Assignment#2, MSCS3045-Fall22,

Chapter 5

Objective: Learn more about pandas, the students need to get comfortable with its two workhorse data structures: Series and DataFrame.

Q1:

(10 points) Why Index objects in pandas are immutable and thus can't be modified by the user? Explain (5 points) with example (5 points)

Answer:

Q2: (10 points) Study the example given in page 163-164 (ch5). Explain the retrieved result after using .corr (5points) and .cov (points) functions.

Answer:

Q3: (10 points) What is the difference between ***corr*** and ***corrwith*** functions (5 points), explain with example (5 points).

Answer:

Q4: (10 points) Assume you have the following datafram, give the code to do the following:

	one	two	three	four
Ohio	0	1	2	3
Colorado	4	5	6	7
Utah	8	9	10	11
New York	12	13	14	15

. Assume also, your dataframe is defined as:

```
data = pd.DataFrame(np.arange(16).reshape((4, 4)),
                    index=['Ohio', 'Colorado', 'Utah', 'New York'],
                    columns=['one', 'two', 'three', 'four'])
```

1. Give the code to delete: 'Colorado' and 'Ohio' from your dataframe (5 points):

Answer:

2. Give the code to drop values from the column two (5 points):

Answer:

Q5: (10 points) Study the output from the following code:

sdata = {'Ohio': 35000, 'Texas': 71000, 'Oregon': 16000, 'Utah': 5000}

states = ['California', 'Ohio', 'Oregon', 'Texas']

obj1 = pd.Series(sdata, index=states)

obj1=

```
California        NaN
Ohio          35000.0
Oregon        16000.0
Texas         71000.0
dtype: float64
```

Now answer the following questions:

1. What is the meaning of “NaN”? (5point)

Answer:

2. Why “NaN” are there in obj1 ? (5points)

Answer: