FP-Final-Project Report

Problem Statement

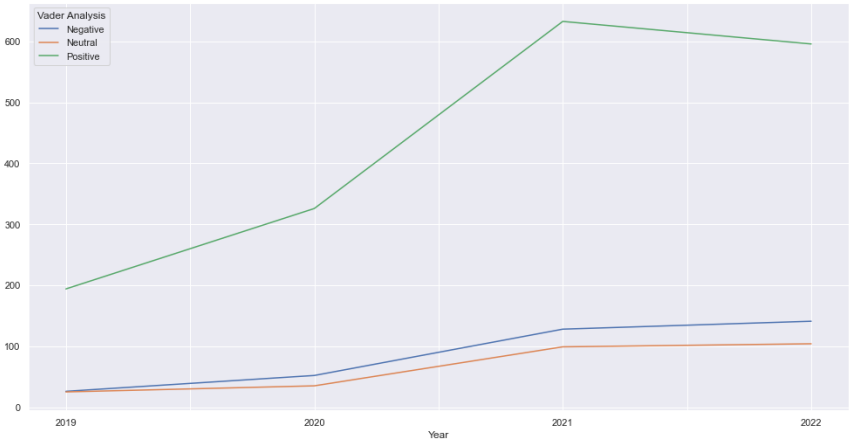
Preparedness for the forthcoming automobile product business opportunity by analyzing consumer comments and sentiments along with OEM sales insight for the period 2019-2022.

Data collection and Preprocessing

Using Python API BeautifulSoup we have collected consumer review records 2359 for automobile products such as tyres, wiper, seatcover, engine oil etc. We also gathered sales turnover of automobile products for the period between 2019-2022.

Methodology and Algorithms

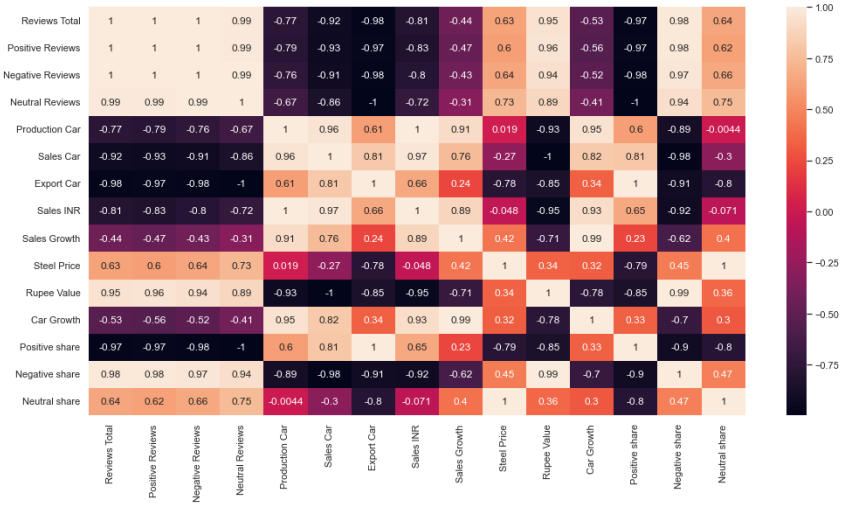
- Sentiment analysis: Using NLP libraries we deduced positive, negative, and neutral sentiments score.



- Economic factors: Steel price, Rupee value
- Production, sales, and export of four wheelers for the period FY2019-FY2022
- Correlation and Feature selection: To identify factors reflecting influence on sales
- Regression: Conducted OLS regression over auto component sales with identified features.

Model Validation and Evaluation

Correlation among factors



Evaluation

OLS Regression Results

Dep. Variable:	Sales Growth	R-squared (uncentered):	0.981
Model:	OLS	Adj. R-squared (uncentered):	0.942
Method:	Least Squares	F-statistic:	25.52
Date:	Sat, 13 Aug 2022	Prob (F-statistic):	0.139
Time:	13:29:14	Log-Likelihood:	8.3215
No. Observations:	3	AIC:	-12.64
Df Residuals:	1	BIC:	-14.45
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Positive share	0.1506	0.029	5.243	0.120	-0.214	0.516
Car Growth	1.4683	0.206	7.144	0.089	-1.143	4.080

Omnibus:	nan	Durbin-Watson:	2.136
Prob(Omnibus):	nan	Jarque-Bera (JB):	0.356
Skew:	0.386	Prob(JB):	0.837
Kurtosis:	1.500	Cond. No.	10.6

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Limitation

Derived ML model has limited references of other factors which influences sales turnover, such as:

- Data pertaining to automobile parts OEMs in public domain
- Government policy on geo economy factors such export/import duty, road tax etc. and infrastructure

Conclusion

Derived ML model does provide future inference of sales in correlation with consumer's sentiments and four-wheeler's sales. Based on sentiment score in conjunction with car sales perceived from the collected dataset, we can conclude FY22 sales turnover will see rise in comparisons to FY21.

References

https://www.amazon.in, https://www.acma.in/auto-component.php, https://marketresearch.biz, https://auto.economictimes.indiatimes.com/tag/auto+parts

Github link (Project artifact repository)

https://github.com/Bhargavi-6/group1_fp1_final_project