

Rotación de Personal

Un proyecto de People Analytics

By :AnaMumaq

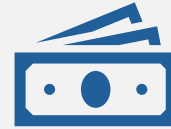
Contenido

- **Introducción**
- Variables consideradas
- Insights
- Correlaciones
- Modelado
- Resultados del modelo
- Importancia de variables
- Resumen

Problemática



Aumento en la rotación del personal en el último año en 10 %



Costos elevados en adquisición de nuevo personal en 30%

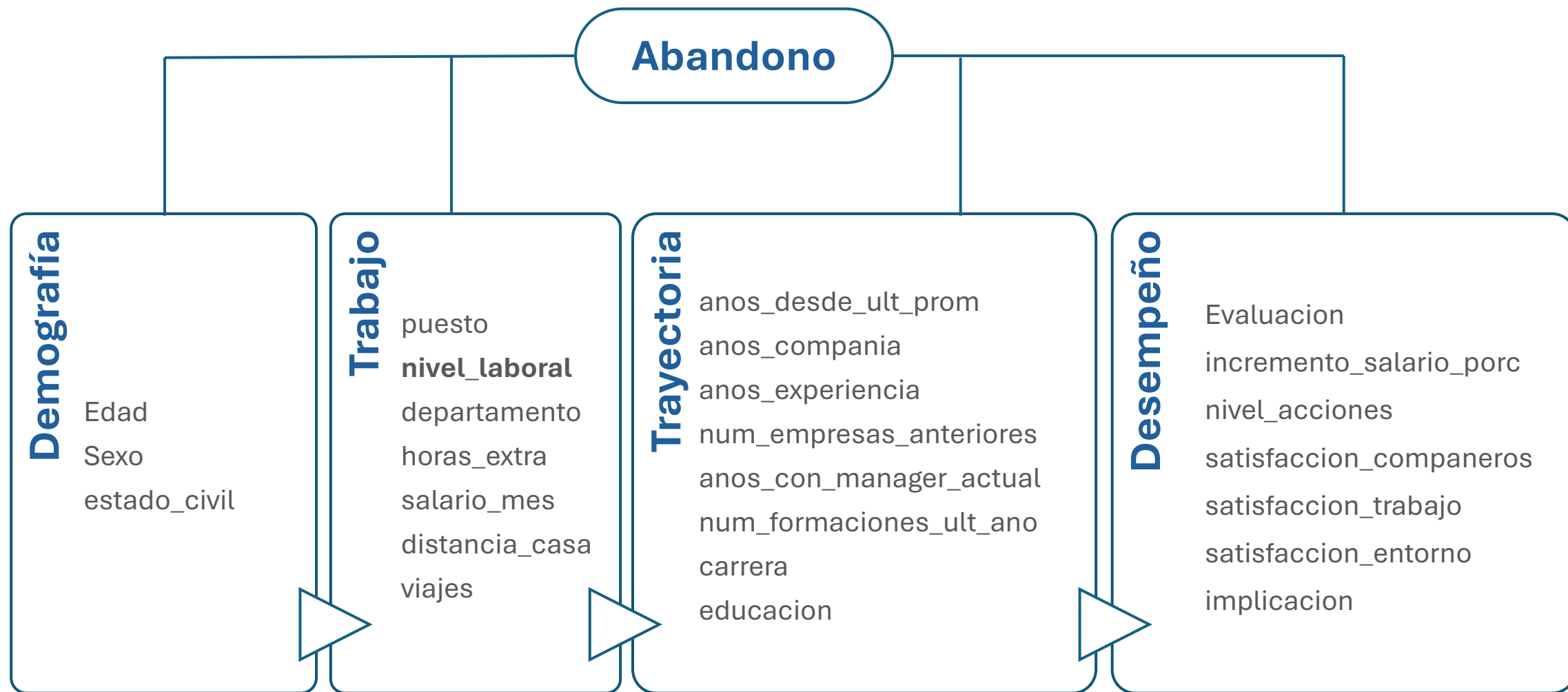


No se están cumpliendo las metas mensuales de rotación en los diferentes departamentos

Contenido

- Introducción
- **Variables consideradas**
- Insights
- Correlaciones
- Modelado
- Resultados del modelo
- Importancia de variables
- Resumen

Variables consideradas



Contenido

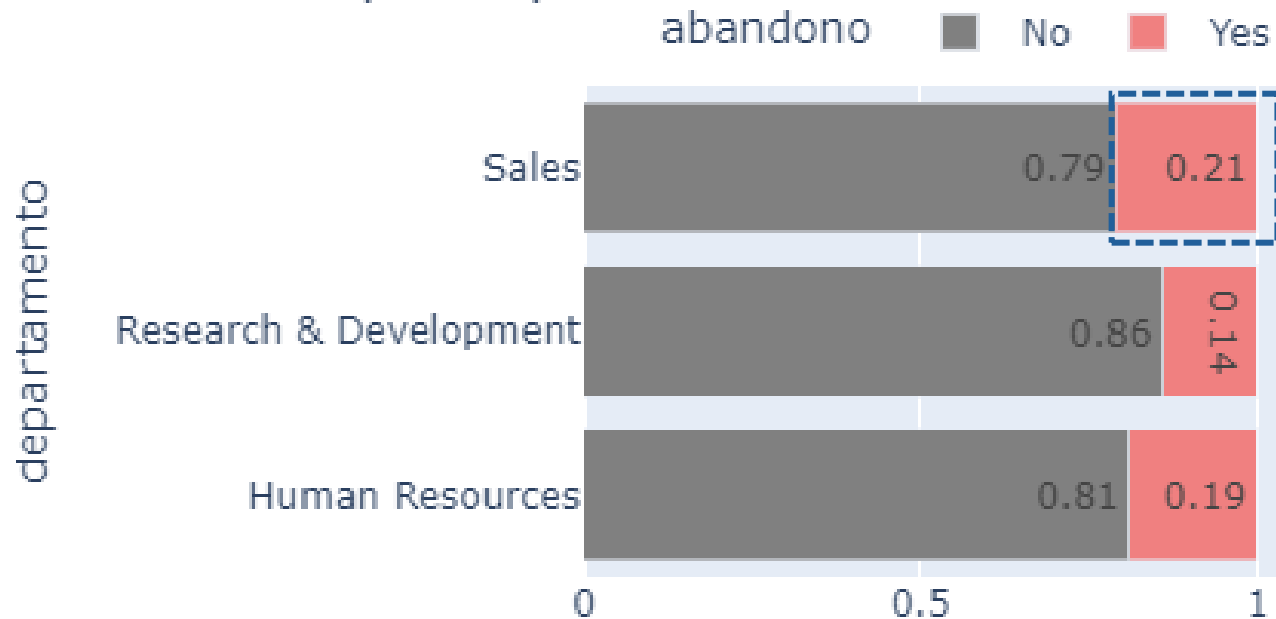
- Introducción
- Variables consideradas
- **Insights**
- Correlaciones
- Modelado
- Resultados del modelo
- Importancia de variables
- Resumen

Abandono global



Abandono
Global: 16%

Abandono por departamento

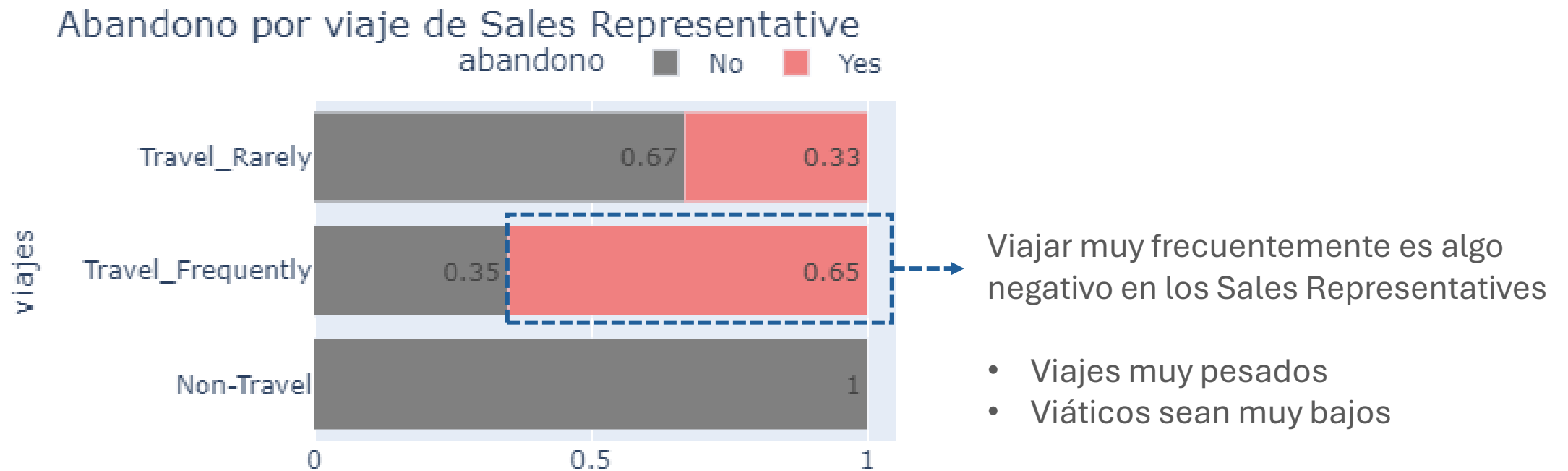


Impacto de rotación global
por el **área de ventas**

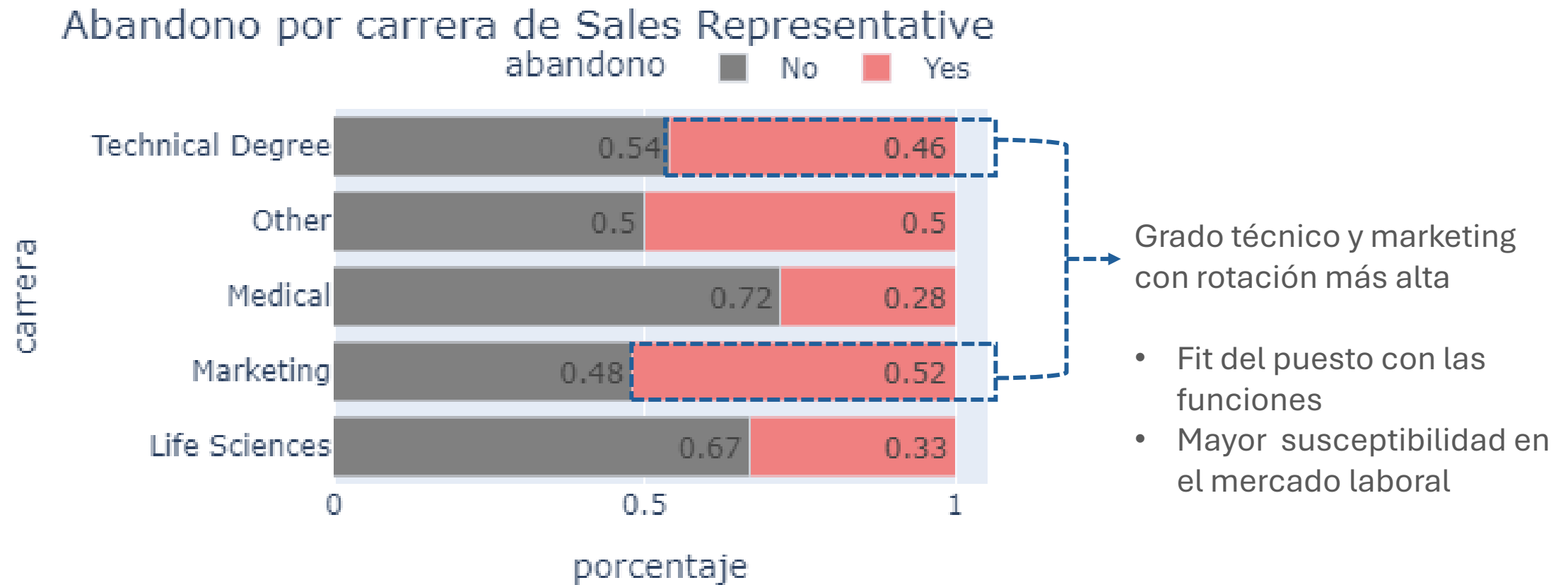
Rotación por puesto:

- Manager 5%
- Sales Executive 17%
- **Sales Representative 40%**

Sales Representatives & Travel

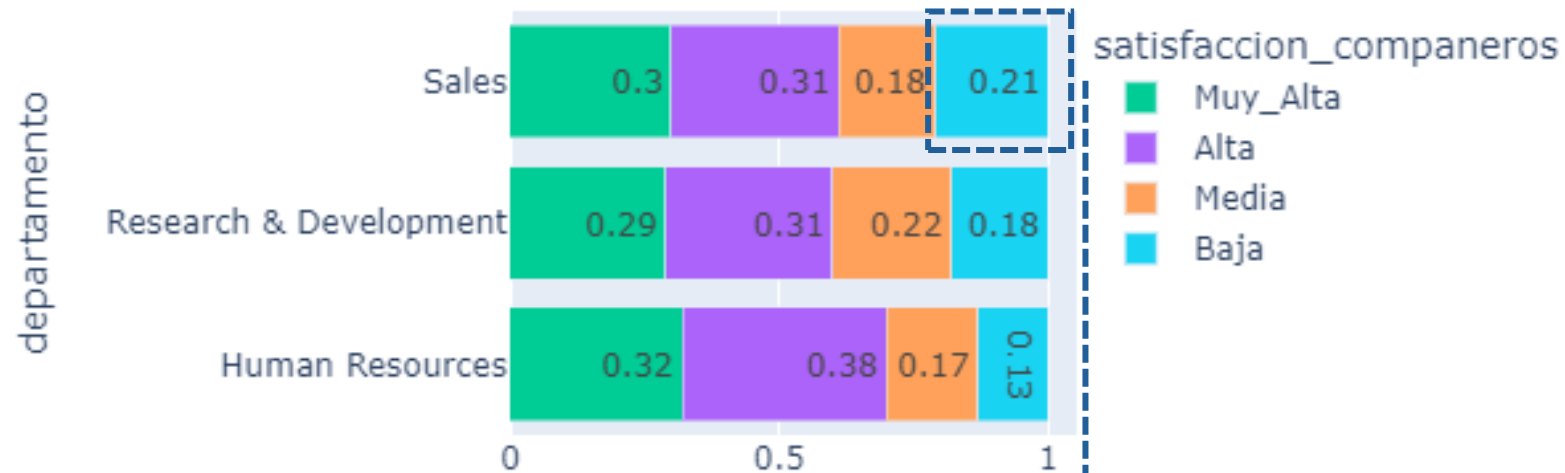


Sales Representatives & Career



Satisfaccion Companeros

Satisfaccion de compañeros por departamento

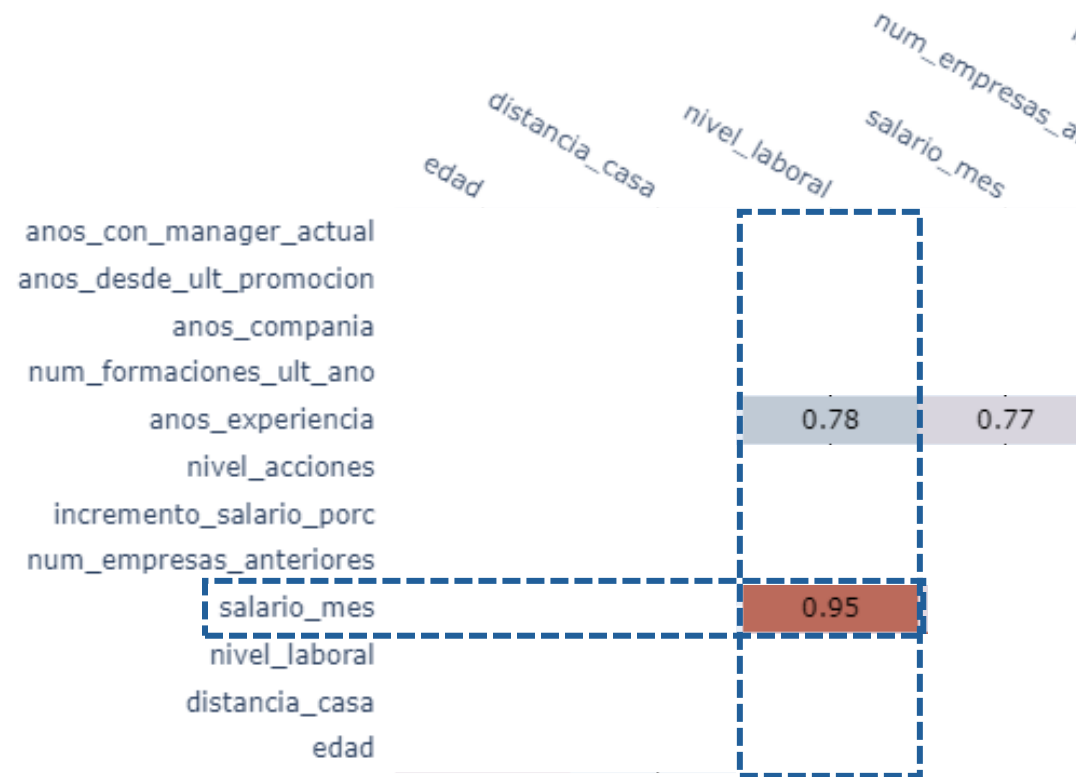


En el departamento de ventas el clima interno es más bajo que el resto

Contenido

- Introducción
- Variables consideradas
- Insights
- **Correlaciones**
- Modelado
- Resultados del modelo
- Importancia de variables
- Resumen

Correlaciones



- Por definición el nivel laboral se define como seniority de un puesto, suele aumentar conforme a los años de experiencia
- Mayores años de experiencia se suele traducir en mejor salario
- **A mayor nivel laboral el salario siempre será mayor en empresas que tengan definida la banda salarial**

Decidimos eliminar la variable problemática de nivel laboral que engloba las otras variables

Contenido:

- Introducción
- Variables consideradas
- Insights
- Correlaciones
- **Modelado**
- Resultados del modelo
- Importancia de variables
- Resumen

Modelos

Modelo Unbalanced

Dataset original con el 16% de casos de renuncia (**180 valores True**)

Modelo Oversampling

Dataset con data sintética con el 50% de casos de renuncia (**922 valores True**)

Modelo Undersampling

Dataset eliminando data con el 50% de casos de renuncia (**180 valores True y 180 valores False**)

Contenido

- Introducción
- Variables consideradas
- Insights
- Correlaciones
- Modelado
- **Resultados del modelo**
- Importancia de variables
- Resumen

Resultados de modelado

Recall (%)

Resultados después de
validación cruzada

75%

**Modelo
Unbalanced**

80%

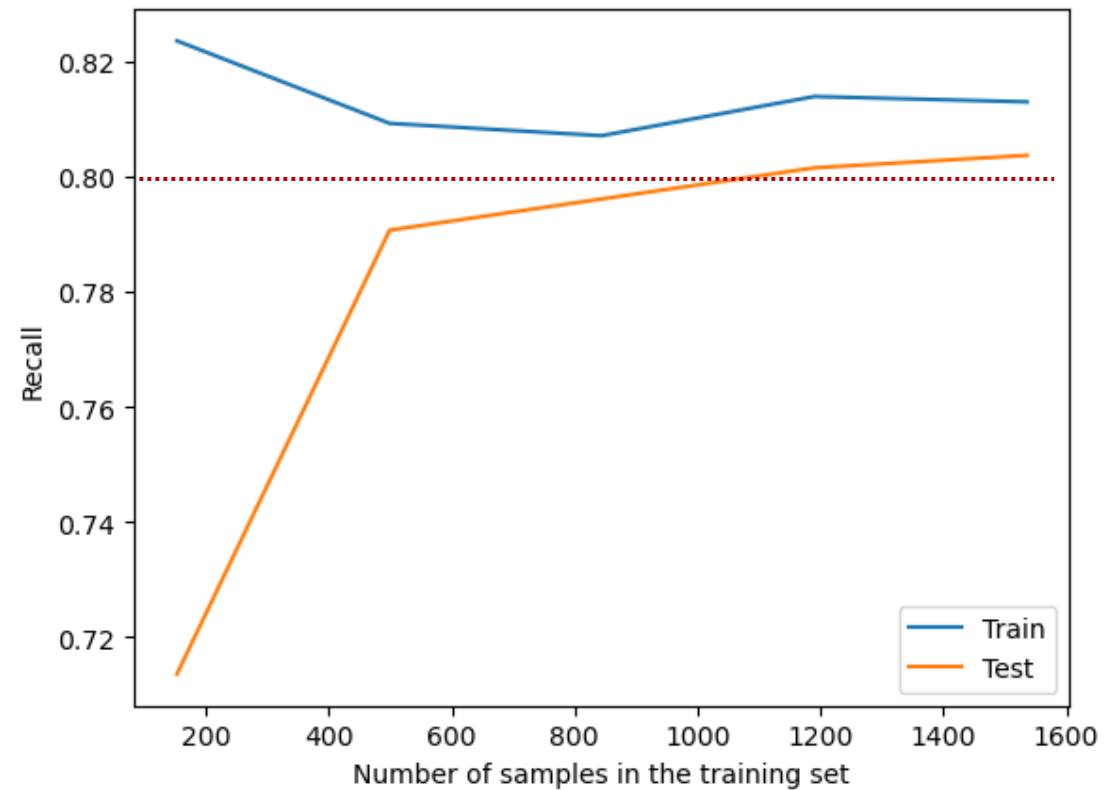
**Modelo
Oversampling**

73%

**Modelo
Undersampling**

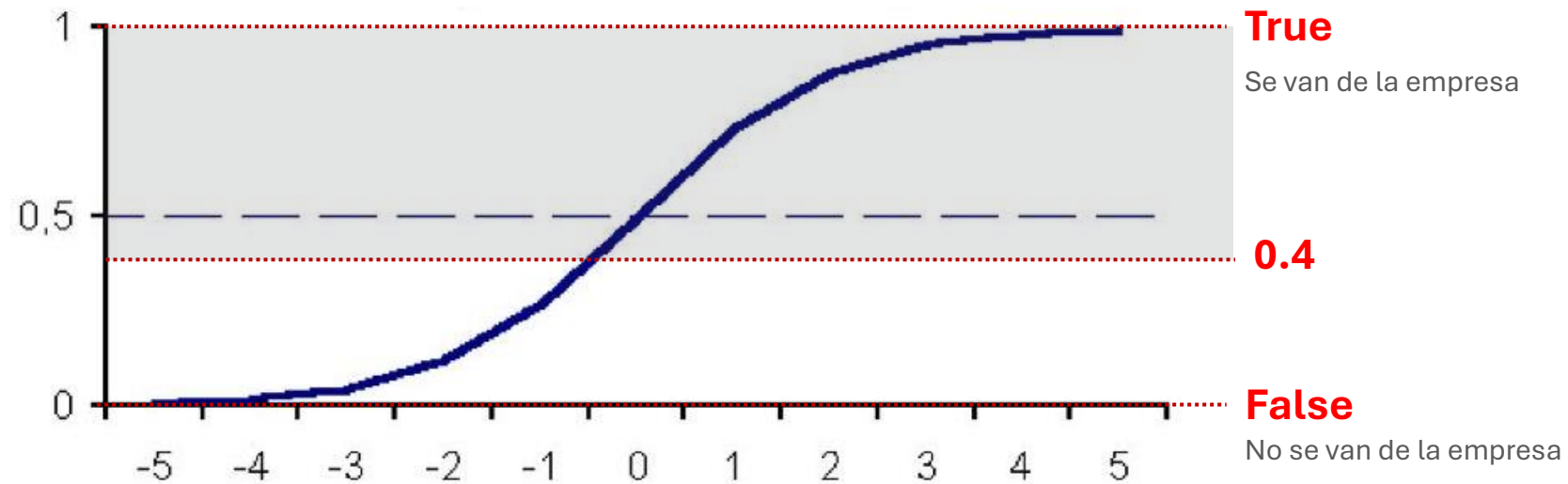
Curva de aprendizaje

Modelo Oversampling



Resultados de modelado

Modelo de Regresión Logística



Modificamos la línea decisoria para tener más True Positives (posibles renunciaciones)

Contenido

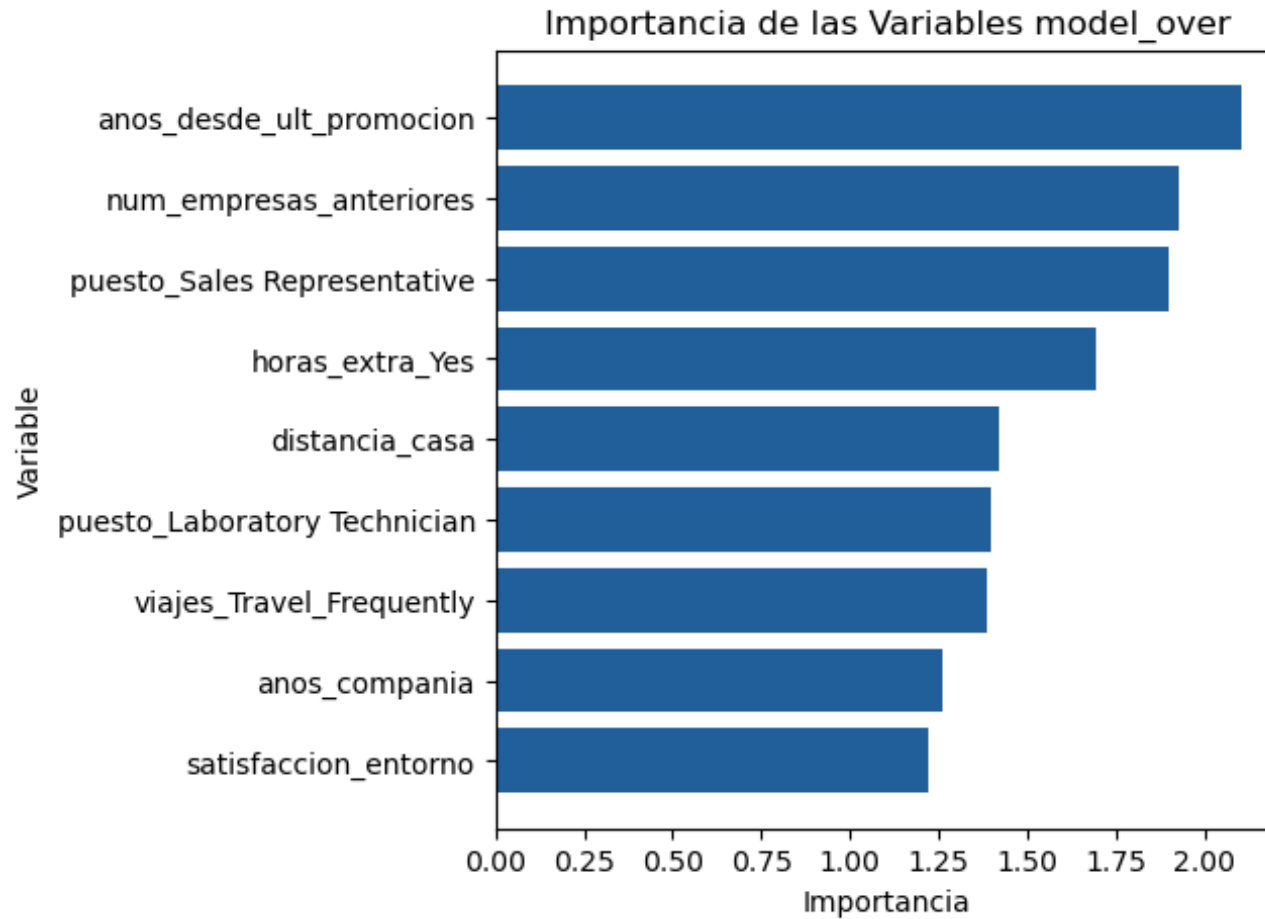
- Introducción
- Variables consideradas
- Insights
- Correlaciones
- Modelado
- Resultados del modelo
- **Importancia de variables**
- Resumen

Importancia de variables (comparativa)

Variable	Importancia_unb	Importancia_over	Importancia_under
anos_desde_ult_promocion	1.873208	2.101926	0.814584
num_empresas_anteriores	1.582318	1.927006	1.346988
puesto_Sales Representative	1.483472	1.896916	1.095766
horas_extra_Yes	1.483770	1.694922	1.450469
distancia_casa	1.224680	1.418275	1.251658
puesto_Laboratory Technician	1.099019	1.395714	1.062324
viajes_Travel_Frequently	1.170474	1.387271	0.735745
anos_compania	0.911709	1.262672	0.168990
satisfaccion_entorno	1.059743	1.222257	0.815274
viajes_Travel_Rarely	0.496423	0.951720	0.384055
estado_civil_Single	0.978580	0.866692	1.339991
educacion_Secundaria	0.526939	0.750705	0.288408
puesto_Human Resources	0.745558	0.736672	0.813303
satisfaccion_companeros	0.519058	0.687115	0.578217
implicacion	0.484011	0.621458	0.344307
satisfaccion_trabajo	0.484011	0.621458	0.344307

No existe diferencia entre las variables más importantes entre los modelos oversampling y unbalanced

Importancia de variables (modelo Oversampling)



* Solo variables con coeficiente mayor a 1

Contenido

- Introducción
- Variables consideradas
- Insights
- Correlaciones
- Modelado
- Resultados del modelo
- Importancia de variables
- **Resumen**

Resumen de resultados

