LAB Logbook

Lab 1

```
sid=np.arange(43)
print(sid)
```

```
[ 0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42]
```

```
sid=sid.reshape(1,43)
sid
```

```
array([[ 0,  1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15,
        16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31,
        32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42]])
```

```
b=sid
print(b)
```

```
[[ 0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
  24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42]]
```

```
print(b.shape)
```

```
(1, 43)
```

Lab 2

[57]:
```
Group_by_relationship=data.groupby(["relationship","hours-per-week"])
Group_by_relationship.size()
```

[57]:
```
relationship    hours-per-week
 Husband        13.0             1
                40.0             2
                45.0             1
                80.0             1
 Not-in-family  16.0             1
                40.0             2
                50.0             2
 Own-child      30.0             1
 Wife           40.0             2
dtype: int64
```

[61]:
```
def reduction(x):
    return x-3
data["hours-per-week"]=data["hours-per-week"].apply(reduction)
data
```

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	per-week	native-country	Answer	IsHomeless
0	39	State-gov	77516.0	Bachelors	13.0	Never-married	Adm-clerical	Not-in-family	White	Male	2174.0	NaN	37.0	United-States	<=50K	False
1	50	Self-emp-not-inc	83311.0	Bachelors	13.0	Married-civ-spouse	Exec-managerial	Husband	White	Male	0.0	0.0	10.0	United-States	<=50K	False
2	38	Private	215646.0	HS-grad	9.0	Divorced	Handlers-cleaners	Not-in-family	White	Male	0.0	NaN	37.0	United-States	<=50K	False
3	53	Private	234721.0	11th	7.0	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0.0	NaN	37.0	United-States	<=50K	False
4	28	Private	338409.0	Bachelors	13.0	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0.0	NaN	37.0	Cuba	<=50K	False
5	37	Private	284582.0	Masters	14.0	Married-civ-spouse	Exec-managerial	Wife	White	Female	0.0	NaN	37.0	United-States	<=50K	False
6	49	Private	160187.0	9th	5.0	Married-spouse-absent	Other-service	Not-in-family	Black	Female	0.0	0.0	13.0	Jamaica	<=50K	False
7	52	Self-emp-not-inc	209642.0	HS-grad	9.0	Married-civ-spouse	Exec-managerial	Husband	White	Male	0.0	0.0	42.0	United-States	>50K	False
8	31	Private	45781.0	Masters	14.0	Never-married	Prof-specialty	Not-in-family	White	Female	14084.0	NaN	47.0	United-States	>50K	False

8	31	Private	45781.0	Masters	14.0	Never-married	Prof-specialty	Not-in-family	White	Female	14084.0	NaN	47.0	United-States	>50K
10	37	Private	280464.0	Some-college	10.0	Married-civ-spouse	Exec-managerial	Husband	Black	Male	0.0	NaN	77.0	United-States	>50K
12	23	Private	122272.0	Bachelors	13.0	Never-married	Adm-clerical	Own-child	White	Female	0.0	NaN	27.0	United-States	<=50K
13	32	Private	205019.0	Assoc-acdm	12.0	Never-married	Sales	Not-in-family	Black	Male	0.0	NaN	47.0	United-States	<=50K
14	40	Private	121772.0	Assoc-voc	11.0	Married-civ-spouse	Craft-repair	Husband	Asian-Pac-Islander	Male	0.0	NaN	37.0	?	>50K
15	25	Private	NaN	Some-college	NaN	NaN	NaN	NaN	White	Male	0.0	NaN	NaN	NaN	NaN

[62]:
```
Group_by_relationship=data.groupby(["relationship","hours-per-week"])
Group_by_relationship.size()
```

[62]:
```
relationship    hours-per-week
 Husband         10.0             1
                 37.0             2
                 42.0             1
                 77.0             1
 Not-in-family   13.0             1
                 37.0             2
                 47.0             2
 Own-child       27.0             1
 Wife            37.0             2
dtype: int64
```


Lab 3

Lab 4

Lab 5

Lab 6

Lab 7

Lab 8

Lab 9

Lab 10

Lab 11

Lab 12