Fatimah Niyas
05.14.2025

Final Report
DATA205-33334

INTRODUCTION

Meteorites offer unique insights into both the history of our solar system and the geological processes of Earth. This project analyzes the [Meteorite Landings] dataset from Data.gov, which documents thousands of meteorite discoveries worldwide. The primary goal is to investigate how meteorite findings have changed over time, how technology and human activity have shaped detection, and whether geographic patterns exist in discovery trends.

To explore the influence of human presence on meteorite reporting, this analysis also incorporates global [Population Density] data from OurWorldInData.org. By comparing landing locations with population density, we aim to evaluate whether areas with more people are more likely to detect and report meteorites.

Data

The Meteorite Landings dataset includes key attributes such as: name and type of meteorite, mass (in grams), year of discovery, latitude and longitude, classification as "Fell" (witnessed) or "Found" (discovered later). To ensure quality analysis, the data was cleaned by: removing missing or invalid entries in critical fields (mass, year, coordinates), filtering data to meteorites findings from 1850s onward, excluding extreme outliers (mass > 1,000,000g or mass ≤ 0), and renaming variables for clarity (e.g., mass..g. to mass). The cleaned dataset contains over 30,000 observations and serves as the foundation for my exploratory and statistical analysis.

The Population Density dataset provides data on the global density (people per square kilometer), grouped by country and updated annually. It serves as a demographic layer in the analysis to investigate: whether densely populated areas report more meteorites and how human settlement patterns correlate with "Fell" vs. "Found" meteorite types.

Goals

The goal of this project is to analyze global meteorite landings to uncover patterns over time, explore how technology has influenced detection, and understand the geographic distribution of discoveries. By examining changes in the number and size of meteorites found since the late 19th century, the project looks at how scientific advancements and increased surveillance may have made it easier to detect smaller or previously overlooked meteorites. It also compares meteorite landings across different regions and population densities to see if more populated or technologically advanced areas report more findings. This analysis is important because it helps us understand how both natural events and human factors shape what we know about meteorites. In particular, it highlights potential gaps like regions where meteorites may fall but go unreported and biases where the data is skewed toward areas with more people or better technology, rather than reflecting where meteorites actually land.

Fatimah Niyas
05.14.2025

Final Report
DATA205-33334

Tools

Data Ingestion and Cleaning: R (tidyverse, dplyr) for cleaning up missing values, renaming variables, and more.
Exploratory Data Analysis: R, ggplot2 for visualizing trends and summarizing data.
Main Analysis: Hypothesis testing, correlation analysis, and modeling trends in R.
Visualization: ggplot2 for data visualizations and Tableau for interactive dashboards and geographic heatmaps.

Data Cleaning

In order to prepare the dataset for effective analysis, I implemented several key cleaning steps using R and the tidyverse package:

- Removed records with missing or invalid values in essential fields: year, mass, reclat (latitude), and reclong (longitude).
- Renamed columns for clarity; most notably, renaming mass..g. to mass.
- Filtered years to include only records from 1850 onward, aligning the analysis with the period of major technological advancements.

Applied geographic constraints, restricting coordinates to valid Earth-based ranges: latitude between -90 and 90, longitude between -180 and 180.

Excluded extreme mass values, filtering out meteorites with mass ≤ 0 or > 1,000,000 grams to reduce outlier bias.

Basic Descriptive Statistics

To better understand the patterns within the dataset, I began with a descriptive overview of the key variables. The cleaned dataset includes approximately 37,000 meteorite entries recorded between 1850 and 2013. Each entry contains several essential attributes that provide a foundation for analysis:

- Name and type of the meteorite
- Mass (in grams)
- Year of discovery
- Latitude and longitude of the landing location
- Status (either "Fell" for witnessed landings or "Found" for meteorites discovered later)
- Classification type, such as L6, H5, Iron, or CM2

Fatimah Niyas
05.14.2025

Final Report

DATA205-33334

The mass variable revealed a highly skewed distribution. While some meteorites weighed hundreds of kilograms, most were significantly smaller. The mean mass was approximately 2,400 grams, but the median mass was only 28.6 grams, reflecting a large number of small meteorites and the influence of a few very large ones. This informed decisions in later visualization steps, where log-scaling was sometimes applied to better represent the data.

The year variable was central to understanding trends over time. Meteorite counts were grouped by year to explore discovery rates, and I overlaid technological milestones to assess their potential influence on reporting frequency. This variable also served to create time series plots and detect patterns in discovery rates and average meteorite mass per year.

Latitude and longitude were used to plot meteorite landings geographically and compare them with global population density maps. These coordinates enabled the creation of heatmaps, choropleths, and regional breakdowns of "Fell" versus "Found" discoveries.

The status variable revealed a clear imbalance: approximately 93% of the meteorites were “Found”, while only 7% were “Fell.” This contrast was key to understanding the role of human observation in meteorite detection and reporting. For instance, "Fell" events tended to cluster near populated regions, suggesting that witness-based reporting is more common where people are present to observe the fall. "Found" meteorites were often discovered in more remote locations such as deserts, where preservation is ideal and search efforts are more deliberate.

Finally, a summary of entries by country showed that the United States, Australia, Libya, Argentina, and Russia had the highest counts. This geographic distribution helped identify global hotspots of meteorite discoveries and highlighted the potential influence of certain factors such as search infrastructure, and land area.

Altogether, this descriptive analysis laid the groundwork for deeper exploration into how meteorite discoveries have changed over time and how they relate to population, geography, and technology.

Final Data Product

My final data product consists of Tableau dashboards featuring three key interactive visualizations: a side-by-side geographic map comparing population density with meteorite landing locations (Figure1), a bar chart of annual meteorite counts with annotated technological milestones(Table 1), and a time series of average meteorite mass per year with a trend line.

Fatimah Niyas
05.14.2025

Final Report

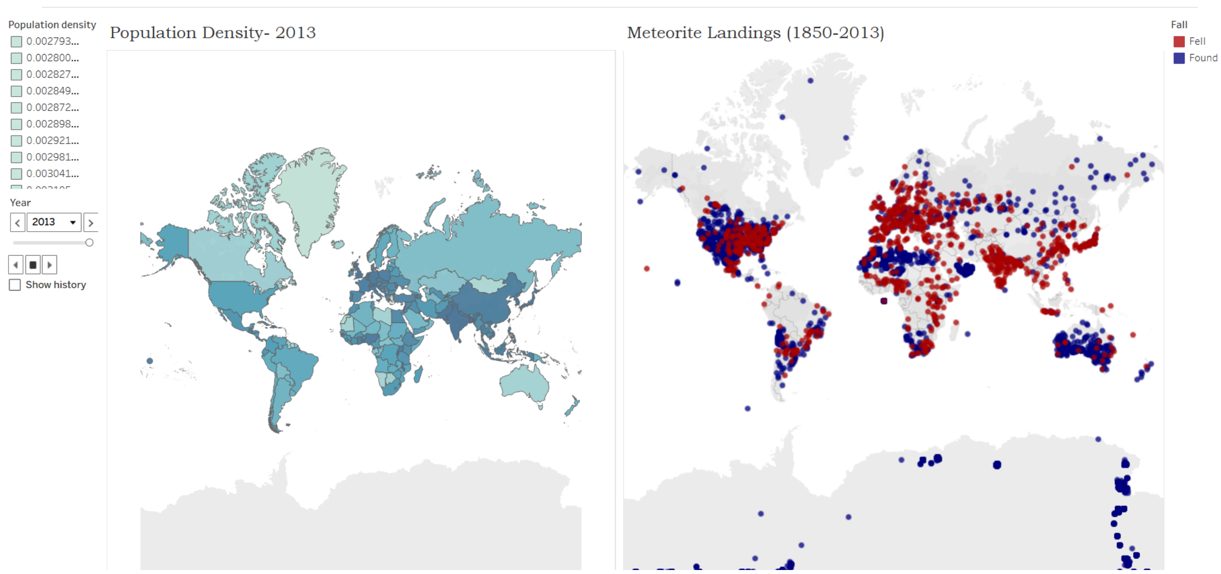
DATA205-33334

Together, these visualizations allow users to explore the influence of human activity on meteorite discovery.

Table 1- Technological Milestones and Meteorite Discovery Trends

Milestone Year	Technological Advancement	Observed Change
1858	Introduction of Photography	Rise in documented meteorite falls
1939	Radar use in WWII	Increased detection, especially during/after war
Post-1970s	Civilian access to military tech	Major spike in discoveries

Figure 1 - side by side geographic map of population density vs meteorite landings



The main results show that meteorite discoveries are not distributed randomly but are strongly influenced by where people live, the tools available to detect them, and the infrastructure for reporting such events. For instance, areas like Australia, despite low population density, show many "found" meteorites due to ideal environmental conditions for preservation and targeted search efforts. Meanwhile, the U.S. reveals patterns where meteorites are "found" in the less populated western regions and "fell" are more often reported in the more populated east. The bar chart visualization further showed that meteorite discoveries increased sharply following key technological developments, supporting the idea that better detection tools expand our capacity

Fatimah Niyas
05.14.2025

to observe space events. Finally, the average mass of discovered meteorites has decreased over time, which suggests an increasing ability to detect smaller meteorites thanks to improved search methods and technologies.

These results are valuable because they demonstrate how technological progress and human geography shape our understanding of meteorite activity. The dashboard provides a dynamic tool for educators, researchers, and science communicators to explore these insights, uncover patterns, and potentially identify gaps or regions for future investigation.

References and acknowledgements

This project incorporated data from multiple sources and benefited from various tools and supporting materials. The primary dataset, Meteorite Landings, was obtained from Data.gov, and the global population density data was from Our World in Data. Tableau Public was used for visualization and dashboard creation, allowing for dynamic and interactive exploration. Background context and supporting information were drawn from reliable articles such as meteororbs.org and University of Plymouth to better understand meteorite behavior, detection challenges, and the influence of human development on discovery patterns. These sources helped frame the analysis and interpret trends in a broader scientific and social context.

A very special thanks to Professor Perine for the guidance and feedback throughout the project. I would also like to thank Professor Saidi and Professor Alraee. Appreciation also goes to classmates for peer reviews.