

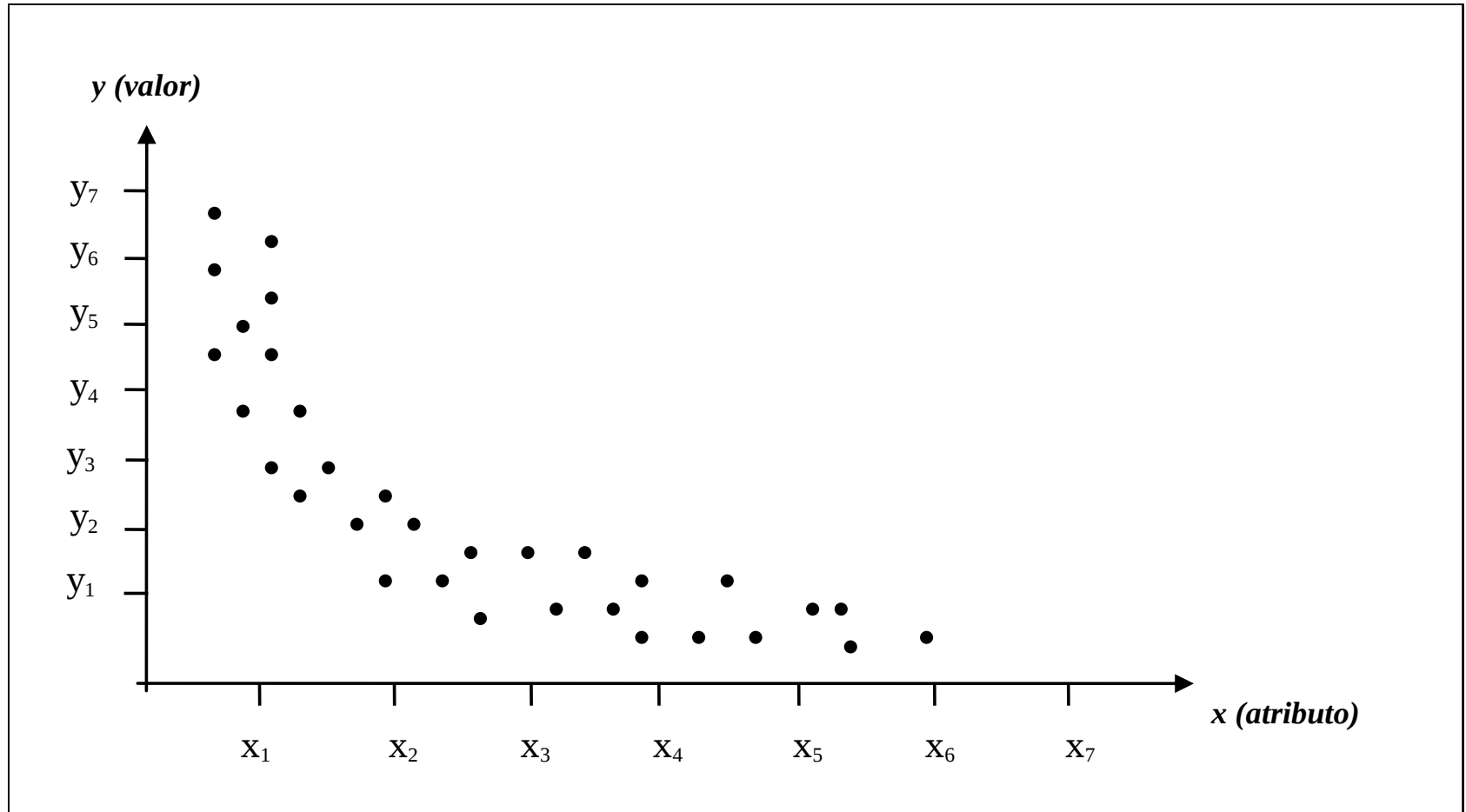
TRANSFORMAÇÃO POR ANÁLISE DE ERRO

Sérgio Alberto Pires da Silva

REGRESSÃO SIMPLES – 2 VARIÁVEIS

- Cálculo de parâmetro da equação por Mínimos Quadrados é aplicável somente para relações Lineares;
- No mercado imobiliário a maioria das relações entre variáveis não são Lineares;
- Transformação de escala de medida das variáveis possibilita a linearizar a relação;
- Na Regressão simples a análise gráfica da dispersão sinaliza a transformação adequada para esta linearização;

EXEMPLO FUNÇÃO HIPERBÓLICA



MUDANÇA DE ESCALAS DOS EIXOS

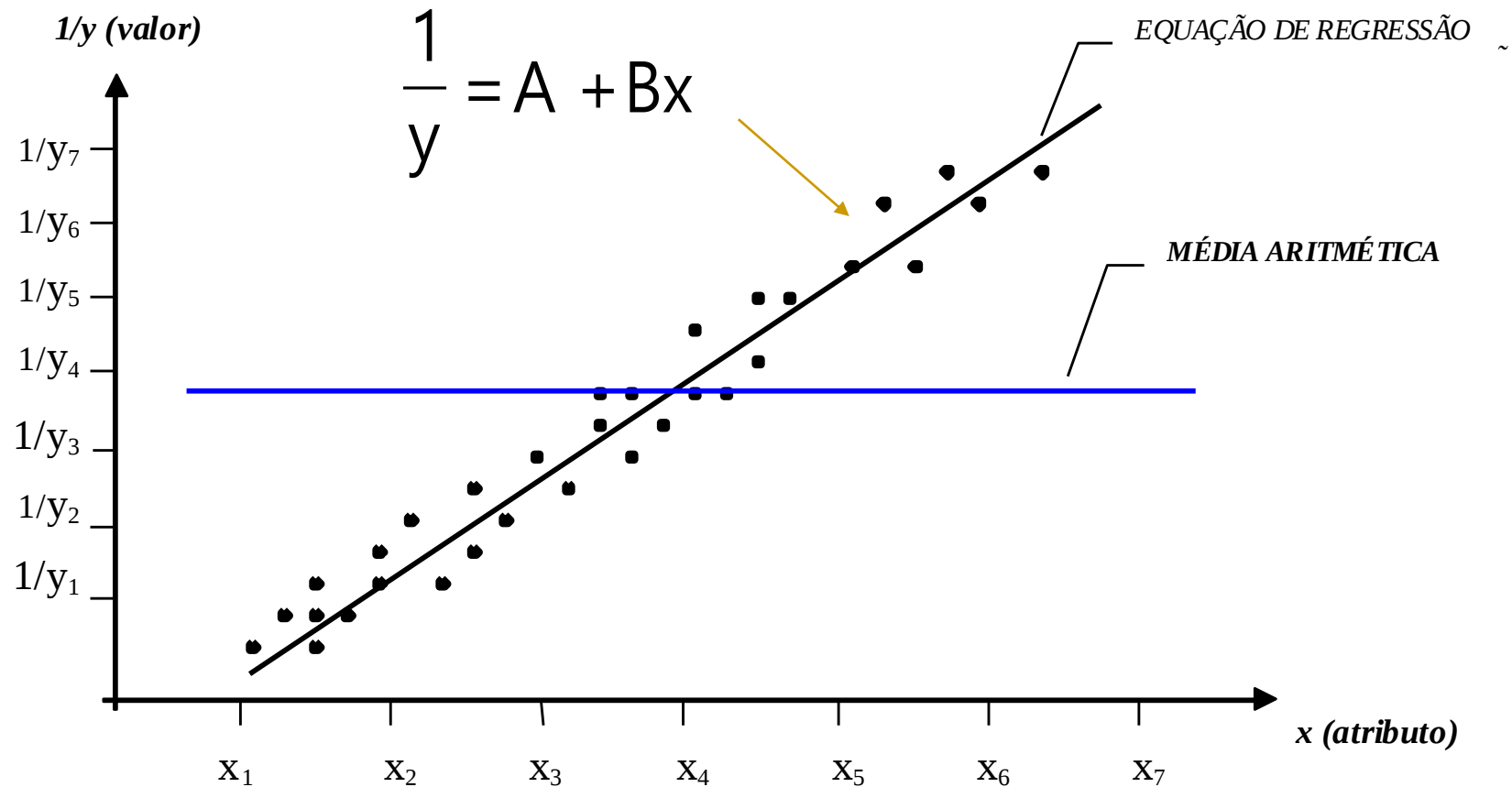
- Transformação adequada indica a Variável Dependente na forma inversa – Equação hiperbólica, ou seja, 1/y no eixo vertical, gerando uma função linear do tipo:

$$\frac{1}{y} = A + Bx$$

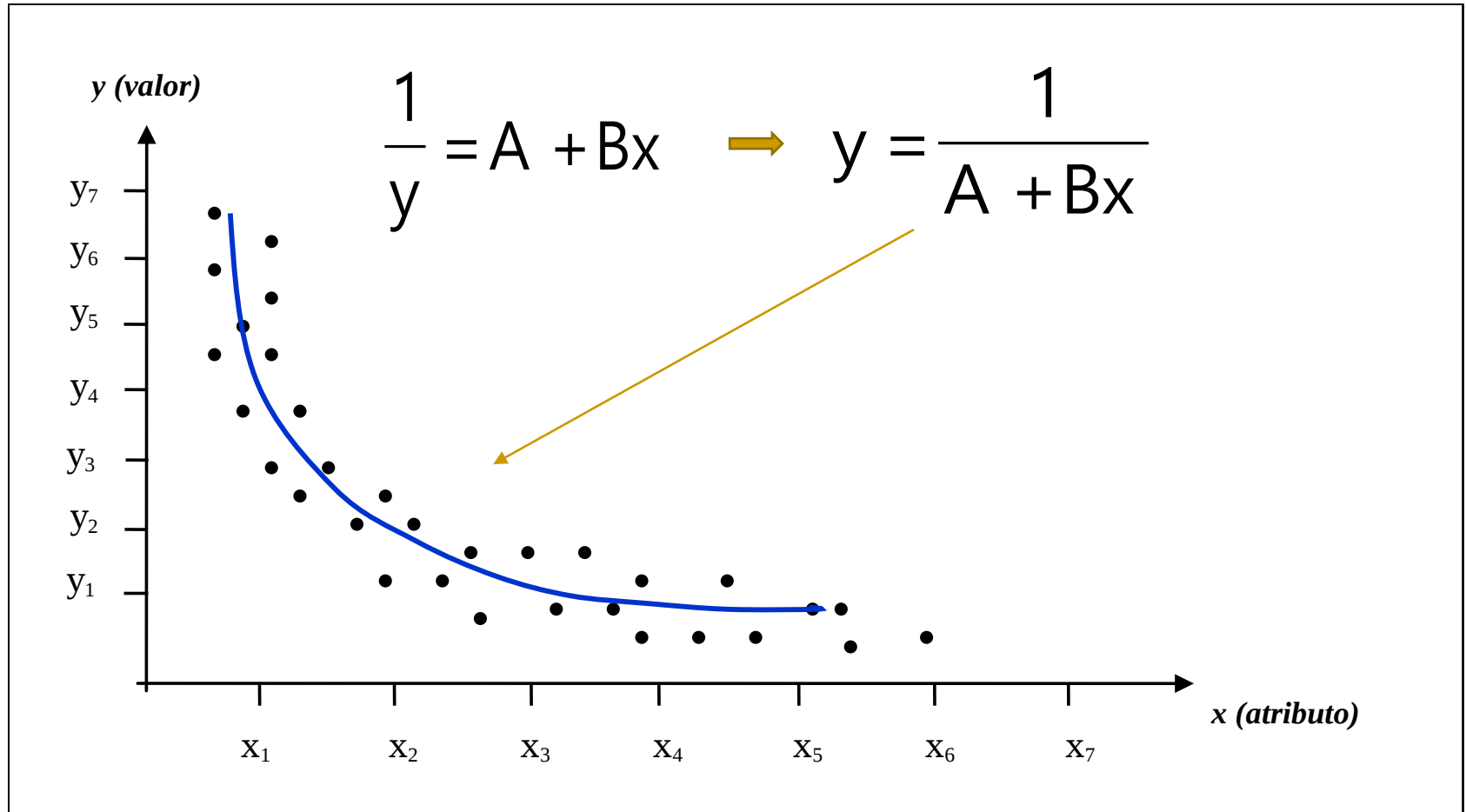
- Resultando Equação Não Linear tipo:

$$y = \frac{1}{A + Bx}$$

EXPECTATIVA DE LINEARIZAÇÃO



EXEMPLO FUNÇÃO HIPERBÓLICA



REGRESSÃO MÚLTIPLA – MUITAS VARIÁVEIS

- O uso de mais de uma variável independente dificulta a análise das relações individuais com a dependente;
- Para identificar a relação adequada entre cada variável com o valor dos imóveis (variável dependente) será necessária a homogeneização do modelo em relação às demais variáveis;

DIFICULDADE DE LINEARIZAÇÃO

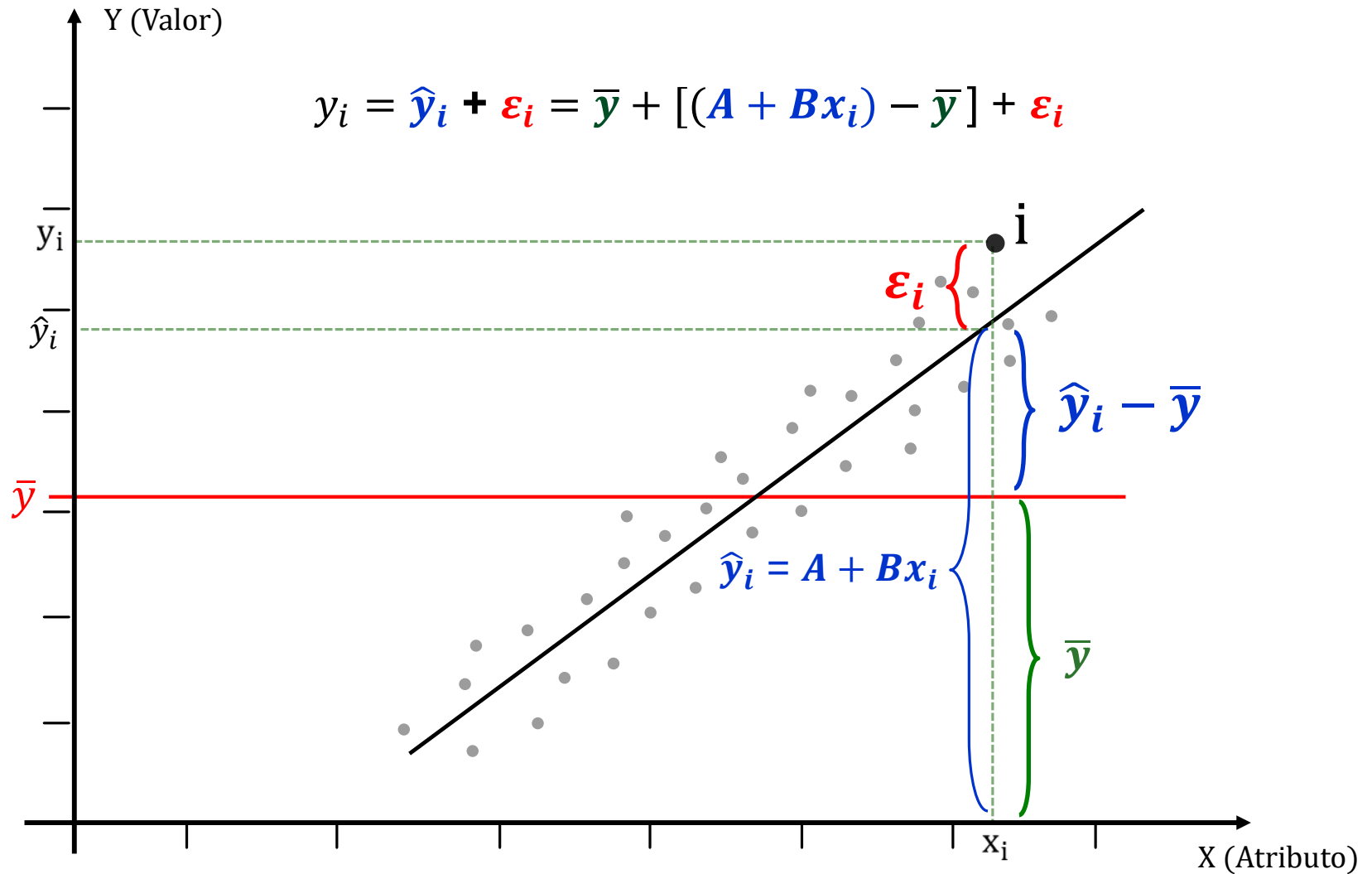
- A parcela da diferença entre os valores dos dados da amostra e sua média aritmética que tem origem no conjunto de características heterogêneas é formada de um bloco único, não sendo possível identificar qual a parcela relativa à cada uma das características isoladamente;
 - Análise Gráfica da Variável Dependente com cada Atributo é inócua e não sinaliza a Transformação de escala para Linearização;
-

DECOMPOSIÇÃO DO VALOR OBSERVADO

- Em amostra homogênea o Valor de cada elemento é composto pela média aritmética e o erro aleatório: $y_i = \bar{y} + \varepsilon_i$
- A Regressão relação Linear Simples (entre duas variáveis) implica na decomposição do Valor dos elementos em 3 parcelas:
 - Média da Amostra;
 - Influência do atributo;
 - Erro aleatório

$$y_i = \bar{y} + (A + Bx_i - \bar{y}) + \varepsilon_i$$

REPRESENTAÇÃO GRÁFICA

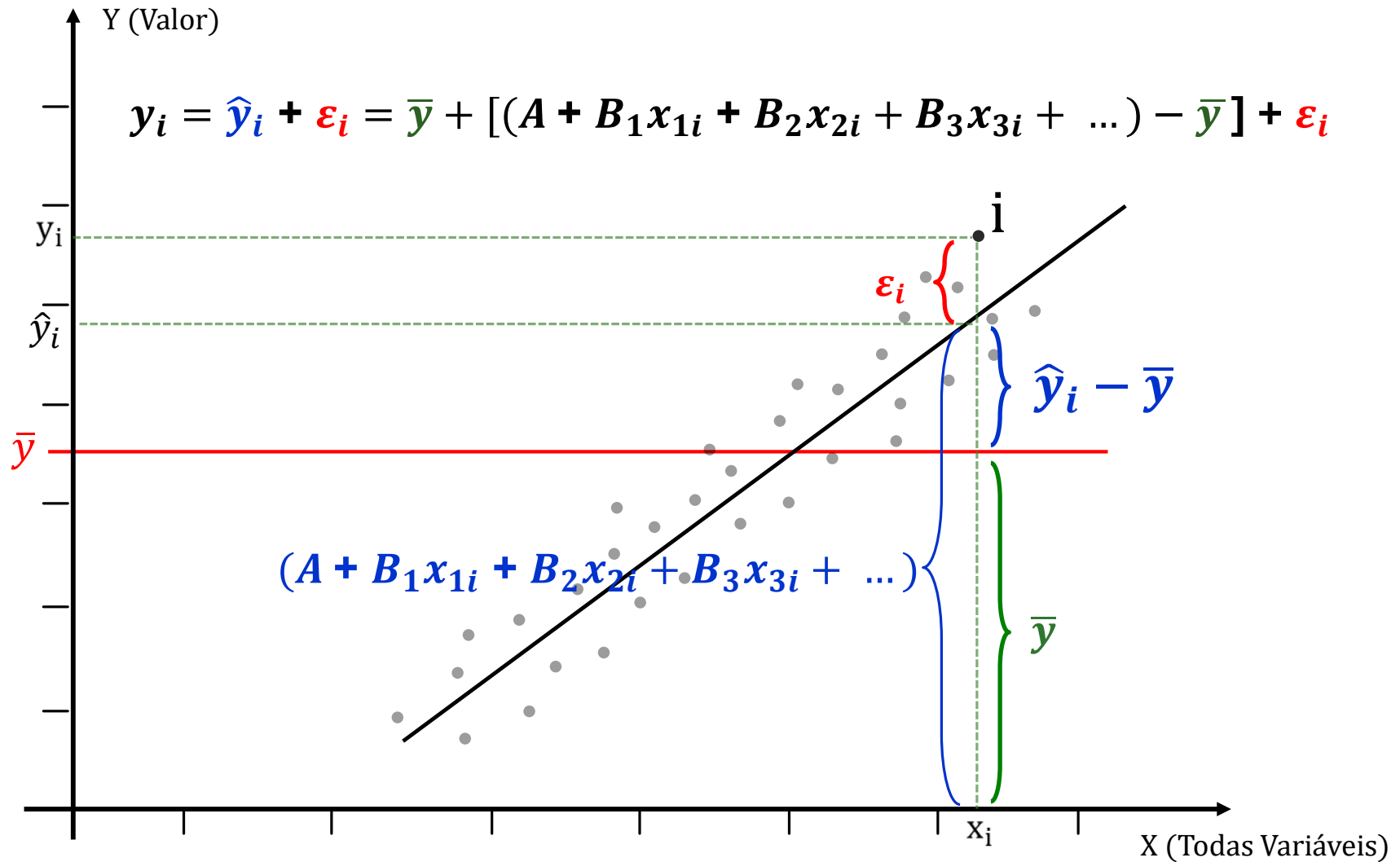


REGRESSÃO COM MAIS VARIÁVEIS

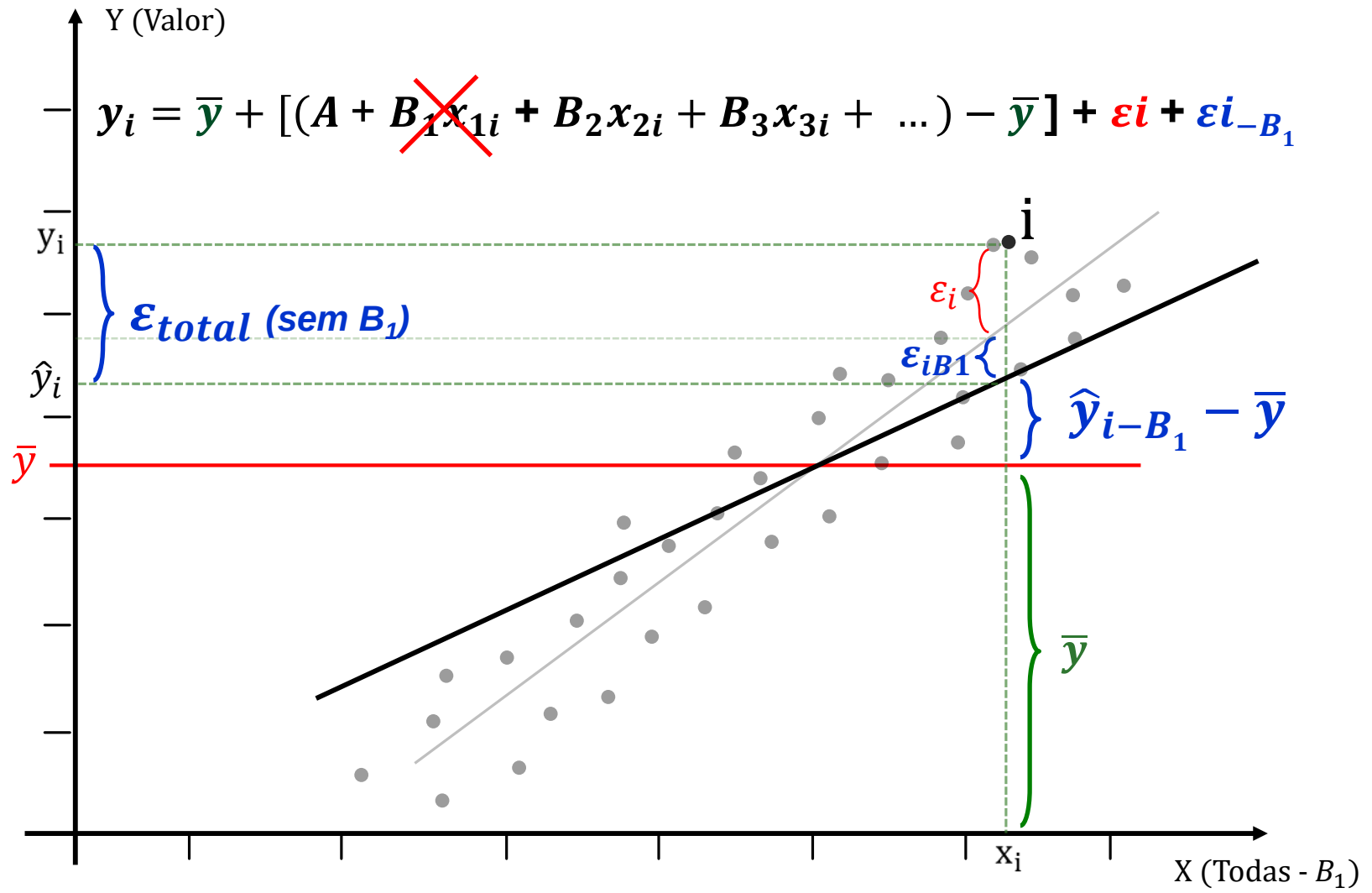
- Na Regressão Múltipla (mais variáveis) o Valor de cada elemento da amostra também estará decomposto em 3 parcelas:
 - Média da Amostra;
 - Influência do conjunto de variáveis utilizadas;
 - Erro aleatório

$$y_i = \bar{y} + (A + B_1x_{1i} + B_2x_{2i} + \dots + B_kx_{ki} - \bar{y}) + \varepsilon_i$$

REGRESSÃO MÚLTIPLA



REGRESSÃO MÚLTIPLA - VARIÁVEL



HOMOGENEIZAÇÃO PARCIAL

- Para homogeneizar os dados em relação a todas as demais variáveis exceto a variável representada por B_1 , retiramos a influência das demais variáveis (diferença entre os eixos da média e equação de regressão), da seguinte forma:

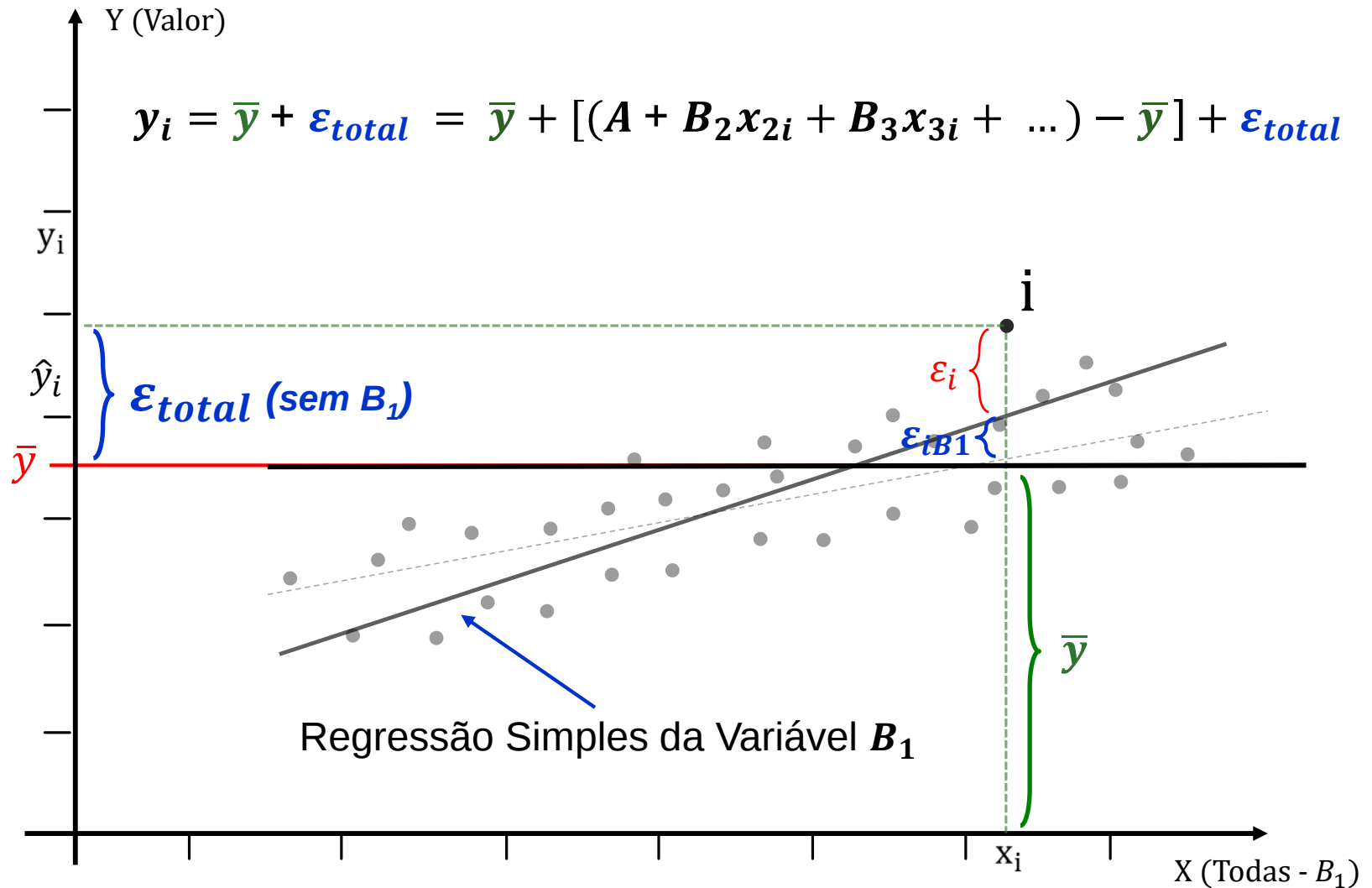
$$y_i \text{ (homogeneizado)} = \bar{y} + \epsilon_{iB_1} + \epsilon_i$$

$$\text{ou seja } \rightarrow y_{ih} = \bar{y} + \epsilon_{total}$$

Modelo sem variável B_1



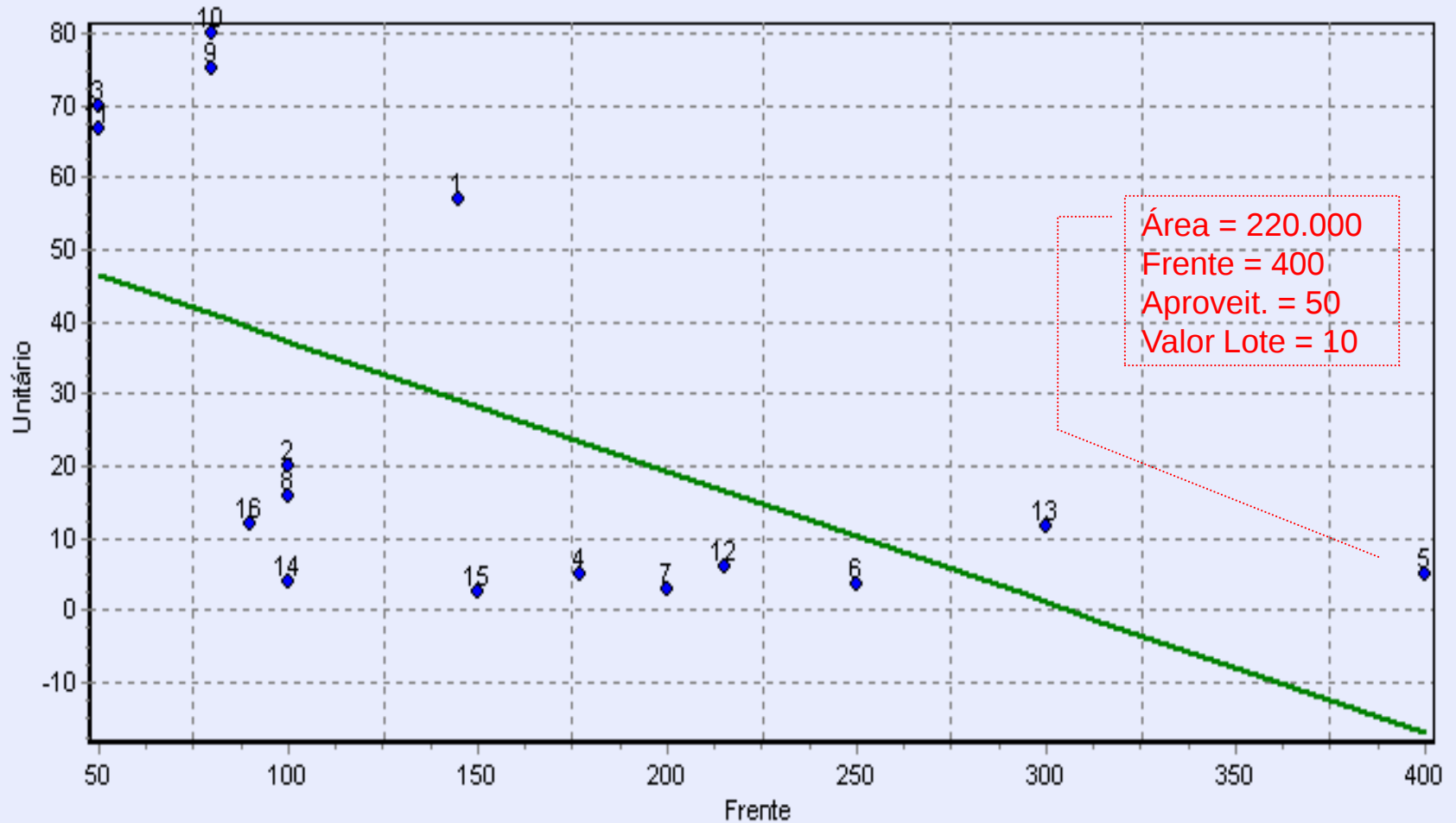
GRÁFICO DE HOMOGENEIZAÇÃO



EXEMLO NUMÉRICO

Dado	Área Total (m²)	Frente (m)	Aproveitamento (%)	Valor de Lote (R\$/m²)	Unitário (R\$/m²)
1	14.000,00	145,00	55	90	57,14
2	25.000,00	100,00	55	40	20,00
3	15.000,00	50,00	65	90	70,00
4	40.000,00	177,00	45	20	5,00
5	220.000,00	400,00	50	10	5,00
6	240.000,00	250,00	45	10	3,62
7	420.000,00	200,00	40	20	3,09
8	25.000,00	100,00	55	60	16,00
9	22.000,00	80,00	60	120	75,00
10	11.000,00	80,00	75	90	80,00
11	19.000,00	50,00	60	100	66,66
12	180.000,00	215,00	55	30	6,11
13	26.000,00	300,00	55	25	11,53
14	200.000,00	100,00	60	25	4,00
15	350.000,00	150,00	40	15	2,50
16	75.000,00	90,00	60	50	12,00

ANÁLISE GRÁFICA - FRENTE



CONCLUSÕES EQUIVOCADAS

- Relação entre a característica (frente) e o Valor Unitário dos imóveis é negativa (quanto maior a frente, menor o valor dos imóveis);
- Dados que prejudicam a hipótese não estão claramente expostos;
- O dado número 5, por exemplo, apresenta uma frente de 400,00 m e Valor Unitário de R\$ 5,00/m², este valor reduzido não está originado na frente de grandes dimensões, mas na área de grande porte, baixa atratividade da região, etc.

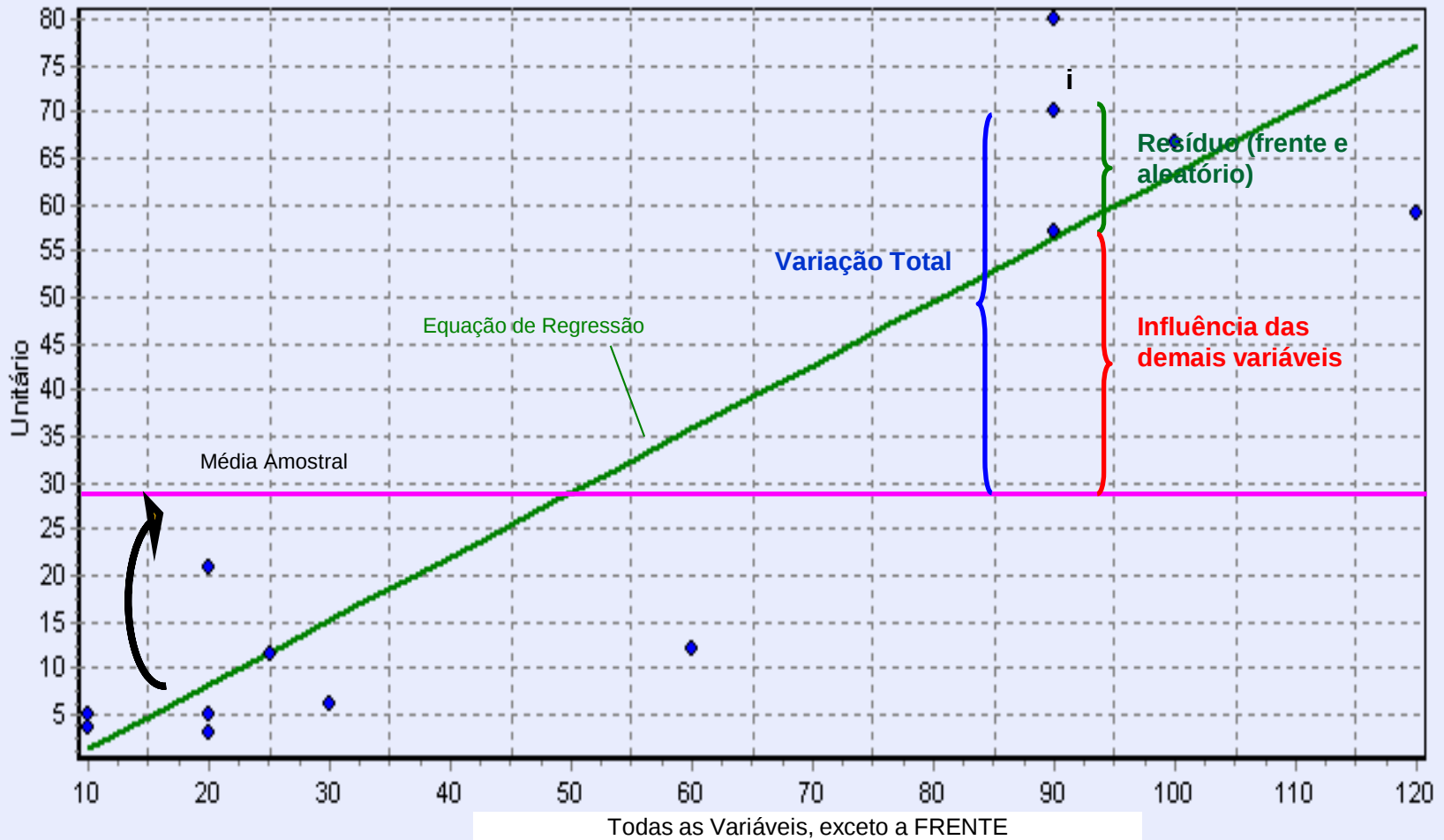
ISOLAR INFLUÊNCIA DE VARIÁVEL

- Elaboração de Equação de Regressão sem a presença da Característica que se pretende identificar a relação de influência no valor;
- Homogeneizar cada elemento da amostra pela equação elaborada;
- Análise da dispersão gráfica da característica e a variável dependente homogeneizada:
 - identificar a relação matemática de influência da característica no valor dos imóveis;
 - visualizar os dados inconsistentes para a hipótese formulada (variável);

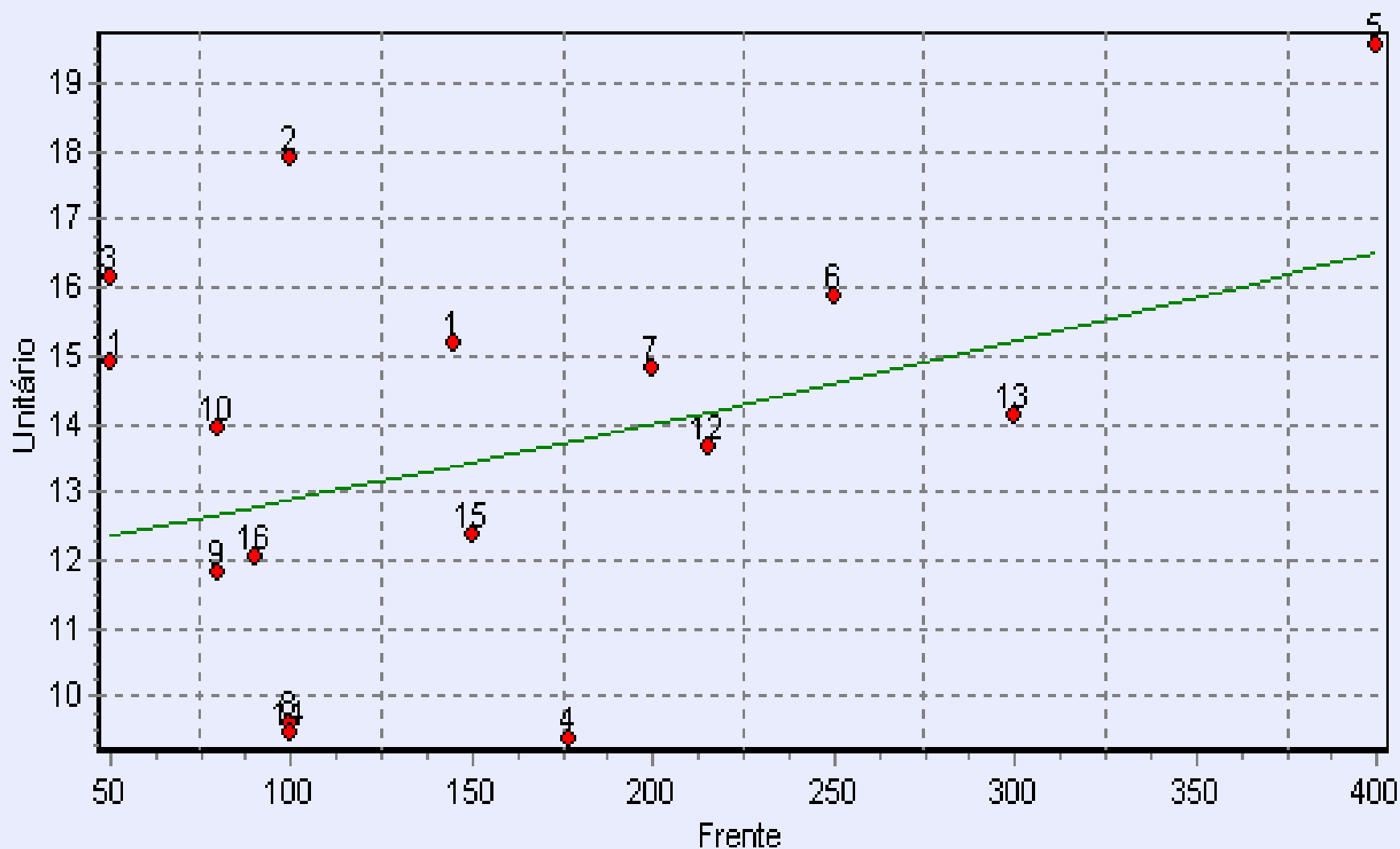
HOMOGENEIZAÇÃO DO VALOR

- Retirando-se do valor de cada dado da amostra a parcela referente à influência das demais variáveis, equivalente a:
 - $(A + B_1 \cdot \text{Área}_i + B_2 \cdot \text{Aproveitamento}_i + B_3 \cdot \text{Valor de Lote}_i - \bar{y})$
 - tem-se: $\text{ValorUnitário}_i = \bar{y} + \varepsilon_i - \text{frente}$
 - Neste caso ε_i contém o erro aleatório mais a influência da *Frente* no valor;
 - Equivale a um giro na equação de regressão:

REPRESENTAÇÃO ESQUEMÁTICA



VALORES AJUSTADOS E FRENTE



ANÁLISE DA DISPERSÃO

- Observa-se agora que a relação apropriada da variável Frente e o Valor Unitário dos imóveis é direta ($y = A + Bx$).
- Sem a influência das demais características verifica-se que os dados mostram de forma mais consistente a influência da variável FRENTE no valor das glebas;
- A análise é semelhante à Regressão Simples onde os dados inconsistentes com a hipótese podem ser identificados com mais clareza;

PRICIPAIS VANTAGENS

- Identificação clara dos dados inconsistentes para cada hipótese;
 - Possibilidade de criação de variável de Localização baseada em média homogênea de vizinhos, evitando correlação espacial;
 - Possibilidade de ajustes de valor estimado em Modelos Colineares através de Regressão Simples de variável colinear retirada e dados homogêneos.
-

CONCLUSÕES E RECOMENDAÇÕES

- O processo representa importante ferramenta na comprovação das hipóteses e identificação de dados inconsistentes;
- Ao se retirar do modelo geral a variável a ser estudada para construção do modelo parcial, não se considera a influência conjunta desta variável com cada uma das demais variáveis do modelo.
- Quanto maiores as relações colineares da variável estudada com as demais variáveis independentes do modelo geral, maior a dificuldade na análise.