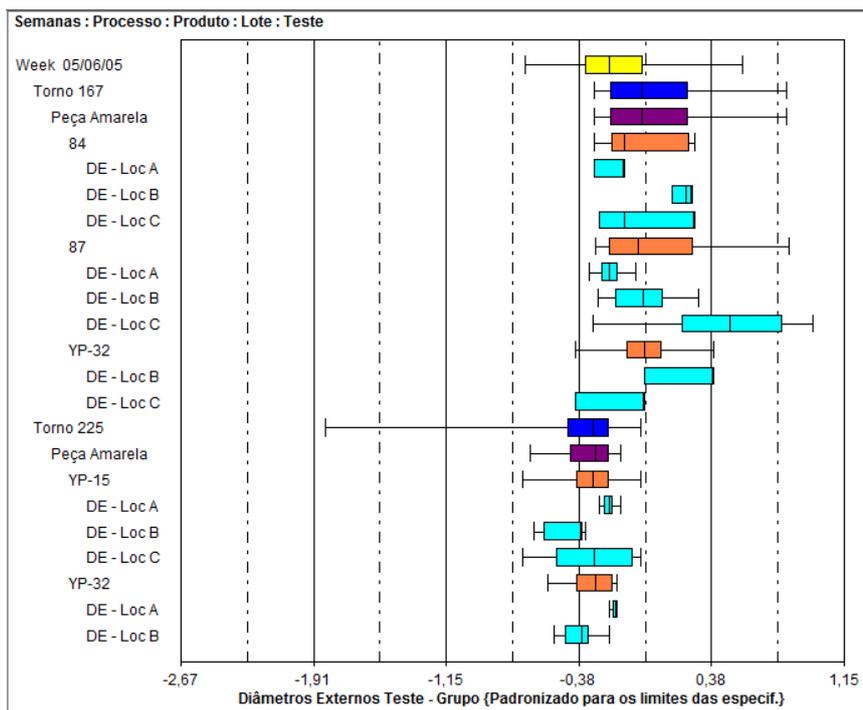


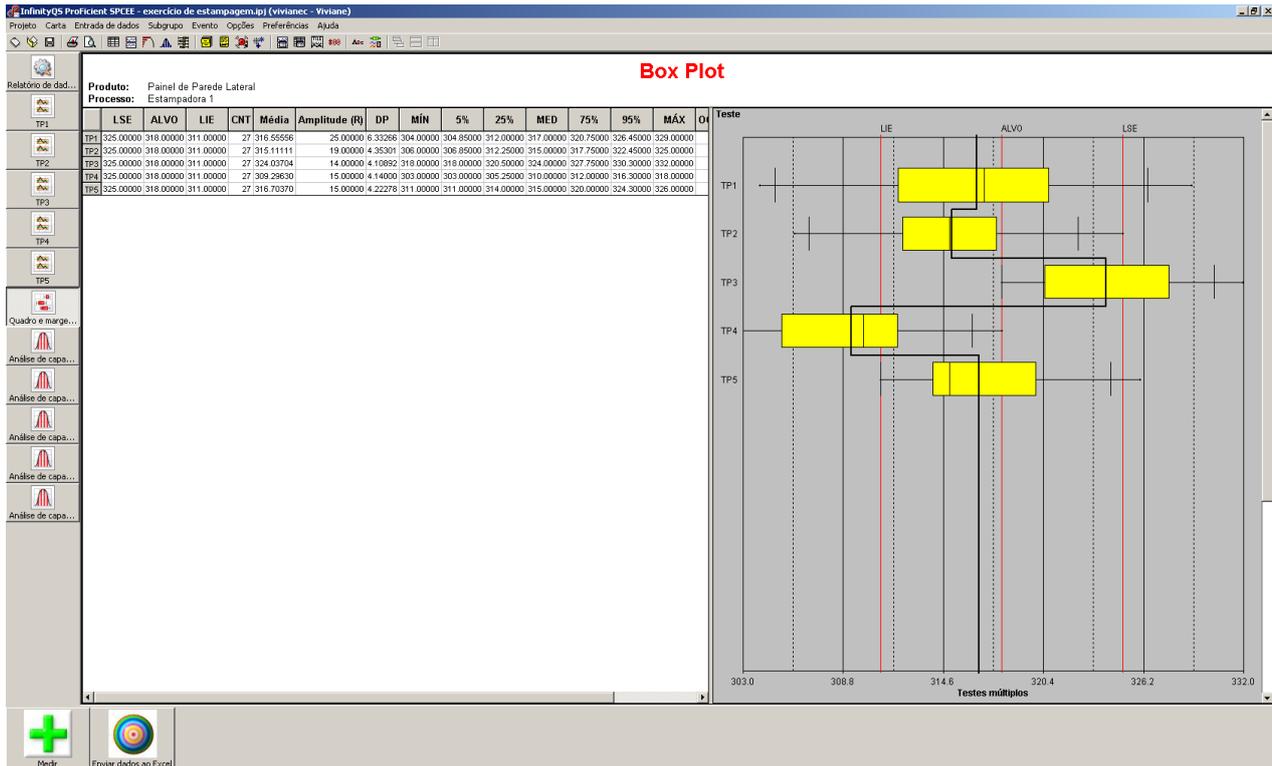
## Gráfico Box Plot

O *Box Plot* ou Quadro e Margem é um gráfico utilizado para auxiliar na identificação de valores discrepantes de um conjunto de dados (chamados de *outliers*), assim como comparar a dispersão entre dois ou mais conjuntos de dados (análise da simetria). Além disso, pode ser analisada a dispersão dos conjuntos de dados em torno dos limites de especificação e do valor alvo.



Este tipo de gráfico baseia-se na localização dos quartis (1º quartil, 2º quartil ou mediana, 3º quartil), que são três valores que dividem um conjunto de dados (ordenado de forma crescente) em quatro sub-intervalos, cada um com 25% das observações.

Abaixo segue um exemplo deste tipo de gráfico:



No exemplo acima, foram realizados cinco testes, os quais possuem as seguintes denominações: TP1, TP2, TP3, TP4 e TP5.

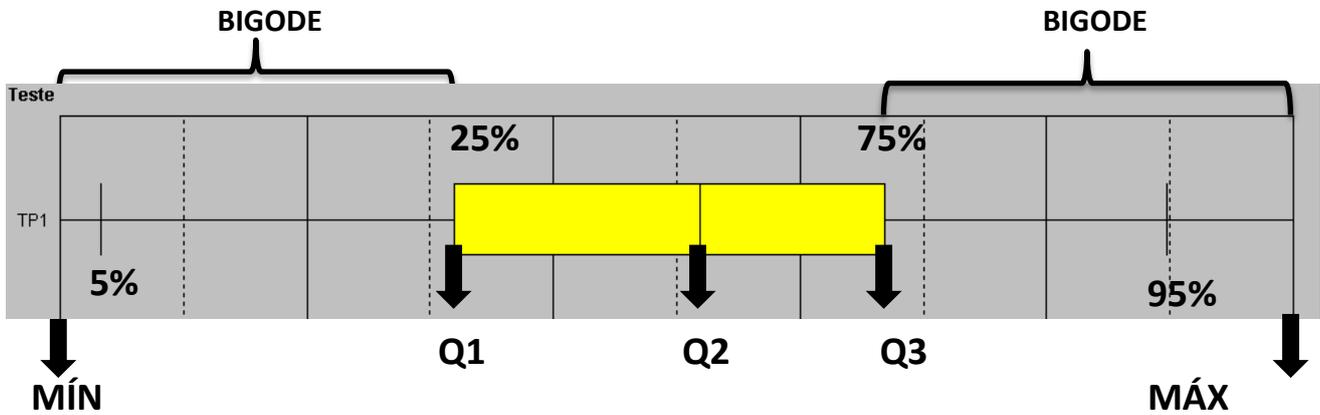
**Observação:** A linha preta que passa por todas as caixas amarelas representa a média dos valores de cada teste, o que facilita a comparação dos testes em torno do alvo.

Para cada teste realizado, o Box plot é construído da seguinte forma:

- **1º intervalo quartílico:** é delimitado por  $x_{\min}$  e  $q_1$ . A partir do menor ponto do conjunto de dados até o segmento vertical, encontram-se 5% dos valores do conjunto de dados; do segmento vertical até o 1º segmento do quadro amarelo, encontram-se 20% dos valores do conjunto de dados; totalizando 25% dos valores. Este intervalo é representado pelo 1º “bigode” do Box Plot.
- **2º intervalo quartílico:** é delimitado pelo  $q_1$  e  $q_2$  (mediana). Neste intervalo encontram-se mais 25% dos valores do conjunto de dados e é representado pela 1ª parte da caixa amarela.
- **3º intervalo quartílico:** é delimitado pelo  $q_2$  (mediana) e  $q_3$ . Neste intervalo encontram-se mais 25% dos valores do conjunto de dados e é representado pela 2ª parte da caixa amarela.
- **4º intervalo quartílico:** é delimitado pelo  $q_3$  e  $x_{\max}$ . A partir do final do quadro amarelo até o segmento vertical, encontram-se 20% dos valores do conjunto de dados; do segmento vertical ao maior ponto do conjunto de dados, encontram-se os 5% restantes dos valores do

conjunto de dados, totalizando os últimos 25% dos valores. Este intervalo é representado pelo 2º “bigode” do Box plot.

Todos os intervalos quartílicos estão representados na imagem abaixo:



Abaixo estão detalhadas todas as informações que podem estar contidas no relatório de dados do box plot, de acordo com o produto, processo e teste:

LSE	ALVO	LIE	CNT	Média	Amplitude (R)	DP	Pp	Ppk	Ppl	Ppu	Pr	MIN	5%	25%	MED	75%	95%	MÁX	FDE	%FDE	SIGL
1.000	----	-1.000	108	0.078	4.140	0.610	1.071	0.746	1.320	0.822	0.934	-2.280	-1.004	-0.150	0.130	0.475	0.784	1.860	8	7.407%	2.237
1.000	0.000	-1.000	36	0.476	0.900	0.184	1.810	0.949	2.670	0.949	0.553	0.100	0.100	0.350	0.500	0.555	0.847	1.000	0	0.000%	2.847
1.000	0.000	-1.000	36	-0.032	1.440	0.329	1.014	0.981	0.981	1.047	0.986	-0.960	-0.875	-0.175	0.000	0.140	0.463	0.480	0	0.000%	2.944
1.000	0.000	-1.000	36	-0.209	4.140	0.859	0.388	0.307	0.307	0.469	2.576	-2.280	-2.212	-0.495	-0.070	0.275	1.486	1.860	8	22.222%	0.921
1.000	-0.010	-1.000	202	0.049	1.560	0.343	2.176	1.801	2.010	2.341	0.460	-0.540	-0.326	-0.202	-0.040	0.240	0.760	1.020	1	0.495%	5.403
1.000	-0.010	-1.000	68	-0.209	0.436	0.087	3.829	3.028	3.028	4.631	0.261	-0.446	-0.386	-0.262	-0.208	-0.149	-0.039	-0.010	0	0.000%	9.083
1.000	0.000	-1.000	67	0.000	0.940	0.185	1.800	1.799	1.799	1.800	0.556	-0.540	-0.292	-0.140	0.020	0.160	0.280	0.400	0	0.000%	5.398
1.000	0.000	-1.000	67	0.361	1.560	0.382	0.873	0.557	1.188	0.557	1.146	-0.540	-0.356	0.040	0.340	0.700	0.940	1.020	1	1.493%	1.672

- **LSE:** limite superior de especificação para cada teste.
- **ALVO:** valor alvo (desejado) previamente especificado para cada teste.
- **LIE:** limite inferior de especificação para cada teste.
- **CNT:** contagem do número de testes, ou seja, quantidade de subgrupos coletados para cada teste.
- **Média:** média aritmética dos valores amostrados.
- **Amplitude (R):** diferença entre o maior e o menor valor entre os valores amostrados.
- **DP:** desvio padrão, também chamado de sigma, que é a medida de variação que representa a média dos desvios em relação a média aritmética dos dados:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

- **Pp:** índice de desempenho:

$$Pp = \frac{LSE - LIE}{6[DP(lp)]}$$

- **Ppk**: índice de desempenho definido como o mínimo entre Ppu e Ppl levando em consideração a centralização do processo:

$$Ppk = \text{Min}[Ppl, Ppu] = \text{Min}\left[\frac{LSE - \mu}{3s}, \frac{LSE - \mu}{3s}\right]$$

- **Ppl**: índice inferior de desempenho.
- **Ppu**: índice superior de desempenho.
- **Pr**: relação de desempenho (ou taxa de desempenho) definida como o inverso de Pp:

$$Pr = \frac{1}{Pp}$$

- **MÍN**: menor valor contido no conjunto de dados amostrados.
- **5%**: valor correspondente a concentração de 5% dos dados amostrados.
- **25%**: valor correspondente a concentração de 25%, ou seja, a  $\frac{1}{4}$  dos dados amostrados. Este valor representa o 1º quartil do Quadro e Margem, o qual é observado na seguinte posição ordenada:

$$Q1 = \frac{n + 1}{4}$$

Onde n é quantidade de subgrupos contidos em cada teste.

- **MED**: é a mediana (2º quartil) do conjunto de dados amostrados, ou seja, é o valor que ocupa a posição central dos dados ordenados e divide o conjunto de dados em duas partes iguais tal que o número de valores iguais ou maiores que a mediana é igual ao número de valores menores ou iguais que a mediana.
- **75%**: valor correspondente a concentração de 75%, ou seja, a  $\frac{3}{4}$  dos dados amostrados. Este valor representa o 3º quartil do Box Plot, o qual é observado na seguinte posição ordenada:

$$Q3 = \frac{3(n + 1)}{4}$$

Onde n é quantidade de subgrupos contidos em cada item do teste.

- **95%**: valor correspondente a concentração de 95% dos dados amostrados.
- **MÁX**: maior valor encontrado no conjunto de dados amostrados.
- **FDE**: número de itens que estão fora das especificações.
- **%FDE**: porcentagem de valores de cada item de teste que estão fora das especificações.
- **SIGL**: nível sigma equivalente baseado na porcentagem fora da especificação.