Project Report

1. **Problem Description**
   What are you trying to predict? Why did you choose this problem?

   I tried to predict the monthly electric bill of residential buildings. The reason for choosing this problem is to contemplate electricity usage without wasting.

2. **Dataset**
   Where is your data from? What features did you use, and why?

   This data was found from kaggle.com and I selected useful features.
   'Area' = the total area of the site in square meters
   'Structure type' = the primary use of the building
   'Water consumption' = daily water usage per day for a month
   'Resident count' = the population of the site

3. **Modeling Approach**
   What model(s) did you use? How did you preprocess or scale the data?

   I used Linear Regression Model , Multiple Linear Model and Polynomial Regression Model to determine which model would be best. Among them , I finally chose the Multiple Linear Model which is the best performance. In the preprocessing stage I chose (Residential) from 'Structure type' as a parameter.

4. **Evaluation**
   How well did your model perform? Which metric(s) did you use?

   My Multiple Linear Regression Model performance is 93% fit .
   The metric values are
   - R squared = 0.93
   - MAE = 234
   - MSE = 82875.

5. **Reflection**
   What challenges did you face? What did you learn from this project? What would you do differently next time?

   Data preprocessing process is the most challenging step that I faced. I learned the differences of these models and the data preprocessing skills. Next time I would be more cautious on the data preprocessing and to find a more efficient way for evaluation.