BUSINESS INTRODUCTION

Emerald Technologies is a multinational mobile telecommunications company, operating in many African and Asian countries with Its head office is in Lagos.

As of December 2022, Emerald Technologies recorded over 1 Billion subscribers, making it the one of the largest mobile network operators in Africa/Asia.

They currently have in-house issues regarding employees promotion and performance rating and they need an experienced Data Scientist to help

PROBLEM STATEMENT

Emerald Technologies is facing challenges related to employee attrition, performance management, and effective utilization of training programs.

There is a pressing need to better understand the factors contributing to employee turnover, to ensure that our promotion policies are fair and effective, and to optimize training programs to enhance employee performance and satisfaction

![10Alytics logo]
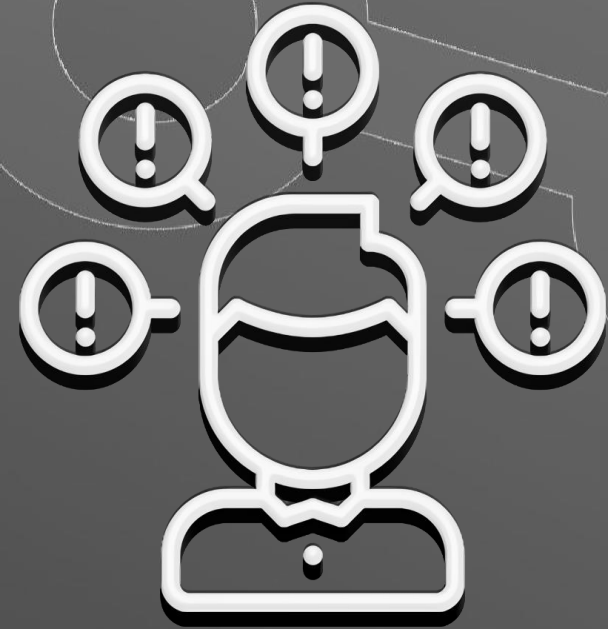
OBJECTIVES OF THE PROJECT

1. **Employee Demographics:** Analyzing the demographic profiles of employees to understand diversity and inclusion within the organization
2. **Attrition Analysis:** Understanding the demographic and job-related factors that contribute to attrition.
3. **Employee Performance:** Evaluating employee performance and ensuring that employees with lower performance ratings are provided with targeted training and development opportunities to improve their skills and overall performance

WHAT IS "ATTRITION"?

Attrition refers to the gradual reduction in the number of
employees within an organization due to various reasons such as **resignations, retirements, layoffs, or voluntary departures.**

It is an important metric for businesses because it can indicate employee satisfaction, organizational health, and
the effectiveness of management practices.

The Data Records The Attrition Value for All **Employees Still Present(Attrition = No)** Vs **Employees That Have Left(Attrition = Yes)**

DATA DESCRIPTION

Emerald Technologies is Providing you with a remote copy(backup) of their database with the following
Tables And Columns:

TABLE 1: Employee_Perf
This table holds information relevant to employee performance and compensation alongside attrition data

- **employee_id**: A unique identifier for each employee in the dataset.
- **HourlyRate**: This represents how much an employee earns per hour worked.
- **JobLevel**: The level or rank of the employee's job within the organization. Higher numbers usually indicate higher levels or positions.
- **Attrition**: Indicates whether an employee has left the company. this is represented as "Yes" for employees who have left and "No" for those who are still with the company

DATA DESCRIPTION

TABLE 1: employee_Perf

- **JobSatisfaction:** A measure of how satisfied an employee is with their job. This could be represented on a scale (e.g., 1 to 4 or 1 to 5) where higher numbers indicate greater satisfaction.
- **MonthlyIncome:** The total income an employee earns in a month.
- **MonthlyRate:** The fixed rate of payment for the employee's job per month. This could be different from **MonthlyIncome** if it includes **additional payments or bonuses**.
- **PerformanceRating:** The rating of the employee's performance, which typically is measured on a scale
- (e.g., 1 to 4 or 1 to 5) where higher numbers indicate better performance.

DATA DESCRIPTION

TABLE 2: employee_test
This Table holds The demographic and training Information on each employee

- **employee_id:** A unique identifier assigned to each employee in the dataset.
- **department:** The specific department within the organization where the employee works (e.g., HR, Sales, IT).
- **region:** The geographical region where the employee is located or assigned.
- **education:** The highest level of education attained by the employee (e.g., High School, Bachelor's, Master's).
- **gender:** The gender of the employee (e.g., Male, Female).

10Alytics

A-Z

DATA DESCRIPTION

TABLE 2: employee_test

- **erecruitment_channel:** The channel through which the employee was recruited (e.g., Direct, Referral, Agency).
- **no_of_trainings:** The number of training programs the employee has attended.
- **age:** The age of the employee.
- **previous_year_rating:** The performance rating of the employee from the previous year.
- **length_of_service:** The number of years the employee has been with the company.
- **awards_won:** Indicates whether the employee has won any awards (e.g., Yes/No or number of awards).
- **avg_training_score:** The average score the employee has achieved in training programs

DATA OVERVIEW & ANALYSIS

The Previous year Rating Column Has Some Missing Values In it

Can You:

1. Deduce the Cause of the missing Values
2. Discuss a good default value for when performing analysis
3. Review your understanding of the coalesce function as a means to help with missing values

Before we move on to Analysis: **Brainstorming session**

Come up with at-least 3 Questions that would be in line with the business problem

CONCEPTS OVERVIEW

CASE
Executes conditional logic by evaluating conditions and returning a value

JOINS
Combine rows from two or more tables based on a related column between them.

COALESCE
Returns the first non-null value from a list of arguments.

CAST
Converts a value from one data type to another.

10Alytics

PROJECT WORKFLOW
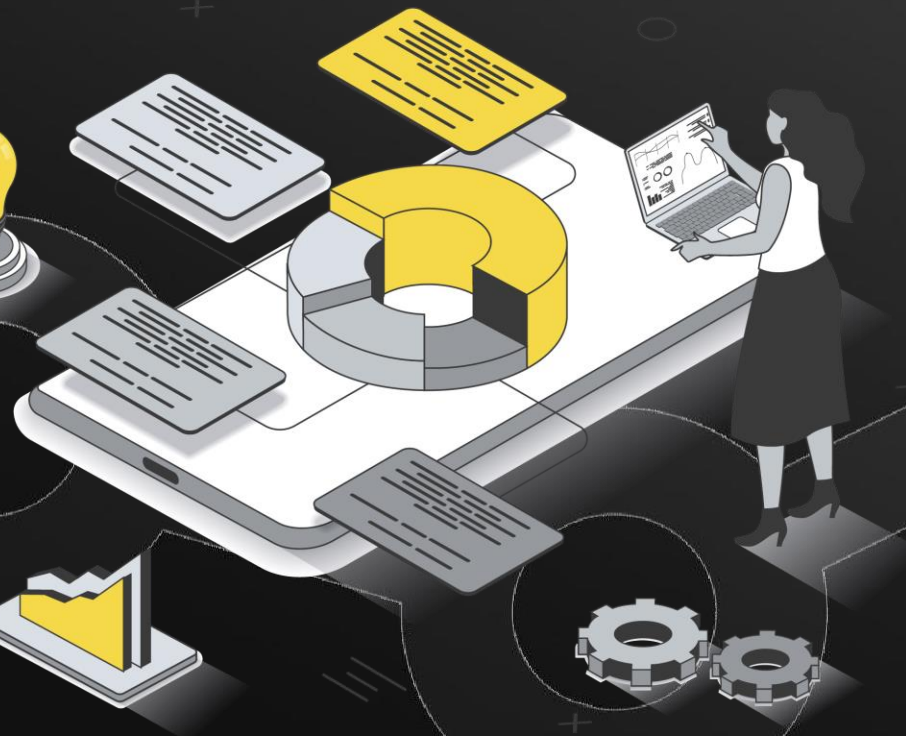
The project workflow follows these steps:

1. **Business Overview:** Understanding the business, and their expectations
2. **Data Overview:** Looking into the data and exploring it's relevance to the business
3. **Data Extraction**: Extract product, order, and customer data from the provided SQL file.
4. **SQL Queries**: Write SQL queries to extract key metrics and insights from the data.
5. **Recommendations**: Provide business recommendations based on the analysis.

TAILORED ANALYSIS

DESCRIPTIVE ANALYTICS: Employee Demography

Using the data answer the following questions:

- How many employees do we have in the organization and what is the maximum length of service?
- Which department has the most employees, and which department has the fewest employees?
- What is the proportion of male to female employees?
- Group Employee age into 5 categories (20 – 29, 30 – 39, 40-49, 50-59, >60).
- What age group has the highest and lowest employee?

10Alytics

DIAGNOSTIC ANALYTICS: Employee Performance

Using the data answer the following questions:

- Who has the highest average training score among all employees?
- What is the average training score of employees in each department
- What is the average previous year rating by department?
- What is the average training score of employees by education type?
- What is the average previous year rating by recruitment channel?
- Based on the age group created what is the average previous year rating and average training score.

TAILORED ANALYSIS

DIAGNOSTIC ANALYTICS: Attrition Analysis

Using the data answer the following questions:

- What is the Average Attrition Rate
- Which regions have the highest rate of departures (employees who have left), and what are the corresponding departments?
- Which regions have the highest average job satisfaction?
- Which departments have the highest average job satisfaction and what is the the rating?
- Which departments have the highest rate of departures?
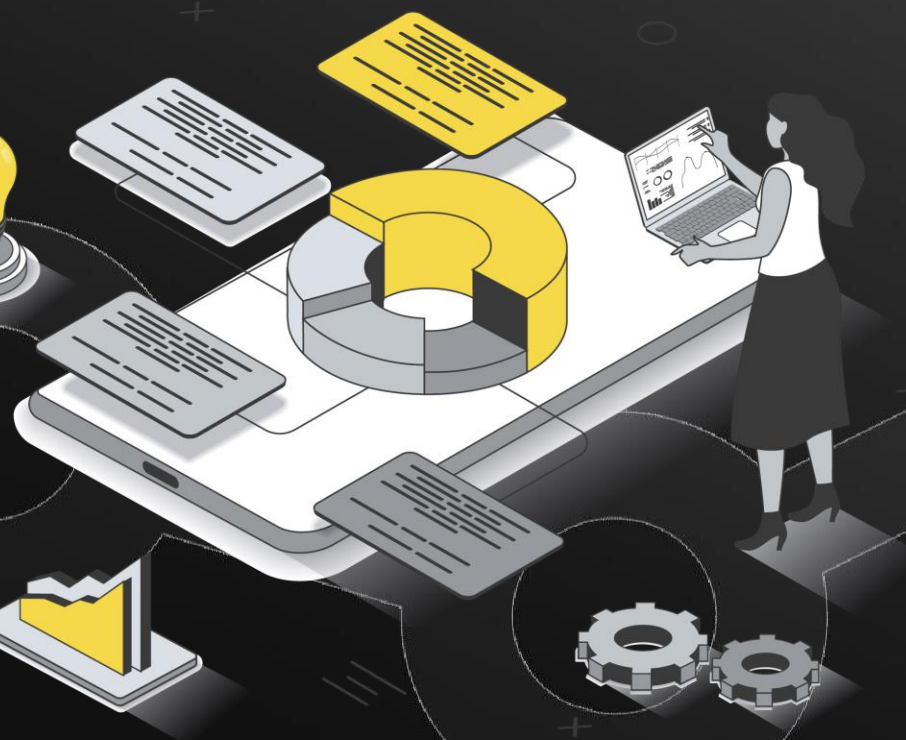
10Alytics

Remember to share your learning on LinkedIn

(Tag @10Alytics)