


```
import pandas as pd

# Load the dataset
covid_data = pd.read_csv("/content/covid_cleaned_data.csv")

covid_data
```



	usmer	medical_unit	sex	patient_type	intubed	pneumonia	age	pregnant	diabetes	copd	...	hipertension	other_disease	c
0	2		1	1	2	2	1	40	2	2	2	...	2	2
1	2		1	1	2	2	2	37	2	1	2	...	1	2
2	2		1	1	2	2	2	25	2	2	2	...	2	2
3	2		1	1	2	2	1	80	2	2	2	...	1	2
4	2		1	1	2	1	1	58	2	2	2	...	1	2
...
78115	1		13	1	2	2	2	79	2	2	2	...	1	2
78116	2		13	1	2	2	1	61	2	1	2	...	1	2
78117	2		13	1	2	2	1	63	2	2	2	...	2	2
78118	1		13	1	2	2	2	23	2	1	2	...	2	2
78119	1		13	1	2	2	2	56	2	1	2	...	2	2

78120 rows × 22 columns

```
covid_data.describe()
```

	usmer	medical_unit	sex	patient_type	intubed	pneumonia	age	pregnant	diabetes	
count	78120.000000	78120.000000	78120.0	78120.0	78120.000000	78120.000000	78120.000000	78120.000000	78120.000000	78120.0
mean	1.450819	7.324411	1.0	2.0	1.847913	1.430415	52.769048	2.378751	2.161738	2.34
std	0.497579	3.691005	0.0	0.0	0.359107	0.495137	20.299598	6.201378	6.786956	6.21
min	1.000000	1.000000	1.0	2.0	1.000000	1.000000	0.000000	1.000000	1.000000	1.00
25%	1.000000	4.000000	1.0	2.0	2.000000	1.000000	40.000000	2.000000	1.000000	2.00
50%	1.000000	6.000000	1.0	2.0	2.000000	1.000000	55.000000	2.000000	2.000000	2.00
75%	2.000000	12.000000	1.0	2.0	2.000000	2.000000	67.000000	2.000000	2.000000	2.00
max	2.000000	13.000000	1.0	2.0	2.000000	2.000000	115.000000	98.000000	98.000000	98.00

8 rows × 21 columns

```
covid_data.columns

Index(['usmer', 'medical_unit', 'sex', 'patient_type', 'intubed', 'pneumonia',
      'age', 'pregnant', 'diabetes', 'copd', 'asthma', 'inmsupr',
      'hipertension', 'other_disease', 'cardiovascular', 'obesity',
      'renal_chronic', 'tobacco', 'clasiffication_final', 'icu', 'death',
      'date_parsed'],
      dtype='object')
```

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

# Scatter plot of Age vs Patient Outcomes
plt.figure(figsize=(15, 5))

plt.subplot(1, 3, 1)
plt.scatter(covid_data['age'], covid_data['intubed'], alpha=0.5)
plt.title('Age vs Intubation')
plt.xlabel('Age')
plt.ylabel('Intubation')

plt.subplot(1, 3, 2)
plt.scatter(covid_data['age'], covid_data['pneumonia'], alpha=0.5)
plt.title('Age vs Pneumonia')
plt.xlabel('Age')
plt.ylabel('Pneumonia')

plt.subplot(1, 3, 3)
plt.scatter(covid_data['age'], covid_data['death'], alpha=0.5)
plt.title('Age vs Death')
plt.xlabel('Age')
plt.ylabel('Death')

plt.tight_layout()
plt.show()

# Bar chart of categorical variables
plt.figure(figsize=(15, 5))

plt.subplot(1, 3, 1)
covid_data['sex'].value_counts().plot(kind='bar')
plt.title('Gender Distribution')
plt.xlabel('Gender')
plt.ylabel('Count')

plt.subplot(1, 3, 2)
covid_data['patient_type'].value_counts().plot(kind='bar')
plt.title('Patient Type Distribution')
plt.xlabel('Patient Type')
plt.ylabel('Count')

plt.subplot(1, 3, 3)
covid_data['icu'].value_counts().plot(kind='bar')
plt.title('ICU Admission')
plt.xlabel('ICU Admission')
plt.ylabel('Count')

plt.tight_layout()
plt.show()
```

