

# Northwind Traders

Hypothesis Testing

Mark Ehler

## **Executive Summary Must-Haves**

Your presentation should:

- Contain between 5-10 professional quality slides detailing:
  - A high-level overview of your methodology
  - The results of your hypothesis tests
  - Any real-world recommendations you would like to make based on your findings (ask yourself--why should the executive team care about what you found? How can your findings help the company?)
- Take no more than 5 minutes to present
- Avoid technical jargon and explain results in a clear, actionable way for non-technical audiences.

# The Tests

---

# Hypothesis

- **Discounts orders ship a higher quantity of products than those without.**
- **Northwind Traders has a preferred shipper by total order \$\$\$. They use one shipping company more than the others.**
- **Meat Orders are \$100 more than other categories on average.**
- **San Francisco ships 80% of the western orders by quantity.**

# Null-Hypothesis

- **There is no difference in between orders with discounts**
- **There is no favored shipper. Northwind uses all of them the same amount.**
- **All categories are stocked the same amount.**
- **The West Region earns more \$\$\$ than other regions.**

# **The Process**

**Ask the Question**

**Tailor the Query**

**Test the Hypothesis**

**Prove Yourself Wrong**

**Use Your Findings to Inform Your Next Question**

# A Word of Caution

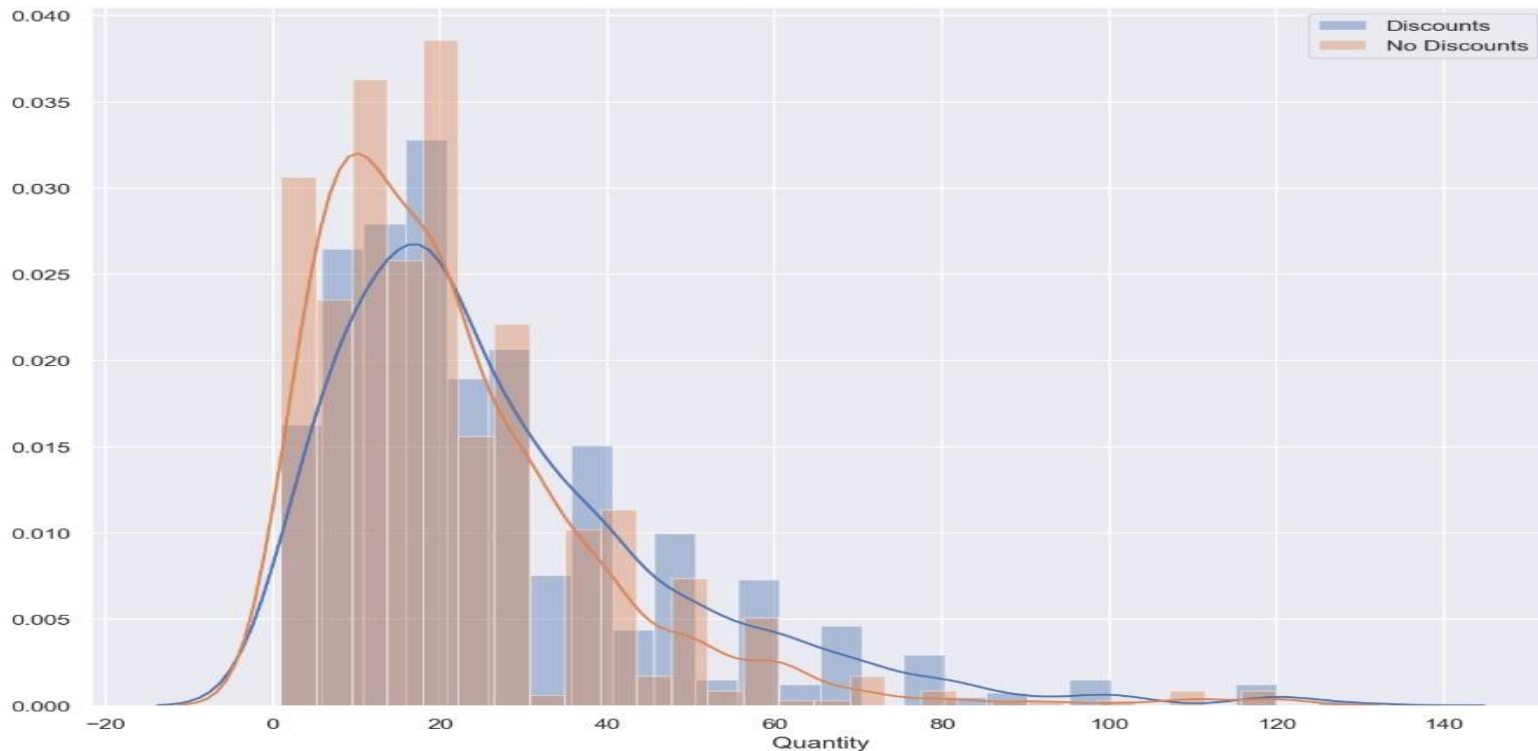
Goodhart's Law and Metric  
Tracking

"Any measure which becomes a  
target ceases to be an effective  
measure!"

---

# Discounts

Discounts orders ship a higher quantity of products than those without.



**Quantity**

**Discounts**

**No Discount**

Means:

27

21

Mean Difference:

6

Standard deviations:

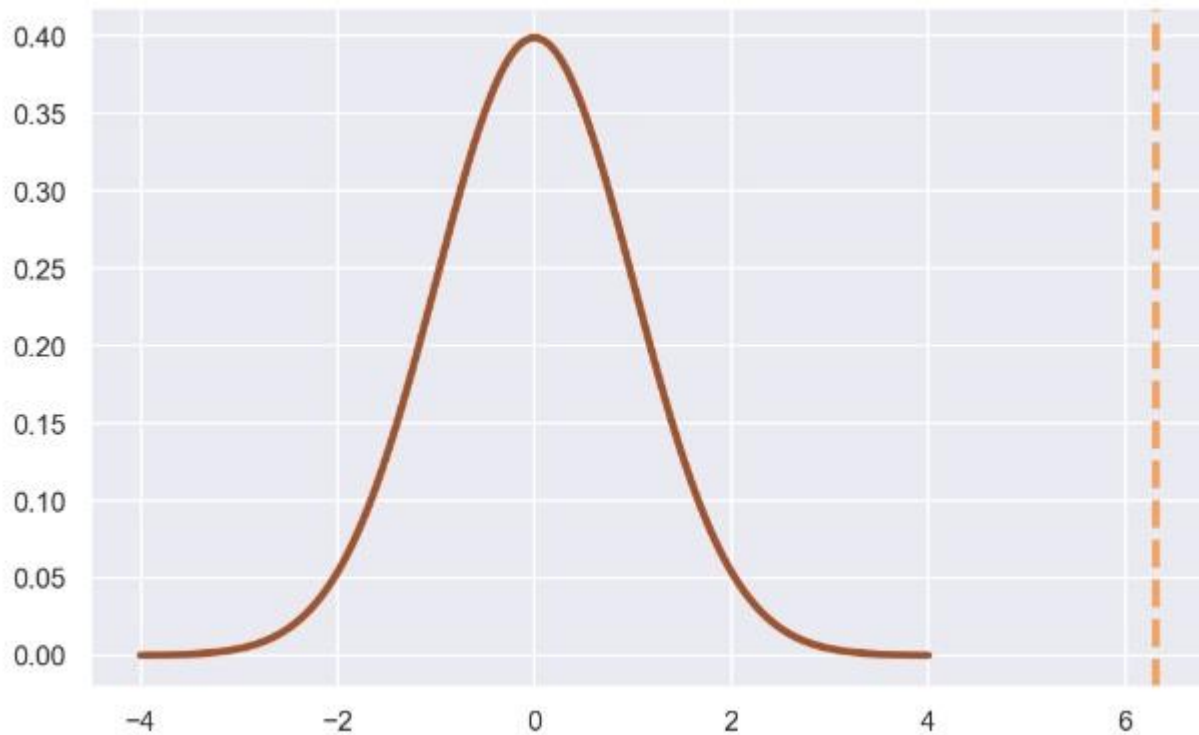
21

17



**T Value : 6.23**

**P Value : 1.6 e-10**



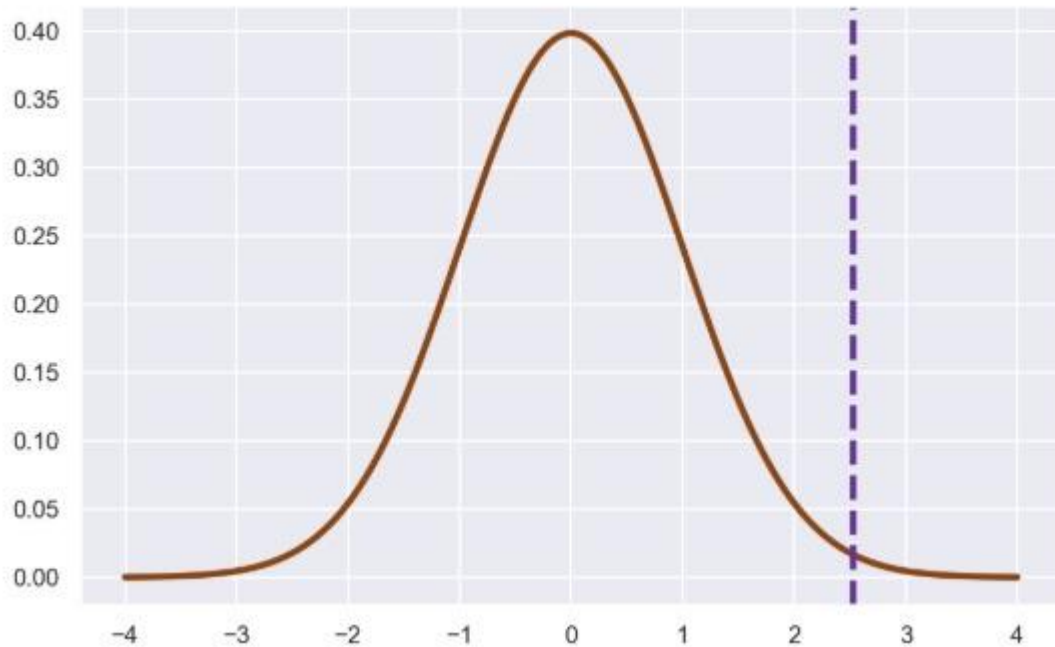
# Summary

There is enough evidence to support the claim that orders with discounts are placed with a larger quantity than orders without discounts.

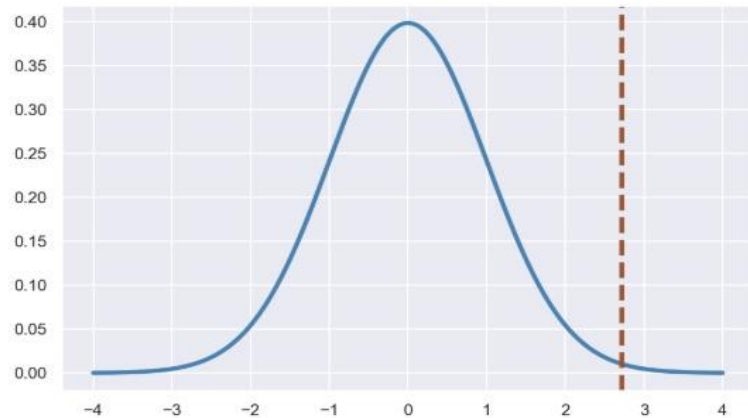


# Discount by Percentage

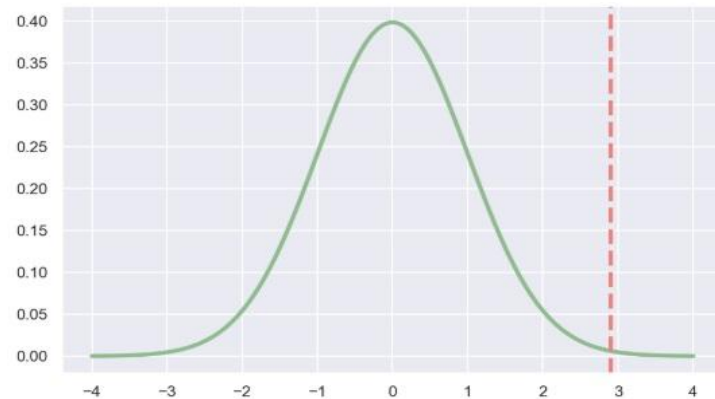
Discount Level: 0.1: Probability of error: 0.01175638486446167



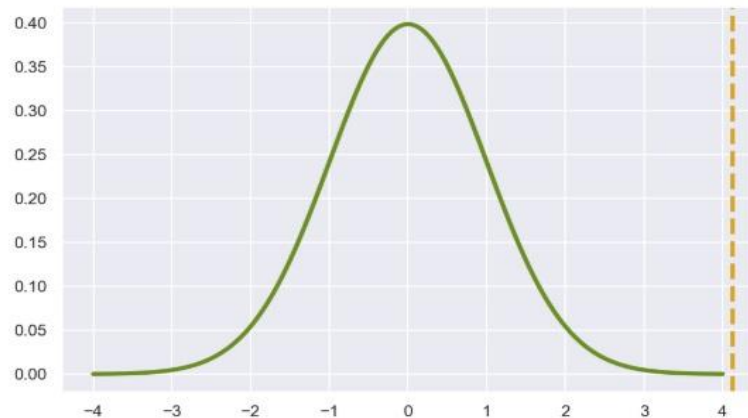
Discount Level: 0.2: Probability of error: 0.006917136252883391



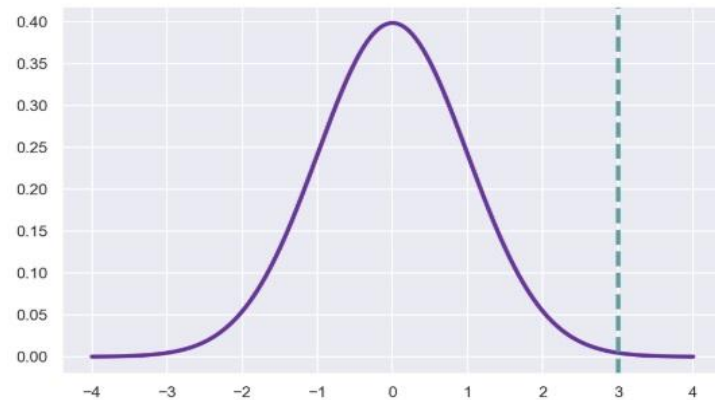
Discount Level: 0.15: Probability of error: 0.003899114297712308



Discount Level: 0.25: Probability of error: 4.77333721632627e-05



Discount Level: 0.05: Probability of error: 0.0028292023288866617

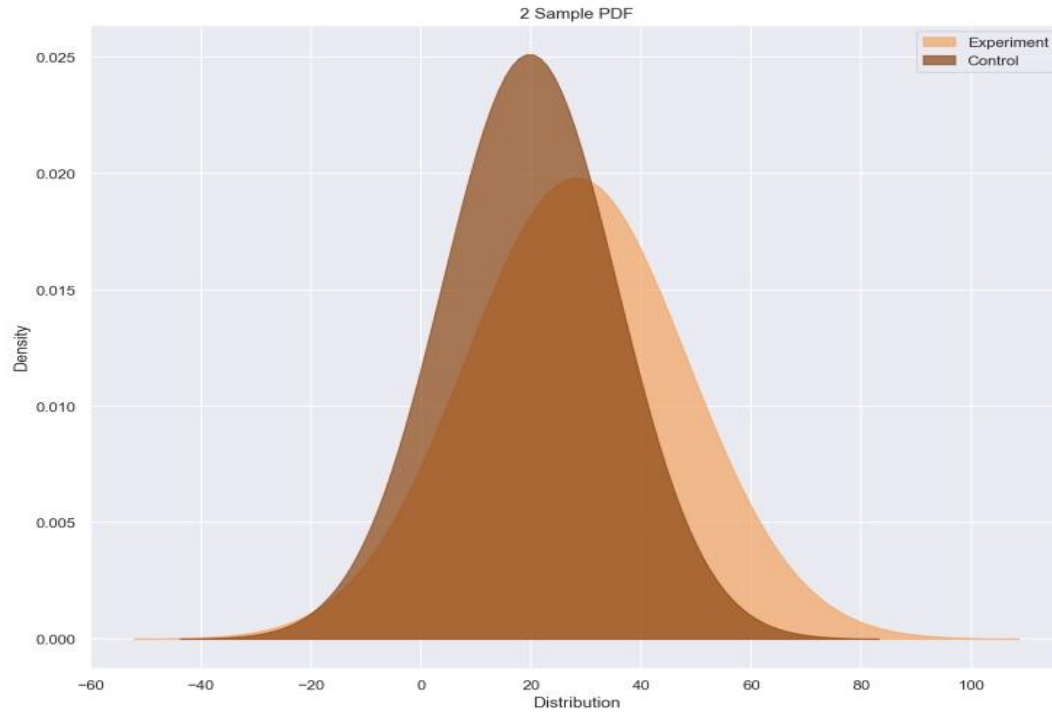


Simple Threshold: 23  
The midpoint between two means.

Overlap: 0.7727  
The total AUC.

Superiority: 62.34  
Probability that a randomly chosen sample from the first group is [higher] than one of the second group.

Misclassification Rate: 0.3864  
The chance of misclassification if using this metric alone as a predictor.

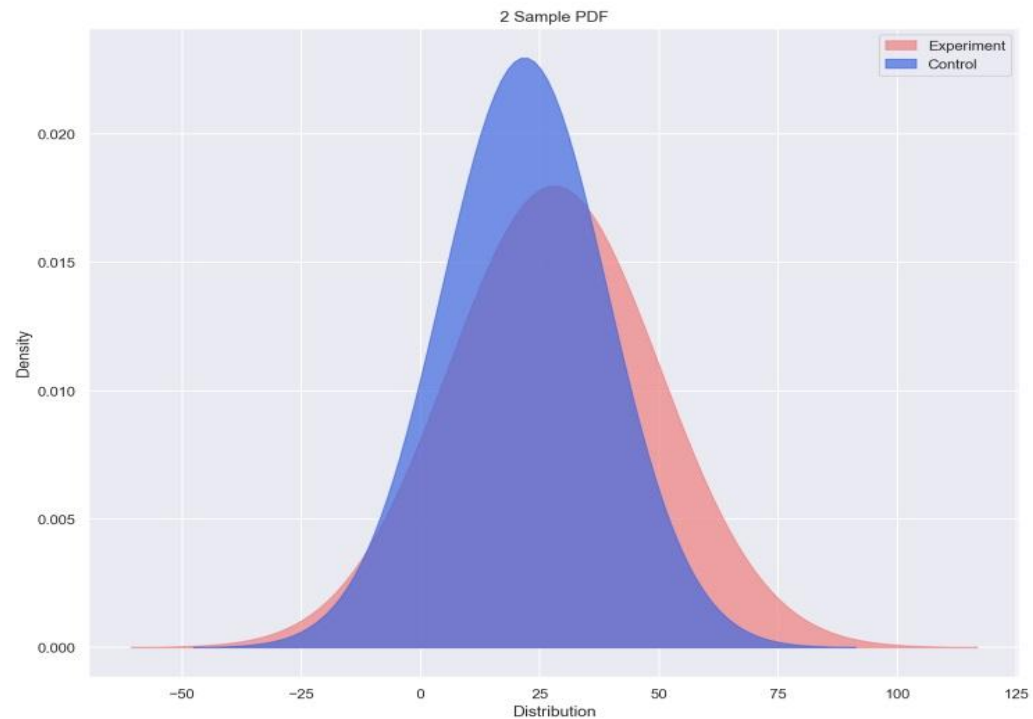


Simple Threshold: 24  
The midpoint between two means.

Overlap: 0.9297  
The total AUC.

Superiority: 52.97  
Probability that a randomly chosen sample from the first group is [higher] than one of the second group.

Misclassification Rate: 0.4649  
The chance of misclassification if using this metric alone as a predictor.

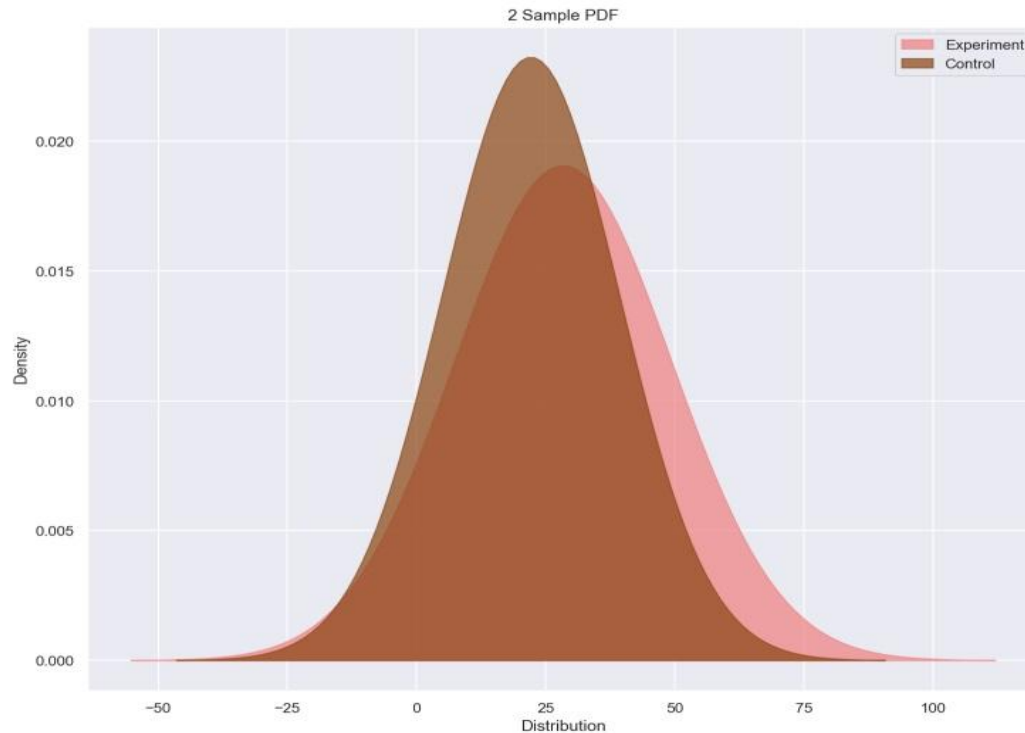


Simple Threshold: 25  
The midpoint between two means.

Overlap: 0.9045  
The total AUC.

Superiority: 54.78  
Probability that a randomly chosen sample from the first group is [higher] than one of the second group.

Misclassification Rate: 0.4522  
The chance of misclassification if using this metric alone as a predictor.

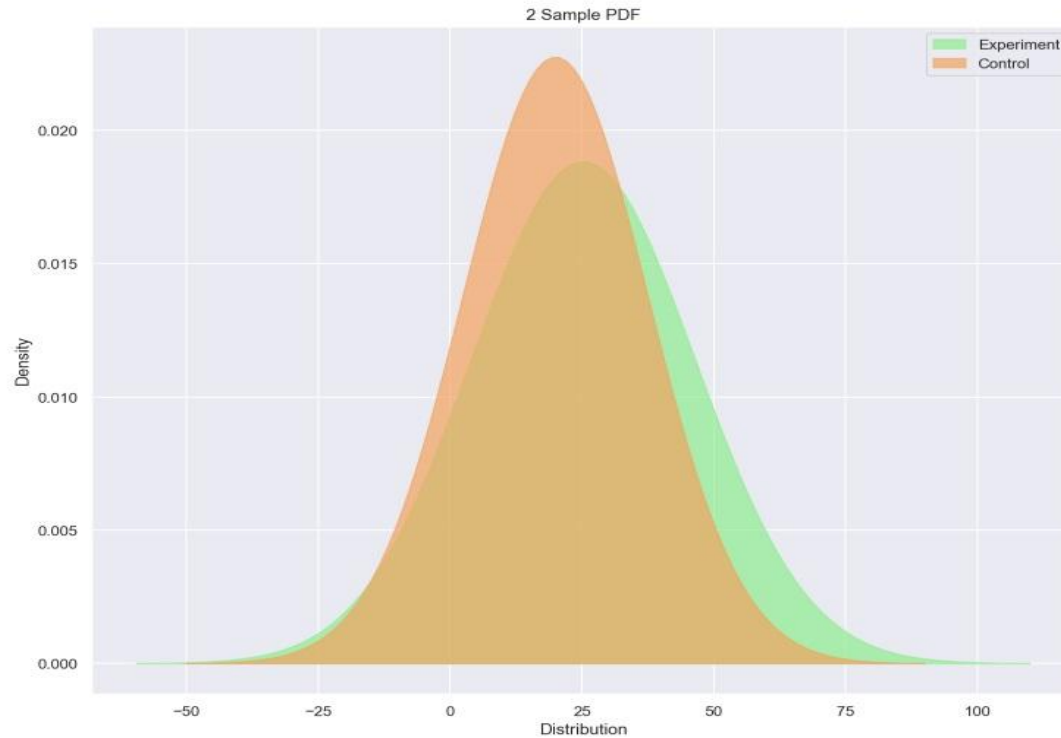


Simple Threshold: 22  
The midpoint between two means.

Overlap: 0.8844  
The total AUC.

Superiority: 56.65  
Probability that a randomly chosen sample from the first group is [higher] than one of the second group.

Misclassification Rate: 0.4422  
The chance of misclassification if using this metric alone as a predictor.



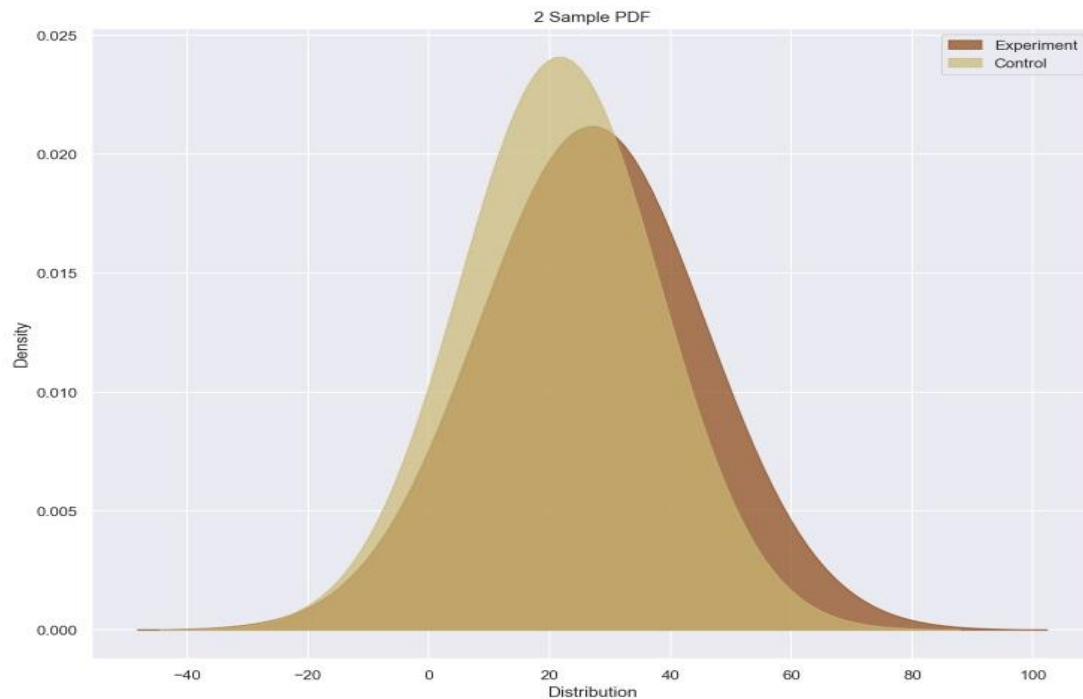


Simple Threshold: 24  
The midpoint between two means.

Overlap: 0.882  
The total AUC.

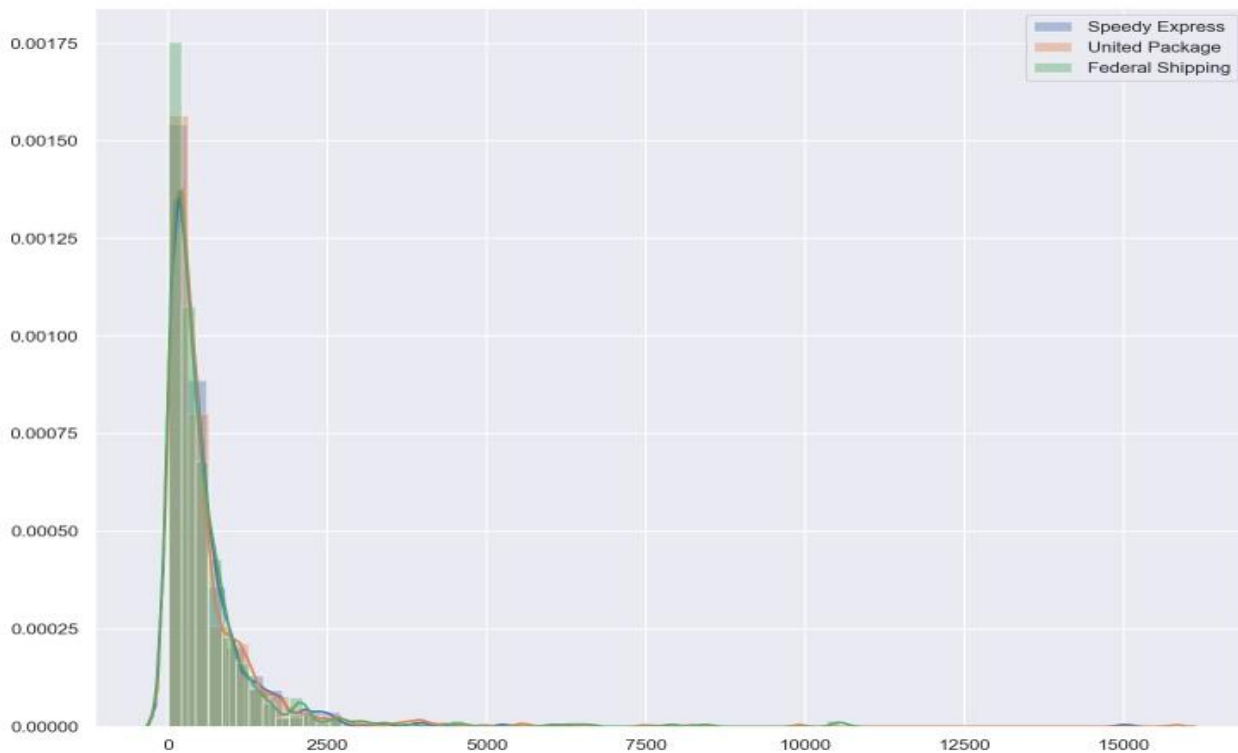
Superiority: 52.8  
Probability that a randomly chosen sample from the first group is [higher] than one of the second group.

Misclassification Rate: 0.441  
The chance of misclassification if using this metric alone as a predictor.



# Best Shipper-

Northwind Traders has a preferred shipper by total order \$\$ . They use one shipping company more than the others.



## Tukey Results

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	lower	upper	reject
Federal Shipping	Speedy Express	-54.4272	-180.9466	72.0922	False
Federal Shipping	United Package	23.1049	-95.1715	141.3813	False
Speedy Express	United Package	77.5321	-40.6919	195.7561	False

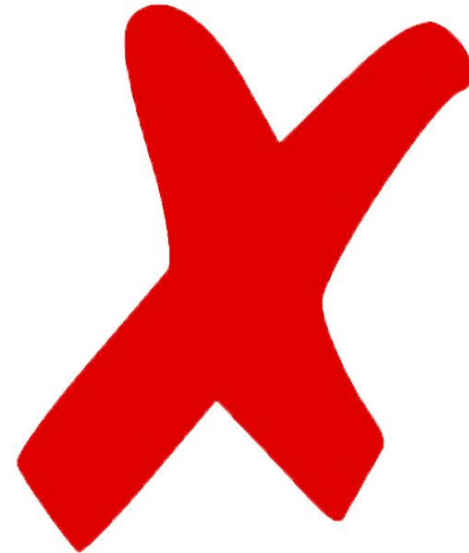
## Budget Shipper

Multiple Comparison of Means - Tukey HSD, FWER=0.05

=====					
group1	group2	meandiff	lower	upper	reject
-----					
Federal Shipping	Speedy Express	-0.0344	-0.1565	0.0877	False
Federal Shipping	United Package	-0.0256	-0.1397	0.0886	False
Speedy Express	United Package	0.0088	-0.1053	0.1229	False
-----					

# Summary

We fail to reject the Null Hypothesis. All shippers appear to handle the same amount of orders. We can say the same for price among shippers.

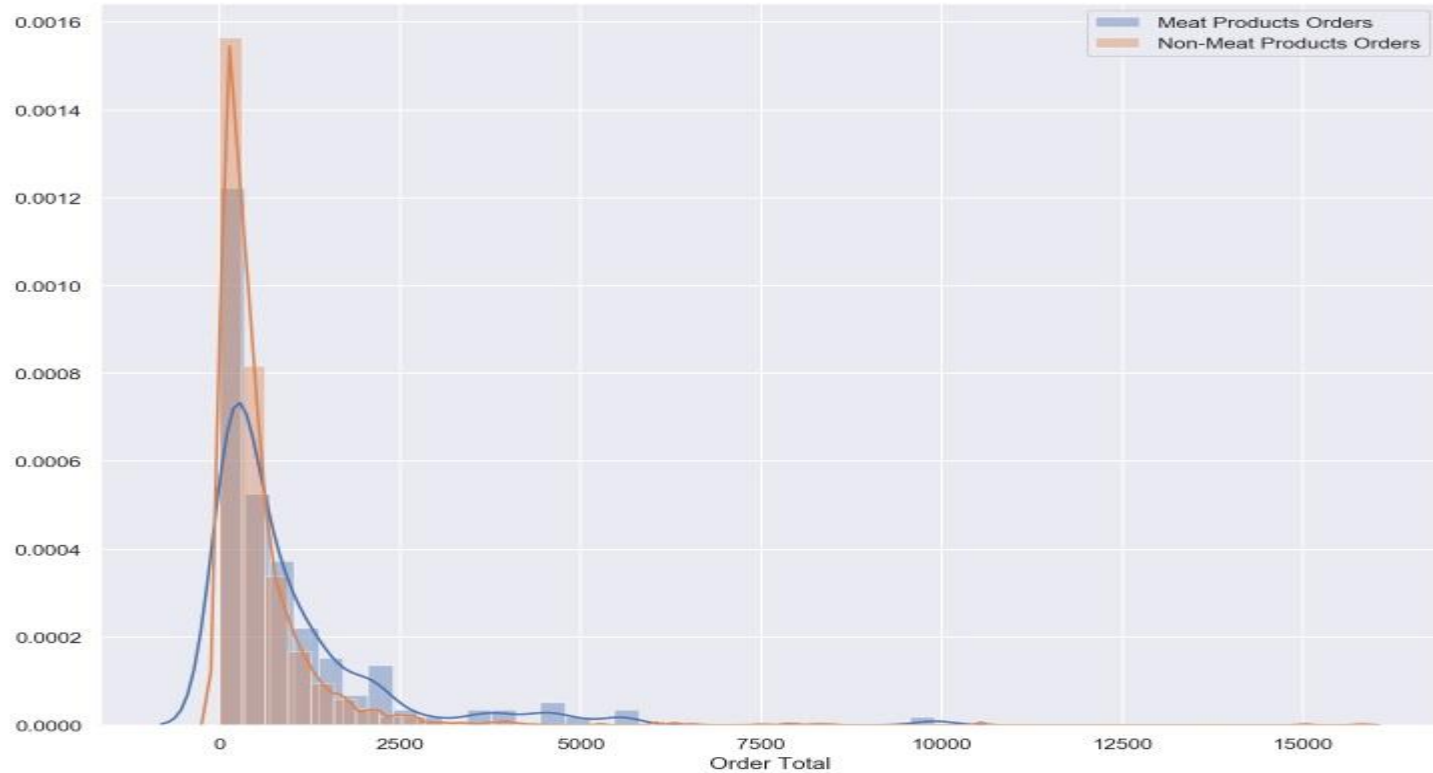


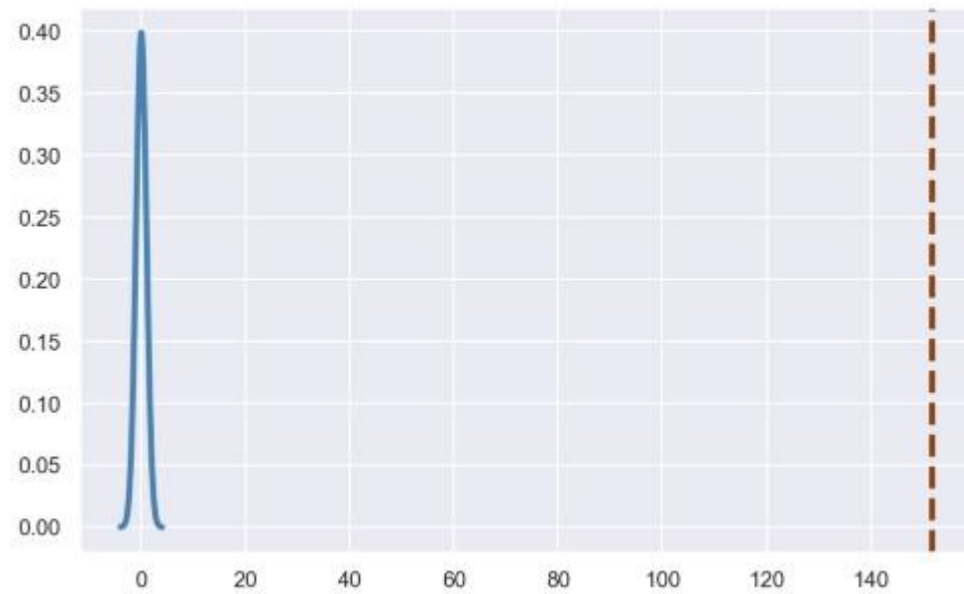
# One Category to Rule Them All

### H<sub>1</sub>: Meat/Poultry is the most lucrative category by total Unit Price per order being at least 100 dollars more than other categories with a confidence of 95%

### H<sub>0</sub>: There is no significant difference between the Meat/Poultry and the other products.

# Categories vs Meat Category





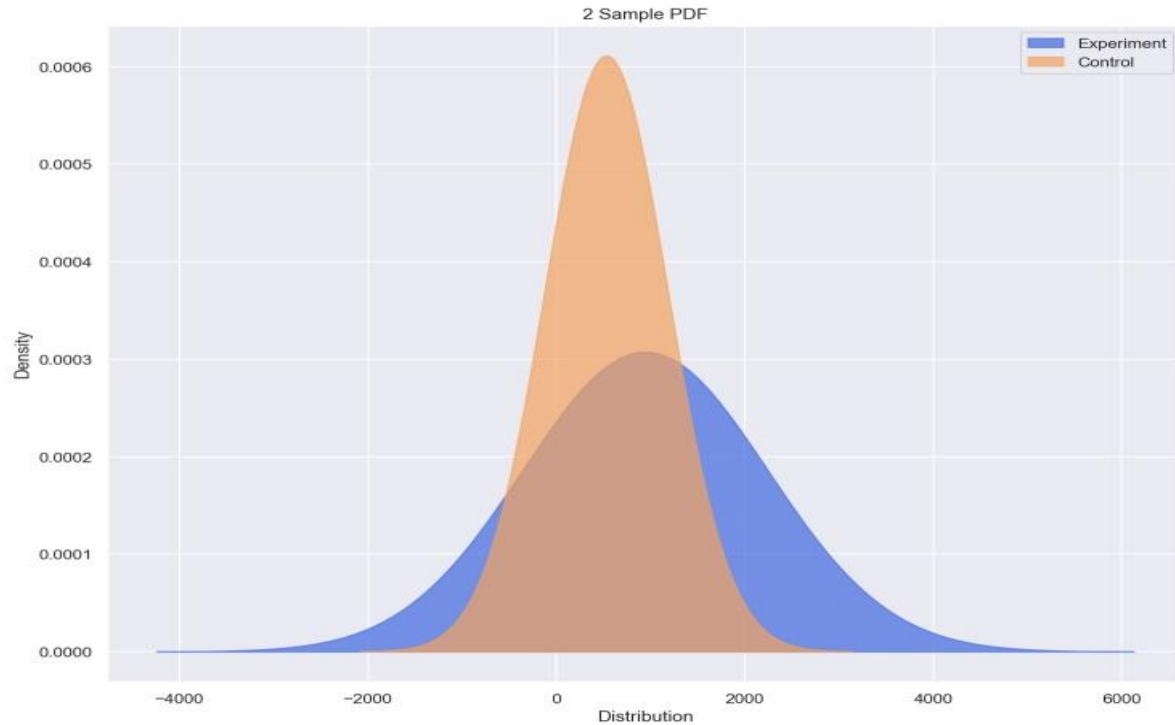


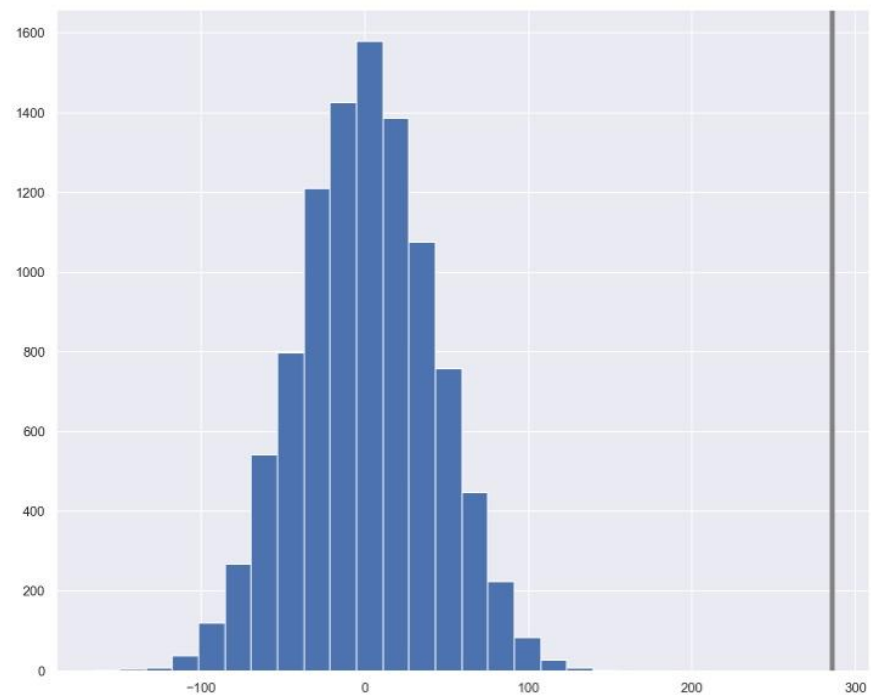
Simple Threshold: 736  
The midpoint between two means.

Overlap: 0.4127  
The total AUC.

Superiority: 31.33  
Probability that a randomly chosen sample from the first group is [higher] than one of the second group.

Misclassification Rate: 0.2063  
The chance of misclassification if using this metric alone as a predictor.





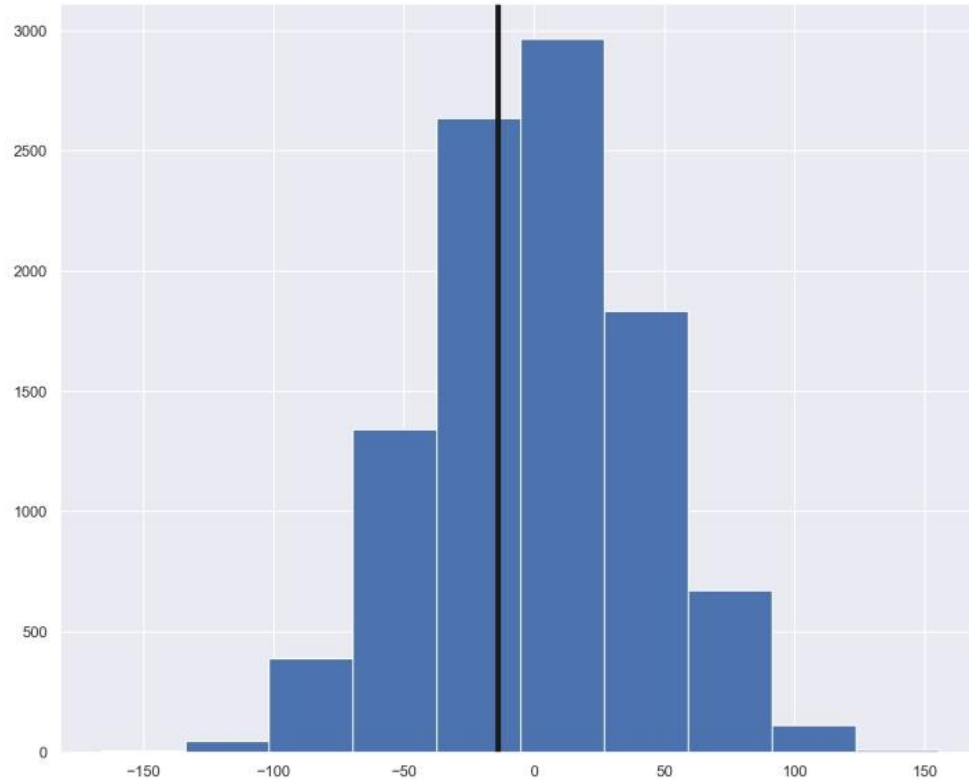
# Summary

There is enough evidence to support the claim that orders on meats are placed with a larger order price than other categories, even by \$100.

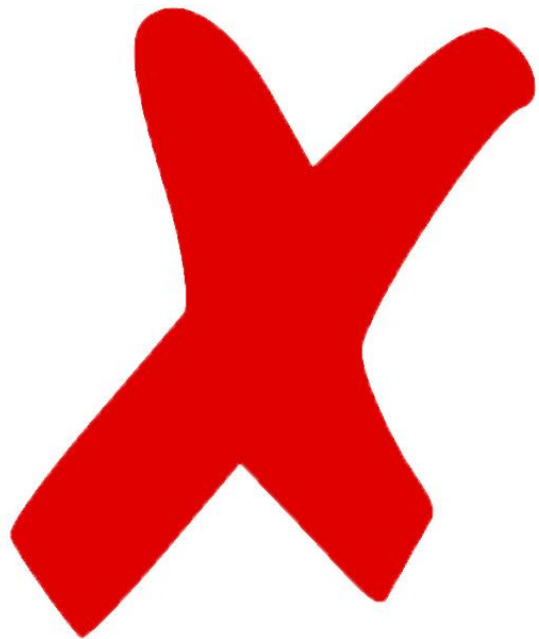
We fail to accept the null hypothesis and accept the Alternative that meat orders bring in more money.



# Are more meat products kept in stock?



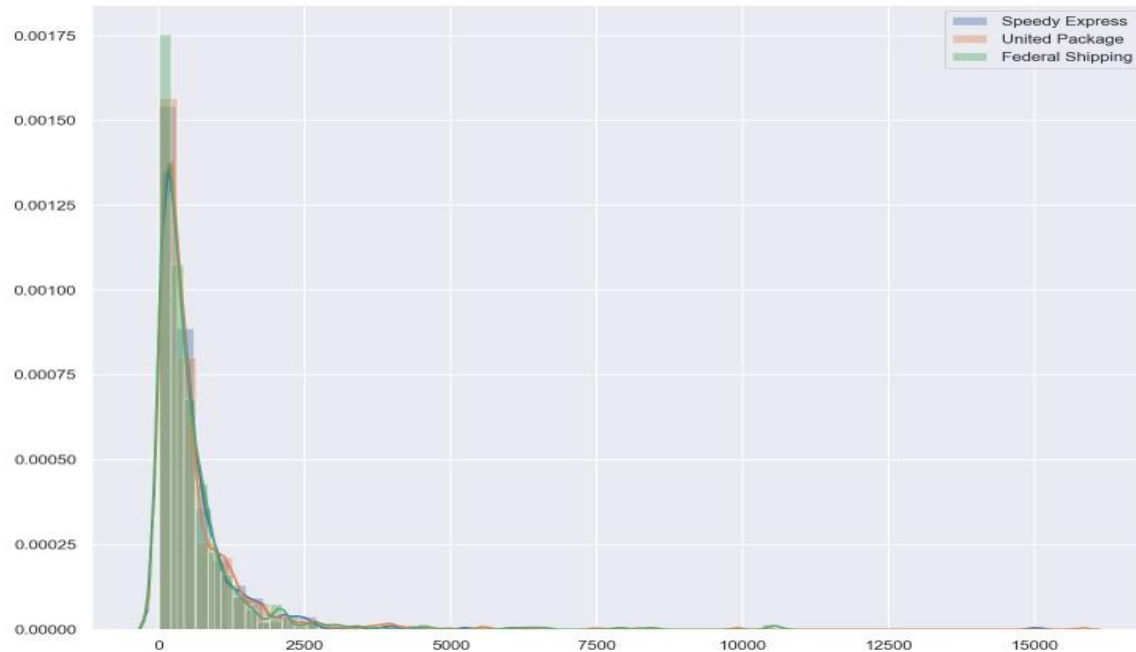
**Fail to reject the null**

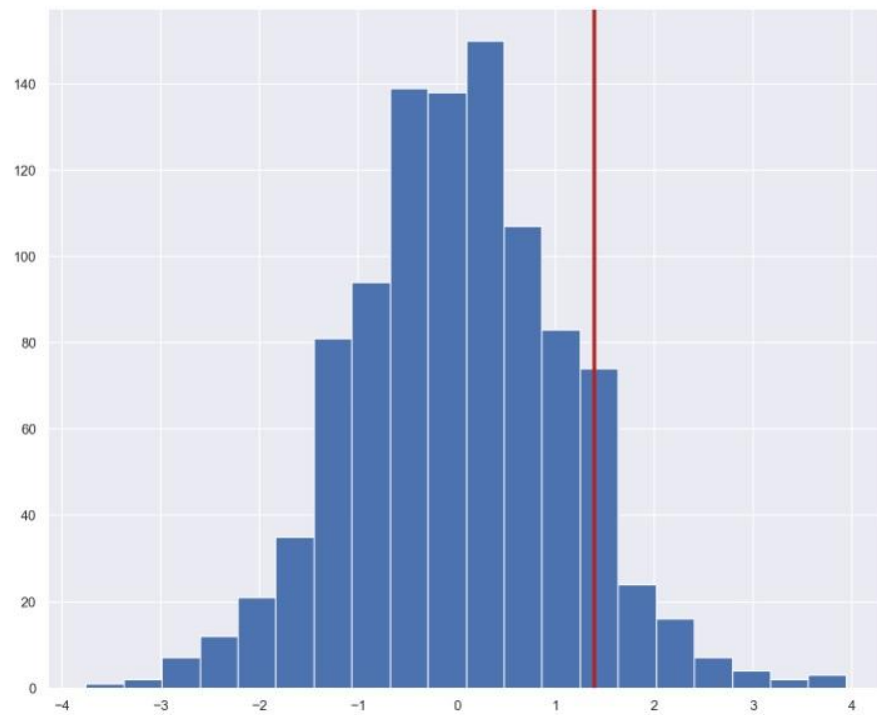


# Home Town -

**H1: The western region is the most profitable region by quantity of products shipped with a confidence of 95%**

**H0: There is no significant difference between the western region and the other regions.**





## Summary

This P value is a little high. Our informal test tells us that we are only shipping more products per order to the west 85% of the time. Nor can we prove that 80% of the product in the west is shipped to San Francisco.





# Hypothesis



**Discounts orders ship a higher quantity of products than those without.**



**Northwind Traders has a preferred shipper by total order \$\$\$. They use one shipping company more than the others.**



**Meat Orders are \$100 more than other categories on average.**



**San Francisco ships 80% of the western orders by quantity.**

# Null-Hypothesis



**There is no difference in between orders with discounts**



**There is a preferred shipper that costs less per order than other shippers.**



**Meat is stocked more than other categories**



**The West Region earns more \$\$\$ than other regions.**

# **The Process**

**Ask the Question**

**Tailor the Query**

**Test the Hypothesis**

**Prove Yourself Wrong**

**Use Your Findings to Inform Your Next Question**

# Northwind Trades

---