Cutting and Sorting Data with the Pandas Library: Asteroid Statistics

1. Go to the dataset linked below and create a new notebook: https://www.kaggle.com/datasets/austinhinkel/asteroid-orbital-elements-and-near-earth-object-tag

2. Read in the data into a pandas DataFrame and use the .head() method to peak at the data. Use the len() function to check how many asteroids are in the DataFrame. How many asteroids do you have?

3. Create a new column called 'aphelion' to store the aphelion distance and calculate it using the following equation: $r_a = 2a - r_p$, where a is the semi-major axis length and r_p is the perihelion distance.

4. Use the .loc[] method for your to obtain only asteroids with perihelion distances less than 1 AU AND aphelion distances greater than 1 AU and create a new DataFrame called slice1. If an asteroid gets both closer to the Sun than the Earth and goes further from the Sun than the Earth, it must cross a point 1 AU from the Sun – but that's where Earth orbits! How many asteroids are in your slice1 DataFrame?

5. What are the three highest aphelion distances in the dataset? What are the diameters of the objects with these large aphelion distances?

6. What is the minimum perihelion distance in the dataset?

7. Among *only* asteroids classified as a Near-Earth Object (neo == Y), find the mean diameter.

8. How many asteroids in the dataset have a diameter greater than 1 kilometer?

9. What is the mean orbital period in the dataset?