

# Proyecto Final Análisis servicio de telefonía virtual CallMeMaybe

---

*Por: Francisco Pinto*



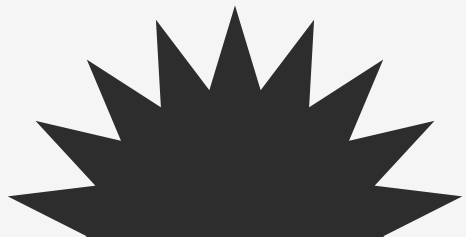
# *Contexto y Necesidad*

---

## Identificando ineficiencias en operadores del servicio de telefonía virtual CallMeMaybe

CallMeMaybe es un servicio de telefonía virtual que permite a empresas gestionar grandes volúmenes de llamadas entrantes, salientes e internas.

La empresa está desarrollando una funcionalidad para supervisores que permita identificar operadores ineficaces, para mejorar la satisfacción del cliente y la eficiencia operativa. Este análisis tiene como objetivo abordar esta necesidad empresarial específica.

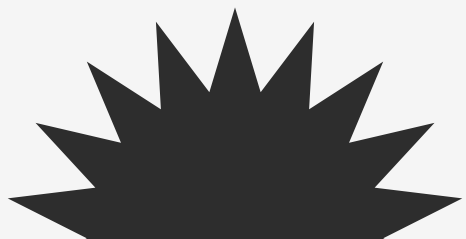


# *Metodología de Ineficiencia*

---

## Establecimiento de umbrales y criterios para identificación de operadores

Para identificar operadores ineficaces, se han definido umbrales utilizando percentiles. Los operadores que presentan al menos dos condiciones críticas, como una tasa de llamadas perdidas superior al P90 y un tiempo de espera superior a 546.65 segundos, son considerados ineficaces. Este enfoque permite focalizar las intervenciones efectivamente.

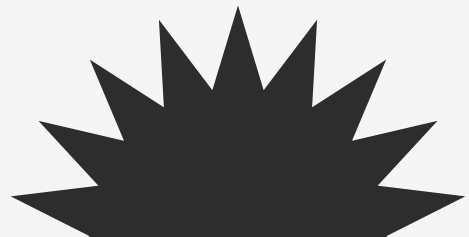


# *Análisis Estadístico*

---

## Validación de Diferencias en Desempeño de Operadores Ineficientes

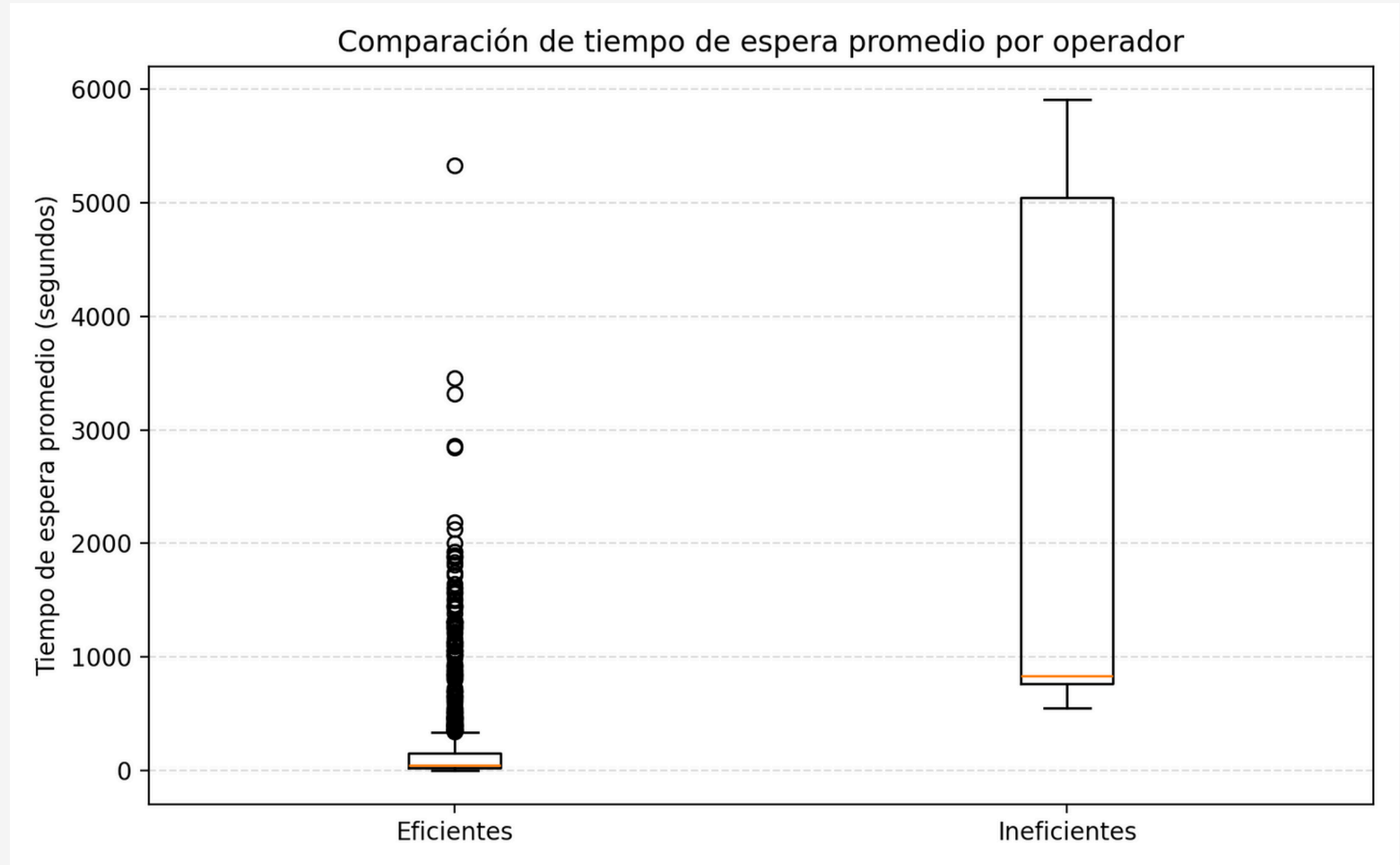
Este análisis presenta un **resumen de las pruebas de hipótesis** realizadas para identificar las diferencias significativas entre los operadores ineficientes y el resto del grupo. Los resultados del test Mann-Whitney U mostraron p-values menores a 0.00001 para el tiempo de espera y la tasa de llamadas perdidas, confirmando la efectividad de los indicadores utilizados.



# Análisis de Desempeño

## Comparación de operadores: tiempos de espera y llamadas perdidas

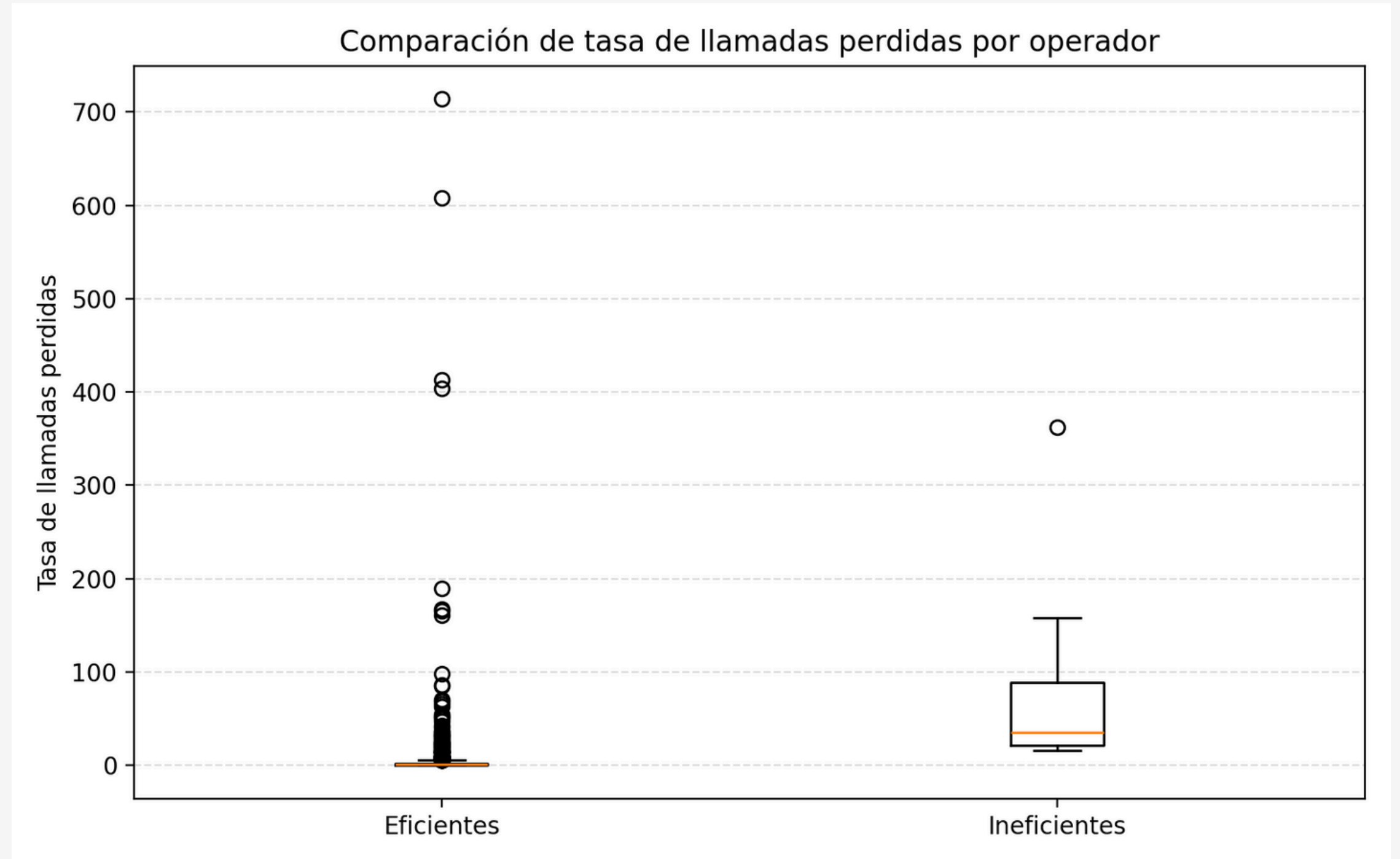
Este análisis presenta gráficos que **ilustran claramente** la diferencia en el tiempo promedio de espera y la tasa de llamadas perdidas entre operadores ineficientes y eficaces, destacando áreas de mejora.



# Análisis de Desempeño

## Relación entre tasa de llamadas y tiempo de espera

Este análisis visualiza la **correlación significativa** entre la tasa de llamadas perdidas y el tiempo de espera de los operadores, revelando patrones cruciales para la mejora del servicio.





# Conclusiones y Recomendaciones

---

## Identificación de patrones de ineficiencia y pasos para mejoras operativas

Este análisis ha revelado que aproximadamente el **10% de los operadores** presentan un rendimiento ineficaz. Se recomienda implementar una herramienta de seguimiento basada en indicadores clave que permitirá a los supervisores monitorear el desempeño y realizar intervenciones efectivas para mejorar la eficiencia operativa.

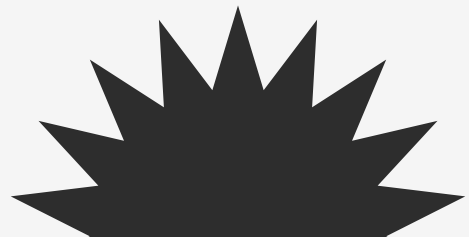
Recomendaciones de negocio:

Implementar alertas automáticas para operadores críticos El sistema debería notificar cuando un operador supere los umbrales.

Revisar cargas de trabajo y asignación de llamadas

Identificar necesidades de capacitación

Un programa de capacitación personalizado podría mejorar el desempeño.



# *Fuentes de referencia utilizadas en el proyecto*

## 1. Pandas – GroupBy y agregaciones

Fuente: Pandas Documentation

Enlace: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.groupby.html>

Esta documentación fue utilizada para comprender y aplicar la agregación de datos por operador (`operator_id`), permitiendo calcular métricas como total de llamadas, llamadas entrantes, llamadas salientes y llamadas perdidas mediante funciones como `groupby()` y `agg()`. Fue clave para construir la tabla `operator_stats`.

## 2. Pandas – Manejo de valores nulos

Fuente: Pandas Documentation

Enlace: [https://pandas.pydata.org/docs/user\\_guide/missing\\_data.html](https://pandas.pydata.org/docs/user_guide/missing_data.html)

Se utilizó para manejar correctamente valores faltantes en columnas como `operator_id` y para aplicar estrategias como `dropna()` antes de realizar cálculos estadísticos y pruebas de hipótesis, garantizando la validez de los resultados.

## 3. Pandas – Función `describe()`

Fuente: Pandas Documentation

Enlace: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.describe.html>

Esta referencia permitió utilizar correctamente la función `describe(include="all")` para obtener estadísticas descriptivas tanto de variables numéricas como categóricas, apoyando la fase de análisis exploratorio de datos (EDA).



# *Fuentes de referencia utilizadas en el proyecto*

## *4. NumPy – Operaciones condicionales vectorizadas*

*Fuente: NumPy Documentation*

*Enlace: <https://numpy.org/doc/stable/reference/generated/numpy.where.html>*

*Se empleó esta documentación para crear columnas auxiliares de forma eficiente, como `incoming_calls`, `outgoing_calls` y `missed_calls`, utilizando operaciones vectorizadas en lugar de bucles, lo que mejoró la claridad y el rendimiento del código.*

## *5. SciPy – Prueba Mann–Whitney U*

*Fuente: SciPy Documentation*

*Enlace: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.mannwhitneyu.html>*

*Fue utilizada para seleccionar y aplicar correctamente la prueba estadística Mann–Whitney U, empleada para contrastar las hipótesis relacionadas con diferencias significativas en el tiempo de espera y la tasa de llamadas perdidas entre operadores eficaces e ineficaces.*

## *6. Khan Academy – Percentiles y distribución de datos*

*Fuente: Khan Academy – Statistics & Probability*

*Enlace: <https://www.khanacademy.org/math/statistics-probability/summarizing-quantitative-data/percentiles>*

*Esta fuente se utilizó para reforzar la interpretación conceptual de los percentiles y justificar su uso como método objetivo para definir umbrales de ineficiencia ( $P_{90}$  y  $P_{10}$ ) dentro de la distribución de los indicadores operativos.*

## *7. Matplotlib – Boxplots*

*Fuente: Matplotlib Documentation*

*Enlace: [https://matplotlib.org/stable/api/\\_as\\_gen/matplotlib.pyplot.boxplot.html](https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.boxplot.html)*

*Se utilizó esta documentación para crear boxplots comparativos entre operadores eficientes e ineficientes, permitiendo visualizar de forma clara las diferencias en tiempo de espera y tasa de llamadas perdidas que luego fueron confirmadas mediante pruebas estadísticas.*