Aris Dikran Nişan - 2020402075

IE 360 - Assignment 1 Report

Introduction

In this report, I present a comprehensive analysis conducted as part of my forecasting homework. The analysis is structured to explore, visualize, and model economic data, focusing on unemployment rates, inflation (TÜFE), exchange rates (Dolar), and commercial credit interest rates (ticari kredi faiz). The goal is to understand the relationships between these variables, their trends, and seasonality, and to build predictive models.

Data Preparation and Initial Exploration

I started by importing necessary R packages such as `data.table` for data manipulation, `lubridate` for date operations, and `forecast`, `ggplot2`, among others, for visualization and forecasting which we also used in our class & PSs. This setup ensures a smooth analysis workflow.

The dataset, stored in "EVDS.xlsx", was loaded into R using the `read.xlsx` function. Initial data exploration involved structure examination using `str(data)`, omitting missing values with `na.omit(data)`, and displaying the first few rows using `head(data)` to understand the dataset's composition.

Data Transformation

I made significant transformations to ensure the data's usability for analysis. Specifically, the 'Tarih' column was converted to a date format to facilitate time series analysis. Other relevant economic indicators (issizlik, Tufe, Dolar, ticari_kredi_faiz) were transformed into numeric formats. These steps were crucial for the subsequent analysis stages.

Visualization of Relationships

To visualize the relationships between time ('Tarih') and each economic indicator, I created line plots. These visualizations are instrumental in revealing the overall trends and guiding the more detailed analysis that follows.

Correlation Analysis

I employed the `ggpairs` function to generate pairwise plots of the dataset, providing a comprehensive overview of the relationships between all pairs of variables. This step is fundamental in identifying potential predictors for the modeling phase. I have chosen a intuitively correlated paris to perform a regression analysis.

Trend and Seasonality Analysis

To explore trends and seasonality, I added a trend component ('trnd') to the dataset and extracted the month from the 'Tarih' column as a categorical variable representing seasonality ('ay'). Visual analysis of these components was conducted through scatter plots with trend lines and faceted plots highlighting seasonal variations in unemployment rates.

Aris Dikran Nişan - 2020402075

Model Building

Two linear models were constructed to forecast unemployment rates. The first model included only the trend component, while the second model was more elaborate, incorporating the trend, seasonality, and other economic indicators (Tufe, Dolar, ticari_kredi_faiz) as predictors. Model diagnostics were performed to evaluate the residuals, ensuring the model's assumptions were met.

Prediction and Model Evaluation

I generated predictions using the developed models and visualized the actual vs. predicted values to assess model performance. This comparison is crucial for understanding the models' effectiveness and identifying areas for improvement.

Conclusion

This assignment allowed me to delve into economic data analysis, emphasizing the importance of thorough data preparation, visualization, and methodical modeling. The linear models developed offer insights into the factors influencing unemployment rates and demonstrate the value of incorporating trend and seasonality in forecasting models. The analysis underscores the interplay between economic indicators and provides a solid foundation for more advanced forecasting endeavors.

This report summarizes my journey and approach through this forecasting homework, highlighting the critical steps and decisions I made to unravel the intricate patterns within the economic data.