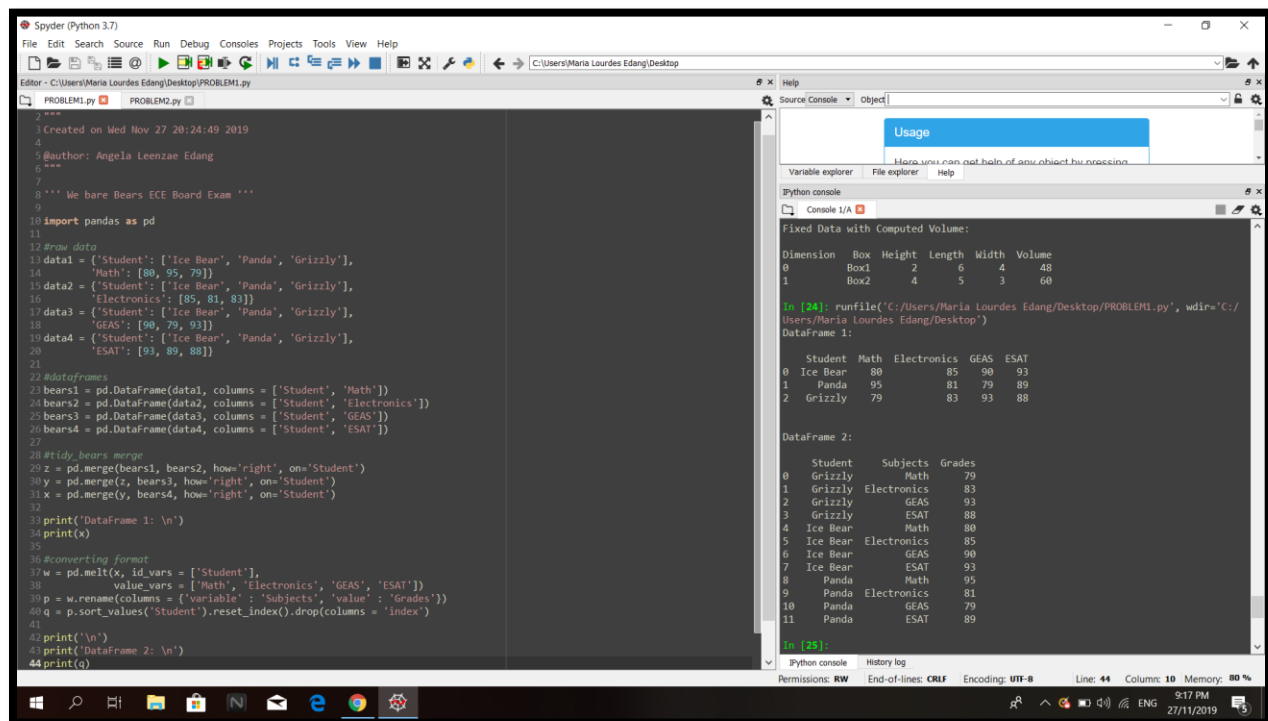


EXPERIMENT 9

DATA WRANGLING USING PANDAS

Problem 1:



The screenshot shows the Spyder Python IDE with a script named `PROBLEM1.py` and its output in the console. The script creates four DataFrames, merges them, and then melts and sorts the result.

```
1 Created on Wed Nov 27 20:24:49 2019
2
3 @author: Angela Leenza Edang
4
5 """ We bare Bears ECE Board Exam """
6
7
8 import pandas as pd
9
10 #raw data
11
12 data1 = {'Student': ['Ice Bear', 'Panda', 'Grizzly'],
13         'Math': [80, 95, 79]}
14
15 data2 = {'Student': ['Ice Bear', 'Panda', 'Grizzly'],
16         'Electronics': [85, 81, 83]}
17
18 data3 = {'Student': ['Ice Bear', 'Panda', 'Grizzly'],
19         'GEAS': [90, 79, 93]}
20
21 data4 = {'Student': ['Ice Bear', 'Panda', 'Grizzly'],
22         'ESAT': [93, 89, 88]}
23
24 #dataframes
25
26 bears1 = pd.DataFrame(data1, columns = ['Student', 'Math'])
27 bears2 = pd.DataFrame(data2, columns = ['Student', 'Electronics'])
28 bears3 = pd.DataFrame(data3, columns = ['Student', 'GEAS'])
29 bears4 = pd.DataFrame(data4, columns = ['Student', 'ESAT'])
30
31 #tidy bears merge
32
33 z = pd.merge(bears1, bears2, how='right', on='Student')
34 y = pd.merge(z, bears3, how='right', on='Student')
35 x = pd.merge(y, bears4, how='right', on='Student')
36
37 print(DataFrame 1: \n')
38 print(x)
39
40 #converting format
41
42 w = pd.melt(x, id_vars = ['Student'],
43            value_vars = ['Math', 'Electronics', 'GEAS', 'ESAT'])
44 p = w.rename(columns = {'variable': 'Subjects', 'value': 'Grades'})
45 q = p.sort_values('Student').reset_index().drop(columns = 'index')
46
47 print('\n')
48 print(DataFrame 2: \n')
49 print(q)
```

The console output shows the execution of the script, including the creation of DataFrames and the final sorted result:

```
In [24]: runfile('C:/Users/Maria Lourdes Edang/Desktop/PROBLEM1.py', wdir='C:/Users/Maria Lourdes Edang/Desktop')
DataFrame 1:
   Student  Math  Electronics  GEAS  ESAT
0  Ice Bear    80          85    90    93
1   Panda    95          81    79    89
2  Grizzly    79          83    93    88

DataFrame 2:
   Student  Subjects  Grades
0  Grizzly    Math    79
1  Grizzly  Electronics    83
2  Grizzly    GEAS    93
3  Grizzly    ESAT    88
4  Ice Bear    Math    80
5  Ice Bear  Electronics    85
6  Ice Bear    GEAS    90
7  Ice Bear    ESAT    93
8   Panda    Math    95
9   Panda  Electronics    81
10  Panda    GEAS    79
11  Panda    ESAT    89
```

Problem 2:

The screenshot shows the Spyder Python IDE interface. The main editor displays a Python script for computing the volume of boxes. The script includes comments, imports pandas, creates a DataFrame, pivots it to tidy format, and calculates the volume. The Python console on the right shows the execution output, including the raw data and the fixed data with computed volume.

```
1 """
2 Created on Wed Nov 27 20:43:31 2019
3
4
5 @author: Angela Leenzae Edang
6 """
7
8 ''' Computing Volume of a Box using Length, Width and, Height Columns '''
9
10 import pandas as pd
11
12 #Raw data
13 data = {'Box': ['Box1', 'Box1', 'Box1',
14               'Box2', 'Box2', 'Box2'],
15         'Dimension': ['Length', 'Width', 'Height',
16                      'Length', 'Width', 'Height'],
17         'Value': [6, 4, 2, 5, 3, 4]}
18
19 messy = pd.DataFrame(data, columns = ['Box', 'Dimension', 'Value'])
20 tidy = messy.pivot_table(index = 'Box', columns = 'Dimension',
21                          values = 'Value').reset_index()
22
23 print('Raw Data: \n')
24 print(messy)
25
26 #Fixed data
27 tidy['Volume'] = tidy.Height*tidy.Length*tidy.Width
28
29 print('\n')
30 print('Fixed Data with Computed Volume: \n')
31 print(tidy)
```

Usage

Here you can get help of any object by pressing

Variable explorer | File explorer | Help

Python console

Console 1/A

```
In [26]: runfile('C:/Users/Maria Lourdes Edang/Desktop/PROBLEM2.py', wdir='C:/Users/Maria Lourdes Edang/Desktop')
Raw Data:
   Box Dimension  Value
0  Box1   Length     6
1  Box1   Width     4
2  Box1   Height     2
3  Box2   Length     5
4  Box2   Width     3
5  Box2   Height     4

Fixed Data with Computed Volume:
   Dimension  Box  Height  Length  Width  Volume
0         Box1    2      6      4     48
1         Box2    4      5      3     60

In [27]:
```

Python console | History log

Permissions: RW End-of-lines: CRLF Encoding: UTF-8 Line: 31 Column: 13 Memory: 77 %

9:23 PM 27/11/2019