

Practical 5

About this unit

Practical 5

Practice Lab Assignment

Unit • 100% completed



Lab Assignment

Unit • 100% completed



CODETANTRA

5.1.1. Stacked Plot

Create a stacked area plot to visualize the temperature variations for three different cities (City A, City B, and City C) across the months of the year. The temperature data is provided for each city in the editor.

Your task is to:

- Create a stacked area plot using the data.
- Label the x-axis as "Month", the y-axis as "Temperature", and provide the title "Temperature Variation" for the plot.
- Display the plot showing the temperature variation for each city throughout the months of the year.

```
stackedplot.py
import matplotlib.pyplot as plt
import pandas as pd

# Data for Months and Temperature for three cities
data = {
    'Month': ['January', 'February', 'March', 'April', 'May', 'June', 'July', 'August', 'September', 'October', 'November', 'December'],
    'City_A_Temperature': [5, 7, 10, 13, 17, 20, 22, 21, 18, 12, 8, 6],
    'City_B_Temperature': [2, 3, 5, 6, 10, 14, 16, 17, 12, 9, 5, 3],
    'City_C_Temperature': [3, 4, 6, 8, 9, 12, 15, 14, 10, 7, 4, 2]
}

# Write your code...
df = pd.DataFrame(data)
plt.stackplot(df['Month'], df['City_A_Temperature'], df['City_B_Temperature'], df['City_C_Temperature'])
plt.title('Temperature Variation')
plt.xlabel('Month')
plt.ylabel('Temperature')
plt.show()
```

Sample Test Cases +

< Prev Reset Submit Next >

CODETANTRA

5.2.1. Titanic Dataset

Write a Python program to analyze and visualize data from the Titanic dataset based on the following instructions:

Dataset Information:

The dataset is stored in a CSV file named titanic.csv and has been loaded using the pandas library. It contains the following columns:

- Pclass: Passenger class (1 = First, 2 = Second, 3 = Third).
- Gender: Gender of the passenger (male/female).
- Age: Age of the passenger.
- Survived: Survival status (0 = Did not survive, 1 = Survived).
- Fare: Ticket fare paid by the passenger.

Visualization:

To represent these trends, you will create 5 visualizations using Matplotlib. The visualizations should be arranged in a 3x2 grid (3 rows and 2 columns).

Visualization Details:

Write the code to create a series of visualizations as follows:

Bar Plot (Pclass Distribution):

- Create a bar plot to show the distribution of passengers across the different passenger classes (Pclass).
- Use the color skyblue for the bars.
- Title the plot as "Passenger Class Distribution".
- Label the x-axis as "Pclass" and the y-axis as "Count".

Pie Chart (Gender Distribution):

- Create a pie chart to display the distribution of male and female passengers.
- Use lightblue for males and lightcoral for females.
- Include percentages on the slices (use autopct='%.1f%').
- Title the plot as "Gender Distribution".

Histogram (Age Distribution):

- Create a histogram to visualize the distribution of passengers' ages.
- Use lightgreen for the bars with black edges (edgecolor = 'black').
- Set the number of bins to 8 for the histogram.
- Title the plot as "Age Distribution".
- Label the x-axis as "Age" and the y-axis as "Frequency".

Bar Plot (Survival Count):

- Create a bar plot to show the count of passengers who survived and those who did not, based on the Survived column.
- Use the colors lightblue for survivors (1) and lightcoral for non-survivors (0).
- Title the plot as "Survival Count".
- Label the x-axis as "Survived (0 = No, 1 = Yes)" and the y-axis as "Count".

Scatter Plot (Fare vs Age):

- Create a scatter plot to visualize the relationship between the Fare and Age of passengers.
- Use orange for the data points.
- Title the plot as "Fare vs Age".
- Label the x-axis as "Age" and the y-axis as "Fare".

Note: Refer to the displayed plot in the sample test cases for better understanding.

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the Titanic dataset from the
5 # CSV file
6 df = pd.read_csv('titanic.csv')
7
8 # Set up the figure for 5 subplots
9 fig, axes = plt.subplots(3, 2,
10 figsize=(12, 12))
11
12 # write the code..
13 # Plot 1: Count of passengers by
14 # class
15 axes[0, 0].bar(df['Pclass'].value_counts().index,
16 df['Pclass'].value_counts(), color='skyblue')
17 axes[0, 0].set_title("Passenger
18 Class Distribution")
19 axes[0, 0].set_xlabel("Pclass")
20 axes[0, 0].set_ylabel("Count")
21
22 # plot 2: Gender distribution
23 axes[0, 1].pie(df['Gender'].value_counts(),
24 labels=df['Gender'].value_counts().index,
25 autopct='%.1f%%', colors=[
26 'lightblue', 'lightcoral'])
27 axes[0, 1].set_title("Gender
28 Distribution")
29
30 # plot 3: Age distribution
31 axes[1, 0].hist(df['Age'].dropna(),
32 bins=8, color='lightgreen',
33 edgecolor='black')
34 axes[1, 0].set_title("Age
35 Distribution")
36 axes[1, 0].set_xlabel("Age")
37 axes[1, 0].set_ylabel("Frequency")
38
39 # plot 4: Survival count
40 axes[1, 1].bar(df['Survived'].value_counts().index,
41 df['Survived'].value_counts(), color=[
42 'lightblue', 'lightcoral'])
43 axes[1, 1].set_title("Survival
44 Count")
45 axes[1, 1].set_xlabel("Survived (0 =
46 No, 1 = Yes)")
47 axes[1, 1].set_ylabel("Count")
48
49 # plot 5: Fare vs Age
50 axes[2, 0].scatter(df['Age'],
51 df['Fare'], color='orange',
52 edgecolors='black')
53 axes[2, 0].set_title("Fare vs Age")
54 axes[2, 0].set_xlabel("Age")
55 axes[2, 0].set_ylabel("Fare")
56
57 plt.tight_layout()
58 plt.show()

```

Sample Test Cases

< Prev Reset Submit Next >

CODETANTRA

5.2.2. Histogram of passenger information o... 02:54 A ⚡ -

Write a Python code to plot a histogram for the distribution of the 'Age' column from the Titanic dataset. The histogram should display the frequency of different age ranges with the following specifications:

1. Use **30 bins** for the histogram.
2. Set the **edge color** of the bars to **black (k)**.
3. Label the x-axis as '**Age**' and the y-axis as '**Frequency**'.
4. Add the title "**Age Distribution**" to the histogram.

The Titanic dataset contains columns as shown below,

P	a	s	S	u	P	N	S	A	S	i	P	T	F	C	E
a	s	s	e	r	c	a	e	g	e	b	a	i	a	a	m
e	n	i	v	e	s	s	x	e	p	S	r	c	r	b	b
I	d														

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note: Refer to the visible test case for better reference.

Sample Test Cases +

Histogram...
Submit
Debugger

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the Titanic dataset
5 data = pd.read_csv('Titanic-
Dataset.csv')
6
7 # Data Cleaning
8 data['Age'].fillna(data['Age'].median
(), inplace=True)
9 data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
10 data.drop('Cabin', axis=1,
inplace=True)
11
12 # Convert categorical features to
13 numeric
13 data['Sex'] =
14 data['Sex'].map({'male': 0,
15 'female': 1})
14 data = pd.get_dummies(data, columns=
15 ['Embarked'], drop_first=True)
16
16 # Write your code here for Histogram
17 plt.hist(data['Age'], bins=30,
edgecolor='k')
18 plt.xlabel('Age')
19 plt.ylabel('Frequency')
20 plt.title('Age Distribution')
21 plt.show()

```

< Prev
Reset
Submit
Next >

CODETANTRA

5.2.3. Bar plot of survival rate of passengers

Write a Python code to plot a bar chart that shows the count of passengers who survived and did not survive in the Titanic dataset. The chart should display the following specifications:

1. Use the **'Survived'** column to show the count of survivors (0 = Did not survive, 1 = Survived).
2. Set the chart type to **'bar'**.
3. Add the title **"Survival Count"** to the chart.
4. Label the x-axis as **'Survived'** and the y-axis as **'Count'**.

The Titanic dataset contains columns as shown below,

P	a	s	S	u	P	N	S	A	S	i	P	T	F	C	E
s	s	r	v	i	c	a	e	g	b	a	a	i	a	a	m
e	e	e	e	s	a	m	x	e	S	c	c	r	r	b	b
P	a	s	S	u	P	N	S	A	S	i	P	T	F	C	E
a	s	r	v	i	c	a	m	e	b	a	a	i	a	a	m
s	e	e	e	s	a	m	x	e	S	c	c	r	r	b	b
e	e	e	e	s	a	m	x	e	S	c	c	r	r	b	b
r	e	e	e	s	a	m	x	e	S	c	c	r	r	b	b
I	d														

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note: Refer to the visible test case for better reference.

Sample Test Cases

BarPlotOf...

Submit Debugger

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the Titanic dataset
5 data = pd.read_csv('Titanic-
Dataset.csv')
6
7 # Data Cleaning
8 data['Age'].fillna(data['Age'].median
(), inplace=True)
9 data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
10 data.drop('Cabin', axis=1,
inplace=True)
11
12 # Convert categorical features to
13 numeric
14 data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
15 data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)
16
17 # Write your code here for Bar Plot
18 for Survival_Rate
19
20 survival_counts =
21 data['Survived'].value_counts()
22 survival_counts.plot(kind='bar')
23 plt.title('Survival Count')
24 plt.xlabel('Survived')
25 plt.ylabel('Count')
26 plt.show()

```

< Prev Reset Submit Next >

CODETANTRA

5.2.4. Bar Plot for Survival by Gender

07:49 A ⚡

Write a Python code to plot a stacked bar chart that shows the count of passengers who survived and did not survive, grouped by gender, in the Titanic dataset. The chart should display the following specifications:

1. Group the data by the 'Sex' column, then use the `value_counts()` function to count the occurrences of survivors (0 = Did not survive, 1 = Survived) for each gender.
2. Use a **stacked bar chart** to display the survival counts.
3. Add the title "Survival by Gender" to the chart.
4. Label the x-axis as 'Gender' and the y-axis as 'Count'.
5. The legend should indicate 'Not Survived' and 'Survived'.

The Titanic dataset contains columns as shown below,

P	a	s	S	u	P	N	S	A	S	P	T	F	C	E
a	s	s	r	i	c	a	e	g	i	a	i	a	b	m
s	s	v	e	a	l	m	x	e	b	r	c	r	a	b
P	a	s	S	u	P	N	S	A	S	P	T	F	C	E
a	s	s	r	i	c	a	e	g	i	a	i	a	b	m
s	s	v	e	a	l	m	x	e	b	r	c	r	a	b
e	e	r	e	s	s	e	s	p	s	c	h	e	t	d
r	e	r	e	s	s	e	s	p	s	c	h	e	t	d
I	d													

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note: Refer to the visible test case for better reference.

Sample Test Cases

BarPlotOf... Explorer Submit Debugger

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the Titanic dataset
5 data = pd.read_csv('Titanic-
Dataset.csv')
6
7 # Data Cleaning
8 data['Age'].fillna(data['Age'].median
(), inplace=True)
9 data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
10 data.drop('Cabin', axis=1,
inplace=True)
11
12 # Convert categorical features to
13 numeric
14 data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
15 data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)
16
17 # Write your code here for Bar Plot
for Survival by Gender
18
19 survival_by_gender =
data.groupby('Sex')
20 ['Survived'].value_counts().unstack()
.fillna(0)
21 survival_by_gender.columns = ['Not-
Survived', 'Survived']
22 survival_by_gender.index = ['0', '1']
23 survival_by_gender.plot(kind='bar',
stacked=True)
24 plt.title('Survival by Gender')
25 plt.xlabel('Gender')
26 plt.ylabel('Count')
27 plt.legend(title=None)
plt.show()
```

< Prev Reset Submit Next >

CODETANTRA

5.2.5. Bar Plot for Survival by Pclass

05:28 A ⚡

Write a Python code to plot a stacked bar chart that shows the count of passengers who survived and did not survive, grouped by passenger class (**Pclass**), in the Titanic dataset. The chart should display the following specifications:

1. Group the data by the **Pclass** column and count the number of survivors (0 = Did not survive, 1 = Survived) for each class using `value_counts()`.
2. Use a **stacked bar chart** to display the survival counts.
3. Add the title "Survival by Pclass" to the chart.
4. Label the x-axis as '**Pclass**' and the y-axis as '**Count**'.
5. The legend should indicate 'Not Survived' and 'Survived'.

The Titanic dataset contains columns as shown below,

P	S	P	N	S	A	S	P	T	F	C	E
a	u	c	a	e	g	b	a	i	a	a	m
s	r	l	m	x	e	s	c	c	r	b	b
e	v	i	a	m	e	x	p	h	e	a	k
n	i	a	m	e	g	e	s	h	e	b	e
g	e	e	s	s	e	e	p	t	e	a	d
e	r	e	s	s	e	e	s	e	e	r	d
I	l	d									

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8,0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8,4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51,86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note:

- Refer to the visible test case for better reference.
- Ensure you use the `groupby()` function with `value_counts()` to count the survivors and non-survivors for each **Pclass**.
- Do not manually use `size()` or `unstack()` without `value_counts()`. Use the `value_counts()` method for counting survival status directly.

Sample Test Cases



BarPlotOf... Submit Debugger

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the Titanic dataset
data = pd.read_csv('Titanic-
Dataset.csv')

# Data Cleaning
data['Age'].fillna(data['Age'].median
(), inplace=True)
data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
data.drop('Cabin', axis=1,
inplace=True)

# Convert categorical features to
numeric
data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)

# Write your code here for Bar Plot
for Survival by Pclass

survival_by_class =
data.groupby('Pclass')
['Survived'].value_counts().unstack()
.fillna(0)
survival_by_class.columns = ['Not
Survived', 'Survived']
survival_by_class.plot(kind='bar',
stacked=True)
plt.title('Survival by Pclass')
plt.xlabel('Pclass')
plt.ylabel('Count')
plt.legend(title=None)
plt.show()
```

< Prev Reset Submit Next >

☰ CODETANTRA

5.2.6. Bar Plot for Survival by Embarked

13:41 AA ☾ ✎ ⌂ -

Write a Python code to plot a stacked bar chart showing the survival count for passengers based on their embarkation location in the Titanic dataset.

The chart should display the following specifications:

1. Use the **Embarked** column to determine the embarkation location. After converting this column into dummy variables (using `pd.get_dummies()`), plot the survival count based on the **Embarked_Q** column (representing passengers who embarked from Queenstown) in relation to survival.
 2. Set the chart type to 'bar' and make it stacked.
 3. Add the title "**Survival by Embarked**" to the chart.
 4. Label the x-axis as '**Embarked**' and the y-axis as '**Count**'.
 5. Include a legend to distinguish between survivors and non-survivors (label the legend as '**Survived**' and '**Not Survived**').

The Titanic dataset contains columns as shown below

Sample Data:

PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Titanic
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thayer)",female,38,1,1,PC/2 3473
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2 3134
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",female,35,1,1,PC/2 3503
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8,4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.06
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)",female,27,1,0,342370
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,1,1,343355

Note: Refer to the visible test case for better reference

BarPlotOf...

Submit

Debugger

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the Titanic dataset
5 data = pd.read_csv('Titanic-
Dataset.csv')
6
7 # Data Cleaning
8 data['Age'].fillna(data['Age'].median()
(), inplace=True)
9 data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
10 data.drop('Cabin', axis=1,
inplace=True)
11
12 # Convert categorical features to
numeric
13 data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
14 data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)
15
16 # Write your code here for Bar Plot
for Survival by Embarked
17
18
19 grouped = data.groupby('Embarked_Q')
['Survived'].value_counts().unstack()
.fillna(0)
20 grouped.columns = ['Not Survived',
'Survived']
21 grouped.plot(kind='bar',
stacked=True)
22 plt.title('Survival by Embarked')
23 plt.xlabel('Embarked')
24 plt.ylabel('Count')
25 plt.legend(title=None)
26 plt.show()
```

Sample Test Cases

+

>-

< Prev Reset Submit Next >

☰ CODETANTRA

5.2.7. Box plot for Age Distribution

04:50 AA ☾ ✎ ⌂ -

Write a Python code to plot a boxplot that shows the distribution of the 'Age' column from the Titanic dataset across different passenger classes. The boxplot should display the following specifications:

1. Use the **Pclass** column to group the data for the boxplot.
 2. Set the title of the plot to "**Age by Pclass**".
 3. Remove the default subtitle with `plt.suptitle("")`.
 4. Label the x-axis as '**Pclass**' and the y-axis as '**Age**'.

The Titanic dataset contains columns as shown below,

Sample Data:

Note: Refer to the visible test case for better reference.

BoxPlotFor...

Submit

Debugger

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the Titanic dataset
data = pd.read_csv('Titanic-Dataset.csv')

# Data Cleaning
data['Age'].fillna(data['Age'].median(), inplace=True)
data['Embarked'].fillna(data['Embarked'].mode()[0], inplace=True)
data.drop('Cabin', axis=1, inplace=True)

# Convert categorical features to numeric
data['Sex'] =
data['Sex'].map({'male': 0, 'female': 1})
data = pd.get_dummies(data, columns=['Embarked'], drop_first=True)

# Write your code here for Box Plot
for Age by Pclass

plt.figure(figsize=(8, 6))
data.boxplot(column='Age', by='Pclass')
plt.suptitle('')
plt.title('Age by Pclass')
plt.xlabel('Pclass')
plt.ylabel('Age')
plt.show()
```

Sample Test Cases

七

1

< Prev Reset Submit Next >

CODETANTRA

5.2.8. Box Plot for Age by Survived

03:58 AA ☾ ☽ -

Write a Python code to plot a boxplot that shows the distribution of the 'Age' column from the Titanic dataset based on whether passengers survived or not. The boxplot should display the following specifications:

1. Use the **Survived** column to group the data for the boxplot (0 = Did not survive, 1 = Survived).
2. Set the title of the plot to "**Age by Survival**".
3. Remove the default subtitle with `plt.suptitle("")`.
4. Label the x-axis as '**Survived**' and the y-axis as '**Age**'.

The Titanic dataset contains columns as shown below,

P	a	s	S	p	c	N	s	A	S	i	p	T	F	C	E
a	s	s	u	r	v	i	e	g	e	b	S	a	a	a	m
s	s	e	l	a	s	a	x	g	e	b	c	c	r	b	b
r	e	e	s	s	s	e									d

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note: Refer to the visible test case for better reference.

Sample Test Cases

Explorer
BoxPlotFor...
Submit
Debugger

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the Titanic dataset
5 data = pd.read_csv('Titanic-
Dataset.csv')
6
7 # Data Cleaning
8 data['Age'].fillna(data['Age'].median
(), inplace=True)
9 data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
10 data.drop('Cabin', axis=1,
inplace=True)
11
12 # Convert categorical features to
13 numeric
14 data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
15 data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)
16
17
18
19 plt.figure(figsize=(8, 6))
20 data.boxplot(column='Age',
by='Survived')
21 plt.suptitle('')
22 plt.title('Age by Survival')
23 plt.xlabel('Survived')
24 plt.ylabel('Age')
25 plt.show()

```

< Prev Reset Submit Next >

CODETANTRA

5.2.9. Box Plot for Fare by Pclass

06:01 A ⚡

Write a Python code to plot a boxplot that shows the distribution of the 'Fare' column from the Titanic dataset based on the passenger class (Pclass). The boxplot should display the following specifications:

1. Use the **Pclass** column to group the data for the boxplot.
2. Set the title of the plot to **"Fare by Pclass"**.
3. Remove the default subtitle with **plt.suptitle("")**.
4. Label the x-axis as **'Pclass'** and the y-axis as **'Fare'**.

The Titanic dataset contains columns as shown below,

P	a	s	S	u	P	N	S	A	S	i	P	T	F	C	E
a	s	s	e	r	c	a	e	g	e	b	a	i	a	a	m
e	n	i	v	e	s	s	x	e	p	S	r	c	r	b	b
I	d														

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note: Refer to the visible test case for better reference.

Sample Test Cases



BoxPlotFor...

Submit



Debugger

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the Titanic dataset
data = pd.read_csv('Titanic-
Dataset.csv')

# Data Cleaning
data['Age'].fillna(data['Age'].median
(), inplace=True)
data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
data.drop('Cabin', axis=1,
inplace=True)

# Convert categorical features to
numeric
data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)

# Write your code here for Box Plot
for Fare by Pclass

plt.figure(figsize=(8, 6))
data.boxplot(column='Fare',
by='Pclass')
plt.suptitle('')
plt.title('Fare by Pclass')
plt.xlabel('Pclass')
plt.ylabel('Fare')
plt.show()
```

< Prev Reset Submit Next >

☰ CODETANTRA

5.2.10. Scatter Plot for Age vs. Fare

Write a Python code to plot a scatter plot showing the relationship between the 'Age' and 'Fare' columns in the Titanic dataset. The scatter plot should display the following specifications:

1. Use the **Age** column for the x-axis and the **Fare** column for the y-axis.
 2. Set the title of the plot to "**Age vs. Fare**".
 3. Label the x-axis as '**Age**' and the y-axis as '**Fare**'.

The Titanic dataset contains columns as shown below,

Sample Data:

PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ticket
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,71
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thayre)",female,38,1,1,C/5 21315,73
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2 313389,73
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",female,31,1,1,PC 17593,74
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8,0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8,4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51,86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34490,9,1
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Helmerina Bergfeldt)",female,27,3,1,34330,9,1
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,1,1,34338,9,1

Note: Refer to the visible test case for better reference.

AgeFareSc...

Submit

Debugger

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the Titanic dataset
data = pd.read_csv('Titanic-
Dataset.csv')

# Data Cleaning
data['Age'].fillna(data['Age'].median(
), inplace=True)
data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
data.drop('Cabin', axis=1,
inplace=True)

# Convert categorical features to
numeric
data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)

# Write your code here for Box Plot
for Fare by Pclass

plt.figure(figsize=(6.4,4.8))
plt.scatter(data['Age'],data['Fare'])
plt.title('Age vs. Fare')
plt.xlabel('Age')
plt.ylabel('Fare')
plt.show()
```

Sample Test Cases

†

1

< Prev Reset Submit Next >

CODETANTRA

5.2.11. Scatter Plot for Age vs. Fare by Survived

Write a Python code to plot a scatter plot showing the relationship between the 'Age' and 'Fare' columns in the Titanic dataset, with points color-coded by survival status. The scatter plot should display the following specifications:

1. Use the **Age** column for the x-axis and the **Fare** column for the y-axis.
2. Color the points based on the **Survived** column: Red for passengers who did not survive (**Survived = 0**). Blue for passengers who survived (**Survived = 1**).
3. Set the title of the plot to "Age vs. Fare by Survival".
4. Label the x-axis as 'Age' and the y-axis as 'Fare'.

The Titanic dataset contains columns as shown below,

P	S	P	N	S	A	S	P	T	F	C	E
a	s	c	a	e	g	b	a	i	a	a	m
s	u	l	m	x	e	S	r	c	r	b	b
e	v	i	a	m	e	p	c	h	e	a	r
n	i	a	m	x	e	S	r	c	e	b	k
g	e	s	e			p	t	i	r	a	e
e	e	s						c	a	r	d
r											
I											
d											

Sample Data:

```
PassengerId,Survived,Pclass,Name,Sex,Age,SibSp,Parch,Ti
1,0,3,"Braund, Mr. Owen Harris",male,22,1,0,A/5 21171,7
2,1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thay
3,1,3,"Heikkinen, Miss. Laina",female,26,0,0,STON/O2. 3
4,1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",fe
5,0,3,"Allen, Mr. William Henry",male,35,0,0,373450,8.0
6,0,3,"Moran, Mr. James",male,,0,0,330877,8.4583,,Q
7,0,1,"McCarthy, Mr. Timothy J",male,54,0,0,17463,51.86
8,0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,34990
9,1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg
10,1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,
```

Note: Refer to the visible test case for better reference.

Sample Test Cases

Submit
Reset

```

import pandas as pd
import matplotlib.pyplot as plt

# Load the Titanic dataset
data = pd.read_csv('Titanic-
Dataset.csv')

# Data Cleaning
data['Age'].fillna(data['Age'].median
(), inplace=True)
data['Embarked'].fillna(data['Embarke
d'].mode()[0], inplace=True)
data.drop('Cabin', axis=1,
inplace=True)

# Convert categorical features to
numeric
data['Sex'] =
data['Sex'].map({'male': 0,
'female': 1})
data = pd.get_dummies(data, columns=
['Embarked'], drop_first=True)

# Write your code here for Scatter
Plot for Age vs. Fare by Survived
colors = data['Survived'].map({0:
'red', 1: 'blue'})
plt.scatter(data['Age'],
data['Fare'], c=colors)

# Set labels and title
plt.xlabel("Age")
plt.ylabel("Fare")
plt.title("Age vs. Fare by Survival")

# Show the plot
plt.show()

```