**Task description:**

The task is to **predict future energy use in a household** based on weather conditions by building time series forecasting models.

**Note**: The ultimate prediction objective is the **same** as in Assignment 1, however, the prediction methods you are expected to use are **different**.

You need to write **Python or R code** to**predict** energy use and **analyse** the impact of different factors based on your model.

**1. Download the**[**dataset**](https://myuni.adelaide.edu.au/courses/91968/files/14319102?wrap=1)[**Download dataset**](https://myuni.adelaide.edu.au/courses/91968/files/14319102/download?download_frd=1)and [code template](https://myuni.adelaide.edu.au/courses/91968/files/14319101/download) provided. Use this template for your assignment.

**2. Read** the [**data descriptionLinks to an external site.**](https://archive.ics.uci.edu/ml/datasets/Appliances+energy+prediction).

**3. Construct a code** in Python Jupyter notebook or R Notebook/Markdown. Python is preferred.

**4. Analyse and visualise the data** (word limit: **200 words**).

* Use suitable time series analysis techniques learned in [Module 6](https://myuni.adelaide.edu.au/courses/91968/pages/module-6-online-learning).
* **Discuss** your findings  and how can they be used to select a time series forecasting model.
* **Compare** this analysis with the one from your analysis in Assignment 1 and comment on any similarities. Consider whether the time series analysis techniques have given you more insights, and comment on what are those insights were.
* Include the **charts and diagrams together with the code**, e.g. in Jupyter Notebook.

**Note**: You can use the pre-processed data from your Assignment 1, if suitable.  
If for some reason you haven’t done it in Assignment 1, then consider applying the following processing techniques: scaling, feature selection and imputation of missing values.  
If any of these techniques are not suitable, explain your decision in less than 100 words. You may use other pre-processing techniques if needed.

**5. Based on your analysis:**

* **Use at least two time series forecasting methods** learned in [Module 6](https://myuni.adelaide.edu.au/courses/91968/pages/module-6-online-learning) (naïve method, seasonal naïve method, ARIMA, LSTM) and **justify your methods selection**.
* You can use the [paper](https://myuni.adelaide.edu.au/courses/91968/files/14319100?wrap=1)[Download paper](https://myuni.adelaide.edu.au/courses/91968/files/14319100/download?download_frd=1)mentioned below and the paper code in your assignment.  
  [Candanedo, LM, Feldheim, V & Deramaix, D 2017, 'Data driven prediction models of energy use of appliances in a low-energy house', *Energy and buildings*, vol. 140, pp. 81-97.](https://myuni.adelaide.edu.au/courses/91968/files/14319110?wrap=1)[Download Candanedo, LM, Feldheim, V & Deramaix, D 2017, 'Data driven prediction models of energy use of appliances in a low-energy house', Energy and buildings, vol. 140, pp. 81-97.](https://myuni.adelaide.edu.au/courses/91968/files/14319110/download?download_frd=1)  
  If you use the code from the paper, **clearly identify which part of the code is used** and where, and how it has been adapted to your task. You can also use common Python and R libraries.
* **If you use any other code** except the code from the seminar, workshop and the abovementioned paper, clearly identify the source (r.g. URL), which part of the code is used, what is the task that the code is used for, and how it has been adapted to your task.
* **Use training/testing methodology suitable** for time series and suitable model performance metrics.

**6.** **Test the models** using the same performance metrics as you selected in Assignment 1, and show the results for all models.

**7. Compare the results** from all candidate models, **choose the best** model, **justify your choice** and **discuss the results** (word limit: **200 words**).

* Show the results of all models in the form of suitable charts and tables.
* Select the best-performing model, show the final results for this model and justify your selection.
* Compare the forecasting results obtained in this assignment with the results you have obtained in Assignment 1. Make sure that the results are comparable in terms of metrics. Use conversion or adjustments if needed.

**8. Reflect** on what you have learned by completing this assignment (word limit: **200 words**).

**Submission requirements:**

You are required to **submit all the runnable code(s), analysis and results**. **Do not** include the dataset in the submission.

Submit one file (e.g., .ipynb), and do not zip it.

**Submitted file name must be in the form “<your id>\_<your\_name>\_assign2”.**