Filas: 1614

Columnas: 41

1. **Age**: La edad del empleado.

- no hay nulos

- tipo de dato: object (str) 🡪 cambiar a int

1. **Attrition**: Indica si el empleado ha dejado la empresa ("No" significa que no ha dejado la empresa y "Yes" significa que ha dejado la empresa).

- no hay nulos

- tipo de dato: object (str)

3. **BusinessTravel**: Describe la frecuencia de los viajes relacionados con el trabajo del empleado (por ejemplo, "Travel\_Rarely" para raramente).

- hay NaN

- 772 nulos

- tipo de dato: object (str)

4. **DailyRate**: La tarifa diaria del empleado.

- sacar $ para convertirlo a float

- no nulos

- tipo de dato: object (str) 🡪

5. **Department**: El departamento en el que trabaja el empleado (por ejemplo, "Research & Development", "Sales", etc.).

- 1312 nulos

- object

6. **DistanceFromHome**: La distancia desde el hogar del empleado hasta su lugar de trabajo.

- Datos negativos

- no nulos

- Int64

7. **Education**: Nivel de educación del empleado (generalmente en una escala del 1 al 5).

- no nulos

- int64

8. **EducationField**: El campo de educación del empleado.

- Nan

- 745 nulos

- object

9. **EmployeeCount**: Un contador que generalmente es 1 y se usa para contar empleados.

- no nulos

- int64

10. **EmployeeNumber**: Un número de identificación único para el empleado.

- transformar a inbt? Completar vacíos (NaN)

- añadir columna auto incremental (id) y eliminar esta

- 431nulos

- object

11. **EnvironmentSatisfaction**: Nivel de satisfacción del empleado en relación con su entorno de trabajo. Con valores que están comprendidos entre el 1 y el 4, siendo el 4 el nivel de máxima satisfacción.

- no nulos

- int64

12. **Gender**: El género del empleado. Donde 0 corresponde con "hombre" y 1 con "mujer".

- cambiar a string

- no nulos

- object

13. **HourlyRate**: La tarifa por hora del empleado.

- no nulos

- object 🡪 cambia a int?

14. **JobInvolvement**: Nivel de implicación del empleado en su trabajo.

- no nulos

- int64

15. **JobLevel**: Nivel jerárquico del empleado en la empresa.

- no nulos

- int64

16. **JobRole**: El rol o puesto de trabajo del empleado.

- tendremos que hacer un lower(), cada elemento escrito diferente

- no nulos

- object

17. **JobSatisfaction**: Nivel de satisfacción del empleado con su trabajo.

- no nulos

- int64

18. **MaritalStatus**: El estado civil del empleado (por ejemplo, "Single", "Married", etc.).

- NaN

- 651 nulos

- object

19. **MonthlyIncome**: Ingresos mensuales del empleado.

- 843 nulos

- object

20. **MonthlyRate**: Tasa mensual del empleado.

- no nulos

- int64

21. **NumCompaniesWorked**: Número de compañías en las que el empleado ha trabajado.

- no nulos

- int64

- cambiar nombre de columna minúsculas

22. **Over18**: Indica si el empleado es mayor de 18 años.

- (Y)

- Nan

- 901 nulos

- object

23. **OverTime**: Indica si el empleado trabaja horas extra ("Yes" para sí o "No" para no).

- 676 nulos

- object

24. **PercentSalaryHike**: El porcentaje de aumento salarial del empleado.

- no nulos

- int64

25. **PerformanceRating**: Calificación de rendimiento del empleado.

- decidir puntuación todas en int o todas en float

- 195 nulos

- object

26. **RelationshipSatisfaction**: Nivel de satisfacción en las relaciones interpersonales del empleado.

- no nulos

- int64

27. **StandardHours**: Las horas estándar de trabajo.

- 1195 nulos

- object

28. **StockOptionLevel**: Nivel de opciones de compra de acciones del empleado.

- no nulos

- int64

29. **TotalWorkingYears**: Total de años de experiencia laboral del empleado.

- 526 nulos

- object

30. **TrainingTimesLastYear**: Número de veces que el empleado recibió capacitación el año pasado.

- no nulos

- int 64

31. **WorkLifeBalance**: Equilibrio entre trabajo y vida personal del empleado.

- 108 nulos

- object

- nombre columna todo en mayúscula

32. **YearsAtCompany**: Años que el empleado ha trabajado en la empresa actual.

- no nulos

- int64

33. **YearsInCurrentRole**: Años que el empleado ha estado en su puesto actual.

- 1580 nulos

- object

34. **YearsSinceLastPromotion**: Años desde la última promoción del empleado.

- no nulos

- int64

35. **YearsWithCurrManager**: Años que el empleado ha estado bajo la supervisión del actual gerente.

- no nulos

- int

36. **SameAsMonthlyIncome**: Ingresos mensuales del empleado.

- comprobar para errores y eliminar

- 843 nulos

- object

37. **DateBirth**: Año de nacimiento del empleado (teniendo en cuenta que los datos fueron recogidos en el 2023)

- no nulos

- int64

38. **Salary**: Salario de los empleados.

- todos son iguales, calcularlo a partir del mensual, diario, ….

- no nulos

- object

39. **RoleDepartament**: El departamento y el rol del empleado.

- usar el rol para completar errores del JobRole y eliminar (dejar solo departamento)

- 1312 nulos

- object

40. **NumberChildren**: Número de hijos de los empleados

- todo nulos (eliminar columna)

- float64

41. **RemoteWork**: Si el empleado puede teletrabajar o no.

- False, 0 ,1, … decidir tipo de dato

- no nulos

- object

Procedimiento:

- Estudiar dataframe: nulos, tipos de dato x columna, datos atípicos que llamen la atención