Price Elasticity and Sales Optimization in Retail Using Historical Data

Data Mining Final Project Presentation
Fall 2024

- Sanjana Vashisth(savash)
- Vinay Rawat(virawat)
- Harsh Gupta(hagupta)

What We Will Cover

- Objective
- Flow of work/Pricing Pipeline
- Dataset
- EDA
- Calculating Price Changes & Revenue
- Final output & Demo
- References
- Future Scope/Improvements

Objective

This project aims to develop a price recommendation system using historical sales data to assess the elasticity of demand for different products and recommend pricing adjustments to increase both revenue and profitability.

Flow of work/Pricing Pipeline

- Collecting data
- Data cleaning & feature generation
- EDA
- Calculating Elasticities
- Finding Optimal price based on constraints
- Price Recommendation

Dataset

- Mexico Toy Sales data set : Sales & inventory data for a fictitious chain of toy stores in Mexico called Maven Toys, including information about products, stores, daily transactions, and current inventory levels at each location. Files used:
  - Sales.csv
  - Products.csv
  - Store.csv
- Created additional dataset using raw data files.
- Added price variation for each day considering discounts to calculate elasticities.
- Created additional features such as actual product price revenue, expected revenue and profit which are essential for calculating new prices.

Exploratory Data Analysis

Basic Checks:

- No null data were found
- Checked for duplicates
- No Outliers were found

```
print("Dataset Shape before deduplicating:",merged_data.shape)
print("Dataset Shape after  deduplicating:",merged_data.drop_duplicates().shape)
```

✓ 0.8s

```
Dataset Shape before deduplicating: (829262, 20)
Dataset Shape after  deduplicating: (829262, 20)
```

```
merged_data.isnull().sum()
```

✓ 0.1s

Sale_ID	0
Date	0
Store_ID	0
Product_ID	0
Units	0
Year	0
Quarter	0
Product_Name	0
Product_Category	0
Product_Cost	0
Product_Price	0
Month	0
Actual_Product_Price	0
Store_Name	0
Store_City	0
Store_Location	0
Store_Open_Date	0

Exploring the dataset

- Based on various factors such as city, store, month etc, we identified:
  - Top selling items (most units sold)
  - Top profitable items
  - Top revenue generating items
  - Frequent sold together

Exploratory Data Analysis

Top Selling Items

Product_ID	Product_Name	Units
6	Colorbuds	104368
25	PlayDoh Can	103128
3	Barrel O' Slime	91663
8	Deck Of Cards	84034
19	Magic Sand	60598

Top Profitable Items

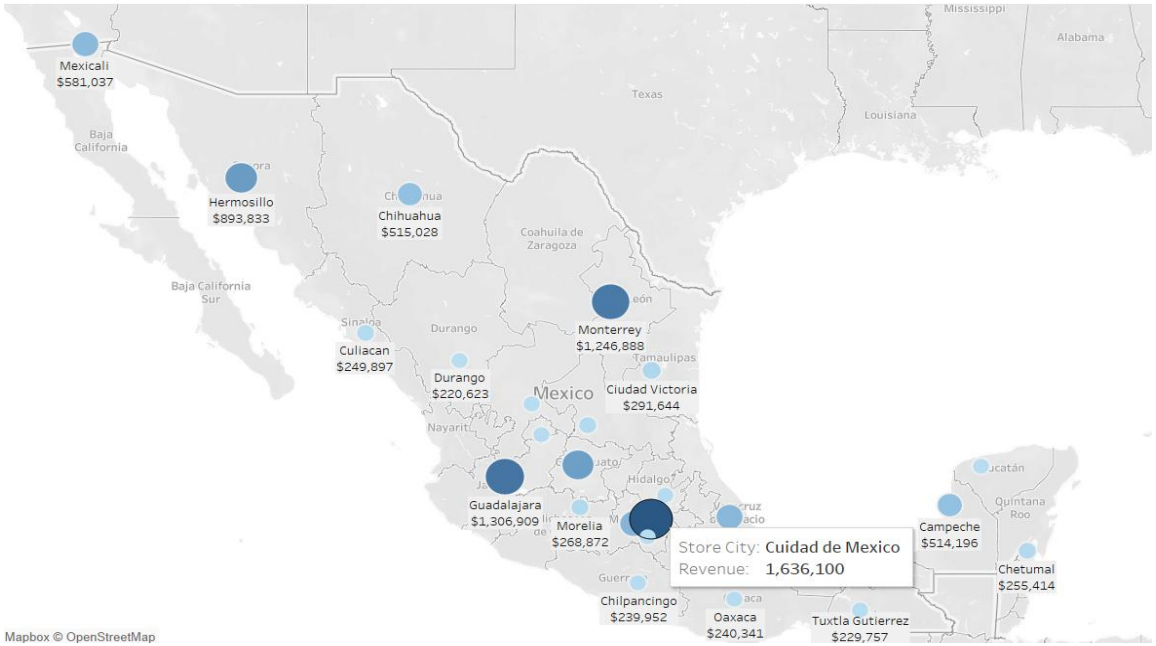
Product_ID	Product_Name	Profit
6	Colorbuds	806518.64
1	Action Figure	347825.31
8	Deck Of Cards	251917.38
3	Barrel O' Slime	217607.83
18	Lego Bricks	191680.99

Top Revenue Generating Items

Product_ID	Product_Name	Revenue
18	Lego Bricks	2281878.62
6	Colorbuds	1553544.56
19	Magic Sand	964886.18
1	Action Figure	963467.73
30	Rubik's Cube	859579.58

Exploratory Data Analysis

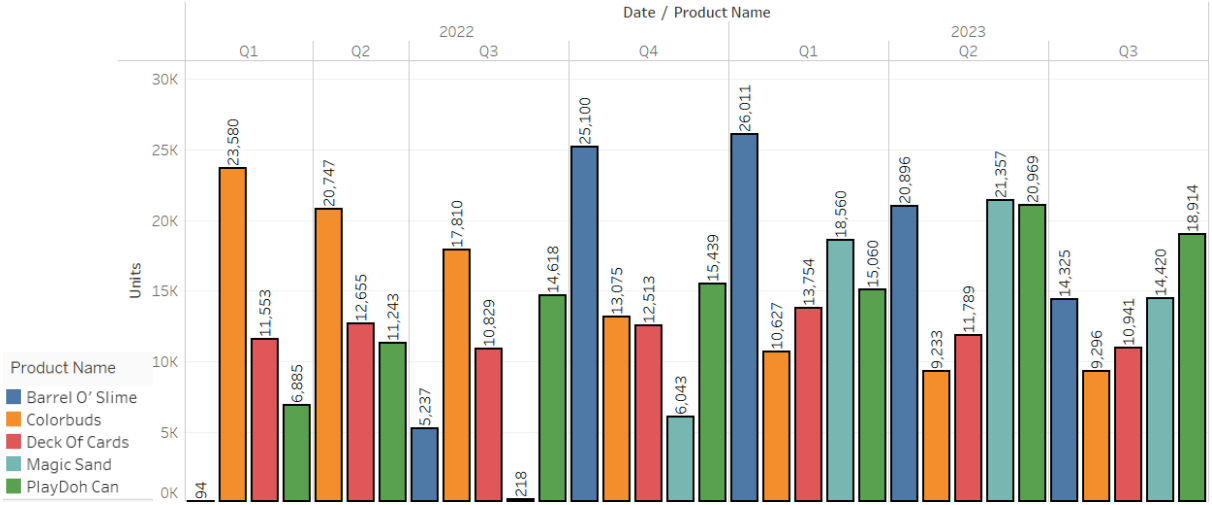
Top Revenue Generating Cities



Store City	
Cuidad de Mexico	$1,636,100
Guadalajara	$1,306,909
Monterrey	$1,246,888
Hermosillo	$893,833
Guanajuato	$859,521
Toluca	$626,588
Xalapa	$601,777
Mexicali	$581,037
Chihuahua	$515,028
Campeche	$514,196
Ciudad Victoria	$291,644
Santiago	$273,791
Morelia	$268,872
San Luis Potosi	$261,181
Villahermosa	$255,739
Chetumal	$255,414
Culiacan	$249,897
Oaxaca	$240,341
Chilpancingo	$239,952
Pachuca	$237,148
Aguascalientes	$236,773
Merida	$229,830
Tuxtla Gutierrez	$229,757
Zacatecas	$228,241
Durango	$220,623
Cuernavaca	$219,789

Exploratory Data Analysis

Units/Quarter of Top Selling Items



Observations

Slime sale was high during 2022 Q4 and 2023 Q1

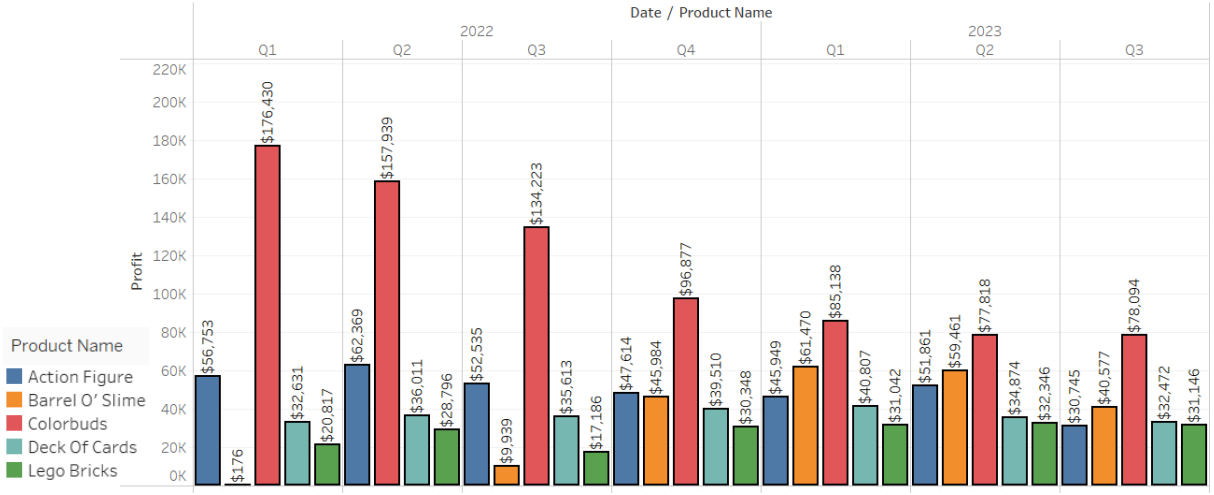
Colorbuds saw decrease in unit sales

Play Doh can saw rise quarter by quarter

	Quarter of Date						
Product Name	2022 Q1	2022 Q2	2022 Q3	2022 Q4	2023 Q1	2023 Q2	2023 Q3
Barrel O' Slime	$3.86		$3.88	$3.82	$4.35	$4.84	$4.82
Colorbuds	$14.46	$14.61	$14.52	$14.38	$15.60	$16.02	$16.00
Deck Of Cards	$6.81	$6.83	$7.28	$7.16	$7.25	$7.26	$7.26
Magic Sand			$15.47	$15.32	$15.98	$16.03	$15.94
PlayDoh Can	$2.90	$2.90	$2.89	$2.87	$3.10	$3.10	$3.09

Exploratory Data Analysis

Profit/Quarter of Top Profitable Items



Observations

Deck Of Cards had very less profit in first quarter 2022 as compared to other quarters

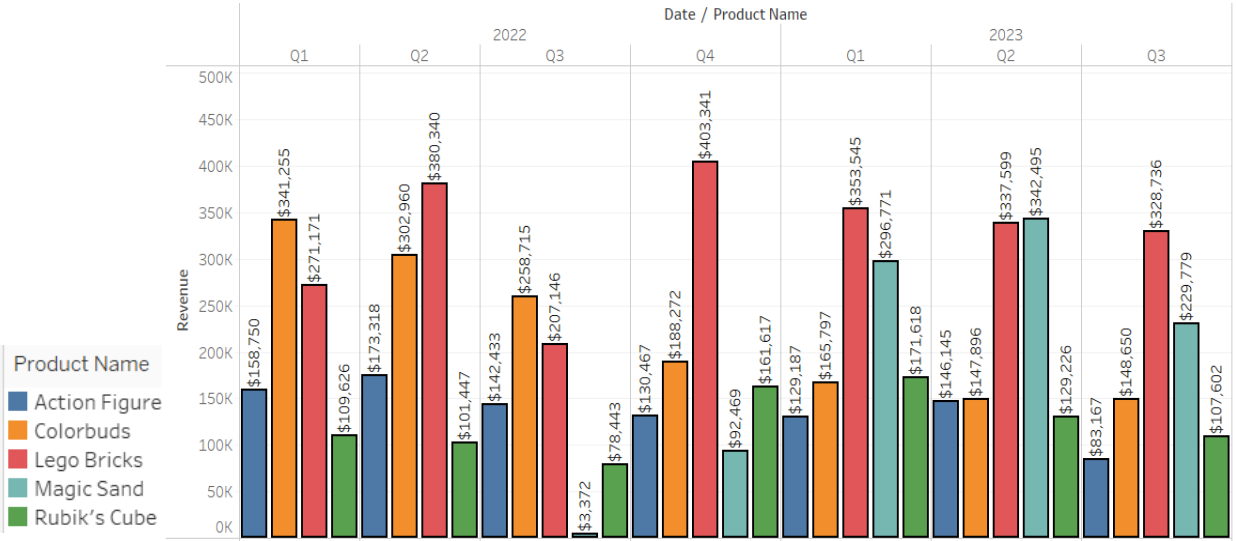
Barrel O' Slime had no profit for first 2 quarters in 2022

Quarter of Date

Product Name	2022 Q1	2022 Q2	2022 Q3	2022 Q4	2023 Q1	2023 Q2	2023 Q3
Action Figure	$15.55	$15.60	$17.41	$17.31	$17.06	$17.04	$17.45
Barrel O' Slime	$3.86		$3.88	$3.82	$4.35	$4.84	$4.82
Colorbuds	$14.46	$14.61	$14.52	$14.38	$15.60	$16.02	$16.00
Deck Of Cards	$6.81	$6.83	$7.28	$7.16	$7.25	$7.26	$7.26
Lego Bricks	$37.93	$37.88	$38.16	$37.83	$38.36	$38.68	$38.67

Exploratory Data Analysis

Revenue/Quarter of Top Revenue Generating Items



Observations

Despite no revenue in the first two quarters of 2022 and low revenue in Q3, Magic Sand became one of the most profitable items overall, peaking in Q2 2023.

Lego Bricks is consistently one of the top revenue generating items.

	Date						
	2022				2023		
Product Name	Q1	Q2	Q3	Q4	Q1	Q2	Q3
Action Figure	$15.55	$15.60	$17.41	$17.31	$17.06	$17.04	$17.45
Colorbuds	$14.46	$14.61	$14.52	$14.38	$15.60	$16.02	$16.00
Lego Bricks	$37.93	$37.88	$38.16	$37.83	$38.36	$38.68	$38.67
Magic Sand			$15.47	$15.32	$15.98	$16.03	$15.94
Rubik's Cube	$18.45	$18.38	$18.45	$18.15	$19.41	$19.39	$19.46

Calculating Price Changes & Revenue

How price change will be determined?

- Using price variation and units sold, we determine price elasticity of items.

$$Elasticity\ of\ an\ item = \frac{\%\ change\ in\ Quantity\ Demanded}{\%\ Change\ in\ Price}$$

- Based on price elasticity we calculate new price such that the new revenue is greatest and more than original revenue.

The log-log model is a popular method in econometrics for estimating price elasticity

$$\ln(Q)=\beta 0+\beta 1\ln(P)+\epsilon$$

β1 is the price elasticity coefficient.

Calculating Price Changes & Revenue

Ground rules for price changes

1. Will only consider price increases.
2. Will increase price at max by either 10% or 50 cents whichever is lower.
3. Will make sure price ends with either 5($x.x5) or 9($x.x9)
4. Will make sure unit sold loss is below 1%

Elasticities can be calculated at different levels and for next slides we have considered it as the type of location in each city.

Final Output & Demo

Store Location	Product Name	Product Price	Units	Elasticity	New Product Price	Delta Price Change	Revenue	New Revenue	Delta Revenue
Airport	Total	$15.38	16,292	-11.642	$472.66	$11.26	208,890	220,626	5,260
	PlayDoh Can	$3.19	1,928	-0.001	$3.55	$0.36	5,998	6,844	693
	Action Figure	$17.99	704	-0.001	$18.49	$0.50	12,049	13,017	640
	Magic Sand	$16.49	1,142	-0.033	$16.99	$0.50	18,297	19,383	552
	Animal Figures	$13.99	624	-0.001	$14.49	$0.50	8,293	9,041	497
	Dino Egg	$11.99	576	-0.001	$12.49	$0.50	6,686	7,194	349
Commercial	Total	$15.38	8,891	-8.859	$473.50	$12.10	117,992	125,725	3,752
	Barrel O' Slime	$4.99	858	-0.001	$5.49	$0.50	4,027	4,710	570
	Dino Egg	$11.99	675	-0.001	$12.49	$0.50	7,788	8,430	406
	PlayDoh Can	$3.19	977	-0.001	$3.55	$0.36	3,022	3,468	351
	Rubik's Cube	$19.99	521	-0.001	$20.49	$0.50	10,110	10,675	260
	Magic Sand	$16.49	520	-0.001	$16.99	$0.50	8,274	8,835	260
Downtown	Total	$15.40	12,616	-13.328	$552.61	$13.76	155,718	164,871	4,800
	Barrel O' Slime	$4.99	2,130	-0.001	$5.49	$0.50	10,173	11,693	1,212
	Magic Sand	$16.49	1,239	-0.001	$16.99	$0.50	19,879	21,050	619
	Splash Balls	$9.49	685	-0.122	$9.99	$0.50	6,336	6,799	299
	Dino Egg	$11.99	983	-0.867	$12.09	$0.10	11,210	11,798	260
	Animal Figures	$13.99	380	-0.095	$14.49	$0.50	5,072	5,487	259
Residential	Total	$15.40	11,777	-6.031	$554.05	$15.20	148,764	158,657	5,390
	Mini Ping Pong Set	$11.99	1,211	-0.001	$12.49	$0.50	14,111	15,125	605
	Dino Egg	$11.99	774	-0.001	$12.49	$0.50	8,809	9,667	533
	Kids Makeup Kit	$21.49	556	-0.001	$21.99	$0.50	11,442	12,226	463
	PlayDoh Can	$3.19	1,049	-0.001	$3.55	$0.36	3,251	3,724	377
	Action Figure	$17.99	485	-0.001	$18.49	$0.50	8,396	8,967	356

Future Scope/Improvements

- Considering additional Data
  - Price elasticities are calculated based on units and price variation, but more features can be considered
    - weather conditions can alter the data since fewer people will go for shopping
    - Age demographic can play a role.
  - Competitor's data.
- Optimizing price based on various factor such as profit maximization.
- Considering price reduction based on requirement.
- Item recommendation for stores
- Inventory recommendation

References

- https://conjointly.com/guides/understanding-price-elasticity-of-demand/
- https://thedatageneralist.com/using-machine-learning-to-estimate-price-elasticity/
- https://towardsdatascience.com/optimization-newtons-method-profit-maximization-part-3-applied-profit-maximization-23a8c16167cd
- "Price Optimization with Practical Constraints"
- Demand Impact for Prices Ending with “9” and “0” in Online and Offline Consumer Goods Retail Trade Channels
- https://mavenanalytics.io/dataplayground?accessType=open&order=date_added%2Cdesc&tags=Retail
- Python Documentation for various modules (streamlit, statmodel, pandas)


Thank You !!