20220517_Wirtschaftsinformatik_FAT2

May 18, 2022

[64]:
```
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

[65]:
```
'''Der k-means Algorithmus sucht nach einer vorgegebenen
Anzahl von Clustern innerhalb eines unbeschriteten mehrdimensionalen␣
 ↪Datensatzes.

Er erreicht durch dies eine "einfache" Vorstellung davon,
wie die optimale Clusterbildung aussieht.

Das Clusterzentrum ist das arythmetische Mittel Aller Punkte die zum Cluster␣
 ↪gehören.

Jeder Punkt liegt näher an seinem eigenen Clusterzentrum als an anderen␣
 ↪Clusterzentren.

Diese Beiden Annahmen sind die Grundlage des k-means Clusters Modells.'''
```

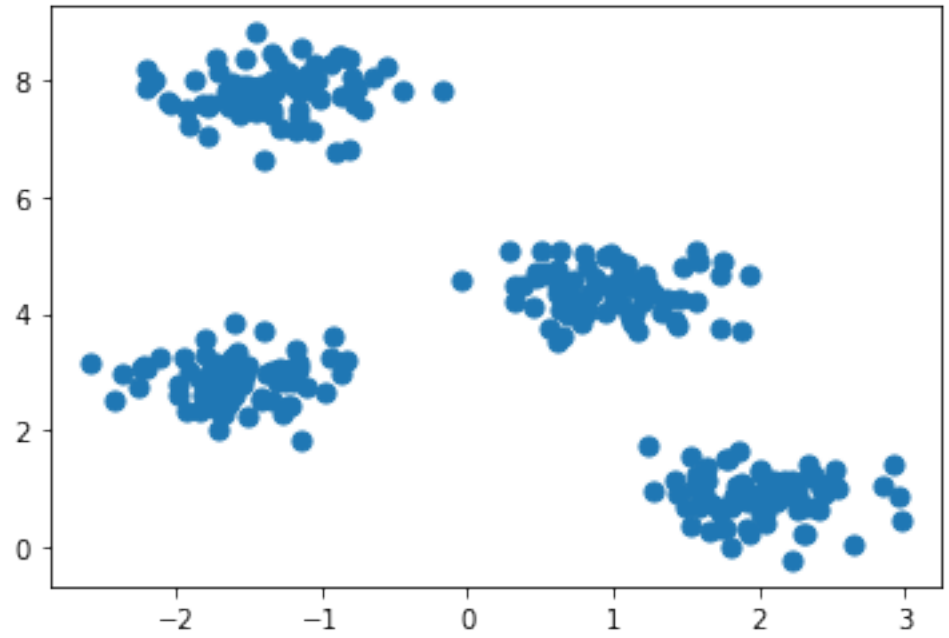
[65]:
```
'Der k-means Algorithmus sucht nach einer vorgegebenen \nAnzahl von Clustern
innerhalb eines unbeschriteten mehrdimensionalen Datensatzes.\n\nEr erreicht
durch dies eine "einfache" Vorstellung davon, \nwie die optimale Clusterbildung
aussieht.\n\nDas Clusterzentrum ist das arythmetische Mittel Aller Punkte die
zum Cluster gehören.\n\nJeder Punkt liegt näher an seinem eigenen Clusterzentrum
als an anderen Clusterzentren.\n\nDiese Beiden Annahmen sind die Grundlage des
k-means Clusters Modells.'
```

[66]:
```
# Beispiel 1. "Blobls k-Means"
```

[67]:
```
# Daten generation

from sklearn.datasets import make_blobs

X, y_true = make_blobs(n_samples=300, centers=4, cluster_std=0.4,␣
 ↪random_state=0)
```

```
plt.scatter(X[:,0], X[:,1], s=50)
```

[67]: <matplotlib.collections.PathCollection at 0x7fe94a060fd0>



[68]:
```
X
```

[68]:
```
array([[ 1.24299373e+00,  1.72346080e+00],
       [-1.35852399e+00,  7.55156883e+00],
       [ 1.09556535e+00,  4.83434169e+00],
       [-1.09516280e+00,  7.82176310e+00],
       [ 1.53265678e+00,  1.56091593e+00],
       [ 2.97683420e+00,  4.73657331e-01],
       [-1.71444969e+00,  2.03730575e+00],
       [ 1.26790630e+00,  4.35538049e+00],
       [-5.52706704e-01,  8.23321777e+00],
       [-8.90151666e-01,  8.38544564e+00],
       [-1.65054979e+00,  2.82449560e+00],
       [-2.19609054e+00,  8.18108094e+00],
       [ 1.15381537e+00,  4.43725706e+00],
       [-1.90130959e+00,  7.20835293e+00],
       [ 1.65518138e+00,  2.79755221e-01],
       [ 6.27951219e-01,  4.07224746e+00],
       [ 1.03824905e+00,  4.45505234e+00],
       [ 2.39799977e+00,  6.37253423e-01],
       [-1.22023874e+00,  3.06039939e+00],
       [-1.22802868e+00,  2.44230428e+00],
```

```
[ 2.49513136e+00,  1.15976915e+00],
[ 1.86982913e+00,  1.09025625e+00],
[ 8.14999300e-01,  4.79276536e+00],
[-1.83995568e+00,  2.87372654e+00],
[-1.78782145e+00,  2.70940654e+00],
[-4.49258478e-02,  4.56523477e+00],
[ 1.81674191e+00,  8.76636741e-01],
[ 7.21931647e-01,  4.57436065e+00],
[-2.42912970e+00,  2.50887952e+00],
[ 4.49907114e-01,  4.11915349e+00],
[-7.58307549e-01,  7.86140246e+00],
[-1.78742594e+00,  7.53083066e+00],
[ 1.57390171e+00,  4.22172402e+00],
[-1.17022790e+00,  7.12283487e+00],
[-1.51227830e+00,  7.90578760e+00],
[-2.02897742e+00,  7.57150332e+00],
[-1.76013144e+00,  7.60128753e+00],
[-2.16370379e+00,  7.93605378e+00],
[-1.97763434e+00,  2.62561116e+00],
[ 1.79430409e+00,  7.41282310e-01],
[ 7.71948024e-01,  3.83153445e+00],
[-1.83256158e+00,  2.34276567e+00],
[ 2.93762002e-01,  5.08409749e+00],
[ 1.35517087e+00,  4.24178329e+00],
[-1.64363896e+00,  2.61328538e+00],
[-1.19025151e+00,  2.81809883e+00],
[-1.33503560e+00,  8.01361732e+00],
[-1.54377259e+00,  2.80312738e+00],
[-1.40211394e+00,  7.65792558e+00],
[ 2.01105126e+00,  1.30573274e+00],
[-1.80878412e+00,  3.29518655e+00],
[ 2.32182075e+00,  7.13375745e-01],
[-1.01230385e+00,  7.68991649e+00],
[ 1.81634109e+00,  8.02494968e-01],
[ 1.85605454e+00,  1.66947648e+00],
[-1.24939475e+00,  2.85405289e+00],
[-4.43753102e-01,  7.81762185e+00],
[-1.26957823e+00,  2.28963290e+00],
[-1.62031838e+00,  7.76402438e+00],
[ 1.61351419e+00,  9.18529692e-01],
[-1.92074306e+00,  7.47987147e+00],
[ 7.06537014e-01,  4.31651955e+00],
[-1.44938246e+00,  8.80044149e+00],
[-1.20109608e+00,  2.94145250e+00],
[-1.91534584e+00,  3.04380914e+00],
[-1.29998790e+00,  2.82881222e+00],
[-2.05481843e+00,  7.61967816e+00],
```

```
[ 2.26482393e+00,  9.33032495e-01],
[-2.19392530e+00,  7.84359369e+00],
[-1.50710602e+00,  3.11541697e+00],
[ 1.20690641e+00,  4.22046783e+00],
[-8.33815538e-01,  3.19168270e+00],
[-1.10754373e+00,  7.98715423e+00],
[-1.57894679e+00,  2.95546355e+00],
[-1.51145179e+00,  2.25519622e+00],
[-9.39693395e-01,  8.24723555e+00],
[-1.68085594e+00,  2.95562290e+00],
[ 1.11881664e+00,  4.58641659e+00],
[ 2.11204879e+00,  7.69932293e-01],
[-1.10635270e+00,  8.08221464e+00],
[ 1.52156413e+00,  3.58976658e-01],
[ 7.22541241e-01,  4.15869086e+00],
[ 1.53812476e+00,  1.00448401e+00],
[ 1.83812293e+00,  1.06408368e+00],
[-1.04877966e+00,  8.25464888e+00],
[ 9.48973436e-01,  4.98912442e+00],
[ 2.26457819e+00,  8.29045127e-01],
[ 6.11141188e-01,  4.75059384e+00],
[-1.78114044e+00,  7.04801014e+00],
[-1.38723604e+00,  7.60295262e+00],
[ 6.78368150e-01,  3.97321191e+00],
[ 1.87305452e+00,  9.04655324e-01],
[-1.71991895e+00,  8.35603124e+00],
[-1.67278143e+00,  2.98056380e+00],
[-1.29549541e+00,  3.05774404e+00],
[ 1.10149716e+00,  3.96214903e+00],
[ 1.59279455e+00,  1.21014290e+00],
[ 2.41933908e+00,  1.02455095e+00],
[ 5.07020717e-01,  5.08123580e+00],
[-2.59057291e+00,  3.16041007e+00],
[-8.08142039e-01,  8.35466880e+00],
[ 1.79782016e+00,  8.30239905e-03],
[-8.71268527e-01,  7.72842212e+00],
[ 2.26369347e+00,  6.67348472e-01],
[ 1.02703492e+00,  4.46458307e+00],
[ 2.16980500e+00,  1.14120119e+00],
[ 2.31132013e+00,  2.50881242e-01],
[ 1.22190183e+00,  4.67267000e+00],
[-1.29245154e+00,  7.17118009e+00],
[ 9.80470087e-01,  5.01813553e+00],
[-1.21760282e+00,  2.44433001e+00],
[-1.58038464e+00,  3.34897978e+00],
[ 1.73390366e+00,  6.21843749e-01],
[-1.86842751e+00,  8.00238754e+00],
```

```
[ 1.90759479e+00,  8.01911989e-01],
[ 1.72953036e+00,  3.76468370e+00],
[-1.51167696e+00,  7.62976643e+00],
[ 1.42209216e+00,  1.14181541e+00],
[ 1.75944232e+00,  1.51486950e+00],
[ 6.31779805e-01,  5.06781331e+00],
[-8.67650040e-01,  2.98357336e+00],
[ 1.80403250e+00,  7.05252813e-01],
[-1.18373444e+00,  3.37432301e+00],
[ 1.57972382e+00,  6.94937118e-01],
[ 1.63716617e+00,  1.38212178e+00],
[ 2.04845936e+00,  1.04932435e+00],
[ 1.80029271e+00,  7.38754934e-01],
[-1.41472733e+00,  2.52063282e+00],
[ 1.43894872e+00,  9.22968458e-01],
[-1.66827671e+00,  2.27129259e+00],
[-6.42924542e-01,  8.05671284e+00],
[-1.60092548e+00,  2.59482287e+00],
[-1.63399743e+00,  2.46667773e+00],
[ 2.30536010e+00,  2.56840598e-01],
[-1.33607219e+00,  7.40184537e+00],
[-1.34393777e+00,  2.53203746e+00],
[-2.10551789e+00,  3.23800144e+00],
[-1.81064094e+00,  7.59943696e+00],
[ 6.51011566e-01,  3.61327429e+00],
[-7.75704334e-01,  7.58269851e+00],
[-2.14409740e+00,  7.99605964e+00],
[-1.20667809e+00,  2.94918633e+00],
[ 1.16253505e+00,  3.68928985e+00],
[-1.58997082e+00,  3.82057166e+00],
[ 1.57157096e+00,  5.06214300e+00],
[-1.68499561e+00,  2.45411405e+00],
[-1.33729805e+00,  7.47022833e+00],
[ 3.30710940e-01,  4.21869122e+00],
[-1.15140859e+00,  7.47997191e+00],
[-1.61176107e+00,  7.66573297e+00],
[-1.42393158e+00,  7.90799539e+00],
[ 9.64997187e-01,  4.47512008e+00],
[-8.16976853e-01,  6.81178615e+00],
[ 4.75151935e-01,  4.61478347e+00],
[ 1.72992182e+00,  3.11093929e-01],
[-1.19826973e+00,  2.91999932e+00],
[-1.03093102e+00,  8.01107720e+00],
[-1.76636559e+00,  2.47152347e+00],
[ 1.93091305e+00,  9.36623726e-01],
[ 1.05958007e+00,  4.69444294e+00],
[-1.70417676e+00,  8.14944302e+00],
```

```
[ 7.72409206e-01,  4.12855761e+00],
[ 8.21339352e-01,  4.18286623e+00],
[ 1.77844758e+00,  1.51221448e+00],
[ 8.69068730e-01,  4.62476989e+00],
[-2.26413183e+00,  2.72669266e+00],
[-1.14968558e+00,  1.82201139e+00],
[ 4.68076079e-01,  4.69154601e+00],
[ 1.64157038e+00,  1.17030147e+00],
[ 1.13467276e+00,  3.86656272e+00],
[ 1.04724054e+00,  4.14307495e+00],
[-1.23345349e+00,  8.14262098e+00],
[ 1.67288952e+00,  7.59270950e-01],
[ 1.58938176e+00,  4.89153084e+00],
[-1.39134236e+00,  3.72629969e+00],
[-1.18268435e+00,  7.87398856e+00],
[ 2.33119479e+00,  1.41840215e+00],
[ 2.22577101e+00,  1.16842687e+00],
[ 7.07285899e-01,  4.15996606e+00],
[-1.15416466e+00,  3.05386825e+00],
[ 2.43503584e+00,  9.32684156e-01],
[ 1.44778191e+00,  4.23181739e+00],
[-1.98984612e+00,  2.79296536e+00],
[-1.56127632e+00,  2.99559944e+00],
[ 3.79767041e-01,  4.47954401e+00],
[ 1.28068517e+00,  4.35245733e+00],
[ 1.03388751e+00,  4.88549673e+00],
[ 8.37105219e-01,  4.36632692e+00],
[ 2.33188302e+00,  1.17556332e+00],
[-1.57050638e+00,  7.38813525e+00],
[ 5.56848893e-01,  3.73578015e+00],
[-1.17657091e+00,  2.87183927e+00],
[ 1.00287697e+00,  4.42477609e+00],
[ 6.98442935e-01,  4.24393351e+00],
[-1.92110831e+00,  2.32914826e+00],
[-1.41709747e+00,  2.56151623e+00],
[-2.36474524e+00,  2.96737103e+00],
[ 1.42692445e+00,  3.87181472e+00],
[-1.80203906e+00,  2.43225130e+00],
[-1.62600317e+00,  7.93070127e+00],
[ 1.12684029e+00,  3.86402701e+00],
[-1.15510197e+00,  3.15077210e+00],
[ 1.93166233e+00,  2.27262137e-01],
[-1.71364223e+00,  2.35111982e+00],
[ 1.13720674e+00,  4.02986329e+00],
[-1.40521137e+00,  6.61700279e+00],
[-1.55155457e+00,  2.87496015e+00],
[-8.73558789e-01,  8.40039110e+00],
```

```
[ 1.04293948e+00,  4.55779990e+00],
[-1.06674301e+00,  7.10356385e+00],
[ 1.24518798e+00,  4.46677206e+00],
[-1.60316541e+00,  2.75994246e+00],
[ 3.24190740e-01,  4.48890023e+00],
[ 1.46838635e+00,  4.78473927e+00],
[-1.35291395e+00,  7.76256210e+00],
[-9.81091274e-01,  2.64210259e+00],
[-1.09950025e+00,  2.73652794e+00],
[ 2.00800190e+00,  6.25592378e-01],
[ 2.65306134e+00,  6.96696499e-02],
[ 6.13350733e-01,  4.32456549e+00],
[-1.66022989e+00,  7.69548267e+00],
[ 2.36979871e+00,  7.11096021e-01],
[ 2.04553707e+00,  6.02451296e-01],
[-2.22926025e+00,  3.09825605e+00],
[ 1.76502857e+00,  3.44318078e-01],
[-2.23431939e+00,  3.06007498e+00],
[ 6.21155779e-01,  3.51146874e+00],
[-7.20601024e-01,  7.48249249e+00],
[-1.16853570e+00,  7.81521860e+00],
[ 1.74544689e+00,  4.89599324e+00],
[ 9.94573485e-01,  4.22891379e+00],
[-1.34506371e+00,  8.44276448e+00],
[ 8.10822486e-01,  4.00480540e+00],
[ 1.28075560e+00,  9.73175099e-01],
[-1.24752115e+00,  2.91939062e+00],
[ 8.51649066e-01,  4.32625346e+00],
[ 2.00211529e+00,  7.78547308e-01],
[-1.34165188e+00,  3.02952057e+00],
[-1.13454391e+00,  8.53252753e+00],
[-1.75462883e+00,  3.02584400e+00],
[ 2.53047944e+00,  1.02444070e+00],
[ 5.10210142e-01,  4.66411792e+00],
[ 2.95899110e+00,  8.80760799e-01],
[-1.46997963e+00,  7.64720495e+00],
[-1.43626893e+00,  7.74876743e+00],
[-1.63245753e+00,  7.51821307e+00],
[-1.53987364e+00,  7.91408298e+00],
[-1.60966548e+00,  3.26995383e+00],
[-1.37857401e+00,  2.97470698e+00],
[-9.03217370e-01,  6.77321238e+00],
[ 1.92952799e+00,  4.68157912e+00],
[-1.53317686e+00,  2.98225353e+00],
[ 1.67748902e+00,  7.33643783e-01],
[ 6.68303649e-01,  4.51948700e+00],
[-1.71876634e+00,  3.16602558e+00],
```

```
       [-1.80050837e+00,  3.58170258e+00],
       [-9.18906069e-01,  3.60571798e+00],
       [ 2.03955439e+00,  4.30426261e-01],
       [ 2.11787014e+00,  9.90536074e-01],
       [-1.32941451e+00,  7.79150690e+00],
       [ 9.34982538e-01,  4.46802673e+00],
       [ 5.47969030e-01,  4.72556802e+00],
       [-1.94875540e+00,  3.24598140e+00],
       [ 1.59628006e+00,  7.22535642e-01],
       [-7.96818149e-01,  8.03386039e+00],
       [-1.17931862e+00,  3.02863102e+00],
       [ 1.32204456e+00,  4.00692132e+00],
       [-8.09717036e-01,  7.74177370e+00],
       [ 8.02208658e-01,  5.04349282e+00],
       [ 2.09435751e+00,  1.13084513e+00],
       [ 1.48564316e+00,  7.00335707e-01],
       [-2.20614634e+00,  3.07279445e+00],
       [-1.81474577e+00,  2.59268507e+00],
       [ 1.88417193e+00,  3.72204106e+00],
       [ 2.92056190e+00,  1.43227484e+00],
       [ 1.91366996e+00,  3.47683143e-01],
       [ 1.56509331e+00,  1.23540885e+00],
       [ 1.72329366e+00,  4.66620519e+00],
       [-1.24480462e+00,  8.04626170e+00],
       [-1.65547252e+00,  7.80431811e+00],
       [ 2.51620015e+00,  1.32951110e+00],
       [ 2.42361105e+00,  1.02515472e+00],
       [ 1.43203035e+00,  3.80985700e+00],
       [ 2.25276423e+00,  8.51222084e-01],
       [ 2.16723736e+00,  8.58403504e-01],
       [ 2.83763244e+00,  1.05370099e+00],
       [-1.45526339e+00,  7.44392807e+00],
       [-9.40272527e-01,  3.25890304e+00],
       [-1.30067738e+00,  8.28869197e+00],
       [ 9.36889069e-01,  4.03839601e+00],
       [ 2.21488606e+00, -2.11373443e-01],
       [ 1.89548791e+00,  1.04568602e+00],
       [-1.16286376e+00,  7.35203055e+00],
       [-1.51946609e+00,  8.35459833e+00],
       [-1.26653736e+00,  7.92366308e+00],
       [ 2.00088764e+00,  1.35242020e+00],
       [ 2.36398374e+00,  1.22706532e+00],
       [ 6.18083454e-01,  4.45854833e+00],
       [-1.69766154e-01,  7.80589015e+00],
       [-1.70327307e+00,  2.80574006e+00]])
```

[69]:
```
X[:,0]
```

```
[69]: array([ 1.24299373, -1.35852399,  1.09556535, -1.0951628 ,  1.53265678,
               2.9768342 , -1.71444969,  1.2679063 , -0.5527067 , -0.89015167,
              -1.65054979, -2.19609054,  1.15381537, -1.90130959,  1.65518138,
               0.62795122,  1.03824905,  2.39799977, -1.22023874, -1.22802868,
               2.49513136,  1.86982913,  0.8149993 , -1.83995568, -1.78782145,
              -0.04492585,  1.81674191,  0.72193165, -2.4291297 ,  0.44990711,
              -0.75830755, -1.78742594,  1.57390171, -1.1702279 , -1.5122783 ,
              -2.02897742, -1.76013144, -2.16370379, -1.97763434,  1.79430409,
               0.77194802, -1.83256158,  0.293762  ,  1.35517087, -1.64363896,
              -1.19025151, -1.3350356 , -1.54377259, -1.40211394,  2.01105126,
              -1.80878412,  2.32182075, -1.01230385,  1.81634109,  1.85605454,
              -1.24939475, -0.4437531 , -1.26957823, -1.62031838,  1.61351419,
              -1.92074306,  0.70653701, -1.44938246, -1.20109608, -1.91534584,
              -1.2999879 , -2.05481843,  2.26482393, -2.1939253 , -1.50710602,
               1.20690641, -0.83381554, -1.10754373, -1.57894679, -1.51145179,
              -0.9396934 , -1.68085594,  1.11881664,  2.11204879, -1.1063527 ,
               1.52156413,  0.72254124,  1.53812476,  1.83812293, -1.04877966,
               0.94897344,  2.26457819,  0.61114119, -1.78114044, -1.38723604,
               0.67836815,  1.87305452, -1.71991895, -1.67278143, -1.29549541,
               1.10149716,  1.59279455,  2.41933908,  0.50702072, -2.59057291,
              -0.80814204,  1.79782016, -0.87126853,  2.26369347,  1.02703492,
               2.169805  ,  2.31132013,  1.22190183, -1.29245154,  0.98047009,
              -1.21760282, -1.58038464,  1.73390366, -1.86842751,  1.90759479,
               1.72953036, -1.51167696,  1.42209216,  1.75944232,  0.6317798 ,
              -0.86765004,  1.8040325 , -1.18373444,  1.57972382,  1.63716617,
               2.04845936,  1.80029271, -1.41472733,  1.43894872, -1.66827671,
              -0.64292454, -1.60092548, -1.63399743,  2.3053601 , -1.33607219,
              -1.34393777, -2.10551789, -1.81064094,  0.65101157, -0.77570433,
              -2.1440974 , -1.20667809,  1.16253505, -1.58997082,  1.57157096,
              -1.68499561, -1.33729805,  0.33071094, -1.15140859, -1.61176107,
              -1.42393158,  0.96499719, -0.81697685,  0.47515193,  1.72992182,
              -1.19826973, -1.03093102, -1.76636559,  1.93091305,  1.05958007,
              -1.70417676,  0.77240921,  0.82133935,  1.77844758,  0.86906873,
              -2.26413183, -1.14968558,  0.46807608,  1.64157038,  1.13467276,
               1.04724054, -1.23345349,  1.67288952,  1.58938176, -1.39134236,
              -1.18268435,  2.33119479,  2.22577101,  0.7072859 , -1.15416466,
               2.43503584,  1.44778191, -1.98984612, -1.56127632,  0.37976704,
               1.28068517,  1.03388751,  0.83710522,  2.33188302, -1.57050638,
               0.55684889, -1.17657091,  1.00287697,  0.69844293, -1.92110831,
              -1.41709747, -2.36474524,  1.42692445, -1.80203906, -1.62600317,
               1.12684029, -1.15510197,  1.93166233, -1.71364223,  1.13720674,
              -1.40521137, -1.55155457, -0.87355879,  1.04293948, -1.06674301,
               1.24518798, -1.60316541,  0.32419074,  1.46838635, -1.35291395,
              -0.98109127, -1.09950025,  2.0080019 ,  2.65306134,  0.61335073,
              -1.66022989,  2.36979871,  2.04553707, -2.22926025,  1.76502857,
              -2.23431939,  0.62115578, -0.72060102, -1.1685357 ,  1.74544689,
               0.99457349, -1.34506371,  0.81082249,  1.2807556 , -1.24752115,
```

```
 0.85164907,  2.00211529, -1.34165188, -1.13454391, -1.75462883,
 2.53047944,  0.51021014,  2.9589911 , -1.46997963, -1.43626893,
-1.63245753, -1.53987364, -1.60966548, -1.37857401, -0.90321737,
 1.92952799, -1.53317686,  1.67748902,  0.66830365, -1.71876634,
-1.80050837, -0.91890607,  2.03955439,  2.11787014, -1.32941451,
 0.93498254,  0.54796903, -1.9487554 ,  1.59628006, -0.79681815,
-1.17931862,  1.32204456, -0.80971704,  0.80220866,  2.09435751,
 1.48564316, -2.20614634, -1.81474577,  1.88417193,  2.9205619 ,
 1.91366996,  1.56509331,  1.72329366, -1.24480462, -1.65547252,
 2.51620015,  2.42361105,  1.43203035,  2.25276423,  2.16723736,
 2.83763244, -1.45526339, -0.94027253, -1.30067738,  0.93688907,
 2.21488606,  1.89548791, -1.16286376, -1.51946609, -1.26653736,
 2.00088764,  2.36398374,  0.61808345, -0.16976615, -1.70327307])
```

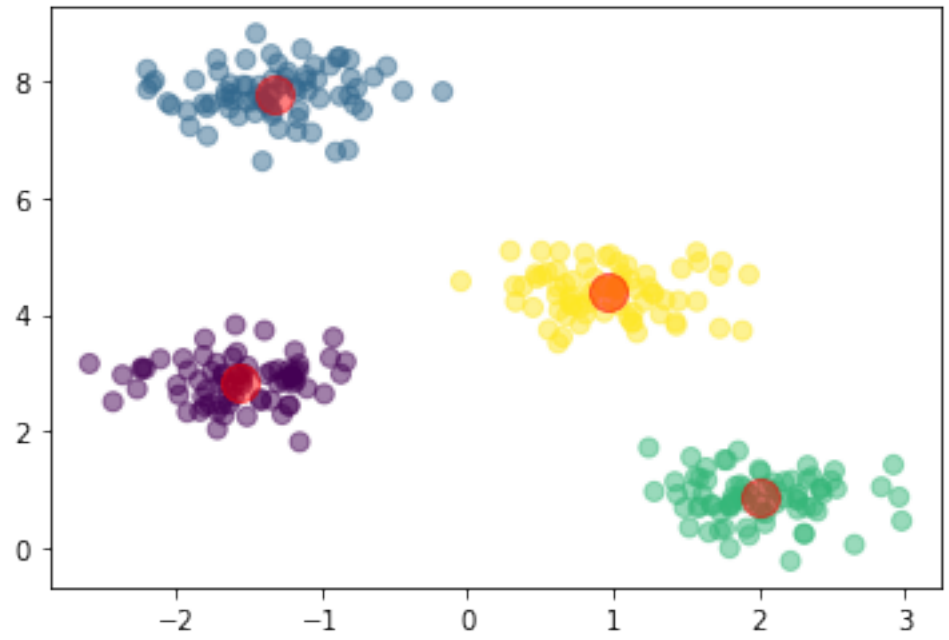
[70]:
```
y_true
```

[70]:
```
array([1, 3, 0, 3, 1, 1, 2, 0, 3, 3, 2, 3, 0, 3, 1, 0, 0, 1, 2, 2, 1, 1,
       0, 2, 2, 0, 1, 0, 2, 0, 3, 3, 0, 3, 3, 3, 3, 3, 2, 1, 0, 2, 0, 0,
       2, 2, 3, 2, 3, 1, 2, 1, 3, 1, 1, 2, 3, 2, 3, 1, 3, 0, 3, 2, 2, 2,
       3, 1, 3, 2, 0, 2, 3, 2, 2, 3, 2, 0, 1, 3, 1, 0, 1, 1, 3, 0, 1, 0,
       3, 3, 0, 1, 3, 2, 2, 0, 1, 1, 0, 2, 3, 1, 3, 1, 0, 1, 1, 0, 3, 0,
       2, 2, 1, 3, 1, 0, 3, 1, 1, 0, 2, 1, 2, 1, 1, 1, 1, 2, 1, 2, 3, 2,
       2, 1, 3, 2, 2, 3, 0, 3, 3, 2, 0, 2, 0, 2, 3, 0, 3, 3, 3, 0, 3, 0,
       1, 2, 3, 2, 1, 0, 3, 0, 0, 1, 0, 2, 2, 0, 1, 0, 0, 3, 1, 0, 2, 3,
       1, 1, 0, 2, 1, 0, 2, 2, 0, 0, 0, 0, 1, 3, 0, 2, 0, 0, 2, 2, 2, 0,
       2, 3, 0, 2, 1, 2, 0, 3, 2, 3, 0, 3, 0, 2, 0, 0, 3, 2, 2, 1, 1, 0,
       3, 1, 1, 2, 1, 2, 0, 3, 3, 0, 0, 3, 0, 1, 2, 0, 1, 2, 3, 2, 1, 0,
       1, 3, 3, 3, 3, 2, 2, 3, 0, 2, 1, 0, 2, 2, 2, 1, 1, 3, 0, 0, 2, 1,
       3, 2, 0, 3, 0, 1, 1, 2, 2, 0, 1, 1, 1, 0, 3, 3, 1, 1, 0, 1, 1, 1,
       3, 2, 3, 0, 1, 1, 3, 3, 3, 1, 1, 0, 3, 2])
```

[71]:
```
# Lösung 1. Direkt

from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=4) # hier sagen wir KMeans wie viele Clusters wir␣
 ↪haben!
kmeans.fit(X) # hier lassen wir die Clusters berechnen
y_means = kmeans.predict(X) # hier geben wir Jedem Punkt einen Label
```

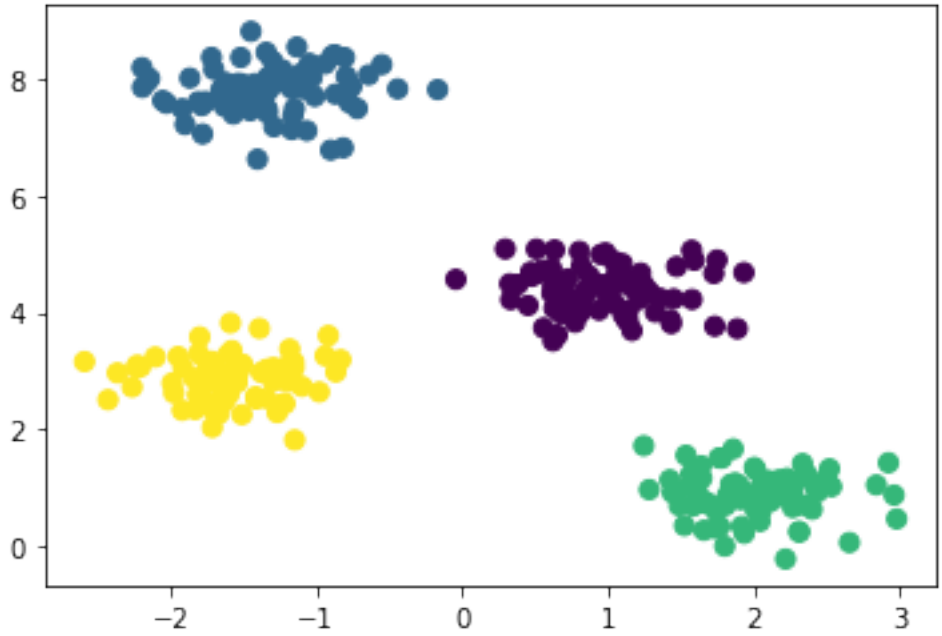
[72]:
```
plt.scatter(X[:,0], X[:,1], c=y_means, s=50, cmap='viridis', alpha=0.5)
# hier stellen wir graphisch die Punkte dar.
# die Farben oder Clusters werden nach den y_labels ausgewählt.
centers=kmeans.cluster_centers_
# hier erzeugen wir eine Variabel welche die Clusterzentren enthält
plt.scatter(centers[:,0], centers[:,1], c='red', s=200, alpha=0.5)
# hier stellen wir graphisch die Clusterzentren dar.
```

[72]: <matplotlib.collections.PathCollection at 0x7fe9495d4850>



[73]:
```
# Lösung 2. Mit Code

from sklearn.metrics import pairwise_distances_argmin

def find_clusters(X, n_clusters, rseed=2):
    # 1. Random auswahl von Clusters
    rng = np.random.RandomState(rseed)
    i = rng.permutation(X.shape[0])[:n_clusters]
    centers = X[i]
    while True:
    # 2. Labels werden an den am nähesten Center zugeordnet
        labels = pairwise_distances_argmin(X,centers)
    # 3 Finden Sie neue Zentren aus den Mittelwerten
        new_centers = np.array([X[labels==i].mean(0)
        for i in range(n_clusters)])
    # 4 Konvergenz finden

        if np.all(centers == new_centers):
            break
        centers = new_centers
    return centers, labels

centers, labels = find_clusters(X,4)
plt.scatter(X[:,0], X[:,1], c= labels, s=50, cmap='viridis')
```

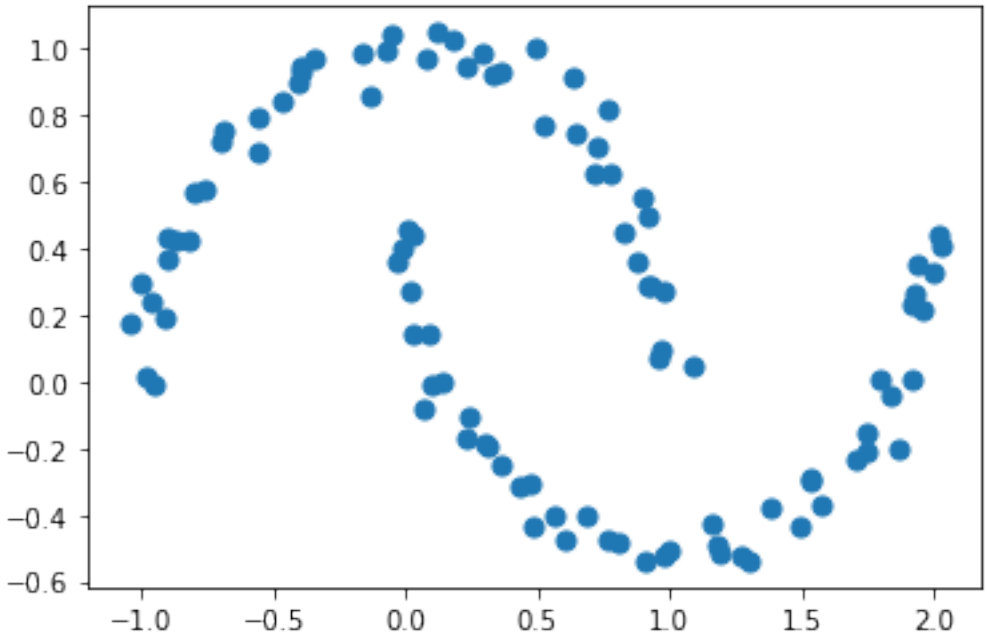
[73]: <matplotlib.collections.PathCollection at 0x7fe94a1155e0>



[74]:
```
# k-means ist NUR für LINEARE Clustergrenzen beschränkt!
```

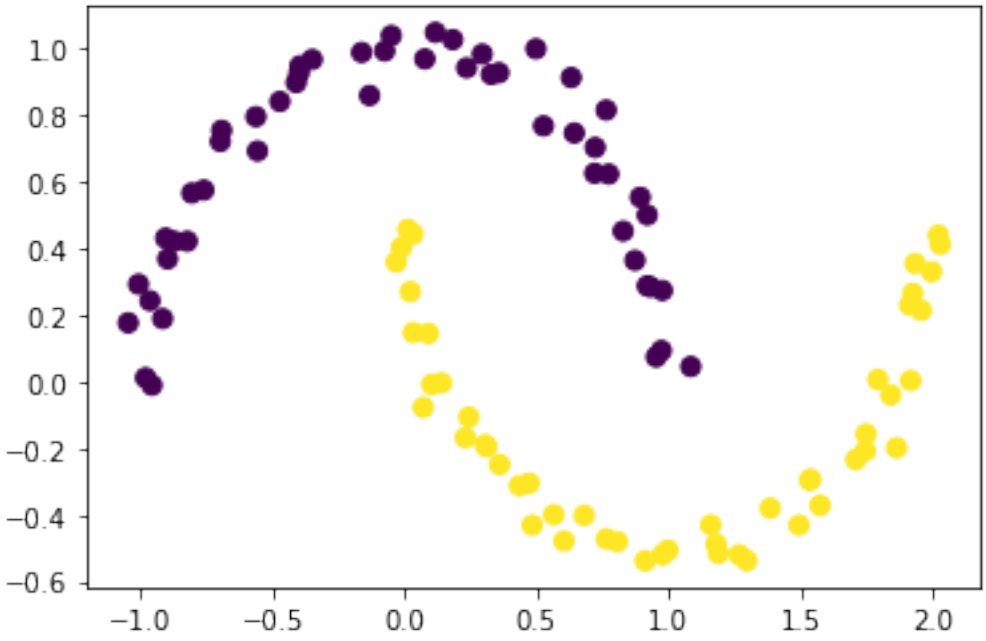
[75]:
```
from sklearn.datasets import make_moons
X,y = make_moons(100, noise=0.05, random_state=0)

plt.scatter(X[:,0], X[:,1], s=50, cmap='viridis')
```

[75]: <matplotlib.collections.PathCollection at 0x7fe9437135e0>



[76]:
```
# Das Clustering funktioniert nur mit einer Version des SkLearns "Spectral␣
↪Clustering"

from sklearn.cluster import SpectralClustering

model = SpectralClustering(n_clusters=2, affinity='nearest_neighbors',
                           assign_labels='kmeans')

labels = model.fit_predict(X)

plt.scatter(X[:,0], X[:,1], c=labels, s=50, cmap='viridis')
```

[76]:
```
<matplotlib.collections.PathCollection at 0x7fe94a279d90>
```


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

[ ]: