Summer Research Objectives

* Discover findings for application on the new engineering building for energy efficiency, using Data Science tools and concepts such as Supervised Learning.
* Identify areas for further research opportunities

Research Process:

* Learning tools, Data Exploration, Data modelling
* Initial steep learning curve. What you learnt in the beginning, and how long you spent doing it. What the importance of doing this was and what you learnt by doing the self learning (specific techniques).

**Summer Research Objectives**

The democratization of data analysis algorithms and computer processing power give a greater ability to organize and make use of scores of data generated by different equipment used in a building. Although buildings’ electricity systems have evolved over several years, to incorporate several state-of-the art equipment, an understanding of intricate electricity systems is exclusive to the electrical trades. Although this is a significant barrier to understanding the underlying principles and foundations of a building’s equipment and distribution of electricity consumption, a marriage between Data Science applications and integrated Electrical Systems will lead to optimization of energy consumption patterns in response to several aspects such as occupancy.

The objective for summer research was to explore the electrical systems and procedures in place for the Claudette MacKay-Lassonde Pavilion (CMLP) Engineering Building. In doing so, key insights were to be provided for making the new Three C+ Engineering Building more energy efficient. The open-ended nature of this task allowed for flexibility and aided in pursuing the most ideal path towards transferrable insights. The following explains the evolution of objectives for summer research:

1. Explore the procedures currently upheld by Facilities Management personnel, and how conducive they are to supporting energy efficient procedures. This would be useful in enabling the research group to make more informed decisions.
2. Perform Time Series analysis on energy consumption data collected over several years for different electrical circuits within CMLP; visualize and manipulate data to generate key insights that can contribute to the research group’s knowledge of energy consumption patterns.
3. Use information found in time series exploration to create Exogenous and Non-Exogenous Seasonal Autoregressive Integrated Moving Averages (ARIMA), so as to forecast energy consumption for the building. This step would be useful in explaining how insights translate to a more accurate model.

**Institutional and Analytical Research**

In order to accomplish the objectives set out for the summer, research and learning had to be performed on two fronts: institutional and analytical. An understanding the building management procedures was developed, which defines the constraints within which new solutions to existing problems can be deployed for the new building. By exploring and analyzing the data, key insights were obtained pertaining to energy usage patterns of CMLP building, which will be similar to that of the new Three C+ Engineering Building.

Over 3 meetings with Facilities Management personnel, several aspects of the building-management procedures and resources were learnt. A great number of sensors make a variety of measurements across campus; however, only a handful are stored permanently. A vast majority of the sensors collect measurements that are only available for live visualization through a Web Platform. Therefore, most of the data that is measured is not actively monitored for optimization, but only for troubleshooting when issues arise. It is also evident that the University has not extracted maximum value from the Building Information System in place, as it is used for reactionary purposes rather than for precautionary and planned activities. Because the electricity metering for the building was not planned at the time according to needs of the future, it remains to be used for functional purposes at best. The system of meters and sensors, and the type of data collected are not tailored to needs of energy efficient buildings, rather investments have been made on the former for trivial purposes such as decisions pertaining to whether or not an electricity ‘line’ has enough capacity to support an entire new building. Nonetheless, this sentiment was echoed by Facilities Management personnel – much cannot be accomplished with the resource constraints on the human resources.

In spite of these persistent problems that have limited the university’s accomplishments pertinent to building efficiency and energy conservation, there exists several opportunities to fill human resource gaps by automating several tasks. These tasks include monitoring abnormal electricity consumption, deploying field engineers to issue-sites, and performing predictive maintenance among others.

The summer research period, during which I was tasked to accomplish the aforementioned objectives, consisted of significant learning. The month of May was primarily spent on learning various tools that would be necessary for Data Science-related work that would ensue for the remainder of the summer. These include learning about the aspects of the Python programming language and environment (through the usage of browser-based Jupyter Notebooks IDE) pertaining to Data Science work. This step proved to be an important one, since the tools were essential to manipulate, edit, clean, and explore the electricity consumption data. These tools are as follows: Python Numpy for Matlab-like multi-dimensional array creation and manipulation, used for improvement of algorithms’ performance on the Array Data Structure; Python Pandas package for arrangement, easy manipulation and exploration of multivariate data, represented by Series and DataFrames (such as those in R programming language); Python Matplotlib package for creation of graph objects for variety visualizations of different aspects of data in DataFrames; Python ‘seaborn’ package, built on top of Matplotlib, for greater ease of creating more sophisticated graphing plots.

Equipped with a basic understanding of data manipulation and exploration tools offered through the several Python packages, data that was collected for the CMLP building was explored to better understand underlying effects attributed to time. Together with meetings with the Building Managers at Western Facilities’ group and application of tools on the building data, the building’s Information Systems were better understood from not only a technical perspective, but an institutional one as well. The latter has been particularly important to understand what is truly needed for improving energy efficiency, and what changes can be sustainably implemented if at all. Data collected for CMLP spanned over 5 years – from 2012 to 2017 – with a granularity of 5-minute intervals, i.e. measurements were made and data points were stored every 5 minutes. The fields in which measurements were made include steam and ethanol flow rates for the heating systems, water flow rates through several pipelines within the building, and instantaneous electricity consumption data (in kW units) and cumulative electricity consumption data (in kWh units) for various circuits within the building. Every 5 minutes for each data field, the timestamp value and the measured value are stored. While a total of greater than 100 000 data fields exist throughout campus (collecting different types of data pertaining to buildings), there are only about 50 such data fields for CMLP, many of which do not involve electricity data. Nevertheless, the data fields that explain electricity consumption over the years were categorized into differently numbered meters, representing separate electricity circuits.

Summary of findings from the processes above:

* Tangible findings pertaining to energy data exploration – culminating into finding that the current data is very much useless due to the data fields collected thus far.
* Findings pertaining to building management
* Ideas for further research, and pertaining fields.