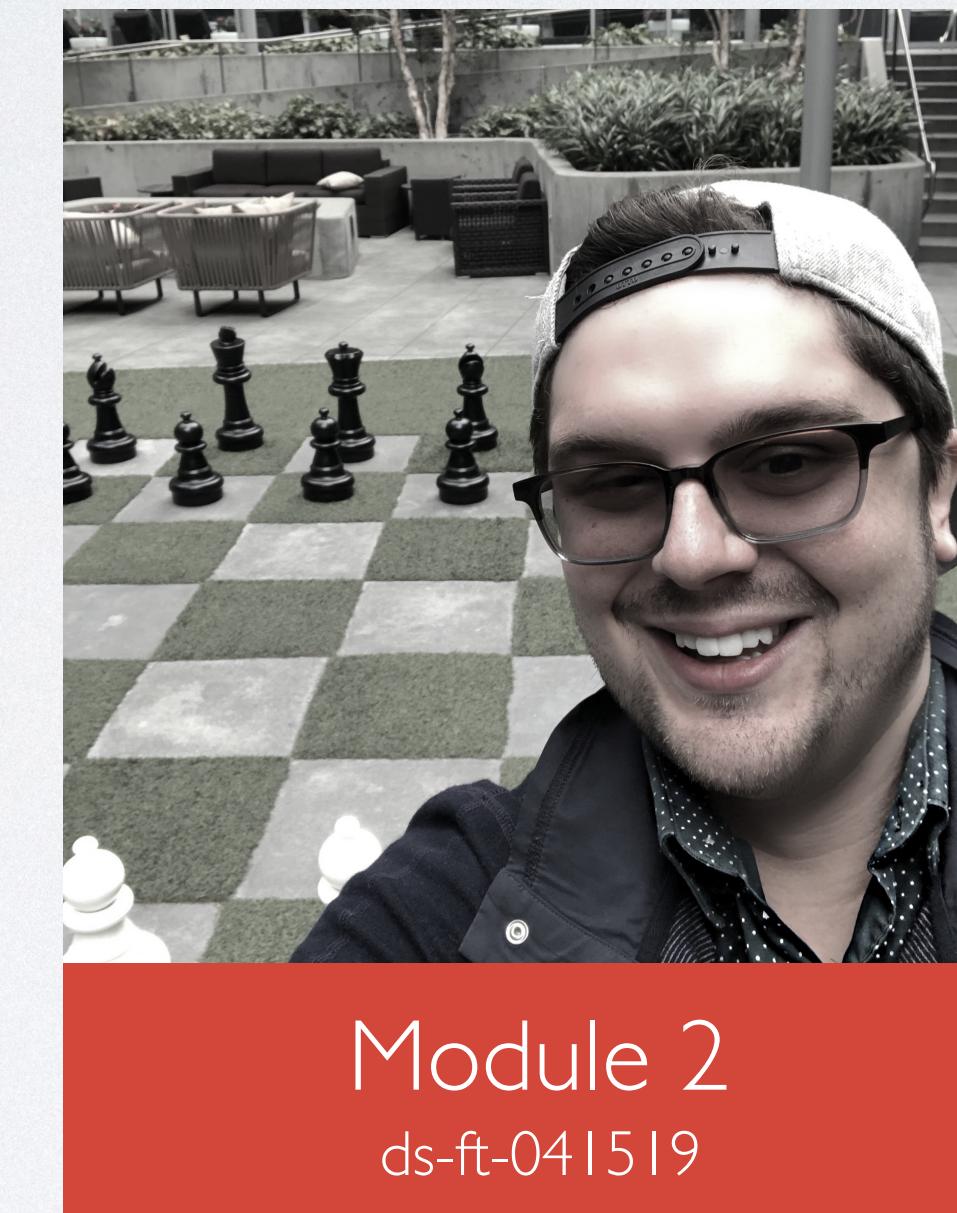


# NORTHWIND DATASET ANALYSIS



Module 2  
ds-ft-041519

Paul Woody

# NORTHWIND Q1 SALES REVIEW

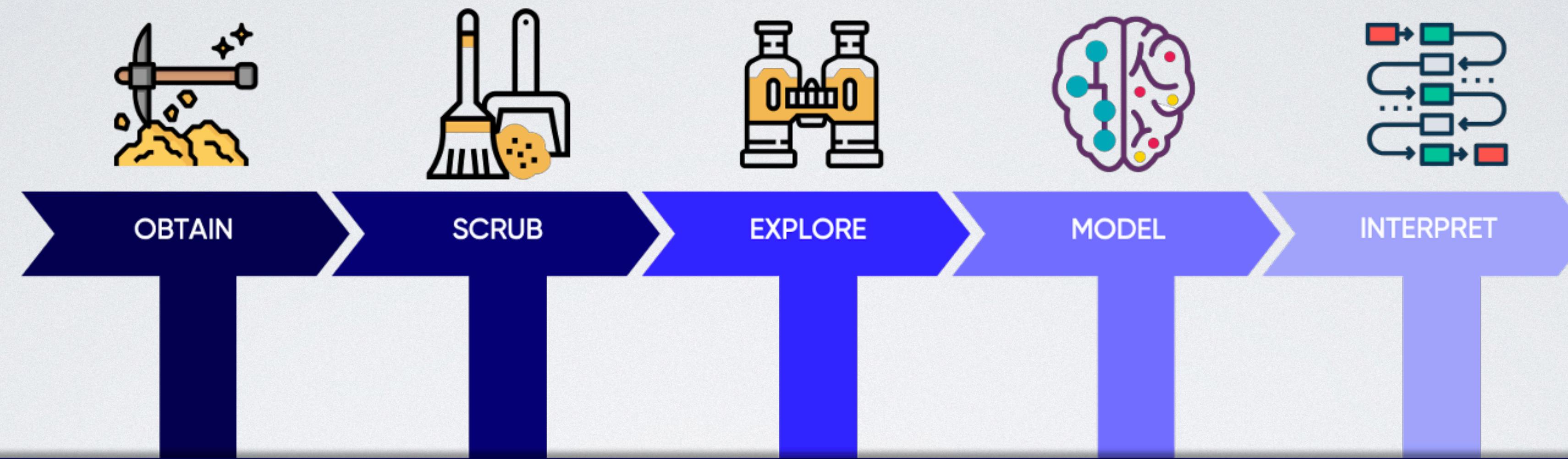
YTD Sales

Discount  
Strategy

Product  
Performance

Sales Team EPR

# METHODOLOGY: OSEMN



O

Gather data from relevant sources

S

Clean data to formats that machine understands

E

Find significant patterns and trends using statistical methods

M

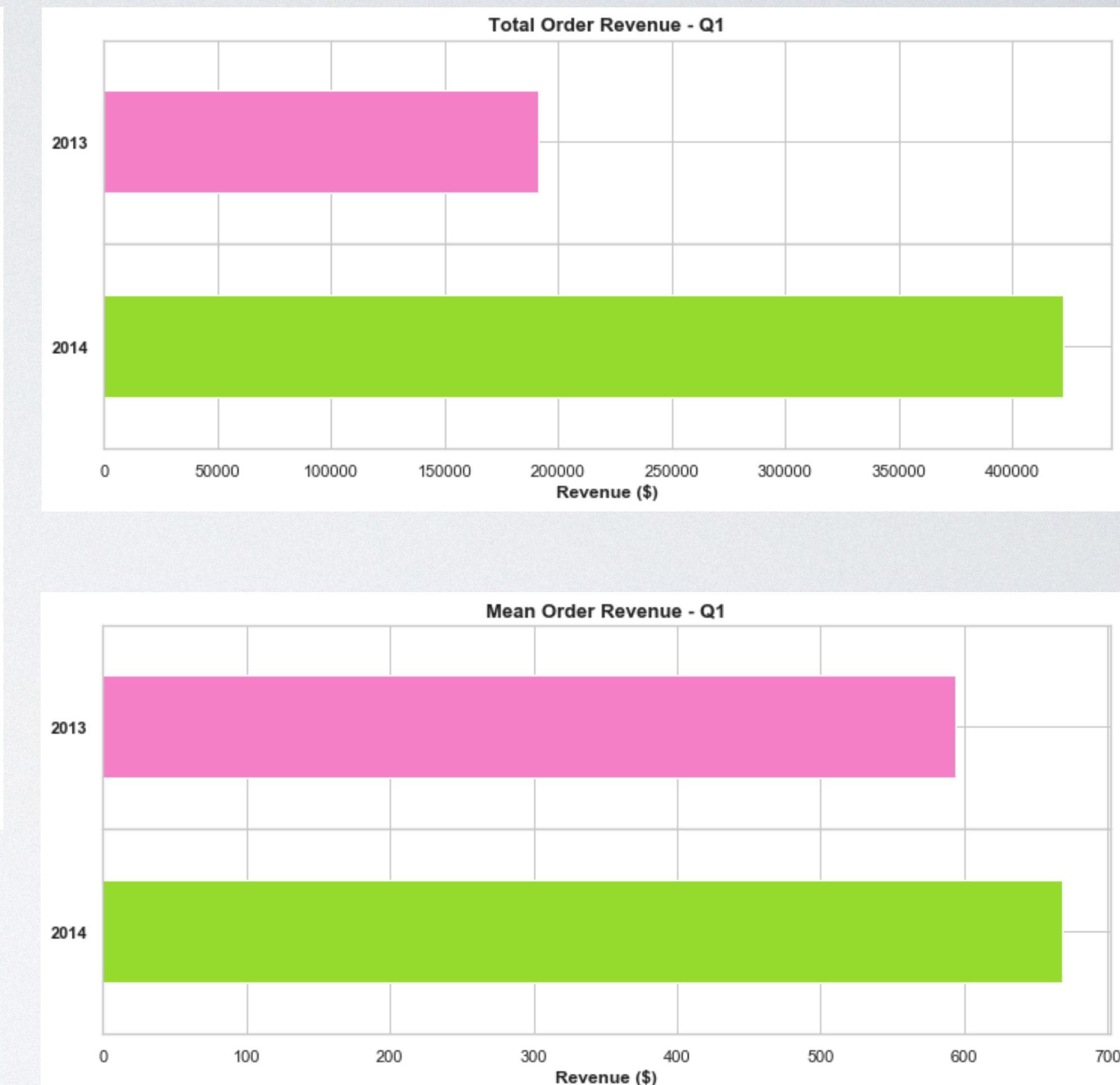
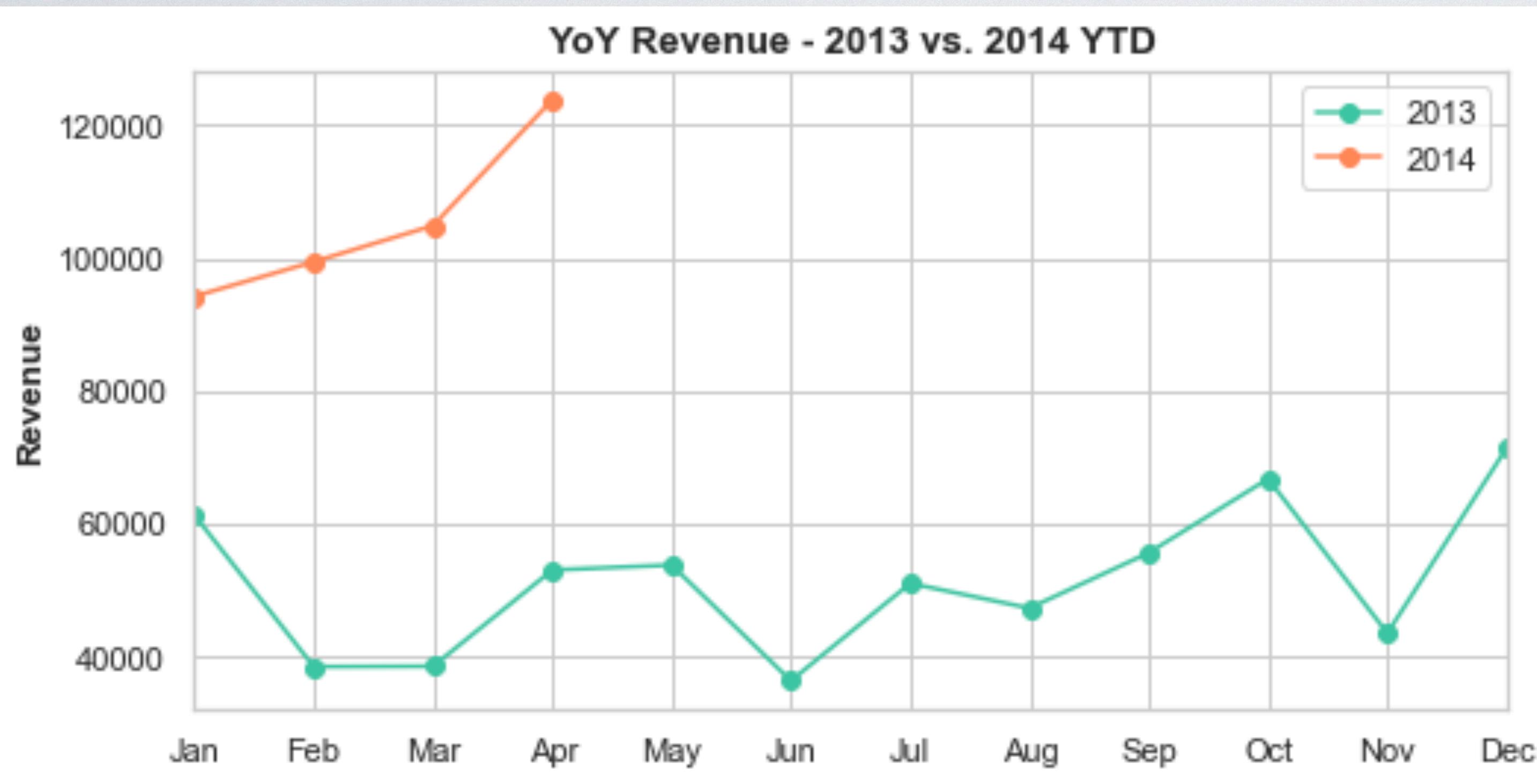
Construct models to predict and forecast

N

Put the results into good use

Perform exploratory data analysis to identify meaningful trends.  
Use models and appropriate hypothesis tests to evaluate significance.

# YOY REVENUE: 2013 - 2014

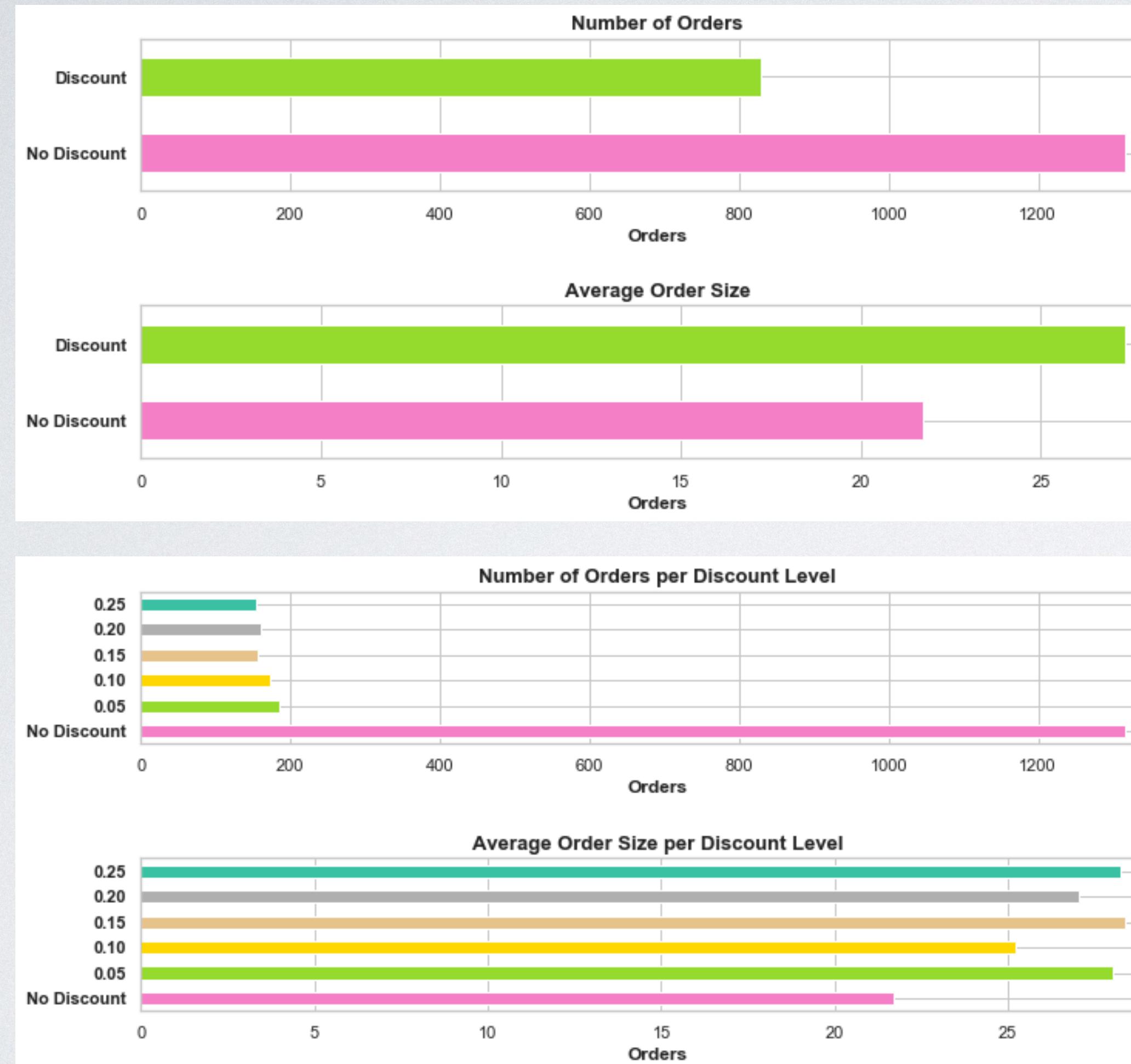


- $H_0$ : Average Order Revenue for Q1 2014 is not greater than that of 2013.
- $H_a$ : Average Order Revenue for Q1 2014 is greater than that of 2013.
- $\alpha = 0.05$

The p-value = 0.04. Because our p-value < 0.05, we reject the null hypothesis that Average Order Revenue for Q1 2014 is not greater than that of 2013.



# DISCOUNTED PRODUCT PERFORMANCE



Tukey Test - Discount Level vs. Order Quantity

Multiple Comparison of			Means	-	Tukey HSD,FWER=0.05
group1	group2	meandiff	lower	upper	reject
0.0	0.05	6.2955	2.081	4	10.5097 True
0.0	0.1	3.5217	-0.818	7	7.8622 False
0.0	0.15	6.6669	2.135	2	11.1986 True
0.0	0.2	5.3096	0.828	5	9.7907 True
0.0	0.25	6.525	1.954		11.096 True
0.05	0.1	-2.7738	-8.45	4	2.9028 False
0.05	0.15	0.3714	-5.452	8	6.1955 False
0.05	0.2	-0.986	-6.77	8	4.7989 False
0.05	0.25	0.2294	-5.625	3	6.0842 False
0.1	0.15	3.1452	-2.77	1	9.0613 False
0.1	0.2	1.7879	-4.089	6	7.6653 False
0.1	0.25	3.0033	-2.94	3	8.9496 False
0.15	0.2	-1.3573	-7.377	5	4.6628 False
0.15	0.25	-0.1419	-6.229	2	5.9454 False
0.2	0.25	1.2154	-4.834	3	7.2652 False

- $H_0$ : Discount does not affect order quantity.
- $H_a$ : Discount does affect order quantity.
- $\alpha = 0.05$
- Test(s) Used: Independent two-tailed t-test

- Significant for all discount levels in comparison to control apart from discount level 0.10.

# PRODUCT CATEGORY PERFORMANCE

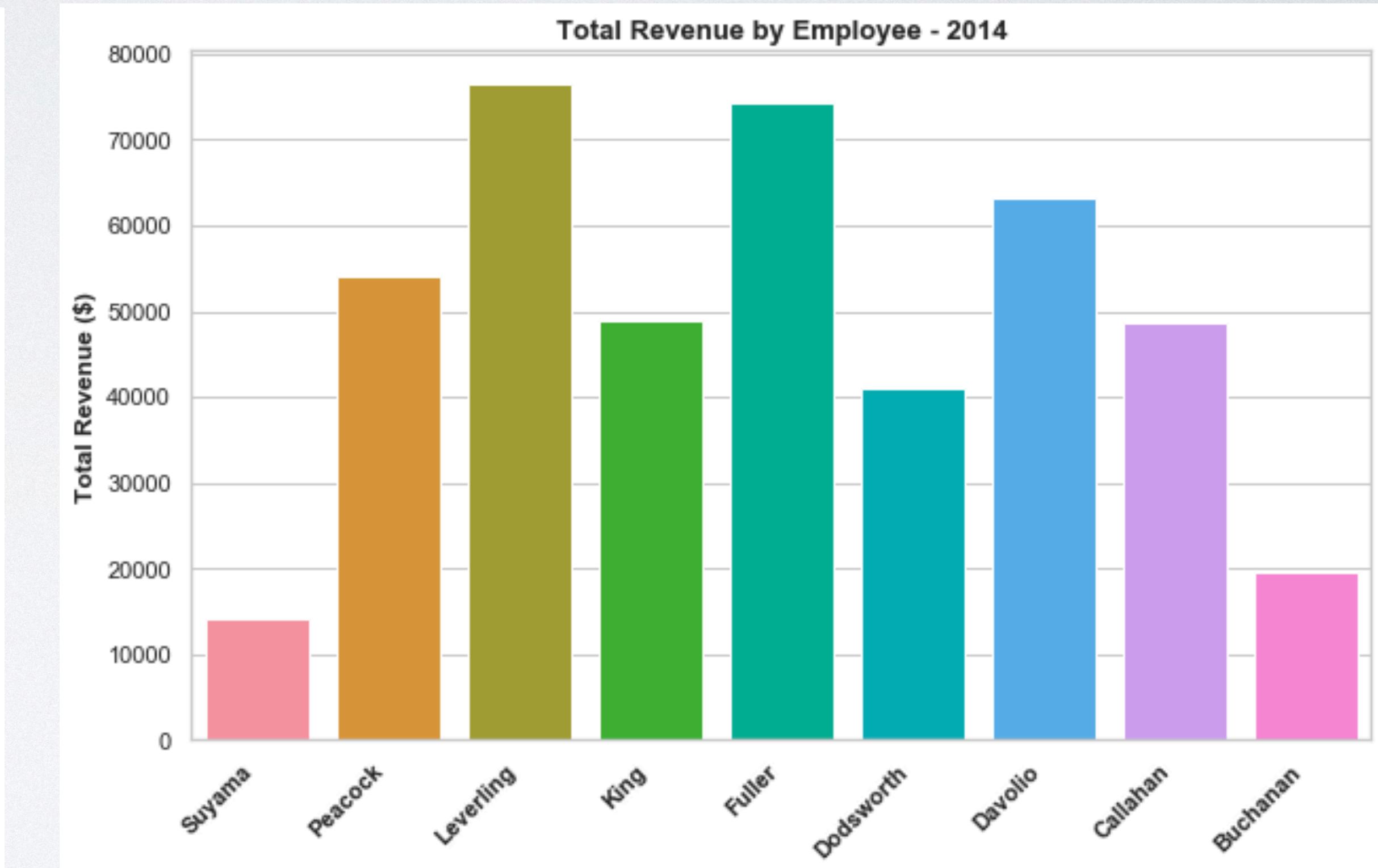
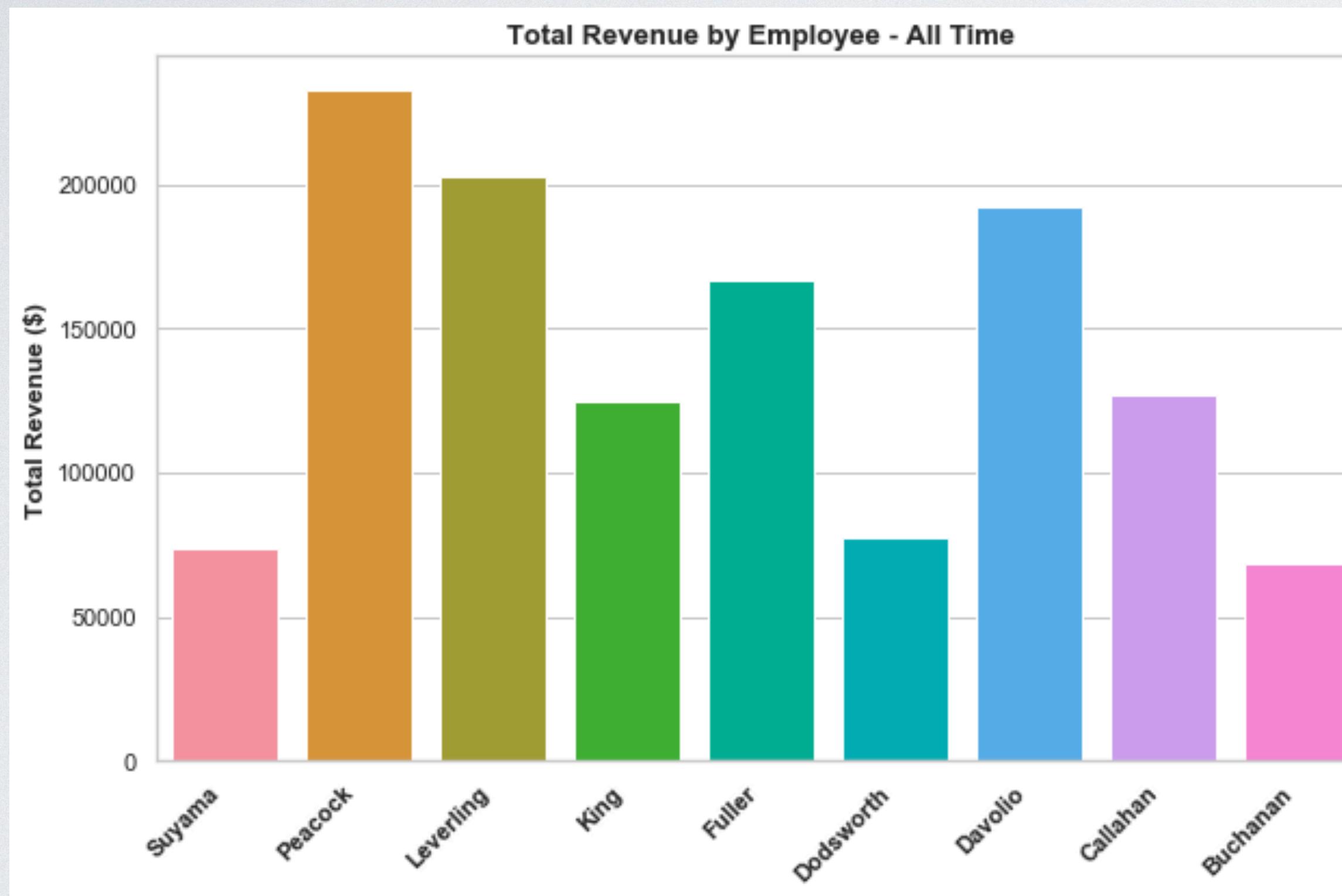
- $H_0$ : Discount do not have a statistically significant effect on number of products customers order by category.
- $H_a$ : Discounts do have a statistically significant effect on number of products customers order by category.
- $\alpha = 0.05$
- Test(s) Used: Tukey test

We fail to reject the null hypothesis that discounts do not have a statistically significant effect on number of products customers order by category.

Tukey Test - Discount Level Vs. Order Quantity by Product Category

Multiple	Comparison	of	Means -	Tukey HSD,FWER=0.05
group1	group2	---	meandi	ff lower upper reject
Beverages	Condiments	ts	1.043	1 -3.8296 5.9159 False
Beverages	Confectionery	ns	0.21	2 -4.0656 4.4859 False
Beverages	Dairy Prod	uct	s 1.466	2 -2.7021 5.6346 False
Beverages	Grains/Cereals	eal	s -0.209	4 -5.2426 4.8237 False
Beverages	Meat/Poultry	try	0.677	6 -4.567 5.9222 False
Beverages	Produce		-1.453	3 -7.1915 4.2849 False
Beverages	Seafood		-0.191	6 -4.4817 4.0985 False
Condiments	Confectionery	ns	-0.83	3 -5.8859 4.2199 False
Condiments	Dairy Prod	uct	s 0.423	1 -4.5393 5.3854 False
Condiments	Grains/Cereals	eal	s -1.252	6 -6.9607 4.4555 False
Condiments	Meat/Poultry	try	-0.365	5 -6.2609 5.5298 False
Condiments	Produce		-2.496	5 -8.8349 3.842 False
Condiments	Seafood		-1.234	8 -6.2998 3.8302 False
Confections	Dairy Prod	uct	s 1.256	1 -3.1216 5.6337 False
Confections	Grains/Cereals	eal	s -0.419	6 -5.6274 4.7882 False
Confections	Meat/Poultry	try	0.467	5 -4.9449 5.8799 False
Confections	Produce		-1.663	5 -7.5554 4.2285 False
Confections	Seafood		-0.401	8 -4.8955 4.0919 False
Dairy Products	Grains/Cereals	eal	s -1.675	7 -6.7956 3.4443 False
Dairy Products	Meat/Poultry	try	-0.788	6 -6.1165 4.5393 False
Dairy Products	Produce		-2.919	5 -8.734 2.8949 False
Dairy Products	Seafood		-1.657	8 -6.0494 2.7337 False
Grains/Cereals	Meat/Poultry	try	0.887	1 -5.1416 6.9157 False
Grains/Cereals	Produce		-1.243	9 -7.7065 5.2188 False
Grains/Cereals	Seafood		0.017	8 -5.2017 5.2374 False
Meat/Poultry	Produce		-2.13	9 -8.7595 4.4977 False
Meat/Poultry	Seafood		-0.869	2 -6.2929 4.5545 False
Produce	Seafood		1.261	7 -4.6406 7.164 False

# EMPLOYEE PERFORMANCE



# INSIGHTS

There is a significant difference in the total Q1 sales revenue and average order revenue between 2013 and 2014.

There is a difference in order quantity, but the difference between order quantity for varying levels of discount is not significant. We may consider discounting products by 5%, given the apparent lack of significant difference in order quantity between products at 5% and greater discounts.

There are no apparent significant differences in order quantity by product category based on discount level, so discounts should be placed on lower cost items to minimize the effect of discount on total revenue.

THANK YOU  
FOR YOUR  
TIME