**Alteryx Case Study 3**

**Dataset:** Sample\_Data with multiple sheets as shown below:

* Department
* EMP2
* Employee
* Joining\_date
* Payslip
* Rank

****

**Cleansing steps to follow:**

* **Populate nulls with 0 if numeric type. Else populate nulls for empty rows**
* **Combine Employee and EMP2 sheet to get all employee list**
* **Join Tool:**
  + Identify employees without department assignment (Left/Right join outputs).
  + Hint (Join the **Employee** sheet with **Department** using the Department\_ID or Department\_Name (if available).)
* **Join Multiple Tool:**
  + Combine **Employee**, **Joining\_date**, and **Payslip** sheets and output a consolidated table with employee name, join date, and latest salary/payslip info.
* **Union Tool:**
  + Combine the **Employee** and **EMP2** sheets (assuming they have similar structures). Use both **Auto Config by Name** and **Manual Configuration** to understand the difference. The resulted table should have headers from EMP2 sheet.
  + Also try **Auto Config by Position**
* **Find Replace Tool:**
  + Use the **Department** sheet to replace department codes/names in the **Employee** sheet using Department\_ID.
* **Date Time Tool:**
  + In the **Joining\_date** sheet, convert joining dates to a standard format like yyyy-MM-dd, then extract:
    - Year
    - Month
    - Day of the week
* **Text to Columns Tool:**
  + Split employee email and create a user name column by splitting it by domain name.
* **RegEx Tool:**
  + Extract domain names from email addresses (if present) in **Employee** or **EMP2** sheets.
  + Identify employee IDs with specific patterns (e.g., starting with "E" followed by 4 digits).
* **Cross Tab Tool:**
  + Use **Payslip** data to show total pay per employee across different months as columns.
* **Transpose Tool:**
  + Take the cross-tab output and transpose it back to long format for further summarization.
* **Summarize Tool:**
  + From **Payslip**, calculate:
    - Total salary paid per department
    - Average salary per rank
    - Maximum and minimum salary across all employees
* **Formula Tool**
  + **Create Full Name:**
    - In the **Employee** sheet, combine First\_Name and Last\_Name into a new field Full\_Name.
* **Calculate Tenure:**
  + Use the **Joining\_date** sheet to calculate the number of years an employee has been with the company based on today’s date.
* **Net Salary Calculation:**
  + In the **Payslip** sheet, create a field Net\_Salary = Gross\_Salary - Tax - Deductions.
* **Filter Tool**
  + **Employees with Salary > ₹50,000:**
    - Filter employees from **Payslip** sheet whose gross salary is greater than 50,000.
  + **Employees joined in a particular year:**
    - Filter employees who joined in the year 2011 using the **Joining\_date** sheet.
* **Unique Tool**
  + **Unique Departments:**
    - From **Employee table**, find unique department names or codes.
  + **Unique Job Titles or Ranks:**
    - In the **Rank** sheet, extract a list of unique employee ranks
* **Record ID Tool**
  + **Assign Row Numbers:**
    - Add a Record\_ID to the **Employee** sheet to simulate serial numbers.
* **Sample Tool**
  + **First N Records:**
    - Get the first 10 employees from the **Employee** sheet.
  + **Random Sample:**
    - Randomly sample 5 records from the **EMP2** sheet for quality checks.
* **Select Records Tool**
  + **Pick Records 11 to 20:**
    - Select records between the 11th and 20th rows from the **Payslip** sheet.
  + **Last 3 Employees:**
    - Pick the last 3 employees based on record order in the **Employee** sheet.
* **Sort Tool**
  + **Sort by Salary Descending:**
    - Sort the **Payslip** sheet by Gross\_Salary from highest to lowest.
  + **Sort by Joining Date:**
    - Sort **Joining\_date** data to find the earliest and latest joiners.