

Trilha 01

Indicações de Soluções das Atividades de Aprofundamento

Atividade de Aprofundamento

A Pelican Stores, uma divisão da National Clothing, é uma cadeia de lojas de roupas femininas que opera em todos os Estados Unidos. Recentemente, a cadeia realizou uma promoção na qual cupons de desconto foram enviados aos clientes de outras lojas da National Clothing. Dados coletados para uma amostra de 100 transações de cartões de crédito nas lojas Pelican Stores durante um dia, durante a promoção, estão no arquivo PelicanStores.csv. A Tabela abaixo mostra uma parte do conjunto de dados. O método de pagamento chamado “Cartão Proprietário” se refere a cobranças feitas utilizando um cartão de crédito da National Clothing. Os clientes que fizeram uma compra utilizando um cupom de descontos são denominados clientes promocionais e os clientes que fizeram uma compra sem utilizar um cupom de desconto são chamados clientes regulares. Como os cupons promocionais não foram enviados aos clientes Pelican Stores regulares, a administração considera as negociações feitas para pessoas apresentando cupons promocionais como vendas que, de outro modo, não teriam sido efetuadas. Naturalmente, a Pelican também espera que os clientes promocionais continuem a comprar em suas lojas.

Cliente	Tipo de Cliente	Itens	Vendas líquidas	Método de Pagamento	Gênero	Estado Civil	Idade
1	Regular	1	39,5	Discover	Masculino	Casado	32
2	Promocional	1	102,4	Cartão Proprietário	Feminino	Casado	36
3	Regular	1	22,5	Cartão Proprietário	Feminino	Casado	32
4	Promocional	5	100,4	Cartão Proprietário	Feminino	Casado	28
5	Regular	2	54	MasterCard	Feminino	Casado	34
6	Regular	1	44,5	MasterCard	Feminino	Casado	44
7	Promocional	2	78	Cartão Proprietário	Feminino	Casado	30
8	Regular	1	22,5	Visa	Feminino	Casado	40
9	Promocional	2	56,52	Cartão Proprietário	Feminino	Casado	46
10	Regular	1	44,5	Cartão Proprietário	Feminino	Casado	36
11	Regular	1	29,5	Cartão Proprietário	Feminino	Casado	48

A maioria das variáveis mostradas na tabela acima são autoexplicativas, mas duas das variáveis exigem algum esclarecimento:

A) Itens – O número total de itens comprados

B) Vendas líquidas – A quantia total (\$) cobrada pelo cartão de crédito

A administração da Pelican gostaria de utilizar esses dados amostrais para aprender a respeito do perfil de seus clientes e avaliar a promoção envolvendo cupons de desconto.

Mas antes de obter estas informações, o analista precisa construir um dicionário de dados, indicando para cada variável (em um formato de tabela, com estas colunas):

- a) Nome da Variável
- b) Descrição da Variável – o que ela mede/significa
- c) Tipo da Variável (quantitativa, qualitativa, discreta, contínua)
- d) Tipo de Mensuração da Variável (nominal, ordinal, intervalar ou racional)
- e) Valores possíveis da variável.

Depois de construído o dicionário de dados, determine as medidas de posição e variabilidade para as variáveis apropriadas. Para isso, utilize o conjunto de dados completo, fornecido no arquivo PelicanStores.csv

Resolução

Dicionário de Dados

É necessário criar este dicionário para todas as variáveis do Dataset Pelican. Preencha cada elemento, de acordo com o modelo da tabela abaixo.

Nome da Variável	Descrição	Tipo de Variável	Tipo de Mensuração	Valores Possíveis
<i>Nome da variável na base de dados</i>	<i>Descrição da variável</i>	<i>A qual tipo esta variável se encaixa, por exemplo: Numérica / Categórica / etc...</i>	<i>Como esta variável pode ser mensurada, por exemplo: Razão / Nominal / etc...</i>	<i>Quais são as representações de dados encontradas no variável, como por exemplo: Classes usadas na loja / Números Inteiros Positivos / Números reais positivos / Métodos de pagamentos aceitos pela loja / Feminino/Masculino/Não declarado (se houver esta opção) / etc...</i>

Estatística Descritiva

É necessário apresentar as Medidas de Posição (média, mediana) e Variabilidade (desvio padrão e variância) das variáveis numéricas: *Itens*, *Vendas Líquidas* e *Idade*.

Moda – O **R** não tem uma função para o cálculo da moda, então criamos uma função específica para isso, lembrando que a *moda* se aplica a variáveis numéricas discretas e variáveis categóricas. Para variáveis contínuas não faz sentido o cálculo da moda, já que tais variáveis podem assumir quaisquer valores no conjunto dos números reais.

Variáveis	Média	Mediana	DesvPadrao	Variancia	Moda
Itens	?	?	?	?	?
Vendas líquidas	?	?	?	?	?
Idade	?	?	?	?	?

Tabulações das Variáveis Categóricas (*Distribuição da Frequência de cada variável categórica encontrada no Dataset*)

```
##
## Promocional      Regular
##           70           30
##
##
## American Express Cartão Proprietário      Discover
##           2           70           4
```

##	MasterCard	Visa
##	14	10
##		
##	Feminino	Masculino
##	93	7
##		
##	Casado	Solteiro
##	84	16

Moda para as variáveis categóricas, explicitamente, utilizando nossa função definida acima.

Variáveis	Moda
Tipo de Cliente	?
Método de Pagamento	?
Gênero	?
Estado Civil	?