Big Data Analytics - 17CS82

Assignment 3 - Two Clusters

Name: Amogh R

USN: 1KS17CS005

Phone: 9972875490

E-mail: amoghpavan5363@gmail.com (mailto:amoghpavan5363@gmail.com)

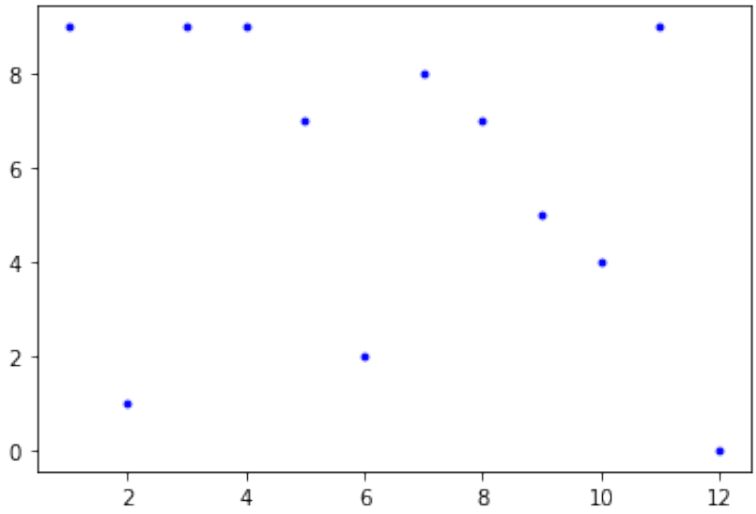
Input

Points	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
X	1	2	3	4	5	6	7	8	9	10	11	12
Y	9	1	9	9	7	2	8	7	5	4	9	0

In [1]:
```
# DRAWING A GRAPH FOR THE ABOVE TABLE
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np

x=[1,2,3,4,5,6,7,8,9,10,11,12]
y=[9,1,9,9,7,2,8,7,5,4,9,0]

plt.scatter(x,y,color='blue',s=8)
data=np.array(list(zip(x,y)))
```



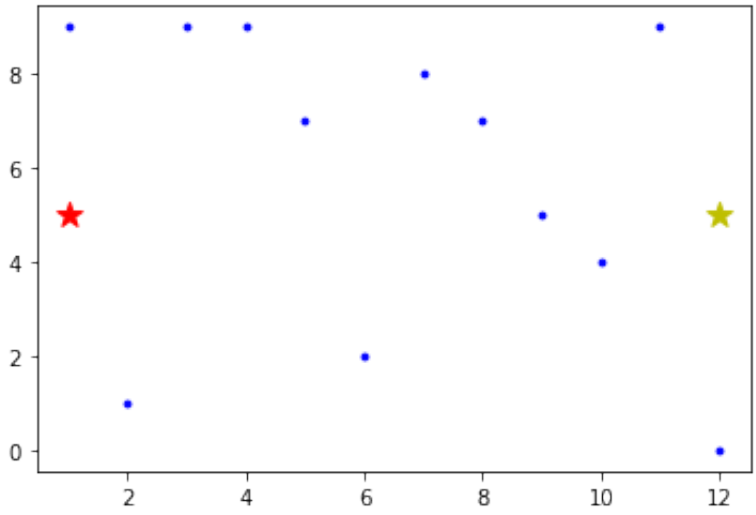
EUCLIDEAN DISTANCE = SQRT [(Xi - Xc) ^ 2 + (Yi - Yc) ^ 2]
Xi = Observed value of X from table
Xc = Centroid value of X
Yi = Observed value of Y from table
Yc = Centroid value of Y

New Centroid Formula = [Sum (Respective Cluster Points (of K1 or K2) + Current Centroid) / Count (Cluster Points + 1)]

In [2]:
```
# FUNCTION FOR CALCULATING EUCLIDEAN DISTANCE
import math
def calculateEuclideanDistance(xi, yi, xc, yc):
    return math.sqrt(math.pow((xi-xc),2) + math.pow((yi-yc),2))
```

Given Initial Centroids are:
C1 = (1,5) which will be represented by RED star
C2 = (12,5) which will be represented by YELLOW star

In [3]:
```
# PLOTTING THE DATAPOINTS ALONG WITH CENTROIDS
plt.scatter(x,y,color='blue',s=8)
plt.scatter(1, 5, s=150, c='r', marker='*')
plt.scatter(12, 5, s=150, c='y', marker='*');
```



ITERATION 1

Centroid 1, C1 = (1,5)
Centroid 2, C2 = (12,5)

In [4]:
```
# EUCLIDEAN DISTANCE FOR POINT P1=(1, 9), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 1, 9)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 1, 9)
print(E2)
```

```
4.0
11.704699910719626
```

Since **E1 < E2**, P1(1,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9)}
K2 = {}

In [5]:
```
# EUCLIDEAN DISTANCE FOR POINT P2=(2, 1), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 2, 1)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 2, 1)
print(E2)
```

```
4.123105625617661
10.770329614269007
```

Since **E1 < E2**, P2(2,1) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1)}
K2 = {}

In [6]:
```
# EUCLIDEAN DISTANCE FOR POINT P3=(3, 9), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 3, 9)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 3, 9)
print(E2)
```

```
4.47213595499958
9.848857801796104
```

Since **E1 < E2**, P3(3,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9)}
K2 = {}

In [7]:
```
# EUCLIDEAN DISTANCE FOR POINT P4=(4, 9), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 4, 9)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 4, 9)
print(E2)
```

```
5.0
8.94427190999916
```

Since **E1 < E2**, P4(4,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9)}
K2 = {}

In [8]:
```
# EUCLIDEAN DISTANCE FOR POINT P5=(5, 7), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 5, 7)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 5, 7)
print(E2)
```

```
4.47213595499958
7.280109889280518
```

Since **E1 < E2**, P5(4,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7)}
K2 = {}

In [9]:
```
# EUCLIDEAN DISTANCE FOR POINT P6=(6, 2), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 6, 2)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 6, 2)
print(E2)
```

```
5.830951894845301
6.708203932499369
```

Since **E1 < E2**, P6(4,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {}

In [10]:
```
# EUCLIDEAN DISTANCE FOR POINT P7=(7, 8), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 7, 8)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 7, 8)
print(E2)
```

```
6.708203932499369
5.830951894845301
```

Since **E2 < E1**, P7(7,8) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8)}

In [11]:
```
# EUCLIDEAN DISTANCE FOR POINT P8=(8, 7), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 8, 7)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 8, 7)
print(E2)
```

```
7.280109889280518
4.47213595499958
```

Since **E2 < E1**, P8(6,2) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7)}

In [12]:
```
# EUCLIDEAN DISTANCE FOR POINT P9=(9, 5), CENTROIDS=(1,5) AND (12,5
E1 = calculateEuclideanDistance(1, 5, 9, 5)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 9, 5)
print(E2)
```

```
8.0
3.0
```

Since **E2 < E1**, P9(9,5) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5)}

In [13]:
```
# EUCLIDEAN DISTANCE FOR POINT P10=(10, 4), CENTROIDS=(1,5) AND (12
E1 = calculateEuclideanDistance(1, 5, 10, 4)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 10, 4)
print(E2)
```

```
9.055385138137417
2.23606797749979
```

Since **E2 < E1**, P10(10,4) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4)}

In [14]:
```
# EUCLIDEAN DISTANCE FOR POINT P11=(11, 9), CENTROIDS=(1,5) AND (12
E1 = calculateEuclideanDistance(1, 5, 11, 9)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 11, 9)
print(E2)
```

```
10.770329614269007
4.123105625617661
```

Since **E2 < E1**, P11(11,9) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4), P11(11,9)}

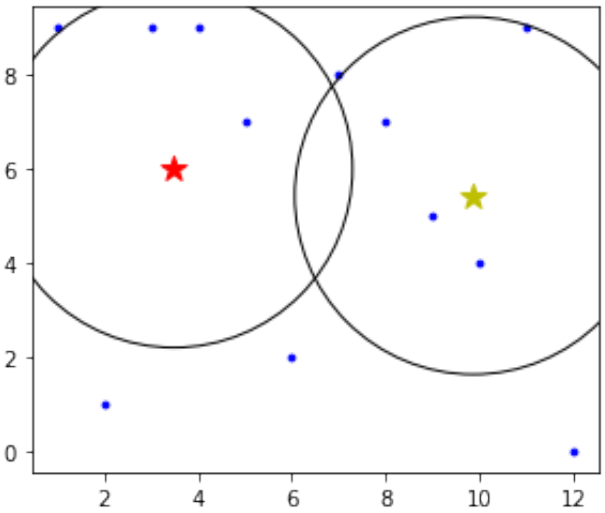
In [15]:
```
# EUCLIDEAN DISTANCE FOR POINT P12=(12, 0), CENTROIDS=(1,5) AND (12
E1 = calculateEuclideanDistance(1, 5, 12, 0)
print(E1)
E2 = calculateEuclideanDistance(12, 5, 12, 0)
print(E2)
```

```
12.083045973594572
5.0
```

Since **E2 < E1**, P12(12,0) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4), P11(11,9), P12(12,0)}

New Centroid Formula = [Sum (Respective Cluster Points (of K1 or K2) + Current Centroid) / Count (Cluster Points + 1)]
Centroid C1(1,5) becomes ((1+2+3+4+5+6+1)/7,(9+1+9+9+7+2+5)/7) => **new C1 is (3.14,6)**
Centroid C2(12,5) becomes ((7+8+9+10+11+12+12)/7,(8+7+5+4+9+0+5)/7) => **new C2 is (9.86,5.43)**

In [16]:
```
# PLOTTING THE DATAPOINTS ALONG WITH NEW CENTROIDS
figure, axes = plt.subplots()
circle1 = plt.Circle((3.49, 6), 3.8, fill = False)
circle2 = plt.Circle((9.86, 5.43), 3.8, fill = False)
axes.set_aspect(1)
axes.add_artist(circle1)
axes.add_artist(circle2)
plt.scatter(x,y,color='blue',s=8)
plt.scatter(3.49, 6, s=150, c='r', marker='*')
plt.scatter(9.86, 5.43, s=150, c='y', marker='*')
plt.show()
```



ITERATION 2

Centroid 1, C1 = (3.14,6)
Centroid 2, C2 = (9.86,5.43)

In [17]:
```
# EUCLIDEAN DISTANCE FOR POINT P1=(1, 9), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 1, 9)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 1, 9)
print(E2)
```

```
3.6850508816025864
9.552198699775879
```

Since **E1 < E2**, P1(1,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9)}
K2 = {}

In [18]:
```
# EUCLIDEAN DISTANCE FOR POINT P2=(2, 1), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 2, 1)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 2, 1)
print(E2)
```

```
5.12831356295615
9.022444236458321
```

Since **E1 < E2**, P2(2,1) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1)}
K2 = {}

In [19]:
```
# EUCLIDEAN DISTANCE FOR POINT P3=(3, 9), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 3, 9)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 3, 9)
print(E2)
```

```
3.0032648900821255
7.733336925286522
```

Since **E1 < E2**, P3(3,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9)}
K2 = {}

In [20]:
```
# EUCLIDEAN DISTANCE FOR POINT P4=(4, 9), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 4, 9)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 4, 9)
print(E2)
```

```
3.120833222073874
6.861814628799002
```

Since **E1 < E2**, P4(4,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9),P4(4,9)}
K2 = {}

In [21]:
```
# EUCLIDEAN DISTANCE FOR POINT P5=(5, 7), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 5, 7)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 5, 7)
print(E2)
```

```
2.1117765033260505
5.107298698920986
```

Since **E1 < E2**, P5(5,7) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9),P4(4,9),P5(5,7)}
K2 = {}

In [22]:
```
# EUCLIDEAN DISTANCE FOR POINT P6=(6, 2), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 6, 2)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 6, 2)
print(E2)
```

```
4.9172756684977506
5.163768004083839
```

Since **E1 < E2**, P6(6,2) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9),P4(4,9),P5(5,7),P6(6,2)}
K2 = {}

In [23]:
```
# EUCLIDEAN DISTANCE FOR POINT P7=(7, 8), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 7, 8)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 7, 8)
print(E2)
```

```
4.347367019242797
3.845061768034422
```

Since **E2 < E1**, P7(7,8) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8)}

In [24]:
```
# EUCLIDEAN DISTANCE FOR POINT P8=(8, 7), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 8, 7)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 8, 7)
print(E2)
```

```
4.961814184348301
2.4340295807569796
```

Since **E2 < E1**, P8(8,7) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7)}

In [25]:
```
# EUCLIDEAN DISTANCE FOR POINT P9=(9, 5), CENTROIDS=(3.14,6) AND (9
E1 = calculateEuclideanDistance(3.14, 6, 9, 5)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 9, 5)
print(E2)
```

```
5.9447119358300275
0.961509230324909
```

Since **E2 < E1**, P9(9,5) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5)}

In [26]:
```
# EUCLIDEAN DISTANCE FOR POINT P10=(10, 4), CENTROIDS=(3.14,6) AND
E1 = calculateEuclideanDistance(3.14, 6, 10, 4)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 10, 4)
print(E2)
```

```
7.145600044782802
1.4368368035375483
```

Since **E2 < E1**, P10(10,4) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4)}

In [27]:
```
# EUCLIDEAN DISTANCE FOR POINT P11=(11, 9), CENTROIDS=(3.14,6) AND
E1 = calculateEuclideanDistance(3.14, 6, 11, 9)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 11, 9)
print(E2)
```

```
8.413061273995334
3.7475992315080866
```

Since **E2 < E1**, P11(11,9) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4), P11(11,9)}

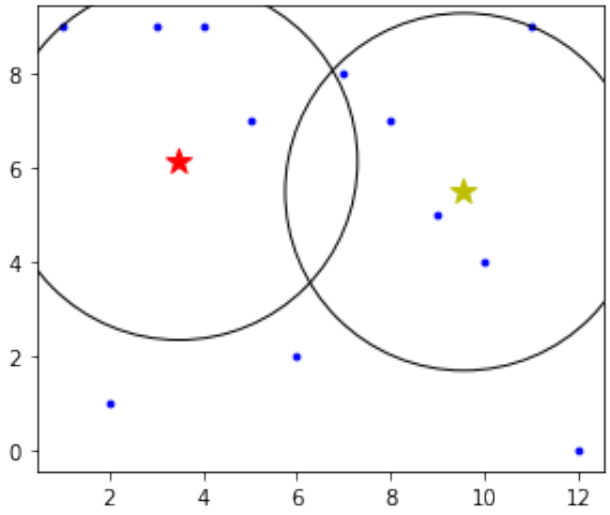
In [28]:
```
# EUCLIDEAN DISTANCE FOR POINT P12=(12, 0), CENTROIDS=(3.14,6) AND
E1 = calculateEuclideanDistance(3.14, 6, 12, 0)
print(E1)
E2 = calculateEuclideanDistance(9.86, 5.43, 12, 0)
print(E2)
```

```
10.700448588727483
5.836480103624101
```

Since **E2 < E1**, P12(12,0) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4), P11(11,9), P12(12,0)}

New Centroid Formula = [Sum (Respective Cluster Points (of K1 or K2) + Current Centroid) / Count (Cluster Points + 1)]
Centroid C1(1,5) becomes ((1+2+3+4+5+6+3.14)/7,(9+1+9+9+7+2+6)/7) => **new C1 is (3.49,6.14)**
Centroid C2(12,5) becomes ((7+8+9+10+11+12+9.86)/7,(8+7+5+4+9+0+5.43)/7) => **new C2 is (9.55,5.49)**

In [29]:
```
# PLOTTING THE DATAPOINTS ALONG WITH NEW CENTROIDS
figure, axes = plt.subplots()
circle1 = plt.Circle((3.49, 6.14), 3.8, fill = False)
circle2 = plt.Circle((9.55, 5.49), 3.8, fill = False)
axes.set_aspect(1)
axes.add_artist(circle1)
axes.add_artist(circle2)
plt.scatter(x,y,color='blue',s=8)
plt.scatter(3.49, 6.14, s=150, c='r', marker='*')
plt.scatter(9.55, 5.49, s=150, c='y', marker='*')
plt.show()
```



ITERATION 3

Centroid 1, C1 = (3.49,6.14)
Centroid 2, C2 = (9.55,5.49)

In [30]:
```
# EUCLIDEAN DISTANCE FOR POINT P1=(1, 9), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 1, 9)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 1, 9)
print(E2)
```

```
3.792057489015693
9.242434744156975
```

Since **E1 < E2**, P1(1,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9)}
K2 = {}

In [31]:
```
# EUCLIDEAN DISTANCE FOR POINT P2=(2, 1), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 2, 1)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 2, 1)
print(E2)
```

```
5.351607235214482
8.784224496220483
```

Since **E1 < E2**, P2(2,1) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1)}
K2 = {}

In [32]:
```
# EUCLIDEAN DISTANCE FOR POINT P3=(3, 9), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 3, 9)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 3, 9)
print(E2)
```

```
2.9016719318351623
7.431191021633074
```

Since **E1 < E2**, P3(3,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9)}
K2 = {}

In [33]:
```
# EUCLIDEAN DISTANCE FOR POINT P4=(4, 9), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 4, 9)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 4, 9)
print(E2)
```

```
2.9051161766786544
6.566780032862377
```

Since **E1 < E2**, P4(4,9) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9),P4(4,9)}
K2 = {}

In [34]:
```
# EUCLIDEAN DISTANCE FOR POINT P5=(5, 7), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 5, 7)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 5, 7)
print(E2)
```

```
1.7377284022539312
4.794017104683713
```

Since **E1 < E2**, P5(5,7) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9),P4(4,9),P5(5,7)}
K2 = {}

In [35]:
```
# EUCLIDEAN DISTANCE FOR POINT P6=(6, 2), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 6, 2)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 6, 2)
print(E2)
```

```
4.841456392450519
4.9782125306177925
```

Since **E1 < E2**, P6(6,2) will become part of cluster K1 representing C1 as centroid
Therefore,
K1 = {P1(1,9),P2(2,1),P3(3,9),P4(4,9),P5(5,7),P6(6,2)}
K2 = {}

In [36]:
```
# EUCLIDEAN DISTANCE FOR POINT P7=(7, 8), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 7, 8)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 7, 8)
print(E2)
```

```
3.9723670525267427
3.578072106595953
```

Since **E2 < E1**, P7(7,8) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8)}

In [37]:
```
# EUCLIDEAN DISTANCE FOR POINT P8=(8, 7), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 8, 7)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 8, 7)
print(E2)
```

```
4.59126344267022
2.163931607052312
```

Since **E2 < E1**, P8(8,7) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7)}

In [38]:
```
# EUCLIDEAN DISTANCE FOR POINT P9=(9, 5), CENTROIDS=(3.49,6.14) AND
E1 = calculateEuclideanDistance(3.49, 6.14, 9, 5)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 9, 5)
print(E2)
```

```
5.626695300085122
0.7366138744281165
```

Since **E2 < E1**, P9(9,5) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5)}

In [39]:
```
# EUCLIDEAN DISTANCE FOR POINT P10=(10, 4), CENTROIDS=(3.49,6.14) A
E1 = calculateEuclideanDistance(3.49, 6.14, 10, 4)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 10, 4)
print(E2)
```

```
6.852714790504563
1.5564703659241317
```

Since **E2 < E1**, P10(10,4) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4)}

In [40]:
```
# EUCLIDEAN DISTANCE FOR POINT P11=(11, 9), CENTROIDS=(3.49,6.14) A
E1 = calculateEuclideanDistance(3.49, 6.14, 11, 9)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 11, 9)
print(E2)
```

```
8.036149575511894
3.7977098362039188
```

Since **E2 < E1**, P11(11,9) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4), P11(11,9)}

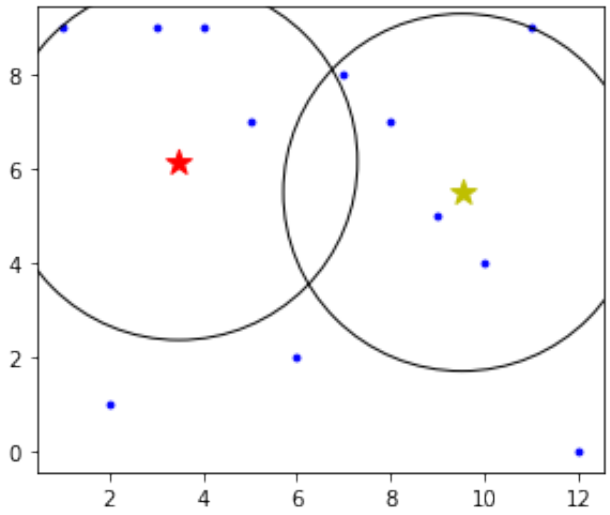
In [41]:
```
# EUCLIDEAN DISTANCE FOR POINT P12=(12, 0), CENTROIDS=(3.49,6.14) A
E1 = calculateEuclideanDistance(3.49, 6.14, 12, 0)
print(E1)
E2 = calculateEuclideanDistance(9.55, 5.49, 12, 0)
print(E2)
```

```
10.49379340372203
6.01187158878165
```

Since **E2 < E1**, P12(12,0) will become part of cluster K2 representing C2 as centroid
Therefore,
K1 = {P1(1,9), P2(2,1), P3(3,9), P4(4,9), P5(5,7), P6(6,2)}
K2 = {P7(7,8), P8(8,7), P9(9,5), P10(10,4), P11(11,9), P12(12,0)}

New Centroid Formula = [Sum (Respective Cluster Points (of K1 or K2) + Current Centroid) / Count (Cluster Points + 1)]
Centroid C1(1,5) becomes ((1+2+3+4+5+6+3.49)/7,(9+1+9+9+7+2+6.14)/7) => **new C1 is (3.49,6.16)**
Centroid C2(12,5) becomes ((7+8+9+10+11+12+9.55)/7,(8+7+5+4+9+0+5.49)/7) => **new C2 is (9.51,5.50)**

In [42]:
```
# PLOTTING THE DATAPOINTS ALONG WITH NEW CENTROIDS
figure, axes = plt.subplots()
circle1 = plt.Circle((3.49, 6.16), 3.8, fill = False)
circle2 = plt.Circle((9.51, 5.50), 3.8, fill = False)
axes.set_aspect(1)
axes.add_artist(circle1)
axes.add_artist(circle2)
plt.scatter(x,y,color='blue',s=8)
plt.scatter(3.49, 6.16, s=150, c='r', marker='*')
plt.scatter(9.51, 5.50, s=150, c='y', marker='*')
plt.show()
```



Initial Centroids

C1 = (1,5)
C2 = (12,5)

Centroids after Iteration 1

C1 = (3.49,6.14)
C2 = (9.55,5.49)

Centroids after Iteration 2

C1 = (3.49,6.16)
C2 = (9.51,5.50)

We see that the centroids are stabilising at this point, hence the iterations have been stopped here.