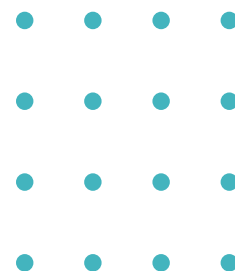


Evasão de profissionais da saúde

BOOTCAMP MIA - GRUPO 2





NOSSO TIME

MENTORA



KAROL FERNANDES
cientista de dados - Credits



AMANDA CALVO
estudante de engenharia
biomédica



LETÍCIA ALVES
médica



**LARISSA DEPA | LARISSE
DEPA**
biomédicas

01



SOBRE O PROJETO

02



DADOS

03



FEATURE
ENGINEERING

04



MODELO
PREDITIVO

05

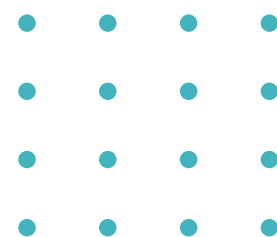


CONCLUSÕES

06



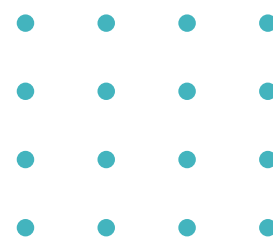
NOSSOS LINKS



01



SOBRE O PROJETO





SOBRE O PROJETO

O Problema

A desistência de profissionais da área da saúde pode gerar um grande impacto em empresas de healthcare e desestruturar sistemas de saúde

O Desafio

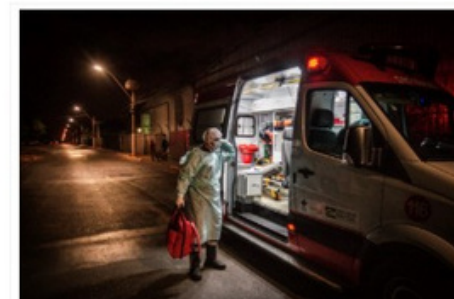
Recebemos um dataset baseado no conjunto de dados do IBM Watson para a saída de funcionários da área

Utilizamos os dados para prever quando um profissional irá sair da área e determinar os fatores relevantes para essa evasão

09/02/2022 Expostos e sem valorização, jovens podem desistir da Enfermagem

Afinal, por que estudantes vão escolher seguir em uma profissão que trabalha à exaustão e recebe salário mínimo?

O esgotamento, a desvalorização e a falta de perspectivas estão levando estudantes e profissionais de Enfermagem a desistir da carreira. Com o aprofundamento da



10/02/2022 Pesquisa inédita revela condições precárias para profissionais de saúde

Segundo levantamento da Fiocruz, 80% dos técnicos e auxiliares que atuam frente à Covid-19 vivem situação de desgaste profissional

Uma pesquisa inédita realizada pela Fiocruz revelou preocupante para os trabalhadores considerados “periféricos” da área da saúde no país.

Segundo o estudo, 80% dos profissionais de nível técnico e



13/07/2021

Cresce o risco de abandono da profissão de enfermagem

Estudo do Conselho Internacional de Enfermagem associa as más condições de trabalho a desistência dos profissionais.

Estudo global do Conselho Internacional de Enfermagem mostra que 90% das entidades temem que o estresse e as más condições de trabalho,



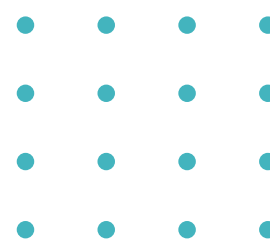
Pesquisa analisa o impacto da pandemia entre profissionais de saúde

22/03/2021

Por: Filipe Leonel (Ensp/Fiocruz)

Compartilhar: [Twitter](#) [Facebook](#) [WhatsApp](#)

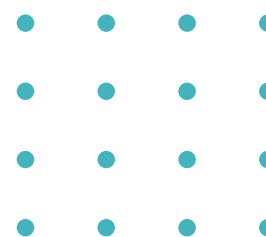
Há mais de um ano atuando na linha de frente contra a Covid-19, os profissionais da área da Saúde estão esgotados! E essa exaustão advém não só da proximidade com o elevado número de casos e mortes de pacientes, colegas de profissão e familiares, como também das alterações significativas que a pandemia vem provocando em seu bem-estar pessoal e vida



02



DADOS





DADOS

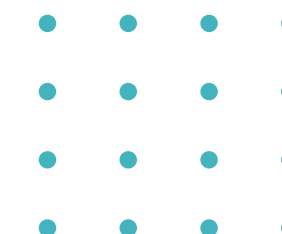
O dataset

dados de 1676
funcionários



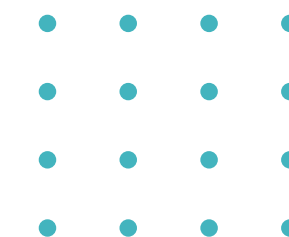
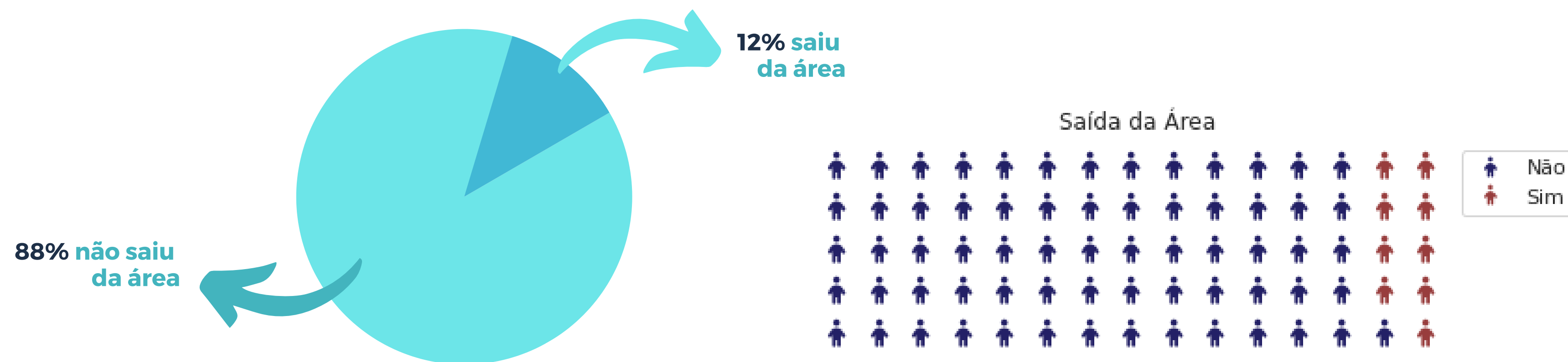
35 variáveis

Dados relacionados à idade, gênero, salário,
educação, estado civil, turno, horas extras,
anos de trabalho...



Evasão geral da área

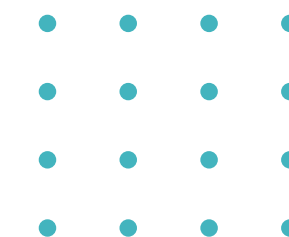
Porcentagem de funcionários que saiu da área





DADOS

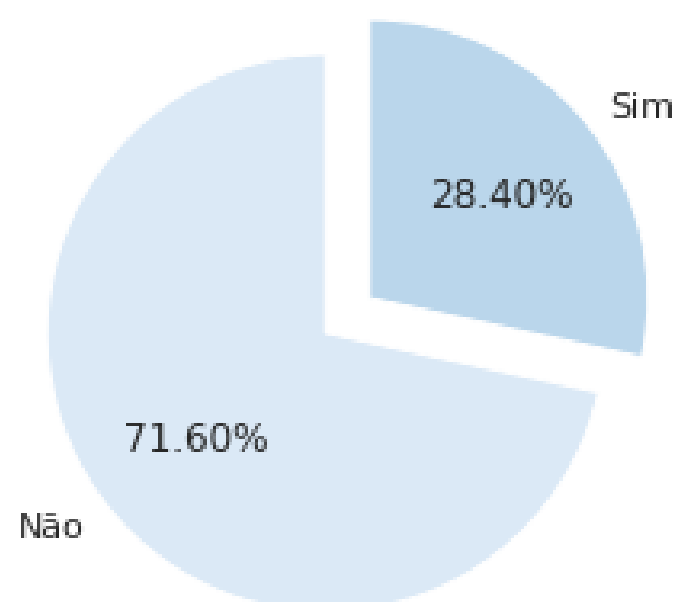
Ok,
mas quais foram os
fatores mais relevantes
para essa evasão



Horas extras

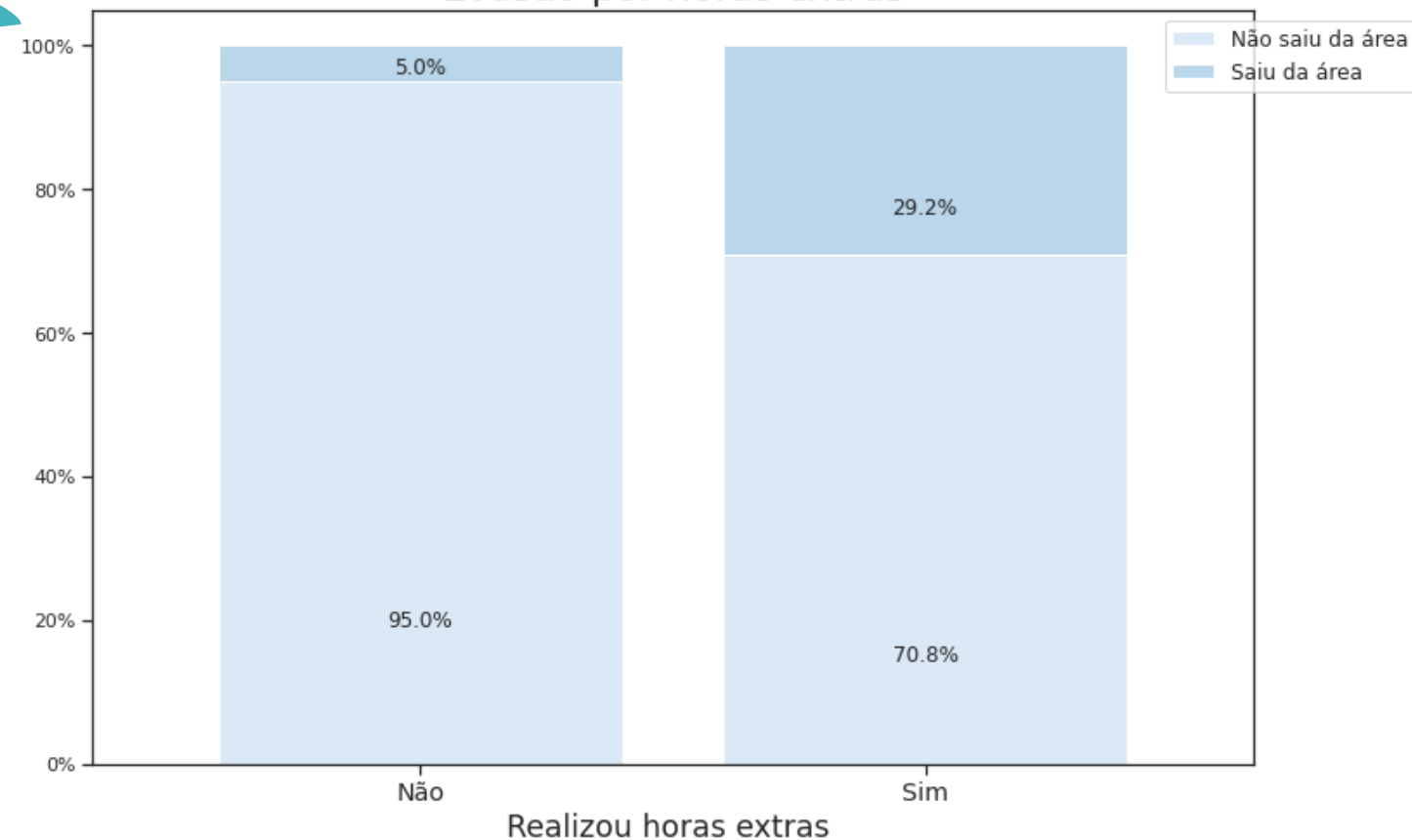
O campeão para determinar a evasão dos funcionários: horas extras!

Profissionais que realizam horas extras

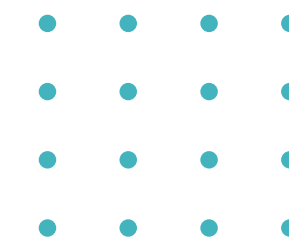


apenas 5% dos profissionais que não realizaram horas extras saíram da área

Evasão por horas extras



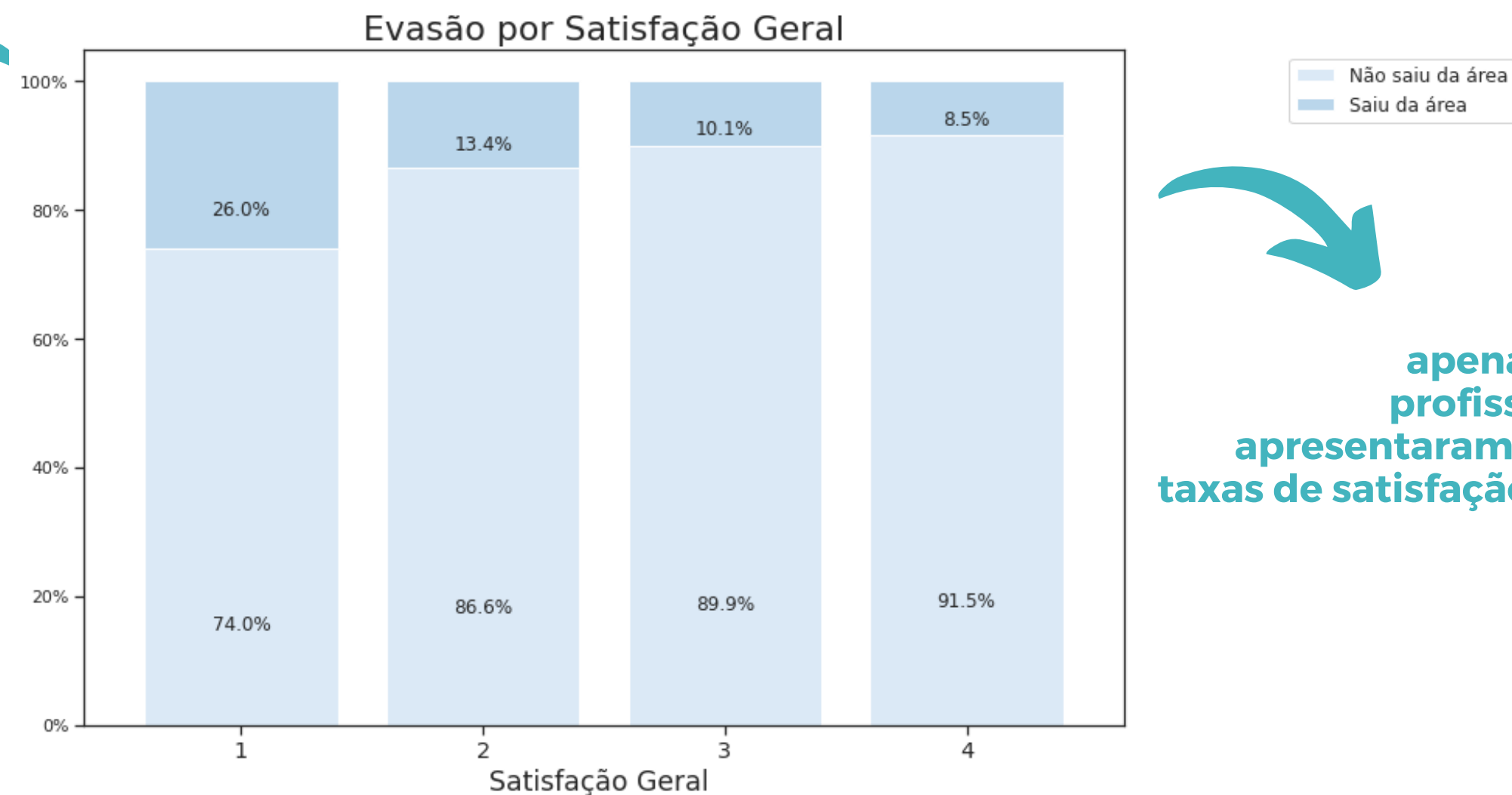
29% dos profissionais que realizaram horas extras saíram da área



Satisfação Geral

Uma média das variáveis "Satisfação com o ambiente", "Satisfação no trabalho" e "satisfação no relacionamento" teve grande peso!

26% dos
profissionais que
apresentaram as
piores taxas de
satisfação saíram
da área



apenas **8.5%** dos
profissionais que
apresentaram as maiores
taxas de satisfação saíram da
área



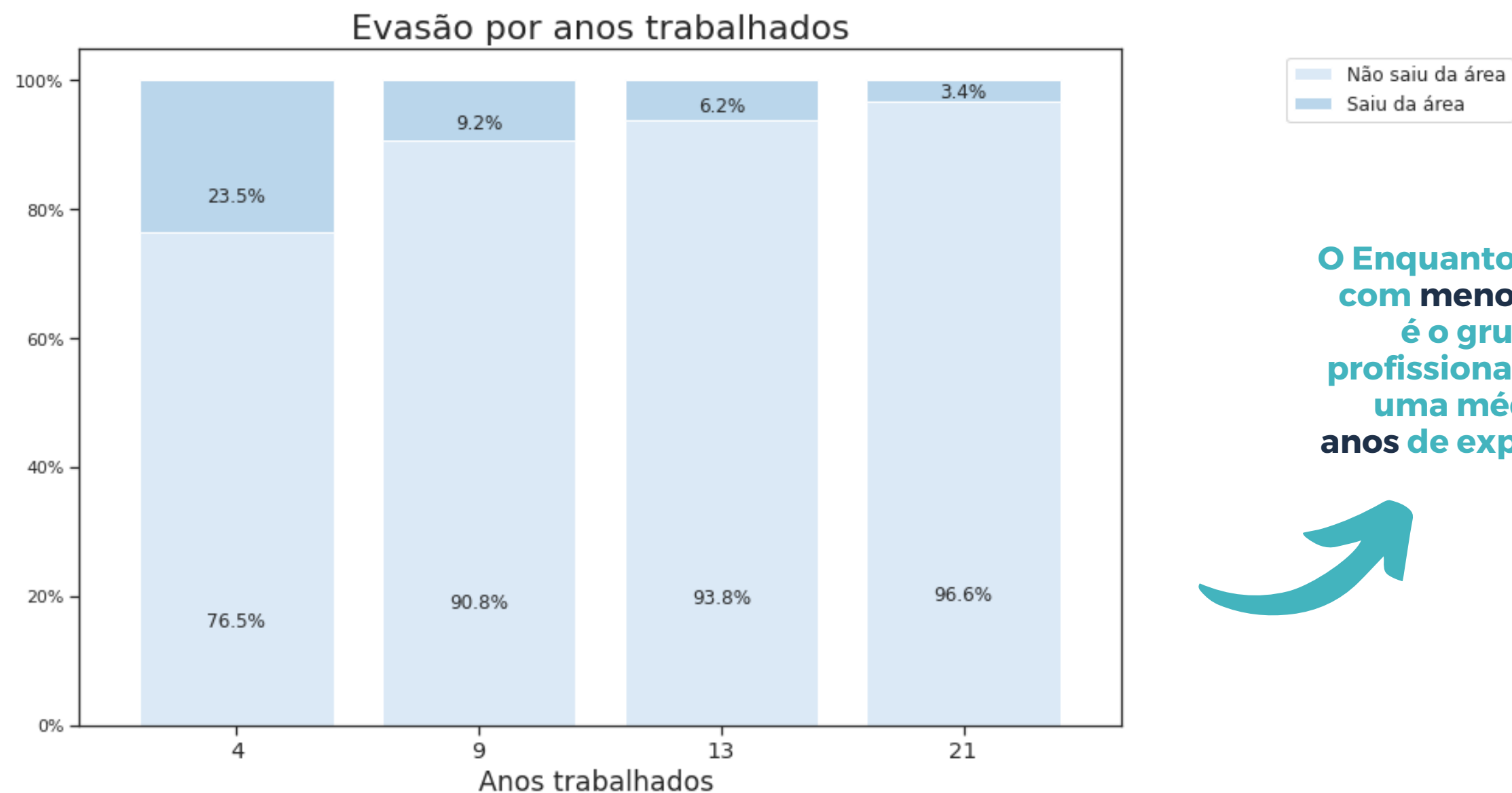


DADOS

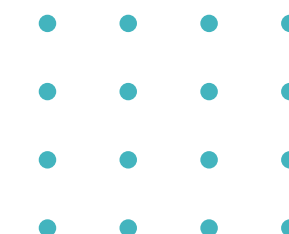
Anos trabalhados

Os anos trabalhados também tiveram forte relação com a evasão

O grupo com maior evasão é o de profissionais que trabalharam uma média de 4 anos.

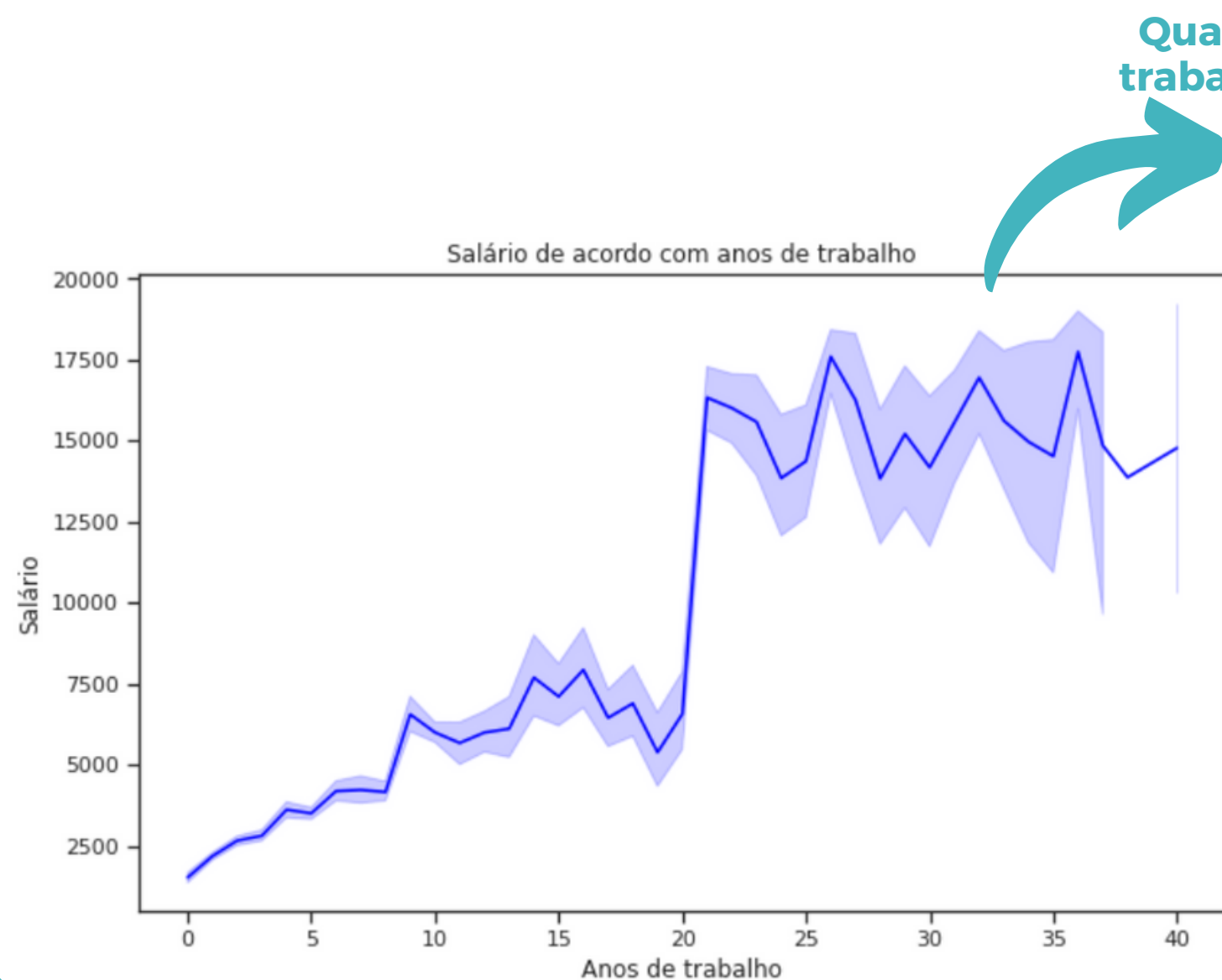


O Enquanto o grupo com menor evasão é o grupo cujos profissionais já têm uma média de 21 anos de experiência

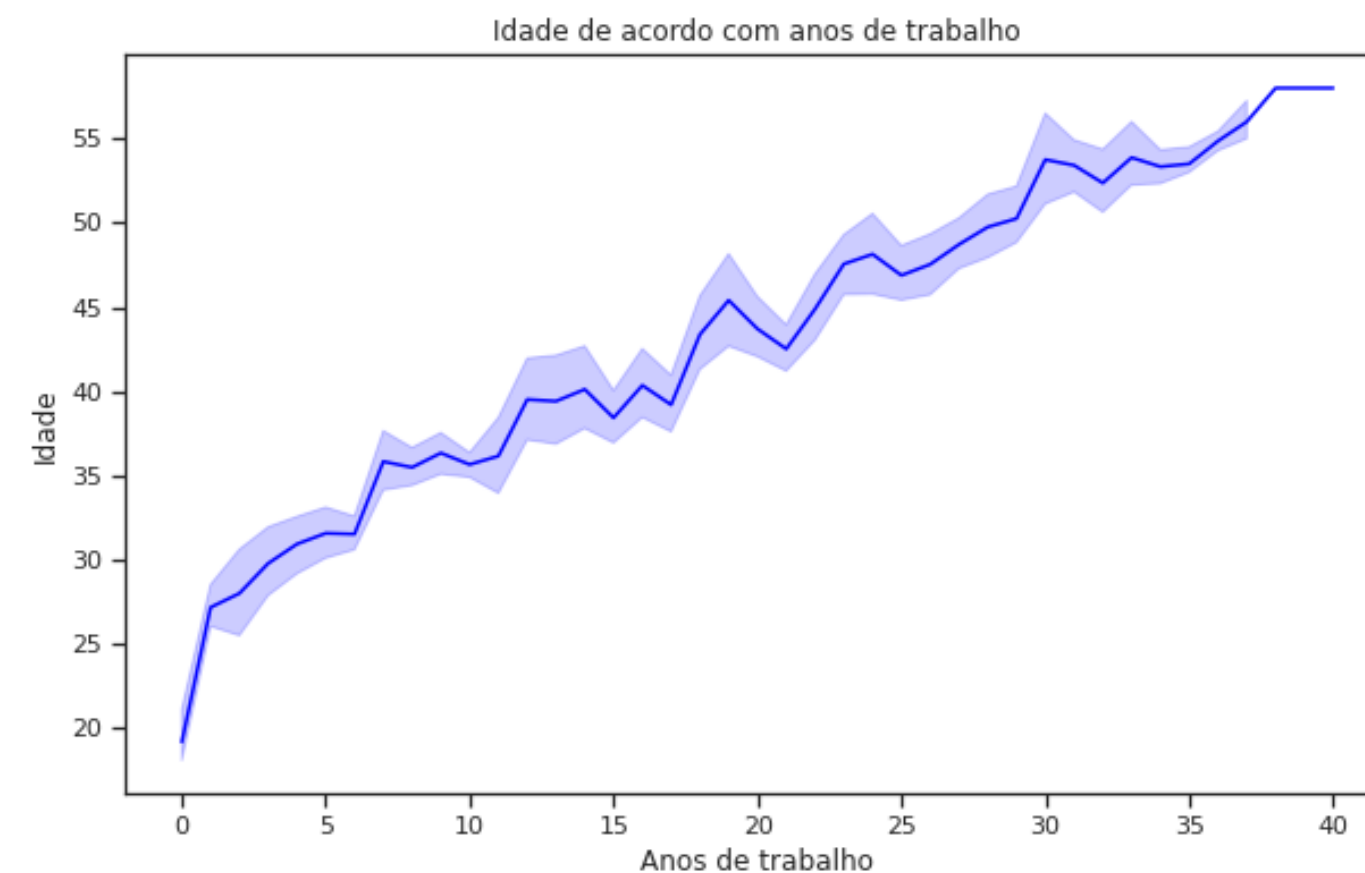


Anos trabalhados - outros fatores relacionados

Porém, não é só isso... A quantidade de anos trabalhados é fortemente relacionada com outras duas variáveis: **salário e idade!**



Quanto mais anos trabalhados, maior o salário



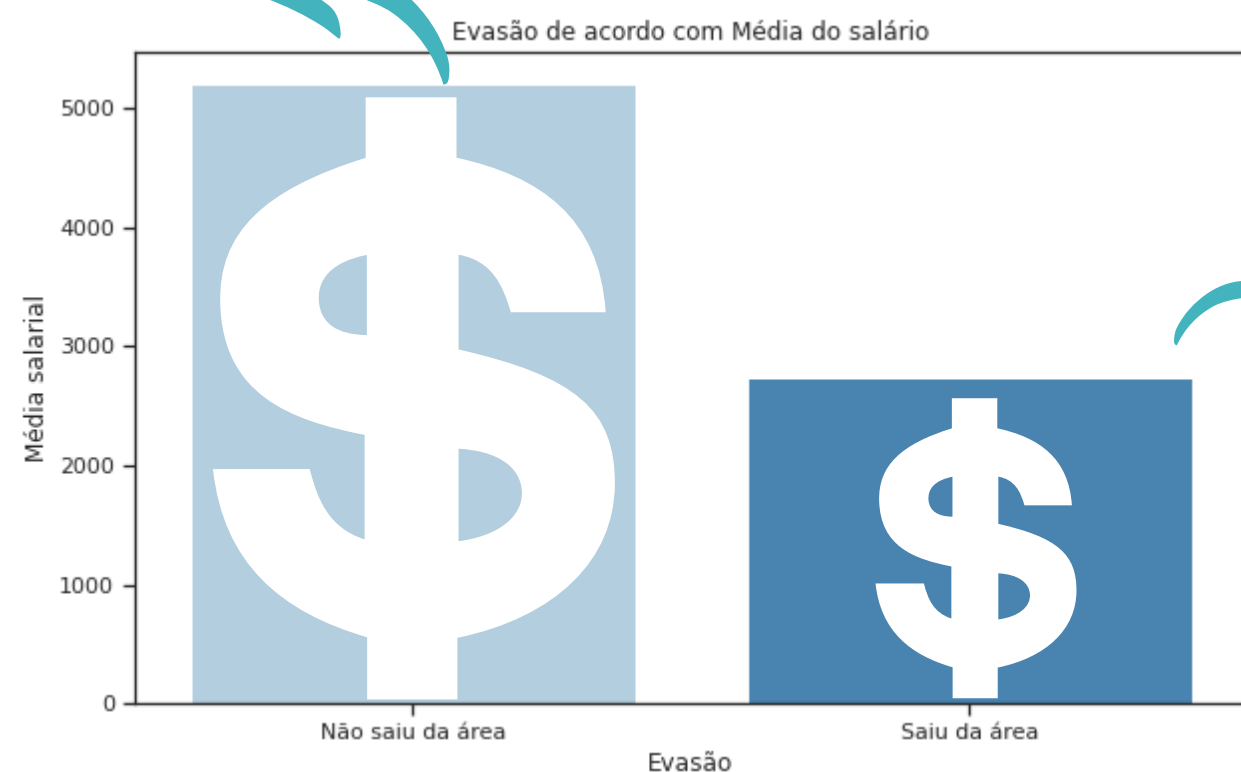


DADOS

Salário mensal

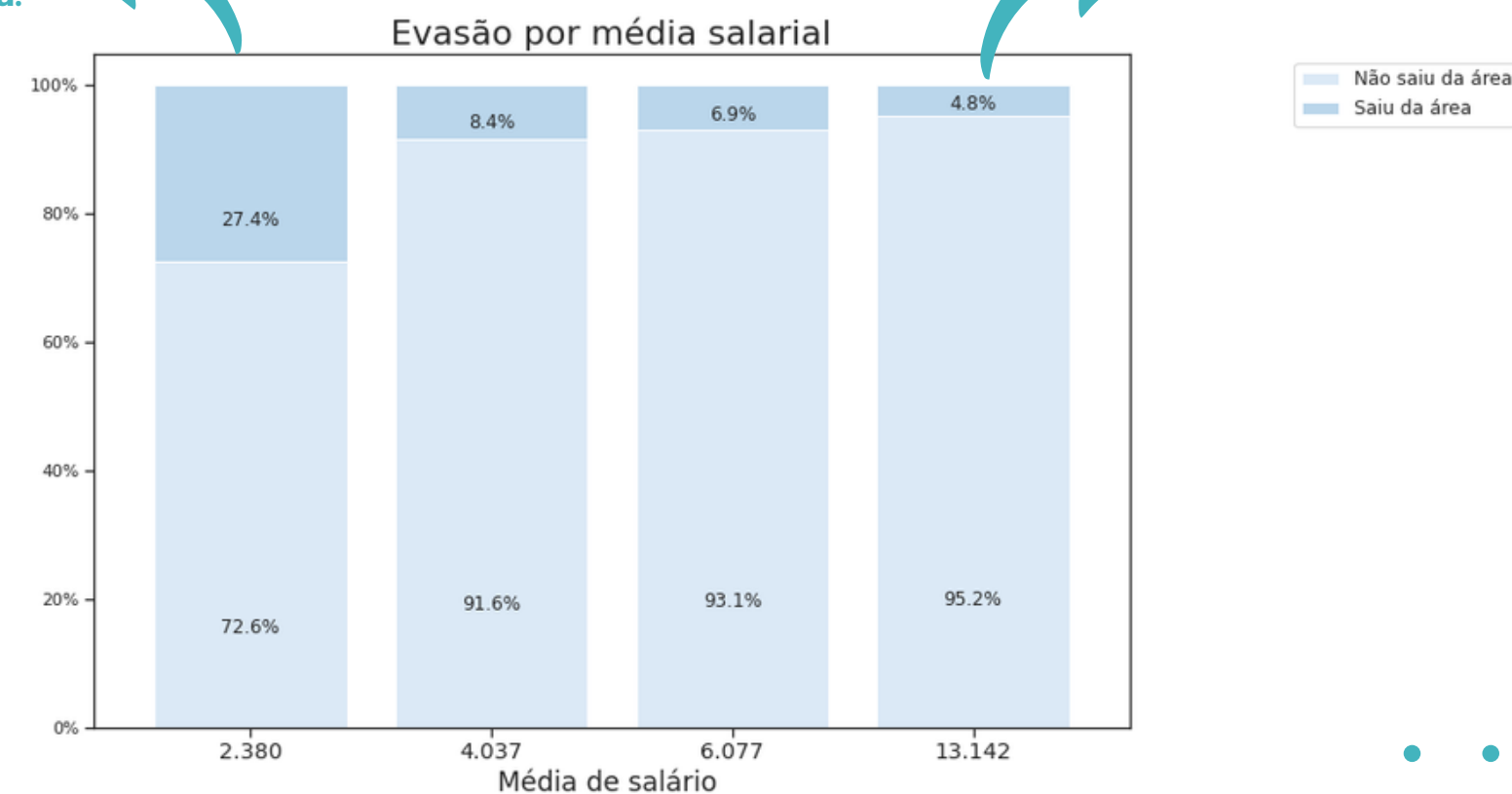
Por isso, pessoas mais novas e que ganham menos são as que apresentam as maiores taxas de evasão também.

Funcionários com uma média salarial de R\$ 5.204 tendem a não sair da área.



27,4% dos funcionários que recebem até R\$ 2.900 tendem a sair da área.

Funcionários com uma média salarial de R\$ 2.741 tendem a sair da área.

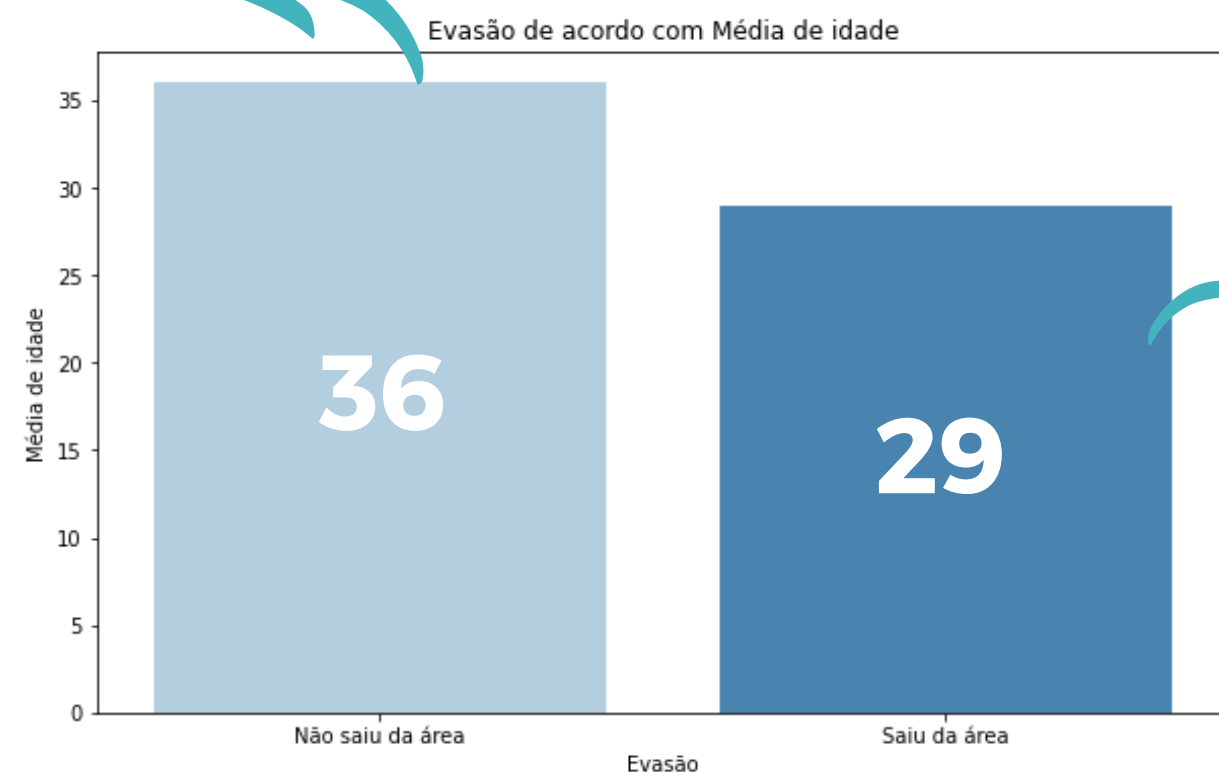


Apenas 4,8% dos funcionários que recebem mais de R\$ 8.400 tendem a sair da área.

Idade

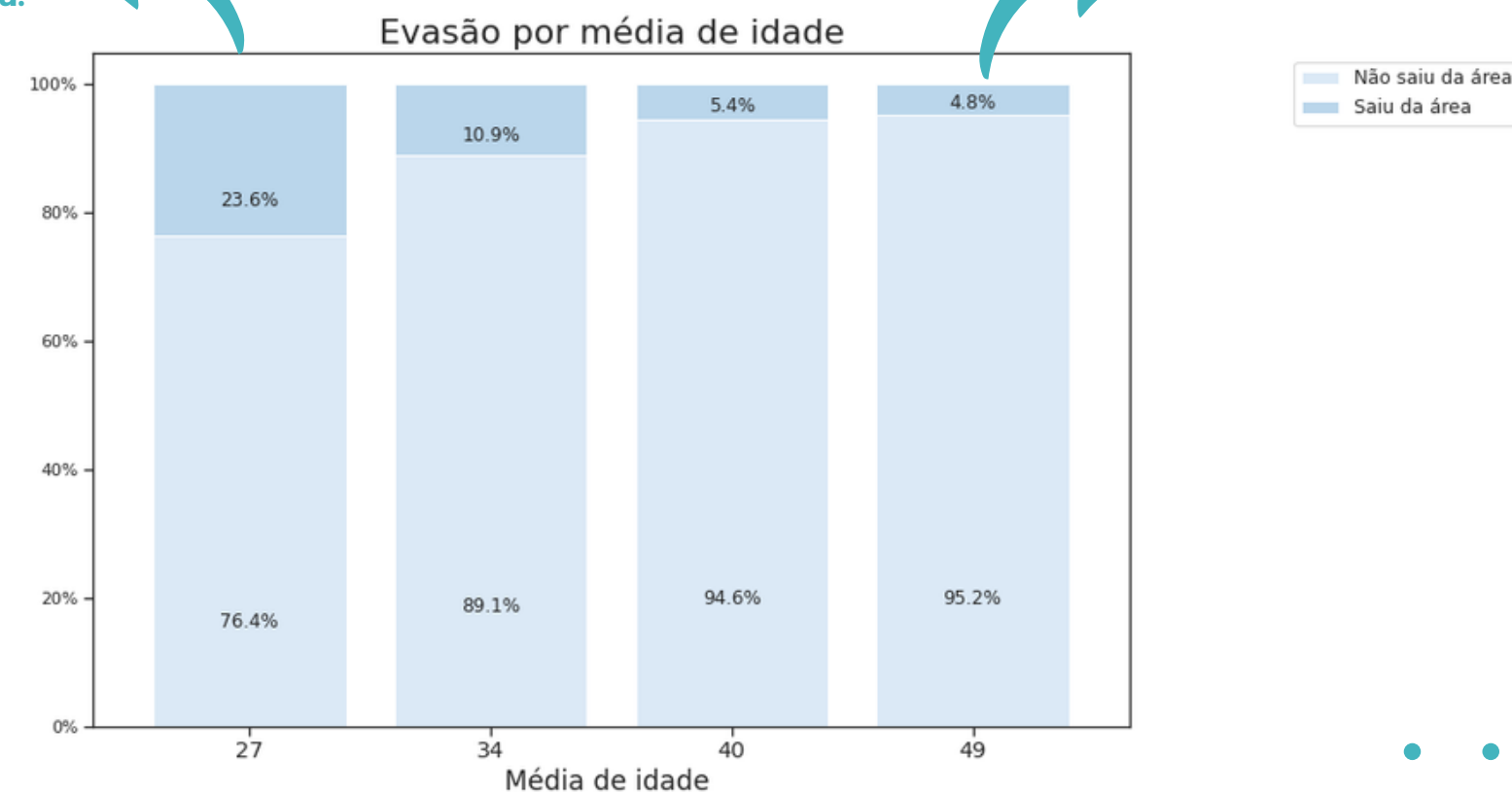
Por isso, pessoas mais novas e que ganham menos são as que apresentam as maiores taxas de evasão também.

Funcionários com uma média de idade de 36 anos tendem a não sair da área.



23,6% dos funcionários que têm entre 18 e 30 anos tendem a sair da área.

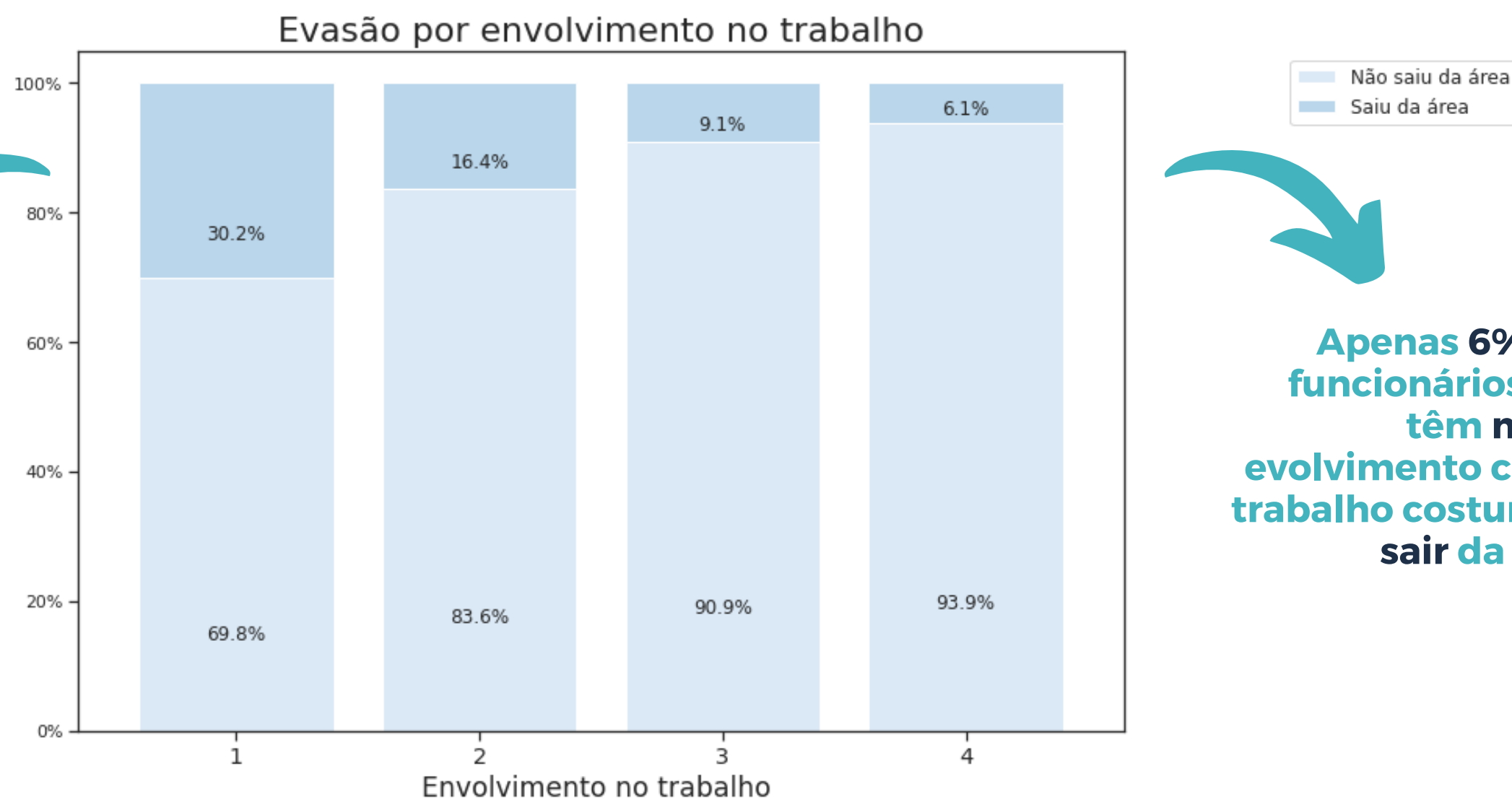
Funcionários com uma média de idade de 29 anos tendem a sair da área.



Apenas 4,8% dos funcionários que têm mais de 43 anos tendem a sair da área.

Envolvimento no trabalho

Quanto mais envolvidos no trabalho...
Menor a chance de evasão!



30% dos funcionários que têm pouco envolvimento com o trabalho costumam sair da área!

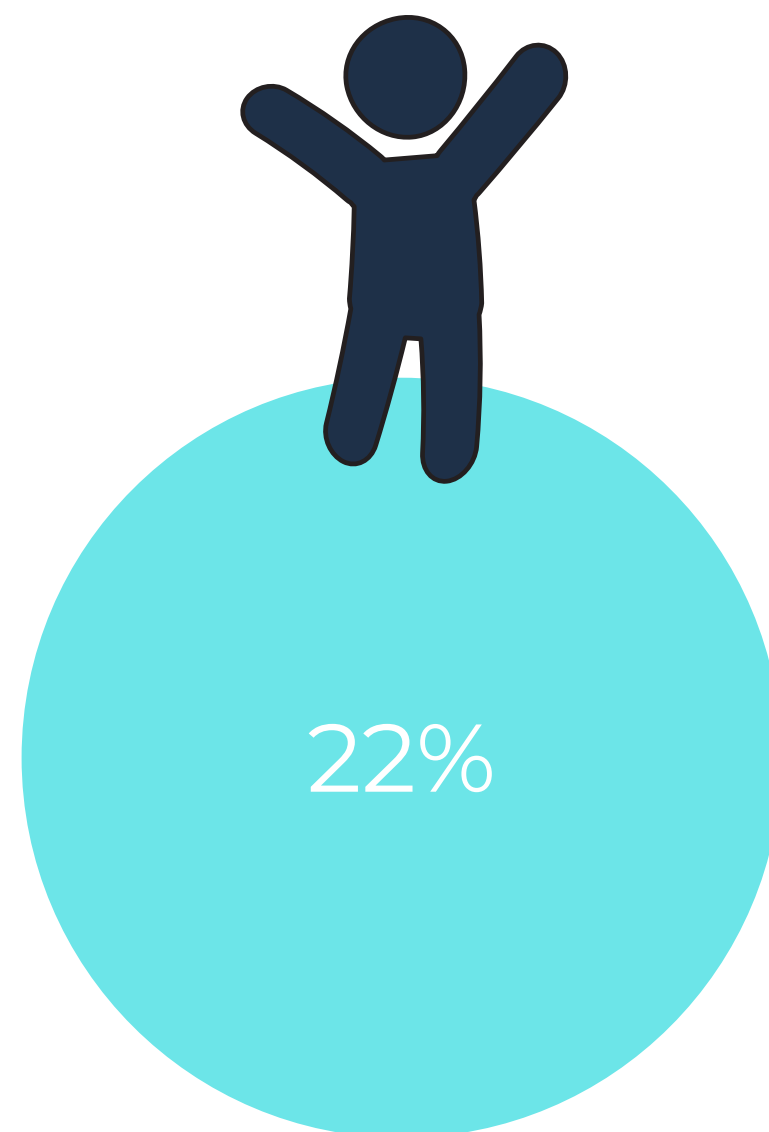
Apenas 6% dos funcionários que têm muito envolvimento com o trabalho costumam sair da área!



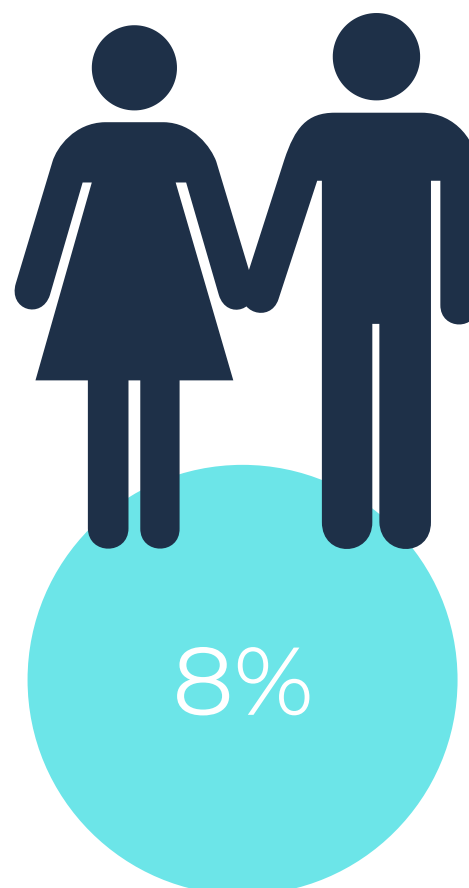
Estado civil

Por essa ninguém esperava!

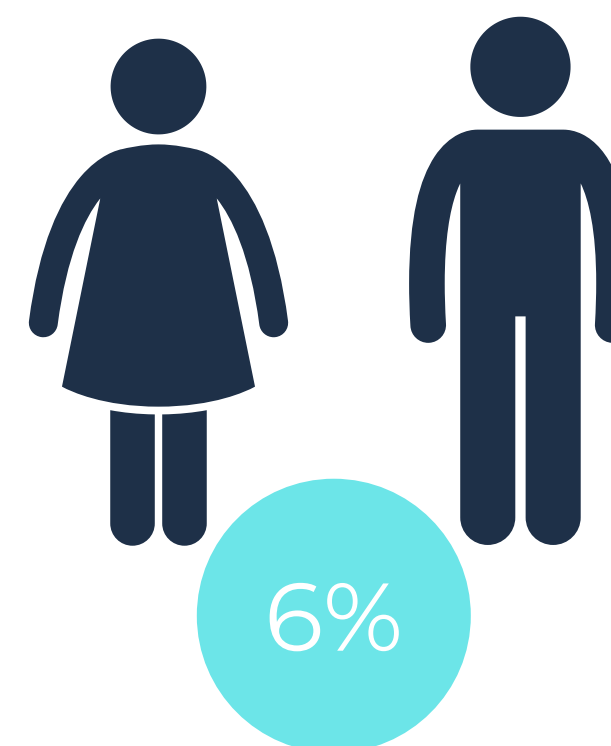
SOLTEIROS



CASADOS



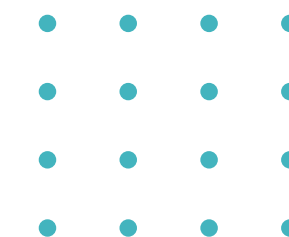
DIVORCIADOS



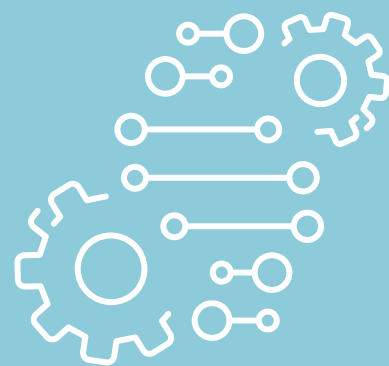
Será que é porque solteiros se sentem mais livres e casados têm mais responsabilidades financeiras?

Distância do trabalho

Evasão dos funcionários de acordo com a distância de casa para o trabalho.



03



FEATURE ENGINEERING





Separação em treino e teste

Remoção de variáveis correlacionadas

Normalização dos dados

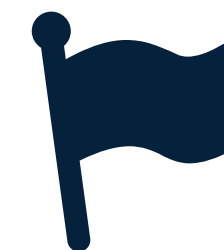
Separação de feature e target

Alteração de variáveis categóricas

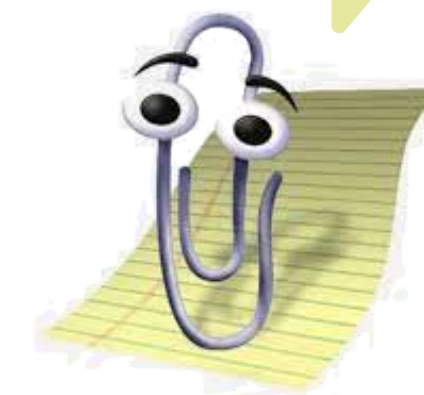
Remoção de variáveis

Criação de variáveis

Dados brutos



Parece que você se
esqueceu de balancear os
dados.
Você quer ajuda ?





FEATURE ENGINEERING

Criação de Variáveis

Starting Age

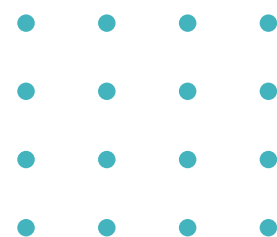
Idade em que o funcionário iniciou o trabalho na empresa:
 $\text{Idade} - \text{Anos na empresa}$

Overall Satisfaction

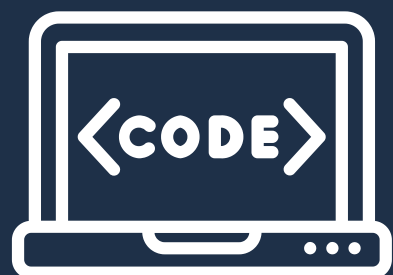
Satisfação geral do funcionário:
Média entre Satisfação no ambiente + Satisfação no trabalho + Satisfação nos relacionamentos

Ratio Job Involvement & Satisfaction

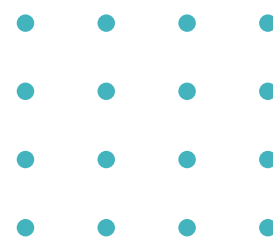
Razão entre nível de envolvimento no trabalho e Satisfação no trabalho



04

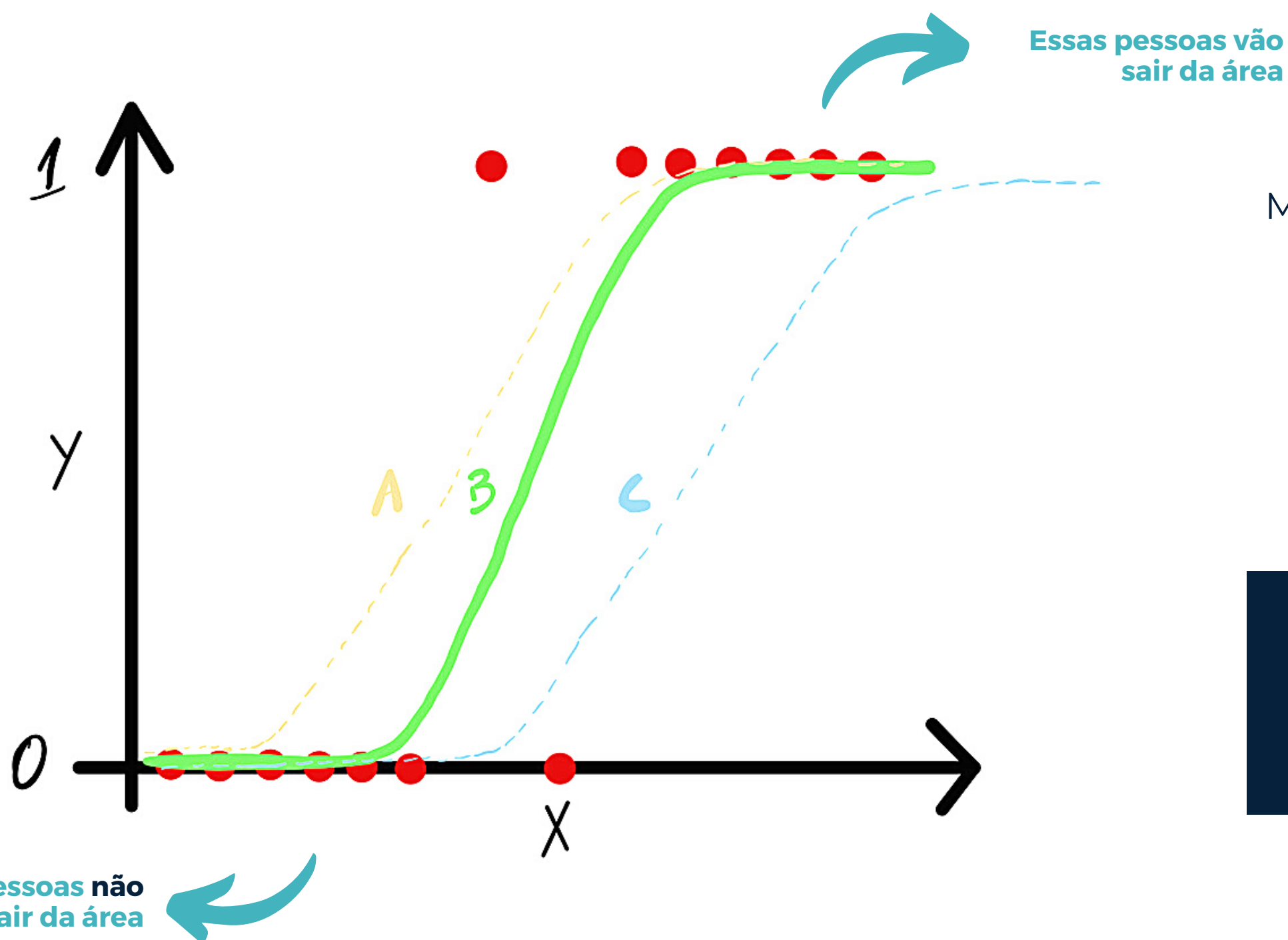


MODELO PREDITIVO





MODELO PREDITIVO

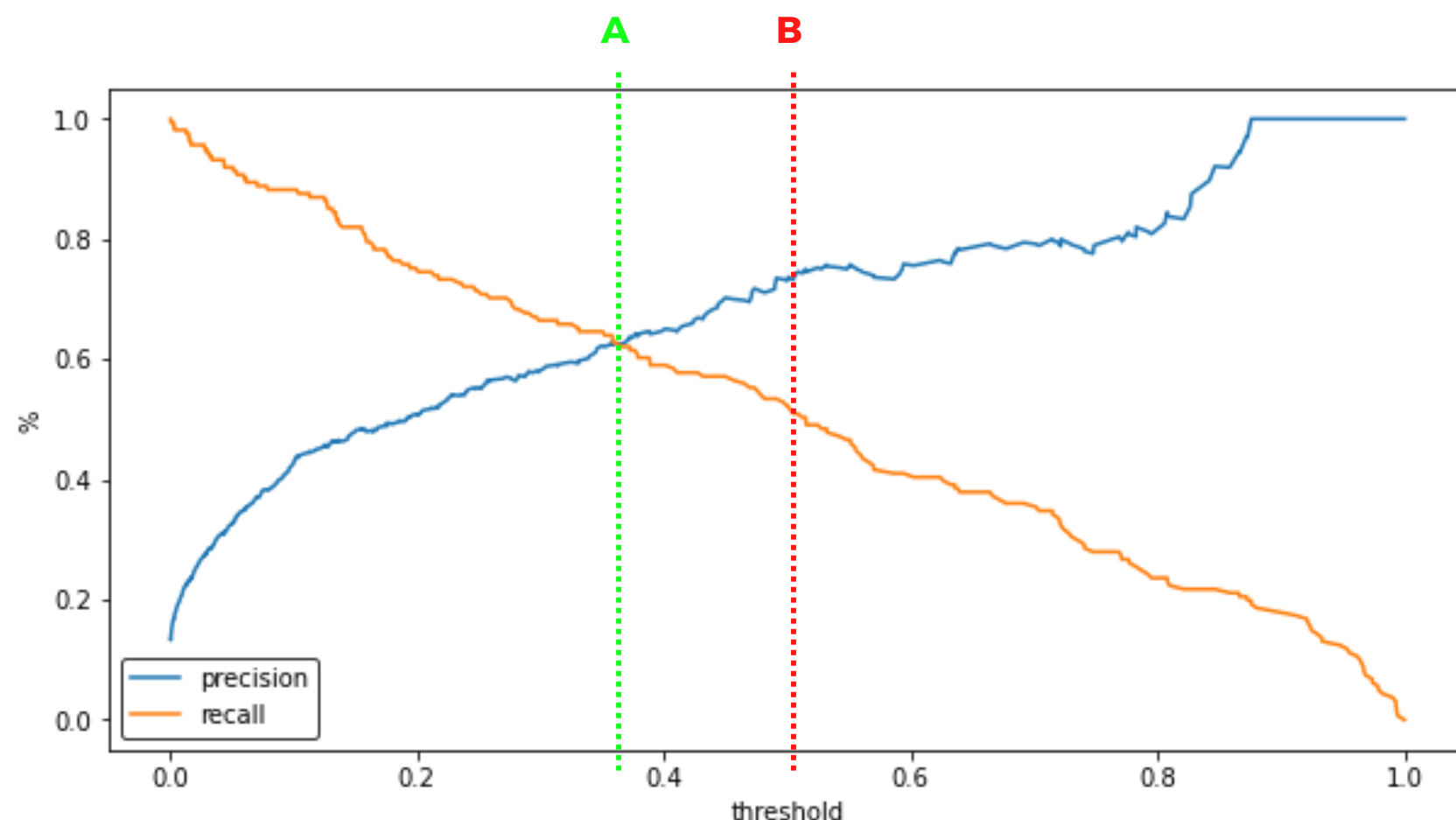


Regressão Logística

Modelo estatístico que mostra a relação entre variáveis e calcula a probabilidade de um determinado resultado binário: 0 ou 1.

Otimização do modelo: Ajuste da curva.

$$P(Y=1) = \frac{1}{1+e^{-g(x)}}$$



O melhor threshold foi obtido a partir do gráfico no ponto em que precisão e recall se iguallassem para que o modelo ficasse mais equilibrado

Ajuste do Threshold

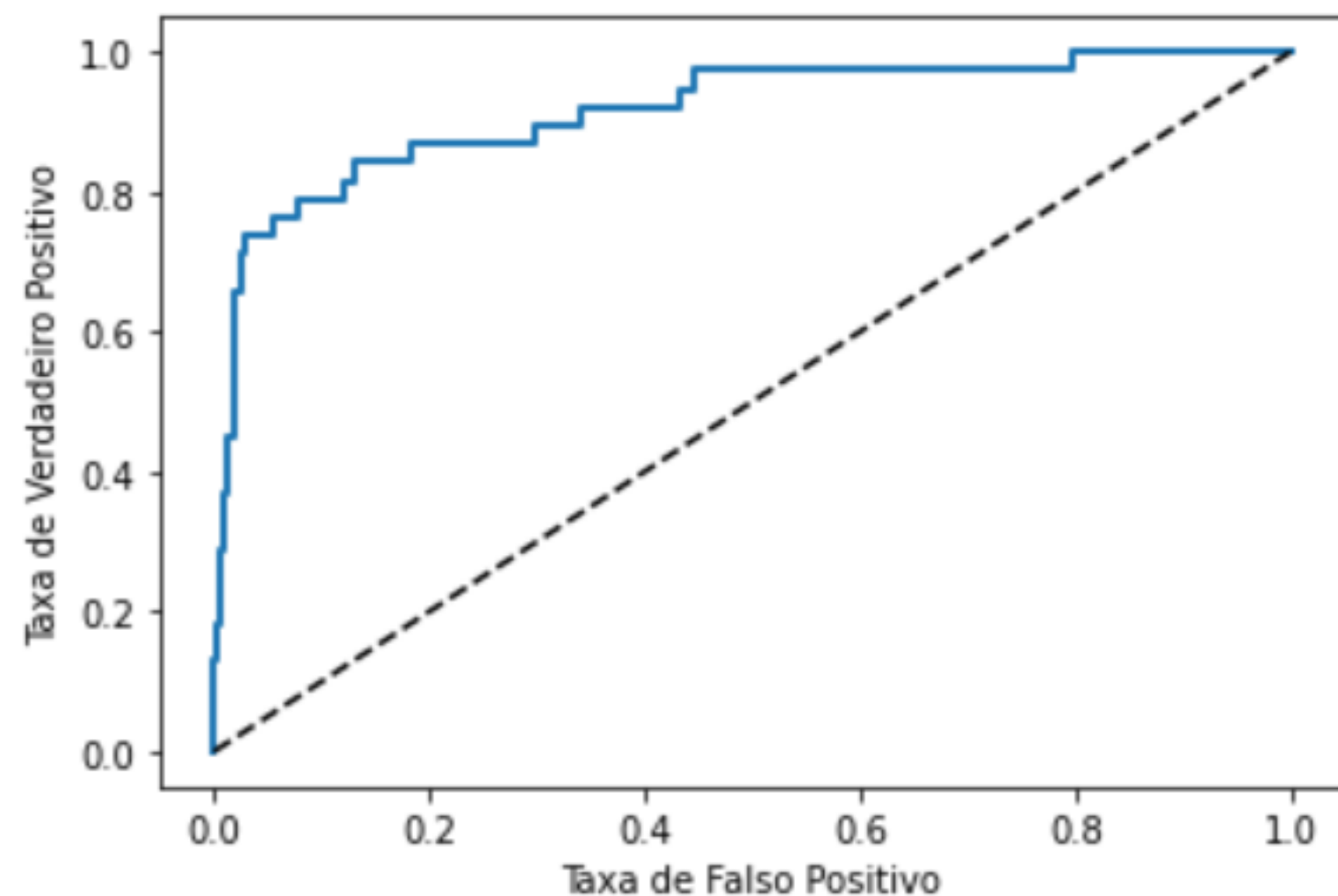
Recall: Falsos Negativos são considerados mais prejudiciais que os Falsos Positivos

Precisão: Falsos Positivos são considerados mais prejudiciais que os Falsos Negativos

Threshold: determina o limite da escolha de transformar uma probabilidade em uma classe.

Priorizamos equilibrar recall e precisão pois, assim, atingimos melhores resultados. Porém, para análises futuras, cabe penalizar mais os falsos negativos.

Resultados e métricas

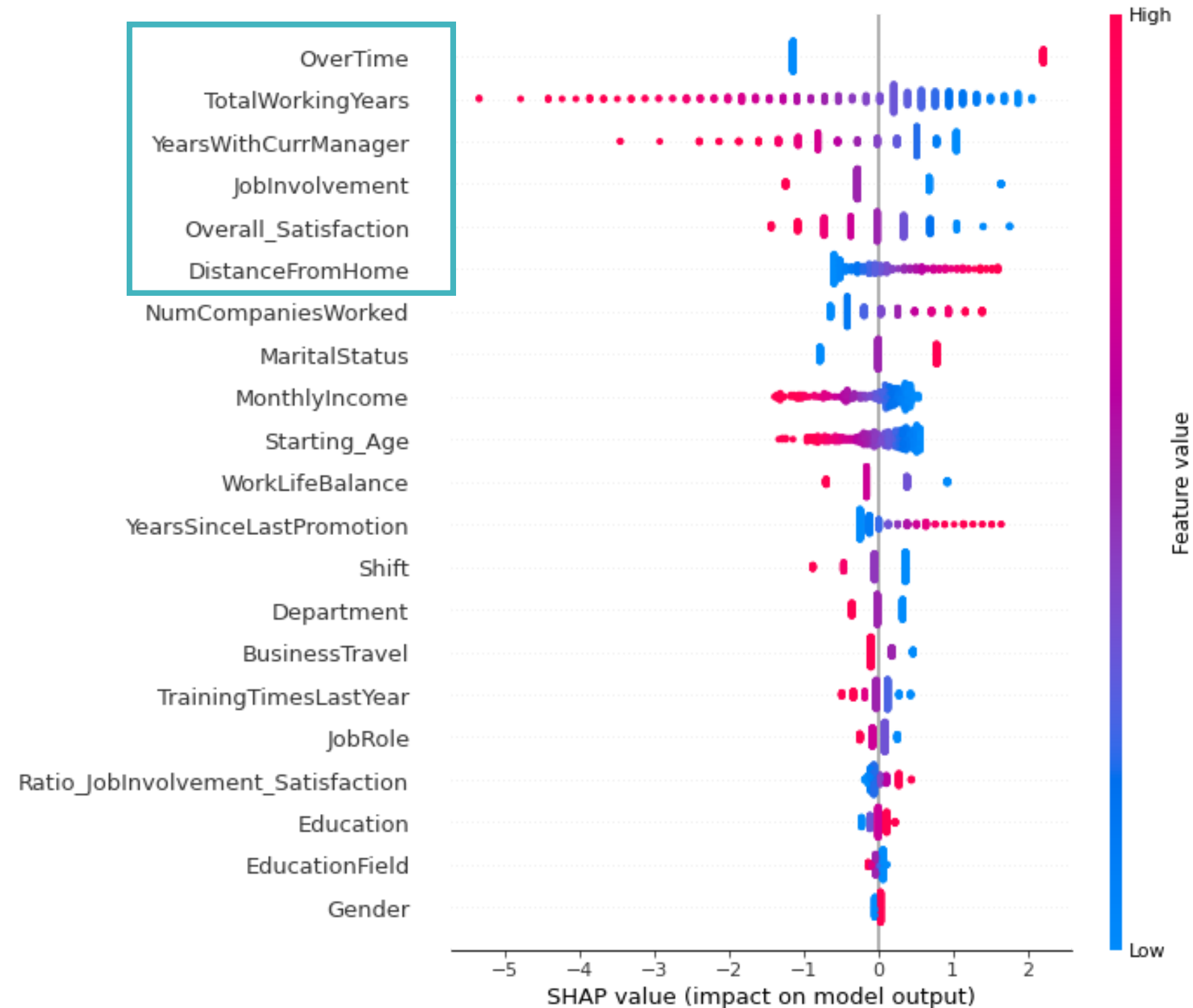


Área Sob a Curva ROC (AUC): 0.9149

Treino: 0,92
Teste: 0,91

	Antes dos ajustes	Depois dos ajustes
RECALL	0,76	0,74
PRECISÃO	0,55	0,74
F1 SCORE	0,64	0,74
ACURÁCIA		0,94

Analizando o peso de cada feature



Valores grandes de "Horas Extras" levam a uma maior evasão, enquanto valores pequenos de "Total de Anos Trabalhados" levam a uma maior evasão.

Analizando o peso de cada feature

SAÍDA

NÃO FAZER HORA EXTRA
↑ TEMPO DE TRABALHO EM ANOS
↑ TEMPO COM O MESMO GERENTE
↑ ENVOLVIMENTO COM O TRABALHO
↑ SATISFAÇÃO COM O TRABALHO

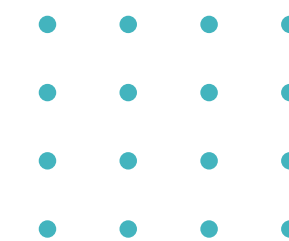


NÃO



SIM

FAZER HORA EXTRA
↓ TEMPO DE TRABALHO EM ANOS
↓ TEMPO COM O MESMO GERENTE
↓ ENVOLVIMENTO COM O TRABALHO
↓ SATISFAÇÃO COM O TRABALHO



Analizando as probabilidades

QUAIS FORAM OS FATORES RELEVANTES
PARA A SUA SAÍDA?

	id	Target	prob_0	prob_1
1659	1622091	1	0.07	0.93
1252	1196870	1	0.11	0.89
759	1631537	1	0.12	0.88
102	1088304	1	0.12	0.88
718	1034220	1	0.12	0.88
...
1051	1705514	0	1.00	0.00
697	1376209	0	1.00	0.00
408	1499671	0	1.00	0.00
231	1470356	0	1.00	0.00
1007	1793695	0	1.00	0.00

336 rows x 4 columns

Funcionário 1622091
Chance de sair da área:

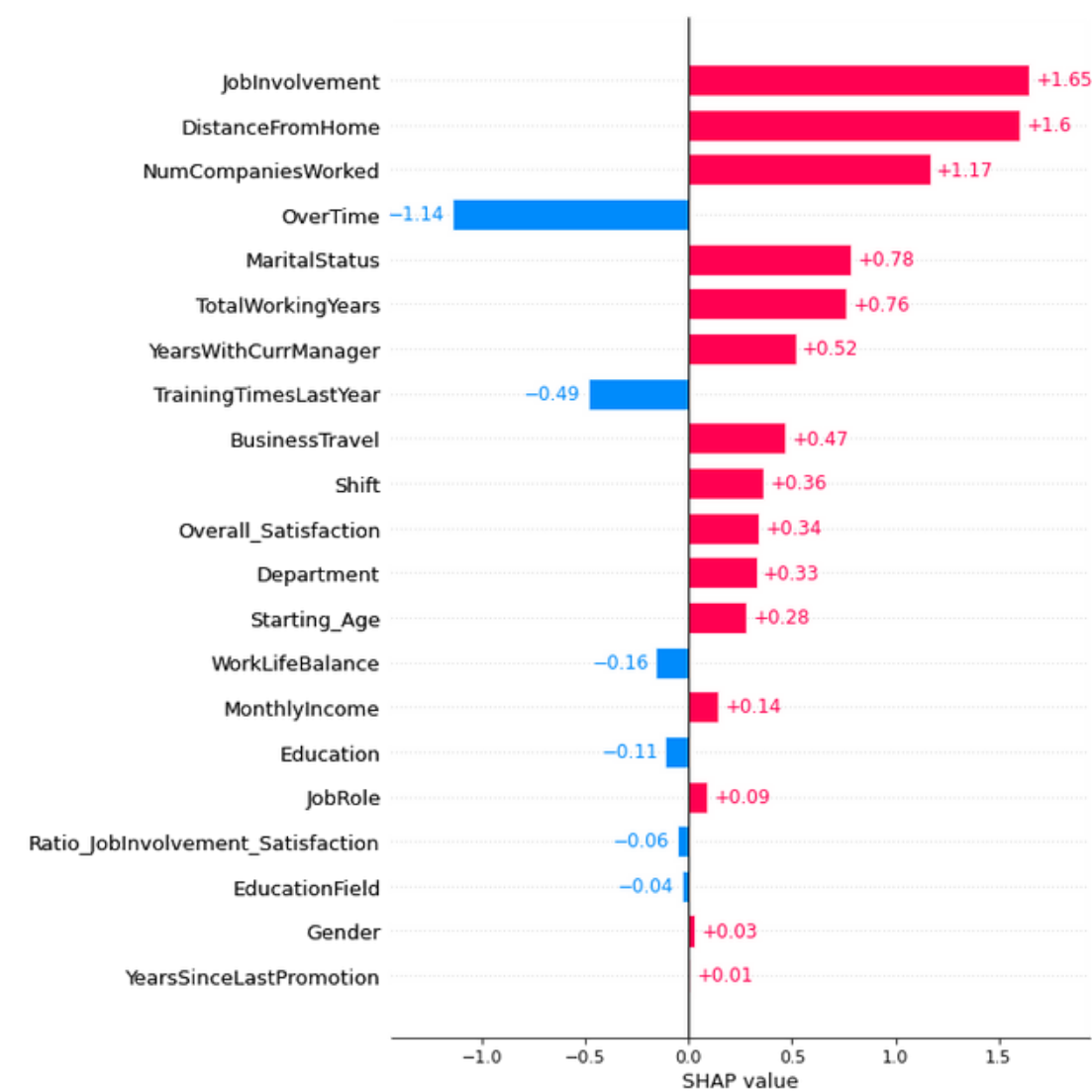
93%

Chance de não sair da área:

7%

Saiu da área?

SIM



05



CONCLUSÕES



O que os dados nos dizem?

▶ Horas extras são o fator que mais influencia na decisão do funcionário de sair da área: funcionários que realizam horas extras tendem a sair.

▶ Fatores relacionados à idade e anos na área também são relevantes para a análise: os mais novos e que têm menos tempo na área também tendem a sair.

▶ Outros fatores relevantes são: salário, nível envolvimento no trabalho, estado civil e distância do trabalho à casa.

Próximos passos

▶ Fatores importantes como: localização do funcionário, benefícios, classe social, informações sobre a família, quantidade de horas extras, podem ser fatores decisivos para a questão.

▶ O modelo de machine learning deve ser melhorado no futuro, considerando diferentes parâmetros para que se penalize mais os falsos negativos.

▶ Um dataset maior e com informações de mais funcionários tornaria o modelo preditivo mais assertivo.

06



NOSSOS LINKS





NOSSOS LINKS

OBRIGADA!

dataset

nosso LinkedIn

apresentação

código do modelo



Acessem nossos links! 😊

