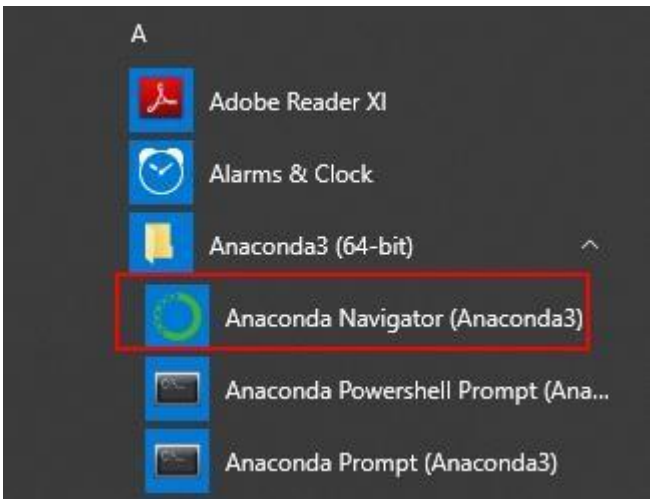
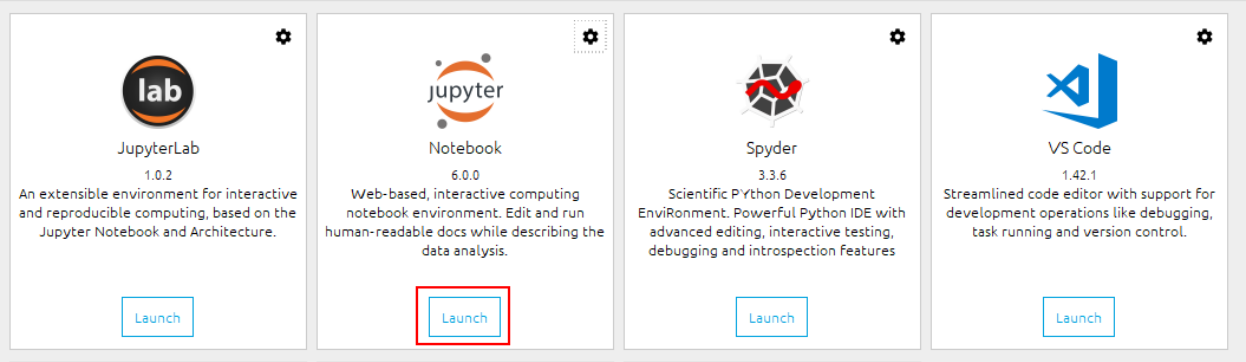
Module 7: Hands-On: 3

Merge Join and Concat.

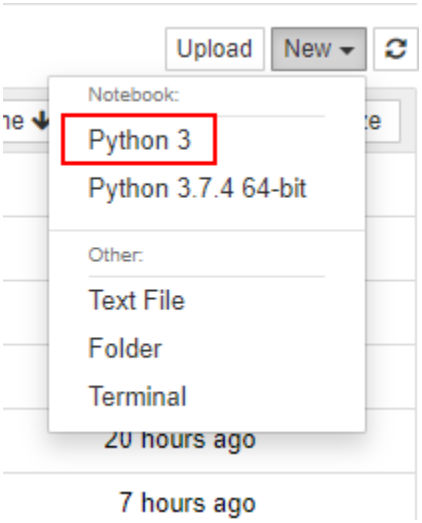
Step 1: Open Anaconda Navigator



Step 2: Click on Launch button under jupyter notebooks.



Step 3: After the notebook opens click on new and Python 3.



Step 4: Import the required packages and create two data frames.

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

```
In [2]: user = pd.DataFrame({
            'id': [1, 2, 3, 4],
            'name': ['a', 'b', 'c', 'd']
        })
```

```
In [3]: qualifications = pd.DataFrame({
            'user_id': [1, 2, 3, 4],
            'qualification': ['e', 'f', 'g', 'h']
        })
```

Step 5: Merge the two dataframes.

```
In [4]: pd.merge(user, qualifications, left_on='id', right_on='user_id')
```

Out[4]:

	id	name	user_id	qualification
0	1	a	1	e
1	2	b	2	f
2	3	c	3	g
3	4	d	4	h

Step 6: Create two data frames.

```
In [5]: df1 = pd.DataFrame({
            'id': [1, 2, 3, 4],
            'name': ['a', 'b', 'c', 'd']
        }).set_index('id')
```

```
In [6]: df2 = pd.DataFrame({
            'id': [1, 2, 3, 4],
            'qualification': ['a', 'b', 'c', 'd']
        }).set_index('id')
```

Step 7: Join two dataframes.

```
In [9]: df1.join(df2)
```

Out[9]:

id	name	qualification
1	a	a
2	b	b
3	c	c
4	d	d

```
In [10]: df2.join(df1)
```

Out[10]:

id	qualification	name
1	a	a
2	b	b
3	c	c
4	d	d

Step 8: Create 3 data frames.

```
In [11]: df1 = pd.DataFrame({
             'id': [0, 1, 2],
             'name': ['a', 'b', 'c']
         }).set_index('id')

         df2 = pd.DataFrame({
             'id': [3, 4, 5],
             'name': ['d', 'e', 'f']
         }).set_index('id')

         df3 = pd.DataFrame({
             'id': [6, 7, 8],
             'name': ['g', 'h', 'i']
         }).set_index('id')
```

Step 9: Concatenate 3 data frames by them as a list in pd.concat.

```
In [15]: pd.concat([df1, df2, df3])
```

Out[15]:

id	name
0	a
1	b
2	c
3	d
4	e
5	f
6	g
7	h
8	i