PROJECT REPORT

Our journey began in the digital realm, where we set up a cozy corner in Python, inviting in some essential libraries like old friends: requests for whispering through the web, json for unraveling data mysteries, and datetime for weaving through the moments.

The data collection process begins with setting up a Python environment and importing necessary libraries: requests for HTTP requests, json for parsing JSON data, and datetime for handling dates. A unique API key is required to authenticate requests to NASA's API.

We crafted a magical spell, a function, fetch_neo_data, is defined to make requests to the NeoWS API for a specified week. This enchantment takes start_date, end_date, and api_key as parameters, constructs the request URL, and fetches the data. If the request is successful (HTTP status code 200), we were rewarded with a JSON scroll of knowledge; otherwise, the spell would sadly whisper an error.

Given the API's limit on the date range for each request, the entire desired date range (in this case, one year from 2023-02-06 to 2024-02-06) is divided into weekly segments. A while loop iterates through the year, week by week, calling fetch_neo_data for each segment and appending the result to the all_data list.

Once we gathered our digital treasures, we tucked them away into a JSON chest, labeling it 'neo_data_year.json' to keep our findings safe and sound for further adventures.

After collecting the data, it is saved into a JSON file named neo_data_year.json for temporary storage. This step ensures that data is not lost and provides a checkpoint for further processing.

Our next chapter involved transforming our collected tales into a CSV scroll, carefully selecting columns to paint the full picture of each Near Earth Object's story.

We then embarked on a quest to loading the collected JSON data from neo_data_year.json and preparing it for conversion into a CSV format.

We chose our heroes, the column names, with care to represent the attributes of interest for each NEO, such as the name, ID, estimated diameter, velocity, miss distance, and potential hazard status. These columns are selected based on their relevance to the project goals and the information available in the JSON structure.

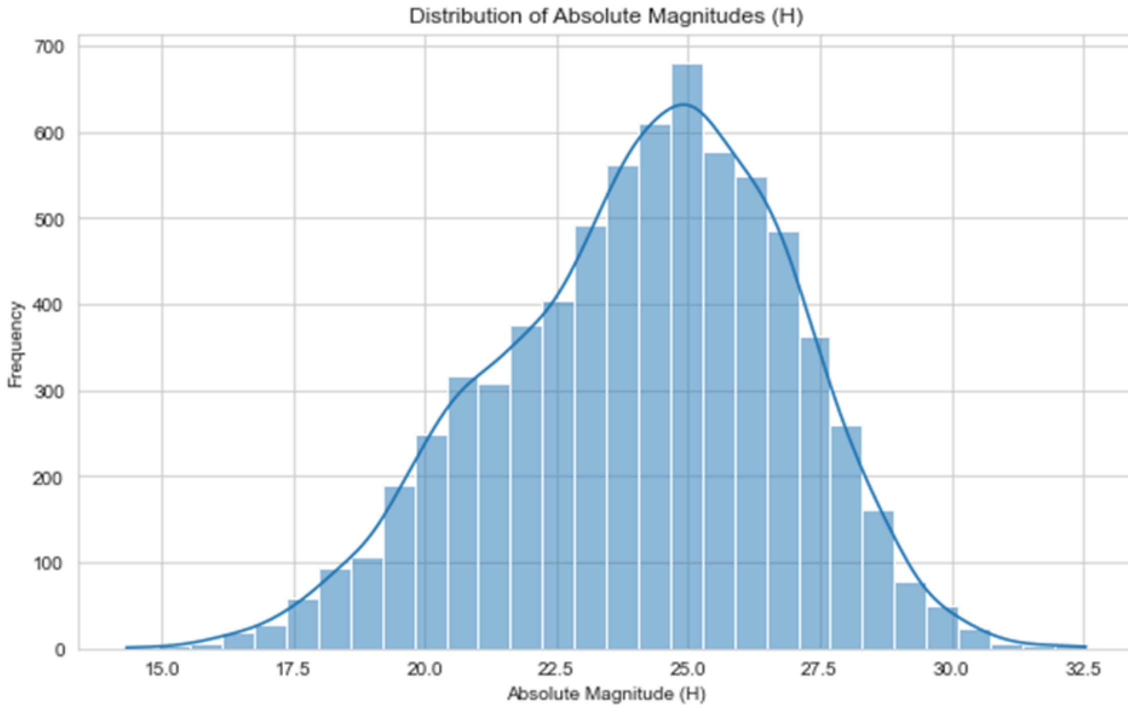
The Python csv library is used to write the data into a CSV file named neo_data_year.csv. For each NEO entry, a dictionary is created with keys corresponding to the CSV column names. The .get() method is utilized to safely access each attribute within the nested JSON structure, with 'N/A' as the default value if any attribute is missing.

In our quest, we encountered 7071 celestial wanderers, each with its own unique story, illuminated by the data like stars in the night sky.

The dataset comprises 7,071 Near Earth Objects (NEOs), each identified by a unique ID and name. These objects have been observed on specific dates, with their properties such as absolute magnitude, estimated diameter, velocity, and miss distance recorded. The dataset also indicates whether an object is considered potentially hazardous.

Distribution of Absolute Magnitudes

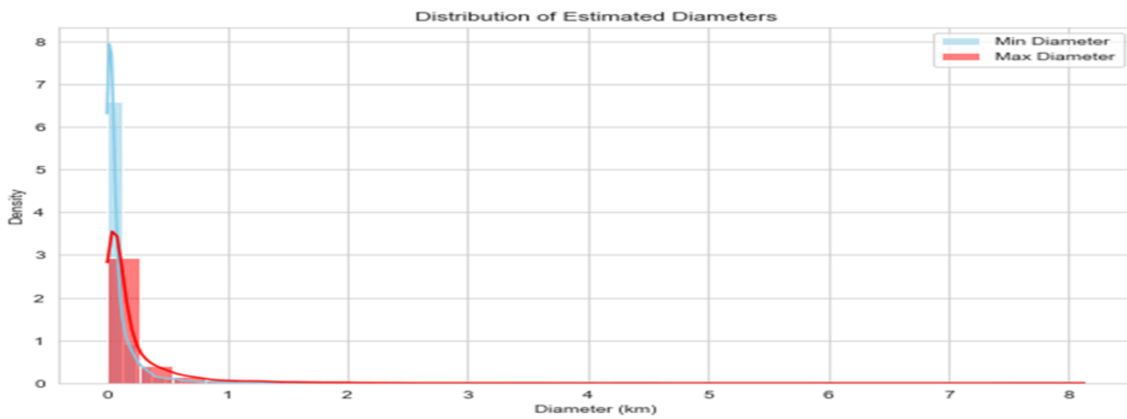
Absolute Magnitude (H) is a measure of an object's brightness. A lower value indicates a brighter and potentially larger object.



The distribution of Absolute Magnitudes (H) shows a wide range, with a concentration of objects having magnitudes between approximately 22 and 26. This suggests that the dataset includes a variety of NEOs, from relatively bright (lower H values) to dimmer objects (higher H values).

Estimated Diameter Analysis

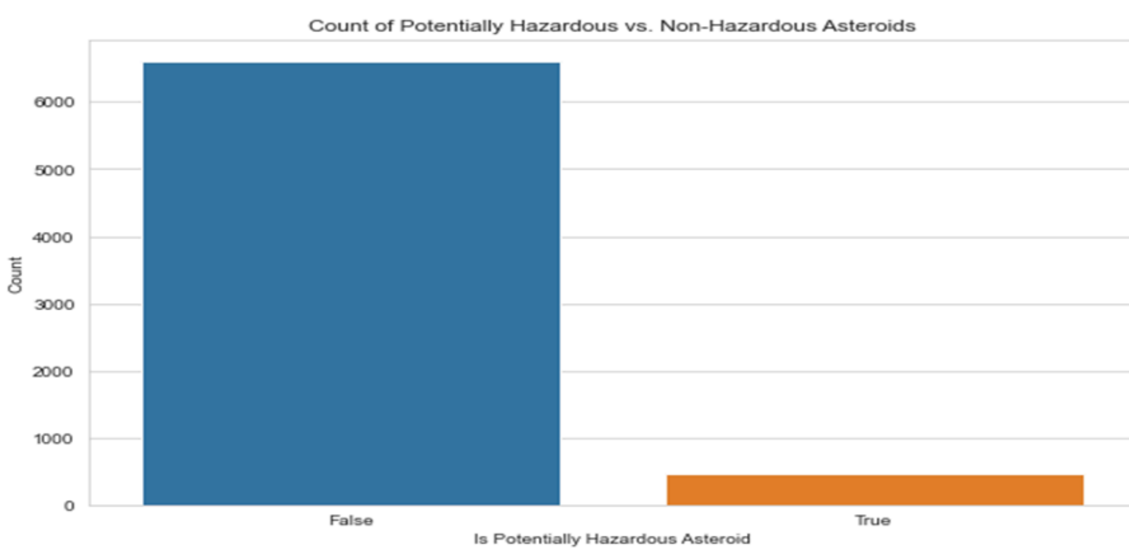
The estimated diameter of NEOs, given in kilometers, provides insight into the size of these objects. The distribution of estimated diameters, both minimum and maximum estimates, reveals that most NEOs are relatively small, with diameters less than 1 km. There's a noticeable skew towards smaller objects, indicating that larger NEOs are less common. The overlap in the density plots for minimum and maximum estimated diameters suggests consistency in size estimation but highlights the range of uncertainty in these measurements.



Velocity and Miss Distance

The relative velocity of NEOs in km/h and their miss distance, both in astronomical units and kilometers, will be analyzed to understand the dynamics of their close approaches to Earth. High velocities and small miss distances may indicate higher risk scenarios.

Potentially Hazardous Asteroids Analysis



The count plot of potentially hazardous versus non-hazardous asteroids shows that a smaller portion of the dataset is classified as potentially hazardous. This indicates that while many NEOs come close to Earth, only a subset is considered to pose a significant risk based on their size and close approach distance.

NEO Analytics Dashboard Report

This dashboard provides a comprehensive overview of Near-Earth Objects (NEOs) based on data from . It includes visualizations of NEO sizes, velocities, distances from Earth, and their potential hazard status, allowing for a detailed analysis over time. The dashboard's interactivity enables users to filter results by date, offering tailored insights.

Our saga culminated in the creation of a magical dashboard, a window to the cosmos, where NEOs' sizes, velocities, and tales of potential hazards danced before our eyes.

Gauge Chart: NEO Average Maximum Diameter

Purpose: This gauge chart represents the average maximum diameter of observed NEOs, giving an immediate sense of the potential size range of these objects.

Creation Process: The chart was created using Power BI's native gauge chart feature, with a range set from 0 to 10 to reflect the observed data. Colors were chosen to represent different thresholds of NEO sizes.

Data Utilized: Data for NEO diameter was pulled from the 'Estimated Diameter Max (km)' column.

Measures and Calculations: A DAX measure was created to calculate the average diameter:

Average Diameter = AVERAGE('NEO_Data'[Estimated Diameter Max (km)])

Interactivity: Users can adjust the date range to see how the average diameter changes over time.

Line Chart: Proportion of Potentially Hazardous NEOs Over Time

Purpose: This line chart tracks the proportion of NEOs classified as potentially hazardous, highlighting trends and potential increases in risk levels.

Creation Process: The line chart was developed using Power BI's line chart visualization, plotting the proportion of hazardous NEOs over the selected time frame.

Data Utilized: The 'Is Potentially Hazardous Asteroid' field was used, along with the date of observation.

Measures and Calculations: A calculated column was created to determine the hazardous status:

Hazardous Status = IF('NEO_Data'[Is Potentially Hazardous Asteroid] = TRUE, "Hazardous", "Non-Hazardous")

Interactivity: The chart responds to the date slicer, allowing users to explore different periods.

Area Chart: Cumulative NEO Observations

Purpose: The area chart displays the cumulative count of NEOs observed over time, showcasing the total number of NEOs detected and how this count has increased.

Creation Process: Constructed with Power BI's area chart tool, this visualization accumulates the total number of NEOs day by day.

Data Utilized: Each NEO's observation date from the 'Date' column was used to build this visualization.

Measures and Calculations: The cumulative count was calculated using a DAX measure:

```
Cumulative Count =
CALCULATE(
  COUNT('NEO_Data'[ID]),
   FILTER(
       ALL('NEO_Data'[Date]),
       'NEO_Data'[Date] <= MAX('NEO_Data'[Date])
   )
)
```

Interactivity: The area chart is interactive with the date slicer, allowing the user to adjust the viewed time range and see how the cumulative count has evolved.

Scatter Plot: NEO Size vs. Miss Distance Correlation

Purpose: This scatter plot correlates the size of NEOs with their miss distance from Earth, offering insights into the potential risk based on these two critical factors.

Creation Process: Implemented using Power BI's scatter chart functionality, this plot maps each NEO's size against its miss distance.

Data Utilized: 'Estimated Diameter Max (km)' for size and 'Miss Distance (kilometers)' for distance.

Measures and Calculations: Additional measures were not necessary for this visualization, as each data point represents an individual NEO's recorded size and miss distance.

Interactivity: Users can filter the scatter plot by the hazardous status of NEOs and by time period using slicers.

Bar Chart: Velocity Distribution of NEOs

Purpose: The bar chart visualizes the distribution of NEO velocities, categorized into bins, to depict the common velocity ranges of these objects.

Creation Process: Created with Power BI's bar chart feature, the velocities were grouped into bins for a histogram-like representation.

Data Utilized: Velocity data from the 'Relative Velocity (km/h)' column was used to categorize NEOs into velocity bins.

Measures and Calculations: The velocity bins were created using a calculated column in Power BI:

```
Velocity Bin =
SWITCH(
    TRUE(),
    'NEO_Data'[Relative Velocity (km/h)] <= 10000, "0-10,000 km/h",
    'NEO_Data'[Relative Velocity (km/h)] <= 20000, "10,001-20,000 km/h",
```

```
    ...
    "50,001+ km/h"
)
```

Interactivity: The bar chart responds to the date slicer, enabling the user to see how velocity distribution changes over selected periods.