

Introdução

O seguinte relatório tem como objetivo analisar as características dos meteoritos documentados em relação a diversas variáveis, a fim de identificar possíveis associações entre elas.

Para isso, serão realizadas análises descritivas das variáveis. O banco de dados utilizado, chamado “Meteorite Landings”, foi obtido na plataforma “Kaggle” e contém 45.716 observações e 10 colunas, com informações como ano e local de queda, tipo de meteorito, nome, peso, entre outras.

O software utilizado para as análises foi o R versão 4.3.1.

Metodologia

Este relatório é composto por técnicas estatísticas que serão descritas a seguir de acordo com o que foi utilizado para a realização das análises.

Média

A média é a soma das observações dividida pelo número total delas, dada pela fórmula:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Com:

- $i = 1, 2, \dots, n$
- n = número total de observações

Mediana

Sejam as n observações de um conjunto de dados $X = X_{(1)}, X_{(2)}, \dots, X_{(n)}$ de determinada variável ordenadas de forma crescente. A mediana do conjunto de dados X é o valor que deixa metade das observações abaixo dela e metade dos dados acima.

Com isso, pode-se calcular a mediana da seguinte forma:

$$med(X) = \begin{cases} X_{\frac{n+1}{2}}, & \text{para } n \text{ ímpar;} \\ \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}, & \text{para } n \text{ par.} \end{cases}$$

Quartis

Os quartis são separatrizes que dividem o conjunto de dados em quatro partes iguais. O primeiro quartil (ou inferior) é o conjunto que delimita os 25% menores valores, o segundo representa a mediana e é o valor que ocupa a posição central (ou seja, metade dos dados estão abaixo dela e a outra metade está acima) e o terceiro delimita os 75% maiores valores. Inicialmente deve-se calcular a posição do quartil:

- Posição do primeiro quartil P_1 :

$$P_1 = \frac{n+1}{4}$$

- Posição da mediana (segundo quartil) P_2 :

$$P_2 = \frac{n+1}{2}$$

- Posição do terceiro quartil P_3 :

$$P_3 = \frac{3 \times (n+1)}{4}$$

Com n sendo o tamanho da amostra. Dessa forma, $X_{(P_i)}$ é o valor do i -ésimo quartil, onde $X_{(j)}$ representa a j -ésima observação dos dados ordenados.

Variância

A variância é uma medida que avalia o quanto que os dados estão dispersos em relação à média, em uma escala ao quadrado da escala dos dados.

Para uma população, a variância é dada por:

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

Com:

- X_i = i -ésima observação da população
- μ = média populacional
- N = tamanho da população

Desvio padrão

O desvio padrão é a raiz quadrada da variância. Avalia o quanto os dados estão dispersos em relação à média.

Para uma população, o desvio padrão é dado por:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}}$$

Com:

- X_i = i -ésima observação da população
- μ = média populacional
- N = tamanho da população

Alterações no banco

Ademais, vale ressaltar a necessidade de correção de valores no banco de dados. O meteorito “Northwest Africa 7701”, observação 30.683 do banco de dados, apresentava o ano em que foi encontrado com o valor 2101. Fazendo uma breve pesquisa, foi possível verificar que o meteorito foi encontrado em 2010. Supondo-se que houve um erro na digitação desse valor no banco, o valor foi devidamente corrigido.

Visualização dos resultados

Com base nos dados contidos no banco, foram construídos gráficos, tabelas e mapas a fim de facilitar a visualização dos resultados obtidos. Assim, será possível interpretar as características das observações de forma mais clara.

Análises

Número de meteoritos achados ou observados por ano

Esta seção tem como objetivo analisar a quantidade de meteoritos documentados a cada ano. Os meteoritos estão classificados como “observado”, caso tenha sido vistos durante a queda, ou “achado”, caso tenham sido encontrados já na superfície terrestre.

O registro mais antigo de um meteorito observado é do ano de 860, enquanto os meteoritos mais recentes documentados no banco de dados são do ano de 2013.

Ademais, vale salientar que, antes de ano de 1972, o ano com mais meteoritos documentados foi 1969, com 70 observações. No entanto, após 1971, o número de meteoritos achados por ano aumentou significativamente, de modo que, para melhor visualização dos dados, foi realizado um gráfico de linhas do número de meteoritos por ano a partir de 1972.

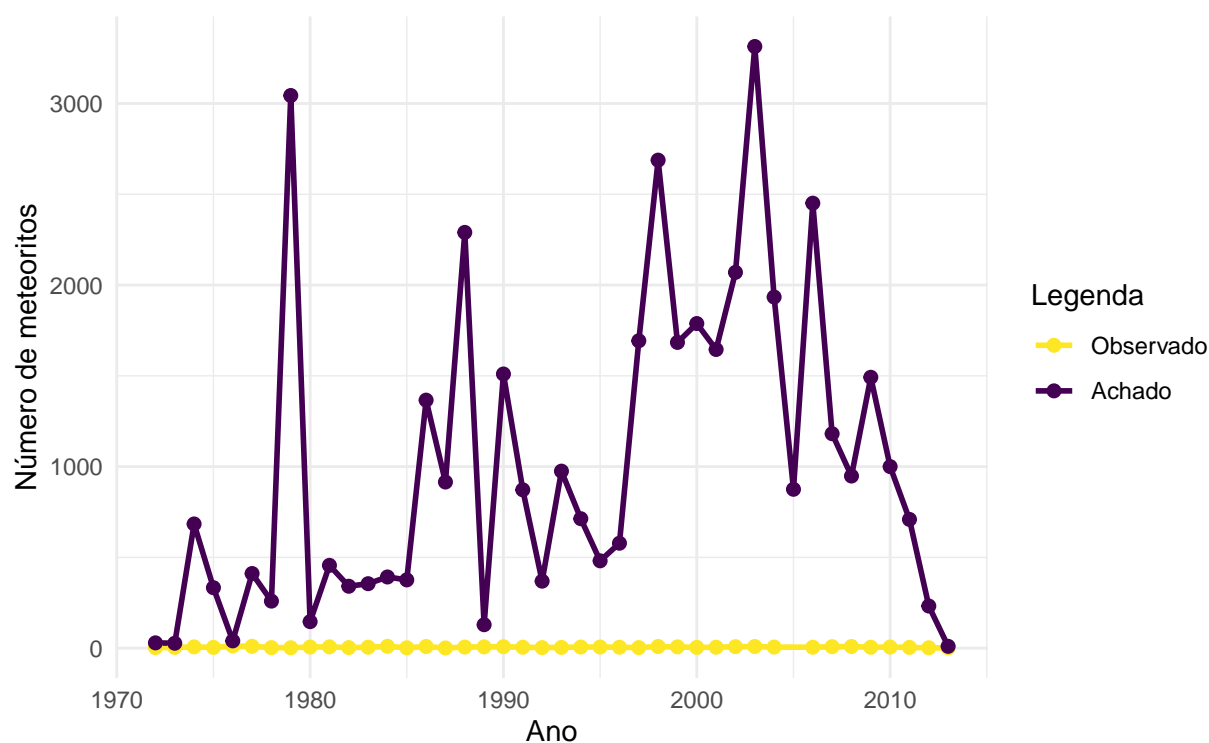


Figure 1: Gráfico de linhas do número de meteoritos registrados por ano

Table 1: Cinco anos com maior registro de meteoritos

Ano	Nº de meteoritos
2003	3323
1979	3046
1998	2697
2006	2456
1988	2296

Com base no gráfico, observa-se que a quantidade de meteoritos achados foi muito maior do que o número de meteoritos observados. Durante o período registrado, a maior quantidade de meteoritos observados em

um ano foi de 10, fazendo com que os valores de meteoritos observados se aproximem muito de 0 no gráfico em função da escala.

Além disso, o gráfico apresenta alguns picos evidentes, que foram representados na tabela acima. Esses picos indicam os cinco anos com mais documentação de meteoritos, sendo o maior deles no ano de 2003.

Classificação dos meteoritos

Nesta seção, serão identificadas as principais classificações dos meteoritos presentes no banco de dados. Os meteoritos são classificados de acordo com sua composição e estrutura. Dentre os dados analisados, existem 466 tipos de meteoritos diferentes. Dessa forma, serão analisados somente os dez tipos mais frequentes.

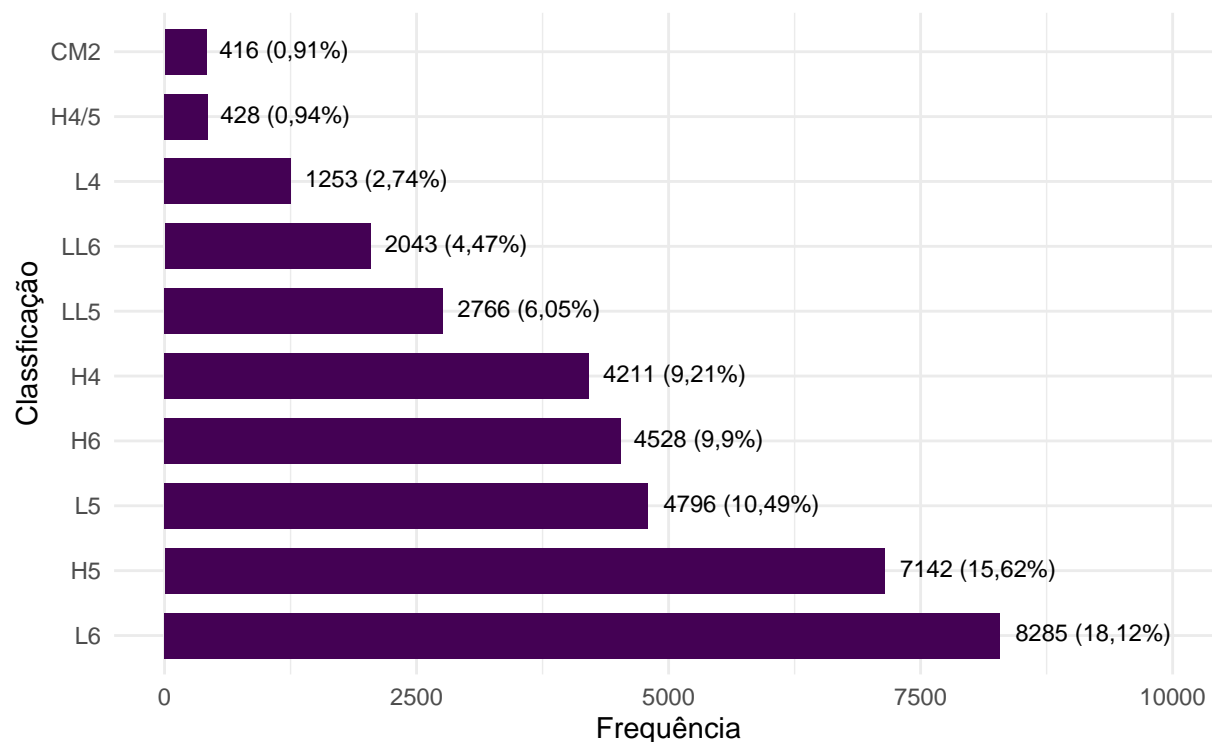


Figure 2: Gráfico de barras das 10 principais classificações de meteorito

Pelo gráfico acima, é possível perceber que os dez meteoritos mais frequentes representam 78,45% do total de observações. Além disso, nota-se que o tipo de meteorito com maior ocorrência é o L6, compondo 18,12% dos dados.

O menor valor apresentado no gráfico foi 0,91%, que representa a porcentagem de meteoritos do tipo CM2. Com isso, pode-se inferir que a porcentagem de todos os demais tipos de meteoritos não mostrados na figura é menor que 1%.

Peso dos meteoritos separado pela classificação

Esta seção tem como objetivo analisar o peso dos meteoritos de acordo com suas dez classificações principais. O peso é uma variável quantitativa contínua que varia de 0,00001 a 4000 gramas.

A seguir, observamos o gráfico e a tabela com a distribuição das variáveis:

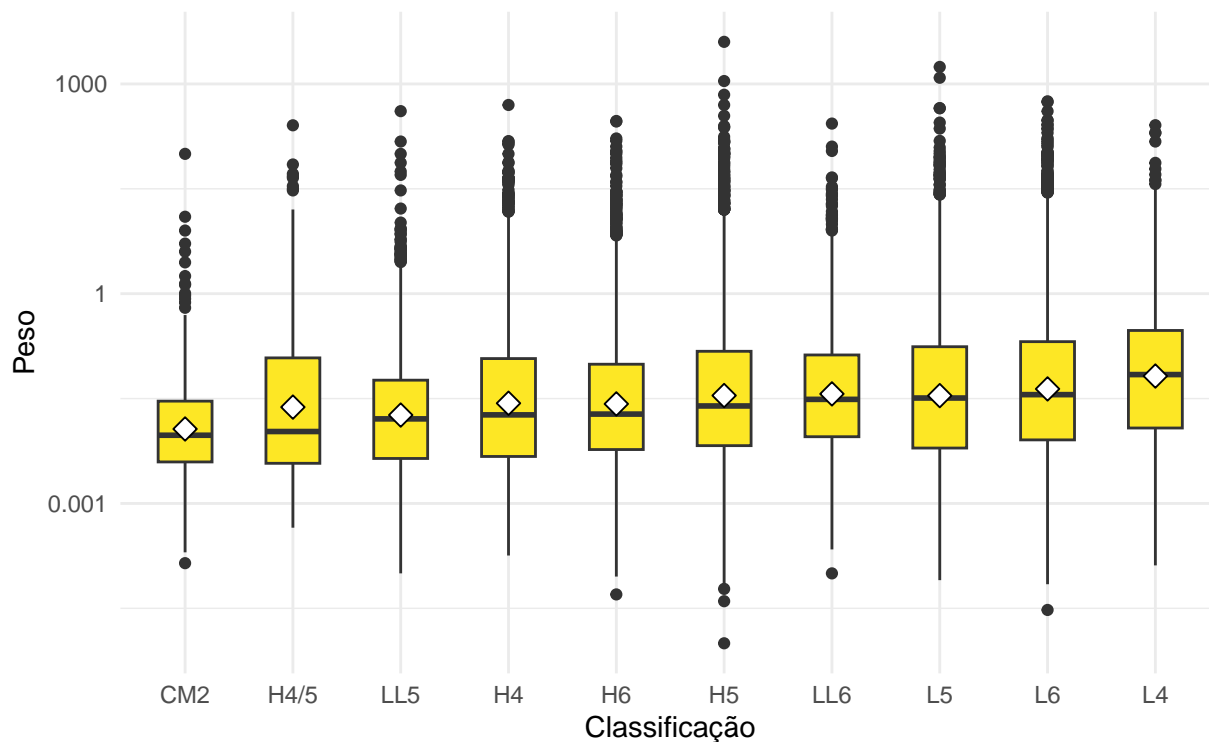


Figure 3: Boxplot do peso dos meteoritos nas 10 principais classificações

Table 2: Medidas resumo

	CM2	H4	H4/5	H5	H6	L4	L5	L6	LL5	LL6
Média	0,37	1,00	1,60	2,17	0,86	1,64	1,80	1,45	0,46	0,69
Desvio Padrão	4,98	9,86	13,70	51,17	8,52	12,43	33,13	14,01	8,75	7,48
Variância	24,77	97,14	187,80	2618,04	72,51	154,56	1097,65	196,29	76,59	55,91
Mínimo	0,00014	0,00018	0,00045	0,00001	0,00005	0,00013	0,00008	0,00003	0,00010	0,00010
1º Quartil	0,0039	0,0047	0,0037	0,0067	0,0059	0,0120	0,0062	0,0081	0,0044	0,0090
Mediana	0,01	0,02	0,01	0,02	0,02	0,07	0,03	0,04	0,02	0,03
3º Quartil	0,03	0,12	0,12	0,15	0,10	0,30	0,17	0,21	0,06	0,13
Máximo	100	500	256	4000	295	257	1750	564	408	271

A partir das informações obtidas, percebe-se que existem algumas variações nos pesos de cada categoria de meteorito. Primeiramente, é interessante destacar o comportamento dos meteoritos tipo H5: essa categoria apresentou a maior média e o maior desvio padrão, com valores de 2,17 e 51,17, respectivamente. Além disso, essa classe possui o menor e o maior peso observados na tabela, e é também a classe com mais outliers, tanto inferiores quanto superiores.

Em relação aos demais grupos, vale notar que a menor média e o menor desvio padrão estão na classe CM2, com valores de 0,37 e 4,98, respectivamente. Pelo gráfico, é evidente que a categoria CM2 apresenta a

menor caixa do boxplot. A maior mediana é de 0,07, na categoria L4, enquanto o menor valor de mediana é de 0,1, para os meteoritos H4/5.

Além disso, observa-se uma grande quantidade de outliers acima do limite superior em todos os grupos, enquanto os outliers inferiores aparecem apenas em CM2, H6, H5, LL6 e L6.

Localização da queda dos meteoritos

O mapa a seguir pode ser utilizado para melhor compreender a distribuição dos locais onde meteoritos foram encontrados ao redor do mundo.

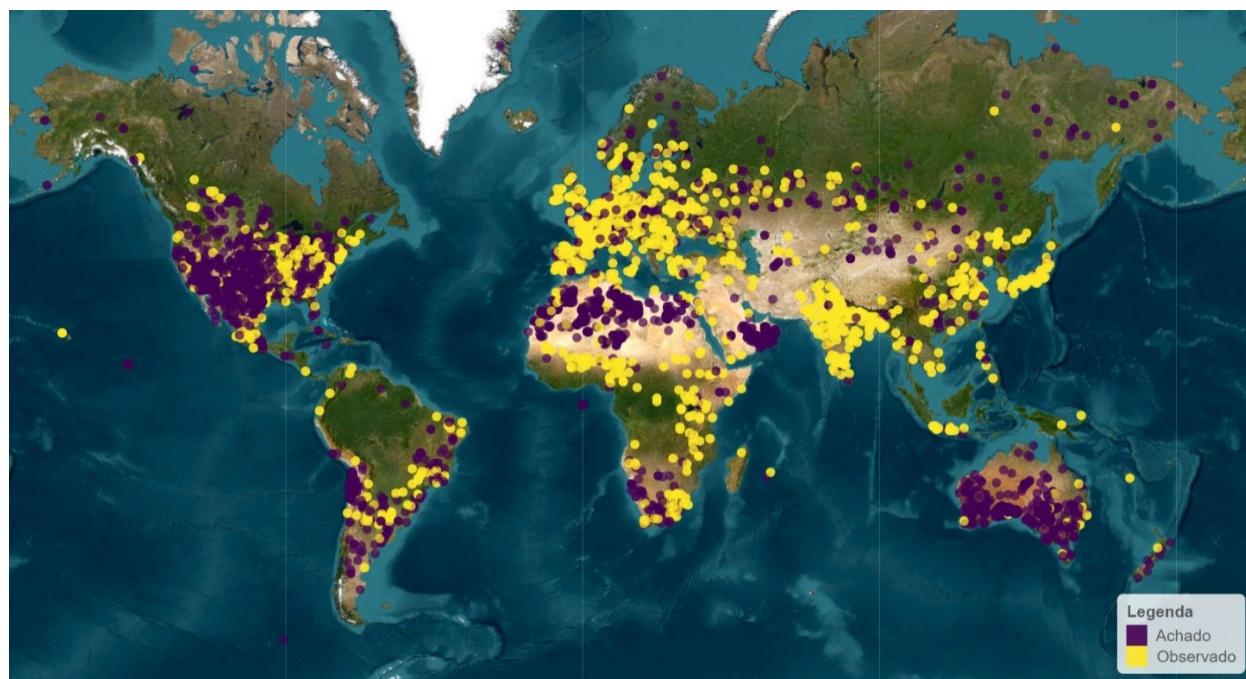


Figure 4: Mapa dos locais de queda de meteoritos

[Link de uma versão interativa](#)

Com base no mapa acima, é possível visualizar um comportamento interessante na distribuição dos locais de queda dos meteoritos documentados. Primeiramente, observa-se que há uma concentração de meteoritos achados na América do Norte, na Oceania, no norte da África e na península arábica, sobretudo na região onde fica o Omã. Já os meteoritos que foram observados em queda se encontram concentrados na Europa, na África Subsaariana e na região onde se encontra Índia, leste da China, Japão e países do sudeste asiático.

Ademais, também pode-se notar que existem regiões onde poucos ou nenhum meteorito foi encontrado. A partir disso, é possível levantar algumas hipóteses sobre esse fenômeno. As regiões norte do Canadá e da Rússia e a Groenlândia, por exemplo, apresentam climas frios extremos e são, portanto, pouco habitadas, de modo que meteoritos que caem nessas regiões dificilmente são encontrados.

Outras regiões em que se observa uma ausência do registro de meteoritos são o centro da América do Sul e da África. Uma das similaridades entre as duas regiões é que ambas apresentam florestas tropicais densas, a Floresta Amazônica e a Floresta do Congo. Assim, as características da vegetação, do clima e do relevo dessas regiões poderia explicar a dificuldade em descobrir meteoritos caídos nesses locais.

Conclusão

Participações

Júlia Lima Nosralla (232014825)

Leonardo Farias Saraiva (232014834)