

Q. The Planets dataset contains details about the 1,000+ extrasolar planets discovered up to 2014. Visualize the distribution of the masses of the planets (expressed as a multiple of the mass of Jupiter), using a histogram and a box plot. Make appropriate modifications to the chart title, axis titles, legend, figure size, font size, colors etc. to make the chart readable and visually appealing.

=>

+ Code

+ Markdown

```
1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 import seaborn as sns
```

[30] ✓ 0.0s

```
1 df = pd.read_csv(r"planets.csv")
```

[31] ✓ 0.0s

```
1 df.head()
```

[32] ✓ 0.0s

...

	method	number	orbital_period	mass	distance	year
0	Radial Velocity	1	269.300	7.10	77.40	2006
1	Radial Velocity	1	874.774	2.21	56.95	2008
2	Radial Velocity	1	763.000	2.60	19.84	2011
3	Radial Velocity	1	326.030	19.40	110.62	2007
4	Radial Velocity	1	516.220	10.50	119.47	2009

```
1 df.shape
```

[33] ✓ 0.0s

... (1035, 6)

```
1 df.info()
[34] ✓ 0.0s

... <class 'pandas.core.frame.DataFrame'>
RangeIndex: 1035 entries, 0 to 1034
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   method          1035 non-null   object
1   number           1035 non-null   int64
2   orbital_period   992 non-null    float64
3   mass             513 non-null    float64
4   distance          808 non-null    float64
5   year            1035 non-null   int64
dtypes: float64(3), int64(2), object(1)
memory usage: 48.6+ KB
```

```
1 df.isnull().sum()
[35] ✓ 0.0s

... method          0
number            0
orbital_period     43
mass              522
distance           227
year              0
dtype: int64
```

```
1 import warnings
2 warnings.filterwarnings("ignore")
[36] ✓ 0.0s
```

```
1 df['orbital_period'].fillna(df['orbital_period'].mean(),inplace=True)💡
[37] ✓ 0.0s
```

```
1 df['mass'].fillna(df['mass'].mean(),inplace=True)
[38] ✓ 0.0s
```

```
1 df['distance'].fillna(df['distance'].mean(),inplace=True)
[39] ✓ 0.0s
```

```
1 df.isnull().sum()
[40] ✓ 0.0s

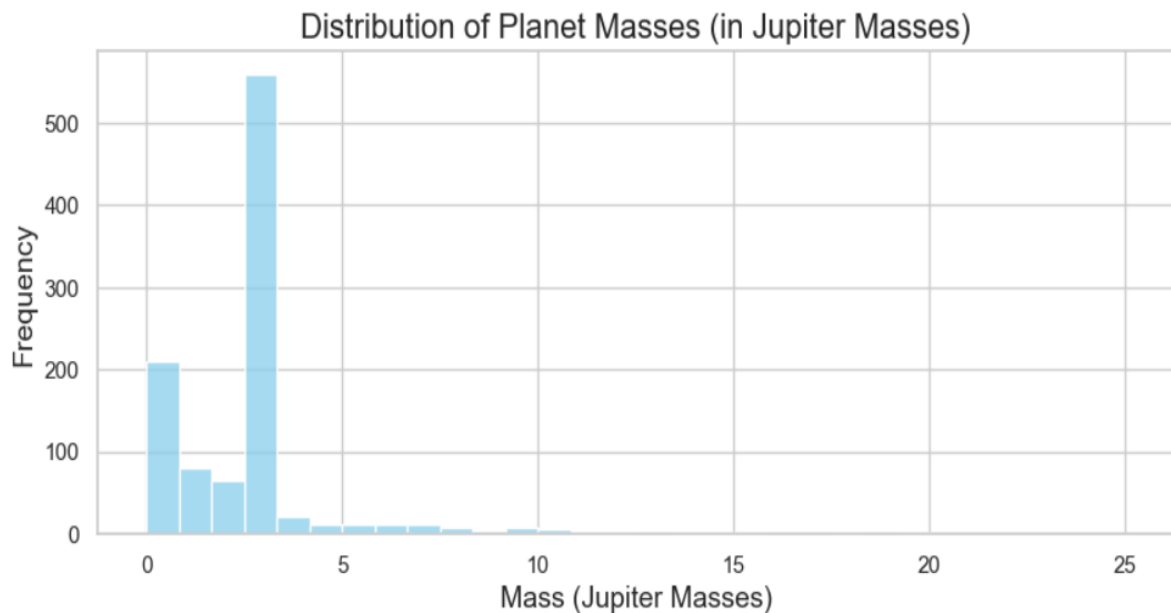
... method          0
number            0
orbital_period     0
mass              0
distance           0
year              0
dtype: int64
```

```

1 # Filter out any rows where the mass is missing or zero (if applicable)
2 planets_df = df[df['mass'].notnull() & (df['mass'] > 0)]
3
4 # Histogram
5 plt.figure(figsize=(10,4))
6 sns.histplot(planets_df['mass'], bins=30,color='skyblue')
7 plt.title('Distribution of Planet Masses (in Jupiter Masses)', fontsize=16)
8 plt.xlabel('Mass (Jupiter Masses)', fontsize=14)
9 plt.ylabel('Frequency', fontsize=14)
10 plt.show()
11

```

[44] ✓ 0.2s



```

1 # Box plot
2 plt.figure(figsize=(10,4))
3 sns.boxplot(x=planets_df['mass'],color='lightgreen')
4 plt.title('Box Plot of Planet Masses (in Jupiter Masses)', fontsize=16)
5 plt.xlabel('Mass (Jupiter Masses)', fontsize=14)
6 plt.show()

```

[45] ✓ 0.1s

