Task 1:- to create a bar chart or histogram to visualize the distribution of a categorical or continuous variable.

In this task, we conducted an analysis of the Titanic dataset to determine the distribution of passengers who survived versus those who did not. This analysis was performed using Python, specifically leveraging the pandas and matplotlib.pyplot libraries within a Jupyter Notebook environment.

Analysis Summary

Objective:- To visualize and quantify the number of passengers who survived and those who did not.

Tools Used- Pandas (for data manipulation and analysis) and Matplotlib (for creating visual representations of data)

The analysis revealed that,

342 passengers survived the titanic disaster and **549 did not survive**.

Out of Survivors, **109 were Females** and **233 were males**

In [15]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [16]:
```
df=pd.read_csv("C:\\Users\\Simarjeet kaur\\Downloads\\titanic.csv")
df
```

Out[16]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows × 12 columns

In [17]:
```
df.describe()
```

Out[17]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

In [18]:
```
df.head()
```

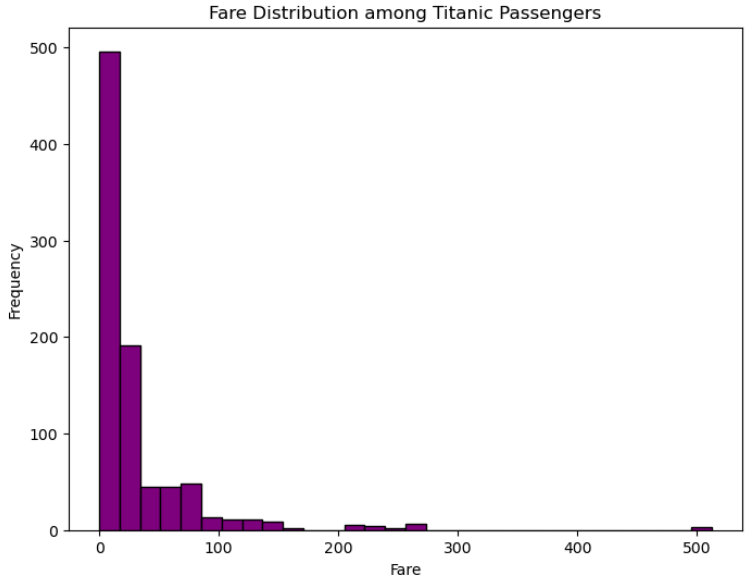
Out[18]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

In [19]:
```
df['Fare'].isnull().sum()
```

Out[19]:
```
0
```

In [29]:
```
plt.figure(figsize=(8,6))
plt.hist(df['Fare'], bins=30, color='purple', edgecolor='black')
plt.title('Fare Distribution among Titanic Passengers')
plt.xlabel('Fare')
plt.ylabel('Frequency')
plt.show()
```

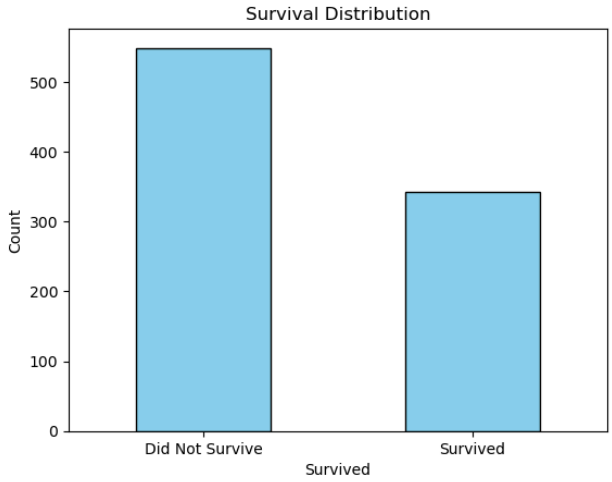


In [21]:
```
Survived=df["Survived"]
Age=df["Age"]
Pclass=df["Pclass"]
```

In [31]:
```
df['Survived'].value_counts().sort_index().plot(kind='bar', title='Survival Distribution', color="skyblue",edgecolor="Black")

plt.xlabel('Survived')
plt.ylabel('Count')

plt.xticks(ticks=[0, 1], labels=['Did Not Survive', 'Survived'], rotation=0)
plt.show()
```



In [23]:
```
Survived=(df["Survived"]).sum()
print(f"{Survived} people survived")
```

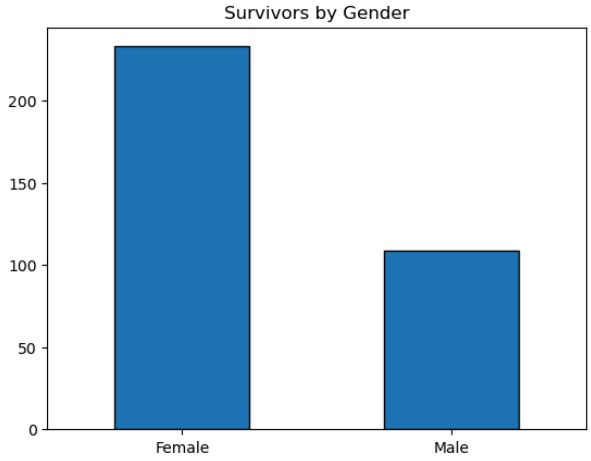
```
342 people survived
```

In [24]:
```
did_not_survive=(df["Survived"]==0).sum()
print(f"{did_not_survive} people did not survived")
```

```
549 people did not survived
```

In [25]:
```
survivors = df[df['Survived'] == 1]
```

In [32]:
```
survivors['Sex'].value_counts().plot(kind='bar', title='Survivors by Gender', edgecolor="Black")
plt.xticks(ticks=[0, 1], labels=['Female', 'Male'], rotation=0)
plt.show()
```



In [27]:
```
survivors=df[df["Survived"]==1]
gender_count=survivors["Sex"].value_counts()
male_survivor=gender_count.get("male",0)
female_survivor=gender_count.get("female",0)
```

In [28]:
```
print(f"Total {male_survivor} males survived")
print(f"Total {female_survivor} females survived")
```

```
Total 109 males survived
Total 233 females survived
```

In [33]:
```
import jovian
```

In [34]:
```
jovian.commit(project="Task 1 Data Science Internship Prodigy InfoTech")
```

```
[jovian] Creating a new project "simarjeetk927/Task 1 Data Science Internship Prodigy InfoTech"
[jovian] Committed successfully! https://jovian.com/simarjeetk927/task-1-data-science-internship-prodigy-infotech
```

Out[34]:
```
'https://jovian.com/simarjeetk927/task-1-data-science-internship-prodigy-infotech'
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: