

# METODOLOGIA DE DESAGREGAÇÃO

Dados faltantes; natureza da amostra; banco de dados; aprendizado de máquina e metodologias.

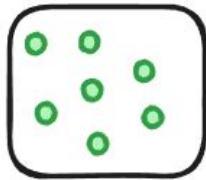
# SUMÁRIO

1. Contextualizando o processo de desagregação do déficit habitacional
  - a. Natureza das amostras: dados faltantes, mudança de escala e problema de extração de domínio;
  - b. Aprendizado de máquina supervisionado e variáveis de saída;
  - c. Dependência das métricas estatísticas em nível agregado.
1. Analisando e discutindo metodologias possíveis para a desagregação dos dados;
1. Processo de propensão: análise de propensão em déficit total e análise residual;
1. Próximos passos.

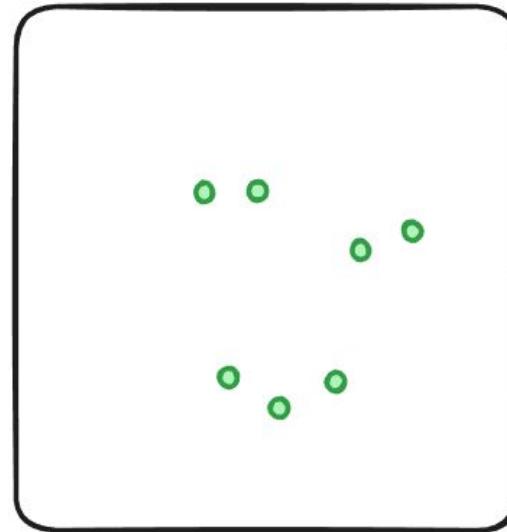
# 1. CONTEXTUALIZANDO O PROCESSO DE DESAGREGAÇÃO DO DÉFICIT HABITACIONAL

# PROBLEMA DE EXTRAPOLAÇÃO DE DOMÍNIO

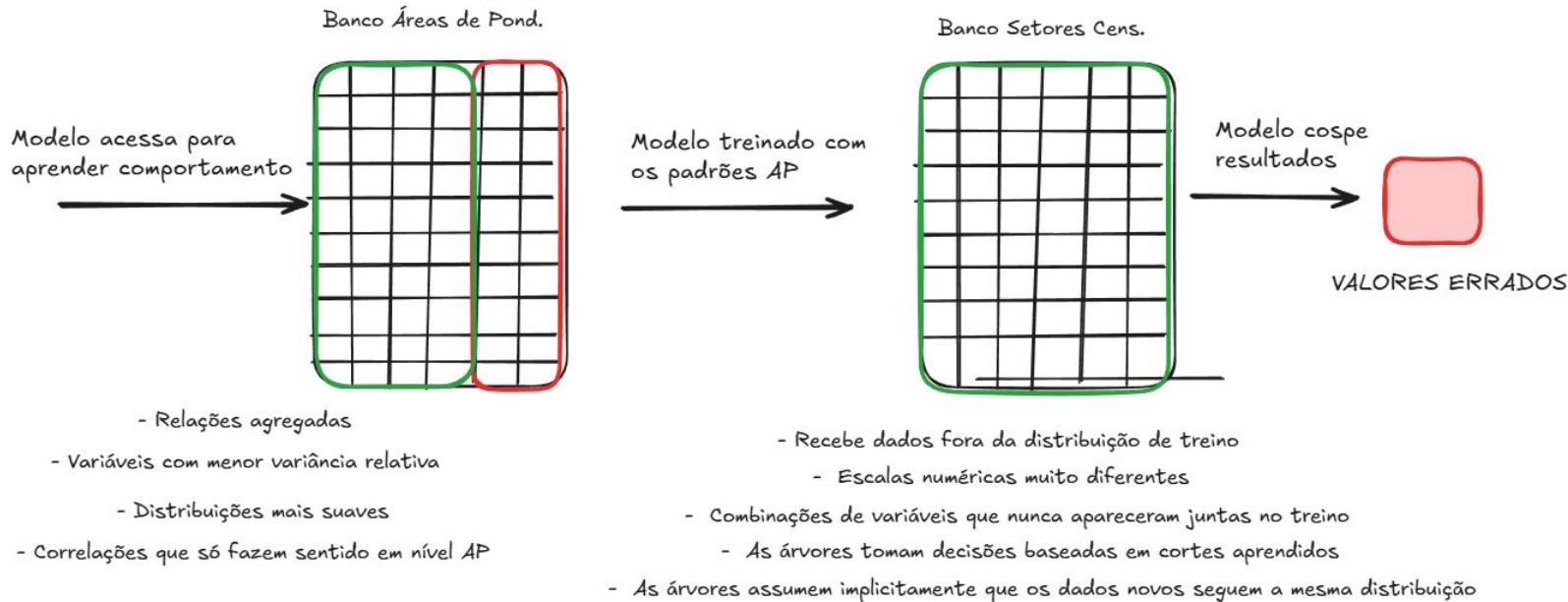
Banco Áreas de Pond.



Banco Setores Cens.

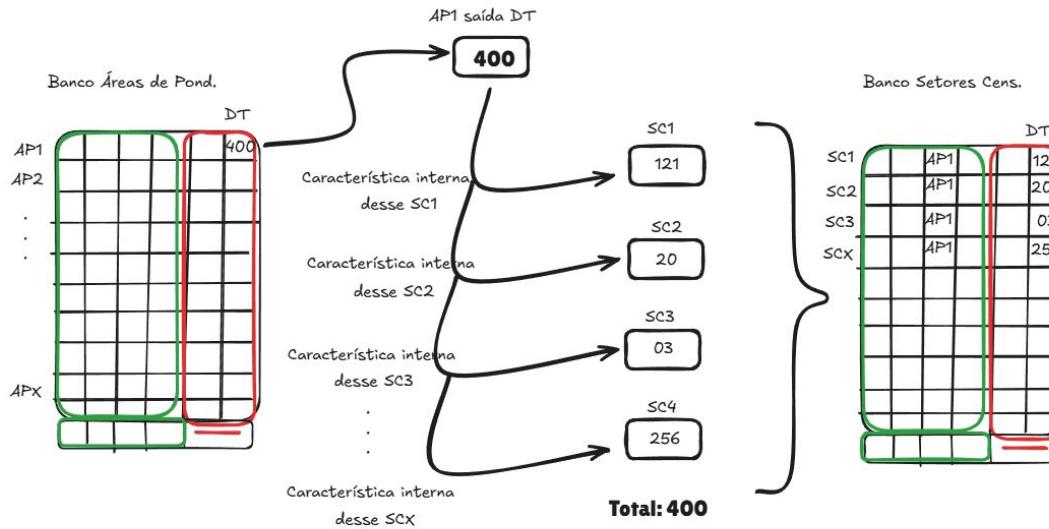


# PROBLEMA DE EXTRAPOLAÇÃO DE DOMÍNIO



## 2. ANALISANDO E DISCUTINDO METODOLOGIAS POSSÍVEIS PARA A DESAGREGAÇÃO DOS DADOS

# MÉTODO DE PROPENSÃO



Mesmo modelo em escalas diferentes:

- Ruim para estimar magnitude absoluta correta (predição)
- Não necessariamente ruim para ordenar, comparar ou ponderar observações dentro de um mesmo contexto



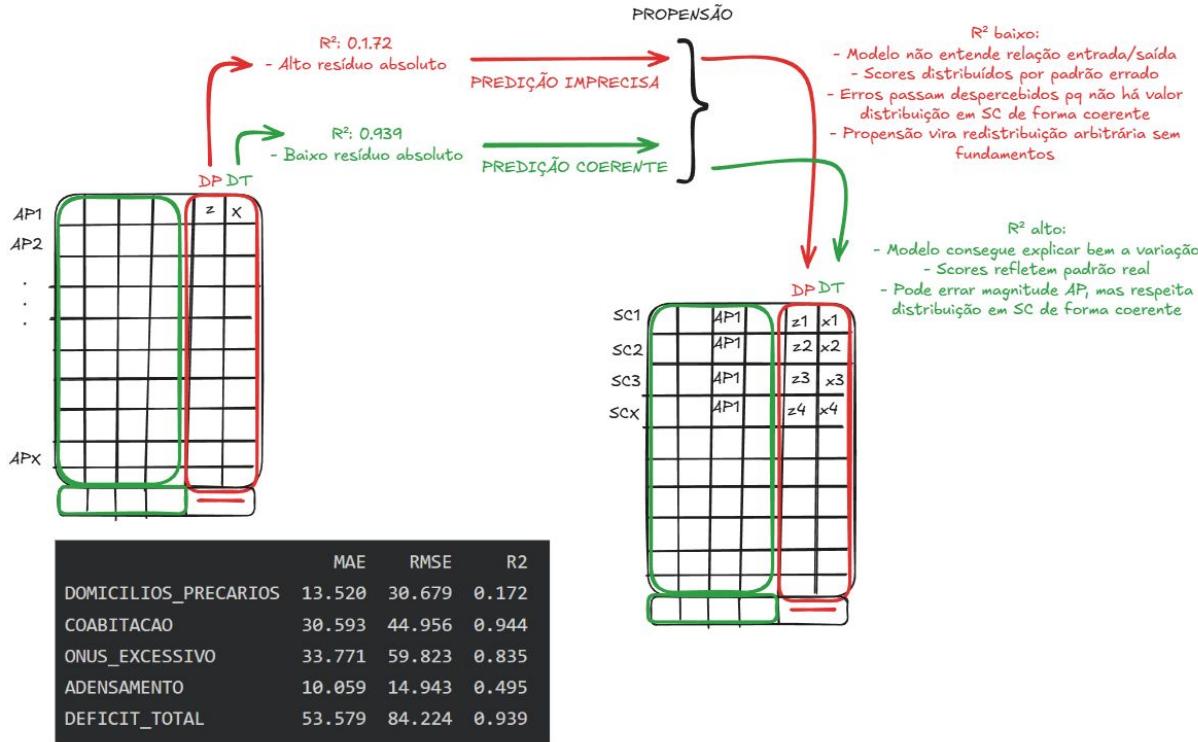
"Mais soma\_X → tende a aumentar o déficit"

"Combinação de variáveis A, B e C tende a ser mais crítica"

"Certos perfis socioeconômicos são mais vulneráveis que outros"

Isso é, não é bom para prever valores absolutos, mas preserva interações entre as entradas (que são as mesmas)

# MÉTODO DE PROPENSÃO



### 3. PROCESSO DE PROPENSÃO: ANÁLISE DE PROPENSÃO EM DÉFICIT TOTAL E ANÁLISE RESIDUAL

...	COD_SC	COD_AP	DEFICIT_TOTAL_PROPENSO	PESO	FONTE_VALOR_AP	RESIDUO_AP
0	310010405000001	3100104001001	20.548823	0.156861	real	6.52000
1	310010405000002	3100104001001	25.143032	0.191932	real	6.52000
2	310010405000003	3100104001001	25.673518	0.195981	real	6.52000
3	310010405000004	3100104001001	29.592107	0.225894	real	6.52000
4	310010405000005	3100104001001	30.042520	0.229332	real	6.52000
5	310020305000001	3100203001001	33.666478	0.041718	real	39.81000
6	310020305000002	3100203001001	18.346697	0.022734	real	39.81000
7	310020305000003	3100203001001	15.967992	0.019787	real	39.81000
8	310020305000004	3100203001001	34.378353	0.042600	real	39.81000
9	310020305000005	3100203001001	36.571854	0.045318	real	39.81000
10	310020305000006	3100203001001	23.688653	0.029354	real	39.81000
11	310020305000007	3100203001001	38.632240	0.047871	real	39.81000
12	310020305000008	3100203001001	29.887174	0.037035	real	39.81000
13	310020305000009	3100203001001	15.331356	0.018998	real	39.81000
14	310020305000010	3100203001001	26.455126	0.032782	real	39.81000
15	310020305000011	3100203001001	25.708526	0.031857	real	39.81000
16	310020305000012	3100203001001	16.425213	0.020353	real	39.81000
17	310020305000013	3100203001001	31.756569	0.039351	real	39.81000
18	310020305000014	3100203001001	36.322987	0.045010	real	39.81000
19	310020305000015	3100203001001	37.879852	0.046939	real	39.81000
20	310020305000016	3100203001001	36.606579	0.045361	real	39.81000

...	COD_AP	DEFICIT_TOTAL_FINAL	SOMA_SC	FONTE_DEFICIT	DIFERENCA	DIF_ABS
3118304003005	1,207.9800	1,207.9800	predito	4.547474e-13	4.547474e-13	
3106200005038	1,149.0000	1,149.0000	real	4.547474e-13	4.547474e-13	
3146107003002	1,324.0000	1,324.0000	real	-4.547474e-13	4.547474e-13	
3136702005008	1,445.0000	1,445.0000	real	-4.547474e-13	4.547474e-13	
3118601005010	1,970.0450	1,970.0450	predito	4.547474e-13	4.547474e-13	
3170206005003	1,634.0000	1,634.0000	real	-4.547474e-13	4.547474e-13	
3170206005011	1,837.0000	1,837.0000	real	-4.547474e-13	4.547474e-13	
3170206005013	2,213.0000	2,213.0000	real	4.547474e-13	4.547474e-13	
3106200005030	1,680.0000	1,680.0000	real	4.547474e-13	4.547474e-13	
3118601005014	1,245.3150	1,245.3150	predito	2.273737e-13	2.273737e-13	
3101706003002	1,116.0000	1,116.0000	real	-2.273737e-13	2.273737e-13	
3115300003003	704.9200	704.9200	predito	-2.273737e-13	2.273737e-13	
3131307005004	1,210.0000	1,210.0000	real	2.273737e-13	2.273737e-13	
3134202003005	859.0000	859.0000	real	2.273737e-13	2.273737e-13	
3136702005010	1,644.0000	1,644.0000	real	-2.273737e-13	2.273737e-13	
3170107005012	724.0000	724.0000	real	-2.273737e-13	2.273737e-13	
3125101003001	1,158.0000	1,158.0000	real	-2.273737e-13	2.273737e-13	
3126109003003	744.0000	744.0000	real	2.273737e-13	2.273737e-13	
3122306005005	1,239.0000	1,239.0000	real	2.273737e-13	2.273737e-13	
3170107005006	1,072.0000	1,072.0000	real	-2.273737e-13	2.273737e-13	
3126109003004	744.0000	744.0000	real	2.273737e-13	2.273737e-13	