EMPLOYEE ATTRITION

E-NUMPY (BATCH 19A)

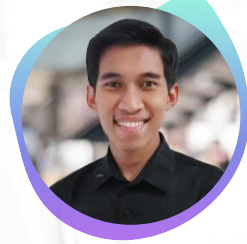# THE TEAMS

Ilham
Taufiqurrohim

M Chosasih
Mahendra

Putri Maylita

Rama
Kuswidyawan

Samuel Akwila

OUTLINE

PROJECT
INTRO

PROJECT **DESCRIPTION**

Background, Context and Problem.

- In one year, the number of employees of Kagg Company decreased from 1470 → 1233 (reduced by **237** or **16%**)

- The **SHRM Human Capital Benchmarking** report found that the average employee turnover rate in 2017 was 18%, and less than 50% of organizations have a succession plan. Organizations should target 10% for employee attrition.

- **Marc Prosser (CEO and co-founder of Choosing Therapy)**, said that, if the employee turnover rate is more than 15% per year, the company has a high employee turnover rate.

Therefore, the problems that this project will solve are:

> 1. What are the factors that drive employees to attrite?

> 2. What steps can be taken to prevent employees from attrition?

PROJECT **DESCRIPTION**

Goals

- Finding the factors behind employees to attrice from the data provided using predictive modeling.

- Provide recommendations to prevent employees from attrition.

Business Metrics

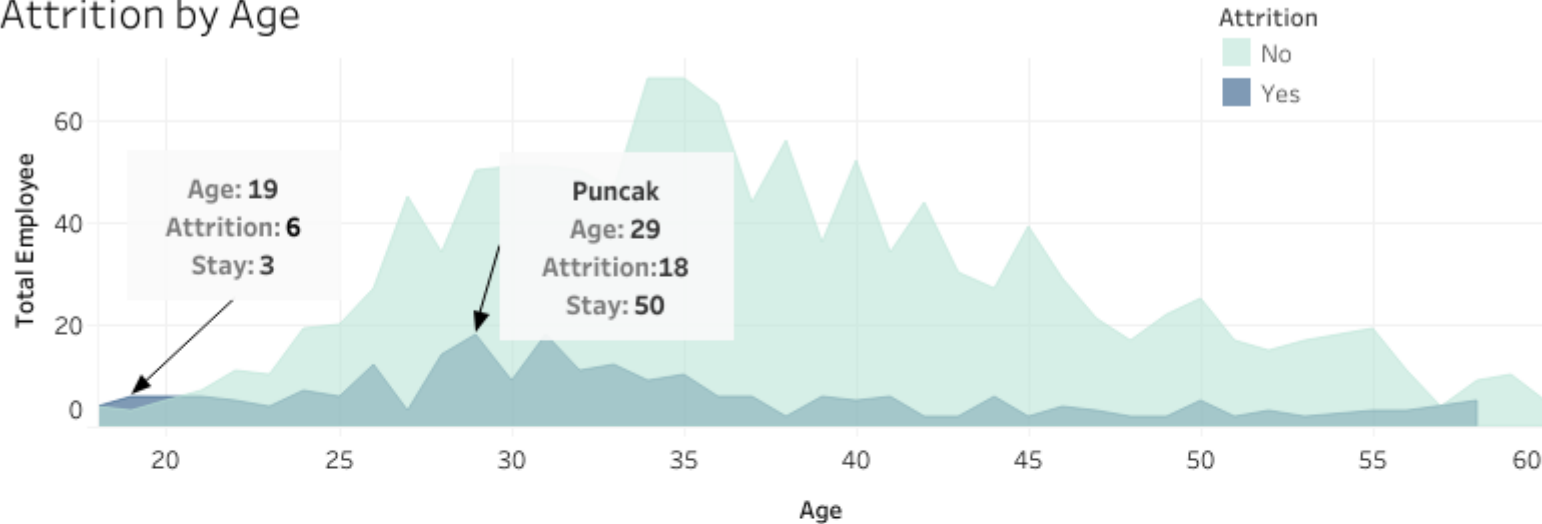**Employee Attrition Rate**
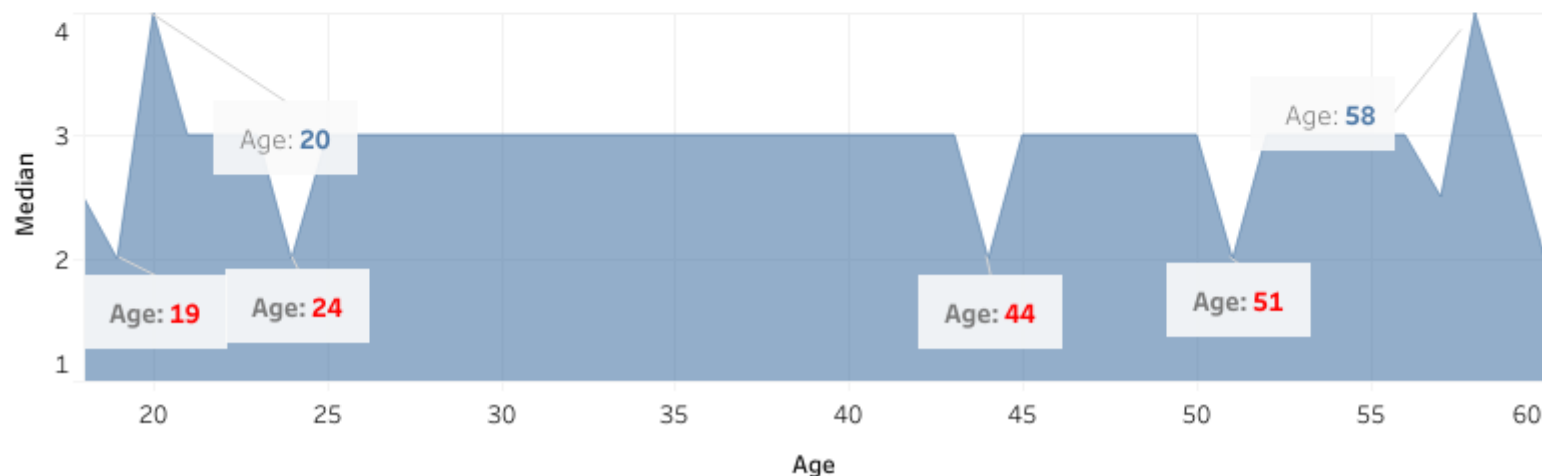
METHODS

DATA **COLLECTION**

Employee Attrition
.CSV

1470 Rows
35 Columns

EXPLORATORY **DATA ANALYSIS**


Attrition by Age

- The number of attrition employees at the age of 19 years is more than those who stayed.

- The level of environmental satisfaction at the age of 19 years is low.

- At the age of 20 years showed a high rate of environment satisfaction.

- Perceptions of environmental satisfaction seem to fluctuate in the young age range (18–25 years).

EXPLORATORY **DATA ANALYSIS**

## **Environment Satisfaction** by Monthly Income

- Semakin rendah median gaji bulanan justru kepuasan lingkungan kerja semakin meningkat.

- Selisih median gaji bulanan antar persepsi kepuasan lingkungan kerja kurang lebih $100.

- Semakin tinggi gaji tidak menjamin karyawan menjadi lebih puas dengan lingkungan perusahaan.

Environment Satisfaction

Environment Satisfaction
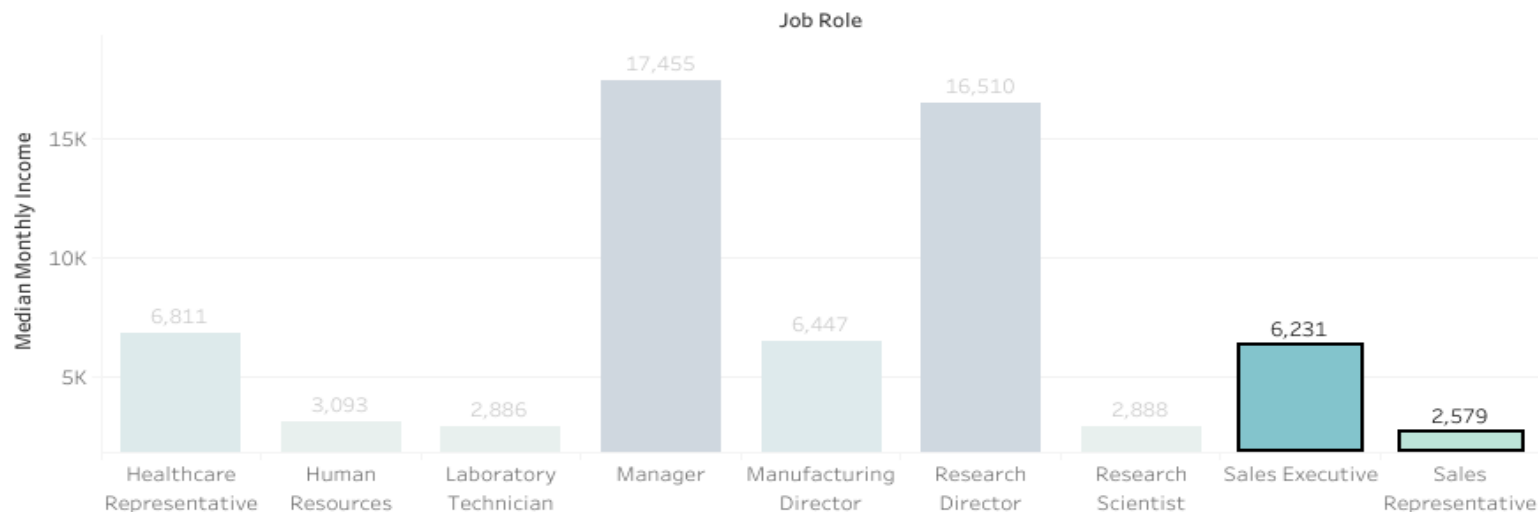1
2
3
4
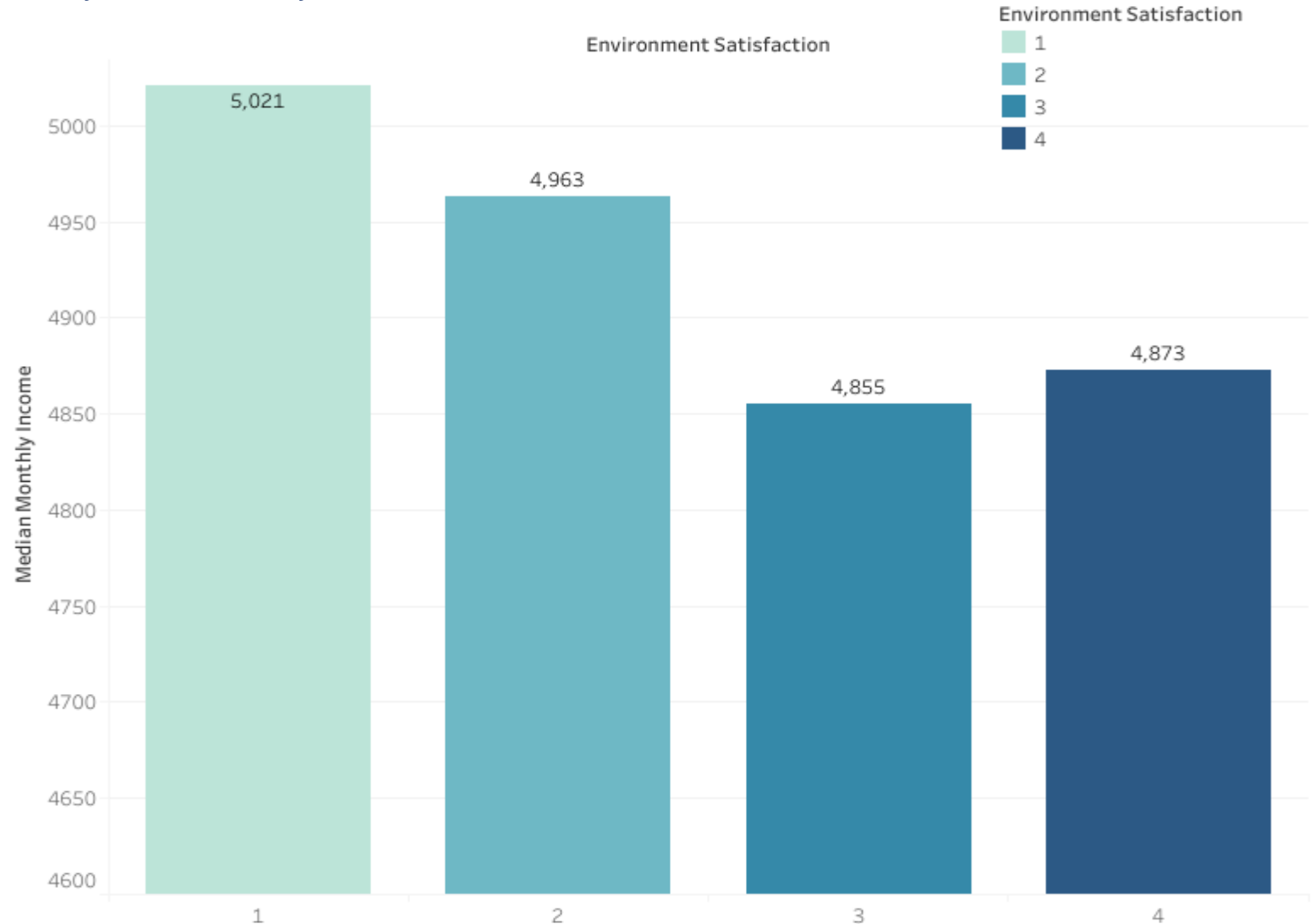
EXPLORATORY **DATA ANALYSIS**

Over Time by Department, Job Role
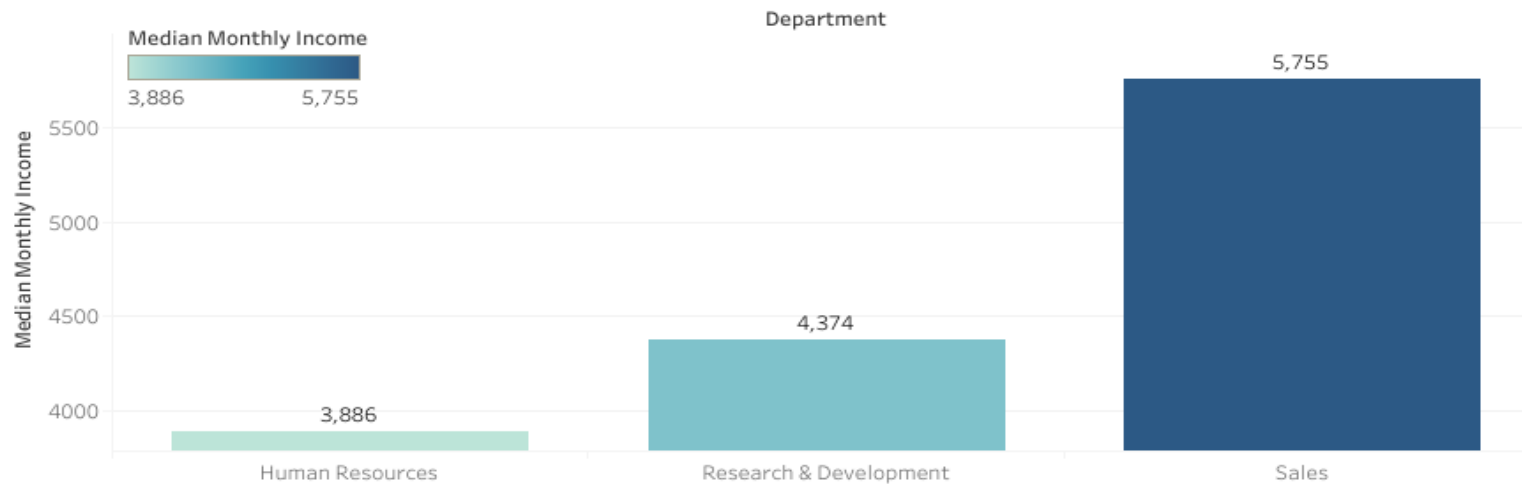
- Departemen Sales dan R&D memiliki persentase Overtime paling tinggi, terutama Sales Representative dan Research Director.

- Departemen Human Resource sangat jarang mendapatkan Overtime.
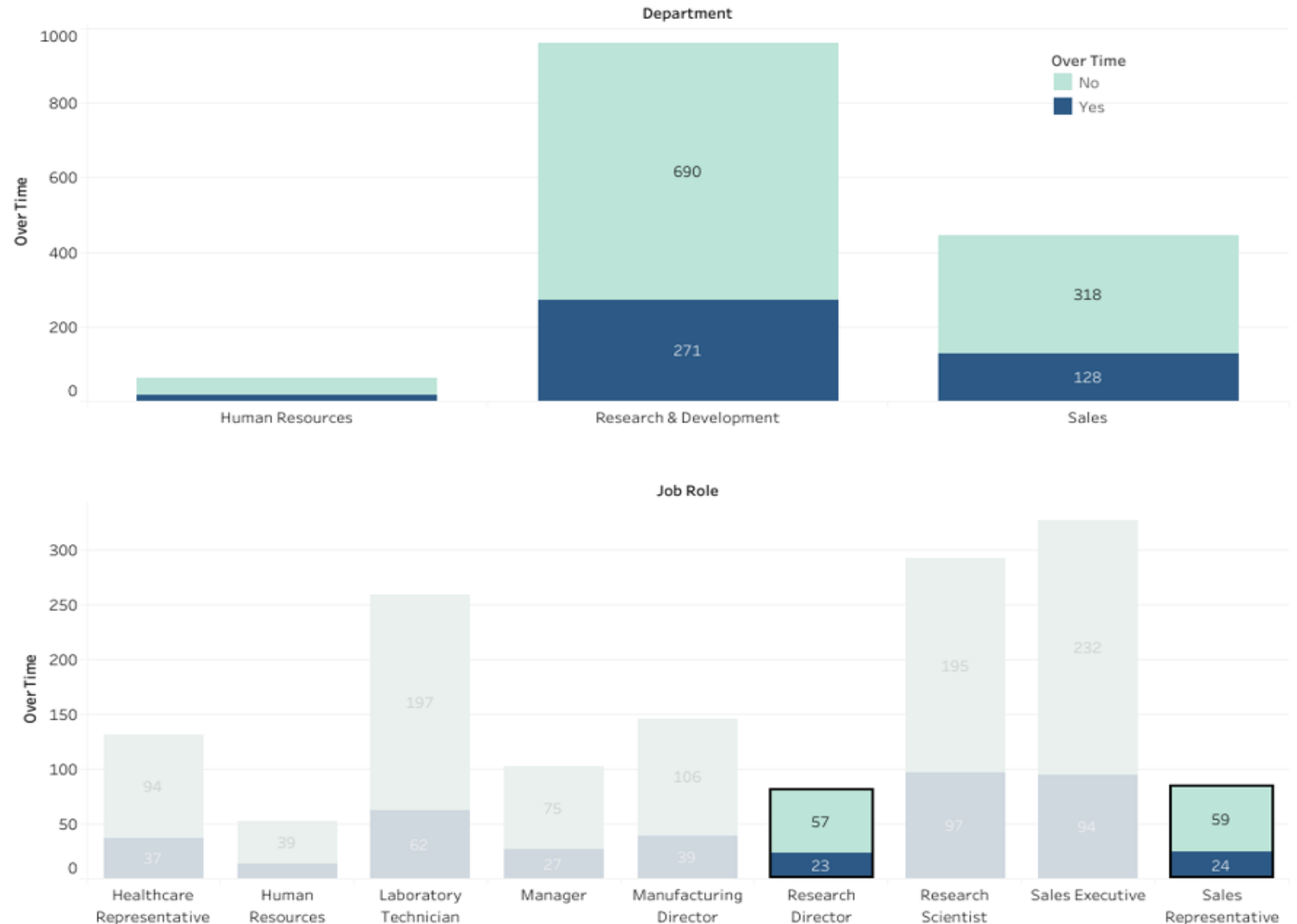
EXPLORATORY **DATA ANALYSIS**

Monthly Income by Department, Job Role

- Departemen Sales memiliki median gaji bulanan tertinggi.

- Terdapat ketimpangan yang signifikan pada Departemen Sales.

EXPLORATORY **DATA ANALYSIS**
Business Travel by Department, Job Role

- Departemen Human Resource sangat jarang mendapat kesempatan perjalanan kerja.

- Departemen R&D cukup sering mendapat kesempatan perjalanan kerja.

- Departemen Sales pada Job Role Sales Representative merupakan posisi yang sangat mobile.
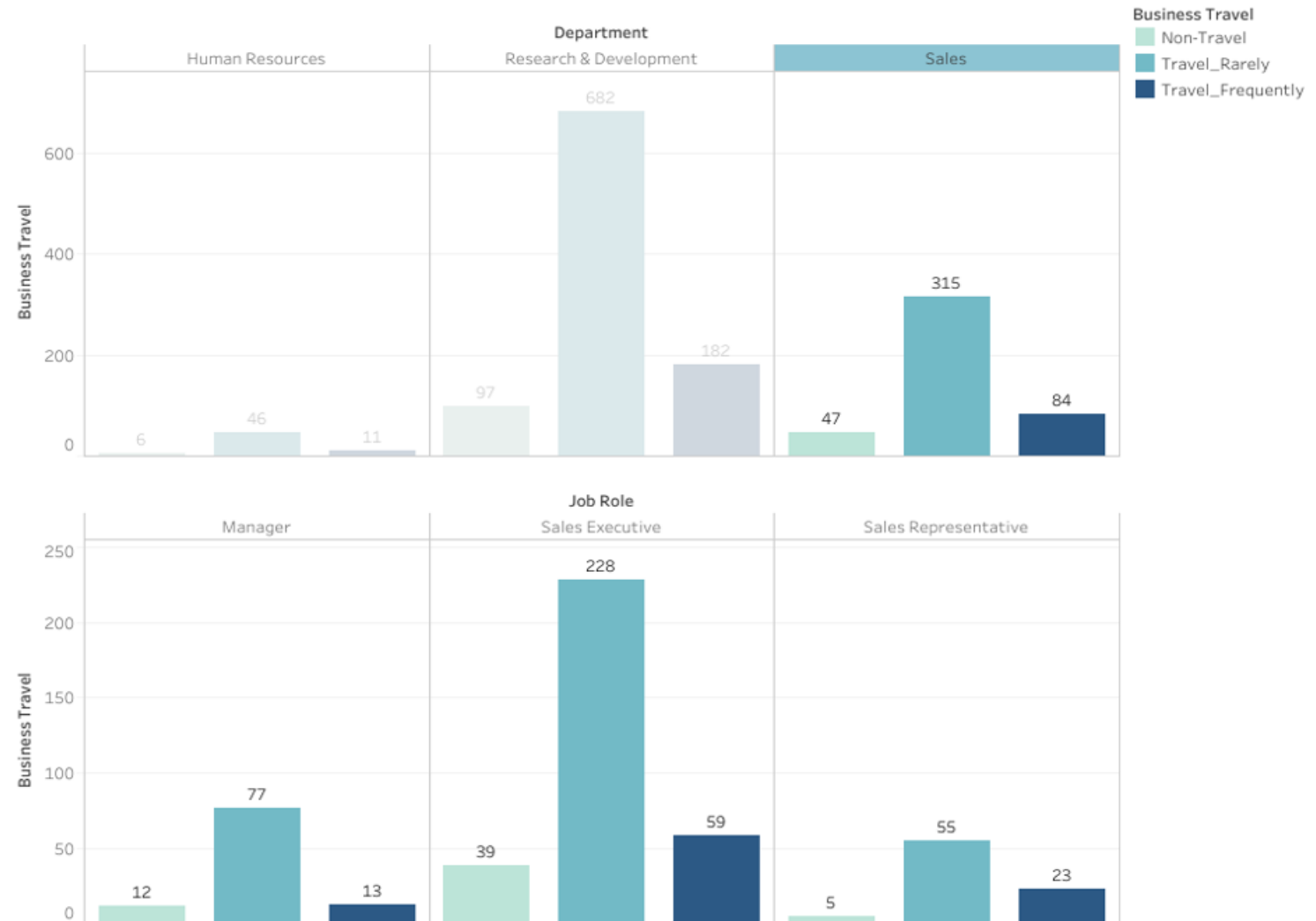
EXPLORATORY **DATA ANALYSIS**

Attrition by Department, Job Role
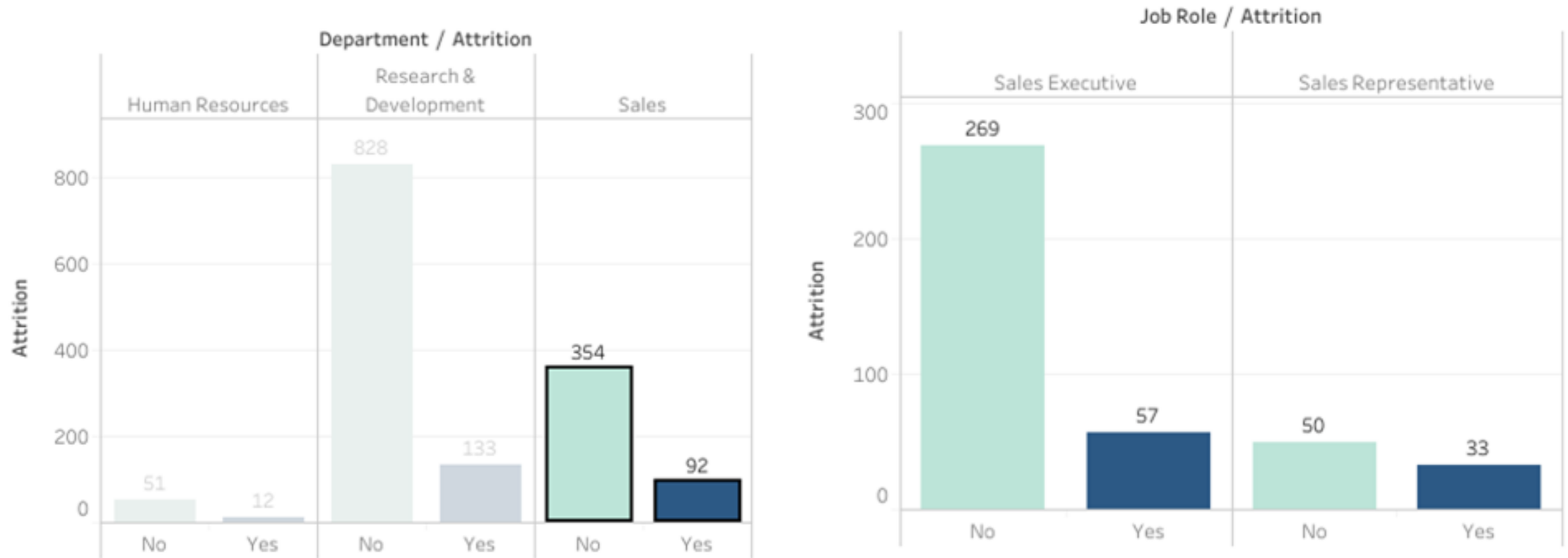
- Departemen Sales perlu mendapatkan perhatian karena sering mendapatkan Over time, terdapat kesenjangan gaji yang besar dan sangat mobile.

- Persentase Atrisi tertinggi terjadi pada Departement Sales dengan Job Role Sales Representative.

DATA **PREPROCESSING**

Handle Irrelevant Value

Menghapus feature:
- EmployeeCount
- StandardHours
- Over18
- EmployeeNumber

Handle Outlier

- Outlier dihapus menggunakan **Z score > 2.**

Feature Engineer

- **Feature Encoding**
Label Encoding : Attrition, OverTime, Gender, Business Travel

- **Feature Selection:**
Menghapus fitur yang redundan dan menghapus fitur yang korelasinya rendah ke Attrition

DATA **PREPROCESSING**

Feature Transformation

1. **Log Transform Feature:**
   - MonthlyIncome
   - TotalWorkingYears
   - YearsAtCompany

2. **Normalisasi**
Melakukan normalisasi numeric data menggunakan StandardScaler,

Modeling

- Menggunakan algoritma **Logistic Regression** dengan **Random Oversampling.**

- Metric yang digunakan **Recall dan ROC–AUC**

Evaluasi

Melakukan evaluasi model dan Feature Importance.

PREDICTIVE **MODELING**
Machine Learning

| Algorithm	Recall (Train)	Recall (Test)	ROC-AUC (Train)	ROC-AUC (Test)	Sampling Type
Logistic Regression	0.50	0.59	0.87	0.82	Normal
	0.82	**0.80**	**0.88**	**0.82**	**Oversampling (Best Fit)**
	0.83	0.84	0.89	0.80	SMOTE
Random Forest	0.71	0.45	1.00	0.81	Normal
	0.99	0.66	1.00	0.82	Oversampling
	0.98	0.66	1.00	0.79	SMOTE
XGBoost	1.00	0.55	1.00	0.78	Normal
	1.00	0.64	1.00	0.80	Oversampling
	1.00	0.66	1.00	0.79	SMOTE

Random Oversampling, SMOTE
Metode dalam mengubah sample minoritas menjadi sebanyak sample mayoritas. Perbandingan 1 : 1.

- **Tuning Hyperparameter**
- **Metric 1 : Recall**
- **Metric 2 : ROC – AUC**

PREDICTIVE **MODELING**
Confusion Matrix
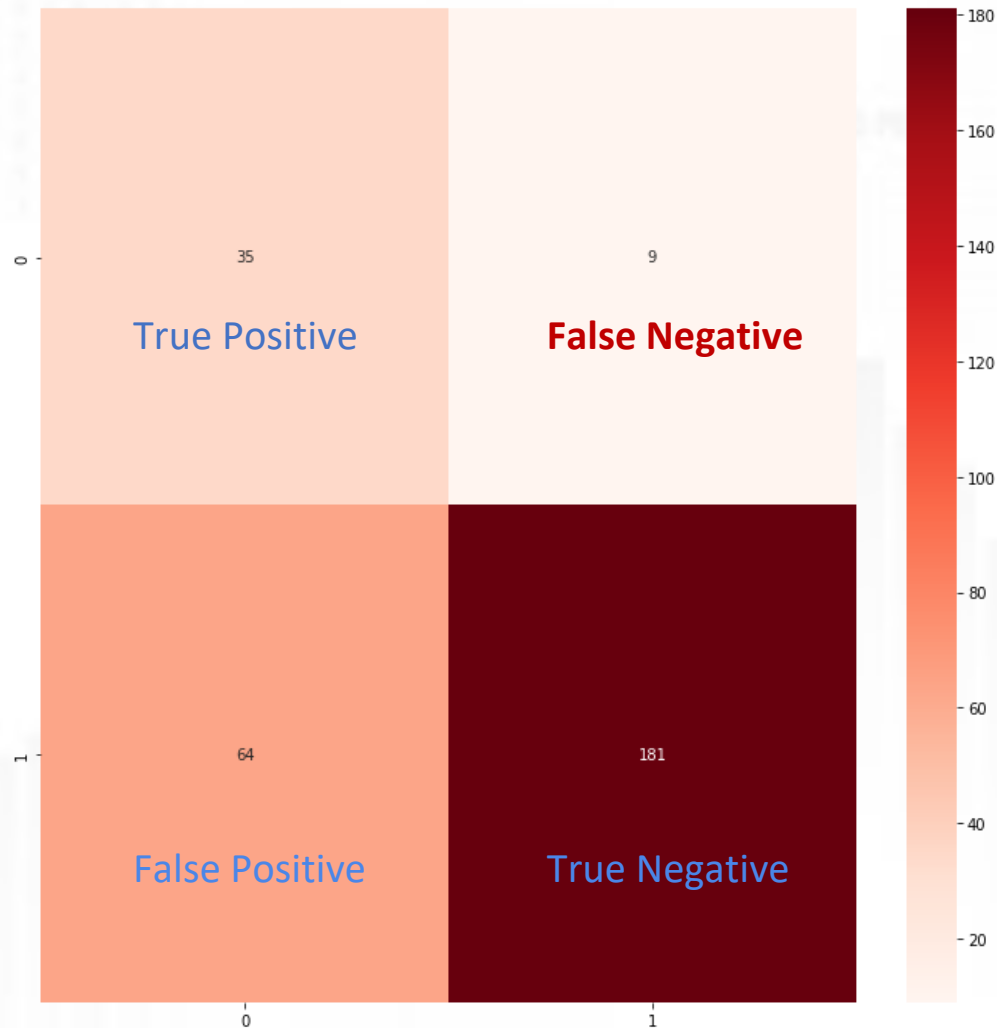
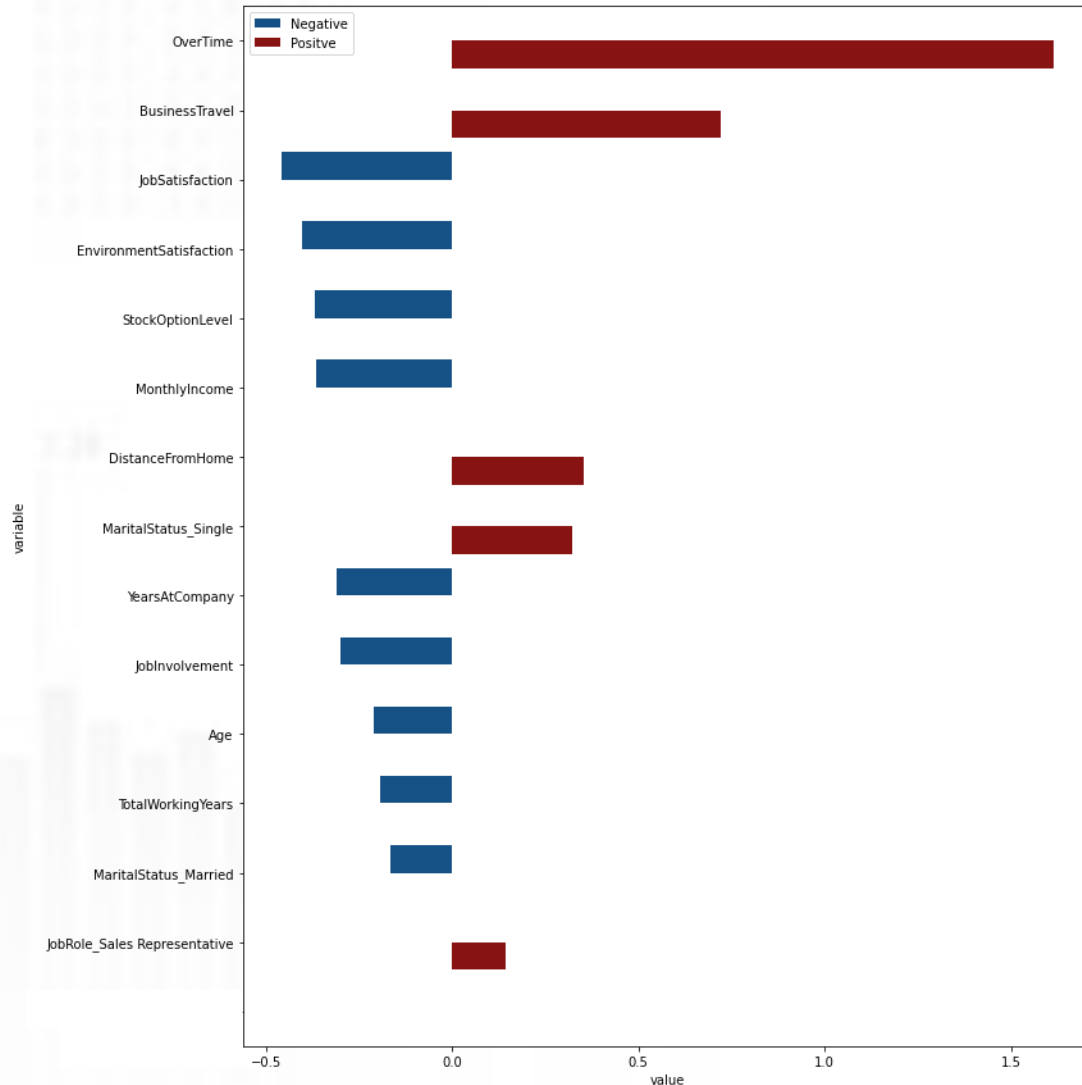**Berfokus pada predictive model yang paling bagus dalam menurunkan false negative.**

False Negative: Hasil prediksi pegawai bertahan di perusahaan tetapi faktanya pegawai keluar dari perusahaan,

RESULTS & RECOMMENDATIONS

RESULT

Features Importance

Faktor berpengaruh positif kepada Atrisi
- **Overtime**
- **Business Travel**
- **Distance From Home**
- **Single**

Faktor berpengaruh negatif kepada Atrisi
- Job Satisfaction
- Environment Satisfaction
- Stock Option Level
- Monthly Income
- Years At Company
- Job Involvement
- Age
- Total Working Years
- Married

Sales Representative merupakan satu-satunya posisi yang memiliki pengaruh kepada Atrisi.

RECOMMENDATIONS

OVERTIME

- Evaluasi penyebab Overtime
- Membuat kebijakan Overtime
- Mengadakan program Cross-Training
- Efisiensi Kerja

JOB INVOLVEMENT & SATISFACTION

- Membangun budaya kerja positif
- Komunikasi dan kepercayaan pada semua level
- Memberikan penghargaan secara berkala

RECOMMENDATIONS

ENVIRONMENT SATISFACTION

- Mengadakan kegiatan keakraban internal.
- Program mentoring oleh senior kepada karyawan muda.

DISTANCE FROM HOME

- Memberikan opsi Living Sponsorship.

MONTHLY INCOME

- Mengevaluasi proporsi pendapatan dan beban kerja.
- Memberikan opsi bonus saham bagi karyawan yang high-perform.

IMPLEMENTATION

IMPLEMENTATION

Model ini dapat diimplementasikan pada evaluasi rutin karyawan pada periode waktu tertentu. Sehingga perusahaan dapat memprediksi karyawan yang akan attrition dan melakukan pencegahan dengan melakukan treatment tertentu.

INTEGRASI MODEL DENGAN APLIKASI/WEBSITE INTERNAL → **DATA EVALUASI PERIODIK KARYAWAN** → **PERSONALISASI KARYAWAN YANG POTENSI ATRISI**

SIMULATIONS

SIMULATION

50%

- Overtime = Yes
- JobSatisfaction = 1 , 2
- EnvironmentSatisfaction = 1, 2

→

- Overtime = No
- JobSatisfaction = 4
- EnvironmentSatisfaction = 4

237
(16%)

81
(5.5%)