```
import pandas as pd
import seaborn as sns
```

```
df=pd.read_csv("insurance.csv")
```

```
df
```

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520
...	...	...	...	...	...	...	...
1333	50	male	30.970	3	no	northwest	10600.54830
1334	18	female	31.920	0	no	northeast	2205.98080
1335	18	female	36.850	0	no	southeast	1629.83350
1336	21	female	25.800	0	no	southwest	2007.94500
1337	61	female	29.070	0	yes	northwest	29141.36030

1338 rows × 7 columns

```
df.head
```

pandas.core.generic.NDFrame.head
```
def head(n: int=5) -> Self
```

```
Return the first `n` rows.

This function returns the first `n` rows for the object based
on position. It is useful for quickly testing if your object
has the right type of data in it.
```

```
df.tail
```

pandas.core.generic.NDFrame.tail
```
def tail(n: int=5) -> Self
```

```
Return the last `n` rows.

This function returns last `n` rows from the object based on
position. It is useful for quickly verifying data, for example,
after sorting or appending rows.
```

```
df['age'].mean()
```

```
np.float64(39.20702541106129)
```

```
df.isnull().sum()
```

	0
age	0
sex	0
bmi	0
children	0
smoker	0
region	0

```
df.nunique()
```

	0
age	47
sex	2
bmi	548
children	6
smoker	2
region	4
charges	1337

dtype: int64

dtype: int64

```
df['age'].value_counts()
```

age	count
18	69
19	68
46	29
52	29
50	29
47	29
48	29
45	29
20	29
27	28
28	28
49	28
54	28
53	28
22	28
21	28
26	28
31	27
58	25
59	25

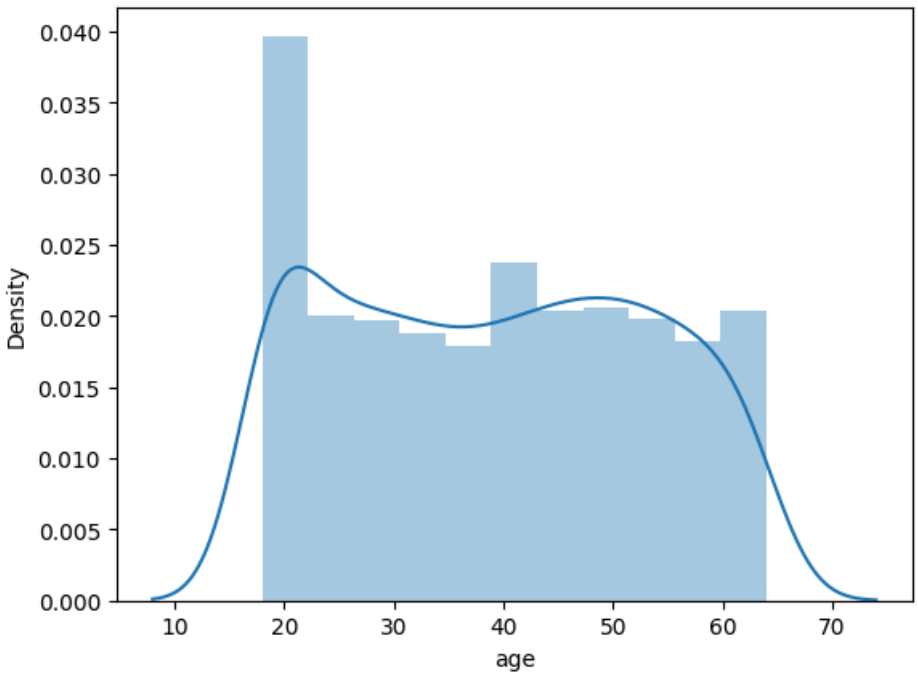
```
df['sex'].value_counts()
```

sex	count
male	676
female	662

dtype: int64

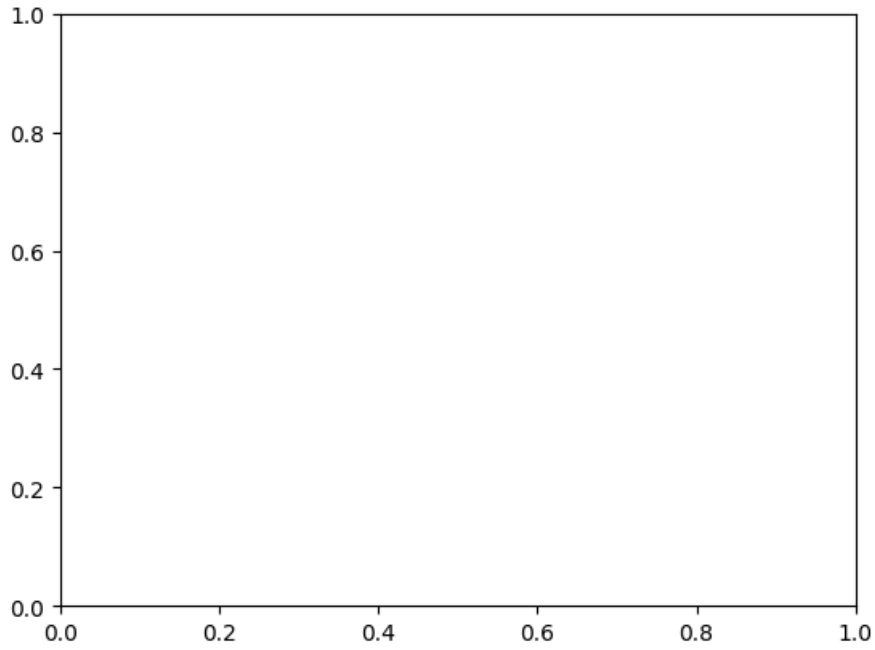
```
sns.distplot(df['age'])
```

```
/tmp/ipython-input-3234920688.py:1: UserWarning:
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.
Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['age'])
<Axes: xlabel='age', ylabel='Density'>
```



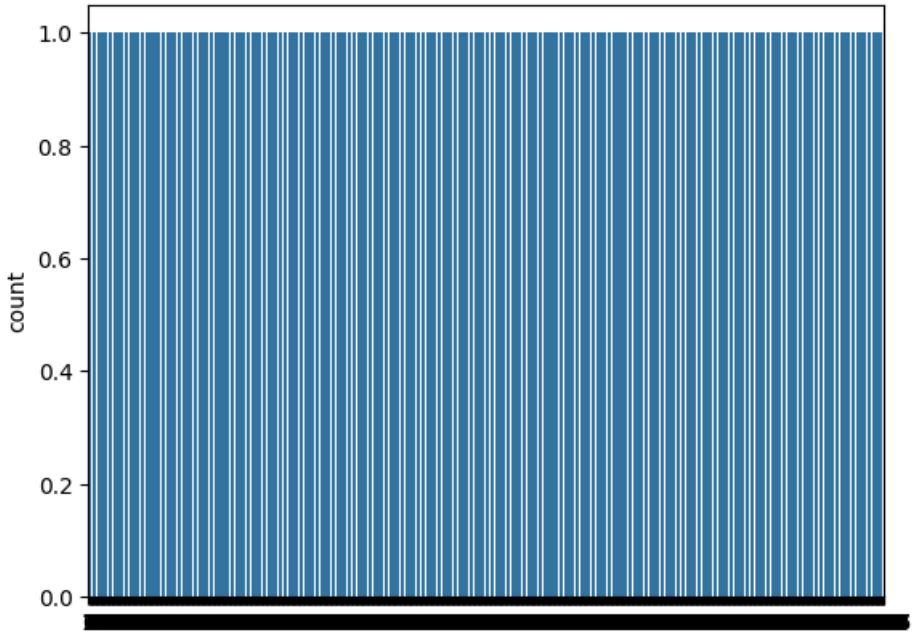

```
sns.distplot(df['sex'])
```

```
/tmp/ipython-input-2062954179.py:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['sex'])
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
/tmp/ipython-input-2062954179.py in <cell line: 0>()
----> 1 sns.distplot(df['sex'])

1 frames
/usr/local/lib/python3.12/dist-packages/pandas/core/series.py in __array__(self, dtype, copy)
   1029         """
   1030         values = self._values
-> 1031         arr = np.asarray(values, dtype=dtype)
   1032         if using_copy_on_write() and astype_is_view(values.dtype, arr.dtype):
   1033             arr = arr.view()

ValueError: could not convert string to float: 'female'
```



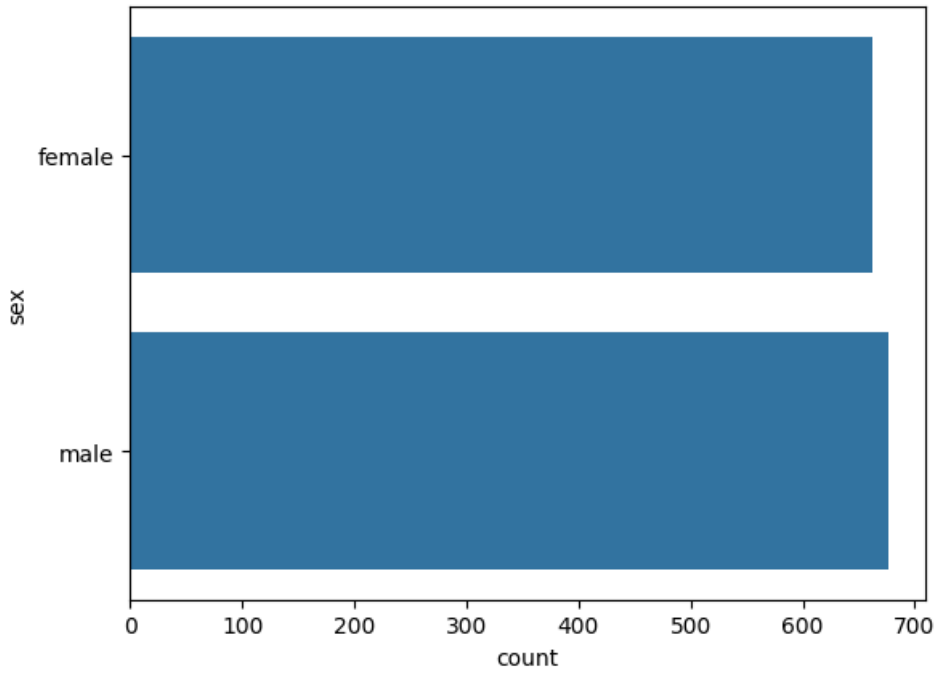

```
sns.countplot(df['age'])
```

```
<Axes: ylabel='count'>
```



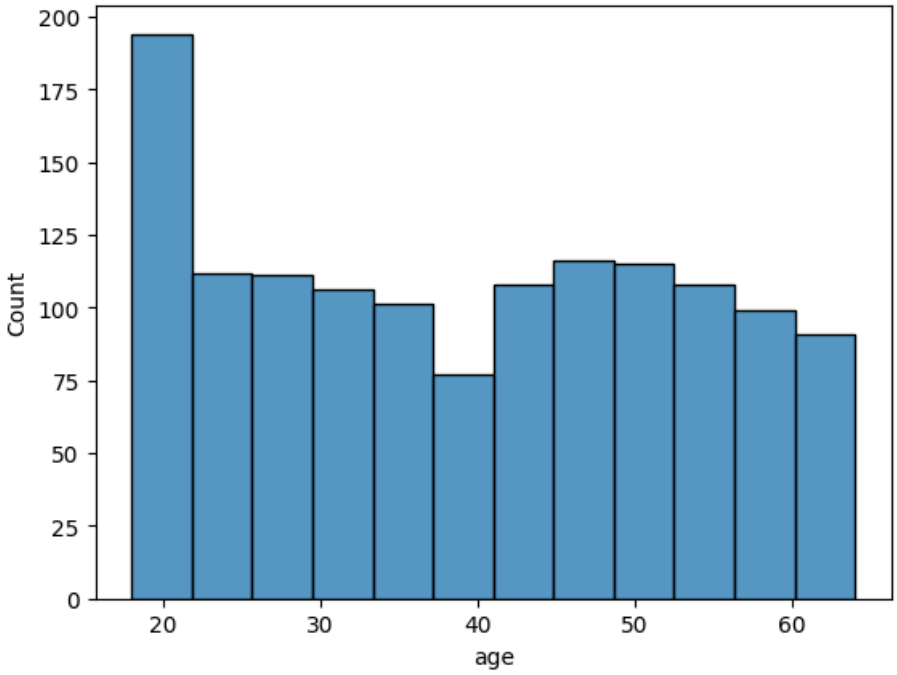

```
sns.countplot(df['sex'])
```

```
<Axes: xlabel='count', ylabel='sex'>
```



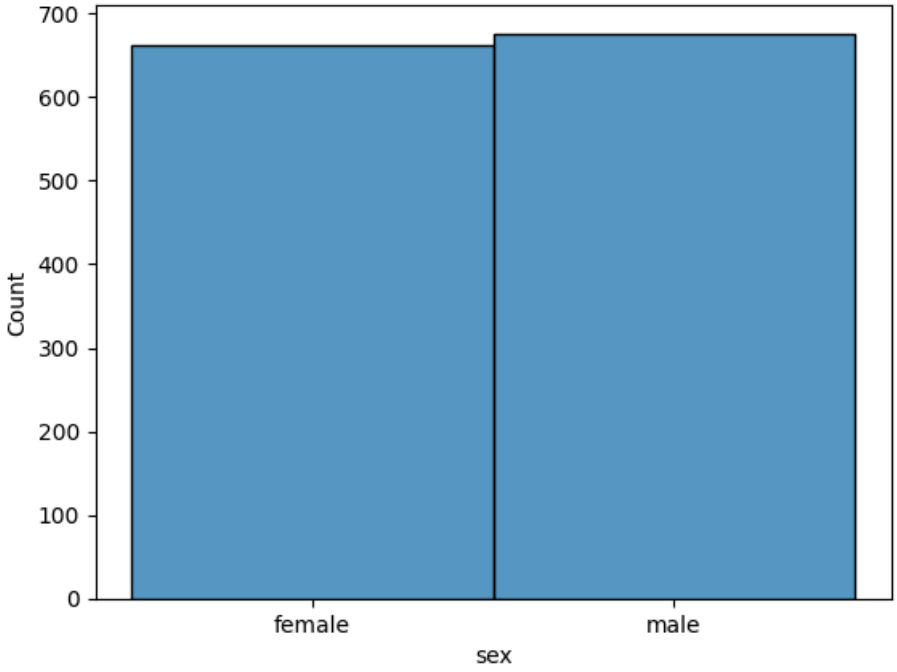

```
sns.histplot(df['age'])
```

```
<Axes: xlabel='age', ylabel='Count'>
```



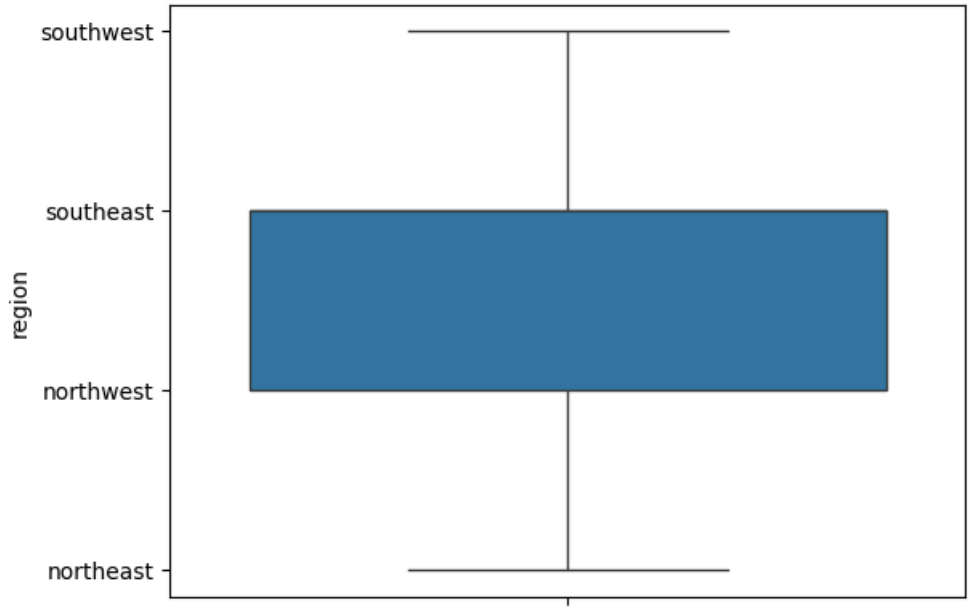

```
sns.histplot(df['sex'])
```

```
<Axes: xlabel='sex', ylabel='Count'>
```



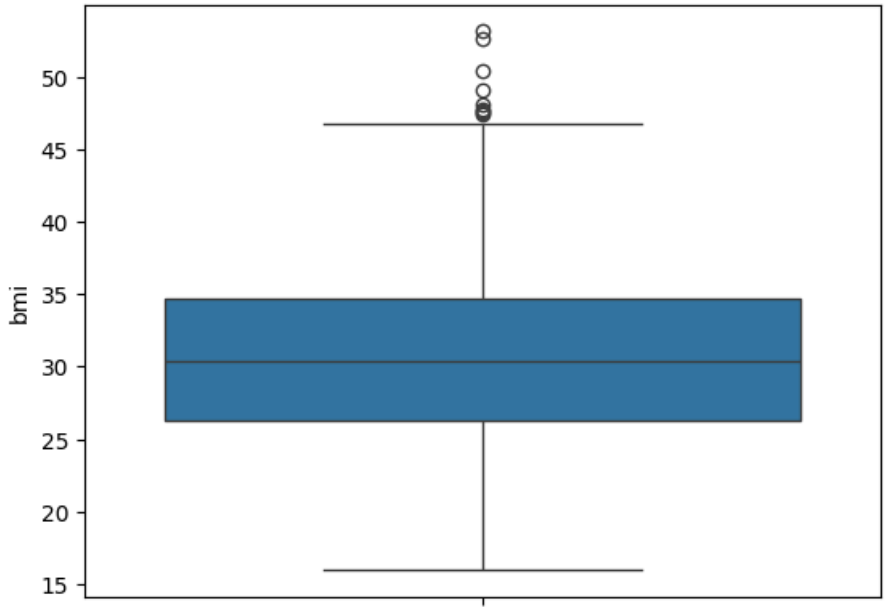

```
sns.boxplot(df['region'])
```

```
<Axes: ylabel='region'>
```



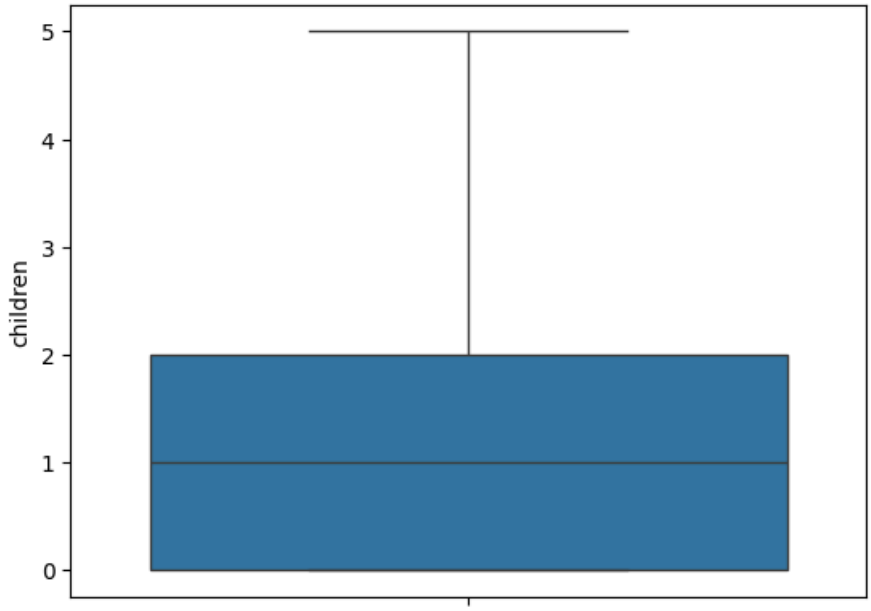

```
sns.boxplot(df['bmi'])
```

```
<Axes: ylabel='bmi'>
```



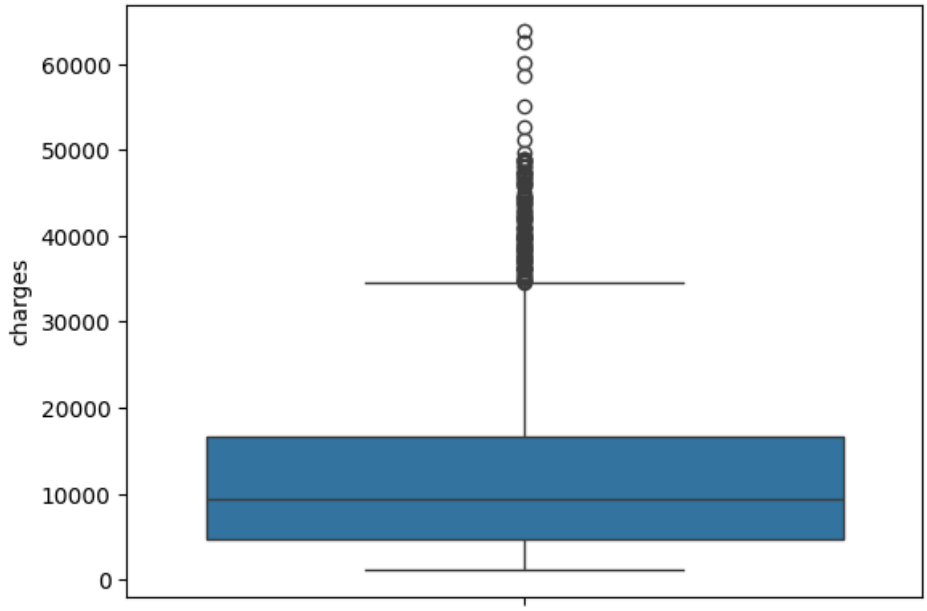

```
sns.boxplot(df['children'])
```

```
<Axes: ylabel='children'>
```




```
sns.boxplot(df['charges'])
```

```
<Axes: ylabel='charges'>
```

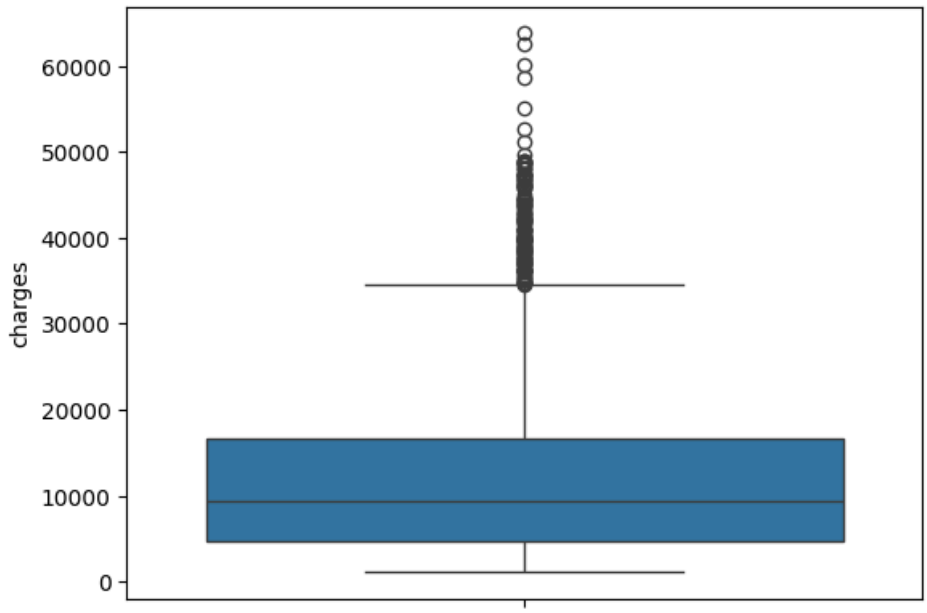



```
sns.boxplot(df['smoker'])
```

```
<Axes: ylabel='smoker'>
```


```
sns.boxplot(df['charges'])
```

```
<Axes: ylabel='charges'>
```




```
sns.boxplot(df['age'])
```

```
<Axes: ylabel='age'>
```