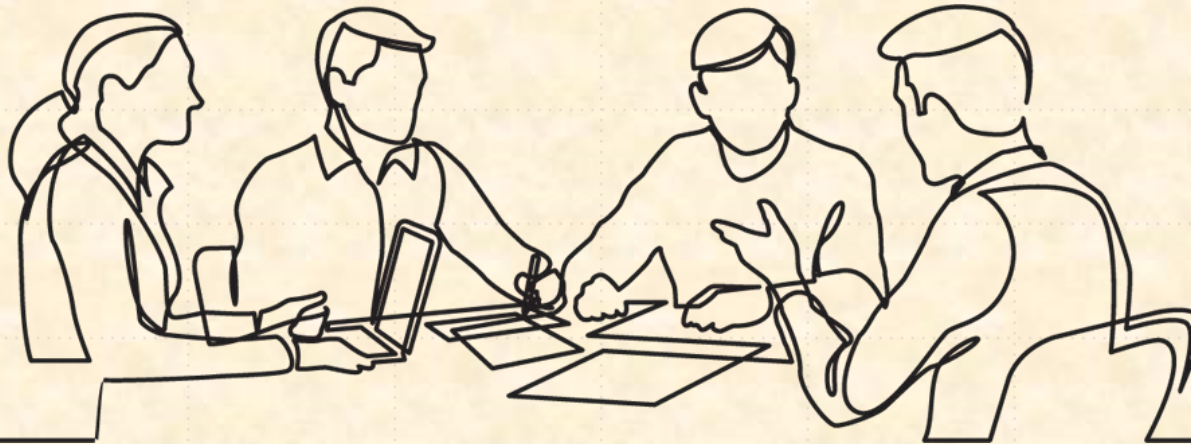


HR Attrition Analysis

By Money Oriented Team



Meet The Duo



Attariq Muhammad Kasfilla

Data Scientist



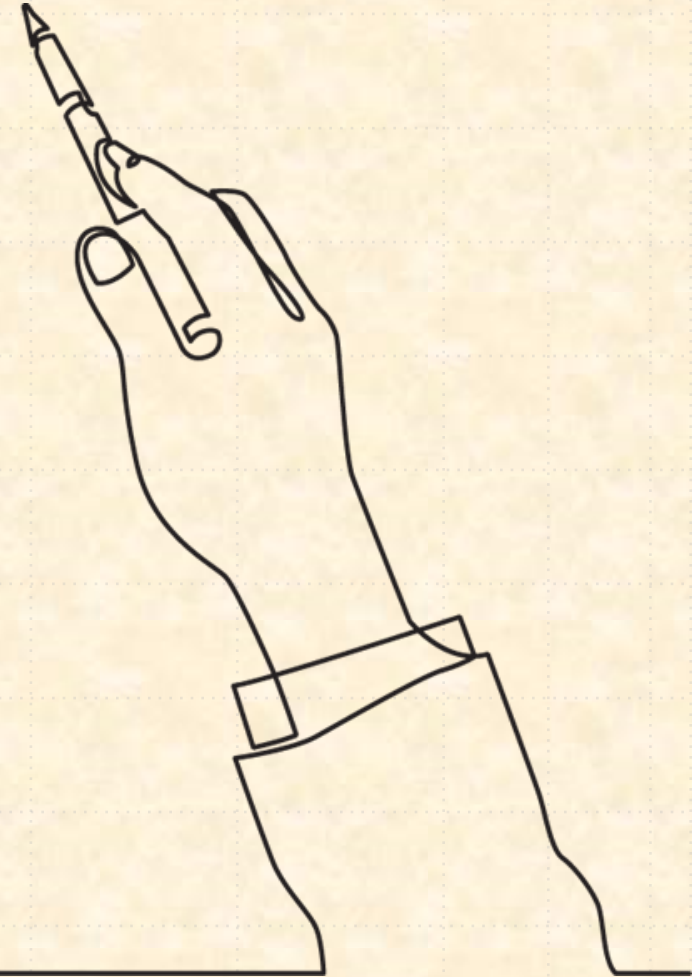
Nizar Mochammad Yusuf

People Analytics



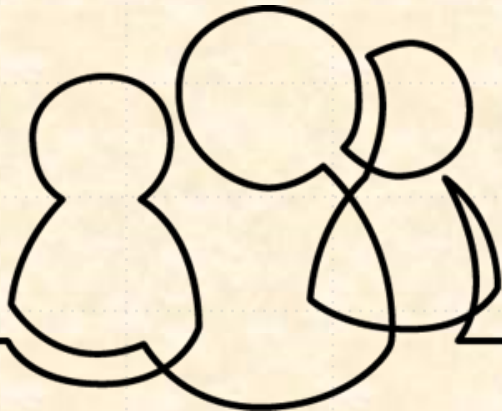
Table Of Contents

- Project Background
- Data Preparation
- Insight
- Modeling
- Business Recommendation



Who We Are

Tim HR Analis pada PT YKGY yang bertanggung jawab untuk memberikan penilaian mengenai bagaimana tingkat kepuasan dari setiap karyawan dengan mengidentifikasi faktor-faktor apa saja yang menjadi komponen dari kepuasan karyawan tersebut sehingga berdampak terhadap loyalitas.



Project Background



Problem Statement

- Perusahaan kita memiliki (*attrition rate*) sebesar 16%
- *TurnOver Rate* 2 tahun terakhir sebesar 34%-36%



Goal

- Mendeteksi karyawan/staff yang ingin mengundurkan diri
- Menurunkan tingkat karyawan yang ingin mengundurkan diri
- Meningkatkan loyalitas karyawan pada PT YGKY



Objective

- Membuat system untuk memprediksi karyawan yang berpotensi untuk *attrition*
- Implementasi kebijakan baru, untuk meningkatkan loyalitas karyawan

Menurut Gillis (1994), turnover karyawan dikatakan normal berkisar antara 5-10% per tahun, dan tergolong tinggi apabila lebih dari 10% per tahunnya

Dataset



Employee Profile

- Age
- Education
- Gender
- MaritalStatus
- DistanceFromHome
- NumCompaniesWorked
- MonthlyIncome
- NumCompaniesWorked
- TotalWorkingYears
- WorkLifeBalance
- YearsAtCompany
- YearsInCurrentRole
- YearsSinceLastPromotion
- YearsWithCurrentManager



Feature

- Department
- DailyRate
- EnvironmentSatisfaction
- HourlyRate
- JobInvolment
- JobLevel
- JobRole
- JobSatisfaction
- OverTime
- PerformanceRating
- RelationshipSatisfaction
- StockOptionLevel



Target

Attrition

Business Metrics

- Attrition Rate
- Turn Over Rate

Missing Value

```
Age 0
BusinessTravel 0
DailyRate 0
Department 0
DistanceFromHome 0
Education 0
EducationField 0
EnvironmentSatisfaction 0
Gender 0
HourlyRate 0
JobInvolvement 0
JobLevel 0
JobRole 0
JobSatisfaction 0
MaritalStatus 0
MonthlyIncome 0
MonthlyRate 0
NumCompaniesWorked 0
OverTime 0
PercentSalaryHike 0
PerformanceRating 0
RelationshipSatisfaction 0
StockOptionLevel 0
TotalWorkingYears 0
TrainingTimesLastYear 0
WorkLifeBalance 0
YearsAtCompany 0
YearsInCurrentRole 0
YearsSinceLastPromotion 0
YearsWithCurrManager 0
Attrition 0
DistanceFromHome_Label 0
dtype: int64
```

Duplicated Data

```
df_mentah.duplicated().sum()

0
```

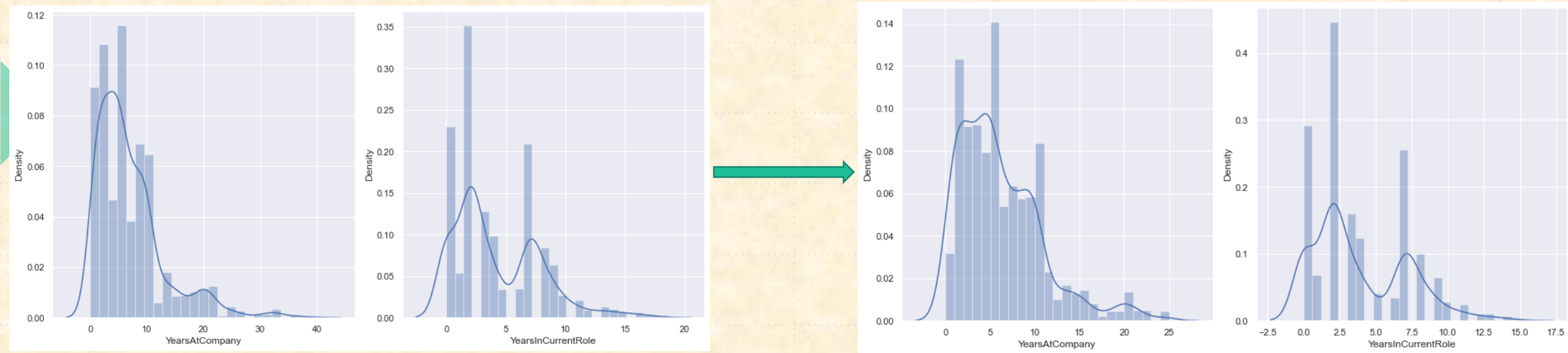
Feature Engineering

Label encoding pada feature *Overtime*, *Gender* dan *Attrition* menggunakan `LabelEncoder()`

Feature Selection

- OverTime_label
- TotalWorkingYears
- JobLevel
- YearsInCurrentRole
- YearsAtCompany
- YearsWithCurrManager
- MonthlyIncome
- MaritalStatus_label
- Age
- WorkLifeBalance

Outliers

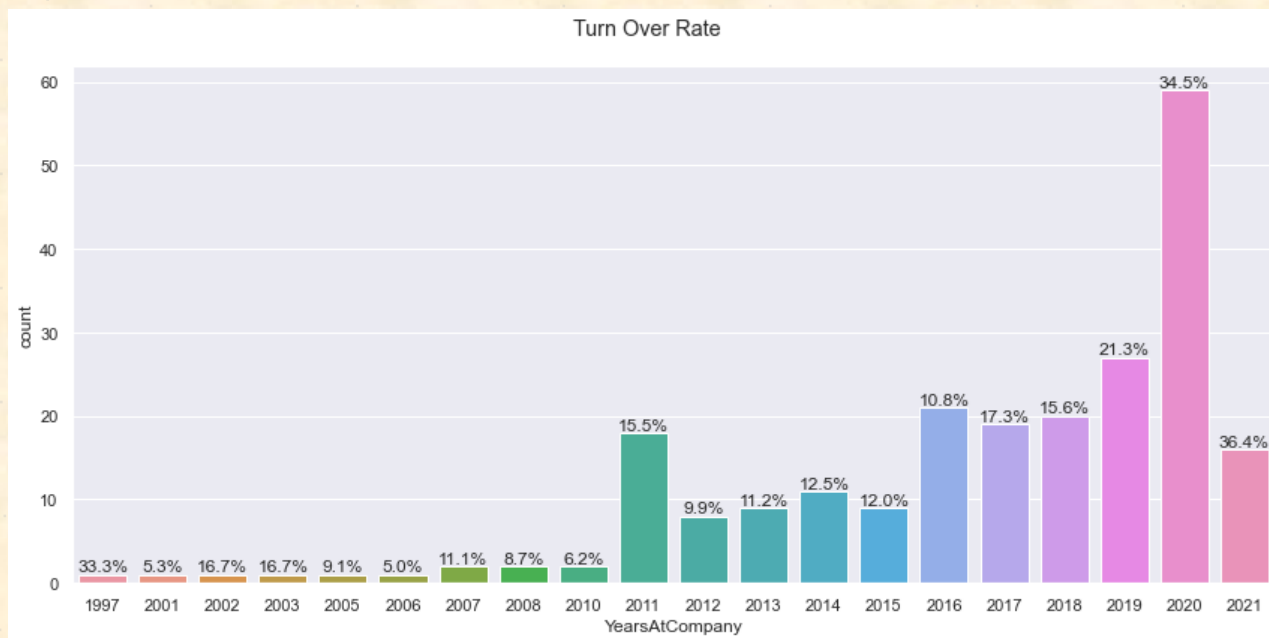
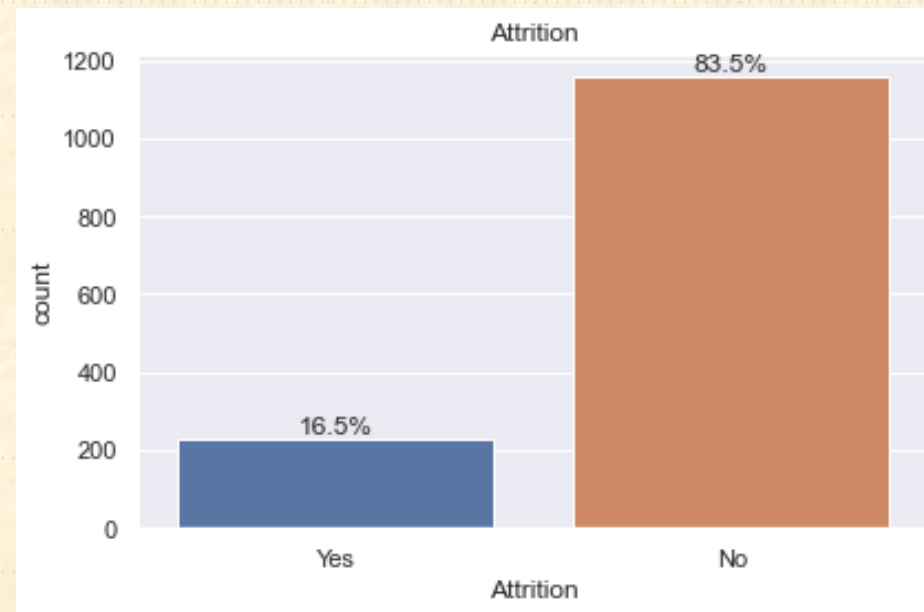
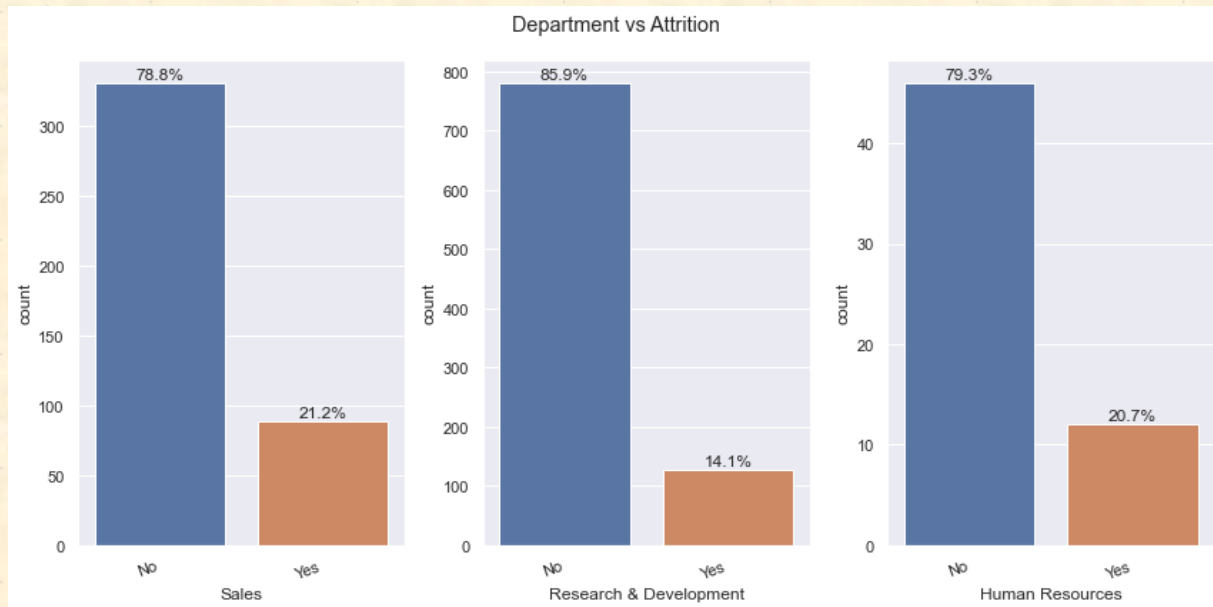


Menghapus data-data yang nilainya teralalu jauh dari data lainnya dengan menggunakan *z-score (standar deviation)* yang dilihat dari distribusi data tersebut

Normalization

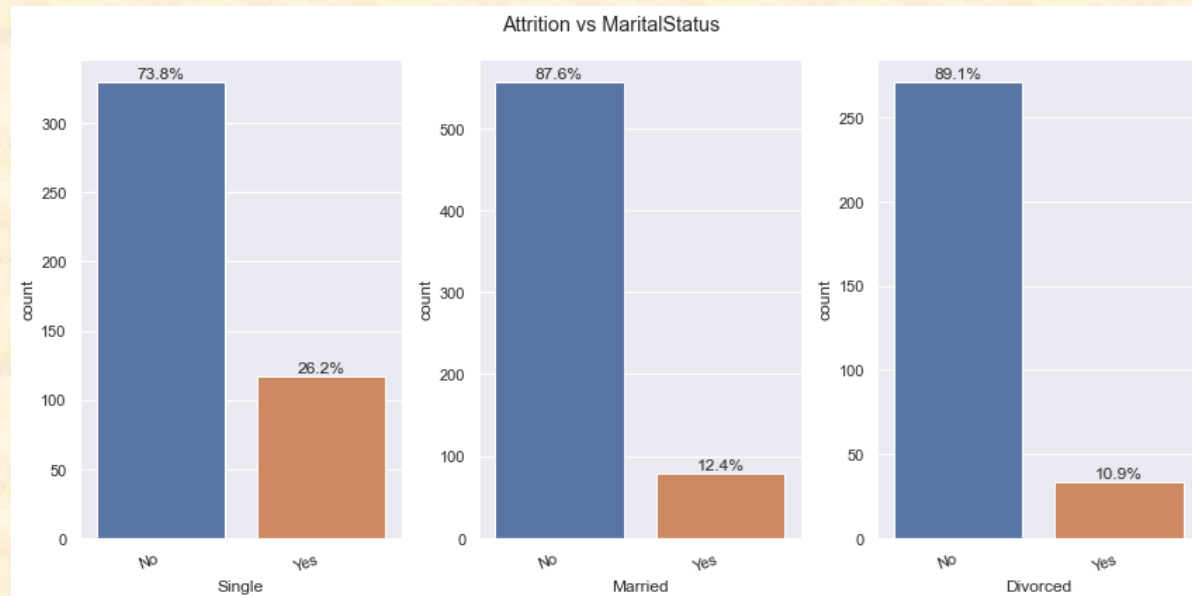
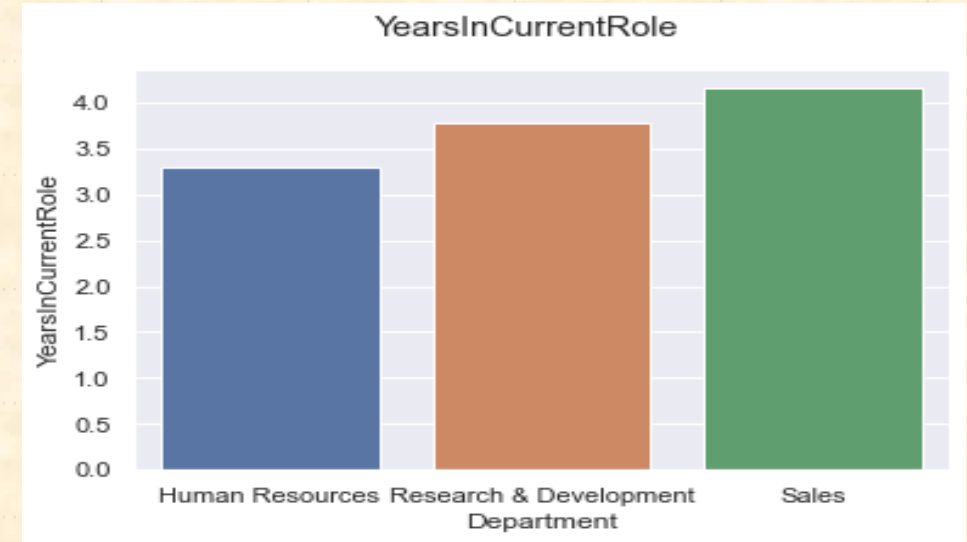
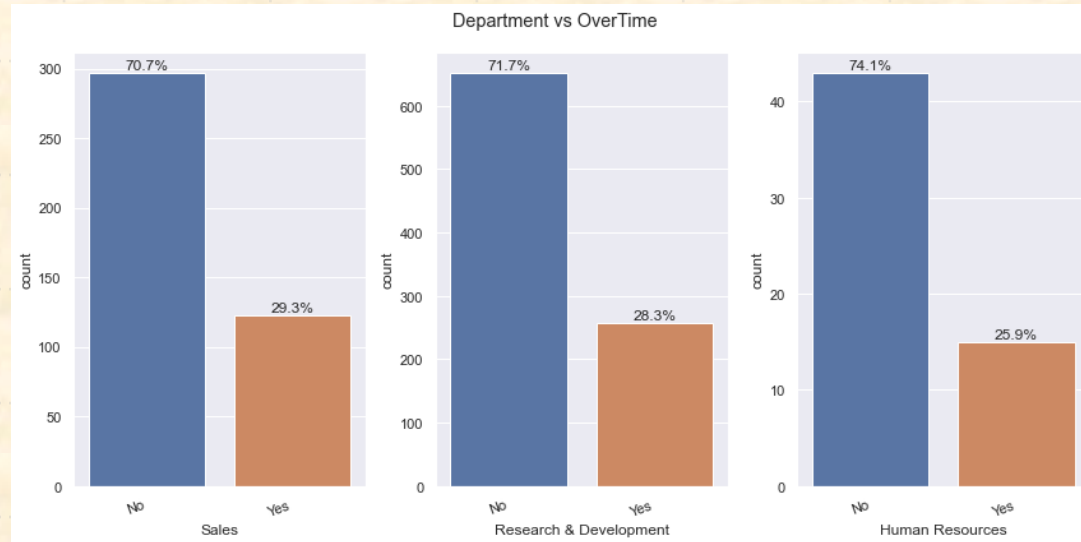
Normalization menggunakan *StandarScaler* dan *RobustScaler*

Bisnis Insight



- *Attrition Rate* perusahaan kita sebesar 16.5%
- Department Sales memiliki tingkat *Attrition Rate* terbesar
- Pada 2 tahun terakhir memiliki *Turn Over Rate* paling besar

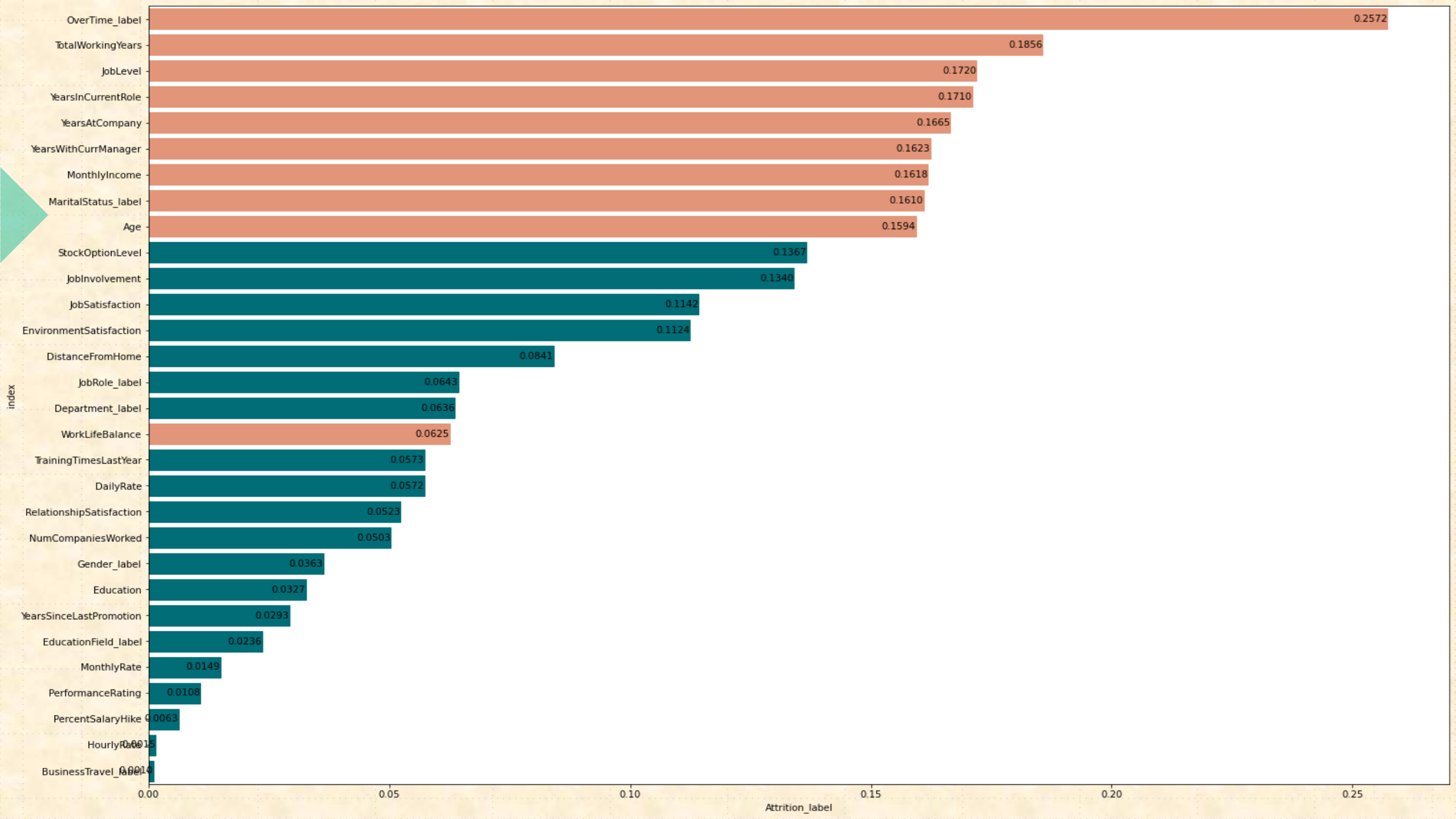
Another Feature



- Departement Sales memiliki tingkat *OverTime* lebih tinggi
- Departement Sales memiliki rata-rata lebih tinggi pada feature tahun lama bekerja
- Karyawan yang memiliki status *Single* memiliki tinggkat *Attrition* yang paling tinggi

Feature Correlation





Train – Test Split

Ratio 70% : 30% (970 : 417)

Class Imbalace

Menggunakan over_sampling (*SMOTE*) menjadi 1 : 1
(813 : 813)

Modeling

- Logistic Regresion
- K Nearest Neighbor
- Decision Tree
- Support Vector Machine
- Random Forest
- Gradient Boosting
- XGBoost

Hyperparameter Tuning

Evaluation

Accuracy	Recall	Precision
F-Score	ROC-AUC	

Top 3 Modelling

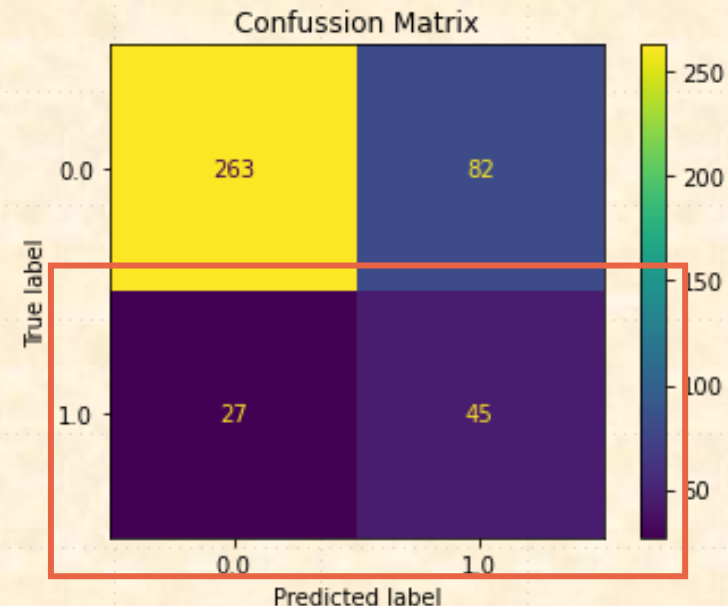
By Recall Test

	Akurasi Train	Akurasi Test	Precision	Recall	F1	ROC/AUC
LR	0.75646	0.73621	0.35156	0.62500	0.45000	0.69221
SVM	0.83825	0.75779	0.36697	0.55556	0.44199	0.67778
KNN	0.89176	0.70743	0.28814	0.47222	0.35789	0.61437

Confusion Matrix

Logistic Regresion - Tuning

- $\text{Recall} = (45)/(45+27) = 0,625$
- Jadi 62,5% merupakan karyawan yang diprediksi *attrition* dibandingkan keseluruhan karyawan yang sebenarnya *attrition*
- Meminimalisir karyawan yang diprediksi tidak *attrition* (False/Negatif) ternyata sebenarnya *attrition* (True/Positif)



Business Recommendation

Menurunkan Attrition Rate & TurnOver Rate

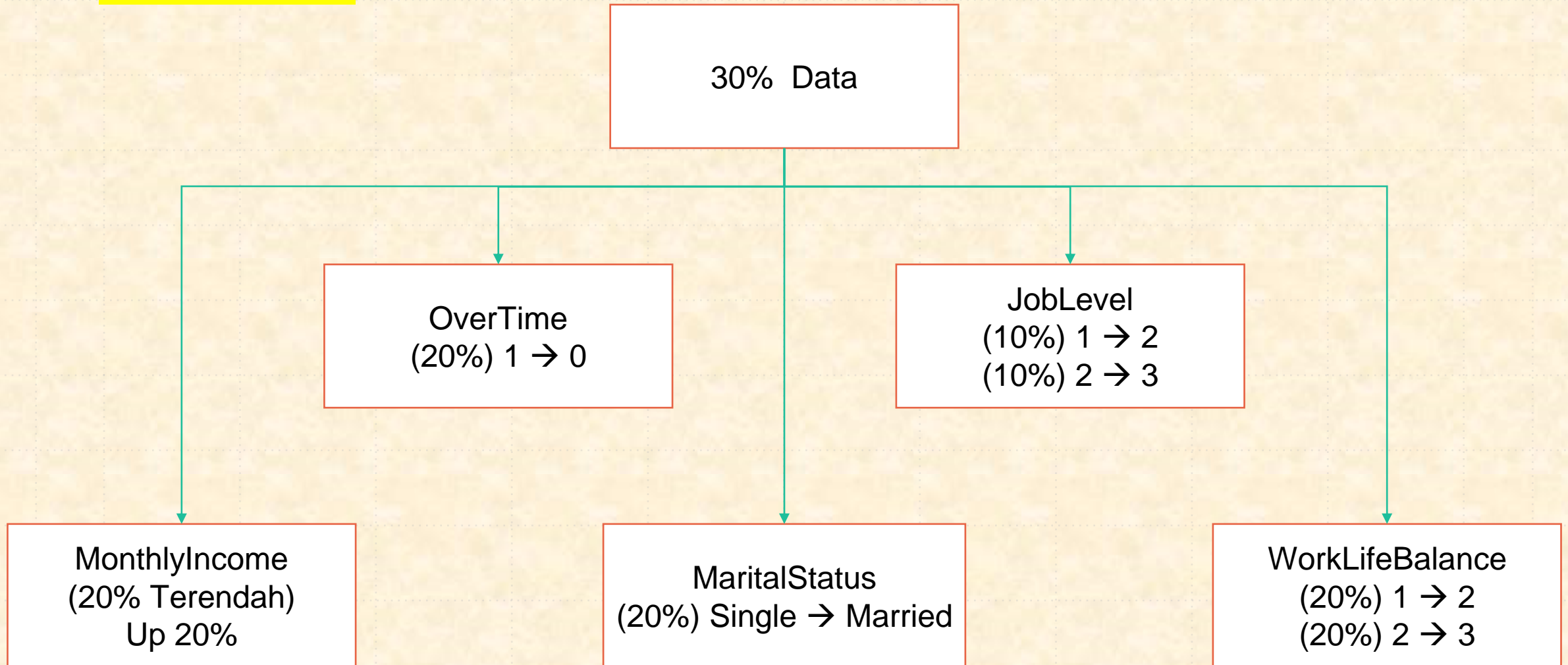
- Membatasi *over time* karyawan
- Menseleksi berdasarkan status pernikahan dari awal
- Menaikan *Monthly Income & Job Level* berdasarkan performa dan kinerja

Timbulnya loyalitas kerja dipengaruhi oleh banyak faktor, beberapa faktor pengaruh loyalitas kerja menurut Streers dan Porter (1983) adalah :

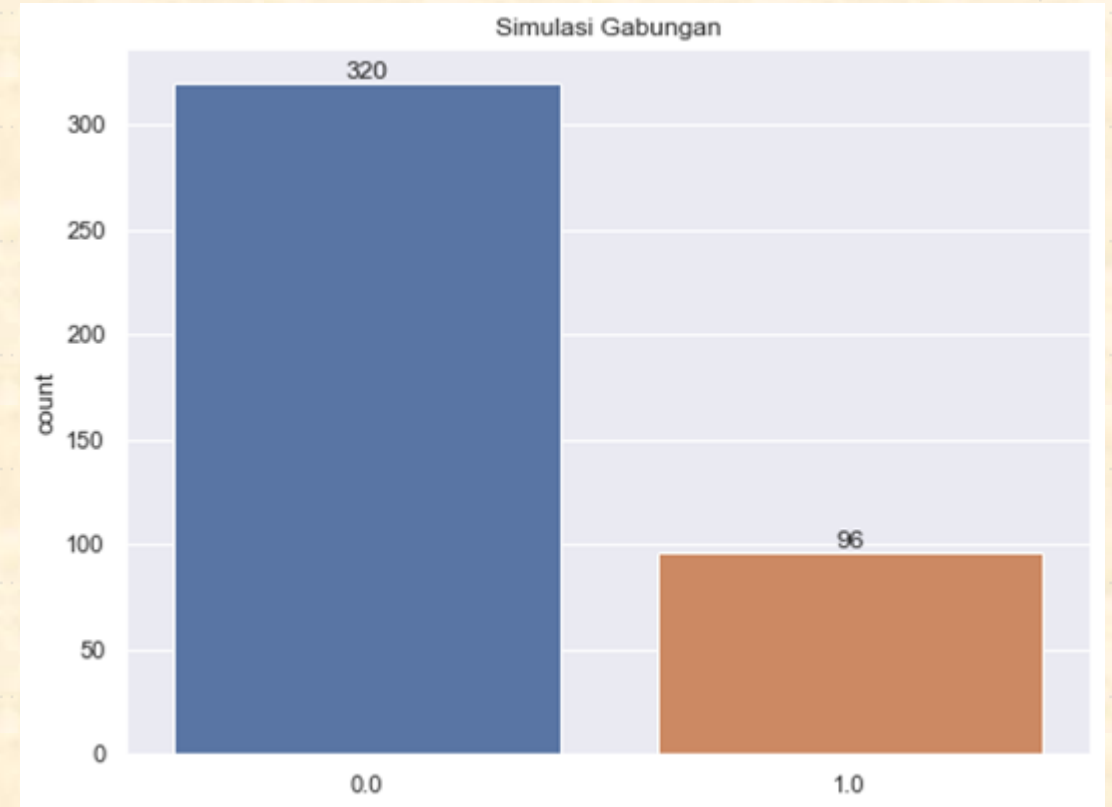
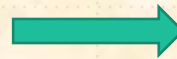
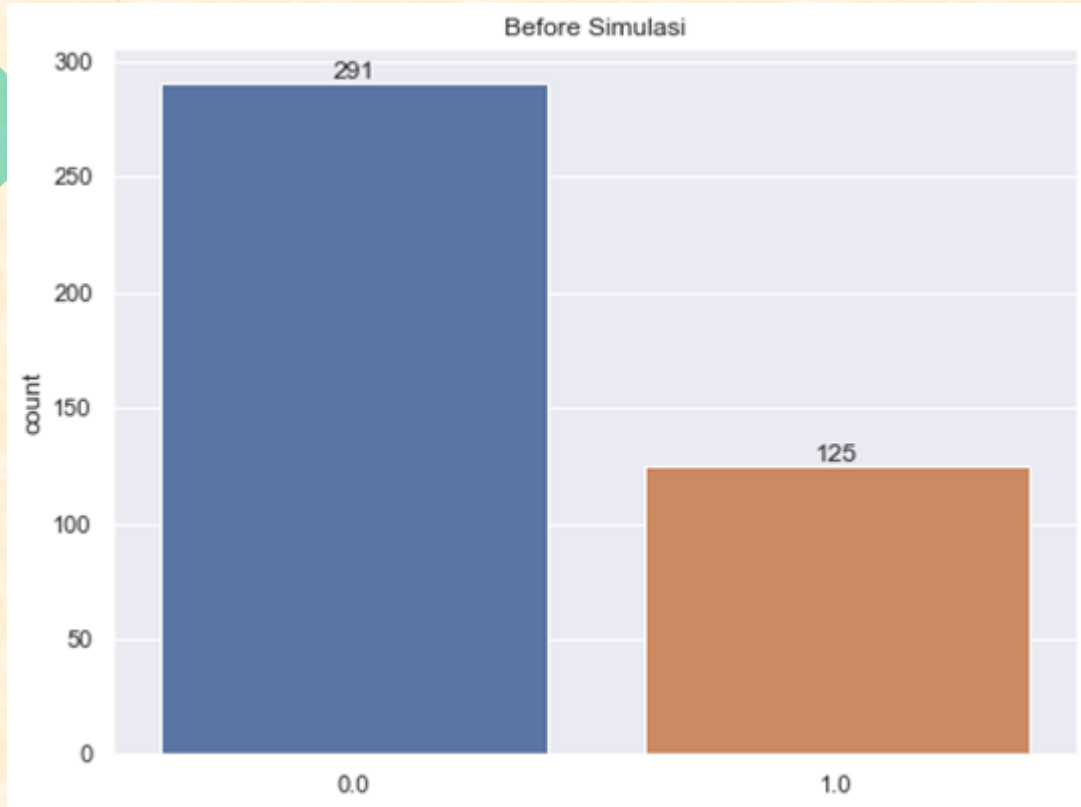
1. *Faktor individu karakteristik pribadi*
2. *Karakteristik pekerjaan*
3. *Karakteristik desain perusahaan*
4. *Pengalaman yang diperoleh dari perusahaan*

Simulasi Bisnis Recommendation

Proses Simulasi



Simulasi Bisnis Rekomendasi



Penurunan Attrition Rate 30% → 23%

Daftar Pustaka

Gibson, James, L., et al. 1987. Organisasi Dan Manajemen: Perilaku, Struktur, Proses. Binarupa Aksara, Jakarta

Gillis, J. R. (1994). Memory and identity: The history of a relationship. Commemorations: The Politics of National Identity, 8

S Keller dan M Meaney. 2017. Attracting and retaining the right talent. McKinsey Global Institute study.