

Proyecto Final:

Telecomunicaciones: identificar operadores ineficaces



Ejercicio

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

Los datasets contienen información sobre el uso del servicio de telefonía virtual CallMeMaybe. Sus clientes son organizaciones que necesitan distribuir gran cantidad de llamadas entrantes entre varios operadores, o realizar llamadas salientes a través de sus operadores. Los operadores también pueden realizar llamadas internas para comunicarse entre ellos. Estas llamadas se realizan a través de la red de CallMeMaybe.

Comenzamos con este dataset

	user_id	date	direction	internal	operator_id	is_missed_call	calls_count	call_duration	total_call_duration
0	166377	2019-08-04 00:00:00+03:00	in	False	NaN	True	2	0	4
1	166377	2019-08-05 00:00:00+03:00	out	True	880022.0	True	3	0	5
2	166377	2019-08-05 00:00:00+03:00	out	True	880020.0	True	1	0	1
3	166377	2019-08-05 00:00:00+03:00	out	True	880020.0	False	1	10	18
4	166377	2019-08-05 00:00:00+03:00	out	False	880022.0	True	3	0	25
...
53897	168606	2019-11-10 00:00:00+03:00	out	True	957922.0	True	1	0	38
53898	168606	2019-11-11 00:00:00+03:00	out	True	957922.0	False	2	479	501
53899	168606	2019-11-15 00:00:00+03:00	out	True	957922.0	False	4	3130	3190
53900	168606	2019-11-15 00:00:00+03:00	out	True	957922.0	False	4	3130	3190
53901	168606	2019-11-19 00:00:00+03:00	in	False	NaN	True	2	0	64

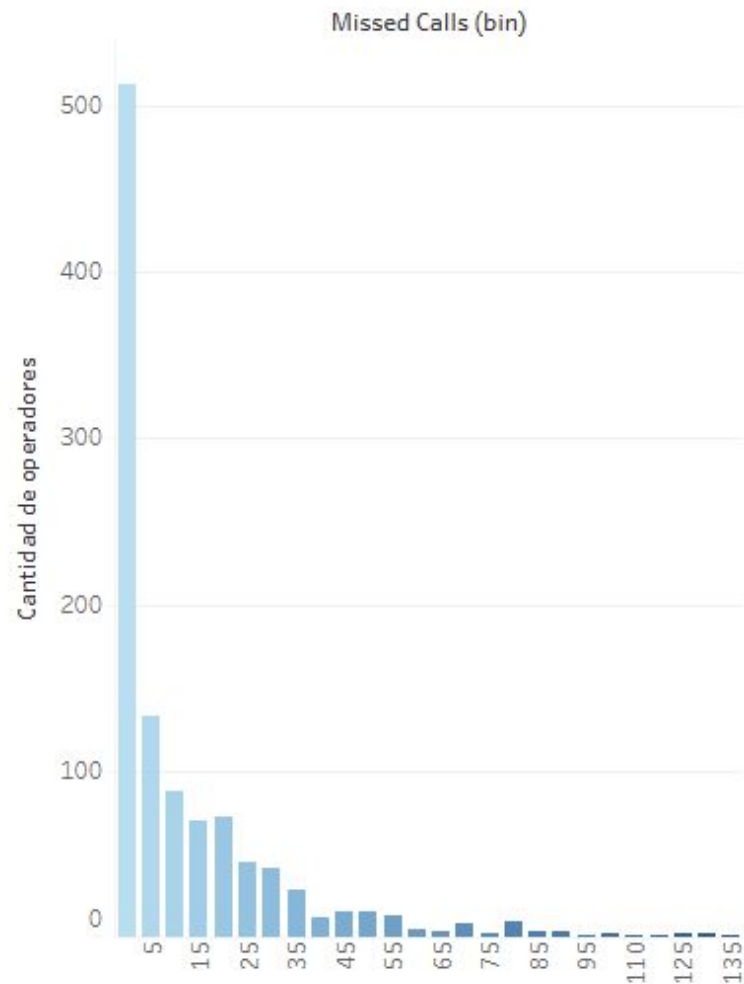
Tras el análisis exploratorio, agrupamos por operador y seleccionamos las métricas que utilizaremos para encontrar a los operadores ineficientes

	missed_calls	total_waiting	incoming_calls	outgoing_calls
operator_id				
879,896	50	14166	60	1071
879,898	100	112972	118	7856
880,020	7	334	8	46
880,022	33	4375	8	211
880,026	94	23992	25	2414
...
972,410	2	1498	0	77
972,412	2	1050	1	60
972,460	3	801	2	68
973,120	1	20	0	3
973,286	0	88	2	0

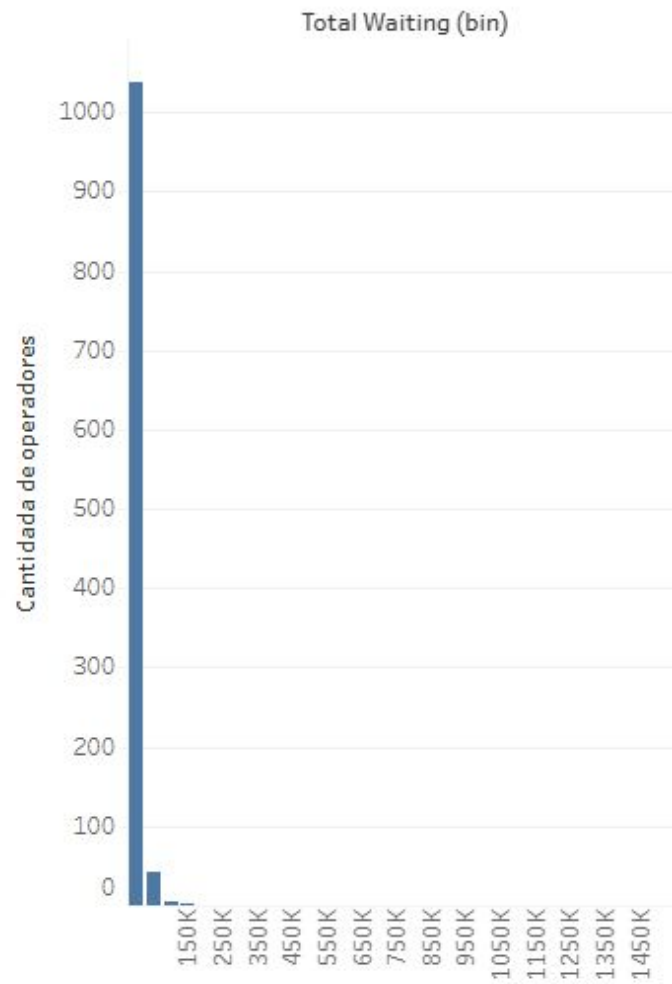
Graficamos las métricas para darnos una idea de sus distribuciones

Nuestra intención con este paso es ir entendiendo un poco mejor los datos con los que estamos trabajando e ir ideando cómo encontrar a los operadores ineficaces.

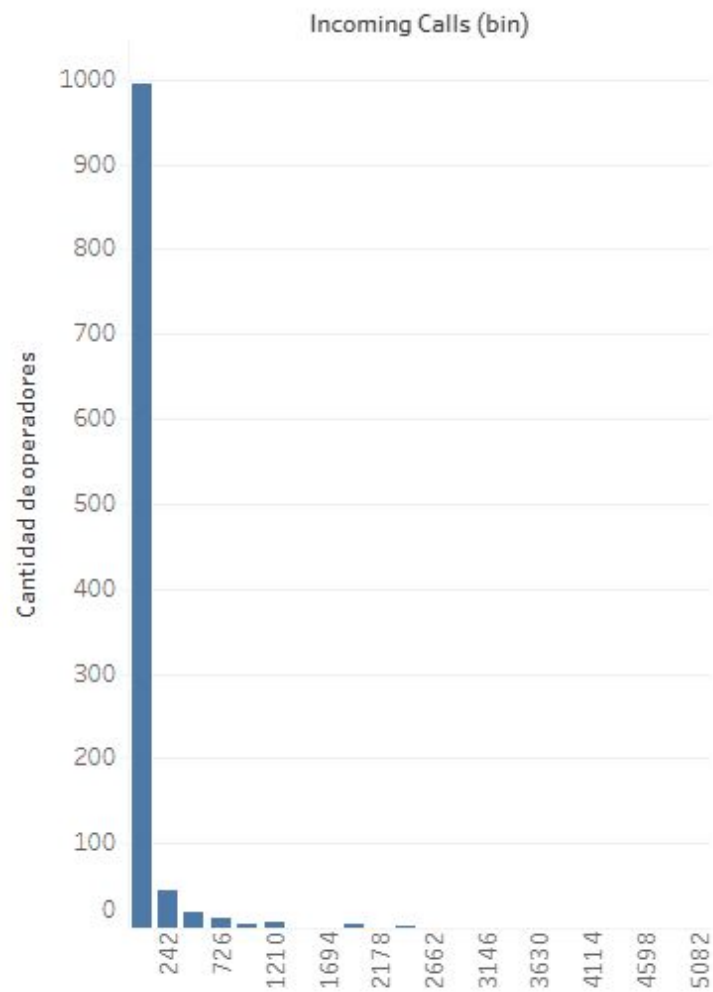
Llamadas perdidas



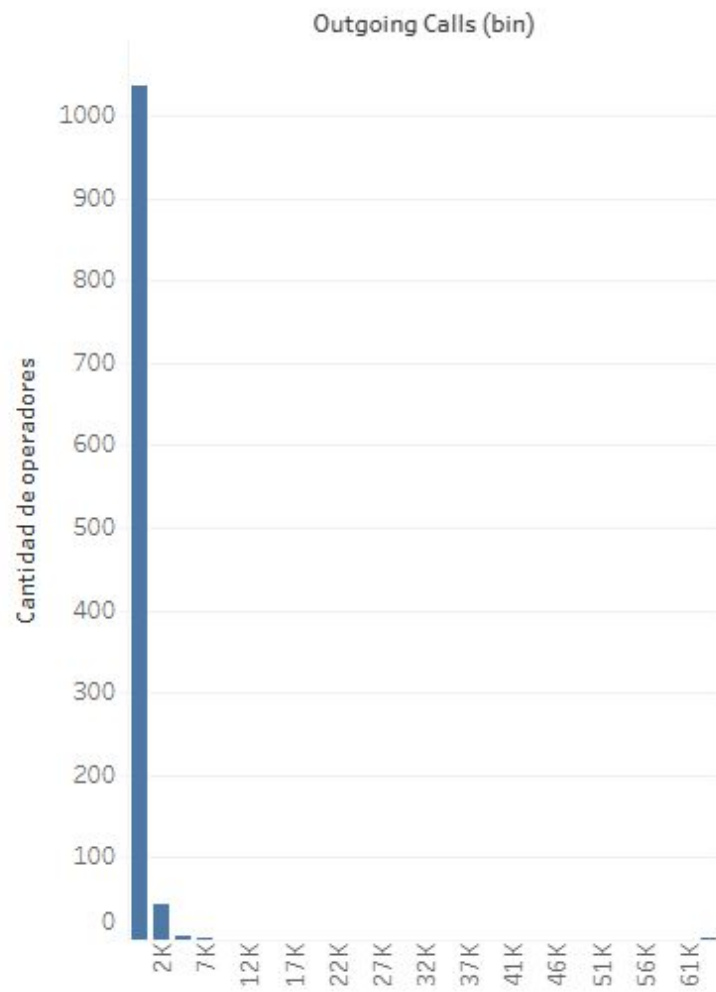
Tiempo de espera (segundos)



Llamadas entrantes



Llamadas salientes



Observamos varios percentiles para cada metrica, en base a esto establecere un sistema de puntajes en el cual cada columna tendra cierto valor:

missed_calls: 24 llamadas perdidas otorgan medio punto. A partir de 38 se otorga punto completo.

total_waiting: 15000 segundos equivalen a medio punto. A partir de 32000 se otorga punto completo.

incoming_calls: 0 llamadas otorgan medio punto.

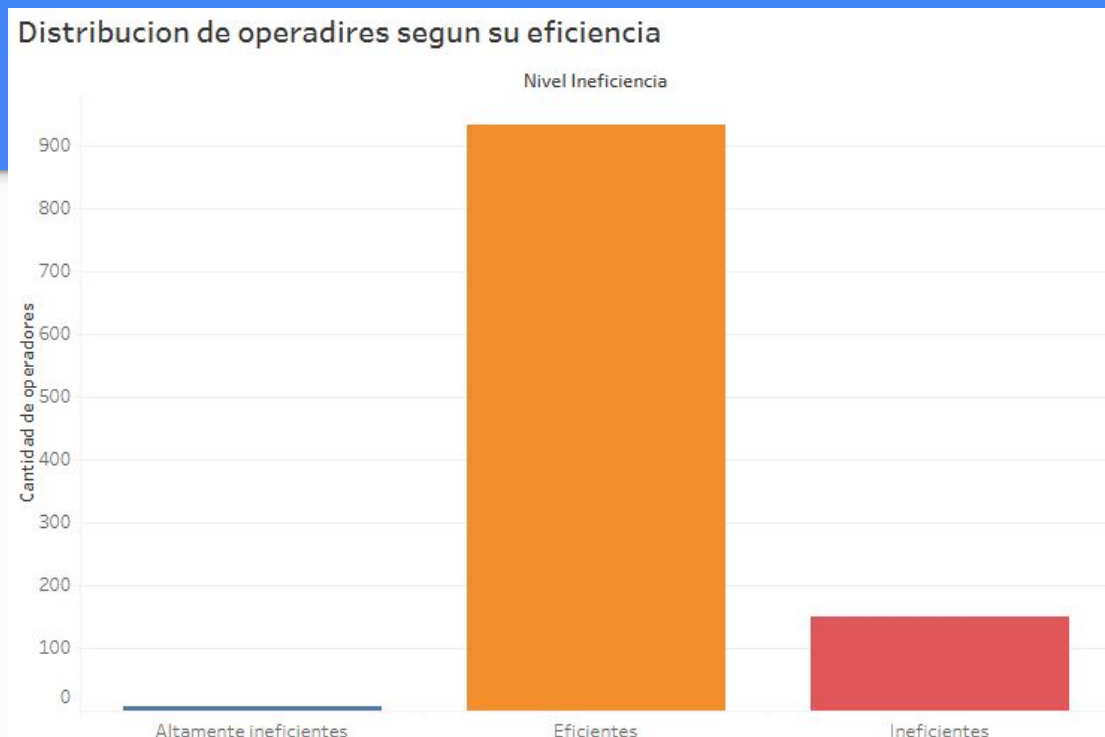
outgoing_calls: 0 llamadas otorgan medio punto.

	missed_calls	total_waiting	incoming_calls	outgoing_calls	puntaje_ineficiencia
0.01	0.0	3.00	0.00	0.00	0.0
0.05	0.0	15.00	0.00	0.00	0.0
0.10	0.0	36.00	0.00	0.00	0.0
0.15	0.0	62.00	0.00	0.00	0.0
0.20	1.0	104.20	0.00	1.00	0.0
0.30	1.0	242.30	0.00	4.00	0.0
0.40	3.0	501.00	2.00	13.00	0.5
0.50	6.0	1113.00	5.00	36.50	0.5
0.60	10.0	2432.60	11.00	105.60	0.5
0.75	21.0	8573.50	38.25	388.75	0.5
0.80	24.0	14141.20	64.00	663.00	1.0
0.90	37.9	32722.60	208.90	1547.20	1.5
0.95	55.0	49424.95	468.50	2466.20	2.0
0.99	94.0	106577.43	1464.49	5022.29	2.0

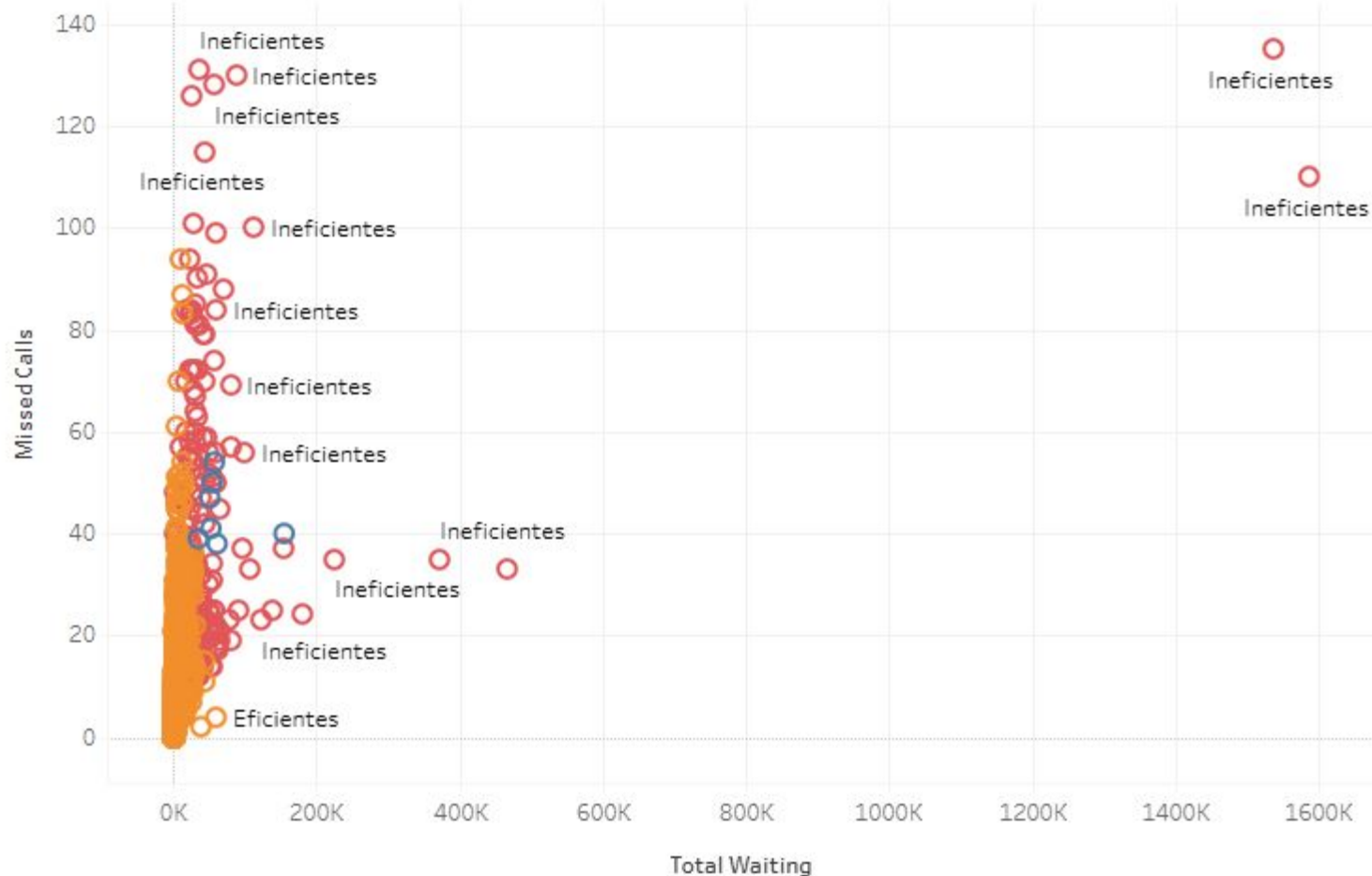
Al aplicar los parametros mencionados logramos encontrar el nivel de eficiencia para cada operador

	missed_calls	total_waiting	incoming_calls	outgoing_calls	puntaje_ineficiencia	nivel_ineficiencia
operator_id						
945302.0	40	154973	0	7059	2.5	Altamente ineficientes
893674.0	47	51203	0	2701	2.5	Altamente ineficientes
902782.0	39	35118	0	1473	2.5	Altamente ineficientes
902748.0	38	61217	0	2831	2.5	Altamente ineficientes
945296.0	41	52591	0	2821	2.5	Altamente ineficientes
...
908080.0	0	108	7	1	0.0	Eficientes
908082.0	2	443	31	20	0.0	Eficientes
939716.0	5	251	14	16	0.0	Eficientes
939708.0	2	694	10	36	0.0	Eficientes
906070.0	16	14614	361	89	0.0	Eficientes

Encontramos que tenemos 935 operadores eficaces, 150 ineficaces y 7 altamente ineficaces



Llamadas perdidas vs. Tiempo de espera



Hipótesis estadísticas

Hipotesis nula (H_0): No existe diferencia significativa entre los grupos.

Hipotesis alternativa (H_1): Existe una diferencia significativa entre los grupos.

Pruebas t para missed_calls:

Eficientes vs Ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Eficientes vs Altamente ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Ineficientes vs Altamente ineficientes: $p = 0.4849 \rightarrow$ No hay diferencia significativa

Pruebas t para total_waiting:

Eficientes vs Ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Eficientes vs Altamente ineficientes: $p = 0.0057 \rightarrow$ Hay diferencia significativa

Ineficientes vs Altamente ineficientes: $p = 0.8652 \rightarrow$ No hay diferencia significativa

Pruebas t para incoming_calls:

Eficientes vs Ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Eficientes vs Altamente ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Ineficientes vs Altamente ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Pruebas t para outgoing_calls:

Eficientes vs Ineficientes: $p = 0.0000 \rightarrow$ Hay diferencia significativa

Eficientes vs Altamente ineficientes: $p = 0.0032 \rightarrow$ Hay diferencia significativa

Ineficientes vs Altamente ineficientes: $p = 0.9524 \rightarrow$ No hay diferencia significativa

Comprobamos que los operadores eficientes se comportan de manera totalmente distinta que los operadores ineficientes y los altamente ineficientes. Entre estos dos grupos el comportamiento es similar.

Conclusión

El desempeño inferior por parte de los operadores ineficientes ha sido comprobado y demostrado en este proyecto; En prácticamente todas las métricas analizadas se pueden apreciar estas afirmaciones.

También me he tomado la libertad de dividir a los operadores en eficaces, ineficaces y altamente ineficaces, esto con el fin de tomar decisiones derivadas de su desempeño, por ejemplo, para empleados ineficaces se podría realizar una auditoría, una llamada de atención o simplemente un aviso de su desempeño, en cambio para los operadores altamente ineficaces se podrían tomar medidas un poco más drásticas, eso ya corresponde a un área especializada dentro de la empresa.