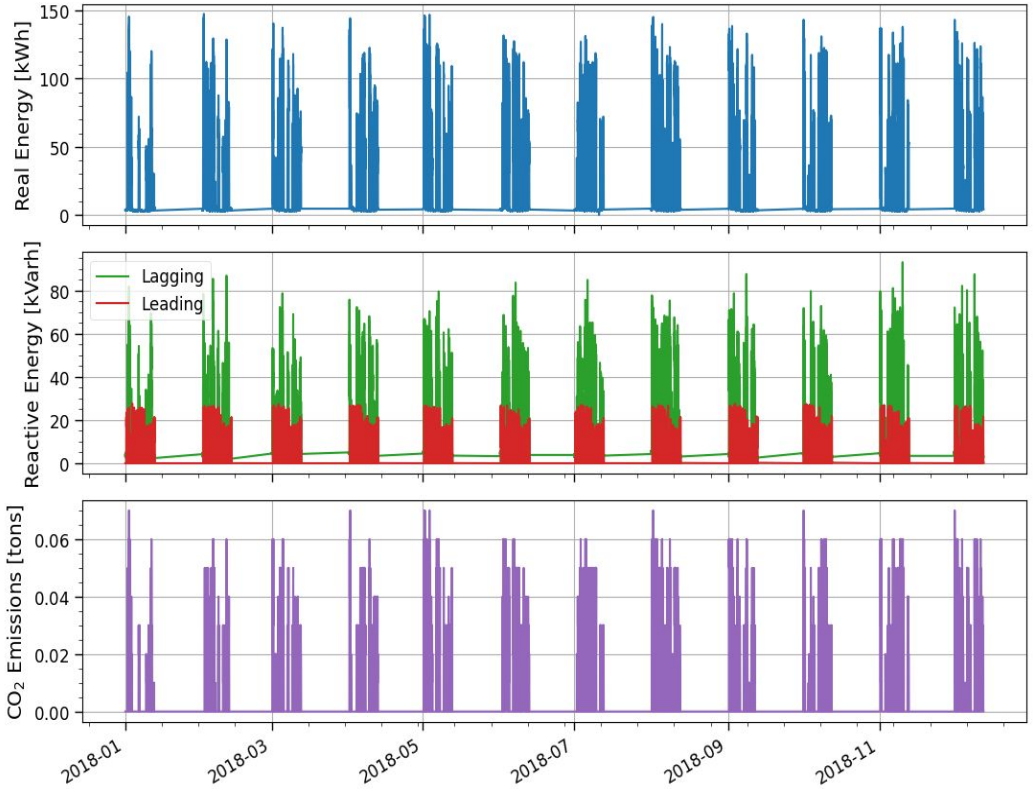
Steel Industry Data Exploration

A Quick EDA and Model Building

Time-Series - Full Data Set

A comprehensive time-series analysis covers the entire observation period, with a detailed view of 2018 segmented into two-week intervals, alternating with data gaps.

The units provided in the data set are units of energy (kWh and kVarh), from which the power is simply given from dividing by the integrated time of each measurement.



Time Series - One Week

Focusing on the week of February 2, 2018, clear correlations emerge between real energy usage, reactive energy lagging, and CO_2 emissions.

This motivates further examination of CO_2 emissions in the space of Real Vs. Reactive Lagging Energy…

Modeling CO_2 Emissions from Power

While other 2d spaces were also explored, the space of real vs. reactive lagging energies did prove to be the most revealing of the distribution of CO_2 emission.

★ **The CO_2 emission may potentially be modeled as a function by the real power P and the reactive lagging power Q_L.**

At a first glance, the data seem to reflect a Bivariate Gaussian distribution, though there is certainly more structure.