```
import numpy as np
import pandas as pd
```

```
data=pd.read_csv('titanic.csv')
```

```
data.head()
```

	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

```
data.shape
```

```
(891, 12)
```

```
data.info
```

```
pandas.core.frame.DataFrame.info
def info(verbose: bool | None=None, buf: WriteBuffer[str] | None=None, max_cols: int | None=None,
memory_usage: bool | str | None=None, show_counts: bool | None=None) -> None

Print a concise summary of a DataFrame.

This method prints information about a DataFrame including
the index dtype and columns, non-null values and memory usage.

Parameters
```

```
data['Age_fwd']=data['Age'].fillna(method='ffill')
data['Age_bkwd']=data['Age'].fillna(method='bfill')
data.head(10)
```

```
/tmp/ipython-input-49-3989185386.py:1: FutureWarning: Series.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.
  data['Age_fwd']=data['Age'].fillna(method='ffill')
/tmp/ipython-input-49-3989185386.py:2: FutureWarning: Series.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.
  data['Age_bkwd']=data['Age'].fillna(method='bfill')
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Age_fwd	Age_bkwd
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S	22.0	22.0
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C	38.0	38.0
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S	26.0	26.0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S	35.0	35.0
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S	35.0	35.0
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	NaN	Q	35.0	54.0
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S	54.0	54.0
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	NaN	S	2.0	2.0
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	NaN	S	27.0	27.0
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	NaN	C	14.0	14.0

```
data['Cabin'].isnull().sum()
```

```
np.int64(687)
```

```
data['Cabin']=data['Cabin'].fillna(value='unknown',limit=5)
data.head(10)
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Age_fwd	Age_bkwd
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	unknown	S	22.0	22.0
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C	38.0	38.0
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	unknown	S	26.0	26.0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S	35.0	35.0
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	unknown	S	35.0	35.0
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	unknown	Q	35.0	54.0
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S	54.0	54.0
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	unknown	S	2.0	2.0
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	NaN	S	27.0	27.0
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	NaN	C	14.0	14.0

```
print(data['Cabin'].isnull().sum())
```

```
682
```

```
data_new=data.drop_duplicates()
data_new=data.reset_index(drop=True)
```

```
data_new.head()
```

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

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

```
encoder=LabelEncoder()
data_new['Sex'] = encoder.fit_transform(data_new['Sex'])
data_new
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Age_fwd	Age_bkwd
0	1	0	3	Braund, Mr. Owen Harris	1	22.0	1	0	A/5 21171	-0.502445	unknown	S	22.0	22.0
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	38.0	1	0	PC 17599	0.786845	C85	C	38.0	38.0
2	3	1	3	Heikkinen, Miss. Laina	0	26.0	0	0	STON/O2. 3101282	-0.488854	unknown	S	26.0	26.0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	35.0	1	0	113803	0.420730	C123	S	35.0	35.0
4	5	0	3	Allen, Mr. William Henry	1	35.0	0	0	373450	-0.486337	unknown	S	35.0	35.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	1	27.0	0	0	211536	-0.386671	NaN	S	27.0	27.0
887	888	1	1	Graham, Miss. Margaret Edith	0	19.0	0	0	112053	-0.044381	B42	S	19.0	19.0
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	0	NaN	1	2	W./C. 6607	-0.176263	NaN	S	19.0	26.0
889	890	1	1	Behr, Mr. Karl Howell	1	26.0	0	0	111369	-0.044381	C148	C	26.0	26.0
890	891	0	3	Dooley, Mr. Patrick	1	32.0	0	0	370376	-0.492378	NaN	Q	32.0	32.0

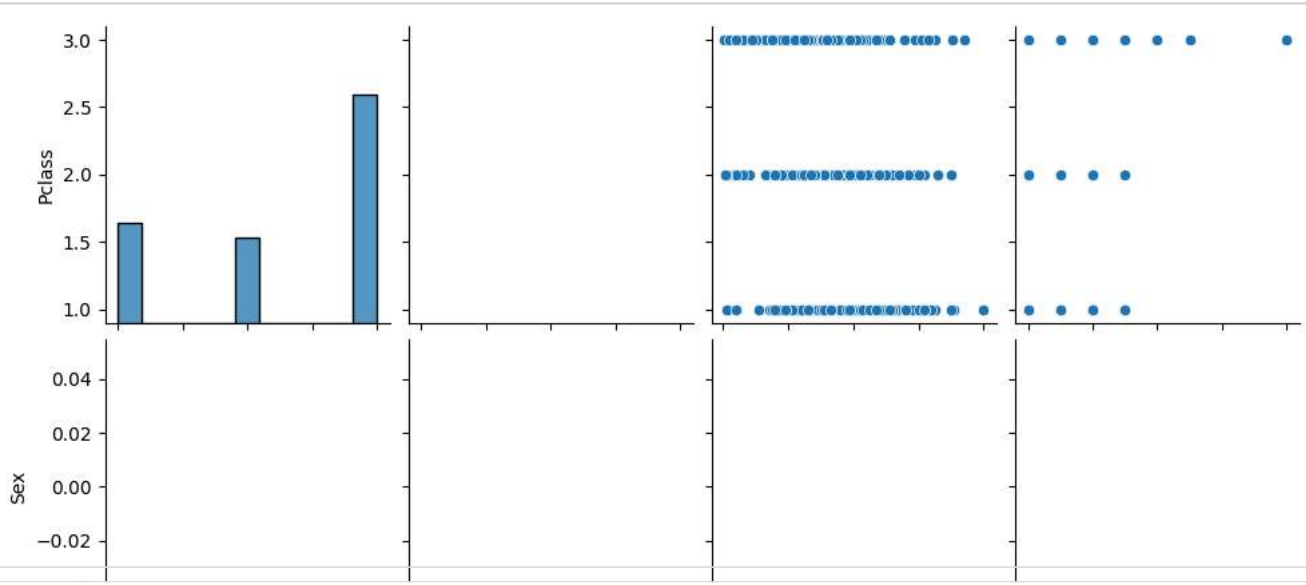
891 rows × 14 columns

```
scaler=StandardScaler()
data_new['Fare']=scaler.fit_transform(data_new[['Fare']])
data_new
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Age_fwd	Age_bkwd
0	1	0	3	Braund, Mr. Owen Harris	1	22.0	1	0	A/5 21171	-0.502445	unknown	S	22.0	22.0
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	38.0	1	0	PC 17599	0.786845	C85	C	38.0	38.0
2	3	1	3	Heikkinen, Miss. Laina	0	26.0	0	0	STON/O2. 3101282	-0.488854	unknown	S	26.0	26.0
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	35.0	1	0	113803	0.420730	C123	S	35.0	35.0
4	5	0	3	Allen, Mr. William Henry	1	35.0	0	0	373450	-0.486337	unknown	S	35.0	35.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	1	27.0	0	0	211536	-0.386671	NaN	S	27.0	27.0
887	888	1	1	Graham, Miss. Margaret Edith	0	19.0	0	0	112053	-0.044381	B42	S	19.0	19.0
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	0	NaN	1	2	W./C. 6607	-0.176263	NaN	S	19.0	26.0
889	890	1	1	Behr, Mr. Karl Howell	1	26.0	0	0	111369	-0.044381	C148	C	26.0	26.0
890	891	0	3	Dooley, Mr. Patrick	1	32.0	0	0	370376	-0.492378	NaN	Q	32.0	32.0

891 rows × 14 columns

```
features=['Pclass', 'Sex', 'Age', 'SibSp']
data_new['Sex']=data_new['Sex'].map({'male': 0, 'female': 1})
sns.pairplot(data_new[features])
plt.show()
```




```
selected_features = ['Pclass', 'Age', 'SibSp', 'Parch', 'Fare']
corr_matrix = data_new[selected_features].corr()
plt.figure(figsize=(8, 6))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt='.2f', linewidths=0.5)
plt.title('Correlation Heatmap of Selected Features')
plt.show()
```