

Assignment 2 Bethan Moyle

1707952

by Bethan Moyle

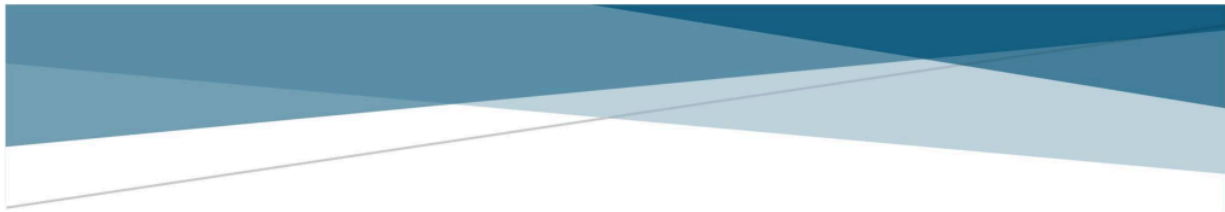
Submission date: 14-Jan-2026 12:49PM (UTC+0000)

Submission ID: 273249306

File name: 57913_Bethan_Moyle_Assignment_2_Bethan_Moyle_1707952_709067_1427138506.docx (129.3K)

Word count: 2418

Character count: 13397



NETWORK SPEED DATA ANALYSIS

Data Analysis & Visualisation

Module Code: ACCB4002

Bethan Moyle

1707952@student.uwtsd.ac.uk

Table of Contents

1	Introduction.....	2
2	Dataset description	2
2.1	Dataset	2
2.2	What is routing algorithm?	2
2.3	Identifying Variables	2
2.4	Why is Statistical testing important?.....	3
2.5	Type I and Type II Errors.....	3
3	Hypothesis.....	4
3.1	Null Hypothesis	4
3.2	Alternative Hypothesis.....	4
4	Testing and result.....	4
	What is a t-test?	4
4.1	Welch's t-test	4
5	Tests.....	5
5.1	0.05	5
5.2	0.1	5
5.3	0.01	6
5.4	Result conclusion	6
6	Outlier.....	6
6.1	How does the outlier impact the result?	6
6.2	Test Results:	7
6.2.1	Outlier included	7
6.2.2	Outlier Excluded.....	8
6.3	Results Conclusion.....	8
7	Conclusion.....	8
8	References	9

1 Introduction

In modern computer networks, efficient data transmission is essential for ensuring fast and reliable communication. One of the key factors influencing network performance is the routing algorithm, which determines the path data packets take from source to destination. Even small improvements in routing decisions can lead to noticeable reductions in latency, improving overall system responsiveness and user experience. As a result, evaluating and comparing routing algorithms is a critical task in network performance analysis.

This report investigates whether a newly implemented routing algorithm offers a statistically significant improvement in latency compared to an existing algorithm. Using response time data collected from two independent servers, each operating under a different routing algorithm, statistical methods are applied to assess differences in mean latency. Welch's two-sample t-test is used to account for potential unequal variances and the presence of an outlier within the dataset. Through hypothesis testing, analysis at multiple significance levels, and an evaluation of outlier impact, this study aims to determine whether the new routing algorithm provides a meaningful performance advantage.

2 Dataset description

2.1 Dataset

The dataset presents response time measurements for two servers using different algorithms. Server 1 operates with the original routing algorithm, while server 2 operates using a newly implemented routing algorithm. For each server, latency was recorded in milliseconds for 42 individual data packets.

2.2 What is routing algorithm?

A routing algorithm is a set of procedures or a software program used by a router to determine the efficient path for a data packet to travel from its source to its destination across a network [1]. The primary role is to calculate the most efficient path to ensure the data packets reach their destination quickly minimizing delays from factors such as distance, bandwidth, delay or network load [1].

2.3 Identifying Variables

The routing algorithm is treated as the independent variable in this data, as it is the factor that differs between the two groups and influences the systems performance [2]. Latency is the dependent variable, as it represents the outcome being measured in response to the routing algorithm used [2]. The two samples are independent of each other, as each latency corresponds to a separate data packet. Latency

measurements may be influenced by other factors such as processing delays, network congestion or packet routing paths [3] which are not explicitly controlled within this dataset.

In addition to identifying the independent and dependent variables, several statistical assumptions must be considered to ensure the validity of the chosen analysis. The latency measurements are treated as continuous numerical data, making them suitable for parametric statistical testing. Each latency value corresponds to a separate data packet and is assumed to be independent, meaning that the response time of one packet does not influence another. This independence is a key assumption of two-sample t-tests and is reasonable in this context given that each packet represents a separate transmission event.

Understanding the dataset allows for appropriate formation of hypothesis and informed selection of suitable statistical testing [4] to compare the mean latency values between both routing algorithms.

2.4 Why is Statistical testing important?

Statistical testing is important as it helps determine if observed patterns in sampled data are real or just random chance, allowing researchers to draw reliable conclusions, preventing misinterpretation, and ensure research integrity by quantifying evidence for or against a hypothesis [5]. It filters out random fluctuations to identify true effects or relationships, preventing decisions based on flukes, reduces the chance on making a decision that could potentially have a Type II or Type I error, it validates findings and improves the quality of studies enhancing research quality, prevents misinterpretation by ensuring consistent understanding among professionals preventing contradictory conclusions from the same data [5]. It also guides decision-making as it allows us to place more trust in the study results allocating sources effectively [5].

2.5 Type I and Type II Errors

Statistical hypothesis testing also involves the consideration of potential decision errors [6]. A Type I error (false positive) occurs when the null hypothesis is incorrectly rejected, meaning that a statistically significant difference is detected when no true difference exists. In the context of this study, a Type I error would involve concluding that the new routing algorithm improves latency when it does not [6]. The probability of committing a Type I error is controlled by the chosen significance level (p-value), such as 0.05 or 0.01. A Type II (false negative) error occurs when the null hypothesis is not rejected despite a true difference being present [6]. In this case, a Type II error would involve failing to detect a genuine improvement in latency offered by the new routing algorithm. By selecting appropriate significance levels and using a statistical test that acknowledge unequal variances, this analysis aims to balance the risks of

both Type I and Type II errors, ensuring that conclusions drawn from the data are both cautious and reliable [6].

3 Hypothesis

3.1 Null Hypothesis

There is no statistically significant difference between the mean latency of server 1 operating under the old routing algorithm and the mean latency of server 2 operating under the new routing algorithm.

3.2 Alternative Hypothesis

There is a statistically significant difference between the mean latency of server 1 using the old routing algorithm and the mean latency of server 2 using the new routing algorithm.

4 Testing and result

What is a t-test?

A t-test is a statistical technique used to assess whether there is a meaningful difference between the average values of two groups [7]. There are three commonly used forms of the t-test. A one-sample t-test evaluates whether a single group's mean differs from a known or expected value, an independent samples t-test examines differences between the means of two unrelated groups, and a paired sample t-test compares measurements taken from the same group under different condition or at different points in time [7]. Selecting the appropriate type of t-test for the given dataset is essential to produce accurate and dependable results [7].

4.1 Welch's t-test

Welch's t-test is a version of the independent samples t-test used to compare the means of two groups that might have unequal variances [8][9], and possibly different sample sizes, making it more robust than the standard students t-test when the assumption of equal variances isn't met [9]. It calculates the degrees of freedom using the Welch-Satterthwaite equation to adjust for these differences [9], providing a more accurate result for comparing means in real world data [9]. When the variances are unequal, the estimate of the difference between means becomes less precise, therefore, reducing the degrees of freedom makes the test more conservative and helps protect against an increased risk of Type I error (false positive) [9].

Welch's t-test [8][9] was selected because the two samples are independent and there is no strong evidence to suggest that their variances are equal. This concern is reinforced by the presence of an extreme outlier in the server 1 latency data (refer to section 6), which may increase the variance and disrupt the equal-variance assumption required for the standard two-sample t-test, unlike the student standard t-test. Therefore, the results of our two-sample independent t-test is reliable and trustworthy.

5 Tests

5.1 0.05

	<i>Algorithm 1</i>	<i>Algorithm 2</i>
Mean	88.87142857	77.61904762
p-value	0.05	
df	53	
t obt	8.216787683	
t Critical two-tail	2.005745995	

A two-sample t-test assuming unequal variances [9] revealed a statistically significant difference between Algorithm 1 and Algorithm 2 using a significance level of 0.05:

$$t(53) = 8.22 \text{ (2dp)}, p < 0.05 \text{ (two tailed)}$$

The mean performance score for Algorithm 1 (88.87(2dp)) was higher than the mean for Algorithm 2 (77.62(2dp)), indicating that Algorithm 2 performed better than Algorithm 1 within this data set.

5.2 0.1

	<i>Algorithm 1</i>	<i>Algorithm 2</i>
Mean	88.87142857	77.61904762
p-value	0.1	
df	53	
t obt	8.216787683	
t Critical two-tail	1.674116237	

When the significance level was increased to a significance level of 0.10, the result remained statistically significant,

$$t(53) = 8.22(2dp), p < 0.10 \text{ (two tailed)}$$

the obtained t-statistic exceeded the critical t value at this level, indicating further strong evidence against the null hypothesis. As with the 0.05 level, Algorithm 2 showed significantly faster mean performance than Algorithm 1.

5.3 0.01

	Algorithm 1	Algorithm 2
Mean	88.87142857	77.61904762
p-value	0.01	
df	53	
t obt	8.216787683	
t Critical two-tail	2.671822636	

Even when the significance level was reduced to 0.01, the difference between the two algorithms remained statistically significant

$$t(53) = 8.22(2dp), p < 0.01 \text{ (two tailed)}$$

this indicates very strong evidence against the null hypothesis and suggests that the observed difference in performance is highly unlikely to have occurred by chance.

5.4 Result conclusion

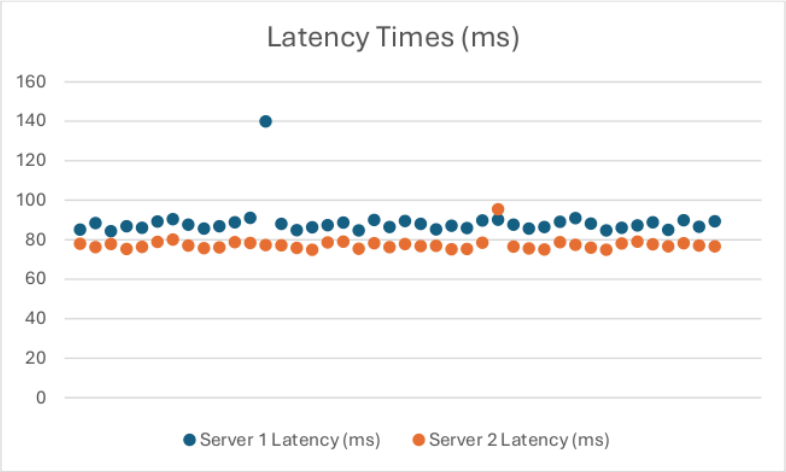
The consistency of significance across all tested alpha levels indicates that the result is robust and not sensitive to changes in the significance threshold.

6 Outlier

6.1 How does the outlier impact the result?

Visualising data is an essential step in analysis, as it allows patterns, trends, and influential factors to be identified that may otherwise go unnoticed in numerical results alone [4]. Data visualizations play a crucial role in exploratory data analysis by enabling analysts to inspect data visually, detect anomalies, and form initial

hypotheses [4]. They also assist in the data-cleaning process by highlighting missing values or unusual observations, helping to ensure the accuracy and reliability of subsequent analysis [4].



A scatter chart is particularly effective for identifying outliers, as it clearly displays data points that deviate significantly from the overall pattern or cluster, appearing isolated from the main group. Using a scatter chart to examine the spread of the data revealed an extreme outlier of 140 ms for Algorithm 1. This outlier may be the result of error factors such as network traffic or power supply interruptions [4]. To prevent skewed results and avoid drawing incorrect conclusions, it is important to conduct the analysis both with and without the outlier and assess its impact on the results [4].

Removing the outlier results in unequal sample sizes between the two algorithms. In this case, Welch's t-test [9] is appropriate, as it is a two-sample t-test that does not assume equal variances and can accommodate unequal group sizes. This is achieved by adjusting the degrees of freedom using [9]. To evaluate the impact of the outlier, the mean values and t-statistics can be compared using a significance level of 0.05.

6.2 Test Results:

6.2.1 Outlier included

The results of the two-sample t-test assuming unequal variances with p=0.05:

	Algorithm 1	Algorithm 2
Mean	88.87142857	77.61904762
t Stat	8.216787683	
t Critical two-tail	2.005745995	

6.2.2 Outlier Excluded

The results of the two-sample t-test assuming unequal variances with $p=0.05$, with the outlier of 140ms excluded from the dataset:

	Algorithm 1	Algorithm 2
Mean	87.685	77.60731707
t Stat	17.36395168	
t Critical two-tail	1.996564419	

6.3 Results Conclusion

The results show that removing the outlier has minimal impact on the mean latencies of the algorithms. However, the t-statistic increases substantially. This suggests that the outlier inflated the variability in the original dataset, and its removal reduced this variability, thereby strengthening the statistical significance of the observed difference in latency. As a result, the outlier does not influence the overall outcome of the analysis. These findings confirm that the earlier conclusion regarding Algorithm 2's superior latency performance remains valid, even after the outlier is removed.

7 Conclusion

This analysis set out to determine whether there is a statistically significant difference in mean latency between two routing algorithms used by separate servers. Using Welch's t-test, a consistent and statistically significant difference was observed across all tested significance levels ($\alpha = 0.10, 0.05$, and 0.01). In each case, the null hypothesis was rejected, providing strong evidence that the routing algorithm has a measurable effect on server latency.

The results demonstrated that Algorithm 2 consistently achieved a lower mean latency than Algorithm 1, indicating superior performance. Further investigation into the impact of an extreme outlier showed that, while the outlier increased variability in the dataset, its removal did not alter the overall conclusion. In fact, excluding the outlier strengthened the statistical evidence by reducing variance and increasing the t-statistic, reinforcing the robustness of the findings.

Overall, the study confirms that the new routing algorithm significantly improves latency performance compared to the original algorithm. The consistency of results across multiple significance levels and under outlier analysis increases confidence in the validity and reliability of this conclusion. These findings highlight the value of appropriate statistical testing in network performance evaluation and support the adoption of the new routing algorithm for improved efficiency.

8 References

- [1] U. Sinha, V. Kumar, and S. Kumar Singh, "Routing Algorithms: a Review," *Int. Res. J. Mod. Eng. Bangalore Inst. Technol.*, vol. 4601, no. 03, pp. 1–10, 2024.
- [2] D. Cramer and D. Howitt, *Sage Dictionary of Statistics: A Practical Resource for Students in the Social Sciences*. London, Thousand Oaks, New Delhi: SAGE Publications, 2004. doi: 10.1108/09504120510580208.
- [3] A. Steed, "Chapter 10: Requirements," in *Networked Graphics : Building Networked Games and Virtual Environments*, Morgan Kaufmann, 2010, ch. 10, pp. 313–350.
- [4] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5th ed. 2018.
- [5] R. A. Cooper, "Making Decisions with Data : Understanding Hypothesis Testing & Statistical Significance Students must be given opportunities to analyze data in their science classes , using statistical methods that are suited to the data and age- The Lady Tasting Tea &," vol. 81, no. 8, pp. 535–542, 2019.
- [6] A. Banerjee, U. B. Chitnis, S. L. Jadhav, J. S. Bhawalkar, and S. Chaudhury, "Hypothesis testing, type I and type II errors," *Ind. Psychiatry J.*, vol. 18, no. 2, pp. 127–131, Dec. 2009, doi: 10.4103/0972-6748.62274.
- [7] D. C. Howell, *Statistical Methods for Psychology*, 5th ed. Duxbury, 2002.
- [8] B. L. Welch, "The Significance of the Difference Between Two Means when the Population Variances are Unequal," *Biometrika*, vol. 29, no. 3/4, pp. 350–361, 1938, doi: 10.2307/2332010.
- [9] L. B. W. Frederick J. Gravetter, *Statistics for the Behavioral Sciences*. 2009.

Assignment 2 Bethan Moyle 1707952

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

58/100

Generic Assessment Criteria	Marks available	Marks awarded
1. Engagement with Literature Skills	10	
Bibliography and referencing present and good.		9
2. Knowledge and Understanding Skills	30	
Reasonably good knowledge demonstrated of the outliers. More detailed discussion and evidence of descriptive stats would have been beneficial as would mentioning IQR for finding the outlier(s). Good visualisation, illuding to correlation however more comparison may have been advantageous with further visualisations		16

and through more detailed discussion of correlation but overall a reasonably good knowledge demonstrated through research.		
3. Cognitive and Intellectual Skills	3 0	
Reasonable interpretation of results however more detailed and a little more explicit discussion would have been desirable. It would have been beneficial to explicitly discuss acceptance or rejection of H0 and also H1 at all three significance levels in an explicit manner. Could the t-test have been one tailed?		1 5
4. Practical Application Skills	2 0	
A reasonable H0 and H1, however it may have been more beneficial to maybe use $H_0 = S_1 \leq S_2$ if $H_1 = S_1 > S_2$ for a one tailed test. Good use of a scatterplot, however more evidence of statistical testing would have been desirable along with further descriptive graphs.		1 1

85. Transferable Skills for Life and Professional Practice	10	
A well laid out and sectioned report with good signposting.		7

PAGE 1
PAGE 2
PAGE 3
PAGE 4
PAGE 5
PAGE 6
PAGE 7
PAGE 8
PAGE 9
PAGE 10