# Future Sales Forecasting Using Machine Learning Techniques

Mohammed Bazlamit, Alan Wang, Sasha Patil, Nur Farahiyah Joon

PURDUE UNIVERSITY | The Data Mine

Helmer SCIENTIFIC

## Background

- Helmer specializes in medical equipment, such as Freezers, Plasma Thawers, Cell Washers, etc.
- Helmer's sales are correlated to seasonal factors as well as outbreaks of new viral diseases such as COVID-19.
- Helmer is competing on the domestic front, but international sales remain a steady percentage of the worldwide business, with a promising future.
- The task is to project future equipment sales using machine learning algorithms based on Helmer's previous sales history.
- Forecasting results will:
  - give Helmer the ability to account for future manufacturing line strategic planning and logistics
  - to have a better understanding of what Helmer should expect in terms of the future growth and
  - analyze the need for more working power to accommodate their growth.

## Research Methodology

- Started with exploratory data analysis.
- Real time prediction models.
- Develop models with further variables correlated with current sales.

Worked on:

- ARIMA models
- Non time series models (Random forest)
- Time series models that are not based on ARIMA approach
- Grid search through best performance model by comparing the RMSE of different models.

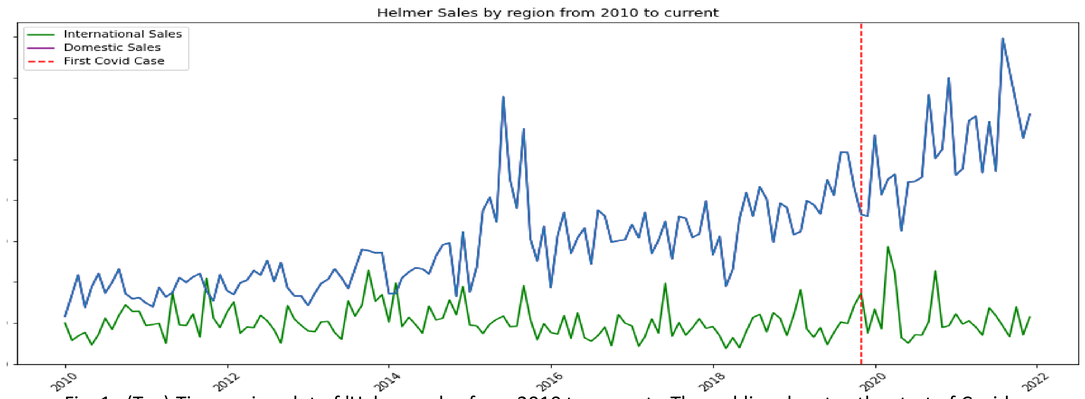


Fig. 1. (Top) Time series plot of 'Helmer sales from 2010 to current. The red line denotes the start of Covid. (Right) Bar histogram showing the overall frequencies of units sold over the entire range of primary group of products. The analysis reveals the dominance of domestic sales as compared to international sales.

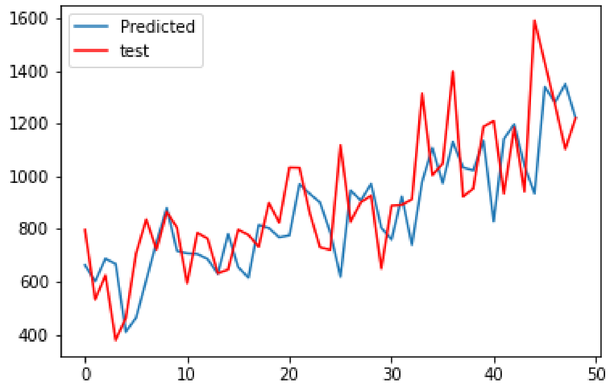


Fig. 3. Cross Validation plot of predicted vs observed (Test) values produced via Arima modeling

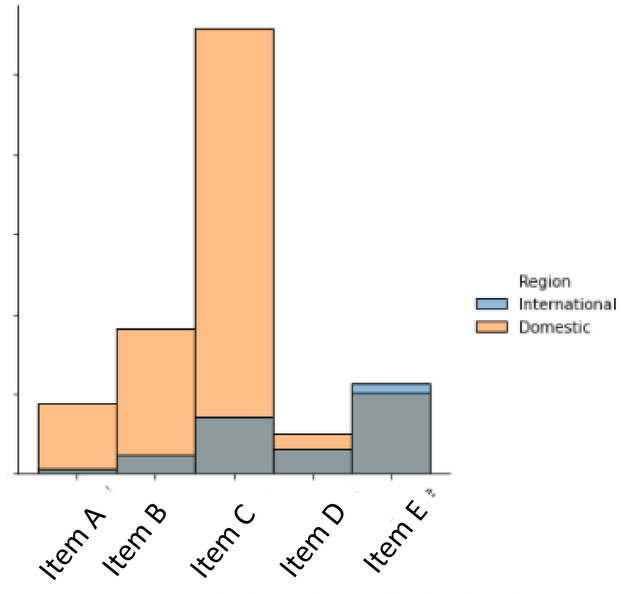


Fig. 2. Bar Histogram displaying the overall sales of products from different family groups from 2010 to current

## Conclusion and Future goals

- ARIMA model analysis onto lower hierarchy data column is necessary to make more accurate predictions.
- Covid Pandemic changed the normal sales trend and adding more complexity in the predictions.

**Future goals:**

- Real time prediction models
- Develop models with further variables correlated with current sales.

## References and Acknowledgement

•step-by-Step Graphic Guide to Forecasting through ARIMA Modeling using R – Manufacturing Case Study Example, Roopam Upadhyay

•How to Develop Multivariate Multi-Step Time Series Forecasting Models for Air Pollution, Jason Brownlee

•How to Make Out-of-Sample Forecasts with ARIMA in Python, Jason Brownlee

•ARIMA Model – Complete Guide to Time Series Forecasting in Python, Selva Prabhakaran

We would like to express our gratitude to the Helmer Scientific team who were kind enough to be patient with the students and their feedback, to the Purdue Data Mine Staff who provided support and help, as well as the team members for the all the work that has been done till this point.