

EMPLOYEE TURNOVER ANALYSIS

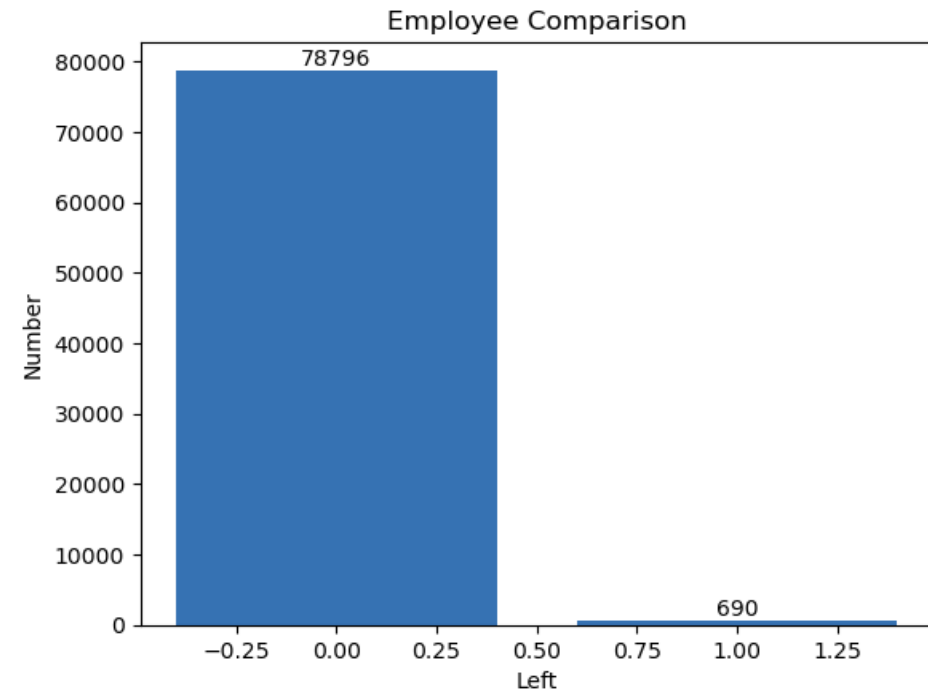
CASE STUDY (INTERN) - HR ADVANCED ANALYTICS - SCHAEFFLER

RAFI BIN ARMAN

DATA EXPLORATION

FEATURE OF INTEREST & TARGET VARIABLE

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 79486 entries, 0 to 79485
Data columns (total 20 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   Age                                         79486 non-null  int64
1   Length of service in months               79486 non-null  int64
2   Gender                                     79486 non-null  int64
3   Nationality                               79486 non-null  object
4   Weekly working hour                       79486 non-null  float64
5   Employee Category                         79486 non-null  object
6   Job                                         79486 non-null  object
7   Region                                     79486 non-null  object
8   Country                                    79486 non-null  object
9   HR location                               79486 non-null  object
10  Board                                      79486 non-null  object
11  Function                                   79486 non-null  object
12  Division // Segment                       79486 non-null  object
13  Segment Unit                             79486 non-null  object
14  Difference to Managers Length of Service (in Years) 79486 non-null  float64
15  Number of Colleagues in the Team          79486 non-null  int64
16  Probation Period                         79486 non-null  int64
17  Commuting distance (in Kilometers)        79486 non-null  float64
18  Commuting duration (in Minutes)          79486 non-null  float64
19  Left                                       79486 non-null  int64
dtypes: float64(4), int64(6), object(10)
memory usage: 12.1+ MB
```



DATA CLEANING

MISSING VALUES & IRREGULARITIES

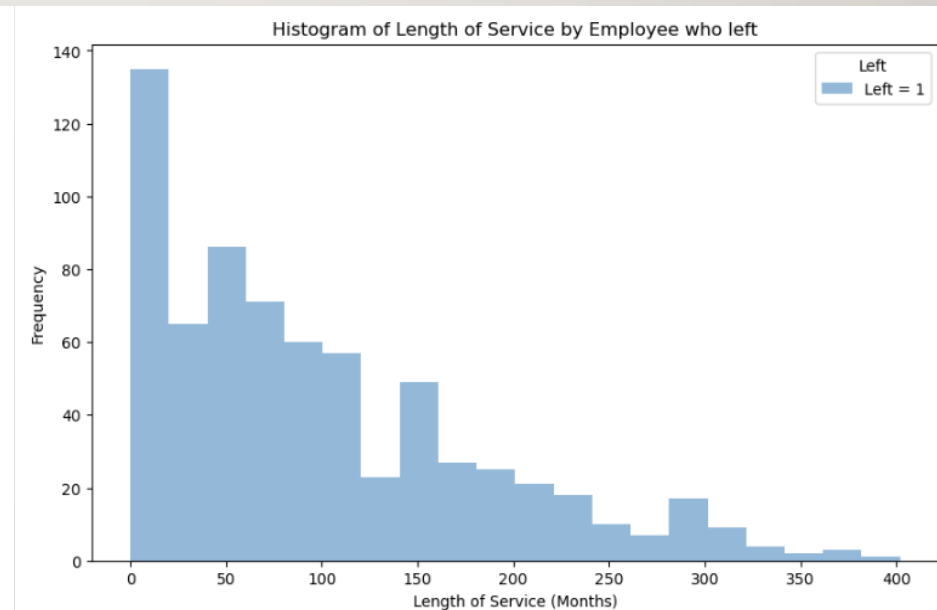
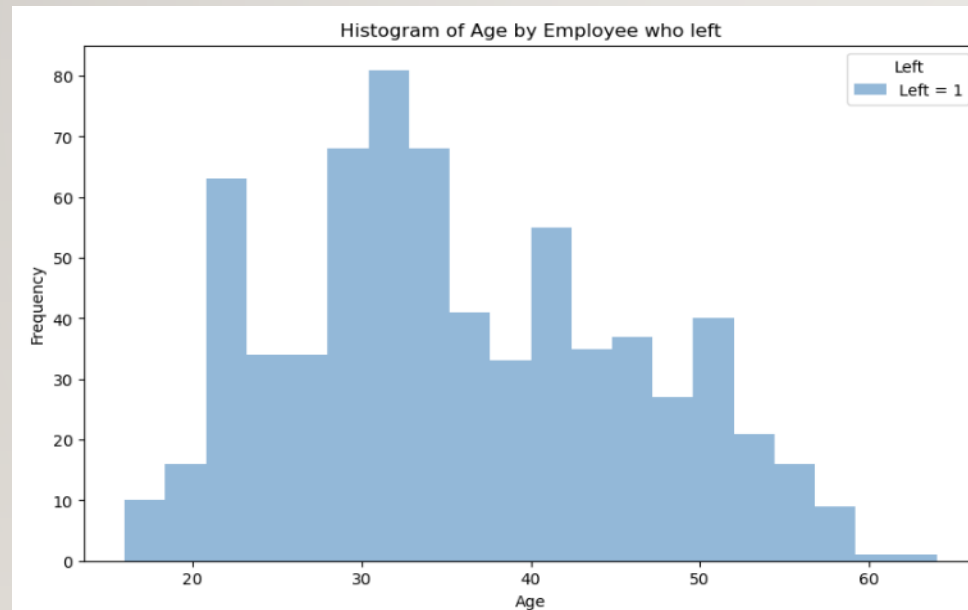
- Missing values in the features are replaced by the mean of the respective feature values.
- There are Irrelevant outlier values in Commuting distance (in Kilometers) & Commuting duration (in Minutes) which will be replaced by the mean of relevant values.

```
Age 0
Length of service in months 0
Gender 0
Nationality 0
Weekly working hour 0
Employee Category 0
Job 0
Region 0
Country 0
HR location 0
Board 0
Function 0
Division // Segment 0
Segment Unit 0
Difference to Managers Length of Service (in Years) 221
Number of Colleagues in the Team 0
Probation Period 0
Commuting distance (in Kilometers) 18871
Commuting duration (in Minutes) 18871
Left 0
dtype: int64
```

PATTERN RECOGNITION & INSIGHT GENERATION

UNIVARIATE ANALYSIS OF AGE & LENGTH OF SERVICE

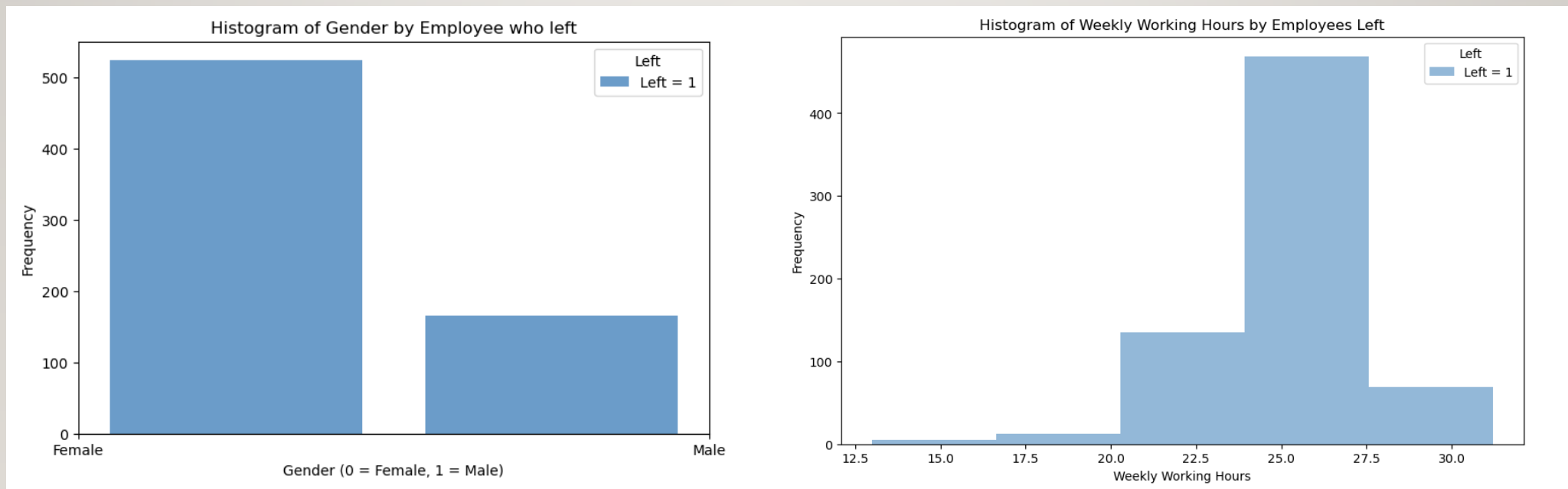
- Employees in early age(20-22) and mid-career(30-34) ones are more prone to leave the company.
- During first 1-2 years, there is high chances to leave the company and with the increase of length of service, the proportion of leaving is reduced.
- ✓ Salary details, Employment benefits, Market compensation benefits of these employees are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION

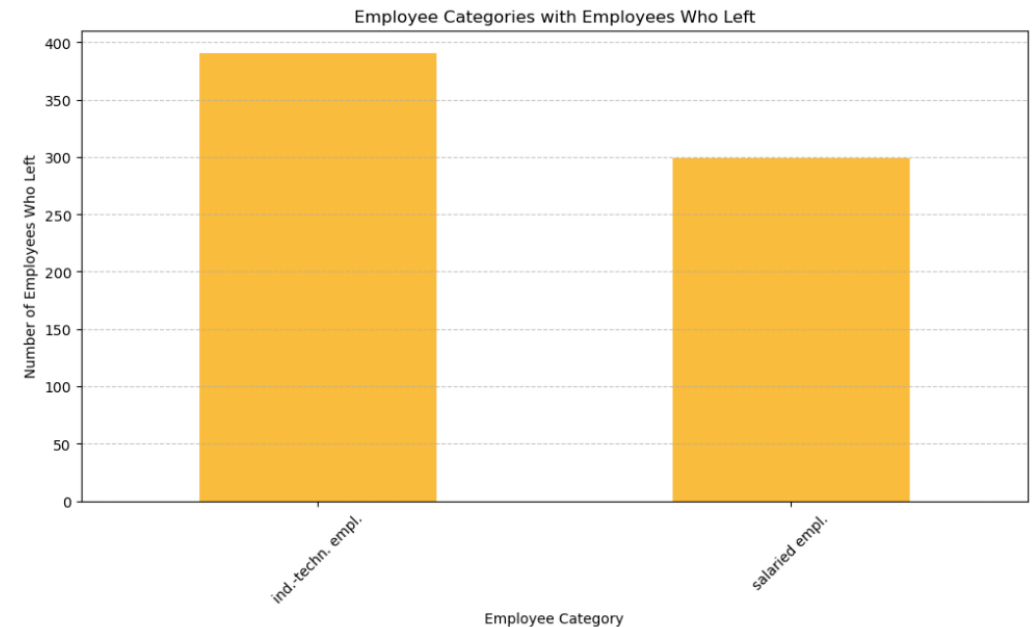
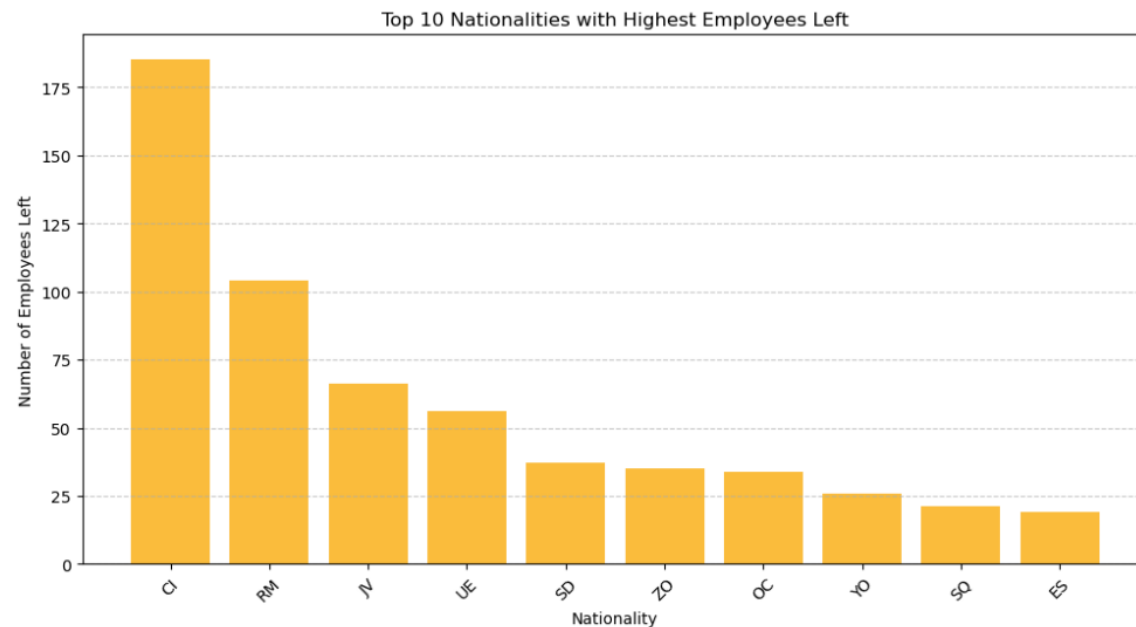
UNIVARIATE ANALYSIS OF GENDER & WEEKLY WORKING HOURS

- Women are more prone to leave the company. The portion is more than double compared to men.
- There is high churn rate for employees who work between 24 and 27.5 hours per week.
- ✓ Employment benefits for women and work stress level of the employees with weekly working hours (24-27.5) are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION UNIVARIATE ANALYSIS OF NATIONALITY & EMPLOYEE CATEGORY

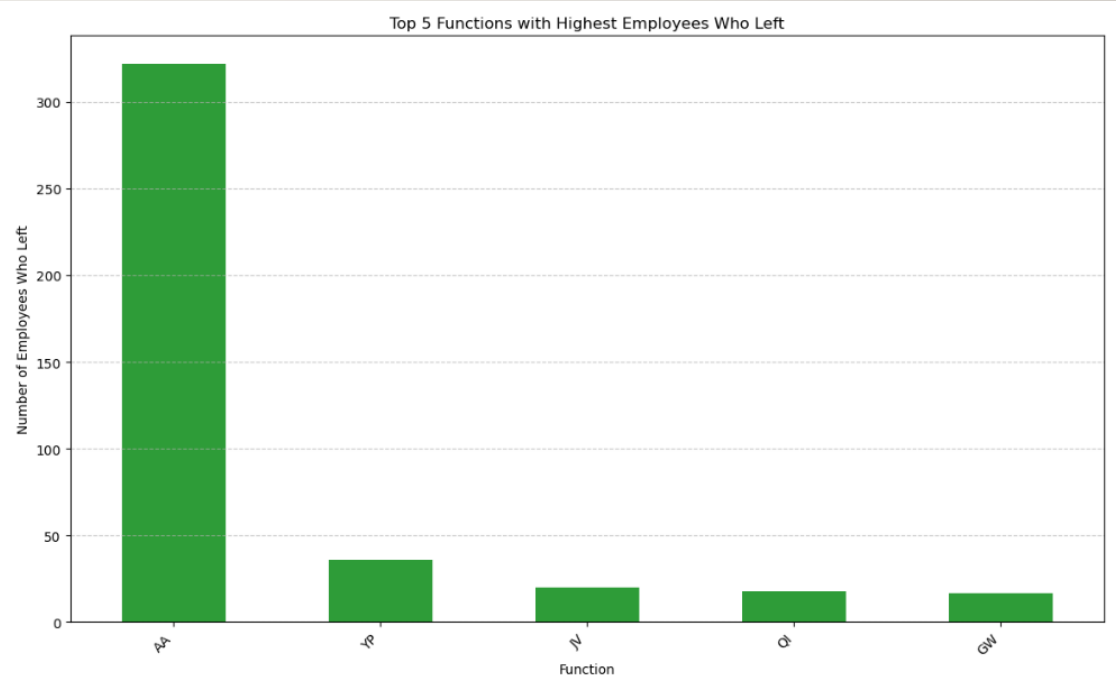
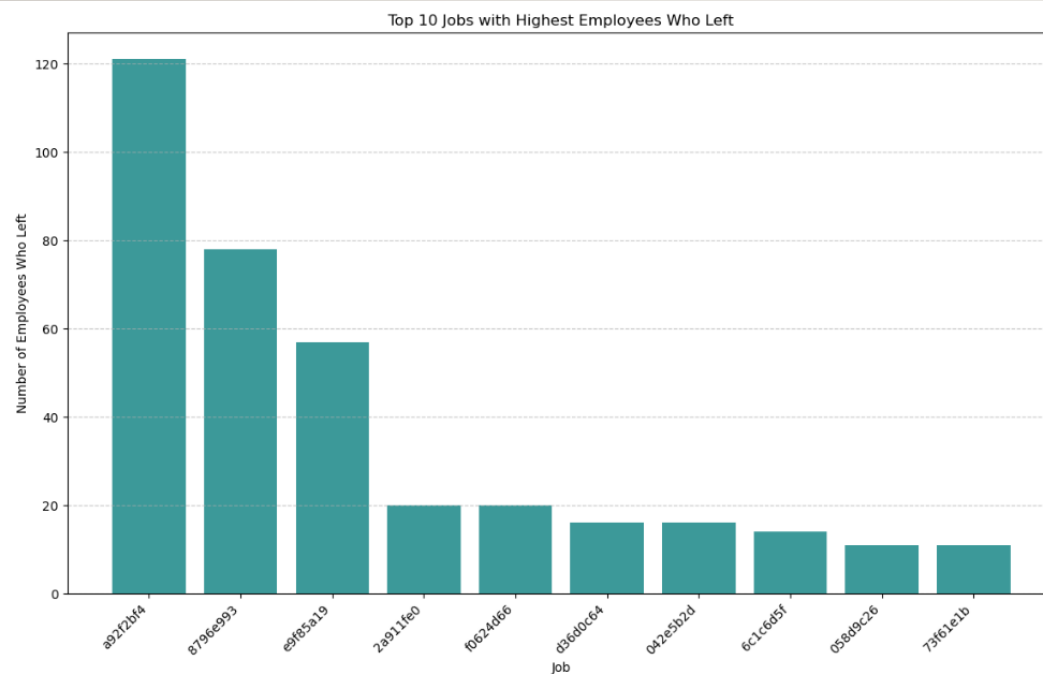
- Employees from 'CI' are showing the highest level of churn compare to other nationalities.
- Ind.-techn. empl. are more prone to leave the company compared to salaried empl.
- ✓ The discriminations focused on nationality 'CI' and Ind.-techn. empl. are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION

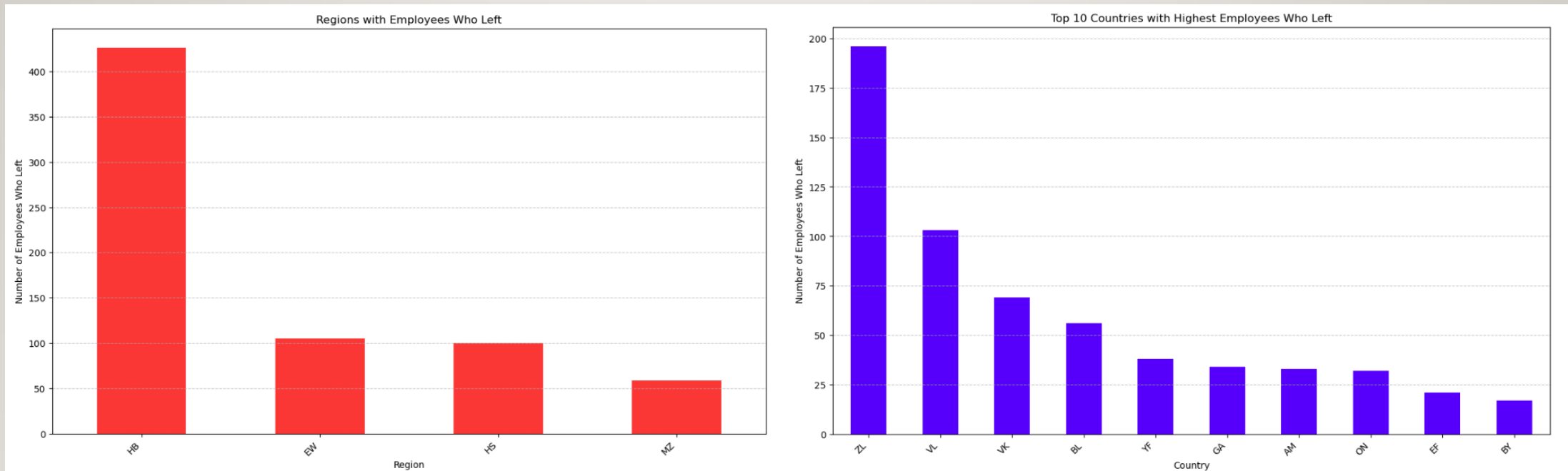
UNIVARIATE ANALYSIS OF JOBS & FUNCTIONS

- Employees with jobs 'a92f2bf4', '8796e993' and 'e9f85a19' have the high chances to leave the company.
- Employees of function AA are more prone to churn.
- ✓ Work pressure and work standards to these Jobs and Functions are needed to analyse further.



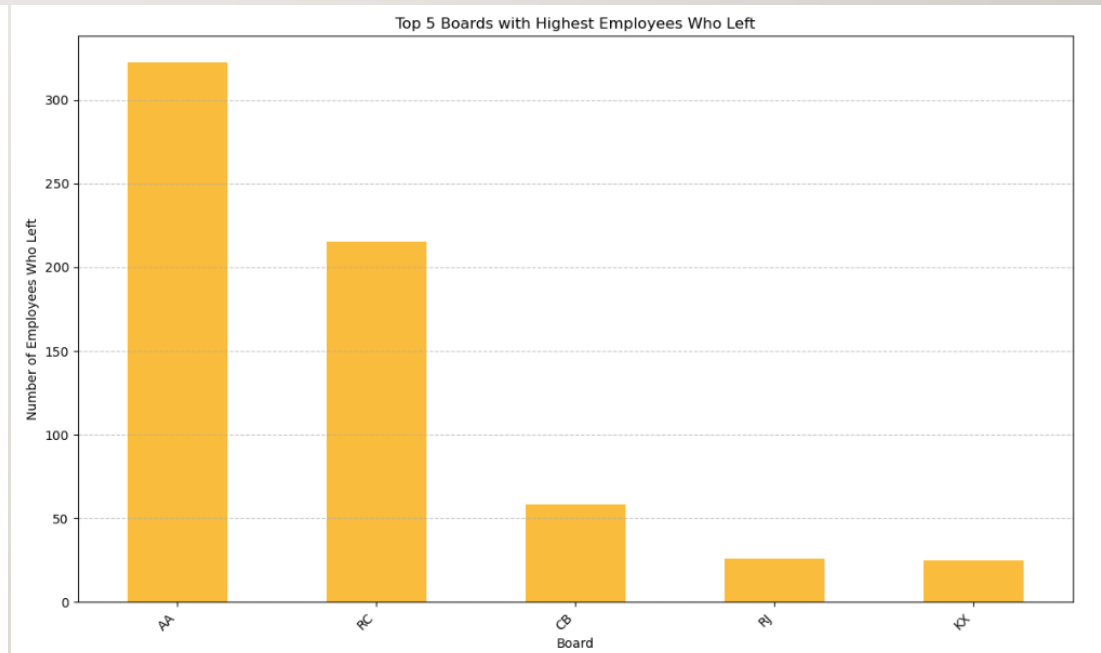
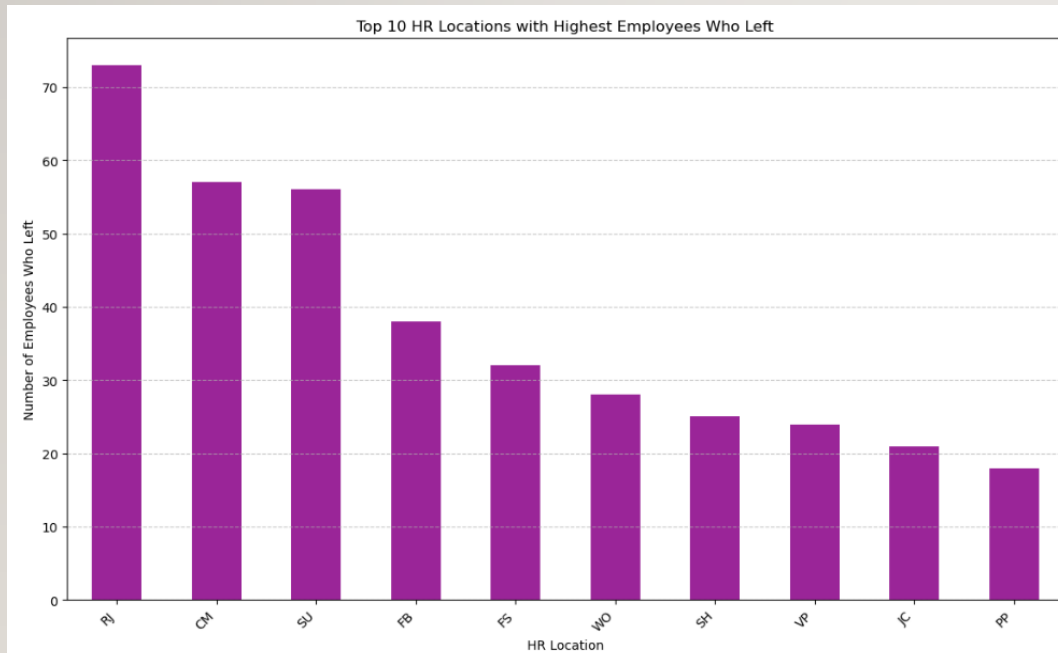
PATTERN RECOGNITION & INSIGHT GENERATION UNIVARIATE ANALYSIS OF REGIONS & COUNTRIES

- Employees with Region HB have the high chances to the leave the company.
- Employees of country ZL, VL, VK, and BL are more prone to churn.
- ✓ Work ethics of Region HB and the discriminations related to these countries are needed to analyse.



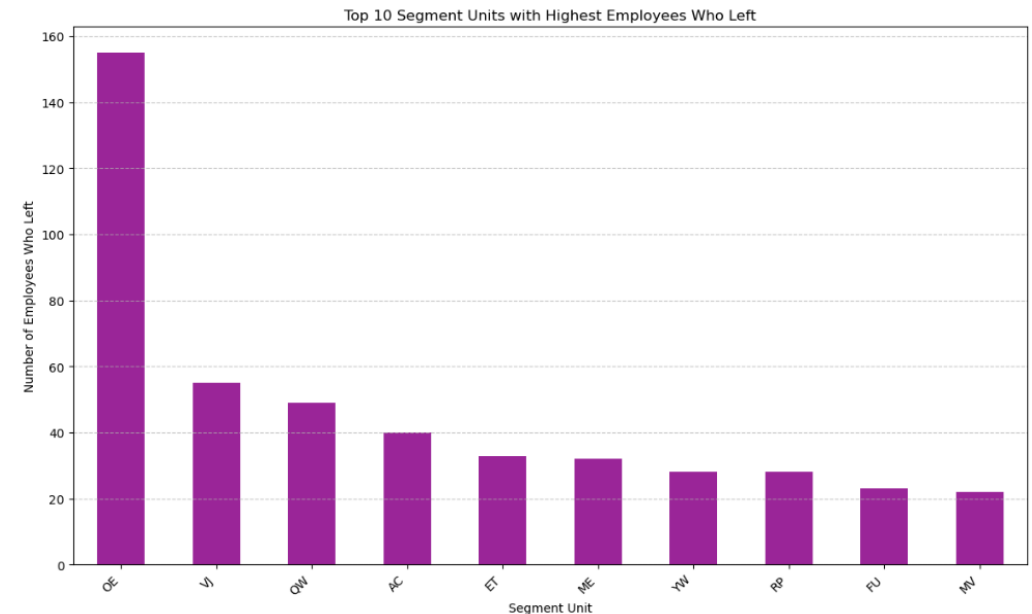
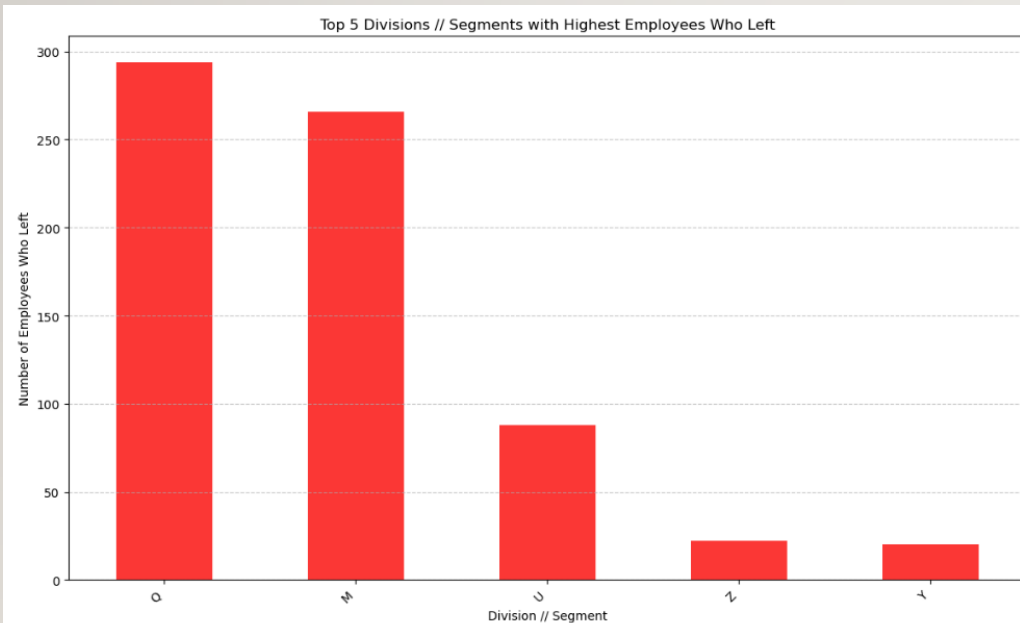
PATTERN RECOGNITION & INSIGHT GENERATION UNIVARIATE ANALYSIS OF HR LOCATIONS & BOARDS

- Employees with HR Location RJ, CM, SU have the high chances to the leave the company.
- Employees of Boards AA, RC are more prone to churn.
- ✓ Employee management of these HR Locations and Boards are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION UNIVARIATE ANALYSIS OF DIVISIONS // SEGMENTS & SEGMENT UNITS

- Employees with HR Divisions // Segments Q and M have the high chances to the leave the company.
- Employees of Segment Units OE are more prone to churn.
- ✓ Work pressure of these Divisions // Segments and Segment Units are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION

UNIVARIATE ANALYSIS OF DIFFERENCE TO MANAGERS LENGTH OF SERVICE

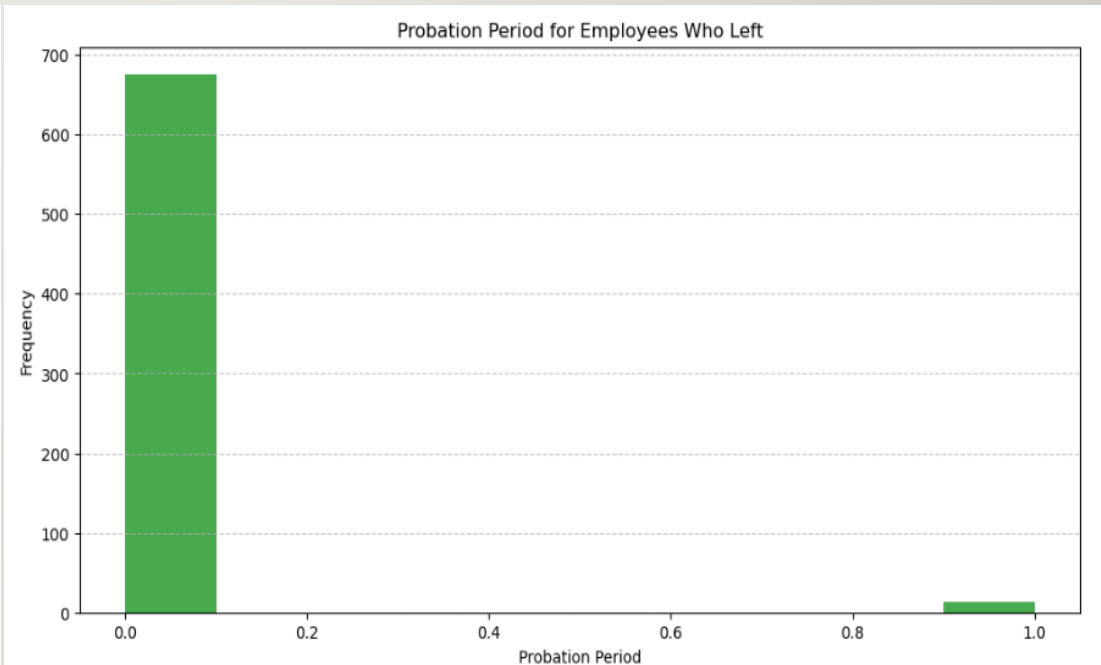
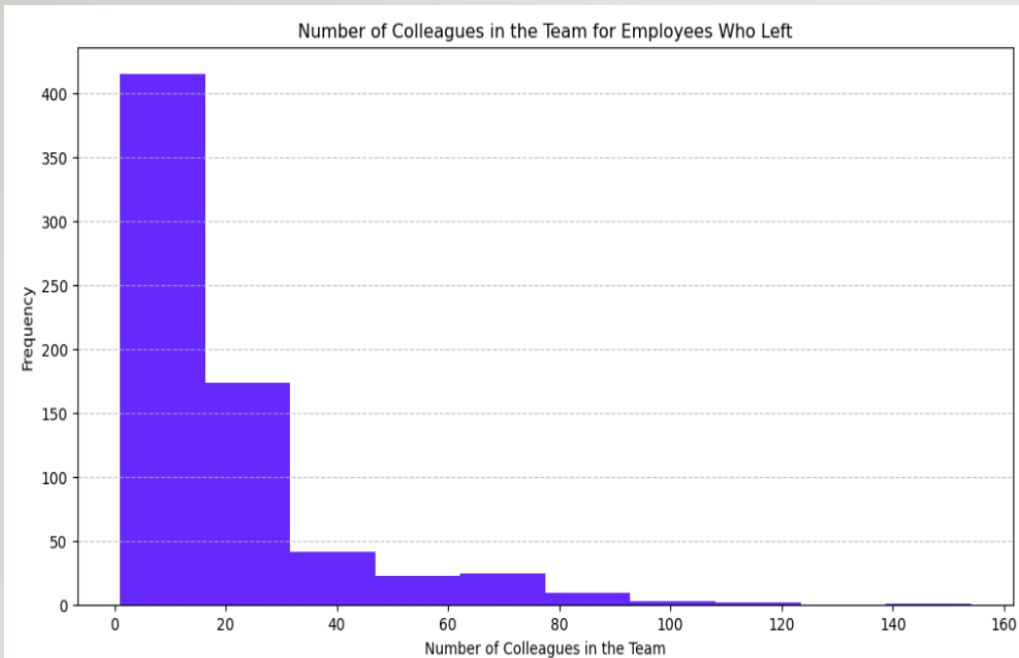
- Employees whose managers are of same age or 2-4 junior or senior than them are more prone to leave.
- ✓ Different work metrics like team work culture, mutual understanding of these scenarios are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION

UNIVARIATE ANALYSIS OF TEAM MEMBERS & PROBATION PERIOD

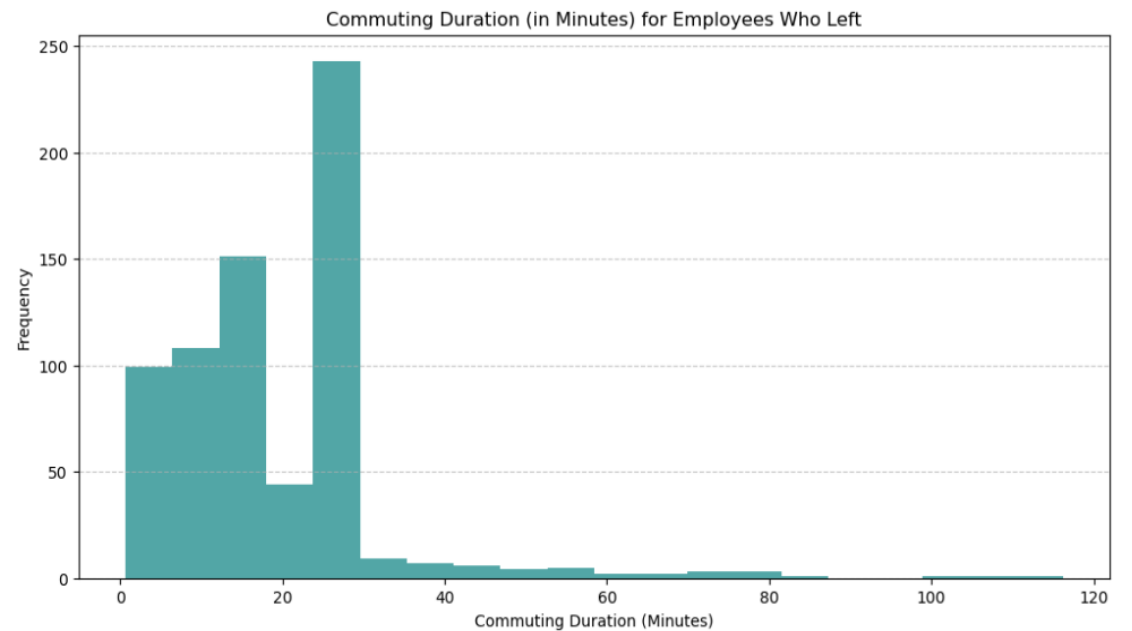
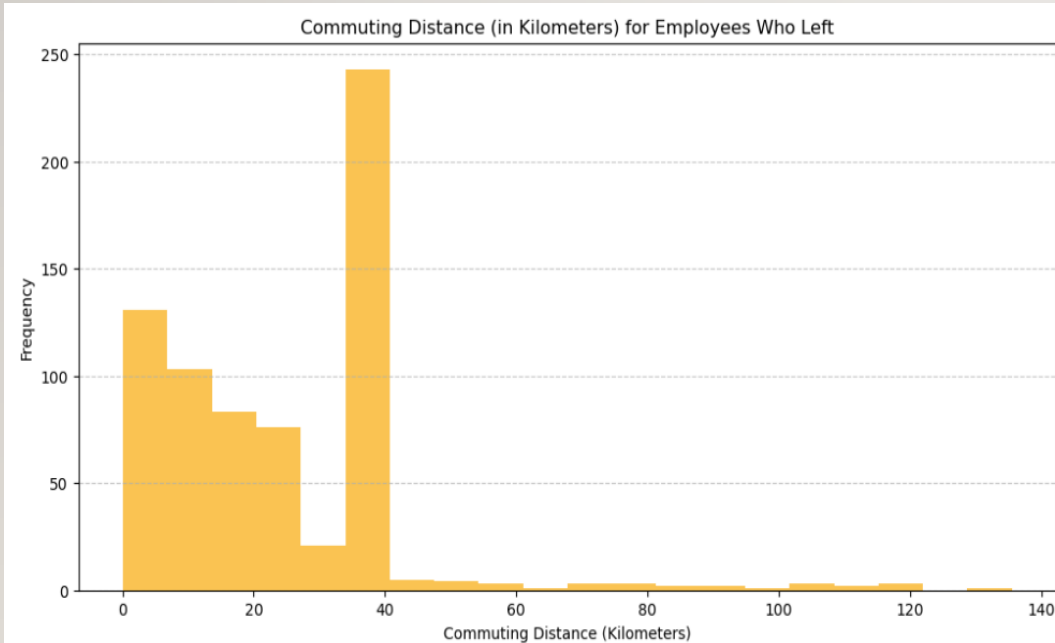
- Employees with small teams like 15 people have the high chances to leave the company.
- There is a trend that employees in probation period are less prone to churn.
- ✓ Work pressure, understanding problems of small teams are needed to analyse.



PATTERN RECOGNITION & INSIGHT GENERATION

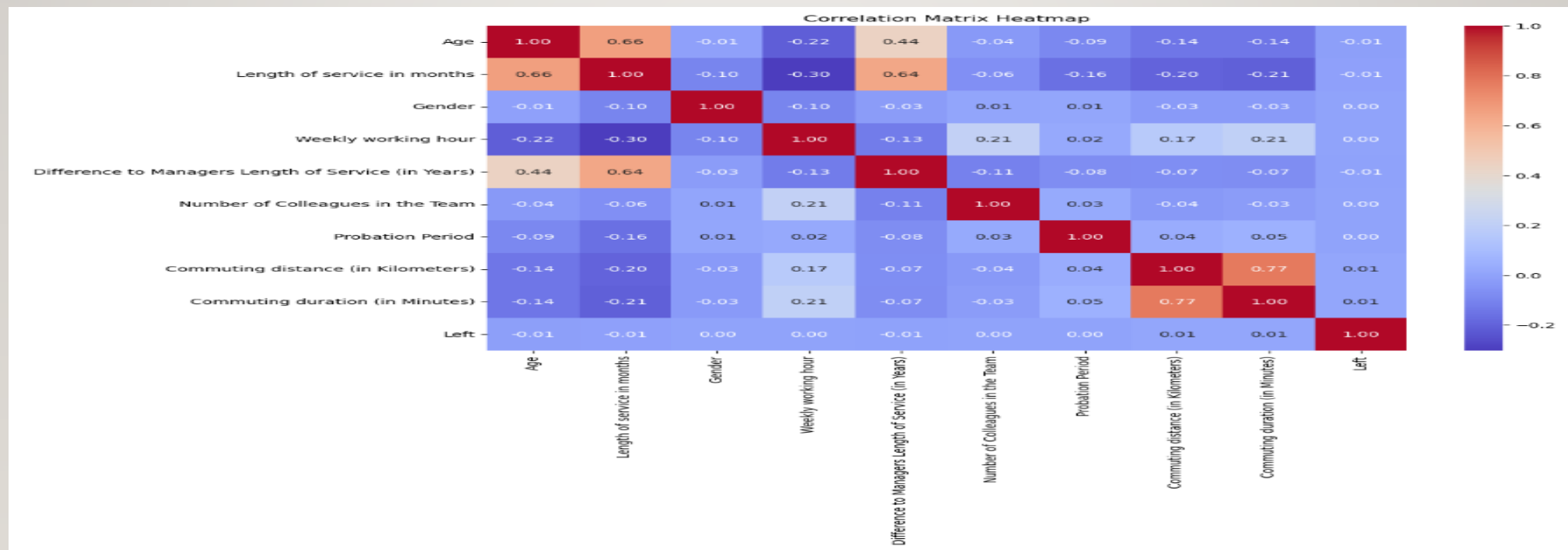
UNIVARIATE ANALYSIS OF COMMUTING DISTANCE AND DURATION

- Employees with commuting distance of less than 40 kms have the high chances to leave the company.
- Employees with commuting duration of less than 30 minutes are more prone to churn.
- ✓ Commuting types and locality can be analysed further to sort this issue.



PATTERN RECOGNITION & INSIGHT GENERATION CORRELATION MATRIX

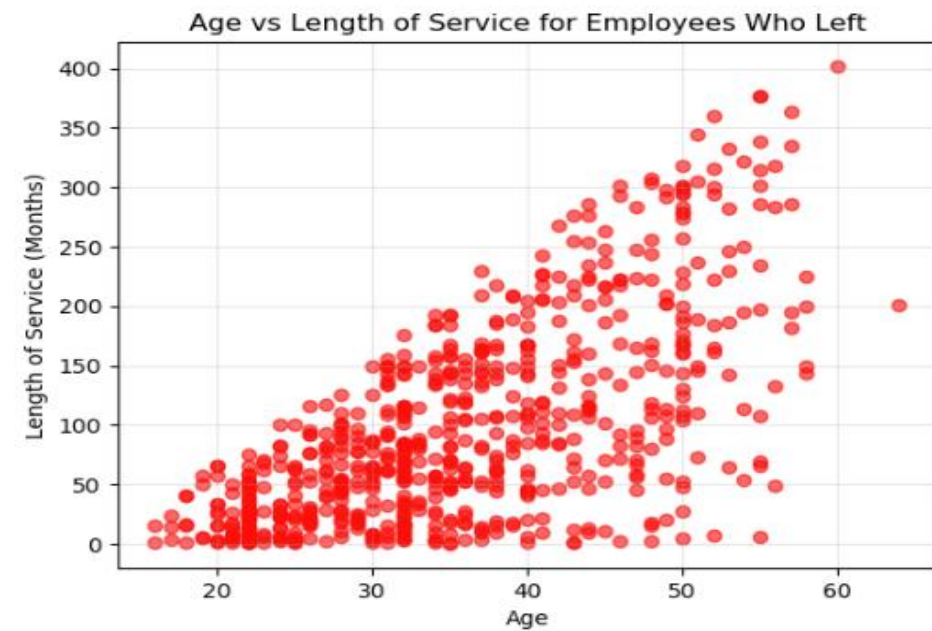
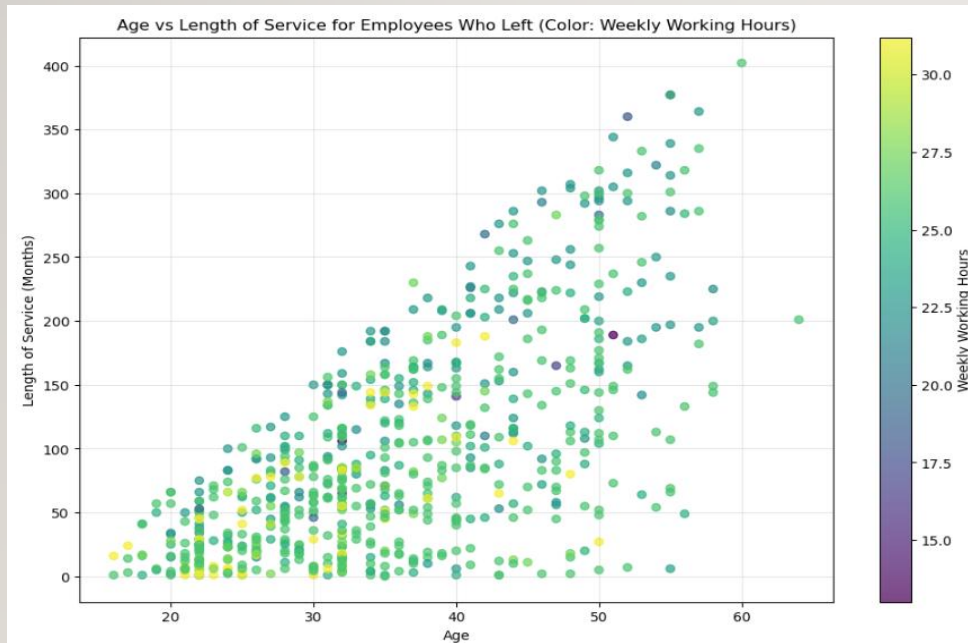
- Considering only the numerical features, the correlation matrix shows that the features independently are not so significant in deciding the employee churn rather the features combinedly impact on the employees leaving the company.



PATTERN RECOGNITION & INSIGHT GENERATION

MULTIVARIATE ANALYSIS

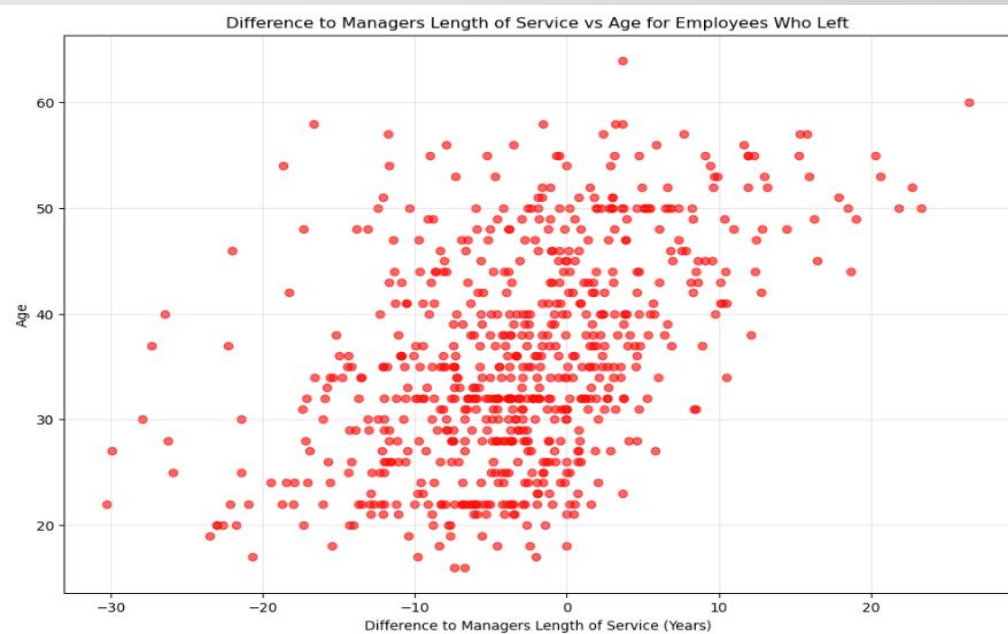
- Employees with less than 150 months of service and less than 40 years of age have the high chances to leave.
- Additionally, around 30 hours of work per week makes the churn more occurring.



PATTERN RECOGNITION & INSIGHT GENERATION

MULTIVARIATE ANALYSIS

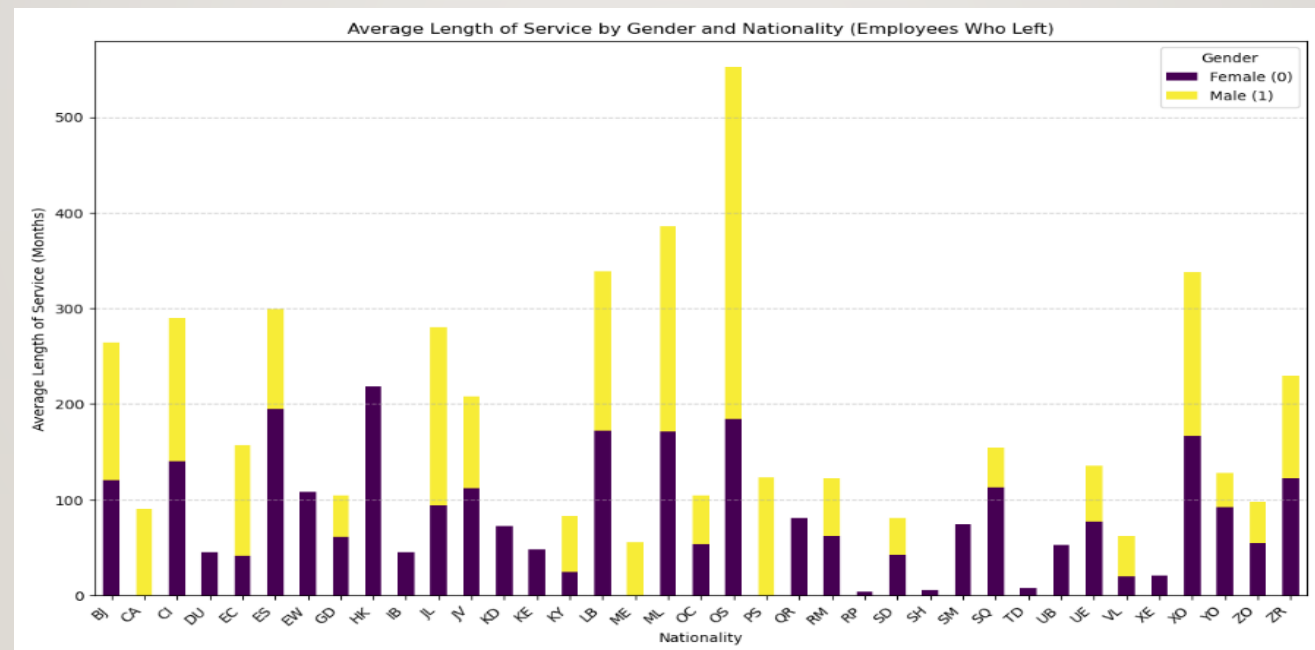
- Employees of less than 40 years of age and whose manager is as same age or 10 years junior than them have the high chances to leave the company.
- Additionally, with less than 100 months of the length of service makes the churn more occurring.



PATTERN RECOGNITION & INSIGHT GENERATION

MULTIVARIATE ANALYSIS

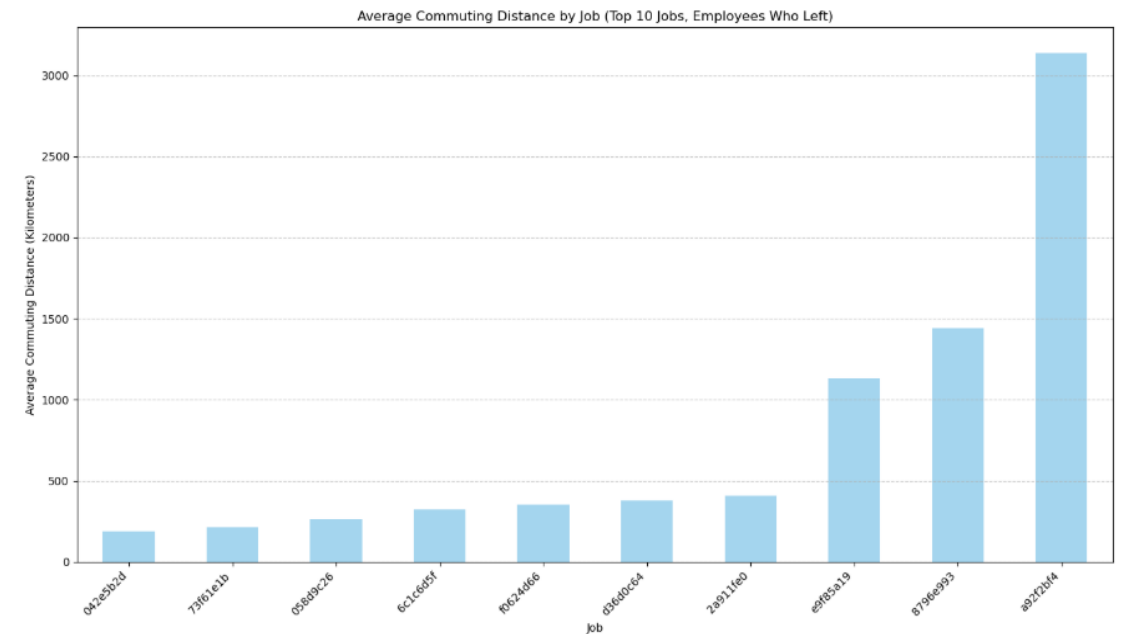
- Male employees tend to have longer average service durations than females across most nationalities, with significant variations among nationalities.
- Nationalities(OS, ML, LB) where the average length of service is substantially higher for both genders, indicating potential cultural or organizational influences.



PATTERN RECOGNITION & INSIGHT GENERATION

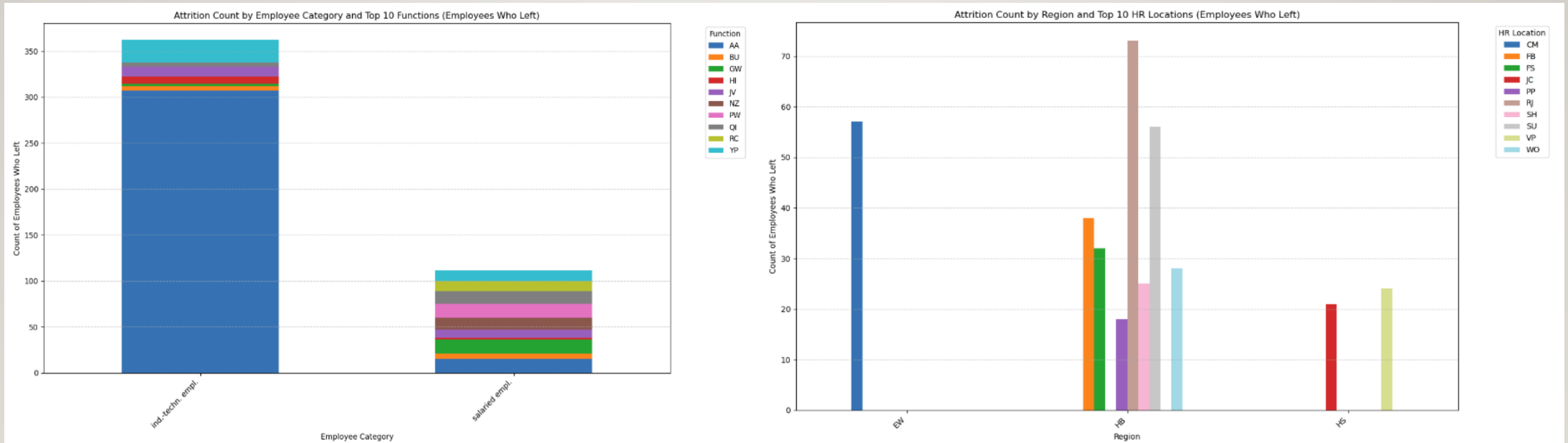
MULTIVARIATE ANALYSIS

- With less commuting distance and duration, employees are more prone to churn. Probably in case of high distance and duration, employees shift their house close to office that reduces the churning rate.
- Specific jobs ('a92f2bf4', '8796e993' and 'e9f85a19') have significantly higher average commuting distances among employees who left.



PATTERN RECOGNITION & INSIGHT GENERATION MULTIVARIATE ANALYSIS

- The "Ind. –techn. empl." category experiences significantly higher attrition across the top 10 functions (especially in AA) compared to the "Salaried Employees" category.
- The attrition is concentrated in specific regions, with "HB" showing the highest attrition rates distributed across multiple HR locations.



THANK YOU!