3

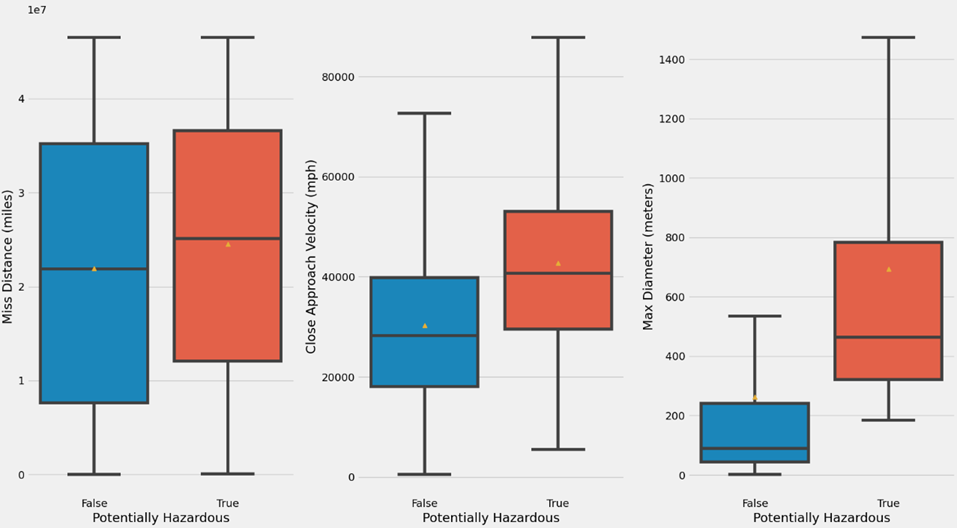
Project 1 Summary – Team 2

•Introduction  
Team 2 consisted of Brooke Robertson, Blake Stephenson, Nayoung Kim, and   
Ilse Styles. We initially considered several possible topics, but after discussion,   
agreed on the topic of asteroids, and the source API as NASA, which Brooke   
proposed, and then narrowed down our questions to: “What is the frequency of   
close encounters between asteroids and Earth?” and “What factors contribute to   
the potential danger of an asteroid?” We were motivated to learn more about   
this topic out of general interest as there have been many movies with asteroids   
heading toward earth and panic ensuing among the population. Also, there was   
personal interest. For example, Nayoung lived in Arizona previously and skied in  
the location of a crater.

•Process Summary  
We worked through the data finding, cleaning and analysis process together.   
The first steps involved reading in the data from a NASA API, creating a   
DataFrame from this information, and then creating a CSV from the DataFrame.   
Data cleaning included checking for null values (there were none, so we did not   
have to exclude them); and changing the data type of Distance variable from   
string to float to be able to use it for graphs and charts.

•After data cleaning, we began the analysis, which included writing codes to   
create a summary stats table, creating a correlation table and correlation plots   
among the 3 predictors in our model (velocity, diameter, distance), and running   
a binomial logistic regression model for our data since the dependent variable   
(potentially hazardous) is T/F. (Nayoung and her amazing modeling skills were   
responsible for the logistic regression model). We then created plots to show   
the relationship between predictors and potentially hazardous asteroids, and   
frequency of close encounters over time for both hazardous and nonhazardous   
asteroids.

•Question 1: What factors contribute to the potential danger of an asteroid?  
Our Findings:   
After the analysis described above, we found that the major factors that   
contributed to an asteroids classification as hazardous were velocity, diameter   
and miss distance from the Earth. Details are shown in the box plot graph   
below. Some asteroids can be very large, for example one named “2008 OS7”   
which is the size of a football stadium and made its closest pass ever past Earth   
on Friday Feb. 2, 2024. It may be stating the obvious to say that a larger   
asteroid at a higher velocity that actually hits the earth will do more damage,   
although of course there are other variables not accounted for in our analysis,   
such as the location and population density of the area of impact.



* *Question 2:* What is the frequency of close encounters between asteroids and Earth?

Our findings, which are illustrated in the graphs titled “Counts of Hazardous Asteroids vs. Year” and “Counts of Not Hazardous Asteroids vs. Year”, show that nonhazardous approaches are more frequent. In the first graph specifically, you can see that there was a period between 1900 and 1920 where the frequency of hazardous approaches shot up. A caveat to all our data, however, is that in order to have a more workable sample size, we used only the first pass data for each asteroid in the time period (an individual asteroid can have a number of passes around the earth).

A graph of a graph of a number of individuals

Description automatically generated with medium confidence

Implications and Lessons Learned

Although asteroids are a fear that exists in the public consciousness, the actual   
impact on our daily lives will be almost nonexistent compared to other natural   
events such as flooding, tornados, hurricanes, and earthquakes.

We enjoyed this process and learned a number of things regarding the topic and   
working on the data and analysis, one of the most memorable being “if you   
make a CSV file from an API, always save the CSV file separately!”