1 Question 1

To calculate the number of edges, we will look at each graph independently.

- In a complete graph, all vertices are connected. For each vertex, there are therefore $n-1$ edges created. We have a total of $(n \times (n-1))/2$ so as not to count the same edges twice.
- In a bipartite graph, we have two subsets of size M and N respectively. The first node of the first subset is connected to the N nodes of the other subset. We have a total of $M \times N$ stops.

We thus have in our case $\frac{(100\times 99)}{2} + 50 \times 50 + 1$ (to link the graphs) stopped.

Regarding the number of triangles, a bipartite graph cannot contain any. And, a complete graph has $\binom{n}{3} = \frac{n!}{3!(n-3)!}$. So we have 161700 triangles.

2 Question 2

$$C = \frac{3 \times |\text{closed triplet}|}{|\text{distinct neighbor pairs}|}$$

Consequently, we have $C \leq 1$. The equality holds if and only if the graph is a set of triangles, which corresponds to a complete graph for a connected graph.

3 Question 3

A vector v is an eigenvector of Laplacian matrix L_{rw} of eigenvalue λ if $L_{rw}v = \lambda v$. The smallest eigenvalue of L_{rw} is 0 corresponding to the eigenvector $\mathbb{1}_N$.

This eigenvalue implies that all nodes end up in the same cluster, which is not interesting for our algorithm. If this eigenvalue is ignored, the result of the clustering is therefore unchanged.

4 Question 4

The result of the Spectral Clustering Algorithm depends on the initializations of KMeans. The eigenvalue decomposition strategy may also depend on its initialization. The result is therefore stochastic by default. It is possible to make this algorithm deterministic by setting a seed in the algorithms.

5 Question 5

In graph a, there is a total $m = 13$ edges:

- The blue cluster forms a community [6, 7, 8, 9] with $l_c = 4$. The degrees of the nodes are respectively [4, 3, 3, 3] so $d_c = 13$. In conclusion, Q = $\frac{4}{13} - (\frac{13}{2\times 13})^2 \approx 0.058$.
- The green cluster forms a community [1, 2, 3, 4, 5] with $l_c = 5$. The degrees of the nodes are respectively [3, 2, 3, 3, 2] so $d_c = 13$. In conclusion, Q = $\frac{5}{13} - (\frac{13}{2\times 13})^2 \approx 0.135$.

For a total of $Q_a \approx 0.192$

In graph b, there is a total $m = 13$ edges:

- The blue cluster forms a community [3, 4, 5, 6, 7] with $l_c = 5$. The degrees of the nodes are respectively [3, 3, 2, 4, 3] so $d_c = 15$. In conclusion, Q = $\frac{5}{13} - (\frac{15}{2\times 13})^2 \approx 0.052$.
- The green cluster forms a community [1, 2, 8, 9] with $l_c = 4$. The degrees of the nodes are respectively [3, 2, 3, 3] so $d_c = 11$. In conclusion, Q = $\frac{4}{13} - (\frac{11}{2\times 13})^2 \approx 0.129$.

For a total of $Q_b \approx 0.180$

6 Question 6

A C_4 is equal to $[4, 2, 0, 0, ...]$ and a P_4 is equal to $[3, 2, 1, 0, ...]$.

So we have :

- $(C_4, C_4) = sum([4^2, 2^2, 0^2, 0^2, ...]) = 20$.
- $(P_4, P_4) = sum([3^2, 2^2, 1^2, 0^2, ...]) = 14$.
- $(C_4, P_4) = sum([3 \times 4, 2 \times 2, 1 \times 0, 0 \times 0, ...]) = 16$.