

AAEC 4484/5484: Applied Economic Forecasting

Your Name Here

Homework #2 - Spring 2026

This assignment aims to enhance your understanding of time series and data patterns. It is intended to be straightforward.

Instructions:

Please ensure that your graphs and visuals have proper titles and axis labels. Recall that you can use `help()` or `?seriesname` in your consoles to get general information on the dataset.

Question 1: Basic Visualization of Time Series Data

- i. Create time plots of the following time series: Bricks from `aus_production`, Lynx from `pelt`, Open for FB from `gafa_stock`, Demand for 2013 from `vic_elec`.

Note: Use the `wrap_plots()` function to arrange your plots in a 2x2 grid.

- ii. Briefly discuss any discernible pattern(s) you noticed in the data.

Question 2: Random Data Sampling

In class, we explored the idea of drawing a random sample of data. Another practical operation is to randomly select a representative sample from a large dataset.

The `aus_retail` data set contains monthly retail turnover data for various industries and states in Australia. Your task is to randomly select an industry from this dataset and explore its time series patterns.

- i. In the code chunk below, set a seed of your choice to ensure a reproducible draw of the data. Next, use some combination of the `filter()` and `sample()` functions to draw a random sample of the data based on a single Series ID.
- ii. Use the `autoplot()`, `gg_season()`, `gg_subseries()`, and `ACF() |> autoplot()` functions to explore possible seasonality in your chosen sample. **Note: You can use the `theme(strip.text.y = element_blank())` to remove the y-axis labels in the `gg_subseries()` plot to make it look cleaner.**
 - Again use the `wrap_plot()` function with the `design` argument to arrange your plots as grids. **You are free to organize them however you wish.**
 - **It might be useful to change the `lag_max` (how about to 3 years of data?) in the ACF to ensure that you can see a fair bit of the pattern in the correlogram.**
- ii. What can you say about the series? Are there any seasonal patterns? Trends?

Question 3: White Noise

The aus_livestock series contains data on monthly Meat production in Australia for human consumption.

- i. Using the `filter()` function, extract the number of **Lambs** slaughtered in **Victoria**. Store this variable as `lambs`.
- ii. Produce the `autoplot()` of `lambs` and its correlogram. **Comment on any pattern noticed in both. Explain** whether this series look like white noise. **Hint: You would find it best to report at least 3 years in the lags of the ACF.**
- iii. Now, using the `difference()` and `mutate()` functions, create a new column, `diff`, that computes the **month-to-month** changes (`lag = 1`) in your `lambs` series. Store this new data as `d.lambs`.
- iv. Produce an `autoplot()` of `diff` along with the associated correlogram. Does the first differenced data now look like white noise? Did differencing remove any potential seasonality in the data? How about any trend? **Recall that a simple yes will not suffice. You will need to explain your conclusion.**
- v. Return to the `lambs` series and again use the `difference()` and `mutate()` functions to create a new column called `diff12` that computes the **year-on-year** changes (`lag = 12`) in your `lambs` series instead. Store this new data as `d12.lambs`.
- vi. Produce an `autoplot` of `diff12` along with the associated correlogram.
 - a. Does this new data, differenced at lag12, now look like white noise? **Recall that a simple yes will not suffice. You will need to explain your conclusion.**
 - vi. Return to the original `lambs` series and create a new column, `diff12_1`, that differences the `lambs` series at both lags 1 and 12, simultaneously (the order does not matter). Store this new data as `d12_1.lambs`.
 - a. Produce an `autoplot` of `diff12_1` along with the associated correlogram. Does this new data, differenced at lags 1 and 12, now look like white noise? **Recall that a simple yes will not suffice. You will need to explain your conclusion.**

Question 4: Gasoline Supplied in the US

You just got your first gig as a junior forecaster at the US Energy Information Administration (EIA) on the gasoline desk. You are given the `us_gasoline` dataset, which contains weekly data on the amount of gasoline supplied in the US (measured in million barrels per day, MBd) from Week 6, 1991 to Week 3, 2017. Your task is to explore the features of this series and identify any patterns that may be present.

Your boss stopped by and apologized for giving you the weekly data instead of the monthly data, but admitted that he is unable to locate that data. You are asked to aggregate the data to monthly totals.

1. Use the `mutate()`, `index_by()`, `summarise()`, and `filter_index()` functions to create a new dataset, `us_gas_monthly`, that contains the monthly totals of gasoline supplied from **March 1991 – Dec 2016**.
Hint: Keep the `Barrels` column as the total amount of gasoline supplied in a month.
2. Next, using the following graphics functions: `autoplot()`, `gg_subseries()`, `ACF()`, explore the features of `Barrels` from `us_gasoline_monthly`. **Hint:** You may want to use the `lag_max` of about 3 years (accounting for the data frequency) in the `ACF()` function to ensure you can see a fair bit of the pattern in the correlogram.
 - Can you spot any seasonality, cycles, and trend?
 - What do you learn about the series?
 - What can you say about the seasonal patterns?