1 Question 1

1.1 Number of edges :

First, we want to compute the number of edge in G, we will denote G_1 the complete graph component, and G_2 the bipartite graph component.

To compute the number of edge in G_1 we can try a few example with a complete graph of 3 nodes, then 4 nodes, then 5.. From here, we can deduce a general formula (which can be prooved by induction) :

$$n_{edges} = \frac{n_{nodes}(n_{nodes} - 1)}{2}$$

So we get for our graph G_1 : $n_{edges} = 100 \times 99/2 = 4950$

Now we can compute edges in the graph G_2. As it is a bipartite graph, we now that each node of the first part is connected to all nodes of the second part. Immediatly, we compute for G_2 : $n_{edges} = 50 \times 50 = 2500$

From the previous 2 calculations, we deduce that the number of edges in G is 7450

1.2 Number of triangle :

The number of triangles in G_1 is a counting problem. We are looking for the number of nodes that can be grouped by 3, which can be calculated using the formula :

$$\binom{100}{3} = 161700$$

For G_2, as it is a bipartite graph, there is no triangle.

Thus, the number of triangles in G is 161700.

2 Question 2

Let's compute the modularity for the first graph (a) :

$$Q_a = \left[\frac{6}{13} - (\frac{13}{2*13})^2\right] + \left[\frac{6}{13} - (\frac{13}{2*13})^2\right] \simeq 0.42$$

For the second graph (b) :

$$Q_b = \left[\frac{2}{13} - (\frac{11}{2*13})^2\right] + \left[\frac{4}{13} - (\frac{15}{2*13})^2\right] \simeq -0.05$$

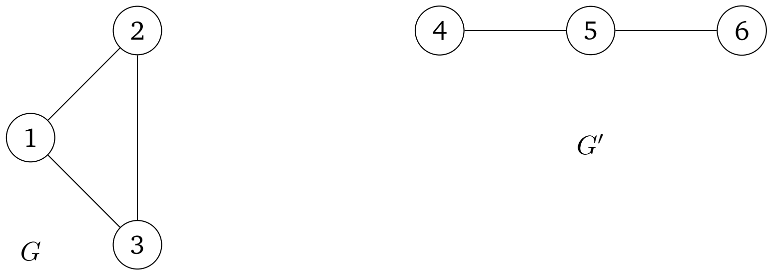
3 Question 3

We have $\phi(P_4) = [3, 2, 1, 0]$ and $\phi(C_4) = [4, 4, 0, 0]$ Thanks to this, we can calculate the shortest path kernel for each pairs :

$$k(C_4, C_4) = \langle\phi(C_4), \phi(C_4)\rangle = 4^2 + 4^2 = 32$$
$$k(C_4, P_4) = \langle\phi(C_4), \phi(P_4)\rangle = 4 \times 3 + 4 \times 2 = 20$$
$$k(P_4, P_4) = \langle\phi(P_4), \phi(P_4)\rangle = 3^2 + 2^2 + 1^2 = 12$$

4 Question 4

The kernel $k(G, G')$ compute the similarity between the 2 graphs G and G' based on graphlets of size 3. Here, if the scalar product of the feature map of each graph is 0, it means that the graphs do not share any common graphlet of size 3.
We can visualize this throught an example.



We can note that this graph are used to create our graphlet kernel in code part 3.