



Formação cientista de dados

Estatística I: Amostragem





Amostra

- Parte de uma população, selecionada usando alguma técnica que de chances iguais a todos os elementos da população de serem selecionados
-

Conceitos



População: alvo do estudo



Amostra: subconjunto da população

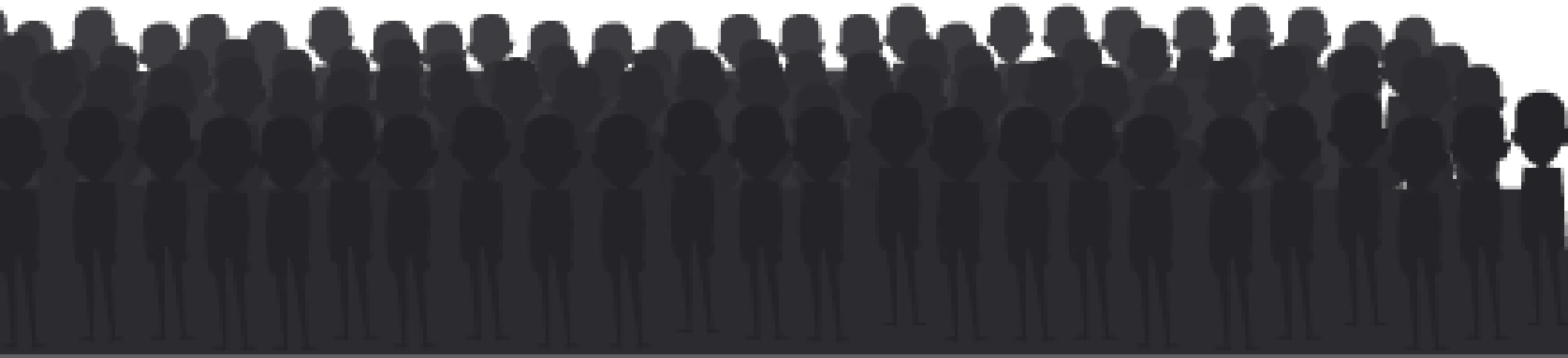


Censo: pesquisa com toda a população



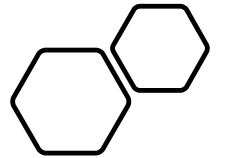
Amostra <> População!





Porque Amostra?

Pode ser caro ou impossível inferir sobre toda a população (censo)





É possível inferir
sobre uma
amostra

- Uma amostra feita corretamente deve representar as mesmas características da população de onde foi retirada.
- Se ela não representa a população, dizemos que ele é enviesada



Enviesamento



Você subestima ou superestima o parâmetro da população



Causas:

Pesquisa de pessoas próximas ou de fácil acesso

Pesquisas pela Internet

Sem uso de mecanismo de seleção aleatório

“Custo” da Amostra

- Margem de Erro e Nível de Confiança
- Variação: amostrar diferentes podem apresentar resultados diferentes
- Podemos “medir” a variação esperada



**MUNDO
REAL**

Amostragem

- Amplamente utilizada em pesquisas, estudos etc.
- Faz parte do dia a dia de um Cientista de Dados