

ETW3420

Principals of Forecasting and Applications

Individual Assignment 1

Semester 1, 2023

DUE DATE: Thursday, 14 September 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 25% of your final mark for this unit. The total number of marks for this assignment is 100.

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 endeavor 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 analyze the impact of the COVID-19 pandemic on US revenue passenger miles and quantify the forecasted loss in revenue passenger miles 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.

Download the data file “US revenue passenger miles.csv” from Moodle.

Stage 1: Define Goal [2 marks]

- Provide a brief background of revenue passenger miles as an airline traffic statistic in the United States by referring to appropriate academic/news article citations.
- Subsequently, provide a narration of how the COVID-19 pandemic has impacted the airline industry in the United States.
- In relation to the Assignment Aim above, define the goal of the forecasting problem that you are going to address in this report.

– You will gain a better understanding of the purpose of this forecasting exercise after completing this assignment. Hence, it is advised to write this part of the introduction after completing the other assignment tasks below.

Stage 2: Explore and Visualize Time Series [2 marks]

- Plot suitable time series graphics. Do not plot graphs for the sake of plotting them – be selective on the number and type of graphs you want to display.
- Describe in detail your time series graphics, making sure to identify any patterns or outliers. If applicable, are there any particular explanations as to why your data plots display certain interesting features (e.g. what event happened to explain a sudden jump or drop in data)?

Stage 3: Pre-Process Data [1 mark]

- Based on your data visualization, determine the choice of your time span. That is, determine how far back into the past you want to consider for this assignment.
- In other words, for the rest of the assignment, decide if you want to use the full data set from January 2000 to April 2023, or if you want to exclude a portion of the data whereby the resulting start date of your reduced data set is later than January 2000.

Be sure to justify your answer.

- Note that the last time point of your data set should still be April 2023. It is the starting date of your data set that you should decide on.
- The resulting data set from this stage will be referred to as the “full dataset”.

Stage 4: Partition Series [1 mark]

First, using your full dataset defined from Stage 3, partition your series into two parts labelled as “Pre-Covid” and “Covid”. (Covid period is from Feb 2020 until July 2021)

- The “Pre-Covid” part comprises of data prior to the start of the pandemic in the United States.
- Ensure that the “Covid” part comprises of data that coincides with the time frame when the United States was affected by the pandemic.
- Next, divide the “Pre-Covid” dataset into a training set and test set. Be sure to state the time periods over which your training and test sets span, respectively. Also report the number of observations in each of these sets.

Stage 5: Apply Forecasting Methods/Models [10 marks]

Using appropriate tools and methods, identify

- (i) ONE of the four simple forecasting methods from Topic 3 and (ii) TWO ETS exponential smoothing models

that you would consider using for forecasting. Justify your answers.

- Using the training set data, fit your identified forecasting method/models from above, and also use the `ets()` function in R to automatically select an ETS model. Report the ETS model selected by R.

You should have a total of FOUR forecasting method/models.

Note: In some cases, it is possible that your identified ETS model selected might coincide with the ETS model chosen by the `ets()` function. In this case, then you will have THREE forecasting method/models.

- For the exponential smoothing models, report the parameter estimates and initial values, along with some summary measures of the within-sample forecast residuals in a table. What can you infer from the magnitude of some of the estimated smoothing parameters?
- Produce separate plots of the training set data with the fitted values from each of the method/models. Ensure that your plots are correctly labelled. Comment on the plots.
- Perform residual diagnostic checks for all the method/models.
- Which of the above method/models has the best goodness of fit? Explain if this is what you expected.
- Use each of the above method/models to produce forecasts for the test set period. Subsequently, plot the “Pre-Covid” data set (i.e. both training and test set data), along with the forecasts produced by the 4 forecasting methods/models.
 - Do not include prediction intervals in the plot.
 - Make sure that your axes are labeled correctly, along with the chart title.
 - Make sure that you insert a legend in the plot, correctly labeling the various forecasts.
- Describe the forecast plots in detail, giving explanations as to why the test set forecasts produced by the above methods/models look as such.

Stage 6 : Evaluate and Compare Forecasting Performance [4 marks]

- Evaluate the out-of-sample forecast accuracy of the 4 forecasting method/models by using the traditional approach. Be sure to articulate the process on how you did this evaluation. Which method/models forecasts best?
- Evaluate the forecasting performance of the 4 method/models using time series crossvalidation with a forecast horizon of $h = 12$. Do you come to the same conclusion?
- Which forecasting method/model would you finally select for producing forecasts? Explain.

Stage 7: Implement Forecasts [3 marks]

Using your selected forecasting method/model from Stage 6, re-estimate the parameters using the “Pre-Covid” data set.

- Report the new estimated parameter values in a table, if applicable.
- Using the new set of estimated parameters, produce forecasts for the “Covid” time period.
- Produce a plot of the full dataset (see Stage 4) with the forecasts. Ensure the plot is labelled correctly.
- Describe the pattern of the forecasts, including the forecast intervals. Do they look reasonable?
- Compare the forecasts with the actual revenue passenger miles in the “Covid” time Period.

Stage 8 : Quantifying the Forecasted Loss in Revenue Passenger Miles [2 marks]

- Using all the results obtained so far,
 - Calculate the forecasted loss in revenue passenger miles (in billions) sustained by the United States since the start of the pandemic.
 - Calculate a range of forecasted loss in revenue passenger miles.
- To conclude the report, provide a summary of what has been done and some policy recommendations based on your results.