Group member: Hao Cheng, Sheng Luo

1. Title

Electric Data Generation Based on Austin, Texas Electric Dataset

1. Purpose of the Study – What you hoping to accomplish

Our purpose is to generate data that is close to the real data. We plan to find out the relationship between the numeric characters of each state and its energy usage distribution.

1. Practical Objectives – What problem are you trying to solve, high level results

Real data of granular appliance energy usage for a household is important for many reasons. For the government, knowledge of the energy usage can help them save energy and reduce emissions; for a company, it can help them find potential customers. Unfortunately, real datasets are very scarce or not accessible for public. If we can generate data whose distribution is similar to the one of real data, we can significantly improve the capability of government and companies to do data-driven decision making.

1. Research Questions

How to generate electric data based on known dataset?

1. Data Source

The main data source is Data port’s Pecan Street.

Data port’s Pecan Street contains a household energy use database. This database has many features of use for this objective.

We are also trying to find numeric characters of other states from other sources, such as Kaggle, NOAA, World Bank and United States Census, etc.

1. Potential Analysis Method

VAE(Variational Autoencoder); InfoGAN(Generative Adversarial Network); Conditional GAN; VAE-GAN; GNN(Graph Neural Network)