

HR Solutions

Análisis y proceso de modelado de datos para determinar las características de los salarios de la empresa

Mayo 2023

Agustina Fernandez Fiant

AGENDA

- **01** | Contexto y Audiencia
- **02** | Hipótesis/Preguntas de Interés
- **03** | Metadata
- **04** | Análisis Exploratorio
- **05** | Machine Learning Model
- **06** | Insights, Next steps y Recomendaciones



CONTEXTO Y AUDIENCIA

Contexto

Hoy en día, uno de los principales problemas que tienen las empresas es el abandono en los puestos de trabajo de las mismas.

La rotación de personal puede tener un impacto negativo en la productividad, el desempeño y la cultura de la empresa.

La pérdida de empleados puede ser costosa para la empresa, ya que puede tener que incurrir en costos de contratación y capacitación para reemplazar a los empleados que se van. Por lo tanto, es importante para las empresas identificar las causas de la rotación y tomar medidas preventivas para retener a sus empleados más valiosos.

Audiencia

Este análisis intenta contestar, con evidencia, las preguntas del párrafo anterior por lo cuál puede ser de utilidad para cualquier persona interesada en conocer si existe alguna causa para la rotación del personal.

Limitaciones

Este análisis está hecho para una empresa en concreto, en EEUU, con cierta cantidad de datos y determinados puestos de trabajo. Lo que significa que, si realizamos este análisis para una empresa de otro rubro o país, el resultado puede ser diferente.

PREGUNTAS DE INTERÉS

Preguntas principales o primarias

- ¿Existen patrones o tendencias en la rotación de personal?
- ¿Hay ciertas áreas o departamentos que tienen una tasa más alta de rotación?
- ¿El género influye en la rotación del personal?
- ¿Y el salario influye?

Preguntas secundarias (nos ayudaran a contestar las principales)

- ¿Cuál es la distribución en los distintos cargos de la empresa entre hombres y mujeres?
- ¿Los hombres tienen salarios más elevados que las mujeres?
- ¿Los hombres tienen cargos más altos que las mujeres?
- A medida que el empleado es más grande, ¿cobra más?

• 03 RESUMEN METADATA

1470



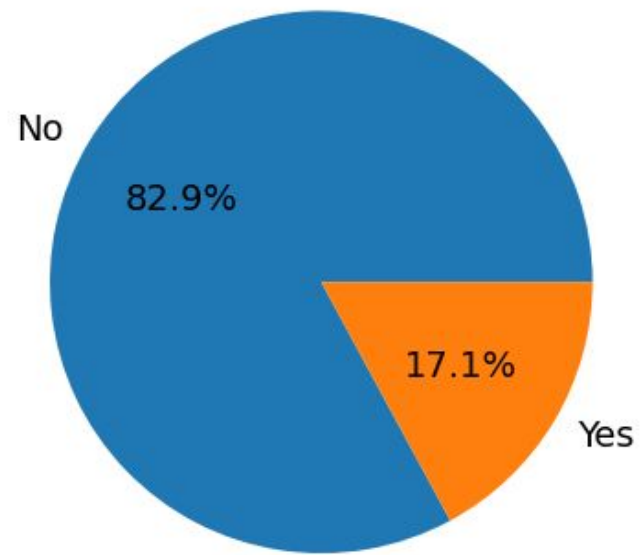
588



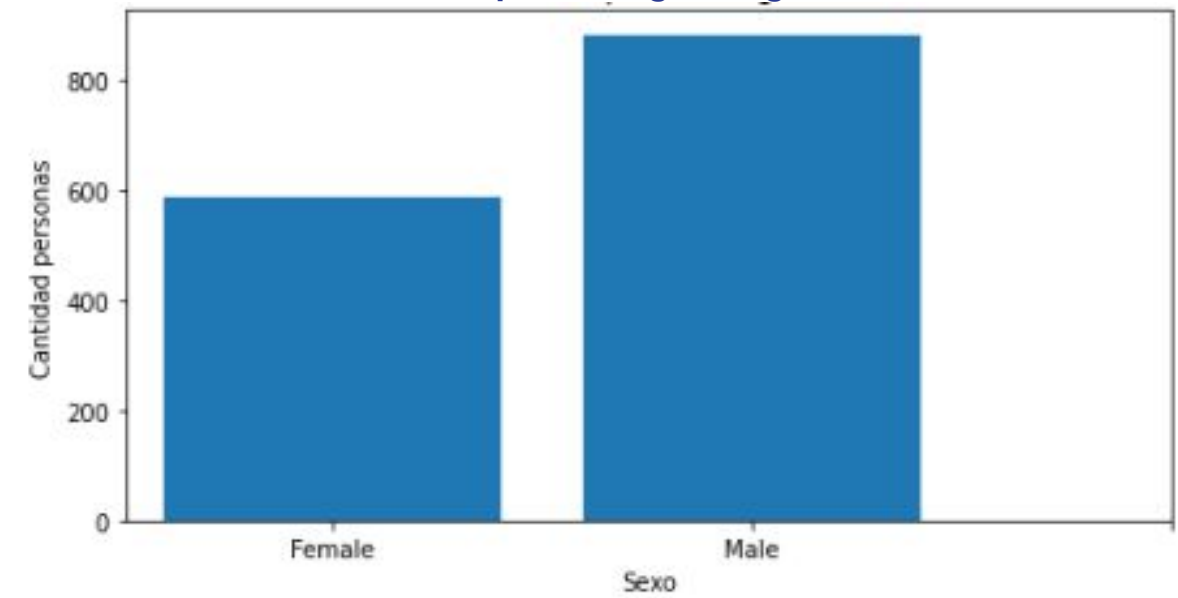
882



% de abandono de los puestos de trabajo



Cantidad de empleados en la empresa según el género



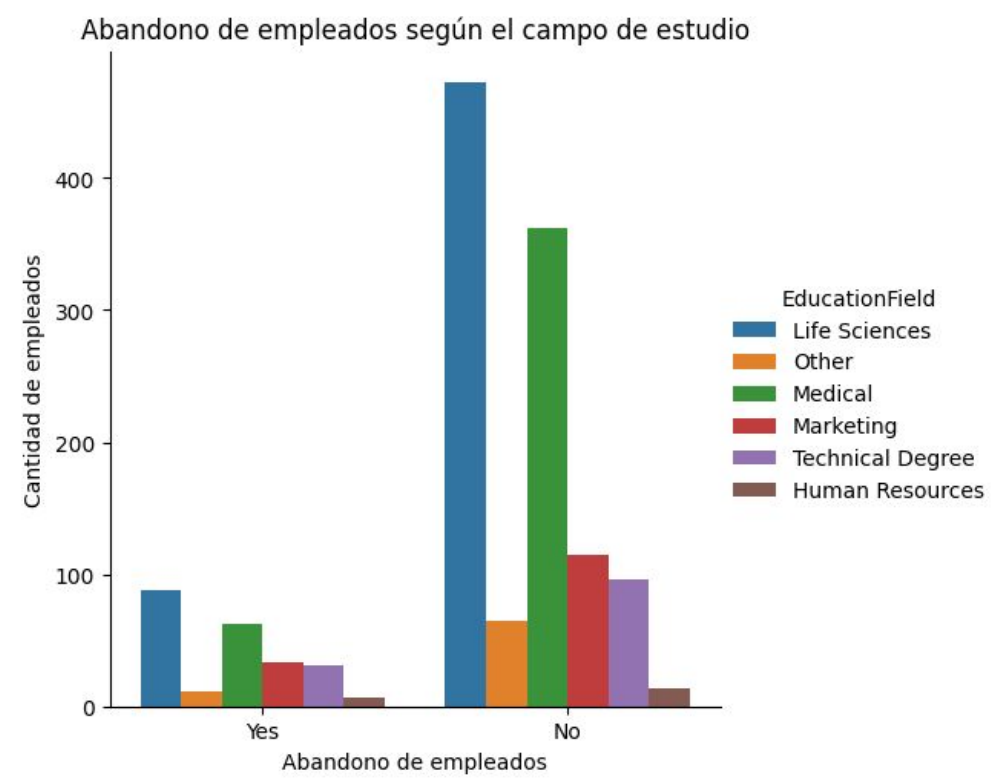
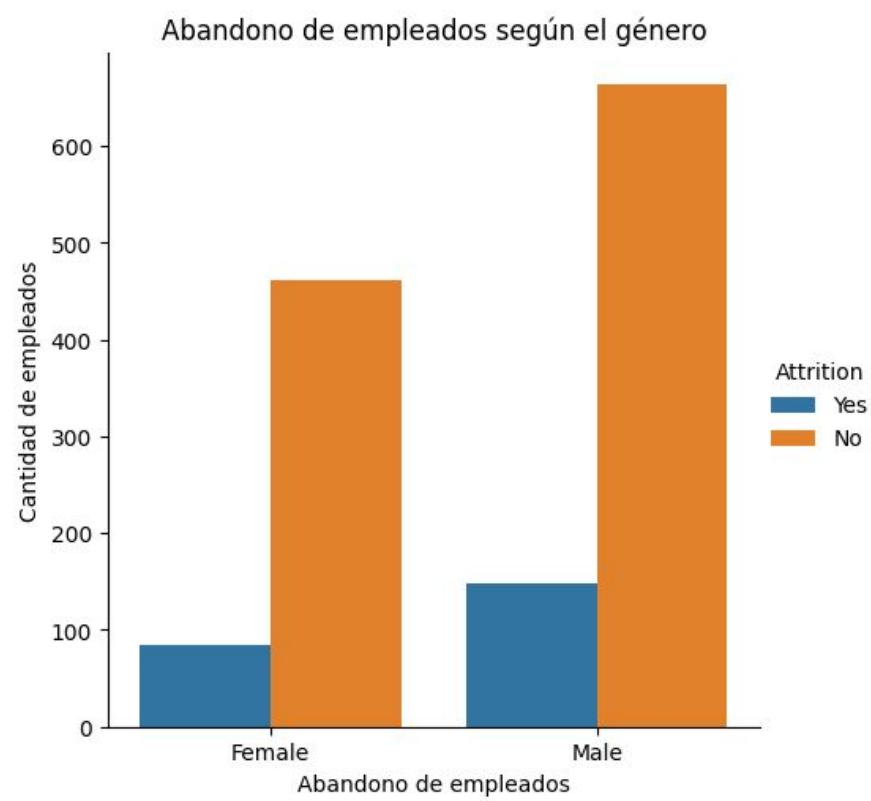
Del total de 1470 empleados, el 83% decide abandonar su puesto de trabajo y el 17% no. Además, el 60% son hombres y el 40% mujeres.

Los datos fueron extraídos de la siguiente base de datos: [DS](#). Accede al siguiente notebook si quieres reproducir el análisis: [link](#).

ANÁLISIS EXPLORATORIO DE DATOS (EDA)

• 04

¿El género influye en la rotación del personal? ¿Hay ciertas áreas o departamentos que tienen una tasa más alta de rotación?



Podemos observar que no hay una relación entre el género y el abandono de los puestos de trabajo y que el departamento tampoco influye.

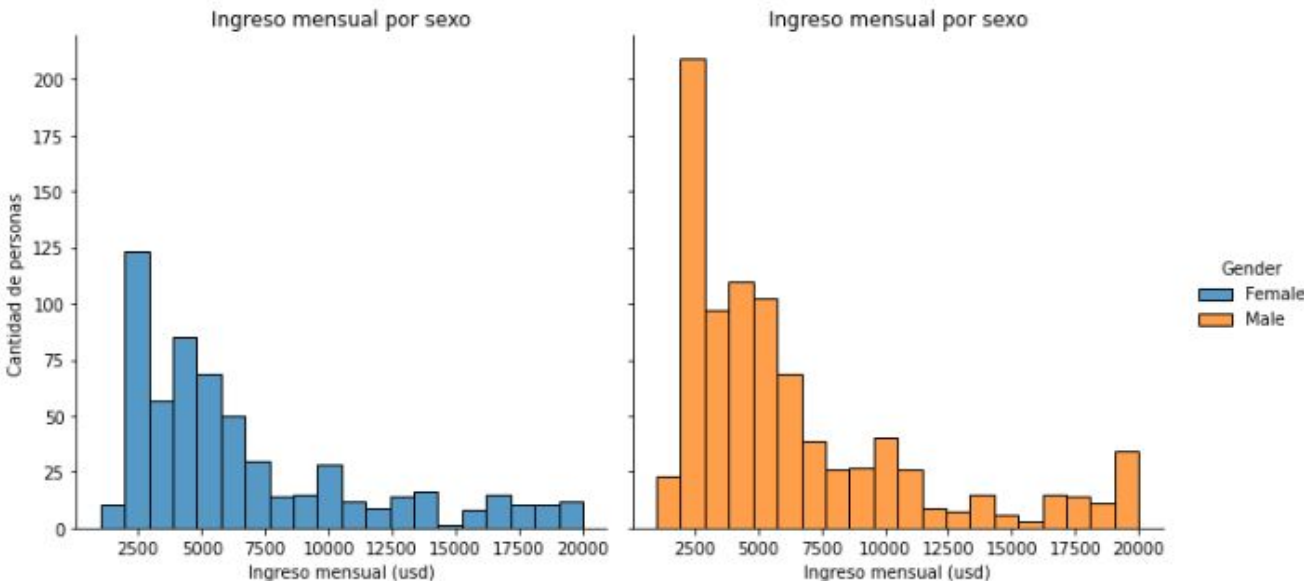
• 04

DISTRIBUCIÓN DEL INGRESO MENSUAL

¿Los hombres tienen salarios más elevados que las mujeres?

Podemos observar como los ingresos mensuales tienen el mismo comportamiento para hombres y para mujeres. Se genera un pico máximo en los 2500usd y después comienza a bajar.

Distribución del salario según el género



Si bien hay más hombres que mujeres, como ya demostramos no hay una correlación entre los salarios de los hombres y mujeres. No se puede observar una brecha salarial entre ambos.

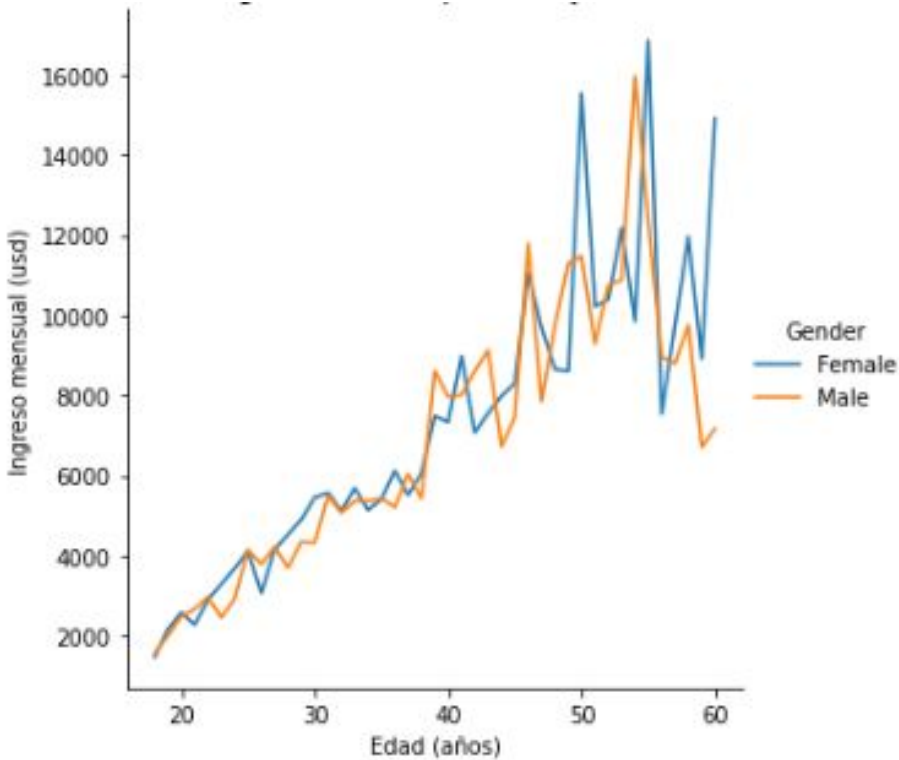
• 04

INGRESO MENSUAL SEGÚN LA EDAD

A medida que el empleado es más grande, ¿cobra más?

En promedio, las mujeres tienen un ingreso mensual más alto que los hombres, por lo que la primera sentencia de la hipótesis no es cierta. A mayor edad, mayor es el cobro mensual. Existe una relación proporcional entre la edad y el ingreso mensual en periodos altos, por lo que podemos asegurar que la segunda parte de la hipótesis sí que es cierta.

Ingreso mensual por edad y sexo



MACHINE LEARNING MODEL

Hemos optado por comparar dos modelos para predecir si el empleado va a abandonar su puesto de trabajo o no.

Baseline model: Regresión logística
Segundo modelo: XGBoost

Los resultados obtenidos son los siguientes:

	Model	Accuracy	Precision	Recall	F1 Score	AUC-ROC score
0	Logistic Regression	0.805310	0.427083	0.788462	0.554054	0.868266
1	XGBoost	0.870206	0.595238	0.480769	0.531915	0.783704

Debido a todos los resultados y conclusiones arrojadas, teniendo en cuenta que el modelo de regresión logística no tiene under/overfitting y que tiene mejor capacidad para distinguir entre clases positivas y negativas, el modelo de regresión logística es el elegido para el problema planteado.



INSIGHTS, NEXT STEPS & RECOMENDACIONES

INSIGHTS, NEXT STEPS & RECOMENDACIONES

Insights

- ❑ Actualmente, el 83% decide continuar en la empresa, mientras que el otro 17% decide renunciar
- ❑ Hay más hombres(60%) que mujeres(40%) en la empresa
- ❑ Los ingresos mensuales, tienen un comportamiento similar para ambos géneros
- ❑ En promedio, las mujeres tienen un ingreso mensual más elevado que los hombres
- ❑ A mayor edad, mayor es el ingreso

Next steps & recomendaciones

- ❑ Realizar un EDA más exhaustivo, analizando si hay más outliers y elegir la forma de tratarlos
- ❑ Ajuste de hiperparámetros: Ver si se pueden mejorar los hp
- ❑ Selección de variables: Realizar un análisis más profundo de las variables seleccionadas
- ❑ Generar nuevas variables: Puede mejorar el rendimiento del modelo al agregar información adicional al conjunto de datos
- ❑ Utilizar modelos más complejos: Se puede probar con modelos más complejos como Redes Neuronales, SVM, Random Forest, etc.
- ❑ Analizar si es necesario balancear el 40% de las mujeres sobre el 60% de los hombres y que la empresa cuente con un 50% de cada género
- ❑ También se podría ver en profundidad si los que tienen ingresos de \$ 2500 la proporción de mujeres y hombres es la misma y sino que es lo que la influencia
- ❑ Analizar si algún departamento tiene más hombres porque no hay muchas mujeres que se postulen para el puesto, por lo que se podría ver como modificar esto
- ❑ Además, se podría examinar si la satisfacción en el trabajo o el equilibrio entre la vida laboral y personal influye en el género



GRACIAS