Northwind Traders

Discounts, Sales, and Employee Performance Analysis

Statistical Testing Methods

t-test: *one variable*

Are the means of two groups statistically different?

one-way ANOVA: *one categorical variable*

Are the means of multiple groups statistically different?

Two-way ANOVA: *two categorical variables*

Does the interaction of two variables affect the means of multiple groups?

ANCOVA: *one categorical, one continuous variable*

Are the means of multiple groups statistically different?

Tukey's HSD: *post-hoc test*

Where exactly do the differences lie?

p-value: *0.05*

Cutoff for whether results can be declared statistically significant

Discount Effectiveness

Do discounts increase the quantity of product ordered?

Does the level of discount affect the quantity of product ordered?

Are discounts more effective for certain product categories?

Is the effectiveness of a discount different depending on the initial unit price?

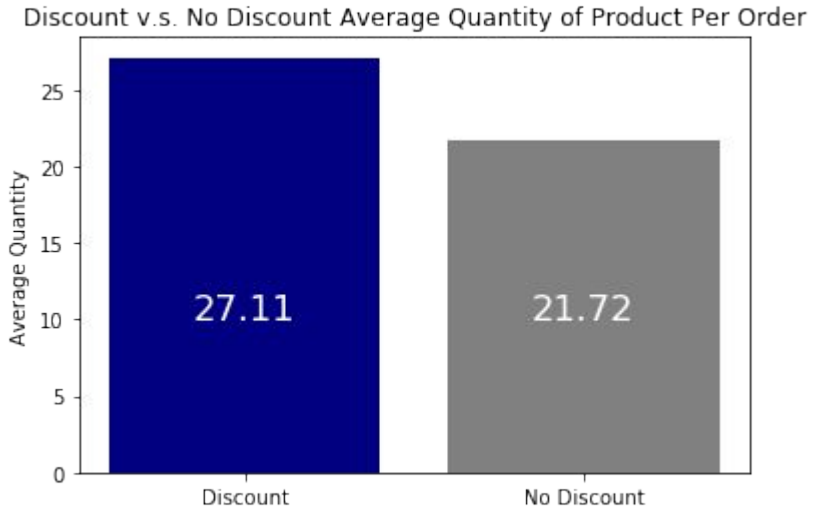
What products do we discount the most?

Employee Sales Performance

Do any employees outperform others?

Are any employees more effective in selling a particular category of product?

Do discounts increase the quantity of product ordered?

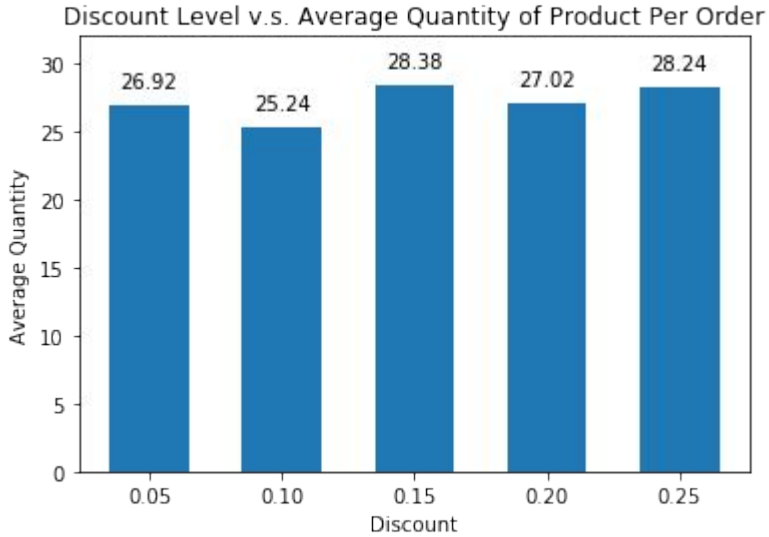


t-test result:

p-value = 0.000000000114

Providing a discount increases the quantity ordered in a statistically significant way

Does the level of discount affect the quantity of product ordered?

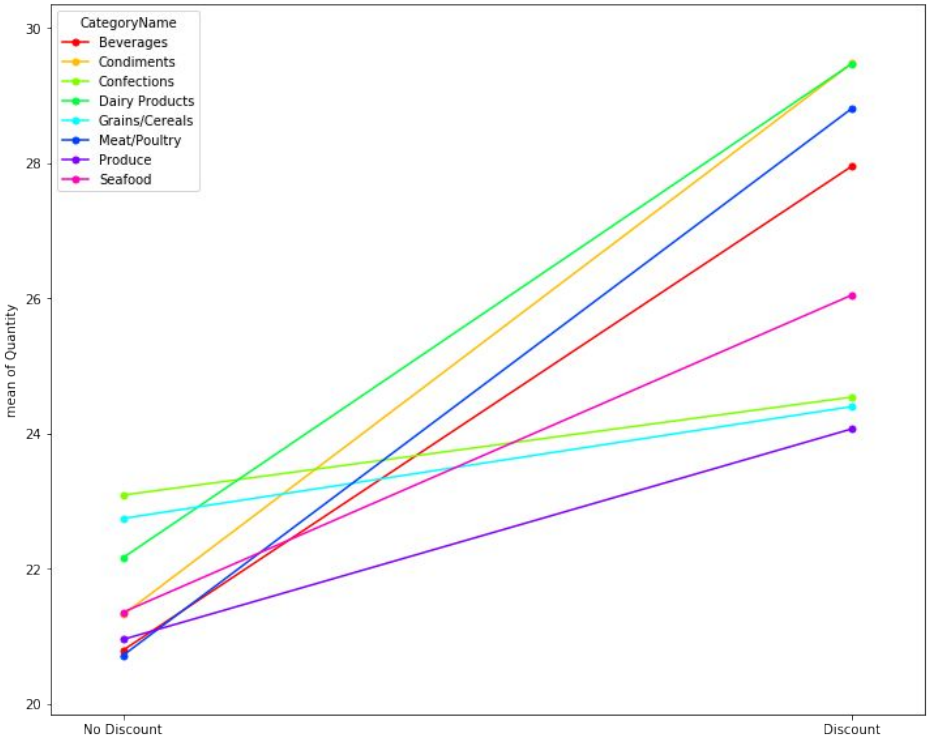


One-way ANOVA result:

p-value = 0.650947

The discount amount is not significant in regards to increasing the quantity of product ordered.

Are discounts more effective for certain product categories?



Two-way ANOVA result:

p-value = 0.2226259

There is no significant difference in the effectiveness of a discount across the categories of products.

Is the effectiveness of a discount different depending on the initial unit price?

ANCOVA result:

OLS Regression Results

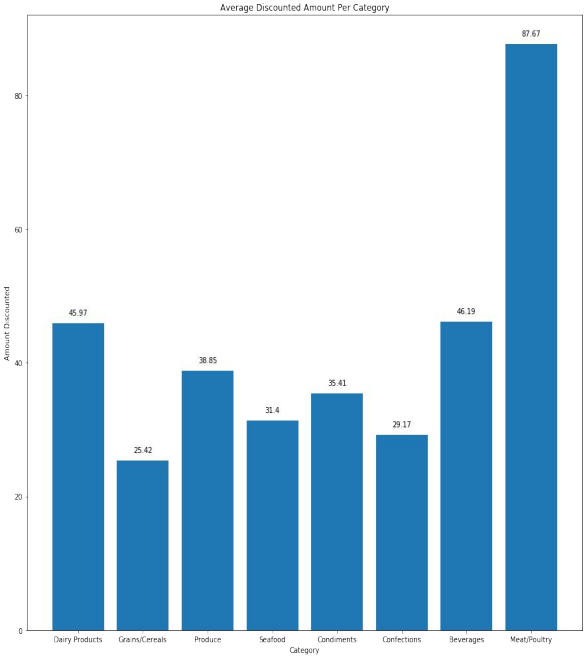
Dep. Variable:	Quantity	**R-squared:**	0.023
Model:	OLS	**Adj. R-squared:**	0.018
Method:	Least Squares	**F-statistic:**	4.592
Date:	Fri, 30 Nov 2018	**Prob (F-statistic):**	6.18e-07
Time:	13:16:53	**Log-Likelihood:**	-9380.0
No. Observations:	2155	**AIC:**	1.878e+04
Df Residuals:	2143	**BIC:**	1.885e+04
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	21.6653	0.692	31.297	0.000	20.308	23.023
C(Discount)[T.0.05]	5.6568	1.839	3.076	0.002	2.051	9.263
C(Discount)[T.0.1]	4.8518	1.981	2.449	0.014	0.967	8.736
C(Discount)[T.0.15]	6.3929	2.720	2.350	0.019	1.058	11.728
C(Discount)[T.0.2]	3.2313	2.154	1.500	0.134	-0.992	7.455
C(Discount)[T.0.25]	4.5212	2.241	2.018	0.044	0.127	8.916
UnitPrice	0.0019	0.017	0.109	0.913	-0.032	0.036
C(Discount)[T.0.05]:UnitPrice	-0.0149	0.038	-0.395	0.693	-0.089	0.059
C(Discount)[T.0.1]:UnitPrice	-0.0528	0.050	-1.05[illegible]	0.291	-[illegible].151	0.045
C(Discount)[T.0.15]:UnitPrice	0.0122	0.096	0.1[illegible]8	0.898	-[illegible].175	0.200
C(Discount)[T.0.2]:UnitPrice	0.0886	0.062	1.43[illegible]	0.153	-[illegible].033	0.210
C(Discount)[T.0.25]:UnitPrice	0.0709	0.056	1.27[illegible]	0.204	-0.039	0.180

All p-values > 0.05

There is no significant interaction between discount level and unit price.

Do we discount some products more than others?



One-way ANOVA result:

p-value = 0.000113

There is some significant difference in average discounted amount across the categories.

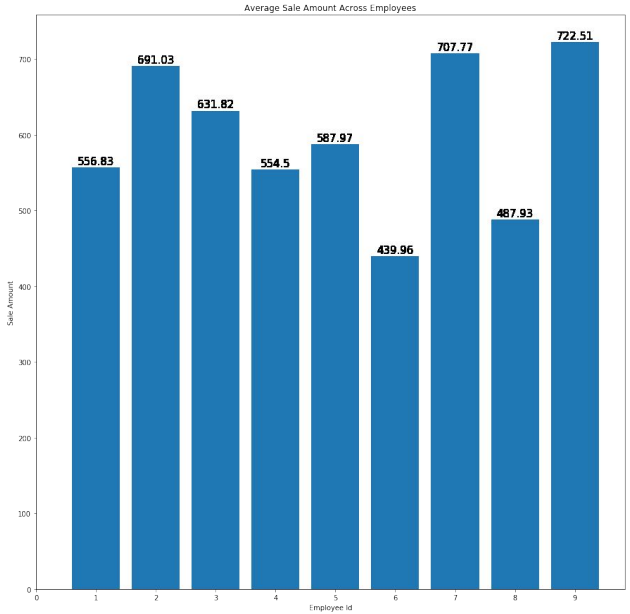
Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
Beverages	Condiments	-10.7792	-44.8653	23.3068	False
Beverages	Confections	-17.0178	-46.9241	12.8885	False
Beverages	Dairy Products	-0.22	-29.4018	28.9618	False
Beverages	Grains/Cereals	-20.7656	-55.9666	14.4354	False
Beverages	Meat/Poultry	41.4822	4.7394	78.225	True
Beverages	Produce	-7.332	-47.4219	32.7579	False
Beverages	Seafood	-14.787	-44.7924	15.2183	False
Condiments	Confections	-6.2386	-41.5471	29.0699	False
Condiments	Dairy Products	10.5592	-24.1378	45.2562	False
Condiments	Grains/Cereals	-9.9864	-49.879	29.9062	False
Condiments	Meat/Poultry	52.2614	11.002	93.5208	True
Condiments	Produce	3.4472	-40.8191	47.7135	False
Condiments	Seafood	-4.0078	-39.4003	31.3846	False
Confections	Dairy Products	16.7978	-13.8031	47.3986	False
Confections	Grains/Cereals	-3.7478	-40.1339	32.6382	False
Confections	Meat/Poultry	58.5	20.6203	96.3796	True
Confections	Produce	9.6858	-31.4485	50.8201	False
Confections	Seafood	2.2307	-29.1564	33.6179	False
Dairy Products	Grains/Cereals	-20.5456	-56.3386	15.2474	False
Dairy Products	Meat/Poultry	41.7022	4.3919	79.0125	True
Dairy Products	Produce	-7.112	-47.7226	33.4987	False
Dairy Products	Seafood	-14.567	-45.2647	16.1306	False
Grains/Cereals	Meat/Poultry	62.2478	20.0626	104.4331	True
Grains/Cereals	Produce	13.4336	-31.6968	58.5641	False
Grains/Cereals	Seafood	5.9786	-30.4889	42.4461	False
Meat/Poultry	Produce	-48.8142	-95.1573	-2.4711	True
Meat/Poultry	Seafood	-56.2692	-94.2271	-18.3113	True
Produce	Seafood	-7.4551	-48.6615	33.7514	False

Tukey's HSD result:

Meat/Poultry products are discounted significantly more than all other categories.

Do some employees outperform others?



One-way ANOVA results:

p-value = 0.046929

There is some significant difference in sales performance amongst employees.

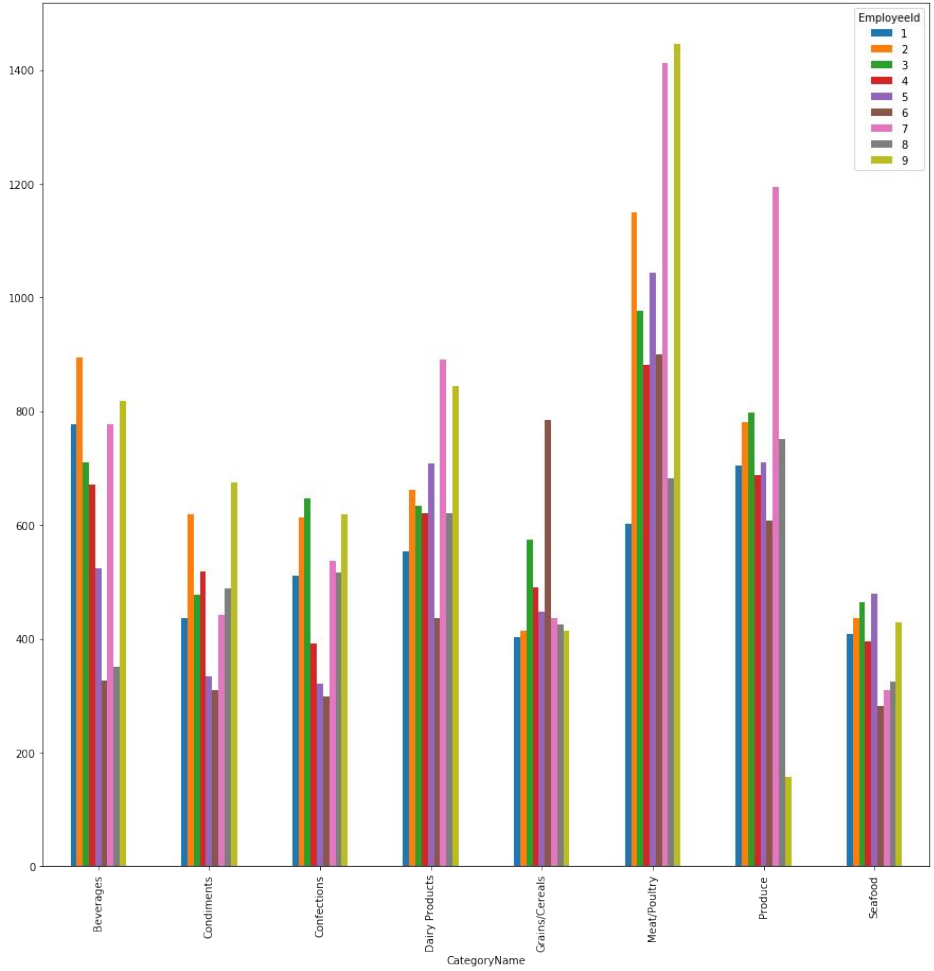
Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
1	2	134.1944	-117.9953	386.3841	False
1	3	74.9821	-157.9727	307.9368	False
1	4	-2.3316	-220.6011	215.9379	False
1	5	31.1346	-290.2426	352.5118	False
1	6	-116.8745	-399.4868	165.7378	False
1	7	150.9404	-127.3188	429.1996	False
1	8	-68.9018	-315.6068	177.8032	False
1	9	165.6717	-166.7309	498.0742	False
2	3	-59.2123	-315.2499	196.8252	False
2	4	-136.526	-379.2787	106.2267	False
2	5	-103.0598	-441.5426	235.423	False
2	6	-251.0689	-552.9911	50.8533	False
2	7	16.746	-281.1054	314.5975	False
2	8	-203.0962	-471.705	65.5126	False
2	9	31.4773	-317.4909	380.4454	False
3	4	-77.3137	-300.0179	145.3906	False
3	5	-43.8475	-368.253	280.558	False
3	6	-191.8566	-477.9078	94.1946	False
3	7	75.9584	-205.7929	357.7096	False
3	8	-143.8839	-394.521	106.7533	False
3	9	90.6896	-244.6417	426.0208	False
4	5	33.4662	-280.5602	347.4926	False
4	6	-114.5429	-388.7672	159.6814	False
4	7	153.272	-116.4638	423.0079	False
4	8	-66.5702	-303.62	170.4797	False
4	9	168.0033	-157.2977	493.3042	False
5	6	-148.0091	-509.7282	213.71	False
5	7	119.8058	-238.5225	478.1341	False
5	8	-100.0364	-434.4528	234.3801	False
5	9	134.5371	-267.2868	536.361	False
6	7	267.815	-56.1998	591.8297	False
6	8	47.9727	-249.3836	345.329	False
6	9	282.5462	-89.0031	654.0954	False
7	8	-219.8422	-513.0644	73.3799	False
7	9	14.7312	-353.5178	382.9802	False
8	9	234.5735	-110.4519	579.5988	False

Tukey's HSD results:

No significant comparisons

No employee performs significantly better in sales than any other employee.



Are any employees more effective in selling a particular category of product?

One-way ANOVA results:

p-value = 0.9995606

There is no significant interaction between the employee and the category of product in relation to average sale price.

Tukey's HSD results:

No significant comparisons

No employee significantly outperforms another employee in any particular category.