

- 1) Load the dataset and display its shape, columns, and data types.
- 2) Calculate the percentage of missing values for each column.
- 3) Identify columns containing mixed data types.
- 4) Clean `user_id` by replacing empty strings with `NaN` and converting values to uppercase.
- 5) Convert `age` to numeric and replace invalid values (negative or `>100`) with `NaN`.
- 6) Clean `salary` by removing currency symbols and converting it to numeric.
- 7) Convert `join_date` into datetime format handling multiple date formats.
- 8) Extract `joining_year` from the cleaned `join_date`.
- 9) Standardize `department` values by trimming spaces and converting to uppercase.
- 10) Normalize `city` names to a consistent case and spelling.
- 11) Convert `performance_score` into numeric by mapping text values to numbers.
- 12) Clean `experience_years` by converting text values to numeric and removing negatives.
- 13) Convert `is_active` values (yes/no/True/False) into boolean format.
- 14) Identify and remove duplicate records based on `user_id`.
- 15) Create a clean final DataFrame containing only valid, standardized records.