

ETW3420

Principals of Forecasting and Applications

Final Assignment

Semester 1, 2023

DUE DATE: Tuesday, 7 November 2023, 4.30pm

The unit learning objectives of this assignment are:

- Motivate the need for obtaining reliable forecasts in business and economics
- Understand and apply appropriate statistical methods for business and economic forecasting
- Develop computer skills for forecasting from business and economic time series data
- Provide practical insights from your forecasts

INSTRUCTIONS

1. This is an individual assignment worth 40% of your final mark for this unit. The total number of marks for this assignment is 40.

2. Make sure that you regularly make back-up copies of your work. Computer, disk, or cloud problems will not be accepted as valid reasons for late submissions or requests for extensions.

3. Students should pay particular emphasis on the narration, and how the results are presented and interpreted. Students should endeavour to ensure that the report is complete and well-composed.

Your report should be no more than 15 pages (excluding Graphs, Reference List, and Appendix). Any part of the report beyond the 15 page limit will be struck out and not marked.

- a) Use default format, paragraph, and margin settings.
- b) Font size: 12
- c) At least 1.2 line spacing between lines.
- d) Graphs should be appropriately sized and easy to read. They should not be made
- e) small to conserve space.
- f) Penalties may apply if the assignment does not conform to the formatting guidelines

5. With regards to graphs and estimation outputs:

- a) All graphs should be in-line with the text for ease of reading, and not placed in an Appendix at the end of the report.
- b) Any raw R output can be labelled and placed in an Appendix at the end of the report. Otherwise, if reporting any estimation output within the report, the output should be professionally presented in a table format.

6. All submissions will be via Moodle.

- a. **Put your full name, student's number, tutorial group number, and day as a file label.**
- b. If you choose to type your assignment in Microsoft Word, you will need to save it as a PDF file and submit (i) the PDF document, (ii) the Excel dataset (.csv format), and (iii) the R-script file consisting of the codes used to perform your analysis.

7. All submissions should be submitted with an Assignment Cover Sheet attached.

ASSIGNMENT TASK

Assignment Aim:

To quantify the forecasted tourist arrival in Malaysia using appropriate forecasting methods.

Instructions:

Proceed with the prescribed Forecasting Process below. Write a research report based on your analysis. Your report should include tables and graphs and an associated narrative. Keep the report concise and clear. Thoughtfulness, clarity of your discussion, and the communication of your results are important.

The Appendix does not constitute part of the 15-page limit.

Assignment:

Forecasting Monthly Tourist Arrivals to Malaysia

Introduction:

Tourism is a crucial sector in Malaysia's economy, ranking as the third largest contributor to the country's GDP. The tourism industry also plays a pivotal role as the second largest source of income from foreign exchange, making a positive impact on the nation's economic landscape. Moreover, it serves as a catalyst for job creation.

The tourism sector significantly bolstered employment in Malaysia, with a noteworthy 2.9% to 3.6 million individuals finding opportunities within it in 2019. This sector contributed a substantial 23.6% to the overall national employment figures. Notably, food and beverage serving services, accounting for a 34.7% share, and retail trade, with a 32.5% share, emerged as the key industries driving employment within the tourism sector during the same year.

In this assignment, our focus is on forecasting tourist arrivals in Malaysia. It's essential to recognize that exchange rates are a fundamental determinant influencing tourist arrivals.

In phase (1) of our analysis, we aim to estimate the impact of exchange rate fluctuations on tourist arrivals in Malaysia. We will illustrate how the depreciation of exchange rates can significantly influence the tourism industry in Malaysia. In phase (2), based on tourist arrival data in Malaysia, we will employ various forecasting methods to estimate tourist numbers, ultimately determining the most accurate approach for our forecasting needs. In phase (3), you will provide a conclusion regarding results obtained in phase (1) and phase (2) of your analysis.

-Download the data file “ Msia Tourist Arrival2.csv” for phase (1) and “Msia Tourist Arrival.csv” for phase (2) from Moodle.

-The first data set contains two series from 01/1/1989 to 01/12/2019, one of the series shows the monthly tourist arrivals in Malaysia while the other series shows the exchange of “RM to USD.”

-The second data set contains two series from 01/1/1989 to 01/1/2020 and shows the monthly tourist arrivals in Malaysia.

Phase (1): Linear Regression [8 Marks]

For this phase, you need to use the entire dataset to conduct your analysis.

- a) Using the entire dataset, plot the data and find the regression model for Tourist arrival with Exchange rate as an explanatory variable. Explain the relationship clearly and provide an interpretation of the constructed model. (2+1m)
- b) Conduct residual diagnostics for your produced regression model. Is the model adequate? (2m)
- c) Use the model to forecast the expected monthly tourist arrival if the exchange rate was 3.5 for January 2020 and 4.9 for February 2020 “RM to USD.” Assess if the forecasted value was reasonable and provide the reasoning behind your assessment. Additionally, present the forecast interval. (2m).
- d) Interpret the coefficients and results of the estimated model. (1m)

Phase (2): [28 Marks]

- a) Plot the time series plot of the tourist and insert the plot. Comment on 3 notable features present in the series. (3m)
- (b) How would you determine if seasonality is present in Malaysia’s monthly tourist arrivals?
List and explain 2 possible ways. (2m)
- (c) Take a subset of the data starting from January 2010 to January 2020. Label this as tourist2. (1m)

- (d) Partition the tourist2 dataset into a training and test set. You need to justify your selection and clearly specify the length of each set. (2m)
- e) Fit a piecewise linear regression model to the training set, train. Specify January 2014 to be the knot at which the trend should bend. Your piecewise linear regression model should also include seasonal dummies. (2m)
- (f) Generate fitted values from the piecewise linear regression model and produce a plot of the training set data with the fitted values. (2m)
- (g) Interpret the beta slope coefficients of the piecewise linear model. (2m)
- (h) This section will focus on producing forecasts for the time period spanning the test set generated from previous section. You will produce forecasts using the piecewise linear regression model, ensemble method, and the bagging procedure.
- I. Produce forecasts for the time period spanning the test dataset using the piecewise linear regression model estimated from the previous section. (2m)
 - II. Produce forecasts for the time period spanning the test dataset using an ensemble forecast comprising the ETS and ARIMA forecasts. In doing so, use the automated functions of ETS and ARIMA to estimate their respective model parameters from the training dataset, train. (1+1+2m)
 - III. Produce forecasts for the time period spanning the test dataset using the bagging procedure. (2m)
Note the following specifications:
 - Generate 10 bootstraps of the train data
 - Use the ets function to estimate the model parameters for the bootstrapped datasets
 - Use the set.seed() function to set the seed value as 123
 - IV. Produce the plot of the tourist2 dataset, along with the point forecasts from the piecewise linear regression model, ensemble method, and the bagging procedure. (2m)
 - V. Use appropriate methods to determine which of the above model/method produces the most accurate forecast. (4m)

Phase (3): Conclusion and policy implications [4 Marks]

Drawing upon the results of Phase (1) and Phase (2), compile a succinct conclusion.