

De Dados a Decisões: Como Evitar Armadilhas Comuns em Pesquisas Estatísticas

Alexandre Loureiros Rodrigues

08/11/2024



**XV Semana de
Estatística**

7 a 8 de novembro de 2024. Vitória, ES.

Carreira profissional

- Graduação em estatística - UFES
- Mestrado em Estatística - UFMG
- PhD em estatística - Lancaster University
- Professor da UFES desde 2010
- Experiência em projetos de pesquisa com grande empresas



Agenda

1. Introdução
2. Pesquisas estatística
3. Dinâmica
4. Qualidade de dados
5. Considerações finais

1 - Introdução

- Segundo Callyampudi Radhakrishna (C.R.) Rao, um dos maiores estatísticos do século XX, a estatística pode ser descrita pela equação:



1 - Introdução

- a. Baixar aplicativo **Coin Toss**
- b. Se der "cara" o responder honestamente à pergunta

Você já tentou colar em alguma prova / atividade do curso de estatística?

- c. Se der "coroa", responder automaticamente "sim".
- d. Resposta usando o celular na plataforma **polleverywhere**

1 - Introdução

- $SIM = (SIM \text{ e Cara}) \text{ ou } (SIM \text{ e Coroa})$
- $P(SIM) = P(SIM | Cara)P(Cara) + P(SIM|Coroa)P(Coroa)$
- $P(SIM | Cara) = [P(SIM) - 1 * \frac{1}{2}] / \frac{1}{2} = 2*[P(SIM) - \frac{1}{2}]$

2 - Pesquisas estatística

- Cada estudo / pesquisa tem um objetivo. Todo foco é para responder uma pergunta de interesse.

Ex.: Pesquisa nos EUA aponta que tomar cervejas todos os dias faz bem para o coração.



2 - Pesquisas estatística

Nova pesquisa mostra que bebedores de cerveja tem maior risco de câncer.



2 - Pesquisas estatística

Uma pesquisa realizada pela universidade de Washington, cidade de Seattle, EUA: Maridos que lavam a louça da casa fazem menos sexo. 7 mil pessoas dos EUA e da Espanha participaram dessa pesquisa. Primeiramente, o casal dividia as tarefas da casa, e depois contavam quantas vezes faziam sexo no mês. Depois, o homem parava de ajudar em casa e aí o casal continuava a contar o número de vezes em que faziam sexo no mês.



2 - Pesquisas estatística

Uma pesquisa feita na universidade de Umea, na Suécia, divulgada em outubro do ano passado, diz o seguinte: homem que lava louça é mais feliz.

AGORA JUNTE OS
RESULTADOS DA DUAS
PESQUISAS



2 - Pesquisas estatística

INFORMAÇÃO:

Intenção de voto de diferentes estados dos EUA: Os dados mostram que estados com uma maior proporção de população negra tendem a votar majoritariamente no Partido Democrata.

QUAL CONCLUSÃO PODEMOS TIRAR ?

2 - Pesquisas estatística

FALÁCIA ECOLÓGICA

erro estatístico que ocorre quando conclusões sobre indivíduos são tiradas usando dados de grupos.



3 - Dinâmica - Cenário 1

EQUIPAMENTOS DE DIAGNÓSTICO

Uma clínica precisa decidir quais novos equipamentos de diagnóstico comprar. Os dados mostram que 70% dos pacientes necessitam de exames de imagem, mas não especificam os tipos de exames necessários

3 - Dinâmica - Cenário 2

REDUÇÃO DO TEMPO DE ESPERA

Um hospital deseja reduzir o tempo de espera na emergência. Os dados mostram um aumento de 40% no tempo de espera médio, mas não identificam as principais causas.

3 - Dinâmica - Cenário 1

EQUIPAMENTOS DE DIAGNÓSTICO

Uma clínica precisa decidir quais novos equipamentos de diagnóstico comprar. Os dados mostram que 70% dos pacientes necessitam de exames de imagem, mas não especificam os tipos de exames necessários

Dos pacientes que necessitam de exames de imagem, 70% precisam de ressonâncias magnéticas, 20% de tomografias computadorizadas e 10% de ultrassons.

3 - Dinâmica - Cenário 2

REDUÇÃO DO TEMPO DE ESPERA

Um hospital deseja reduzir o tempo de espera na emergência. Os dados mostram um aumento de 40% no tempo de espera médio, mas não identificam as principais causas.

"O aumento no tempo de espera é devido à falta de pessoal em 10%, aumento no número de pacientes em 15% e problemas de no processo de classificação em 75%.

4 - Qualidade dos dados

- Dados de alta qualidade são cruciais para a tomada de decisões informadas.
- Dados de baixa qualidade podem levar a conclusões erradas, custos elevados e oportunidades perdidas.



4 - Qualidade dos dados

Problemas Comuns

- Decisões Erradas: Dados imprecisos levam a decisões incorretas.
- Ineficiência Operacional: Processos baseados em dados ruins são ineficazes.
- Perda Financeira: Custos associados à correção de erros e ineficiências.

4 - Qualidade dos dados

Exemplo real de uma operadora de saúde

- Identificação de prestadores de serviço
 - Identificar prestadores de serviços (clínicas / médicos) em diferentes atendimentos
 - CNPJ do prestador muitas vezes não preenchido
 - Quando preenchido, pode apresentar erros
 - Associação dada tipicamente pelo nome do prestador digitado

EX: Dados 1: angiovasc bahi - Dados 2: clinica angiovascular bahia

SÃO OS MESMOS PRESTADORES ?

4 - Qualidade dos dados

Como avaliar algum procedimento ?

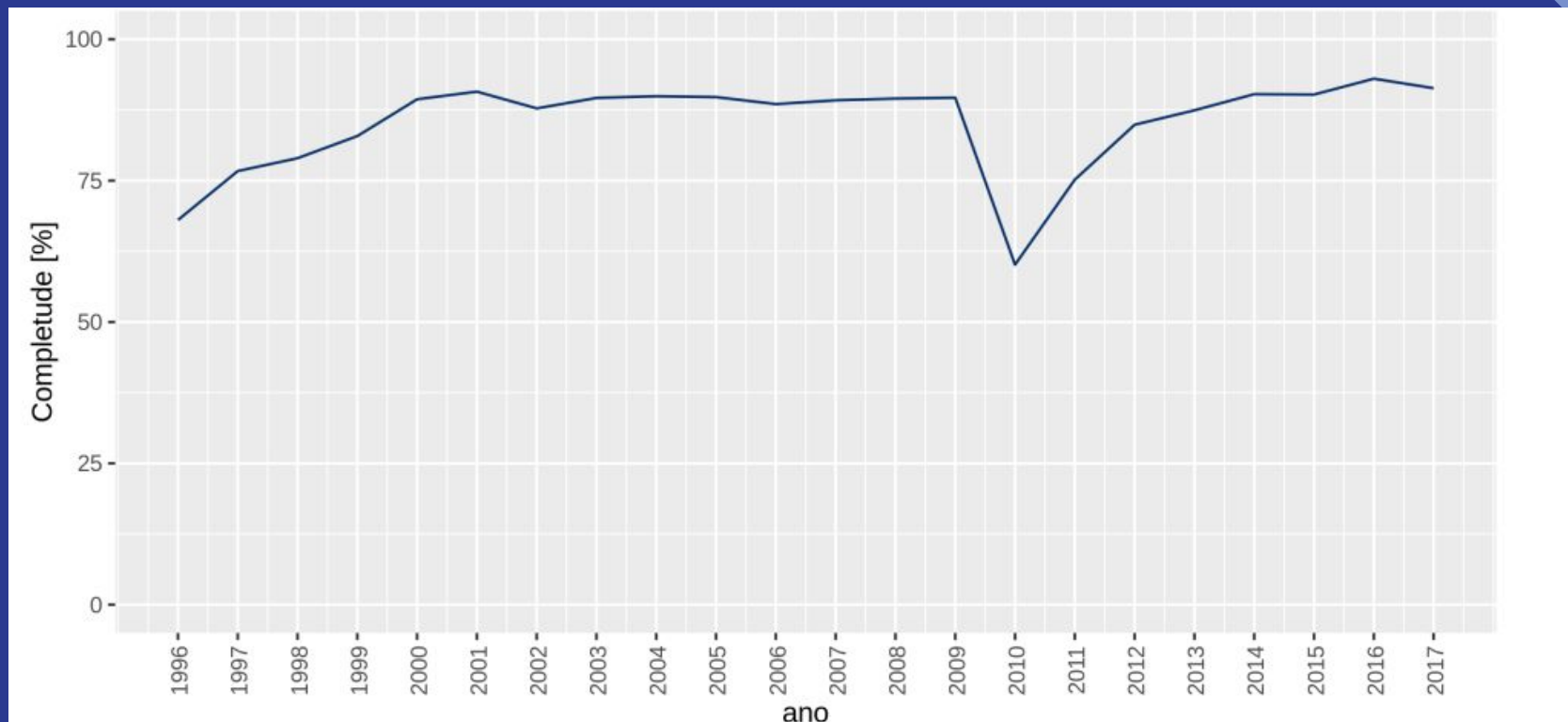
Pense em uma dieta, como saber se ela está dando resultado ?



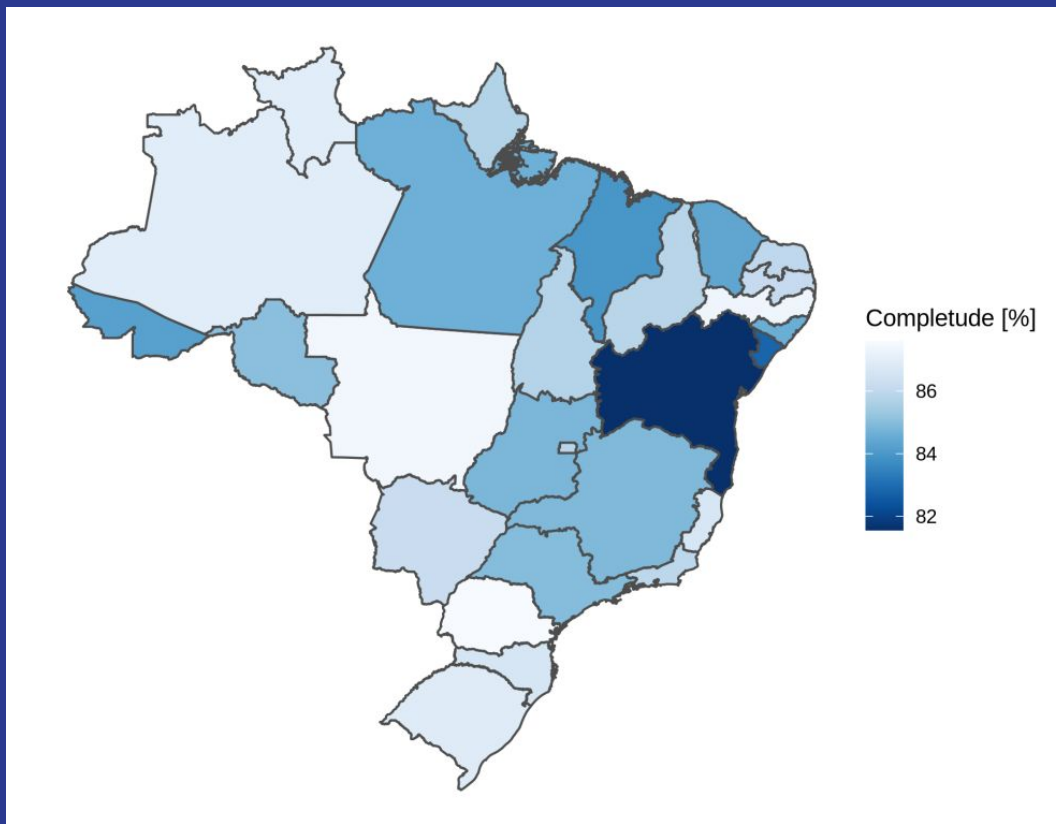
4 - Qualidade dos dados

Base de dados	SINASC
Fonte	< http://ftp.datasus.gov.br/dissemin/publicos/SINASC >
Data de obtenção dos dados	10 de março de 2020
Período	1996 a 2017
Região geográfica	Todas as 27 Unidades Federativas
Volume	1,8 GB
Número máximo de variáveis	74
Número de registros	65.873.856

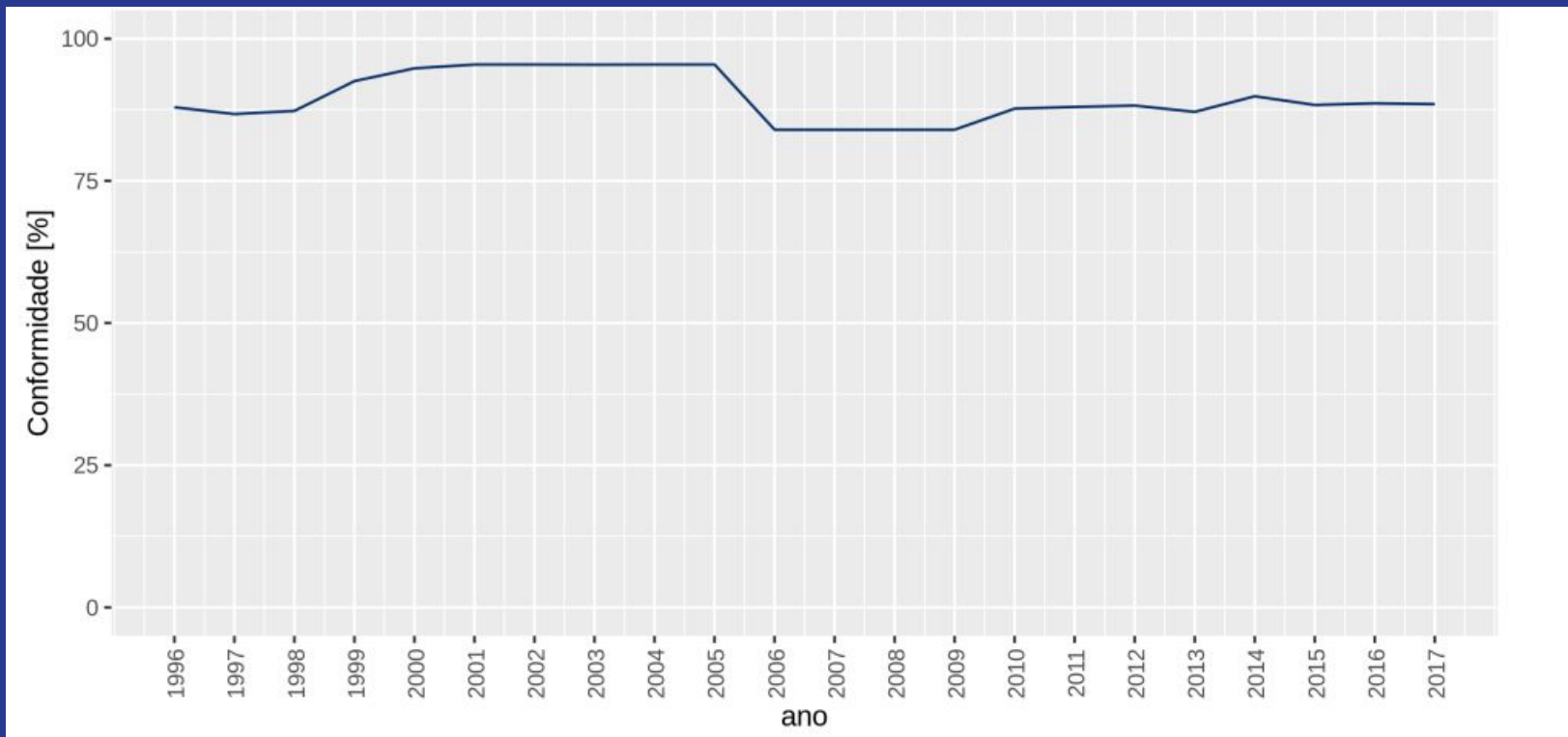
4 - Qualidade dos dados



4 - Qualidade dos dados



4 - Qualidade dos dados



4 - Qualidade dos dados

Exemplo de Inconsistências

dtnasc > dtcadastro

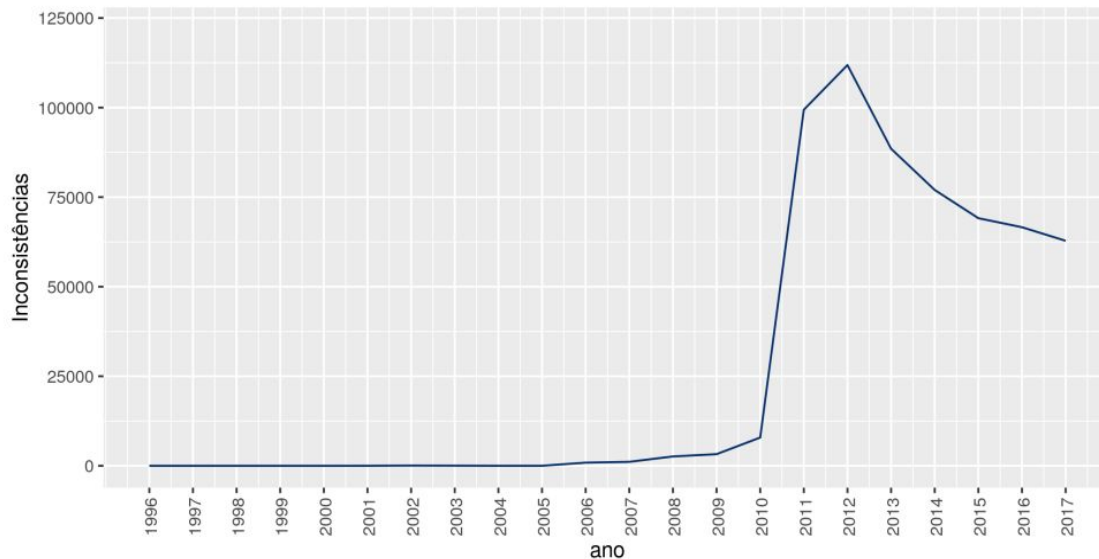
dtnasc < dtultmenst

dtnasc > dtrecoriga

dtnasc > dtrecebim

dtnasc < dtnascmae

escmae > idade_mae



4 - Qualidade dos dados

Temporalidade

- Diferença entre data de nascimento e as variáveis *dfcadastro* e *dtrecebim*

Variável	Média	Mediana
<i>dfcadastro</i>	36,35	17
<i>dtrecebim</i>	52,13	41

4 - Qualidade dos dados

Estratégias para garantir qualidade dos dados

- Governança de Dados: Políticas e procedimentos claros.
- Validação e Limpeza: Ferramentas e processos para verificar e corrigir dados.
- Treinamento: Capacitar funcionários na importância e técnicas de gestão de dados.
- Monitoramento Contínuo: Auditorias regulares e sistemas de monitoramento.

5 - Considerações finais

- A qualidade dos dados é fundamental para o sucesso das organizações.
- Investir em qualidade de dados traz inúmeros benefícios e evita riscos significativos.
- Adotar boas práticas de gestão de dados é crucial para manter a integridade e a utilidade dos dados.

