

6_Central Limit Theorem

```
In [22]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [23]: # Generate random data by using List comprehension

pop_data = [np.random.randint(10,100) for i in range(10000)]
pop_data
```

```
Out[23]: [85,  
56,  
80,  
74,  
62,  
54,  
20,  
70,  
51,  
13,  
80,  
15,  
92,  
23,  
34,  
67,  
98,  
64,  
48,  
97,  
10,  
72,  
47,  
48,  
80,  
56,  
27,  
28,  
85,  
39,  
11,  
97,  
21,  
28,  
64,  
65,  
80,  
70,  
16,  
78,  
41,  
29,  
60,  
23,  
82,  
93,  
51,  
59,  
60,  
60,  
43,  
12,  
35,  
49,  
10,  
19,
```

37,
52,
42,
68,
22,
61,
30,
23,
52,
98,
40,
19,
49,
38,
48,
67,
71,
22,
41,
30,
78,
14,
27,
78,
97,
29,
59,
82,
73,
97,
87,
45,
93,
46,
21,
63,
59,
79,
27,
36,
49,
23,
91,
96,
95,
79,
43,
89,
26,
75,
73,
15,
61,
72,
99,
64,

38,
20,
72,
11,
85,
67,
88,
86,
70,
35,
22,
38,
68,
48,
69,
20,
12,
48,
54,
30,
73,
79,
73,
31,
46,
40,
58,
86,
36,
12,
12,
52,
50,
28,
32,
68,
74,
88,
80,
51,
55,
42,
35,
81,
78,
49,
10,
16,
31,
22,
41,
87,
58,
63,
66,
80,

82,
55,
47,
58,
93,
12,
82,
52,
51,
21,
58,
45,
87,
60,
49,
81,
82,
21,
15,
49,
67,
61,
16,
69,
85,
40,
12,
57,
98,
13,
46,
14,
62,
13,
83,
31,
75,
64,
82,
35,
51,
63,
47,
90,
90,
14,
41,
10,
27,
93,
20,
10,
67,
46,
40,
67,

42,
90,
86,
81,
22,
95,
65,
42,
74,
89,
81,
48,
80,
17,
98,
94,
91,
14,
68,
17,
63,
34,
27,
32,
91,
40,
76,
81,
52,
61,
94,
46,
23,
87,
18,
87,
60,
93,
11,
39,
27,
30,
30,
22,
77,
47,
88,
17,
33,
27,
57,
33,
20,
10,
19,
40,

15,
95,
72,
85,
36,
34,
64,
71,
29,
46,
27,
43,
58,
69,
85,
87,
54,
96,
89,
81,
64,
64,
30,
93,
60,
17,
51,
79,
39,
87,
72,
67,
30,
84,
53,
59,
54,
25,
75,
71,
82,
21,
12,
17,
29,
46,
33,
73,
99,
22,
54,
16,
39,
58,
46,
69,

17,
32,
25,
44,
96,
75,
97,
50,
11,
49,
48,
25,
40,
21,
54,
78,
61,
58,
12,
67,
62,
86,
49,
74,
18,
35,
56,
62,
68,
15,
40,
29,
95,
39,
72,
98,
69,
66,
56,
41,
90,
84,
26,
63,
33,
16,
72,
22,
99,
40,
81,
64,
37,
32,
31,
38,

40,
47,
16,
75,
17,
60,
68,
90,
85,
25,
25,
83,
78,
33,
14,
23,
51,
54,
56,
11,
47,
90,
17,
43,
33,
87,
13,
65,
84,
54,
29,
20,
89,
45,
81,
65,
34,
34,
80,
64,
69,
79,
36,
15,
49,
56,
38,
81,
61,
27,
51,
89,
34,
84,
23,
16,

46,
59,
51,
27,
17,
18,
29,
81,
67,
81,
83,
73,
46,
14,
82,
31,
80,
74,
24,
93,
12,
98,
93,
53,
35,
50,
15,
71,
61,
33,
54,
71,
67,
29,
39,
44,
61,
88,
30,
53,
48,
32,
92,
22,
65,
67,
42,
66,
73,
45,
99,
57,
73,
29,
28,
54,

67,
47,
14,
10,
70,
16,
32,
93,
12,
53,
52,
63,
13,
91,
65,
52,
94,
28,
36,
33,
49,
83,
48,
15,
93,
25,
55,
24,
32,
67,
95,
61,
53,
45,
43,
47,
38,
28,
37,
98,
21,
92,
84,
74,
58,
69,
90,
68,
12,
98,
34,
21,
91,
89,
36,
23,

98,
73,
93,
86,
22,
57,
34,
36,
81,
17,
58,
47,
96,
12,
38,
95,
63,
65,
41,
51,
71,
65,
23,
49,
65,
43,
75,
91,
38,
10,
83,
11,
21,
46,
99,
75,
66,
52,
88,
82,
16,
20,
56,
92,
32,
87,
37,
32,
68,
39,
16,
32,
90,
72,
81,
56,

10,
77,
70,
62,
64,
61,
29,
22,
16,
40,
24,
40,
71,
49,
21,
93,
60,
57,
12,
58,
59,
73,
64,
25,
70,
37,
70,
29,
93,
68,
32,
11,
85,
79,
51,
92,
62,
78,
54,
55,
67,
49,
66,
18,
56,
25,
66,
44,
38,
68,
92,
61,
95,
21,
86,
22,

94,
16,
85,
34,
19,
21,
22,
22,
78,
89,
18,
57,
86,
44,
75,
31,
99,
13,
80,
47,
26,
22,
89,
69,
52,
94,
86,
80,
18,
25,
28,
32,
52,
14,
37,
40,
57,
33,
53,
73,
25,
77,
77,
53,
76,
68,
10,
25,
11,
41,
65,
21,
43,
92,
62,
79,

95,
29,
88,
79,
78,
66,
82,
95,
91,
17,
51,
65,
54,
23,
10,
23,
90,
48,
18,
44,
47,
63,
74,
84,
58,
77,
62,
90,
39,
41,
54,
91,
97,
79,
42,
66,
35,
25,
32,
94,
18,
32,
28,
63,
95,
24,
20,
57,
73,
89,
86,
69,
86,
35,
39,
96,

82,
92,
75,
91,
38,
83,
84,
27,
25,
62,
12,
40,
90,
97,
59,
47,
30,
66,
87,
10,
12,
78,
77,
38,
54,
60,
24,
29,
12,
69,
48,
51,
39,
97,
68,
59,
98,
46,
15,
50,
60,
89,
48,
81,
79,
54,
15,
57,
96,
35,
25,
14,
38,
68,
82,
50,

58,
19,
67,
34,
18,
85,
26,
35,
59,
90,
60,
68,
77,
80,
78,
76,
95,
92,
39,
24,
50,
33,
32,
64,
12,
40,
15,
79,
84,
43,
51,
28,
63,
49,
42,
29,
22,
36,
67,
57,
86,
61,
61,
39,
81,
49,
91,
22,
35,
72,
81,
31,
17,
57,
81,
80,

73,
55,
32,
81,
76,
71,
60,
11,
88,
65,
77,
64,
28,
11,
75,
23,
71,
30,
48,
56,
26,
82,
45,
44,
65,
39,
45,
77,
69,
58,
35,
92,
59,
45,
78,
25,
95,
71,
75,
26,
13,
63,
10,
81,
77,
27,
63,
92,
23,
37,
84,
48,
57,
67,
82,
57,

```
83,  
75,  
22,  
79,  
24,  
36,  
97,  
83,  
18,  
51,  
59,  
19,  
94,  
99,  
15,  
39,  
97,  
52,  
55,  
97,  
94,  
57,  
66,  
76,  
95,  
91,  
93,  
39,  
48,  
12,  
64,  
80,  
28,  
28,  
49,  
80,  
32,  
42,  
88,  
88,  
86,  
87,  
51,  
19,  
85,  
72,  
62,  
84,  
...]
```

```
In [24]: # the above line of code could be written as:  
pop_data = []  
for i in range(10000):  
    pop_data.append(np.random.randint(10,100))  
pop_data
```

```
Out[24]: [71,  
45,  
29,  
38,  
64,  
82,  
64,  
35,  
74,  
76,  
63,  
26,  
10,  
87,  
15,  
57,  
54,  
38,  
52,  
97,  
32,  
87,  
49,  
93,  
39,  
83,  
33,  
11,  
20,  
85,  
94,  
79,  
87,  
76,  
53,  
89,  
25,  
93,  
33,  
29,  
12,  
92,  
91,  
73,  
28,  
21,  
81,  
15,  
24,  
67,  
61,  
65,  
13,  
46,  
92,  
46,
```

35,
10,
50,
55,
16,
50,
93,
19,
98,
83,
60,
84,
16,
59,
98,
38,
65,
78,
27,
43,
80,
76,
67,
70,
91,
57,
50,
42,
88,
95,
35,
93,
25,
65,
32,
39,
79,
37,
20,
58,
71,
61,
75,
96,
60,
39,
98,
79,
32,
68,
12,
24,
47,
54,
38,
97,

40,
40,
52,
84,
26,
61,
96,
92,
11,
32,
62,
31,
76,
35,
43,
33,
89,
94,
68,
11,
85,
79,
53,
53,
56,
67,
20,
93,
50,
83,
25,
66,
43,
69,
43,
70,
39,
42,
95,
30,
52,
46,
48,
86,
68,
30,
94,
27,
47,
33,
78,
50,
27,
15,
66,
38,

65,
72,
78,
39,
93,
90,
86,
19,
13,
76,
19,
56,
53,
82,
54,
77,
81,
16,
56,
92,
75,
39,
45,
25,
28,
39,
51,
91,
32,
85,
43,
17,
41,
19,
53,
24,
50,
14,
12,
92,
51,
85,
33,
39,
54,
66,
58,
24,
19,
29,
25,
89,
60,
57,
33,
76,

61,
49,
80,
69,
18,
16,
93,
19,
20,
57,
66,
38,
13,
71,
49,
32,
84,
88,
92,
84,
87,
11,
14,
83,
66,
71,
46,
59,
44,
16,
40,
13,
42,
48,
17,
14,
11,
86,
28,
11,
62,
87,
33,
99,
61,
40,
23,
77,
77,
80,
25,
50,
16,
75,
87,
37,

36,
97,
75,
39,
47,
32,
87,
58,
93,
21,
66,
34,
67,
32,
33,
84,
93,
61,
19,
95,
34,
80,
20,
36,
12,
31,
60,
81,
19,
97,
56,
11,
10,
70,
39,
30,
32,
33,
80,
85,
52,
27,
56,
19,
74,
21,
68,
45,
75,
55,
66,
16,
84,
92,
71,
59,

93,
78,
27,
16,
76,
18,
29,
52,
25,
63,
52,
80,
57,
82,
33,
85,
67,
60,
92,
43,
42,
93,
27,
42,
21,
51,
20,
74,
32,
37,
89,
56,
92,
98,
59,
77,
85,
22,
32,
15,
86,
94,
78,
80,
92,
91,
97,
18,
38,
86,
53,
68,
41,
82,
63,
60,

52,
98,
12,
43,
19,
41,
49,
48,
17,
99,
95,
43,
54,
99,
78,
81,
20,
66,
72,
45,
46,
34,
51,
96,
90,
98,
10,
57,
52,
64,
35,
76,
43,
42,
59,
54,
45,
44,
20,
34,
90,
22,
50,
36,
27,
54,
70,
55,
41,
23,
42,
19,
49,
19,
59,
68,

87,
99,
78,
72,
91,
74,
63,
35,
24,
24,
23,
36,
73,
50,
31,
89,
85,
67,
62,
50,
99,
21,
65,
23,
96,
53,
35,
22,
47,
26,
40,
57,
93,
24,
73,
68,
63,
86,
60,
50,
73,
51,
56,
29,
59,
64,
24,
57,
12,
84,
38,
89,
33,
30,
78,
78,

41,
23,
40,
89,
90,
17,
42,
61,
13,
77,
41,
29,
25,
75,
50,
35,
36,
92,
92,
88,
57,
12,
76,
40,
80,
50,
24,
87,
36,
80,
98,
20,
13,
82,
18,
90,
43,
57,
92,
28,
37,
89,
56,
26,
18,
84,
55,
34,
60,
88,
68,
47,
20,
67,
27,
78,

95,
76,
53,
76,
50,
88,
44,
97,
54,
25,
24,
21,
86,
77,
25,
23,
92,
24,
14,
96,
63,
13,
46,
95,
16,
25,
71,
61,
65,
13,
58,
92,
93,
88,
74,
29,
84,
82,
88,
38,
52,
53,
56,
78,
14,
53,
39,
42,
66,
14,
34,
72,
12,
32,
17,
95,

84,
24,
55,
71,
60,
84,
85,
85,
44,
14,
39,
38,
76,
83,
16,
85,
78,
56,
98,
65,
68,
42,
15,
57,
64,
23,
39,
78,
73,
84,
90,
63,
64,
65,
58,
95,
91,
33,
83,
14,
23,
63,
71,
14,
44,
14,
86,
75,
28,
97,
77,
60,
27,
39,
68,
67,

22,
40,
73,
89,
90,
20,
90,
17,
90,
69,
63,
15,
35,
16,
78,
68,
36,
53,
70,
71,
64,
72,
74,
73,
97,
78,
45,
27,
69,
21,
63,
50,
27,
15,
94,
49,
22,
15,
55,
42,
25,
10,
74,
76,
19,
65,
91,
61,
32,
15,
69,
46,
35,
14,
18,
14,

55,
85,
67,
88,
61,
30,
64,
45,
74,
16,
96,
97,
78,
90,
71,
22,
82,
26,
15,
69,
61,
31,
96,
97,
72,
78,
91,
83,
18,
14,
56,
98,
55,
84,
44,
37,
84,
78,
55,
13,
73,
97,
68,
52,
99,
83,
86,
23,
90,
29,
48,
89,
22,
84,
65,
43,

83,
48,
11,
71,
53,
48,
99,
73,
24,
93,
75,
50,
36,
66,
35,
65,
82,
84,
65,
27,
53,
63,
96,
70,
55,
34,
75,
93,
47,
12,
64,
94,
28,
23,
83,
51,
22,
23,
28,
99,
28,
92,
48,
63,
79,
24,
66,
65,
66,
57,
78,
10,
43,
70,
39,
42,

39,
27,
57,
37,
66,
37,
90,
77,
80,
10,
99,
36,
79,
44,
88,
16,
99,
39,
37,
52,
98,
51,
31,
72,
38,
31,
10,
28,
86,
13,
38,
36,
61,
52,
26,
96,
93,
31,
44,
22,
73,
87,
17,
13,
72,
82,
18,
77,
31,
24,
33,
98,
97,
22,
78,
46,

16,
65,
58,
46,
83,
14,
23,
87,
23,
94,
34,
86,
45,
80,
26,
40,
41,
55,
67,
87,
14,
83,
50,
48,
11,
43,
70,
54,
20,
54,
26,
44,
38,
77,
55,
95,
95,
86,
82,
93,
16,
51,
56,
33,
76,
78,
87,
10,
23,
13,
66,
75,
16,
53,
68,
98,

```
95,  
60,  
18,  
77,  
85,  
62,  
39,  
71,  
11,  
78,  
54,  
31,  
69,  
64,  
76,  
70,  
28,  
79,  
86,  
78,  
12,  
79,  
87,  
38,  
36,  
81,  
18,  
17,  
78,  
17,  
46,  
20,  
57,  
98,  
48,  
66,  
21,  
25,  
84,  
25,  
18,  
84,  
57,  
63,  
17,  
24,  
76,  
70,  
...]
```

```
In [25]: len(pop_data)
```

```
Out[25]: 10000
```

```
In [26]: # TO convert population data into a csv file  
pop_table = pd.DataFrame({'pop_data':pop_data})
```

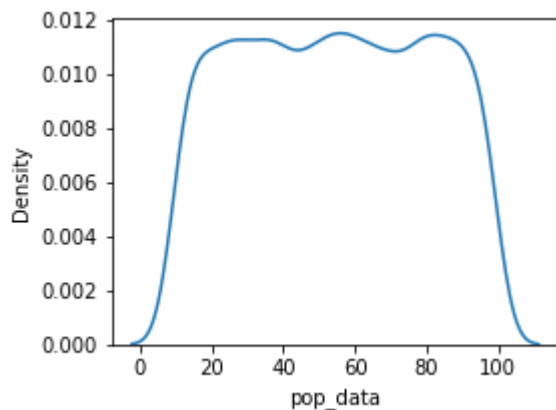
```
pop_table
```

Out[26]: **pop_data**

0	71
1	45
2	29
3	38
4	64
...	...
9995	55
9996	20
9997	43
9998	74
9999	66

10000 rows × 1 columns

```
In [27]: plt.figure(figsize=(4,3))  
sns.kdeplot(x='pop_data', data=pop_table)  
plt.show()
```



above graph shows that our data is not normally distributed, so we will apply CLT

```
In [33]: # First we will pick up random samples from population data  
# Pre-req: Sample should not be more than 10% population and more than 30 samples s  
  
# so calculate 10% of 10000 data  
  
10/100 * 10000
```

Out[33]: 1000.0

That means i.e. $n > 30$ and $n < 1000$, so are taking $n = [50, 500]$

```
In [32]: # To pick random data from population data  
np.random.choice(pop_data)
```

Out[32]: 37

```
In [37]: # So will take sample data less than 1000  
sample_mean = []  
# to take number of sample data 50 (to meet requirement  $n > 30$ )  
for no_of_sample in range(50):  
    sample_data = []  
    # to take number of sample data less than 1000 (so will take 500 sample)  
    for i in range(500):  
        sample_data.append(np.random.choice(pop_data))  
    # To calculate mean of sample data  
    sample_mean.append(np.mean(sample_data))
```

```
In [41]: len(sample_data), len(sample_mean)
```

Out[41]: (500, 50)

```
In [42]: sample_data
```

```
Out[42]: [78,  
78,  
16,  
32,  
24,  
68,  
74,  
89,  
44,  
21,  
57,  
61,  
27,  
70,  
34,  
84,  
29,  
77,  
77,  
24,  
13,  
88,  
37,  
43,  
17,  
50,  
98,  
36,  
87,  
24,  
15,  
69,  
77,  
52,  
89,  
24,  
93,  
56,  
42,  
42,  
67,  
36,  
20,  
50,  
45,  
43,  
88,  
46,  
13,  
17,  
95,  
68,  
10,  
26,  
95,  
37,
```


79,
69,
35,
97,
57,
86,
77,
76,
97,
23,
63,
60,
12,
54,
37,
27,
33,
89,
70,
58,
65,
55,
91,
35,
12,
86,
64,
59,
59,
64,
50,
36,
98,
64,
19,
22,
77,
32,
69,
95,
73,
12,
82,
42,
37,
90,
72,
43,
83,
32,
80,
85,
59,
38,
82,
50,

56,
91,
83,
73,
24,
66,
10,
84,
61,
76,
35,
80,
21,
70,
57,
30,
82,
50,
55,
23,
88,
90,
96,
56,
67,
98,
11,
61,
27,
10,
92,
37,
22,
71,
34,
90,
50,
98,
36,
36,
39,
94,
58,
40,
91,
91,
43,
61,
81,
50,
33,
47,
61,
41,
34,
81,

76,
76,
18,
41,
87,
99,
45,
65,
86,
21,
69,
83,
62,
19,
61,
66,
19,
17,
89,
46,
48,
95,
54,
65,
18,
49,
29,
88,
74,
38,
28,
93,
29,
54,
93,
23,
79,
64,
27,
68,
67,
75,
79,
77,
16,
51,
82,
69,
34,
80,
24,
17,
89,
37,
48,
12,

63,
76,
98,
91,
46,
61,
60,
75,
33,
95,
25,
91,
52,
26,
76,
38,
46,
13,
50,
20,
75,
99,
81,
91,
57,
87,
81,
69,
87,
39,
61,
83,
79,
85,
81,
38,
62,
71,
56,
21,
79,
61,
28,
58,
56,
30,
27,
89,
47,
91,
76,
27,
17,
35,
88,
97,

50,
32,
68,
72,
96,
91,
70,
23,
23,
63,
45,
63,
26,
62,
96,
84,
55,
65,
60,
29,
14,
80,
26,
50,
99,
65,
53,
39,
84,
56,
95,
66,
53,
62,
98,
68,
65,
65,
20,
27,
96,
91,
43,
36,
42,
79,
66,
59,
78,
71,
58,
36,
38,
17,
93,
85,

87,
90,
33,
25,
94,
58,
57,
60,
45,
53,
79,
97,
22,
85,
24,
53,
87,
93,
50,
31,
12,
31,
45,
91,
35,
58,
17,
52,
51,
38,
72,
37,
71,
13,
76,
82,
27,
19,
18,
22,
73,
19,
67,
89,
92,
23,
85,
87,
29,
95,
85,
71,
97,
72,
47,
89,

52,
32,
42,
56,
50,
44,
82,
90,
84,
62,
49,
41,
80,
21,
72,
72,
67,
27,
22,
96,
87,
25,
27,
96,
75,
42,
10,
34,
87,
62,
12,
86,
44,
43,
54,
81,
80,
84,
43,
59,
29,
20,
83,
59,
23,
37,
41,
53,
79,
25,
51,
25,
54,
53,
20,
90,

83,
30,
46,
96,
36,
16,
29,
77,
87,
65,
36,
14,
68,
29,
54,
96,
98,
85,
62,
92,
41,
43,
22,
89,
73,
89,
97,
96,
44,
44,
35,
84,
74,
50,
44,
18,
55,
31,
94,
18,
39,
72,
87,
34,
11,
63,
62,
99,
17,
56,
60,
29]

In [43]: sample_mean


```
Out[43]: [53.812,
54.18,
56.164,
56.114,
53.844,
54.842,
54.934,
54.79,
53.856,
55.86,
53.654,
57.08,
53.74,
54.942,
53.634,
54.91,
55.612,
54.014,
55.036,
54.076,
54.232,
54.018,
52.262,
55.06,
52.824,
52.568,
54.11,
55.224,
52.204,
54.616,
53.556,
54.684,
54.642,
54.806,
54.81,
53.248,
55.382,
54.486,
53.684,
52.336,
55.712,
53.55,
53.78,
54.64,
55.746,
54.278,
53.292,
56.57,
53.29,
57.052]
```

```
In [44]: # To see data in sample_mean is normally distributed or not
sample_mean_DF = pd.DataFrame({"Sample_mean":sample_mean})
```

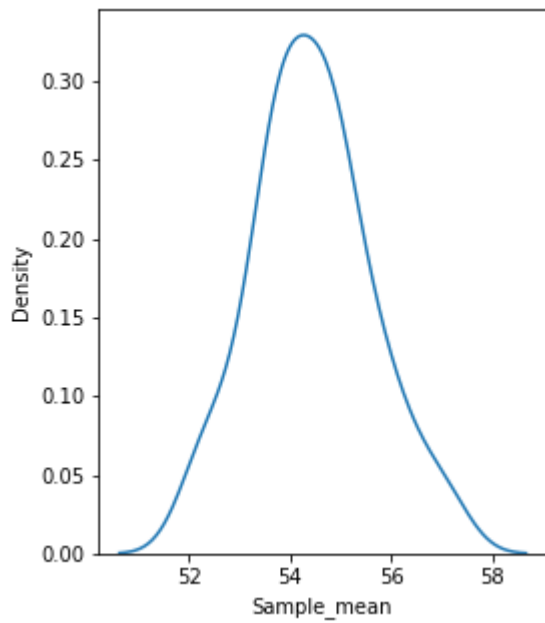
```
In [45]: sample_mean_DF
```

Out[45]:

Sample_mean	
0	53.812
1	54.180
2	56.164
3	56.114
4	53.844
5	54.842
6	54.934
7	54.790
8	53.856
9	55.860
10	53.654
11	57.080
12	53.740
13	54.942
14	53.634
15	54.910
16	55.612
17	54.014
18	55.036
19	54.076
20	54.232
21	54.018
22	52.262
23	55.060
24	52.824
25	52.568
26	54.110
27	55.224
28	52.204
29	54.616

Sample_mean	
30	53.556
31	54.684
32	54.642
33	54.806
34	54.810
35	53.248
36	55.382
37	54.486
38	53.684
39	52.336
40	55.712
41	53.550
42	53.780
43	54.640
44	55.746
45	54.278
46	53.292
47	56.570
48	53.290
49	57.052

```
In [50]: plt.figure(figsize=(4,5))
sns.kdeplot(x="Sample_mean", data=sample_mean_DF)
plt.show()
```



So the data is normally distributed

```
In [53]: # To meat another requirement of CLT that is the mean of population data and the me  
# so we will check the both means  
np.mean(pop_data), np.mean(sample_mean)
```

```
Out[53]: (54.3654, 54.43512)
```

```
In [ ]:
```