Business Insights and Recommendations using Statistical Significance Testing

Data Science Consultation for Northwind Traders

Overview of Business Insights: 5 Wishes

I wish Northwind Traders could...

1. Increase sales quantities

2. Motivate the sales team

3. Develop the European market

4. Better distribute the sales team's workload

5. Improve the supply chain

1.1 Discounts and Quantity

Discount v.s. No Discount Average Quantity of Product Per Order

Question:

Do applying discounts impact the quantity of the sale?

Results:

Providing a discount does increase the amount of quantity in the order.

EDA shows higher average sales quantities with a discount

1 Sample T-test, T-value = 7.52, Critical t-value = 1.647, P-value = 0.0, Null hypothesis rejected.

1.2 Discounts and Quantity

Average Quantity With and Without Discounts

Legend: 5%, 10%, 15%, 20%, 25%
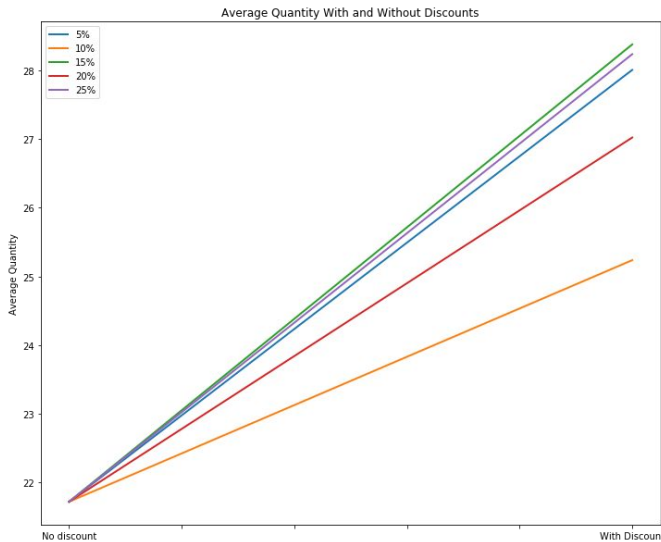
Question:

Do different discount levels impact quantity?

Results:

The impact of discounts on quantity vary by discount level, with a %5 discount being the most impactful on quantity.

ANOVA test to compare each level of discount to the average.
5%, 15%, 25%, 20%, 10%

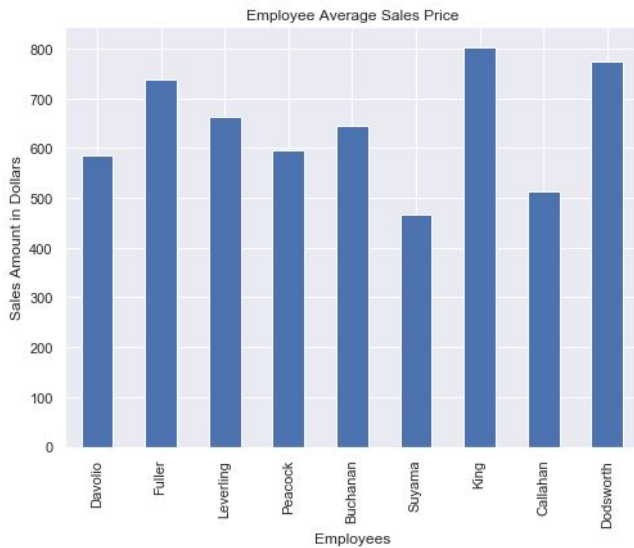# 1.1/1.2 Discounts and Quantity - Business Recs

Wish #1: Increase sales quantities.

Recommendations:

- Find more opportunities to provide customers with 5% and 15% discounts (ad campaigns, email blasts, customer reward programs)

2. Employee Sales

Employee Average Sales Price

Question:

Are there any employees who are averaging higher sales prices per sale?

Results:

There are no significant differences between employees average sales price.

Kruskal-Wallis H-test, T-value = 0.511, P-value = 0.231, Fail to reject the null hypothesis.
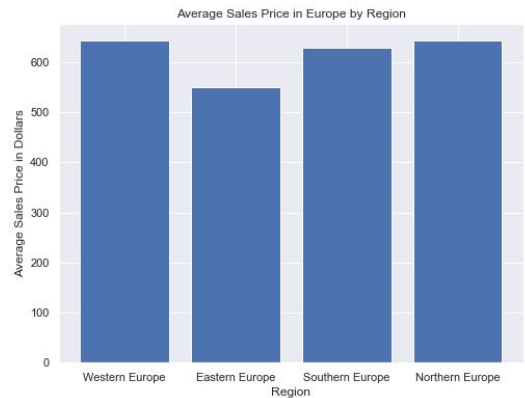
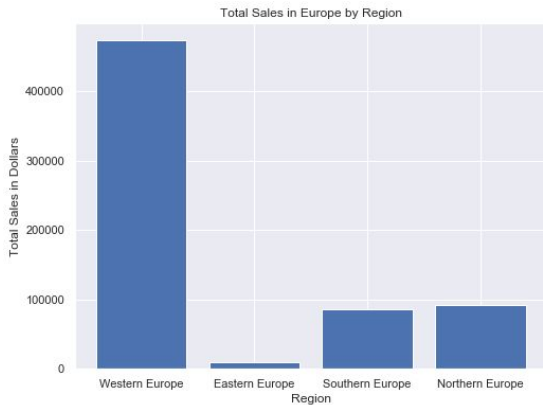# 2. Employee Sales - Business Recs

Wish #2: Motivate our sales team.

Recommendations:

- Use average sale by employee (~$628) as a marker for success, rewarding employees who increase their average over time (commissions, extra PTO, special events)

3. The European Market

Total Sales in Europe by Region

Average Sales Price in Europe by Region

Question:

Are any of the European regions averaging higher sales prices?

Results:

There are no significant differences in average sales price between regions.

2 Sample T-tests between individual regions and all of Europe, Null hypothesis rejected for each test.

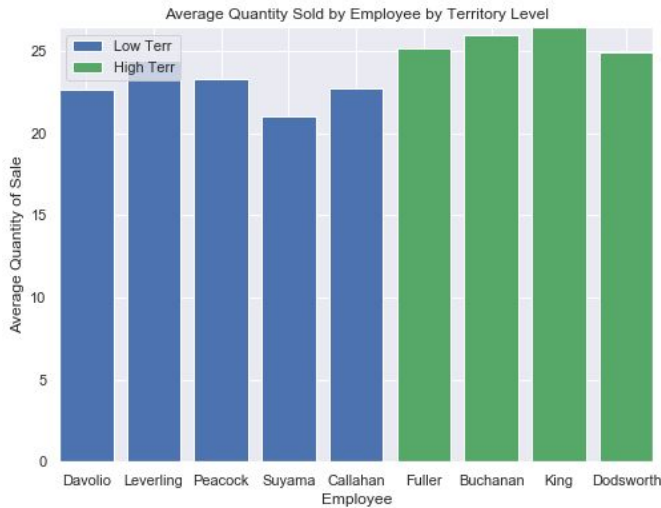# 3. European Market - Business Recs

Wish #3: Develop the European market

<u>Recommendations:</u>

- Focus ad campaigns in eastern, northern and southern European regions.
- Focus on further development western Europe.

4. The Distribution of Sales Territories

Average Quantity Sold by Employee by Territory Level

Question:

Do employees with a high number of sales regions sell higher quantities?

Results:

There are significant differences in employee sales quantities depending on territory amounts.

Are regions distributed evenly? Some employees may have a higher number of less productive regions whereas some employees may have a lower number of very productive regions.
2 Sample T-test, T Stat = -6.960, p = 0.000, Null hypothesis rejected.

4. The Distribution of Sales Territories - Business Recs

*Wish #4: Better distribute our sales team's workload*

Recommendations:

- Develop young and promising employees by adding sales regions to their responsibilities.

Give employees with lower work loads (territories to cover) the added territories as the company expands. This will drive up quantity of sale and distribute responsibility evenly.

5.1 Improve the Supply Chain

Average Number of Items on Backorder by Product Category
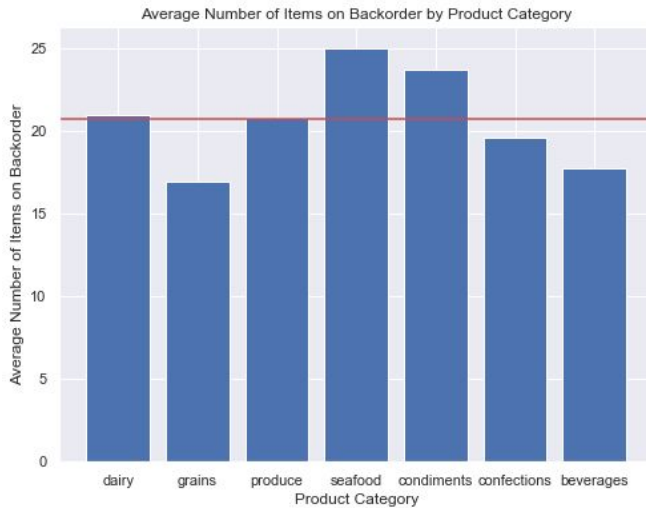
Question:

Do certain product categories have a higher number of products going on backorder?

Results:

***Backordering varies by product, but no one category is backordered more than average.***

Are more dairy products backordered because of a short shelf life? Deeper testing can provide insights into warehouse storage efficiency, supplier demand and the time it takes to get product to a customer.

Analysis 1: Kruskall-Wallis H-test, Null hypothesis rejected.

Analysis 2:, 2 Sample T-test, Fail to reject the null hypothesis.

5.2 Improve the Supply Chain

Non backordered v.s. Backordered Processing Time
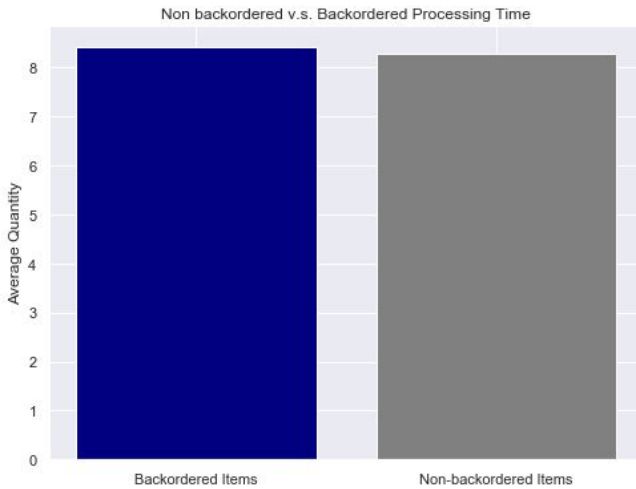
Question:

Do items on backorder take longer to get to customers?

Results:

Items that go on backorder do not significantly impact the average processing time of the order.

Data showed non-normality, non-parametric tests needed, Mann-Whitney U Test, Statistics=512383.000, p=0.208, Same distribution (fail to reject H0)

5.1/5.2 Improve the Supply Chain - Business Recommendations

Wish #5: Improve our supply chain

Recommendations:

- Further investigate how often items are backordered for future growth

Future Work

- Questions focusing heavily on supply chain

- Predictive modeling for new markets

- Need for deeper analysis of employee sales

(1) Discrepancies found in the reorder level, quantity, units in stock data shows further investigation is needed to develop the supply chain. This work will become increasingly more important as the company grows and expands into further markets.
(2) Regression models could predict potential in new markets
(3) With small employee numbers, growth and distribution of responsibilities will be an important factor
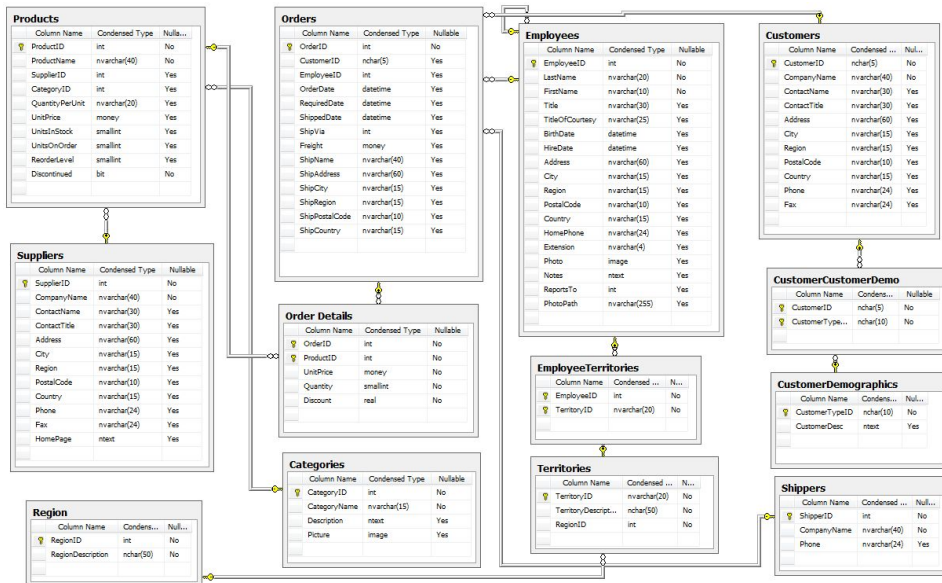
Thank You

www.lucaskellydataportfolio.weebly.com
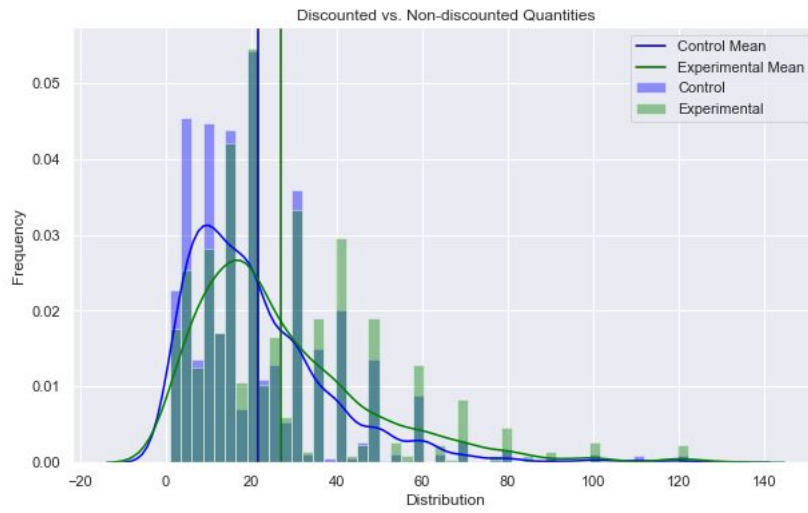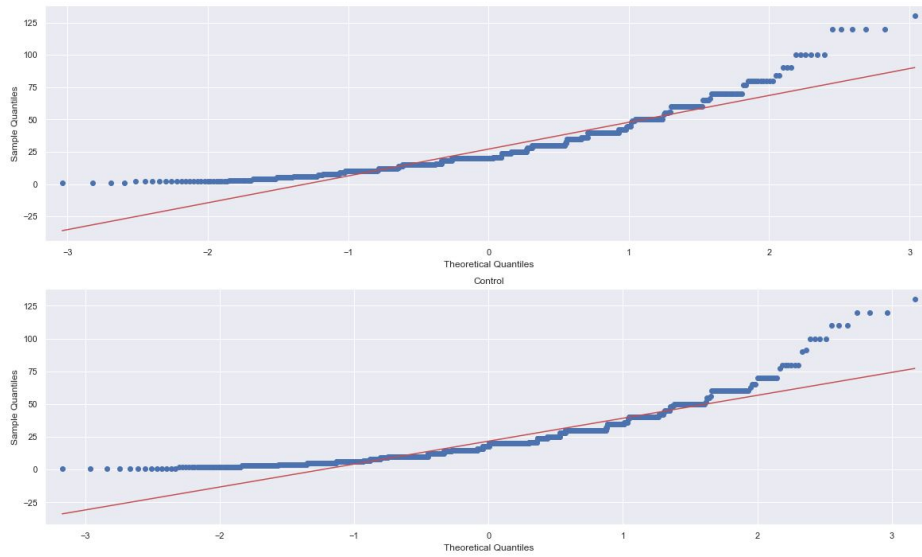
www.github.com/lucaskelly49

https://www.linkedin.com/in/lucaskelly49/

Index ERD

Products

Column Name	Condensed Type	Nulla...
ProductID	int	No
ProductName	nvarchar(40)	No
SupplierID	int	Yes
CategoryID	int	Yes
QuantityPerUnit	nvarchar(20)	Yes
UnitPrice	money	Yes
UnitsInStock	smallint	Yes
UnitsOnOrder	smallint	Yes
ReorderLevel	smallint	Yes
Discontinued	bit	No

Suppliers

Column Name	Condensed Type	Nullable
SupplierID	int	No
CompanyName	nvarchar(40)	No
ContactName	nvarchar(30)	Yes
ContactTitle	nvarchar(30)	Yes
Address	nvarchar(60)	Yes
City	nvarchar(15)	Yes
Region	nvarchar(15)	Yes
PostalCode	nvarchar(10)	Yes
Country	nvarchar(15)	Yes
Phone	nvarchar(24)	Yes
Fax	nvarchar(24)	Yes
HomePage	ntext	Yes

Region

Column Name	Condens...	Null...
RegionID	int	No
RegionDescription	nchar(50)	Yes

Orders

Column Name	Condensed Type	Nullable
OrderID	int	No
CustomerID	nchar(5)	Yes
EmployeeID	int	Yes
OrderDate	datetime	Yes
RequiredDate	datetime	Yes
ShippedDate	datetime	Yes
ShipVia	int	Yes
Freight	money	Yes
ShipName	nvarchar(40)	Yes
ShipAddress	nvarchar(60)	Yes
ShipCity	nvarchar(15)	Yes
ShipRegion	nvarchar(15)	Yes
ShipPostalCode	nvarchar(10)	Yes
ShipCountry	nvarchar(15)	Yes

Order Details

Column Name	Condensed Type	Nullable
OrderID	int	No
ProductID	int	No
UnitPrice	money	No
Quantity	smallint	No
Discount	real	No

Categories

Column Name	Condensed Type	Nullable
CategoryID	int	No
CategoryName	nvarchar(15)	No
Description	ntext	Yes
Picture	image	Yes

Employees

Column Name	Condensed Type	Nullable
EmployeeID	int	No
LastName	nvarchar(20)	No
FirstName	nvarchar(10)	No
Title	nvarchar(30)	Yes
TitleOfCourtesy	nvarchar(25)	Yes
BirthDate	datetime	Yes
HireDate	datetime	Yes
Address	nvarchar(60)	Yes
City	nvarchar(15)	Yes
Region	nvarchar(15)	Yes
PostalCode	nvarchar(10)	Yes
Country	nvarchar(15)	Yes
HomePhone	nvarchar(24)	Yes
Extension	nvarchar(4)	Yes
Photo	image	Yes
Notes	ntext	Yes
ReportsTo	int	Yes
PhotoPath	nvarchar(255)	Yes

EmployeeTerritories

Column Name	Condensed ...	N...
EmployeeID	int	No
TerritoryID	nvarchar(20)	No

Territories

Column Name	Condensed ...	N...
TerritoryID	nvarchar(20)	No
TerritoryDescript...	nchar(50)	No
RegionID	int	No

Customers

Column Name	Condensed ...	Nul...
CustomerID	nchar(5)	No
CompanyName	nvarchar(40)	No
ContactName	nvarchar(30)	Yes
ContactTitle	nvarchar(30)	Yes
Address	nvarchar(60)	Yes
City	nvarchar(15)	Yes
Region	nvarchar(15)	Yes
PostalCode	nvarchar(10)	Yes
Country	nvarchar(15)	Yes
Phone	nvarchar(24)	Yes
Fax	nvarchar(24)	Yes

CustomerCustomerDemo

Column Name	Condens...	Nullable
CustomerID	nchar(5)	No
CustomerType...	nchar(10)	No

CustomerDemographics

Column Name	Condens...	Nul...
CustomerTypeID	nchar(10)	No
CustomerDesc	ntext	Yes

Shippers

Column Name	Condensed ...	Null...
ShipperID	int	No
CompanyName	nvarchar(40)	No
Phone	nvarchar(24)	Yes

Index Q1. Dist

Discounted vs. Non-discounted Quantities
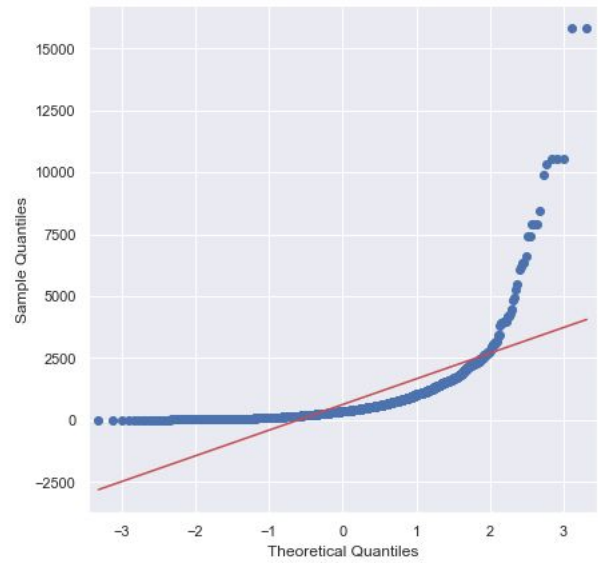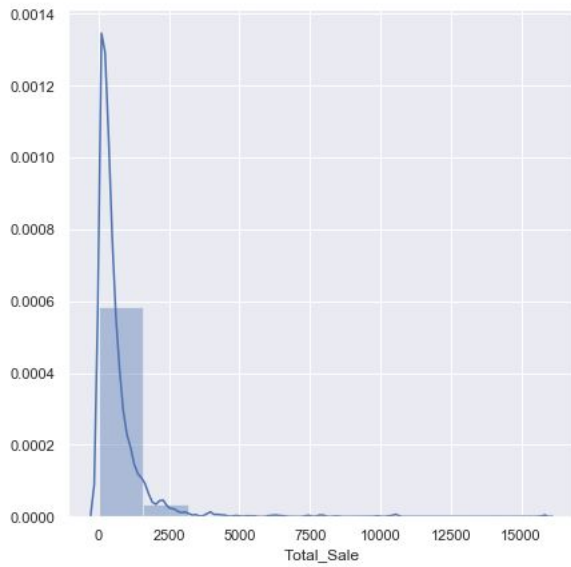
Index Q1 - Normality check, control/experimental

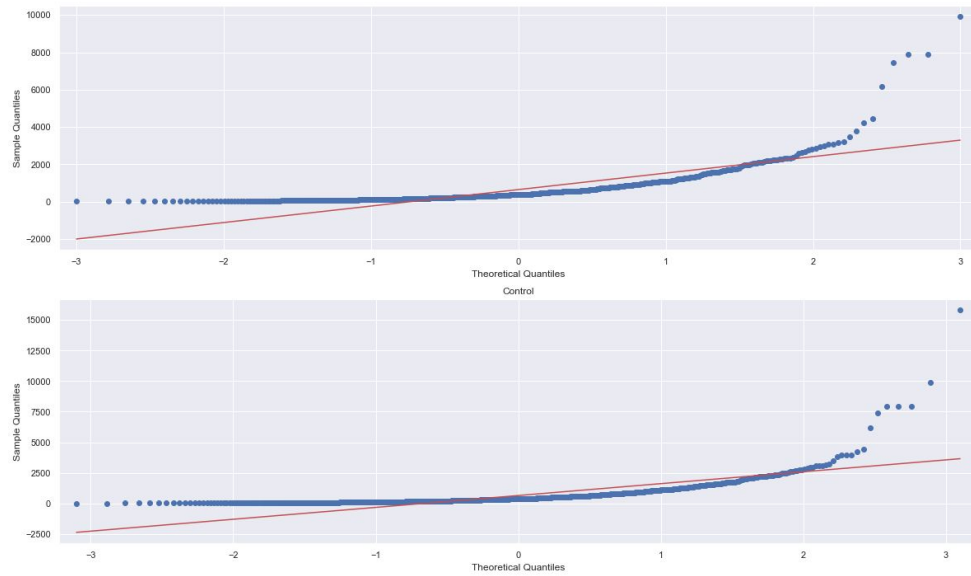Index Q2 - ANOVA Test Results

```
                  PR(>F)
C(five)          0.000015
C(ten)           0.016761
C(fifteen)       0.000020
C(twenty)        0.000558
C(twentyfive)    0.000035
Residual            NaN
```
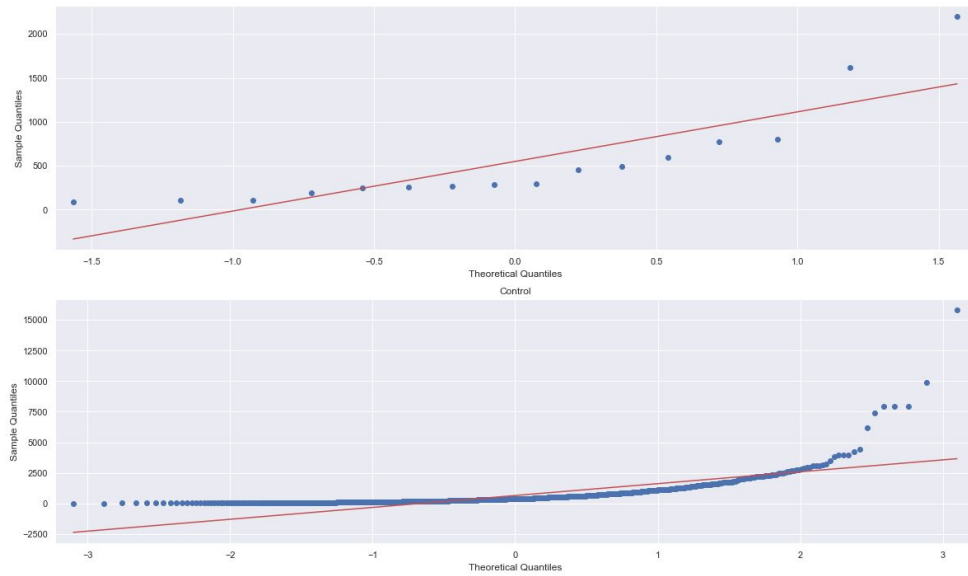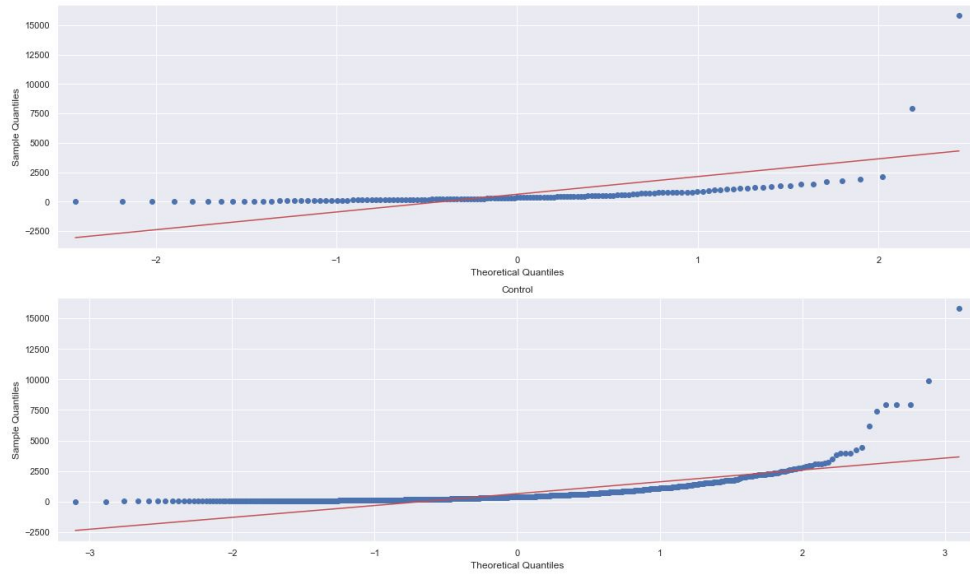
Index Q2 - Normality Check
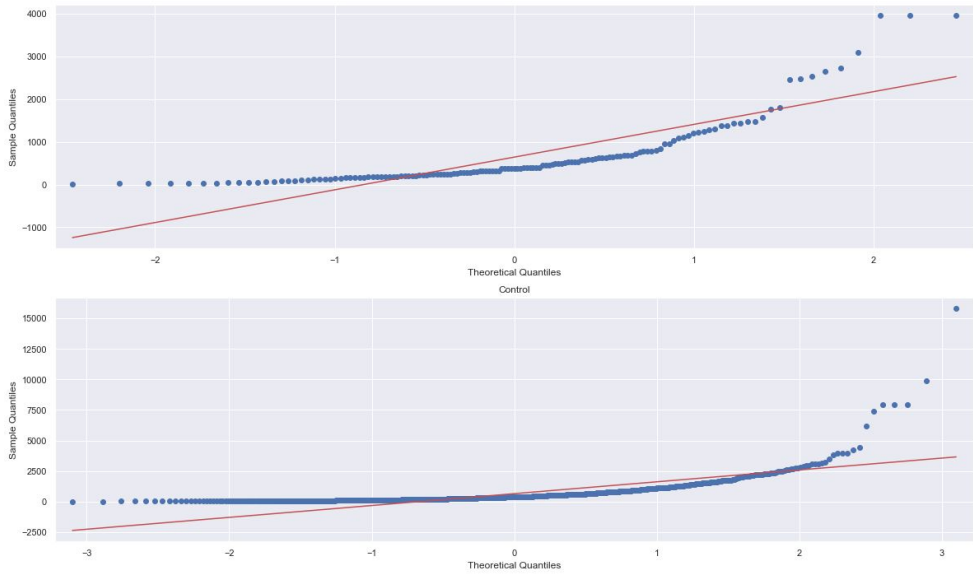
Index Q3 - Western Europe vs Europe Normality Check
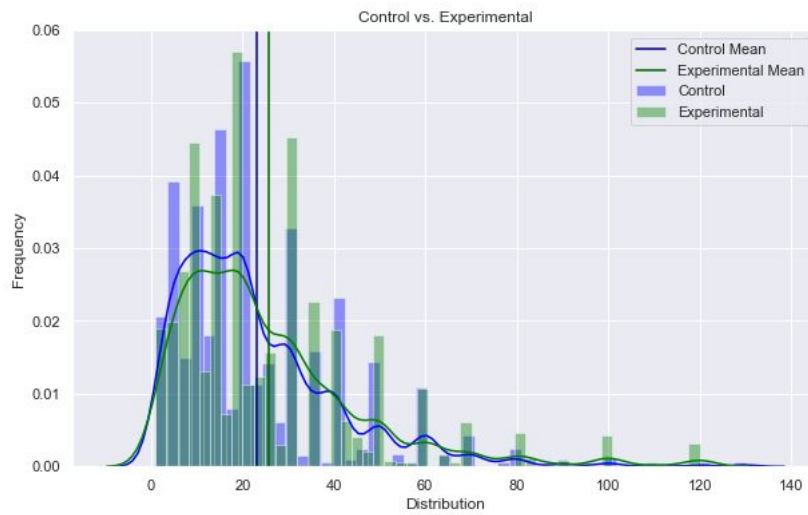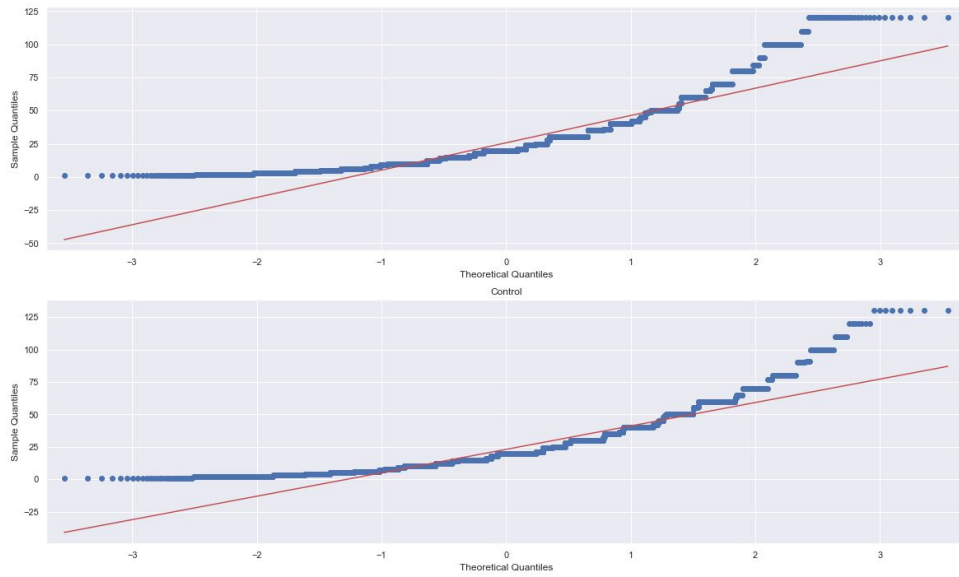
Index Q3 - Eastern Europe vs Europe Normality Check

Control

Index Q3 - Southern Europe vs Europe Normality Check

Control

Index Q3 - Northern Europe vs Europe Normality Check

Control

Index Q4 - Low Territory/High Territory Distribution

Control vs. Experimental

Index Q4 - Low Territory/High Territory Normality Check

Index Q5 - Dairy Backorder vs. Average Normality Check

Control
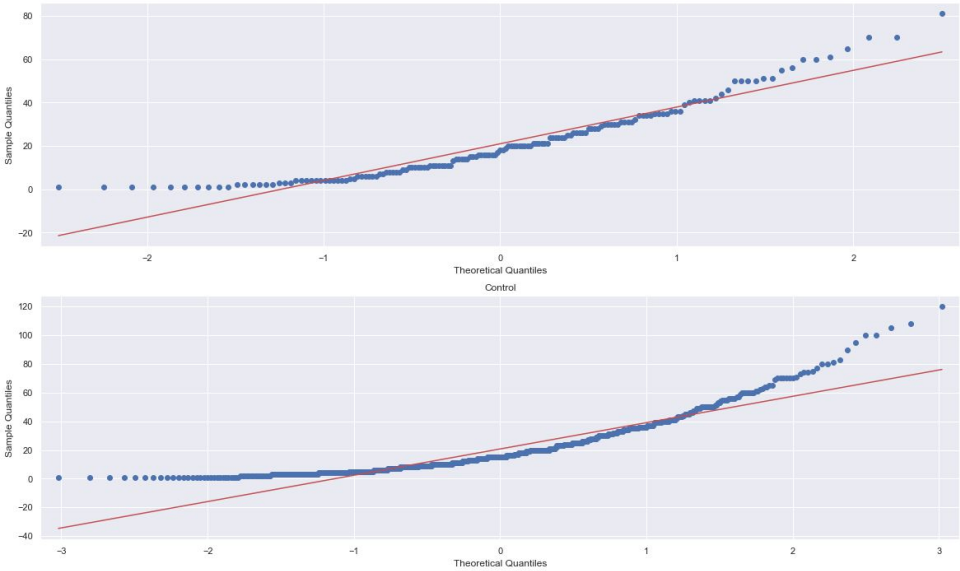
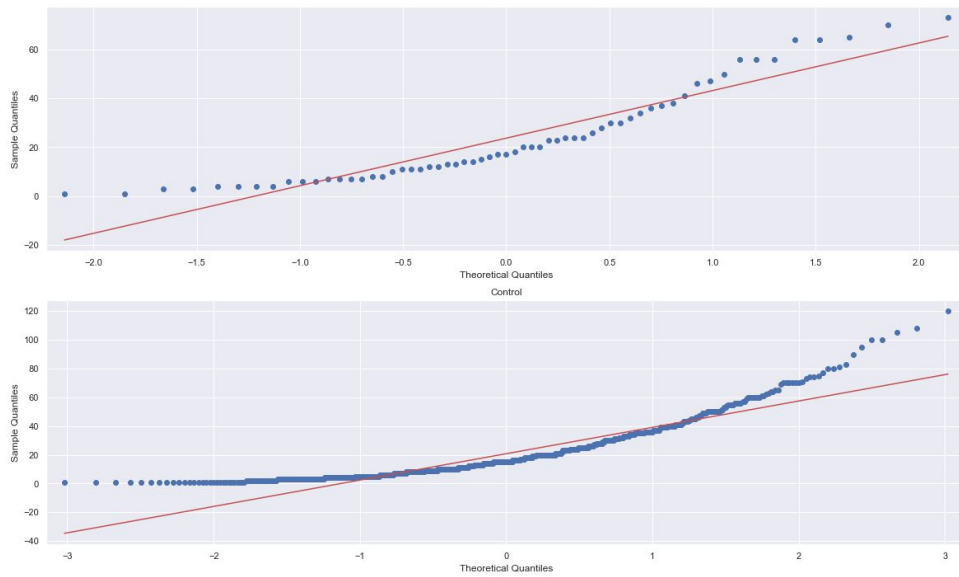# Index Q5 - Grains Backorder vs. Average Normality Check
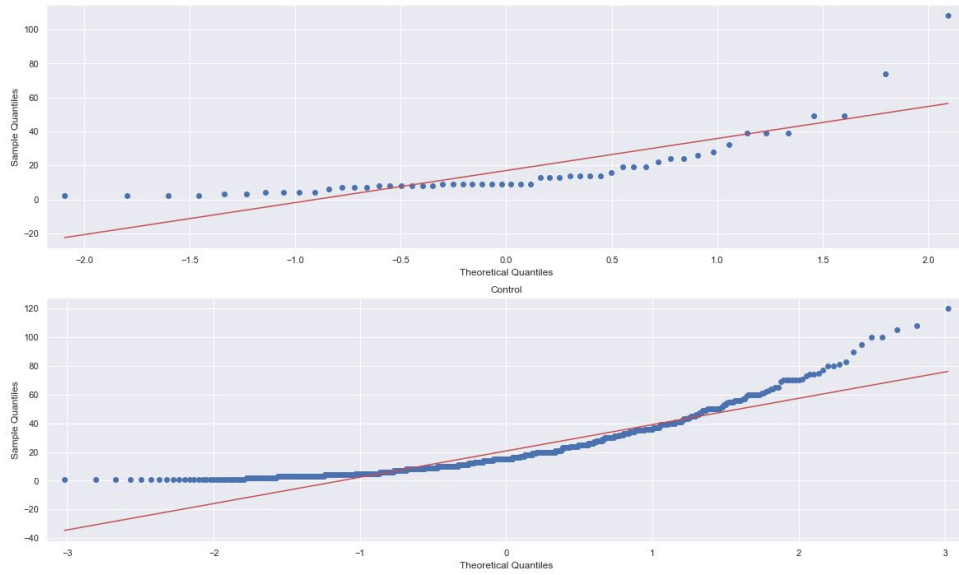
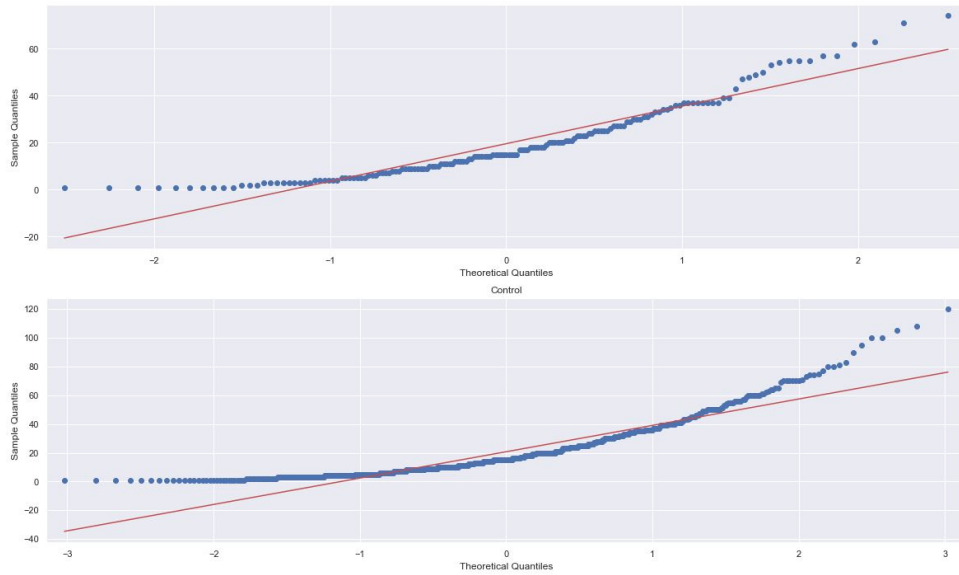# Index Q5 - Produce Backorder vs. Average Normality Check
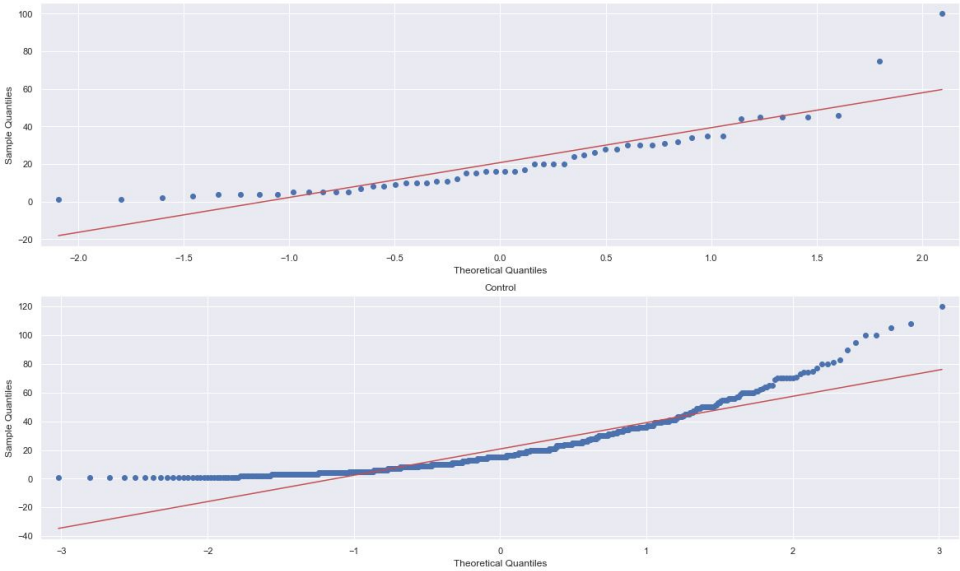
Index Q5 - Seafood Backorder vs. Average Normality Check
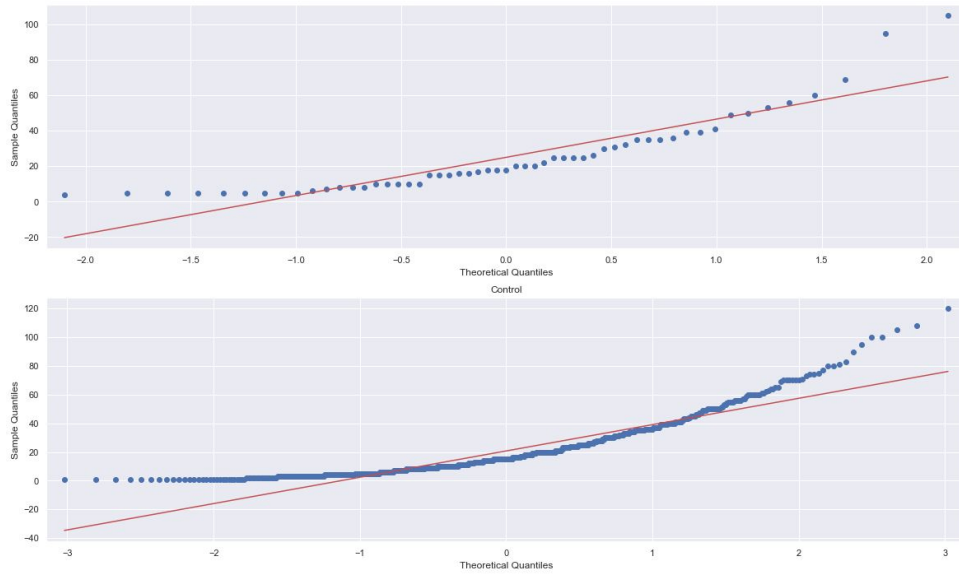
Index Q5 - Condiments Backorder vs. Average Normality Check

Index Q5 - Confections Backorder vs. Average Normality Check

Control

Index Q5 - Beverages Backorder vs. Average Normality Check

Control