```
import pandas as pd
```

```
df=pd.read_csv('/content/titanic.csv')
```

```
Start coding or generate with AI.
```

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

	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

Next steps: Generate code with df | View recommended plots | New interactive sheet

Generate | a slider using jupyter widgets | Close

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

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.00	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.00	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.45	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.00	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.75	NaN	Q

```
df.isna()
```

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

891 rows × 12 columns

```
df.describe()
```

	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

Generate print hello world using rot13 Close

```
df['Age']
```

	Age
0	22.0
1	38.0
2	26.0
3	35.0
4	35.0
...	...
886	27.0
887	19.0
888	NaN
889	26.0
890	32.0

891 rows × 1 columns

dtype: float64

```
df['Survived']
```

	Survived
0	0
1	1
2	1
3	1
4	0
...	...
886	0
887	1
888	0
889	1
890	0

891 rows × 1 columns

dtype: int64

```
df['Age'].fillna(0)
```

	Age
1	0.0
3	0.0
6	0.0
10	0.0
11	0.0
...	...
871	0.0
872	0.0
879	0.0
887	0.0
889	0.0

202 rows × 1 columns

dtype: float64

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

```
29.69911764705882
```

```
df['Age'].median()
```

```
28.0
```

Generate create a dataframe with 2 columns and 10 rows Close

```
df['Age'].mode()
```

	Age
0	24.0

dtype: float64

```
df['Age'].fillna(df['Age'].mean())
```

	Age
0	22.000000
1	38.000000
2	26.000000
3	35.000000
4	35.000000
...	...
886	27.000000
887	19.000000
888	29.699118
889	26.000000
890	32.000000

891 rows × 1 columns

dtype: float64

```
df['Age'].fillna(df['Age'].mean(),inplace=True)
```

```
df['Age'].isna()
```

	Age
0	False
1	False
2	False
3	False
4	False
...	...
886	False
887	False
888	False
889	False
890	False

891 rows × 1 columns

dtype: bool

```
df.dropna(inplace=True)
```

```
df
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000	G6	S
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	C103	S
...	...	...	...	...	...	...	...	...	...	...	...	...
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.5542	D35	S
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.0000	B51 B53 B55	S
879	880	1	1	Potter, Mrs. Thomas Jr (Lily	female	56.0	0	1	11767	83.1583	C50	C

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
df['Age'].max()
```

```
80.0
```

```
df['Age'].astype
```

```
pandas.core.generic.NDFrame.astype
def astype(dtype, copy: bool_t | None=None, errors: IgnoreRaise='raise') -> Self
```

```
>>> ser_date = pd.Series(pd.date_range('20200101', periods=3))
>>> ser_date
0   2020-01-01
1   2020-01-02
2   2020-01-03
dtype: datetime64[ns]
```

```
df['Age']=(df['Age']- df['Age'].min())//(df['Age'].max() - df['Age'].min())
```

```
df['Age']
```

	Age
1	0.0
3	0.0
6	0.0
10	0.0
11	0.0
...	...
871	0.0
872	0.0
879	0.0
887	0.0
889	0.0

Generate print hello world using rot13 Close

```
df_scaled = df.copy()

numeric_columns = df.select_dtypes(include=['float64', 'int64']).columns
df_scaled[numeric_columns] = (df[numeric_columns] - df[numeric_columns].min()) / (df[numeric_columns].max() - df[numeric_columns].min())

print(df_scaled.head())
```

```
     PassengerId  Survived  Pclass  \
1       0.000000       1.0     0.0
3       0.002252       1.0     0.0
6       0.005631       0.0     0.0
10      0.010135       1.0     1.0
11      0.011261       1.0     0.0

                                                 Name     Sex  Age     SibSp  \
1   Cumings, Mrs. John Bradley (Florence Briggs Th...  female  0.0  0.333333
3        Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  0.0  0.333333
6                             McCarthy, Mr. Timothy J    male  0.0  0.000000
10                   Sandstrom, Miss. Marguerite Rut  female  0.0  0.333333
11                          Bonnell, Miss. Elizabeth  female  0.0  0.000000

     Parch    Ticket      Fare Cabin Embarked
1     0.00  PC 17599  0.139136   C85        C
3     0.00    113803  0.103644  C123        S
6     0.00     17463  0.101229   E46        S
10    0.25   PP 9549  0.032596    G6        S
11    0.00    113783  0.051822  C103        S
```

```
from sklearn.preprocessing import StandardScaler

numeric_columns = df.select_dtypes(include=['float64', 'int64']).columns

scaler = StandardScaler()

df_scaled = scaler.fit_transform(df[numeric_columns])

df_scaled_df = pd.DataFrame(df_scaled, columns=numeric_columns)

df_scaled_df.head()
```

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
0	-1.820641	0.712364	-0.375823	-0.070535	0.881589	-0.603158	-0.064633
1	-1.812612	0.712364	-0.375823	-0.070535	0.881589	-0.603158	-0.308460
2	-1.800568	-1.403776	-0.375823	-0.070535	-0.708420	-0.603158	-0.325054
3	-1.784509	0.712364	3.419989	-0.070535	0.881589	0.765807	-0.796561
4	-1.780494	0.712364	-0.375823	-0.070535	-0.708420	-0.603158	-0.664479

Next steps: Generate code with df_scaled_df View recommended plots New interactive sheet