

```
import numpy as np
import pandas as pd

data_set = pd.read_csv('/neo.csv')
data_set
```

	id	name	est_diameter_min	est_diameter_max	relative_velocity	miss_
0	2162635	162635 (2000 SS164)	1.198271	2.679415	13569.249224	5.483974e+07
1	2277475	277475 (2005 WK4)	0.265800	0.594347	73588.726663	6.143813e+07
2	2512244	512244 (2015 YE18)	0.722030	1.614507	114258.692129	4.979872e+07
3	3596030	(2012 BV13)	0.096506	0.215794	24764.303138	2.543497e+07
4	3667127	(2014 GE35)	0.255009	0.570217	42737.733765	4.627557e+07
...
90831	3763337	(2016 UV11)	0.026580	0.059435	52078.886692	1.235204e+07

```
print(data_set.head())
print()
print(data_set.describe())
```

3D		id	name	est_diameter_min	est_diameter_max	\
	0	2162635	162635 (2000 SS164)	1.198271	2.679415	
	1	2277475	277475 (2005 WK4)	0.265800	0.594347	
	2	2512244	512244 (2015 YE18)	0.722030	1.614507	
	3	3596030	(2012 BV13)	0.096506	0.215794	
	4	3667127	(2014 GE35)	0.255009	0.570217	
		relative_velocity	miss_distance	orbiting_body	sentry_object	\
	0	13569.249224	5.483974e+07	Earth	False	
	1	73588.726663	6.143813e+07	Earth	False	
	2	114258.692129	4.979872e+07	Earth	False	
	3	24764.303138	2.543497e+07	Earth	False	
	4	42737.733765	4.627557e+07	Earth	False	
		absolute_magnitude	hazardous			
	0	16.73	False			
	1	20.00	True			
	2	17.83	False			
	3	22.20	False			
	4	20.09	True			
		id	est_diameter_min	est_diameter_max	relative_velocity	\
count	9.083600e+04	90836.000000	90836.000000	90836.000000	90836.000000	
mean	1.438288e+07	0.127432	0.284947	48066.918918	25293.296961	
std	2.087202e+07	0.298511	0.667491	0.001362	203.346433	
min	2.000433e+06	0.000609	0.001362	0.043057	28619.020645	
25%	3.448110e+06	0.019256	0.043057	0.108153	44190.117890	
50%	3.748362e+06	0.048368	0.108153	0.320656	62923.604633	
75%	3.884023e+06	0.143402	0.320656	84.730541	236990.128088	
max	5.427591e+07	37.892650	84.730541			
		miss_distance	absolute_magnitude			
count	9.083600e+04	90836.000000				
mean	3.706655e+07	23.527103				
std	2.235204e+07	2.894086				
min	6.745533e+03	9.230000				
25%	1.721082e+07	21.340000				
50%	3.784658e+07	23.700000				
75%	5.654900e+07	25.700000				
max	7.479865e+07	33.200000				

```
print(data_set.dtypes)
print()
print(data_set.isnull().sum())
```

id	int64
name	object
est_diameter_min	float64
est_diameter_max	float64
relative_velocity	float64

```
miss_distance      float64
orbiting_body      object
sentry_object      bool
absolute_magnitude float64
hazardous          bool
dtype: object

id      0
name    0
est_diameter_min  0
est_diameter_max  0
relative_velocity 0
miss_distance    0
orbiting_body    0
sentry_object    0
absolute_magnitude 0
hazardous        0
dtype: int64
```

```
#accessing the dataset
```

```
c = data_set['id']
print(c)
```

```
sd = data_set[['orbiting_body','absolute_magnitude']]
print(sd)
```

```
0      2162635
1      2277475
2      2512244
3      3596030
4      3667127
...
90831    3763337
90832    3837603
90833    54017201
90834    54115824
90835    54205447
Name: id, Length: 90836, dtype: int64
      orbiting_body absolute_magnitude
0      Earth      16.73
1      Earth      20.00
2      Earth      17.83
3      Earth      22.20
4      Earth      20.09
...      ...      ...
90831    Earth      25.00
90832    Earth      26.00
90833    Earth      24.60
90834    Earth      27.80
90835    Earth      24.12
```

```
[90836 rows x 2 columns]
```

```
a1 = data_set[data_set['absolute_magnitude'] > 20.00]
print(a1)
```

```
id      name  est_diameter_min  est_diameter_max \
3      3596030 (2012 BV13)      0.096506      0.215794
4      3667127 (2014 GE35)      0.255009      0.570217
5      54138696 (2021 GY23)      0.036354      0.081291
6      54189957 (2021 PY40)      0.171615      0.383743
7      54230078 (2021 XD6)      0.005328      0.011914
...      ...      ...
90831    3763337 (2016 VX1)      0.026580      0.059435
90832    3837603 (2019 AD3)      0.016771      0.037501
90833    54017201 (2020 JP3)      0.031956      0.071456
90834    54115824 (2021 CN5)      0.007321      0.016370
90835    54205447 (2021 TW7)      0.039862      0.089133

relative_velocity  miss_distance  orbiting_body  sentry_object \
3      24764.303138      2.543497e+07      Earth      False
4      42737.733765      4.627557e+07      Earth      False
5      34297.587778      4.058569e+07      Earth      False
6      27529.472307      2.906912e+07      Earth      False
7      57544.470083      5.511502e+07      Earth      False
...      ...      ...      ...
90831    52078.886692      1.230039e+07      Earth      False
90832    46114.605073      5.432121e+07      Earth      False
90833    7566.807732      2.840077e+07      Earth      False
90834    69199.154484      6.869206e+07      Earth      False
90835    27024.455553      5.977213e+07      Earth      False

absolute_magnitude  hazardous
3      22.20      False
4      20.09      True
5      24.32      False
```

6	20.95	False
7	28.49	False
...
90831	25.00	False
90832	26.00	False
90833	24.60	False
90834	27.80	False
90835	24.12	False

[79477 rows x 10 columns]

```
#sorting the dataset
sorting = data_set.sort_values(by='absolute_magnitude')
print(sorting)
```

	id		name	est_diameter_min	est_diameter_max	\
36418	2001036	1036	Ganymed (A924 UB)	37.892650	84.730541	
37651	2000433	433	Eros (A898 PA)	23.043847	51.527608	
12709	2000433	433	Eros (A898 PA)	23.043847	51.527608	
56533	2000433	433	Eros (A898 PA)	23.043847	51.527608	
16895	2001866	1866	Sisypheus (1972 XA)	8.640820	19.321462	
...
45661	54106298		(2021 B0)	0.000683	0.001528	
82493	54106298		(2021 B0)	0.000683	0.001528	
75986	54106298		(2021 B0)	0.000683	0.001528	
56366	54106298		(2021 B0)	0.000683	0.001528	
24949	3430497		(2008 TS26)	0.000609	0.001362	

	relative_velocity	miss_distance	orbiting_body	sentry_object	\
36418	51496.923293	5.372124e+07	Earth	False	
37651	21402.705247	2.672952e+07	Earth	False	
12709	15884.252623	5.468808e+07	Earth	False	
56533	21761.703426	3.120592e+07	Earth	False	
16895	96530.774890	7.402691e+07	Earth	False	
...
45661	69293.963546	6.947159e+07	Earth	False	
82493	27805.638117	2.399589e+04	Earth	False	
75986	30725.628392	1.686312e+07	Earth	False	
56366	43432.989900	3.436341e+07	Earth	False	
24949	56751.017090	1.263816e+04	Earth	False	

	absolute_magnitude	hazardous
36418	9.23	False
37651	10.31	False
12709	10.31	False
56533	10.31	False
16895	12.44	False
...
45661	32.95	False
82493	32.95	False
75986	32.95	False
56366	32.95	False
24949	33.20	False

[90836 rows x 10 columns]

```
x = data_set.dropna(inplace = True)
print(x)
```

None

```
y = data_set.drop_duplicates()
print(y)
```

	id		name	est_diameter_min	est_diameter_max	\
0	2162635	162635	(2000 SS164)	1.198271	2.679415	
1	2277475	277475	(2005 WK4)	0.265800	0.594347	
2	2512244	512244	(2015 YE18)	0.722030	1.614507	
3	3596030		(2012 BV13)	0.096506	0.215794	
4	3667127		(2014 GE35)	0.255009	0.570217	
...
90831	3763337		(2016 VX1)	0.026580	0.059435	
90832	3837603		(2019 AD3)	0.016771	0.037501	
90833	54017201		(2020 JP3)	0.031956	0.071456	
90834	54115824		(2021 CN5)	0.007321	0.016370	
90835	54205447		(2021 TW7)	0.039862	0.089133	

	relative_velocity	miss_distance	orbiting_body	sentry_object	\
0	13569.249224	5.483974e+07	Earth	False	
1	73588.726663	6.143813e+07	Earth	False	
2	114258.692129	4.979872e+07	Earth	False	
3	24764.303138	2.543497e+07	Earth	False	
4	42737.733765	4.627557e+07	Earth	False	
...
90831	52078.886692	1.230039e+07	Earth	False	
90832	46114.605073	5.432121e+07	Earth	False	
90833	7566.807732	2.840077e+07	Earth	False	

90834	69199.154484	6.869206e+07	Earth	False
90835	27024.455553	5.977213e+07	Earth	False

	absolute_magnitude	hazardous
0	16.73	False
1	20.00	True
2	17.83	False
3	22.20	False
4	20.09	True
...
90831	25.00	False
90832	26.00	False
90833	24.60	False
90834	27.80	False
90835	24.12	False

[90836 rows x 10 columns]

```
r = data_set.iloc[0]
print(r)
```

```
id          2162635
name        162635 (2000 SS164)
est_diameter_min    1.198271
est_diameter_max    2.679415
relative_velocity    13569.249224
miss_distance       54839744.082846
orbiting_body       Earth
sentry_object       False
absolute_magnitude    16.73
hazardous          False
Name: 0, dtype: object
```

```
n_arr = data_set['absolute_magnitude'].values
print(n_arr)
```

```
[16.73 20.  17.83 ... 24.6 27.8 24.12]
```

```
m = np.mean(n_arr)
print("The mean is:",m)
```

```
st = np.std(n_arr)
print("The standar values:",st)
```

```
mn = np.median(n_arr)
print("The medain value is:",mn)
```

```
The mean is: 23.52710347219164
The standar values: 2.8940695766099647
The medain value is: 23.7
```

```
b = data_set['absolute_magnitude']
print(b)
print("The modified new data...")
z = data_set['absolute_magnitude_new']=data_set['absolute_magnitude'].apply(lambda x :x*2)
print(z)
```

```
0      16.73
1      20.00
2      17.83
3      22.20
4      20.09
...
90831   25.00
90832   26.00
90833   24.60
90834   27.80
90835   24.12
Name: absolute_magnitude, Length: 90836, dtype: float64
The modified new data...
0      33.46
1      40.00
2      35.66
3      44.40
4      40.18
...
90831   50.00
90832   52.00
90833   49.20
90834   55.60
90835   48.24
Name: absolute_magnitude, Length: 90836, dtype: float64
```

```
#visualization
import pandas as pd
import matplotlib.pyplot as plt
d = pd.DataFrame(data_set)
d
```

	id	name	est_diameter_min	est_diameter_max	relative_velocity	miss.
0	2162635	162635 (2000 SS164)	1.198271	2.679415	13569.249224	5.4%
1	2277475	277475 (2005 WK4)	0.265800	0.594347	73588.726663	6.1%
2	2512244	512244 (2015 YE18)	0.722030	1.614507	114258.692129	4.9%
3	3596030	(2012 BV13)	0.096506	0.215794	24764.303138	2.5%
4	3667127	(2014 GE35)	0.255009	0.570217	42737.733765	4.6%
...
90831	3763337	(2016 VX1)	0.026580	0.059435	52078.886692	1.2%
90832	3837603	(2019 AD3)	0.016771	0.037501	46114.605073	5.4%

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the data
d = pd.read_csv('/neo.csv')

# Plotting
plt.figure(figsize=(12, 6))
plt.bar(d['id'], d['est_diameter_max'], color='skyblue')
plt.title("Maximum Estimated Diameter of Near-Earth Objects")
plt.xlabel('NEO ID')
plt.ylabel('Estimated Diameter (meters)')
plt.xticks(rotation=45)
plt.tight_layout() # Adjust layout to prevent clipping of axis labels
plt.show()
```



