

Valoración de modelos en campañas. Medición del churn en tarjetas de crédito

26 de mayo de 2021



Francisco J. Rodríguez Aragón

Responsable de Analytics en AFI

frrodriquez@afi.es





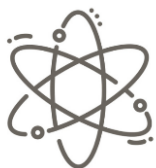
CONCEPTOS DE NEGOCIO

- Problema de negocio
- Definición de Churn
- Estrategia de negocio ¿Cómo reducir el Churn?



CONCEPTOS ESTADÍSTICOS

- Introducción a un modelo estadístico
- El concepto de lift acumulada
- Lift en el primer decil



SIMULACIONES DE RESULTADOS

- Hipótesis de partida
- Listado de ingresos y costes
- Escenarios de simulación



CALCULADORA SHINY DE SIMULACIÓN DEL CHURN

- Muy breve recordatorio de shiny
- Resolución de escenarios
- Detalle del código y de la implementación

CONCEPTOS DE NEGOCIO

CONCEPTOS DE NEGOCIO



- Problema de Negocio:
 - Muchos clientes dejan de usar nuestra tarjeta de crédito, durante 1 año o más
 - En general quién no usa durante más de 1 año la tarjeta, es raro que la use en un futuro inmediato
 - Vivimos del crédito: ¡Necesitamos que usen el modo crédito de la tarjeta! Los intereses están en torno al 20%
 - Es cierto que llegan nuevos clientes, pero la tasa de abandono del crédito supera a la captación

¿Cómo evitar la sangría de clientes?



- Definición de Churn
 - El Churn no es exactamente una fuga del cliente, el cliente definitivamente se fuga cuando se da voluntariamente de baja de una entidad, lo cual, no siempre es precisamente fácil (pregunten a los apóstatas o díglele a su compañía de teléfono que se quieren dar de baja, verán ...)
 - El Churn realmente es un abandono del uso de un servicio, en este caso se trata de algo muy particular como es el uso de una tarjeta a crédito
 - Un problema que aún el big-data no resuelve es ¿Qué cliente de mi base de datos está vivo o muerto? Los clientes Churn ¡Nos suelen avisar cuando mueren!, pero ahí están, en la base de datos ocupando su espacio y aún vivos (para que después digan que no hay vida después de la muerte, ...)
 - Para definir el churn en una entidad de crédito, hay que bajar al nivel transaccional. No suelen ser definiciones de lenguaje complejo, pero como bien saben lo que me conocen: “¡No hay nada fácil en la vida!, así que no os fieis del comercial que os diga, pero si eso es muy fácil, ... Si lo fuera lo haría él/ella, no os preguntaría”



- Definición de Churn: Problemas de construcción
 - En general su definición implica mover datos pesados y en ocasiones se está ante un problema de Big-Data, ya que hay que bajar a nivel de transacción y ver para cada cliente qué transacciones y de qué tipo realiza a lo largo de 1 año
 - Las tarjetas que se ofrecen a los clientes en general tienen modalidades a débito, a crédito y en ocasiones opciones mixtas que pueden generar o no algún tipo de interés. Por tanto debe existir la posibilidad de distinguir por tipología de transacción
 - Otra dificultad que se va a encontrar el DS acerca de la validez de la transacción a crédito: ¿Qué ocurre con las transacciones erróneas? ¿Cómo tratar los abonos cuando ha habido correcciones? ¿Cómo se corrigen duplicados en las transacciones? ¿Y si ha habido un fraude? ¿Cómo se detecta?

El punto final de todo esto debe ser algo así como saber si por ejemplo un cliente que ha sido cliente activo durante el 2019, lo sigue siendo en el 2020 y para ello hay que ver si al menos ha hecho una transacción correcta en el 2020, además en ocasiones se impone algún límite de que al menos la suma de todas las transacciones llegue al menos a 100€ o 200€



- Definición de Churn: Cliente Activo
 - Intrínsecamente unido a la definición de Churn, está el concepto de cliente activo y por tanto Churn se puede re-definir como aquel cliente activo que deja de estarlo a lo largo de 1 año, el siguiente (en ocasiones el período puede cambiar según la entidad)
 - ¿Qué es un cliente activo? En este caso hay que definir ya una ventana temporal y conviene fijar una fecha como por ejemplo el 31-12-2019:
 - Una posible definición de Cliente Activo es como aquel que desde el 01-01-2019 hasta el 31-12-2019 ha realizado al menos 5 transacciones de crédito y la suma de los importes de ésta superan los 100€
 - Consecuentemente una posible definición de Churn sería como un cliente es churn si desde el 01-01-2020 y el 31-12-2020 ha hecho menos de 4 transacciones o el importe de ambas es menor a 100€

CONCEPTOS DE NEGOCIO



- Estrategia de negocio I

Aquí llega el momento en el que me da la risa tonta !!!!!

- Posibles estrategias de negocio de los más reputados clientes:
 - Quiero que no se vayan los clientes, apliquemos nuestros modelos de alta predictividad para evitarlo



Vamos como si los modelos fueran una especie de mafia siciliana que se pasan por tu casa y si no compras te sueltan un rigostrio

- Paso de modelos, ¡son cosas del demonio!, démosle un descuento del 20% en su próxima compra a todo cliente

La gran mayoría de los clientes tienen un lío de datos que impiden construir la target adecuada, o sencillamente son vagos y ¡ala! A gastar pasta como si no hubiera mañana

CONCEPTOS DE NEGOCIO



Aquí llega el momento en el que me da la risa tonta !!!!!

- Estrategia de negocio II

- Hay que definir bien qué se pretende conseguir en la campaña y cómo se va a medir
- La primera pregunta es ¿Qué motiva en general que un cliente siga utilizando nuestros servicios? En el tema de crédito hay que tener en cuenta que el uso es por necesidad de capital transitoria, obviamente, si su economía mejora, es posible que no lo necesite, por lo que hay una bolsa de clientes que sí o sí van a hacer churn
- Sobre lo que realmente se puede impactar es sobre una “palanca” que ante una competencia de productos similares, el nuestro se diferencie en algo. ¿Ofrezco un bono de 10€ de regalo en la próxima compra? ¿Ofrezco un descuento en el tipo de interés a lo largo de 1 año? ... Claramente esta acción tiene un coste que debe conocerse e implementarse en la simulación
- Y una pregunta importante resulta ser ¿Voy a toda la población o a parte de ella? Obviamente, si vas a lanzar la campaña a toda la población, no necesitas modelos, ni DS ni nada, ¡¡¡Te gastas la pasta y a rular...!!!!



CONCEPTOS DE NEGOCIO



- Estrategia de negocio III

Aquí llega el momento en el que me da la risa tonta !!!!!

- No siempre es factible el uso de modelos, sólo cuando se quiere actuar sobre una parte de la población, bien porque se tenga un presupuesto limitado, bien porque se quiera y desee ahorrar costes tratando de maximizar ingresos

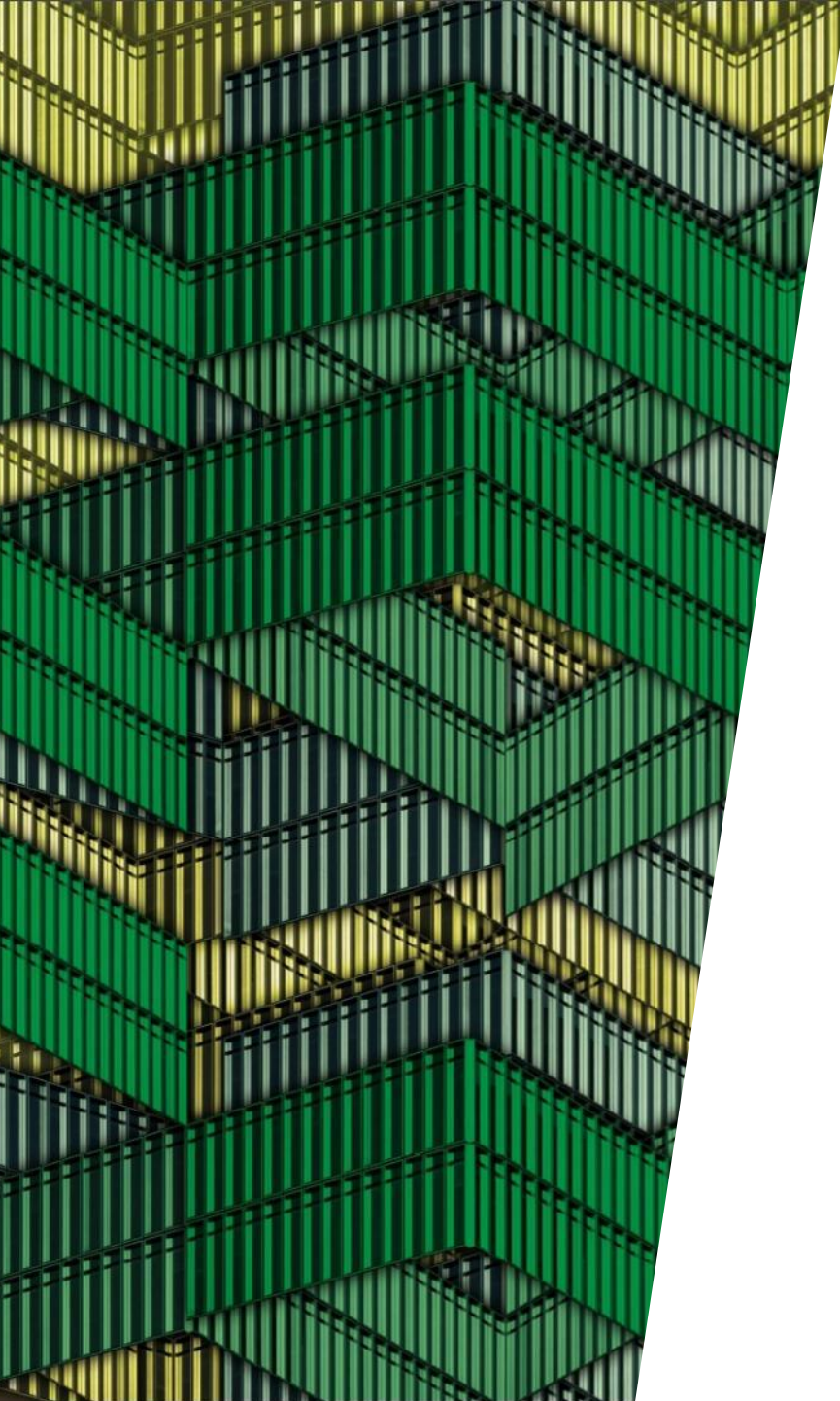
En general resulta triste ver cómo en muchas empresas, lo único que se hace es que te dan un presupuesto de x€ y te lo gastas para enviar push, bien con sms o con mails a los correos de tus clientes y achicharrarlos sin un objetivo claro en la vida

Política Comercial:

“La flagoneta la tengo llena de malacotones mu buenos y baratos”



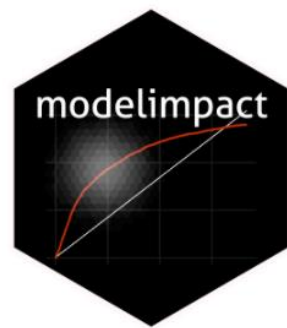
- DS: Nuestros modelos tienen una ROC del 90% son la po...
- CEO: Sí, pero los clientes se van
- Director Comercial: ¡Ja, me maaaten!



CONCEPTOS ESTADÍSTICOS



- Introducción a un Modelo estadístico
 - Aquí se van a considerar sólo la modelización de una variable binaria tipo 0 – 1, donde 1 es que se produce churn y 0 es que no tiene lugar lo anterior a lo largo de un horizonte temporal
 - Unos posibles modelos a considerar son toda la amplia gama de aquellos que generan para un determinado registro (cliente) un número entre 0 y 1 que se identifica con la probabilidad de ocurrencia del evento churn
 - Hace poco salió una librería en R que se denomina *modelimpact* que hace parte de lo que aquí se va a tratar pero desde un enfoque aún más purista al considerar la típica tabla 2 x 2 <https://github.com/PeerChristensen/modelimpact>



	+	-
+	Verdaderos positivos (VP)	Falsos positivos (FP)
-	Falsos negativos (FN)	Verdaderos negativos (VN)

En ocasiones el score del modelo se transforma a un 0 – 1 predicho, por lo que “a toro corrido”, se puede ver la efectividad del acierto en una tabla 2x2, sin embargo ... ¡Ojo con la interpretación! ¿Qué ocurre si el 99% de la muestra es 0 y el 1% es 1?



- Introducción a un Modelo estadístico: Una medida clásica, la curva ROC I
 - Un primer punto es cómo medir la efectividad de un modelo. Si se está ante la situación anterior, se podría crear un modelo que dijese que todo el churn es 0 y tendríamos un resultado aproximado al siguiente

Dataset con un churn del 1%		Real	
		0	1
Predicción	0	99%	0,50%
	1	0,50%	0%

- Por tanto, medir el mero acierto bajo una tabla 2x2 no es una buena idea para saber si un modelo funciona o no y es que lo que se quiere es detectar efectivamente ese 1%
¿Cómo saber entonces si el modelo los detecta o no?

Sensibilidad	99%			
1 - Especificidad	1			
¡Con el 100% de 1-especificidad, no se consigue el 100% de sensibilidad!!!!!!!!!!				



- Introducción a un Modelo estadístico: Una medida clásica del accuracy, la curva ROC
 - Construcción ROC:
 - Supongamos que se divide el espacio de probabilidad en 100 partes iguales. De este modo se obtienen 99 valores entre 0 y 1 que separan esos datos
 - Considérese cada punto como un umbral, el que sea que puede ser por ejemplo 99% y calcúlese los siguientes parámetros a partir de una tabla 2x2

	+	-
+	Verdaderos positivos (VP)	Falsos positivos (FP)
-	Falsos negativos (FN)	Verdaderos negativos (VN)

Ratio VPR = Verdaderos Positivos / Total

Ratio FPR = Falsos Positivos / Total

Sensibilidad = VPR

1 - Especificidad = FPR

- Para cada umbral, habrá un recorrido en 100 tablitas 2x2 donde se calculan los pares Sensibilidad y 1 - Especificidad, expresan el intercambio de errores de clasificar como evento churn, aquello que no es y por tanto existirá un intercambio de costes

CONCEPTOS ESTADÍSTICOS



- Introducción a un Modelo estadístico: Una medida clásica, la curva ROC II
 - Ejemplo sencillo en Excel con 4 umbrales para comparar un buen modelo, uno mediocre y aleatoriedad

Cliente	PD	Target casi-Perfecto	Target 2	Target 3
1	0,8	1	1	0
2	0,7	1	1	0
3	0,7	1	0	0
4	0,5	0	0	1
5	0,3	0	1	0
6	0,2	0	0	0
7	0,1	0	0	1
8	0,09	0	0	0
9	0,08	0	0	0
10	0,03	0	0	1
11	0,02	0	0	0
12	0,01	0	0	0

Suponiendo (lo cual es imposible), que un modelo escupe las mismas probabilidades para 3 targets distintas ¿Qué modelo es mejor? ¿Cómo se dibuja la curva ROC? ¿Cómo se mide el intercambio acierto-fracaso?

CONCEPTOS ESTADÍSTICOS



- Concepto de lift acumulada. El concepto
 - La idea que se explota en este trabajo entra mejor en la cabecita de los CEOs, CIOs, CTOs ... y todos aquellos que piensan que hacer algo más allá de una media es algo al alcance de unos pocos iluminados
 - Lift en inglés significa algo así como elevador, ascensor, ... No me refiero aquí a hacernos un lifting que a más de uno no nos iría mal, aunque alguna relación tiene el concepto



- En el caso de los modelos, la lift tiene que ver con “cuánto se eleva el grado de acierto de aplicar un modelo frente a no aplicar nada”



- Concepto de lift acumulada. La construcción no acumulada
 - Se supone que existe un modelo y por tanto una ordenación de datos por ejemplo en términos de probabilidad de mayor a menor ocurrencia del evento churn
 - Posteriormente se crea un número de segmentos, un caso particular es si dividimos la población por deciles con 9 puntos de corte: $[0; p_1)$, $[p_1; p_2)$ $[p_9; 1]$. Claramente existe prácticamente la misma cantidad de individuos con probabilidad entre 0 y p_1 que entre p_6 y p_7
 - En cada tramo se puede calcular la tasa de churn sobre el total que existe por lo que se obtendrían 10 tasas: t_1, t_2, \dots, t_{10}
 - Por otro lado, ante la inexistencia de un modelo, cabe suponer que la cantidad de Churn ante cualquier selección aleatoria que divida a la población en 10 grupos, debería distribuirse homogéneamente, por tanto si hay una cantidad de churn igual a c_{Total} y se escoge una muestra del 10% de los datos SIN USAR Modelo, la cantidad de datos que debería haber en esa muestra es $c_{Total} / 10$
 - Pues bien se denomina lift no acumuladas a las 10 lifts que se crean del siguiente modo:
$$L1_{na} = t_1 / (c_{Total} / 10), L2_{na} = t_2 / (c_{Total} / 10), \dots, L10 = t_{10} / (c_{Total} / 10)$$



- Concepto de lift acumulada. La construcción en versión acumulada
 - La lift acumulada se realiza de análogo modo sólo que se va acumulando las cantidades de datos con forme se avanza en los deciles
 - $Ln_a = (\text{Tasa de Churn acumulada desde el decil 1 hasta el decil } n) / (n \cdot c\text{Total}/10)$

Con $n = 1, 2, 3, \dots, 10$

¿Cuánto debe valer $L10_a = ?$ ¿Cómo se espera que sea la forma de la curva lift acumulada? ¿Creciente o decreciente ante un buen modelo?



- Lift en el primer decil
 - Una medida habitual que te entienden algunos directores, tras arduo esfuerzo de comprensión es la Lift en el primer decil, es decir, la lift en ese decil conviene que sea lo mayor posible porque estaría indicando sobre cómo es capaz de “elevar” la tasa de acierto el modelo en un 10% de los datos, frente a elegir una cantidad análoga de población de modo totalmente aleatorizado
 - Por supuesto también se puede pensar en divisiones más finas como vigentiles o percentiles y entonces tomar la lift del primer vigentil y del primer percentil, pero tomándola del primer decil, en la mayoría de los modelos tiene su sentido
 - Un modelo cabe ser considerado como que es capaz de ordenar cuando, en general, la lift del primer decil supera el valor de 2.5, en algunos casos (que no se discuten aquí), ese valor hay que cambiarlo porque exista una elevada tasa de evento a predecir

CONCEPTOS ESTADÍSTICOS



- Lift en el primer decil Ejemplo de una distribución realizada por un modelo ficticio en deciles

decil	population	non_acum_cl	churn	lift
1	50000	20000	20000	4,0
2	100000	10000	30000	3,0
3	150000	7500	37500	2,5
4	200000	500	38000	1,9
5	250000	1000	39000	1,6
6	300000	1000	40000	1,3
7	350000	2000	42000	1,2
8	400000	2000	44000	1,1
9	450000	3250	47250	1,1
10	500000	2750	50000	1,0

Preguntas de examen: ¿Cuántos datos hay? ¿Cuál es la tasa de fuga? ¿Qué proporción de fugados se detectan en el primer decil frente al total de fugas? ¿Cuántos fugados habría en el primer decil si no existiese modelo estadístico?



SIMULACIÓN DE RESULTADOS

SIMULACIÓN DE RESULTADOS



- Hipótesis de Partida. Hipótesis exógenas
 - Muchas de las hipótesis que se van a considerar aquí suelen ser conocidas por los usuarios de negocio y es conveniente utilizarlas para tener al menos una primera aproximación a al valor económico que puede aportar un modelos
 - En este caso se parte de los siguientes hipótesis de negocio que vienen dadas:
 - Existe una población objetivo total de 1.000.000 clientes activos en el momento de lanzar la campaña
 - Se espera, por experiencia, que un 20% de esos clientes, es decir, unos 200.000 harán churn
 - Se supone que un cliente que no hace churn, en general hace uso de unos 1.000€ de compra a crédito

SIMULACIÓN DE RESULTADOS



- Hipótesis de Partida. Parámetros de escenario
 - Se introducen algunos parámetros que el usuario puede manipular con incidencia en el resultado final coste-beneficio de la acción:
 - Se supone un presupuesto máximo de unos 1.000.000€ para la campaña
 - Costes: La campaña consiste en enviar un regalo al cliente que se va a fugar como un bono a utilizar a través del crédito de valorado en unos 20€ (se supone que no hay más costes que influyen en el presupuesto y no se incluyen por tanto los costes fijos)
 - Ingresos: En estas compras a crédito cabe aplicar un interés TAE del 20% sobre el total del solicitado a crédito

SIMULACIÓN DE RESULTADOS



- Hipótesis de Partida. Hipótesis comercial
 - La dirección comercial piensa que la efectividad en la retención de aquel a quién le llegue el obsequio está en torno al 5%, es decir, si alguien si iba a fugar le llega la campaña, hay un 5% de probabilidad de que dicho cliente no se llegue a fugar
 - Además suponen que aquellos clientes que no se fugan debido a la campaña, van a tener en media, un gasto similar al caso sin fuga, es decir, de al menos unos 1.000€

SIMULACIÓN DE RESULTADOS



- Listado de Ingresos y Costes
 - Ingresos:
 - Se supone que el ingreso de la campaña se mide a partir del gasto por tipo de interés que harían aquellos que se iban a fugar y que dejan de hacerlo
 - Puede ser que los clientes que acepten la campaña hagan algún gasto adicional vía crédito que reporte beneficio pero no se considera
 - Puede existir un ingreso por vía de mejora reputacional a los impactados pero no se considera
 - Gastos:
 - El gasto principal es el coste de la promoción que se envía al cliente
 - Puede existir un coste adicional por gastos administrativos de gestión de la campaña que no se considera aquí
 - Se supone que el cliente no tiene un equipo de DS y por tanto habrá que añadir un coste de desarrollo de un modelo de fuga que se estima en unos 25.000€

SIMULACIÓN DE RESULTADOS



- Escenarios de simulación. Un poco de publicidad sobre el blog de AFI ...


<http://analytics.afi.es/AfiAnalytics/blog/1491511/1491517/0/0/0/analyticsblog.html>



Escenario 1: Se gasta todo el dinero de la campaña en una muestra aleatoria de clientes, escogiéndose por tanto a $1.000.000\text{€} / 20\text{€/cli} = 50.000$ clientes

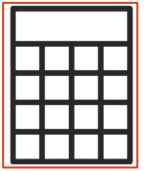
Escenario 2: El cliente se plantea construir un modelo y lo hace con AFI, que le ofrece un modelo con una lift de 4 en el primer decir por el módico precio de 25.000€

¿Qué hace el cliente? ¿Nos compra o no?



CALCULADORA SHINY DE SIMULACIÓN DEL CHURN

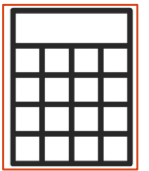
CALCULADORA SHINY DE SIMULACIÓN DEL CHURN



- Muy breve recordatorio de Shiny

<https://shiny.rstudio.com/gallery/>

CALCULADORA SHINY DE SIMULACIÓN DEL CHURN



- Resolución de Escenarios. Se realiza el cálculo sin tener en cuenta el coste de los servicios de consultoría

Parámetros

Número de Clientes

Presupuesto (en euros) Campaña

Gasto mínimo por cliente y año

Tipo de Interés revolving

0 20 25

0 3 6 9 12 15 18 21 24

Coste por impacto aceptado

% Fugas en la población

0 20 100

0 10 20 30 40 50 60 70 80 90 100

% de activación del impacto

0 5 100

0 10 20 30 40 50 60 70 80 90 100

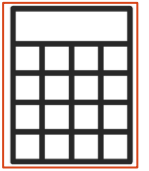
	variables	value
1	Ingreso sin fugas y sin campañas (todos usan la tarjeta)	200000000
2	Coste de oportunidad debido a las fugas	40000000
3	Número máximo de clientes a impactar	50000

Showing 1 to 3 of 3 entries

Previous 1 Next

Hay unos 40.000.000 € en juego, es decir, esos 200.000 clientes que se fugan, si se gastasen 1000€ a tipo de interés el 20%, es una auténtica pasta !!!!!!!

CALCULADORA SHINY DE SIMULACIÓN DEL CHURN

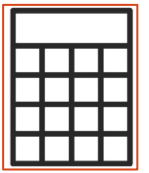


- Resolución de Escenarios. Beneficios sin modelo vs modelo con lift 4

Escenario 1 (no modelo)			Escenario 2 (lift = 4)		
Show 10 ▾ entries			Show 10 ▾ entries		
Search: <input type="text"/>			Search: <input type="text"/>		
	variables	value		variables	value
1	Fugas detectadas	10000	1	Fugas detectadas	40000
2	Fugas evitadas	500	2	Fugas evitadas	2000
3	Clientes sin fuga	40500	3	Clientes sin fuga	12000
4	Coste final	810000	4	Coste final	240000
5	Ingreso esperado extra	100000	5	Ingreso esperado extra	400000
6	Benefico esperado campaña	-710000	6	Benefico esperado campaña	160000

Con el modelo consigo ganar 160.000€, claramente **¡Esto no es Jauja!**, pero tras gastar 1.000.000€ cubro costes, pago 25.000 al consultor, yo contento, el cliente contento y nos vamos de cañas con los 135.000€ limpios

CALCULADORA SHINY DE SIMULACIÓN DEL CHURN

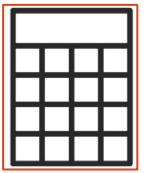


- Resolución de Escenarios.

Escenario 1 (no modelo)			Escenario 2 (lift = 4)		
Show 10 ▾ entries			Show 10 ▾ entries		
Search: <input type="text"/>			Search: <input type="text"/>		
	variables	value		variables	value
1	Fugas detectadas	10000	1	Fugas detectadas	40000
2	Fugas evitadas	500	2	Fugas evitadas	2000
3	Cientes sin fuga	40500	3	Cientes sin fuga	12000
4	Coste final	810000	4	Coste final	240000
5	Ingreso esperado extra	100000	5	Ingreso esperado extra	400000
6	Benefico esperado campaña	-710000	6	Benefico esperado campaña	160000

Un beneficio adicional del modelo es se ha conseguido evitar 2.000 de las 200.000 fugas posibles, se ha detenido significativamente la sangría de fugas en un 1%, mientras que la selección aleatoria sólo habría conseguido parar sólo a 500, es decir, a un 0.25% de las fugas totales

CALCULADORA SHINY DE SIMULACIÓN DEL CHURN



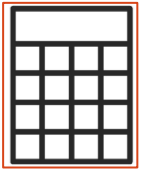
- Resolución de Escenarios. El gasto óptimo ¿Cuánto tengo que gastar para maximizar el beneficio bajo estas hipótesis?

El óptimo bajo estas hipótesis estaría en torno a 2.000.000€

Escenario 1 (no modelo)			Escenario 2 (lift = 4)		
Show 10 ▾ entries			Show 10 ▾ entries		
Search: <input type="text"/>			Search: <input type="text"/>		
variables		value	variables		value
1	Fugas detectadas	20000	1	Fugas detectadas	80000
2	Fugas evitadas	1000	2	Fugas evitadas	4000
3	Cientes sin fuga	81000	3	Cientes sin fuga	24000
4	Coste final	1620000	4	Coste final	480000
5	Ingreso esperado extra	200000	5	Ingreso esperado extra	800000
6	Benefico esperado campaña	-1420000	6	Benefico esperado campaña	320000

¿Por qué el óptimo está justo en 2.000.000€? ¿Qué está ocurriendo realmente? ¿Cómo afinar la calculadora?

CALCULADORA SHINY DE SIMULACIÓN DEL CHURN



- Detalle del código y de la implementación. Un poco de tripas en R ...



© 2021 Afi. Todos los derechos reservados.