



Atividade: A maratona

A maratona é uma prova de atletismo que consiste em correr uma distância de **42.195 km**. Pelas suas características, este tipo de prova é realizada nas ruas de uma grande cidade ou na estrada. As principais cidades do mundo realizam um destes eventos anualmente, recebendo milhares de atletas profissionais e amadores que encaram o desafio e almejam finalizar a corrida ou melhorar o próprio tempo do passado.

Uma das mais famosas é a Maratona da Cidade de Nova Iorque, nos Estados Unidos. Veja na figura [Corredores participando da Maratona de Nova York, Wikipedia](#) realização de uma maratona em Nova Iorque. Com mais de 50.000 participantes cada ano, é um dos principais eventos do atletismo mundial, junto com as maratonas de Chicago, Londres, Boston, Berlim e Tóquio.



Figura 1: Corredores participando da Maratona de Nova York, [Wikipedia](#)

Os resultados do evento são divididos nas categorias de homens e mulheres, além disso, no evento participam cadeirantes e pessoas usando triciclos de mão (*handcycle*), categorias cujos resultados são premiados e publicados separadamente. Qual das categorias você acha que terá os melhores resultados na maratona? Em quanto tempo você acha que uma pessoa percorre os **42.195 km**? O que você acha ser mais rápido: correr em cadeira de rodas ou em triciclo de mão?



A seguir analisaremos os tempos de corrida das 100 melhores atletas na categoria de Mulheres da Maratona de Nova York do ano 2017, dados disponíveis no [site oficial da competição](#).

Observe no quadro a seguir, que os tempos já estão ordenados do menor para o maior e que para identificar o tempo da quadragésima sétima chegada, basta tomar a interseção da linha **7** com a coluna **+40** para obter o tempo **2 : 55 : 36**, ou seja, duas horas 55 minutos e 36 segundos.

Tabela 1: 100 melhores tempos de finalização da Maratona de Nova Iorque 2017 para mulheres (hora:minuto:segundo)

	+0	+10	+20	+30	+40	+50	+60	+70	+80	+90
1	2:26:53	2:32:01	2:42:52	2:49:44	2:53:59	2:56:58	2:58:35	2:59:36	3:01:24	3:03:43
2	2:27:54	2:32:09	2:44:26	2:49:59	2:54:42	2:57:05	2:58:36	2:59:41	3:01:26	3:03:46
3	2:28:08	2:33:18	2:44:48	2:50:04	2:54:52	2:57:10	2:58:50	2:59:43	3:01:28	3:04:02
4	2:29:36	2:34:10	2:45:20	2:50:05	2:55:04	2:57:40	2:58:52	2:59:46	3:01:44	3:04:04
5	2:29:39	2:34:23	2:45:52	2:51:11	2:55:25	2:57:49	2:58:56	2:59:51	3:02:09	3:04:17
6	2:29:39	2:36:38	2:46:45	2:53:01	2:55:34	2:57:49	2:59:01	2:59:56	3:02:15	3:04:26
7	2:29:41	2:37:22	2:47:04	2:53:02	2:55:36	2:57:50	2:59:03	3:00:02	3:02:39	3:04:42
8	2:29:56	2:37:33	2:47:30	2:53:02	2:55:39	2:58:08	2:59:10	3:00:05	3:02:41	3:04:49
9	2:31:21	2:39:01	2:47:35	2:53:19	2:56:47	2:58:23	2:59:16	3:00:49	3:02:56	3:04:58
10	2:31:44	2:40:09	2:49:37	2:53:38	2:56:57	2:58:26	2:59:23	3:01:18	3:03:32	3:05:09

Para calcular a média destes dados é conveniente reduzi-los a uma única unidade de medida, pois, caso contrário, seria necessário calcular três médias e, ainda fazer conversões apropriadas para obter a resposta em hora:minuto:segundo. Convertendo todos os tempos para minutos, obtemos o seguinte quadro de tempos arredondados para duas casas decimais.

	+0	+10	+20	+30	+40	+50	+60	+70	+80	+90
1	146,88	152,02	162,87	169,73	173,98	176,97	178,58	179,60	181,40	183,72
2	147,90	152,15	164,43	169,98	174,70	177,08	178,60	179,68	181,43	183,77
3	148,13	153,30	164,80	170,07	174,87	177,17	178,83	179,72	181,47	184,03
4	149,60	154,17	165,33	170,08	175,07	177,67	178,87	179,77	181,73	184,07
5	149,65	154,38	165,87	171,18	175,42	177,82	178,93	179,85	182,15	184,28
6	149,65	156,63	166,75	173,02	175,57	177,82	179,02	179,93	182,25	184,43
7	149,68	157,37	167,07	173,03	175,60	177,83	179,05	180,03	182,65	184,70
8	149,93	157,55	167,50	173,03	175,65	178,13	179,17	180,08	182,68	184,82
9	151,35	159,02	167,58	173,32	176,78	178,38	179,27	180,82	182,93	184,97
10	151,73	160,15	169,62	173,63	176,95	178,43	179,38	181,30	183,53	185,15

- a) Construa um histograma dos dados convertidos para horas, completando a [tabela 2](#), que indica os intervalos de classe (fechados no limite inferior e abertos no limite superior).

Intervalo	Frequência
[146,0; 150,0[
[150,0; 154,0[
[154,0; 158,0[
[158,0; 162,0[
[162,0; 166,0[
[166,0; 170,0[
[170,0; 174,0[
[174,0; 178,0[
[178,0; 182,0[
[182,0; 186,0[

Tabela 2: Intervalos de classes

b) Calcule o tempo médio dos 100 melhores tempos das corredoras, sabendo que a soma dos tempos é 17.191,66 minutos. Localize o valor encontrado no eixo horizontal do histograma. Em que posição ficaria uma corredora cujo tempo no qual completou a maratona é igual ao tempo médio calculado neste item?

c) Trace linhas verticais no histograma de modo a separar as classificações em 4 grupos: uma linha vertical que identifica o 25º lugar, separando os 25 primeiros colocados dos demais; outra, que identifica a 50ª classificação e, por fim, uma que marca o 75º tempo na classificação geral.

As marcações dos tempos das 25ª, 50ª e 75ª posições neste conjunto de 100 observações são chamadas de quartis da distribuição, este conceito será formalizado adiante.

d) Calcule os comprimentos dos intervalos de tempo determinados pela proposta de divisão no item d) e compare-os.

Intervalo	Comprimento
1º a 25º	
25º a 50º	
50º a 75º	
75º a 100º	

Observa-se que os comprimentos dos intervalos são bem diferentes, sendo maiores no início e mais estreitos no final.

Observe que apesar das diferenças de comprimento desses intervalos, cada um deles corresponde a cerca de 25 tempos.

e) O valor obtido para o tempo médio coincide com alguma das outras marcas feitas no histograma?

f) Observe que o tempo médio dos 100 melhores tempos para mulheres e o tempo da corredora que chegou em 50º. lugar são diferentes. Qual deles você escolheria como medida resumo destes dados? Por quê?

g) Que características da distribuição dos 100 melhores tempos para mulheres podem ser destacadas, analisando-se o histograma construído?

(Construção de boxplot simplificado) Considerando os valores mínimo (146,88 min.) e máximo (185,15 min.), trace uma reta (vertical ou horizontal), incluindo esse intervalo de variação. Por exemplo, de 146 a 186.

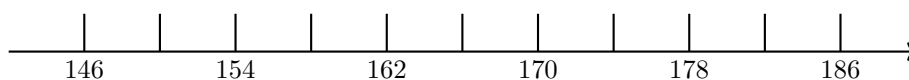


Figura 2: Sugestão de escala para a construção do boxplot dos 100 melhores tempos (em minutos) para a categoria mulheres

Em seguida, trace um retângulo, cujas bases estão nas posições referentes aos 25º e 75º tempos de chegada, cortando o retângulo por um segmento na posição referente ao 50º tempo de chegada. Para terminar a construção, trace um segmento, partindo do ponto médio da parte inferior até o valor mínimo e repita para o ponto médio da parte superior até o valor máximo.

A figura obtida é conhecida como boxplot (gráfico-caixa) do 100 melhores tempos sem a sinalização de valores discrepantes.

O boxplot é uma representação gráfica de dados quantitativos alternativa ao histograma. Esse gráfico é útil na comparação do comportamento de uma variável, considerando diferentes grupos, por exemplo, 100 melhores tempos de chegada entre homens e mulheres.