Sabneet Bains

Data Visualization – EN.605.662

Project #3: Interactive Visualization using Tableau

Data Description

Near-Earth objects or NEOs are neither a planet, a dwarf planet, nor natural satellite, and the International Astronomical Union states, "All other objects, except satellites, orbiting the Sun shall be referred to collectively as 'Small Solar System Bodies'." [1] Since 1998, NASA and other world space agencies have collected data on NEOs under project Spaceguard. Thus, one area of interest that piqued my attention from this massive amount of data is studying NEOs that have close approaches to Earth, i.e., Earth-grazing meteoroids.

NASA's Jet Propulsion Laboratory at Caltech keeps a public record of NEOs' close approaches to the Earth from 1900 CE to 2200 CE. [2] The data is available to download as a CSV, Excel file, or a tabular print of the website table.

I utilized the Excel option and looked at the data from the past year (2021-Jun-27 to 2022-Jun-27). With trivial reformatting, I corrected the faulty units from "Km to m" or vice versa and fragmented the "Diameter" ranges into two discrete attributes defining the estimated minimum and maximum diameters. Likewise, the "Date" and "Time" were split for more accessible statistical analyses.

[1] https://solarsystem.nasa.gov/planets/in-depth/
[2] https://cneos.jpl.nasa.gov/ca/

Thus, the data stands at 11 variables x 1560 records. With individual variable descriptions as follows,

Variable	***Description***
Object	Primary designation
Close-Approach (CA) Date	Date of closest Earth approach
Close-Approach (CA) Time	Time of closest Earth approach
CA Distance Nominal	The most likely close-approach distance
CA Distance Minimum	The minimum possible close-approach distance
V relative	Object velocity relative to Earth
V infinity	Object velocity relative to a massless Earth
H	Asteroid absolute magnitude
Diameter Minimum	Estimated minimum diameter
Diameter Maximum	Estimated maximum diameter
Rarity	0 means an average frequency of 100 per year, 2 is roughly once a year, 3 is roughly once a decade

Table 1*:* Table column variable descriptions

Data Exploration

Variable	*Data Type*	*Mean*	*Standard Error*	*Median*	*Mode*	*Standard Deviation*	*Skewness*	*Minimum*	*Maximum*
Object	Nominal	-	-	-	-	-	-	-	-
Close-Approach (CA) Date	Interval	2021-Dec-24	-	2021-Dec-10	2021-Nov-03	-	-	2021-Jun-27	2022-Jun-26
Close-Approach (CA) Time	Interval	12:12	-	12:19	17:55	-	-	0:00	23:59
CA Distance Nominal	Nominal	3070159.91	54534.75	2801164.00	-	2153950.22	0.39	9427.00	7471513.00
CA Distance Minimum	Ratio	3017477.06	53784.61	2761558.00	-	2124322.07	0.40	1037.00	7439811.00
V relative	Ratio	10.62	0.12	9.68	8.47	4.82	1.02	0.65	35.28
V infinity	Ratio	10.57	0.12	9.62	5.65	4.83	1.02	0.27	35.28
H	Ratio	26.27	0.04	26.30	26.20	1.76	-0.51	16.60	31.80
Diameter Minimum	Ratio	0.02	0.00	0.02	0.01	0.06	18.41	0.00	1.80
Diameter Maximum	Ratio	0.05	0.00	0.03	0.02	0.10	12.10	0.00	1.80
Rarity	Ordinal /Discrete	0.12	0.01	0.00	0.00	0.37	3.31	0.00	3.00

Table 2*:* A subset of the statistical analysis that properly fits the paper aspect ratio is given here; for a more holistic analysis, please refer to the included Excel file.

Based on Table 2, it is evident that different statistical descriptors are possible/practical for specific data types; for example, it does not make much sense to find the mean of the ”Object” names as they are purely nominal. However, for something like “V relative,” finding the minimum and maximum provides the much-needed range for further calculations or visualizations.

Analytical Questions

Initially, I had the following five analytical questions that I wanted to be answered visually:

1. *Which Near-Earth objects had the closest approaches in the past year, and what were their relative sizes?*
2. *Which Near-Earth object had the closest approaches in the past year, and at what date and time?*
3. *How many rare close approaches were recorded in the past year?*
4. *Which Near-Earth objects have the fastest relative velocities?*
5. *Rank the hazardous Near-Earth objects by size that had close approaches in the past year?*

Visualization Techniques

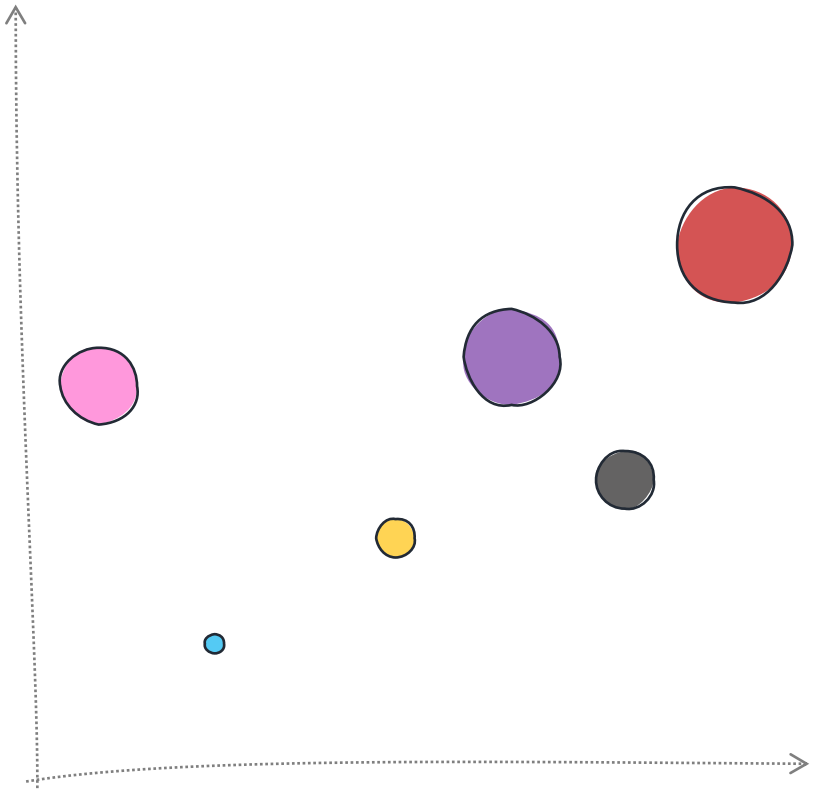
After further analysis, I wanted to visualize the following three analytical questions:

- *Which Near-Earth objects had the closest approaches in the past year, and what were their relative sizes?*
- *How many rare close approaches were recorded in the past year?*

- *Rank the hazardous Near-Earth objects by size that had close approaches in the past year?*

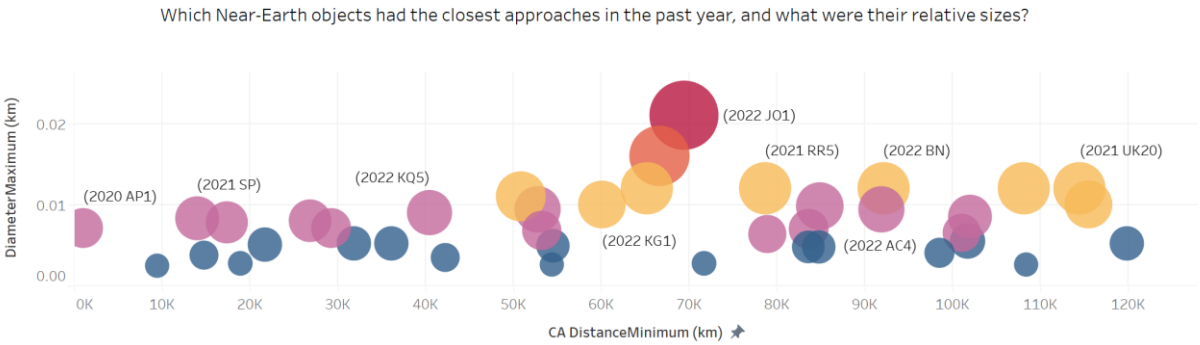
⊕ **To answer**: *Which Near-Earth objects had the closest approaches in the past year, and what were their relative sizes?*

I had proposed the following bubble chart in Project 2:



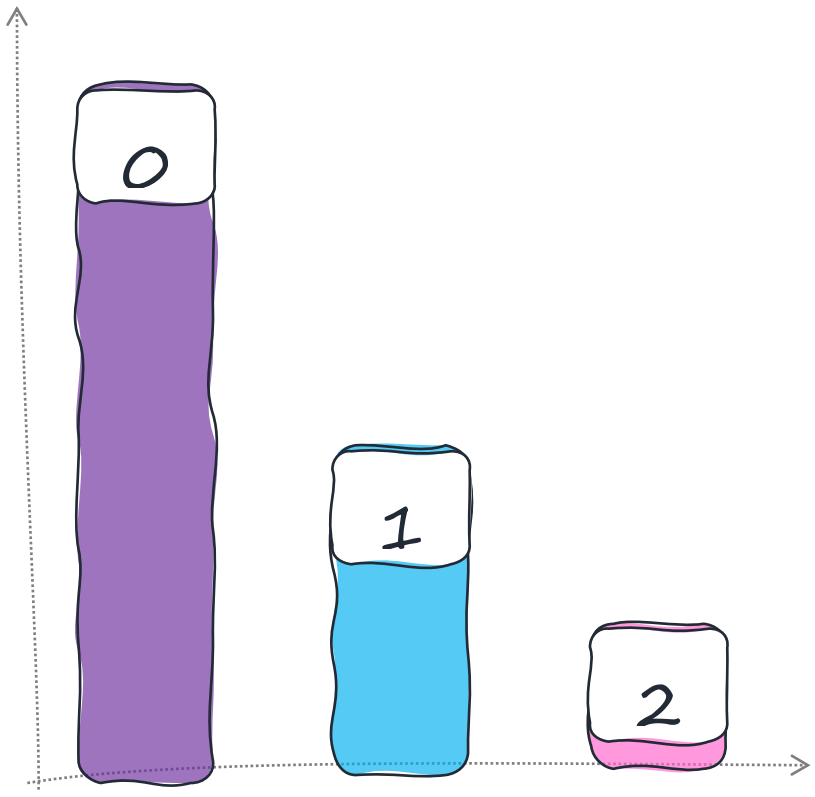
"A bubble chart also depicts the position of each NEO compared to the origin (Earth), and if the scope of the question included relative sizes, then that could be answered immediately here as well."

Analogously in Tableau, I came up with the following visualization:



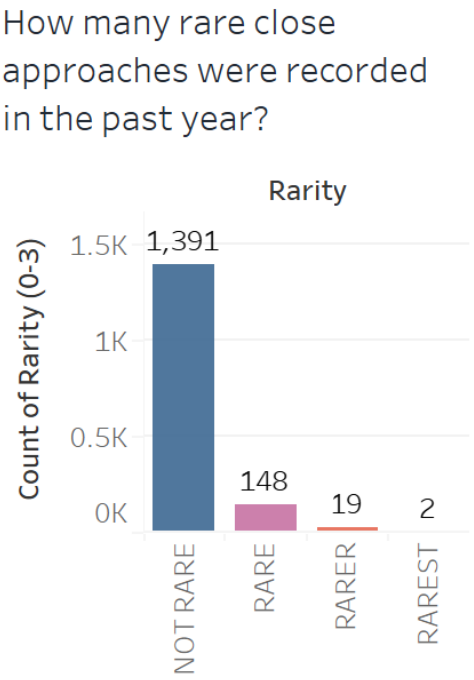
⊕ **To answer**: *How many rare close approaches were recorded in the past year?*

I had proposed the following histogram in Project 2:



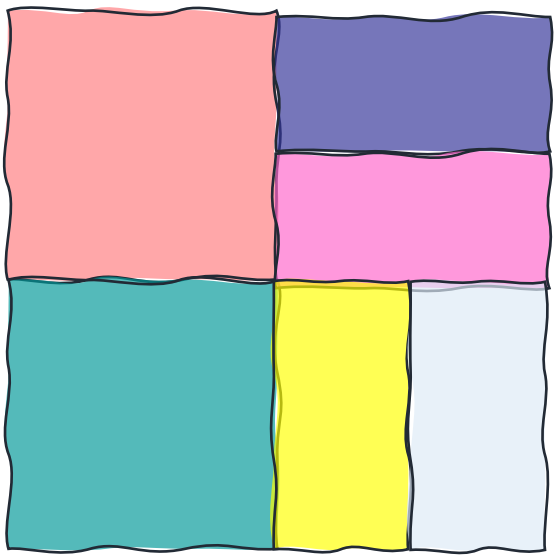
"A simple histogram best depicts each category and their respective counts. One can quickly realize if there were any ultrarare close approaches in the past year or not. This is indeed extremely powerful and simple to construct at the same time."

Analogously in Tableau, I came up with the following visualization:



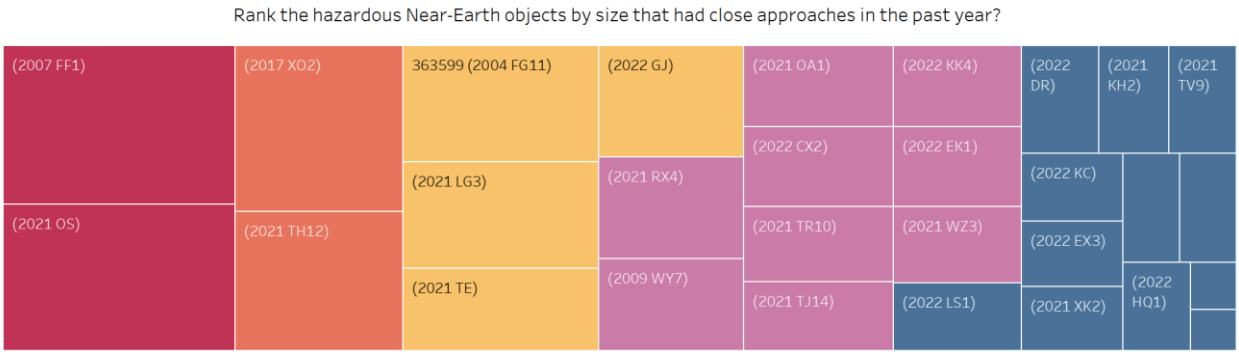
- **To answer**: *Rank the hazardous Near-Earth objects by size that had close approaches in the past year?*

I had proposed the following area graph in Project 2:



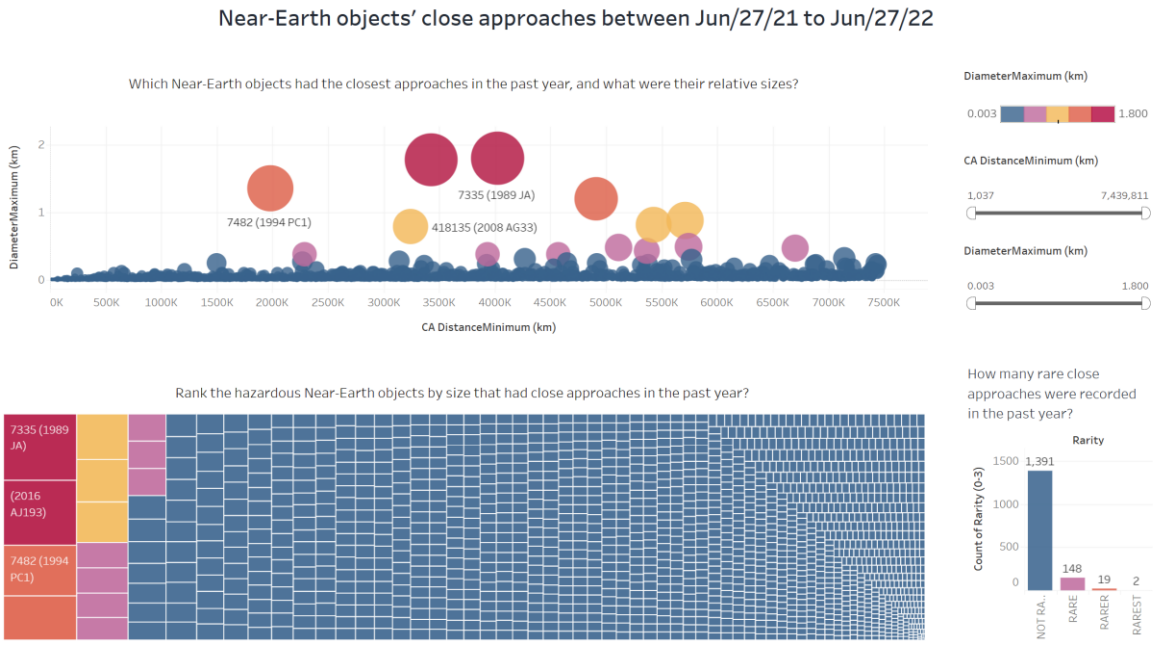
“Another useful visualization would be an area graph to quickly scan the sizes of the objects, again limited to a lower count, like the top 10.”

Analogously in Tableau, I came up with the following visualization:



Conclusion

The ultimate dashboard ended up looking something like this:



In fact, I was able to answer all my questions as I could quickly tell from the bubble chart which Near-Earth objects had the closest approaches in the past year and what were their relative sizes while manipulating the global filters to alter the minimum distance range or even the diameter of the object.

From the histogram, it was evident how many rare close approaches were recorded in the past year and at what degree (see variable description for rarity)

Likewise, the area chart ranked the hazardous Near-Earth objects by size that had close approaches in the past year, while the global filters provided interactivity to drill down the data. Moreover, each chart acts as a filter for the rest, as clicking on one element of a particular chart filters all the charts accordingly.