AN7914 Week 7 Python

March 22, 2023

1 Week 7 Python

1.1 Filtering/Conditional Selection

Let's create a toy dataframe.

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

people = {
    "first": ["Sakib", 'Jane', 'John', 'Harry'],
    "last": ["Anwar", 'Doe', 'Doe', None],
    "email": ["SakibAnwar@winchester.ac.uk", 'JaneDoe@email.com',␣
 ↪'JohnDoe@email.com', 'HarryPotter@email.com'],
    "age":[100,24,32,14],
    "degree":['Economics','Economics','Management', None],
    "role":['Programmer','Analyst','HR', np.nan],
}
df = pd.DataFrame(people)
df
```

[1]:

	first	last	email	age	degree	role
0	Sakib	Anwar	SakibAnwar@winchester.ac.uk	100	Economics	Programmer
1	Jane	Doe	JaneDoe@email.com	24	Economics	Analyst
2	John	Doe	JohnDoe@email.com	32	Management	HR
3	Harry	None	HarryPotter@email.com	14	None	NaN

[2]:
```
df[df['role'].isnull()]
```

[2]:

	first	last	email	age	degree	role
3	Harry	None	HarryPotter@email.com	14	None	NaN

[3]:
```
df[df['last'].notnull()]
```

[3]:

	first	last	email	age	degree	role
0	Sakib	Anwar	SakibAnwar@winchester.ac.uk	100	Economics	Programmer
1	Jane	Doe	JaneDoe@email.com	24	Economics	Analyst
2	John	Doe	JohnDoe@email.com	32	Management	HR

[4]:
```
df.dropna()
```

[4]:

	first	last	email	age	degree	role
0	Sakib	Anwar	SakibAnwar@winchester.ac.uk	100	Economics	Programmer
1	Jane	Doe	JaneDoe@email.com	24	Economics	Analyst
2	John	Doe	JohnDoe@email.com	32	Management	HR

You can use `df.dropna(inplace=True)` if you want to save this dataframe after dropping missing values.

[5]:
```
import pandas as pd
import numpy as np

# Create a list of 20 student names
students = ['Student {}'.format(i) for i in range(1, 21)]

# Create a dictionary with keys as subjects and values as random grades for␣
 ↪each student
grades = {'DataAnalytics': np.random.randint(10, 101, 20),
          'Finance': np.random.randint(10, 101, 20),
          'Management': np.random.randint(10, 101, 20)}

# Create a DataFrame with student names as index and grades for each subject as␣
 ↪columns
df = pd.DataFrame(data=grades, index=students)
```

[6]:
```
df
```

[6]:

	DataAnalytics	Finance	Management
Student 1	81	41	77
Student 2	31	12	85
Student 3	93	38	20
Student 4	40	34	85
Student 5	26	99	83
Student 6	70	59	99
Student 7	46	25	94
Student 8	15	11	18
Student 9	19	11	44
Student 10	40	18	27
Student 11	58	53	13
Student 12	27	39	43
Student 13	79	84	82
Student 14	43	58	65
Student 15	43	55	48
Student 16	66	22	91
Student 17	85	57	85
Student 18	54	67	42

Student 19	97	100	76
Student 20	38	25	60

Task 1. Filter the dataset for all the students who passed DataAnalytics. Pass mark is 50. 2. Filter the dataset for all the students who passed all. Pass mark is 50. 3. Filter the dataset for all the students who passed DataAnalytics or Finance. Pass mark is 50. 4. Filter the dataset for all the students who passed DataAnalytics and Finance. Pass mark is 50.

[7]:
```
cond1=(df['DataAnalytics']>=50)
df[cond1]
```

[7]:

	DataAnalytics	Finance	Management
Student 1	81	41	77
Student 3	93	38	20
Student 6	70	59	99
Student 11	58	53	13
Student 13	79	84	82
Student 16	66	22	91
Student 17	85	57	85
Student 18	54	67	42
Student 19	97	100	76

[8]:
```
df[( df['DataAnalytics']>=50) & (df['Finance']>=50) & (df['Management']>=50)]
```

[8]:

	DataAnalytics	Finance	Management
Student 6	70	59	99
Student 13	79	84	82
Student 17	85	57	85
Student 19	97	100	76

1.2 Assigning

Assigning data

[9]:
```
df['Degree']='MSc_IFM'
```

[10]:
```
df
```

[10]:

	DataAnalytics	Finance	Management	Degree
Student 1	81	41	77	MSc_IFM
Student 2	31	12	85	MSc_IFM
Student 3	93	38	20	MSc_IFM
Student 4	40	34	85	MSc_IFM
Student 5	26	99	83	MSc_IFM
Student 6	70	59	99	MSc_IFM
Student 7	46	25	94	MSc_IFM
Student 8	15	11	18	MSc_IFM
Student 9	19	11	44	MSc_IFM

```
Student 10          40       18          27  MSc_IFM
Student 11          58       53          13  MSc_IFM
Student 12          27       39          43  MSc_IFM
Student 13          79       84          82  MSc_IFM
Student 14          43       58          65  MSc_IFM
Student 15          43       55          48  MSc_IFM
Student 16          66       22          91  MSc_IFM
Student 17          85       57          85  MSc_IFM
Student 18          54       67          42  MSc_IFM
Student 19          97      100          76  MSc_IFM
Student 20          38       25          60  MSc_IFM
```

[ ]: