TASK#2

February 7, 2024

1 AICP Internship Task Week 2

[1]:
```
pip install pandas
```

```
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
Requirement already satisfied: pandas in /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages (2.1.4)
Requirement already satisfied: numpy<2,>=1.23.2 in /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages (from pandas) (1.24.3)
Requirement already satisfied: python-dateutil>=2.8.2 in /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages (from pandas) (2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages (from pandas) (2023.3)
Requirement already satisfied: six>=1.5 in /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages (from python-
dateutil>=2.8.2->pandas) (1.16.0)
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
Note: you may need to restart the kernel to use updated packages.
```

[1]:
```
import pandas as pd
```

1.0.1 Question#1: Writer pandas series code to get following output without using dictionary:

a 1

x 4

c 9

2 6

e 7

dtype: int 64

[2]:
```
values = [1, 4, 9, 6, 7]

index = ['a', 'x', 'c', '2', 'e']

series = pd.Series(values, index=index)

print(series)
```

```
a    1
x    4
c    9
2    6
e    7
dtype: int64
```

1.0.2 Question#2: Writer pandas series code to get following output using dictionary:

Bilal 42

Ayesha 38

Hadia 30

dtype: int64

[3]:
```
data = {'Bilal': 42, 'Ayesha': 38, 'Hadia': 30}

series = pd.Series(data)

print(series)
```

```
Bilal     42
Ayesha    38
Hadia     30
dtype: int64
```

1.0.3 Question#3: Write pandas dataframe code to get following output using python dictionary

[6]:
```
data = {'day': ['1/1/2017', '1//2/2017', '1/3/2017', '1/4/2017', '1/5/2017', '1/
↪6/2017'],
        'temperature': [32, 35, 28, 24, 32, 31],
        'windspeed': [6,7,2,7,4,2],
        'event': ['Rain', 'Sunny', 'Snow', 'Snow', 'Rain', 'Sunny']}

df = pd.DataFrame(data)

print(df)
```

```
         day  temperature  windspeed  event
0   1/1/2017           32          6   Rain
1  1//2/2017           35          7  Sunny
2   1/3/2017           28          2   Snow
3   1/4/2017           24          7   Snow
4   1/5/2017           32          4   Rain
5   1/6/2017           31          2  Sunny
```

1.0.4 Question#4: In extension to above question, you are required to replace index by [‘a’,‘b’,‘c’,‘d’,‘e’,‘f’]

[7]:
```
data = {'day': ['1/1/2017', '1//2/2017', '1/3/2017', '1/4/2017', '1/5/2017', '1/
↪6/2017'],
        'temperature': [32, 35, 28, 24, 32, 31],
        'windspeed': [6,7,2,7,4,2],
        'event': ['Rain', 'Sunny', 'Snow', 'Snow', 'Rain', 'Sunny']}

custom_index = ['a', 'b', 'c', 'd', 'e', 'f']

df = pd.DataFrame(data, index=custom_index)

print(df)
```

```
         day  temperature  windspeed  event
a   1/1/2017           32          6   Rain
b  1//2/2017           35          7  Sunny
c   1/3/2017           28          2   Snow
d   1/4/2017           24          7   Snow
e   1/5/2017           32          4   Rain
f   1/6/2017           31          2  Sunny
```

1.0.5 Question#5: In extension to above Q.3, calculate mean, miximum and minimum for label "temperature"

[8]:
```
import pandas as pd

data = {'day': ['1/1/2017', '1/2/2017', '1/3/2017', '1/4/2017', '1/5/2017', '1/
↪6/2017'],
        'temperature': [32, 35, 28, 24, 32, 31],
        'windspeed': [6, 7, 2, 7, 4, 2],
        'event': ['Rain', 'Sunny', 'Snow', 'Snow', 'Rain', 'Sunny']}
custom_index = ['a', 'b', 'c', 'd', 'e', 'f']

df = pd.DataFrame(data, index=custom_index)

temperature_mean = df['temperature'].mean()
temperature_max = df['temperature'].max()
temperature_min = df['temperature'].min()

print("Mean temperature:", temperature_mean)
print("Maximum temperature:", temperature_max)
print("Minimum temperature:", temperature_min)
```

```
Mean temperature: 30.333333333333332
Maximum temperature: 35
Minimum temperature: 24
```

1.0.6 Question#6: Import CSV 'people.csv' in the given folder. Keep in mind the following instructions:

You're required to import only specific columns ["First Name", "Sex", "Email", "Phone", "Job Title"]

Set the following columns ["Sex", "Job Title"] as index columns

Skip following rows [1,5]

Export the CSV as "NewPeople.csv"

[16]:
```
df = pd.read_csv('people.csv')
df.head()
```

[16]:

	Index	User Id	First Name	Last Name	Sex	\
0	1	88F7B33d2bcf9f5	Shelby	Terrell	Male	
1	2	f90cD3E76f1A9b9	Phillip	Summers	Female	
2	3	DbeAb8CcdfeFC2c	Kristine	Travis	Male	
3	4	A31Bee3c201ef58	Yesenia	Martinez	Male	
4	5	1bA7A3dc874da3c	Lori	Todd	Male	

	Email	Phone	Date of birth	\
0	elijah57@example.net	001-084-906-7849x73518	1945-10-26	

```
1       bethany14@example.com       214.112.6044x4913    1910-03-24
2       bthompson@example.com            277.609.7938    1992-07-02
3   kaitlinkaiser@example.com            584.094.6111    2017-08-03
4  buchananmanuel@example.net       689-207-3558x7233    1938-12-01

            Job Title
0     Games developer
1      Phytotherapist
2           Homeopath
3   Market researcher
4  Veterinary surgeon
```

[17]:
```
df = pd.read_csv('people.csv',
                 usecols=["First Name", "Sex", "Email", "Phone", "Job Title"],
                 skiprows=[1, 5],
                 index_col=["Sex", "Job Title"])

df.to_csv('SelectedPeople.csv')
```

[18]:
```
df.head()
```

[18]:
```
                                First Name                        Email  \
Sex    Job Title
Female Phytotherapist             Phillip      bethany14@example.com
Male   Homeopath                 Kristine      bthompson@example.com
       Market researcher          Yesenia  kaitlinkaiser@example.com
       Waste management officer       Erin        tconner@example.org
Female Intelligence analyst      Katherine    conniecowan@example.com

                                                    Phone
Sex    Job Title
Female Phytotherapist                  214.112.6044x4913
Male   Homeopath                            277.609.7938
       Market researcher                    584.094.6111
       Waste management officer  001-171-649-9856x5553
Female Intelligence analyst      +1-773-151-6685x49162
```

1.0.7 Question#7: Import excel sheet 'SampleWork.xlsx' in the given folder. Keep in mind the following instructions:

Import sheet 1

Import only first and last column from sheet 1

Skip row 2 while importing the sheet

Set row 2 as header

export as new sheet.

[21]:
```
pip install openpyxl
```

```
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
Collecting openpyxl
  Downloading openpyxl-3.1.2-py2.py3-none-any.whl (249 kB)

250.0/250.0 kB 290.2 kB/s eta
0:00:00[36m0:00:01[36m0:00:01:01
Collecting et-xmlfile (from openpyxl)
  Downloading et_xmlfile-1.1.0-py3-none-any.whl (4.7 kB)
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
Installing collected packages: et-xmlfile, openpyxl
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
Successfully installed et-xmlfile-1.1.0 openpyxl-3.1.2
WARNING: Skipping /home/malik-m-shahmeer-
rashid/miniconda3/lib/python3.11/site-packages/webencodings-0.5.1.dist-info due
to invalid metadata entry 'name'
Note: you may need to restart the kernel to use updated packages.
```

[2]:
```
import pandas as pd

df = pd.read_excel('SampleWork.xlsx', sheet_name=0, usecols=[0, -1],␣
 ↪skiprows=[1], header=1)

df.to_excel('NewSheet.xlsx', index=False)
```

1.0.8 Question#8: Create the following dataframe as AICP_DF then implement different operations as described below:

select 'Name', 'Qualification' coloumns and save to df1

add a new column to AICP_DF "Height" with the following values: [5.1, 6.2, 5.1, 5.2,5.1]

set column “Name” as the index column.

retrieve row with index “Hifza”

retrieve row with index 3

drop row with index “Bilal”

[5]:
```
import pandas as pd

data = {
    'Name': ['Ali', 'Bilal', 'Hifza', 'Ahmed', 'Ayesha'],
    'Qualification': ['MBA', 'BSc', 'PhD', 'MSc', 'MA']
}

AICP_DF = pd.DataFrame(data)
```

[6]:
```
df1 = AICP_DF[['Name', 'Qualification']]
```

[7]:
```
AICP_DF['Height'] = [5.1, 6.2, 5.1, 5.2, 5.1]
```

[8]:
```
AICP_DF.set_index('Name', inplace=True)
```

[9]:
```
row_hifza = AICP_DF.loc['Hifza']
```

[10]:
```
row_index_3 = AICP_DF.iloc[2]
```

[11]:
```
AICP_DF.drop('Bilal', inplace=True)
```