Capstone III Project Proposal

1. *What dataset or datasets do you plan to use? What are the features, rows, and data types of each?*

The dataset that I plan to use is a CSV file that [comes from Kaggle.](https://www.kaggle.com/datasets/thedevastator/residential-home-energy-efficiency) It provides information about energy efficiency meter evaluated projects for residential existing homes in New York state from 2007-2012. Below is each column along with its variable type.

|  |  |  |
| --- | --- | --- |
| Variable | Quantitative or qualitative? | Variable type |
| Project County | Qualitative | Nominal |
| Project City | Qualitative | Nominal |
| Size of Home | Quantitative | Continuous |
| Number of Units | Quantitative | Discrete |
| Total Project Cost | Quantitative | Continuous |
| Baseline Electric Usage | Quantitative | Continuous |
| Reported Electric Usage | Quantitative | Continuous |
| Estimated Annual Electric savings | Quantitative | Continuous |

1. *What research or business questions do you want to answer?*

I would like to answer two questions using the data. The first is to see if there is any correlation between cost of work and reporting energy costs. I will check for calculated field for the difference in energy saved.

The second question is Does the size of the house or the location affect the project cost more?

1. *What are your hypotheses going in?*

My first hypothesis is that the cost of work and the energy savings done are correlated. This means that houses that build for energy savings are expected to save more energy in the future. To confirm this, I would expect the correlation coefficients to be at least 0.5.

My second hypothesis is that house size or region affects costs. I will test this at the 95% significance level.

1. *How will you use your data to test your hypotheses?*

I will use the Pearson r test for relationship between difference in energy costs. I will use the independent samples t-test to check for the location and size of building.

1. *Who will find your findings valuable, and how will they use them?*

Energy saving investments are costly and one of the most important issues for homeowners. If they know about this subject. They know what awaits them. It allows them to help them make the right decision before investing.