

Relatório de Análise Exploratória de Dados (EDA)

1. Descrição do Dataset

O dataset contém 200 observações e 6 colunas, com as seguintes variáveis:

Idade: Idade dos usuários.

Renda Anual (em \$): Renda anual dos usuários em dólares.

Gênero: Gênero dos usuários (Feminino ou Masculino).

Tempo no Site (min): Tempo que os usuários passaram navegando no site.

Anúncio Clicado: Indica se o usuário clicou em um anúncio (Sim ou Não).

Compra (0 ou 1): Indica se o usuário comprou uma casa (1 para sim, 0 para não).

2. Análise de Dados Faltantes

Idade: 10 valores faltantes.

Renda Anual (em \$): 11 valores faltantes.

Gênero: 7 valores faltantes.

Anúncio Clicado: 10 valores faltantes.

Tempo no Site (min) e Compra (0 ou 1): Nenhum valor faltante.

Recomendação: Tratar os valores faltantes antes de prosseguir com a modelagem, utilizando técnicas como imputação de média/mediana ou exclusão de linhas.

3. Distribuição das Variáveis

3.1. Idade

O histograma da variável Idade mostra uma distribuição aproximadamente normal, com a maioria dos usuários concentrados entre 20 e 40 anos.

A curva de densidade (KDE) confirma essa tendência, indicando que a idade média dos usuários está em torno de 30-35 anos.

3.2. Renda Anual (em \$)

O histograma da Renda Anual foi dividido em 4 bins, mostrando que a maioria dos usuários tem renda entre

30.000

e

30.000 e 70.000.

Há uma concentração maior de usuários na faixa de \$50.000.

3.3. Tempo no Site (min)

O histograma do Tempo no Site foi dividido em 16 bins, mostrando que a maioria dos usuários passa entre 5 e 30 minutos no site.

O tempo médio no site é de aproximadamente 15-20 minutos.

4. Estatísticas Descritivas

4.1. Tempo no Site (min)

Tempo máximo: 29.39 minutos.

Tempo mínimo: 5.74 minutos.

Número de observações: 200.

Diferença média entre tempos: 0.154 minutos por usuário.

4.2. Compra (0 ou 1)

A variável Compra é binária, com valores 0 (não comprou) e 1 (comprou).

Aproximadamente 33% dos usuários compraram uma casa.

5. Relação entre Variáveis

5.1. Pairplot (Idade, Renda Anual, Tempo no Site, Compra)

O Pairplot mostra a relação entre as variáveis Idade, Renda Anual, Tempo no Site e Compra.

Compradores tendem a ter uma renda anual mais alta e passar mais tempo no site.

Não há uma relação clara entre Idade e Compra.

5.2. Gênero vs Compra

A proporção de compras é semelhante entre Feminino (32.98%) e Masculino (33.33%).

O gráfico de barras agrupadas mostra que a distribuição de compras é equilibrada entre os gêneros.

5.3. Anúncio Clicado vs Compra

Usuários que clicaram em anúncios têm uma taxa de compra mais alta (36.45%) em comparação com os que não clicaram (24.10%).

O gráfico de barras agrupadas confirma que clicar em anúncios está associado a uma maior probabilidade de compra.

6. Insights e Recomendações

Perfil dos Compradores:

Compradores tendem a ter uma renda anual mais alta e passar mais tempo no site.

Clicar em anúncios aumenta a probabilidade de compra.

Estratégias de Marketing:

Focar em anúncios para usuários com renda anual acima de \$50.000.

Incentivar os usuários a passar mais tempo no site, oferecendo conteúdo relevante e interativo.

7. Conclusão

A análise exploratória revelou padrões importantes no comportamento dos usuários, como a influência da renda anual e do tempo no site na decisão de compra. Além disso, clicar em anúncios mostrou-se um fator significativo para aumentar a taxa de conversão.