

Tasa de  
matriculación

Alejandro  
Acosta León

Introducción

Datos

Var. Objetivo

Var. explicativas

Estadística  
descriptiva

Ciclo

Colegio

Modelos

Resultados

Accuracy

Importancia variables

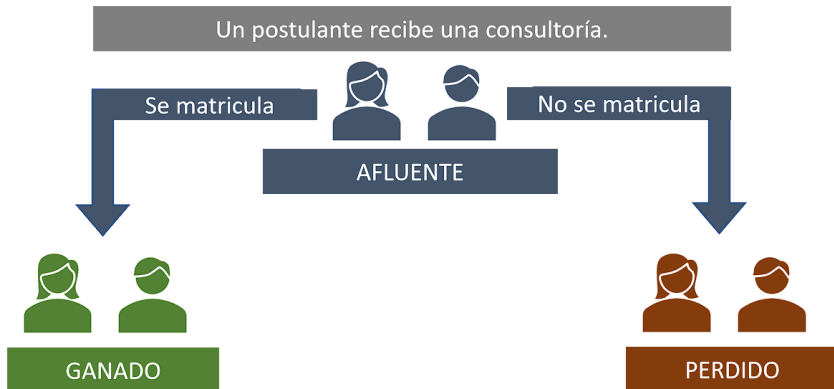
Conclusiones

# Tasa de matriculación de alumnos universitarios pregrado

Alejandro Acosta León

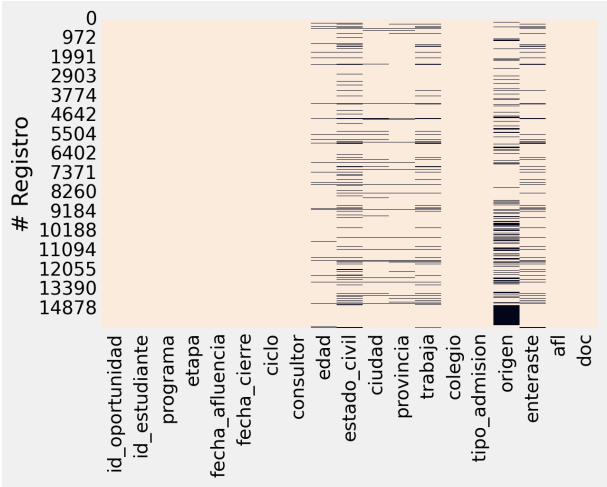
30 de marzo de 2022

# El proceso de admisión



$$conv = \frac{Ganados}{Afluentes}$$

# Datos e imputación



Tras la limpieza se tienen 12626 Observaciones, 9.3 % de registros con valores perdidos.

Introducción

Datos

- Var. Objetivo
- Var. explicativas

Estadística  
descriptiva

- Ciclo
- Colegio

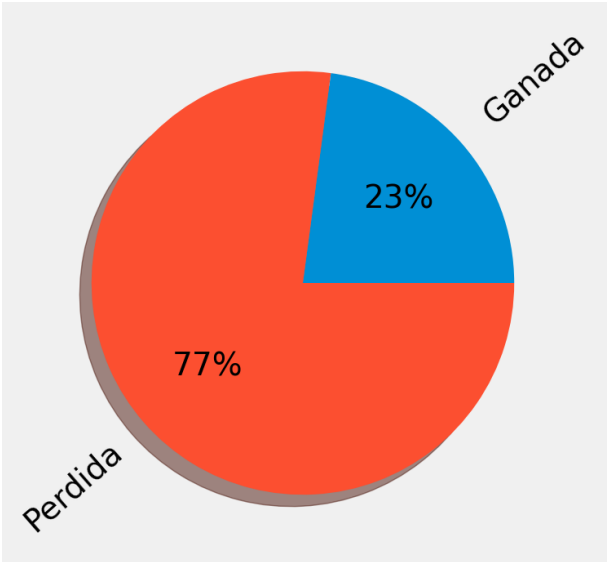
Modelos

Resultados

- Accuracy
- Importancia variables

Conclusiones

Variable objetivo



## Variables del proceso de admisión:

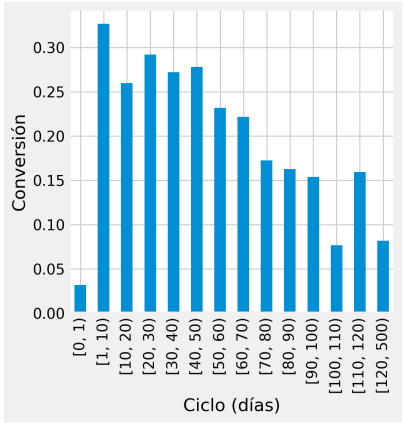
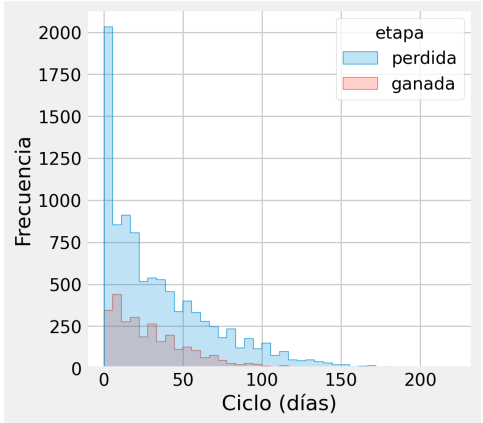
- ➊ **Programa:** código del programa académico.
- ➋ **Ciclo:** tiempo en días transcurrido entre la fecha de afluencia y la de cierre.
- ➌ **Consultor:** código del consultor que le atendió.
- ➍ **Tipo\_Admision:** Normal o convalidado.
- ➎ **Enteraste:** ¿cómo te enteraste de la universidad?

## Variables demográficas del postulante

- ➊ **Estado\_Civil**
- ➋ **Edad**
- ➌ **Provincia**
- ➍ **Trabaja:** 1 si trabaja, 0 caso contrario.
- ➎ **Colegio:** código del colegio donde estudió.

# Hallazgos importantes

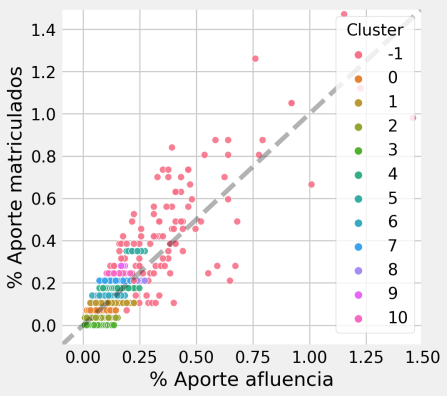
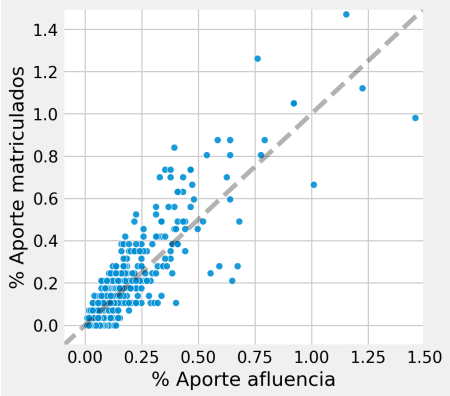
Ciclo



Un ciclo corto (¡pero no tan corto!) incrementa la probabilidad de matriculación.

# Hallazgos importantes

Colegio



Los colegios comparten características respecto al aporte en afluencia y matriculación.

Ya que la variable objetivo existe en el dataset y es binaria, se requiere un modelo supervisado de clasificación.

- Ratio de 70 % para entrenamiento.
- El número de estimadores óptimo se encontrará mediante un Grid Search.

---

## Random Forest

N-Estimators: 800

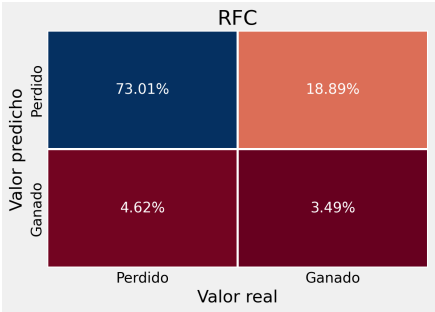
## XGBoost

N-Estimators: 200



# Resultados

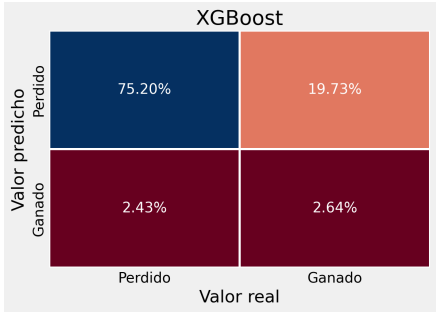
Accuracy del modelo



**Accuracy: 76 %**

**Sensitivity: 16 %**

**Specificity: 94 %**



**Accuracy: 77 %**

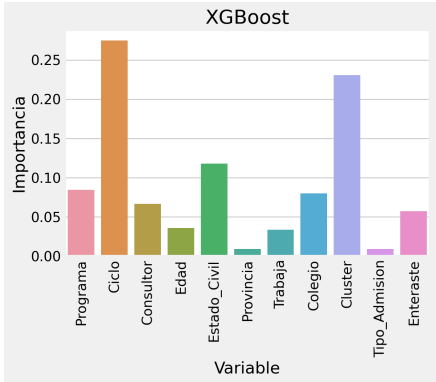
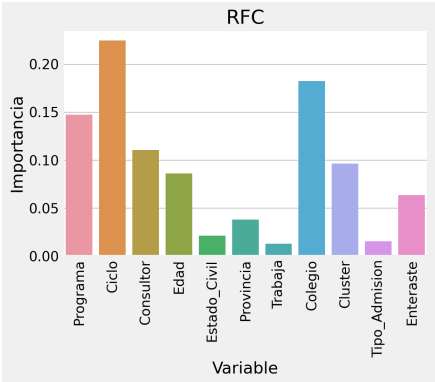
**Sensitivity: 12 %**

**Specificity: 97 %**

El modelo XGBoost mejora la predicción de perdidos pero empeora la predicción de ganados.

# Resultados

## Importancia de la variable



Ciclo y colegio son las variables mas importantes.

# Conclusiones

- ¡El ciclo de venta importa!
- Mantener buenas relaciones y convenios con colegios claves puede mejorar la conversión.
- La provincia aparentemente no es relevante, pero esto puede deberse a que Pichincha está sobre representado.

## Siguientes pasos

- La variable objetivo podría cambiar al % de beca ofrecida.
- De esta forma se podría estimar una especie de "elasticidad" del postulante.
- Es posible integrar el modelo predictivo CRM, de tal forma que el consultor tenga en tiempo real una recomendación de beca y probabilidad de matriculación basado a las características del postulante.