Customer Satisfaction and Growth

GROUP 6

Customer Satisfaction and Growth

Our primary focus will be to address the high-level questions posed by our Head of Business Development:

- How are we doing with customer satisfaction?
- What does the data say about how we can grow our business?

Feature Engineering

'Product Profit': This calculates the profit per product (sales price - cost), which is crucial for understanding which products contribute most to the company's bottom line.

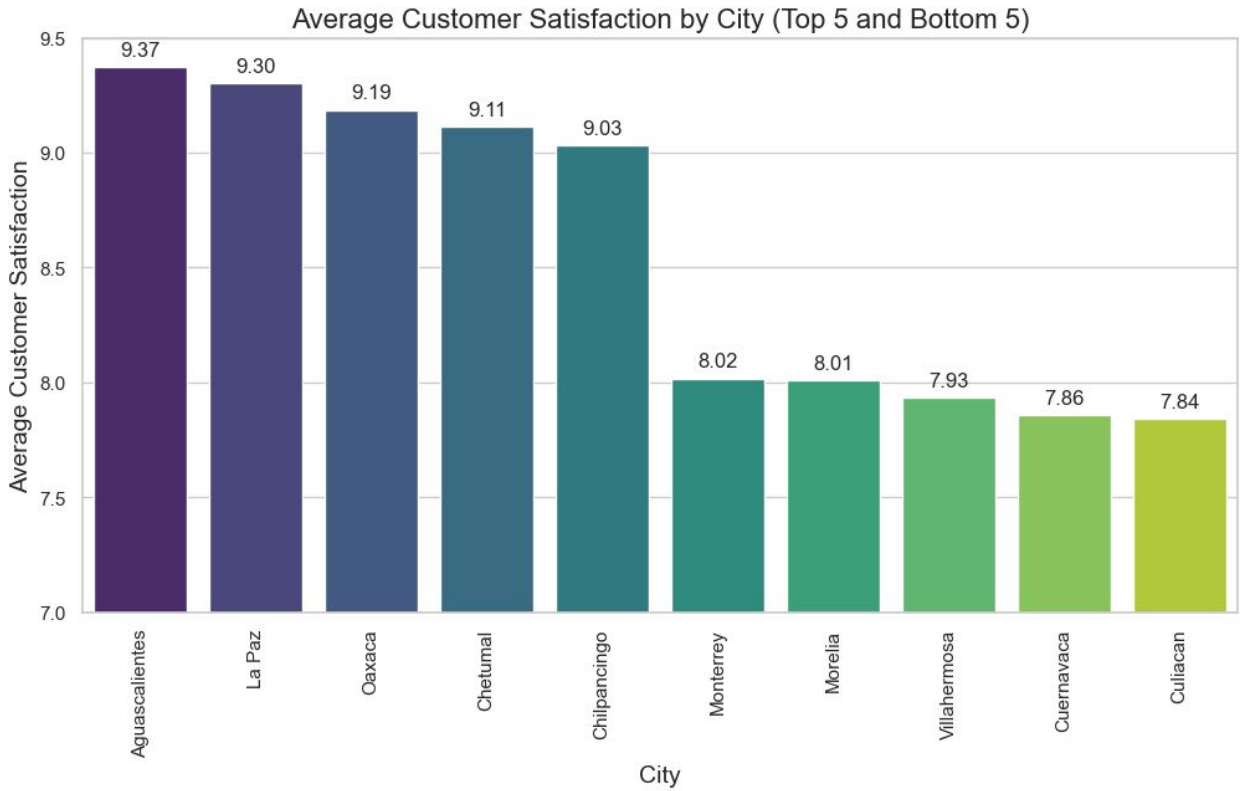
To Evaluate Growth potential we wanted combine the information in Customer Satisfaction and Profit to estimate how likely each location would be grow in the future. High values in these metrics could indicate stores or products that are both profitable and well-liked by customers, suggesting potential areas for business growth.

- Total Growth Potential: Total Profit * Customer Satisfaction
- Efficient Growth Potential: Profit Ratio * Customer Satisfaction

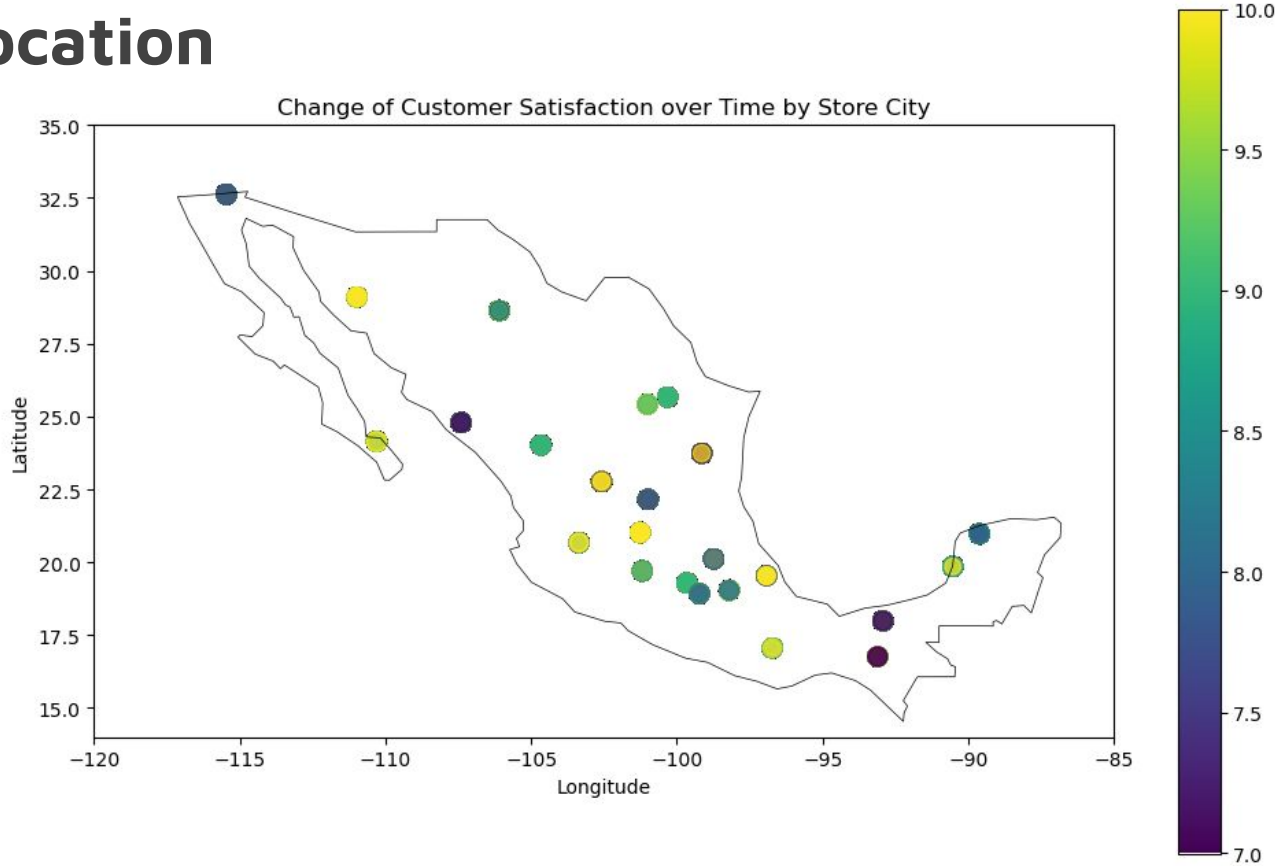
Location


Average Customer Satisfaction by Location
Average Customer Satisfaction
9.0
8.8
8.6
8.4
8.2
8.0
7.8
7.6
8.84
8.62
8.45
7.96
Residential
Commercial
Downtown
Airport
Location

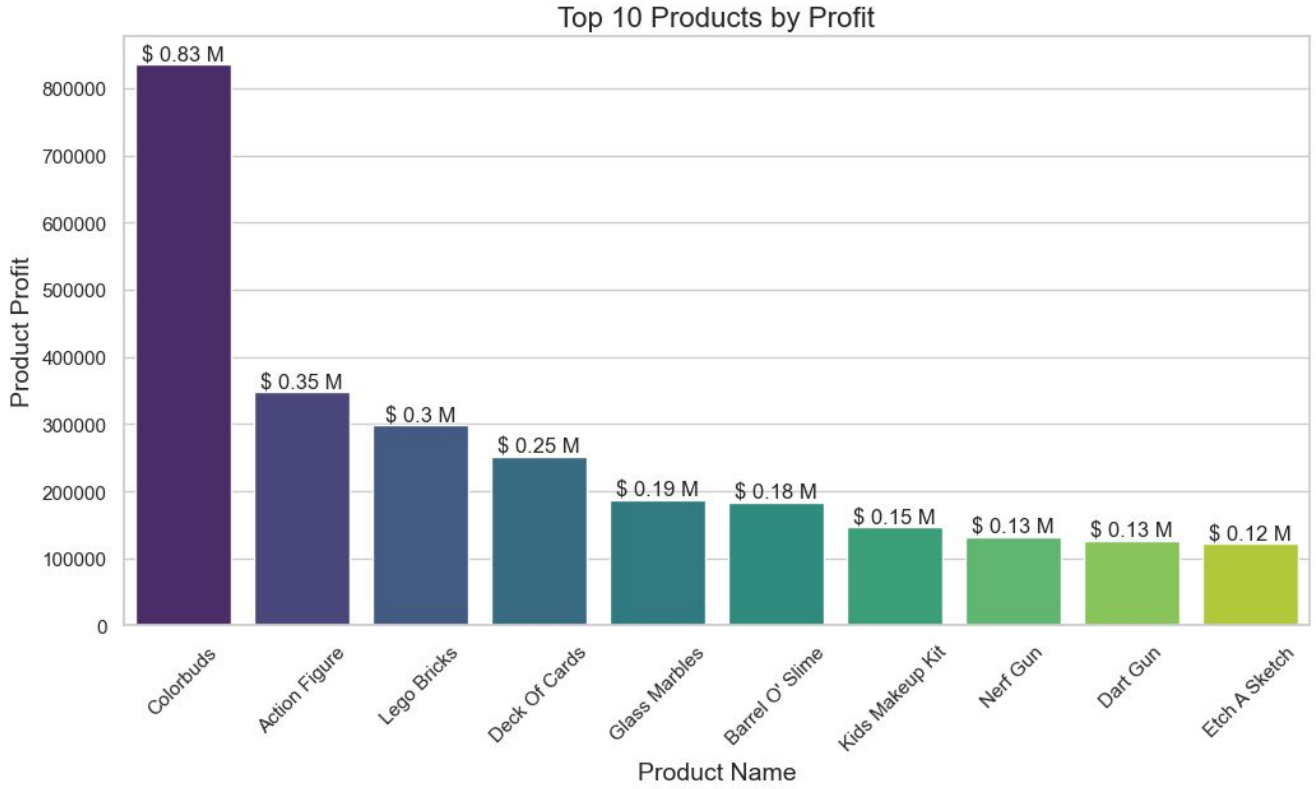
Location


Average Customer Satisfaction by City (Top 5 and Bottom 5)
Average Customer Satisfaction
9.5
9.0
8.5
8.0
7.5
7.0
9.37
9.30
9.19
9.11
9.03
8.02
8.01
7.93
7.86
7.84
Aguascalientes
La Paz
Oaxaca
Chetumal
Chilpancingo
Monterrey
Morelia
Villahermosa
Cuernavaca
Culiacan
City

Location


Change of Customer Satisfaction over Time by Store City
Latitude
Longitude
35.0
32.5
30.0
27.5
25.0
22.5
20.0
17.5
15.0
−120
−115
−110
−105
−100
−95
−90
−85
10.0
9.5
9.0
8.5
8.0
7.5
7.0

Most Profitable Products



SARIMAX Results

SARIMAX Results

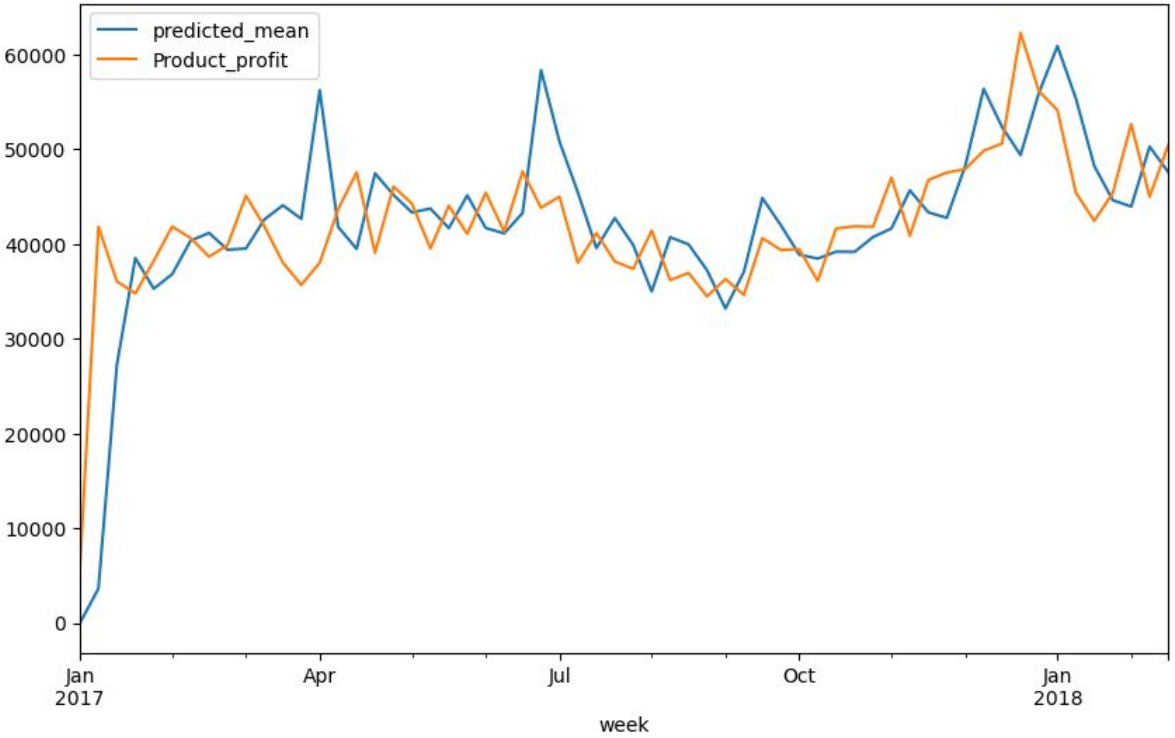
Dep. Variable:	Product_profit	No. Observations:	92
Model:	SARIMAX(1, 1, 0)x(0, 1, [1], 12)	Log Likelihood	-796.434
Date:	Fri, 23 Jun 2023	AIC	1598.868
Time:	12:23:11	BIC	1605.976
Sample:	01-01-2017	HQIC	1601.715
	- 09-30-2018		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.4162	0.104	-3.999	0.000	-0.620	-0.212
ma.S.L12	-0.9105	0.142	-6.419	0.000	-1.189	-0.632
sigma2	2.922e+07	4.16e-09	7.03e+15	0.000	2.92e+07	2.92e+07

Ljung-Box (L1) (Q):	0.35	Jarque-Bera (JB):	1.64
Prob(Q):	0.55	Prob(JB):	0.44
Heteroskedasticity (H):	1.01	Skew:	0.18
Prob(H) (two-sided):	0.99	Kurtosis:	3.61

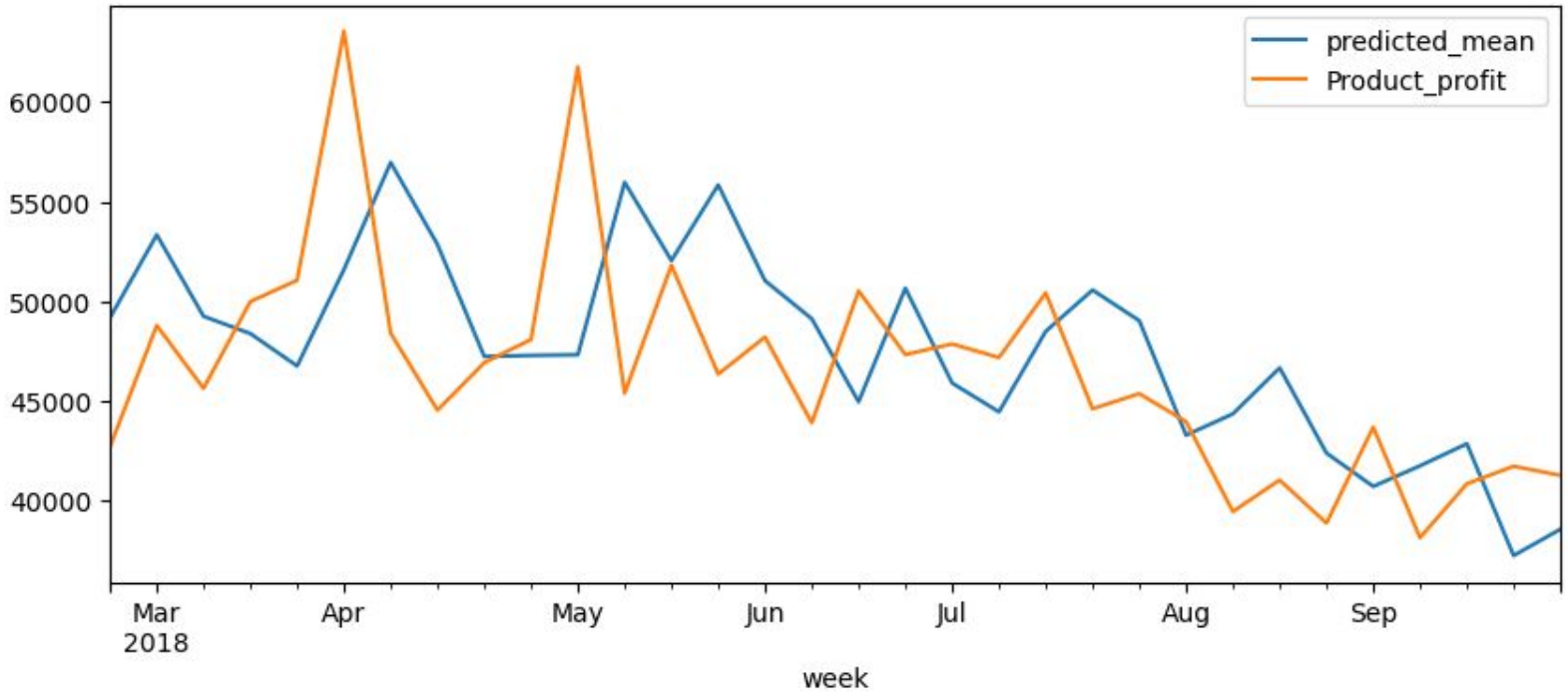
The dataset was grouped into weekly profits and placed into a seasonal ARIMA Model.
The autoregressive term has a p-value that is less than the significance level of 0.05. You can conclude that the coefficient for the autoregressive term is statistically significant, and you should keep the term in the model.

Prediction on Training Model (not overfitted)



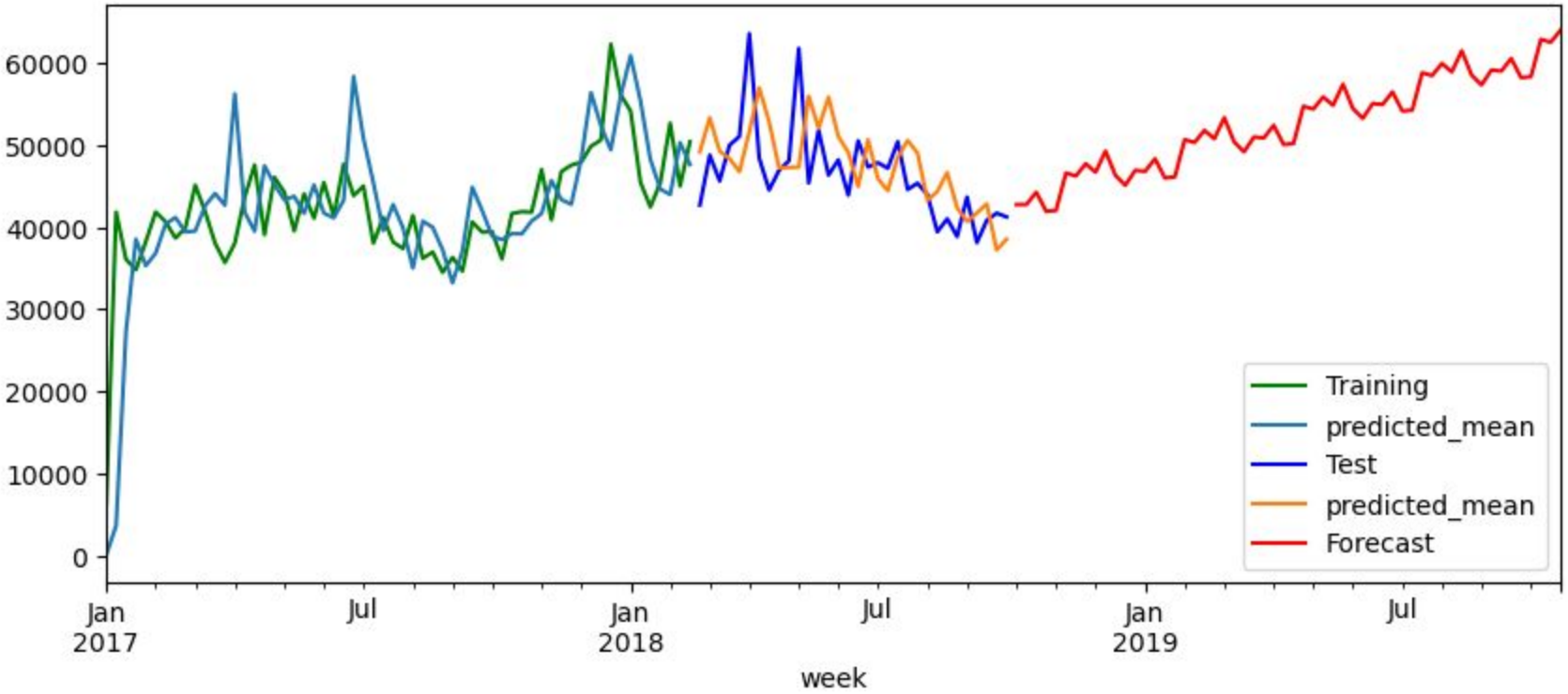
Product profits compared against the predicted mean for our training model. Since there is a decent level of error between the two graphs, the model is not overfitted.

Prediction on Test Model

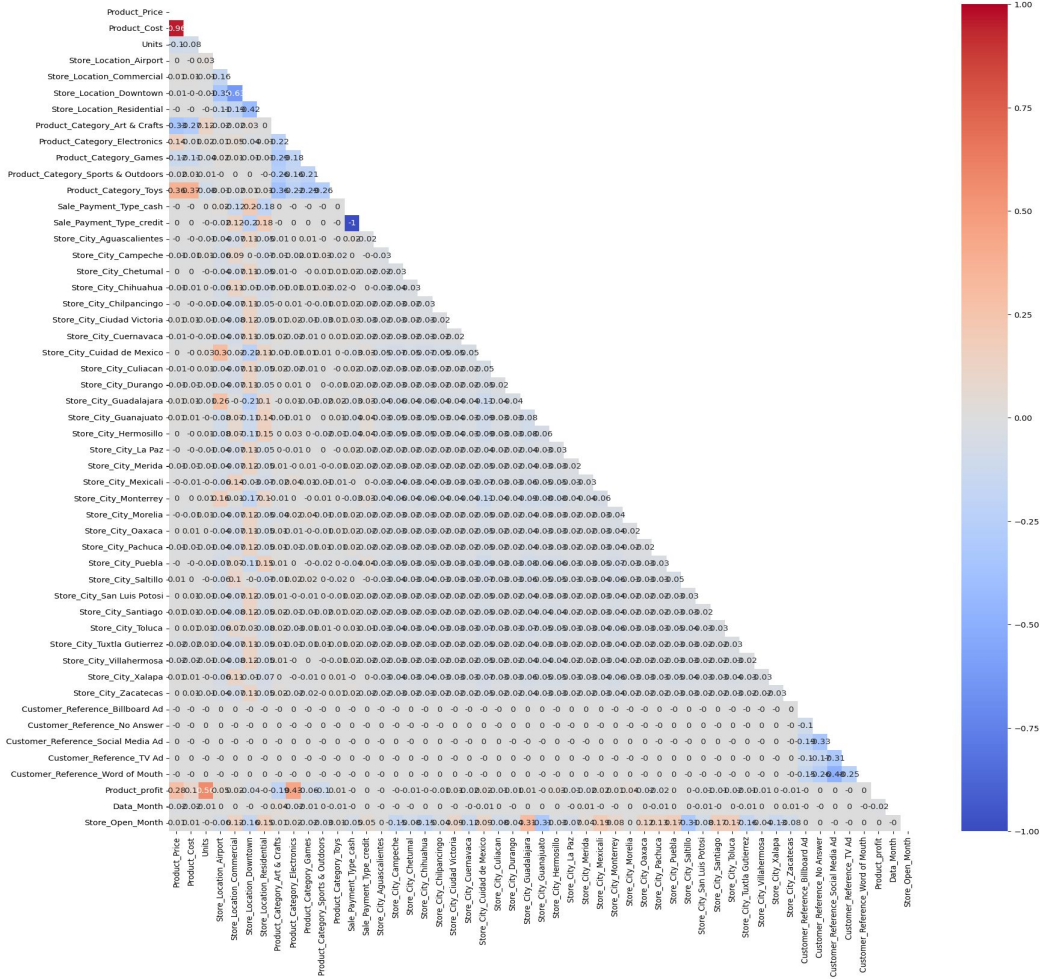


Product profits compared against the predicted mean for our test model.

Forecasted based on the Model



Forecast using our model.



47 variables

Threshold = 0.5

	Variable 1	Variable 2	Correlation Coefficient
1	Product_Price	Product_Cost	0.961101
2	Product_Cost	Product_Price	0.961101
5	Units	Product_profit	0.562843
8	Store_Location_Commercial	Store_Location_Downtown	-0.627905
9	Store_Location_Downtown	Store_Location_Commercial	-0.627905
18	Sale_Payment_Type_cash	Sale_Payment_Type_credit	-1.000000
19	Sale_Payment_Type_credit	Sale_Payment_Type_cash	-1.000000
55	Product_profit	Units	0.562843

Linear Reg Model

- **R-squared = 0.006**

Dependent variable: Customer satisfaction

Independent variables:

'Store_Location_Airport',
'Store_Location_Downtown',
'Store_Location_Residential',
'Product_profit',
'Store_Open_Month'

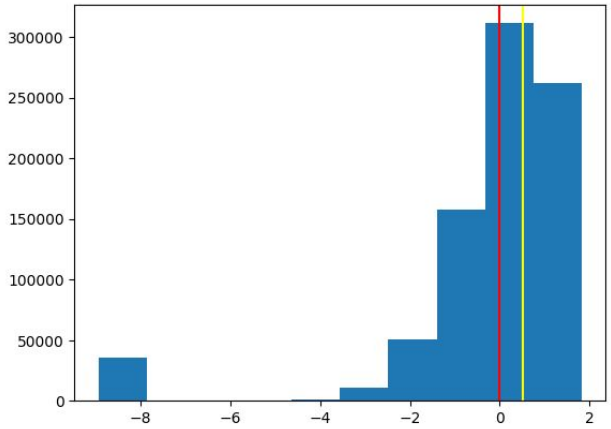
OLS Regression Results

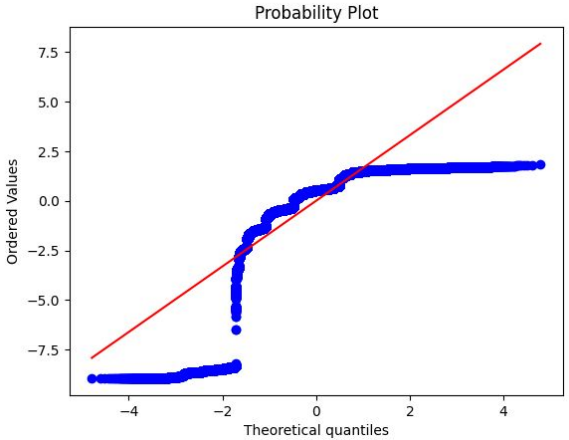
Dep. Variable:	Customer_Satisfaction	**R-squared:**	0.006
Model:	OLS	**Adj. R-squared:**	0.006
Method:	Least Squares	**F-statistic:**	950.9
Date:	Fri, 23 Jun 2023	**Prob (F-statistic):**	0.00
Time:	13:20:01	**Log-Likelihood:**	-1.7810e+06
No. Observations:	829262	**AIC:**	3.562e+06
Df Residuals:	829256	**BIC:**	3.562e+06
Df Model:	5		
Covariance Type:	nonrobust		

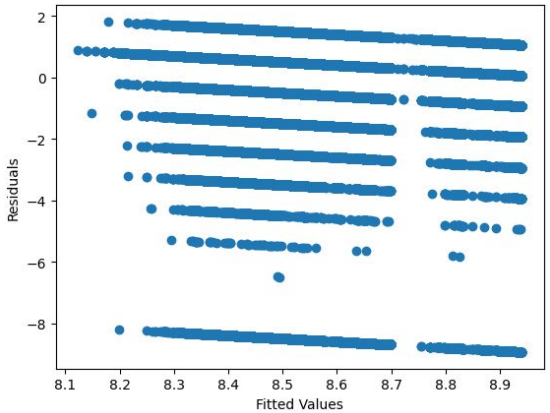
	coef	std err	t	P>|t|	[0.025	0.975]
const	8.7471	0.008	1139.783	0.000	8.732	8.762
Store_Location_Airport	-0.3239	0.009	-34.710	0.000	-0.342	-0.306
Store_Location_Downtown	-0.1856	0.006	-32.360	0.000	-0.197	-0.174
Store_Location_Residential	0.2558	0.008	30.970	0.000	0.240	0.272
Product_profit	-0.0020	0.001	-4.019	0.000	-0.003	-0.001
Store_Open_Month	-0.0152	0.001	-21.645	0.000	-0.017	-0.014

NOT a good linear model - very weak relationship

- Shapiro-Wilk test on residuals: p-value = 0.0
- Not a normal distribution
- Left-skewed
- Potential presence of heteroscedasticity

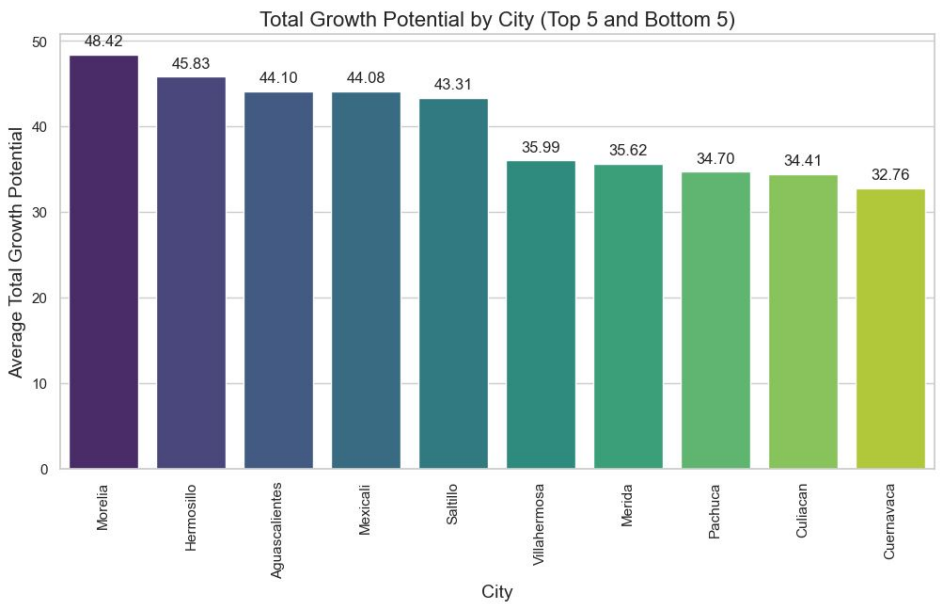






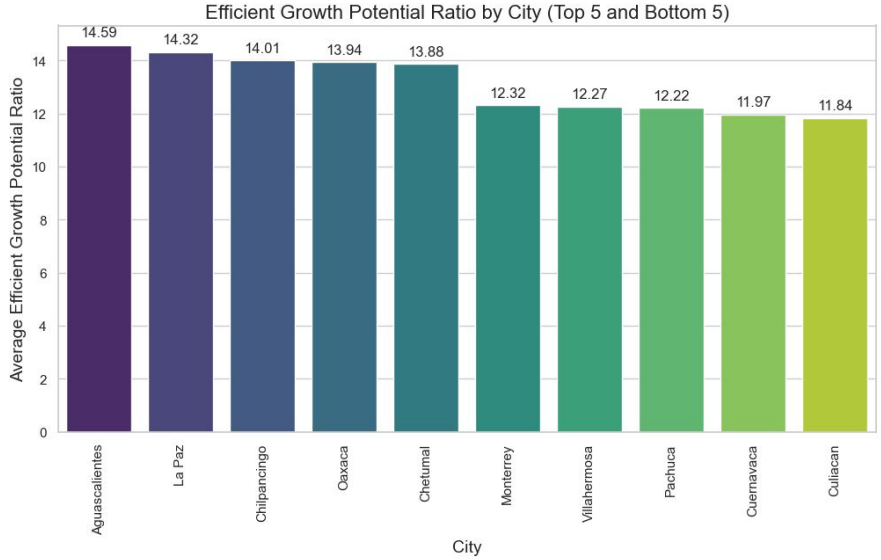
City Growth Potential

Total Profit * Customer Satisfaction



The top 5 are Morelia, Hermosillo, Aguascalientes, Mexicali, Saltillo

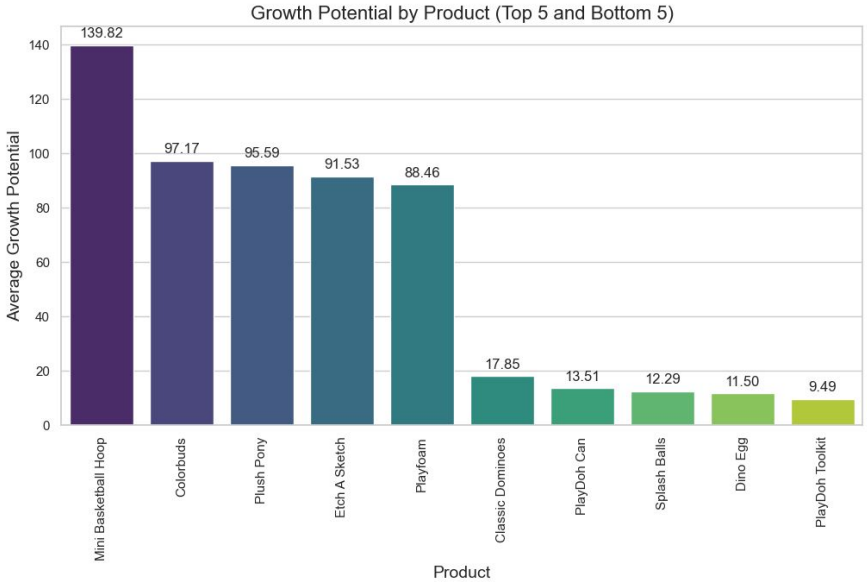
Profit Ratio * Customer Satisfaction



The top 5 are Aguascalientes, La Paz, Chilpancingo, Oaxaca, Chetumal

Product Growth Potential

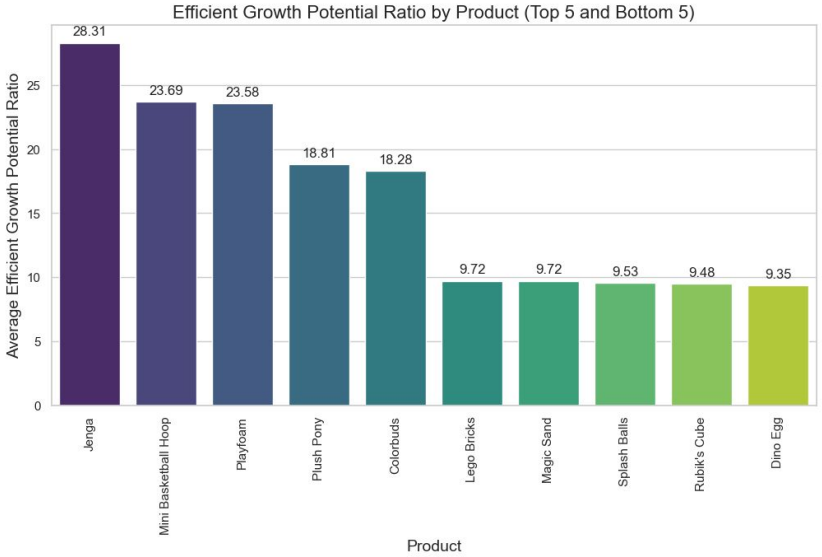
Total Profit * Customer Satisfaction



The top 5 are Mini Basketball Hoop, Colorbuds, Plush Pony, Etch A Sketch, Playfoam,

The bottom 5 Classic Dominoes, PlayDoh Can, Splash Balls, Dino Egg, PlayDoh Toolkit

Profit Ratio * Customer Satisfaction



The top 5 are Jenga, Mini Basketball Hoop, Playfoam, Plush Pony, Colorbuds

The bottom 5 Lego Bricks, Magic Sand, Splash Balls, Rubik's Cube, Dino Egg