



# A Geospatial Competitive Intrusion and Market Expansion Solution for Large Chain Retailers

Agraj Agarwal, Mohana Kodipaka, Sessa Sai Tumma, Tanaya Chitaldurg, Matthew Lanham

Purdue University Krannert School of Management

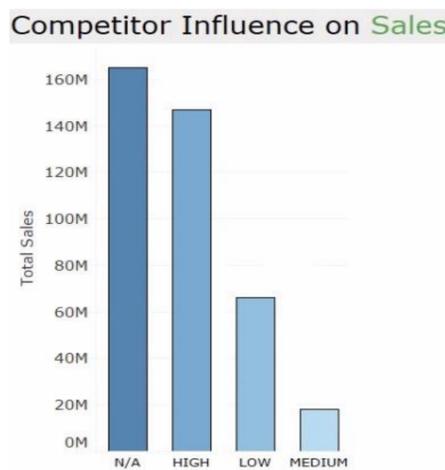
agarw214@purdue.edu; mkodipak@purdue.edu; tumma@purdue.edu; tchitald@purdue.edu; lanhamm@purdue.edu

## Abstract

There is a critical need for strategic expansion across all industries. This poster talks about data-driven approach to identify store expansions for parts retailers. Our prediction model aided by detailed competitor analysis, on top of an easily assessable geospatial tool, can help a company's real estate team identify opportunities and most likely future threats. The audience for our work is anybody using data science for a retailer chain.

## Introduction

For retail chains to try and stay ahead of the curve they need to be leaders at identifying expansion opportunities of their business and their competitors. Analyzing competitor's presence is essential to get better insights on sales patterns and volumes. Outcomes from this work assists the retailer to predict the intrusion of its competitors in different markets and strategically identify the most ideal locations for its new stores, giving them an edge over the competition.



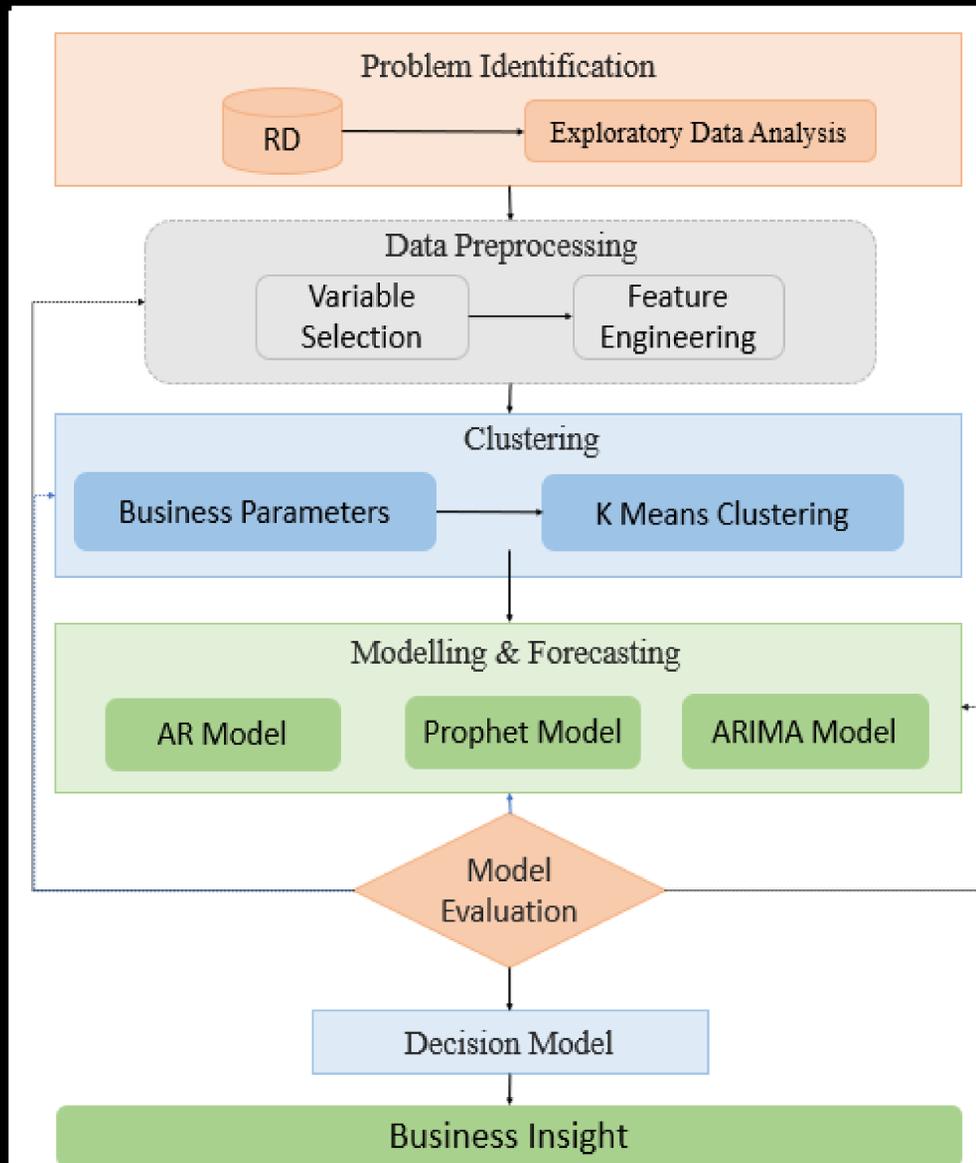
## Research Questions

- 1. Identify best suited locations for store expansion in a state.
2. Predict competitive intrusion in different markets for given geographical locations.

## Literature Review

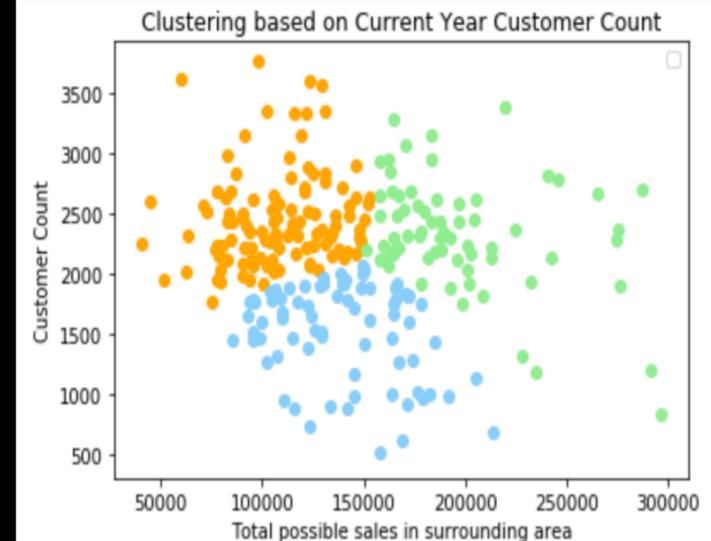
Table with 3 columns: Author, Paper, Methodology. It lists two papers: 'Optimal Retail Location - Empirical Methodology & Application' and 'General Sales Forecast Models for Automobile Markets and their analysis'.

## Methodology



EDA was used to identify parameters that impact sales volumes. Based on the factors identified, variable selection was performed, and data was scaled. Additionally, business context was leveraged to identify the parameters to be used for KMeans clustering which was implemented to identify optimal clusters for the given data. AR/ARIMA Modelling was used to forecast sales for each store. A sequence of transformations was employed to address non-stationarity (e.g. log, rolling means subtraction, exponential decay differencing). After considering weighted average of the selected parameters, we identified the top 5 market expansion sectors.

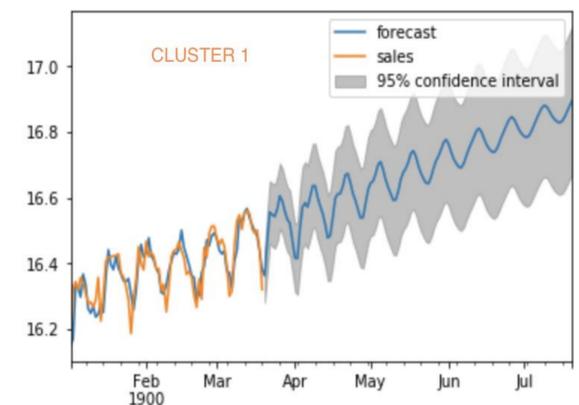
## Results



After considering various significant parameter combinations, the optimal clusters were obtained by clustering data using customer count and total possible sales in surrounding areas. The decision of the optimal number of clusters i.e. 3, was obtained using the elbow plot.

## Cluster Forecasts

The best cluster was found by forecasting sales using several time series techniques. AR model resulted in the best fit sales forecast for next 2 years with 95% confidence interval. The locations in this cluster were then analyzed using weighted factors and sensitivity analysis was performed to arrive at top 5 locations for expansion.



## Conclusion

We were able to develop a model that takes several statistical parameters from the raw data and identifies the best sectors for market expansion. Through detailed tableau visualizations, we can further identify sectors with high potential for competitor expansion. This model can be replicated across different industries and can be leveraged by various brick & motor retailers to identify market expansion locations.

## Acknowledgement

We thank Prof. Matthew Lanham for his constant guidance on this project.