

CASE STONE

Análise de performance e quebra de uma
carteira de crédito

Arthur de Figueiredo Gusmão





01

TRATANDO OS DADOS

Resolvendo inconsistências
nos dados

02

SAÚDE DA CARTEIRA

Análise exploratória dos
dados

03

CLUSTERIZAÇÃO

Utilizando K-Means Elbow
Method

04

CONSIDERAÇÕES FINAIS

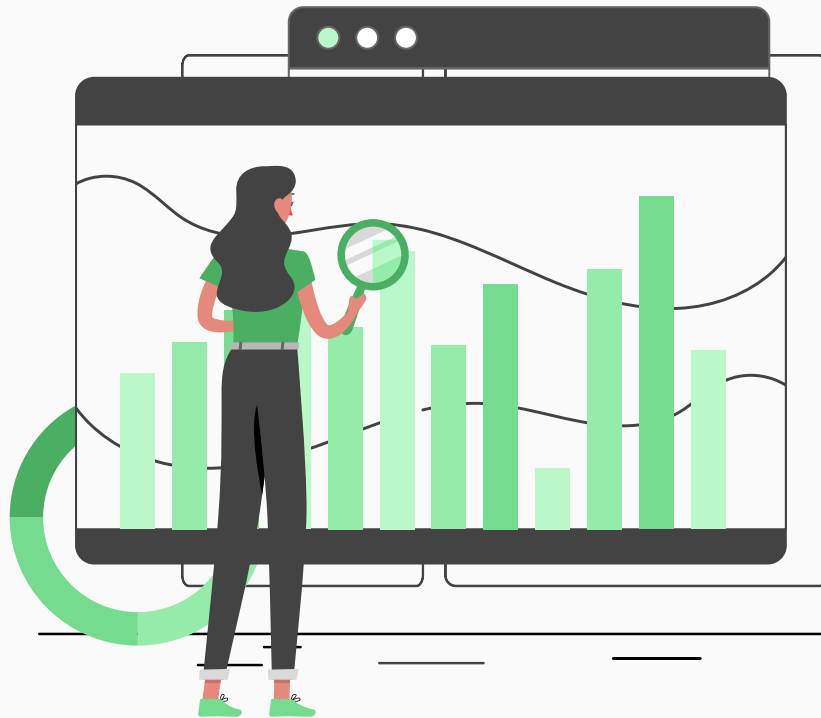
Insights e recomendações

INTRODUÇÃO



OI. TRATANDO OS DADOS

Resolvendo inconsistências e
erros de imputação



TRATANDO OS DADOS

PROBLEMA:

As colunas: **dsc_tipo empresa**, **geographykey**, **mcckey** possuem inconsistências

SOLUÇÃO:

Função **interpolate()** da biblioteca Pandas

RESULTADOS:

67791

É o total de NaN's na coluna **geographykey** antes da limpeza

22000

É o total de NaN's na coluna **geographykey** após a limpeza

969

É o total de clientes que estão sem dados nas colunas **dsc_tipo empresa**, **geographykey**, **mcckey**

67791

É o total de NaN's na coluna **mcckey** antes da limpeza

22000

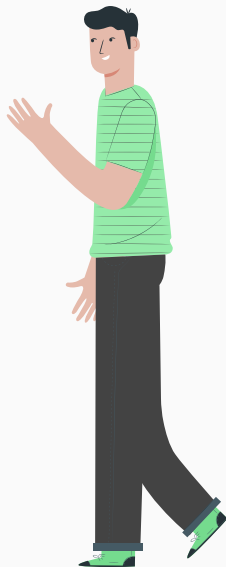
É o total de NaN's na coluna **mcckey** após a limpeza

18184

É o total de NaN's na coluna **dsc_tipo_empresa** antes da limpeza

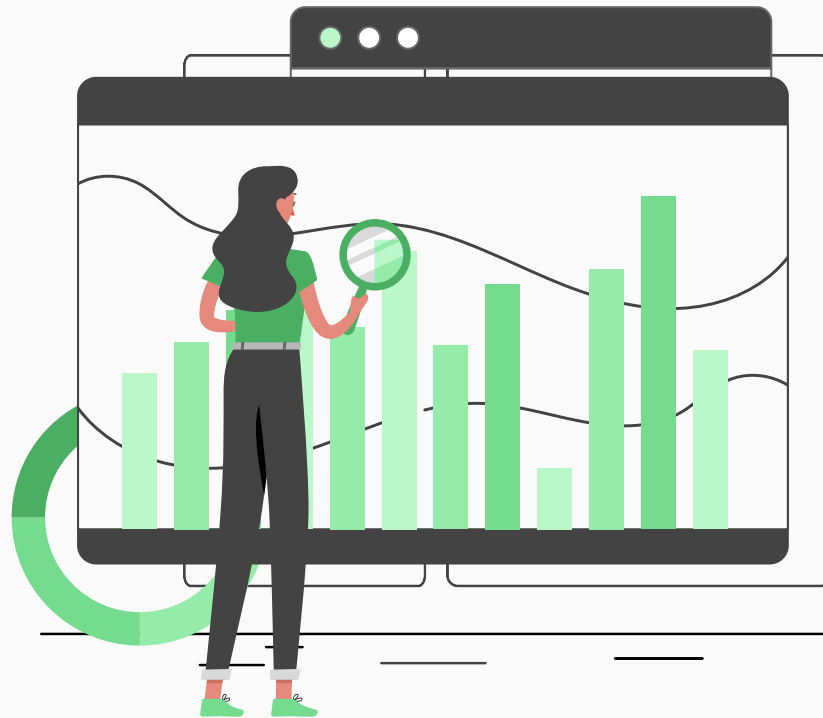
17890

É o total de NaN's na coluna **dsc_tipo_empresa** após a limpeza



02. SAÚDE DA CARTEIRA DE CRÉDITO

Análise exploratória dos
dados

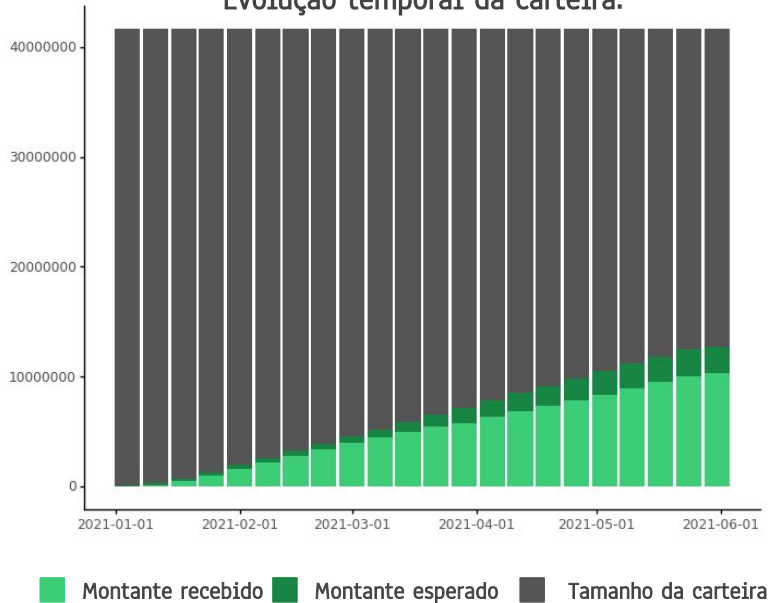


SAÚDE DA CARTEIRA DE CRÉDITO

Tamanho da carteira de crédito:

R\$: 41,640,459.59

Evolução temporal da carteira:



Principais estatísticas descritivas:

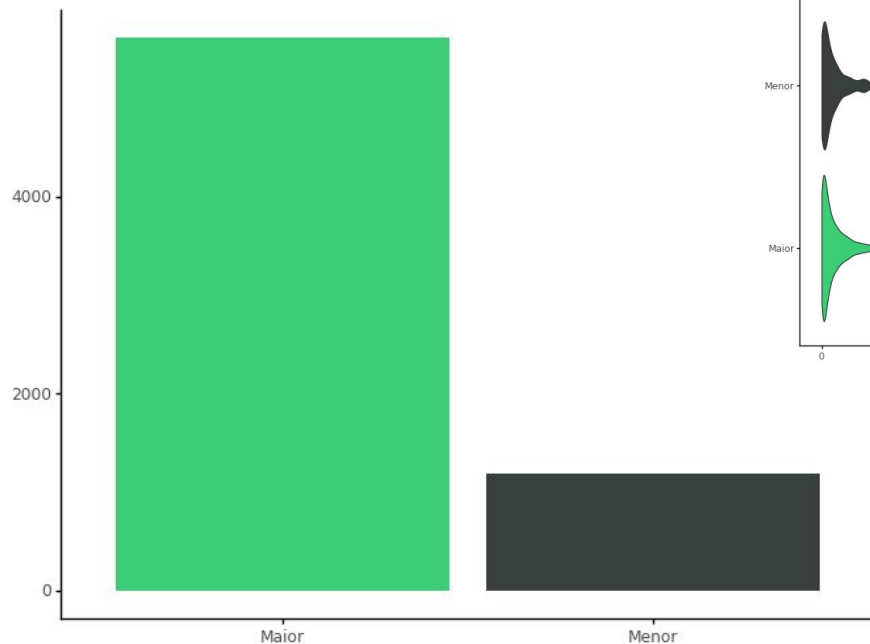
	PRAZO_MESES	VLR_DESEMBOLSO	PERC_RETENCAO	VLR_PMT_SEMANAL	VLR_PAGAMENTO_SEMANAL	VLR_TPV_ACUMULADO_7D	VLR_SD_OPERACIONAL_FINAL
MEAN	16,379417	R\$ 6.131,63	12%	R\$ 92,50	R\$ 75,12	R\$ 630,52	R\$ 5.612,69
MIN	6	R\$ 520,	1%	R\$ 0,00	-R\$ 135,75	-R\$ 4.018,81	-R\$ 148,48
0,25	15	R\$ 745,45	8%	R\$ 7,82	R\$ -	R\$ -	R\$ 594,91
0,5	15	R\$ 2.388,82	12%	R\$ 28,43	R\$ 9,03	R\$ 1,89	R\$ 2.060,10
0,75	18	R\$ 6.573,65	15%	R\$ 90,25	R\$ 58,29	R\$ 426,86	R\$ 5.858,78
MAX	21	R\$ 129.316,97	100%	R\$ 2.851,69	R\$ 35.095,55	R\$ 64.775,77	R\$ 133.227,14

SAÚDE DA CARTEIRA DE CRÉDITO

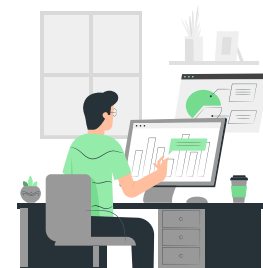
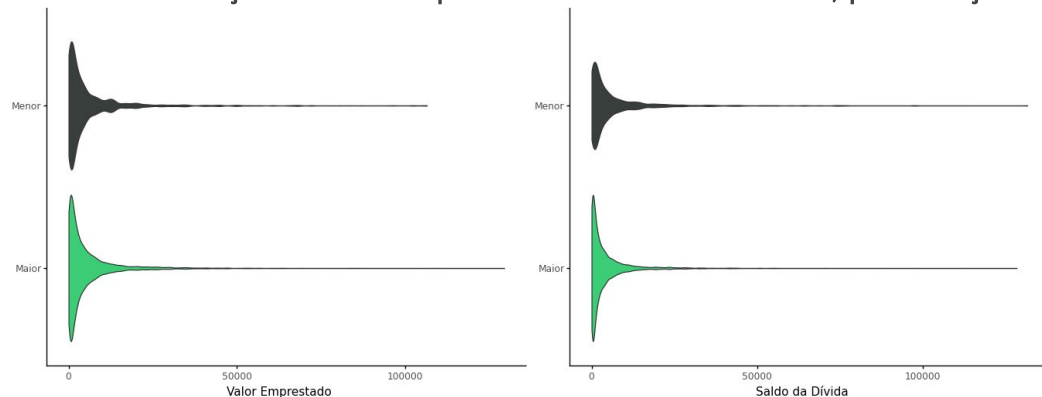
PERGUNTA CHAVE: OS CLIENTES ESTÃO COM O VALOR EMPRESTADO MAIOR OU MENOR QUE O SALDO DA DÍVIDA?

■ Maior ■ Menor

Número de clientes, por situação:

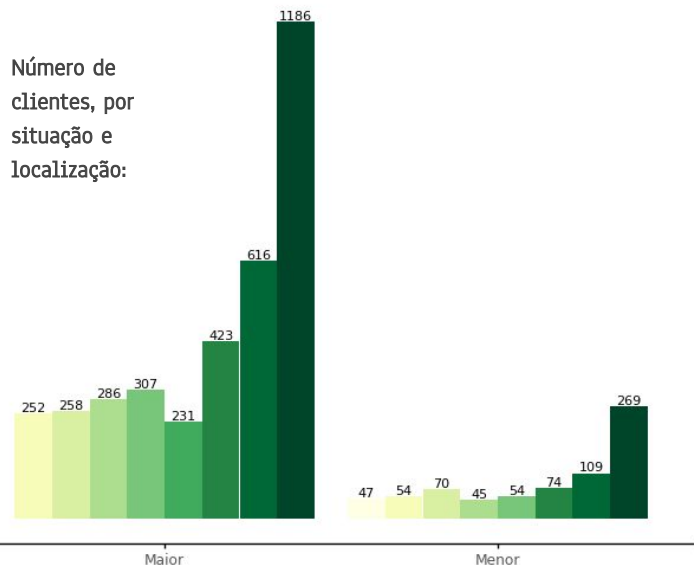


Distribuição do Valor emprestado e do Saldo da dívida, por situação:



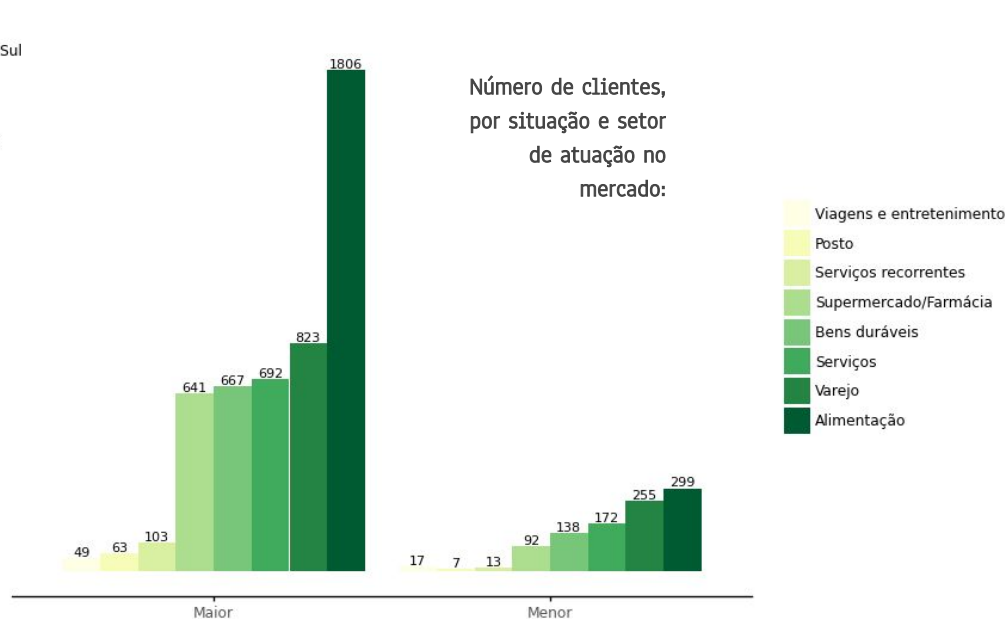
SETOR, LOCALIZAÇÃO?

Número de clientes, por situação e localização:



SAÚDE DA CARTEIRA DE CRÉDITO

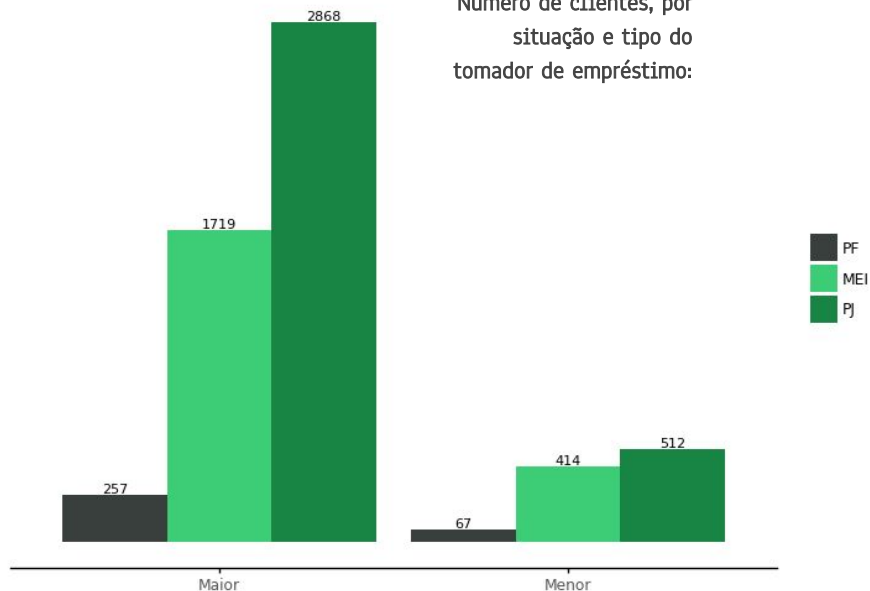
Número de clientes, por situação e setor de atuação no mercado:



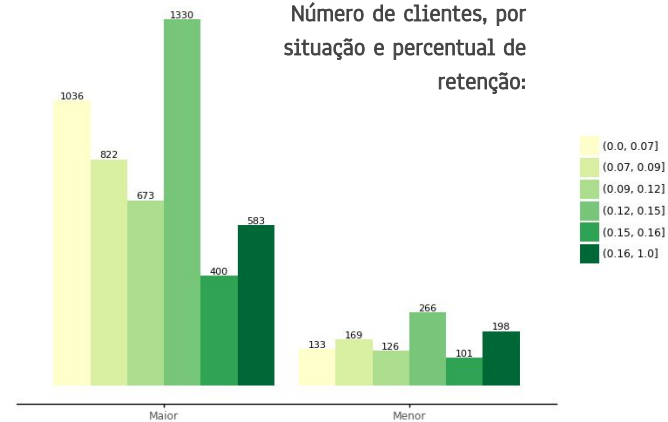
TIPO DE TOMADOR, PERCENTUAL DE RETENÇÃO, PRAZO?

SAÚDE DA CARTEIRA DE CRÉDITO

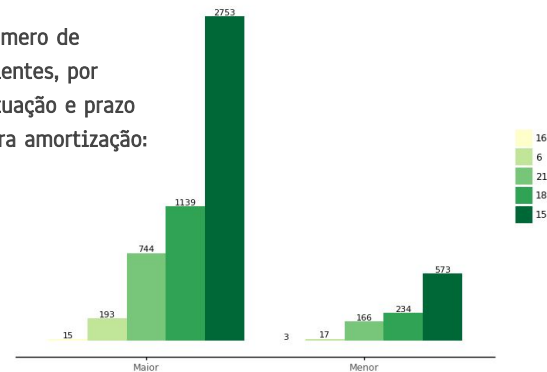
Número de clientes, por situação e tipo do tomador de empréstimo:



Número de clientes, por situação e percentual de retenção:



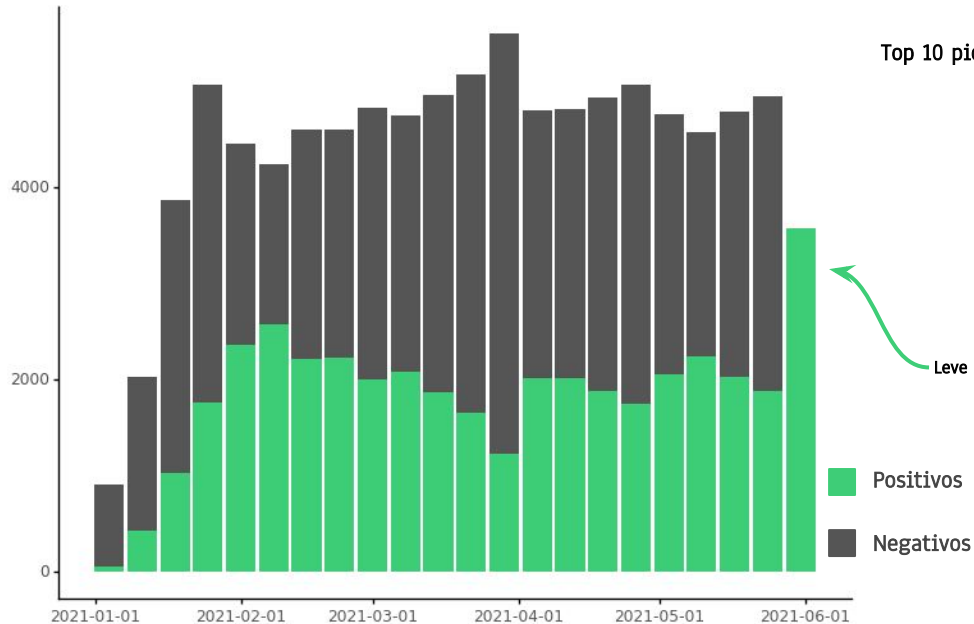
Número de clientes, por situação e prazo para amortização:



SAÚDE DA CARTEIRA DE CRÉDITO

PERGUNTA CHAVE: OS CLIENTES ESTÃO PAGANDO MAIS OU MENOS POR SEMANA?

Número de clientes negativos e positivos, por semana:



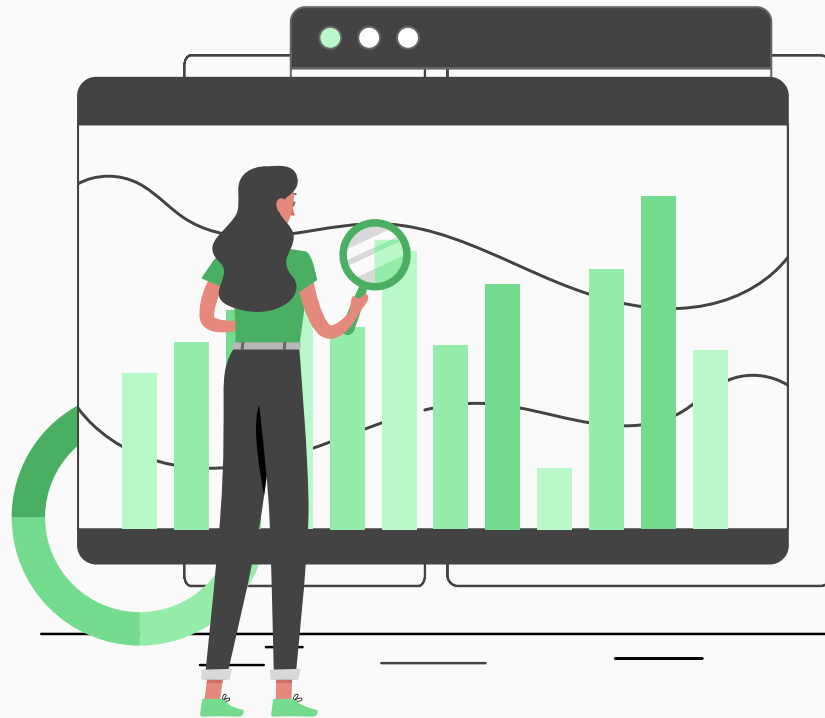
Top 10 piores pagadores

Leve melhora em 31/05/2021

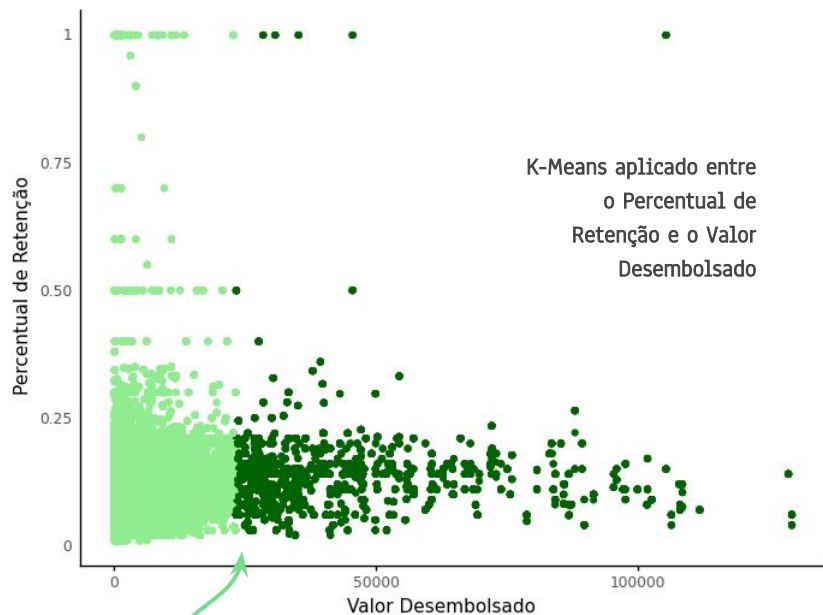
ID	SOMA DE VLR_PAGAMENTO_SEMANAL
I53352	R\$ -42.764,42
447172	R\$ -42.732,29
I64606	R\$ -31.103,47
341125	R\$ -25.821,16
697658	R\$ -24.715,99
607451	R\$ -22.805,24
619959	R\$ -21.808,31
302166	R\$ -21.479,50
I70432	R\$ -20.798,34
I25498	R\$ -19.993,97

03. CLUSTERIZAÇÃO

Utilizando K-Means Elbow Method

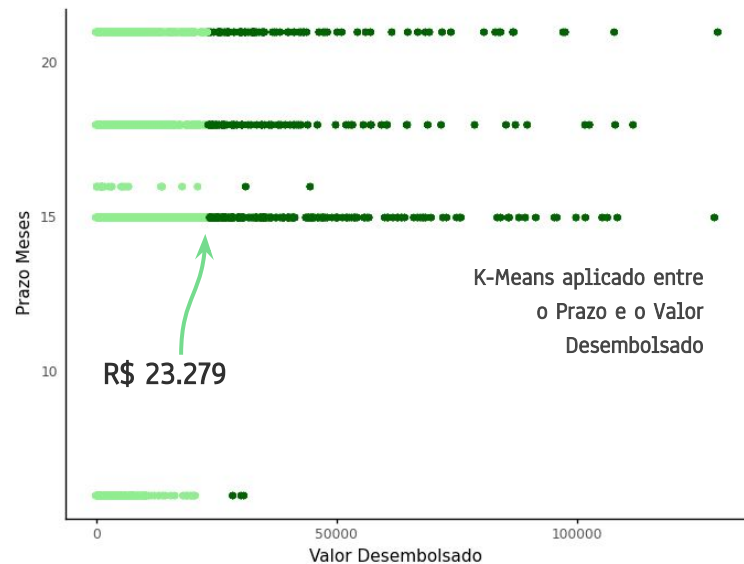


K-MEANS ELBOW METHOD



R\$ 23.279

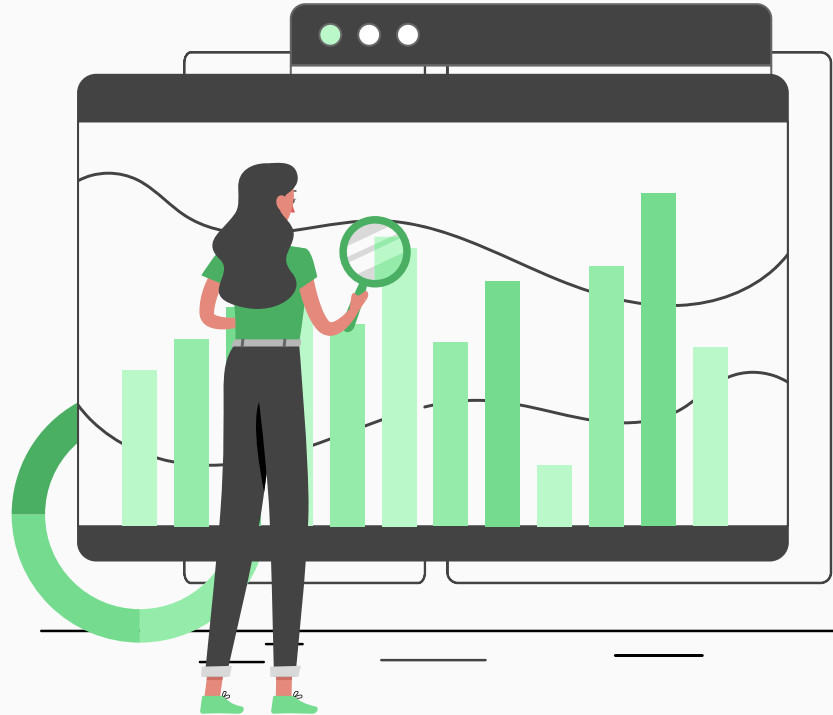
CLUSTERIZAÇÃO



R\$ 23.279

04. CONSIDERAÇÕES FINAIS

Insights e recomendações



OBRIGADO!

