**Diccionario de datos**

1. **Age** → Edad del empleado (igual DROP)
   * Convertir a int
   * Sin nulos
2. **Attrition** → Indica si dejó la empresa (Yes/No)
   * Columna completa
   * Mantener
3. **BusinessTravel** → Frecuencia de viajes laborales
   * Categorías: Travel\_Rarely, Travel\_Frequently, Non-Travel
   * Rellenar nulos con "Unknown"
   * Mantener para análisis de Attrition
4. **DailyRate** → Tarifa diaria (estimada)
   * Tipo float
   * Redondear a 2 decimales
   * Sin nulos
5. **Department** → Departamento
   * **Drop**
6. **DistanceFromHome** → Distancia al trabajo (km)
   * Tipo int
   * Revisar valores negativos (error de carga)
   * Sin nulos
7. **Education** → Nivel educativo (1–5)
   * Tipo int
   * Escala: 1 bajo – 5 alto
   * Sin nulos
8. **EducationField** → Campo de estudio
   * Tipo object
   * Rellenar nulos con la moda "Life Sciences"
9. **EmployeeCount** → Constante (1)
   * **Drop**
10. **EmployeeNumber** → ID empleado

* Revisar duplicados
* Puede servir como ID

1. **EnvironmentSatisfaction** → Satisfacción con el ambiente (1–4) \*\*\*\*\*\*\*+

* Tipo int
* Corregir valores erróneos (49 → 4, etc.)
* Sin nulos

1. **Gender** → Género

* Normalizar a M/F o 0/1 (0 = Male, 1 = Female)
* Sin nulos

1. **HourlyRate** → Tarifa por hora (DailyRate/8)

* **Drop** (derivada y con nulos)

1. **JobInvolvement** → Compromiso en el trabajo (1–4)

* Tipo int
* Sin nulos

1. **JobLevel** → Nivel jerárquico (1–5)

* Tipo int
* Sin nulos

1. **JobRole** → Rol específico\*\*

* Tipo object
* Normalizar (.str.lower().str.strip())
* Sin nulos

1. **JobSatisfaction** → Satisfacción en el puesto (1–4)

* Tipo int
* Sin nulos

1. **MaritalStatus** → Estado civil

* Tipo object
* Rellenar nulos si necesario
* Normalizar con .str.title()

1. **MonthlyIncome** → Ingreso mensual

* Tipo float
* 468 nulos
* Borrar??

1. **MonthlyRate** → Tarifa mensual

* Revisar relevancia (mantener de momento)

1. **NumCompaniesWorked** → Número de empresas anteriores

* Tipo int
* Renombrar num\_empresas\_anteriores

1. **Over18** → Mayor de edad

* **Drop**

1. **OverTime** → Horas extra (Yes/No)

* Convertir a bool (True/False)
* Renombrar horas\_extra

1. **PercentSalaryHike** → Incremento salarial (%)

* Tipo int
* Renombrar incremento\_salario

1. **PerformanceRating** → nivel de trabajo

* Tipo int
* Renombrar nivel\_trabajo

1. **RelationshipSatisfaction** → Satisfacción relaciones interpersonales (1–4)

* Tipo int
* Renombrar satisfaccion\_relaciones\_interpersonales
* No hay nulos

1. **StandardHours** → Tipo de jornada (Full/Part Time)

* Tipo object
* Rellenar nulos con "Full Time"
* Renombrar calificacion\_jornada

1. **StockOptionLevel** → Opciones sobre acciones

* **Drop**

1. **TotalWorkingYears** → Años de experiencia total

* Tipo int
* Renombrar años\_experiencia

1. **TrainingTimesLastYear** → Sesiones de entrenamiento último año

* Tipo int
* Renombrar sesiones\_entrenamiento\_ultimo\_año

1. **WorkLifeBalance** → Balance vida/trabajo (1–4)

* Tipo int
* Renombrar balance\_vida\_trabajo

1. **YearsAtCompany** → Años en la empresa

* Tipo int
* Renombrar años\_compañia

1. **YearsInCurrentRole** → Años en el rol actual

* Tipo int
* Renombrar años\_mismo\_puesto
* Muchos nulos (revisar si mantener o dropear)

1. **YearsSinceLastPromotion** → Años desde última promoción

* Tipo int
* Mantener

1. **YearsWithCurrManager** → Años con mismo manager

* Tipo int
* Mantener

1. **DateBirth** → Año nacimiento

* **Drop** (ya está Age)

1. **Salary** → Salario anual

* Tipo float
* Limpiar símbolo $ con .str.replace()

1. **RoleDepartament** → Combinación rol/departamento

* Muchos nulos → revisar
* O bien normalizar (.str.lower()) o **Drop**

1. **NumberChildren** → Número de hijos

* **Drop** (sin datos)

1. **RemoteWork** → Teletrabajo (Yes/No)

* Convertir a bool
* Renombrar remotework