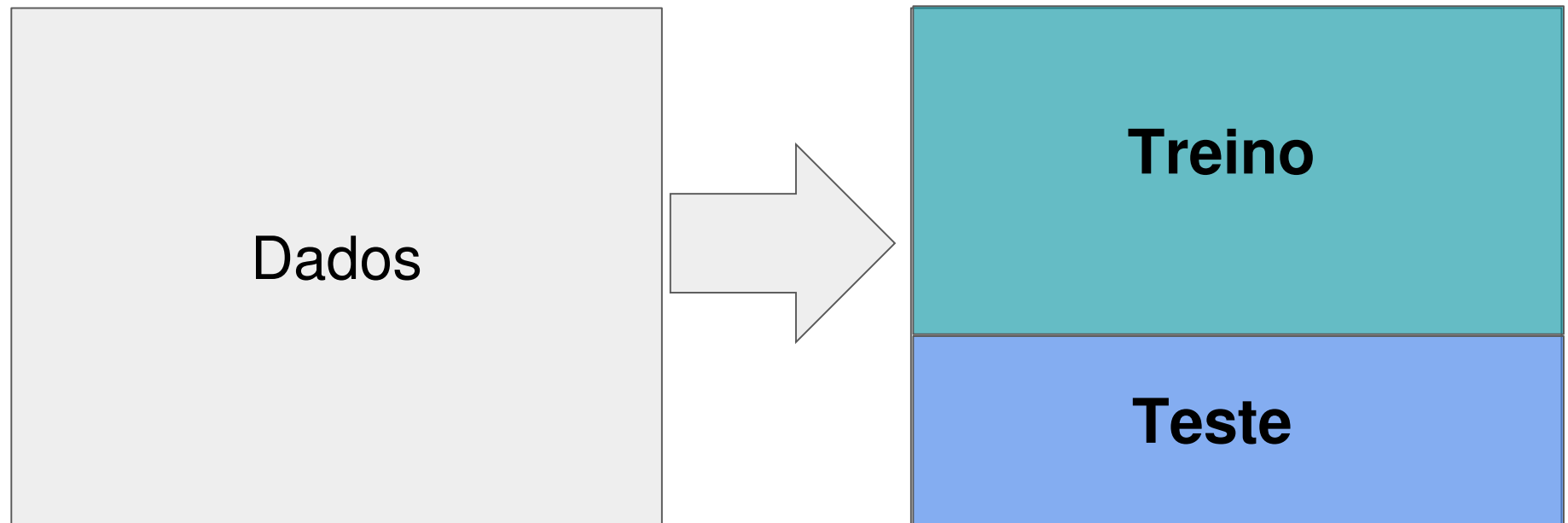


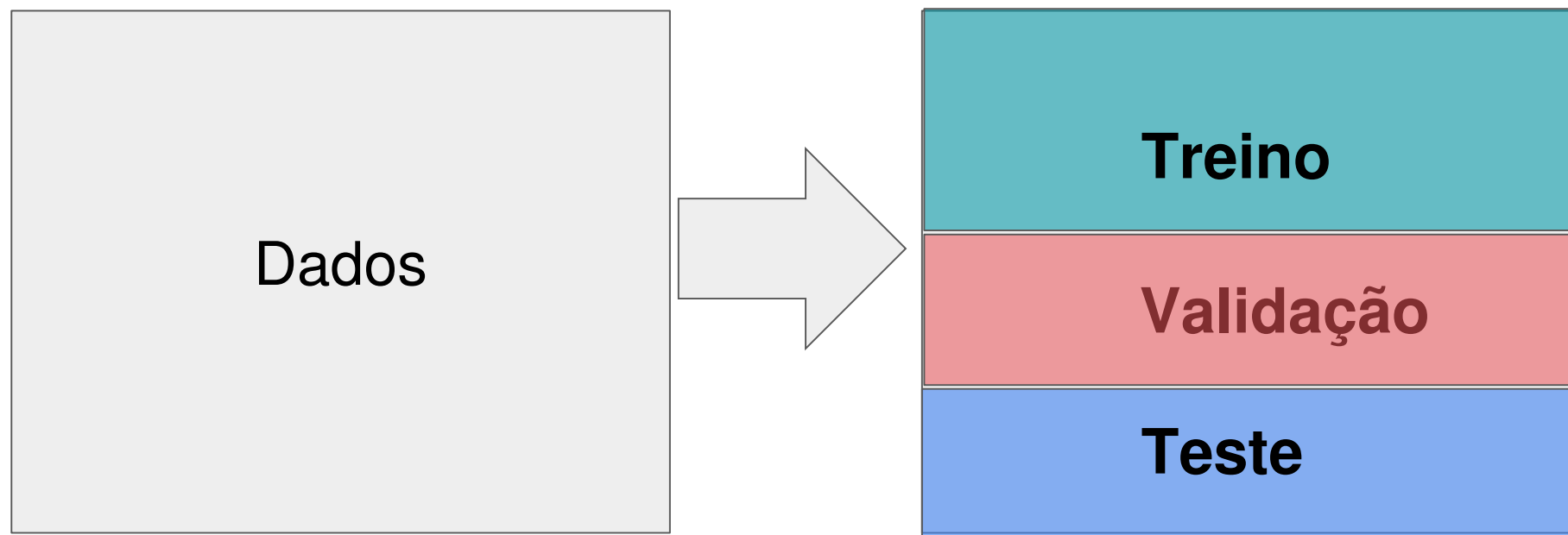
MAIS SOBRE DATA-SPLITTING



Data Splitting



Data Splitting



Nested CV

Dados



Tamanho dos bancos



LASSO



Esparsidade

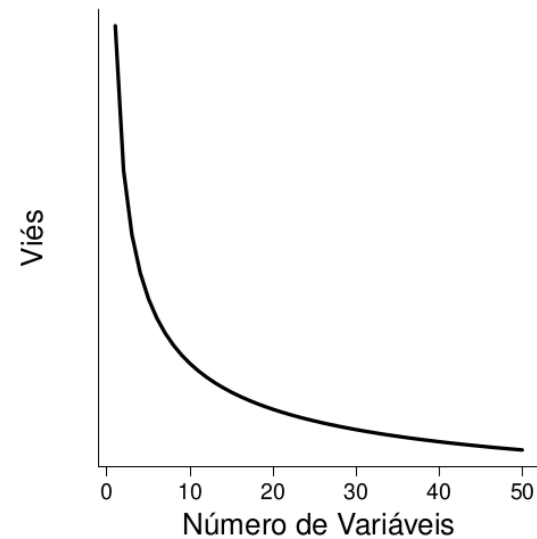
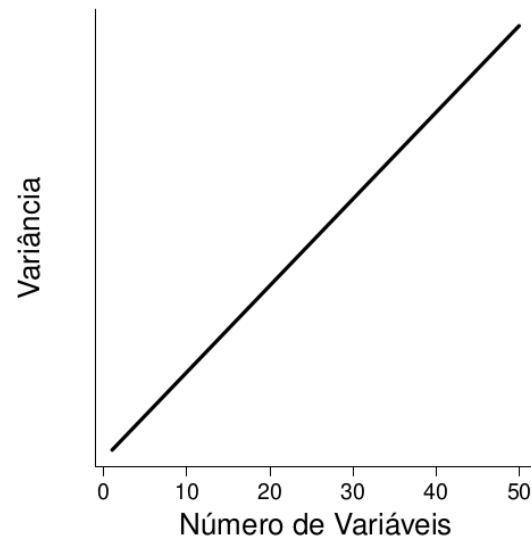
$$\text{Risco} = \text{Variabilidade Intrínseca} + \text{Variância} + \text{Viés}$$

- Redução na variância para melhorar o risco
- Usual que muitos features não sejam relevantes



Esparsidade

$$\text{Risco} = \text{Variabilidade Intrínseca} + \text{Variância} + \text{Viés}$$



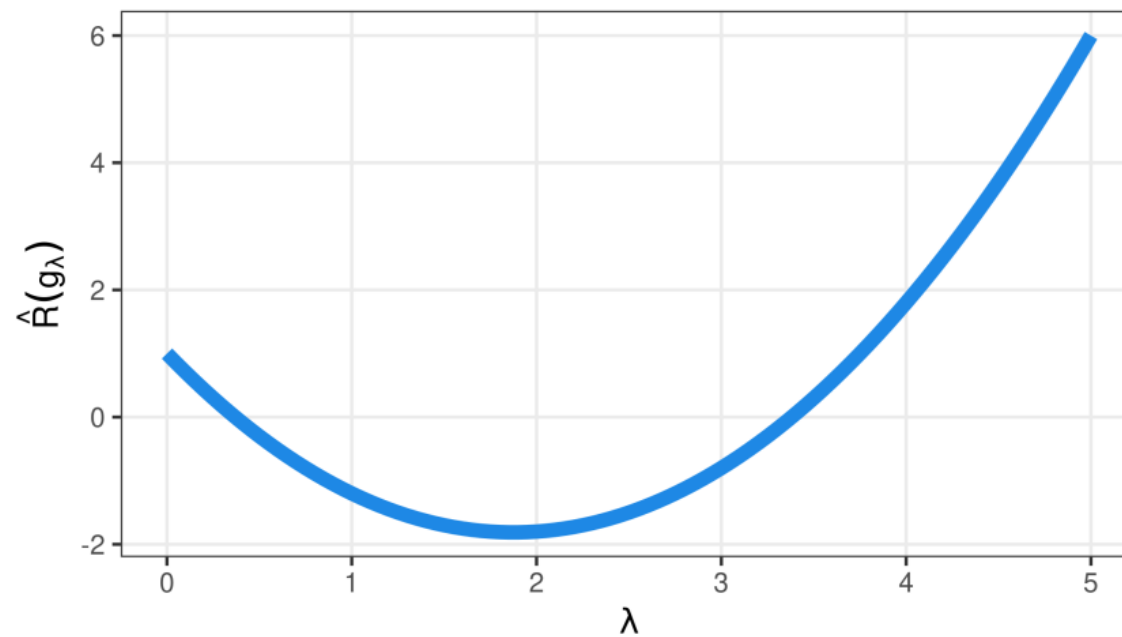
Lasso

$$\arg \min_{\beta} \text{EQM}(g_{\beta}) + \lambda \sum_{j=1}^d |\beta_j|$$

Shiny



Como escolher λ ?



Amazon



Outros Métodos (overview)

