

Project

What are we working on?

How climate, energy consumption, and cost of energy affects individuals in cities across Texas

Why is this data important?

We wanted to look into how the average temperatures in certain cities across the state affect the average households cost of energy for their homes.

If some cities are paying a higher percentage of their income towards energy for their home we wanted to compare it with the temperatures of those cities to try and understand why this may be happening

For the more extreme months energy consumption obviously will increase. We wanted to compare these months across the cities to gauge how big these increases may be for specific years.

What We Hoped to Find:

- We wanted to find outlier months or years in which electricity consumption or prices were abnormally high compared to previous numbers.
- We also wanted to see how different cities across the state consume energy based on the temperatures of specific months. If average temperatures and energy consumption were similar we wanted to see if certain cities pay a higher percentage of their income towards the cost of their energy

Datasets We Used:

- Average Annual income per city
- Monthly Climate per city
- Texas Monthly Energy rates
- Texas Monthly Energy consumption

Climate of Texas 2010-2020:

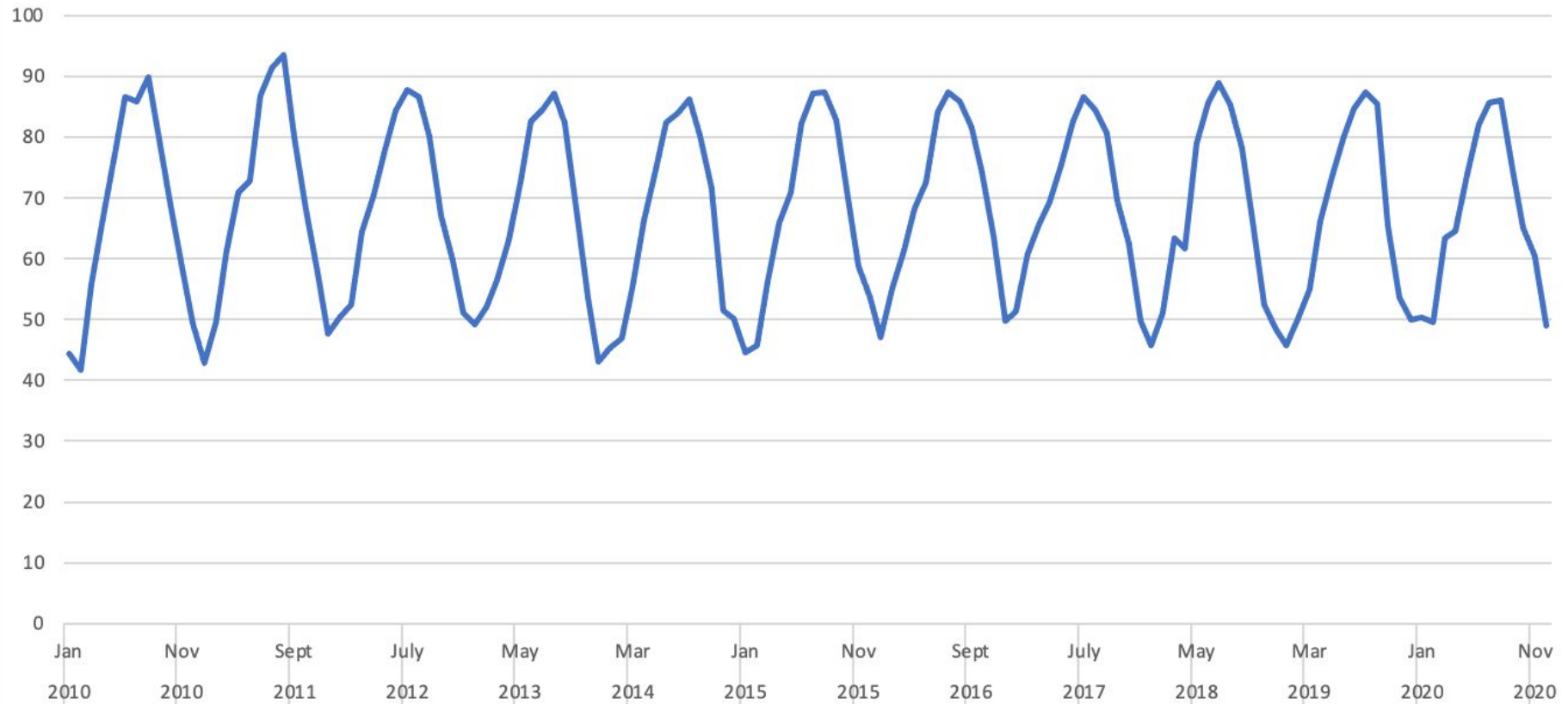
<https://www.weather.gov/wrh/Climate?wfo=lub>

Used to find the monthly average temperatures for the years 2010-2020 for each city

Uses NOWdata - NOAA Online Weather Data

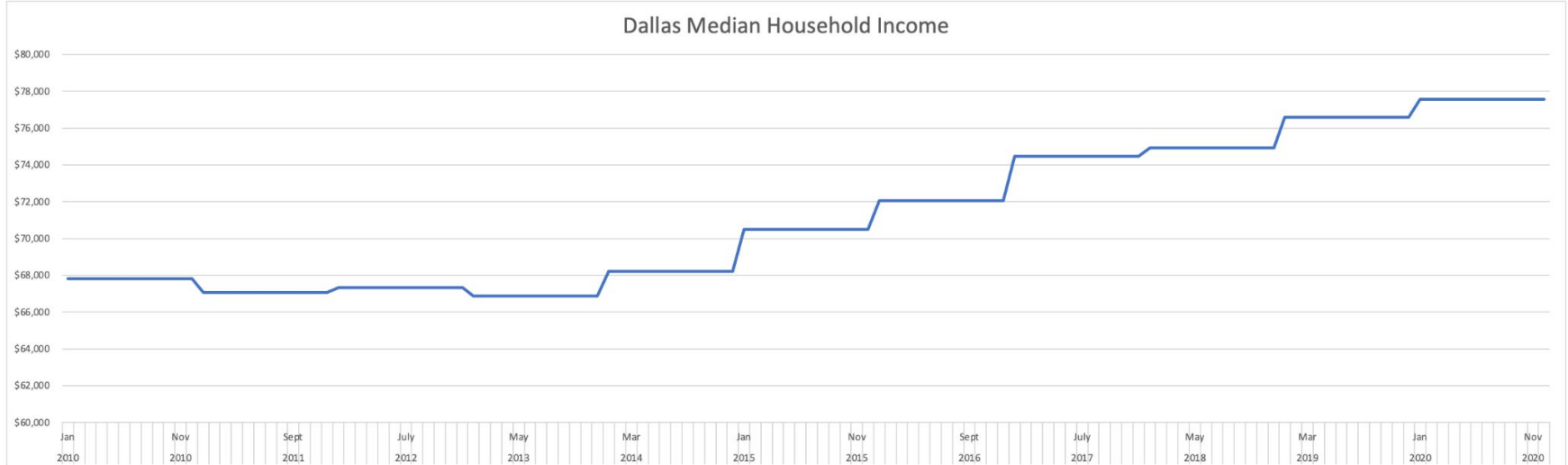
The Applied Climate Information System (ACIS) is a joint project of the Regional Climate Centers, the National Centers for Environmental Information (NCEI) and the National Weather Service. Official data and data for additional locations are available from the Regional Climate Centers and NCEI.

Dallas Average Temperature



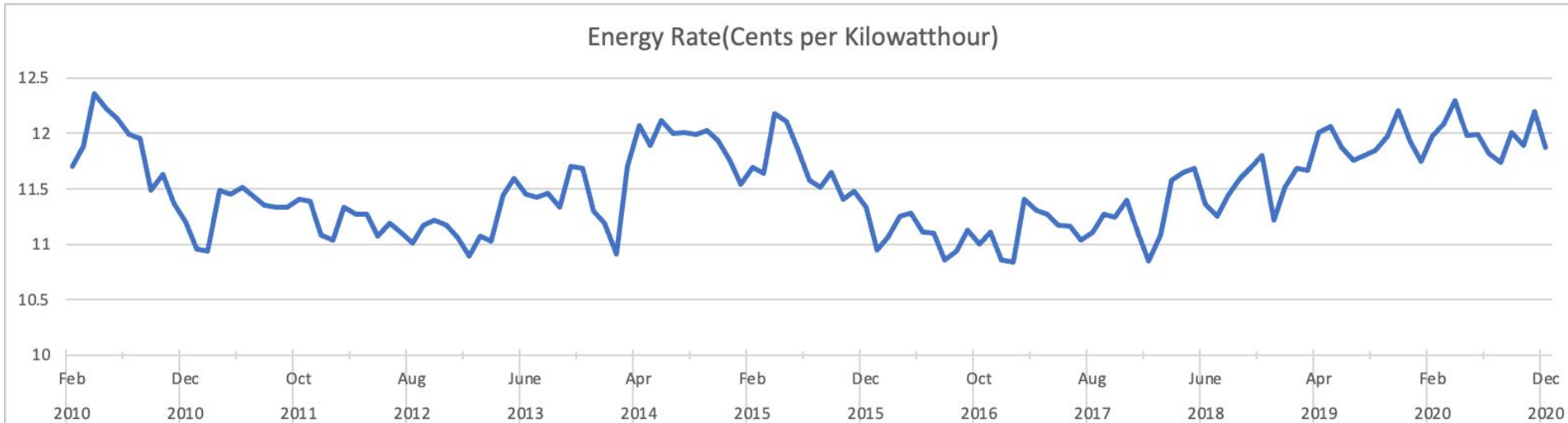
Income in Texas 2010-2020:

- <https://www.deptofnumbers.com/>
- Has the census ACS annual survey of median household income
- Average Median Income
- Used Lubbock, Houston, Austin, Dallas, Amarillo, Corpus Christi



Energy Rates 2010-2020:

- <https://www.eia.gov/electricity/monthly/> (Energy Information Administration)
- Gives Monthly Reports on Electricity
- Average Price for Residential Consumers
- Cents Per Kilowatt Hour



Energy Consumption 2010-2020:

- <https://www.eia.gov/electricity/monthly/>
- This data is Residential Customers Energy Consumption in Thousand Megawatt Hours.
- The PDF data gave monthly usage of every state.
- August had the most usage with an average of 17,355
- April had the least usage with an average of 8,632

30000 foot overview

What We Found:

Average temperatures for Texas summers are more extreme than the average temperature for the winter months.

When the temperatures reach their peak of the year for the months of July and August more energy is consumed to keep a home cool compared to the energy consumed to keep a home warm in the winter months.

Throughout the years the price has remained within close to 11-12 cents per kilowatt hour. With the peak happening in April of 2010 at 12.36 and the lowest being in January of 2017 at 10.84

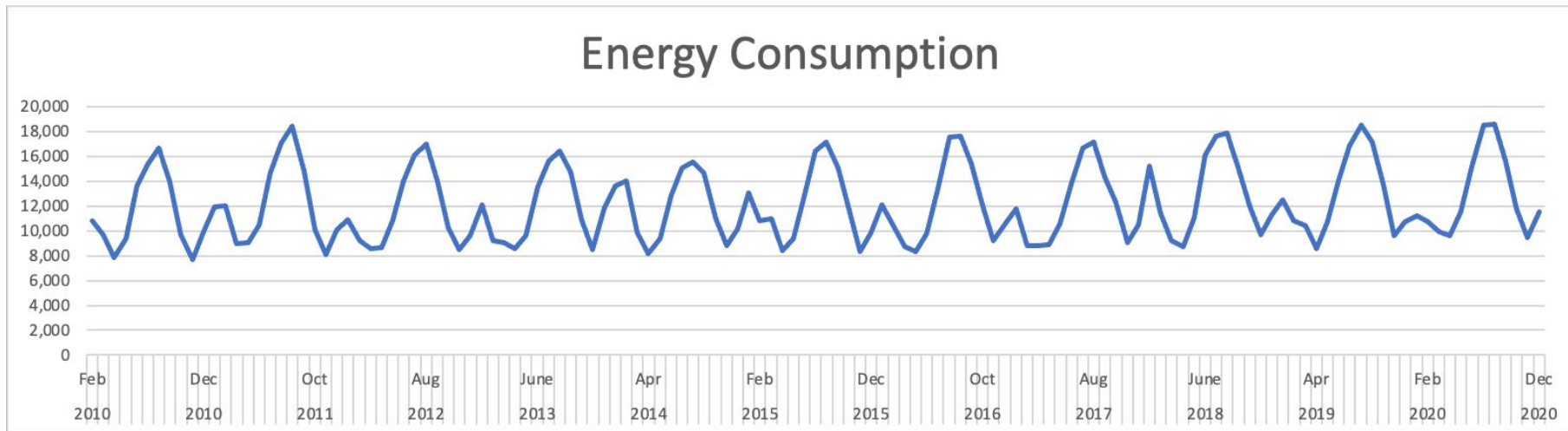
Energy consumption also remains pretty steady with the highest peak occurring in August of 2020 at 18,595 and the lowest occurring in November of 2010 with 7667

With energy consumption and prices being the average for the state of Texas we can only assume that the cities with higher average incomes are better off during the high consumption months because a lower percentage of their monthly income goes towards the cost of energy for their homes.

Further Conclusions from the Data:

The seasonality speaks for itself in the graph below that depicts energy consumption in thousands megawatts-hours. Though the seasonality remains consistent across the data set, some spikes can certainly be seen in years such as the first quarter of 2014 in which there is an unusual large spike in the winter-time months.

In addition to this, the winter of 2017 represents another outlier worth looking into. Looking at the data in a surface-level way such as this can give us insights into where we need to dig deeper by combining with data sets such as income and weather by city.



Data Quality

- Couldn't find Energy Rate and Consumption specified by cities
 - Only found it for the whole state of Texas
- Average Household Income is annually
- The data for temperatures is the monthly average. The data quality for weather is very high. We were able to find even the daily temperatures for the entire year per city but chose to go with monthly in order to keep the data small enough to work with but large enough to provide us with enough information to draw conclusions from

Issues with Merging:

- Trouble merging the energy datasets with the income datasets
 - Lost a lot of rows
 - Identified: The Month Columns in the datasets were different
- We had some issues merging the monthly columns into a “month” column that contained the months iterating through the rows

The Code:

p2projectcode.py x project_update.py x

```
1  ## Final Project Code Merging the Data Sets ##
2
3  ## Imports ##
4  import pandas as pd
5
6  ## Filepath and CSVs ##
7  energy_rate = 'energy_rate.csv'
8  new_income = 'new_income.csv'
9  monthly_energy = 'monthly_tx_energy.csv'
10 avg_temp = 'Avg_Month_Temp.csv'
11 filepath = 'C:/Users/gonza/Downloads/'
12
13 ## Reading the CSVs ##
14 energy_rate_path = filepath + energy_rate
15 df_enrate = pd.read_csv(energy_rate_path, delimiter=',')
16 tx_income_path = filepath + new_income
17 df_income = pd.read_csv(tx_income_path, delimiter=',')
18 monthly_energy_path = filepath + monthly_energy
19 df_mont_energy = pd.read_csv(monthly_energy_path, delimiter=',')
20 temp_path = filepath + avg_temp
21 df_avg_temp = pd.read_csv(temp_path, delimiter=',')
22
23
24 ## Merging the Data ##
25 df_enrate["Month"][df_enrate["Month"]=="January"]="Jan"
26 df_enrate["Month"][df_enrate["Month"]=="February"]="Feb"
27 df_enrate["Month"][df_enrate["Month"]=="March"]="Mar"
28 df_enrate["Month"][df_enrate["Month"]=="April"]="Apr"
29 df_enrate["Month"][df_enrate["Month"]=="May"]="May"
30 df_enrate["Month"][df_enrate["Month"]=="June"]="June"
31 df_enrate["Month"][df_enrate["Month"]=="July"]="July"
32 df_enrate["Month"][df_enrate["Month"]=="August"]="Aug"
33 df_enrate["Month"][df_enrate["Month"]=="September"]="Sept"
34 df_enrate["Month"][df_enrate["Month"]=="October"]="Oct"
35 df_enrate["Month"][df_enrate["Month"]=="November"]="Nov"
36 df_enrate["Month"][df_enrate["Month"]=="December"]="Dec"
37
38 df_mont_energy["Month"][df_mont_energy["Month"]=="January"]="Jan"
39 df_mont_energy["Month"][df_mont_energy["Month"]=="February"]="Feb"
40 df_mont_energy["Month"][df_mont_energy["Month"]=="March"]="Mar"
41 df_mont_energy["Month"][df_mont_energy["Month"]=="April"]="Apr"
42 df_mont_energy["Month"][df_mont_energy["Month"]=="May"]="May"
43 df_mont_energy["Month"][df_mont_energy["Month"]=="June"]="June"
44 df_mont_energy["Month"][df_mont_energy["Month"]=="July"]="July"
45 df_mont_energy["Month"][df_mont_energy["Month"]=="August"]="Aug"
46 df_mont_energy["Month"][df_mont_energy["Month"]=="September"]="Sept"
47 df_mont_energy["Month"][df_mont_energy["Month"]=="October"]="Oct"
48 df_mont_energy["Month"][df_mont_energy["Month"]=="November"]="Nov"
49 df_mont_energy["Month"][df_mont_energy["Month"]=="December"]="Dec"
50
51 df_income_energy_merge = df_income.merge(df_enrate, how='inner', on=['Month', 'Year'])
52 df_merge_2 = df_income_energy_merge.merge(df_mont_energy, how='inner', on=['Month', 'Year'])
53 df_merge_all = df_avg_temp.merge(df_merge_2, how='inner', on=['Month', 'Year', 'City'])
54
55 ## Write to csv ##
```

Name	Type	Size	Value
avg_temp	str	18	Avg_Month_Temp.csv
df_avg_temp	DataFrame	(792, 4)	Column names: City, Year, Month, Avg_Temp
df_enrate	DataFrame	(132, 3)	Column names: Energy Rate, Year, Month
df_income	DataFrame	(792, 4)	Column names: City, Year, Month, Income
df_income_energy_merge	DataFrame	(726, 5)	Column names: City, Ye...
df_merge_2	DataFrame	(726, 18)	Column names: City, Ye...
df_merge_all	DataFrame	(132, 19)	Column names: City, Ye...
df_mont_energy	DataFrame	(132, 15)	Column names: Year, Mo...
energy_rate	str	15	energy_rate.csv
energy_rate_path	str	48	C:/Users/gonza/Downloa...
filepath	str	25	C:/Users/gonza/Downloads/
monthly_energy	str	21	monthly_tx_energy.csv

Console 1/A x

```
c:\users\gonza\downloads\project_update.py:48: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df_mont_energy["Month"][df_mont_energy["Month"]=="November"]="Nov"
c:\users\gonza\downloads\project_update.py:49: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
df_mont_energy["Month"][df_mont_energy["Month"]=="December"]="Dec"

In [29]:
```


Major Obstacles:

- Navigating through 300 page pdfs for each monthly energy consumption data
 - Same for energy rate
- Finding datasets that matched each other in order to make proper comparisons