IT2313 - Programming for Data Science

Practical 3

NumPy for Numerical Computations (Part 1)

Let try to practise our NumPy knowledge. We will start off with a few simple tasks in this Practical

Creating NumPy Arrays

Q1 - Import NumPy as np

```
In [1]:  # Type your answer here
```

Q2 - Create an array of 10 zeros

```
In [2]:  # Type your answer here
```
```
Out[2]:  array([0., 0., 0., 0., 0., 0., 0., 0., 0., 0.])
```

Q3 - Create an array of 10 ones

```
In [3]:  # Type your answer here
```
```
Out[3]:  array([1., 1., 1., 1., 1., 1., 1., 1., 1., 1.])
```

Q4 - Create an array of 10 fives

```
In [4]:  # Type your answer here
```
```
Out[4]:  array([5., 5., 5., 5., 5., 5., 5., 5., 5., 5.])
```

```
In [5]:  # Type your answer here
```
```
Out[5]:  array([5., 5., 5., 5., 5., 5., 5., 5., 5., 5.])
```

Q5 - Create an array of the integers from 20 to 40

```
In [6]:  # Type your answer here
```
```
Out[6]:  array([20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36,
                37, 38, 39, 40])
```

Q6 - Create an array of all the Even integers from 20 to 40

```
In [7]:  # Type your answer here
         )
```
```
Out[7]:  array([20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40])
```

Q7 - Create an 10 evenly spaced values from 20 to 40

The formula for the step size (step) is calculated as:

step = (stop - start) / (num - 1), num is the number of evenly spaced values to generate

In [8]:
```
# Type your answer here
```

Out[8]:
```
array([20.        , 22.22222222, 24.44444444, 26.66666667, 28.88888889,
       31.11111111, 33.33333333, 35.55555556, 37.77777778, 40.        ])
```

Q8 - Create a 3x3 matrix with values ranging from 0 to 8

In [9]:
```
# Type your answer here
```

Out[9]:
```
array([[0, 1, 2],
       [3, 4, 5],
       [6, 7, 8]])
```

In [10]:
```
# Type your answer here
```

Out[10]:
```
array([[0, 1, 2],
       [3, 4, 5],
       [6, 7, 8]])
```

Q9 - Create a 3x3 matrix with values of 6

In [11]:
```
# Type your answer here
```

Out[11]:
```
array([[5, 5, 5],
       [5, 5, 5],
       [5, 5, 5]])
```

Q10 - Use NumPy to generate 10 random number between 0 and 1

In [12]:
```
# Type your answer here
```

Out[12]:
```
array([0.30841917, 0.33134378, 0.86502522, 0.23380227, 0.42341794,
       0.30821873, 0.56637339, 0.69045236, 0.87683985, 0.45919358])
```

Q11 - Use NumPy to create the following matrix

Note - Use 2 different approaches

In [13]:
```
# Type your answer here
```

Out[13]:
```
array([[0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1 ],
       [0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.2 ],
       [0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3 ],
       [0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4 ],
       [0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5 ],
       [0.51, 0.52, 0.53, 0.54, 0.55, 0.56, 0.57, 0.58, 0.59, 0.6 ],
       [0.61, 0.62, 0.63, 0.64, 0.65, 0.66, 0.67, 0.68, 0.69, 0.7 ],
       [0.71, 0.72, 0.73, 0.74, 0.75, 0.76, 0.77, 0.78, 0.79, 0.8 ],
       [0.81, 0.82, 0.83, 0.84, 0.85, 0.86, 0.87, 0.88, 0.89, 0.9 ],
       [0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 1.  ]])
```

In [14]:
```
# Type your answer here
```

Out[14]:
```
array([[0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1 ],
       [0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.2 ],
       [0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3 ],
       [0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4 ],
       [0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5 ],
       [0.51, 0.52, 0.53, 0.54, 0.55, 0.56, 0.57, 0.58, 0.59, 0.6 ],
       [0.61, 0.62, 0.63, 0.64, 0.65, 0.66, 0.67, 0.68, 0.69, 0.7 ],
       [0.71, 0.72, 0.73, 0.74, 0.75, 0.76, 0.77, 0.78, 0.79, 0.8 ],
       [0.81, 0.82, 0.83, 0.84, 0.85, 0.86, 0.87, 0.88, 0.89, 0.9 ],
       [0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 1.  ]])
```

Creating NumPy Indexing and Selection

Question 12 - You will be giving a few matrices, and required to to produce the resulting matrix outputs

```
In [15]: mat = np.arange(1,26).reshape(5,5)
         mat
```

```
Out[15]: array([[ 1,  2,  3,  4,  5],
                [ 6,  7,  8,  9, 10],
                [11, 12, 13, 14, 15],
                [16, 17, 18, 19, 20],
                [21, 22, 23, 24, 25]])
```

```
In [16]: # Q12(a)
         # Type your answer here
```

```
Out[16]: array([[11, 12, 13, 14, 15],
                [16, 17, 18, 19, 20],
                [21, 22, 23, 24, 25]])
```

```
In [17]: # Q12(b)
         # Type your answer here
```

```
Out[17]: array([[12, 13, 14, 15],
                [17, 18, 19, 20],
                [22, 23, 24, 25]])
```

```
In [18]: # Q12(c)
         # Type your answer here
```

```
Out[18]: 20
```

```
In [19]: # Q12(e)
         # Type your answer here
```

```
Out[19]: array([ 2,  7, 12])
```

```
In [20]: # Q12(e)
         # Type your answer here
```

```
Out[20]: array([[ 2],
                [ 7],
                [12]])
```

```
In [21]: # Q12(f)
         # Type your answer here
```

```
Out[21]: array([21, 22, 23, 24, 25])
```

```
In [22]: # Q12(g)
         # Type your answer here
```

```
Out[22]: array([[16, 17, 18, 19, 20],
                [21, 22, 23, 24, 25]])
```

Q13 - Get the sum of all the values in the matrix

```
In [23]: # Type your answer here
         mat = np.arange(5,14).reshape(3,3)
         mat
```

```
Out[23]: array([[ 5,  6,  7],
                [ 8,  9, 10],
                [11, 12, 13]])
```

```
In [24]:  # Type your answer here

Out[24]:  81


In [25]:  # Type your answer here

Out[25]:  81
```

Sorting and Filtering of NumPy Arrays

Q13 - The medals_2023,csv file contains all the medals won by countries in SEA Games 2023

```
In [26]:  import csv
          csvfile = open("medals_2023.csv")
          for row in csv.reader(csvfile):
              print(row)
          csvfile.close()
```

```
['Rank', 'Country', 'Gold', 'Silver', 'Bronze', 'Total']
['1', 'Vietnam', '136', '105', '114', '355']
['2', 'Thailand', '108', '96', '109', '313']
['3', 'Indonesia', '87', '80', '109', '276']
['4', 'Cambodia', '81', '74', '127', '282']
['5', 'Philippines', '58', '86', '116', '260']
['6', 'Singapore', '51', '42', '64', '157']
['7', 'Malaysia', '34', '45', '96', '175']
['8', 'Myanmar', '21', '25', '68', '114']
['9', 'Laos', '6', '22', '60', '88']
['10', 'Brunei', '2', '1', '6', '9']
['11', 'Timor-Leste', '0', '0', '8', '8']
```

Q13(a) - Use the data you read from csv file and convert it into NumPy array and display the data type

```
In [27]:  # Q13(a)
          # This array was created
          medals = [
              ['Rank', 'Country', 'Gold', 'Silver', 'Bronze', 'Total'],
              ['1', 'Vietnam', '136', '105', '114', '355'],
              ['2', 'Thailand', '108', '96', '109', '313'],
              ['3', 'Indonesia', '87', '80', '109', '276'],
              ['4', 'Cambodia', '81', '74', '127', '282'],
              ['5', 'Philippines', '58', '86', '116', '260'],
              ['6', 'Singapore', '51', '42', '64', '157'],
              ['7', 'Malaysia', '34', '45', '96', '175'],
              ['8', 'Myanmar', '21', '25', '68', '114'],
              ['9', 'Laos', '6', '22', '60', '88'],
              ['10', 'Brunei', '2', '1', '6', '9'],
              ['11', 'Timor-Leste', '0', '0', '8', '8']
          ]

          # Convert the data to NumPy array
          # Type your answer here

          # Print the NumPy array
          # Type your answer here
```

```
[['Rank' 'Country' 'Gold' 'Silver' 'Bronze' 'Total']
 ['1' 'Vietnam' '136' '105' '114' '355']
 ['2' 'Thailand' '108' '96' '109' '313']
 ['3' 'Indonesia' '87' '80' '109' '276']
 ['4' 'Cambodia' '81' '74' '127' '282']
 ['5' 'Philippines' '58' '86' '116' '260']
 ['6' 'Singapore' '51' '42' '64' '157']
 ['7' 'Malaysia' '34' '45' '96' '175']
 ['8' 'Myanmar' '21' '25' '68' '114']
 ['9' 'Laos' '6' '22' '60' '88']
 ['10' 'Brunei' '2' '1' '6' '9']
 ['11' 'Timor-Leste' '0' '0' '8' '8']]
```

In [28]:
```
# Q13(a)
# Type your answer here
```

Out[28]:
```
numpy.ndarray
```

Q13(b) - Only Display the Top 3 medal ranking countries

In [29]:
```
# Q13(b)
# Type your answer here
# Extract the header and the top 3 rows


# Print the header and top 3 countries
```

```
['Rank' 'Country' 'Gold' 'Silver' 'Bronze' 'Total']
[['1' 'Vietnam' '136' '105' '114' '355']
 ['2' 'Thailand' '108' '96' '109' '313']
 ['3' 'Indonesia' '87' '80' '109' '276']]
```

Q13(c) - Display all the medals won by Singapore

In [30]:
```
# Q13(c)
# Type your answer here
# Create a boolean mask for rows where 'Country' is 'Singapore'


# Use boolean indexing to extract the rows for Singapore


# Print the medals won by Singapore
```

```
['Rank' 'Country' 'Gold' 'Silver' 'Bronze' 'Total']
[['6' 'Singapore' '51' '42' '64' '157']]
```

Q13(d) - Display only the Gold and Total Medals won by the Countries

In [31]:
```
# Q13(d)
# Type your answer here
# Keep 'Rank', 'Country', 'Gold', and 'Total' columns


# Print the selected columns for all countries
```

```
[['Rank' 'Country' 'Gold' 'Total']
 ['1' 'Vietnam' '136' '355']
 ['2' 'Thailand' '108' '313']
 ['3' 'Indonesia' '87' '276']
 ['4' 'Cambodia' '81' '282']
 ['5' 'Philippines' '58' '260']
 ['6' 'Singapore' '51' '157']
 ['7' 'Malaysia' '34' '175']
 ['8' 'Myanmar' '21' '114']
 ['9' 'Laos' '6' '88']
 ['10' 'Brunei' '2' '9']
 ['11' 'Timor-Leste' '0' '8']]
```

Q13(e) Remove the Silver and Bronze column

In [32]:
```python
# Q13(e)
# Type your answer here
# Remove 'Silver' and 'Bronze' columns (columns 3 and 4)


# Print the modified NumPy array
```

```
[['Rank' 'Country' 'Gold' 'Total']
 ['1' 'Vietnam' '136' '355']
 ['2' 'Thailand' '108' '313']
 ['3' 'Indonesia' '87' '276']
 ['4' 'Cambodia' '81' '282']
 ['5' 'Philippines' '58' '260']
 ['6' 'Singapore' '51' '157']
 ['7' 'Malaysia' '34' '175']
 ['8' 'Myanmar' '21' '114']
 ['9' 'Laos' '6' '88']
 ['10' 'Brunei' '2' '9']
 ['11' 'Timor-Leste' '0' '8']]
```

Q13(f) - Sort the Medal Tally Table based on the Total Medals won by countries

In [33]:
```python
# Q13(f)

import numpy as np

medals = [
    ['Rank', 'Country', 'Gold', 'Silver', 'Bronze', 'Total'],
    ['1', 'Vietnam', '136', '105', '114', '355'],
    ['2', 'Thailand', '108', '96', '109', '313'],
    ['3', 'Indonesia', '87', '80', '109', '276'],
    ['4', 'Cambodia', '81', '74', '127', '282'],
    ['5', 'Philippines', '58', '86', '116', '260'],
    ['6', 'Singapore', '51', '42', '64', '157'],
    ['7', 'Malaysia', '34', '45', '96', '175'],
    ['8', 'Myanmar', '21', '25', '68', '114'],
    ['9', 'Laos', '6', '22', '60', '88'],
    ['10', 'Brunei', '2', '1', '6', '9'],
    ['11', 'Timor-Leste', '0', '0', '8', '8']
]

numpy_array = np.array(medals[1:], dtype=object)

# Convert 'Total' column to integers for sorting
# Type your answer here

# Get indices that would sort the 'Total' column in descending order
# Type your answer here

# Sort the NumPy array based on the sorted indices
# Type your answer here

# Add the header row back to the sorted array
```

```python
# Type your answer here

# Print the sorted NumPy array
# Type your answer here
```

```
[['Rank' 'Country' 'Gold' 'Silver' 'Bronze' 'Total']
 ['1' 'Vietnam' '136' '105' '114' '355']
 ['2' 'Thailand' '108' '96' '109' '313']
 ['4' 'Cambodia' '81' '74' '127' '282']
 ['3' 'Indonesia' '87' '80' '109' '276']
 ['5' 'Philippines' '58' '86' '116' '260']
 ['7' 'Malaysia' '34' '45' '96' '175']
 ['6' 'Singapore' '51' '42' '64' '157']
 ['8' 'Myanmar' '21' '25' '68' '114']
 ['9' 'Laos' '6' '22' '60' '88']
 ['10' 'Brunei' '2' '1' '6' '9']
 ['11' 'Timor-Leste' '0' '0' '8' '8']]
```

In [34]:
```python
# Q13(f)

import numpy as np

medals = [
    ['Rank', 'Country', 'Gold', 'Silver', 'Bronze', 'Total'],
    ['1', 'Vietnam', '136', '105', '114', '355'],
    ['2', 'Thailand', '108', '96', '109', '313'],
    ['3', 'Indonesia', '87', '80', '109', '276'],
    ['4', 'Cambodia', '81', '74', '127', '282'],
    ['5', 'Philippines', '58', '86', '116', '260'],
    ['6', 'Singapore', '51', '42', '64', '157'],
    ['7', 'Malaysia', '34', '45', '96', '175'],
    ['8', 'Myanmar', '21', '25', '68', '114'],
    ['9', 'Laos', '6', '22', '60', '88'],
    ['10', 'Brunei', '2', '1', '6', '9'],
    ['11', 'Timor-Leste', '0', '0', '8', '8']
]

# Convert the data to NumPy array, excluding the first row
# Type your answer here

# Remove 'Silver' and 'Bronze' columns (columns 3 and 4)
# Type your answer here

# Convert 'Total' column to integers for sorting
# Type your answer here

# Get indices that would sort the 'Total' column in descending order
# Type your answer here

# Sort the NumPy array based on the sorted indices
# Type your answer here

# Add the header row back to the sorted array
# Type your answer here

# Print the sorted NumPy array without 'Silver' and 'Bronze'
# Type your answer here
```

```
[['Rank' 'Country' 'Gold' 'Total']
 ['1' 'Vietnam' '136' '355']
 ['2' 'Thailand' '108' '313']
 ['4' 'Cambodia' '81' '282']
 ['3' 'Indonesia' '87' '276']
 ['5' 'Philippines' '58' '260']
 ['7' 'Malaysia' '34' '175']
 ['6' 'Singapore' '51' '157']
 ['8' 'Myanmar' '21' '114']
 ['9' 'Laos' '6' '88']
 ['10' 'Brunei' '2' '9']
 ['11' 'Timor-Leste' '0' '8']]
```

```
[['Rank' 'Country' 'Gold' 'Total']
 ['1' 'Vietnam' '136' '355']
 ['2' 'Thailand' '108' '313']
 ['4' 'Cambodia' '81' '282']
 ['3' 'Indonesia' '87' '276']
 ['5' 'Philippines' '58' '260']
 ['7' 'Malaysia' '34' '175']
 ['6' 'Singapore' '51' '157']
 ['8' 'Myanmar' '21' '114']
 ['9' 'Laos' '6' '88']
 ['10' 'Brunei' '2' '9']
 ['11' 'Timor-Leste' '0' '8']]
```