

Projeto Interdisciplinar: Ciências de Dados – Entrega 1

Antonio Petri 24026144

Daniel Moribe 24026509

Gabriel Paschoal 21010656

Vitor Kanashiro 24026113

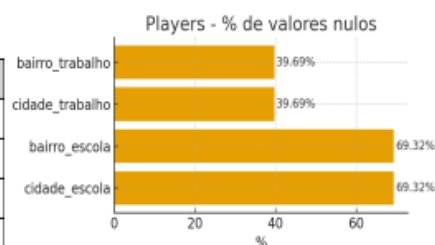
Base de transações – Cupons Capturados Descrever os Dados:

- Celular: Número de telefone do player que realizou a captura.
- Data: Dia, mês e ano em que o cupom foi registrado.
- Hora: Momento exato da captura, no formato horas, minutos e segundos.
- Nome_Estabelecimento: Nome do local onde ocorreu a captura do cupom.
- Bairro_Estabelecimento: Bairro em que o estabelecimento está localizado.
- Categoria_Estabelecimento: Classificação do tipo de negócio do estabelecimento.
- ID_Campanha: Código que identifica a campanha vinculada ao cupom.
- ID_Cupom: Identificação única de cada cupom registrado.
- Tipo_Cupom: Modalidade do cupom capturado (cashback, desconto ou produto).
- Produto: Item adquirido por meio do cupom utilizado.
- Valor_Cupom: Valor monetário do cupom resgatado.
- Repasse_Picmoney: Montante destinado à PicMoney referente ao cupom.

Verificação da Qualidade dos Dados

Players

Coluna	Valores Nulos	% Nulo
cidade_escola	6932	69.32%
bairro_escola	6932	69.32%
cidade_trabalho	3969	39.69%
bairro_trabalho	3969	39.69%



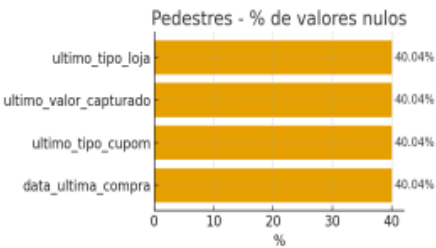
Cupons

Coluna	Valores Nulos	% Nulo
produto	66672	66.67%



Pedestres

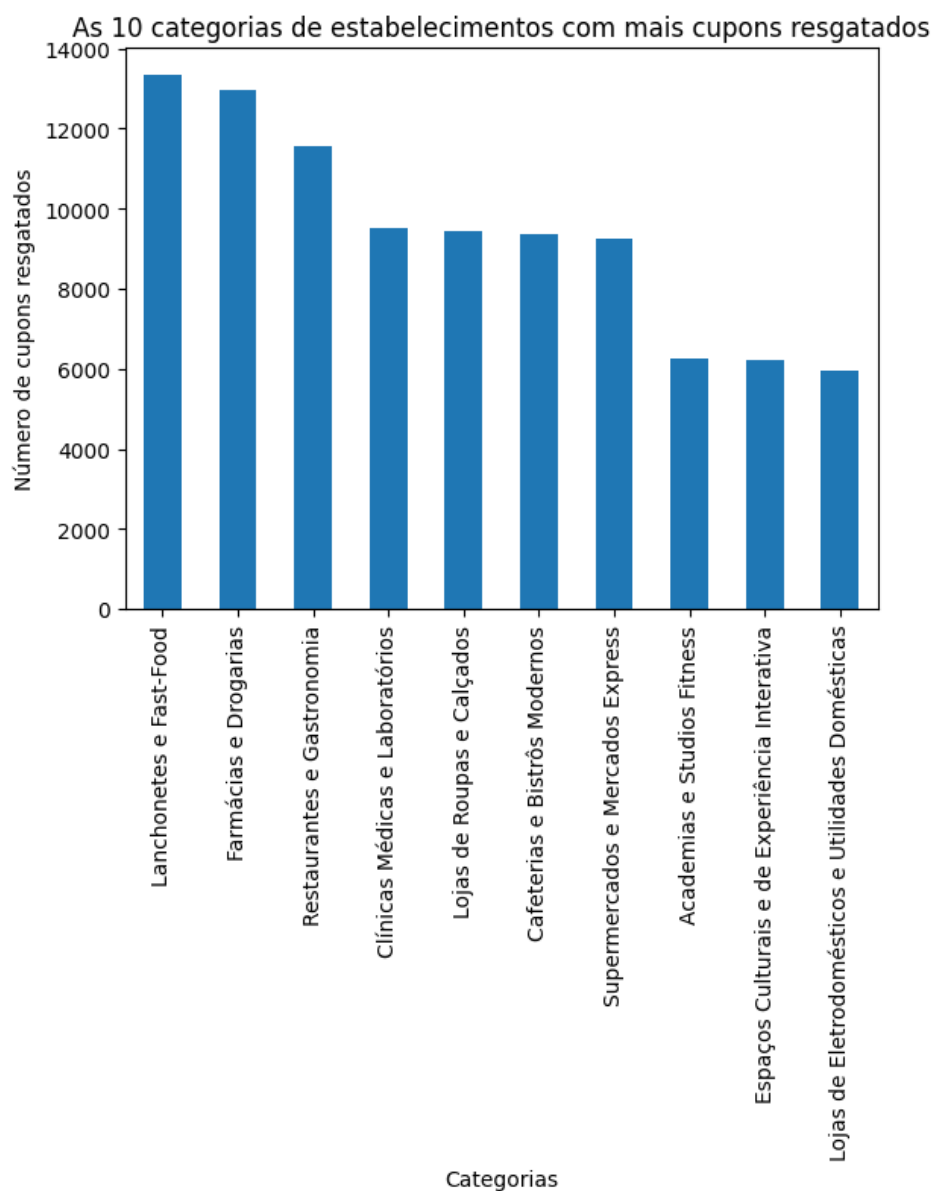
Coluna	Valores Nulos	% Nulo
data_ultima_compra	40043	40.04%
ultimo_tipo_cupom	40043	40.04%
ultimo_valor_capturado	40043	40.04%
ultimo_tipo_loja	40043	40.04%



Lojas

Não foram encontrados valores nulos.

Explorar Dados



Base Cadastral de Players

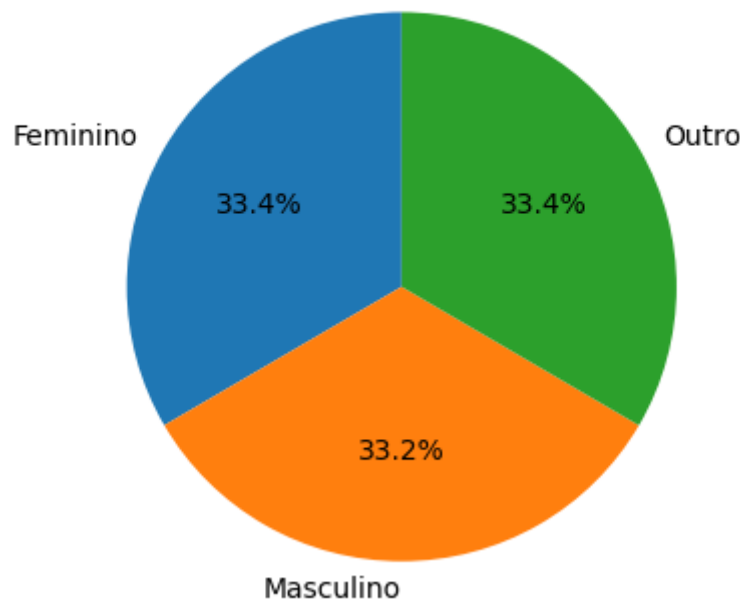
Descrever os Dados:

- Celular: O número do celular cadastrado pelo player.
- Data de nascimento: A data de nascimento registrada, no formato dia/mês/ano.
- Idade: Idade do player.
- Sexo: Sexo informado pelo player.
- Cidade_Residencial: Cidade onde o player reside.
- Bairro_Residencial: Bairro onde o player reside.
- Cidade_Trabalho: Cidade onde o player trabalha, se aplicável.

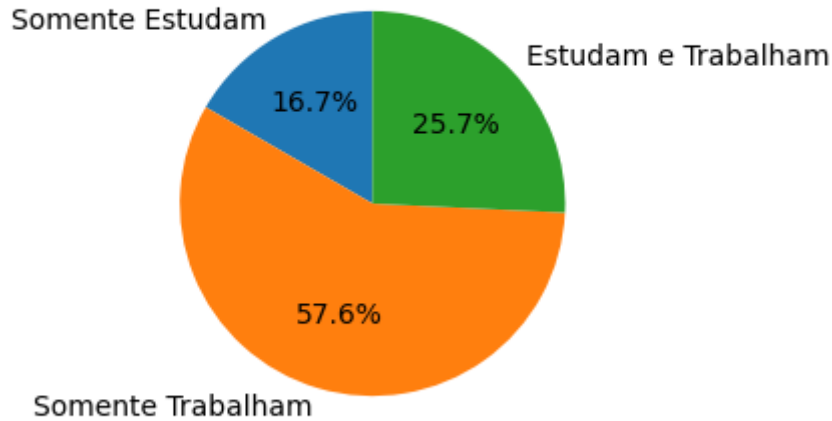
- Bairro_Trabalho: Bairro onde o player trabalha, se aplicável.
- Cidade_Escola: Cidade onde o player estuda, se aplicável.
- Bairro_Escola: Bairro onde o player estuda, se aplicável.
- Categoria_Estabelecimento: Categoria de estabelecimento mais frequentada pelo player.

Explorando os dados

Distribuição de cupons capturados por sexo



Distribuição de Clientes por Atividade



Base Simulada de Pedestres – Av. Paulista

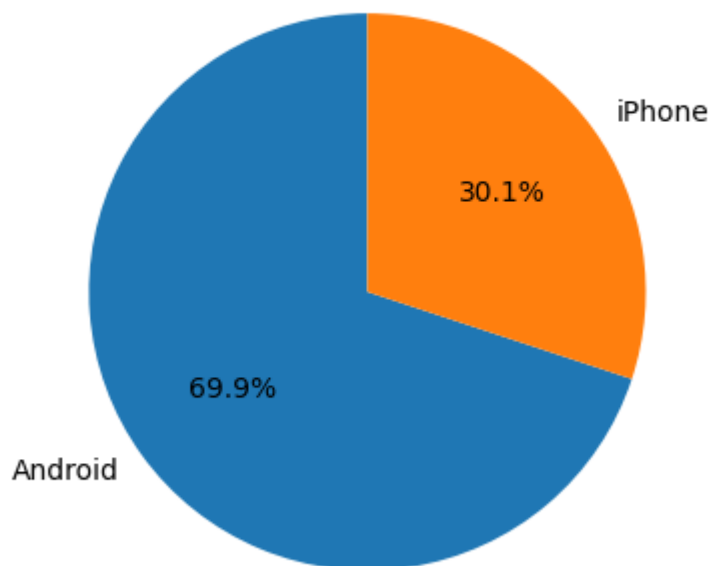
Descrever os dados

- Celular: O número de telefone cadastrado.
- Data: A data do resgate dos cupons (22/07/2025).
- Horário: O horário em que o cupom foi capturado.
- Local: Descrição do local onde o cupom foi capturado.
- Latitude: Latitude exata do ponto de captura.
- Longitude: Longitude exata do ponto de captura.
- Tipo_Celular: Sistema operacional do celular (Android ou iPhone).
- Modelo_Celular: Marca e modelo do celular do player.
- possui_app_picmoney: Indica se o player possui o app.
- data_ultima_compra: Data da última compra com a PicMoney.
- ultimo_tipo_cupom: Tipo do cupom na última compra (cashback, desconto ou produto).
- ultimo_valor_capturado: Valor do último cupom capturado.
- ultimo_tipo_loja: Tipo de loja da última compra com cupom.

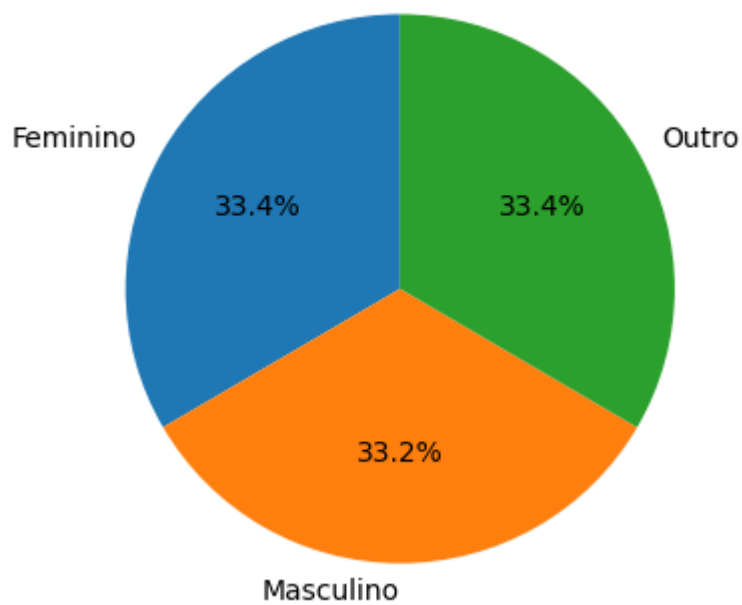
- Idade: Idade do player.
- Sexo: Sexo do player.

Explorando os dados

Quantidade de cupons capturados por tipo de celular



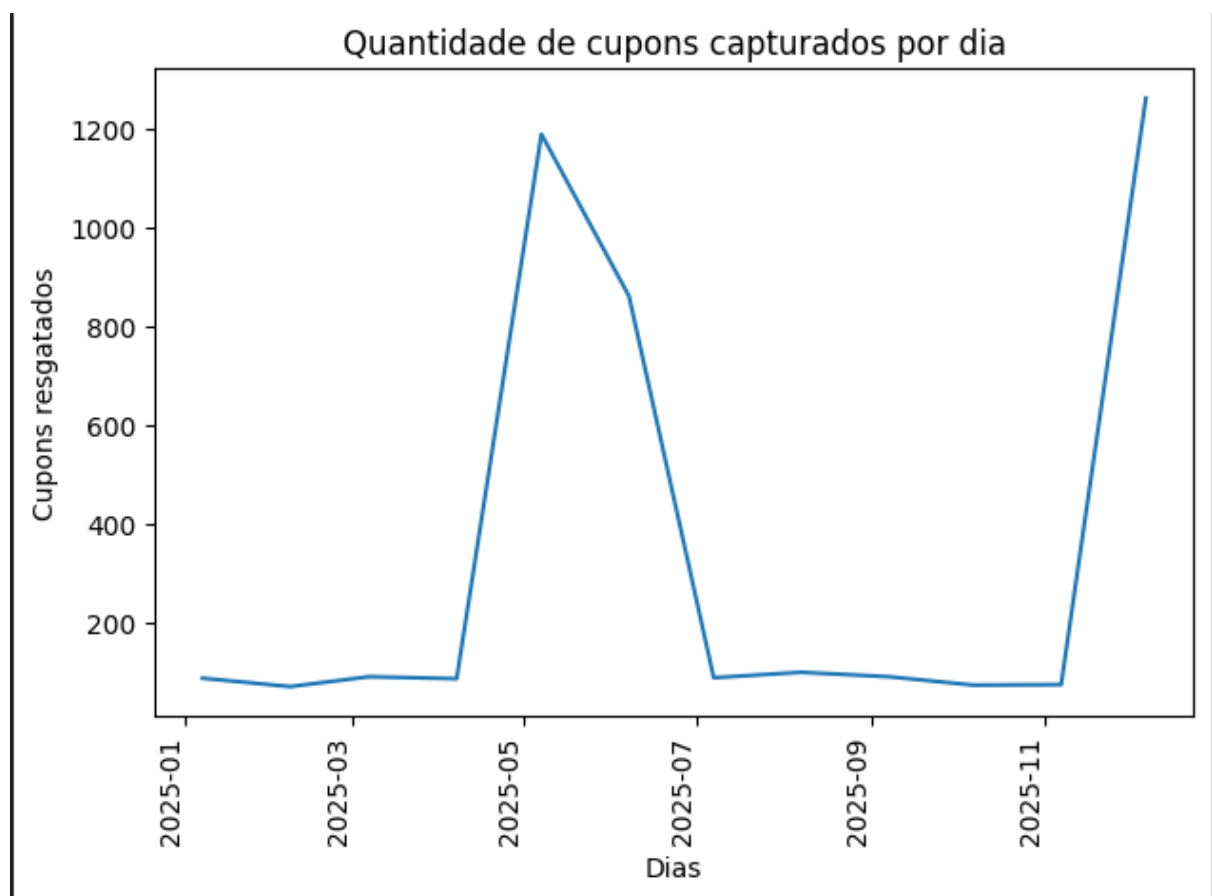
Distribuição de cupons capturados por sexo



Massa de teste com valores e lojas

- numero_celular: O número do celular do player.
- data_captura: Data da captura do cupom.
- tipo_cupom: Tipo do cupom resgatado (cashback, produto ou desconto).
- tipo_loja: Tipo da loja onde o cupom foi resgatado.
- local_captura: Local onde o cupom foi capturado.
- Latitude: Latitude exata de onde o cupom foi capturado.
- Longitude: Longitude exata de onde o cupom foi capturado.
- nome_loja: Nome da loja onde o cupom foi capturado.
- endereco_loja: Endereço da loja (rua e CEP).
- valor_compra: Valor da compra realizada.
- valor_cupom: Valor do cupom descontado

Exploração dos Dados



Conclusão

A análise interdisciplinar dos dados permitiu identificar padrões relevantes no comportamento dos usuários da PicMoney e nas características das transações. A

limpeza e padronização das bases foram fundamentais para garantir a confiabilidade dos resultados, eliminando inconsistências como categorias incorretas, coordenadas inválidas e informações irrelevantes. Com os dados tratados, observou-se predomínio de categorias de consumo cotidiano (lanchonetes, farmácias e restaurantes), uma base equilibrada por sexo e média etária próxima de 40 anos, e predominância do sistema Android. Esses resultados reforçam a importância da preparação de dados e abrem caminho para análises mais profundas, como estudos preditivos, segmentação e estratégias direcionadas para aumentar a eficiência da PicMoney.