5.1.1. Stacked Plot



stackedpl...

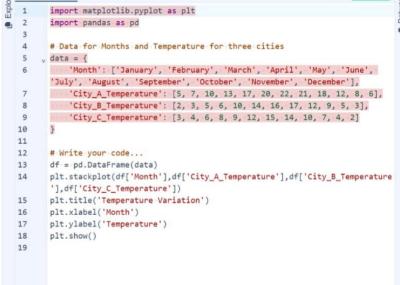
> Terminal

⊞ Test cases

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 vear.



### 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

# Visualization Details:

Sample Test Cases

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

be arranged in a 3x2 grid (3 rows and 2 columns).

import pandas as pd

atitanicDat...

10

11

12

13

14

15

16

19

21

import matplotlib.pyplot as plt

# Load the Titanic dataset from the CSV file

axes[0, 0].set xlabel("Pclass")

df = pd.read csv('titanic.csv')

# Set up the figure for 5 subplots

fig, axes = plt.subplots(3, 2, figsize=(12, 12))

axes[0, 0].bar(df['Pclass'].value counts().index,

df['Pclass'].value counts(), color='skyblue') axes[0, 0].set title("Passenger Class Distribution")

axes[0, 0].set ylabel("Count") axes[0, 1].pie(df['Gender'].value counts(),

labels=df['Gender'].value\_counts().index, autopct='%1.1f%%', colors= ['lightblue', 'lightcoral'])

17 axes[0, 1].set\_title("Gender Distribution") 18

axes[1, 0].hist(df['Age'].dropna(), bins=8, color='lightgreen', edgecolor='black')

axes[1, 0].set\_title("Age Distribution") axes[1, 0].set\_xlabel("Age")

Terminal Test cases

#### 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

# Visualization Details:

Sample Test Cases

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

be arranged in a 3x2 grid (3 rows and 2 columns).

19 20

17

18

21

22

23

24

25

26

27

28

29

30

Expl

16

atitanicDat...

axes[0, 1].pie(df['Gender'].value counts(),

labels=df['Gender'].value counts().index, autopct='%1.1f%%', colors= ['lightblue', 'lightcoral'])

axes[0, 1].set title("Gender Distribution")

axes[1, 0].hist(df['Age'].dropna(), bins=8, color='lightgreen',

edgecolor='black') axes[1, 0].set\_title("Age Distribution")

axes[1, 0].set\_xlabel("Age") axes[1, 0].set vlabel("Frequency")

axes[1, 1].bar(df['Survived'].value\_counts().index,

df['Survived'].value\_counts(), color=['lightblue', 'lightcoral']) axes[1, 1].set\_title("Survival Count") axes[1, 1].set xlabel("Survived (0 = No, 1 = Yes)") axes[1, 1].set ylabel("Count")

axes[2, 0].scatter(df['Age'], df['Fare'], color='orange', edgecolors='black') axes[2, 0].set\_title("Fare vs Age")

31 axes[2, 0].set\_xlabel("Age") 32 axes[2, 0].set vlabel("Fare") 33

34 plt.tight\_layout() 35 plt.show()

> Terminal Test cases

Prev Reset Submit

### 5.2.2. Histogram of passenger information of Titanic



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

Pas sen gerl d	Sur vive d	Pda ss	Na me	Sex	Age	Sib Sp	Par ch	Tick et	Far e	Cab In	Em bark ed	
-												ĺ

#### Sample Data:

PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked 1,0,3,"Braund, Mr. Owen Harris", male, 22,1,0,A/5 21171,7.25,,S 2,1,1, "Cumings, Mrs. John Bradley (Florence Briggs Thayer)", female, 38,1,0,PC 17599,71.2833,C85,C

Sample Test Cases

Histogra...

import pandas as pd

import matplotlib pyplot as plt

# Load the Titanic dataset

# Convert categorical features to numeric

plt.hist(data['Age'], bins=30, edgecolor='k')

# Write your code here for Histogram

data['Sex'] = data['Sex'].map({'male': 0. 'female': 1})

data = pd.get dummies(data, columns=['Embarked'], drop first=True)

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) 10 data.drop('Cabin', axis=1, inplace=True) 11

12

13 14

15

16 17

18

19

20

21

22

> Terminal

plt.xlabel('Age')

plt.show()

plt.vlabel('Frequency')

plt.title('Age Distribution')

Test cases

< Prev

Reset

Submit

## CDETANTRA # Home

# 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.

Pas Sur Cab Pcla Na Sib Par Tick Far sen vive Age Sex Sp gerl me ch in

# Sample Data:

PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked 1,0,3, "Braund, Mr. Owen Harris", male, 22,1,0,A/5 21171,7.25,,S 2.1.1, "Cumings, Mrs. John Bradley (Florence Briggs Thaver)", female, 38, 1, 0, PC 17599, 71, 2833, C85, C

Sample Test Cases

18 19 20 21

> 15 16

17

22

23

> Terminal

bark ed

10

Em

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

import pandas as pd import matplotlib.pyplot as plt

# Data Cleaning

BarPlotOf...

Explor

data.drop('Cabin', axis=1, inplace=True)

# Convert categorical features to numeric

survival counts.plot(kind='bar')

plt.title('Survival Count')

plt.xlabel('Survived')

⊞ Test cases

plt.ylabel('Count')

plt.show()

data['Age'].fillna(data['Age'].median(), inplace=True)

data['Sex'] = data['Sex'].map({'male': 0, 'female': 1})

# Write your code here for Bar Plot for Survival Rate survival counts = data['Survived'].value counts()

data['Embarked'].fillna(data['Embarked'].mode()[0], inplace=True)

data = pd.get\_dummies(data, columns=['Embarked'], drop\_first=True)

# C∯DETANTRA # Home 5.2.4. Bar Plot for Survival by Gender

1214 A C B 8 -

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'.

The legend should indicate 'Not Survived' and 'Survived'.

The Titanic dataset contains columns as shown below.

Pas Sur Em Na Par Tick Far Cab sen Pcla Sib vive Sex Age bark gerl ch in me ed

# Sample Data:

Sample Test Cases

Explor

10

11

17

19

20

21

22

23

24

25

BarPlotOf...

# 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 12 13 data['Sex'] = data['Sex'].map({'male': 0, 'female': 1}) 14

data = pd.get dummies(data, columns=['Embarked'], drop first=True) 15 16

survival by gender = data.groupby('Sex')

18

['Survived'].value counts().unstack().fillna(0) survival by gender.columns = ['Not Survived', 'Survived']

survival by gender.index = ['0', '1']

survival by gender.plot(kind='bar', stacked=True) plt.title('Survival by Gender')

plt.xlabel('Gender')

import matplotlib.pyplot as plt

data = pd.read csv('Titanic-Dataset.csv')

# Load the Titanic dataset

plt.ylabel('Count')

plt.legend(title=None) plt.show()

> Terminal 

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

202401070175@mitaoe.ac.in Support Logout (\*)



### 5.2.5. Bar Plot for Survival by Pclass



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.

	Pas sen gerl	Sur vive d	Pcla ss	Na me	Sex	Age	Sib Sp	Par ch	Tick et	Far e	Cab in	Em bark ed
--	--------------------	------------------	------------	----------	-----	-----	-----------	-----------	------------	----------	-----------	------------------

## Sample Data:

Sample Test Cases

BarPlotOf...

import pandas as pd

import matplotlib.pvplot as plt

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

# Data Cleaning

10

11

16

17

18 19

20

12 13

14 15

21 plt.xlabel('Pclass') 22

plt.ylabel('Count') plt.legend(title=None) 23

plt.show()

Terminal Test cases

plt.title('Survival by Pclass')

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 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)

# CDETANTRA # Home

#### 5.2.6. Bar Plot for Survival by Embarked





Write a Python code to plot a stacked bar chart showing the survival count for passengers based on

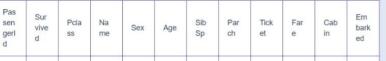
their embarkation location in the Titanic dataset.

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 Test Cases



ger

BarPlotOf...

9

13

14

15

22

23

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)

10 data.drop('Cabin', axis=1, inplace=True) 11 12

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

data = pd.get dummies(data, columns=['Embarked'], drop first=True)

16 # Write your code here for Bar Plot for Survival by Embarked 17 grouped = data.groupby('Embarked Q')

18 grouped.columns = ['Not Survived', 'Survived'] 19 grouped.plot(kind='bar', stacked=True) 20 plt.title('Survival by Embarked')

['Survived'].value\_counts().unstack().fillna(0)

plt.xlabel('Embarked') plt.vlabel('Count') plt.legend(title=None)

< Prev

BoxPlotF...

24

25

> Terminal

Test cases

#### 5.2.7. Box plot for Age Distribution A L E 2 -

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.

```
Pas
        Sur
                                                                                             Em
sen
                 Pcla
                         Na
                                                   Sib
                                                           Par
                                                                    Tick
                                                                            Far
                                                                                     Cab
        vive
                                  Sex
                                          Age
                                                                                             bark
gerl
                         me
                                                           ch
                                                                                     in
                                                                                             ed
```

## Sample Data:

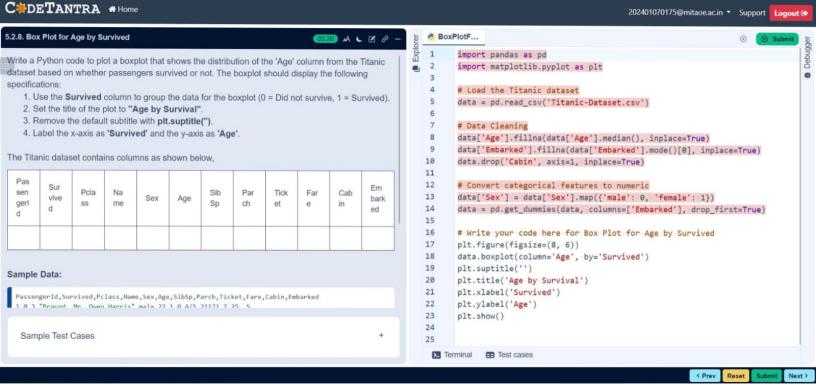
```
PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked
1,0,3, "Braund, Mr. Owen Harris", male, 22,1,0, A/5 21171,7,25,.5
```

2.1.1, "Cumings, Mrs. John Bradley (Florence Briggs Thayer)", female, 38.1,0,PC 17599,71.2833,C85,C

Sample Test Cases

1 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) 9 data['Embarked'].fillna(data['Embarked'].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 Box Plot for Age by Pclass 17 plt.figure(figsize=(8, 6)) 18 data.boxplot(column='Age', by='Pclass') plt.suptitle('-') plt.title('Age by Pclass') 20 21 plt.xlabel('Pclass') 22 plt.ylabel('Age') 23 plt.show()

Logout (\*)



# CDETANTRA # Home

# (02211) A L Z 2 -5.2.9. Box Plot for Fare by Pclass 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.

```
Pas
        Sur
                                                                                            Em
                Pcla
                                                  Sib
                                                           Par
                                                                           Far
                                                                                    Cab
                         Na
                                                                   Tick
sen
        vive
                                 Sex
                                         Age
                                                                                            bark
gerl
                                                  Sp
                                                           ch
                                                                                    in
                SS
                         me
                                                                           e
                                                                                            ed
```

# Sample Data:

```
PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked
 1,0,3,"Braund, Mr. Owen Harris", male, 22,1,0,A/5 21171,7.25,,S
2,1,1, "Cumings, Mrs. John Bradley (Florence Briggs Thayer)", female, 38,1,0,PC 17599,71.2833,C85,C
```

Sample Test Cases

```
BoxPlotF...
       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)
10
       data.drop('Cabin', axis=1, inplace=True)
11
       # Convert categorical features to numeric
12
      data['Sex'] = data['Sex'].map({'male': 0, 'female': 1})
13
14
      data = pd.get dummies(data, columns=['Embarked'], drop first=True)
15
16
      # Write your code here for Box Plot for Fare by Pclass
      plt.figure(figsize=(8, 6))
17
      data.boxplot(column='Fare', by='Pclass')
18
19
      plt.suptitle('')
20
      plt.title('Fare by Pclass')
      plt.xlabel('Pclass')
21
22
      plt.ylabel('Fare')
23
      plt.show()
```

















Terminal





Test cases























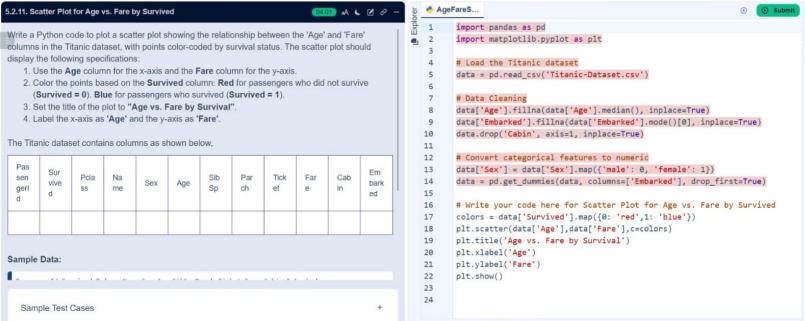








Doubles almostic



 ☐ Test cases

> Terminal

Logout (\*)