**Análise dos problemas apresentados para aula Hands-on Data Science**

**🧠 1. Enunciado do Problema**

Queremos saber se a **preferência por tipo de transporte** (carro, ônibus ou bicicleta) é **diferente entre dois bairros** da cidade: **Centro** e **Periferia**.

**🔍 2. Objetivo da Análise**

Avaliar se existe **associação** entre o bairro em que a pessoa mora e sua preferência de transporte. Em outras palavras:

“A escolha de transporte depende do bairro?”

**📊 3. Tipo de Dados e Variáveis**

| **Elemento** | **Descrição** | **Tipo de Variável** | **Papel** |
| --- | --- | --- | --- |
| Bairro | Centro ou Periferia | **Qualitativa Nominal** | **Variável Independente** (fator) |
| Transporte | Carro, Ônibus, Bicicleta | **Qualitativa Nominal** | **Variável Dependente** |

* Estamos observando a distribuição de uma variável **dependente categórica** (preferência de transporte) **em relação**a uma variável **independente categórica** (bairro).

**👥 4. Tipo de Grupos**

* **Grupos não pareados** (independentes):
  + Os indivíduos do bairro Centro não são os mesmos do bairro Periferia.
  + Não há correspondência ou relação direta entre as unidades amostrais dos dois grupos.

**🧪 5. Hipóteses Estatísticas (Teste do Qui-quadrado de Independência)**

**Hipótese Nula (H₀):**

A preferência por tipo de transporte é independente do bairro.  
Ou seja, a distribuição de transporte é igual nos dois bairros.

**Hipótese Alternativa (H₁):**

A preferência por tipo de transporte depende do bairro.  
Ou seja, a distribuição de transporte difere entre Centro e Periferia.

**🧰 6. Teste Estatístico Utilizado**

**📌 Teste do Qui-quadrado de Independência**

* **Motivo da escolha:**  
  Comparar frequências observadas e esperadas em uma tabela de contingência de **duas variáveis categóricas nominais**.
* **Pré-requisitos atendidos:**
  + Dados categóricos nominais
  + Tamanho da amostra adequado (n=300)
  + Grupos independentes
* **Ferramenta usada:**  
  scipy.stats.chi2\_contingency

**📈 7. Resumo dos Resultados**

| **Bairro** | **Bicicleta** | **Carro** | **Ônibus** |
| --- | --- | --- | --- |
| **Centro** | 22 | 75 | 53 |
| **Periferia** | 15 | 29 | 106 |

* **Valor-p**: 0.0000

**✅ 8. Conclusão da Análise**

Como **p < 0.05**, rejeitamos a hipótese nula:

**Conclusão**: Existe uma **diferença estatisticamente significativa** na preferência por transporte entre os bairros.

**📚 9. Interpretação Didática para Sala de Aula**

Imagine que você está analisando uma política de mobilidade urbana. Este teste mostra que bairros diferentes **usam o transporte de formas diferentes**. Portanto, **ações públicas não podem ser genéricas**, pois o perfil da população muda conforme o bairro.