



A DATA MINING APPROACH ON ATTRITION RATE ANALYSIS OF EMPLOYEES

GROUP 4: HARSHINI KEERTHI VASAN, VINU SREENIVASAN, SANJANA ARAVINDAN



PROJECT GOAL

“A GREAT EMPLOYEE IS LIKE A FOUR LEAF CLOVER, HARD TO FIND AND LUCKY TO HAVE”

- This makes employee retention a very important aspect for the growth of the organization.
- **Attrition rate** is a measure of the number of individuals or items moving out of a collective group over a specific period.
- **Goal** : Develop an evaluation model to determine the most important factors that influence the attrition rate of employees.

DATASET

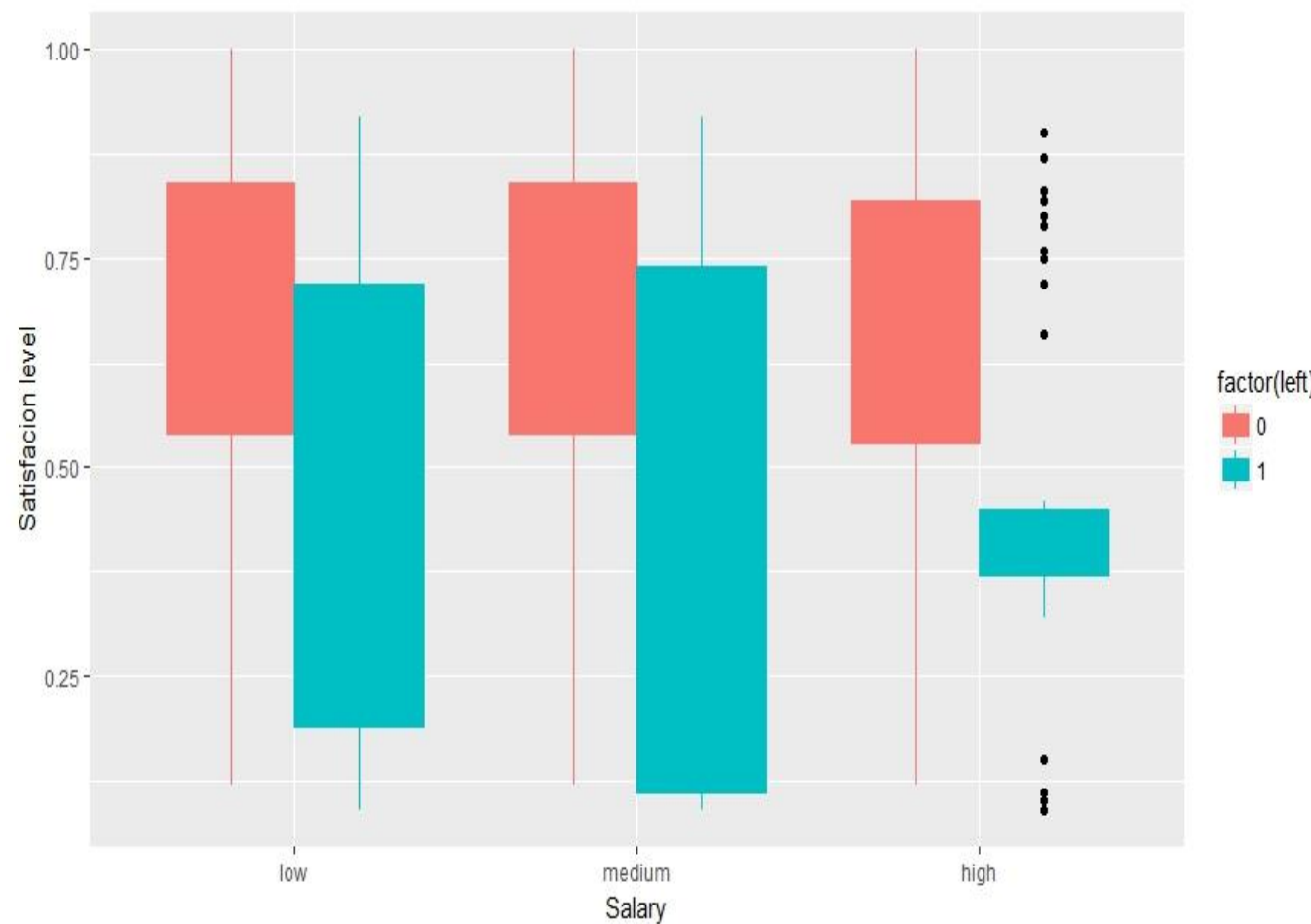
- Used the HR Analytics dataset available in Kaggle.
- Factors that describe each employee's record : employee satisfaction level, last evaluation, number of projects, average monthly hours, time invested for the company, work accident, promotion in the last 5 years, department, salary, current employee or ex-employee

DATA PREPROCESSING

- **Improper Evaluation** : Employees with last evaluation > 0.87 and salary = low , assign “YES”
- **Overrated** : Employees with satisfaction level, last evaluation and number of project less than median value, and, promoted= “yes”, assign “YES”
- **Average Daily Hours** = Average Monthly Hours/22

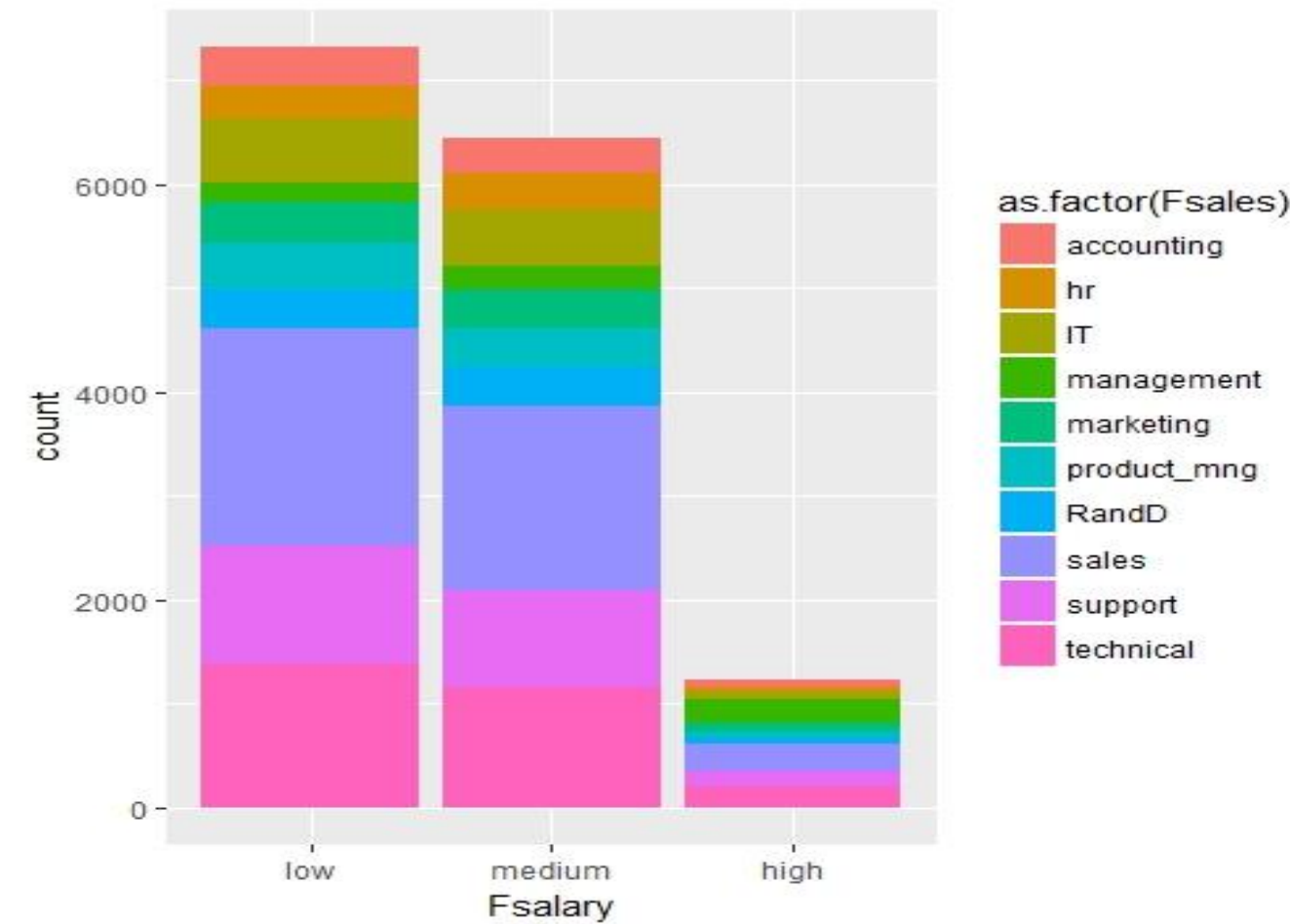
DATA EXPLORATION

Exploration done with seaborn plots in Python and gg plots in R.



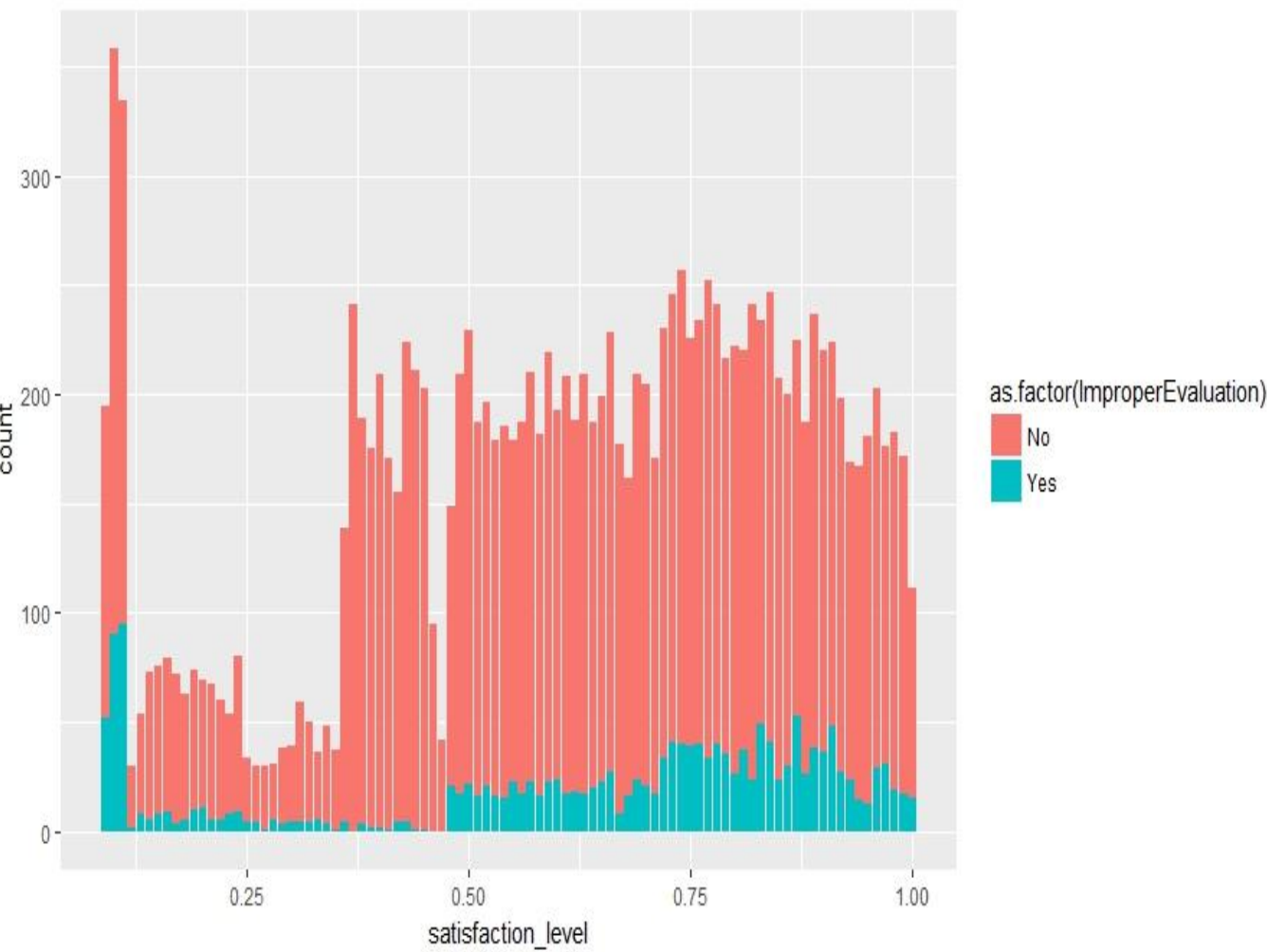
ANALYSIS

Employees leaving the company have spent more years with salary levels “low ” and “medium”



ANALYSIS

The sales department has majority of employees falling under salary levels “low” and “medium” category.

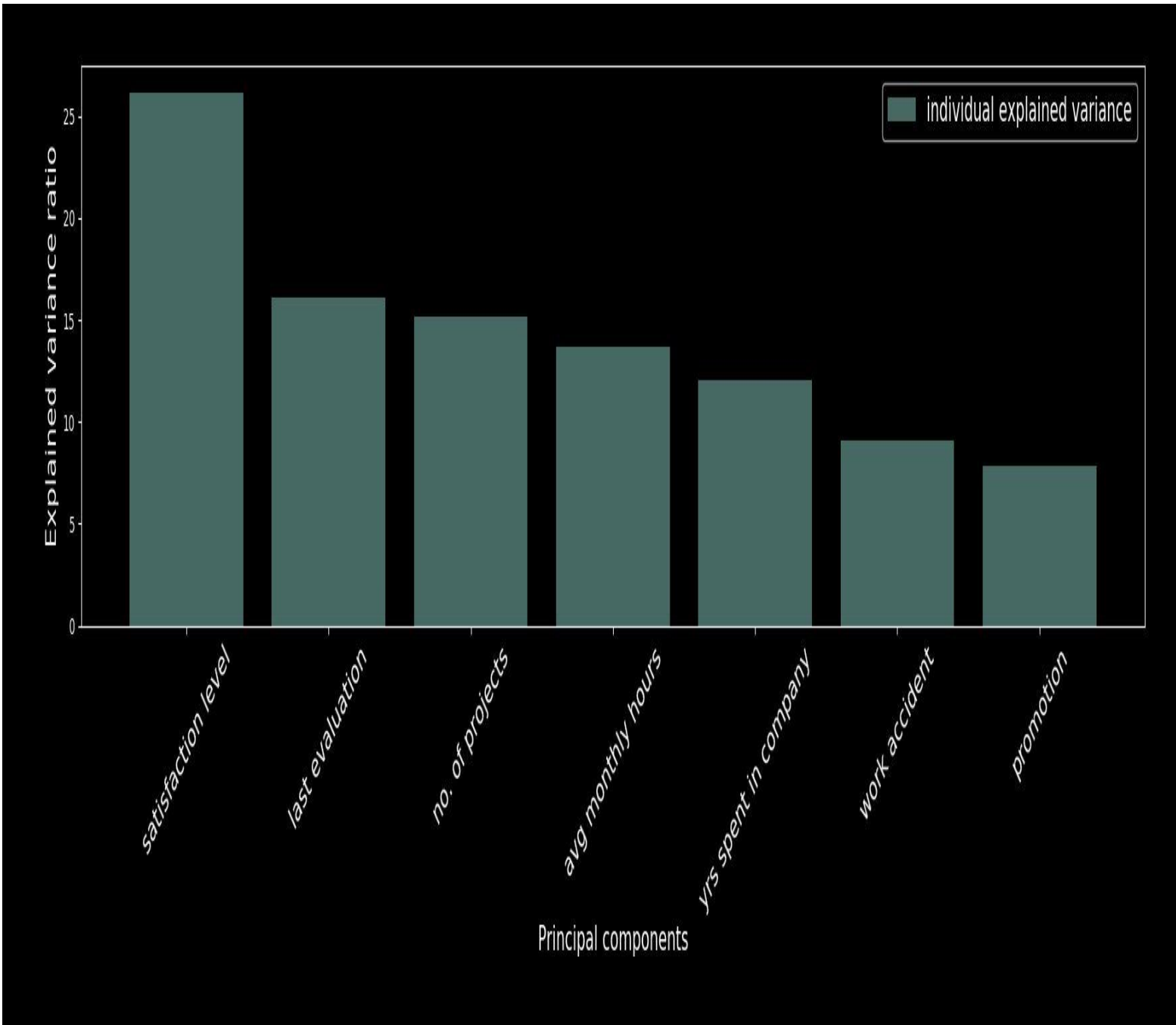


ANALYSIS

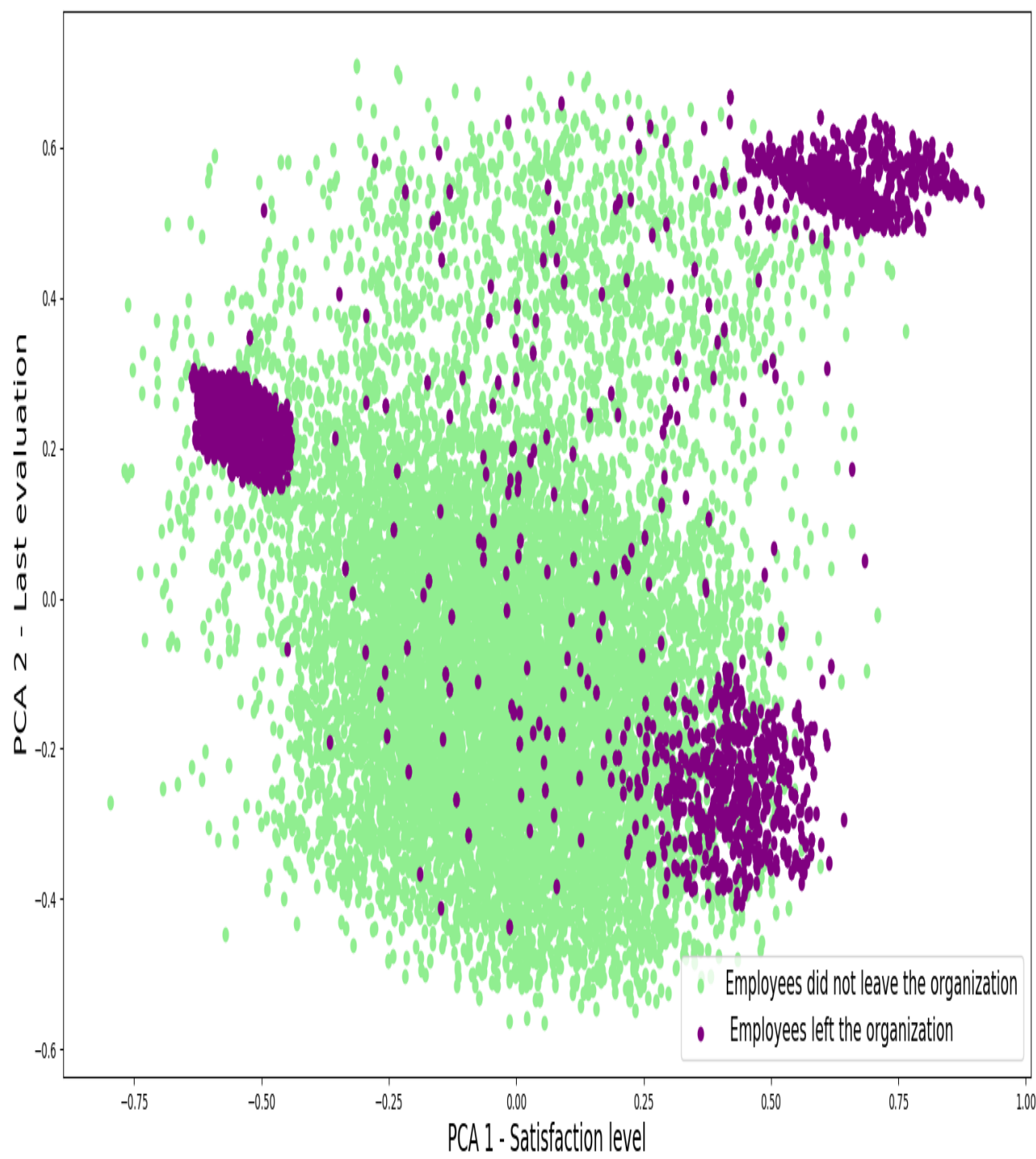
Employees with higher “satisfaction level” are evaluated incorrectly.

PRINCIPAL COMPONENT ANALYSIS

The high dimensional employee attributes went through dimensional reduction via PCA.

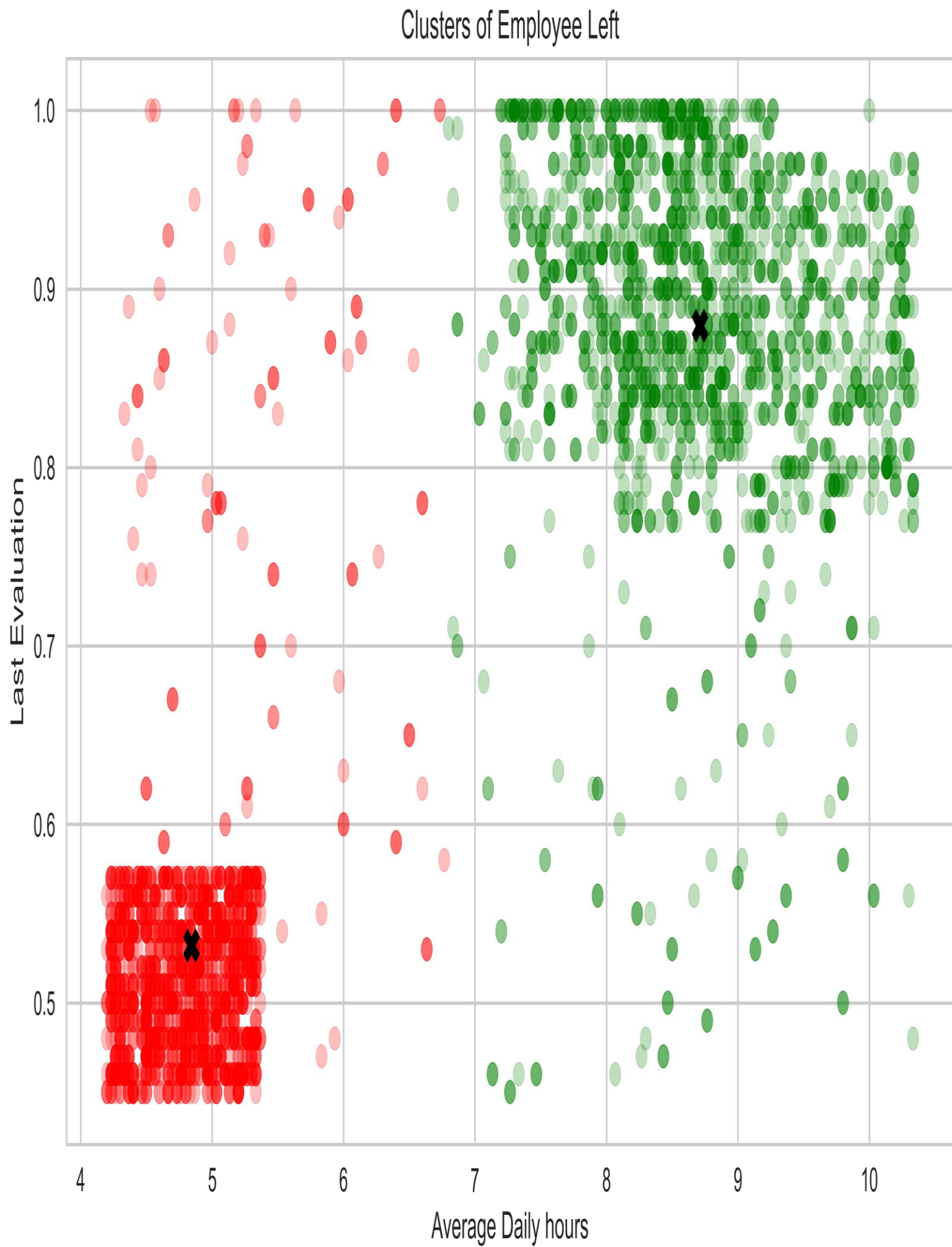


2-COMPONENT PCA



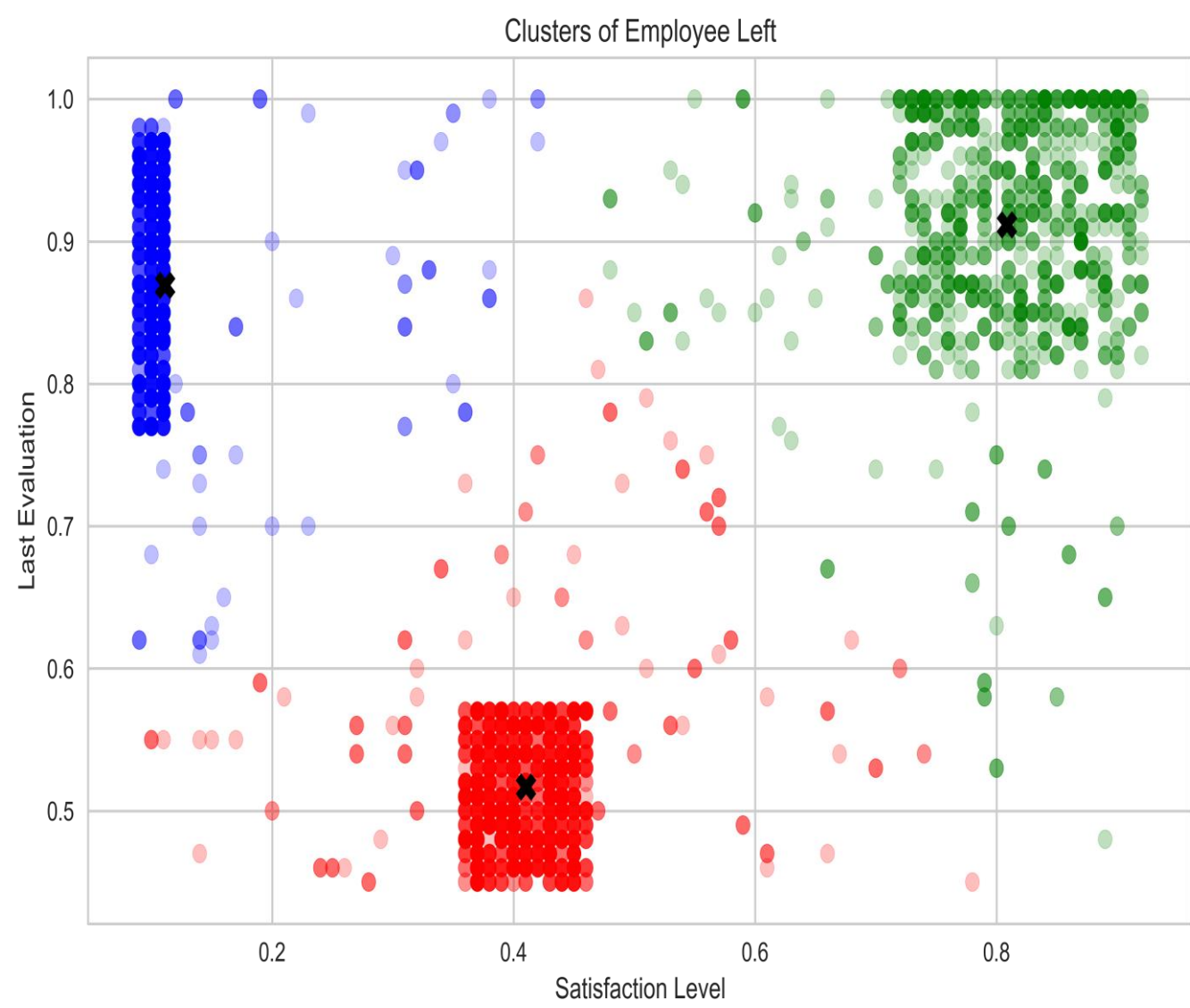
K-MEANS CLUSTERING

- Clustering dimensional reduction attributes
- Analysed clusters for continuous valued attributes



ANALYSIS

Employees with high “average daily hours” and high “last evaluation” , “low” average daily hours and “low” evaluation tend to leave the organization.

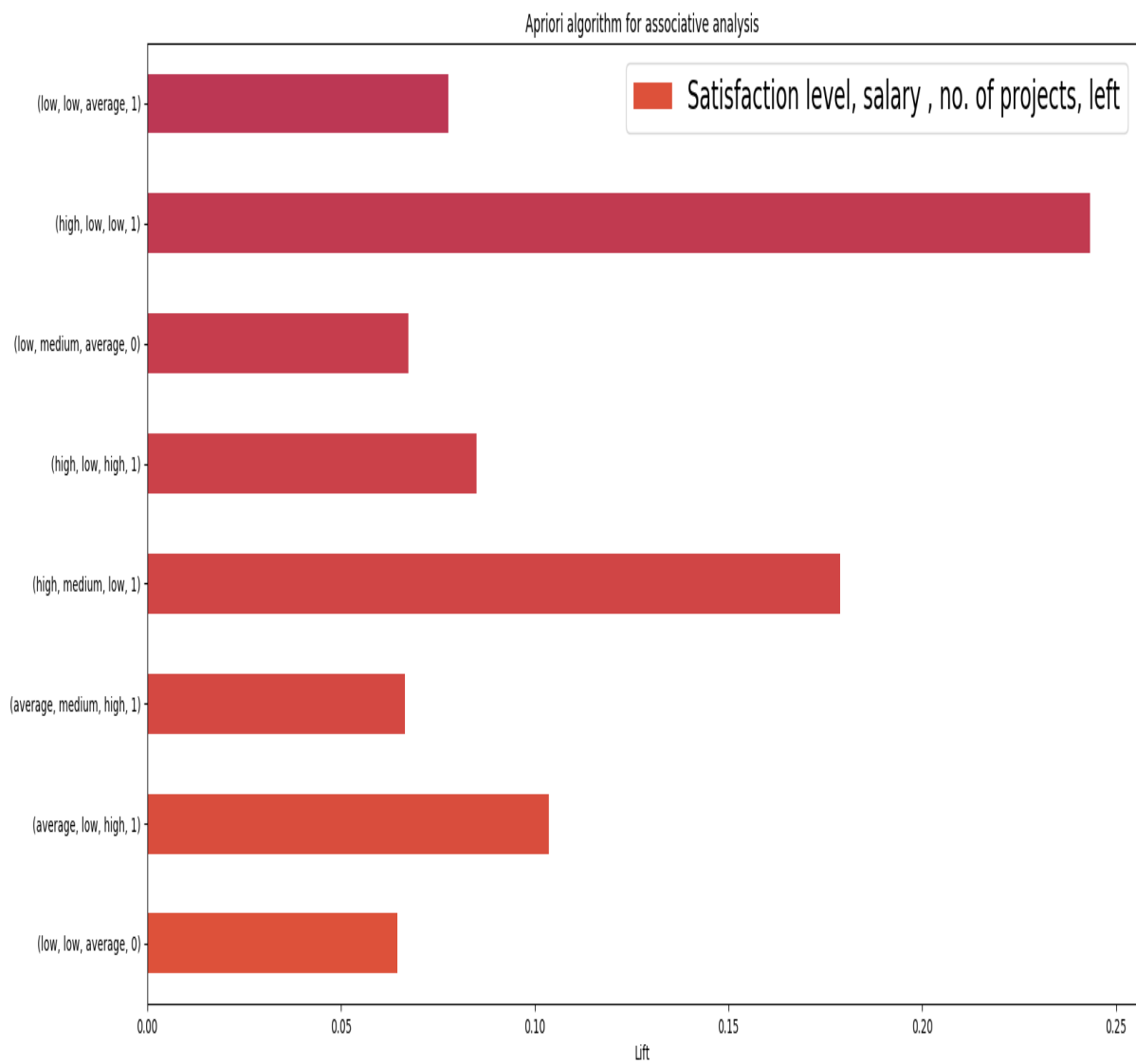


ANALYSIS

Employees with “average” satisfaction level and low “last evaluation” , “low” satisfaction level and “high” last evaluation, “high” satisfaction level and “high” last evaluation tend to leave the organization.

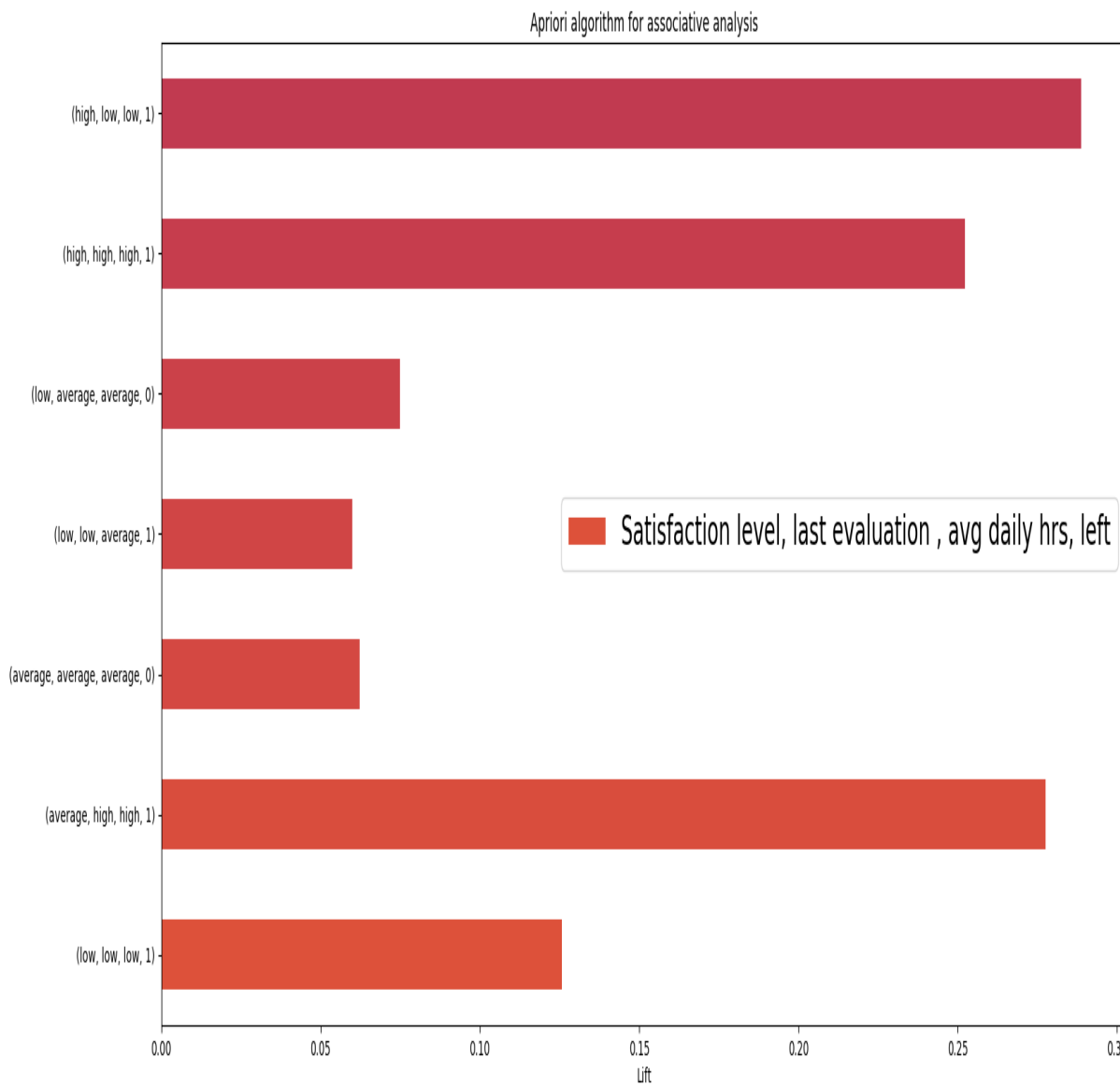
ASSOCIATIVE ANALYSIS (USING APRIORI ALGORITHM)

- Implemented Apriori algorithm with minimum support and confidence.
- Determined the associativity between attributes (not clusterable attributes) and employees who left the organization.



ANALYSIS

Employees with “high” satisfaction level, “low” salary and “low” number of projects mostly tend to leave the organization.



ANALYSIS

Employees with “high” satisfaction level, “low” evaluation and “low” average daily hours mostly tend to leave the organization.

CONCLUSION

- The major factors that affect the employees attrition rate are **satisfaction level, last evaluation, number of projects, salary** and **average daily hours**.
- The result obtained above would help an organization make positive changes and deliver better experiences to meet the employee's expectations.
- This will ultimately reduce the attrition rate of the employees.