

Untitled

March 15, 2022

```
[1]: #momente der statistik
```

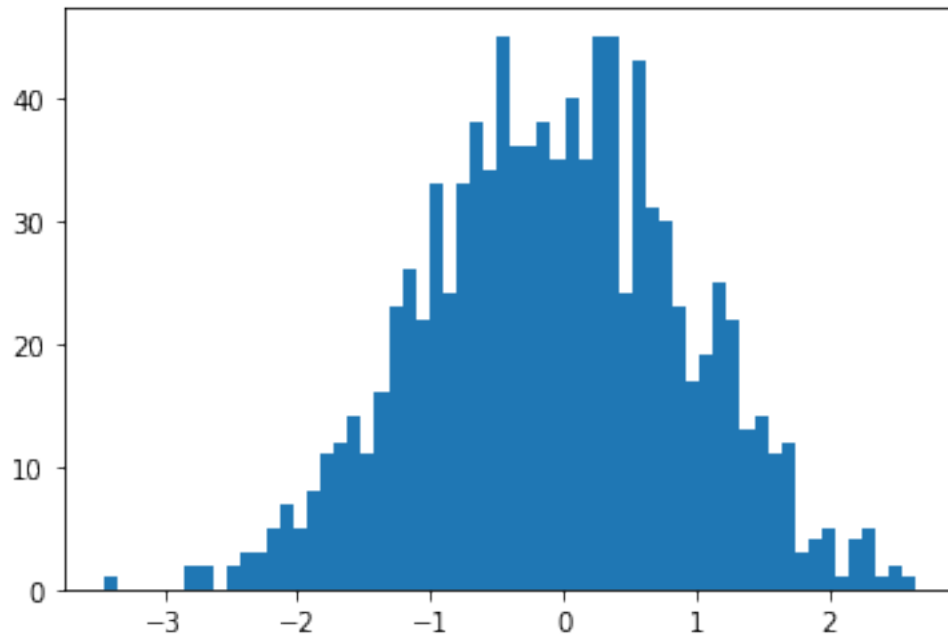
```
[2]: import numpy as np
import pandas as pd
%matplotlib inline
from scipy.stats import kurtosis #woelbung
from scipy.stats import skew # schiefe
import matplotlib.pyplot as plt
```

```
[3]: # data generierung in Normaldistribution (Gauss)
```

```
[4]: data = np.random.normal(0,1,1000)
```

```
[5]: plt.hist(data, bins=60)
```

```
[5]: (array([ 1.,  0.,  0.,  0.,  0.,  0.,  2.,  2.,  0.,  2.,  3.,  3.,  5.,
           7.,  5.,  8., 11., 12., 14., 11., 16., 23., 26., 22., 33., 24.,
          33., 38., 34., 45., 36., 36., 38., 35., 40., 35., 45., 45., 24.,
          43., 31., 30., 23., 17., 19., 25., 22., 13., 14., 11., 12.,  3.,
           4.,  5.,  1.,  4.,  5.,  1.,  2.,  1.]),
 array([-3.45226419, -3.35053276, -3.24880133, -3.1470699 , -3.04533847,
        -2.94360704, -2.84187561, -2.74014418, -2.63841275, -2.53668132,
        -2.43494989, -2.33321846, -2.23148703, -2.1297556 , -2.02802417,
        -1.92629274, -1.82456131, -1.72282988, -1.62109845, -1.51936702,
        -1.41763559, -1.31590416, -1.21417273, -1.1124413 , -1.01070987,
        -0.90897844, -0.80724701, -0.70551558, -0.60378415, -0.50205272,
        -0.40032129, -0.29858986, -0.19685843, -0.095127 ,  0.00660443,
         0.10833586,  0.21006729,  0.31179872,  0.41353015,  0.51526158,
         0.61699301,  0.71872444,  0.82045587,  0.9221873 ,  1.02391873,
         1.12565016,  1.22738159,  1.32911302,  1.43084445,  1.53257588,
         1.63430731,  1.73603874,  1.83777017,  1.9395016 ,  2.04123303,
         2.14296446,  2.24469589,  2.34642732,  2.44815875,  2.54989018,
         2.65162161]),
 <BarContainer object of 60 artists>)
```



```
[6]: print('mean', np.mean(data))
```

```
mean -0.0749082540984842
```

```
[7]: print('Varianz', np.var(data))
```

```
Varianz 0.9685977035397909
```

```
[9]: print('Schiefe', skew(data))
```

```
Schiefe -0.051420113651045624
```

```
[10]: print('Wölbung', kurtosis(data))
```

```
Wölbung -0.2076383215163462
```

```
[11]: Studierenden = {'hours':
    ↳ [29,9,10,38,16,26,50,10,30,33,43,2,39,15,44,29,41,15,24,50],
        'test_results':
    ↳ [65,7,8,76,23,56,100,3,74,48,73,0,62,37,74,40,90,42,58,100]}
```

```
[12]: import scipy.stats
```

```
[13]: students_data = pd.DataFrame(data=Studierenden)
```

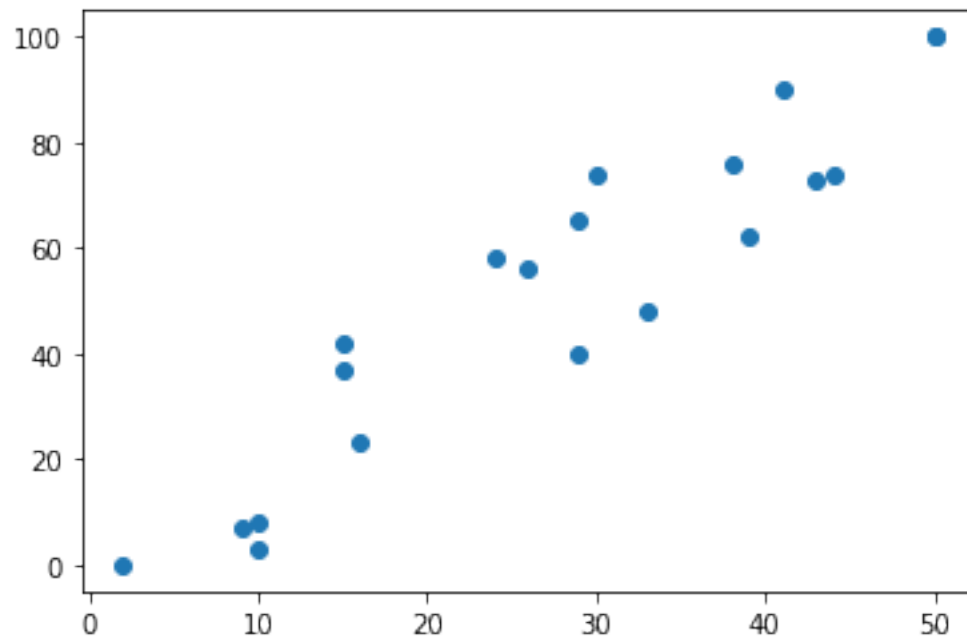
```
[14]: students_data
```

```
[14]:      hours  test_results
      0      29           65
      1       9           7
      2      10           8
      3      38          76
      4      16          23
      5      26          56
      6      50         100
      7      10           3
      8      30          74
      9      33          48
     10      43          73
     11       2           0
     12      39          62
     13      15          37
     14      44          74
     15      29          40
     16      41          90
     17      15          42
     18      24          58
     19      50         100
```

```
[15]: # Übung: bitte Mittelwert, Varianz, Schiefe und Wölbung  
# sowohl mit Bleistift als auch mit Python errechnen.
```

```
[16]: x = students_data.hours  
      y = students_data.test_results  
  
      plt.scatter(x,y)
```

```
[16]: <matplotlib.collections.PathCollection at 0x7f802092f4c0>
```

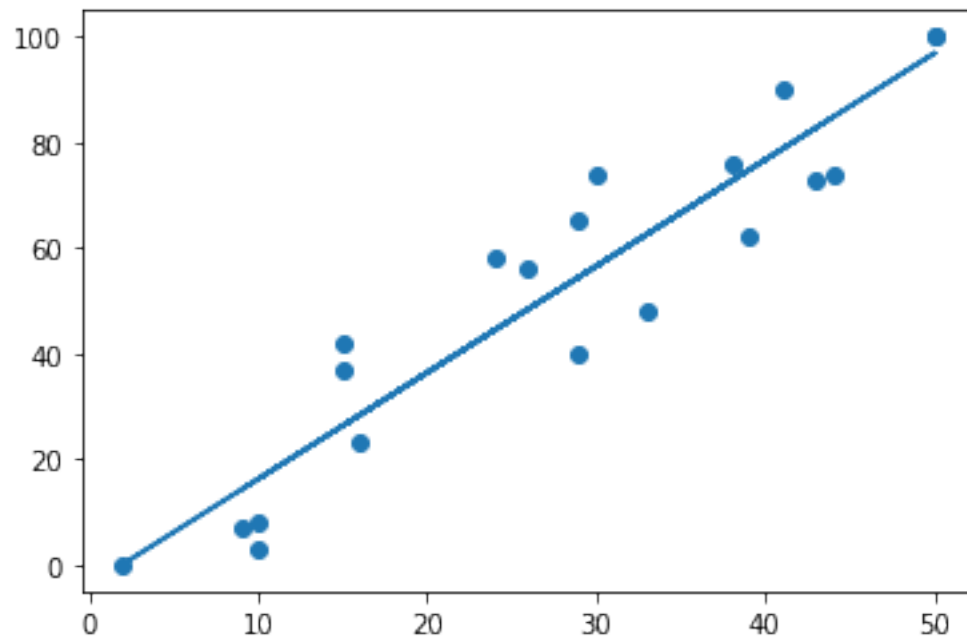


```
[17]: model = np.polyfit(x,y,1)
```

```
[18]: b0=model[1]
      b1=model[0]

      plt.scatter(x,y)
      plt.plot(x,(b0+b1*x))
```

```
[18]: [<matplotlib.lines.Line2D at 0x7f802096b8b0>]
```



```
[21]: predict = np.poly1d(model)
      hours_studied=45
      predict(hours_studied)
```

```
[21]: 86.75460901949661
```

```
[ ]:
```