

Assignment 6: Global convexity (fitness-distance/similarity correlations) tests

Authors: Kiril Andrukh, 162069; Uladzislau Lukashevich, 155671.

Source code: [link](#)

Description of the problem

For each instance generate 1000 random local optima obtained from random solutions using greedy local search. For each solution calculate its similarity either to the best solution (could be the best out of the 1000 local optima or an even better solution generated by another method) or the average similarity to all other local optima. Make charts, x-axis – value of the objective function, y-axis (average) similarity. In the results with similarity to the best solution do not include the best solution (it would be an outlier with 100% similarity to itself).

Use (separately) two measures of similarity:

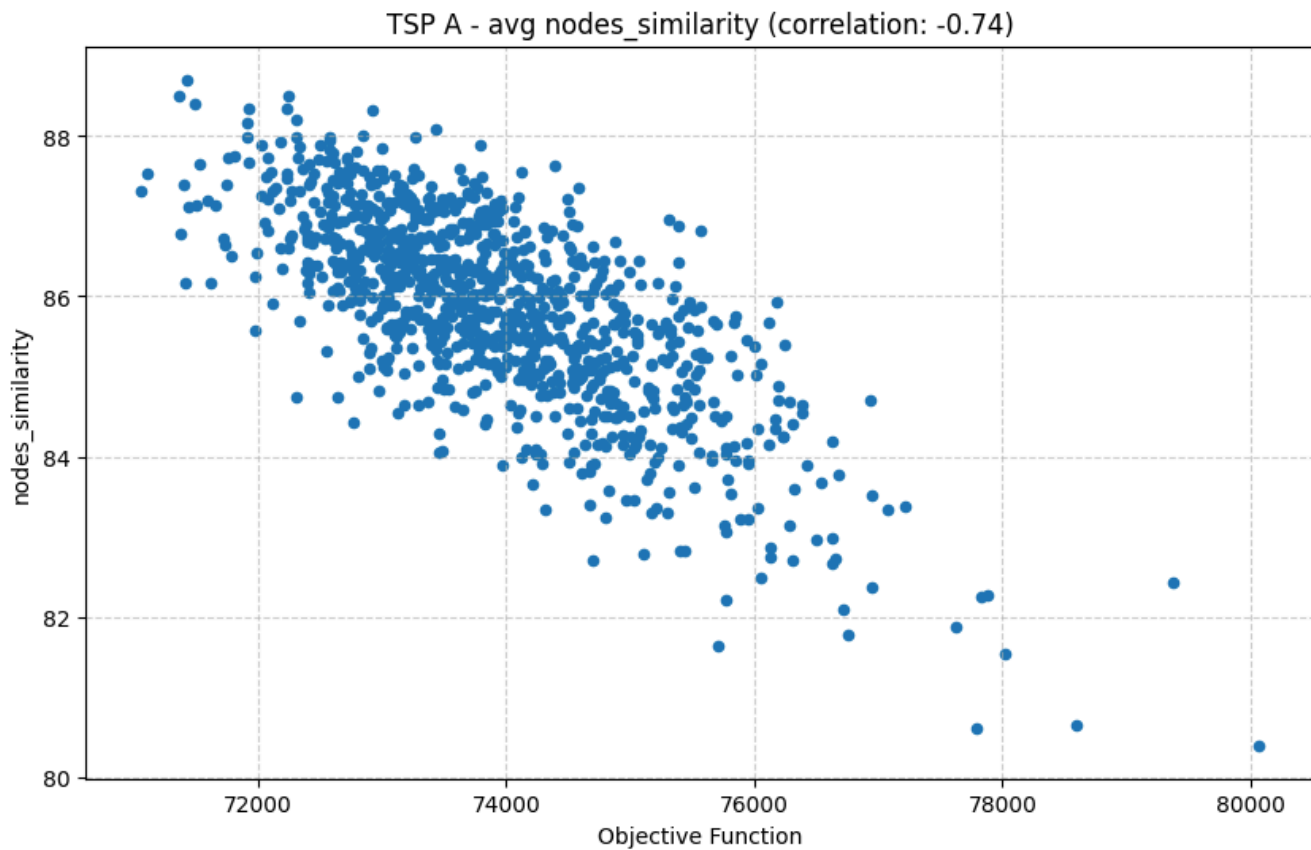
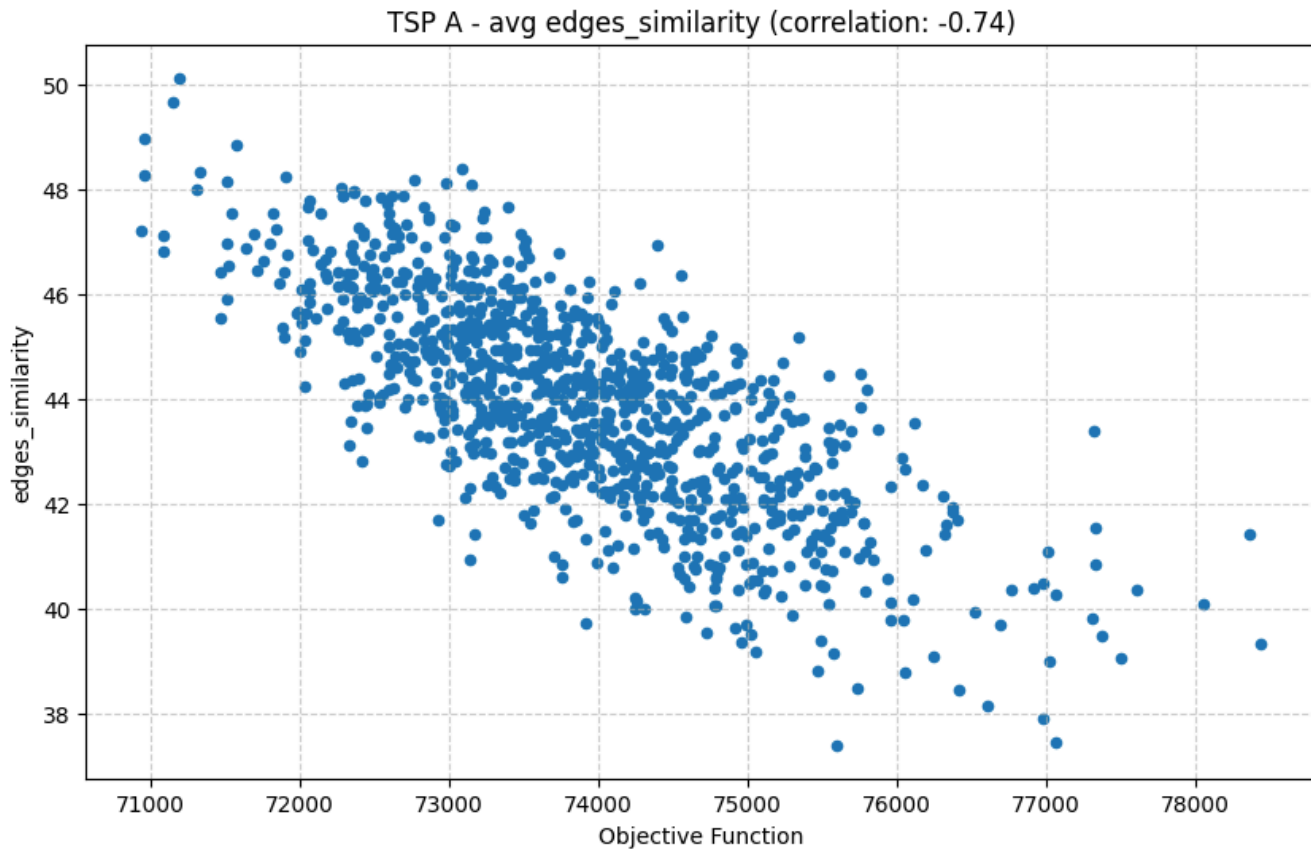
- The number of common edges.
- The number of common selected nodes.

Finally we have 8 charts: 2 instances, 2 versions of similarity (either to the best or average), 2 similarity measures (either common edges or common selected nodes).

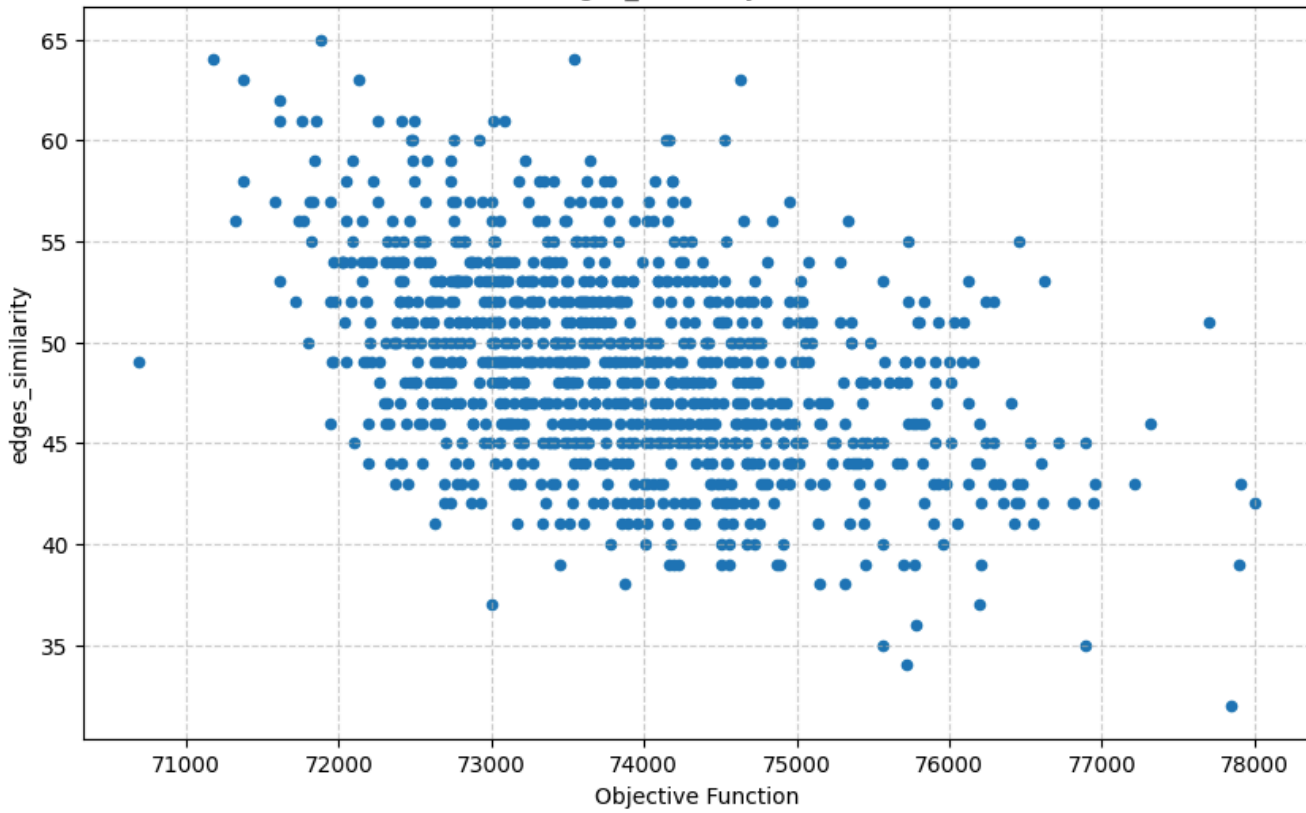
For each chart calculate also the correlation coefficient.

Plots

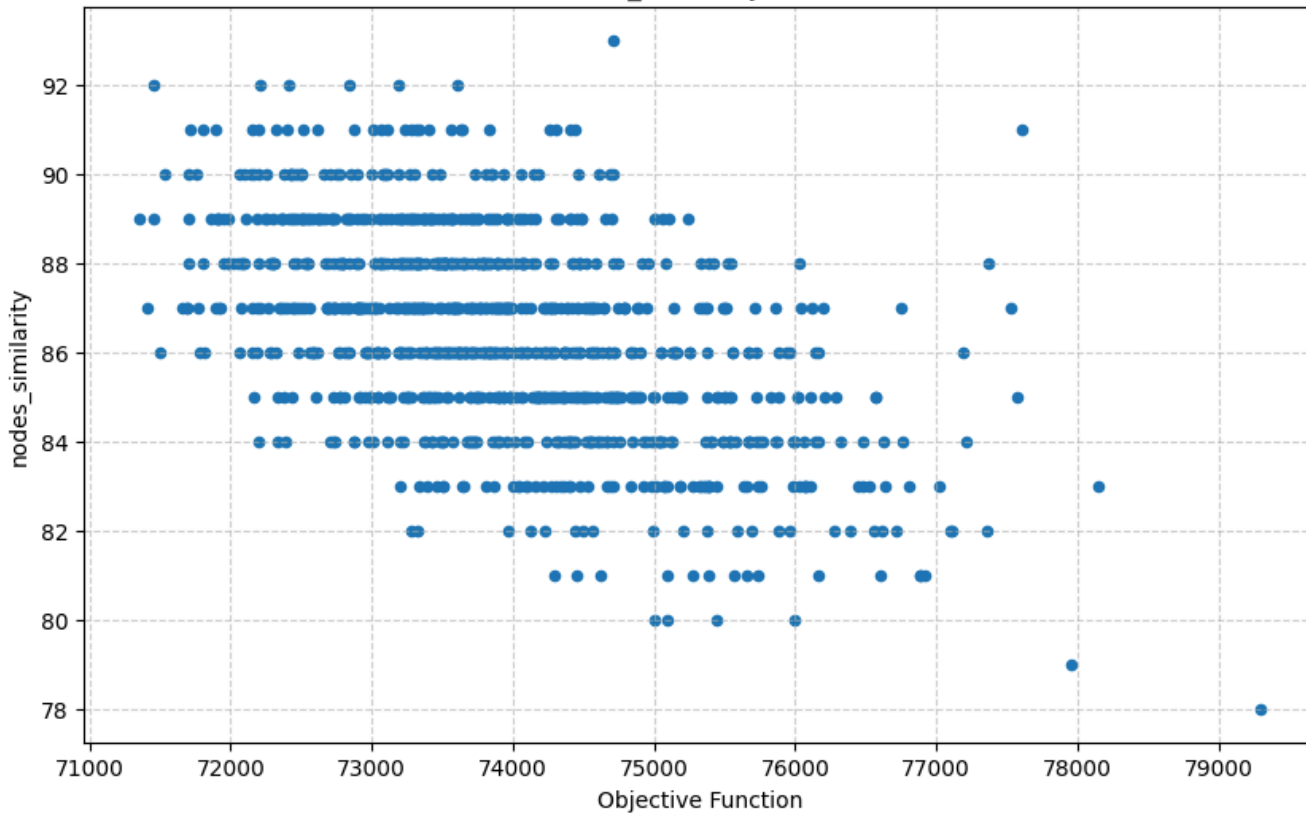
Dataset A:



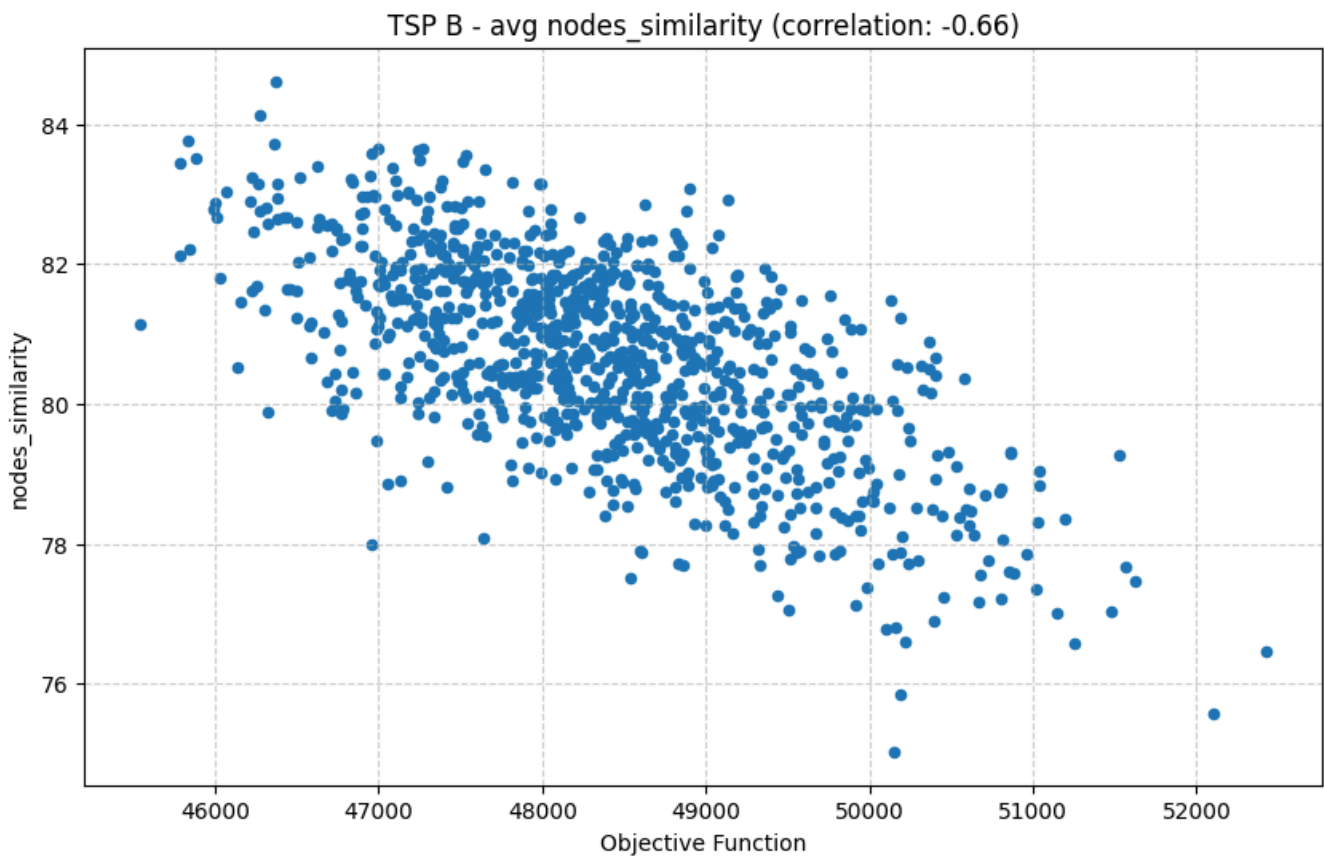
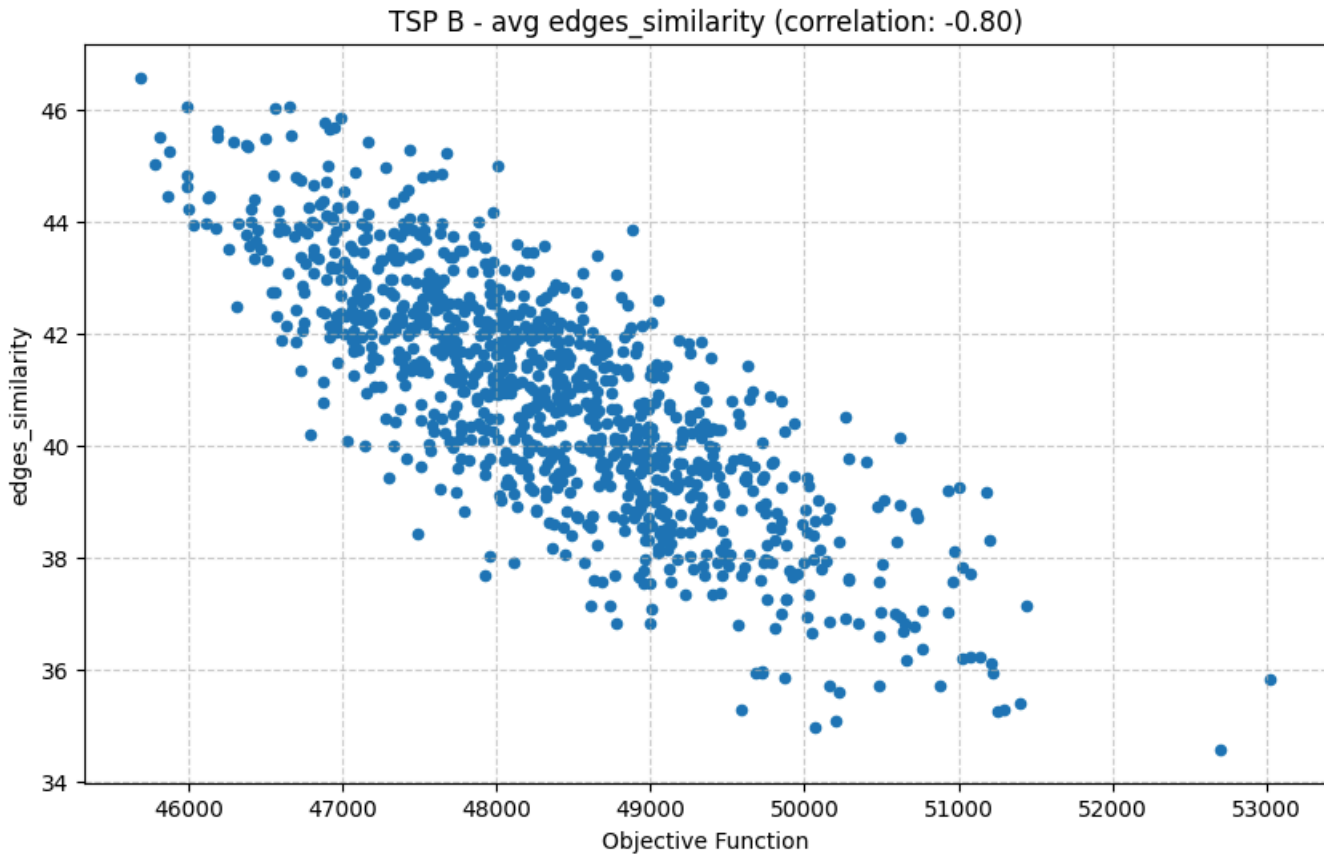
TSP A - best edges_similarity (correlation: -0.44)



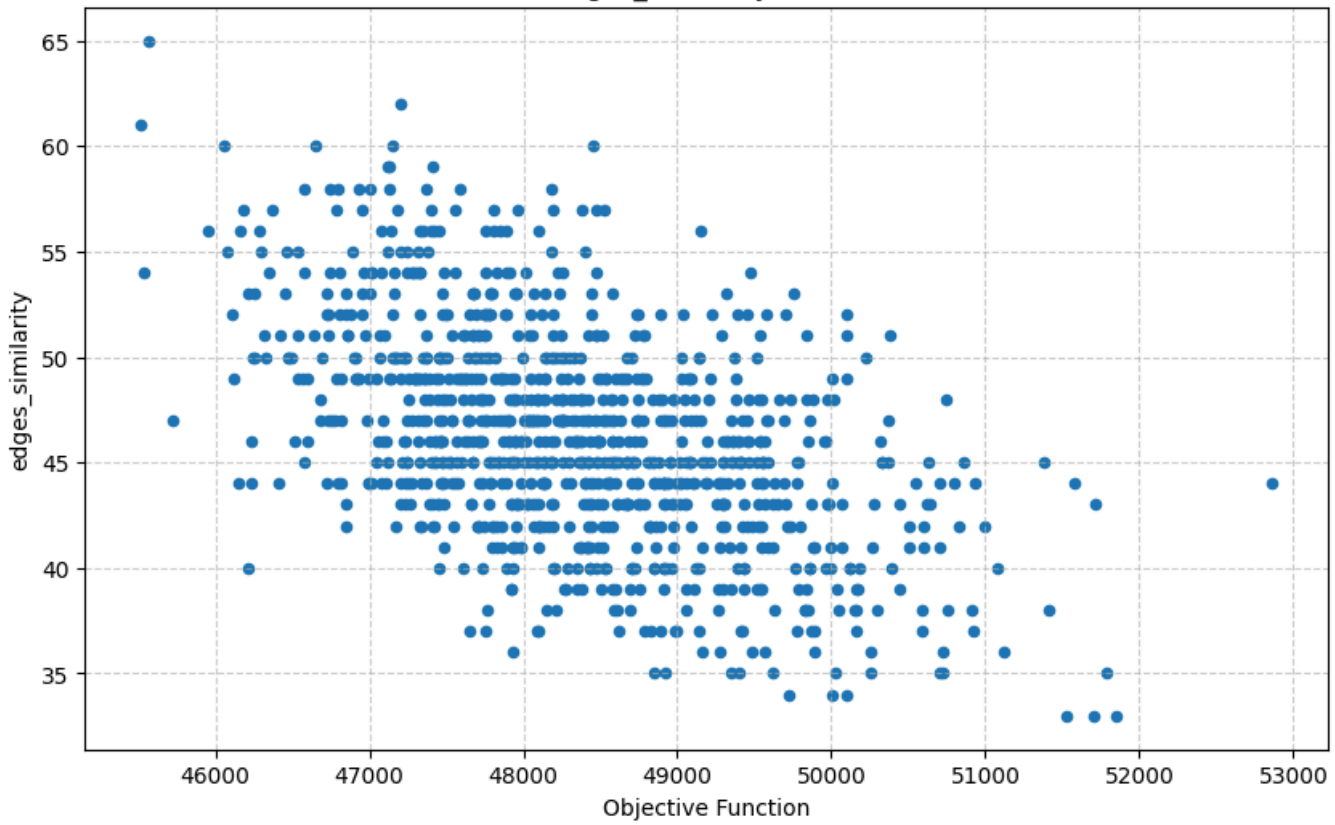
TSP A - best nodes_similarity (correlation: -0.53)



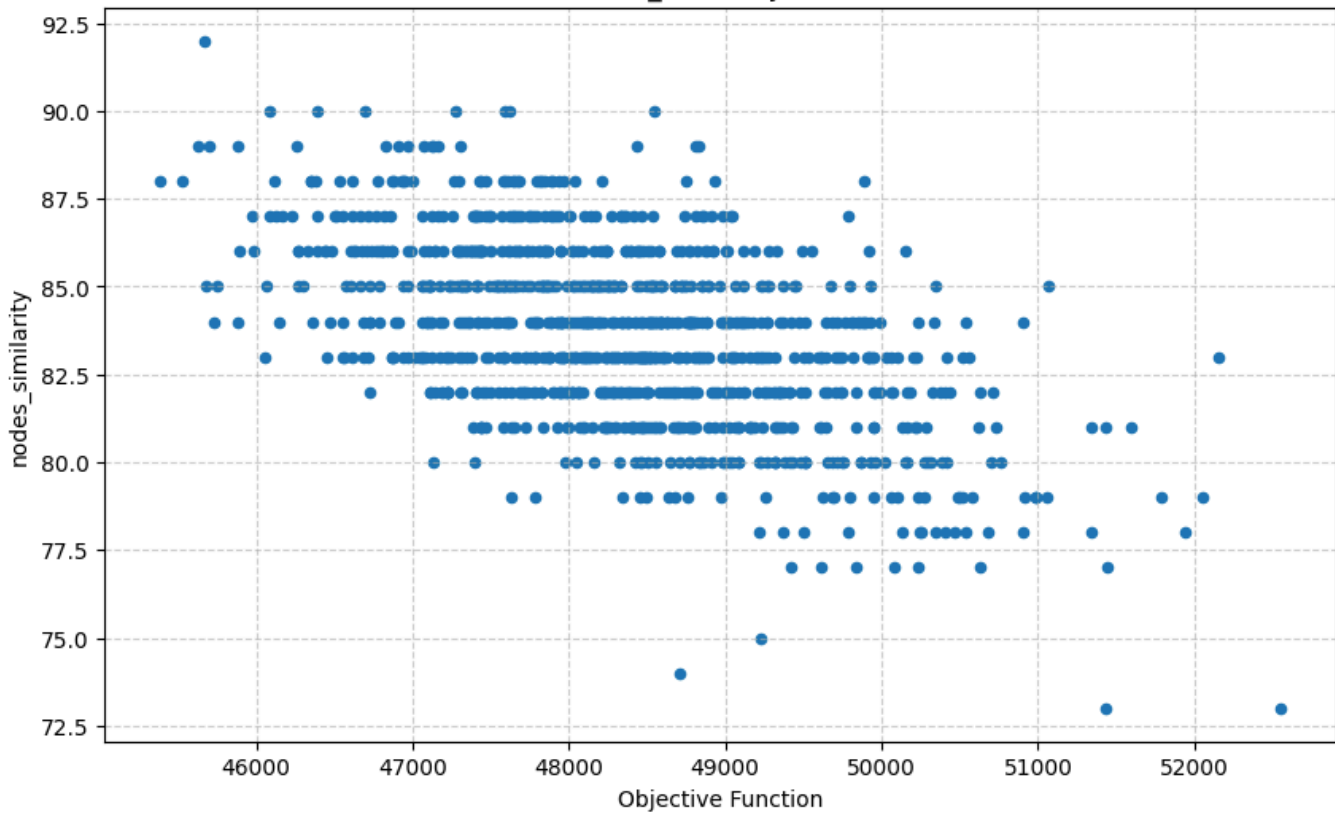
Dataset B:



TSP B - best edges_similarity (correlation: -0.52)



TSP B - best nodes_similarity (correlation: -0.56)



Conclusion:

There are more similarities between nodes than between edges

The correlation between average similarities and objective function is much higher than to the best.
However, the similarities are higher when compared to the best solution, not average.