

Dataton 2024

Data Ninjas

Estimación del ROI de las
compensaciones



Planteamiento del problema

Nuestra política de compensaciones busca mitigar el impacto negativo en el reorder y la frecuencia de usuarios tras una mala experiencia.

Dado el presupuesto limitado, no todas las experiencias negativas se compensan; la política actual utiliza características tanto de la orden como del usuario para decidir. ¿Cómo podríamos evaluar en el corto plazo si esta política realmente reduce la pérdida de órdenes?

En cifras Latam

Acumulado 2024

	Monto USD
Compensaciones	\$7,267,475
Refunds	\$23,484,859
Total	\$30,752,334



Hipótesis y alcances

- ¿Una compensación o refund ayuda a mantener o aumentar la frecuencia de consumo de un usuario?
- ¿Ese delta en frecuencia es rentable o no?
- ¿Es el ROI diferente para los tipos de usuario?
- ¿Qué recomendaciones podemos dar al esquema de compensación?

El límite del proyecto será primero tratar el problema a nivel país (El Salvador) y posteriormente, replicarlo a distintos mercados.





Metodología

Proponemos hacer un backtesting antes de realizar un A/B tradicional. Podemos aprovechar que en la historia hemos compensado algunas malas experiencias y otras no.

- ★ Construir la base de usuarios -compensaciones que dimos en julio-2024. Why Jul? porque julio y agosto fueron muy similares en frecuencia de orden por usuario. Septiembre es atípico en SV
- ★ Construir para esos usuarios la frecuencia en órdenes +/- 1 mes de la compensación
- ★ Construir la base de usuarios que no compensamos en julio-2024
- ★ Construir para estos usuarios la frecuencia en órdenes +/- 1 mes del evento no compensado
- ★ Comparar peras con peras: crear una segmentación de usuario que nos permita ser homogéneos en la medición de la comparación
- ★ Construir un modelo estadístico que nos diga si es significativo el re-order entre un usuario compensado y no, por tipo de usuario segmentado
- ★ Analizar y entender próximos pasos (ej, un A/B testing)













Segmento, luego existo

Creamos una segmentación para los usuarios de SV - Ampliamente usada en todas las industria

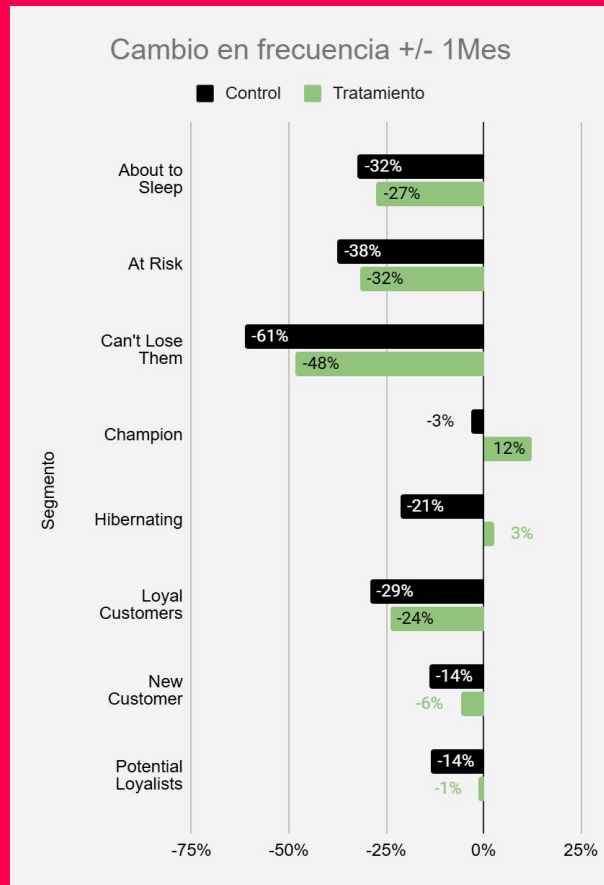
RFM

Recency Frequency Monetary

Segmento	# Usuarios	% Usuarios	Frecuencia	Ticket
 New Customer	12,040	5.1%	1.04	\$ 9.34
 Champion	29,617	12.5%	10.28	\$ 13.73
 Loyal Customers	40,547	17.1%	6.45	\$ 14.61
 Potential Loyalists	29,306	12.4%	2.26	\$ 13.39
 Can't Lose Them	2,726	1.1%	6.29	\$ 15.95
 About to Sleep	30,786	13.0%	1.76	\$ 12.13
 At Risk	26,439	11.1%	2.75	\$ 15.49
 Hibernating	65,699	27.7%	1.15	\$ 12.30
Total	237,160	100.0%	3.64	\$ 13.23



Exploratorio (peras con peras)



Recibieron comp (Tratamiento)

Segmento	Usuarios	Freq Antes	Freq Después	Var
About to Sleep	281	2.0	1.5	-27%
At Risk	94	2.5	1.7	-32%
Can't Lose Them	22	4.8	2.5	-48%
Champion	8720	9.7	10.9	12%
Hibernating	44	1.6	1.6	3%
Loyal Customers	4447	5.7	4.3	-24%
New Customer	69	1.9	1.8	-6%
Potential Loyalists	877	2.2	2.2	-1%

No Recibieron comp (Control)

Segmento	Usuarios	Freq Antes	Freq Después	Var
About to Sleep	164	2.4	1.7	-32%
At Risk	90	2.9	1.8	-38%
Can't Lose Them	14	6.1	2.4	-61%
Champion	2195	8.3	8.1	-3%
Hibernating	79	2.0	1.6	-21%
Loyal Customers	1283	5.0	3.5	-29%
New Customer	27	1.7	1.5	-14%
Potential Loyalists	338	2.5	2.1	-14%





Nos ponemos los lentes y serios

Modelo de diferencias en diferencias para cada segmento, para entender si estadísticamente es significativa la variación en frecuencia

```
# Champion
import statsmodels.formula.api as smf
model_champion = smf.ols('num_orders ~ group * period', data=data_long[data_long['categoria'] == 'Champion']).fit()
print(model_champion.summary())
```

OLS Regression Results

Dep. Variable:	num_orders	R-squared:	0.010
Model:	OLS	Adj. R-squared:	0.010
Method:	Least Squares	F-statistic:	103.2
Date:	Thu, 07 Nov 2024	Prob (F-statistic):	1.73e-66
Time:	02:21:34	Log-Likelihood:	-1.1212e+05
No. Observations:	30625	AIC:	2.242e+05
Df Residuals:	30621	BIC:	2.243e+05
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	8.3482	0.190	43.932	0.000	7.976	8.721
group[T.treatment]	1.3478	0.207	6.497	0.000	0.941	1.754
period	0.2610	0.269	-0.971	0.331	-0.788	0.266
group[T.treatment]:period	1.4716	0.293	5.018	0.000	0.897	2.046

Omnibus:	18009.489	Durbin-Watson:	0.451
Prob(Omnibus):	0.000	Jarque-Bera (JB):	219815.114
Skew:	2.632	Prob(JB):	0.00
Kurtosis:	15.023	Cond. No.	12.6

Para los champion ej, es significativo y el efecto de la compensación se estima en 1.47 órdenes adicionales vs el que no recibe compensación



Amigos, ¿Sabían que...?

El modelo de diferencias en diferencias fue usado por los ganadores del premio nobel de economía de 2024, en su famoso paper donde demuestran que la estabilidad de las instituciones de una sociedad afectan en el largo plazo la prosperidad del pueblo...

Comparan países con estabilidad en un periodo de tiempo con otros que no...



Bueno, regresemos... resumen de resultados

	categoria	users_control	users_tratamiento	Incremental de Ordenes	Estd. Sig.	GPO Average	Monto compensado	ROI
0	Champion	2195	8720	1.47161	sí	1.19458	1.82709	96.22%
1	Loyal Customers	1283	4447	0.0886073	no	1.29675	2.06557	5.56%
2	About to Sleep	164	281	0.229698	no	0.916541	2.27264	9.26%
3	At Risk	90	94	0.286915	no	1.44374	4.11111	10.08%
4	Can't Lose Them	14	22	1.40846	no	2.65575	4.48302	83.44%
5	Hibernating	79	44	0.462599	no	0.803392	4.63594	8.02%
6	Potential Loyalists	338	877	0.311825	no	0.832176	1.93375	13.42%
7	New Customer	27	69	0.126396	no	0.314094	1.98336	2.00%

ROI de corto plazo= (Incremento en ordenes (segun modelo) * GPO Avg) / Monto Compensado Avg



Respondiendo hipótesis

→ ¿Una compensación o refund ayuda a mantener o aumentar la frecuencia de consumo de un usuario?

R// Sí ayuda. Aunque no logramos encontrar significancia estadística en todos los segmentos (sospechamos por tamaños de muestra), en los usuarios que son más importantes es notable y significativo el efecto

→ ¿Ese delta en frecuencia es rentable o no?

R// el ROI es positivo para todos los segmentos!

→ ¿Es el ROI diferente para los tipos de usuario?

R// Sí, notablemente diferente para los usuarios que más usan la app

→ ¿Qué recomendaciones podemos dar al esquema de compensación?

Siguiente lámina...



Conclusiones

- La estrategia de compensar más a los usuarios que valen más es la correcta, lo dice el ROI
- Recomendamos pensar en una mejor segmentación de usuario de la actual basada en reglas duras (criterios fijados en el tiempo) a una más dinámica, ej RFM
- Con RFM es interesante como ciertos segmentos de clientes leales no parecen mostrar un incremento significativo en la tasa de re-orden tras recibir compensaciones. Recomendamos realizar un A/B testing en esos segmentos para entender si podemos ser más eficientes en los montos de compensación.
- Con esos posibles ahorros, podemos invertir más en clientes “champion”, maximizando así el retorno de nuestra inversión.

