# Report on Statistical Analysis of Two Implementations

## 1. Hypotheses

**Null Hypothesis (H0):** There is no significant difference in performance between the two implementations when tokens are randomly placed in the environment.

**Alternative Hypothesis (H1):** There is a significant difference in performance between the two implementations when tokens are randomly placed in the environment.

## 2. Experimental Setup

**Objective:** To comprehensively compare the performance of two implementations (mine and my colleague's) in terms of average time required to complete the task, time required to find a box, and time required to deliver a box.

**Variables:**

- Independent Variable: The implementation used (mine vs. my colleague's).
- Dependent Variables:
  - ✓ Average time required to find the box.
  - ✓ Average time required to deliver the box.
  - ✓ Total execution time.

**Experimental Design:**

1. Token Placement: Tokens randomly placed in the environment.
2. Number of Tokens: Varying from 4 to 8 to test scalability and robustness.
3. Number of Repetitions: 5 repetitions for each scenario for statistical significance.
4. Total number of Execution: 25

## 3. Results

The below table highlights the average time taken for to find a box, time to deliver it and the total execution time for both me and my colleague.

| Number of Boxes | Avg. Time Finding Box (Me) | Avg. Time Finding Box (Colleague) | Avg. Time Delivering Box (Me) | Avg. Time Delivering Box (Colleague) | Total Execution Time (Me) | Total Execution Time (Colleague) |
|---|---|---|---|---|---|---|
| 4 | 3.8392 | 0.00108 | 7.1246 | 10.22256 | 74.2912 | 55.40991 |
| 4 | 3.8403 | 0.00075 | 7.2118 | 10.0744 | 73.8848 | 54.21028 |
| 4 | 3.8404 | 0.00058 | 7.0159 | 10.04327 | 72.7981 | 55.32878 |
| 4 | 3.8405 | 0.00068 | 7.1704 | 10.10568 | 71.8546 | 55.40147 |
| 4 | 3.8406 | 0.00073 | 7.0078 | 10.16062 | 73.9217 | 55.32169 |
| 5 | 1.2532 | 0.0008 | 8.6088 | 9.68905 | 80.7349 | 64.59154 |
| 5 | 1.2529 | 0.00079 | 8.3675 | 9.71939 | 81.8155 | 64.75527 |
| 5 | 1.2526 | 0.00076 | 8.0929 | 9.88943 | 81.573 | 63.94536 |
| 5 | 1.2528 | 0.00075 | 7.4248 | 9.71376 | 80.3744 | 65.65486 |
| 5 | 1.2528 | 0.00082 | 8.303 | 9.96104 | 82.717 | 64.96388 |
| 6 | 2.106 | 0.00075 | 7.1894 | 5.67447 | 94.5756 | 93.03451 |
| 6 | 2.1045 | 0.00074 | 7.583 | 5.67247 | 96.7856 | 91.34882 |
| 6 | 2.1038 | 0.00075 | 7.3568 | 5.5669 | 95.151 | 93.26598 |
| 6 | 2.1151 | 0.00079 | 7.6185 | 5.75345 | 97.4291 | 89.66574 |
| 6 | 2.1066 | 0.00071 | 7.6565 | 5.76824 | 97.1363 | 91.998 |
| 7 | 1.5065 | 0.0008 | 7.82 | 6.45192 | 112.4981 | 83.02083 |
| 7 | 1.5064 | 0.00079 | 7.5531 | 6.62114 | 110.2404 | 83.7007 |
| 7 | 1.5064 | 0.00076 | 8.5636 | 6.26633 | 112.2676 | 85.87087 |
| 7 | 1.5065 | 0.00078 | 7.841 | 6.26905 | 113.0941 | 86.15219 |
| 7 | 1.5065 | 0.00079 | 8.0721 | 6.26114 | 114.8151 | 83.87605 |
| 8 | 1.5055 | 0.00075 | 8.9275 | 6.86793 | 138.4851 | 92.86547 |
| 8 | 1.5058 | 0.00075 | 8.2052 | 6.87494 | 136.8342 | 92.79515 |
| 8 | 1.5055 | 0.00074 | 7.6574 | 7.00758 | 131.0297 | 91.20435 |
| 8 | 1.5059 | 0.00078 | 7.8098 | 6.89074 | 132.0554 | 92.33948 |
| 8 | 1.5058 | 0.00077 | 8.469 | 6.92644 | 138.5852 | 92.99349 |

**Table:** Execution time, avg time of finding box and delivering box data of me vs my colleague

**Execution Time:**

- My Implementation: Mean = 98.86s, Std Dev = 23.71s.
- Colleague's Implementation: Mean = 80.13s, Std Dev = 18.23s.

T-test for Execution Time:

t-statistic = 4.003416183300305

p-value = 0.00021566851110559757

- Result: My execution time is significantly worse than my colleague's ($p < 0.05$).

**Finding the Box:**

- My Implementation: Mean = 1.63s, Std Dev = 1.60s.
- Colleague's Implementation: Mean = 0.88s, Std Dev = 0.77s.

T-test for Average Time Finding Box:

t-statistic = 10.6204162640387

p-value = 3.402425780871767e-14

- Result: My average time finding the box is significantly worse than my colleague's ($p < 0.05$).

**Delivering the Box:**

- My Implementation: Mean = 8.18s, Std Dev = 1.33s.
- Colleague's Implementation: Mean = 7.82s, Std Dev = 1.33s.

T-test for Average Time Delivering Box:

t-statistic = 0.020423517355574756

p-value = 0.9837901874332756

- Result: No significant difference found in average time delivering the box ($p > 0.05$).

## 4. Discussion

**Statistical Analysis:**

We can conclude that null hypothesis is invalid for total execution time and time required for finding a box as $p < 0.05$ but null hypothesis is valid for delivering the box as $p > 0.05$.

- My execution time and average time finding the box are significantly worse than my colleague's.
- There is no significant difference in average time delivering the box or in success and failure rates between the two implementations.

Interpretation:

- My colleague's implementation is more efficient in execution time and finding the box.

## 5. Conclusion

The statistical analysis indicates that my colleague's implementation outperforms mine in execution time and finding the box. However, there is no significant difference in delivering the box or in success and failure rates between the two implementations. Thus, we reject the null hypothesis and conclude that my colleague's implementation performs better in this scenario.