Assignment 4 due 11/4 at midnight

Stat 416

Remember, you are allowed to use AI on this assignment. However, you must provide a link to a transcript. If you are working with a classmate, please include the name of that classmate.

Snowshoe hare furs

Consider the number of snowshoe hare furs sold by the Hudson Bay Company between 1845 and 1935. You can access the dataset, `pelt`, in `R` by loading the `fpp3` package.

1. Produce a well-labeled time series plot of the Snowshoe Hare furs.
2. Describe the plot produced in number 1.
3. Make an acf and pacf of the data. Suggest a model for the snowshoe hare fur *based on the acf/pacf*.
4. Fit the model you suggested in number 3.
5. Write the equation of the model you fit in number 4.
6. Provide and interpret 3-5 diagnostic plots or tests and comment on whether the model fits well.
7. **By hand**, calculate forcasts for the years 1936-1939.
8. Using R, calculate forecasts for the years 1936-1939.
9. Plot the forecasts on the same plot as the data, and interpret the forecast.
10. Use the automatic algorithm in the `ARIMA()` function to fit a model to the snowshoe hare time series.
11. Compare the models and provide a recommendation on choice of model (use a 2 pieces of "evidence" for your recommendation). Make sure to mention differences in forecasts, if any.

Transforming

1. For the following series, find an appropriate transformation and order of differencing to obtain stationary data.

- Turkish GDP data from `global_economy`
- Accommodation takings in the state of Tasmania from `aus_accommodation`.
- Monthly sales from `souvenirs`

2. For the transformed Turkish GDP series, perform a KPSS unit root test.

- Write the null and alternative hypotheses
- Include the code and output
- Interpret the results.

Australian Arrivals

Consider the `aus_arrivals` data set.

1. Use the information from the data documentation (see `?aus_arrivals`) to create a nice plot of the arrivals from just Japan.
2. Describe the trend and seasonal components (including the period) in the data, if any.
3. Use differencing to obtain stationary data.
4. Plot the acf and pacf of the data.
5. Identify and fully specify the order of a potential SARIMA model based on the ACF and PACF of the data.
6. Use the automatic model selection process in `ARIMA()` to select a model.
7. Plot the fitted values from the automatically fitted model over the series and comment on the quality of the fit.
8. Explain what the following code does and interpret the results (look a the documentation!!!).

```
# install.packages("microbenchmark") # run in console just once to install
library(microbenchmark)
library(fpp3)

start <- microbenchmark::get_nanotime()
  aus_arrivals |>
  filter(Origin == "Japan") |>
```

```r
  model(ARIMA(Arrivals, approximation = TRUE))
time_approx <- microbenchmark::get_nanotime() - start

start <- microbenchmark::get_nanotime()
  aus_arrivals |>
  filter(Origin == "Japan") |>
  model(ARIMA(Arrivals, approximation = FALSE))
time_noapprox <- microbenchmark::get_nanotime() - start

(time_noapprox - time_approx)/1e9
```

7. Can you use AICc to compare you manually chosen model to the automatically chosen model?

Successful upload

- Remember to use `embed-resources: true` in the options at the top of the document if you are using .html output (you may also turn in a pdf)
- Check that your uploaded file contains all the images and latex math you expect to see.