Data Bootcamp

Project 1 Report

NASA Potentially Hazardous Objects (PHO)

Team 6

Contents

Introduction

DART was the first-ever mission dedicated to investigating and demonstrating one method of asteroid deflection by changing an asteroid's motion in space through kinetic impact. NASA's Double Asteroid Redirection Test (DART) investigation team showed the spacecraft's kinetic impact with its target asteroid, Dimorphos, successfully altered the asteroid's orbit. This marks humanity's first time purposely changing the motion of a celestial object and the first full-scale demonstration of asteroid deflection technology.

Potentially Hazardous Asteroids (PHAs) are currently defined based on parameters that measure the asteroid's potential to make threatening close approaches to the Earth. Specifically, all asteroids with a minimum orbit intersection distance (MOID) of 0.05 au or less and an absolute magnitude (H) of 22.0 or less are considered PHAs.

An astronomical unit is the average distance between the Earth and the Sun, meaning 0.05 au is approximately 7.5 million km.

The Absolute Magnitude is a measure of the apparent size of an asteroid or comet based on the luminosity of the celestial body on an inverse logarithmic scale, i.e., Venus has a magnitude of -5, Jupiter has a luminosity of -3, Pluto has a magnitude of 14, meanwhile the Near-Earth Objects (NEOs) have a magnitude above 22.

Null Hypothesis (H0):

There is no correlation between the distance of Near-Earth Objects (NEOs) from Earth and their apparent size. The danger presented by the NEOs, as observed from Earth, is not influenced entirely by their distance or apparent size.

Alternate Hypothesis (H1):

There is a correlation between the distance of Near-Earth Objects (NEOs) and their size. We expect to find a positive correlation between the distance of NEOs and their apparent size, indicating that NEOs in closer proximity to Earth will appear larger in the sky when observed from our vantage point, and therefore will be a potentially hazardous object.

To work on our hypothesis, we will be focusing on objects with a proximity equal or less than 0.05 au and a magnitude below 22.

Analysis

Code in .IPYNB file.

Data cleaning

Initially we determined that the dataset required cleaning:

- there were duplicate rows (same ID, same name) which were removed
- only the occurrence with minimum 'miss distance' was kept
- columns with no statistically relevant information were removed

Data viz

Pie charts

- Proportion of hazardous objects
- Proportion of hazardous objects within maximum miss distance range
- Proportion of hazardous objects within maximum magnitude range

Bar charts

- 10 closest (minimum miss distance) hazardous objects
- 10 brightest (minimum magnitude) hazardous objects

Scatter plot

- Clean dataset, miss distance vs magnitude

Box plot

- Clean dataset was divided and represented into three subsets related to miss distance (millions of km)
    - Close objects: (0, 2.5)
    - Intermediate objects: (2.5, 5)
    - Far objects: (5, ∞)
- 10 closest hazardous objects variation in miss distance
- Clean dataset variation in magnitude
- 10 brightest hazardous objects variation in magnitude

Statistics

Basic calculations

	miss_distance	absolute_magnitude
mean	1.770540e+07	23.361668
median	1.145114e+07	23.600000
var	2.972905e+14	8.236446
std	1.724212e+07	2.869921
sem	1.041197e+05	0.017331

Linear regression:

slope	intercept	rvalue	pvalue	stderr
-1.1685e-07	25.4305	-0.7020	0.0	7.1582e-10

For ANOVA, the segments followed box plot segments

p-value	Close mean	Intermediate mean	Far mean
0.0	26.7841	25.2168	22.2277

Conclusions

Confirmation of Hypothesis

The analysis conducted on the dataset supports our hypothesis, indicating a significant correlation between the distance of Near-Earth Objects (NEOs) from Earth and their apparent size. The data reveals a negative correlation, meaning that as the magnitude (apparent size) of NEOs decreases, their miss distance increases. This finding aligns with the expectation that NEOs in closer proximity to Earth appear larger in the sky when observed from our vantage point.

Identification of Potentially Hazardous Objects

Out of the entire dataset, 7.9% of the objects were identified as Potentially Hazardous Asteroids (PHAs) based on the criteria of a minimum orbit intersection distance (MOID) of 0.05 astronomical units (au) or less and an absolute magnitude (H) of 22.0 or less. This highlights the importance of continued monitoring and tracking of these objects to assess potential threats.

Magnitude and Distance Relationships

Further analysis reveals that 38.6% of hazardous objects are within the specified distance range, emphasizing the significance of monitoring NEOs within this proximity. Additionally, 32.7% of hazardous objects exhibit a magnitude below the hazardous threshold, emphasizing the importance of considering both distance and apparent size when categorizing PHAs.

Regression Analysis

The negative correlation observed in the regression analysis indicates that NEOs with smaller apparent sizes tend to have closer miss distances to Earth. The calculated *r-value = -0.7020* strengthens this correlation, suggesting a moderately strong inverse relationship between magnitude and miss distance.

It is important to highlight that we have observed this correlation at the scale of the full population, however when we test a smaller dataset, the relationship is lost.

We executed a statistical analysis using additional data like velocity and diameters of the NEOs; we performed the calculation of the mean, median, variance, SEM and ANOVA test.

None of those calculations brought additional statistically significant information. The standard deviation is too high, for instance:

- Miss distance standard deviation = 1.724212e+07
- Magnitude standard deviation = 2.869921

The decision was made to focus on regression analysis results.

Variance and Trends

The high variance observed in the box plots suggests a wide range of miss distances for NEOs with similar apparent sizes. This variability could explain the observed trend in the regression analysis despite the *r-value = 0.7020*, indicating a strong correlation but not a perfect one.

The presence of outliers, particularly in the farthest objects, contributes to the high variance in the dataset, however the nearest and the intermediate objects show a distribution without a significant number of outliers.

References (Retrieved October 2023):

Dataset

NASA - Nearest Earth Objects, A cumulative data for Nearest Earth Objects by NASA

https://www.kaggle.com/datasets/sameepvani/nasa-nearest-earth-objects/code

Definitions

PHA (Potentially Hazardous Asteroid)

https://cneos.jpl.nasa.gov/glossary/PHA.html

Potentially hazardous object

https://en.wikipedia.org/wiki/Potentially_hazardous_object

Double Asteroid Redirection Test (DART)

https://science.nasa.gov/mission/dart

NASA Confirms DART Mission Impact Changed Asteroid's Motion in Space

https://www.nasa.gov/news-release/nasa-confirms-dart-mission-impact-changed-asteroids-motion-in-space/