(i) Plot histogram of different colours of 7 days temperature of your city.

(ii) Set title as "7 Days Temperature".

(iii) Set label for X-axis "Weekdays"

(iv) Set label for Y-axis "Temperature"

(v) Save the image as "temp-pick.jpg"

7. (a) Write short note on histograms and density plots, with suitable examples. (4)

(b) Explain the functions : (8)

(i) value_counts()

(ii) stack()

(iii) barh()

(iv) arange()

(c) Write output of the following code : (3)

```
import pandas as pd
ages=[15, 47, 56, 23, 78, 90]
bins=[12, 30, 70, 95]
categories = pd.cut(ages, bins)
print(categories)
```


[This question paper contains 8 printed pages]

Your Roll No.............

Sr. No. of Question Paper : 1461 **F**

Unique Paper Code	: 2342201202
Name of the Paper	: Data Interpretation and Visualization using Python
Name of the Course	: **B.A.(Prog.)**
Semester	: II

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory
3. Attempt any **4 (four)** questions from **Section B**.
4. Parts of a question must be answered together.

Section A
(Compulsory)

1. (a) Write the output of the following code : (3)

```
import numpy as np
a=np.array([[[1,2,3], [3,4,2]], [[4,6,5], [3,1,6]]])
b=np.random.randn(4,3)
print(a.shape)
print(b.shape)
```

(b) Explain any two functions that can be used to create numpy array objects, with suitable examples. (4)

(c) Consider the given pandas series object and write the output of the following code statements : (5)

```
import pandas as pd
obj2=pd.Series([4, 6, 5, 8, 7],index=['d', 'b', 'c', 'a', 'e'])
```

(i) print (obj2['d'])

(ii) print (obj2 [obj2>5])

(iii) print (obj2.index)

(iv) print (obj2*3)

(v) print (obj2['b' : 'a'])

(d) Write the output of the following code : (2)

```
import numpy as np
arr = np.array([3.7, –1.2, –2.6, 0.5, 12.9, 10.1])
```



```
year;qtr1;qtr2;qtr3;qtr4
2019;3000;3200;;3500
2020;2900;-9999;;
2021;1800;2100;1900;1950
2021;1800;2100;1900;1950
2022;1850;-9999;2900;2550
2023;3400;3200;;
```

(i) Load the file sales.csv into a dataframe.

(ii) Fill all the NaN values with 0.

(iii) Remove the duplicate rows from the original dataframe.

(iv) Replace all the -9999 values with a 0.

(v) Print the average sales made during qtrl across all the years

6. (a) Write Python statements to plot a horizontal and vertical bar chart of any assumed dataset. (5)

(b) Write the Python statements to do the following using matplotlib package : (10)

(c) Explain fillna and dropna functions in pandas dataframes, with example code. (5)

5. (a) Consider the given dataframe df and write the Python statements to perform the following operations: (5)

	ID	marks
B	23	52
A	34	67
C	25	60

(i) Set the title of the row index as name

(ii) Add a column 'semester' with values 2,1,2

(iii) Sort the dataframe in descending order of the marks column.

(iv) Reindex the dataframe in the order A, B and C.

(v) Drop the row corresponding to row index C.

(b) Explain covariance and correlation with suitable examples. (5)

(c) Consider the fde sales.csv as given below and answer the following questions. (5)


```
arr2 = arr.astype(np.int32)
print(arr2)
```

(e) Explain reindex() function for a series object. (2)

(f) Explain *skiprows* and *na_values* attributes of the read_csv() function. (3)

(g) Explain *bool* and *string* data types of ndarray object. (4)

(h) Write a short note on matplotlib library in python. (3)

(i) Write output of the following code : (4)

```
import pandas as pd
df = pd.DataFrame({'Animal' : ['Falcon', 'Falcon',
                               'Parrot', 'Parrot'],
                   'Max Speed' : [380., 370., 24., 26.]})
print(df.groupby(['Animal']).mean())
```

Section B

2. (a) Write output of the following code : (6)

```
import pandas as pd
d1 = {'Name': ['Pankaj', 'Meghna', 'Lisa'],
      'ID': [1, 2, 3],
      'Country': ['India', 'India', 'USA'],
      'Role': ['CEO', 'CTO', 'CTO']}
df1 = pd.DataFrame(d1)
df2 = pd.DataFrame({'ID': [1, 2, 3],
                    'Name': ['Pankaj', 'Anupam', 'Amit']})
print(df1.merge(df2, on='ID'))
print(df1.merge(df2, on='Name'))
print (df1.merge(df2, how='left')
```

(b) Differentiate between loc and iloc operators giving suitable examples. (4)

(c) Write code in python to create a pandas series object of all even numbers from 2 to 30, including 2 and 30, using arange(). (5)

3. (a) Explain hierarchical indexing in dataframes. (4)

(b) Write the names of functions and give example code for each in matplotlib : (6)

(i) for setting tick labels on x axes

(ii) for adding legend to a plot

(c) Given the code below (5)

```
import numpy as np
import pandas as pd
data = pd.DataFrame(np.arange(12).reshape((3, 4)),
index=['A', 'B', 'C'], columns=['one', 'two', 'three',
'four'])
```

Write the output for following :

(i) print(data)

(ii) print(data.sum())

(iii) print (data.sum(axis= 'columns'))

(iv) print (data.idxmax())

(v) print (data.cumsum())

4. (a) Explain the functions used for : (5)

(i) Removing duplicates in a dataframe

(ii) Filling missing entries in a dataframe

(b) Explain how to create subplots using matplotlib, with suitable example. (5)