



Universidad de las Américas

TTMZO411-93-PROYECTO MBD  
CAPSTONE PROJECT

# EVALUACIÓN DE LA EFICACIA DE CAMPAÑAS TELEFÓNICAS PARA PROSPECTOS EN INSTITUCIONES DE EDUCACIÓN SUPERIOR EN ECUADOR MEDIANTE MODELOS PROBABILÍSTICOS DE CLASIFICACIÓN BINARIA

Presentado por:  
**Luis Felipe Contreras**  
**Ana Belén Erazo**

# INTRODUCCIÓN

01



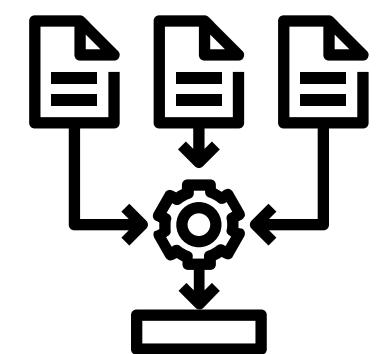
## Antecedentes

- En Ecuador, las universidades enfrentan una alta competencia por atraer estudiantes y usan campañas telefónicas, pero su eficacia es incierta.
- La finalidad es identificar las estrategias más efectivas para convertir prospectos en inscripciones, ya que sin un modelo predictivo, se corre el riesgo de una inversión ineficiente y pérdida de oportunidades.

# Objeto de estudio



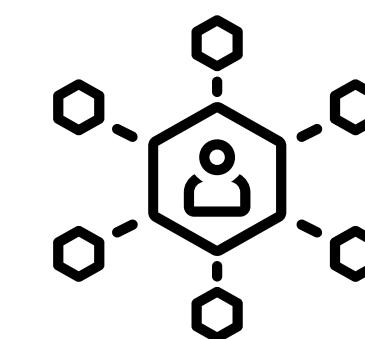
**Evaluar Eficacia  
de interacciones  
de prospectos**



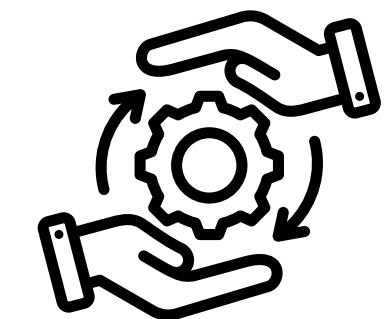
**Determinar los  
modelos  
predictivos más  
efectivos**



**Factores más  
influyentes**



**Identificar  
patrones y  
características de  
los prospectos**



**Transformar de  
prospectos a  
estudiantes**

# Objetivos

## Objetivo Principal

Evaluar la eficacia de las campañas telefónicas dirigidas a prospectos en instituciones de educación superior en Ecuador, utilizando modelos probabilísticos de clasificación binaria para identificar y predecir los factores que influyen en la conversión de prospectos a estudiantes.

03

1

Analizar datos históricos de campañas telefónicas en universidades para detectar patrones y tendencias en la respuesta de los prospectos.

2

Identificar las características demográficas y conductuales de los prospectos que influyen en su decisión de inscribirse tras una llamada telefónica.

3

Aplicar y comparar modelos probabilísticos de clasificación binaria para predecir la respuesta positiva de un prospecto a una campaña telefónica.

4

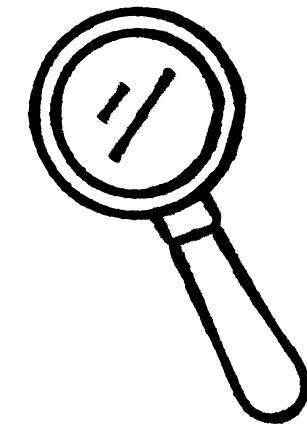
Proponer recomendaciones para optimizar campañas telefónicas y maximizar la conversión, basadas en el análisis y modelado de datos.



# Datos

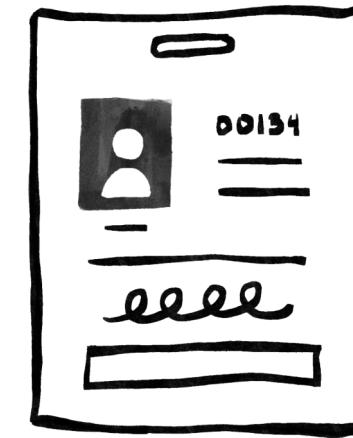
1

**BASE DE DATOS DE  
UNA INSTITUCIÓN  
DE EDUCACIÓN  
SUPERIOR EN  
ECUADOR**



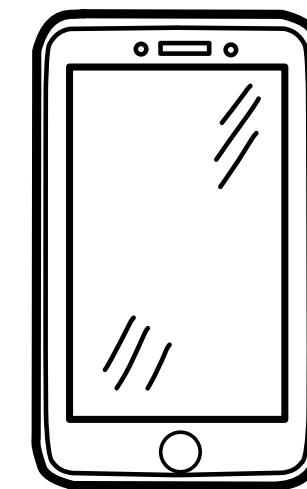
3

**INCLUYE  
REGISTROS  
ANONIMIZADOS,  
DATOS  
DEMOGRÁFICOS,  
HISTÓRICOS Y DE  
LLAMADAS  
TELEFÓNICAS**



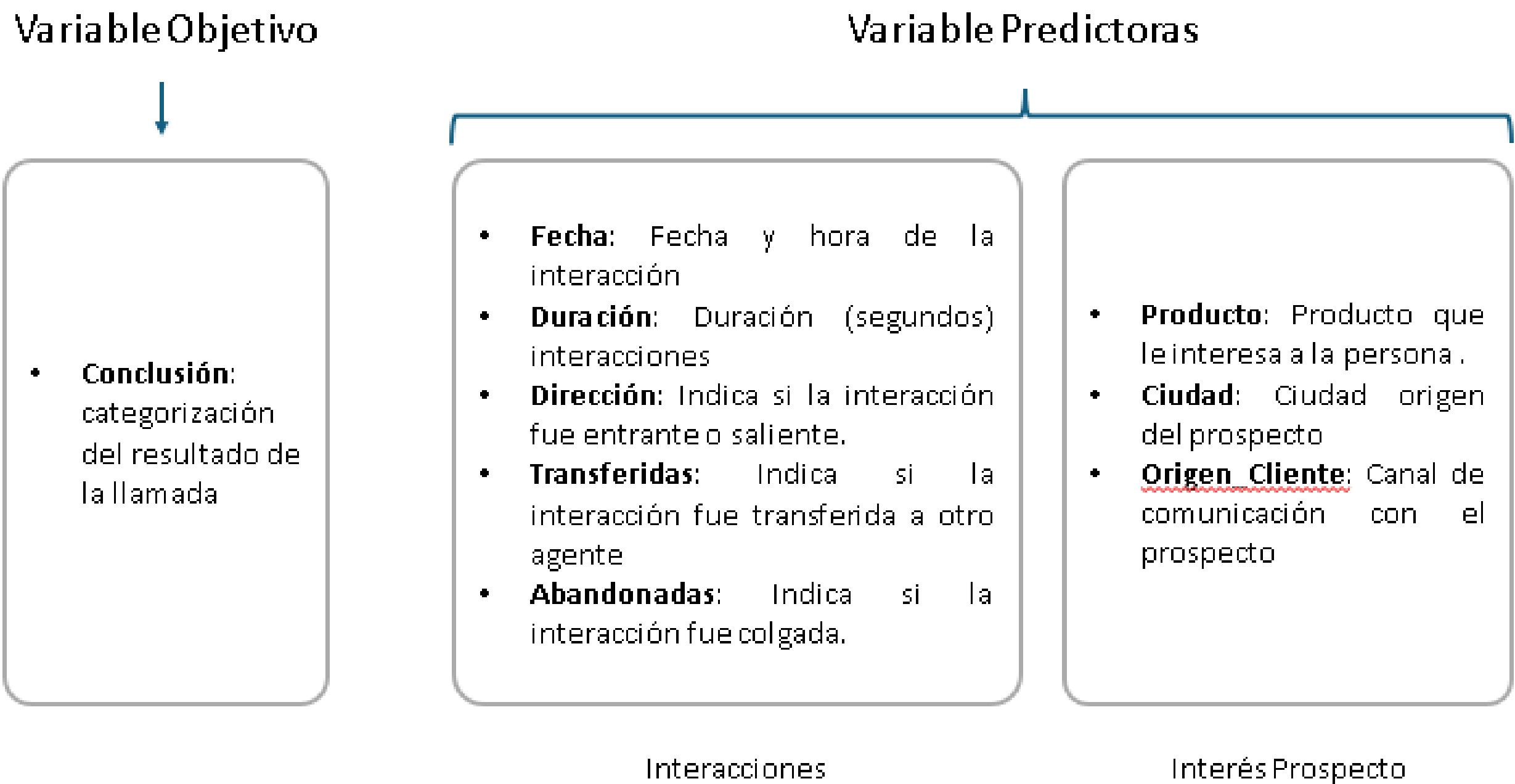
2

**INFORMACIÓN  
SOBRE  
INTERACCIONES  
TELEFÓNICAS**



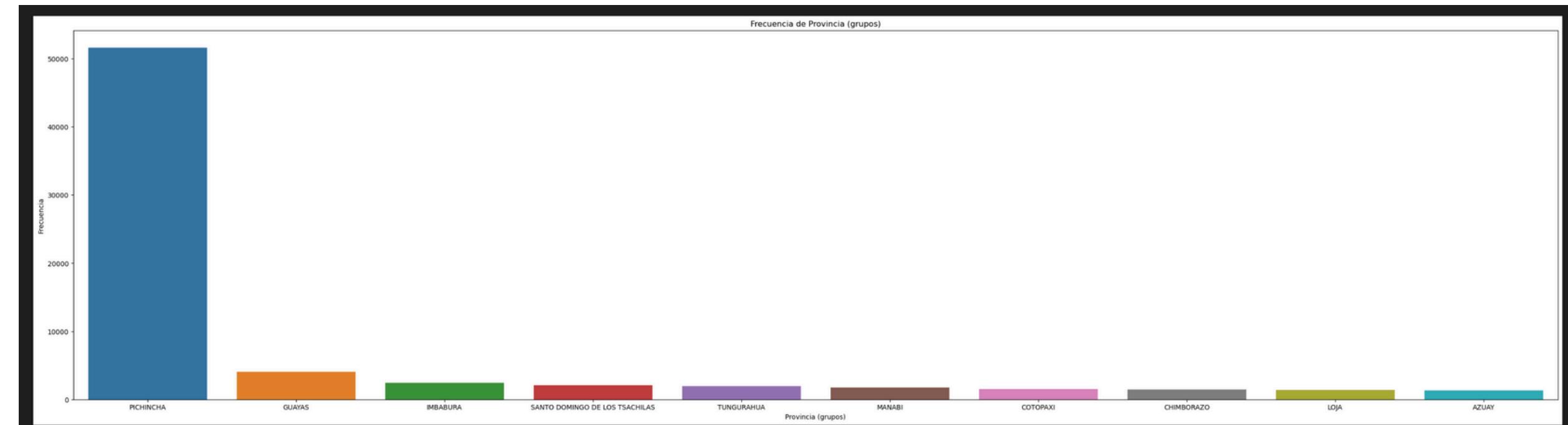
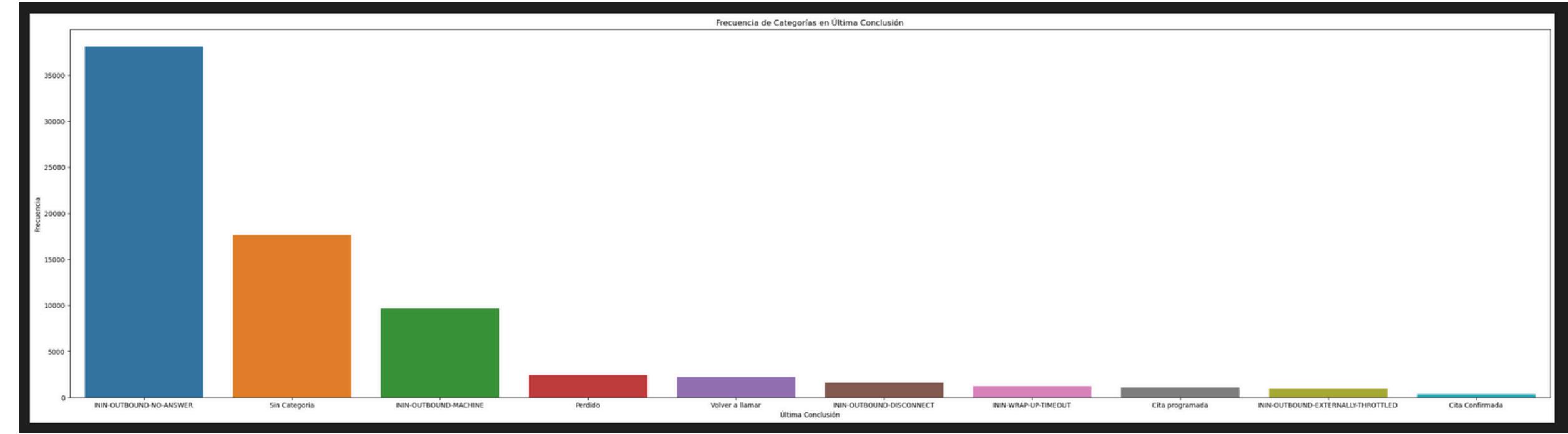
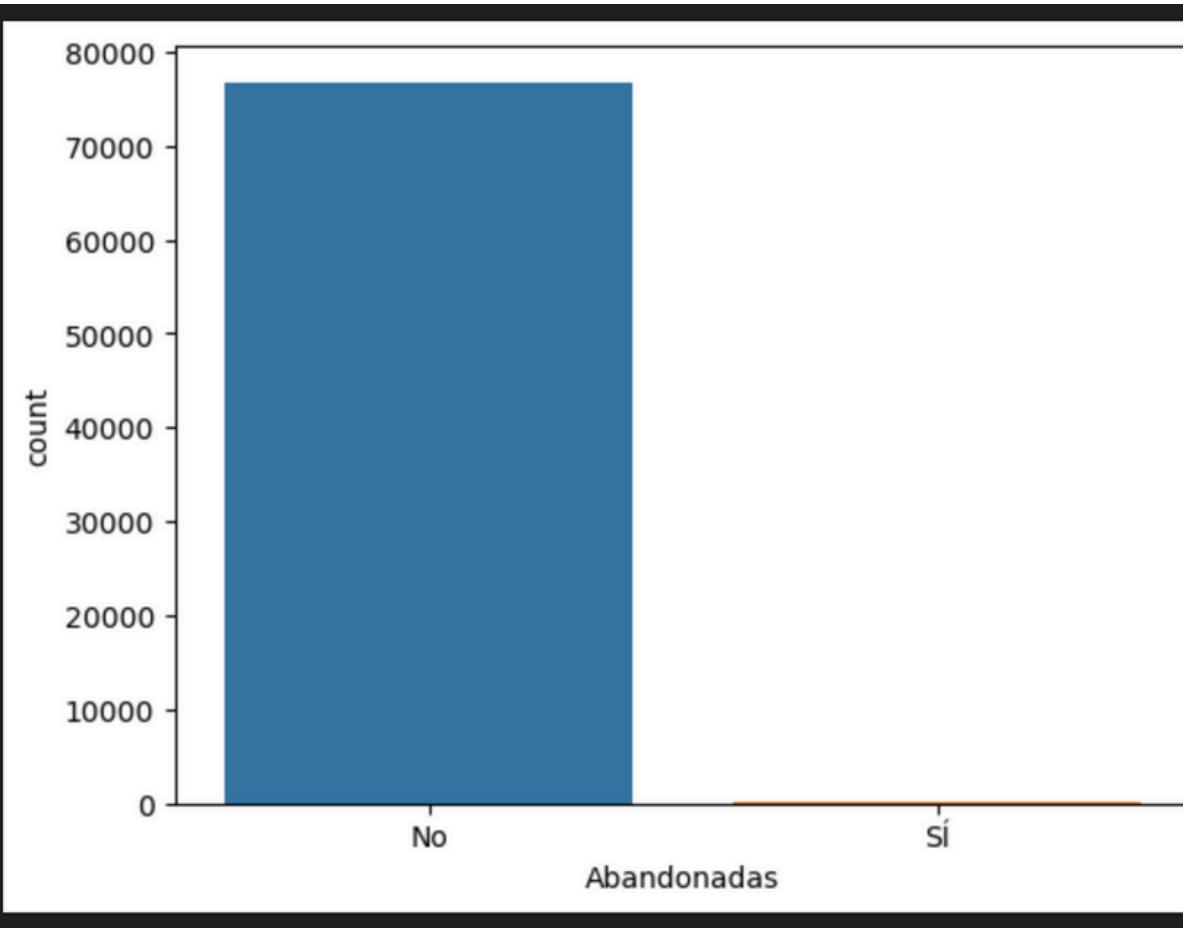
# Selección de la base de datos

La base de datos original contiene 502.745 registros clasificados en 15 campos



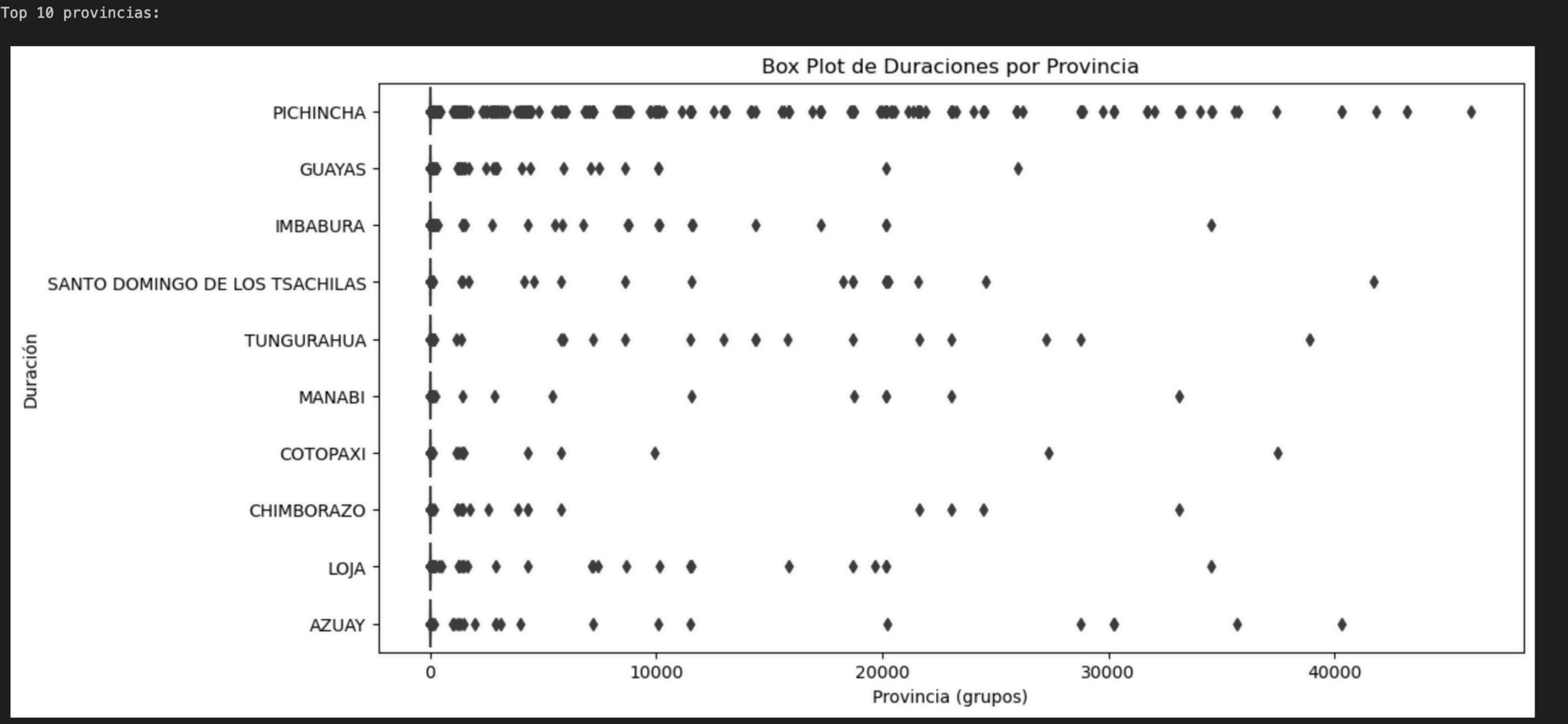
06

# Análisis Exploratorio de Datos

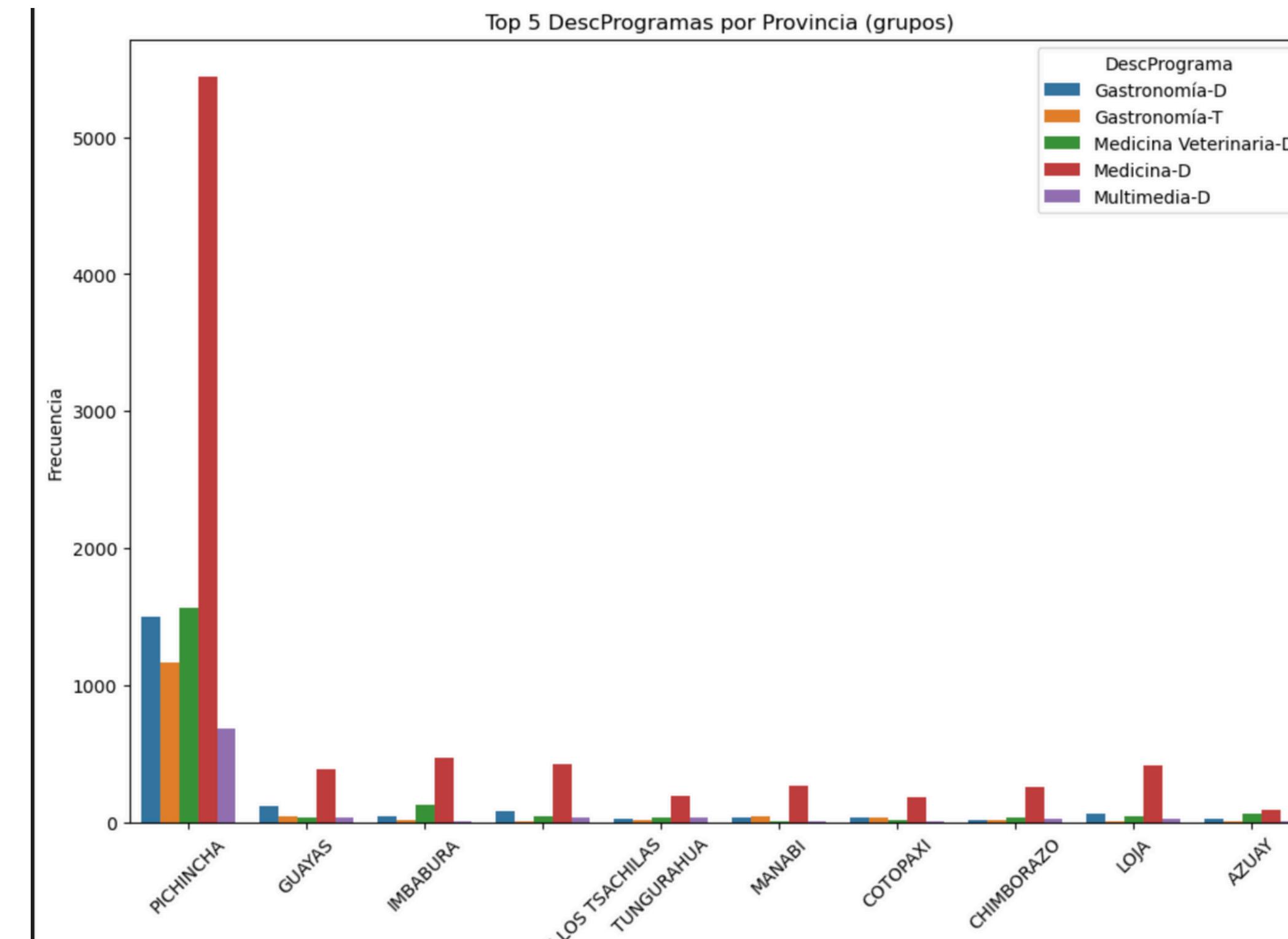


# Análisis Exploratorio de Datos

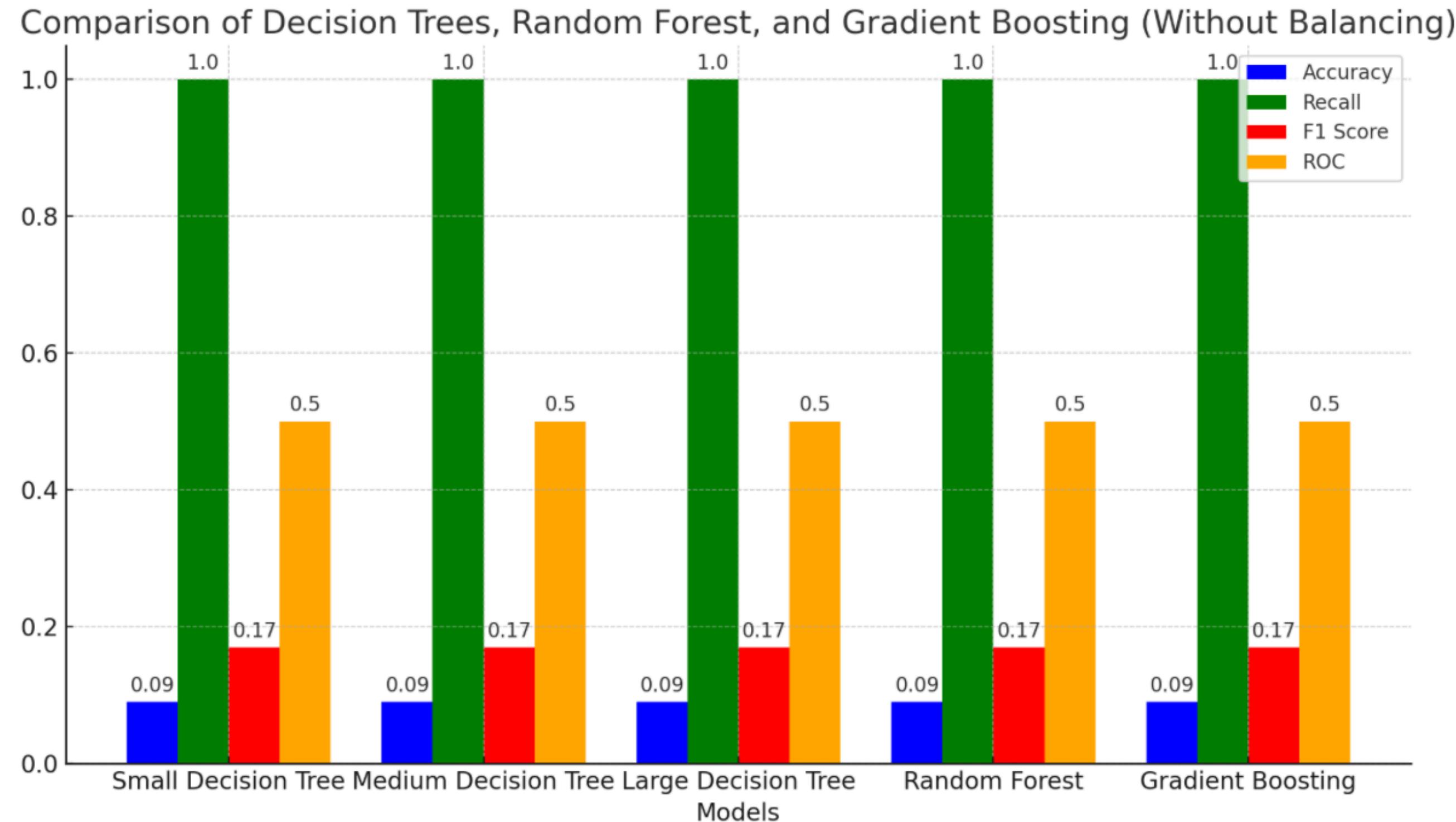
Top 10 provincias:



# Análisis Exploratorio de Datos

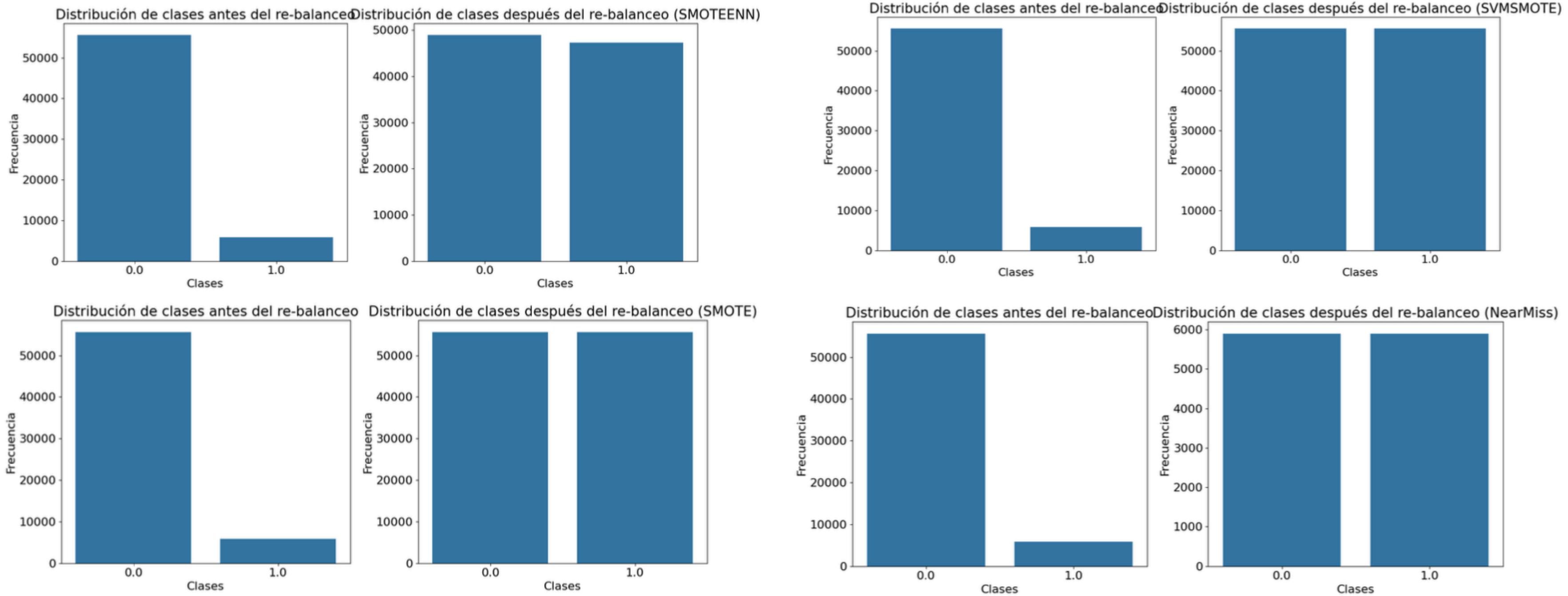


# Resultados

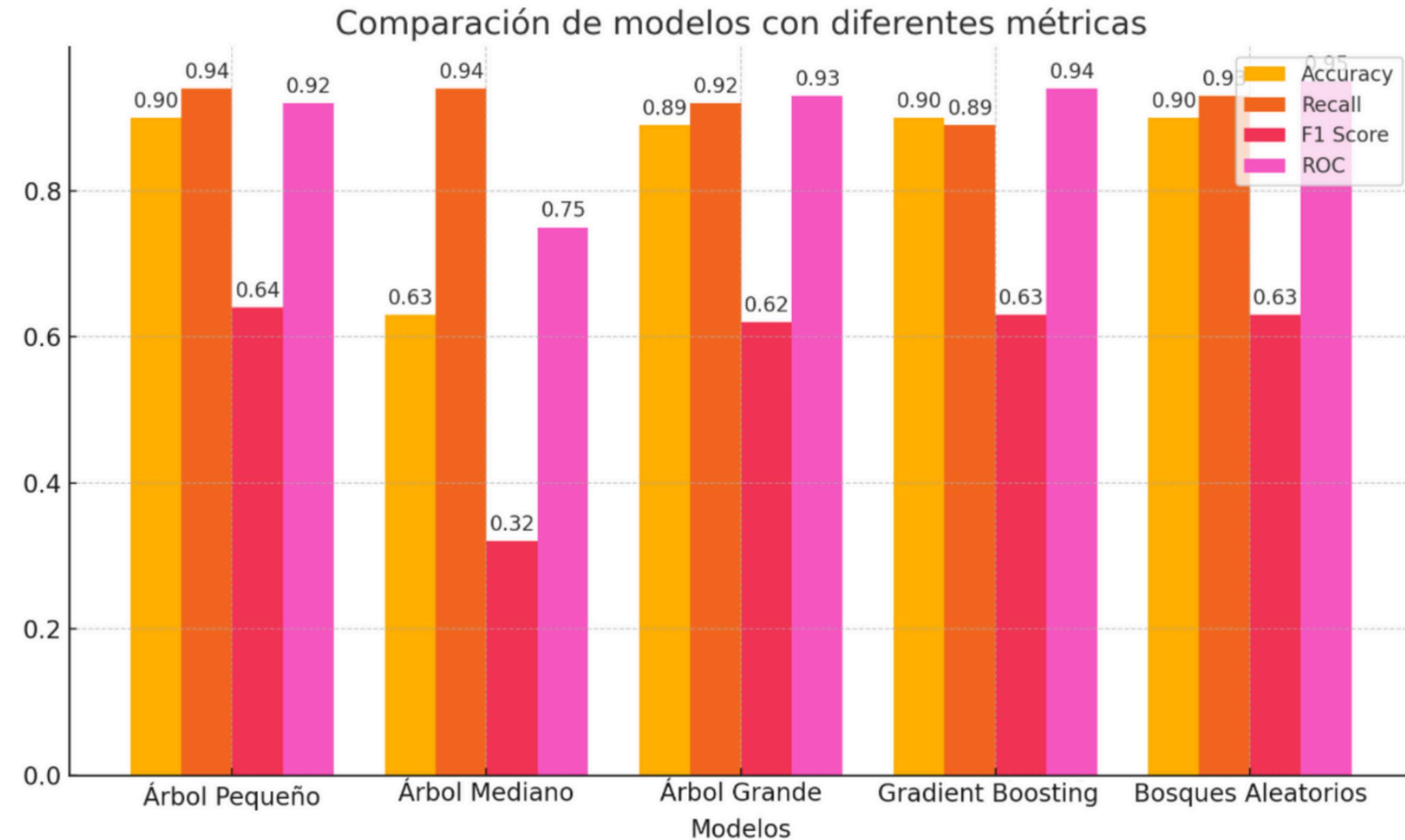


# Resultados

## Balanceo



# Resultados



# Discusión de los resultados y propuesta de solución

08

## MEJOR HORARIO PARA REALIZAR LLAMADAS

**Datos:** Mayor tasa de respuesta entre 9:00-11:00 AM y 4:00-6:00 PM.

**Recomendación:** Enfocar campañas telefónicas en esos períodos para un 25% más de éxito.

MEJORA DE ESTRATEGIAS PARA PROGRAMAS ACADÉMICOS ESPECÍFICOS

**Datos:** Medicina, Economía y Marketing tienen mayor tasa de respuesta.

**Recomendación:** Personalizar campañas para estos programas, con estrategias más directas.

## OPTIMIZACIÓN DEL SEGUIMIENTO DE PROSPECTOS

**Datos:** Llamadas repetidas tras 2-3 horas aumentan la respuesta un 20%.

**Recomendación:** Realizar seguimientos dentro de esas horas, en los mejores horarios identificados.

AUTOMATIZACIÓN PARA LLAMADAS ABANDONADAS

**Datos:** Altas tasas de abandono por espera prolongada.

**Recomendación:** Implementar mensajes automatizados y programación de seguimientos automáticos.

## MEJORES DÍAS PARA REALIZAR LLAMADAS

**Datos:** Miércoles y jueves tienen un 30% más de éxito que lunes y viernes.

**Recomendación:** Concentrar esfuerzos de llamadas en estos dos días clave.

PRIORIZACIÓN DE PROSPECTOS CON MAYOR PROBABILIDAD DE RESPUESTA

**Datos:** Modelo de Bosques Aleatorios con AUC-ROC de 0.95.

**Recomendación:** Priorizar prospectos con mayor probabilidad de éxito en llamadas.

## SEGMENTACIÓN GEOGRÁFICA Y TEMPORAL

**Datos:** Pichincha, Guayas e Imbabura tienen mayor éxito en las campañas.

**Recomendación:** Priorizar estas provincias en los horarios más efectivos.

OPTIMIZACIÓN DEL TIEMPO DE CONVERSACIÓN

**Datos:** Llamadas de 1-3 minutos son las más efectivas.

**Recomendación:** Capacitar a agentes para mantener las llamadas dentro de este rango.

# CONCLUSIONES

1. El Random Forest fue el más efectivo con un AUC de 95% y precisión del 90%. Equilibró bien entre precisión y recall, siendo mejor que otros modelos.
2. Muchas llamadas no fueron contestadas, afectando la efectividad de las campañas. Los modelos predictivos como Random Forest y Gradient Boosting permiten predecir con precisión cuáles llamadas tienen más probabilidad de ser contestadas.
3. Los prospectos de Medicina, Derecho y Gastronomía de áreas urbanas tienen más probabilidades de responder. La duración de llamadas previas y la última categoría de conclusión son indicadores clave.

# RECOMENDACIONES

1. **Priorización de prospectos:** Usar el modelo predictivo para clasificar prospectos en tres grupos:
  - Grupo A (Alta probabilidad): Priorizar llamadas en horarios clave.
  - Grupo B (Media probabilidad): Intentar contacto adicional por correo o SMS.
  - Grupo C (Baja probabilidad): Reducir prioridad y usar canales no telefónicos.

2. **Automatización:** Configurar recordatorios automáticos para prospectos de media y baja prioridad, optimizando recursos del call center.

**Ventajas:** Ahorro de recursos, mejora en tasa de conversión y optimización del tiempo de los agentes.

# Bibliografía

- Backenköhler, M., & Wolf, V. (2017). Student performance prediction and optimal course selection: An MDP approach. International Conference on Software Engineering and Formal Methods, 40–47.
- Cui, Y., Chen, F., Shiri, A., & Fan, Y. (2019). Predictive analytic models of student success in higher education: A review of methodology. *Information and Learning Sciences*, 120(3/4), 208–227.
- Gray, C. C., & Perkins, D. (2019). Utilizing early engagement and machine learning to predict student outcomes. *Computers & Education*, 131, 22–32.
- Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. *The Internet and Higher Education*, 37, 66–75.
- Miguéis, V. L., Freitas, A., Garcia, P. J., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems*, 115, 36–51.
- Zhao, C., Yang, J., Liang, J., & Li, C. (2016). Discover learning behavior patterns to predict certification. In International Conference on Computer Science & Education, IEEE, 69–73.
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284.
- Cui, Y., Howard, J., & Perkins, M. (2019). Modelos predictivos en campañas telefónicas educativas: Un análisis comparativo. *Journal of Educational Research*, 8(4), 256–275.
- Zhao, X., Backenköhler, M., & Wolf, R. (2016). Random Forest en la segmentación de prospectos: Un enfoque predictivo en educación superior. *Journal of Data Science*, 14(2), 125–145.

# **Gracias a todos**