Meteoric

Landings

Analysis and Prediction

Report

Presented by-

Team “Kuch Bhi”

*Chitkara University*

Index

**Topic Page Number**

Synopsis 3

About the dataset 4

Elements used… 5

Logic and Approaches… 6

Approach to Analysis 7

Analysis and Conclusions 13

Analysis Attributes 18

Conclusion 19

*Synopsis*

Scientists predict over 44,0000 tonnes of meteoritic mass falls on the earth daily. As the earth is being bombarded with astronomical debris.

We were provided with the dataset containing Information on the geographic location as well as structure of the meteorite.

We conducted an analysis on the data, visualizing it and predict areas with the highest showers.

The dataset was provided by the organization, which we presumed to be cleaned and ready for use beforehand.

*About the Dataset*

The Meteoritical Society collects data on meteorites that have fallen to Earth from outer space. This dataset includes the location, mass, composition, and fall year for over 45,000 meteorites that have struck our planet.

The dataset contains the following variables:

* **name**: the name of the meteorite (typically a location, often modified with a number, year, composition, etc)
* **id**: a unique identifier for the meteorite
* **nametype**: one of:  
  -- *valid*: a typical meteorite  
  -- *relict*: a meteorite that has been highly degraded by weather on Earth
* **recclass**: the class of the meteorite; one of a large number of classes based on physical, chemical, and other characteristics (see the Wikipedia article on meteorite classification for a primer)
* **mass**: the mass of the meteorite, in grams
* **fall**: whether the meteorite was seen falling, or was discovered after its impact; one of:  
  -- *Fell*: the meteorite's fall was observed  
  -- *Found*: the meteorite's fall was not observed
* **year**: the year the meteorite fell, or the year it was found (depending on the value of **fell**)
* **reclat**: the latitude of the meteorite's landing
* **reclong**: the longitude of the meteorite's landing
* **GeoLocation**: a parentheses-enclose, comma-separated tuple that combines **reclat** and **reclong**

*Elements Used for Analysis and prediction*

Dataset used:

* + <https://tinyurl.com/yc553x2z>

Technologies used:

* + Microsoft Excel
  + Python
  + Kaggle
  + Plotly
  + Pandas
  + Matplotlib
  + Numpy
  + GeoPy
  + Scatter plots
  + Heatmaps

Kaggle notebook link

<https://www.kaggle.com/code/shubhankermehta/meteor-shower>

*Logic and approaches used*

The problem statement is as follows:

***Scientists predict over 44,0000 tonnes of meteoritic mass falls on the earth daily. As the earth is being bombarded with astronomical debris, here is a dataset containing Information on the geographic location as well as structure of the meteorite. Conduct An analysis on the data, visualize it and predict areas with the highest showers.***

The team considered may approaches to the problem and after a discussion we decided to implement a hybrid approach of machine accuracy and human intuition.

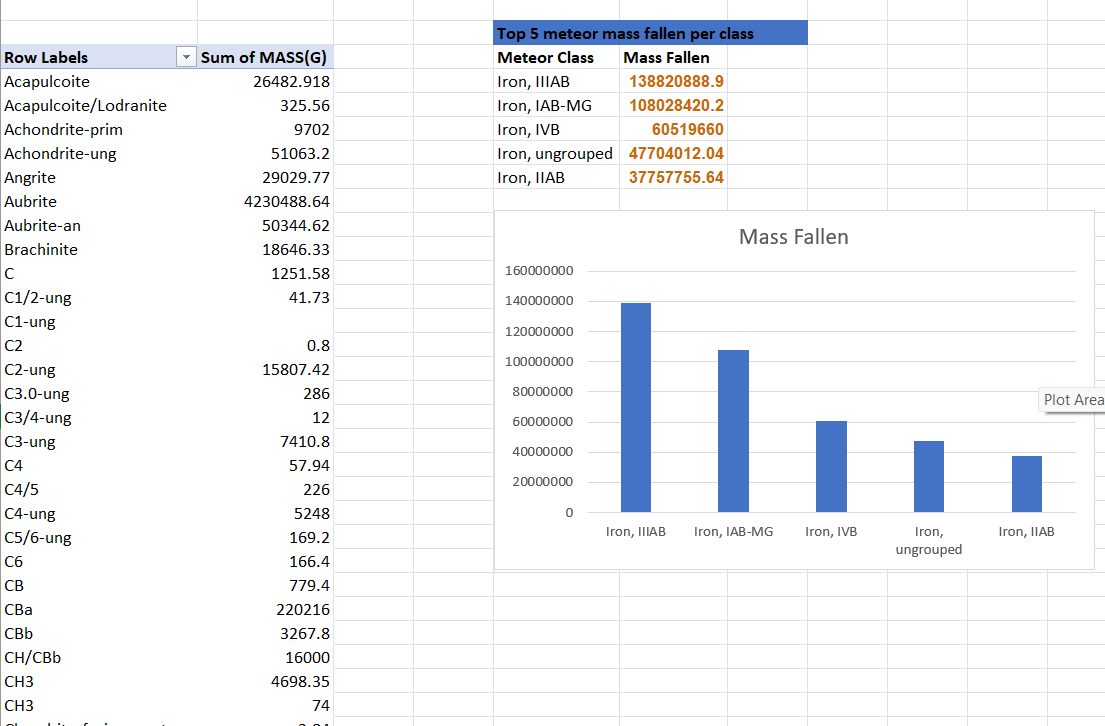
The logic used here is:

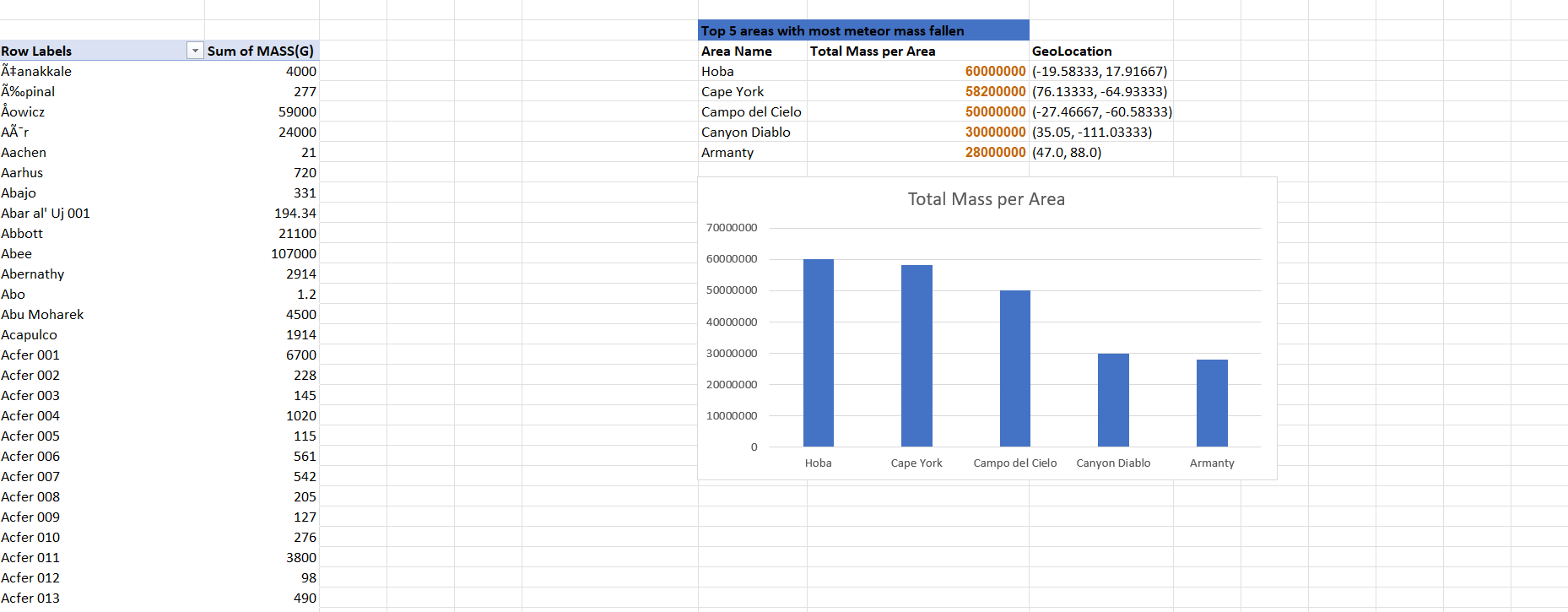
*There are particular areas which have a significantly higher amount of meteoric showers as compared to other areas. We will deduce which these areas are. Then we will observe which out of the shortlisted areas have the most recent meteoric showers.*

*The combination of high frequency and time frame recency will enable us to predict which areas will receive the most meteoric showers.*

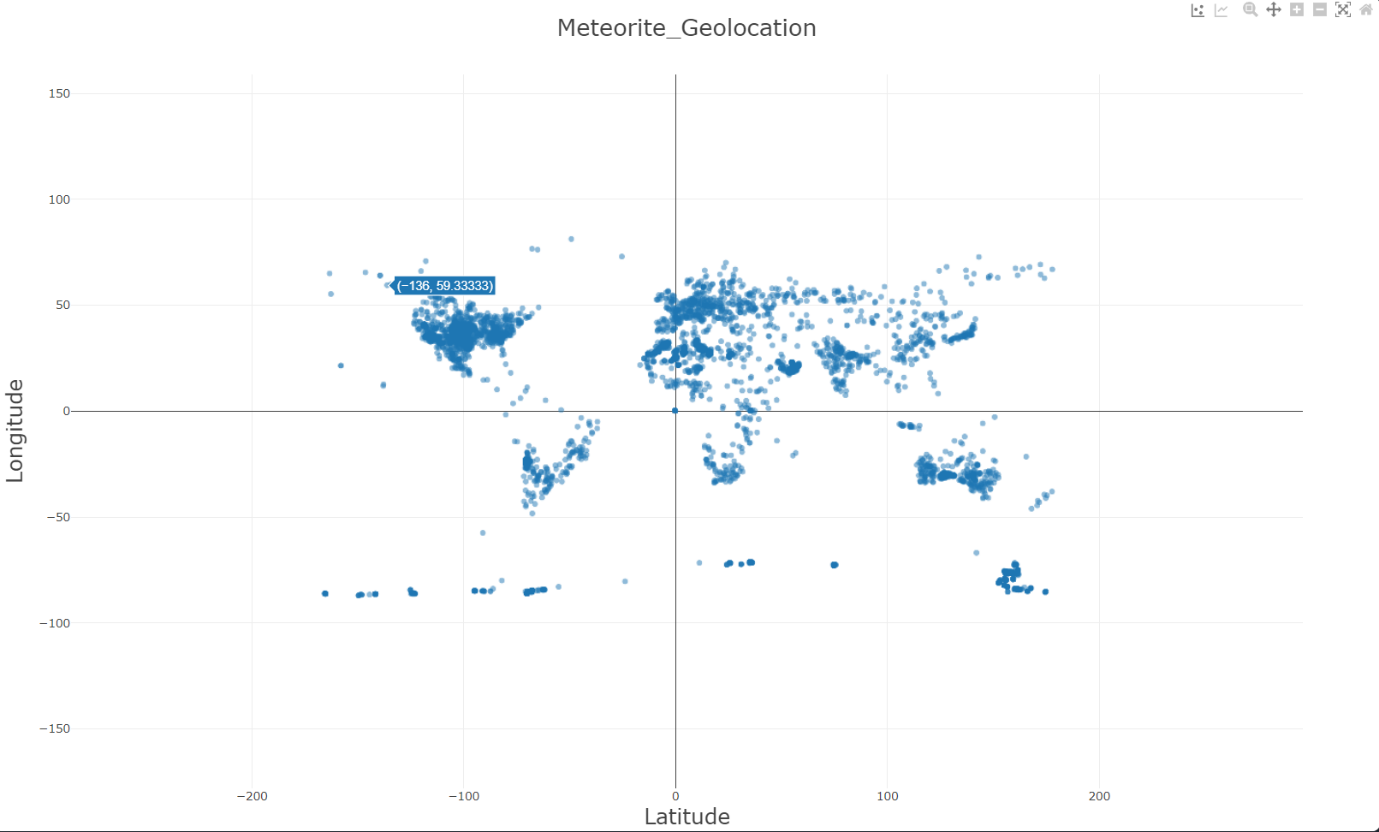
*Approach to Analysis*

The given dataset had a significantly large number of data entries. The team first condensed the 45,717-record data into usable pivot tables using Excel.





After condensing the data into more insightful data blocks, we then considered the frequency of the meteoric falls in given areas. For this the provided geolocation coordinate were used. Afterwards, the team considered the geolocation coordinates provided within the dataset to plot a density scatter plot.

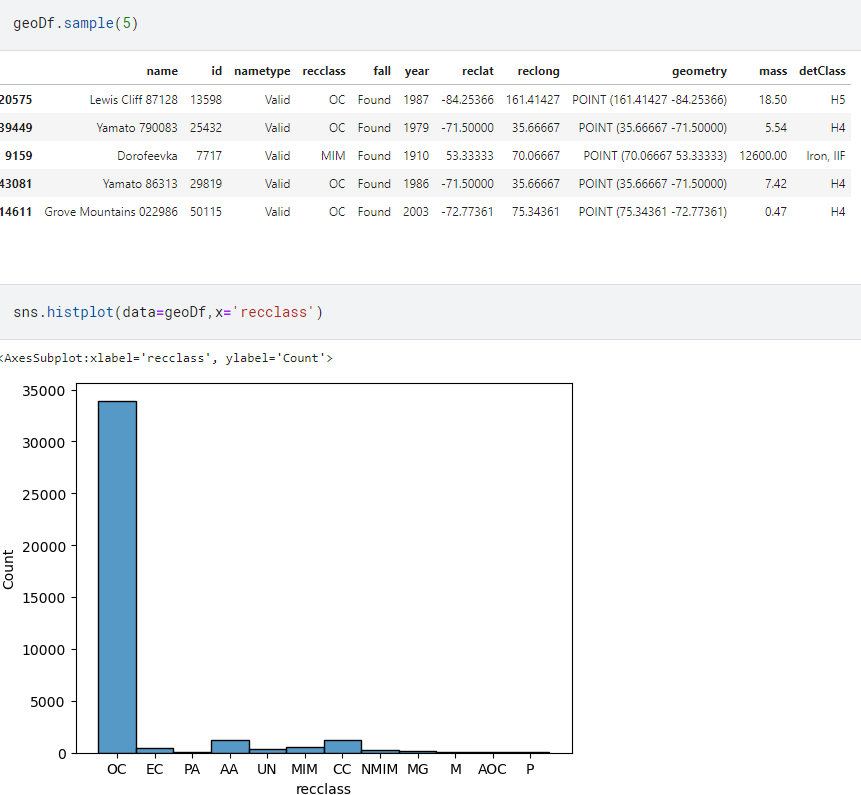


We could now observe which areas have a high amount of meteoric showers using the coordinates.

The dataset contained meteorite classes, which were in a very high degree of detail.

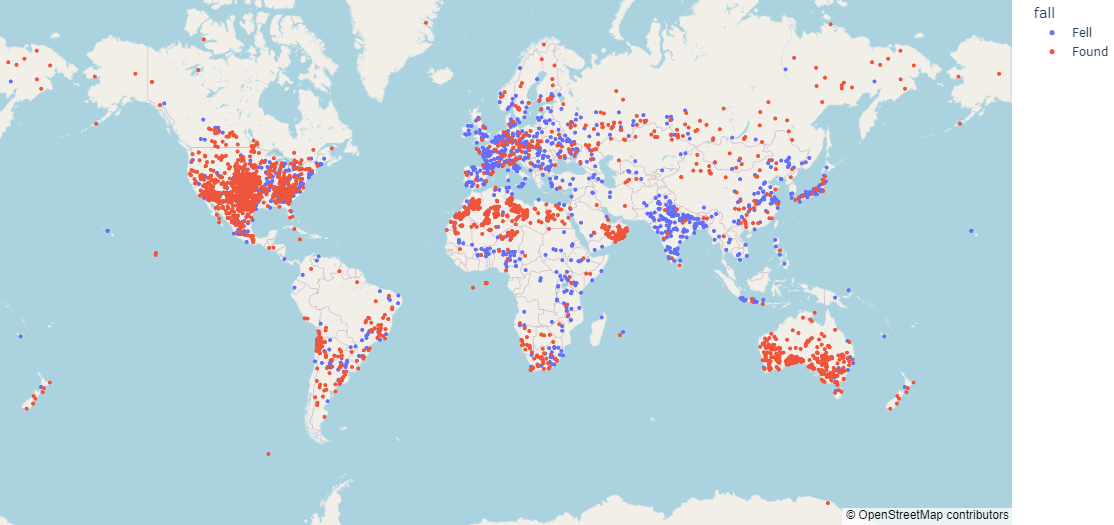
Since this would’ve hindered our approach and not result in any significant analytical results, we decide to further classify the classes into major categories.

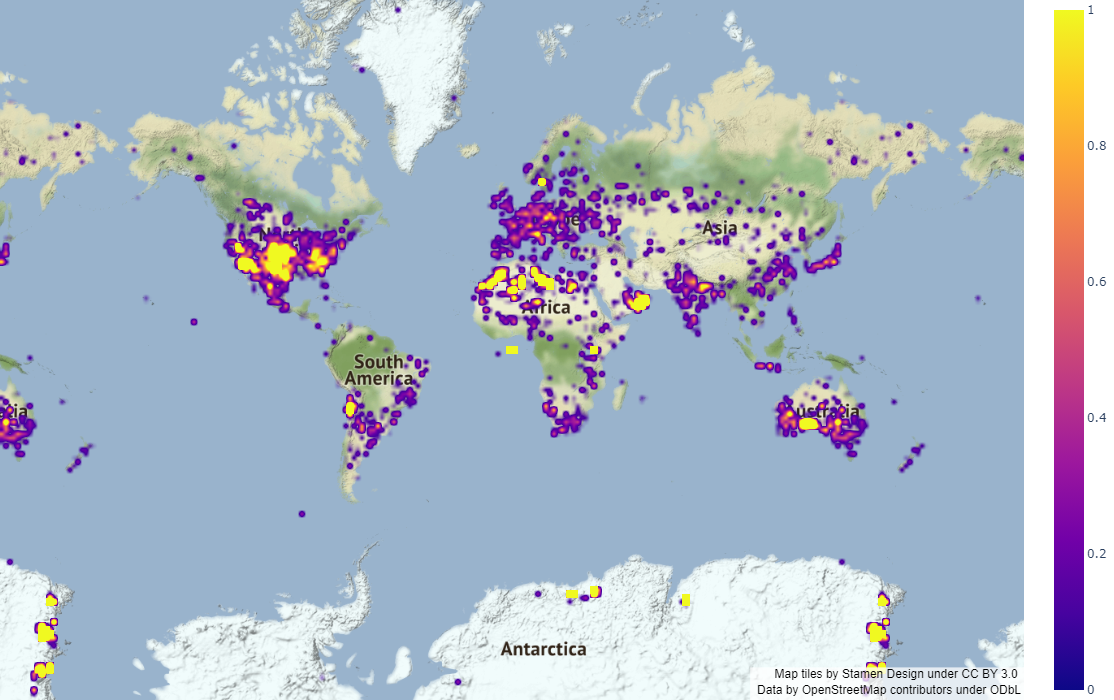
Here, python libraries such as pandas and numPy were used. Wikipedia as reference was used (reference link [here](https://en.wikipedia.org/wiki/Meteorite_classification)).



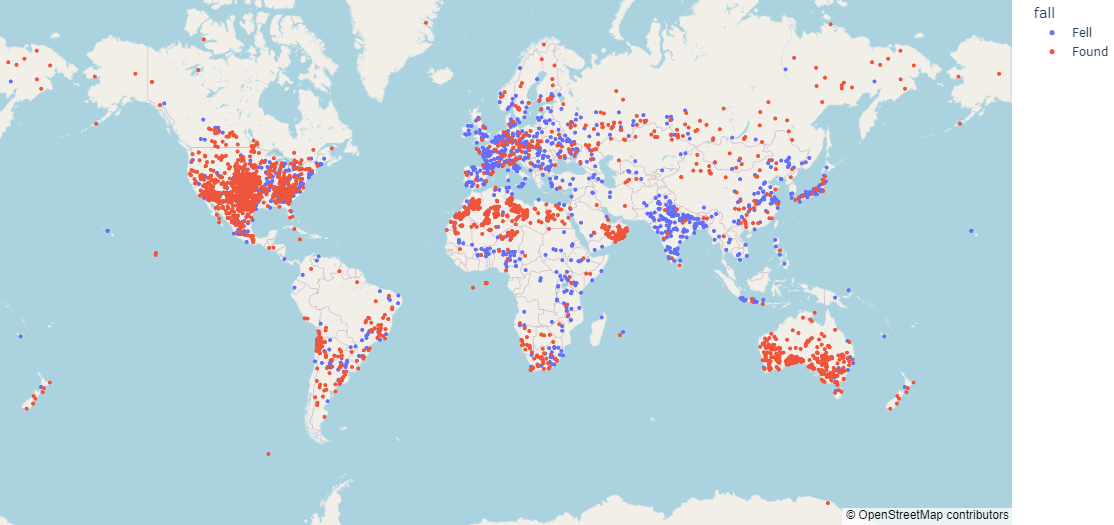
After the team condensed all the given data into concise and more useful datasets we then proceeded to create visualizations that will help us observe and analyse the data.

Using plotly and matplotlib, python libraries to visualize data we created two unique maps based on the geolocation frequency of the fallen meteorites.



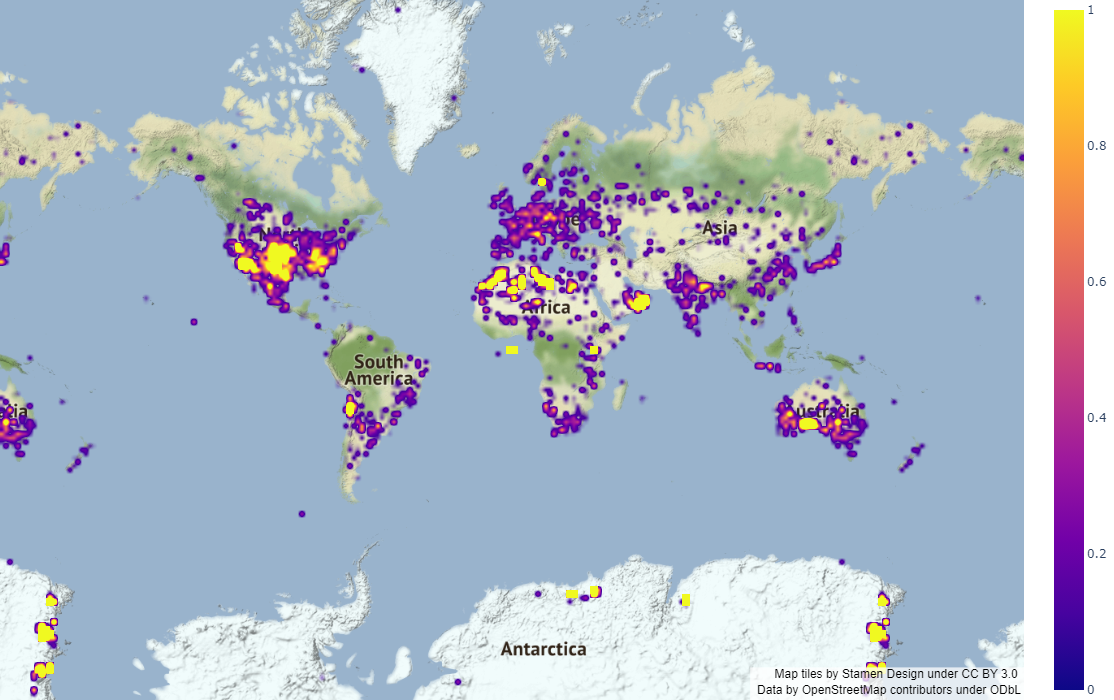


The first map was a geolocation frequency map that showed us the frequency of meteorites fallen in regards to the metrics of ***fell*** and ***found*** properties of the meteorites(for information on ***fell*** and ***found*** refer to the *About the Dataset* page).



This map gave us the opportunity to deduce which areas were the ones where active meteorite showers were taking place.

The second map that was generated was frequency based heatmap, that allowed us to better understand which areas receive the most amount of Astro-debris.



This map enabled us to observe which areas had the most density of meteorite showers and filter those particular areas out for further analysis.

*Analysis and Conclusions*

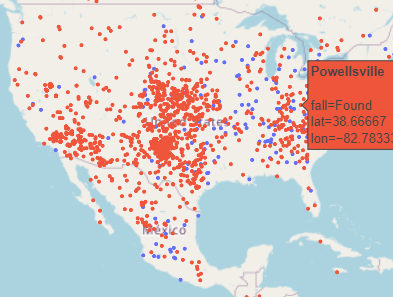
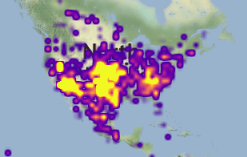
After generating various pivot tables, dataframes, plots and maps, the team felt confident in the refined information produced from the data provided to now commence with the Exploratory Data Analysis and draw conclusions.

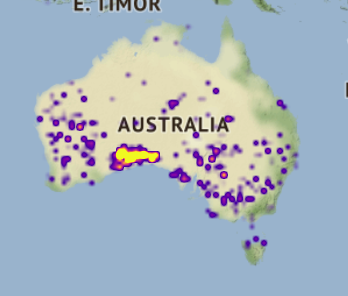
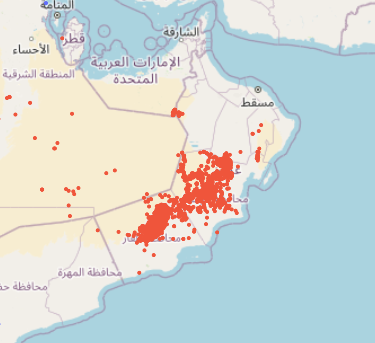
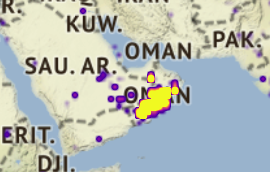
This was done by a visual analysis by the members of the team first, then cross referencing it with more data in the database, primarily year and frequency-based data.

Through visual analysis of both the ***fall****-*category based maps and frequency heatmap based maps we concluded with confidence that areas with the most meteor showers were as follows:

* + *The United States of America*
  + *Oman*
  + *Northern Africa*
  + *Southern Australia*

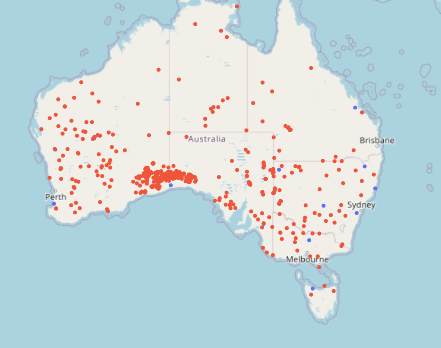
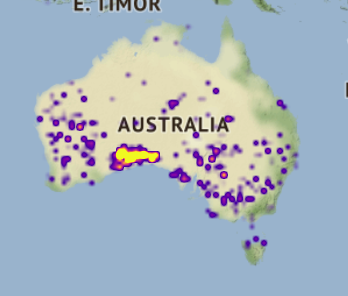
All of the above listed areas had the most concentrated areas of data points generated by the maps.



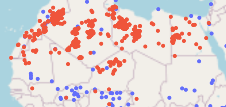


**USA**

**Oman**



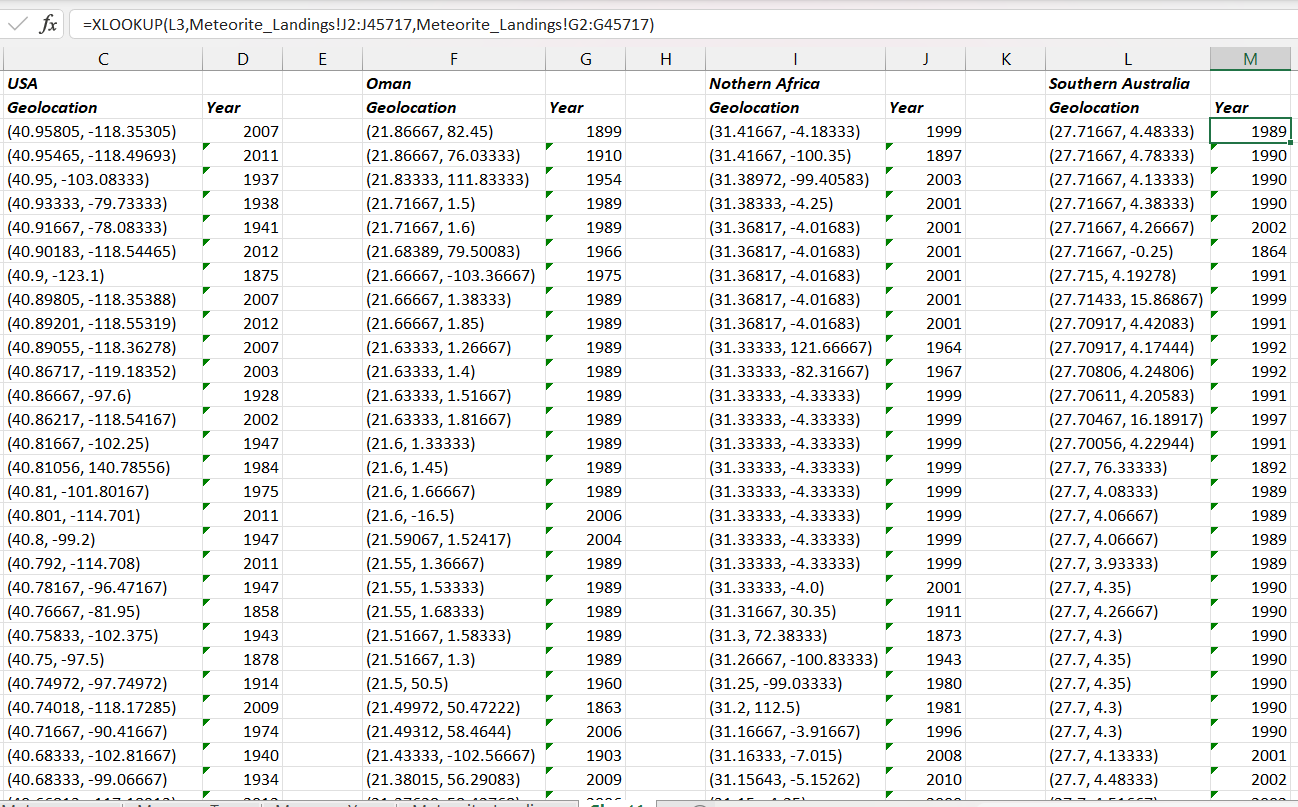
**Southern Australia**

****

**Northern Africa**

Since we also wanted the areas to have a meteor shower in a recent timeframe i.e., 1800s and above we checked the areas geolocation coordinates with years the meteorites fell.

After the visual analysis and location confirmation, we then proceeded to cross reference each major geolocation coordinate with the year it fell to find out the recency of the meteorites fallen.



With this we were able to confirm the recency of the meteorite showers in the deduced areas, hence confirming our prediction

After some more fine tuning of the data and a significant amount of data analysis, we have concluded the following.

The areas which received the most amount of meteorite showers are:

*The USA, Oman, Northern Africa, Southern Australia, and European continent.*

The areas which received Astro-debris most recently were:

*The USA, Oman, Northern Africa, Southern Australia*

Keeping in mind the frequency of meteorite showers and the recent time frame, we can confidently predict that the USA, Oman, Southern Australia and Northern Africa will be the areas which will receive the most amount of meteorite showers.

The European Continent, while receiving a significant amount of showers as well, has majority of its Astro-debris in a ***fell*** state and hasn’t received any showers in recent times as well. Therefore, we can conclude that it is not a hot zone for meteoric showers.

*Analysis Attributes*

This analysis and prediction project was a hybrid of machine accuracy and human intuition, drawing the best properties of both sides and delivering a result that the team is confident in.

We estimate a prediction accuracy of 90 percent with this analysis, keeping in mind the random erratic nature of Astro-debris and acts of God.

No particular Machine Learning model or algorithm was implemented as, given the dataset, there were far too many differentiating variables for any particular ML model or algorithm to accurately process and give satisfactory results.

We have confidence that this hybrid we have taken is one of the better approaches to this particular problem case.

*Conclusion*

The meteorite landings analysis and prediction project aims to provide a significantly accurate model that can analyze past trends in meteorite showers and predict the areas which have the most amount of showers.

The team has worked rigorously and have delivered results that are accurate and can be implemented with fair confidence.

*TEAM “Kuch Bhi” (Chitkara University)*

*Members-*

* + *Shubhanker Mehta*
  + *Shaurya Vashisht*
  + *Aseespreet Singh*
  + *Aakash Dixit*