

# Apresentação dos Resultados

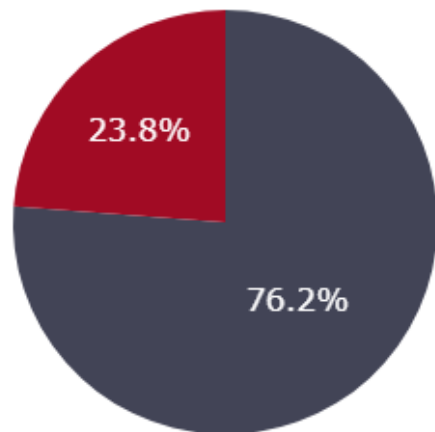
**People Analytics – Previsão se o colaborador vai ou não deixar a empresa.**

Realizador por: Henrique K. Secchi

# Análise Descritiva

# Colaboradores que deixaram a empresa

Proporção de colaboradores que deixaram a empresa



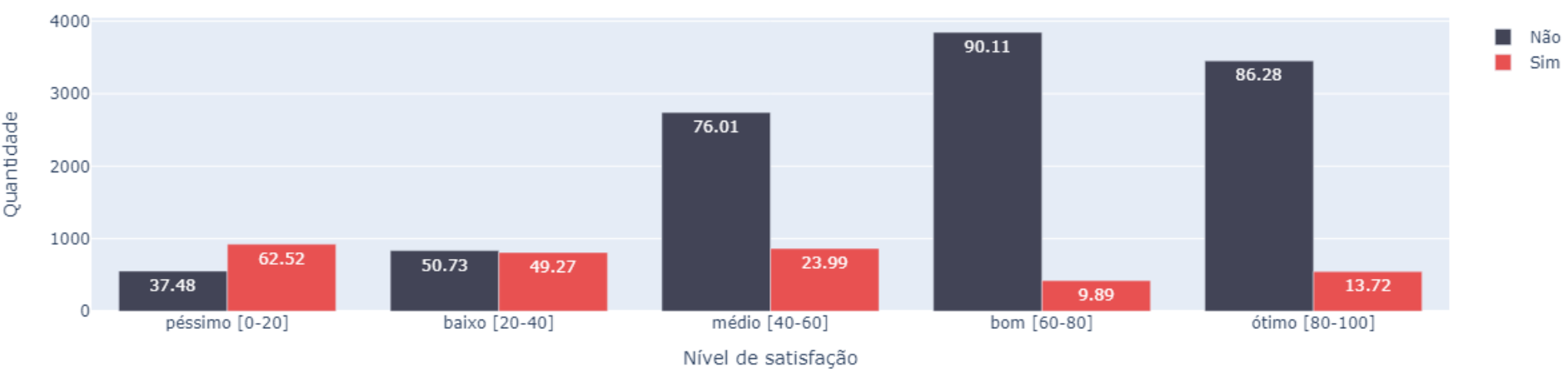
■ Não  
■ Sim

Deixou a empresa	Não Deixou
3571	11428

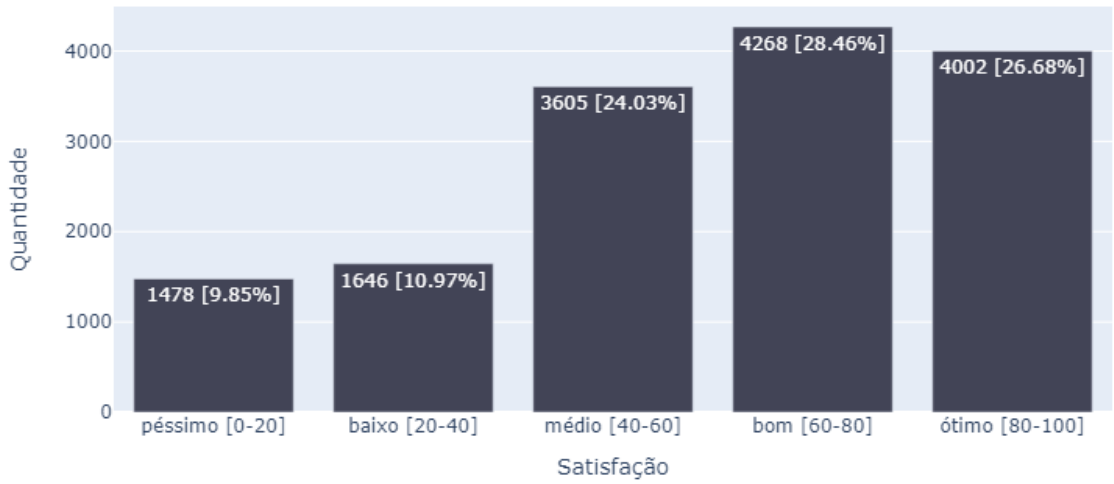
- A empresa possui média 6 em nível de satisfação;
- Na última avaliação a média foi de 7 pontos, mostrando um aumento;
- Pela matriz de correlação e pelo método Random Forest, observei que o principal atributo do conjunto de dados é o nível de satisfação, vamos analisar:

# Nível de Satisfação

Quantidade de demissões e não demissões por grupo de satisfação dos colaboradores



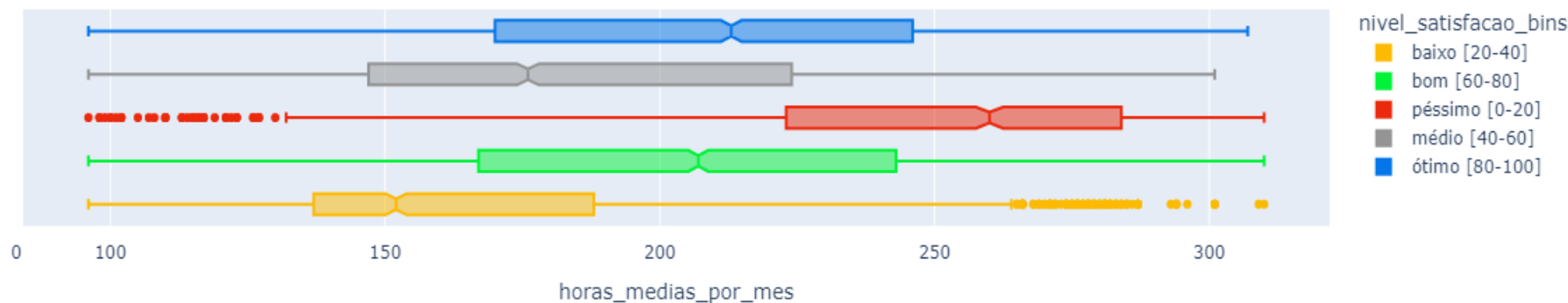
Quantidade de nível de satisfação dos colaboradores



- Quanto menor o nível de satisfação, mais propenso o colaborador está a deixar a empresa;
- 9.85% do total de colaboradores estão insatisfeitos e 20% possuem avaliação negativa;

# Nível de Satisfação

Distribuição de horas médias trabalhadas por mês

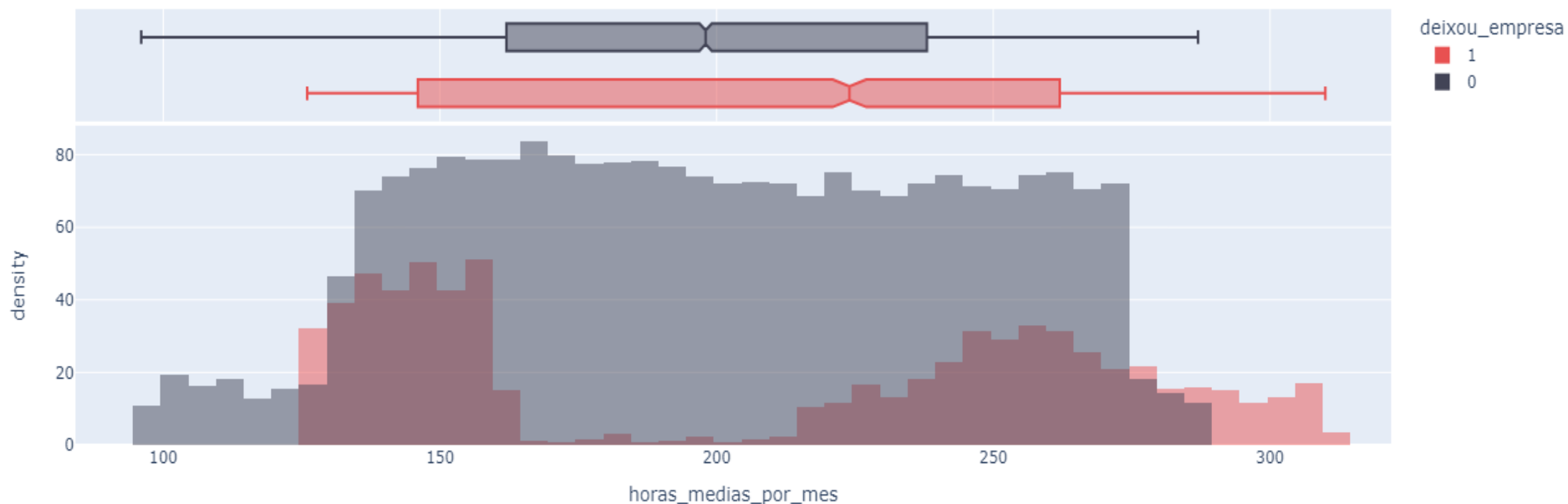


Pela visualização por boxplot fica nítido que a quantidade de horas trabalhadas por mês afeta diretamente na satisfação dos colaboradores, observe que:

- **Satisfação péssima:** São os colaboradores que trabalham muitas horas semanais, são os trabalhadores que fazem muita hora extra no serviço;
- **Satisfação baixa:** Se concentra mais nos colaboradores que trabalham poucas horas por semana;
- **Satisfação média:** Colaboradores que trabalham em média 40 horas semanais;
- **Satisfação boa e ótima:** Se concentra principalmente nos trabalhadores que tem a jornada semanal de 44~ horas em média.

# Horas Trabalhadas e Deixou a Empresa

Distribuição de horas médias trabalhadas por mês



- Os colaboradores que trabalham entre 125~164 horas semanais são mais propensos a deixarem a empresa;
- A partir das 215 horas semanais o histograma de colaboradores que deixaram a empresa volta a crescer;
- Poucos funcionários que fazem mais de 275 horas semanais continuam na empresa;

# Nível de Satisfação

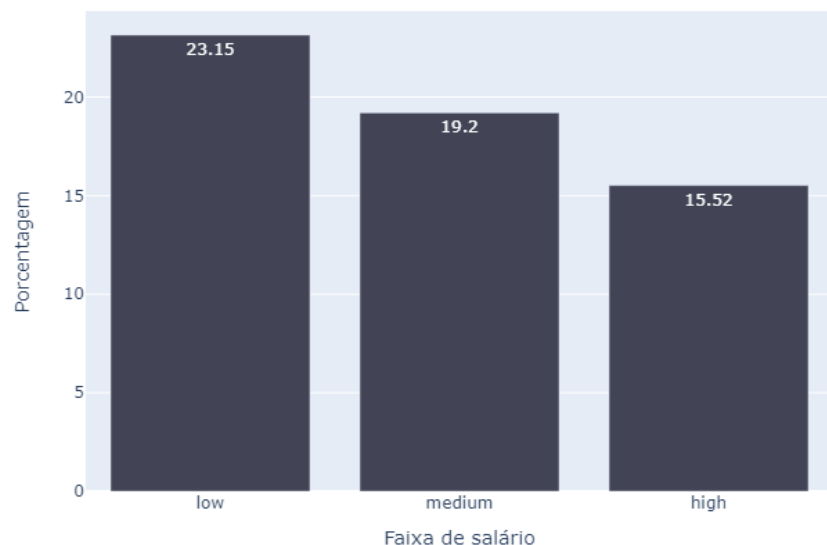
Nível de satisfação baixa ou péssima por área

area	% do grupo
accounting	26,2
hr	22,19
technical	21,8
sales	20,65
IT	20,62
support	20,46
marketing	20,28
product_mng	19,29
RandD	19,06
management	16,51

- A própria área que utiliza o People Analytics é uma das maiores em nível de insatisfação na empresa.

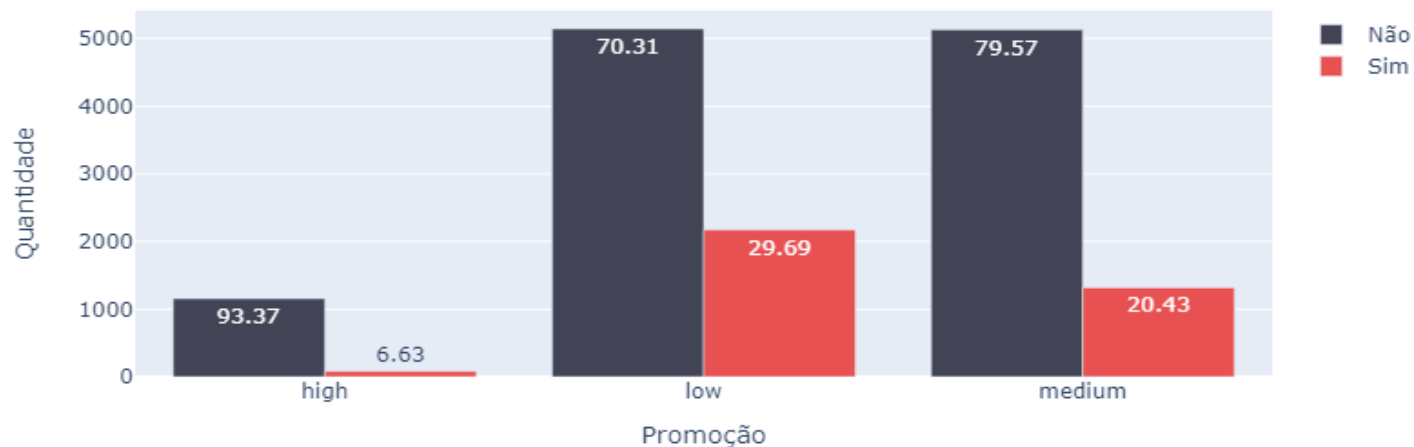
# Categoria de Salário

Porcentagem por grupo de salario de colaboradores insatisfeitos



- Há uma relação direta entre categoria salarial e insatisfação dos colaboradores;
- A categoria salarial impacta diretamente no evento do colaborador deixar ou não a empresa.

Quantidade de demissões e não demissões por categoria de salario





# Análise Preditiva

## Seleção de atributos

Atributo	Importância
nivel_satisfacao	0.290887
tempo_empresa	0.229679
horas_medias_por_mes	0.155505
numero_projetos	0.141435
ultima_avaliacao	0.122763
acidente_trabalho	0.013874
salario_high	0.006540
salario_low	0.006295
area_technical	0.005101
area_sales	0.004219
salario_medium	0.004052

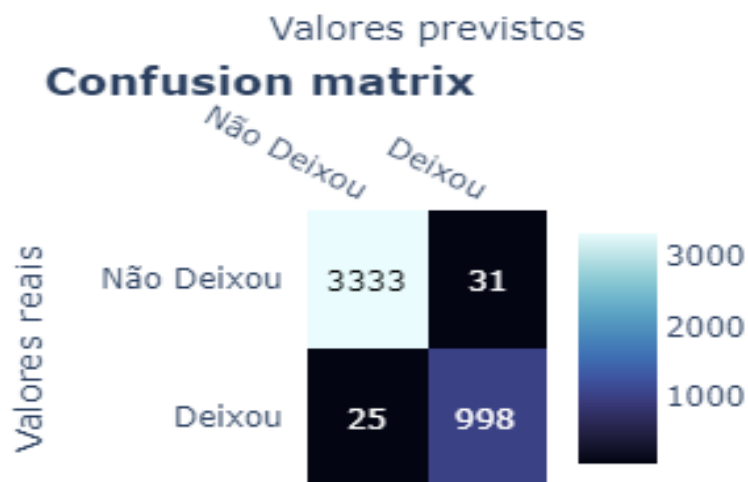
- Utilizei o algoritmo de Random Forest para selecionar os atributos mais relevantes para o modelo;
- Apliquei o critério: importância > 0.1 para a seleção;

# Avaliação dos Algoritmos

Nome do Algoritmo	Método de Escala	Avaliação f1 score
Random Forest	StandardScaler	97,27
Random Forest	RobustScaler	96,88
Random Forest	Normalizer	94,46
XGB - Extreme Boosting Classifier	RobustScaler	91,5
XGB - Extreme Boosting Classifier	StandardScaler	91,5
SVM - Support Vector Machine	StandardScaler	90,85
KNN	RobustScaler	90,45
KNN	StandardScaler	90,41
SVM - Support Vector Machine	RobustScaler	89,89
KNN	Normalizer	87,91
XGB - Extreme Boosting Classifier	Normalizer	86,87
Logistic Regression	RobustScaler	66,44
Logistic Regression	StandardScaler	66,15
Naive Bayes	StandardScaler	65,45
Naive Bayes	RobustScaler	65,45
Logistic Regression	Normalizer	63,68
Naive Bayes	Normalizer	59,5
SVM - Support Vector Machine	Normalizer	43,92

# Algoritmo final

Random Forest obteve a previsão final no conjunto de testes de 97.27%.



Algoritmo: [Random Forest]

- Escala:[StandardScaler] [ACCURACY] treino:[99.59%] testes:[98.72%]
- Escala:[StandardScaler] [F1 SCORE] treino:[99.59%] testes:[97.27%]

Fim.