CITY UNIVERSITY
LONDON

Individual Project Report

Table of Contents

Abstract

This report presents an extensive analysis of sales forecasting methods applied to retails datasets. Through the use of a range of predictive models such as ARIMA, SARIMA, and Prophet, complex dynamics of sales forecasting were explored. Each model was chosen and adjusted for the purpose of accurately predicting future sales performance. In this study Multiple datasets were used focusing on aggregate sales and top-performing store/departments and forecasting was improved by integrating outside variables which include holidays and promotions. To guarantee the reliability of the models, thorough model diagnostics, cross-validation, and error analysis were included. This analysis provided insightful information about predictive analytics models, which highlighted their usefulness in retail sales forecasting. Furthermore, it does not not only aid in the understanding of time series analysis, but also demonstrates how different approaches can be used to address real world business problems.

Cover Sheet

Degree and Program:

Data Science (MSci)

Project Title:

Python Trading Bot for Decision-Making in Financial Markets

Name and email:

Youssef Abdelmoamen

Youssef.abdelmoamen@city.ac.uk

Consultant:

Dr Jo Wood

Who proposed the project:

Myself

Word Count:

11,523 words

Chapter 1: *Introduction*

In the current state of the continually changing retail industry, the capability of accurately predicting future sales is not only essential to gain a sustainable competitive advantage but to grow at all. The specific project was designed to forecast sales with the help of predictive analytics, having the following goal: to solve a major problem faced by many retailers - that is, to predict sales as accurately as possible. Based on historical sales information for 45 different stores, collected via an extensive dataset on Kaggle, this analysis was aimed at enhancing sales predictions significantly. The use of advanced forecasting tools - ARIMA, SARIMA, and Prophet - made it possible to improve sales predictions in a variety of ways.

1.1 Problem Statement

The primary inspiration behind this project was the immediate need of organizations for better decision-making when it comes to resource and inventory allocation and decision-making at large. Since reported forecasts were almost never accurate and highly inflexible towards quickly changing market factors, an innovative strategy to approach this situation was needed. Therefore, the necessity identified the main objective of this project, which is providing organizations with reliable and accurate forecasts of projected sales volumes in the future with a novel approach to predictive analytics.

1.2 Objectives and Sub-objectives

The sub-objectives of this project were outlined to aid in achieving the main objective (to create a sophisticated predictive analytics model for sales forecasting, these included comprehensive data collection to ensure a strong foundation for the analysis, data cleaning to make it suitable and detailed exploratory data analysis to understand the patterns, trends and outliers, and developing multiple time-series forecasting models using multiple methodologies to understand which accurately predicts the future sales. In addition, external factors were integrated like holidays and promotions to improve predictive accuracy and interactive visualizations were used to give sales projections and insights that would be insightful to any apparent stakeholders. Finally, each model was evaluated and fine tuned to ensure top performance and to aid in comparing it to other models.

1.3 Anticipated Beneficiaries

A wide range of stakeholders within the corporate ecosystem stand to benefit, both internal and external entities will gain from this project each of which will derive value in a unique way. The internal stakeholders include executive management, operations and supply chain teams, sales and marketing teams, and the finance department. Each of them will benefit the following: -

- **Executive management**: Executives and business leaders will learn important information about potential sales trends in the future which will facilitate knowledgeable decision-making, enabling them to make the correct decisions and execute successful expansion and resource optimization plans.
- **Operations and Supply chain teams**: More efficient resource allocation will help the teams working in operations and supply chain. Predicting future sales accurately should enable effective inventory control and help in avoiding stock outs or surplus inventory.
- **Sales and Marketing teams**: To efficiently reach the projected sales, these teams will have to find high-demand products, improve consumer engagement and optimize marketing efforts.
- **Finance department:** By precisely forecasting future sales, it should support this department in cost control, revenue planning and financial stability by showing accurate estimates of financial estimates.

On the other hand, there are external beneficiaries that will also benefit from this approach which include the customers, shareholders and investors through higher product availability, Improved financial performance, and strengthened operational alignment.

1.4 Work Plan

The work plan selected for the analysis was comprehensive, well-structured, and clear. The steps included the literature review with the analysis of time-series forecasting and predictive analytics as the basic topics, data collection, data cleaning, exploratory data analysis (EDA), stationarity testing, identification of seasonality, model development, an integration of external factors, evaluation of the models, and implementation of interactive visualizations. Every stage could be seen as a continuation of the previous one, and it was possible to follow a logical and systematic course to complete the project's objectives.

1.5 Assumptions

Throughout this project, multiple assumptions were made which included the stability of market conditions and the relevance of the selected external factors to sales performance.Within the limitations of the available data and the selected analytical tools, the scope was carefully restricted to concentrate on providing actionable insights.

Chapter 2: *Output Summary*

The work of this project, aimed at advancing the methodologies of sales forecasting in the retail field, has resulted in a successful and all-inclusive output, which is a Jupyter Notebook compilation. This output is a single yet composite representation of the advanced predictive models combined with application in the field of retail analytics.

2.1 Jupyter Notebook:Integrated Sales Forecasting system

2.1.1 Overview Description

The output of this project, transformed into a Jupyter Notebook, represents a single, all-inclusive platform that contains all of the advanced sales forecasting models, such as ARIMA, SARIMA, and Prophet for two datasets which were aggregate sales and the top_performing store. Moreover, the interactive nature of this notebook does not limit it to be merely a compendium of code but represents an educational tool for users, which guides them through and shows the retail sales forecasts models.

2.1.2 Type of Output and Size

This output represents a Jupyter Notebook that is embedded in a structured document of annotations, estimating models, visualizations, and analysis conclusions. The notebook offers the combined value of code, visual output, and the narrative text of the process of sales forecasting.

2.1.3 Intended Recipients and End-Users

The audience for this notebook are data scientists or business analysts working in the retail field who aim to predict the amounts of sales; forecast inventory planning or market approach effect on their sales among optimized decision-making using the generated retail sales knowledge.

2.1.4 Usage and Benefits

The end-users can implement the results of this notebook compilation to enhance the business decision-making process by creating more accurate sales forecasts. In addition, the commentary, and annotations, along with the visualizations, provide them to comprehend which model to choose for a sales forecast and the probable model's outcomes. This can help retailers to understand their market needs, provide resources accordingly, and organize the operations based on the new understanding.

2.1.5 Links to Results and Appendices

The work conducted, the forecasting models comparison, and the results themselves are provided in Chapter 5: Results. Also, multiple important screenshots from the full Jupyter notebook will be included in the appendix.

Chapter 3: *Literature Review*

The dynamic nature of the market often makes traditional forecasting methods unsuitable, resulting in poor operational and strategic outcomes. Incorrect sales forecasting is one of the main reasons for inadequate inventory management, resource allocation and competitiveness in business that can be averted by accurate predictions The current project is aimed at creating an advanced predictive analytics based model that will provide for a much more sophisticated approach to sales forecasting than any of the existing models.

3.1 Current Sales Forecasting Methods

Traditional methods to predict future sales go from basic moving averages to exponential smoothing. These methods have been key to predicting in retail stores and other areas, but they depend too much on historical sales numbers and assume that trends will always be linear (Hyndman & Athanasopoulos, 2018). To beat these weak spots, statistical models like ARIMA and SARIMA came up, using patterns of data over time to see what sales will do next (Box & Jenkins, 1976). These models have made predictions more accurate, but they sometimes miss sudden changes in the market and external factors that change how sales move.

3.2 Shortcomings of Existing Approaches

Literature often shows that multiple methods to guess the predict future sales in markets cannot adapt to changing market conditions (Lawrence, Edmundson, & O'Connor, 1985). Wrong predictions about sales can cause too much stock,stock outs, or waste of resources. This costs a lot and may slow down a business when it tries to meet what buyers demand. These shortcomings show why there is a need for models that can quickly include changes in the market and outside factors.

3.3 Advancements in Predictive Analytics

New steps forward in predictive analytics gives hope for fixing the problems with the traditional ways of trying to predict the future sales. Using machine learning has made a large difference in how well we can predict the future (Carbonneau, Laframboise, & Vahidov, 2008). Algorithms such as the Prophet, which deal with repeating trends, promotions and special days, provide more detailed and nuanced guesses (Taylor & Letham, 2018). Adding in outside factors like economic indicators and promotional calendars to the predictive model has shown to make usual guesses even more accurate (Shi et al., 2009).

3.4 Proposed Methodology

This project utilized a retail data analytics dataset from Kaggle, featuring detailed historical sales information from 45 different stores. This dataset offers various characteristics that are helpful for building predictive models (Kaggle, 2020). Different techniques such as Prophet, SARIMA, and ARIMA will be employed in developing a time series forecasting model, all of which have been proven to handle intricate sales patterns effectively according to research by Hyndman & Khandakar (2008). The model's performance will be assessed using metrics like

MAE, MSE, and RMSE, which are widely respected for their ability to measure forecast accuracy (Willmott & Matsuura, 2005).

3.5 Visualizations and Decision Support

The significance of interactive visualizations is in their ability to convert intricate data into valuable visualizations for the stakeholders (Few, 2009). This will be possible by employing such instruments as Jupyter notebooks to generate dynamic simulations that can give projections on sales in a way that is easy and meaningful enough in turn supporting decision making based on data all over the organization.

3.6 Impact on Stakeholders

Various stakeholders can feel the wide-ranging impact of refining sales forecasting models. Accurate forecasting allows better resource optimization and strategic planning for internal beneficiaries, such as the executive management and operations team (Makridakis, Wheelwright & Hyndman 1998). On the other hand, improved forecasting could help in improving levels of service standards for customers and suppliers through enhanced coordination in the supply chain (Christopher 2016). Correct forecasts also support financial stability and growth opportunities; thus benefiting shareholders and investors (Fama & French, 2004).

3.7 Defining Model Accuracy

A model is said to be accurate if it produces forecasts very near the actual outcomes. The accuracy can be said to be very subjective in relation to the context in question but generally refers to the level of both precision and recall of the forecasted data (Fildes & Goodwin, 2007). Precision ensures that the predictions are very close to the actual truth. In contrast, recall is the quantity of a model's capacity to capture all relevant instances within the data without leaving them out. Usually, using historical data whereby the actual values are known and compared with the predictions made by a model (Armstrong and Collopy, 1992).

3.8 Comparison and Contrast of Forecasting Models

All the different forecasting models each have their pluses and minuses. For example, the ARIMA model is very flexible, of great generality, and has many applications in time series data for many other types. Much praise, however, with failings when dealing with the data of seasonality or some extra adjustments for non-stationary (Hyndman & Khandakar, 2008). On the other hand, SARIMA models generalize the ARIMA models with the incorporation of seasonality. As such, SARIMA becomes quite valuable for data sets that clearly show patterns within the seasons (De Gooijer & Hyndman, 2006).

On the other hand, the Prophet model has been designed to be capable of dealing with the irregularity of the business time series that holds strong seasonality effects and double seasonality. However, sometimes its black-box nature can be applied to hide the rationale of the forecast (Taylor & Letham, 2018).

3.9 Understanding Performance Metrics

Performance metrics are quantifiable measures to determine the accuracy of forecasting models. Standard metrics considered in this regard include the Mean Absolute Error (MAE), among others, being the average magnitude of errors in a set of forecasts without regard to their direction (Willmott & Matsuura, 2005). Another popular metric is the root mean squared error (RMSE), which provides more weight to more significant errors and is, therefore, potentially more useful when significant errors are very undesirable (Hyndman & Koehler, 2006). As a mean absolute percentage error, the measure (MAPE) is humanly valuable because its relative error is much more communicable and understandable even by non-experts (Lewis, 1982). That is why in this analysis MAPE is the main performance metric used for evaluation.

3.10 Model Selection and Its Impact on Project Direction

Thus, the nature of the dataset, specific goals of the forecasting exercise, and trade-offs one might be willing to accept among bias and variance, complexity and interpretability, precision and computational cost, will somewhat guide the choice of the model for any forecasting project.These insights from the literature make it clear that the approach is to be taken into consideration; statistical metrics in moderation, while giving practical utility and interpretability of the model importance.The current project takes these considerations on board, integrating a comparative analysis between traditional and advanced models, weighing pros and cons to come to a selection combining relative simplicity with a high level of predictability so that they are likely to deliver forecasts that are both accurate and usable within the context of retail sales.

3.11 Relevance of the Literature to the Present Project

Furthermore, this gamut of literature about methods for forecasting provides an exciting and intriguing background against which the drafted methodology for the project can be seen in light of the limitations and potential limitations brought out in the literature. Thus, the project's methodology, crafted against the identified limitations and probable limitations of the literature, synthesized the strengths of the different models while ensuring that their weaknesses were mitigated. This should create a rational approach and lead to the development of a robust model that aligns with the actual intricacies of making retail sales forecasts. The literature review, therefore, sets the stage for an academically rigorous and practically applicable forecasting model. It served as the theoretical base on which the project methodology was built, to continue contributing to the field of predictive analytics by closing gaps that are found in current forecasting practices and using the latest advancements made in this field.

Chapter 4: *Method*

4.1 Project Overview and Planning

The foundation of this project was in identifying a pressing issue in the retail sector – the need to improve the accuracy of sales forecasting. Specifically, the primary goal was to explore possibilities outside traditional forecasting methods, which have shown limitations in handling sudden market changes and the impact of external variables. This endeavor involved the application and evaluation of advanced predictive models, including ARIMA, SARIMA, and Prophet, utilizing a Kaggle dataset that provided historical sales data across various stores.

Given the project's exploratory nature, a flexible and dynamic project management approach was crucial. Though not strictly adhering to corporate Agile methodologies, the project embodied an agile-like philosophy, focusing on adaptability and responsiveness to new insights. Initial steps included setting achievable goals centered around the development and testing of forecasting models, and importantly, integrating external variables such as promotions and economic indicators to enhance forecasting accuracy.

The project proceeded through several iterations, incorporating data collection and insights from existing literature to refine and innovate the forecasting approach. A significant shift from the original plan involved a greater focus on the role of external variables, underscored by early results and relevant studies.

Challenges in data preprocessing and model optimization were met, particularly in cleaning the dataset to accurately reflect the sales dynamics targeted for modeling. Regular self-evaluations ensured alignment with objectives and facilitated necessary adjustments, emphasizing a self-guided, adaptable management style crucial for navigating the project's complexities.

Reflecting on the project, the flexible and iterative methodology proved instrumental in achieving success, allowing for project evolution in sync with emerging discoveries and methodological challenges. This approach ensured the fulfillment of objectives, laying a solid foundation for advancing retail sales forecasting with sophisticated predictive modeling. Through careful planning and a results-oriented execution strategy, significant progress was made towards providing retail businesses with more precise and actionable sales forecasts.

4.2 Data Collection and Preprocessing

4.2.1 Data Sourcing

This project was built using a robust dataset obtained from kaggle, a known website used for data science projects and competitions.This dataset contains historical sales data across 45 stores and is divided into 3 datasets which include a features dataset, a sales dataset, and a stores dataset. The features dataset contains a store number column, a date column, and other external factors like temperature, fuel price, different markdowns, CPI, unemployment rate and whether the week is a holiday. Moreover, the sales dataset also contains the store number, department number, a date column, a column that contains the weekly sales of the given department and a column that shows if that week was a holiday or not. Finally, the stores dataset contains all the details of the stores which are the store number, the type of store ,and the size. Each store has around 90 departments and it was detailed over around 3 years. The comprehensiveness of the dataset allowed for an in-depth exploration of sales forecasting with the retail industry, providing a strong base for applying and evaluating multiple predictive models.

4.2.2 Preprocessing Objectives

To apply advanced forecasting models effectively, the first task was to ensure the readiness of the dataset by applying data-preprocessing which mainly focused on data cleaning and transformation. The objectives were to refine the dataset into two clean, well-structured datasets ready for time-series forecasting. The data cleaning steps consisted of checking for missing values, checking the data types and creating a function to detect outliers.

4.2.3 Cleaning and Structuring

The first step of preprocessing was data cleaning, which was an important step in removing inaccuracies and inconsistencies. This consisted of checking for missing values ,which were present in a big portion of the promotional markdown columns in the features dataset and also economic indicators like CPI and unemployment rate. Imputation was made based on the nature of missing data; all missing values in these columns were replaced with a value of 0.

The second step was checking the data types of each of the columns and in each dataset and whatever returns an ‘object’ data type is transformed into the right data type like in the date column in all the datasets had to be transformed to date-time format and also the store type had to be converted into ‘category’.

Another main step in the data cleaning process was the removal of outliers. Retail sales in general usually have seasonality which causes extreme values due to seasonal peaks or promotions. Sometimes removing these outliers could affect the models performance or accuracy in predicting the future values, so creating a function to remove outliers was the best option to check if later on it will improve the accuracy of the model.

The dataset was also restructured to facilitate time-series analysis. This involved aggregating sales data to a weekly level across all stores and departments, ensuring a uniform time scale for forecasting. Additionally, the data was segmented into separate series for aggregate sales analysis and the top-performing department, enabling a focused approach to model application and evaluation.

4.2.4 Reflections on Preprocessing

This preprocessing phase was iterative and reflective where each step was evaluated on how it would impact the dataset's suitability for forecasting. In doing that, this ensured that the data would retain its integrity and representativeness of sales dynamics around the world and also ensured the alignment of the data with the requirements of the predictive models. Including external factors in the following sections in the analysis will aim to transcend the traditional methods.

In conclusion, the data collection and preprocessing phase laid a solid foundation for the subsequent modeling efforts. Through diligent cleaning, structuring, and enrichment of the dataset, this phase encapsulated the project's commitment to leveraging data-driven insights for advancing retail sales forecasting. The processed dataset, now ready for sophisticated time-series analysis, promised a fertile ground for exploring the potential of ARIMA, SARIMA, and Prophet models in capturing and predicting the complex patterns of retail sales.

4.3 Model Selection Rationale

The choice of models was narrowed down between three main models which were ARIMA (AutoRegressive Integrated Moving Average), SARIMA (Seasonal AutoRegressive Integrated Moving Average) and prophet models. These models were chosen based on the literature review and an evaluation of the specific needs and objectives of retail sales prediction. The main reasons these models were chosen is due to their proficiency in handling time-series data, their ability to incorporate seasonality, and their ability to adapt to diverse forecasting scenarios. Each model will be used twice for two datasets , aggregate weekly sales of all the stores and departments and the top performing store, for separate reasons. The aggregate weekly sales analysis will be used to understand overall trends and set a benchmark for forecasting models, while the top performing store/departments will be used to focus on the segment contributing most to sales which can refine a model's accuracy and usefulness, Also Since the time series could not be done as the dataset was because of the overlapping dates of the different stores and departments, this split was conduct to observe which would be the most accurate method and will yield the best accuracy.

4.3.1 ARIMA Model

The ARIMA (AutoRegressive Integrated Moving Average) model has been the traditional model used in the time-series forecasting realm, usually chosen for its effectiveness in modeling a wide range of time-series data but without the necessity for seasonal components (Box & Jenkins, 1976).Its capacity to model data that exhibits non-stationarity through differencing makes it a versatile choice for sales data, which often reflects trends and cyclicality (Hyndman & Athanasopoulos, 2018). This model was chosen due to the ability to produce short-term forecasts accurately.

4.3.2 SARIMA Model

SARIMA (Seasonal ARIMA) is the same as an ARIMA model except it integrates seasonal differencing, making it suitable for datasets where seasonality is present; In some cases retail sales data (Box & Jenkins, 1976). The retail sector being prone to seasonal fluctuations like holidays, promotions, and consumer behavior demands a model that has the capability to handle these effects accurately. The structure of the model offers modeling for both non-seasonal and seasonal data, which helps as a more nuanced approach to forecasting retail sales data. The inclusion of SARIMA was also motivated by literature highlighting its superiority over ARIMA in dealing with seasonal time-series data, which is commonplace in retail sales forecasting (De Livera, Hyndman, & Snyder, 2011).

4.3.3 Prophet Model

Developed by Taylor and Letham (2018) at Facebook, the Prophet model represents the most recent advancement in time-series forecasting, and it was designed to address the limitations of the traditional ARIMA-based approaches. It has the ability to incorporate holiday effects and other external regressors, detect change points and handle the outliers robustly which makes the most fit for the objectives of the project. The flexibility and simplicity of this model combined with its capability of producing accurate forecasts even if there is missing data and large outliers, ensured that it would be one of the models chosen for this project.This model's effectiveness in various business forecasting scenarios, as documented in the literature, underscores its suitability for predicting retail sales (Taylor & Letham, 2018).

These models were chosen not only for their individual merits but also for their complementary strengths in modeling different aspects of the sales data. The integration of ARIMA, SARIMA, and Prophet models provides a holistic approach to sales forecasting, capable of addressing the complex dynamics of retail sales with improved accuracy and reliability.

4.4 Model Development and Implementation

4.4.1 Development and Implementation of the ARIMA Model:

The ARIMA model development was initiated with the help of ACF and PACF plots in determining plausible values of hyperparameters. These plots gave insight into temporal dependencies in the sales data, which was very critical for the forming of the order of the autoregressive (p) and moving average (q) components of the ARIMA model, together with the degree of differencing (d) needed to take to obtain stationarity.

The model implementation was done in Python, using the Pandas library for data processing. The prima library applies the 'auto_arima' function in search of a model that would minimize the AIC for the best parameters. Integration of the model in the Jupyter Notebook was also influential since it enhanced the testing and validation of forecasts in an interactive and iterative approach. Some of the significant influences on sales patterns—promotional events and holiday periods—were exogenous variables of the model. In this way, it included the model with these exogenous variables: the promotional events and the holiday periods—events of significant influence on the sales pattern. These external indicators were included in the model as one more step to tune forecast accuracy and provide a more realistic projection of sales figures during such impactful periods.

4.4.2 Development and Implementation of the SARIMA model:

The seasonal and non-seasonal ingredients of the development of this model SARIMA were incorporated together. With the 'seasonal' parameter set to true, the ‘auto_arima’ function automatically obtained the best combination of seasonal (P, D, Q, m) and non-season seasonal (p, d, q) hyperparameters to allow the model to capture the intrinsic seasonality of the sales data. The implementation was carried out in the Python programming environment, using the same libraries as applied in the ARIMA model. Special attention was drawn to the parameter 'm', representing the seasonal periodicity, which had great importance to the cyclical nature of retail sales data related to regular seasonal fluctuations.

The model includes an opportunity of external factors, with the reference to seasonal dummies and specific event indicators. Such incorporation in the model allows the forecasts to be adjusted ahead for known variations in selling activity.

4.4.3 Development and Implementation of the Prophet Model:

Retail sales data usually describes a flexible growth curve that may model non-linear trends within the sales data and further capture weekly and yearly seasonality that led to the development of the Prophet model. Flexible time series data management with high seasonality effects has made Facebook's Prophet library an ideal choice for this analysis. The same was implemented in Python, where, in this case, the Prophet library provided convenient functions for modeling and forecasting. Special events and holidays were explicitly included in the model through a supplemental data frame so that these cases are considered by the model whenever they are likely to cause spikes or drops in sales. It was, therefore, important that those determinants are included in the forecast, and not just to improve its accuracy, to ensure that the model is adaptive to an ever-changing retail environment. The Prophet model explicitly modeled such events and resulted in forecasts that could adapt to the ebb and flow of retail demand patterns.

4.4.4 Code and Materials Reuse:

Throughout the project, minimum code reuse of individual lines of code and functions was practiced, ensuring that the produced code came with originality and was tailor-made to only specific sales forecasting task requirements. Nevertheless, standard libraries and functions, mostly recognized and validated by the strong data science community, have been used to ensure result reliability. In fact, the methodological approach towards developing and implementing models was systematic, iterative, based on established statistical practices, and very much augmented by the power of modern computational tools. The detailed account given in this series reflects the careful detailed planning and very thoughtful consideration applied at each stage of the project life cycle.

4.5 Evaluation Metrics and Model Testing

The developed predictive models for this project have been tested within a rigid evaluation framework, leveraging widely recognized statistical metrics. Identification of the evaluation metrics was the major role played in the assessment of the performance of ARIMA, SARIMA, and Prophet models developed. Major primary metrics used in this system include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and mainly Mean Absolute Percentage Error (MAPE).

The Mean Absolute Error (MAE) measures the error between paired observations presenting the same phenomenon. It sums up the absolute difference between a real value and a predicted value, providing a direct mean of these absolute differences. This is why MAE is used; it shows the average prediction error in the same units as the data, and thus its magnitude is clearly and directly interpretable (Willmott & Matsuura, 2005).

Mean Squared Error (MSE) is another measure; it is simply the sum of the squared differences between estimated values and the actual value, then averaged. MSE is of great help, for it exaggerates large errors that may be more harmful to the forecast. Squaring the errors makes sure that the model deviations are being taken in a robust way, so MSE also becomes an unsurprising measure of good quality.

Root mean square error (RMSE) extends the mean square error (MSE) through its square root, scaling the error to be in the same unit as that of the data. RMSE was used because it is sensitive to larger errors, offering conservative variance. RMSE is particularly valuable: "RMSE is especially useful (Chai and Draxler, 2014) because, with dimensional compatibility with the output variable, it communicates to the stakeholders in a way they will find comprehensible to the extent to which there is an error."

MAPE (Mean Absolute Percentage Error) is a special case in the performance rating of forecasting models. This is mainly so because it provides a normalized measure of accuracy through which models with varying scales can be compared directly. This metric gives an average absolute percentage error between the forecasted and actual values. It comes in handy for cases in which the magnitude of comparative error is more insightful than the value of absolute error like with this analysis.The MAPE is calculated as the average of the absolute differences between predicted and actual values, divided by the actual values, and most often expressed as a percentage. This makes it humanly intuitive since it simply relates the error to the sizes of the numbers being forecast, presenting a clear percentage-based indication of model accuracy. For example, a MAPE of 5% indicates that the error in average prediction will be at most 5% around the actual values. Therefore, this easily guides a given stakeholder in case he or she is trying to assess the performance of the model.

In this project, the principal metric used to distinguish the effectiveness of the various models deployed—ARIMA, SARIMA, and Prophet—was MAPE. This was due to its relevance in showing the proportionality of the prediction errors in relation to the actual figures of sales—a matter imperative in making any strategic decisions within the context of retail management. This relative error measure allowed the comparison of the different models with an aim to show which one provides more precise prediction and forecast—actually, following the tendency of sales more closely, without being influenced by a big numerical value that is present in large volumes in big data sets.The test that matters most in precision forecasting for inventory and planning sales, MAPE, was the litmus test of which model was most likely to be the one that would benefit the business operationally. For instance, MAPE values for the models with low values would indicate likely improved stock order accuracy with optimized levels of inventory, hence likely improved customer satisfaction from enhanced product availability.This focus on MAPE was supported by the mention in the literature as a fundamental tool of comparison between forecast accuracies for different models and industries, hence adding reliability and relevance for use (Hyndman and Koehler, 2006). MAPE further helped in putting the clear identification of strengths and weaknesses into perspective for every forecasting methodology considered, hence giving guidance for further refinement in the phase of developing the model to the detail that guarantees the chosen models perfectly conform to the operational objectives of the organization.

Fitting the models to a training set of the data and testing it on another piece that was heretofore unseen, in fact, this simulates the real situation in which forecasts are made for future unknown values. The comprehensive testing framework helps make sure the models are not only statistically accurate per the MAPE but also practically viable in a true forecasting environment. The parameters of each model were fine-tuned using insights identified during this validation process, ensuring that the final applied models are both accurate and generalizable, giving good effectiveness in front of new data as it becomes available. This high level of methodological detail underscores how seriously the project is taking the need to deliver sturdy and reliable tools for strategic decision-making in retail management.

4.6 Visualization and Decision Support Tools

Interactive visualizations have played a critical role in translating complex predictive data for decision-makers into actionable insights. The visualizations were developed interactively using Plotly, which is really a very good and powerful library friendly with Jupyter notebooks. This was an informed choice due to the capabilities in Plotly's ability to produce highly interactive and visually aesthetic graphical representations of data, which are a must-have in engaging stakeholders in the interpretation of trends and forecasts.

Dynamic charting capability was brought out using Plotly in the Python-based Jupyter Notebook environment so that the user can be able to interact with the piece of data presented in real time. This comes in handy with the need to explore different forecasting scenarios by making adjustments to parameters and immediately observing the effects on forecasts. For example, it was possible to zoom to any timeframe, hover over a data point to get exact information, switch off and on any data series, among a host of other features.

These visualizations would include times series forecasts. Every one of these charts would be designed and built to support some definite aspects of the retail management decision-making process. For instance, forecasting visualizations give management the ability to predict future sales volumes within the upcoming periods—an aspect that is a necessity for inventory and workforce planning. Above error metric graphs will help to understand accuracy and reliability between different models and hence further will help in refining the best performing models.

Such visual representations would have armed the management with better strategies by displaying the effects of such external factors as holidays and promotional events on sales. In such ways, they may schedule promotions or changes in stock levels due to projected volume increases over this period, thereby maximizing revenue and minimizing stocking costs.

Overall, Integration of Plotly in this work offered a strong platform for the development of data-driven strategies in retail management. These tools improved not only readability in complex datasets but also empowerment of stakeholders toward making informed, priority-aligned decisions with the organization's operational goals and market dynamics. So, they underline that the project is in the line of the trends of our time in business intelligence, where data visualization is taking an increasingly important role and is recognized as the absolute cornerstone of effective, goal-oriented, and evidence-based strategic planning.

Chapter 5: *Results*

The results chapter shows and discusses the outcomes of each model and the project as a whole, covering the spectrum starting from data preprocessing to the evaluation of the models and visualization. This chapter also presents the findings derived from the application of ARIMA, SARIMA and Prophet models to forecast retail sales data, showing what each method is stronger at and what the limitation for each method is. The analysis done at first highlighted the complexity of the data, paving the way for a detailed exploration of the performance of the models. The section following that critically examines the forecasting accuracy of each model, the impact of applying holidays and promotions to that model, and the effectiveness of each model in addressing the challenges faced in traditional sales forecasting. Moreover, the data visualizations applied as part of the project help in providing insights into the forecasts of the models, facilitating a deeper understanding of future sales trends and making it easier for stakeholders to understand the future trends. The aim of this chapter is to delve into the analytical journey of this project, offering a clear view of the results obtained and their significance within the broader context of retail analytics.The culmination of this analysis not only benchmarks the models' performances but also sets the stage for future advancements in sales forecasting.

5.1 Overview of Data and Initial Analysis

Any predictive analysis is founded on a strong dataset that would reflect subtlety and patterns adherent to the subject in question with preciseness. For this project, large databases of retail sales were collected in the most delicate manner, from which to set fertile ground for producing accurate sophisticated models of forecasting.

5.1.1 Summary of Data Collection

The following is an analysis that has compiled datasets derived from a dataset of a highly respected retail analytics competition posted on Kaggle. The data has years of weekly sales numbers from 45 different stores, with many departments in each store.A few pre-processing steps were taken before any sort of analysis to make sure the data was of quality and integrity. They involved cleaning missing values, correction of anomalies, and normalization of the data into a suitable format for time series analysis. After these steps, the result was two datasets where the first one was called 'aggregate_sales' which contained all sum of weekly sales of all the stores and departments, and the other was 'top_performer' where the top performing department in the highest performing store was selected. The reason for choosing these two datasets instead of the original sales datasets was because the dates of stores overlapped which did not provide a clear analysis. The models where aggregate sales was used should be essential for understanding overall trends and setting a benchmark for forecasting models while the model for the top performing store focuses on a certain segment to refine the model's usefulness and accuracy. Moreover, this division was used to observe which dataset would yield better MAPE and better overall.

5.1.2 Initial key insights

Initial visualizations, particularly time series plots, laid bare the overall sales trends across the retail chain (refer to the following figure A "Overall Sales Over Time"). These plots also revealed the cyclical character of sales—it has its maximums and minimums during determined seasons of year. They are quite important for any forecasting model. The seasonality detection through time-series decomposition (Seasonality decomposition) provides a clear picture of the patterns that repeat and are not necessarily visible in the aggregation of data (refer to the following figure B).This decomposition helped in separating trend, seasonality, and the residuals to further analyze the systematic and unsystematic parts of the sales data.

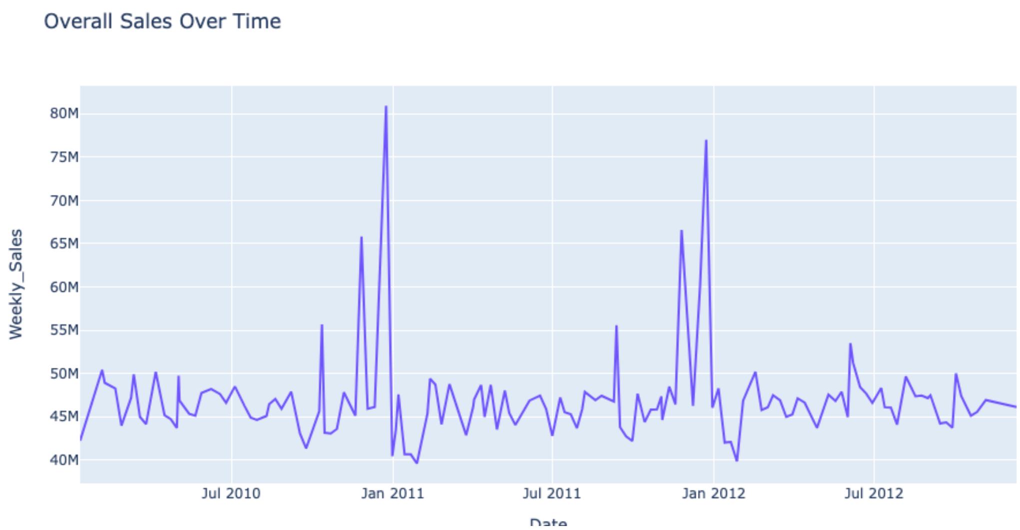


Figure A

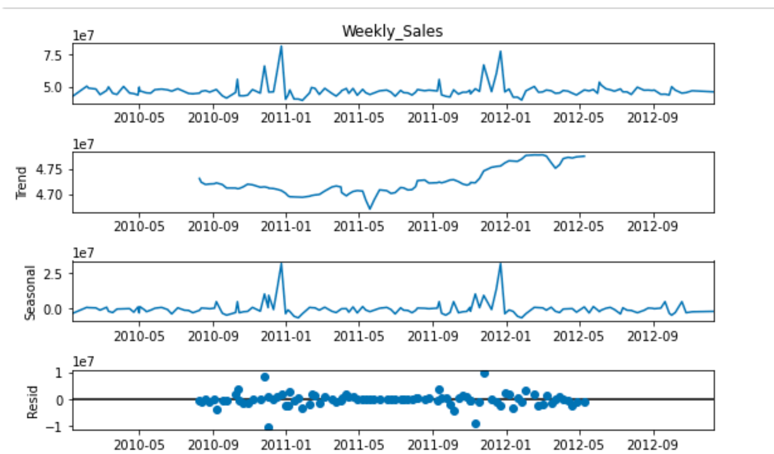


Figure B

Anomaly detection was therefore part of a very important EDA, considering rolling statistics of the data and testing for stationarity using tests like the Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (view Images with the exact labeled titles in the Appendix), irregularities were identified in the top performer dataset using the KPSS test and a further investigation of them was carried out using rolling statistics to prevent bias in the forecasting models.

Rolling statistics analysis In time-series analysis, rolling statistics provide a tool to judge graphically about the stability of the mean and the variance. The moving window helps to damp short-term changes in the level of data and picks out long-term movements, seasonal effects, and potential breaks.

The rolling mean and standard deviation are obtained from the aggregate sales data for a 12-week window that gives significant spikes related to seasonal promotional periods and holidays (refer to the following figure C "Rolling Statistics for Aggregate Sales"). These peaks in the rolling mean indicate periods of high consumer activity that are likely attributed to seasonal buying behaviors and marketing campaigns.

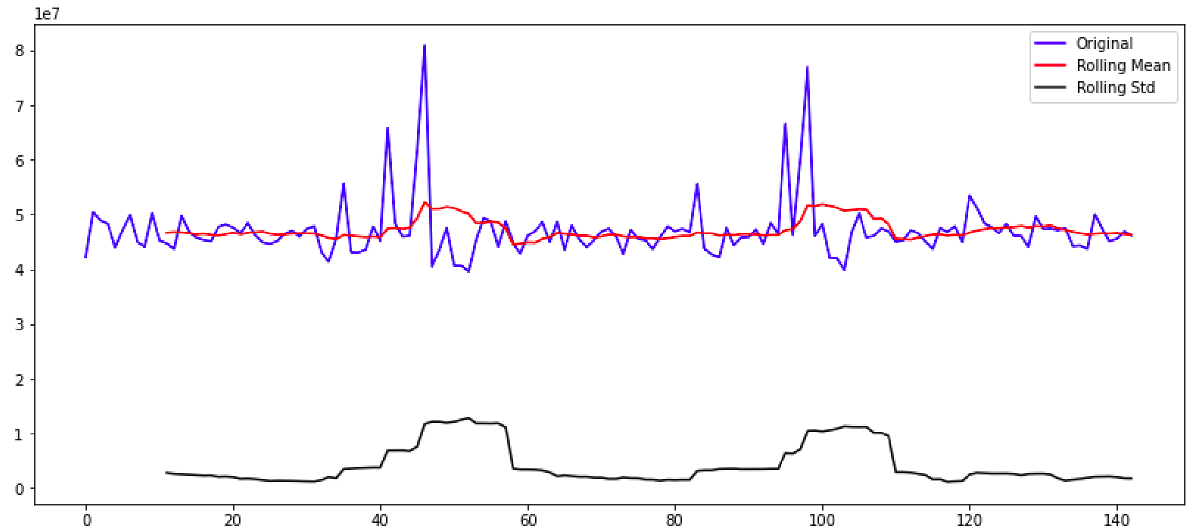


Figure C

While the rolling standard deviation presents variation, this must be considered, attributable to the same events since the average is showing the same behavior, even though it is relatively stable. In the case of the highest sales (refer to the following figure D "Rolling Statistics for Top Performing Store"), the rolling mean is taking a more pronounced cyclical pattern that aligns with recurring sales events and seasonal promotions.

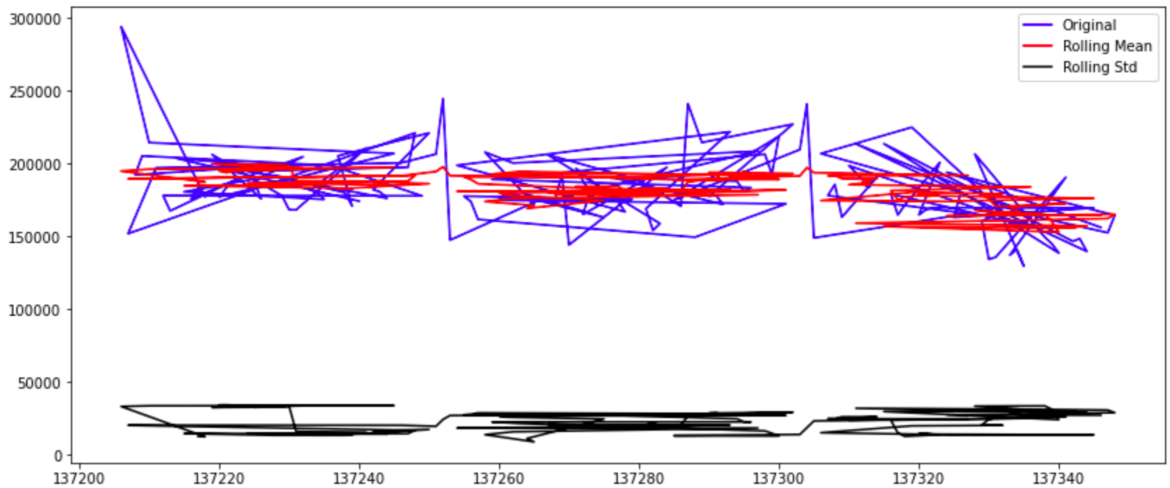


Figure D

Fluctuations in the rolling standard deviation therefore suggest that there's greater variability in sales and may be explained by factors such as localized marketing campaigns, store-specific events, or changes within the inventory stock. The analysis with rolling statistics is not only indicative of the predictability of the sales but also portrays the importance of taking into consideration the external factors in forecasting models. As an example, some peaks in weekly sales may need another modeling approach whereby the public holidays or the store promotions are included in explaining the weekly sales.

Moreover, the rolling statistics show stationarity in the aggregate sales data—a factor later confirmed by the results of the ADF test, which implies that a constant mean factor over time is in existence. However, using the KPSS test, it may result that the best store does not follow a stationary process and therefore may follow a different underlying process. Thus, such a dichotomy between the two datasets explains why later on models are incapable of incorporating behavior differences in data, like ARIMA models. For example, SARIMA models are capable of accounting for both stationary and non-stationary data components with seasonal elements. Thus, the analysis of these rolling statistics equips us with good temporal structure knowledge of the data that will enable guiding the right pre-processing steps of the data series and, in the end, help to choose the right model for accurate forecasting. In the second model development and implementation stages, patterns and statistical properties observed by these visual tools prove useful.

Armed with dynamic visualizations and rigorous statistical testing, these are the basis upon which model selection and tuning are done during the predictive modeling. The insights derived aim from these to guide not only the model selection and tuning process but also to enhance the interpretability and applicability of the forecasting results.

5.2 Model Development and Evaluation

5.2.1 ARIMA Model Results

Model Specification: ARIMA models have parameters that are usually identified by ACF (autocorrelation function) and PACF (partial autocorrelation function) diagrams which were plotted before going into the model developments. These functions show the intensity of the temporal correlation, each in their own manner. A classical model is considered adjusted when both, ACF and PACF, show no significance. The ACF is obtained from the linear correlation of each xt value of the series to the others in different lags, as xt−1, xt−2 and so on. The PACF, however, obtains almost the same result, but removing the interference of the other values. In the ACF, for example, the correlation between xt and xt−2 suffers the interference of xt−1(Flores, Engel and Pinto, 2012). In the context of this analysis, these functions were used to identify the parameters for the two ARIMA models used, ARIMA model for the top performing store and ARIMA model for the aggregate sales.

After these functions and their diagrams, shown in the appendix as *figures 1, 2, 3 and 4*, were created the data was split into training and testing sets where the ratio was 80:20. This meant for aggregate sales data frame the training set size was 114 and the test set size was 29, on the other hand the top performing store data frame training set size was also 114 and the test set was also 29.

Using the previously stated functions ACF and PACF, the parameters for the ARIMA model which are denoted as **p,d and q** were decided, meaning there were two forecasts done for each data frame. The first parameter, **p** is the number of autoregressive terms, then **d** is the number of differencing needed for data and that usually is for data that is non-stationary and finally **q** is the number of lagged forecast errors in the prediction equation (people.duke.edu, n.d.). The first ARIMA forecast was done for the aggregate sales using p-value of 0, d-value of 0 and q-value of 22 and the ARIMA forecast done for the top performing store had the parameters of p-value of 8, d-value of 0, q-value of 22. After forecasting these models and evaluating them which will be discussed in the following sections, the parameters were hypertuned using the 'auto_arima' function imported from the 'pmdarima' library. This function seeks to identify the most optimal parameters for an ARIMA model (alkaline-ml.com, n.d.) and for the aggregate sales hypertuned model it gave the p-value of 0, d-value of 0, and q-value of 1, on the other hand for the hypertuned top performing store ARIMA model it gave the parameters of p-value of 0,d-value of 0, and q-value of 0. This meant that the total number of ARIMA models developed in this analysis was 4, 2 of which were for the aggregate sales data frame and the other two were for the top performing store data frame.

Performance Evaluation: The evaluation of the model will be done using the metrics: Mean Absolute Error (MAE) for average absolute errors, (MSE), Root Mean Squared Error (RMSE) to highlight larger errors, and Mean Absolute Percentage Error (MAPE) used as a relative evaluation metric for time series forecasting models. The main metric used for measurement in this assignment is the MAPE metric where the average percentage of absolute errors is measured between the actual and predicted values. MAPE was calculated using the following function (GeeksforGeeks, 2021) :

$$MAPE = \frac{1}{n} \times \sum \left| \frac{actual\ value\ - forecast\ value}{actual\ value} \right|$$

n denotes the number of observations. For the first ARIMA model, which was for the total aggregate sales data frame, the MAPE value obtained for that model was 0.045 which is 4.5%, which indicates highly accurate forecasting but after the 'auto_arima' function was used better results were yielded where the MAPE had a value of 3.5% but as shown in the previous section the forecast somehow only showed a stationary line which meant the using the hypertuned parameters did yield better performance metrics but worse forecasts which leaded to the conclusion that a SARIMA model might be more accurate. Furthermore, the two ARIMA models implemented for the top performing store yielded good results as well, but not as well as the aggregate sales ARIMA models. The first ARIMA model got 21% for MAPE results which meant that the result was reasonable but not good for predicting future sales, after hypertuning the parameters the model yielded 19.9% which meant the model good for forecasting but had the same issue as the aggregate sales where the forecast showed a stationary l;ine in the hypertuned model but not on the normal model. To sum it up after observing these results, it seems that aggregate sales would be better for predicting future sales but also a SARIMA model would be better due to the seasonality of retail sales data.

Forecasting Results: Forecasting results serve as a critical component in the validation and application of predictive models. The primary objective in forecasting is not only to predict future values but also to understand the behavior and dynamics of the data under study. In the context of sales forecasting, this becomes particularly essential, where businesses leverage these predictions to make informed decisions on inventory management, resource allocation, and strategic planning.

For the aggregate sales data, the ARIMA model's forecast was anchored in historical sales trends, aiming to extend the pattern into the future. The model's parameters were chosen based on the ACF and PACF analysis, with an initial non-hyper-tuned model displaying a p-value of 0, a d-value of 0, and a q-value of 22. The forecast, as shown in the figure 1, reveals a certain degree of fluctuation with spikes and troughs, suggesting the model's sensitivity to the historical variance in the data. This variation, although indicative of the model capturing the inherent sales volatility, also raises questions about overfitting and the model's ability to generalize to unseen data.

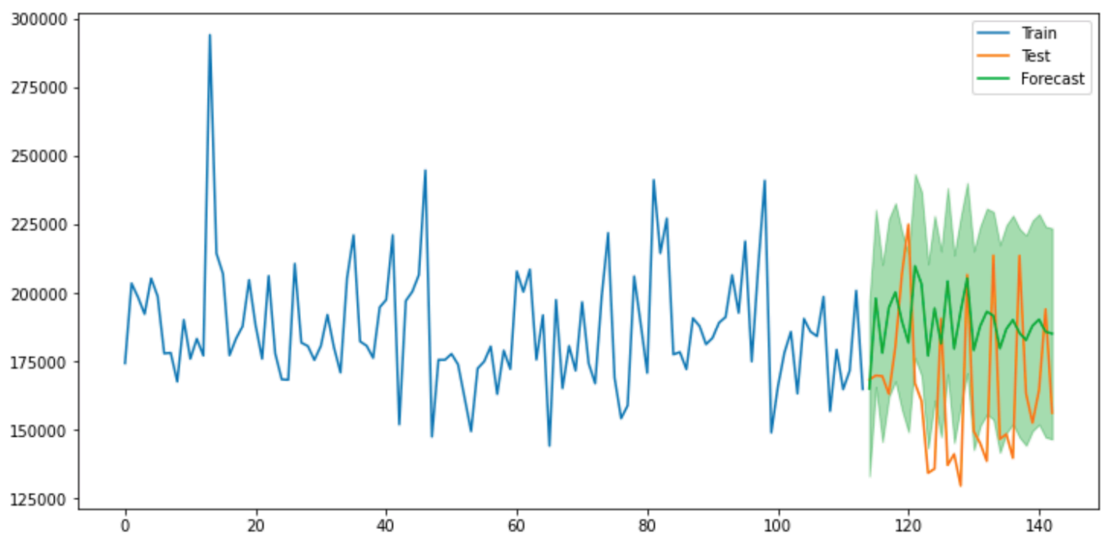


Figure 1

Upon hypertuning the ARIMA parameters using the 'auto_arima' function, the model simplified to a p-value of 0, a d-value of 0, and a q-value of 1. The hypertuned model, depicted in figure 2, resulted in a forecast that presents as a stationary line. While the performance metrics improved, with MAPE dropping to 3.5%, indicating a more accurate average prediction error, the graphical representation of the forecast suggests a lack of dynamic response to the seasonality and trend components of the sales data. This incongruence between the performance metrics and visual forecast underscores the complexity of model selection and the trade-off between statistical accuracy and practical applicability.

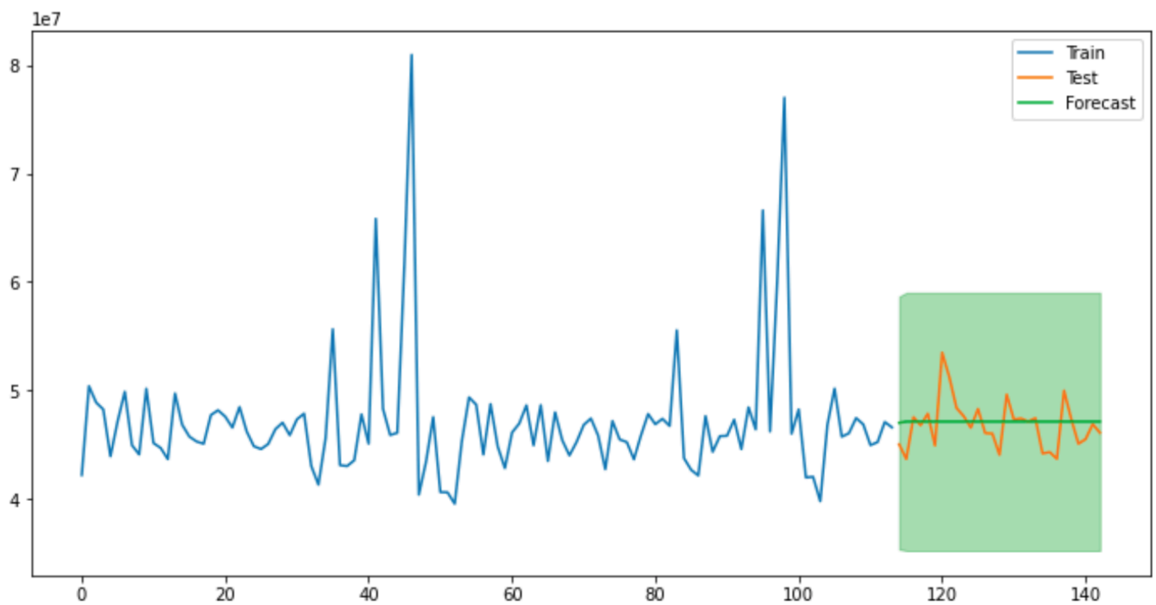


Figure 2

Turning to the top-performing store, the ARIMA model results exhibited a different behavior. The non-hyper-tuned model parameters were p=8, d=0, and q=22, showing a forecast that included variability, as shown in figure 3. This variability suggests that the model attempted to mirror the sales behavior with a moderate level of accuracy, as evidenced by a MAPE of 21%.

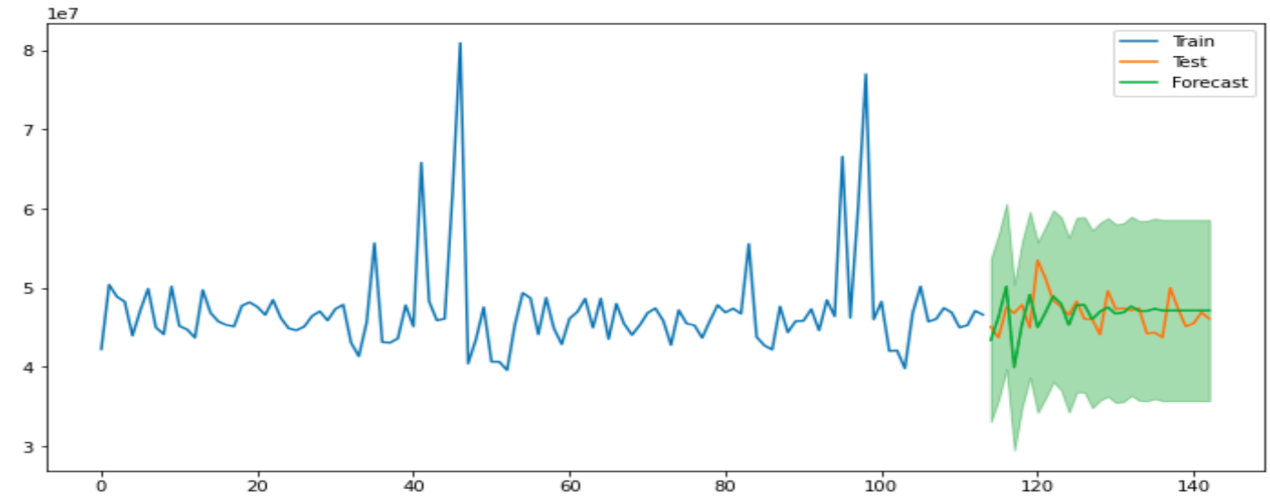


Figure 3

Post-hypertuning, with parameters reducing to p=0, d=0, and q=0, the model forecast, much like the aggregate sales hyper-tuned model, rendered a stationary prediction line illustrated will be shown in the appendix as *figure 5*. The performance metrics saw a marginal improvement with MAPE decreasing to 19.9%; however, similar to the aggregate sales model, the predictive line failed to display the seasonal swings expected in retail sales data.

In summary, while hyper-tuning resulted in statistical improvement according to selected metrics, the visual forecast analysis raises concerns about the models' utility in real-world scenarios. The stationary forecast line contradicts the expected cyclical nature of sales, indicating that these models may not capture the full scope of underlying sales patterns. Therefore, while the MAPE improvement suggests a superficial enhancement, the practical use of these forecasts for decision-making in a dynamic retail environment is questionable. These observations point to the need for a more nuanced approach, potentially involving models that explicitly account for seasonality, trend components, and external factors—a role that SARIMA models are better suited to fulfill.

5.2.2 SARIMA Model Results

Model Specification: the SARIMA model, also known as seasonal ARIMA, is an extension of ARIMA where seasonality is incorporated in addition to the non-seasonal components. In addition to the three parameters of ARIMA, which were listed in the previous section (**p, d and q**), SARIMA has three other parameters that represent the seasonal components denoted as **P, D, Q,** and **s** represents the seasonal period. For the SARIMA models, a slightly different approach was taken than that of ARIMA where the 'auto_arima' function was used right away, and changed the 'is_seasonal' parameter of the function to true. In light of that information, for SARIMA only two models in total were created where one was for aggregate sales and the other was for the top performing store. The parameters that were given by the function for the aggregate sales data frame model were a p-value of 2, a d-value of 0, a q-value of 1 and for the seasonal components a P-value of 0, a D-value of 1, a Q-value of 1 and an s-value of 12. On the other hand for the top performing data frame model had parameters which consisted of a p-value of 0, a d-value of 0, a q-value of 0 and for the seasonal components a P-value of 0, a D-value of 1, a Q-value of 0 and an s-value of 12.

SARIMA models have some advantages over normal ARIMA models which include capturing seasonality, like in the case of retail sales, some seasons have higher sales due to holiday or promotions and SARIMA can capture that and adjust the forecast accordingly. While it is superior to ARIMA in that sense, ARIMA may be more robust in that it would not assume that seasonal patterns are stable over time.

Performance evaluation: The SARIMA model for aggregate sales displayed a Mean Absolute Percentage Error (MAPE) of 4.5%, a metric that, while marginally higher than the 3.5% score of the hypertuned ARIMA, indicating a robust predictive capability as well as reflecting that in the forecasting which will be shown in the following section. Notably, the SARIMA model managed to get the same score as the basic ARIMA model's MAPE value but also reflected it better on the forecast, underscoring the importance of accommodating seasonality in the predictive process. The forecasts generated by the SARIMA model showcased a dynamic line that mirrored the seasonality and trends present in the historical data more closely than the hypertuned ARIMA's stationary projection. This suggests that while the hypertuned ARIMA model exhibited tighter performance metrics, it fell short in terms of generating actionable insights for future sales trends.

The SARIMA model developed for the top-performing store data revealed a MAPE of 18%, a considerable improvement over the 21% score of the original ARIMA model. Although this marked an advancement in forecast accuracy, it remained inferior to the aggregate sales SARIMA model. The implication of this discrepancy could be rooted in the nature of the data set for the top-performing store, which might possess unique characteristics or non-standard seasonal patterns that are more challenging to model accurately.

When juxtaposing the SARIMA models’ performance with their ARIMA counterparts, a clear pattern emerges: the SARIMA models, designed to factor in seasonality, tend to offer enhanced forecasting capabilities for retail sales data. The improvement in MAPE values between non-seasonal ARIMA models and their seasonal SARIMA counterparts was pronounced for the aggregate sales dataset, less so for the top-performing store. However, the SARIMA model still offered a more nuanced forecast that could potentially translate to more strategic stock replenishment and inventory management decisions.

In conclusion, while the hypertuned ARIMA models did provide a marginal improvement in point forecast accuracy as measured by MAPE, their utility was limited due to the lack of seasonal variation in the forecasts. On the other hand, the SARIMA models proved to be superior in capturing the cyclical nature of retail sales, making them more suitable for real-world application where understanding periodicity is crucial for anticipating sales volumes. This analysis reaffirms the pivotal role of selecting a model congruent with the data’s inherent structure and the business context in which it will be applied.

Forecasting results:

SARIMA Forecast Analysis for Aggregate Sales:

The SARIMA model designed for the aggregate sales data implemented specific parameters to effectively capture the inherent seasonality of the dataset. The model's non-seasonal components were set at (p=2, d=0, q=1) and seasonal components at (P=0, D=1, Q=1, s=12), aiming to accurately reflect the cyclical fluctuations observed in retail sales cycles. The forecast, visualized as a dynamic curve that followed the data's peaks and valleys, achieved a MAPE of 4.5%, indicating a model that not only presents statistically low average error but also provides practical and actionable forecasts.(shown in figure 4).

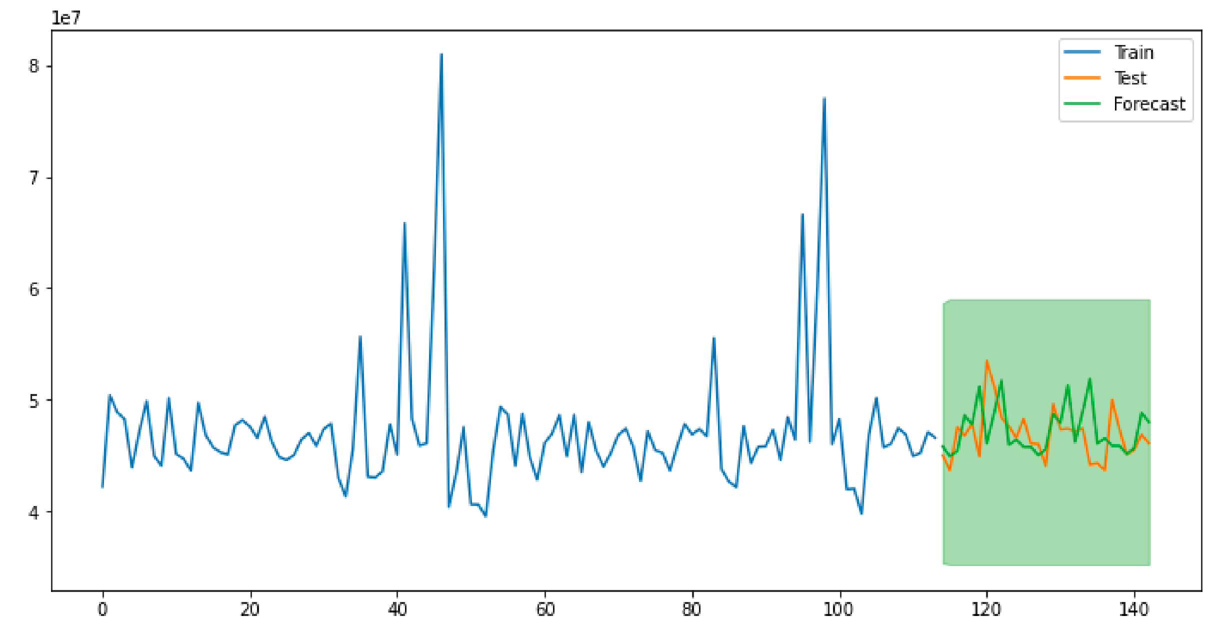


Figure 4

SARIMA Forecast Analysis for Top Performer:

For the top performer dataset, the parameters of the model were adjusted to (p=0, d=0, q=0) for non-seasonal and (P=0, D=1, Q=0, s=12) for seasonal components, tailored to capture specific and potentially more erratic seasonal trends. Despite these adjustments, the resulting forecast depicted in the analysis showed a MAPE of 18%, highlighting a significant discrepancy in predictive accuracy compared to the aggregate sales SARIMA model (the forecast is shown in figure 5).

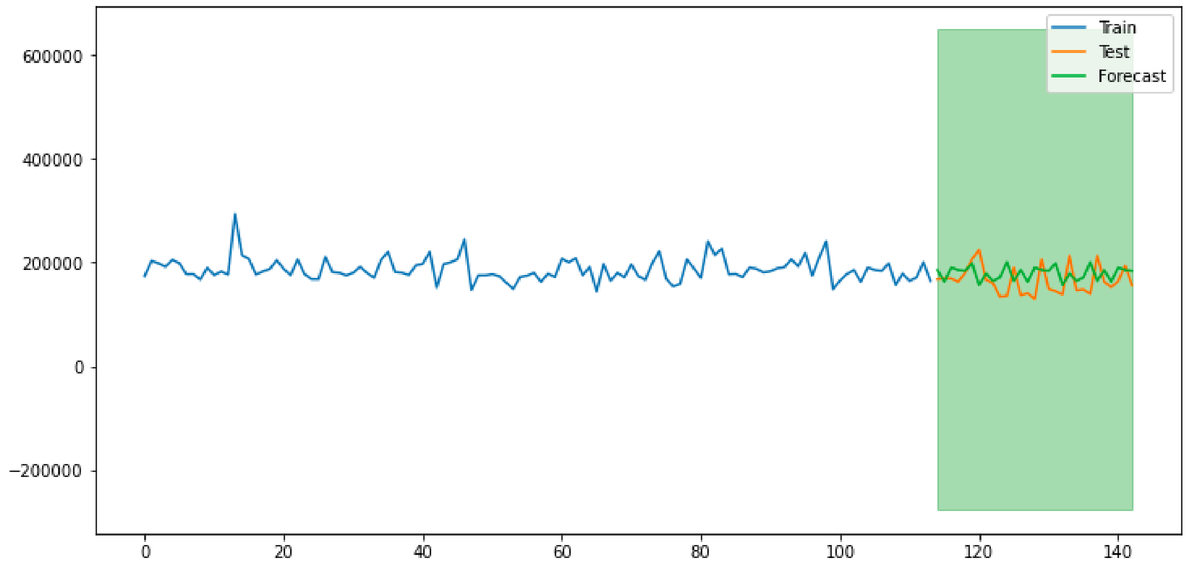


Figure 5

Discussion of Comparative Forecasting Performance:

The direct comparison of SARIMA models across both datasets reveals that model performance is closely tied to the data characteristics it is applied to. The aggregate sales data, with broader and more general sales patterns, appeared to be more conducive to this type of seasonal modeling. In contrast, the top performer's data, which may exhibit more unique and idiosyncratic patterns, poses challenges that are not as straightforwardly modeled by SARIMA. This analysis underscores the importance of model parameter customization and selection based on specific dataset properties to optimize forecasting accuracy.

5.2.3 Prophet Model Results

Model Adaptations for Aggregate Sales Data:

The Prophet model tailors the aggregate sales data with the focus to detect weekly seasonality, while the daily seasonality was set to False. This reflects the Prophet nature of working with weekly sales data. The interval width was set in hyperparameters to be 0.95, such that when set to a 95% prediction interval, we get the resultant confidence intervals of the forecasts, as shown on the plot below (Figure 6).

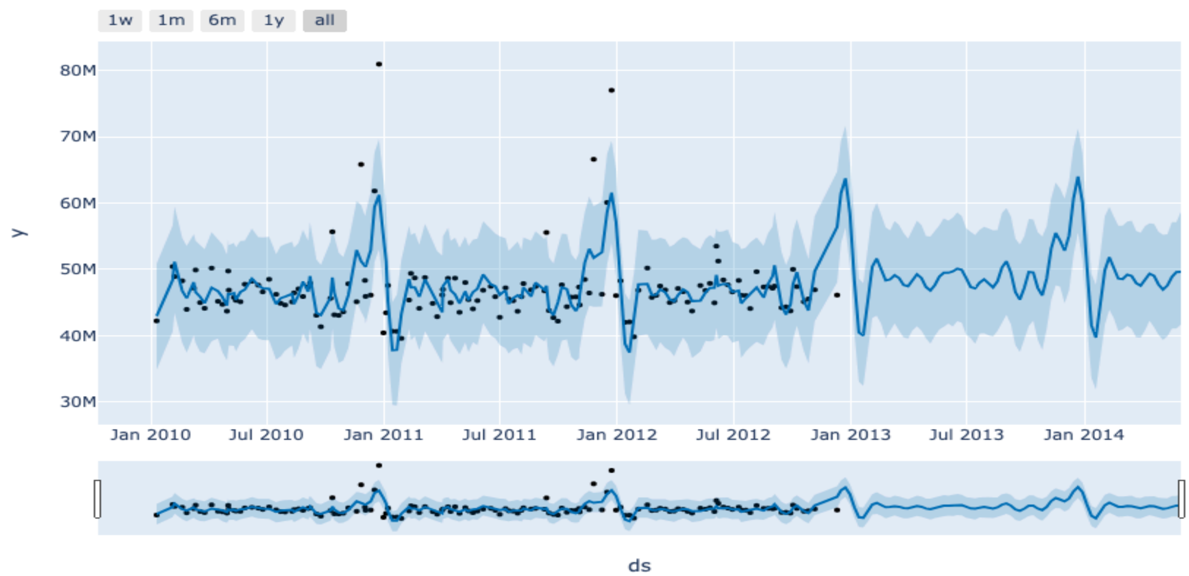


Figure 6

From component plots (Figure 7), it is observed that the model has handled non-linear trends very effectively, as there is yearly seasonality in the model. These are plotted, which helps identify the trend, yearly seasonality, and weekly seasonality present in the dataset. Thus, this level of decomposition has allowed a very accurate level of refinement — most particularly, in capturing the unique weekly patterns a normal retail sales cycle will show.

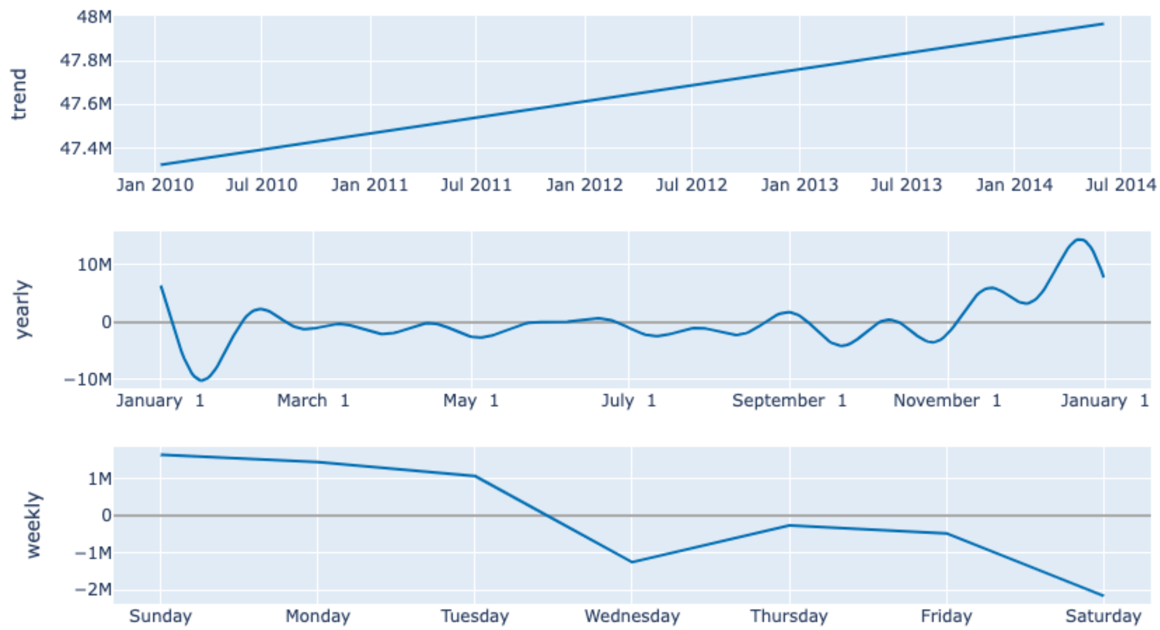


Figure 7

Forecasting Visualizations:

The generated visualization by Prophet (figure 1) included not only the forecast but also the confidence intervals; these are very important for risk assessment in sales planning. Moreover, the ability to forecast the future 1.5 years ahead provides enough runway for strategic decisions.

Model Components Dissection: in this model the component analysis indicates that the components brought out inherent behavior over the years. The trend component shows a steady increment over time, showing growth in the sales data (figure 7).

The yearly seasonality summarized the patterns taking place annually, which may be linked to the holidays or the seasonal sales campaigns. The weekly component explained the weekly fluctuations and might indicate the actual customers' behavior on certain days.

Inclusion of Holiday Effects:

Another Prophet model test for adaptability was done for an addition to the forecasting model (Figure 8). Here, the US holidays were added into the model in order to find the effect such events may create on sales.

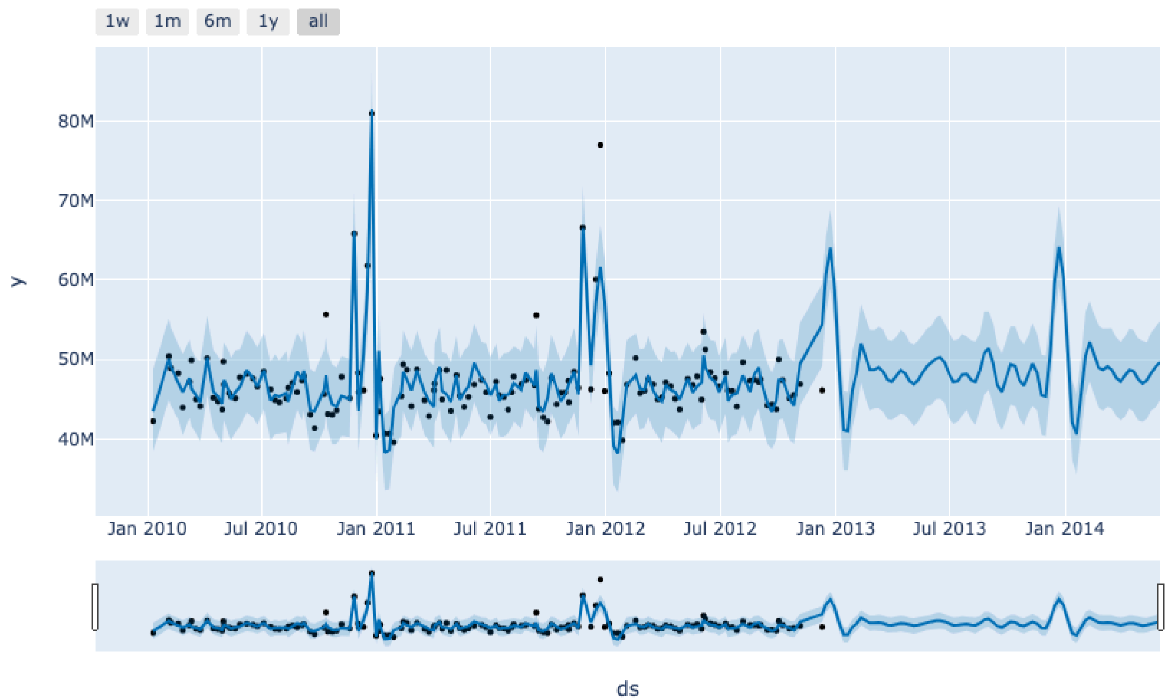


Figure 8

This is clearly shown in the plots during these periods, reaffirming the common hypothesis of increased sales during holidays.

Key Performance Metrics: The Prophet model's performance metrics with aggregate sales data were quite stringent. The MAPE of the aggregate sales data under Prophet was 0.047601, which suggests that the forecast is off from the actual, on average, by 4.76%. Which is overall very good, and that is within the right range for the forecasting purposes (Image 6). Generally, the Prophet model has shown enormous promise in forecasting retail sales data and performed excellently in handling the complex seasonal patterns, providing very sensible forecasts, which are key to strategic planning.

Model Adaptations for Top Performing Store Data:

This was a similar approach to the prophet model for aggregate sales weekly seasonality turned on, daily seasonality turned off, and the confidence interval was kept at 0.95 to maintain consistency in the confidence level of the predictions. This setting was reflected in the forecast visualization which did not bring out much detail in the sales forecast since the level of fluctuation is very high at the individual store performance.

Forecasting Visualizations:

The forecast plots show the range of expected sales as confidence intervals, both showing the expected values and the range of uncertainty. These visualizations, up to 77 weeks ahead, attempt to exhibit the model's ability in extrapolating trends while taking into consideration the inherent uncertainties present in the sales data.

Model Components Dissection:

The component plots for the top performing store showed a slightly different trend than the mean aggregate sales data; the latter had a clear downward trajectory with time, showing either falling sales or a different pattern of buying taking place. In the seasonal components of the model, one notices intricate patterns in sales behavior, with the graph for the yearly seasonality pointing to certain periods of increased activity, probably due to promotions or season behavior.

Inclusion of Holiday Effects:

The model had included holiday effects in the procedure of assessing its impacts on sales. Conversely, the component plot including holidays showed spikes that were conforming to known holiday periods and, therefore, indicated an expected sales lift during such times.

However, the level of detail within the model allowed this amount of detail and a more engaged realization of what was driving sales. This attention to detail in the models and extra potential variables gives a sophisticated analysis that can lead to sharp marketing approaches at key holiday times.

Performance Metrics:

The performance metrics at store level for the top-performing store showed a larger error margin with a MAPE of 0.161497 as compared to the aggregate sales model. It means that predictions were off by 16.15% from actual sales on average—of course, still useful for forecasting but suggesting a little bit more unpredictability or volatility of the sales patterns of the particular store. In conclusion, Prophet was applied on the sales data of the top performing store, and very intriguing and unique dynamics of Prophet were met. However, a larger error in the forecasted metrics might seem to signal a difficult environment for perfect predictions but is, however, detailed in component decomposition and further, in the incorporation of holiday effects. These insights can then be of great value in tailoring store-specific strategies and operations.

5.3 Comparative Analysis of Models

Overall Performance:

The comparison made for the ARIMA, SARIMA, and Prophet models shows the edge of all those models in improving forecasting accuracy for all the two datasets: aggregate sales and top performing store.

ARIMA & SARIMA Models: For aggregate sales, the ARIMA model served as a good baseline because tuning of hyperparameters did not make any difference in the score. SARIMA improved on that with the accommodation of seasonality fluctuation with the same MAPE but correctly forecasted seasonality in the pattern. For the top performer store, both ARIMA and SARIMA models suffer from problems, for it is evident with the initial MAPE values at 21% and 18%, respectively. It draws perhaps a potential limitation to the models on their ability to capture volatilities in individual store sales data.

Prophet Model: The Prophet model was adaptive in nature, as it accommodated effects from holidays and therefore could produce a smoother forecast. With a MAPE value of around 4.76% in aggregate sales and 16.15% in the top performer store, it became statistically significant in modeling complex seasonalities and trends. Though it showed higher predictive error for the more volatile store-specific dataset, the model actually did predict the sale value better.

Computational Efficiency:

ARIMA & SARIMA models: Usually, these models are lighter computationally but may get cumbersome during the parameter choosing stage, more so when methods are applied to hyper-parameter tuning. Simple data will rarely have a lot of computation overhead; however, high complexity in the data may arise in a series that is not stationary and needs differencing and seasonal adjustment.

Prophet Model: Prophet is designed to work well with big datasets, however, sensitivities to parameter changes are rare, making it faster to compute results. On the other hand, this might be computationally expensive when adding many regressors or holiday effects, as was the case with the best performer store.

Adaptability and Scalability:

ARIMA & SARIMA Models: The included ARIMA model contains somewhat less adaptability against the changes in data patterns without having to re-estimate the parameters, whereas SARIMA includes seasonal parameters that offer little better adaptability. Both models, however, suffer from the limitation of scalability, because they operate on historical data points and can thus be unwieldy as the dataset becomes large.

Prophet Model: Prophet model is a robust model that leverages its capability of adapting to new data and being scalable. The model has been developed in a manner that automatically changes points in trends, thereby dynamically changing the respective forecasts. It is well designed to accept daily data for the test years and can be highly scalable for large datasets. In brief, ARIMA and SARIMA models have the ability to give great forecasts when applied to very stable datasets. On the other hand, these models might fail in other cases of irregularities and volatilities in individual store data. Where the Prophet model stands out is its flexibility and scalability. However, in terms of accuracy, the performance is definitely not outstanding—especially when working with such volatile datasets like the Top Performer Store. Based on the characteristics of the dataset and the forecasting needs of the business, the computational requirements and performance metrics of each model suggest that the choice of model should be aligned.

5.5 Future Work and Limitations

Model Limitations:

- **ARIMA/SARIMA Models**: These models primarily rely on the assumption of stationary data and consistent seasonal patterns, making them less effective in scenarios with sudden market shifts or non-typical events. Additionally, they require extensive historical data for accurate predictions, which may not be readily available or reflective of future patterns due to changes in consumer behaviors or market conditions.
- **Prophet Model**: While adaptable and capable of integrating external factors such as holidays, the Prophet model may suffer from overfitting due to its flexibility. It also relies heavily on daily data to capture seasonality, which might not be practical for all retail settings, particularly where data granularity varies.

Future Work:

- **Enhanced Model Robustness**: Future studies could focus on increasing the robustness of ARIMA/SARIMA models to better handle non-stationary data. Developing hybrid models that incorporate sudden data changes, perhaps by integrating modern machine learning algorithms capable of predicting shifts in consumer behavior, could be beneficial.
- **Algorithm Optimization**: Efforts could be directed towards optimizing the Prophet model to minimize the risk of overfitting. This might include employing regularization techniques or developing criteria for selecting which external factors to include in the model to ensure a balance between fit and generalization.
- **Exploration of New Models**: Research could explore alternative time series models such as Vector Autoregression (VAR) and Long Short-Term Memory (LSTM) networks, which may offer better capabilities for capturing complex interactions among different product lines and store locations.

Chapter 6: *Conclusion and Discussion*

Project Objectives:

The broad objective of the project was to enhance the accuracy and reliability of retail sales forecasting through the use of advanced time series models, aimed at mitigating inventory mismanagement like stock outs or overstocking and optimizing resource allocation. By implementing ARIMA, SARIMA, and Prophet models, this project addressed the key aspects of sales forecasting and improved decision-making processes within the retail context.

Objective Assessment:

- **ARIMA and SARIMA models** effectively provided baseline forecasts and captured seasonal fluctuations in the data, meeting the objective of utilizing time series data to forecast future sales with a solid foundation in predictive analytics.
- **The Prophet model** excelled in incorporating holidays and promotional events, demonstrating a high level of adaptability to external factors that significantly impact retail sales.

Research Questions and Hypotheses:

- Research questions focused on the models' capacity to accurately predict sales and their computational efficiency. Analysis showed that while ARIMA and SARIMA are competent, the Prophet model provided a more nuanced understanding of sales patterns.
- Hypotheses regarding the integration of external factors leading to enhanced forecasting were confirmed, particularly with the Prophet model's adaptability to holidays and promotional events.

General Conclusions:

- The comparative analysis revealed that no single model unequivocally outperforms the others. Each model has distinct strengths and limitations, suggesting that a hybrid or ensemble approach might yield the most robust forecasting method.

Individual Conclusion Discussion:

- **ARIMA/SARIMA Conclusion:** These models are reliable for data with clear seasonal patterns but show limitations in responding to abrupt market changes.
- **Prophet Conclusion:** The model's flexibility and extensibility are its core strengths, offering detailed insights that can significantly benefit retail sales forecasting.

Link to Literature Review:

- The findings align with literature advocating for the use of SARIMA in capturing seasonal behaviors and ARIMA for more stationary trends. Prophet's performance supports recent advancements in predictive analytics, emphasizing the importance of external factors in forecasting.

Contributions to Knowledge:

- This project contributes empirical evidence to the debate on the effectiveness of traditional time series models versus more modern approaches like Prophet in retail sales forecasting. It also offers a new perspective on the practical implications of model selection in a real-world retail setting.

Implications for Future Work:

- The conclusions drawn from this project serve as a foundation for future research, particularly in exploring ensemble methods that combine the strengths of individual models. The project also highlights the need for continuous model refinement to adapt to evolving market dynamics.

Personal Reflection:

- **Project Management and Control:** The project required meticulous planning and iterative testing, emphasizing the importance of agility in project management.
- **Learning and Challenges:** Exploring different forecasting models provided invaluable insights into data analytics, while the challenges faced underscored the need for flexibility in problem-solving.

Closing Remarks:

This project has illuminated the intricacies of sales forecasting within the retail industry. The iterative process of model selection and evaluation has not only enriched the understanding of predictive modeling but also underscored the importance of continuous innovation in this domain to keep pace with the changing retail landscape.

Glossary

- **ARIMA (AutoRegressive Integrated Moving Average):** A forecasting algorithm that uses time-series data to predict future points by accounting for trends and cycles.
- **SARIMA (Seasonal AutoRegressive Integrated Moving Average):** An extension of ARIMA that also models seasonality, making it ideal for data with seasonal patterns.
- **Prophet:** A forecasting tool designed for handling time-series data that is irregular, has strong seasonal effects, and may include historical data with missing values or outliers.
- **MAE (Mean Absolute Error):** A measure of errors between paired observations that represent the same phenomenon.
- **MSE (Mean Squared Error):** The average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value.
- **RMSE (Root Mean Squared Error):** The square root of the average of squared differences between prediction and actual observation.
- **MAPE (Mean Absolute Percentage Error):** A measure of prediction accuracy of a forecasting method in statistics, for example, in trend estimation, representing the average absolute percent error between predicted and actual values.
- **ACF (AutoCorrelation Function):** A tool used to find the correlation of the time series with its own lagged values.
- **PACF (Partial AutoCorrelation Function):** A tool used to find the correlation of the time series with its own lagged values, after removing the effects of earlier lags.
- **Kaggle:** An online community of data scientists and machine learners, owned by Google, that offers datasets and competitions to encourage data science work.
- **Outlier:** A data point that differs significantly from other observations, which could due to variability in measurement or it may indicate experimental error; the latter are sometimes excluded from the data set.
- **Stationarity:** A statistical assumption that a time series is stationary if its statistical properties such as mean, variance, autocorrelation, etc. are all constant over time.
- **Seasonality:** Any predictable and recurring pattern in a time series data set that occurs at regular intervals less than a year.
- **Hyperparameter Tuning:** The process of finding the set of optimal hyperparameters for a learning algorithm, which leads to the best performance.
- **Time Series Decomposition:** A process by which a time series is broken down into trend, seasonality, and noise components.

- **Forecast Accuracy:** The degree of closeness of the quantity forecasted to the actual outcome.
- **Rolling Statistics:** Mathematical techniques, which apply a statistical measure to a rolling subset of data points to analyze the time series data's trends and patterns over time.
- **Augmented Dickey-Fuller (ADF) Test:** A statistical test used to determine whether a unit root is present in an autoregressive model.
- **Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test:** A statistical test used to test for the presence of a unit root in a time series sample.
- **Jupyter Notebook:** An open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text.
- **Plotly:** A graphing library that makes interactive, publication-quality graphs online.
- **Auto_arima:** A function within statistical software that automates the process of ARIMA model selection by identifying the most optimal parameters for the model.

References

1. Armstrong, J.S. and Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8(1), pp.69-80.
2. De Gooijer, J.G. and Hyndman, R.J., 2006. 25 years of time series forecasting. *International Journal of Forecasting*, 22(3), pp.443-473.
3. Fildes, R. and Goodwin, P., 2007. Against your better judgment? How organizations can improve their use of management judgment in forecasting. *Interfaces*, 37(6), pp.570-576.
4. Hyndman, R.J. and Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), pp.1-22.
5. Hyndman, R.J. and Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), pp.679-688.
6. Lewis, C.D., 1982. *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth Scientific.
7. Shmueli, G. and Koppius, O.R., 2011. Predictive analytics in information systems research. *MIS Quarterly*, pp.553-572.
8. Taylor, S.J. and Letham, B., 2018. Forecasting at scale. *The American Statistician*, 72(1), pp.37-45.

9. Willmott, C.J. and Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, pp.79-82.
10. Anon, (n.d.). Hyndman, R. J. and Khandakar, Y. (2008) 'Automatic Time Series Forecasting: The forecast Package for R', Journal of Statistical Software, 27(3), pp. 1–22. doi: 10.18637/jss.v027.i03..
11. Box, G.E.P. and Jenkins, G.M. (1976). *Time series analysis : forecasting and control*. Hoboken, New Jersey: John Wiley & Sons.
12. Carbonneau, R., Laframboise, K. and Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research*, 184(3), pp.1140–1154. doi:https://doi.org/10.1016/j.ejor.2006.12.004.
13. Christopher, M. (2016). *Logistics & Supply Chain Management*. 5th ed. Harlow, England ; New York: Pearson Education.
14. Fama, E.F. and French, K.R. (2004). *Capital asset pricing model : theory and evidence*. Chicago: Center For Research In Security Prices, Graduate School Of Business, University Of Chicago.
15. Few, S. (2009). *Now you see it : simple visualization techniques for quantitative analysis*. Oakland, Calif.: Analytics Press.
16. Hyndman, R.J. and Athanasopoulos, G. (2018). *Forecasting : Principles and Practice*. 2nd ed. Heathmont, Vic.: Otexts.
17. Lawrence, M.J., Edmundson, R.H. and O'Connor, M.J. (1985). An examination of the accuracy of judgmental extrapolation of time series. *International Journal of Forecasting*, 1(1), pp.25–35. doi:https://doi.org/10.1016/s0169-2070(85)80068-6.
18. Makridakis, S.G., Hyndman, R.J. and Wheelwright, S.C. (1998). *Forecasting : methods and applications*. New York: Wiley.
19. Shi, Y., Shouyang Wang, Peng, Y., Li, J., Zeng, Y. and Springerlink (Online Service (2009). *Cutting-Edge Research Topics on Multiple Criteria Decision Making : 20th International Conference, MCDM 2009, Chengdu/Jiuzhaigou, China, June 21-26, 2009. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg.
20. Taylor, S.J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, [online] 72(1)(37-45). doi:https://doi.org/10.7287/peerj.preprints.3190v2.
21. Willmott, C.J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance.

Climate Research, [online] 30(1), pp.79–82. Available at: https://www.jstor.org/stable/24869236.
22. www.kaggle.com. (n.d.). *Retail Data Analytics*. [online] Available at: https://www.kaggle.com/datasets/manjeetsingh/retaildataset?select=sales+data-set.csv [Accessed 10 Feb. 2024].
23. Chai, T. and Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), pp.1247-1250.
24. Hyndman, R.J. and Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), pp.679-688.
25. Willmott, C.J. and Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, pp.79-82.
26. Hyndman, R.J. and Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), pp.679-688.
27. Facebook (2019). *Prophet*. [online] Prophet. Available at: https://facebook.github.io/prophet/.
28. Graves, A. (2020). *Time Series Forecasting with a SARIMA Model*. [online] Medium. Available at: https://towardsdatascience.com/time-series-forecasting-with-a-sarima-model-db051b7ae459#:~:text=SARIMA%20Model%20Parameters%20%E2%80%94%20ACF%20and%20PACF%20Plots&text=Per%20the%20formula%20SARIMA(p [Accessed 29 Apr. 2024].
29. kaggle.com. (n.d.). *Retail_Analytics*. [online] Available at: https://www.kaggle.com/code/soorajgupta7/retail-analytics [Accessed 29 Apr. 2024].
30. kaggle.com. (n.d.). *Time Series: Interpreting ACF and PACF*. [online] Available at: https://www.kaggle.com/code/iamleonie/time-series-interpreting-acf-and-pacf#%F0%9F%9A%80-Cheat-Sheet [Accessed 29 Apr. 2024].

Appendices and other supporting material

Appendix A: *PDD*

Problem to be solved

Accurately predicting future sales is a difficulty that most businesses face at the moment and it will be addressed by this project: sales forecasting project with predictive analytics. The current approaches to sales forecasting may be imprecise and neglect to take into account changing market conditions, which can result in ineffective strategic decision-making, inventory management, and resource allocation.

Inaccurate Sales Predictions: Companies frequently struggle to forecast future sales, which can result in overstocking or stock outs, inefficient resource use, and higher carrying costs.

Impact on Resource Optimization: Businesses' capacity to effectively allocate resources is hampered by the absence of trustworthy sales forecasts. This involves difficulties in controlling personnel needs, production schedules, and inventory levels.

Implications for Business Strategy: Poor sales forecasting leads to poor decision-making. Companies may find it difficult to match their plans to the expected demand of the market, which would impede expansion and profitability.

The project proposal seeks to enhance the accuracy and dependability of sales forecasting by tackling the shortcomings of current techniques and integrating knowledge from other recognized predictive analytics studies [1][2].

The dataset that is going to be used in this analysis is retail data analytics dataset found on Kaggle [4] which contains historical sales data from 45 stores. it contains multiple features that will aid in my analysis, and fits perfectly with the project objectives.

<u>Project Objectives</u>

<u>Main Objective:</u> To give organizations precise and trustworthy estimates of future sales volumes, this project will create a sophisticated predictive analytics model for sales forecasting.

<u>Sub-objectives:</u>

1. **Collect data:**

   Objective: The goal of this objective is to a reliable and full dataset that would help in the analysis of the sales of a certain company and if needed merging more than one dataset.

   Testable: when the data is ready to be cleaned, explored and analyzed, this sub-objective will be deemed successful.

2. **Exploratory data analysis and data cleaning:**

   Objective: the goal of this objective is to clean the data and make it suitable for analysis and then explore the data and its properties and try to identify patterns, trends and outliers that would soon help us with the forecasting.

   Testable: successfully exploring the data and cleaning it will be enough evidence in identifying and handling anomalies.

3. **Develop Time Series Forecasting Model:**
   Objective: Create a time series by applying cutting-edge predictive analytics methods like Prophet, SARIMA, and ARIMA.
   Testable: The created model will be put to the test on the data, and pertinent metrics (such MAE, MSE, and RMSE) will be used to assess how accurate it is.
4. **Integrate External Factors for Enhanced Predictions:**
   Objective: improve the forecasting model's predicted accuracy by integrating external inputs, such as economic data and promotions.
   Testable: Through comparative analysis, the effect of outside factors on model correctness will be evaluated.
5. **Implement Interactive Visualizations for Stakeholders:**
   Objective: put in place interactive visualizations that will give business stakeholders sales projections and insights.
   Testable: The visualization's usefulness, level of visualization, and efficiency in communicating anticipated data will all be taken into account when determining how successful it is.
6. **Optimize Resource Allocation Based on Forecasts:**
   Objective: By using the generated sales projections, this project will help firms allocate resources (such as personnel, inventory, and manufacturing capacity) as efficiently as possible.
   Testable: The degree to which the actual resource allocation matches the anticipated demands will serve as the benchmark for resource optimization success.
7. **Evaluate Model Performance and Fine-Tune:**
   Objective: Using test data, this project will assess the generated model's performance and make any necessary adjustments.
   Testable: The model's performance will be evaluated by comparing it to the test sales data and making any required modifications as needed.

Project Beneficiaries

A wide range of stakeholders within the corporate ecosystem stand to gain from the predictive analytics-based sales forecasting project. Both internal and external entities will receive concrete benefits from the project, each of which will derive value in a unique way.

Internal beneficiaries:

Executive Management:

Executives and business leaders will learn important information about potential sales trends in the future. Precise projections facilitate knowledgeable decision-making, enabling executives to create and execute successful expansion and resource optimization plans.

Operations and Supply Chain Teams:

Teams working in the supply chain and operations will gain from more efficient resource allocation. Precise sales projections enable effective inventory control, averting stockouts or surplus inventory, and simplifying manufacturing plans and transportation.

Sales and Marketing Teams:

By using accurate sales forecasts, marketing and sales teams may modify their plans accordingly. To efficiently reach sales targets, this involves finding high-demand products, improving consumer engagement, and optimizing marketing efforts.

Finance Department:

Precise financial planning and budgeting will help the finance department. Precise sales forecasts support cost control, revenue planning, and overall financial stability by offering a strong basis for financial estimates.

External beneficiaries:
Customers:

Customers will enjoy better product availability and service standards. Precise sales forecasting guarantees that companies properly fulfill client demand, minimizing instances of product unavailability or service delays.
Suppliers and Partners:

By aligning their operations with precise sales forecasts, suppliers and business partners can reap the benefits. This promotes better relationships, more coordination, and supply chain optimization.

Shareholders and Investors:

The business's total success, which is fueled by wise decision-making and resource utilization, benefits investors and shareholders. The financial stability and growth potential of the organization are enhanced by precise sales forecasts.

By expanding the use of predictive analytics in sales forecasting, the project advances scientific and technological community in an academic setting. Data scientists, researchers, and students may find use for the approaches created and the knowledge acquired.

Work plan [3] (used chatgpt to help in time management in this section)

Project Initialization (Weeks 1-2):
Outputs: Project Definition Document
Define project scope, objectives, and deliverables.
Conduct an initial literature review for relevant methodologies.
Draft the Project Definition Document.

Data Collection (Weeks 3-4):
Outputs: Data Collection
Research and select suitable data collection methods.
Begin searching for the suitable dataset.

Data Exploration and Cleaning (Weeks 5-6):
Outputs: Cleaned Sales Data and visualizations
Explore sales data for patterns and outliers.
Clean and preprocess the data for analysis.

Time Series Forecasting Model Development (Weeks 7-10):
Outputs: Time Series Forecasting Model
Implement time series forecasting algorithms.
Fine-tune and validate the model with data
.
Integration of External Factors (Weeks 11-12):
Outputs: Updated Forecasting Model with External Factors
Identify and integrate external factors influencing sales.
Evaluate the impact on model accuracy.

Visualizations Design and Implementation (Weeks 13-14):
Outputs: Interactive Sales Forecast visualizations
Design and implement interactive visualizations using Jupyter notebook
Integrate machine learning predictions into the visualizations.

Resource Allocation Optimization (Weeks 15):
Outputs: Optimized Resource Allocation Guidelines
Utilize the forecasts to develop resource allocation guidelines.
Align inventory, staffing, and production capacity with forecasted demands.

Model Evaluation and Fine-Tuning (Weeks 16):
Outputs: Evaluated and Fine-Tuned Forecasting Model
Evaluate the model's performance on recent data.
Fine-tune the model based on real-world outcomes.

Finalize Project Outputs and Documentation (Weeks 17-18):
Outputs: Final Project Submission, Documentation

Finalize project outputs, including the interactive dashboard and model documentation.
Prepare the Final Project Submission document.

Research ethics checklist

Research Ethics Review Form for BSc and MSci Projects

Computer Science Research Ethics Committee (CSREC)
http://www.city.ac.uk/department-computer-science/research-ethics

Undergraduate students undertaking their final project in the Department of Computer Science must consider the ethics of their project work and ensure that it complies with research ethics guidelines and the law for data protection. In some cases, a project will need approval from an ethics committee before it can proceed. Usually, but not always, this will be because the student is involving other people ("participants") in the project.
To ensure that they give appropriate consideration to ethical issues, all students must complete this form and attach it to their project definition document (PDD). There are two parts:

PART A: Ethics Checklist. All students must complete this part.
The checklist identifies whether the project requires ethical approval and, if so, where to apply for approval.

PART B: Ethics Proportionate Review Form. Students who have answered "no" to all questions in A1, A2 and A3 and "yes" to question 4 in A4 in the ethics checklist must complete part B as well. The project supervisor or consultant has delegated authority to provide approval in such cases that are considered to involve MINIMAL risk. The approval may be ***provisional*** – *identifying the planned work with human end user participants as* likely to involve MINIMAL RISK. In such cases you must additionally seek ***full approval*** from the supervisor or consultant as the project progresses and details are established. You must obtain ***full approval*** in writing, before recruiting and engaging with human end users participants for your project.

A.1 If you answer YES to any of the questions in this block, your consultant/supervisor must have obtained approval for the project from an appropriate external ethics committee, and you need to have received written confirmation of this from him/her. Students cannot themselves apply for ethics approval in this case as the project is considered high risk". This type of research is not covered by City's	*Delete as appropriate*

process, and external approval from an appropriate institution is required.		
1.1	Does your research require approval from the National Research Ethics Service (NRES)?	**NO**
1.2	Will you recruit participants who are covered by the Mental Capacity Act 2005?	**NO**
1.3	Will you recruit any participants who are covered by the Criminal Justice System, for example, people on remand, prisoners and those on probation?	**NO**
A.2 If you answer YES to any of the questions in this block your consultant/supervisor must have obtained appropriate ethics committee approval		*Delete as appropriate*
2.1	Does your research involve participants who are unable to give informed consent? For example, people who may have a degree of learning disability or mental health problem, that means they are unable to make an informed decision on their own behalf.	**NO**
2.2	Is there a risk that your research might lead to disclosures from participants concerning their involvement in illegal activities?	**NO**
2.3	Is there a risk that obscene and or illegal material may need to be accessed for your research study (including online content and other material)?	**NO**
2.4	Does your project involve participants disclosing information about protected characteristics (as identified by the Equality Act 2010)? *For example: racial or ethnic origin; political opinions; religious beliefs; trade union membership; physical or mental health; sexual life; criminal offences and proceedings*	**NO**
2.5	Does your research involve you travelling to another country outside of the UK, where the Foreign & Commonwealth Office has issued a travel warning that affects the area in which you will study? *Please check the latest guidance from the FCO - http://www.fco.gov.uk/en/*	**NO**
2.6	Does your research involve invasive or intrusive procedures? These may include, but are not limited to, electrical stimulation, heat, cold or bruising.	**NO**
2.7	Does your research involve animals?	**NO**
2.8	Does your research involve the administration of drugs, placebos or other substances to study participants?	**NO**

A.3 If you answer YES to any of the questions in this block, then unless you are applying to an external ethics committee or the Senate Research Ethics Committee (SREC), you must apply for approval from the Computer Science Research Ethics Committee (CSREC) through Research Ethics Online - https://researchmanager.city.ac.uk/. Depending on the level of risk associated with your application, it may be referred to the Senate Research Ethics Committee (SREC).		*Delete as appropriate*
3.1	Does your research involve participants who are under the age of 18?	**NO**
3.2	Does your research involve adults who are vulnerable because of their social, psychological or medical circumstances (vulnerable adults)? This includes adults with cognitive and / or learning disabilities, adults with physical disabilities and older people.	**NO**
3.3	Are participants recruited because they are staff or students of City, University of London? For example, students studying on a particular course or module. If yes, then approval is also required from the Head of Department or Programme Director.	**NO**
3.4	Does your research involve intentional deception of participants?	**NO**
3.5	Does your research involve participants taking part without their informed consent?	**NO**
3.5	Is the risk posed to participants greater than that in normal working life?	**NO**
3.7	Is the risk posed to you, the researcher(s), greater than that in normal working life?	**NO**
A.4 If you answer YES to the following question and your answers to all other questions in sections A1, A2 and A3 are NO, then your project is deemed to be of MINIMAL RISK. **If this is the case, then you can apply for approval through your supervisor under PROPORTIONATE REVIEW. You do so by completing PART B of this form.** **If you have answered NO to all questions on this form, then your project does not require ethical approval. You should submit and retain this form as evidence of this.**		*Delete as appropriate*
4	Does your project involve human participants or their identifiable personal data? *For example, as interviewees, respondents to a survey or participants in testing.*	**NO**

Reference list

1. Hyndman, R.J. and Athanasopoulos, G. (2018). *Forecasting : Principles and Practice*. 2nd ed. Heathmont, Vic.: Otexts.

2. Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, [online] 13(3), p.e0194889. doi:https://doi.org/10.1371/journal.pone.0194889.

3. OpenAI (2022). *ChatGPT*. [online] chat.openai.com. Available at: https://chat.openai.com/.

4. www.kaggle.com. (n.d.). *Retail Data Analytics*. [online] Available at: https://www.kaggle.com/datasets/manjeetsingh/retaildataset?select=sales+data-set.csv [Accessed 10 Feb. 2024].

Appendix

Chatgpt prompt and response for the work plan section: -

Prompt: -

I need to create a workplan to follow for the following project:

Project Objectives

Main Objective: To give organizations precise and trustworthy estimates of future sales volumes, this project will create a sophisticated predictive analytics model for sales forecasting.

Sub-objectives:

1. Collect data:

Objective: The goal of this objective is to a reliable and full dataset that would help in the analysis of the sales of a certain company and if needed merging more than one dataset.

Testable: when the data is ready to be cleaned, explored and analyzed, this sub-objective will be deemed successful.

2. Exploratory data analysis and data cleaning:

Objective: the goal of this objective is to clean the data and make it suitable for analysis and then explore the data and its properties and try to identify patterns, trends and outliers that would soon help us with the forecasting.

Testable: successfully exploring the data and cleaning it will be enough evidence in identifying and handling anomalies.

3. Develop Time Series Forecasting Model:

Objective: Create a time series by applying cutting-edge predictive analytics methods like Prophet, SARIMA, and ARIMA.

Testable: The created model will be put to the test on the data, and pertinent metrics (such MAE, MSE, and RMSE) will be used to assess how accurate it is.

4. Integrate External Factors for Enhanced Predictions:

Objective: improve the forecasting model's predicted accuracy by integrating external inputs, such as economic data and promotions.

Testable: Through comparative analysis, the effect of outside factors on model correctness will be evaluated.

5. Implement Interactive Visualizations for Stakeholders:

Objective: put in place interactive visualizations that will give business stakeholders sales projections and insights.

Testable: The visualization’s usefulness, level of visualization, and efficiency in communicating anticipated data will all be taken into account when determining how successful it is.

6. Optimize Resource Allocation Based on Forecasts:

Objective: By using the generated sales projections, this project will help firms allocate resources (such as personnel, inventory, and manufacturing capacity) as efficiently as possible.

Testable: The degree to which the actual resource allocation matches the anticipated demands will serve as the benchmark for resource optimization success.

7. Evaluate Model Performance and Fine-Tune:

Objective: Using test data, this project will assess the generated model's performance and make any necessary adjustments.

Testable: The model's performance will be evaluated by comparing it to the test sales data and making any required modifications as needed.

Chat gpt response:

Work Plan:

The work plan for the sales forecasting project with predictive analytics is structured to systematically achieve the project objectives. The plan outlines key activities, their corresponding outputs, allocated resources, and start-to-end dates. The timeline is presented as a Gantt chart for visual clarity.

Activities:

1. Project Initialization (Weeks 1-2):

- Outputs: Project Definition Document
- Define project scope, objectives, and deliverables.
- Conduct initial literature review for relevant methodologies.
- Draft the Project Definition Document.

2. Data Collection Framework Design (Weeks 3-4):

- Outputs: Data Collection Framework

- Research and select suitable data collection methods.
- Design the data collection framework.
- Begin implementing data gathering mechanisms.

3. Data Exploration and Cleaning (Weeks 5-6):

- Outputs: Cleaned Historical Sales Data
- Explore historical sales data for patterns and outliers.
- Clean and preprocess the data for analysis.

4. Time Series Forecasting Model Development (Weeks 7-10):

- Outputs: Time Series Forecasting Model
- Implement time series forecasting algorithms.
- Fine-tune and validate the model with historical data.

5. Integration of External Factors (Weeks 11-12):

- Outputs: Updated Forecasting Model with External Factors
- Identify and integrate external factors influencing sales.
- Evaluate the impact on model accuracy.

6. Dashboard Design and Implementation (Weeks 13-16):

- Outputs: Interactive Sales Forecast Dashboard
- Choose a data visualization tool (e.g., Tableau).
- Design and implement an interactive dashboard.
- Integrate machine learning predictions into the dashboard.

7. Resource Allocation Optimization (Weeks 17-18):

- Outputs: Optimized Resource Allocation Guidelines
- Utilize the forecasts to develop resource allocation guidelines.
- Align inventory, staffing, and production capacity with forecasted demands.

8. Model Evaluation and Fine-Tuning (Weeks 19-20):

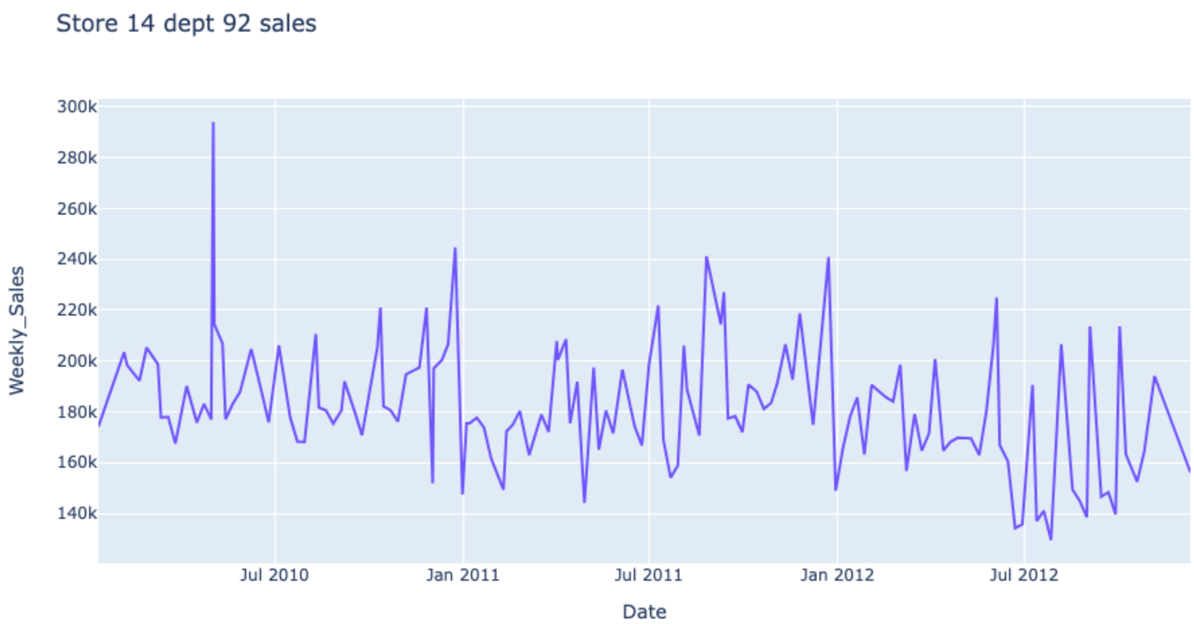
- Outputs: Evaluated and Fine-Tuned Forecasting Model
- Evaluate the model's performance on recent data.
- Fine-tune the model based on real-world outcomes.

9. Finalize Project Outputs and Documentation (Weeks 21-22):

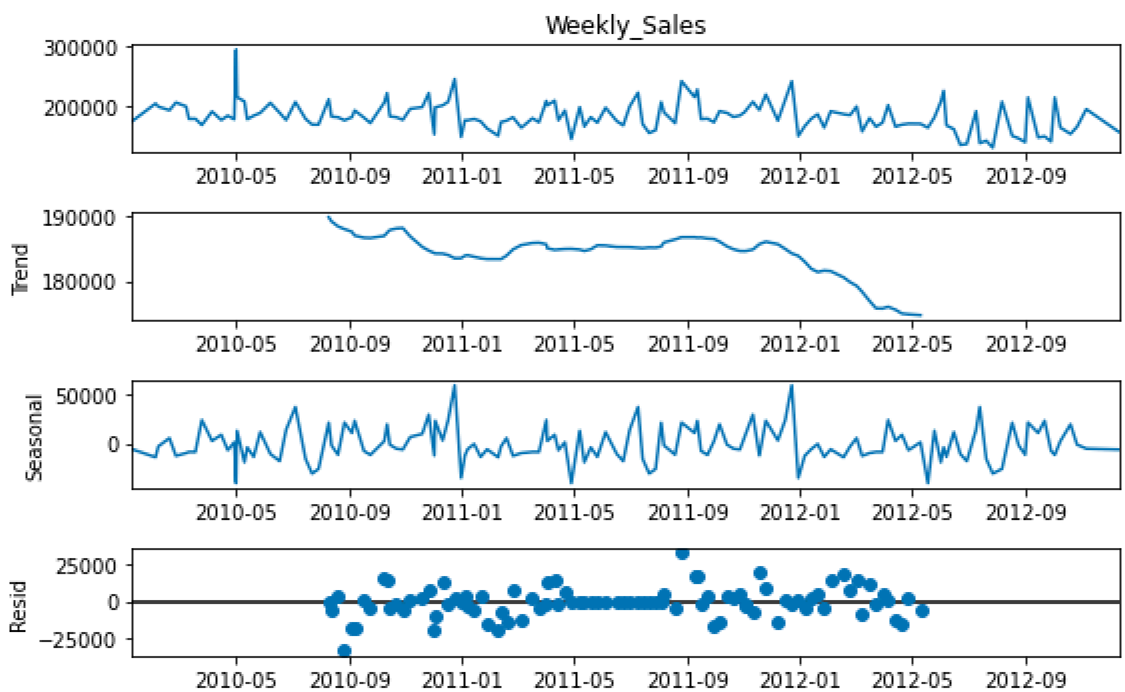
- Outputs: Final Project Submission, Documentation
- Finalize project outputs, including the interactive dashboard and model documentation.
- Prepare the Final Project Submission document.

Appendix: Figures not included in main report

Time series of the top performer weekly sales



Time series decomposition of the top performer time series



ADF Test for both datasets

Augmented Dickey-Fuller Test

In [27]:
```python
# Create a function that performs the ADF test
def adf_test(timeseries):
    result = adfuller(timeseries, autolag='AIC')
    print('p-value: %f' % result[1])
    print('Critical Values:')
    for key, value in result[4].items():
        print('\t%s: %.3f' % (key, value))
```

In [28]:
```python
# Apply the function to the aggregate_sales['Weekly_Sales']
adf_test(aggregate_sales['Weekly_Sales'])
```

```
p-value: 0.000000
Critical Values:
	1%: -3.477
	5%: -2.882
	10%: -2.578
```

In [29]:
```python
# Apply the function to the top_performer['Weekly_Sales']
adf_test(top_performer['Weekly_Sales'])
```

```
p-value: 0.001183
Critical Values:
	1%: -3.478
	5%: -2.883
	10%: -2.578
```

KPSS Test for both datasets

Kwiatkowski-Phillips-Schmidt-Shin Test

In [30]:
```python
# Create a function that performs the KPSS test
def kpss_test(timeseries):
    statistic, p_value, lags, critical_values = kpss(timeseries, 'c')
    print(f'KPSS Statistic: {statistic}')
    print(f'p-value: {p_value}')
    print('Critical Values:')
    for key, value in critical_values.items():
        print(f'\t{key}: {value}')
```

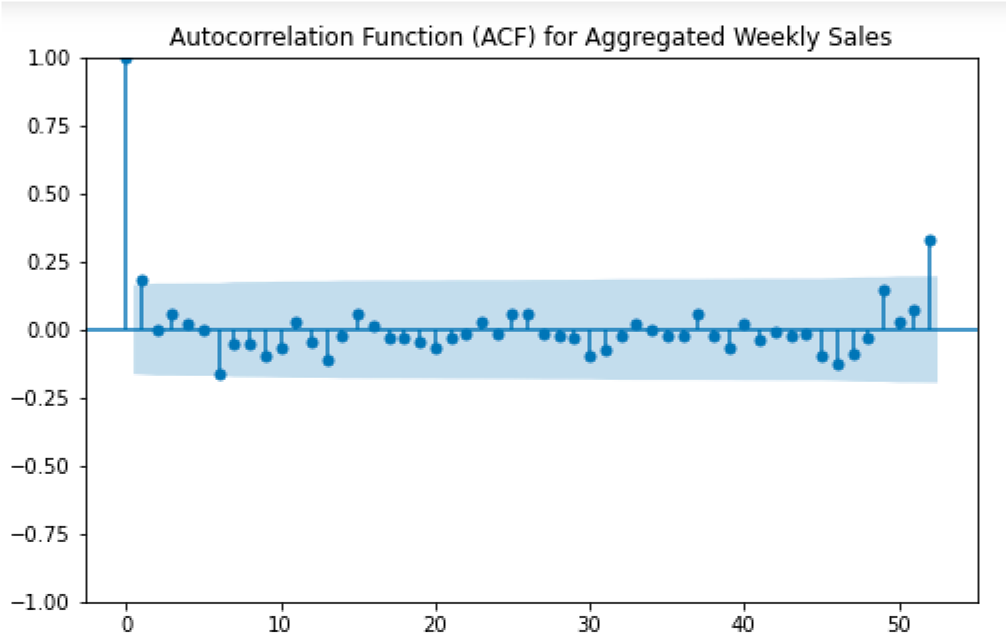
In [31]:
```python
# Apply the function to the aggregate_sales['Weekly_Sales']
kpss_test(aggregate_sales['Weekly_Sales'])
```

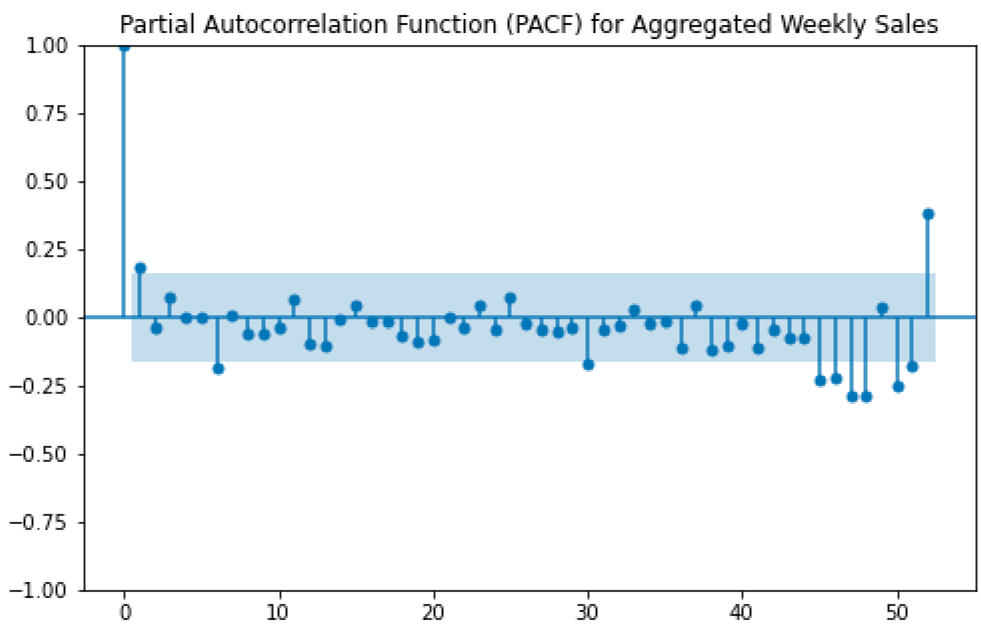
```
KPSS Statistic: 0.045319698887147895
p-value: 0.1
Critical Values:
	10%: 0.347
	5%: 0.463
	2.5%: 0.574
	1%: 0.739
```

In [32]:
```python
# Apply the function to the top_performer['Weekly_Sales']
kpss_test(top_performer['Weekly_Sales'])
```

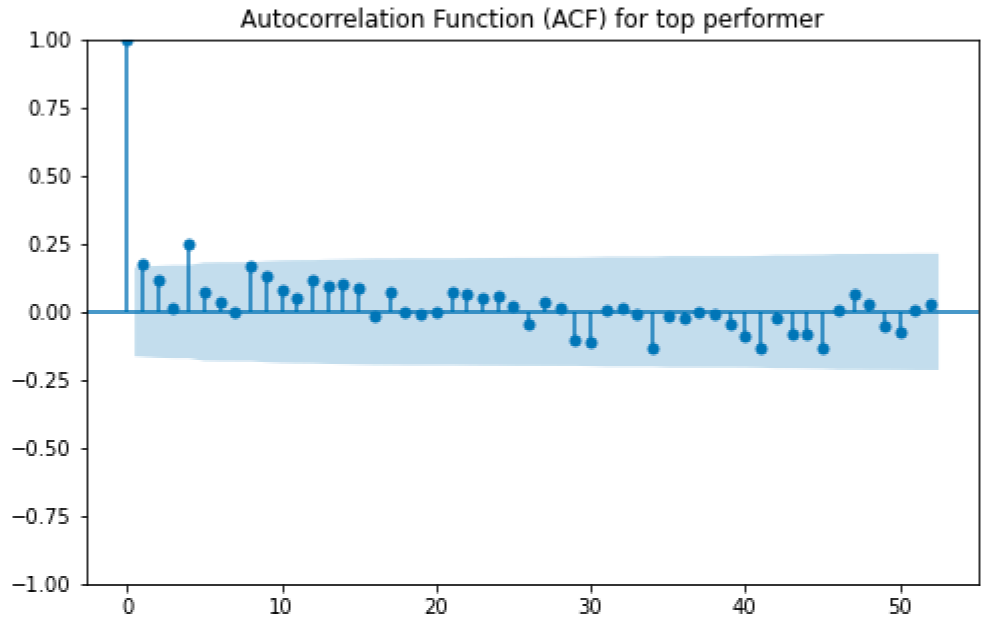
```
KPSS Statistic: 0.8702164672470161
p-value: 0.01
Critical Values:
	10%: 0.347
	5%: 0.463
	2.5%: 0.574
	1%: 0.739
```

ACF and PACF plots for aggregate sales dataset

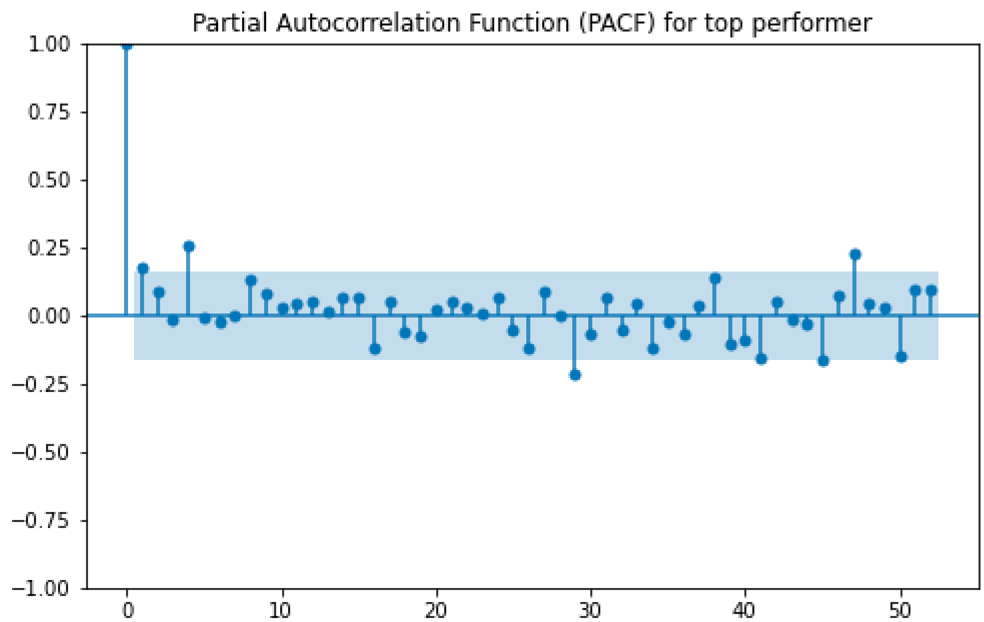




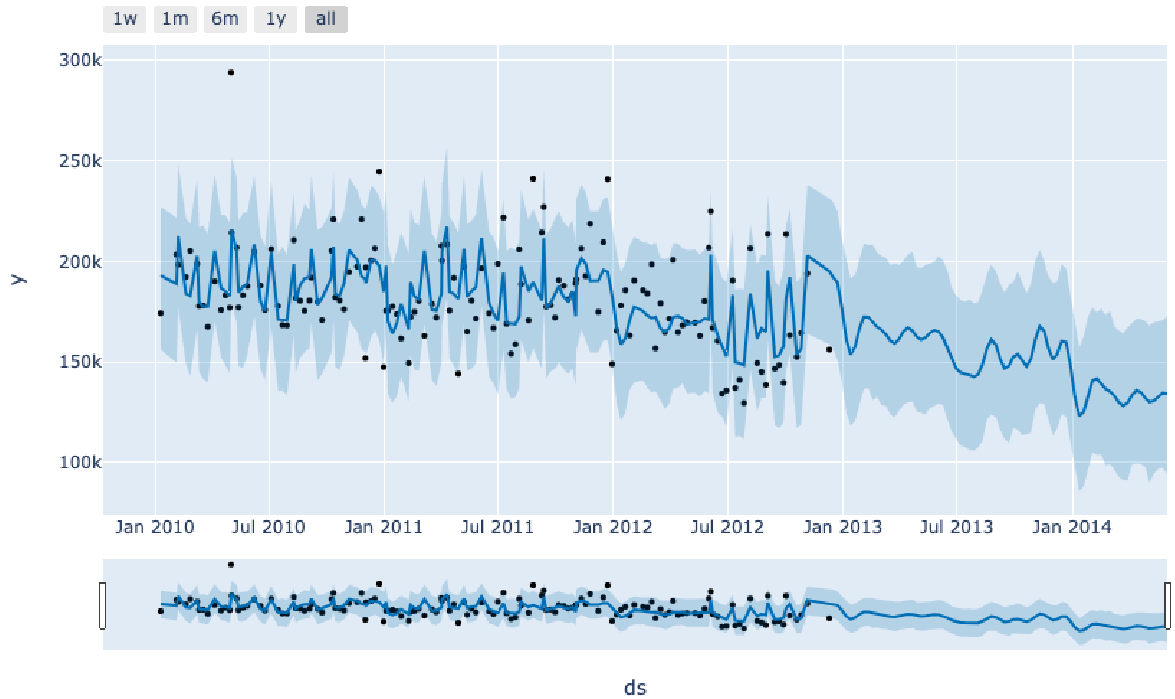
ACF and PACF plots for Top performer dataset

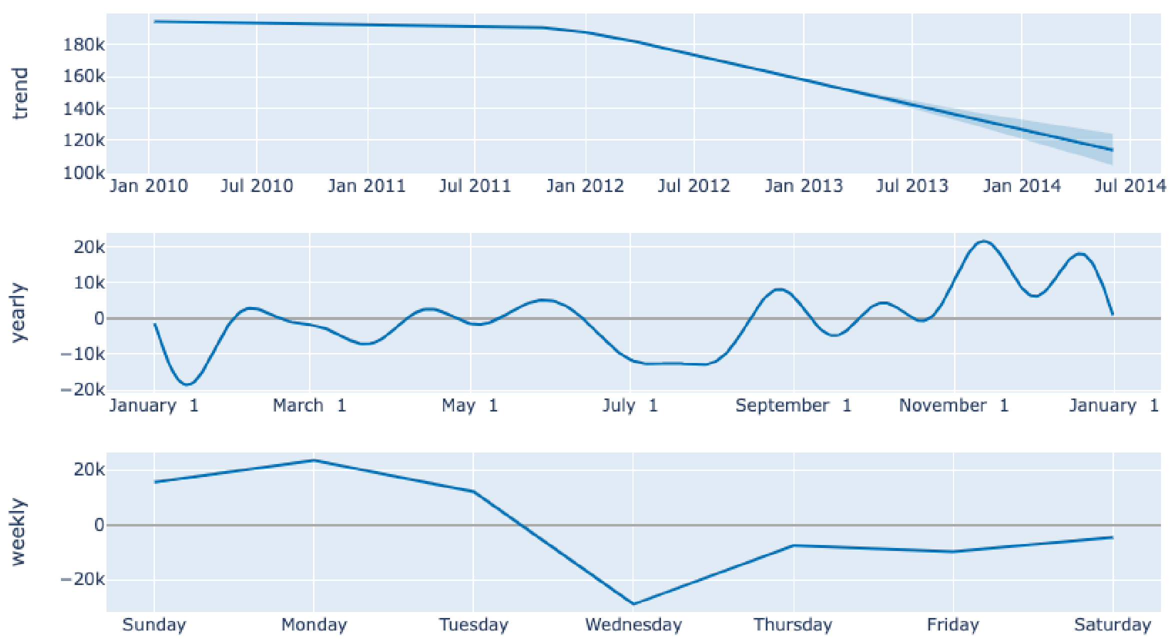


<Figure size 504x288 with 0 Axes>

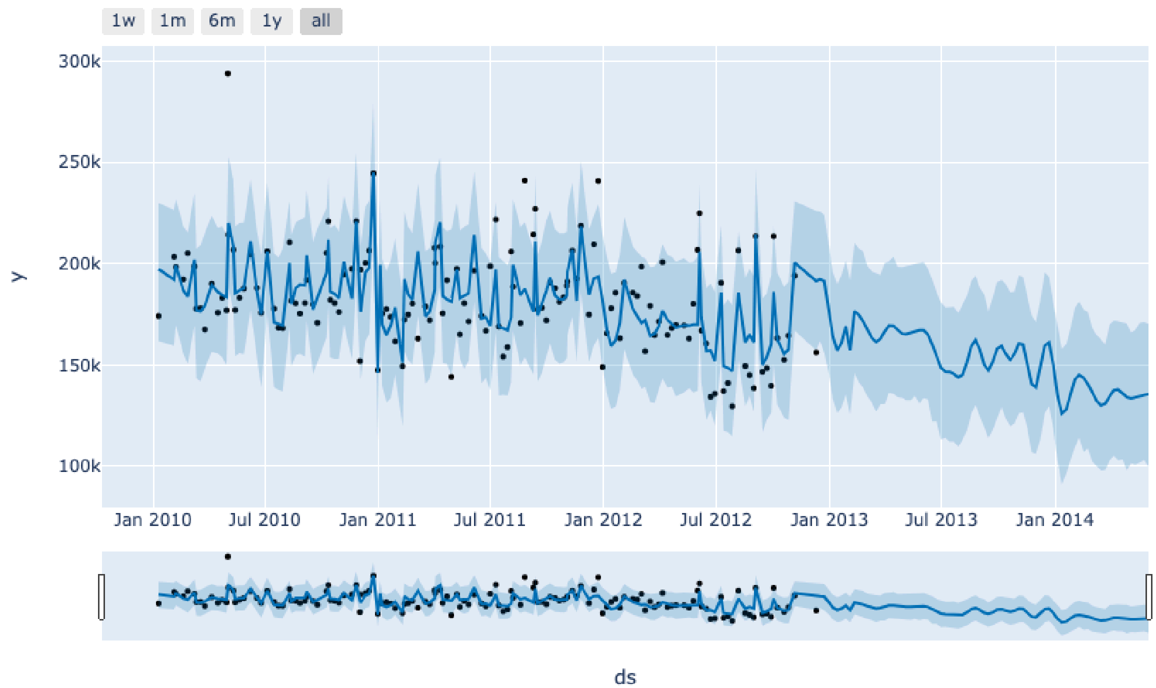


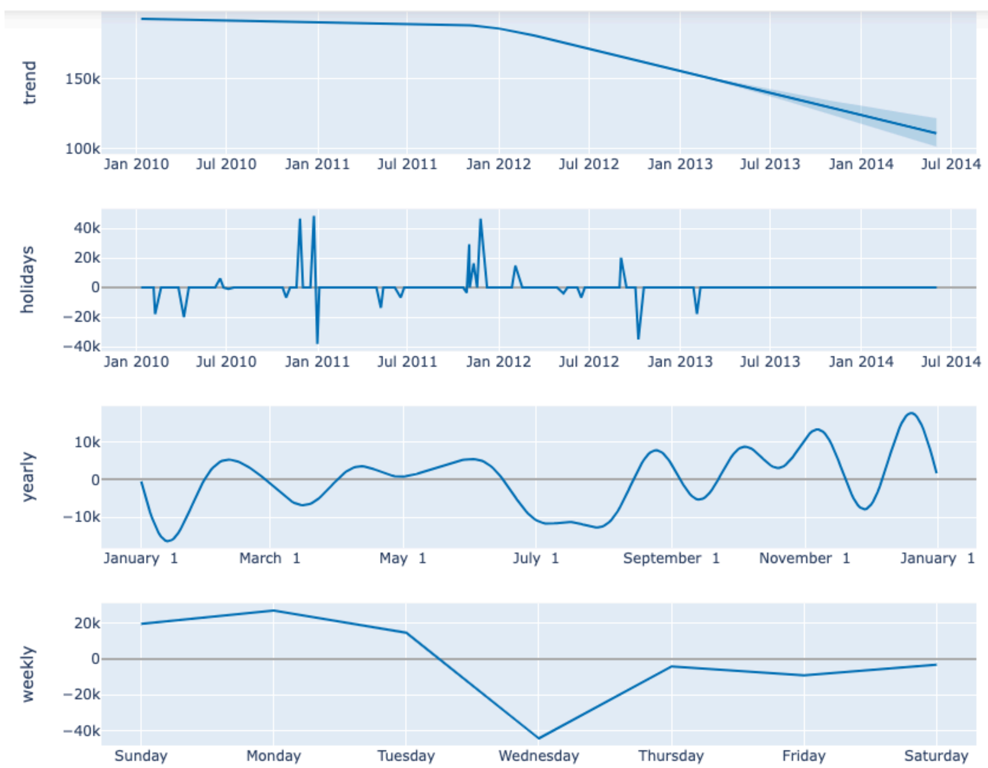
Prophet model for top performer dataset and decomposition



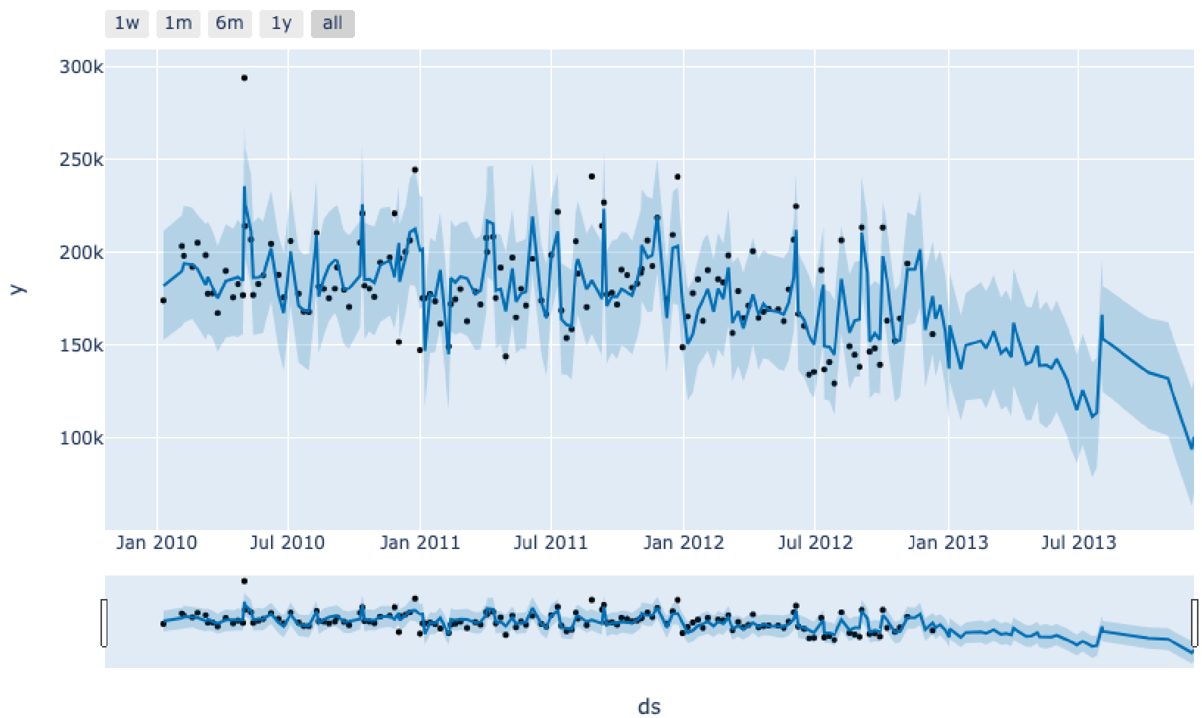


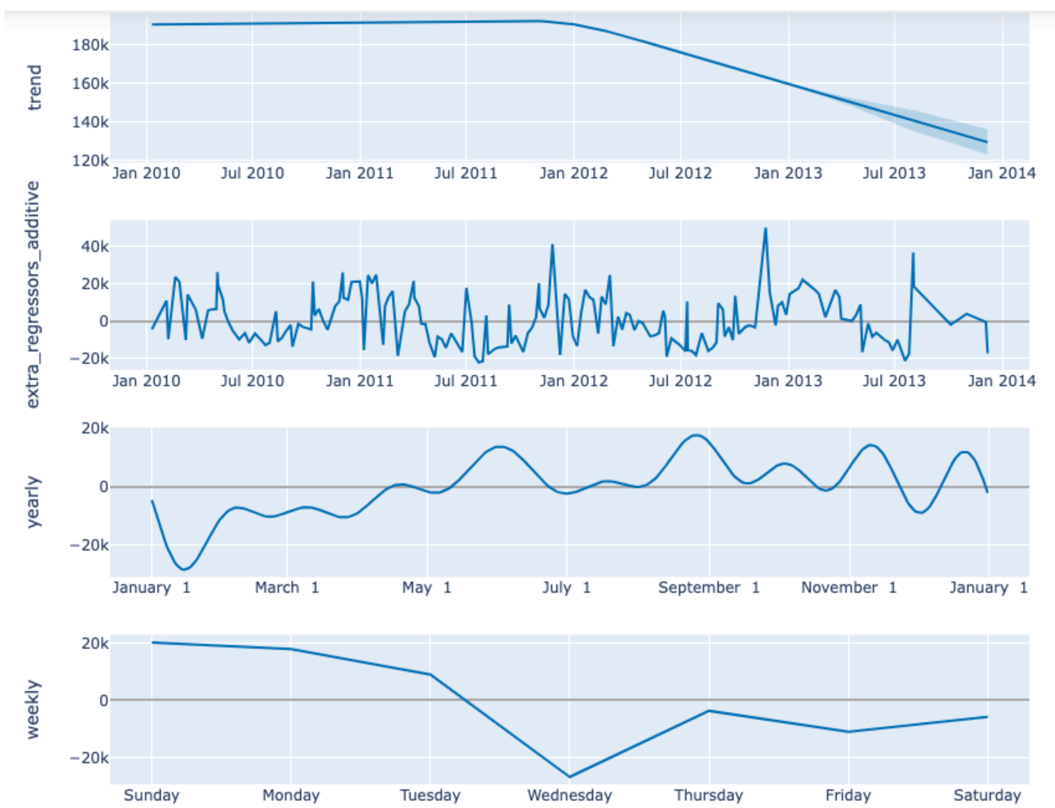
<u>Prophet model for top performer dataset and decomposition (including holidays)</u>





Prophet model for top performer dataset and decomposition (including regressors)


1w
1m
6m
1y
all
300k
250k
200k
150k
100k
y
Jan 2010
Jul 2010
Jan 2011
Jul 2011
Jan 2012
Jul 2012
Jan 2013
Jul 2013
ds


trend
180k
160k
140k
120k
Jan 2010
Jul 2010
Jan 2011
Jul 2011
Jan 2012
Jul 2012
Jan 2013
Jul 2013
Jan 2014
extra_regressors_additive
40k
20k
0
−20k
yearly
January 1
March 1
May 1
July 1
September 1
November 1
January 1
weekly
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday