REVIVING SALES GROWTH FOR MYPICK

Developing a comprehensive strategy to address the slowing sales growth and market share erosion faced by MyPick in the competitive market.

Data Analytics

GUNASINDHU

BUSINESS PROBLEM: SLOWING SALES GROWTH

- **Declining Sales Growth**

  MyPick needs to identify and address the factors hindering their sales growth

- **Market Share Erosion in the competitive market**

  The market is becoming increasingly competitive, with new players entering and existing competitors intensifying their efforts, putting pressure on MyPick's sales and profitability.

BUSINESS GOAL:

- Use of data for better marketing
- Increase sales

PROBABLE CAUSES

- ## Rise in competition

  Increased competition in the market with **competitors offering** similar or better products at competitive prices, leading to a decline in MyPick's market share.

- ## Positioning in customers' minds

  Weak brand positioning and lack of a clear brand identity in the minds of customers, making it difficult for MyPick to **stand out in the competitive** landscape.

- ## Ineffective sales promotion

  Ineffective or poorly executed s**ales promotion strategies**, failing to effectively attract and retain customers, leading to a decline in sales.

- ## Wrong distribution of stores

  Suboptimal distribution of MyPick's stores, with potential **misalignment between store locations and target customer demographics**, resulting in underutilization of resources and missed sales opportunities.

PRIORITIZING THE PROBLEMS

Identify Key Problems → Impact/Effort Matrix → Assess Impact and Effort → Prioritize Problem

The key problems identified from the context are:

- 1) Rise in competitor,
- 2) Brand positioning on customers' minds,
- 3) Ineffective sales promotion, and
- 4) Wrong distribution of stores.

To prioritize the problems and **identify the most important** ones to address. The matrix plots each problem based on its potential impact on the goals and the effort required to implement a solution.

Each problem is assessed based on its potential impact on the goals of using data analytics for **better marketing and increasing sales**, as well as the effort required to implement a solution.

The problems are plotted on the Impact/Effort matrix, and the ones with **high impact and low effort** are identified as the top priorities to address.

APPROACH:EXTRACTING INSIGHTS & RECOMMENDING ACTIONS

Based on :Mypick's goals and data available on Sale and ratings

Descriptive Analysis	Diagonistic analysis	Predictive Analysis	Prescriptive Analysis
• EDA/Visualization	• Neighborhood-Based Clustering (For Brand identity rating) • Linear Regression (Sales data)	• Model with Linear Regression (Sales data)	• Recommendation

HYPOTHESIS FROM DATA ANALYSIS

1. **New Product development :** Mypick's Brand positioning can be improved as preferred brand of to make products similar to MDH-If this Hypothesis is proven true, MyPick will become ideal brand of customers.

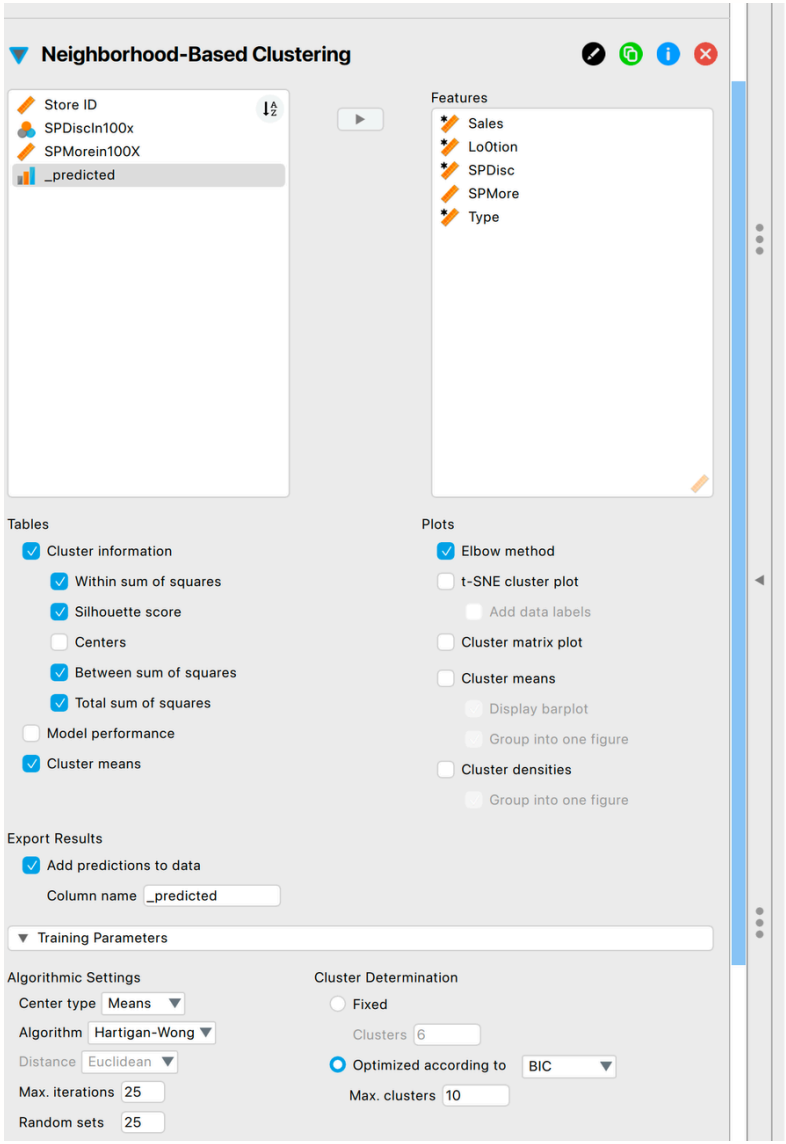
2. **Distribution :** Increasing Mypick's distribution based on location(e.g., residential area, commercial area) will lead to significant sales growth-If this Hypothesis is proven true, MyPick can focus on specific location to keep its products.

3. **Channel partners :**Increasing Mypick's distribution on channel partners (e.g., supermarkets, hypermarkets) will lead to significant sales growth-If this Hypothesis is proven true, MyPick can invest in expanding its presence in modern trade channels.

ANALYSIS

CLUSTER ANALYSIS: K- MEANS BASEDCLUSTERING

(On Sales data)



Neighborhood-Based Clustering

Model Summary: K-Means Clustering

Clusters	N	R^2	AIC	BIC	Silhouette
5	100	0.754	171.890	237.020	0.450

Note. The model is optimized with respect to the *BIC* value.

Cluster Information

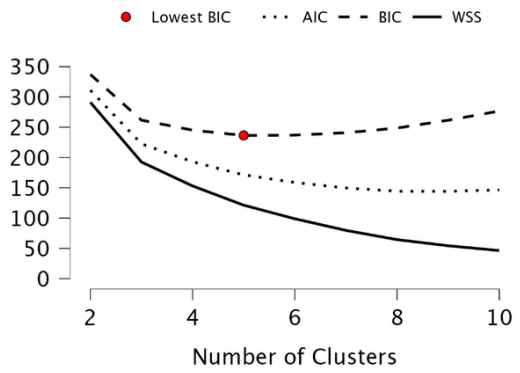
Cluster	1	2	3	4	5
Size	29	15	23	11	22
Explained proportion within-cluster heterogeneity	0.135	0.293	0.179	0.025	0.368
Within sum of squares	16.509	35.655	21.833	3.090	44.805
Silhouette score	0.599	0.276	0.495	0.727	0.193

Note. The Between Sum of Squares of the 5 cluster model is 373.11
Note. The Total Sum of Squares of the 5 cluster model is 495

Cluster Means

	Sales	Lo0tion	SPDisc	SPMore	Type
Cluster 1	0.700	0.900	−0.517	−0.488	−1.015
Cluster 2	1.213	0.900	1.566	1.417	−0.352
Cluster 3	−1.100	−1.100	−0.696	−0.678	0.629
Cluster 4	0.585	0.900	−0.677	−0.615	0.975
Cluster 5	−0.893	−1.100	0.680	0.694	0.432

Elbow Method Plot



SEGMENTATION: CLUSTER ANALYSIS (ON SALES DATA)

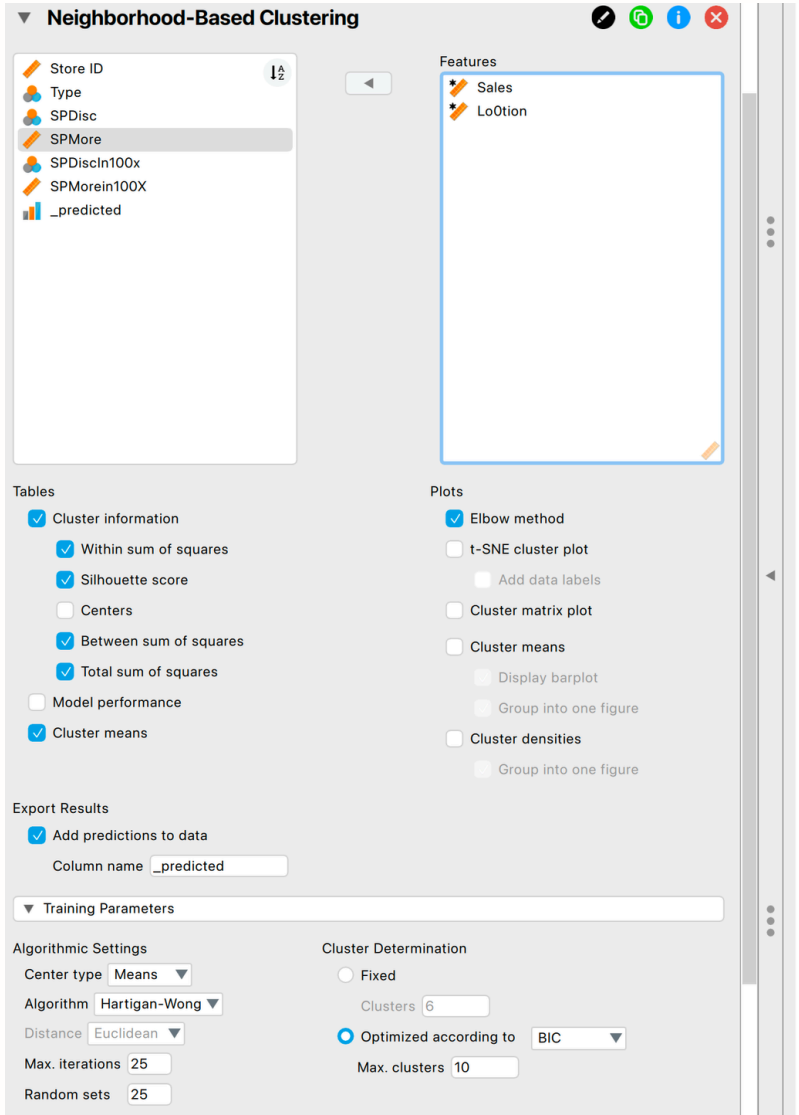
Recommendation: Cluster 1 has good sale without even Discounts and promotions . And only consideration was the store location. Cluster 4 : has good sale also without taking consideration of Discounts and promotions . And only consideration was the store location and type of store. Cluster 2: has highest sell but lesser than its promotional and discounts .So we are not recommending cluster 2

CLUSTER ANALYSIS: CLUSTERING(IMPROVED BIC VALUE)

BIC value becomes minimum only with **Location** discarding rest Type, SPDisc and SPMore

Recommended : **Location only**

Optimum no of clusters: 2



Neighborhood-Based Clustering

Model Summary: K-Means Clustering

Clusters	N	R^2	AIC	BIC	Silhouette
2	100	0.912	25.370	35.790	0.820

Note. The model is optimized with respect to the *BIC* value.

Cluster Information

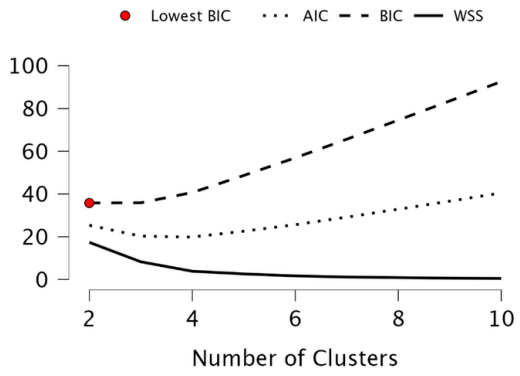
Cluster	1	2
Size	45	
Explained proportion within-cluster heterogeneity	0.655	0.3
Within sum of squares	11.376	5.9
Silhouette score	0.781	0.8

Note. The Between Sum of Squares of the 2 cluster model is 180.63
Note. The Total Sum of Squares of the 2 cluster model is 198

Cluster Means

	Sales	Lo0tion
Cluster 1	−0.999	−1.100
Cluster 2	0.817	0.900

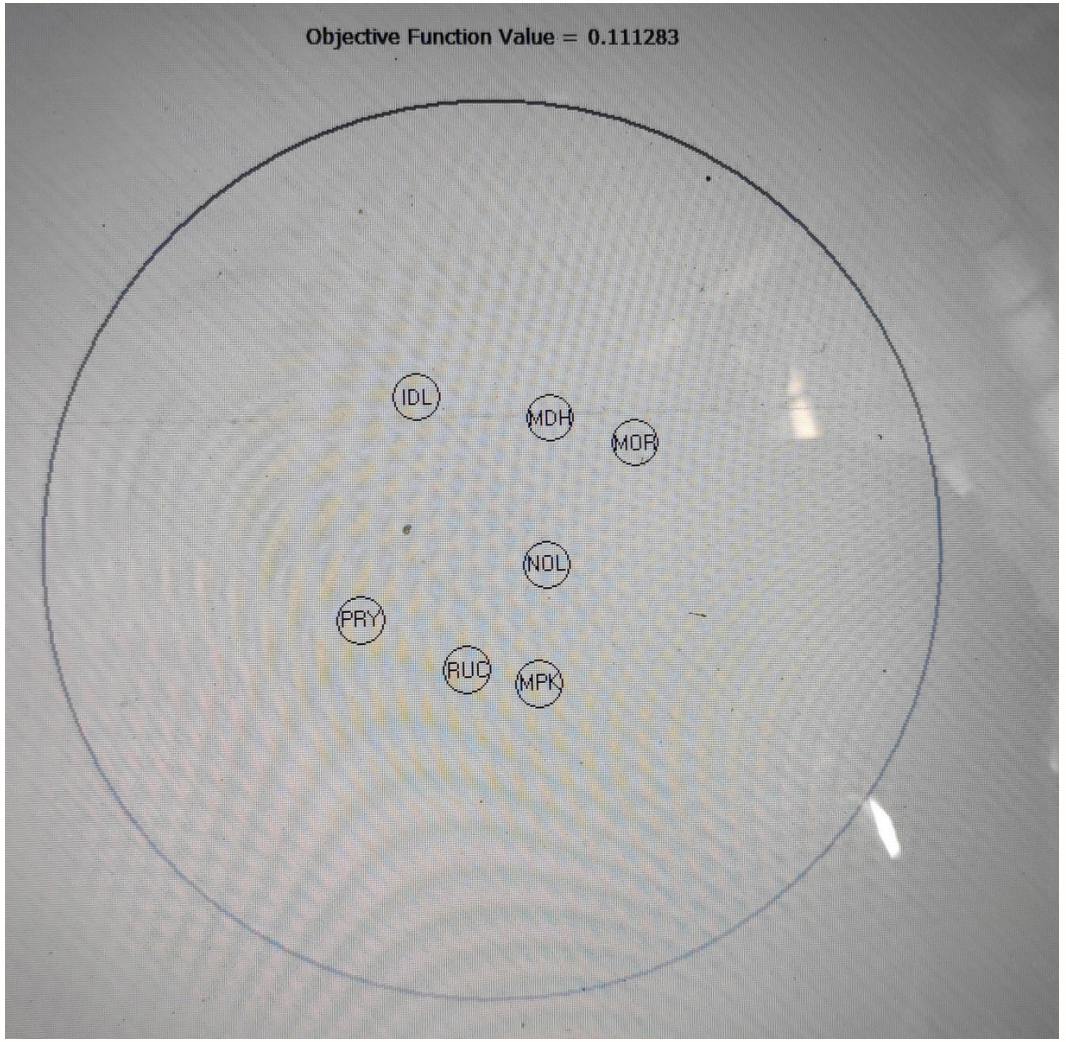
Elbow Method Plot



POSITIONING USING PERCEPTUAL MAP:(ON SIMILARITY RATING DATA)

Recommendation : Enhance 'Freshness' in the product like brand 'MDH' to be preferred brand of customers

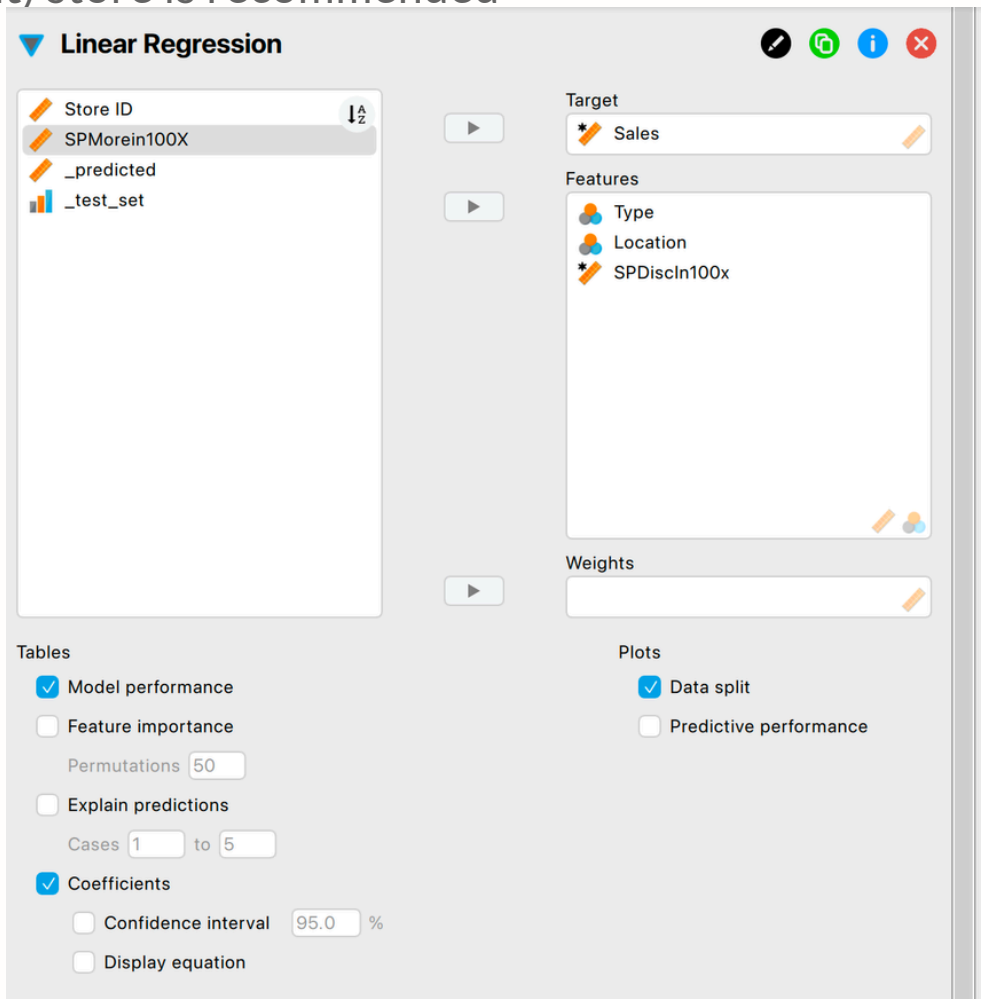
```
NOBJECTS=7
SIMILARITYLIST
MDH 7
PRY 3 7
MPK 3 4 7
RUC 3 5 6 7
NOL 4 4 5 4 7
MOR 6 2 3 3 4 7
IDL 5 3 2 3 4 3 7
```



REGRESSION ANALYSIS:

Analyzing the Impact on Sales from Location, Type of stores and Promotions and Discounts

Recommended : **Location** as **primarily residential area** has more impact on sale and **Store type** Multi-product (**department**) store is recommended



Linear Regression

Model Summary: Linear Regression

n(Train)	n(Test)	Test MSE	R^2	Adjusted R^2
80	20	321.787	0.900	0.896

Note. The trained model is saved as *predicated_sales_data_SEE.jaspML*.

Data Split

Train: 80 | Test: 20 | Total: 100

Model Performance Metrics

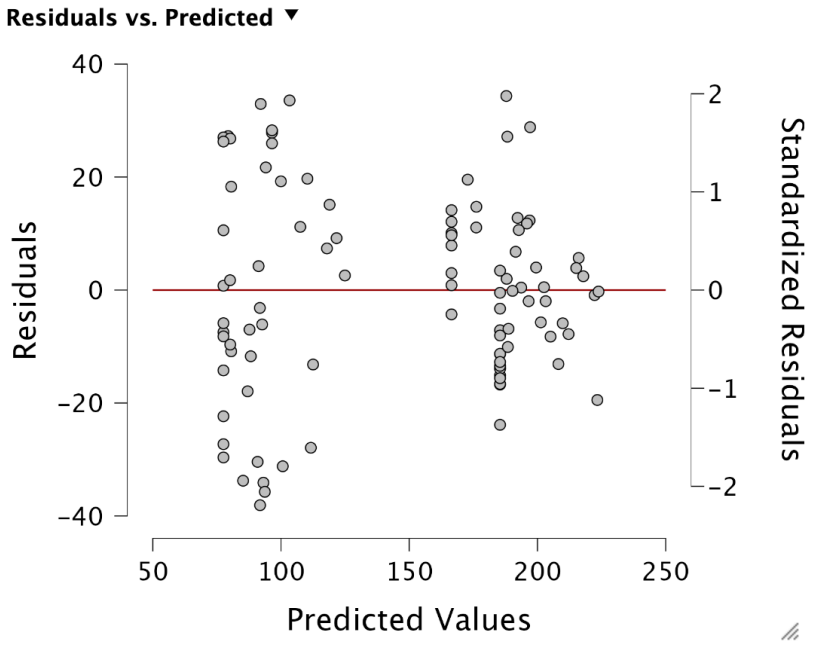
	Value
MSE	321.787
MSE(scaled)	0.129
RMSE	17.938
MAE / MAD	13.585
MAPE	12.93%
R^2	0.869

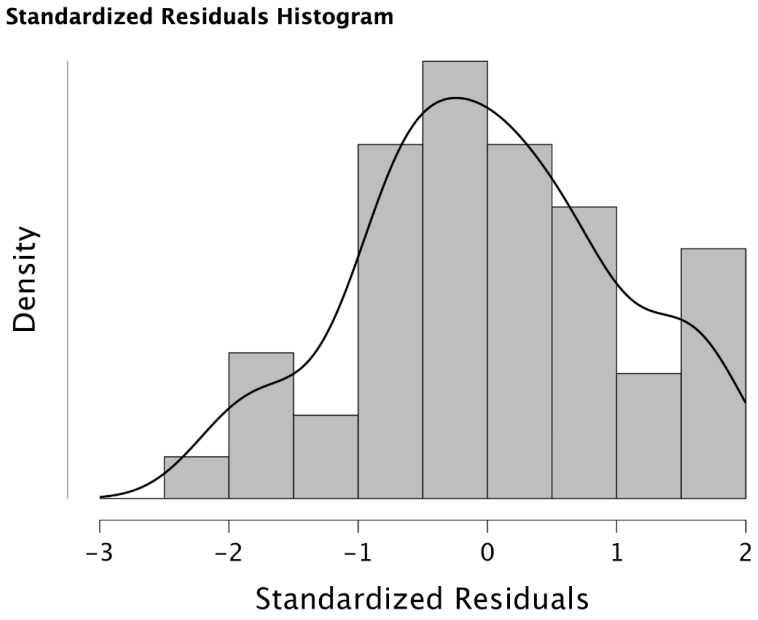
Regression Coefficients

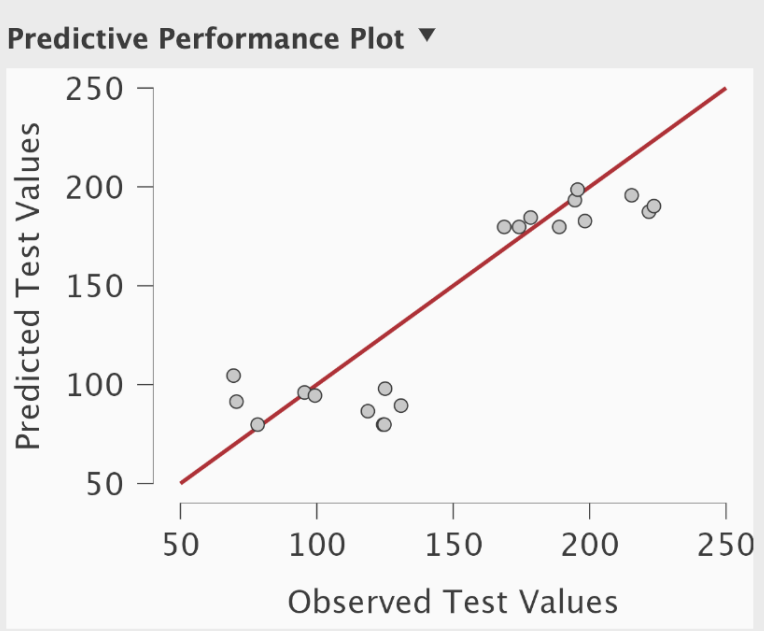
	Coefficient (β)	Standard Error	t	p
(Intercept)	102.051	4.922	20.736	< .001
Type (1)	−13.648	4.681	−2.916	0.005
Location (1)	92.330	4.732	19.512	< .001
SPDiscIn100x	13.620	2.089	6.521	< .001

REGRESSION ANALYSIS:

Residuals vs Predicted







ANOVA

Model		Sum of Squares	df	Mean Square	F	p
M_1	Regression	248614.894	2	124307.447	277.461	< .001
	Residual	43457.726	97	448.018		
	Total	292072.620	99			

Note. M_1 includes Type, Lo0tion

Note. The intercept model is omitted, as no meaningful information can be shown.

RECOMMENDATION

Data-Driven Marketing

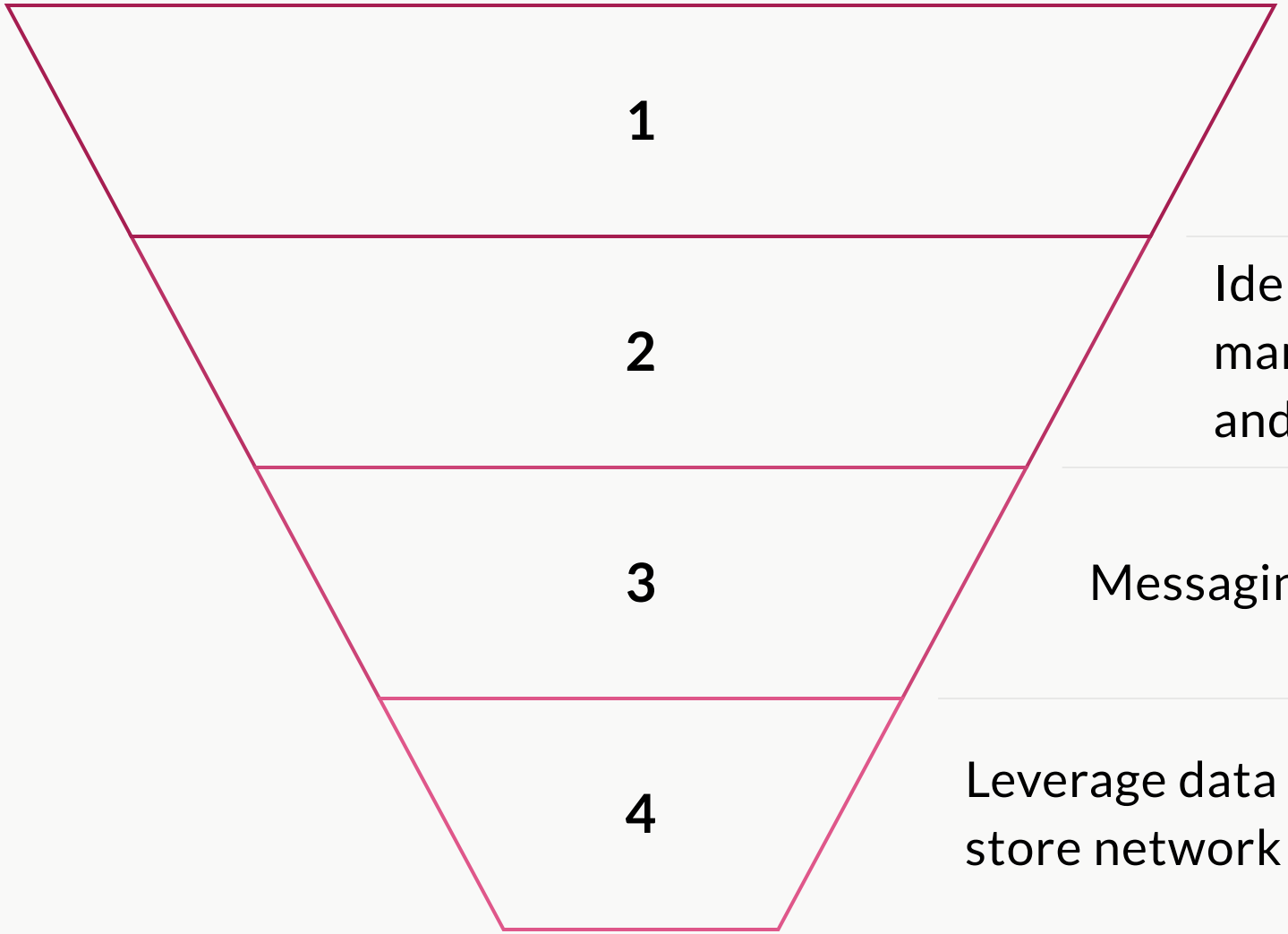
Market Segmentation Insights : Location based clusters

Enhance Brand Positioning: Make closure to 'MDH' which is known for freshness.

Choose Store type **Multi-product (department)** store

Prefer Location as **primarily residential area** has more impact on sale

OPTIMIZING STORE DISTRIBUTION



Analyze current store locations and distribution patterns

Identify underserved or high-potential market areas as **Residential area** and **Multi-product (department) store**

Messaging strategies as "**freshness**"

Leverage data analytics to monitor and adjust the store network based on evolving customer needs

THANKS

STEPS - SALES DATA ANALYSIS:

- Clean the data: Removed missing values from DiscountPromotions and MorePromotions
- Converted Categorical data (store type and store location) to Nominal data (as ‘0’ and ‘1’)
- Created a Linear Regression to check the Impact of (independent variables)Store location, store type , Discount promo and more promo on (dependent variables) Sales.
- Checked interchanging different ‘covariates’ with the ‘r^2’ value for its highest ‘r^2’ till 0.9.In the coefficient table , as ‘p’ of ‘MoreDiscounts’ > ‘0.05’ so ‘MoreDiscounts’ is removed from the model.
- From the ‘coefficient’ table, we also realized that , as coefficient of 'store Location’ is maximum and highest impact on sale, so recommendation will be to focus on choosing the ‘store location’ rather than other ‘covariates’ like ‘storeType’ or ‘MoreDiscounts’.
- Checked from the regression equation which location will have more Sale such that we can open more stores in those type of locations.Regression equation tells Location "residential area"has more sells, so recommendation is to stock more in location "residential area" to get more sales.