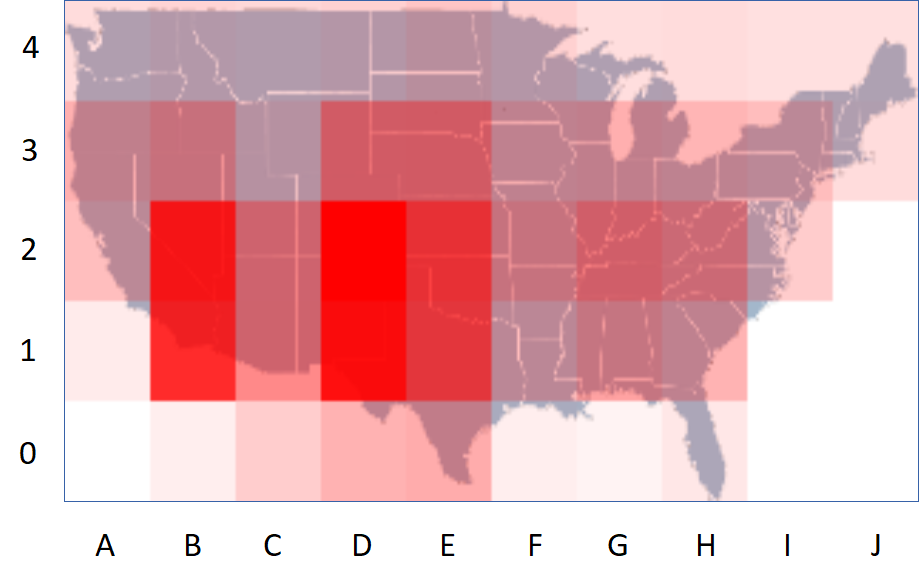
ABSTRACT

We have identified various pockets of the contiguous United States with higher than average meteorite activity. We have developed a system to rank these areas in terms of "dangerousness", frequency, or mass characteristics. Using these characteristics, our system performs calculations on over 40000 entries to determine statistics such as the mean, the interquartile range of mass, standard deviation, and the chance of meteorite landings in the future. Our intention is for the reader to gain an understanding of meteorite activity within US territory and to develop a predictive model for this data.

INTRO

Meteorites are meteors that survive atmospheric entry and land somewhere on the Earth. The main benefit of meteorites is their scientific research value. Meteorites can provide us information about the early formation of our galaxy. A meteorite can provide data about planetary bodies that would otherwise be inaccessible except through a mission to that place. Meteorites allow us to gather information about the formation of our universe at an exponentially smaller cost over sending out equipment to places outside of earth. They also make great collectors’ items and can fetch a high price due to their reputation. Our data is pulled from NASA's open archive of datasets that are publicly available online. This dataset was originally started by The Meteorological Society and contains an entry for every known meteorite landing dating back all the way to the early 861 AD. A link to our specific dataset is here: <https://data.nasa.gov/Space-Science/Meteorite-Landings/gh4g-9sfh/data>. We made a contemporary dashboard website that can perform statistic analysis of the dataset and provide visualizations of several different data focuses. The primary reason for creating this dashboard is to study and analyze the patterns that meteorites follow in their landing and construction. A falling meteorite may cause damage depending on its size and material. By examining our dashboard, we can find the most common areas for meteorites to land or even the areas with the highest concentration of very large (greater than 1000kg mass) meteorites. Using our dashboard, we found that in the US, before much expansion westward happened, a lot of meteorite activity was focused in the southern United States, roughly in the Tennessee area. However, as time moves on, we noticed a shift towards the New Mexico area, which became the #1 area for meteorite activity in the United States. I theorize that it may have to do with less forested conditions of the area, making it easier to find a meteorite once it has fallen to the ground. Another explanation may be simply that the expansion of the United States and the drastically increasing population leads to a higher chance of preexisting meteorite landings to be found and recorded.

We decided to mainly focus on the meteorite landings in the contiguous United States, as that information is the most pertinent to our situation. We broke the area into 50 evenly sized areas in a 10 by 5 grid as shown in Figure 1. Individual areas of the map will henceforth be referred to by the grid axis depicted in Figure 1.



**Figure 1**

One of the first measures that we implemented was the average mass. The average mass gave us a very rough idea of what kind of mass numbers we would be dealing with. While the dataset has its masses measured in grams, we decided to label all mass values in kilograms because of the relatively large average mass values. Figure 2 shows the top 10 areas for average mass.

|  |  |  |
| --- | --- | --- |
| Region | Average Mass (kg) | Entries |
| A4 | 1732.923 | 9 |
| D0 | 1449.360 | 43 |
| C0 | 1401.453 | 18 |
| C1 | 372.960 | 100 |
| A3 | 372.736 | 45 |
| C2 | 355.874 | 98 |
| F1 | 49.177 | 32 |
| E2 | 45.855 | 253 |
| F3 | 43.346 | 42 |
| C3 | 41.795 | 20 |

**Figure 2**

As you can tell from the spread of this data, there are too many outliers to practically use average mass as any sort of viable figure. Therefore, we decided to take the approach of using interquartile range values for studying the masses of meteorites. Interquartile range will give us a much more accurate method representing the mass data that we have. Figure 3 shows the top 10 regions sorted by average mass using only values that fall in the interquartile range for their region.

|  |  |  |  |
| --- | --- | --- | --- |
| Region | Average Mass (kg) | Entries | Mass Range (kg) |
| C0 | 53.826 | 10 | 7.632 - 140.000 |
| D0 | 27.061 | 24 | 0.610 - 85.000 |
| H4 | 13.702 | 5 | 1.090 - 42.650 |
| A4 | 11.821 | 5 | 0.143 - 28.397 |
| G0 | 11.300 | 1 | 11.300 - 11.300 |
| F3 | 10.672 | 22 | 0.880 - 25.300 |
| F0 | 9.960 | 2 | 0.664 - 19.256 |
| E4 | 9.406 | 32 | 2.130 - 21.180 |
| G1 | 9.356 | 35 | 1.455 - 34.500 |
| F2 | 9.264 | 25 | 1.600 - 30.000 |

**Figure 3**

When you compare the values in Figure 2 with the values in Figure 3, you see clearly just how far the outliers were skewing the data. There is a change of 2 significant digits in the highest ranking area. Compare the values in figure 3 to those in figure 4, where instead of eliminating values based on interquartile range, we used the standard deviation to eliminate extreme values that fall out of the range of μ ± ơ. in Figure 4, we see values that closer to that of Figure 3. They are still a whole order of magnitude smaller than our initial average mass calculations. However, when we consider that the data in Figure 3 is dependent on the average mass as part of its calculation, the interquartile range data seems like a much better representation.

|  |  |  |  |
| --- | --- | --- | --- |
| Region | Average Mass (kg) | Entries | Mass Range (kg) |
| C0 | 189.774 | 17 | 0 - 423.623 |
| D0 | 148.062 | 40 | 0 - 467.563 |
| C1 | 73.697 | 99 | 0 - 297.415 |
| C2 | 50.265 | 97 | 0 - 239.062 |
| A3 | 28.935 | 44 | 0 - 176.263 |
| E2 | 18.390 | 247 | 0 - 41.359 |
| G1 | 16.858 | 57 | 0 - 30.173 |
| C3 | 16.153 | 19 | 0 - 33.674 |
| F3 | 16.145 | 38 | 0 - 30.486 |
| H1 | 15.096 | 40 | 0 - 28.143 |

**Figure 4**

NOTE: CALCULATE STDDEV INSIDE IQR

Next, we’ll look at the standard deviation of the masses in the regions. We will use the mass values of the entries that fall in the interquartile range to calculate the standard deviation.

A lot more information is available at this resource. http://www.lpi.usra.edu/science/kring/epo\_web/meteorites/toc.html