

UNIVERSIDADE FEDERAL
DA PARAÍBA

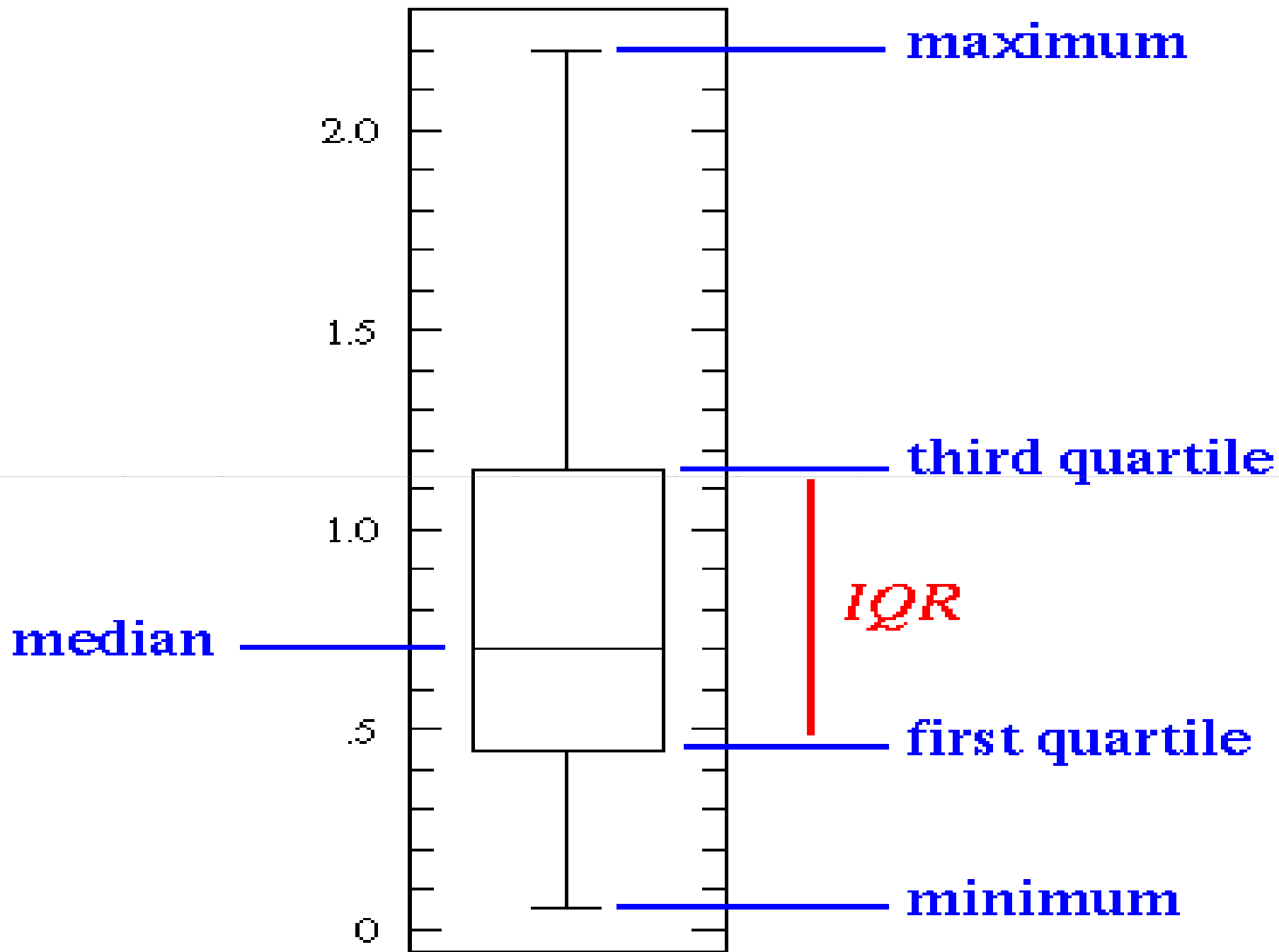
BOXPLOT

Departamento de Estatística
Luiz Medeiros

<http://www.de.ufpb.br/~luiz/>

BOXPLOT

- ▶ Este tipo de gráfico estatístico representa os dados através de um retângulo construído com os quartis.
- ▶ O objetivo deste gráfico é fornecer informações sobre a variabilidade dos dados e valores atípicos que podem influenciar o cálculo de medidas como a média aritmética, por exemplo.
- ▶ O boxplot utiliza cinco medidas estatísticas: mínimo, máximo, mediana, primeiro quartil, terceiro quartil.
- ▶ O conjunto destas medidas fornece evidência acerca da posição, dispersão, assimetria e valores extremos (atípicos).



BOXPLOT

- ▶ A posição central é dada pela mediana e a dispersão pelo chamado desvio interquartílico, denotado por $dq = Q_3 - Q_1$.
- ▶ As posições dos quartis Q_1 , Q_2 e Q_3 fornecem evidência sobre o nível de assimetria da distribuição dos dados.
- ▶ Os comprimentos das caudas da distribuição são dados pelas linhas que vão do retângulo aos valores atípicos. Estes valores atípicos são chamados de outliers.
- ▶ De modo geral, um ponto será considerado outlier quando estiver fora do intervalo denotado por $(LI ; LS)$, onde

$$LI = Q_1 - 1,5 (Q_3 - Q_1) \quad \text{e} \quad LS = Q_3 + 1,5 (Q_3 - Q_1)$$

CONSTRUÇÃO DO BOXPLOT

- 1) Determinação das cinco medidas: mínimo, máximo, mediana, primeiro quartil, terceiro quartil;
- 2) Construção do rol;
- 3) Construir um eixo de escalas que contenha o valor mínimo e máximo;
- 4) Construir uma caixa (retangular), estendendo-se de Q_1 até Q_3 e construir uma linha na caixa, linha esta correspondente à localização do valor da mediana;
- 5) Traçar uma reta paralela ao eixo do passo 2, e que passe por dentro da caixa, com extremidades correspondentes aos valores LI e LS.
- 6) Identifique os pontos discrepantes por um símbolo (*, por exemplo).

Exemplo:

Considere que 22 pessoas foram entrevistadas e forneceram a informação de idade, conforme dados informados a seguir.

18	18	19	20	20	20	20	20	20	21	21
22	23	24	25	25	25	26	29	30	35	37

Seguiremos os passos descritos anteriormente para construir o BoxPlot.

Exemplo:

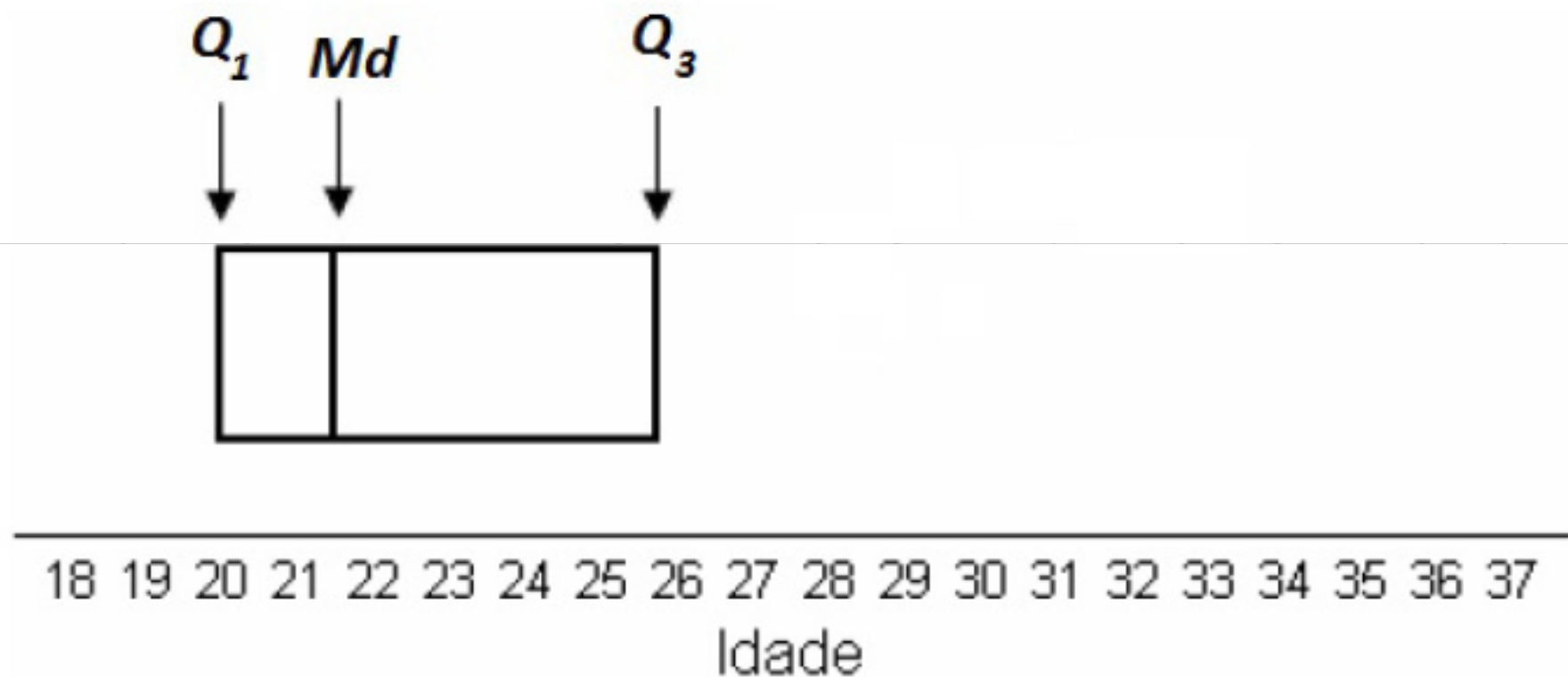
1. Temos que $Md = 21,50$, $Q_1 = 20$, $Q_3 = 25,75$, $dq = 5,75$

2. $LI = Q_1 - 1,5$, $dq = 11,375$ e $LS = Q_3 + 1,5dq = 34,375$

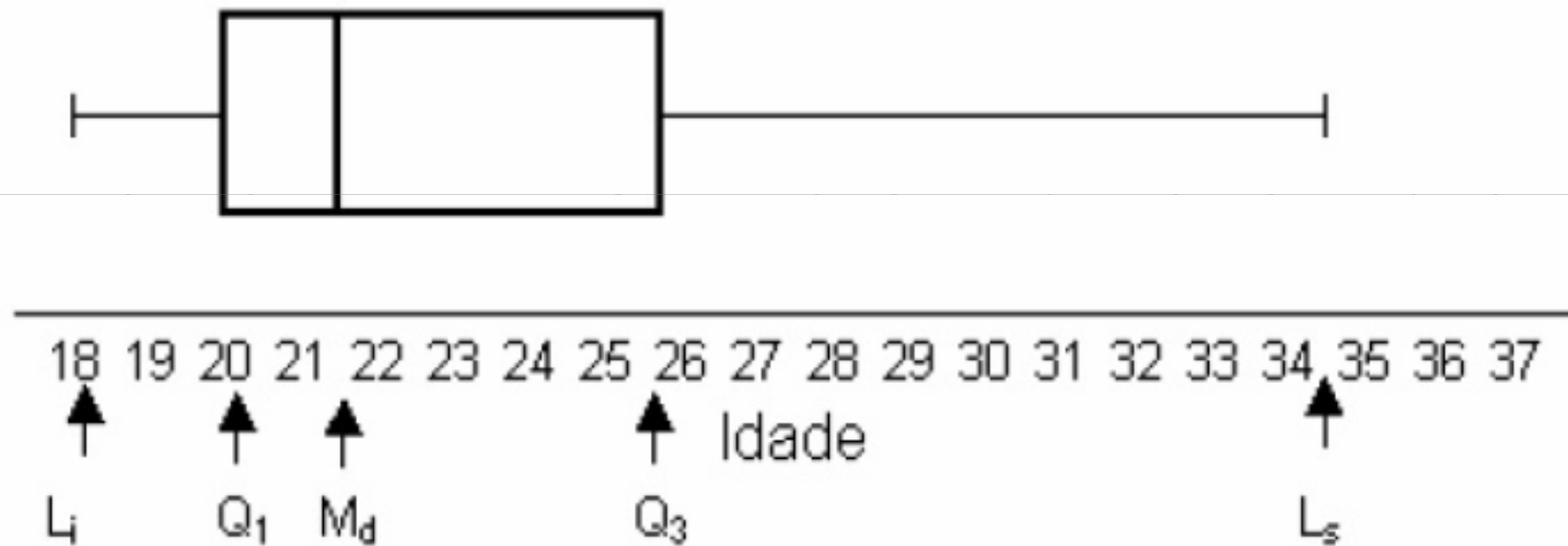
Exemplo:

18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
Idade																			

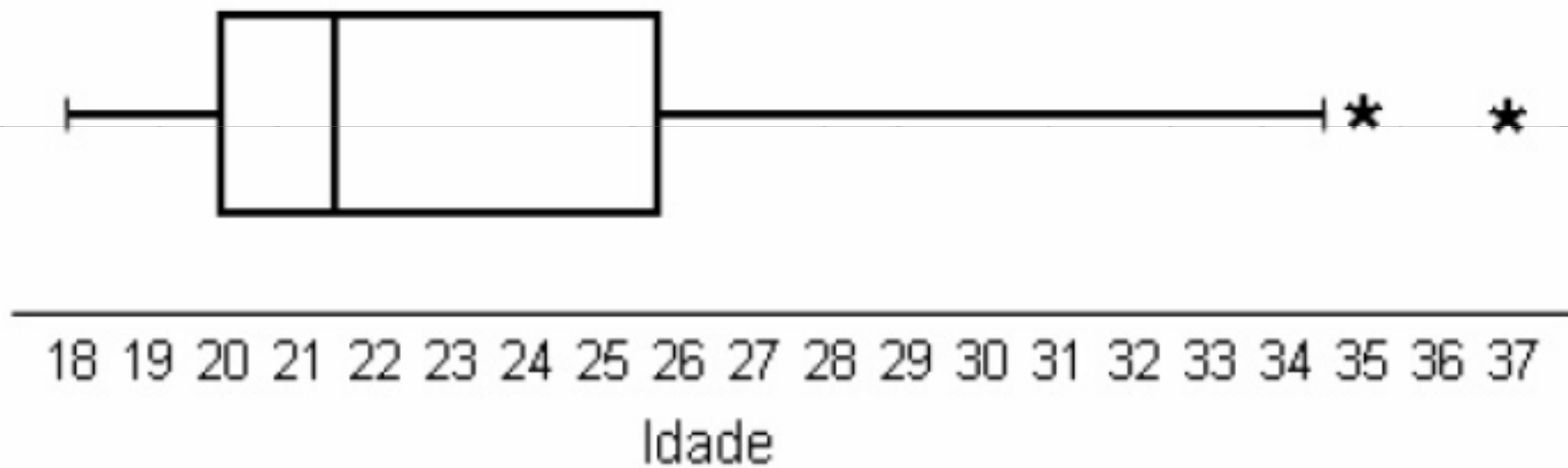
Exemplo:



Exemplo:



Exemplo:



Exemplo 2: Os dados abaixo representam o Preço (R\$) do produto A vendido em 25 diferentes estabelecimentos. A partir da distribuição de frequência que iremos construir podemos observar diversas características deste conjunto de dados. Construa um Boxplot do preço do produto A.

Dados Brutos: 20,5 - 19,5 - 15,6 - 24,1 - 9,9 - 15,4 - 12,7 - 5,4 - 17,0 - 28,6 - 16,9 - 7,8 - 23,3 - 11,8 - 18,4 - 13,4 - 14,3 - 19,2 - 9,2 - 16,8 - 8,8 - 22,1 - 20,8 - 12,6 - 15,9

Rol: 5,4 - 7,8 - 8,8 - 9,2 - 9,4 - 9,9 - 11,8 - 12,6 - 12,7 - 13,4 - 14,3 - 15,4 - 15,6 - 15,9 - 16,8 - 16,9 - 17,0 - 18,4 - 19,2 - 19,5 - 20,5 - 20,8 - 22,1 - 24,1 - 28,6