```
import pandas as pd
dataframe = pd.DataFrame()

url = "/content/titanic.csv"
df = pd.read_csv(url)
# show first two rows
print(df.head (2))
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1

                                                Name     Sex   Age
SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0
1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1

   Parch     Ticket     Fare Cabin Embarked
0      0  A/5 21171   7.2500   NaN        S
1      0   PC 17599  71.2833   C85        C
```

```
# show dimensions
print("Dimensions: {}".format(df.shape))
```

```
Dimensions: (891, 12)
```

```
# show statistics
df.describe()
```

```
       PassengerId    Survived      Pclass         Age       SibSp  \
count   891.000000  891.000000  891.000000  714.000000  891.000000
mean    446.000000    0.383838    2.308642   29.699118    0.523008
std     257.353842    0.486592    0.836071   14.526497    1.102743
min       1.000000    0.000000    1.000000    0.420000    0.000000
25%     223.500000    0.000000    2.000000   20.125000    0.000000
50%     446.000000    0.000000    3.000000   28.000000    0.000000
75%     668.500000    1.000000    3.000000   38.000000    1.000000
max     891.000000    1.000000    3.000000   80.000000    8.000000

            Parch        Fare
count  891.000000  891.000000
mean     0.381594   32.204208
std      0.806057   49.693429
min      0.000000    0.000000
25%      0.000000    7.910400
50%      0.000000   14.454200
75%      0.000000   31.000000
max      6.000000  512.329200
```

```
# select the first row
print(df.iloc[0])
```

```
PassengerId                          1
Survived                             0
Pclass                               3
Name           Braund, Mr. Owen Harris
Sex                               male
Age                               22.0
SibSp                                1
Parch                                0
Ticket                       A/5 21171
Fare                              7.25
Cabin                              NaN
Embarked                             S
Name: 0, dtype: object
```

```
# select three rows
print(df.iloc[1:4])
```

```
   PassengerId  Survived  Pclass  \
1            2         1       1
2            3         1       3
3            4         1       1

                                                Name     Sex   Age
SibSp  \
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1
2                             Heikkinen, Miss. Laina  female  26.0
0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
1

   Parch             Ticket      Fare Cabin Embarked
1      0           PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
3      0             113803  53.1000  C123        S
```

```
# all rows up to and including the fourth row
print(df.iloc[:4])
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1

                                                Name     Sex   Age
SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0
```

```
1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1
2                             Heikkinen, Miss. Laina  female  26.0
0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
1

   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
3      0            113803  53.1000  C123        S

# set index
df = df.set_index(df ['Name'])

# show row
df.loc['Braund, Mr. Owen Harris']

PassengerId                          1
Survived                             0
Pclass                               3
Name           Braund, Mr. Owen Harris
Sex                               male
Age                               22.0
SibSp                                1
Parch                                0
Ticket                       A/5 21171
Fare                              7.25
Cabin                              NaN
Embarked                             S
Name: Braund, Mr. Owen Harris, dtype: object

# select top two rows where column 'sex' is 'female'
df[df ['Sex'] == 'female'].head(2)

                                                    PassengerId
Survived  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...            2
1
Heikkinen, Miss. Laina                                        3
1

                                                    Pclass  \
Name
Cumings, Mrs. John Bradley (Florence Briggs Tha...       1
Heikkinen, Miss. Laina                                   3
```

```
Name  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...
Heikkinen, Miss. Laina
Heikkinen, Miss. Laina

                                                       Sex   Age
SibSp  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...  female  38.0
1
Heikkinen, Miss. Laina                              female  26.0
0

                                                    Parch
Ticket  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...      0           PC
17599
Heikkinen, Miss. Laina                                  0  STON/O2.
3101282

                                                       Fare Cabin
Embarked
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...  71.2833   C85
C
Heikkinen, Miss. Laina                               7.9250   NaN
S

# multiple conditions
df [(df['Sex'] == 'female') & (df ['Age'] >= 65)]

Empty DataFrame
Columns: [PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch,
Ticket, Fare, Cabin, Embarked]
Index: []

# replace any instance of 'female' with Woman
df['Sex'].replace('female', 'Woman').head (2)

Name
Braund, Mr. Owen Harris                                   male
Cumings, Mrs. John Bradley (Florence Briggs Thayer)     Woman
Name: Sex, dtype: object
```

```
# replace any instance of 'female' with Woman
df['Sex'].replace(['female', 'male'], ['Woman', 'Man']).head(5)

Name
Braund, Mr. Owen Harris                                  Man
Cumings, Mrs. John Bradley (Florence Briggs Thayer)    Woman
Heikkinen, Miss. Laina                                 Woman
Futrelle, Mrs. Jacques Heath (Lily May Peel)           Woman
Allen, Mr. William Henry                                 Man
Name: Sex, dtype: object

df.replace(1,"One").head(2)

                                                   PassengerId
Survived  \
Name

Braund, Mr. Owen Harris                                    One
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...           2
One

                                                   Pclass  \
Name
Braund, Mr. Owen Harris                                 3
Cumings, Mrs. John Bradley (Florence Briggs Tha...    One


Name  \
Name

Braund, Mr. Owen Harris
Braund, Mr. Owen Harris
Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...

                                                       Sex   Age SibSp
Parch  \
Name

Braund, Mr. Owen Harris                                 male  22.0   One
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...  female  38.0   One
0

                                                       Ticket     Fare
Cabin  \
Name

Braund, Mr. Owen Harris                             A/5 21171   7.2500
NaN
```

```
Cumings, Mrs. John Bradley (Florence Briggs Tha...   PC 17599  71.2833
C85

                                                   Embarked
Name
Braund, Mr. Owen Harris                                   S
Cumings, Mrs. John Bradley (Florence Briggs Tha...        C
```

```
df.rename(columns={'PClass': 'Passenger Class'}).head (2)
```

```
                                                    PassengerId
Survived  \
Name

Braund, Mr. Owen Harris                                       1
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...            2
1

                                                    Pclass  \
Name
Braund, Mr. Owen Harris                                  3
Cumings, Mrs. John Bradley (Florence Briggs Tha...       1


Name  \
Name

Braund, Mr. Owen Harris
Braund, Mr. Owen Harris
Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...

                                                       Sex   Age
SibSp  \
Name

Braund, Mr. Owen Harris                               male  22.0
1
Cumings, Mrs. John Bradley (Florence Briggs Tha...  female  38.0
1

                                                    Parch     Ticket
Fare  \
Name

Braund, Mr. Owen Harris                                 0  A/5 21171
7.2500
Cumings, Mrs. John Bradley (Florence Briggs Tha...      0   PC 17599
71.2833
```

```
                                                   Cabin Embarked
Name
Braund, Mr. Owen Harris                              NaN        S
Cumings, Mrs. John Bradley (Florence Briggs Tha...   C85        C

df.rename(columns={'PClass': 'Passenger Class', 'Sex':
'Gender'}).head(2)

                                                    PassengerId
Survived  \
Name

Braund, Mr. Owen Harris                                       1
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...            2
1

                                                    Pclass  \
Name
Braund, Mr. Owen Harris                                  3
Cumings, Mrs. John Bradley (Florence Briggs Tha...       1


Name  \
Name

Braund, Mr. Owen Harris
Braund, Mr. Owen Harris
Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...

                                                    Gender   Age
SibSp  \
Name

Braund, Mr. Owen Harris                               male  22.0
1
Cumings, Mrs. John Bradley (Florence Briggs Tha...  female  38.0
1

                                                    Parch     Ticket
Fare  \
Name

Braund, Mr. Owen Harris                                 0  A/5 21171
7.2500
Cumings, Mrs. John Bradley (Florence Briggs Tha...      0   PC 17599
71.2833

                                                   Cabin Embarked
Name
```

```
Braund, Mr. Owen Harris                              NaN        S
Cumings, Mrs. John Bradley (Florence Briggs Tha...   C85        C
```

```
print('Maximum: {}'.format(df ['Age'].max()))
print('Minimum: {}'.format(df['Age'].min()))
print('Mean: {}'.format(df ['Age'].mean()))
print('Sum: {}'.format(df ['Age'].sum()))
print('Count: {}'.format(df ['Age'].count()))
```

```
Maximum: 80.0
Minimum: 0.42
Mean: 29.69911764705882
Sum: 21205.17
Count: 714
```

```
print("Variance: {}".format(df.var()))
print("Standard Deviation: {}".format(df.std()))
print("Kurtosis: {}".format(df.kurt()))
print("Skewness: {}".format(df.skew()))
```

```
Variance: PassengerId    66231.000000
Survived           0.236772
Pclass             0.699015
Age              211.019125
SibSp              1.216043
Parch              0.649728
Fare            2469.436846
dtype: float64
Standard Deviation: PassengerId    257.353842
Survived         0.486592
Pclass           0.836071
Age             14.526497
SibSp            1.102743
Parch            0.806057
Fare            49.693429
dtype: float64
Kurtosis: PassengerId    -1.200000
Survived       -1.775005
Pclass         -1.280015
Age             0.178274
SibSp          17.880420
Parch           9.778125
Fare           33.398141
dtype: float64
Skewness: PassengerId    0.000000
Survived       0.478523
Pclass        -0.630548
Age            0.389108
SibSp          3.695352
Parch          2.749117
```

```
Fare           4.787317
dtype: float64
```

```
<ipython-input-23-bd8bae3b2eee>:1: FutureWarning: The default value of
numeric_only in DataFrame.var is deprecated. In a future version, it
will default to False. In addition, specifying 'numeric_only=None' is
deprecated. Select only valid columns or specify the value of
numeric_only to silence this warning.
  print("Variance: {}".format(df.var()))
<ipython-input-23-bd8bae3b2eee>:2: FutureWarning: The default value of
numeric_only in DataFrame.std is deprecated. In a future version, it
will default to False. In addition, specifying 'numeric_only=None' is
deprecated. Select only valid columns or specify the value of
numeric_only to silence this warning.
  print("Standard Deviation: {}".format(df.std()))
<ipython-input-23-bd8bae3b2eee>:3: FutureWarning: The default value of
numeric_only in DataFrame.kurt is deprecated. In a future version, it
will default to False. In addition, specifying 'numeric_only=None' is
deprecated. Select only valid columns or specify the value of
numeric_only to silence this warning.
  print("Kurtosis: {}".format(df.kurt()))
<ipython-input-23-bd8bae3b2eee>:4: FutureWarning: The default value of
numeric_only in DataFrame.skew is deprecated. In a future version, it
will default to False. In addition, specifying 'numeric_only=None' is
deprecated. Select only valid columns or specify the value of
numeric_only to silence this warning.
  print("Skewness: {}".format(df.skew()))
```

```
# unique will return an array of all unique values in a column
df ['Sex'].unique()
```

```
array(['male', 'female'], dtype=object)
```

```
# value_counts will display all unique values with the number of times
each value appears
df['Sex'].value_counts()
```

```
male      577
female    314
Name: Sex, dtype: int64
```

```
# select missing values, show 2 rows
df[df ['Age'].isnull()].head(2)
```

```
                              PassengerId  Survived  Pclass  \
Name
Moran, Mr. James                        6         0       3
Williams, Mr. Charles Eugene           18         1       2

                                                      Name   Sex  Age
SibSp  \
Name

Moran, Mr. James                          Moran, Mr. James  male  NaN
0
Williams, Mr. Charles Eugene  Williams, Mr. Charles Eugene  male  NaN
0

                              Parch  Ticket     Fare Cabin Embarked
Name
Moran, Mr. James                  0  330877   8.4583   NaN        Q
Williams, Mr. Charles Eugene      0  244373  13.0000   NaN        S

df.drop('Age',axis=1).head(2)

                                                    PassengerId
Survived  \
Name

Braund, Mr. Owen Harris                                       1
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...            2
1

                                                    Pclass  \
Name
Braund, Mr. Owen Harris                                  3
Cumings, Mrs. John Bradley (Florence Briggs Tha...       1


Name  \
Name

Braund, Mr. Owen Harris
Braund, Mr. Owen Harris
Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...

                                                       Sex  SibSp
Parch  \
Name

Braund, Mr. Owen Harris                               male      1
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...  female      1
0

                                                       Ticket     Fare
Cabin  \
Name

Braund, Mr. Owen Harris                             A/5 21171   7.2500
```

```
NaN
Cumings, Mrs. John Bradley (Florence Briggs Tha...   PC 17599  71.2833
C85

                                                   Embarked
Name
Braund, Mr. Owen Harris                                   S
Cumings, Mrs. John Bradley (Florence Briggs Tha...        C

# create new dataframe excluding the rows you want to delete
df[df ['Sex']!='male'].head(2)

                                                    PassengerId
Survived  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...            2
1
Heikkinen, Miss. Laina                                        3
1

                                                    Pclass  \
Name
Cumings, Mrs. John Bradley (Florence Briggs Tha...       1
Heikkinen, Miss. Laina                                   3


Name  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...
Heikkinen, Miss. Laina
Heikkinen, Miss. Laina

                                                       Sex   Age
SibSp  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...  female  38.0
1
Heikkinen, Miss. Laina                              female  26.0
0

                                                    Parch
Ticket  \
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...      0          PC
17599
Heikkinen, Miss. Laina                                  0  STON/O2.
```

```
3101282

                                                       Fare Cabin
Embarked
Name

Cumings, Mrs. John Bradley (Florence Briggs Tha...  71.2833   C85
C
Heikkinen, Miss. Laina                               7.9250   NaN
S

# delete a row by matching a unique value
df [df['Name'] != 'Allison, Miss Helen Loraine'].head(2)

                                                    PassengerId
Survived  \
Name

Braund, Mr. Owen Harris                                       1
0
Cumings, Mrs. John Bradley (Florence Briggs Tha...            2
1

                                                    Pclass  \
Name
Braund, Mr. Owen Harris                                  3
Cumings, Mrs. John Bradley (Florence Briggs Tha...       1


Name  \
Name

Braund, Mr. Owen Harris
Braund, Mr. Owen Harris
Cumings, Mrs. John Bradley (Florence Briggs Tha...  Cumings, Mrs. John
Bradley (Florence Briggs Th...

                                                       Sex   Age
SibSp  \
Name

Braund, Mr. Owen Harris                               male  22.0
1
Cumings, Mrs. John Bradley (Florence Briggs Tha...  female  38.0
1

                                                    Parch     Ticket
Fare  \
Name

Braund, Mr. Owen Harris                                 0  A/5 21171
```

```
7.2500
Cumings, Mrs. John Bradley (Florence Briggs Tha...      0   PC 17599
71.2833

                                                   Cabin Embarked
Name
Braund, Mr. Owen Harris                              NaN        S
Cumings, Mrs. John Bradley (Florence Briggs Tha...   C85        C

# delete a row by index
[df.index==0]

[array([False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False,
False])]
```

```
df.groupby('Sex').mean()
```

```
<ipython-input-33-fe69c9b95872>:1: FutureWarning: The default value of
numeric_only in DataFrameGroupBy.mean is deprecated. In a future
version, numeric_only will default to False. Either specify
numeric_only or select only columns which should be valid for the
function.
  df.groupby('Sex').mean()
```

```
        PassengerId  Survived    Pclass        Age     SibSp     Parch
\
Sex

female   431.028662  0.742038  2.159236  27.915709  0.694268  0.649682

male     454.147314  0.188908  2.389948  30.726645  0.429809  0.235702


             Fare
Sex
female  44.479818
male    25.523893
```

```
df.groupby('Survived')['Name'].count()
```

```
Survived
0    549
1    342
Name: Name, dtype: int64
```

```
df.groupby(['Sex', 'Survived'])['Age'].mean()
```

```
Sex     Survived
female  0           25.046875
        1           28.847716
male    0           31.618056
```

```
        1           27.276022
Name: Age, dtype: float64
```

```
# for .. in.. loop
for name in df ['Name'][0:2]:
  print(name.upper())
```

```
BRAUND, MR. OWEN HARRIS
CUMINGS, MRS. JOHN BRADLEY (FLORENCE BRIGGS THAYER)
```

```
# list comprehension (more "pythonic")
[name.upper() for name in df ['Name'] [0:2]]
```

```
['BRAUND, MR. OWEN HARRIS',
 'CUMINGS, MRS. JOHN BRADLEY (FLORENCE BRIGGS THAYER)']
```

```
def uppercase(x):
  return x.upper()

df['Name'].apply(uppercase)[0:2]
```

```
Name
Braund, Mr. Owen Harris
BRAUND, MR. OWEN HARRIS
Cumings, Mrs. John Bradley (Florence Briggs Thayer)    CUMINGS, MRS.
JOHN BRADLEY (FLORENCE BRIGGS TH...
Name: Name, dtype: object
```

```
df.groupby('Sex').apply(lambda x: x.count())
```

```
        PassengerId  Survived  Pclass  Name  Sex  Age  SibSp  Parch
Ticket  \
Sex

female          314       314     314   314  314  261    314    314
314
male            577       577     577   577  577  453    577    577
577

        Fare  Cabin  Embarked
Sex
female   314     97       312
male     577    107       577
```