Campaign's Effect on Sales

Dataset + Sheets: *Click here*

Background

A fast-food chain company plans to add a new item to its menu. The company is still undecided between the three possible marketing campaigns for promoting the new product. In order to determine which promotion has the greatest effect on sales, the new items is introduced at location in several randomly selected markets. A different promotion is used at each location, and the weekly sales of the new item are recorded for the first four weeks. As a data analyst with the raw data we have, I am expected to perform analysis and extract insights to determine which promotions have the greatest effect on sales.

Objective

Create an exploratory data analysis & statistical measurement to give some insights and recommendation on:

Store Character

How is the character of the store in each type of market size?

Campaign

Wich promotion/campaign brings the greatest effect on sales?

Market Size

Does the type of market size affect the results of the campaign?

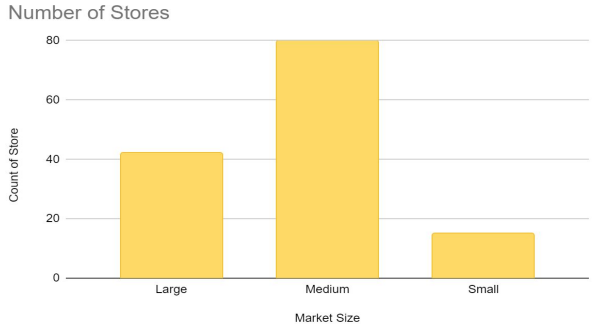
Effect on Sales

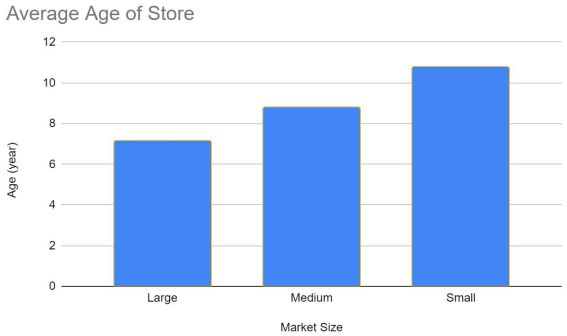
Is there any effect of age of store and type of market size on the sales generated?

Data Dictionary

Column Name	Definition
MarketID	Unique identifier for market
MarketSize	Size of market area by sales
LocationID	Unique identifier for store location
AgeOfStore	Age of store in years
Promotion	One of three promotions that were tested
Week	One of four weeks when the promotions were run
SalesInThousands	Sales amount for a specific LocationID, Promotion, and week

Store Characteristics

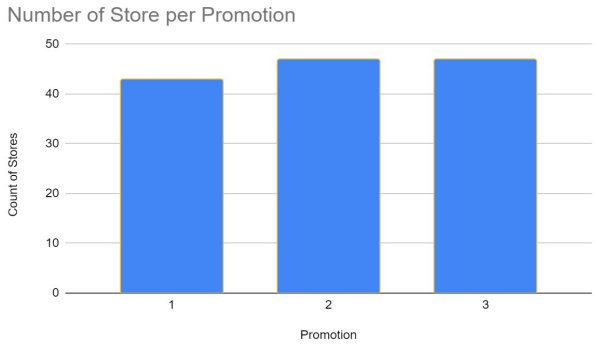


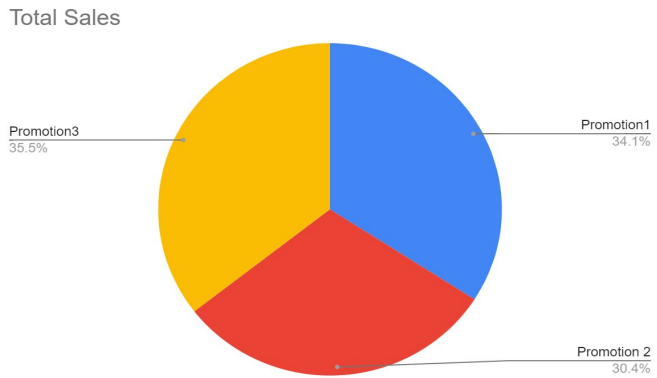


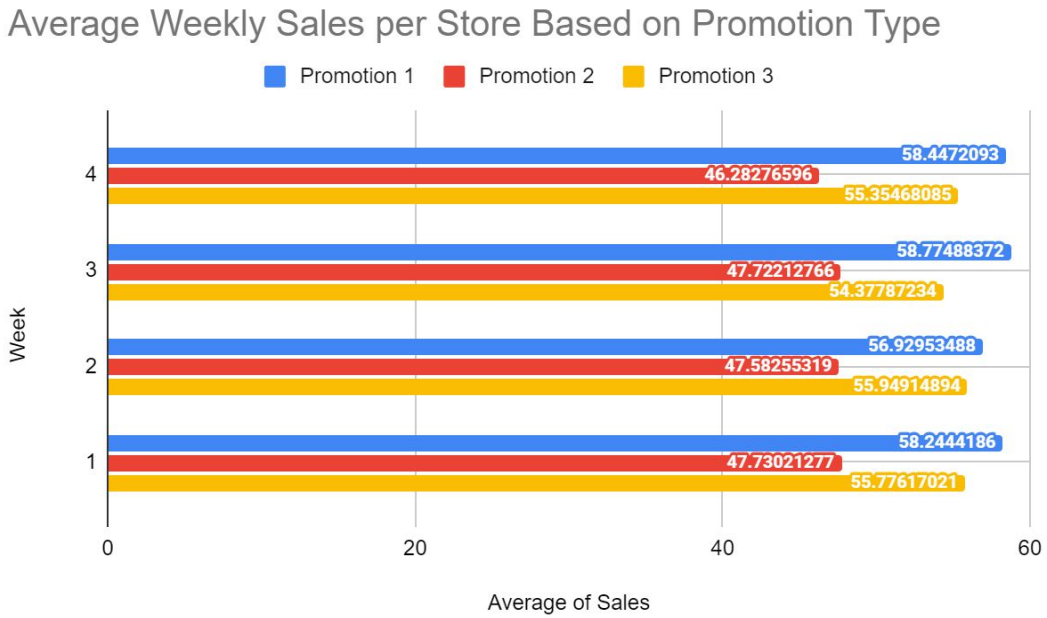
Store Characteristics

Small	Medium	Large
The number of stores for the small market size category is 15 stores.	The number of stores for the medium market size category is 80 stores.	The number of stores for the market size category of large is 42 stores.
The most used type of promotion in the small market size category is Promotion 3.	The most used type of promotion in the medium market size category is Promotion 3.	The most used type of promotion in the large market size category is Promotion 2.
The average age of the store in the small market size category is 10.8 years.	The average age of the store in the medium market size category is 7.1 years.	The average age of the store in the large market size category is 8.8 years.
The average weekly sales in the small category of market size is 57.41 thousand	The average weekly sales in the medium market size category is 43.99 thousand	The average weekly sales in the large category market size is 70.12 thousand.

Campaign Performance Review







Campaign Performance Review

Promotion type 1 was implemented in 43 stores, while promotion type 2 and 3 were both implemented in 47 stores.

Stores that implemented promotion type 1 **generates the highest average sales**.

In total, promotion type 3 generates the **highest total sales** as there are more stores that implement promotion type 3 compared to promotion type 1

Promotion type 2 had the **lowest performance** based on sales compared to the other promotion types.

Campaign Performance Review

Campaign that brings the greatest effect on sales?

To answer this question we can use the **t-test**, because we will **compare the performance between the two variables.**

Hypothesis

H0 : sales promotion i = sales promotion j

H1 : sales promotion i != sales promotion j

I,j = 1, 2, 3 and i!=j

Level of Significant

Alpha = 5% = 0.05

Test Criterion

Reject H0 if p-value < alpha

Campaign Performance Review

Promotion 1 vs Promotion 2		
	Promotion 1	*Promotion 2*
Mean	58.09901163	47.32941489
Variance	274.0276885	228.2805146
Observations	172	188
Pooled Variance	250.1318183	
Hypothesized Mean Difference	0	
df	358	
t Stat	6.453671702	
P(T<=t) one-tail	0.0000000001775334837	
t Critical one-tail	1.649121007	
P(T<=t) two-tail	0.0000000003550669673	< 0.05
t Critical two-tail	1.966612447	

Because the resulting p-value is **<0.05**, there is a **significant difference in performance** between promotion 1 and promotion 2.

Campaign Performance Review

Promotion 2 vs Promotion 3		
	Promotion 2	*Promotion 3*
Mean	47.32941489	55.36446809
Variance	228.2805146	281.1064944
Observations	188	188
Pooled Variance	254.6935045	
Hypothesized Mean Difference	0	
df	374	
t Stat	-4.881392711	
P(T<=t) one-tail	0.0000007814471443	
t Critical one-tail	1.648937988	
P(T<=t) two-tail	0.000001562894289	< 0.05
t Critical two-tail	1.96632711	

Because the resulting p-value is **<0.05**, there is a **significant difference in performance** between promotion 2 and promotion 3.

Campaign Performance Review

Promotion 1 vs Promotion 3		
	Promotion 1	*Promotion 3*
Mean	58.09901163	55.36446809
Variance	274.0276885	281.1064944
Observations	172	188
Pooled Variance	277.725277	
Hypothesized Mean Difference	0	
df	358	
t Stat	1.555138369	
P(T<=t) one-tail	0.06039833525	
t Critical one-tail	1.649121007	
P(T<=t) two-tail	0.1207966705	> 0.05
t Critical two-tail	1.966612447	

Because the resulting p-value is **>0.05**, there is a **no significant difference in performance** between promotion 1 and promotion 3.

Campaign Performance Review

Conclusion

From the results of the t-test analysis, it can be seen that **Promotion 1 has the best influence on sales**, compared to Promotion 3 and Promotion 2, **with an average sales of 58.09 thousand**.

Market Size

Does type of market size affect the result of the campaign?

To answer this question we can use the **t-test**, because we will **compare the performance between the two variables.**

Hypothesis

H0 : sales promotion i = sales promotion j

H1 : sales promotion i != sales promotion j

I,j = Small, Medium, Large and i!=j

Level of Significant

Alpha = 5% = 0.05

Test Criterion

Reject H0 if p-value < alpha

Market Size

Promotion 1

Small vs Medium		
	Small	*Medium*
Mean	60.1625	47.67260417
Variance	26.28898816	65.08182157
Observations	20	96
Pooled Variance	58.61634933	
Hypothesized Mean Difference	0	
df	114	
t Stat	6.63698831	
P(T<=t) one-tail	0.000000000565 0314906	
t Critical one-tail	1.658329909	
P(T<=t) two-tail	0.000000001130 062981	<0.05
t Critical two-tail	1.980992234	

Medium vs Large		
	Medium	*Large*
Mean	47.67260417	75.235892 86
Variance	65.08182157	240.17241 37
Observations	96	56
Pooled Variance	129.2817054	
Hypothesized Mean Difference	0	
df	150	
t Stat	-14.41684847	
P(T<=t) one-tail	0	
t Critical one-tail	1.65507544	
P(T<=t) two-tail	0	<0.05
t Critical two-tail	1.975905298	

Small vs Large		
	Small	*Large*
Mean	60.1625	75.23589 286
Variance	26.28898816	240.1724 137
Observations	20	56
Pooled Variance	185.2563991	
Hypothesized Mean Difference	0	
df	74	
t Stat	-4.251347689	
P(T<=t) one-tail	0.00003061701482	
t Critical one-tail	1.665706833	
P(T<=t) two-tail	0.00006123402964	<0.05
t Critical two-tail	1.992543466	

Market Size

Promotion 1 Conclusion

From the results of the t-test analysis, it can be seen that the p-values of all market sizes are **<0.05**, therefore there **is an influence between all market sizes and promotion 1**. The average sales on the **largest size are the highest**, namely 70.23 thousand , followed by small at 60.16 thousand, and moderate at 47.67 thousand.

Market Size

Promotion 2

Small vs Medium		
	Small	*Medium*
Mean	50.810625	39.11435185
Variance	34.45473958	77.57041733
Observations	16	108
Pooled Variance	72.26930941	
Hypothesized Mean Difference	0	
df	122	
t Stat	5.136077592	
P(T<=t) one-tail	0.000000538449 9251	
t Critical one-tail	1.657439439	
P(T<=t) two-tail	0.000001076899 85	<0.05
t Critical two-tail	1.979599808	

Medium vs Large		
	Medium	*Large*
Mean	39.11435185	60.322031 25
Variance	77.57041733	247.38663 55
Observations	108	64
Pooled Variance	140.5023099	
Hypothesized Mean Difference	0	
df	170	
t Stat	-11.34198474	
P(T<=t) one-tail	0	
t Critical one-tail	1.653866257	
P(T<=t) two-tail	0	<0.05
t Critical two-tail	1.974016669	

Small vs Large		
	Small	*Large*
Mean	50.810625	60.32203 125
Variance	34.45473958	247.3866 355
Observations	16	64
Pooled Variance	206.438194	
Hypothesized Mean Difference	0	
df	78	
t Stat	-2.368398029	
P(T<=t) one-tail	0.0101705487	
t Critical one-tail	1.664624585	
P(T<=t) two-tail	0.0203410974	<0.05
t Critical two-tail	1.990847036	

Market Size

Promotion 2 Conclusion

From the results of the t-test analysis, it can be seen that the p-values of all market sizes are **<0.05,** therefore **there is an influence between all market sizes and promotion 2**. The average sales on the **large size are the highest**, namely 60.32 thousand, followed by small of 50.81 thousand, and medium of 39.11 thousand.

Market Size

Promotion 3

Small vs Medium		
	Small	*Medium*
Mean	59.51416667	45.46887931
Variance	27.17210362	65.47020656
Observations	24	116
Pooled Variance	59.0871894	
Hypothesized Mean Difference	0	
df	138	
t Stat	8.148066417	
P(T<=t) one-tail	0	
t Critical one-tail	1.655970322	
P(T<=t) two-tail	0	<0.05
t Critical two-tail	1.977303459	

Medium vs Large		
	Medium	*Large*
Mean	45.46887931	77.20395833
Variance	65.47020656	207.3687606
Observations	116	48
Pooled Variance	106.6383056	
Hypothesized Mean Difference	0	
df	162	
t Stat	-17.90649787	
P(T<=t) one-tail	0	
t Critical one-tail	1.654313896	
P(T<=t) two-tail	0	<0.05
t Critical two-tail	1.974715749	

Small vs Large		
	Small	*Large*
Mean	59.51416667	77.20395833
Variance	27.17210362	207.3687606
Observations	24	48
Pooled Variance	148.1612876	
Hypothesized Mean Difference	0	
df	70	
t Stat	-5.813200976	
P(T<=t) one-tail	0.0000000834181171	
t Critical one-tail	1.666914419	
P(T<=t) two-tail	0.0000001668362354	<0.05
t Critical two-tail	1.994437086	

Market Size

Promotion 3 Conclusion

From the results of the t-test analysis, it can be seen that the p-values of all market sizes are **<0.05,** therefore **there is an influence between all market sizes and promotion 3**. The average sales on the **large size are the highest**, namely 77.20 thousand, followed by small of 59.51 thousand, and medium of 45.44 thousand.

Market Size

Does type of market size affect the result of the overall sales?

To answer this question we can use the **t-test**, because we will **compare the performance between the two variables.**

Hypothesis

H0 : sales i = sales j

H1 : sales i != sales j

I,j = Small, Medium, Large and i!=j

Level of Significant

Alpha = 5% = 0.05

Test Criterion

Reject H0 if p-value < alpha

Market Size

Overall Sales

Small vs Medium		
	Small	*Medium*
Mean	57.40933333	43.98534375
Variance	43.99803006	81.92745255
Observations	60	320
Pooled Variance	76.00725168	
Hypothesized Mean Difference	0	
df	378	
t Stat	10.94492797	
P(T<=t) one-tail	0	
t Critical one-tail	1.64889466	
P(T<=t) two-tail	0	< 0.05
t Critical two-tail	1.966259562	

Medium vs Large		
	Medium	*Large*
Mean	43.98534375	70.11672619
Variance	81.92745255	290.776009
Observations	320	168
Pooled Variance	153.6922857	
Hypothesized Mean Difference	0	
df	486	
t Stat	-22.12361388	
P(T<=t) one-tail	0	
t Critical one-tail	1.647994915	
P(T<=t) two-tail	0	< 0.05
t Critical two-tail	1.96485709	

Small vs Large		
	Small	*Large*
Mean	57.40933333	70.11672619
Variance	43.99803006	290.776009
Observations	60	168
Pooled Variance	226.3516693	
Hypothesized Mean Difference	0	
df	226	
t Stat	-5.616004515	
P(T<=t) one-tail	0.000000028546048 84	
t Critical one-tail	1.651623799	
P(T<=t) two-tail	0.000000057092097 69	< 0.05
t Critical two-tail	1.970516191	

Market Size

Overall Sales Conclusion

From the results of the t-test analysis, it can be seen that **there is an influence between market size and sales** because the p-values of all market sizes are **<0.05.** With the **highest average sales for the large size**, namely 70.11 thousand, followed by small at 57.40 thousand, and medium at 43.98 thousand.

Effect on Sales

To find out is there any effect of age of store and type of market size on the sales generated. **We can use linear regression**. But because the market size variable is a category, an encode is needed to convert the variable to numeric (**small: 1, medium: 2, large: 3**)

Effect on Sales

Correlation

	MarketSize	AgeOfStore	SalesInThousands
MarketSize	1		
AgeOfStore	*-0.1641378349*	1	
SalesInThousands	0.4544945279	-0.0285328811	1

Insight

There is a correlation between Market Size and Sales, but the correlation is moderate. Meanwhile, the correlation between age of stores and sales is very weak.

Effect on Sales

Linear Regression

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	32055.74934	16027.87467	71.89012623	0
Residual	545	121507.5304	222.949597		
Total	*547*	*153563.2797*			

Insight

If seen from the Significance F value is less than 0.05, it can be concluded that the regression is quite good and the variables Market Size and Ages of Store collectively have a significant influence on Sales.

Effect on Sales

Linear Regression

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95%	Upper 95%
Intercept	24.76827472	2.659721884	9.312355129	0	19.54371332	29.99283613	19.54371332	29.99283613
Market Size	12.59936171	1.052803861	11.96743494	0	10.5313115	14.66741193	10.5313115	14.66741193
Ages of Store	0.1194922639	0.09749454298	1.225630279	0.2208671881	-0.07201882195	0.3110033498	-0.07201882195	0.3110033498

Insight

However, when viewed from the P-value, the ages of store variable has a p-value > 0.05, which indicates that the ages of store variables individually have no influence on sales. While the market variable has a p-value <0.05, in which the individual market size variables have an influence on sales.

Effect on Sales

Model for Predicting Sales

$$y = 24,76 + 12,59x_1 + 0,11x_2$$

$x_1 = Market\ Size\ (small = 1, medium = 2, large = 3)$

$x_2 = Age\ of\ Store$

It should be noted that the age of store variable has no significant effect on sales. Therefore the model is still not perfect and needs to be supplemented by other variables outside the dataset.

Thanks!

Contact me:

rafiqnaufal97@gmail.com
+62 812 809 05778
linkedin.com/in/rafiqnaufal