

In [1]:

```
pip install pandas
```

Requirement already satisfied: pandas in c:\users\acer\anaconda3\lib\site-packages (1.4.2)  
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\acer\anaconda3\lib\site-packages (from pandas) (2.8.2)  
Requirement already satisfied: numpy>=1.18.5 in c:\users\acer\anaconda3\lib\site-packages (from pandas) (1.21.5)  
Requirement already satisfied: pytz>=2020.1 in c:\users\acer\anaconda3\lib\site-packages (from pandas) (2021.3)  
Requirement already satisfied: six>=1.5 in c:\users\acer\anaconda3\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)  
Note: you may need to restart the kernel to use updated packages.

In [2]:

```
pip install seaborn
```

Requirement already satisfied: seaborn in c:\users\acer\anaconda3\lib\site-packages (0.11.2)  
Requirement already satisfied: pandas>=0.23 in c:\users\acer\anaconda3\lib\site-packages (from seaborn) (1.4.2)  
Requirement already satisfied: scipy>=1.0 in c:\users\acer\anaconda3\lib\site-packages (from seaborn) (1.7.3)  
Requirement already satisfied: matplotlib>=2.2 in c:\users\acer\anaconda3\lib\site-packages (from seaborn) (3.5.1)  
Requirement already satisfied: numpy>=1.15 in c:\users\acer\anaconda3\lib\site-packages (from seaborn) (1.21.5)  
Requirement already satisfied: fonttools>=4.22.0 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (4.25.0)  
Requirement already satisfied: pillow>=6.2.0 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (9.0.1)  
Requirement already satisfied: cycler>=0.10 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (0.11.0)  
Requirement already satisfied: packaging>=20.0 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (21.3)  
Requirement already satisfied: pyparsing>=2.2.1 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (3.0.4)  
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (1.3.2)  
Requirement already satisfied: python-dateutil>=2.7 in c:\users\acer\anaconda3\lib\site-packages (from matplotlib>=2.2->seaborn) (2.8.2)  
Requirement already satisfied: pytz>=2020.1 in c:\users\acer\anaconda3\lib\site-packages (from pandas>=0.23->seaborn) (2021.3)  
Requirement already satisfied: six>=1.5 in c:\users\acer\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib>=2.2->seaborn) (1.16.0)  
Note: you may need to restart the kernel to use updated packages.

In [3]:

```
pip install scipy
```

Requirement already satisfied: scipy in c:\users\acer\anaconda3\lib\site-packages (1.7.3)

Requirement already satisfied: numpy<1.23.0,>=1.16.5 in c:\users\acer\anaconda3\lib\site-packages (from scipy) (1.21.5)

Note: you may need to restart the kernel to use updated packages.

In [4]:

```
#import python Libraries  
import numpy as np  
import pandas as pd
```

In [5]:

```
df=pd.read_csv(r"C:\Users\acer\Desktop\train.csv")
```

In [ ]:

```
#r"C:\Users\acer\Desktop\train.csv"
```

In [6]:

```
df.columns.values
```

Out[6]:

```
array(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',  
      'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked', 'familysize',  
      'Title start', 'Title end', 'left', 'right', 'Title'], dtype=object)
```

In [15]:

```
df.dtypes
```

Out[15]:

```
PassengerId      int64
Survived          int64
Pclass           int64
Name             object
Sex              object
Age             float64
SibSp            int64
Parch            int64
Ticket           object
Fare             float64
Cabin            object
Embarked         object
familysize       int64
Title start      int64
Title end        int64
left            object
right           object
Title            object
dtype: object
```

In [16]:

```
df.shape
```

Out[16]:

```
(891, 18)
```

In [17]:

```
df.shape[0]
```

Out[17]:

```
891
```

In [7]:

```
df.shape[1]
```

Out[7]:

```
18
```

In [18]:

```
df.head(5)
```

Out[18]:

ge	SibSp	Parch	Ticket	Fare	Cabin	Embarked	familysize	Title start	Title end	left	right
0.0	1	1	WE/P 5735	71.0000	B22	S	3	7	13	Crosby,	ifford
0.0	0	0	13213	35.5000	A26	C	1	16	21	Simoni	Alfons
0.0	0	0	113800	26.5500	NaN	S	1	5	10	Weir, C	. John
0.0	0	0	PC 17601	27.7208	NaN	C	1	10	15	Uruchur	nuel E
0.0	2	0	19928	90.0000	C78	Q	3	8	12	Minahan	Edward

In [19]:

```
df.tail(5)
```

Out[19]:

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
886	0	2	Kirkland, Rev. Charles Leonard	male	57.0	0	0	219533	12.3500
887	0	2	Harper, Rev. John	male	28.0	0	1	248727	33.0000
888	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000
889	1	1	Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")	male	49.0	1	0	PC 17485	56.9292
890	1	1	Roths, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	0	110152	86.5000

In [20]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 18 columns):
 #   Column             Non-Null Count  Dtype  
---  --
 0   PassengerId        891 non-null    int64  
 1   Survived           891 non-null    int64  
 2   Pclass             891 non-null    int64  
 3   Name               891 non-null    object  
 4   Sex               891 non-null    object  
 5   Age               714 non-null    float64 
 6   SibSp             891 non-null    int64  
 7   Parch             891 non-null    int64  
 8   Ticket            891 non-null    object  
 9   Fare             891 non-null    float64 
10   Cabin            204 non-null    object  
11   Embarked         889 non-null    object  
12   familysize       891 non-null    int64  
13   Title start      891 non-null    int64  
14   Title end        891 non-null    int64  
15   Title start      891 non-null    int64  
16   Title end        891 non-null    int64  
17   Title start      891 non-null    int64  
18   Title end        891 non-null    int64
```

In [22]:

```
df.Age.count()
```

Out[22]:

714

In [21]:

```
df.Age.value_counts()
```

Out[21]:

```
24.0    30
22.0    27
18.0    26
19.0    25
28.0    25
..
14.5     1
55.5     1
36.5     1
70.5     1
20.5     1
Name: Age, Length: 88, dtype: int64
```

In [23]:

```
df.isnull()
```

Out[23]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0		False	False	False	False	False	False	False	False	False	False
1		False	False	False	False	False	False	False	False	False	False
2		False	False	False	False	False	False	False	False	False	True
3		False	False	False	False	False	False	False	False	False	True
4		False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...
886		False	False	False	False	False	False	False	False	False	True
887		False	False	False	False	False	False	False	False	False	True
888		False	False	False	False	False	False	False	False	False	True
889		False	False	False	False	False	False	False	False	False	False
890		False	False	False	False	False	False	False	False	False	False

891 rows × 12 columns



In [26]:

```
df.notnull()
```

Out[26]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0		True	True	True	True	True	True	True	True	True	True	True
1		True	True	True	True	True	True	True	True	True	True	True
2		True	True	True	True	True	True	True	True	True	False	True
3		True	True	True	True	True	True	True	True	True	False	True
4		True	True	True	True	True	True	True	True	True	True	True
...	...	...	...	...	...	...	...	...	...	...	...	...
886		True	True	True	True	True	True	True	True	True	False	True
887		True	True	True	True	True	True	True	True	True	False	True
888		True	True	True	True	True	True	True	True	True	False	True
889		True	True	True	True	True	True	True	True	True	True	True
890		True	True	True	True	True	True	True	True	True	True	True

891 rows × 13 columns



In [27]:

```
df[df.notnull()]
```

Out[27]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	746	0	1	Crosby, Capt. Edward Gifford	male	70.0	1	1	WE/P 5735	71.0000
1	648	1	1	Simonius- Blumer, Col. Oberst Alfons	male	56.0	0	0	13213	35.5000
2	695	0	1	Weir, Col. John	male	60.0	0	0	113800	26.5500
3	31	0	1	Uruchurtu, Don. Manuel E	male	40.0	0	0	PC 17601	27.7208
4	246	0	1	Minahan, Dr. William Edward	male	44.0	2	0	19928	90.0000
...	...	...	...	...	...	...	...	...	...	...
886	627	0	2	Kirkland, Rev. Charles Leonard	male	57.0	0	0	219533	12.3500
887	849	0	2	Harper, Rev. John	male	28.0	0	1	248727	33.0000
888	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000
889	600	1	1	Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")	male	49.0	1	0	PC 17485	56.9292
890	760	1	1	Roths, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	0	110152	86.5000

891 rows × 18 columns



In [28]:

```
df[df.isnull()]
```

Out[28]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...
886	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
887	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
888	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
889	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
890	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

891 rows × 18 columns

In [30]:

```
df.describe()
```

Out[30]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200



In [31]:

```
df.sum()
```

C:\Users\acer\AppData\Local\Temp\ipykernel\_9984\1703867807.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric\_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
df.sum()
```

Out[31]:

```
PassengerId      397386
Survived          342
Pclass            2057
Name      Crosby, Capt. Edward GiffordSimonius-Blumer, C...
Sex      malemalemalemalemalemalemalemalefemale...
Age      21205.17
SibSp      466
Parch      340
Ticket      WE/P 573513213113800PC 17601199282901124427813...
Fare      28693.9493
familysize      1697
Title start      6990
Title end      11244
left      Crosby,SimoniuWeir, CUruchurMinahanMorawecPain...
right      iffordAlfons. Johnnuel EEdwardErnestAlfredr. M...
Title      Capt Col Col Don Dr Dr Dr Dr Dr Dr Dr Jonkhee...
dtype: object
```

In [32]:

```
df.mean()
```

C:\Users\acer\AppData\Local\Temp\ipykernel\_9984\3698961737.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric\_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
df.mean()
```

Out[32]:

```
PassengerId      446.000000
Survived          0.383838
Pclass            2.308642
Age              29.699118
SibSp            0.523008
Parch            0.381594
Fare             32.204208
familysize        1.904602
Title start        7.845118
Title end        12.619529
dtype: float64
```

In [33]:

```
df.corr()
```

Out[33]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare	far
PassengerId	1.000000	-0.005007	-0.035144	0.036847	-0.057527	-0.001652	0.012658	-0
Survived	-0.005007	1.000000	-0.338481	-0.077221	-0.035322	0.081629	0.257307	0
Pclass	-0.035144	-0.338481	1.000000	-0.369226	0.083081	0.018443	-0.549500	0
Age	0.036847	-0.077221	-0.369226	1.000000	-0.308247	-0.189119	0.096067	-0
SibSp	-0.057527	-0.035322	0.083081	-0.308247	1.000000	0.414838	0.159651	0
Parch	-0.001652	0.081629	0.018443	-0.189119	0.414838	1.000000	0.216225	C
Fare	0.012658	0.257307	-0.549500	0.096067	0.159651	0.216225	1.000000	0
familysize	-0.040143	0.016639	0.065997	-0.301914	0.890712	0.783111	0.217138	1
Title start	-0.024081	0.032149	0.005029	-0.011401	-0.046958	-0.076696	-0.044505	-0
Title end	-0.038914	0.214525	-0.014596	-0.194519	0.102819	0.075980	0.019896	0

In [34]:

```
df.cov()
```

Out[34]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	F
PassengerId	66231.000000	-0.626966	-7.561798	138.696504	-16.325843	-0.342697	161.883:
Survived	-0.626966	0.236772	-0.137703	-0.551296	-0.018954	0.032017	6.221'
Pclass	-7.561798	-0.137703	0.699015	-4.496004	0.076599	0.012429	-22.830
Age	138.696504	-0.551296	-4.496004	211.019125	-4.163334	-2.344191	73.849!
SibSp	-16.325843	-0.018954	0.076599	-4.163334	1.216043	0.368739	8.748'
Parch	-0.342697	0.032017	0.012429	-2.344191	0.368739	0.649728	8.661!
Fare	161.883369	6.221787	-22.830196	73.849030	8.748734	8.661052	2469.436!
familysize	-16.668539	0.013063	0.089028	-6.507525	1.584782	1.018467	17.409'
Title start	-14.230337	0.035921	0.009655	-0.378715	-0.118904	-0.141955	-5.078:
Title end	-25.237079	0.263058	-0.030753	-7.126687	0.285730	0.154337	2.491!

In [8]:

```
df.Age.mean()
```

Out[8]:

29.69911764705882

In [9]:

```
df.Fare.sum()
```

Out[9]:

28693.9493

In [10]:

```
df.fillna(0)
```

Out[10]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	746	0	1	Crosby, Capt. Edward Gifford	male	70.0	1	1	WE/P 5735	71.0000
1	648	1	1	Simonius- Blumer, Col. Oberst Alfons	male	56.0	0	0	13213	35.5000
2	695	0	1	Weir, Col. John	male	60.0	0	0	113800	26.5500
3	31	0	1	Uruchurtu, Don. Manuel E	male	40.0	0	0	PC 17601	27.7208
4	246	0	1	Minahan, Dr. William Edward	male	44.0	2	0	19928	90.0000
...	...	...	...	...	...	...	...	...	...	...
886	627	0	2	Kirkland, Rev. Charles Leonard	male	57.0	0	0	219533	12.3500
887	849	0	2	Harper, Rev. John	male	28.0	0	1	248727	33.0000
888	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000
889	600	1	1	Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")	male	49.0	1	0	PC 17485	56.9292
890	760	1	1	Roths, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	0	110152	86.5000

891 rows × 18 columns



In [11]:

```
df.isnull()
```

Out[11]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	True
3	False	False	False	False	False	False	False	False	False	False	True
4	False	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...
886	False	False	False	False	False	False	False	False	False	False	True
887	False	False	False	False	False	False	False	False	False	False	True
888	False	False	False	False	False	False	False	False	False	False	True
889	False	False	False	False	False	False	False	False	False	False	False
890	False	False	False	False	False	False	False	False	False	False	False

891 rows × 18 columns



In [48]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 18 columns):
#   Column          Non-Null Count  Dtype
---  -
0   PassengerId     891 non-null   int64
1   Survived        891 non-null   int64
2   Pclass          891 non-null   int64
3   Name            891 non-null   object
4   Sex             891 non-null   object
5   Age             714 non-null   float64
6   SibSp           891 non-null   int64
7   Parch           891 non-null   int64
8   Ticket          891 non-null   object
9   Fare            891 non-null   float64
10  Cabin           204 non-null   object
11  Embarked        889 non-null   object
12  familysize      891 non-null   int64
13  Title start     891 non-null   int64
14  Title end       891 non-null   int64
15  left            891 non-null   object
16  right           891 non-null   object
17  Title           891 non-null   object
dtypes: float64(2), int64(8), object(8)
memory usage: 125.4+ KB
```

In [49]:

```
df[df.isnull()]
```

Out[49]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...
886	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
887	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
888	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
889	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
890	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

891 rows × 18 columns



In [13]:

```
df1=df.drop(['Fare'],axis=1)
df1
```

Out[13]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Cabin	E
0	746	0	1	Crosby, Capt. Edward Gifford	male	70.0	1	1	WE/P 5735	B22	
1	648	1	1	Simonius- Blumer, Col. Oberst Alfons	male	56.0	0	0	13213	A26	
2	695	0	1	Weir, Col. John	male	60.0	0	0	113800	NaN	
3	31	0	1	Uruchurtu, Don. Manuel E	male	40.0	0	0	PC 17601	NaN	
4	246	0	1	Minahan, Dr. William Edward	male	44.0	2	0	19928	C78	
...	...	...	...	...	...	...	...	...	...	...	
886	627	0	2	Kirkland, Rev. Charles Leonard	male	57.0	0	0	219533	NaN	
887	849	0	2	Harper, Rev. John	male	28.0	0	1	248727	NaN	
888	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	NaN	
889	600	1	1	Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")	male	49.0	1	0	PC 17485	A20	
890	760	1	1	Roths, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	0	110152	B77	

891 rows × 17 columns



In [14]:

```
df1.drop(['Cabin', 'Parch'],axis=1)
```

Out[14]:

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Ticket	Embarked	family
0	746	0	1	Crosby, Capt. Edward Gifford	male	70.0	1	WE/P 5735	S
1	648	1	1	Simonius- Blumer, Col. Oberst Alfons	male	56.0	0	13213	C
2	695	0	1	Weir, Col. John	male	60.0	0	113800	S
3	31	0	1	Uruchurtu, Don. Manuel E	male	40.0	0	PC 17601	C
4	246	0	1	Minahan, Dr. William Edward	male	44.0	2	19928	Q
...	...	...	...	...	...	...	...	...	...
886	627	0	2	Kirkland, Rev. Charles Leonard	male	57.0	0	219533	Q
887	849	0	2	Harper, Rev. John	male	28.0	0	248727	S
888	887	0	2	Montvila, Rev. Juozas	male	27.0	0	211536	S
889	600	1	1	Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")	male	49.0	1	PC 17485	C
890	760	1	1	Roths, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	110152	S

891 rows × 15 columns





In [63]:

```
del df['Ticket']
```

In [64]:

```
df
```

Out[64]:

	PassengerId	Survived	Name	Sex	Age	SibSp	Parch	Fare	Cabin	familysize
0	746	0	Crosby, Capt. Edward Gifford	male	70.0	1	1	71.0000	B22	3
1	648	1	Simonius-Blumer, Col. Oberst Alfons	male	56.0	0	0	35.5000	A26	1
2	695	0	Weir, Col. John	male	60.0	0	0	26.5500	NaN	1
3	31	0	Uruchurtu, Don. Manuel E	male	40.0	0	0	27.7208	NaN	1
4	246	0	Minahan, Dr. William Edward	male	44.0	2	0	90.0000	C78	3
...	...	...	...	...	...	...	...	...	...	...
886	627	0	Kirkland, Rev. Charles Leonard	male	57.0	0	0	12.3500	NaN	1
887	849	0	Harper, Rev. John	male	28.0	0	1	33.0000	NaN	2
888	887	0	Montvila, Rev. Juozas	male	27.0	0	0	13.0000	NaN	1
889	600	1	Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")	male	49.0	1	0	56.9292	A20	2
890	760	1	Roths, the Countess. of (Lucy Noel Martha Dye...	female	33.0	0	0	86.5000	B77	1

891 rows × 15 columns



In [65]:

df.T

Out[65]:

	0	1	2	3	4	5	6	7	
PassengerId	746	648	695	31	246	318	399	633	
Survived	0	1	0	0	0	0	0	1	
Name	Crosby, Capt. Edward Gifford	Simonius-Blumer, Col. Oberst Alfons	Weir, Col. John	Uruchurtu, Don. Manuel E	Minahan, Dr. William Edward	Moraweck, Dr. Ernest	Pain, Dr. Alfred	Stahelin-Maeglin, Dr. Max	Fraue Dr. I W
Sex	male	male	male	male	male	male	male	male	
Age	70.0	56.0	60.0	40.0	44.0	54.0	23.0	32.0	
SibSp	1	0	0	0	2	0	0	0	
Parch	1	0	0	0	0	0	0	0	
Fare	71.0	35.5	26.55	27.7208	90.0	14.0	10.5	30.5	1
Cabin	B22	A26	NaN	NaN	C78	NaN	NaN	B50	
familysize	3	1	1	1	3	1	1	1	
Title start	7	16	5	10	8	9	5	17	
Title end	13	21	10	15	12	13	9	21	
left	Crosby,	Simoniu	Weir, C	Uruchur	Minahan	Morawec	Pain, D	Staheli	Fr
right	ifford	Alfons	John	nuel E	Edward	Ernest	Alfred	r. Max	
Title	Capt	Col	Col	Don	Dr	Dr	Dr	Dr	

15 rows × 891 columns



In [66]:

```
df.transpose()
```

Out[66]:

	0	1	2	3	4	5	6	7	
PassengerId	746	648	695	31	246	318	399	633	
Survived	0	1	0	0	0	0	0	1	
Name	Crosby, Capt. Edward Gifford	Simonius-Blumer, Col. Oberst Alfons	Weir, Col. John	Uruchurtu, Don. Manuel E	Minahan, Dr. William Edward	Moraweck, Dr. Ernest	Pain, Dr. Alfred	Stahelin-Maeglin, Dr. Max	Fraue Dr. I W
Sex	male	male	male	male	male	male	male	male	
Age	70.0	56.0	60.0	40.0	44.0	54.0	23.0	32.0	
SibSp	1	0	0	0	2	0	0	0	
Parch	1	0	0	0	0	0	0	0	
Fare	71.0	35.5	26.55	27.7208	90.0	14.0	10.5	30.5	1
Cabin	B22	A26	NaN	NaN	C78	NaN	NaN	B50	
familysize	3	1	1	1	3	1	1	1	
Title start	7	16	5	10	8	9	5	17	
Title end	13	21	10	15	12	13	9	21	
left	Crosby,	Simoniu	Weir, C	Uruchur	Minahan	Morawec	Pain, D	Staheli	Fr
right	ifford	Alfons	John	nuel E	Edward	Ernest	Alfred	r. Max	
Title	Capt	Col	Col	Don	Dr	Dr	Dr	Dr	

15 rows × 891 columns



In [67]:

```
df['new']=12
df
```

Out[67]:

Sex	Age	SibSp	Parch	Fare	Cabin	familysize	Title start	Title end	left	right	Title
male	70.0	1	1	71.0000	B22	3	7	13	Crosby,	ifford	Capt
male	56.0	0	0	35.5000	A26	1	16	21	Simoniu	Alfons	Col
male	60.0	0	0	26.5500	NaN	1	5	10	Weir, C	. John	Col
male	40.0	0	0	27.7208	NaN	1	10	15	Uruchur	nuel E	Don
male	44.0	2	0	90.0000	C78	3	8	12	Minahan	Edward	Dr
...	...	...	...	...	...	...	...	...	...	...	...
male	57.0	0	0	12.3500	NaN	1	9	14	Kirklan	eonard	Rev
male	28.0	0	1	33.0000	NaN	2	7	12	Harper,	. John	Rev
male	27.0	0	0	13.0000	NaN	1	9	14	Montvil	Juozas	Rev
male	49.0	1	0	56.9292	A20	2	12	17	Duff Go	rgan")	Sir
female	33.0	0	0	86.5000	B77	1	7	21	Rothes,	wards)	the Countess



In [68]:

```
df['new'] = [3.0,1.3,2.2,0.8,1.1]  
df
```

```

-----
-
ValueError                                Traceback (most recent call las
t)
Input In [68], in <cell line: 1>()
----> 1 df['new'] = [3.0,1.3,2.2,0.8,1.1]
      2 df

File ~\anaconda3\lib\site-packages\pandas\core\frame.py:3655, in DataFram
e._setitem__(self, key, value)
    3652     self._setitem_array([key], value)
    3653 else:
    3654     # set column
-> 3655     self._set_item(key, value)

File ~\anaconda3\lib\site-packages\pandas\core\frame.py:3832, in DataFram
e._set_item(self, key, value)
    3822 def _set_item(self, key, value) -> None:
    3823     """
    3824     Add series to DataFrame in specified column.
    3825
    3826     (...)
    3830     ensure homogeneity.
    3831     """
-> 3832     value = self._sanitize_column(value)
    3834     if (
    3835         key in self.columns
    3836         and value.ndim == 1
    3837         and not is_extension_array_dtype(value)
    3838     ):
    3839         # broadcast across multiple columns if necessary
    3840         if not self.columns.is_unique or isinstance(self.columns,
MultiIndex):

File ~\anaconda3\lib\site-packages\pandas\core\frame.py:4535, in DataFram
e._sanitize_column(self, value)
    4532     return _reindex_for_setitem(value, self.index)
    4534 if is_list_like(value):
-> 4535     com.require_length_match(value, self.index)
    4536 return sanitize_array(value, self.index, copy=True, allow_2d=True)

File ~\anaconda3\lib\site-packages\pandas\core\common.py:557, in require_l
ength_match(data, index)
    553 """
    554 Check the length of data matches the length of the index.
    555 """
    556 if len(data) != len(index):
--> 557     raise ValueError(
    558         "Length of values "
    559         f"({len(data)}) "
    560         "does not match length of index "
    561         f"({len(index)})"
    562     )

```

**ValueError:** Length of values (5) does not match length of index (891)

