

MANIPAL ACADEMY OF HIGHER EDUCATION
MANIPAL SCHOOL OF INFORMATION SCIENCES

THIRD SEMESTER MASTER OF SCIENCE – M.Sc. (INFORMATION SCIENCE) DEGREE
EXAMINATION – DECEMBER 2020

SUBJECT: MIS 609.8 – PRINCIPLES OF DATA VISUALIZATION (ELECTIVE)

Wednesday, 9th December 2020

Time: 10.00 – 13.00 Hrs.

Max. Marks: 100

Answer ALL the questions

All questions carry 10 marks.

1. Define Data Visualization. What makes good data visualization? Explain the importance of Data Analysis, Storytelling and design for effective data visualization
(2+2+6 = 10 Marks)
2. Explain the following data Visualization Categories. (2.5x4 = 10 Marks)
 - A. Multidimensional
 - B. Hierarchical
 - C. Network
 - D. Geospatial/ Planar
3. Explain Informative v/s Persuasive v/s Visual Art with neat diagram.
4. What are the three ingredients of data visualization? Explain the points to consider while building the data visualization in the context of Data and Designers objective.
(2+4+4 = 10 Marks)
5.
 - 5A. Explain Mutable and immutable types in Python
 - 5B. Explain with an example each how List Comprehension and Map function reduces the code.
(5+5 = 10 Marks)
6.
 - 6A. Using the method count, write an expression that produces number of 's' in 'manipal school of information sciences'.

6B. Using the method find, write a single expression that produces the index of second 's' in 'manipal school of information sciences'.

6C. variable 'fruit' refers to value 'pineapple'. Run the following code and explain the order of subexpression evaluation a. fruit.find('p',fruit.count('p')) b. fruit.count(fruit.upper().swapcase()) c. fruit.replace(fruit.swapcase(), fruit.lower())

(2+2+6 = 10 Marks)

7.

```
import numpy as np
y = np.arange(2,14)
y.shape
y.itemsize
y = y.reshape(3,4)
np.vstack([y,y])
y.T
```

Write the output after each line of code

8. Assume that "train.csv" contains schedule of trains. Fields in csv are "Train Number", "Train Name", "Source", "Destination" and "Distance". Write the script using pandas, to perform following tasks.

- A. List Train Number.
- B. List Train name and number which originates from a given city.
- C. List the longest running train.
- D. Check whether train runs between given pair of cities.

(2.5x4 = 10 Marks)

9.

```
html_doc = """<html><head><title>Elsie's Details</title></head>
<body>
<p class="title"><b>My Details</b></p> <div>
<p class="Details">My contact information
<a href="http://example.com/Name" class="name" id="link">Elsie</a>,
<a href="http://example.com/Mobile" class="mobile" id="link2">+91
99223344556</a>
<a href="http://example.com/email" class="email"
id="link3">elsie.george@xyz.com</a>;
and they lived at the bottom of a well.</p></div>
```

<p

Using beautiful soup, write the script to extract following details

- A. To display "My Details" from <p> tag.
- B. To display the Name, mobile number and email
- C. To display the Mobile number if it is valid (+<3 digits> < 10 digit number>

(3+3+4 = 10 Marks)

10.

```
from matplotlib import pyplot as plt
import numpy as np
fig,ax = plt.subplots(1,1)
a = np.array([22,87,5,43,56,73,55,54,11,20,51,5,79,31,27])
ax.hist(a, bins = [0,25,50,75,100])
ax.set_title("histogram of result")
ax.set_xticks([0,25,50,75,100])
ax.set_xlabel('marks')
ax.set_ylabel('no. of students')
plt.show()
```

Document every line of code and display the output in graphical format. 10 Marks
