Exercise 2

Time series graphics

Prof. Dr. Sarah Brockhaus

Exercises and solutions are taken from

Hyndman, R.J., & Athanasopoulos, G. (2021) Forecasting: Principles and Practice, 3rd edition, OTexts: Melbourne, Australia. OTexts.com/fpp3, see https://otexts.com/fpp3/graphics-exercises.html.

Ex 1

Download the file `tute1.csv` from the book website, http://OTexts.com/fpp3/extrafiles/tute1.csv, open it in Excel (or some other spreadsheet application), and review its contents. You should find four columns of information. Columns B through D each contain a quarterly series, labelled Sales, AdBudget and GDP. Sales contains the quarterly sales for a small company over the period 1981-2005. AdBudget is the advertising budget and GDP is the gross domestic product. All series have been adjusted for inflation.

You can download the data and read it into R with the following script:

```
download.file("http://OTexts.com/fpp3/extrafiles/tute1.csv", tute1 <- tempfile())
tute1 <- readr::read_csv(tute1)
```

a. Convert the data to the time series data format `tsibble`.

b. Construct time series plots of each of the three time series using `autoplot()` and using `ggplot()` directly.

c. Construct a scatter-plot of Sales and AdBudget.

Ex 2

a. Download `tourism.xlsx` from the book website, http://OTexts.com/fpp3/extrafiles/tourism.xlsx, and read it into R using `read_excel()` from the `readxl` package.
b. Create a tsibble which is identical to the `tourism` tsibble from the `tsibble` package.
c. Find what combination of `Region` and `Purpose` had the maximum number of overnight trips on average.
d. Create a new tsibble which combines the Purposes and Regions, and just has total trips by State.

Ex 3

Monthly Australian retail data is provided in aus_retail. Select one of the time series as follows (but choose your own seed value):

```r
set.seed(12345678)
myseries <- aus_retail |>
filter(`Series ID` == sample(aus_retail$`Series ID`,1))
```

Explore your chosen retail time series using the following functions:

```r
autoplot(), gg_season(), gg_subseries(), gg_lag(), ACF() |> autoplot()
```

Ex 4

Use the following graphics functions: `autoplot()`, `gg_season()`, `gg_subseries()`, `gg_lag()`, `ACF()` and explore features from the following time series:

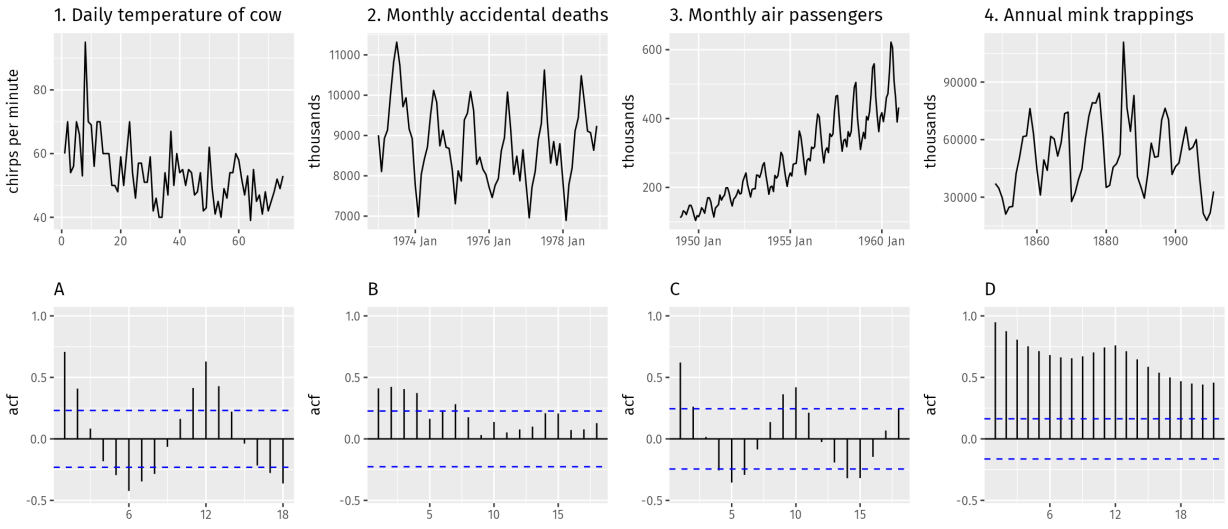
- "Total Private" `Employed` from `us_employment`,
- `Bricks` from `aus_production`,
- `Hare` from `pelt`,
- "H02" `Cost` from `PBS`, and
- `us_gasoline`.

You can use the following questions:

- Can you spot any seasonality, cyclicity and trend?
- What do you learn about the series?
- What can you say about the seasonal patterns?
- Can you identify any unusual years?

Ex 5

The following time plots and ACF plots correspond to four different time series. Your task is to match each time plot in the first row with one of the ACF plots in the second row.



Ex 6

a. Use the following code to compute the daily changes in Google closing stock prices.

```r
dgoog <- gafa_stock |>
  filter(Symbol == "GOOG", year(Date) >= 2018) |>
  mutate(trading_day = row_number()) |>
  update_tsibble(index = trading_day, regular = TRUE) |>
  mutate(diff = difference(Close))
```

b. Why was it necessary to re-index the tsibble?
c. Plot these differences and their ACF.
d. Do the changes in the stock prices look like white noise?

Ex 7

Use the data from your case study and answer the following questions:

- Can you spot any seasonality, cyclicity and trend?
- What can you say about the seasonal patterns?
- Can you identify any unusual years or days?

- Visualize the data using, e.g., the following plots:
  - Time plot
  - Seasonal plot (yearly and weekly; maybe monthly?)
  - Seasonal subseries plots for monthly aggregated data
  - Autocorrelation plots for daily and monthly data
- Tidy up your code, such that you would be happy to share it in class. And make sure that your code can be easily adapted to run on the data-sets from the other case studies.