AMERICAN UNIVERSITY OF BEIRUT

OPTIMIZING OPERATIONS OF LEBANESE STEEL COMPANY: FOZ TRADING USING DEMAND FORECASTING AND OPTIMAL ORDERING POLICY

by
AHMAD AYMAN CHEHAB

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Engineering Management
to the Department of Industrial Engineering and Management
of the Maroun Semaan Faculty of Engineering and Architecture
at the American University of Beirut

Beirut, Lebanon
June 2020

AMERICAN UNIVERSITY OF BEIRUT

OPTIMIZING OPERATIONS OF LEBANESE STEEL COMPANY: FOZ TRADING USING DEMAND FORECASTING AND OPTIMAL ORDERING POLICY

by

AHMAD AYMAN CHEHAB

Approved by:

[Signature]

Dr. Hussein Tarhini, Assistant Professor — Advisor
Industrial Engineering and Management

[Signature]

Dr. Bacel Maddah, Professor and Chair — Member of Committee
Industrial Engineering and Management

[Signature]

Dr. Walid Nasr, Associate Professor — Member of Committee
Business Information and Decision Systems

Date of thesis defense: June 18, 2020

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name:	Chehab	Ahmad	Ayman
	Last	First	Middle

⊗ Master's Thesis ◯ Master's Project ◯ Doctoral Dissertation

☐ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☒ I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes after:

Three ---- years from the date of submission of my thesis, dissertation, or project.

July 3, 2020

Signature Date

ACKNOWLEDGMENTS

This Thesis would have not been possible without the support of many people. Henceforth, I would like to extend my appreciation to the following:

Thanks to my parents, family, and every loved one who endured with me and supported me throughout this endeavor. It would not have been possible without their backing and support.

Special thanks to Dr. Hussein Tarhini, for making this thesis conceivable. His guidance, support, advice, brilliance, and tremendous effort are greatly appreciated. Without his assistance this case study would not have been possible.

Thanks to Dr. Bacel Maddah and Dr. Walid Nasr for their contribution and provision throughout this thesis and for suggesting some of the most beneficial resources and remarks.

Thanks to Mr. Fawzi Foz and Mr. Fawaz Foz for supporting and assisting us with the needed information and data regarding FOZ Trading.

AN ABSTRACT OF THE THESIS OF

Ahmad Ayman Chehab for Master of Engineering Management
Major: Engineering Management

Title: Optimizing Operations of Lebanese Steel Company: FOZ Trading Using Demand Forecasting and Optimal Ordering Policy

This thesis examines the various industry-used methods in time series forecasting for the Lebanese tool steel company: FOZ Trading. Previous sales data of FOZ Trading showed signs of intermittent demand in all of the categories examined. The limitations of usual forecasting methods in the case of intermittent demand, such as simple exponential smoothing and simple moving average, has prompted the use of intermittent demand-specific approaches along with several basic and traditional forecasting methods.

Basic forecasting methods such as last period demand and simple moving average are benchmarked against traditional and alternate forecasting methods such as Box-Jenkins, Croston, Croston TSB and the simple exponential smoothing. Moving block bootstrapping and circular block bootstrapping were also applied on top of Croston TSB and simple moving average in order to check if they would perform better. Locally weighted linear regression and gaussian process regression were tested and evaluated for suitability as part of the machine learning methods. Additionally, temporal aggregation was applied to reduce the intermittency aspect of the data so that basic and traditional forecasting methods would perform better.

As part of the evaluation process, bias correction and prediction intervals were generated to improve the performance and usability of the forecasts. All the methods were implemented using python and then tested and validated using the walk forward optimization. The best methods to use, based on the root mean squared error and the mean absolute error, were selected and then applied in the order up-to model to optimize the inventory level of the company. The whole demand forecasting and optimal ordering process was automated using python and Jupyter Notebook. The final application was provided to the company to use in their demand planning process. This allowed FOZ Trading to adjust their stocks and orders according to the forecasted demand and in line with the recommendations of the optimal ordering policy thus optimizing the inventory level and reducing cost in the short and long term.

CONTENTS

ILLUSTRATIONS

TABLES

CHAPTER I

INTRODUCTION AND MOTIVATION

This thesis represents a production planning and inventory management case study performed for the Lebanese tool steel company: FOZ Trading.

A. Motivation

FOZ Trading Establishment is a Lebanese company that serves the mechanical and steel industries, covering almost all their needs regarding equipment, machinery and consumables in steel and welding.

The choice to work with FOZ Trading came to be after careful consideration and market research which led to the realization that the steel industry market is increasing in value and capacity. Upon further research, it turned out that the steel industry is growing at a significant pace that would make the region one of the most prominent figures in this industry when compared to other regions (East Asia, Latin America, etc.…). "Steel demand in the region is growing at a significant pace thanks to the various ambitious government plans related to the construction and infrastructure sector" (Zakharova, 2016).

The various projects in the MENA region - especially those associated with the current visions of some Gulf countries: UAE 2021 vision, SA 2030 vision, Qatar 2030 vision, Kuwait 2035 vision, and Oman 2040 vision - will serve a great deal in reviving the steel sector and improve demand of steel in the region. “Apparent steel usage in the Middle East increased from 34 million tons in 2005 to 48.1 million tons in 2011, representing a compound annual growth rate of nearly 6 per cent” (Gavin, 2013).

This continuous growth and expansion of the steel industry in the MENA region proved to be the initial spark that prompted this case study to take place as a way to get FOZ Trading to compete locally and internationally in an ever-expanding market.

B. Introduction

Demand forecasting and optimal ordering policies are considered to be one of the most important aspects of operations research and production management. "An organization's ability to reduce the demand dispersion and consequently improve the forecasting performance will be a key to the effective production planning and inventory management system, Because of the benefits it can bring; many industries have paid great attention to it" (Chen et al., 2007).

This case study focuses on demand forecasting techniques that aim to generate reasonable forecasts to reduce demand uncertainty and improve inventory. The forecasting techniques will be applied to 18 different product categories[1] utilizing 5 years of previous data. Basic forecasting methods such as last period demand and best previous period demand, are used as benchmarks against traditional and alternate forecasting methods. The traditional forecasting methods include simple moving average (SMA), simple exponential smoothing (SES) and Box-Jenkins (ARMA and ARIMA). The alternate forecasting methods proposed are Croston, adjusted Croston (SBA and TSB), moving block bootstrapping (MBB) and circular block bootstrapping (CBB). Machine learning methods such as locally weighted linear regression (LOESS) and gaussian processes regression (GPR) are briefly tested as well to check whether they are suitable for our case study or not.

---

[1] The reason and process of choosing the 18 categories is discussed later on in the paper, refer to CHAPTER 4, Section A. Aggregation of Stock Keeping Units (SKUs).

The steel industry in Lebanon is directly affected by the number of active industrial projects going on in the country. So, fluctuations in demand are very normal. That's why past data exhibited points of zero and others of very high demand, otherwise known as intermittent demand. This discontinuity makes predicting the demand harder with conventional forecasting techniques, hence the inclusion of Croston's and bootstrapping methods. These methods are implemented and tested to check whether they will produce more reasonable forecasts than the traditional forecasting methods. The forecasts will be evaluated and assessed based on mean absolute error (MAE), and root mean square error (RMSE). The smaller the RMSE and the MAE are, the smaller the error.

C. Objectives

- Analysis of the sales data to ensure reliable inferences can be generated from the data's underlying process.
- Reduction of demand dispersion to improve forecasting accuracy and performance.
- Generation of reasonable demand forecasts using various industry proven methods.
- Application of a suitable optimal ordering policy to optimize inventory and minimize cost.

CHAPTER II

LITERATURE REVIEW

Time series analysis predicts future demand from past internal data where historical data are analyzed to determine temporal patterns. Usually, the analysis of times series is accomplished by decomposing the data into five components: level (α), trend (b), seasonal variations (F), cyclical movements (C), and irregular random fluctuations (ε). "Level captures the scale of a time series (if only a level was present, the series would be constant with time). Trend identifies the rate of growth or decline of a series over time" (Silver et al., 2016). Seasonal variations correspond to periodic, repetitive and predictable patterns in the levels of a time series that occur at specific regular intervals less than a year, such as weekly, monthly, or quarterly. Cyclical variations consist of rises and falls in demand that are not of a fixed period. These fluctuations are usually due to economic conditions and are often related to the business cycle. "Irregular fluctuations are the residue that remain after the effects of the other four components are identified and removed from the time series" (Silver et al., 2016). Using these concepts, Silver et al. (2016) formulated an additive model of a time series:

Demand in period t = (level) + (Trend) + (Seasonal) + (Cyclic) + (Irregular)

$$X_t = \alpha + b_t + F_t + C_t + \varepsilon_t$$

The choice of the general type of underlying model to use for a particular item depends very much on cost consideration. An Analysis of historical data should suggest the general types of models that would be appropriate for a given time series (Silver et al., 2016).

Silver et al. (2016) discuss several types of forecasting methods for individual-item, short-term forecasts. The most widely used forecasting techniques are the simple moving average and the simple exponential Smoothing. Additionally, Tersine (1982) suggests two more methods which are the last period demand and the arithmetic average. These four forecasting methods are based on a level demand model with low variability and fluctuations around the mean. Therefore, they are inappropriate when the underlying demand pattern involves intermittent demand, significant trend or seasonal variation.

Intermittent demand, sometimes known as sporadic demand, occurs when a product or a stock keeping unit experiences several periods of high variability in demand or of zero demand. Intermittent demand (ID) is usually characterized by high variability in demand between successive periods. It usually occurs in aviation and manufacturing industries. The high variations in demand between periods along with the many zero values in ID time-series render most usual forecasting methods unreasonable and difficult to apply (Waller, 2015). Companies often use traditional forecasting techniques, such as simple exponential smoothing and moving averages. These techniques work best for normal, high-volume demand, however, when it comes to intermittent demand, it is recommended to stay away from these techniques. Simple exponential smoothing was the first forecasting method to be applied to intermittent demand. However, it was soon confirmed that SES performs poorly when it comes to intermittent demand. Operations Research experts reckon that "SES usually perform poorly in the case of intermittent demand due to the fact that the forecast exhibit an upward bias in the period directly after a non-zero demand" (Waller, 2015). They also believe that traditional forecasting methods fail in the case of intermittent demand because they try to identify recognizable patterns in the demand data, such as trend and seasonality. However, what sets intermittent demand apart is that its data don't exhibit such regular patterns, in fact, ID's only main characteristic is its multitude of zero values. Companies usually use traditional forecasting techniques without taking into consideration the effect of the numerous zero values found in the historical data. By doing that, they are not properly accounting for the timing and frequency of the zero demand, thus affecting the whole forecasting process and damaging the estimation accuracy of the items analyzed.

According to Chen et al. (2007), demand aggregation is one of the most efficient methods to reduce demand dispersion and improve forecast accuracy. Demand aggregation can take many forms and can be implemented in many different ways. One approach, is to combine periods into pre-existing time blocks. So instead of having the data in days or months, it is grouped into quarters or years, thereby reducing the number of zeros in the series and minimizing the intermittency within the data. "This aggregation approach is often referred to, in academic literature, as Temporal Aggregation" (Nikolopoulos et al., 2011).

In 2009, Athanasopoulos et al. investigated the advantages of temporal aggregation on fast and slow-moving demand data using 366 monthly series tested in the M3 competition. They aggregated the monthly time series into quarters, and then aggregated them further into years. Upon comparing the forecasts generated from the aggregated series (quarterly and yearly) with the original unaggregated series, they found that the aggregated series forecasts were more accurate than the forecasts produced at the monthly level. "Athanasopoulos et al. study provided considerable empirical evidence in support of temporal aggregation" (Tabar et al., 2013).

ARIMA and ARMA models, also called Box-Jenkins models, are general time series forecasting techniques that usually provide reasonable and accurate forecasts. An ARIMA model is a generalization of an ARMA model. ARIMA stands for Autoregressive Integrated Moving Average. The ARIMA model differentiates (or integrates) the data first in order to make it stationary, while the ARMA model requires it to be beforehand. Both methods model future points based on past points and residuals. The future points are modeled using the autoregressive function on previous points and using the moving average function on previous residuals. Box-Jenkins methods don’t usually perform well with intermittent demand since they are based on moving average and regression. According to Waller (2015), they tend to allow negative integer values when regressing on previous points which is inappropriate for intermittent data.

“Croston (1972) was the first to suggest that traditional forecasting methods such as moving average (MA) and simple exponential smoothing (SES) may be inappropriate for slow-moving items” (Teunter and Duncan, 2009). According to Teunter and Duncan (2009), Croston demonstrated that using traditional forecasting methods can lead to sub-optimal stocking decisions. In order to overcome this problem, Croston implemented an alternative forecasting procedure that “separately updates the demand interval and the demand size (exponentially, and with the same smoothing constant for both), and only does so in periods with positive demand” (Teunter and Duncan, 2009). According to Croston, the demand follows a Bernoulli’s process while the inter-demand intervals follow a geometric distribution. “Croston also accepted that (non–zero) demand sizes are normally distributed and independent from inter–demand intervals” (Doszyń, 2019).

The main limitation to Croston's method was biasness. In 2001, Syntetos and Boylan argued that the original method is biased and suggested later on in 2005 a modification to correct it, which was later called the Syntetos-Boylan Approximation (SBA). Syntetos and Boylan suggest multiplying the Croston forecast by $(1-\frac{\alpha}{2})$, in order to get rid of biasness, where α is the smoothing constant. The SBA approximation corrects Croston's biasness issue in the case of intermittent demand, however it introduces new bias to non-intermittent demand due to the fact that now non-intermittent forecasts are multiplied by $(1-\frac{\alpha}{2})$.

Croston and its SBA variant also face criticism due to the fact that they do not take into consideration product obsolescence (i.e. a situation where an item is not in demand anymore). So, when a product is no longer demanded they continue to forecast a fixed non-zero demand forever without ever gradually decreasing it to zero (the inter-demand intervals are not updated unless non-zero demand take place).

Teunter, Syntetos, and Babai suggested a Croston variant, named Croston TSB, that is designed to deal with product obsolescence (Teunter et al., 2011). "Croston TSB estimates the probability of non-zero demand (rather than interval size), and in that the estimates are updated every period, rather than just when demand occurs" (Waller, 2015).

Bootstrapping is a statistical technique that involves random sampling with replacement. It is only valid for IID observations, and since this assumption is not met in time series analysis, the bootstrapping technique has to be adjusted to work with time series. Carlstein (1986) was the first to discuss the notion of bootstrapping blocks of observations rather than individual observations. However, it was Künsch (1989) who introduced the concept of the Moving Block Bootstrapping (MBB) and Politis and

Romano (1992) who introduced the concept of the Circular Block Bootstrapping (CBB) which are applicable to stationary time series data. “Bootstrapping’s main advantage is that (the mean and variance) of the lead time demand distribution is forecasted directly by repeated sampling from realized demands” (Teunter and Duncan, 2009). Many scholars used bootstrapping and adjusted it in order to yield reasonable and accurate classification and forecasting results, that’s why there are many variants of the bootstrapping method.

CHAPTER III

METHODS

A. Basic Forecasting Methods

1. Last Period Demand (Naïve Forecast)

The last period demand method states that the next period has the same level of demand that occurred in the previous period, so the forecasted values lag behind actual demand by one period.

Mathematically,

$\hat{Y}_t$ = *estimate of mean sales in period t; i.e. the forecast made in period t-1 for period t.*

Y_{t-1} = *actual sales in period t-1.*

$$\hat{Y}_t = Y_{t-1}$$

$$\hat{Y}_{t+1} = Y_t$$

2. Best Previous Period Demand

The best previous period demand method states that the forecasted demand (next period demand) should be similar to the demand of a previous period (a specific previous period, not necessarily the one directly before it).

Mathematically,

$\hat{Y}_t$ = *estimate of mean sales in period t; i.e. the forecast made in period t-1 for period t.*

n = *number of periods.*

$\boldsymbol{k}$ = *index of the period k,*

$\boldsymbol{Y_k}$ = *actual sales in period k; where k =1, 2, …, n.*[2]

$$\hat{Y}_t = Y_k$$

$$\hat{Y}_{t+1} = Y_{k+1}$$

In order to check which previous period should be chosen as a next period demand, RMSE was evaluated for all previous periods, and the one with the least RMSE was chosen as a next period demand.

B. Traditional Forecasting Methods

1. Simple Moving Average (SMA)[3]

The simple moving average method generates the next period's forecast by averaging the actual demand for the last N time periods. This technique gives equal weight to the specified periods included in the average.

Mathematically,

$\hat{Y}_t$ = *estimate of mean sales in period t; i.e. the forecast made in period t-1 for period t.*

$\boldsymbol{N}$ = *the order of the simple moving average; i.e. the window.*

$\boldsymbol{D_i}$ = *actual sales in period i.*

$$\hat{Y}_t = \left(\frac{1}{N}\right) \sum_{i=t-N}^{t-1} D_i$$

$$= \left(\frac{1}{N}\right) (D_{t-1} + D_{t-2} + \cdots + D_{t-N})$$

---

[2] Only one "k" is chosen. "k" is chosen by evaluating $\hat{Y}_t = Y_k$ for all k =*1, 2, …, n*. And then taking the "k" of the Ŷt with lowest RMSE score.

[3] Adapted from Nahmias and Olsen (2015).

In other words, this formula states that the mean of the ***N*** most recent observations is used to forecast for the next period. For the sake of our case study, two simple moving average models were considered with window, ***N*** = 3 and 4 periods.

2. *Simple Exponential Smoothing (SES)*[4]

Exponential smoothing is a special kind of moving average that does not require the keeping of a long historical record and where past data are not given equal weight. The weight given to past data decreases geometrically with the increasing age of the data. Therefore, the main advantage of exponential smoothing is that forecasted demand depends more on recent data than on old data.

Mathematically,

$\hat{Y}_t$ = *estimate of mean sales in period t; i.e. the forecast made in period t-1 for period t.*

Y_{t-1} = *actual sales in period t-1.*

$\hat{Y}_{t-1}$ = *estimate of mean sales in period t-1; i.e. the forecast made in period t-2 for period t-1.*

α = *smoothing constant between 0 and 1*

$$\hat{Y}_t = \alpha Y_{t-1} + (1 - \alpha)\hat{Y}_{t-1}$$

For the sake of our case study, only a few smoothing constants were taken to avoid too much unnecessary complexity. In total five SES models were considered.

[4] Adapted from Silver et al. (2016).

Four SES models with smoothing constant α between 0.1 and 0.4 with 0.1 increments, and one special SES model where the smoothing constant α was optimized by maximizing the log-likelihood.

Hyndman and Athanasopoulos (2018) state that maximizing the likelihood, in the case of an additive error model, will give the same results as those generated from minimizing the sum of squared errors. However, in the case of multiplicative error models, maximizing the likelihood will give different results than minimizing sum of squared errors (Hyndman and Athanasopoulos, 2018). That would be worth exploring if Holt-Winters method with multiplicative seasonality was applied. However, double exponential smoothing (Holt's Linear Trend Model) and triple exponential smoothing (Holt-Winters exponential smoothing with multiplicative seasonality) were not used in this case study. Holt's linear trend model is used when the data analyzed show signs of an existing trend, and Holt-Winters exponential smoothing model is used when the data show signs of an existing trend and seasonality. FOZ Trading data showed heavy signs of intermittent demand with little to no signs of trend or seasonality which prompted us to exclude them from the forecasting methods used.

3. Box-Jenkins (ARMA and ARIMA)[5]

The ARIMA model is a generalization of the ARMA model. ARIMA stands for Autoregressive Integrated Moving Average. Both are theoretical, model-based approaches that aim to generate reasonable forecasts by modeling future points based on past points and residuals. The main difference between the two is that ARIMA

[5] Adapted from Brockwell and Davis (2016) and Kang (2017).

differentiates (or integrates) the data first in order to make it stationary, while the ARMA requires it to be beforehand.

AR (Autoregression): A model that uses the dependent relationship between an observation and some number of lagged observations. $\boldsymbol{\varphi_i}$ are parameters of the model**,** $\boldsymbol{p}$ is a parameter of how many lagged observations to be taken in, $\boldsymbol{c}$ is a constant, and $\boldsymbol{\varepsilon_t}$ is random noise.

$$\boldsymbol{X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t}$$

I (Integrated): A model that uses the differencing of raw observations (e.g. subtracting an observation from the previous time step). Differencing in statistics is a transformation applied to time-series data in order to make it stationary. This allows the properties do not depend on the time of observation, eliminating trend and seasonality and stabilizing the mean of the time series.

MA (Moving Average): A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations. $\boldsymbol{\theta_i}$ are parameters of the model**,** $\boldsymbol{q}$ is a parameter of how many lagged observations to be taken in, $\boldsymbol{\mu}$ is a constant, and $\boldsymbol{\varepsilon_t}$ is random noise. Contrary to the AR model, the finite MA model is always stationary.

$$\boldsymbol{X_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t}$$

When combined, the full ARIMA ***(p,d,q)*** model is as follows:

$$\boldsymbol{X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}}$$

In other words, we model the points in our series as being dependent on the previous ***p*** points (auto-regressive) and on the previous ***q*** residuals (moving-average).

Parameters of the ARIMA ***(p,d,q)*** model:

- **p (lag order):** number of lag observations included in the model.
- **d (degree of differencing):** number of times that the raw observations are differenced.
- **q (order of moving average):** size of the moving average window.

For the purpose of this case study, a grid search of over 294 ARIMA and ARMA possible combinations were exhausted (147 ARIMA, and 147 ARMA), and the one with least RMSE was used for forecasting.

Seasonal Autoregressive Integrated Moving Average (SARIMA) was not used in this case study to forecast next period demand. SARIMA is used when the data analyzed show signs of an existing seasonality. As mentioned before, the data being analyzed showed little to no signs of seasonality, only signs of intermittent demand which prompted us to exclude SARIMA from the forecasting methods used.

C. Alternate Forecasting Methods

1. Croston[6]

Croston (1972) was the first to propose an intermittent demand-specific method. He suggested estimating demand probability (via interval size) and demand size separately in order to make the forecast more intuitive and accurate.

[6] Adapted from Silver et al. (2016) and Waller (2015)

Let $\boldsymbol{Z_t}$ be the estimate of mean non-zero demand size for time t, $\boldsymbol{V_t}$ the estimate of mean interval size between non-zero demands, and $\boldsymbol{X_t}$ the actual demand observed at time t. $\boldsymbol{q}$ is the current number of consecutive zero-demand periods, $\boldsymbol{\alpha}$ is the smoothing constant between 0 and 1 and $\boldsymbol{Y_t}$ will denote an estimate of mean demand size (i.e. taking zero demands into the calculation). Then:

$$\textbf{If } \boldsymbol{X_t \neq 0} \textbf{ then } \begin{cases} \boldsymbol{Z_{t+1} = \alpha X_t + (1-\alpha) Z_t} \\ \boldsymbol{V_{t+1} = \alpha q + (1-\alpha) V_t} \\ \boldsymbol{Y_{t+1} = \frac{Z_{t+1}}{V_{t+1}}} \end{cases}$$

$$\textbf{If } \boldsymbol{X_t = 0} \textbf{ then } \begin{cases} \boldsymbol{Z_{t+1} = Z_t} \\ \boldsymbol{V_{t+1} = V_t} \\ \boldsymbol{Y_{t+1} = Y_t} \end{cases}$$

For the sake of our case study, five Croston models were considered. Four with smoothing constant $\boldsymbol{\alpha}$ between 0.1 and 0.4 with 0.1 increments, and one special Croston model where the smoothing constant $\boldsymbol{\alpha}$ was optimized by maximizing the log-likelihood.

2. Adjusted Croston[7]

The most important limitation to Croston's method is biasness. Syntetos and Boylan (2005) showed that Croston's Method is biased since: $\boldsymbol{E[\overline{X_t}] = E\left[\frac{Z_t}{V_t}\right] \neq E[Z_t]\frac{1}{E[V_t]}}$.

[7] Adapted from Silver et al. (2016) and Waller (2015)

Syntetos and Boylan (2005) proposed an adjustment to Croston's forecast $\boldsymbol{Y_t}$ that would eliminate the biasness in the forecast, which was later known as Croston-SBA. They stated that if $\boldsymbol{Y_t}$ is multiplied by a factor of $(\mathbf{1}-\frac{\alpha}{2})$ the new forecast will be approximately unbiased, since:

$$\boldsymbol{E}\left[\left(\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}}\right)\left(\frac{\boldsymbol{Z_t}}{\boldsymbol{V_t}}\right)\right] = \left(\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}}\right)\boldsymbol{E}\left[\frac{\boldsymbol{Z_t}}{\boldsymbol{V_t}}\right]$$

$$= \left(\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}}\right)\left(\frac{\boldsymbol{\mu}}{\boldsymbol{p}}+\frac{\mathbf{1}}{\mathbf{2}}\frac{\boldsymbol{\partial}^2\left(\frac{\boldsymbol{\mu}}{\boldsymbol{p}}\right)}{\boldsymbol{\partial p^2}}\boldsymbol{Var}(\boldsymbol{p})\right)$$

$$= \left(\frac{\mathbf{2}-\boldsymbol{\alpha}}{\mathbf{2}}\right)\left(\frac{\boldsymbol{\mu}}{\boldsymbol{p}}+\frac{\boldsymbol{\alpha}}{\mathbf{2}-\boldsymbol{\alpha}}\boldsymbol{\mu}\frac{\boldsymbol{p}-\mathbf{1}}{\boldsymbol{p^2}}\right)$$

$$= \left(\frac{\boldsymbol{\mu}}{\boldsymbol{p}}\right)\left(\frac{\mathbf{2}-\boldsymbol{\alpha}}{\mathbf{2}}+\frac{\boldsymbol{\alpha}}{\mathbf{2}}\frac{\boldsymbol{p}-\mathbf{1}}{\boldsymbol{p}}\right) \approx \frac{\boldsymbol{\mu}}{\boldsymbol{p}}$$

That proved to be true for the case of intermittent demand, however, when applying the SBA approximation to non-intermittent demand, new bias was noticed due to the fact that now the non-intermittent forecasts are multiplied by $(\mathbf{1}-\frac{\alpha}{2})$. Due to the fact that our data includes both intermittent demand and non-intermittent demand, we decided to not use Croston with SBA approximation.

Teunter et al. proposed a Croston variant that dealt with biasness and obsolescence. Their variant, called Croston TSB, estimates the probability of non-zero demand instead of interval size and the estimates are updated every period instead of updating just when demand occurs as in Croston and SBA's case.

Let $\mathbf{D_t}$ be an indicator of a non-zero demand at time t and $\mathbf{P_t}$ the demand estimates for time period t. The smoothing parameter $\boldsymbol{\alpha}$ is used for smoothing demand size, while the smoothing parameter $\boldsymbol{\beta}$ is used for smoothing demand probability. The Teunter, Syntetos and Babai (TSB) method is as follows:

$$\textbf{If } D_t = 1 \textbf{ then } \begin{cases} P_{t+1} = \beta + (1-\beta)P_t \\ Z_{t+1} = \alpha X_t + (1-\alpha)Z_t \\ Y_{t+1} = P_{t+1}\ Z_{t+1} \end{cases}$$

$$\textbf{If } D_t = 0 \textbf{ then } \begin{cases} P_{t+1} = (1-\beta)P_t \\ Z_{t+1} = Z_t \\ Y_{t+1} = P_{t+1}\ Z_{t+1} \end{cases}$$

For the sake of our case study, five Croston TSB models were considered. Four with smoothing constant $\boldsymbol{\alpha}$ between 0.1 and 0.4 with 0.1 increments, and one special Croston TSB model where the smoothing constant $\boldsymbol{\alpha}$ was optimized by maximizing the log-likelihood. $\boldsymbol{\beta}$ is chosen to be 0.15. The decision - for choosing $\boldsymbol{\beta}$ = 0.15 - was taken after testing values of $\boldsymbol{\beta}$ ranging from 0.01 to 0.3. $\boldsymbol{\beta}$ = 0.15 showed the most promise for our case study. However, it should be noted that after testing the whole range of values between 0.01 and 0.3, $\boldsymbol{\beta}$ = 0.08, 0.15, and 0.23 showed the most promise, however, the difference they made to the forecasts were almost negligible. That's why $\boldsymbol{\beta}$ was chosen to be equal to 0.15 only.

3. Bootstrap[8]

a. <u>Moving Block Bootstrap (MBB)</u>

The Moving Block Bootstrap (MBB) models the data into overlapping blocks of observations by dividing the data of $\boldsymbol{n}$ observations into blocks of length $\boldsymbol{l}$**.** Then, $\boldsymbol{b}$

[8] Adapted from Künsch (1989) and Politis and Romano (1992)

of these blocks are selected (repetition allowed) by resampling with replacement all the possible blocks. By doing so, the total amount of blocks will be $\boldsymbol{n-l-1}$.

b. <u>Circular Block Bootstrap (CBB)</u>

The Circular Block Bootstrap (CBB) is very similar to the Moving Block Bootstrap. However, in CBB we "wrap the time series around a circle", so that it goes $\boldsymbol{X_1, X_2, ..., X_{n-1}, X_n, X_1, X_2, \ldots}$ **etc.** We then sample the $\boldsymbol{b}$ blocks of length $\boldsymbol{l}$. The this gives us a total amount of blocks equal to $\boldsymbol{n}$.

Both methods, MBB and CBB, are usually implemented on top of a previous method in order to forecast demand over the lead time. For the purpose of our case study, we generated 1000 MBB samples and 1000 CBB samples, and implemented Croston TSB and the simple moving average on each of those samples. Finally, we calculated the average of the generated predictions to get our point estimate.

D. Machine Learning Methods

1. Locally Weighted Linear Regression (LOWESS)

LOWESS, also known as LOESS, is a non-parametric regression method that combines multiple regression models in a k-nearest-neighbor-based meta-model.

Implementing LOWESS on the FOZ Trading historical data gave very good results especially when interpolating (using them on a given set of data) however when extrapolating (trying to predict values for a new unseen dataset) they failed to make a reasonable extrapolation. That is due to the fact that LOWESS is based on local regression so when trying to predict a new value it will only take the nearby values into consideration thus making our projection extremely sensitive to the neighboring data. That’s why we opted not to use the LOWESS method in our case study.

2. Gaussian Processes Regression (GPR)

Gaussian process regression (GPR) is a nonparametric, Bayesian approach towards regression problems that can be utilized in exploration and exploitation scenarios.

When using GPR to extrapolate for new unseen data, it failed to make a reasonable extrapolation as well. The GPR almost always predicted the next value as the mean of the training data. That’s why it failed in capturing inference from previous data. "Anywhere away from the training points, a GP regression model will drop off to the mean”. That’s why we opted not to use the GPR method in our case study.

3. Artificial and Recurrent Neural Networks (ANNs and RNNs)

After careful consideration it turned out that Artificial Neural Networks (ANNs) and Recurrent Neural Networks (RNNs) won’t be suitable to implement for our case study because of the low amount of data points that we have. If we disaggregate the data into daily data instead of quarterly data, the neural networks won’t produce very reliable forecasts as well since most of the data will be zeros. That’s why we opted not to use neural networks in our case study.

E. Evaluation Parameters

To evaluate the accuracy of the forecasts generated by the different methods, two evaluating parameters where considered. The Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE).

Mean Absolute Error (MAE): "MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. It's the average over the test sample of the absolute differences between forecast and corresponding actual observation where all individual differences have equal weight" (Wesner, 2016).

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right|$$

Root Mean Squared Error (RMSE): "RMSE is a quadratic scoring rule which measures the average magnitude of the error. It's the square root of the average of squared differences between prediction and actual observation" (Wesner, 2016).

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}$$

The main advantage of using RMSE over MAE is that it penalizes large errors. RMSE gives a relatively high weight to large errors due to the fact that the errors are squared before they are averaged. Thus, making RMSE more useful when large errors are undesirable.

CHAPTER IV

DATA EXPLORATION AND ANALYSIS

A. Aggregation of Stock Keeping Units (SKUs)

FOZ Trading currently sells more than 500 Stock Keeping Units (SKUs), however, upon analyzing the units we noticed that these SKUs are not independent. A lot of the SKUs are intertwined and can be replaced by one another. For example, if a customer requested an order of SKU 12, but the warehouse did not have SKU 12 in stock, the management will contact the customer and offer to replace SKU 12 with SKU 13 since 12 is not in stock. Management explained that almost always the customer doesn't mind since the differences between SKU 12 and SKU 13 are negligible, and most of the time the customer prefers to get his product right away rather than wait for the company to restock from that product. The negligible differences are usually in diameter, e.g. SKU 12 might have a diameter 12 mm, while SKU 13's diameter is 13 mm. The differences are usually in small millimeter increments which explains why customers usually don't mind when asked to replace one SKU with another.

This problem meant that previous data might infer unreliable information when analyzing previous sales. The previous sales of an SKU do not represent anymore the portion of the true demand that is usually captured by the previous sales, since the sales data of an SKU now represents the previous demand of the SKU along with a portion of the previous demand of similar SKUs. In order to avoid unrealistic results when forecasting demand for future periods based on past data, we decided to aggregate the

SKUs that are interchangeable, i.e. the ones that share minor differences. By doing so we captured the true demand portion present in the past data.

Upon further consultations with FOZ Trading Management we found that the best way to group the SKU is by aggregating the individual similar SKUs into one unit being the parent category to which they belong. This approach is both sensible and practical for the following reasons. First, every item naturally already belongs to its parent category as previously explained and has the same applications. Second, the aggregation approach makes calculating the forecasted demand, and the ordering policies computationally efficient. Instead of computing forecasts and policies for 500 items, they are computed for the parent 18 categories and then disaggregated using industry and market knowledge.

Category Number	Category Name	Category Number	Category Name
1	E 200/230 Round	**10**	M 238 Flat
2	K 100 Round	**11**	M 300/303 Round
3	K 100 Flat	**12**	M 300/303 Flat
4	K 110 Round	**13**	M 310 Round
5	K 110 Flat	**14**	M 310 Flat
6	K 510 Round	**15**	V 320 Round
7	K 720 Round	**16**	V 320 Flat
8	K 720 Flat	**17**	W 302 Round
9	M 238 Round	**18**	W 302 Flat

Table 1: The 500+ SKUs Aggregated into The Parent 18 Categories

B. Data Exploration

The analyzed data are real sales data provided to us by FOZ Trading. The data represents the sales quantities in Kilograms of the 18 categories over a span of 5 years, from 2013 till 2017.

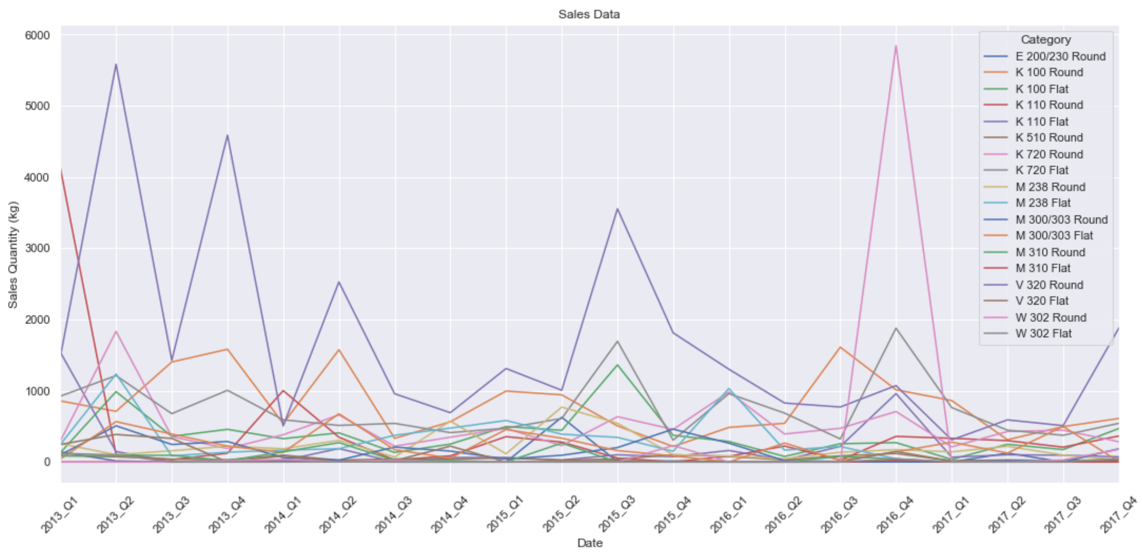

Figure 1: Sales Data of the 18 categories from 2013-2017

Before this paper, the company used the naïve forecast (last period demand) to forecast future sales in order to control inventory and to manage customer service level. To obtain acceptable forecasts it is important to know something about the nature of the data. Upon analyzing the data, we found out that the sales data of all the categories show signs of intermittent demand. Hence our decision to include specific methods used usually for intermittent demand. When it comes to time series forecasting, Syntetos and Boylan suggest classifying the products being forecasted (Syntetos et al. 2005). If the data did not include intermittent demand, we would have used the ABC categorization suggested by Teunter et al. (2010) and Reid (1987). However, since intermittent demand is crucial to the process of deciding which method to use for forecasting, it is recommended to categorize the data based on the coefficient of variation and the averaged inter-demand intervals (Syntetos et al. 2005).

The histograms in Error! Reference source not found. give us a clear idea r egarding the spread of the data and how it is distributed on the 18 categories. The histograms of the 18 categories seem to be generally skewed to the right which implies the mean is greater than the median for most categories. Usually, when the mean is greater than the median, it means that the data has major outliers in the high end of the distribution.

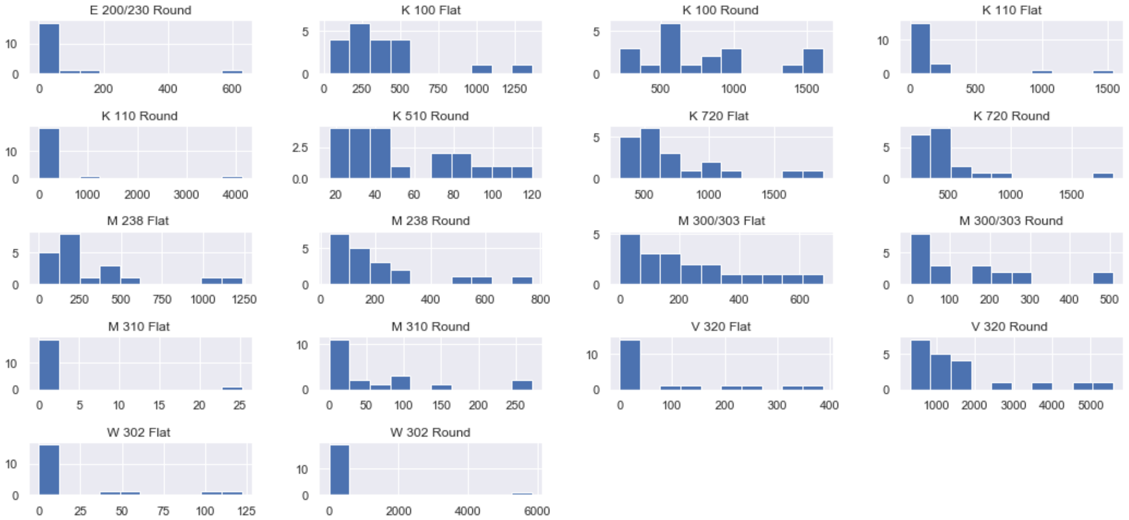

Figure 2: Histogram Representations of the 18 Categories

To confirm this hypothesis, we analyzed the data further using box and whisker plots. Upon examining the box and whisker plots in Error! Reference source not found., w e noticed very high and extreme coefficient of variations of some categories due to the effect of the outliers. The outliers were significantly raising the forecasting levels of the categories.

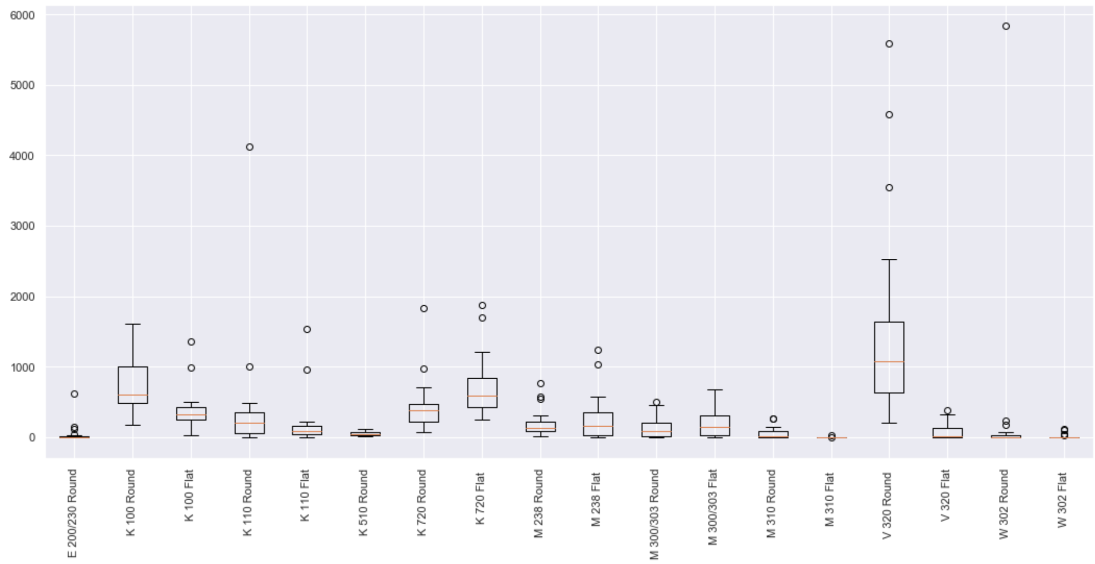

Figure 3: Box and Whisker plots of the 18 Categories

In such cases, Doszyn (2019) suggests using the sales frequency instead of the inter-demand intervals. The sales frequency is defined as the share of periods with zero or near zero sales in all periods. The advantage of using sales frequency in this case lies in the fact that sales frequency can be updated in every period even when the sale is zero (Doszyn, 2019). The sales frequency is the fraction of quarters[9] with non-zero sales. So, when the sales frequency is 1, it means that the category analyzed does not show signs of intermittent demand. If the sales frequency is 0.75, it means that the category is sold 75% of the time (the category shows little or no sales 25% of the time). If the sales frequency is 0.25, it means that the category is sold 25% of the time, and so on. We decided to classify our categories into 4 classes according to the sales frequency they exhibit:

Class SF-25: Categories with sales frequency between 0% and 25%
Class SF-50: Categories with sales frequency between 25% and 50%
Class SF-75: Categories with sales frequency between 50% and 75%
Class SF-100: Categories with sales frequency between 75% and 100%

[9] We aggregated our monthly data into quarters to reduce the intermittency of the demand data. The aggregation process is discussed later on in the paper, refer to "Temporal Aggregation" Section.

Category	Sales Frequency	Sales Frequency Interval	Class Label
M 310 Flat	0.04	0-0.25	SF-25
E 200/230 Round	0.22	0-0.25	SF-25
W 302 Flat	0.22	0-0.25	SF-25
W 302 Round	0.35	0.25-0.5	SF-50
V 320 Flat	0.48	0.25-0.5	SF-50
M 310 Round	0.48	0.25-0.5	SF-50
M 300/303 Round	0.7	0.5-0.75	SF-75
M 238 Flat	0.74	0.5-0.75	SF-75
M 300/303 Flat	0.83	0.75-1.0	SF-100
M 238 Round	0.91	0.75-1.0	SF-100
K 110 Flat	0.91	0.75-1.0	SF-100
K 510 Round	0.91	0.75-1.0	SF-100
K 110 Round	0.96	0.75-1.0	SF-100
V 320 Round	1	0.75-1.0	SF-100
K 100 Round	1	0.75-1.0	SF-100
K 720 Flat	1	0.75-1.0	SF-100
K 720 Round	1	0.75-1.0	SF-100
K 100 Flat	1	0.75-1.0	SF-100

Table 2: Sales Frequencies and Classes

C. Temporal Aggregation

As apparent from the research papers and studies presented in the literature review, temporal aggregation can help reduce intermittency in demand data and limit demand dispersion. That's why we opted to implement temporal aggregation on all the time series present in the data. By doing so we reduced the intermittency of the demand by more than 50%; now only eight out of eighteen categories show signs of intermittent demand (Sales Frequency <75%). This will allow traditional forecasting methods, such as ARIMA and SES, to work better on our data, since the number of zeros was greatly reduced.

D. Stationarity Check

Some of the methods tested requires the data to be stationary before it is applied such as ARMA and ARIMA. In order for the data to be considered stationary the statistical properties of the data such as the mean, variance and autocorrelation should be constant over time. Usually, real data are almost never stationary. So, in order to figure out if our data is stationary, we tested our data using three methods.

1. *The Graphical Method (Using Rolling Statistics)*

We plotted the moving average and moving standard deviation of all the time series to get a visual representation of the data's rolling statistics. If the moving average and the moving standard deviation vary with time, it means the data is not stationary.

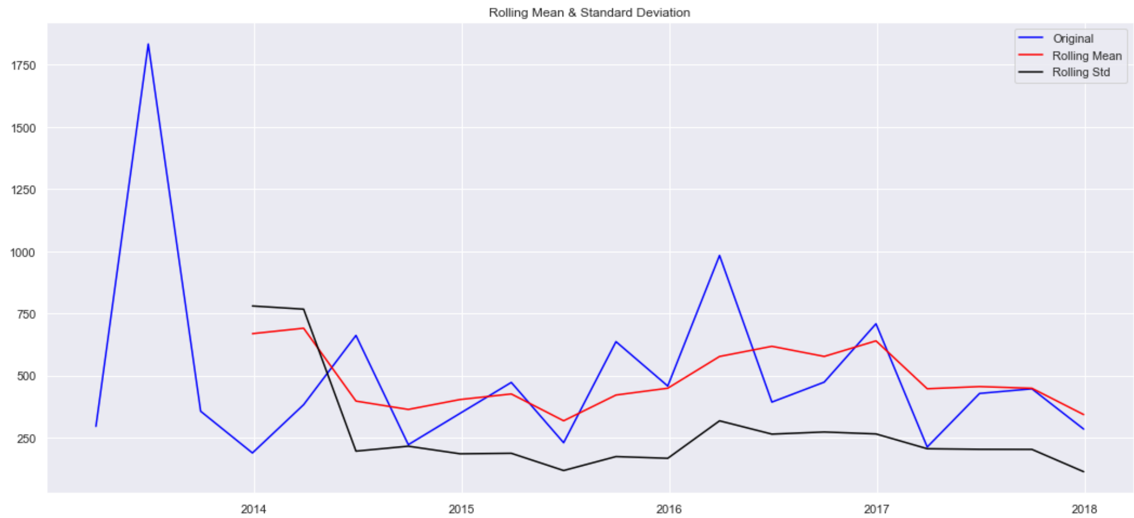

Figure 4: Rolling Statistics (Moving Average and Moving Standard Deviation) of K 720 Round

2. *The Augmented Dickey-Fuller Test (ADF)*

The Augmented Dickey-Fuller test is a type of statistical unit root test that determines how strongly a time series is defined by its intercept and time trend. It is based on the regression of the observed variable on its multiple lag values (Xiao and Phillips, 1998).

In ADF, the null hypothesis is that the Time Series is **non-stationary.**

- ***Null Hypothesis (H_0):*** *If failed to be rejected, it suggests the time series has a unit root, meaning it is non-stationary. It has some time dependent structure.*
- ***Alternate Hypothesis (H_1):*** *The null hypothesis is rejected; it suggests the time series does not have a unit root, meaning it is stationary. It does not have time-dependent structure.*

We interpret this result using the p-value from the test. A p-value below a threshold (such as 5% or 1%) suggests we reject the null hypothesis.

- *p-value > 0.05: Fail to reject the null hypothesis (H_0), the data has a unit root and is non-stationary.*
- *p-value <= 0.05: Reject the null hypothesis (H_0), the data does not have a unit root and is stationary.*

```
Results of Dickey-Fuller Test:
ADF Statistic: -5.100940
p-value: 0.000014
Critical Values:
 1%: -3.833
 5%: -3.031
 10%: -2.656
```

Figure 5: Results of the ADF Test Showing that the K 720 Round Time Series is Stationary

3. The Kwiatkowski-Phillips-Schmidt-Shin Test (KPSS)

The Kwiatkowsku-Phillips-Schmidt-Shin test also checks the stationarity of the time series. Its main difference from the ADF test is that the null and alternate hypothesis of the KPSS is opposite to that of the ADF. The KPSS can be used to test for trend and level stationarity. “Inference from this test is complementary to that derived from those based on the Dickey-Fuller distribution. The KPSS test is often used in conjunction with those tests to investigate the possibility that a series is fractionally integrated” (Baum, 2018).

In KPSS, the null hypothesis is that the Time Series is **stationary.**

- ***Null Hypothesis (H_0):*** *If failed to be rejected, it suggests the time series is stationary.*
- ***Alternate Hypothesis (H_1):*** *The null hypothesis is rejected; it suggests the time series is not stationary.*

We interpret this result using the p-value from the test. A p-value below a threshold (such as 5% or 1%) suggests we reject the null hypothesis.

- *p-value > 0.05: Fail to reject the null hypothesis (H_0), the data is stationary.*
- *p-value <= 0.05: Reject the null hypothesis (H_0), the data is non-stationary.*

```
Results of KPSS Test:
Test Statistic            0.310
p-value                   0.081
Lags Used                 9.000
Critical Value (10%)      0.347
Critical Value (5%)       0.463
Critical Value (2.5%)     0.574
Critical Value (1%)       0.739
```

Figure 6: Results of the KPSS Test Showing that the K 720 Round Time Series is Stationary

Upon analyzing the 18 categories using the tests highlighted above, it turned out that 16 out of 18 categories are non-stationary. However, some methods, such as ARMA and ARIMA, require stationarity before applying them. Dickey and Pantula

(2002) suggests using differencing or log transformation to make the series stationary. It should be mentioned that other methods such as STL, SEATS, Box-Cox Transformation or even Time Dummies can be used to make the series stationary. However, for our case study differencing seems to do the trick. Differencing was applied to the non-stationary time series only for the methods that require them, and then an un-differencing process was done to the series after the methods were applied.

E. Walk Forward Optimization

Walk forward optimization is an out of sample method that is used for validation and testing. It is a special type of cross-validation that's widely popular in trading and finance.

In the walk forward method, 60% to 70% of the entire data are held back to train the model, usually they are the first portion of the data. The remaining 30% to 40% are called the out-of-sample data and are used for validation and testing (Kirkpatrick and Dahlquist, 2010).

We start the procedure by training the model on the in-sample data (training dataset), and computing the next period prediction. We record this prediction and then move the training set one step forward, while preserving the window size, to include the first entry of the tested dataset (the out-of-sample data). This results in dropping the oldest entry in the training dataset in favor of the new entry (Masters, 2013). A new model will be trained on the new training dataset (original training dataset, plus first testing entry, minus oldest training entry) and a new prediction will be generated for the next period. This process is repeated until we reach the end of the testing dataset. At the end, the predictions are compared to the true observations and evaluation parameters are computed to assess the model.

For the purpose of our case study, we adjusted the walk forward optimization - to suit the small size of our data - and implemented it in the following manner:

1. The first 60% of the data were held back to train the model (in-sample data).
2. The remaining 40% of the data were used as out-of-sample data.
3. The out-of-sample data was used to iterate and test the model.
4. For every iteration in the test dataset:
   - A model was trained.
   - A one-step prediction was generated.
   - The one-step prediction was stored for later evaluation.
   - The actual observation from the test dataset was added to the training dataset for the next iteration without dropping the old iteration.
5. The predictions made during the iteration of the test dataset were assessed and evaluation parameters were computed.

Given the small size of our data, the oldest entry in the training dataset was not dropped when iterating over the test dataset. This allowed us to re-train the model using all available data to generate the predictions.

F. Bias Correction

Upon evaluating the forecasting methods, some of them exhibited signs of bias. In order to combat the biasness, a review of the residual forecast error was done. “If the model is unbiased, then the mean residual (error) must be equal to zero” (St-Pierre, 2003). The distribution of residual errors was plotted to figure out patterns in the distribution and ways to bias-correct the prediction.

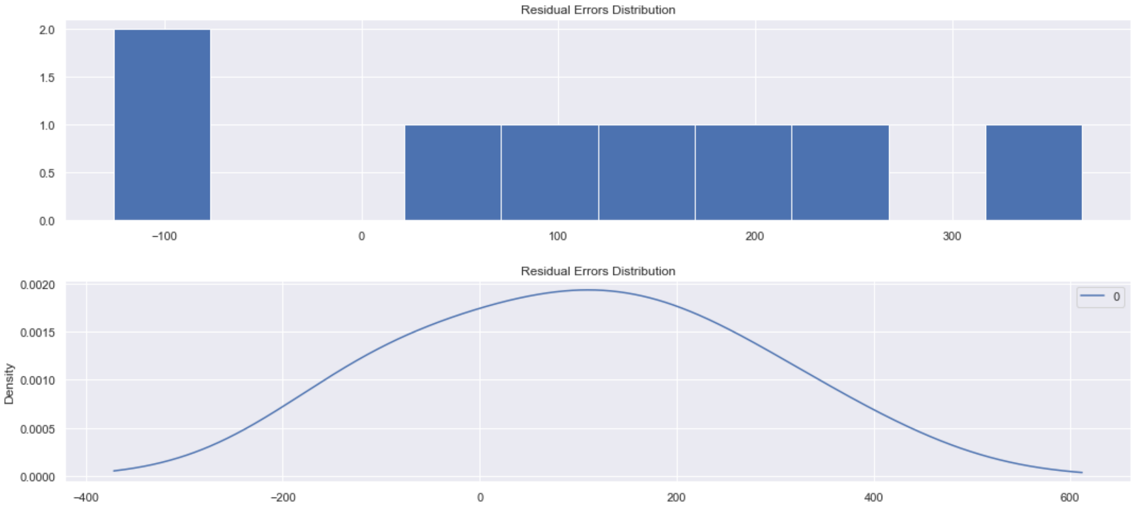

Figure 7: Residual Error Distribution of V 320 Round (Before Bias Correction)

```
count     8.000
mean    100.606
std     166.248
min    -125.738
25%      -3.226
50%      95.926
75%     215.044
max     365.961
```

Figure 8: Statistical Properties of V 320 Round's Residual Error Distribution (Before Bias Correction)

We used this information to Bias-correct predictions by adding the mean residual error to each biased forecast in each iteration of the walk forward optimization. By doing so the performance of the predictions improved slightly for all the categories.

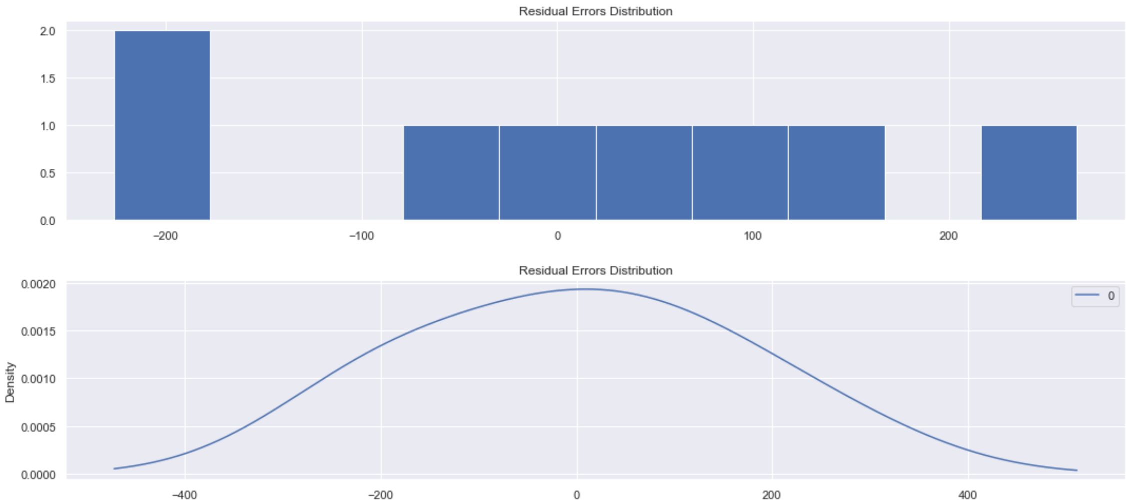

Figure 9: Residual Error Distribution of V 320 Round (After Bias Correction)

```
count   8.000e+00
mean    1.421e-14
std     1.662e+02
min    -2.263e+02
25%    -1.038e+02
50%    -4.680e+00
75%     1.144e+02
max     2.654e+02
```

Figure 10: Statistical Properties of V 320 Round's Residual Error Distribution (After Bias Correction)

The statistical properties of the residual error distribution after bias correction shows that the mean moved to a value very close to zero. And the density plots of the residual error shifted towards zero as well confirming that the model was indeed bias corrected.

G. Prediction Intervals

After calculating a point estimate of the next period demand, it is important to calculate a prediction interval to capture most of the uncertainty regarding the next period forecast. “A prediction interval gives an interval within which we expect $\mathbf{y_t}$ to lie with a specified probability” (Hyndman and Athanasopoulo, 2018).

Prediction intervals describes the uncertainty of a single point forecast, whereas a confidence interval describes the uncertainty of a model's parameter such as mean or standard deviation. Our aim is to assess the accuracy of our forecasted next period demand, that's why prediction intervals are preferred over confidence intervals for the purpose of our study.

Prediction intervals use the standard deviation of the residuals along with a "*t*-multiplier". The residual errors are assumed to follow a normal distribution. This assumption is backed by plotting the residual error distributions for all the methods used. The plots show normal distribution characteristics which confirms our assumption. Thus, instead of using a "*t*-multiplier" to compute the prediction interval, the "z-multiplier" is utilized in its place.

Below are examples of the residual error distribution for two of the categories included in our study.

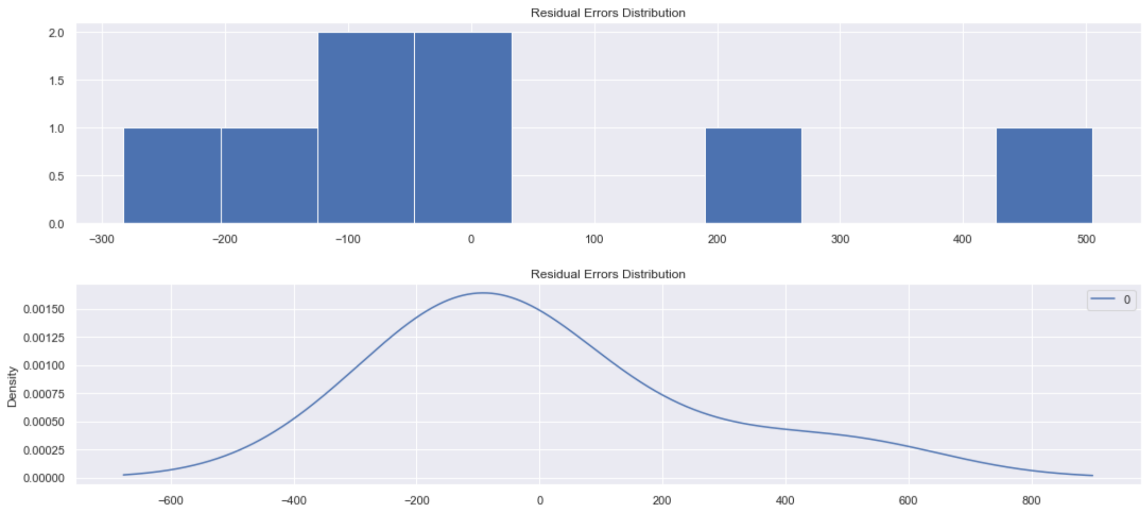

Figure 11: Residual Errors Distribution for K 720 Round

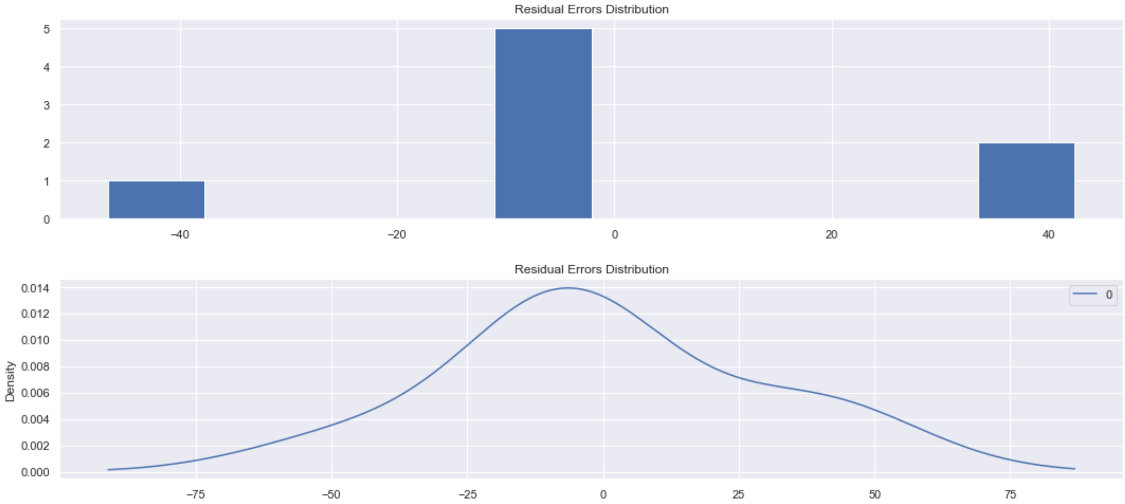

Figure 12: Residual Errors Distribution for W 302 Flat

When using significance level of 95%, the prediction interval will be = $\hat{y}_t \pm 1.96\ \sigma_{residuals}$ where 1.96 is the "z-multiplier" used to get 95% significance level. The Root Mean Squared Error is also called the standard deviation of the residuals, which means that the prediction interval now becomes $\hat{y}_t \pm 1.96 \times RMSE$.

It should be noted however that the standard deviation of the residuals can be substituted by RMSE due to the fact that the forecasts has been bias corrected, otherwise, RMSE would overestimate the standard deviation of the residuals (Silver et al., 2016).

Upon evaluating the 95% prediction intervals for the different categories we noticed that they are very broad, that is due to the fact the 95% intervals aim to include the true observation with 95% confidence. The broad intervals are also due to the fact that the RMSE values for the categories are close to the forecasted next period demands, which is similar to when we have very broad confidence intervals due to the fact that the standard deviations are close to the means.

For the purpose of our study, we decided to use 50% prediction intervals to ground our intervals closer to reality. By using the 50% prediction intervals we are sacrificing prediction confidence but we are gaining implementation realism which allows the company to use them better in practice. The prediction interval now becomes $\boldsymbol{\widehat{y}_t \pm 0.67 \times RMSE}$.

CHAPTER V

DEMAND FORECASTING RESULTS

A. Preliminary Results

Upon evaluating the results, we noticed that the Best Previous Period Demand Method performed surprisingly well, however, when we tried extrapolating (forecasting on new unseen data points) we noticed that it doesn't perform well at all. Upon digging deeper into the method, it became clear why this method didn't perform well on unseen data points. The Best Previous Period Demand turned out to be extremely biased to the training and testing data. It mimics the previous data and tries to predict new data points by finding the previous datum in the training dataset with the least error relative to the testing dataset. By doing so, the new predicted value can only take the form of a previous one from the training dataset, thus failing to capture inference from the training dataset and applying it to the testing dataset.

In light of these findings we decided to disregard the Best Previous Period Demand Method from the final evaluated methods.

Category	Sales Frequency	Sales Frequency Interval	Class Label	Theoretical Best Method to Use	RMSE	MAE	Forecast	Theoretical 2nd Best Method to Use	RMSE	MAE	Forecast	Theoretical 3rd Best Method to Use	RMSE	MAE	Forecast
M 310 Flat	0.04	**0-0.25**	**SF-25**	CBB-Croston-TSB (alpha = 0.1, beta = 0.15)	4.204	2.725	1.433	MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	4.610	3.000	1.593	CBB-Moving Average (window = 4 periods)	4.954	2.636	1.327
E 200/230 Round	0.22	**0-0.25**	**SF-25**	ARMA	41.20	27.23	16.31	Simple Exponential Smoothing (alpha = 0.039)	45.98	37.65	24.90	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	59.726	56.102	29.786
W 302 Flat	0.22	**0-0.25**	**SF-25**	Best Previous Period Demand	15.91	6.63	0.00	Simple Exponential Smoothing 1 (alpha = 0.1)	21.25	18.35	14.32	Croston 4 (alpha = 0.4)	21.899	18.319	14.620
W 302 Round	0.35	**0.25-0.5**	**SF-50**	Moving Average 2 (window = 4 periods)	1611.40	745.66	52.73	ARMA	1925.52	1273.00	75.54	Croston-TSB (alpha = 0.017, beta = 0.15)	2064.93	763.62	67.47
V 320 Flat	0.48	**0.25-0.5**	**SF-50**	Best Previous Period Demand	43.00	30.83	0.00	Croston 4 (alpha = 0.4)	47.87	37.78	21.55	Croston (alpha = 0.594)	49.22	31.85	17.48
M 310 Round	0.48	**0.25-0.5**	**SF-50**	ARMA	28.56	22.33	28.83	Last Period Demand	37.14	24.81	39.90	Simple Exponential Smoothing 4 (alpha = 0.4)	36.28	31.48	21.40
M 300/303 Round	0.7	**0.5-0.75**	**SF-75**	ARMA	61.83	42.42	151.59	ARIMA	87.43	71.61	90.53	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	124.52	120.96	109.96
M 238 Flat	0.74	**0.5-0.75**	**SF-75**	Moving Average 2 (window = 4 periods)	284.28	243.22	52.38	Croston-TSB (alpha = 0.231, beta = 0.15)	336.90	284.10	149.65	Simple Exponential Smoothing 3 (alpha = 0.3)	336.98	281.85	123.64
M 300/303 Flat	0.83	**0.75-1.0**	**SF-100**	Best Previous Period Demand	158.10	137.08	498.90	Croston 1 (alpha = 0.1)	169.31	142.37	185.61	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	170.55	144.77	186.08
M 238 Round	0.91	**0.75-1.0**	**SF-100**	Best Previous Period Demand	68.81	60.01	31.40	Last Period Demand	68.813	60.012	31.400	ARMA	71.83	63.17	18.80
K 110 Flat	0.91	**0.75-1.0**	**SF-100**	Moving Average 2 (window = 4 periods)	269.66	177.03	84.35	Moving Average 1 (window = 3 periods)	279.84	176.91	89.93	Simple Exponential Smoothing 1 (alpha = 0.1)	305.39	193.42	185.07
K 510 Round	0.91	**0.75-1.0**	**SF-100**	ARMA	27.02	23.47	7.49	Moving Average 2 (window = 4 periods)	33.41	27.83	24.18	Simple Exponential Smoothing 1 (alpha = 0.1)	37.33	35.10	50.52
K 110 Round	0.96	**0.75-1.0**	**SF-100**	ARMA	125.14	98.58	357.11	Simple Exponential Smoothing (alpha = 0.051)	125.56	110.48	224.79	Simple Exponential Smoothing 3 (alpha = 0.3)	126.97	114.63	279.34
V 320 Round	1	**0.75-1.0**	**SF-100**	Last Period Demand	624.86	479.44	1880.90	ARMA	663.28	511.12	1839.42	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	702.33	587.92	1130.36
K 100 Round	1	**0.75-1.0**	**SF-100**	Best Previous Period Demand	261.97	174.10	1611.91	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	409.00	324.43	751.80	ARMA	409.21	333.41	708.45
K 720 Flat	1	**0.75-1.0**	**SF-100**	Best Previous Period Demand	367.30	341.64	1878.58	MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	433.243	342.734	754.055	CBB-Moving Average (window = 4 periods)	459.665	352.286	742.771
K 720 Round	1	**0.75-1.0**	**SF-100**	Moving Average 2 (window = 4 periods)	227.18	171.06	342.88	ARMA	232.74	178.64	478.14	Simple Exponential Smoothing 1 (alpha = 0.1)	241.38	195.46	473.13
K 100 Flat	1	**0.75-1.0**	**SF-100**	ARMA	155.51	130.55	387.95	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	185.47	161.21	310.18	Best Previous Period Demand	186.18	163.61	470.15

Table 3: Preliminary Results

B. Final Results

Category	Sales Frequency	Sales Frequency Interval	Class Label	Theoretical Best Method to Use	RMSE	MAE	Forecast	Theoretical 2nd Best Method to Use	RMSE	MAE	Forecast	Theoretical 3rd Best Method to Use	RMSE	MAE	Forecast
M 310 Flat	0.04	**0-0.25**	**SF-25**	CBB-Croston-TSB (alpha = 0.1, beta = 0.15)	4.204	2.725	1.433	MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	4.610	3.000	1.593	CBB-Moving Average (window = 4 periods)	4.954	2.636	1.327
E 200/230 Round	0.22	**0-0.25**	**SF-25**	ARMA	41.20	27.23	16.31	Simple Exponential Smoothing (alpha = 0.039)	45.98	37.65	24.90	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	59.726	56.102	29.786
W 302 Flat	0.22	**0-0.25**	**SF-25**	Simple Exponential Smoothing 1 (alpha = 0.1)	21.25	18.35	14.32	Croston 4 (alpha = 0.4)	21.899	18.319	14.620	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	21.988	18.981	16.308
W 302 Round	0.35	**0.25-0.5**	**SF-50**	Moving Average 2 (window = 4 periods)	1611.40	745.66	52.73	ARMA	1925.52	1273.00	75.54	Croston-TSB (alpha = 0.017, beta = 0.15)	2064.93	763.62	67.47
V 320 Flat	0.48	**0.25-0.5**	**SF-50**	Croston 4 (alpha = 0.4)	47.87	37.78	21.55	Croston (alpha = 0.594)	49.22	31.85	17.48	Croston 3 (alpha = 0.3)	51.270	44.58	27.702
M 310 Round	0.48	**0.25-0.5**	**SF-50**	ARMA	28.56	22.33	28.83	Last Period Demand	37.14	24.81	39.90	Simple Exponential Smoothing 4 (alpha = 0.4)	36.28	31.48	21.40
M 300/303 Round	0.7	**0.5-0.75**	**SF-75**	ARMA	61.83	42.42	151.59	ARIMA	87.43	71.61	90.53	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	124.52	120.96	109.96
M 238 Flat	0.74	**0.5-0.75**	**SF-75**	Moving Average 2 (window = 4 periods)	284.28	243.22	52.38	Croston-TSB (alpha = 0.231, beta = 0.15)	336.90	284.10	149.65	Simple Exponential Smoothing 3 (alpha = 0.3)	336.98	281.85	123.64
M 300/303 Flat	0.83	**0.75-1.0**	**SF-100**	Croston 1 (alpha = 0.1)	169.31	142.37	185.61	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	170.55	144.77	186.08	Croston-TSB 2 (alpha = 0.2, beta = 0.15)	180.734	153.53	190.298
M 238 Round	0.91	**0.75-1.0**	**SF-100**	Last Period Demand	68.813	60.012	31.400	ARMA	71.83	63.17	18.80	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	124.62	95.88	100.74
K 110 Flat	0.91	**0.75-1.0**	**SF-100**	Moving Average 2 (window = 4 periods)	269.66	177.03	84.35	Moving Average 1 (window = 3 periods)	279.84	176.91	89.93	Simple Exponential Smoothing 1 (alpha = 0.1)	305.39	193.42	185.07
K 510 Round	0.91	**0.75-1.0**	**SF-100**	ARMA	27.02	23.47	7.49	Moving Average 2 (window = 4 periods)	33.41	27.83	24.18	Simple Exponential Smoothing 1 (alpha = 0.1)	37.33	35.10	50.52
K 110 Round	0.96	**0.75-1.0**	**SF-100**	ARMA	125.14	98.58	357.11	Simple Exponential Smoothing (alpha = 0.051)	125.56	110.48	224.79	Simple Exponential Smoothing 3 (alpha = 0.3)	126.97	114.63	279.34
V 320 Round	1	**0.75-1.0**	**SF-100**	Last Period Demand	624.86	479.44	1880.90	ARMA	663.28	511.12	1839.42	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	702.33	587.92	1130.36
K 100 Round	1	**0.75-1.0**	**SF-100**	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	409.00	324.43	751.80	ARMA	409.21	333.41	708.45	Simple Exponential Smoothing 1 (alpha = 0.1)	411.294	325.445	753.292
K 720 Flat	1	**0.75-1.0**	**SF-100**	CBB-Croston-TSB (alpha = 0.1, beta = 0.15)	419.051	330.037	743.202	MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	433.243	342.734	754.055	CBB-Moving Average (window = 4 periods)	459.665	352.286	742.771
K 720 Round	1	**0.75-1.0**	**SF-100**	Moving Average 2 (window = 4 periods)	227.18	171.06	342.88	ARMA	232.74	178.64	478.14	Simple Exponential Smoothing 1 (alpha = 0.1)	241.38	195.46	473.13
K 100 Flat	1	**0.75-1.0**	**SF-100**	ARMA	155.51	130.55	387.95	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	185.47	161.21	310.18	ARIMA	198.142	178.983	462.752

Table 4: Final Results

When analyzing the forecasting methods used, we noticed that the Box-Jenkins Methods allow negative integer values which is inappropriate for intermittent demand. The forecasts also exhibit an upward bias directly after a non-zero demand and that is the case for several evaluated methods, not just Box-Jenkins, such as Simple Moving Average and Simple Exponential Smoothing. These issues are not apparent when intermittent demand is not present.

C. Results by Class

1. Class SF-25 (High Intermittency)

Category	Theoretical Best Method to Use	RMSE	MAE	Forecast	Theoretical 2nd Best Method to Use	RMSE	MAE	Forecast	Theoretical 3rd Best Method to Use	RMSE	MAE	Forecast
M 310 Flat	CBB-Croston-TSB (alpha = 0.1, beta = 0.15)	4.204	2.72	1.433	MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	4.610	3.0	1.593	CBB-Moving Average (window = 4 periods)	4.954	2.636	1.327
E 200/230 Round	ARMA	41.20	27.23	16.31	Simple Exponential Smoothing (alpha = 0.039)	45.98	37.65	24.90	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	59.726	56.102	29.786
W 302 Flat	Simple Exponential Smoothing 1 (alpha = 0.1)	21.25	18.35	14.321	Croston 4 (alpha = 0.4)	21.899	18.319	14.620	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	21.988	18.981	16.308

Table 5: Class SF-25 (High Intermittency)

Croston or one of its variants proved to be the best method to use in the SF-25 Class with **probability = 55%** of being one of the top three methods. While simple exponential smoothing and bootstrapping came in second with **probability = 22%** each.

2. *Class SF-50 (Medium Intermittency)*

Category	Theoretical Best Method to Use	RMSE	MAE	Forecast	Theoretical 2nd Best Method to Use	RMSE	MAE	Forecast	Theoretical 3rd Best Method to Use	RMSE	MAE	Forecast
W 302 Round	Moving Average 2 (window = 4 periods)	1611.40	745.66	52.73	ARMA	1925.52	1273.00	75.54	Croston-TSB (alpha = 0.017, beta = 0.15)	2064.93	763.62	67.47
V 320 Flat	Croston 4 (alpha = 0.4)	47.87	37.78	21.55	Croston (alpha = 0.594)	49.22	31.85	17.48	Croston 3 (alpha = 0.3)	51.270	44.58	27.702
M 310 Round	ARMA	28.56	22.33	28.83	Last Period Demand	37.14	24.81	39.90	Simple Exponential Smoothing 4 (alpha = 0.4)	36.28	31.48	21.40

Table 6: Class SF-50 (Medium Intermittency)

Croston or one of its variants proved to be the best method to use in the SF-50 Class with **probability = 44%** of being one of the top three methods. While ARMA and ARIMA came in second with **probability = 22%**.

3. *Class SF-75 (Low Intermittency)*

Category	Theoretical Best Method to Use	RMSE	MAE	Forecast	Theoretical 2nd Best Method to Use	RMSE	MAE	Forecast	Theoretical 3rd Best Method to Use	RMSE	MAE	Forecast
M 300/303 Round	ARMA	61.83	42.42	151.59	ARIMA	87.43	71.61	90.53	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	124.52	120.96	109.96
M 238 Flat	Moving Average 2 (window = 4 periods)	284.28	243.22	52.38	Croston-TSB (alpha = 0.231, beta = 0.15)	336.90	284.10	149.65	Simple Exponential Smoothing 3 (alpha = 0.3)	336.98	281.85	123.64

Table 7: Class SF-75 (Low Intermittency)

Both Croston-TSB and Box-Jenkins (ARMA and ARIMA) proved to be the best methods to use in the SF-75 Class with **probability = 33%** of being one of the top three methods.

4. *Class SF-100 (No Intermittency)*

Category	Theoretical Best Method to Use	RMSE	MAE	Forecast	Theoretical 2nd Best Method to Use	RMSE	MAE	Forecast	Theoretical 3rd Best Method to Use	RMSE	MAE	Forecast
M 300/303 Flat	Croston 1 (alpha = 0.1)	169.31	142.37	185.61	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	170.55	144.77	186.08	Croston-TSB 2 (alpha = 0.2, beta = 0.15)	180.734	153.53	190.298
M 238 Round	Last Period Demand	68.813	60.012	31.400	ARMA	71.83	63.17	18.80	Croston-TSB 4 (alpha = 0.4, beta = 0. 15)	124.62	95.88	100.74
K 110 Flat	Moving Average 2 (window = 4 periods)	269.66	177.03	84.35	Moving Average 1 (window = 3 periods)	279.84	176.91	89.93	Simple Exponential Smoothing 1 (alpha = 0.1)	305.39	193.42	185.07
K 510 Round	ARMA	27.02	23.47	7.49	Moving Average 2 (window = 4 periods)	33.41	27.83	24.18	Simple Exponential Smoothing 1 (alpha = 0.1)	37.33	35.10	50.52
K 110 Round	ARMA	125.14	98.58	357.11	Simple Exponential Smoothing (alpha = 0.051)	125.56	110.48	224.79	Simple Exponential Smoothing 3 (alpha = 0.3)	126.97	114.63	279.34
V 320 Round	Last Period Demand	624.86	479.44	1880.90	ARMA	663.28	511.12	1839.42	Croston-TSB 4 (alpha = 0.4, beta = 0.15)	702.33	587.92	1130.36
K 100 Round	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	409.00	324.43	751.80	ARMA	409.21	333.41	708.45	Simple Exponential Smoothing 1 (alpha = 0.1)	411.294	325.445	753.292
K 720 Flat	CBB-Croston-TSB (alpha = 0.1, beta = 0.15)	419.051	330.037	743.202	MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	433.243	342.734	754.055	CBB-Moving Average (window = 4 periods)	459.665	352.286	742.771
K 720 Round	Moving Average 2 (window = 4 periods)	227.18	171.06	342.88	ARMA	232.74	178.64	478.14	Simple Exponential Smoothing 1 (alpha = 0.1)	241.38	195.46	473.13
K 100 Flat	ARMA	155.51	130.55	387.95	Croston-TSB 1 (alpha = 0.1, beta = 0.15)	185.47	161.21	310.18	ARIMA	198.142	178.983	462.752

Table 8: Class SF-100 (No Intermittency)

Croston or one of its variants proved to be the best method to use in the SF-100 Class with **probability = 30%** of being one of the top three methods. Box-Jenkins (ARMA and ARIMA) and simple exponential smoothing came in second and third with **probability = 27%** and **probability = 20%**, respectively.

D. Results Validation

In order to make sure the results of the case study conform to the reality of the actual market, and essentially improves the forecasting performance of the company, we asked FOZ Trading to provide us with new data to see if our findings can be confirmed. The company provided us with the 2018 data to validate our findings.

The results from implementing the models on the new dataset confirm our findings. Croston or one of its variants remained the favorable method to use in the four classes. And Box-Jenkins and simple exponential smoothing continued to show impressive results. However, simple exponential smoothing seemed to out-perform Box-Jenkins in the SF- 50 class which was not the case when the models were implemented on the old dataset. Regardless of which methods came in second or third, it seems that the best three methods to use are almost always Croston-TSB, ARMA, and simple exponential smoothing.

CHAPTER VI

OPTIMAL ORDERING POLICY

For our case study we decided to implement the order up-to model to optimize the order quantity and inventory level.

The order up-to model is a periodic review ordering policy that has a fixed review period. The reason for choosing a periodic ordering policy over a continuous one is because it allows the company to better forecast the orders made over a period of time. So instead of continuously reviewing every product every time it goes below the reorder point level, the order up-to model specifies a certain time period where the inventory is compared to a stock up level (S). "This periodic review service level model is very useful in retail settings, in particular. It is common in retailing to place orders at fixed points in time to take advantage of bundling multiple orders together" (Nahmias and Olsen, 2015).

The order up-to model also states that when an order is placed, there should be enough inventory to cover uncertain demand while waiting for delivery; i.e., inventory is needed because of both uncertain demand and lead time. According to Nahmias and Olsen (2015), The stock is replenished if the inventory is less than the stock up level (S) at the time period specified. This time period is dependent on the review period and the lead time.

In order to implement the order up-to model we have to define a distribution for the forecasted demand. A practical approach suggested by Silver et al. (2016) is to use the RMSE of the unbiased forecast as the standard deviation of the demand, and to use a normal distribution of lead time demand. In other words, assume that the demand is normal with mean as the forecasted value over the lead time and the RMSE as standard deviation.

A. Notation[10]

P	period length
l	lead time
S	stock up level
μ	expected demand
σ	standard deviation of demand
μ_{l+1}	expected demand over $l+1$ periods
σ_{l+1}	standard deviation of the demand over $l+1$ periods
$N(\mu_{l+1}, \sigma_{l+1})$	normal distribution with mean μ_{l+1}, and std. dev σ_{l+1}
$n(S)$	the expected number of demands that stock out at the end of the period
β	fill rate
$f(t)$	pdf of demand distribution at instance t

B. Computation[10]

$l = 1$ period

$\beta = 0.99$

$\mu =$ forecasted demand over lead time

$\sigma =$ RMSE of the unbiased forecast

Expected demand over 2 periods, $\mu_2 = 2 \times \mu$

Standard deviation of demand over 2 periods, $\sigma_2 = \sqrt{2} \times \sigma$

[10] Adapted from Nahmias and Olsen (2015)

demand distribution $= N(\mu_2, \sigma_2)$

$$n(S) = (1-\beta)\mu \qquad (1)$$

$$n(S) = \int_S^\infty (t-S) f(t)dt \qquad (2)$$

S is computed by finding the root of *(1) – (2):*

$$(1-\beta)\mu - \int_S^\infty (t-S) f(t)dt = 0 \qquad (1) - (2)$$

C. Results

CATEGORY	INVENTORY LEVEL	S
W 302 ROUND	15051	211
K 720 ROUND	10078	1300
K 100 FLAT	9107	1550
M 300/303 ROUND	9051	607
V 320 ROUND	8742	7600
K 110 ROUND	7830	1440
M 300/303 FLAT	5929	760
K 100 ROUND	4619	3000
M 238 ROUND	4073	400
V 320 FLAT	3385	86
E 200/230 ROUND	2694	66
M 310 ROUND	2432	115
K 110 FLAT	2342	740
W 302 FLAT	2095	57
K 720 FLAT	1390	2900
K 510 ROUND	1162	200
M 310 FLAT	65	6
M 238 FLAT	0	600

Table 9: Stock Up Level (S) Computed for Every Category Using a Fill Rate of 0.99

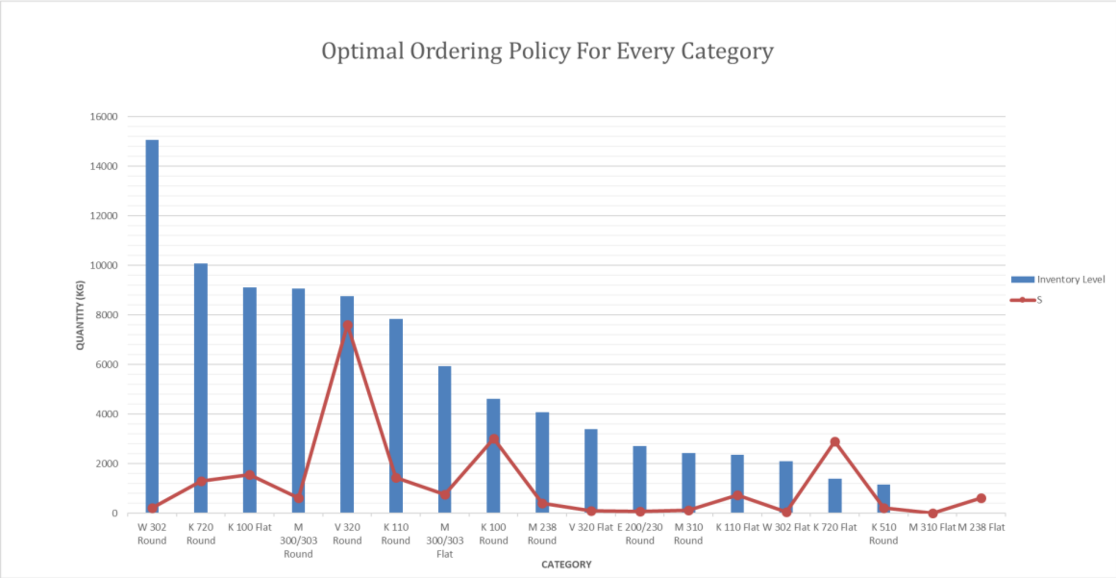

Figure 13: Order Up-To Model for the 18 Categories

Nahmias and Olsen (2015) suggest comparing the inventory level ***u*** to the order up-to point ***S***. If ***u*** turned out to be smaller than ***S***, an order of size ***S* - *u*** should be placed, otherwise no order should be placed.

We noticed, after computing the order up to level ***S*** of every category using a fill rate of 0.99, that the implemented optimal ordering policy advocates restocking 600 kg (600-0) of M 238 Flat, and 1510 kg (2900-1390) of K 720 Flat in order to avoid stockouts. This is due to the fact that the forecasted demand over the lead time for the next two periods exceeds the amount currently available in stock. The implemented optimal ordering policy also recommends the user to pay close attention to V 320 Round and K 100 Round, since their inventory levels are very close to their order up to levels, thus they both might require restocking in the near future.

D. Disaggregation of the Categories

After computing the forecasted demand (D_i) and the stock up level (S_i) for every category, it is important to know the forecasted demand $(D_{i,j})$ and stock up level $(S_{i,j})$ for every SKU j in parent category i. In order to do so, a disaggregation rule was introduced based on the percentage of demand (sales) for each SKU from its parent category. The percentage is assigned to every SKU as a weight factor.

Mathematically,

$$w_{i,j} = \frac{\lambda_{i,j}}{\sum_{j=1}^{n} \lambda_{i,j}} \quad \forall i = 1,2,\dots,18$$

Where,

$w_{i,j} =$ weight assigned to SKU j in product family i

$\lambda_{i,j} =$ demand for SKU j in product family i

Then, the individual forecasts and the stock up level for every SKU is computed as follows:

$$D_{i,j} = w_{i,j} D_i$$

$$S_{i,j} = w_{i,j} S_i$$

It should be noted however, that instead of applying this aggregate rule, FOZ Trading can apply market knowledge in order to figure out every SKU's individual demand, and subsequently its stock up level. If the disaggregation process was done using the formula, market knowledge and human input are not needed anymore anywhere in the demand forecasting process, thus making this forecasting process reliant only on previous data.

Silver et al. (2016) argues that human judgment and input represents a crucial step in the forecasting framework and should not be neglected. That's why it's preferable to disaggregate the categories based on market knowledge rather than the mathematical formula stated above. By doing so, user input and market knowledge become an integral part in this forecasting process, thus enhancing its credibility and usefulness.

CHAPTER VII

APPLICATION OUTPUT

All the models were implemented using Jupyter Notebook (Kluyver et al., 2016) as the hosting application and Python as the coding language. The finished product is a fully automated Jupyter web application that automatically implements all the methods, checks for stationarity, applies walk forward optimization, implements order up-to model to optimize order quantity, displays and plots the all the outputs, and exports the best methods to use. The only thing the user has to do is to import the data and run the program.

The following graphs and tables are the outputs the application displays for K 720 Round after it finishes implementing and evaluating the methods.

A. Last Period Demand (Naïve Forecast)

Figure 14: Forecasting on K 720 Round Using Last Period Demand

B. Best Previous Period Demand

Figure 15: Forecasting on K 720 Round Using Best Previous Period Demand

C. Simple Moving Average

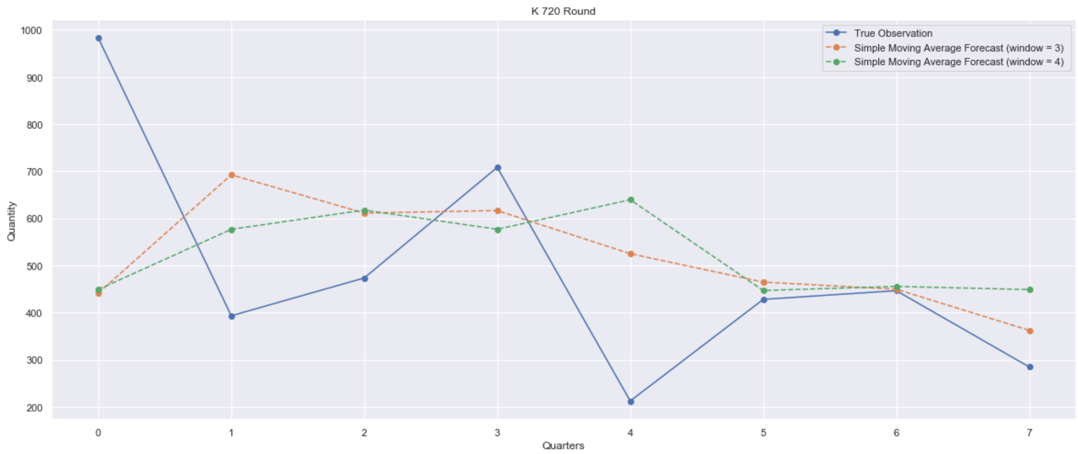

Figure 16: Forecasting on K 720 Round Using Simple Moving Average

D. Simple Exponential Smoothing

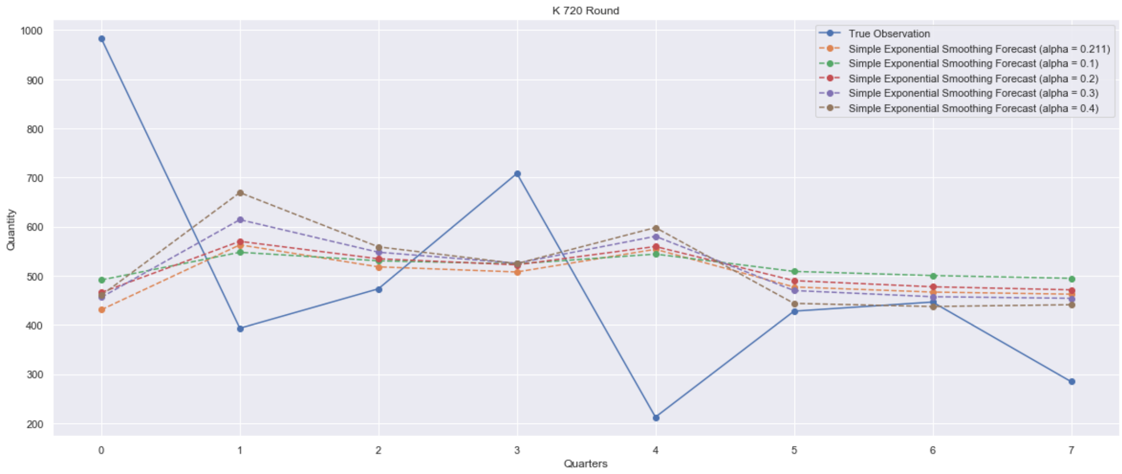

Figure 17: Forecasting on K 720 Round Using Simple Exponential Smoothing

E. Box-Jenkins

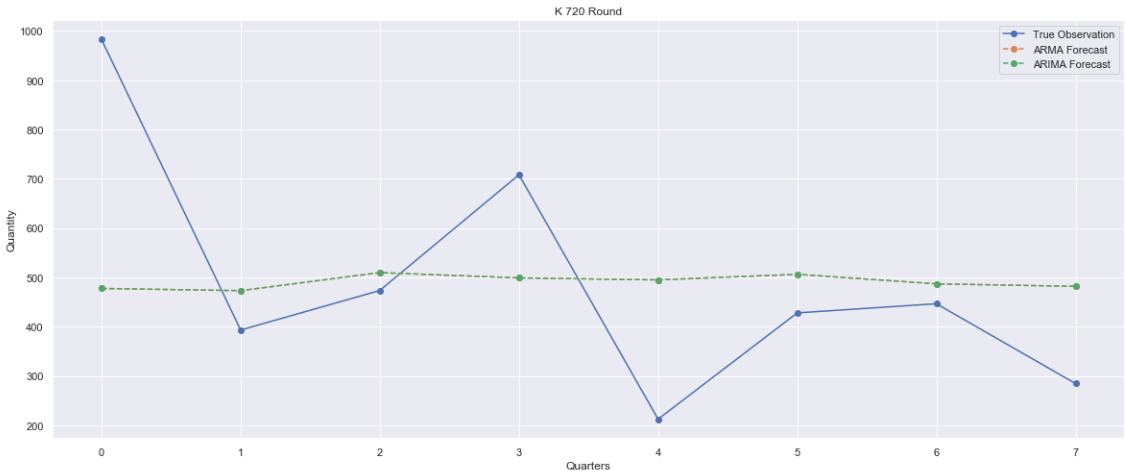

Figure 18: Forecasting on K 720 Round Using Box-Jenkins

F. Croston

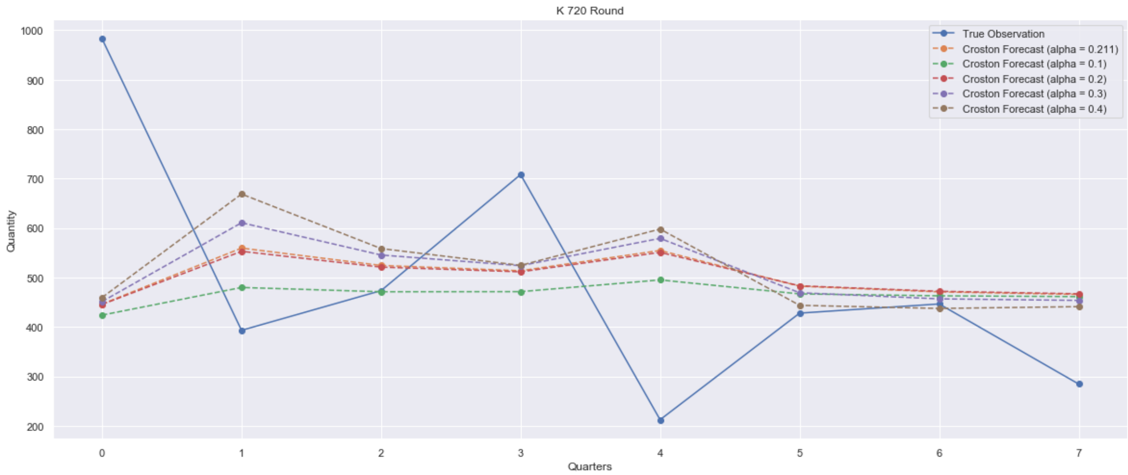

Figure 19: Forecasting on K 720 Round Using Croston

G. Croston TSB

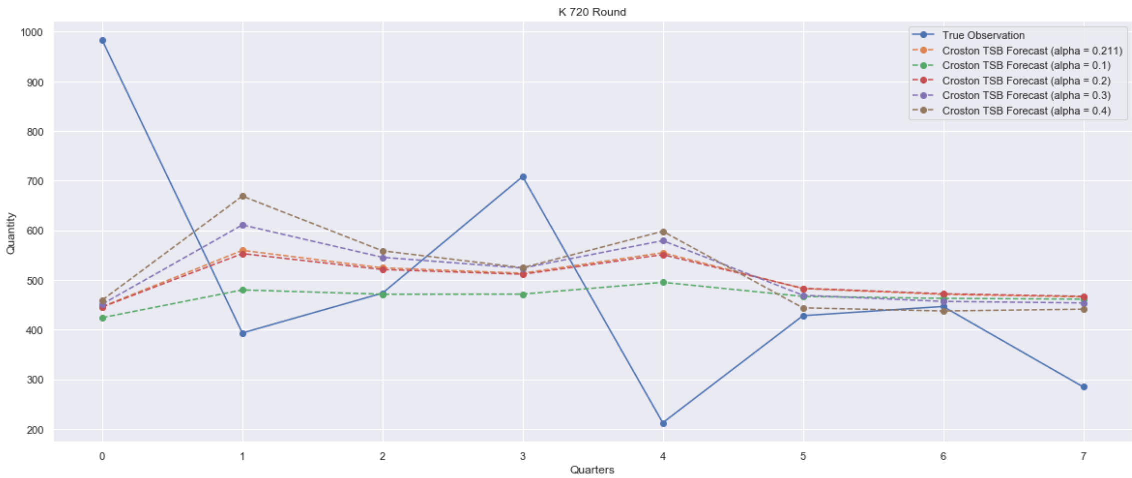

Figure 20: Forecasting on K 720 Round Using Croston TSB

H. Bootstrap

The bootstrap method cannot be plotted, because it is based on the fact that it generates a number of bootstrapping samples from the time series (for our case: 1000 bootstrapping samples using MBB and 1000 bootstrapping samples using CBB). And then it implements the same method (in our case: Croston-TSB and Simple MA) on every one of those samples. Thus, we would have 1000 generated plots for every method used (4000 in total).

	0	1	2	3	4	5	6	7	8	9	...	990	991	992	993	994	995	996	997	998	999
0	708.200	457.700	446.530	348.03	356.400	427.990	427.990	393.200	284.600	393.200	...	229.70	382.700	457.700	188.600	348.030	427.990	348.030	284.60	393.200	356.400
1	212.400	982.900	284.600	472.30	188.600	446.530	446.530	473.464	295.310	473.464	...	636.30	661.310	982.900	382.700	472.300	446.530	472.300	295.31	473.464	188.600
2	427.990	393.200	295.310	229.70	382.700	284.600	284.600	708.200	1833.400	708.200	...	457.70	221.900	393.200	661.310	229.700	284.600	229.700	1833.40	708.200	382.700
3	284.600	982.900	212.400	457.70	284.600	473.464	393.200	212.400	708.200	221.900	...	188.60	982.900	212.400	982.900	295.310	457.700	708.200	1833.40	295.310	221.900
4	295.310	393.200	427.990	982.90	295.310	708.200	473.464	427.990	212.400	348.030	...	382.70	393.200	427.990	393.200	1833.400	982.900	212.400	356.40	1833.400	348.030
5	1833.400	473.464	446.530	393.20	1833.400	212.400	708.200	446.530	427.990	472.300	...	661.31	473.464	446.530	473.464	356.400	393.200	427.990	188.60	356.400	472.300
6	221.900	212.400	982.900	295.31	982.900	284.600	295.310	229.700	427.990	348.030	...	427.99	661.310	1833.400	356.400	661.310	473.464	295.310	382.70	188.600	473.464
7	348.030	427.990	393.200	1833.40	393.200	295.310	1833.400	636.300	446.530	472.300	...	446.53	221.900	356.400	188.600	221.900	708.200	1833.400	661.31	382.700	708.200
8	472.300	446.530	473.464	356.40	473.464	1833.400	356.400	457.700	284.600	229.700	...	284.60	348.030	188.600	382.700	348.030	212.400	356.400	221.90	661.310	212.400
9	982.900	661.310	661.310	229.70	1833.400	348.030	982.900	472.300	446.530	382.700	...	188.60	982.900	393.200	472.300	1833.400	212.400	356.400	472.30	457.700	636.300
10	393.200	221.900	221.900	636.30	356.400	472.300	393.200	229.700	284.600	661.310	...	382.70	393.200	473.464	229.700	356.400	427.990	188.600	229.70	982.900	457.700
11	473.464	348.030	348.030	457.70	188.600	229.700	473.464	636.300	295.310	221.900	...	661.31	473.464	708.200	636.300	188.600	446.530	382.700	636.30	393.200	982.900
12	708.200	212.400	356.400	457.70	457.700	636.300	661.310	982.900	473.464	188.600	...	188.60	188.600	212.400	472.300	295.310	457.700	295.310	1833.40	212.400	188.600
13	212.400	427.990	188.600	982.90	982.900	457.700	221.900	393.200	708.200	382.700	...	382.70	382.700	427.990	229.700	1833.400	982.900	1833.400	356.40	427.990	382.700
14	427.990	446.530	382.700	393.20	393.200	982.900	348.030	473.464	212.400	661.310	...	661.31	661.310	446.530	636.300	356.400	393.200	356.400	188.60	446.530	661.310
15	427.990	229.700	221.900	212.40	284.600	457.700	427.990	221.900	356.400	661.310	...	284.60	1833.400	393.200	472.300	446.530	636.300	636.300	348.03	295.310	473.464
16	446.530	636.300	348.030	427.99	295.310	982.900	446.530	348.030	188.600	221.900	...	295.31	356.400	473.464	229.700	284.600	457.700	457.700	472.30	1833.400	708.200
17	284.600	457.700	472.300	446.53	1833.400	393.200	284.600	472.300	382.700	348.030	...	1833.40	188.600	708.200	636.300	295.310	982.900	982.900	229.70	356.400	212.400
18	229.700	708.200	188.600	188.60	457.700	473.464	661.310	982.900	661.310	473.464	...	708.20	708.200	393.200	446.530	473.464	661.310	473.464	472.30	982.900	382.700
19	636.300	212.400	382.700	382.70	982.900	708.200	221.900	393.200	221.900	708.200	...	212.40	212.400	473.464	284.600	708.200	221.900	708.200	229.70	393.200	661.310

20 rows × 1000 columns

Figure 21: 1000 Bootstrapping Samples Generated Using CBB for K 720 Round

I. Summary Graph

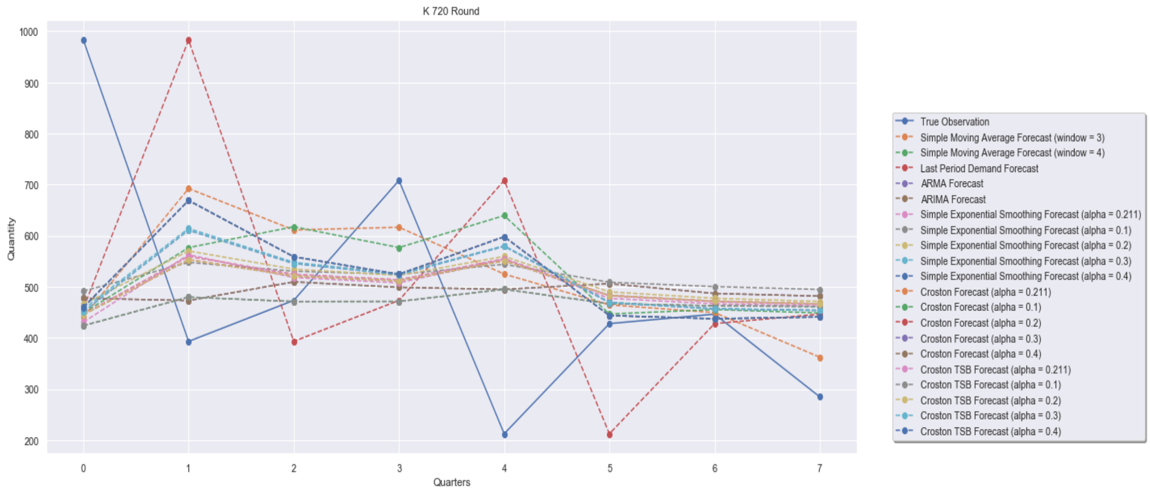

Figure 22: Summary Graph of K 720 Round Forecasts

J. Summary Table

Method	RMSE	MAE	Next Period Forecasted Demand	50% Prediction Interval
Last Period Demand	354.257	290.220	284.600	[47.248, 521.952]
Best Previous Period Demand	296.878	269.291	446.530	[247.622, 645.438]
Moving Average 1 (window = 3 periods)	282.608	214.268	386.373	[197.026, 575.721]
Moving Average 2 (window = 4 periods)	227.184	171.056	342.880	[190.667, 495.093
Simple Exponential Smoothing (alpha = 0.211)	256.615	194.485	424.878	[252.946, 596.811]
Simple Exponential Smoothing 1 (alpha = 0.1)	241.377	195.457	473.127	[311.404, 634.849]
Simple Exponential Smoothing 2 (alpha = 0.2)	249.241	195.988	434.034	[267.043, 601.026]
Simple Exponential Smoothing 3 (alpha = 0.3)	257.681	199.394	403.268	[230.621, 575.914]
Simple Exponential Smoothing 4 (alpha = 0.4)	265.502	204.342	378.504	[200.618, 556.39]
ARMA	232.739	178.640	478.143	[322.208, 634.078]
ARIMA	232.739	178.640	478.143	[322.208, 634.078]
Croston (alpha = 0.211)	252.405	193.866	427.571	[258.459, 596.682]
Croston 1 (alpha = 0.1)	247.208	174.914	443.641	[278.011, 609.27]

Croston 2 (alpha = 0.2)	251.690	192.885	430.529	[261.897, 599.162]
Croston 3 (alpha = 0.3)	258.367	199.108	402.984	[229.878, 576.09]
Croston 4 (alpha = 0.4)	265.625	204.349	378.490	[200.521, 556.458]
Croston-TSB (alpha = 0.211, beta = 0.15)	252.405	193.866	427.571	[258.459, 596.682]
Croston-TSB 1 (alpha = 0.1, beta = 0.15)	247.208	174.914	443.641	[278.011, 609.27]
Croston-TSB 2 (alpha = 0.2, beta = 0.15)	251.690	192.885	430.529	[261.897, 599.162]
Croston-TSB 3 (alpha = 0.3, beta = 0.15)	258.367	199.108	402.984	[229.878, 576.09]
Croston-TSB 4 (alpha = 0.4, beta = 0.15)	265.625	204.349	378.490	[200.521, 556.458]
CBB-Croston-TSB (alpha = 0.1, beta = 0.15)	330.675	243.426	493.982	[272.43, 715.535]
MBB-Croston-TSB (alpha = 0.1, beta = 0.15)	313.469	233.335	493.388	[283.364, 703.413]
CBB-Moving Average (window = 4 periods)	372.945	268.853	479.893	[230.02, 729.766]
MBB-Moving Average (window = 4 periods)	343.024	252.430	494.871	[265.045, 724.697]

Table 10: Summary Table of K 720 Round Forecasts

K. Optimal Ordering Policy

```
    fjac: array([[-1.]])
     fun: array([7.54951657e-15])
 message: 'The solution converged.'
    nfev: 16
     qtf: array([-3.86573618e-09])
       r: array([0.02784094])
  status: 1
 success: True
       x: array([1300.54674356])

the stock up level (S) = [1300.54674356]
```

Figure 23: Computation of Stock Up Level (S) for K 720 Round

CHAPTER VIII

IMPACT OF DEMAND FORECASTING AND OPTIMAL ORDERING POLICY

Building the application, implementing different forecasting methods, and applying the order up-to ordering policy to the historical data gathered from FOZ Trading yielded several improvements to the company's operations and supply chain.

Upon reviewing the inventory data in 2019 and 2020, we noticed that the company is now maintaining lower end of the month inventory on average, lesser number of stockout days, lower inventory turnover rate, and improved perfect order fulfillment rate which indicates the effectiveness of the demand forecasting methods and optimal ordering policy adapted in the application.

The aforementioned application optimized the inventory level while reducing demand uncertainty and cost. Now the company has a better understanding of its customer needs and can implement a more efficient strategy in terms of warehouse management, equipment and labor.

Some KPIs that show the rate and extent of improvement of the company's operations and supply chain are highlighted below:

- Inventory level before implementation = 395,535 in US Dollar Value

  Inventory level after implementation = 385,690 in US Dollar Value
- Perfect order fulfillment rate before implementation =N/A (likely lower than 90%)

  Perfect order fulfillment rate after implementation = 93.8 %

- Inventory turnover rate before implementation = 18.5 %

  Inventory turnover rate after implementation = 10.63 %

- Backorder rate before implementation = N/A (likely higher than 8%)

  Backorder rate after implementation = 6%

CHAPTER IX

CONCLUSION AND RECOMMENDATION

A. Conclusion

In this case study, we have applied basic, traditional and alternate forecasting techniques to reduce demand uncertainty and optimize inventory level of the Lebanese steel company FOZ Trading.

Analysis of the historical data revealed heavy intermittent demand characteristics. In order to reduce the intermittency of the demand, temporal aggregation was applied to every time series present in the data. Basic forecasting methods such as last period demand and best previous period demand were benchmarked against simple moving average, simple exponential smoothing, and Box-Jenkins from the traditional forecasting methods; and against Croston, Croston TSB, moving block bootstrapping and circular block bootstrapping from the alternate forecasting methods.

Results analysis indicated that the alternate methods such as Croston, Bootstrapping, and their variants tend, most of the times, to provide better and more reasonable forecasts than traditional and basic methods especially when intermittent demand is present. The results also showed that the Box-Jenkins performed better than (or at least as good as) simple exponential smoothing in most cases.

The evaluation and testing of the different demand forecasting methods allowed us to predict the next period demand with least amount of error, which in return helped FOZ Trading maintain optimal levels of inventory using the order up-to model as part of the optimal ordering policy.

The implementation of a suitable optimal ordering policy, aided by reasonable demand forecasts, improved the company's inventory level, perfect order fulfillment rate, inventory turnover and backorder rate while at the same time increased the company's understanding of its customer needs.

B. Recommendation

Several forecasting researches such as Armstrong (2001), Hibon and Evgeniou (2005) suggest combining the best forecasts instead of taking only one forecast. Their research proved that in practice it is less risky to combine forecasts than to select an individual forecasting method.

That's why we recommend FOZ Trading to combine the top three performing forecasts when they are uncertain about the situation, unsure about the method, or when they want to avoid large errors. "Compared with errors of the typical individual forecast, combining reduces errors" (Armstrong, 2001).

BIBLIOGRAPHY

Armstrong, J. S. (2001). Combining forecasts. In *Principles of forecasting* (pp. 417 439). Springer, Boston, MA.

Athanasopoulos, G., Hyndman, R. J., Song, H. & WU, D. C. 2009. The tourism forecasting competition. Working Paper 10/08. Monash University, Australia.

Baum, C. F. (2006). KPSS: Stata module to compute Kwiatkowski-Phillips-Schmidt-Shin test for stationarity.

Boylan, J. E., & Syntetos, A. A. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting*, *21*(2), 303-314.

Brockwell, P. J., Davis, R. A., & SpringerLink (Online service). (2016). *Introduction to time series and forecasting* (3rd 2016; ed.). Cham: Springer International Publishing.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The annals of statistics*, *14*(3), 1171-1179.

Chen, A., Hsu, C. H. & Blue, J. 2007. Demand planning approaches to aggregating and forecasting interrelated demands for safety stock and backup capacity planning. *International Journal of Production Research*, 45, 2269-2294.

Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Journal of the Operational Research Society*, *23*(3), 289-303.

Dickey, D. A., & Pantula, S. G. (2002). Determining the order of differencing in autoregressive processes. *Journal of Business & Economic Statistics*, *20*(1), 18-24.

Doszyń, M. (2019). Intermittent demand forecasting in the enterprise: Empirical verification. *Journal of Forecasting*, *38*(5), 459-469.

Gavin, J. (2013). Demand for steel continues to climb: Investment in Middle East and North Africa region's steel production is increasing as demand from infrastructure, reconstruction and real estate projects rises. *MEED Middle East Economic Digest*, *57*(9), 32.

Hibon, M., & Evgeniou, T. (2005). To combine or not to combine: selecting among forecasts and their combinations. *International journal of forecasting*, *21*(1), 15-24.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Kang, E. (2017). Time Series: ARIMA Model. Retrieved from https://medium.com/@kangeugine/time-series-arima-model-11140bc08c6.

Kirkpatrick II, C. D., & Dahlquist, J. A. (2010). *Technical analysis: the complete resource for financial market technicians*. FT press.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., ... & Ivanov, P. (2016, May). Jupyter Notebooks-a publishing format for reproducible computational workflows. In *ELPUB* (pp. 87-90).

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The annals of Statistics,* 17(3), 1217-1241.

Masters, T. (2013). *Assessing and Improving Prediction and Classification*. Create Space Independent Publishing Platform.

Nahmias, S., & Olsen, T. L. (2015). *Production and operations analysis*. Waveland Press.

Nikolopoulos, K., Syntetos, A. A., Boylan, J. E., Petropoulos, F., & Assimakopoulos, V. (2011). An aggregate-disaggregate intermittent demand approach (ADIDA) to forecasting: An empirical proposition and analysis. *Journal of the Operational Research Society, 62*(3), 544-554.

Politis, D., & Romano, J. P. (1992). Exploring the limits of bootstrap. *A circular block resampling procedure for stationary data. Wiley-Interscience*, 263-270.

Reid, R. A. (1987). the ABC method in hospital inventory management a practical. *Production and Inventory Management Journal, 28*(4), 67.

Silver, E. A., Pyke, D. F., & Thomas, D. J. (2016;2017;). *Inventory and production management in supply chains* (Fourth ed.). Baton Rouge: CRC Press Inc.

St-Pierre, N. R. (2003). Reassessment of biases in predicted nitrogen flows to the duodenum by NRC 2001. *Journal of Dairy Science, 86*(1), 344-350.

Tabar, B. R., Babai, M. Z., Syntetos, A. A., & Ducq, Y. (2013). The impact of temporal aggregation on demand forecasting of ARMA(1,1) process: Theoretical analysis. *IFAC Proceedings Volumes, 46*(9), 1696-1701.

Tersine, R. J. (1982). *Principles of inventory and materials management* (2nd ed.). New York: North-Holland.

Teunter, R. H., & Duncan, L. (2009). Forecasting intermittent demand: A comparative study. *Journal of the Operational Research Society,* 60(3), 321-329.

Teunter, R. H., Babai, M. Z., & Syntetos, A. A. (2010). ABC classification: Service levels and inventory costs. *Production and Operations Management, 19(3), 343-352*.

Teunter, R. H., Babai, M. Z., & Syntetos, A. A. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, *214*(3), 606-615.

Waller, D. (2015). Methods for Intermittent Demand Forecasting. *Lancaster University*. Retrieved from https://www.lancaster.ac.uk/pg/waller/pdfs/Intermittent_Demand_Forecasting.pdf.

Wesner, J. (2016). MAE and RMSE — Which Metric is Better? Retrieved from https://medium.com/human-in-a-machine -world/mae-and-rmse-which-metric-is-better-e60ac3bde13d.

Xiao, Z., & Phillips, P. C. (1998). An ADF coefficient test for a unit root in ARMA models of unknown order with empirical applications to the US economy. *The Econometrics Journal*, *1*(2), 27-43.

Zakharova, A. P., Vinichuk, O. Y., & Maksimova, D. A. (2016). Analysis and forecasting of sales volumes of the regional insurance company: The Russian experience. *Journal of Internet Banking and Commerce : JIBC, 21*(S3), 1.