**Group Project Report**

**University of New Brunswick, Saint John**

**CS 2704: Data Analytics using Python**

**Professor Jong-Kyou Kim**

**Group Members:**

**Aryan Bansal**

**Aniket Shah**

**Dhruv Thumar**

### **Dataset**

I will use two publicly available datasets:

* **Electricity Consumption Data** (kWh per state per year)
  + Source: [U.S. Energy Information Administration (EIA)](https://www.eia.gov/electricity/data.php)
* **Temperature Data** (Average yearly temperature by state)
  + Source: [NOAA Climate Data](https://www.ncdc.noaa.gov/cdo-web/)

These datasets provide yearly electricity consumption and temperature data by U.S. state and will be merged using “State” and “Year.”

### **GitHub Repository**

**Repository Link**: [https://github.com/dhruvthumar/2704project](https://github.com/dhruvthumar/2704project?tab=readme-ov-file) The repo will include:

* Raw dataset and cleaned dataset
* Source code for analysis
* This proposal
* Final report and slides

### **Original Hypothesis**

Higher average temperatures in a region correlate with increased electricity consumption due to air conditioning use.