

Identificación de Operadores Ineficaces en CallMeMaybe

Juan Carlos Pedroza

Objetivo:

Identificar operadores con bajo rendimiento (10% peor en cada métrica) para ayudar a supervisores a tomar decisiones informadas.

Métricas clave:

- Llamadas perdidas (internas/externas).
- Tiempo de espera en llamadas entrantes.
- Cantidad de llamadas salientes.

Enfoque: Análisis por rubros separados para precisión y acciones personalizadas.



Proceso de Análisis

- 01 **Preprocesamiento:** Eliminación de duplicados y valores nulos (operator_id). Transformación de tipos de datos (ej. convertir de object a datetime).
- 02 **Agrupación por mes y operador:** Llamadas perdidas (count). Tiempo de espera (sum). Llamadas salientes (count).
- 03 **Percentiles:** P90 (llamadas perdidas/tiempo de espera). P10 (llamadas salientes).

Proceso de Análisis

Se agregaron más de 400 nuevos operadores en 3 meses



Desde el primer registro al último, el número de operadores únicos incrementó en más de 400. Es por ello que decidí realizar reportes mensuales y trabajé con el último mes.



Hallazgos Principales (Noviembre 2019)

Operadores ineficaces por rubro:

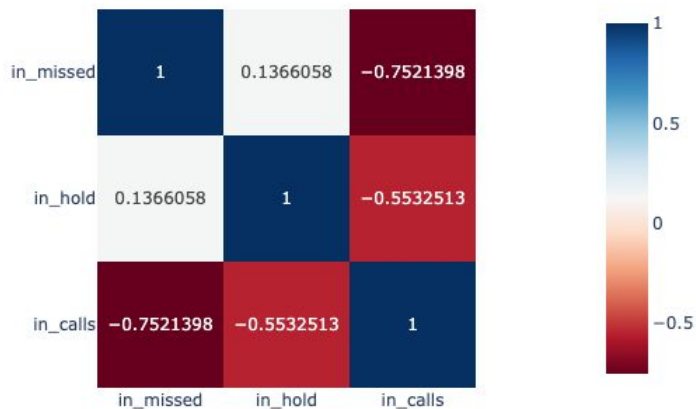
81 operadores con más llamadas perdidas (≥ 59).

66 operadores con menos llamadas salientes (≤ 1).

32 operadores con alto tiempo de espera (> 2492 segundos).

Correlaciones

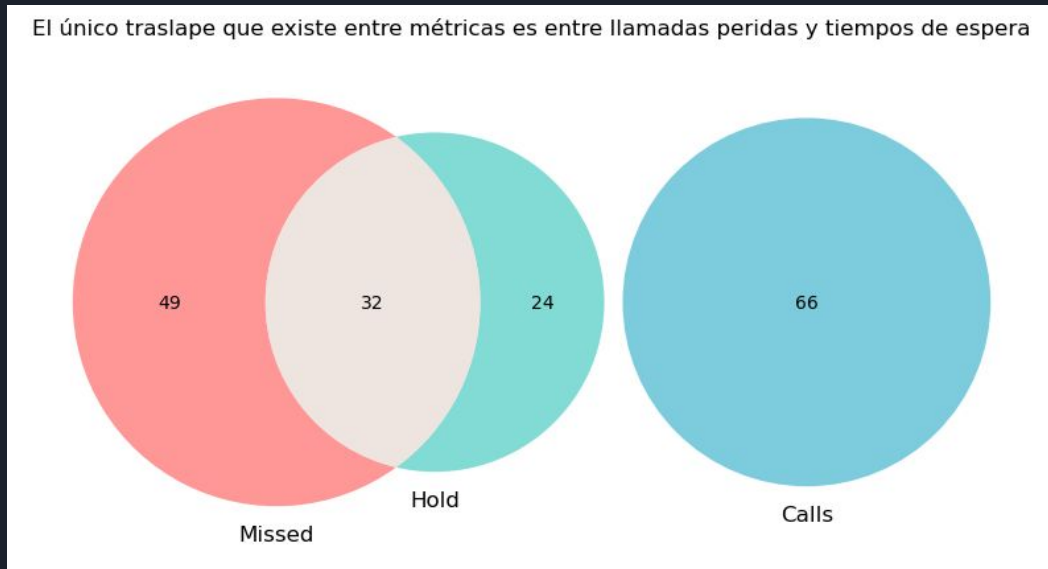
Parece haber una correlación negativa entre las llamadas salientes



Los operadores con pocas llamadas salientes no suelen fallar en otros rubros.

En general hay los operadores que tienen pocas llamadas salientes, suele ser porque tienen mejores resultados en sus otras dos métricas.

Correlaciones



En este periodo (noviembre 2019) solo 32 operadores fallaron en llamadas perdidas y tiempo de espera. Ninguno en los 3 rubros.



Pruebas estadísticas

```
# Prueba Mann-Whitney U
stat, p_value = stats.mannwhitneyu(group_peor, group_resto)

print("Tiempos de espera" Comparación de grupos.)
print(f"Estadístico U: {stat}")
print(f"p-value: {p_value}")

# Interpretación
alpha = 0.05
if p_value < alpha:
    print("✅ Rechazamos H₀: El grupo peor es significativamente peor (p < 0.05).")
else:
    print("❌ No hay evidencia para rechazar H₀ (p ≥ 0.05).")
```

✓ 0.0s

"Tiempos de espera" Comparación de grupos.

Estadístico U: 31136.0

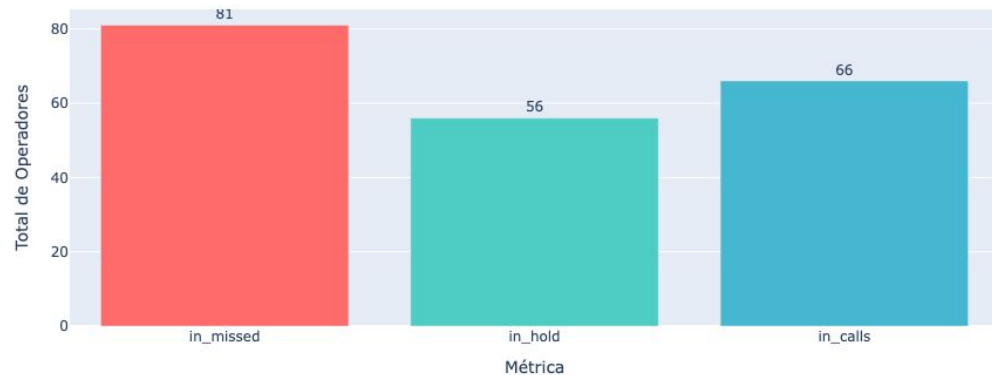
p-value: 4.2848490497286975e-156

✅ Rechazamos H₀: El grupo peor es significativamente peor (p < 0.05).

Se demostró con la prueba de Mann-Whitney que existen diferencias significativas ($p < 0.05$) entre los grupos ineficaces vs. resto.

Hallazgos

Total de operadores ineficientes por Rubro



- Llamadas perdidas es el rubro con más incidencias.
- Los operadores ineficaces son heterogéneos; pocos fallan en múltiples áreas.
- Llamadas perdidas es el rubro con más incidencias.



Tabla resumen

dataframe: operadores_ineficaces

Columnas: operator_id, in_missed, in_hold, in_calls (valores 1/0).

En esta tabla se encuentran los id's de los operadores que fueron ineficaces en el mes de noviembre, y en las columnas se indica con un 1 el rubro o rubros en los que fallaron.

	operator_id	in_missed	in_hold	in_calls
0	882686	1	0	0
1	887276	1	1	0
2	888406	0	0	1
3	888540	0	0	1
4	889410	1	0	0
...
166	969288	0	0	1
167	969380	0	0	1
168	970240	0	0	1
169	970250	0	0	1
170	970258	0	0	1

171 rows x 4 columns



Acciones sugeridas

- Capacitación específica según rubro (ej. gestión de llamadas entrantes).
- Revisar carga laboral para operadores con pocas llamadas salientes.

Detalles técnicos:

- **Pruebas estadísticas:** Mann-Whitney U (no paramétrica).
- **Librerías:** Pandas, SciPy, Math, Plotly.
- **Dataset final:** 41,546 registros limpios (agosto-noviembre 2019).



Fuentes

- Reyes, J. A., & Torres, F. J. (2019). Estadística aplicada a la investigación en salud (2ª ed.). Editorial Médica Panamericana.

Me ayudó a comprender qué es una prueba paramétrica y una no paramétrica y cuándo debería usar una u otra.

- Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2018). Metodología de la investigación (7ª ed.). McGraw-Hill.

Para saber qué es una hipótesis e hipótesis alternativa, y cómo formularlas.



Fuentes

- VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media.

Para escribir código correctamente y de manera legible.

- Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5ª ed.). Sage.

Para saber de qué manera redactar y organizar una investigación relacionada con datos.

- Moore, D. S., Notz, W. I., & Fligner, M. A. (2021). The basic practice of statistics (9ª ed.). W.H. Freeman.

Sobre los alcances y limitaciones de las pruebas T Student.