

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: Annie Sadler Kamanya |
| Course Date: 03/02/2025 |
|  |

**Table of contents**

[Day 2: Task 1 3](#_Toc189838669)

[**Day 3: Task 1** 4](#_Toc189838670)

[**Exercise 1: Loading and Exploring the Data** 4](#_Toc189838671)

[**Exercise 2: Indexing and Slicing** 4](#_Toc189838672)

[**Exercise 3: Data Manipulation** 5](#_Toc189838673)

[**Day 4: Task 1** 6](#_Toc189838674)

[**Day 4: Task 2** 7](#_Toc189838675)

[**Course Notes** 8](#_Toc189838676)

[**Additional Information** 9](#_Toc189838677)

# Day 2: Task 1

It is a common software development interview question to create the below with a certain programming language. Create the below using Python syntax, test it and past the completed syntax and output below.

FizzBuzz:

Go through the integers from 1 to 100.

If a number is divisible by 3, print "fizz."

If a number is divisible by 5, print "buzz."

If a number is both divisible by 3 and by 5, print "fizzbuzz."

Otherwise, print just the number.

|  |  |
| --- | --- |
| Paste your completed work to the right | i = 1  while i<=100:  if i%3==0 and i%5==0:  print("FizzBuzz")  elif i%3==0:  print("Fizz")  elif i%5==0:  print("Buzz")  else:  print(i)  i +=1 |

# **Day 3: Task 1**

Using the ‘student.csv’ which can be downloaded [here](https://justit831-my.sharepoint.com/:x:/g/personal/danpe_justit_co_uk/ER92LoQB1PpNqWj07fnfO4EBh9HB7CiI-i4RH273HoqY6A?e=mVdIeY), complete the below exercises and paste your input and output.

### **Exercise 1: Loading and Exploring the Data**

1. Question: "Write the code to read a CSV file into a Pandas DataFrame."
2. Question: "Write the code to display the first 5 rows of the DataFrame."
3. Question: "Write the code to get the information about the DataFrame."
4. Question: "Write the code to get summary statistics for the DataFrame."

|  |
| --- |
| 1. # Read the CSV(excel)file into a DataFrame  df = pd.read\_excel('/content/student (2).xlsx')  print(df.head())  2. # Display the first 5 rows of the DataFrame  df = pd.read\_excel('/content/student (2).xlsx')  print(df.head())  3. # code for information about DataFrame  df = pd.read\_excel('/content/student (2).xlsx')  df.info()  4. # Read the excel file into a DataFrame  df = pd.read\_excel('/content/student (2).xlsx')  summary\_statistics = df.describe()  print(summary\_statistics) |

### **Exercise 2: Indexing and Slicing**

1. Question: "Write the code to select the 'name' column."
2. Question: "Write the code to select the 'name' and 'mark' columns."
3. Question: "Write the code to select the first 3 rows."
4. Question: "Write the code to select all rows where the 'class' is 'Four'."

|  |
| --- |
| 1. # Select the 'name' column   name\_column = df['name']  # Display the 'name' column  print(name\_column)   1. # Select the 'name' and 'mark' columns   selected\_columns = df[['name', 'mark']]    print(selected\_columns)   1. # code: select the first 3 rows."   first\_3\_rows = df.head(3)    print(first\_3\_rows)   1. # select all rows where the 'class' is 'Four'."   class\_Four\_rows = df[df['class'] == 'Four']  print(class\_Four\_rows) |

### 

### **Exercise 3: Data Manipulation**

1. Question: "Write the code to add a new column 'passed' that indicates whether the student passed (mark >= 60)."
2. Question: "Write the code to rename the 'mark' column to 'score'."
3. Question: "Write the code to drop the 'passed' column."

|  |
| --- |
| 1. # add new column 'passed' mark >=60   df['passed'] = df['mark'] >=60  print(df)   1. # Rename 'mark' column as 'Score'   df = df.rename(columns = {'mark' : 'Score'})  print(df)   1. # drop 'passed' column   df.drop(columns=['passed'], inplace=True)  print(df) |

# **Day 4: Task 1**

Using the ‘GDP (nominal) per Capita.csv’ which can be downloaded [here](https://justit831-my.sharepoint.com/:x:/g/personal/danpe_justit_co_uk/EV1Xzb5eNENHmOVMDssxyoMBqTCVcLg18U4qOLUDZZHSkw?e=PAbKfN), complete the below exercises and paste your input and output. Work individually, but we will work and support each other in the room.

* Read and save the ‘GDP (nominal) per Capita’ data to a data frame called “df” in Jupyter notebook
* Print the first 10 rows
* Print the last 5 rows
* Print ‘Country/Territory’ and ‘UN\_Region’ columns

|  |
| --- |
| Input  print(df.head())  Out-put    # print the first 10 rows  print(df.head(10)) |

# **Day 4: Task 2**

Back with ‘GDP (nominal) per Capita’. As a group, import and work your way through the Day\_4\_Python\_Activity.ipynb notebook which can be found [here](https://justit831-my.sharepoint.com/:u:/g/personal/danpe_justit_co_uk/Ede5Pb1JwPNMj49hTDzeEUMB7GZWBP7SVidCo0Gt6tnP1w?e=D83SIR). There are questions to answer, but also opportunities to have fun with the data – paste your input and output below.

Once complete, and again as a group, work with some more data and have some fun –there is no set agenda for this section, other than to embed the skills developed this week. Paste your input and output below and upon return we’ll discuss progress made.

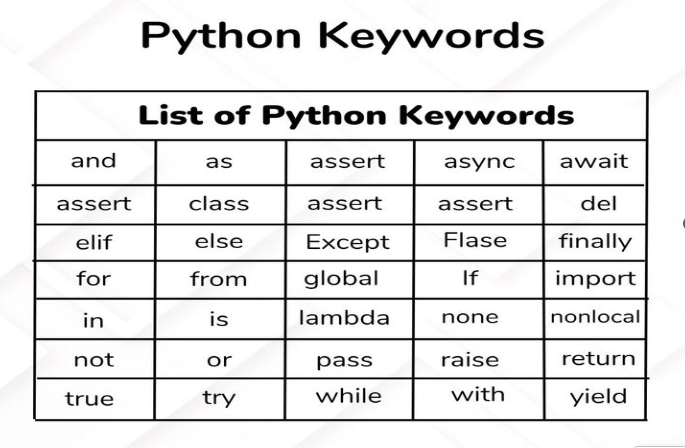
[Additional data found here.](https://justit831-my.sharepoint.com/:f:/g/personal/danpe_justit_co_uk/Er0ybU9i0AZKiuGaCWZyj2ABoqKD23zwLGdJf3WlaixpRA?e=QVj2Bs)

|  |
| --- |
|  |

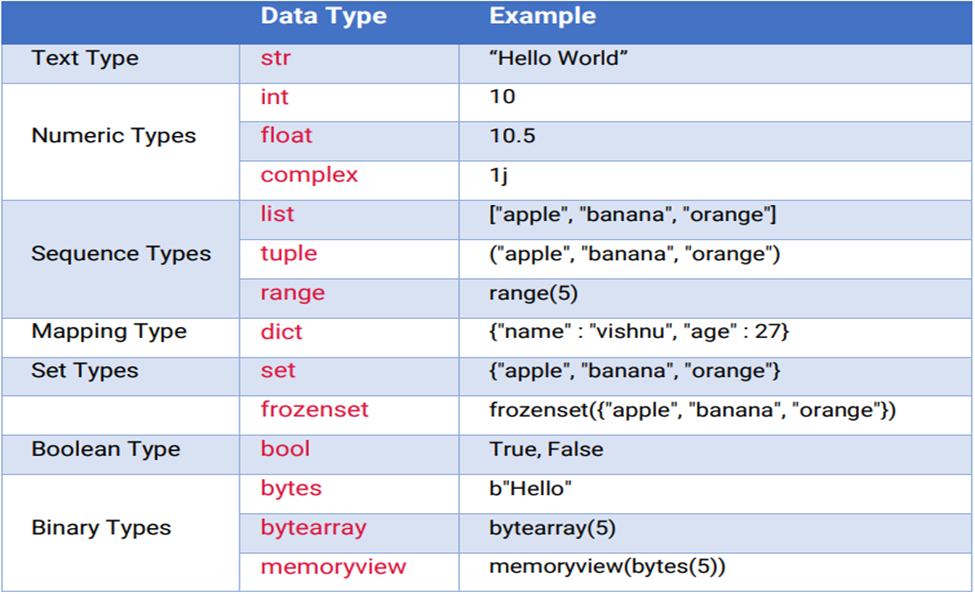
|  |
| --- |
| **Course Notes** |

It is recommended to take notes from the course, use the space below to do so, or use the revision guide shared with the class:

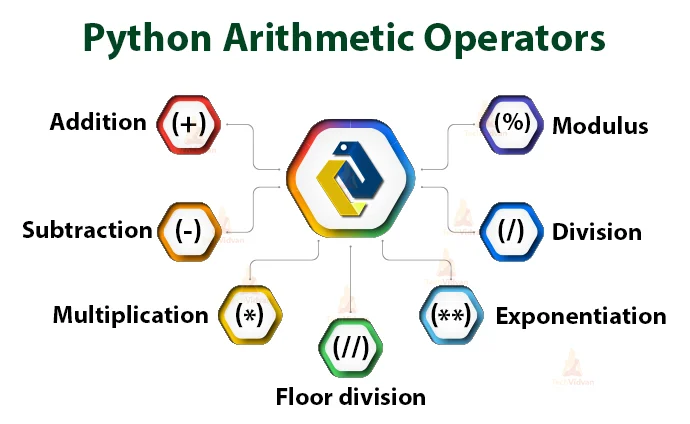
**Python is Case sensitive**

**Python Keywords**

**DATA TYPES- Built-in Python by default**



**Arithmetic Operators**



**Other escape character used in Python**

|  |  |
| --- | --- |
| **code** | **Result** |
| \’ | Single Quote |
| \\ | Backslash |
| \n | New Line |
| \t | Tab |
| \r | Carriage Return |
| \b | Backspace |

|  |
| --- |
| **Additional Information** |

We have included a range of additional links to further resources and information that you may find useful, these can be found within your revision guide.

**END OF WORKBOOK**

**Please check through your work thoroughly before submitting and update the table of contents if required.**

**Please send your completed work booklet to your trainer.**