Fast Food Chain Analysis

RevoU Full Stack Data Analytics
Week 2 - Week 3, January 2023

Ichsan Maulana

Background

A fast-food chain plans to add a new item to its menu. You are the data analyst working for this fast-food chain company. This 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 item is introduced at locations 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.

Task Creating / Methodology

1

2

3

4

Point 1

Understanding Business Problem

We need to set the problem definition, datasets and output delivery

Point 2

Basic Data Cleaning

Do the data cleaning such a remove and replace, and descriptive statistics

Point 3

Exploratory Data Analytics

Do the EDA to see the pattern of data in case to generate findings and insight

Point 4

Hypothesis Testing

Hypothesis testing in purpose to test the promotions by using T - test

Fast Food Chain Analysis - Business Problem

Fast food chain plans to add new item at the menu with product promotions but they still undecided which promotions have greatest effect on sales. By the data we have we can help the company to decide the right promotions by market share and store age. We can use descriptive statistics, hypothesis testing and EDA to generate findings and insight

We can apply remove and replace for basic data cleaning and understand the variables with data dictionary

Name	Explaination
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

Fast Food Chain Analysis - Descriptive Statistics

Sales have 0.8 skewness means highly positive skewed, here we used median instead of mean because there is an outlier but it doesn't have to removed

| LocationID		AgeOfStore		Promotion		week		SalesInThousands	
Mean	479.6569343	Mean	8.503649635	Mean	2.02919708	Mean	2.5	Mean	53.46620438
Standard Error	12.30162591	Standard Error	0.2835760571	Standard Error	0.03463261973	Standard Error	0.04780368141	Standard Error	0.7157473472
Median	504	Median	7	Median	2	Median	2.5	Median	50.2
Mode	920	Mode	1	Mode	3	Mode	4	Mode	51.09
Standard Deviation	287.9736794	Standard Deviation	6.638345301	Standard Deviation	0.8107288421	Standard Deviation	1.119055491	Standard Deviation	16.75521582
Sample Variance	82928.84004	Sample Variance	44.06762834	Sample Variance	0.6572812554	Sample Variance	1.252285192	Sample Variance	280.7372572
Kurtosis	-1.155923439	Kurtosis	0.3785708379	Kurtosis	-1.476448078	Kurtosis	-1.361461438	Kurtosis	0.1647924508
Skewness	-0.01600146514	Skewness	1.041542877	Skewness	-0.05339018062	Skewness	0	Skewness	0.806496413
Range	919	Range	27	Range	2	Range	3	Range	82.31
Minimum	1	Minimum	1	Minimum	1	Minimum	1	Minimum	17.34
Maximum	920	Maximum	28	Maximum	3	Maximum	4	Maximum	99.65
Sum	262852	Sum	4660	Sum	1112	Sum	1370	Sum	29299.48
Count	548	Count	548	Count	548	Count	548	Count	548

Fast Food Chain Analysis - Correlation

We do correlation analysis to see what variables have strong correlation between other variable and we can see sales has strong correlation with market size.

	LocationID	*AgeOfStore*	*Promotion*	*week*	*SalesInThousands*	*MarketSizeNumber*
LocationID	1					
AgeOfStore	-0.05073557138	1				
Promotion	-0.04991500583	0.0597648402	1			
week	0	0	0	1		
SalesInThousands	-0.1878518268	-0.0285328811	-0.05921195055	-0.01098353709	1	
MarketSizeNumber	0.2704086788	-0.1641378349	-0.05558417376	0	0.4544945279	1

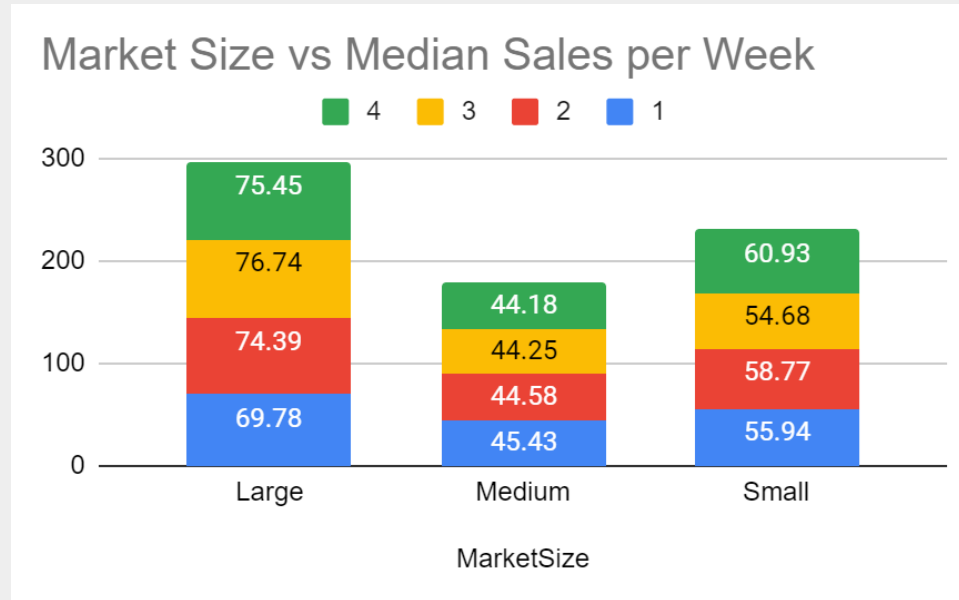**Fast Food Chain Analysis - Exploratory Data Analytics**

Market size is dominated by Medium market size with 320 location or 58.4% of total market. Medium and Large market size have 89% market share which mean this market share is major revenue contributor for company
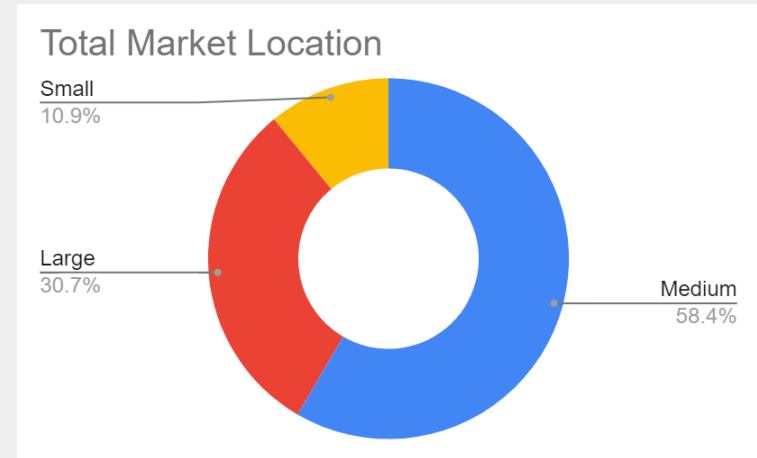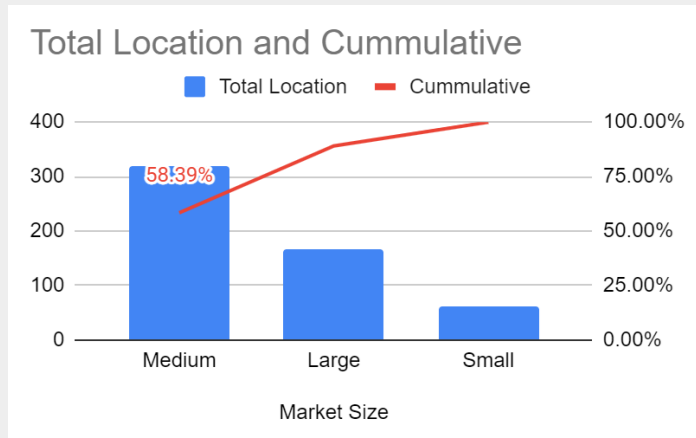
Fast Food Chain Analysis - Exploratory Data Analytics

The different median sales variation is slightly different for medium market size. Large market size had most sales in a month with $296.36k following with small market size with $230.32k and medium market size with $178.44k

Fast Food Chain Analysis - Exploratory Data Analytics

As we know we have 3 type promotions, each promotions is applied to each market share and we can see

- Promotion 1 is works for large & small market size and have the biggest median sales
- Promotion 2 is also works for large market size
- At large market size, all promotion is works

MEDIAN of SalesInThousands	Promotion			
MarketSize	1	2	3	Grand Total
Small	61.145	51.6	59.745	57.555
Medium	47.425	39.765	46.295	44.59
Large	72.835	53.01	82.345	75.02
Grand Total	55.385	45.385	51.165	50.2

Also we grouped by Age of Store and we finds

- Based on age grouping store, group 1-10 at large market have a great impact for promotion number 1
- Group 11-20 at large market size also have great impact for promotion number 1
- Group 21-30 at small market size is suitable for promotion number 1 and number 3

Grouped AgeOfStore	COUNTA of MarketSize
1 - 4	176
5 - 8	160
9 - 12	92
13 - 16	40
17 - 20	36
21 - 24	32
25 - 28	12
Grand Total	548

Market Size vs Age Of Store
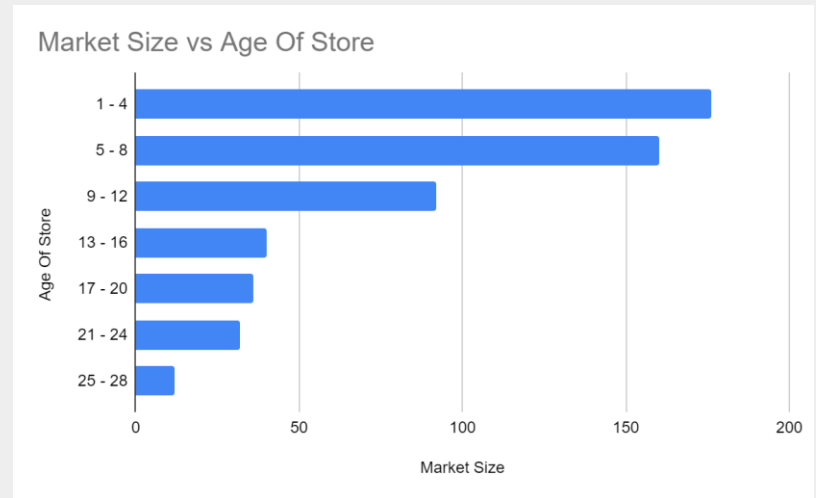
Overall, promotion 1 has the biggest sales with $55.385k or equal to 36.5% of total revenue

Promotion	Median Sales In Thousands
1	55.385
2	45.385
3	51.165

MEDIAN of Sales In Thousands
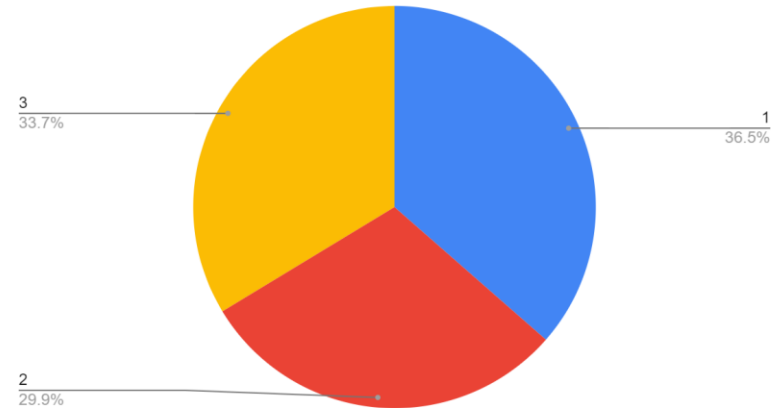
Fast Food Chain Analysis - Hypothesis Testing

For hypothesis testing, we use t-test two sample that assuming equal variance. We state the hypothesis (H0), as promotion (n) is have same result as promotion (n) if the P score is more 5%

	Promotion 1	Promotion 2
Mean	58.09901163	47.32941489
Variance	274.0276885	228.2805146
Observations	172	188
Pooled Variance	250.1318183	
Hypothesized M	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	0.00%	
t Critical two-tail	1.966612447	

P score is <5%, reject H0 Promotion 1 doesnt have same result as promotion 2

	Promotion 1	Promotion 3
Mean	58.09901163	55.36446809
Variance	274.0276885	281.1064944
Observations	172	188
Pooled Variance	277.725277	
Hypothesized M	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	12.08%	
t Critical two-tail	1.966612447	

P score is >5%, accept H0 Promotion 1 have same result as promotion 3

	Promotion 2	Promotion 3
Mean	47.32941489	55.36446809
Variance	228.2805146	281.1064944
Observations	188	188
Pooled Variance	254.6935045	
Hypothesized M	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	0.00%	
t Critical two-tail	1.96632711	

P score is <5%, reject H0 Promotion 2 doesnt have same result as promotion 3

Statistics

Thanks

Advanced Assignment
Week 2&3 - RevoU FSDA JAN23