

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File Explorer: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
[2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

Matplotlib is building the font cache; this may take a moment.

[3]: data = pd.read_csv('Titanic-Dataset.csv')

[4]: 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

```
[5]: data.dtypes

[5]: PassengerId    int64
Survived         int64
Pclass           int64
Name             object
Sex              object
Age              float64
SibSp            int64
Parch            int64
Ticket           object
Fare             float64
Cabin            object
Embarked         object
dtype: object
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle Mode: Command Ln 1, Col 12 Intro.ipynb 2 32°C Haze 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File Explorer: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
[10]: data['Sex'].value_counts()

[10]: Sex
      male    577
      female  314
      Name: count, dtype: int64

[11]: print('% of male who survived', 100*np.mean(data['Survived'][data['Sex']=='male']))
      print('% of female who survived', 100*np.mean(data['Survived'][data['Sex']=='female']))

      % of male who survived 18.890814558058924
      % of female who survived 74.28382165605095

[12]: np.mean(data['Survived'][data['Sex']=='male'])

[12]: np.float64(0.18890814558058924)

[13]: data['Pclass'].value_counts()

[13]: Pclass
      3    491
      1    216
      2    184
      Name: count, dtype: int64

[14]: print('% of passengers who survived in first class', 100*np.mean(data['Survived'][data['Pclass'] == 1]))
      print('% of passengers who survived in second class', 100*np.mean(data['Survived'][data['Pclass'] == 2]))
      print('% of passengers who survived in third class', 100*np.mean(data['Survived'][data['Pclass'] == 3]))

      % of passengers who survived in first class 62.96296296296296
      % of passengers who survived in second class 47.28260869565217
      % of passengers who survived in third class 24.236252545824847

[15]: data[["Pclass", "Survived"]].groupby(['Pclass']).mean()

[15]:      Survived
Pclass
3      0.629629
1      0.472827
2      0.242363
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle

Windows taskbar: Type here to search | 32°C Haze | 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File browser: / notebooks /

- Untitled Folder
- Untitled Folder1
- Untitled Folder2
- Untitled Folder3
- Untitled Folder4
- Untitled Folder5
- Intro (copy).ipynb
- Intro.ipynb
- Lorenz.ipynb
- sqlite.ipynb
- Titanic-Dataset.csv
- titatic
- Untitled3.ipynb
- Untitled4.ipynb

Intro.ipynb

Code

```
[15]: Survived
      Pclass
      1 0.629630
      2 0.472826
      3 0.242363

[16]: data.shape
[16]: (891, 12)

[17]: data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 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
dtypes: float64(2), int64(5), object(5)
memory usage: 66.2+ KB

[18]: data['Age'].value counts()
```

Activate Windows
Go to Settings to activate Windows

Simple 0 1 Python (Pyodide) | Idle

Mode: Command Ln 1, Col 12 Intro.ipynb 2

32°C Haze 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File browser: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
1 female
2 female
3 female
4 male
...
886 male
887 female
888 female
889 male
890 male
Name: Sex, Length: 891, dtype: object
```

[21]:

```
df2 = data.copy()
df2['Sex'] = data['Sex'].apply(lambda x: 1 if x == 'male' else 0)
df2['Sex']
```

[21]:

```
0    1
1    0
2    0
3    0
4    1
..
886  1
887  0
888  0
889  1
890  1
Name: Sex, Length: 891, dtype: int64
```

[22]:

```
df2 = data.copy() #dataframe copy
df2.isnull().sum()
```

[22]:

```
PassengerId    0
Survived        0
Pclass          0
Name            0
Sex             0
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle

Mode: Command Ln 1, Col 12 Intro.ipynb 2

32°C Haze 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P, Data-Analysis/Data-Preprocess, JupyterLite, how to take screenshot in lapto

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File browser: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
[22]: PassengerId    0
      Survived    0
      Pclass     0
      Name       0
      Sex        0
      Age       177
      SibSp      0
      Parch      0
      Ticket     0
      Fare       0
      Cabin     687
      Embarked   2
      dtype: int64

[23]: int(data['Age'].mean())

[23]: 29

[24]: df2['Age'] = df2['Age'].fillna(np.mean(df2['Age']))
      df2.isnull().sum()

[24]: PassengerId    0
      Survived    0
      Pclass     0
      Name       0
      Sex        0
      Age       0
      SibSp      0
      Parch      0
      Ticket     0
      Fare       0
      Cabin     687
      Embarked   2
      dtype: int64

[25]: df2.Embarked.value_counts()
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle Mode: Command Ln 1, Col 12 Intro.ipynb 2

Type here to search 32°C Haze 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... Data-Analysis/Data-Preprocess... JupyterLite how to take screenshot in lapto... | jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File Explorer: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
C23 C25 C27 4
C22 C26 3
F33 3
..
E34 1
C7 1
C54 1
E36 1
C148 1
Name: count, Length: 147, dtype: int64
```

[32]: df2['Cabin'].mode()

```
0      B96 B98
1  C23 C25 C27
2      G6
Name: Cabin, dtype: object
```

[36]: df2['Cabin'].fillna(df2['Cabin'].mode()[0], inplace=True)

[37]: df2.isnull().sum()

```
PassengerId  0
Survived     0
Pclass       0
Name         0
Sex          0
Age         0
SibSp       0
Parch       0
Ticket      0
Fare        0
Cabin        0
Embarked    0
dtype: int64
```

[]:

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle Mode: Command Ln 1, Col 12 Intro.ipynb 2 32°C Haze 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File browser: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
Emarked      object
dtype: object

[6]: data.columns

[6]: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
          'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
          dtype='object')

[7]: data['Survived'].value_counts()

[7]: Survived
0     549
1     342
Name: count, dtype: int64

[8]: data['Survived']==0

[8]: 0      True
1      False
2      False
3      False
4       True
...
886     True
887     False
888     True
889     False
890     True
Name: Survived, Length: 891, dtype: bool

[9]: print('Total number of passangers in the training data...', len(data))
print('Number of passangers who survived...', len(data[data['Survived'] == 1]))
print('Number of passangers who didn't survived...', len(data[data['Survived'] == 0]))

Total number of passangers in the training data... 891
Number of passangers who survived... 342
Number of passangers who didn't survived... 549
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle

Mode: Command Ln 1, Col 12 Intro.ipynb 2

32°C Haze 09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File Explorer: / notebooks /

- Untitled Folder
- Untitled Folder1
- Untitled Folder2
- Untitled Folder3
- Untitled Folder4
- Untitled Folder5
- Intro (copy).ipynb
- Intro.ipynb
- Lorenz.ipynb
- sqlite.ipynb
- Titanic-Dataset.csv
- titatic
- Untitled3.ipynb
- Untitled4.ipynb

Intro.ipynb

```
[25]: df2.Embarked.value_counts()

[25]: Embarked
S    644
C    168
Q     77
Name: count, dtype: int64

[26]: embark = df2['Embarked'].dropna()

[27]: df2[df2['Embarked'].isnull()]

[27]:
```

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	
61	62	1	1	Icard, Miss. Amelie	female	38.0	0	0	113572	80.0	B28	NaN
829	830	1	1	Stone, Mrs. George Nelson (Martha Evelyn)	female	62.0	0	0	113572	80.0	B28	NaN

```
[28]: df2['Embarked'].mode()

[28]: 0    S
Name: Embarked, dtype: object

[35]: df2['Embarked'].fillna(df2['Embarked'].mode()[0], inplace=True)

[30]: df2.isnull().sum()

[30]: PassengerId    0
Survived        0
Pclass          0
Name            0
Sex             0
Age             0
SibSp           0
Parch           0
Ticket          0
Fare            0
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle Mode: Command Ln 1, Col 12 Intro.ipynb 2

Type here to search

09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File browser: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

```
Parch      0
Ticket     0
Fare       0
Cabin      687
Embarked    0
dtype: int64
```

[31]: df2['Cabin'].value_counts()

```
Cabin
B96 B98      4
66           4
C23 C25 C27  4
C22 C26      3
F33          3
..          .
E34          1
C7           1
C54          1
E36          1
C148         1
Name: count, Length: 147, dtype: int64
```

[32]: df2['Cabin'].mode()

```
0      B96 B98
1   C23 C25 C27
2         66
Name: Cabin, dtype: object
```

[36]: df2['Cabin'].fillna(df2['Cabin'].mode()[0], inplace=True)

[37]: df2.isnull().sum()

```
PassengerId  0
Survived     0
Pclass       0
Name         0
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle Mode: Command Ln 1, Col 12 Intro.ipynb 2

Type here to search

09:40 PM 26-05-2025

Browser tabs: (322) Data Preprocessing with P... | Data-Analysis/Data-Preprocess... | JupyterLite | how to take screenshot in lapto... | +

Address bar: jupyter.org/try-jupyter/lab/

File Edit View Run Kernel Tabs Settings Help

File browser: / notebooks /

Name	Modified
Untitled Folder	1h ago
Untitled Folder1	1h ago
Untitled Folder2	1h ago
Untitled Folder3	1h ago
Untitled Folder4	1h ago
Untitled Folder5	1h ago
Intro (copy).ipynb	now
Intro.ipynb	2m ago
Lorenz.ipynb	3mo ago
sqlite.ipynb	3mo ago
Titanic-Dataset.csv	16m ago
titatic	1h ago
Untitled3.ipynb	1h ago
Untitled4.ipynb	1h ago

Intro.ipynb

memory usage: 66.2+ KB

```
[18]: data['Age'].value_counts()
```

```
[18]: Age
24.00    30
22.00    27
18.00    26
19.00    25
28.00    25
..
36.50     1
55.50     1
0.92      1
23.50     1
74.00     1
Name: count, Length: 88, dtype: int64
```

```
[19]: data['Cabin']
```

```
[19]: 0      NaN
1      C85
2      NaN
3     C123
4      NaN
...
886    NaN
887    B42
888    NaN
889    C148
890    NaN
Name: Cabin, Length: 891, dtype: object
```

```
[20]: data['Sex']
```

```
[20]: 0      male
1     female
2     female
```

Activate Windows
Go to Settings to activate Windows.

Simple 0 1 Python (Pyodide) | Idle

Mode: Command Ln 1, Col 12 Intro.ipynb 2

32°C Haze 09:40 PM 26-05-2025