22/7/25 Load the Titanic dataset and convert
Ex 1 it into a data frame.

Aim:

To perform basic preprocessing and exploring data analysis on the Titanic dataset using pandas, seaborn and sklearned.

Procedure / Algorithm:-

Step 1: Load the Titanic dataset

Step 2: display the first few rows using head

Step 3: Explore column data types and check for missing values.

Step 4: Apply forward fill and backward fill methods.

Step 5: fill missing column values with 'unknown'.

Step 6: Encode the sex col using label Encoder.

Step 7: Create a pair plot for 'polars'

Step 8: Display conclusion heatmap

o/p:

first 5 Rows.

Survived	pclass	sex	age	Sibsp	Parch	fare	embarked	class
0	3	M	22	1	0	7.25	S	Third
1	1	F	38	1	0	71.283	C	first
1	3	f	26	0	0	7.925	S	Third
1	1	f	35	1	0	53.10	S	first
0	3	M	35	0	0	8.050	S	Third

Data info:

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 15 columns)

#	Column	non-null count	Dtype
0	Survived	891 non-null	int64
1	pclass	891 non-null	int64
2	sex	891 non-null	object
3	age	891 non-null	float64
4	sbsp	891 non null	int64
5	parch	891 non-null	int64
6	fare	891 non null	float64.

Code:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder, StandardScaler

df = sns.load_dataset('titanic')
print("First 5 Rows:")
display(df.head())
print("\nData Info:")
df.info()
df['age'] = df['age'].fillna(method='ffill')
df['age'] = df['age'].fillna(method='bfill')
df['deck'] = df['deck'].cat.add_categories('Unknown').fillna('Unknown')
df = df.drop_duplicates()
le = LabelEncoder()
df['sex'] = le.fit_transform(df['sex'].astype(str))
scaler = StandardScaler()
df['fare'] = scaler.fit_transform(df[['fare']].fillna(0))
```

```
plt.figure (figsize = (8,6))
    corr().annot = True. cmap = 'coolwarm'
      linewidths = 0.5)
plt.title ("Correlation Heatmap")
plt.show.
```

Result:

Load the Titanic dataset and convert it into a dataframe is completed successfully.