Energy Forecasting with minimal data

Team Members:

Arjun Singh Bidesi
Dinkar Juyal
Lawrence Zhang
Reehan Shah

Cold start prediction for energy usage of new buildings

- Energy forecasting is a critical problem, with ramifications ranging from environmental to financial
- Traditionally, algorithms used high volume of past data to predict future trends.
- We aim at tackling the more difficult problem of making energy predictions with limited data

Key areas of focus

- Performance metrics
- Cold-start prediction algorithm
- Feature Engineering and ARIMA Analysis
- Data Imputation
- User Interface

Input data description

Dataset 1: Historical consumption

- Time series data of consumption and temperature data identified by Building ID
- There are ~510K records at a day-level

Dataset 2: Building Metadata

- Captures information on building specifics such as surface area, temperature inside building, etc.
- There is data for ~1380 unique buildings

Dataset 3: Cold Start Test Data

- Test data used to start a forecast
- Includes metadata about prediction window as well as time series data on consumption
- There are ~112K records at a day-level

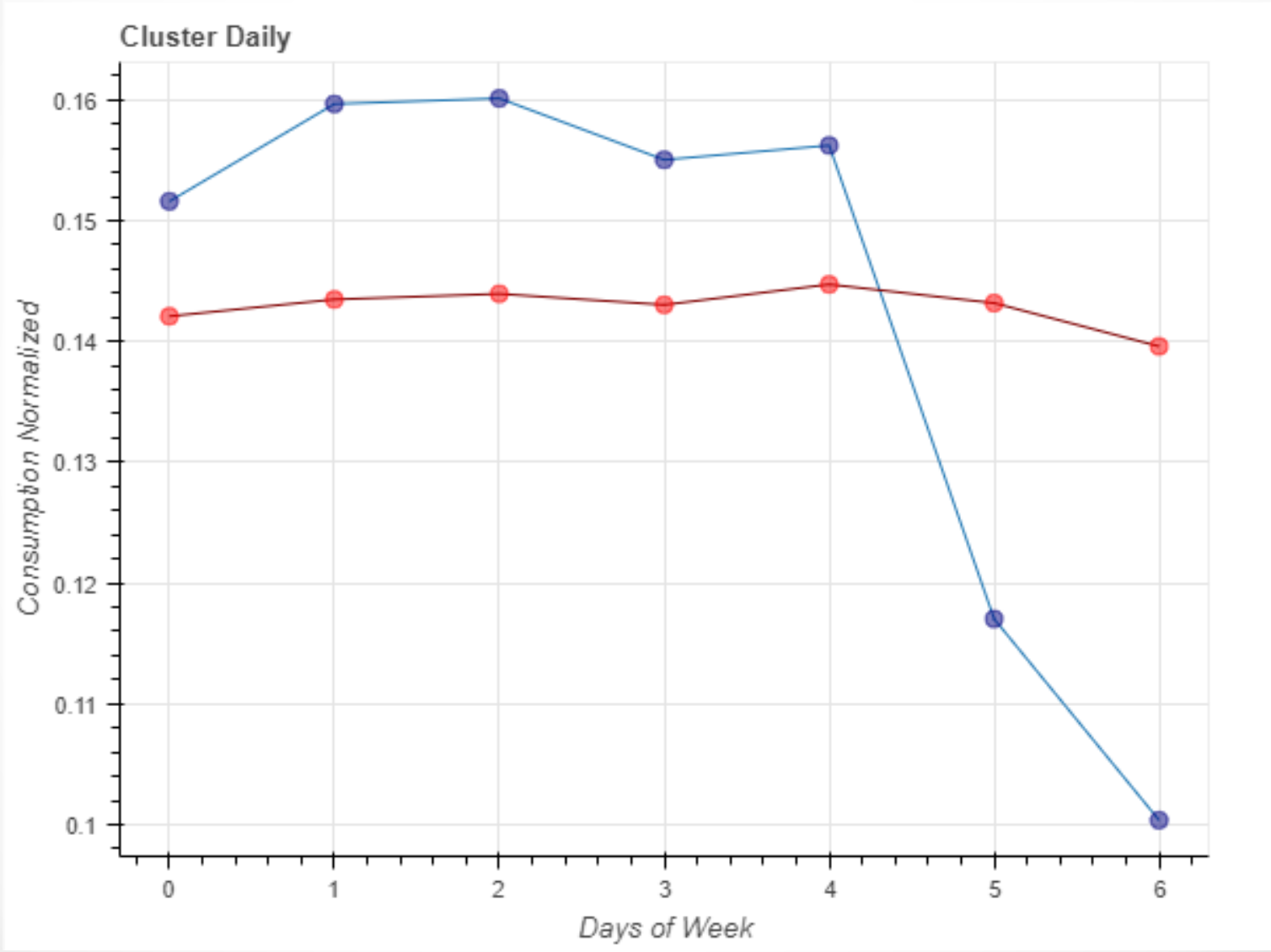
Modelling Approach

- Import data
- Raw data plots
- Data cleaning
- Data imputation
- Data factorization
- Data transformation
- Single variable analysis

- Stage the data
- ARIMA
- Time Series Regression (PanelAR)
- Deep neural networks
- LSTM

- Model selection
- Performance metric based iterative tuning
- Cross-validation
- Grid search

Data Exploration - Clustering

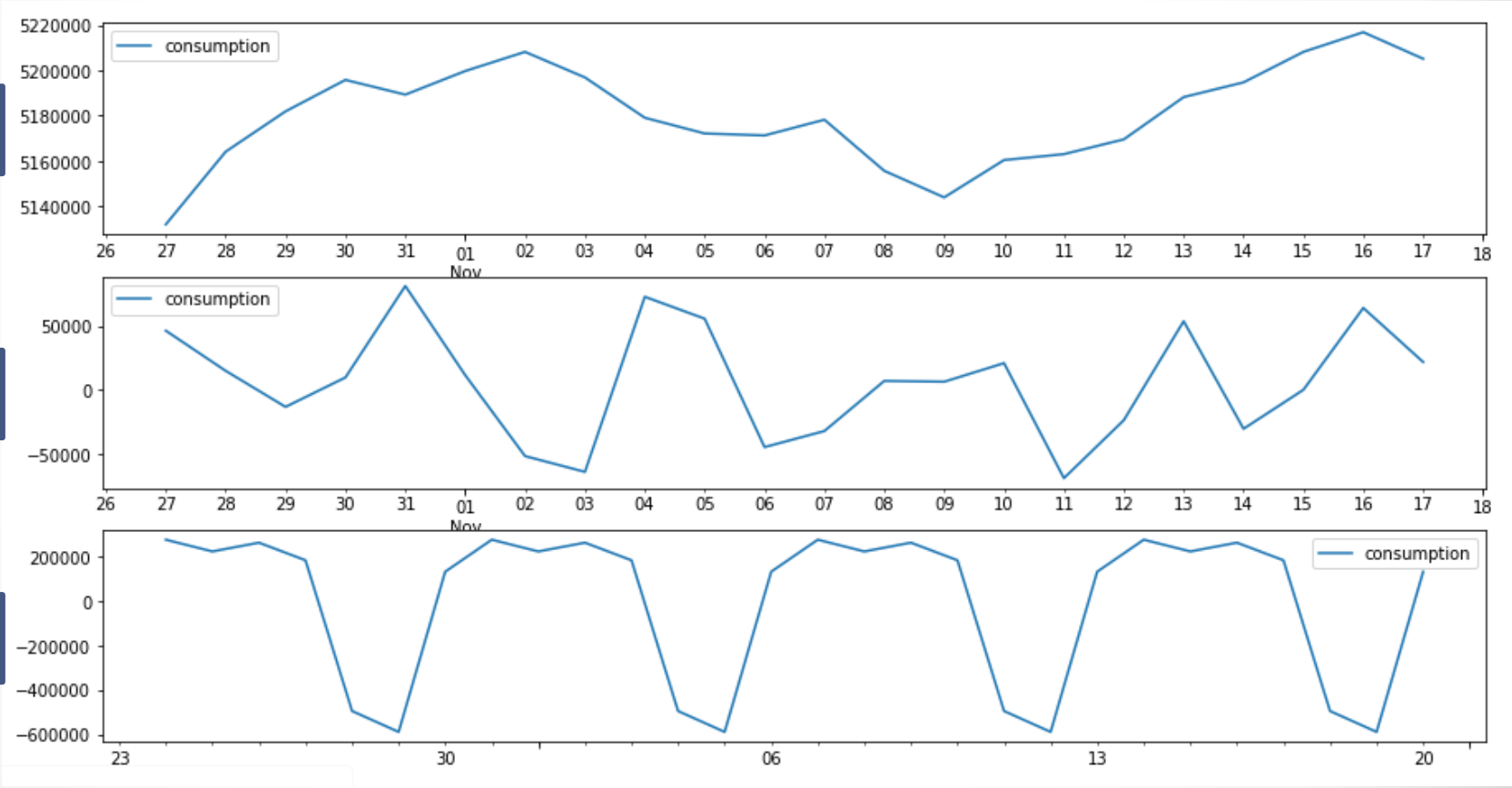


Data Exploration – Seasonal Decomposition

Trend

Residual

Seasonality



Models Used

Linear Regression

Model	R^2
Linear	0.88
Polynomial	0.48
Ridge	0.90
Lasso	0.89

Time-Series Models

Model	R^2
Panel OLS	0.87
ARIMA	0.6

Other Models

Model	R^2
Random Forest	0.95
AdaBoost	0.8
LSTM	0.52

Future work

- With more data, we can shift focus from our existing approach (linear models) to more sophisticated models
- Retraining including clustering will have to be done as we collect more data to ensure that any recent trend is incorporated
- Continue to build on existing UI