**Nature and Sources of Data Set:**

The dataset has been downloaded from Kaggle. The below is the link for the dataset: <https://www.kaggle.com/datasets/taranvee/smart-home-dataset-with-weather-information/data>.

The primary dataset for this dissertation includes detailed records of household energy consumption along with weather conditions. The variables in the dataset are: timestamp, energy consumption of various household appliances (such as dishwasher, fridge, and furnace), overall household energy usage, solar energy generation, and weather parameters like temperature, humidity, wind speed, and cloud cover.

This data has been collected from smart meters and IoT devices in a residential setting. The weather data comes from local meteorological stations or online weather services and is synchronized with the energy consumption data.

This rich dataset allows for a thorough analysis, making it possible to identify patterns and relationships between energy usage and weather conditions. The high-resolution data also aids in developing accurate predictive models.

To maintain data integrity and quality, pre-processing will involve handling missing values, detecting outliers, and normalization. Data segmentation will be used to create training and validation sets for machine learning models.

Combining detailed energy consumption data with comprehensive weather information provides a strong foundation for research, enabling the development of advanced systems for predictive modelling.