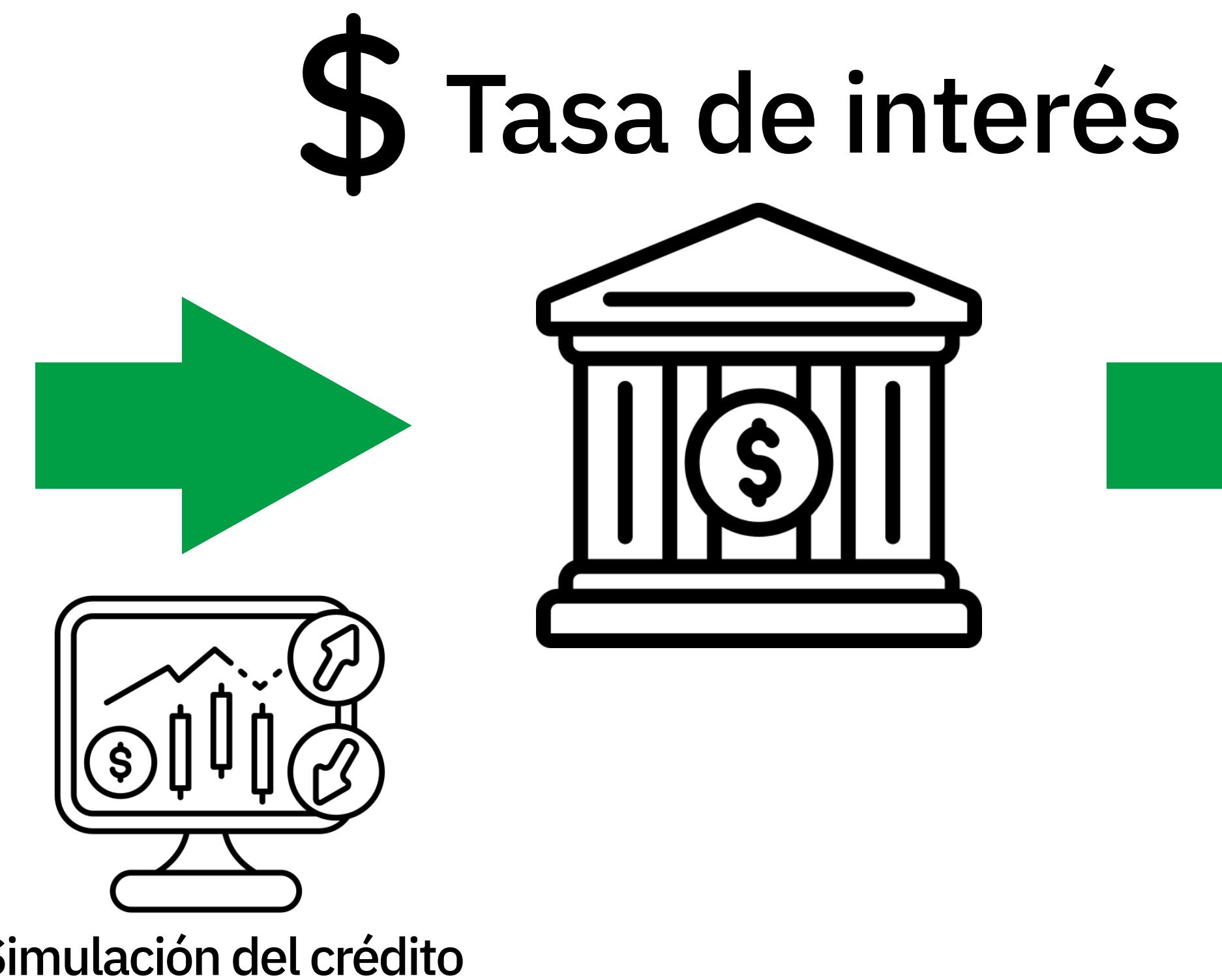


Simulación del crédito

Descripción del problema



Descripción del problema



Maximizar las ganancias del banco

Asignando los mejores tratamientos(Ejecutivos y mails)
y los mejores precios (Tasas de interés)

Objetivo del problema

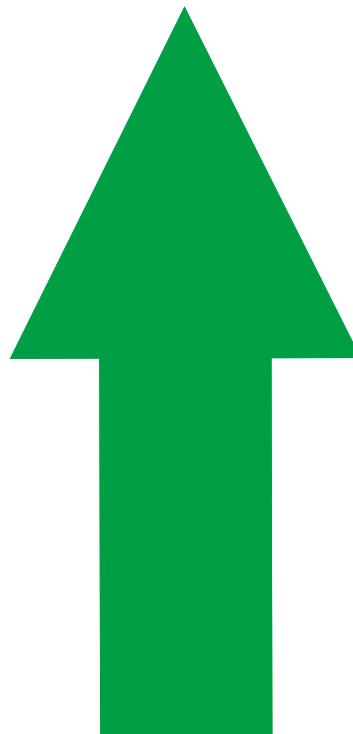
Asignar tratamientos a clientes:

- Asignar ejecutivos
- Asignar número de mails

Objetivo del problema

Asignar tratamientos a clientes:

- Asignar ejecutivos
- Asignar número de mails



Simulación de créditos
de consumo



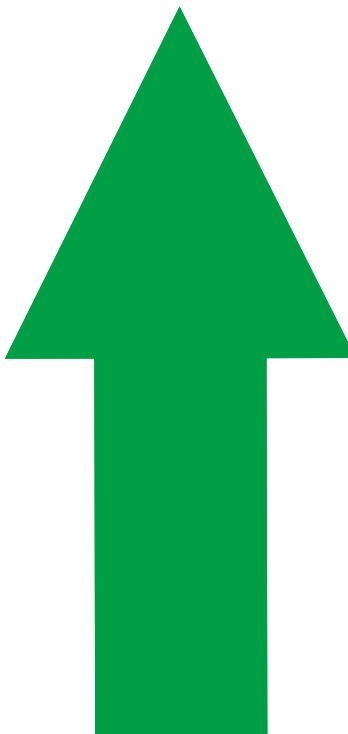
Objetivo del problema

Asignar tratamientos a clientes:

- Asignar ejecutivos
- Asignar número de mails

Asignar precios:

- Encontrar el mejor precio para maximizar las ganancias



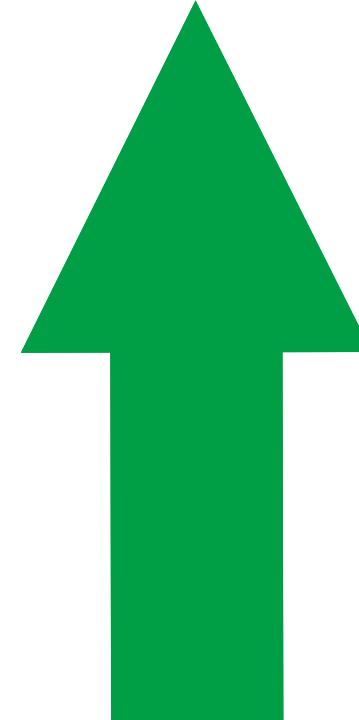
Simulación de créditos
de consumo



Objetivo del problema

Asignar tratamientos a clientes:

- Asignar ejecutivos
- Asignar número de mails

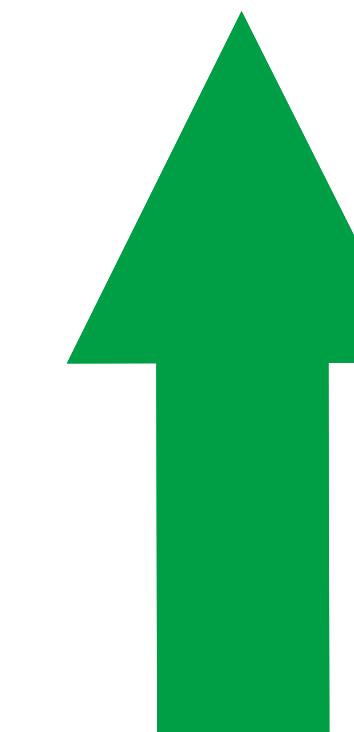


Simulación de créditos
de consumo



Asignar precios:

- Encontrar el mejor precio para maximizar las ganancias



Probabilidad de
que los clientes
acepten el crédito



Información con la que contamos

Datos de los clientes

- Género
- Edad
- Nacionalidad
- Renta
- Elasticidad
- Etc...

Datos de ventas históricas

- Fecha de la venta
- Rut que compró

Datos de simulaciones históricas

- Fecha de la simulación
- Rut que simuló
- Monto que simuló
- Plazo que simuló
- Tasa de interés

Datos de tratamientos otorgados por el banco históricamente

- Fecha del tratamiento
- Rut tratado
- Tratamiento

Desafíos de la gestión de créditos

Número elevado de clientes

- Segmentación compleja
- Diversas necesidades y requerimientos

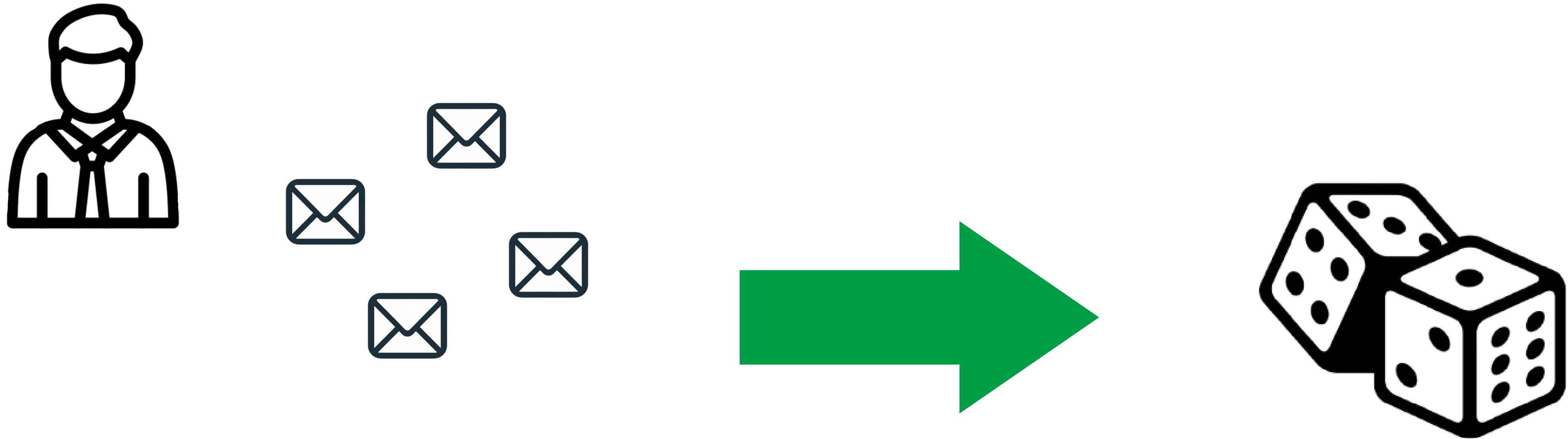
Gran número de datos

- Tratamientos: 35 Mill x 5
- Simulaciones: 8.5 Mill x 6
- Ventas: 2,8 Mill x 3
- Clientes: 550 mil x 13

Problema estocástico

- Variables aleatorias
- Probabilidad de aceptación de crédito
- Probabilidad de no pago
- Montos simulados
- Plazos simulados

Caso Base

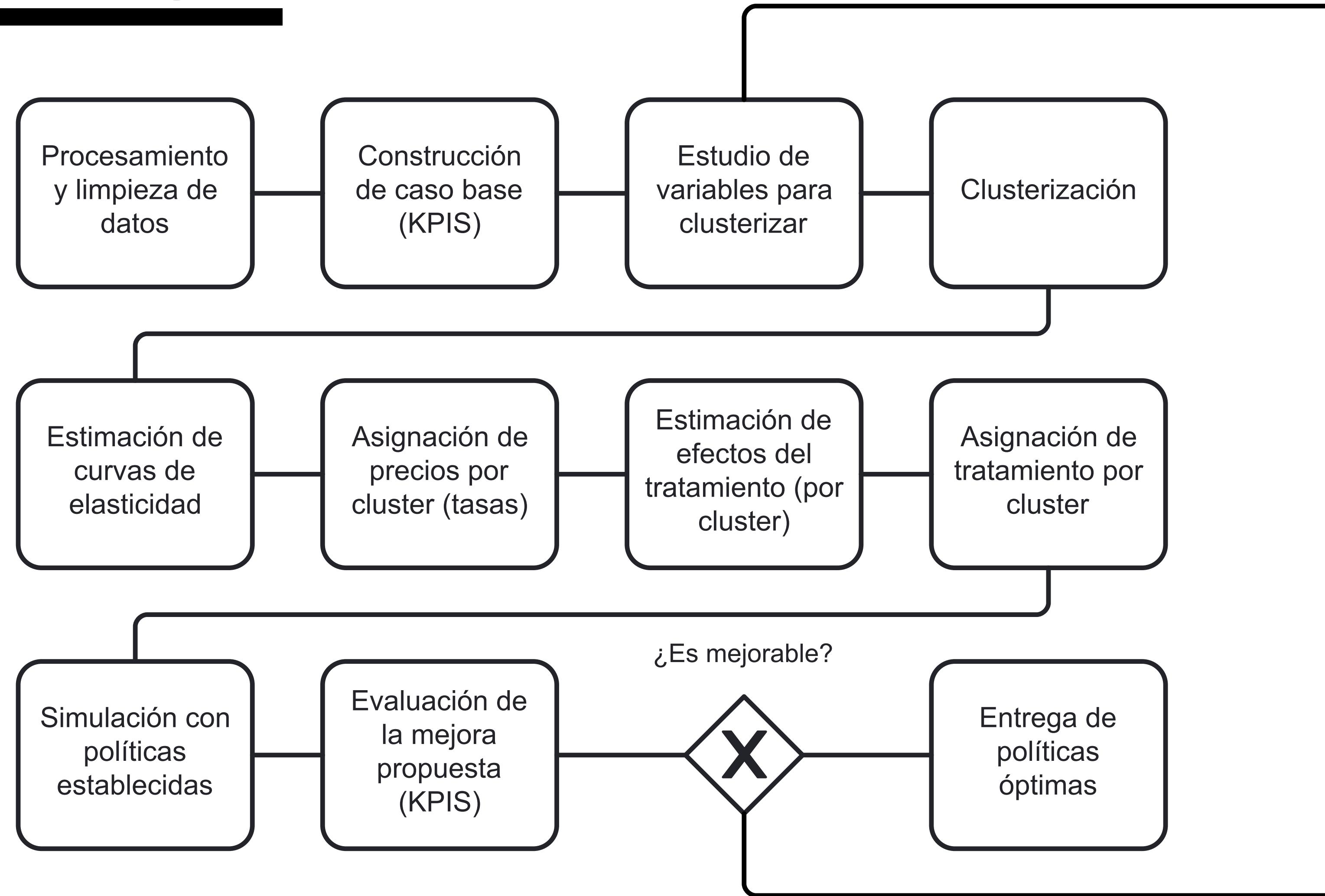


\$ Tasa de interés

Metodología e Implementación

Conclusiones y resultados

Metodología



Datos iniciales

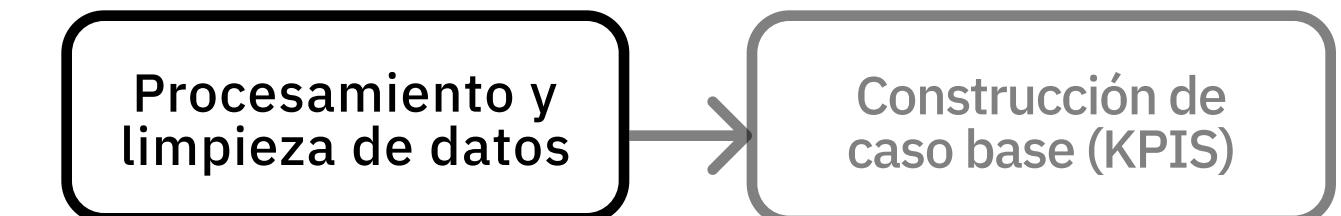
Datos determinísticos: Características demográficas de los clientes

Datos estimados: Comportamientos y preferencias de los clientes frente a los créditos

Tratamientos: De solicitud de créditos

Simulaciones: Montos, plazos y tasas de los créditos

Ventas: Fecha de venta de créditos



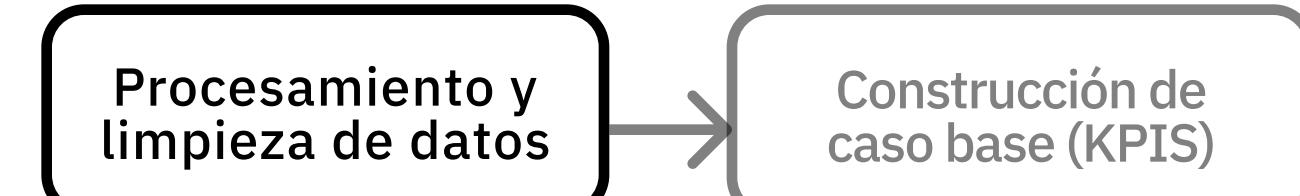
Datos iniciales

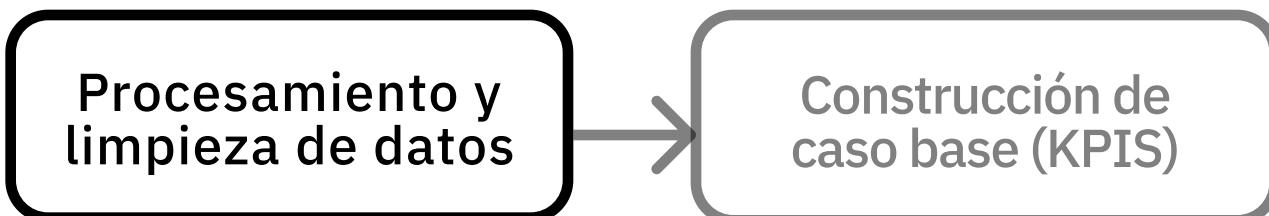
Eliminación de datos no útiles:

- Columnas sin variación entre clientes
- Datos nulos o inválidos.

Variables clave analizadas:

- Monto simulado
- Renta
- Oferta de consumo
- Deuda CMF





Datos relevantes (énfasis en los tratados)

- Elasticidad_Precios - Alta, Media o Baja
- Propensión - Valor decimal entre 0 y 1
- Probabilidad_No_Pago - Valor decimal entre 0 y 1
- Oferta_Consumo - Monto máximo que el cliente puede solicitar basado en antecedentes.
- Renta - Ingreso promedio del cliente (últimos 12 meses)
- Deuda_CMF - Deuda en otros bancos
- Tiempo_como_cliente - Tiempo en años que lleva en el banco
- Monto simulado - Monto que el cliente simula al evaluar un crédito

Scatter plots

- Relación lineal entre las variables Deuda, Monto Simulado, Oferta Consumo y Renta.

Elasticidad Precios

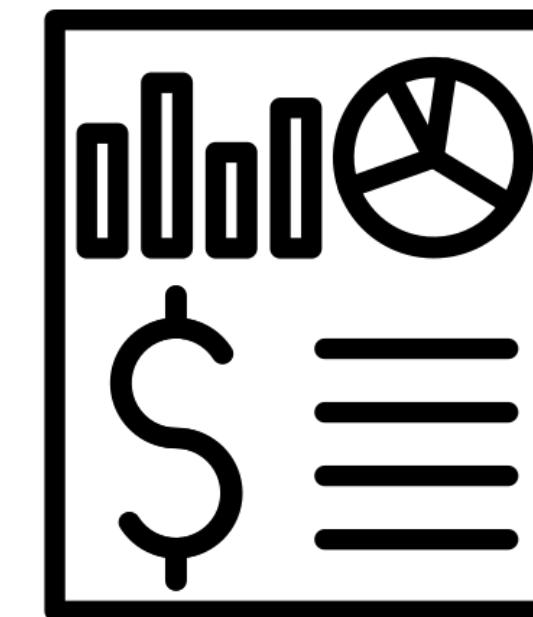
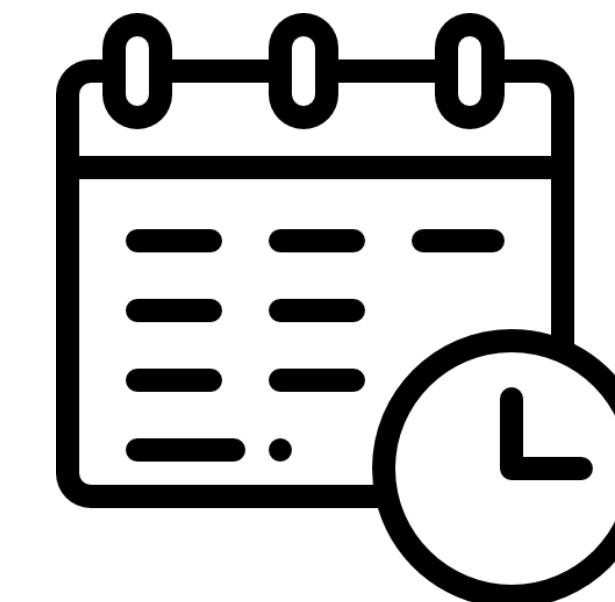
- Alta
- Baja
- Media



Procesamiento y
limpieza de datos → Construcción de
caso base (KPIs)

KPIS Caso Base

- Análisis **mensual** del rendimiento financiero
- Análisis de la ejecución operativa inicial del banco (sin tratamiento)
- Base para evaluar el impacto de la solución propuesta por el grupo.



KPIS iniciales



- **Revenue esperado**

$$\text{Revenue esperado} = \sum \text{Revenue}_i * (1 - P_{\text{no_ pago_}i}) = \textcolor{red}{58.000 - 65.000 millones de pesos por mes}$$

- **Tasa de créditos cursados**

$$\text{Tasa de créditos cursados} = \frac{\text{Créditos cursados}}{\text{Créditos simulados}} \times 100 = \textcolor{red}{32\% - 34\%}$$

- **Promedio de correos enviados**

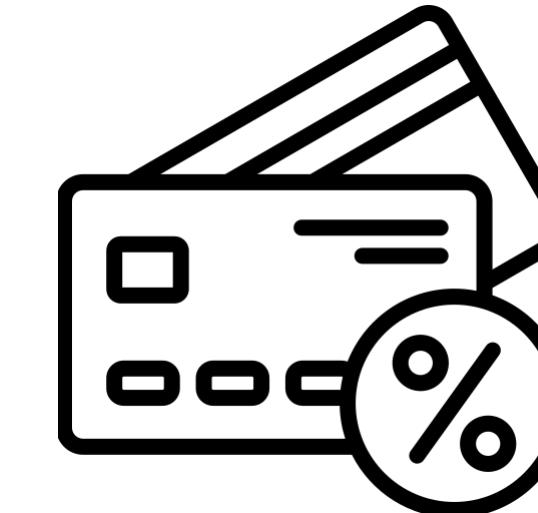
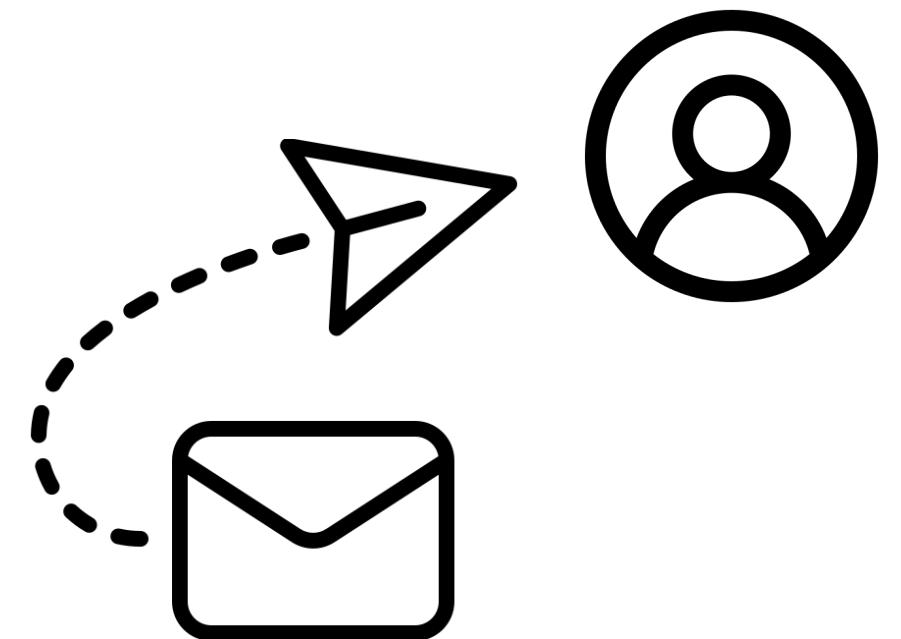
$$\frac{\sum n_{\text{correos_}i}}{n} = 2.15 \text{ correos}$$

- **Revenue simulado (100 simulaciones)**

Valores muy cercanos al revenue esperado

- **Desviación Estándar de las simulaciones**

39 - 57 millones de pesos (0.1% de las ganancias mensuales)



Resultados KPIs



KPI	Cluster inicial
Revenue Esperado	61.5 mil millones de pesos
Tasa de Curse	33%
Correos Promedio	2.15
Revenue Simulado	61.4 mil millones de pesos
Desviación Estándar	48 millones de pesos (<0.1%)

Conclusiones Caso base



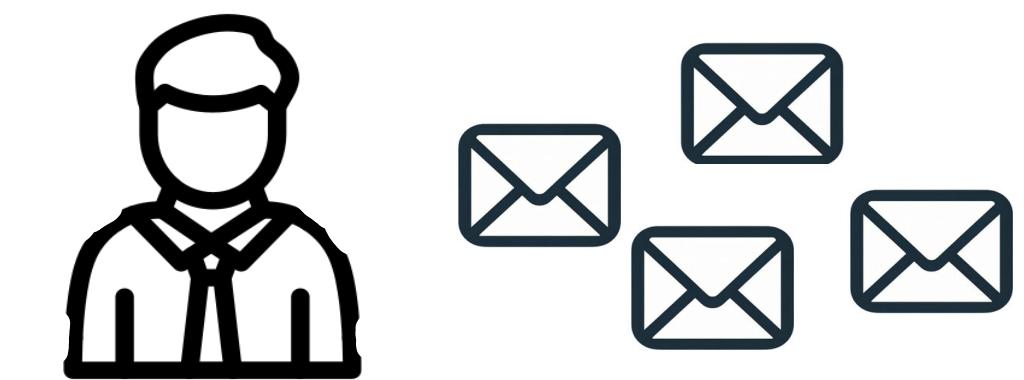
- **Contexto:** El banco ofrece créditos con **tasas variables**, bajo incertidumbre por probabilidad de curse y no pago.
- **Factores clave:** La tasa afecta la **aceptación** del cliente de un crédito. El tratamiento afecta las **simulaciones** que el cliente hace.
- **Problema actual:** Los recursos del banco se asignan **aleatoriamente**, generando un revenue base que **no es óptimo**.
- **Objetivo:** La nueva metodología busca **maximizar el revenue** ajustando estratégicamente las **variables clave**, asignando precios y tratamientos.

¿Para qué queremos clusterizar?



2 preguntas claves

- 1 ¿Qué tanto **esfuerzo queremos invertir** en un cliente en términos de la **ganancia que nos puede proporcionar**?
- 2 ¿Qué tanto **esfuerzo queremos invertir** en un cliente en términos de la **respuesta que tenga al tratamiento**?

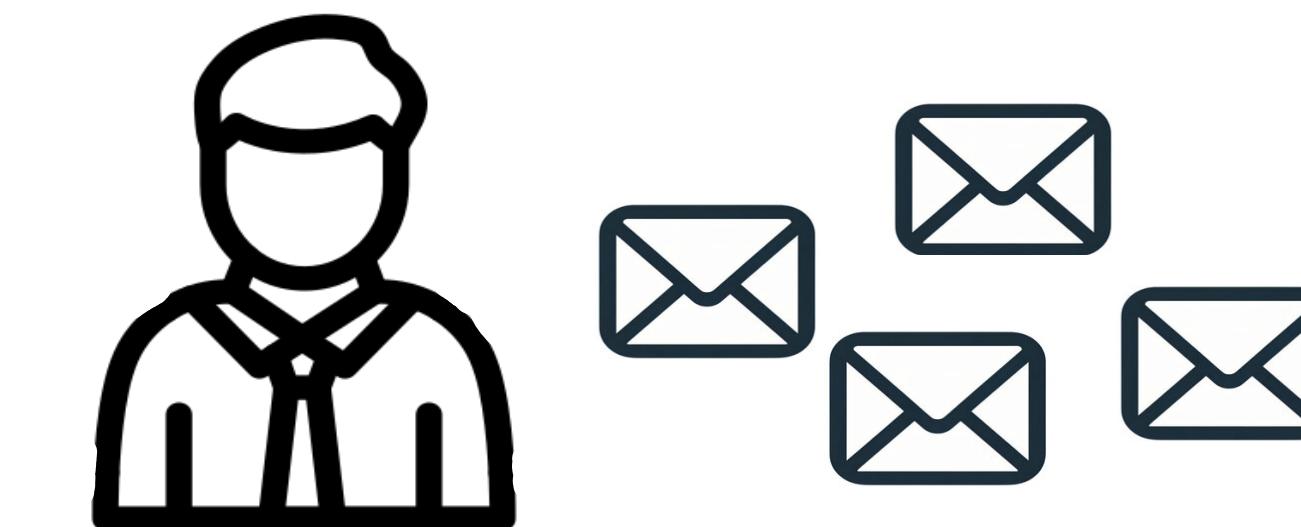
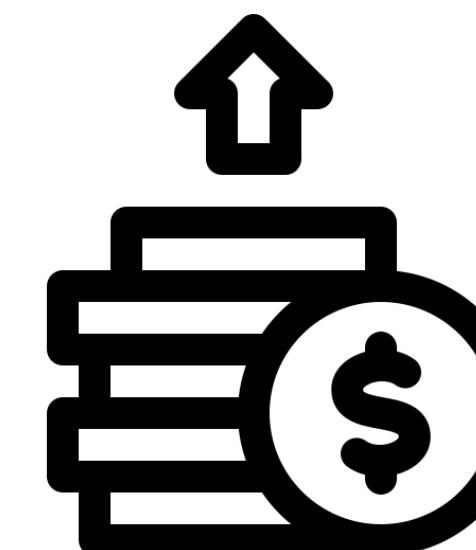


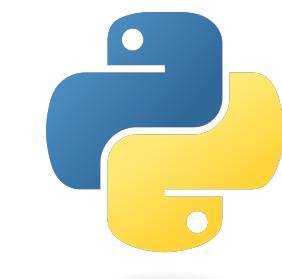
Estudio de variables para clusterizar



Se busca identificar qué variables son útiles para clusterizar.

Se construyeron tablas por mes con los datos más relevantes para calcular el revenue potencial y ver cómo influyó el tratamiento.



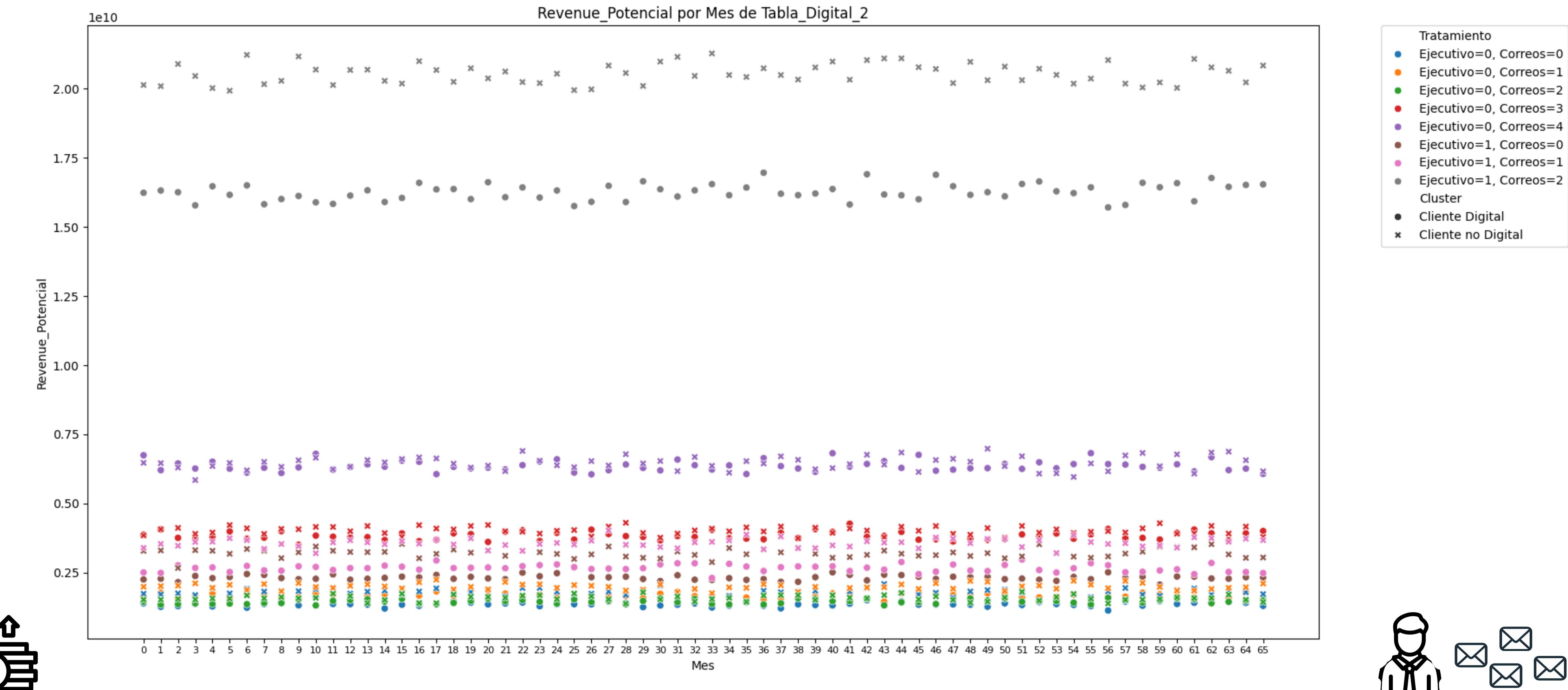


Variables y Tratamientos

Construcción de caso base (KPIs)

Estudio de variables para clusterizar

Clusterizar



Conclusiones de variables

Conclusiones clave:

Variables generadoras de revenue:

Género (Hombres >>> Mujeres)

Edad (Adultos >>> Jóvenes, Adultos mayores)

Renta (R_Altas >>> R_Medias >>> R_Bajas)



Variables reactivas al tratamiento:

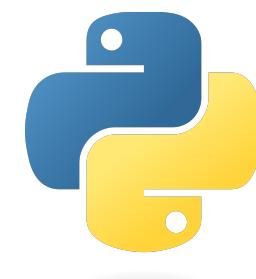
Edad: Adultos

Género: Hombres

Rentas: Rentas Altas

Prob_No_Pago: Buenos pagadores (<1.5%)





Clusterización

Se construyó una primera clusterización **preliminar y arbitraria.**

Variables:

Categoría_Digital: [Digital, no digital] → 2

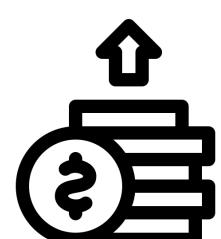
Género: [Masculino, Femenino] → 2

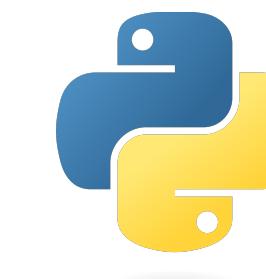
Edad: [Joven (18 a 35), Adulto (35 a 60), Adulto mayor (60+)] → 3

Renta: [Alta (20% superior) , media (50% del medio) , baja (30% inferior)] → 3

Propensión: [Alta (tercio superior) , media (tercio del medio) , baja (tercio inferior)] → 3

$$2 \cdot 2 \cdot 3 \cdot 3 \cdot 3 = 108$$





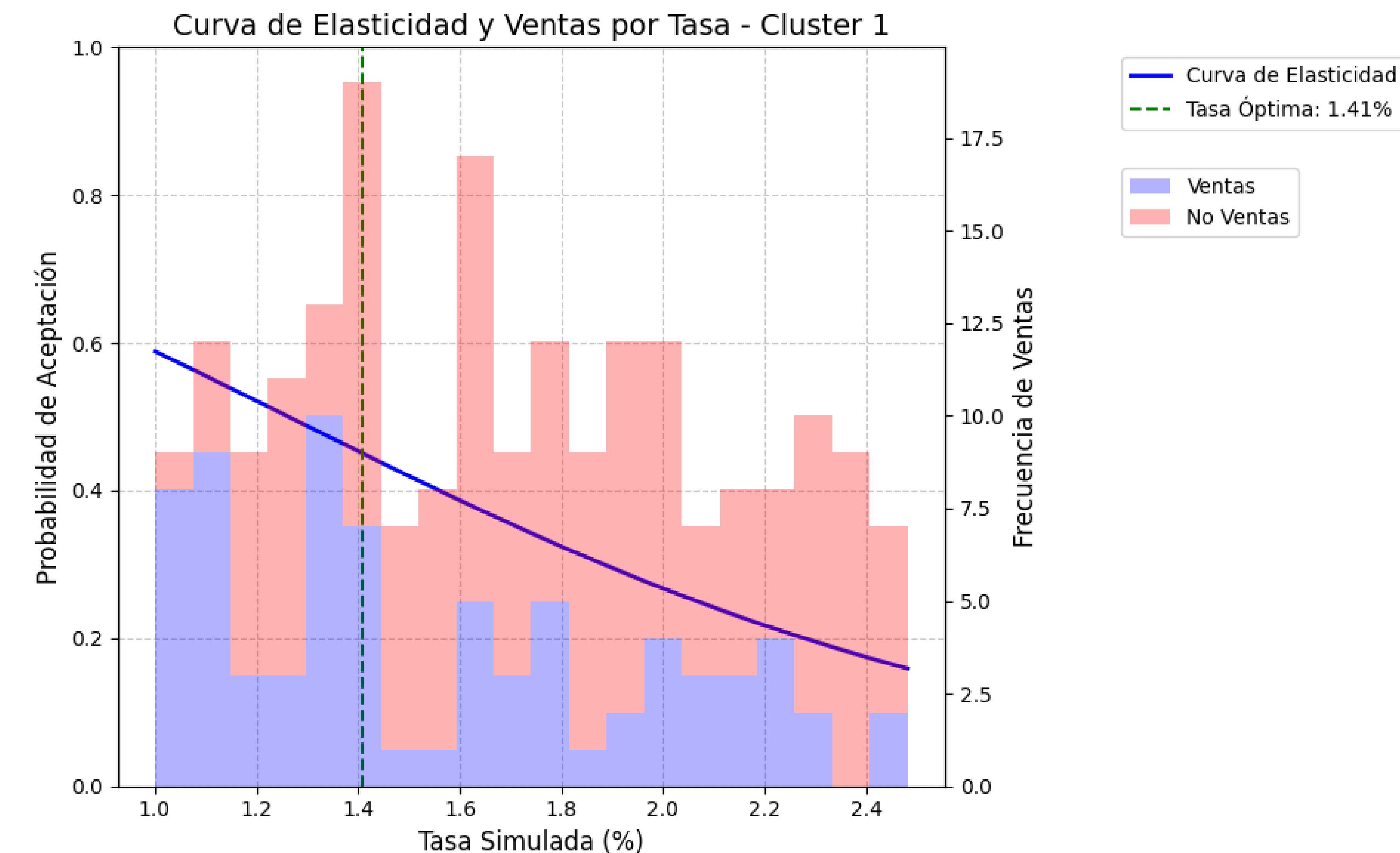
Metodología de solución

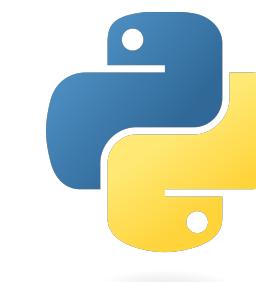
REGRESIÓN LOGÍSTICA

Variable dependiente: Venta

Predictor: Tasa_Simulado.

Se generan predicciones para una cuadrícula de tasas (Tasa_Simulado), permitiendo trazar una curva de elasticidad que relaciona probabilidad de aceptación con tasa.





Metodología de solución

Rmp: Revenue potencial medio

MP: Monto prestado del crédito medio mensual del cluster

PC: Plazo del credito medio en meses del cluster

Ti: Tasa en el mes i

$$Rmp_i = \left(PC \times MP \times \left(\frac{T_i \times (1+T_i)^{PC}}{(1+T_i)^{PC}-1} \right) \right) - MP$$

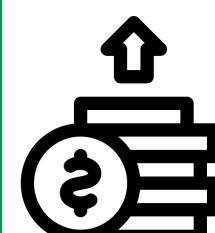
REi: Revenue esperado del banco

Pnp: La probabilidad media de no pago del cluster

Smm: Cantidad de simulaciones medias mensuales al cluster

PAi: Probabilidad de aceptacion de la tasa i para el cluster

$$RE_i = Rmp_i \times (1 - Pnp) \times Smm \times PA_i$$



Metodología de solución

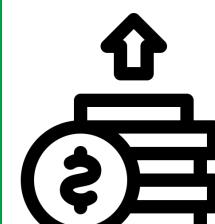


- 1 ¿Qué tanto **esfuerzo** queremos **invertir** en un cliente en términos de la **ganancia** que nos puede proporcionar?



Obtenemos el revenue esperado por cluster y podemos responder la primera pregunta.

Tenemos una jerarquía con respecto a que cluster nos puede generar más ganancias



Metodología de solución

CASOS FAVORABLES / CASOS TOTALES

Casos Totales

Para cada uno de los tratamientos posibles contamos cuantos hubo a lo largo de toda la historia

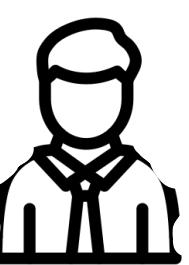
Casos Favorables

Para cada uno de los tratamientos posibles contamos cuantas simulaciones hubo.

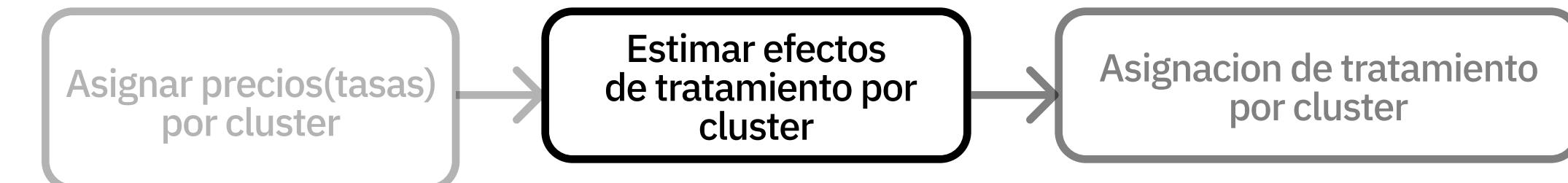
Probabilidad de simular dado tratamiento

Casos favorables / Casos totales

Tratamiento	Personas que recibieron dicho tratamiento (Casos totales)	Simulaciones relacionadas con ese tratamiento (Casos favorables)	Probabilidad de simular dado el tratamiento
Tratamiento 1	1000	200	0.2
Tratamiento 2	2000	30	0.015
Tratamiento 3	1500	12	0.008
⋮	⋮	⋮	⋮
Tratamiento 8	50	49	0.98

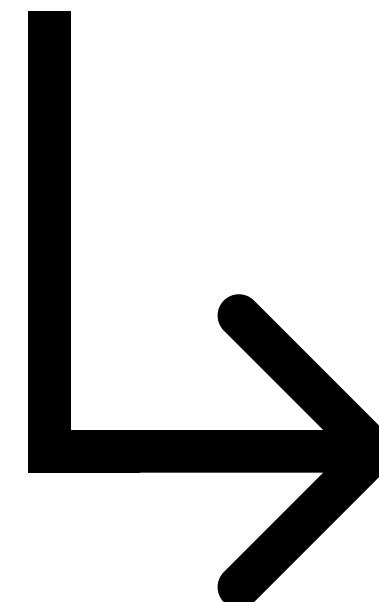


Metodología de solución

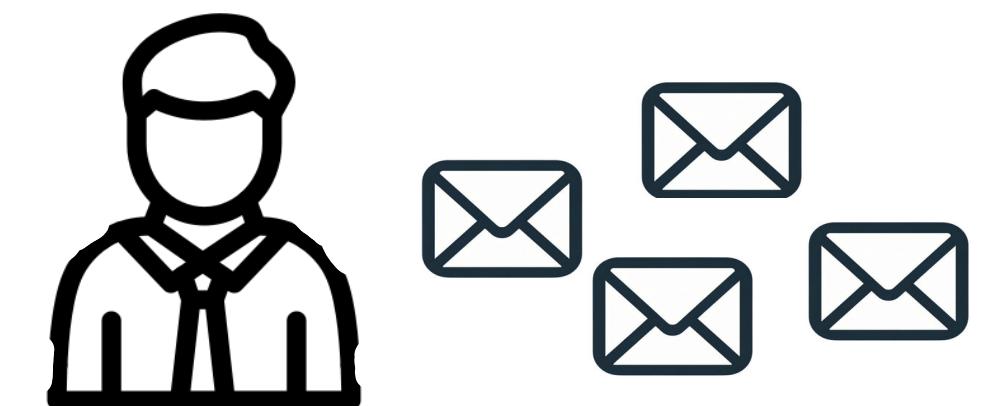


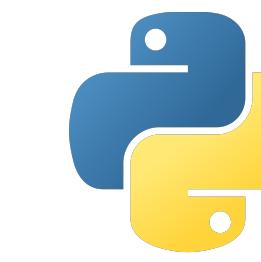
2

¿Qué tanto **esfuerzo** queremos **invertir** en un cliente en términos de la **respuesta que tenga al tratamiento**?



Obtener la **probabilidad de simulación** para un **tratamiento por cluster**





Modelo de optimización

$$\text{Max} \quad \sum_{i \in I} \sum_{t \in T} z_{i,t} \cdot r_i \cdot pc_i \cdot ps_i - \sum_{c \in C} \sum_{i \in I_c} \sum_{t \in T} z_{i,t} \cdot C_c \cdot n_correos_t$$

Sujeto a:

Cada cliente recibe exactamente un tratamiento:

$$\sum_{t \in T} z_{i,t} = 1 \quad \forall i \in I$$

Los clientes del mismo cluster reciben el mismo tratamiento:

$$z_{i,t} = z_{j,t} \quad \forall i, j \in I_c, \forall c \in C, \forall t \in T$$

Restricción sobre la capacidad de ejecutivos:

$$\sum_{c \in C} \sum_{i \in I_c} (z_{i,t_6} + z_{i,t_7} + z_{i,t_8}) \leq E$$

Relación binaria para asignación de tratamientos:

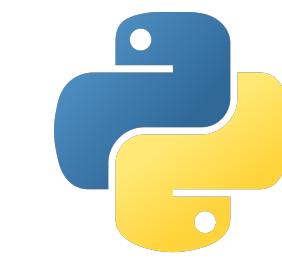
$$z_{i,t} \in \{0, 1\} \quad \forall i \in I, \forall t \in T$$

Cada cluster tiene un tratamiento único:

$$\sum_{t \in T} z_{i,t} = \sum_{t \in T} z_{j,t} \quad \forall i, j \in I_c, \forall c \in C$$



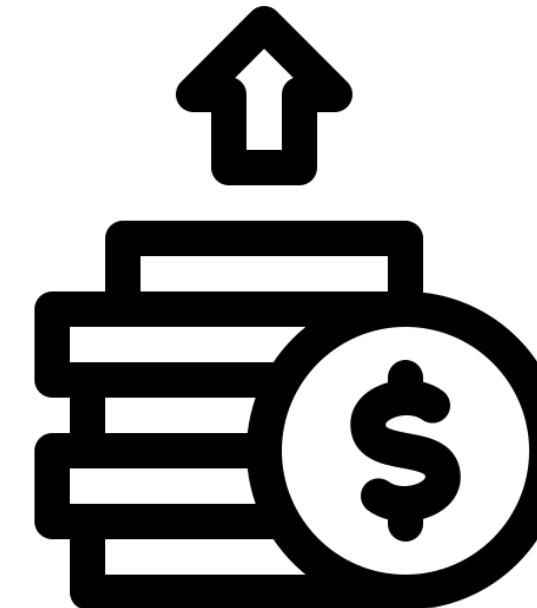
- I : Conjunto de clientes.
- C : Conjunto de clusters, con $c \in C$ representando un clúster específico.
- I_c : Conjunto de clientes en el clúster $c \in C$, donde $I = \bigcup_{c \in C} I_c$.
- T : Conjunto de tratamientos disponibles $\{t_1, t_2, \dots, t_8\}$.
- r_i : Revenue potencial por cliente $i \in I$.
- pc_i : Probabilidad de conversión para el cliente $i \in I$, dada una simulación, que es función del tratamiento $t \in T$.
- ps_i : Probabilidad asociada a la simulación para el cliente i , que es función del precio del clúster al que pertenece.
- C_c : Costo de enviar un correo en el clúster $c \in C$.
- $z_{i,t}$: Variable binaria que indica si el cliente i recibe el tratamiento $t \in T$.
- $n_correos_t$: Número de correos asociados al tratamiento t .
- e_t : Número de ejecutivos asociados al tratamiento t .
- E : Capacidad máxima de ejecutivos disponibles.

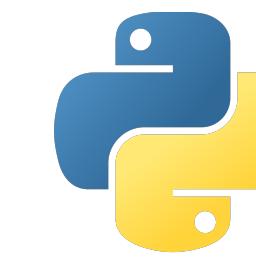


Resultados del Modelo

El modelo nos entrega:

- Revenue máximo del banco según restricciones
- Tratamiento óptimo por Cluster para maximizar el Revenue





Simulación y KPIs



Variables aleatorias:

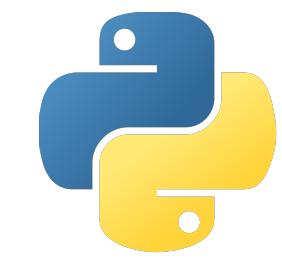
- Probabilidad_No_Pago → Binomial
- Probabilidad_Simular → Binomial
- Probabilidad_Aceptación → Binomial
- Plazo_Simulado → Distribución Triangular

Aproximaciones:

- Monto_Simulado → Regresión Lineal Múltiple
- Tasa_Optima → Obtenido en pasos previos
- N_Correos → Tratamiento de cada cliente según el resultado del modelo

Simulación (100 veces):

- Calculamos el Revenue Esperado (previo a simular)
- Calculamos el número de correos promedio (previo a simular)
- Simulamos la tasa de curse
- Simulamos el Revenue generado
- Calculamos la desviación estándar del Revenue simulado



Simulación y KPIs

Aproximaciones:

- **Monto_Simulado → Regresión Lineal Múltiple**

Relación lineal entre las variables Deuda, Monto Simulado, Oferta Consumo y Renta.



$$\text{Monto_Simulado} = -866900 + 0.8845 * \text{Renta} + 0.7231 * \text{Oferta_Consumo} - 0.105 * \text{Deuda_CMF}$$



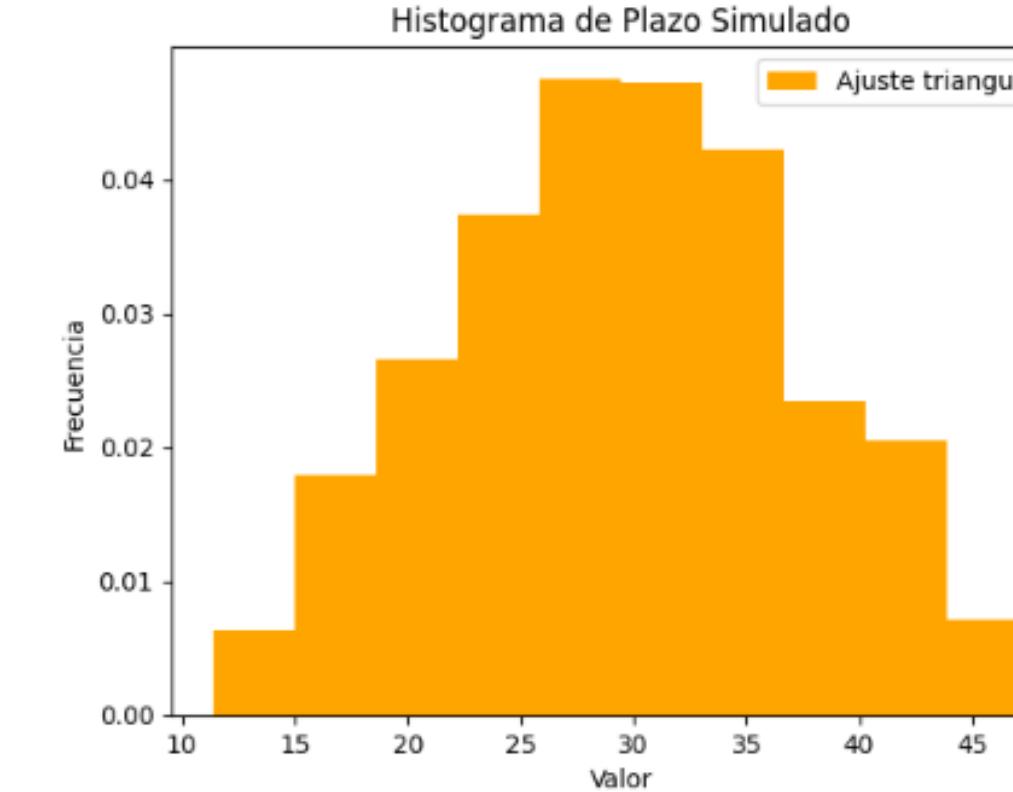
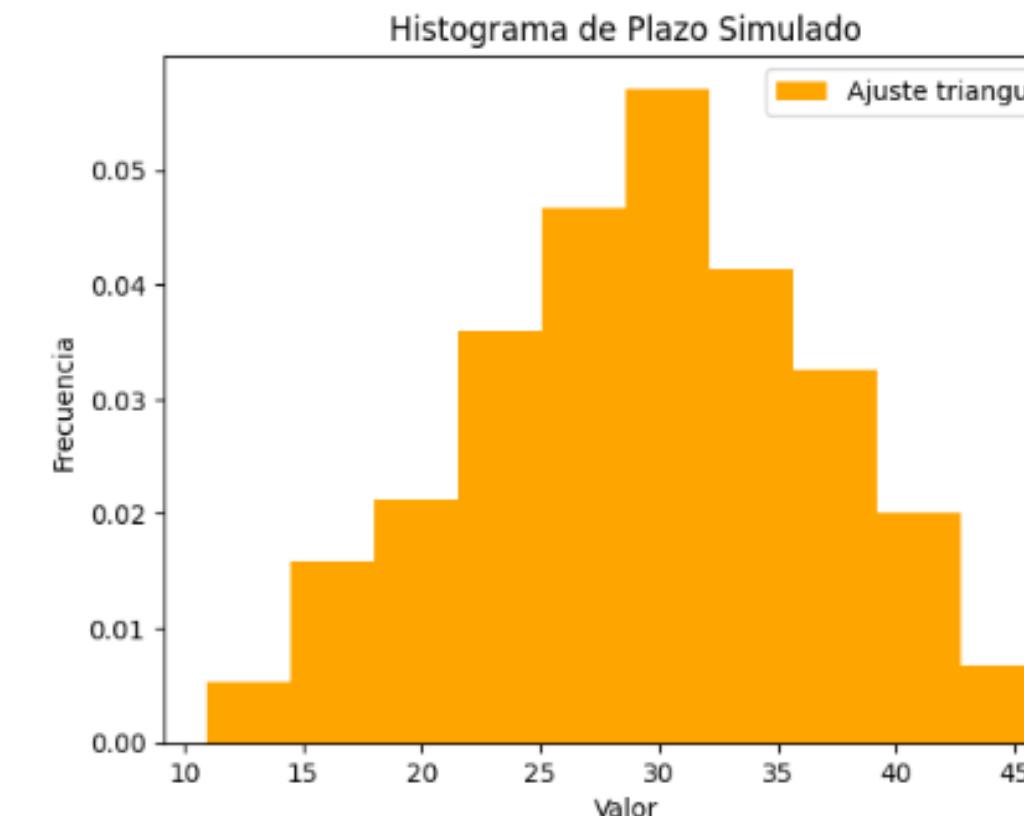
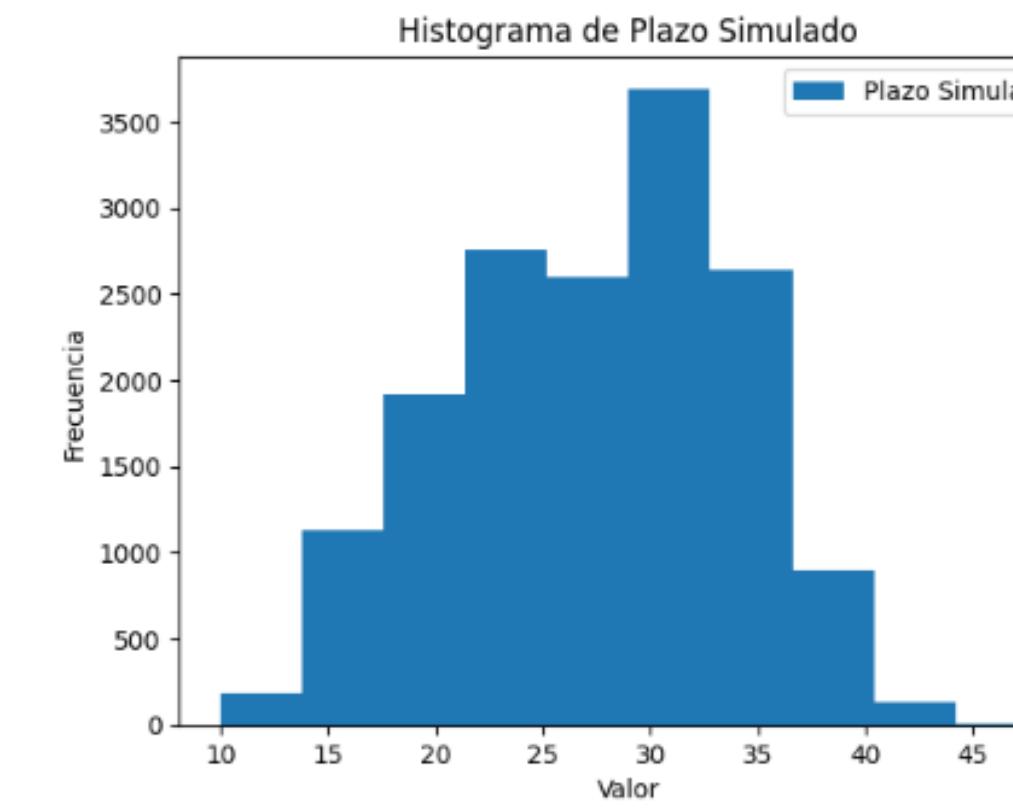
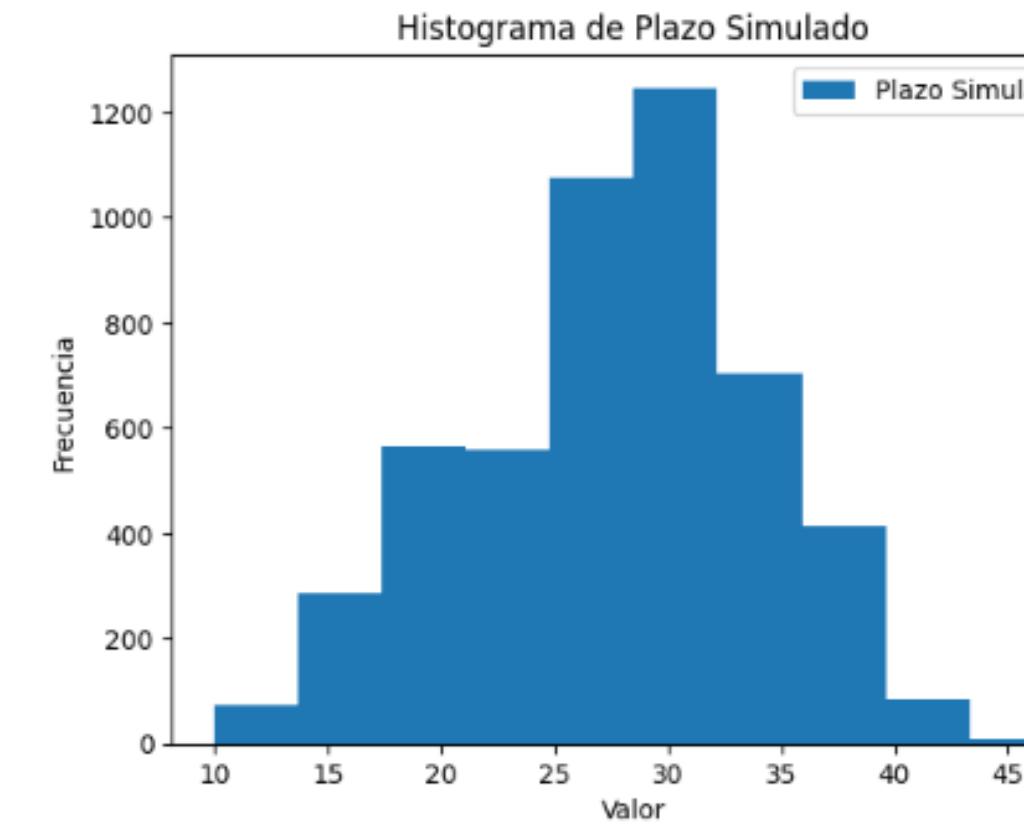
R Studio®



Simulación y KPIs

Variable aleatorias:

- Plazo_Simulado → Distribución Triangular



Comparación KPIs

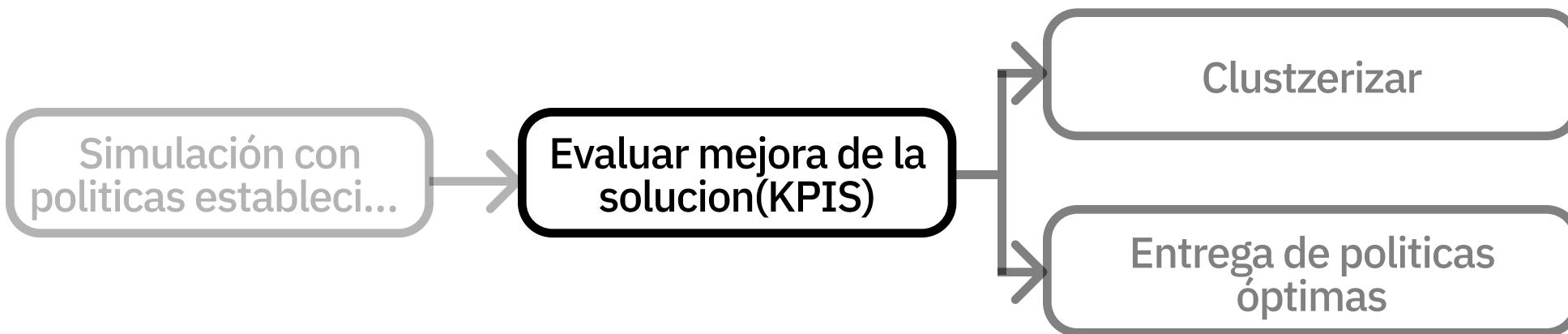
CASO BASE

KPI	Caso Base
Revenue Esperado [Millones de pesos]	61.500
Tasa de Curse	33%
Correos Promedio	2,15
Revenue Simulado [Millones de pesos]	61.400
Desviación Estándar [Millones de pesos]	48 (<0.1%)

CLUSTERIZACIÓN INICIAL

KPI	Cluster inicial
Revenue Esperado [Millones de pesos]	124.100
Tasa de Curse	53%
Correos Promedio	3,08
Revenue Simulado [Millones de pesos]	124.600
Desviación Estándar [Millones de pesos]	1.188(<1%)





Comparación KPIs

1. Ejecutivos aleatorios, envío máximo de correos a todos los clientes
2. Ejecutivos aleatorios, envío mínimo de correos a todos los clientes
3. Ejecutivos asignados a los clientes de mayor renta, envío máximo de correos
4. Ejecutivos asignados a los clientes mejores pagadores, envío máximo de correos

Comparación KPIs

Simulación con
políticas estableci...

Evaluar mejora de la
solucion(KPIs)

Clustzerizar

Entrega de políticas
óptimas

1. Ejecutivos aleatorios, envío máximo de correos a todos los clientes

KPI	Intuitivo 1
Revenue Esperado [Millones de pesos]	74.300
Tasa de Curse	33%
Correos Promedio	3,25
Revenue Simulado [Millones de pesos]	74.600
Desviación Estándar [Millones de pesos]	1.154(<2%)

Comparación KPIs

Simulación con
políticas estableci...

Evaluar mejora de la
solucion(KPIs)

Clustzerizar

Entrega de políticas
óptimas

2. Ejecutivos aleatorios, envío mínimo de correos a todos los clientes

KPI	Intuitivo 2
Revenue Esperado [Millones de pesos]	65.600
Tasa de Curse	33%
Correos Promedio	0
Revenue Simulado [Millones de pesos]	65.900
Desviación Estándar [Millones de pesos]	1.051(<2%)

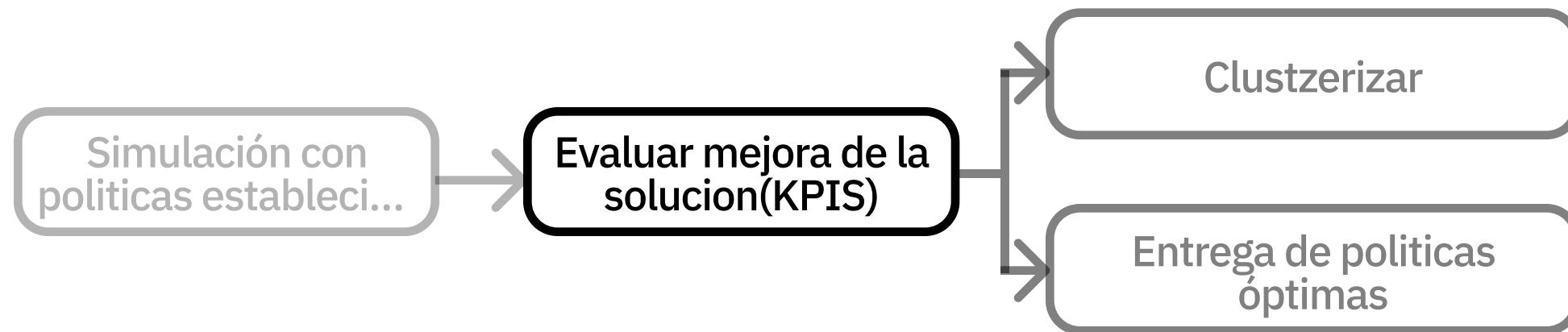
Comparación KPIs



3. Ejecutivos asignados a los clientes de mayor renta, envío máximo de correos

KPI	Intuitivo 3
Revenue Esperado [Millones de pesos]	130.000
Tasa de Curse	53%
Correos Promedio	3,25
Revenue Simulado [Millones de pesos]	130.700
Desviación Estándar [Millones de pesos]	1.106(<1%)

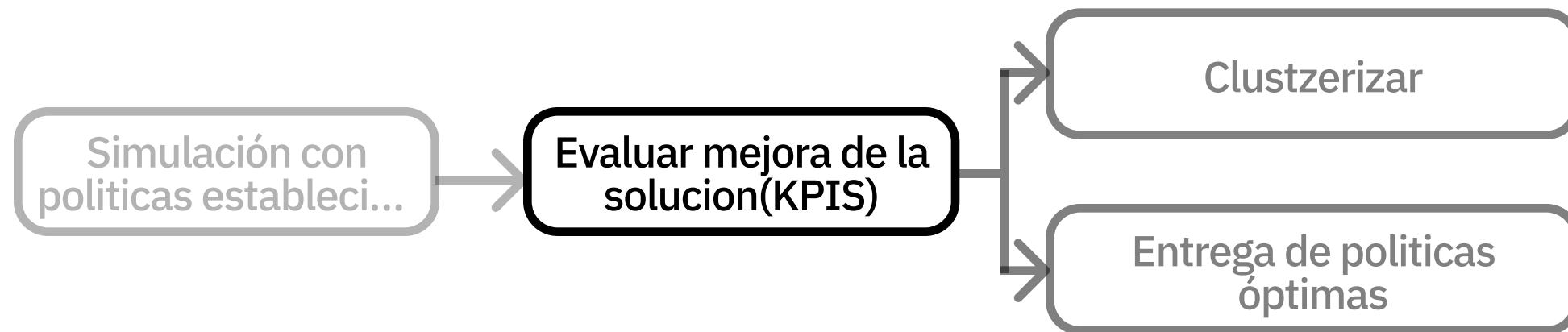
Comparación KPIs



4. Ejecutivos asignados a los clientes mejores pagadores, envío máximo de correos

KPI	Intuitivo 4
Revenue Esperado [Millones de pesos]	113.300
Tasa de Curse	53%
Correos Promedio	3,25
Revenue Simulado [Millones de pesos]	114.000
Desviación Estándar [Millones de pesos]	1.055(<1%)

Comparación KPIs



Mejor clusterización encontrada

Renta: 205,000 mejores rentas | Resto.

Probabilidad de No Pago: 205,000 mejores pagadores | Resto.

Elasticidad: Alta | Media | Baja.

Total: 12 grupos.

KPI	Mejor
Revenue Esperado [Millones de pesos]	144.500
Tasa de Curse	64%
Correos Promedio	3,27
Revenue Simulado [Millones de pesos]	145.123
Desviación Estándar [Millones de pesos]	1.336(<1%)

Comparación KPIs

Simulación con
políticas estableci...

Evaluar mejora de la
solucion(KPIs)

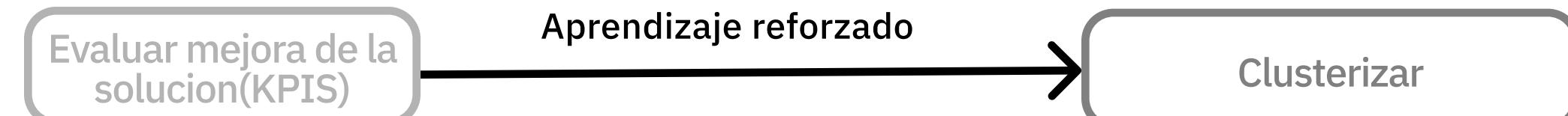
Estudio de variables
para clusterizar

Entrega de políticas
óptimas

KPI	Caso Base	Cluster inicial	Intuitivo 1	Intuitivo 2	Intuitivo 3	Intuitivo 4	Mejor
Revenue Esperado [Millones de pesos]	61.500	124.100	74.300	65.600	130.000	113.300	144.500
Tasa de Curse	33%	53%	33%	33%	53%	53%	64%
Correos Promedio	2.15	3.08	3.25	0	3.25	3.25	3,27
Revenue Simulado [Millones de pesos]	61.400	124.600	74.600	65.900	130.700	114.000	145.123
Desviación Estándar [Millones de pesos]	48 (<0.1%)	1.188<br (<1%)<="" b=""/>	1.154<br (<2%)<="" b=""/>	1.051<br (<2%)<="" b=""/>	1.106<br (<1%)<="" b=""/>	1.055<br (<1%)<="" b=""/>	1.336<br (<1%)<="" b=""/>

Tratamientos intuitivos:

1. Ejecutivos aleatorios, envío máximo de correos a todos los clientes
2. Ejecutivos aleatorios, envío mínimo de correos a todos los clientes
3. Ejecutivos asignados a los clientes de mayor renta, envío máximo de correos
4. Ejecutivos asignados a los clientes mejores pagadores, envío máximo de correos



Metodología de solución

APRENDIZAJE REFORZADO - MEJORAR CLUSTERIZACION

Acciones del agente:

- Agregar/quitar variables para la clusterización
- Elegir por cuantos tramos dividir y en qué lugares, algunas variables que lo permitan

Recompensa:

- Revenue esperado

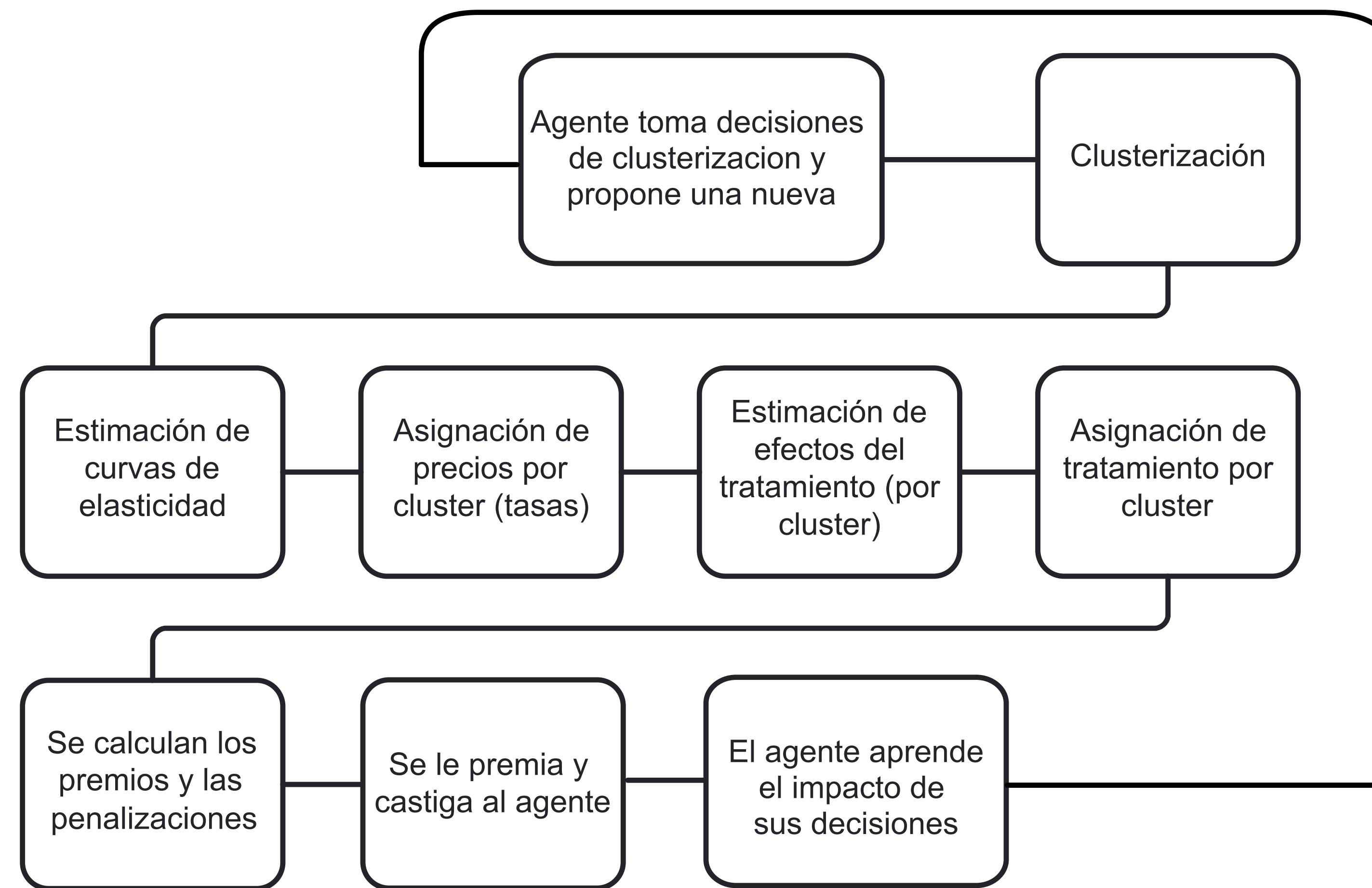
Penalización:

- Cantidad de clusters
- Ejecutivos no asignados



Metodología de solución

APRENDIZAJE REFORZADO - MEJORAR CLUSTERIZACION



26 DE NOVIEMBRE DEL 2024

Campaña comercial de retail banking

Rafael Antonio Fabio Aguilera

Itzae Abigail Flores Lara

Tarek Elías Hirmas Aboid

Felipe Santiago Nicolás Norambuena Palacios

Alon Menahem Rosenblum Keller

Ilan Enrique San Martín Kleinkopf

Profesor: Agustín Chiu - Ayudante: Lucas Salcedo

José Pedro Vargas Figueroa

Anexo

Estimación KPIs iniciales (formulas)

Revenue esperado: Ganancia esperada del banco con los créditos.

Revenue = **Costo total del crédito** - Monto Prestado

Revenue = (Plazo del crédito * **Valor fijo Cuota mensual**) - Monto Prestado

Revenue = (Plazo del crédito * Monto Prestado * $\frac{Tasa * (1+Tasa)^{Plazo_del_credito}}{(1+Tasa)^{Plazo_del_credito} - 1}$) - Monto Prestado

***Revenue esperado** = $\sum Revenue_i * (1 - P_{no_pago_i})$*

Revenue Esperado = 58.000 - 65.000 millones de pesos por mes

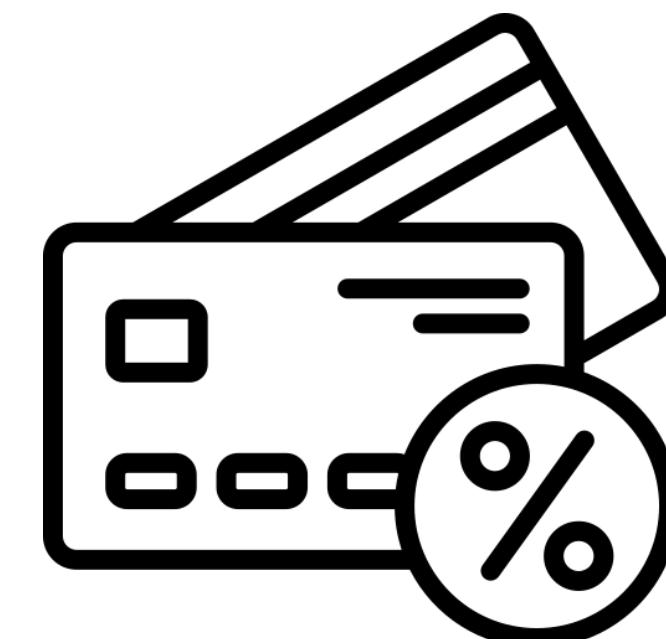
Estimación KPIs iniciales

Tasa de créditos cursados:

Créditos cursados = 42.000 - 43.000

Créditos simulados = 128.000 - 130.000

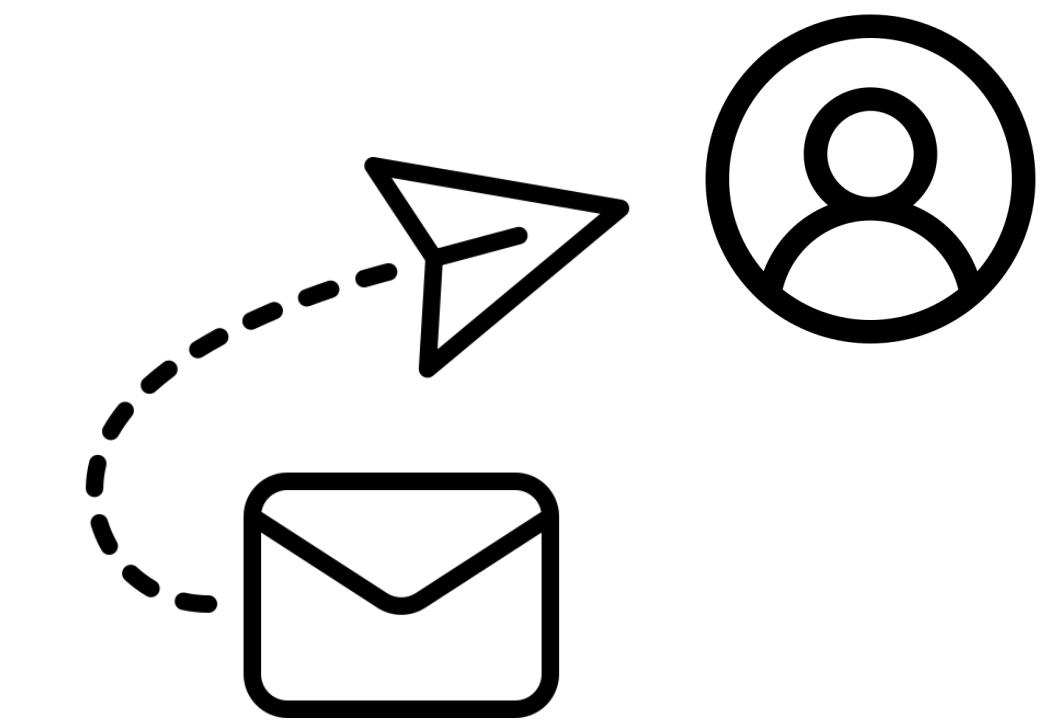
$$\text{Tasa de créditos cursados} = \frac{\text{Créditos cursados}}{\text{Créditos simulados}} \times 100 = 32\% - 34\%$$



Estimación KPIs iniciales

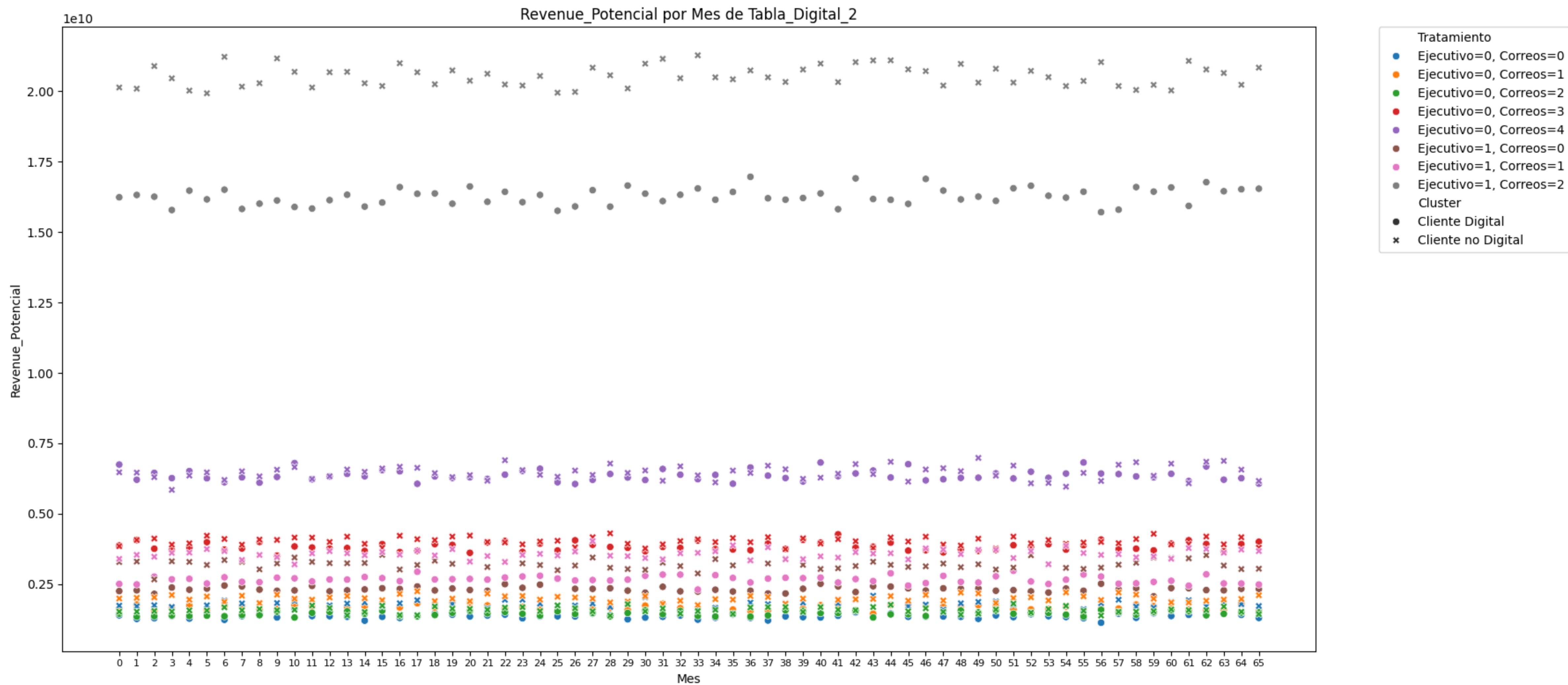
Correos promedio enviados:

$$\frac{\sum n_{correos_i}}{n} = 2.15 \text{ correos}$$

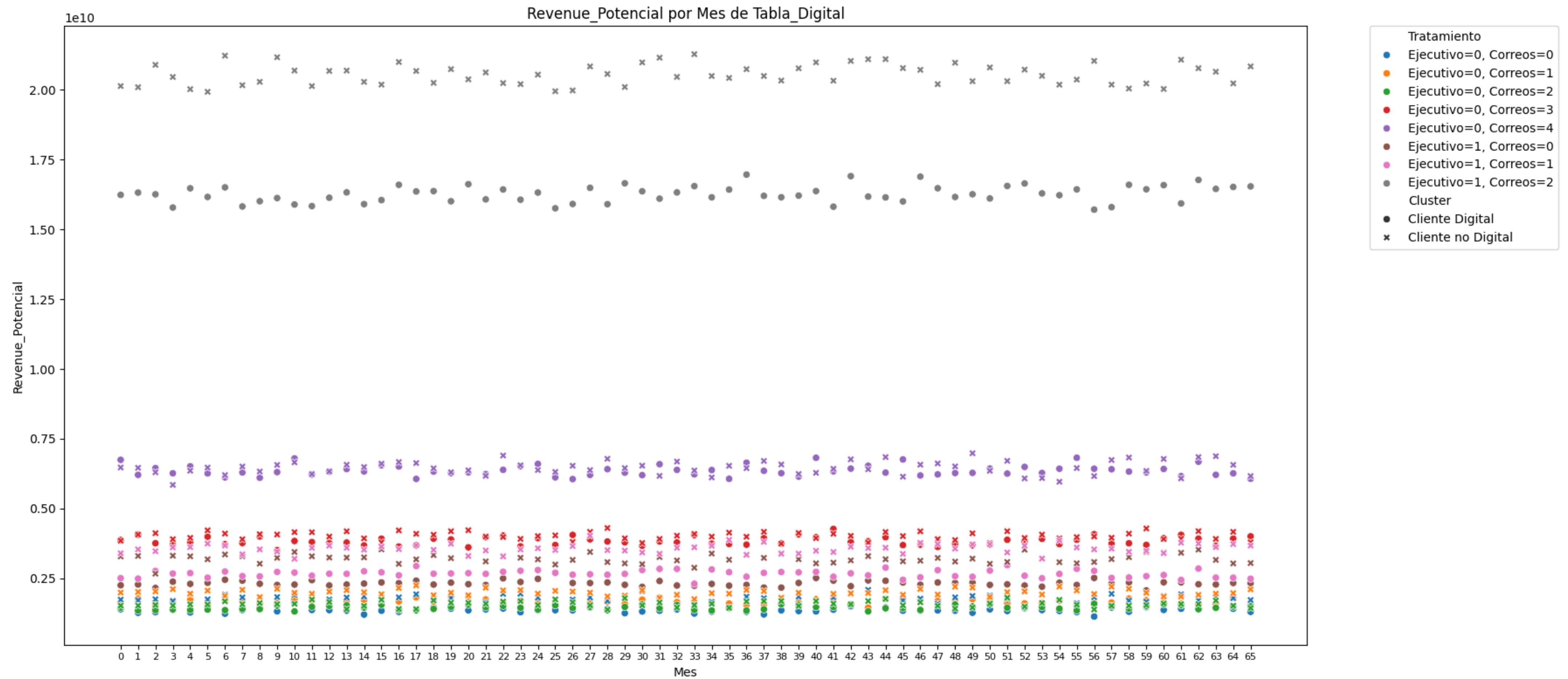


Revenue simulado y su desviación estándar

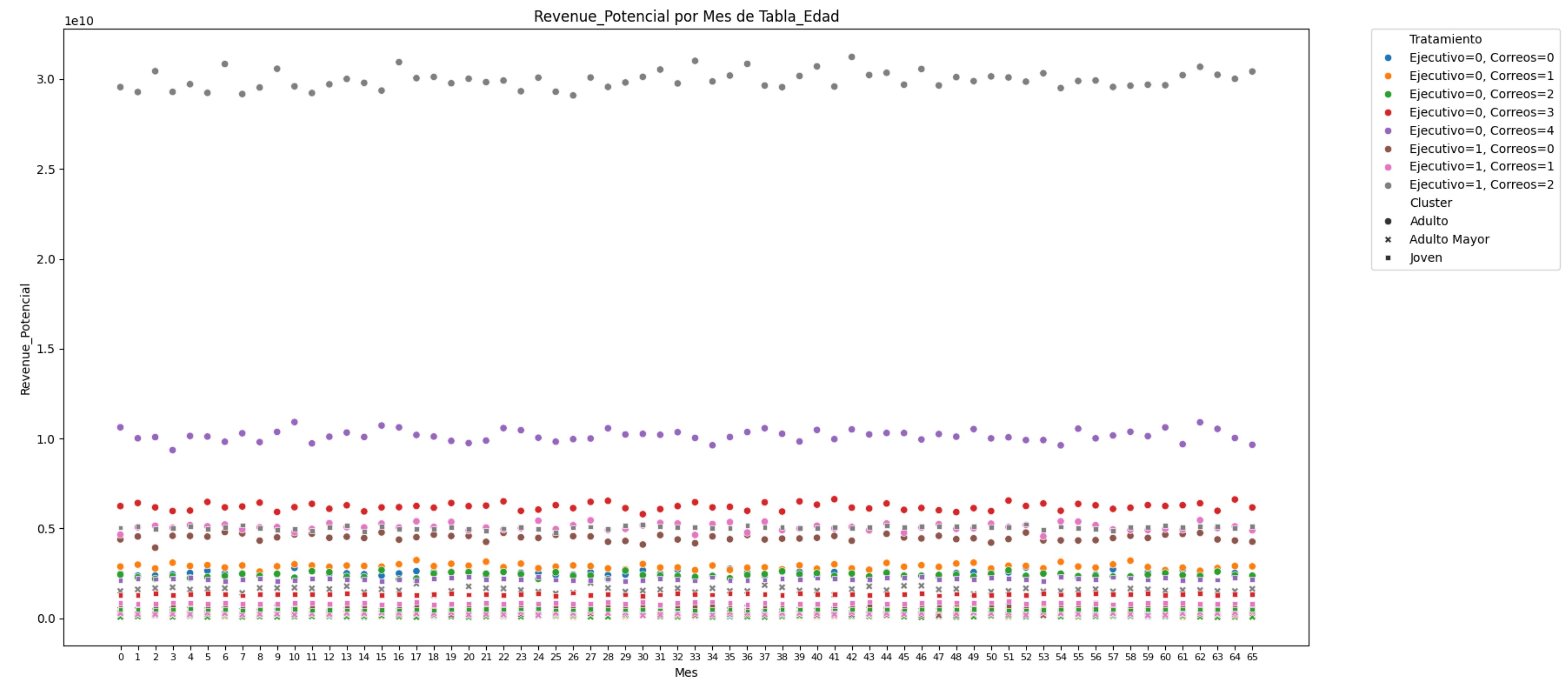
- **Revenue simulado:** Valores muy cercanos al revenue esperado
- **Desviación estándar:** 39 - 57 millones de pesos (0.1% de las ganancias mensuales)



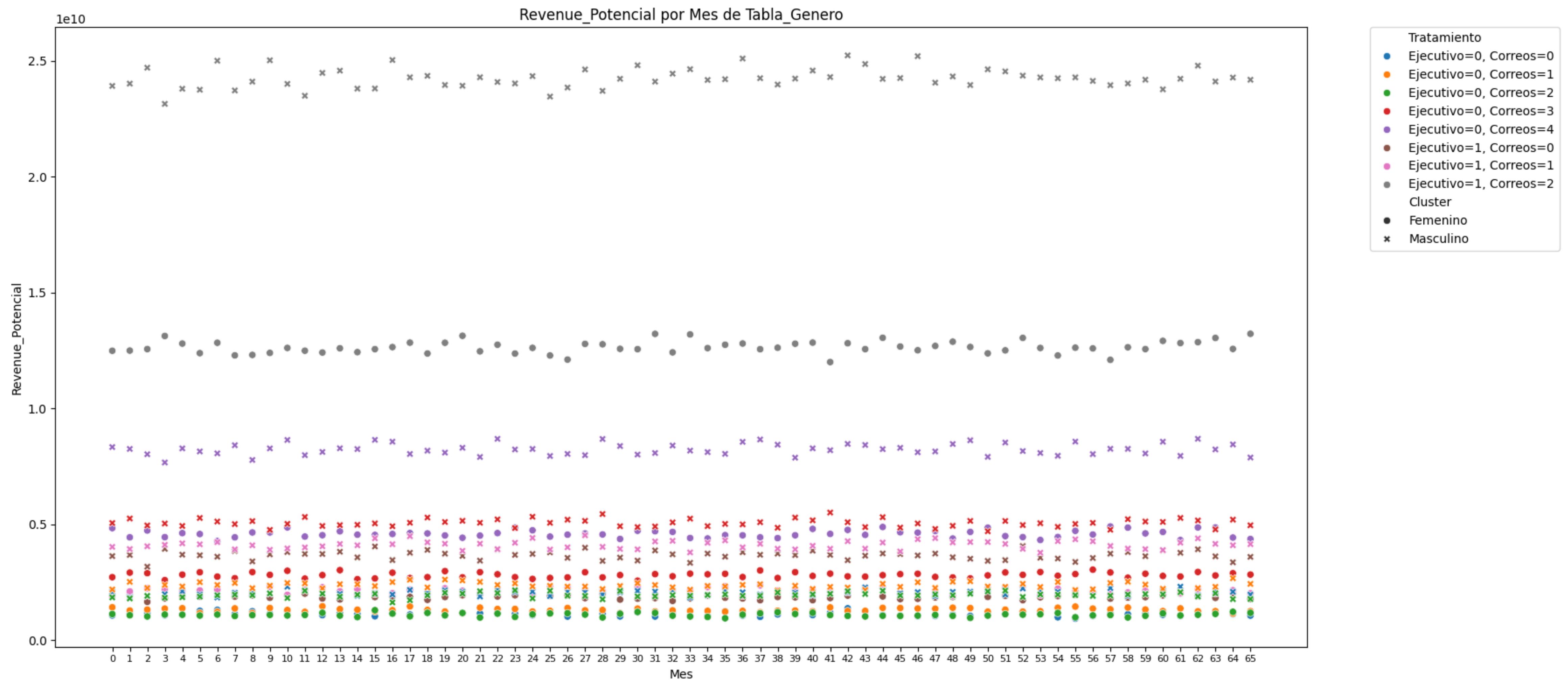
Cliente digital vs Revenue potencial



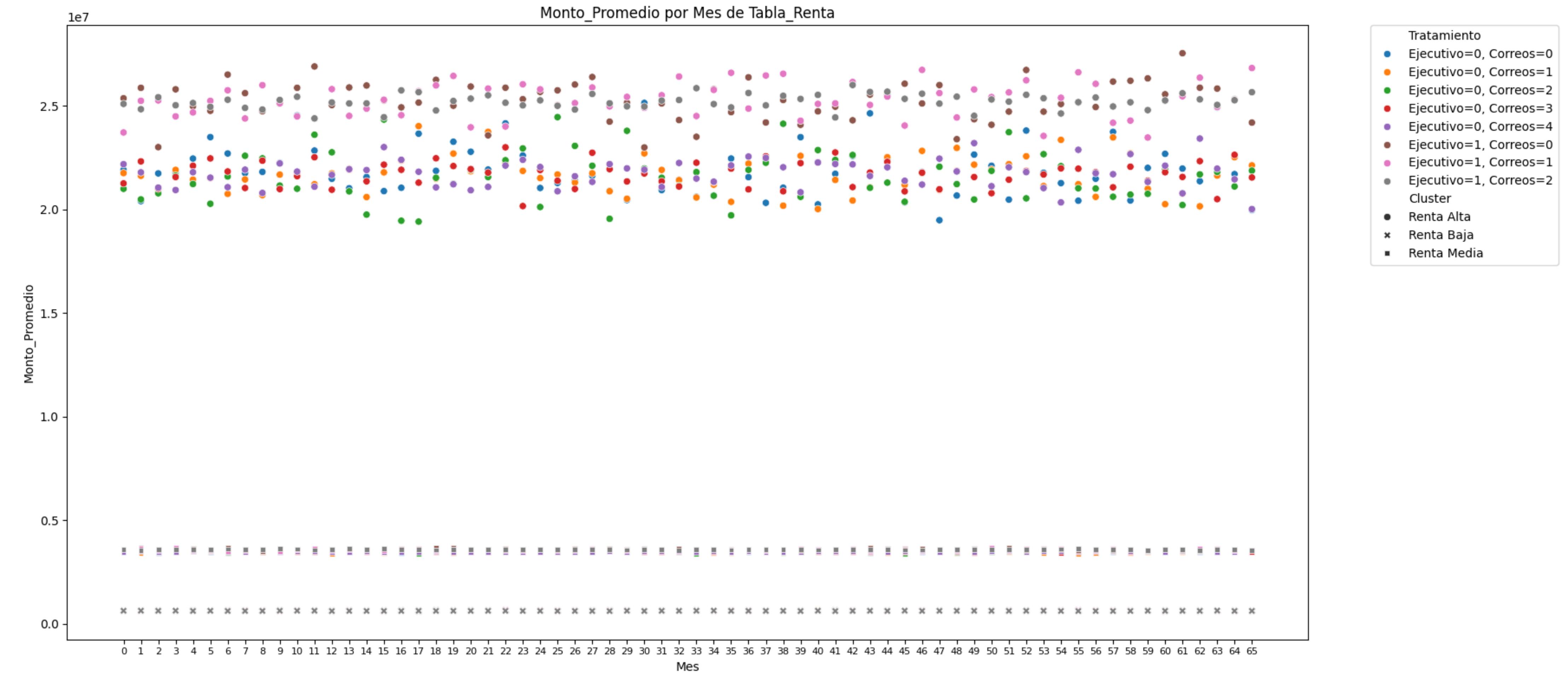
Edad vs Revenue potencial



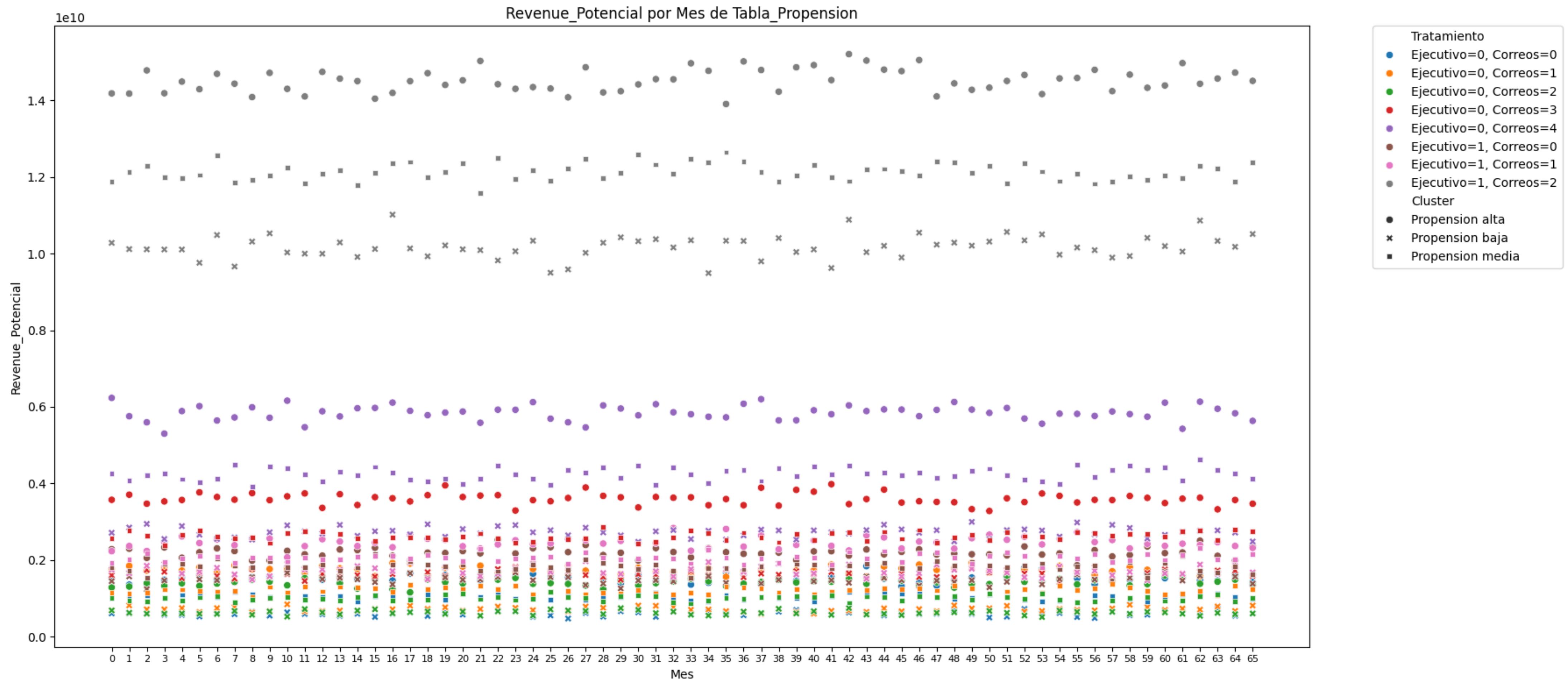
Genero vs Revenue potencial



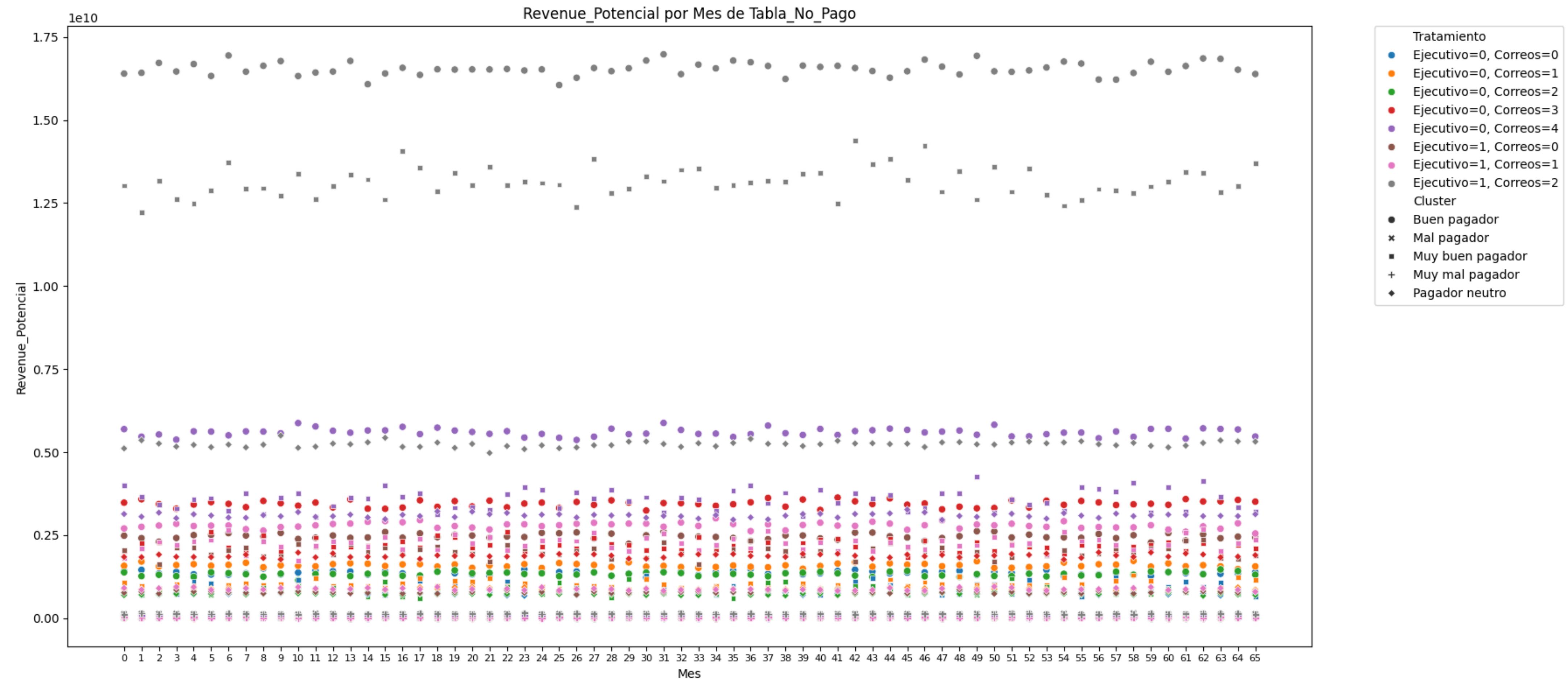
Renta vs Monto promedio



Propensión vs Revenue potencial



Prob. No pago vs Revenue potencial



Elasticidad vs créditos cursados

