**Portfolio Milestone**

Madhav Sharma

CSU Global Campus

22WA-MIS581-1

Dr. Orenthio Goodwin

01-08-2023

**Table of Contents**

[Abstract 3](#_Toc124008670)

[Introduction 4](#_Toc124008671)

[Objectives 5](#_Toc124008672)

[Overview of study 6](#_Toc124008673)

[Research Questions and Hypothesis 7](#_Toc124008674)

[Literature Review 9](#_Toc124008675)

[Methodology 11](#_Toc124008676)

[Limitation 12](#_Toc124008677)

[Ethical Consideration 13](#_Toc124008678)

[Finding 14](#_Toc124008679)

[Conclusion 17](#_Toc124008680)

[Recommendations 18](#_Toc124008681)

[References 19](#_Toc124008682)

# **Abstract**

Inflation fuel modern business environment affect the business entities big or small one way or other. Rising cost, and accessibility to the modern technology significantly impact consumer purchasing behaviors. Consumer spending more time in virtual shopping rather than instore visit. Virtual shopping allows the retailer to see the data from different perspective. This paper analyzes the shift in customer purchasing pattern and purpose the standard procedure that rapidly improve business performance and exceed the customers expectation. The research finding reveal the interdependence between the variables and revenue maximization as well as outline the strategy to mitigate reverse shipment or product return. By utilizing this finding retail leader easily outline optimal product mix and price mix that improve business performance and create the business value by exceeding customer expectation.

# **Introduction**

Modern business operates in dynamic environment surrounded by uncertainty and complexity. Success of business hinge on their ability to cooperate with uncertain business environment and implement the data driven strategic decision. On the other hand, business entities are in constant pressure to maximize the profitability, retailer has no exception. Business leader always looking for innovative ideas to maximize sales and room to minimize operating expenses. Inflation fuel current business environment affect business big or small one way or other, in a meantime it also affects customer purchasing power, customers are spending more in essential product, than household product. Supplying too little means risking profit loss on the other hand supplying too much cause the waste of the resource. Striking the balance between the demand and the supply are the fundamental process for business success. With the rise of modern technology that directly associated with the Internet of Things (IoT), organization gather excessive amount of data than ever before—and at a lightning speed i.e., Walmart collects 2.5 petabytes of unstructured data from over 1 million customers every hour. It also opens the window of opportunity for the customers, i.e., customer spending more time in virtual shopping rather than instore visit. Detecting the consumer purchasing pattern and implementing data driven strategic decision are the most challenging activity in retail success. Big data analytics is an important tool to understand shift in consumer purchasing. Combining big data with sophisticated analytics tools organization rapidly improve their performance and exceed the customer expectation. Value of data is not the data itself – it is what we do with the data. Data to be useful first we need to know what data we need, otherwise we just get tempted to know everything and that is not a strategy, it is an act of desperation that is doomed to end in failure. (Marr, 2015). K- Mart is a great example of data alone does nothing. One of the prominent business empires in US retail market. K- Mart consistently decline since the 1980s is it unable to compete with large retailer such as Walmart, target. Even though they completely restructure their dept ratio as well as implement their own business intelligence platform. Lack of effective leadership Kmart unable implement data driven strategic decision, even though they implement their own analytics platform. Their unwilling to analyze the customer data and act to market trends lead Kmart failure in the marketplace. On the flip side retailer like, Walmart relies on big data analytics to leverage its ability to manage supply chains and keep inventory levels optimal, perform research on constantly changing customer trends and demographics, and market across a multitude of social media outlets. Liu & et al (2018).

# **Objectives**

In modern data driven environment, establishing effective business analytics tools that utilize the data analytics and implement strategies business decision to organizational growth is the primary objective of any business organization. Creating the business value and improving the business performance by combining customer data (purchasing pattern) with sophisticated analytics tools is the primary objective of this research work.

# **Overview of study**

In modern business environment predicting or forecasting future growth/sales become more and more challenging. Retailer just can’t rely on traditional sales report and predict the future sales. They need to collect the data from different structure and unstructured sources such as online and offline sales report, POS, social media, trends and etc. Big data sets bring a challenging situation for retail enterprises in terms of their management accordingly and professionally, which needs to focus on specific strategies to solve the issue and create competencies for businesses (Davenport, 2014).

It is important for the retailer to identify the way to collect the data, analyze the data, store the data and utilize the data get maximum out of it. This research benefit retailer to understand the best way to collect the customer data from the sources such as social media, IoT, POS (point of sales), instore/online sales transaction and outline the way to maximize data inside. Retailer utilize this report to create optimal inventory level and minimize the operating expenses. This research work helps the retailer to identify the way to create value and improve the organization performance by evaluating the shift in purchasing pattern of the customer.

# **Research Questions and Hypothesis**

Online retailing is an opportunity for the business entities to see the data inside from different perspective. Data inside help business organization to identify potential opportunity i.e., market expansion, price adjustment for product/services as well as room for improvement i.e., understanding customer prospective towards the product and competitor position in market. Raising inflation, global economic slowdown and COVID outbreak affect retailers big or small one way or other. The impact of global economic slowdown also seen in consumer behaviors and their buying power. They are spending more in essential product, than household product as well as frequency of store visit also decline. After observing current business environment retailer / online retailer are more concern with following business question that directly and indirectly affect organizational strategic goal.

How the promotion affect customer buying behavior/How customer react with promotion/ Is it the significate factor to drive sales?

Time spent by the customer in web page/ product they looking for?

Does the relationship between unit price and ordering quantity are significant factor to boost sales?

How often the customer visit/ what are the product they order/ how often they order? How they react with alternative product/ do they follow the recommendation, how often they buy recommended product?

After identifying potential business question researcher need to understand the relationship between business problem and strategic goal. Hypothesis is an interpretation of a practical situation or condition taken as the ground for action. (Merriam-Webster, 2022). A statistical hypothesis test is a method of statistical inference used to decide whether the data at hand sufficiently support a particular hypothesis. (Wikipedia, 2022). The purpose of a hypothesis test is to demonstrate that the sample of values comes or does not come from a population whose mean is known. It allows researcher to draw the probabilistic statements about population parameters. In statistics two different test (Null hypothesis and Alternative) will be perform to measure the probability of the population parameter. Null hypothesis indicates that there is no association between two parameters, whereas alternative hypothesis indicates that there is a relationship (statistically significance) exists between two dependent and independent variables. Primary objective of this project work is to understand customer purchasing pattern and identify the opportunity to maximize the revenue. Selling price and selling quantity has significant effect in revenue maximization. To understand impact of selling price and selling quantity in revenue maximization author perform hypothesis test.

Null hypothesis variables (unit price, quantity) have no effect in store sales.

Alternative hypothesis variables (unit price, quantity) have significant effect in store sales.

# **Literature Review**

Online retail sales expected to grow exponentially, according to data from Statista, e-commerce sales amounted to approximately 5.2 trillion U.S. dollars worldwide in 2021. Expected to increase by 56% next year and, reach to 8.1 trillion dollars by 2026. This indicates customers purchasing pattern shifting from instore purchasing to online purchasing. Online shopping allows the retailers to gather exponent amount of customer data than ever before in matter of time. It is an opportunity for the retailer to observe their customers purchasing pattern more precisely. Combining data with sophisticated analytics tools help the retailer to improve their performance as well as find the way to exceed the customers expectation. Big data analytics make it possible for companies to meet emerging business demands that allow them to stay ahead of the competition. Retailers have long recognized that data-driven decision-making could improve decision quality. (Kowalczyk & Buxmann, 2015).

B. Rachford & et al (2022) examine the impact of online retailing on consumer behavior, by observing the different component such as search habit, price dispersion, product reviews and etc. Product with sensory attribute required instore inspection to satisfied the customer desire, other than that online shopping eliminate travel cost. They also conclude that online product reviews had significant impact on product sales. Retailers need to find the way to address negative reviews more precisely and effective manner to retain their customer.

Gauri & et al., (2021) argue that Consumers seek to choose a format and price level that best suits their needs in a given shopping situation. It is an opportunity for the retailer to attract customer by offering different types product/service mix at competitive price. As well as examine the service requested by customer are equally importance. If the online retailer misses the service window there is possibility to lose those customer and more likely service and product purchased may be return by customer.

If the Product/service does not meet the customer expectation more likely the consumer returns the product/service, has negative impact in revenue maximization. To mitigate product return, retail should offer the quality product to the customer. As well as allow them physical inspection for product quality. It is also essential to establish generous return policies incase product does not meet the fit criterial. Online retailers such as Amazon, Stitch Fix, Warby Parker utilize Try-Before-You-Buy” (TBYB) programs to mitigate product fit uncertainty (product return) by letting consumers to try multiple items at home before buying online to fully mitigate PFU at no cost of shipping and return. (Wingfield, 2017).

Li & et al, 2019 argue CSM (consumer self-mending) behavior plays an important role in an online retailer’s TBYB strategy. With CSM, the adoption of the strategy critically depends on the operation cost, product value, and the size of member segments. Specifically, the strategy is viable if: 1) the operation cost is low enough; or 2) the operation cost is relatively high, but the product value is large enough; or 3) although the operation cost is relatively high, and the product value is relatively small, the size of member segment is large enough.

Bradlow & et al (2017) data accuracy is a critical component in online retailing, big data analytics improve the data quality (‘better’ data) rather than merely a rise in data volumes that drives improved outcomes. Authors outline five data dimension such as customer, product, time, location and channel are the key component in online retailing. Big data analytic and predictive analytics help the retailer optimize product mix and improve the business outcome. Data gather from online retailing allows retailer to see data inside such as, trend, frequency, amount spent, time spent. Advance analytics modeling and predictive analytics allow them to create strategic decision by utilize customer data gather from different sources.

# **Methodology**

In literature review author examine different scholarly article to understand the different factor that influence customer purchasing decision or behavior. For this research project author evaluate pre-existing dataset “online retail” from UCI machine learning repository. Dataset capture the transaction occurred between 01/12/2010 to 09/12/2011, from UK-based and registered non-store online retailer. Author believes combining qualitative and quantitative research method generate excellent report that clearly outline the significate factor that influence consumer purchasing pattern. As well as set the alternative strategies to maximize sales and create business value. A dataset contains eight distinct variable such as invoice number, stock code, description, quantity, unit price, invoice data, customer-id, country and date with nominal output.

**Methods**

In order to evaluate relationship between different dependent and independent variable author perform t -test. After setting null hypothesis and alternative hypothesis for business question next sept step is to select the significant level for this project work, standard significant level 5% or 0.05 is used for this analysis. Author will utilize SAS studio and perform one way t-test to observe T of the test statistics T. Once the P- value is calculated author compare the calculated P-value with significate level 0.05. If the calculated p-value is less then significant level author will reject the null hypothesis (there no similarity between variables) and accept the alternative hypothesis (significate relation exist between variables). Multivariate analysis, as well as data visualization tool also utilize to see data inside.

# **Limitation**

Retail management is combination of different function with human interferences. People are directly or indirectly involved in record keeping and there is always possibility of human error. Incomplete report had a significant impact in strategic decision making. Even the right decision with wrong data, lead the business toward failure. It is an important for business to evaluate source of data, trustworthiness of data and reliability of data. Data was gathered from particular retailer in stated time, author believe data set contain accurate record. As of today, no unusual pattern are detected in dataset.

# **Ethical Consideration**

Combination of data driven strategic decision and effective utilization of the available resources lead the business to the new height. It is an important for the retailer to identify an effective way to collect, analyze and utilize the data inside get maximum out of it. Marr argue data to be useful first we need to know what data we need, otherwise we just get tempted to know everything and that is not a strategy, it is an act of desperation that is doomed to end in failure. (B. Marr, 2015). After identifying potential data, it is important to outline how we going to collect the data, what are the ethical measure are taken in account. Ethical consideration outlines the process how the data will be gathered store and exchange for further used. It may require to obtain consent and permission from the participant to use and share the data. For this research project author utilizing “online retail” data set from UCI machine learning repository. Dataset capture the transaction occurred between 01/12/2010 to 09/12/2011, from UK-based and registered non-store online retailer. To avoid the exposure of personal information author, replace customer personal information with uniquely design 5 digit integral. Since the dataset publicly available to reuse author believe dataset address all legal and ethical issues.

# **Finding**

Online-retail success depends in an effective utilization product and price mix, it helps business entities to implement data driven strategic decision. Since the primary objective of this research work is to evaluate purchasing pattern and identify the opportunity to maximize the revenue. Selling price and selling quantity has significant effect in revenue maximization. To understand impact of selling price and selling quantity in revenue maximization author perform hypothesis test.

Figure 1. hypothesis testing for selling quantity

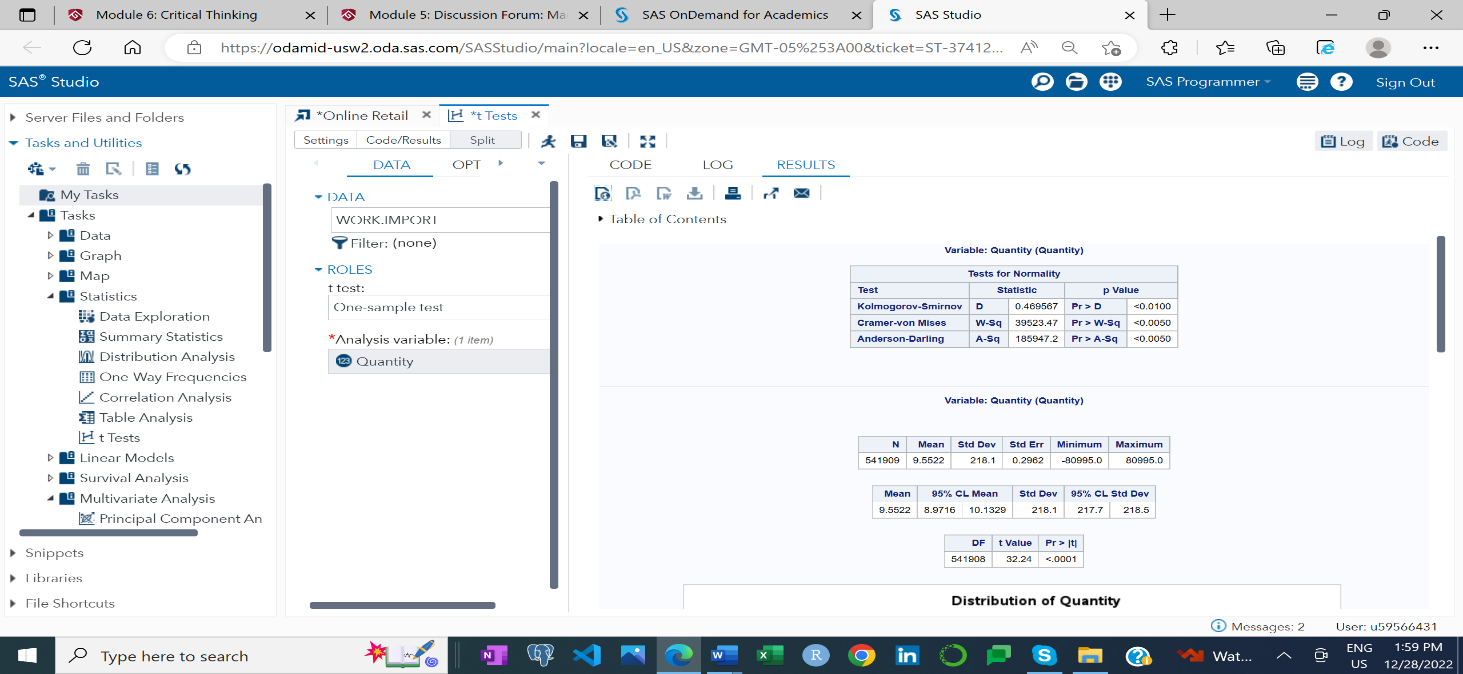
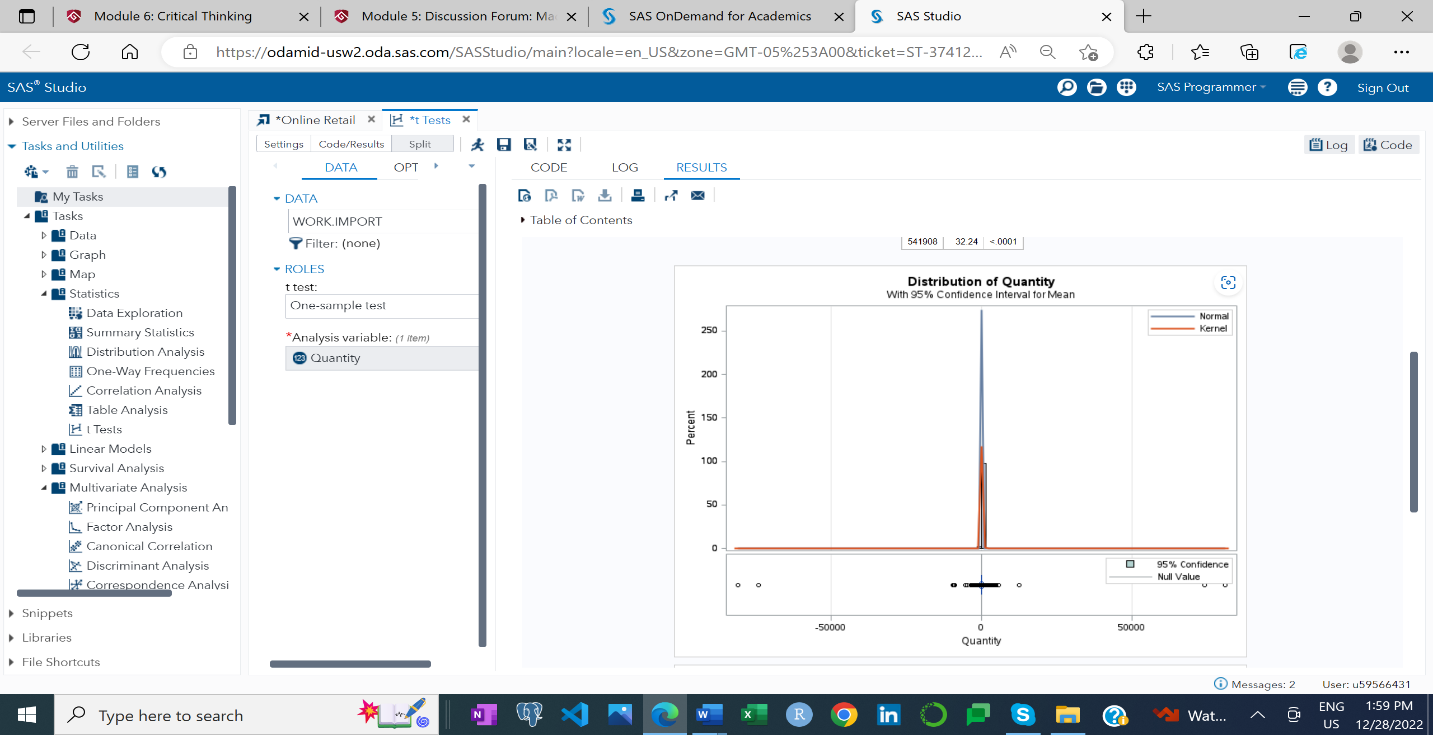
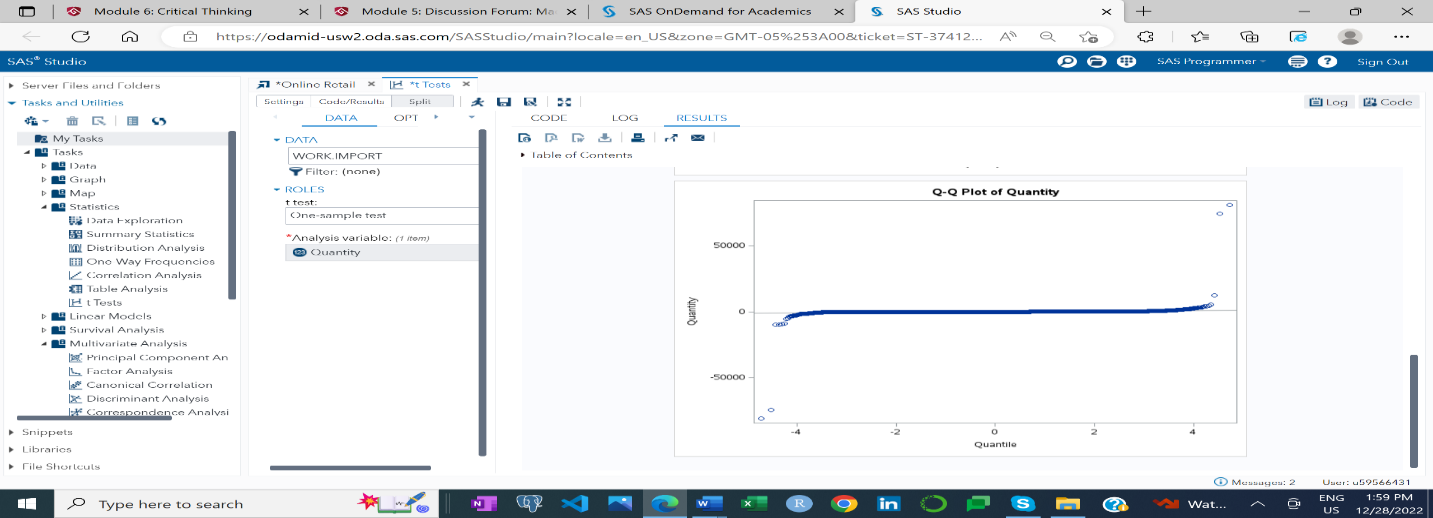


Figure 1 outline summary statistics as well as result of one-sample t-test. Number of transactions 541909 with average ordering (for customer) or selling quantity (retailer) is 9.55. Which represent 9.55 quantity of the product or service place in order for single transaction. Calculated p-value for the quantity is 0.0001 which is less then significant level 0.005, author reject the null hypothesis and accept the alternative hypothesis, means that significant relation exist between the quantity and revenue maximization.

Figure 2. Probability distribution for quantity

Author utilizes the figure two to evaluate the probability distribution for the quantity. By observing the above figure, it is clear that data are evenly distribute to both directions left and right and there is no skewed which represent the normality of distribution curve. Where mean, median and mode are equal.

Figure 3. Q and Q plot of quantity

Author observe points are fall along a line in the center of the graph, but the curve away from the line at the end with opposite direction. This outline that the distribution has extreme values than expected.

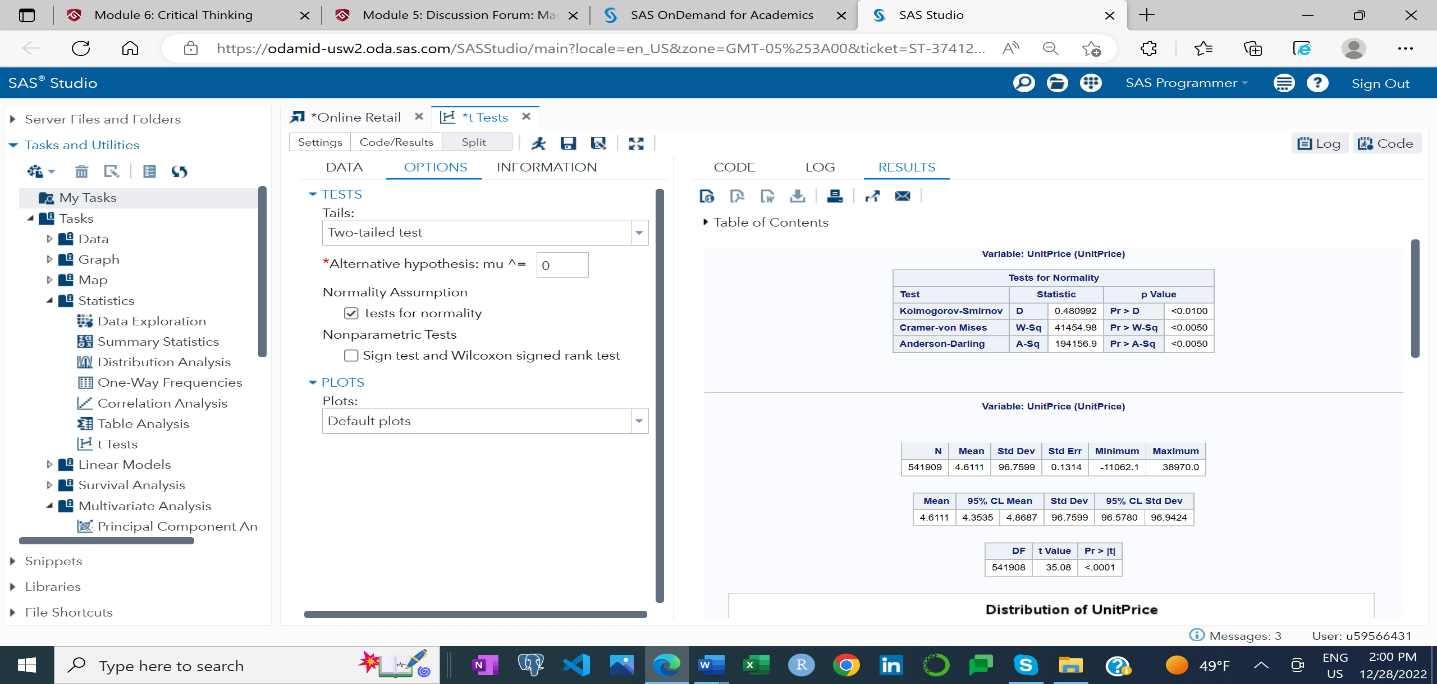
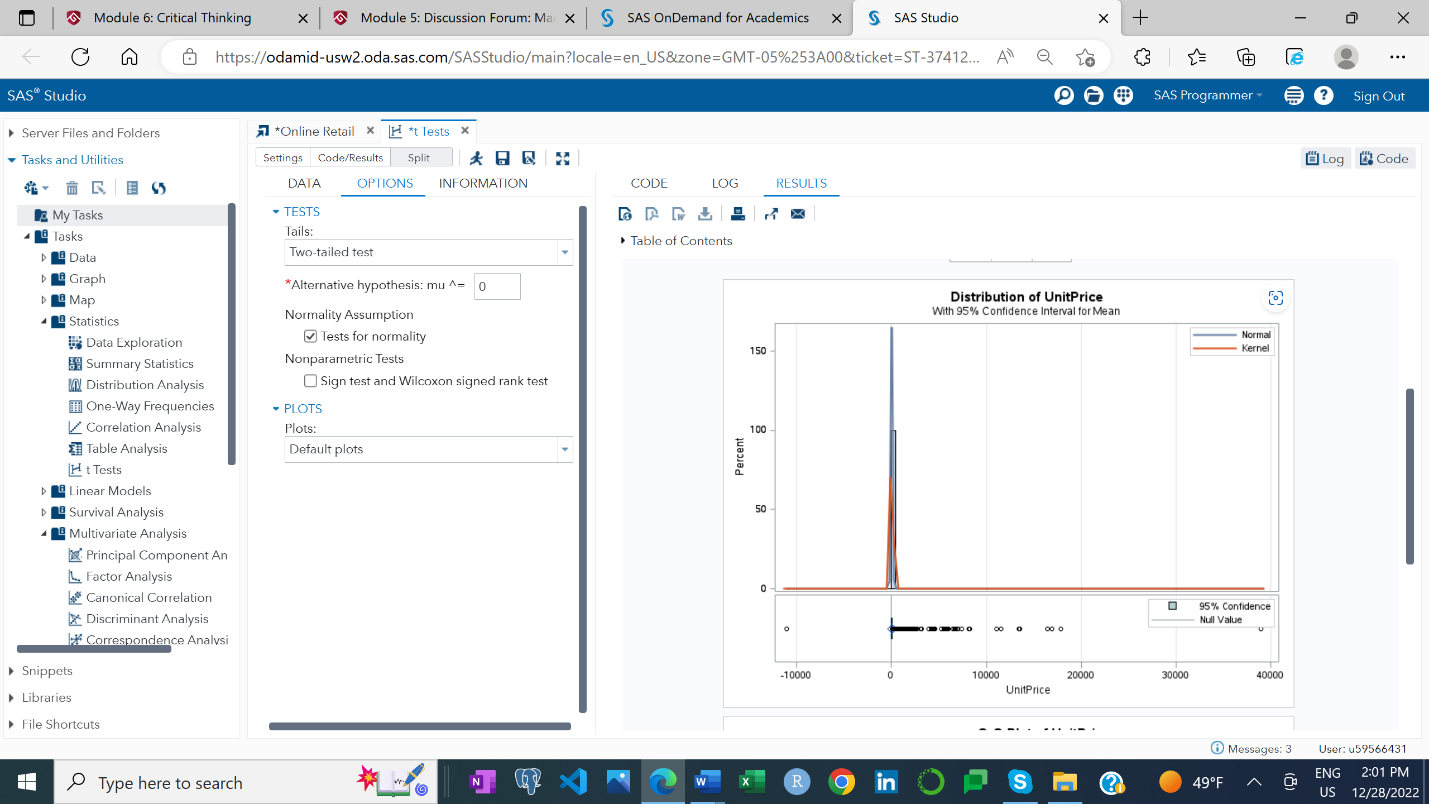
Figure 4. Hypothesis testing for unit price

Figure 4 outline summary statistics as well as result of one-sample t-test for unit price. Number of observations 541909 with average selling price per unit is $ 4. 61, with standard deviation of 96.76, on average unit price of product deviates from the mean by $96.76. Calculated p-value for the quantity is 0.0001 which is less then significant level 0.005, so the author reject the null hypothesis and accept the alternative hypothesis, means that the significant relation exists between the unit price and revenue maximization.

Figure 5. probability distribution of Unit price



Author utilizes the figure 5 to evaluate the probability distribution for unit price. By observing the above figure, it is clear that data are evenly distribute to both directions left and right and there is no skewed which represent the normality of distribution curve.

# **Conclusion**

The purpose of a hypothesis test is to demonstrate that the sample of values comes or does not come from a population whose mean is known. It allows researcher to draw the probabilistic statements about population parameters. Null hypothesis indicates that there is no association between two parameters, whereas alternative hypothesis indicates that there is a relationship (statistically significance) exists between two dependent and independent variables. By observing the figure 1 and figure 4 it is clear that calculated P- value for quantity (0.0001) and Unit-price (0.0001) is less than the significate level 0.05 (95% confidence interval) base on this author reject or disprove the null hypothesis and accept alternative hypothesis (means there is a statistical significance exist between quantity/unit price with store sales.). Retailer are in constant pressure to maximize profitability. It is obvious that unit price and unit sold has significant impact in revenue generation, but not necessarily contribute in profit maximization. If the customer not satisfied with the product or the product arrived after special occasion there is possibility that customer reject the shipment or return the product, by observing the summary statistics there is situation where whole shipment of the product was return.

# **Recommendations**

With an access to the modern technology that directly associated with the Internet of Things (IoT) customer search window widen. Customer search for the product without leaving their places, there are several online retailers competing to provide the same goods and service. In this situation attracting new customer and retaining existing customer is very critical for retailer. To create the win- win online retailer need to come up with new idea to exceeding the customer expectation and maximizing profitability. Compare to the past there is significate difference the way company do the business. To create the win- win situation in modern days business online retailer need to focus not only in sales but also in after sales services such as product delivery window, reliability of the product, customer satisfaction, customer preference. In order to succeed in this dynamic business environment online retailer, need to pay close attention what customer thinking about their product over the competitors product. Every feedback requires thorough investigation and accept this as an opportunity to improve in order to remain competitive.

# **References.**

B. Marr (2015). Big data: Using SMART big data, analytics, and metrics to make better decisions and improve performance. Wiley.

B. Ratchford., G. Soysal., A. Zentner., D. K. Gauri. (2022). Online and offline retailing: What we know and directions for future research. Journal of Retailing. 98.1. P.152-177. https://doi.org/10.1016/j.jretai.2022.02.007.

D. K. Gauri., R. P. Jindal., B. Ratchford., E. Fox., A. Bhatnagar., A. Pandey., J. R. Navallo., J. Fogarty., S. Carr., E. Howerton. (2021). Evolution of retail formats: Past, present, and future. *Journal of Retailing. Vol. 97. 1. P.42-61*. https://doi.org/10.1016/j.jretai.2020.11.002.

E. T. Bradlow., M. Gangwar., P. Kopalle., S. Voleti. (2016). The role of big data and predictive analytics in retailing. Journal of retail. Vol. 93. (1). https://doi-org.csuglobal.idm.oclc.org/10.1016/j.jretai.2016.12.004

N. Wingfield. (2017). Amazon will let customers try on clothes before buying. Retrieved on December 9, 2022, from, https://www.nytimes.com/2017/06/20/business/amazon-shopping-retail-fashion.html

Online Retail. (2015). *UCI Machine Learning Repository*. UCI. https://archive.ics.uci.edu

R. Cody. (2021). *A Gentle Introduction to Statistics Using SAS Studio in the Cloud*. SAS Institute Inc.

S. Chevalier. (2022). Global retail e- commerce sales 2014-2026. Statista. Retrieved on December 8, 2022, from. https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/#statisticContainer.

Y. Kayikci., S. Demri., S. K. Mangal., N. Subramanian., B. Koc. (2022). Data driven optimal dynamic pricing strategy for reducing perishable food waste at retailers. *Journal of cleaner production*. *Vol. 334* https://doi-org.csuglobal.idm.oclc.org/10.1016/j.jclepro.2022.131068.

Y. Liu., H. Han., J.E. DeBello. (2018). *The challenges of Business analytics: success and failure*. Proceedings of the 51st Hawaii International Conference on System Sciences. https://core.ac.uk/download/pdf/143480944.

Z. Sisi., W. Feng., J. Tao., S. Lei. (2018). The impact of product returns on price and delivery time competition in online retailing*.* *Computers & Industrial engineering. Vol.125. P 658-667.* https://doi-org.csuglobal.idm.oclc.org/10.1016/j.cie.2018.01.007