Data In Motion Week 9 Pandas Challenge

Week 9 Task explanation.

Create the data

raw_data_1 = { 'subject_id': ['1', '2', '3', '4', '5'], 'first_name': ['Alex', 'Amy', 'Allen', 'Alice', 'Ayoung'], 'last_name': ['Anderson', 'Ackerman', 'Ali', 'Aoni', 'Atiches']}

raw_data_2 = { 'subject_id': ['4', '5', '6', '7', '8'], 'first_name': ['Billy', 'Brian', 'Bran', 'Bryce', 'Betty'], 'last_name': ['Bonder', 'Black', 'Balwner', 'Brice', 'Btisan']}

raw_data_3 = { 'subject_id': ['1', '2', '3', '4', '5', '7', '8', '9', '10', '11'], 'test_id': [51, 15, 15, 61, 16, 14, 15, 1, 61, 16]}

Challenge Tasks

Task 1: Assign each to a variable called data1, data2, data3.

In [1]:

```
data1 = { 'subject_id': ['1', '2', '3', '4', '5'],
         'first_name': ['Alex', 'Amy', 'Allen', 'Alice', 'Ayoung'],
         'last_name': ['Anderson', 'Ackerman', 'Ali', 'Aoni', 'Atiches']}
```

In [2]:

```
data1
```

Out[2]:

```
{'subject_id': ['1', '2', '3', '4', '5'],
 'first_name': ['Alex', 'Amy', 'Allen', 'Alice', 'Ayoung'],
 'last_name': ['Anderson', 'Ackerman', 'Ali', 'Aoni', 'Atiches']}
```

In [3]:

```
data2 = { 'subject_id': ['4', '5', '6', '7', '8'],
         'first_name': ['Billy', 'Brian', 'Bran', 'Bryce', 'Betty'],
         'last_name': ['Bonder', 'Black', 'Balwner', 'Brice', 'Btisan']}
```

In [4]:

```
data2
```

Out[4]:

```
{'subject_id': ['4', '5', '6', '7', '8'],
 'first_name': ['Billy', 'Brian', 'Bran', 'Bryce', 'Betty'],
 'last_name': ['Bonder', 'Black', 'Balwner', 'Brice', 'Btisan']}
```

In [5]:

```
data3 = { 'subject_id': ['1', '2', '3', '4', '5', '7', '8', '9', '10', '11'],
         'test_id': [51, 15, 15, 61, 16, 14, 15, 1, 61, 16]}
```

In [6]:

```
data3
```

Out[6]:

```
{'subject_id': ['1', '2', '3', '4', '5', '7', '8', '9', '10', '11'],
 'test_id': [51, 15, 15, 61, 16, 14, 15, 1, 61, 16]}
```

In [7]:

```
import pandas as pd
```

In [8]:

```
#Converting data1 from dictionary to a dataframe
data1 = pd.DataFrame.from_dict(data1)
```

In [9]:

```
data1
```

Out[9]:

	subject_id	first_name	last_name
0	1	Alex	Anderson
1	2	Amy	Ackerman
2	3	Allen	Ali
3	4	Alice	Aoni
4	5	Ayoung	Atiches

In [10]:

```
# Converting data2 and data3 from dictionary to a dataframe
data2 = pd.DataFrame.from_dict(data2)
data3 = pd.DataFrame.from_dict(data3)
```

Task 2: Join data1 and data2 along rows and assign all_data.

In [11]:

```
all_data = pd.concat([data1, data2], axis=0)
```

In [12]:

```
all_data
```

Out[12]:

	subject_id	first_name	last_name
0	1	Alex	Anderson
1	2	Amy	Ackerman
2	3	Allen	Ali
3	4	Alice	Aoni
4	5	Ayoung	Atiches
0	4	Billy	Bonder
1	5	Brian	Black
2	6	Bran	Balwner
3	7	Bryce	Brice
4	8	Betty	Btisan

Task 3: Join the two dataframes along columns and assign to all_data_col.

In [13]:

```
all_data_col = pd.concat([data1, data2], axis=1)
```

In [14]:

```
all_data_col
```

Out[14]:

	subject_id	first_name	last_name	subject_id	first_name	last_name
0	1	Alex	Anderson	4	Billy	Bonder
1	2	Amy	Ackerman	5	Brian	Black
2	3	Allen	Ali	6	Bran	Balwner
3	4	Alice	Aoni	7	Bryce	Brice
4	5	Ayoung	Atiches	8	Betty	Btisan

Task 4: Print data3.

In [15]:

```
data3
```

Out[15]:

	subject_id	test_id
0	1	51
1	2	15
2	3	15
3	4	61
4	5	16
5	7	14
6	8	15
7	9	1
8	10	61
9	11	16

Task 5: Merge all_data and data3 along the subject_id value.

In [16]:

```
pd.merge(all_data, data3, on='subject_id')
```

Out[16]:

	subject_id	first_name	last_name	test_id
0	1	Alex	Anderson	51
1	2	Amy	Ackerman	15
2	3	Allen	Ali	15
3	4	Alice	Aoni	61
4	4	Billy	Bonder	61
5	5	Ayoung	Atiches	16
6	5	Brian	Black	16
7	7	Bryce	Brice	14
8	8	Betty	Btisan	15

Task 6: Merge only the data that has the same 'subject_id' on both data1 and data2.

In [17]:

```
pd.merge(data1, data2, on='subject_id')
```

Out[17]:

	subject_id	first_name_x	last_name_x	first_name_y	last_name_y
0	4	Alice	Aoni	Billy	Bonder
1	5	Ayoung	Atiches	Brian	Black

Task 7: Merge all values in data1 and data2, with matching records from both sides where available.

In [18]:

```
pd.merge(data1, data2, on='subject_id', how = 'outer')
```

Out[18]:

	subject_id	first_name_x	last_name_x	first_name_y	last_name_y
0	1	Alex	Anderson	NaN	NaN
1	2	Amy	Ackerman	NaN	NaN
2	3	Allen	Ali	NaN	NaN
3	4	Alice	Aoni	Billy	Bonder
4	5	Ayoung	Atiches	Brian	Black
5	6	NaN	NaN	Bran	Balwner
6	7	NaN	NaN	Bryce	Brice
7	8	NaN	NaN	Betty	Btisan