



Understanding Attrition

Implementation of Machine Learning & Data Analytics for HR Management



Consulting Team



Adrian Ku



Rita Orvalho



Moritz Zoepffel



Roberta Troccoli



Alan Haro



Pedro Olivares



Sergio Ospina

CONTENT

1

Context

2

EDA

3

Feature
Engineering

4

Models

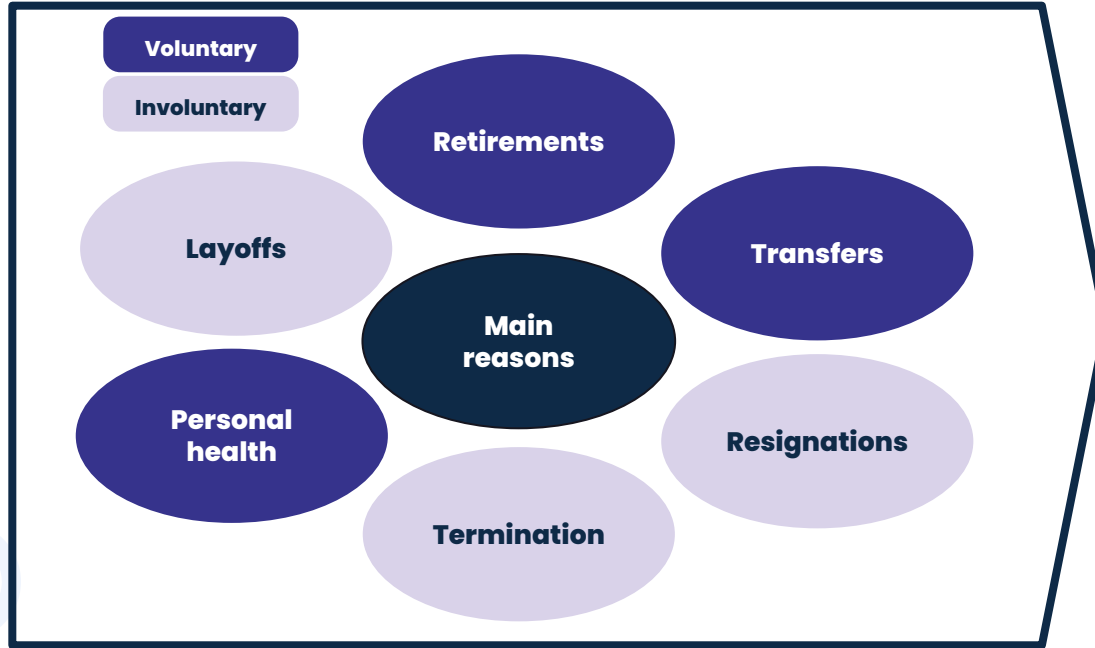
5

Results

6

Conclusions

CONTEXT – Attrition



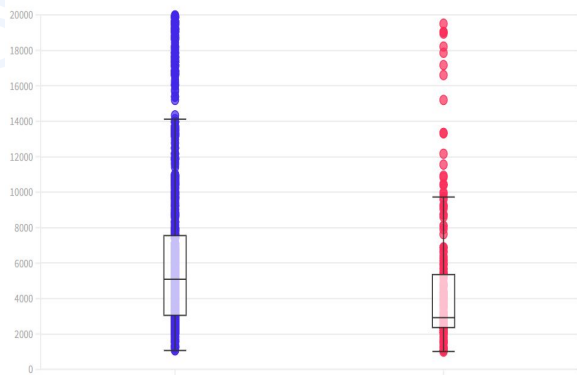
CONS

- High costs
- Loss of talent and knowledge
- Decreased morale
- Training gaps
- Lower overall productivity



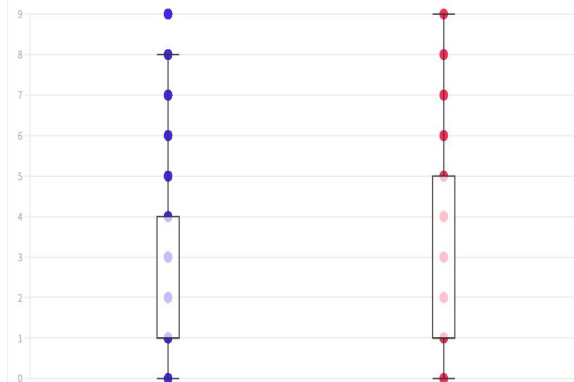
SITUATION – UNDERSTANDING THE DATA

Income Distribution



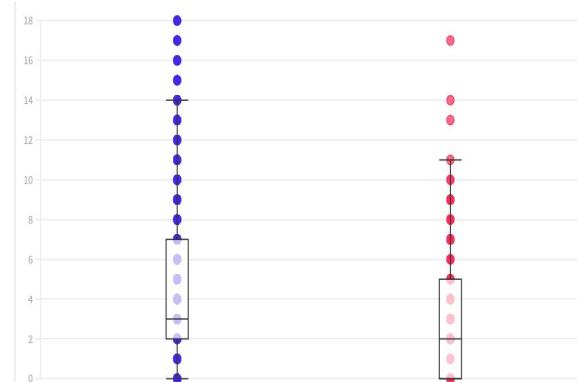
Employees with **lower salaries** produce **higher attrition rates**.

Number of Companies Worked



Employees with **high rotation** between companies have a **higher attrition rate**.

Years at Current Role



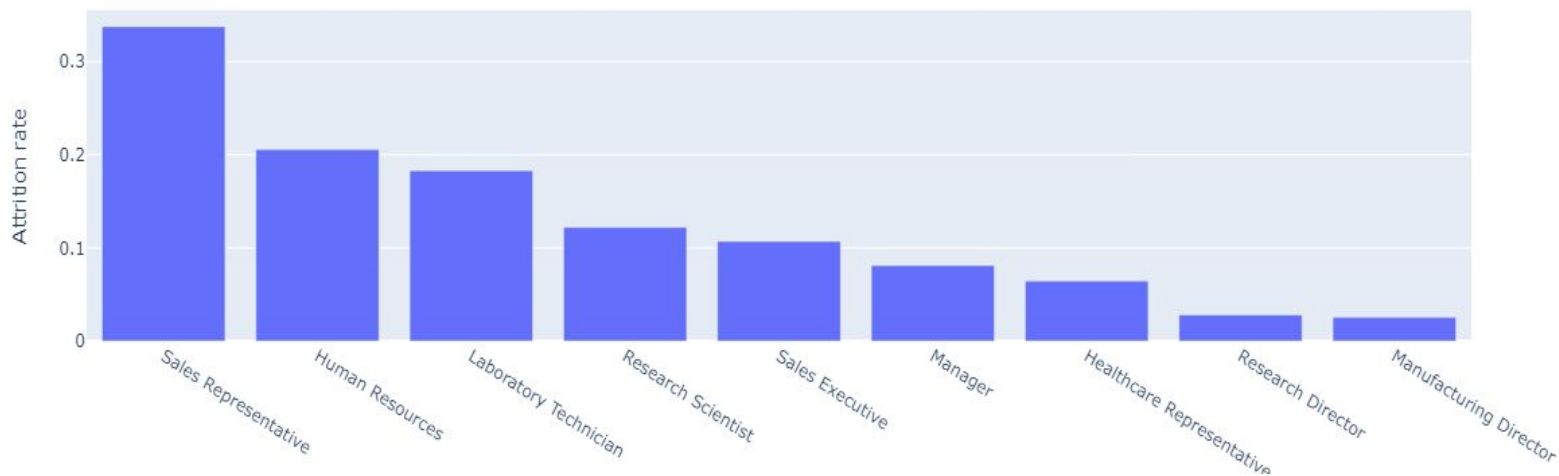
Employees with **more years** at **current role** have **lower attrition rates**.

Attrition ● No ● Yes



SITUATION – UNDERSTANDING THE DATA

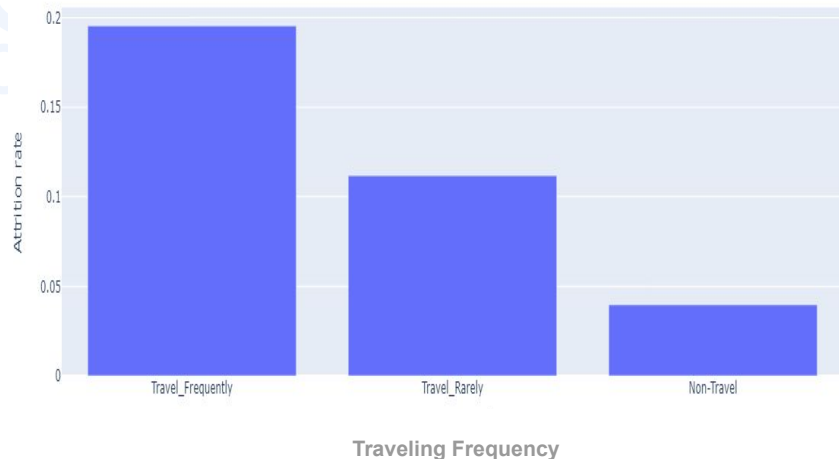
Job Roles



Different **Job Roles** present different **Attrition levels**; being Sales Representative the one with highest rate.

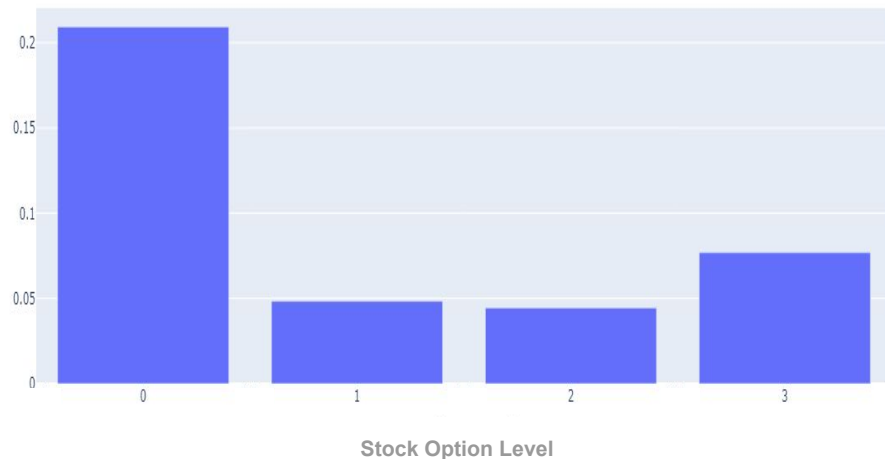
SITUATION – UNDERSTANDING THE DATA

Business Traveling



Employees that **travel frequently** have a much **higher incidence** of attrition.

Stock Options



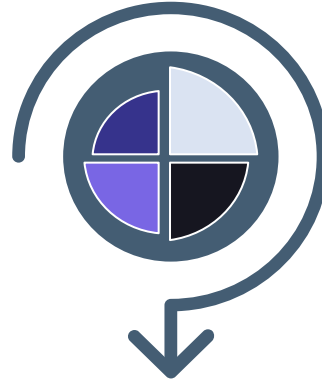
Employees with **NO stock options** have a much **higher incidence** of attrition.

FEATURE ENGINEERING



Unnecessary data

Remove **data** that is **irrelevant** for our model prediction.



Categorical

Since we are facing a **classification problem**, we applied **One-hot encoding**- a method that converts labels into numbers.



Numerical

For **quantifiable** characteristics, **Standard Scaling** method has been applied. It deals with the magnitudes.

IMPACT OF PREDICTIONS

	Prediction	Outcome	Result	
TP	Attrition	Attrition	Actions to keep employee	
FP	Attrition	No Attrition	Wrong allocation of resources	
TN	No Attrition	No Attrition	Nothing happens	
FN	No Attrition	Attrition	Loose employees	



IMPLEMENTED MODELS

Cross validation / Random search

Model performance

Hyperparameters' optimization

Class Weights - Balancing Data

Models / Score Methods

Random Forest

Logistic Regression



Support Vector Machines

Extreme Gradient Booster

K Nearest Neighbors

ROC-AUC

F-1

0.855

0.092

0.854

0.446

0.833

0.000

0.813

0.396

0.751

0.000



MAIN VARIABLES AFFECTING ATTRITION

INCREASE ATTRITION



No **Stock Options**



Low **Job Involvement**



Poor **Work-Life Balance**



Marketing Background



Low Satisfaction in **work environment**

DECREASE ATTRITION



Few **Business Traveling**



High **Performance Rating**



Low **Work Overtime**



High **Job Positions**



Higher **Salaries**



ACTIONS

If...

Offer...

No Stock Options



Increase salary, bonuses, competitive benefits (**retirement, healthcare...**)

Low Job Involvement



Additional **training**, team building, coaching options

Poor Work-Life Balance



Increase **flexibility, remote work**, adjust working hours, **wellness programs**

Marketing Background



Create a clear **career path, benchmark analysis** of other **marketing companies**

Low satisfaction in work environment



Team building activities, **mental health** check, **feedback** sessions



NEXT STEPS



- Increase **employee feedback**
- Implement **retention metrics** in relation with observed variables
- **Integrate** recruiting, HR and C-level



- **Track actions and results over next period** to understand positive impact on Attrition Levels

12.3%
Attrition

*Estimated rate for next period

~10%
Attrition

EoY Target





THANKS!

Do you have any questions?

consulting@dynamite.com

+34 626 778 682

dynamiteconsulting.com

