

ETW3420

Principals of Forecasting and Applications

Group Assignment 2

Semester 2, 2022

DUE DATE: Friday, 14 October 2022, 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

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

INSTRUCTIONS

1. This is an **DUO group assignment** worth 20% of your final mark for this unit. The total number of marks for this assignment is 80.
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. **Poor presentation, poor command of English writing and/or failure to comply with instructions may result in a mark penalty.**

4. 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 small to conserve space.
 - (e) 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 R or 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. Students must uphold academic integrity at all times. Any students caught for contract cheating, plagiarizing or permitting others to plagiarize their work will be reported to the Responsible Officer for academic misconduct in accordance to the Student Academic Misconduct Procedure. Severe penalties may apply resulting from the investigation.
7. All submissions will be via Moodle.
 - (a) 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.
 - (b) If you choose to type your assignment using Rmarkdown, you will need to submit (i) the **PDF** form of the assignment, (ii) the Excel dataset (.csv format), and (iii) the RMD file.
 - (c) You will also be required to put your assignment through a **Turnitin** report. The similarity index should not be more than 15%. Note that this is only a rough guideline - we understand that some common usage of phrases and sentences may contribute to the similarity index. Students should not be worried for this particular instance.

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

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder

ASSIGNMENT TASK

Assignment Aim

Use the ARIMA framework to quantify the forecasted loss in US revenue passenger miles since the COVID-19 pandemic.

Instructions

Unlike Group Assignment 1, this assignment will not utilize all the steps in the Forecasting Process. In particular, it will not consider Steps 1, 2, and 4 of the Forecasting Process (see Topic 1 Lecture Notes).

For this assignment, write a research report based on your analysis. Your report should include tables, graphs, and 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.

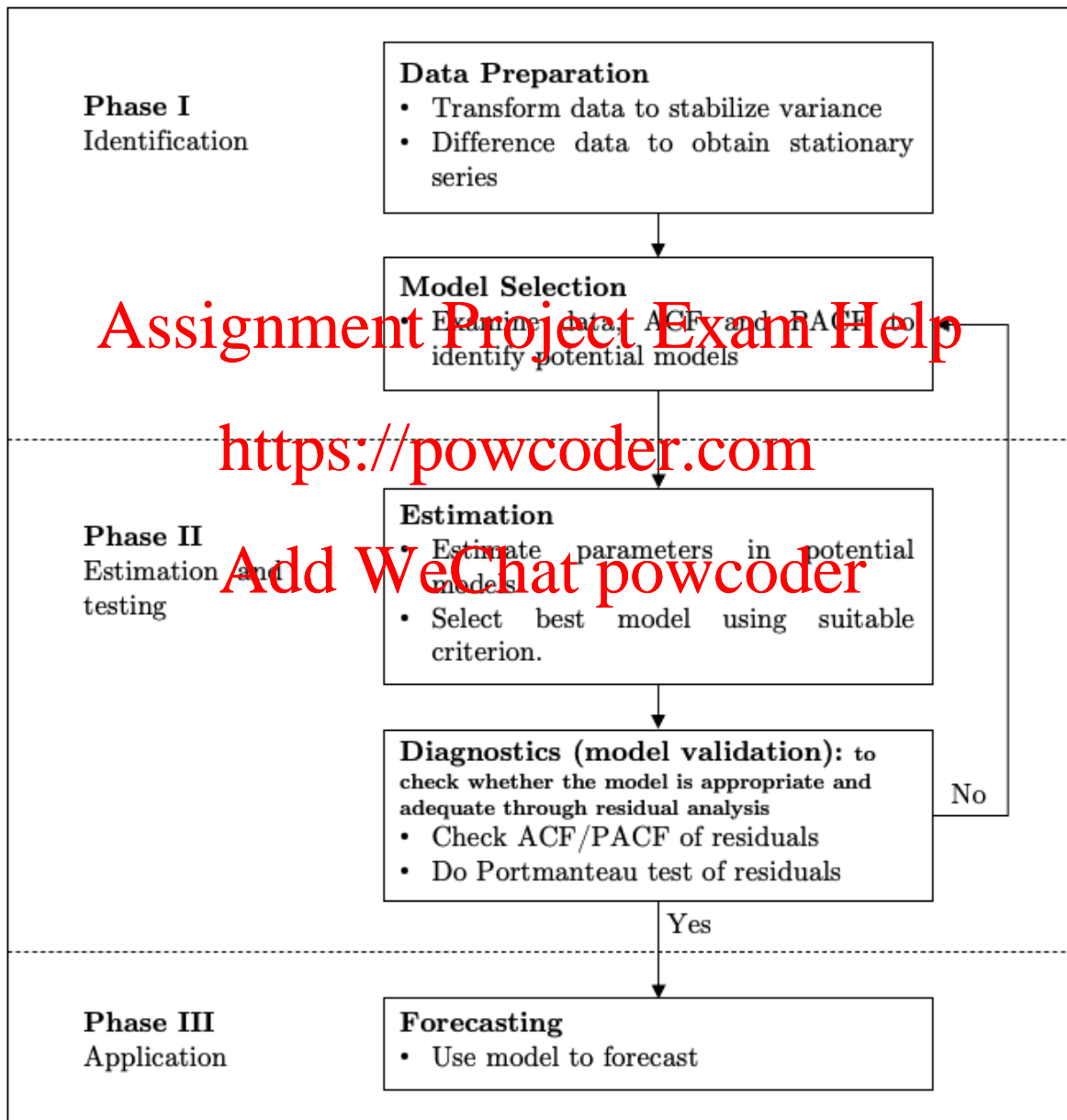
<https://powcoder.com>

Add WeChat powcoder

Method

The Box Jenkins methodology is a popular tool used in forecasting. Using the revenue passenger miles data from your Group Assignment 1, your task is to identify, estimate and select the best forecasting method from the ARIMA class, using the software R.

Your report should clearly highlight and explain the following phases of the Box Jenkins methodology:



Data Preliminaries

- For this Assignment, use the full dataset that you have obtained from “Stage 4: Pre-Process Data” of **Group Assignment 1**.
- As with Stage 5 of Group Assignment 1, partition your full dataset into 2 parts labelled as “Pre-Covid” and “Covid”.
 - 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.

Introduction [5 marks]

- Write a brief introduction to the report before commencing the tasks below. It should include the contextual background of the assignment.

Phase 1: Model Identification [30 marks]

- Using the “Pre-Covid” dataset, partition your data into a training and test set, specifying the time duration over which the respective sets span. (The training and test sets should be the same as your Group Assignment 1, unless they were commented as unsuitable in the feedback).
- Plot the training set data and determine if any transformation is necessary.
- If deemed necessary, transform the data using appropriate tools and techniques prior to analysis, taking heed to provide the relevant details.
- Using appropriate graphic plots for the (transformed) training set data, **identify at least TWO possible ARIMA models** for subsequent comparison purposes. Also, determine and explain if your specified ARIMA models should have a constant.
- Be sure to articulate the process and the reasoning behind the identification of the models.

Phase 2: Estimation and testing [20 marks]

- Estimate your identified ARIMA models from Phase 1 using the training set data.
- Estimate another ARIMA model using the `auto.arima()` function. Set the `stepwise` argument as `FALSE`.
- Report the parameter estimates of the ARIMA models, along with other relevant statistics, in a table.
- Rank the ARIMA models according to an appropriate information criterion. We will proceed to compare your **top 2** models.
- Perform diagnostic checks on the residuals for these 2 models. Determine if both models are adequate.
- Does any of your 2 models need to be re-identified? In other words, if the model residuals are not white noise, try a modified model by modifying the values of the AR and/or MA component. If your re-identified model still does not return a white noise residual, comment that the model does not return a white noise residual process and proceed on with the rest of the assignment. Note that you MUST articulate and demonstrate this process, and not reveal your final 2 models without first showing how you obtained them.
- Determine which of these 2 models is a better forecasting model and write out this model using backshift notation.

Phase 3: Application [10 marks]

- Re-estimate your selected model from Phase 2 using the Pre-Covid data. Report the estimated model using the backshift notation.
- Produce forecasts for the “Covid” time period.
- Produce a plot of the full dataset (i.e., “Pre-Covid” and “Covid”) with the forecasts. Ensure the plot is labeled 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.

Comparison with an ETS model [10 marks]

- Using the “Pre-Covid” dataset, produce forecasts for the “Covid” time period using the ets and forecast functions in R. (Let R select the ETS model automatically).
- Plot the full dataset, along with the ARIMA and ETS forecasts (with their corresponding prediction intervals) in the same plot. You may also play around with the ‘alpha’ argument to change the opaqueness of the prediction interval. Alpha values range from 0 to 1, with larger values indicating greater levels of opaqueness.
- Compare the 2 sets of forecasts and comment.
- Quantify and compare the forecasted loss in revenue passenger miles arising from the 2 models.

Conclusion [5 marks]

- Provide a brief summary of what was done and a report of the forecasted loss in revenue passenger miles.

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder