**Entrega 3: Análise Descritiva de Dados**

**Análise Exploratória**

**Análise Exploratória dos Dados de Roubos de Veículos nos Estados**

Os dados utilizados nesta análise representam o número de ocorrências de roubos de veículos por estado/cidade no Brasil. Essas informações estão organizadas na planilha Excel juntamente com o gráfico Box Plot, que ilustra visualmente a distribuição desses dados.

Foi realizada uma análise exploratória sobre a quantidade de roubos de veículos em diferentes localidades dos estados/cidades pelo Brasil. O objetivo é entender a distribuição dos dados, identificar padrões e valores extremos (outliers), e concluir se os dados estão concentrados em valores baixos, médios ou altos.

* **Mediana**

Como a mediana é o valor central. Quando o número de observações é par (**324**), a mediana será a média dos valores nas posições 162 e 163.

* **Valor da posição 162:** 128
* **Valor da posição 163:** 129

A mediana é 128,5, indicando que metade das cidades tem até esse número de ocorrências.

* **1º Quartil (Q1)**

O primeiro quartil é a mediana da metade inferior dos dados (valores da posição 1 a 162). Como são 162 valores, Q1 será a média dos valores nas posições 81 e 82.

* **Valor da posição 81:** 54
* **Valor da posição 82:** 54

Portanto, o 1º quartil (Q1) é 54, indicando que 25% das cidades registram até esse número de ocorrências.

* **3º Quartil (Q3)**

O terceiro quartil é a mediana da metade superior dos dados (valores da posição 163 a 324).

Como são 162 valores, Q3 será a média dos valores nas posições 243 e 244.

* **Valor da posição 243:** 336
* **Valor da posição 244:** 337

Portanto, o 3º quartil (Q3) é 336,5, indicando que 75% das cidades registram até esse número de ocorrências.

* **Identificação e Exclusão de Outliers**

Para identificar os valores fora da curva, utilizamos os limites calculados a partir do IQR:

* **IQR (intervalo)**
* **Limite Inferior:**

Como o limite inferior calculado resultou em um valor negativo, não há outliers inferiores, pois não existem quantidades negativas de ocorrências.

* **Limite Superior:**

Todos os valores acima de 760,25 foram considerados outliers e, portanto, excluídos da análise final, por estarem significativamente acima da distribuição principal dos dados.

**Conclusão da Análise Exploratória**

Após a exclusão dos outliers (valores acima de 760,25), observou-se que a maioria das cidades apresenta entre 54 e 336,5 ocorrências de roubos de veículos. A mediana de 128,5 indica que metade desses estados registra valores iguais ou inferiores a esse número.

Os dados estão concentrados principalmente entre o primeiro e o terceiro quartil, e a distribuição mostra assimetria à direita, pois ainda há valores elevados próximos ao limite superior.

**Conclusão sobre a concentração dos dados (valores altos, baixos, médios)**

Os dados indicam uma concentração em valores médios a baixos. A maioria dos estados apresenta até 336,5 ocorrências, com mediana de 128,5. Alguns possuem valores elevados de roubos, que se destacam como outliers. Isso indica que a criminalidade está mais concentrada em poucas regiões, enquanto a maior parte registra índices relativamente baixos.

**Probabilidade de ocorrência**

Com base no conjunto de dados analisado, que compreende as ocorrências de roubos de veículos em 324 cidades no Brasil, deseja-se calcular a probabilidade de uma cidade apresentar mais de 759 ocorrências de roubos de veículos.

**Resolução:**

* **Total de cidades analisadas:** 324
* **Número de cidades com mais de 759 ocorrências:** 40

Para encontrar essa probabilidade, aplica-se a fórmula básica de probabilidade:

**Conclusão:**

A chance de uma cidade registrar mais de 759 ocorrências de roubos de veículos é de aproximadamente 12,35%, o que confirma que a maioria apresenta números abaixo desse valor.