**ANÁLISE EXPLORATÓRIA DE DADOS**

**DESCRIBE():** Permite gerar estatísticas descritivas, tais como Contagem, Média, Desvio Padrão, etc.

**MAANTHUI:** Quantidade de Casas (*Number of houses) 1 – 10*

**MEAN()**: Média de 1.1 casas por cliente registrado.

**STD():** Desvio padrão de 0.4, considerado baixo.

**25%, 50%, 75%:** Todos são 1, porém o **MAX()** valor registrado é 10, o que indica que a grande maioria dos clientes registrados possuem apenas 1 casa.

**APERSONG:** Quantidade de Seguros Privados Contra Acidentes (*Number of private accident insurance policies*)

**MEAN():** Média de **0.005** Seguros Privados Contra Acidentes por cliente.

**STD():** Desvio padrão de **0.07**, considerado extremamente baixo.

**75%:** 0. Indica que **ao menos** 75% dos clientes registrados não possuem esse tipo de seguro.

**MAX:** 1. Indica que nenhum cliente possui mais que 1 Seguro Privado Contra Acidentes.

**ABRAND:** Quantidade de Seguros Contra Incêndios (*Number of fire policies*)

**MEAN():** Média de **0.57** Seguros Contra Incêndios por cliente.

**STD():** Desvio padrão de **0.56**, considerado baixo.

**25%:** 0. Indica que **ao menos 25%** dos clientes registrados não possuem esse tipo de apólice.

**50%:** Indica que **ao menos 50%** dos clientes registrados possuem esse tipo de apólice.

**CARAVAN:** Quantidade de apólices de mobile homes/”casas móveis”/trailers.

-A diferença entre os que possuem apólice e os que **não** possuem apólice é enorme.

-Dos 5822 registros de clientes, **5474 não possuem** apólice. **348 possuem** apólice.

**94.02%** de todas as pessoas não possuem apólice de mobile homes. **5.98%** possuem apólices de mobile homes.

**1 - GRÁFICO BARRA:**

Separa as pessoas que **possuem** ou **não possuem** apólices de mobile homes de acordo com diversas propriedades/variáveis.

**MOSTYPE:** Sub-categoria dos clientes.

**CARAVAN X MOSTYPE:**

**0: Não** possuem apólices de mobile homes. **1: Possuem** apólices de mobile homes.

**0 – NÃO POSSUEM APÓLICE:**

-A sub-categoria mais recorrente entre as pessoas que **não possuem** apólice é a 33 (Famílias grandes de classe baixa).

**764** pessoas de famílias grandes de classe baixa **não possuem** apólice.

Isso representa **13.95%** de todas as pessoas que **não possuem** apólice,

e **13.12%** de todas as pessoas no geral.

**\*13.12% de todas as pessoas registradas são de famílias grandes de classe baixa E não possuem apólice de mobile homes.**

-A segunda sub-categoria mais recorrente entre as pessoas que **não possuem** apólice é a 38 (Famílias tradicionais).

**316** pessoas de famílias tradicionais **não possuem** apólice.

Isso representa **5.77%** de todas as pessoas que **não possuem** apólice,

e **5.42%** de todas as pessoas no geral.

\***5.42% de todas as pessoas registradas são de famílias tradicionais E não possuem apólice de mobile homes.**

**1 – POSSUEM APÓLICE:**

-A sub-categoria mais recorrente entre as pessoas que **possuem** apólice é a

8 (Famílias de classe média).

**51** pessoas de famílias de classe média **possuem** apólice.

Isso representa **14.65%** de todas as pessoas que **possuem** apólice,

e **0.87%** de todas as pessoas no geral.

**\*0.87% de todas as pessoas registradas são de famílias de classe média E possuem apólice de mobile homes.**

-A segunda sub-categoria mais recorrente entre as pessoas que **possuem** apólice é a 33 (Famílias grandes de classe baixa).

**46** pessoas de famílias grandes de classe baixa **possuem** apólice.

Isso representa **13.21%** de todas as pessoas que **possuem** apólice,

e **0.79%** de todas as pessoas no geral.

**\*0.79% de todas as pessoas registradas são de famílias grandes de classe baixa E possuem apólice de mobile homes.**

**MOSHOOFD:** Tipo Principal do cliente (*Customer main type*).

**CARAVAN X MOSHOOFD:**

**0: Não** possuem apólices de mobile homes. **1: Possuem** apólices de mobile homes.

**0 – NÃO POSSUEM APÓLICE:**

**1º** -O Tipo Principal do Cliente mais recorrente entre as pessoas que **não possuem** apólice de mobile homes é a 8 (Família com adultos - *Family with grownups*).

**1474** pessoas de Famílias com Adultos **não possuem** apólice de mobile homes.

Isso representa **26.92%** de todas as pessoas que **não possuem** apólice de mobile homes, e **25.31%** de todas as pessoas no geral.

**\*25.31% de todas as pessoas registradas são de Famílias com Adultos E não possuem apólice de mobile homes.**

**2º** -O segundo Tipo Principal do Cliente mais recorrente entre as pessoas que **não possuem** apólice de mobile homes é a 3 (Família Média/Comum – *Average family*).

**827** pessoas de Família Média/Comum **não possuem** apólice de mobile homes.

Isso representa **15.1%** de todas as pessoas que **não possuem** apólice de mobile homes, e **14.2%** de todas as pessoas no geral.

\***14.2% de todas as pessoas registradas são de Família Média/Comum E não possuem apólice de mobile homes.**

**1 – POSSUEM APÓLICE:**

**1º** – O Tipo Principal do Cliente mais recorrente entre as pessoas que **possuem** apólice de mobile homes é a 8 (Família com Adultos – *Family with grownups*).

**89** pessoas de Família com Adultos **possuem** apólice de mobile homes.

Isso representa **25.57%** de todas as pessoas que **possuem** apólice de mobile homes, e **1.52%** de todas as pessoas no geral.

**\*1.52% de todas as pessoas registradas são de Família com Adultos E possuem apólice de mobile homes.**

**2º –** O segundo Tipo Principal do Cliente mais recorrente entre as pessoas que **possuem** apólice de mobile homes é a 2 (Cultivadores Ambiciosos

– *Driven growers*).

**66** pessoas Cultivadores Ambiciosos **possuem** apólice de mobile home.

Isso representa **18.96%** de todas as pessoas que **possuem** apólice de

mobile homes, e **1.13%** de todas as pessoas no geral.

**\*1.13% de todas as pessoas registradas são Cultivadores Ambiciosos**

**E possuem apólice de mobile homes.**

**2 – GRÁFICO RETÂNGULO**

Separa todas as pessoas de acordo com uma propriedade/variável.

**MOSTYPE:** Sub-categoria do cliente (*Customer subtype*).

No Treemap feito em cima da variável MOSTYPE, podemos ver que o maior retângulo de todos tem a label “33”, que representa a sub-categoria “Famílias Grandes de Classe Baixa” (*Lower class large families*). Com a contagem de **810**, esse é o número total de clientes registrados que se enquadram nessa sub-categoria.

**Com essa informação, sabemos que 810 das 5822 pessoas registradas fazem parte de Famílias Grandes de Classe Baixa, que representam 13.91% do total.**

*Levando em conta o gráfico em barra, podemos verificar que essas* ***810*** *pessoas contabilizam a soma do Número de Pessoas que* ***NÃO POSSUEM APÓLICE DE MOBILE HOMES (764)*** *e o Número de Pessoas que* ***POSSUEM APÓLICE DE MOBILE HOMES (46).***

O menor retângulo de todos tem a label “19”, que representa a sub-categoria “Juventude do Subúrbio” (*Suburban youth*). Com a contagem de **3**, esse é o número total de clientes registrados que se enquadram nessa sub-categoria.

**Com essa informação, sabemos que 3 das 5822 pessoas registradas fazem parte de Juventude do Subúrbio, que representam 0.051% do total.**

Levando em conta os dados do gráfico em barra, todas as **3** pessoas da sub-categoria “Juventude do Subúrbio” **POSSUEM APÓLICE DE MOBILE HOMES**. Ou seja, **100%** das pessoas que se enquadram nessa categoria possuem uma apólice de mobile homes.

**MGEMOMV:** Tamanho médio da residência (1-6) (*Average size household).*

A variável **MGEMOMV** representa o tamanho médio da residência do cliente, que foram divididos de 1 a 6. Ou seja, quanto menor o número, menor o tamanho médio da residência.

No Treemap podemos ver que a maior parte dos clientes moram em residências de tamanho médio de 2 a 4. Juntos, esses clientes somam **5470**. Ou seja, **93.95%** de todos os clientes registrados moram em residências de tamanho médio entre 2 a 4.

**MGEMLEEF:** Idade média do cliente (1-6) (*Average age*).

A variável **MGEMLEEF** representa a faixa etária média do cliente, que foram divididos de 1 a 6:

1: entre 20 a 30 anos

2: entre 30 a 40 anos

3: entre 40 a 50 anos

4: entre 50 a 60 anos

5: entre 60 a 70 anos

6: entre 70 a 80 anos

De acordo com o gráfico, podemos que ver que os 3 maiores retângulos representam as labels 2, 3, e 4, as faixas etárias de 30 a 60 anos. Essas três faixas etárias totalizam **5525** clientes, que representam **94.89%** do total de clientes registrados.

**MOSHOOFD:** Tipo Principal do Cliente (*Customer main type*).

A variável **MOSHOOFD** representa o Tipo Principal do Cliente, divididos em 10 categorias.

Como podemos ver no gráfico, as três categorias mais recorrentes são as de label 8, 3 e 9, respectivamente. Elas representam as categorias Família com Adultos (*Family with grownups*), Família Média/Comum (*Average family*), Famílias Conservadoras (*Conservative families*).

Juntos, essas três categorias totalizam **3116** clientes, que são **53.52%** de todos os clientes registrados.