

Monoso, Jose Miguel D.
2ECE – C

Data Wrangling with Pandas

Problem 1:

The screenshot shows the Spyder Python IDE with the following code in `DataWrangling_1.py`:

```
1 import pandas as pd
2
3 a = {"Student": ["Ice Bear", "Panda", "Grizzly"],
4      "Math": [80, 95, 79]}
5 a = pd.DataFrame(a, columns = ["Student", "Math"])
6
7 b = {"Student": ["Ice Bear", "Panda", "Grizzly"],
8      "Electronics": [85, 81, 83]}
9 b = pd.DataFrame(b, columns = ["Student", "Electronics"])
10
11 c = {"Student": ["Ice Bear", "Panda", "Grizzly"],
12      "GEAS": [90, 79, 93]}
13 c = pd.DataFrame(c, columns = ["Student", "GEAS"])
14
15 d = {"Student": ["Ice Bear", "Panda", "Grizzly"],
16      "ESAT": [93, 89, 88]}
17 d = pd.DataFrame(d, columns = ["Student", "ESAT"])
18
19 e = pd.merge(a, b, how="outer", on="Student")
20
21 f = pd.merge(e, c, how="outer", on="Student")
22
23 print(" ")
24 g = pd.merge(f, d, how="outer", on="Student")
25 print(g)
26 print(" ")
27
28 longformat = pd.melt(g, id_vars = "Student",
29                      value_vars = ["Math", "Electronics", "GEAS", "ESAT"])
30
31 print(longformat.rename(columns={"variable": "Subject", "value": "Grades"}))
32
33
34
```

The IPython console shows the output of the code:

```
In [5]: runfile('C:/Users/Miguel Monoso/.spyder-py3/DataWrangling_1.py', wdir='C:/Users/Miguel Monoso/.spyder-py3')

  Student  Math  Electronics  GEAS  ESAT
0  Ice Bear    80         85    90    93
1   Panda     95         81    79    89
2  Grizzly    79         83    93    88

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

Permissions: RW End-of-lines: CRLF Encoding: UTF-8 Line: 33 Column: 1 Memory: 44 %

Problem 2:

The screenshot shows the Spyder Python IDE with the following code in `DataWrangling_2.py`:

```
1 import pandas as pd
2
3 messy = {"Box": ["Box1", "Box1", "Box1", "Box2", "Box2", "Box2"],
4          "Dimension": ["Length", "Width", "Height", "Length", "Width", "Height"],
5          "Value": [6, 4, 2, 5, 3, 4]}
6 messy = pd.DataFrame(messy, columns=["Box", "Dimension", "Value"])
7
8 tidy = messy.pivot_table(index = ["Box"],
9                           columns = "Dimension",
10                          values = "Value")
11 tidy["Volume"] = tidy.Length*tidy.Height*tidy.Width
12
13 print(" ")
14 print(tidy)
15
```

The IPython console shows the output of the code:

```
In [6]: runfile('C:/Users/Miguel Monoso/.spyder-py3/DataWrangling_2.py', wdir='C:/Users/Miguel Monoso/.spyder-py3')

Dimension  Height  Length  Width  Volume
Box
Box1       2      6      4      48
Box2       4      5      3      60
```

Permissions: RW End-of-lines: CRLF Encoding: UTF-8 Line: 14 Column: 12 Memory: 45 %