

Análisis de Operadores Ineficaces

Descripción

Identificar operadores ineficaces. El servicio de telefonía virtual CallMeMaybe está desarrollando una nueva función que brindará a los supervisores y las supervisoras información sobre los operadores menos eficaces. Se considera que un operador es ineficaz si tiene una gran cantidad de llamadas entrantes perdidas (internas y externas) y un tiempo de espera prolongado para las llamadas entrantes. Además, si se supone que un operador debe realizar llamadas salientes, un número reducido de ellas también será un signo de ineficacia.

Descripción de los Datos

El dataset comprimido telecom_dataset_us.csv contiene las siguientes columnas:

- user_id: ID de la cuenta de cliente
- date: fecha en la que se recuperaron las estadísticas
- direction: "dirección" de llamada (out para saliente, in para entrante)
- internal: si la llamada fue interna (entre los operadores de un cliente o clienta)
- is_missed_call: si fue una llamada perdida
- calls_count: número de llamadas
- call_duration: duración de la llamada (sin incluir el tiempo de espera)
- total_call_duration: duración de la llamada (incluido el tiempo de espera)

El conjunto de datos telecom_clients_us.csv tiene las siguientes columnas:

- user_id: ID de usuario/a
- tariff_plan: tarifa actual de la clientela
- date_start: fecha de registro de la clientela

Análisis Exploratorio de Datos

- ✓ El total de llamadas entrantes (OUT) es de 31917 y de llamadas salientes (IN) es de 21985, el total de llamadas perdidas es de 23568, el total de llamadas realizadas es de 30639, el total de llamadas en espera es de 52547.
- ✓ El promedio de la duración de llamadas es de 866.68, la mediana de la duración de llamadas es de 38, el promedio de la duración total de llamadas es de 1157.13, la mediana de la duración total de llamadas es de 210, el promedio de la llamada en espera es de 290.44, la mediana de llamadas en espera es de 55.
- ✓ La fecha mínima es el 02-08-2019 y la fecha máxima es el 28-11-2019.
- ✓ Existen tres planes tarifarios A, B, C.
- ✓ El promedio total de número de llamadas por día es de 16.45, el promedio diario de número de llamadas es el valor mínimo de 1 y el valor máximo de 32.58 entre el 02-08-2019 y el 28-11-2019.
- ✓ Las llamadas entrantes (IN) es el 40.8% y las salientes (OUT) es el 50.2% del total de llamadas.
- ✓ Los principales diez clientes tienen un porcentaje mayor de llamadas salientes (OUT) en comparación a llamadas entrantes (IN). Cinco clientes tienen por encima del 95% las llamadas salientes (OUT), dos clientes por encima del 70% las llamadas salientes (OUT), dos clientes por encima del 55% las llamadas salientes (OUT) y un cliente tiene el 47% las llamadas salientes (OUT).

- ✓ Los principales diez clientes tienen en promedio 4.55 operadores y siete clientes están por encima del promedio. Los principales diez clientes tienen ocho clientes con más número de llamadas salientes (OUT) que llamadas entrantes (IN), un cliente con igual número de llamadas entrantes (IN) y salientes (OUT) y un cliente con más número de llamadas entrantes (IN) que llamadas salientes (OUT). Los principales diez clientes tienen como el operador más usado a operadores no identificados.
- ✓ El porcentaje de operadores ineficientes es del 80.5% y de 19.5% de operadores eficientes para la duración total de llamadas, operadores ineficientes es del 80.4% y de 19.6% de operadores eficientes para la duración de llamadas, operadores ineficientes es de 83.3% y de 16.7% de operadores eficientes para las llamadas en espera.
- ✓ Las llamadas entrantes (IN) tienen 89.3% de operadores ineficientes y 10.7% de operadores eficientes para la duración de llamadas, el 91.1% de operadores ineficientes y 8.9% de operadores eficientes para la duración total de llamadas y el 97.4% de operadores ineficientes y 2.6% de operadores eficientes para las llamadas en espera. Las llamadas salientes (OUT) tienen 77.2% de operadores ineficientes y 22.8% de operadores eficientes para la duración de llamadas, el 77.5% de operadores ineficientes y 22.5% de operadores eficientes para la duración total de llamadas y del 78.6% de operadores ineficientes y 21.4% de operadores eficientes para las llamadas en espera.
- ✓ La correlación de las llamadas en espera con la duración de la llamada tiene un 0.5 y la correlación de las llamadas en espera con la duración total de las llamadas es de 0.68. (Se considera que 1 es una fuerte correlación y 0 es una nula correlación)
- ✓ La tasa de retención de clientes aumenta en el primer mes, pero a partir del segundo empieza a disminuir.
- ✓ Los histogramas de distribución para la duración de llamadas, la duración total de llamadas, llamadas en espera y número de llamadas por día presentan un sesgo hacia la izquierda.

- ✓ Las llamadas perdidas se concentran en una duración promedio de llamadas por operador por debajo de 1000.
- ✓ La participación de mercado para las llamadas entrantes (IN) por plan tarifario es plan A el 45.7%, plan B el 26.8% y plan C el 27.5% y para las llamadas salientes (OUT) por plan tarifario es plan A el 57.6%, plan B el 18.1% y plan C el 24.3%.

Prueba de Hipótesis

Se realizan las siguientes pruebas para dos distribuciones independientes:

PRUEBA 1 Diferencia entre llamadas entrantes y salientes en duración

PRUEBA 2 Correlación entre tiempo de espera y duración

PRUEBA 3 Plan tarifario vs dirección de llamada

PRUEBA 4 Dirección vs eficiencia

PRUEBA 5 Eficiencia por plan tarifario

PRUEBA 6 Tendencia en cantidad diaria de llamadas

PRUEBA 7 Duración baja vs llamadas perdidas

Se obtienen los siguientes resultados:

Prueba 1 - Duración IN vs OUT: p-value = 0.0810

Prueba 2 - Espera vs Duración: $p = 0.25$, $p = 0.0000$ Prueba 2b - Espera vs Duración Total: $p = 0.78$, $p = 0.0000$

Prueba 3 - Plan Tarifario vs Dirección: p-value = 0.0000

Prueba 4 - Dirección vs Eficiencia: p-value = 0.0000

Prueba 5 - Plan Tarifario vs Eficiencia: p-value = 0.0000

Prueba 6 - Tendencia diaria: $R^2 = 0.01$, p-value = 0.2815

Prueba 7 - Duración baja vs Perdidas: p-value = 0.0000

Conclusiones

1. Distribución y comportamiento de las llamadas

El volumen de llamadas salientes (OUT) supera al de llamadas entrantes (IN), lo cual sugiere una estrategia de comunicación activa hacia los clientes.

El 44.5% del total de llamadas son perdidas , un porcentaje de alarma que puede afectar la percepción de servicio.

La media de duración de llamadas es alta (866.68 segundos), pero la mediana (38 segundos) muestra una distribución altamente sesgada , indicando que muchas llamadas son cortas y unas pocas muy largas distorsionan el promedio.

2. Eficiencia operativa

La eficiencia general de los operadores es muy baja : solo entre el 16,7% y 22,8% de los casos se consideran eficientes, dependiendo de la métrica.

Las llamadas entrantes muestran peores indicadores de eficiencia (hasta un 97.4% de ineficiencia en llamadas en espera), en comparación con las salientes.

Esto sugiere problemas en la atención reactiva (recibir llamadas), posiblemente por falta de recursos, entrenamiento o sistemas automatizados.

3. Clientes clave y concentración

Los principales 10 clientes concentran una alta proporción de llamadas salientes y presentan un perfil proactivo. La mayoría de estos clientes operan con múltiples operadores (promedio: 4.55), y algunos tienen una fuerte dependencia de operadores no identificados, lo cual limita el control de calidad y análisis operativo.

4. Planes tarifarios

El plan A domina en términos de participación, tanto en llamadas IN como OUT.

Las pruebas de hipótesis revelan una dependencia significativa entre el plan tarifario y el tipo de llamada ($p < 0.0001$).

También hay diferencias en eficiencia entre planes tarifarios, lo que podría indicar que algunos planes están mal adaptados a los perfiles de los clientes.

5. Correlaciones e insights estadísticos

Hay una fuerte estimación entre el tiempo de espera y la duración total de la llamada ($\rho = 0.78$), lo cual puede indicar problemas en la resolución rápida de casos.

La prueba de hipótesis muestra que las llamadas con duración < 1000 están significativamente asociadas a llamadas perdidas ($p < 0.0001$).

No se observa una tendencia significativa en el promedio de llamadas diarias ($R^2 = 0.01$, $p = 0.28$), lo cual sugiere un comportamiento estable en el volumen de uso diario.

6. Retención de clientes

La tasa de retención mejora al principio , pero luego cae a partir del segundo mes , lo que sugiere posibles fallos en el seguimiento o satisfacción post-venta.

Recomendaciones

1. Operaciones y eficiencia

Reentrenar o reforzar el equipo que atiende llamadas entrantes , donde los indicadores de eficiencia son notablemente bajos.

Implementar sistemas de gestión de colas para reducir tiempos de espera y evitar pérdidas de llamadas.

Usar inteligencia artificial (bots o IVR) para filtrar o resolver solicitudes simples , mejorando tiempos de respuesta y eficiencia global.

2. Clientes estratégicos

Revisar y optimizar la gestión de los 10 principales clientes: asegurar buena cobertura de operadores, asignar ejecutivos dedicados, y formalizar operadores identificados para tener trazabilidad de atención.

Evaluar si los clientes con un alto porcentaje de llamadas salientes están generando retorno (ventas, retención, satisfacción).

3. Modelos tarifarios

Realizar un estudio de ajuste de planos tarifarios según el comportamiento y eficiencia del cliente.

Evaluar si los clientes del plan A, aunque dominantes, son también los que generan más llamadas perdidas o ineficientes.

4. Duración y calidad

Explorar las razones detrás de llamadas extremadamente cortas (mediana de 38 seg): ¿Cuelgan?

¿Problemas técnicos? ¿Falta de preparación?

Analizar la relación entre llamadas de baja duración y llamadas perdidas para identificar puntos críticos de abandono.

5. Seguimiento al cliente

Cree un programa de seguimiento al cliente que arranque en el segundo mes , momento en el que la retención comienza a caer.

6. Monitoreo continuo

Implementar cuadros de mando automáticos con indicadores de:

Eficiencia por operador y dirección de llamada

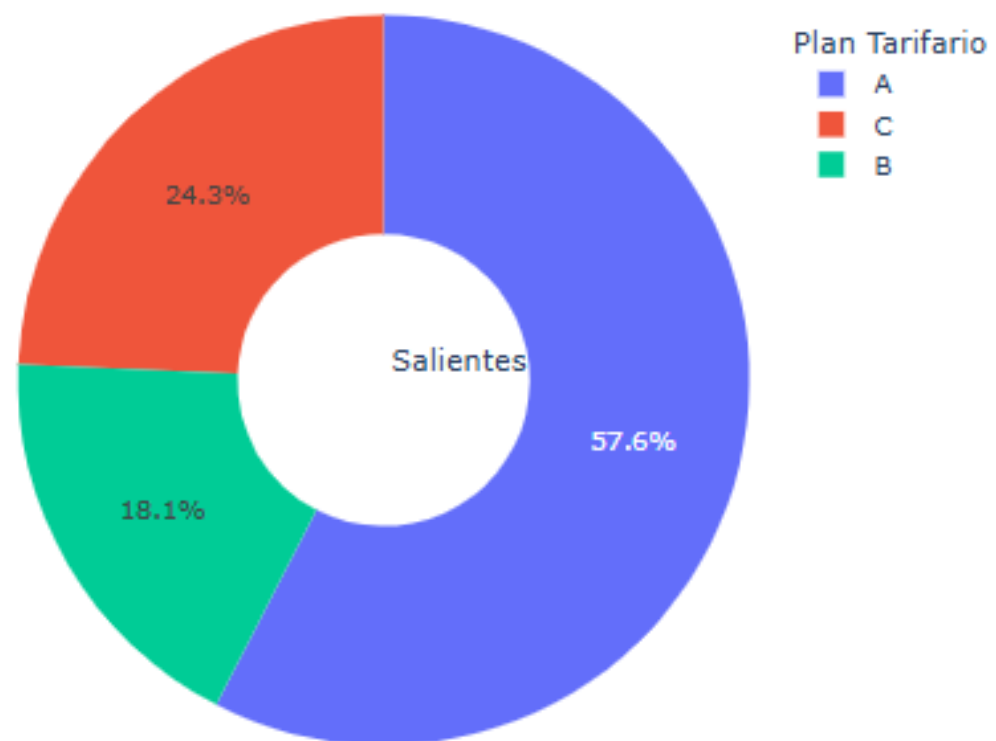
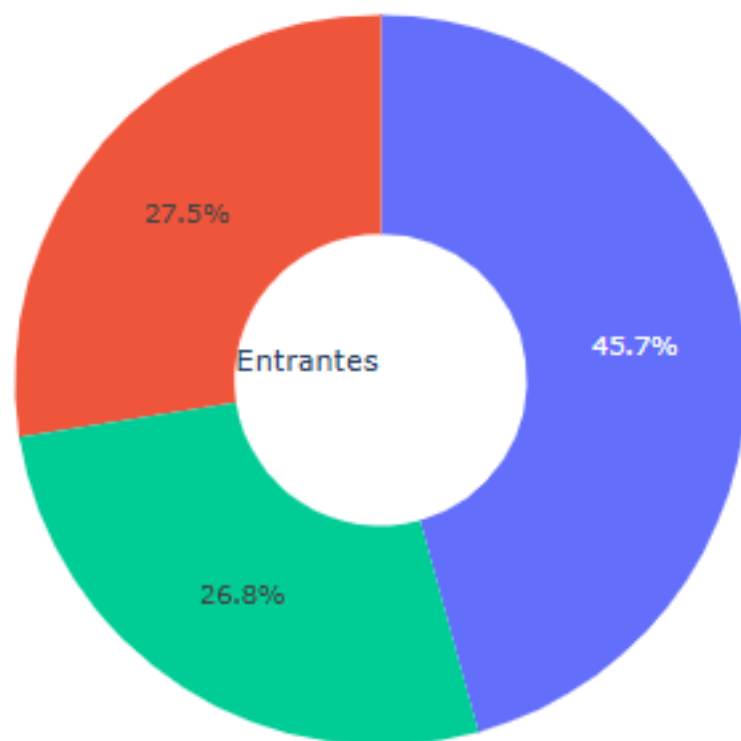
Tasa de llamadas perdidas

Evolución de llamadas diarias

Correlaciones y anomalías

Anexos

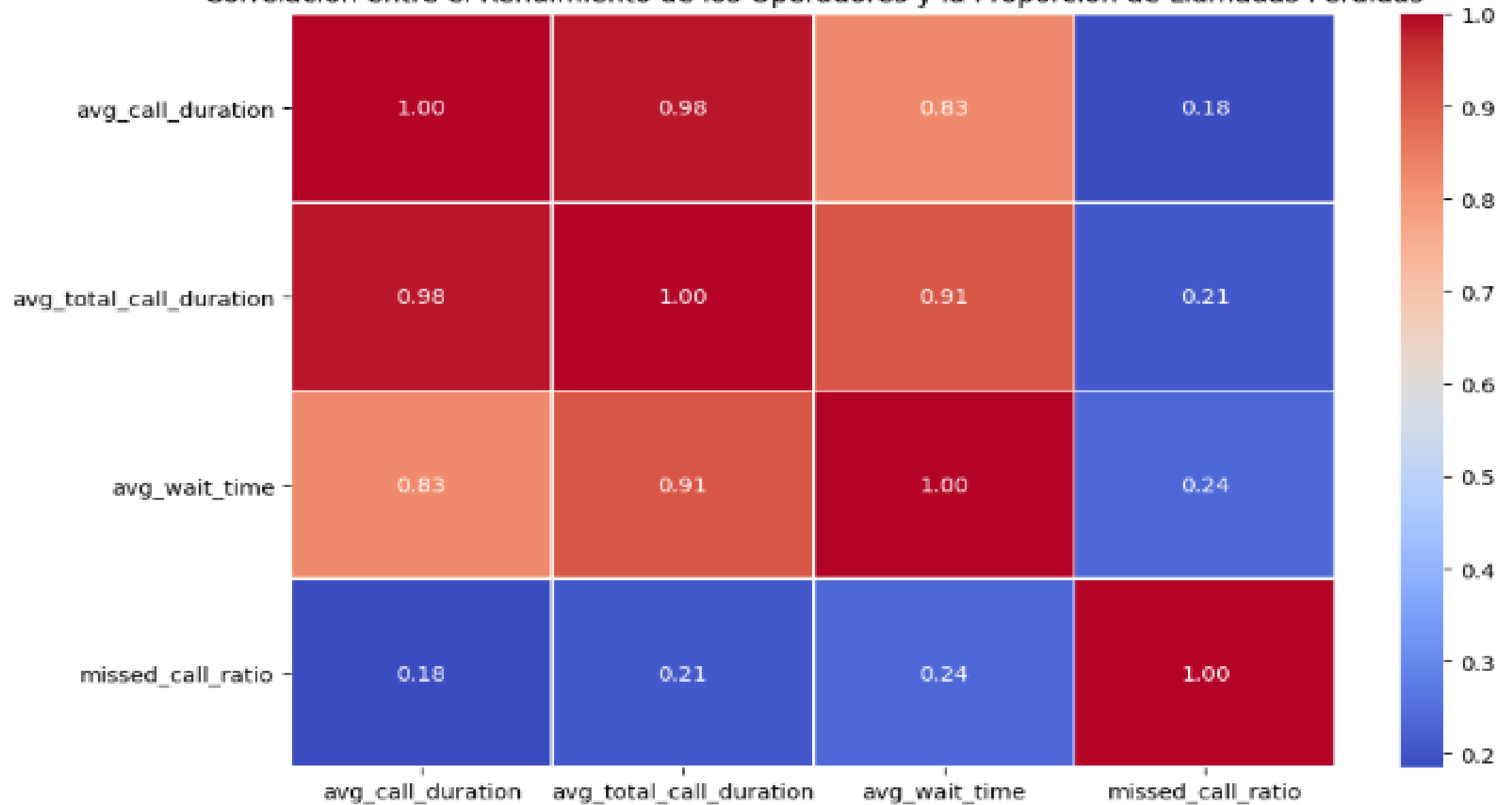
Distribución de llamadas por tipo y plan tarifario



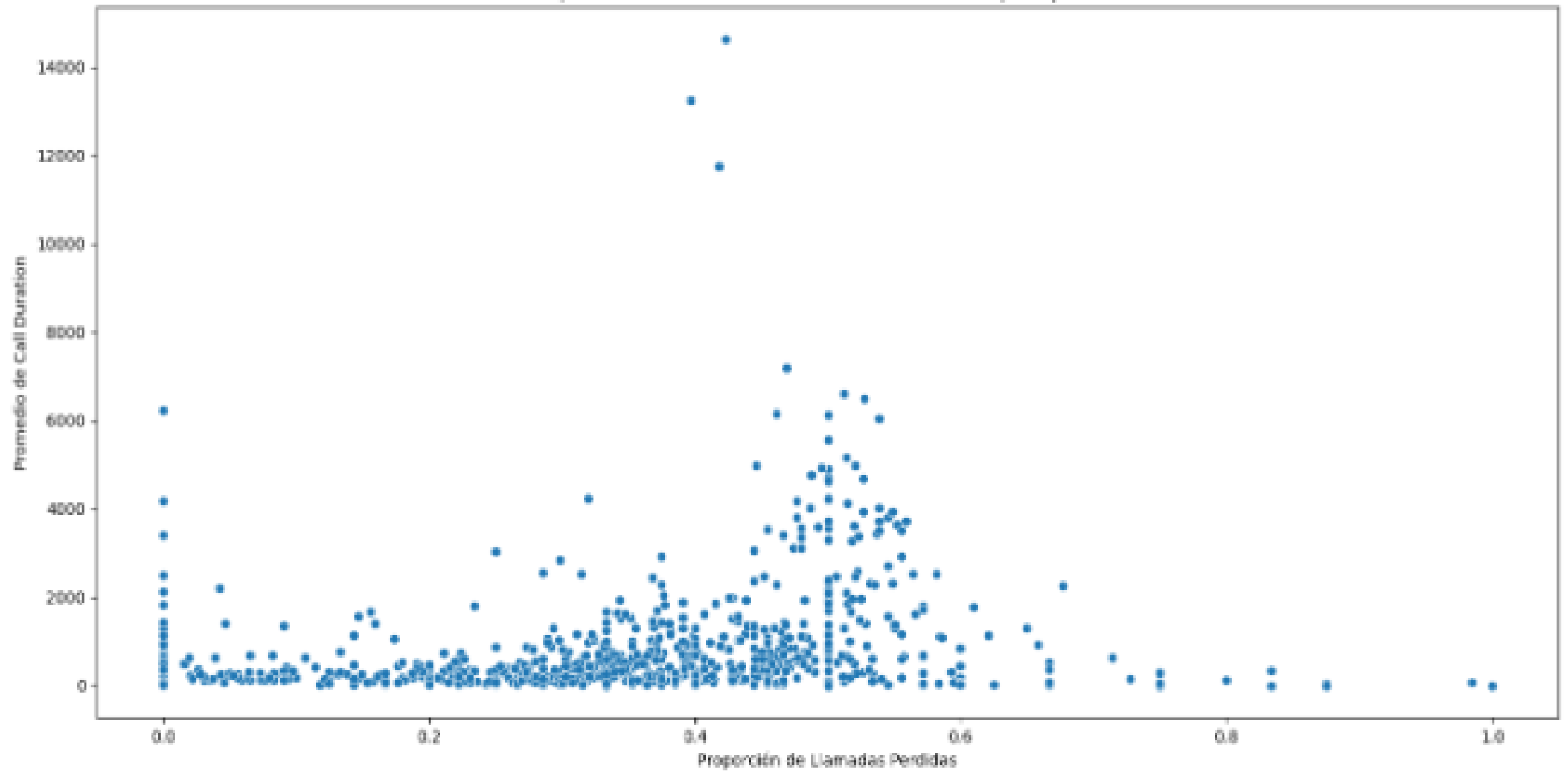
Rendimiento Mensual - Top 10 Operadores



Correlación entre el Rendimiento de los Operadores y la Proporción de Llamadas Perdidas

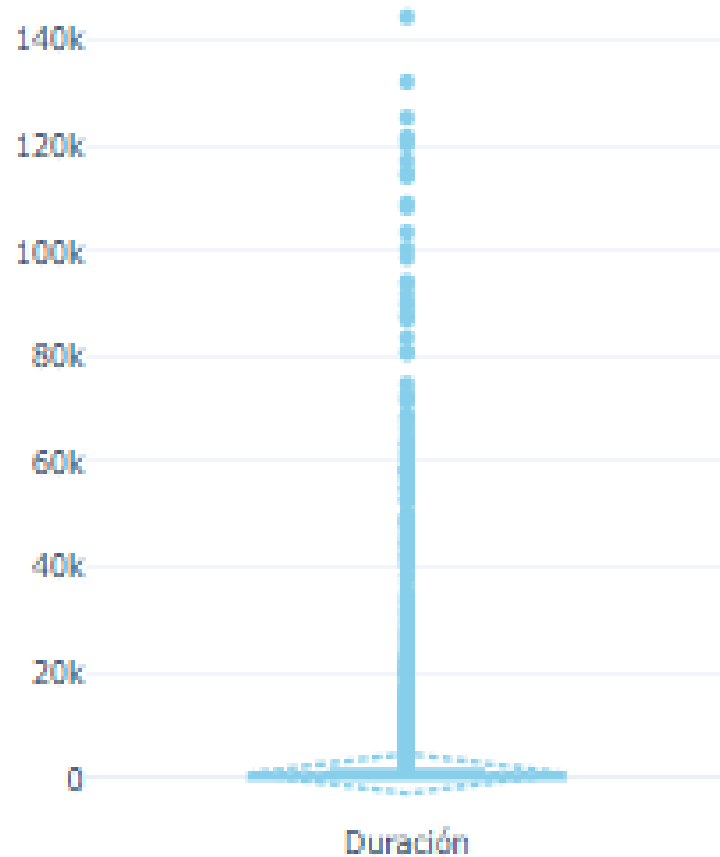


Proporción de Llamadas Perdidas vs Call Duration por Operador

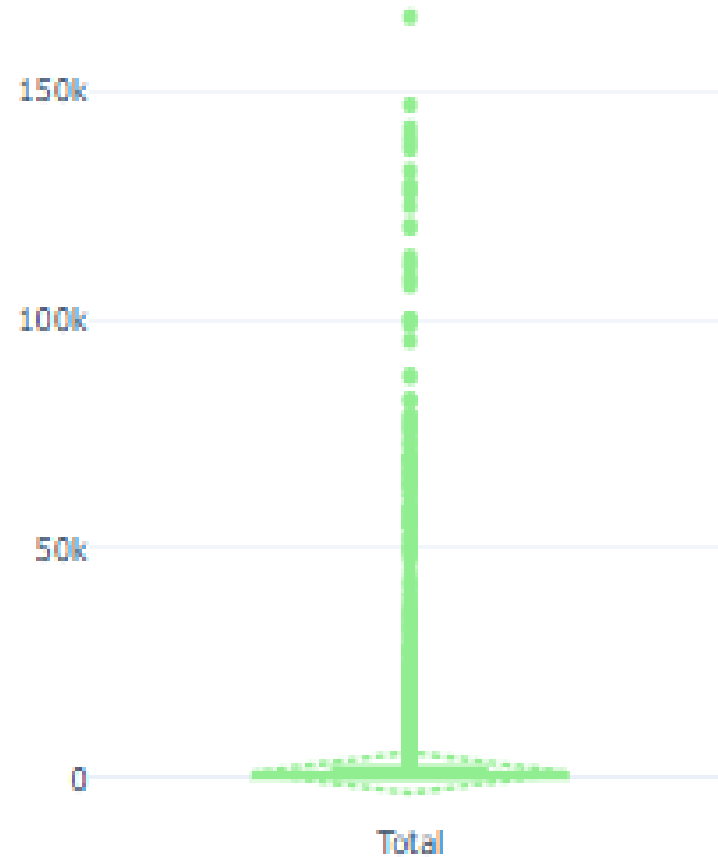


Diagramas de Caja de Duración y Tiempo de Espera en la Llamada

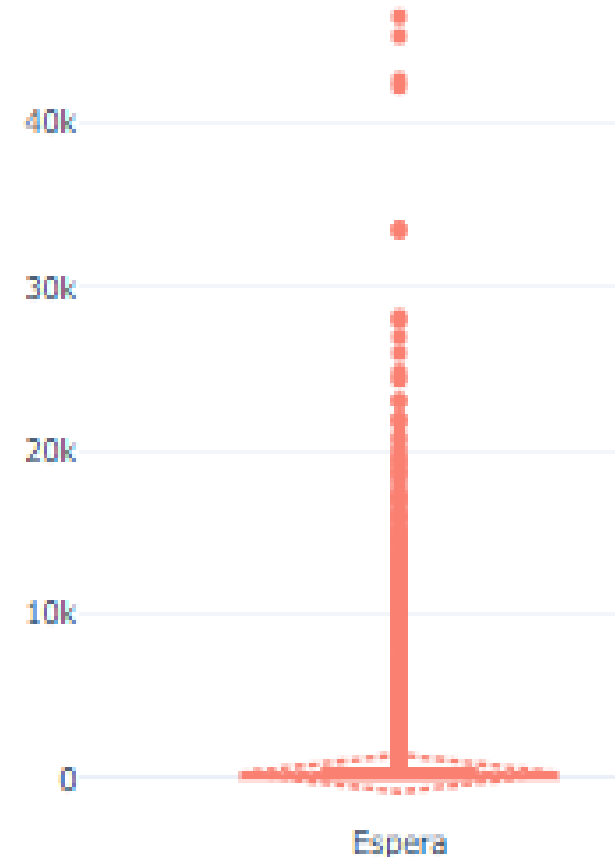
Duración de la Llamada

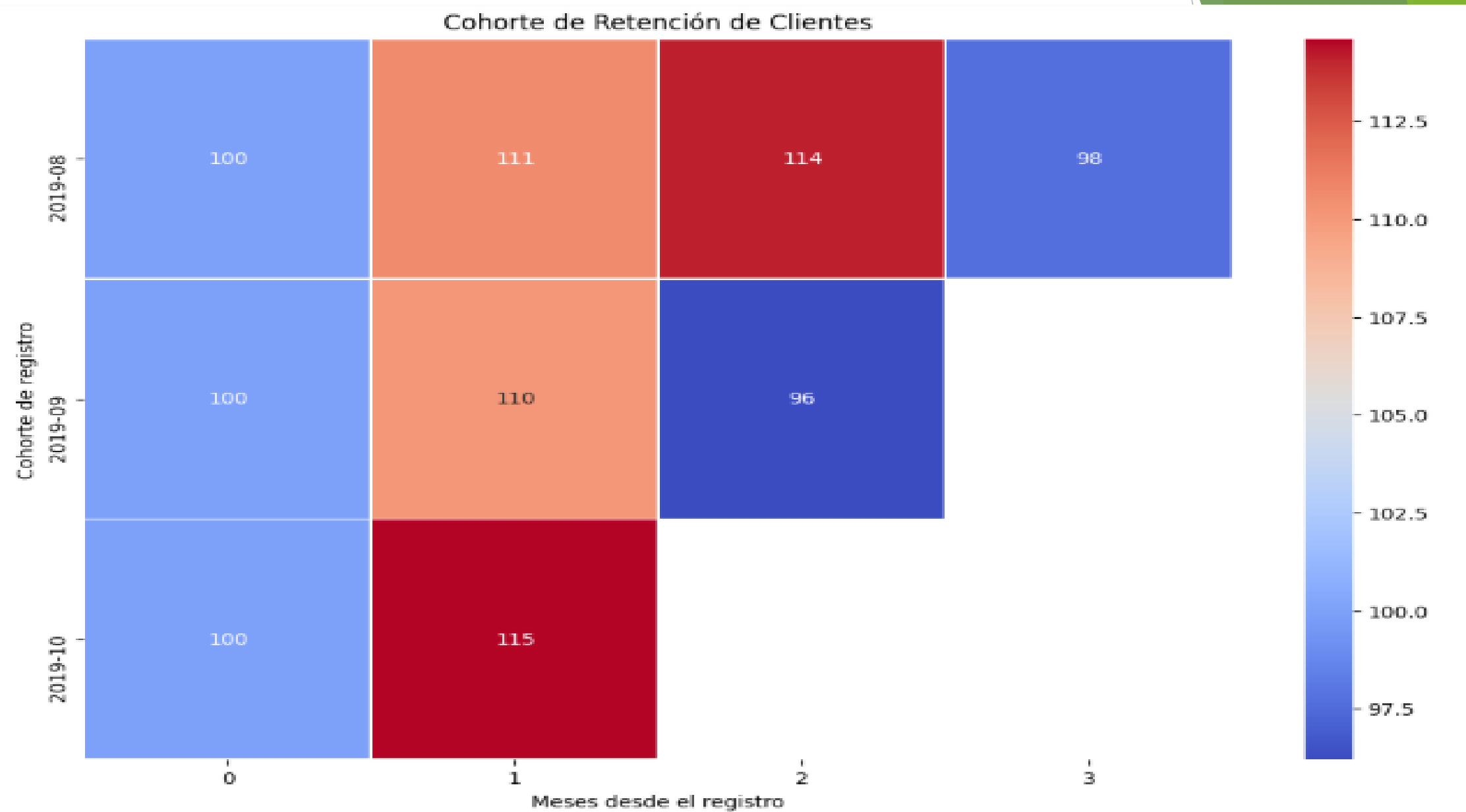


Duración Total de la Llamada

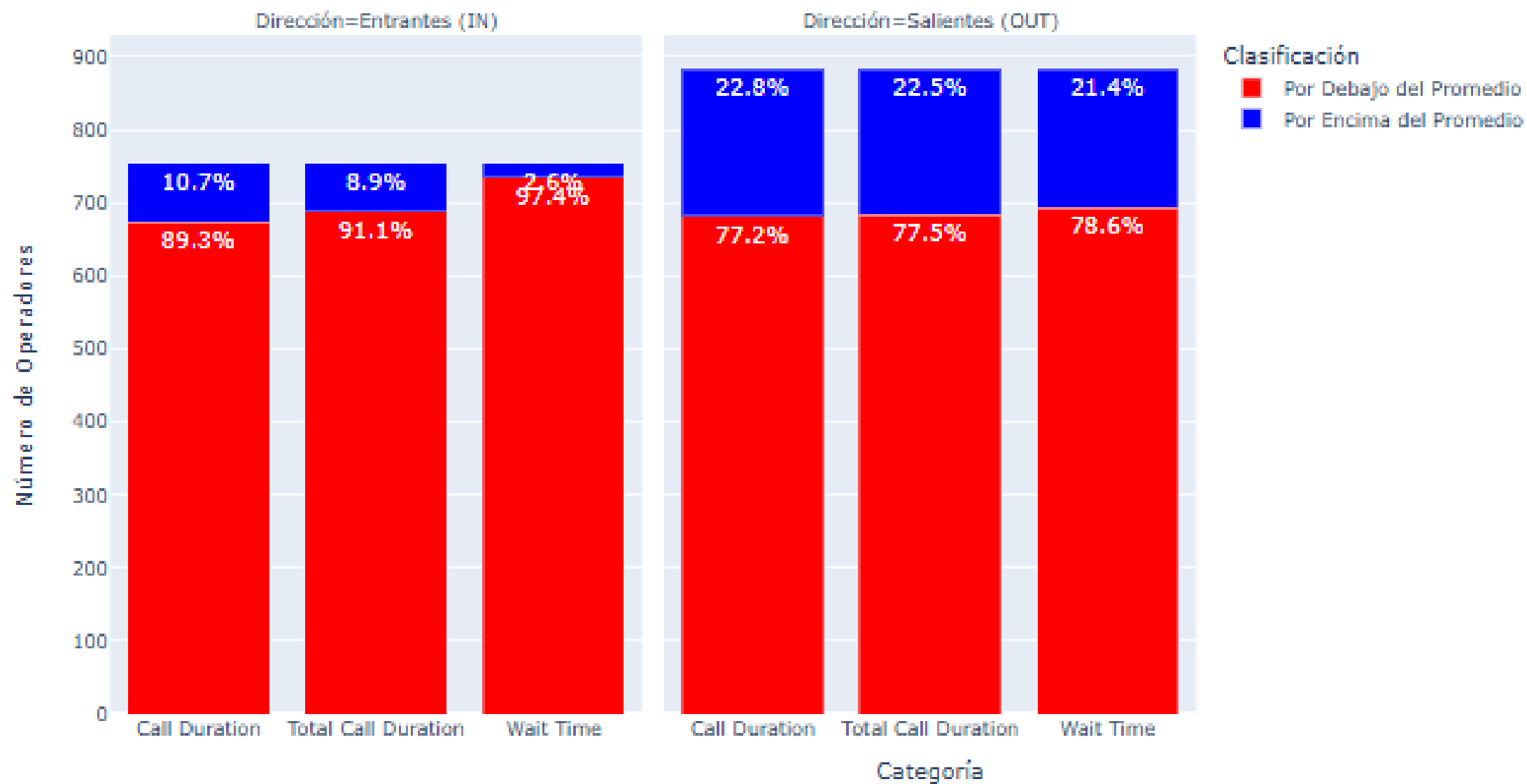


Tiempo de Espera

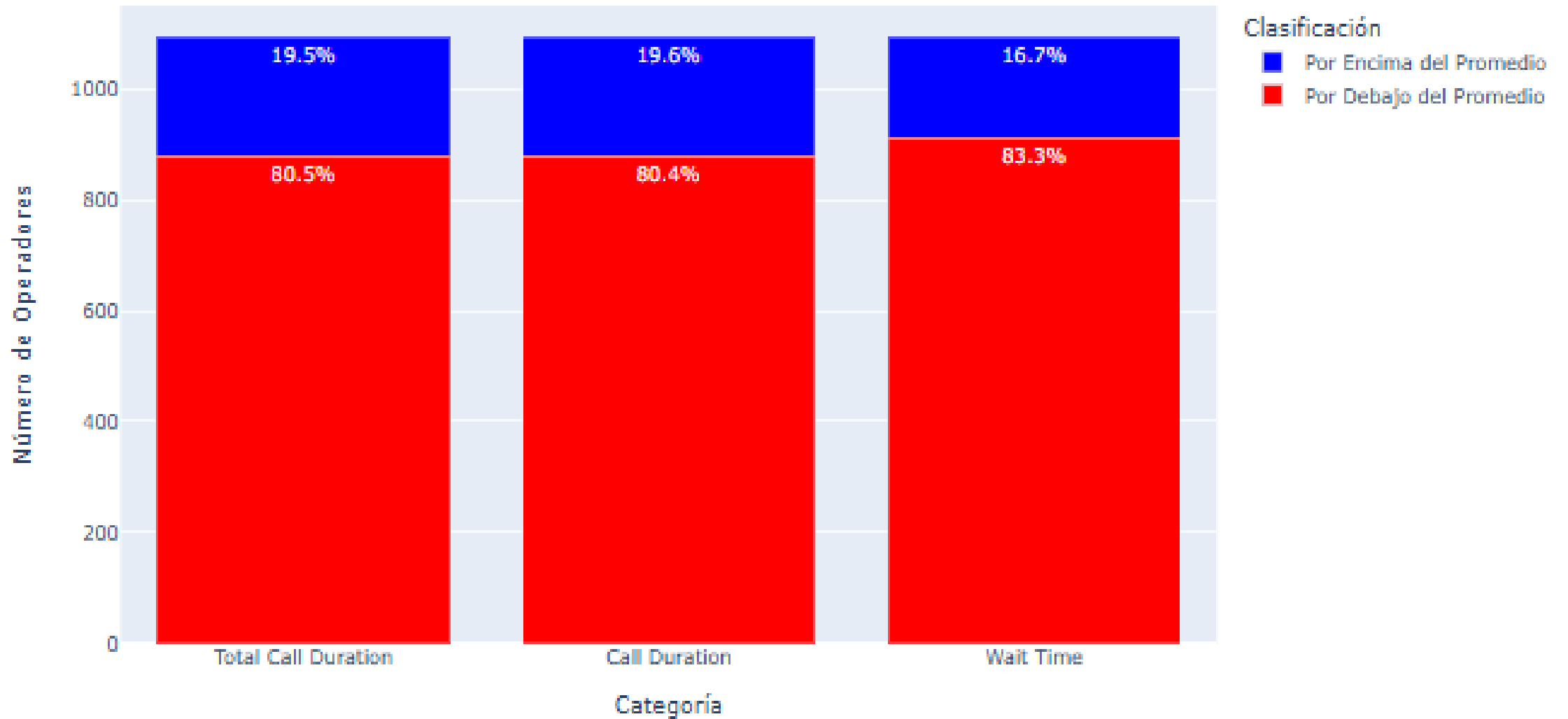




Comparación de Operadores por Métrica y Dirección



Operadores por Encima y por Debajo del Promedio



Prueba A/B

Estructura del embudo:

event_name		PRODUCT_CART	PRODUCT_PAGE	PURCHASE
ab_test	group			
RECOMMENDER_SYSTEM_TEST	A	2351	5043	2430
	B	643	1297	649

Tasas de conversión en cada etapa del embudo:

event_name		PRODUCT_CART	PRODUCT_PAGE	PURCHASE
ab_test	group			
RECOMMENDER_SYSTEM_TEST	A	0.239312	0.513335	0.247353
	B	0.248358	0.500966	0.250676

Resultado de la prueba Z entre los grupos A y B para 'purchase':

Z-score: 0.17366144026265296

P-value: 0.8621315431940824

No se ha encontrado una diferencia estadísticamente significativa entre los grupos A y B.

Conclusiones

1. Tasas de Conversión en el Embudo A/B :

Se ha analizado el embudo de conversión para los grupos A y B de la prueba A/B del "RECOMMENDER_SYSTEM_TEST".

En la etapa de PRODUCT_PAGE , el grupo A tiene una tasa de conversión del 51,33% , mientras que el grupo B tiene una tasa del 50,10% . Aunque hay una ligera diferencia a favor del grupo A, esta diferencia no parece ser significativa (ver análisis de la prueba Z a continuación).

En la etapa de PRODUCT_CART , la tasa de conversión es similar, con el grupo A en 23.93% y el grupo B en 24.84% , lo que indica que los dos grupos tienen comportamientos muy similares en esta etapa.

En la etapa final de COMPRA , las tasas de conversión entre los dos grupos son también muy cercanas: 24,74% para el grupo A y 25,07% para el grupo B.

2. Prueba Z y Significancia Estadística :

Se ha realizado una prueba Z para evaluar la diferencia de proporciones entre los grupos A y B en la etapa de COMPRA .

El puntaje Z resultante es 0.17 , con un valor p de 0.86 . Esto indica que no hay una diferencia estadísticamente significativa entre los grupos A y B en cuanto a la conversión en la etapa de compra. Dado que el valor p es mucho mayor que el umbral de 0.05 , podemos concluir que no hay evidencia suficiente para afirmar que el grupo B es superior al grupo A o viceversa en la tasa de conversión.

3. Eventos de interés :

Los eventos PRODUCT_PAGE , PRODUCT_CART y PURCHASE son los que se han analizado a lo largo de la prueba.

A pesar de las diferencias en las tasas de conversión entre los dos grupos, las diferencias no son sustanciales y no indican que uno de los grupos haya tenido un desempeño claramente superior. La cantidad de eventos en general no muestra una variación significativa entre los grupos, lo que indica que la interacción con el sistema fue bastante similar entre ambos grupos.

Recomendaciones

1. Reevaluar la Hipótesis de la Prueba A/B :

Aunque se ha evaluado la diferencia de tasas de conversión entre los grupos A y B, el análisis no ha mostrado una diferencia estadísticamente significativa. Esto sugiere que el cambio implementado (en el sistema de recomendación) no tiene un impacto claro en las tasas de conversión.

Es recomendable reevaluar la hipótesis de la prueba y considerar si las modificaciones en el sistema de recomendación están realmente impactando el comportamiento de los usuarios de manera significativa. Tal vez se necesiten ajustes adicionales o pruebas más específicas.

2. Ampliar el Período de Prueba :

Una posible razón por la que no se observe una diferencia significativa podría ser que el período de prueba no fue lo suficientemente largo como para que los efectos de la modificación se reflejen completamente. Sería útil ampliar el período de la prueba para ver si con más datos surgen diferencias significativas.

Alternativamente, aumentar el tamaño de la muestra podría proporcionar una mayor confianza en los resultados.

3. Análisis de subgrupos :

Puede ser útil realizar un análisis segmentado basado en variables adicionales, como región o dispositivo , para ver si el efecto de la prueba A/B varía en función de estas características.

Por ejemplo, es posible que los usuarios de diferentes dispositivos (PC, Android, iPhone) o regiones (UE, N.AMERICA) respondan de manera diferente a las recomendaciones del sistema. Este análisis podría revelar si hay diferencias que no son evidentes cuando se agrupan todos los usuarios.

4. Optimización de los Eventos de Conversión :

A pesar de que no se encontró una diferencia significativa en la tasa de conversión, se debe revisar si los eventos clave en el embudo (PRODUCT_PAGE, PRODUCT_CART y PURCHASE) están optimizados.

Se podrían realizar ajustes para mejorar la tasa de conversión en cada etapa del embudo. Por ejemplo, mejorar la experiencia de usuario en la página de producto o en el carrito de compras podría ayudar a aumentar las conversiones, independientemente del grupo A/B.

5. Explorar otras métricas :

En lugar de centrarse únicamente en las tasas de conversión, se podrían explorar otras métricas como tiempo en el sitio , número de páginas vistas o valor medio del pedido para obtener una visión más completa del impacto del sistema de recomendación.

Incluso si no hay una diferencia significativa en la

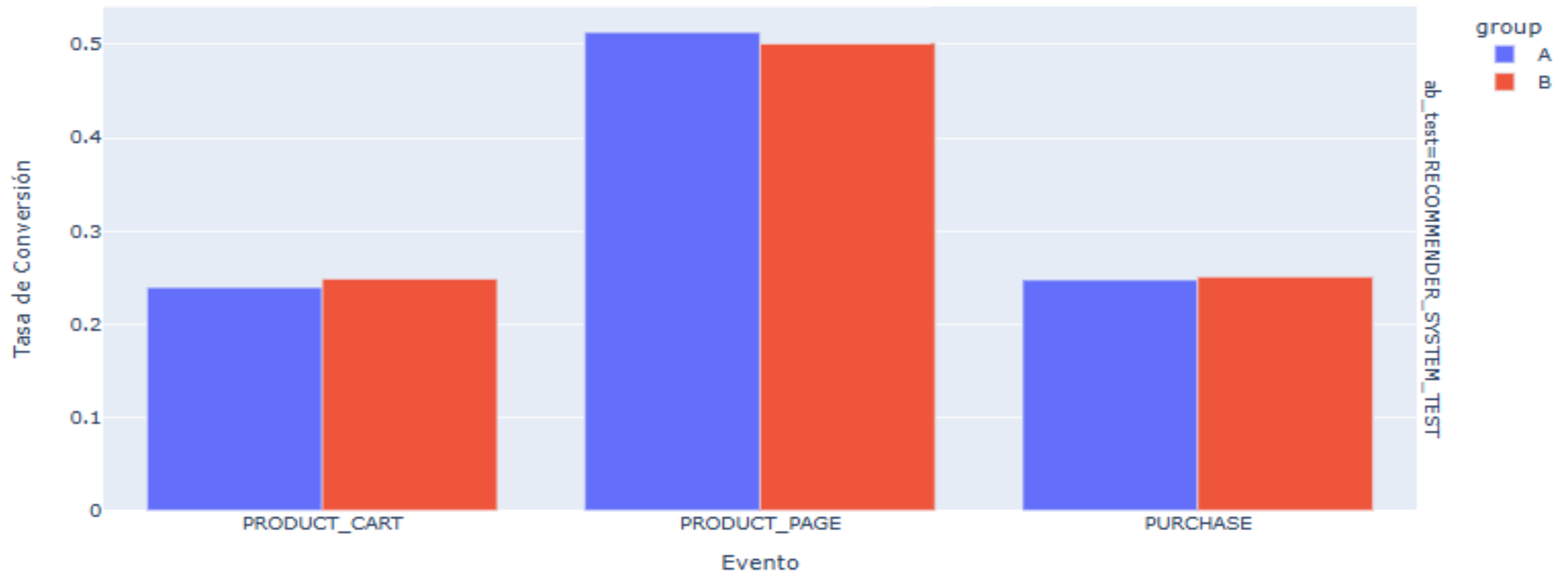
6. Realizar pruebas iterativas :

Dado que no se observará una diferencia significativa en esta prueba A/B, sería prudente realizar pruebas adicionales con ajustes en los parámetros del sistema de recomendación.

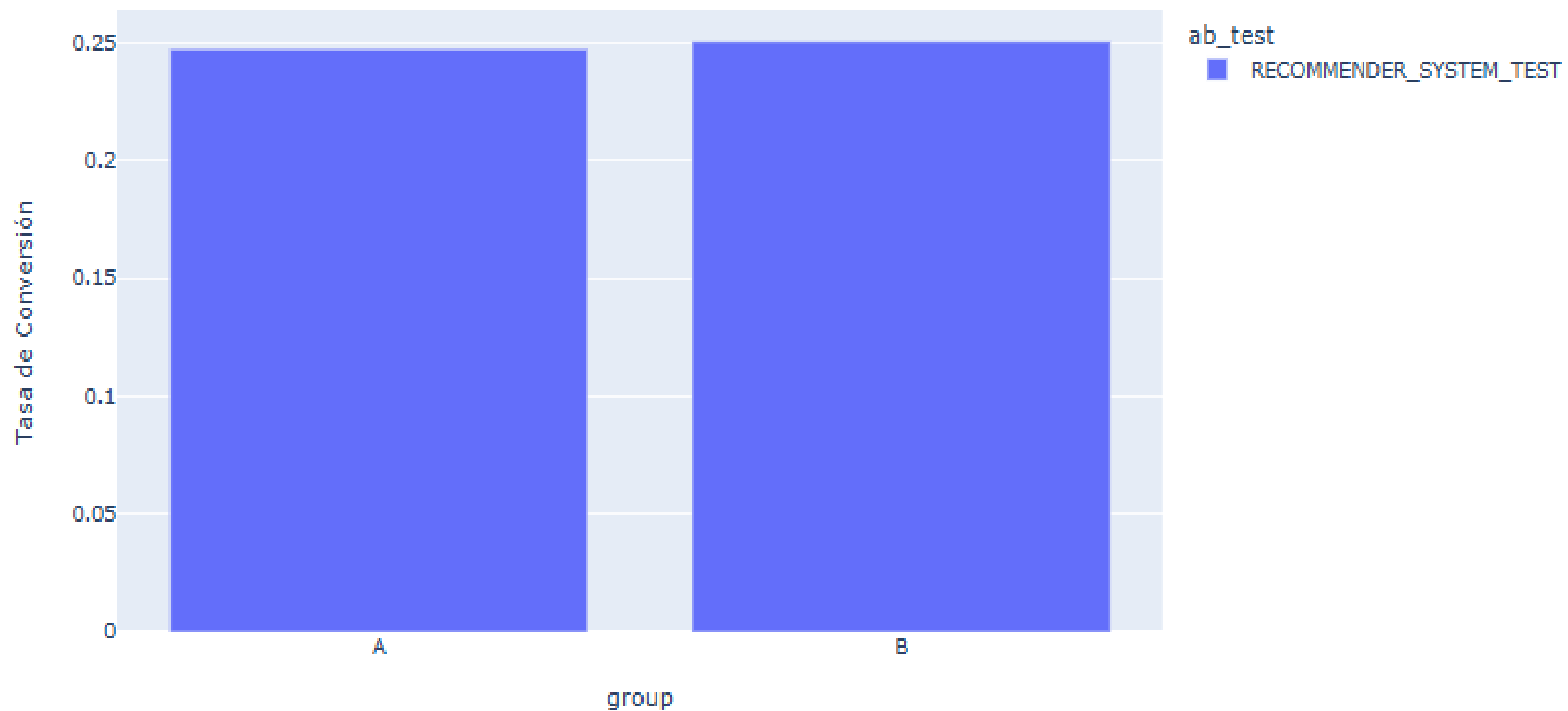
Esto podría incluir modificar los algoritmos, probar nuevas variables de entrada o personalizar más las recomendaciones, lo que podría llevar a resultados más favorables en futuras pruebas

Anexos

Tasa de Conversión por Evento para Grupos A y B



Tasa de Conversión en la Etapa de Compra ("PURCHASE") por Grupo



Distribución de Eventos por Grupo A/B

