EMPLOYEE TURNOVER ANALYSIS

Idris Mahamat

2025MIS637-WZ

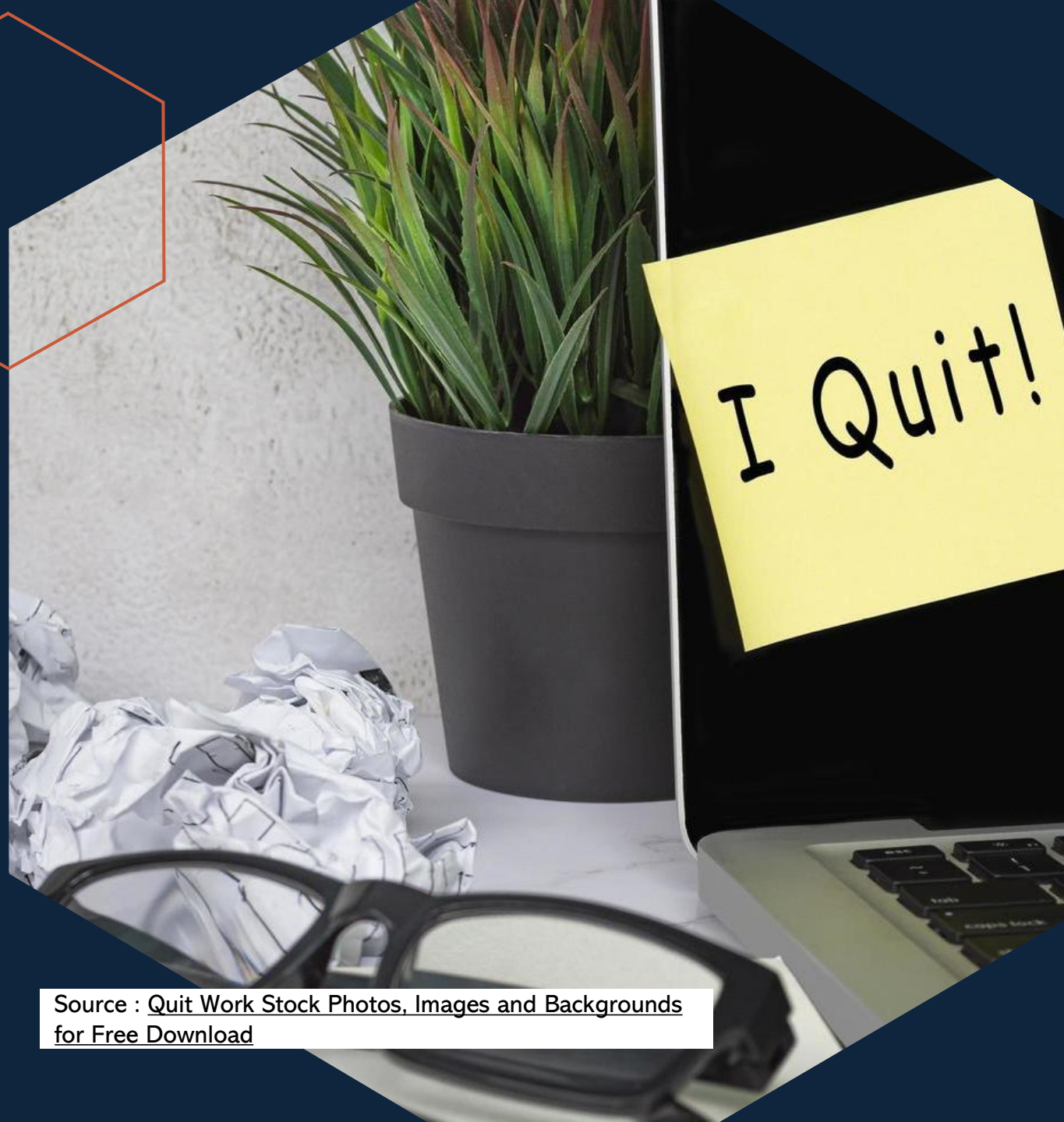
Agenda

- ✓ Introduction
- ✓ Exploratory Data Analysis
  - ✓ Dataset description
  - ✓ Bar chart
  - ✓ Histogram
- ✓ Modeling
- ✓ Model performance
- ✓ Implementation
- ✓ Summary

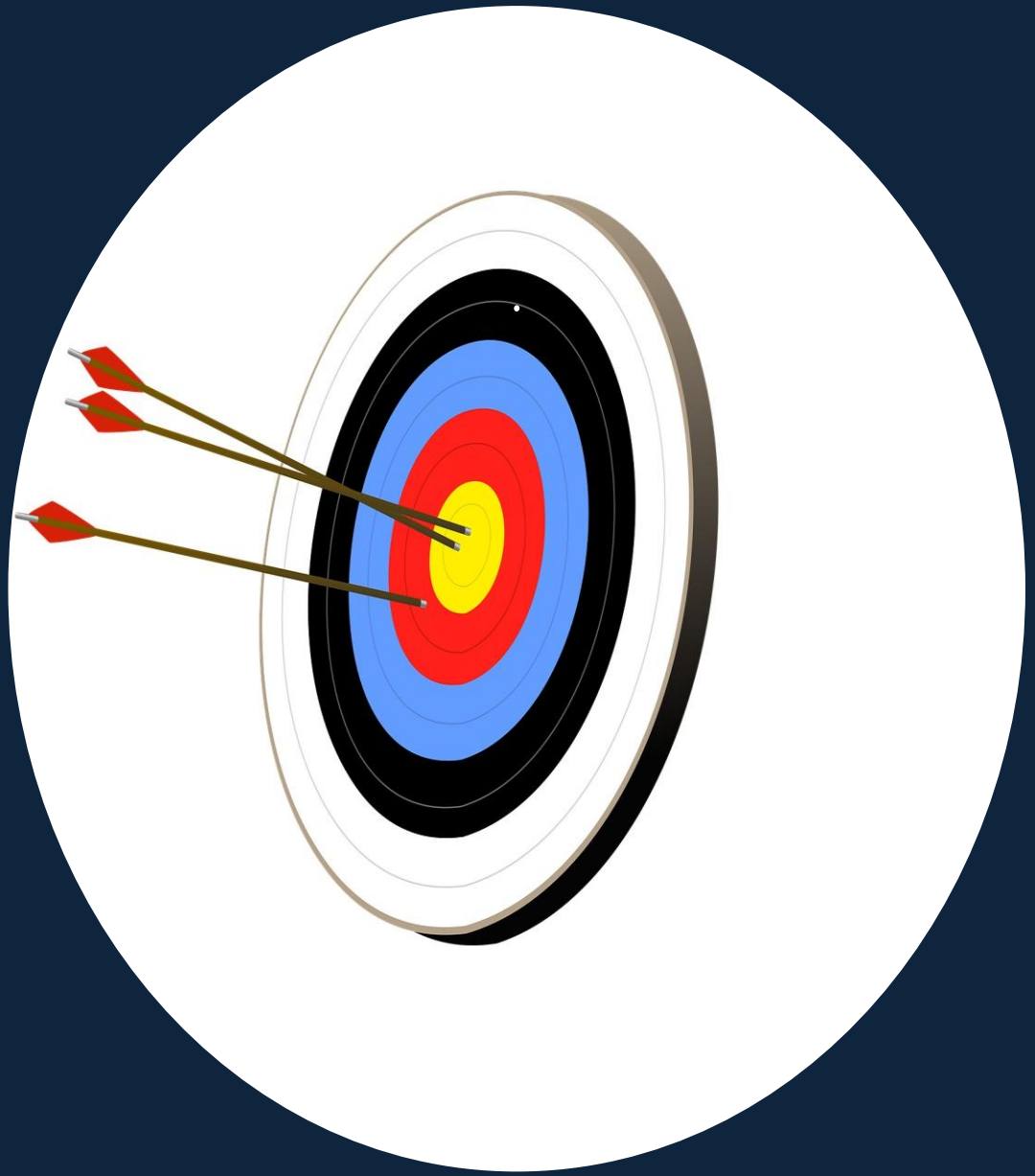
Introduction

Employee turnover analysis is an HR analysis that involves collecting data, analyzing and reporting HR data to help understand the company's turnover rate. In our use case, we will study the dataset name "**data_generated**" which was randomly generated from mackaroo.com.

The implementation of this use case will be as follow : first we are going to conduct a basic data exploratory analysis, then build different machine learning models, compare each model's performance and finally choose the model with the highest accuracy.



Source : Quit Work Stock Photos, Images and Backgrounds for Free Download



- Find out the reason that employee is resigning from the company.

- Determine the key factor that led to the turnover.

Employee turnover analysis

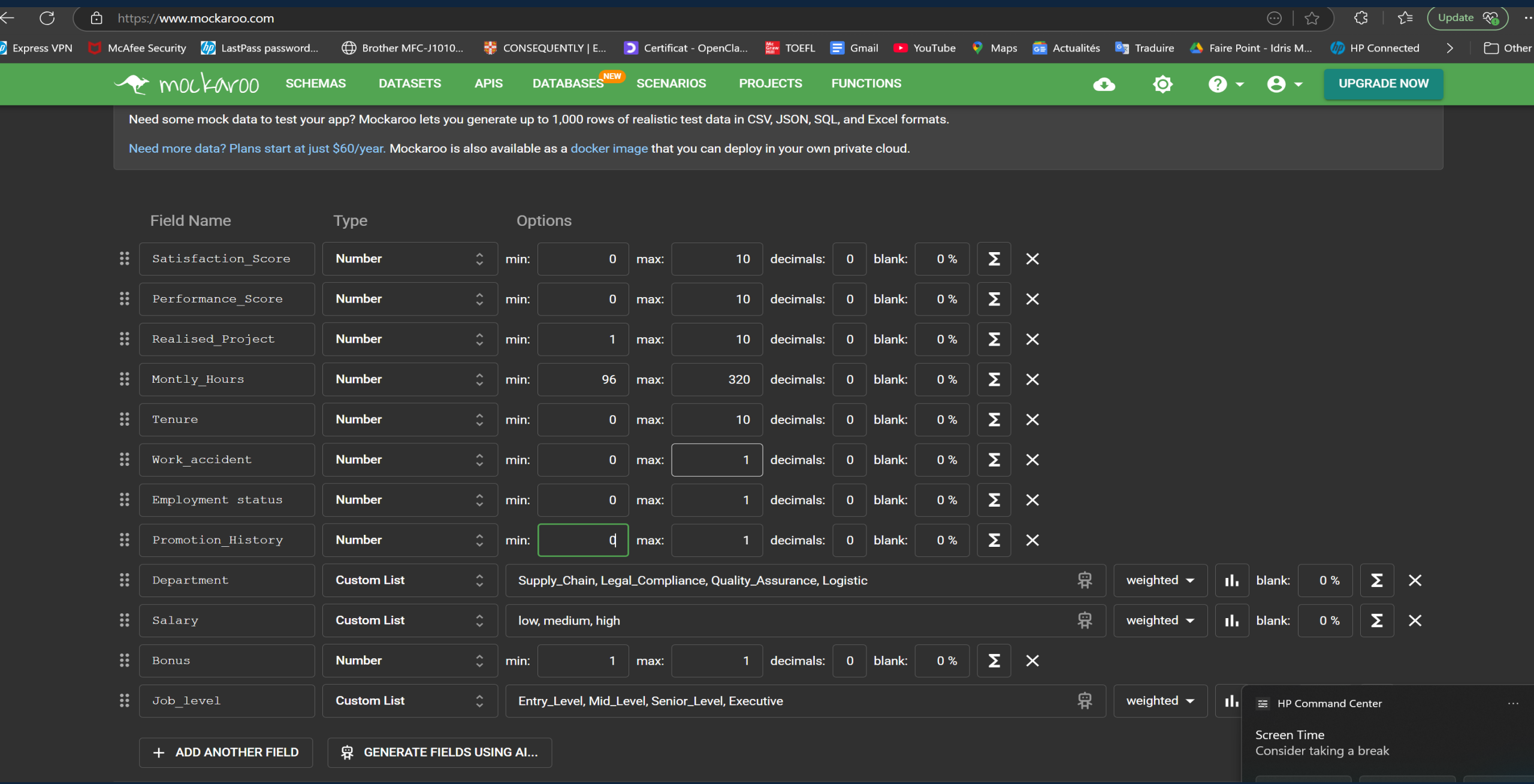
Exploratory Data Analysis

Employee turnover analysis

Dataset description

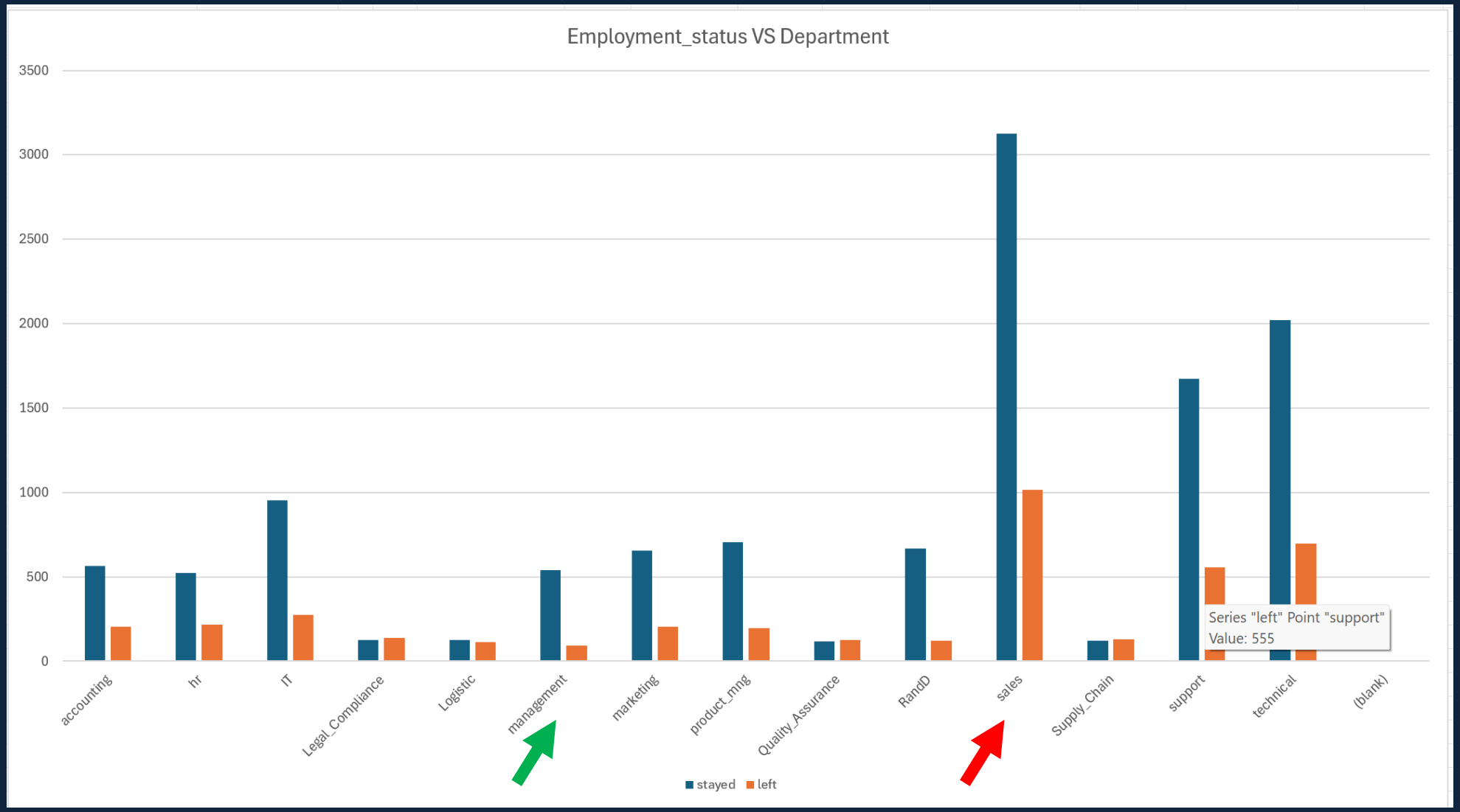
Satisfaction_Score	Realised_Project	Performance_Score	Work_accident	Promotion_History	Department
Numerical (0-10)	Numerical (1-10)	Numerical (0-10)	Numerical (1-0) True= 1 False = 0	Numerical (1-0) True= 1 False = 0	Categorical

Montly_Hours	Tenure	Employment_status	Job_level	Bonus	Salary
Numerical (96-320)	Numerical (0-10)	Numerical (0-1) True= 1 False = 0	Categorical	Numerical (0-1) True= 1 False = 0	Categorical

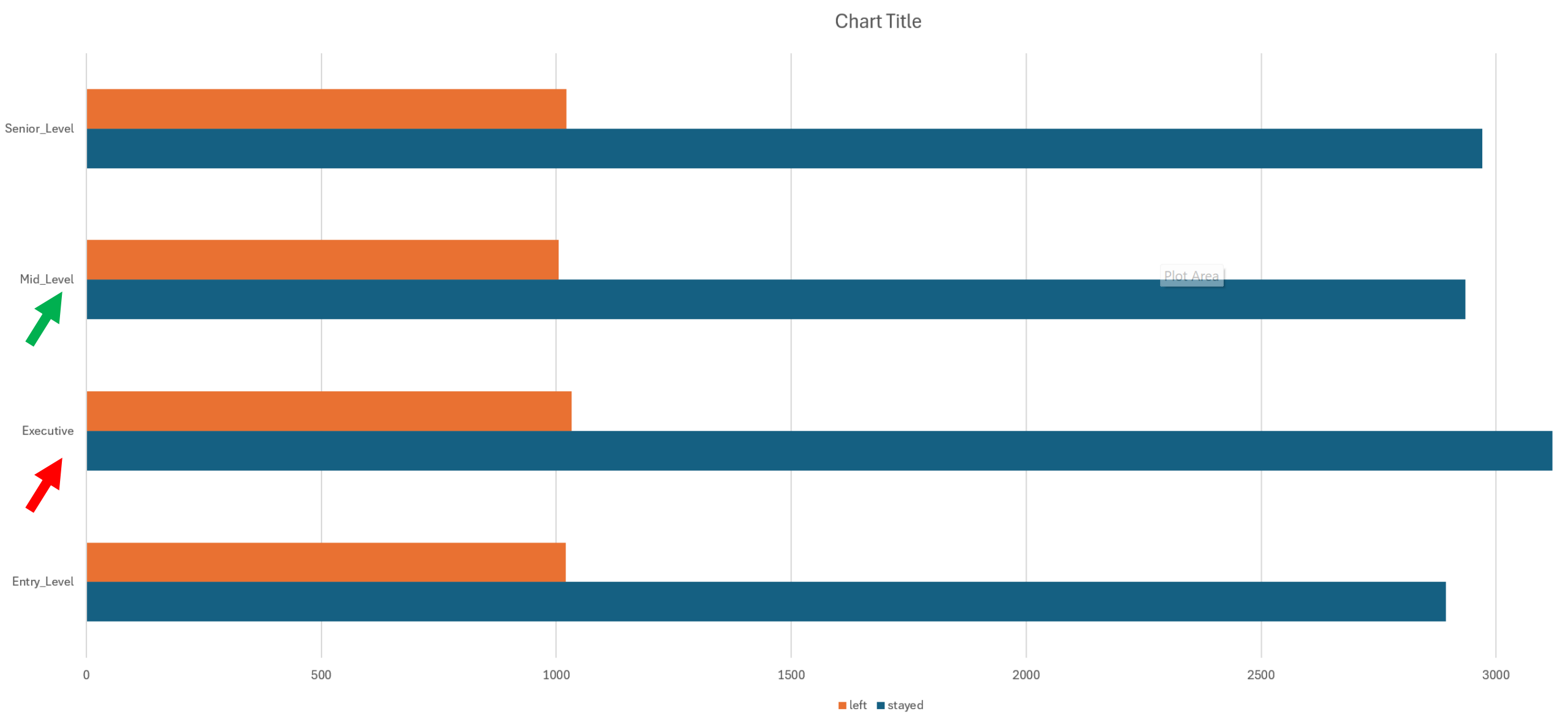


Employee turnover analysis

Bar chart



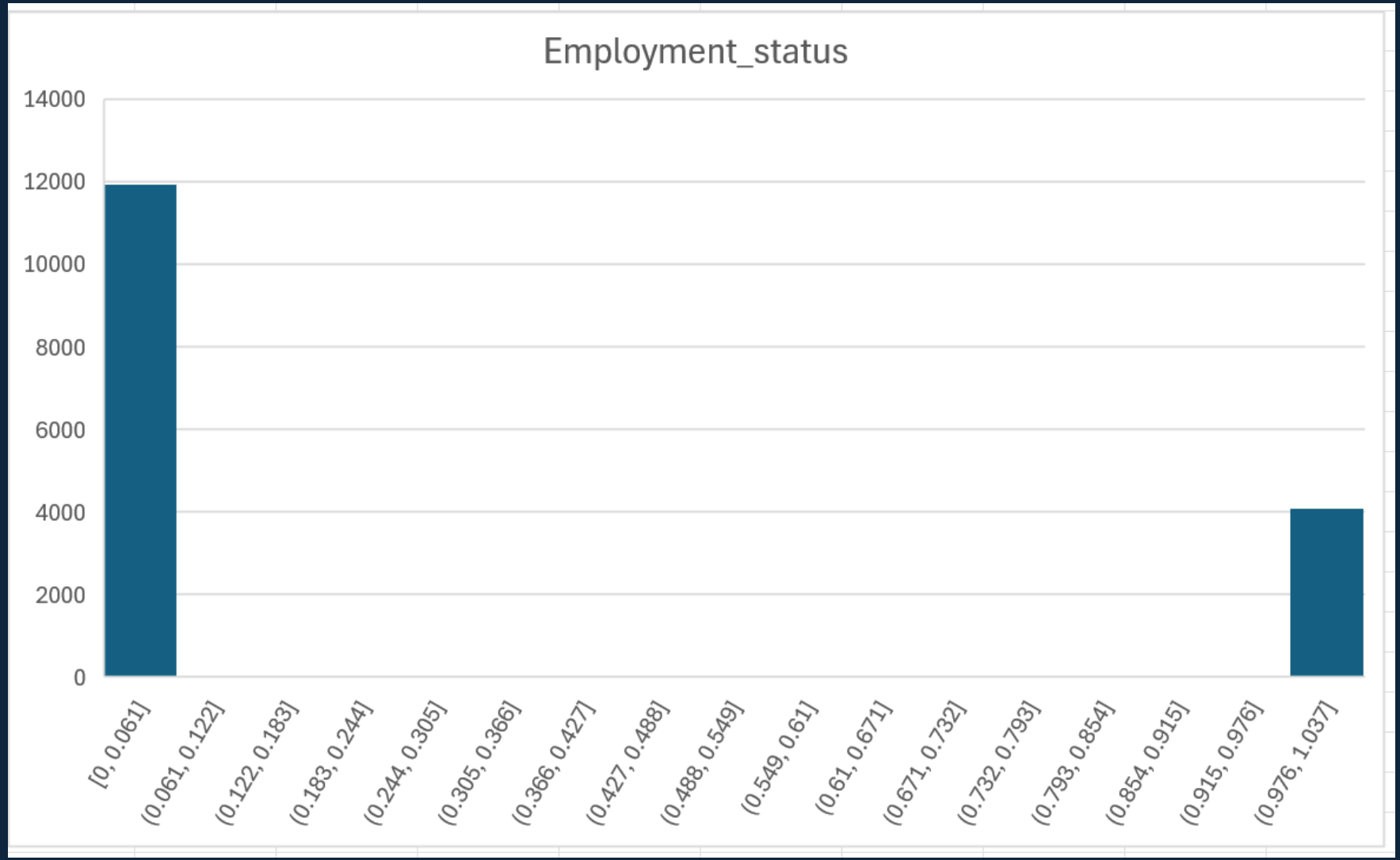
Employee turnover analysis



Employee turnover analysis

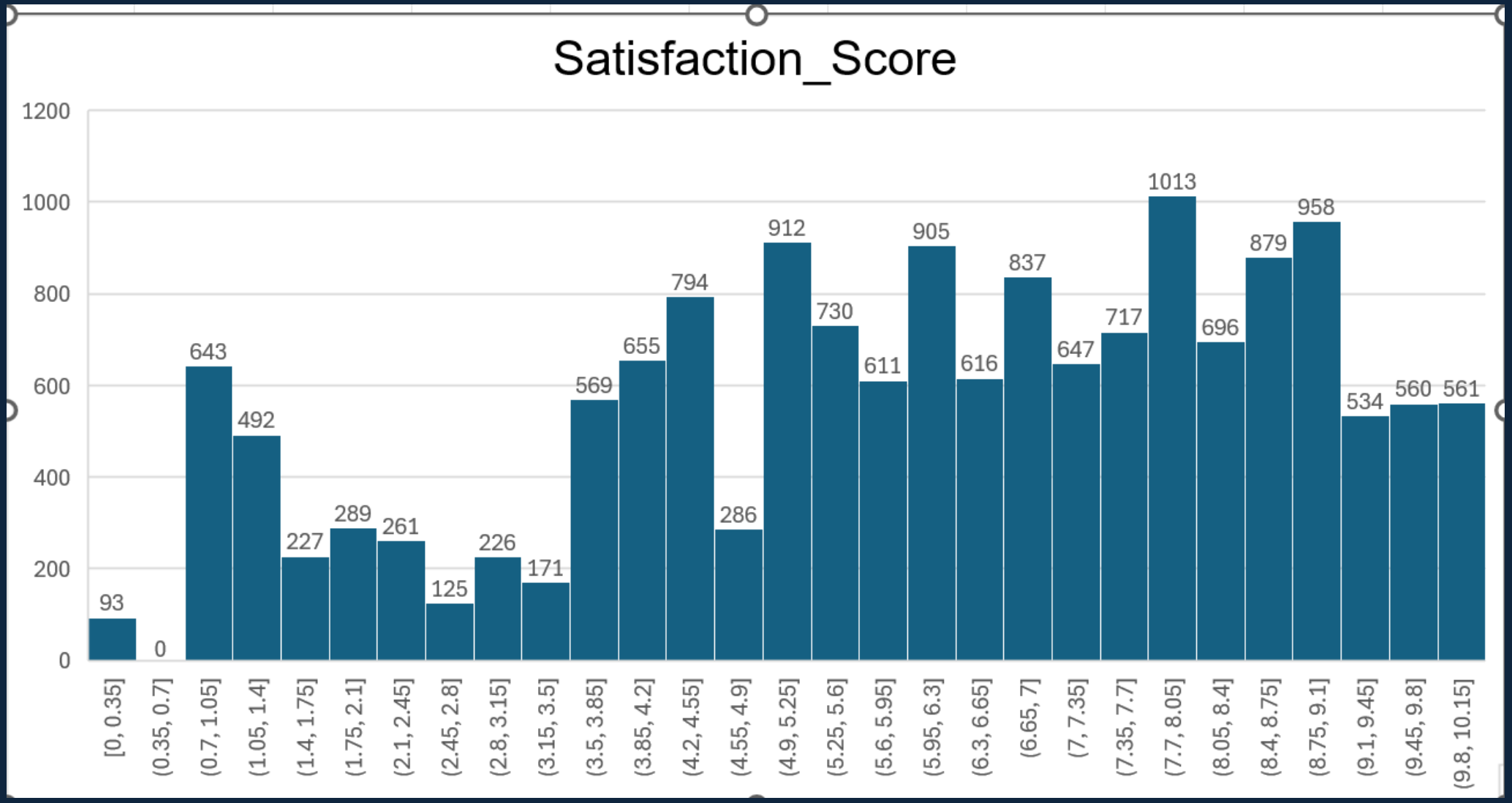
Employee turnover analysis

Histogram

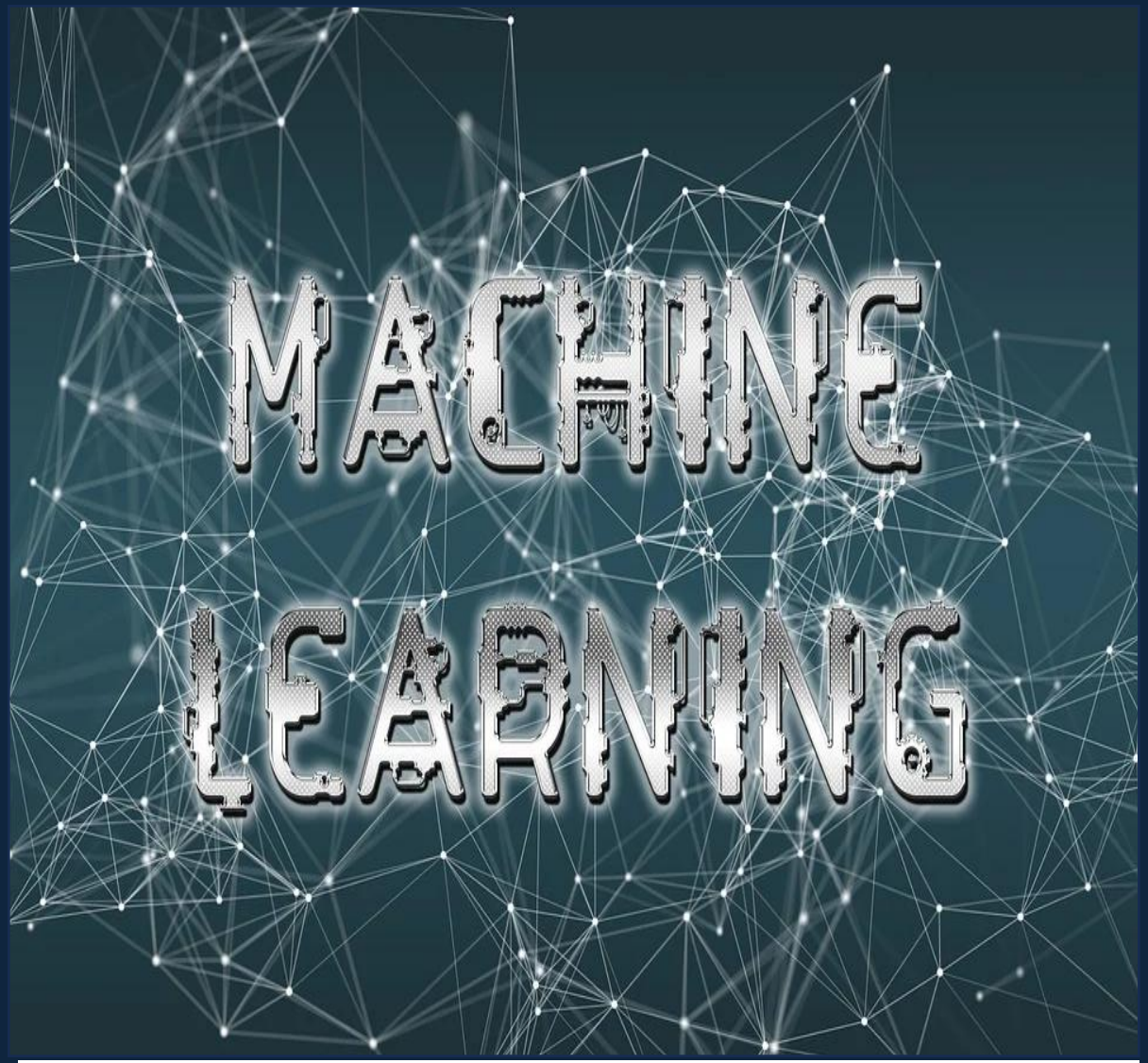


Employee turnover analysis

Histogram



Employee turnover analysis

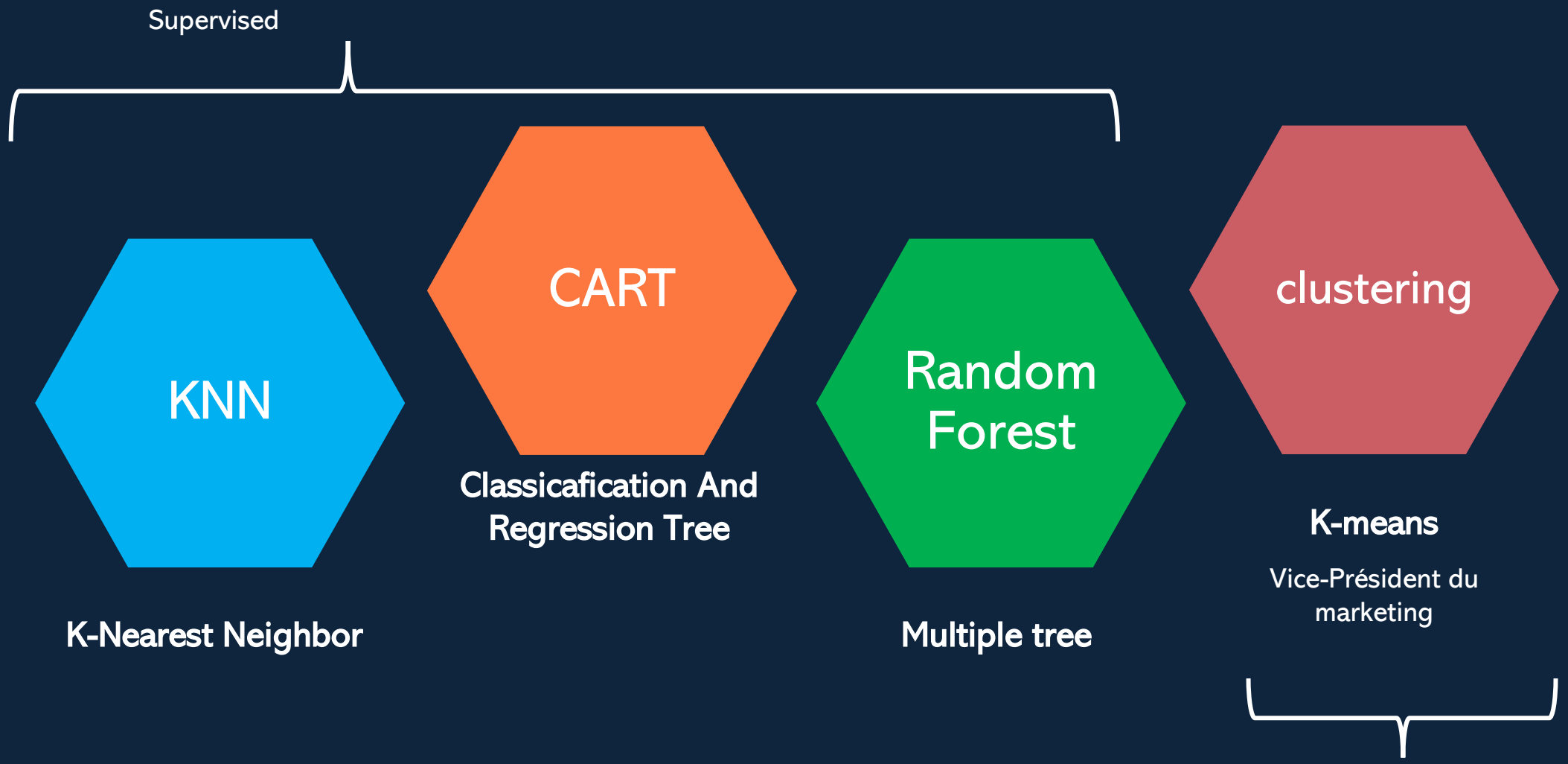


Source : https://cdn.pixabay.com/photo/2019/04/15/12/09/machine-learning-4129175_960_720.jpg

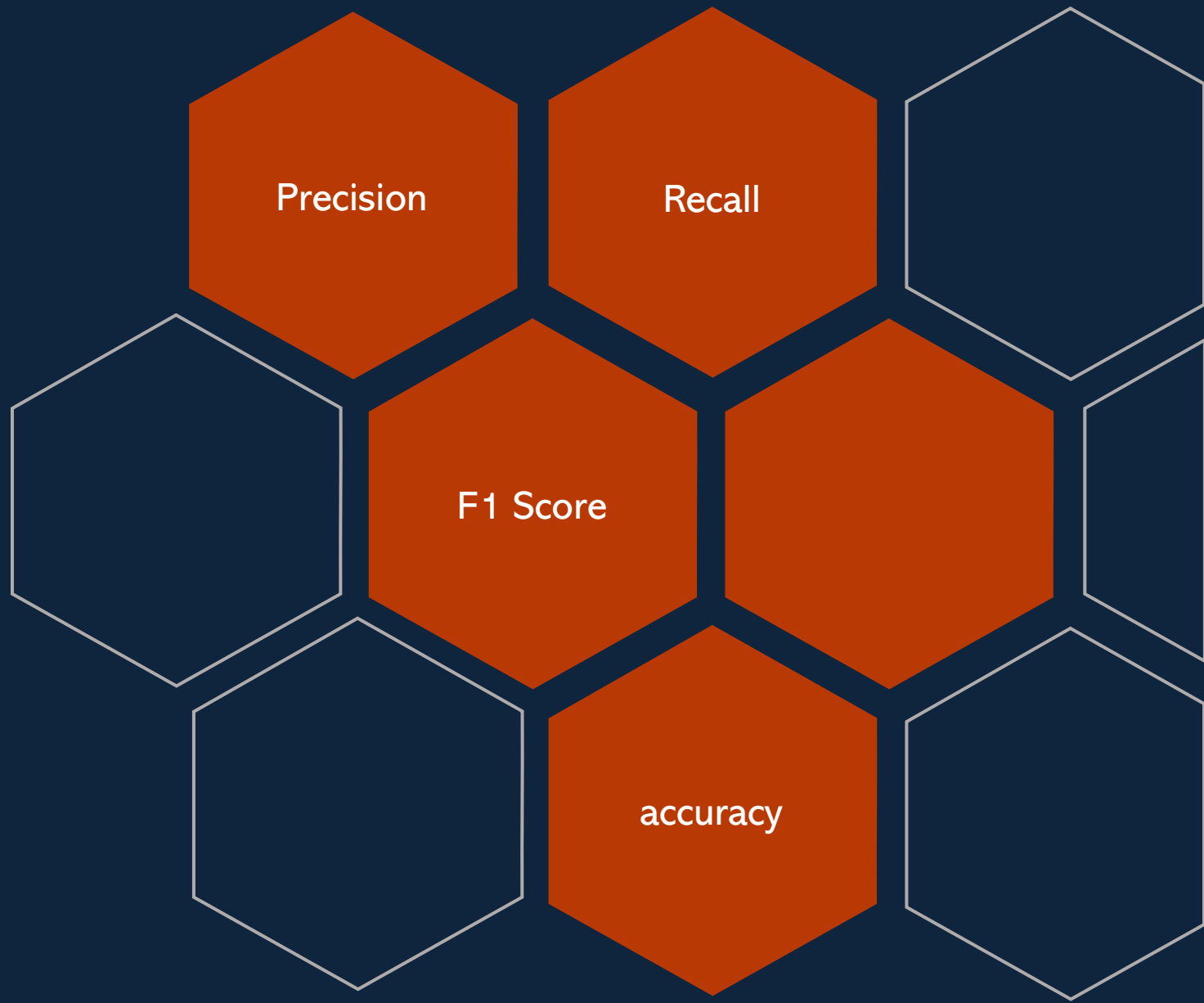
Modeling

Employee turnover analysis

Model


Supervised
KNN
CART
Random Forest
clustering
Classicafication And Regression Tree
K-means
Vice-Président du marketing
K-Nearest Neighbor
Multiple tree
Unsupervised

Model Performance



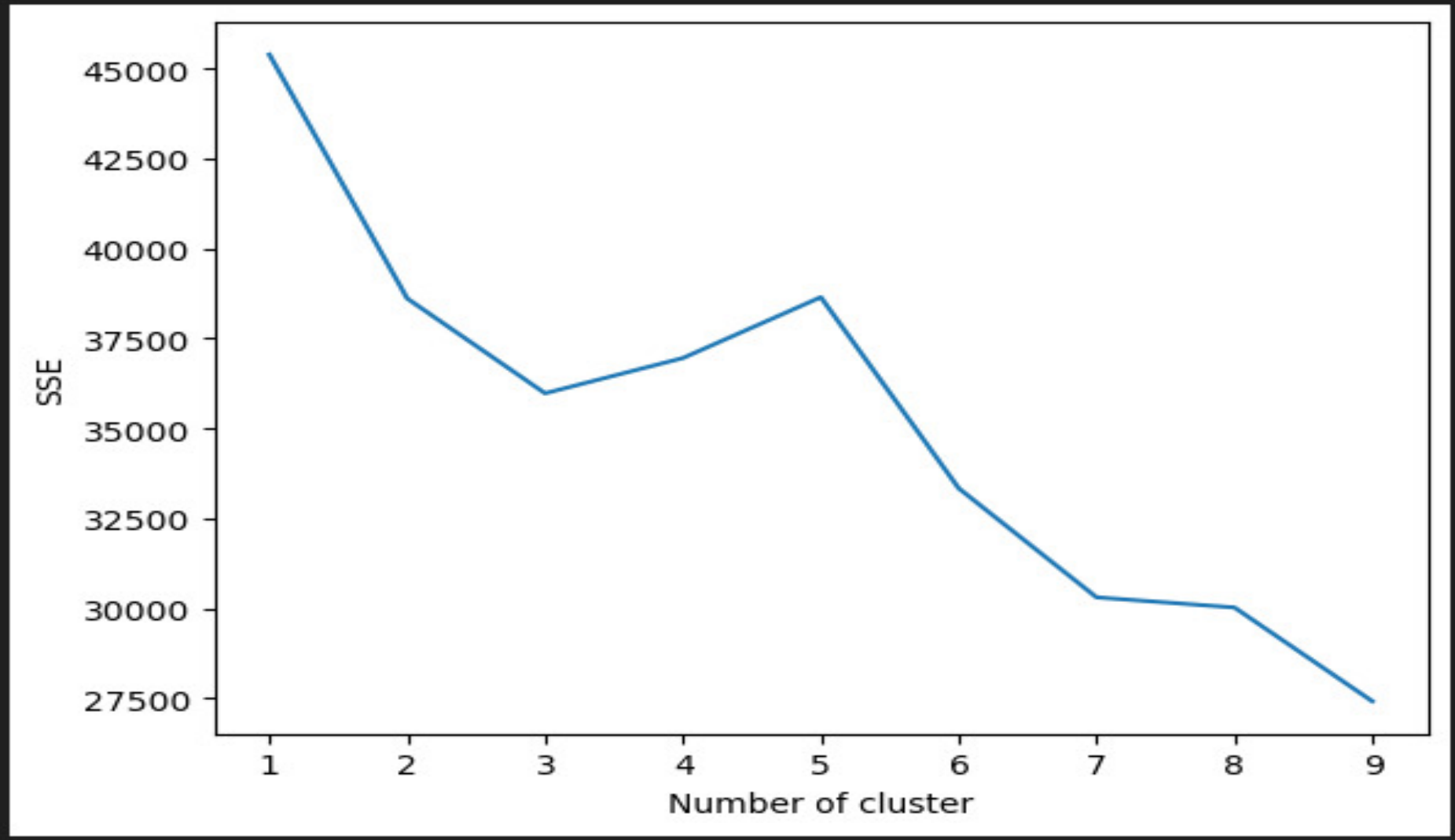
Implementation

Model measurement - test = 30%

model	Precision [0 – 1]	Recall [0 – 1]	F1-score [0 – 1]	accuracy	Support [0 – 1]
KNN (K = 2)	0.95 – 0.85	0.95 – 0.84	0.95 – 0.84	0.92	3593 - 1207
CART	0.96 – 0.88	0.96 – 0.90	0.96 – 0.89	0.94	3593 - 1207
RANDOM FOREST	0.95 – 0.94	0.98 – 0.86	0.97 – 0.90	0.95	3593 - 1207
					4800
K_MEAN					

Out of all model we tested so far, random forest has the highest accuracy score about 95% and F1 scrore which is better than KNN, CART. So, we will choose random forest for our project.

Cluster	0	1
Actual		
0	4678	7241
1	1529	2551



Summary

According to the model, the key feature which determine an employee turnover in our use case is SATISFACTION. From that we understand that the chance of an employee to leave is unlikely if he his satisfied with his job.

references

- ✓ Mockaroo - Random Data Generator and API Mocking Tool | JSON / CSV / SQL / Excel
- ✓ Quit Work Stock Photos, Images and Backgrounds for Free Download
- ✓ Employee Turnover Data Analysis: 8 Tips for Success – AIHR
- ✓ DeepSeek - Into the Unknown
- ✓ https://cdn.computerhoy.com/sites/navi.axelspringer.es/public/media/image/2023/04/raspberry-lanza-editor-codigo-aprender-python-lenguaje-ia-3008158.jpg
- ✓ https://d3srxiunz7lgh6.cloudfront.net/nhrdh8t9d2rox4ezsizzhyvm1y5l

Merci

Idris Mahamat