Statistical Modelling : Project

Important Remarks : You can work the project in pair with a colleague or individually. In order to avoid duplicate projects, once you choose the dataset(s) to work on, please send me for agreement an email with dataset(s) and a short description of the problem. If the dataset(s) is/are appropriate and have not been taken by any other group, you will have my agreement ; if not, you have to change the data. The report of the project has to be send by email to `barbu@univ-rouen.fr`, in `pdf` format, by June 10, 2021, together with the associated `R` programs in `.txt` or `.R` format and the datasets used in the project. The programs have to be commented.

The objective of the project is to study a time series (or several time series) that you can freely find on internet (see below some comments on available datasets). Starting from the dataset that you choose, you have to : (1) adjust one or several models using Holt-Winters' model and/or regression techniques and/or ARMA models ; (2) predict the evolution of the time series using the model that you have adjusted ; provide also confidence intervals.

Here you have some points that you can take into account, if necessary.

— Choice of the size of the sample : a too small sample size can give an estimation that is not reliable ; a too large sample can determine a change of regime in the data (more complex/difficult analysis, but more interesting).
— The first thing to decide is whether there is a change of pattern (regime) in your data. This can be done by simple graphical inspection of the data. If this is the case, one needs to propose different models for each region.
— If necessary, preliminary transform the data.
— It would be interesting to adjust several concurrent models on the same data and then to choose the best model, using a certain (information) criteria.
— Generally, it is not useful to make prediction at a horizon that is too large, because there is an important risk to be inaccurate.

Choice of the dataset : you can choose any time series dataset that you want. Some available datasets :

a. At `http://www.seriestemporelles.com` you have several nice datasets with description (in French) and associated programs.
b. At

   `http://users.stat.umn.edu/~kb/classes/5932/BDFiles.html#FileList`

   you have a large amount of datasets. The same can be found at

https://www.springer.com/fr/book/9783319298528#aboutBook

where you can download ITSM2000 Software package and any dataseta that appears with the extension .tsm can be copied/pasted and the extension changed in .txt (for instance, from airpass.tsm you get airpass.txt).