

Final Project Proposal

The group members (names and UNIs):

Jiayi Ge jg4890, Yi Su ys3924, Zitao Zhang zz3308

The tentative project title:

Meteorite Landing

The motivation for this project:

This project explores meteorite landing data to understand the spatial and temporal distribution of impacts on Earth, with a focus on patterns in landing locations, mass distributions, and trends over time. Our interest was sparked by recent celestial events, like the Atlas Comet's close approach and the "two moons" phenomenon, we became curious about how meteorites interact with Earth. By studying global landing sites, we hope to gain insights into these cosmic fragments, connecting us more closely with the universe's diversity.

The intended final products:

We will create a website with an interactive map of meteorite landings, a dashboard with key visualizations like geographic distribution and time series, and a brief introduction and summary of findings for concise, interactive exploration of meteorite data.

The anticipated data sources:

Our main dataset, "Exploring Meteorite Landing Sites," is sourced from Kaggle and contains detailed information on meteorite landings recorded worldwide. This dataset includes columns such as the name, mass, classification, and geographic coordinates of each meteorite, along with the year of discovery. Here is the link: <https://www.kaggle.com/code/kinguistics/exploring-meteorite-landing-sites/report>

The planned analyses / visualizations / coding challenges:

Planned analyses: How are the meteorites distributed across the world all years from 1900-2013? (map) How is the mass of meteorite related to type? (scatter plot) What is the number of meteorite landings change through years? (line plot)

Visualizations: Create an interactive map, a scatter plot displaying the relationship between meteorite mass and type, and a line plot illustrating changes in meteorite landings over the years.

Challenge: Given a dataset with a wide range of values, normalize the column using min-max scaling or log transformation to handle skewed mass values. Simplify plotting by combining low-frequency types into an "Other" category and reducing crowded type names for readability.

The planned timeline:

We plan to meet once or twice per week on Wednesday and Thursday afternoons to keep our project on track.

November 7: Finalize project topic, find a suitable dataset, and complete the proposal.

November 17: Complete coding and data analysis in RStudio, and begin the webpage.

November 26: Finish the project report.

December 5: Finalize all deliverables (report, webpage, video).

December 7: Submit peer assessments.