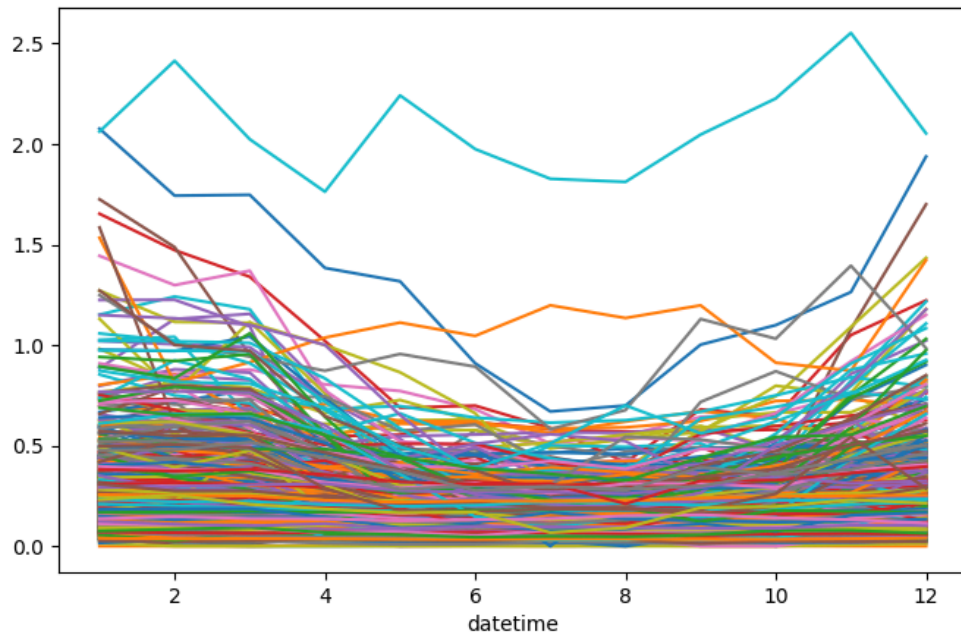


4. Baseline model performance

- Use your data in the proposed baseline model from last week.

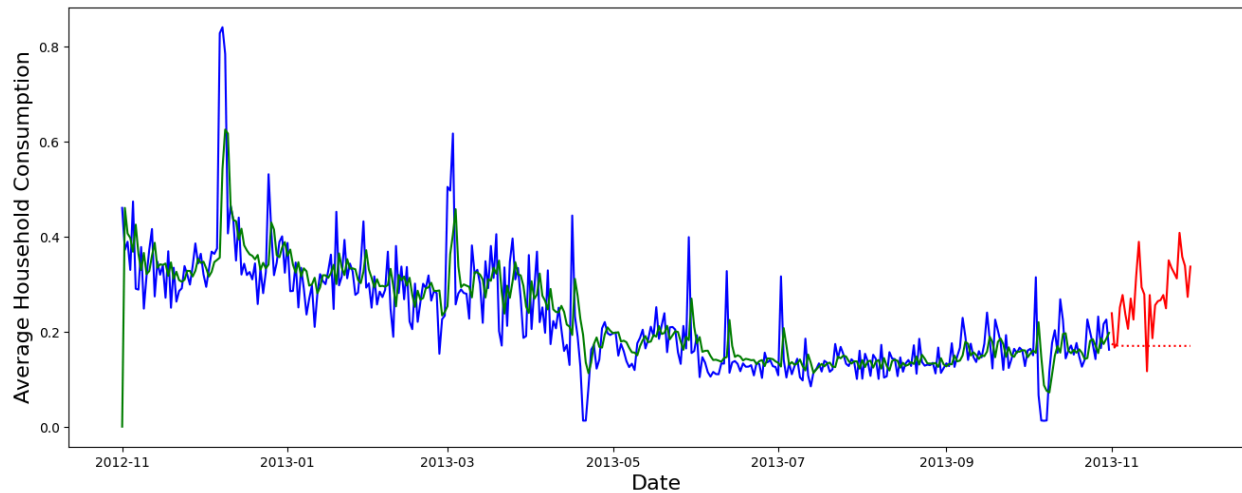
We fitted the data to auto_arima models for some households. In order to reduce the amount of memory needed, we computed the daily average of the 30 min granulated data and the baseline model aims to predict daily average energy consumption of the selected households.



The vertical axis represents Kwh. In the above figure, we determine the monthly average for all 900 UK households. We notice some outliers, such as the top household in blue, which we may remove from our dataset. We observe that energy consumption is greatest in winter months.

- List the performance metrics and any useful diagnostic figures

The forecasting in every model was unsuccessful. In particular, the model mimics the pattern of the training data (2012-11 to 2013-10) well, but performs poorly on the testing data (2013-11). We suspect that further fine-tuning is needed for the model and auto_arima might not be well-suited for this task.



Describe where the model worked well and where it didn't work well

Example: Model inference time took <5 minutes. Audio classification accuracy was 90% for “clean” data (no background noise), but was lower in accuracy for audio data with background noise. Across different background noise types (e.g., car horns, static, other people talking), mean accuracy was 60%, with a standard deviation of 25%

Without making huge simplifications, choosing the parameters and fitting the models was surprisingly slow and ineffective. Without investing a lot of effort into the baseline model, the forecasting seems extremely poor.

On the good side, a few adjustments should make the model applicable to independent households that have not been considered in the data.