

27__assignment

March 14, 2022

```
[ ]: # Assignment
# 1. Create a new dataframe. 20 rows, 5 Series [columns [A-E]], containing
↳ random numerical data
# 2. Extend a new Series [column [mean]] with mean of all Series [columns] in
↳ the row
# 3. Take a subset of the newly created Series [column[mean]] into another
↳ variable.

# Importing required libraries
import pandas as pd
import numpy as np

# generating indexing array
indexing_arr = np.arange(1, 21)

# Generating a sample data frame
data_frame = pd.DataFrame(data=np.arange(1, 101).reshape(20,5),
↳ index=indexing_arr, columns=list('ABCDE'))
# data_frame
```

```
[ ]: # Add a new column mean (First simply add a column), then replace it with
data_frame["mean"] = indexing_arr
data_frame
```

```
[ ]:
```

	A	B	C	D	E	mean
1	1	2	3	4	5	1
2	6	7	8	9	10	2
3	11	12	13	14	15	3
4	16	17	18	19	20	4
5	21	22	23	24	25	5
6	26	27	28	29	30	6
7	31	32	33	34	35	7
8	36	37	38	39	40	8
9	41	42	43	44	45	9
10	46	47	48	49	50	10
11	51	52	53	54	55	11
12	56	57	58	59	60	12

13	61	62	63	64	65	13
14	66	67	68	69	70	14
15	71	72	73	74	75	15
16	76	77	78	79	80	16
17	81	82	83	84	85	17
18	86	87	88	89	90	18
19	91	92	93	94	95	19
20	96	97	98	99	100	20

```
[ ]: # calculating mean of first row
np.mean(data_frame.loc[1, "A":"E"])
```

```
[ ]: 3.0
```

```
[ ]: # setting value in the mean Series [column]
data_frame.loc[1, "mean"] = np.mean(data_frame.loc[1, "A":"E"])
data_frame
```

```
[ ]:
```

	A	B	C	D	E	mean
1	1	2	3	4	5	3
2	6	7	8	9	10	2
3	11	12	13	14	15	3
4	16	17	18	19	20	4
5	21	22	23	24	25	5
6	26	27	28	29	30	6
7	31	32	33	34	35	7
8	36	37	38	39	40	8
9	41	42	43	44	45	9
10	46	47	48	49	50	10
11	51	52	53	54	55	11
12	56	57	58	59	60	12
13	61	62	63	64	65	13
14	66	67	68	69	70	14
15	71	72	73	74	75	15
16	76	77	78	79	80	16
17	81	82	83	84	85	17
18	86	87	88	89	90	18
19	91	92	93	94	95	19
20	96	97	98	99	100	20

```
[ ]: # looping through indices and changing value to mean
for x in indexing_arr:
    data_frame.loc[x, "mean"] = np.mean(data_frame.loc[x, "A":"E"])

data_frame
```

```
[ ]:
```

	A	B	C	D	E	mean
1	1	2	3	4	5	3
2	6	7	8	9	10	8
3	11	12	13	14	15	13
4	16	17	18	19	20	18
5	21	22	23	24	25	23
6	26	27	28	29	30	28
7	31	32	33	34	35	33
8	36	37	38	39	40	38
9	41	42	43	44	45	43
10	46	47	48	49	50	48
11	51	52	53	54	55	53
12	56	57	58	59	60	58
13	61	62	63	64	65	63
14	66	67	68	69	70	68
15	71	72	73	74	75	73
16	76	77	78	79	80	78
17	81	82	83	84	85	83
18	86	87	88	89	90	88
19	91	92	93	94	95	93
20	96	97	98	99	100	98

```
[ ]: # storing mean subset into a different variable
mean_series = data_frame.loc[0:, "mean"]
mean_series
```

```
[ ]:
```

1	3
2	8
3	13
4	18
5	23
6	28
7	33
8	38
9	43
10	48
11	53
12	58
13	63
14	68
15	73
16	78
17	83
18	88
19	93
20	98

Name: mean, dtype: int32

```
[ ]: # Assignment Completed
```