

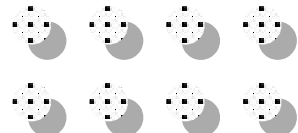


Proyeto final

01/04/25

Obetivo de proyecto:

- Llevar a cabo el análisis exploratorio de datos
- Identificar operadores ineficaces
- Prueba las hipótesis (a reserva que el area interesada tenga necesidades de informacion particulares) se sugieren las siguientes hipotesis y analisis
 - Hay una mayor cantidad de llamadas internas que externas.
 - La mayoría de las llamadas se conectaron (es decir no son llamadas perdidas)
 - Las llamadas internas tienen una duracion mayor con respecto a las externas
- Esto utilizando dos datasets sobre el uso del servicio de telefonía virtual CallMeMaybe.



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)
 - operator_id: identificador del operador
 - 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)

Datos:

- 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

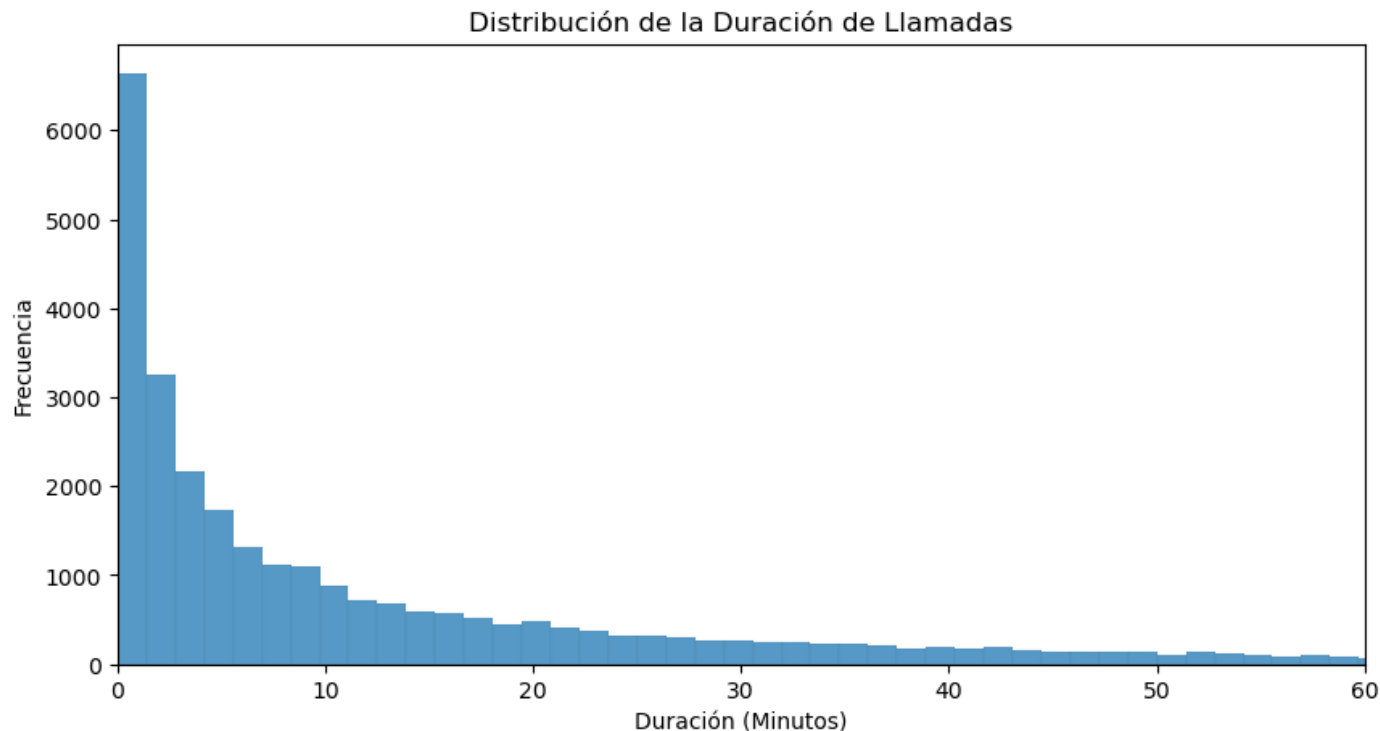
EDA

- No se encuentran particularidades especialmente relevantes en los datos.
- Se hacen correcciones de valores vacíos y formatos.

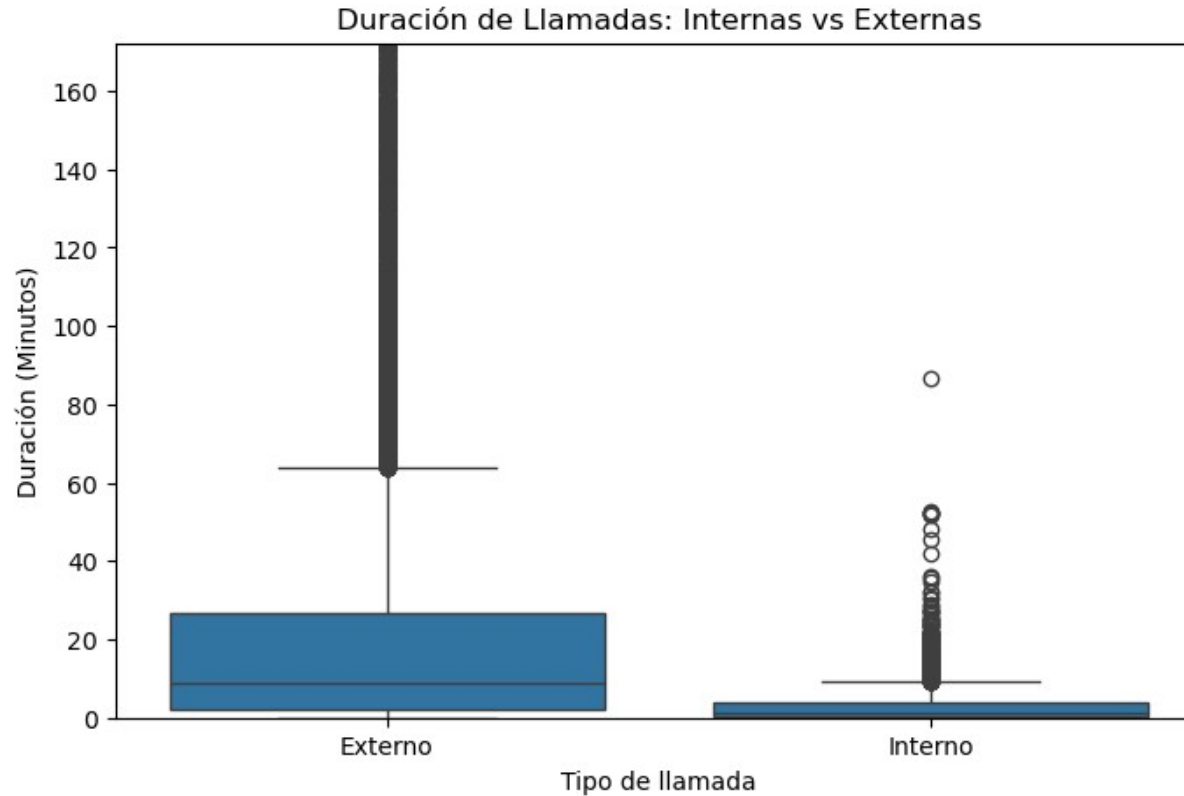
Resultados

//

//



Se observa que la mayoría de las llamadas tienen una duracion de aproximadamente menor a 10 minutos.



- Con base en la grafica y corroborado en la prueba estadística confirmamos la hipótesis que las llamadas externas duran mas que las internas

T-statistic: -46.96548101210464, P-value: 0.0
Existe una diferencia estadísticamente significativa.

# operator_id	# total_llamadas	# llamadas_perdidas	# porcentaje_perdidas
941826.0	1	1	100.0
937872.0	1	1	100.0
885682.0	3	3	100.0
914272.0	1	1	100.0
970258.0	1	1	100.0
924572.0	1	1	100.0
914626.0	1	1	100.0
937778.0	1	1	100.0
958458.0	1	1	100.0
937772.0	1	1	100.0

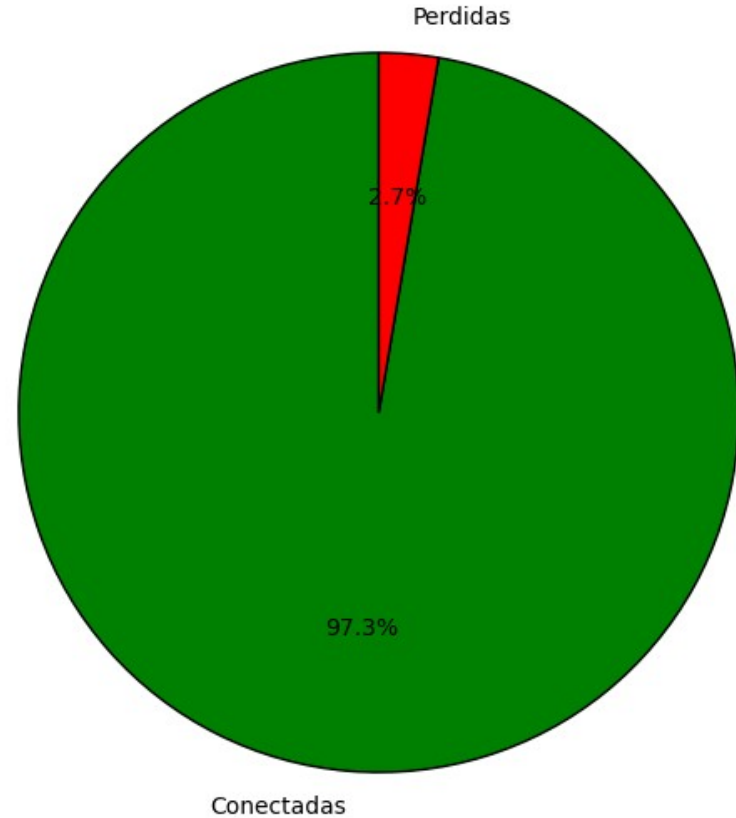
- Extracto de tabla de los operadores con mayor cantidad de llamadas perdidas

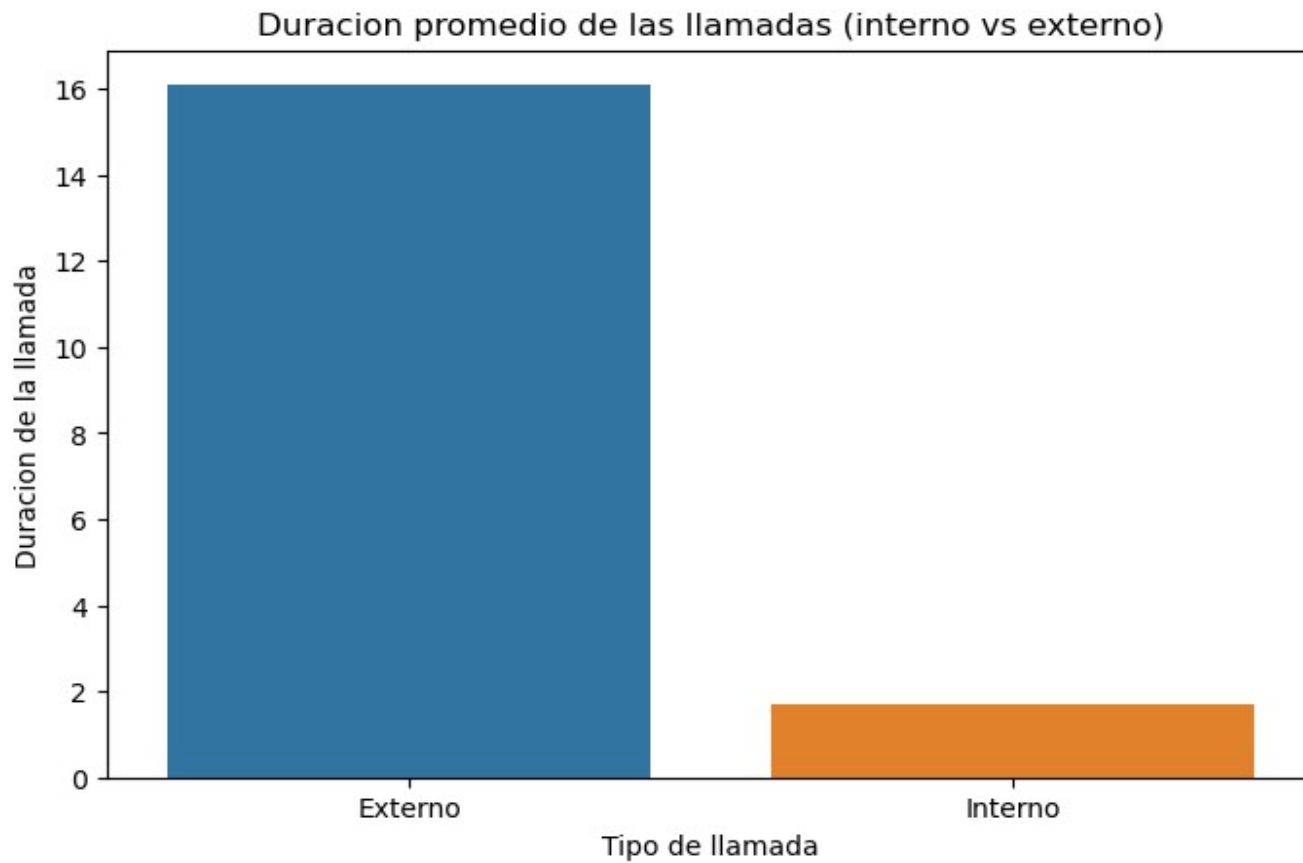
# operator_id	# total_llamadas	# llamadas_perdidas	# porcentaje_perdidas
941826.0	1	1	100.0
911104.0	1	0	0.0
946468.0	1	0	0.0
946032.0	1	0	0.0
899906.0	1	0	0.0
945710.0	1	0	0.0
945274.0	1	0	0.0
950632.0	1	0	0.0
947596.0	1	0	0.0
948282.0	1	0	0.0

- Extracto de tabla de los operadores con menor actividad

- Se observa que el 97.3% de las llamadas son enlazadas lo cual muestra una eficiencia general en ese sentido.

Proporción de Llamadas Conectadas vs. Perdidas



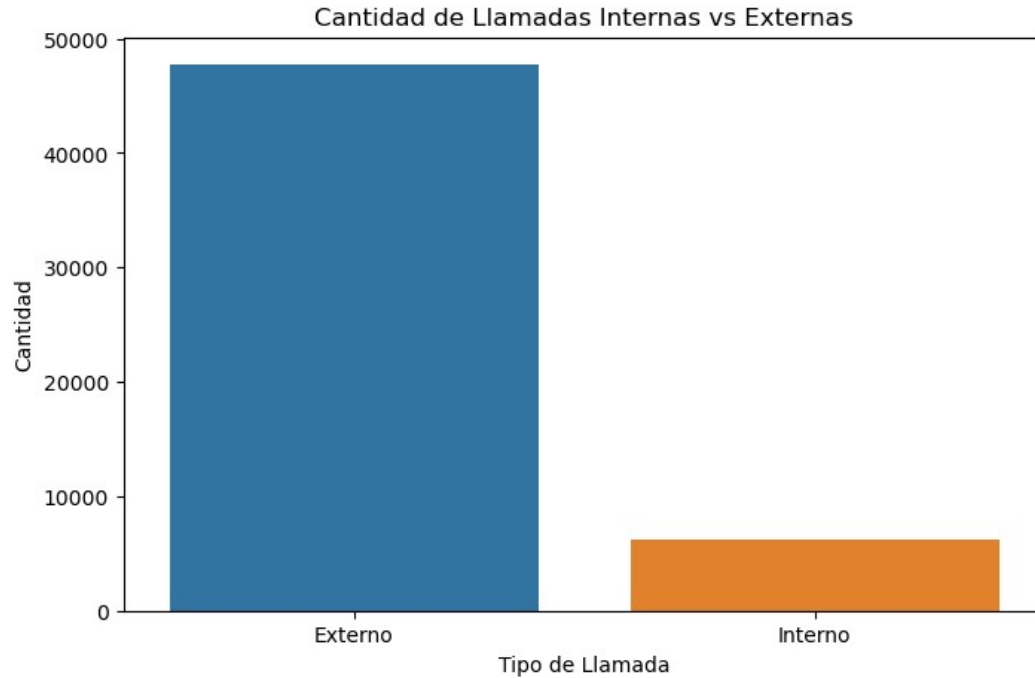


- Observamos que la duración promedio de las llamadas externas es mayor que las internas. (Confirmado con prueba estadística)

T-statistic: -46.96548101210464, P-value: 0.0

Existe una diferencia estadísticamente significativa: las llamadas internas y externas tienen duraciones diferentes.

📌 Las llamadas externas tienen una duración mayor en promedio.



- Confirmamos de manera visual que las llamadas externas son mayores que las internas.

- El presente proyecto tuvo como objetivo proveer informacion que pueda ayudar a tomar desiones empresariales.
 - Se confirma que hay una mayor cantidad de llamadas externas y estas tienen una duracion mayor
 - Se confirma que la mayoria de las llamadas fueron enlazadas (97%)
 - Se identifican operadores mas inactivos y con mayor cantidad de llamadas perdidas

