**Project proposal**

Basic programming

**Project topic:** Understanding the local electricity market (Datacamp competition)

**Background and inspiration:**

Electricity is becoming general use for our everyday gadgets such as cell phones, computers, robots, electric cars and some medical wearable devices, other than in the industrial. Solar power plants and solar power panels are becoming important energy generator, generating directly from the sunlight source. More residencies are looking forward to decreasing their electrical pay by reaching solar energy.

A solar cell is made of p-type and n-type silicon. The near junction of two layers n-type electron move into the holes on p-type layer, creates an area around the junction called the depletion zone in which the electrons fill the holes. When the sunlight strikes a solar cell, electrons in the silicon are ejected forming of holes. Connect the two silicon layers with metallic wire, the electrons will travel and create a flow of electricity.

Solar panel receives energy from the sunlight and converted into electricity. Meanwhile, the amount of energy produced depends on some factors such as the panel orientation and weather. Even though solar panel work more efficiently on the cold weather, summer which is the sunniest weather and has fewer clouds is more likely to produce electricity. In the raining season, cloud might obscure the sunlight, but rainstorm washout the panel dusts so they are able to produce more energy.

From Datacamp competition, the company build solar panel arrays and they sell the energy to industrial customers. Now they will expand sales to the city of Melbourne. They want to estimate energy prices for the next 12-18 months to negotiate with their customers and for their research strategy in storage capacity. Their plan is to store some of the energy produced by the solar panels when pricing conditions are unfavorable and sell it by the next day.

Competition source: https://app.datacamp.com/learn/competitions/victoria-energy-demand

**Expected scope of knowledge to solve the problem:**

* Numpy
* Pandas: dataframe and plot, data analysis
* Machine learning: Time series forecasting, Use of Scikit-learn regression model package
* Timeseries models: Recurrent Neural Network (RNN), Long Short-term Memory (LSTM), Multilayer perceptron (MLP), Auto Regressive Integrated Moving Average (ARIMA), Support Vector Regression, Convolutional Neural Network (CNN), XGBoost
* Matplotlib, Seaborn: data visualization

**Timeline things to do**

* Acquire data
  + Data source from Kaggle:   
    https://www.kaggle.com/datasets/aramacus/electricity-demand-in-victoria-australia
* Learn about time series forecasting models
* Extract test data, plot the demand to compare with recommended retail price and other
* Clean data
* Models training
  + Do multiple model trainings by using some of the time series model mentioned in the scope of knowledge to compare algorithms by forward chaining strategy or K-Fold cross validation.
* Models’ optimization and evaluation
  + Testing stage after running multiple models to compare testing data, choose the most accurate and optimal performance algorithm.
* Work on story telling
  + Describe data process and relate to the given situation and connect to the data recommendation results.
  + Clarify the data visualization
* End point: Plot and forecasting
  + Compare and plot the prices change throughout the year.
  + Look for price pattern in seasonal or month of the year.
  + Build a forecast of daily energy prices the company can use as the basis of its financial planning.
  + Provide guidance on how much revenue the energy storage venture could generate per year using retail prices and a 70MWh storage system.