

MACHINE FIVE.

GRUPO 5

www.reallygreatsite.com

MODELO PREDITIVO



Vamos!



Sumário

Progress Step
By Step



Dúvidas e Melhorias

Passos para a próxima
Sprint 4

Teste e modelagem

Apresentação da acurácia
e precisão do modelo

Análise de Dados

Gráficos de Dispersão

Algoritmo utilizado


Explicação do SVM



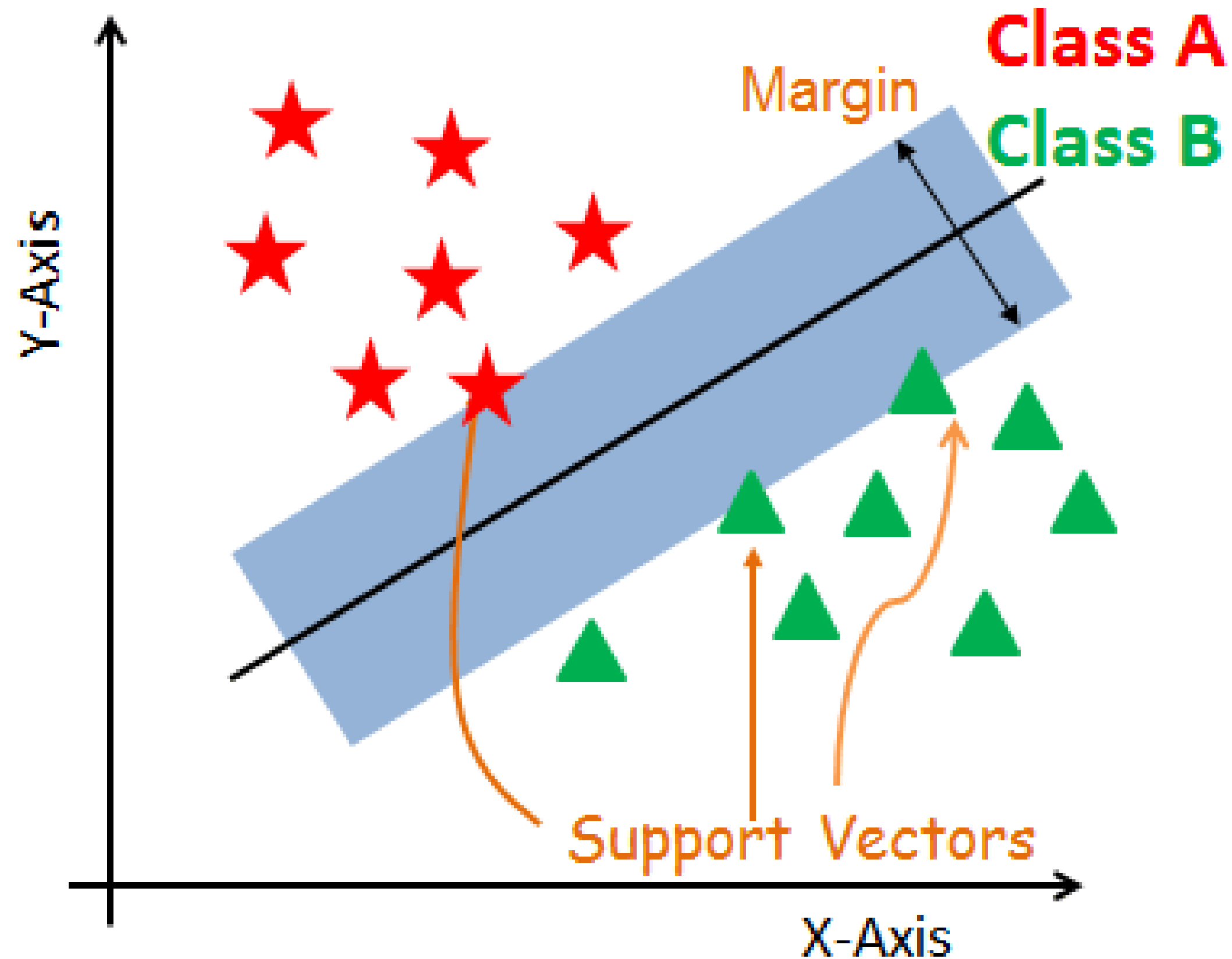
Eficaz em espaços de alta dimensão (várias variáveis).

Ainda eficaz nos casos em que o número de dimensões é maior que o número de amostras.

É bom comportamento em base de dados mais reduzida

- 
- Gênero
 - Salário Mês
 - Escolaridade
 - Idade

Modelo SVM(Support Vector Machines)



▼ Legendas

Tipo de Saída

- 0: Dispensa sem justa causa
- 1: Pedido de demissão
- 2: Recisão contrato - Exp/Dispensa (Saiu em menos de 3 meses por opção da empresa)
- 3: Recisão contrato - Exp/Pedido (Saiu em menos de 3 meses por pedido)
- 4: Não saiu

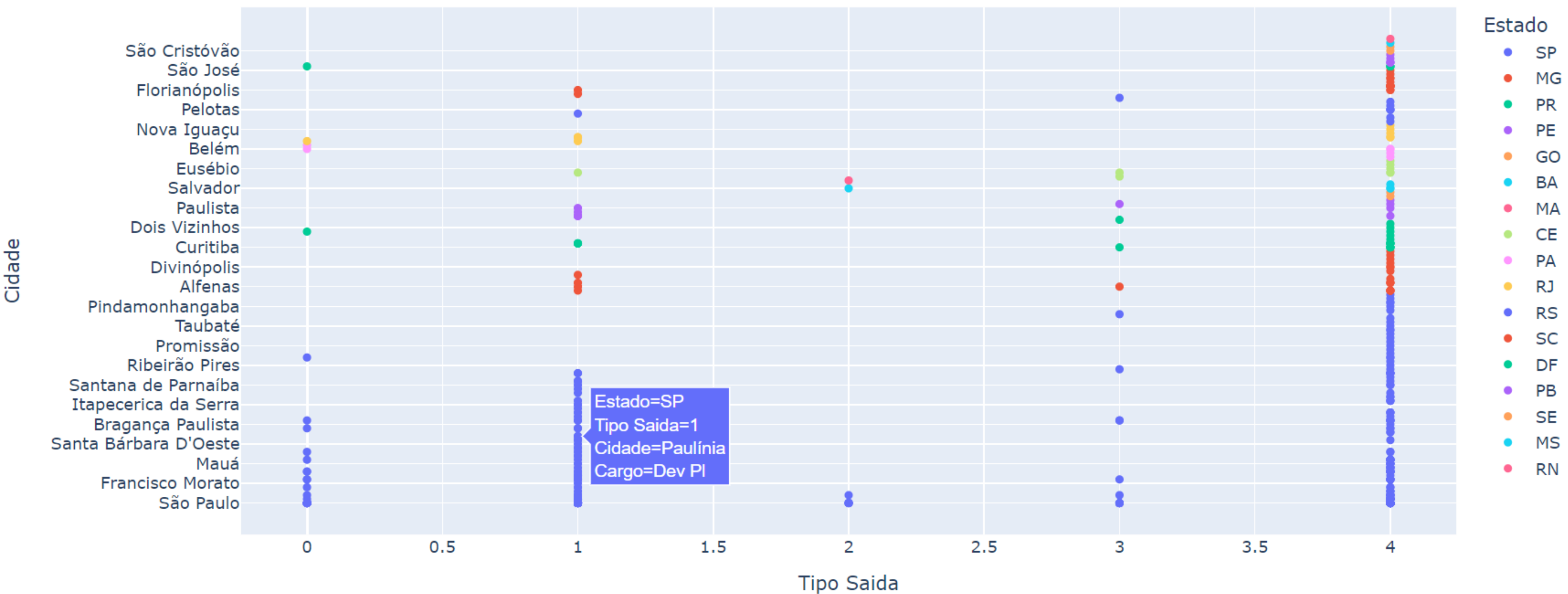
Escolaridade

- 0: Ensino médio
- 1: Ensino médio incompleto
- 2: Graduação
- 3: Mestrado
- 4: Pós graduação
- 5: Superior incompleto
- 6: Técnico

Estado civil

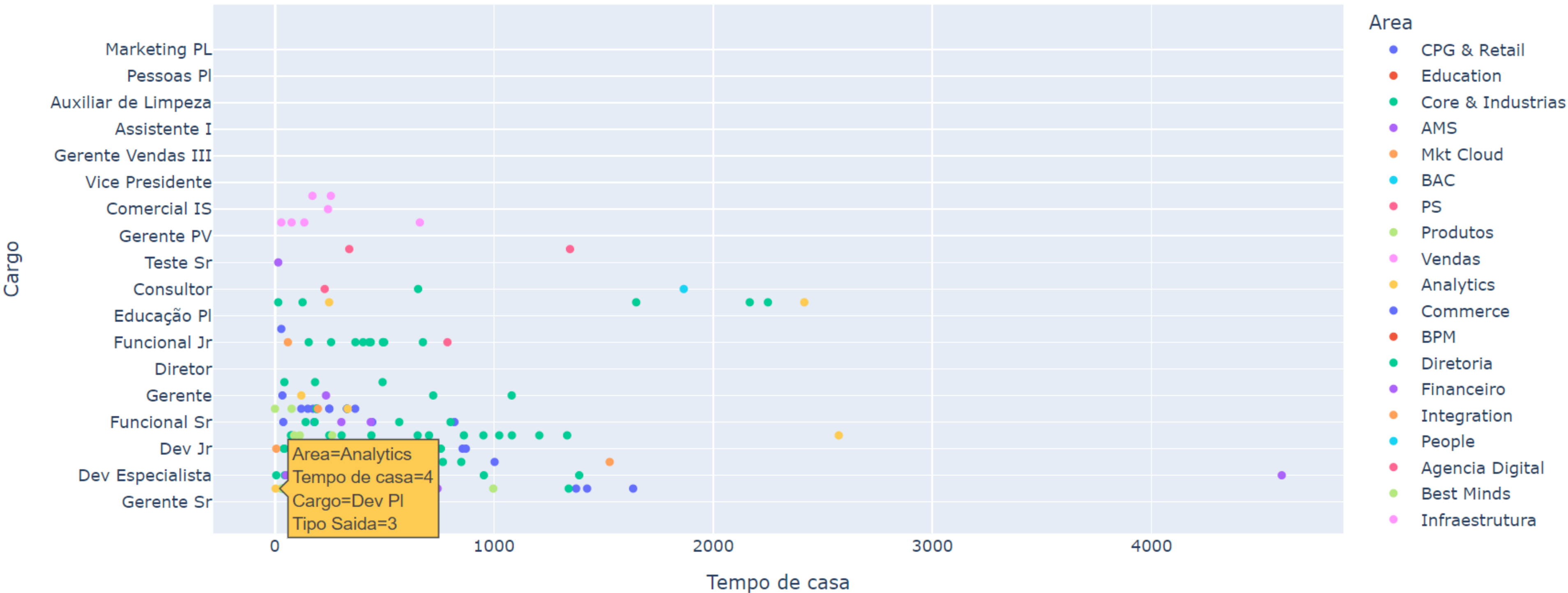
- 0: Casado
- 1: Divorciado
- 2: Separado
- 3: Solteiro
- 4: União estável

Tipo de Saída vs Cidade e Estado



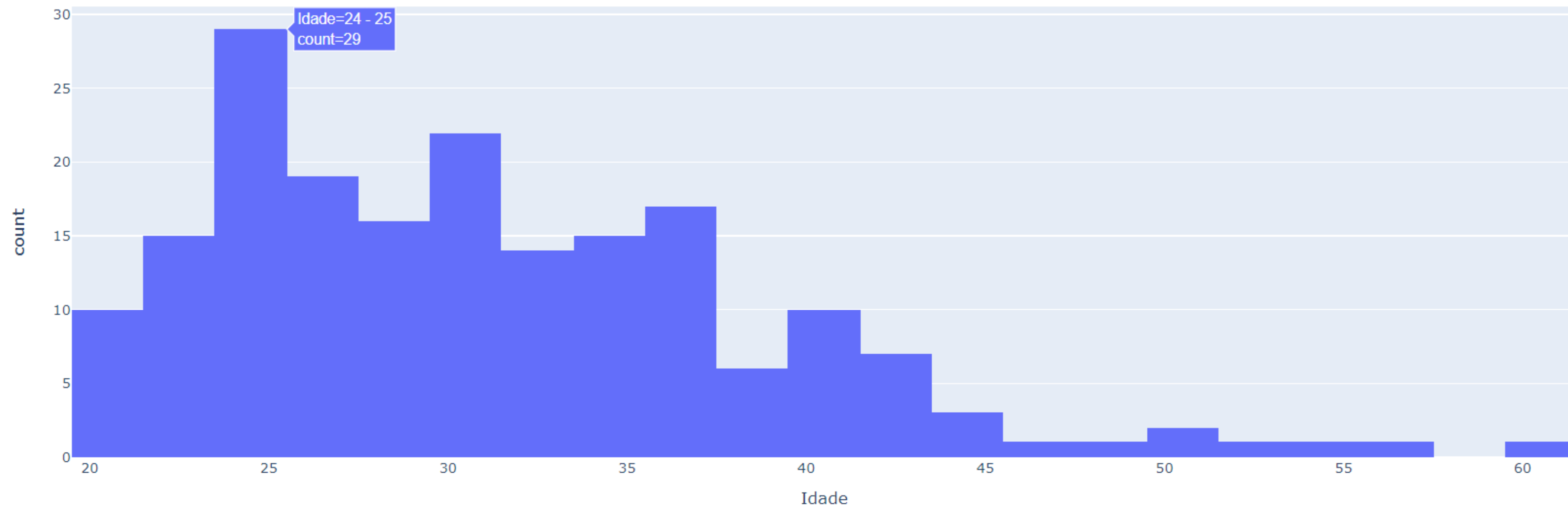
Tipo de Saída "1" - Pedido de Demissão

Tempo de Casa vs Cargo e Área

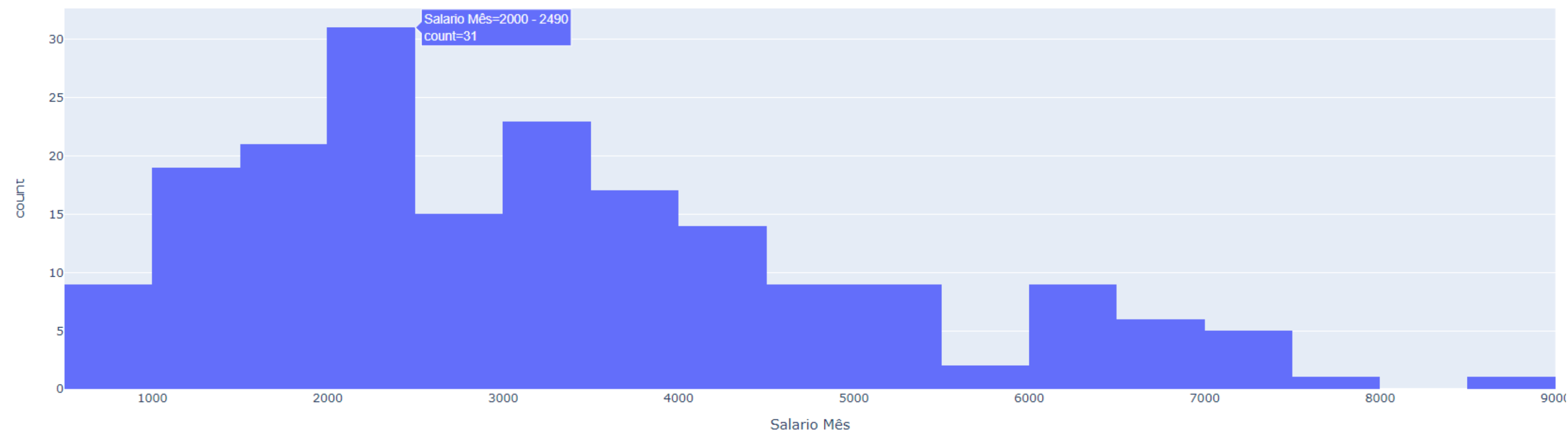


Tipo de Saída "3" - Recisão contrato - Exp/Pedido (Saiu em menos de 3 meses por pedido)

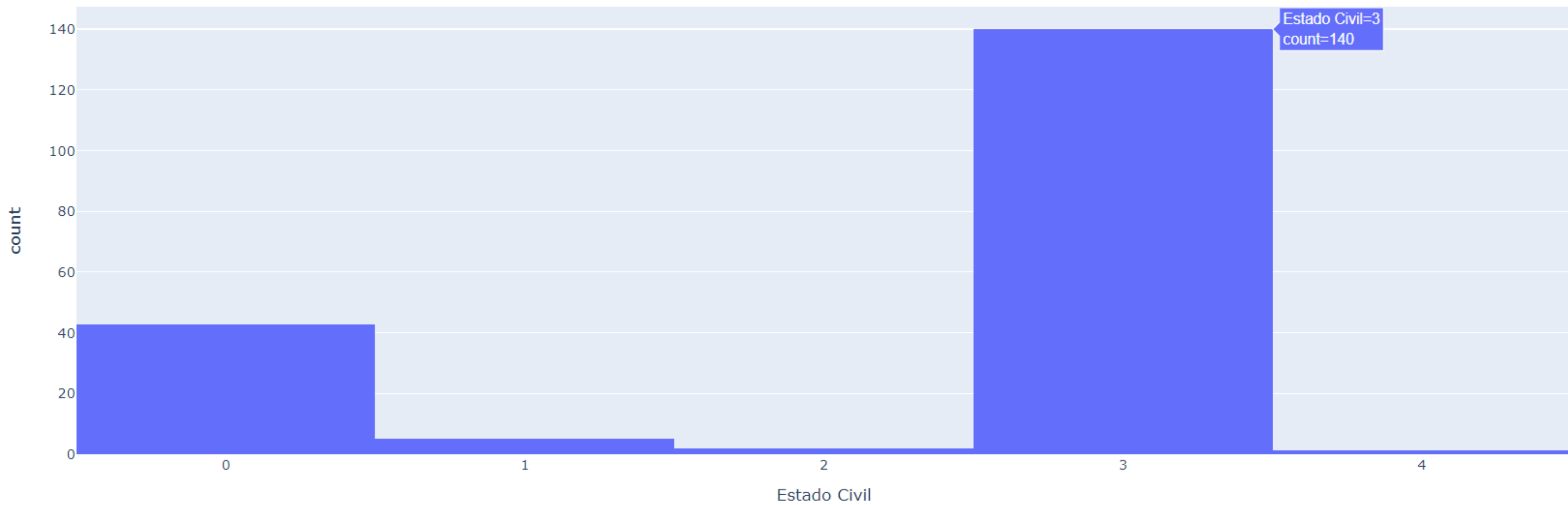
Faixa etária de pessoas desligadas



Salário mês dos funcionários inativos



Estado Civil pessoas inativas



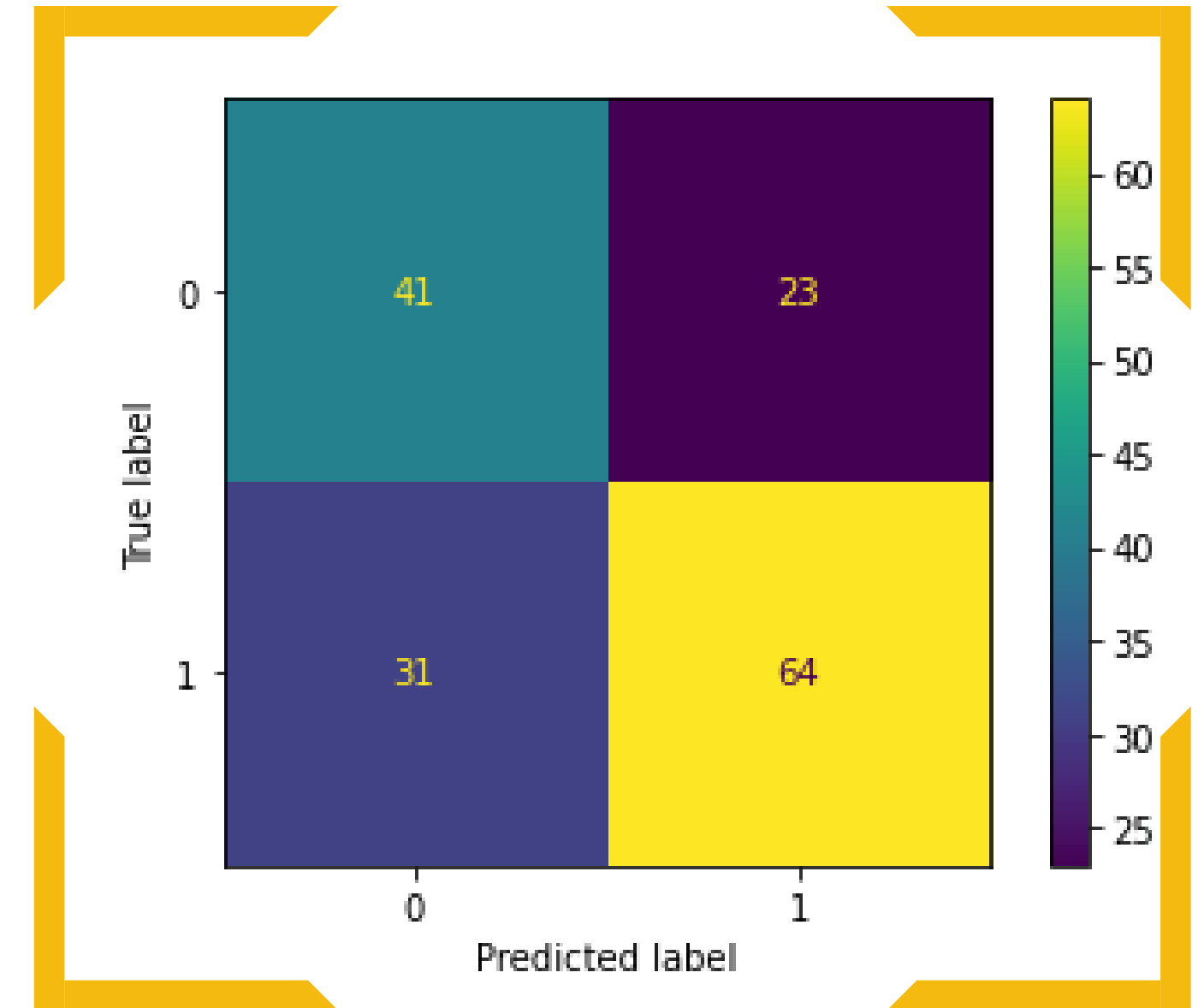
Métricas de Avaliação

- Acurácia: 66%
- Revogação: 67%
- Precisão: 73%

Porcentagem de dados utilizados:

- Pessoas que já pediram demissão: **50%**
- Pessoas que trabalham na empresa atualmente: **50%**

Matriz de Confusão



- Acurácia: indica uma performance geral do modelo. Dentre todas as classificações, quantas o modelo classificou corretamente;
- Precisão: dentre todas as classificações de classe Positivo que o modelo fez, quantas estão corretas;
- Revocação: dentre todas as situações de classe Positivo como valor esperado, quantas estão corretas;

Acurácia	$\frac{VP + VN}{VP + VN + FP + FN}$
Precisão	$\frac{VP}{VP + FP}$
Revocação	$\frac{VP}{VP + FN}$
F1-Score	$2 * \frac{\text{Precisão} * \text{Revocação}}{\text{Precisão} + \text{Revocação}}$



Dúvidas e Melhorias para a próxima Sprint

Foco apenas na área "Core & Industrias" da empresa??

Atuação em algum cargo específico?

Adição de alguma amostra nova de "Etnia"?

Alguma variável