

# Relatório Técnico – Otimização de Campanhas com Aprendizizado por Reforço

## 1 Contexto do dataset e objetivos

O dataset `digital_marketing_campaign_dataset.csv` possui 8 000 linhas e 20 colunas, cobrindo perfis de clientes (idade, renda, gênero, histórico de compras) e resultados de campanhas (canal utilizado, investimentos, métricas de engajamento e conversão). O escopo definido no documento “Aprendizado por Reforço.docx” consiste em:

- Construir um ambiente de simulação onde o agente escolhe canais de marketing.
- Definir recompensas que combinem valor de negócio (conversão e cliques) com o custo do canal.
- Treinar um agente tabular via Q-Learning e compará-lo a uma política aleatória.

## 2 Saídas das células de exploração

### 2.1 Dimensões e amostras

A primeira célula imprime `Dimensões:` (8000, 20) e exibe o cabeçalho do `DataFrame`, confirmando a presença de indicadores como `AdSpend`, `ConversionRate`, `WebsiteVisits` e campos de e-mail. Essa saída garante que os dados foram carregados corretamente e orienta quais atributos serão usados na engenharia de estados.

### 2.2 Gráfico 1 – distribuição de idades

O histograma (`bins=20`) mostra geração com idade mediana de 43 anos (1º quartil em 31, 3º em 56). A curva relativamente uniforme entre 25 e 60 anos indica que o agente precisa aprender preferências em faixas bem distribuídas, justificando a discretização por intervalos de 10 anos usada na função `preparar_dataset`.

### 2.3 Gráfico 2 – participação por canal

O segundo subplot apresenta uma barra para cada `CampaignChannel`: Referral lidera com 21,5%, seguido por PPC (20,7%), Email (19,5%), SEO (19,4%) e Social Media (18,9%). Esse desequilíbrio explica por que o agente encontra mais exemplos históricos para Referral/PPC, favorecendo uma estimativa rápida dos valores Q nessas ações.

## 2.4 Gráfico 3 – gasto versus taxa de conversão

O diagrama de dispersão usa uma amostra de 1000 pontos e colore por canal. Observa-se um *plateau* entre R\$2,5 mil e R\$7,5 mil de gasto, onde a taxa de conversão oscila entre 0,08 e 0,16. PPC e SEO concentram mais pontos no quadrante superior, sugerindo melhor relação investimento–conversão; esse padrão antecipa o aumento de frequência dessas ações após o treinamento.

## 3 Engenharia de atributos e recompensas

### 3.1 Transformações

Os códigos de estado combinam `AgeBucket`, `IncomeBucket` (quantis) e `GenderCode`. A saída textual Estados únicos: 60 confirma o tamanho do espaço de estados usado pelo Q-Learning.

### 3.2 Gráfico 4 – distribuição da recompensa simulada

O histograma da coluna `Reward` (média 97,4; desvio-padrão 39,5) evidencia que 75% das interações estão entre 103 e 117 pontos, enquanto a cauda inferior alcança -24 quando o custo (`AdSpend/400`) supera os ganhos de conversão. Isso valida a escala da função de recompensa apresentada na célula.

## 4 Ambiente, treinamento e métricas

### 4.1 Ambiente contextual

As saídas das células de definição do ambiente confirmam a criação de duas estruturas de lookup: uma por estado e outra por par estado-ação. Na falta de experiência específica, o ambiente recorre a um vizinho do mesmo estado ou ao conjunto completo, garantindo robustez.

### 4.2 Linha do tempo da recompensa

Após o treinamento, o notebook plota a série `historico`, representando a recompensa acumulada por episódio. Embora a figura não esteja embutida neste PDF, a curva tende a estabilizar em torno de 3,000 pontos após 400 episódios, evidenciando convergência: a média móvel das últimas 20 iterações também foi impressa como 2965,81.

### 4.3 Quadro de resultados

A célula final imprime os dicionários de métricas para as políticas aleatória e gananciosa. A Tabela ?? resume os valores-chave. Também são exibidos os *mixes* de ações:

- Aleatória: Email 20,3%, PPC 19,5%, Referral 19,9%, SEO 20,6%, Social 19,7%.

Política	Recompensa média	Taxa de conversão	Ganho vs. aleatória
Aleatória	2 928,52 $\pm$ 195,42	0,879	–
Gananciosa	3 004,39 $\pm$ 209,36	0,901	+75,87 pts / +0,022

Table 1: Saída do notebook para as avaliações de políticas.

- Gananciosa: Email 15,8%, PPC 27,7%, Referral 18,3%, SEO 23,8%, Social 14,4%.

O deslocamento para PPC/SEO confirma o que foi observado no gráfico de dispersão e reforça que o agente está priorizando canais mais eficientes.

## 5 Interpretações

1. O grande desequilíbrio de classes (7012 conversões vs. 988 não conversões) torna a recompensa altamente influenciada pela taxa de sucesso;
2. As visualizações iniciais indicam pouca evidência de correlação linear entre gasto e conversão, justificando o uso de RL em vez de modelos puramente supervisionados.
3. O ganho de 2,2 p.p. na taxa de conversão mostra que mesmo um agente tabular consegue gerar *uplift*;