

Anscombe

August 1, 2025

```
[1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[2]: df1 = pd.read_excel("../..//datasets/Anscombe.xlsx", sheet_name="prvi")
df2 = pd.read_excel("../..//datasets/Anscombe.xlsx", sheet_name="drugi")
df3 = pd.read_excel("../..//datasets/Anscombe.xlsx", sheet_name="treći")
df4 = pd.read_excel("../..//datasets/Anscombe.xlsx", sheet_name="četvrti")
```

```
[3]: df1.head()
```

```
[3]:      x      y
0  10  8.04
1   8  6.95
2  13  7.58
3   9  8.81
4  11  8.33
```

```
[4]: df2.head()
```

```
[4]:      x      y
0  10  9.14
1   8  8.14
2  13  8.74
3   9  8.77
4  11  9.26
```

```
[5]: df3.head()
```

```
[5]:      x      y
0  10  7.46
1   8  6.77
2  13 12.74
3   9  7.11
4  11  7.81
```

```
[6]: df4.head()
```

```
[6]:      x      y
      0  8  6.58
      1  8  5.76
      2  8  7.71
      3  8  8.84
      4  8  8.47
```

```
[7]: df1.describe()
```

```
[7]:      x      y
count  11.000000  11.000000
mean    9.000000   7.500909
std     3.316625   2.031568
min     4.000000   4.260000
25%     6.500000   6.315000
50%     9.000000   7.580000
75%    11.500000   8.570000
max    14.000000  10.840000
```

```
[8]: df2.describe()
```

```
[8]:      x      y
count  11.000000  11.000000
mean    9.000000   7.500909
std     3.316625   2.031657
min     4.000000   3.100000
25%     6.500000   6.695000
50%     9.000000   8.140000
75%    11.500000   8.950000
max    14.000000   9.260000
```

```
[9]: df3.describe()
```

```
[9]:      x      y
count  11.000000  11.000000
mean    9.000000   7.500000
std     3.316625   2.030424
min     4.000000   5.390000
25%     6.500000   6.250000
50%     9.000000   7.110000
75%    11.500000   7.980000
max    14.000000  12.740000
```

```
[10]: df4.describe()
```

```
[10]:      x      y
count  11.000000  11.000000
mean    9.000000   7.500909
```

std	3.316625	2.030579
min	8.000000	5.250000
25%	8.000000	6.170000
50%	8.000000	7.040000
75%	8.000000	8.190000
max	19.000000	12.500000

```
[11]: df1.corr()
```

```
[11]:
```

	x	y
x	1.000000	0.816421
y	0.816421	1.000000

```
[12]: df2.corr()
```

```
[12]:
```

	x	y
x	1.000000	0.816237
y	0.816237	1.000000

```
[13]: df3.corr()
```

```
[13]:
```

	x	y
x	1.000000	0.816287
y	0.816287	1.000000

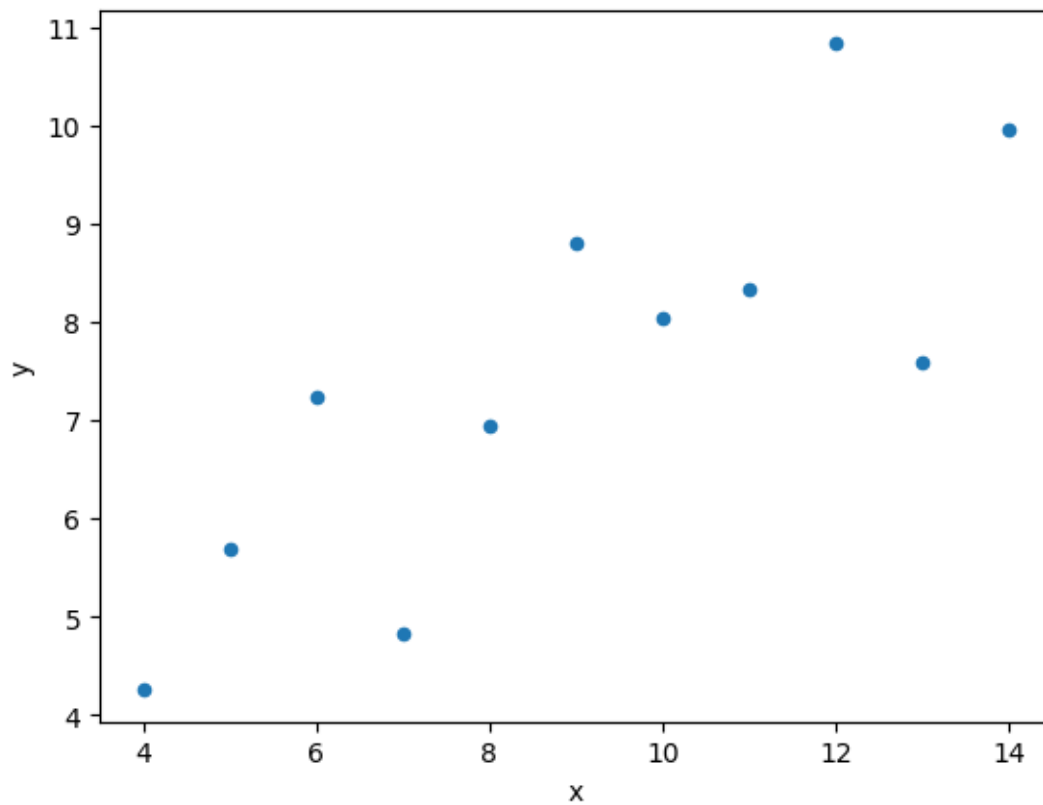
```
[14]: df4.corr()
```

```
[14]:
```

	x	y
x	1.000000	0.816521
y	0.816521	1.000000

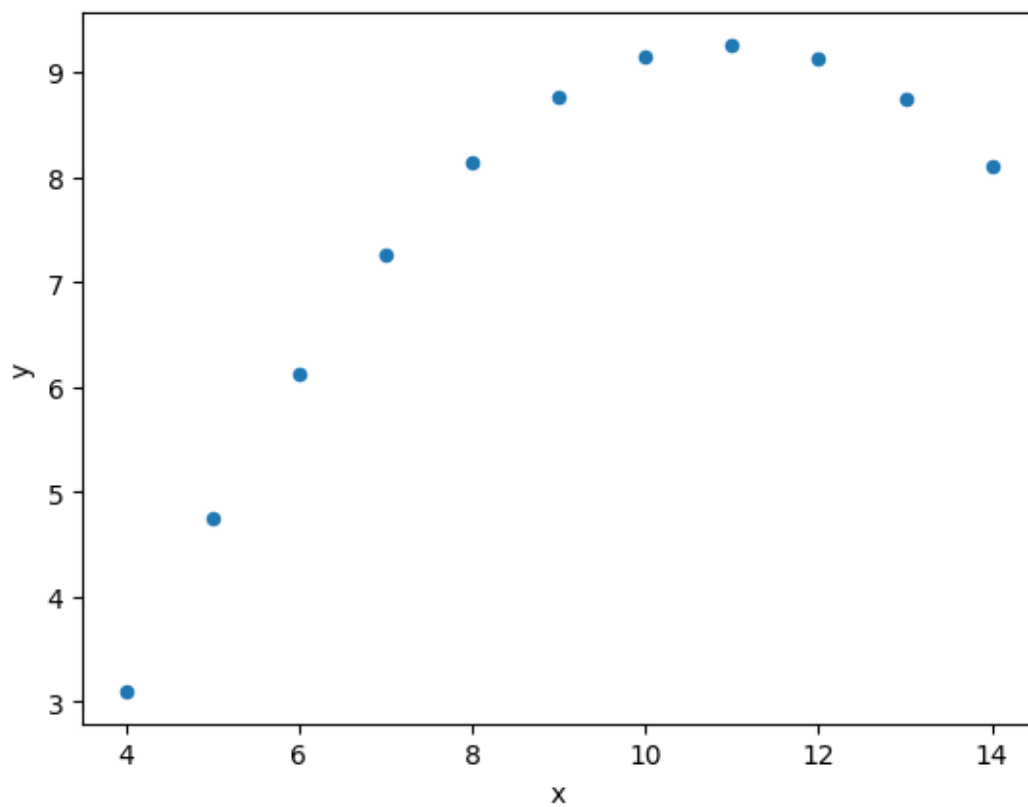
```
[15]: df1.plot(kind="scatter", x='x', y='y')
```

```
[15]: <Axes: xlabel='x', ylabel='y'>
```



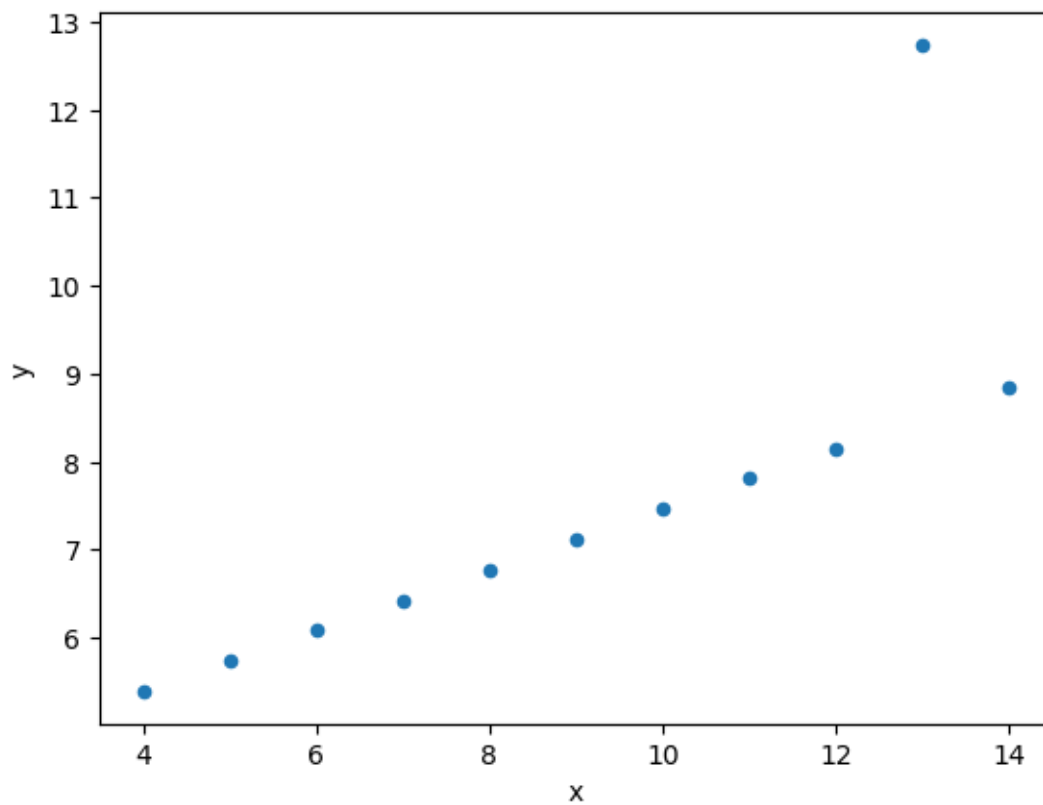
```
[16]: df2.plot(kind="scatter", x='x', y='y')
```

```
[16]: <Axes: xlabel='x', ylabel='y'>
```



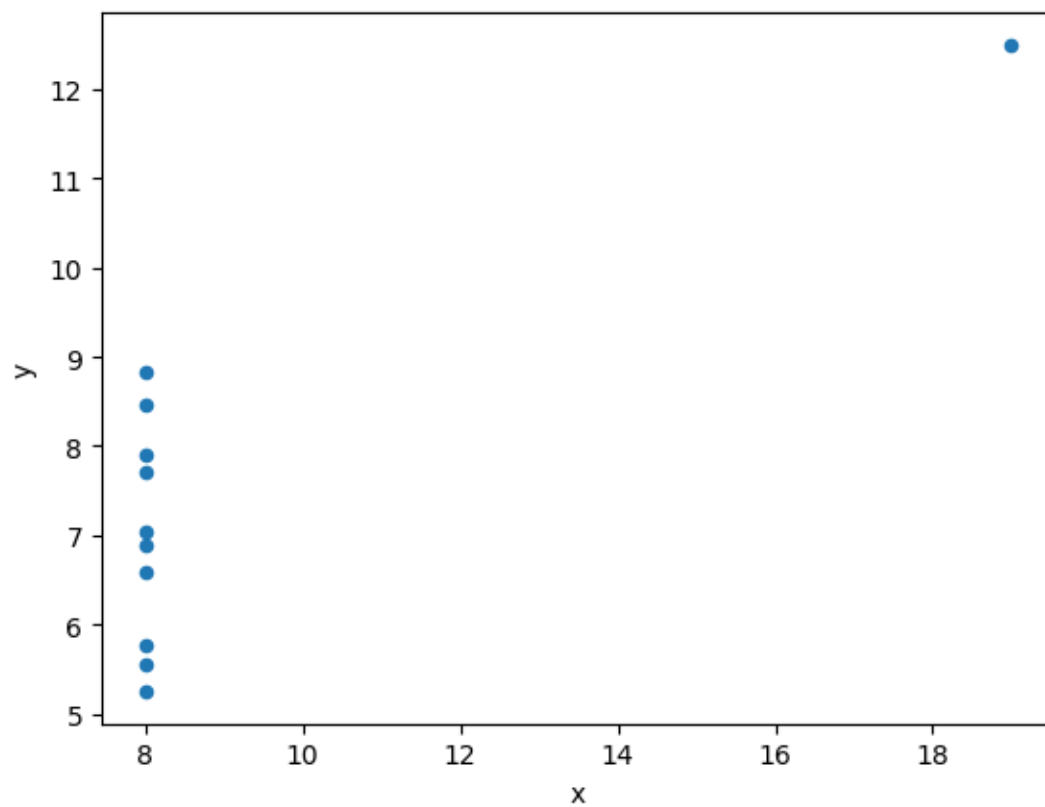
```
[17]: df3.plot(kind="scatter", x='x', y='y')
```

```
[17]: <Axes: xlabel='x', ylabel='y'>
```



```
[18]: df4.plot(kind="scatter", x='x', y='y')
```

```
[18]: <Axes: xlabel='x', ylabel='y'>
```



[]: