EXCEL PROJECT REPORT

TOPIC: NEAR-EARTH OBJECT ANALYSIS: ASTEROID DETECTION AND HAZARDOUS OBJECT MONITORING, RISKS & TRENDS

NAME: SAYANDIP GHOSH

REGISTER NUMBER: 2448008

SUBJECT NAME: APPLIED EXCEL

SUBJECT CODE: MDS151

INTRODUCTION

There are too many different types of materials in space, like asteroids, non-identifying space objects, nearest earth objects, fallen matter, and meteors, but NEO (Nearest Earth Objects) are potentially dangerous to the Earth's Surface. NASA flags them as hazardous objects in the Earth's Orbit. Due to the gravitational pull of the other planets, often asteroids and meteors, comets come closer to the Earth. Our Project, **"NEAR-EARTH OBJECT ANALYSIS: ASTEROID DETECTION AND HAZARDOUS OBJECT MONITORING, RISKS & TRENDS,"** refers to the comprehensive analysis of these celestial wanders from NASA's dataset; we aim to find the insights to uncover the deep understanding about NEO's. The analysis encloses various patterns for asteroid detection, its characteristics, and potential risk for the earth's orbit. From this analysis, we can make many powerful key points like NEO's dynamics, Hazardous object monitoring, Orbiting body influences etc.

DATA DESCRIPTION

Attached is a CSV file containing NEO's observation data from the '90s to 2022, carrying nearly 90,000 plus rows and nine columns, representing data for 27375 unique NEO objects with multiple and more times observation. The columns in the dataset are as follows:

1. **ID:** Denoting each object.
2. **Name:** Name of the objects.
3. **Estimated Diameter Max:** Estimated minimum diameter.
4. **Estimated Diameter Min:** Estimated maximum diameter.
5. **Relative Velocity:** Relative velocities of Neo's.
6. **Hazardous:** The object is risky or not.
7. **Miss Distance**: How much distance it misses from the Earth's surface.

8. **<u>Sentry Object:</u>** The object is sentry or not.
9. **<u>Orbiting Body:</u>** In which body's orbit it enters.
10. **<u>Absolute Magnitude:</u>** The absolute magnitudes of Neo's

OBJECTIVE

Analyze the data to identify key factors that help denote whether the space object is risky for Earth, along with the seasonal trend of coming of asteroids towards Earth's orbit and characteristics of each NEO to predict the risk impact for the future.

METHODOLOGY

1. DATA COLLECTION AND CLEANING:

Based on the objective, I collect the data from the Kaggle library, import it into Excel through the Data option, and then load CSV. In the data, each number is observed multiple times, so in the data cleaning process, I had to remove the duplicates first, and then I did the actual alignment for text and numbers that were not in the data. I also removed unnecessary columns like orbiting bodies because all objects concerning the Earth and Absolute magnitudes are observed.

2. DATA VALIDATION:

When checking for null values, I don't get any null values in any columns using the filter. There were many repeated rows, so I only took unique values. All columns contain text values and numerical values according to the observed entries. However, the text value was not in text format for the hazardous and sentry object cases, so I had to correct the alignment first. There are no data entry errors like formula or text errors. There are no logical outliers in the data. All data ranges are in the expected range.

3. SUMMARY STATISTICS:

The summary statistics of my dataset are given below:

MEAN: For diameter min and max, the mean is 0.1390,49 and 0.310099, the relative velocity and missed distance is 49280km/h, and 33139285 km denotes average values for each individual.

MEDIAN: For diameter min and max, the median is 0.0500 and 0.1132, the relative velocity and missed distance is 45144km/h, and 31366346 km denotes the median values for each individual.

MODE: For diameter min and max, the mean is 0.029144 and 0.0651688, the relative velocity and missed distance is 36473km/h, and 70019515km denotes mode values for each individual.

MIN: For diameter min and max, the mean is 0.006089013 and 0.00136157, the relative velocity and missed distance is 516km/h, and 9316km denotes minimum values for each individual.

MAX: For diameter min and max, the mean is 37 and 84, the relative velocity and missed distance is 236990km/h, and 74798651km denotes maximum values for each individual.

STANDARD DEVIATION: For diameter min and max, the mean is 0.374420271and 0.837229178, the relative velocity and missed distance is 26258km/h, and 23562423km denotes standard deviation values for each individual.

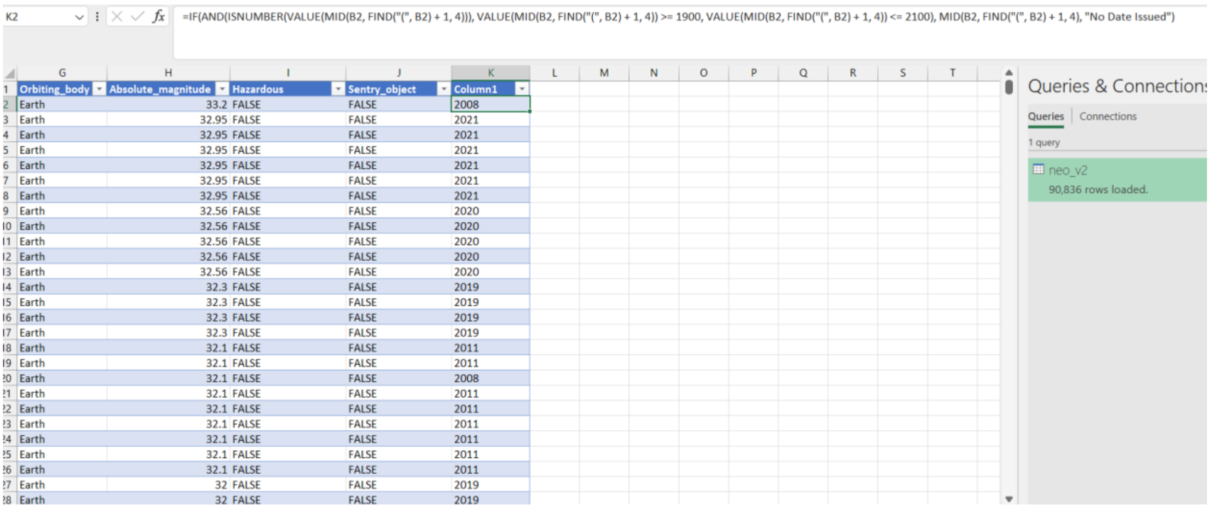
VARIANCE: For diameter min and max, the mean is 0.140190539 and 0.700952696, the relative velocity and missed distance is 689518717km/h, and 5.51km denotes variance values for each individual.

Statistic	Est_diameter_min	Est_diameter_max	Relative_velocity	Miss_distance
Mean	0.139081658	0.310996041	49280.64801	33139285.09
Median	0.050647146	0.113250461	45144.89523	31366348.18
Mode	0.029144391	0.065168838	36473.19106	70019515.65
Minimum	0.000608913	0.00136157	561.6956016	9316.925424
Maximum	37.89264984	84.73054089	236990.1281	74798651.45
Standard Deviation	0.374420271	0.837229178	26258.68841	23562423.42
Variance	0.140190539	0.700952696	689518717	5.55188E+14

4. ANALYSIS, INTERPRETATION AND VISUALIZATION:

1. SIZE AND VELOCITY ANALYSIS:

Concerning the sentry object and orbiting body, when I analyze year-wise discovery over the neo-objects, the size vs velocity graph shows that when the size increases, the velocity also increases. For this analysis, first, I have to find the years from the discovery object's name by using this formula,



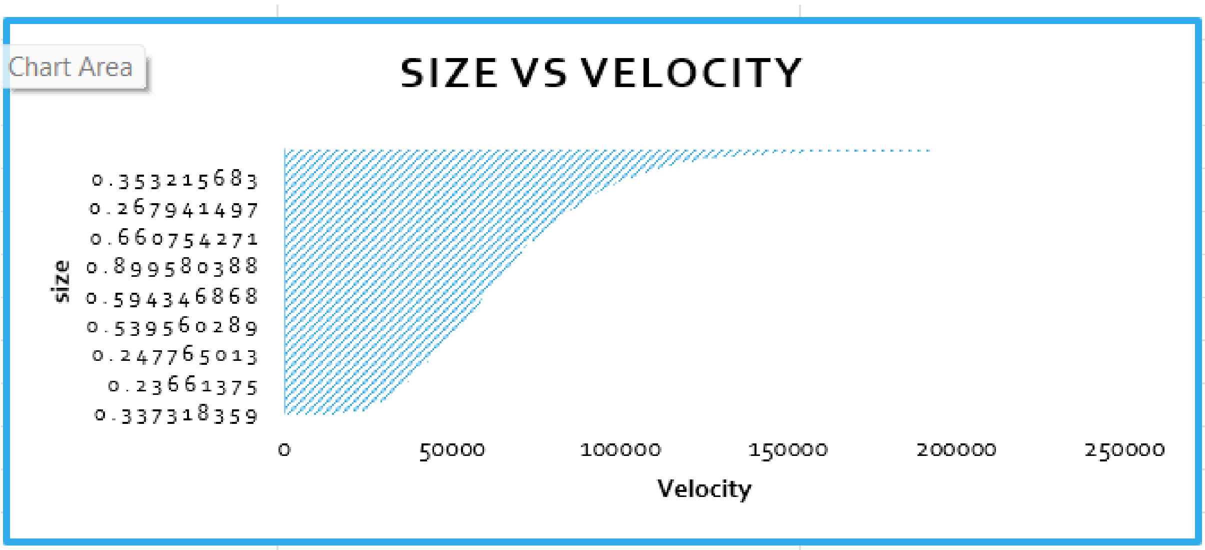
K2

```
=IF(AND(ISNUMBER(VALUE(MID(B2, FIND("(", B2) + 1, 4))), VALUE(MID(B2, FIND("(", B2) + 1, 4)) >= 1900, VALUE(MID(B2, FIND("(", B2) + 1, 4)) <= 2100), MID(B2, FIND("(", B2) + 1, 4), "No Date Issued")
```

	G	H	I	J	K
1	Orbiting_body	Absolute_magnitude	Hazardous	Sentry_object	Column1
2	Earth	33.2	FALSE	FALSE	2008
3	Earth	32.95	FALSE	FALSE	2021
4	Earth	32.95	FALSE	FALSE	2021
5	Earth	32.95	FALSE	FALSE	2021
6	Earth	32.95	FALSE	FALSE	2021
7	Earth	32.95	FALSE	FALSE	2021
8	Earth	32.95	FALSE	FALSE	2021
9	Earth	32.56	FALSE	FALSE	2020
10	Earth	32.56	FALSE	FALSE	2020
11	Earth	32.56	FALSE	FALSE	2020
12	Earth	32.56	FALSE	FALSE	2020
13	Earth	32.56	FALSE	FALSE	2020
14	Earth	32.3	FALSE	FALSE	2019
15	Earth	32.3	FALSE	FALSE	2019
16	Earth	32.3	FALSE	FALSE	2019
17	Earth	32.3	FALSE	FALSE	2019
18	Earth	32.1	FALSE	FALSE	2011
19	Earth	32.1	FALSE	FALSE	2011
20	Earth	32.1	FALSE	FALSE	2008
21	Earth	32.1	FALSE	FALSE	2011
22	Earth	32.1	FALSE	FALSE	2011
23	Earth	32.1	FALSE	FALSE	2011
24	Earth	32.1	FALSE	FALSE	2011
25	Earth	32.1	FALSE	FALSE	2011
26	Earth	32.1	FALSE	FALSE	2011
27	Earth	32	FALSE	FALSE	2019
28	Earth	32	FALSE	FALSE	2019

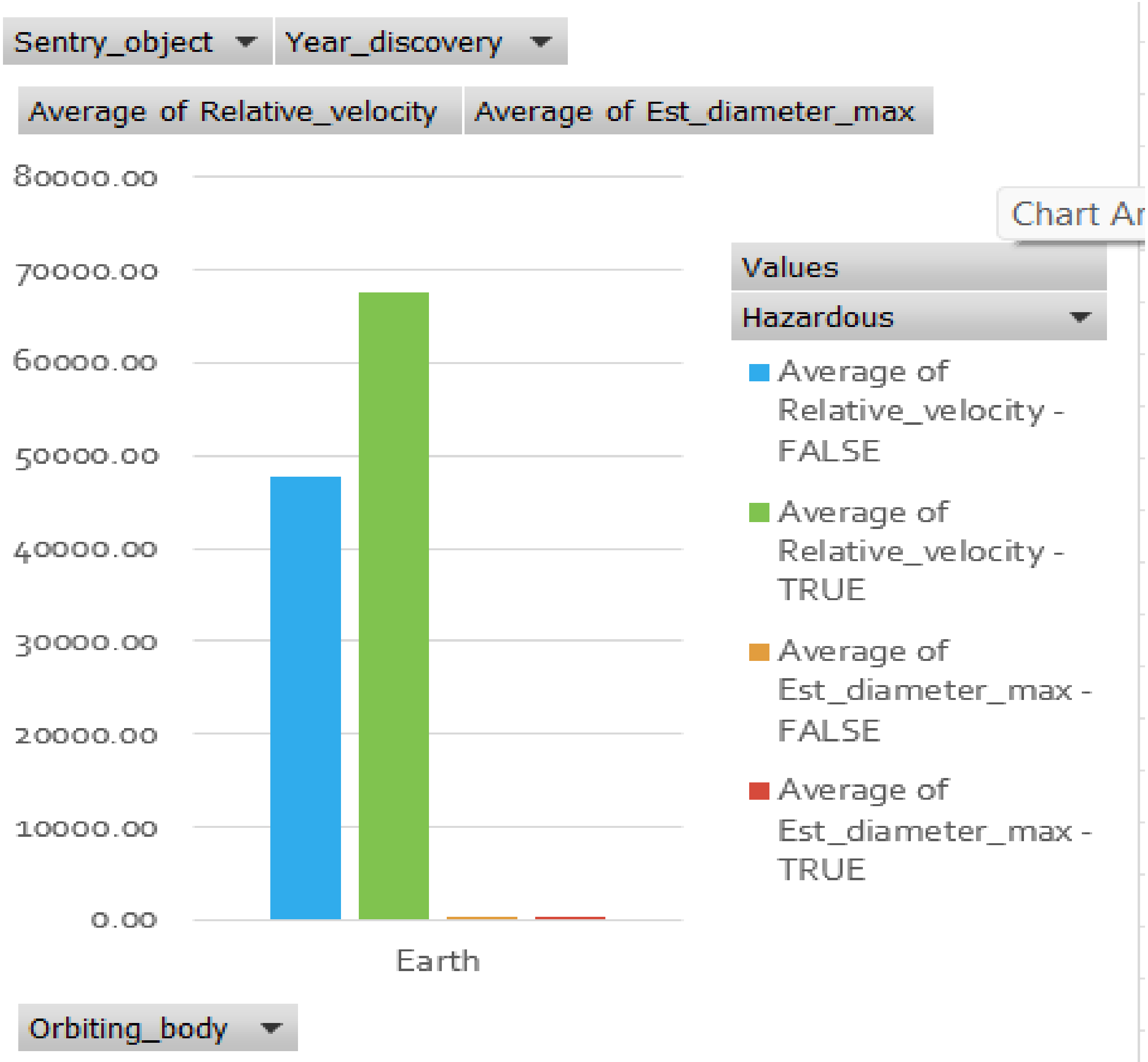
Queries & Connections
Queries | Connections
1 query
neo_v2
90,836 rows loaded.

After finding the years, I analyzed the velocity and size in detail. I got the below graph, which gives the proportional relation analyzing the pivot table; I get that the average relative velocity and estimated diameter are 47707km/h, and 0.277024 km is the signal line, which refers to the hazardous level and risk potentiality for the new objects. Less than it, this value remarks on safety while over than denotes the risk index. For example, if the asteroid comes with 67565 km/h and is 0.70574 km, it becomes a hazardous object for the Earth.

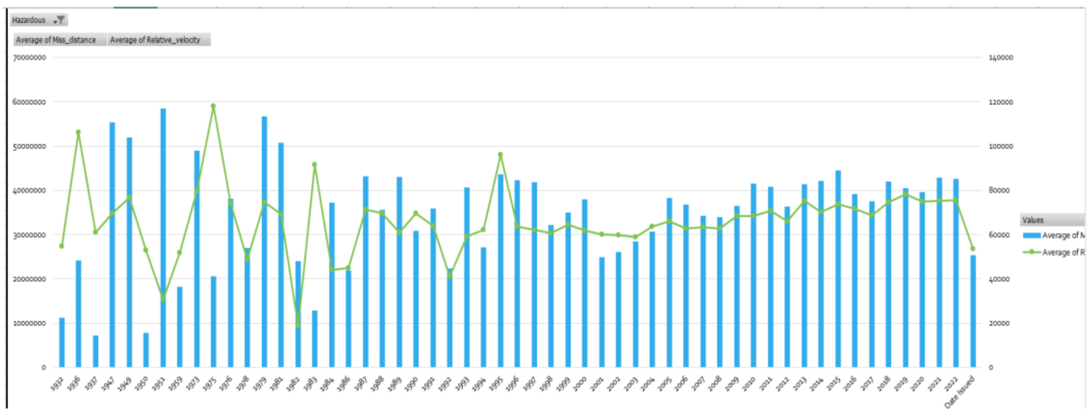


This data analysis gives insight into predicting the future neo-object's risk factors from

Earth. From the analysis, I interpret that to avoid the risk factor, the out-space objects should have less velocity and size than the average.



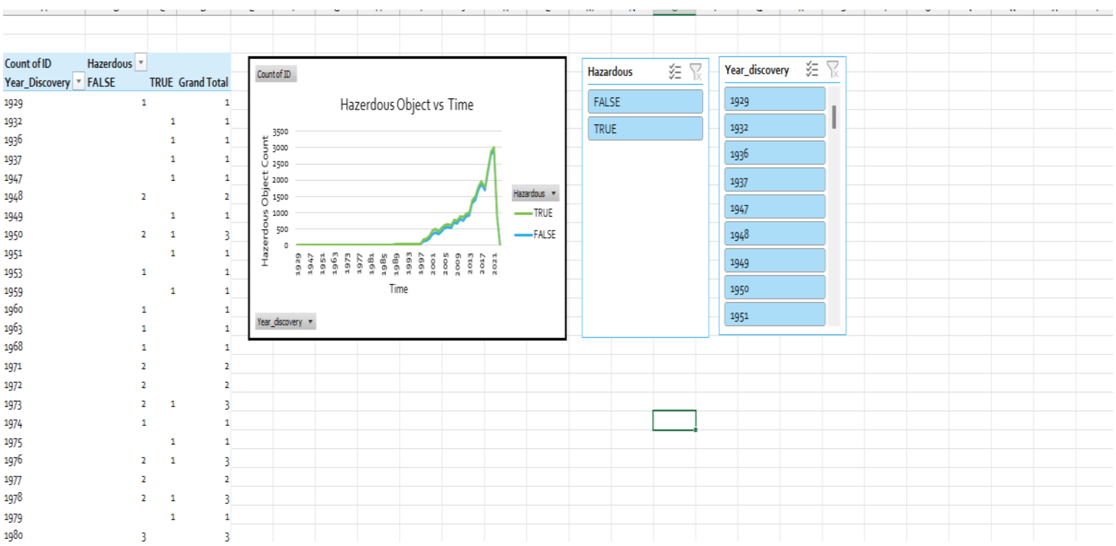
2. HAZARDOUS OBJECT MONITORING:



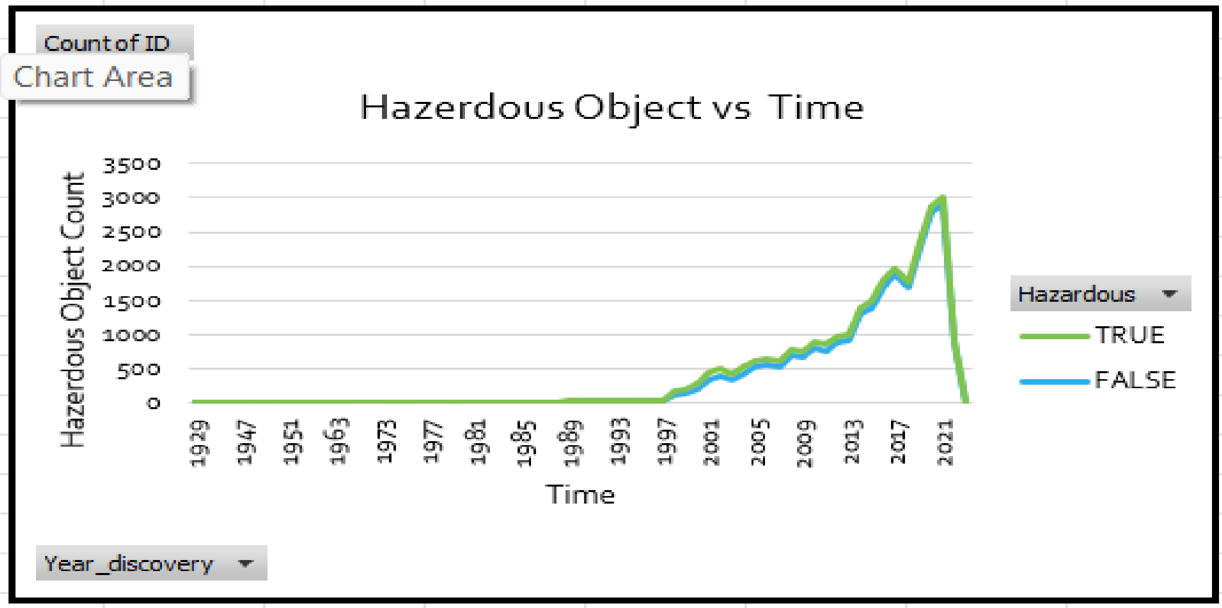
Monitoring NEO's years by making pivotable and detailed graphical analysis based on average miss distance and relative velocity based on hazardous or not, I get that miss distance, i.e., the distance from the Earth's orbit and its relative velocity is not proportional. Year-wise detail analysis shows that when the miss distance increases from its average, the relative velocity decreases, and potential risk increases.
In 1951, the NEO contained a high average relative velocity, and the distance's average was very low, which denotes the potential risks from the objects to Earth.

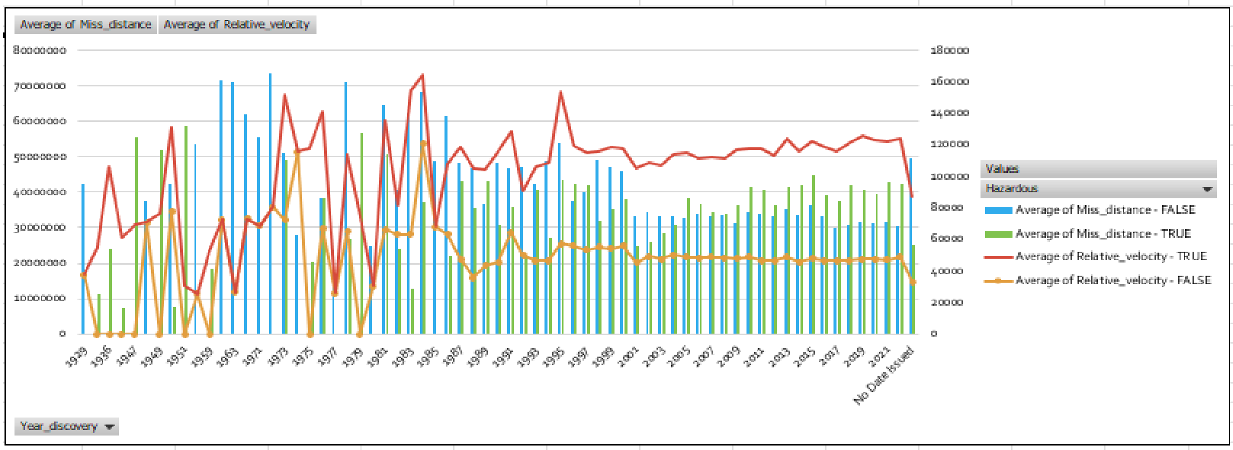
So, when monitoring hazardous objects, it's clear that we have to focus on the miss distances and relative velocities to predict the risk impact for future predictions.

3. RISK IMPACT AND SEASONAL TREND:


Count of ID	Hazerdous		
Year_Discovery	FALSE	TRUE	Grand Total
1929	1		1
1932		1	1
1936		1	1
1937		1	1
1947		1	1
1948	2		2
1949		1	1
1950	2	1	3
1951		1	1
1953	1		1
1959		1	1
1960	1		1
1963	1		1
1968	1		1
1971	2		2
1972	2		2
1973	2	1	3
1974	1		1
1975		1	1
1976	2	1	3
1977	2		2
1978	2	1	3
1979		1	1
1980	3		3

Analyzing year-wise NEO, I get that after 1989, the coming of our space objects toward Earth's surface increased, and 2021 reached its peak with 2999 objects coming, and before 1989, the count was meager. So, after 1989, i.e., at the start of the 20th century, the count of space objects increased, which may happen in the future. The count will increase cause after 2001 to 2022, the trend of observing the objects suddenly increases. However, seeing the previous data, the hazardous object count is very low compared to the non-hazardous objects. These two years, like 1983 and 1989, have only a much higher count of dangerous objects than the nonhazardous corresponding years. Despite these two, in the rest cases, the non-hazardous count is high, so it can be predicted that though per year we get NEOs, we have a low risk from those things until they cross the risk line and get a low miss distance.





=(F1126 - MIN(F:F)) / (MAX(F:F) - MIN(F:F))

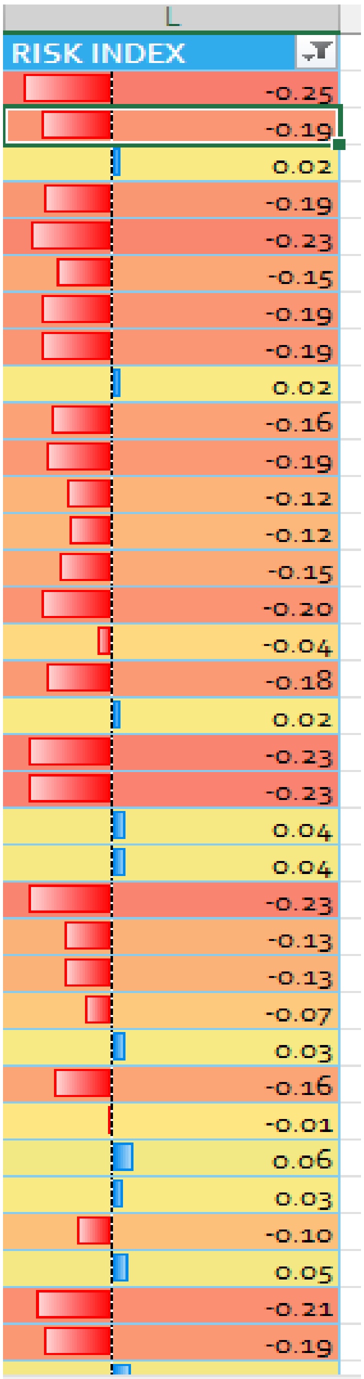
=(E1126 - MIN(E:E)) / (MAX(E:E) - MIN(E:E))

To analyze the risk impact from the NEO object, I've to make two additional columns named normalized miss distance and normalized relative velocity and create the risk index between them, which lies between -0.30 to 0.63 through the formula.

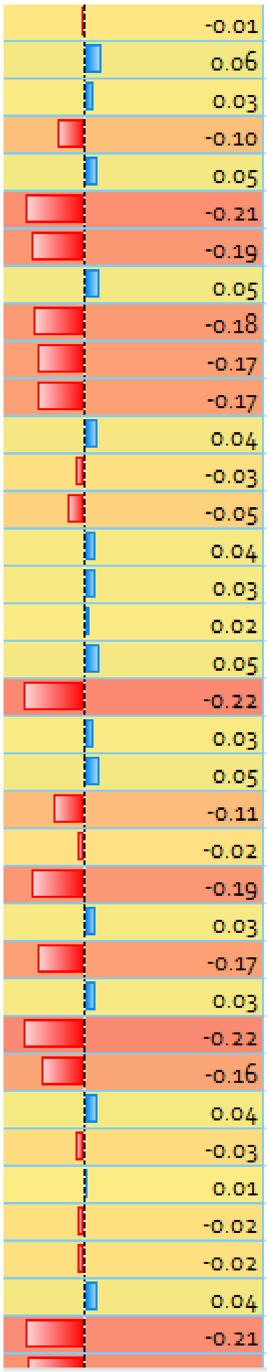
=(0.7 * J1126 - 0.3 * K1126)

This index denotes that when the index goes towards -0.30, the risk impact becomes so high, and its positive values have less effect on the safety range. The following charts show the non-proportional relation between average miss distance and relative velocity, which affects the risk impact. Risk impact goes to the non-proportional of relative velocity and proportional of miss distance.

Total Average of Miss_	Total Average of Relative_
42485508.81	37802.19171
11256645.6	54823.84085
24086731.63	106221.7733
7157012.617	60891.21815
55334045.19	69559.72129
37440698.06	70794.9978
51882752.85	76768.62725
30930628.48	69432.62616
58505275.48	30748.31162
53319315.27	25568.50595
18224809.67	51827.19209
71559524.61	72640.32301
70947620.41	26427.69253
61922309.05	72903.68786
55489467.21	68537.88039
73248171.85	80876.30954
50316437.13	74790.42946
28033970.12	115828.9194
20567325.3	117951.4076
38378620.24	69068.93169
23590393.8	25798.44272
56264042.99	59854.49546
56724390.08	74739.55508
24901932.44	30412.80927
59943902.32	67253.43588
35046404.51	48374.5101
32580338.99	80149.13076
52767178.94	82290.82695
48686701.04	68310.1233
54974276.02	60132.59161
47237742.6	52076.50157
42863787.38	47140.24319
40935686.23	55018.95797
43691777.42	52073.30357
42717977.3	64389.409
43720352.93	48509.26153
42141790.08	49043.41488
41303774.5	51924.96184
51944491.62	64130.47626
[illegible]	[illegible]



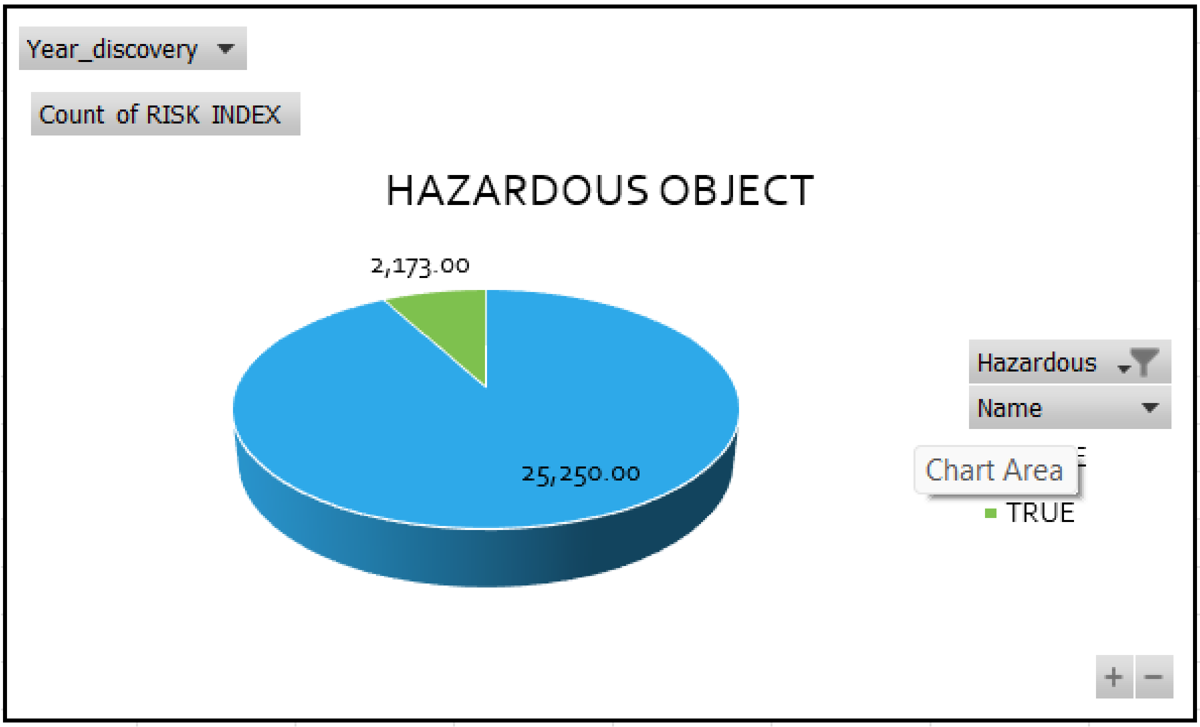
L
RISK INDEX
-0.25
-0.19
0.02
-0.19
-0.23
-0.15
-0.19
-0.19
0.02
-0.16
-0.19
-0.12
-0.12
-0.15
-0.20
-0.04
-0.18
0.02
-0.23
-0.23
0.04
0.04
-0.23
-0.13
-0.13
-0.07
0.03
-0.16
-0.01
0.06
0.03
-0.10
0.05
-0.21
-0.19



-0.01
0.06
0.03
-0.10
0.05
-0.21
-0.19
0.05
-0.18
-0.17
-0.17
0.04
-0.03
-0.05
0.04
0.03
0.02
0.05
-0.22
0.03
0.05
-0.11
-0.02
-0.19
0.03
-0.17
0.03
-0.22
-0.16
0.04
-0.03
0.01
-0.02
-0.02
0.04
-0.21

The yellow color denotes the low risk, and the red one gives the high risk**.** Depending on the miss distance it varies to the risk impact of NEO**.** However, a point is noticeable: although some objects touch the risk line, they don't become hazardous. Maybe the diameter size, relative velocity, or the missed distance fluctuation.

It's interpreted that based on previous data, in the future, we will have less risk impact from the upcoming NEO because as the risk index varies along with the increase of relative velocity and decrease of relative velocity, it may be the reason for the potentially high-risk zone.



The above chart shows that hazardous cases are much higher than non-hazardous cases depending on individually counting the risk index.

4. DATA STORYTELLING

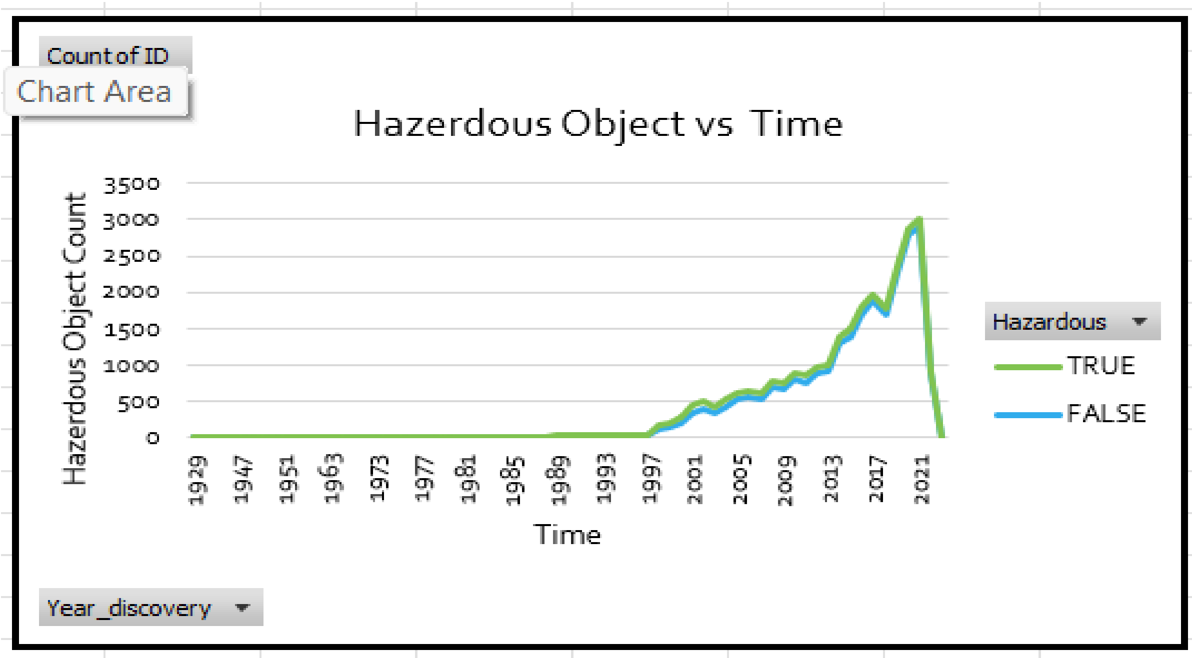
In this Data storytelling, I construct a narrative that conveys findings, insights into the trends, and key conclusions in an insightful and easily understandable manner. Let's build a data story based on your NEO dataset, Near-Earth Objects, from 1929 to 2022.

1. The Growing Awareness of NEOs

Before 1939, astronomers paid increasing attention to identifying and tracking NEOs amid growing apprehensions about their possible impact on Earth.

Plotting the number of annual NEOs discovered or observed. I see a trend in the increase over time, especially in the later years when technology and interest in space monitoring rose.

A line graph showing the annually growing number of NEOs discovered from 1929 – 2022.



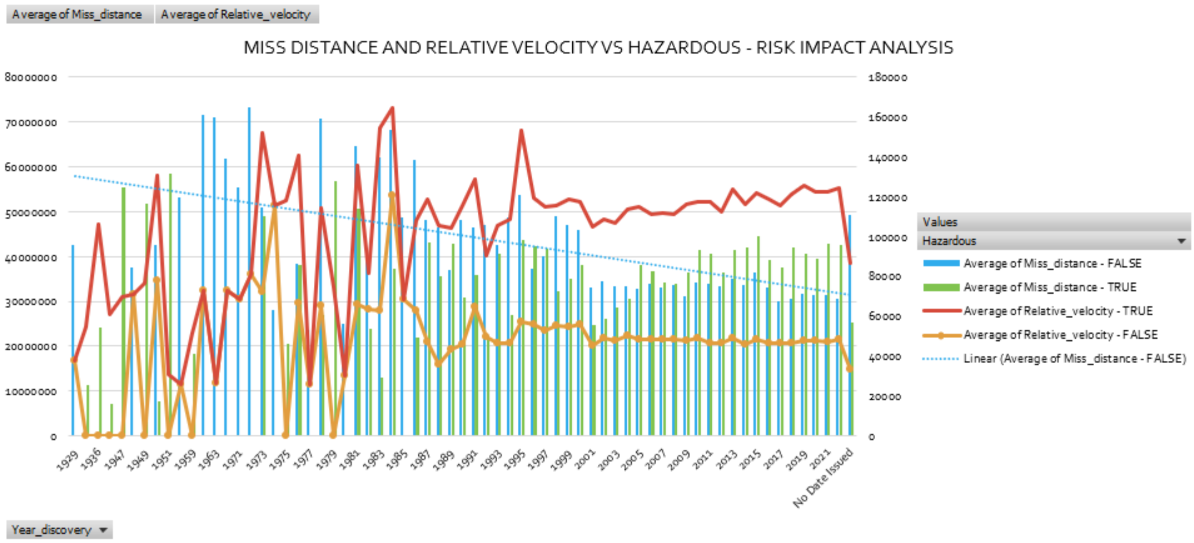
As humanity entered the mid-20th century, observation and follow-up of NEOs started to gain momentum. This post-war era saw an unprecedented development of technologies that could enable astronomers to spot more NEOs than ever. By the 1990s, the quantity of NEO discoveries had grown, reflecting improved detection capabilities and increasing awareness of the potential hazards that these kinds of objects can pose.

2. Proximity of NEOs

As NEO discoveries increased, attention turned to how close these objects were passing by Earth.

Data Point Plot the trend in miss distances over the years. Were NEOs passing closer and closer to Earth over time, or were the closest approaches random?

The missed distances for each year, including an optional trendline to show the overall trend.



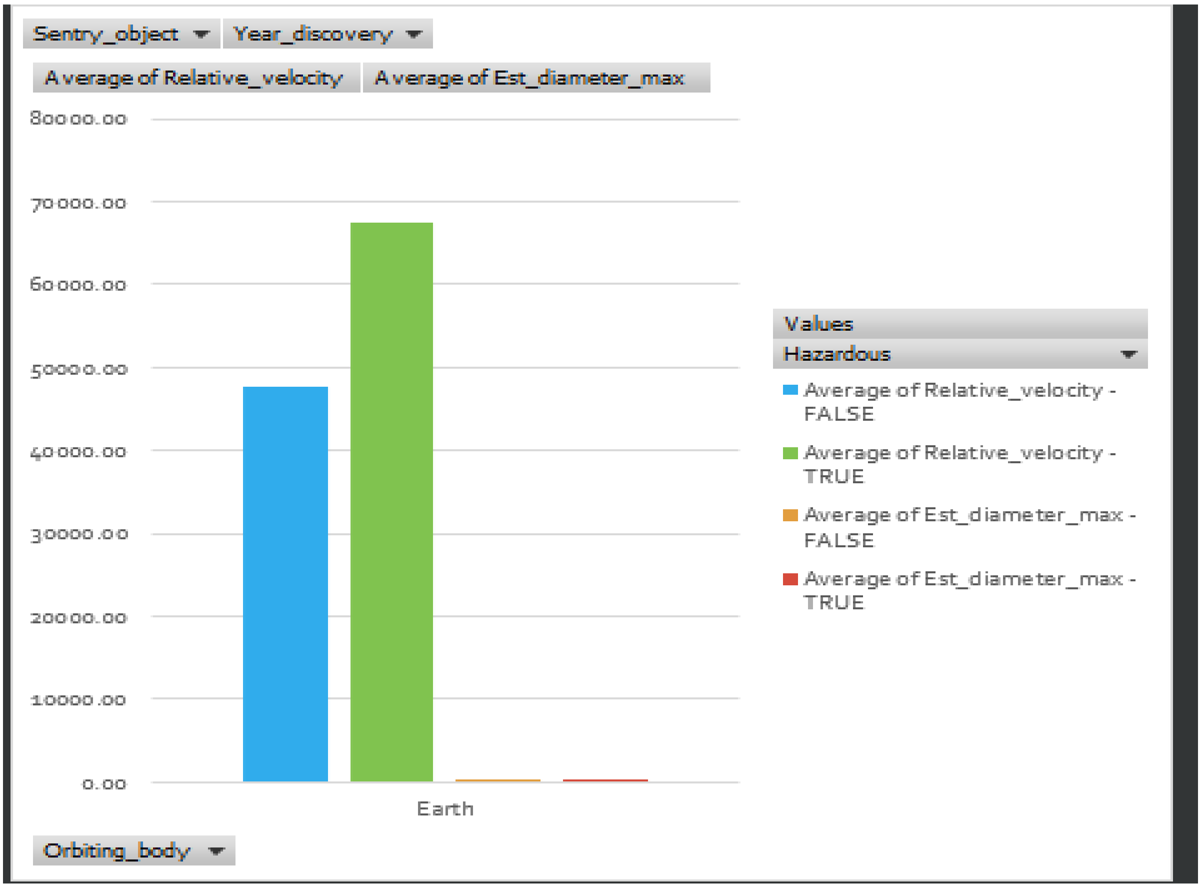
While NEO discoveries were on the rise, so was scrutiny of their orbits. The data indicates that during those decades, several NEOs passed reasonably close to Earth, which raised concerns over the possibility of impacts well into the future. However, the miss distances show huge variations and remind one of the unpredictability of these celestial bodies.

3. Speed of Danger: NEOs-Relative Velocity

The NEO relative velocity at the Earth passage is one of the main parameters for assessing the seriousness of the possible impact.

Trend Data Point Study the trends in relative velocities over time. NEOs approaching the Earth increasingly faster than the constant level.

A plot of relative velocities in the form of a histogram denotes this topic closely.



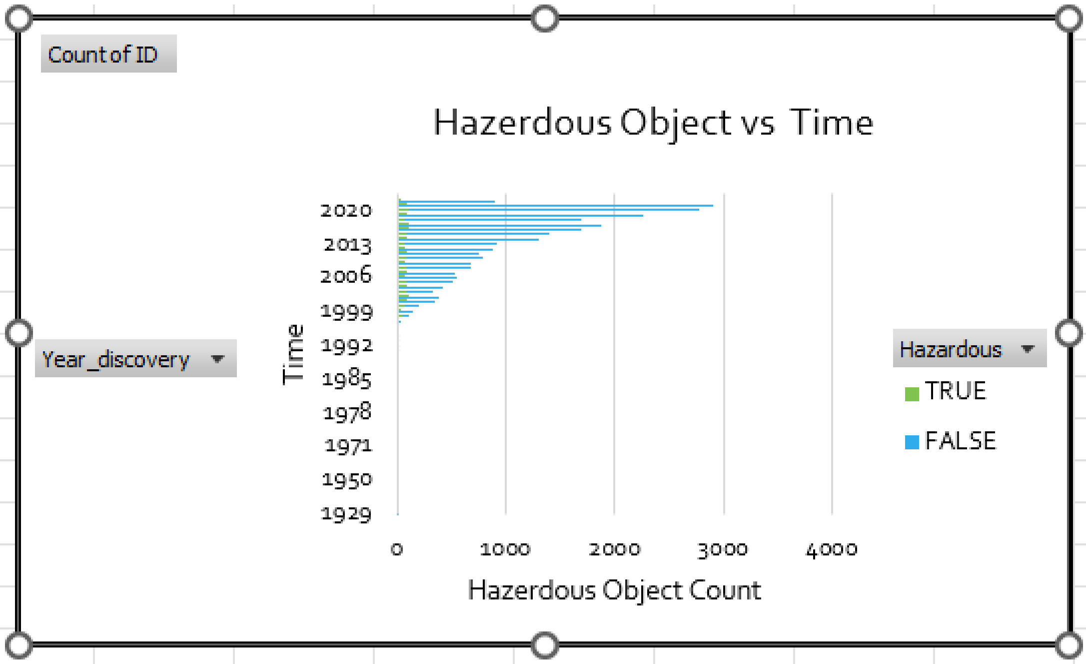
Not all NEOs are equally hazardous. Some fly across the face of our planet in a swift blink of an eye, yet others creep along. The data from 1929 to 2022 encompasses almost every conceivable relative velocity; thus, whereas some NEOs would provide little warning in the event of a potential collision, others may allow a more orderly approach.

4. The Hazard Factor: Identifying the Most Hazardous NEOs

The designation of NEOs as "hazardous" indicates the seriousness with which scientists view their potential threat.

Data Point Plot the number of NEOs classified as hazardous over the years. It seems from the data that hazardous NEOs are relatively stable and in a very low-count zone.

Plotting a bar graph showing hazardous NEOs discovered by year.



As astronomers refined their understanding of NEOs, certain objects were flagged as particularly dangerous. The data indicates that by the latter half of the 20th century, many NEOs were classified as hazardous, indicating the need for further and sustained vigilance and research in the area.

5. CONCLUSION

Considering how the data from 1929 through 2022 set the stage for today's NEO tracking and planetary defense initiatives. Outlook for the Future: Briefly address how the trends observed in this period continue to influence NEO observation strategies today. Story Element: The data from 1929 to 2022 narrates a story of increased awareness and sophistication in NEO observation. Moving into the 21st century, these grounds laid during these decades have proved vital in these ongoing endeavors for Earth's protection from a possible cosmic threat. So, our journey still goes on as we keep refining and sharpening our tools and strategies to answer whatever comes from the universe.

Indeed, the NEOs carefully observed over the period from 1939 to 2022 increased the catalog of such celestial bodies and helped scientists get a feel for the associated risks. The data shows a trend of growing discovery rates, variable miss distances, and a steady watch on possibly hazardous objects. This period laid the foundation for today's planetary defense initiatives and ensured, to a large extent, that we are in a better position to detect and probably mitigate any future threats.

Based on the analysis of Neo data from 1929 to the year 2022, we can advocate for better NEO detection, tracking, and mitigations. These would strengthen our understanding of the NEO menace while improving methods in planetary defense.

1. Improve investment in NEO-detection ground-based and space-based telescopes.

2. Improve Data Sharing and Global Collaboration. In connection with this, it is recommended that the international collaboration and data-sharing policy between space agencies, research institutions, and observatories be enhanced.

3. Plan and Test Methods of Planetary Defense, and priority should be given to developing and testing deflection and disruption technologies capable of deflecting a NEO from a hazardous orbit.

4. Refine Computational Models and Predictive Analytics with more sophisticated computational models should be used to predict NEOs' future path.

5. REFERENCES

DATASET: https://www.kaggle.com/datasets/ivansher/nasa-nearest-earth-objects-1910-2024/data

OTHER SOURCES: https://data.nasa.gov/