



M10\_T1\_E2

# Final Project – Problem Statement

## HR Data Analytics at Acme TelCo

### Situation

Acme Telco is a small telecom company that develops, sells, and installs telecommunication devices such as network routers, conference call hardware, etc. for business needs. Acme has been operating for over a decade with a modest current headcount of close to 200 people and it has organized itself around the following 5 departments: Operations, IT, Sales, Software engineering, and Admin.

Acme has been having a stressed financial situation for quite some time resulting in general unhappiness among its employees.

Acme has a long-standing tradition of conducting employee satisfaction surveys but it has not paid attention to the detailed results coming out of the survey. Also, same with the attrition situation.

In order to boost employee confidence to turn around the company, the President & CEO, Mrs. Jan Ramsey asked that the HR department conducts a detailed study of the employee situation, satisfaction, attrition reasons, etc.

Currently, there is no clarity on information such as gender equality e.g. whether female employees are paid the same as male employees. She realizes that it is a good time to explore and find answers to such questions and take actions to improve the company image among its employees and potential future hires.

The Chief Information Officer, Mrs. Jen Heinz has asked her Business Intelligence Director to pull data to conduct a detailed study. After spending a considerable amount of time, the BI team unearthed three different data sets containing (1) Basic Employee Data, (2) Employee performance and attendance data, and (3) Employee salary data

The employee ID has been randomized and encoded, and the employee names have been masked in order to protect employee's personal information. The data is presented in three different sets in order to protect confidential information such as salary or performance scores from being circulated to a wider audience. Employee ID is a unique number that is present across all three datasets to allow joining the tables.

You have been hired by the President & CEO, Mrs. Jan Ramsey to conduct a detailed study and present your findings to her and the board of directors. Based on your recommendations, the company may take several actions that will help improve the company image among the public and the employees, increase employee satisfaction and control attrition, etc.

## Expectations

Mrs. Ramsey has thought about the implications of the study carefully and in order to be specific about the findings and save time she wrote down a few areas that needed to be explored carefully.

As a Data Scientist, you are expected to conduct analysis on the lines of these managerial questions and find answers.

Here are the questions:

- 1) Group size and demographics
  - a. How many employees are currently employed by each department?
  - b. What are the demographics of our current employees?
    - i. Age
    - ii. Gender
    - iii. Marital Status
- 2) Salary structure
  - a. What was the current total salary expense for each department?
  - b. What is the salary structure for each demography of our current employees?
    - i. Age
    - ii. Gender
    - iii. Marital Status
- 3) Performance results
  - a. What was the distribution of employees in terms of their performance?
  - b. Could we do a deep dive per group?
    - i. Department
    - ii. Age
    - iii. Gender
    - iv. Marital Status
- 4) Satisfaction scores
  - a. How satisfied our employees are?
  - b. Could we do a deep dive per group?
    - i. Department
    - ii. Position
    - iii. Age
    - iv. Gender
    - v. Marital Status
- 5) Attrition
  - a. How many employees have left the company in total?
  - b. What were the main reasons for them to leave?
  - c. How many of those reasons are voluntary and non-voluntary?
  - d. Is there some pattern related to employee groups?
    - i. Department
    - ii. Age
    - iii. Gender
    - iv. Marital Status

## Data Set

As mentioned earlier the dataset contains three tables:

1. General HR information gathered from the HR information system I
  - a. Employee ID: Unique employee record number used as a key across different databases
  - b. Position: Role the employee held
  - c. DOB: Date of birth of the employee
  - d. Gender: Employee gender (Male / Female)
  - e. Marital Status of the employee: Single / Married / Divorced / Widowed
  - f. Date of hire: Date when the employee was hired by the company
  - g. Date of termination: If the employee has been terminated, the date on which they were terminated or the last official working day
  - h. Termination reason: The reason why the employee left the organization
  - i. Employment status: Whether the employee is currently with the company, terminated for a reason, or they left themselves
2. Performance and Employee satisfaction survey results gathered from the HR information system II
  - a. Employee ID: Unique employee record number used as a key across different databases
  - b. Dept: which department the employee worked in – Admin / Executive Office / IT / Operations / Sales / Software Engineering
  - c. Manager name: Employee's manager name
  - d. Perf Score: Most recent employee performance rating – Exceeded expectations / Met Expectations / Improvement needed. Employees with 'Improvement needed' are generally moved towards a performance improvement program in which they are monitored for about 3 – 6 months after which a decision is made whether to continue to employ them let them go depending on the performance during the improvement program.
  - e. Emp. Sat: Employee satisfaction score – rating given by the employee on how much they are satisfied with their employment with the company. The survey is conducted once a year in a confidential way. A rating of 1 means the employee is highly dissatisfied and a rating of 5 means that the employee is highly satisfied. However, it is noticed that the employees very rarely gave a score of 1 or 2
  - f. Date of the last perf. Review – Date on which the most recent performance review was conducted
  - g. Late days: how many days the employee came late to work in the last one month time
  - h. Absent days: how many leave days the employee availed in the last one year
3. Employee salary details gathered from the financial information system
  - a. Employee ID: Unique employee record number used as a key across different databases
  - b. Salary: Annual salary of the employee in USD