

# Statistical Analysis

HOUSE PRICES

# Conhecendo e Preparando os Dados

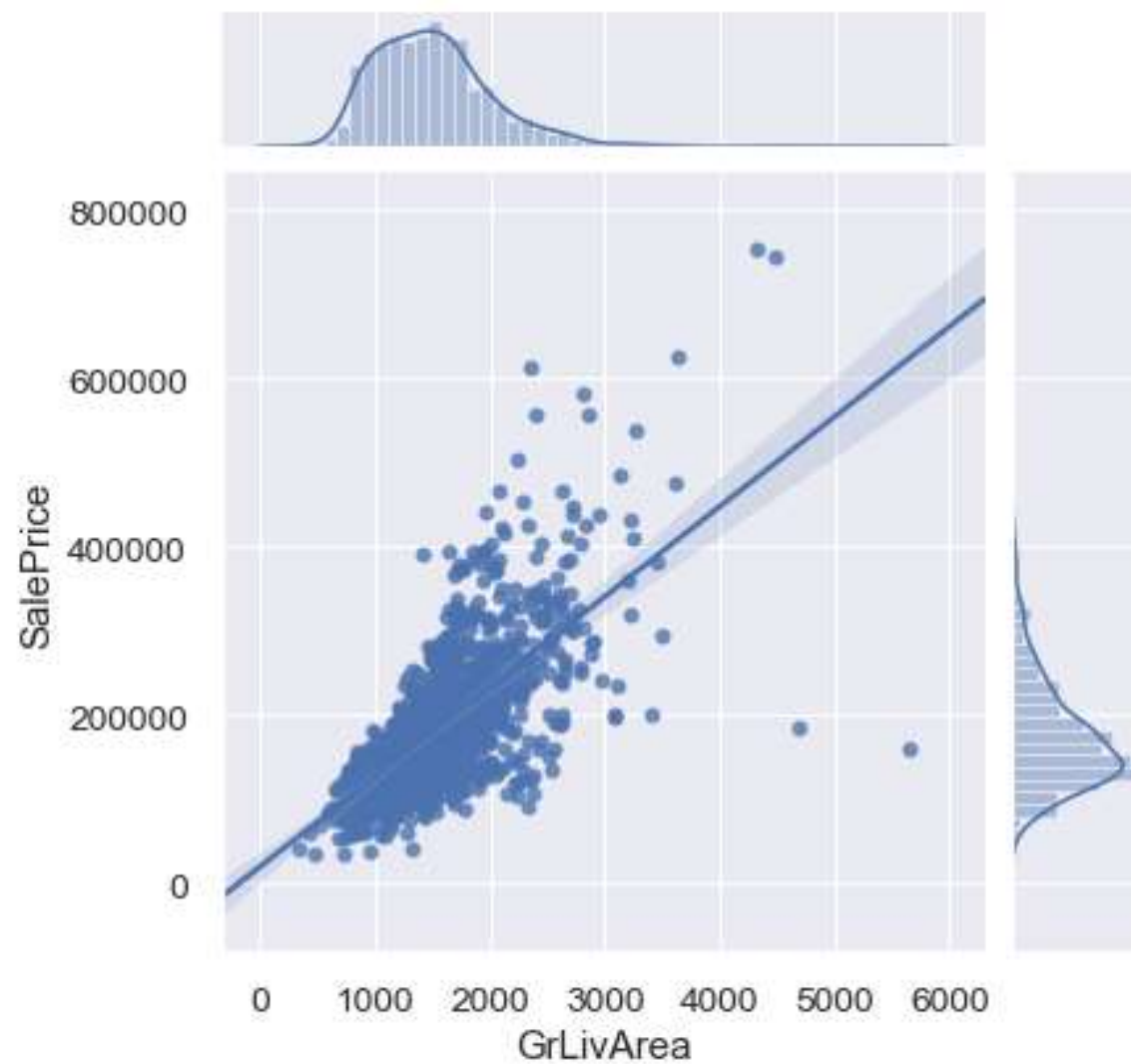
---

- 81 → 43 Qualitativas 41 Quantitativas
- Valores nulos
- Dados “duplicados”
- 70 → 41 Qualitativas 29 Quantitativas

Missing Ratio	
PoolQC	100.0
MiscFeature	96.0
Alley	94.0
Fence	81.0



Correlação com o Preço de Venda

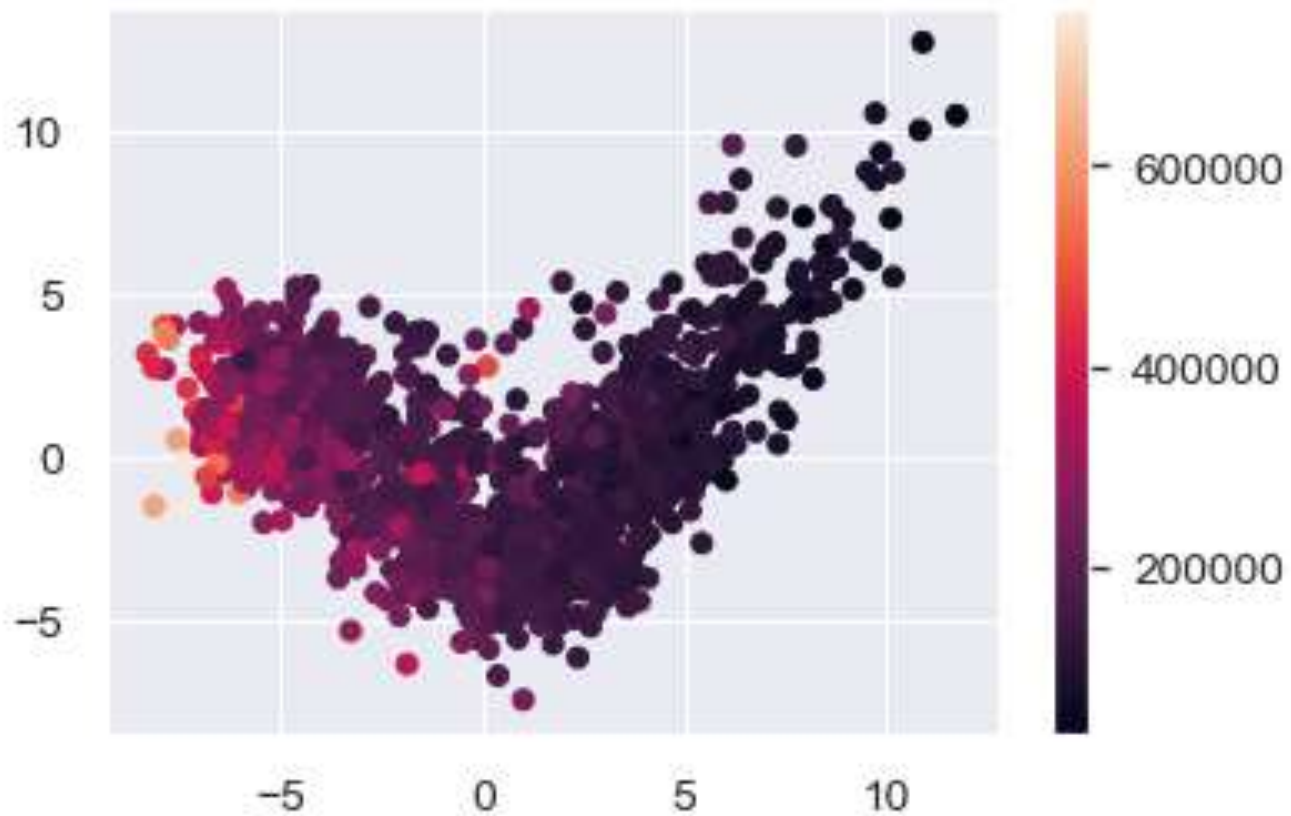


Sala de Estar  
(tamanho) x Preço

# Pré-regressão

---

- Dummies 1458 x 260
- PCA



# Aplicando o Modelo

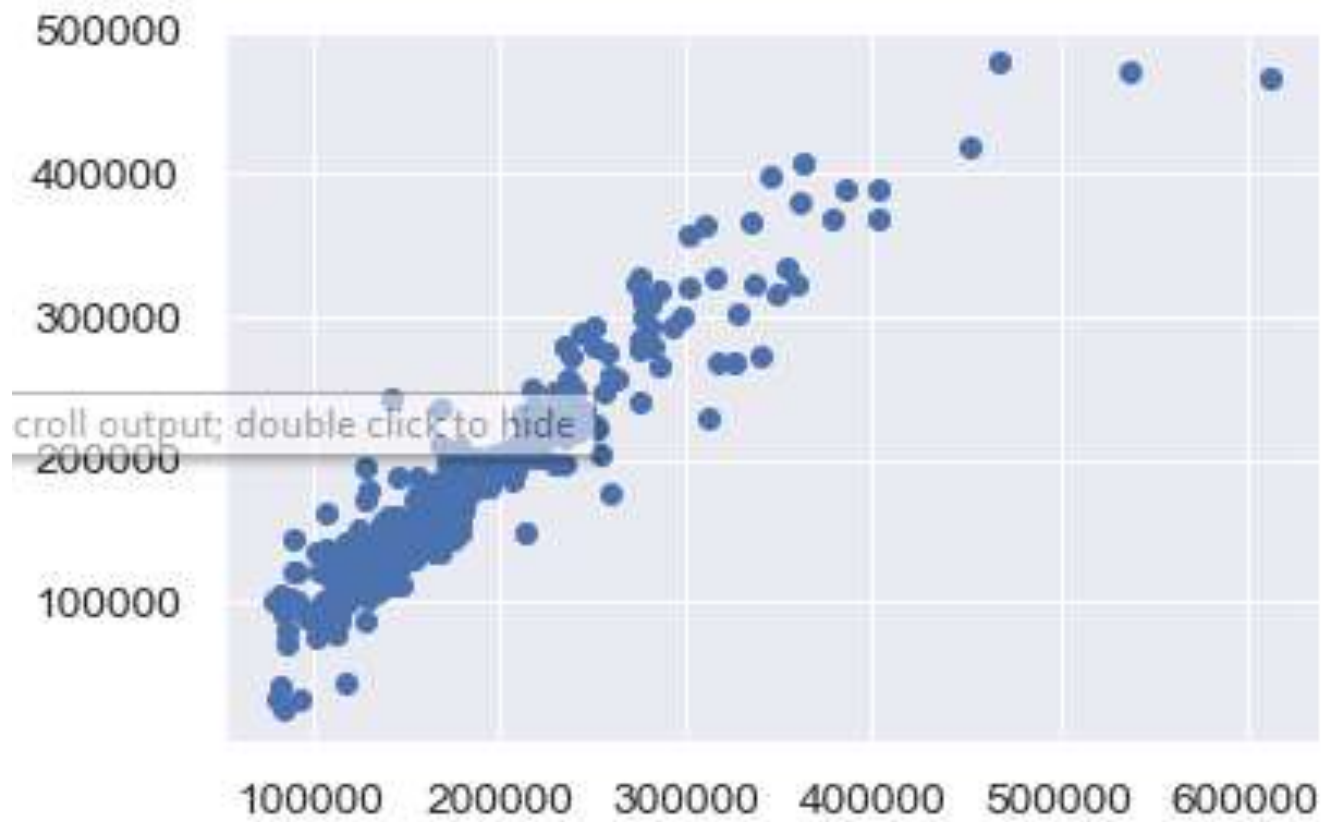
---

- Regressão Linear
- Set de Treinamento 40%
- R2 no Set de Treinamento: 0.95
- R2 no Set de Teste: 0.87
- RMSE (erro da predição): 27.732,02

# Coeficientes

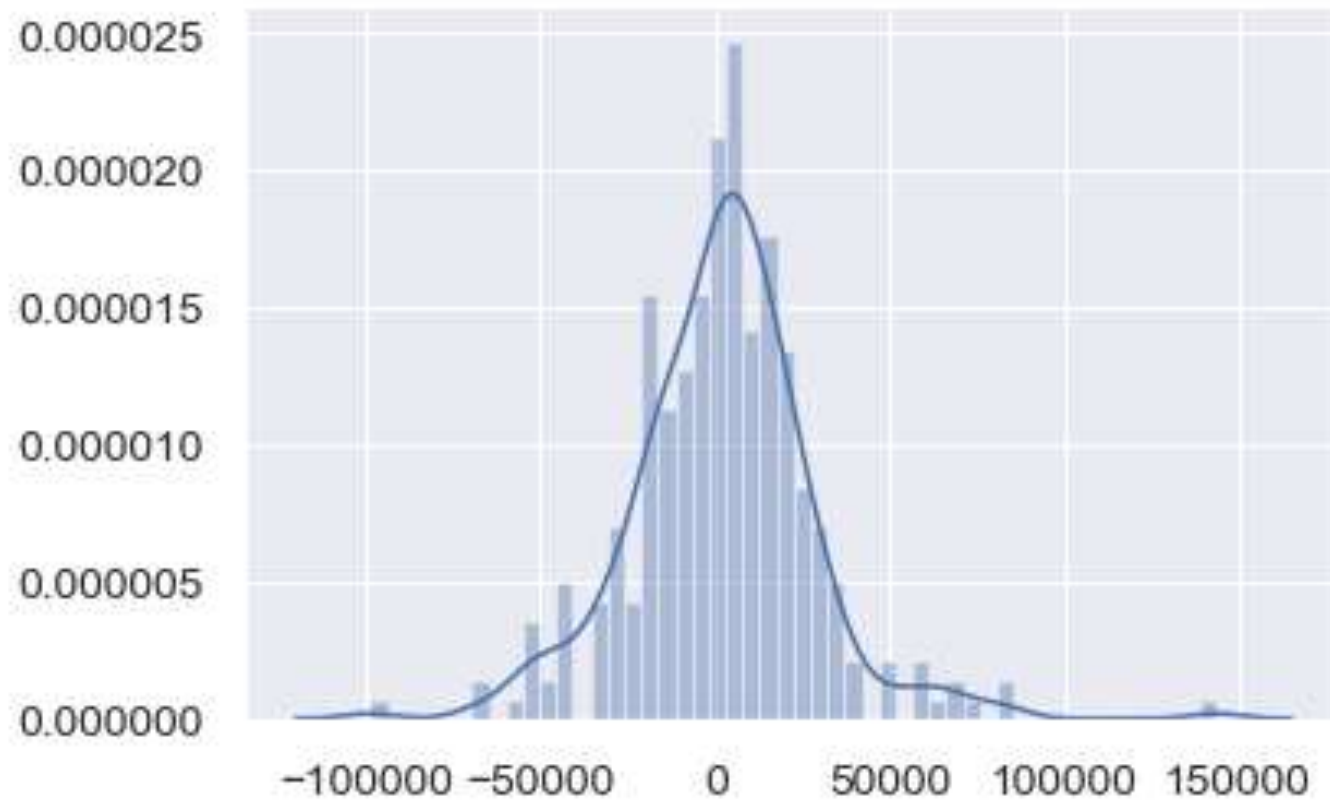
---

○ Telhado de madeira	RoofMatl_WdShngl	60641.870110
○ Telhado no estilo galpão (?)	RoofStyle_Shed	54096.959935
○ Telhado de membrana (?)	RoofMatl_Membran	49965.835622
○ Área (agrícola/comercial/residêncial)	MSZoning_FV	49811.681844
○ 2° piso inacabado	HouseStyle_1.5Unf	47431.741903
○ Condição de venda: parcial (imóvel novo)	SaleCondition_Partial	47390.503667
○ Garagem com padrão normal	GarageCond_TA	43007.927383
○ Revestimento externo em telhas de asfalto	GarageCond_Fa	42562.593818
○ Tipo de venda: Contrato de 15% para pagamento de prazos regulares	Exterior2nd_AsphShn	42456.236620
	SaleType_Con	42451.621747



Preço Real x  
Predições





Histograma Residual