

# SALIFORT MOTORS: Employee Retention Analysis

**Objective:** From data from all around the world, analyze which factors influence an employee's decision to stay in the company or leave.

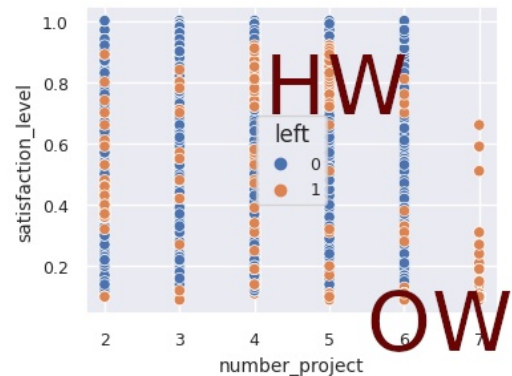
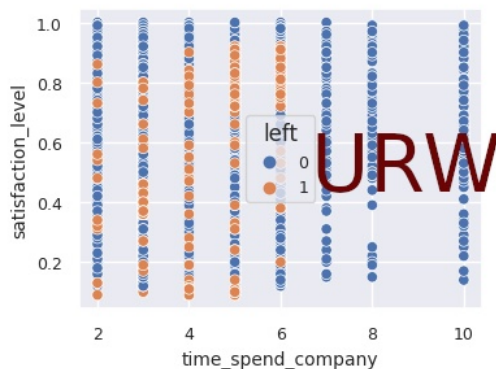
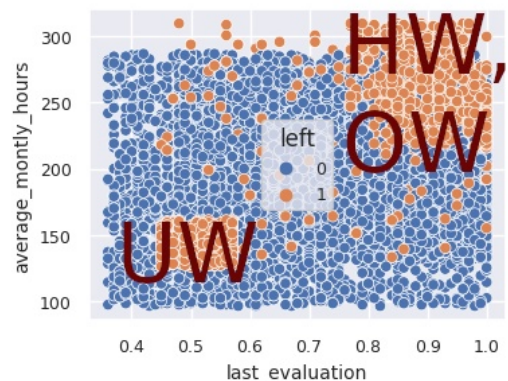
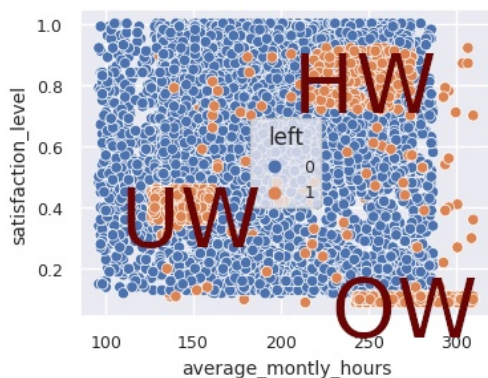
## Methodology

Build different machine learning models from the data and test them. Two classifier algorithms were tested: *logistic classification* (LC) and *random forest* (RF), an aggregation of decision trees.

## Observations about the models and statistics

- RF performs better than LC with the given data. However it relies too strongly on the feature "satisfaction\_level"
- LC provides better explainability of feature influences on people deciding to stay or leave
- 'time\_spend\_company', 'number\_project', 'satisfaction\_level' are **very important features** for both models
- The **salary is not a very predominant feature** in any of the models. But still the higher the salary the higher the retention rate.
- The **number of projects an employee is working on is quite relvant**. People having 7 projects or more in parallel leave to 100%. But also people with 2 or less projects leave to 60%. Such "**extremes**" leaving is also observed for 'average\_monthly\_hours' and 'last\_evaluation'
- (Hypothetical) Groups of employess are observed in plots : HW: hard workes at risk of leaving. OW: Overloaded, unhappy workers. URW: workers who probably stay until rent.

	f1_score	accuracy_score	precision_score	recall_score
Random Forest Balanced on Test Set	0.961733	0.982333	0.992548	0.932773
Logistic Regression Balanced on Test Set	0.734719	0.841837	0.624091	0.893016



## Further Considerations

- A multiclass algorithm can be used for futher analysis of hypothetical groups (ethical considerations: segregation of employees).
- It would be importatnt to revisit the way the employees' evaluations are made and whether this score takes into account all relevant aspects in a weighted manner.
- Balanced employees (regarding the studied features) are not prone to leave. However some relevant features change automatically with time (time in company).