REVA UNIVERSITY
Bengaluru, India

A Project Report on

Automated Forecasting & Capacity Solution for Back-office Operations

Submitted in partial fulfilment for award of degree of

PGDM

In **Business Analytics**

Submitted by

Anuraag Anand

2500014

Under the Guidance of

Dr. J.B. Simha

Chief Mentor- Analytics

REVA Academy for Corporate Excellence

REVA University

Rukmini Knowledge Park, Kattigenahalli,

Yelahanka, Bangalore – 560064

October, 2020

Candidate's Declaration

I, Anuraag Anand hereby declare that I have completed the project work towards the aster of Technology in Cybersecurity (PGDM Business Analytics) at, REVA University on the topic entitled Automated Forecasting & Capacity Solution for Backoffice-Operations under the supervision of Dr. J.B. Simha- Chief Mentor Analytics and Dr. Shinu Abhi- Director, Director Corporate Trainings. This report embodies the original work done by me in partial fulfilment of the requirements for the award of degree for the academic year 2020.

Place: Bengaluru

Date:

Name of the Student: Anuraag Anand

Signature of Student

REVA UNIVERSITY
Bengaluru, India

Certificate

This is to Certify that the PROJECT work entitled **Automated Forecasting & Capacity Solution for Backoffice-Operations** carried out by Anuraag Anand with SRN 2500014, is a bonafide student of REVA University, is submitting the project report in fulfilment for the award of PGDM in Business Analytics during the academic year 2020. The Project report has been tested for plagiarism, and has passed the plagiarism test with the similarity score less than 15%. The project report has been approved as it satisfies the academic requirements in respect of PROJECT work prescribed for the said Degree.

<Signature of the Guide>
<Name of the Guide>
Guide

<Signature of the Director>
<Name of the Director>
Director

External Viva

Names of the Examiners

1. <Name> <Designation> <Signature>
2. <Name> <Designation> <Signature>

Place: Bengaluru

Date:

Acknowledgement

I would like to express my gratitude to Dr. Simha for his guidance throughout as well as Dr. Shinu Abhi. They have helped me in my research and get thorough understating of the subject. Secondly I would like to thanks my friends and my wife who has helped me a lot in finalising the report out within short span of time.

Also want to express my gratitude to Hon’ble Chancellor, Dr. P Shayma Raju, Vice Chancellor, Dr. K. Mallikharjuna Babu, and Registrar, Dr. M. Dhanamjaya, who gave me the opportunity to undertake the course, helping me enhanced my skills.

Place: Bengaluru

Date: 21-10-2020

REVA UNIVERSITY
Bengaluru, India

Similarity Index Report

This is to certify that this project report titled **Automated Forecasting & Capacity Solution for Back-office Operations** was scanned for similarity detection. Process and outcome is given below.

Software Used: Turnitin
Date of Report Generation:
Similarity Index in %: 21%
Total word count: 3412
Name of the Guide: Dr. J. B Simha

Place: Bengaluru
Date:29th Oct 2020

Name of the Student: Anuraag Anand
Signature of Student

Verified by:

Signature
Dr. Shinu Abhi,
Director, Corporate Training

List of Abbreviations

Sl. No	Abbreviation	Long Form
1	ODE	Operating Deal Economics
2	SLA	Service Level Agreement
3	P2P	Procure to Pay
4	O2C	Order to Cash
5	R2R	Record to Report
6	FP&A	Financial Planning and Analysis
7	POC	Proof of Concept
8	MAPE	Mean Absolute Percentage Error

List of Figures

List of Tables

Abstract

In virtual world, every decision they create, executives today consider some quite forecast. Sound forecasting of demands and variations are not any longer extravagance items, but a necessity, since organisations have to deal with the seasonality, sudden changes in capacity management, cost-cutting strategies of the competition, and enormous swipes of the economy. Forecasting can help them affect these troubles; but it can help them more, the more they realize the overall principles of forecasting, what it can and can't do for them currently, and which techniques are suited to their needs of the instant.

Faced with extraordinary pressure to supply better outcomes at a lower cost, objective is to develop a sustainable model for cost reductions exceeding 20% across the system, along with improving quality, capacity management and outcomes. With labour costs typically comprising up to 60% of total expenses, the simplest way to scale back operating costs is to optimize forecasting and capacity planning. Whether there's a variation, decline or spike in volumes, optimum staffing is critical to operational performance. However, the goal is to sustainably allocate staff based on the forecasted volumes.

This **Forecasting tool** is a incessant planning tool that enables operations to consistently forecast incoming volume for scheduling/rostering. A combination of past process specific data, algorithmic forecasting, SME (Subject Matter Expert) inputs and modeling results in forecast with an absolute daily accuracy of up to 85% per month out and up to 95-98% one week out. The tool leverages the generated forecast to predict capacity and staffing requirements. This will result in more accurate volume forecasting and capacity planning, cost efficiency and increased client satisfaction.

Contents

Chapter 1: Introduction

In an organisation, majority of the deals lack ability to gauge the process variation caused due to seasonality /cyclical components present in the process, as there are no accurate volumetric details received from the client. This has led to inaccurate capacity estimation causing longer cycle time/ breach in SLAs and eventually in client dissatisfaction. This is the common issue observed in all delivery units across all the domains like Finance and accounting, Resources, Utilities, Insurance, Healthcare.

In a current scenario, most of the deals do not receive forecast volume data from the client and operations is being capacitized to the best efforts to meet the client requirement, this is leading to inefficient resource management impacting directly ODE (Operating Deals Economic). Also, directly impacting client satisfaction.

Need of the hour is to create an end-to-end system, which not only forecasts, but can also estimate best fit capacity and create a visual dashboard for Eye on Glass analysis and send early warnings to mitigate the risk.

Chapter 2: Problem Statement

In an organisation, majority of the deals lack ability to gauge the process variation caused due to seasonality /cyclical components present in the process, as there are no accurate volumetric details received from the client. This leads to inaccurate capacity estimation causing longer cycle time/ breach in SLAs and eventually in client dissatisfaction. This is the common issue observed in all Backoffice operations delivery units irrespective of domains like Finance and accounting, Resources, Utilities, Insurance, Healthcare.

Currently, 32% of deals are at High Risk list due to ODE (Operating Deals Economics), out of which >50% is due to capacitized misalignment. Based on the recent client satisfaction analysis it is also one of the leading indicators contributing to ~15% of dissatisfied clients.

Therefore, the study is to develop an end-to-end Forecasting and Capacity Management tool which will help in more accurate forecasting daily volumes, leading to efficient capacity management.

Chapter 3: Objectives of the Study

The objective of the study is to develop an end-to-end Forecasting and Capacity Management tool which will help in more accurate volume forecasting leading to efficient capacity management. This will be improving ODE (Operating Deal Economics) and eventually help in reducing client noise caused due to capacity misalignment. Expected target is to reduce current client target from 15% to 5%.

Phase 1 [Data Preparation – Prediction] → (Current Scope)

- Data Collection (Automated/Manual)
- Recommend best fit forecasting model
- Deploy tool in operations

Phase 2 [Proactive – Reporting] → (Future Scope)

- Develop Capacity Model
- Digital dashboard
- Alert mechanism

There are 500+ accounts in the organization, as a current scope, Finance and accounting processes were chosen for the tool deployement. Once the model is succeessful deployedt in this area, tool will replicated in other domains like healthcare, utilities etc.

Chapter 4: Project Methodology

Basis the industry experience, working guidelines were formulated to develop the working model.

There are roughly speaking four approaches to forecasting:

- Using a tool where advanced mathematical forecasting methods are implemented assist with data-gathering, data exploration, incorporation of causal factors and automated best fit modelling and advanced machine learning techniques and long short-term memory neural networks.
- Using a simpler but systematic approach that is executed by the forecaster.
- Using a non-systematic approach largely based on human judgement.
- Using a combination of the above three options.

Below structured approach is considered to develop Automated forecasting tool.

Forecasting Methods(*McKesson White Paper on Capacity Planning and Forecasting - Google Scholar*, n.d.)
"Historically seen, organizations have included some degree of forecasting abilities. There have been many technological advances in the forecasting techniques since years, however, most systems still produce very simplistic forecasts. The actual benefits of forecasting is when it is used over multiple time horizons (daily, weekly, quarterly, annual and real time and is integrated into the daily workflow of frontline operations, and planning is fully aligned with process improvement efforts. There are five key demand forecasting methods that together enable these benefits in a business operations environment:

- ❑ **Predictive Modeling:** Predictive modeling is the use of historical data to determine factors that might cause something to happen.
- ❑ **Algorithmic Modeling:** Averages and distributions of past activity are combined in an algorithm to supply a forecast or prediction. for instance, averaging the last 12 weeks of process-wise data, shift and day of week would be an easy algorithmic model that predicts future volume. it'd also include a trend factor if numbers are rising or falling. Understanding the accuracy of the model is important, especially if that's to be wont to

allocate resources. With statistics, it's easy to point out a high level of accuracy level of accuracy.

- **Pattern Identification.** Pattern identification is one among the building blocks of basic forecasts. Identifying trends and seasonal or day-of-week patterns in historical data may be a proven forecasting technique. Likewise, removing anomalies from past data and flattening trends in areas that are reaching capacity is all a part of the utilization of pattern identification in forecasting. Once patterns are identified, multiple patterns are often combined into a forecast. this sort of base forecast doesn't include things that haven't occurred, and it'd need further refinement to switch past activity.

- **Scenario Modeling:** Once a base forecast has been developed, it must be refined by operational resources who can identify upcoming changes which will impact the forecast. so as to form the forecast more representative of what is going to happen, global adjustments are often made to the variables. Other adjustments might only affect a date range or seasonality.

- **Preventive Visual Management (Dashboard):** By combining incoming volume (either employing a distribution or a known event) with predictive analytics about expected inflow volume, a simulation are often developed to model and assess the impact of varied scenarios and a true time dashboard are going to be generated." (*McKesson White Paper on Capacity Planning and Forecasting - Google Scholar*, n.d.)

Chapter 6: Business Understanding

In a current scenario, there are considerable number of processes in the organization where forecast volume data is not received from client and operations is being capacitized to the best effort basis to meet the client requirement. Which leads to inefficient resource management, as the processes are not sufficiently capacitized to handle variation in volume spikes.

Due to high risk involved here, there is an urgency for a tool to be developed which will not only help in predicting the FTE numbers required in a deal at process level, but also provide efficient resource management with proactive alert system in case of outliers.

Based on pre data analysis the organisation have more them 500 clients out of which 70% are Finance and accounting processes. Broadly Finance and Accounting processes are divided into four major categories, they are:

- Procure to Pay (P2P)
- Order to Cash (O2C)
- Record to Report (R2R)
- Financial Planning and Analysis (FP&A)

P2P and O2C are mainly daily and weekly transaction-based processes where 80% of the work is received and processed on daily & weekly basis. However, R2R and FPNA are majorly monthly and quarterly activities.

For this study and tool creation, we are considering transaction activities pertaining to P2P and O2C and a POC is carried out in one of the P2P subprocess i.e. Invoice processing.

Chapter 7: Data Understanding

Each process is unique in the organisation, where it was tough to identifying all of the relevant variables, figuring out how to measure them, and getting those variables into the data warehouse was the most difficult parts of establishing a successful modelling program. The relevant point here is to maintain the data warehouse over time. The models must be calibrated and reweighted on a regular basis, here the model should be calibrated at least every 2 to 3 months.

Since the nature of the processes are majorly transaction based with daily or weekly frequency data. So, the approach was to try different combinations of models, using techniques like Exponential smoothing model, Holt Winter's classification models, time series decomposition model etc. and then pick the best fit model.

The data gathering stage involves identification of what data is needed and what data is present with the process. Additionally, many patterns were observed in the available data sets and it was important to identify these patterns to select the appropriate forecasting model later.

Time Series data was provided by the Operations team involving at least 30 datapoints and this data was used to forecast the number of transactions to be received. In many areas data was not sufficient and accurate to provide any insight.

Data collection plan was laid down to collect the data and further cleanse it to be captured in the data warehouse.

Chapter 8: Data Preparation

Each process in the organisation is unique and identifying the relevant data to be captured and figuring out how to measure them is the toughest part of developing a successful model. Data collected should be maintained over the period of time. As a best practice, the models must be calibrated and reweighted at a regular frequency interval. In the current scenario, the model is required to be calibrated at least every 2 to 3 months' time.

Data collection and understanding

- Design the data collection warehouse to support the modelling.
- Collecting the data based on the data collection plan designed as a part of the project.
- Data cleansing- required to cleanse the past data and capture it into the newly designed data warehouse
- Data capturing and cleansing on a continuous basis.

Cleansed, specific and accurate data is mandatory for successful modelling. The data should be process specific not the company as a whole. To run the model, at least 30 datapoints are required from the past data to run the model.

Operations team provided Time Series data involving at least 30 datapoints which was required to generate forecast volume i.e. transactions to be received on a given day/week).

Note: Only inflow volume is required to be considered to run the model and not the processed volume.

Chapter 9: Data Modeling

Forecasting Tool is required to works without any assumptions or knowledge of the data's characteristics. Different combinations of models were tried, using techniques like **Pegel's classification, Holt Winter's classification models** to pick the winners.

"Outcome of the analysis showed how the different methods of obtaining initial values can affect the accuracy of the forecast generated. Post parameter optimization, the differences were reduced to minimal." (Forecasting & 2010, n.d.; Garen, 1992; *Methods to Perform Time Series Forecasting Gurchetan Singh - Google Scholar*, n.d.; Singh, n.d.; Taylor, 2010)

Framework – End to End Forecasting and Capacity Model

Data Preparation **Predictive Modelling** **Proactive Analytics** **Proactive Alerts & Visual managament**

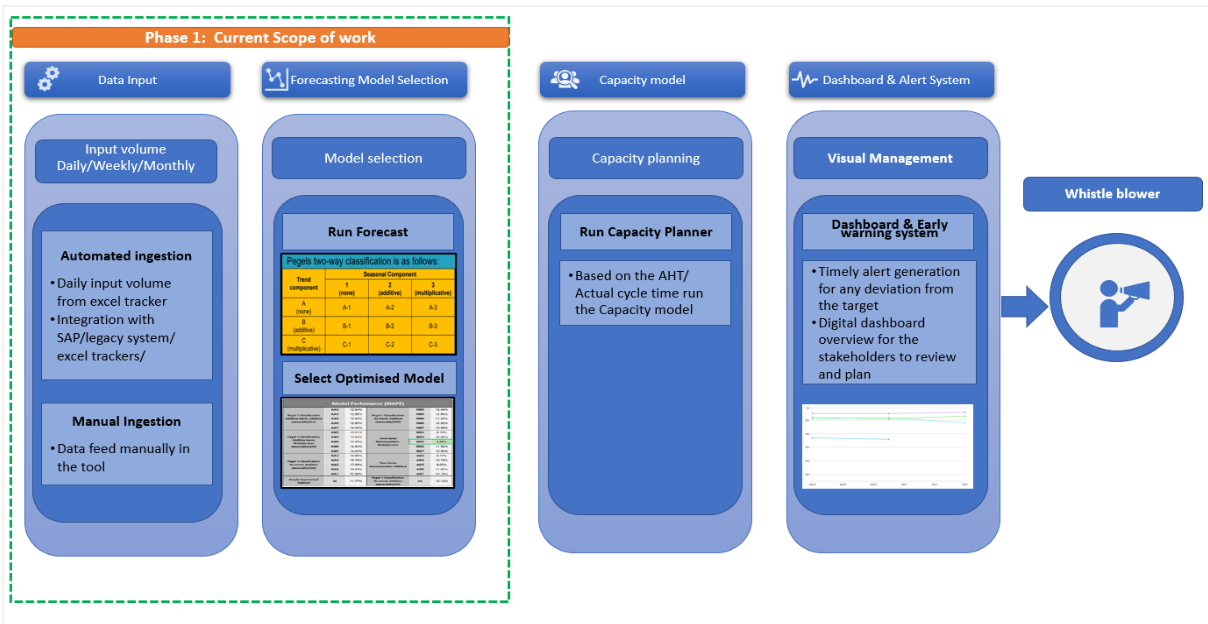


Figure1: Framework- End to End Forecasting and Capacity Model

"A method was required that can map the variation and trend accurately without any assumptions in such a way that undertakes the consideration of the trend of the dataset which is named Holt's Linear Trend method. Each Time series dataset have three components which are Trend, Seasonality and Residual. The dataset which shows a trend can use Holt's linear trend method for forecasting. Holt extended simple exponential smoothing to permit forecasting of knowledge with a trend. It is nothing quite exponential smoothing applied to both level (the average value within the series) and trend.

Owing to the seasonality factor of the data, using Holt's winter method will be the best option among the rest of the models. The Holt-Winter's seasonal method comprises the forecast equation and three smoothing equations — one for the extent ℓt, one for trend bt and one for the seasonal component st, with smoothing parameters α, β and γ". ." (Forecasting & 2010, n.d.; Garen, 1992; *Methods to Perform Time Series Forecasting Gurchetan Singh - Google Scholar*, n.d.; Singh, n.d.; Taylor, 2010)

$$
\begin{aligned}
\text{level} \quad L_t &= \alpha(y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1}); \\
\text{trend} \quad b_t &= \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}, \\
\text{seasonal} \quad S_t &= \gamma(y_t - L_t) + (1-\gamma)S_{t-s} \\
\text{forecast} \; F_{t+k} &= L_t + kb_t + S_{t+k-s},
\end{aligned}
$$

"The level equation shows a weighted average between the seasonally adjusted observation and therefore the non-seasonal forecast for time. The trend equation is identical to Holt's linear method. The seasonal equation shows a weighted average between the present seasonal index, of an equivalent season last year i.e. "s" time periods ago." (Forecasting & 2010, n.d.; Garen, 1992; *Methods to Perform Time Series Forecasting Gurchetan Singh - Google Scholar*, n.d.; Singh, n.d.; Taylor, 2010)

Chapter 10: Data Evaluation

Additive and multiplicative methods were used in the tool:

- The additive method is generally preferred when seasonal variations are roughly constant throughout the series

- Multiplicative method is used when seasonal variation is changes proportionally to the level of the series.

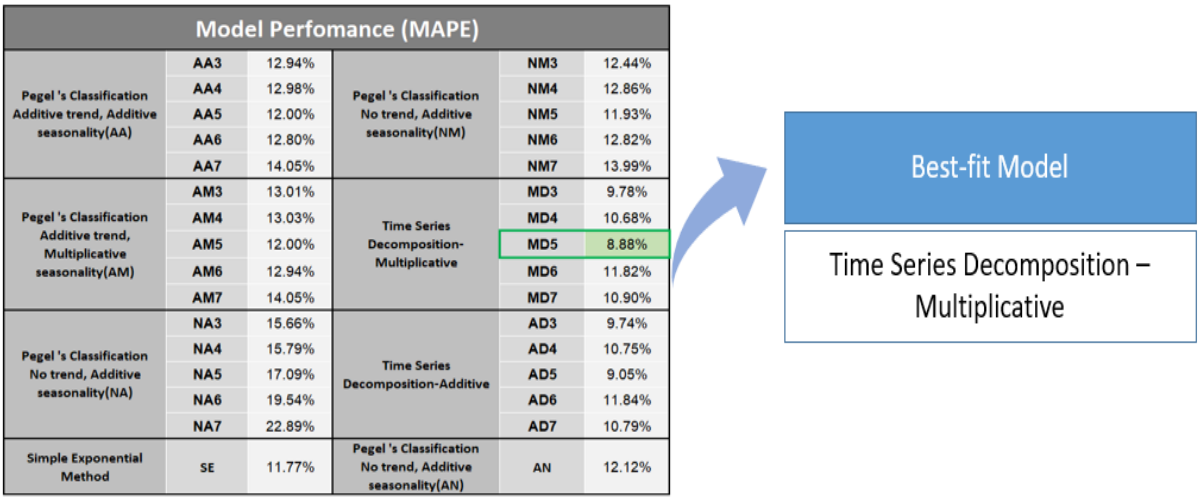
- During the data analysis it was observed that how the use of one method of obtaining initial values can affect the forecast accuracy. After optimizing the values of the parameters these differences can be minimized and clearly shows their effect on accuracy. As a result, parameters might not be optimal with initial values, which complicates optimization on model in the adjustment and eventually the forecast.

- In the draft model MAPE score was analyzed as a response variable with different initialization methods of smoothing equations to impact the response variable. The results show the impact of the methods of obtaining initial values is greater on multiple season Holt-Winters models with an additive trend.

Chapter 11: Analysis and Results for different Models

In the tool, data is captured (Min 30 data points / transaction) and model is run. The tool generates MAPE score and the user can pick up the model from the MAPE score. Model with the lowest MAPE score is used to forecast the capacity and volume.

Note*: Each model behaves differently for different set of data. Hence, the model which performs well on one set of data might not work for different set of data.*



Model Perfomance (MAPE)					
Pegel 's Classification Additive trend, Additive seasonality(AA)	AA3	12.94%	Pegel 's Classification No trend, Additive seasonality(NM)	NM3	12.44%
	AA4	12.98%		NM4	12.86%
	AA5	12.00%		NM5	11.93%
	AA6	12.80%		NM6	12.82%
	AA7	14.05%		NM7	13.99%
Pegel 's Classification Additive trend, Multiplicative seasonality(AM)	AM3	13.01%	Time Series Decomposition-Multiplicative	MD3	9.78%
	AM4	13.03%		MD4	10.68%
	AM5	12.00%		MD5	8.88%
	AM6	12.94%		MD6	11.82%
	AM7	14.05%		MD7	10.90%
Pegel 's Classification No trend, Additive seasonality(NA)	NA3	15.66%	Time Series Decomposition-Additive	AD3	9.74%
	NA4	15.79%		AD4	10.75%
	NA5	17.09%		AD5	9.05%
	NA6	19.54%		AD6	11.84%
	NA7	22.89%		AD7	10.79%
Simple Exponential Method	SE	11.77%	Pegel 's Classification No trend, Additive seasonality(AN)	AN	12.12%

Figure 2: MAPE score

Chapter 12: Conclusions and Recommendations for future work

In the absence of accurate forecasts, it is challenging for the operations to optimized number of resources. By proactively planning for the variations, operations can plan to support volume spikes with the limited resources and can plan accordingly to cater to the change in requirement. Forecast generated by the tool (proactive & highlighting any sudden volume spikes), operations teams can work proactively to reducing potential delay in SLA execution and avoid supplementary costs associated with the inaccurate resource alignment.

After successful deployment of the model in Finance and accounting process, the model will be deployed in other domains like Utilities, healthcare etc.

Appendix – Graphical Output

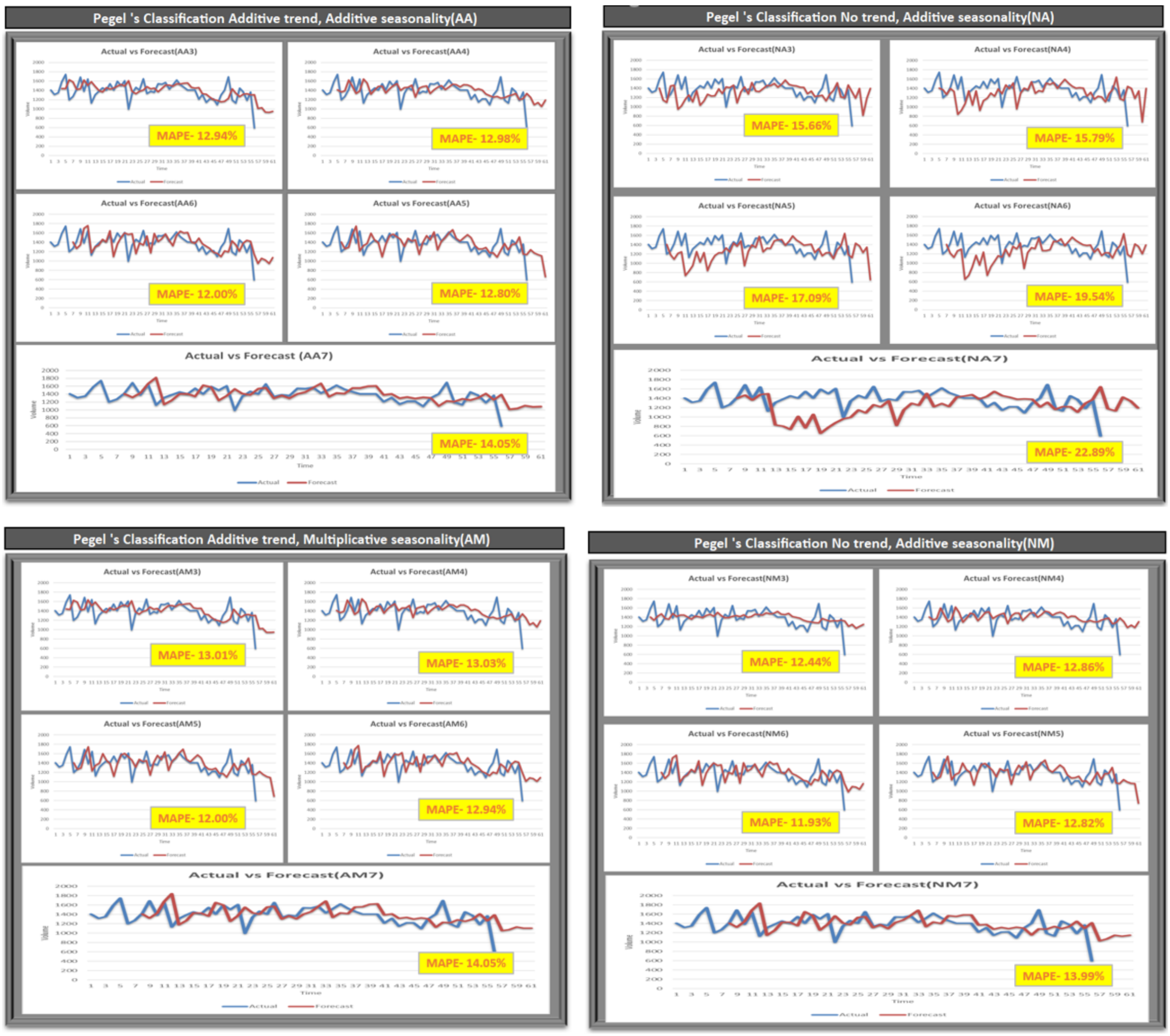
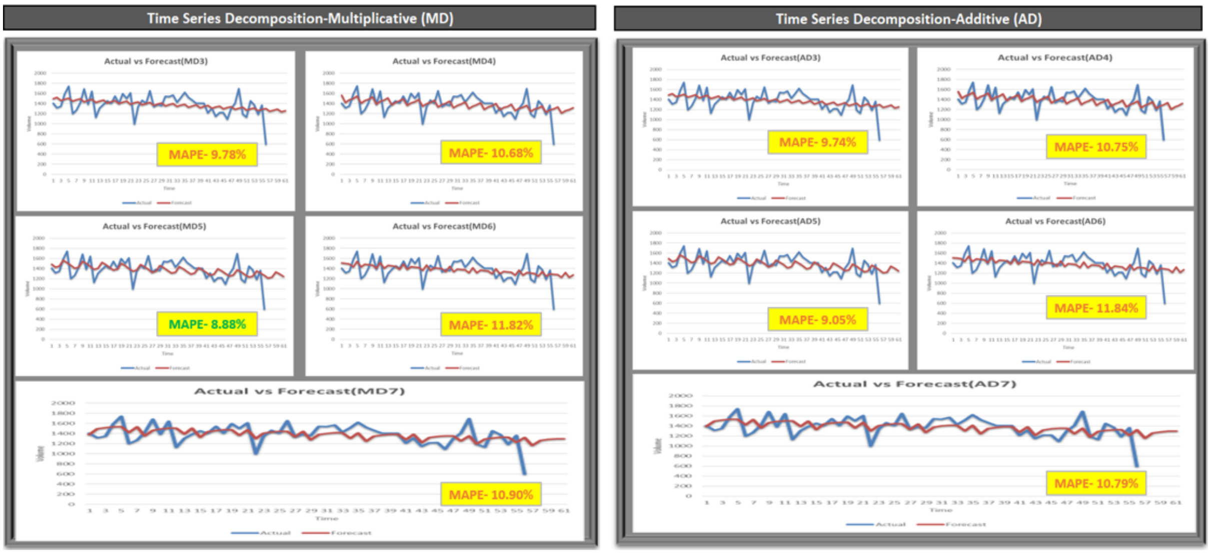
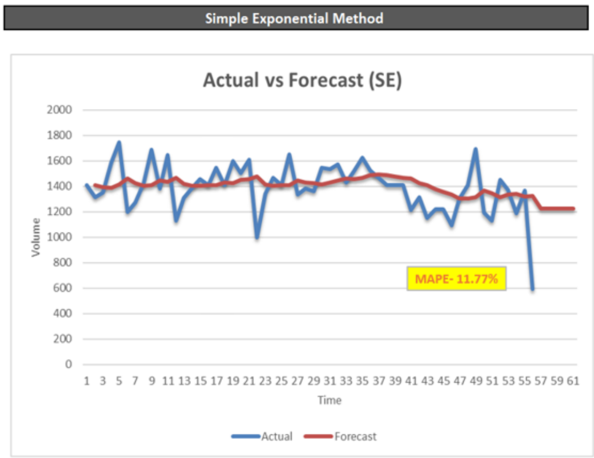


Figure 3: Graphical output





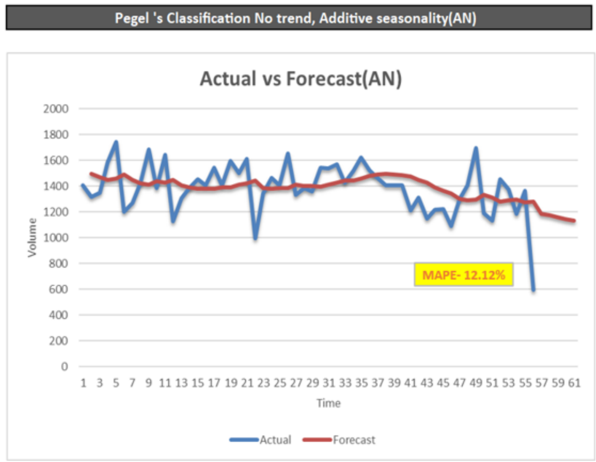


Figure 4: Graphical Output

Appendix – Similarity Report

Automated Forecasting & Capacity Solution for Back-office Operations

by Anuraag Anand

Submission date: 30-Oct-2020 05:21PM (UTC+0530)
Submission ID: 1431144924
File name: Forecasting_Model_Capstone_Project_Report_Format.docx (1.35M)
Word count: 3412
Character count: 19172

Automated Forecasting & Capacity Solution for Back-office Operations

ORIGINALITY REPORT

21%	18%	4%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.longwoods.com Internet Source	7%
2	www.analyticsvidhya.com Internet Source	2%
3	Oscar Trull, Juan Carlos García-Díaz, Alicia Troncoso. "Initialization Methods for Multiple Seasonal Holt–Winters Forecasting Models", Mathematics, 2020 Publication	2%
4	www.broadwaymerchandiseshop.com Internet Source	2%
5	Submitted to Sogang University Student Paper	2%
6	docplayer.net Internet Source	2%
7	www.termpaperwarehouse.com Internet Source	1%
	getsix.eu	

8	Internet Source		1%
9	Ramazan Ünlü, Ersin Namlı. "Machine Learning nd Classical Forecasting Methods Based Decision Support Systems for COVID-19", Computers, Materials & Continua, 2020	Publication	1%
10	Submitted to Regional Centre for Biotechnology	Student Paper	1%
11	jacobbradjohnson.com	Internet Source	1%
12	www.inmybangalore.com	Internet Source	<1%
13	Submitted to uva	Student Paper	<1%
14	Submitted to Visvesvaraya Technological University, Belagavi	Student Paper	<1%

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		

Bibliography

Forecasting, H. S.-I. J. of, & 2010, undefined. (n.d.). Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles: Comments. *Ideas.Repec.Org*. Retrieved October 18, 2020, from https://ideas.repec.org/a/eee/intfor/v26yi4p652-654.html

Garen, D. C. (1992). Improved techniques in regression-based streamflow volume forecasting. *Journal of Water Resources Planning and Management*, *118*(6), 654–670. https://doi.org/10.1061/(ASCE)0733-9496(1992)118:6(654)

McKesson White paper on Capacity Planning and Forecasting - Google Scholar. (n.d.). Retrieved October 18, 2020, from https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&q=McKesson+White+paper+on+Capacity+Planning+and+Forecasting&btnG=

Methods to perform time series forecasting gurchetan singh - Google Scholar. (n.d.). Retrieved October 18, 2020, from https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&q=Methods+to+perform+time+series+forecasting+gurchetan+singh&btnG=

Singh, G. (n.d.). *methods to perform Time Series forecasting (with Python codes).*

Taylor, J. W. (2010). Exponentially Weighted Methods for Forecasting Intraday Time Series with Multiple Seasonal Cycles. In *International Journal of Forecasting* (Vol. 26). https://www.sciencedirect.com/science/article/pii/S0169207010000464