

Base de E-commerce

ESTUDO DE BASE

CLIENTES

Março de 2024

OBJETIVO

O objetivo deste projeto é realizar uma segmentação estratégica da base de clientes de um e-commerce, com foco em otimizar ações de CRM (Customer Relationship Management).

A proposta é identificar diferentes perfis de consumo, avaliar o comportamento de compra, definir 'o que é um churn' e propor ações personalizadas para retenção, reativação e aumento de valor dos clientes.

A análise foi realizada com base em dados transacionais de um dos maiores e-commerces do Brasil, abrangendo pedidos, produtos, pagamentos, avaliações e perfis de clientes. O período analisado compreende de setembro de 2016 a outubro de 2018.

Por meio de técnicas de estatísticas e aprendizado de máquina, os clientes foram agrupados com base em seu histórico de compras, permitindo a criação de estratégias de comunicação mais eficientes.

Panorama da Base

Visão Geral dos Clientes

95.406

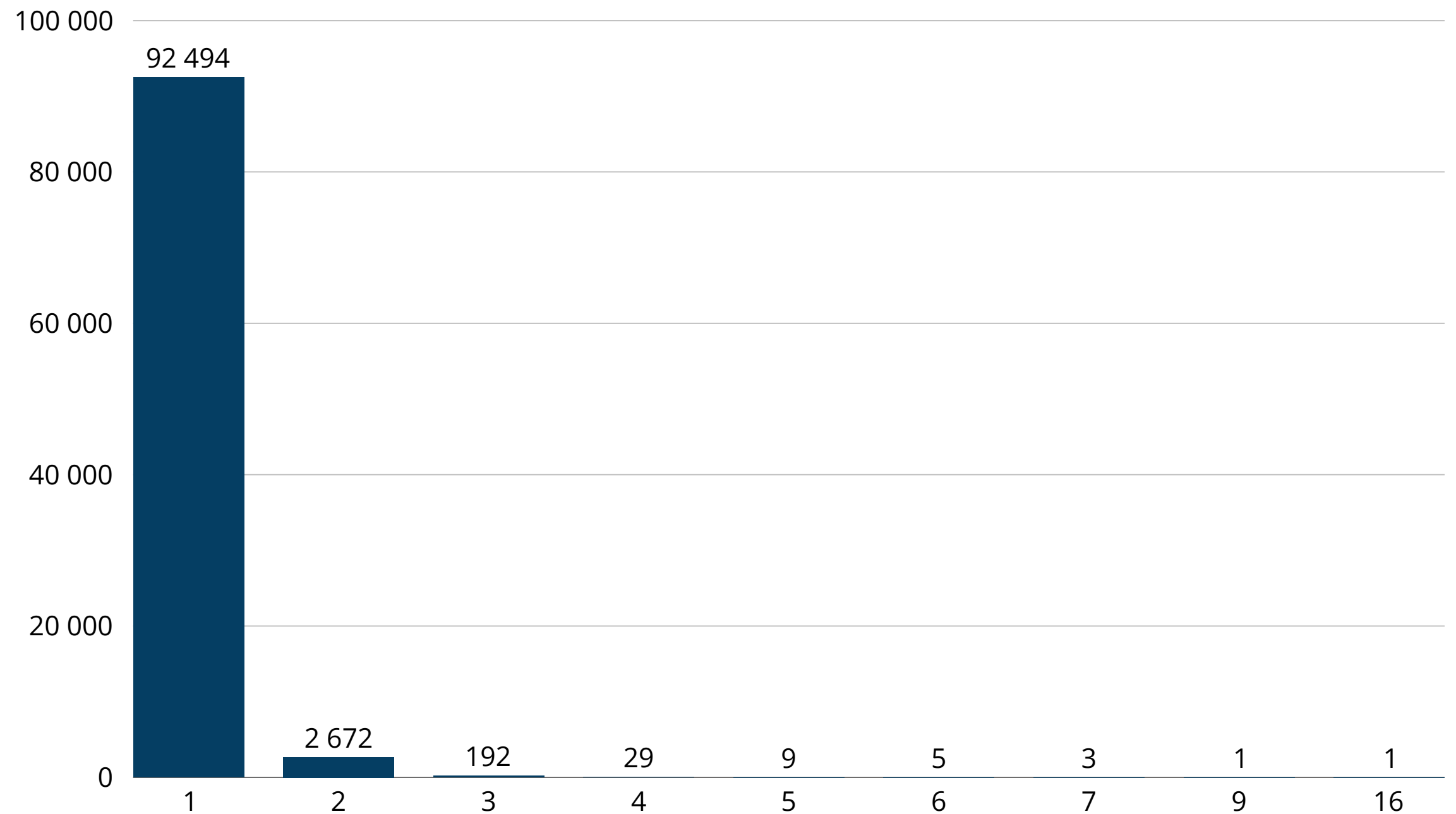
TOTAL CLIENTES
DISTINTOS

97%

% QUE COMPRARAM
SÓ 1 VEZ

Cerca de 97% dos clientes realizaram apenas uma compra, indicando alto volume de compradores pontuais e grande **potencial de reativação**.

Número de Clientes por Qnt. de Pedidos



Panorama da Base

Visão Geral das Vendas

98.651

TOTAL DE PEDIDOS

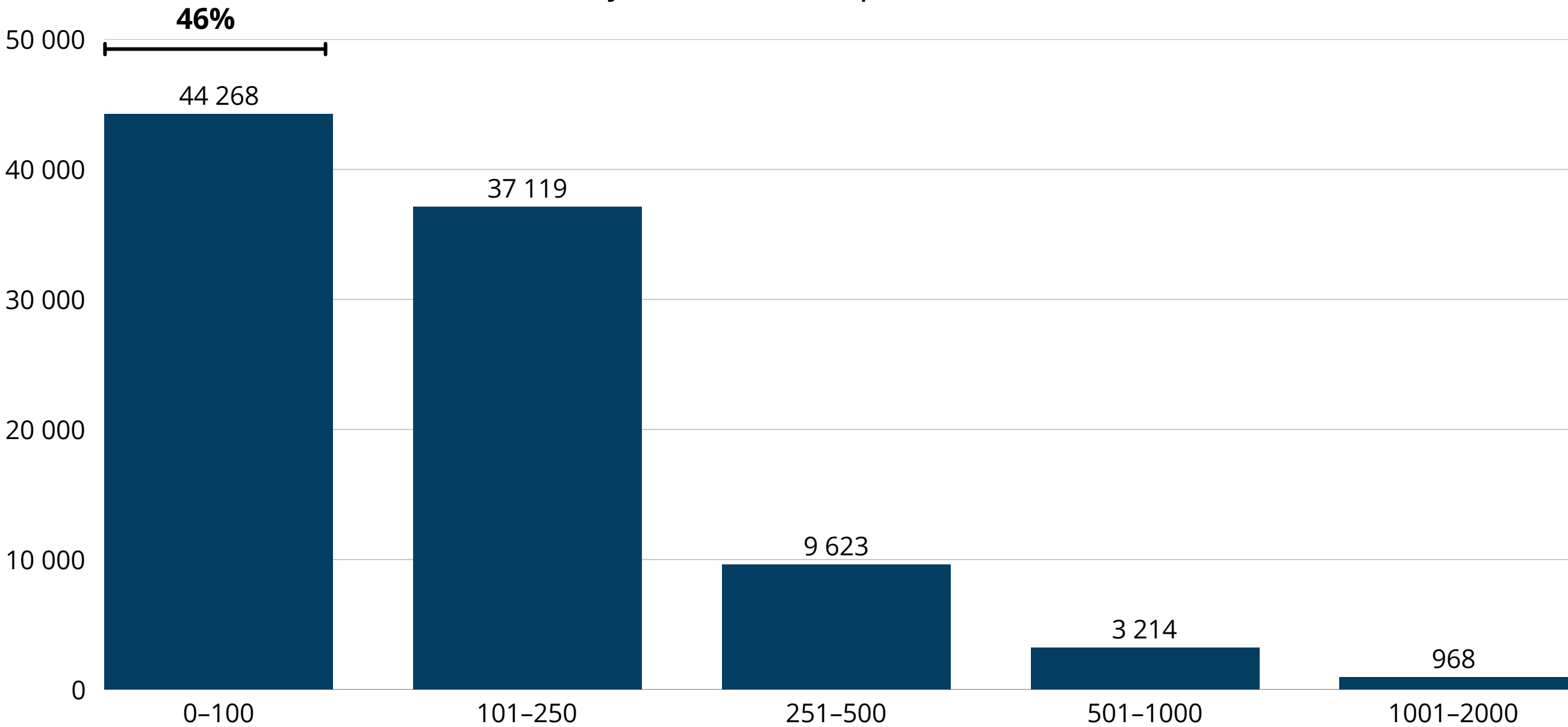
R\$ 158Mi

VALOR TOTAL DAS VENDAS*

R\$104

TICKET MÉDIO (MÉDIA POR CLIENTE*)

Distribuição de Clientes por Valor Total Gasto



*Apenas pedidos com data de pagamento aprovado.

*Para o valor médio foram utilizados os valores referente a mediana, pois a base apresenta valores que são discrepantes da média.

Panorama da Base

Visão Geral dos Produtos

Distribuição de Clientes por Qnt. de Produtos Comprados

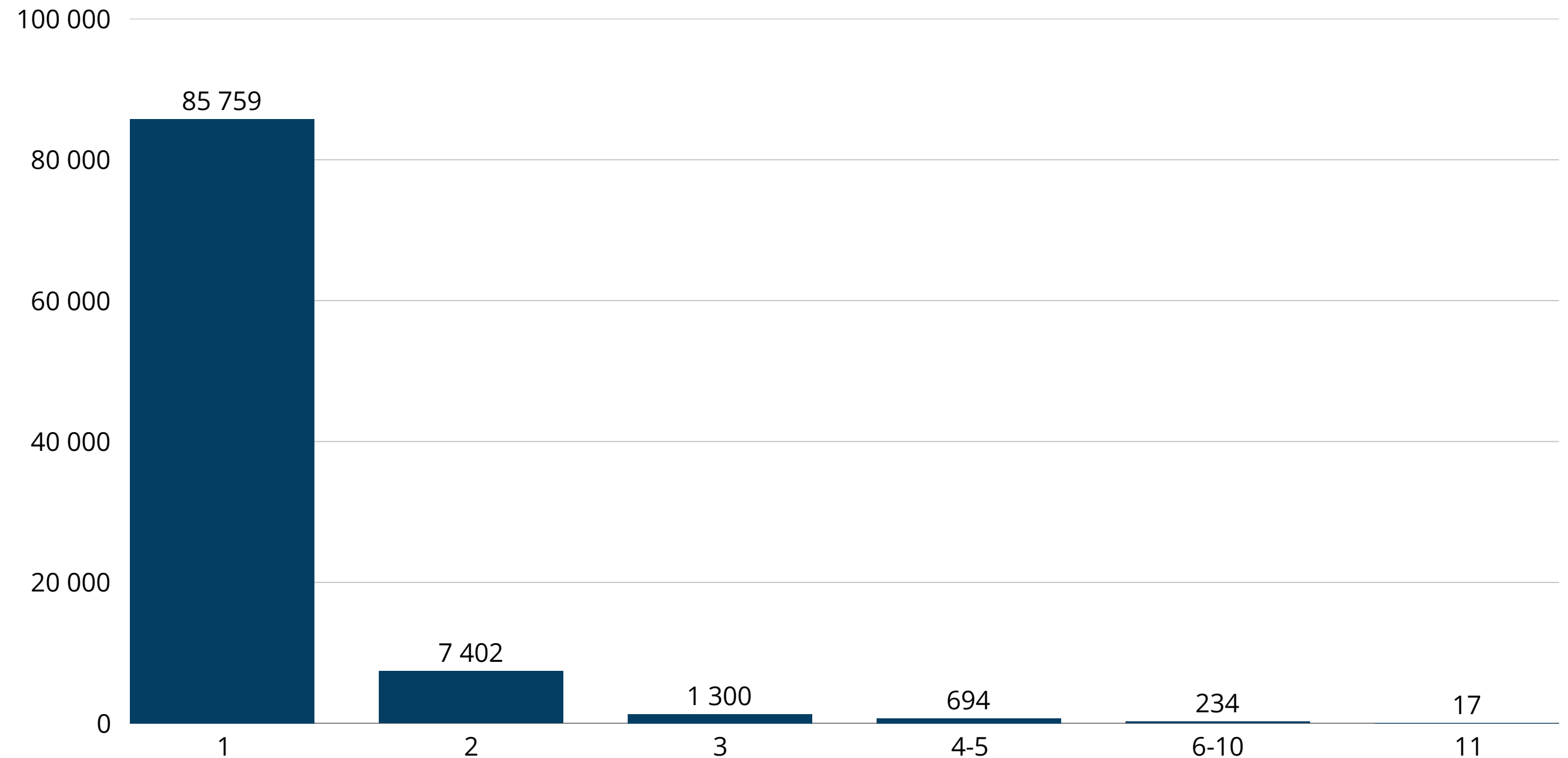
109.156

TOTAL DE ITENS
VENDIDOS

90%

CLIENTES QUE COMPRARAM
APENAS UM PRODUTO

A maioria das compras é
composta por apenas 1 item,
revelando oportunidades para
campanhas de recompra.



Base de E-commerce

DEFINIÇÃO DE
CHURN

Março de 2024

Definição de Churn

Segmentando Ativos e Inativos

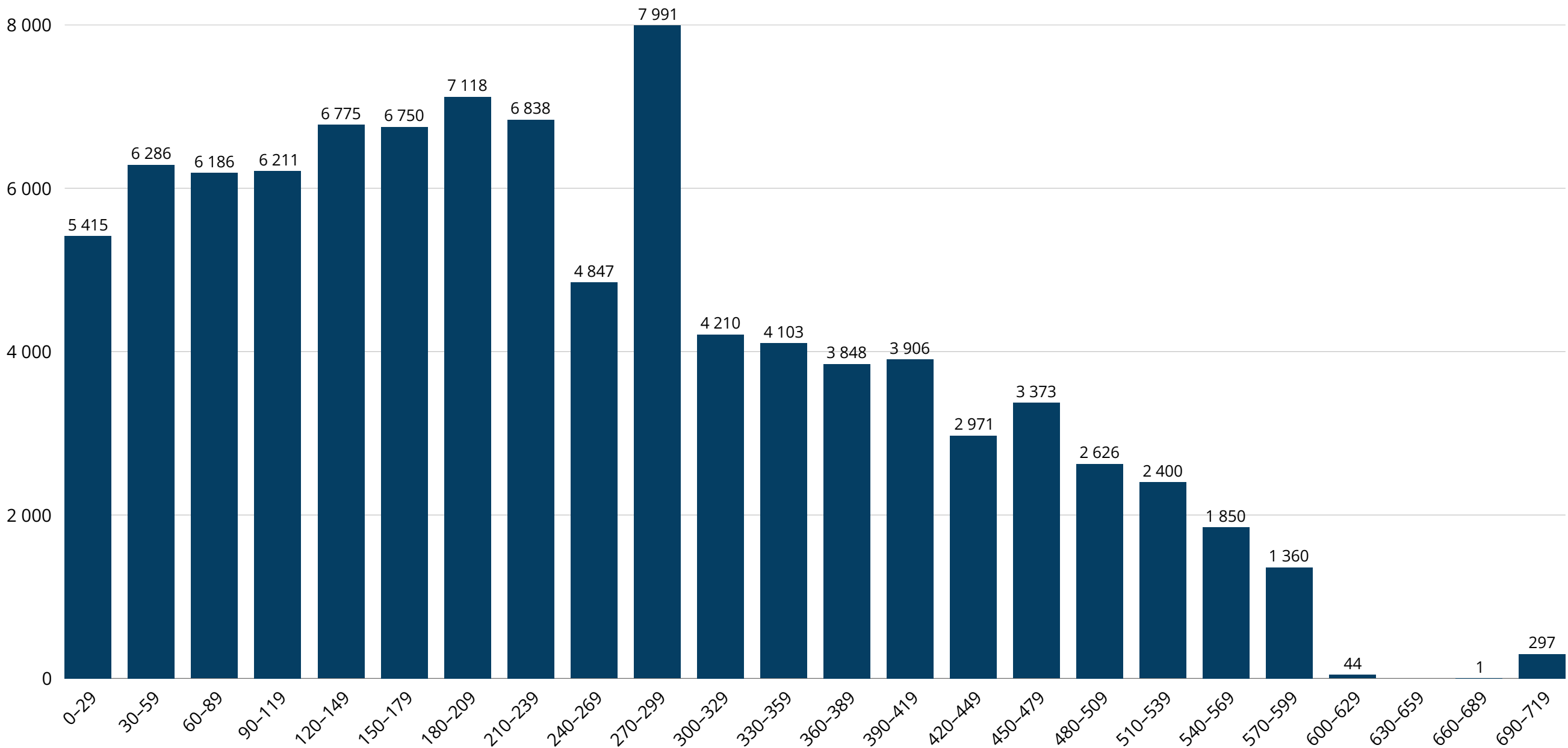
50% DA BASE FICA ATÉ

223

DIAS SEM COMPRA
(MEDIANA DA RECÊNCIA)

O cálculo de recência, utilizado para identificar churn, foi feito com base na diferença entre a data da última compra de cada cliente e a data da compra mais recente registrada na base de pedidos.

Distribuição de Clientes por Dias Sem Compra



Definição de Churn

Segmentando Ativos e Inativos

75% DA BASE FICA ATÉ

324

DIAS SEM REALIZAR
UMA COMPRA

Para definir o risco de churn, consideramos o comportamento histórico dos **clientes que realizaram mais de uma compra**. A partir da distribuição de recência desse grupo, foi adotado como ponto de **corte o percentil 75%**, que corresponde a **324 dias sem realizar uma nova compra**.

Clientes com recência superior a esse limite foram classificados como churn, por estarem fora do intervalo esperado de recompra observado na maioria dos casos.

29%

DE CLIENTES
EM CHURN

Definindo churn como **clientes que ficam mais de 324 dias sem compra** e classificando toda a base de clientes, vemos que 29% dos clientes foram considerados churn pela inatividade além do tempo limite.

Base de E-commerce

APRESENTAÇÃO DOS
CLUSTERS

Março de 2024

Metodologia

Segmentação dos Clientes

Para entender melhor os diferentes perfis de clientes, aplicamos uma segmentação baseada em comportamento de compra, utilizando cinco variáveis:

Total de Pedidos

Total de Itens

Valor Total

Ticket Médio

Recência

As variáveis com grande variação ou assimetria passaram por transformação logarítmica, e todas foram padronizadas para garantir equilíbrio na análise.

Foi utilizado o **Método do Cotovelo** para determinar o número ideal de clusters. Esse método avalia o desempenho do modelo com diferentes quantidades de clusters e identifica o ponto em que adicionar mais grupos deixa de trazer ganho significativo. A partir disso, foi definido **3 clusters** como a segmentação ideal para os dados.

Em seguida, foi aplicado o algoritmo **K-Means**, uma técnica de aprendizado de máquina não supervisionada. Isso significa que ele identifica padrões e divide os clusters sem precisar inserir dados de resultado para treino.

Resultados

Segmentação dos Clientes

A segmentação gerou três clusters distintos de clientes, com base no volume de pedidos, valor gasto, itens por pedido, ticket médio e recência da última compra. Abaixo, estão os principais perfis identificados:

Alta Receita

31%

Esse grupo concentra os clientes mais valiosos.

- Maior Ticket Médio
- Mais Itens por Pedido
- Maior Valor Total

Apesar de não serem os mais recentes, ainda demonstram comportamento ativo.

Baixo Valor Recente

41%

- Compraram Recentemente
- Baixo Ticket Médio
- Menor Volume de Itens
- Baixo Valor Total

Esse grupo representa uma oportunidade estratégica de crescimento, sendo ideal para campanhas de engajamento, cross-sell e estímulo à recompra.

Inativos Estratégicos

28%

- Alta Recência
- Ticket Médio Intermediário

Eles representam um risco de churn elevado, sendo recomendada a aplicação de campanhas de reativação personalizadas, focadas em resgatar esse público antes da perda definitiva.

1- Alta Receita

Segmentação dos Clientes

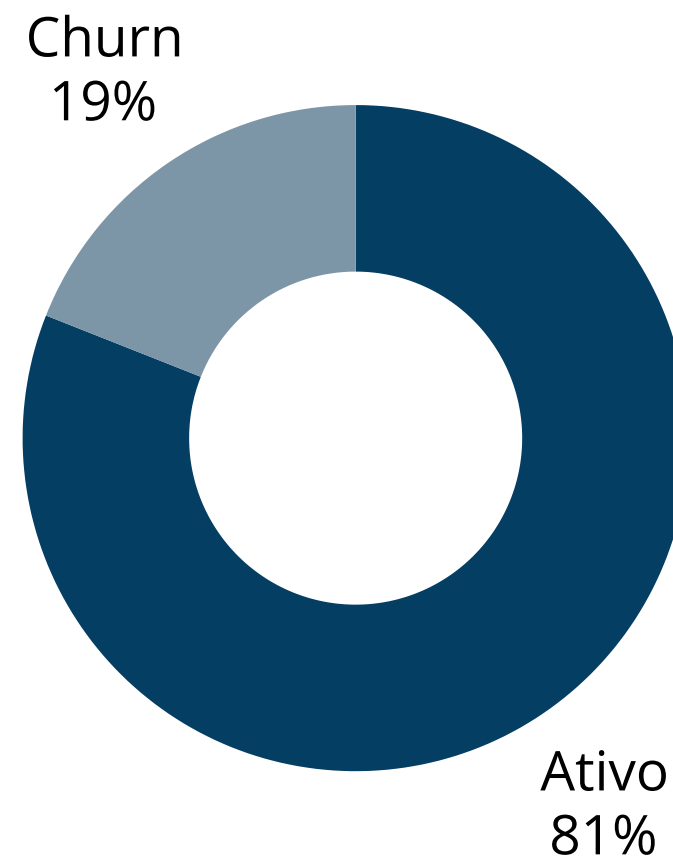
31%

DA BASE SÃO CLIENTE
ALTA RECEITA

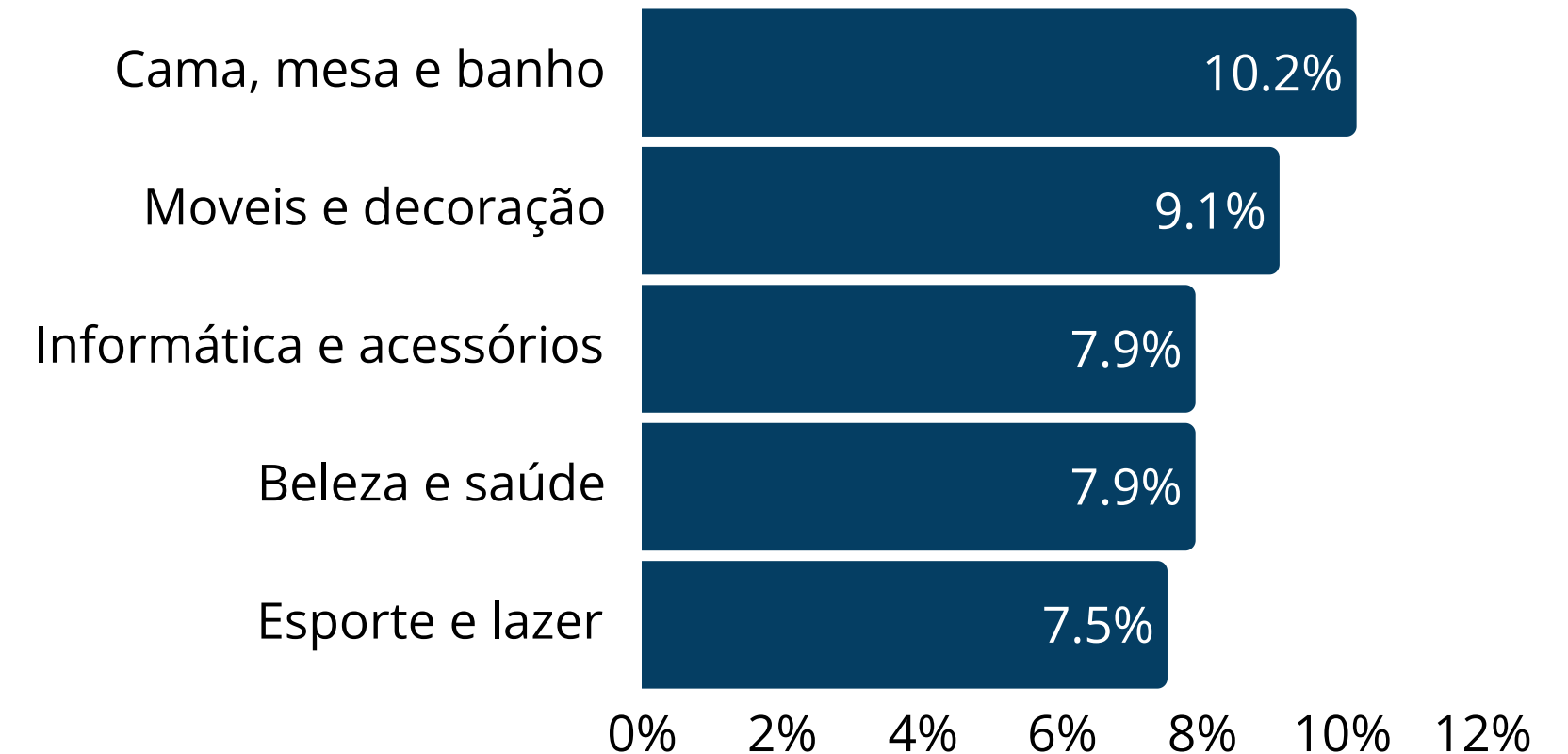
R\$ 236

MÉDIA DO VALOR
TOTAL GASTO*

% de Churn



Top 5 Categorias Mais Compradas

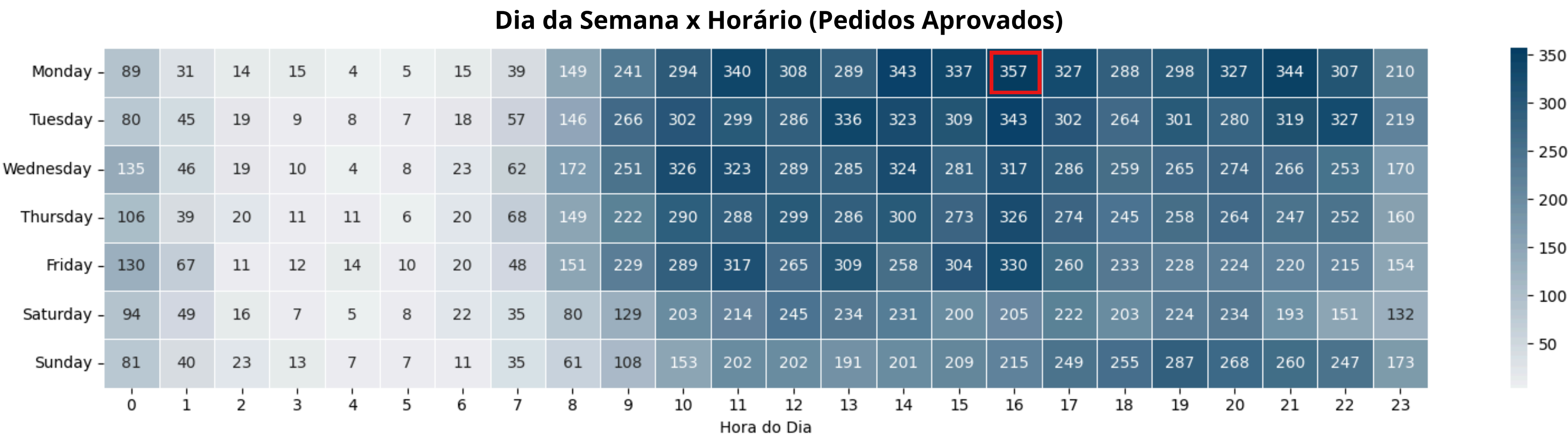


O cluster **Alta Receita** representa 31% da base e reúne os clientes mais valiosos, com alto ticket médio e maior volume de itens por pedido. Apesar de seu impacto positivo na receita, 19% apresentam risco de churn, reforçando a importância de ações de fidelização. As principais categorias compradas (Cama, mesa e banho, Móveis e decoração e Informática) indicam um perfil voltado ao lar e tecnologia.

*Para o valor médio foram utilizados os valores referente a mediana, pois a base apresenta valores que são discrepantes da média.

1- Alta Receita

Segmentação dos Clientes



Ao analisarmos o comportamento de compra do cluster **Alta Receita**, observamos que os clientes concentram suas compras principalmente entre **10h e 16h, de segunda a quinta-feira**. O pico absoluto acontece às **16h de segunda-feira**, com 357 pedidos aprovados, seguido por horários similares nas **terças e quartas-feiras**, especialmente entre **10h e 14h**.

2- Baixo Valor Recente

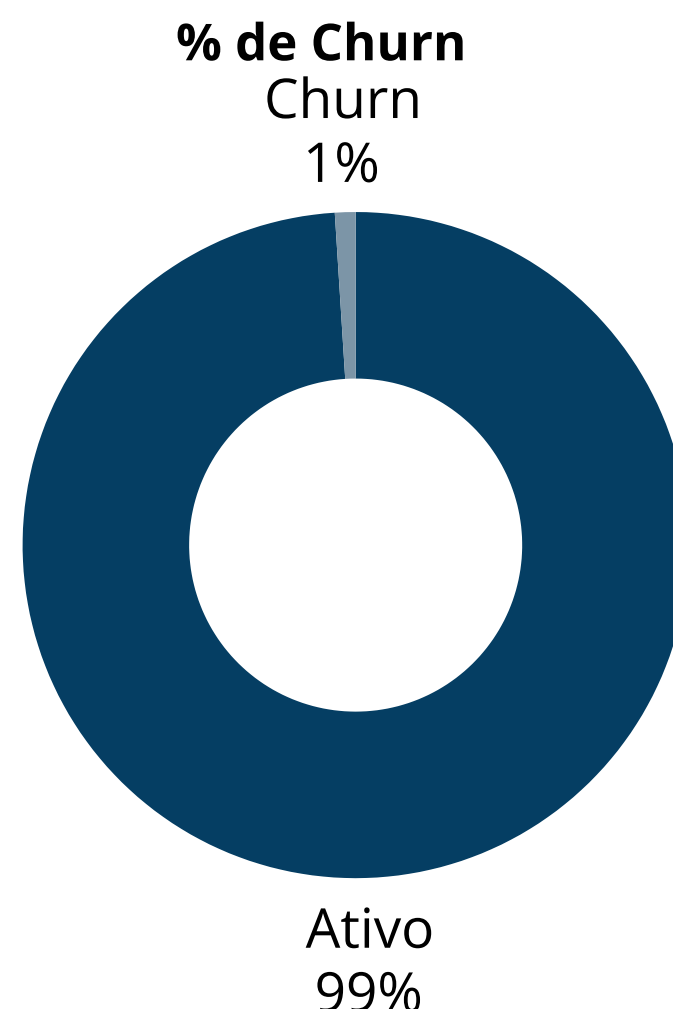
Segmentação dos Clientes

41%

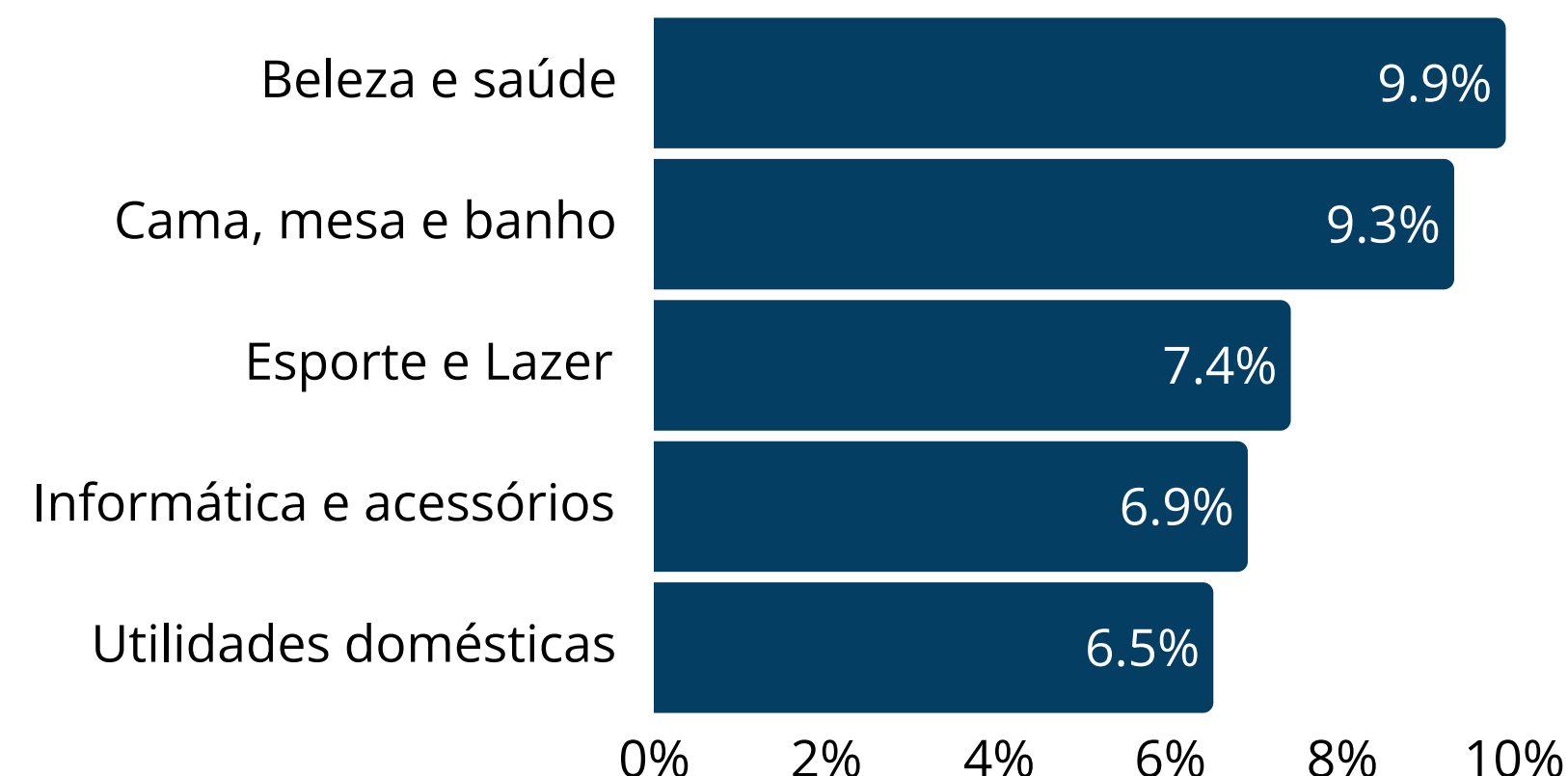
DA BASE SÃO CLIENTE
BAIXO VALOR RECENTE

R\$ 68

MÉDIA DO VALOR
TOTAL GASTO*



Top 5 Categorias Mais Compradas

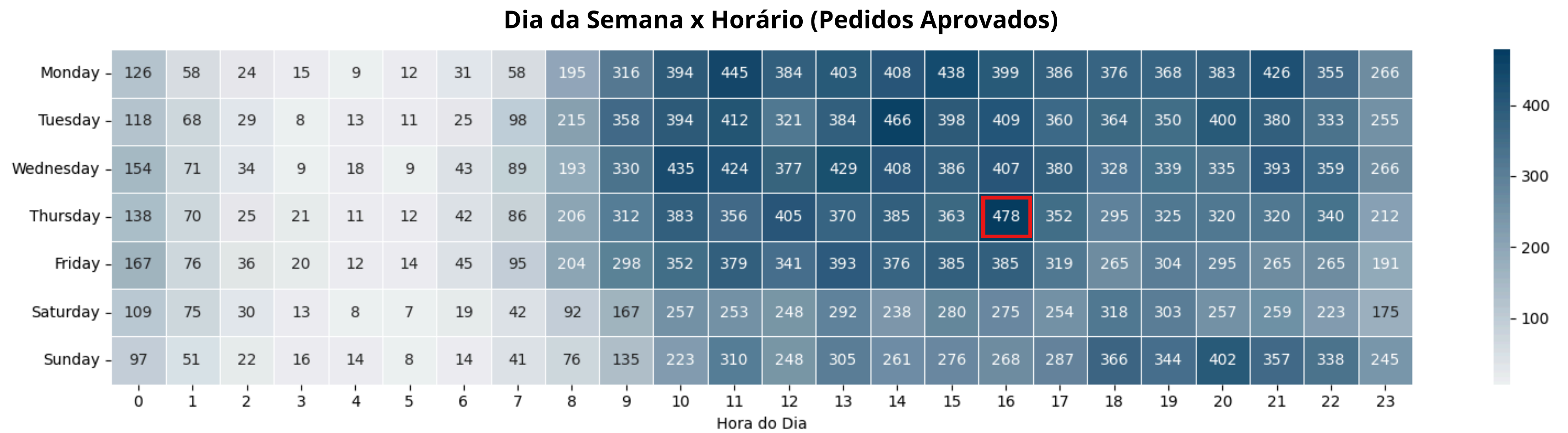


O cluster **Baixo Valor Recente** representa 41% da base de clientes, sendo o grupo mais numeroso. Apesar do baixo valor médio de compra (R\$ 75), apresenta um comportamento promissor, com 99% dos clientes ainda ativos e potencial para estímulo de recorrência. As categorias mais compradas, Beleza e Saúde, Cama, mesa e banho e Esporte e Lazer, indicam um perfil focado em itens de uso pessoal e domésticas.

*Para o valor médio foram utilizados os valores referente a mediana, pois a base apresenta valores que são discrepantes da média.

2- Baixo Valor Recente

Segmentação dos Clientes



Os clientes classificados como **Baixo Valor Recente** apresentam um volume consistente de compras ao longo da semana, com destaque para o período entre **10h e 16h, de segunda a sexta-feira**. O pico de pedidos ocorre às **16h de quinta-feira**, com forte presença também nas quintas e quartas-feiras nos mesmos horários.

3- Inativos Estratégicos

Segmentação dos Clientes

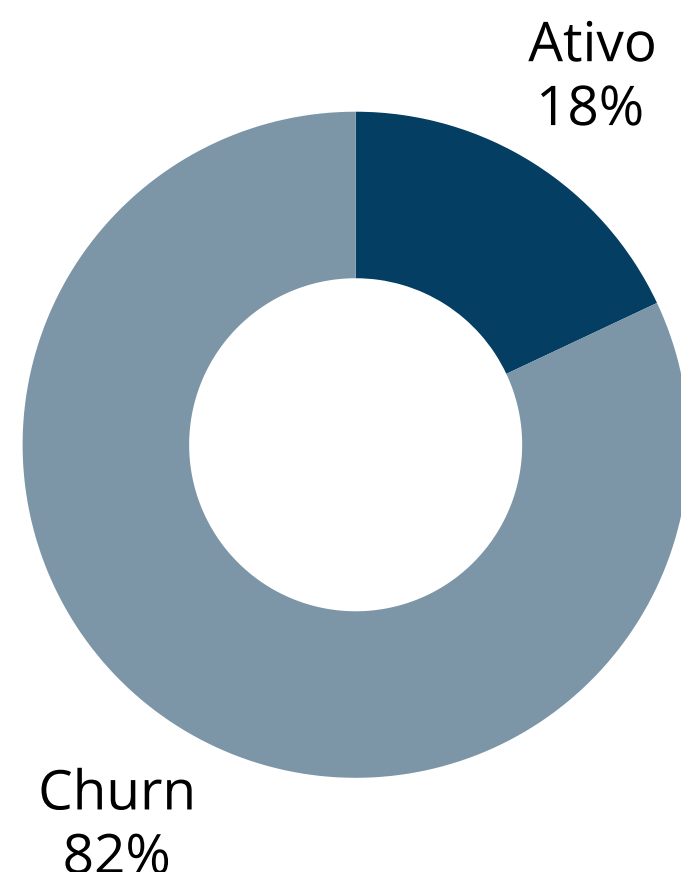
28%

DA BASE SÃO CLIENTE
INATIVOS ESTRATÉGICOS

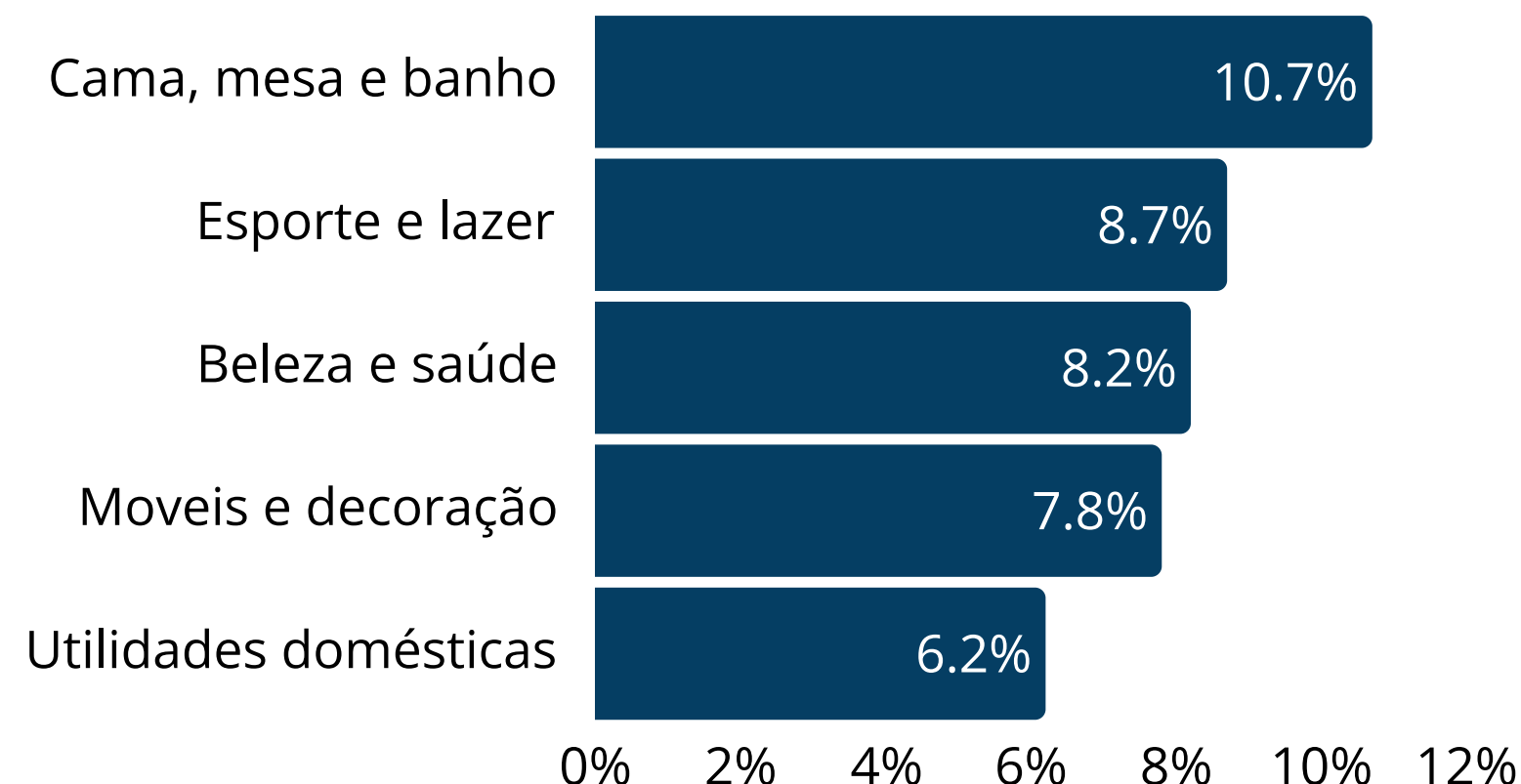
R\$ 91

MÉDIA DO VALOR
TOTAL GASTO*

% de Churn



Top 5 Categorias Mais Compradas

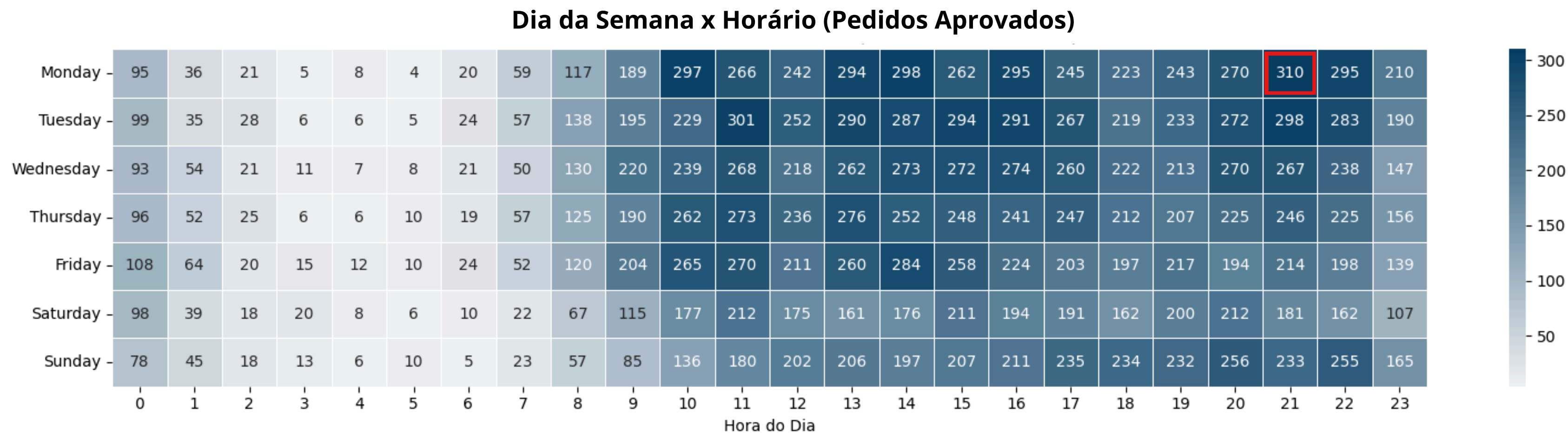


O cluster **Inativos Estratégicos** representa 28% da base e reúne clientes com gasto médio razoável (R\$ 99), mas com alta taxa de inatividade (82%). Apesar de estarem distantes da última compra, já demonstraram bom potencial de valor e devem ser priorizados em ações de reativação segmentada. As principais categorias consumidas, Cama, mesa e banho, Esporte e lazer e Beleza e saúde, revelam interesse em itens duráveis e de uso pessoal

*Para o valor médio foram utilizados os valores referente a mediana, pois a base apresenta valores que são discrepantes da média.

3- Inativos Estratégicos

Segmentação dos Clientes



O cluster **Inativos Estratégicos** apresenta um padrão de compras mais concentrado entre os dias úteis, especialmente nas **segundas, terças e quartas-feiras**, com picos de volume entre **10h e 14h**. O maior volume registrado foi às **21h de segunda-feira**, o que demonstra uma atividade expressiva no início da semana

Base de E-commerce

APRESENTAÇÃO DOS
INSIGHTS

Março de 2024

Insights

Plano de Ação

Alta Receita

Churn: 19%

Média valor: R\$236

Melhor horário: 16h de segunda-feira

Principais categorias: Cama, mesa e banho, Móveis e decoração e Informática e Acessórios.

Objetivo: Retenção desses clientes.

Ações:

- Ofertas Exclusivas e Personalizadas
 - Produtos das categorias favoritas
 - Recomendar produtos premium
- Programas de Fidelidade ou Benefícios VIP
 - Descontos acumulativos
 - Cupons Mensais
 - Cashback

Baixo Valor Recente

Churn: 1%

Média valor: R\$68

Melhor horário: 16h de quinta-feira

Principais categorias: Beleza e saúde, Cama, mesa e banho, Móveis e decoração e Esporte e lazer.

Objetivo: Aumentar o valor de compra.

Ações:

- Incentivo à Recorrência
 - Campanha de recompra
 - Upsell e Cross Sell
- Aumentar Ticket
 - Ofertas de frete grátis acima de R\$X
 - Recomendação de maior valor

Inativos Estratégicos

Churn: 82%

Média valor: R\$91

Melhor horário: 21h de segunda-feira

Principais categorias: Cama, mesa e banho, Esporte e lazer e Beleza e saúde.

Objetivo: Converter inativos.

Ações:

- Campanhas de reativação
 - Comunicação personalizada
 - Descontos exclusivos para inativos
 - Cupom de retorno
- Recomendação de produtos
 - Com base nas categorias mais compradas

OBRIGADA!



ELABORADO POR:

Mariana Caetano Vidal

Cientista de Dados e Economista



Linkedin



Medium



Github