

Coleta de Dados - Parte 2

Projeto Interdisciplinar I

Gestão da Tecnologia da Informação (GTI)
Instituto Federal do Paraná (IFPR)
Campus Pinhais

1 Introdução

2 Viés

- Na aula passada falamos sobre alguns métodos de coleta de dados
- Documental, Observacional e Entrevista



- Erro sistemático ou distorção no processo de estudo (na coleta, análise ou interpretação de dados) que afasta o resultado da realidade
- Como um viés, geralmente, surge:
 - Na escolha dos participantes
 - Na forma de perguntar ou medir
 - Na interpretação do cientista

- Ocorre quando os participantes selecionados não representam adequadamente o grupo que queremos estudar
- *"Como estudantes universitários percebem o uso de Inteligência Artificial na educação?"*
- Você divulga o questionário apenas em um grupo de WhatsApp sobre IA e provavelmente responderão mais pessoas que:
 - já usam IA
 - têm interesse no assunto
 - possuem opinião forte sobre IA
 - têm maior familiaridade tecnológica
 - ...

- *"Como estudantes universitários percebem o uso de Inteligência Artificial na educação?"*
- A pergunta está ampla e complexa
- Deve-se repensar a pergunta ou repensar o modo de seleção de participantes

- Acontece quando a maneira como uma informação é perguntada ou medida influencia sistematicamente o resultado
- Exemplos:
 - Pergunta que sugere uma resposta
 - Opções de resposta desequilibradas
 - Instrumento de medição inadequado
 - Escala mal construída
- Consequência: os dados podem não representar corretamente o fenômeno estudado.

- *"Você concorda que a IA pode **melhorar** a educação?"*
- *"Você acredita que o uso de IA pode trazer **riscos** para a educação?"*
- A presença de determinados termos pode influenciar o participante durante a resposta
- Mesmo que não tenha nenhuma crítica ou benefício, pode fazer com que o participante liste algo para responder a pergunta

- Para evitar esse tipo de viés, é importante postular perguntas seguindo uma diretriz de neutralidade:
- "Você não acha que o novo sistema é mais eficiente?" – > "Como você avalia a eficiência do novo sistema?"

- Caso seja impossível introduzir um termo, pode-se sempre colocar as duas opções:
- "Quais são os principais desafios e os principais benefícios do sistema?"

- A medida das respostas pode não ser adequada
- Por exemplo, se você quer medir a satisfação dos usuários de um sistema, mas oferece apenas:
 - Muito satisfeito
 - Satisfeito
 - Neutro
- Não consegue coletar visões críticas do sistema

- A definição das medidas deve ser sempre adequada ao que se está medindo
- Como regra, pode-se observar a dimensão do objeto a ser mensurado e adotar alguma escala (como a escala Likert)

Atenção

A escala deve ser adequada àquilo que se pretende medir. Uma escala mal construída pode introduzir **viés na coleta dos dados**.

Concordância

Concordo totalmente
Concordo
Nem concordo, nem discordo
Discordo
Discordo totalmente

Frequência

Sempre
Frequentemente
Às vezes
Raramente
Nunca

Importância

Extremamente importante
Muito importante
Moderadamente importante
Pouco importante
Nada importante

Probabilidade

Muito provável
Provável
Nem provável, nem improvável
Improvável
Muito improvável

- Mesmo que os dados tenham sido coletados corretamente, a interpretação que fazemos deles pode ser influenciada pelas nossas expectativas, crenças ou interesses
- Exemplos:
 - Procurar evidências que confirmem uma hipótese
 - Ignorar resultados que contradizem a expectativa
 - Dar destaque apenas aos resultados favoráveis
 - Confundir correlação com causalidade
 - Interpretar dados ambíguos de maneira conveniente

- Exemplo: 60% dos usuários avaliaram um sistema como "bom" ou "muito bom".
 - Pesquisador A: "A maioria dos usuários aprovou o sistema"
 - Pesquisador B: "40% dos usuários não avaliaram o sistema positivamente"
- Como reduzir?
 - Definir critérios de análise antes de observar os resultados
 - Considerar evidências contrárias à hipótese
 - Tornar métodos e dados transparentes
 - Submeter a análise à revisão de outras pessoas