Part a

```
import numpy as np
import seaborn as sns
import matplotlib.pylab as plt
import pandas as pd

samples=np.random.multinomial(200, [0.5,0.15,0.35], size=1)[0]

print(samples)
a = np.array([1,0,0])


b = np.array([0,1,0])

c=np.array([0,0,1])
for i in range(samples.shape[0]):
  if i==0:
    a=np.tile(a,(samples[i],1))
  if i==1:
    b=np.tile(b,(samples[i],1))
  if i==2:
    c=np.tile(c,(samples[i],1))

theta=np.concatenate((a, b,c), axis=0)
```

```
[105  28  67]
```

```
print("Expected number of members of first cluster," ,200*0.5)
print("Expected number of members of second cluster,",200*0.15)
print("Expected number of members of third cluster,",200*0.35)
```

```
Expected number of members of first cluster, 100.0
Expected number of members of second cluster, 30.0
Expected number of members of third cluster, 70.0
```

```
plt.figure(figsize=(10,10))
heat_map = sns.heatmap( theta,cmap="Greys",xticklabels=False,yticklabels=False)

#heat_map = sns.heatmap( theta,cmap=sns.cubehelix_palette(as_cmap=True))
plt.title( "Community membership ${\Theta}$ matrix" )
plt.show()
```

Answer

What is the expected number of members for each community?

Expected number of members of first cluster, 100.0

Expected number of members of second cluster, 30.0

Expected number of members of third cluster, 70.0

For your instance of Θ, what is your observed number of members for each community?

Observed number of members of first cluster, 105

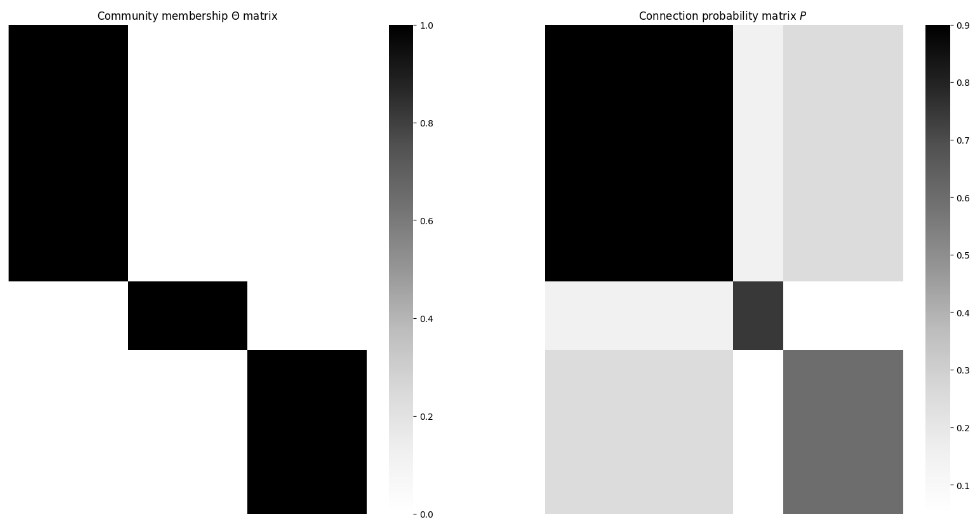
Observed number of members of second cluster, 28

Observed number of members of third cluster, 67

Part b

```
theta_T=theta.transpose()
B=[[0.9,0.15,0.25],[0.15,0.75,0.05],[0.25,0.05,0.6]]
P=theta.dot(B).dot(theta_T)
```

```
fig, (ax1, ax2) = plt.subplots(1, 2,figsize=(20, 10))

heat_map1 = sns.heatmap(theta,cmap="Greys",ax=ax1,xticklabels=False,yticklabels=False)
ax1.title.set_text('Community membership ${\Theta}$ matrix ')


heat_map2 = sns.heatmap( P,cmap="Greys",ax=ax2,xticklabels=False,yticklabels=False)
ax2.title.set_text('Connection probability matrix $P$  ')
```




```
print("P[1,3]=",P[1,3])
print("P[197,199]=",P[197,199])
print("P[192,4]=",P[192,4])
```

```
P[1,3]= 0.9
P[197,199]= 0.6
P[192,4]= 0.25
```

Answer

P[1,3]= 0.9

P[197,199]= 0.6

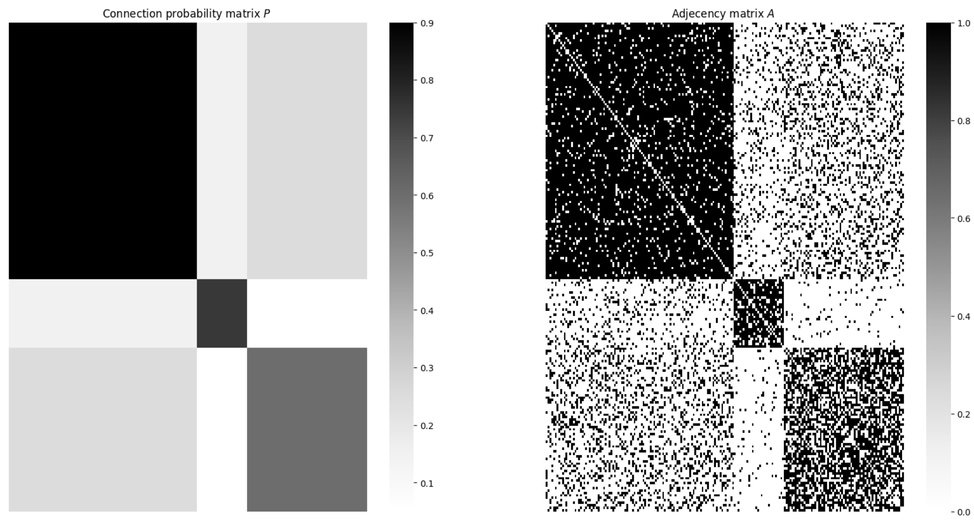
P[192,4]= 0.25

Part c

```
R=np.random.uniform(size=[P.shape[0],P.shape[0]])
R1=np.triu(R)+np.transpose(np.triu(R))
A=1*(R1<P)
A=A-np.diag(np.diag(A))
```

```
fig, (ax1, ax2) = plt.subplots(1, 2,figsize=(20, 10))

heat_map1 = sns.heatmap(P,cmap="Greys",ax=ax1,xticklabels=False,yticklabels=False)
ax1.title.set_text('Connection probability matrix $P$  ')


heat_map2 = sns.heatmap( A,cmap="Greys",ax=ax2,xticklabels=False,yticklabels=False)
```

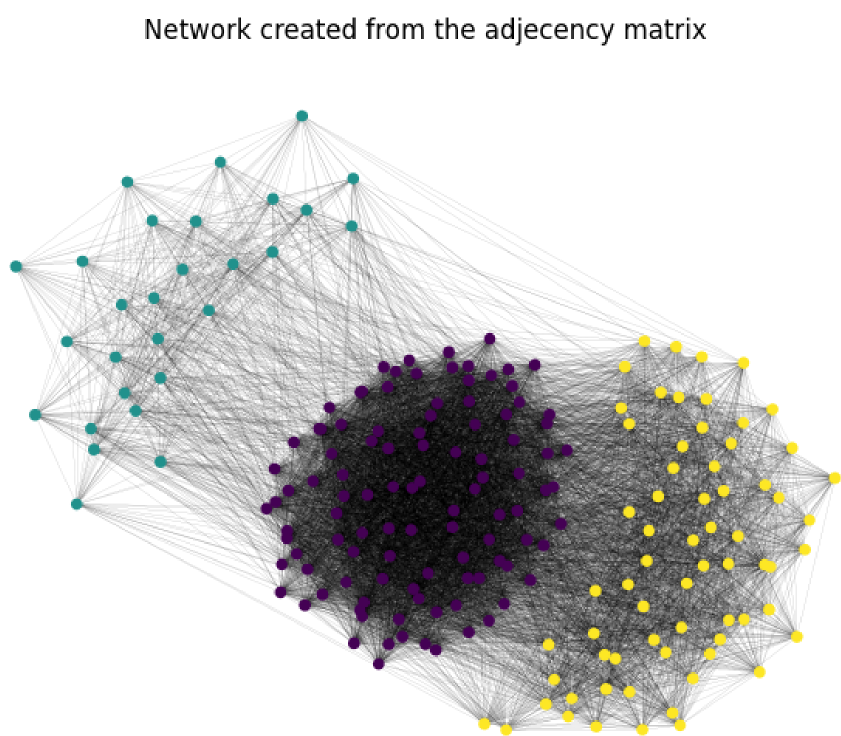
```
ax2.title.set_text('Adjecency matrix $A$  ')
```



Part d

```
import networkx as nx

network=nx.from_numpy_array(A)


import matplotlib.cm as cmx
import matplotlib.colors as colors

np.random.seed(1)
colors1_=[0]*a.shape[0]
colors2_=[1]*b.shape[0]
colors3_=[2]*c.shape[0]
colors_=np.concatenate((colors1_, colors2_,colors3_), axis=0)
print(colors_)
f = plt.figure(1)
#ax = f.add_subplot(1,1,1)
#cNorm  = colors.Normalize(vmin=0, vmax=max(values))

#scalarMap = cmx.ScalarMappable(norm=cNorm, cmap=jet)
#ColorLegend = {'red': 0,'blue': 1,'yellow': 2}
nx.draw(network,node_color=colors_,node_size=20,width=0.05)
plt.title( "Network created from the adjecency matrix" )

#for label in ColorLegend:
#    ax.plot([0],[0],color=scalarMap.to_rgba(ColorLegend[label]),label=label)
```

```
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2]
Text(0.5, 1.0, 'Network created from the adjecency matrix')
```



Answer

How do the connection densities of the different communities compare within communities? **The purple community is extremely dense connected within nodes of the community. Then follows the yellow community with also a dense within connections (but not as dense as the purple. Finally the green community, which is not as dense as the other two communities.**

How do the connection densities of the different communities compare across communities? **The purple community is the most connected across other communities than the other two communities. The yellow and green community are NOT very connected between them. On the other hand, the yellow community is more connected across other communities than the green community. The yellow community has many connections to the purple community. The green community is the less connected across other communities.**

+ Code + Text

✓ 0s completed at 12:55 PM