Social Forecasting — Assignment 1

Thomas Chadefaux

When submitting your answers, please 1) show your code and 2) make sure your presentation is clear and easy to follow. Rmarkdown will make it easy to produce a suitable report, but you are welcome to use another approach if you prefer.

1. Import in R the dataset AFG_IRL_refugeeData.csv. Call the object "rd". The data include information on the number of refugees from Afghanistan to Ireland from 1999 to 2021 (one observation per year), as recorded by the UNHCR.
2. Create a time series object out of the column "refugees under UNHCR mandate". Call this object rdts. Make sure the frequency of your object matches the data's. Print rdts and report it here.
3. Produce a suitable plot for this time series. Make sure your axes are properly labeled. Note that if using the function autoplot, you will first want to load (and install if necessary) the "forecast" and "ggplot2" packages.
4. Using only data from 1999 to 2015 (i.e., not the entire period available), produce and report forecasts for years 2016 to 2021 using the following methods:

- Average method
- Naive method
- Drift method

5. Produce a plot which displays all three forecasts combined. Make sure your axes, title and legend are appropriate. Also include what actually happened from 2016–21 in the plot.
6. Based on these plots, which method do you conclude is best? Can you suggest an improvement to that method that would improve the predictions?