

## 20PAIC51J – Python for Data Science

### Project Questions – Units I and II

Question 1:

Given the sets of events,

```
sports_winners = set(["Jake", "John", "Eric", "Geetha", "Femi", "Kala"])
```

```
quiz_winners = set(["John", "Jill", "Eric", "Geetha", "Femi"])
```

1. print the winners who got sports prize and quiz prize
2. print the winners who got sports prize and not quiz prize
3. print the winners who either sports prize or quiz prize
4. print the winners who got quiz prize and not sports prize

Question 2:

Consider the set with some integers given below

```
myset = {12, 34, 56, 3, 45, 67, 89, 1, 6}
```

1. Create set by filtering even numbers
2. Create set by filtering odd numbers
3. Create set by filtering numbers > 20
4. remove the item 56 and create a new set
5. state the difference between `remove()` & `discard()` on the set.

Question 3:

Date and Time

1. Print the Date after 1 week and 6 days from the current date.
2. Given the birthday as `date(2000, 10, 2)`, find the age
3. Subtract five days from the current date and display it.
4. Display the dates of yesterday, today and tomorrow

Question 4:

Consider the following strings, `string1="Great"`, `string2="Learning"` and `string3="Python"`.

Perform the following operations on the given strings.

- a. Concatenate them to create a single string called `"single_string"`
- b. Access and print the first and last characters of `"single_string"`
- c. Reverse the `"single_string"` and save it in `"rev_string"`
- d. Extract the word `"Learning"` from `"single_string"`

Question 5:

Consider the given sentence "Hello! Great Learning. Now start learning Python."

- a. In the given sentence where is the word "Learning"?:
- b. In the given sentence where is the first occurrence of the letter "e"?:
- c. In the given sentence where is the first occurrence of the letter "e" when you only search between position 5 and 10?:
- d. Display the list of words in the sentence

Question 6:

Build a program that analyzes a text document (a long string) and uses a dictionary to count the frequency of each word. You can then display the 5 most common words and their counts.

Question 7:

Given a list of tuples, consisting of `employee_name` and `total_bonus_points`. Find the best 3 employees who are having the highest score. Write a function *find\_best\_employer*, which takes the list of tuples as argument and prints the names of top three employees

Sample Input:

```
[('Sanju',321), ('Reva',671), ('Darvin',233), ('Velan',922), ('Satya',223), ('Vanathi',102)]
```

Expected Output:

The top three employees for the year are Velan, Reva and Sanju