|  |
| --- |
|  |

|  |
| --- |
| Name: |
| Course Date: |
|  |

**Table of contents**

[Day 2: Task 1 3](#_Toc1765240926)

[Day 3: Task 1 3](#_Toc873857632)

[Exercise 1: Loading and Exploring the Data 4](#_Toc1363711086)

[Exercise 2: Indexing and Slicing 4](#_Toc1679584900)

[Exercise 3: Data Manipulation 4](#_Toc998449442)

[Exercise 4: Aggregation and Grouping 5](#_Toc267294297)

[Exercise 5: Advanced Operations 5](#_Toc1232452443)

[Exercise 6: Exporting Data 6](#_Toc1251390377)

[Exercise 7: If finished early try visualising the results 6](#_Toc66584577)

[Day 4: Task 1 7](#_Toc1452452084)

[Day 4: Task 2 8](#_Toc934594145)

[Course Notes 8](#_Toc1227711066)

[Additional Information 9](#_Toc1836543406)

# 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 | 1, 2, fizz, 4, buzz, fizz, 7, 8, fizz, buzz, 11, fizz, 13, 14, fizz, buzz, 16, 17, fizz, 19, buzz, fizz, 22, 23, fizz, buzz, 26, fizz, 28, 29, fizzbuzz, 31, 32, fizz, 34, buzz, fizz, 37, 38, fizz, buzz, 41, fizz, 43, 44, fizzbuzz, 46, 47, fizz, 49, buzz, fizz, 52, 53, fizz, buzz, 56, fizz, 58, 59, fizzbuzz, 61, 62, fizz, 64, buzz, fizz, 67, 68, fizz, buzz, 71, fizz, 73, 74, fizzbuzz, 76, 77, fizz, 79, buzz, fizz, 82, 8 3, fizz, buzz, 86, fizz, 88, 89, fizzbuzz, 91, 92, fizz, 94, buzz, fizz, 97, 98, fizz, buzz. |

# **Day 3: Task 1**

Download the ‘student.csv’, complete the below exercises as a group and paste your input and output. Although this is a group activity, everyone should have the below answered so it supports your portfolio:

### **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."

|  |
| --- |
| import pandas as pd  df = pd.read\_csv('your\_file.csv') # Replace with your actual file path  print(df.head())  print(df.info())  print(df.describe())  print(df.describe(include='all')) |

### **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'."

|  |
| --- |
| df['name']  df[['name', 'mark']]  df.head(3)  df[df['class'] == 'Four'] |

### **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."

|  |
| --- |
| df['passed'] = df['mark'] >= 60  df.rename(columns={'mark': 'score'}, inplace=True)  df.drop(columns=['passed'], inplace=True) |

### **Exercise 4: Aggregation and Grouping**

1. Question: "Write the code to group the DataFrame by the 'class' column and calculate the mean 'mark' for each group."
2. Question: "Write the code to count the number of students in each class."
3. Question: "Write the code to calculate the average mark for each gender."

|  |
| --- |
| df.groupby('class')['mark'].mean()  df['class'].value\_counts()  df.groupby('gender')['mark'].mean() |

### **Exercise 5: Advanced Operations**

1. Question: "Write the code to create a pivot table with 'class' as rows, 'gender' as columns, and 'mark' as values."
2. Question: "Write the code to create a new column 'grade' where marks >= 85 are 'A', 70-84 are 'B', 60-69 are 'C', and below 60 are 'D'."
3. Question: "Write the code to sort the DataFrame by 'mark' in descending order."

|  |
| --- |
| df.pivot\_table(values='mark', index='class', columns='gender', aggfunc='mean')  df['grade'] = pd.cut(df['mark'], bins=[0, 59, 69, 84, 100], labels=['D', 'C', 'B', 'A'])  df.sort\_values(by='mark', ascending=False, inplace=True) |

### **Exercise 6: Exporting Data**

1. Question: "Write the code to save the DataFrame with the new 'grade' column to a new CSV file."

|  |
| --- |
| df.to\_csv('students\_with\_grades.csv', index=False) |

### **Exercise 7: If finished early try visualising the results**

|  |
| --- |
| I am unsure about how to do visualising results. |

# **Day 4: Task 1**

Using the ‘GDP (nominal) per Capita.csv’ which can be downloaded from the shared Folder, 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 Jyputer notebook
* Print the first 10 rows
* Print the last 5 rows
* Print ‘Country/Territory’ and ‘UN\_Region’ columns

|  |
| --- |
| First 10 rows:  Unnamed: 0 Country/Territory UN\_Region IMF\_Estimate IMF\_Year WorldBank\_Estimate WorldBank\_Year UN\_Estimate UN\_Year  0 1 Monaco Europe 0 0 234316 2021 234317 2021  1 2 Liechtenstein Europe 0 0 157755 2020 169260 2021  2 3 Luxembourg Europe 132372 2023 133590 2021 133745 2021  3 4 Ireland Europe 114581 2023 100172 2021 101109 2021  4 5 Bermuda Americas 0 0 114090 2021 112653 2021  5 6 Norway Europe 101103 2023 89154 2021 89242 2021  6 7 Switzerland Europe 98767 2023 91992 2021 93525 2021  7 8 Singapore Asia 91100 2023 72794 2021 66822 2021  8 9 Isle of Man Europe 0 0 87158 2019 0 0  9 10 Cayman Islands Americas 0 0 86569 2021 85250 2021  Last 5 rows:  Unnamed: 0 Country/Territory UN\_Region IMF\_Estimate IMF\_Year WorldBank\_Estimate WorldBank\_Year UN\_Estimate UN\_Year  218 219 Malawi Africa 496 2023 635 2021 613 2021  219 220 South Sudan Africa 467 2023 1072 2015 400 2021  220 221 Sierra Leone Africa 415 2023 480 2021 505 2021  221 222 Afghanistan Asia 611 2020 369 2021 373 2021  222 223 Burundi Africa 249 2023 222 2021 311 2021 |

# **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 on the shared Folder. 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)

|  |
| --- |
| I am unsure about this activity and would prefer to ask my tutor if we can do this together as a group. |

|  |
| --- |
| **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:

|  |
| --- |
| **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.**