

**Case Study :** People First Inc. is a consulting multinational company which specialization in Human Resource Analytics and their department Human Capital Advisory (HCA) is faced with the following problem:

They have a client who is facing quite high rates of attrition in their company. They have approached People First Inc. to find ways to resolve this issue. The client has provided the HCA department with the employee data for the last 10 years. You are an Analytics Consultant in the HCA department of People First Inc. and have been provided with a random sample of the clients' data to understand the data. Dataset Attached.

Your tasks are as under:

(Ensure that you are interpreting the operations at all steps as you shall be presenting this report to your Team Leader)

1. Perform the basic Exploratory Data Analysis on the sample data.
  - a. Check for missing values in "Hourly Rate" and "Performance Rating" and impute them. Explain your logic behind the imputation.
  - b. What percentage of employees left the company?
2. The client claims that employees who spend 15 years or more with the same manager do not leave the company. Can you validate the claim?
3. Create a pivot table of the variables "Marital Status" and "Attrition". Show your data as percentages of the count. What insight(s) do you gather?
4. The client has reported that there was more attrition among Females as compared to Males in Sales Representative and Human Resources Job roles. Cross-check this claim and state your observations.
5. Answer the following questions:
  - a. Generate the measure of central tendency, dispersion and symmetry for Variable "Monthly Income". Share your insights.
  - b. Generate the measure of central tendency, dispersion and symmetry for Variable "Age". Also, create a histogram for this variable. Share your insights.
  - c. What is the most common salary hike?
  - d. Which is the more appropriate metric to represent the central tendency for "Monthly Income" variable, mean or median? Explain why.
  - e. The client claims of having a diverse workforce in terms of the total experience of employees for both male and female. Analyse the range and standard deviation of the Total Working Years variable and state your observation. (Hint: Use filter for Gender Variable)
  - f. Client claims that "Marital Status" has no impact on "Job Satisfaction". Analyse this and share your insights. (Hint- Use Mean function)