

Objective: To forecast a real time series using ETS and ARIMA models.

Data: The data are monthly measures of retail trade volume in Australia, [obtained from the ABS](#). Each student will use a different time series, selected using their student ID number as follows. This is the same series that you used in previous assignments.

```
library(fpp3)
get_my_data <- function(student_id) {
  set.seed(student_id)
  all_data <- readr::read_rds("https://bit.ly/monashretaildata")
  while(TRUE) {
    retail <- filter(all_data, `Series ID` == sample(`Series ID`, 1))
    if(!any(is.na(fill_gaps(retail)$Turnover))) return(retail)
  }
}
# Replace the argument with your student ID
retail <- get_my_data(12345678)
```

Assignment value: This assignment is worth 12% of the overall unit assessment. You may copy and paste material from previous assignments, but you must take into account any feedback that you have received on these assignments.

Report:

You should produce forecasts of the series using ETS and ARIMA models. Write a report of your analysis in Quarto format explaining carefully what you have done and why you have done it. You may use [this file](#) as a starting point. Your report should include the following elements.

- Produce some plots of your series, and describe what you learn from each plot. [2 marks]

- Discuss the statistical features of the data, including the effect of COVID-19 on your series. [2 marks]
- Find an appropriate Box-Cox transformation for your data and explain why you have chosen the particular transformation parameter λ . [2 marks]
- Produce a plot of an STL decomposition of the transformed data. What do you learn from the plot? [2 marks]
- What differencing would be required to make the data stationary? You should use a unit-root test as part of the discussion. [2 marks]
- Describe the methodology used to create a short-list of appropriate ARIMA models and ETS models. Include discussion of AIC values as well as results from applying the models to a test-set consisting of the last 24 months of the data provided. [6 marks]
- Choose one ARIMA model and one ETS model based on this analysis and show parameter estimates, residual diagnostics, forecasts and prediction intervals for both models. Diagnostic checking for both models should include ACF graphs and the Ljung-Box test. [8 marks]
- Compare the results obtained from each of your preferred models. Which method do you think gives the better forecasts? Explain with reference to the test-set. [2 marks]
- Apply your two chosen models to the full data set, re-estimating the parameters but not changing the model structure. Produce out-of-sample point forecasts and 80% prediction intervals for each model for two years past the end of the data provided. [4 marks]
- Obtain up-to-date data from the [ABS website](#) (Table 11). You may need to use the previous release of data, rather than the latest release. Compare your forecasts with the actual numbers. How well did you do? [4 marks]
- Discuss the benefits and limitations of the models for your data. [3 marks]
- Ensure graphs are properly labelled, including appropriate units of measurement. [3 marks]

ETC5550 students

- Suppose forecasts of your series were required every year based on the most recent data available at the time. Describe the steps you would undertake each year to produce these forecasts. [5 marks]

Notes

- Your submission must include the Quarto file (`.qmd`), and should run without error.

- There will be a 5 marks penalty if file does not run without error.
- When using the updated ABS data set, do not edit the downloaded file in any way.
- There is no need to provide the updated ABS data with your submission.

Due: 30 May 2025

Submit