

Pregunta objetivo del modelo de predicción desarrollado:

¿Un beneficiario de los programas de Sapiencia desertará en el semestre en curso?

Variables más relevantes para el modelo de predicción de deserción.

Se analizó la dispersión entre el **promedio** y la **edad** de los beneficiarios, encontrando que sería adecuado tomar una muestra de la población para un análisis más exhaustivo. Se observó que la mayoría de los desertores se encuentran en un rango de edad de 15 a 23 años, y se identificaron posibles valores atípicos en edades superiores a 26 años. La distribución de los datos y la forma de la función de densidad respaldaron esta interpretación. Se sugirió la eliminación de los valores atípicos para un análisis más preciso.

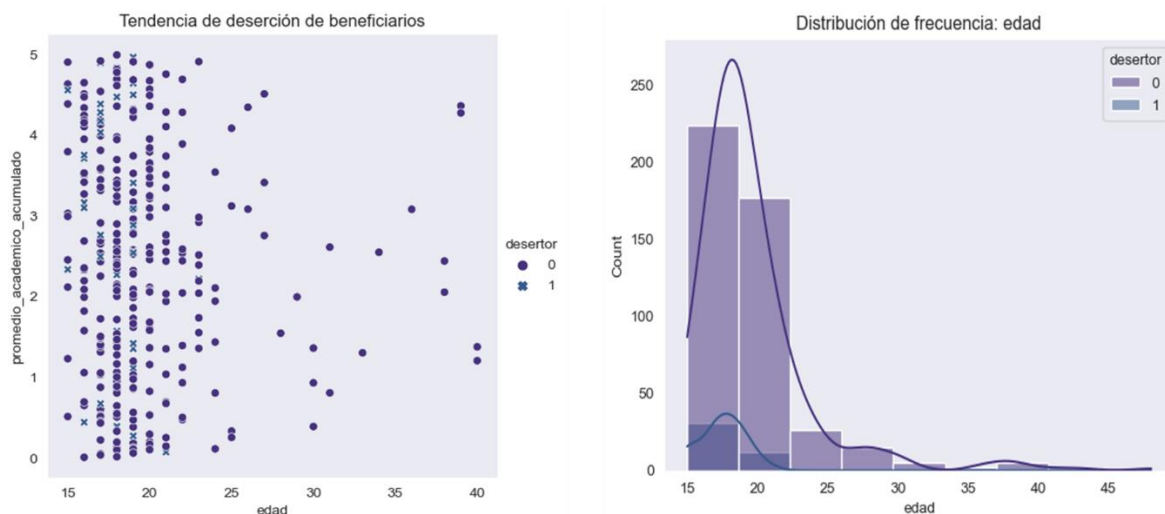


Figura 1. Tendencia de la deserción de los beneficiarios. Figura 2. Distribución de frecuencia de edades con respecto a la deserción.

Por último, es importante tener en cuenta que el conjunto de datos contiene una baja cantidad de beneficiarios que han desertado. Esto puede generar sesgos en el modelo y se debe prestar atención al manejo adecuado de los datos para evitar conclusiones erróneas o sobreestimaciones de la tasa de deserción.

Además, también se observó que, en cuanto a cantidad de desertores con respecto al estado del beneficiario, en el cual se identifican que hay gran cantidad de **NO** desertores con respecto a los desertores. Para ello, lo que se plantea es analizar las diferentes formas de transformación del conjunto de datos, es decir, aquellos valores de desertores "1", se pueden sintetizar más datos para contrarrestar la distribución o acudir a la eliminación de registros en cuanto a los no desertores "0".

Se debe tener en cuenta que estas metodologías parten de un estudio exhaustivo para determinar ciertas analogías e implementar alternativas diferentes según convenga en el estudio.

Por otro lado, algunas de las otras variables en las que se desarrolló el modelo de predicción según EDA fueron:

1. **estrato**: El estrato socioeconómico puede ser un factor importante que influya en la decisión de desertar. Las personas de diferentes estratos pueden enfrentar diferentes desafíos y oportunidades en su educación.
2. **tip_ocupa_vivienda**: El tipo de ocupación de vivienda podría tener relación con la estabilidad y las condiciones económicas del beneficiario, lo cual podría afectar su decisión de desertar.
3. **sexo_persona**: El género de la persona podría tener influencia en la deserción, ya que pueden existir diferencias socioeconómicas, culturales o de acceso a la educación entre hombres y mujeres.
4. **edad**: La edad puede ser un factor relevante, ya que las personas en diferentes etapas de su vida pueden tener diferentes motivaciones y circunstancias que influyan en su decisión de continuar o abandonar sus estudios.
5. **tip_estado_civil**: El estado civil puede estar relacionado con las responsabilidades y compromisos familiares, lo cual podría tener un impacto en la deserción.
6. **ind_discap_ninguna**: La presencia o ausencia de discapacidad puede influir en la forma en que una persona afronta los desafíos académicos y puede ser relevante para comprender la deserción.
7. **tip_seg_social**: El tipo de seguridad social al que pertenece el beneficiario puede indicar su nivel de acceso a servicios y apoyos sociales, lo cual puede ser un factor influyente en la deserción.
8. **niv_educativo**: El nivel educativo alcanzado por el beneficiario antes de ingresar al programa puede tener un impacto en la probabilidad de deserción.
9. **tip_actividad_mes**: El tipo de actividad realizada durante el mes por el beneficiario puede reflejar su situación laboral o de compromisos adicionales, lo cual podría influir en la deserción.
10. **vlr_ingr_salario**: El valor del ingreso salarial del beneficiario puede estar relacionado con su situación económica y tener un impacto en su decisión de continuar o abandonar sus estudios.
11. **promedio_academico_acumulado**: El promedio académico acumulado puede ser un indicador del desempeño académico del beneficiario y puede tener influencia en la decisión de abandonar o continuar.

Características con posibilidad de mejora en el modelo

- Características geográficas o culturales (ya las habían, pero no se tuvieron en cuenta).
- Factores psicosociales, así como variables relacionadas con el bienestar emocional, la motivación y la resiliencia del beneficiario, como la autoestima, la satisfacción con la educación, el apoyo social percibido o la presencia de factores de estrés o dificultades personales.
- Historial educativo, variables relacionadas con el rendimiento en exámenes previos, la asistencia a clases, la participación en actividades extracurriculares o el número de suspensiones o expulsiones anteriores.