Service Learning Practicum Requirements Document
LEAP (Local Energy Alliance Program)

Overview: The Local Energy Alliance Program (LEAP; http://leapva.org/) helps homeowners and businesses increase the energy efficiency of their buildings, through such things as insulation, increased efficiency appliances, etc. They perform energy audits of buildings, and use the results to recommend changes to said building.

Contact: Jon Proffitt (jon@leapva.org). Lesley Fore (lesley@leapva.org) will frequently be involved as well, but Jon is the primary contact.

Summary: The project is a data management system that allows LEAP to keep the data from multiple disparate sources in one database; some of the data will be noisy. Data visualization and extracting statistics from the data will be supported as well.

Notes: The data is confidential, so students will have to sign a nondisclosure agreement with LEAP.

Description: After a house has improvements, LEAP keeps track of the energy usage. In particular, the period of 12 months prior to the improvements and 12 months after the improvements is recorded to analyze the actual energy efficiency improvement. Customers sign an agreement that allows LEAP to obtain the data directly from the local utility companies (currently only electricity and only from a few power companies but will eventually be expanded to other utilities, such as gas).

The data that arrives from the utilities is noisy. A file will likely have data for many different addresses and for many different months; the data itself may contain how much their bill was, for which the energy usage will have to consider the current electric rate; otherwise, it may contain the actual amount of the energy used. Although the required fields are present, depending on which utility company employee generated the data, there could be many additional fields, or the fields could be named something different. If the utility company has a delay in providing the data to LEAP, there may be "holes" in the period covered, which the system will have to provide an alert for. Furthermore, addresses may be represented differently: "120 N. 5th St", "120 North 5th Street", "120 5th St N" are all the same address, but different data sets may represent the same address differently. Likewise, the data in some of the date columns, for example could be in different formats.

Electric data comes only from Dominion, and there are two companies that provide natural gas information. LEAP has a separate database that contains all the house and project information; this is in a system called Optix, and that data can be exported for use in the system being developed.

Among other information, each house has a unique house ID field. Houses also have 20 (or so) Boolean fields that describe what was done on that house: insulation, new heat pump, etc.

There will likely be other data fields that will need to be stored, but which are not necessary to go into detail in this document. All these fields will need to be configurable in the system, as they will change over time (new ones added, in particular).

All data is in CSV format, or a format that can be exported to CSV format prior to loading into the system (such as XLS). Currently, the two types of data files to be read in would be the housing data (from Optix) and the utility data. As the utility data will have the addresses in different formats, the system will have to remember that "120 N. 5th St" is the same as "120 5th St North". Any "unknown" addresses should present to the user so s/he can identify which address that is for.

An upload data set could contain data from a previous set, so there will need to be a means to identify data that is identical to what was uploaded before.

Requirements: Minimum
The minimum requirements are what is minimally necessary to create a system that is of value to the customer. It is expected that this system will be completed by the end of the fall semester (i.e., December).
- User login
- Allow for the upload of both types of data (utility)
- Simple data visualization
     1. Gaps in Data
          1. Color coded Gaps Reports to visually flag weird places in data
          2. Different color in data visualization for missing data, and data that is potentially outliers
- Extracting data from the system, in a series of predefined reports and queries
     1. Need to capture meter read date
     2. Queries / Search functionality

Requirements: Desired
The desired requirements, along with the minimum requirements, constitute the hoped for system. It is expected that this system will be completed by the end of the spring semester (i.e., May).
- More & better data visualization (reports)
     o Reports
          ▪ Utility Requests
               - Info needed on this report: Name, Customer ID, Date
          ▪ Completed Job Data sets
               - Name
               - Jon will get back to with more details
- No one outside the system can "sign up"
- Display the number of days in the reading period so data can be better analyzed
- More general queries to allow for better data extraction and data mining
     a.   Have auto-complete / partial match
     b.   Search By:
                    1. First Name
                    2. Last Name
                    3. Property Address
     c.   Filter:
                    1. ascending / descending
                    2. by name
                    3. by test out date

4. by installed measures
- List of installed measures: include fields for possible installed measures-- each property can have measures associated with it. There should be a free text field associated with each properties measure where the user can enter contractor data/other information
- Ability to update customer name and account number
- Have filtering results downloadable in a PRISM ready format

Development notes
Platform preference: Ruby on Rails

Approved by Team and Jon Proffitt 3/5/2014