**Employee Attrition**

**Aim:**

To understand the concepts of business intelligence insights, by analysing the employee resignation behaviour.

Data cleaning and EDA through python and identifying the KBRs and KPIs for this specific business problem.

**Business Problem:**

The key to success in any organization is attracting and retaining top talent. As an HR analyst one of the key task is to determine which factors keep employees at the company and which prompt others to leave. Given in the data is a set of data points on the employees who are either currently working within the company or have resigned. The objective is to identify and improve these factors to prevent loss of good people.

Our goal is to provide a complete analysis about all the factors that has impact on employee’s resignation patterns, the current insights of the organisation and its employee demographics.

**Learning Outcomes:**

1. Understand the KBRs and KPIs of an Attrition, HR report.
2. Cleaning data and feature selection.
3. Learn EDA through Python, to find all the relevant metrics that would help us to track our KPIs.
4. Dynamic Charts creation and Dashboarding through Tableau using various metrics to provide a clear analysis of the company’s attrition factors.

**Initial Skill Requirement:**

Proficiency in Python libraries: pandas, numpy, matplotlib and seaborn.

Along with that Tableau dynamic charts creation and Dashboarding skills.

**Data Dictionary:**

|  |  |
| --- | --- |
| **Variables** | **Description** |
| AGE | Numerical Value |
| ATTRITION | Employee leaving the company (0=no, 1=yes) |
| BUSINESS TRAVEL | (1=No Travel, 2=Travel Frequently, 3=Travel Rarely) |
| DAILY RATE | Numerical Value - Salary Level |
| DEPARTMENT | (1=HR, 2=R&D, 3=Sales) |
| DISTANCE FROM HOME | Numerical Value - THE DISTANCE FROM WORK TO HOME |
| EDUCATION | Numerical Value. (1 'Below College' 2 'College' 3 'Bachelor' 4 'Master' 5 'Doctor') |
| EDUCATION FIELD | (1=HR, 2=LIFE SCIENCES, 3=MARKETING, 4=MEDICAL SCIENCES, 5=OTHERS, 6= TECHNICAL) |
| EMPLOYEE COUNT | Numerical Value |
| EMPLOYEE NUMBER | Numerical Value - EMPLOYEE ID |
| ENVIRONMENT SATISFACTION | Numerical Value - SATISFACTION WITH THE ENVIRONMENT (1 'Low' 2 'Medium' 3 'High' 4 'Very High') |
| GENDER | (1=FEMALE, 2=MALE) |
| HOURLY RATE | Numerical Value - HOURLY SALARY |
| JOB INVOLVEMENT | Numerical Value - JOB INVOLVEMENT (1 'Low' 2 'Medium' 3 'High' 4 'Very High') |
| JOB LEVEL | Numerical Value - LEVEL OF JOB |
| JOB ROLE | (1=HR REP, 2=HR, 3=LAB TECHNICIAN, 4=MANAGER, 5= MANAGING DIRECTOR, 6= RESEARCH DIRECTOR, 7= RESEARCH SCIENTIST, 8=SALES EXECUTIVE, 9= SALES REPRESENTATIVE) |
| JOB SATISFACTION | Numerical Value - SATISFACTION WITH THE JOB (1 'Low' 2 'Medium' 3 'High' 4 'Very High') |
| MARITAL STATUS | (1=DIVORCED, 2=MARRIED, 3=SINGLE) |
| MONTHLY INCOME | Numerical Value - MONTHLY SALARY |
| MONTHLY RATE | Numerical Value - MONTHLY RATE |
| NUMCOMPANIES WORKED | Numerical Value - NO. OF COMPANIES WORKED AT |
| OVER 18 | (1=YES, 2=NO) |
| OVERTIME | (1=NO, 2=YES) |
| PERCENT SALARY HIKE | Numerical Value - PERCENTAGE INCREASE IN SALARY |
| PERFORMANCE RATING | Numerical Value - PERFORMANCE RATING |
| RELATIONS SATISFACTION | Numerical Value - RELATIONS SATISFACTION |
| STANDARD HOURS | Numerical Value - STANDARD HOURS |
| STOCK OPTIONS LEVEL | Numerical Value - STOCK OPTIONS (Higher the number, the more stock option an employee has) |
| TOTAL WORKING YEARS | Numerical Value - TOTAL YEARS WORKED |
| TRAINING TIMES LAST YEAR | Numerical Value - HOURS SPENT TRAINING |
| WORK LIFE BALANCE | Numerical Value - TIME SPENT BETWEEN WORK AND OUTSIDE |
| YEARS AT COMPANY | Numerical Value - TOTAL NUMBER OF YEARS AT THE COMPANY |
| YEARS IN CURRENT ROLE | Numerical Value -YEARS IN CURRENT ROLE |
| YEARS SINCE LAST PROMOTION | Numerical Value - LAST PROMOTION |
| YYEARS WITH CURRENT MANAGER | Numerical Value - YEARS SPENT WITH CURRENT MANAGER |

**Road map:**

Phase: 1

Understanding the business problem and the problems associated to this project.

Domain Knowledge about various components of the project. Develop initial level knowledge about how attrition rates effect business.

1. Explaining the Business Problem and our goals in detail.
2. Understanding the dataset by discussing about its various features (columns). Plan our approach
3. Study the data initially from Excel and create initial set of KBRs and KPIs for tracking attrition rate in the company and maintain proper notes for it.
4. Required Installations – Python, MS Excel, Tableau.

Phase: 2

After planning and studying our business requirement, it’s time to perform further data cleaning and EDA in Python using various Python libraries.

1. Loading the excel file data into a python pandas data frame, understanding the types of files and their connecting function in Python
2. Studying the data with python data frame features, checking for the exact numbers of null values in all the columns.
3. Clean the data using various functions within pandas’ package that helps in proper feature selection.
4. Plotting various features in a graph to see the correlation of metrics to our goals (KPIs).
5. Finalizing all our findings and converting the updated data frame into excel sheet to be further used in Tableau

Phase 3:

Finally in this phase our focus would be to visualize the cleaned data in Tableau.

Develop a dashboard that would best highlight the insights for attrition in the company (by using the metrics in our document in Phase 1 to track our created KPIs)

1. Connect the data in Tableau.
2. Create the required charts, according to the features and metrics selected during EDA.
3. Add dynamic features to the charts to make it more insightful and user friendly with the help of parameters and calculated fields.
4. Design a Dashboard that perfectly presents all the business insights with the dynamic chats created above in step 3.