Prodigy Infotech Data Science Internship

TASK-1

Create a bar chart or histogram to visualize the distribution of a categorical or continuous variable, such as the distribution of ages or genders in a population.

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

In [12]:
```
df=pd.read_csv("indian_liver_patient.csv")
```

In [13]:
```
df
```

Out[13]:

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protie
0	65	Female	0.7	0.1	187	16	18	
1	62	Male	10.9	5.5	699	64	100	
2	62	Male	7.3	4.1	490	60	68	
3	58	Male	1.0	0.4	182	14	20	
4	72	Male	3.9	2.0	195	27	59	
...	...	...	...	...	...	...	...	
578	60	Male	0.5	0.1	500	20	34	
579	40	Male	0.6	0.1	98	35	31	
580	52	Male	0.8	0.2	245	48	49	
581	31	Male	1.3	0.5	184	29	32	
582	38	Male	1.0	0.3	216	21	24	

583 rows × 11 columns

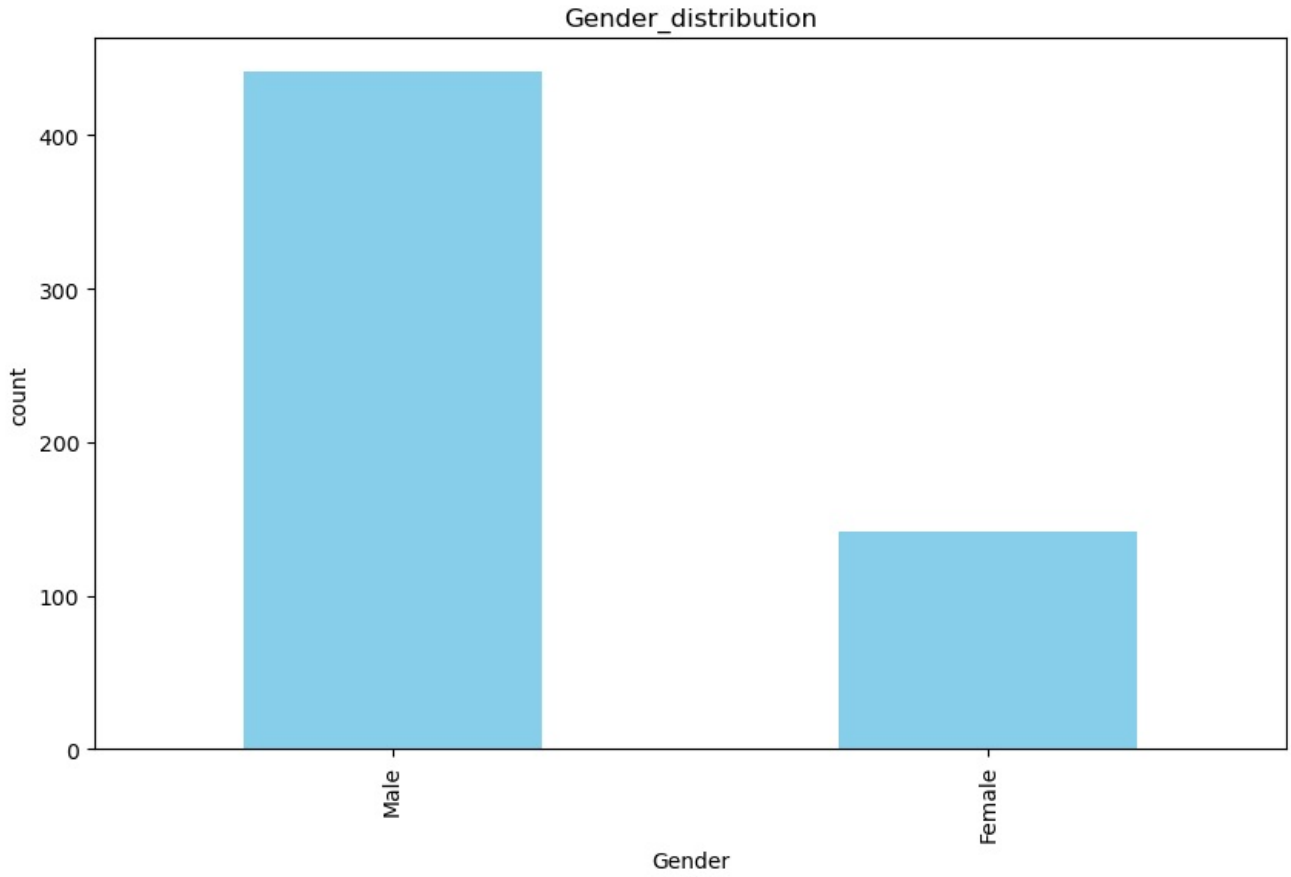
In [14]:
```
df.head()
```

Out[14]:

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens
0	65	Female	0.7	0.1	187	16	18	6.8
1	62	Male	10.9	5.5	699	64	100	7.5
2	62	Male	7.3	4.1	490	60	68	7.0
3	58	Male	1.0	0.4	182	14	20	6.8
4	72	Male	3.9	2.0	195	27	59	7.3

In [15]:
```
gender_counts=df['Gender'].value_counts()
```

In [20]:
```
#distributiopn of gender

plt.figure(figsize=(10,6))
gender_counts.plot(kind="bar",color='skyblue')
plt.title("Gender_distribution")
plt.ylabel('count')
plt.xlabel("Gender")
plt.show()
```



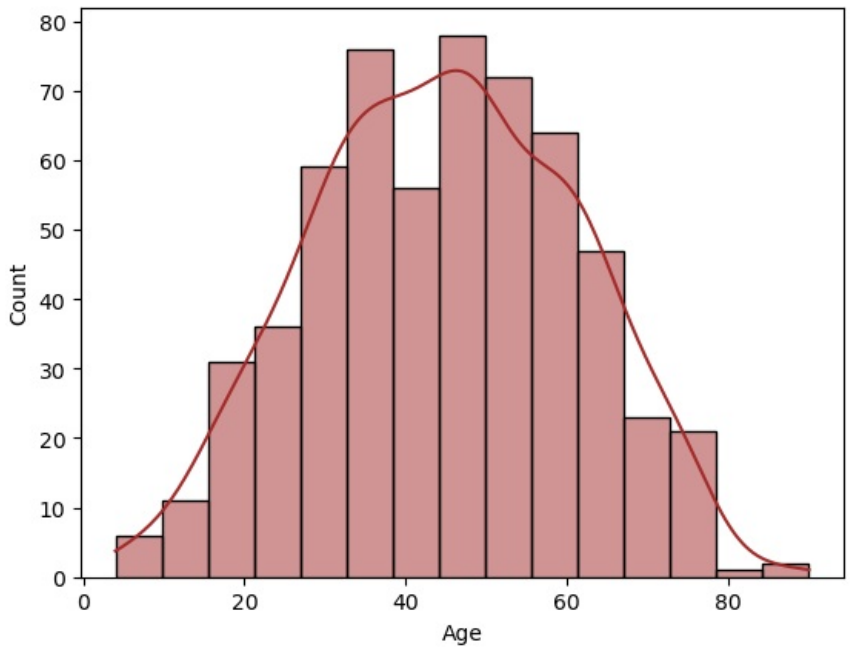
In [21]:
```
#distribution of age

age_counts=df["Age"].value_counts()
```

In [22]:
```
age_counts
```

Out[22]:
```
Age
60    34
45    25
50    23
42    21
38    21
      ..
78     1
11     1
67     1
10     1
90     1
Name: count, Length: 72, dtype: int64
```

In [24]:
```
sns.histplot(data=df,x="Age",kde=True,color="brown")
```

Out[24]:
```
<Axes: xlabel='Age', ylabel='Count'>
```



In [ ]: