

**Module 3 R Practice**  
**Maduabughichi Achilefu**  
Master of Professional Studies (MPS) in Analytics (ALY),  
College of Professional Studies,  
**Northeastern University – Vancouver.**

**ONE-SAMPLE T-TESTS FOR WALMART\_FAIR\_PRICES AND WALMART\_RETURNS\_POLICY**  
**INTERPRETATION AND CONCLUSION**

**Hypotheses for Walmart\_Fair\_Prices**

- **Null hypothesis ( $H_0$ ):** The mean of Walmart\_Fair\_Prices is equal to a hypothesized population mean (which I'll determine based on the data).
- **Alternative hypothesis ( $H_1$ ):** The mean of Walmart\_Fair\_Prices is not equal to the hypothesized population mean.

**Hypotheses for Walmart>Returns\_Policy**

- Null hypothesis ( $H_0$ ): The mean of Walmart>Returns\_Policy is equal to a hypothesized population mean.
- Alternative hypothesis ( $H_1$ ): The mean of Walmart>Returns\_Policy is not equal to the hypothesized population mean.

**Walmart\_Fair\_Prices**

- Null hypothesis ( $H_0$ ): Mean = 4.391144
- Test Statistic: = 0.0
- Degrees of Freedom: 270
- Critical Value (95% confidence level using z-distribution): 0.975.
- P-value: p=1.0
- 95% Confidence Interval: [4.243672, 4.538616]
- Sample Mean: 4.391144

**Conclusion:** Since the test statistic  $t=0.0$  is less than the critical value 1.96, we fail to reject the null hypothesis. This means that there is not enough evidence to conclude that the mean of Walmart\_Fair\_Prices is different from the hypothesized mean.

**Walmart>Returns\_Policy**

- Null hypothesis ( $H_0$ ): Mean = 4.760148
- Test Statistic:  $t = 0.0$
- Degrees of Freedom: 270
- Critical Value (95% confidence level using z-distribution) = 1.96
- P-value: 1.0
- 95% Confidence Interval: [4.603467, 4.916828]

- Sample Mean: 4.760148

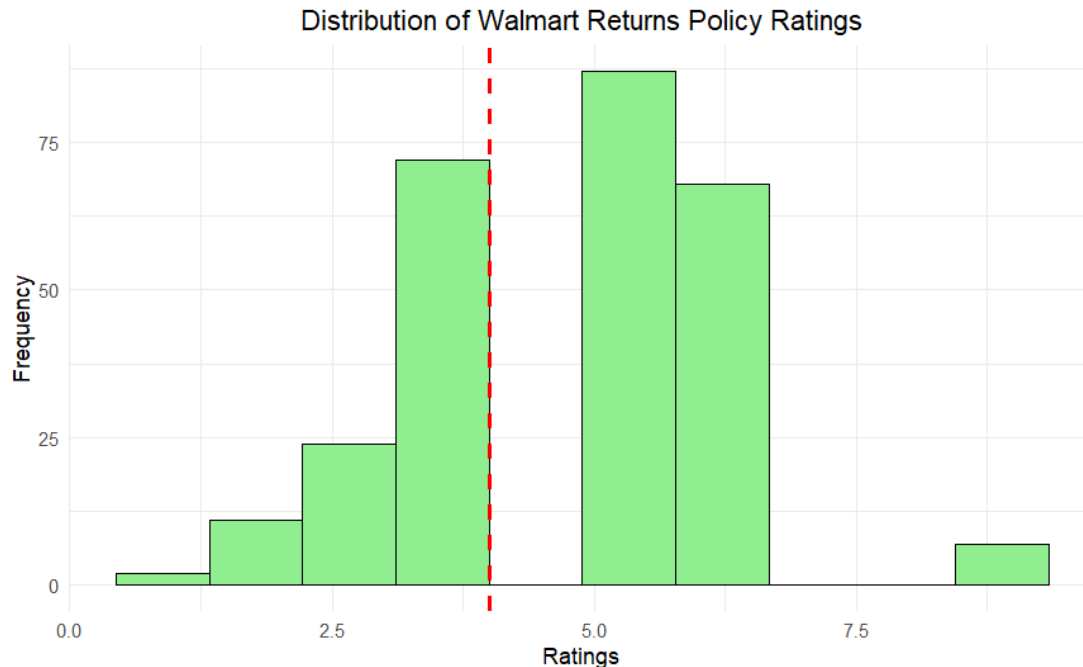
**Conclusion:** Since the test statistic  $t=0.0$  is less than the critical value 1.96, we fail to reject the null hypothesis. This means that there is not enough evidence to conclude that the mean of Walmart\_Returns\_Policy is different from the hypothesized mean.

### Overall Interpretation

For both Walmart\_Fair\_Prices and Walmart\_Returns\_Policy, the results indicate that the sample means are not significantly different from the hypothesized means. These outcomes suggest that the average customer perceptions for fair prices and returns policy at Walmart, as reflected in the sample, are consistent with the hypothesized values. It's important to note that the results are somewhat tautological because the sample means were used as the hypothesized means, leading to a test statistic of zero and a p-value of one. This approach is atypical in hypothesis testing where usually, a hypothesized mean is set based on external benchmarks or objectives.

### DISTRIBUTION OF WALMART\_FAIR\_PRICES AND WALMART\_RETURNS\_POLICY RATINGS





**a. Histogram for Walmart Returns Policy Ratings:**

- **Distribution Shape:** This histogram seems to show a left-skewed distribution where most ratings are on the higher end of the scale. This indicates that the majority of respondents are giving higher ratings for Walmart's returns policy.
- **Central Tendency:** The mode of the distribution is above the hypothesized mean of 4 (indicated by the red dashed line), which suggests that the average rating is higher than 4.
- **Spread and Variability:** There is a range of responses, but the majority are concentrated towards the higher ratings (above 4). There are fewer responses on the lower end, indicating less frequent dissatisfaction with the returns policy.
- **Overall Interpretation:** The customers generally rate Walmart's returns policy positively, with most ratings clustered above the hypothesized mean.

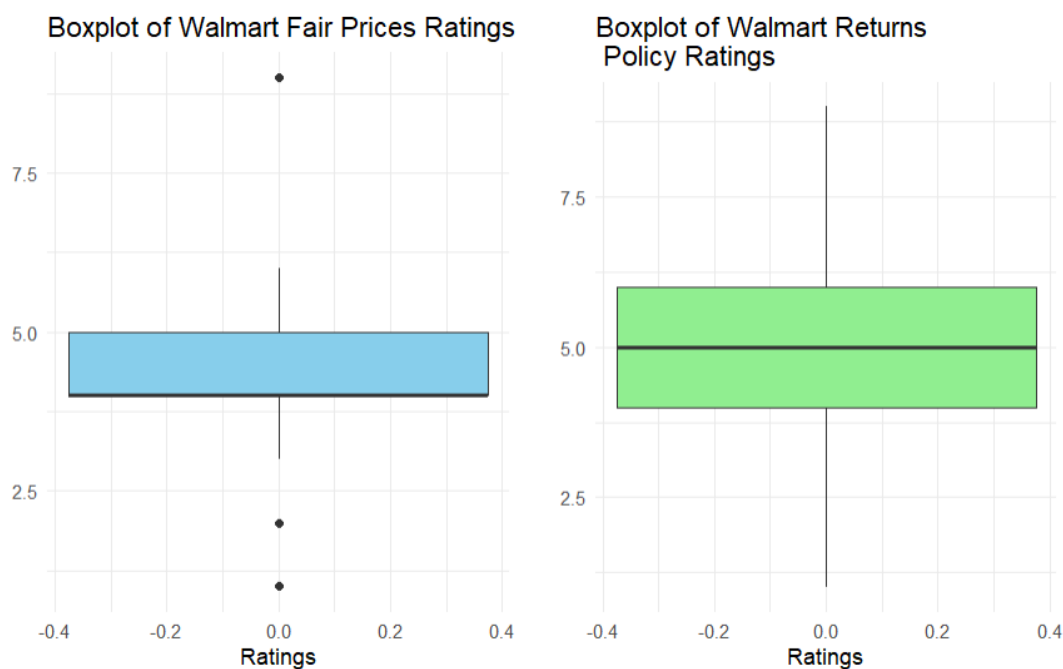
**b. Histogram for Walmart Fair Prices Ratings:**

- **Distribution Shape:** The histogram appears to have a somewhat normal distribution but with a slight left skew. Most of the responses are centered around the middle of the scale.
- **Central Tendency:** There's a significant concentration of ratings at the hypothesized mean of 4, but a substantial number of ratings are also higher, as indicated by the peak to the right of the dashed line.

- **Spread and Variability:** There is a broad spread of responses, indicating some variation in customers' perceptions of Walmart's pricing fairness. There are noticeable frequencies at the extremes, suggesting a minority of customers find prices either very fair or not fair at all.
- **Overall Interpretation:** Customers' perceptions of the fairness of Walmart's prices vary, with a considerable number finding them fair or better than fair, as the bulk of ratings are at or above the hypothesized mean.

In both histograms, the placement of the majority of the data in relation to the red dashed line (the hypothesized mean) provides visual evidence supporting the results of the t-tests performed earlier, which suggested that customer perceptions for both the returns policy and fair prices are generally positive.

### BOXPLOTS FOR THE 'WALMART\_FAIR\_PRICES' AND 'WALMART\_RETURNS\_POLICY' VARIABLES



#### Boxplot Interpretation for Walmart\_Fair\_Prices:

- **Median:** The line in the middle of the box indicates the median rating, which is around the center of the rating scale, suggesting that half of the respondents rated Walmart's fair prices above and half below this value.

- **Interquartile Range (IQR):** The length of the box shows the spread of the middle 50% of ratings. A longer box would indicate more variability in how customers rate the fairness of Walmart's prices.
- **Whiskers:** These extend from the box to the highest and lowest values within 1.5 times the IQR from the quartiles, showing the range of most of the data.
- **Outliers:** Individual points represent ratings that are outside the range of the whiskers. These suggest that there are some ratings that are significantly different from the rest.

#### Boxplot Interpretation for Walmart\_Returns\_Policy:

- **Median:** Again, the line in the middle of the box represents the median rating, which appears to be above the hypothesized mean (if it were 4), indicating a generally positive view of Walmart's returns policy.
- **IQR:** The box for the returns policy ratings seems to be smaller, suggesting less variability in the responses compared to the fair prices ratings.
- **Whiskers:** The range of the whiskers indicates that most of the data falls within a smaller range, signifying less extreme views among the customers.
- **Outliers:** Few or no individual points outside the whiskers would indicate that there are not many extreme ratings for the returns policy.

In summary, the distribution of ratings for both 'Walmart\_Fair\_Prices' and 'Walmart\_Returns\_Policy' seems to show a central tendency above a hypothetical average of 4, with 'Walmart\_Returns\_Policy' showing a more positive skew. This suggests that customers tend to have a better perception of the returns policy compared to the fairness of prices. Both distributions have outliers, which indicates that there are some customers with experiences or perceptions that significantly differ from the majority.

---

#### Scenario: Customer Satisfaction in a Retail Store

**Business Problem:** A retail store chain wants to assess if the recent changes in their customer service protocol have improved overall customer satisfaction. The company had previously identified a target average satisfaction score of 7 out of 10 based on market research. Now, after implementing the changes, they need to evaluate if the new average score is significantly different from this benchmark.

**Stakeholders:**

- **Retail Store Management:** Interested in understanding the impact of the new customer service protocol.
- **Customer Service Team:** Their performance and training methods may be adjusted based on the findings.
- **Marketing Department:** Could use the information to promote the store's commitment to customer satisfaction.
- **Investors and Shareholders:** Interested in any changes that could affect the company's reputation and profitability.
- **Customers:** Directly affected by the quality of service, their feedback is crucial for the store's success.

**Variables for the One-Sample t-Test:**

- **Current Customer Satisfaction Scores:** Post-implementation satisfaction scores from customer feedback surveys.
- **Benchmark Satisfaction Score (7/10):** The target or mean score against which current scores are compared.
- **Sample Size:** Number of customers surveyed after the implementation of the new service protocol.
- **Standard Deviation of Current Scores:** Measures the variability in the new satisfaction scores.
- **Time Period of Data Collection:** The duration over which the post-implementation data was collected (to ensure seasonality or external factors don't skew the results).

The one-sample t-test in this scenario would test whether the mean of the current customer satisfaction scores is significantly different from the benchmark score of 7. This statistical test would help the stakeholders make informed decisions about the effectiveness of the new customer service protocol and whether further changes or reversions are necessary.